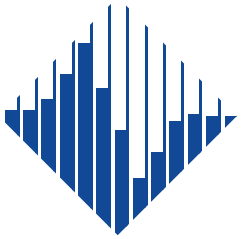


PACIFIC EARTHQUAKE ENGINEERING
RESEARCH CENTER

Central and Eastern North America Ground-Motion Characterization

NGA-East Final Report

Christine Goulet
University of Southern California

Yousef Bozorgnia
University of California, Los Angeles

Norman Abrahamson
University of California, Berkeley

Nicolas Kuehn
University of California, Los Angeles

Linda Al Atik
Independent Consultant

Robert Youngs
Wood Environment & Infrastructure Solutions, Inc.

Robert Graves
U.S. Geological Survey

Gail Atkinson
Western University

PEER Report No. 2018/08
Pacific Earthquake Engineering Research Center
Headquarters at the University of California, Berkeley

December 2018

PEER 2018/08
December 2018

Disclaimer

The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the study sponsor(s), the Pacific Earthquake Engineering Research Center, or the Regents of the University of California.

Central and Eastern North America Ground-Motion Characterization

NGA-East Final Report

Christine Goulet
University of Southern California

Yousef Bozorgnia
University of California, Los Angeles

Norman Abrahamson
University of California, Berkeley

Nicolas Kuehn
University of California, Los Angeles

Linda Al Atik
Independent Consultant

Robert Youngs
Wood Environment & Infrastructure Solutions, Inc.

Robert Graves
U.S. Geological Survey

Gail Atkinson
Western University

PEER Report No. 2018/08
Pacific Earthquake Engineering Research Center
Headquarters at the University of California, Berkeley

December 2018

Abstract

This document is the final project report of the Next Generation Attenuation for Central and Eastern North America (CENA) project (NGA-East). The NGA-East objective was to develop a new ground-motion characterization (GMC) model for the CENA region. The GMC model consists of a set of new ground-motion models (GMMs) for median and standard deviation of ground motions and their associated weights to be used with logic-trees in probabilistic seismic hazard analyses (PSHA).

NGA-East is a large multidisciplinary project coordinated by the Pacific Earthquake Engineering Research Center (PEER), at the University of California. The project has two components: (1) a set of scientific research tasks, and (2) a model-building component following the framework of the "Seismic Senior Hazard Analysis Committee (SSHAC) Level 3" (Budnitz et al. 1997; NRC 2012). Component (2) is built on the scientific results of component (1) of the NGA-East project. This report documents the tasks under component (2) of the project.

Under component (1) of NGA-East, several scientific issues were addressed, including: (a) development of a new database of ground motion data recorded in CENA; (b) development of a regionalized ground-motion map for CENA, (c) definition of the reference site condition; (d) simulations of ground motions based on different methodologies; and (e) development of numerous GMMs for CENA. The scientific tasks of NGA-East were all documented as a series of PEER reports.

The scope of component (2) of NGA-East was to develop the complete GMC. This component was designed as a SSHAC Level 3 study with the goal of capturing the ground motions' center, body, and range of the technically defensible interpretations in light of the available data and models. The SSHAC process involves four key tasks: evaluation, integration, formal review by the Participatory Peer Review Panel (PPRP), and documentation (this report).

Key tasks documented in this report include review and evaluation of the empirical ground-motion database, the regionalization of ground motions, and screening sets of candidate GMMs. These are followed by the development of new median and standard deviation GMMs, the development of new analyses tools for quantifying the epistemic uncertainty in ground motions, and the documentation of implementation guidelines of the complete GMC for PSHA computations. Appendices include further documentation of the relevant SSHAC process and additional supporting technical documentation of numerous sensitivity analyses results. The PEER reports documenting component (1) of NGA-East are also considered "attachments" to the current report and are all available online on the PEER website (https://peer.berkeley.edu/).

The final NGA-East GMC model includes a set of 17 GMMs defined for 24 ground-motion intensity measures, applicable to CENA in the moment magnitude range of 4.0 to 8.2 and covering distances up to 1500 km. Standard deviation models are also provided for site-specific analysis (single-station standard deviation) and for general PSHA applications (ergodic standard deviation). Adjustment factors are provided for consideration of source-depth effects and hanging-wall effects, as well as for hazard computations at sites in the Gulf Coast region.

References

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Clu, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. 1997. “Recommendations for probabilistic seismic hazard analysis: Guidance on uncertainty and use of experts.” Report NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington D.C.

NRC. 2012. “Practical implementation guidelines for SSHAC Level 3 and 4 hazard studies.” Technical Report NUREG-2117, U.S. Nuclear Regulatory Commission, Washington D.C.

Acknowledgments

This study was sponsored by the Pacific Earthquake Engineering Research Center (PEER), as part of the NGA-East research project, and was funded by the U.S. Nuclear Regulatory Commission (NRC), the U.S. Department of Energy (DOE), and the Electric Power Research Institute (EPRI), with the participation of the U.S. Geological Survey (USGS). Some individual researchers have also been funded by other agencies. Supports of NGA-East at all levels are gratefully acknowledged. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsoring agencies.

Support and encouragement of the following representatives of the sponsoring agencies are gratefully acknowledged: Annie Kammerer (formerly with NRC), Clifford Munson (NRC), Steve McDuffie (DOE), Larry Salomone (EPRI), and Mark Petersen (USGS). The project could not start without the initial supports, encouragement, and enthusiasm of Annie Kammerer.

Members of the Participatory Peer Review Panel (PPRP) spent a tremendous amount of time reviewing all pieces of the project. PPRP's input and feedback have resulted in a great improvement of the project and its outcome. We appreciate all efforts and cooperation of the PPRP members: Gabriel Toro (PPRP Chair), John Adams, Jon Ake, John Ebel, Jeff Kimball, Rich Lee, and Julian Bommer (PPRP former Chair).

We recognize the dedication and unwavering support from our editor, Claire Johnson, on all the reports produced by the project.

Like with past NGA research programs, NGA-East is another manifestation of the extraordinary teamwork and collaboration among numerous individuals tackling various tasks in the project. We could not finish the project without the efforts and collaboration of the following individuals (listed alphabetically): Nayeem Al Noman, Trevor Allen, Timothy Ancheta, John Anderson, Rasool Anooshehpour, Ralph Archuleta, Karen Assatourians, Annemarie Baltay, Jeff Bayless, Greg Beroza, Jack Boatwright, David Boore, Oliver Boyd, Alexander Bykovstev, Kenneth Campbell, Brian Carlton, Martin Chapman, Brian Chiou, Ariel Conn, Christopher Cook, Kevin Coppersmith, Chris Cramer, Jorge Crempien, Luis Dalguer, Bob Darragh, Sahar Derakhshan, Carola Di Alessandro, Doug Dreger, Jennifer Dreiling, Arthur Frankel, Richard Godbee, Vladimir Graizer, Nick Gregor, Aybars Gurpinar, Christina Hale, Jeff Hamel, Tom Hanks, Joseph Harmon, Ruth Harris, Youssef Hashash, Behzad Hassani, Robert Herrmann, Justin Hollenback, Marius Isken, Ronnie Kamai, Byungmin Kim, Tadahiro Kishida, Albert Kottke, Olga Ktenidou, Maurice Lamontagne, Phil Maechling, Steve Mahin, Jimmy Martin, Silvia Mazzoni, Arthur McGarr, Robin McGuire, Walter Mooney, Cheryl Moss, Charles Mueller, Sifat Muin, Michael Musgrove, Sissy Nikolaou, Kim Olsen, Grace Parker, Shahram Pezeshk, Arben Pitarka, Ellen Rathje, Leon Reiter, Philippe Renault, Sanaz Rezaeian, Andreas Rietbrock, Richard Rivera-Lugo, Adrian Rodriguez-Marek, Justin Rubistein, Frank Scherbaum, Alireza Shahjouei, Fabio Silva, Walt Silva, Alireza Haji Soltani, Paul Somerville, Paul Spudich, Jonathan Stewart, Ken Stokoe, Fleur Strasser, Sarah Tabatai, Rumi Takedatsu, Behrooz Tavakoli, Eric Thompson, David Wald, Melanie Walling, Rob Williams, Katie Wooddell, Emrah Yenier, Alan Yong, Arash Zandieh, Yuehua Zeng, and John Zhao.

Contents

Appendices

List of Tables

CHAPTER 1

CHAPTER 2

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 8

CHAPTER 9

CHAPTER 10

CHAPTER 11

CHAPTER 12

CHAPTER 13

CHAPTER 14

List of Figures

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 8

CHAPTER 9

CHAPTER 10

CHAPTER 11

CHAPTER 12

CHAPTER 13

CHAPTER 14

1. Introduction

1.1 Background

The objective of the Next Generation Attenuation for Central and Eastern North America (CENA) project (NGA-East) has been to develop a new ground-motion characterization (GMC) model for the CENA region. The GMC model consists of a set of new ground-motion models (GMMs)—also known as ground-motion prediction equations (GMPEs)—for median and standard deviation of ground motions and their associated weights in the logic-trees for use in probabilistic seismic hazard analysis (PSHA).

The NGA-East Project is a multidisciplinary project coordinated by the Pacific Earthquake Engineering Research Center (PEER), at the University of California. The project was originally developed as a "science-based" research project (Bozorgnia 2008). In 2010, the sponsors of the project recommended that the final results and models of NGA-East be developed under the framework of the Seismic Senior Hazard Analysis Committee (SSHAC) (Budnitz et al. 1997; NRC 2012) Level 3. Thus, the project has two components: (1) a set of scientific research tasks, and (2) a model-building component following the SSHAC Level 3 process. Component (2) of the NGA-East project was developed using the products of various scientific research tasks completed under Component (1) of NGA-East. This document is the final project report for Component (2) of the NGA-East project. A list and summary of scientific reports published under Component (1) of NGA-East are provided in this chapter. An overview of the SSHAC process, as it was implemented in this project, is presented in Chapter 2.

Integrating the SSHAC Level 3 framework to the NGA-East resulted in a considerable additional complexity to the scope of work of the project, as is evident in the project organization summarized in Chapter 2 of this report. The end users of NGA-East, especially for nuclear facilities, will likely also use the products of a recently completed project, the Central and Eastern US (CEUS) Seismic Source Characterization project, CEUS SSC (EPRI/DOE/NRC 2012). The CEUS SSC project was also developed under the SSHAC Level 3 framework; thus, the end users of a combined CEUS SSC and NGA-East projects will have access to a combined high-quality SSHAC Level 3 product.

The NGA-East project involved a large number of participants from various organizations in academia, private industry, and government. Including both researchers and practitioners in the overall effort has increased the scientific quality of the results; at the same time, it also enhanced their usefulness for the end users.

The NGA-East project was jointly sponsored by the U.S. Nuclear Regulatory Commission (NRC), the U.S. Department of Energy (DOE), the Electric Power Research Institute (EPRI), and the U.S. Geological Survey (USGS). The sponsoring agencies' interest has been to develop a GMC model for the CEUS. Nonetheless, because the tectonic region of interest reaches across Canada and a large number of earthquake data was also recorded in Canada, the GMC developed in NGA-East is applicable to the larger CENA region.

1.2 Project Objectives and Limitations

1.2.1 Project Objective

The NGA-East project objective is: to provide the best estimate of the distribution (median and standard deviation) of $RotD_{50}$ (Boore 2010) horizontal ground motions for the peak ground acceleration (PGA), peak ground velocity (PGV), and 5%-damped pseudo-acceleration response spectra (PSA) for oscillator periods ranging from 0.01 to 10 sec on "hard-rock" sites located up to 1500 km from future earthquakes in CENA with moment magnitudes in the 4.0–8.2 range, and to provide the epistemic uncertainty associated with this estimate. Details of the NGA-East project plan, as originally envisioned, are provided in Goulet, et al. (2011) as provided in Appendix A.

As indicated previously, for the "model-building" component of the NGA-East, the objective stated above is to be achieved in the context of a SSHAC Level 3 study. More specifically, the SSHAC objectives are associated with the development of a recorded ground-motion database, the full development of the GMC that captures the center, body and range (CBR) of technically defensible interpretations (TDI) of the available data, models and methods, the development of exhaustive documentation, and a thorough peer review (see also Chapter 2 of this report).

1.2.2 Study Region

The study region is CENA is shown in yellow in Figure 1–1. Details of regionalization boundaries for NGA-East are presented in Chapter 4 of this report.

1.2.3 Reference Site Conditions

The reference site conditions have been defined by the NGA-East Geotechnical Working Group as corresponding to shear-wave velocity V_S = 3000 m/sec and a kappa (κ) of 0.006 sec (Hashash et al. 2014; Campbell et al. 2014). The significance of the reference rock definition is that it represents the site condition for which ground motions will be predicted using GMMs, and it represents the site condition to which site amplification factors are referenced (i.e., site amplification is unity for reference rock).

1.2.4 Magnitude and Distance Ranges of Model Applicability

Final models in NGA-East are applicable to moment magnitudes (**M**) in the 4.0–8.2 range, and for sites up to 1500 km from earthquakes source, using rupture distance as a distance metric.

1.2.5 Ground-Motion Intensity Measures (GMIMs)

In NGA-East, the preferred horizontal GMIM is $RotD_{50}$ (Boore 2010). $RotD_{50}$ is the median value of resultants of two horizontal components of ground motions as computed over each angle of rotation from 1 to 180°. $RotD_{50}$ is computed independently for each oscillator period/frequency. The minimum requested GMIMs are PGA, PGV, and 5%-damped elastic PSA for oscillator periods listed in Table 1-1. Development of GMM for vertical ground motion was not in the scope of the NGA-East project. Because the value of the peak ground displacement (PGD) is very sensitive to the signal processing (filtering correction) of the ground-motion record, it is not a robust ground-motion parameter. Therefore, it is not included in the GMIMs considered in

NGA-East. This was also the case in the NGA-West2 project (Bozorgnia et al. 2014). It should be noted that the NGA-East database includes values of $RotD_{50}$ at the periods of interest. The NGA-East database is summarized in Chapter 5.

1.2.6 Interface with CEUS SSC

Input to PSHA computations requires both SSC and GMC models. As mentioned above, NGA-East is a GMC project.

The complementary CEUS-SSC Project was a separate project conducted from April 2008 to December 2011 to provide the nuclear industry with a new, regional seismic source model for use in conducting PSHA for nuclear facilities (EPRI/DOE/NRC 2012). The CEUS SSC was also conducted as a SSHAC Level 3 study. Since the NGA-East and CEUS SSC products will possibly be used together in PSHA assessments, there needed to be an interface between the two projects. A timely dialogue between key participants of both projects, as well as having common key participants between the two projects, ensured compatibility of the source characterization and ground-motion characterization. This is reflected in the implementation guidance developed by the NGA-East Project team as described in Chapters 13–15.

1.2.7 Categories of Median Ground-Motion Models

As indicated previously, the scope of NGA-East includes the development of a set of new GMMs for median and standard deviation and their associated weights in the logic-trees for use in PSHA in CENA. In the process of developing median GMMs, the project team selected and developed the following nomenclature, which is further detailed in Chapters 6-9:

- "**Candidate GMMs**": A total of 30 median GMMs were initially considered. Ten of the models are from the recent EPRI Review Project (EPRI 2013) and are listed in Chapter 7 (Table 7–1). These models are described in detail in EPRI (2013). Ultimately, none of the EPRI (2013) models was used in the development of the final GMMs; this was primarily because they have been updated and replaced by more recent models (see Chapter 7, Section 7.3). The remaining 20 candidate GMMs were developed specifically for the NGA-East project by various teams of experts in earthquake community, and full descriptions of these models are provided in a PEER report (2015a). These 20 GMMs are denoted as "candidate" GMMs.
- "**Seed GMMs**": Following the evaluation of the candidate GMMs, the TI team selected 19 GMMs for further investigation. Some of the selected GMMs required adjustments, mostly to cover the desired 1500-km-distance range. The NGA-East Technical Integrator (TI) Team extrapolated the candidate GMMs as needed. A certain level of smoothing of the extrapolated GMMs was also carried out. Some GMMS were selected for a limited range of frequencies. This set of 19 selected, extrapolated, and smoothed GMMs are denoted as "seed GMMs." This GMM set was used to generate a continuous distribution of GMMs for each GMIM (Chapter 8, Section 8.1).

- "**Sampled GMMs**": A set of 10,000 GMMs was sampled from the continuous GMM distribution developed above. These GMMs are referred to as sampled GMMs. At this stage, Sammon's maps (1969) were used to visualize the ground-motion space spanned by the seed and sampled GMMs.
- "**Final NGA-East GMMs**": The 10,000 "sampled GMMs" were "grouped' into 17 cells in the Sammon's map space. For each of these 17 cells, a process to define a representative model in each cell was defined. The 17 resulting GMMs are the final NGA-East GMMs, which are also referred to as the "representative sampled GMMs" (chapter 8). These 17 GMMs were then assigned weights (Chapter 9) by the TI team for logic tree and PSHA applications.

1.2.8 Limitations

There are several limitations in the ground-motion data recorded in the CENA. For example, these include lack of large magnitude data and very little data in magnitude-distance range for engineering applications, among other limitations. Chapter 5 of this report as well as PEER Report 2014/07 (Goulet et al. 2014) elaborate on such limitations. As a consequence of this lack of relevant data, the GMMs categorized in the previous section do not explicitly parameterize attributes such as style of faulting and effects of directivity of ground motions. The effects from seismic source depth and hanging-wall conditions are treated separately, to be applied globally to the final NGA-East GMMs. This is also the case of "adjustments" for the Gulf Coast Region (GCR). The models developed for these three types of adjustments are described in Chapter 13.

1.3 NGA-East Science Component

1.3.1 Introduction

As indicated, besides the "SSHAC component", the NGA-East project has also a science component addressing various new research topics (Bozorgnia 2008). The science component of NGA-East consisted of strongly focused efforts, with a specific objective of providing valuable input on data and models to be used in model development or to be evaluated by the TI team depending on the task. Some tasks led to the publication of reports used to inform model development. This is the case, for example, of the magnitude-area report by Somerville (2014) described in the next section. The candidate median GMMs developed under the "Science Component" of NGA-East (PEER 2015a) were considered by the TI team as proponent models, and, as such, the TI team evaluated the models, modified them (if needed), and assigned weights for integration into the final NGA-East GMC. This evaluation and integration was performed under the SSHAC umbrella.

In the following subsections, the products of the science component of the NGA-East are listed and summarized. These subsections indicate which parts of the science products have been adopted by the TI team for the SSHAC component of NGA-East.

1.3.2 Products

The following reports have addressed various science topics of the NGA-East Project. These reports are also considered as "attachments" to this report and are all available online on the PEER website (https://peer.berkeley.edu/).

1.3.2.1 PEER NGA-East Database Report

PEER Report 2014-07 (Goulet et al. 2014) serves as a documentation of the recorded ground-motion database development for the NGA-East Project. The ground-motion database includes the two- and three-component ground-motion recordings from numerous selected events (**M** > 2.5, distances up to 1500 km) recorded in the CENA region since 1988, up to December 2011. The database contains over 29,000 records from 81 earthquake events and 1379 recording stations. The events include "mainshocks," "aftershocks," and tectonic events, as well as potentially induced events (PIEs). The time series and metadata collected went through numerous rounds of quality assurance and review. The NGA-East database constitutes the largest database of processed recorded ground motions in stable continental regions (SCRs). The motivation behind the development of the recorded ground-motion database is the same as for other NGA projects (NGA-West1 and NGA-West2), which is to be used, along with other information and data, for the development of GMMs (also known as ground-motion prediction equations or GMPEs). The NGA-East ground-motion database, similar to those from the NGA-West projects, includes PGA, PGV, and PSA for the 5%-damped elastic oscillators with periods ranging from 0.01 to 10 sec. In addition, the NGA-East database includes Fourier amplitude spectra (FAS) of the processed ground motions. Chapter 5 of this report summarizes the attributes of the database. The database has been extensively used in the SSHAC component of NGA-East, notably for the evaluation of candidate and seed models, as well as in the weight assignment of the final NGA-East GMMs. Site correction to reference rock (next two sections) are discussed in Chapter 7 and were the topic of report 2015/06 described below.

1.3.2.2 PEER NGA-East Reference Rock Report: Part 1: Velocity

PEER Report 2014/11 (Hashash et al. 2014) addresses the definition of the reference P- and S-wave velocities. The significance of the reference-rock definition is that it represents the site condition for which ground motions will be predicted using GMMs. Moreover, it represents the site condition to which site amplification factors are referenced (i.e., site amplification is unity for reference rock). There are significant differences in the reference rock site conditions between active tectonic regions, such as western North America (WNA), and mid-plate regions, such as CENA. Using velocity measurements reported in the license applications at nuclear power plants, as well as published data, a set of criteria was developed to assess the presence of the reference rock site condition that is based on the seismic velocities and their gradient with respect to depth. The criteria were applied to the available profiles from which the following seismic velocities for reference rock in CENA were defined:

- $V_{s,ref}$ = 3000 m/sec or 9800 ft/sec (2700 to 3300 m/sec or 8900 to 10,800 ft/sec)
- $V_{p,ref}$ = 5500 m/sec or 1000 ft/sec (5000 to 6,100 m/sec or 16,400 to 20,000 ft/sec)

The range given for seismic velocities is based on a ±5% change in amplification using quarter wavelength theory (Hashash et al. 2014). This range is provided as guidance to regulators on whether or not site-specific site response may be required relative to past studies (where reference rock was historically defined for 2800 m/sec). The study did not find evidence for regional dependence of the reference velocities, which were derived principally from three general geographic regions: (1) the Atlantic coast; (2) the continental interior; and (3) the Appalachian Mountains. The data do not provide reference velocities for the GCR. In this region the depth to the CENA reference-rock condition is expected to be much greater than other CENA regions due to several kilometers of overlying sediments. The study did not provide a reference-rock condition for the GCR. The recommendation is to adopt a consistent reference-rock condition for the entire CENA region, as given above, and then estimate transfer functions to a softer reference condition (such as 760 m/sec) for application of the NGA-East ground-motion models. The reference-rock site conditions defined in this study represent the reference site condition in CENA in the NGA-East project.

1.3.2.3 PEER NGA-East Reference Rock Report: Part 2: Kappa

PEER Report 2014/12 (Campbell et al. 2014) presents the results of a comprehensive literature search and limited additional studies that support the recommendation of a probability distribution for the shear-wave site attenuation parameter κ_0, or site kappa, associated with a reference-rock site condition in CENA. This study was conducted as part of the Geotechnical Working Group (GWG) activities of the NGA-East. The recommended reference-rock site condition, which is documented in PEER Report 2014/11 (Hashash et al. 2014, section above) is defined as a hard-rock site with a shear-wave velocity of 3000 ± 300 m/sec. The recommended distribution of the reference-rock site kappa κ_{0ref} for this reference-rock site condition is lognormal, with a median value of 0.006 sec, an aleatory standard deviation of 0.43, and epistemic standard deviations of 0.12, in natural log units, when uncertainty in source, path, and site-amplification effects are included in simulations used to develop GMPEs, and 0.20 when they are not. This distribution is intended to represent the center, body, and range of the technically defensible interpretations of the wider scientific community as defined in the SSHAC guidelines. The reference-rock site conditions defined in this study represent the reference site condition in CENA in the NGA-East project.

1.3.2.4 PEER NGA-East Regionalization Report

In this study, documented in PEER Report No. 2014/15 (Dreiling et al. 2014), the CENA is subdivided into four regions based on the geologic and tectonic setting. The regions are the Central North America (CNA), the Appalachian Province (APP), the Atlantic Coastal Plain (ACP), and the Gulf Coast/ Mississippi Embayment Region (GCR). Each region is described by a statistically representative crustal seismic velocity-depth structure and Q-factor model. The crustal structural models are for very hard rock conditions and do not include any sediments.

The purpose of this study was to evaluate similarities and differences in attenuation for these regions and to assess whether regions needed to remain separate or if they could be grouped based on their attenuation properties. This was achieved through a series of ground-motion simulations. Seismic-wave propagation was simulated for earthquakes at focal depths of 5, 10,

20, and 30 km, using two different ground-motion simulation codes. Synthetic time series and the 5% damped PSA provide insight into the attenuation of ground motions that are typical for each region. The calculated PSA covers a hypocentral distance range of 7.5–500 km and oscillator frequencies ranging from 0.5 to 20 Hz. Spectral accelerations were compared both within and between regions. The CNA is the biggest region geographically and offers the largest variety of crustal seismic velocity-depth structures associated with the unique geologic evolution of its sub-regions.

The CNA is defined as the base region; it is used for both comparisons and to estimate a reference range of within-region variability. After generalizing the 417 profiles available for CNA into one representative profile (CNA_{Rep}), ground motions were calculated for the four aforementioned focal depths. The within-region variability was also assessed using ground-motion simulations for a selected set of 18 alternative velocity models developed for the region (CNA_{Alt}). The PSA values calculated for CNA_{Rep} were compared to the PSA values for the 18 alternative crustal structures, CNA_{Alt}. To determine which of the four regions should be assigned to a common attenuation group, we compared the ACP, APP, and GCR to the CNA base region. Statistical distributions (histograms) of the PSAs for specific distance and frequency bands were used to investigate if there were significant differences between the regions. Additional analysis tools, such as moving window average of PSA versus distance for specific frequency bands, were also used in these comparisons. This analysis demonstrates that there are two distinct attenuation groups:

- Mid-continent region (MCR, GROUP 1 in the 2014-15 report): Central North America, Appalachians, and the Atlantic Coastal Plain
- Gulf Coast Region (GCR, GROUP 2 in the 2014-15 report): Gulf Coast including the Mississippi Embayment

The GCR was found to clearly belong to a separate attenuation group. This result is in agreement with previous analyses that have found that the GCR has unique attenuation characteristics.

This regionalization scheme was considered by the TI team as a proponent regionalization model as it is elaborated in Chapter 4 of the current report.

1.3.2.5 PEER NGA-East Magnitude-Area Report

PEER Report 2014/14 (Somerville 2014) describes the development of scaling relations between seismic moment and rupture area of earthquakes in stable continental regions (SCRs). The report reviews existing relations, develops new relations, and compares the new relations with the existing relations. It also compares the scaling relations of SCR earthquakes with those in tectonically active continental regions (ACR). Three different methods of estimating rupture area—based on aftershocks, slip models, and duration—were used to analyze the relation between seismic moment and rupture area, using earthquake source parameters compiled from published literature. For each category of data, the relations obtained were not significantly different from those obtained by constraining them to be self-similar. Accordingly, these self-similar relations were adopted in this study. The stress drops corresponding to these scaling relations range from 51 to 86 bars, with an average of 65 bars. This value is comparable to the

value of 58 bars obtained by Leonard (2010). Because Leonard (2010) did not document his data and used an undifferentiated mixture of different ways of measuring fault area, the relation that he developed is less soundly based than that developed in this study. However, the two relations are not significantly different, and the Leonard (2010) relations have the advantage of having been derived in a self-consistent manner for a wide range of earthquake categories, including crustal earthquakes in tectonically active regions. Consequently, it is recommended that the Leonard (2010) scaling relations for SCR earthquakes be used for the NGA-East Project. To a first approximation, the results of this study indicate that the rupture areas of SCR earthquakes are about half those of ACR earthquakes, and their stress drops are about 2.8 times higher. Allmann and Shearer (2009) find less of a difference, presumably because their intraplate category includes some earthquakes that the NGA-East Project would assign to ACR instead of SCR. Their study indicates that the rupture areas of intraplate earthquakes are about two-thirds those of ACR earthquakes, and their stress drops are about two times higher.

The magnitude-area scaling documented in PEER report 2014/14 was used in the finite-fault simulations for which the results were shared with the candidate GMM developers. As indicated previously, the candidate GMMs were subsequently used by the TI team to select and develop seed GMMs, and to ultimately develop the final NGA-East GMMs.

1.3.2.6 PEER NGA-East Median GMM Report

PEER Report 2015/04 (PEER 2015a) documents the development of 20 new median candidate GMMs. As indicated previously, the 20 candidate GMMs were developed specifically for the NGA-East project by various teams of experts in earthquake community. Models for standard deviations of ground motions were developed through a separate set of tasks within NGA-East and were published separately. The new 20 GMMs have been developed using various tasks previously completed in NGA-East: notably, the path regionalization, finite-fault simulations, and database development tasks. This report consists of eleven chapters. Each chapter documents the work of one GMM developer team and may include multiple new GMMs. The 20 GMMs described in this report cover a range of alternative approaches for modeling ground motions, building on empirical relations for CENA and WNA, using recorded ground motions and collected intensity data, and incorporating point-source and finite-fault simulations.

These 20 new candidate GMMs were reviewed and adjusted (if needed) to select and develop the seed GMMs as described in Chapter 7 of this report. As indicated before, the seed GMMs were used to develop the Final NGA-East GMMs as described in Chapter 8 of this report.

1.3.2.7 PEER NGA-East Adjustments to Median GMMs Report

PEER Report 2015/08 (PEER 2015b) documents a series of adjustments developed for the median GMMs as part of the NGA-East project. The adjustments to median GMMs are necessary so that NGA-East (1) is applicable to rupture distances in the range from 0 to 1500 km; (2) allows source depth effects to be incorporated; and (3) is applicable to the vast CENA region to include the Gulf Coast and the Mississippi Embayment. The three corresponding adjustment models are documented in this report. This report can be considered as a supplemental report to the PEER Report 2015-04 (PEER 2015a).

The adjustments to the median GMMs were considered by the TI team to adjust the candidate GMMs. As indicated previously, the candidate GMMs were evaluated to select and develop the seed GMMs. Subsequently, the seed GMMs were used by the TI team to develop the final median GMMs for NGA-East.

1.3.2.8 PEER NGA-East Sigma Report

In the study documented in PEER Report 2015/07 (Al Atik 2015) the recorded ground-motion data from CENA were used to analyze the components of ground-motion variability in CENA. Trends of ground-motion variability with parameters such as magnitude, distance, and V_{S30} were analyzed and compared to trends of ground-motion variability in other regions, particularly the western United States (WUS) using the NGA-West2 dataset. The CENA dataset is limited in magnitude range to small-to-moderate magnitudes and in frequency content to frequencies between 1 and 10 Hz due to the bandwidth limitations of the recordings. Therefore, standard deviation models developed using the CENA ground-motion data cannot be reliably extrapolated to large magnitudes and to frequencies outside of 1 to 10 Hz. As a result, standard deviation models from other regions such as WUS and Japan were used to inform the extrapolation of CENA standard deviations and overcome data limitations. Candidate models for between-event standard deviation (τ), single-station within-event standard deviation (ϕ_{SS}), and site-to-site variability (ϕ_{S2S}) were developed for CENA. In turn, these models were combined to develop single-station sigma (σ_{SS}) and ergodic sigma models for CENA.

Sigma models developed and documented in PEER report 2015-7 (Al Atik 2015) were considered by the TI team as proponent sigma models as elaborated in Chapters 10 and 11 of this report.

1.3.2.9 PEER NGA-East Report on Site-Correction Factors

In the PEER Report 2015/06 (Boore 2015) adjustment factors that can be used to convert ground-motion intensity measures at sites with V_{S30} = 760 m/sec and V_{S30} = 2000 m/sec to a reference-rock site, defined as one with V_{S30} = 3000 m/sec, were provided as tables: (1) for moment magnitudes from 2 to 8; (2) rupture distances from 2 km to 1200 km; (3) response spectra at periods from 0.01 sec to 10.0 sec; and (4) PGA and PGV. Ten velocity models used in ground-motion studies in CENA with V_{S30} values very close to 760 m/sec were considered, and adjustment factors were provided for two of those models that effectively span the range of models; for the convenience of the user, adjustment factors were provided for an average of a representative set of models with V_{S30} = 760 m/sec. For models with this velocity, adjustment factors were provided for four values of the diminution parameter κ, ranging from 0.005 sec to 0.030 sec. The adjustment factors were based on stochastic-method simulations of ground motion.

The adjustments developed under this task were used to correct the data to the reference rock condition and subsequently used for the computation of residuals as discussed in Chapters 7 and 9 of this report.

1.4 Organization of this Report

This report is the final project report of the tasks in the NGA-East project that fall under the SSHAC Level 3 framework. The report contains 15 chapters as follows.

- Chapter 1, this chapter, provides an overview of the NGA-East project as a combination of a science-based component and of a formal SSHAC Level 3 study.
- Chapter 2 provides a brief overview of the SSHAC process, along with the organization of the NGA-East project under the SSHAC Level 3 framework.
- Chapter 3 provides a summary of the Central and Eastern United States Seismic Source Characterization (CEUS SSC) model developed by EPRI/DOE/NRC (2012), and describes the basis for selection of the range of magnitudes and distances for which the NGA-East GMC and the individual GMMs were developed.
- Chapter 4 presents an overview of the regionalization of ground motions in CENA; proponent regionalization models are summarized and evaluated.
- Chapter 5 summarizes the databases available for the NGA-East project. A key component of the chapter is to summarize the development of the empirical NGA-East ground-motion database of events recorded in CENA. The NGA-East database was used by numerous working groups, and individual researchers and practitioners who worked on various aspects of the NGA-East project. The chapter also summarizes the NGA-West2 database as well as a database of simulated records, which were made available to the NGA-East researchers and participants.
- Chapter 6 presents a summary of the methodology used by NGA-East for the ground-motion characterization in CENA. It summarizes the process selected for quantifying the epistemic uncertainty, which is an important part of the SSHAC process.
- Chapter 7 describes the candidate GMMs, including those that were developed by the members of the NGA-East GMM working group, that were considered for use for the next step as “seed” GMMs in populating the ground-motion space. Site correction issues are also discussed in this chapter.
- Chapter 8 elaborates on the process involved in three steps in the NGA-East methodology, as briefly described in Section 1.2.7 of this chapter: (a) development of continuous distributions of GMM using the “seed” models from Chapter 7, (b) visualization of the ground-motion space via Sammon’s maps, and (c) re-discretization of the ground-motion space.
- Chapter 9 discusses the median ground-motion logic tree, which consists of 17 branches, each corresponding to a GMM. This chapter only addresses the weights for the final 17 median models.

- Chapter 10 reviews the candidate proponent ground-motion aleatory variability models and provides the framework used to develop standard deviation models for CENA.
- Chapter 11 presents the evaluation of the candidate models for the between-event variability, the single-station within-event variability, and the site-to-site variability (τ, ϕ_{SS}, and ϕ_{S2S}, respectively). In this chapter, logic trees are developed for each of τ, ϕ_{SS}, and ϕ_{S2S}, and the assigned weights are discussed.
- Chapter 12 presents PSHA results of NGA-East GMC as compared to those of the EPRI (2013) GMC. Additionally, the EPRI (2013) GMMs were used as "seed" models for the NGA-East methodology and the results were compared with those of the NGA-East GMC and EPRI (2013).
- Chapter 13 presents the development of adjustment models to be applied to the 17 final median GMMs to address: (1) adjustments for the Gulf Coast region and (2) source-depth adjustments; and discusses hanging-wall adjustments.
- Chapter 14 presents the implementation of the Gulf Coast Region and source-depth effects in PSHA calculations, provides implementation guidance for practitioners, and compares PSHA results computed with the full NGA-East GMM to those computed using the EPRI (2013) GMM.
- Appendices: several appendices are used to further document the work described in the various chapters. Although the material provided in appendices is essential for complete documentation, it was deemed to be distracting to the main text if it had been included directly in the chapters.
    - Appendix A supports Chapter 1 and contains the project plan and the related PPRP correspondence.
    - Appendix B supports Chapter 2 and contains the workshops summaries and the related PPRP correspondence. It also includes the PPRP comments and the TI team responses to report drafts.
    - Appendix C supports Chapter 5 and consists of electronic database documents.
    - Appendix D supports Chapter 7 and the development and evaluation of candidate GMMs.
    - Appendix E supports Chapters 8 and 9 and provides additional evaluations and sensitivity analyses related to the epistemic uncertainty in median GMMs.
    - Appendix F supports Chapters 10 and 11 on the aleatory variability quantification.

- Appendix G supports Chapter 13 and document additional investigation of depth effect issues.
- Appendix H provides the hazard input document (HID) describing the implementation of the full GMM.
- Appendix I provides tables and plots of hazard results for the seven CEUS demonstration sites, as completed using Appendix H.

- Attachments: all the PEER reports described in Section 1.3.2 are included as "attachments" to this report.

1.5 References

Al Atik, L. 2015. "NGA-East: Ground Motion Standard Deviation Models for Central and Eastern North America." PEER Report No. 2015/07, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Allmann, B.P., and P.M. Shearer. 2009. "Global Variations of Stress Drop for Moderate to Large Earthquakes." J. Geophys. Res., 114, B01310, doi:10.1029/2008JB005821.

Boore, D.M. 2010. "Orientation-Independent, Nongeometric-Mean Measures of Seismic Intensity from Two Horizontal Components of Motion." Bull. Seismol. Soc. Am., 100(4): 1830–1835.

Boore, D.M. 2015. "Adjusting Ground-Motion Intensity Measures to a Reference Site for which V_{S30} = 3000 m/sec," PEER Report No. 2015/06, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Bozorgnia, Y. 2008. "A Roadmap for the Next Generation Attenuation models for Central and Eastern North America (NGA-East)." Letter report to the U.S. NRC, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Bozorgnia, Y., N.A. Abrahamson, L.A. Atik, T.D. Ancheta, G.M. Atkinson, J.W. Baker, A. Baltay, D.M. Boore, K.W. Campbell, B.S.J. Chiou, R. Darragh, S. Day, J. Donahue, R.W. Graves, N. Gregor, T. Hanks, I.M. Idriss, R. Kamai, T. Kishida, A. Kottke, S.A. Mahin, S. Rezaeian, B. Rowshandel, E. Seyhan, S. Shahi, T. Shantz, W. Silva, P. Spudich, J.P. Stewart, J. Watson-Lamprey, K. Wooddell, and R. Youngs. 2014. "NGA-West2 Research Project." Earthq. Spectra, 30: 973–987.

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Clu, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. 1997. "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts." Report NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington D.C.

Campbell, K.W., Y.M.A. Hashash, B. Kim, A.R. Kottke, E.M. Rathje, W.J. Silva, and J.P. Stewart. 2014. “Reference-Rock Site Conditions for Central and Eastern North America: Part II – Attenuation Definition.” PEER Report No. 2014/12, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Dreiling, J., M.P. Isken, W.D. Mooney, M.C. Chapman, and R.W. Godbee. 2014. “NGA-East Regionalization Report: Comparison of Four Crustal Regions within Central and Eastern North America using Waveform Modeling and 5%-Damped Pseudo-Spectral Acceleration Response.” PEER Report No. 2014/15, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

EPRI. 2013. “EPRI 2004-2006 Ground-Motion Model (GMM) Review Project 3002000717.” Final Report. June 2013, Electric Power Research Institute, Palo Alto, CA.

EPRI/USDOE/USNRC. 2012. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Goulet, C.A., N.A. Abrahamson, and Y. Bozorgnia. 2011. “NGA-East Final Project Plan.” Pacific Earthquake Engineering Research Center, July 2011.

Goulet, C.A., N.A. Abrahamson, P.G. Somerville, and K.E. Wooddell. 2015. “The SCEC Broadband Platform Validation Exercise: Methodology for Code Validation in the Context of Seismic-Hazard Analyses.” Seismol. Res. Lett., 86 (1): 17–26.

Goulet, C.A., T. Kishida, T.D. Ancheta, C.H. Cramer, R.B. Darragh, W.J. Silva, Y.M.A Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. 2014. “PEER NGA-East Database.” PEER Report No. 2014/17, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Hashash, Y.M.A., A.R. Kottke, J.P. Stewart, K.W. Campbell, B. Kim, E.M. Rathje, and W.J. Silva. 2014. “Reference Rock Site Condition for Central and Eastern North America.” Bull. Seismol. Soc. Am., 104: 684–701.

Leonard, M. 2010. “Earthquake Fault Scaling Self-Consistent Relations of Rupture Length, Width, Average Displacement and Moment Release.” Bull. Seismol. Soc. Am., 100: 1971−1988.

NRC. 2012. “Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies.” Technical Report NUREG-2117, U.S. Nuclear Regulatory Commission, Washington D.C.

PEER. 2015a. “NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

PEER. 2015b. “NGA-East: Adjustments to Median Ground-Motion Models for Central and Eastern North America.” PEER Report No. 2015/08, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

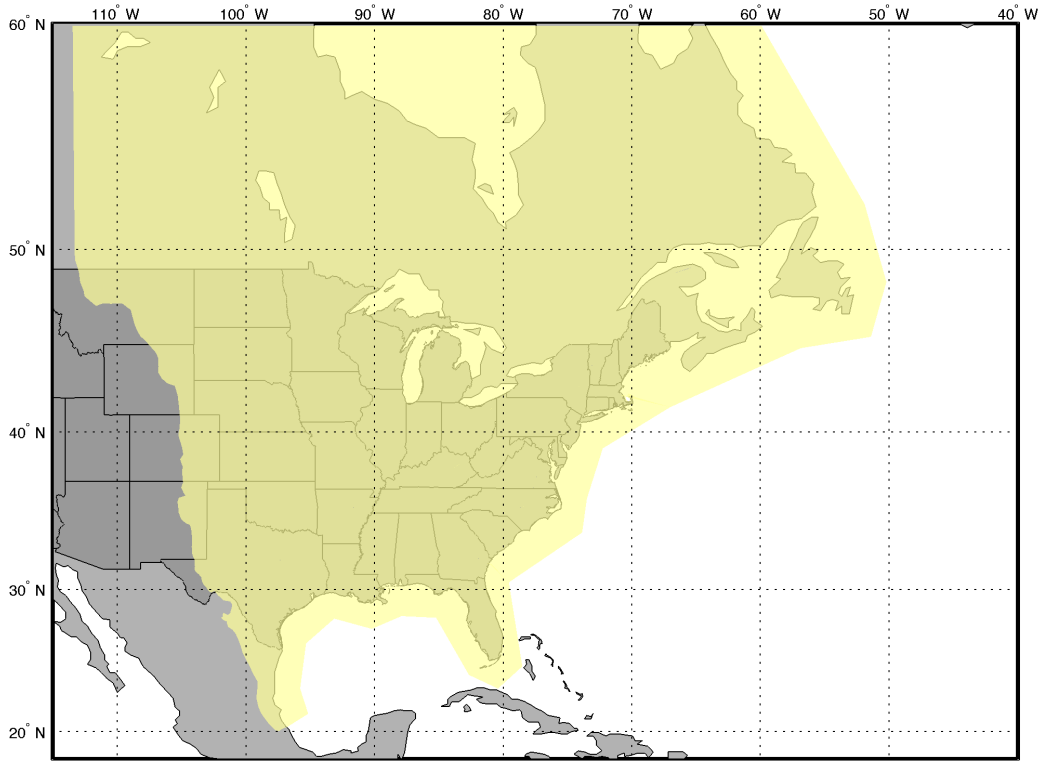
Sammon, J.W. 1969. “A Nonlinear Mapping for Data Structure Analysis.” Trans. Comp. C13(5): 401-409.

Somerville, P.G. 2014. “Scaling Relations between Seismic Moment and Rupture Area of Earthquakes in Stable Continental Regions.” PEER Report No. 2014/14, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Table 1–1Selected 5%-damped PSA periods, T (and frequencies, f) for NGA-East. Peak ground acceleration (PGA) and peak ground velocity (PGV) are also GMIMs of interest.

T (sec)	F (Hz)
0.01	100
0.02	50
0.025	40
0.03	33.33
0.04	25
0.05	20
0.075	13.33
0.1	10
0.15	6.67
0.2	5
0.25	4
0.3	3.33
0.4	2.5
0.5	2
0.75	1.33
1	1
1.5	0.67
2	0.5
3	0.33
4	0.25
5	0.2
7.5	0.13
10	0.1

Figure 1–1 NGA-East CENA study region in yellow. The study area extends into the oceans to indicate that earthquake sources could be considered offshore. The model is meant to be applicable to the shaded land areas in the U.S. and Canada only. The western boundary was borrowed from Dreiling et al. (2014). Discussion of regionalization is provided in Chapter 4.



2. Implementation of SSHAC Level 3 Process and NGA-East Project Organization

2.1 SSHAC Level 3 for NGA-East Project

As indicated in Chapter 1, the NGA-East project has two “components”: (1) a set of scientific research tasks, and (2) a model-building component, to develop a new ground-motion characterization (GMC) model under the Senior Seismic Hazard Analysis Committee (SSHAC) Level 3 framework. The tasks under Component (2) include taking advantage of various research findings and products developed under research Components (1) of the project. The SSHAC Guidelines are detailed in the NUREG/CR-6372 document (Budnitz et al. 1997) and in NUREG-2117 (NRC 2012). The SSHAC assessment process can be used in the development probabilistic seismic hazard analysis (PSHA) studies, as elaborated next.

A typical PSHA study requires two important components: a seismic source characterization (SSC) and a GMC. For applications to design, the NGA-East GMC is to be used in conjunction with the Central and Eastern United States (CEUS) SSC products (EPRI/DOE/NRC 2012) The CEUS SSC project was also conducted under the SSHAC Level 3 framework; thus, having the GMC component of the NGA-West2 under SSHAC Level 3 framework will ensure that both projects are of the same high-quality level for the end-uses, especially for nuclear facilities.

The fundamental goal of a SSHAC assessment process is to carry-out properly and document completely the activities of evaluation and integration, defined as:

- **Evaluation:** The consideration of all the data, models, and methods proposed by the larger technical community that are relevant to the hazard analysis.
- **Integration:** Representing the center, body, and range of technically defensible interpretations in light of the evaluation process.

The process has the advantages to increase the assurance that uncertainties have been captured. Additionally, from a practical point of view, the SSHAC assessment process provides (NRC 2012): (1) determination of more accurate and consistent assessments of seismic hazard and the associated uncertainty; (2) standardization and complete and transparent documentation of the assessment process undertaken, the input data, and the basis for the resulting model and findings; (3) increased regulatory assurance based on the transparency of the study's technical basis; and (4) the increased longevity of a study as a result of the ability to assess new data against the existing model and its basis and assumptions. All of these goals lead to greater regulatory assurance and stability.

2.2 SSHAC Level 3 Process for NGA-East

The SSHAC guidelines define four study levels, with each higher level corresponding to an increase in complexity. Higher study levels are associated with a higher confidence that the center, body, and range of technically defensible interpretations of the available data, models, and methods have been captured in the final products. The SSHAC Level 3 was selected as appropriate to ensure the stability and transparency of the NGA-East products given the complexity, importance, and regulatory concerns associated with the study.

The SSHAC Level 3 assessment process requires a level of documentation and review that is much more demanding than what was carried out in other comparable research projects, such as NGA-West for example. This implies larger resources in both time and capital investment. In an effort to optimize the needs of the different agencies with the available resources, the SSHAC Level 3 study was assigned to the GMC model building tasks associated to Nuclear Regulatory Commission (NRC) and the Department of Energy (DOE) objectives only. The tasks that are part of the scientific research phase have been addressed as typical research tasks coordinated by PEER and documented in PEER reports (see list in Chapter 1).

In a SSHAC Level 3 assessment project, quantification of epistemic uncertainty is a key component of the process, as the epistemic uncertainty of ground motion models is a significant contributor to the final hazard, especially in Stable Continental Regions (SCRs). In this regard, the Technical Integrator (TI) Team is responsible for the documentation of the technical bases for accepting, rejecting, and assigning weights to models.

2.3 SSHAC Workshops

SSHAC workshops play a vital role in the SSHAC Level 3 assessment process. The SSHAC workshops provide opportunities for key interactions to occur; for models and interpretations to be presented, debated, and defended; and for sponsors and reviewers to observe and comment on the progress being made on the study. For a SSHAC Level 3 assessment process, there are three mandatory SSHAC workshops or workshop themes, each serving a specific purpose. The objectives and goals of each SSHAC workshop are briefly described below for convenience. The NRC document (2012) provides more details.

First and foremost, the workshops are held to provide information to assist the TI team and the project team in their technical assessments (see next section for the definition of various levels of the project team). The workshops were open to the public, and videos of the presentations and discussions were posted on the PEER website following each workshop (links in Appendix B).

The goals and attributes of the required SSHAC workshops are organized into three themes that are summarized below for convenience. The term "theme" is used because the specific NGA-East SSHAC workshops may cover more than a single SSHAC workshop element. The summary and scopes of the *specific workshops* organized and held for the NGA-East project are listed in the next section of this report and elaborated in Appendix B of this report.

2.3.1 Workshop Theme 1: Significant Issues and Data Needs

The goals of this workshop are as follows: (1) to identify the technical issues of highest significance to the hazard analysis; and (2) to identify the available data and information that are needed to address those issues. The discussions of the available data should be made by a series of presentations by resource experts who have developed specific datasets.

From the standpoint of the SSHAC assessment process, the evaluation of the data for use in the hazard analyses is led by the TI team. In the case of the NGA-East project, a significant amount of technical development work was required and led by the Working Groups (WGs). As

a result, the WGs support the TI team by performing a number of critical evaluations and proposing a variety of technical choices.

2.3.2 Workshop Theme 2: Proponent Discussions of Alternative Interpretations

The goals of Workshop 2 are as follows: (1) to present, discuss, and debate alternative viewpoints regarding key technical issues; (2) to identify the technical bases for the alternative hypotheses and to discuss the associated uncertainties; and (3) to provide a basis for the subsequent development of preliminary hazard models that consider these alternative viewpoints. The workshop also provides an opportunity to review the progress being made on the database development and to elicit additional input, as needed, regarding this activity.

A key attribute of this workshop is the discussion and debate of the merits of alternative viewpoints regarding key technical issues. Proponents and Resource Experts [see SSHAC (1997) to how roles are defined] in present their interpretations and the data supporting them. Alternative viewpoints are to be juxtaposed and facilitated discussions to be focused on implications to the inputs to the hazard analysis (not just on scientific viability) and on uncertainties (e.g., what conceptual models would capture the range of interpretations and what weights should be applied). The Proponent Experts need to be prepared to discuss the uncertainties in their interpretations, the strengths and weaknesses in their arguments, and their view of where their interpretations lie with regard to the larger technical community. When organizing the proponent workshops, the TI team leads circulated the proposed list of participants to ensure that the agenda incorporates all viable views and hypotheses.

2.3.3 Workshop Theme 3: Presentation and Feedback on Proposed Models

Following the workshop (or workshops), which is focused on proponent discussions of alternative interpretations, the TI team members develop their preliminary models, and preliminary calculations and sensitivity analyses are conducted. The goal of Workshop 3 is to present and discuss the preliminary models and calculations in a forum that provides the opportunity for feedback to the TI team. Feedback can be given in the form of input by technical specialists or in the form of hazard results and sensitivity analyses to shed light on the most important technical issues. The feedback gained at this workshop ensures that no significant issues have been overlooked and allows the TI team to understand the relative importance of their models, uncertainties, and assessments of weights. At this time, the Participatory Peer Review Panel (PPRP) is invited to interrogate the TI team on the models and weights they are proposing. This information provides a basis for the finalization of the models following the workshop.

The workshop typically consists principally of the TI team presenting their preliminary models, with particular emphasis on the manner in which alternative viewpoints and uncertainties have been captured. The technical bases for the assessments and weights are described to allow for a reasoned discussion of the constraints provided by the available data. The invited experts are responsible to question and probe aspects of the preliminary model to understand the manner in which the range of technically defensible interpretations has been captured.

2.4 Specific SSHAC Workshops Organized in the NGA-East Project

This section provides a short summary of each of the SSHAC workshops held after the NGA-East project was assigned as SSHAC Level 3. Due to the project evolution and the large number of research tasks initiated as part of the "science part" of NGA-East, the SSHAC workshops were also an opportunity to present research results and discuss various issues that were deemed important and relevant to the development of ground-motion models (GMMs). This strategy allowed the project to increase the level of transparency throughout the process and to keep everyone (SSHAC participants and researchers alike) informed and involved in the discussions. This is evidenced by the range of topics covered in the workshops' presentations linked below, which is broader than expected for a typical SSHAC workshop.

In the following sections, a short introduction and the context of each workshop are provided. Appendix B of this report provides links to the NGA-East workshops web pages, which include the workshop agenda and links to all the presentations. Appendix B also provides letter correspondences between the PPRP and TI team.

2.4.1 Workshop 1, November 15-18, 2010

This workshop followed the SSHAC Theme 1: Critical Issues and Data Needs (see Section 2.3.1 of this report). The workshop was also the opportunity to further discuss the project organization and to initiate the development of the official project plan (Appendix A), which built extensively on the original "Roadmap" document (Bozorgnia 2008).

The workshop started with an overview of the project, as well as an introduction to the SSHAC process. The morning session of the first day highlighted critical issues identified for the project, such as geometrical spreading in Central and Eastern North America (CENA), variability and magnitude dependence of the stress parameter, regionalization, and the inclusion of simulations. The afternoon session of the first day was about the NGA East database, with presentations about the current state of the database—highlighting different aspects such as number of data, processing of recordings, and metadata uncertainty—and the consideration of possibly adding of new data, both from CENA and from other SCRs. Discussions revolved around the input parameters to simulations, in particular for large magnitudes, and geometrical spreading at short distances. Other discussion points involved the tails of the ground-motion distribution and the verification of GMMS.

The morning session of the second day focused on site effects issues, with presentations on reference-rock conditions and issues regarding the inclusion of kappa as a parameter in NGA-East. Discussion points were the range of both reference-rock shear-wave velocity and kappa conditions, and the estimation of site conditions for sites with observed ground motion.

The second day continued with sessions on simulations, with one part highlighting issues for point-source models [mainly informed by experience from western North America (WNA)]. This was followed by an overview of finite-fault simulations, which presented different methods and their evaluation based on observed data. Main discussion points were about the stress parameter in point-source simulations, its range, magnitude, and depth dependence.

The third day started with an investigation of CENA regionalization based on crustal structure and the inclusion of point-source models using different attenuation models. The afternoon

session focused on the range of the stress parameter as input for point source models. Discussions revolved around the inclusion of the path and source parameters (and their correlation) in point-source simulations.

The fourth day's morning session was about inputs to finite-fault models and featured presentations on different modeling approaches. Main discussion points were the range and correlation of different input parameters, as well as validating simulation methods against observed data.

The focus moved to aleatory variability, with an overview of available data/data needs regarding the estimation of single-station sigma. The effect of aftershocks/swarms on sigma were discussed, as well as the possible inclusion of variability from simulations. The workshop concluded with an overview of the approach/status of work to modeling vertical ground motions. Finally, a presentation of evaluation methods/metrics for finite-fault simulations was given in the last part of the morning session.

2.4.2 Workshop 1B-2A, October 11–13, 2011

This workshop followed the SSHAC Theme 1: Critical Issues and Data Needs and initiated discussions on Theme 2: Proponent Discussions (see Sections 2.3.1 and 2.3.2 of this report).

The workshop started with an overview of preliminary hazard feedback results using point-source stochastic models, showing the sensitivity of hazard to different input parameters, such as the median stress parameters, duration model, and geometrical spreading. The afternoon session covered the status of the NGA-East database and data from SCRs. Discussions revolved around how the data from other regions can be incorporated into NGA-East. This was followed by a presentation on the 2011 Mineral, Virginia, earthquake, and its comparison to other existing CENA data. The first day concluded with the proposed reference-rock shear-wave velocity and kappa values, and models for simple corrections of recorded motions to reference-rock conditions.

The morning session of the second day focused on regionalization, with presentations on different aspects (source, path) and the identified regions. Discussion points were empirical evaluation of ground motions from different regions, the regional differences of median stress parameter, and the inclusion of depth/style-of-faulting into GMMs. Subsequently, hazard feedback analyses with respect to path effects were shown.

The afternoon session of the second day featured proponent median GMMs, using different approaches (point-source stochastic simulations, hybrid empirical, empirical). Inputs to all methods were discussed, and the question of testing the different models was approached. This was followed by a presentation of a study regarding single-station sigma for Switzerland; it was discussed whether the available sigma models cover the range of interpretation allowed by available data.

The final day of the workshop was devoted to finite fault simulations. The morning session presented results of validation from different simulation methods against observed events. Discussion revolved around extending the set of events that are used for validation and the treatment of site effects. This was followed by a presentation/discussion of magnitude-area

relationships, which compared CENA data against available models. Discussion points were the range of epistemic uncertainty and the treatment of aleatory variability. The afternoon session was about inputs to different finite-fault simulation methods, and the question how to modify available GMMs to accommodate a different value of kappa. The workshop concluded with presentations about inputs (in particular regarding the stress parameter) to point-source stochastic models.

2.4.3 Workshop 2B, July 14-16, 2014

This workshop followed the SSHAC Theme 2: Proponent Discussions (see Section 2.3.2 of this report).

This workshop was focused on proponent discussions of candidate models. It also summarized the conclusions of various science tasks, including the development of the ground-motion database, the regionalization of path effects, the status of finite-fault simulations and update on the modeling of standard deviation.

The morning session of the first day focused on the development of median response spectral estimates using a new approach, which is based on the combination of a Fourier spectral model and a duration model through random vibration theory (RVT). In particular, this approach was called for a consistent, calibrated duration, and initial models for both the Fourier spectrum and the duration were shown. In addition, investigations into kappa, the parameter controlling the high-frequency spectrum, and issues relating to its estimation from CENA stations, were presented. Discussions focused on depth scaling issues for the regression and the problem of extrapolation, in particular for the duration model, to large magnitudes. The issue of consistency in the complete model building for the duration model was also raised.

The afternoon session of the first day focused on database issues. The status of source, site, and regionalization tasks were presented. Discussion points were in particular the issue of sites with poorly constrained predictor values (V_{S30}) and their inclusion in the regressions, as well as the extrapolation to hard-rock conditions. In addition, the robustness of the conclusion that there are only two main regions was discussed. Furthermore, discussion focused on the stress parameter and its depth dependence, in particular, the regional differences between the depth dependence.

The morning session of the second day started with a discussion on epistemic uncertainty with respect to median approaches, and an overview of previous approaches was presented. Then, the approach used in the Southwestern United States Ground Motion Characterization SSHAC Level 3 Project (SWUS) was presented (based on a continuous distribution of median predictions), and its application to NGA East was discussed. Discussions focused on the method (Sammon's maps), including redundant models, and the inclusion of simulation results.

Focus shifted to finite-fault simulations and their validation. Different simulation methods and their evaluation against data were presented. Discussion points were some of the events that the simulations were compared against (in particular, the Saguenay event), as well as how the simulations can be incorporated into NGA-East. In addition, the input parameters for the forward simulation runs were discussed.

2.4.4 Workshop 3A-2C, October 29-30, 2014

This workshop followed the SSHAC Theme 2: Proponent Discussions and Theme 3: Feedback Analyses (see Sections 2.3.2 and 2.3.3 of this report).

The key new element for this workshop was the summary of new GMMs developed as part of the GMM Working Group. Preliminary versions of those models were presented. This workshop focused on proponent discussions of candidate models and their impact on hazard.

The morning session of the first day presented some updates regarding the NGA-East database, in particular, advances regarding assigning metadata (V_{S30}, kappa) to different stations. The rest of the first day focused on proponent median GMMs.

Different models were proposed, based on different methodologies (for example, stochastic point-source and referenced empirical). This was carried on in the morning session of the second day, which also featured an overview of the proposed hanging-wall model, adopted from the SWUS project, and an introduction to RVT used in one of the median proponent models. The different approaches to median GMMs were discussed.

In the afternoon session of the second day presented a comparison of the different proponent models, both in terms of scaling differences and in terms of their distance in ground-motion space. This then was followed by a presentation of the NGA-East approach to capturing epistemic uncertainty, based on an underlying probability distribution. The initial set of results using the proponent models were presented. Discussions followed about ensuring a physical spectral shape, about scenarios underlying the projection to two dimensions, and how to measure similarities between different models.

The workshop concluded with the approach of the TI teams to models regarding aleatory variability. The general approach, the underlying data for NGA-East, and existing models (from shallow active tectonic regions) were presented. Discussions focused on differences between tau models relative to those from NGA-West2.

2.4.5 Workshop 3B, March 4–5, 2015

This workshop followed the SSHAC Theme 2: Proponent Discussions and Theme 3: Feedback Analyses (see Section 2.3.3 of this report)

The focus of the workshop was to review the proponent models for both the median and the aleatory variability, and preliminary hazard feedback regarding those models. The first day started with a presentation of the median proponent GMMs that were considered for inclusion in NGA-East. Then, an overview of the NGA-East approach for integration of the median GMMs was given, and preliminary results using the NGA-East median GMMs were presented. Discussion revolved around various details of the NGA-East integration approach, such as the physicality of sampled models, and the selection of models on the two-dimensional projection of ground-motion space. In addition, it was suggested to apply the NGA-East process for median models to the proponent models that were used in the EPRI project to obtain a comparison of the approaches used in both projects.

The morning session of the second day finished the presentation of median models, with an overview of approaches to assigning weights to the median models. Different possibilities of

weighting the models based on their fit to data (e.g., residuals and likelihood), and their range of covered probability density were summarized.

Focus then moved on to models for the standard deviation. The CENA data, which can be used to evaluate the models, was presented. The resulting models for the different components of aleatory variability (based on CENA) were compared to standard deviation model for shallow active tectonic regions (NGA-West2, Japan). The dependence of the components on magnitude and distance was investigated, and challenges of building a model based on limited CENA data were discussed. The preliminary logic tree for the various components of standard deviation was presented, together with an investigation into their uncertainties, which is important for assigning weights.

2.4.6 Workshop 3C, June 17–18, 2015

This workshop followed the SSHAC Theme 3: Feedback Analyses (see Section 2.3.3 of this report).

The workshop's focus was on the proposed NGA-East models for the median predictions and the standard deviations. The workshop started with the median models. First, the evaluation of available median models used in the integration of NGA-East was presented. Then, the extrapolation to large distances for these models in a reliable, consistent fashion was presented. The integration of the models via the NGA-East process was presented next, with a focus on the reasoning behind the process. This involved an overview of the complete approach, where the first part reiterated the idea of a continuous distribution of median predictions that is essential to the NGA-East characterization of median models. The focus on median models concluded with a presentation on the process applied to the NGA-East seed models, which covered the visualization of the ground-motion space covered by the NGA-East distribution and the discretization into a manageable subset of models. This part also covered the weight assignment approach, based on the probability density covered by each model and the respective fit to CENA records. Discussion focused on the selection of underlying magnitude and distance scenarios for the visualization, and on improved method for displaying the weights.

The rest of the first day was devoted to the proposed models for the standard deviations. The underlying CENA data was presented and approaches to model building for the different components of aleatory variability were laid out. The uncertainty of the different component models, important for the building of the logic tree, was discussed.

The second day dealt with adjustments to the NGA-East median models, in particular, regional adjustments for the Gulf Coast, adjustments for varying source depth, and hanging-wall adjustments. The presentation on source-depth effects showed differences between the CENA median models that include scaling with hypocentral depth and the NGA-West2 models. The hanging-wall model from SWUS was presented as a proponent model for NGA-East. The day wrapped up with a feedback on hazard calculations using the median and standard deviation models.

2.5 Working Meetings

As indicated previously, a significant amount of technical development work was required and led by the WGs. As a result, the WGs supported the TI team by performing a number of evaluations and proposing a variety of technical choices. Besides the formal SSHAC workshops listed in the previous section, there have been numerous working meetings for various WGs. A selected list of such working meetings is provided at Appendix B of this report. A subset of these working meetings formally involved the PPRP, who was invited to participate and interact with the TI team and other meeting participants on targeted topics.

2.6 Project Documentation

A critical task in SSHAC Level 3 is the project documentation, which is vital to the successful completion of any project. The need for comprehensive documentation is especially important for studies conducted within the regulatory arena. The SSHAC guidelines document devotes a full chapter on the type and required level of documentation (Budnitz et al. 1997).

For the NGA-East project, the overall project documentation includes:

- The set of PEER reports documenting the science part of the NGA-East (as listed in Chapter 1 of this report)
- The project plan (provided in Appendix A of this report)
- SSAHC workshop agendas and presentations (Section 2.3 and Appendix B of this report)
- The set of comments provided by the PPRP and the project resolutions to the comments (Appendix A of this report)
- This final report.

2.7 Project Team

As a result of the project evolution, the NGA-East organization features all the components of a SSHAC Level 3 project, but it also features groups from the original project model (Figure 2–1). An important feature of NGA-East is the inclusion of WGs that supported the TI team and focused on specific technical areas. Some NGA-East WGs and technical tasks are not formally part of the SSHAC Level 3 process, but they were nonetheless important to the overall project. These are the Geotechnical and Vertical WGs shown in Figure 2–1. The role of the different groups and participants in Figure 2–1 are briefly summarized below. In the context of the SSHAC process, the WGs essentially played the role of Resource Experts and the sub-award researchers and contractors played the role of Specialty Contractors. Some individuals from these two groups also played a Proponent Expert role at specific times during the project. Refer to SSHAC (1997) and U.S. NRC (2012) for the key attributes and requirements associated to the SSHAC roles.

The Project Manager (PM), with the assistance of the TI team, informs the Joint Management Committee (JMC) and the PPRP of process and technical developments. The TI team Leads (e.g., co-chairs) are responsible for ensuring that all the technical participants have the required information to support the project. The TI team, with input from the whole project team, is

responsible for identifying and providing invitations to the Resource and Proponent Experts proposed for the workshops. Project-wide email distributions are to be coordinated by the PM and the TI leads and channeled through the PEER staff. Table 2–1 provides the list of key project participants.

2.7.1 Sponsors

The NGA-East project is jointly sponsored by the U.S. Nuclear Regulatory Commission (NRC), the U.S. Department of Energy (DOE), the Electric Power Research Institute (EPRI), and the U.S. Geological Survey (USGS).

2.7.2 Project Management

The project is managed by the PM and the JMC, which is composed of representatives of the key sponsoring organizations. These organizations are the NRC, EPRI, DOE, and USGS. The PM and JMC authorize the use of project resources on various tasks and are responsible for the overall direction of the project.

The SSHAC Guidelines Process Manager (SGPM) provides further guidance on the implementation of the SSHAC Level 3 assessment process. The SGPM is also responsible to maintain discussion and communication with the NRC staff to assure the ongoing regulatory acceptability of the NGA-East SSHAC Level 3 approach.

2.7.3 Technical Integrator Team (TI Team)

The TI team is ultimately responsible for all GMC technical products and technical assessments, and for defending their bases, as well as for the associated documentation. The TI team also oversees and coordinates the technical work performed by the WGs, sub-award researchers, and contractors, and supports the PM ensuring that the project scope and schedule are maintained. The TI team participated regularly in WG meetings to monitor the progress on technical tasks. The TI team, in collaboration with the PM, is responsible for the development of the project plan and for the organization of the workshops. The two TI co-chairs are also referred to as TI Leads in the current document.

2.7.4 Participatory Peer Review Panel

The PPRP provides the overall process and technical review as required by the SSHAC Level 3 process. The PPRP reports directly to the Project Manager, as shown in Figure 2–1. The role of the PPRP is by definition participatory and continual from project inception to project completion. The PPRP is responsible for assuring that the overall process is consistent with the objectives of the SSHAC guidelines. The PPRP is not responsible for the review of NGA-East tasks that are not defined in the SSHAC Level 3 process.

2.7.5 Working Groups

The NGA-East project includes seven WGs, each of which is focused on a specific technical area. The WGs were originally created to address the key technical issues identified in Bozorgnia (2008). The WGs are an essential part of NGA-East. They support the TI team by providing guidance on research needs and/or research products. Some research tasks are

performed directly by the WG members, while other tasks are performed by other researchers outside the WG. The WGs work closely with the TI team. Below is a short overview of the main tasks associated with each WG.

- **Database WG:** responsible for developing an exhaustive database of recorded motions in Central and Eastern North America (CENA), with the associated metadata.
- **Path/Source WG:** responsible for developing regionalized models for correlated sets of source (stress-drop) and path parameters (attenuation and quality factor, Q).
- **Simulations WG:** responsible for developing coordinating the validation and forward modeling of ground-motion simulation, considering different methods for finite-fault and stochastic simulations. Since simulations are important for achieving the project goals, the TI team assumed a very active role in the Simulations WG.
- **Geotechnical WG:** responsible for developing a simplified model to remove site effects at the recording stations, define the reference-rock shear-wave velocity and kappa values, and the range of conditions to which they apply. The Geotechnical WG is also tasked to develop a site effects model for NGA-East, but this task is not formally part of the SSHAC Level 3 process.
- **Ground-Motion Models (GMMs) WG:** responsible for developing a suite of "candidate" GMMs. These models are subsequently used by the TI team to develop the final suite of GMMs and their associated logic tree weights.
- **Sigma WG:** responsible for developing a suite of candidate standard deviation models for the project. This WG uses both recorded data and numerical simulations from CENA and western U.S. (WUS) to develop the models. The Sigma WG tasks are integrated through both the NGA-East and the NGA-West2 projects.
- **Vertical Motions WG:** responsible for developing models for V/H ratios to be applied to the horizontal GMMs. This task is not part of the SSHAC Level 3 process, and will be coordinated by PEER in conjunction with other NGA projects.

2.7.6 Specialty Contractors

This category represents researchers and practitioners who will contribute data, models, or methods outside of the WGs. Some of the contractors will be coordinated by WGs, while others will be directly coordinated by the TI team. In the context of Figure 2–1, this category also represents the different Resource Experts that provide data, models, or methods and the Proponent Experts as defined in the SSHAC assessment process documentation (SSHAC 1997).

2.8 References

Bozorgnia, Y. (2008), “A Roadmap for the Next Generation Attenuation Models for Central & Eastern North America (NGA-East).” Letter report to the U.S. NRC. Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Clu, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. 1997. “Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts.” Report NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington D.C.

EPRI/USDOE/USNRC. 2012. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

NRC. 2012. “Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies.” Technical Report NUREG-2117, U.S. Nuclear Regulatory Commission, Washington D.C.

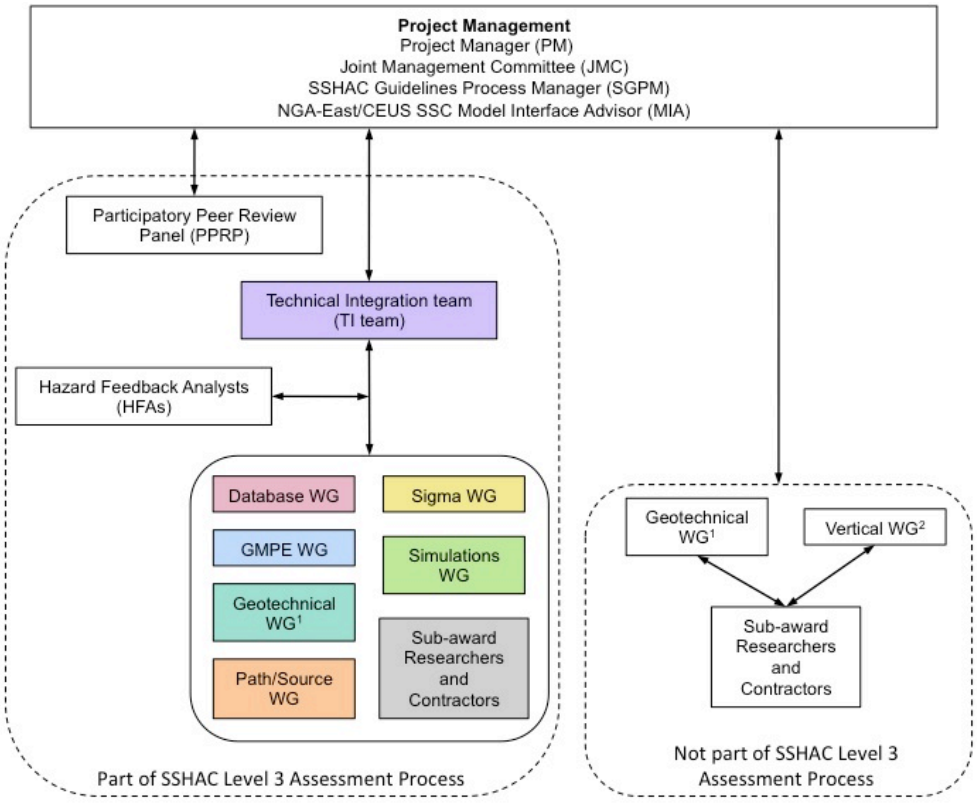
Senior Seismic Hazard Analysis Committee (SSHAC). 1997. “Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts.” U.S. Nuclear Regulatory Commission Report, NUREG/CR-6372, Washington, D.C.

Table 2–1 NGA-East project team and points of contact (group chairs are marked in boldface)

Committee or Group	Members		
	Name	Last Name	Organization
Project Manager (PM)	**Yousef**	**Bozorgnia**	**UC Berkeley, UCLA**
Joint Management Committee (JMC)	Clifford	Munson (*)	NRC
	Steve	McDuffie	DOE
	Lawrence	Salomone	EPRI
	Mark	Petersen	USGS, Denver
SSHAC Guidelines Process Manager	Clifford	Munson (*)	NRC
NGA-East/CEU SSC Model Interface Advisor (MIA)	Robert	Youngs	AMEC
Technical Integrator Team (TI)	**Norman**	**Abrahamson**	**UC Berkeley**
	Christine	**Goulet**	**UC Berkeley, USC**
	Linda	Al Atik	Linda Alatik Consulting
	Gail	Atkinson	Western University
	Robert	Graves	USGS, Pasadena
	Robert	Youngs	AMEC
Hazard Feedback Analysts (HFAs)	Robert	Youngs	AMEC
Participatory Peer Review Panel (PPRP)	**Gabriel**	**Toro (**)**	**Lettis Consultants Intl**
	John	Adams	Geological Survey of Canada
	Jon	Ake	NRC
	John	Ebel	Boston College
	Jeff	Kimball	National Nuc. Security Admin., Ret.
	Rich	Lee	Los Alamos National Lab
Database Working Group	**Chuck**	**Mueller**	**USGS, Denver**
	David	Boore	USGS, Menlo Park
	Kenneth	Campbell	EQECAT, Inc.
	Chris	Cramer	Univ. of Memphis
	Robert	Herrmann	Saint Louis Univ.
	Walt	Silva	Pacific Engineering, Inc.
Geotechnical Working Group	**Youssef**	**Hashash**	**Univ. of Illinois**
	Kenneth	Campbell	CoreLogic, Inc.
	Albert	Kottke	UC Berkeley
	Ellen	Rathje	Univ. of Texas - Austin
	Walt	Silva	Pacific Engineering, Inc.
	Jonathan	Stewart	UCLA
GMPE Working Group	See Chapter 7 of this report for the list of candidate ground motion models and their developers		
Path/Source Working Group	**Martin**	**Chapman**	**Virginia Tech.**
	Gail	Atkinson	Univ. of Western Ontario
	David	Boore	USGS, Menlo Park
	Kenneth	Campbell	CoreLogic, Inc.
	Robert	Herrmann	Saint Louis Univ.
	Shahram	Pezeshk	Univ. of Memphis
	Walt	Silva	Pacific Engineering, Inc.
Sigma Working Group	**Linda**	**Al Atik**	**Linda Alatik Consulting**
	Jack	Baker	Stanford Univ.
	Kenneth	Campbell	CoreLogic, Inc.
	Brian	Chiou	Caltrans
	Ellen	Rathje	Univ. of Texas - Austin
	Adrian	Rodriguez-Marek	Virginia Tech.
	Gabriel	Toro	Lettis Consultants Intl
	Melanie	Walling	USGS, Golden
Simulation Working Group	**Ralph**	**Archuleta**	**UC Santa Barbara**
	Norman	Abrahamson	UC Berkeley
	Martin	Chapman	Virginia Tech.
	Walter	Silva	Pacific Engineering, Inc.
	Paul	Somerville	AECOM
	Yuehua	Zeng	USGS, Denver

(*) Originally, Annie Kammerer
(**) Originally, Julian Bommer

Figure 2–1 NGA-East organization flowchart and lines of communications.



Note: NGA-East has two sets of objectives: one for the SSHAC Level 3 objectives and a set of additional objectives outside of the SSHAC Level 3 assessment process.
[1] The Geotechnical WG has tasks associated with both sets of objectives.
[2] All the tasks associated with the Vertical WG are associated with non-SSHAC objectives.

3. Use of CEUS SSC Information and NGA-East Hazard Implications

3.1 Introduction

It is expected that the primary use of the NGA-East Ground-Motion Characterization (GMC) model will be in conjunction with the Central and Eastern United States Seismic Source Characterization (CEUS SSC) model developed by EPRI/DOE/NRC (2012). Therefore, the NGA-East GMC development was focused on characterization of ground motions for the types of seismic sources represented in the CEUS SSC model. This chapter provides a brief overview of the CEUS SSC model and describes the basis for selection of the range of magnitudes and distances for which the GMC and the individual ground-motion models (GMMs) were developed.

3.2 Regional CEUS SSC Source Model

The CEUS SSC project was conducted from April 2008 to December 2011 to provide a regional seismic source model for use in probabilistic seismic hazard analyses (PSHA) for nuclear facilities (EPRI/DOE/NRC 2012). The CEUS SSC project was conducted using Senior Seismic Hazard Analysis Committee (SSHAC) Study Level 3 methodology (Budnitz et al. 1997; NRC 2012) to provide high levels of confidence that the data, models, and methods of the larger technical community had been considered, and the center, body, and range of technically defensible interpretations had been included. Hazard sensitivity calculations were conducted at seven test sites representative of different CEUS hazard environments. The study region and the location of these sites are shown in Figure 3–1.

3.2.1 Types of Seismic Sources

Two types of seismic sources are included in the CEUS-SSC model: distributed seismicity sources and Repeated Large Magnitude Earthquake (RLME) sources. Distributed seismicity sources represent future seismicity that is broadly distributed and not related to specific features known to rupture repeatedly in large earthquakes. The distributed seismicity source zones allow for the occurrence of earthquakes at all locations in the CEUS. The distributed seismicity source zones delineate large regions of the CEUS within which the characteristics of future earthquakes are expected to be similar. The rate of seismicity is allowed to vary spatially within these large regions to capture the observed patterns of earthquake activity.

In contrast to the distributed seismicity zones, the RLME sources are geographically constrained to areas identified as capable of generating repeated large (**M** ≥ 6.5) earthquakes as inferred from the historical or paleoseismic record (e.g., geologically-young fault displacement and paleoliquefaction features). The RLME sources represent the potential occurrence of additional, large-magnitude earthquakes at specific locations in the CEUS and the hazard from these sources is to be added to the seismic hazard computed from the distributed seismicity sources.

3.2.2 Modeling of Epistemic Uncertainty in CEUS SSC Model

The framework for the CEUS SSC model is established using a logic tree. The logic tree indicates the alternative approaches and conceptual models that are used, and establishes the relative weights assigned to the main alternatives [see Figure 4.2.2–1 of EPRI/DOE/NRC (2012)]. The first-order levels of the tree address the basic conceptual models related to the alternative approaches; these are followed by levels that represent the uncertainties in the implementation of each of the alternative approaches. At each level (node) of the logic tree, the set of branches defines the discrete alternative models or model parameters considered potentially applicable. These alternatives are assigned weights that represent the relative credibility assigned to each alternative in the model.

3.2.3 Epistemic Uncertainty in CEUS SSC Model Seismic Source Geometry

The first assessment on the master logic tree [see Figure 4.2.2–1 of EPRI/DOE/NRC (2012)] is the choice between two conceptual models used to define the distributed seismicity sources: $\mathbf{M}_{max}$ zones and seismotectonic zones. Figure 3–2 shows the preferred division of the CEUS into the MESE-N and NMESE-N $\mathbf{M}_{max}$ zones, with "-N" indicating the "narrow" interpretation of the width of the MESE zone along the Atlantic margin. An alternative "wide" interpretation of the width of the MESE zone is included in the CEUS SSC model along with a third interpretation that the entire CEUS consists of a single $\mathbf{M}_{max}$ zone, delineated by the study region shown on Figure 3–1.

The seismotectonic zones approach involves the use of additional seismotectonic data beyond differences in $\mathbf{M}_{max}$ to define spatial variations in the characteristics of future earthquakes. Figure 3–3 shows the preferred model for the seismotectonic source zones. Other alternatives represent differences in the width of the zones along the Atlantic margin and in zone geometries in the vicinity of New Madrid.

The locations of the RLME sources are shown on Figure 3–4. These sources confine the location of repeated large-magnitude earthquakes to specific geographical locations. A number of the RLME sources have alternative geometries reflecting epistemic uncertainty in the source boundaries.

3.2.4 Epistemic Uncertainty in CEUS SSC Model Seismicity Parameters

The primary seismicity parameters are the recurrence rates of earthquakes associated with each seismic source and the maximum magnitude that each source can produce. The earthquake recurrence rates for the distributed seismicity sources are provided for each source for either ¼° × ¼° or ½° × ½° cells. Epistemic uncertainty in the activity rates and b-values is modeled by three alternative sets of weighted magnitude intervals and eight alternative sets of parameters for each source zone and for each set of weighted magnitude intervals. Two sets of recurrence parameters were provided. One set specifies the seismicity rates in terms of the frequency of earthquakes of **M** = 5 and larger. These parameters are for use when hazard is computed using a fixed minimum magnitude of **M** = 5.0. The second set of parameters specifies the seismicity rates in terms of the frequency of earthquakes of **M** = 4 and larger.

Uncertainty in maximum magnitude was assessed using a combination of statistical techniques that resulted in broad distributions ranging from a minimum value of the larger of **M** = 5.5 or the largest observed earthquake to a maximum in the range of **M** = 7.9 to 8.1. The specified upper limit on CEUS SSC earthquakes is **M** = 8.25.

The earthquake recurrence for the RLME sources was modeled using a characteristic magnitude only, with the uncertainty in recurrence rate and magnitude assessed from the available data for timing of historical and paleo earthquakes and the spatial extent of felt effects.

3.2.5 Earthquake Rupture Characterization in the CEUS SSC Model

The CEUS SSC model provides a description of the rupture characteristics for each seismic source in Tables 5.4-1 and 5.4-2 of EPRI/DOE/NRC (2012). These are defined in terms of aleatory distributions for faulting type (strike–slip versus reverse), and aleatory distributions of fault strike and fault dip for each style of faulting. The thickness of the seismogenic crust in each source is treated as an epistemic uncertainty modeled by two or three values with assigned relative weights. The EPRI/DOE/NRC (2012) suggests the use of the model of Leonard (2010) to define rupture area as a function of magnitude and use of an aspect ratio of 1:1 until the seismogenic crustal thickness is reached. In Section 13.3.6.3 of this report, the relationships for rupture area and aspect ratio are updated using more recent information for the purpose of evaluation of hazard. The EPRI/DOE/NRC (2012) report provides information on the depth distribution of earthquakes, but no specific model is given. Section 13.3.6.2 presents the development of depth distribution models for the CEUS SSC sources for use in hazard calculations.

3.3 Demonstration Sites

For purposes of demonstrating the CEUS SSC model, seven sites were selected for hazard calculations by EPRI/DOE/NRC (2012); their locations are shown on Figure 3–1. The site locations were selected to span a range of seismic source types and levels of seismicity. These sites were subsequently used by EPRI (2013) to demonstrate the effects of the EPRI (2013) GMC model. For the NGA-East study, these sites were again used to perform hazard sensitivity analyses and to demonstrate the hazard produced by the NGA-East GMMs (see Chapters 12 and 14).

3.4 Earthquake Contributions to Hazard in the CEUS

Figures 3–5 and 3–6 show example magnitude-distance deaggregation of hazard for and annual frequency of exceedance (AFE) of 10^{-4} at two of the CEUS SSC demonstration sites located at large distances from RLME sources. These hazard calculations were performed using the CEUS SSC source model and preliminary versions of the NGA-East ground-motion models. On each figure, the top plot shows the results for 10 Hz spectral acceleration and the bottom plot shows the results for 1 Hz spectral acceleration. The Manchester site is located in a region of moderate seismicity approximately 500 km from the Charlevoix RLME source. At this site, earthquakes from a wide range of distances and magnitudes contribute to the hazard, with significant contributions to the 1 Hz hazard from distances beyond 500 km. The Topeka site is located in an area of relatively low seismicity approximately 600 km from the New Madrid faults

RLME. At this site, small to moderate earthquakes in the distance range of 0 to 200 km produce most of the 10 Hz hazard while the 1 Hz hazard is dominated by the contributions from the New Madrid faults RLME source.

Figures 3–5 and 3–6 show that there can be a significant contribution to the hazard from large earthquakes at large distances. As a perhaps extreme example, Figure 3–7 shows 1 Hz hazard curves computed for a site in northern Minnesota in an area of very low local seismicity, located approximately 1400 km from the New Madrid RLME. The calculations for this site were performed using the EPRI (2013) GMM. These results indicate that in areas of very low local seismicity, there can be large contributions to hazard from very distant, large earthquakes.

3.5 Implications of Existing Seismic Source Models to NGA-East Project

The CEUS SSC model that used currently for nuclear applications requires specification of ground motions from earthquakes as large as **M** = 8.1 for the distributed seismicity sources and **M** = 8.15 for the RLME sources. This upper magnitude limit is similar to that developed by Petersen et al. (2014) for the CEUS. The current standard of practice for nuclear facilities is to compute seismic hazard using a minimum magnitude of **M** 5, although one is permitted to use the CAV filter for lower bound magnitudes in hazard [e.g., EPRI (2006)], requiring hazard integration from magnitudes as low as **M** 4. Based on these observations, the magnitude range used in the development of the NGA-East GMM is **M** = 4 to 8.2. The example calculations presented in Section 3.4 indicate contributions to hazard at two demonstration sites from earthquakes at distances beyond 500 km. As indicated by the results shown for the site in northern Minnesota, in areas of very low local seismic activity, hazard computed using existing ground-motion models may indicate significant contributions from earthquakes at a distance of nearly 1500 km. Although it is expected that at most sites the majority of contributions to hazard will be from earthquakes within 1000 km, it is important that the resulting ground-motion models developed in the NGA-East project extrapolate to larger distances in a reasonable manner. These considerations were used to establish the magnitude and distance range for model development described in Section 6.

The CEUS SSC model provided explicit characterization of the orientation of future earthquake ruptures as well as guidance on the dimensions of ruptures. It is anticipated that other source characterization models that might be developed in the near future would provide similar characterizations. These characterizations influences the selection of the distance metric for parameterizing ground-motion variation with distance discussed in Section 6.

Existing US NRC Guidance generally require the development of design or evaluation response spectra over a broad frequency range, typically from about 0.1 to 100 Hz. However, the current EPRI (2013) GMM that is used for seismic hazard evaluations for nuclear plants provides ground-motion assessments for only seven ground-motion frequencies, thus requiring users to develop interpolation and extrapolation schemes in order to produce design and evaluation response spectra for the full-frequency range of interest. In order to alleviate the need to develop elaborate interpolation methods, The NGA-East project characterized ground motions for a sufficient number of frequencies over the frequency range of 0.1–100 Hz [plus peak ground acceleration (PGA) and peak ground velocity (PGV)] at a sufficient number of frequencies such that smooth design and evaluation response spectra for reference rock

conditions can be developed from the hazard results using simple interpolation, as discussed in Section 6.

The SSHAC process used in the NGA-East Project requires that the TI team receives feedback during the evaluation and integration process that includes hazard sensitivity calculations. Because the CEUS SSC model is the current standard for use in PSHA for nuclear facilities, it represents the natural choice for developing appropriate seismic source scenarios for sensitivity calculations. These calculations are presented in Section 12.

3.6 References

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Clu, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. 1997. "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts." Report NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington D.C.

EPRI. 2006. "Program on technology innovation: Use of Cumulative Absolute Velocity (CAV) in Determining Effects of Small Magnitude Earthquakes on Seismic Hazard Analyses." Electric Power Research Institute, Palo Alto, CA, and the U.S. Department of Energy, Germantown, MD: 2006 1014099.

EPRI. 2013. "Ground-Motion Model (GMM) Review Project." Electrical Power Research Institute, EPRI Report 3002000717, Palo Alto, CA.

EPRI/USDOE/USNRC. 2012. "Central and Eastern United States Seismic Source Characterization For Nuclear Facilities." U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Leonard, M. 2010. "Earthquake Fault Scaling Self-Consistent Relations of Rupture Length, Width, Average Displacement and Moment Release." Bull. Seismol. Soc. Am., 100: 1971−1988.

NRC. 2012. "Practical implementation guidelines for SSHAC Level 3 and 4 Hazard Studies." NUREG-2117, Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, 138 pgs., Washington, D.C.

Petersen, M.D., M.P. Moschetti, P.M. Powers, C.S. Mueller, K.M. Haller, A.D. Frankel, Y. Zeng, S. Rezaeian, S.C. Harmsen, O.S. Boyd, N. Field, R. Chen, K.S. Rukstales, N. Luco, R.L. Wheeler, R.A. Williams, and A.H. Olsen. 2014. "Documentation for the 2014 Update of the United States National Seismic Hazard Maps." U.S. Geological Survey Open-File Report 2014–1091, 243 pgs., http://dx.doi.org/10.333/ofr20141091

Figure 3–1 Map showing the CEUS study region and the seven demonstration sites (red stars).

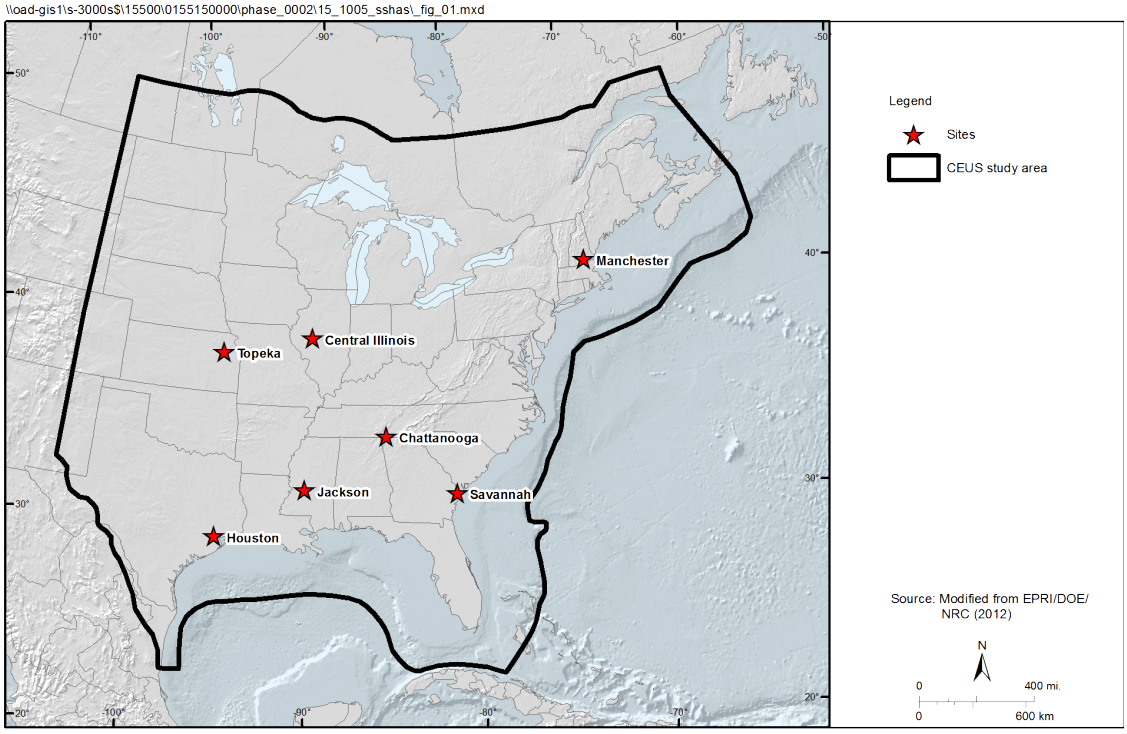


Figure 3–2 Map of preferred geometries for the M_{max} source zones.

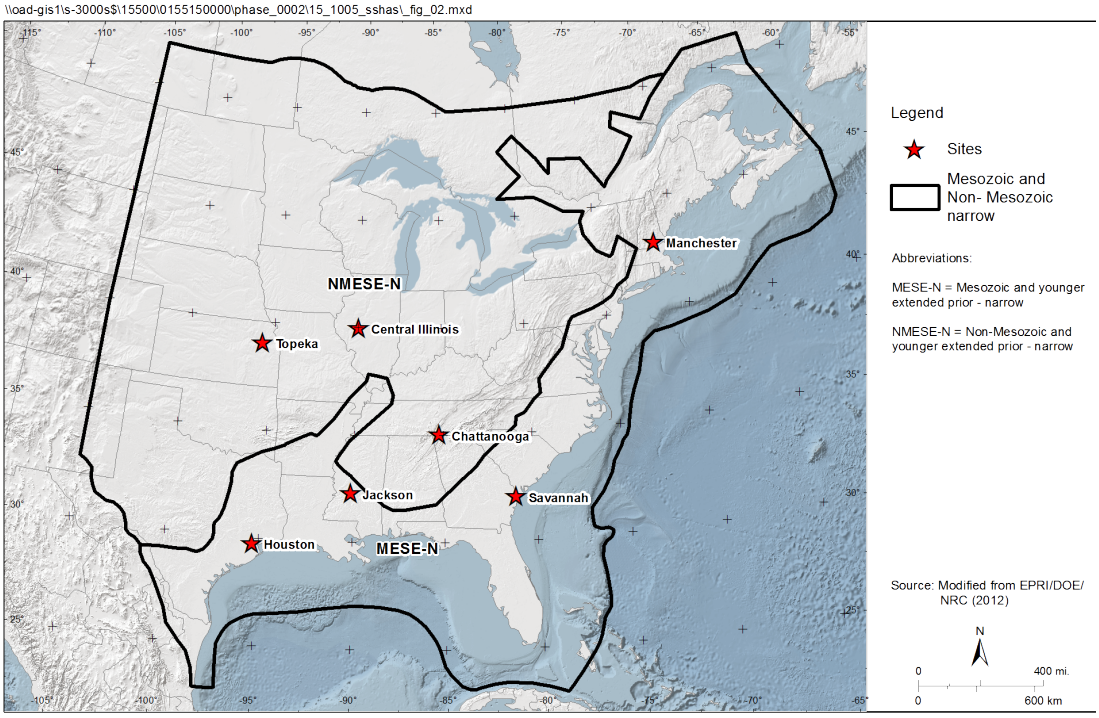


Figure 3–3 Map of preferred geometries for the seismotectonic source zones.

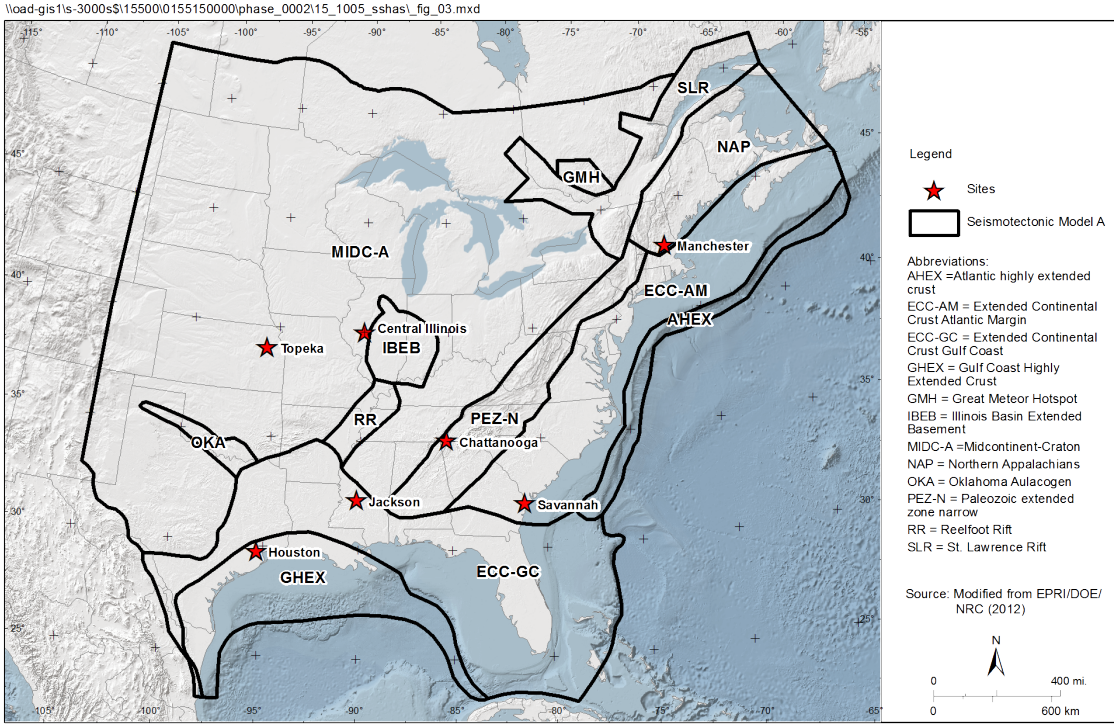


Figure 3–4 Locations of RLME sources.

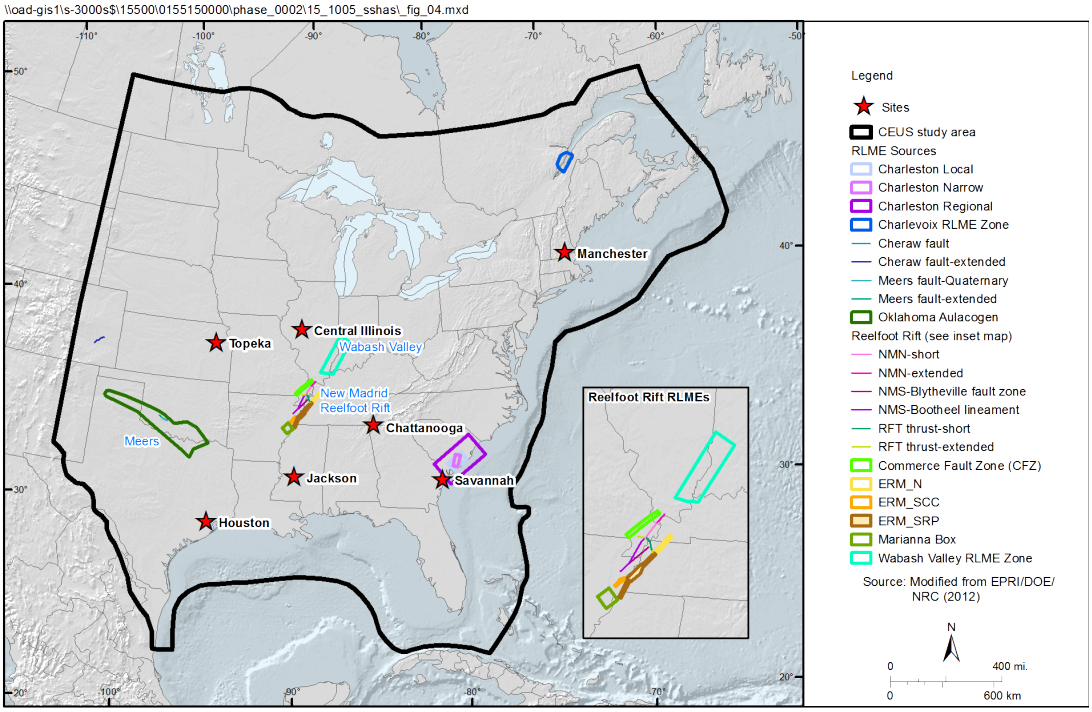


Figure 3–5 Example magnitude-distance deaggregation seismic hazard at the Manchester CEUS SSC demonstation site.

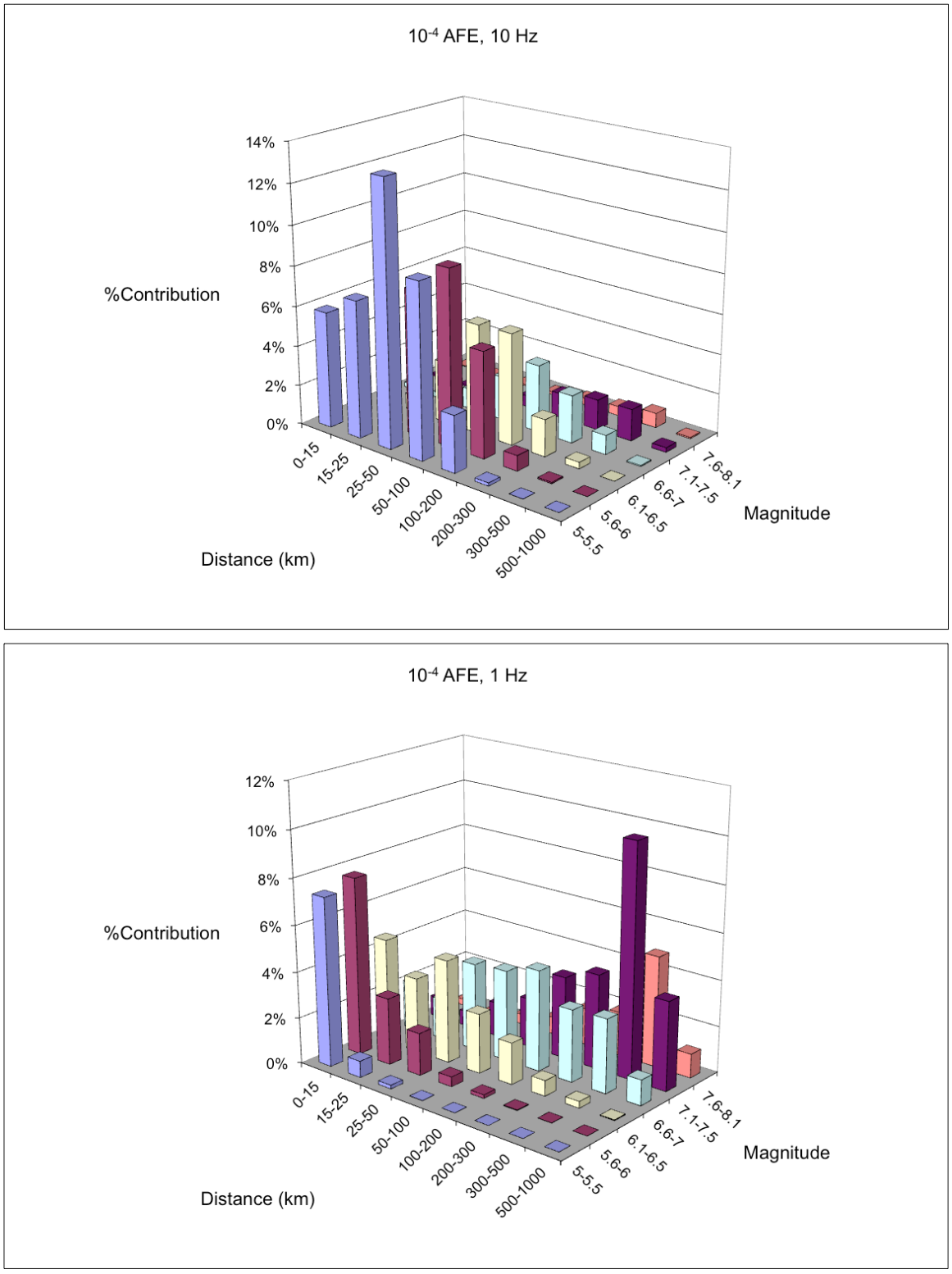


Figure 3–6 Example magnitude-distance deaggregation seismic hazard at the TopekaCEUS SSC deomonstration site.

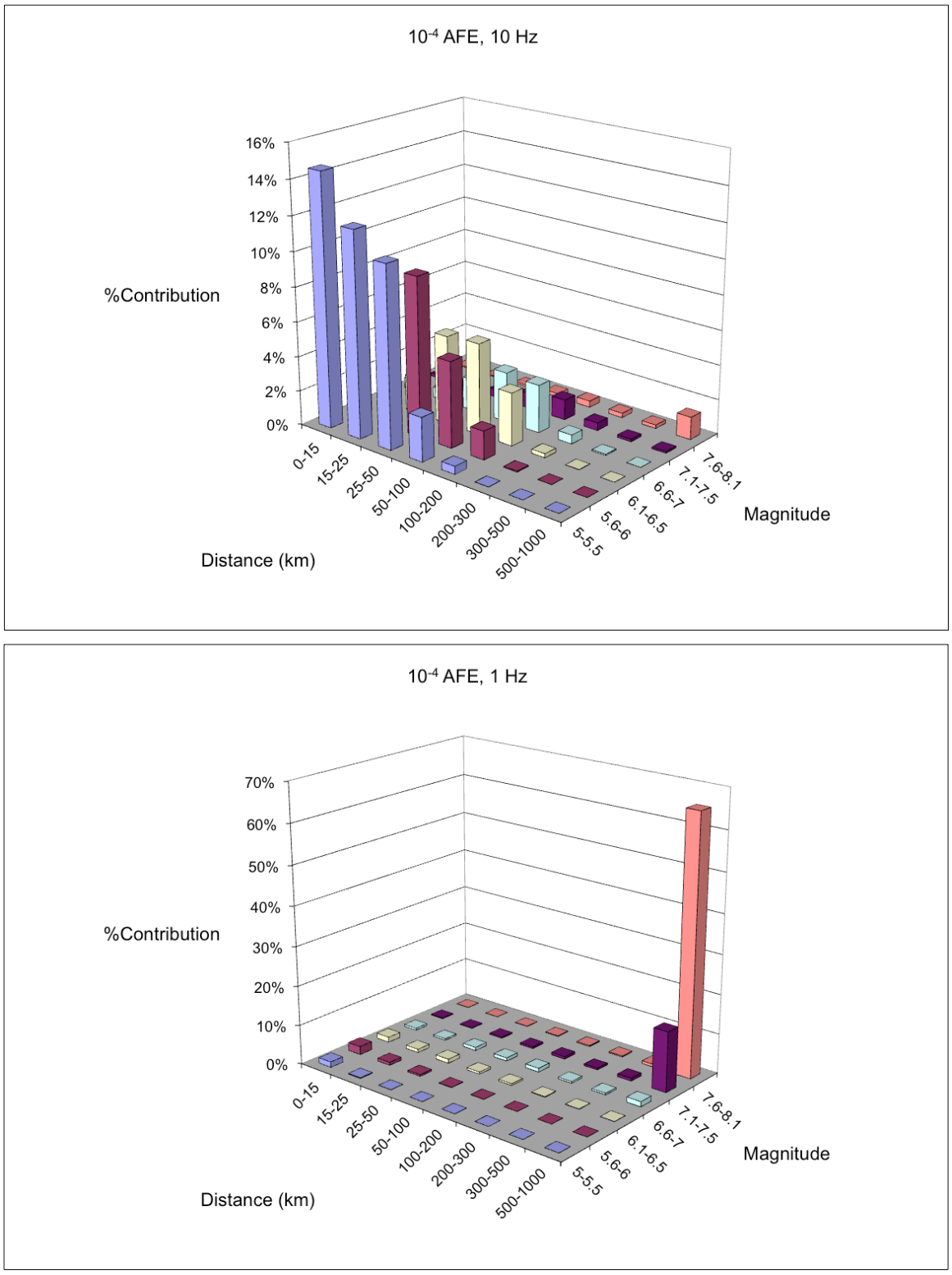
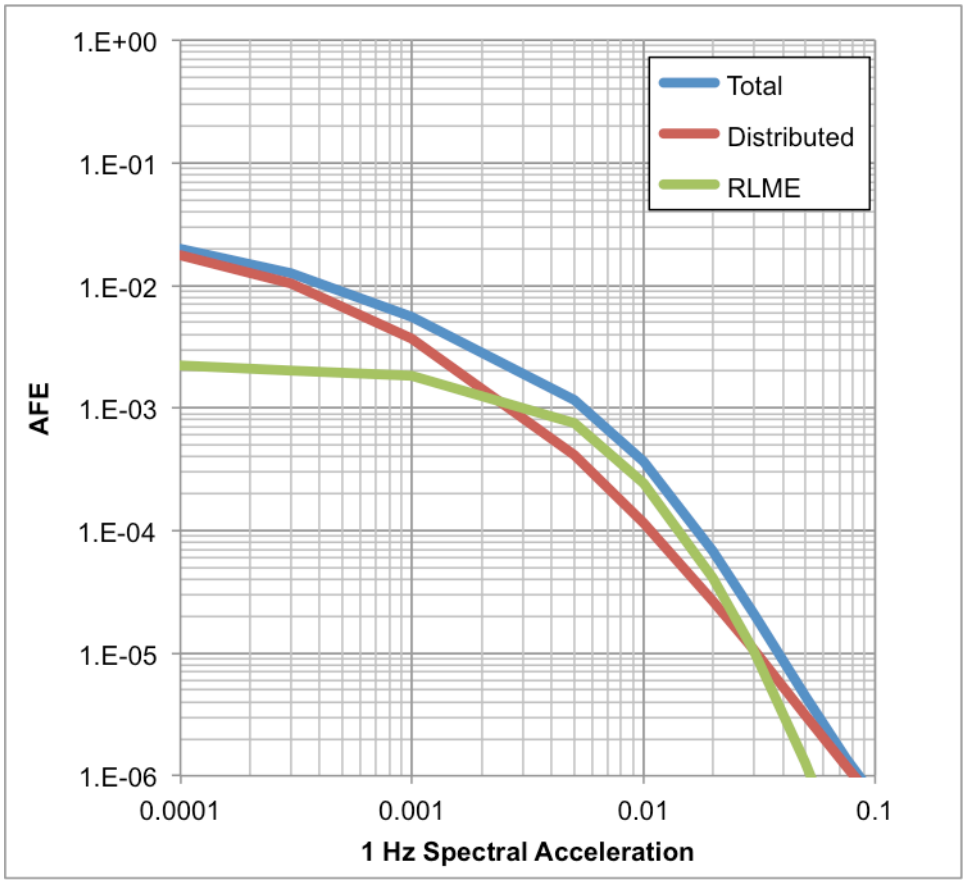


Figure 3–7 One-Hz spectral acceleration hazard curves for a site in northern Minnesota.



4. Regionalization

4.1 Introduction

The ground-motion characterization (GMC) documented in this report was developed for application to the CEUS region covered by the EPRI/DOE/NRC (2012) CEUS SSC model. This region extends from the Rocky Mountains to the Atlantic Coast and from southern Texas into southeastern Canada. The GMC developed in Chapters 9 and 11 uses as a basis candidate ground-motion models (GMMs) developed for general application in the CENA region that encompasses eastern Canada along with the CEUS. The issue addressed in this chapter is the extent to which this model can be applied throughout the CEUS region and the necessary adjustments needed for application in specific sub-regions. In the next two sections, we provide a summary of regionalization projects and the TI team recommendation for regionalization.

4.2 Summary of Studies of GMC Regionalization in the CEUS

4.2.1 EPRI (1993)

Prior to the NGA-East Project, the most comprehensive evaluation of GMC regionalization in the CEUS was performed as part of the EPRI (1993) study, which investigated the effects of differences in crustal characteristics across the CEUS on the rate of attenuation of earthquake ground motions. The CEUS was divided into 16 regions on the basis of crustal velocity structure and 11 regions on the basis of differences in the intrinsic attenuation rate, as parameterized by the quality factor Q. Figure 4–1 shows the resulting regionalizations. EPRI (1993) performed numerical simulations to investigate the degree to which differences in crustal structure, Q, and earthquake focal depth in the various sub-regions produce significant differences in the attenuation of earthquake ground motions; see Figure 4–1. Based on analyses of these simulations, EPRI (1993) concluded that the CEUS should be separated into two primary sub-regions: a Mid-Continent Region (MCR) and a Gulf Coast Region (GCR), as shown on Figure 4–2. The EPRI (1993) GCR corresponded to their crustal region 4, shown on the top panel of Figure 4–1, and the MCR encompasses the rest of the CEUS. EPRI (1993) then developed GMMs for these two regions. EPRI (1993) noted that a third sub-region could be defined (corresponding to crustal regions 7 and 13 on Figure 4–1) in which ground-motion amplitudes may be significantly lower than those for the MCR in the limited distance range of 80 to 150 km, and suggested that this could be accounted for in site-specific studies. For application in cases where the source-to-site travel path crosses the sub-region boundary, EPRI (1993) suggested comparing the results obtained using the models for the two sub-regions and using the higher of the two.

4.2.2 EPRI (2004)

EPRI (2004) developed an updated GMC for the CEUS in one of the first applications of the SSHAC Level 3 process defined in Budnitz et al. (1997). The primary focus of the EPRI (2004) study was to incorporate the significant effort on GMM development that occurred since the completion of the EPRI (1993) study into an updated GMC. The EPRI (2004) study adopted the EPRI (1993) regionalization model shown on Figure 4–2 primarily on the basis of its acceptance

by the participants in the project workshops. EPRI (2004) developed GMMs for the two sub-regions, with the Gulf Coast model created by applying scaling factors to the mid-continent model to account for differences in crustal structure and Q. EPRI (2004) indicated that for cases where the source-to-site path crosses the sub-region boundary, one could either follow the suggestion of EPRI (1993) of using the higher of the two predictions or could select the model that was assessed to best represent the travel path.

4.2.3 EPRI (2013)

The EPRI (2013) performed a SSHAC Level 2 study to update the 2004 EPRI GMC to incorporate the results of an additional decade of research on ground-motion modeling and the results generated by the NGA East project. The EPRI (2013) study adopted the concept developed by EPRI (1993) of two primary sub-regions, the MCR and GCR, but performed additional evaluations to refine the sub-region boundaries. EPRI (2013) found that the preliminary results of crustal regionalization of the CEUS being performed for the NGA East Project (Mooney et al. 2012) defined a Gulf Coast crustal region that corresponded closely with the boundaries of three seismotectonic source zones for distributed seismicity developed by EPRI/DOE/NRC (2012). This result is not surprising as the EPRI/DOE/NRC (2012) seismotectonic source zones were characterized in part by differences in crustal structure. Accordingly, EPRI (2013) proposed the modified Gulf Coast GMM sub-region shown on Figure 4–3 for use in seismic hazard analyses in conjunction with the EPRI/DOE/NRC (2012) SSC model. One notable difference between the EPRI (2013) and EPRI (1993) GCR is that the EPRI (2013) region encompasses all of the Mississippi Embayment, extending to the New Madrid region. The EPRI (2013) GCR also encompasses all of Florida. Previously, the EPRI (1993) characterization was ambiguous about the regionalization in southern Florida.

EPRI (2013) performed analyses of the available ground-motion data recorded by the Earthscope Transportable Array (TA) in the Mississippi Embayment and the central GCR, confirming differences in ground-motion attenuation characteristics in this region compared to the surrounding regions of the central U.S. EPRI (2013) used the results of these analyses to develop adjustments to the updated mid-continent GMM for application in the GCR. For those cases where the source-to-site path crosses the sub-region boundary, EPRI (2013) suggested that one could compute a weighted logarithmic average of the ground-motion predictions from the two sub-region GMMs based on the relative path lengths in the two sub-regions.

4.2.4 Dreiling et al. (2014)

As part of the science component of the NGA-East Project, Dreiling et al. (2014) performed an extensive evaluation of GMM regionalization following the general approach used by EPRI (1993). The CEUS was divided into four primary sub-regions on the basis of differences in crustal structure and tectonic history. These regions are shown on Figure 4–4. Dreiling et al. (2014) developed assessments of crustal velocity structure and Q for each sub-region, including variability in these parameters. The distributions for crustal properties within each region as well as distributions of earthquake focal depths were used in ground-motion simulations to develop distributions of earthquake ground motions for a range of magnitudes and distances. Similar Q models were used for all four regions, in contrast to many previous studies that utilized a lower

Q model for the GCR compared to other CENA regions. The ground-motion distributions for the four sub-regions defined on Figure 4–4 were compared to determine if significant differences exist. The results of the analysis confirm past conclusions. The attenuation characteristics in the Atlantic Coastal Plain and Appalachian Province were found to be similar to those for central North America, while those for the Mississippi Embayment/GCR were significantly different. Dreiling et al. (2014) concluded that differences in the seismic velocity structure of the crust, rather than the Q-factor, had the largest effect on the differences in attenuation of ground motions in the distance range they evaluated (35 to 500 km). This conclusion is based on the fact that they found significant differences in attenuation between the GCR and MCR despite using similar Q models for the two regions. One important note is that the crustal profile characterization for the GCR developed by Dreiling et al. (2014) did not include crustal velocity profiles from sites east of Mississippi.

4.2.5 Gallegos et al. (2014)

Gallegos et al. (2014) performed a Q tomography study for the CEUS using the Earthscope TA recordings filtered to the narrow frequency band of 0.5 to 1.5 Hz. They produced the Q_0 maps shown in Figures 4–5 and 4–6 using the two-station method (TSM) and reverse two-station method (RTSM), respectively. Their Q_0 maps show low values along the Gulf Coast (GCP) and scattered areas of lower Q in other parts of the CEUS, including the Reelfoot Rift Region (RFR). Gallegos et al. (2014) state that the TSM and RTSM results are generally similar, but there is lower resolution in the RTSM results due to less extensive ray–path coverage. They do indicate that both methods show lower Q in the GCP and RFR regions.

4.2.6 Cramer and Al Noman (2016)

Cramer and Al Noman (2016) performed an analysis of TA recordings to identify regional differences in Q within the CEUS. The dataset used included recordings obtained over much of the southeastern U.S. They fit a model of the form:

$$n(PSA) = A + CR - 0.5\ln(R) \qquad (4–1)$$

to 1 Hz and 5 Hz spectral accelerations recorded by the TA stations at distances beyond 150–200 km from the causative earthquakes, with the coefficients A and C determined by regression. The apparent Q as a function of frequency f is then computed by the relationship

$$Q(f) = -\pi f / C\beta \qquad (4–2)$$

where β is the crustal shear-wave velocity, taken to be 3.5 km/sec. Cramer and Al Noman (2016) use the dense grid of TA recordings to identify locations of significant changes in the slope parameter *C* of Equation (4–1) that indicate the transition from the higher Q of the GCR to the lower Q of the GCR. Figure 4–7 shows their proposed Q regionalization boundaries for earthquakes occurring outside of the GCR, and Figure 4–8 shows their proposed Q boundaries for earthquakes occurring within or near the GCR. They defined separate boundaries based on 1 Hz and 5 Hz motions. The Cramer and Al Noman (2016) Gulf Coast Q Region is smaller than the GCR proposed by Dreiling et al. (2014), extending only part way up the Mississippi Embayment and including only the western portion of Florida.

4.3 Evaluation of Regionalization

The studies summarized above consistently defined a division of the CEUS into two sub-regions with significant differences in attenuation characteristics: a GCR and a MCR. The recent advances in mapping crustal structure and the dense recordings from the Earthscope TA project have produced refinements to the boundary between these two regions from those originally proposed by EPRI (1993), which was based on more limited data. The regionalization proposed by Dreiling et al. (2014) is based primarily on crustal structure, which they conclude is the most important factor in producing differences in attenuation. The Cramer and Al Noman (2016) regionalization is based solely on differences in *Q*, as parameterized by a linear distance attenuation factor derived from recorded ground-motion data from a more extensive coverage of the southeastern U.S. than was available at the time of the Dreiling et al. (2014) study. The Cramer and Al Noman (2016) GCR represents a subset of the Dreiling et al. (2014) GCR in which the Reelfoot Rift portion of the upper Mississippi Embayment and the Florida Peninsula are considered to be part of the MCR. Although Cramer and Al Noman (2016) state that their results are similar to those of Gallegos et al. (2014), one could argue that the Gallegos et al. (2014) Q_0 maps shown on Figures 4–5 and 4–6 indicate that the low Q region on the Gulf Coast extends up the Mississippi Embayment to the New Madrid Region, which is consistent with the Dreiling et al.'s (2014) definition of the GCR.

The difference in the proposed location of GCR/MCR boundary between the Dreiling et al. (2014) and Cramer and Al Noman (2016) studies is considered to be representative of the current state of epistemic uncertainty in GMM regionalization for the CEUS. This uncertainty is taken into account in applying the GCR GMM adjustments presented in Chapter 13. Specific application recommendations by the TI team are provided in Chapter 13.

4.4 References

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Clu, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. 1997. "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts." Report NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington D.C.

Cramer, C.H., and N. Al Noman. 2016. "Improving Regional Ground Motion Attenuation Boundaries and Models in the CEUS and Developing a Gulf Coast Empirical GMPE using Earthscope US Array Data for Use in the National Seismic Hazards Mapping Project." US Geological Survey External Grants Program Final Technical Report, Award Number G14AP00049, August, 52 pgs.

Dreiling, J., M.P. Isken, W.D. Mooney, M.C. Chapman, and R.W. Godbee. 2014. "NGA-East Regionalization Report: Comparison of Four Crustal Regions within Central and Eastern North America using Waveform Modeling and 5%-Damped Pseudo-Spectral Acceleration Response." PEER Report No. 2014/15, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

EPRI. 1993. “Guidelines for Determining Design Basis Ground Motions. Volume 1: Method and Guidelines for Estimating Earthquake Ground Motion in Eastern North America.” Electric Power Research Institute, Technical Report, Palo Alto, CA.

EPRI. 2004. “CEUS Ground Motion Project Final Report.” Technical Report 1009684. Electric Power Research Institute, Palo Alto, CA.

EPRI. 2013. “EPRI 2004-2006 Ground-Motion Model (GMM) Review Project.” Final Report Project 3002000717. June 2013, Electric Power Research Institute, Palo Alto, CA.

EPRI/USDOE/USNRC. 2012. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Gallegos, A., N. Ranasinghe, J. Ni, and E. Sandvol. 2014. “Lg Attenuation in the Central and Eastern United States as Revealed by the Earthscope Transportable Array.” Earth Planetary Sci. Lett., 402: 187–196.

Mooney, W.D., G. Chulick, A. Ferguson, A. Radakovich, K. Kitaura, and S. Detweiler. 2012. “NGA-East: Crustal Regionalization.” Presentation at NGA-East Workshop on Path and Source Issues, University of California, Berkeley, October 16,; available at http://peer.berkeley.edu/ngaeast/2012/08/working-meeting-path-and-source-issues-oct-16-2012/.

Figure 4–1 Regionalization of the CEUS studied by EPRI (1993). Top: regionalization of crustal seismic velocity structure into 16 regions. Bottom: Intrinsic seismic damping *Q*-value regionalization into eight regions.

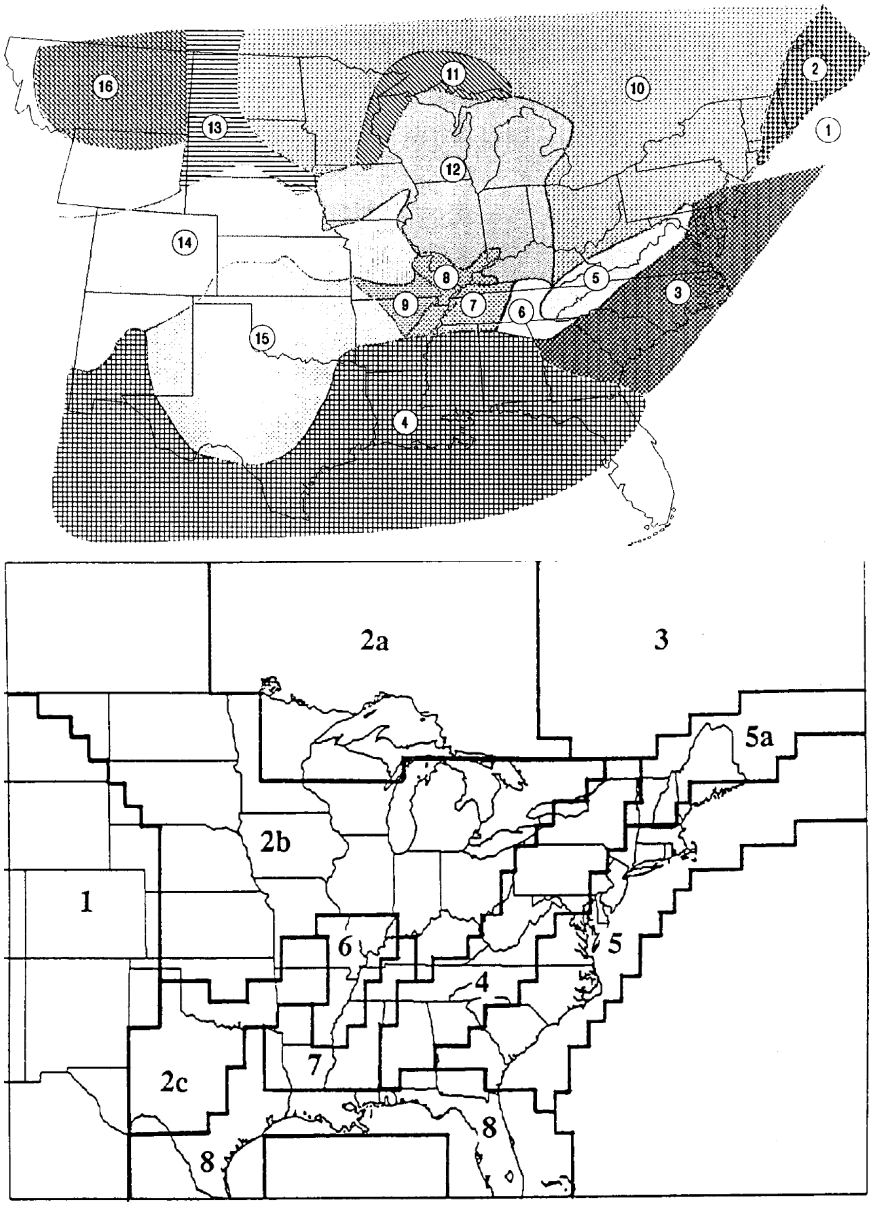


Figure 4–2 Mid-continent and Gulf Coast GMC regionalization defined by EPRI (1993), from EPRI (2004).

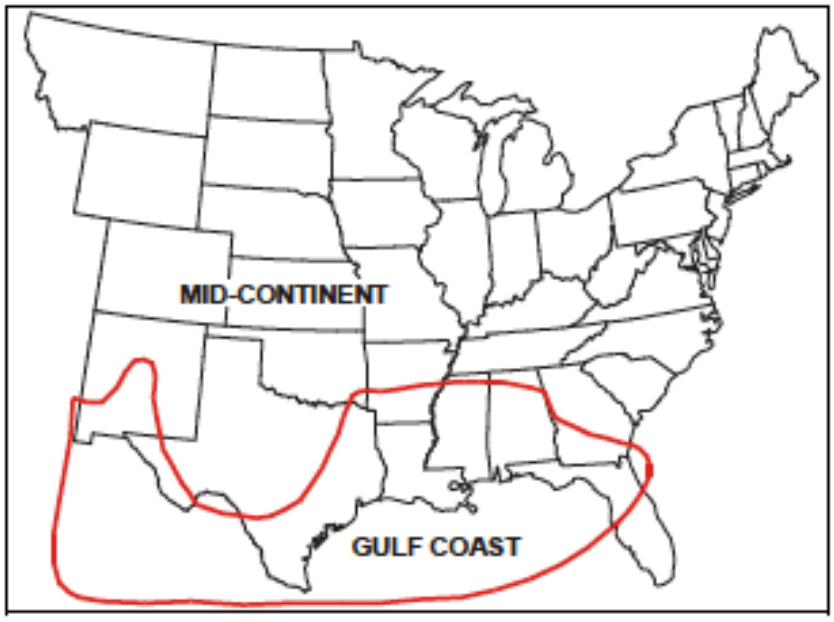


Figure 4–3 Gulf Coast GMC regionalization defined by EPRI (2013).

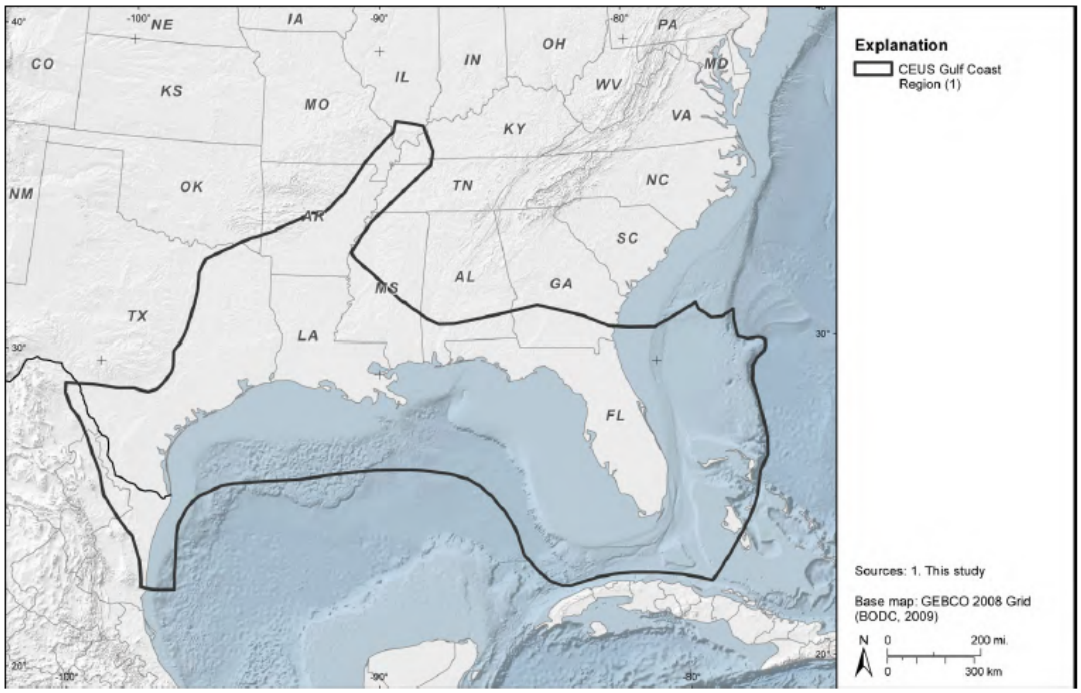


Figure 4–4 Four regions defined for CENA by Dreiling et al. (2014). The regions have been numbered as follows for the NGA-East database: (1) Mississippi Embayment/Gulf Coast Region; (2) Central North America; (3) the Appalachian Province; and (4) the Atlantic Coastal Plain. Together, Regions 2, 3, and 4 form the larger Mid-Continent Region.

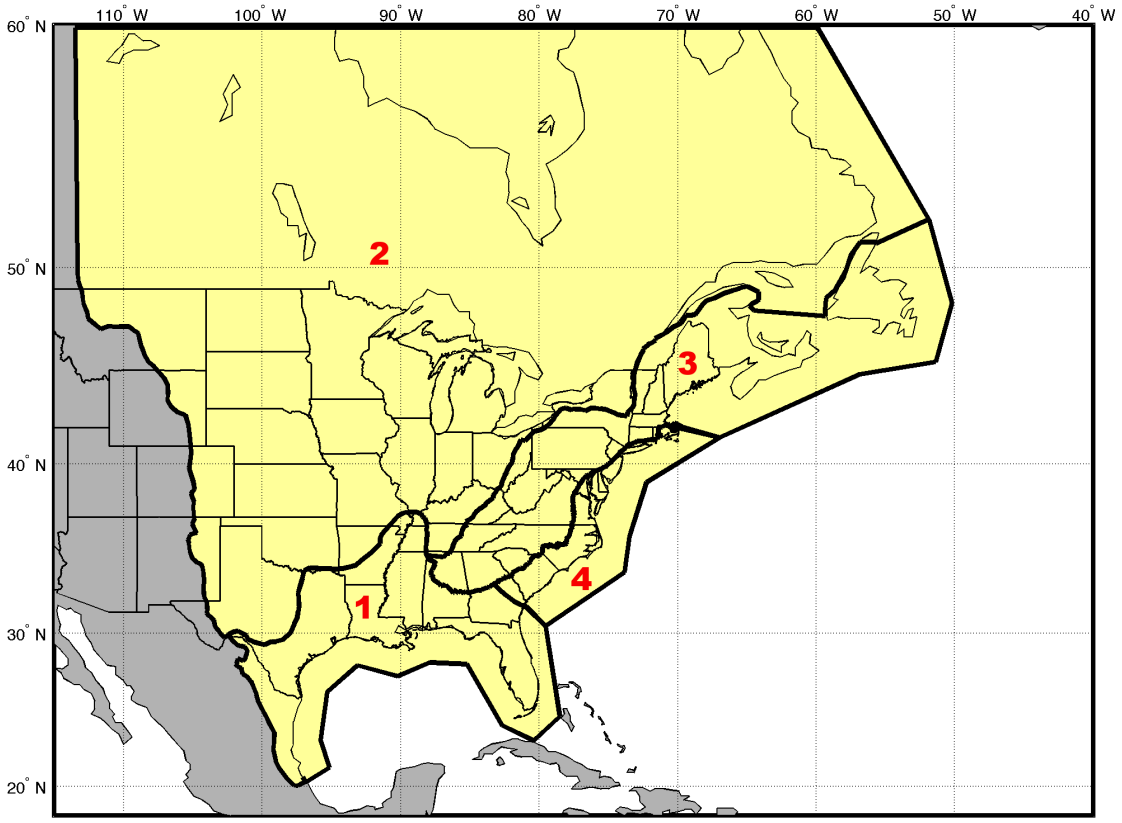


Figure 4–5 Two-station Lg attenuation map of the CEUS (Gallegos et al. 2014).

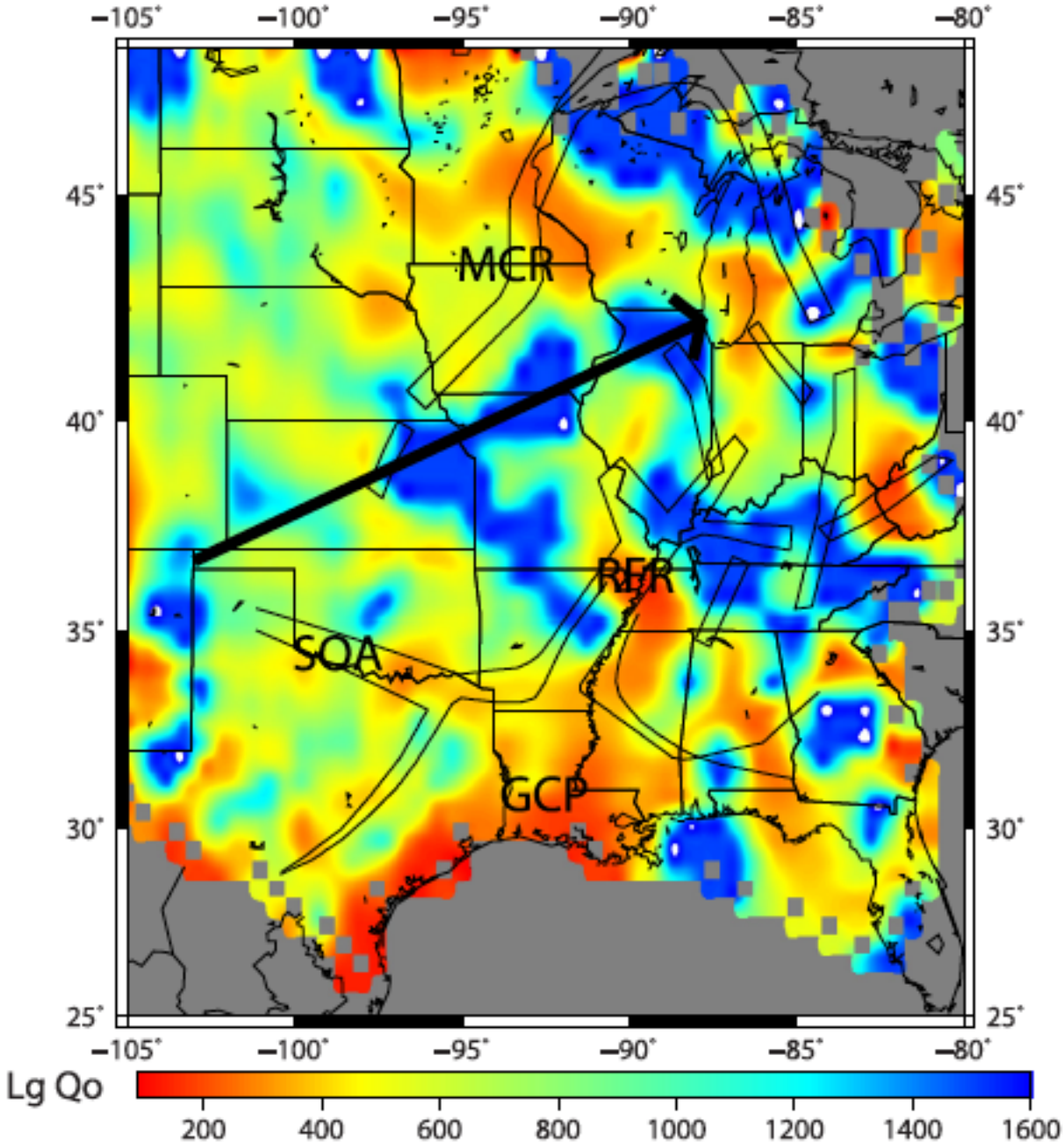


Figure 4–6 Reverse two-station Lg attenuation map of the CEUS (Gallegos et al. 2014).

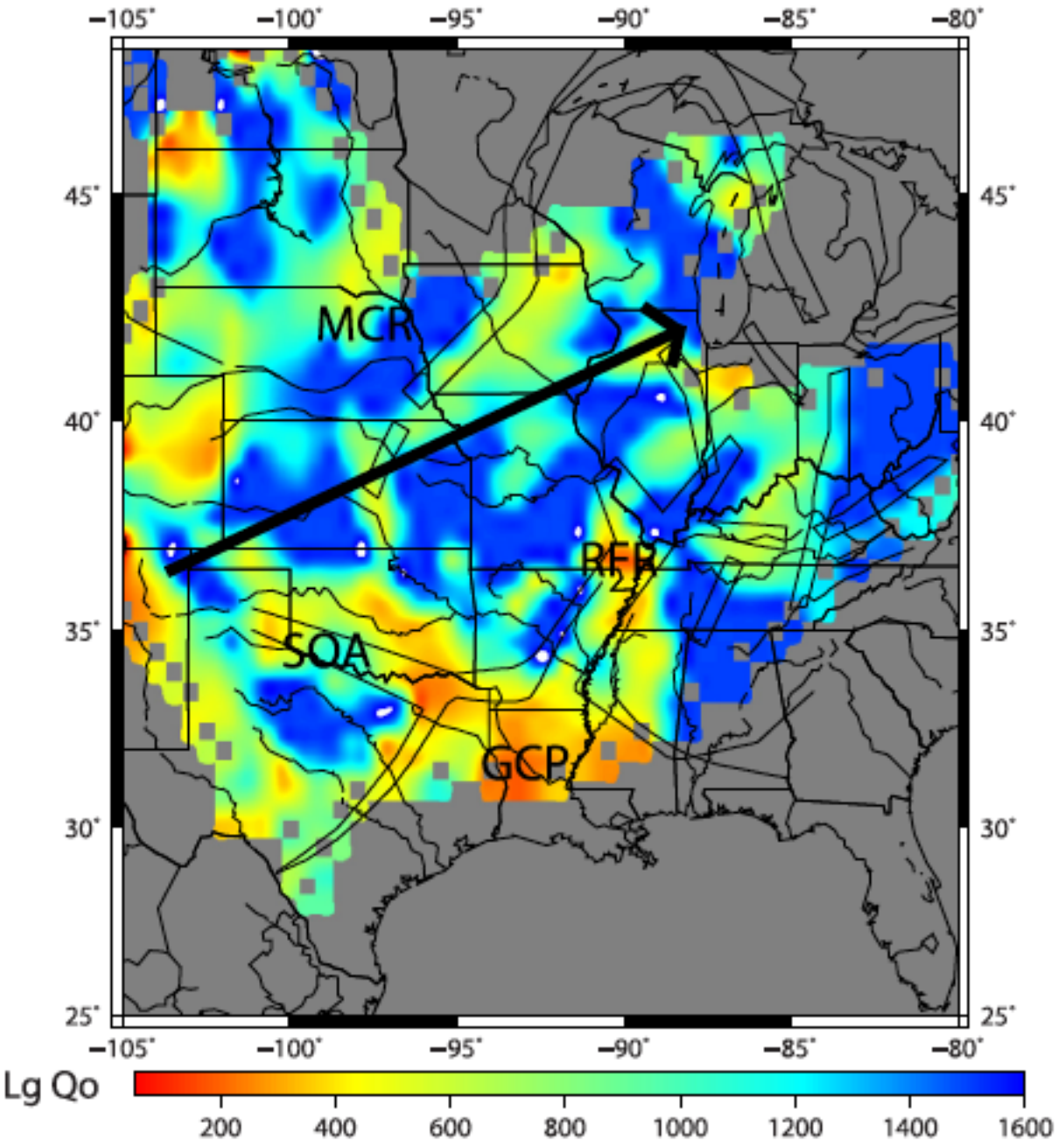


Figure 4–7 Boundary between Mid-Continent and Gulf Coast Q Regions for earthquakes occurring outside of the Gulf Coast Region proposed by Cramer and Al Noman (2016) for 5 Hz and 1 Hz spectral accelerations.

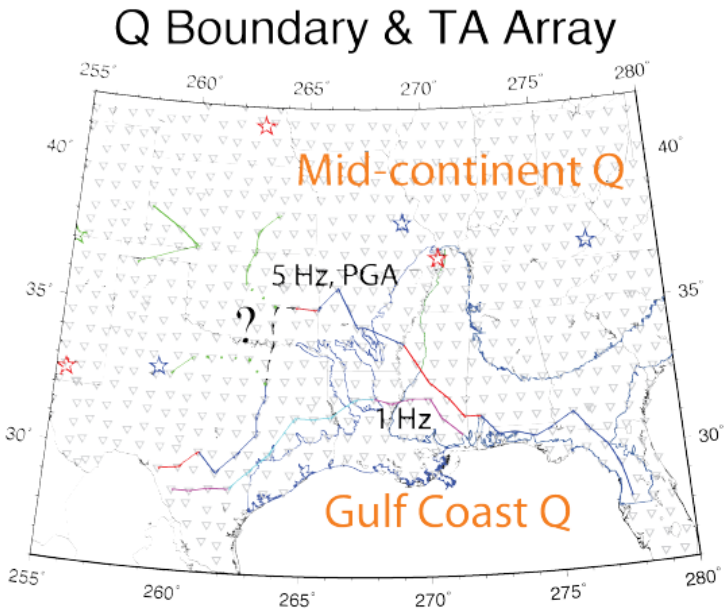
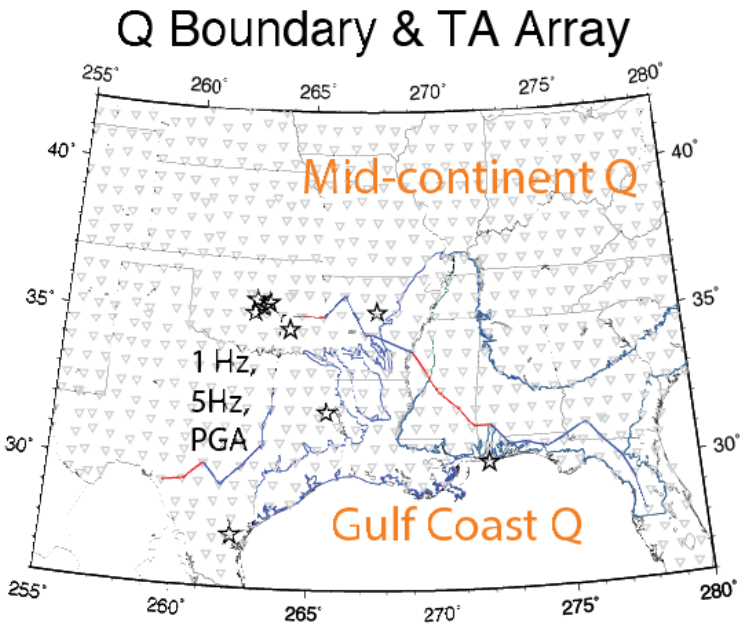


Figure 4–8 Boundary between Mid-Continent and Gulf Coast *Q* Regions for earthquakes occurring within or near the Gulf Region proposed by Cramer and Al Noman (2016) for 5 Hz and 1 Hz spectral accelerations.



5. Summary of Databases

5.1 Introduction

As in the case of the NGA-West1 (Power et al. 2008) and NGA-West2 (Bozorgnia et al. 2014) projects, one of the first key tasks of the NGA-East project was to develop a high-quality empirical ground-motion database recorded in Central and Eastern North America (CENA) to be shared with the various ground-motion model (GMM) developers. The database development was carried out as a science task by the Pacific Earthquake Engineering Research Center (PEER); however, database development and dissemination are also important parts of the Senior Seismic Hazard Analysis Committee (SSHAC) process (Budnitz et al. 1997; NRC 2012). The main objective for NGA-East was to develop GMMs and associated logic trees for horizontal ground motions. Therefore, the NGA-East database development focused on recordings that have at least two horizontal components. If a third vertical component was available, it was also processed, but single vertical records were not included in the initial release of database products. This choice was made as a matter of priority given the scope of the project and addressed in SSHAC Workshop 1. The majority of the current chapter is devoted to documentation of the NGA-East database (Goulet et al. 2014). Section 5.2 is largely based on content from Goulet et al. (2014); more details on record processing is provided in the original report, which documents the work of the Database and Geotechnical working groups. Changes to the original report text were made to address comments from the Participatory Peer Review Panel (PPRP).

Two additional datasets are briefly summarized in Sections 5.3 and 5.4. Section 5.3 is devoted to data resulting from finite-fault simulations that were made available to the NGA-East GMM developers and is largely based on PEER (2015). Section 5.4 summarizes the key attributes of the NGA-West2 database (Ancheta et al. 2013; 2014). Although the NGA-West2 database was developed for active tectonic regions (ATRs), it covers wider magnitude and frequency ranges than the NGA-East database. Various GMM developers investigated that dataset to constrain ground-motion scaling rules for application to CENA. Finally, Section 5.5 summarizes an additional dataset from Japan, which was used in the standard deviation modeling task.

5.2 NGA-East Database

This section summarizes the development of a large, high-quality dataset of earthquake ground motions recorded in CENA. The database also includes metadata associated with the events, recording stations, and ground motions. This combination of carefully processed, high-quality records with metadata into a single repository serves as an important building block for GMM development.

Data challenges for stable continental regions (SCRs) such as CENA are different from those of ATRs. The data are sparser in magnitude, distance, and frequency range compared to those of ATRs, and are generally of lower amplitudes, thus requiring different thresholds for an acceptable range of signal-to-noise ratio. The low attenuation rate in SCRs also requires longer durations to fully capture the motions until the end of the coda waves. The NGA-East Project addressed those issues early on and set the appropriate requirements to develop the database.

The NGA-East project is also unique in that GMMs were developed using two different approaches, which required additional data products. The first approach, consistent with what was carried out in the NGA-West projects, was to develop models using recorded pseudo-spectral acceleration (PSA). The second approach involved the initial development of GMMs in the Fourier amplitude spectral (FAS) space; see Chapter 11 in PEER (2015).

The NGA-East database consists of three groups of complementary products: (1) the summary file referred to as the flatfile, which contains metadata, ground-motion information, and intensity measures on a record-per-record basis; (2) the time series (acceleration, velocity, and displacement); and (3) the corresponding Fourier spectra files. This section summarizes the key elements of the data collection, processing, and development of data products for the NGA-East database. Again, more details are provided in Goulet et al. (2014), which also has compiled the long list of collaborators involved in the database development.

5.2.1 Summary of Data Collection

Data selection criteria were developed by the NGA-East Database Working Group documented in Cramer (2008) and Cramer et al. (2013). The criteria are to include all **M** 4+ earthquakes and add selected well-recorded **M** ≥ ~2.5 earthquakes with five or more records within 100 km. The five or more records within 100 km rule for **M** < 4 earthquakes is to avoid selecting smaller earthquakes with too few records for GMM development and to focus on selecting earthquakes with more important records at distances less than 100 km. Due to recording station coverage, effectively these criteria apply to CENA earthquakes from 1988 onwards, with the bulk of the records comprising earthquakes since 2000.

Table 5–1 lists the earthquakes selected using the criteria mentioned above for inclusion in the NGA-East ground-motion database. All earthquakes were assigned a unique integer number (EQID) and event names that include the event location (or a shorthand version of it) and the date, which have been used to organize all the data into event-specific folders (e.g., CapRouge97-11-06). Fortunately, recording-station density has increased significantly in the last decade, and several recent **M** > 4 earthquakes have been better recorded; the EarthScope Transportable Array (TA) traversing the U.S. has contributed significantly to the better recording of these recent earthquakes. However, due to poor station coverage (e.g., the St-Teresa, Mexico, earthquake) and other factors, not all of the **M** 4+ CENA earthquakes have been; this has led to gaps in the EQID sequence after events were rejected. It is believed that the NGA-East ground-motion database represents well the source regions in CENA.

After discussions in Workshop 1, the 1982 Miramichi and 1985 Nahanni strong-motion records have also been included, in addition to the data meeting the selection criteria since 1988 (Figure 5–1). Also, key large-magnitude data from CENA and analogous regions have been included: the 1925 Charlevoix, 1929 Grand Banks, 1935 Timiskaming, 1944 Cornwall-Massena, 1976 Gazli, and 2001 Bhuj earthquakes.

Additional datasets were evaluated for inclusion, such as those from Australia, Europe (Norway and the United Kingdom), South Africa, Russia, and Korea. These datasets for regions considered to be analogous to CENA were discussed in several Database WG meetings and in SSHAC Workshop 1. The Technical Integrator (TI) team deemed that the data was quite limited

and that it was not clear that these other SCRs behaved like CENA. In addition, difficulties—both institutional (raw data not always openly shared) and technical (such as the limited access to station information for processing)—regarding data acquisition outweighed their potential benefit. Following Workshop 1 (see Appendix B), the effort to collect recordings from other SRC regions was halted. As mentioned earlier, as a matter of resource allocation priority the focus of the project was to process and include horizontal components first.

A total of 30,167 horizontal (H) and vertical (V) component records (19,817 H, 10,358 V) from 89 CENA **M** $\geq$2.5 earthquakes (Figures 5–1, 5–2, and 5–3) from 1925 through 2011 were collected and prepared (instrument response removed, initial filtering to acceptable-signal band, quality assurance, etc.) for the database. This includes records from the April 18, 2008, **M** 5.2 Mt. Carmel, Illinois, mainshock and three **M** 4 aftershocks, the February 28, 2011, **M** 4.7 Greenbrier, Arkansas, earthquake, the August 23, 2011, **M** 5.7 Mineral, Virginia, mainshock and largest aftershock, and the November 6, 2011, **M** 5.6 Sparks, Oklahoma, mainshock and foreshock. In addition, an accelerograph recording of the May 17, 1976, **M** 6.8 Gazli, USSR, earthquake and engineering seismoscope observations for the January 26, 2001, **M** 7.6 Bhuj, India, earthquake have been included in the NGA-East ground-motion database. The Geological Survey of Canada (GSC) strong-motion recordings, previously not available, have also been added as they became available through funding from the NGA-East project. In particular, broadband station coverage has improved greatly in the last decade and, more recently, the EarthScope TA has added significantly to the density of observations in CENA (Figure 5–2). Additional earthquakes since 2000 have increased the number of ground-motion recordings in the 10–100 km range (Figure 5–3), in particular, the Mt. Carmel events, the 2005 **M** 5.0 Rivière-du-Loup, Québec, Canada, earthquake, and the Arkansas and Virginia events. Records from soil sites (with time-averaged shear-wave velocity in the upper 30 m V_{S30} < 1500 m/sec) have also been added to the database, which are needed for developing V_{S30} scaling terms in NGA-East GMMs. Available source (location, magnitude, focal mechanism, etc.) and site (geology, V_S profile, V_{S30}, etc.) information has been gathered as part of this effort and included in the ground-motion database. The information was grouped into three tables: the initial record catalog, the earthquake source database, and the station database. The record table was used as the basis for the development of the final project flatfile.

The broadband and accelerometer records were assembled from several sources:

- IRIS Data Center (IRIS)
- Canadian National Data Centre (CNDC)
- Geological Survey of Canada (GSC)
- Center for Earthquake Research and Information (CERI)
- St. Louis University Earthquake Center (SLU)
- U.S. Geological Survey's Advanced National Seismic System (USGS ANSS)
- Lamont Doherty Cooperative Network (LDCN)
- Virginia Polytechnic Institute's Southeast Network

- Weston Observatory's Northeast Network
- Hydro Québec (HQ)
- NetQuakes
- Quake Catchers

Also included in the database are hand-digitized records for long-period ground motions only from four pre-1950 large magnitude CENA earthquakes (Atkinson and Chen 1997), spectral acceleration data from the 2001 **M** 7.6 Bhuj, India, earthquake (Cramer and Kumar 2003), and a close-in accelerometer record from the 1976 **M** 6.8 Gazli, USSR, earthquake. All the records were obtained in digital form, but some have been originally recorded on analog systems.

The naming convention for files in the NGA-East database follows the IRIS convention, which is based on the station identifier of the downloaded data. The form of the station identifier is NN.SSS.IIO.AA. NN is the network identifier (see Table 5–2), SSS is a 3–5 character unique station id for that network, and II is a two-letter code for the instrument type. O is the instrument orientation (generally E, N, or Z for East, North, and Vertical), and AA is a two-character designator among different instruments at the same station (which is sometimes optional if only one recording instrument is present; these usually comprise datasets from older earthquakes). An AA of 00 (zero zero) is usually the same instrument as earlier records from a given station without the optional designation.

Table 5–2 lists details of the data sources by network. Included are network name, network owner/manager, data source, data type, and instrumentation information. The table is organized alphabetically by two-letter network code associated with the data files. The network name and owner/manager columns list this information as provided by the IRIS Network Codes website. The actual source (data source) accessed to obtain the data files for each network is also listed as some data did not come directly from the owner/manager with the most common acronyms defined above; the other sources are spelled out. Under the Data Type column, V is for velocity (broadband or short period), and A is for acceleration (strong motion). As to data format, SAC is for SAC files, SEED for SEED files (read using RDSEED from IRIS), MiniSeed for MiniSeed files, and ASCII files for text format files (specific to the provider). Instrument Type codes are the standard codes from the IRIS SEED manual. The first letter is the band code (B for broadband, H for high-gain broadband, E for extremely short period, and S for short period). The second letter is an instrument code (H for high-gain seismometer, L for low-gain seismometer, and N for accelerometer). The third symbol or letter is the orientation code (Z for vertical, N for north, E for east, and a dash for all three). The last two columns are Time Interval and Sampling Frequency, and are related (one the inverse of the other).

5.2.2 Summary of Time Series Processing

Time series were initially processed by the CERI team [as documented in Cramer et al. (2013)]. However, the NGA-East project required additional products, necessitating the re-processing of the time series. For example, the FAS for specific windows (for noise, P-, S-, Lg- and coda waves) needed to be extracted at the processing stage. The CERI team shared the collected raw time series and their processing package (codes and workflow including RDSEED and

SAC) for computing instrument-corrected acceleration time series. PEER then produced the instrument-corrected time series in ASCII format for use by the NGA-East processing team. The processing team built on software developed for the NGA-West2 database (Ancheta et al. 2013, 2014) and included additional computational capabilities to address NGA-East's specific data needs. Section 3.4 in Goulet et al. (2014) describes the processing scheme in detail, which consisted of the following key steps:

5.2.2.1 Inspection for Time Series Acceptance

The first step in the process involved determining the acceptability of time-aligned, instrument-corrected time series that were not otherwise processed or filtered, as described in the previous section. An initial visual inspection of the time series was used to select records with no apparent issues, e.g., clipping in amplitude or truncation in time. At this stage, each potential recording was either accepted or rejected. This approach is consistent with that applied in the NGA-West1 (Chiou et al. 2008) and NGA-West2 (Ancheta et al. 2013) projects.

5.2.2.2 Determination of Time Windows for Data Processing

Six distinct time windows were then determined for each of the accepted acceleration time series. Figure 5–4 shows these six windows schematically. The first time window contains only pre-event noise (magenta box). The second time window contains P-waves (yellow box) and includes a small buffer prior to the first arrival. The third time window contains S- and Lg-waves (green box) and is hereafter termed the "SLg-wave window." The fourth time window contains the coda waves (gray box). The fifth and sixth time windows are combinations of these first four windows. The fifth time window contains the P-, S- and Lg-waves, and represents the combined second and third time windows (orange box). The sixth time window encompasses all the windows in what we call the "entire recording" (blue box); FAS were computed for each time windows for all the recordings. The method for definition of each the six windows is based on expected travel times of the various phases and on visual inspection; the process presented in detail in Section 3.4 of Goulet et al. (2014).

5.2.2.3 Mean Removal, Tapering, and Computation of Fourier Amplitude Spectra

Mean removal is a standard step of record processing that allows for correction of the record for systematic trends in the time domain over the duration of each time window. Following the mean removal, a cosine taper is applied to the end of the time series, which is then padded by a series of zeros. Tapering and zero-padding are essential steps in time-series processing to prevent numerical issues with the fast Fourier transform (FFT) computations. This extended duration avoids the possible wrap-around effects that can occur in the time domain after applying acausal filters. The lengths of pre-event noise, P-, SLg-, and coda-wave windows were all increased to 50 min. with zeroes. This process provides a consistent frequency step (df) among the different windows and recordings, making the FAS usable directly for many applications without the need for users to decimate or interpolate each record individually.

5.2.2.4 Filtering

High-pass (low-cut) and low-pass (high-cut) filters were applied to only the window encompassing the entire record. The other four time windows discussed above were not filtered. The filtering was performed in the frequency domain. Causal and acausal Butterworth filters were applied in which five and four poles were used for the high-pass and low-pass filters, respectively. A discussion of these two types of filters (causal and acausal) in the context of NGA-East is presented in Section 3.4.4 of Goulet et al. (2014).

Corner frequencies were selected based on the theoretical acceleration decay at low frequencies consistent with the f^2 model (Brune 1970, 1971; Boore and Bommer 2005); Figure 5–5 is an example of a line consistent with the f^2 assumption. The FAS calculated from the accelerations in the pre-event noise window also helps in the selection of corner frequencies, in defining the transition from low to acceptable signal-to-noise levels.

Figure 5–5 shows the selected high-pass corner frequency ($f_{c\text{-}HP}$) for an example recording. This frequency was chosen based on a smoothed FAS. In the figure, the selected $f_{c\text{-}HP}$ is shown by the vertical line at 0.018 Hz, which is the frequency at which the entire spectrum deviates from an f^2 model. It also corresponds to the intersection of the entire spectrum with the increasing noise spectrum. A low-pass corner frequency ($f_{c\text{-}LP}$) was not selected for this record because no additional filtering was needed. However, for a large number of records, the noise spectrum was used to define the low-pass filter frequency in a fashion similar to that described for the high-pass filter. All the records in NGA-East database were processed following this method of determining $f_{c\text{-}HP}$ and $f_{c\text{-}LP}$. The usable frequency range was defined to accommodate unbiased computations of PSA. Since single degree oscillators sample a FAS bandwidth of roughly 25% below and above its natural frequency, and considering the smoothness of the Butterworth filter around the filter frequency, we define the usable range as 1.25 times the lowest frequency and 0.80 time the highest frequency as is PEER standard practice (Ancheta et al. 2013; 2014). This prevents “undesirable” parts of the filtered signal to be used in the PSA computations.

5.2.2.5 Baseline Correction

After filtering was completed, the zero pads were removed from the series (hereafter called the “pad-stripped” time series). The acceleration time series was then double integrated in the time domain to compute the displacement time series.

Figure 5–6 (top) shows the pad-stripped acceleration time series after applying an acausal filter. The middle and bottom figures the show velocity and displacement time series calculated by integrating the acceleration time series shown in the top figure. The displacement time series shown in the bottom plot displays a linear trend of baseline drift, which indicates that the velocity time series includes a constant offset throughout the recording. This drift occurred because an initial velocity (prior to the start of the actual signal) was added to the pad-stripped recording as a result of applying an acausal filter. Boore et al. (2012) showed that this drift does not occur if the correct initial velocity and displacement are used in integrating acceleration time series. They also discussed that this does not occur if zero-padded filtered time series are provided.

Records from PEER databases are often used by engineers in response history analyses of various systems; for this application, very long time series were considered impractical. To

address this issue, a baseline correction as described by Boore et al. (2012) was applied. The NGA-East project produced pad-stripped, baseline-corrected acceleration time series such that initial velocity and displacement can be assumed to be zero.

To process the pad-stripped time series, first a beginning taper is added to the acceleration series, which are then integrated to displacement. A sixth-order polynomial fit to the trend is computed as the baseline correction that is removed from the displacement time series. The time series is then differentiated back to acceleration. Figure 5–7 shows the baseline corrected, pad-stripped acceleration, velocity, and displacement time series corresponding to Figure 5–6. The process of baseline correction also ensures compatibility of the processed acceleration, velocity, and displacement time series as discussed in Chiou et al. (2008) and Boore et al. (2012). Figure 5–8 provides a comparison of PSA before and after applying baseline corrections. The figure shows that the difference in PSA between pad-stripped baseline-corrected and fully-padded time series is negligible.

5.2.2.6 Treatment of Microseisms

Microseisms not related to earthquake processes can sometimes dominate in time series when these are recorded from small magnitudes events at large distances. Figure 5–9 shows an example FAS from a record affected by microseisms. A comparison of the FAS for the entire window to that of the noise window shows that the time series is affected by microseisms for frequency ranges between 0.13–0.23 Hz (magenta line). Outside of that frequency range, the noise FAS is mostly flat. If the $f_{c\text{-}HP}$ is selected at 0.23 Hz, the processed records lose useable frequencies from 0.055–0.15 Hz, which is below the microseisms range. On the other hand, if the $f_{c\text{-}HP}$ is selected at 0.055 Hz, the processed record includes the microseisms. For NGA-East, the range of usable frequency may include microseisms; therefore, the recommendation is that, when applicable, for the user to also look up the microseism bandwidths in further restraining data usability. The frequency ranges affected by microseisms are also provided in the flatfile, allowing users to bypass that range of frequencies in their analyses depending on their specific application. Figure 5–10 shows the acceleration, velocity, and displacement time series for the example record in Figure 5–9 after applying the high-pass filter with $f_{c\text{-}HP}$ of 0.055 Hz.

5.2.3 Earthquake Source Metadata Table

The earthquake source table contains basic information about the seismic source, including earthquake name, origin time, magnitude, location, mechanism, and finite-fault geometry (strike, dip, length, width, top of rupture, etc.). As discussed in Section 5.3.1, the initial collection of data led to a list of 94 candidate earthquakes for the NGA-East database. The general source information is included in the database flatfile and provides the GMM developers with consistent source parameters and classifications for all the earthquakes in the NGA-East dataset. The complete source table contains additional detailed information—such as alternate source information—that is provided as part of the documentation (Appendices C.1 and C.2).

The initial earthquake source table developed by Cramer et al. (2013) was reviewed and modified by the EPRI (2013) project team. The final NGA-East source table includes those modifications, but it was further developed by the project team and documented in Goulet et al. (2014). The earthquake source table contains alternative moment tensor solutions along with a

set of preferred ones, assuming that such a judgment could be made from the literature review. Appendix C.1 summarizes how the earthquake source parameters included in the source table and described in the subsections below were obtained.

5.2.3.1 Finite-Fault Models

An earthquake's finite-fault model is a critical piece of information from which numerous other source and path data were derived. The finite-fault geometry can be defined by the end points of the top edge of rupture, the depth to the bottom edge of rupture, the fault dip angle, and the strike direction (Figure 5–11). We used rectangular fault planes, which are consistent with past GMM development and parameterization. Information about each finite-fault model was extracted from the available publications and converted to a uniform format and a latitude/longitude coordinate system, following an approach similar to that of previous projects (see references below). The areal extent of the rupture was a main issue in evaluating the finite-fault models. Finite-fault models for three CENA earthquakes were derived from published models (Saguenay, Rivière-du-Loup, and Mineral), and one was taken from the PEER NGA-West1 project (Nahanni). The NGA-West1 finite-fault models were built on three model collections previously used in ground-motion studies: PEER-NEAR (Silva et al. 1999), USGS-Yucca Mountain (YM) (Spudich et al. 1996), and Chiou et al. (2000). When a model included regions of zero or low level of slip near the edges, the model area was reduced or trimmed back [see Silva et al. (1999) for details].

5.2.3.2 Earthquake Source Parameters

In general, finite-fault inversions are not available for events with a moment magnitude less than about five. As most events in the NGA-East database have a magnitude less than five, the source parameters were chosen from earthquake and moment tensor catalogs or publications as available. Parameters that are typically available in the catalogs or in the literature included the event time, magnitude, strike, dip, rake, hypocenter location, and hypocenter depth. Appendix A of Goulet et al. (2014) documents the assessments for each earthquake in the NGA-East database.

Magnitude assignments for each event were based on data collection and magnitude estimates from both the NGA-East and EPRI (2013) database working groups. Both working groups collected magnitude estimates from the literature, and the final comprehensive list is contained in the source table.

The earthquake size measure selected for use in the NGA-East project is moment magnitude (**M**) related to seismic moment, M_0, by the Hanks and Kanamori (1979) relationship:

$$\mathbf{M} = 2/3 \log(M_0) - 10.7 \qquad (5–1)$$

where M_0 is the seismic moment (in dyne-cm). When multiple estimates of M_0 were available, each was converted to **M** using Equation (5–1) and the resulting values averaged. If estimates of M_0 were not available for an earthquake, then **M** was estimated from other magnitude scales using the relationships developed in Chapter 3 of NUREG-2115 (EPRI/DOE/NRC 2012). The earthquakes with magnitudes estimated from other size measures are listed Table 5–3.

Epicentral locations were selected from special studies of individual earthquakes published in the literature if available. Otherwise, the location was taken from online catalogs (the ANSS catalog for earthquakes located in the U.S. or the NRCan catalog for earthquakes located in Canada).

Two types of depth data are available. The standard catalog depths are the reported hypocentral depths. However, for many earthquakes, a moment centroid depth is reported as part of the central moment tensor solution. In most cases, the difference between these two depths is larger than the typical dimensions of rupture for the reported earthquakes. Where available, the moment centroid depth was selected as it is consistent with the calculated seismic moment used to define the magnitude. The hypocentral depths from seismic network location reporting typically have significant uncertainty because of sparse station density. Both depths are reported in the source table when available. When multiple depths are reported based on different moment tensor solutions, the average depth was selected, which is consistent with the averaging of the magnitude estimates.

Focal mechanisms are taken from the SLU website or published literature. If a preferred plane is indicated by the source, it is selected as the only focal plane. Otherwise, both planes are used in estimating distances. Where multiple focal mechanisms are reported, a single focal mechanism solution is selected for use in estimating distance measures. It is expected that minor differences in focal mechanism parameters will produce only small differences in estimated distances. Unresolved fault planes are treated as a source of uncertainty in the calculation of distance measures (Section 5.2.3.3).

If a high-quality moment tensor solution (or fault-plane solution) is not available for an earthquake, then its strike and dip are estimated from other associated events in the same region. Unknown fault strike is treated in the calculation of distance measures as a uniformly distributed random variable between -180° and 180°. The selected unknown fault strike is based on the methodology described in Section 5.2.3.3.

5.2.3.3 Method of Simulating a Finite-Fault Geometry for Distance Computation Purposes

In the interim NGA-East flatfile reviewed by the EPRI (2013) study, distance metrics such as the closest distance to the surface projection of the fault rupture plane—the Joyner-Boore distance (R_{JB}) —and the closest distance to the fault rupture plane (R_{RUP}), were missing. Similar to NGA-West2, a decision was made by the NGA-East project to adopt a method to simulate finite-fault planes for events without a finite-fault model but with minimal information of hypocenter, magnitude, and fault plane solution (or style of faulting). The goal of the simulation routine is to obtain an approximate fault-rupture geometry that may be used to compute distance metrics from the source model parameters (such as the fault-rupture width, W, and the depth to the top of the rupture, Z_{TOR}) that require knowledge of finite-fault geometry.

The simulation methodology described here is a modified procedure from Appendix B of Chiou and Youngs (2008). During the NGA-West2 project the methodology was modified to provide a more stable simulated finite fault when recordings are removed or added to the event. In this methodology, the missing fault plane information is filled in by random sampling of pertinent probabilistic distributions of fault-ruptured area, aspect ratio of ruptured area, and hypocenter

position on the fault plane, as summarized in Table 5–4. The simulation routine generates a set of 101 random fault ruptures that are rotated and translated in space, but fixed on the given hypocenter location. The strike and dip is fixed if reported. The routine computes the median R_{RUP} value of the 101 simulated ruptures using a grid of pseudo stations (700 spaced around the epicenter at epicentral distances from 0 to 300 km). A simulated rupture is selected based on a minimization of the squared difference between each pseudo-station median R_{RUP} value for all ruptures and that rupture. For the four final fault models listed in Section 5.2.3.1, the distances were computed directly from the geometry defined in Appendix C.1, without going through the randomization discussed above.

5.2.4 Event Classification: Potentially Induced Events

The CENA region has seen a sharp increase in earthquake activity rates since the mid-2000s (Ellsworth 2013), and it has been suggested that much of the upsurge is related to underground fluid injection. Research is currently on-going to better understand and characterize those earthquake events relative to those traditionally associated with natural processes. There is currently no clear method for discriminating between induced and natural (or tectonic) events. In addition, it is still unclear at this point if (or how) ground motions from this type of events are different from naturally occurring events. The study of this issue will continue well beyond the NGA-East project completion date. The NGA-East approach is to provide a flag for those potentially induced events (PIEs) identified as such by the USGS National Seismic Hazard Mapping Program (NSHMP) 2014 development team as described below.

Because the USGS maps underpin long-term engineering and public-policy guidelines like building codes, the NSHMP concluded it was undesirable for the seismic hazard to be conditional on unpredictable industrial activity that can start or stop for commercial or policy reasons. On this basis, the USGS identified and deleted these earthquakes from the catalogs that were used for the building-code hazard maps. Sequences were identified using information from the literature, and their extents were estimated by looking for suspicious activity within local sub-catalogs, parameterized by simple time windows and spatial polygons (Petersen et al. 2014). The USGS PIE flags were added to the NGA-East earthquake source table, but contrary to the 2014 NHSMP, NGA-East did not exclude records from these events from the final database, leaving this choice to the database user. The earthquake source table is provided in Appendix C.1.

5.2.5 Station Database and V_{S30} Assignment

Site parameters for the stations in the flatfile are taken from the station database (StDB), which contains information on site condition and housing for the 1379 stations with recordings in the project flatfile. The site information was originally compiled from the station table developed by the CERI team at the University of Memphis and dated 2013-02-28. That table was merged with the updated station metadata table released within NGA-East on 2014-04-15. The site parameter investigated is the time-averaged shear-wave velocity in the upper 30 m (V_{S30}). This entire section documents the work performed by the Geotechnical Working Group.

Values of V_{S30} have been measured for 84 of the 1379 stations (6%). For the other 1295 stations, V_{S30} was estimated from various sources available at that time. Values of V_{S30} are

estimated for a majority of sites using proxies, which is also common practice for active crustal regions (ACRs) (Chiou et al. 2008; Seyhan et al. 2014). The most commonly available proxies for the stations in the database are geology, terrain, and slope, which are available for 1375 (99.7%), 1367 (99.1%), and 1375 (99.7%) of the sites, respectively. Section 5.2.5.3 describes proxy-based methods for estimating V_{S30} from this data, which use geology only, slope only, terrain only, and a hybrid of slope and geology. In addition to proxies, the StDB also contains estimates of V_{S30} using a P-wave seismogram method (Kim et al. 2016), which is based on recorded ground motions and increases in accuracy as more ground motions are available for a site. The StDB contains P-wave estimates derived from single recordings for 42 sites and from multiple recordings for 81 sites (6%).

Subsequent to the original preparation of this report, the NGA-East Geotechnical Working Group performed a more thorough evaluation of proxy-based V_{S30} estimation in CENA (Parker et al. 2017). Our current recommendation is to use values from that document. However, the older estimates are retained here because of their use in the NGA-East project, as documented in the current report.

The categories of information contained in the StDB are as follows:

- Station metadata including name, location, and housing
- V_{S30} values derived from measurements
- Available proxies for the site, including geology, terrain category, and slope
- Inferred values of V_{S30} from known geology and local conditions, predominantly for sites we have received personal accounts of station installation or visit
- Proxy-based estimates of V_{S30} using ground slope, surface geology, terrain, and hybrid slope-geology methods
- Estimates of V_{S30} derived from P-wave methods (utilizing ground-motion data)
- Recommended values of V_{S30} with identification of the basis for the recommendation
- Recommended values of V_{S30} dispersion

The objectives of Section 5.2.5 (and its subsections) are as follows: (1) document key changes relative to the original CERI-developed station table; (2) describe principal sources of V_{S30} measurements; (3) identify proxy-based methods for estimating V_{S30} [prior to the Parker et al. (2017) work]; and (4) explain the protocols developed for assigning preferred V_{S30} values and their uncertainties to a site conditioned on the amount of information available from proxies and measurements in CENA [again, prior to the Parker et al. (2017) work]. The methods and formulations are modified only slightly from those developed during the NGA-West2 Project, as presented in Seyhan et al. (2014) and Ancheta et al. (2013, 2014). Appendix C.1 contains the complete StDB in Excel format and an explanation of column headings.

5.2.5.1 Station Metadata

Considerable effort was dedicated to updating the station metadata assembled by Cramer et al. (2013). The original station table was developed on an instrument-specific basis, which can be impractical for GMM development as discussed below. The organization of the station metadata data was therefore modified to location-specific basis.

In the CENA region, including parts of Canada and the U.S., there are often several seismic instruments at a single location (such as broadband or high-gain seismometers and/or strong-motion accelerometers). Some instruments were permanently removed at some point. The evolution of the specific instrumentation layout is continuous, involving changes such as sensor type, gain level settings, and/or digitizers over time, but the location of these instruments tends to remain essentially fixed, which we refer to as a station. In addition, a single instrument can also have different names, depending on the network operator responsible for distributing the data at a given time. This is especially true for stations shared by Canada and the U.S. or, for example, when data is distributed from the GSC, IRIS, and/or SLU data centers (see Section 5.2.1 for details).

In the context of NGA-East GMM development, each record was processed individually, and its specific bandwidth is reported in the project flatfile. Keeping instruments separate in the station flatfile is not only unnecessary but can lead to incorrect statistics during model development. That would be the case, for example, if recordings of a single event from co-located instruments are treated as independent measurements. Therefore, one key change relative to the original station file was to designate a unique station ID number to each “station” that is independent of the number of instruments installed at that location. This station ID was selected from those available in the original station table. Alternate station IDs and names for stations shared between networks are also provided in the current version of the StDB, and alternate instrument designations were also collected and documented whenever possible. Keeping the alternate station IDs for each station allows cross-referencing between the older station files and the officially released ones; therefore, the different V_{S30} estimates discussed below are associated with a single site (station) at which one or more instrument(s) recorded earthquake ground motions.

5.2.5.2 Data Sources and Distribution of V_{S30}

The V_{S30} values in the NGA-East StDB are from a variety of sources and are summarized in Table 5–5. Many sources of V_{S30} values are available in the open literature. The distribution of V_{S30} values derived from measurements are shown in Figure 5–12. The small number of strong-motion sites with geophysical measurements (84) is in contrast to the situation for ACRs, where approximately 50% of recording stations have V_{S30} from measurements (Seyhan et al. 2014).

5.2.5.3 Proxy-Based Estimation of V_{S30}

5.2.5.3.1 Available Proxies

Various descriptors or quantitative metrics of site conditions have been proposed to develop estimates of V_{S30} in the absence of measurement. In the NGA-East project, the development of these estimates is critical because only 6% of stations in the database have measurements.

The methods and proxies considered in this study are as follows: (1) V_{S30} estimation using ground slope at 30 arc-sec resolution (Wald and Allen 2007; Allen and Wald 2009); (2) V_{S30} estimation using a geomorphological terrain proxy (Yong et al. 2012); (3) V_{S30} estimation based on surface geology (Kottke et al. 2012); (4) V_{S30} estimation by a hybrid slope-geology proxy (Thompson and Silva 2013) and a slope-only proxy in that same report; and (5) V_{S30} estimation by P-wave seismogram method (Kim et al. 2016).

Kottke et al. (2012) developed a procedure to estimate V_{S30} from surface geology. We utilized relatively small-scale (1:2,000,000) digital maps for the U.S. (Fullerton et al. 2003) and 1:5,000,000 scale maps for Canada (Fulton 1986). There are two potential issues with these maps: (1) in some cases the maps represent true surface geology, whereas in others they are bedrock geology maps (with no soil included); and (2) due in part to the small scale, there are occasions where sites of interest (strong-motion stations) plot in bodies of water. The issue of variable meaning of the mapped geology has not been resolved as of this writing, and current applications are based on the mapped geologic unit regardless of map type. When the geologic unit is not available (typically because a site plots in a body of water), a global average value of V_{S30} is assigned based on the database compiled by Kottke et al. (2012). For stations in this database, 1083 (79%) have estimates based on mapped geology and 292 (21%) use the global estimate.

The hybrid slope-geology proxy by Thompson and Silva (2013) provides two estimates for V_{S30}, which are referred to as ‘actual’ and ‘effective.’ The ‘actual’ value is based on a V_{S30} measurement that is in the vicinity of the station. The ‘effective’ value is adjusted from the ‘actual’ based on inferences of different surface geology at the measurement location and the instrument location (typically, shallow soil at measurement and inferred rock at the surface for the instrument). We used the ‘actual’ V_{S30} over the ‘effective’ V_{S30} to allow for comparison to V_{S30} measurements that may not have been corrected. The electronic supplement of this document provides both ‘effective’ and ‘actual’ V_{S30} estimates by hybrid slope-geology, but only the ‘actual’ estimates are considered here.

The P-wave seismogram method (Kim et al. 2016) allows for a V_{S30} estimate to be made at a location if a ground-motion recording is available. The accuracy of this approach improves when more than one ground-motion recording is available at a site with sufficiently high signal/noise ratio to reliably evaluate P-wave arrivals and amplitudes (this generally results in removal of data at distances > 300 km). The electronic supplement of this document provides all available V_{S30} estimates by the P-wave seismogram method, but only estimates derived from more than one ground-motion recording are evaluated.

5.2.5.3.2 Proxy Evaluation

The reliability of proxy-based estimations of V_{S30} in CENA is evaluated by comparing predictions of V_{S30} to values derived from geophysical testing. For each site ‘i’ where a value of V_{S30} is available from geophysical measurements, the residual, R_i, is calculated as:

$$R_i = \ln(V_{S30})_i - \overline{\ln(V_{S30})}_{(proxy,i)} \quad (5\text{–}2)$$

where $\ln(V_{S30})_i$ is the measurement-based V_{S30} for a site, i, and $\overline{\ln(V_{S30})}_{(proxy,i)}$ is the proxy-based estimate of V_{S30} for site 'i'. Model bias is estimated as the mean of the residuals (μ_{lnV}), and the standard deviation of the residuals is represented as (σ_{lnV}). This analysis of residuals is based on a small dataset of 84 strong-motion sites having geophysical data. A much larger profile database compiled by Kottke et al. (2012), and later updated by Parker et al. (2017) was subsequently used for proxy development and residuals analysis. Figure 5–13 shows the distribution of residuals for each proxy-based estimation method considered in this study.

The relative efficacy of the different proxy-based estimation techniques can be judged on the basis of bias and standard deviation of the residuals (μ_{lnV} and σ_{lnV}). Figure 5–14 shows these quantities at sites distinguished by proxy availability. Only 34 stations in the NGA-East StDB have available all considered V_{S30} estimates and measured V_{S30} values. The most substantial limitation is for sites for which multiple estimates by the P-wave seismogram method are available. Note that estimates by the terrain, slope, and geology-hybrid techniques appear to have substantial bias, whereas geology and P-wave methods are relatively unbiased. Standard deviations of the proxy-based estimates are much higher than is typical in ACRs (0.35–0.45) (Seyhan et al. 2014). Among the methods considered, the P-wave seismogram estimates of V_{S30} have the smallest dispersion. Because the size of the database is small, we did not adjust the proxies for application here to remove bias.

5.2.5.4 Preferred V_{S30} and Its Uncertainty

5.2.5.4.1 Method of Selecting Preferred V_{S30}

The process by which 'preferred' V_{S30} values are assigned in the StDB is as follows (the number corresponds to codes in the StDB file):

0. Assign V_{S30} from measurements
1. Assign V_{S30} from known site conditions and geology based on measurements at different locations but the same geologic conditions. This assignment is only used based on a recommendation or site visit from a geologist.
2. Estimate by P-wave seismogram method (Kim et al. 2016) for sites having multiple ground-motion recordings and corresponding V_{S30} estimates. This code is applied exclusively for sites where the estimated $V_{S30} \geq 760$ m/sec.
3. Estimate by slope-based method of Thompson and Silva (2013). This code is applied exclusively for sites where the estimated $V_{S30} \geq 760$ m/sec.
4. Use weighted average of V_{S30} estimates from all available proxies, including those from the P-wave seismogram method when multiple ground motions are available for the P-wave estimate.
5. Use weighted average of V_{S30} estimates from all available proxies. This code is used when estimates from the P-wave seismogram method are not available.

The codes are ranked in ascending order of priority for V_{S30} assignment. Code 1 requires knowledge about the placement and location of the specific station in question. As part of the study by Beresnev and Atkinson (1997), some stations in the CENA region had an inferred V_{S30} of 2000 m/sec based on firm rock measurements and site investigations; therefore, for this study the inferred V_{S30} values were used where a Code 1 assignment was available. There were

also sites with V_{S30} inferred from known site condition and documented in a series of personal communications between the database developers. For the use of Code 1, the geologic condition is typically verified by site visit by geologist; this is the most common code for firm-rock conditions.

Codes 2 and 3 use the estimation of V_{S30} by the P-wave seismogram method (Kim et al. 2016) and slope approach of Thompson and Silva (2013). Other proxy methods evaluated in this study do not provide large values of V_{S30} corresponding to hard-rock site conditions. It is for this reason that these two methods are essential in providing V_{S30} estimates higher than 760 m/sec.

The weights of V_{S30} estimates for Codes 4 and 5 are based on the proxy mean and standard deviation of residuals (μ_{lnV} and σ_{lnV}). A relative weight $(W_R)_i$ for proxy 'i' is taken from the inverse of the residual sum of squares of the mean and standard deviation of that proxy as follows:

$$(w_R)_i = \frac{1}{\mu_{lnV}^2 + \sigma_{lnV}^2} \tag{5–3}$$

This weighting functional form, which is related to the mean square error (MSE), assesses the quality of the proxy methods in terms of both their variation and degree of bias. The actual weight is computed from the relative weights through adjustment to ensure they sum to one. Tables 5–6 and 5–7 show the relative weights and proxy weights for use with Codes 4 and 5 preferred V_{S30} assignment, respectively. If only a subset of proxies is available for V_{S30} assignment, the weights are adjusted to sum to unity for those proxies. For example, if V_{S30} estimates are only available from terrain and geology proxies, the relative weights would be 0.37 and 0.63, respectively.

Figure 5–15 shows the distribution of code assignments for assigning recommended values of V_{S30} at all stations in the StDB. Figure 5–16 shows the distribution of the recommended values.

5.2.5.5 V_{S30} Uncertainty

The preferred V_{S30} value for a site is taken as the exponent of the mean estimate in natural log units. The assignment of epistemic uncertainty to V_{S30} is dependent on the estimation method, and the dispersion is expressed as a log standard deviation, σ_{lnV}. The uncertainty values are used to communicate the level of precision associated with the preferred V_{S30} values. These uncertainties can be included in Bayesian regression analyses when developing the site component of ground-motion models (Moss 2011).

5.2.5.5.1 Code 0 sites (sites with geophysical measurement of V_{S30})

Extensive work was undertaken in the NGA-West2 project to quantify the uncertainty of V_{S30} at locations with multiple V_S profiles (Seyhan et al. 2014). That study found that for relatively consistent terrain, σ_{lnV} was fairly stable and not particularly sensitive to site stiffness. For such conditions, a value of $\sigma_{lnV} = 0.1$ (roughly equivalent to ± 10% error) was found to reasonably represent the available data. This result is adopted for CENA as well, where we lack data to independently evaluate this dispersion.

5.2.5.5.2 Code 1 sites

Code 1 depends on inferred site and station condition. The recommended dispersion is σ_{lnV} = 0.3. This value is based on engineering judgment of the Geotechnical Working Group.

5.2.5.5.3 Code 2 and 3 sites

We recommend use of the σ_{lnV} values from the residuals analysis in Table 5–2, which are 0.46 and 0.57 for Code 2 and Code 3, respectively.

5.2.5.5.4 Code 4 and 5 sites

For the weighted proxy assignments of V_{S30}, uncertainty is assigned based on the analysis of the residuals. It is recommended that the site dispersion be calculated by weighting the σ_{lnV} of the site's available proxies by the values listed in Tables 5–6 and 5–7 for Codes 4 and 5, respectively.

5.2.6 NGA-East Database Products

The NGA-East database consists of several complementary products that are linked together by the summary table referred to as the flatfile. The flatfile contains a list of all the selected and processed strong-motion records, metadata information, selected ground-motion intensity measures and the paths to a series of text files for time series, and FAS products.

The processed acceleration, velocity, and displacement time series are provided for each available component of each recording in a format that is consistent with previous NGA database releases. The processed acceleration time series serve as input in the calculation of numerous ground-motion products, including the as-processed PSA for each component and $RotD_{50}$.

Finally, acceleration Fourier spectra files including both the amplitudes and phases are available for the processing windows described in Section 5.2.2. For all intermediate windows only the instrument-corrected unfiltered Fourier spectral information is available. For the complete time series, the corresponding Fourier spectral information is available for the instrument-corrected unfiltered time series as well as for the fully processed time series.

5.2.6.1 Flatfile Components and Organization

The flatfile lists all the available records in the database. Each record is assigned a Record Sequence Number (RSN), which is a unique identifier associated with a single instrument-event pair (e.g., record). The RSN refers to all the components of the record. Similarly, each event is assigned an earthquake identifier (EQID), and each station is assigned its own Station ID, as described above. For NGA-East specifically, RSNs, EQIDs, and Station IDs were assigned by Cramer et al. (2013) at the time the team designed the database. As records were evaluated and processed, some entries were dropped, leading to gaps in the numbering. Some records were also added to the database at a later date, so RSNs and EQIDs do not necessarily follow a natural time progression.

The flatfile information is grouped into the following categories (see Appendix C.1 for the complete flatfile fields list and documentation), with each line starting with the RSN:

- SOURCE Metadata: EQID and earthquake event metadata (extracted from the earthquake source table, Section 5.2.3)
- STATION Metadata: Station ID, station information, and site characterization metadata (extracted from the station database, Section 5.2.4)
- PATH Metadata: Path metadata, including various distance metrics and regionalization information (Chapter 4)
- RECORD-SPECIFIC Metadata: Record-specific metadata, including paths and file names for time series, time steps, processing corner frequencies, and quality flags (extracted from the processing information table, Section 5.2.2)
- GROUND-MOTION INTENSITY MEASURES: intensity measures obtained from the final processed time series are PGA, PGV and PSA.

The earthquake source table, the station database and the flatfile (along with a detailed description of fields in each file) are provided electronically in Appendix C.1. Most of the fields included in the flatfile do not require lengthy explanations, but a short description of selected intensity measures is provided below.

5.2.6.2 Available Ground-Motion Intensity Measures

The NGA-East project has selected the rotated ground-motion intensity measure called $RotD_{nn}$ described by Boore (2010). $RotD_{50}$ is the 50th percentile (or median) PSA value computed from two horizontal components, over all non-redundant rotations. It is the main PSA product considered by NGA-East. The minimum and maximum rotated response spectra, $RotD_{00}$ and $RotD_{100}$, were also computed for the project but are not directly used in GMM development. $RotD_{nn}$ provides an average horizontal PSA measure that is independent of sensor orientation. The NGA-East database released for use in GMM development provides 5%-damped PSA.

5.2.6.2.1 $RotD_{nn}$

As described in Boore (2010), the $RotD_{nn}$ spectra are a set of response spectra over all non-redundant rotation angles where 'nn' represents the fractile of the spectra sorted by amplitude. The 'D' indicates that rotation angle will be specific to the period of the oscillator. The $RotD_{nn}$ can be computed from the rotation of the two as-recorded orthogonal horizontal ground motions. For any rotation angle, θ, the rotated time series, a_{ROT}, can be computed from the orthogonal horizontal-component time series, $a_1(t)$ and $a_2(t)$, using Equation (5–4):

$$a_{ROT}(t,\theta) = a_1(t)\cos(\theta) + a_2(t)\sin(\theta) \qquad (5\text{–}4)$$

The response spectra for the rotated time series are calculated for the non-redundant rotation angles 0−180°. The median fractile (nn = 50) was used for NGA-East. The maximum component (nn = 100) was also provided on the PEER online tool (PEER website) at the time of publication.

5.3 CENA Finite-Fault Simulation Results

This section is partly based on sections from PEER report 2015-04 (Chapter 1 appendix to PEER 2015).

Following a large finite-fault validation exercise, three finite-fault simulation models were selected for the generation of CENA ground-motion data. The methodologies are implemented on the Southern California Earthquake Center (SCEC) Broadband Platform (BBP), version 14.3.6, and are documented, along with the validation exercise itself, in a Focus Section in Seismological Research Letters (Volume 86, Issue 1). We briefly summarize the BBP, validation exercise, its results, and how methods were selected to generate datasets for NGA-East.

5.3.1 SCEC Broadband Platform: Overview

The BBP has been developed and released as an open-source scientific software package that can generate broadband (0–100 Hz) ground motions for earthquakes by integrating complex scientific modules that include rupture generation, deterministic and stochastic seismogram synthesis, nonlinear site effects, and visualization into a software system that supports on-demand computation of seismograms. The BBP was developed in a collaborative effort that involves geoscientists, engineers, and computer scientists, and is constantly being refined and upgraded.

The BBP uses finite-fault models, which account for complex, yet simplified, source effects in a computationally efficient manner. The user specifies a hypothetical earthquake description, a list of station names and locations, and a 1D velocity model for the region of interest, and the BBP software then calculates ground motions for the specified stations. The BBP operates in two modes: validation simulations and forward scenario simulations. In both modes, the BBP runs earthquake rupture and wave propagation modeling software to calculate seismograms.

In validation mode, two types of schemes are available: (1) Scheme A compares simulated ground motions to past recorded ground motions in one case; and (2) Scheme B compares the simulated ground motions to GMMs. A suite of goodness of fit (GOF) parameters that quantify how well the synthetics match the observations are also generated. This validation exercise evaluation identified which simulation method(s) are deemed reliable for specific tectonic regions, magnitude, and distance ranges and ground-motion parameters. The validation model has been used to evaluate and validate a variety of numerical ground-motion modeling techniques that are now built into the platform. Once the validation is completed, the platform can be used to simulate events that have not yet occurred. This is referred to as ‘forward’ (scenario) simulations.

5.3.2 Validation Exercise Summary

The BBP validation was performed based on the Goulet et al. (2015) procedure, which was developed in large part as a response to NGA-East’s needs. The validation completed for NGA-East and BBP V.14.3.6 is documented in Dreger et al. (2015). Five simulations were considered in the evaluation (see Table 5–8). The simulation methodologies were evaluated for applicability to Western U.S. (WUS), Japanese, and CENA events. The BBP evaluation exercise design is detailed in Goulet et al. (2015) and the evaluation results are presented in Dreger et al. (2015).

The BBP itself is described in Maechling et al. (2015). The five methods considered for evaluation are also documented in the special issue: Atkinson and Assatourians (2015) describes the EXSIM (EX) method, Graves and Pitarka (2015) describes the Graves and Pitarka (GP) method, Olsen and Takedatsu (2015) describes the San Diego State University (SD) method, Anderson (2015) describes the Composite Source Model (CSM), and Crempien and Archuleta (2015) describes the University of California, Santa Barbara method. The methodology papers describe the physics behind each simulation approach along with the rules for parameter selection in each of the region considered.

The validation exercise was designed to accommodate the simulation needs from the Southern Western U.S. utilities project (GeoPentech 2015) and the needs of NGA-East. The two projects have different specific needs, but they share a similar interest in the simulations: to fill the gap in recorded datasets for PSA to support their ground-motion characterization (GMC) model building. The validation exercise had the limited scope of assessing the ability of different simulation methods to reproduce reasonable average PSA values only. It is understood that many other metrics would be necessary to fully assess ability of simulations methods to produce reasonable ground motions as a whole. These won't be addressed here. However, because the focus of herein is on the design of a validation exercise conducted for a specific application, the key elements described here are portable to other metrics and applications. The validation evaluation, described in detail in Dreger et al. (2015), is based on the performance of the different methodologies in matching the PSA of recorded ground motions and empirical relations for a set of earthquake scenarios and stations. Where data was available, the target validation metric is the RotD50 5%-damped PSA for spectral periods in the 0.01–10 sec range.

The validation is carried out for cases where data are available; however, the intended use is for extrapolated cases (e.g., considering larger magnitude, closer distance, hanging wall or directivity effects). The validation results can be quantified by the goodness-of-fit of results with observed ground motions and aggregated so that an objective set of criteria is used for the evaluation (Dreger et al. 2015). The confidence in using methods beyond the tested limits must also be assessed in light of the science behind each method. For this validation exercise, each modeler provided the technical documentation of their method and a self-assessment of the expected performance for cases for which no data are available. The review panel's role was to provide final recommendations based reviewing the quantitative results, the documentation, and the modelers' assessment.

The validation was performed against historical events' data (Part A validation) and against CENA GMMs (Part B validation). The scenarios considered for Part A are summarized in Table 5–9. The validation panel considered various metrics, all based on PSA, for Part A. A suite of plots were generated for review:

- Simulated times series and Husid plots of simulated versus recorded ground motions: the purpose of which is to provide a qualitative level of review that the time series "looked like" seismograms.
- Simulated versus as-recorded PSA: again, a visual assessment was performed that the response spectra looked reasonable in their shape and levels.

- Mean PSA bias served as the main metric for the evaluation and was referred to as the goodness-of-fit (GOF) measure during the validation exercise. A suite of plots were produced to aggregate trends of GOF with
    - Response period. Purpose: evaluate trends with response periods, potentially identifying systematic effects within and across scenarios.
    - Station location on a map for given periods. Purpose: evaluate systematic trends due to directivity, site effects, distance, etc.
    - Distance for given periods. Purpose: detect systematic bias in observed attenuation.
- Mean PSA bias plots were also generated using four of the NGA-West2 GMPEs as the model (Abrahamson and Silva 2008; Boore and Atkinson 2008; Campbell and Bozorgnia 2008; and Chiou and Youngs 2008). Purpose: provide an average event term estimation as a function of period.

In addition to the review of the above plots, aggregated metrics were computed for the GOF for each method relative to that of the GMPEs. Overall, the purpose was to evaluate whether the simulation matched the data better, about the same, or not as good as the GMPEs, using pre-assigned thresholds defined by the review panel. This was assessed for all the events considered in Part A and in a matrix of four distance bins (0–5, 5–20, 20–70, and 70–200 km) and four PSA bandwidths (0.01–0.1, 0.1–1.0, 1.0–3.0 and 3–10 sec).

For the CENA Part B validation, the SCEC BBP software team implemented three GMMs for validation purposes:

- Atkinson (2008), with Atkinson and Boore (2011) revisions, and as modified by EPRI (2013)
- Pezeshk et al. (2011)
- Silva et al. (2002) Single-Corner, Variable Stress Parameter model

The validation against GMMs (Part B) was intended to provide a general guidance into the performance of the simulation methods, and an exact fit to the GMMs was not desired nor imposed. [See Dreger et al. (2015) for more details.] The summary of the validation events is provided in Tables 5–9 through 5–11 and illustrated in Figure 5–17.

The review panel considered all the products described above and the documentation of the methodologies provided by the developers in its assessment. Dreger et al. (2015) found that the ground motions from four of the methods (EX, GP, SDSU, and UCSB) could be used to inform extrapolation of ground motions to large magnitudes, but they should not be used for their absolute values. The CSM method was not deemed ready for use at the time.

The validation exercise was rerun for BBP V.14.3.6, which integrated minor modifications relative to 14.3, and the original evaluation panel led by Dreger, as well as the TI team, reviewed the results from this new validation, with a special attention to the fit to CENA events. NGA-East Workshop 2B in July 2014 featured presentations and discussions on this topic

(Appendix B). Final recommendations from these interactions are reflected in the published version of Dreger et al. (2015), which are pasted below:

"The results from the part A and part B validations indicate that, under the criteria of this panel's evaluation, the UCSB, EXSIM, G&P, and SDSU methods can provide acceptable estimates of median PSA from 0.01 to 3 s oscillator period within the validation magnitude range (M 5.4–7.2) for earthquakes in California or in comparable active crustal regions. The available data are too limited to claim the methods are adequately validated for stable continental regions (i.e., CENA). For both tectonic environments (active crustal and stable continental), it is the opinion of the panel that the simulations can be used to provide insights into relative effects, or scaling relationships, that are unresolvable from data alone. For example, in active regions, this may include analysis of changes in ground motion due to changes in source geometry, rupture direction, presence of secondary slip on splays, and hanging-wall effects. In stable continental regions, the recommendation is to use simulations for the development of fundamental within-method scaling relationships only (with magnitude and distance). Regardless of region, these relative effects are best utilized in combination with base case motions not derived solely from simulations (e.g., from a semi-empirical GMPE)."

The UCSB method could also not model the small Rivière-du-Loup event due to its low magnitude, underscoring a limitation of the method. Based on this limitation and the marginal fit to the data from the UCSB method, the TI team did not recommend the use of the method for NGA-East purposes.

5.3.3 Forward Simulations for CENA Ground-Motions

Following the recommendations from above, the EXSIM (EX), Graves and Pitarka (GP), and San Diego State University (SD) methodologies were selected to compute forward simulations for NGA-East. The compiled version of the platform (V.14.3.6), forward simulation ground motions and the post-processing products described in this section are the key products from this task. The NGA-East project was in agreement with Dreger et al. (2015) that the ground motions from these methods should not be used for their absolute values but rather for their relative magnitude scaling effects on ground motions; NGA-East developed a set of simulation scenarios for that purpose. The different earthquake scenarios and station layouts were defined to capture the effect of **M**-scaling relative to **M** = 5 for a range of distances.

The three selected simulation methods (EX, GP, and SD) were used to generate forward simulations for NGA-East. Finite-fault simulations on the BBP generally require large computation resources, and a series of sensitivity simulations were conducted in the summer of 2014 in an attempt to limit the number of runs required for NGA-East. It was found that ground-motion ratios were not significantly sensitive to the fault mechanism or to the velocity structure used in the simulations. The forward simulations were therefore designed to use a single-velocity structure and a single-fault mechanism and dip (reverse with a dip of 45° and an average rake of 90°). The fault dimensions were scaled to generate **M** of 5.0, 5.5, 6.5, 7.5, and 8.0, using the Leonard (2010) relations for SCRs. The depth-to-top of rupture plane (Z_{TOR}) was varied to generate additional motions to 0 and 5 km for all the scenarios, as well as to 10 km for **M** < 8.0. A total of 14 earthquake scenarios were defined by these parameters (see Table 5–11). The simulations were completed for footwall conditions.

Stations were arranged in 12 half-rings set at specific distances (R_{RUP}) to the fault trace on the footwall side. Each half-ring contained ten equally-spaced stations at R_{RUP} values of: 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, and 100 km whenever possible (given the Z_{TOR} of the specific scenarios). Sample simulation layouts are shown in Figure 5–18.

For each methodology, the hypocenter location and source properties were randomized so as to generate 16 realizations of the source for a given event scenario. This configuration led to the generation of approximately 160,000 time series (14 scenarios, 120 stations per scenario, 16 source realizations per scenario, 3 simulation methods, and 2 horizontal components per record).

5.3.4 Data Processing and Products

NGA-East required the output of forward simulations in terms of three basic metrics: acceleration time series, FAS, PSA, and with FAS and PSA ratios. The BBP outputs are time series (acceleration, velocity, and displacement) in ASCII format. In addition, NGA-East needed FAS and PSA ratios at various magnitudes relative to a reference **M** = 5 at comparable distances. A set of post-processors were implemented on the BBP to compute the 5%-damped PSA and $RotD_{50}$ PSA from the acceleration time series. $RotD_{50}$ PSA was computed from each horizontal pair of time series for GP and SD, while EX produced a single average component and a single PSA value. Acceleration FAS were computed outside of the BBP; further data aggregation was completed to develop distance-dependent **M**-scaling ratios. The data processing is summarized in Table 5–12. Figure 5–19 shows an example of processed data and computed ratios obtained after these calculations.

The data included the direct output from the BBP and the PSA and FAS values, and the ratios described above. The processed data (PSA and FAS ratio) as it was provided to the GMM developers is available in Appendix C.2.

5.4 NGA-West2 Database

A subset of GMM developers used data from ATRs and developed parts of their model using the NGA-West2 database (Ancheta et al. 2013; 2014). The NGA-West2 database includes earthquake events from multiple ATRs, such as from the WUS, Middle East, Japan, and China, among others. The key NGA-West2 product used in the NGA-East GMM development was the flatfile, which includes metadata on source, propagation, and site effects, as well as 5%-damped PSA $RotD_{50}$ values for 111 oscillator periods ranging from 0.01 to 20 sec and 11 different damping ratios. The database includes 21,336 three-component records from more than 700 shallow crustal events, covering a magnitude range of 3 to 7.9, source-site distances from 0.05 to 1533 km, and a range of V_{S30} of 94 to 2100 m/sec. Figure 5–20 shows the **M** and R_{RUP} data coverage for the NGA-West2 database. Figure 5–21 shows the NGA-East data coverage overlaid on Figure 5–20. The NGA-West2 database was developed outside of NGA-East and is available on the PEER website (peer.berkeley.edu) as an electronic appendix to Ancheta et al. (2013).

5.5 Japanese Dataset

The NGA-East study made use of the single-station analysis performed on Japanese crustal ground-motion data by Dawood (2014) to support the development of the aleatory variability model for CENA, particularly the site-to-site standard deviation ϕ_{S2S}. Insights obtained from the Japanese dataset were useful because of the general similarity in the site conditions between Japan and CENA (shallow soil cover overlying rock). The availability of ground-motion recordings at the surface and deep in the borehole in the Japanese dataset helped in the extrapolation of the ϕ_{S2S} for CENA for application to hard rock conditions with V_{S30} of 3000 m/sec.

A detailed description of the Japanese dataset and the development of a ground-motion model for active crustal earthquakes in Japan are detailed in Dawood (2014). A brief summary of the dataset relevant to the NGA-East study is presented herein. The Japanese dataset consists of ground-motion recordings from active crustal earthquakes recorded on the KiK-net stations. The dataset is comprised of 13,735 six-component (three at the surface and three at the borehole) ground-motion recordings from 679 active crustal earthquakes recorded at 643 stations. The V_{S30} values calculated for the KiK-net stations are based on seismic velocity profiles from downhole PS logging. Dawood (2014) used a GMPE functional form adopted from ASK14 for their regression analysis. Their ϕ_{S2S} results for the surface and the borehole levels were used for comparisons in the NGA-East study. Moreover, the within-event residuals were provided by the authors and used in the NGA-East study to perform single-station regressions and evaluate trends in ϕ_{S2S} for different aspects of the dataset.

Figure 4.47 of Al Atik (2015) (shown as Figure 5–22 here) shows the magnitude and distance distribution of the Japanese dataset. It consists of ground-motion recordings with **M** range of 4.0 to 7.0 and distance R_{RUP} of up to about 350 km. The number of stations histogram for the different V_{S30} bins is shown in Figure 5–23. The average V_{S30} for the surface recordings is 499 m/sec. The Japanese dataset provides borehole depth and the shear-wave velocity (V_S) at the borehole level, which is denoted $V_{S,Zhole}$. Figures 5–24 and 5–25 show the histograms of the number of stations versus bins of borehole depth and $V_{S,Zhole}$, respectively. The majority of the Japanese stations have a borehole depth of 98 to 150 m, with 31 stations having a borehole depth greater than 500 m.

5.6 Summary

This chapter summarized the various dataset used in the NGA-East project. The NGA-East database development was undertaken by PEER as a science task involving several working groups and participants. The preliminary version of the NGA-East database was shared with the EPRI (2013) project to support their work. The generation of simulation data was also performed as a science task. The TI team was involved in various steps of the development of these new datasets and recommended their use for GMM development. Additional datasets were also described above. The datasets described above were used for model development both by the GMM developers and the TI team, and for the evaluation and integration tasks by the TI team.

5.7 References

Abrahamson N.A., and W.J. Silva. 2008. “Summary of the Abrahamson & Silva NGA Ground-Motion relations.” Earthq. Spectra, 24 (1): 67–97.

Al Atik, L. 2015. “NGA-East: Ground Motion Standard Deviation Models for Central and Eastern North America.” PEER Report No. 2015/07, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Allen, T.I., and D.J. Wald. 2009. “On the Use of High-Resolution Topographic Data as a Proxy for Seismic Site Conditions (V_{S30}).” Bull. Seismol. Soc. Am., 99: 935–943.

Ancheta, T.D., R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.-S.J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida, and J.L. Donahue. 2013. “PEER NGA-West2 Database.” PEER Report No. 2013/03, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Ancheta, T.D., R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.-S.J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida, and J.L. Donahue. 2014. “NGA-West2 Database.” Earthq. Spectra, 30: 989–1005.

Anderson, J.G. 2015. “The Composite Source Model for Broadband Simulations of Strong Ground Motions.” Seismol. Res. Lett., 86(1): 68–74.

Assatourians, K. 2011. Personal communication.

Atkinson, G.M. 2008. “Ground Motion Prediction for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty.” Bull. Seismol. Soc. Am., 98(3): 1304–1318.

Atkinson, G.M., and K. Assatourians. 2015. “Implementation and Validation of EXSIM (a Stochastic Finite-Fault Ground-Motion Simulation Algorithm) on the SCEC Broadband Platform.” Seismol. Res. Lett., 86(1): 48–60.

Atkinson, G.M., and D.M. Boore. 2011. “Modifications to Existing Ground-Motion Prediction Equations in Light of New Data.” Bull. Seismol. Soc. Am., 101: 1121–1135.

Atkinson, G.M., and S. Chen. 1997. “Regional Seismograms from Historic Earthquakes in Southeastern Canada.” Seismol. Res. Lett., 68: 797–807.

Beresnev, I.A., and G.M. Atkinson. 1997. “Shear-Wave Velocity Survey of Seismographic Sites in Eastern Canada: Calibration of Empirical Regression Method of Estimating Site Response.” Seismol. Res. Lett., 68: 981–987.

Boore, D.M. 2010. “Orientation-Independent, Nongeometric-Mean Measures of Seismic Intensity from Two Horizontal Components of Motion.” Bull. Seismol. Soc. Am., 100(4): 1830–1835.

Boore D.M., and G.M. Atkinson. 2008. “Ground-Motion Prediction Equations for The Average Horizontal Component of PGA, PGV, and 5%-Damped PSA at Spectral Periods between 0.01 s and 10.0 s.” Earthq. Spectra, 24 (1): 99–138.

Boore, D.M., A. Azari Sisi, and S Akkar. 2012. “Using Pad-Stripped Acausally Filtered Strong-Motion Data.” Bull. Seismol. Soc. Am., 102: 751–760.

Boore, D.M., and J.J. Bommer. 2005. “Processing of Strong-Motion Accelerograms: Needs, Options and Consequences.” Soil Dyn. Earthq. Eng., 25: 93–115.

Bozorgnia, Y., N.A. Abrahamson, L. Al Atik, T.D. Ancheta, G.M. Atkinson, J. Baker, A. Baltay, D.M. Boore, K.W. Campbell, B.-S.J. Chiou, R.B. Darragh, S. Day, J. Donahue, R.W. Graves, N. Gregor, T. Hanks, I.M. Idriss, R. Kamai, T. Kishida, A. Kottke, S.A. Mahin, S. Rezaeian, B. Rowshandel, E. Seyhan, S. Shahi, T. Shantz, W.J. Silva, P. Spudich, J.P. Stewart, J. Watson-Lamprey, K.E. Wooddell, and R.R. Youngs. 2014. “NGA-West2 Research Project.” Earthq. Spectra, 30(3): 973–987.

Brune, J.N. 1970. “Tectonic Stress and the Spectra of Seismic Shear Waves from Earthquakes.” J. Geophys. Res., 75: 4997–5009.

Brune, J.N. 1971. “Correction.” J. Geophys. Res., 76: 5002.

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Clu, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. 1997. “Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts.” Report NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington D.C.

Campbell K.W., and Y. Bozorgnia. 2008. “NGA Ground Motion Model for the Geometric Mean Horizontal Component of PGA, PGV, PGD and 5% Damped Linear Elastic Response Spectra for Periods Ranging from 0.01 to 10 s.” Earthq. Spectra, 24 (1): 139–171.

Chiou, B.S.-J., F.I. Makdisi, and R.R. Youngs. (2000). “Style-of-faulting and Footwall/Hanging Wall Effects on Strong Ground Motion.” FY 1995 NEHRP Award Number 1434-95-G-2614, Final Report, 21 pgs.

Chiou, B.S.-J., and R.R. Youngs. 2008. “An NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra.” Earthq. Spectra, 24(1): 173–215.

Cramer, C.H. 2013. Personal communication.

Cramer, C.H. 2008. “Final Report on Initial Database Design for a Database of CEUS Ground Motions.” Final Technical Report to the USGS for first year of Cooperative Agreement 07CRAG0015, 26 pgs., https://umdrive.memphis.edu/ccramer/public/NGAeast/Documentation/CEUSGMInitialDatabaseDesignFinalRpt.doc.

Cramer, C.H., J.R. Kutliroff, D.T. Dangkua, and M.N. Al Noman. 2013. “Completing the NGA-East CENA/SCR Ground Motion Database.” Report prepared for the Pacific Earthquake

Engineering Research Center by the Center for Earthquake Research and Information, University of Memphis under Sub-agreement 7140, April 5, 2013, 18 pgs. https://umdrive.memphis.edu/ccramer/public/NGAeast/Documentation/NGAEastDbPEERFinalReport.docx

Cramer, C.H., and A. Kumar. 2003. "2001 Bhuj, India, Earthquake Engineering Seismoscope Recordings and Eastern North America Ground-Motion Attenuation Relations." Bull. Seismol. Soc. Am., 93: 1390–1394.

Crempien, J.G.F., and R.J. Archuleta. 2015. "UCSB Method for Simulation of Broadband Ground Motion from Kinematic Earthquake Sources." Seismol. Res. Lett., 86 (1): 61–67.

Dames and Moore. 1974. "GASSAR Seismic Design for General Atomic Company, Site Parameter Study." Dames & Moore Job No. 2395-001-02.

Dawood, H.M. 2014. Partitioning Uncertainty for Non-Ergodic Probabilistic Seismic Hazard Analyses. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Dreger, D.S., G.C. Beroza, S.M. Day, C.A. Goulet, T.H. Jordan, P.A. Spudich, and J.P. Stewart. 2015. "Validation of the SCEC Broadband Platform V14.3 Simulation Methods using Pseudospectral Acceleration Data." Seismol. Res. Lett., 86 (1): 39–47.

Ellsworth, W.L. 2013. "Injection-Induced Earthquakes." Science, 341 (6142), doi:10.1126/science.1225942.

EPRI. 2013. "EPRI 2004-2006 Ground-Motion Model (GMM) Review Project 3002000717." Final Report. June 2013, Electric Power Research Institute, Palo Alto, CA.

EPRI/USDOE/USNRC. 2012. "Central and Eastern United States Seismic Source Characterization for Nuclear Facilities." U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Fullerton, D., C. Bush, and J. Pennell. 2003. "Surficial Deposits and Materials in the Eastern and Central United States (East of 102 Degrees west longitude)." U.S. Geological Survey Geologic Investigations Series I-2789.

Fulton, R.J. (Compiler). 1986. Surficial Materials of Canada, Geological Survey of Canada. Natural Resources Canada. Ottawa. Map 1880A, Scale 1:5 000 000.

GeoPentech. 2015. "Southwestern United States Ground Motion Characterization SSHAC Level 3 – Technical Report Rev.2."

Geusebroek, J.M., G.J. Burghouts, and A.W.M. Smeulders. 2005. "The Amsterdam library of object images." Int. J. Comput. Vision, 61(1): 103–112.

Ghofrani, H.G., and G.M. Atkinson. 2014. “Site Condition Evaluation using Horizontal-To-Vertical Spectral Ratios of Earthquakes in the NGA-West 2 and Japanese Databases.” Soil Dyn. Earthq. Eng., 67: 30–43, DOI: 10.1016/j.soildyn.2014.08.015.

Goulet, C.A., N.A. Abrahamson, P.G. Somerville, and K.E. Wooddell. 2015. “The SCEC Broadband Platform Validation Exercise: Methodology for Code Validation in the Context of Seismic-Hazard Analyses.” Seismol. Res. Lett., 86 (1): 17–26.

Goulet, C.A., T. Kishida, T.D. Ancheta, C.H. Cramer, R.B. Darragh, W.J. Silva, Y.M.A Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. 2014. “PEER NGA-East Database.” PEER Report No. 2014/17, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Graves, R., and A. Pitarka. 2015. “Refinements to the Graves and Pitarka (2010) Broadband Ground-Motion Simulation Method.” Seismol. Res. Lett., 86 (1): 75–80.

Hanks, T.C., and H. Kanamori. 1979. “A Moment Magnitude Scale.” J. Geophys. Res., 84(B5): 2348–2350.

Herrmann, R.B., and D.J. Crossley. 2008. “Shear-Wave Velocity Determination for Metropolitan St. Louis Focal Mechanism.” USGS Grant 05HQGR0046, Final Technical Report.

Jaume, S.C. 2006. “Shear-Wave Velocity Profiles via Seismic Cone Penetration Test and Refraction Microtremor Techniques at ANSS Strong Motion Sites in Charleston, South Carolina.” Seismol. Res. Lett., 77: 771–779.

Kayen, R.E., B.A. Carkin, S.C. Corbett, A.S. Zangwill, L. Lai, and L. Estevez. 2015. “Shear Wave Velocity and Site Amplification Factors for 25 Strong-Motion Instrument Stations Affected by the M5.8 Mineral, Virginia, Earthquake of August 23, 2011.” U.S. Geological Survey, USGS Open-File Report 1099.

Kim, B., Y.M.A. Hashash, E.M. Rathje, J.P. Stewart, S. Ni, P.G. Somerville, A.R. Kottke, W.J. Silva, and K.W. Campbell. 2016. “Subsurface Shear-Wave Velocity Characterization using P-Wave Seismograms in Central and Eastern North America.” Earthq. Spectra, 32, 143–169.

Konno, K., and T. Ohmachi. 1998. “Ground-Motion Characteristics Estimated from Spectral Ratio between Horizontal and Vertical Components of Microtremors.” Bull. Seismol. Soc. Am., 88(1): 228–241.

Kottke, A.R., Y.M.A. Hashash, J.P. Stewart, C.J. Moss, S. Nikolaou, E.M. Rathje, W.J. Silva, and K.W. Campbell. 2012. “Development of Geologic Site Classes for Seismic Site Amplification for Central and Eastern North America.” Proceedings, 15th World Conference in Earthquake Engineering, Paper 4557, Lisbon, Portugal.

Leonard, M. 2010. “Earthquake Fault Scaling Self-Consistent Relations of Rupture Length, Width, Average Displacement and Moment Release.” Bull. Seismol. Soc. Am., 100: 1971−1988.

Lin, L., and J. Adams. 2010. “Strong motion records of the Val-des-Bois, Québec, Earthquake of June 23, 2010.” Canadian Hazard Information Service Internal Report 2010–1.1, 20 pgs. plus digital Appendix.

Maechling, P.J., F. Silva, S. Callaghan, and T.H. Jordan. 2015. “SCEC Broadband Platform: System Architecture and Software Implementation.” Seismol. Res. Lett., 86 (1): 27–38.

Mayne, P. 2011. Personal communication.

Moss, R.E.S., 2011. “Reduced Sigma of Ground-Motion Prediction Equations through Uncertainty Propagation.” Bull. Seismol. Soc. Am., 101, 250-257.

NRC. 2012. “Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies.” Technical Report NUREG-2117, U.S. Nuclear Regulatory Commission, Washington D.C.

NRCAN. Natural Resources Canada, Earthquakes Canada, http://www.earthquakescanada.nrcan.gc.ca/stndon/NEDB-BNDS/bull-eng.php, accessed June, 2014.

Odum, J. K., W.J. Stephenson, and R. A. Williams. 2010. “Predicted and Observed Spectral Response from Collocated Shallow, Active and Passive-Source Vs Data at Five ANSS Sites. Illinois and Indiana, USA.” Seismol. Res. Lett., 81, 955–964.

Olsen, K., and R. Takedatsu. 2015. “The SDSU Broadband Ground-Motion Generation Module bbtoolbox Version 1.5.” Seismol. Res. Lett., 86 (1): 81–88.

Parker, G.A., J.A. Harmon, J.P. Stewart, Y.M.A. Hashash, A.R. Kottke, E.M. Rathje, W.J. Silva, and K.W. Campbell. 2017. “Proxy-Based V S 30 Estimation in Central and Eastern North America.” Bull. Seismol. Soc. Am., 107 (1): 1–15.

PEER. 2015. “NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Petersen, M.D., M.P. Moschetti, M. Power, C.S. Mueller, K.M. Haller, A.D. Frankel, Y. Zeng, S. Rezaeian, S.C. Harmsen, O.S. Boyd, E.H. Field, R. Chen, K.S. Rukstales, N. Luco, R.L. Wheeler, R.A. Williams, and A.H. Olson. 2014. “Documentation for the 2014 Update of the United States National Seismic Hazard Maps.” U.S. Geological Survey, USGS Open-File Report 2014−1091, Reston, VA.

Pezeshk, S., A. Zandieh, and B. Tavakoli. 2011. “Hybrid Empirical Ground-Motion Prediction Equations for Eastern North America Using NGA Models and Updated Seismological Parameters.” Bull. Seismol. Soc. Am., 101: 1859–1870.

Power, M., B.-S.J. Chiou, N.A. Abrahamson, Y. Bozorgnia, T. Shantz, and C. Roblee. 2008. “An Overview of the NGA Project. Earthq. Spectra, 24: 3–21.

Read, K., H. El. Naggar, and D. Eaton. 2008. “Site-Response Spectra for POLARIS Station Sites in Southern Ontario and Québec.” Seismol. Res. Lett., 79: 776–784.

Saint-Louis University Earthquake Center Network (SLU), http://www.eas.slu.edu/eqc/, last accessed October 20, 2014.

Seyhan, E., J.P. Stewart, T.D. Ancheta, R.B. Darragh, and R.W. Graves. 2014. “NGA-West2 Site Database.” Earthq. Spectra, 30: 1007–1024.

Silva, W.J., N. Gregor, and R.B. Darragh. 1999. “Near Fault Ground Motions, A PEARL Report to PG&E/CEC/Caltrans.” Award No. SA2193-59652.

Silva, W.J., N. Gregor, and R. B. Darragh. 2002. “Development of Regional hard rock attenuation relations for Central and Eastern North America, Mid-continent and Gulf Coast areas.” Pacific Engineering and Analysis Report, 80 pgs.

Somerville, P.G., N.F. Collins, N.A. Abrahamson, R.W. Graves, and C. Saikia. 2001. “Earthquake Source Scaling and Ground Motion Attenuation Relations for the Central and Eastern United States.” Final Report to the U.S. Geological Survey, Contract No. 99HQGR0098.

Spudich, P., J.B Fletcher., M. Hellweg, J. Boatwright, C. Sullivan, W.B. Joyner, T.C. Hanks, D.M Boore, A. McGarr, L.M. Baker, and A.G. Lindh. 1996. “Earthquake Ground Motions in Extensional Tectonic Regimes.” U.S. Geological Survey, USGS Open-File Report 96-292.

Thompson, E., and W.J. Silva. 2013. “Empirical Assessment of Site Amplification and Development of NEHRP factors for CEUS: Collaborative Research with Pacific Engineering and Tufts University.” Report to USGS, http://earthquake.usgs.gov/research/external/reports/G12AP20003.pdf.

USGS. 2013. United States Geological Survey National Advanced National Seismic System website, http://earthquake.usgs.gov/monitoring/anss/, last accessed October 20, 2014.

USGS. 2014. United States Geological Survey National Digital Map Database, http://ngmdb.usgs.gov/ngmdb/ngmdb_home.html, last accessed October 20, 2014.

Wald, D.J., and T.I. Allen. 2007. “Topographic Slope as a Proxy for Seismic Site Conditions And Amplification.” Bull. Seismol. Soc. Am., 97: 1379–1395.

Williams R.A., W.J. Stephenson, J.K. Odum, and D.M. Worley. 2003. Seismic Velocities from High-Resolution Surface Seismic Imaging at Six ANSS Sites Near Memphis, Tennessee, U.S. Geological Survey, USGS Open-File Report 03–218, Reston, VA.

Yong, A., S.E. Hough, J. Iwahashi, and A. Braverman. 2012. “A Terrain-Based Site Characterization Map of California with Implications for the Contiguous United States.” Bull. Seismol. Soc. Am., 102: 114–128.

Table 5–1 Earthquakes considered for inclusion in the NGA-East ground-motion database. All were retained, except as indicated in the Comment column. Potentially-induced events (PIEs) are discussed in Section 5.2.4 and are flagged as such in the Comment column.

EQID	Earthquake	Location	M	Comment
1	Charlevoix_1925-03-01	Charlevoix QC	6.43	No data
2	GrandBanks_1929-11-18	Grand Banks NL	7.25	No data
3	Temiskaming_1935-11-01	Timiskaming QC	6.21	No data
4	CornwallMassena_1944-09-05	Cornwall Massena ON	5.79	No data
5	Saguenay_1988-11-25	Saguenay QC	5.85	
6	LaMalbaie_1997-08-20	La Malbaie QC	3.27	
7	LaMalbaie_1997-10-28	La Malbaie QC	4.29	
8	CapRouge_1997-11-06	Cap-Rouge QC	4.45	
9	CoteNord_1999-03-16	Cote-Nord QC	4.43	
10	Kipawa_2000-01-01	Kipawa QC	4.62	
11	LaMalbaie_2000-06-15	La Malbaie QC	3.29	
12	Laurentide_2000-07-12	Laurentide QC	3.65	
13	Laurentide_2000-07-12A	Laurentide QC	3.11	
14	Ashtabula_2001-01-26	Ashtabula OH	3.85	PIE*
15	Enola_2001-05-04	Enola AR	4.37	
16	AuSableForks_2002-04-20	Au Sable Forks NY	4.99	
17	LacLaratelle_2002-06-05	Lac Laratelle QC	3.81	
18	Caborn_2002-06-18	Caborn IN	4.55	
19	Boyd_2002-11-03	Boyd NE	4.18	
20	Charleston_2002-11-11	Charleston SC	4.03	
21	FtPayne_2003-04-29	Ft Payne AL	4.62	
22	Blytheville_2003-04-30	Blytheville AR	3.60	
23	Bardwell_2003-06-06	Bardwell KY	4.05	
24	LaMalbaie_2003-06-13	La Malbaie QC	3.53	
25	BarkLake_2003-10-12	Bark Lake QC	3.82	
26	Jefferson_2003-12-09	Jefferson VA	4.25	
27	StTeresa_2004-04-06	St Teresa MX	4.31	Rejected, outside of CENA, limited station coverage

EQID	Earthquake	Location	M	Comment
28	LaBaie_2004-05-04	La Baie QC	2.87	
29	PrairieCntr_2004-06-28	Prairie Center IL	4.18	
30	PortHope_2004-08-04	Port Hope ON	3.12	
31	MilliganRdg_2005-02-10	Milligan Ridge AR	4.14	
32	RiviereDuLoup_2005-03-06	Riviere Du Loup QC	4.65	
33	ShadyGrove_2005-05-01	Shady Grove AR	4.25	
34	Miston_2005-06-02	Miston TN	4.01	
35	Thurso_2006-02-25	Thurso ON	3.70	
36	Hawkesbury_2006-02-26	Hawkesbury ON	2.59	
37	BaieStPaul_2006-04-07	Baie Saint Paul QC	3.72	
38	Ridgely_2006-09-07	Ridgely TN	3.35	
39	GulfMexico_2006-09-10	Gulf of Mexico	5.85	Rejected, outside of CENA
40	Acadia_2006-10-03	Acadia ME	3.87	
41	Marston_2006-10-18	Marston MO	3.41	Rejected, mine collapse
42	Marvin_2006-11-02	Marvin VA	4.00	Rejected, mine collapse
43	Skeggs_2006-11-23	Skeggs VA	4.00	
44	Cobourg_2007-07-19	Cobourg ON	2.80	
45	BaieStPaul_2008-01-03	Baie Saint Paul QC	2.77	
46	MtCarmel_2008-04-18	Mt Carmel IL	5.30	
47	MtCarmel_2008-04-18a	Mt Carmel IL	4.64	
48	MtCarmel_2008-04-21	Mt Carmel IL	4.03	
49	MtCarmel_2008-04-25	Mt Carmel IL	3.75	
50	Buckingham_2008-06-11	Buckingham QC	2.97	
51	RiviereDuLoup_2008-11-15	Riviere Du Loup QC	3.57	
52	PineForest_2008-12-16	Pine Forest SC	3.16	
53	RoseHill_2009-01-29	Rosehill SC	2.77	
54	Palmetto_2009-05-06	Palmetto SC	2.18	
55	ConstanceBay_2009-05-08	Constance Bay ON	2.57	
56	Jones_2010-01-15	Jones OK	3.84	PIE
57	Lincoln_2010-02-27	Lincoln OK	4.18	PIE

EQID	Earthquake	Location	M	Comment
58	Whiting_2010-03-02	Whiting MO	3.40	
59	Lebanon_2010-05-21	Lebanon IL	2.62	
60	ValDesBois_2010-06-23	Val-des-Bois QC	5.10	
61	StFlavien_2010-07-23	St. Flavien QC	3.51	
62	Buhj_2001-01-26	Bhuj India	7.60	Rejected, outside of CENA, low quality data
63	MontLaurier_1990-10-19	Mont Laurier QC	4.47	
64	Montgomery_2010-07-16	Montgomery MD	3.42	
65	Gazli_1976-05-17	Gazli USSR	6.80	
66	Slaughterville_2010-10-13	Slaughterville OK	4.36	
67	Guy_2010-10-15	Guy AR	3.86	
68	Concord_2010-09-26	Concord NH	3.05	
69	Nahanni_1985-11-09	Nahanni NWT	4.40	
70	Nahanni_1985-12-23	Nahanni NWT	6.76	
71	Nahanni_1985-12-23a	Nahanni NWT	5.10	
72	Nahanni_1985-12-25	Nahanni NWT	5.15	
73	Arcadia_2010-11-24	Arcadia OK	3.96	PIE
74	BethelAcres_2010-12-12	Bethel Acres OK	3.23	PIE
75	Greentown_2010-12-30	Greentown IN	3.85	
76	Guy_2010-11-20	Guy AR	3.90	PIE
77	Greenbrier_2011-02-17	Greenbrier AR	3.83	PIE
78	Greenbrier_2011-02-18	Greenbrier AR	3.91	PIE
79	Greenbrier_2011-02-18a	Greenbrier AR	4.10	PIE
80	Greenbrier_2011-02-28	Greenbrier AR	4.68	PIE
81	Sullivan_2011-06-07	Sullivan MO	3.89	
82	EagleLake_2006-07-14	Eagle Lake ME	3.46	
83	ValDesBois_2010-06-24	Val-des-Bois QC	2.57	
84	ValDesBois_2010-07-22	Val-des-Bois QC	2.37	
85	Hawkesbury_2011-03-16	Hawkesbury ON	3.59	
86	Charlevoix_2001-05-22	Charlevoix QC	3.60	
87	BaieStPaul_2002-08-17	Baie Saint Paul QC	3.24	
88	Mineral_2011-08-23	Mineral VA	5.74	

EQID	Earthquake	Location	M	Comment
89	Mineral_2011-08-25	Mineral VA	3.97	
90	Sparks_2011-11-05	Sparks OK	4.73	PIE
91	Sparks_2011-11-06	Sparks OK	5.68	PIE
92	Comal_2011-10-20	Comal TX	4.71	PIE
93	Miramichi_1982-03-31	Miramichi NB	4.46	
94	Miramichi_1982-05-06	Miramichi NB	3.55	
116	Saguenay_1988-11-23	Saguenay	4.19	
117	Saguenay_1988-11-26	Saguenay	3.53	

Table 5–2 Data per network.

Network code	Network name	Network owner/manager	Data source	Data type	Data format	Instrument type	Time interval (sec)	Sampling frequency (Hz)
AG	Arkansas Seismic Network	Arkansas Geological Survey/CERI	CERI, IRIS	V,A	SAC, SEED	HH-/HN-	0.01	100
AO	Arkansas Seismic Observatory	University of Arkansas at Little Rock	CERI	V,A	SAC	HH-.00, HN-.10	0.01	100
CN	Canadian National Seismic Network	Geological Survey of Canada	CNDC, IRIS, GSC	V,A	SEED, ASCII Files	BH-, HH-, HN- , EHZ, SHZ	0.025, 0.01, 0.005, 0.01, 0.0166	40, 100, 200, 100, 60
CO	South Carolina Seismic Network	University of South Carolina, Columbia	IRIS	V	SEED	BH-/HH-	0.01	100
EP	UTEP Seismic Network	University of Texas, El Paso	IRIS	V	SEED	BH-	0.025	40
ET	CERI Southern Appalachian Seismic Network	CERI, University of Memphis	CERI, IRIS	V,A	SEED, SAC	HH-/HN-	0.01	100
GS	US Geological Survey Networks	USGS NEIC/NetQuakes	IRIS	V/A	SEED/ MiniSeed	BH-/HN-	0.025/ 0.005	40/200
HQ	Hydro-Québec Network (unofficial designation)	Hydro-Québec	Hydro-Québec	A	ASCII Files	HN-	0.005	200
II	IRIS/IDA Network	Scripts Institute of Oceanography, UC San Diego	IRIS	V	SEED	BH-	0.05	20

Network code	Network name	Network owner/manager	Data source	Data type	Data format	Instrument type	Time interval (sec)	Sampling frequency (Hz)
IU	IRIS/USGS Global Seismograph Network (GSN)	USGS Albuquerque Seismological Laboratory	IRIS	V	SEED	BH-	0.05/ 0.025	20/40
IU	IRIS/USGS Global Seismograph Network (GSN)	USGS Albuquerque Seismological Laboratory	IRIS	V	SEED	HH	0.01	100
IW	Intermountain West Seismic Network	USGS/ANSS Golden, CO	IRIS	V	SEED	BH-	0.025	40
LD	Lamont-Doherty Cooperative Seismographic Network	Lamont-Doherty Earth Observatory of Columbia University	IRIS	V	SEED	BH-	0.025	40
LD	Lamont-Doherty Cooperative Seismographic Network	Lamont-Doherty Earth Observatory of Columbia University	IRIS	V	SEED	HH-	0.01	100
LI	Laser Interferometer Gravitational-Wave Experiment	Caltech/USGS, Southern California Seismic Network	IRIS	V	SEED	BH-/HH-	0.05/ 0.01	20/100
NE	New England Seismic Network	Weston Observatory, Boston College/MIT	IRIS, Weston Obs.	V	SEED, SAC	BH-	0.025	40

Network code	Network name	Network owner/manager	Data source	Data type	Data format	Instrument type	Time interval (sec)	Sampling frequency (Hz)
NM	Cooperative New Madrid Seismic Network	St. Louis University/ University of Memphis	CERI, SLU, IRIS	V,A	SAC, SEED	BH-	0.05, 0.025, 0.02	20/40/50
NM	Cooperative New Madrid Seismic Network	St. Louis University/ University of Memphis	CERI, SLU, IRIS	V,A	SAC, SEED	HH-/HN-	0.01/ 0.005	100/200
NP	United State National Strong-Motion Network	USGS National Strong-Motion Program, Menlo Park	IRIS	V	SEED	HN-	0.005	200
NQ	NetQuakes	USGS Menlo Park	IRIS, CERI	A	SEED, SAC	HN-.01	0.005	200
OK	Oklahoma Seismic Network	Oklahoma Geological Survey	IRIS	V	SEED	HH-	0.01	100
PE	Penn State Network	Penn State University	IRIS	V	SEED	BH-	0.02, 0.01	50, 100
PO	POLARIS	Geological Survey of Canada and Canadian Universities	IRIS	V	SEED	HH-	0.01	100
QC	Quake Catcher Network	Stanford University	Stanford	A	SAC	BN-	0.1 to 0.02	10 to 50
RU	USSR Strong Motion Station (unofficial designation)	Unknown	Vladimir Graizer	A	ASCII files	HN-	0.005	200

Network code	Network name	Network owner/manager	Data source	Data type	Data format	Instrument type	Time interval (sec)	Sampling frequency (Hz)
SC	New Mexico Tech Seismic Network	New Mexico Tech, Socorro	IRIS	V	SEED	BH-	0.025	40
SE	Southeastern Appalachian Cooperative Seismic Network	Virginia Tech	CERI, Virginia Tech	V	SAC	HH-, EH-	0.01	100
SP	South Carolina Earthq Physics Project	University of South Carolina, Columbia	IRIS	V	SEED	BH-	0.05	20
TA	USArray Transportable Array	EarthScope Project, IRIS	IRIS	V	SEED	BH-	0.025	40
US	US National Seismic Network	USGS/NEIC, USGS/ASL, EarthScope Project of IRIS	IRIS	V	SEED	BH-/HN-	0.05, 0.025 /0.005	20, 40 /200
UU	University of Utah Regional Network	University of Utah, Salt Lake City	IRIS	V	SEED	BH-/HH-	0.025/ 0.01	40/100
XC	IRIS Temporary Network	IRIS	IRIS	V	SEED	BH-	0.1	10
XF	IRIS Temporary Network	IRIS	IRIS	V	SEED	BH-	0.025	40
XK	IRIS Temporary Network	IRIS	IRIS	V	SEED	BH-	0.05	20
XM	IRIS Temporary Network	IRIS	IRIS	V	SEED	BH-	0.05	20

Network code	Network name	Network owner/manager	Data source	Data type	Data format	Instrument type	Time interval (sec)	Sampling frequency (Hz)
XO	IRIS Temporary Network	IRIS	IRIS	V	SEED	BH-	0.025	40
XR	IRIS Temporary Network	IRIS	IRIS	V	SEED	BH-	0.025	40
Y8	IRIS Temporary Network	IRIS	IRIS	V	SEED	HH-	0.01	100
YC	IRIS Temporary Network	IRIS	IRIS	V	SEED	HH-	0.01	100
Z3	IRIS Temporary Network	IRIS	IRIS	V	SEED	BHZ, HH-	0.05, 0.01	20, 100
Z9	Temporary Network	IRIS	IRIS	V	SEED	BH-	0.02	50

Table 5–3 Earthquakes in NGA-East with M estimated using relationships from NUREG-2115.

EQID	Earthquake	Date	M
27	St Teresa, MX	4/6/2004	4.31
28	La Baie, QC	5/4/2004	2.87
36	Hawkesbury, ON	2/26/2006	2.59
44	Cobourg, ON	7/19/2007	2.80
45	Baie Saint Paul, QC	1/3/2008	2.77
50	Buckingham, QC	6/11/2008	2.97
52	Pine Forest, SC	12/16/2008	3.16
53	Rosehill, SC	1/29/2009	2.77
54	Palmetto, SC	5/6/2009	2.18
55	Constance Bay, ON	5/8/2009	2.57
59	Lebanon, IL	5/21/2010	2.62
61	St. Flavien, QC	7/23/2010	3.51
68	Concord, NH	9/26/2010	3.05
69	Nahanni, NWT foreshock	11/9/1985	4.4
71	Nahanni, NWT aftershock	12/23/1985	5.1
83	Val-des-Bois, QC aftershock	6/24/2010	2.57
84	Val-des-Bois, QC aftershock	7/22/2010	2.37

Table 5–4 Summary of assumed probabilistic distributions of parameters used in finite-fault simulation procedure.

Parameter	Distribution/reference
Rupture area	Somerville et al. (2001, 2014)
Aspect ratio	Chiou and Youngs (2008)
Hypocenter location down dip	Chiou and Youngs (2008)
Hypocenter location along strike	Chiou and Youngs (2008)

Table 5–5 Sources of geophysical data for CENA stations.

Measurement source	Count
Beresnev and Atkinson (1997)	9
Paul Mayne (personal communication, 2011)	9
Dames and Moore (1974)	1
EPRI (2013)	33
Ghofrani and Atkinson (2014)	1
GSC CNSN online station book (2014)	1
Herrmann and Crossey (2008)	1
Jaume (2006)	5
Lin and Adams (2010)	1
Karen Assatourians (personal communication, 2011)	2
Chris Cramer (personal communication, 2013)	2
Odum et al. (2010); USGS NGMD	5
Read et al (2008)	2
Saint Louis University Earthquake Center (2014)	4
USGS ANSS (2013)	2
Kayen et al. (2015)	1
Williams et al. (2003)	5

Table 5–6 Relative proxy weights by region and applied weights for estimation of V_{S30} when all estimates are available (Code 4).

Proxy	μ_{lnV}	σ_{lnV}	Count	Relative weight	Weight
Geology	0.079	0.479	34	4.24	0.27
Terrain	0.375	0.567	34	2.16	0.14
Slope	0.300	0.593	34	2.27	0.15
Hybrid	0.297	0.571	34	2.41	0.15
P-wave	-0.109	0.456	34	4.54	0.29

Table 5–7 Relative proxy weights by region and applied weights for estimation of V_{S30} when no estimate by *P*-wave proxy is available (Code 5).

Proxy	μ_{lnV}	σ_{lnV}	Count	Relative weight
Geology	0.068	0.508	84	3.80
Terrain	0.373	0.612	84	1.94
Slope	0.254	0.613	84	2.27
Hybrid	0.245	0.592	84	2.43

Table 5–8 Simulations methodologies considered for evaluation.

Method name	Short-hand identifier(s)	Latest reference
Composite Source Model	CSM	Anderson (2015)
EXSIM	EXSIM, EX	Atkinson and Assatourians (2015)
Graves and Pitarka	GP	Graves and Pitarka (2015)
San Diego State University	SDSU, SD	Olsen and Takedatsu (2015)
U.C. Santa Barbara	UCSB, SB	Crempien and Archuleta (2015)

Table 5–9 Summary of validation events (Part A. Comparison to historical events).

Region	Event name	Year	Mw	Number of records available < 200 km (* <1000 km)	Note on selection	# Selected number of records
WUS	Loma Prieta	1989		59		40
WUS	Northridge	1994	6.73	124	All stations within 10 km selected	39
WUS	Landers	1992	7.22	69	All stations within 100 km selected	40
WUS	Whittier Narrows	1987	5.89	95	Truncate stations at 40 km	39
WUS	North Palm Springs	1986	6.12	32		31
JAPAN	Tottori	2000	6.59	171		40
JAPAN	Niigata	2004	6.65	246		40
WUS	Alum Rock		5.45	40		40
WUS	Chino Hills		5.39	40		40
CENA	Saguenay	1988	5.81	14*	All records selected, use only within 200 km	11
CENA	Riviere-du-Loup	2005	4.6	98*	All records selected, use only within 200 km	21
CENA	Mineral, VA	2011	5.68	94*	All records selected, use only within 300 km	10

Table 5–10 Summary of validation events (Part B. Comparison to GMMs estimates). All the simulations are complete for distances of 20 and 50 km, on the footwall side of the fault.

Magnitude	Mechanism (dip°)	Z_{TOR} (km)
5.5	Reverse (45)	6
6.2	Strike-slip (90)	4
6.6	Reverse (45)	3
6.6	Strike-slip (90)	0

Table 5–11 Summary of earthquake scenarios for NGA-East simulations, Set 1. All the events are for reverse fault mechanism, dipping at 45°.

Magnitude	Length (km)*	Width (km)*	Area (km^2)*	Z_{TOR} (km) considered
5.0	2.5 (2.55)	2.5 (2.58)	6.25 (6.46)	0, 5, 10
5.5	5 (5.08)	4 (4.02)	20 (20.4)	0, 5, 10
6.5	20 (20.2)	10 (10.1)	200 (204)	0, 5, 10
7.5	80 (80.2)	25 (25.4)	2000 (2041)	0, 5, 10
8.0	160 (159.8)	40 (40.4)	6400 (6456.5)	0, 5

* The first numbers listed were used to define the ruptures and the numbers in parenthesis are the actual numbers obtained from computations.

Table 5–12 Summary of simulations data processing to obtain PSA and FAS magnitude-scaling ratios.

Step #	Fourier amplitude spectra (FAS)	Pseudo-spectral acceleration (PSA)
1	Compute the FAS of each simulated time series using Dave Boore's package TSPP_v4.8 (Boore 2014).	Start with the RotD50 obtained from each pair of horizontal simulated time series, as computed in the BBP workflow. This is denoted PSA.
2	Average (geometric mean) of FAS at each station over 16 realizations; FAS_i	Average (geometric mean) of PSA at each station over 16 realizations; PSA_i
3	Average (geometric mean) of FASi at each R_{rup} band (j); FAS_j Note that there are 12 R_{RUP} distances: 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, and 100 km	Average (geometric mean) of PSAi at each R_{RUP} band (j); PSA_j Note that there are 12 R_{RUP} distances: 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100 km
4	Smooth FAS_j using Konno and Ohmachi (1998)	No smoothing is applied.
5	Repeat for Magnitudes 5.0, 5.5, 6.5, 7.5, and 8.0	Repeat for Magnitudes 5.0, 5.5, 6.5, 7.5, and 8.0
6	Calculate FAS Ratios relative to M 5.0 scenario	Calculate PSA Ratios relative to **M** 5.0 scenario
7	Repeat steps 2-6 for each Z_{TOR} case (Z_{TOR} = 0, 5, and 10 km) and simulation method.	Repeat steps 2-6 for each Z_{TOR} case (Z_{TOR} = 0, 5, and 10 km) and simulation method.

Figure 5–1 Central and Eastern North America earthquakes selected for inclusion in the NGA-East ground-motion database. The 1929 Grand Banks and 1985 Nahanni earthquakes are off this map and hence are not shown.

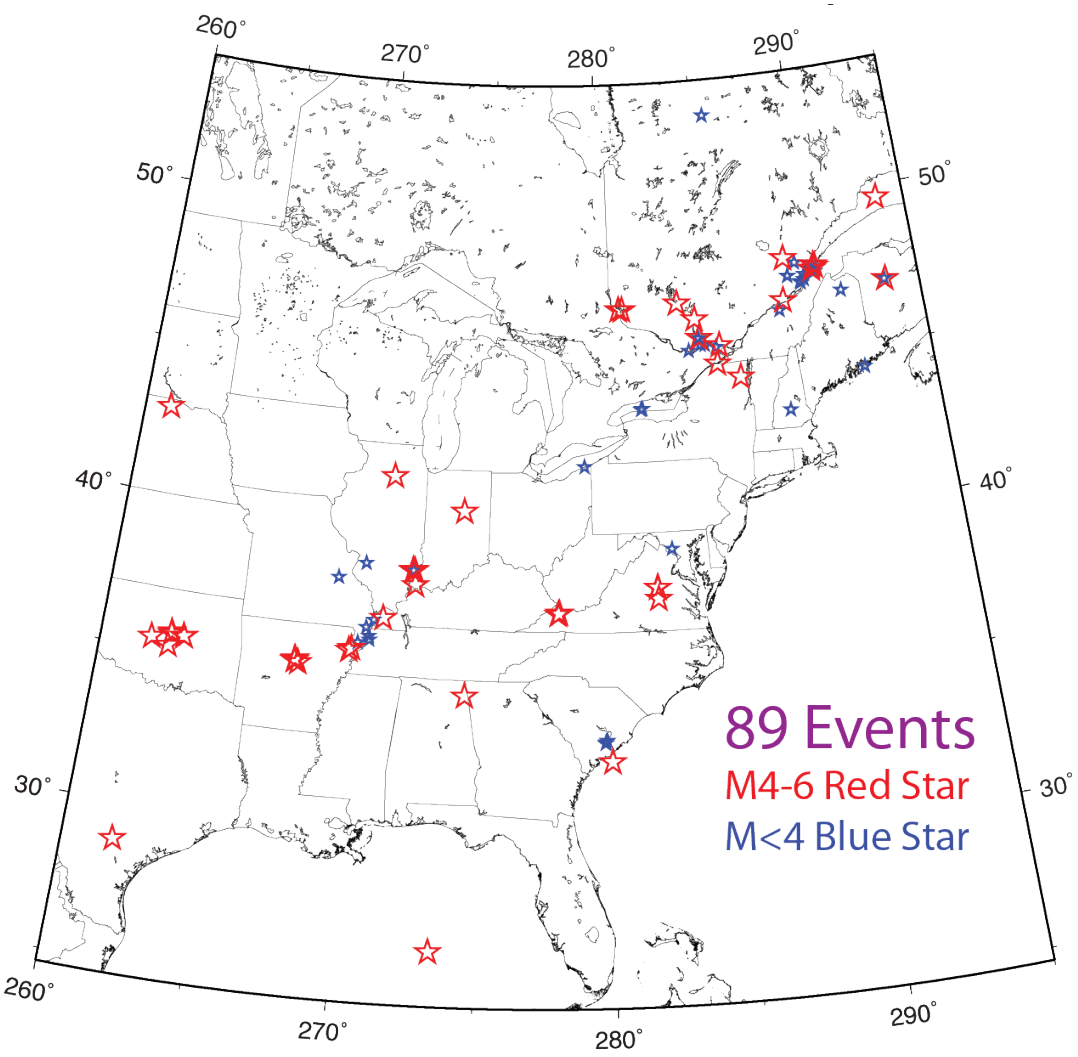


Figure 5–2 Seismograph and accelerograph stations with at least one record in the NGA-East ground-motion database. Some station coverage extends beyond the borders of this map.

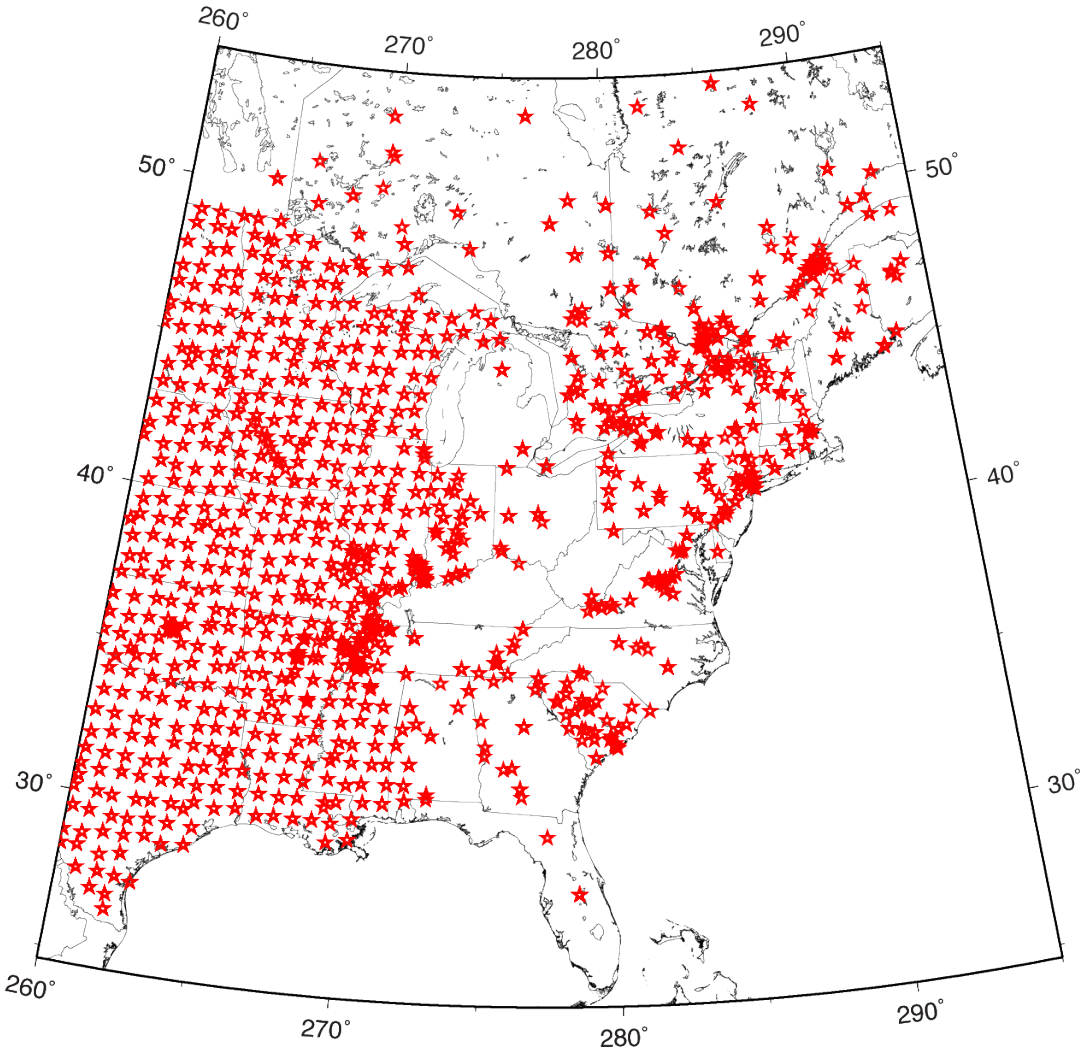


Figure 5–3 Magnitude versus distance coverage of recordings in the NGA-East ground-motion database.

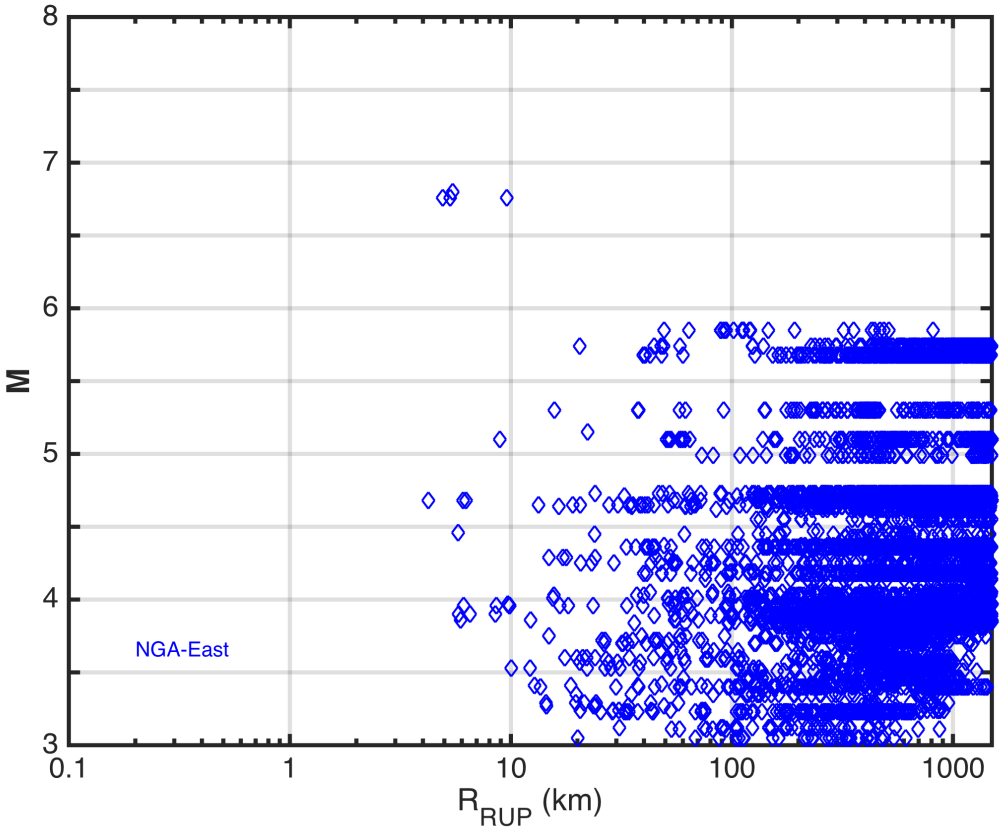


Figure 5–4 Schematic drawing of the six time windows used.

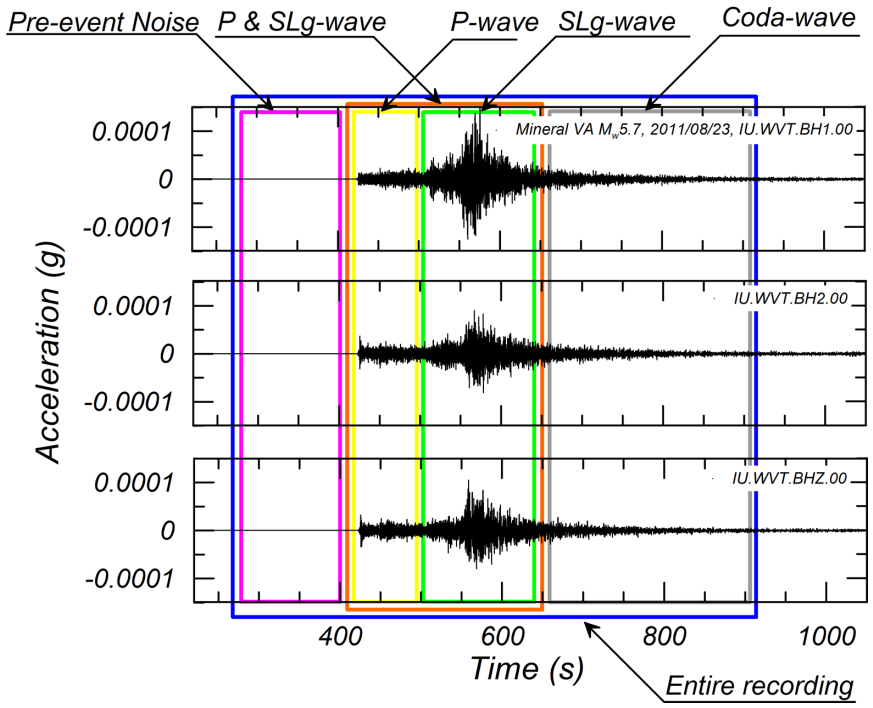


Figure 5–5 Example Fourier amplitude spectra from acceleration time series for different windows.

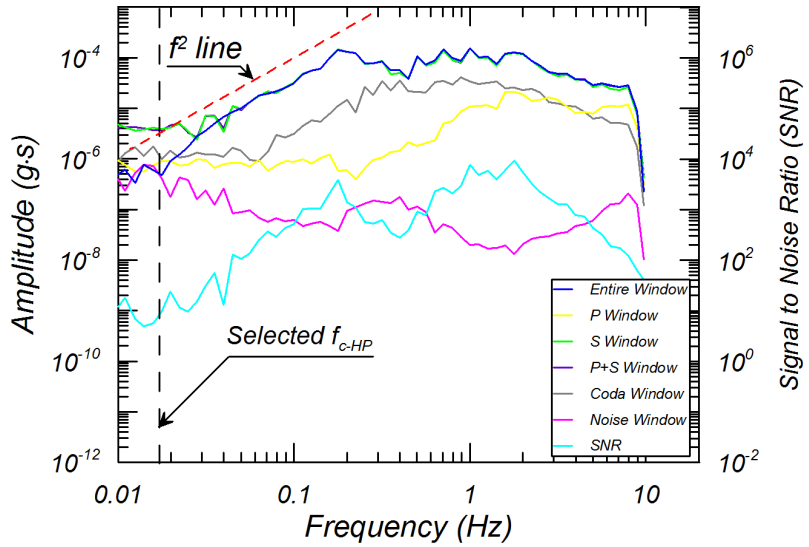


Figure 5–6 Time series of acausally filtered pad-stripped record. Shown are (top) acceleration (middle) velocity, and (bottom) displacement time series. The initial velocity and displacement are assumed to be zero in the integration.

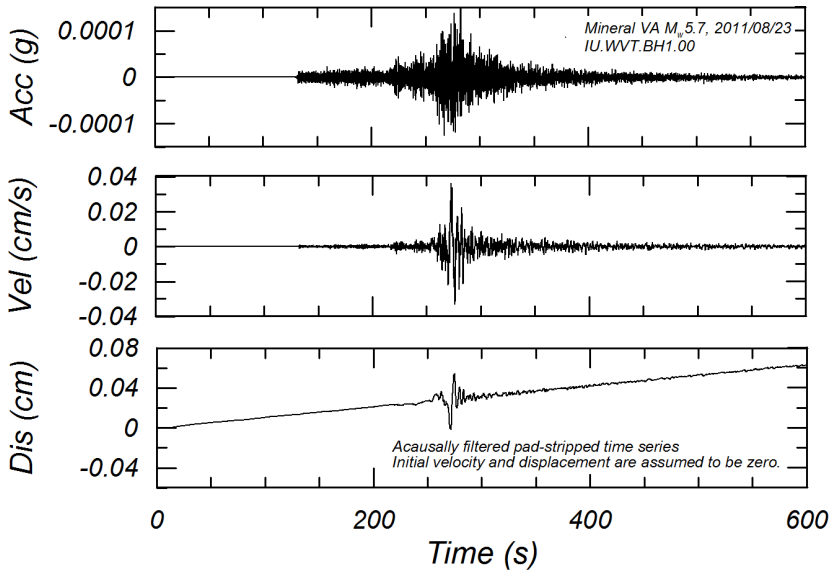


Figure 5–7 Acausally filtered time series with baseline correction. Shown are (top) acceleration, (middle) velocity, and (bottom) displacement time series.

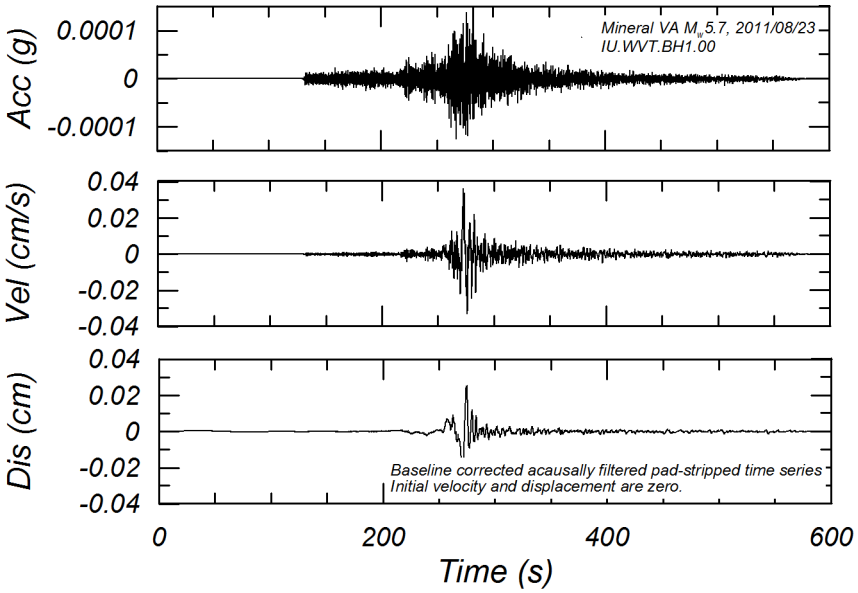


Figure 5–8 Effect of baseline correction on PSA for acausally filtered time series.

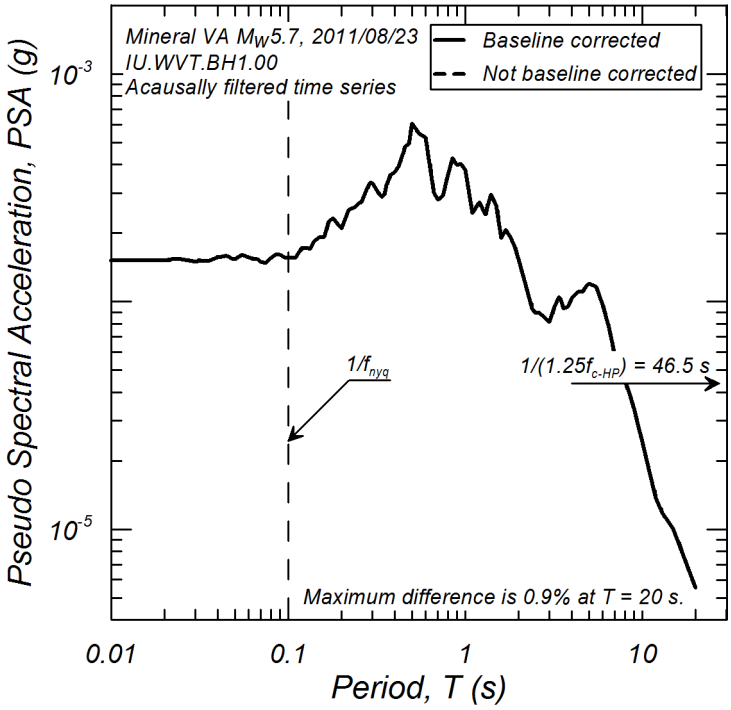


Figure 5–9 FAS affected by microseisms.

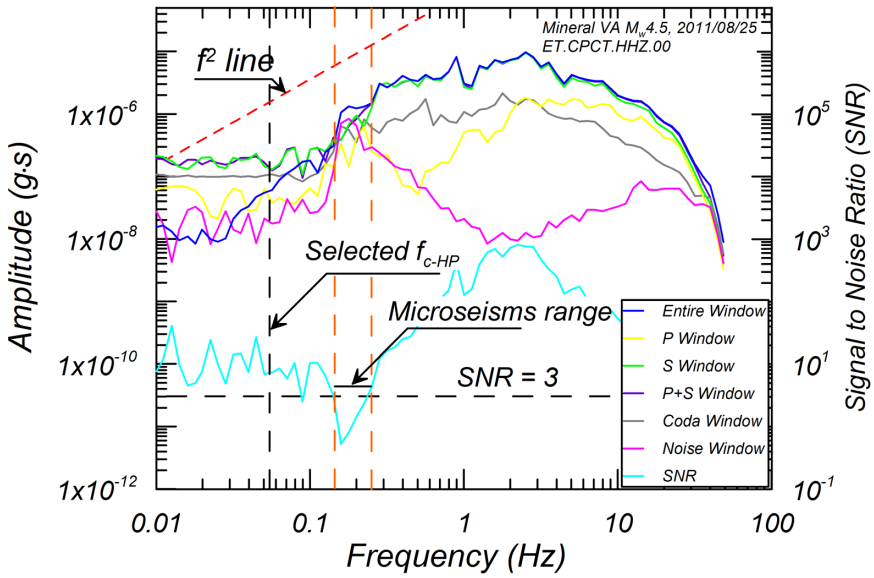


Figure 5–10 Acceleration, velocity, and displacement time series.

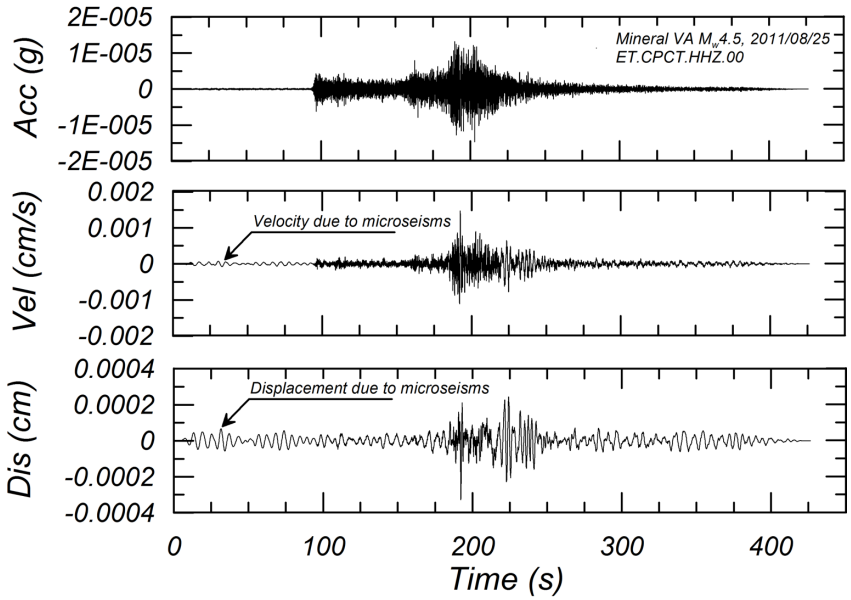


Figure 5–11 Schematic representation of strike, dip, rake, depth to top of rupture (Z_{TOR}), down dip width (W), and length (L).

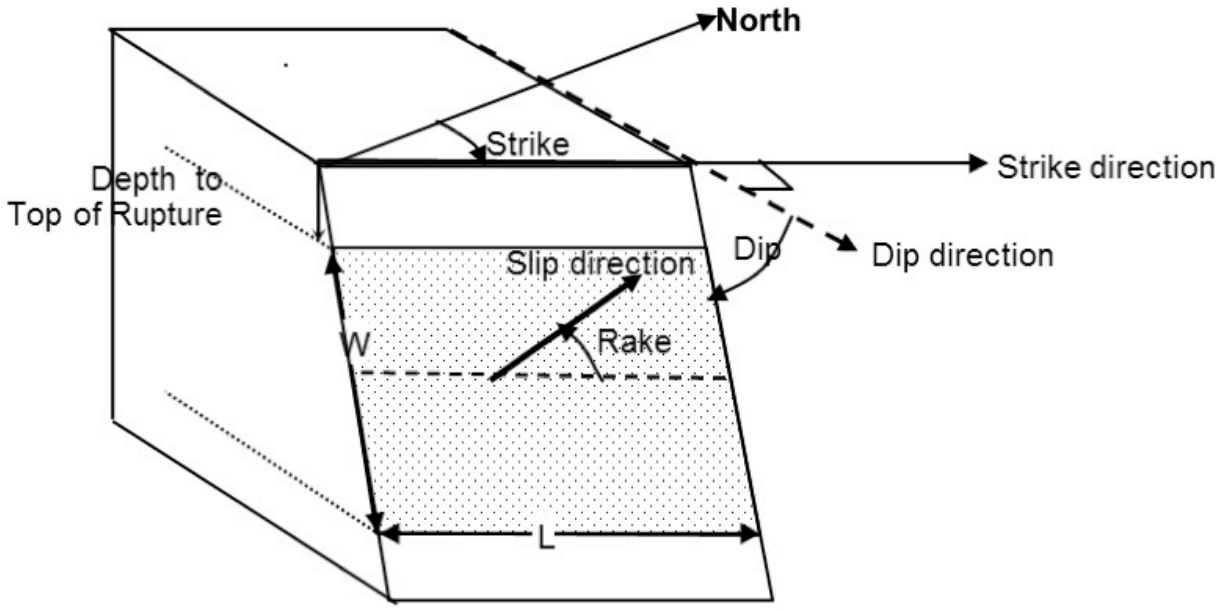


Figure 5–12 Histogram of V_{S30} measurements in the NGA-East station database.

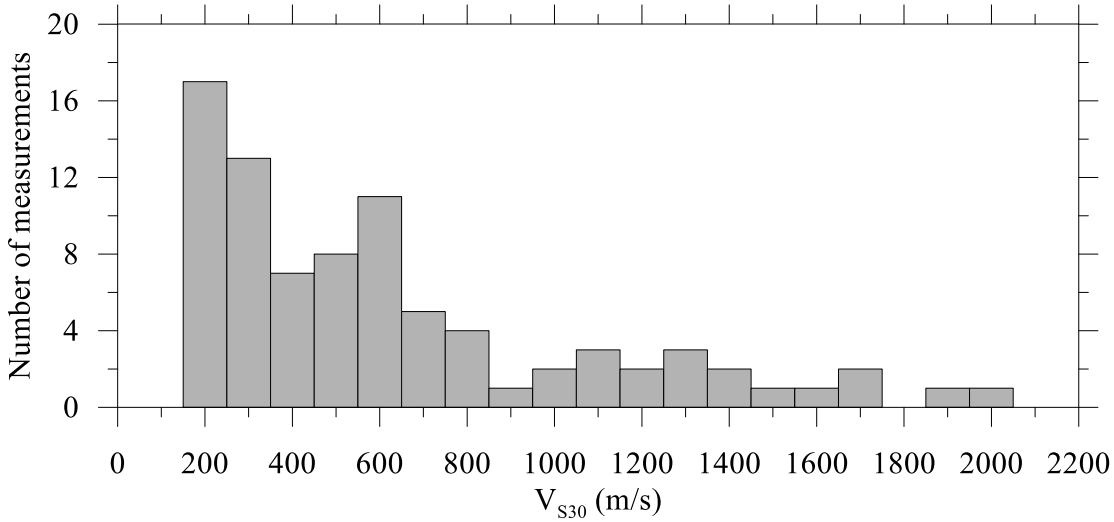


Figure 5–13 Residual of proxy-based estimations at sites where available for (a) geology (Kottke et al. 2012), (b) terrain (Yong et al. 2012), (c) slope (Wald and Allen 2007), (d) P-wave seismogram with multiple estimates (Kim et al. 2016), and (e) hybrid slope-geology (Thompson and Silva 2013).

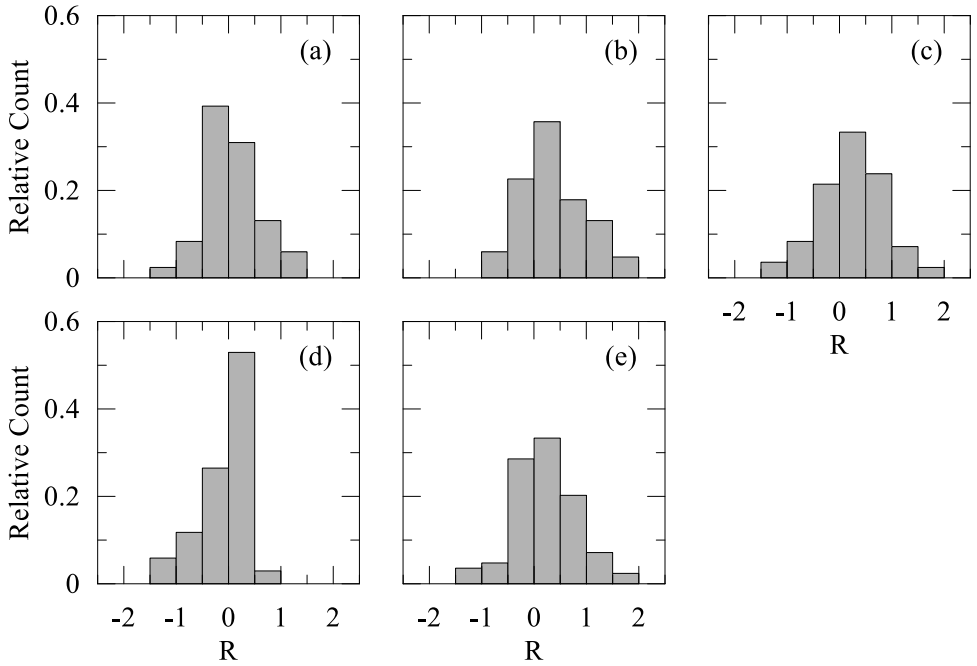


Figure 5–14 Average residual and standard deviation on the residual for geology (Kottke et al. 2012), terrain (Yong et al. 2012), slope (Wald and Allen 2007), hybrid (Thompson and Silva 2013), and P-wave (Kim et al. 2016), of V_{S30} estimates by proxy at recording stations. Circles represent the residual on the V_{S30} at the 34 recording stations that have estimates of V_{S30} by all proxy methods. Triangles represent the residual on the V_{S30} over all stations where the proxy-based estimate is available.

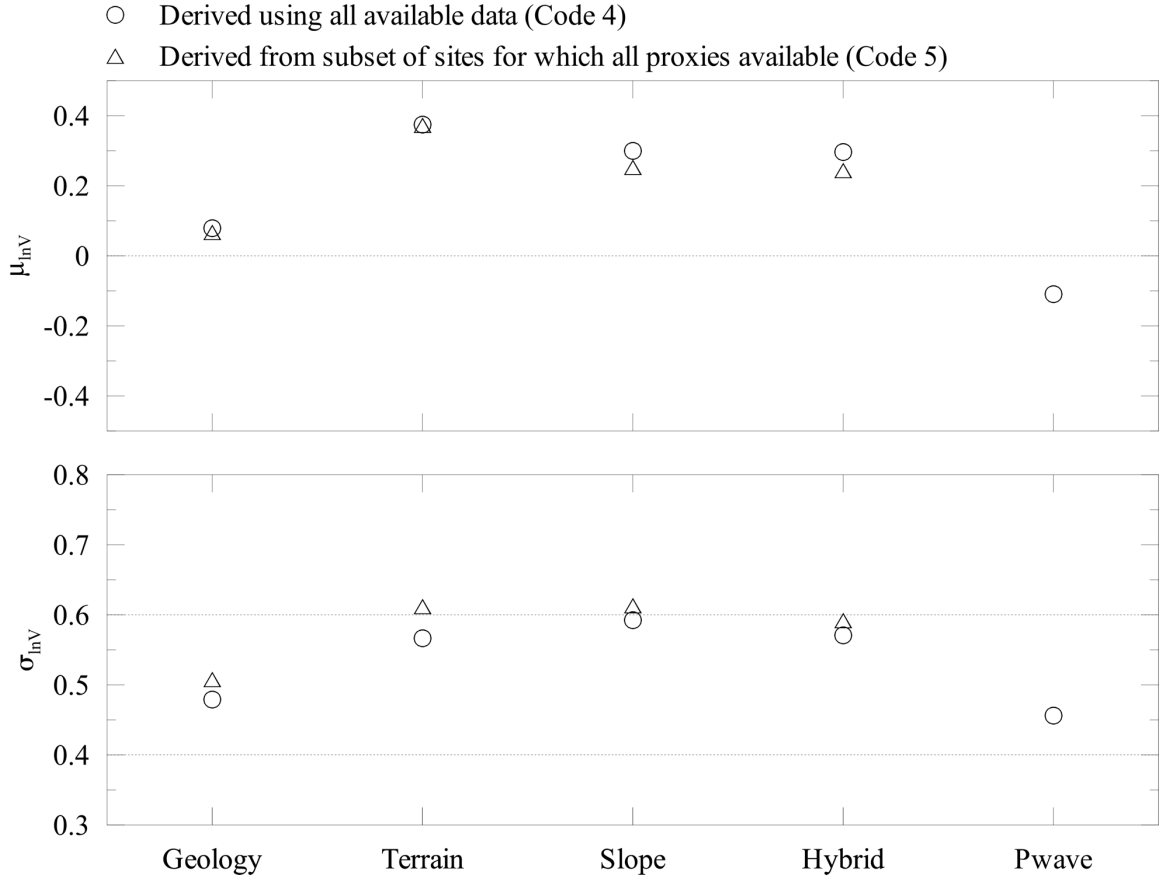


Figure 5–15 Distribution of code assignment for recommendation of V_{S30}.

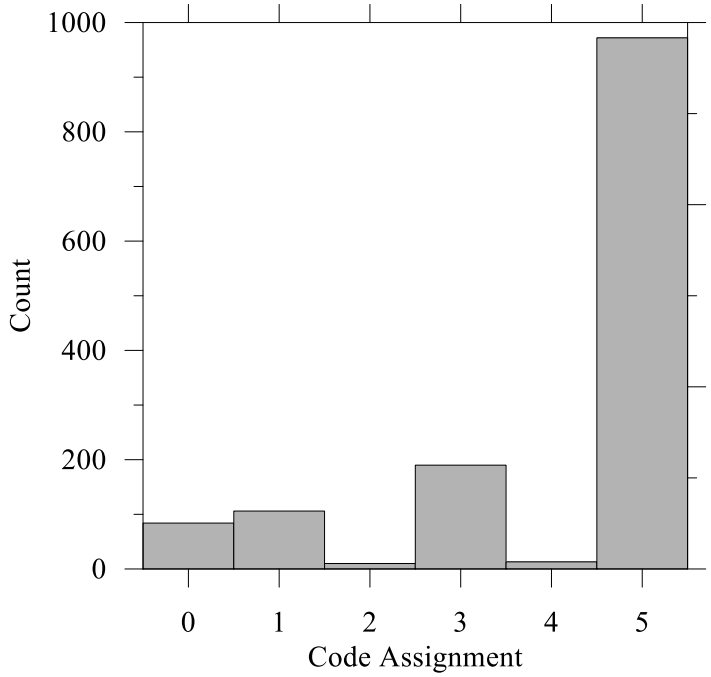


Figure 5–16 Recommended V_{S30} and σ_{lnV} by from Codes 0–5 assignment at all stations.

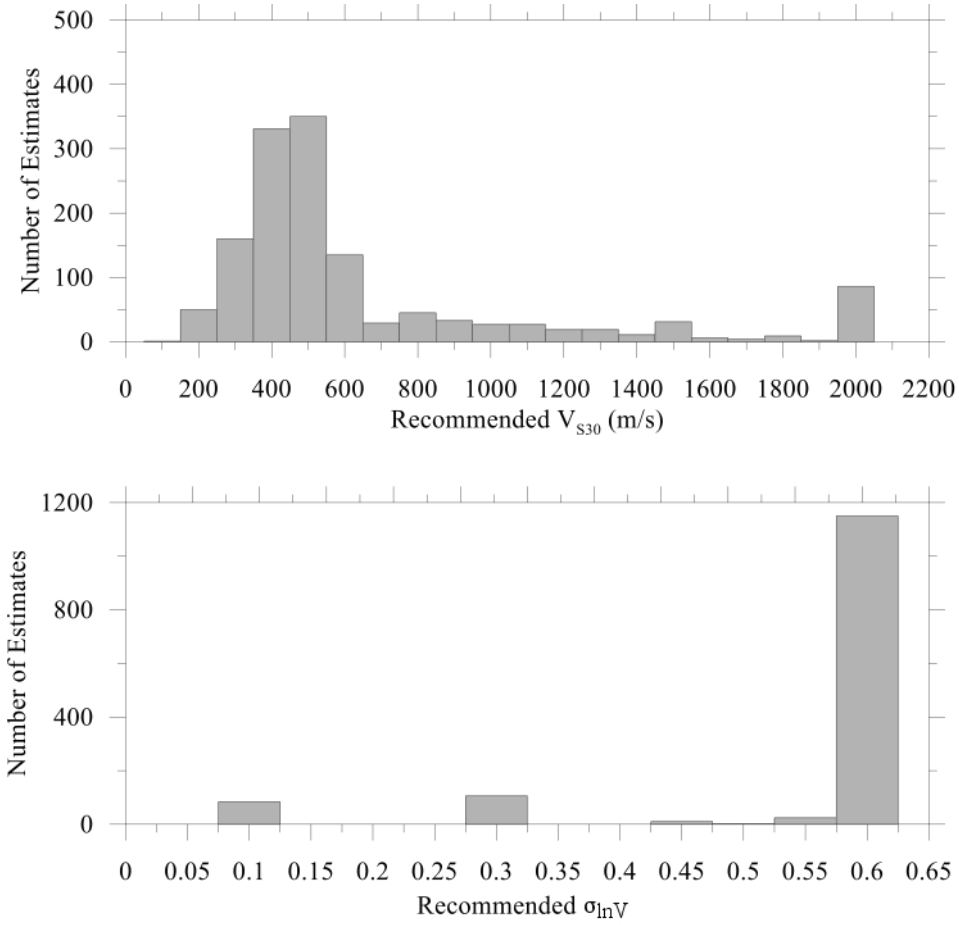


Figure 5–17 Summary of validation events magnitude (M) and station closest rupture distance (R_{RUP}). Part A events are color-coded by tectonic regions.

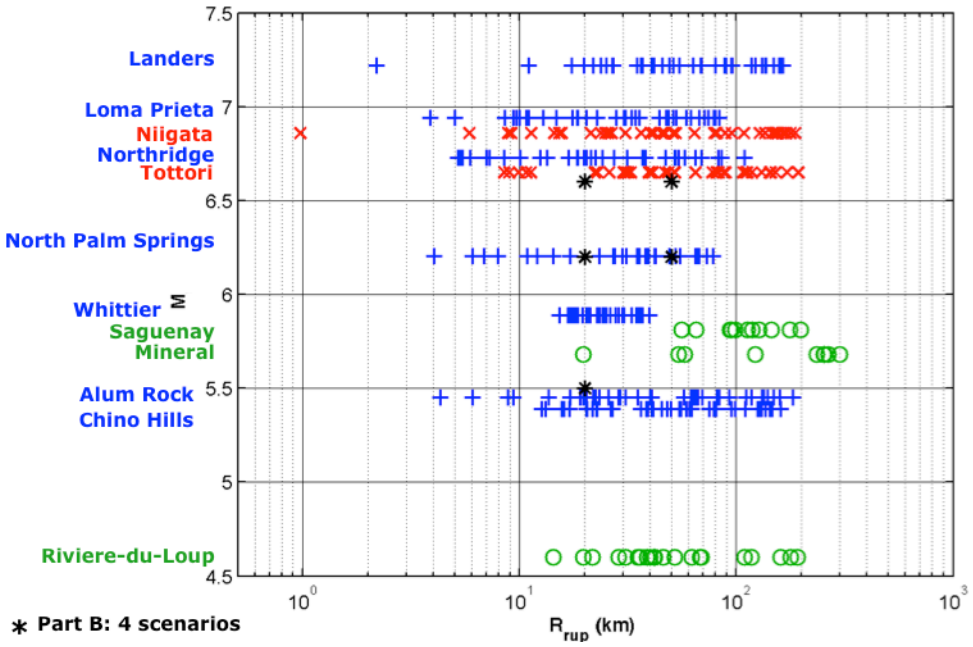


Figure 5–18 Sample station layouts for finite-fault simulations. Top is for M = 6.5, Z_{TOR} = 5 km, bottom is for M = 8.0, Z_{TOR} = 0 km.

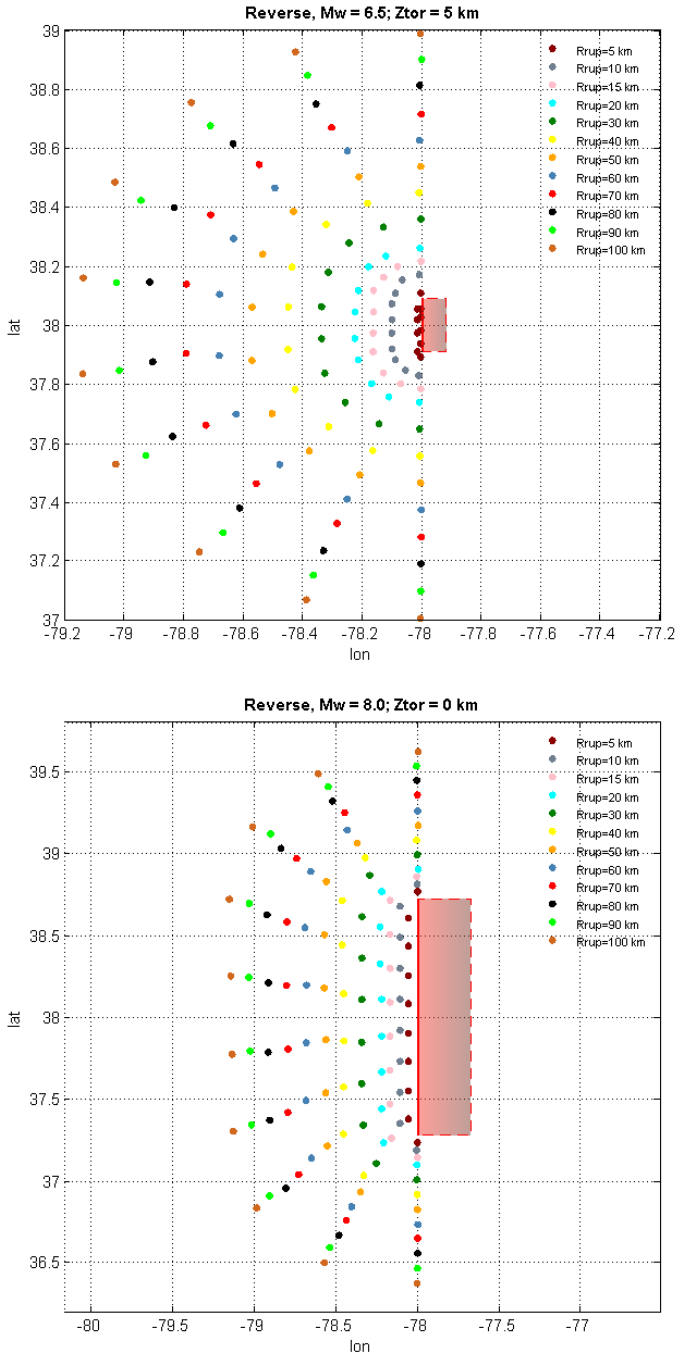


Figure 5–19 Example of M-scaling ratio computations from simulations. This example is for EXSIM simulations, 20 km away from a buried fault (Z_{TOR} = 5 km). Top frames show output FAS (left) and PSA (right). Each thin grey line corresponds to the ground motions from a single station at the 20 km distance. The thick green and blue lines show the smoothed mean ground motions for each M = 5.0, 5.5, 6.5, 7.5, and 8.0 scenarios. Bottom frames show computed ratios for FAS (left) and PSA (right), relative to the M = 5.0 scenario.

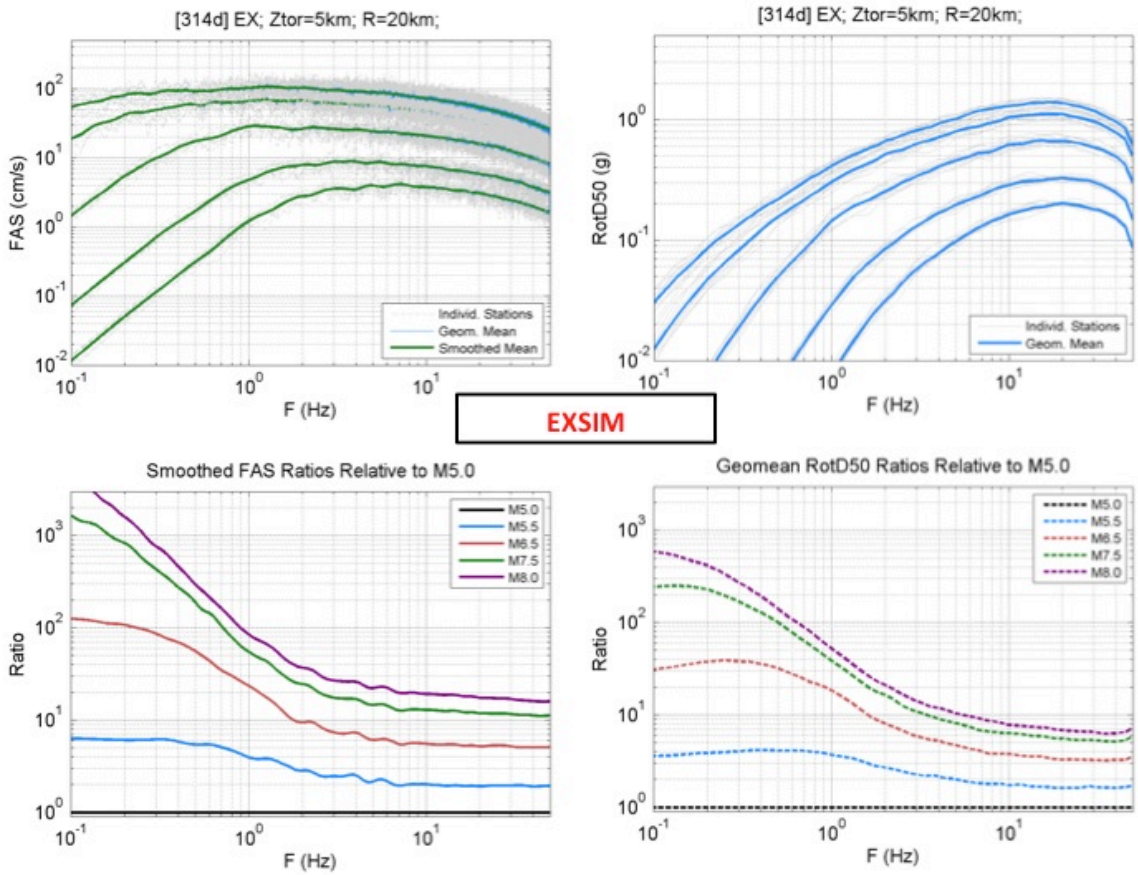


Figure 5–20 Magnitude versus distance coverage of recordings in the NGA-West2 ground-motion database.

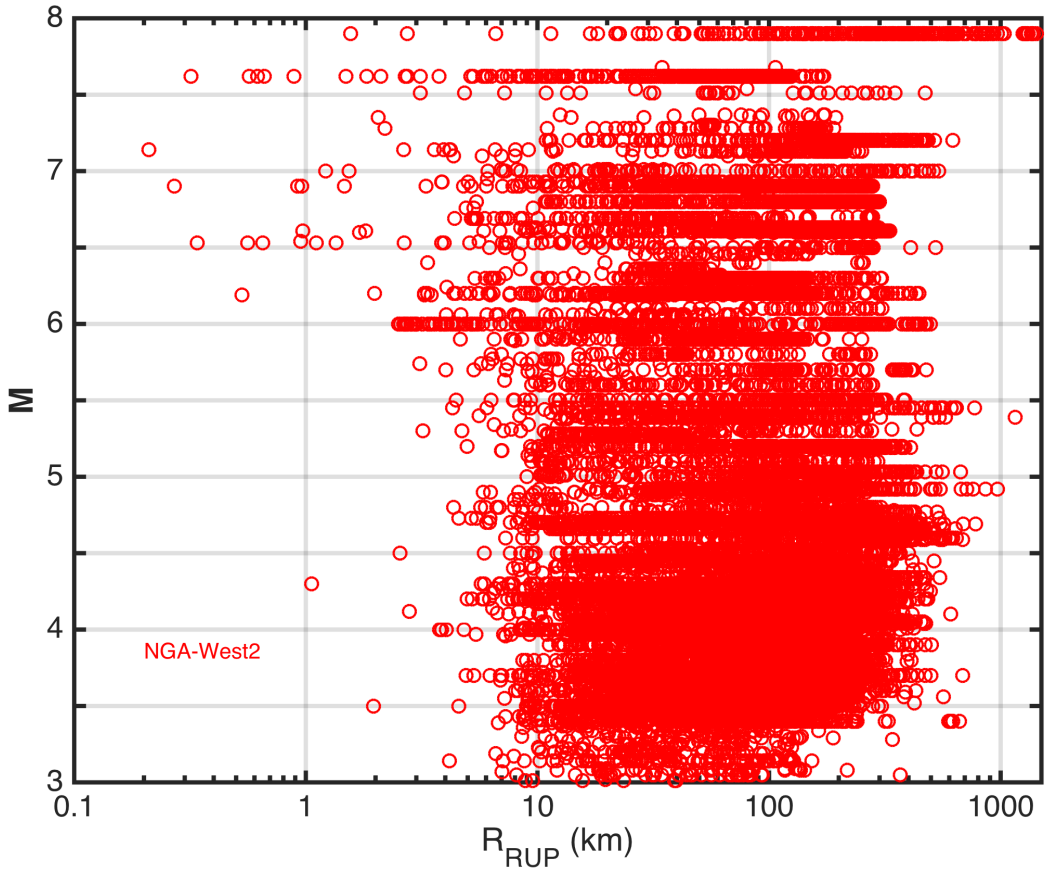


Figure 5–21 Magnitude versus distance coverage of recordings in the NGA-East and NGA-West2 ground-motion databases.

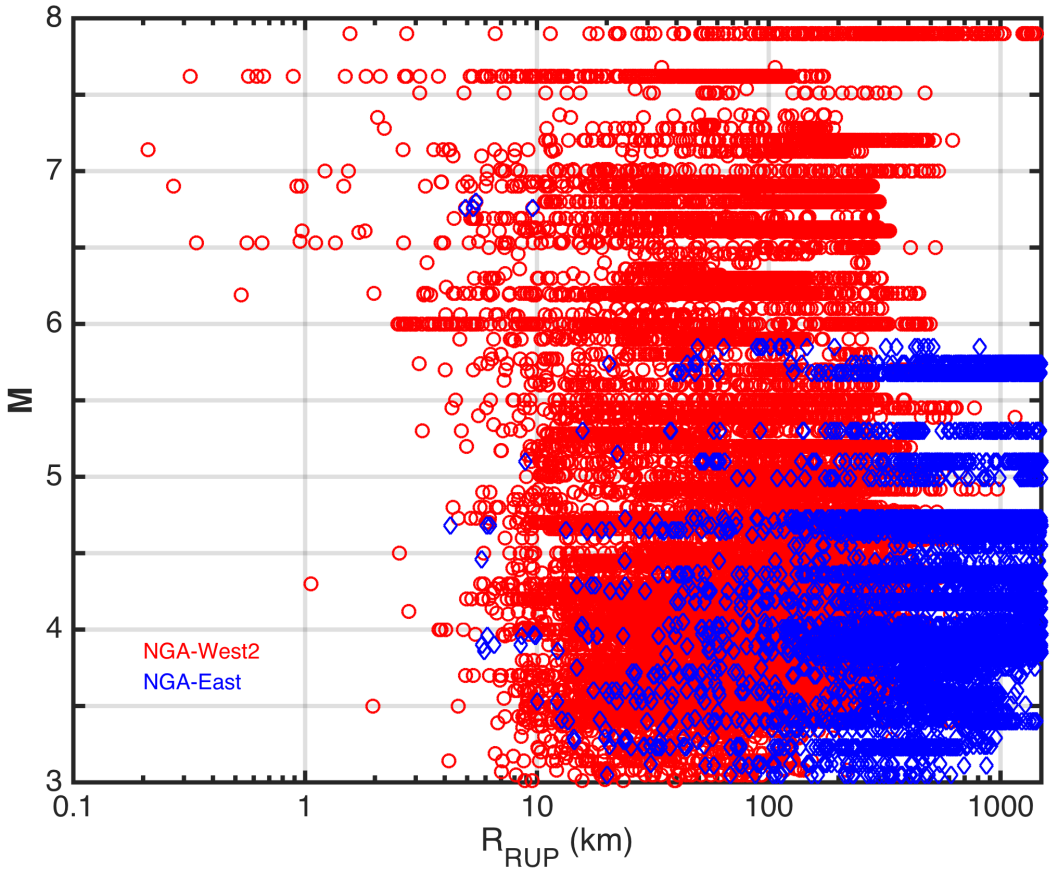


Figure 5–22 Metadata for Japan single-station sigma study. M_w and R_{RUP} obtained from the F-net catalog (blue crosses) and previously published finite fault source models (red circles) [see Dawood (2014)].

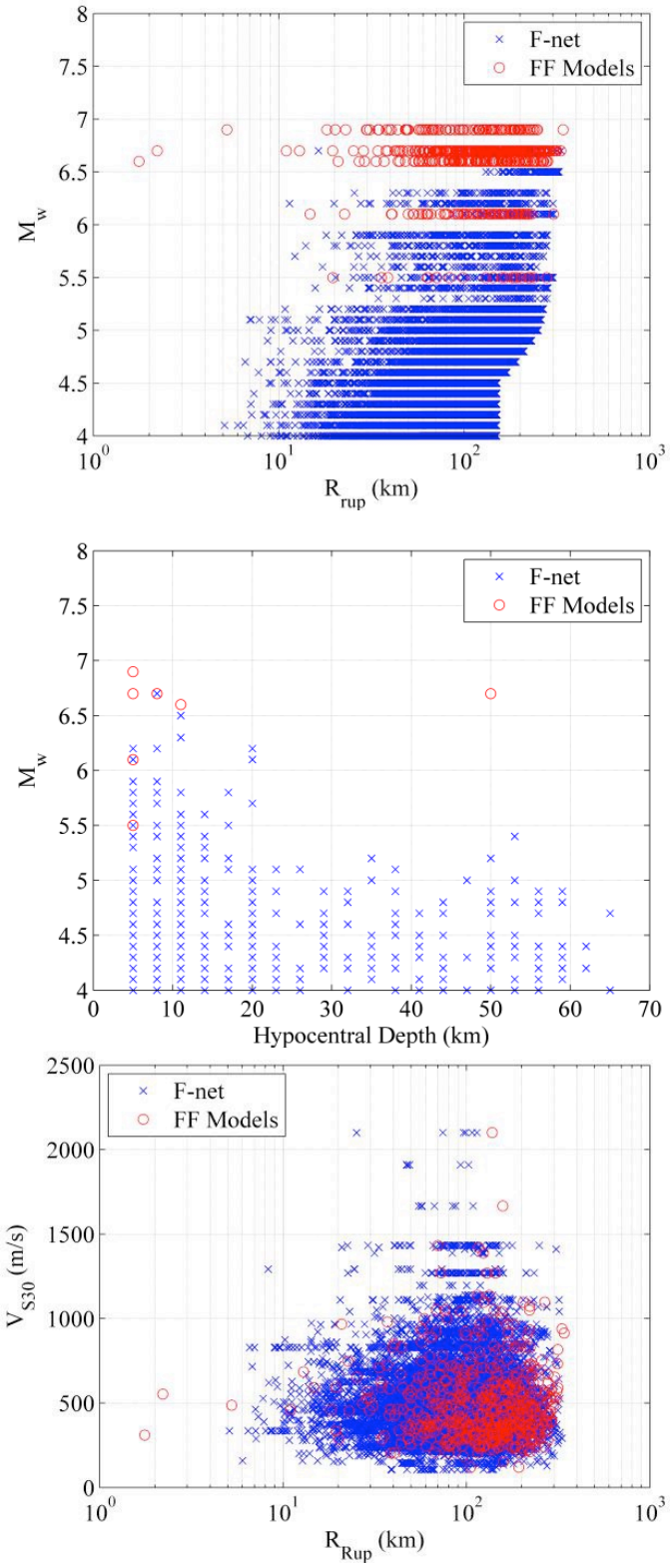


Figure 5–23 Histogram of the number of stations in the V_{S30} bins for the Japanese ground-motion dataset.

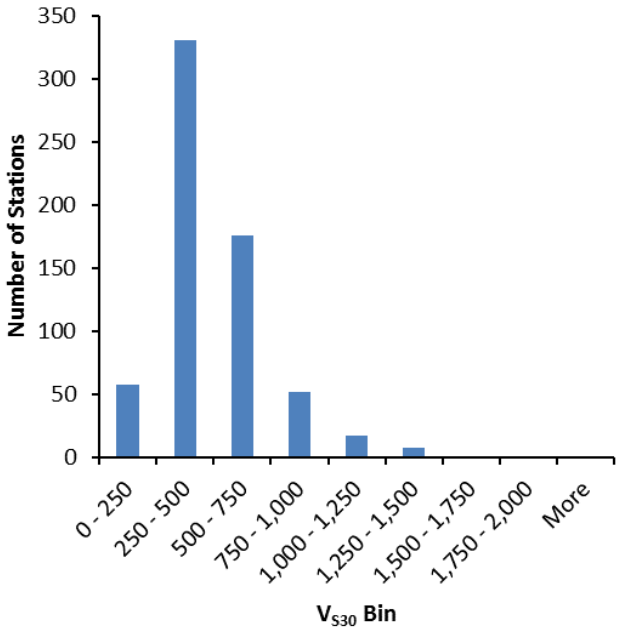


Figure 5–24 **Histogram of the number of stations in the $V_{S,Zhole}$ bins for the Japanese ground-motion dataset.**

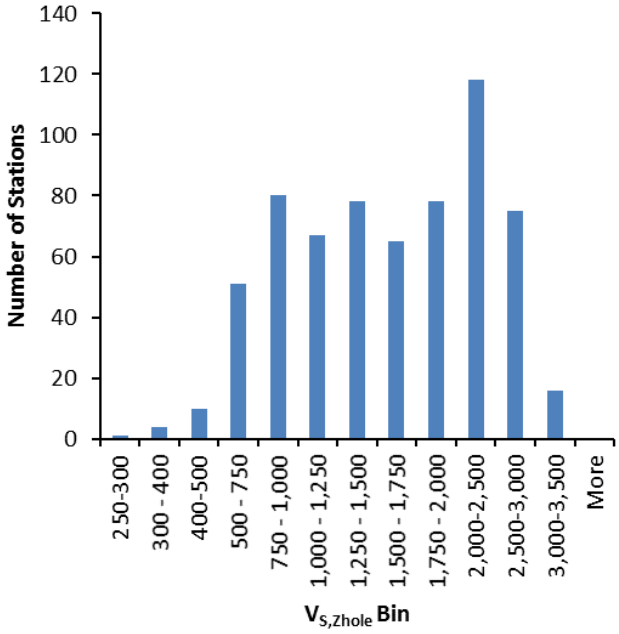
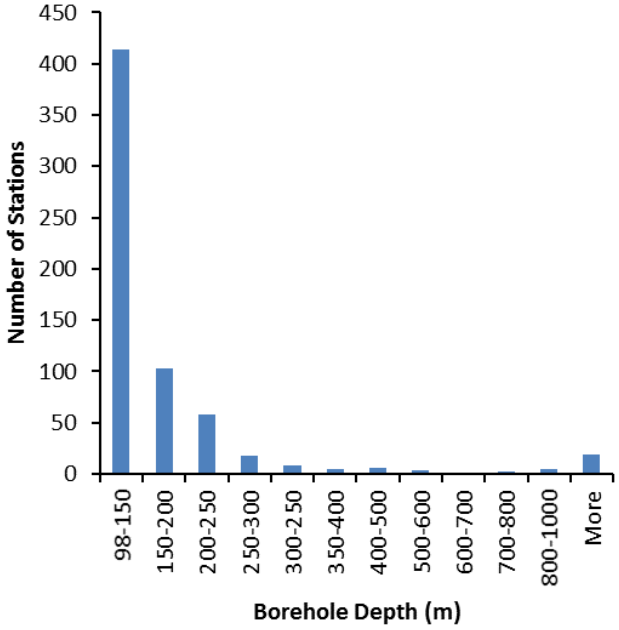


Figure 5–25 **Histogram of the number of stations in the borehole depth bins for the Japanese ground-motion dataset.**



6. NGA-East GMC Approach Overview

This chapter summarizes the conceptual approach behind the ground-motion characterization (GMC) model building. The detailed implementation is described in Chapters 7–10. A large portion of the current chapter (Sections 6.1–6.3) focuses on the selected approaches for the quantification of epistemic uncertainty in median ground motions. The last section (6.4) summarizes the modeling of aleatory variability. Generic notations for magnitude (**M**) and distance (R) are used throughout this chapter.

6.1 Importance of Epistemic Uncertainty Quantification

Epistemic uncertainty in the input parameters of probabilistic seismic hazard analysis (PSHA) studies is typically incorporated through the use of logic trees. In this format, epistemic uncertainty in the GMC is represented by a set of alternative ground-motion models (GMMs) with assigned weights. The epistemic uncertainty in median ground motions, which is usually the largest contributor to the epistemic uncertainty in the hazard, can be characterized separately from the epistemic uncertainty in the aleatory variability or the two can be linked. The alternative models for median motions and aleatory variability are assigned weights by analysts to represent the center, body, and range of the technically defensible interpretations (CBR of TDI). The uncertainty in the median ground motions and aleatory variability is propagated in the analysis by computing the hazard for each alternative GMM and assigning the associated GMM weight to the resulting hazard curve. The suite of weighted hazard curves is then used to compute the mean hazard and hazard fractiles. Past experience has shown that the epistemic uncertainty distributions for seismic hazard are skewed such that the location of the mean hazard within the distribution is sensitive to the shape of the distribution. The proper definition and quantification of epistemic uncertainty is therefore of critical importance to hazard studies, and this recognition lead to the development of the SSHAC guidelines (Budnitz et al. 1997; NRC 2012). The SSHAC process is intended to make the assessments of epistemic uncertainty transparent and defensible.

6.2 Motivation and Brief Summary of Selected Previous Approaches

The epistemic uncertainty is intended to represent the range of a continuous distribution of median ground motions. This uncertainty is often captured by the combination of various GMMs into a logic tree. The mathematical formulation behind the logic-tree representation uses the axioms of probability such that, at a node of the tree, each branch must contribute information that aims to, at least conceptually, make the set of branches at that node mutually exclusive and collectively exhaustive (MECE). In other words, *the weights assigned to the branches are treated as probabilities*. When the underlying uncertainty is quantified by a distribution, then it can be represented discretely to varying degrees of accuracy [e.g., Miller and Rice, (1983)] such that the individual discrete alternatives are mutually exclusive; they are collectively exhaustive to the extent that the analyst wishes to capture the moments and fractiles of the underlying distribution.

However, when the epistemic uncertainty is represented by a set of alternative models for a process with weights assigned often on the basis of relative merit, adherence to MECE may become problematic. As discussed in Bommer and Scherbaum (2008), it is often the case that the alternative models have been derived from the same or overlapping datasets using similar conceptual models such that there may be significant model redundancy. They argued that in these cases the epistemic uncertainty may not be fully captured, and the treatment of the alternative models as MECE in calculating the hazard distribution may not be appropriate. In other words, the mutual exclusivity (ME) condition is violated when two or more of the GMMs may be partial clones of each other because they use the same data or similar assumptions (in contrast to other GMMs in the logic tree, which are more distinct). This violation complicates the assignment of weights and makes equal weights difficult to justify. The complete exhaustiveness (CE) condition is violated if the logic tree includes a finite number of GMPEs because intermediate values of ground-motion amplitudes are artificially excluded. In practice, these CE violations may not create a serious problem as long as one avoids the following two pitfalls: (1) a discretization that is too coarse and does not provide a representation of the CBR of the TDI (and may lead to inaccurate mean or fractile values of the hazard); and (2) unjustified exclusion of tail values (e.g., values higher than those predicted by the highest GMM even though they belong to the CBR of the TDI). Moreover, there are other difficulties in the weighted GMM approach, such as potential lack of suitability of some or all of the models for the magnitude/distance/region of interest. These issues have resulted in a general understanding that in most cases the multiple-GMM approach will not meet the objective of describing the CBR of the TDI of the data (Atkinson et al. 2014).

The following sub-sections briefly summarize the strategies that were used in selected PSHA and GMC projects for the quantification of epistemic uncertainty in GMMs, with each sub-section corresponding to a specific project. All the projects referenced in this section were conducted as SSHAC Level 3 or 4. In all these examples, the epistemic uncertainty is quantified through logic trees, but the composition of each branch and the meaning of their associated weight evolved over time. We present a very short summary of each approach and then comment on the degree to which MECE was achieved. This section provides the background and motivation for the proposed approach for the NGA-East project, which is introduced in Section 6.3.

6.2.1 SSHAC Trial Implementation Project (TIP) (1995–2002)

The TIP project was completed in the mid-late 1990s and documented in 2002 (Savy et al. 2002). The project was launched to test the implementation of the SSHAC Level 4 guidelines for SSC and GMC projects, which were developed around the same time and published in 1997 (Budnitz et al. 1997).

The approach was based on expert elicitation using a panel of five experts, who were asked to assess the median ground motion and its uncertainty for 132 scenarios defined by magnitude, distance, and style of faulting for five frequencies. In addition, the experts were asked to assess the aleatory variability in ground motions for each scenario and its epistemic uncertainty, and then to provide distributions for these parameters defined by a median and standard deviation for each scenario. The resulting assessments were fit by algebraic expressions for the median

ground motions and its epistemic uncertainty, along with algebraic expressions for the aleatory variability and its uncertainty for implementation in the PSHA calculations. The uncertainty distributions for median ground motions and aleatory variability were then represented by discrete three-point approximations in the PSHA calculations. Calculations were performed using each expert's assessment, and a composite model was derived from combining the assessments. These numerical estimates were scalars that corresponded to a continuous distribution. Three points (5^{th}, 50^{th}, and 95^{th} percentiles) were used to generate a simple, coarse MECE representation of a continuous distribution.

6.2.2 Yucca Mountain PSHA (1998)

This study was part of the analyses conducted for the Yucca Mountain (YM) nuclear waste repository project (CRWMS M&O 1998). The PSHA conducted through expert elicitation was performed followed a SSHAC Level 4 process.

The approach followed that utilized by Savy et al. (2002) using a panel of seven experts. The experts were asked to assess the median ground motion and its uncertainty for about 60 scenarios defined by **M**, R, and style of faulting, and nine frequencies for horizontal and vertical ground motions. In addition, the experts were asked to assess the aleatory variability in ground motions for each scenario along with its epistemic uncertainty, and then to provide distributions (instead of a three-point estimate) defined by a median and standard deviation (σ) for each scenario. The resulting assessments were fit by algebraic expressions for the median ground motions and its epistemic implementation, along with algebraic expressions for the aleatory variability and its uncertainty for implementation in the PSHA calculations. The joint uncertainty distribution for median motion and aleatory variability was represented by a 10-point discrete approximation in the PSHA calculations.

The distributions were allowed to be asymmetric so that σ^+ could be distinguished from σ^-. Such distributions could then be down-sampled to be used in a logic tree, with each branch remaining MECE.

The number of requested distributions (~540) was too large to be practical. In the end, the experts used algorithms to define weights based on the merit of alternative models (e.g., empirical GMMs and simulations), straying away from the objective of providing estimates of median logarithmic standard deviation for 540 scenarios. An additional limitation of such an approach is that each (**M**, R) pair is treated independently; therefore, the correlation in ground motions is not preserved.

6.2.3 EPRI Ground Motion Study (2004, updated in 2013)

The Electric Power Research Institute (EPRI) Ground Motion Study was conducted as a SSHAC Level 3 GMC for central and eastern North America (EPRI 2004). The project organized the available GMMs into clusters based on their underlying technical basis (e.g., point-source simulations, hybrid-empirical, finite-fault simulations, etc.). Each cluster was intended to represent a class of methodology, with the uncertainty in implementation/assumptions within the class being addressed within each cluster. Relative weights were assigned to each GMM within a cluster based on how consistent they were with the available empirical ground-motion data.

These weighted models were then used to compute a standard deviation of the log of median ground motions to represent the within-cluster epistemic uncertainty as a lognormal distribution. The within-cluster epistemic standard deviation was further expanded by evaluating how well the cluster GMMs accounted for the range of current interpretations of parameters such as median stress drop and regional Q. The final composite lognormal distribution for each cluster was then represented by a three-point discrete distribution in the form of median, 5th percentile, and 95th percentile GMMs. As such, the within-cluster weights can be considered to represent probabilities.

The final step was to assign relative weights to the four ground-motion cluster distributions. The cluster weights were largely assigned based on the merit of approaches—as assessed by a group of experts—in combination with consistency of the cluster medians with the limited data. The cluster weights did not correspond to probabilities in the same manner as the within-cluster weights, as the clusters cannot be considered mutually exclusive and are not an MECE discretization of the continuous ground-motion space.

The EPRI model was updated in 2013 as most of the GMM models that were included in 2004 had since been updated by their authors. The overall approach of using model clusters was retained, but the approach used to develop within-cluster and cluster-to-cluster weights was refined. In addition, the available empirical ground-motion database had been greatly expanded as part of the NGA-East project and was used to develop the model weighting functions. The empirical data were adjusted for site conditions in this application. Again, the goodness-of-fit to the empirical data was used to assign relative weights to the individual GMMs within a cluster. These weighted models were then used to compute the model-to-model variability in terms of a lognormal standard deviation.

This epistemic uncertainty was increased by two factors: (1) An assessment of a minimum level of epistemic uncertainty based on the amount of data available to constrain any model was performed. Recognizing that the available data was primarily for magnitude of **M** 5 and less, a second source of epistemic uncertainty was added based on range in magnitude scaling represented by the available GMMs; and (2) The combined within-cluster epistemic uncertainty was again represented by constructing of median, 5th percentile, and 95th percentile GMMs. As such, the within-cluster weights can be considered to represent probabilities. The cluster weights were again assigned based on a combination of how consistent the cluster medians were with the empirical data and on the Technical Integrator (TI) team's assessment regarding the robustness of the GMM underlying each cluster as extrapolators for magnitude and distance ranges of engineering interest.

6.2.4 PEGASOS PSHA (2004)

This PSHA study for Switzerland nuclear power plants (NPPs) gets its name from the German acronym Probabilistische Erdbeben-Gefährdungs-Analyse für KKW StandOrte in der Schweiz (NAGRA 2004). Based on observations from the SSHAC Level 4 study of the Yucca Mountain-PSHA project efforts, it was recognized that analysis had evolved so that now experts were using GMMs to develop the large suite of ground motions required and to ultimately capture the epistemic uncertainty.

The distributions were again vectors of ground motions with **M**, R, and f correlated through GMMs. Different weights could be applied to different sets of **M**, R, and f. The correlation across scenarios was preserved through the magnitude and distance scaling built into the GMMs. Again, the weights were based on the relative merit of the GMMs and did not represent a discretization of a continuous ground-motion distribution. Therefore, the branches are not MECE. Although allowing the weights to change with **M**, R, and f allowed more flexibility, it caused breaks in the PSA values at the boundaries that were propagated as breaks in the PSHA fractiles, leading to hazard curves that were not smooth.

6.2.5 BCHydro PSHA (2012)

This PSHA study, led by British Columbia Hydro Power (BCHydro), was conducted as a SSHAC Level 3 project to assess seismic hazard from both shallow crustal and subduction earthquakes, for their dam portfolio (Addo et al. 2012). For shallow crustal earthquakes, the TI Team utilized the then-common approach of weighting the available models, which in this case was the first set of NGA relationships published in 2008 (Power et al. 2008). Epistemic uncertainty in the median amplitude of each of the NGA models was included using the approach later published by Al Atik and Youngs (2014).

For subduction zone earthquakes, the TI Team concluded that the existing set of GMMs was inadequate and developed a new characterization. Learning from other SSHAC projects, and in an attempt to return to logic tree weights to represent discretized distributions, the project adopted the scaled-backbone approach. The weights are again on GMMs, but there was an attempt to define the weights such that the branches are MECE and represent the full ground-motion space. As for EPRI and PEGASOS, the distributions were again vectors of ground motions with **M**, R, and f correlated through GMMs. Instead of individual GMMs, a single "backbone" model was developed that was then scaled up and down to span the range of ground motions; this is the so-called scaled-backbone approach (Figure 6–1). Uncertainty in magnitude scaling for magnitudes larger than about **M** 8 was applied to the backbone model to account for the lack of empirical data at these magnitudes. Uncertainty in distance scaling (e.g., the effect shown on Figure 6–2) was not incorporated. The weights on the scaled GMMs allowed a complete MECE description of the ground-motion space, and the weights returned to representing discrete approximations of a continuous distribution of ground motions.

6.2.6 Canadian National Seismic Hazard Maps (Atkinson and Adams, 2013)

The development of the GMM suite for use in the national seismic hazard maps of Canada, as incorporated into the National Building Code of Canada (2015), was based on the use of a weighted representative GMM suite, specified for each event type and region (Atkinson and Adams, 2013). For crustal events in eastern Canada, a set of five published GMMs representing different model classes was used to define the range of epistemic uncertainty. The five models were those of Pezeshk et al. (2011) (hybrid empirical); Atkinson and Boore (2006, 2011) (stochastic finite-fault, constant stress); Atkinson (2008) and Atkinson and Boore, (2011) (referenced empirical); and Silva et al. (2002) (single-corner point-source, variable stress) and Silva et al. (2002) (double-corner point source with saturation). The mean predicted ground motion (in log space) and its standard deviation were defined as a function of magnitude and

distance using the five GMMs. The central GMM was defined based on the mean prediction of the five GMMs. Upper and lower-branch GMMs were defined to represent epistemic uncertainty as expressed by the standard deviation of the mean estimates. The resulting epistemic uncertainty model is a function of magnitude and distance (because the standard deviation of estimates varies with magnitude and distance). This uncertainty was smoothed and then checked against the epistemic uncertainty adopted for western crustal events; this check resulted in the decision to increase the implemented uncertainty at short distances. The rationale for checking uncertainty across regions was that the eastern GMMs should carry larger epistemic uncertainty than western GMMs (due to poor data constraints), and that epistemic uncertainty should be greatest for large magnitudes at short distances. This approach is similar to the scaled-backbone approach, except that here the backbone was based on a number of GMMs, and its epistemic uncertainty was considered to be variable in both magnitude and distance according to the state-of-knowledge and data constraints.

6.2.7 Hanford (2014)

The Hanford Sitewide PSHA was conducted as a SSHAC Level 3 project for five hazard calculation sites at Hanford, Washington (Coppersmith et al. 2014). The purpose of the study was to update the PSHA for the U.S. Department of Energy (DOE) facilities as well as to fulfill the requirement from the U.S. NRC that Energy Northwest conduct a SSHAC Level 3 PSHA for the Columbia Generating Station. The PSHA results were provided for a defined base rock horizon. Using the scale-backbone approach, ground-motion logic trees were developed to capture the distribution of median ground motions for both shallow crustal earthquakes GMMs and subduction earthquakes GMMs.

Project-specific criteria were defined for the selection of the backbone GMM for crustal earthquakes. As a result, four NGA-West2 GMMs (Abrahamson et al. 2014; Boore et al. 2014; Campbell and Bozorgnia 2014; and Chiou and Youngs 2014) passed the selection criteria, and the Chiou and Youngs (2014) GMPE was selected as the backbone model. For subduction earthquakes, a backbone GMM was developed for this project based on revisions to the BCHydro model (Addo et al. 2012). These revisions were necessary to accommodate project-specific constraints, such as predictions from sources at distances of 200–400 km from the Hanford site, and to incorporate additional ground-motion data that have become available since the introduction of the BCHydro model.

The objective of the developed ground-motion models was to capture the full range of epistemic uncertainty in median ground-motion predictions consisting of two components: inherent uncertainty in the predictions within the host region and uncertainty in the adjustments between the host and target regions.

For crustal earthquakes, the scaled-backbone model was developed in two steps. The first step consisted of developing the model for footwall conditions. This allowed for the development of appropriate factors to center the backbone model and capture the range in magnitude and distance scaling. Based on the evaluations of the candidate GMMs over the range of distances important to hazard at the sites, it was concluded that differences in distance scaling were adequately captured by the up and down amplitude scaling of the central backbone. The candidate models were then used as a sample to compute a bi-variate lognormal distribution for

scaling the backbone to model epistemic uncertainty in amplitude and magnitude scaling at each period. These bi-variate distributions were represented discretely by a set of 9 scaled backbone models. Additional amplitude scaling uncertainty was applied to address uncertainty in potential differences between the host region for the candidate models and the site target region. The second step was to develop a separate distribution of hanging-wall effects using the hanging wall factors from Abrahamson et al. (2014), Campbell and Bozorgnia (2014), and Chiou and Youngs (2014).

To characterize subduction zone earthquake ground motions, the scaled backbone model developed by BC Hydro (2012) (Addo et al. 2015) was adopted and modified to address additional ground-motion data. The scaled backbone model was then extended to address uncertainty in distance scaling because of the large distance of the site from the primary subduction zone earthquake source.

6.2.8 SWUS (2015)

The South-Western U.S. utilities project (SWUS) is a PSHA study funded by private utilities for NPP seismic hazard assessment of facilities in California and Arizona (GeoPentech 2015). Similar to the projects discussed above, SWUS was targeting site-specific analyses. Building on previous efforts, SWUS was first to adopt and further develop a new approach based on visualization techniques of high-dimensional ground-motion space. The concept behind the approach is documented in Scherbaum et al. (2010). The SWUS project ran in parallel to NGA-East, and many of the concepts presented herein were shared by both projects, although their implementations are quite different.

The SWUS approach used a novel method to determine true characterization and sampling of the ground-motion space. The principles behind the method and its implementation for the NGA-East case are presented in Chapter 8. Basically, ground motions from each GMM are mapped into a high-dimensional (or multi-dimensional) space. This is done one frequency at-a-time, and each (**M**, R) pair corresponds to a dimension of the high-dimension space. Each GMM can then be represented by a single point in the high-dimension space. The key to the method is that differences in GMMs can be assessed by their separation (i.e., ground-motion distance) in the high-dimension space. This ground-motion space distance is also referred to as a GMM-distance, and it is a measure of difference in ground motions (as opposed to a measure in the sense of travel distance). For a 3D space consisting of ground motions for three different (**M**, R) scenarios, the GMM-distance would be measured along a straight line. We can imagine computing GMM-distances between points for more than three dimensions in a curved Euclidian space (this is described in Chapter 8). By taking those GMM-distances and projecting them on a 2D plane [using Sammon (1969) maps], we can represent the complete ground-motion space, spanning the complete ground-motion distribution for multiple (**M**, R) pairs for a specific frequency at once. The 2D representation of the GMM-distance is approximate, but the distortion is minimized through the Sammon's mapping process (Chapter 8). By defining branches for specific regions or points on the Sammon's map and assigning weights to them, we can effectively re-discretize the ground-motion space and make the set of branches of the logic tree align more closely with the concept of MECE. The principle is the same as originally

used in the TIP (1995) project, but applied to vectors (correlated sets of ground motions) as opposed to scalars (distribution of ground motions for a given (**M**, R) scenario).

The scaled-backbone approach, where each GMM may produce a given set of ground motions following its own built-in correlations, considers multiple rules for scaling. The SWUS approach, however, uses high-dimension visualization technique that relies on Sammon's map projections as a way to return to the original objective of *quantifying epistemic uncertainty as a discretization of a continuous ground-motion space*.

6.2.9 Summary and Motivation

The proper quantification of epistemic uncertainty in ground motions is a critical task in the development of a GMC for PSHA projects. From the beginning of large PSHA projects in the 1990s, the epistemic uncertainty in ground motions was intended to represent a discretization of the ground-motion space. This can be achieved using logic trees if each branch of the logic tree represents parts of the distribution that are MECE. The various approaches discussed in this and the previous sections are summarized in Table 6–1. This list is not meant to be exhaustive, as other PSHA projects exist, but the subset was selected to summarize the motivation for NGA-East's approach.

For a simple scenario [e.g., a given (**M**, R) pair], this can be easily done through expert elicitation. When trying to consider multiple scenarios, as is usually required for PSHA, it becomes difficult to aggregate the information in a simple, meaningful way. To achieve this, the approach went from expert elicitation of expected ground-motion values (scalars) to expert elicitation of sets of values (vectors), which may or may not be correlated. When experts were confronted with a large number of scenarios, they often used GMMs as algorithms for defining the ground motions and their correlations. The epistemic logic tree branches now consisted of GMMs instead of ground-motion values, and the weights were a representation of the relative merit of each GMM approach as assessed by a group of experts. This approach has the advantage of maintaining the correlation in magnitude and distance scaling between scenarios through the GMMs, but the concept of "weight as a probability" (and the goal for the distribution to be MECE) may be lost. In other words, on the one hand, two GMMs that clone each other and represent the same ground-motion space will not be mutually exclusive, allowing for some double-counting of the occurrence of certain ground-motion values. On the other hand, there may not be enough GMMs to capture the full range of ground motions, leading to gaps in ground motions or to incomplete distributions; the GMMs may not collectively exhaust the ground-motion space.

One solution to this problem was to use the backbone GMM approach (Bommer 2012; Addo et al. 2012; and Atkinson et al. 2014). This is arguably a defensible approach to capture the center, body, and range of epistemic uncertainty associated with GMMs. In its simplest form, a single GMM is selected as representative of the correlation of magnitude and distance scaling for a region. The GMM is scaled up and down to cover the range of ground motion, such as shown in Figure 6–1. As discussed in the examples above, backbone scaling can accommodate differences in magnitude and distance scaling as appropriate for the contributions to the site hazard [e.g., BC Hydro (2012), Atkinson and Adams (2013), and Hanford (2014)]. One can also consider the EPRI (2004; 2013) models as scaled backbones representing the within-cluster

epistemic uncertainty. Now discretized, the distribution consists of a suite of branches, each of which is associated with a scaling factor for the GMM, and all branches are defined to collectively be as close as possible to a discretized MECE representation of the full distribution of ground motions. The backbone approach represents an improvement relative to previous approaches in capturing the epistemic uncertainty, but may not fully capture the range of alternative interpretations of the data that affect magnitude and distance scaling without careful consideration of the range of possible interpretations.

The challenge for the quantification of epistemic uncertainty is to allow a discretization of the continuous ground-motion space into sets of GMMs that include alternative magnitude and distance-scaling correlations in a comprehensive way. The goal is to develop methods to generate, select, and combine discrete GMMs so that they represent the complete ground-motion space and are mutually exclusive, which essentially treats GMM weights as probabilities. The proposed NGA-East approach summarized in this report aims to achieve this goal.

6.3 NGA-East Approach to Median Ground Motions and Epistemic Uncertainty

6.3.1 Overview

This NGA-East approach to median ground motions is based on the understanding that the associated epistemic uncertainty can be described by a *continuous* probability distribution. This is a view that has been used in some previous approaches (e.g., TIP, BCHydro, SWUS, and others), a methodology recently employed again in Atkinson et al. (2014). This key point here is that it is not sufficient to put weights on a discrete number of *existing* GMMs to capture the *center*, *body,* and *range* of epistemic uncertainty associated with median ground-motion estimates.

There are different ways to model a continuous distribution for median ground motions. Working closely with the SWUS (GeoPentech 2015) project since its inception, NGA-East followed a similar conceptual approach but with noted differences. The SWUS project modeled a continuous distribution over GMMs by estimating a (joint) distribution of coefficients assuming a single GMM functional form, which in turn led to a distribution of ground-motion values. In contrast, the NGA-East TI team decided to estimate a distribution over ground-motion values directly, without the detour of a distribution over coefficients. In this approach, selected “seed” GMMs were used to generate ground-motion estimates, which then formed the probability distribution P(**Y**), where **Y** is a vector of median ground-motion estimates at different (**M**, R) scenarios, **Y**= $\{Y_1, Y_2,...,Y_N\}$. Here, Y_i denotes the logarithmic response spectral value at a particular frequency, and the index goes over the (**M**, R) pairs {M, R}. **Y** is a high-dimensional vector of ground motions, each dimension corresponding to a specific (**M**, R) scenario.

In the context of the TIP project, for example, the experts would have been asked to provide their estimate of the distribution for each Y_i individually and working with one dimension (i.e., one (**M**, R) scenario) at-a-time. For NGA-East, both the Y_i marginal distributions and the full joint distribution P(**Y**), which includes the correlations between the individual Y_i, were considered. Given that P(**Y**) represents the epistemic uncertainty associated with median ground-motion estimates, it should therefore be the basis of the center, body, and range. This is different from

the EPRI 2013 project, which was based on the 5% and 95% quantiles for different clusters on the marginal distributions at each (**M**, R) scenario.

Ideally for hazard calculations, one would work directly with the continuous distribution P(**Y**); however, this is computationally not feasible. Therefore, for practical reasons P(**Y**) needs to be discretized. The NGA-East approach is to discretize P(**Y**) into a manageable number of representative GMMs based on drawing from a large number of samples from P(**Y**), which are then evaluated to determine how well they cover the expected range. In this context, it is important to recall that a sample from P(**Y**) is a vector of ground-motion values, but due to the correlation inherent in P(**Y**), the entries of the vector behave like a GMM; for additional discussion see Chapter 8.

The assessment of the samples from P(**Y**), each of which is a high-dimensional vector, is carried-out via high-dimensional visualization tools (Scherbaum et al. 2010), which makes it possible to collapse the analysis down to two dimensions. These tools provide a way to assess the samples in a global sense—as opposed to a single scenario at-a-time—and allow a more intuitive definition of the center, body, and range of P(**Y**); see Chapter 8. Thus, the selection of a set of representative GMMs is based on these visualization tools. The visualization tools also provide a way to calculate weights based on prior information and data for the set of representative GMMs.

The NGA-East approach is based on the premise that epistemic uncertainty in median ground motions can be represented by a continuous distribution in ground-motion space. The approach can be summarized by the following key five steps:

1. Develop a suite of seed GMMs
2. Develop parameters for continuous distributions of GMMs
3. Visualize the ground-motion space and sample GMMs
4. Re-discretize the ground-motion space
5. Assign weights

The following two sub-sections present simple examples (1D and 2D, respectively) of the application of the NGA-East approach. The NGA-East project is dealing with “ND”, or a large N-dimensional version of the problem, with each (**M**, R) scenario pair represents a dimension. For these 1D and 2D examples, these steps are trivial, but their formal application serves as a simple entry to the problem in higher dimensions.

6.3.2 One-Dimensional Example

In this section, a one-dimensional (1D) problem is used to illustrate the NGA-East approach. The problem is centered around quantification of the epistemic uncertainty in median PSA (1 Hz) estimates for a **M** = 6 and R = 200 km earthquake scenario. The example was selected to illustrate the *process* only, and some critical decision steps that would be based on the TI evaluation have been replaced by assumptions.

Step 1. Develop a suite of seed GMMs

This step involves evaluating available GMMs for suitability to the problem at hand. This example assumes that 18 (fictitious) GMMs were deemed applicable and selected as seed models. Table 6–2 lists the median ground-motion estimates for the 18 models.

Step 2. Develop parameters for continuous distributions of GMMs

The characterization of the full ground-motion distribution is performed by estimating the means, variances, and correlations between median values at the different (**M**, R) scenarios considered. For this single scenario, the estimation of the continuous distribution is obtained by calculating the mean and variance (or standard deviation) of the seed GMM estimates. Figure 6–3 shows the median estimates from Table 6–2. Figure 6–3 also shows the estimated normal distribution describing the epistemic uncertainty, with mean μ_Y = -5.166 and σ_Y = 0.245. Since there is only one scenario, we can't compute the correlation of ground motions across scenarios, which will be needed in the multiple dimension case.

Step 3. Visualize the ground-motion space and sample GMMs

The visualization step is critical in the real application of the NGA-East methodology. However, for the 1D example, it is trivial and the visualization step is achieved through Figure 6–3. A suite of 5000 ground-motion values were then sampled from the distribution defined above. They are represented (visualized) by the histogram in Figure 6–4.

Step 4. Re-discretize the ground-motion space

The next steps involve the definition of the range in ground-motion space, the partitioning of that range, and the selection of representative sets of GMMs. There are many tasks required to complete this step, which involve the TI team evaluation.

As an illustration, let's assume that the range of median ground motions to be captured is bounded by the mean plus/minus two standard deviations, and that the goal is to obtain a representative set of five models at the end of the process. The truncation to plus/minus two standard deviations and the definition of five models were selected for illustration purposes of the process only. Accordingly, the distribution P[Y(**M** = 6, R = 200 km)], is discretized into five bins, as shown in Figure 6–5. In this case, bins of equal widths were selected, but this is not a prerequisite. Another assumption is made that the representative model can be defined as the mean of all the models in a particular bin. These values are shown in Figure 6–6 as blue dots.

Step 5. Assign weights

The next step is to calculate weights for the set of representative models. There are different ways to calculate weights: they can be based on information from the PDF of P[Y(**M** = 6, R = 200 km)], (i.e., "prior information") or on the model fit to existing data. Examples of these approaches are presented below.

To calculate weights based on the PDF, two different approaches can be used. The first is to calculate the probability density for each bin by integrating over the PDF:

$$w_{PDF,i} \propto \int_{LB_i}^{UB_i} p(y)\,dy \tag{6–1}$$

where LB_i and UB_i are the lower bound and upper bound of the i^{th} cell, and p(y) is the PDF of P[Y(**M** = 6, R = 200 km)]. In this example, the distribution is essentially truncated, and because the weights calculated do not sum to one, they would need to be normalized.

Weights can also be obtained by calculating the number of samples in each bin using an approximation, as opposed to using Equation (6–1):

$$w_{N,i} \propto N_i \tag{6–2}$$

where N_i is the number of models in the i^{th} cell.

An alternative weight computation process is based on the ability of the models to fit an observational data set. Let's assume we have a dataset available for this scenario, presented as green dots in Figure (6–7). The example data has different mean and standard deviation from the seed median predictions. Since this is just an illustrative example, the actual value is not particularly important. For each sampled model, we calculate the mean residual to the data, as well as the likelihood of the data under each model. The latter is computed by:

$$L_i = \prod_{j=1}^{N} p(y_j \mid m_i) \tag{6–3}$$

where N is the number of data points, $p(y_j \mid m_i)$ is the value of a normal density function with mean m_i, where i indexes the sampled models.

The log-likelihood and mean residual for each sampled model are plotted in Figure 6–8. They both exhibit a clear trend, which is not surprising, since each of the sampled models consists of only one parameter. The mean residual value is zero at the mean of the data; this is also where the curve of the log-likelihood has its maximum. The residuals and the likelihood can then be used to weigh the models selected earlier. Each of the selected models is representative of a given bin, so it needs to reflect the likelihoods and residuals of all sampled models in that particular bin. To calculate weights for each bin based on the data, the following rationale is used:

1. Models with low residual should receive higher weight.
2. Models with high likelihood should receive higher weight.

Hence, weights are calculated based on the following formulae

$$w_{res,i} \propto \frac{A_i^1}{\left| \frac{1}{N_i} \sum_{j=1}^{N_i} r_j \right|} \tag{6–4}$$

$$w_{L,i} \propto A_i \frac{1}{N_i} \sum_{j=1}^{N_i} L_j \tag{6–5}$$

where N_i is the number of samples in the i^{th} bin, A_i is its width, and j indexes over the samples in bin i. The weight based on residuals is inverse-proportional to the mean residual of one cell, whereas the weight based on the likelihood is proportional to the mean likelihood in one cell. Basically, the weights can be understood as an approximation to an expectation of a function over one particular bin:

$$w_{res,i} \propto \sum_{LB_i}^{UB_i} f(y)p(y)dy \tag{6–6}$$

where LB_i and UB_i are again the lower bound and upper bound of the i^{th} cell, and f(y) is the function over which the expectation is computed (such as the likelihood).

Figure 6–9 compares the weights calculated using the four different approaches. As is expected, weights based on the PDF and the number of models are very close. They are also symmetric since they are based on the normal distribution inferred from the seed GMMs, and they favor (give the highest weight to) the center model. On the other hand, weights based on the data are skewed because the mean of the data is not at the mean of the seed distribution.

The TI-team's next step, if they were to complete the 1D example, would then be to combine the weights based on the different approaches by providing "weights on weights." If the data are deemed very reliable, one might argue to give very high weight to the likelihood and/or the residuals. On the other hand, if the data are thought to be uncertain or not very representative, one would give more weight to the weights based on the PDF.

6.3.3 Extension to Two-Dimension Example

In the previous section, the principle behind the NGA-East approach was illustrated with a one-dimensional example. Now, this example is extended to two (**M**, R) scenarios. In this case, it is important to consider the joint distribution $P(\mathbf{Y}) = P(Y_1,Y_2)$ of median predictions for these two scenarios. The two-dimensional (2D) example is done for scenarios ($\mathbf{M}_1$, R_1) = (6, 200 km) and ($\mathbf{M}_2$, R_2) = (8, 200 km) and 1 Hz PSA.

Imagine the one-dimensional example as a case for a specific (admittedly very unrealistic) hazard application, where one wants to know the expected ground-motion at one site that is only affected by one source (thus, only one distance), and this source only produces one magnitude. The 2D example is an extension of that scenario in that now the source can produce two magnitudes. If one assumes that there is a linear scaling of ground motion with a magnitude between $\mathbf{M}_1$ and $\mathbf{M}_2$, then the median predictions Y_1 and Y_2 for the two scenarios are sufficient to describe this situation as well, since one can interpolate between these two ground motions.

For this simplified 2D example, a GMM can be thought of as a 2D vector, containing the two median predictions Y_1 and Y_2. If one considers a 2D space, where one coordinate is Y_1 and the other coordinate is Y_2, then a GMM is a point in this space (see Table 6–3 and red dots in Figure 6–10).

Step 1. Develop a suite of seed GMMs

Again, this example assumes that there are 18 suitable GMMs for 1 Hz PSA. The median estimates from those 18 GMMs for both scenarios are presented in Table 6–3 and plotted in Figure 6–10.

Step 2. Develop parameters for continuous distributions of GMMs

The continuous distribution for the 2D example is modeled as a 2D normal distribution, where the means, standard deviations and correlations are estimated from the median predictions of the 18 seed GMMs for the two scenarios. A contour plot of the probability density function of this 2D distribution is shown in Figure 6–11. For the first scenario, the mean and standard deviation of the 18 median predictions are μ_{Y1} = -5.166 and σ_{Y1} = 0.246; for the second scenario, they are μ_{Y2} = -2.598 and σ_{Y2} = 0.381. The correlation coefficient between the two scenarios is ρ = 0.53. The positive correlation between Y_1 and Y_2, i.e., between the median predictions at the two scenarios, means that a GMM that predicts a relatively high PSA for the first scenario will also likely predict a high PSA for the second scenario (Figure 6–11).

If the correlation were perfect (ρ = 1), then the continuous distribution would approach a straight line in Figure 6–11. This would correspond to a scaled-backbone model (typically, the standard deviations in a scaled-backbone model are assumed to be the same for all scenarios). As mentioned earlier, in a scaled backbone model, the median prediction of a reference GMM is scaled up/down by a set of constant factors to capture the range of epistemic uncertainty associated with median ground-motion predictions; there is no difference in the magnitude/distance scaling between the scaled models. Let's assume that the reference prediction for the two scenarios corresponds to the means μ_{Y1} and μ_{Y2}. If the reference motion for scenario 1 (μ_{Y1}) is increased by a certain amount, then the amount by which the reference motion for scenario 2 (μ_{Y2}) needs to be increased is already determined relative to that amount.

By contrast, in the case of a non-perfect correlation ($\rho \neq 1$) as depicted in Figure 6–11, if the reference motion for scenario 1 is increased by a certain amount, then there is a range of possible Y_2 values, as evidenced by the ellipsoidal contours of the probability density function. This means that there are variations in the possible magnitude/distance scalings.

One can also interpret Figure 6–11 in terms of a two-dimensional ground-motion space, defined by the two variables Y_1 and Y_2. Each point in this two-dimensional space corresponds to a joint median prediction (Y_1, Y_2) for the two scenarios; however, not all of these pairs (Y_1, Y_2) make physical sense. Only the region of high-probability density yields physically reasonable predictions. Traditionally, the joint range of physically reasonable median predictions (Y_1, Y_2) is constrained by different yet reasonable values of the magnitude/distance scaling. Alternatively, the physically possible region can be constrained by the mean and covariance of the joint distribution P(Y_1, Y_2).

Step 3. Visualize the ground-motion space and sample models

As was the case for the 1D example, the visualization step is only mentioned for completeness. Since the ground-motion space is two-dimensional in this example, the location of any vector

contained in the 2D space can easily be plotted; see the red dots in Figure 6–10. The next steps of the full process [i.e., working in a high-dimensional ground-motion space, defined by a large number of (**M**, R)] are very similar to the next steps of the 2D example; the visualization step projects the high-dimensional ground-motion space to a two-dimensional map (Scherbaum et al. 2010; Sammon 1969).

A suite of 5000 ground-motion values were then sampled from the 2D distribution defined above. They are represented (visualized) by the histogram and by the cloud of points in Figure 6–12.

Step 4. Re-discretize the ground-motion space

The next step involves discretizing the ground-motion space, thereby reducing the number of models to a manageable subset representing the continuous distribution. This step involves judgment on behalf of the TI team, who must decide how many discrete models are sought, and how much of the 2D space should be covered. Theoretically, the density of the 2D normal distribution is non-zero everywhere (similarly, in the 1D case the normal distribution reaches zero only in the limit of plus/minus infinity); however, to capture the CBR of the TDI, one has to define limits for the median predictions.

In the current example, the 2D ground-motion space is partitioned into cells; see Figure 6–13. These cells should be chosen such that they cover approximately the high-density region, which corresponds to physically meaningful models. Again, this example is provided for illustration purposes only. The cells are the 2D equivalent of the intervals shown in Figure 6–5 for the 1D example. Each cell covers a part of the ground-motion space, and one needs to find a representative point/model for each cell. Here, the representative example is selected as the mean of all points within one cell (black points in Figure 6–14); note that the model definition takes into account that there is a gradient of the density within each cell.

Step 5. Assign weights

The assignment of weights in the 2D case follows the 1D example. Weights can be assigned based on the probability density function, or, if data are available, based on the fit of a model to data.

Weights based on the probability density function are calculated as in the 1D example. The weights for the different models are proportional to the integral of the probability density function over the cells (cf. Equation 6–1). Alternatively, one can sample from the underlying distribution, and count the number of samples within each cell (Equation 6–2).

If observed data are available, one can calculate weights in the same way as for the 1D example. For illustrative purposes, a synthetic dataset was simulated, and weights were calculated based on residuals and likelihood of the simulated data for all sampled models. The synthetic data set was generated as follows:

- The “data” (synthetic data simulated for this example) model is selected to generate median predictions of Y_1 = -5.00 for scenario 1 (**M** = 6, R = 200 km) and Y_2 = -2.61 for scenario 2 (**M** = 8, R = 200 km).

- For intermediate magnitudes between **M** = 6 and **M** = 8, it is assumed that the prediction of the model can be calculated from scenarios 1 and 2 by a linear interpolation between them. We start by randomly selecting 19 magnitudes with values between 6 and 8, and then calculating median predictions for these 19 magnitudes from the "data" model based on a linear interpolation between the predictions at the two scenarios. Then, for each magnitude a PSA value is sampled from a normal distribution whose mean is the interpolated median prediction for that particular magnitude. The standard deviation is set to 0.6; again, this corresponds to the aleatory variability. Thus, we end up with a set of 21 PSA values at different magnitudes (since the distances for the two scenarios are both 200 km, the distance for each data point is also 200 km), which we treat as "data" for this example, as shown in Figure 6–15.

For each sampled model (gray point in Figure 6–12), one can calculate median predictions at the simulated magnitudes by a linear interpolation between the median predictions for that particular model (in the 2D case, the median predictions for scenarios 1 and 2 correspond to the two coordinates). Thus, for each sampled model one can calculate a residual for each simulated PSA value. Similarly, one can calculate the likelihood of the simulated PSA values. Hence, for each sampled model, one can calculate the mean residual and the likelihood of the simulated data—Figure 6–16—which shows that the center of the cloud sample is skewed, especially relative to the likelihood. Then, for each cell one can calculate weights based on the mean residuals and likelihoods of all sampled models within each cell, according to Equations 6–4 and 6–5; see Figure 6–17. The last step for the TI team, if it was to complete the 2D example, would be to define the weights to be applied to each of these weights so as to obtain a single weight for each of the representative models.

6.3.4 Actual Process

The previous section provided two simple illustrations of NGA-East's conceptual approach to epistemic uncertainty quantification. In this section, the same process is generalized for application to multiple dimensions (scenarios). The same five basic steps described above are followed, expanded upon, and illustrated in the flowcharts shown in Figure 6–18. In these charts, computations are shown as gray rectangles, while TI team decisions are depicted as purple diamonds. Results of computations/decisions are in yellow parallelograms. The details of each step are presented in Chapters 7–9.

The key difference from the 1D and 2D examples, is that in the actual NGA-East process, the number of (**M**, R) scenario increases, resulting in a high-dimensional ground-motion distribution P(**Y**), the task becomes more difficult. For the NGA-East GMC model development, several scenarios were considered, which led to a vector of ground motions **Y** whose entries corresponded to the median predictions for each individual (**M**, R) scenario. Consequently, the associated epistemic uncertainty is described by a high-dimensional ground-motion distribution P(**Y**). Each entry of the vector **Y** can be thought of as a coordinate in a high-dimensional space, where the number of dimensions is the length of the vector [i.e., the number of (**M**, R) scenarios], i.e., how a 3D vector describes a point in 3D space. Hence, a particular ground-motion vector **y** can be thought of as a point in a high-dimensional ground-motion space. To make this space accessible, high-dimensional visualization tools (Scherbaum et al. 2010) were

used, in particular, Sammon's mapping (1969). This is done to approximate the high-dimensional ground-motion space through projection in two dimensions. The resulting product is a 2D map representation of the ground-motion space, making it easy to follow the steps outlined in the current chapter, with a slight modification to go from one dimension to two dimensions. This is explained in detail in Chapter 8.

Step 1. Develop a suite of seed GMMs

The first step, described in detail in Chapter 7, was the compilation of a number of GMMs that were deemed applicable to CENA. These "candidate GMMs" were then evaluated to assess their scaling with magnitude, distance, and frequency based on criteria developed by the TI team. Following this initial screening, a set of feasible (technically defensible) GMMs were selected from the candidate models. The set of selected models may be different for different frequencies. It was found that for very short and large distances, most models required adjusting/extrapolations to meet the {**M**, R} requirements of NGA-East. After these adjustments were made, sets of seed GMMs were selected to define the continuous distribution or ground motions over a large suite of (**M**, R) scenarios. This seed set is analogous to the red points in the 1D (Figure 6–3) and 2D (Figure 6–10) examples.

Step 2. Develop continuous distributions of GMMs

The selected seed GMM set was used to generate ground-motion estimates. The variance and correlation (the covariance) between the different (**M**, R) scenarios could then be calculated. For a ground-motion vector **Y**, the correlations across the various (**M**, R) scenarios were required to ensure that a sample from P(**Y**) would be a valid GMM.

The covariance model, together with the seed GMMs, defines the continuous distribution describing median ground motions, P(**Y**). Hence, samples can be drawn from it (as it was the case for the 2D example, in Figure 6–12). The development of the continuous distribution of ground motions over the scenarios of interest is presented in Section 8.1. Generally, the correlation between different Y_i should ensure that a sample from P(**Y**) is a physically valid GMM. However, it is still possible to generate samples that violate physical assumptions. To avoid samples from P(**Y**) that are unphysical, the TI team defined criteria for physicality required to satisfy a GMM; see Section 8.3.

Step 3. Visualize the ground-motion space and sample models

The result of Step 2 is a number of samples from P(**Y**), each of which meets the physicality constraints and constitutes a valid GMM. Each sample from P(**Y**) is a high-dimensional vector **y** and thus is a point in a high-dimensional space. Samples for the distribution developed in Step 2 were then used to populate the ground-motion space. These points can be projected into 2D using visualization techniques. This process constitutes the essence of Step 3, which is described in detail in Sections 8.2 to 8.4.

Step 4. Re-discretize the ground-motion space

Starting with the 2D projection of the high-dimensional ground-motion space from the step above, the center, body, and range are defined. The underlying assumption is that the 2D

projection calculated in the previous step is a reasonable representation of the ground-motion space and hence of P(**Y**); this assumption is supported and defended in Section 8.4. The TI team's decision involved defining the bounding shape in two dimensions corresponding to Figure 6–13 in the 2D example. The result is a set of cells that cover the center, body, and range on the 2D representation of P(**Y**). Each of the cells comprises several samples from P(**Y**), which are combined into one representative model for that cell (Figure 6–13). The discretization step is detailed in Section 8.5.

Step 5. Assign weights

This last step involved taking the selected set of representative GMMs from the previous step and assigning weights to them. The process is directly comparable to that of the 1D and 2D examples, and is presented in Chapter 9. The total weights are a combination of weights based on "prior information" from the GMM distribution itself and from the available data.

As will be shown in Chapters 7, 8, and 9, the complete NGA-East process resulted in a suite of 17 GMMs per frequency, each with an assigned weight.

6.4 Aleatory Variability of Ground Motions

The model development for the aleatory variability (i.e., standard deviation) of ground motions followed a simpler process, consisting of two key steps: (1) model building, and (2) weight assignment (Chapters 10 and 11). The model building part involved analysis of the various components of ground-motion variability using recorded data from CENA. Trends of ground-motion variability with parameters such as magnitude, distance, and V_{S30} were analyzed and compared to observed trends of ground-motion variability in other regions, with special consideration of trends noted in the WUS using the NGA-West2 dataset.

The magnitude range in the CENA dataset is limited to small-to-moderate earthquakes and frequencies between 1 and 10 Hz due to the bandwidth limitations of the recordings. Therefore, aleatory variability models developed using the CENA ground-motion data cannot be reliably extrapolated to large magnitudes and frequencies outside of 1 to 10 Hz. As a result, aleatory variability models from other regions such as WUS and Japan are used to inform the extrapolation of CENA models and overcome data limitations. Candidate models for between-event standard deviation (τ), single-station within-event standard deviation (ϕ_{SS}), and site-to-site variability (ϕ_{S2S}) were developed for CENA. In turn, these models were combined to develop single-station sigma (σ_{SS}) and ergodic sigma (σ) models for CENA.

The model-building phase was followed by the evaluation of the models and the weight assignment. This last task was performed by the TI team and involved expert judgment based on the data available.

6.5 Complete Model Development and Organization of the Report

The complete model development documentation is covered in Chapters 7 to 11 and summarized topically as follows:

Median ground motions – Figure 6–18 flowchart items

- Step 1. Develop a suite of seed GMMs: Chapter 7
- Step 2. Develop parameters for continuous distributions of GMMs: Section 8.1
- Step 3. Visualize the ground-motion space and sample GMMs: Sections 8.2-8.4
- Step 4. Re-discretize the ground-motion space: Section 8.5
- Step 5. Assign weights: Chapter 9

Standard deviation of ground motions:

- Development of candidate models: Chapter 10
- Evaluation and weight assignment: Chapter 11

6.6 References

Abrahamson, N.A., W.J. Silva, and R. Kamai. 2014. “Summary of the ASK14 Ground-Motion Relation for Active Crustal Regions.” Earthq. Spectra, 30(3): 1025–1055.

Addo, K.O., N.A. Abrahamson, and R.R. Youngs. 2012. “Ground Motion Characterization (GMC) Model.” BCHydro SSHAC Level 3 Probabilistic Seismic Hazard Analysis Report.

Al Atik, L., and R.R. Youngs. 2014. “Epistemic Uncertainty for NGA-West2 Models.” Earthq. Spectra, 30(3): 1301–1318.

Atkinson, G.M. 2008. “Ground Motion Prediction for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty.” Bull. Seismol. Soc. Am., 98 (3): 1304–1318.

Atkinson, G.M., and J. Adams. 2013. “Ground Motion Prediction Equations for Application to the 2015 National Seismic Hazard Maps of Canada.” Can. J. Civil Eng., 40: 988–998.

Atkinson, G.M., J.J. Bommer, and N.A. Abrahamson. 2014. “Alternative Approaches to Modeling Epistemic Uncertainty in Ground Motions in Probabilistic Seismic-Hazard Analysis.” Seismol. Res. Lett., 85(6): 1141–1144.

Atkinson, G.M., and D.M. Boore. 2006. “Earthquake Ground-Motion Prediction Equations for Eastern North America.” Bull. Seismol. Soc. Am., 96(6): 2181–2205.

Atkinson, G.M., and D.M. Boore. 2011. “Modifications to Existing Ground-Motion Prediction Equations in Light of New Data.” Bull. Seismol. Soc. Am., 101(3): 1121–1135.

Bommer, J.J. (2012). “Challenges of Building Logic Trees for Probabilistic Seismic Hazard Analysis.” Earthq. Spectra, 28(4): 1723–1735.

Bommer, J.J., and F. Scherbaum. 2008. “The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis.” Earthq. Spectra, 24(4): 997–1009.

Boore, D.M., J.P. Stewart, E. Seyhan, and G.M. Atkinson. 2014. “NGA-West 2 Equations for Predicting PGA, PGV, and 5%-Damped PSA for Shallow Crustal Earthquakes.” Earthq. Spectra, 30: 1057–1085.

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Clu, K.J. Coppersmith, C.A. Cornell, and P.A. Morris. 1997. “Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts.” Report NUREG/CR-6372, U.S. Nuclear Regulatory Commission, Washington D.C.

Campbell, K.W., and Y. Bozorgnia. 2014. “NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration Response Spectra.” Earthq. Spectra, 30(3): 1087–1115.

Chiou, B.-S.J., and R.R. Youngs. 2014. “Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra.” Earthq. Spectra, 30(3): 1117–1153.

Coppersmith, K.J., J.J. Bommer, K. Hanson, R. Coppersmith, J.R. Unruh, L. Wolf, R.R. Youngs, L. Al Atik, A. Rodriguez- Marek, G. Toro, and V. Montaldo-Falero. 2014. “Hanford Sitewide Probabilistic Seismic Hazard Analysis.” PNNL- 23361, Pacific Northwest National Laboratory, Richland, WA.

CRWMS M&O. 1998. “Probabilistic Seismic Hazard Analyses for Fault Displacement and Vibratory Ground Motion at Yucca Mountain, Nevada.” Milestone SP32IM3, September 23, 1998. Three volumes. Las Vegas, NV.

EPRI. 2004. “CEUS Ground Motion Project Final Report.” EPRI Report 1009684, Palo Alto, CA.

EPRI. 2013. “EPRI 2004-2006 Ground-Motion Model (GMM) Review Project.” Final Report Project 3002000717. June 2013, Electric Power Research Institute, Palo Alto, CA.

GeoPentech. 2015. “Southwestern United States Ground Motion Characterization SSHAC Level 3 – Technical Report Rev.2.”

Miller, A.C., and T.R. Rice. 1983. “Discrete Approximations of Probability Distributions.” Management Sci., 29: 352−362.

NAGRA. 2004. “Probabilistic Seismic Hazard Analysis for Swiss Nuclear Power Plant Sites (PEGASOS Project).” Report to Swiss Nuclear prepared by Nationale Genossenschaft für die Lagerung radioaktiver Abfälle, Wettingen, 358 pgs.

NBC. 2015. National Building Code of Canada, National Research Council Canada, Ottawa, Canada 1404 pgs.

NRC. 2012. “Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies.” Technical Report NUREG-2117, U.S. Nuclear Regulatory Commission, Washington D.C.

Pezeshk, S., A. Zandieh, and B. Tavakoli. 2011. “Hybrid Empirical Ground-Motion Prediction Equations for Eastern North America Using NGA Models and Updated Seismological Parameters.” Bull. Seismol. Soc. Am., 101(4): 1859–1870.

Power, M., B.-J. Chiou, N.A. Abrahamson, Y. Bozorgnia, T. Shantz, and C. Roblee. 2008. “An Overview of the NGA Project.” Earthq. Spectra, 24: 3–21.

Sammon, J.W. 1969. “A Nonlinear Mapping for Data Structure Analysis.” IEEE Trans. Computers, C-18: 401–409.

Savy, J.B., W. Foxall, N.A. Abrahamson, and D. Bernreuter. 2002. “Guidance for Performing Probabilistic Seismic Hazard Analysis for a Nuclear Plant Site: Example Application to the Southeastern United States.” Report NUREG/CR-6607, Lawrence Livermore National Laboratory (UCRL-ID-133494), Livermore, CA.

Scherbaum, F., N.M. Kuehn, M. Ohrnberger, and K. Koehler. 2010. “Exploring the Proximity of Ground-Motion Models Using High-Dimensional Visualization Techniques.” Earthq. Spectra, 26(4): 1117–1138.

Silva, W.J., N. Gregor, and R. Darragh. 2002. Development of Regional Hard Rock Attenuation Relations for Central and Eastern North America, Pacific Engineering and Analysis, El Cerrito, Calif., November; available at www.pacificengineering.org/CEUS/Development%20of%20Regional%20Hard_ABC.pdf.

Table 6–1 Summary of approaches used in selected PSHA projects.

Project	Approach	Resulting model attributes
TIP	Point estimates (M,R,F)	MECE for each point estimate
Yucca Mtn	Point estimates (M,R,F)	MECE for each point estimate
EPRI	GMM clusters Suite of GMMs within cluster	Merit weights for clusters MECE within cluster
PEGASOS	GMMs adjusted to site-specific conditions	Merit weights for GMMs
BCHYDRO	Suite of GMMs for Crustal Scaled backbone for subduction	For crustal, equal weight on selected models with additional epistemtic uncertainty. For subduction, MECE, including corrlation of magnitude scaling at large magnitude but ignored correlation of distance scaling
Canadian National Seismic Hazard Maps	Scaled backbone	MECE including correlation of magnitude and distance scaling
Hanford	Scaled backbone with alternative magnitude scaling	MECE including correlation of M scaling for crustal and magnitude and distance scaling for subduction
SWUS and NGA-East	Sampling GMMs from Sammon's maps, independent for each frequency	MECE including correlation of magnitude and distance scaling

Table 6–2 Ground-motion values for the 1D example. Median estimates of seed models at M = 6, R = 200, for ln[PSA(1 Hz)].

Model number	ln PSA(1 Hz, M = 6, R = 200 km)
1	-4.517
2	-4.834
3	-5.258
4	-5.433
5	-4.926
6	-5.164
7	-4.850
8	-5.362
9	-5.020
10	-5.230
11	-5.332
12	-5.331
13	-5.028
14	-5.416
15	-4.888
16	-4.958
17	-5.167
18	-5.198

Table 6–3 Ground-motion values for the 2D example. Median estimates of seed models for ln[PSA(1 Hz)], for M = 6, R = 200 km, and M = 8, R = 200 km.

Model number	ln PSA (1 Hz, M = 6, R = 200 km)	ln PSA (1 Hz, M = 8, R = 200 km)
1	-4.517	-2.056
2	-4.834	-2.961
3	-5.258	-2.587
4	-5.433	-3.220
5	-4.926	-2.203
6	-5.164	-2.351
7	-4.850	-2.492
8	-5.362	-2.982
9	-5.020	-2.609
10	-5.230	-2.327
11	-5.332	-2.551
12	-5.331	-3.131
13	-5.028	-3.313
14	-5.416	-3.228
15	-4.888	-2.470
16	-4.958	-2.735
17	-5.167	-3.0251
18	-5.198	-2.492

Figure 6–1 Concept of backbone approach where a single GMM (blue line) is scaled up and down to capture the epistemic range in ground motions (solid and dashed black lines). The ground motions (PGA in this case) are represented by vectors, the values of which are correlated through magnitude and distance scaling terms built into the GMM.

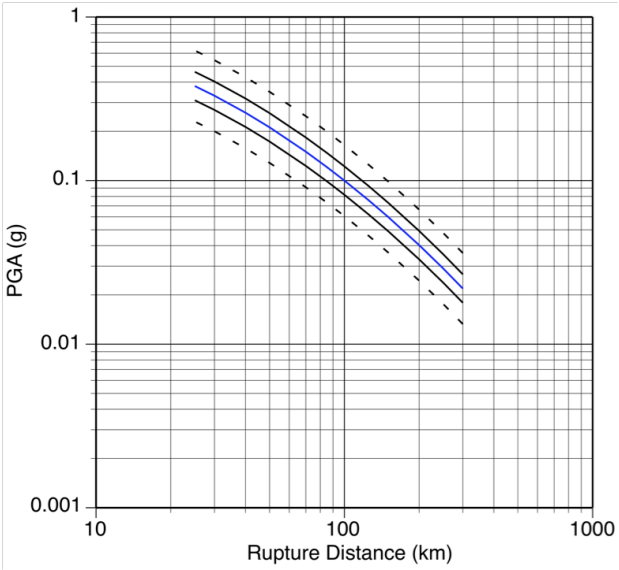


Figure 6–2 Illustration of epistemic uncertainty in distance scaling approaches. Different models allow a different correlation of PSA with distance (Addo et al. 2012).

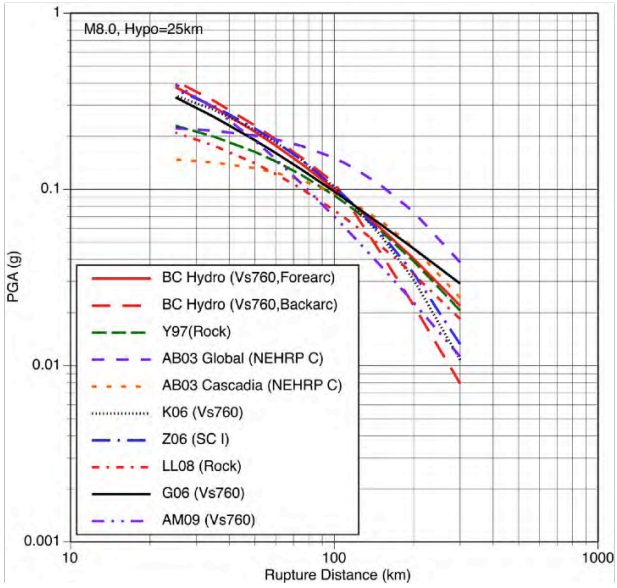


Figure 6–3 Median estimates (red symbols) of seed models at M = 6, R = 200, for PSA (1 Hz), with y in ln units of PSA (*g*). The blue curves are the normal distribution (PDF on top and CDF on bottom) fitted to the 18 median estimates from Table 6–2.

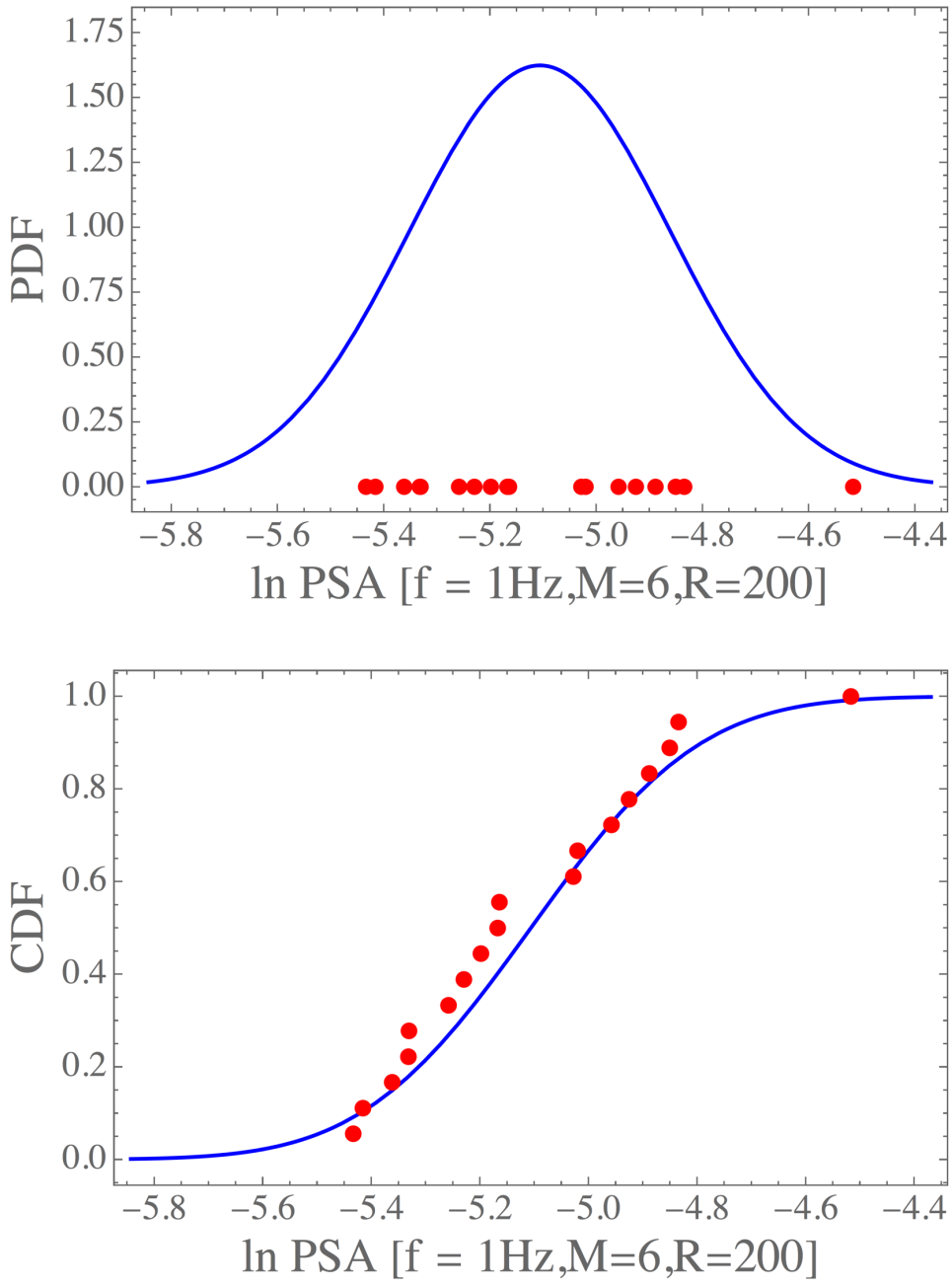


Figure 6–4 Histogram of 5000 models sampled from the distribution shown in Figure 6–3 (blue line).

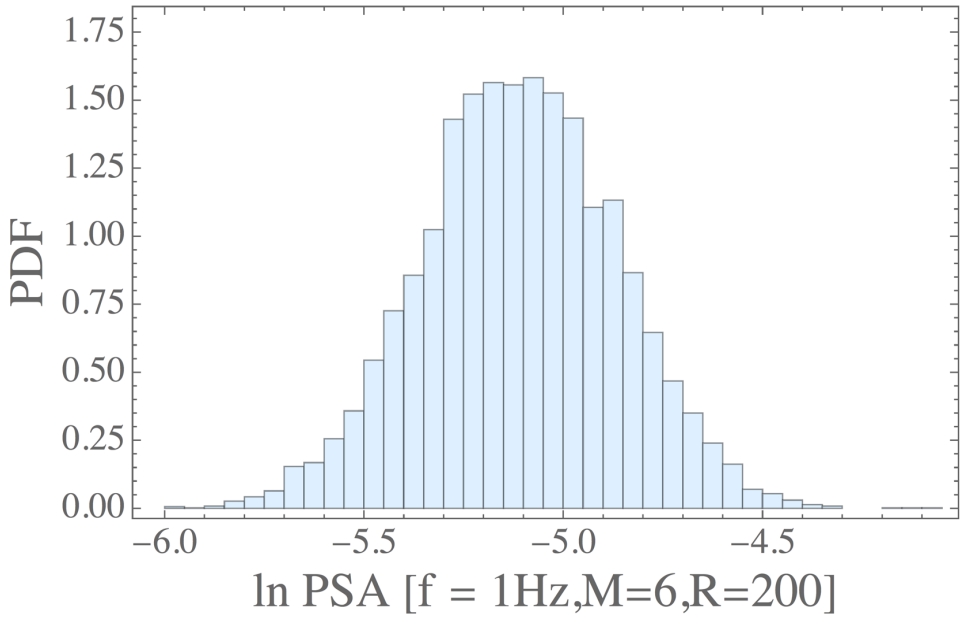


Figure 6–5 Partitioned one-dimensional ground-motion distribution based on the PDF shown in Figure 6-3 (blue lines) and histogram of 5000 sampled models.

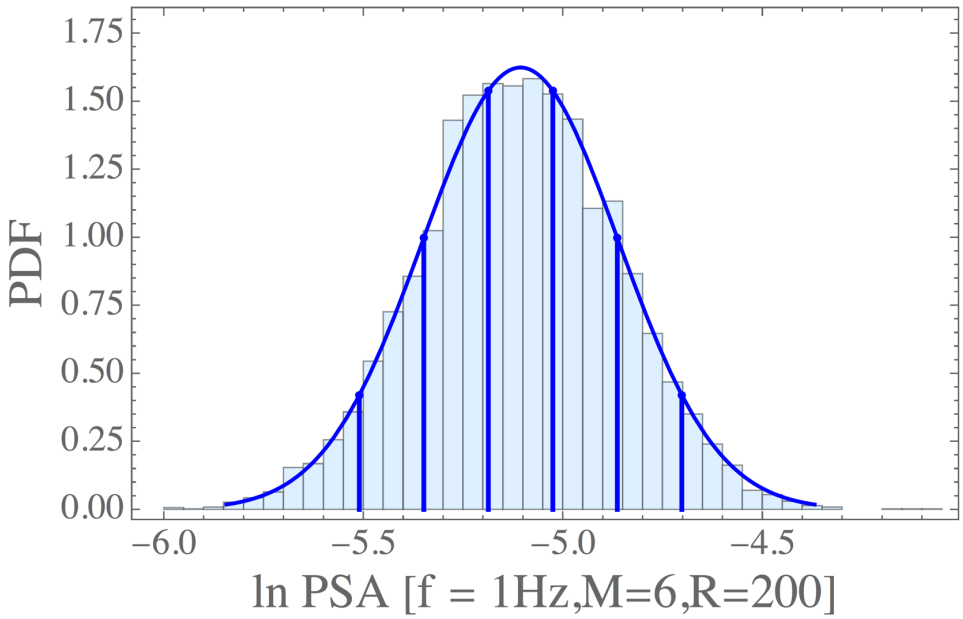


Figure 6–6 Partitioned one-dimensional ground-motion distribution and selected models for each bin (blue points), calculated as mean over samples in each bin.

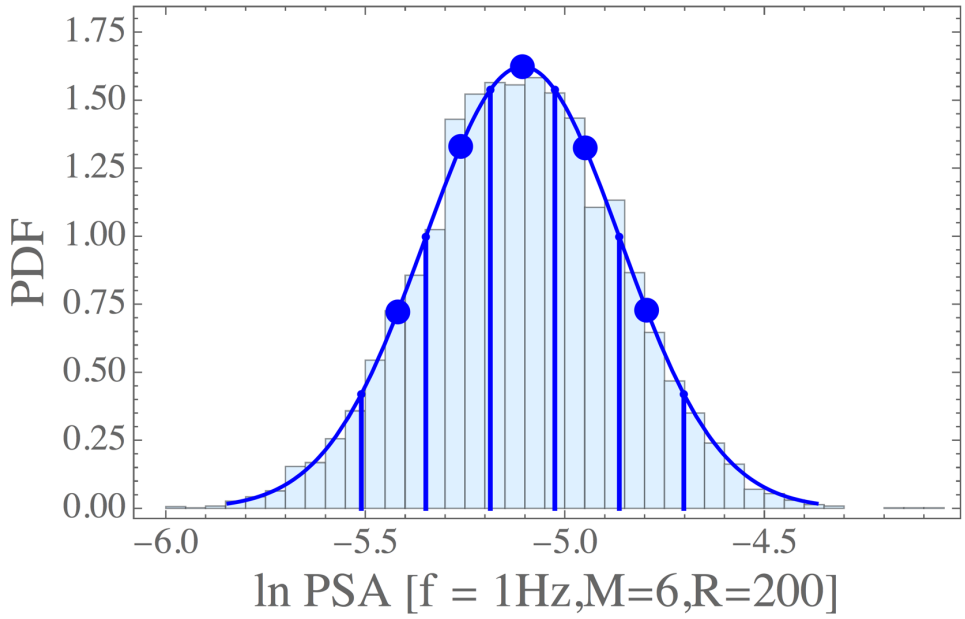


Figure 6–7 Partitioned one-dimensional ground-motion distribution and selected models for each bin (blue points), calculated as mean over samples in each bin. Green dots are 10 "data" points, which are used to evaluate the sampled models.

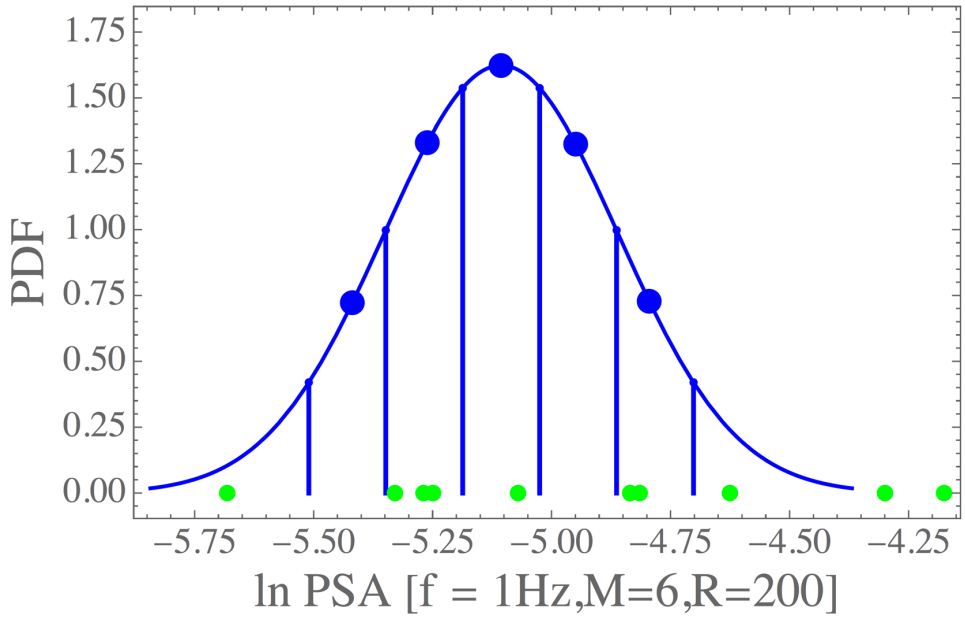


Figure 6–8 Top: mean residual of each sample vs. the sampled data value. Bottom: log-likelihood of data for each sample vs. sampled data value (top of the parabola showing the maximum likelihood). The blue lines show the location of the bin's edges from Figures 6–5 to 6–7.

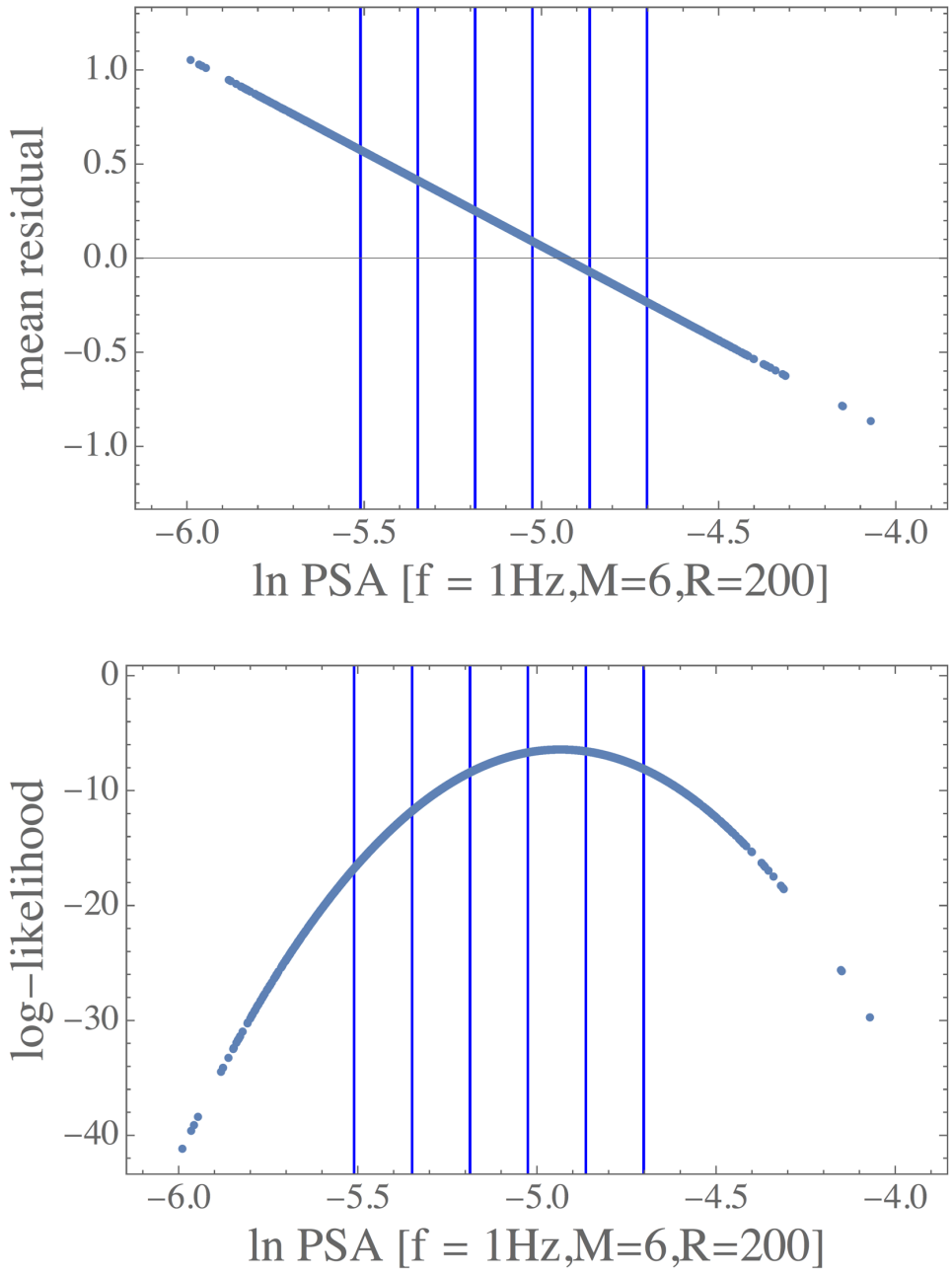


Figure 6–9 **Weights for the different bins, based on different approaches. Model index is from low to high y-values.**

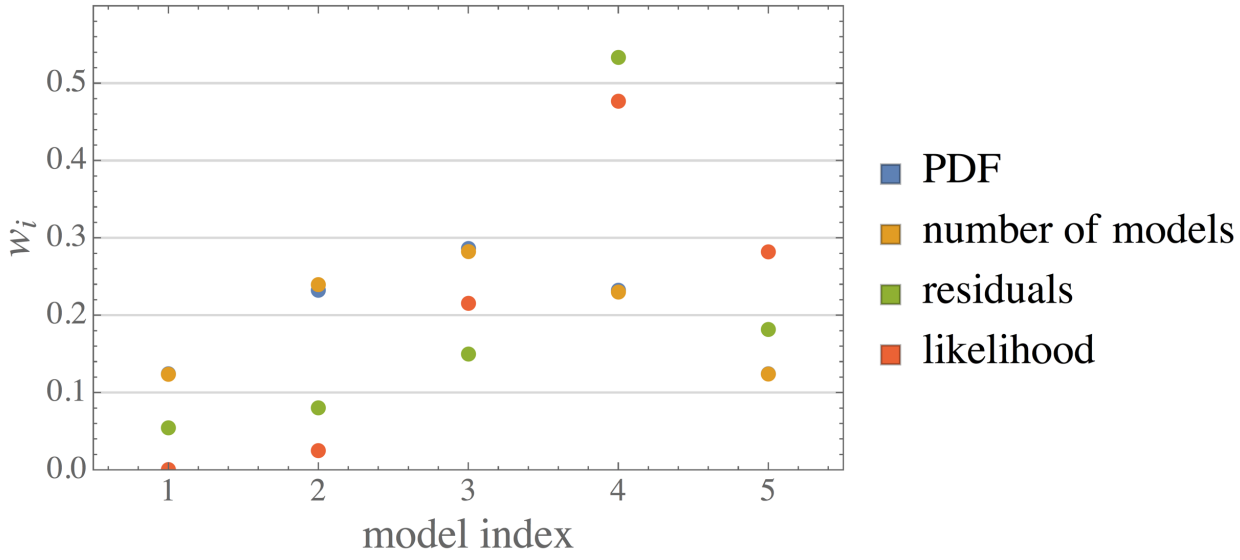


Figure 6–10 **Median estimates (red symbols) of seed models at M = 6, R = 200 and at M = 8, R = 200 for PSA (1 Hz), with axes in ln units of PSA (g).**

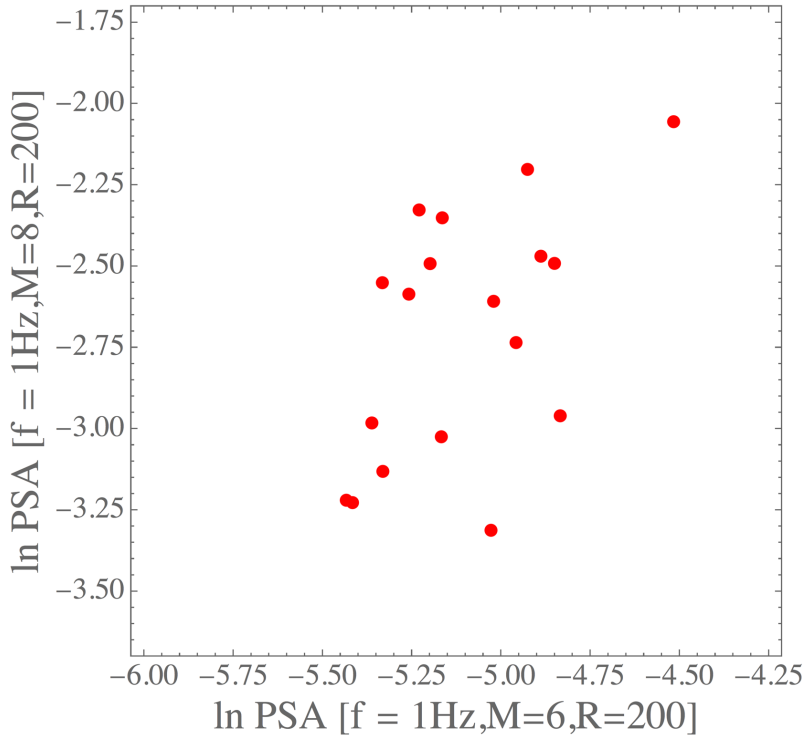


Figure 6–11 Median estimates (red symbols) of seed models at M = 6, R = 200 and at M = 8, R = 200 for PSA (1 Hz), with axes in ln units of PSA (g) and fitted bivariate normal distribution.

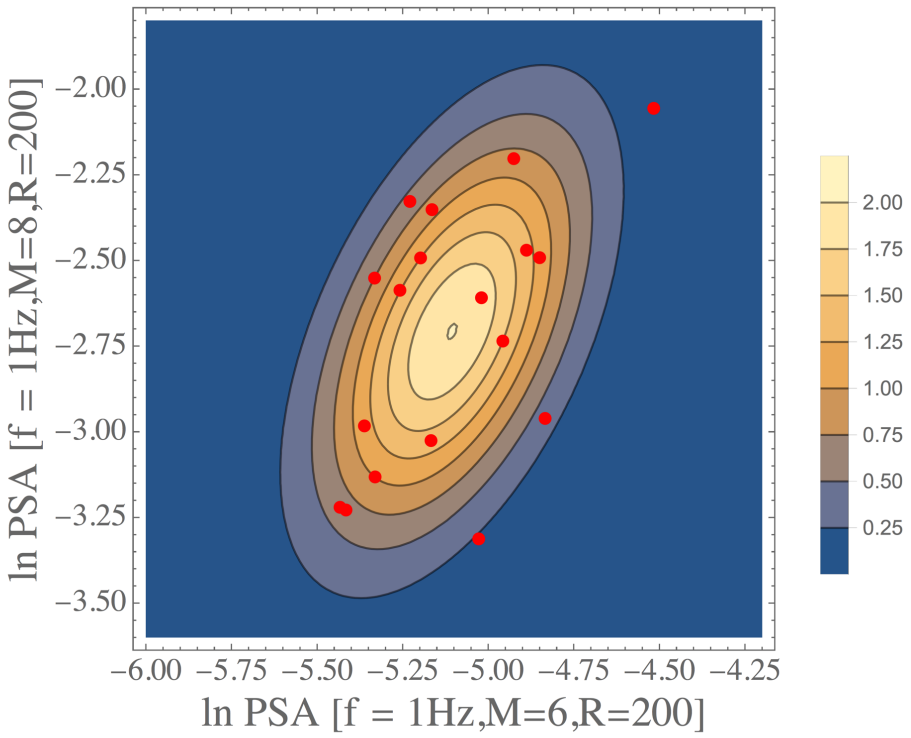


Figure 6–12 Histogram (top) and projected cloud of point (bottom) of 5000 models sampled from the distribution shown in Figure 6–11 (ellipses).

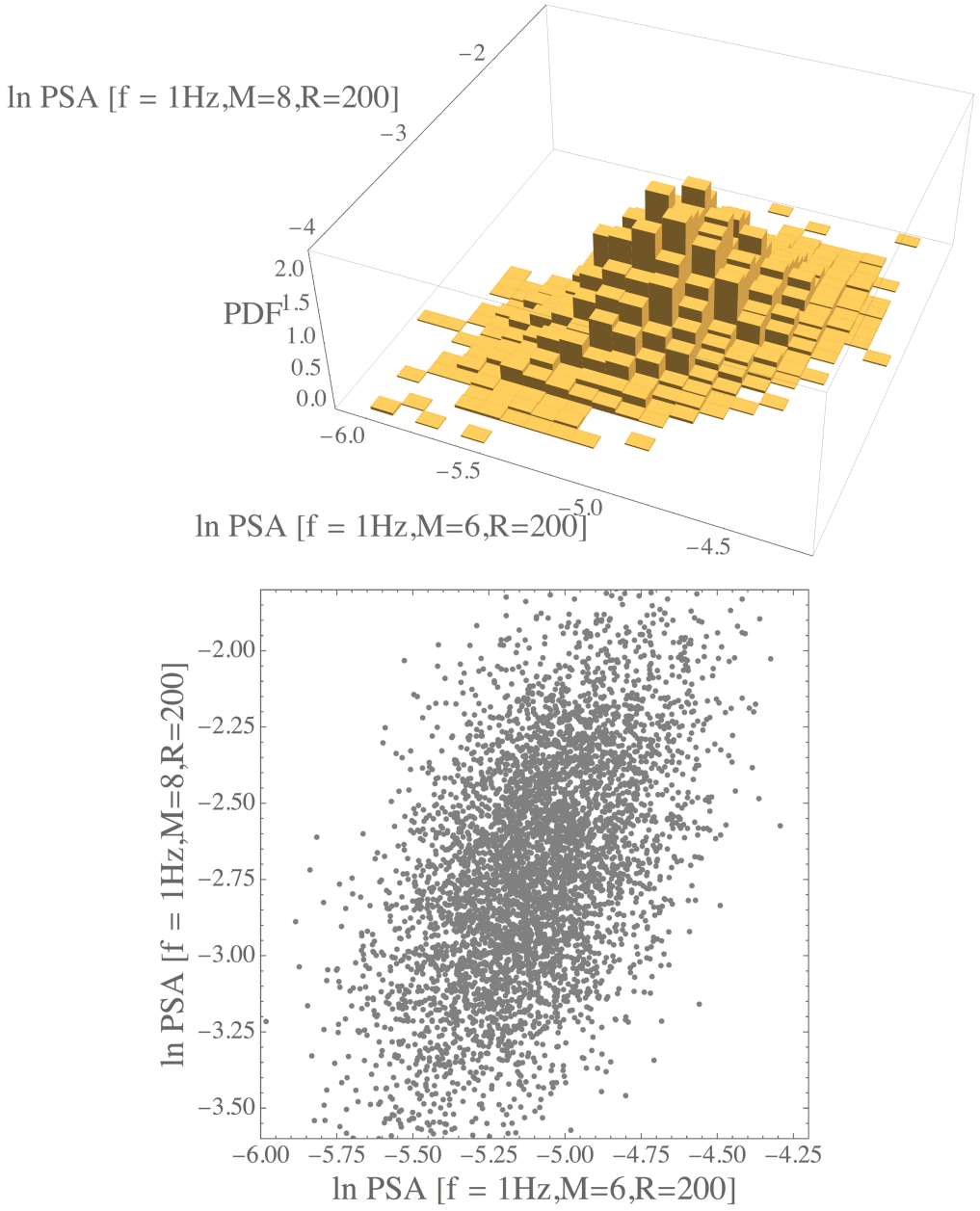


Figure 6–13 Partitioned two-dimensional ground-motion distribution: overlay on the 2D distribution (top) and cell numbers (bottom).

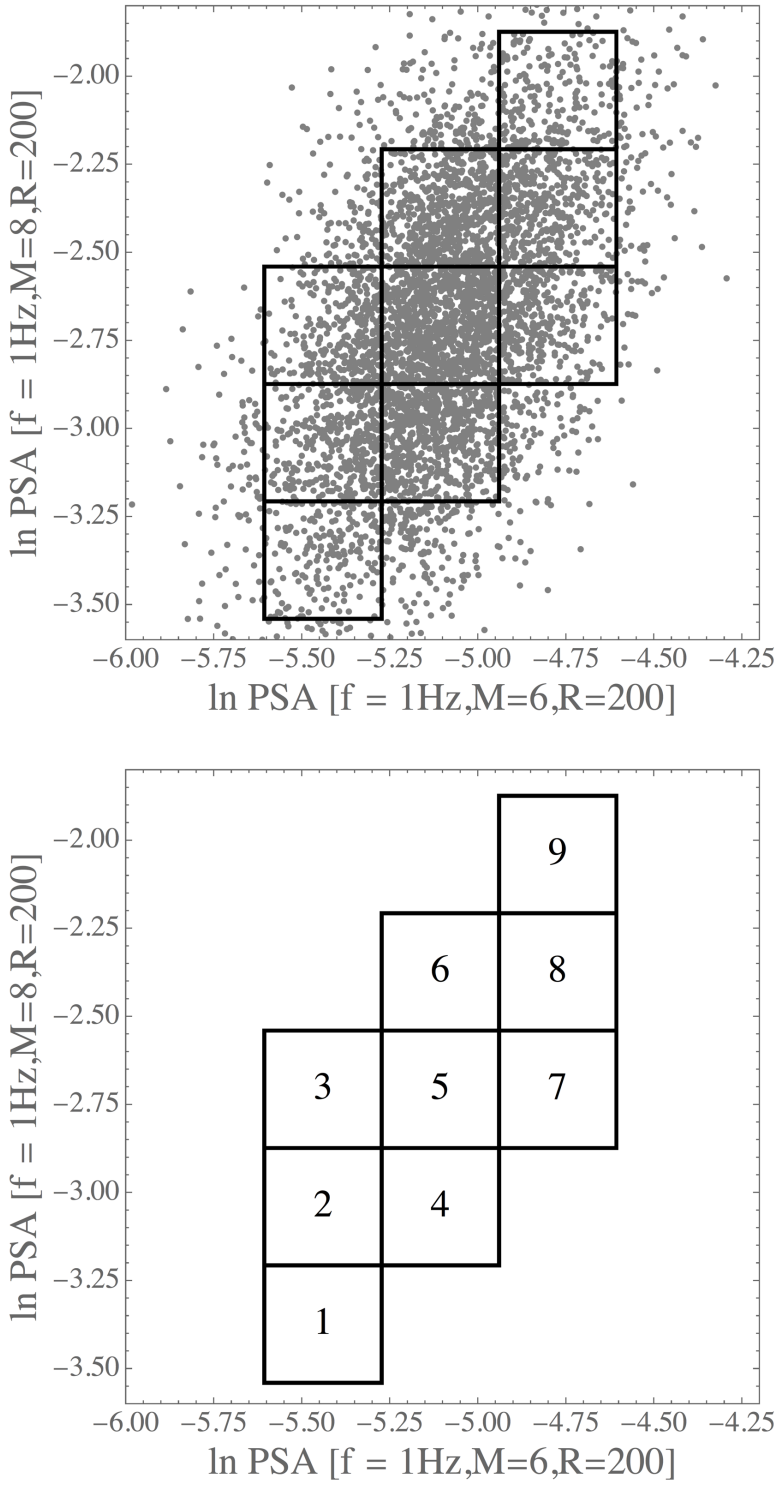


Figure 6–14 Partitioned one-dimensional ground-motion distribution and selected representative models for each bin (black points), calculated as the mean of all the samples samples in each bin. The red dot corresponds to the line in Figure 6–15.

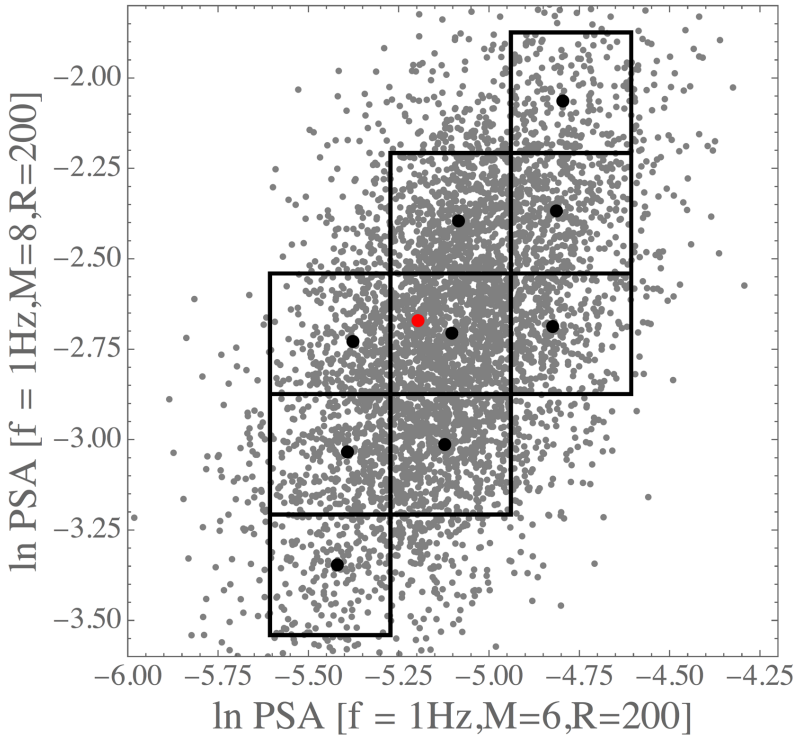


Figure 6–15 Simulated "data" to illustrate the computation of residuals and likelihood for the models shown in Figure 6–14.

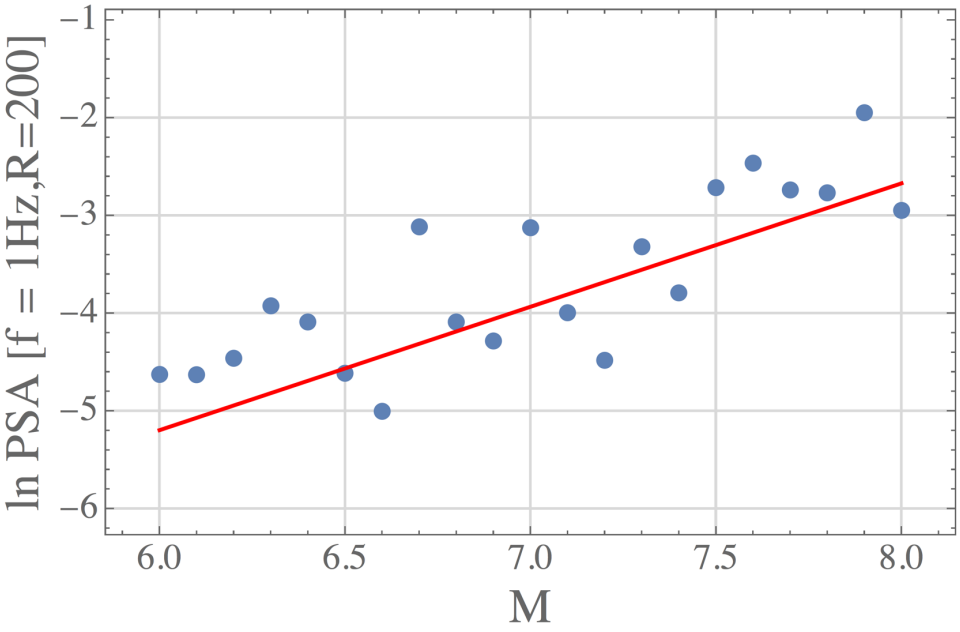


Figure 6–16 Top: mean residual of each sample vs. the sampled data value (the line between yellow and blue shades shows the zero residual value). Bottom: log-likelihood of data for each sample vs. sampled data value (pale beige top of the parabola showing the maximum likelihood). The axis is centered at the mean of the seed GMM distribution.

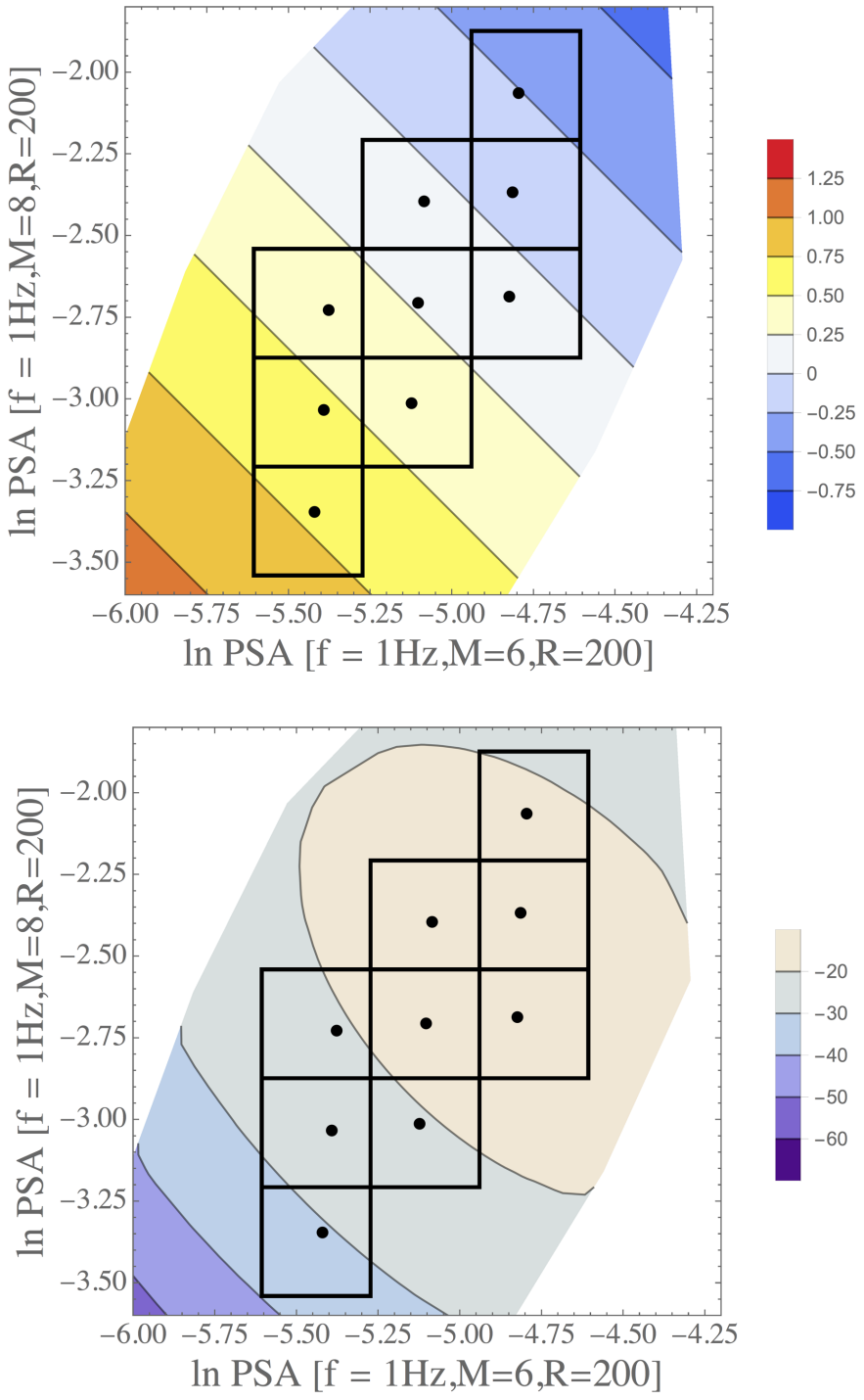


Figure 6–17 Weights for the different bins, based on different approaches. Model index correspond to Figure 6–13.

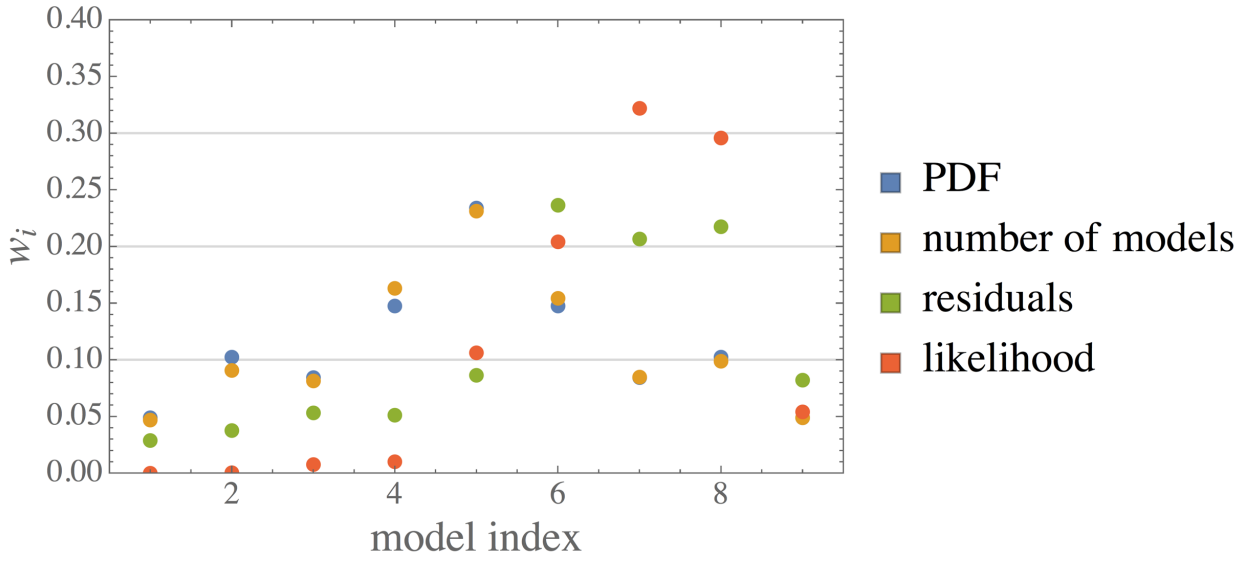
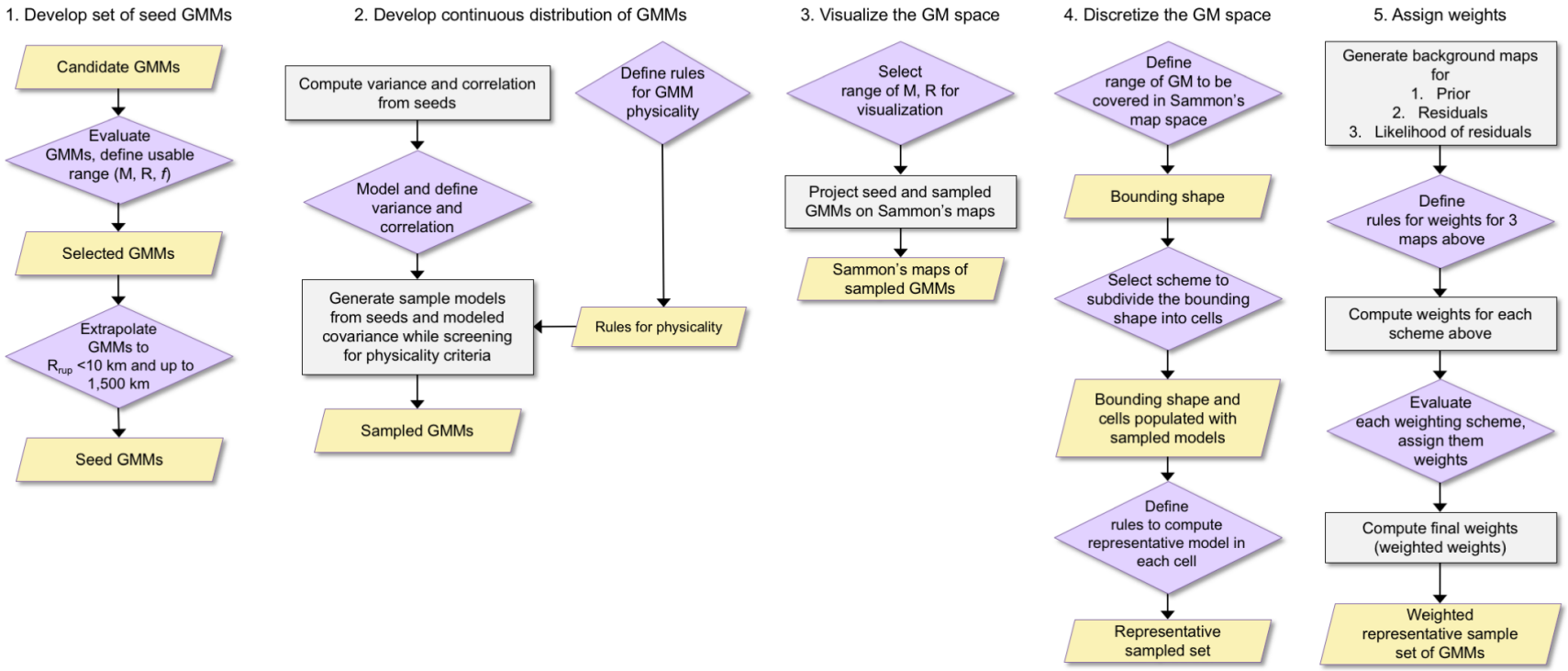


Figure 6–18 Flowchart summarizing the TI team approach for capturing the epistemic uncertainty in median ground motions. Step 1 is described in Chapter 7, Steps 2-4 are detailed in Chapter 8 and Step 5 is covered in Chapter 9.



7. Candidate Proponent Median Ground-Motion Models

The current chapter describes Step 1 of the NGA-East approach to quantifying the epistemic uncertainty in median ground motions (Figure 6–9). The five steps are repeated here for convenience:

1. Develop a suite of seed ground-motion models (GMMs)
2. Develop parameters for continuous distributions of GMMs
3. Visualize the ground-motion space and sample GMMs
4. Re-discretize the ground-motion space
5. Assign weights

More specifically, the chapter describes the candidate GMMs that were considered for use as seed models in developing the final median GMMs. The seed models form the basis for developing a representation of continuous distributions of GMMs, as outlined in Section 6.3. In this process, there are no explicit weights applied to the individual seed models. The "weighting" occurs implicitly as part of the selection of parameters to describe the continuous distributions, which is presented in detail in Chapter 8.

The first part of this chapter (Sections 7.1 to 7.6) provides a brief description of the individual candidate GMMs along with the criteria utilized to screen the candidate models and to determine the usable bandwidth of those models passing the screening criteria. Section 7.7 summarizes short- and large-distance modifications applied to the models to extend their applicability range; Section 7.8 covers the correction of pseudo-spectral acceleration (PSA) spectral shape applied to the extended models. These extended models define the seed models utilized in the median ground-motion characterization (GMC) development as discussed in Chapters 8.

7.1 Development of Candidate Median Ground-Motion Models for the NGA-East Project

A total of 30 median GMMs were initially considered as candidate models. Ten of the models are from the recent EPRI Review Project (EPRI 2013) and are listed in Table 7–1. These models are described in detail in EPRI (2013) and are not summarized here. Ultimately, none of the EPRI (2013) models was used in the development of the final GMMs; this was primarily because they have been updated and replaced by more recent models (see Section 7.3).

The remaining 20 candidate GMMs were developed specifically for the NGA-East project, and full descriptions of these models are provided in a PEER report (2015a). While the development process of these new GMMs involved numerous interactions between TI team and developers, the TI team as a team did not directly develop any candidate GMMs. Rather, the NGA-East Project solicited researchers to develop their own models and submit them for consideration.

In the development of the NGA-East Project plan, models solicited from the broader PEER community were historically described as "PEER" models and, a research group within PEER led by Justin Hollenbeck developed their own models and submitted those for consideration just as other developers did. For convenience, these models are referred to as PEER_EX and

PEER_GP, because all the participants were UC Berkeley employees, but the models should not to be considered official PEER Center models and PEER has not endorsed any specific model. In hindsight, those models should have been named HKGA for the initials of the developer team (Hollenback, Kuehn, Goulet, and Abrahamson). Nonetheless, since the models have already been described in existing PEER reports (PEER 2015a, b), it would be confusing for readers familiar with the project to change the names at this stage.

In addition, as part of the development process, the TI team recognized that numerous geometric spreading and Q models have been proposed for Central and Eastern North America (CENA) in recent years. In order to sample an appropriate a range of attenuation shapes that capture the center, body, and range of technically defensible models using a consistent methodology, the TI team asked Dave Boore to develop a suite of models using the SMSIM approach. A total of 56 models were originally identified from the literature, and from this a subset of six complete attenuation models was selected for consideration as candidate GMMs utilizing the SMSIM modeling approach. Details of this selection process of attenuation models are given in Appendix D. In this case as well, the models have been referred to as “Boore” models in PEER (2015a, b), but could have been named SMSIM models instead. These models were generated at the request of the TI team. Two of the attenuation models employ steep geometric attenuation ($R^{-1.3}$) at near-source distances. Dave Boore has expressed his personal view that these $R^{-1.3}$ models may not be the most appropriate for CENA. However, the TI team felt that since these published models cannot be invalidated by the available data, then they should be included as candidate GMMs in order to adequately sample the range of epistemic uncertainty.

The target conditions for GMM development were specified such that the models should attempt to cover the following conditions:

- Moment magnitudes of **M** = 4–8.2
- Distances of R = 0–1500 km
- Response frequencies of 0.1–100 Hz (plus PGA and PGV). A list of target frequencies was provided by the NGA-East Project (Chapter 1).
- Site condition of kappa = 0.006 sec and V_{S30} = 3000 m/sec (Campbell et al. 2014; Hashash et al. 2014a, b)

Modelers were provided with the complete NGA-East database (Goulet et al. 2014) and the forward finite-fault simulation results described in Chapter 5. The modelers, TI-Team members, and available members of the PPRP and JMC participated in weekly and bi-weekly phone calls extensively in the August 2014–December 2015 period of the GMM development. Additional interactions continued on an as-needed basis between the TI team and the modelers until the production of this report.

The NGA-East data are recorded on a variety of site conditions. The TI team did not require a specific approach for the treatment of site effects in the development of the candidate GMMs. Again, this decision was left to the individual modelers and is part of the epistemic uncertainty that these models are designed to capture. Brief descriptions of the site treatment approach used by each developer are given in the Section 7.2 summaries, with more complete descriptions given in the chapters of PEER Report (2015a).

Source depth was not formally specified to be part of the GMM development; as such, it was left to the individual modelers to decide how to handle this parameter. None of the seed GMMs contain an explicit source depth parameter. Adjustments of the final GMMs for source depth are discussed further in Chapter 13.

7.2 Description of Candidate Median Ground-Motion Models Developed for NGA-East

Table 7–2 lists the 20 GMMs developed for the NGA-East project with specific reference to the title, authorship, and chapter number in the PEER report (2015a). Table 7–2 also provides the acronym for each model used throughout the remainder of the current report. All GMMs are for footwall conditions, with adjustments for hanging-wall conditions developed in a separate task (Chapter 13). None of the models, in the form used here, includes a term for source depth; rather, they were designed to be applicable to the range of depths expected in Central and Eastern North America (CENA). A global depth-adjustment model was developed to be applied to the final NGA-East GMMs and is described in Chapter 13. In addition, the candidate GMM developers have focused on limiting their developed models to the Mid-Continent Region (MCR), i.e., the CENA region that excludes the Gulf Coast Region (GCR), as defined in Chapter 4. Adjustments to the models for the GCR were developed in a separate task, as elaborated in Chapter 13.

The following sections provide a brief summary of the key features for each of the candidate models. A series of tables (Tables 7–4 to 7–13, one table for each group of models) aggregate the basic model development information—including key assumptions—and range of applicability. In this and the following chapters, the following notations are used: the moment magnitude is denoted **M,** the Joyner-Boore distance (shortest distance to the horizontal projection of the rupture plane) is denoted R_{JB}, and the rupture distance (shortest distance to the rupture plane) is denoted R_{RUP}.

7.2.1 SMSIM (Boore) Models

(B_a04, B_ab14, B_ab95, B_bca10d, B_bs11, B_sgd02; see Table 7–4)

The six SMSIM (Boore) GMMs are documented in Chapter 2 of the PEER report (2015a), with the original tables in the corresponding electronic appendices (2B-2G) of the same report.

Modeling Approach and Key Features

The SMSIM (“Boore”) models were developed by Dave Boore at the request of the TI team. The goal was to sample a range of six published Q and geometric spreading models using a consistent simulation methodology, in this case the stochastic point source (PS) approach SMSIM (Boore 2003; 2005). The six attenuation models and their features are listed in Table 7–3, and a description of the TI team selection process for those models is provided in Appendix D.1. Two of the models (B_a04 and B_ab14) are characterized by a geometrical spreading of $1/R^{1.3}$ within the first 70 km and 50 km, respectively, whereas most of the other models have a decay of 1/R for these distance ranges. The simplest model is the B_bca10d one, which has 1/R geometrical spreading at all distances. Details regarding the derivation of the durations are in Boore and Thompson (2015). The attenuation-model-dependent stress parameters used in the stochastic-method simulations were derived from

inversion of PSA data from eight earthquakes in eastern North America (ENA). The simulations used a value of 0.55 for the average radiation pattern, shear-wave velocity and density in the source region of 3.7 km/sec and 2.8 g/cc, and a single-corner frequency (SCF), constant stress parameter model.

Modeling Constraints

Database. Data from nine earthquakes (**M** 4.4–6.8) were used for the stress parameter inversions. Only recordings from hard rock sites were considered, with an average V_{S30} of about 2000 m/sec as given by the NGA-East database. Inversions were run for two cases: R < 200 km and R < 600 km, with the final stress parameter results coming from the R < 200 km case in order to better constrain the closer distance behavior of the models.

Large Magnitude Extrapolation. Implicit in PS model.

Small distance extrapolation. Models were developed for R_{RUP} > 2 km. The TI team made adjustments (described in Section 7.7.3) to extend these models to R_{RUP} = 0 km.

Treatment of Site Effects. Only recordings from hard rock sites were considered with an average V_{S30} of about 2000 m/sec. In the inversions for stress parameter, a crustal amplification model with V_{S30} = 2000 m/sec was used in order to be consistent with the observations. The simulations performed for the final ground-motion models were done using a crustal amplification model with V_{S30} = 3000 m/sec. In both cases, kappa was set to 0.006 sec.

Applicability Range

The simulated motions are provided for distances between of 2 and 1200 km, **M** from 4–8, and 25 ground-motion intensity measures: peak ground velocity (PGV), peak ground acceleration (PGA), and 5%-damped pseudo-absolute response spectral acceleration (PSA) from 0.1 to 100 Hz.

7.2.2 Darragh, Abrahamson, Silva, and Gregor (DASG) Models
(1CCSP, 1CVSP, 2CCSP, 2CVSP; see Table 7–5)

The four GMMs are documented in Chapter 3 of the PEER report (2015a), with the original tables in the corresponding electronic appendices (3C–3F) of the same report.

Modeling Approach and Key Features

Darragh et al. utilized the PS stochastic simulation method to develop new GMMs for CENA. The employed PS model parameters including the stress parameter (both constant and variable models), source depth, bilinear geometrical spreading, path damping (Q), shallow crustal damping (κ), and crustal amplification. Their methodology is based on three steps: (1) the inversion of recorded data using selected assumptions for a subset of PS parameters; (2) generating earthquake simulations beyond the magnitude and distance range covered by the empirical data; and (3) parameterization of the simulations into GMMs.

Four new sets of GMMs are developed: (1) single-corner PS model with a constant stress parameter (referred to as 1CCSP model), (2) single-corner PS model with a variable stress parameter as a function of magnitude (1CVSP), (3) double-corner PS model with a constant

stress parameter (2CCSP), and (4) double-corner PS model with a variable stress parameter as a function of magnitude (2CVSP).

Modeling Constraints

Database. To perform the inversions, they used a subset of the NGA-East database including 53 earthquakes (10 potentially-induced events (PIES)) at 241 different sites for a total of 1133 recordings spanning the hypocentral distance range of about 10 km to 1000 km. They included the full range in site conditions using the PEER specified V_{S30} values for each recording site.

Large Magnitude Extrapolation. PS model.

Small distance extrapolation. PS model.

Treatment of Site Effects. Recording sites were grouped into National Earthquake Hazard Reduction Program (NEHRP) site categories based on their V_{S30} value. Elastic amplification correction factors were computed for each category using category-specific shear-wave velocity profiles originally developed for western North America (WNA) crustal structure and placed on top of a hard rock crustal model for application to CENA. The inversions resulted in kappa estimates of about 0.005 sec for each of the hard and firm rock site NEHRP categories A and B. Since these values were close to the typically adopted value of 0.006 sec, the latter was used in their simulations.

Applicability Range

The GMMs are developed for PGA, PGV, and 5%-damped PSA for frequencies 0.1 to 100 Hz for a magnitude range of **M** 4.5–8.5 at distances up to 1000 km.

7.2.3 Yenier and Atkinson (YA15) Model
(YA15; see Table 7–6)

The GMM is documented in Chapter 4 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (4A) of the same report.

Modeling Approach and Key Features

The YA15 model is based on the referenced empirical approach introduced by Atkinson (2008). It is similar to the hybrid empirical method in concept, but adjustment factors are determined empirically using spectral ratios of observed motions in the target region to predictions of an empirical ground-motion prediction equation (GMPE) in the host region.

The referenced empirical approach is an equivalent point-source simulation model whose parameters have been calibrated to empirical data in California in such a way as to determine the decoupled effects of basic source and attenuation parameters on ground-motion amplitudes. The generic GMPE is formulated as a function of earthquake magnitude, source to site distance, stress parameter, geometrical spreading rate, and anelastic attenuation coefficient.

Modeling Constraints

Database. For simulations, they perform time-domain equivalent point-source stochastic simulations using the widely-cited SMSIM software (Boore 2003; 2005), for **M** 3–8 and

distances from 1 km to 400 km, for a fixed stress parameter of $\Delta\sigma$ = 100 bar. For each simulated time series, PGA, PGV, and PSA were first calculated at 31 periods from 0.01 sec to 10 sec. Then, the geometric mean for each parameter over the 100 simulations was taken.

For calibration, they used the average orientation independent horizontal-component ground motions calculated based on the RotD50 measure (Boore 2010), as provided in the NGA-East flatfile for PGA, PGV, and 5%-damped PSA for CENA earthquakes of **M** ≥ 3.0 that were recorded by at least three stations within 600 km. Both natural and potentially induced earthquakes were considered. However, ground motions recorded in the GCR were excluded due to considerably different attenuation attributes in this region (Chapter 4 and 13).

Large Magnitude Extrapolation. PS model.

Small distance extrapolation. PS model.

Treatment of Site Effects. The ground-motion model was first developed for a reference condition of NEHRP B/C boundary with V_{S30} = 760 m/sec. Recordings are adjusted to the Vs30 = 760 m/s reference condition using the site effects model of Boore et al. (2014). The final model is derived by converting the ground-motion predictions for V_{S30} = 760 m/sec to a site condition of V_{S30} = 3000 m/sec using the site factors of Atkinson (2012).

Applicability Range

Their GMM is developed for average horizontal-component peak ground motions, and 5%-damped pseudo spectral acceleration (0.1 to 100 Hz, plus PGA), for magnitudes **M**3–8 and distance up to 600 km.

7.2.4 Pezeshk, Zandieh, Campbell, and Tavakoli (PZCT) Models *(PZCT15_M1SS, PZCT15_M2ES; see Table 7–7)*

The two GMMs are documented in Chapter 5 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (5A-5B) of the same report.

Modeling Approach and Key Features

Pezeshk et al. utilized a hybrid empirical method (HEM) to develop new GMMs for CENA that are based on the five latest NGA-West2 GMPEs (Bozorgnia et al. 2014). The HEM method calibrates an empirically-constrained GMPE in a data-rich host region (WNA) for use in a data-poor target region (CENA) based on adjustment factors obtained from response-spectral ratios of stochastic PS simulations in the host and target regions. PS inversions are limited to **M** < 6.0 in order to avoid finite-fault effects and to stay within the magnitude range of earthquakes in the NGA-East database. Seismological parameters are adopted from the most recent research and published information available for CENA. For WNA, the developers performed a set of point-source inversions to match the median NGA-West2 GMPEs for **M** ≤ 6.0, R_{RUP} ≤ 200 km, V_{S30} = 760 m/sec, strike–slip faulting, and sediment-depth parameters equal to the default values recommended by each of the NGA-West2 GMM developers.

Modeling Constraints

Database. They used a subset of the NGA-East database for comparison and calibration in which the earthquakes within the GCR and PIEs are excluded. They excluded NEHRP site class E (soft-soil) sites with V_{S30} > 180 m/sec because of their complex site-response characteristics and their potential for significant nonlinear site effects.

Large Magnitude Extrapolation. Two approaches are considered: (1) HEM-based adjustments are used for all magnitudes (referred to as the stochastic-scaling approach model 1 or M1SS), and (2) HEM adjustments are used for **M** ≤ 6.0 and the magnitude-scaling predicted by the NGA-West2 GMPEs is used for **M** > 6.0 (referred to as the empirical-scaling approach model 2 or M2ES)

Small distance extrapolation. Based on PS simulations and HEM adjustment factors.

Treatment of Site Effects. Recordings at sites with V_{S30} < 1500 m/sec are first corrected to a site condition of V_{S30}= 760 m/sec using the amplification factors from Boore et al (2014). Recordings at sites with V_{S30} > 1500 m/sec are assumed to be adequately represented with a value of V_{S30} = 2000 m/sec. Assuming representative velocity and density profiles for the V_{S30} conditions of 760 m/sec, 2000 m/sec, and 3000 m/sec, amplification functions are computed and applied to the data to adjust them to the CENA condition of V_{S30} = 3000 m/sec.

Applicability Range

The two CENA GMMs are derived for PGA and 5%-damped PSA for a frequency range of 0.1 to 100 Hz, a magnitude range of **M** 3.0–8.0, and R_{RUP} ≤ 1000 km. The style-of-faulting is not considered explicitly in the model. The GMMs are for a CENA reference hard-rock site condition defined by site parameters V_{S30} = 3000 m/sec and κ_0 = 0.006 sec (Campbell et al. 2014; Hashash et al. 2014a, b).

7.2.5 Frankel Model

(Frankel; see Table 7–8)

The GMM is documented in Chapter 6 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (6A) of the same report.

Modeling Approach and Key Features

The Frankel model is based on hybrid broadband finite-fault simulations assuming constant stress drop scaling. The broadband synthetics are derived from combining deterministic synthetics for a plane-layered velocity model at low frequencies with stochastic synthetics at high frequencies using matched filters. Finite-fault rupture models were developed using the approach of Frankel (2009). He used the southeastern Canada velocity and Q model of Hartzell et al. (1994) to construct the low-frequency synthetics. For the high-frequency stochastic synthetics, he used a geometrical spreading of R^{-1} out to 70 km, R^{0} from 70–130 km, and $R^{-0.5}$ for distances greater than 130 km, and a Q = 680 $f^{0.36}$ based on Atkinson and Boore (1995). Simulations were run for discrete magnitudes of **M** = 4.5, 5.5, 6.5, 7.5, and 8 for both vertical strike–slip faults and 45°dipping thrust faults, and rupture distances of 2–

1000 km. Simulations for all magnitudes were run with a depth to top of rupture (Z_{TOR}) value of 5 km. In addition, a subset of the **M** = 7.5 and 8.0 runs were also run with Z_{TOR} =1 km.

Modeling Constraints

Database. Limited comparisons were done using the 1 Hz and 5 Hz PSA values from recordings of ENA earthquakes, as provided in the NGA-East database (Goulet et al. 2014). Simulation results for **M** 4.5 were compared to the subset of observations for **M** 4.0 to **M** 5.0, and simulation results for **M** 5.5 were compared with the subset of observations for **M** 5.0 to **M** 5.8.

Large Magnitude Extrapolation. Implicit in simulation model.

Small distance extrapolation. Implicit in simulation model.

Treatment of Site Effects. Simulations are performed for a V_{S30} of 2.8 km/sec and a κ_0 of 0.006. The hard-rock shear-wave velocity profile used in the stochastic simulations is described in Frankel et al. (1996). For comparison of the model with recordings, only data from sites with V_{S30} > 760 m/sec were considered. No site corrections were applied for these comparisons.

Applicability Range

The GMM is developed for average horizontal-component peak ground motions, and 5%-damped PSA (0.1 to 100 Hz), for **M** 4.5 to **M** 8, rupture distance of 2–1000 km, and site condition of V_{S30} = 2.8 km/sec and κ_0 of 0.006.

7.2.6 Shahjouei and Pezeshk (SP15) Model

(SP15; see Table 7–9)

The GMM is documented in Chapter 7 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (7B) of the same report.

Modeling Approach and Key Features

The SP15 model utilizes the HEM to adjust ground-motion models for the host region (WNA) to make them applicable to the target region (CENA). The host region models are taken as the five GMPEs developed by PEER for NGA-West2. The regional adjustment factors are derived from synthetic seismograms computed for both regions using a hybrid broadband finite-fault simulation technique. In the broadband simulation procedure, the low-frequency portion of synthetics are obtained through a deterministic approach, implementing kinematic source models and the discrete wavenumber-finite element method for wave propagation using the program COMPSYN (Spudich and Xu 2003). The high-frequency portions are derived through a finite-fault stochastic simulation where the heterogeneous stress distribution over the fault—which is correlated to the slip distribution used in the low frequency portions— is used. Additionally, two alternative sets of parameters (such as geometrical spreading, anelastic attenuation, stress parameter, etc.) are considered, following the studies of Atkinson and Boore [2014] and Pezeshk et al. [2011]. The results from these two sets are equally weighted and incorporated in this study. Following Boore and Thompson (2015), site amplification in the synthetics was modeled using the quarter-wavelength impedance approach and kappa. The WNA site factors are: V_{s30} = 620 m/sec

and kappa = 0.04 sec; and the CENA factors are: V_{s30} = 3000 m/sec and kappa = 0.005 sec. The intensity measures of empirical GMMs were attained for the generic rock site of NEHRP B-C site condition with V_{S30} = 760 m/sec. This study used a generic style of faulting to evaluate the empirical ground motions, and the hanging-wall effect was excluded.

Modeling Constraints

Database. They compared their results with the NGA-East database (Goulet et al. 2014). The data from the GCR and PIEs were excluded. In addition, they used the data with quality flag of zero recorded at stations with V_{S30}= 180 m/sec.

Large Magnitude Extrapolation. Included in host GMPEs and simulations.

Small distance extrapolation. Included in host GMPEs and simulations.

Treatment of Site Effects. For data comparisons, intensity measures of the NGA-East database are adjusted using Boore and Thompson (2015) amplification factors to scale the motions to the reference rock site condition used in this study, V_{S30} = 3000 m/sec.

Applicability Range

The new GMPEs are developed for R_{JB} distances of 2–1000 km for the moment magnitude range of **M** 5–8, frequencies of 0.1–100 Hz and for the suggested generic hard-rock site condition with V_{S30}= 3000 m/sec (Hashash et al. 2014a, b) for CENA.

7.2.7 Al Noman and Cramer (ANC15) Model
(ANC15; see Table 7–10)

The GMM is documented in Chapter 8 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (8A) of the same report.

Modeling Approach and Key Features

The ANC15 empirical model is based on the two-stage regression approach of Joyner and Boore (1993; 1994). For the regressions, the observed intensities were converted to ground-motion estimates using the ground-motion intensity conversion equations of Ogweno and Cramer (2014). Geometrical spreading is modeled by a single term due to the limited observations at distances less than 50 km. They included a magnitude-dependent geometrical spreading term to help model magnitude saturation at large magnitudes.

Modeling Constraints

Database. The main source of ground-motion observations is the NGA-East ground-motion database, which is limited to observations from **M** < 6.0 earthquakes. Empirical observations were restricted to the MCR, avoiding the GCR for both earthquakes and recording stations. Observations from the 1976 **M** = 6.8 Gazli and 2001 **M** = 7.6 Bhuj earthquakes, plus intensities converted to ground-motion estimates by Ogweno and Cramer (2014) for **M** > 6 ENA historical earthquakes (1811–1812 New Madrid, 1886 Charleston, South Carolina, 1925 Charlevoix, and 1929 Grand Banks earthquakes) have been included in the development of the model.

Large Magnitude Extrapolation. Constrained by data and model form.

Small distance extrapolation. Constrained by data and model form.

Treatment of Site Effects. The site effects are modeled as a linear function of V_{S30} using the NGA-East values for recording sites. For the 2001 Bhuj observations, the V_{S30}estimates are based on geology information in Cramer and Kumar (2003) with Quaternary, Tertiary, and Older site geology represented as NEHRP site class D, C, and B mid-point V_{S30} values, respectively.

Applicability Range

The proposed GMM has observational coverage from **M** 2.5–7.7 over a distance range of less than 10 km to over 2000 km. The prediction equations are for peak ground acceleration, peak ground velocity, and 5% damped pseudo-absolute acceleration spectra at 21 frequencies between 0.1 to 10 Hz. The reference V_{S30} is 760 m/sec.

7.2.8 Graizer Model

(Graizer; see Table 7–11)

The GMM is documented in Chapter 9 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (9A) of the same report.

Modeling Approach and Key Features

Graizer followed a traditional empirical modeling approach based on the modular filter based approach developed by Graizer and Kalkan (2007; 2009). The functional form of the GMPE is derived from filters: each filter represents a particular physical phenomenon affecting the seismic wave radiation from the source. The number of predictors used in the model is limited to a few measurable parameters: moment magnitude (**M**), closest distance to fault rupture plane (R_{RUP}), average shear-wave velocity in the upper 30 m of the geological profile (V_{S30}) style of faulting, anelastic attenuation factor (Q_0), and if available basin depth defined as depth to the 1.5 km/sec shear-wave velocity isosurface ($Z_{1.5}$).

Modeling Constraints

Database. The model is based on the NGA-East database for the horizontal PGA and 5%-damped PS) RotD50 component (Goulet et al. 2014). A subset of 5026 data points from this database with **M** ≥ 3.75 and fault distances R_{RUP} ≤ 1000 km is used to constrain the model. The dataset includes 48 earthquakes from different regions in the CENA with no distinction made between the MCR and GCR. He also included six data points from the **M** 4.8 11/12/2014 Kansas earthquake recorded by high-quality strong-motion instruments at epicentral distances of 18 to 162 km.

Large Magnitude Extrapolation. Constrained by data and model form.

Small distance extrapolation. Constrained by data and model form.

Treatment of Site Effects. He developed site corrections based on multiple runs of different representative V_{S30} profiles through SHAKE-type equivalent-linear programs using time histories and random vibration theory approaches. Site amplification functions are calculated for different V_{S30} relative to hard rock definition used in nuclear industry (V_{S30} = 2800 m/sec).

Applicability Range

The model covers the range of magnitudes 4.0 < **M** <8.2, distances of 0 < R_{RUP}< 1000 km, S-wave velocities in the upper 30 m of 450 < V_{S30}< 2800 m/sec and frequency range of 0.1 to 100 Hz.

7.2.9 Hassani and Atkinson (HA15) Model
(HA15; see Table 7–12)

The GMM is documented in Chapter 10 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (10A) of the same report.

Modeling Approach and Key Features

The HA15 model is based on the referenced empirical approach of Atkinson (2008). This method is similar in concept to the hybrid empirical method. The difference is that the adjustment factors are purely empirical, estimated using the ratio of the observed regional ground motions in the target region to the predicted values for the host region. The main assumption of this method is that the magnitude scaling and overall near-source behavior of ground motions are the same in the host and target regions, although the overall ground-motion levels at the source might be different (if the source parameters differ), and the distance attenuation might be different. The technique is based on the use of residual analysis that models differences between regional ground-motion observations and a reference GMPE developed for a data-rich region. They used the reference model of Boore et al. (2014; denoted BSSA14) as the reference GMPE which was developed as part of the NGA-West2 Project (Bozorgnia et al. 2014).

Modeling Constraints

Database. The considered database consists of PGA, PGV, and 5%-damped PSA at frequencies from 0.1 to 100 Hz for RotD50 of ground motions. They considered records with R_{epi} ≤ 400 km, to be consistent with the BSSA14 model development (Boore et al. 2014).

Large Magnitude Extrapolation. Used scaling from host region model (WNA – BSSA2014).

Small distance extrapolation. Used scaling from host region model (WNA – BSSA2014).

Treatment of Site Effects. Site amplification effects are incorporated using Vs30 and following the format of the reference BSSA14 GMPE. V_{s30} information is extracted from the updated NGA-East database (Goulet et al. 2014).

Applicability Range

The model is developed for PGA, PGV, and PSA at frequencies from 0.1 to 100 Hz for horizontal components of ground motions, and is applicable for **M** = 3–8.5, 0 ≤ RJB ≤ 400 km and V_{S30} = 3000 m/sec.

7.2.10 Hollenback, Kuehn, Goulet, and Abrahamson (PEER) Models
(PEER_EX, PEER_GP; see Table 7–13)

The two GMMs are documented in Chapter 11 of the PEER report (2015a), with the original tables in the corresponding electronic appendix (11A-11B) of the same report.

Modeling Approach and Key Features

The Hollenback et al. team followed the Bora et al. (2014) approach in which a model is first developed in acceleration Fourier Amplitude Spectral (FAS) space; then random vibration theory (RVT) is used to obtain the final PSA, PGA, and PGV quantities. The approach consists of the following general steps: (1) perform an empirical regression on FAS of acceleration in ranges for which the recorded data is reliable; (2) extend the empirical model (in terms of magnitude, distance and frequency coverage) using a combination of point-source (PS) stochastic models and finite-fault (FF) simulations; (3) develop an empirically-calibrated RVT-duration (DRVT) model; and (4) compute GMIMs for a wide range of magnitude and distances using RVT. Two GMMs were developed using alternative finite-fault simulation approaches to constrain the magnitude scaling beyond **M** 5.5. One approach uses EXSIM (Atkinson and Assatourians 2015), and the other uses Graves and Pitarka (2015).

Modeling Constraints

Database. The main dataset used for the GMM development is the NGA-East database (Goulet et al. 2014). They chose to limit the events to only those with **M** > 2.5 and records to closer than 300 km rupture distance (R_{RUP}), with at least one recording per event within 200 km. They have included events flagged as PIE in the selected dataset. The model is developed for path region 2 (Central North America, CNA), a subset of the MCR. Data from other CENA regions were also used to better constrain the regionally varying parameters. To overcome the problem of sparse recorded data available in CENA, a subset of the NGA-West2 database (Ancheta et al. 2014) was included in the analysis, covering a broad range of **M** and R_{RUP}. The various datasets included a regionally dependent statistical analysis so as not to map attenuation and site effects from WNA into CENA, in order to only retain the magnitude scaling effects.

Large Magnitude Extrapolation. NGA-West2 observations and finite-fault simulations.

Small distance extrapolation. Constrained by model form and simulations.

Treatment of Site Effects. The model contains a site term based on V_{S30}, and the estimated V_{S30} values for the recording site were used in the model regressions. The reference condition for the mean model is V_{S30} = 760 m/sec and kappa = 0.025 sec. Adjustment factors for kappa (Al Atik et al. 2014) and V_{S30} (Boore and Thompson 2015) are provided to attain the target CENA values of kappa = 0.006 sec and V_{S30} = 3000 m/sec.

Applicability Range

The GMMs provide estimates of RotD50 5%-damped PSA for frequencies 0.1–100 Hz, as well as PGA and PGV from events in the magnitude (**M**) range of 4 to 8.2 within a distance of 1200 km. The models are applicable to the large MCR (excluding the GCR).

7.3 Criteria for Selecting Candidate Ground-Motion Models

The TI team evaluation of candidate models was essentially completed in two passes. The first pass involved an initial screening of existing GMMs (Section 7.4), and the second pass involved a more in-depth evaluation of the retained GMMs (Section 7.5 and Section 7.6).

The TI team developed a suite of criteria for the evaluation process, and a candidate GMM was excluded from further consideration if it met one or more of the following conditions:

1. The model has been superseded by a more recent model (as confirmed by the model developer).
2. The model is more than 10 years old, unless the lead developer(s) can provide a compelling reason for its inclusion as a candidate model.
3. The model cannot cover, or be reasonably extrapolated to cover, the **M** = 4-8.2, R = 0–1500 km ranges.
4. The model cannot cover, or be reasonably interpolated to cover, the required range of frequencies (0.1 to 100 Hz plus PGA).
5. The model is not based on applicable data or utilizes data that is too uncertain to be diagnostic.
6. The model exhibits magnitude (**M**), distance and/or frequency scaling that appears unphysical or is inconsistent with the applicable data.

Conditions (1) to (4) were used primarily for the initial screening phase (Section 7.4), and conditions (5) and (6) were used primarily for the second phase (Sections 7.5 and 7.6). We note that several of the developed GMMs did not cover the entire range of distances and/or frequencies as specified in (3) and (4); however, these models were extrapolated to cover these ranges as described in Sections 7.7 and 7.8.

The use of data residuals was not employed as an explicit criterion in the evaluation process. The primary reasons for this are: (1) the available data only cover a very limited range of the magnitudes, distances and frequencies required by the GMMs; and (2) the computation of data residuals requires the application of site response corrections, which contain a large degree of uncertainty and variability. However, the TI team did perform consistency checks using the data residuals for each of the GMMs that passed the two evaluation phases, and results of these analyses were presented at Workshop 3B and 3C in March and June 2015 respectively. This consistency check using residual analysis is described in Section 7.6.

7.4 Initial Screening of Candidate Median Ground-Motion Models

The initial screening removed from consideration the ten GMMs from the EPRI (2013) Review Project (Table 7–1). The primary reason for excluding these models was that the models have been superseded by more recent versions; this applies to all eight GMMs in clusters 1, 2, and 3. The remaining two EPRI models (Somerville et al. 2001) were excluded because they had not been updated from their original form. The lead developer (Somerville) advised the TI team that there were no plans to update these models in a time frame that would work for the NGA-East project. In addition, one of the NGA-East GMM (PEER_GP) includes simulations from the method developed for the Somerville et al. (2001) GMMs, capturing some of their important features.

7.5 Systematic Evaluation of Remaining Candidate Median Ground-Motion Models

For the remaining 20 GMMs that passed the initial screening process, the TI team performed a more detailed and systematic analysis of the models to examine their behavior over a range of magnitudes, distances, and frequencies of interest to the NGA-East project. This phase of the assessment focused on criteria (5) and (6) as listed in the previous section. Again, the objective of this process was not to provide a ranking of individual models, but rather to ensure that all the selected models were grounded in physically sound and defendable principles for the region of interest (CENA), and that the data and methodologies used to develop the models were appropriate, reproducible, and reliable. A key consideration was that all retained models need to be suitable for the specific purposes of this project.

7.5.1 Evaluation of Technical Bases and Range of Applicability

With 20 GMMs and such a wide range of magnitude, distance, and frequencies to consider, the TI team focused the primary screening analysis on the magnitude range **M** = 4.5–7.5, distances R < 200 km, and frequency range 0.1 < f < 100 Hz. The rationale behind this was as follows: if a model exhibited systematically unjustifiable behavior (i.e., appears unphysical or is inconsistent with the applicable data) across this subset of key magnitudes and distances that was deemed unacceptable, it was likely the model would exhibit this behavior across other magnitudes and distances. Furthermore, throughout the entire model-building process, the behavior of the selected GMMs was continually checked to make sure that the results provided by these models were appropriate, understandable, and defendable. A series of plotting tools were created at various stages of the process to investigate the scaling trends of the various GMMs. These plotting tools were used to develop intermediate products that the members of the TI team could view and adjust interactively and included plots of all candidate GMMs showing magnitude scaling, distance scaling and spectral shape.

Additionally, this stage of the screening process determined the range of frequencies over which a candidate model could be used as a seed model across the entire range of **M** and R values. To ensure a practical, efficient and consistent model-building process, the TI team excluded models that could not cover, or be reasonably extrapolated to cover the required **M** and/or R ranges. This constraint ensures a consistent number of (**M**, R) scenarios at each frequency, which simplifies the process described in Chapter 8. As the TI team went through this evaluation process, there were numerous interactions with the developer teams to make sure that the features the TI team identified in the GMMs were interpreted correctly. Furthermore, the TI team discussed with each developer team their decisions to exclude a GMM, or to use it only over a subset of frequencies, to ensure the developers understood the basis for the decision and were comfortable with the outcome.

There are two sets of figures described in the following sections. Both sets of figures show 5%-damped PSA for the 20 candidate GMMs as a function of oscillator frequency for the magnitude range **M** = 4.5–7.5 and distances R < 200 km.

The first set of figures (Figures 7–1 to 7–16) shows all 20 candidate GMMs on each panel. The plots are divided into four magnitude bins: **M** 4.5 (Figures 7–1 to 7–4), **M** 5.5 (Figures 7–5 to 7–8), **M** 6.5 (Figures 7–9 to 7–12), and **M** 7.5 (Figures 7–13 to 7–16). Furthermore, within each

magnitude bin, there are separate plots for distances of R = 20 km, 50 km, 100 km, and 200 km. Examination of these plots provided insight on which GMMs exhibited behavior that warranted further investigation.

Features in Figures 7–1 to 7–16 identified by the TI team are briefly summarized here:

- GMMs 1CCSP and 1CVSP show elevated amplitudes at low frequencies (f < 1 Hz) for magnitudes of **M** 6.5 and larger (e.g., Figure 7-10, discussed further in Section 7.5.2.2).
- GMM YA15 shows a relative drop in ground-motion level at 50 Hz and higher (e.g., Figure 7-7, discussed further in Section 7.5.2.3)
- GMM Frankel shows somewhat jagged spectral shape particularly for magnitudes of **M** 5.5 and smaller (e.g., Figure 7–7, discussed further in Section 7.5.2.5).
- GMM ANC15 shows somewhat jagged spectral shape and relatively small magnitude scaling at distances of 50 km and less in the frequency range 1 to 10 Hz (e.g., Figure 7-13, discussed further in Section 7.5.2.7).
- GMM Graizer shows relatively low ground-motion levels for frequencies higher than 5 Hz and also for frequencies lower than 0.2 Hz (e.g., Figure 7–15, discussed further in section 7.5.2.8).
- GMM HA15 shows a relatively strong spectral peak for **M** 5.5 (e.g., Figure 7–8, discussed further in Section 7.5.2.9).
- GMM PEER_EX shows relatively large ground-motion levels for frequencies of 2 Hz and lower particular for magnitudes of **M** 6.5 and larger at distances of 100 km and larger (e.g., Figure 7–16, discussed further in Section 7.5.2.10).

Again, we note that the fact that a particular GMM exhibits behavior that is different from the other GMMs does not in and of itself mean the GMM should not to be trusted or used.

The second set of plots (Figures 7–17 to 7–36) further illustrate the key features of the models that were identified during the interactive screening process. The format of these figures is the same for each GMM. Each figure has four panels, one each for 20 km, 50 km, 100 km, and 200 km distance. The distance metric used is the same as provided by each model development team, without any conversion at this point (R_{JB} or R_{RUP}). Additionally, curves in each distance panel are plotted for magnitudes: **M** = 4.5, 5.5, 6.5, and 7.5. The red curves show the response of the individual GMM under consideration, and the grey curves show the response determined by averaging across all 20 GMMs. The grey curves are included to provide a smooth, common reference from which to compare each of the individual GMMs.

The TI team evaluated the GMMs' responses shown in Figures 7–1 to 7–36 to determine the frequency range over which each candidate model could be reliably used. In making these evaluations, a higher level of importance was given to the spectral shape than to the absolute level of the response. The reason being that the variation in the absolute level across the GMMs can be rather large, especially for the larger magnitudes, whereas systematic deviations in the spectral shape may point to inconsistencies of the GMM with respect to the data. Additionally,

the TI team considered whether features seen in the relative shifts of the spectra across the ranges of magnitude and distance behave in a physically consistent and defendable manner.

7.5.2 Selection of Median Candidate Ground-Motion Models

The majority of the candidate GMMs exhibit reasonable behavior over the complete range of magnitudes, distances, and frequencies considered in this stage of the evaluation process (**M** = 4.5–7.5, R < 200 km, and 0.1 < f < 100 Hz). Based on the TI-Team evaluations, 15 of the 20 candidate GMMs were selected to be "accepted as-is" over the full frequency range (0.1 < f < 100 Hz). Four candidate GMMs were selected for use over limited ranges of frequencies, and one GMM was not selected for use as a seed model as summarized in Section 7.6 and Table 7–14.

Brief summaries for each GMM group ("groups" as defined in Table 7–2) are provided in the following sub-sections. In addition, Table 7–14 summarizes the TI team evaluations for each of the 20 candidate GMMs.

7.5.2.1 SMSIM (Boore) Models
(B_a04, B_ab14, B_ab95, B_bca10d, B_bs11, B_sgd02)

The ground-motion spectra for the six SMSIM GMMs are shown in Figures 7–17 to 7–22. Since these models share several features, the computed spectra also show some similarities, e.g., they all produce relatively smooth spectra and do not exhibit any strong breaks in scaling as a function of distance or magnitude. Nonetheless, the differences in geometric spreading and Q used in these models do produce some noticeable differences. Note the tendency for the $R^{-1.3}$ models (B_ab04 and B_ab14) to produce relatively large motions at close distances (see 20-km panels in Figures 7–17 and 7–18) as compared to the other models. Dave Boore has expressed his personal view that these $R^{-1.3}$ models may not be the most appropriate for CENA. The TI team discussed this issue with Dave Boore and reached the agreement that since these are published models and they cannot be strictly invalidated by the limited available data, then they should be included as candidate GMMs in order to adequately sample the range of epistemic uncertainty. Also, the differences in Q can have a strong impact for the higher frequencies at larger distances, as seen in comparing the relatively low high-frequency motions at R = 200 km for B_bs11 (Figure 7–21) with the relatively large high-frequency motions at the same distance for B_sgd02 (Figure 7–22). All of these models are based on previously published geometric spreading and Q models, and thus provide a sampling of the epistemic uncertainty in these parameters. Based on the evaluations of the TI team, all six of the SMSIM-based GMMs were accepted without restriction for use as seed models.

7.5.2.2 Darragh, Abrahamson, Silva, and Gregor (DASG) Models
(1CCSP, 1CVSP, 2CCSP, 2CVSP)

The ground-motion spectra for the four DASG GMMs are shown in Figures 7–23 to 7–26. The primary difference among these models is the use of either a "single corner" (Figures 7–23 and 7–24) or "double corner" (Figures 7–25 and 7–26) for the stochastic point-source spectrum. The single-corner models (both constant and variable stress parameter cases) tend to produce elevated levels of low-frequency (f < 1 Hz) ground motions, particularly for larger magnitudes (**M** > 6), which lead to a noticeable bump in the spectra. The developers have noted a similar

feature for WNA single-corner models when compared against recorded ground motions. Their recommendation is to use the single-corner models only for frequencies above 1 Hz, and the TI team agrees with this recommendation. The double-corner models do not exhibit this feature, and the TI team accepted them for use as seed models without any restrictions.

7.5.2.3 Yenier and Atkinson (YA15) Model

The ground-motion spectra for the YA15 GMM are shown in Figure 7–27. This model generally produces relatively smooth spectra and does not exhibit any strong breaks in scaling as a function of distance or magnitude. The main feature identified in the evaluation process is the relative drop in ground-motion level around 50 Hz accompanied by a relatively flat spectrum extending to 100 Hz. The relative reduction in the higher frequency motions is related to the site correction factors used to adjust the final model from V_{S30} = 760 m/sec to V_{S30} = 3000 m/sec (Atkinson 2012). These factors are not well constrained for frequencies above 40 Hz due to the lack of empirical data. While there is larger uncertainty in the site corrections for f > 40 Hz, this was not viewed as a compelling reason to reject the model in this bandwidth. This decision was discussed with and approved by the developers. The TI team does not consider this to be a significant issue and accepted this model for use without any restrictions.

7.5.2.4 Pezeshk, Zandieh, Campbell, and Tavakoli (PZCT) Models *(PZCT15_M1SS, PZCT15_M2ES)*

The ground-motion spectra for the PZCT GMMs are shown in Figures 7–28 and 7–29. These models produce relatively similar spectra across the range of magnitudes and distances considered for the evaluation. The main difference in the two models is the stronger high-frequency scaling for PZCT15_M2ES (Figure 7–29) for distances less than about 50 km. This difference is considered part of the epistemic uncertainty (the PZCT_M1SS large magnitude scaling is based on simulations and PZCT_M2ES is based on WUS data), and these models were accepted for use without any restrictions.

7.5.2.5 Frankel Model

The ground-motion spectra for the Frankel GMM are shown in Figure 7–30. This model produces spectra that are relatively jagged, particularly at the lower frequencies. This jaggedness primarily results from the limited number of realizations that were used in simulating the ground motions used to constrain the model. In addition, the use of full waveform Green's functions for the lower frequency portion of this model also contributes to the characteristics of the spectral shape, e.g., the elevation in lower frequency (f < 0.5 Hz) motions in the distance range 50–100 km is likely due to surface wave energy carried by these Green's functions. This feature is more prominent for the smaller magnitudes (which have spatially compact sources) and becomes washed-out at larger magnitudes due to the much larger faults. The TI team did not consider any of these observations to be a significant issue and accepted this model for use without any restrictions.

7.5.2.6 Shahjouei and Pezeshk (SP15) Model

The ground-motion spectra for the SP15 GMM are shown in Figure 7–31. This model produces relatively smoothly varying spectra across the range of magnitudes and distances considered

for the evaluation. The TI team noted the slight relative elevation in spectra around 10 Hz for distances of 50 and 100 km, which likely results from the choice of geometric spreading and attenuation used in generating the synthetic ground motions used to constrain this model. The TI team did not consider this to be a significant issue and accepted this model for use without any restrictions.

7.5.2.7 Al Noman and Cramer (ANC15) Model

The ground-motion spectra for the ANC15 GMM are shown in Figure 7–32. This model exhibits anomalous magnitude scaling at low frequencies, suggesting a possible bias due to the use of ground-motion intensity data used in constraining its development. The model uses a fixed "fictitious depth" term, and the developers believe this may limit the ability of the model to extrapolate well up to larger magnitudes. The developers and the TI team deemed that the model was not ready to be used as a seed model for this project in its current form. Thus, this model was not selected as one of the final seed models; reducing the number of seed GMMs to 19.

7.5.2.8 Graizer Model

The ground-motion spectra for the Graizer (2015) GMM are shown in Figure 7–33. A key feature identified in this model is the spectral peak that occurs around 3–5 Hz for all magnitudes and distances, which is much lower than expected in CENA for a site condition of V_{S30} = 3000 m/sec and κ = 0.006 sec. This results in a noticeable drop in the ground-motion levels for frequencies above 10 Hz. Additionally, the scaling at low frequencies (f < 0.2 Hz) leads to relatively low ground-motion levels at large distances (R > 100 km) and large magnitudes (**M** > 7). Members of the TI-Team had a series of discussions with the developer regarding these features of the GMM, and based on these discussions, and with agreement from the developer, the TI team selected this model for use in the bandwidth 0.2 ≤ f ≤ 5 Hz.

7.5.2.9 Hassani and Atkinson (HA15) Model

The ground-motion spectra for the Hassani and Atkinson (2015) GMM are shown in Figure 7–34. This model exhibits magnitude scaling inherent to the reference model (BSSA14 from NGA-West2, Boore et al. 2014) that was used as basis for this GMM. These features include the relative reduction in spectral scaling for 0.2 Hz < f < 5 Hz at large magnitudes (**M** > 6), and the change in slope of magnitude scaling at **M** = 5.5 for high frequencies (f > 5 Hz). The developers of this model feel these features are warranted by the presence of such data trends in data-rich western regions (e.g., in the reference BSSA14 model), and they cannot be ruled out for CENA with the present set of observations. Therefore, the TI team selected this model for use without any limitations.

7.5.2.10 Hollenback, Kuehn, Goulet, and Abrahamson (PEER) Models *(PEER_EX, PEER_GP)*

The ground-motion spectra for the Hollenbeck et al. GMMs are shown in Figures 7–35 and 7–36. These two models differ only in the approach used to scale the results to **M** ≥ 5.5 (the developers felt this was the upper limit for which the CENA data can be used to constrain the model). PEER_EX uses EXSIM (Atkinson and Assatourians 2015) for this scaling, and

PEER_GP uses the Graves and Pitarka (2015) simulation approach for the scaling. At higher frequencies, PEER_EX shows relatively stronger scaling with increasing magnitude compared to PEER_GP, particularly at larger distances. Nonetheless, both models produce reasonable high-frequency motions. At lower frequencies (f < 2 Hz), the models show larger differences. In particular, the spectra for PEER_EX show a noticeable elevation beginning around 1 Hz and extending to lower frequencies. The developers felt that this feature was probably not well constrained by the simulation approach due the very limited available data, and their recommendation was to only use the PEER_EX model for frequencies above 2 Hz. The TI team agreed with this recommendation for the PEER_EX model. For the PEER_GP GMM, the TI team selected its use as a seed model without any restrictions.

7.6 Consistency Checks of Candidate GMMs using Data Residuals

The available data for NGA-East cover only a small subset of the magnitude, distance, and frequency range of the GMM targets. Hence, it is difficult to make a comprehensive quantitative assessment of the candidate GMMs using the available set of ground-motion records. Nonetheless, we have performed checks of the candidate GMMs to ensure that they are at a minimum consistent with the available data. Figure 7–37 shows the total number of NGA-East recordings versus frequency and indicates that frequencies outside 1 to 10 Hz suffer from limited number of data (due to limitations on the useable frequency bandwidth of the recordings) and cannot be reliably used to assess the consistency of the candidate GMMs. We only used data from Region 2 (see Chapter 4); that is, both the station and the event were located in CNA, and PIEs were not excluded. The magnitude range of these events is **M** 4.0 to 6.8. Furthermore, in performing these checks, we only considered sites with V_{S30} > 300 m/sec and distances R < 400 km in order to avoid potential issues with sampling biases. That is, at large distances, only stronger than average events can exceed the signal-to-noise ratio, effectively biasing the dataset to events with larger ground motions for a given magnitude. Histograms showing the sampling of magnitude, distance and V_{s30} for this subset of the data are shown in Figure 7-38.

The process utilized to perform these consistency checks is to first adjust the recorded data to the generic hard rock site conditions of the GMMs (V_{S30} = 3000 m/sec). Clearly, these adjustments have large uncertainty, which makes quantitative interpretation of the results difficult. The TI team conferred with the GMM developers to perform site corrections in a manner that is consistent (or at least not inconsistent) with their own corrections as they were used for calibrating their models. This led to the selection of a common model for the correction upward to 760 m/sec (for softer sites only) and four representative site correction models for the site correction from 760 m/sec to 3000 m/sec. The correction of the recorded data is therefore done in a two-step process. The first step is to correct the data from the as-recorded site condition to a condition of V_{S30} = 760 m/sec using the NGA-West2 site response model of Boore et al. (2014). The second step is to apply an adjustment from V_{S30} = 760 m/sec to V_{S30} = 3000 m/sec. Four different site response models were used for this adjustment: B_avg1 and B_Fea96 (Boore 2015), A12 Atkinson (2012) and Graizer (2014), resulting in four site-corrected datasets. Figure 7–39 plots these site adjustment models as a function of frequency. Note that the B_Fea96, and B_avg1 models are dependent on both magnitude and distance, and the A12 model is distance dependent as well. However, within the 1–10 Hz frequency range the

magnitude and/or frequency dependence is small, and all of the models are fairly consistent with one another.

The next step is to then compute residuals for each GMM and fit these with a mixed-effects model. The final step is a visual inspection of the fitted model bias to ensure most points lie within +/- 0.5 ln units of zero bias over the 1–10 Hz frequency range. This is the criterion used to assess consistency with the available data.

For each candidate GMM, residuals are computed for each of the four site-corrected data sets. These residuals are then fit using a mixed-effects model:

$$Res(f) = c_0(f) + \sigma_{WE}(f) + \sigma_{BE}(f)$$

where the coefficient c_0 represents the model bias as a function of frequency (f), and σ_{WE} and σ_{BE} are the within-event and between event standard deviations.

Figures 7–40 to 7–59 present the model bias of the fitted residuals for the 20 candidate GMMs. The format of these figures is the same for each GMM. Each figure shows the results obtained using the four site correction models. The GMM developers used a variety of site-treatment approaches in their model generation process, and the TI team checked with the developers to ensure they were comfortable with the use of these four site-adjustment models in the assessment of data residuals. Nonetheless, we note that across all of the GMMs, there is not a large difference in the results obtained with the four different site correction approaches.

With just a few exceptions, the model bias falls within +/- 0.5 ln units of zero bias over the 1–10 Hz frequency range of interest for all 20 candidate GMMs. The few cases where the bias falls outside this range are not viewed as significant deviations by the TI team. These include the largest exceedances, which occur for models 1CCSD (Figure 7–46) and 1CVSD (Figure 7–47) starting around 2 Hz with the level of exceedance continuing to increase with decreasing frequency down to 1 Hz. However, while this trend suggests that these models might have significant bias at lower frequencies, the models are not used for frequencies less than 1 Hz following the developers' recommendations (Table7–14). Based on these analyses, the TI team concluded that the candidate GMMs are consistent with the bulk of the currently available NGA-East data.

7.7 Extrapolation of Candidate GMMs to Very Close and Very Large Distances

This section largely borrows from PEER (2015b), Chapters 1 and 2, and is repeated here for completeness.

7.7.1 Motivation

The GMC is to be implemented in hazard analysis codes and should cover the complete range of distances in the Central and Eastern U.S. Seismic Source Characterization for Nuclear Facilities Project (CEUS SSC), including a distance of 0 km or very close to 0 km, in terms of closest distance to the rupture plane, R_{RUP}. As source depth tends to be magnitude-dependent, it is unlikely to have an **M** = 4 event rupture to the surface, but making this source depth and **M** combination possible and defined in the GMC may prevent computational problems in hazard codes.

The various GMMs listed in Table 7–2 cover a wide range of distances based on two main distance metrics: the Joyner-Boore Distance, R_{JB}, and the closest distance to the rupture plane, R_{RUP}. The R_{JB} distance is defined as the horizontal distance from the projection of the earthquake rupture plane on the Earth's surface. It is therefore zero everywhere above the rupture plane and does not include depth. However, the GMMs that use R_{JB} often incorporate a "fictitious depth" to allow for the near saturation of ground-motion levels at very short distances observed in empirical data. Note that GMMs that use R_{RUP} also include such terms to model ground-motion saturation at very short distances.

In the NGA-East project, both R_{RUP} and source depth are included as parameters for the final GMC, implying that the ground motions from all the R_{JB}-based models had to be converted to R_{RUP}. This is a simple correction if the average depths used in the assumption of the models are known or are in the dataset used to constrain the models. However, when this simple conversion is applied, it is possible that the resulting extensions to R_{RUP} near and at 0 km may not be reasonable, depending on the specific modeling assumptions. For that reason, an extrapolation scheme based on fitting of ground-motion values with (converted) R_{RUP} values at larger distances, typically 10 km and larger, was developed.

As documented in the PEER report (2015a), the NGA-East team requested that all the GMMs provide ground motions up to 1200 km. Following subsequent hazard analysis runs and based on experience in assessment of hazard for nuclear facilities, the NGA-East project team extended the distance range to 1500 km. This large distance allows for application of the GMMs at sites in regions with low seismicity where ground motions from a large, distant earthquake (such as from the New Madrid seismic zone) contribute significantly to the site hazard. Because the intent for the NGA-East project team was to evaluate the model over the full range of distances to which it is likely to be applied, extrapolation of all the models to 1500 km was therefore required.

In addition, most models were not strongly calibrated for distances beyond about 400 or 600 km. The difficulty of calibration at large distances using the empirical data (NGA-East database, Goulet et al. 2014) is primarily due to the lack of ground-motion records at large magnitudes. Because most of the dataset is from earthquakes smaller than **M** = 6, the ground motions at large distances are fairly weak and often below the noise threshold for a wide frequency band. The only ground motions that remain are those motions that are "exceptionally high" for the given magnitude and distance, effectively biasing the dataset. This effect is denoted as distance censoring of the data and has been discussed in the literature with respect to development of GMPEs from empirical data; e.g., see Abrahamson et al. (2014) and Chiou and Youngs (2014). Therefore, NGA-East proposed to develop "reasonable" extrapolation rules for large distances as an alternative to the original GMM ground motions. The conversion of R_{JB}-based models to R_{RUP} is presented in Section 7.7.2, and the extrapolation at short and large distance are covered in Sections 7.7.3 and 7.7.4

7.7.2 Conversion from R_{JB} to R_{RUP}

The median GMMs listed in Table 7–2 are available as electronic appendices to the PEER report (2015a). Each GMM includes an EXCEL workbook with a worksheet that provides ground motions for either PGA, PGV, or PSA at a given frequency. Each worksheet includes ground-

motion values for different magnitudes and distances. The full magnitude range covers **M** = 4–8.2, and the distance range covers 0 to 1500 km. Different modeling groups provided ground motions for two different distance metrics: R_{JB} and R_{RUP}. Because the final NGA-East models are required to provide ground-motion values for R_{RUP}, ground motions from GMMs using R_{JB} need to be converted to R_{RUP}.

The following GMMs provide ground-motion values for R_{JB}:

- The Darragh et al. (DASG) suite of models:
    - Single-Corner Constant Stress Parameter (1CCSP)
    - Single-Corner Variable Stress Parameter (1CVSP)
    - Double-Corner Constant Stress Parameter (2CCSP)
    - Double-Corner Variable Stress Parameter (2CVSP)
- Hassani and Atkinson (HA15)
- Shahjouei and Pezeshk (SP15)

The conversion from R_{JB} to R_{RUP} assumes that the ground motions provided by the developers are valid for the footwall (confirmed with GMM developers). For a given value for the depth to the top of the rupture, Z_{TOR}, the corresponding rupture distance value for a given R_{JB} value can be calculated by

$$R_{RUP} = \sqrt{R_{JB}^{\ 2} + Z_{TOR}^{\ 2}} \qquad (7\text{–}1)$$

This provides a corresponding R_{RUP} value for each prediction. Then, interpolation is conducted in logR_{RUP}/logY space to calculate ground-motion values for other values of R_{RUP} (where logY is the natural logarithmic value of ground motions).

A Z_{TOR} value is calculated for each magnitude as follows: for a given a hypocentral depth, the average hypocenter depth ratios, dip of the rupture, rupture area relationship, aspect ratio of the rupture, and the nominal depth to the top and bottom can be calculated. The rupture area relationship is from Somerville (2014):

$$\mathrm{Log}_{10}A = M - 4.25 \qquad (7\text{–}2)$$

The aspect ratio is 1:1 (Table 5.4–1 in EPRI/DOE/NRC [2012]). For strike–slip events, a dip of 75° is used; for reverse events the dip is 45° [Table 5.4–1 in EPRI/DOE/NRC (2012)]. The average depth ratio is computed using the Chiou and Youngs (2008) model (Appendix B in their report); they are 0.6375 for strike–slip events and 0.628 for reverse events.

For the DASG models, a fixed hypocentral depth of 8 km is assumed, which is concordant with the development of their model (Chapter 3 of PEER 2015a). For the HA15 model, a fixed hypocentral depth of 10 km is used as an average of the depths of events in the NGA-East flatfile (12 km) and the depth suggested by Chapter 5 of EPRI/DOE/NRC (2012) (8 km). With these inputs, Z_{TOR} values can be calculated for strike–slip and reverse faulting events. For the conversion, the mean value of the two is used. In Figure 7–60, the Z_{TOR} values for each magnitude are shown for two hypocentral depths. Figures 7–61 through 7–63 show the models originally developed for R_{JB} after their conversion to R_{RUP} (magnitudes span the range from 4 to 8.2 on all the plots, with specific values as provided by the developers, see Section 7.2). The

SP15 model developers used the relationships of Scherbaum et al. [2004] to convert their model predictions from R_{JB} to R_{RUP}.

7.7.3 Distance Extrapolation: General Approach

7.7.3.1 Short-Distance Extrapolation (< 10 km)

For models that require extrapolation to R_{RUP} = 0 km, one of two functional forms listed below is fit to the median predictions in the distance range where the geometric spreading term in the original model is approximately constant (except for near-source saturation effects), and where the conversion from R_{JB} has not introduced unphysical effects. The distance where this occurs varies depending on the specific GMM, but is usually within 50–70 km. For large magnitudes, the original model data at very short distances were used, while for smaller magnitudes, the original data was typically limited to distances of 10–15 km and greater. The two functional forms are as follows:

$$\ln\left(\text{GMIM}\right) = c_1 + c_2 \ln\left(R_{RUP} + c_3\right) \tag{7–3}$$

$$\ln\left(\text{GMIM}\right) = c_1 + c_2 \ln\left(R_{RUP}^2 + h_2\right)^{1/2} \tag{7–4}$$

where GMIM is the ground-motion intensity measure (such as PGA, PGV, or PSA), h is a "fictitious depth" used for ground-motion saturation at close distances, and c_1, c_2, and c_3 are coefficients obtained from regression. For each GMM, the equation that best fits the predictions is selected.

7.7.3.2 Large-Distance Extrapolation (up to 1500 km)

For models that require extrapolation to R_{RUP} =1500 km, a single functional form is fit to the median predictions in a relatively large-distance range where they are deemed usable; again where the geometric spreading term in the original model is approximately constant and where the effects of Q appear to be physically reasonable. This distance varies depending on the specific GMM, but is usually beyond 400 km. The functional form for large distance extrapolation is provided by:

$$\ln\left(\text{GMIM}\right) = c_1 + c_2 \ln\left(R_{RUP}\right) + c_3 R_{RUP} \tag{7–5}$$

7.7.4 Extrapolation Specific to Each GMM

The following sub-sections summarize the process used to extrapolate each specific median GMM. The GMM-specific constraints described below were developed through extensive interaction with each GMM development team. Issues discussed included close-in attenuation shape and whether or not to allow oversaturation, and are reflected below. The intermediate and final extrapolated curves were all reviewed and approved by the GMM developers.

For most of the GMM extensions described below, figures are presented to illustrate the process. For a given frequency, the left panel shows the extension process, and the right panel shows the final extended model results. Each line in the plots corresponds to a specific magnitude. Most of the plots span a range from **M** = 4 or 4.5 to **M** = 8 or 8.2, with spacing of

either 0.1 or 0.5 magnitude units. This variable set of lines was selected to highlight specific features of the GMMs. Note that in all the plots provided below, ground-motion values at R_{RUP} = 0 km are plotted at R_{RUP} = 0.1 km.

7.7.4.1 Boore Models

The R_{RUP} versions of the Boore GMMs require extrapolation to zero distance for the smaller magnitudes. This extrapolation was performed by fitting Equations (7–3) and (7–4) to the predictions for distances in the range of 15–65 km. The blue curves in Figure 7–64 show example results. As indicated on the figure, the conversion from R_{JB} to R_{RUP} and the process of extrapolation led to crossing of attenuation curves at small distances in which **M** = 4 predictions are larger than **M** = 8 predictions at distances less than a few kilometers (e.g., see left panel of Figure 7–65). In general, this occurred for higher-frequency ground motions, especially those using an $R^{-1.3}$ geometrical spreading at short distances, such as B_a04 and B_ab14. Dr. Boore indicated that this is an unintended effect, and that the ground-motion predictions should increase monotonically with magnitude. To implement this intent, the following algorithm was applied to all the Boore GMMs (see example right panel of on Figure 7–65):

- The **M** = 8 ground motions are defined as the upper limit of ground motions for all lower magnitudes.
- Proceeding downward in magnitude from **M** = 8, if the ground motions at zero distance (extended as described above) exceed those for larger magnitudes, they are capped at the predictions for the next largest magnitude; if not, they are retained.
- Having fixed the value at R_{RUP} = 0 from the previous step, the data in the distance range of 20–65 km was refit with Equation (7–4), applied with added constraint of the fixed value at zero distance. The fitted curves were adjusted to match the original predictions at 20 km. The results are the red curves shown on Figure 7–64.

For large distances, Equation (7–5) was fitted from predictions in the R_{RUP} = 140–1000 km range.

7.7.4.2 Darragh, Abrahamson, Silva, and Gregor (DASG) Models

The four GMMs in this suite already provided ground motions up to 1500 km and were converted from R_{JB} to R_{RUP}, as described in Section 7.4.2. The process used for short distances was to fit the data for individual magnitudes over the whole distance range using a combination of Equations (7–3) and (7–5). This model was then used to predict ground motions at distances less than the minimum provided.

The predictions at short R_{RUP} were then scaled so that the prediction at the minimum distance provided by the developed model matched the value provided by the developer to remove any offset. For large magnitudes, the minimum distance provided by the developers was R_{RUP} = 0.5 km, and the fitted model matches the provided model very well over all distances. For smaller magnitudes, there is an upward curvature in the provided data at very short distances. These points with upward curvature were discarded. In order to maintain smooth model, the number of

discarded distances increased for decreasing magnitude. Figure 7–66 shows an example of the process and final extrapolated model for f = 1 Hz.

7.7.4.3 Yenier and Atkinson (YA15) Model

This model is only defined up to 600 km. Figure 7–67 shows an example of extrapolation from 600 to 1500 km. For each frequency and each magnitude, the ground-motion values beyond 50 km (vertical green line) were fitted with Equation (7–5). The red dashed curves show the fit, and the red solid curves show the extrapolation. The extrapolated values were then scaled so that the value at 600 km predicted by the fitted model matched the value provided by YA15 at 600 km.

The YA15 model provided ground-motion values at R_{RUP} = 0 km for all magnitudes. However, for frequencies of 25 Hz and higher, the values for the lower magnitudes show kinks in the distance scaling at R_{RUP} < 15 km (see left pane of Figure 7–68), which are likely artifacts of the adjustments from R_{JB} to R_{RUP}; these adjustments were completed during the model development by Yenier and Atkinson. These breaks were removed as follows (using the 25 Hz PSA as an example):

- The highest frequency without the kinks is 20 Hz. Predictions for 20 Hz at R_{RUP} < 15 km were used as a template for the distance scaling at higher frequencies, using the predictions for **M** = 8.2 as the reference ground motions.
- The ratio of ground motions for any magnitude **M**(i), and the ground motions for **M** = 8.2 is computed at 15 km for 20 Hz PSA:

$$\mathrm{PSARatio}_{20\mathrm{Hz},15\mathrm{km}}[\mathrm{M(i)}] = \ln\left\{\mathrm{PSA}_{20\mathrm{Hz},15\mathrm{km}}[\mathrm{M(i)}]/\mathrm{PSA}_{20\mathrm{Hz},15\mathrm{km},\mathbf{M}8.2}\right\} \quad (7\text{–}6)$$

An equivalent ratio can be computed using the 25 Hz ground motions:

$$\mathrm{PSARatio}_{25\mathrm{Hz},15\mathrm{km}}[\mathrm{M(i)}] = \ln\left\{\mathrm{PSA}_{25\mathrm{Hz},15\mathrm{km}}[\mathrm{M(i)}]/\mathrm{PSA}_{25\mathrm{Hz},15\mathrm{km},\mathbf{M}8.2}\right\} \quad (7\text{–}7)$$

Taking the ratio of Equations (7–6) and (7–7), we obtain a magnitude-scaling factor (MSF):

$$\mathrm{MSF}_{25\mathrm{Hz}}[\mathrm{M(i)}] = \ln\left\{\mathrm{PSARatio}_{25\mathrm{Hz},15\mathrm{km}}[\mathrm{M(i)}]/\mathrm{PSARatio}_{20\mathrm{Hz},15\mathrm{km}}[\mathrm{M(i)}]\right\} \quad (7\text{–}8)$$

which provides the difference in magnitude scaling between 20 and 25 Hz.

- Then for each magnitude **M**(i) < 8.2 and distance R_{RUP}(j) less than 15 km, the ground motions are computed from the prediction for **M** = 8.2 as follows

$$\begin{aligned}\mathrm{PSA}_{25\mathrm{Hz}}[\mathrm{M(i)},R_{RUP}(\mathrm{j})] = {} & \mathrm{PSA}_{25\mathrm{Hz}}[\mathbf{M}8.2,R_{RUP}(\mathrm{j})]* \\ & \exp\left\{\ln\left[\mathrm{PSA}_{20\mathrm{Hz}}[\mathrm{M(i)},R_{RUP}(\mathrm{j})]/\mathrm{PSA}_{20\mathrm{Hz},15\mathrm{km},\mathbf{M}8.2}\right]* \right. \\ & \left. \mathrm{MSF}_{25\mathrm{Hz}}\left[\mathrm{M(i)}\right]\right\}\end{aligned} \quad (7\text{–}9)$$

The same computations are applied to frequencies above 25 Hz and to PGA. This process preserves the predictions for **M** = 8.2 and produces similar magnitude scaling at each distance less than 15 km, as that seen in the 20 Hz motions (Figure 7–68).

7.7.4.4 Pezeshk, Zandieh, Campbell, and Tavakoli (PZCT) Models

Figure 7–69 shows the large-distance extrapolation of these models from 1000 to 1500 km. For each frequency and each magnitude, the ground-motion values beyond 120 km (vertical green line) were fitted with Equation (7–5). The red dashed curves show the fit, and the red solid curves show the extrapolation. The extrapolated values were then scaled so that the value at 1000 km predicted by the fitted model matched the value provided by PZCT at 1000 km.

Although difficult to see on the left frame of Figure 7–69, there is some oversaturation at close distance for the upper magnitude range, which was not a feature intended by the GMM developers. The right frame in Figure 7–69 shows fewer curves, and one can see the adjustment applied at close distances to prevent oversaturation. This is achieved as follows: for each frequency, the magnitude curve that produces the highest ground motions at R_{RUP} = 0 km is identified as the upper-limit ground motions. Then starting in sequence with the next highest magnitude, the ground motions at each distance are taken to be the maximum of the values for that magnitude and the magnitude below. In this way, full saturation is achieved. Both models were extrapolated using Equation (7–5) and the approach described above.

7.7.4.5 Frankel Model

The Frankel model is simulations-based and provided for a series of discrete magnitudes (4.5, 5.5, 6.5, 7.5, and 8.0) and distances. The final ground-motion tables developed and documented in Chapter 6 of the PEER report (2015a) are based on a suite of simulations with Z_{TOR} = 5 km. Following discussions with the developer, and in order to capture the effect of shallower ruptures for larger magnitudes, NGA-East opted to use a modified version of the GMM. The Frankel GMM, as used by NGA-East, is defined by the simulations for Z_{TOR} = 1 km for **M** ≥ 7.5 and for Z_{TOR} = 5 km for all the other magnitudes.

The large distance extrapolation from 1000 to 1500 km was performed by fitting Equation (7–5) to the ground motions from the model in the 200–1000 km range. Because of the limited number of data points and their somewhat irregular nature, coefficient c_2 in Equation (7–5) was fixed at -0.5.

The following steps were taken for the extension to R_{RUP} = 0 km (see Figures 7–70 and 7–71):

- Fit the **M** = 7.5 and 8 ground motions for distances ≤ 10 km with Equation (7–4). Parameter h is set to a small value (0.2 km) that prevents singularity at R_{RUP} = 0 km while producing essentially a linear trend with ln(R_{RUP}).
- Case 1: the ground-motion value at R_{RUP} = 0 km for **M** = 7.5 is smaller or equal to the equivalent ground motion at **M** = 8:
    - Both ground-motions values are kept for R_{RUP} = 0 km (one value for each magnitude, as shown by green dots in Figure 7–70).
    - For each of the two magnitudes, use the fit to predict ground motions to obtain missing values for distances less than those provided by the developer (red dots in Figure 7–70).
- Case 2: the ground-motion value at R_{RUP} = 0 km for **M** = 7.5 is larger than the equivalent ground motion at **M** = 8:

- Combining the data from both magnitudes and refit the data for distances ≤ 10 km with the added constraint that the ground motion R_{RUP} = 0 km is the same for both magnitudes (full saturation is shown as a single green dot in Figure 7–71).
- For each of the two magnitudes, apply the constrained fit to the combined data to predict ground motions at all distances less than those provided by the developer (see red dots in Figure 7–71).

- For **M** = 6.5, 5.5, and 4.5, the ground motions at R_{RUP} < 5 km are computed as the extended values for **M** = 7.5 (from above) multiplied by the ratio of the predictions for **M** = 6.5, 5.5, or 4.5 divided by the prediction for **M** = 7.5 from the Z_{TOR} = 5 km simulations. In other words, the slope in ground motions from the **M** = 7.5 is applied to the lower magnitudes for R_{RUP} < 5 km. The results are shown by the red open circles in Figures 7–70 and 7–71.

The ground-motion values were then extrapolated to **M** = 4 and **M** = 8.2, and interpolated to fill-up the intermediate magnitudes and distances composing the table.

The final step was to smooth the model predictions to remove the irregularities introduced by averaging the results from a limited number of simulations. The smoothing was performed by fitting a flexible function form individually to the predictions for each magnitude and spectral frequency. The function form allowed for changes in the geometric spreading term in four distance ranges: 0 to 5 km, 5 to approximately 50 km, 50 to approximately 180 km, and beyond 180 km. The distance breakpoints near 50 and 180 km were adjusted individually by eye to follow the trends in the simulated results for the individual magnitudes and spectral periods.

7.7.4.6 Shahjouei and Pezeshk (SP15) Model

Before SP15 GMM was extrapolated, it was first converted from R_{JB} to R_{RUP}, as described in Section 7.4.2. Figure 7–72 shows the extrapolation from 1000 to 1500 km. For each frequency and each magnitude, the ground-motion values beyond R_{RUP} = 120 km (vertical green line) were fitted with Equation (7–5). The red dashed curves show the fit, and the red solid curves show the extrapolation. The extrapolated values were then scaled so that the value at R_{RUP} = 1000 km predicted by the fitted model matched the value provided by SP15 at 1000 km.

Also shown on Figure 7–72 is the extrapolation from the minimum R_{RUP} provided by SP15 to R_{RUP} = 0 km. This is performed assuming that there is no change in amplitude at R_{RUP} values smaller than the smallest R_{RUP} provided by SP15.

7.7.4.7 Graizer Model

Median ground motions were provided for the complete distance range and did not require extrapolations (Figure 7–73).

7.7.4.8 Hassani and Atkinson (HA15) Model

Before HA15 GMM was extrapolated, it was first converted from R_{JB} to R_{RUP}, as described in Section 7.4.2. Figure 7–74 shows an example of extrapolation. Starting with **M** = 8, the curves for **M** = 8 to 6.9 from the HA15 results were extended from R_{RUP} = 0.5 km to R_{RUP} = 0 km, using

a fit to Equation (7–4), with the ground motions for each magnitude capped by the predictions for the next largest magnitude.

After capping, predictions for **M** = 6.9 were fitted with Equation (7–4) to obtain the value of h for **M** = 6.9. For **M** = 6.8 and lower, h was set using the following equation from YA15:

$$h = h(\mathbf{M}\,6.9) * 10^{0.235(\mathbf{M}-6.9)} \qquad (7\text{–}10)$$

Equation (7–4) was fitted (for each magnitude below 6.8), using the h obtained with Equation (7–10) and the ground motions for R_{RUP} values between 20 and 50 km. The resulting fitted model was extrapolated back to R_{RUP} = 0 km, capping the predictions with those for the next highest magnitude (Figure 7–74).

7.7.4.9 PEER Models

For both PEER GMMs, ground-motion predictions beyond 400 km were discarded, and the ground-motion values in the 140– 400 km R_{RUP} range were fitted with Equation (7–5). The fitted model was adjusted to match the ground motions at R_{RUP} = 400 km and then used to extrapolate to 1500 km. The ground-motion values for R_{RUP} < 10 km were discarded, and those in the distance range of 10–50 km were fitted by Equation (7–4). The fitted model was scaled to match the ground motions at 10 km and used to extrapolate to R_{RUP} = 0 km. At higher frequencies, some degree of oversaturation (negative magnitude scaling at very small values of R_{RUP}) is predicted by the PEER models. The PEER development team indicated that this was acceptable behavior of their model. The degree of oversaturation is indicated on Figures 7–35 and 7–36. Ground-motion predictions are shown on these figures in 0.5 magnitude increments. As indicated, oversaturation only occurs at the highest magnitudes (**M** > 7.5) and at distances less than about 10 km.

7.8 Spectral-Shape Adjustments

The adjustment of the GMMs to large distances was performed independently for each frequency. This sometimes led to an unphysical spectrum at large distances (beyond 600 km), with a trough occurring at high frequencies well beyond the peak of the spectrum (> 10 Hz). This observation was also made for some of the initial NGA-East GMMs (before extrapolation) and was noticed for large distances in some of the NGA-West2 GMMs (Youngs and Abrahamson, personal communication). Examples are shown for the 1CCSP and the PEER_GP models (blue lines in Figures 7–77 and 7–78). It can also happen that a spectrum of a seed model exhibits some smaller secondary peaks beyond the main peak, also at large distances, which is a remnant of the frequency-independent extrapolation to large distances (Figure 7–79). This effect is not observed for all models, but when it occurs, it is typically for distances larger than 600 km (the specific range is model dependent). Additionally, the effect can occur at all magnitudes, but it is smaller at low magnitudes.

To correct these unphysical spectra, the high-frequency part is fit by a linear trend. If there is a trough at frequencies beyond the main spectral peak (the main spectral peak typically occurs around 1–2 Hz for these distances), the high-frequency ground-motion values are fixed to a specific value, following two possible cases. If there is a secondary peak beyond the trough, then the ground-motion values are fixed at the geometric mean of the trough value and the

secondary peak value (Figures 7–77 and 7-79). Otherwise, the spectrum is fixed at the geometric mean of the trough value and the PGA value (anchored at 200 Hz in the plots, for illustration), as shown in Figure 7–78. The reason the high-frequency PSA values are set to the geometric mean of the maximum/minimum value beyond 10 Hz—and not to the PGA value—is that the TI team does not judge any of the high-frequency values to be more reliable than others. In particular, there is no reason to believe that the PGA value is more reliable than the PSA value at 50 Hz or 15 Hz. After the distance extrapolation is completed, the correction process described above is applied systematically to all the spectra from all the GMMs for distances of 600 km or larger at all magnitudes.

7.9 Summary of Selected Seed Models

The 19 selected GMMs from Section 7.5, extended in distance as described in Section 7.7 and with spectral shapes modified following the rules in Section 7.8 (as needed), constitute the "seed models." The seed models, as utilized in Chapter 8, are available in electronic Appendix D.2. Appendix D.3 provides electronic plotting tools to illustrate the magnitude, distance and frequency characteristics of the seed models. The seed models are used in the development of the median GMMs for the reference site condition of kappa = 0.006 sec and V_{S30} = 3000 m/sec, as described in Chapter 8. Additionally, Section 8.3.2 describes physicality constraints applied to the seed models during the process of sampling new GMMs from this distribution.

7.10 References

Abrahamson, N.A., W.J. Silva, and R. Kamai. 2014. "Summary of the ASK14 Ground-Motion Relation for Active Crustal Regions." Earthq. Spectra, 30(3): 1025–1055.

Al Atik, L., and R.R. Youngs. 2014. "Epistemic Uncertainty for NGA-West2 Models." Earthq. Spectra, 30(3): 1301–1318.

Ancheta, T.D., R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S.-J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida, and J.L. Donahue. 2014. "NGA-West2 Database." Earthq. Spectra, 30: 989–1005.

Atkinson, G.M. 2004. "Empirical Attenuation of Ground-Motion Spectral Amplitudes in Southeastern Canada and the Northeastern United States." Bull. Seism. Soc. Am., 94, 1079–1095.

Atkinson, G.M. 2008. "Ground Motion Prediction for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty." Bull. Seismol. Soc. Am., 98(3): 1304–1318.

Atkinson, G.M. 2012. "White Paper on Proposed GMPEs for 2015 National Seismic Hazard maps." http://www.seismotoolbox.ca/GMPEtables2012/r12_GMPEs9b.pdf (last accessed January 2015).

Atkinson, G.M., and K. Assatourians. 2015. "Implementation and Validation of EXSIM (a Stochastic Finite-Fault Ground-Motion Simulation Algorithm) on the SCEC Broadband Platform." Seismol. Res. Lett., 86(1): 48–60.

Atkinson, G.M., and D.M. Boore. 1995. “Ground Motion Relations for Eastern North America.” Bull. Seismol. Soc. Am. 85(1): 17–30.

Atkinson, G.M., and D.M. Boore. 2006. Earthquake Ground-Motion Prediction Equations for Eastern North America.” Bull. Seismol. Soc. Am. 96, 2181–2205. (also see the erratum published in Vol. 97, No. 3, p. 1032).

Atkinson, G.M., and D.M. Boore. 2011. “Modification to Existing Ground-Motion Prediction Equations in Light of New Data.” Bull. Seismol. Soc. Am., 101(3): 1121–1135.

Atkinson, G.M., and D.M. Boore. 2014. “The Attenuation of Fourier Amplitudes for Rock Sites in Eastern North America.” Bull. Seismol. Soc. Am., 104: 513–528.

Boatwright, J., and L. Seekins. 2011. “Regional Spectral Analysis of Three Moderate Earthquakes in Northeastern North America.” Bull. Seismol. Soc. Am., 101(4): 1769–1782.

Boore, D.M. 1996. “SMSIM-Fortran Programs for Simulating Ground Motions from Earthquakes: Version 1.0.” U.S. Geol. Survey Open-File Rept. 96-80-A, 73 pgs.

Boore, D.M. (2003). “SMSIM: Stochastic Method SIMulation of Ground Motion from Earthquakes, Chapter 85.13 in: International Handbook of Earthquake and Engineering Seismology.” (edited by W.H.K. Lee, H. Kanamori, P.C. Jennings, and C. Kisslinger), 81B. and CD 3, Academic Press, Amsterdam, pp. 1631–1632.

Boore, D.M. (2005). “On Pads and Filters: Processing Strong-Motion Data.” Bull. Seismol. Soc. Am., 95: 745–750.

Boore, D.M., K.W. Campbell, and G.M. Atkinson. 2010. “Determination of Stress Parameters for Eight Well-Recorded Earthquake in Eastern North America.” Bull. Seismol. Soc. Am., 100(4): 1632–1645.

Boore, D.M., J.P. Stewart, E. Seyhan, and G.M. Atkinson. 2014. “NGA-West 2 Equations for Predicting PGA, PGV, and 5%-Damped PSA for Shallow Crustal Earthquakes.” Earthq. Spectra, 30(3): 1057–1085.

Boore, D.M., and E.M. Thompson. 2015. “Revisions to Some Parameters used in Stochastic-Method Simulations of Ground Motion.” Bull. Seismol. Soc. Am., 105: 1029–1041.

Bora, S.S., F. Scherbaum, N. Kuehn, and P. Stafford. 2014. “Fourier Spectral- and Duration Models for the Generation of Response Spectra Adjustable to Different Source-, Propagation-, and Site Conditions.” Bull. Earthq. Eng., 12: 467–493.

Bozorgnia, Y., N.A. Abrahamson, L. Al Atik, T.D. Ancheta, G.M. Atkinson, J. Baker, A. Baltay, D.M. Boore, K.W. Campbell, B.-S.J. Chiou, R.B. Darragh, S. Day, J.L. Donahue, R.W. Graves, N. Gregor, T. Hanks, I.M. Idriss, R. Kamai, T. Kishida, A. Kottke, S.A. Mahin, S. Rezaeian, B. Rowshandel, E. Seyhan, S. Shahi, R. Shantz, W.J. Silva, P. Spudich, J.P. Stewart, J. Watson-Lamprey, K.E. Wooddell, R.R. Youngs. 2014. “NGA-West2 Research Project.” Earthq. Spectra, 30(3): 973–987.

Campbell, K.W., Y.M.A. Hashash, B. Kim, A.R. Kottke, E.M. Rathje, W.J. Silva, and J.P. Stewart. 2014. “Reference-Rock Site Conditions for Central and Eastern North America: Part II – Attenuation Definition.” PEER Report No. 2014/12, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Chiou, B.S.-J., and R.R. Youngs. 2008. “NGA Model for Average Horizontal Component of Peak Ground Motion and Response Spectra.” PEER Report No. 2008/09, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Chiou, B.S-J., and R.R. Youngs. 2014. “Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra.” Earthq. Spectra, 30(3): 1117–1153.

Cramer, C.H., and O.S. Boyd. 2014. “Why the New Madrid earthquakes are M7-8 and the Charleston earthquake is ~M7.” Bull. Seismol. Soc. Am., 104: 2884–2903.

Cramer, C.H., and A. Kumar. 2003. “2001 Bhuj, India, Earthquake Engineering Seismoscope Recordings and Eastern North America Ground-Motion Attenuation Relations.” Bull. Seismol. Soc. Am., 93: 1390–1394

EPRI. 2013. “EPRI 2004-2006 Ground-Motion Model (GMM), Final Report Project 3002000717.” June 2013, Electric Power Research Institute, Palo Alto, CA.

EPRI/USDOE/USNRC. 2012. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Frankel, A. 2009. “A Constant Stress-Drop Model for Producing Broadband Synthetic Seismograms—Comparison with the Next Generation Attenuation Relations.” Bull. Seismol. Soc. Am., 99: 664–681.

Frankel, A., C. Mueller, T. Barnhard, D. Perkins, E.V. Leyendecker, N. Dickman, S. Hanson, and M. Hopper. 1996. “National Seismic-Hazard Maps: Documentation June 1996.” U.S. Geological Survey, Open-File Report 96-532.

Goulet, C.A., T. Kishida. T.D. Ancheta, C.H. Cramer, R.E. Darragh, W.J. Silva, Y.M.A. Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. 2014. “PEER NGA-East Database.” PEER Report No. 2014/17, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Graizer, V., and E. Kalkan. 2007. “Ground-Motion Attenuation Model for Peak Horizontal Acceleration from Shallow Crustal Earthquakes.” Earthq. Spectra, 23(3): 585–613.

Graizer, V., and E. Kalkan. 2009. “Prediction of Response Spectral Acceleration Ordinates Based on PGA Attenuation.” Earthq. Spectra, 25(1): 39–69.

Graves, R.W., and A. Pitarka. 2015. “Refinements to the Graves and Pitarka (2010) Broadband Ground Motion Simulation Method.” Seismol. Res. Lett., 86(1): 75–80.

Hartzell, S.W., C. Langer, and C. Mendoza. 1994. “Rupture Histories of Eastern North American Earthquakes.” Bull. Seismol. Soc. Am., 84(6): 1703–1724.

Hashash, Y.M.A., A.R. Kottke, J.P. Stewart, K.W. Campbell, B. Kim, E.M. Rathje, and W.J. Silva. 2014. “Reference Rock Site Condition for Central and Eastern North America.” Bull. Seismol. Soc. Am., 104: 684–701.

Hassani, B., and G.M. Atkinson. 2015. “Referenced Empirical Ground-Motion Model for Eastern North America.” Seismol. Res. Lett., 86(2): 477–491.

Joyner, W.B., and D.M. Boore. 1993. “Methods for Regression Analysis of Strong Motion Data.” Bull. Seismol. Soc. Am., 83: 469–487.

Joyner, W.B., and D.M. Boore. 1994. “Errata: Methods for Regression Analysis of Strong Motion Data.” Bull. Seismol. Soc. Am., 84: 955–956.

Ogweno, L.P., and C.H. Cramer. 2014. “Regression Relationships between Modified Mercalli Intensities and Ground Motion Parameters (Abstract).” Program and Abstracts, 86th Annual Meeting, Eastern Section of the Seismological Society of America, Charleston, SC.

PEER. 2015a. “NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

PEER. 2015b. “NGA-East: Adjustments to Median Ground-Motion Models for Central and Eastern North America.” PEER Report No. 2015/08, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Pezeshk, S., A. Zandieh, and B. Tavakoli. 2011. “Hybrid Empirical Ground-Motion Prediction Equations for Eastern North America using NGA Models and Updated Seismological Parameters.” Bull. Seismol. Soc. Am., 101(4): 1859–1870.

Scherbaum, F., J. Schmedes, and F. Cotton. (2004).” On the Conversion of Source-To-Site Distance Measures for Extended Earthquake Source Models.” Bull. Seismol. Soc. Am., 94: 1053–1069.

Silva, W.J., N. Gregor, and R.B. Darragh. 2002. “Development of Regional Hard Rock Attenuation Relations for Central And Eastern North America.” Pacific Engineering and Analysis Report, El Cerrito, CA.

Somerville, P.G. 2014. “Scaling relations between seismic moment and rupture area of earthquakes in stable continental regions.” PEER Report No. 2014/14, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Somerville, P.G., N. Collins, N.A. Abrahamson, R.W. Graves, and C. Saikia. 2001. “Ground Motion Attenuation Relations for the Central and Eastern United States.” Report to U.S. Geological Survey, NEHRP Award No. 99HQGR0098, 36 pgs.

Spudich, P., and L. Xu (2003). “Software for Calculating Earthquake Ground Motions from Finite-Faults in Vertically Varying Media.” In IASPEI International Handbook of Earthquake and Engineering Seismology, W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger (Editors), Chapter 85.14, Academic Press, New York, New York, 1633–1634.

Toro, G.R., N.A. Abrahamson, and J.F. Schneider. 1997. A Model of Strong Ground Motions from Earthquakes in Central and Eastern North America: Best Estimates and Uncertainties. Seismol. Res. Lett., 68(1): 41–57.

Toro, G.R. 2002. “Modification of the Toro et al. (1997) Attenuation Equations for Large Magnitudes and Short Distances.” Risk Engineering Technical Report, 10 pgs.

Yenier, E., and G.M. Atkinson. 2015. “Regionally-Adjustable Generic Ground-Motion Prediction Equation based on Equivalent Point-Source Simulations: Application to Central and Eastern North America.” Bull. Seismol. Soc. Am., 105: 1989–2009.

Youngs, R.R., and N.A. Abrahamson, personal communication.

Table 7–1 EPRI (2013) review project GMMs.

Cluster	Model type	Reference
1	Single corner Brune source	• Silva et al. (2002), SCCS-Sat • Silva et al. (2002), SCVS • Toro et al. (1997) • Frankel et al. (1996)
2	Complex/empirical source; ~R^{-1} geometrical spreading R < 70 km	• Silva et al. (2002), DC-Sat • A08': Atkinson (2008) with modifications from Atkinson and Boore (2011) and EPRI
3	Complex/empirical source; ~$R^{-1.3}$ geometrical spreading R < 70 km	• AB06': Atkinson and Boore (2006) with modifications from Atkinson and Boore (2011) • Pezeshk et al. (2011)
4	Finite-source; Full waveform Green's functions	• Somerville et al. (2001), two models for rifted and non-rifted regions

Table 7–2 Summary of NGA-East median GMMs.

Title (Authorship), chapter number in *PEER Report 2015/04* (2015a)	Acronym(s)
Point-Source Stochastic-Method Simulations of Ground Motions for the PEER NGA-East Project (D.M. Boore), Chapter 2.	Six GMMs from Boore, each based on a different Q and geometrical spreading model: B_a04 B_ab14 B_ab95 B_bca10d B_bs11 B_sgd02
Development of Hard Rock Ground-Motion Models for Region 2 of Central and Eastern North America (R.B. Darragh, N.A. Abrahamson, W.J. Silva, and N. Gregor), Chapter 3.	Four GMMs from DASG, each based single or double corner point source and on constant or variable stress parameter: 1CCSP 1CVSP 2CCSP 2CVSP
Regionally-Adjustable Generic Ground-Motion Prediction Equation based on Equivalent Point-Source Simulations: Application to Central and Eastern North America (E. Yenier and G.M. Atkinson), Chapter 4.	YA15
Ground-Motion Prediction Equations for Eastern North America using a Hybrid Empirical Method (S. Pezeshk, A. Zandieh, K.W. Campbell, and B. Tavakoli), Chapter 5.	Two GMMs from PZCT based on different large **M**-scaling (simulation- and empirical-based): PZCT15_M1SS PZCT15_M2ES
Ground-Motion Predictions for Eastern North American Earthquakes Using Hybrid Broadband Seismograms from Finite-Fault Simulations with Constant Stress-Drop Scaling (A. Frankel), Chapter 6.	Frankel
Hybrid Empirical Ground-Motion Model for Central and Eastern North America using Hybrid Broadband Simulations and NGA-West2 GMPEs (A. Shahjouei and S. Pezeshk), Chapter 7.	SP15
Empirical Ground-Motion Prediction Equations for Eastern North America (M.N. Al Noman and C.H. Cramer), Chapter 8.	ANC15
Ground-Motion Prediction Equations for the Central and Eastern United States (V. Graizer), Chapter 9.	Graizer
Referenced Empirical Ground-Motion Model for Eastern North America (B. Hassani and G.M. Atkinson), Chapter 10.	HA15
PEER NGA-East Median Ground-Motion Models (J. Hollenback, N. Kuehn, C.A. Goulet and N.A. Abrahamson), Chapter 11.	Two GMMs from PEER, based on alternate finite-fault models: PEER_GP PEER_EX

Table 7–3 Representative geometric spreading and Q models selected for the Boore GMMs development.

Model and reference	Geometric spreading G(R)	What is "R"?[1]	Attenuation exp(-πfR/Qβ)	Applicable range[2]
B_ab95 Atkinson and Boore [1995]	$G(R) = \begin{cases} R^{-1}, & R \le 70 \text{ km} \\ C_0 R^{0}, & 70 \text{ km} < R \le 130 \text{ km} \\ C_1 R^{-0.5}, & R > 130 \text{ km} \end{cases}$ $C_0 = (1/70)$, $C_1 = (130^{0.5}/70)$	$R = R_{hyp}$	$Q(f) = 680f^{0.36}$ $\beta = 3.8$ km/sec	$4.0 \le \mathbf{M} \le 7.25$ $10 \le R \le 500$ km $0.5 \le f \le 20$ Hz
B_sgd02 Silva et al. [2002]	$G(R) = \begin{cases} R^{-(a+b(\mathbf{M}-6.5))}, & R \le 80 \text{ km} \\ C_0 R^{-0.5(a+b(\mathbf{M}-6.5))}, & R > 80 \text{ km} \end{cases}$ $a = 1.0296$, $b = -0.0422$, $C_0 = 80^{-0.5(a+b(\mathbf{M}-6.5))}$	$R = R_{hyp}$	$Q(f) = 351f^{0.84}$ $\beta = 3.52$ km/sec	$4.5 \le \mathbf{M} \le 8.5$ $1 \le R \le 400$ km $0.1 \le f \le 100$ Hz
B_a04 Atkinson [2004]	$G(R) = \begin{cases} R^{-1.3}, & R \le 70 \text{ km} \\ C_0 R^{0.2}, & 70 \text{ km} < R \le 140 \text{ km} \\ C_1 R^{-0.5}, & R > 140 \text{ km} \end{cases}$ $C_0 = (70^{-0.2}/70^{1.3})$, $C_1 = C_0(140^{0.5}/140^{-0.2})$	$R = R_{hyp}$	$Q(f) = \max(1000, 893f^{0.32})$ $\beta = 3.7$ km/sec	$4.4 \le \mathbf{M} \le 6.8$ $10 \le R \le 800$ km $0.05 \le f \le 20$ Hz
B_bca10d Boore et al. [2010]	$G(R) = R^{-1}$ all R	$R = R_{PS}$	$Q(f) = 2850$ $\beta = 3.7$ km/sec	$4.4 \le \mathbf{M} \le 6.8$ $10 \le R \le 800$ km $0.05 \le f \le 20$ Hz
B_bs11 Boatwright and Seekins [2011]	$G(R) = \begin{cases} R^{-1}, & R \le 50 \text{ km} \\ C_0 R^{-0.5}, & R > 50 \text{ km} \end{cases}$ $C_0 = (50^{0.5}/50)$	$R = R_{hyp}$	$Q(f) = 410f^{0.5}$ $\beta = 3.5$ km/sec	$4.4 \le \mathbf{M} \le 5.0$ $23 \le R \le 602$ km $0.2 \le f \le 20$ Hz
B_ab14 Atkinson and Boore [2014]	$G(R) = \begin{cases} 10^{T_c C_{LF}} R^{-1.3}, & R \le 50 \text{ km} \\ C_0 R^{-0.5}, & R > 50 \text{ km} \end{cases}$ $T_c = \begin{cases} 1, & f \le 1 \text{ Hz} \\ 1 - 1.429 \log_{10}(f), & 1 \text{ Hz} < f < 5 \text{ Hz} \\ 0, & f \ge 5 \text{ Hz} \end{cases}$ $C_{LF} = \begin{cases} 0.2 \cos\left[\frac{\pi}{2}\left(\frac{R-h}{1-h}\right)\right], & R \le h \\ 0.2 \cos\left[\frac{\pi}{2}\left(\frac{R-h}{50-h}\right)\right], & h < R < 50 \text{ km} \end{cases}$ h = focal depth (km), $C_0 = (50^{0.5}/50^{1.3})$	$R = R_{PS}$	$Q(f) = 525f^{0.45}$ $\beta = 3.7$ km/sec	$3.5 \le \mathbf{M} \le 6$ $10 \le R \le 500$ km $0.2 \le f \le 20$ Hz

[1] R_{hyp} = hypocentral distance; R_{PS} = effective point source distance; $R_{PS} = [R_{hyp}^2 + h_{FF}^2]^{1/2}$, $\log_{10}(h_{FF}) = -0.405 + 0.235\mathbf{M}$ (Yenier and Atkinson, 2015)

[2] When applicable range not explicitly stated in paper it was inferred from data comparisons.

Table 7–4 Summary table: Boore GMMs

Boore Models: B_a04, B_ab14, B_ab95, B_bca10d, B_bs11, B_sgd02	
Key feature(s) of model	Suite of 6 GMM derived from point-source stochastic simulations and based on 6 different combinations of Q and geometrical spreading models. A single corner source spectral shape was used for all models. Stress parameter was estimated by inverting the PSA at 0.1 and 0.2 s from 9 ENA earthquakes recorded within 200 km. The geometric mean of the two stress parameters was used along with Yenier and Atkinson (2015) finite-fault factor, and the Boore and Thompson (2015) path duration.
Applicability	
Distance metric used	R_{RUP}
Site Conditions	V_{S30} = 3000 m/sec, κ = 0.006 sec
PSA metric predicted	Geometric mean of horizontal components.
GM Intensity Measures	PSA at 0.1–100 Hz, PGA, PGV
Magnitude range	4.0–8.0
Distance range	0–1200 km
Constraints	
Summary of selected data used in model development.	Events: Nahanni, Saguenay, Mt. Laurier, Cap Rouge, St. Anne, Kipawa, Rivière du Loup, Val des Bois. V_{S30} = 2 km/sec, R_{RUP} < 200 km
Large M extrapolation	Point-source stochastic model, Brune spectrum.
Small distance extrapolation	Point-source stochastic model, Brune spectrum
Extrapolation to high frequencies (>15 Hz) and PGA	Point-source stochastic model, Brune spectrum
Constraints on Q	Used 6 preselected attenuation models
Constraints on geometrical spreading	Used 6 preselected attenuation models
Style of faulting	Not explicitly considered.

Table 7–5 Summary table: Darragh et al. (DASG) GMMs

Darragh et al. (2015) Models: 1CCSP, 1CVSP, 2CCSP, 2CVSP	
Key feature(s) of model	Single and double corner point-source model with G(R), stress parameter, and Q(f) empirically constrained for NGA-East Region 2. Used both constant and variable stress models.
Applicability	
Distance metric used	R_{JB}
Site Conditions	V_{S30} = 3000m/sec, κ = 0.006 sec
PSA metric predicted	Geometric mean of horizontal components.
GM Intensity Measures	PSA at 0.1–100 Hz, PGA, PGV
Magnitude range	4.5–8.2
Distance range	0–1000 km
Constraints	
Summary of selected data used in model development.	NGA-East database: Region 2, 53 events (Including 10 PIEs), 10 ≤ R_{HYP} ≤ 1000 km all sites
Large M extrapolation	Magnitude scaling is determined from equivalent point-source simulations using NGA-West2 GMPEs.
Small distance extrapolation	Close distance scaling is determined from equivalent point-source simulations using NGA-West2 GMPEs.
Extrapolation to high frequencies (>15 Hz) and PGA	Point source, 1-, 2- corner.
Constraints on Q	Determined empirically from selected ground motions in Region 2.
Constraints on geometrical spreading	Geometrical spreading function features magnitude dependent attenuation based on NGA-West2 GMPEs
Style of faulting	Not explicitly considered.

Table 7–6 Summary table: Yenier and Atkinson (YA15) GMM.

Yenier and Atkinson (2015) Model: YA15	
Key feature(s) of model	Developed based on equivalent point-source simulations with parameters calibrated to California motions. Considers decoupled effects of magnitude, distance, stress parameter, geometrical spreading rate and anelastic attenuation coefficient. Adjusted to CENA based on the analysis of residuals between generic model predictions and observed motions.
Applicability	
Distance metric used	R_{RUP}
Site Conditions	Originally derived for V_{S30} = 760 m/sec, κ = 0.025 sec. Corrected to VHR site condition (V_{S30} = 3000 m/sec, κ = 0.006 sec) using site adjustment factors from Atkinson (2012, White Paper).
PSA metric predicted	RotD50
GM Intensity Measures	PSA at 0.1–100 Hz, PGA, PGV
Magnitude range	3.0–8.2
Distance range	1–600 km
Constraints	
Summary of selected data used in model development.	NGA-East database (2014-09-12): Earthquakes of **M** ≥ 3.0 that were recorded by at least three stations within 600 km (including PI Events). Excluded ground motions obtained in the Gulf Coast region. Maximum usable period as specified in the provided data table.
Large M extrapolation	Magnitude scaling is determined from equivalent point-source simulations using parameters calibrated to California motions.
Small distance extrapolation	Short distance scaling is determined from equivalent point-source simulations using parameters calibrated to California motions. Saturation effect is defined based on the recent findings of empirical studies.
Extrapolation to high frequencies (>15 Hz) and PGA	Determined from selected ground motions. No limit on high frequency considered for response spectra.
Constraints on Q	No constraints. Determined empirically from selected ground motions.
Constraints on geometrical spreading	Geometrical spreading function features magnitude- and period- dependent attenuation rates determined from equivalent point-source simulations where Fourier domain geometrical spreading is defined as $R^{-1.3}$ within the first 50 km and $R^{-0.5}$ at R > 50 km.
Style of faulting	Not explicitly considered.

Table 7–7 Summary table: Pezeshk et al. (PZCT) GMMs.

Pezeshk et al. (2015) Models: PZCT15_M1SS, PZCT15_M2ES	
Key feature(s) of model	Traditional Hybrid Empirical approach with NGA-West 2 GMM used to define the "host" region and the NGA-East Database used to calibrate the "host-to-target" adjustments.
Applicability	
Distance metric used	R_{RUP}
Site Conditions	V_{S30} = 3000 m/sec, κ = 0.006 sec
PSA metric predicted	RotD50
GM Intensity Measures	PSA at 0.1–100Hz, PGA
Magnitude range	3.0–8.0
Distance range	0–1000 km
Constraints	
Summary of selected data used in model development.	Used 5 NGA-West2 GMPEs as the empirical ground-motion models in the WNA host region. NGA-East database is used for comparison and calibration: 1. Excluded Gulf Coast 2. Excluded PIEs 3. Used V_{S30} > 180 m/sec (adjusted to 3000 m/sec) 4. Used data with quality flag of 0
Large M extrapolation	1. Used Hybrid Empirical Method (HEM) for M ≤ 6, then used magnitude scaling predicted by NGA-West2 GMPEs. 2. Used (HEM) for all magnitudes.
Small distance extrapolation	Used HEM (controlled by NGA-West2 GMPEs)
Extrapolation to high frequencies (>15 Hz) and PGA	Used HEM with point-source stochastic simulation model.
Constraints on Q	1. Used most recent database to determine Q for ENA. 2. Performed inversion of NGA-West2 GMPEs to determine Q for WNA.
Constraints on geometrical spreading	Same constraints that were used for Q
Style of faulting	Not explicitly considered.

Table 7–8 Summary table: Frankel GMM.

Frankel (2015) Model: Frankel	
Key feature(s) of model	Uses finite-fault deterministic synthetics at long periods and stochastic, finite-fault synthetics at short periods; combined with matched filter using a crossover frequency dependent on moment. Uses constant dynamic and static stress drops with moment. Method described in Frankel (2009 BSSA). Uses 1/R geometrical spreading for stochastic part, for distances out to 70 km, based on Charlevoix observations (Frankel, submitted to BSSA).
Applicability	
Distance metric used	R_{RUP}
Site Conditions	V_{S30} = 2800 m/sec; κ = 0.006 sec
PSA metric predicted	Geometric mean of horizontal components.
GM Intensity Measures	PSA at 0.1–100Hz, PGA, PGV
Magnitude range	4.5, 5.5, 6.5, 7.5, 8.0
Distance range	2-1000km
Constraints	
Summary of selected data used in model development.	Observed PSA's for **M** 4.5–5.5 were compared to PSA's from simulations
Large M extrapolation	Explicitly included in finite-fault simulations
Small distance extrapolation	Explicitly included in finite-fault simulations
Extrapolation to high frequencies (>15 Hz) and PGA	Explicitly included in stochastic finite-fault simulations
Constraints on Q	Q from Atkinson and Boore (1995)
Constraints on geometrical spreading	For stochastic portion, used geometrical spreading model of Atkinson and Boore (1995). 1/R true geometrical spreading (< 70 km) based on observations of Charlevoix earthquakes (Frankel, submitted to BSSA)
Style of faulting	Simulations were done for vertical strike–slip fault and 45°dipping thrust fault

Table 7–9 Summary table: Shahjouei and Pezeshk (SP15) GMM.

Shahjouei and Pezeshk (2015) Model: SP15	
Key feature(s) of model	Hybrid empirical ground-motion model developed based on the ratios of intensity measures of WNA and CENA regions and considering the empirical GMPEs developed in NGA-West2 project. Synthetics are generated from the hybrid broadband simulation technique and the finite-fault method.
Applicability	
Distance metric used	R_{JB}
Site Conditions	V_{S30} = 3000 m/sec, κ = 0.006 sec
PSA metric predicted	RotD50
GM Intensity Measures	PSA at 0.1–100Hz, PGA, PGV
Magnitude range	5.0–8.0
Distance range	2–1000 km
Constraints	
Summary of selected data used in model development.	Used 5 NGA-West2 GMPEs as the empirical ground-motion models in the WNA host region. NGA-East database is used for comparison and calibration: 1. Excluded Gulf Coast 2. Excluded PIEs 3. Used V_{S30} > 180 m/sec (adjusted to 3000 m/sec) 4. Used data with quality flag of 0
Large M extrapolation	Explicitly included in finite-fault simulations
Small distance extrapolation	Explicitly included in finite-fault simulations
Extrapolation to high frequencies (>15 Hz) and PGA	Explicitly included in stochastic finite-fault simulations
Constraints on Q	Equal weighting of Q models from Atkinson and Boore (2014) and Pezeshk et al. (2011), respectively: $Q = \max(1000, 893f^{0.32})$ $Q = 525f^{0.45}$
Constraints on geometrical spreading	Equal weighting of geometric spreading terms from Atkinson and Boore (2014) and Pezeshk et al. (2011), respectively: $Z(R) = R^{-1.3}$ for R < 50 km and $R^{-0.5}$ for R > 50 km $Z(R) = R^{-1.3}$ for R < 70 km, $R^{0.2}$ for 70 km < R < 140 km and $R^{-0.5}$ for R > 140 km
Style of faulting	Not explicitly considered.

Table 7–10 Summary table: Al Noman and Cramer (ANC15) GMM.

Al Noman and Cramer (2015) Model: ANC15	
Key feature(s) of model	Empirical model with intensity data converted to ground motion for **M** > 6
Applicability	
Distance metric used	R_{RUP}
Site Conditions	V_{S30} = 760 m/sec with correction term for other values
PSA metric predicted	RotD50
GM Intensity Measures	PSA at 0.1 – 10 Hz, PGA, PGV
Magnitude range	3.0–8.2
Distance range	0–2000 km
Constraints	
Summary of selected data used in model development.	NGA East flatfile (2014-11-18) for earthquakes north of 35N and east of 100W and 1985 Nahanni eqks (4), plus 1976 **M** 6.8 Gazli record and 2001 **M** 7.6 Bhuj observations (Cramer and Kumar 2003, BSSA). PIEs included if in selected area. Also used intensity data converted to ground-motion estimates (Ogweno and Cramer 2014, ES-SSA abstract) for the 1925 **M** 6.2 Charlevoix, 1929 **M** 7.2 Grand Banks, 1886 **M** 7.0 Charleston SC, Dec. 1811 **M** 7.5 New Madrid, Jan. 1812 **M** 7.3 New Madrid, and Feb. 1812 **M** 7.7 New Madrid earthquakes. Magnitude for the 1886 and 1811–1812 earthquakes are from Cramer and Boyd (2014, BSSA).
Large M extrapolation	Historical earthquake intensities converted to ground motions as listed above.
Small distance extrapolation	Limited by few observations.
Extrapolation to high frequencies (>15 Hz) and PGA	Observations used for PGA. No values for frequencies above 10 Hz Sa provided by GMPE.
Constraints on Q	Selected data from Mid-continental regions avoiding Gulf Coast and WUS recordings and earthquakes. Bhuj Q is similar to ENA Q at distances less than 300 km (Bodin et al. 2004, BSSA), which corresponds to the Cramer and Kumar (2003, BSSA) dataset.
Constraints on geometrical spreading	Used unilinear model as most data beyond 50 km.
Style of faulting	Explicit terms for reverse, strike–slip, and undefined.

Table 7–11 Summary table: Graizer GMM.

Graizer (2015) Model: Graizer	
Key feature(s) of model	Model developed with modular filter based approach [Graizer and Kalkan (2007; 2009)], technique developed for the western U.S.). Spectral shapes developed for the WUS are adjusted for CEUS using recorded data.
Applicability	
Distance metric used	R_{RUP}
Site Conditions	V_{S30} = 2800 m/sec, κ = 0.006 sec. Include adjustment to lower V_{S30}.
PSA metric predicted	RotD50
GM Intensity Measures	PSA at 0.1– 100Hz, PGA
Magnitude range	4.0–8.2
Distance range	0–1000 km
Constraints	
Summary of selected data used in model development.	Used subset of the NGA-East database of Sept. 2014 for **M** > =3.75 and distances R_{RUP} < 1000 km with addition of 6 Kansas **M** = 4.8 records (5032 data points). All events, including PIE (if **M** >=3.75). All V_{S30} used.
Large M extrapolation	Scaling with the same approximation function type as for WUS (Graizer and Kalkan 2007) with coefficients adjusted based on a combination of (1) ratios of 3.5 < **M** <6 from the NGA-East database,(2) average stress-drop ratio between WUS and CEUS; and 3) checked against recent simulations of **M** >= 5.0.
Small distance extrapolation	Same approach as for the WUS. Used same corner distances (flat area in the near-fault area) as in WUS.
Extrapolation to high frequencies (> 15 Hz) and PGA	Based on info from EPRI 2013 and NGA-East database.
Constraints on Q	Used fixed Q_o = 650 (based on Erickson, D., D. E. McNamara, and H. M. Benz, 2004, Bull. Seism. Soc. Am., **94**, 1630-1643). Potentially can be adjusted to other areas in CEUS with different Q_o.
Constraints on geometrical spreading	Geometrical spreading of R^{-1} is based on data.
Style of faulting	Not explicitly considered.

Table 7–12 Summary table: Hassani and Atkinson (HA15) GMM.

Hassani and Atkinson (2015) Model: HA15	
Key feature(s) of model	Referenced empirical approach was used to adjust the NGA-West2 BSSA14 model to the observed CENA data. Residuals were defined as the ration of the observed CENA data to the equivalent predicted values of BSSA14. Mixed-effect regression on residuals was used to estimate the adjustment factors coefficients.
Applicability	
Distance metric used	R_{JB}
Site Conditions	Originally derived for V_{S30}= 760 m/sec, κ = 0.025 sec. Corrected to VHR site condition (V_{S30} = 3000 m/sec, κ = 0.006 sec) using site adjustment factors from Atkinson (2012, White Paper).
PSA metric predicted	RotD50
GM Intensity Measures	PSA at 0.1–100Hz, PGA, PGV
Magnitude range	3.0–8.5
Distance range	0–400 km
Constraints	
Summary of selected data used in model development.	1. NGA-East database: All $R_{JB} \leq$ 400 km and **M** ≥ 3. All events, including PIE. Exclude Gulf Coast. All V_{S30} used. Limited to usable bandwidth as specified in the flatfile, and only considered data in 0.1–20 Hz range. 2. Seismotoolbox and southern Ontario database. All $R_{JB} \leq$ 400 km and **M** ≥ 3. All V_{S30} used. Limited to usable bandwidth as defined by Signal/Noise ≥ 2, Nyquist frequency, and 0.1–20 Hz range.
Large M extrapolation	Same magnitude scaling as BSSA14
Small distance extrapolation	Same close distance extrapolation as BSSA14.
Extrapolation to high frequencies (>15 Hz) and PGA	Adjustment factors defined up to 20 Hz and also PGA. Extrapolation above 20 Hz uses factors defined at f = 20 Hz.
Constraints on Q	No constraint, not solving for Q.
Constraints on geometrical spreading	No constraint, not solving for geometrical spreading.
Style of faulting	Not explicitly considered.

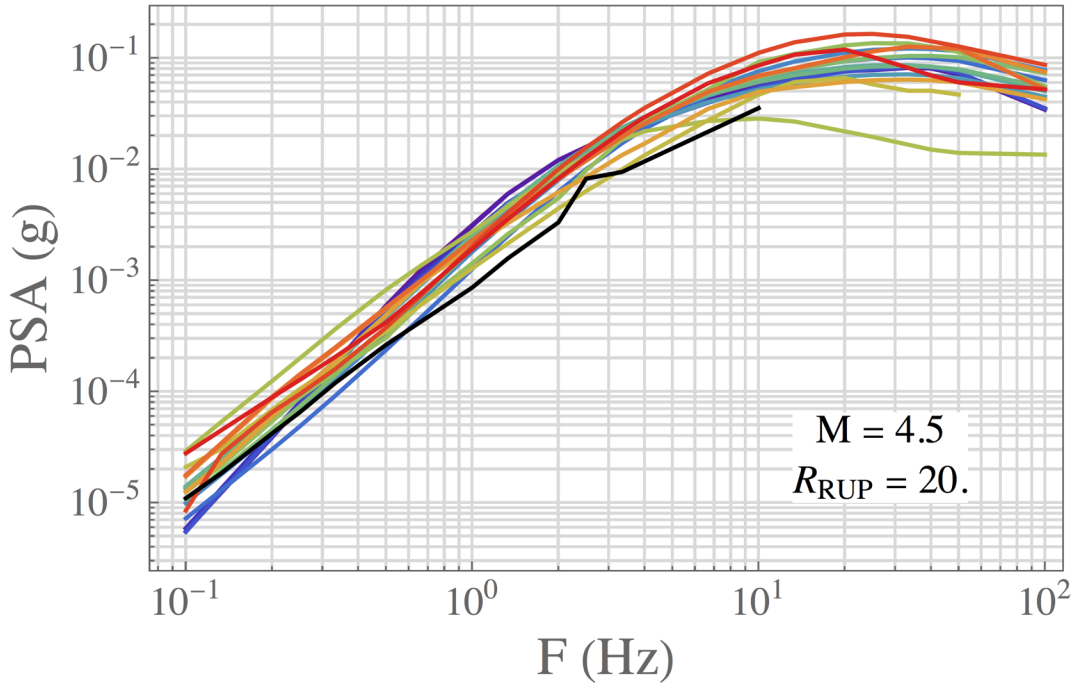
Table 7–13 Summary table: PEER GMMs.

PEER (Hollenback et al., 2015) Models: PEER_EX, PEER_GP	
Key feature(s) of model	Empirical mixed-effect regression on FAS, frequency-by-frequency. PSA predicted through RVT with a calibrated duration model. Extrapolation outside of data range using both finite-fault simulations and point-source model. Two magnitude scaling models based on different finite-fault simulations models.
Applicability	
Distance metric used	R_{RUP}
Site Conditions	Model centered at V_{S30} ~700 m/sec, analytically corrected to V_{S30} = 3000 m/sec, κ = 0.006 sec
PSA metric predicted	RotD50
GM Intensity Measures	PSA at 0.1– 100 Hz, PGA
Magnitude range	4.0–8.2
Distance range	0–1200 km
Constraints	
Summary of selected data used in model development.	1. NGA-East database: All R_{RUP} < 300 km and **M** > 2.5. All events, including PIE. Regions 2–5 (exclude Gulf Coast). All V_{S30} used. Limited to usable bandwidth, as specified in flatfile. 2. NGA-West2 database: subset of ~1400 records covering **M** 3–7.9, R_{RUP} < 250 km. Used for duration model constraint.
Large M extrapolation	Finite-fault simulations for distance-dependent M-scaling up to **M** 8 and up to 10 Hz. Point-source model for higher frequencies, using parameters from inversion.
Small distance extrapolation	Explicitly included in simulation model.
Extrapolation to high frequencies (>15 Hz) and PGA	Point Source, Brune spectrum
Constraints on Q	No constraint, not solving for Q.
Constraints on geometrical spreading	Tri-linear model; hinge locations fixed at 50 and 150 km, slope for 50 km > R_{RUP} and 50 km < R_{RUP} < 150 km regressed from data, slope for R_{RUP} > 150km fixed at $R^{-0.5}$ (for FAS).
Style of faulting	Not explicitly considered.

Table 7–14 Evaluation summary of NGA-East median GMMs.

Model	Acceptance	Comments
B_ab04	As is.	
B_ab14	As is.	
B_ab95	As is.	
B_bca10d	As is.	
B_bs11	As is.	
B_sgd02	As is.	
1CCSP	Partial.	Elevated level of low-frequency (f < 1 Hz) spectra particularly for larger magnitudes (**M** > 6). Developers noted similar bias for WUS single-corner models when compared against recorded motions. Developers recommend only using f > 1 Hz.
1CVSP	Partial.	Elevated level of low-frequency (f < 1 Hz) spectra particularly for larger magnitudes (**M** > 6). Developers noted similar bias for WUS single-corner models when compared against recorded motions. Developers recommend only using f > 1 Hz.
2CCSP	As is.	
2CVSP	As is.	
YA15	As is.	Relative drop in response around 50 Hz (not considered an issue by TI team).
PZCT15_1MSS	As is.	
PZCT15_M2ES	As is.	
Frankel	As is.	Spectral shape a bit jagged due to limited number of simulations (not considered an issue by TI team).
SP15	As is.	
ANC15	No.	Magnitude scaling at low-frequencies suggests possible bias due to use of intensity data, particularly at shorter distance range. Fixed h term (fictitious depth), doesn't extrapolate well with magnitude. Developers recommend not including this model as a seed model.
Graizer	Partial.	Peak in the spectra occurs around 3-5 Hz for all magnitudes and distances, which is much lower than expected for CENA for site condition of V_{S30}=3000 and κ=0.006sec. Scaling of low frequency levels (f < 0.2 Hz) lead to relatively low values at large distance (R>100 km) and large magnitude (**M** > 7). Recommend using only in the bandwidth 0.2 <= f <= 5 Hz.
HA15	As is.	Magnitude scaling exhibits features inherent to the reference model (BSSA14) used as basis for this GMM. These features include relative reduction in scaling for 0.2 Hz < f < 5 Hz at large magnitude (**M** > 6), and change in slope of magnitude scaling at **M** 5.5 for high-frequencies (f>5 Hz). Developers feel these features are warranted by data used to develop reference model (BSSA14), and they cannot be ruled out for CENA with present set of observations.
PEER_EX	Partial.	Magnitude scaling at low-frequency suggests possible bias. Developers recommend only using f > 2 Hz.
PEER_GP	As is.	

Figure 7–1 **5%-damped PSA for candidate GMMs at M = 4.5 and distance of 20 km.**



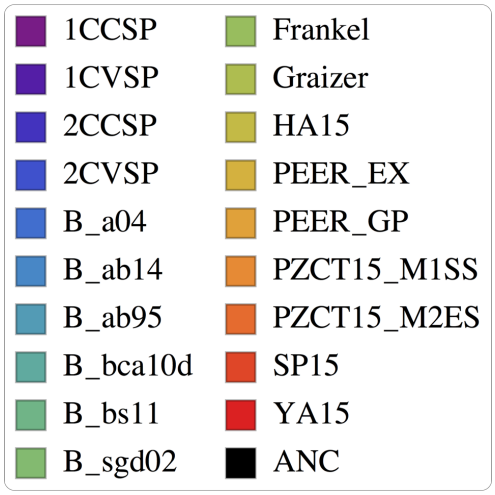


Figure 7–2 **5%-damped PSA for candidate GMMs at M = 4.5, R = 50 km (see legend on Figure 7–1).**

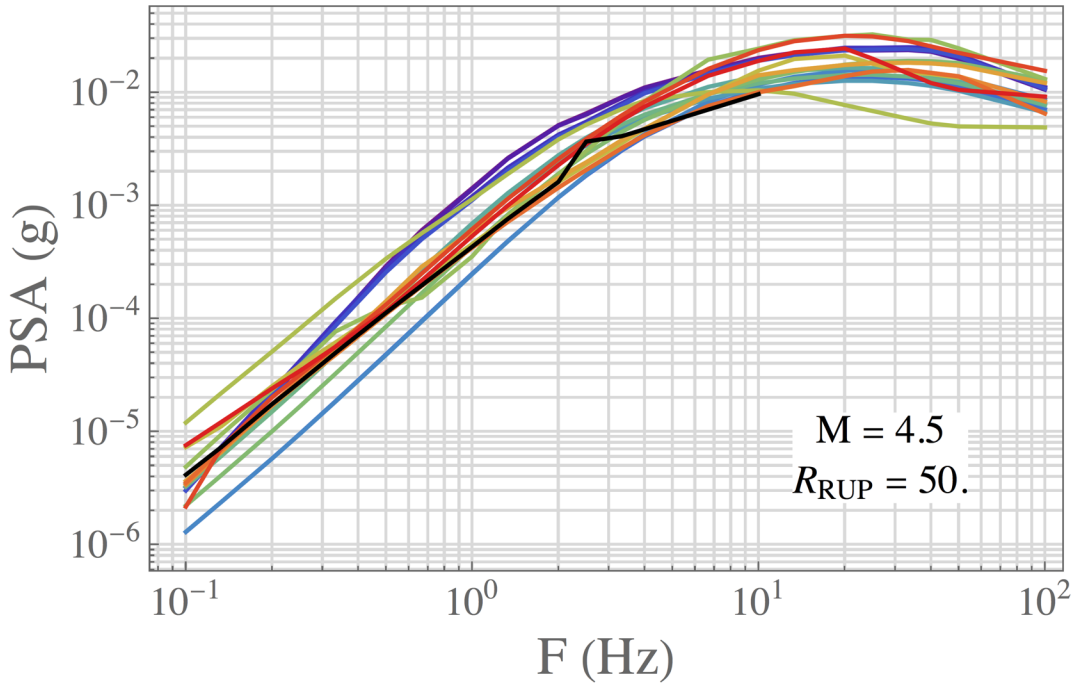


Figure 7–3 **5%-damped PSA for candidate GMMs at M = 4.5, R = 100 km (see legend on Figure 7–1).**

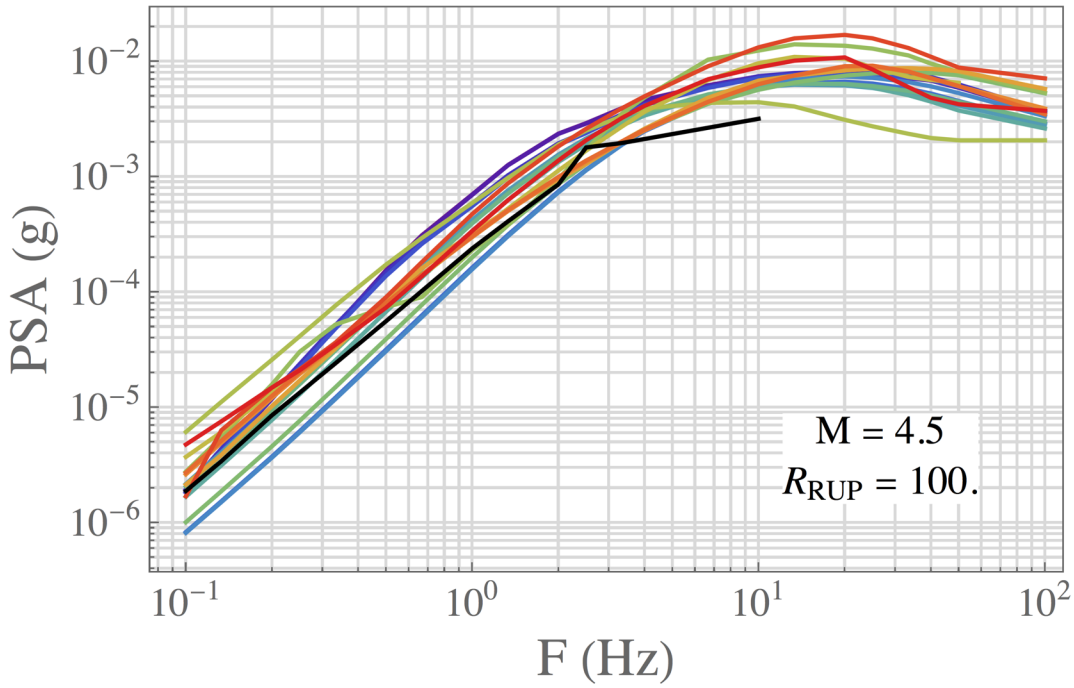


Figure 7–4 **5%-damped PSA for candidate GMMs at M = 4.5, R = 200 km (see legend on Figure 7–1).**

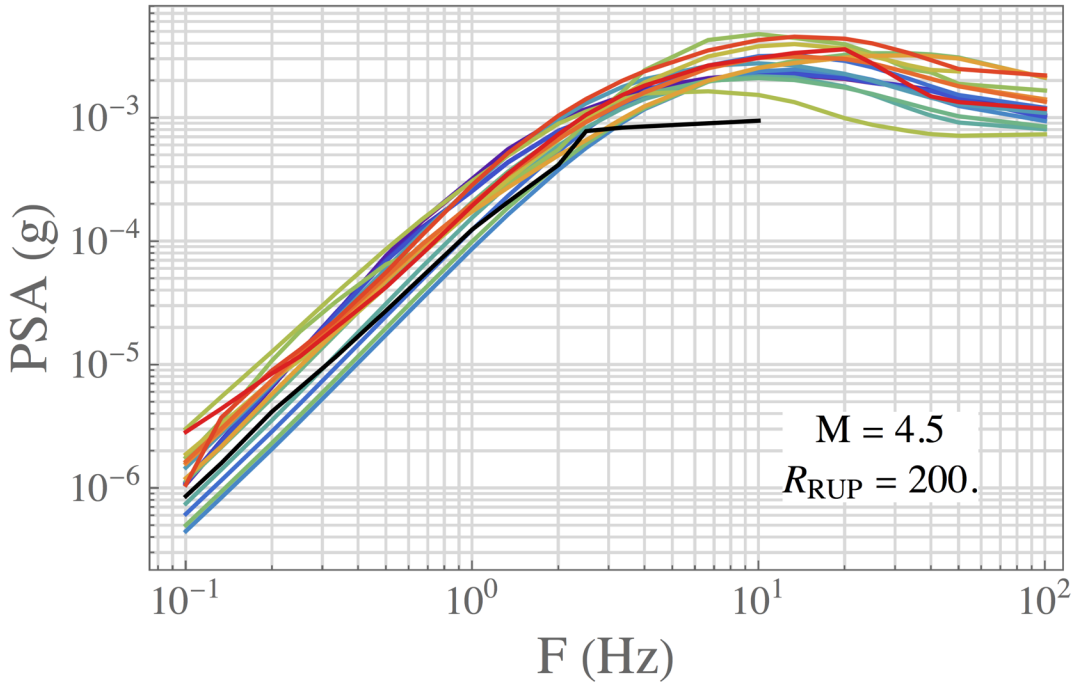


Figure 7–5 **5%-damped PSA for candidate GMMs at M = 5.5 and distance of 20 km (see legend on Figure 7–1).**

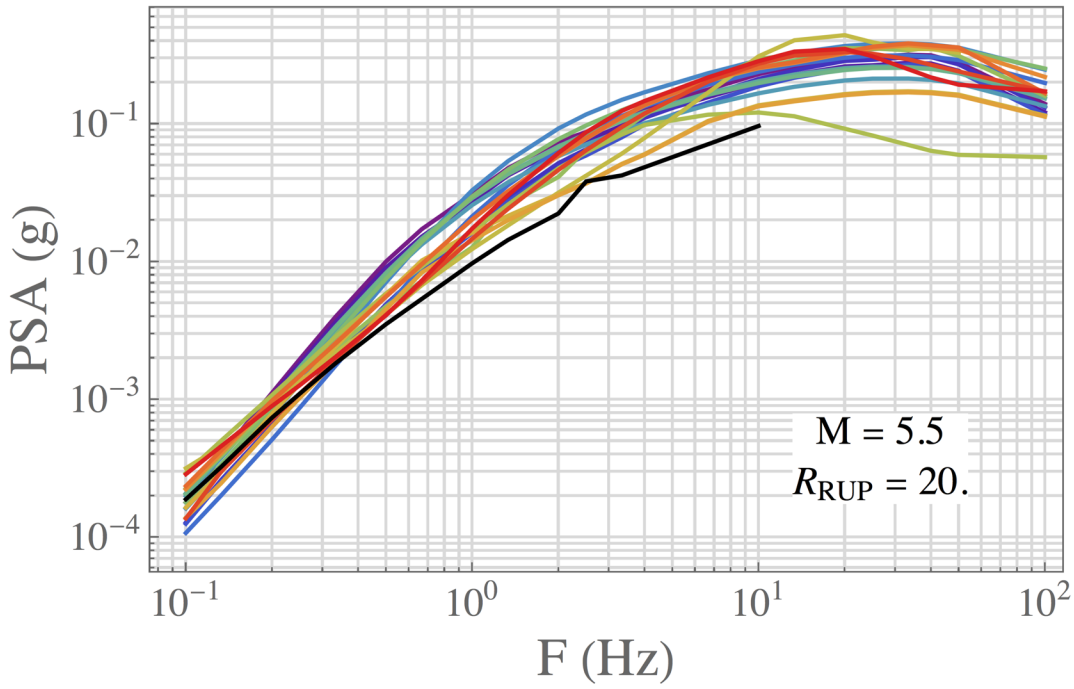


Figure 7–6 **5%-damped PSA for candidate GMMs at M = 5.5, R = 50 km (see legend on Figure 7–1).**

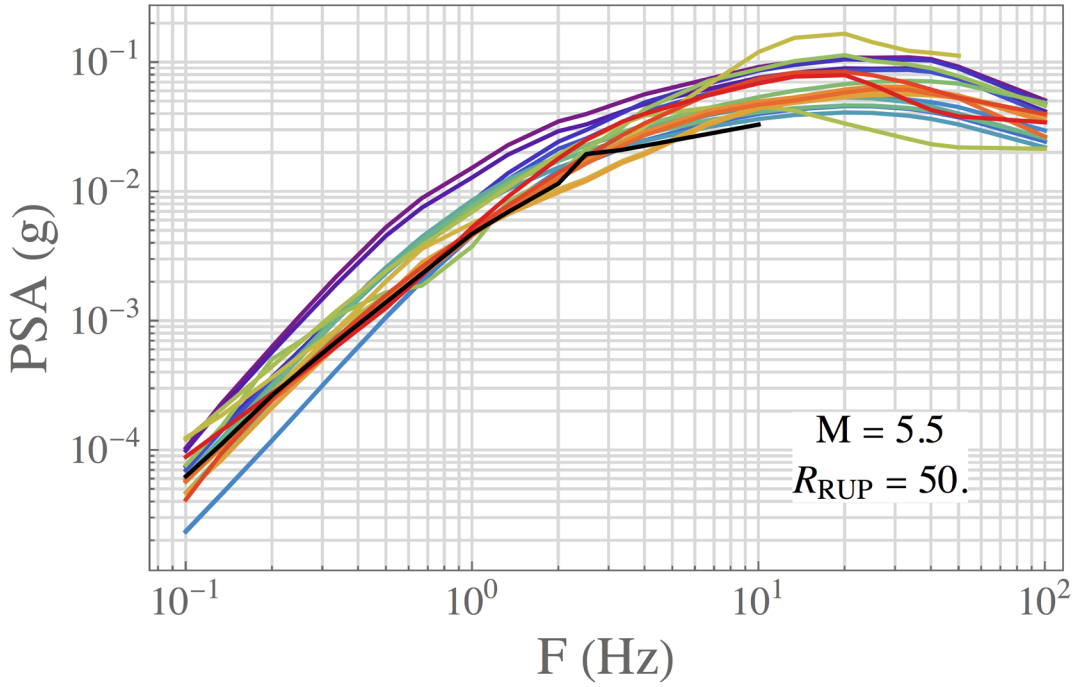


Figure 7–7 **5%-damped PSA for candidate GMMs at M = 5.5, R = 100 km (see legend on Figure 7–1).**

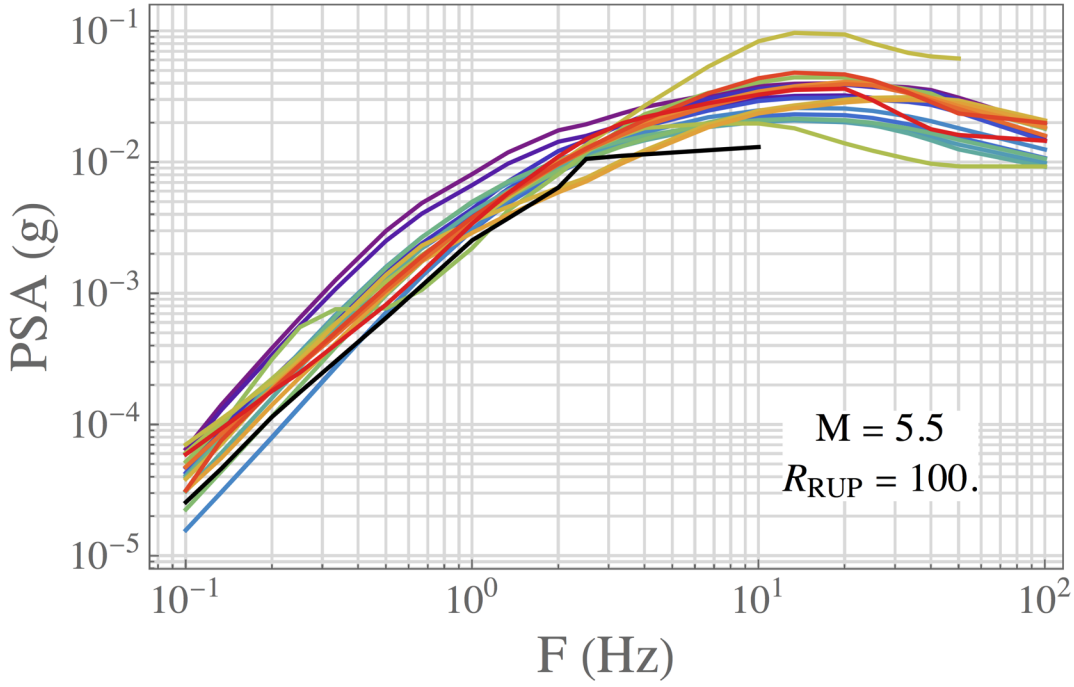


Figure 7–8 **5%-damped PSA for candidate GMMs at M = 5.5, R = 200 km (see legend on Figure 7–1).**

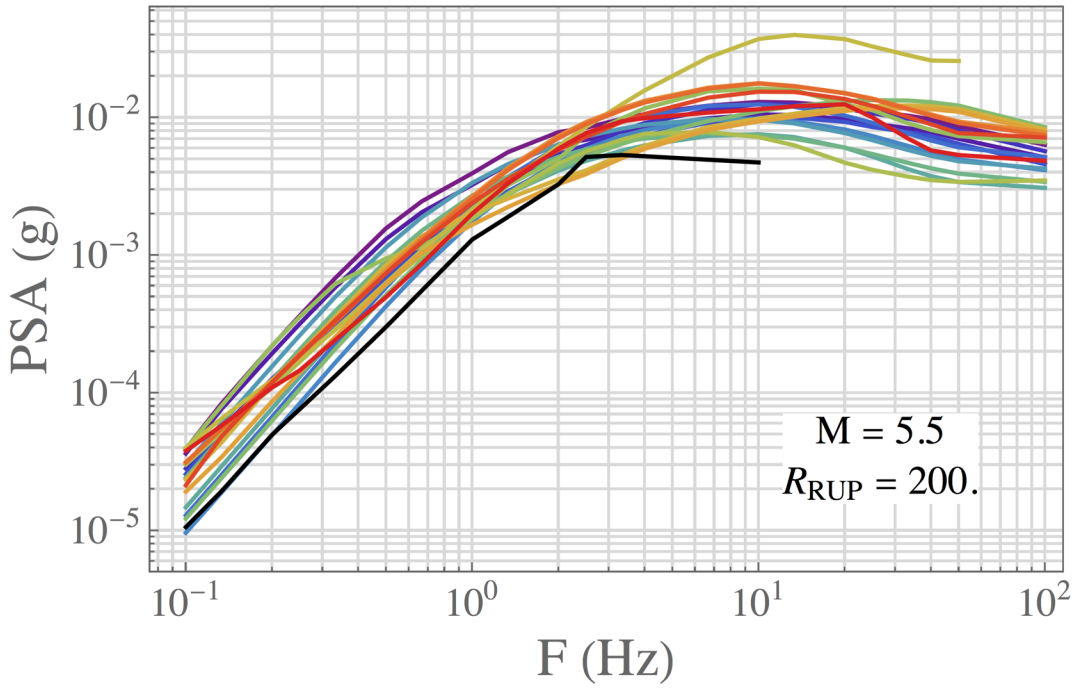


Figure 7–9 **5%-damped PSA for candidate GMMs at M = 6.5 and R = 20 km (see legend on Figure 7–1).**

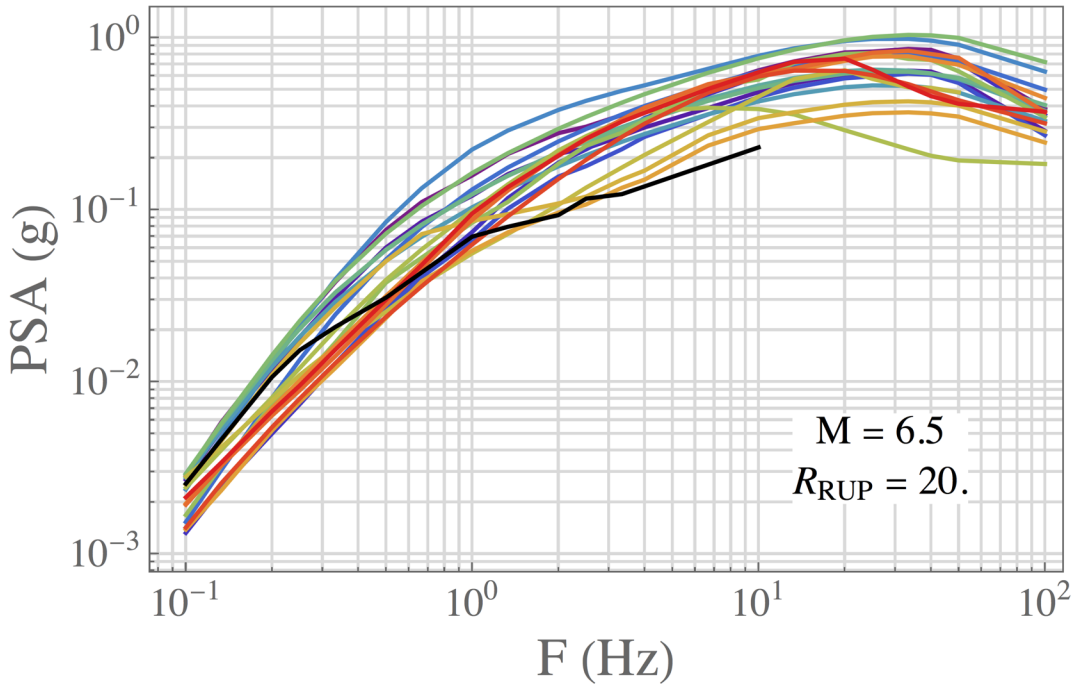


Figure 7–10 5%-damped PSA for candidate GMMs at M = 6.5, R = 50 km (see legend on Figure 7–1).

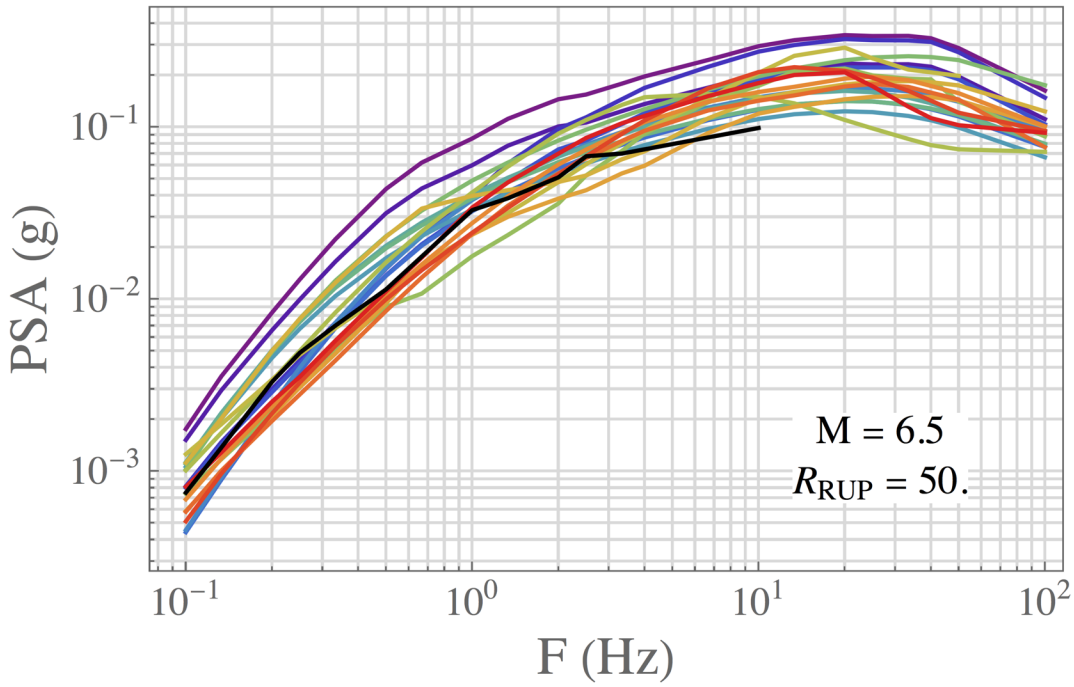


Figure 7–11 5%-damped PSA for candidate GMMs at M = 6.5, R = 100 km (see legend on Figure 7–1).

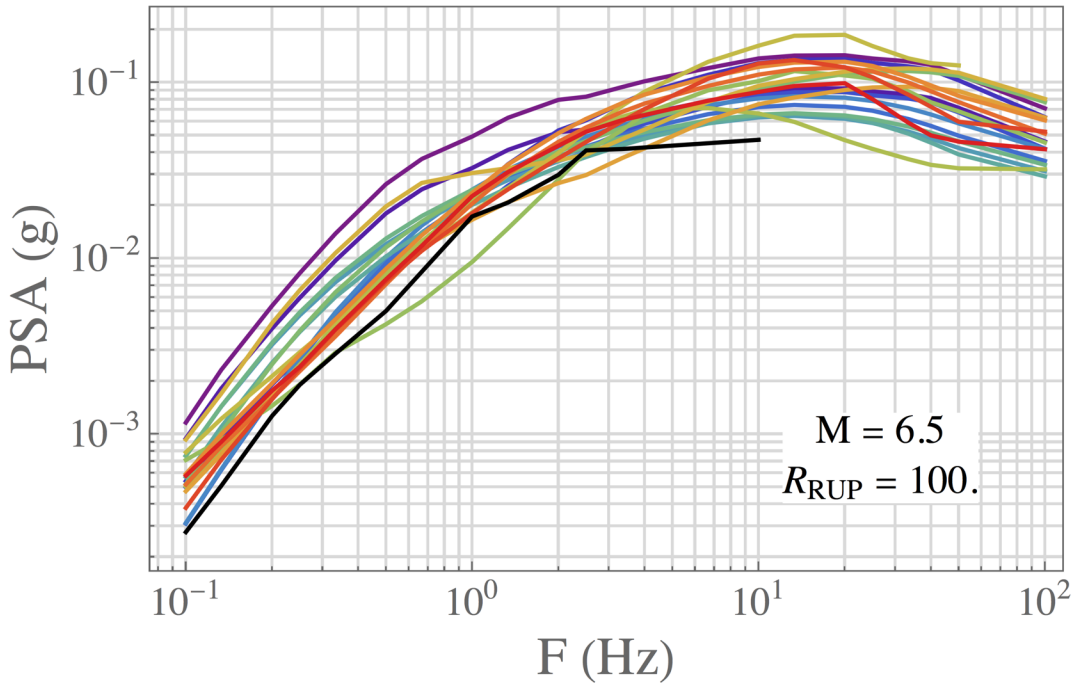


Figure 7–12 **5%-damped PSA for candidate GMMs at M = 6.5, R = 200 km (see legend on Figure 7–1).**

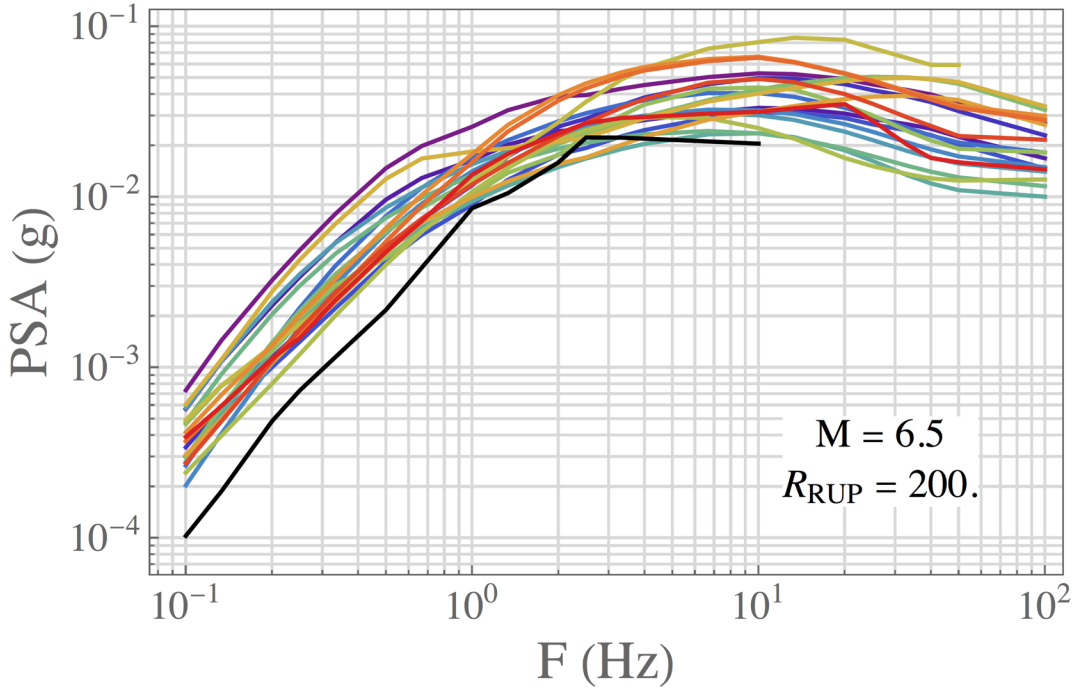


Figure 7–13 **5%-damped PSA for candidate GMMs at M = 7.5 and R = 20 km (see legend on Figure 7–1).**

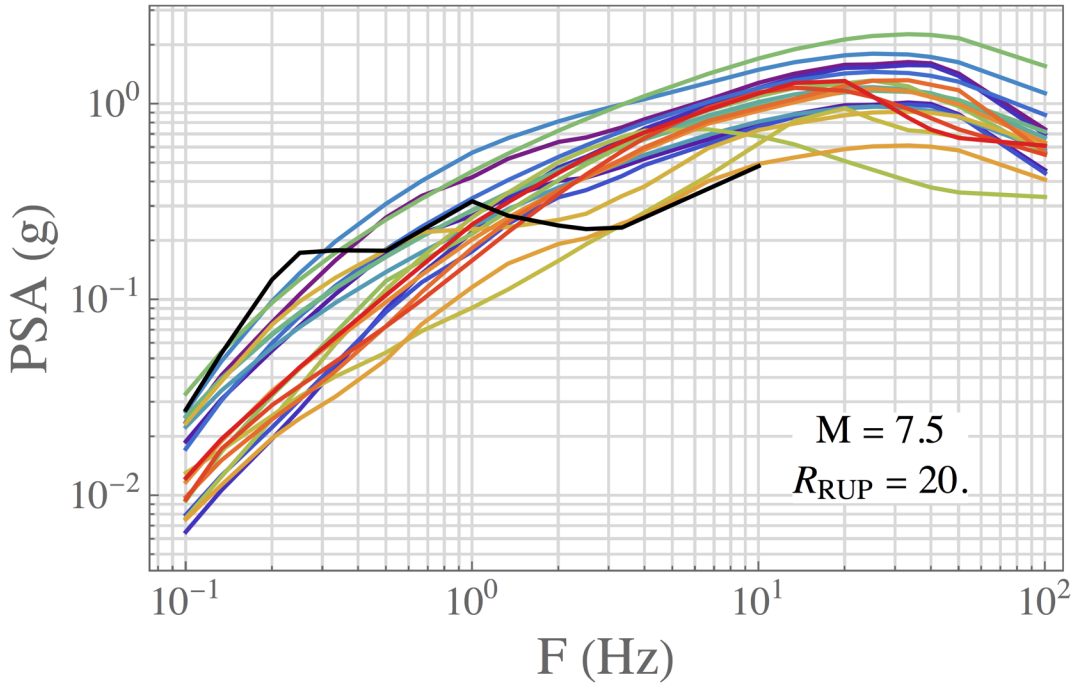


Figure 7–14 **5%-damped PSA for candidate GMMs at M =7.5, R = 50 km (see legend on Figure 7–1).**

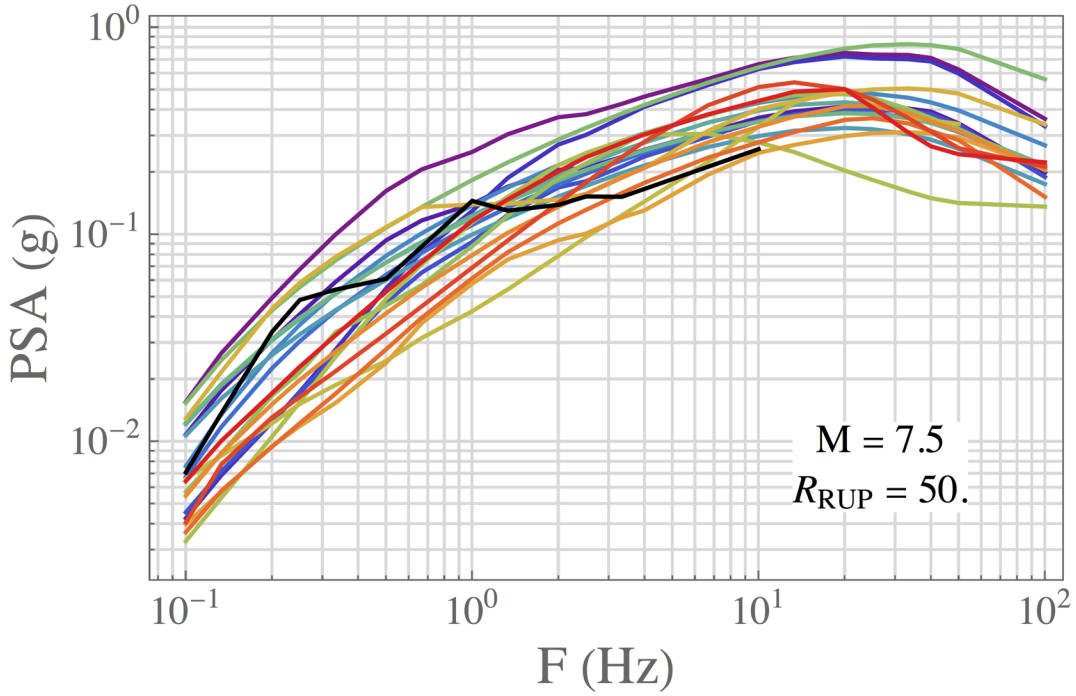


Figure 7–15 **5%-damped PSA for seed GMMs at M = 7.5, R =100 km. (see legend on Figure 7–1).**

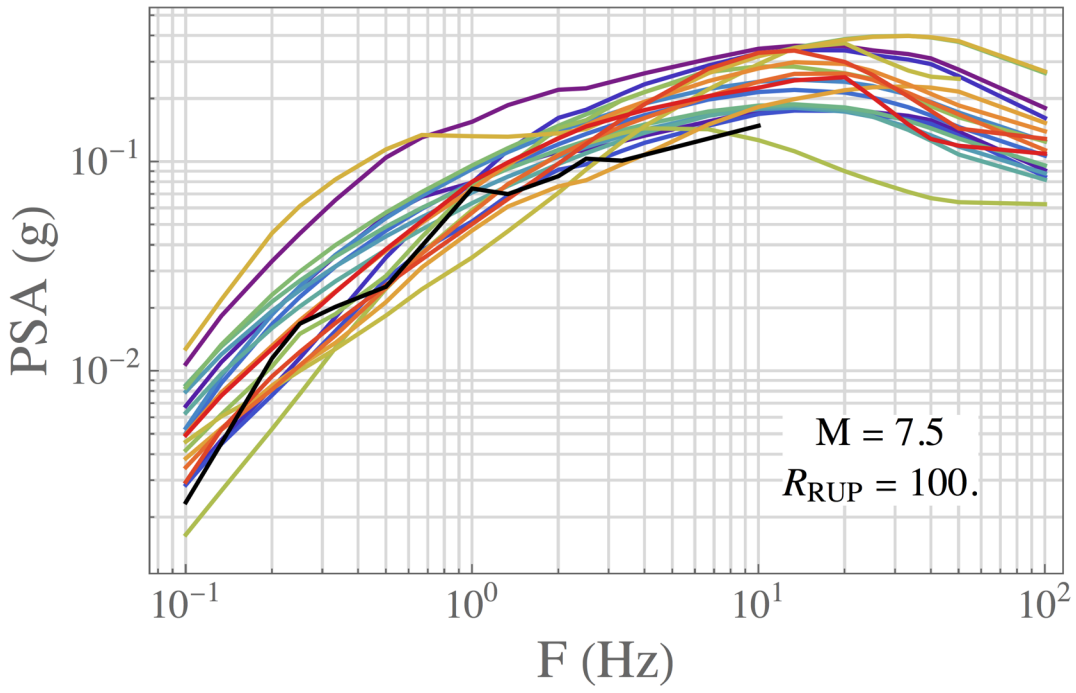


Figure 7–16 5%-damped PSA for seed GMMs at M = 7.5, R = 200 km (see legend on Figure 7–1).

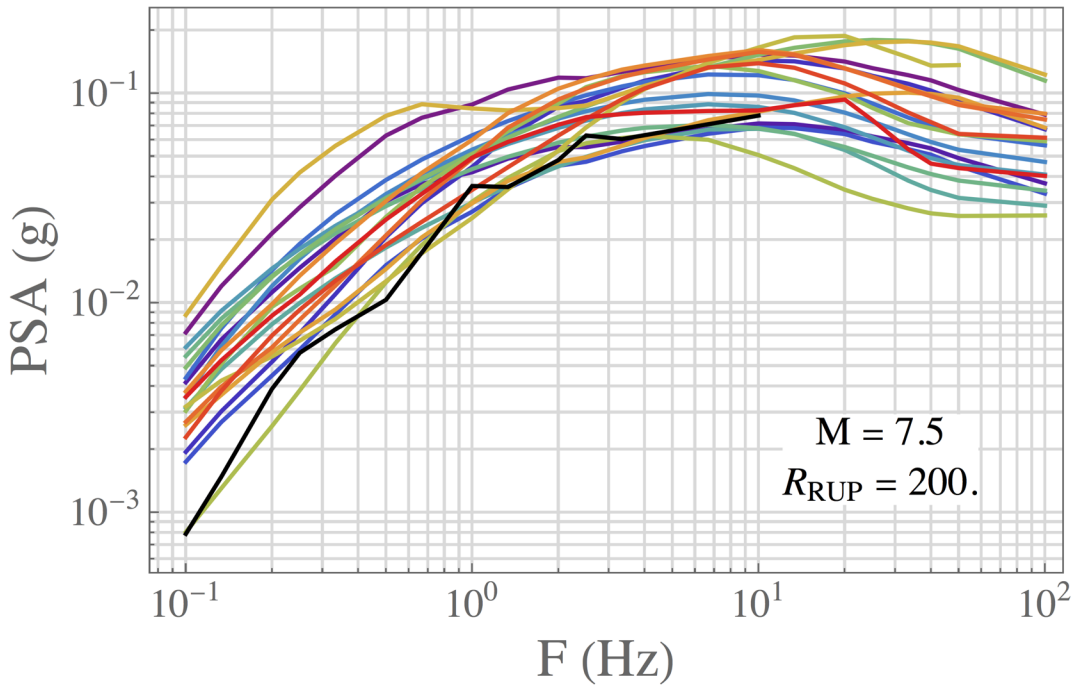


Figure 7–17 Plots of 5% damped PSA for B_a04 GMM at distances of 20 km (upper left), 50 km (lower left), 100 km (upper right), and 200 km (lower right). In each panel, the response is shown for magnitudes M = 4.5, 5.5, 6.5, and 7.5, as indicated by the labels and line sizes. The red curves are for B_a04 GMM. The grey curves are the response determined by averaging over all 20 GMMs under consideration.

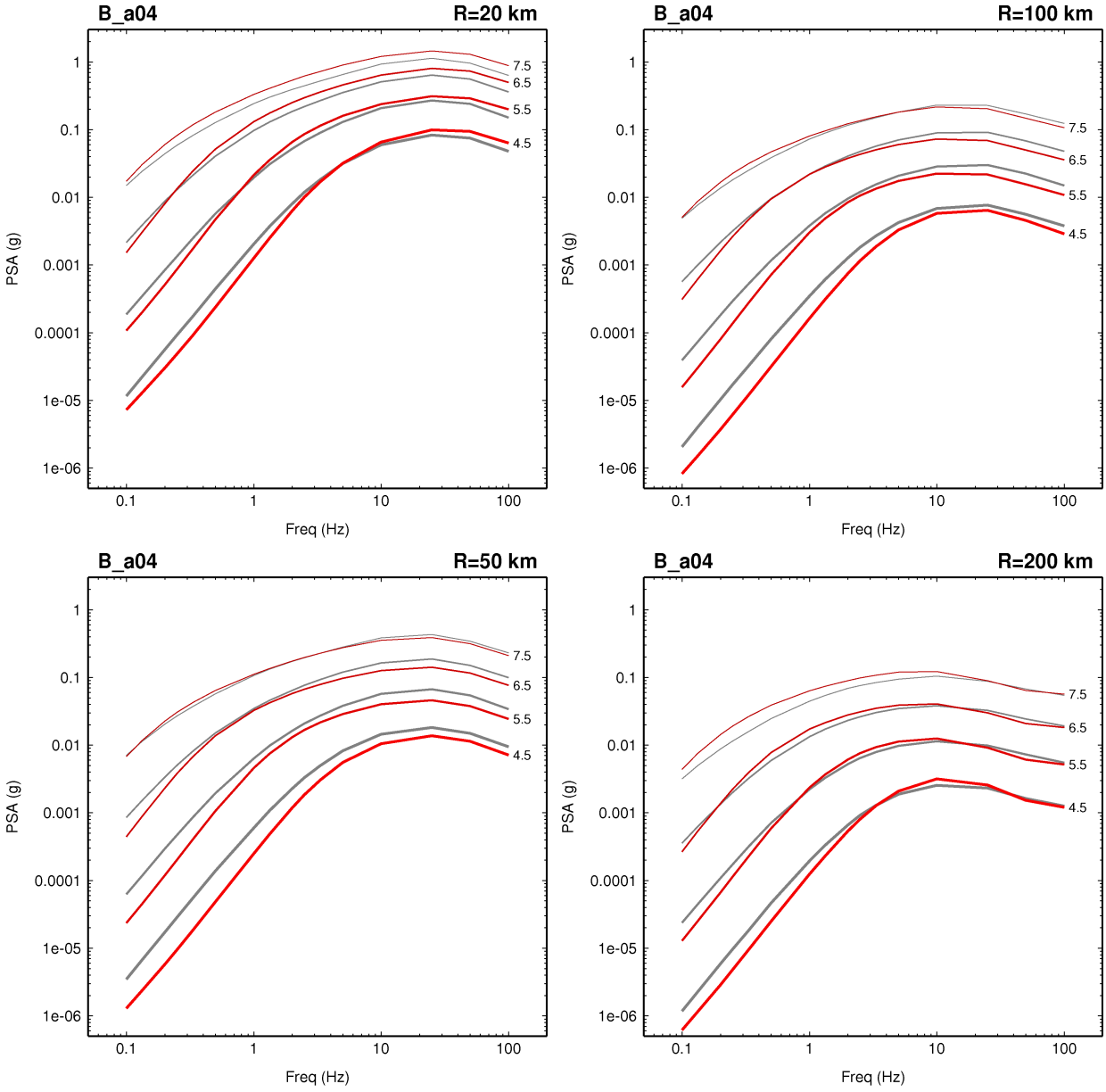


Figure 7–18 B_ab14 GMM. Format is same as Figure 7–17.

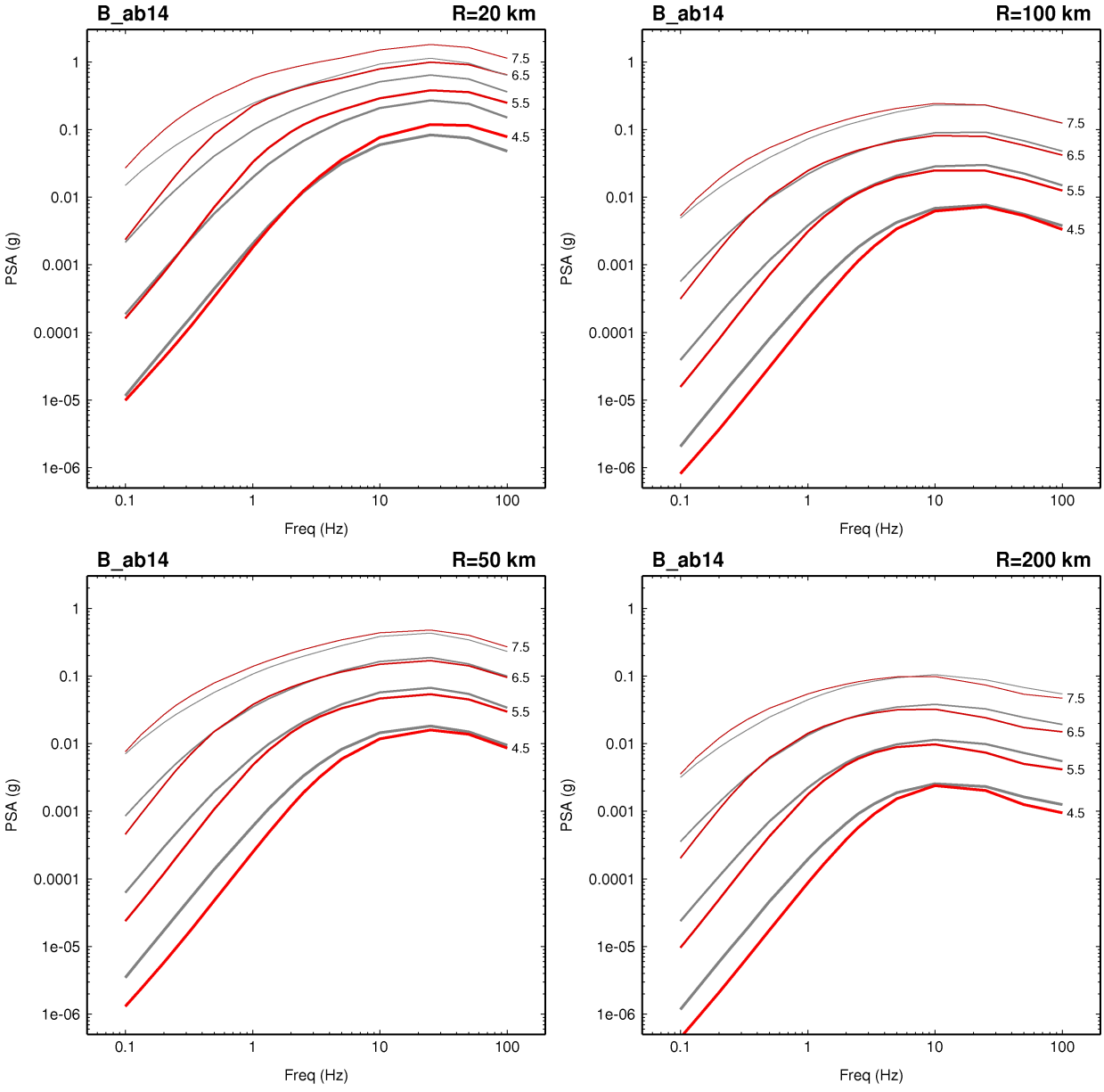


Figure 7–19 B_ab95 GMM. Format is same as Figure 7–17.

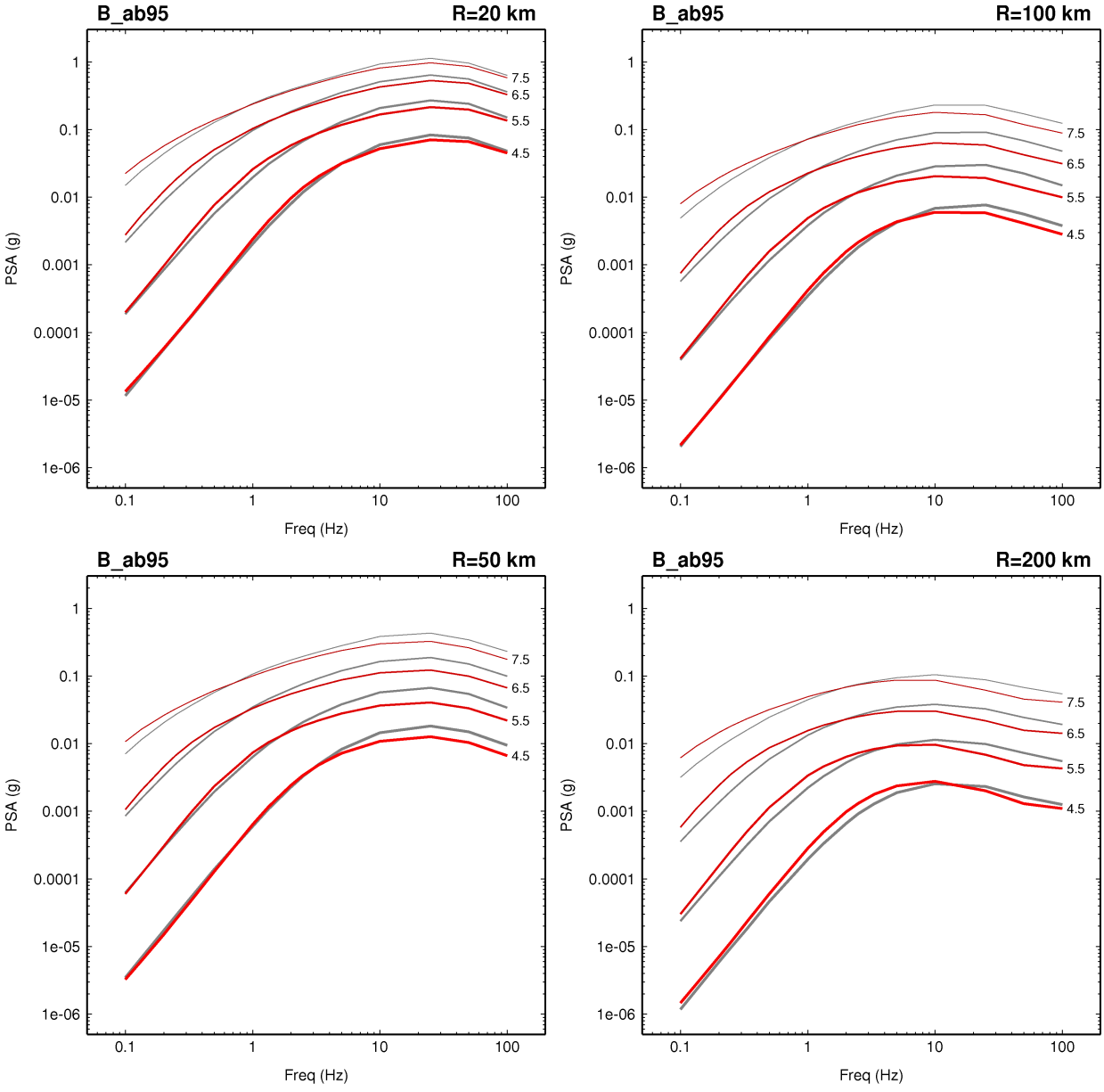


Figure 7–20 B_bca10d GMM. Format is same as Figure 7–17.

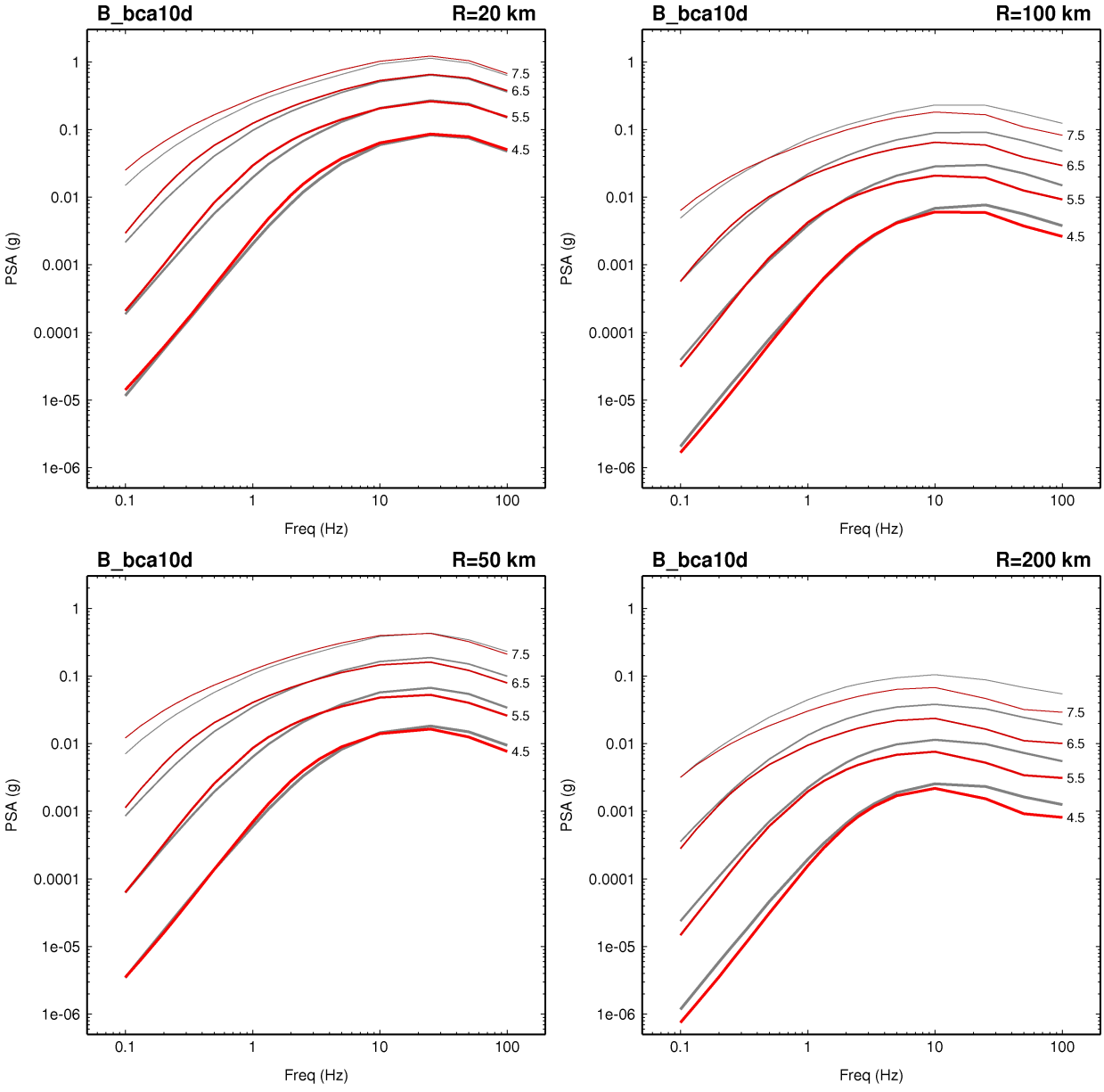


Figure 7–21 B_bs11 GMM. Format is same as Figure 7–17.

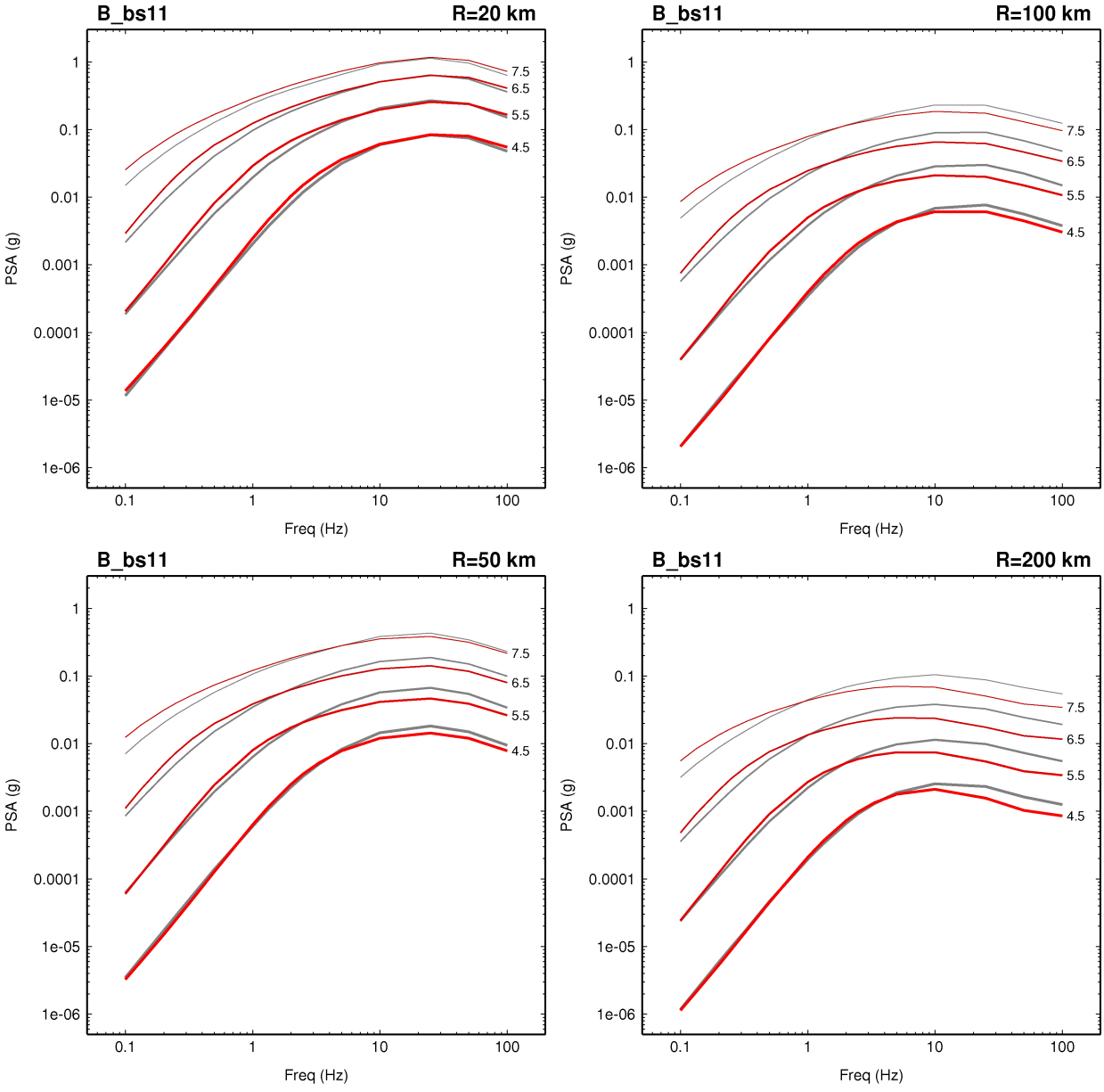


Figure 7–22 B_sgd02 GMM. Format is same as Figure 7–17.

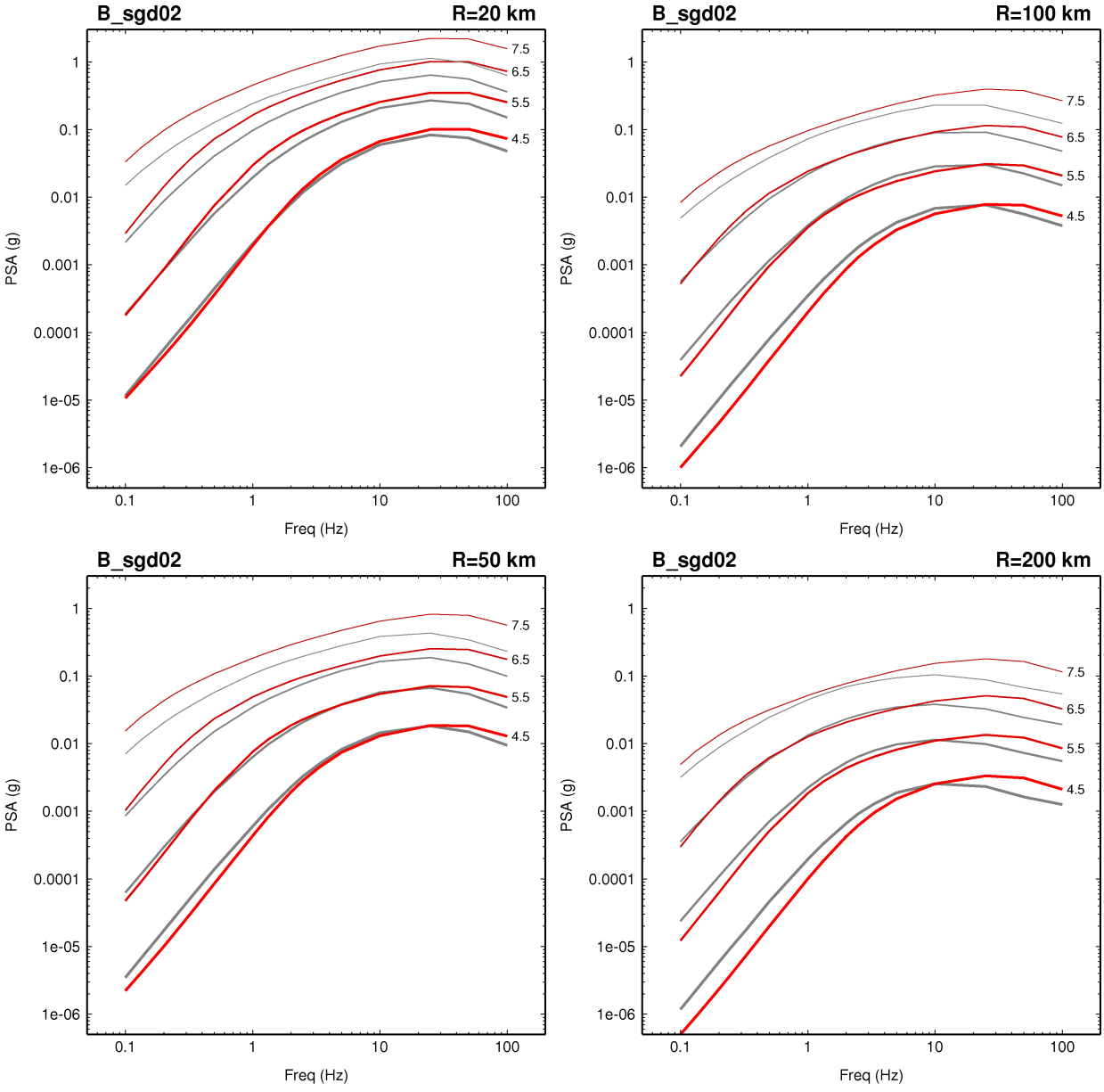


Figure 7–23 1CCSP GMM. Format is same as Figure 7–17.

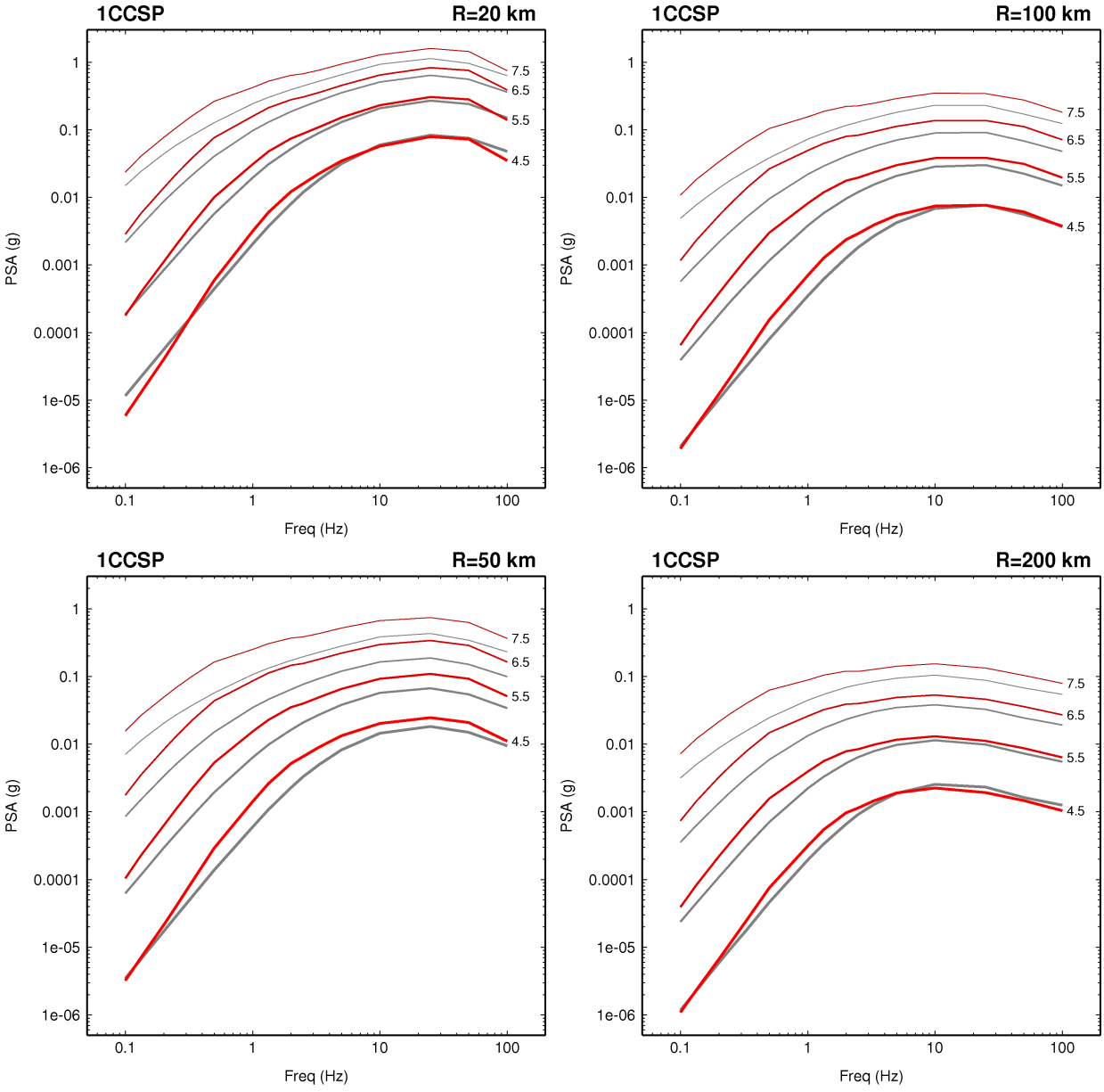


Figure 7–24 1CVSP GMM. Format is same as Figure 7–17.

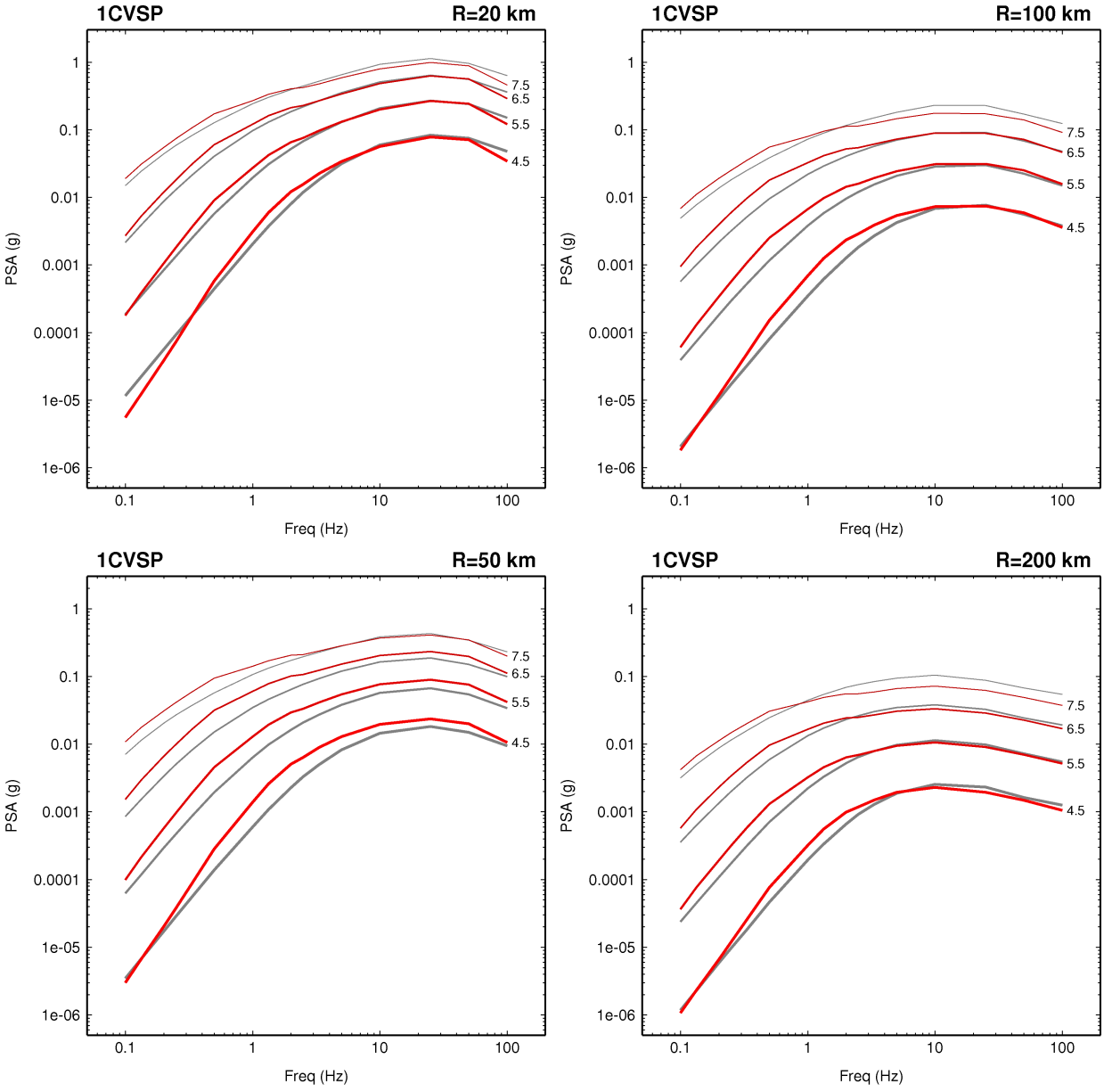


Figure 7–25 2CCSP GMM. Format is same as Figure 7–17.

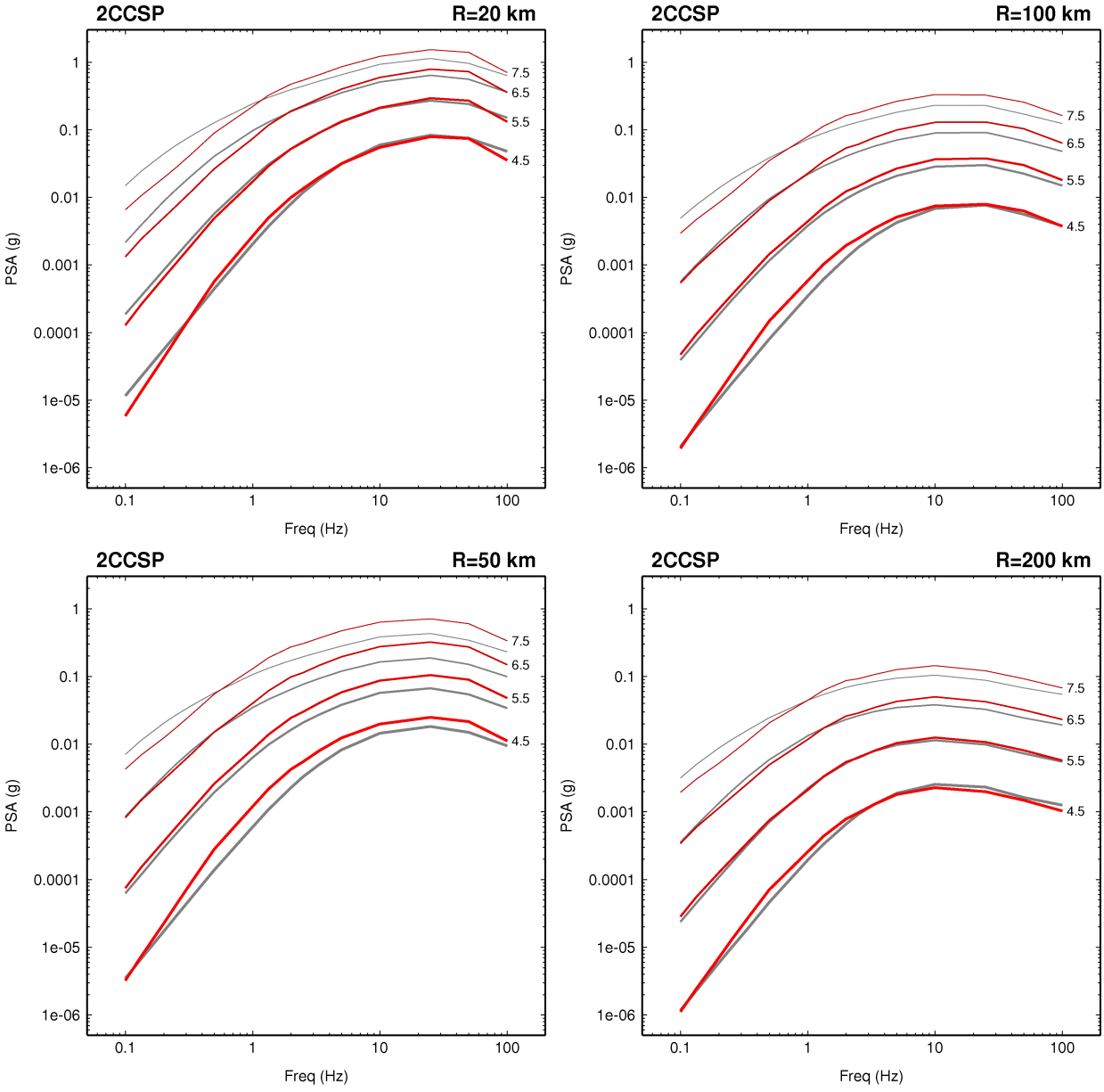


Figure 7–26 2CVSP GMM. Format is same as Figure 7–17.

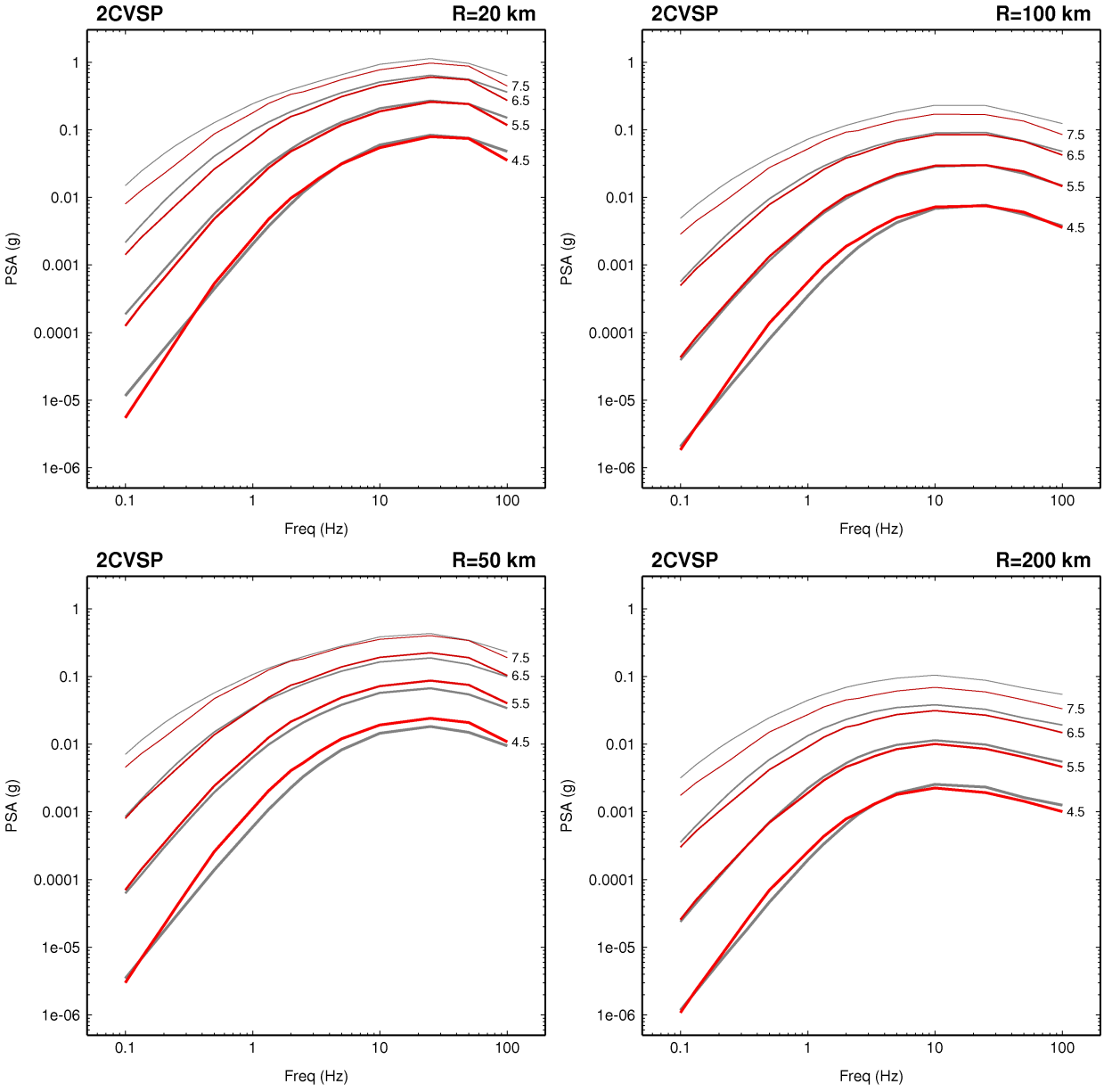


Figure 7–27 YA15 GMM. Format is same as Figure 7–17.

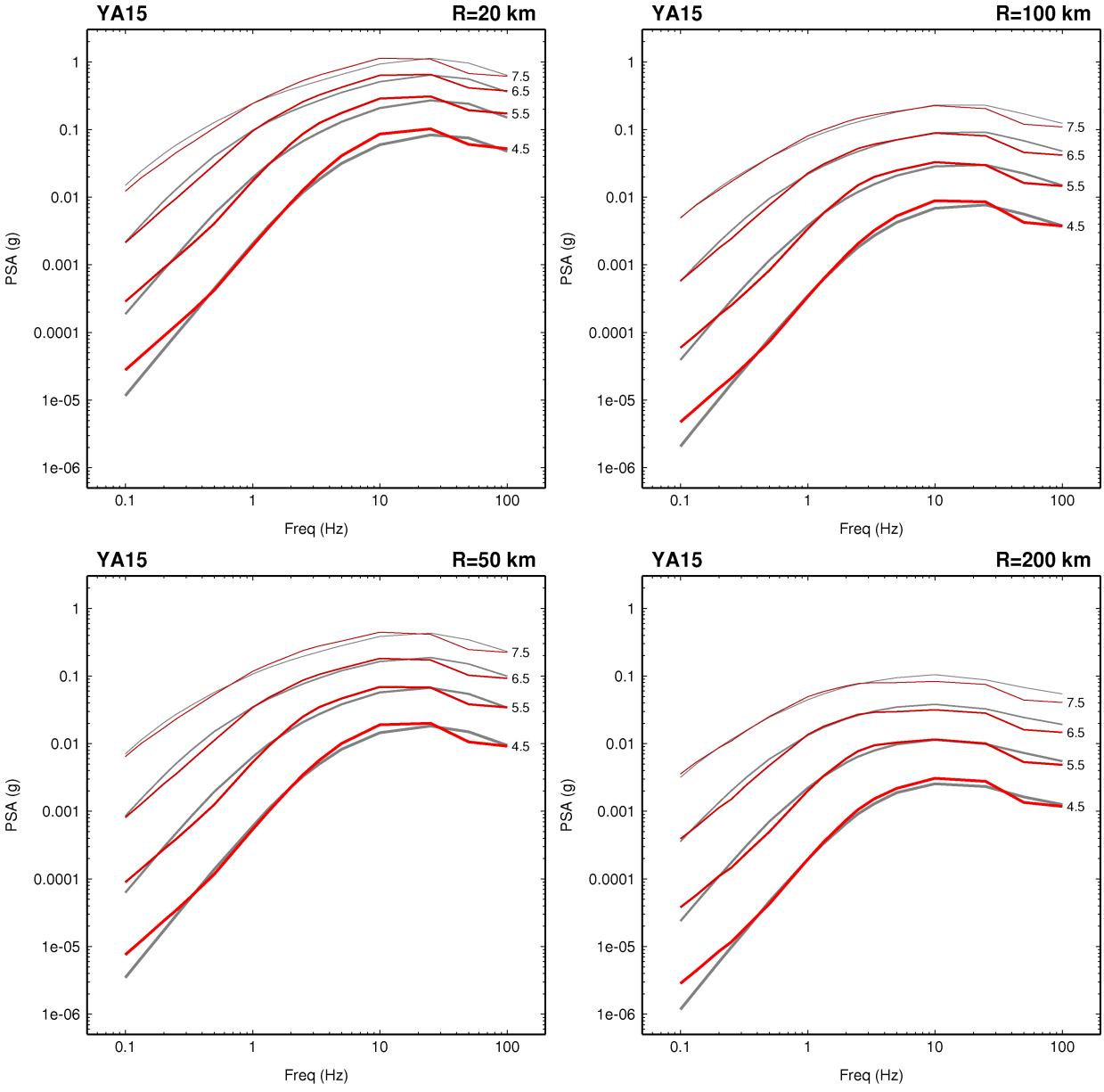


Figure 7–28 PZCT15_M1SS GMM. Format is same as Figure 7–17.

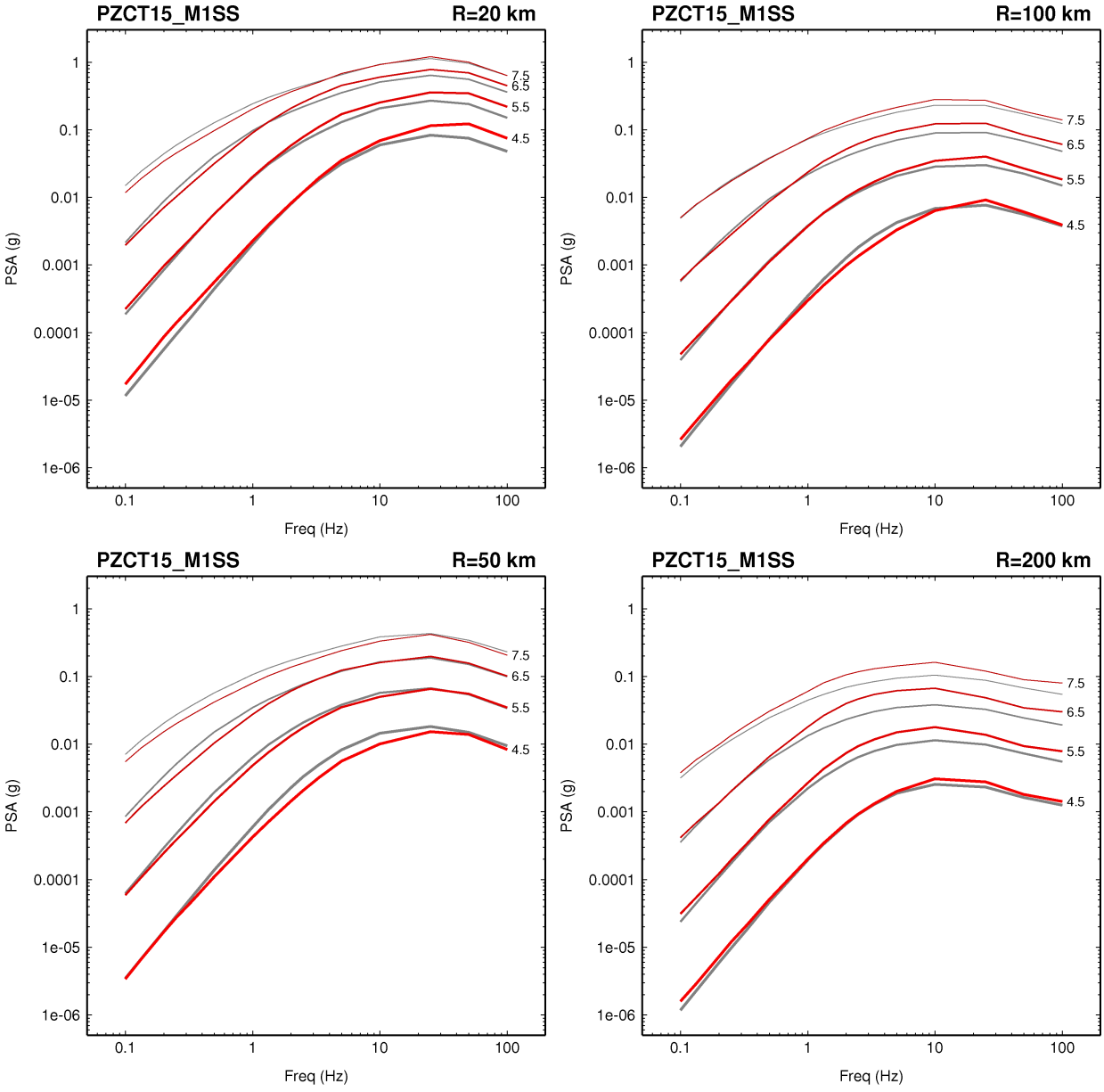


Figure 7–29 PZCT15_M2ES GMM. Format is same as Figure 7–17.

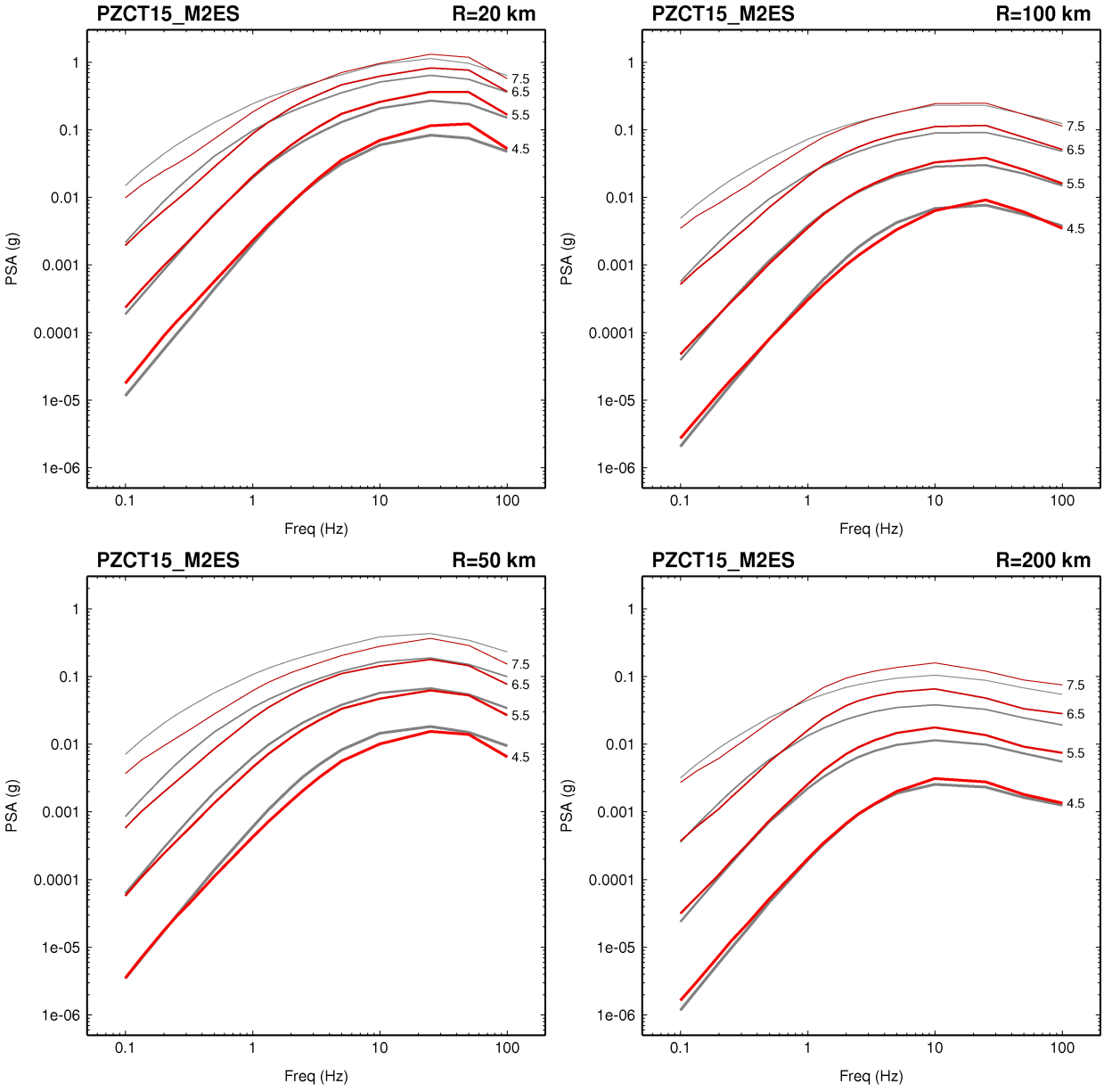


Figure 7–30 Frankel GMM. Format is same as Figure 7–17.

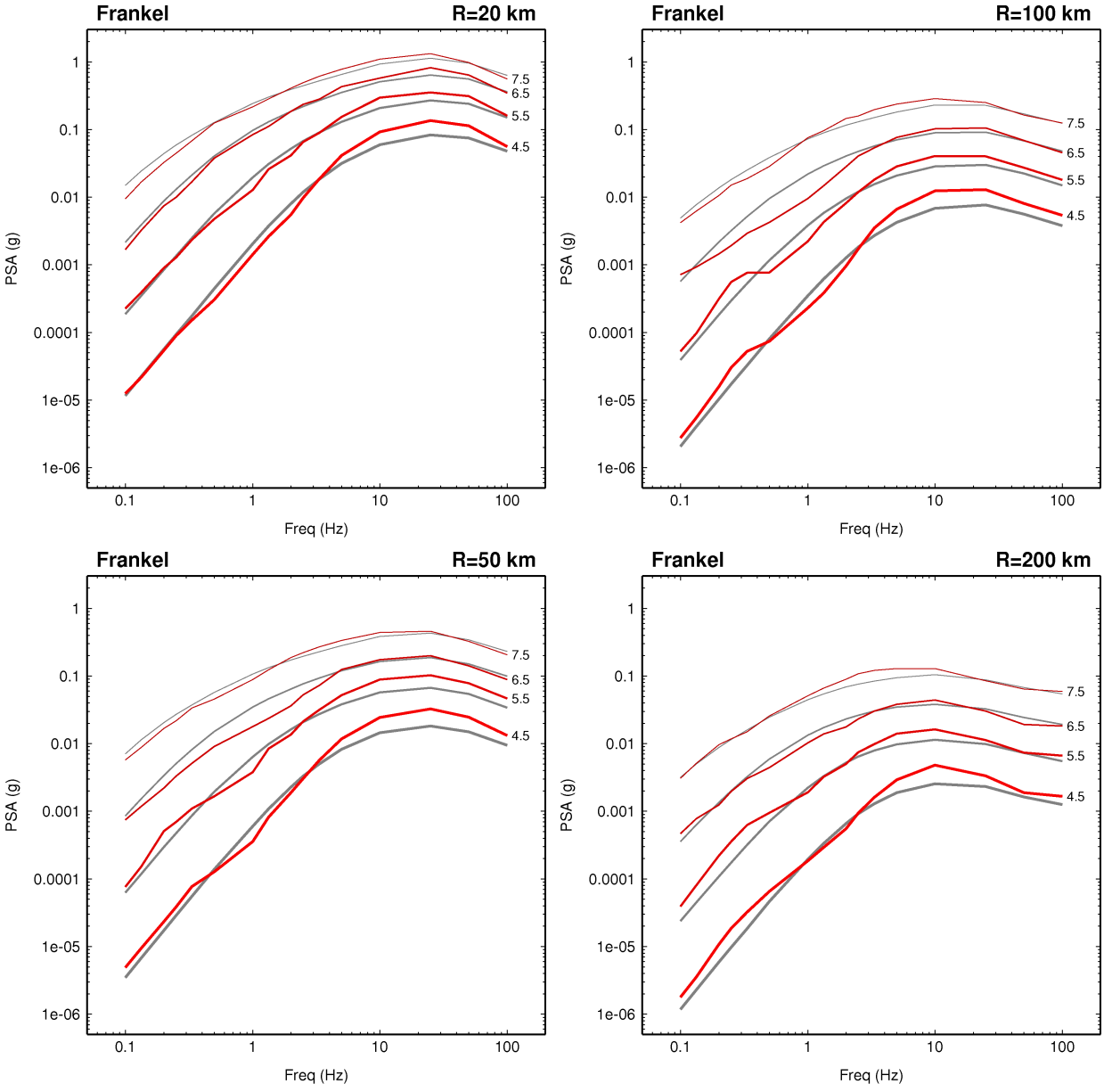


Figure 7–31 SP15 GMM. Format is same as Figure 7–17.

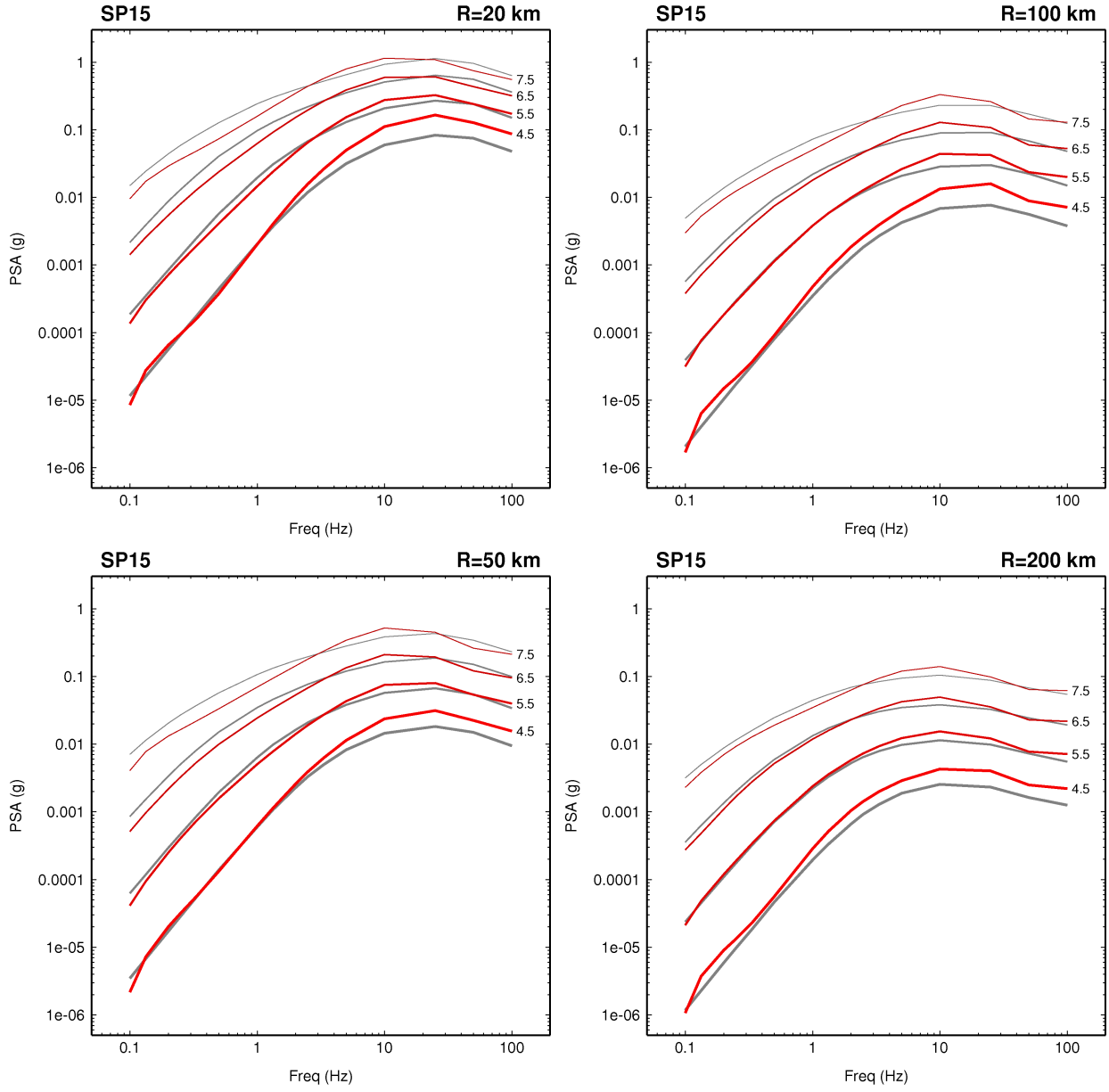


Figure 7–32 ANC15 GMM. Format is same as Figure 7–17.

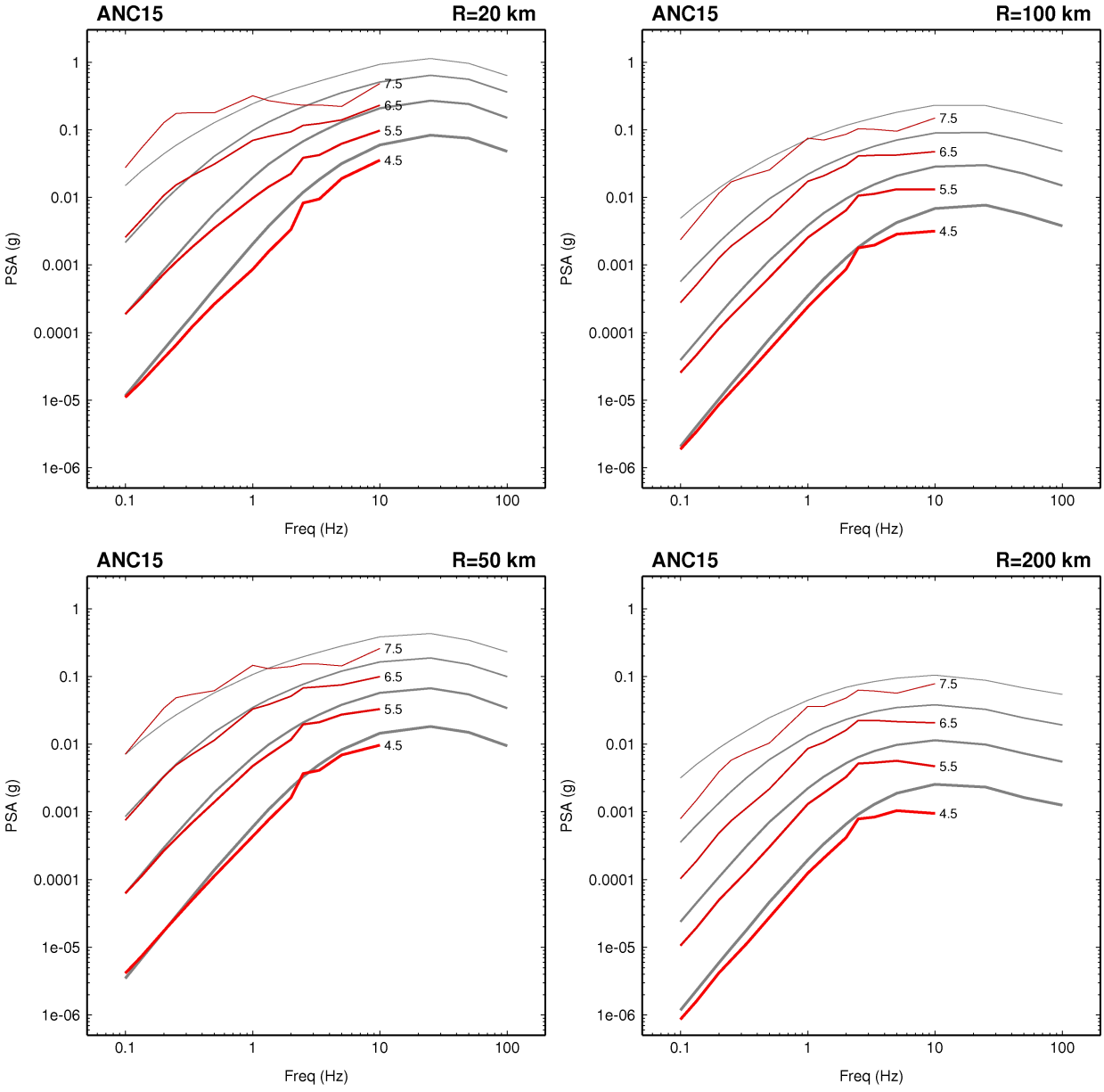


Figure 7–33 Graizer GMM. Format is same as Figure 7–17.

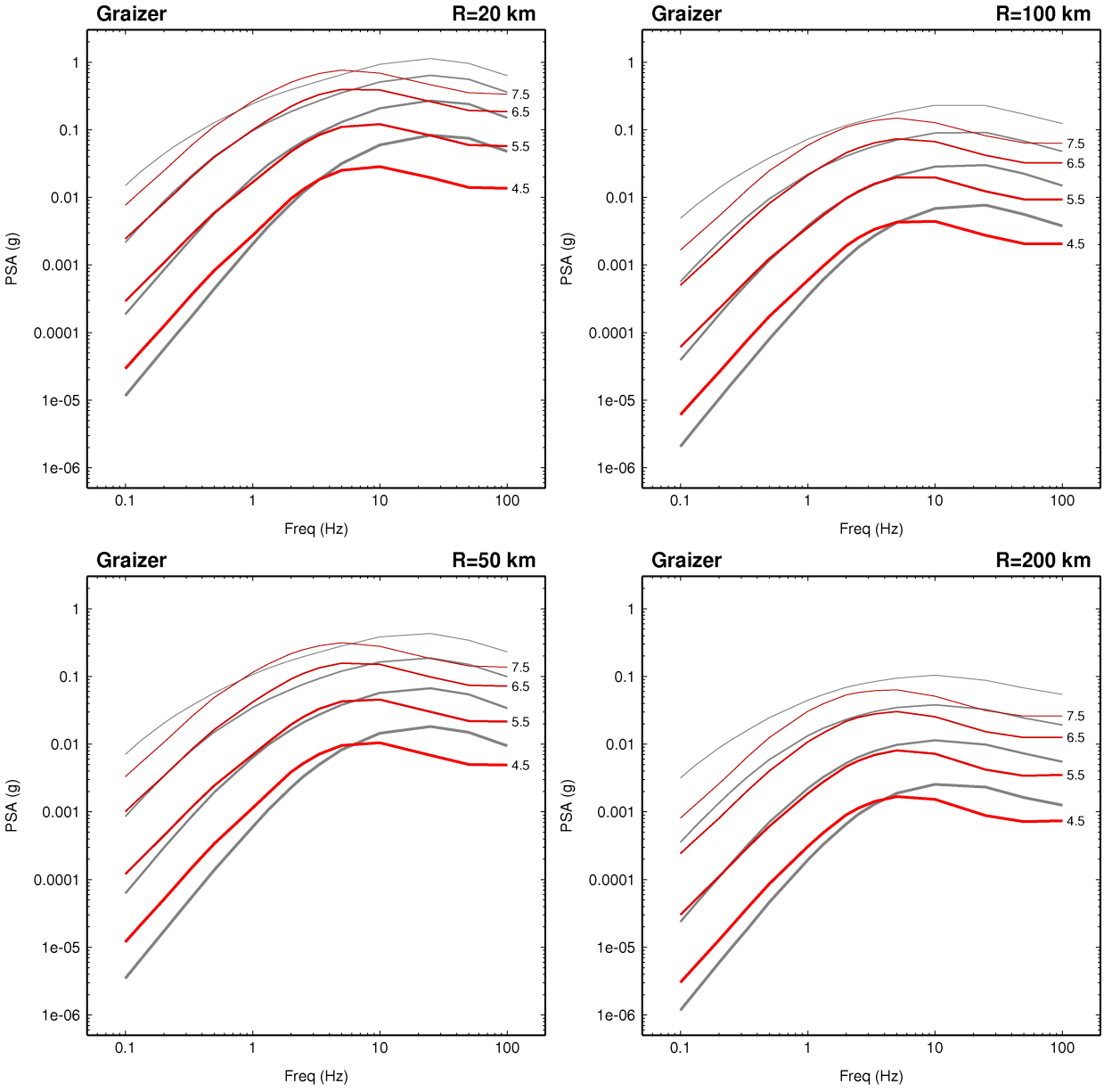


Figure 7–34 HA15 GMM. Format is same as Figure 7–17.

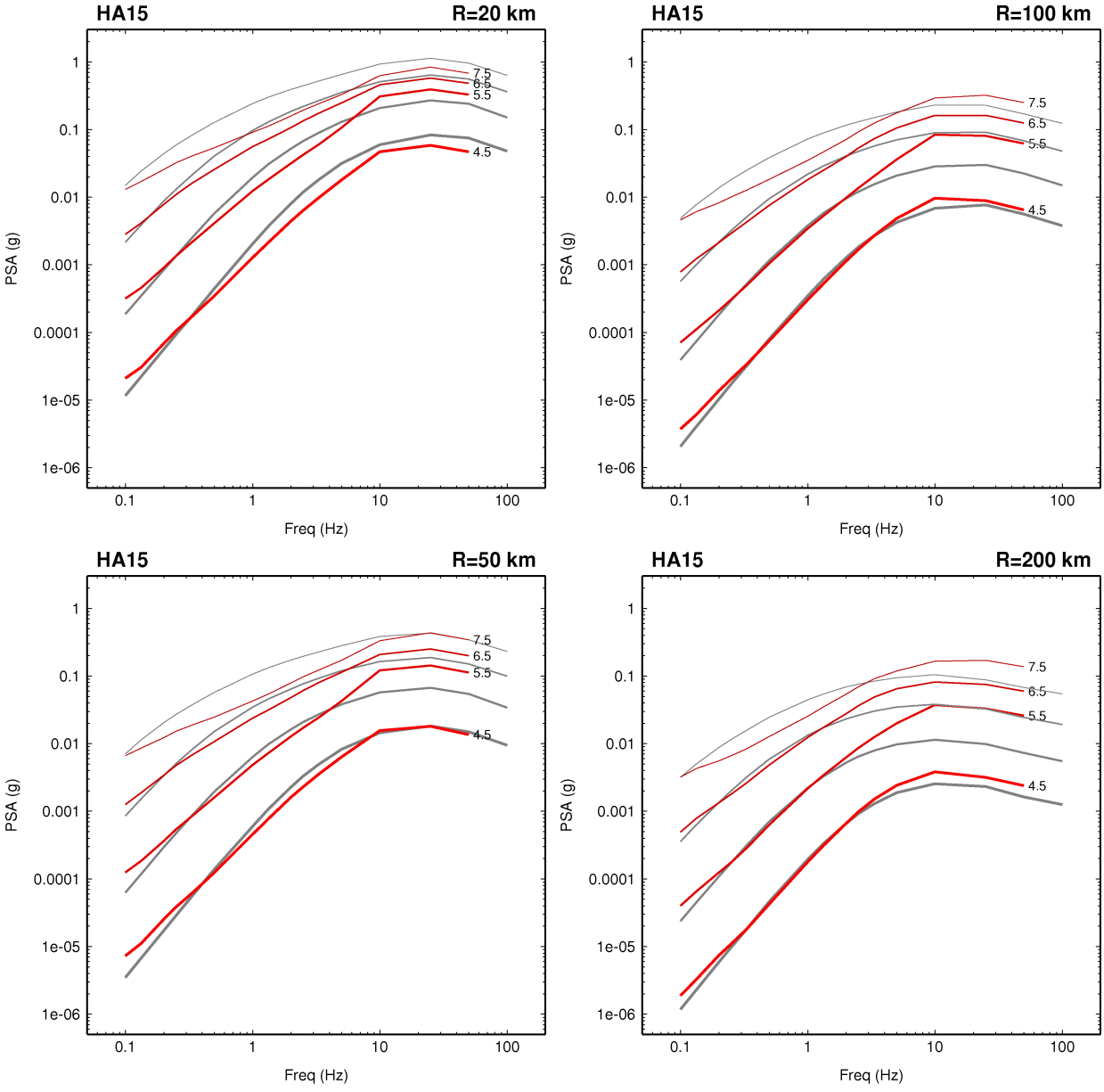


Figure 7–35 PEER_EX GMM. Format is same as Figure 7–17.

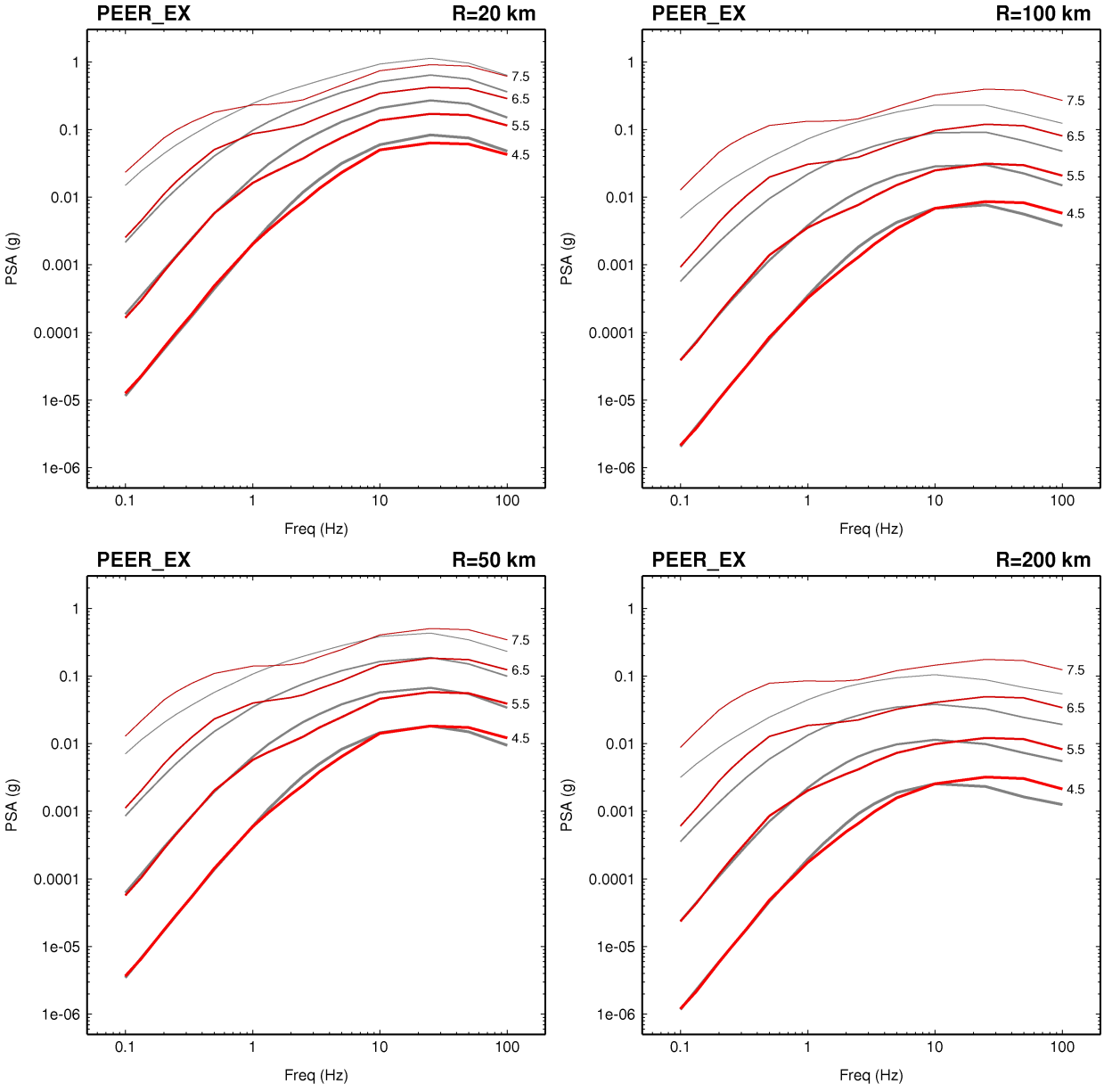


Figure 7–36 PEER_GP GMM. Format is same as Figure 7–17.

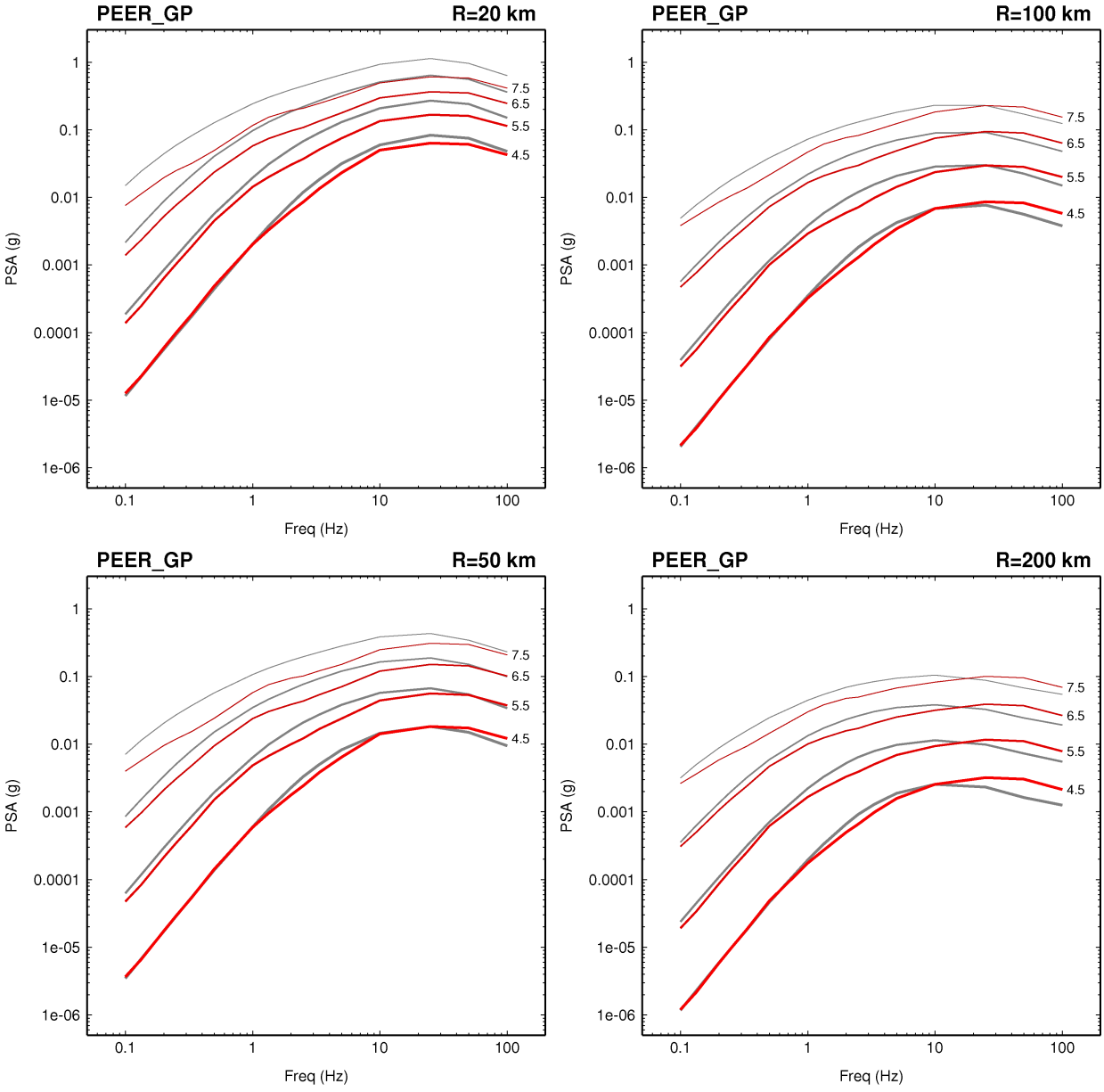


Figure 7–37 Number of recordings versus frequency for the CENA dataset used in the residual assessment of the candidate GMMs.

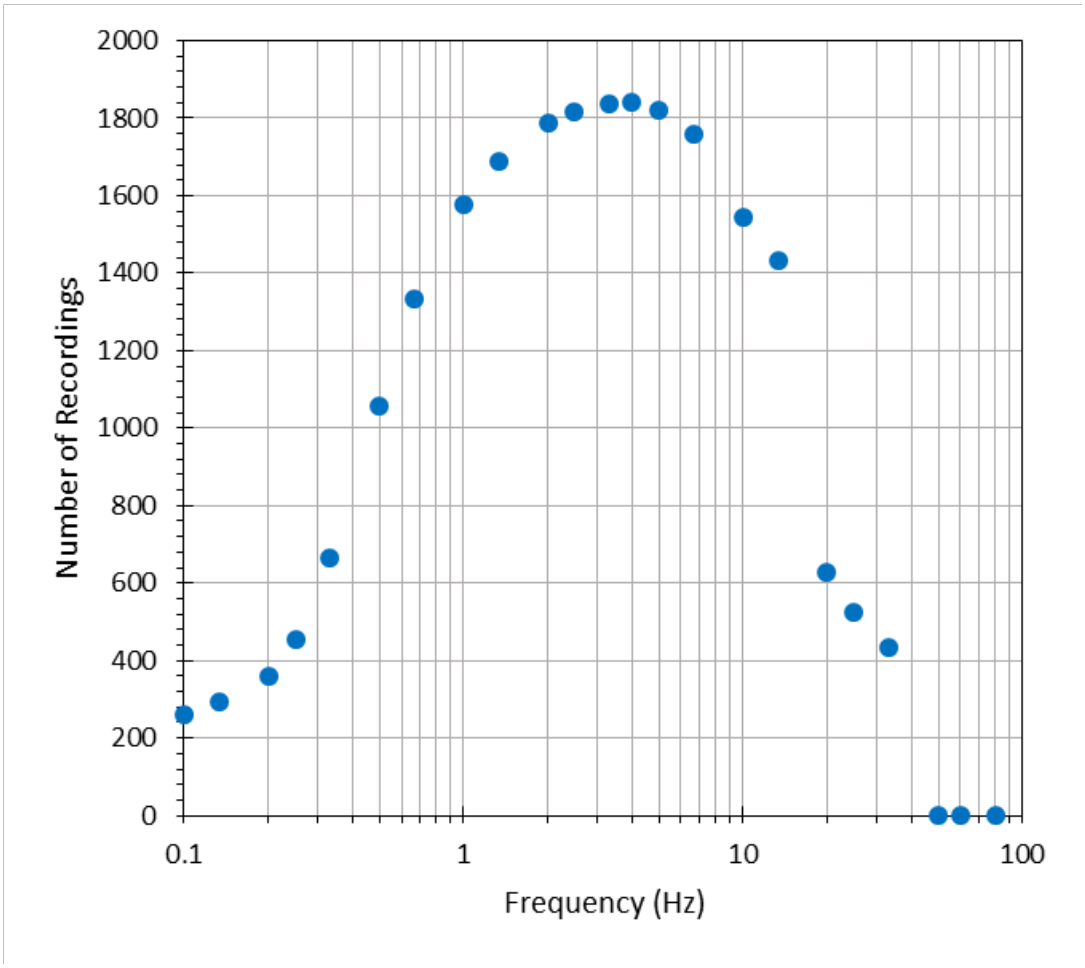


Figure 7–38 Magnitude, R_{RUP} and V_{S30} histograms for the CENA dataset used in the residual assessment of the candidate GMMs.

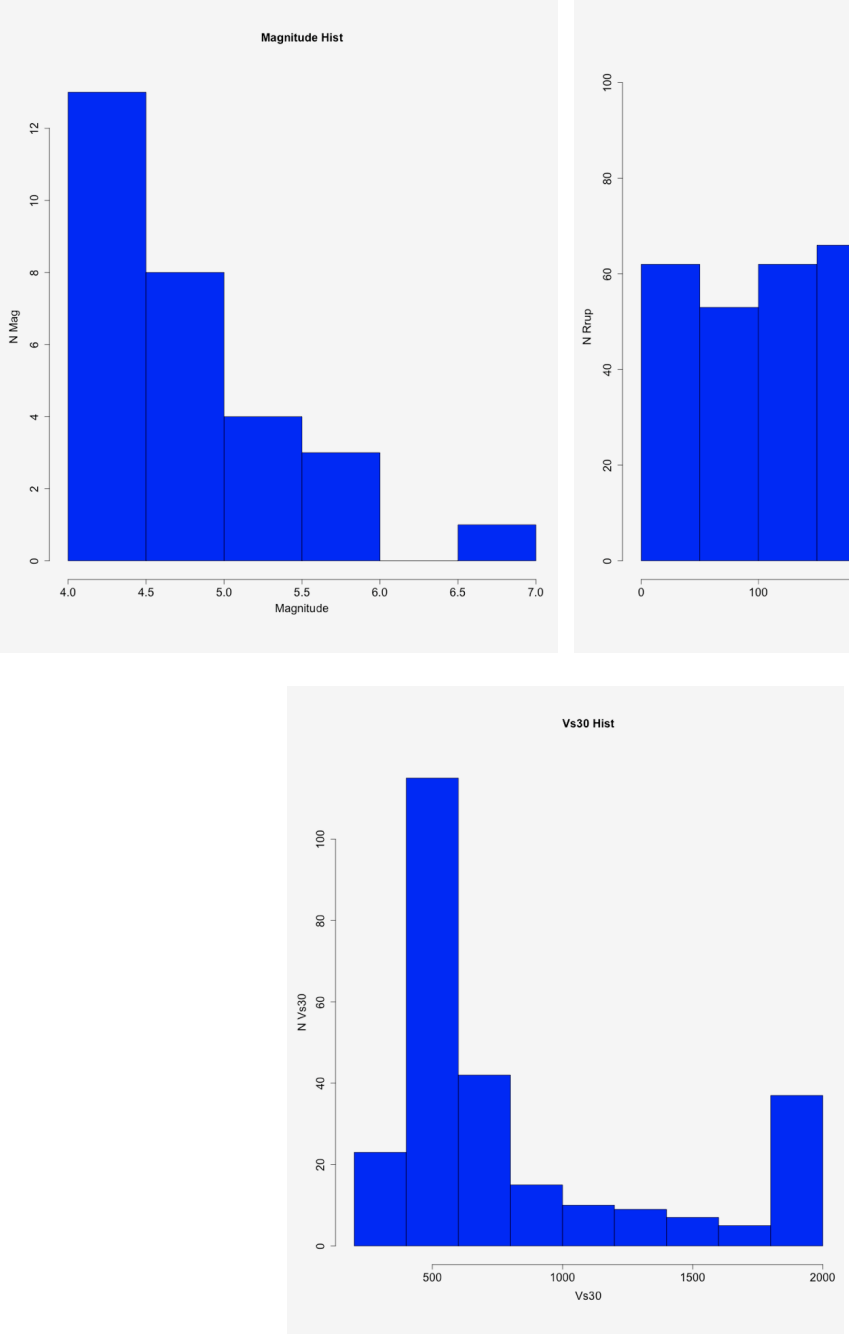


Figure 7–39 Site adjustment factors of V_{S30} = 760 m/sec relative to V_{S30} = 3000 m/sec used in the residual assessment of the candidate GMMs.

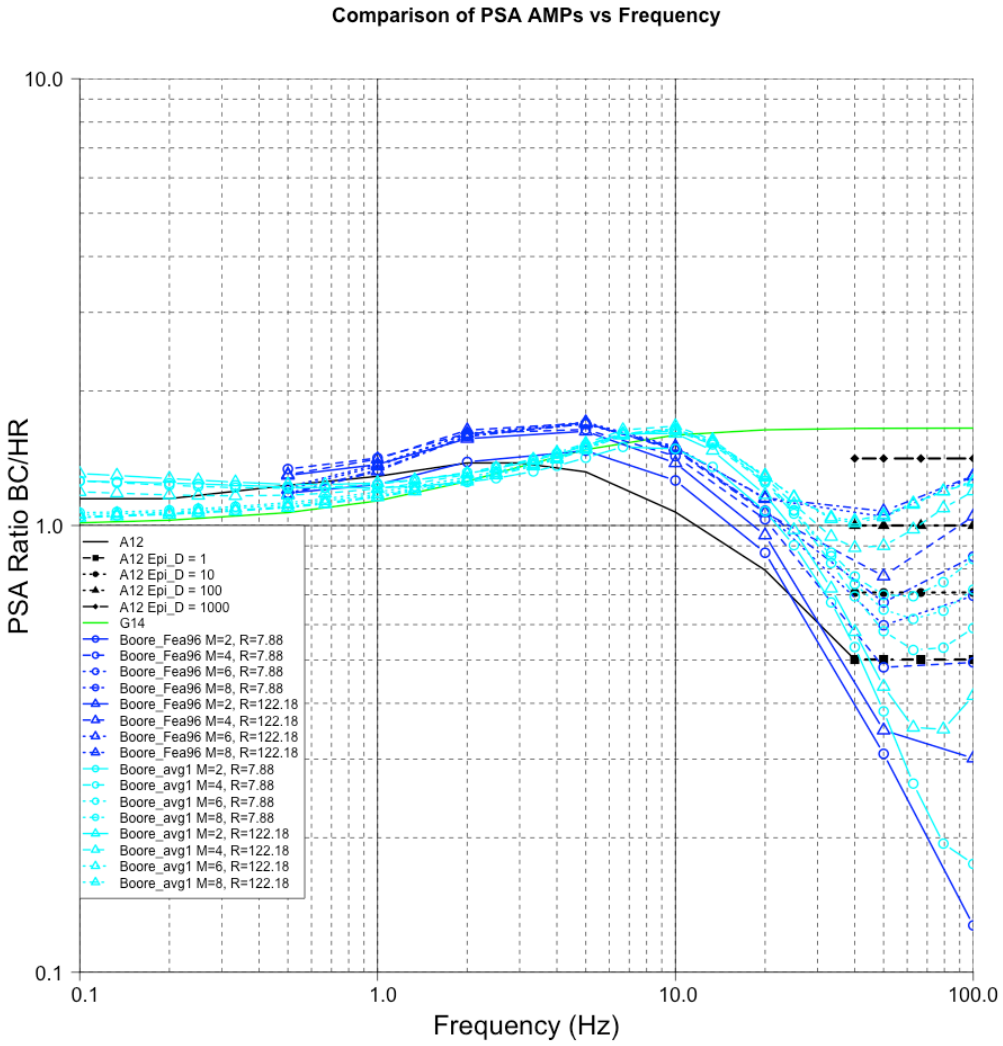


Figure 7–40 Model bias of fit to residuals for B_a04. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

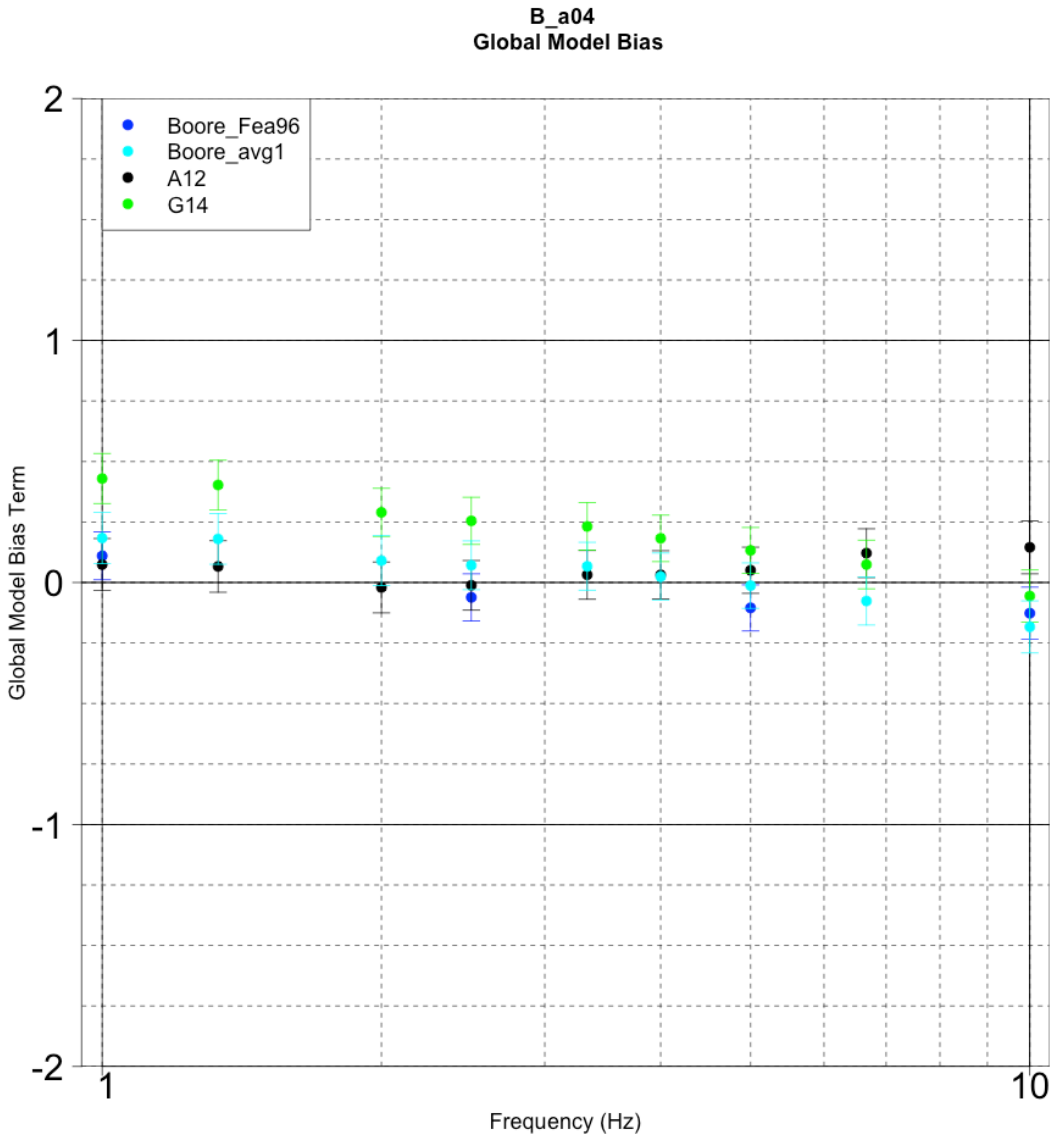


Figure 7–41 Model bias of fit to residuals for B_ab14. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

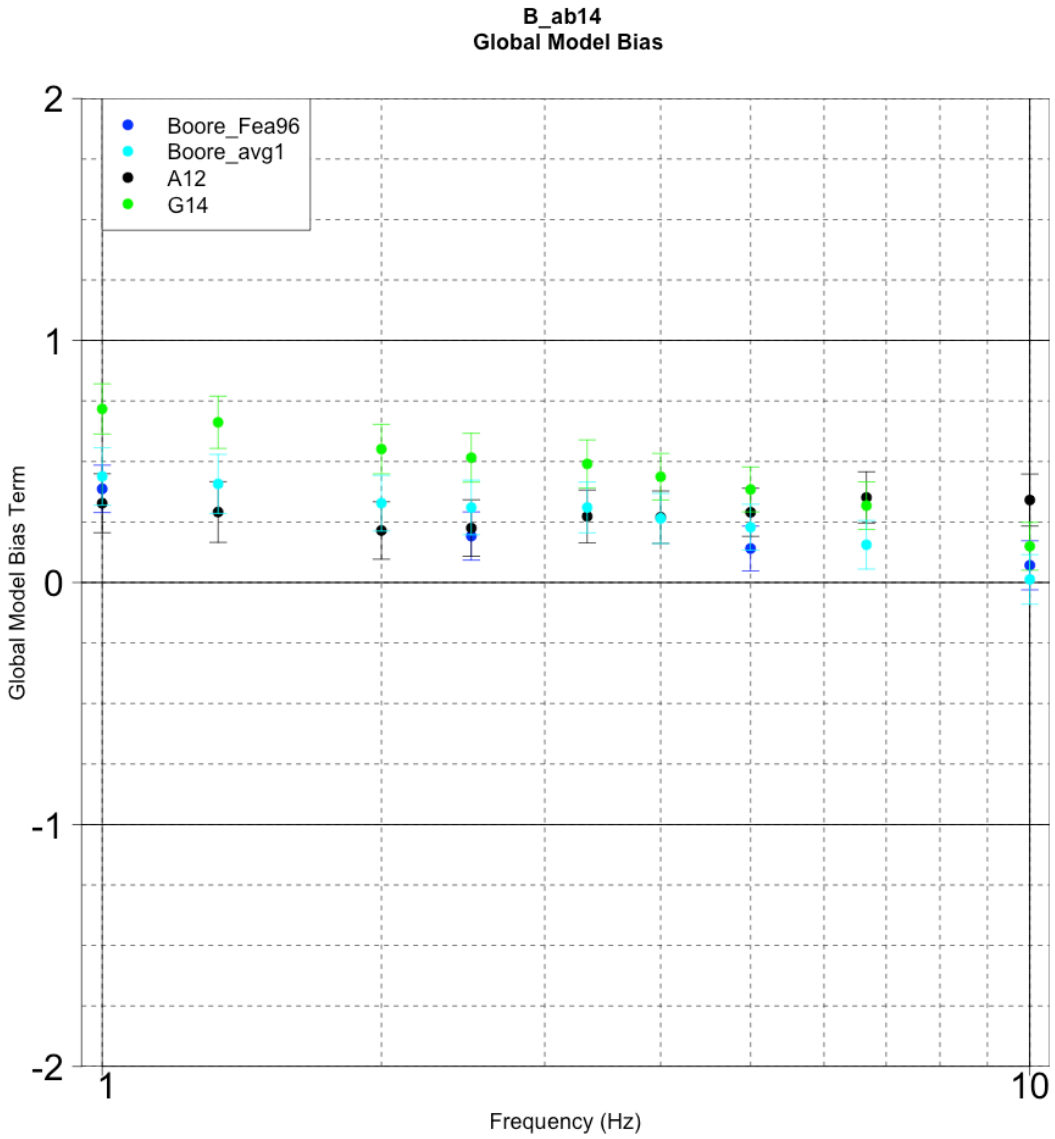


Figure 7–42 Model bias of fit to residuals for B_ab95. The different color symbols indicate the four different site adjustment models used to correct the data to the Vs30 = 3000 m/sec site condition.

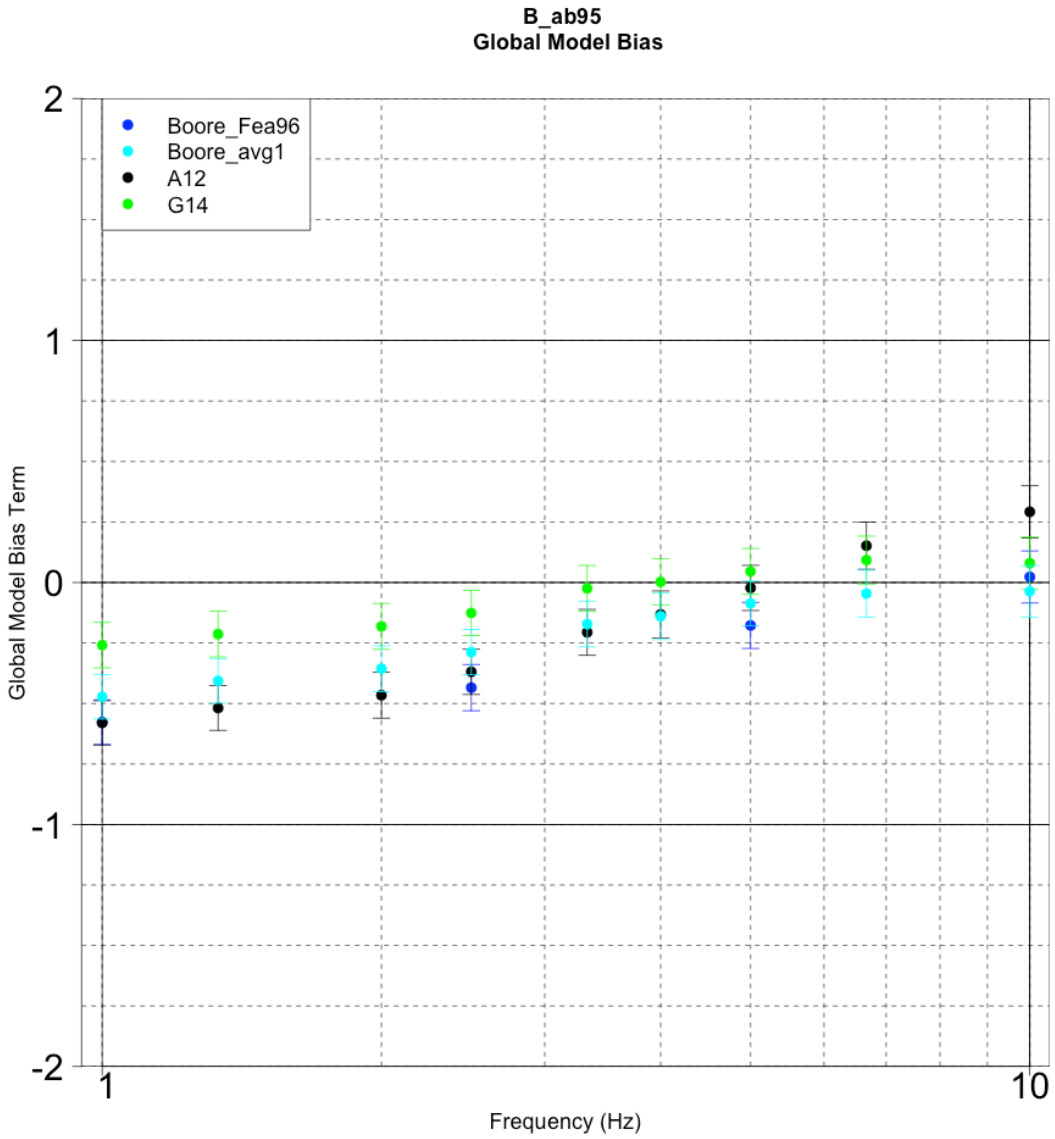


Figure 7–43 Model bias of fit to residuals for B_bca10d. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

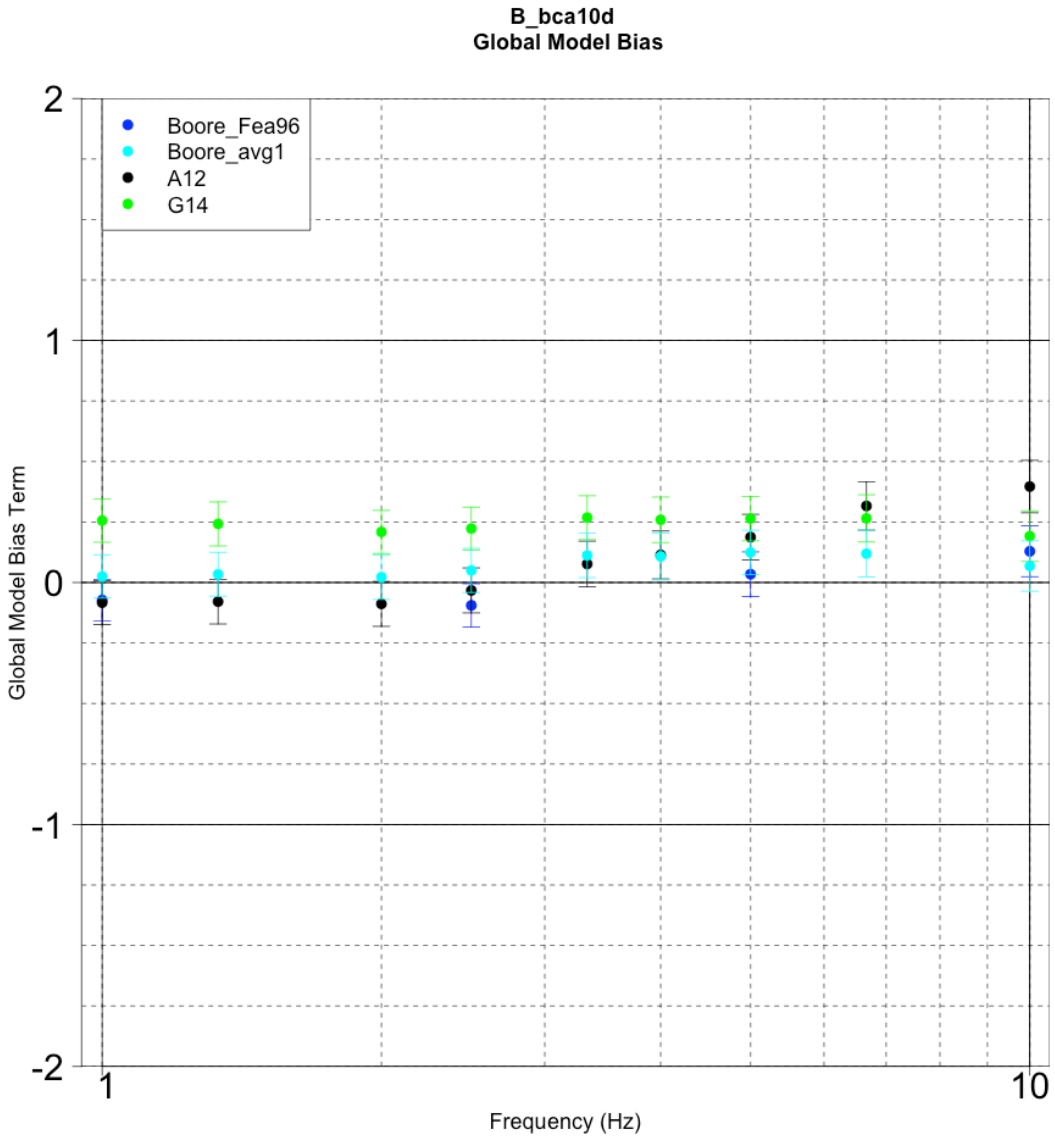


Figure 7–44 Model bias of fit to residuals for B_bs11. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

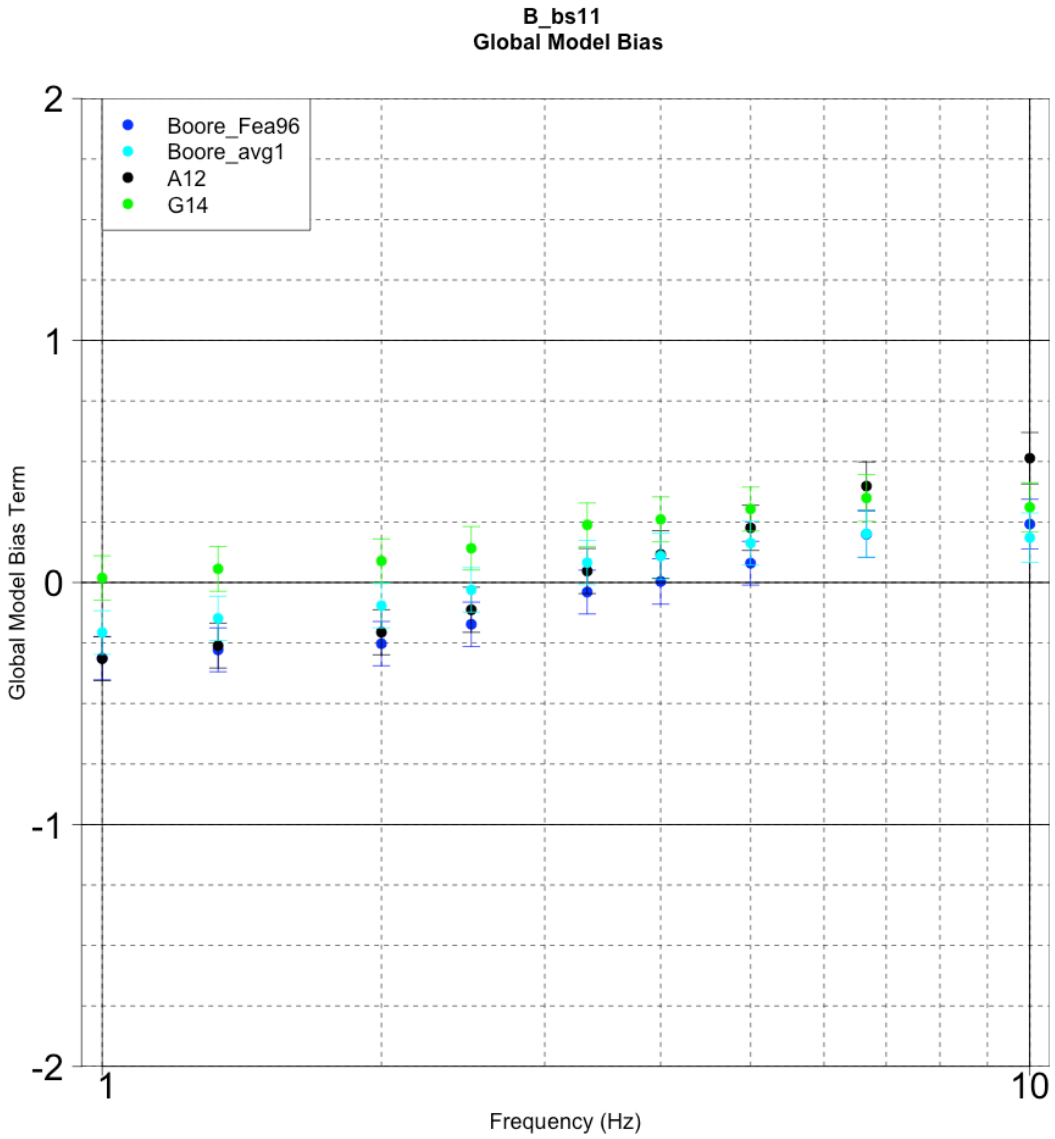


Figure 7–45 Model bias of fit to residuals for B_sgd02. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

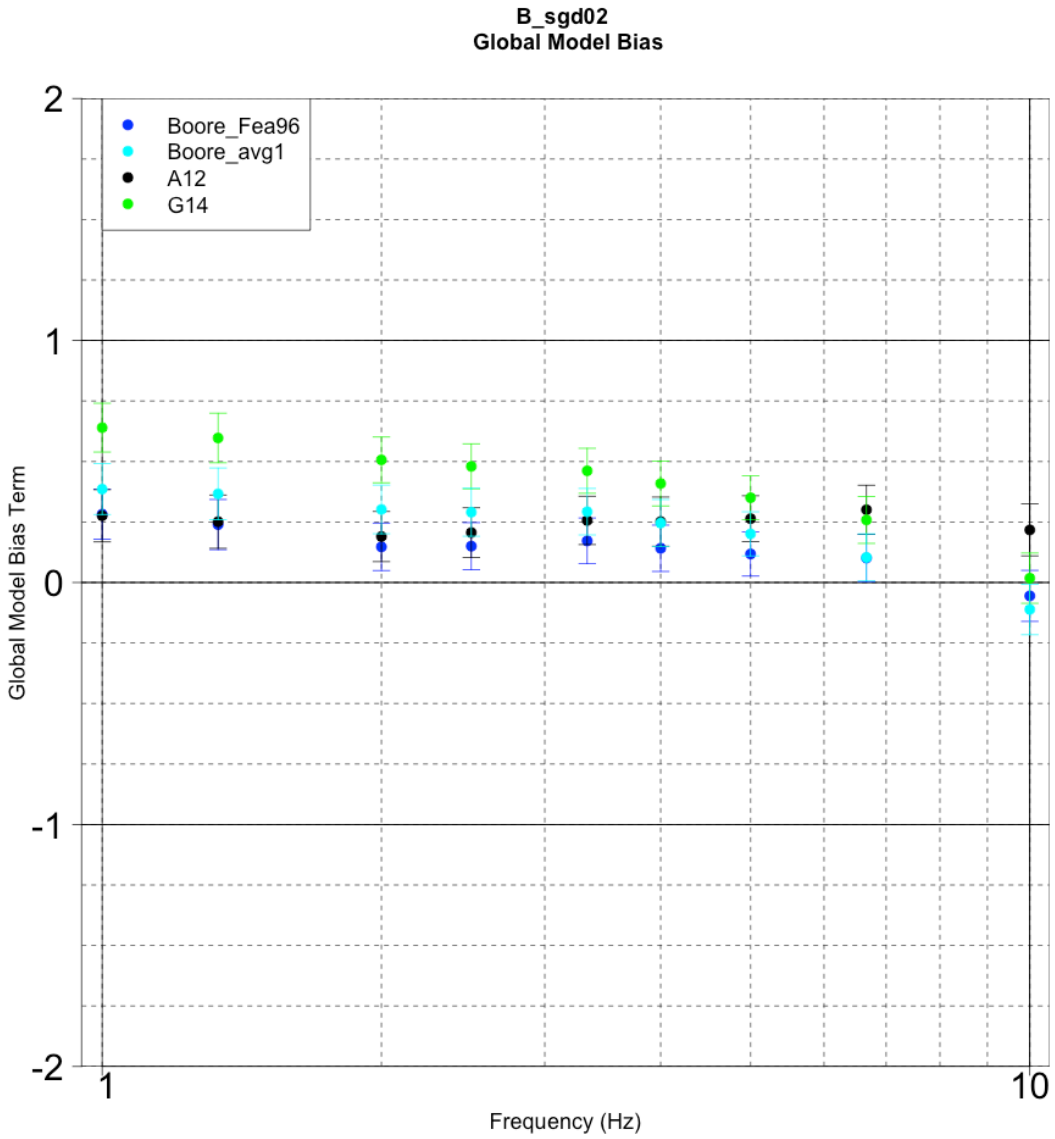


Figure 7–46 Model bias of fit to residuals for 1CCSP. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

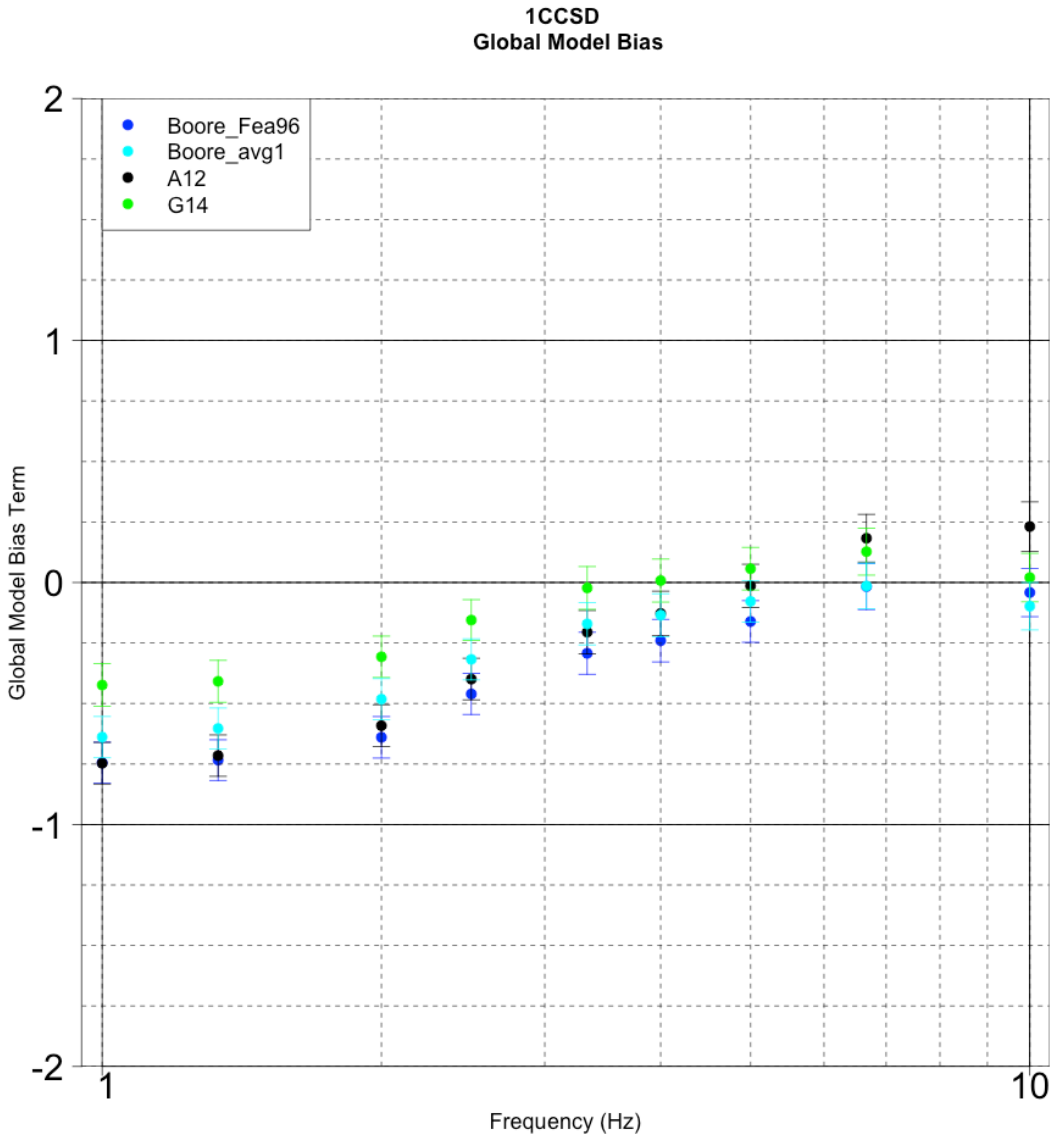


Figure 7–47 Model bias of fit to residuals for 1CVSP. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

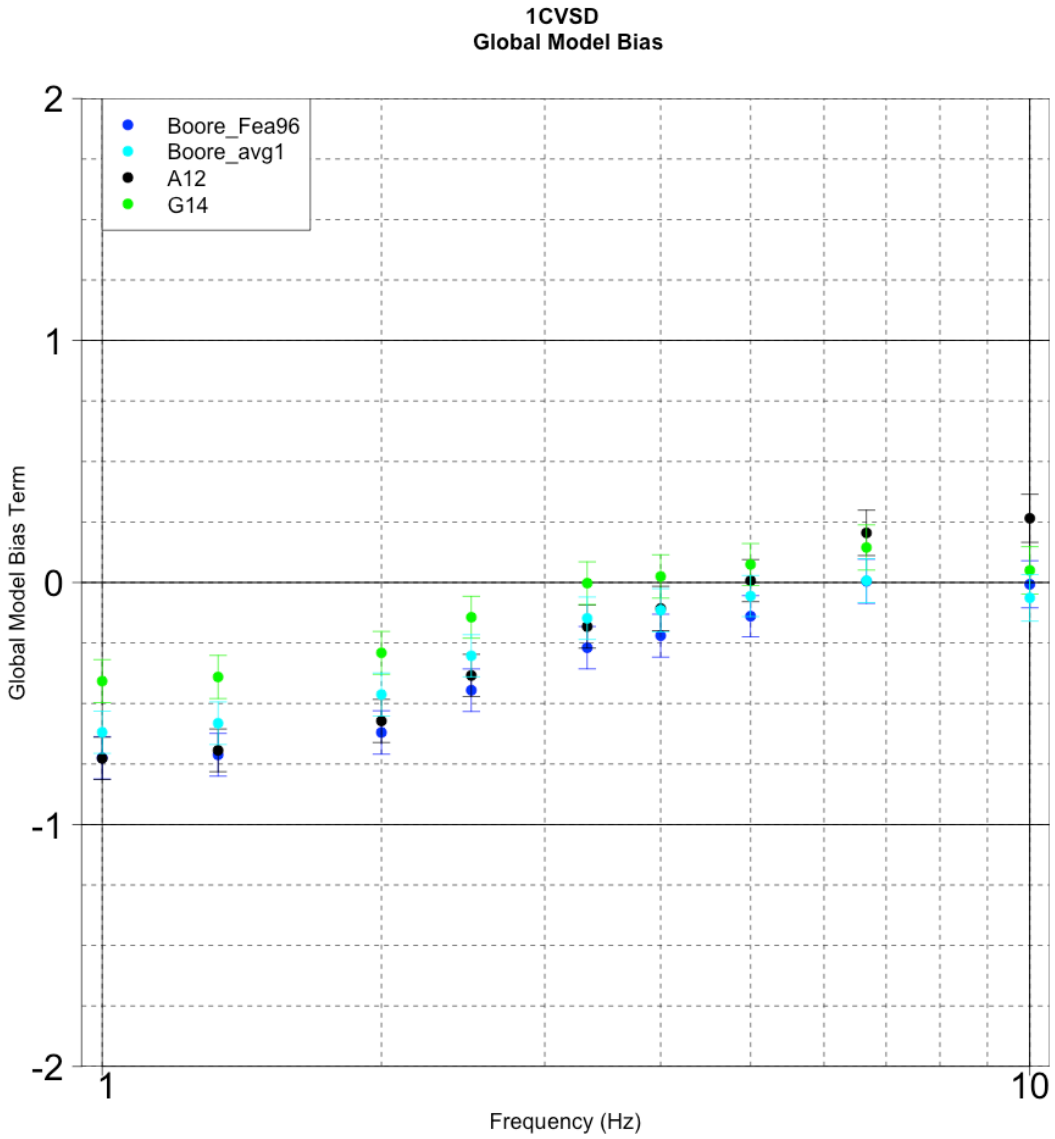


Figure 7–48 Model bias of fit to residuals for 2CCSP. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

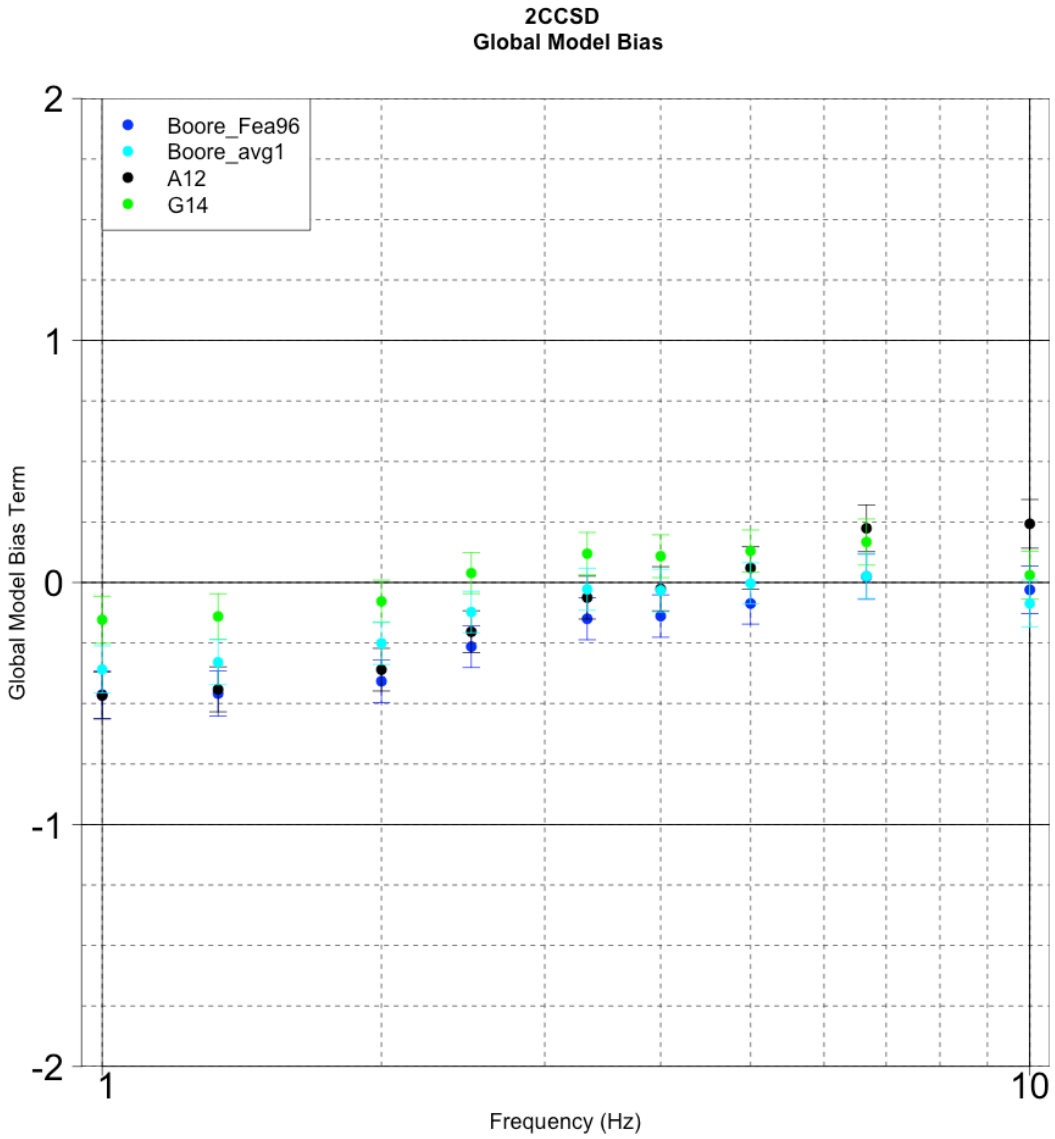


Figure 7–49 Model bias of fit to residuals for 2CVSP. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

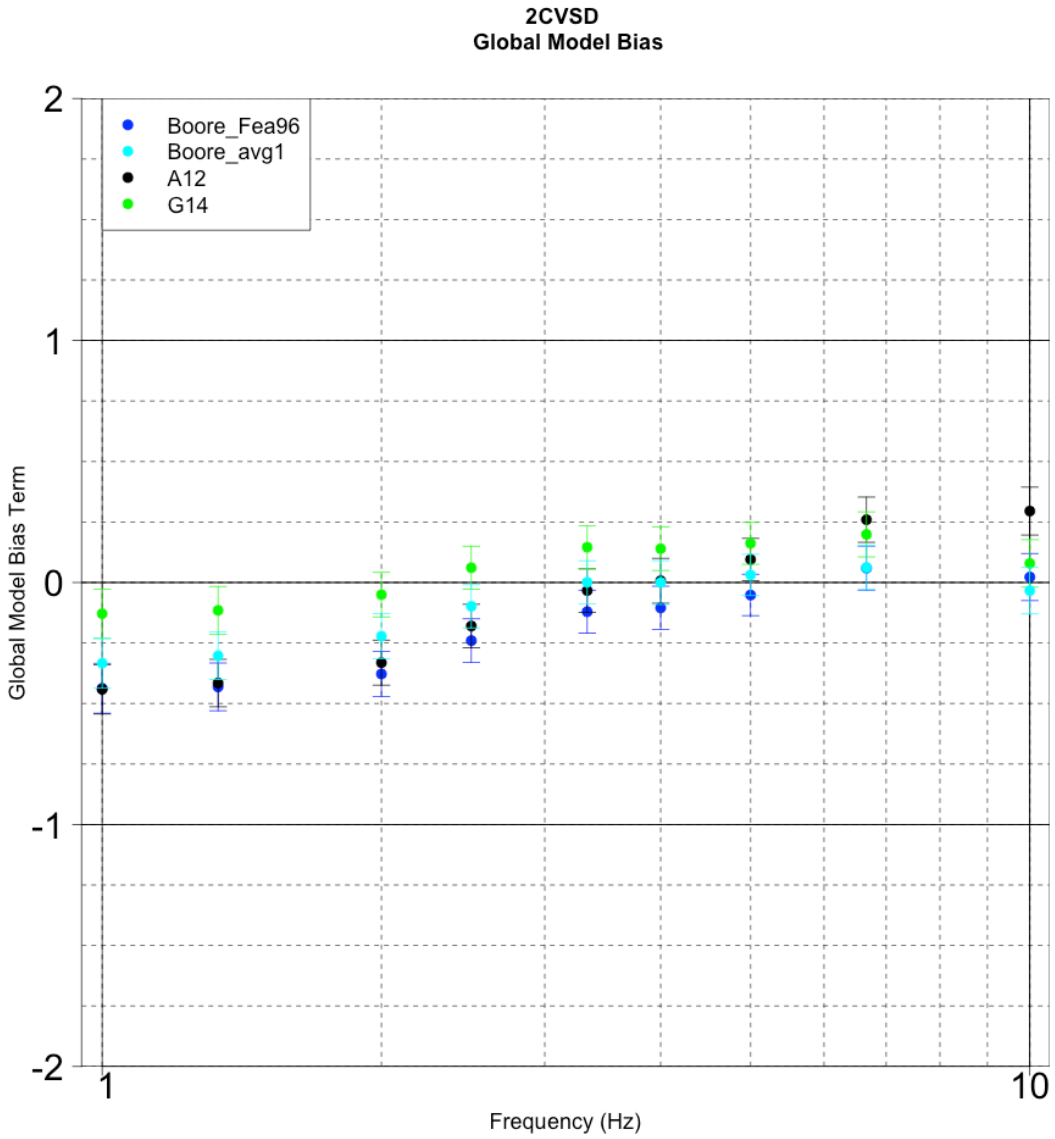


Figure 7–50 Model bias of fit to residuals for YA15. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

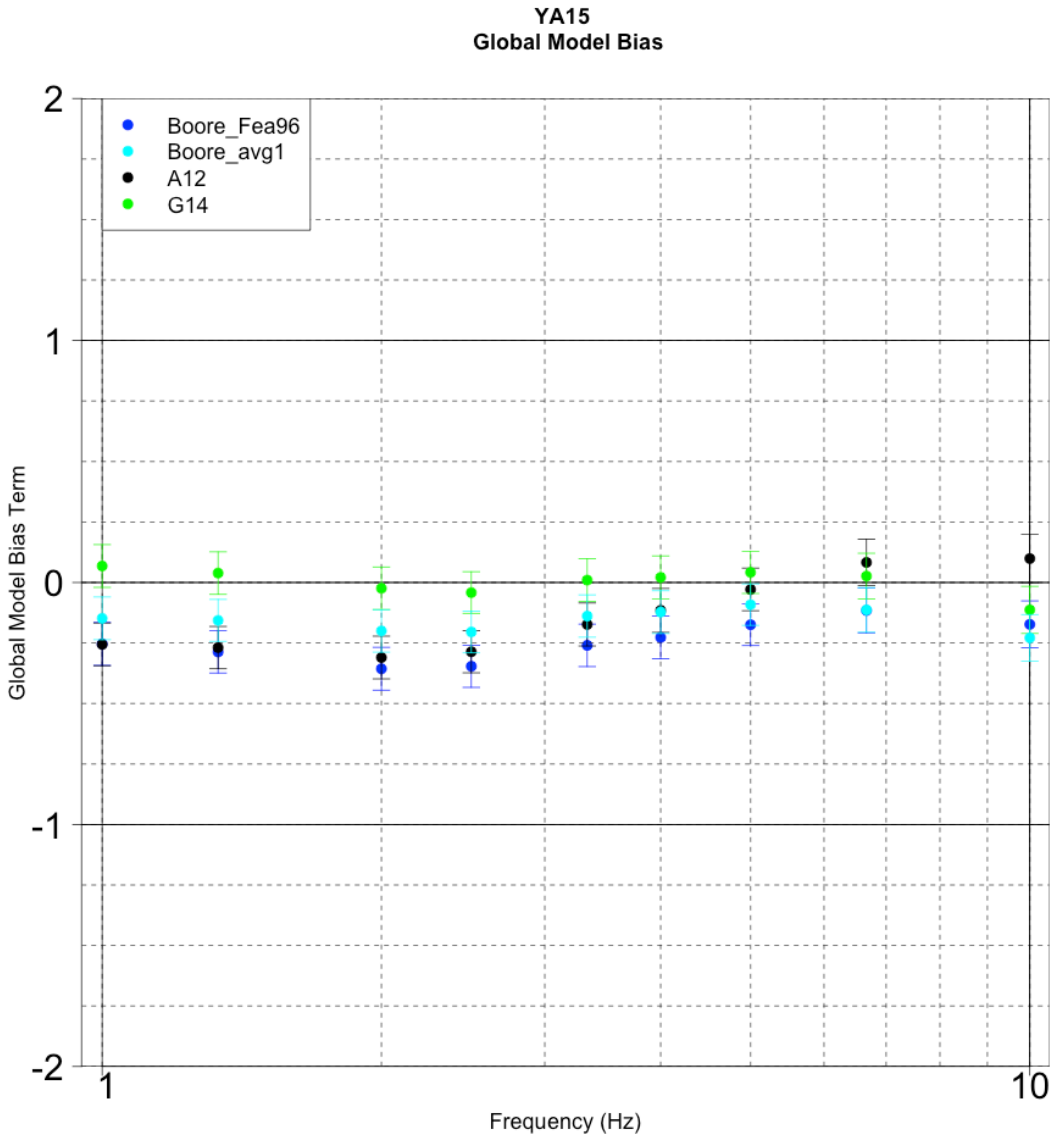


Figure 7–51 Model bias of fit to residuals for PZCT15_M1SS. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

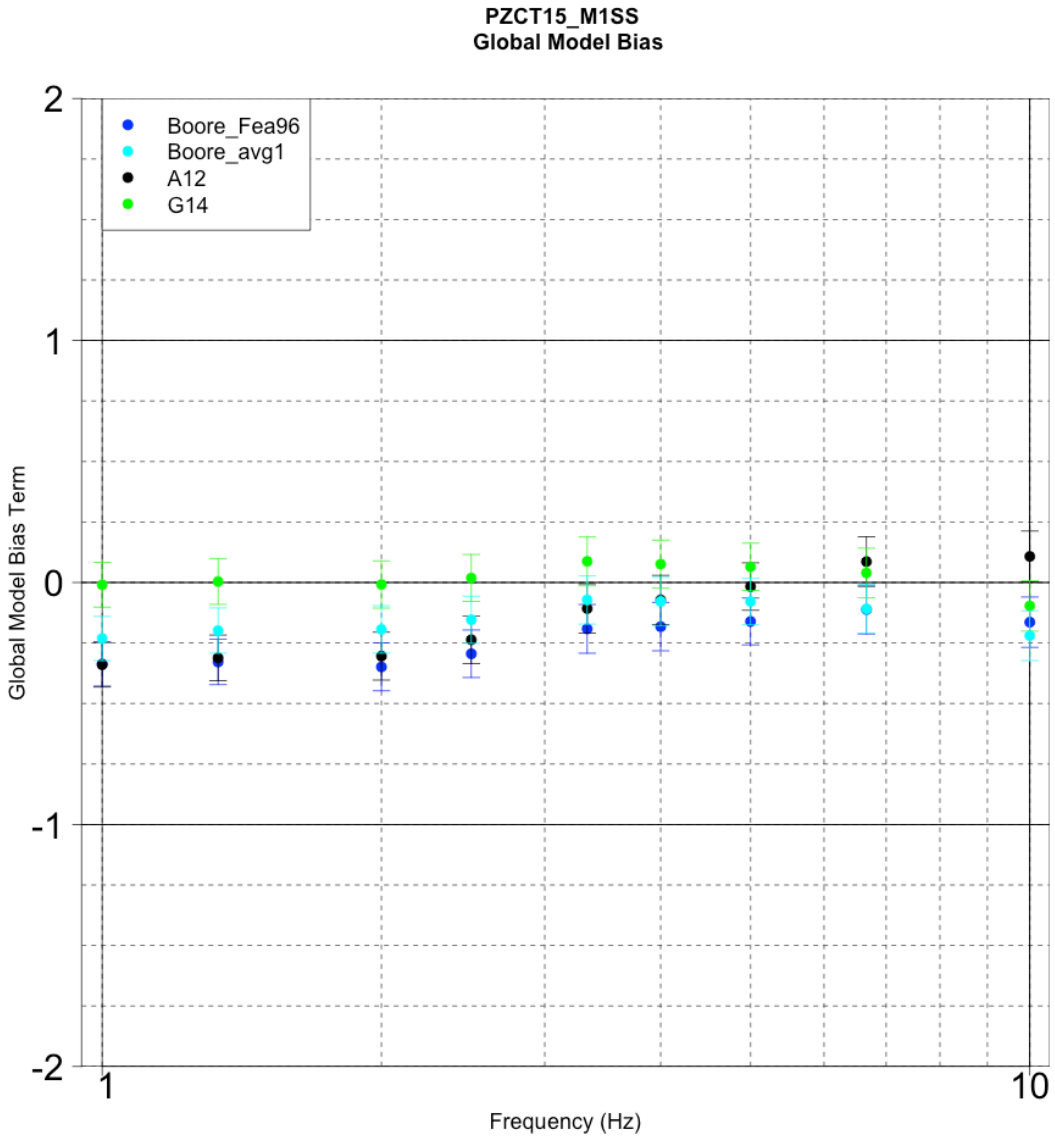


Figure 7–52 Model bias of fit to residuals for PZCT15_M2ES. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

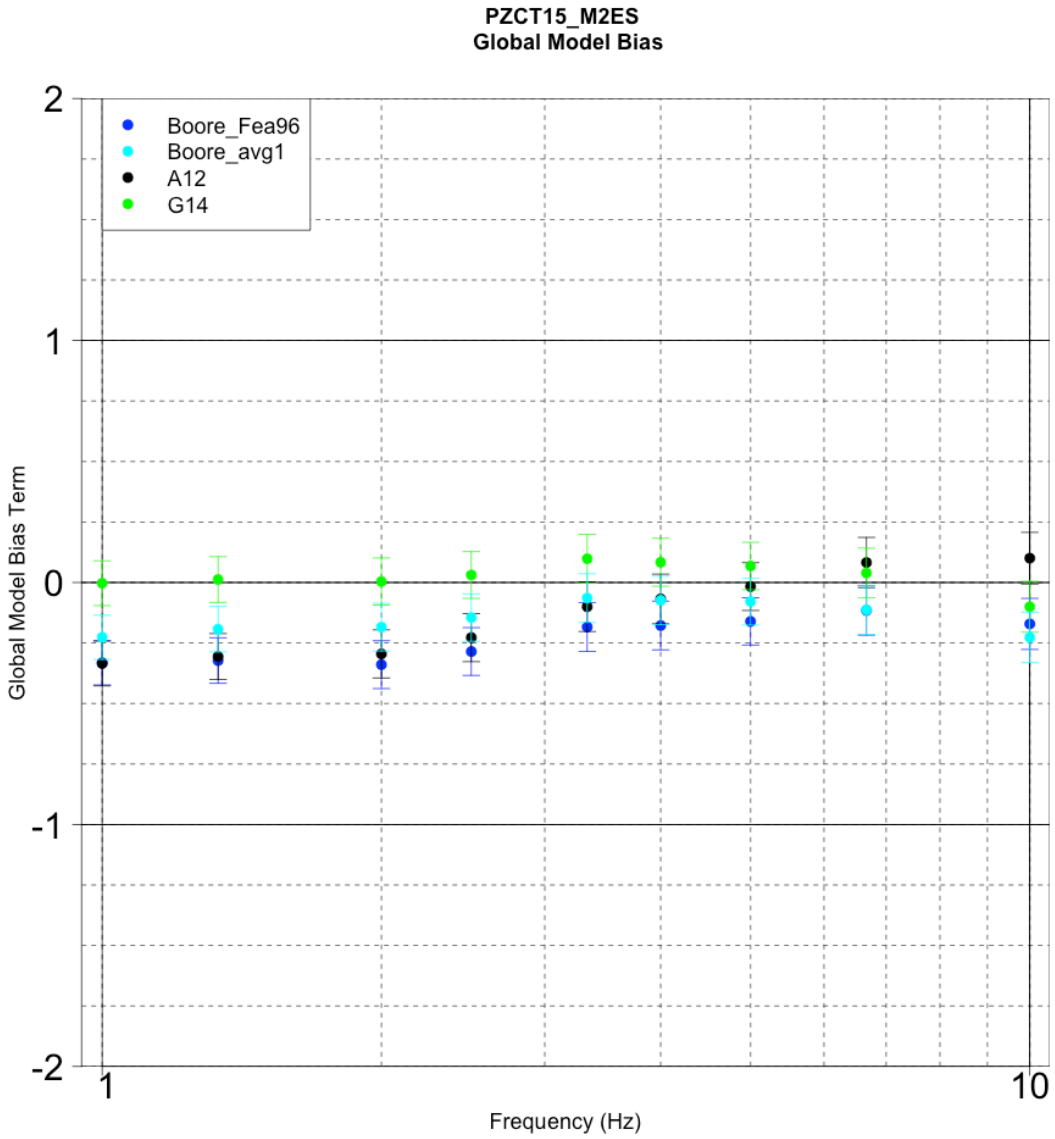


Figure 7–53 Model bias of fit to residuals for Frankel. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

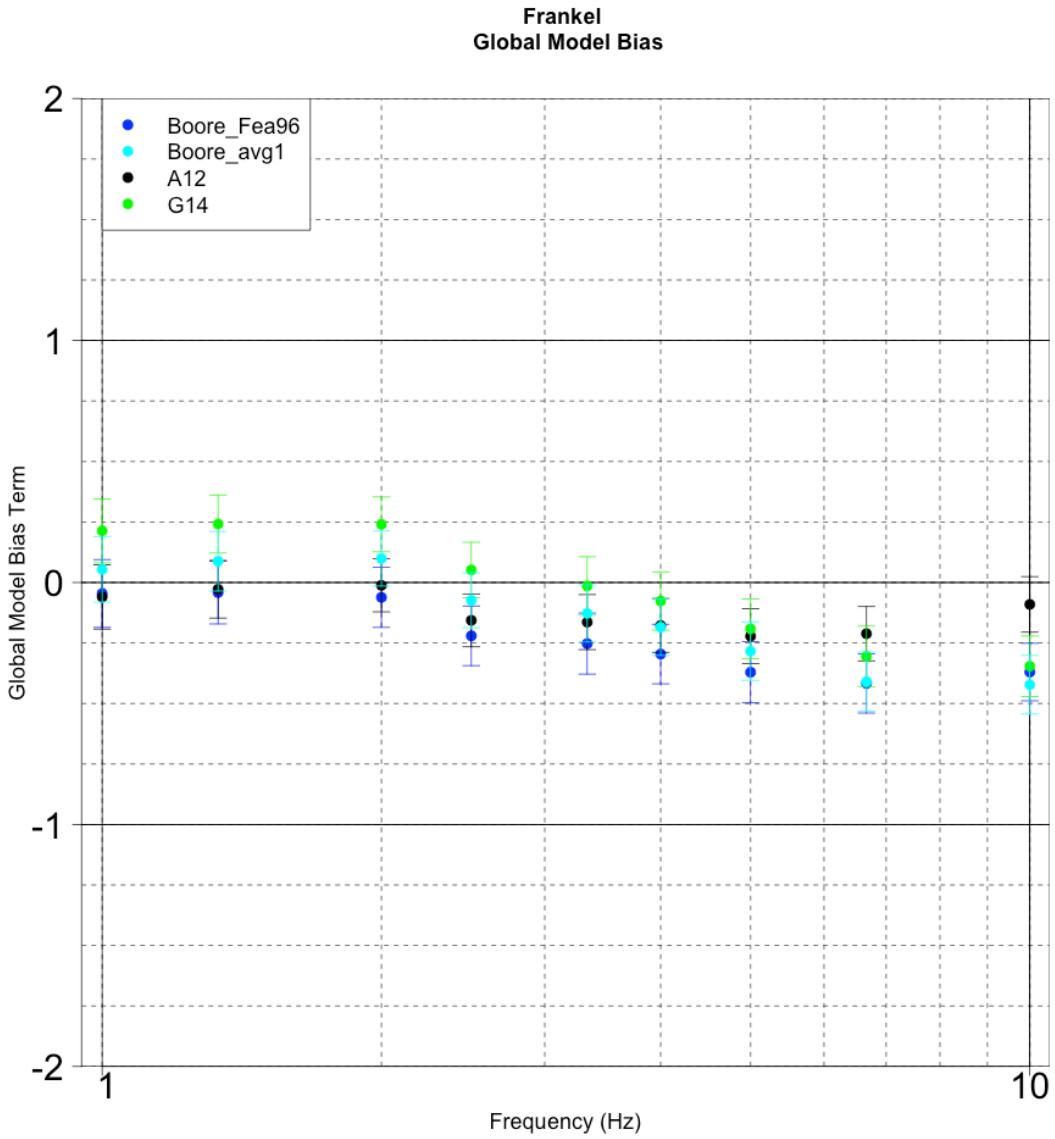


Figure 7–54 Model bias of fit to residuals for SP15. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

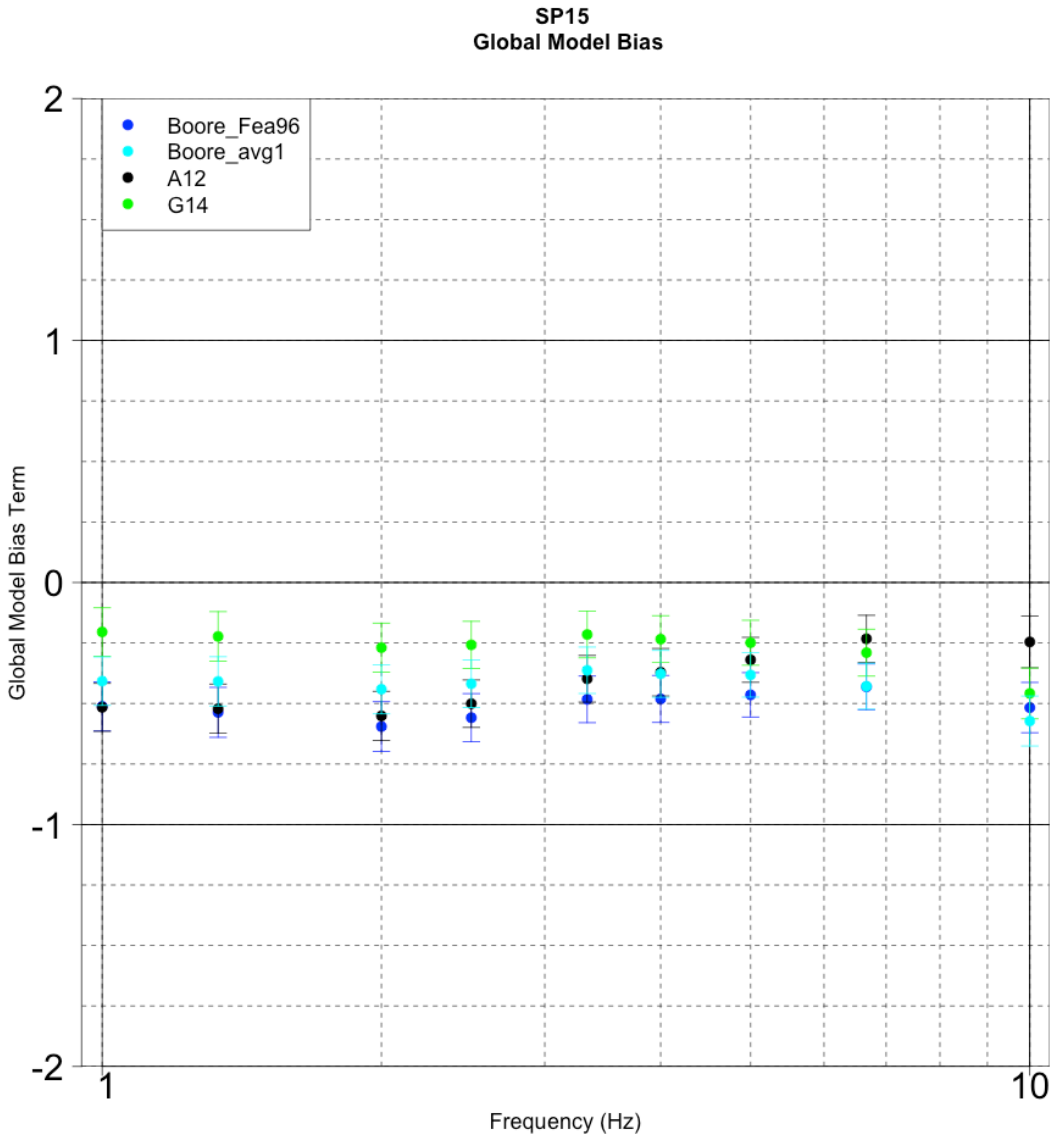


Figure 7–55 Model bias of fit to residuals for ANC15. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

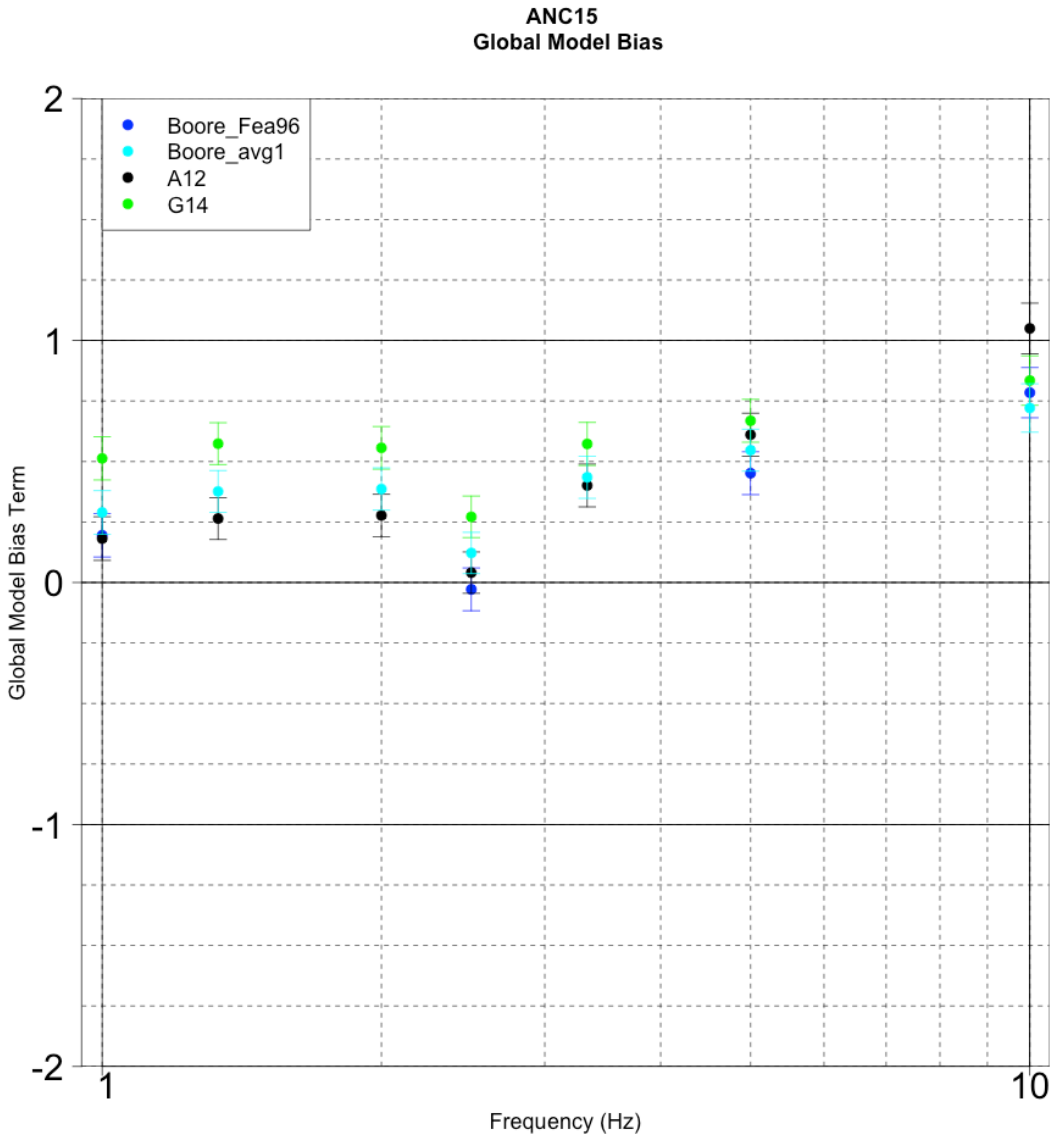


Figure 7–56 Model bias of fit to residuals for Graizer. The different color symbols indicate the four different site adjustment models used to correct the data to the Vs30 = 3000 m/s site condition.

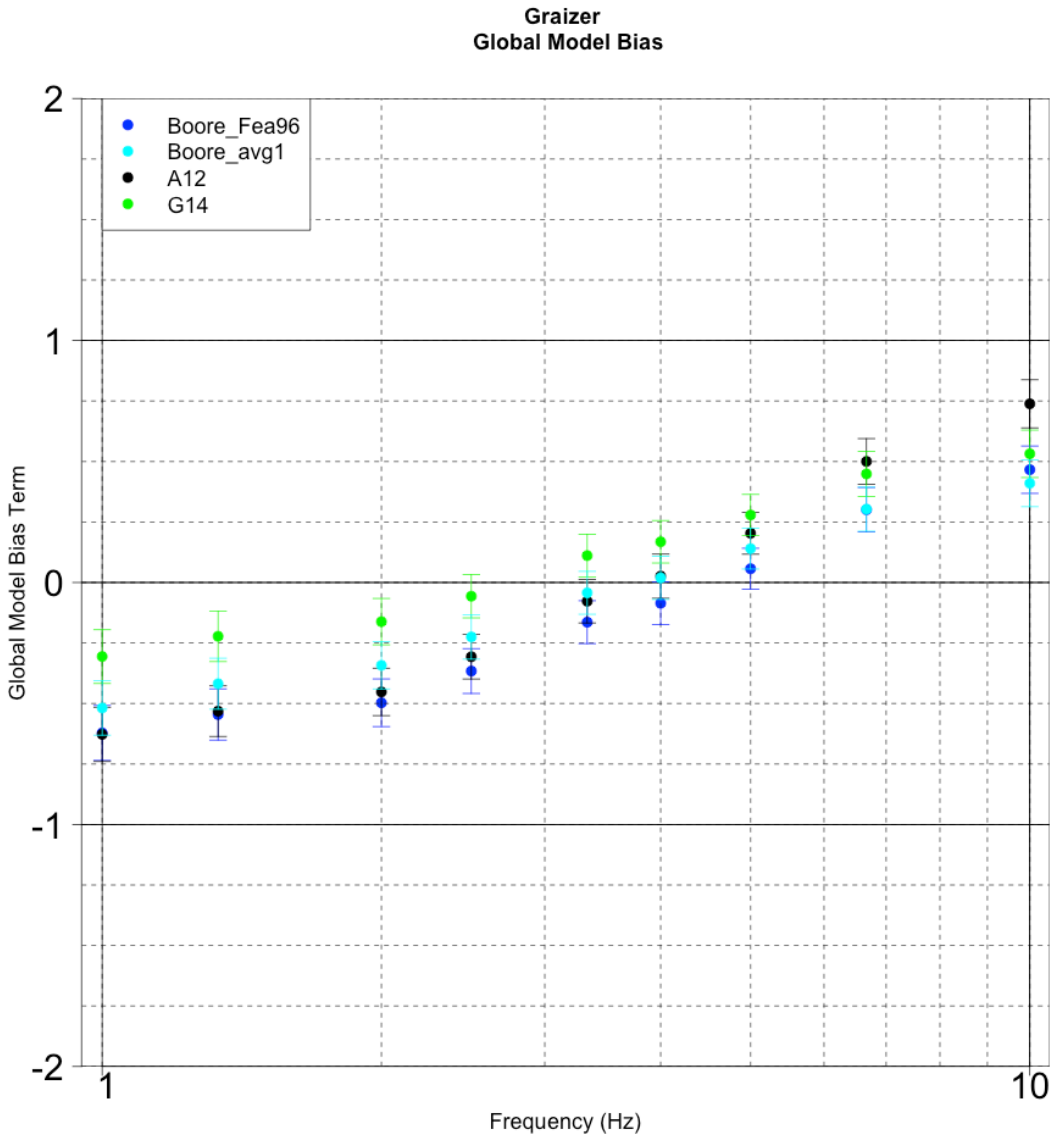


Figure 7–57 Model bias of fit to residuals for HA15. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

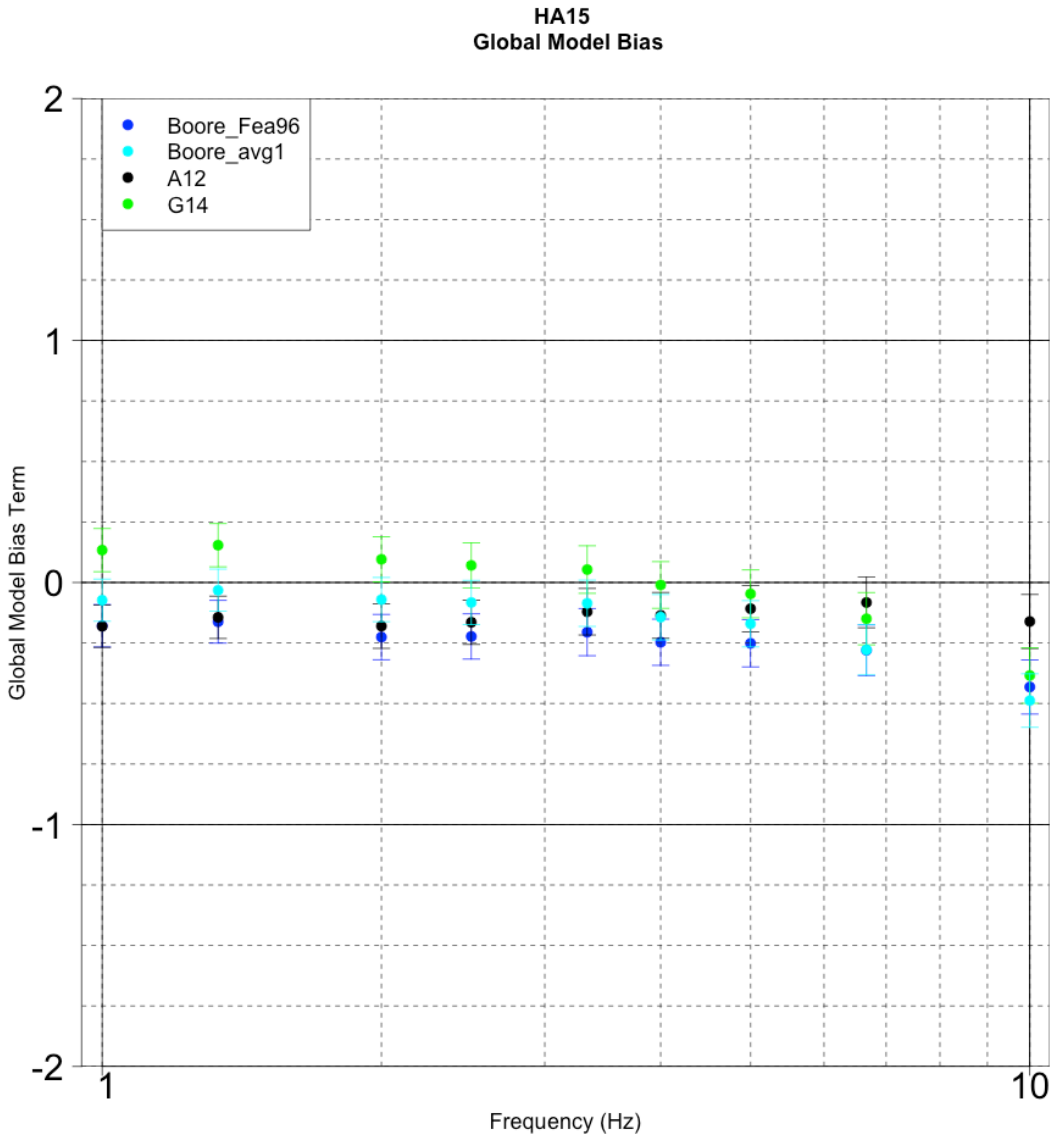


Figure 7–58 Model bias of fit to residuals for PEER_EX. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

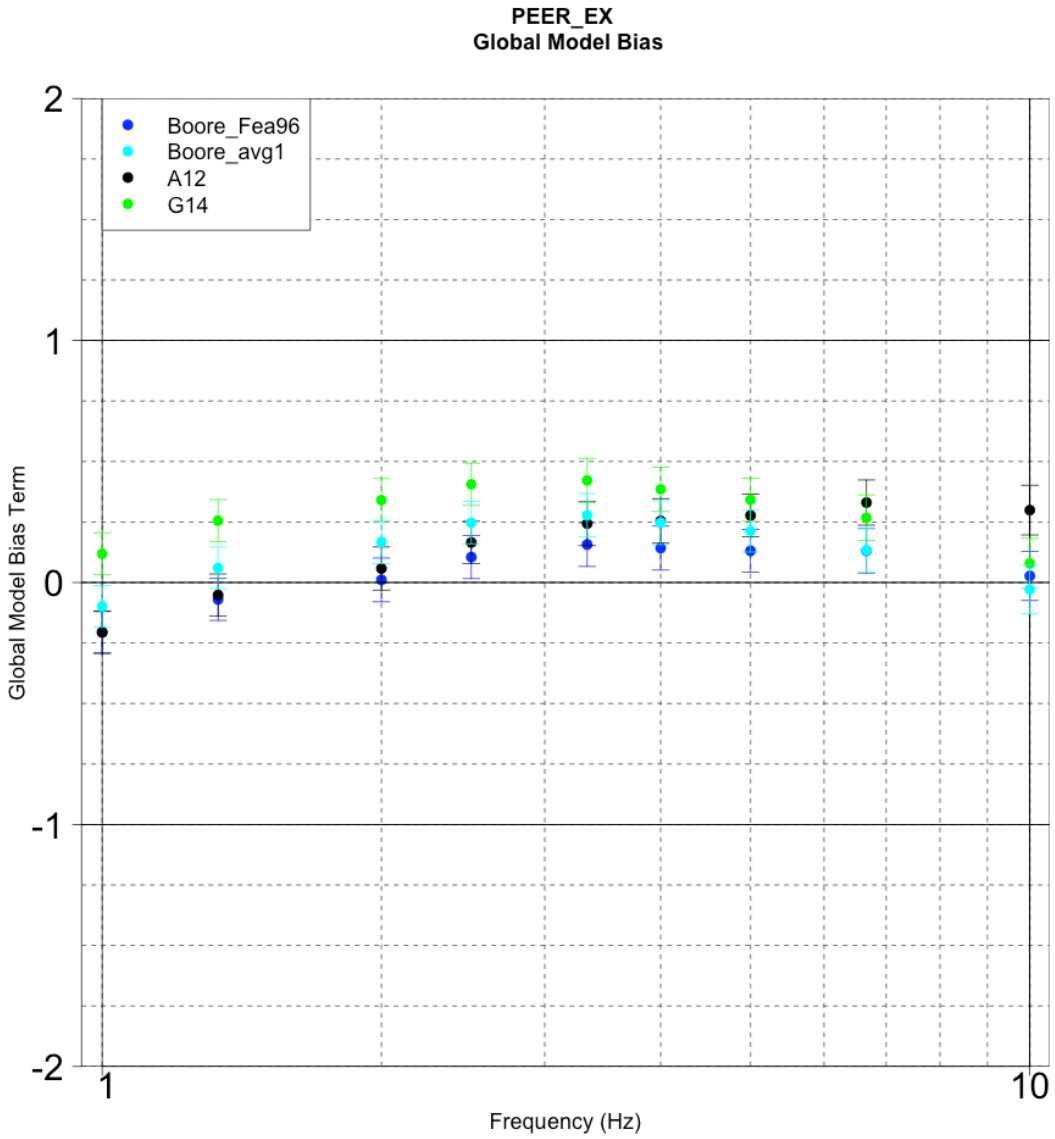


Figure 7–59 Model bias of fit to residuals for PEER_GP. The different color symbols indicate the four different site adjustment models used to correct the data to the V_{S30} = 3000 m/sec site condition.

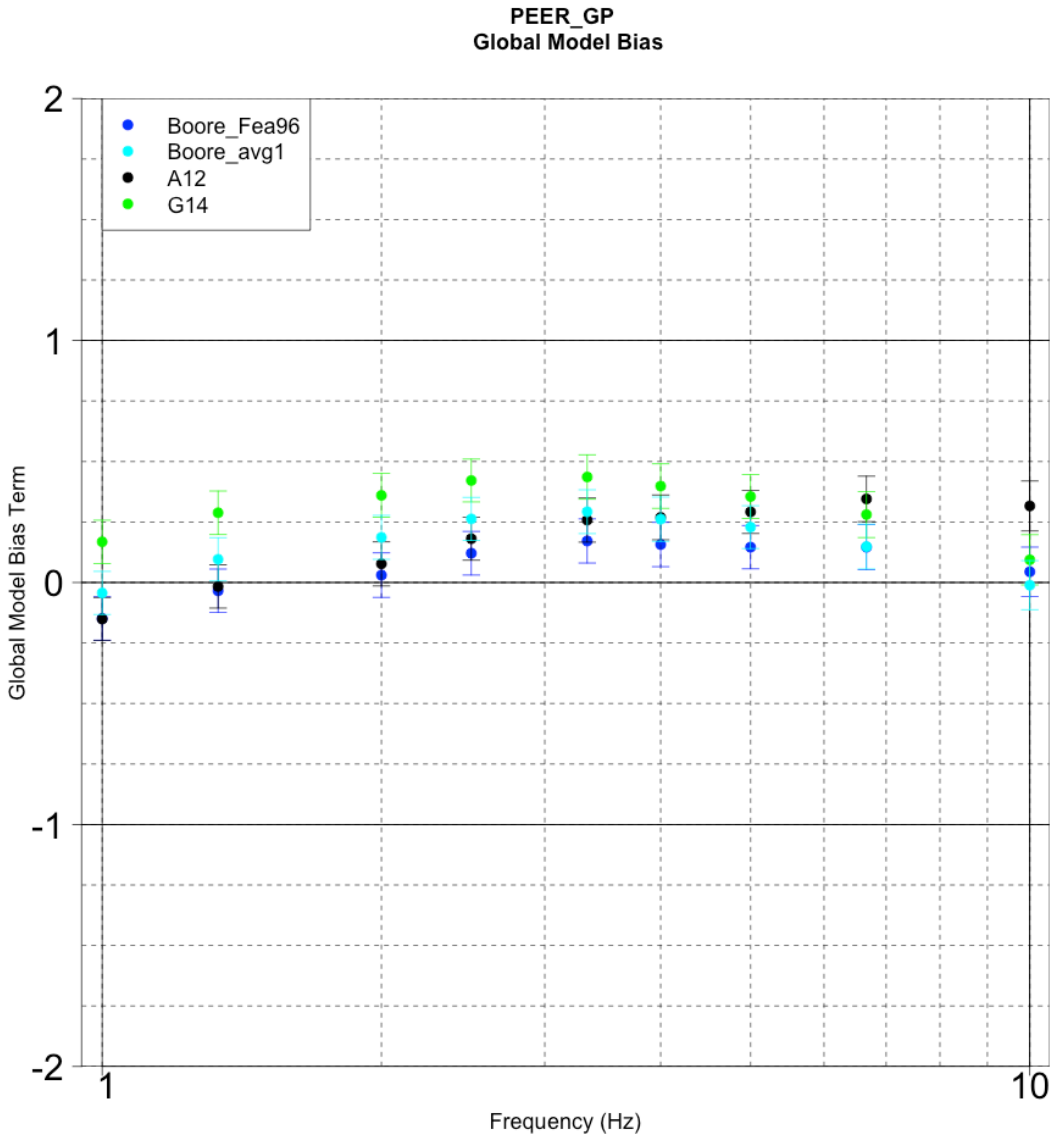


Figure 7–60 **Z_{TOR} values calculated for strike–slip and reverse faulting events for two hypocentral depths.**

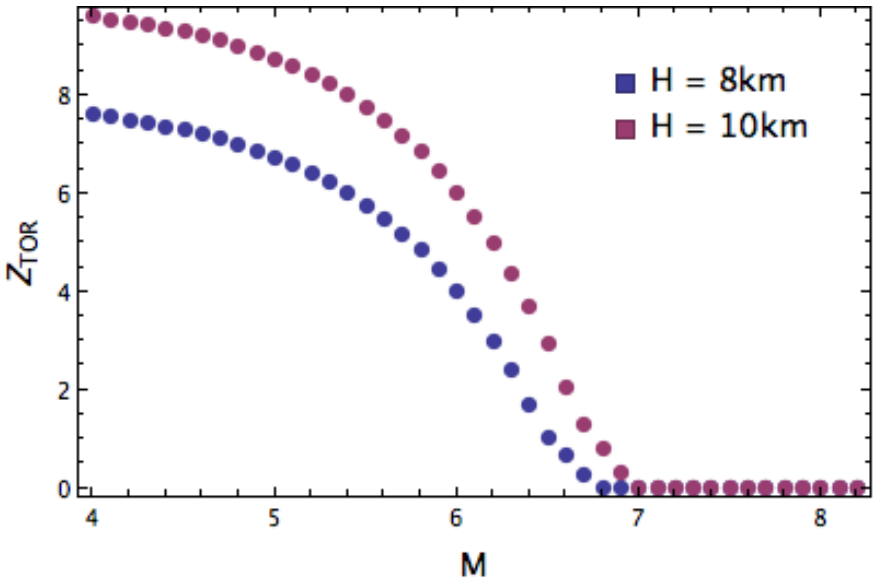


Figure 7–61 **Illustration of R_{RUP} conversion for the DASG 1CVSP model at 1 Hz.**

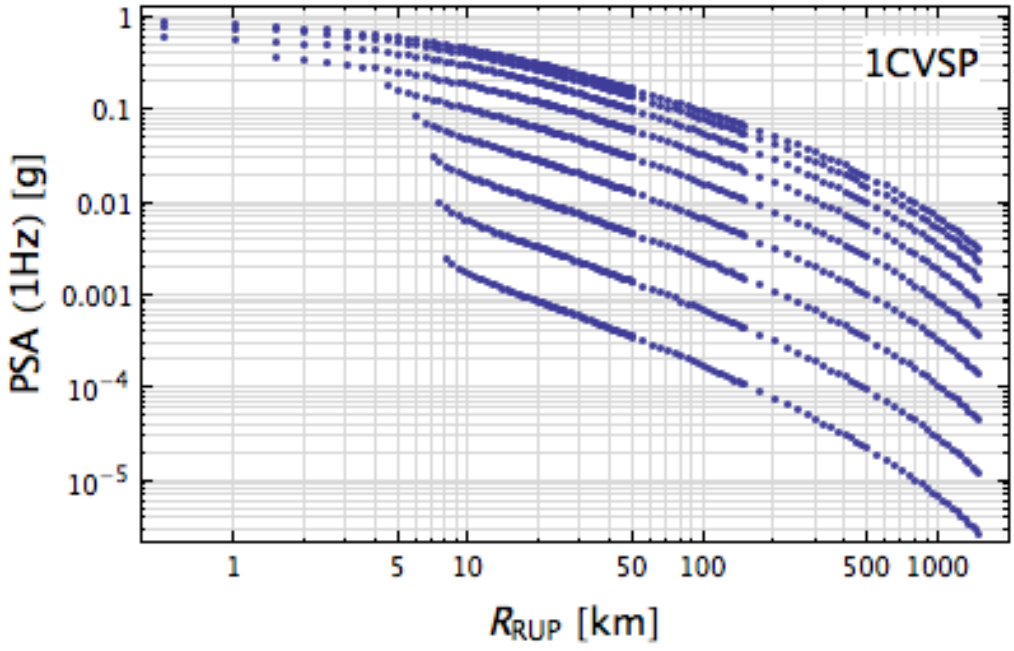


Figure 7–62 Illustration of R_{RUP} conversion for the HA15 model at 1 Hz.

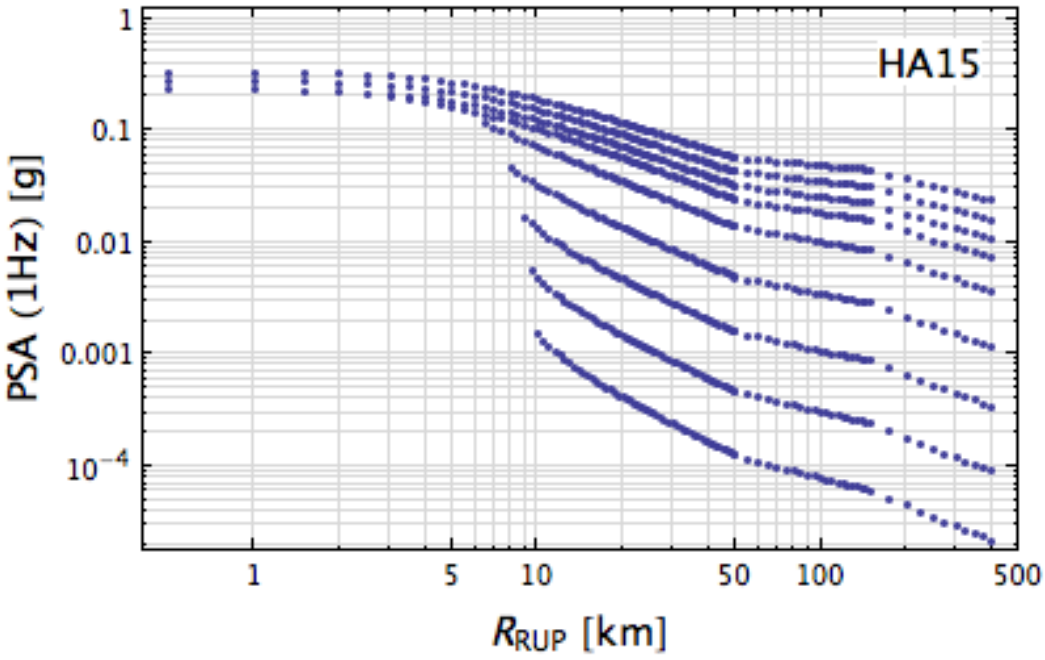


Figure 7–63 Illustration of R_{RUP} conversion for the SP15 model at 1 Hz.

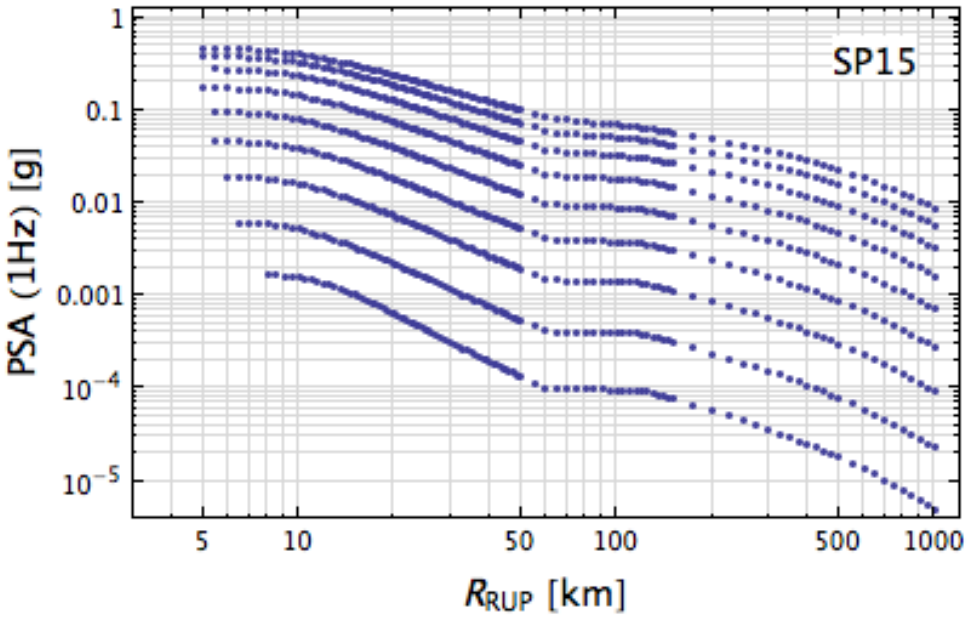


Figure 7–64 Distance extensions for the B_a04, 1 Hz PSA. The left frame shows the extended regions in blue and red; the right frame shows the final model.

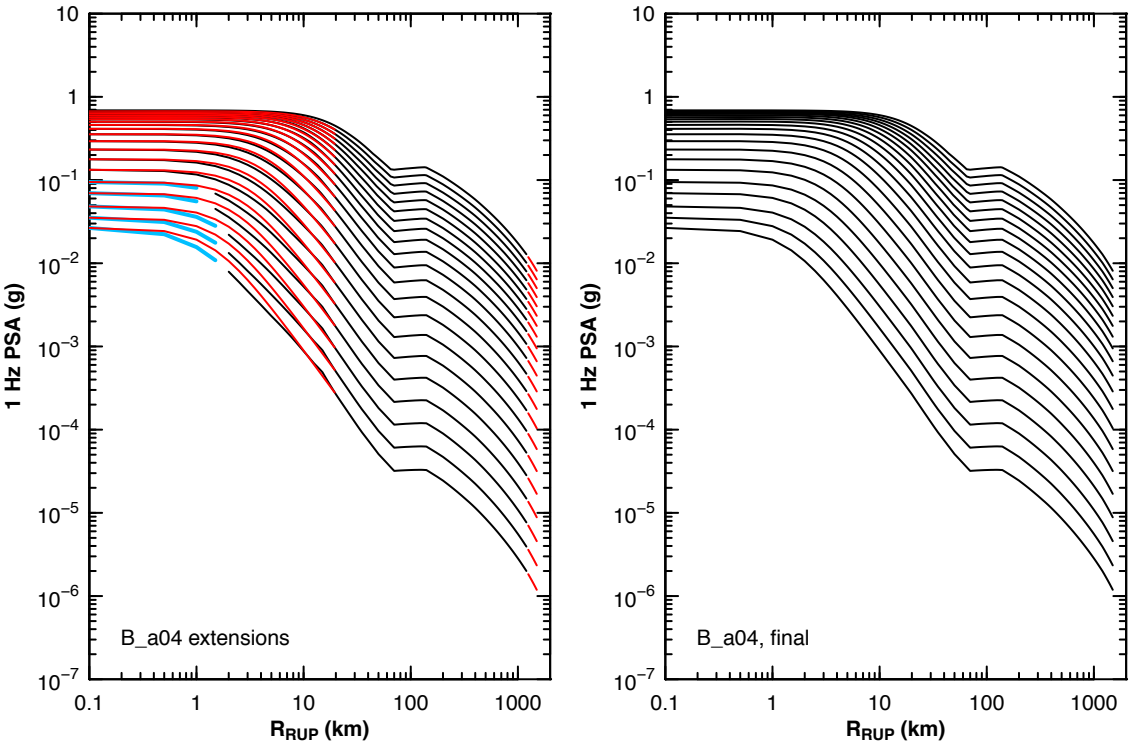


Figure 7–65 Distance extensions for the B_a04, 10 Hz PSA. The left frame shows the original predictions in black and the extended regions in blue and red; the right frame shows the final model. The right panel shows the results of the algorithm used to prevent the crossover.

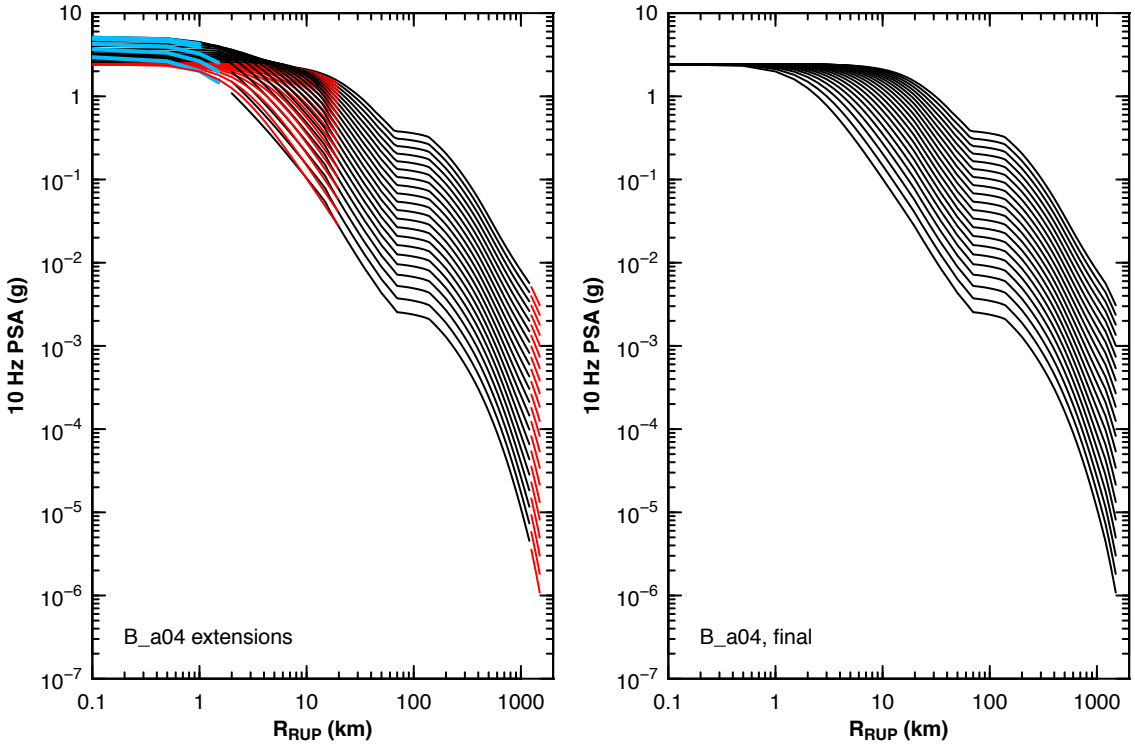


Figure 7–66 Distance extensions for the DASG 1CCSP GMM, 1 Hz PSA. The left frame shows the original predictions in black and the fitted model in red; the right frame shows the final model.

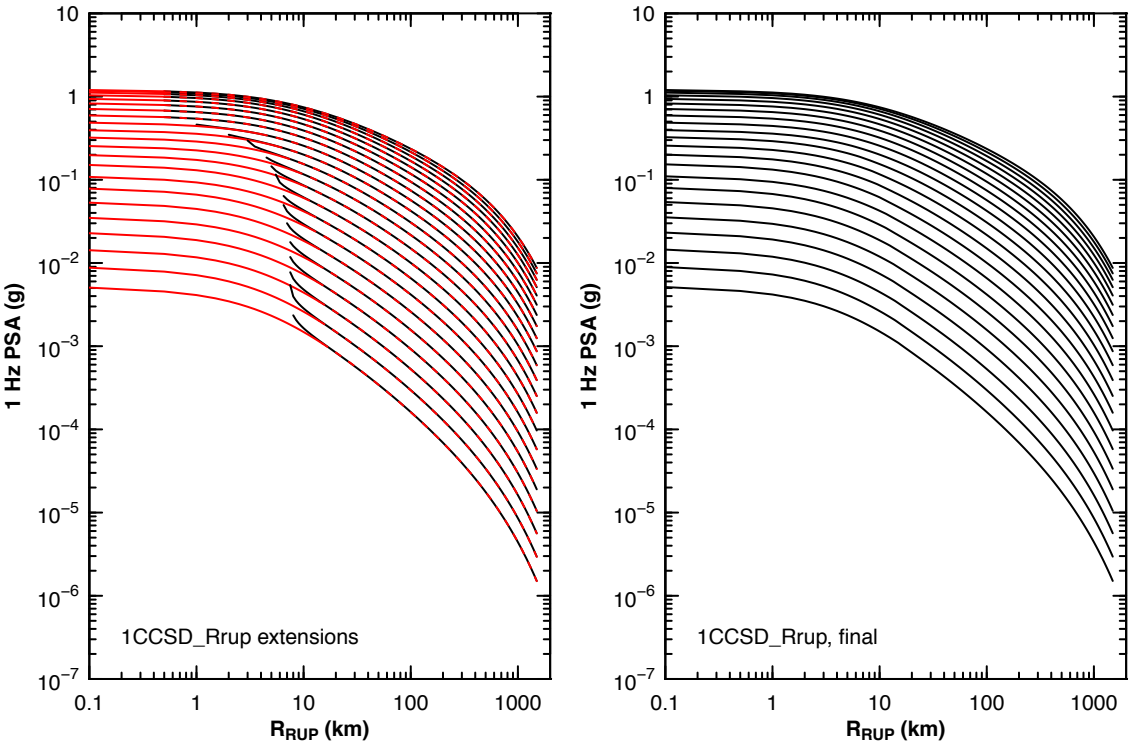


Figure 7–67 Large distance extensions for the YA15 GMM, 1 Hz PSA. The left frame shows the original predictions in black, and the green line shows the lower-bound distance limit used to fit the model (in red); the right frame shows the final model.

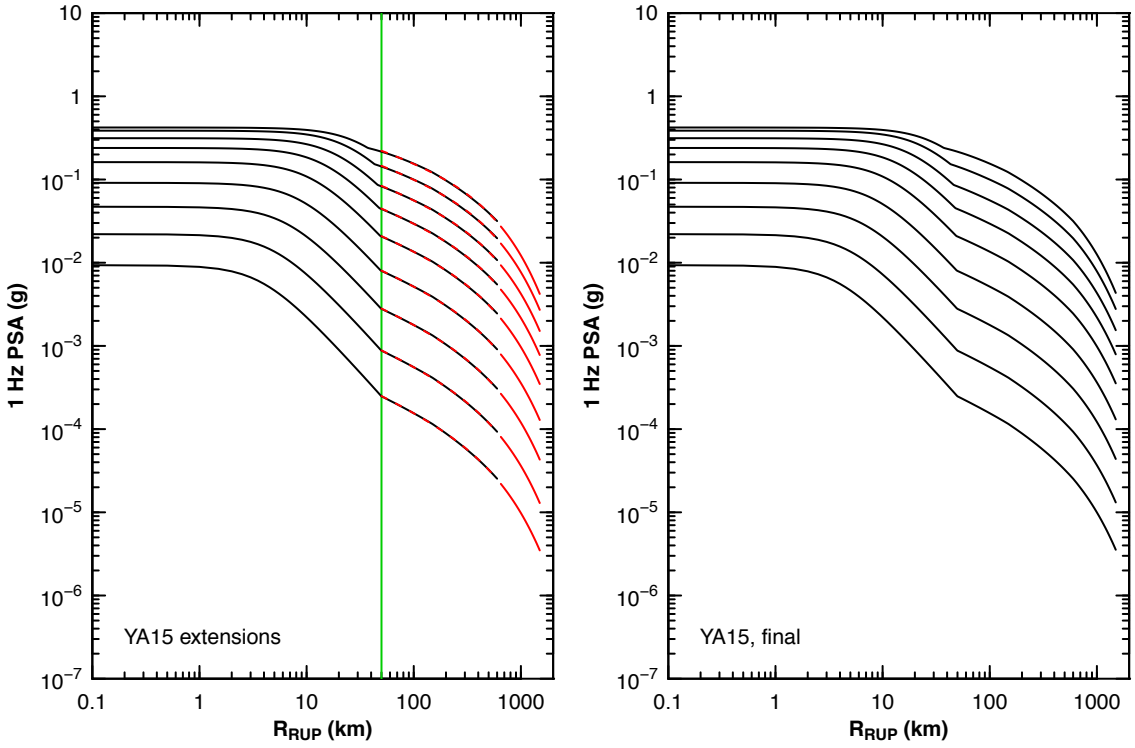


Figure 7–68 Illustration of short-distance correction for the 25 Hz PSA example of YA15. The left frame shows the original predictions in black, the dashed green line is anchored at 15 km, the solid green line shows the lower-bound distance limit used to fit the model at large distances, and the red line shows the model; the right frame shows the final model, including the short distance corrections and the large distance extrapolations.

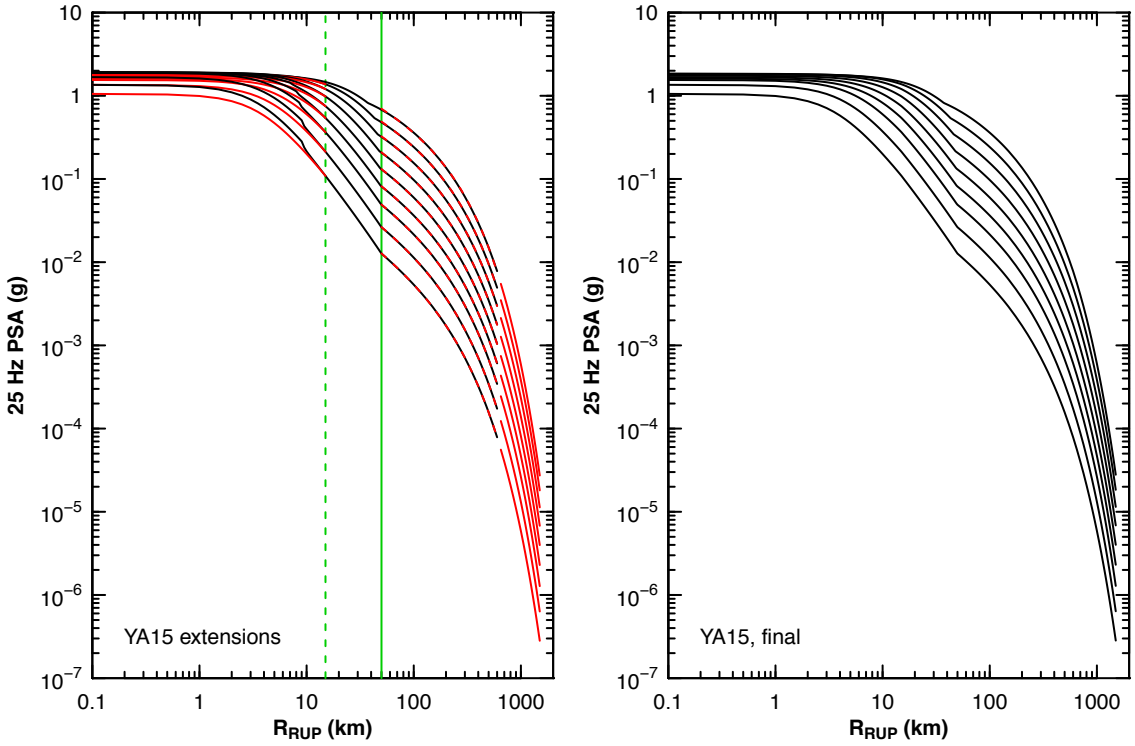


Figure 7–69 Large distance extrapolations for the PZCT_M1SS GMM, 1 Hz PSA. The left frame shows the original predictions in black, and the green line shows the lower-bound distance limit used to fit the model (dashed red lines show the model fit and solid red lines show the extrapolations); the right frame shows the final model.

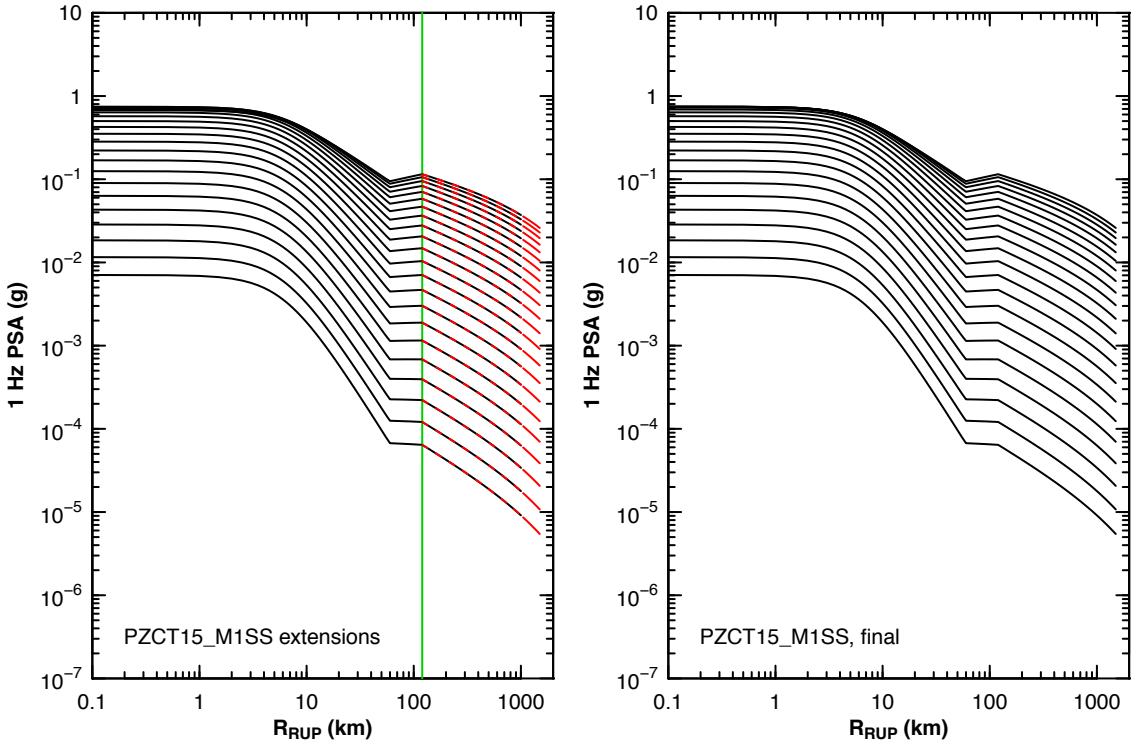


Figure 7–70 Extrapolation of Frankel GMM, 0.1 Hz PSA, Case 1 example. The left frame shows open black circles for Z_{TOR} = 5 km simulations, solid black circles for Z_{TOR} = 1 km simulations, green circles for ground motions extrapolated to R_{RUP} = 0 km for M = 7.5 and M = 8 and open red circles for the extrapolations (short and large distances); the right frame shows the final model.

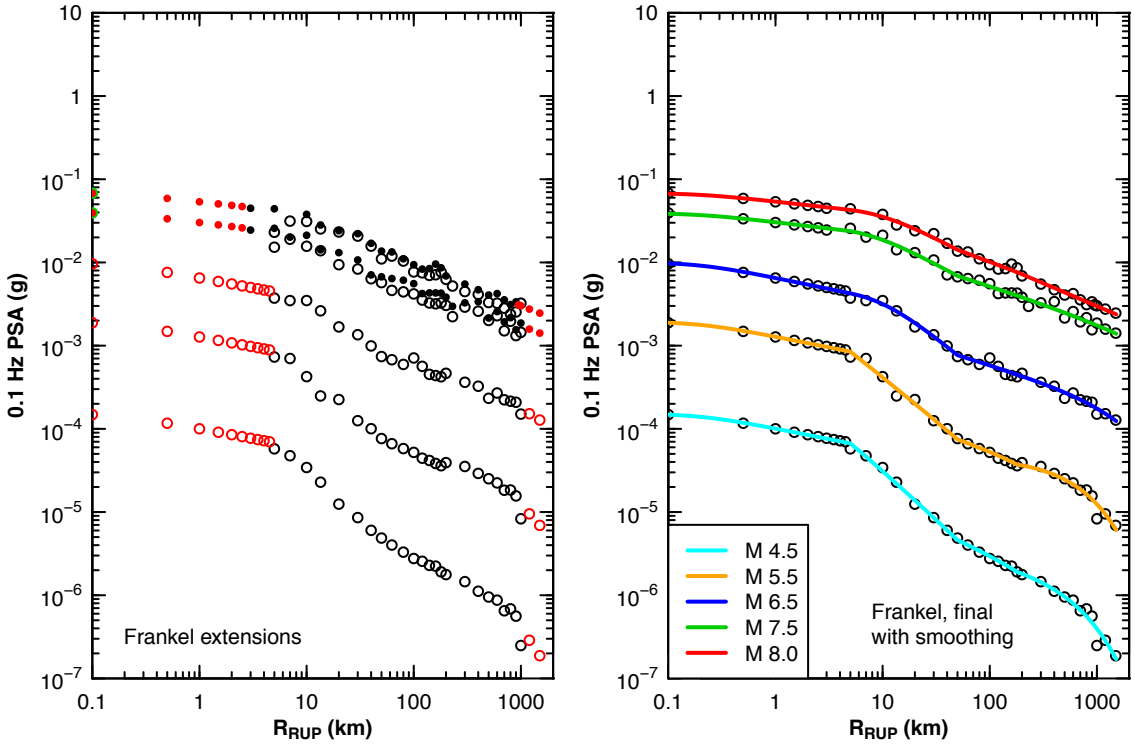


Figure 7–71 Extrapolation of Frankel GMM, 0.2 Hz PSA, Case 2 example. The left frame shows open black circles for Z_{TOR} = 5 km simulations, solid black circles for Z_{TOR} = 1 km simulations, green circles for ground motions extrapolated to R_{RUP} = 0 km for M = 7.5 and M = 8 and open red circles for the extrapolations (short and large distances); the right frame shows the final model.

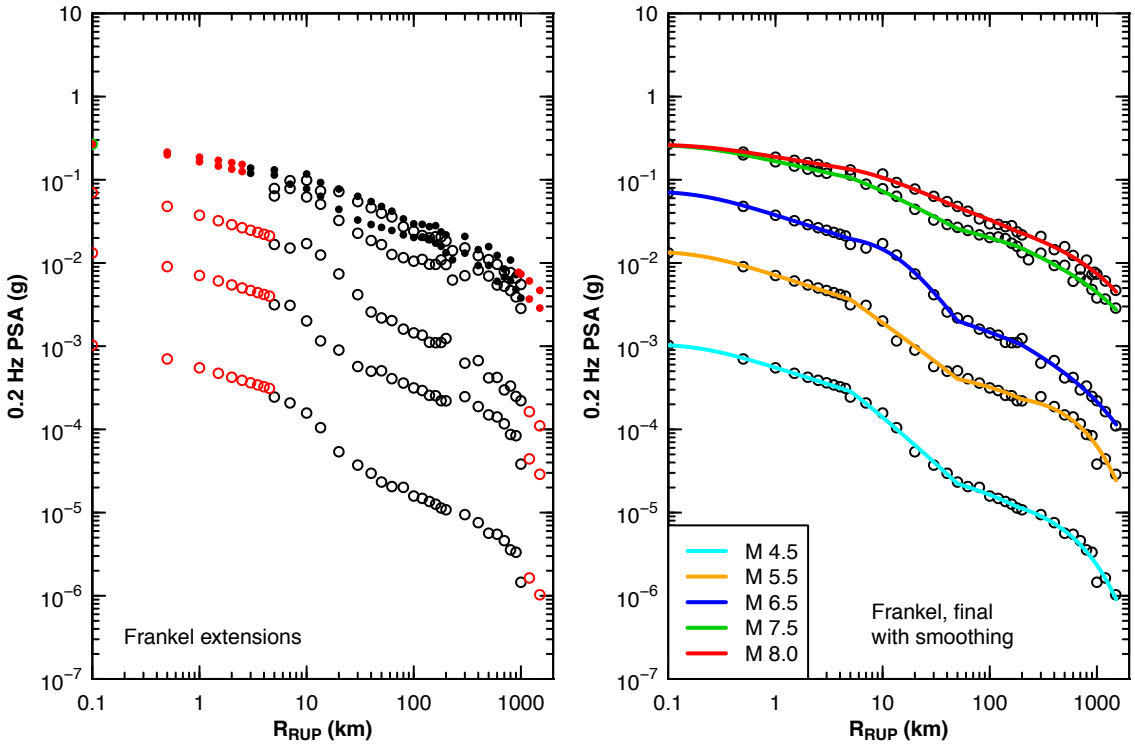


Figure 7–72 Extrapolation of SP15 GMM, 1 Hz PSA. The left frame shows the original predictions in black, and the green line shows the lower-bound limit of ground motions used to fit the model at large distances (dashed red lines show the model fit and solid red lines show the extrapolations); the right frame shows the final model.

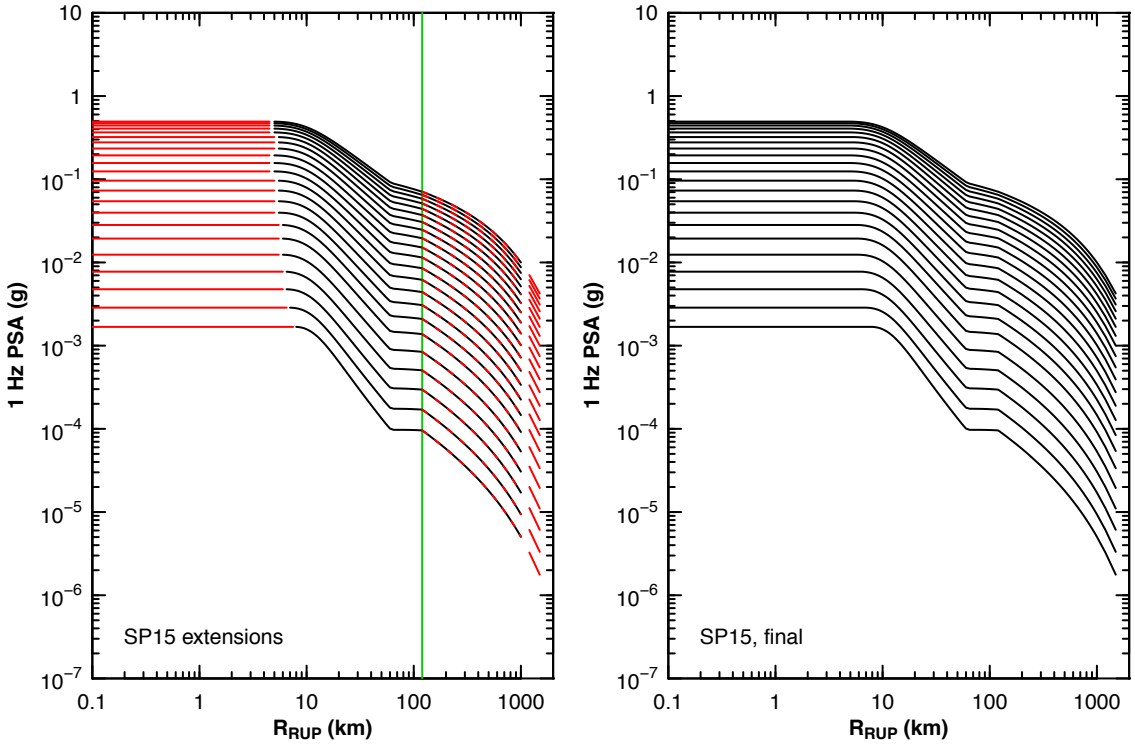


Figure 7–73 Graizer GMM ground motions for 1 Hz PSA. The model has not been extrapolated and is used as-is over the full distance range.

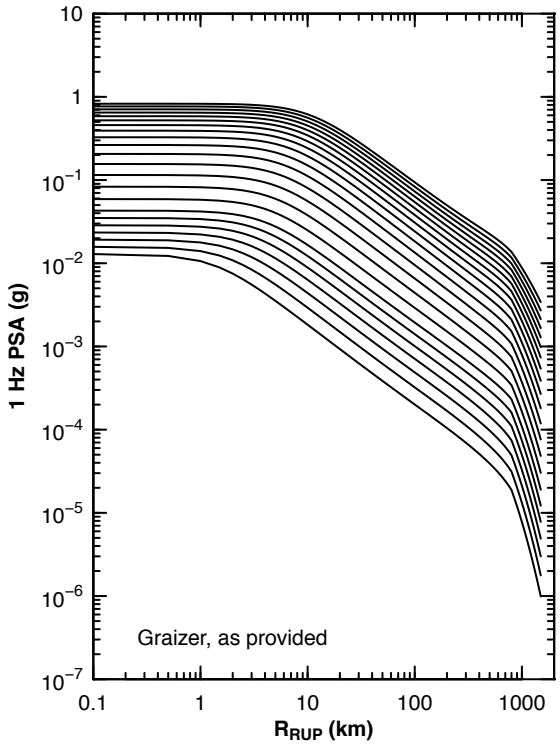


Figure 7–74 Illustration of extrapolation for the HA15 GMM, PSA 1 Hz. The left frame shows the original predictions in black, the dashed green line is anchored at 20 km, the solid green line shows the lower-bound distance limit used to fit the model at large distances (anchored at 120 km), and the red lines show the model; the right frame shows the final model.

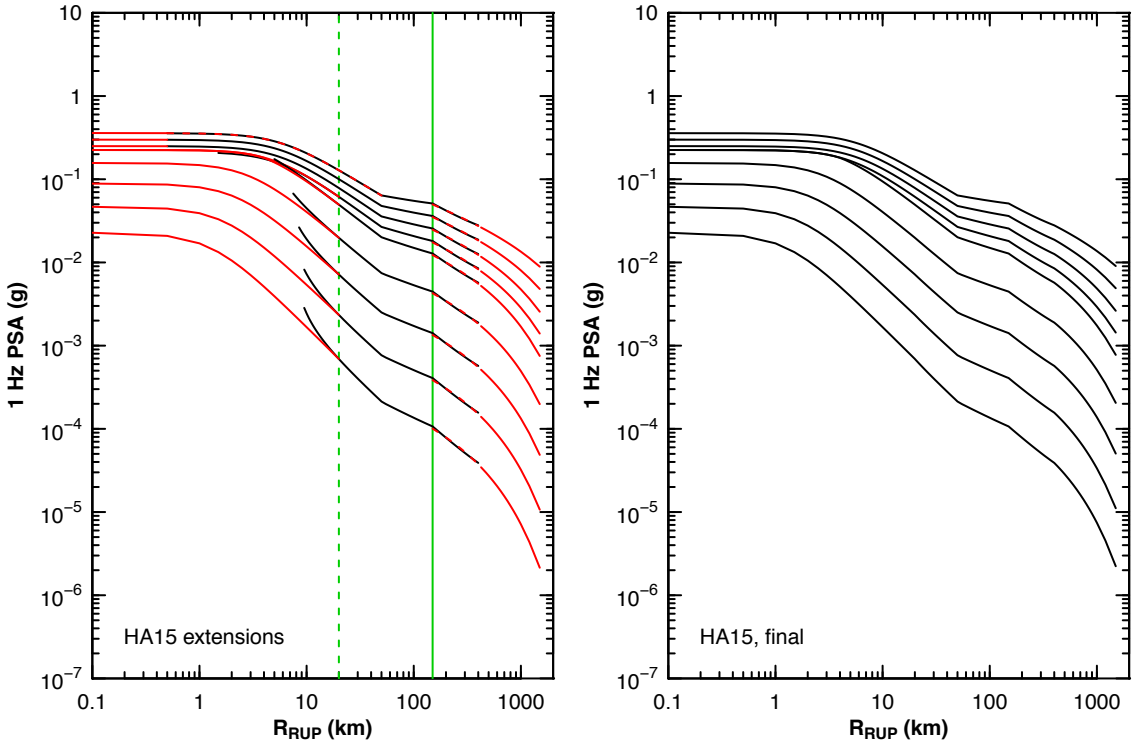


Figure 7–75 Illustration of extrapolation for the PEER_GP GMM, PSA 1 Hz. The left frame shows the original predictions in black and the extrapolations in red; the right frame shows the final model.

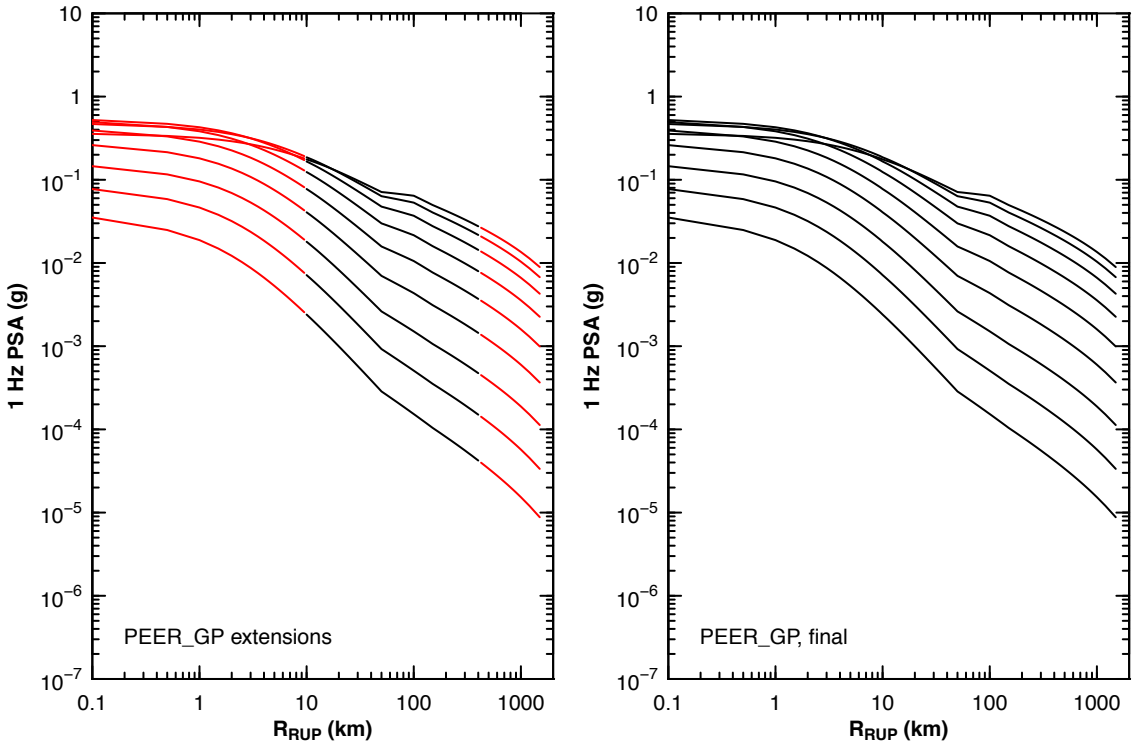


Figure 7–76 Illustration of extrapolation for the PEER_GP GMM, PSA 10 Hz. Format the same as Figure 7-74.

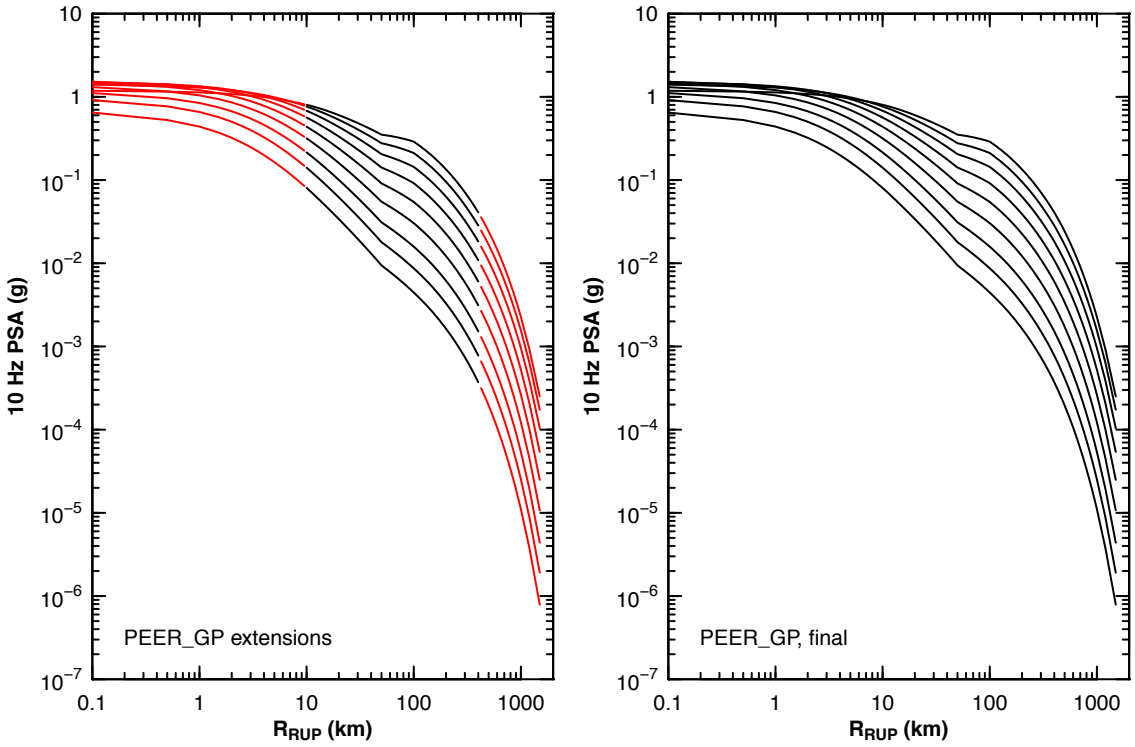


Figure 7–77 Example of spectral-shape correction showing the 1CCSP spectrum for M = 8 and R_{RUP} = 1200 km. The blue line is the original spectrum showing a trough-and-peak pair at high frequencies. The purple line shows the adjusted spectrum.

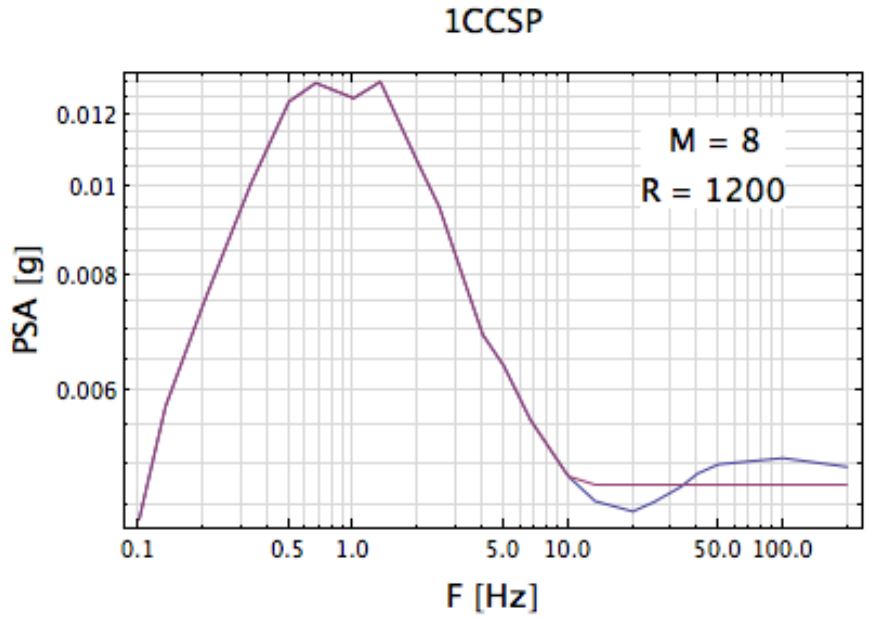


Figure 7–78 Example of spectral-shape correction showing the PEER_GP spectrum for M = 8 and R_{RUP} = 1200 km. The blue line is the original spectrum showing a trough at high frequencies. The purple line shows the adjusted spectrum.

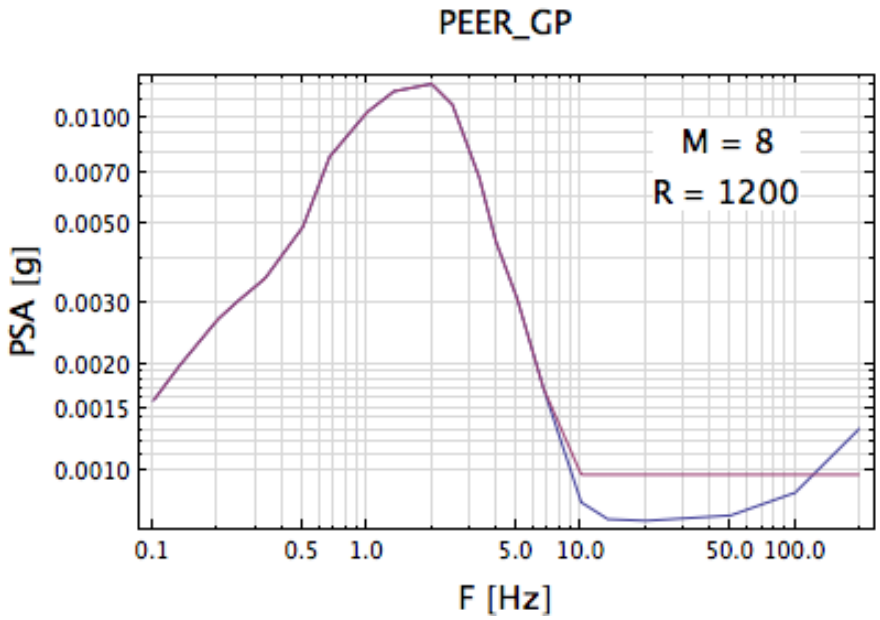
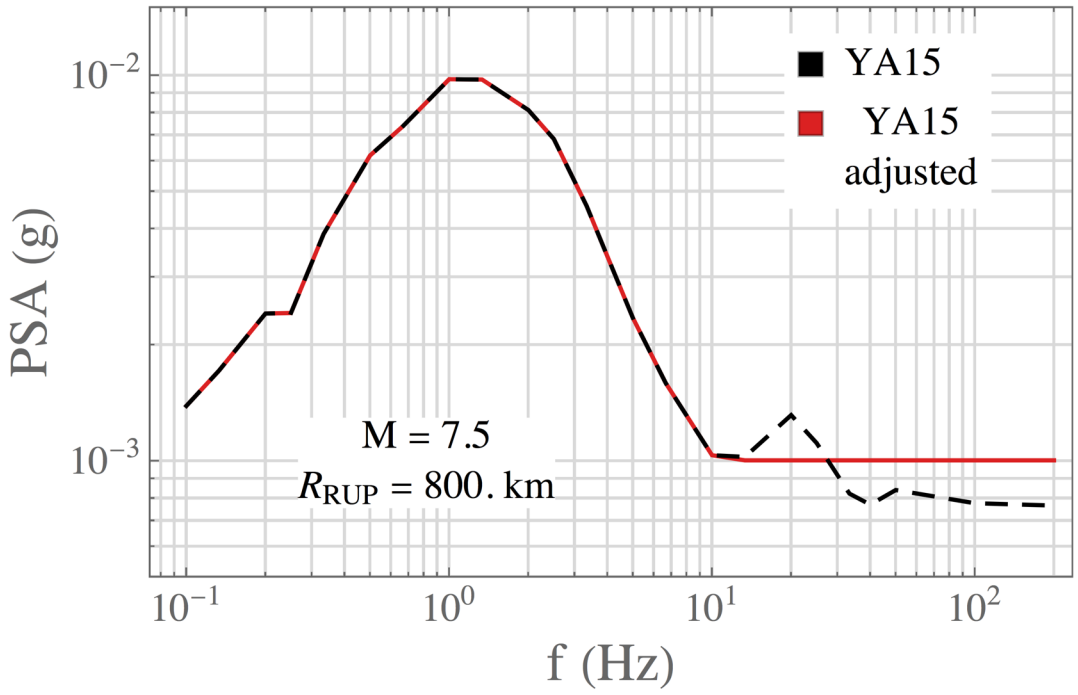


Figure 7–79 Example of spectral-shape correction showing the YA15 spectrum for M = 7.5 and R_{RUP} = 800 km. The dashed black line is the original spectrum showing oscilliations at high frequencies. The red line shows the adjusted spectrum.



8. NGA-East Median Ground-Motion Characterization

The current chapter describes Steps 2-4 of the NGA-East approach to quantify the epistemic uncertainty in median ground motions (Figure 6.9, and steps listed below). The five steps are repeated here for convenience:

1. Develop a suite of seed ground-motion models (GMMs)
2. Develop parameters for continuous distributions of GMMs
3. Visualize the ground-motion space and sample GMMs
4. Re-discretize the ground-motion space
5. Assign weights

The general principles were introduced in Chapter 6. Step 1 was covered in Chapter 7, and Steps 2–4 are elaborated in this chapter. Step 5 is detailed in Chapter 9.

8.1 Parameters for Continuous Distribution of Ground-Motion Models

The NGA-East approach to median ground motions is based on the principle that their associated epistemic uncertainty can be described by a continuous distribution. This was illustrated for one particular moment magnitude and rupture distance (**M**, R_{RUP}) scenario in Figure 6–4, where the (discrete) estimates from individual seed models were fitted by a continuous normal distribution. We are proposing to extend this approach from one to multiple (**M**, R_{RUP}) scenarios, by assuming that the values describing median ground-motion (in log space) would follow a multivariate normal distribution, described by a mean and a covariance matrix. This extension of the 1D approach to a large number of scenarios is based on the general assumptions applicable to all multivariate normal distributions, notably that each marginal distribution is normal (i.e., a distribution-shape assumption) and that each conditional distribution is also normally distributed. The assumption of a multivariate normal distribution helps to motivate the following discussion, because it makes the derivation of the variance and correlation model easier. Later, the multivariate distribution is extended to be a mixture distribution of multivariate normal distributions, which allows deviations in the marginal distributions at each (**M**, R_{RUP}) scenario from a normal distribution.

For the moment, we assume that for each of the NGA-East frequencies, the joint distribution of median ground-motion estimates at different (**M**, R_{RUP}) scenarios is a multivariate normal distribution:

$$P(\mathbf{Y}) \sim N(\mu, \Sigma) \tag{8–1}$$

where $\mathbf{Y} = \{Y_1, \ldots Y_{N_D}\}$ is a vector of random variables describing the (uncertain) median ground-motion values at N_D different (**M**, R_{RUP}) scenarios, μ is a vector characterizing the mean of the uncertainty distribution at each (**M**, R_{RUP}) scenario, and Σ is the covariance matrix between the median ground-motion estimates at the different (**M**, R_{RUP}) scenarios. The probability distribution of **Y** describes the epistemic uncertainty in median estimates. The **Y** distribution can be written as

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_{N_D} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_{N_D} \end{bmatrix}, \left(\begin{bmatrix} \Sigma_{11} & \cdots & \Sigma_{1N_D} \\ \vdots & \ddots & \vdots \\ \Sigma_{N_D1} & \cdots & \Sigma_{N_DN_D} \end{bmatrix}\right)\right) \quad (8\text{–}2)$$

where the entries of the covariance matrix are

$$\Sigma_{ij} = \rho_{ij}\sigma_i\sigma_j \quad (8\text{–}3)$$

with σ_i as the standard deviation of the ground-motion distribution for the ith (**M**, R_{RUP}) scenario, and ρ_{ij} describing the correlation between the ith and jth (**M**, R_{RUP}) scenario.

A sample from the distribution described by Equations (8–1) or (8–2) is a vector of median ground-motion estimates of **Y** at the considered (**M**, R_{RUP}) scenarios. The values from a more traditional GMM at these (**M**, R_{RUP}) scenarios also lead to a vector of median ground-motion estimates **Y**. For a sufficiently large (ideally infinite) number of (**M**, R_{RUP}) scenarios, a sample from the median ground-motion distribution can be considered a continuous function of **M** and R_{RUP}; therefore, samples selected from P(**Y**) are GMMs.

The P(**Y**) distribution described by Equations (8–1) or (8–2) is a full description of the center, body, and range of epistemic uncertainty associated with median ground-motion estimates. The main task required to estimate the P(**Y**) distribution is to estimate its mean and develop the covariance matrix, Σ, which consists of two parts:

- a model for the diagonal elements of Σ, which are the variances σ_i^2 for the ith (**M**, R_{RUP}) scenario.
- a model for the correlation coefficients, ρ_{ij}, describing the correlation between median estimates for the ith and jth scenario.

Given the covariance model, and thus P(**Y**), the approach to discretize the median ground-motion distribution relies on drawing a large number of samples (e.g., a large number of sampled models; see 6–5 which shows the PDF of those scenarios).

For the NGA-East project, the (**M**, R_{RUP}) combinations considered for the generation of new models are:

- **M** = 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 7.8, 8, 8.2
- R_{RUP} = 0,1, 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 175, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 1000, 1200,1500 km

These magnitude and distance ranges are selected to capture all important trends in scaling and their potential influences on the seismic hazards at various sites. The magnitude scaling for magnitudes up to about 7.5 follows a near-linear trend, with a tendency to curve down at larger magnitudes (saturation that is more pronounced at close distance for high frequencies). To capture the linear trend, we deemed bins of 0.5 magnitude units to be sufficient. For the large magnitudes, we reduced the bin size to 0.2, to capture the change in slope due to potential

magnitude saturation. The scaling of ground motion with distance exhibits a more nonlinear behavior, especially within 50 km (e.g., due to geometrical spreading and near source saturation) and near the Moho bounce zone, which extends roughly from about 70 to 150 km. Thus, the distance bins are narrower in these zones. Beyond 500 km, the distance scaling is dominated by a linear trend, thus, fewer distance values are needed.

In total, there are 374 (**M**, R_{RUP}) scenarios at which the seed GMMs are considered for the estimation of the covariance matrix. As is discussed later in this chapter (Section 8.4), subsets from this scenario list were considered for the sampling of new models. Even in cases where fewer scenarios were used for the sampling process, the larger set is important in the generation of the covariance matrix, as it ensures the empirical correlation structure is preserved for the full range of model applicability.

8.1.1 Variance Model

For one (**M**, R_{RUP}) scenario, it is easy to fit a continuous normal distribution from the median estimates coming from the seed GMMs. The normal distribution is described by the sample mean and sample variance of the seed GMMs (see Figure 6–4). For the multivariate case, the diagonal entries of the covariance matrix $\sum$ can be estimated by the sample variance

$$\sigma_j = \frac{1}{N_S - 1} \cdot \sum_{i=1}^{N_D} (y_{ij} - \bar{y}_j)^2 \qquad (8\text{–}4)$$

Here, σ_j is the sample variance for the j^{th} (**M**, R_{RUP}) scenario, and N_S is the number of seed GMMs. Figure 8–1 shows eight contour plots of the diagonal entries of the seed GMMs (i.e., the sample) covariance matrix for a subset of representative frequencies. Each point in the contour plot corresponds to the sample variance. In other words, Figure 8–1 shows the variance from the range of seed GMMs for each (**M**, R_{RUP}) scenario. Figure 8–2 shows the data (recordings) available in the NGA-East database, capped to 400 km to prevent bias in the dataset, as discussed in Section 7.6. It is expected that the variance would be lowest for (**M**, R_{RUP}) scenarios well populated with data. The variance values shown in Figure 8-1 vary greatly with distance, but remain generally constant with increasing magnitude. This is counter-intuitive since the largest recorded event is at **M** = 5.8, yet the epistemic uncertainty represented by the variance is the same for **M** = 7 at the 200 km distance as it is for lower magnitudes. Figure 8–3 shows the same trend in the form of a magnitude-scaling plot at R_{RUP} = 200 km for 1 Hz motions (compare to Figure 8–1b). In addition, the variances do not always vary smoothly with magnitude and distance for all frequencies. This is due to the finite sample size and strong differences in functional forms for some models, which accentuate differences in estimates at certain (**M**, R_{RUP}) scenarios, and thus, lead to jagged changes in variance across scenarios.

The issues with the variance (and ultimately with the epistemic uncertainty range covered by the seed GMMs) were addressed by the TI team through the development of a target variance to be combined with that of the seed GMMs. The variance model expresses epistemic uncertainty in GMMs, considering the alternative GMMs developed during the project and other factors, such as the data constraints and insights gained from analyses in more data-rich regions. This additive variance model is effectively what we develop. We prefer to work in variance space, as

variances are additive and, in that sense, their combination into our model is easier to interpret than standard deviations, which need to be squared, added, and then square rooted.

The following guiding principles were used to develop the variance model.

1. The epistemic uncertainty is partly reflected by the variance computed from seed GMMs; however, it is not perfectly captured. As was discussed earlier, low dispersion at certain (**M**, R_{RUP}) scenarios can be due to the development process (through a point source model, for example) and not necessarily confirmatory that the ground motions are modeled correctly.
2. The epistemic uncertainty in GMMs in CENA should be as large or larger than the corresponding uncertainty in GMMs for shallow crustal events in active tectonic regions (such as WNA). The rationale is that the WNA database is much richer than that for CENA, especially in the magnitude-distance ranges of most hazard-interest. Also, knowledge of empirical ground-motion behavior is greater in WNA than in CENA.
3. The epistemic uncertainty should vary relatively smoothly in magnitude-distance space, being larger in areas where there are limited observed data to constrain the CENA GMM seed development.

Ultimately, the goal was to develop a variance model that did not completely obliterate the variance contained in the seeds, but that also represented a larger level of epistemic uncertainty for the full range of (**M**, R_{RUP}) scenarios in the NGA-East scope. This was a challenging task to achieve, and consequently we revised our initial approach, which consisted of taking the mean of all the seed GMMs to generate samples from the distribution (with that mean and the modeled variance). Instead, as is discussed in more details in Section 8.3.1, we favored using each seed GMM to generate samples, preserving more of their trends in magnitude and distance scaling. This means that instead of using a pure normal multivariate distribution as discussed above, we are using a mixture model. The TI team's evaluation is that this approach is more appropriate to better capture the epistemic uncertainty from the seeds in terms of magnitude and distance scaling. The selected target variance allowed the TI Team to achieve a final variance model that reflected our collective understanding of the epistemic uncertainty expressed by the models and developers. Those results are shown in Chapter 9, because the final median variance must be computed using the final GMMs and their weights. The remaining text in this section provides the process by which the TI team developed the target variance.

The development of the variance model was a critical task for the TI team and it relied in part on recent work conducted in WNA, notably from the South Western U.S. (SWUS) utilities project (GeoPentech 2015). The SWUS project followed a SSHAC Level 3 process to constrain ground motions for two nuclear power plants (NPPs) and represents an extensive study of epistemic uncertainty for active crustal regions. We summarize a few key attributes of the SWUS process and findings and use the model from the Diablo Canyon power plant (DCPP) as a constraint for the target variance for NGA-East (results from the other NPP, Palo Verde, are not significantly different).

For DCPP, the SWUS team evaluated the epistemic uncertainty from the following GMMs:

- Abrahamson et al. (2014) (ASK14)

- Boore et al. (2014) (BSSA14)
- Campbell and Bozorgnia (2014) (CB14)
- Chiou and Youngs (2014) (CY14)
- Idriss (2014) (I14)
- Akkar et al. (2014a, b)
- Zhao et al. (2014)
- Zhao and Lu (2011), used for adjustment to magnitude scaling

The first five models were developed as part of the NGA-West2 project (Bozorgnia et al. 2014), and the three additional shallow crustal earthquakes models were developed independently for Europe and the Middle East (Akkar et al. 2014a, b) and Japan (Zhao and Lu 2011; Zhao et al. 2014). The inclusion of a range of GMMs in addition to the NGA-West2 GMMs was deemed necessary as the NGA-West2 GMMs are based on a common database, come from a highly-collaborative GMM development project and follow similar approaches, which could underestimate the epistemic uncertainty. The focus of the DCPP model development was for scenarios of **M** up to 8.5 and R_{RUP} up to 70 km, although the models listed above were developed for distances up to 200 km.

Figure 8–4 shows the variance from the eight models above for the range of applicability of the GMMs for a subset of frequencies. Again, the epistemic uncertainty would be expected to be lower in areas where data are plentiful, but this is not the case. Figure 8–5 shows a summary of the data available in the NGA-West2 database and for which subsets were selected by the various GMM development teams, while Figure 8–6 shows the data used by the SWUS project for the computation of residuals (this subset is consistent with the desire to obtain a model valid for distances up to 70 km, as mentioned above). The largest variance in the 10 Hz plot (Figure 8–4e) corresponds roughly to around **M** = 5.5 and R_{RUP} in the 50–150 km range, where there is plenty of data and yet the variance at **M** > 7.5—where data are limited—is lower.

This trend is illustrated using the NGA-West2 subset of GMMs in Figure 8–7. Figure 8–8 shows the variance achieved from the same SWUS GMMs when the epistemic uncertainty model from Al Atik and Youngs (2014) is added. The increase in variance is minimal relative to the variance from the GMMs alone (Figure 8–4) for most scenarios. The SWUS project team believed that the epistemic uncertainty captured by these models was not appropriate, and they took steps to increase the variance range in the development of the NPP models. The process for better representing the epistemic uncertainty is described in GeoPentech (2015) and not repeated here; only results for the DCPP model are shown in Figure 8–9. We consider this SWUS model (Section 6.2.8) to provide a better representation of the range of ground motions in WNA and use them as the starting point for constraining the variance in CENA.

Target variances for NGA-East at the tail of the magnitude and distance ranges (8.2 and 1500 km) were guided by the results from SWUS (Figure 8–9). The SWUS models are best constrained at distances of 50 km and yet were designed to work up to 70 km. Considering these distances, we compiled the variances at **M** = 8 (toward the end of the range) for all the frequencies. The results are summarized in Figure 8–10. The average variance across

frequencies was 0.27 for the 50 km and 0.35 for the 70 km case. The TI team, therefore, assigned a target variance of 0.4 for **M** = 8.2, which is the upper end of the NGA-East magnitude range—across all distances beyond 40 km—based on the assumption that the SWUS values represent a lower bound at these large magnitudes. At distances of zero km, even for large magnitudes, we lowered the target variance to 0.15 to allow models to remain physical (and to allow for saturation). The maximum distance of 1500 km was also assigned a 0.4 variance although the seeds already span a larger variance at distances beyond 1000 km. Finally, the range for which we expect the smallest variance is in the data-rich scenarios spanning the **M** 4 to 5 range and distances between about 150 and 400 km (Figure 8–2), to which we assigned a target variance of 0.1. The same variance model was used for all the frequencies. The TI team could not justify the use of frequency-dependent variance models (based on the limited state of knowledge and constraints), and, therefore, used the average at relevant scenarios for guidance, as described above. The variance values described above are highlighted in blue in Figure 8–11. The intermediate values were linearly interpolated to provide a smooth variance target between scenarios, which is illustrated in Figure 8–12. In Appendix E.1, we provide hazard sensitivity results using the final NGA-East models, the seeds only, and the seeds when the Al Atik and Young (2014) uncertainty model is used. The definition of the variance values to add to the distribution of seed GMMs is a critical step of the NGA-East approach (Step 2 in Figure 6–18) that required several tests and sensitivity studies. Appendix E.2 presents sensitivity analyses when the variance model is constrained for different maximum values: 0.3, 0.4 (selected values for the final models), and 0.5. Additional considerations and discussions are presented on this topic in Section 12.3 in relation to the application of the NGA-East calibrated model to the EPRI seeds. All the other steps to generate new models and compute hazard for the aforementioned sensitivity analyses follow the procedures described in Chapters 8 and 9.

8.1.2 Correlation Model

In this section, the model for non-diagonal entries of the covariance matrix, i.e., the correlation coefficients ρ_{ik} [Equation (8–3)], is developed. As for the diagonal entries, it is possible to calculate the sample covariance—and hence the sample correlation. The sample here consists of the ground motion estimates from the seed GMMs. The sample covariance can be calculated as

$$q_{jk} = \frac{1}{N_S - 1} \sum_{i=1}^{N_S} (y_{ij} - \bar{y}_j)(y_{ik} - \bar{y}_k) \qquad (8\text{–}5)$$

where y indicates mean values. The correlation coefficients ρ_{jk} can be computed from q_{jk}. The correlation coefficients ρ_{jk} describe the correlation between the median ground-motion predictions at the j^{th} and k^{th} (**M**, R_{RUP}) scenarios. The sample correlation is shown in Figure 8–13, for three (**M**, R_{RUP}) scenarios and f =1 Hz. In each of those plots, the contours show the correlation between all (**M**, R_{RUP}) scenarios relative to a reference scenario indicated by a black dot.

It is possible to directly use the sample correlation between the different (**M**, R_{RUP}) scenarios to sample from the ground-motion distribution P(**Y**). However, the sample correlation coefficients for Figure 8–13 do not vary smoothly over the full (**M**, R_{RUP}) range due to the finite sample size provided by the seed GMMs. This leads to samples that do not behave in a physically predictable way, producing jagged scaling in magnitude and distance. To overcome this problem, the TI team preferred to model the correlation structure.

As mentioned below Equation (8–5), the correlation coefficients depend on magnitude and distance, and can thus be modeled as a function of **M**, R_{RUP} and can be written as

$$\rho_{jk} = \rho\left[\left(\mathbf{M}, R_{RUP}\right)_j, \left(\mathbf{M}, R_{RUP}\right)_k\right] = f\left(\mathbf{x}_j, \mathbf{x}_k\right) \tag{8-6}$$

where $\mathbf{x} = \{\mathbf{M}, \ln R_{RUP}\}$ is a vector describing the (**M**, R_{RUP}) scenario, and $f\left(\mathbf{x}_j, \mathbf{x}_k\right)$ is a function whose output is the correlation coefficient between scenario **x** and **x′**. For Equation (8–3) to provide a valid covariance matrix, $f\left(\mathbf{x}, \mathbf{x}'\right)$ needs to lead to a positive definite matrix for all combinations of **x** and **x′**. This can be accomplished by borrowing from the field of Gaussian Process (GP) regression. Chapter 4 of Rasmussen and Williams (2006) provides an extensive overview of covariance functions used for GP regression, which always result in a positive definite covariance matrix. To estimate the correlation coefficients ρ_{jk}, the parameters of a covariance function $k\left(\mathbf{x}, \mathbf{x}'\right)$ suitable for the problem are estimated, and the correlation coefficients are calculated according to

$$\rho_{jk} = \frac{k(\mathbf{x}_i, \mathbf{x}_j)}{\sqrt{k(\mathbf{x}_i, \mathbf{x}_i)}\sqrt{k(\mathbf{x}_j, \mathbf{x}_j)}} \tag{8-7}$$

$$k(\mathbf{x}, \mathbf{x}') = k_{stat}(\mathbf{x}, \mathbf{x}') + k_{lin}(\mathbf{x}, \mathbf{x}') \tag{8–8}$$

where the first part is the isotropic covariance function, and the second part is the linear covariance function. This covariance function provides a compromise between assuming no correlation, i.e., independent expert elicitation for different (**M**, R_{RUP}) scenarios and assuming full correlation (i.e., a scaled-backbone approach). The correlation coefficients corresponding to the no-correlation approach are ρ_{jk}=1 if j = k and zero otherwise, whereas the correlation coefficients for the scaled-backbone approach are ρ_{jk}=1 for all j and k. The covariance function in Equation (8–8), however, populates the off-diagonal entries of the correlation matrix with values that can range from zero to one (but not all zero or all one), providing the desired range of correlation.

For NGA East, a covariance function is estimated from the predictions of the seed GMMs. Using this covariance function, the covariances between all (**M**,R_{RUP}) scenarios for NGA-East are calculated. The covariances are converted to correlations using Equation (8–7). Then, the correlations between the different scenarios are combined with the NGA-East target variance model (Figure 8–11 and 8–12) to calculate the desired covariance matrix [cf. Equation (8-3)].

This covariance matrix is used to sample new GMMs. We will now describe the functional form of the covariance function and the estimation of its coefficients.

The functional form selected for the covariance function is

$$k(\mathbf{x}, \mathbf{x}')^{T} = \theta_1 \left[1 + (\mathbf{x} - \mathbf{x}')^{T} \begin{Bmatrix} \theta_2 & 0 \\ 0 & \theta_3 \end{Bmatrix}^{-2} \frac{(\mathbf{x} - \mathbf{x}')}{(2\theta_4)} \right]^{-\theta_4} + \mathbf{x}^{T} \begin{Bmatrix} \theta_5 & 0 \\ 0 & \theta_6 \end{Bmatrix} \mathbf{x}' \qquad (8\text{–}9)$$

where, again, the first part corresponds to the isotropic covariance function, and the second part depends on the dot-product $(\mathbf{x}, \mathbf{x}')$ modeling the linear trend of ground motion with magnitude and distance.

The isotropic part, which is also called a rational quadratic covariance function and contains four parameters: θ_1, θ_2. θ_3, and θ_4. Parameter θ_1 describes the amount of variance described by the rational quadratic part, θ_2 and θ_3 describe the two length-scales relative to magnitude and distance, respectively, and θ_4 describes a mixture of length-scales. The length-scale informally describes how much the correlation is preserved between two different values of the predictor variables (**M** or R_{RUP}). For example, considering two scenarios with **M** = 4 and 6: for a small length-scale, the ground motions corresponding to those two scenarios will be less correlated than if the length-scale was larger. If there is only one dimension, the rational-quadratic covariance function becomes:

$$k_{rq}(x, x') = \left(1 + \frac{r^2}{2L^2 \alpha} \right)^{-\alpha} \qquad (8\text{–}10)$$

where $r = (\mathbf{x} - \mathbf{x}')$. In this case, the parameter α corresponds to parameter θ_4 in Equation (8–9), and L is the length scale, corresponding to θ_2 and θ_3 in the two-dimensional case.

Figure 8–14 shows an example of samples created using Equation (8–10). It shows the effect of the length-scale L on sampled functions with a rational-quadratic covariance function. Here, to generate the sampled functions, the covariance function of Equation (8–10) is evaluated at $x_i \in \{0, 0.1, 0.2, \ldots, 10\}$, which results in a 101 $\times$ 101 dimensional covariance matrix. Then, function values are sampled from a multivariate normal distribution with mean zero and the calculated covariance matrix. These are plotted as function values at the x_i values in Figure 8–14, which shows that with increasing length-scale L, the sampled functions become smoother. In other words, L controls the influence of a particular point x_i: the larger L, the larger the correlation at longer distances r [defined above, representing the distance between two scenarios]. The (point-wise) variability of the sampled functions in the y-direction is controlled by the variance parameter θ_1, which was dropped (set equal to one) for Equation (8–10); a larger overall variance means a larger spread of the sampled function values at one particular x_i.

Figure 8–15 shows the effect of different values for α on the sampled function values. The effect of different values for α is smaller than for the length-scale parameter L. Figure 8–16 shows k_{rq} as a function of the input distance r for different values of L and α.

The second part of the covariance function is a dot-product term, which models a linear dependence between the inputs and the outputs. This part is not isotropic, which means that the correlation does not depend on the distance between the inputs. For a one-dimensional input x, this part reduces to

$$k_{lin} = (x, x') = \sigma_{lin}^2 x\, x' \tag{8–11}$$

Figure 8–17 shows five sampled functions from a GP with linear covariance function. The sampled functions are straight lines [see Rasmussen and Williams (2006)], and the parameter σ_{lin}^2 controls the spread of the sampled function values at each x_i. In contrast to the rational quadratic covariance function (which is isotropic), the spread is different at each x_i. It is important to remember that the functions in Figure 8–17 are sampled from a multivariate normal distribution with 101 dimensions; the function values at each x_i are sampled.

The covariance function used in NGA-East is a sum of a rational quadratic part and a linear part. This can be thought of as a superposition of a general linear trend of the function values (i.e., ground-motion values) with magnitude and distance, while the rational quadratic part models any nonlinearities. Parameters θ_2 and θ_3 control the influence of a particular **M** and R value and correspond to the length-scale L in Equation (8–10). Similarly, θ_5 and θ_6 control the possible slopes of the linear functions with **M** and R_{RUP}.

The rational quadratic covariance was chosen because it can be thought of as a mixture of squared exponential covariance functions and thus can accommodate several length-scales (Rasmussen and Williams 2006).

The parameters $\theta = \{\theta_1, \ldots, \theta_6\}$ are estimated by maximizing the marginal likelihood $p(\mathbf{y}|\mathbf{X}, \theta)$ of the mean estimates of all the seed GMMs. Here, $\mathbf{y} = \{y_1, \ldots, y_N\}$ is the vector containing the median predictions of the mean seed model, and $\mathbf{x} = \{x_1, \ldots, x_N\}$ is the vector containing all (**M**, R_{RUP})-scenarios. The log marginal likelihood can be written as

$$\ln p(\mathbf{y}|\mathbf{x}, \theta) = -\frac{1}{2}\mathbf{y}^T K_Y^{-1}\mathbf{y} - \frac{1}{2}\log|K_Y| - \frac{N_D}{2}\ln 2\pi \tag{8–12}$$

where $K_Y = K_f + \beta^2 I$; I is the identity matrix. The matrix K_f is the covariance function evaluated at the different (**M**, R_{RUP}) scenarios, so that the elements of K_f are $K_{f,ij} = k(\mathbf{x}_i, \mathbf{x}_j)$. Depending on the values of the parameters, it can happen that the matrix K_f is numerically not positive definite. In this case, adding a small variance term, β^2, on the diagonal elements ensures that K_y is positive definite; therefore, parameter β^2 is only used for numerical stability.

The marginal likelihood is calculated with Equation (8–12). For numerical stability, the predictors **X** and targets **y** are all standardized such that they have a mean of zero and a standard deviation of one. The parameters of the correlation function are calculated for a frequency of f = 1 Hz, and frequency dependence between sampled models is taken into account via the mixture

model of using the seed models as the mean for sampling, as described in Section 8.3. Hence, the estimation of the parameters for the correlation model is done in the following way:

1. At the specified (**M,** R_{RUP}) scenarios for NGA-East, calculate the mean of the seed GMMs for f = 1 Hz.
2. Standardize the magnitudes, distances and median ground-motions (this is only to increase numerical stability).
3. Calculate K_f for some starting values of the parameters θ.
4. Calculate $\ln\left[p\left(\mathbf{y}|\mathbf{X},\theta\right)\right]$.
5. Maximize $\ln\left[p\left(\mathbf{y}|\mathbf{X},\theta\right)\right]$ with respect to θ and β^2, where β^2 is a small value that ensures that the covariance matrix is numerically positive definite.

Given the estimated parameters θ, one can insert them into Equation (8–9) and calculate a covariance matrix for the (**M***,* R_{RUP}) scenarios of interest [or any set of (**M**, R) values]; however, the parameters, in particular, θ_1, are optimized to fit the mean of the seed set using a GP with a mean function of zero. Thus, the diagonal entries of the calculated covariance matrix reflect the spread of the estimates across the (**M***,* R_{RUP}) scenarios. Correlation coefficients can be calculated via Equation (8–3), which reflects the changing of estimates with magnitude and distance. Examples of the modeled correlation are shown in Figure 8–18. The correlation coefficients are combined with the NGA-East variance model to calculate a full covariance matrix, which is then used to sample new GMMs. Together with the variance model (Figures 8–11 and 8–12), the correlation coefficients estimated in this section provide the full covariance matrix. The estimated parameters θ and β^2 are given in Table 8–1. Sensitivity analyses based on the correlation model are provided in Appendix E.3.

Since the linear part of the covariance function (Figures 8–8 and 8–9) is non-stationary, the coefficient θ_1 is not equal to one, even though the mean seed predictions were standardized to have a standard deviation of one. This is due to the fact that there is an interplay between the parameters θ_1, θ_5, and θ_6, controlling the spread of the resulting sampled function (Figures 8–5 to 8–17). If only the rational-quadratic part is used to estimate the parameters θ_1 to θ_4, then the resulting value is $\theta_1 = 1.18$, which is close to the standardized variance of the fitted data.

8.2 Visualization of the Ground-Motion Space

8.2.1 Challenges in Evaluation of Multiple GMMs

There are different approaches for selecting, using and assigning weights on GMMs for hazard analysis, such as selecting different GMMs or a scaled-backbone approach (see Chapter 6 for an overview of approaches). Regardless of the methodology, the center, body, and range of epistemic uncertainty should be assessed. Scherbaum et al. (2010) proposed using high-dimensional visualization techniques to provide a graphical representation of this uncertainty. They used tools such as self-organizing maps (SOMs) (Kohonen 2001) and Sammon's maps (Sammon 1969) to project GMMs onto a two-dimensional map, which can be thought of as a two-dimensional projection of the GMM model space. Ultimately, the full, multidimensional GMM

model space is the space upon which the assessment of center, body, and range should be based on.

High-dimensional visualization tools such as Sammon's maps are a good, practical substitute to compare different GMMs. This is illustrated in Figure 8–19, and it is closely related to the use of an ensemble of GMMs (Atkinson et al. 2014). In the left panel of Figure 8–19, the estimates of the 18 candidate GMMs for one particular scenario are shown as a histogram, for f = 1 Hz. In total, there are 19 seed GMMs; however, PEER_EX is used only above 2 Hz, so there are 18 discrete GMMs in this case (see Table 7–14).

Since there are 18 discrete GMMs, there are 18 discrete median ground-motion estimates. However, as discussed previously, there should be a continuous distribution of estimates. The intermediate values are also likely valid; they are simply not generated with the given models. For a single scenario (one dimension) one can fit a continuous distribution to the estimates (in Figure 8–19 the best fitting normal distribution is shown) and evaluate the estimates and the continuous model graphically. For two scenarios, this is still possible, as shown in the right panel of Figure 8–19. For more than three scenarios, the simultaneous (graphical) evaluation of different GMMs becomes almost impossible. This poses a problem because an assessment of the models and their corresponding continuous distribution is required for many different (**M**, R_{RUP}) scenarios [such as those sampled by probabilistic seismic hazard analysis (PSHA)].

For the single scenario case, the problem is straightforward: the difference between the GMMs is simply the difference between their estimates (see also Section 6.3). If one keeps all predictor variables fixed and changes only the distance, one can plot the ground-motion estimates versus distance. In that case, it is possible to distinguish trends and differences in distance scaling if only a few GMMs are evaluated. With an increasing number of GMMs, it becomes more difficult to assess differences between GMMs. As an example, Figure 8–20 shows the magnitude and distance scaling of the 18 aforementioned GMMs for 1 Hz. Although the range in ground motions can be assessed, it is difficult to distinguish differences between individual models. A systematic evaluation would require plotting all the combinations of only two models at a time for one particular magnitude or distance, leading to a large number of plots. Even more plots are needed to compare the models at other magnitudes and distances.

To capture the center, body, and range of median GMM estimates, it is necessary to assess the similarity of GMMs over a wide range of magnitudes and distances that are relevant to the hazard at the site. Projecting the GMMs into two dimensions is a convenient way to achieve this. Basically, the GMMs are mapped onto a projection that allows them to be assessed *visually*.

The basic idea is that GMMs (for any specified value of frequency) "live" in some abstract model space. This space can be thought of as the space of all the possible functions of magnitude and distance. The GMMs form a subspace associated the physically realizable instances of this function space. The basic assumption is that the GMMs occupy a lower dimensional manifold in the larger model space. If this manifold is two-dimensional, it is possible to plot it on a map and it is then easy to visualize.

The assessment of the model space involves the following assumptions:

- The median estimate from a GMM is a function.
- By evaluating the GMM at certain values of its predictor variables [(**M**, R_{RUP}) scenarios], it is discretized.
- If the GMM is evaluated at N_D different values of its predictor variables (**M**, R_{RUP}), it can be represented as an N_D -dimensional vector of ground-motions, one entry for each (**M**, R_{RUP}) pair.
- Each GMM is evaluated at the same N_D values of the predictor variables (**M**, R_{RUP}). This means that each GMM corresponds to a point in the N_D -dimensional ground-motion space. This is an approximation of the GMM model space.
- The N_D -dimensional space can be projected to two dimensions.

There are several different methods to project high-dimensional data onto a lower-dimensional space, such as those such as described in Hotelling (1933), Gianniotis and Riggelsen (2013), and Lawrence (2004). Scherbaum et al. (2010) used SOMs (Kohonen 2001) and Sammon's mapping (Sammon 1969). NGA-East uses Sammon's mapping, which is described in the following sub-section. It is a relatively simple method, and, in contrast to SOMs, it works on a continuous scale.

8.2.2 Introduction to Sammon's Maps

Sammon's mapping is a nonlinear dimensional reduction technique. In Sammon's mapping, a configuration of points in two dimensions is sought that resembles the distance distribution in high dimensions. In this case, distance describes the difference between coordinates (i.e., between their median estimates), and is not a physical distance like R_{JB} or R_{RUP}. Throughout this report, the term GMM-distance or $\overline{\Delta_{GM}}$ is used to describe differences between GMMs in ground-motion space. The misfit function is the difference between the GMM-distances in high dimensions and two dimensions (which is called Sammon's stress):

$$E = \frac{1}{\sum_{i<j} \overline{\Delta_{GM\,ij}}} \sum_{i<j} \frac{(\overline{\Delta_{GM\,ij}} - \overline{\Delta_{map\,ij}})^2}{\overline{\Delta_{GM\,ij}}} \qquad (8\text{–}13)$$

where $\overline{\Delta_{GM_{ij}}}$ is the GMM-distance between GMMs i and j in high-dimensions, and $\overline{\Delta_{map_{ij}}}$ is the corresponding shortest distance on the map (in two dimensions), respectively. Δ_{map} is the shortest path between two points on the map, which is the Euclidean distance in two dimensions. To produce a two-dimensional projection, E is minimized with respect to the positions in two dimensions through an iterative process, e.g., via gradient descent. As a starting configuration, a random set of points on the map can be used. Another common choice is to use the output of principal component analysis (PCA) (Hotelling 1933) as a starting configuration.

From the description of Sammon's mapping, one can appreciate that only the relative positions of GMMs on the map are important, since only the relative distances $\overline{\Delta_{GM_{ij}}}$ and $\overline{\Delta_{map_{ij}}}$ are used in the calculation of the Sammon's stress E. These maps can be rotated or mirrored in any way and still preserve the high-dimensional distance on the map—the information content of the maps is the same. The absolute coordinates depend on the starting configuration of the points on the map, as illustrated in the examples below.

8.2.3 Sammon's Map Conceptual Examples

A real-life analogy to this concept can be made by considering the geographical distribution of three cities. For example, the distance between San Francisco and Los Angeles is 552 km, the distance between San Francisco and Las Vegas is 662 km, and Los Angeles is 377 km away from Las Vegas. This is sufficient information to infer the relative geometry of those three cities on a map; however, it does not tell us which of the cities is the northernmost. For some applications, like estimating the cost of jet fuel, this can be all that is needed. If proper orientation is needed, the map can be mirrored and rotated without affecting the distances between the cities.

The second example is similar in spirit to the situation of comparing GMMs, but the map can be evaluated intuitively. We show an example visualization that exemplifies the power of high-dimensionalization tools. Figure 8–21 shows 18 pictures of rubber ducks as an intuitive example borrowed from Geusebroek et al. (2005). The rubber ducks in the 18 individual pictures show the same duck, each time rotated by 20°. Each of the pictures is a grid of 32×32 pixels, and each pixel is associated with a red, green, and blue value. Hence, each picture can be represented by a 3072-dimensional vector, that is, each picture is a point in a 3072-dimensional "rubber duck"-space. The analogy to the situation for a GMM is that each pixel corresponds to a particular magnitude/distance scenario. Figure 8–22 shows a Sammon's map calculated for the 18 pictures of the ducks. The map is easily interpretable: pictures that are close (one rotation apart) are close on the map, because all their dimensions (pixels) are more similar pair-wise than ducks that plot on different sections of the map (larger rotation angle). Hence, Figure 8–22 shows the potential of using visualization techniques to reveal structure in a high-dimensional dataset. Again in this case, the maps can be mirrored or rotated, but a given duck will remain at the same map-distance from all the other ducks.

In a Sammon's map for GMMs, the map distances correspond to high-dimensional GMM-distances, which have the same units as the ground-motion estimates themselves. As described earlier, the high-dimensional coordinates of a GMM correspond to its estimates at different (**M**, R_{RUP}) scenarios. Hence, GMM-distances in high dimensions can be easily calculated from the differences in their estimates. There are, however, various GMM-distance metrics that one can use. For NGA-East, the Euclidean distance between two GMMs is used, which is based on the L_2-norm:

$$\overline{\Delta_{GMij}}(L_2) = \sqrt{\frac{1}{N_D}\sum_{k}^{N_D}\left(GMM_{ik} - GMM_{jk}\right)^2} \qquad (8\text{–}14)$$

where k indexes the different values of the predictor variables (**M**, R_{RUP}), N is the number of dimensions, $\overline{\Delta_{GM_{ij}}}(L_2)$ is the L_2-distance between GMMs i and j, and GMM_{ik} is the ground-motion prediction of the ith GMM corresponding to the kth (**M**, R_{RUP}) scenario. The difference $GMM_{ik} - GMM_{jk}$ between two GMMs is illustrated in Figure 8–23. The factor $\sqrt{1/N_D}$ in the definition of the L_2-distance is there to normalize the GMM-distance such that it has units of ground motions and can be more easily interpreted.

The Sammon's map for the 18 NGA-East GMMs (the same ones that are shown in Figures 8–19 and 8–20) can then be generated. The input to the Sammon's map are the median prediction for PSA at a frequency of f = 1 Hz for magnitudes **M** =4.5, 5.,…,7.5 and distances R_{RUP} = 10, 20,…, 400 km. (This example has a different range than the one defined for the final computations.) Because each GMM is evaluated at 161 (**M**, R_{RUP}) scenarios, it can be represented as a point in 161-dimensional ground-motion space, where each coordinate corresponds to the prediction for one particular magnitude/distance pair. Under the assumption that physically plausible GMMs reside on a two-dimensional manifold (i.e., a surface that is locally flat), we can estimate this manifold using Sammon's maps.

To facilitate interpretation of the map, reference models are added to the set of 18 GMMs:

- the arithmetic average of all model predictions (log-space), hereafter called the *mix*
- scaled versions of the average model: *mix*+lnα, with α=0.67,0.8,1.25,1.5, called −−, −, +, ++
- the average model with changed magnitude scaling: *mix* + β(**M**−6), with β =−0.4,−0.2, 0.2, and 0.4, called M−−, M−, M+, and M++
- the average model with changed distance scaling: *mix* + γ(lnR − ln100), with γ=−0.5,−0.25, 0.25, and 0.5, called R−−, R−, R+, R++

The scaling of these reference models is shown in Figure 8–24. The reference models can also be used to orient the maps in a consistent way. Figure 8–25 shows the following adjustments: the map is centered with the *mix* model at the point {0, 0}, the map is then rotated such that the model ++ is to the right and the line from −− to ++ is horizontal. In a last step, the map is mirrored about the y-axis such that the M++ model is in the upper half of the model.

In general, the reference models help the interpretation of directions in which GMMs change in a systematic way. Because the different reference models (up/down-scaled, changed magnitude scaling, changed distance scaling) align in different directions, they allow a quick assessment of differences between GMMs over a wide magnitude/distance range–the range that was used to generate the map–in a qualitative manner. In particular, the three Boore models that have a near-source distance scaling proportional to $R^{-1.3}$ line up in the upper half of the map, where also the reference model with R−− resides, which has a steeper attenuation.

8.3 Sampling the Ground-Motion Space

Now that we have developed the covariance model and that we can use high-dimension visualization tools, the next step is to sample models from that distribution and to visualize them.

8.3.1 Retaining the Scaling and Modeling Assumptions of the Original Seed Models

The selected correlation model allows for some variation in the magnitude and distance scaling around the mean function. This behavior is a desired feature of the correlation model because it retains the characteristics of the members of the seed GMMs set. This is shown in Figure 8–26, which shows ten sampled functions each, using three different seed models as mean function. This can be written as

$$\mathbf{f} \sim N\left(\mu_S, \Sigma\right) \tag{8–15}$$

where **f** is a vector of (sampled) ground-motion values at the (**M**, R_{RUP}) scenarios of interest, Σ is the covariance matrix, calculated as described before (or computed from ground-motion estimates at these (**M**, R_{RUP}) scenarios), and μ_S is a vector containing the estimates of the seed model that is used as the mean function for Figure 8–19. In each of the three cases, the sampled functions are randomized versions of their respective mean (seed) functions. Hence, this lends itself to an approach of sampling that produces sampled functions that are flexible in their scaling and follow the (point-wise) variance prescribed by the variance model (see Figure 8–26) but whose scaling on the other hand is bound by the original seed models. Basically, the ground-motion distribution P(**Y**) is a mixture distribution

$$w_{N,i} \propto N_i \tag{8–16}$$

where N_s is the number of seed models, and w_i are the weights of the individual seed models that need to sum up to one. Each individual distribution is a multivariate normal distribution with the same covariance matrix and a different mean.

The weights w_i are calculated to ensure that similar seed models do not bias the resulting ground-motion distribution P(**Y**). Such a bias could be introduced if two models are developed using very similar methods and subsets of data. These two models would give very similar predictions; however, the similarity is not a confirmation that the predictions are more likely to be correct. Therefore, the weight of each seed model is proportional to how similar its predictions are to other models; if two models provide similar estimates, they receive less weight individually. The process of calculating weights for the individual seed models is based on the Sammon's map process described in Section 8.2, which provide aggregated information on how similar the estimates are for the seed GMMs. For each frequency, a Sammon's map is calculated for the seed models. The Sammon's map is partitioned into a grid of squares with length 0.25 ln units, and each square that contains a seed model is given the same weight. All models inside each square are given equal weight. The weights are computed by:

$$w_j = \frac{1}{N_{sq}} \frac{1}{N_j} \tag{8–17}$$

where N_{sq} is the number of squares that are occupied, and N_j is the number of seeds in the j^{th} square. For example, if there are six squares that are occupied, and the j^{th} square contains four seed GMMs, then each of these four seed GMMs gets a weight of 1/24.

An example of the Sammon's map for the seeds, together with the grid that is used for defining clusters of models, is shown in Figure 8–27 for f = 1 Hz. We have conducted hazard sensitivity analyses for various weighting schemes and determined that the approach described here met the intent of the TI team. Sample results for the Manchester site are shown in Figure 8–28 and additional results are provided in Appendix E.4. Figure 8–28 shows the difference in hazard for the initial grid (final model in black), for applying equal weights to all the seed GMMs (red) and for an alternate grid (blue). Results for the 1 Hz case (Figure 8–28a) show minimal differences in the mean hazard. The largest difference is seen in the upper fractile between the red curve and the black curve for the RLME and total hazard. The results are as expected and indicate that giving similar GMMs the same weight instead of dividing the weight among them tends to narrow the distribution—and hence the hazard—for certain scenarios. Differences between alternate grids are not very significant in terms of shape (blue), although the lower fractiles can move relative to the black lines. For the 10 Hz case, the results are not very different between the three alternatives. This type of observation was consistent among the three sites used for sensitivity analyses (Central Illinois, Manchester, and Savannah), which are presented in Appendix E.4. This shows that the initial weighting scheme does not have an impact on mean hazard, but that dividing the weights among similar models achieves the goal of the TI team to avoid assuming that similar models are confirmatory. This is consistent with the intent of achieving a distribution which is close to mutually exclusive and completely exhaustive (MECE).

Figure 8–29 shows the result of sampling from the mixture distribution and compares samples from the mixture distribution in the top row with samples from a distribution where the mean function is the mean of the seed GMMs in the bottom row. In each case, 100 samples are drawn. In the top row, the samples follow the scaling of the individual GMMs, while in the bottom row there is less variability in scaling, due to having only one mean (no mixture distribution). On the other hand, the (point-wise) range covered at each (**M**, R_{RUP}) scenario is similar.

8.3.2 Screening Models for Physicality

The top row in Figure 8–29 shows 100 sampled functions, which are drawn from the mixture distribution defined in Equation (8–16). Each sample generally follows the respective seed model used to center it at the mean; however, because the samples are drawn from the distribution (correlation coefficients different than 1), there is no guarantee that every sample will exhibit a plausible physical behavior. An example of how an unphysical model could occur would be to draw from a seed model that exhibits a steep scaling with distance; if we were to sample the tail of the distribution for such a model, that sample might exhibit an attenuation so steep as to become unphysical. The TI team's approach was to remove such unphysical models to prevent bias in the NGA-East GMC model. A set of constraints was defined whereby a sampled model needs to meet in order to be considered physical. The constraints were developed by the TI team based on expert judgment regarding considerations in magnitude and distance scaling from published models, and are defined as follows:

- The ground motion at **M** = 7 must be larger than the ground-motion at **M** = 6, for distances R_{RUP} = 10 km and R_{RUP} = 1500 km
- The ground motion at **M** = 6 must be larger than the ground-motion at **M** = 5, for distances R_{RUP} = 10 km and R_{RUP} = 1500 km
- The ground motions at distances R_{RUP} = 0, 1 and 5 km must be monotonically increasing between magnitudes **M** = 4 and **M** = 7.
- The ground-motion distance slope (sl) is defined as $sl(R_{RUP1},R_{RUP2}) = y(R_{RUP1}) - y(R_{RUP2}) / \ln R_{RUP2} - \ln R_{RUP1}$ between distance R_{RUP1} and R_{RUP2}, and must meet the following criteria:
  - sl (10,40)>0.4 (i.e., the distance slope in ground motions between R_{RUP} = 10 km and R_{RUP} = 40 km must be larger than 0.4)
  - sl (40,150)>−0.1
  - sl (150,400) > min[0.45,0.9min(GMM)], where min(GMM) is the minimum slope of the seed GMMs in this distance range.

The values that are used for the constraints are based on physical considerations and the range implied by the seed models, as well as expert judgment of the TI team. The magnitude-scaling constraints are based on the general notion that larger magnitudes should generate larger ground motions but allows for oversaturation at very large magnitudes—as is inherent in some of the seed models. The distance-scaling constraints allow for increasing ground motions with distance in the intermediate distance range of 40–150 km, related to the Moho bounce effect. Within 40 km and beyond 150 km, the constraints ensure that the scaling of the sampled models is negative with distance but not too steep. For the slope beyond 150 km, it is assumed that the attenuation with distance must at least have a slope of 0.5, corresponding to the geometrical spreading; however, at some frequencies, some of the seed models violate these constraints. In those cases, the constraints are adjusted to be larger than the minimum value for the seed GMMs. The minimum slopes of the seed GMMs for the three different distance bins are shown in Figure 8–30. It can be seen that at low frequencies, the minimum slopes of the GMMs are lower than 0.5 for the small and large distance ranges.

The physicality constraints affect the selection of models. For example, if a seed model has characteristics that are close to the physicality criteria (e.g., if it has a slope that is close to 0.45 between 40 km and 150 km), then a sample with this seed as mean is more likely to violate the physicality criteria. This is a desirable feature, since while the seed and its samples should contribute to the ground-motion distribution, they represent the ranges of physically possible models and should thus be given lower weight. By rejecting more samples from the seed models that are close to unphysical, these models are implicitly down-weighted.

Figure 8–30 also shows the minimum difference between estimates at **M** = 7 and **M** = 6 of the seed GMMs. For low frequencies, the differences are large, indicating a strong scaling with magnitude, whereas for higher frequencies, the differences are small for R_{RUP} =10 km, representing almost no scaling. Note that the values in Figure 8–30 are the extreme values obtained from the seed GMMs and do not represent the center or body of their distribution.

Any sampled model that does not meet the physicality criteria is rejected. Hence, the final algorithm to sample new models is as follows:

1. Randomly select one of the seed GMMs using the weights defined in Equation (8–17).
2. Sample from a mixture distribution consisting of a normal multivariate distribution with modeled correlation and imposed minimum variance around a centered seed GMM as mean.
3. Check if sample passes criteria for physicality.
   • if yes, proceed to step 4
   • if no, go to step 1
4. Add sample to list of sampled models.

This is repeated until a predefined number of physical samples are drawn from all the seed GMMs. For NGA-East, the number of samples is set to 10,000. The TI team evaluated the number of samples needed to represent the ground-motion space and found that 10,000 was large enough to be sufficient, yet computationally efficient enough to be practical. Figure 8–31 shows the Sammon's space coverage offered by 1000 models and 5000 superimposed on the 10,000 models.

Figure 8–32 shows the number of samples that fail the criteria for physicality at each frequency. For low frequencies, most samples are rejected due to a failure to meet the constraint on distance scaling, whereas for higher frequencies, rejection is due to magnitude scaling. This is consistent with Figure 8–30, which shows that the minimum slopes of the seed GMMs are close to the constraint at low frequencies. For high frequencies, the minimum magnitude-scaling ratios of the seed GMMs are close to their respective constraint.

Figure 8–33 shows the number of samples, based on each seed model, in the final set of 10,000 sampled models, for selected frequencies in the range 0.1–100 Hz and PGA. The weights defined in Equation (8–14) and Figure 8–27 lead to large differences in the number of models based on a specific seed. Again, when models produce similar ground motions, they collectively share a weight; when models are close to the physicality criteria, less of their spawn are accepted. This is all by design. An example of this is the PZCT_M2ES model at f = 1 Hz, which contributes fewer samples. For this model, the distance scaling from 150 to 400 km is relatively flat and close to the physical constraint [$sl(150,400)_{PZCT\_M2ES} = 0.504$],], leading to more rejections of (flatter) variants based on that model. As discussed before, this is desirable, as seed models that only barely pass the physicality criteria are implicitly down-weighted.

8.3.3 Discussion of the NGA-East Approach to Develop Continuous Distributions of GMM Median Predictions

The NGA-East approach is different from a scaled-backbone approach in several respects. As was mentioned earlier, differences in scaling with magnitude and distance are captured with the NGA-East approach, whereas all the curves are parallel with the typical scaled-backbone concept. Thus, models generated by the NGA East approach can cross when plotted against magnitude or rupture distance. This is something we perceive as an advantage, especially in cases where people use only a subset of the final model set for their analyses. The mean hazard may not change, but the difference in scaling is part of what is normally captured by

different modeling approaches. In addition, in the scaled backbone approach, the model that produces the largest median estimates at **M** = 4 and R_{RUP} = 10 km also produces the largest median estimates at M = 8 and R_{RUP} = 1500 km. That behavior is not realistic, and may lead to an overestimation of the range of epistemic uncertainty. Such a behavior does not occur with the NGA-East approach. For example, when sampling 10,000 models from the NGA-East distribution, with the mean of the seed set as the mean, no single model generated is the largest for all 374 (**M**, R_{RUP}) scenarios considered for NGA-East. When selecting models, one needs to make sure to select models that represent the full epistemic uncertainty with respect to magnitude and distance scaling, as well as the overall level of ground motion. This is ensured by selecting models via an ellipse on a Sammon's map (Sammon 1969), as described in the next section.

8.4 Application of Sammon's Maps to Distribution of Sampled GMMs

For the NGA-East application, a suite of 10,000 models are projected onto Sammon's maps (one map per frequency) and used as a visual aid for the subsequent steps. These models are sampled from the mixture model defined in Equation (8–16). Figure 8–34 shows the magnitude and distance scaling of a subset of 200 sampled models, for f = 1 Hz. Even for this small subset of models, it is inconceivable to assess the center, body, and range by comparing the 200 samples based on scaling plots. However, there is structure in the high-dimensional ground-motion space that these models occupy. Each model is sampled at the 374 (**M**, R) scenarios, representing a point in a 374-dimensional (ground-motion) space, where each coordinate is the prediction at one particular (**M**, R_{RUP}) pair. In Figure 8–35, a three-dimensional subspace is plotted (three dimensions can be plotted), for {**M**, R_{RUP}} = {4, 10}, {**M**, R_{RUP}} = {5, 100}, and {**M**, R_{RUP}} = {7, 400}. Each of the 10,000 sampled models is a point in this three-dimensional space and is plotted as a gray point. In addition, the 18 seed GMMs are plotted as red dots.

As shown in Figure 8–35, the sampled models and the seed GMMs lie approximately on a plane in the three-dimensional ground-motion space, although this plane is not oriented in a clear way relative to the axes. However, Figure 8–35 also shows that there is structure in the ground-motion space; i.e., there is a two-dimensional manifold (a thin cloud) on which the sampled models lie, and thus the idea of using Sammon's mapping to visualize the models in two dimensions makes sense. As described above, each GMM is represented as a point in an N_D-dimensional ground-motion space, which is an approximation of the GMM model space. The GMMs form a thin cloud of points in this high-dimensional space. The thickness of this cloud is ignored, and they are projected into two dimensions.

By contrast, samples from a distribution where the correlation coefficients are zero ($\rho_{jk} = 0$) for $j \neq k$ form a sphere in the three-dimensional ground-motion space (Figure 8–36); hence, they cannot be represented in two dimensions. On the other hand, samples from a scaled backbone model ($\rho_{jk} = 1$) form 18 straight lines in ground-motion space, one for each seed model (Figure 8–37). In this context, because the NGA-East covariance is a compromise between the no-correlation and full-correlation distributions, the models reside on an intermediate high-dimensional shape. Hence, the plane from Figure 8–35 cuts the sphere ($\rho_{jk} = 0$) and contains

the straight lines ($\rho_{jk} = 1$). Recall that the Sammon's maps are created through a minimization of the Sammon's stress (Equation 8–13). We are not aware of rules or guidance on the acceptable level of Sammon's stress. However, we ran computations for a fully uncorrelated set (sphere in space), and the Sammon's stress was on the order of 0.8. If the cloud of points formed a perfect plane, the Sammon's stress would theoretically be zero. In the NGA-East case, the Sammon's stress ranges from 0.035–0.087 depending on the frequency. This range is much lower than in the no correlation case and much closer to the full correlation case, in which there would be no distortion at all.

Capitalizing on the fact that there is structure in the ground-motion space, the 10,000 sampled models per frequency are projected to two dimensions using Sammon's mapping. For the NGA-East project, the (**M**, R_{RUP}) scenarios considered for the Sammon's maps analyses are

- **M** = 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 7.8, and 8
- R_{RUP} = 0, 1, 5, 10, 20, 30, 40, 50, 70, 100, 130, 150, 200, 300, 400, 500, 600, 700, 800, 1000, 1200, and 1500 km

Hence, there are in total 220 (**M**, R_{RUP}) scenarios at which the seed GMMs are considered and from which the sampled models are drawn. This is a smaller subset than was considered for the generation of new models. This range comprises the bulk of the NGA-East scope of scenarios and the lower number allowed for faster computations.

The TI team formally tested the use of a smaller subset of hazard-relevant scenarios, defined as contributing to more than 1% of hazard at the 10^{-4} annual frequency of exceedance (this led to 160 scenarios). We also tested using different weights for those hazard-relevant scenarios relative to the remaining ones from the list of the 220. The results from these hazard sensitivity analyses, the corresponding achieved variance and the detailed interpretation are presented in Appendix E.5. The important conclusion from this set of analyses is that considering only the hazard-relevant scenarios effectively obliterated the variance for large-distance scenarios (beyond roughly 800 km), making it much smaller than what the seed GMMs predict. This was an unintended effect of not considering all the scenarios. The variance for hazard-relevant scenarios was also not as close to the target variance described in Section 8.1. The TI team felt that there was no reason to effectively reduce the epistemic uncertainty at large distances. Moreover, because the hazard was not really affected by those scenarios, it was better to use the 220 scenarios above for the generation of the Sammon's maps. Appendix E.5 provides more details.

In addition to the 10,000 sampled models, the seed GMMs are also used to calculate the Sammon's maps. Several reference models are also added to this set. Three models are used to map the mean of the seed GMMs and the mean up and down-scaled by a factor of two. These reference models align on a straight line on the map (see Figure 8–25) and are used to orient the maps in the same way across frequencies. In addition, four reference models for magnitude scaling and four reference models for distance scaling are added. The reference models for the magnitude scaling are calculated as mix + β(**M** − 6), with β =−0.4, −0.2, 0.2, and 0.4, where mix is the mean of all seed models. The reference models for the distance scaling are calculated as mix + γ (R − 400), with γ =−0.002, −0.001, 0.001, and 0.002.

The basic strategy for generating Sammon's maps that are comparable across frequencies is as follows:

1. For the set of selected models (sampled models, seed models, and reference models), select the ground-motion estimates at the defined relevant (**M**, R_{RUP}) scenarios.
2. Calculate GMM-distance or $\overline{\Delta_{GM}}$ between all models based on Equation (8–14).
3. Calculate Sammon's map, with PCA as the starting configuration.
4. Shift the resulting points in two dimensions such that the mean is at {0, 0}.
5. Rotate the resulting points such that the up/down-scaled mean models align horizontally (parallel to the x-axis); the rotation is carried out such that the up-scaled model has a larger x-value.
6. Flip the resulting coordinates along the x-axis such that the distance scaling reference model with γ =0.002 has a positive y-value.

The last three steps are taken to ensure that the maps are comparable for different frequencies. (The map only preserves GMM distances; therefore, its orientation has no meaning.)

Rotated/mirrored maps are shown in Figure 8–38 for four frequencies. The rotation and mirroring ensures that the up/down-scaled reference models are always at the same place for each frequency. Since models that are close in the high-dimensional space are close on the map, this means that sub-regions on the map can be traced across frequencies. This encourages the generation of smoother spectra. Additionally, Figure 8–38 shows that with increasing frequency, the distribution of samples in two dimensions becomes less like a circle and more like ellipses. The reason for this is seen in Figure 8–39, where the variances of the seed models with respect to magnitude and distance scaling are shown. With increasing frequency, the variance of the magnitude scaling decreases, whereas in comparison the variance of the distance scaling stays relatively constant. Because differences in scaling are manifested in the y-direction on the map, the lower variance in magnitude scaling in the seed models for high frequencies leads to a lower variance in the sampled models and thus a lower variance in the y-direction on the map.

The up/down-scaled models in the rotated maps (see Figure 8–38) have x-coordinates that are close to ln(2)=0.693, which is the factor by which they are scaled. Hence, the maps can adequately capture the ground-motion distances in high-dimensional ground-motion space in two dimensions.

Overall, the sampled set of GMMs captures the seed models on the maps and covers a continuous region in two dimensions. Since the 10,000 sampled GMMs are representative of the continuous distribution P(**Y**), and the Sammon's map is a reasonable projection of the 10,000 sampled GMMs, the TI team concluded that the map is a reasonable representation of the continuous distribution of P(**Y**).

8.5 Discretization of Ground-Motion Space

8.5.1 Definition of Range in Ground-Motion Space

In the previous section, the visualization of the sampled models from the continuous ground-motion distribution P(**Y**) was presented. This results in Sammon's maps, one for each frequency, which are a representation of the high-dimensional ground-motion model space—each map is an approximation of P(**Y**). Since P(Y) describes epistemic uncertainty, and the map is an approximation of P(Y), then the center, body, and range (CBR) on the map is an approximation to CBR of P(Y). Thus, defining the CBR of the models in the two-dimensional space (on the map) is an approximation to the center, body, and range of P(**Y**). The definition of the range is done similar to a one-dimensional distribution; in that case, often a range of $\pm 2\sigma$ is chosen (in the case of a normal distribution). Such a range covers 95.45% of the total probability of a normal distribution:

$$\int_{-2}^{2} N(x; \mu = 0, \sigma = 1)\, dx = 0.9545 \qquad (8\text{–}18)$$

However, the Sammon's maps represent a two-dimensional distribution. Based on the definition of a two-dimensional normal distribution and the distribution of the sampled models (gray points in Figure 8–38), an ellipse was selected to represent the range on the map. The half-axes of the ellipses are determined by the standard deviations of the distribution of points in x- and y-direction. These are calculated and then scaled by a factor α such that the resulting ellipse covers 95% of the total probability of a two-dimensional normal distribution. The factor α is calculated in the following way: The probability density function of the two-dimensional normal distribution is converted to polar coordinates r, θ, and the angle θ is marginalized out. This results in a Rayleigh distribution, and the scaling factor α can be calculated from

$$p = F(s) = 1 - e^{\left(-s^2/\sigma^2\right)} \qquad (8\text{–}19)$$

where F is the cumulative distribution of the Rayleigh distribution, and p is the amount of probability that should be enclosed by the ellipse (p = 0.95). The resulting scale factor is α = 2.45. This is equivalent to use a second order chi-squared distribution at the 0.95 bilateral level.

The ellipse is centered on the point {0,0}, which corresponds to the mean of the seed GMMs (see Section 8.3). Figure 8–40 shows the ellipse defining the range for two frequencies. This range covers large portion of the map, and thus also in the ground-motion model space.

8.5.2 Discretization of the Ground-Motion Space into Cells

With the range as defined in the previous section, the ellipse encloses the subset of ground-motion model space (Figure 8–40) that the TI team intends to capture. As described above, this range covers 95% of the total probability on the map. The range needs to be discretized into a manageable number of GMMs. Therefore, the ellipse defining the range is partitioned into several cells, and a representative model for each cell is developed.

The space inside the ellipse is further partitioned via two ellipses (forming rings), with different (smaller) scale factors. Thus, the considered range is partitioned into a central ellipse and two outer rings. These rings are further partitioned into eight cells, based on equal angular distances

(45°), to capture epistemic uncertainty in magnitude and distance scaling. Thus, in total there are 17 cells-one central cell (ellipse), and eight cells each on the central and outer ring.

The scale factors to calculate the semi-axes of the inner ellipses are again based on a two-dimensional Gaussian distribution. The center cell represents the center, the middle ring the body, and the outer ring the range of the distribution. Based on a two-dimensional normal distribution, the center should correspond to 10% of the density, the body should capture 75% (including the center), and the full range should capture 95% of the distribution, as stated in the previous section. Thus, the body corresponds to 65% (75–10), and the range corresponds to 20%. Hence, the scale factors for the inner ellipses are calculated such that the cumulative distribution function of a two-dimensional normal distribution equals 0.1 and 0.75. The scale factors are calculated from the cumulative distribution function of the Rayleigh distribution, according to Equation (8-19). The resulting scale factors are α = 0.46 and 1.65 for p = 0.1 and 0.75 respectively.

The TI team tested a different discretization, using the same outer ellipse and three rings to define a total of 13 and 29 models. The 29 models scheme was the original approach presented at the SSHAC workshops. The TI team concluded that there is not enough information to defend a more complex discretization, and that the 17 models scheme was appropriate (and practical) to capture a range of ground-motion values. The number of models is sufficient to capture alternate magnitude and distance scaling behaviors represented in the space away from the center. Sensitivity results to the different discretizations of the space are presented in Appendix E.6.

An example of the discretized ground-motion space is shown in Figure 8–41, for two frequencies. Figure 8–42 shows the fraction of areas of the different cells with respect to the cell in the center.

8.5.3 Selection of Representative GMM for Each Cell

Each of the 17 cells defined above covers a fraction of the area on the map, which is the full representation of P(**Y**). The next step is to define a representative GMM for each cell. Various candidate representative models were considered by the TI team, with preference given to an approximation to the expectation of **Y** over each cell. Since P(**Y**) is a distribution over vectors of ground-motion estimates (an approximation of a continuous GMM), this results in a valid GMM. The approximation to the expectation is calculated by averaging over all models inside one cell:

$$y_k = \frac{1}{N_k} \sum_{i=1}^{N_k} y_i \qquad (8\text{–}20)$$

where k indexes the cell, and N_k is the number of samples inside a cell. This is an approximation of:

$$E[Y]_{A_k} = \int_{A_k} yP(y)dA \qquad (8\text{–}21)$$

The representative model for each cell is therefore an average of samples from P(**Y**), which allows the representation of the systematic trends in that cell. For example, if the representative

model was randomly selected from any model in the cell, it could lead to extreme model realizations and larger variations from frequency to frequency (and very jagged spectra). Because the maps are all oriented the same way, the average metric has the advantage of producing smoother spectra.

All the samples passed the criteria of physicality established by the TI team, ensuring that the selected models also pass the physicality constraints. Figure 8–43 illustrates the scaling of the 17 selected models against distance and magnitude.

The spectra of the selected models are shown in Figure 8–44, for a single (**M**, R_{RUP}) values. For each frequency, the models with the same model index (see right-hand panel of Figure 8–42) are combined into one GMM. Considering that the process is performed for each frequency independently, the spectra are reasonably smooth. The smoothness is ensured by the fact that the Sammon's maps, on which the selected models are based, are all rotated and flipped in the same way. Figure 8–45 shows an example 3D plot for two different frequencies. The three axes correspond to ground motions from three different (**M**, R_{RUP}) scenarios. The seeds and samples align into plane-like cloud structures, one for each frequency. These two planes cover different ground-motion values, but the location of the seeds and samples remain in similar positions relative to each plane. This is especially true for close-by frequencies and can explain the relative smoothness of the spectra. Once the high-dimensional space is mapped in two dimensions, the structure is preserved across frequencies by the rotation and reflection of the Sammon's maps in a consistent way. Hence, the regions on the Sammon's maps for the different frequencies correspond roughly to the same scaling properties (see Figure 8–25). However, smoother individual models may be desired for specific applications. The smoothing of the 17 models is described in the next section.

8.6 Final Models and Smoothing Process

Figure 8–44 shows that although the spectra are relatively smooth, they may be too jagged for certain applications. This is especially true for models near the tail ends of the distributions (the outer cells in ellipse), which are averaged over a smaller number of models in each cell. Therefore, the TI team decided to smooth the 17 selected models to ensure a reasonable expected shape of all spectra for all (**M**, R_{RUP}) scenarios.

We want to fit the spectrum and PGA, which is anchored at a frequency dependent on distance based on simulations by David Boore (see Appendix E.7). The anchoring frequencies for PGA are shown in Figure 8–46.

The smoothing is done by fitting each of the 17 models to a function that depends on frequency, magnitude and distance.

$$Y_k = g_k(f, \mathbf{M}, R_{RUP}) \tag{8–22}$$

where Y is the logarithmic PSA value at one of the 24 NGA East frequencies (f = 0.1 to 100 Hz plus PGA as described above), and k indexes the 17 models. The function g(f, **M**, R_{RUP}) has the following form (modified from McGuire et al. (2001):

$$g(f, \mathbf{M}, R_{RUP}) = a_0 + \frac{a_1}{\cosh[\exp(a_2) f^{a_3}]} + \frac{\exp(a_5 f)}{f^{a_6}} \tag{8–23}$$

where some of the coefficients a* depend on magnitude and distance. The dependence of the coefficients on **M** and R_{RUP} is as follows

$$
\begin{aligned}
a_4 &= c_{40} + \mathbf{M}\left[c_{41} + c_{42}\mathbf{M} + c_{43}\ln\left(c_{44}R_{RUP} + 1\right)\right] \\
a_6 &= c_{60} + c_{61}R_{RUP} \\
a_7 &= \exp\left[c_{70} + \mathbf{M}\left(c_{71} + c_{72}\mathbf{M} + c_{73}R_{RUP}\right)\right] \\
a_8 &= c_{80} + \mathbf{M}\left(c_{81} + c_{82}R_{RUP}\right) \\
a_9 &= c_{90} + c_{91}\mathbf{M}
\end{aligned}
\qquad (8\text{–}24)
$$

where the coefficients c* are estimated. The other coefficients (a_1, a_2, a_3, and a_5) are held constant across all magnitudes and distances. The coefficient a_0 corresponds to the PGA value. It is different for each of the 374 **M**-R_{RUP} scenarios.

All coefficients (a_x ,c_x) are estimated via Bayesian inference using the program Stan (Carpenter et al. 2017; Stan Development Team 2017). The prior distributions for the coefficients θ (save a_0) are set to be weakly informative--they are normal distributions with mean zero and standard deviation 10, or $\theta \sim N(0,10)$.

The prior distribution for a_0 is a normal distribution whose mean is the PGA value of the unsmoothed model, with a standard deviation of 0.1. This ensures that the estimated coefficient a_0 is similar to the unsmoothed PGA value.

The parameters are estimated by maximum-a-posteriori (MAP) optimization. Since the inference is sensitive to the starting values, four different starting values are used, and the final model with the highest log-probability is used.

Figure 8–47 shows the fit of one particular spectrum, and Figure 8–48 shows the smoothed version of the spectra from Figure 8–44. Figure 8–49 shows comparisons of the hazard curve distribution for f = 1 Hz, calculated the smoothed and unsmoothed models. Figure 8–50 shows a sample uniform hazard spectrum (UHS) calculated for both the smoothed and unsmoothed models. Figures 8–49 and 8–50 were produced using the final weights defined in Chapter 9. Although there are some differences at some frequencies, there is no defensible argument to maintain the “kinks,” which are relatively small. The benefit of smooth spectra was deemed to outweigh benefits of preserving the actual spectral shapes. Plotting tools are also provided in Appendix E as electronic appendices, which allow users to generate plots of magnitude scaling, distance scaling and spectral shape through an interactive interface.

8.7 References

Abrahamson, N.A., W.J. Silva, and R. Kamai. 2014. “Summary of the AKS14 Ground-Motion Relation for Active Crustal Regions.” Earthq. Spectra, 30(3): 1025–1055, DOI 10.1193/070913EQS198M.

Akkar, S., M.A. Sandikkaya, and J.J. Bommer. 2014a. “Empirical Ground-Motion Models for Point- and Extended-Source Crustal Earthquake Scenarios in Europe and the Middle East.” Bull. Earthq. Eng. 12(1): 359–387, DOI 10.1007/s10518-013-9461-4.

Akkar, S., M.A. Sandikkaya, and J.J. Bommer. 2014b. “Erratum to: Empirical Ground-Motion Models for Point- and Extended-Source Crustal Earthquake Scenarios in Europe and the Middle East.” Bull. Earthq. Eng. 12(1): 389–390, DOI 10.1007/s10518-013-9508-6.

Al Atik, L., and R.R. Youngs. 2014. “Epistemic Uncertainty for NGA-West2 Models.” Earthq. Spectra, 30(3): 1301–1318.

Atkinson, G.M., J.J. Bommer, and N.A. Abrahamson. 2014. “Alternative Approaches to Modeling Epistemic Uncertainty in Ground Motions in Probabilistic Seismic-Hazard Analysis.” Seismol. Res. Lett., 85(6): 1141–1144.

Boore, D.M., J.P. Stewart, E. Seyhan, and G.M. Atkinson. 2014. “NGA-West 2 Equations for Predicting PGA, PGV, and 5%-Damped PSA for Shallow Crustal Earthquakes.” Earthq. Spectra, 30(3): 1057–1085. DOI: 10.1193/070113EQS184M.

Bozorgnia, Y., N.A. Abrahamson, L. Al Atik, T.D. Ancheta, G.M. Atkinson, J. Baker, A. Baltay, D.M. Boore, K.W. Campbell, B.-S.J. Chiou, R.B. Darragh, S. Day, J.L. Donahue, R.W. Graves, N. Gregor, T. Hanks, I.M. Idriss, R. Kamai, T. Kishida, A. Kottke, S.A. Mahin, S. Rezaeian, B. Rowshandel, E. Seyhan, S. Shahi, R. Shantz, W.J. Silva, P. Spudich, J.P. Stewart, J. Watson-Lamprey, K.E. Wooddell, R.R. Youngs. 2014. “NGA-West2 Research Project.” Earthq. Spectra, 30(3): 973–987.

Campbell, K.W., and Y. Bozorgnia. 2014. “NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5%-Damped Linear Acceleration Response Spectra.” Earthq. Spectra, 30(3): 1087–1115. DOI: 10.1193/062913EQS175M

Carpenter, R., A. Gelman, M.D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A Probabilistic Programming Language. J. Statistical Software, 76(1). DOI 10.18637/jss.v076.i01.

Chiou, B.S-J., and R.R. Youngs. 2014. “Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra.” Earthq. Spectra, 30(3): 1117–1153. DOI: 10.1193/072813EQS219M.

GeoPentech. 2015. “Southwestern United States Ground Motion Characterization SSHAC Level 3 – Technical Report Rev.2.” March 2015.

Geusebroek, J.M., G.J. Burghouts, and A.W.M. Smeulders. 2005. “The Amsterdam Library of Object Images.” Int. J. Comput. Vision, 61(1): 103–112.

Gianniotis, N., and C. Riggelsen. 2013. “Visualisation of High-Dimensional Data using an Ensemble of Neural Networks.” Proceedings, IEEE Symposium on Computational Intelligence and Data Mining, SSCI 2013, Singapore.

Hotelling, H. 1933. “Analysis of a Complex of Statistical Variables into Principal Components.” J. Educ. Psychol., 24(6): 417–441.

Idriss, I.M. 2014. An NGA-West2 Empirical Model for Estimating the Horizontal Spectral Values Generated by Shallow Crustal Earthquakes, Earthq. Spectra, 30(3): 1155–1177. DOI: 10.1193/070613EQS195M

Kohonen, T. 2001. Self-Organizing Maps. Springer Verlag, New York.

Lawrence, N.D. 2004. “Gaussian Process Latent Variable Models for Visualization of High Dimensional Data.” Adv. Neur. In., 16: 329–336.

McGuire, R.K., W.J. Silva, and C.J. Costantino. 2001. “Technical Basis for Revision of Regulatory Guidance on Design Ground Motions: Hazard-a Risk-Consistent Ground Motion Spectra Guidelines.” NUREG/CR-6728, U.S. Nuclear Regulatory Commission, Washington, D.C.

Rasmussen, C.E., and C.K.I. Williams. 2006. Gaussian Processes for Machine Learning. MIT Press, Cambridge, MA.

Sammon, J.W. 1969. “A Nonlinear Mapping for Data Structure Analysis.” IEEE Trans. Computers, C-18: 401–409.

Scherbaum, F., N.M. Kuehn, M. Ohrnberger, and K. Koehler. 2010. “Exploring the Proximity of Ground-Motion Models using High-Dimensional Visualization Techniques.” Earthq. Spectra, 26(4): 1117–1138.

Stan Development Team. 2017. Stan Modeling Language Users Guide and Reference Manual, Version 2.16.0, http://mc-stan.org.

Zhao, J.X., and M. Lu. 2011. Magnitude-Scaling Rate in Ground-Motion Prediction Equations for Response Spectra from Large, Shallow Crustal Earthquakes, Bull. Seism. Soc. Am., 101: 2643–2661.

Zhao, J.X., J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H.K. Thio, P.G. Somerville, and Y. Fukushima. 2006. Attenuation Relations of Strong Ground Motion in Japan using Site Classification based on Predominate Period. Bull. Seism. Soc. Am., 96: 898–913.

Table 8–1 Values of estimated parameters θ and β^2 for f = 1 Hz.

Parameter	Value
θ_1	4.76
θ_2	8.64
θ_3	4.35
θ_4	0.0421
θ_5	2.93
θ_6	0.288
β^2	3.25 E -06

Figure 8–1 Diagonal entries of the NGA-East seed GMMs covariance matrix plotted against M and R_{RUP} for a suite of frequencies.

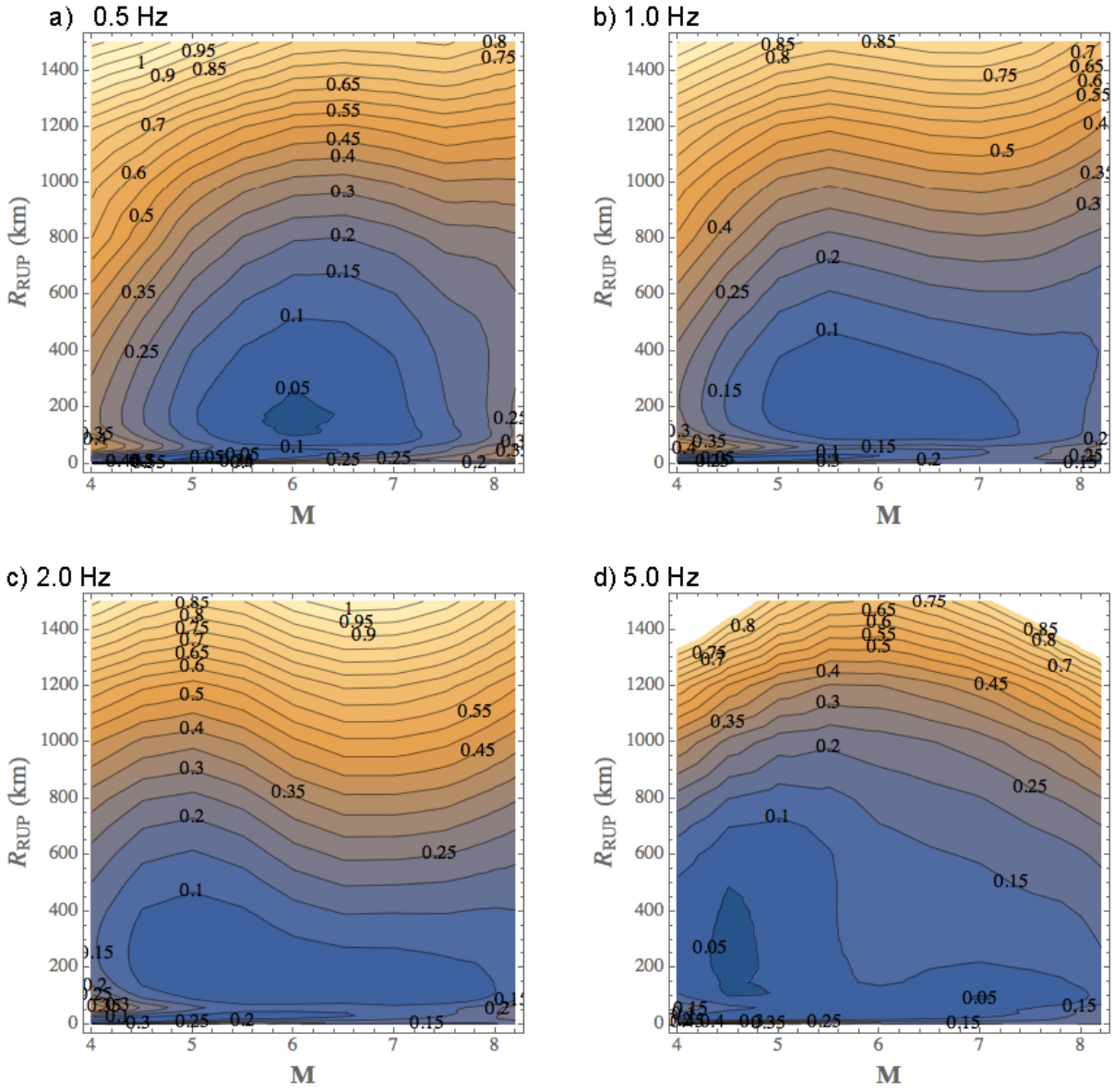


Figure 8–1 (continued) Diagonal entries of the NGA-East seed GMMs covariance matrix plotted against M and R_{RUP} for a suite of frequencies.

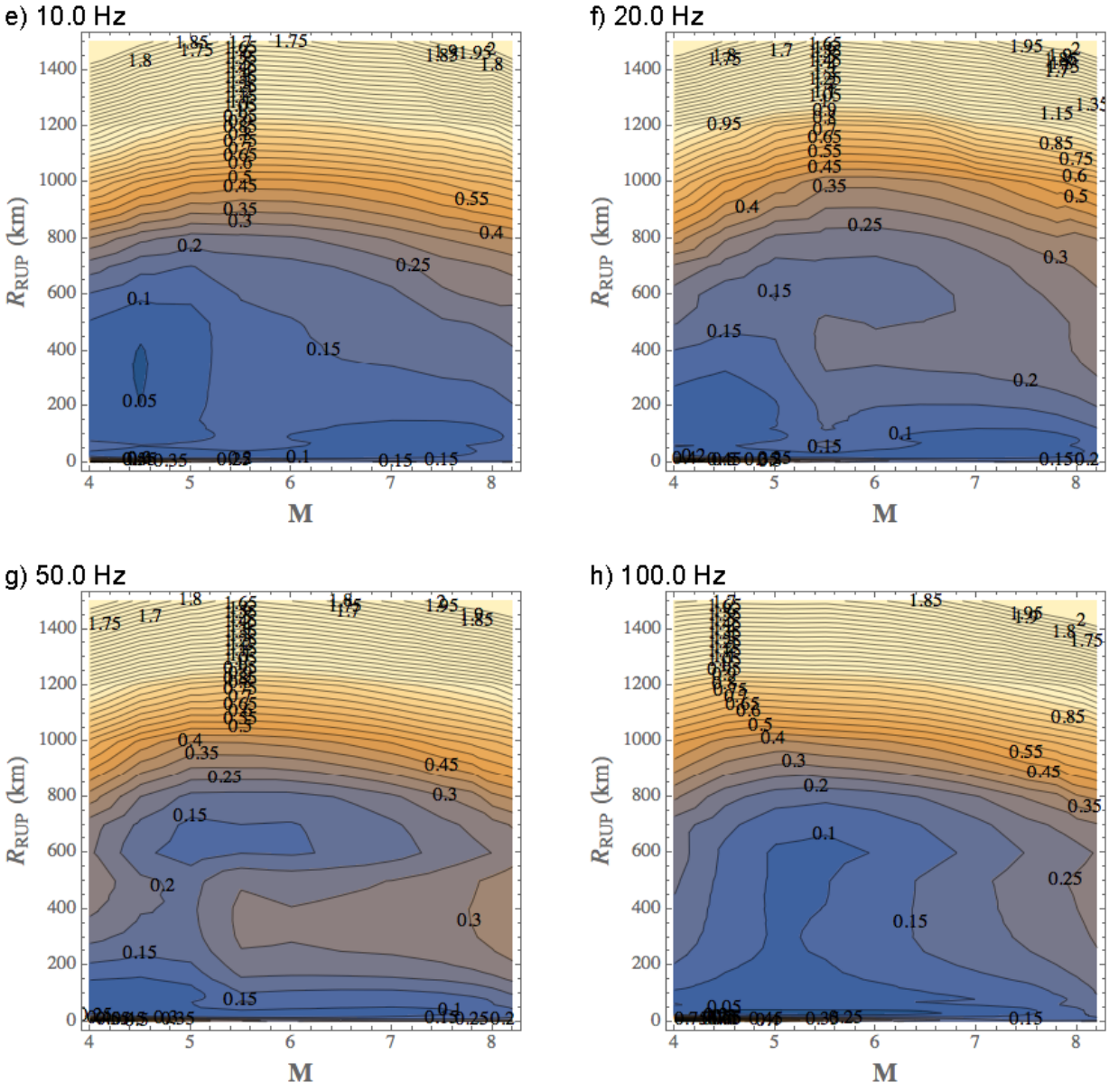


Figure 8–2 Summary of ground-motion data (recordings) available for GMM development, as provided by the NGA-East database (up to 400 km to prevent biased data, as discussed in Section 7.6).

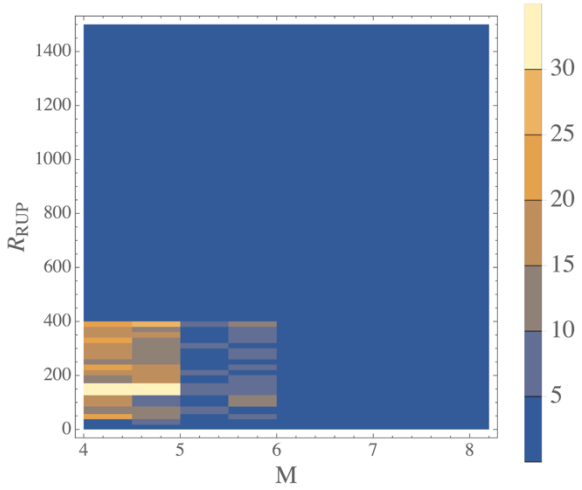


Figure 8–3 Sample magnitude scaling of NGA-East seed GMMs (1 Hz, 200 km): spread of models narrowest around M = 6 and smaller for M = 8 (no data) than M = 4 (see Figure 8–1 left for actual variance values at 200 km).

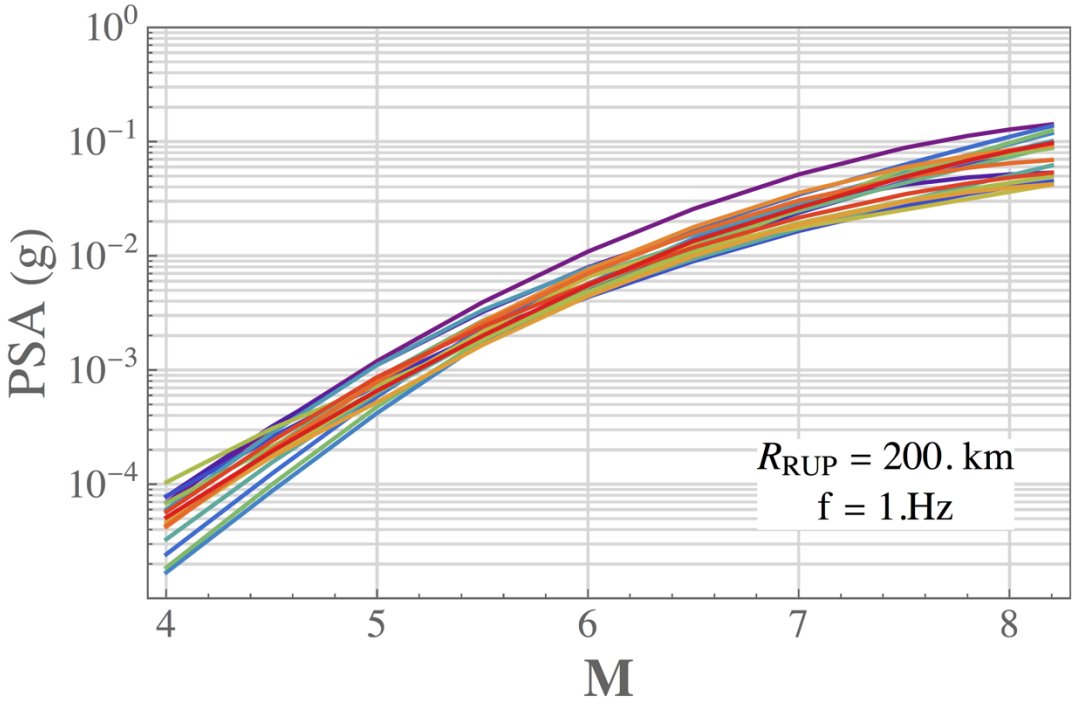


Figure 8–4 Variance from SWUS GMMs for a suite of frequencies.

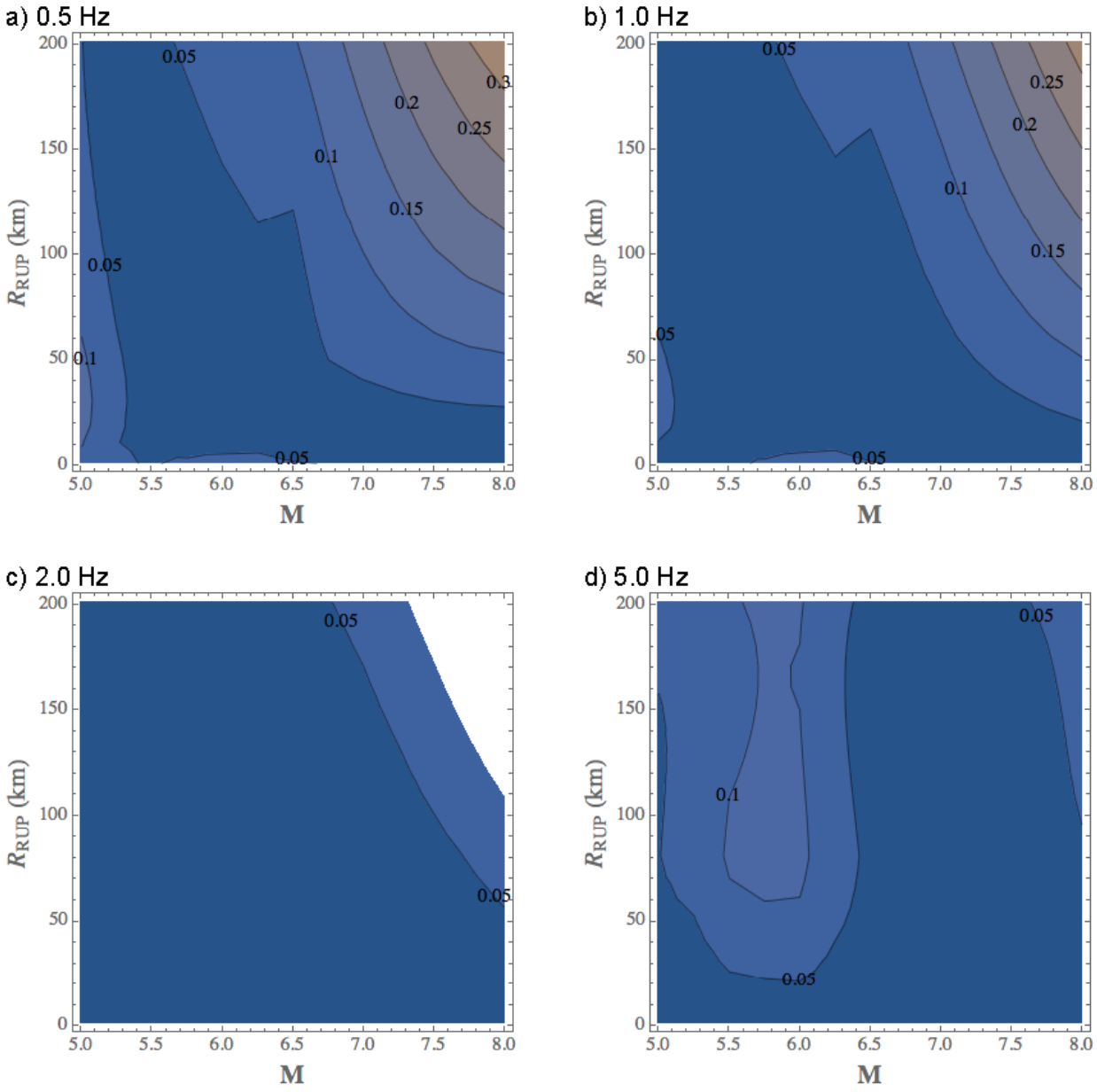


Figure 8–4 (continued) Variance from SWUS GMMs for a suite of frequencies.

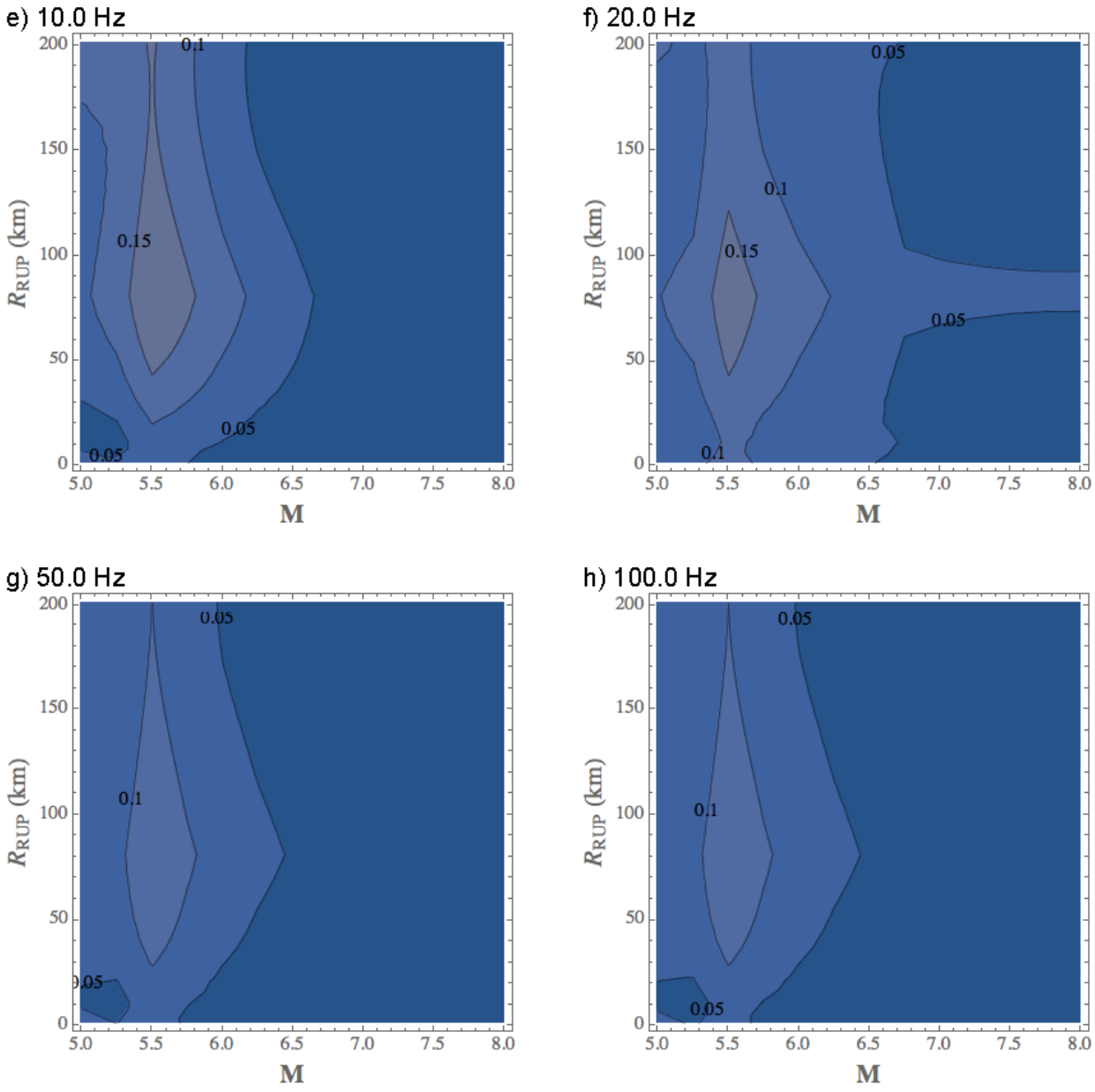


Figure 8–5 **Summary of ground-motion data (recordings) available for GMM development, as provided by the NGA-West2 database.**

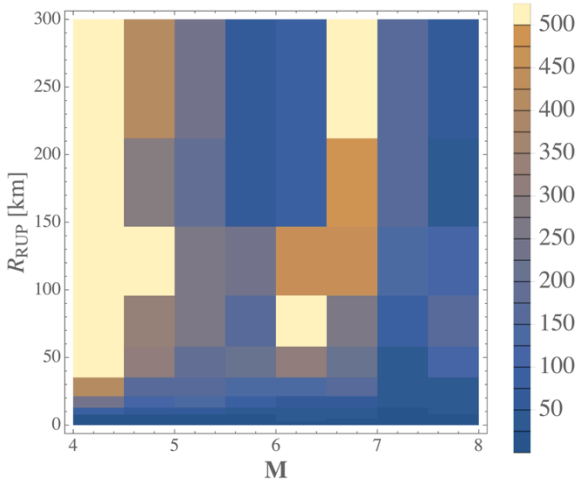


Figure 8–6 **Summary of ground-motion data (recordings) used for residual analyses for DCPP in the SWUS project.**

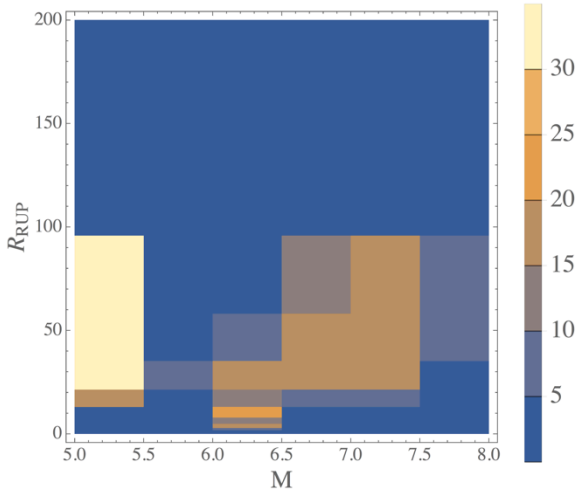


Figure 8–7 Sample magnitude scaling of NGA-West2 GMMs (10 Hz, 25 km): spread of models smaller for M = 8 than M = 5 (see Figure 8–1 left for actual variance values at 200 km).

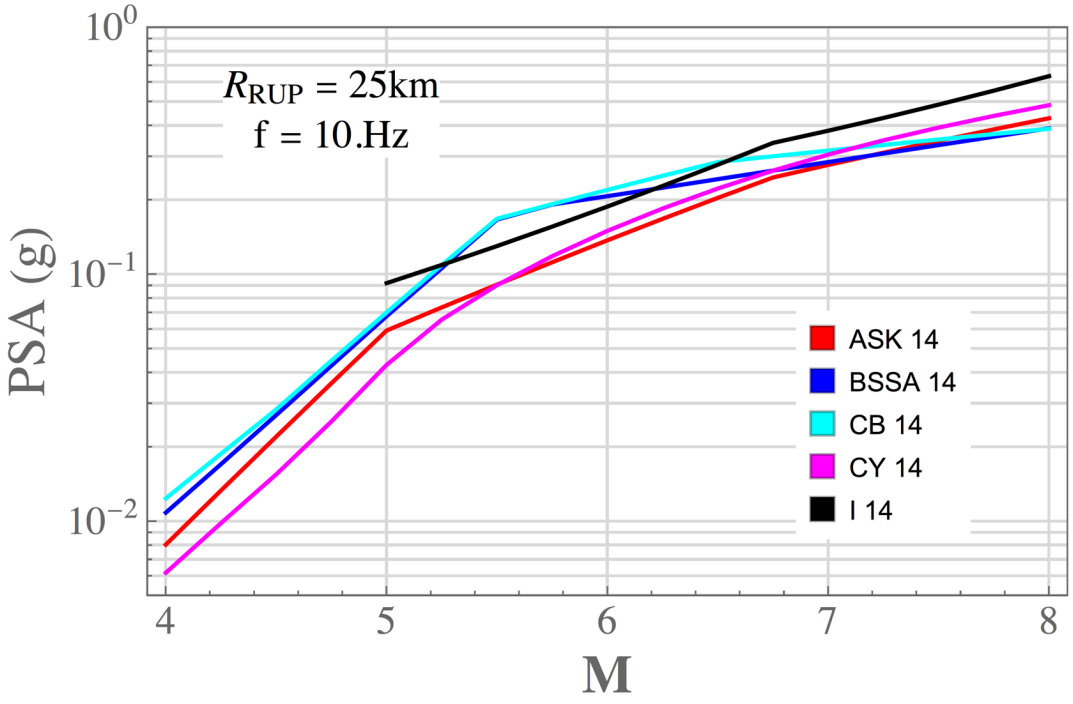


Figure 8–8 Variance from SWUS GMMs when the epistemic uncertainty model from Al Atik and Youngs is included for a suite of frequencies.

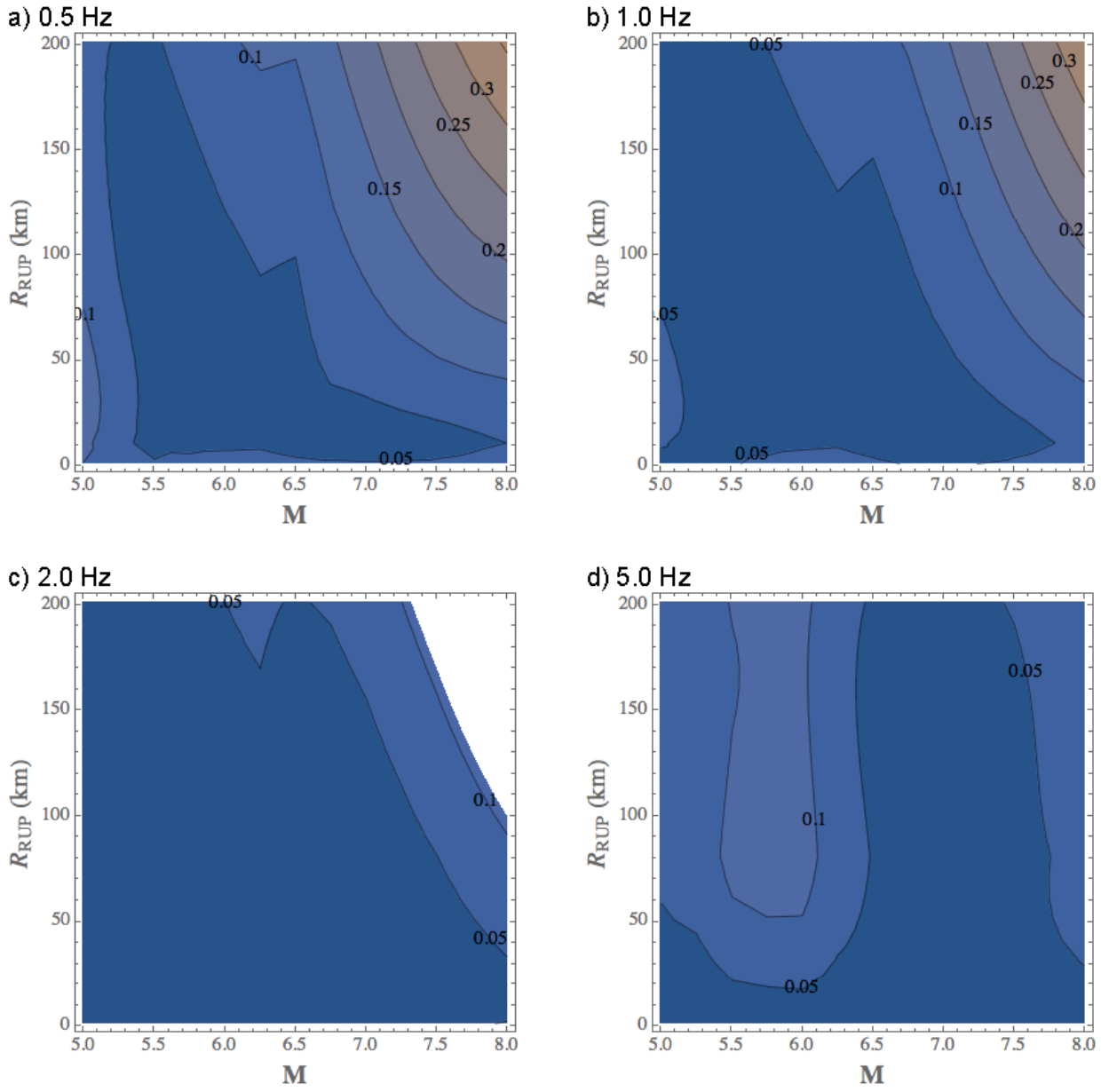


Figure 8–8 (continued) Variance from SWUS GMMs when the epistemic uncertainty model from Al Atik and Youngs is included for a suite of frequencies.

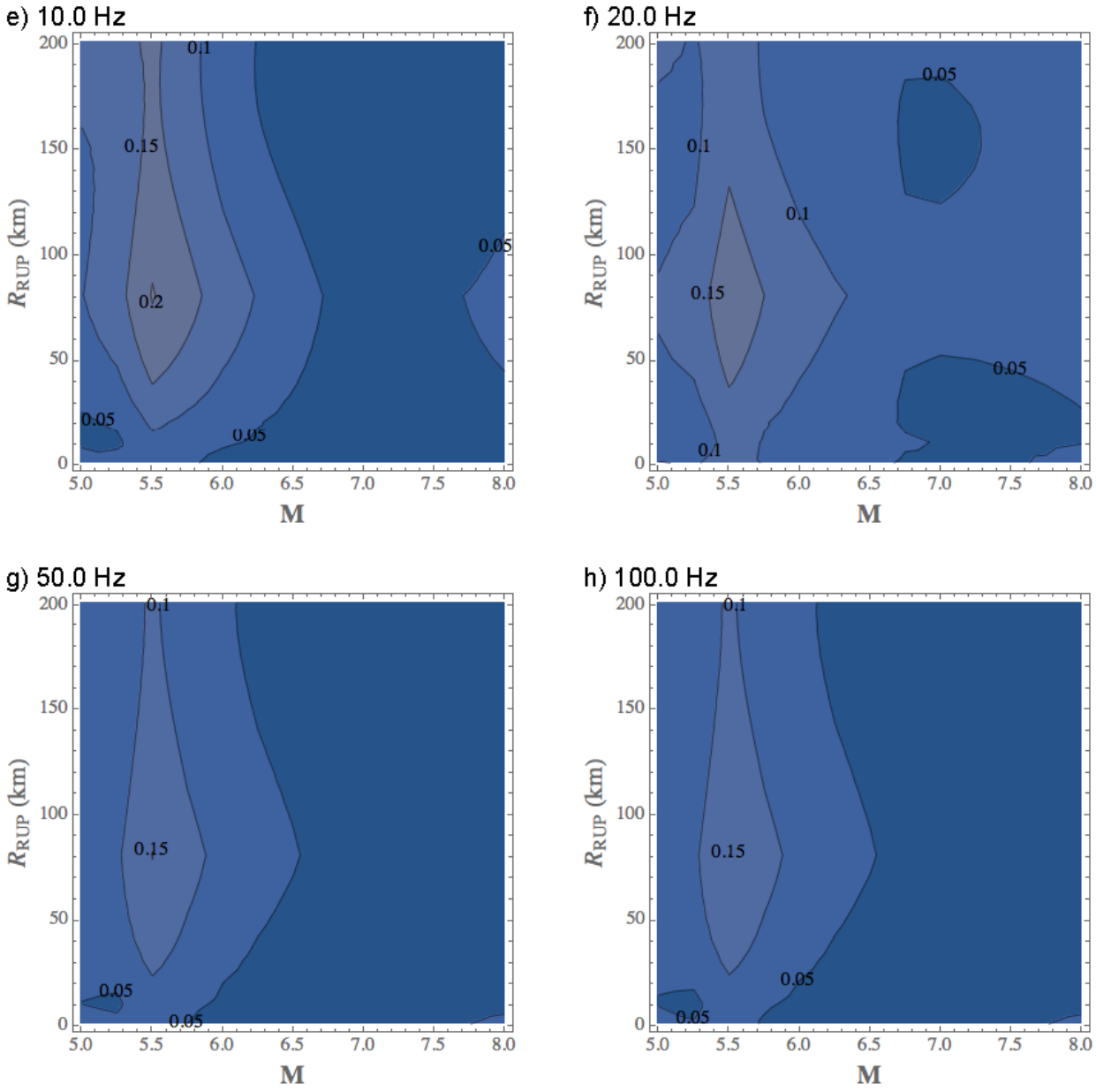


Figure 8–9 Variance from SWUS DCPP final models for a suite of frequencies.

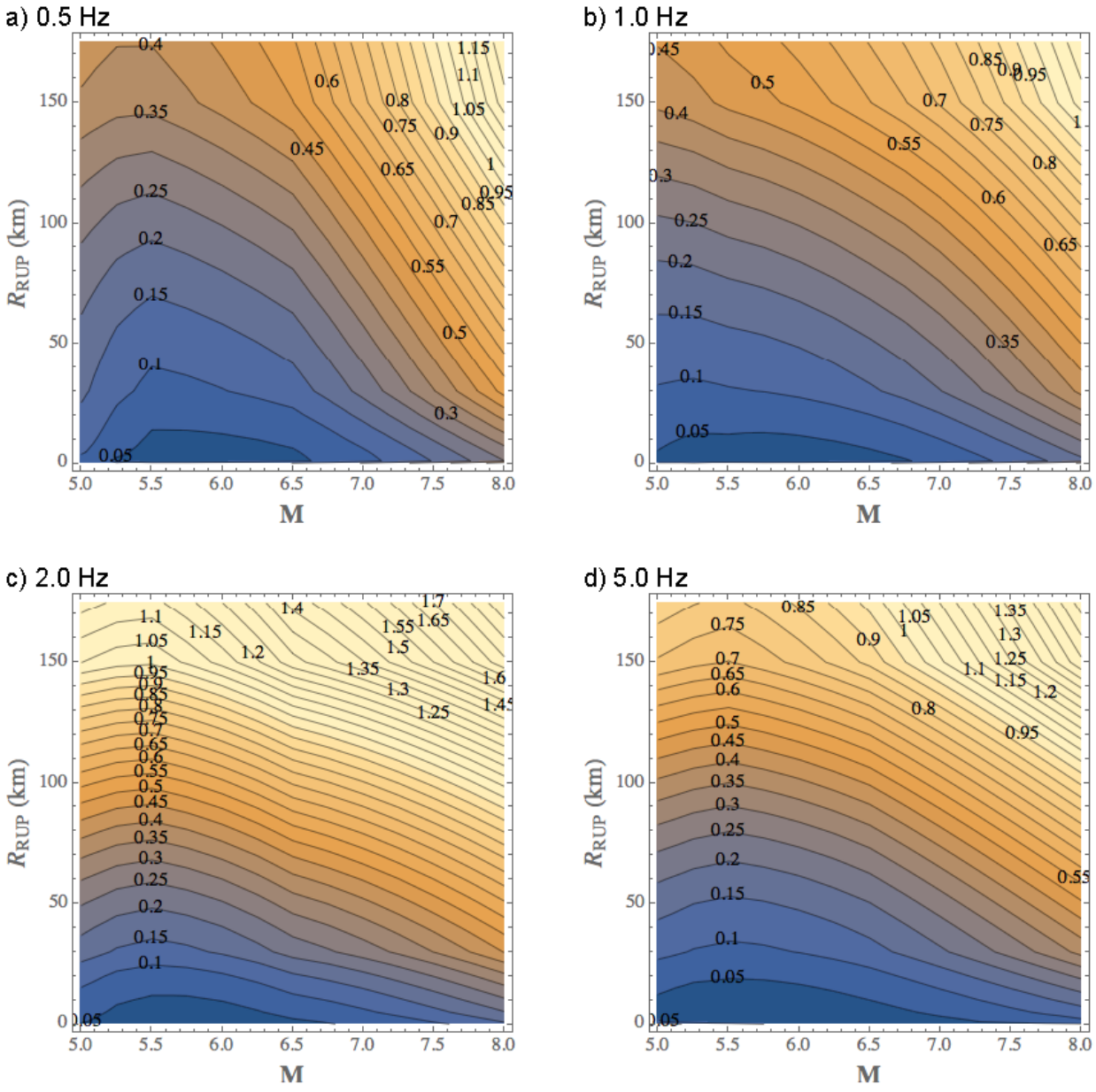


Figure 8–9 (continued) Variance from SWUS DCPP final models for a suite of frequencies.

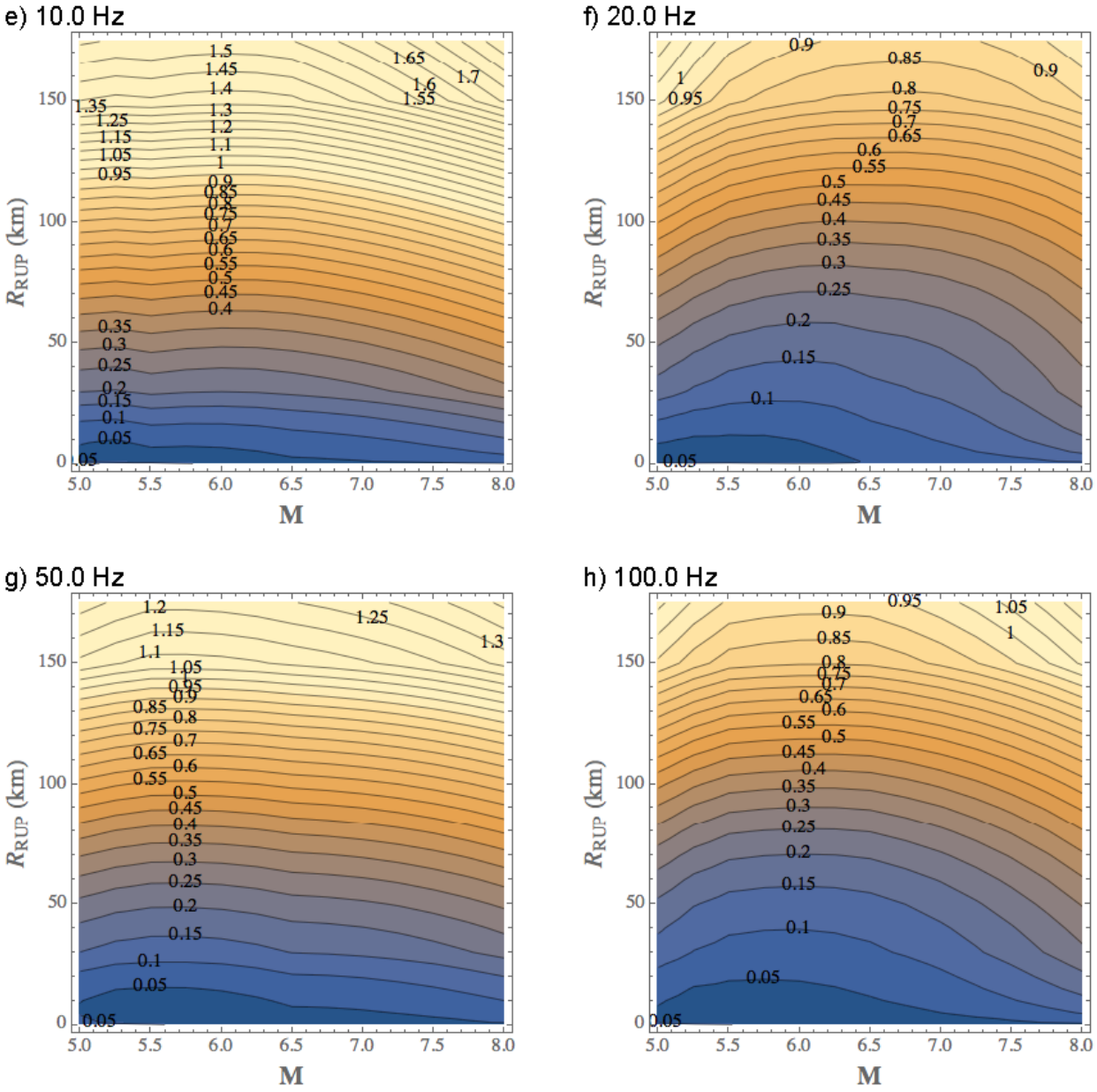


Figure 8–10 Variance from SWUS DCPP final models across all frequencies for M = 8 and R_{RUP} = 50 km (top) and R_{RUP} = 70 km (bottom).

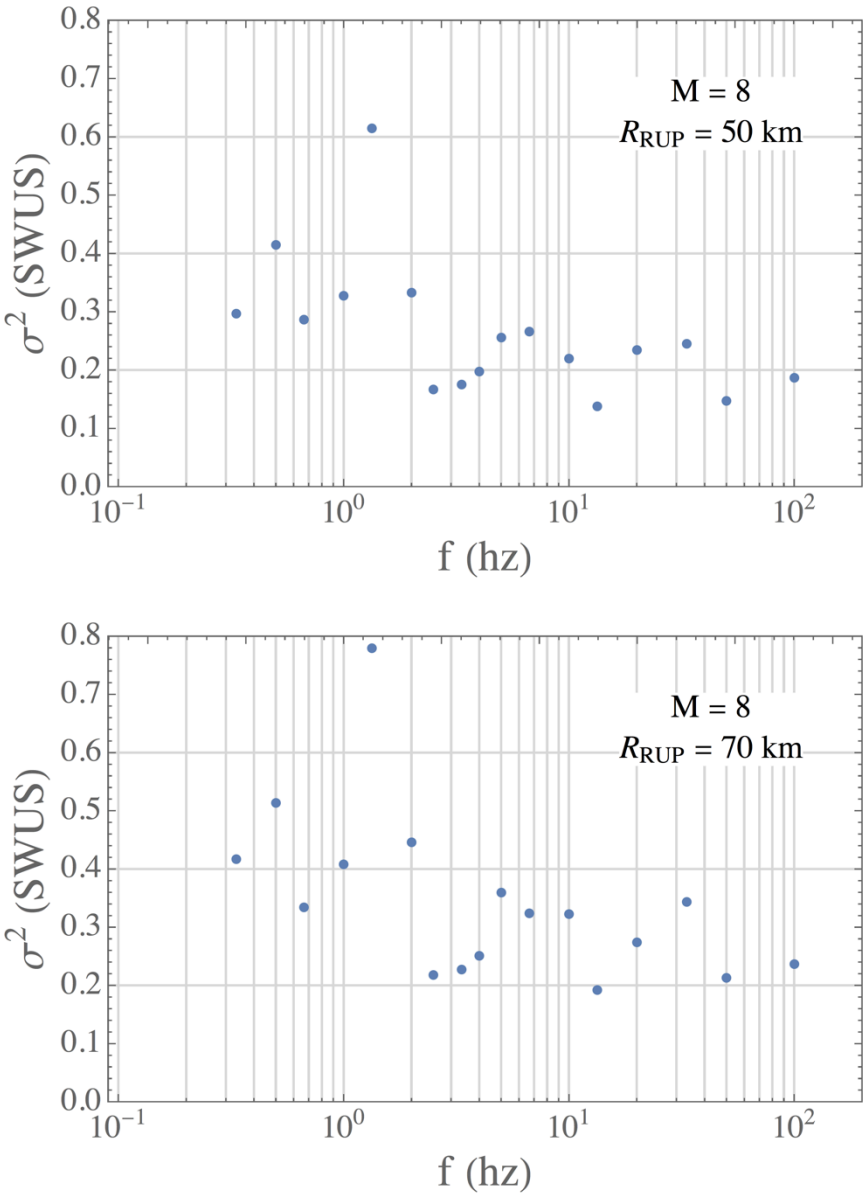


Figure 8–11 CENA target variance model. Values in blue are imposed boundaries developed using the SWUS DCPP final model as guidance and values in white were interpolated to achieve the smooth variance shown in Figure 8–12.

R_{RUP} (km)\ **M**	M4	M4.5	M5	M5.5	M6	M6.5	M7	M7.5	M7.8	M8.2
0	0.2030	0.2030	0.2030	0.2030	0.2030	0.2030	0.2030	0.15	0.15	0.15
1	0.1718	0.1718	0.1718	0.1718	0.1718	0.1718	0.1718	0.15	0.15	0.15
5	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
10	0.1406	0.1406	0.1406	0.1406	0.1406	0.1406	0.1794	0.2187	0.1815	0.15
15	0.1351	0.1351	0.1351	0.1351	0.1351	0.1450	0.1966	0.2483	0.2570	0.20
20	0.1312	0.1312	0.1312	0.1312	0.1312	0.1572	0.2088	0.2605	0.2915	0.25
25	0.1282	0.1282	0.1282	0.1282	0.1282	0.1667	0.2183	0.2700	0.3009	0.30
30	0.1257	0.1257	0.1257	0.1257	0.1257	0.1744	0.2261	0.2777	0.3087	0.35
40	0.1218	0.1218	0.1218	0.1218	0.1350	0.1866	0.2383	0.2899	0.3275	0.40
50	0.1188	0.1188	0.1188	0.1188	0.1444	0.1961	0.2477	0.3120	0.3625	0.40
60	0.1163	0.1163	0.1163	0.1163	0.1522	0.2038	0.2555	0.3344	0.3625	0.40
70	0.1142	0.1142	0.1142	0.1142	0.1587	0.2104	0.2799	0.3344	0.3625	0.40
80	0.1124	0.1124	0.1124	0.1128	0.1644	0.2160	0.2875	0.3344	0.3625	0.40
90	0.1108	0.1108	0.1108	0.1178	0.1694	0.2330	0.2875	0.3344	0.3625	0.40
100	0.1094	0.1094	0.1094	0.1222	0.1739	0.2406	0.2875	0.3344	0.3625	0.40
110	0.1081	0.1081	0.1081	0.1263	0.1779	0.2406	0.2875	0.3344	0.3625	0.40
120	0.1069	0.1069	0.1069	0.1300	0.1924	0.2406	0.2875	0.3344	0.3625	0.40
130	0.1058	0.1058	0.1058	0.1334	0.1938	0.2406	0.2875	0.3344	0.3625	0.40
140	0.1048	0.1048	0.1048	0.1365	0.1938	0.2406	0.2875	0.3344	0.3625	0.40
150	0.1039	0.1039	0.1039	0.1406	0.1938	0.2406	0.2875	0.3344	0.3625	0.40
175	0.1018	0.1018	0.1018	0.1469	0.1938	0.2406	0.2875	0.3344	0.3625	0.40
200	0.10	0.10	0.10	0.1469	0.1938	0.2406	0.2875	0.3344	0.3625	0.40
250	0.10	0.10	0.10	0.1469	0.1938	0.2406	0.2875	0.3344	0.3625	0.40
300	0.10	0.10	0.10	0.1654	0.2123	0.2591	0.3060	0.3529	0.3810	0.40
350	0.10	0.10	0.10	0.1811	0.2279	0.2748	0.3217	0.3686	0.3967	0.40
400	0.10	0.10	0.10	0.1946	0.2415	0.2884	0.3353	0.3821	0.4000	0.40
450	0.13	0.13	0.13	0.2000	0.2535	0.3003	0.3472	0.3941	0.4000	0.40
500	0.16	0.16	0.16	0.2000	0.2642	0.3111	0.3579	0.4000	0.4000	0.40
600	0.2107	0.2107	0.2107	0.2312	0.2827	0.3296	0.3764	0.4000	0.4000	0.40
700	0.2528	0.2528	0.2528	0.2528	0.2984	0.3452	0.3921	0.4000	0.4000	0.40
800	0.2893	0.2893	0.2893	0.2893	0.3013	0.3588	0.4000	0.4000	0.4000	0.40
1000	0.3502	0.3502	0.3502	0.3502	0.3556	0.3815	0.4000	0.4000	0.4000	0.40
1200	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
1500	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40

Figure 8–12 NGA-East target variance model, plotted against M and R_{RUP}.

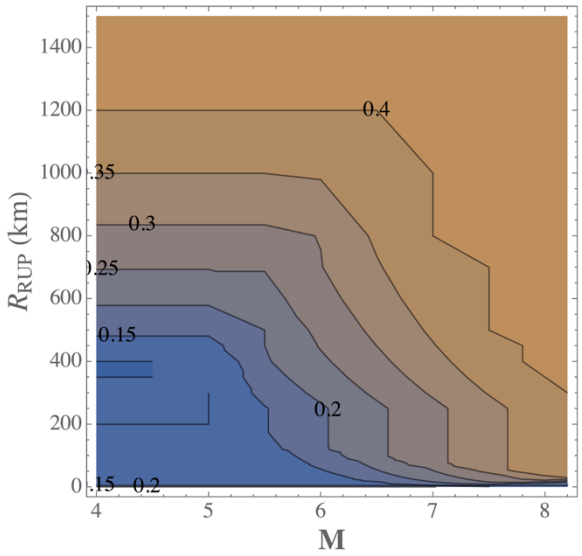


Figure 8–13 Sample correlation coefficients for f = 1 Hz, plotted against M and $\log_{10}R_{RUP}$. Left; M = 4.5, R_{RUP} = 10 km, Center: M = 6, R_{RUP} = 100 km, Right: M = 7.5, R_{RUP} = 400 km.

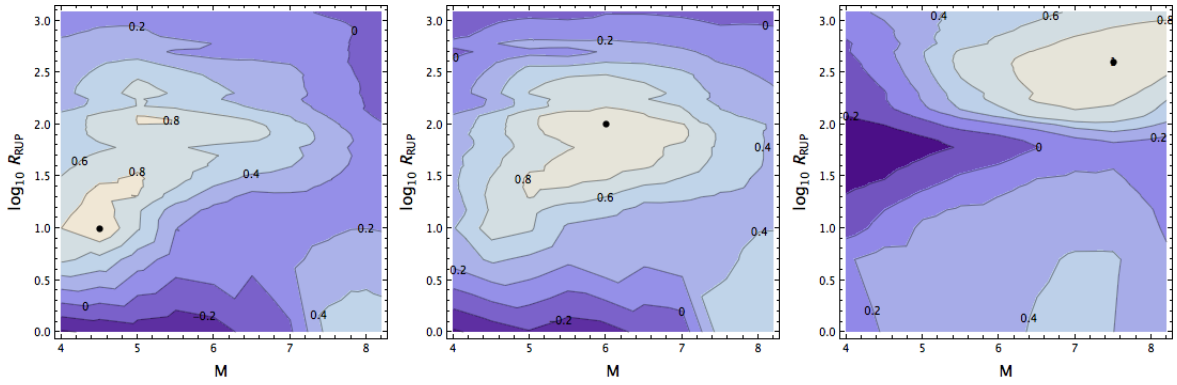


Figure 8–14 Five sampled functions from a 1D rational-quadratic covariance function [Equation (8–9)] with mean zero, for different value of the length-scale L.

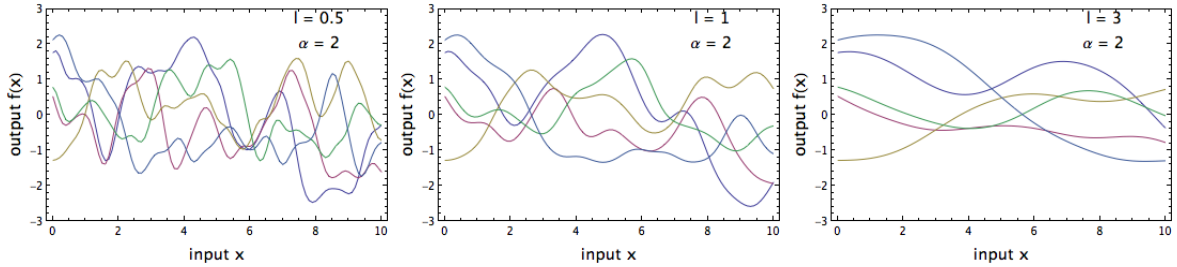


Figure 8–15 Five sampled functions from a 1D Gaussian process [Equation (8–9)] with mean zero and rational-quadratic covariance function, for different value of α and L = 1.

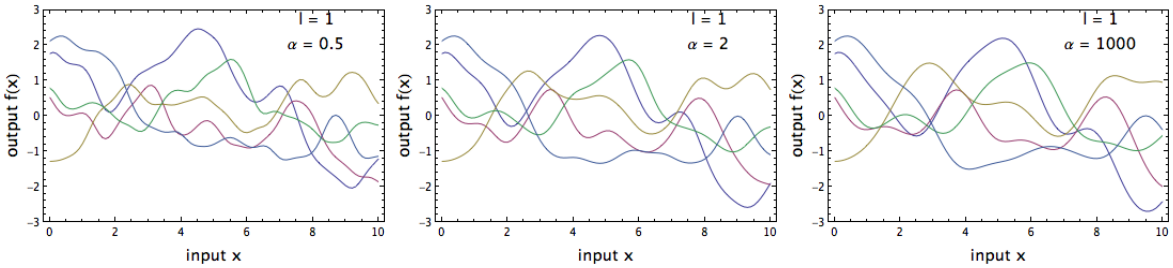


Figure 8–16 Covariance function as a function of the input distance r = |x−x'|.

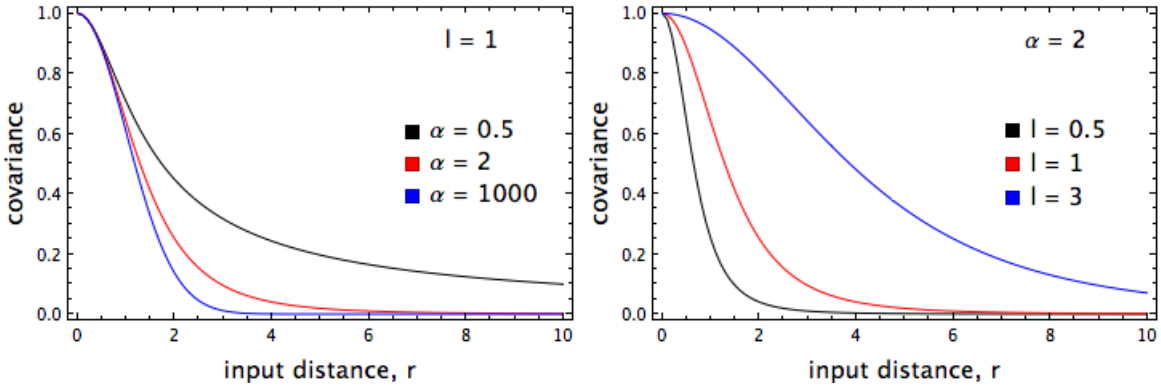


Figure 8–17 Five sampled functions from a GP with mean zero and linear covariance function with $\sigma^2_{\text{lin}} = 0.1$.

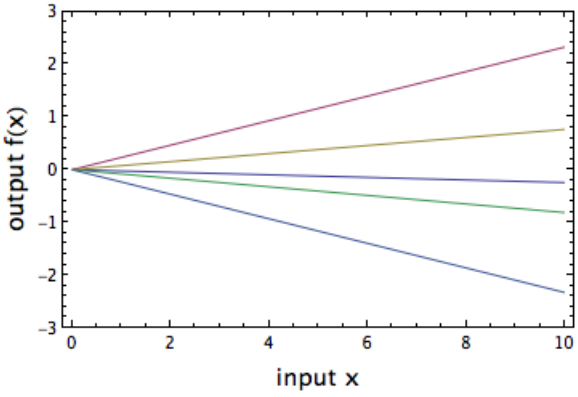


Figure 8–18 Modeled correlation coefficients for f = 1 Hz, plotted against M and $\log_{10} R_{RUP}$. Left; M = 5, R_{RUP} = 1000 km, Center: M = 6, R_{RUP} = 100 km, Right: M = 8, R_{RUP} = 20 km.

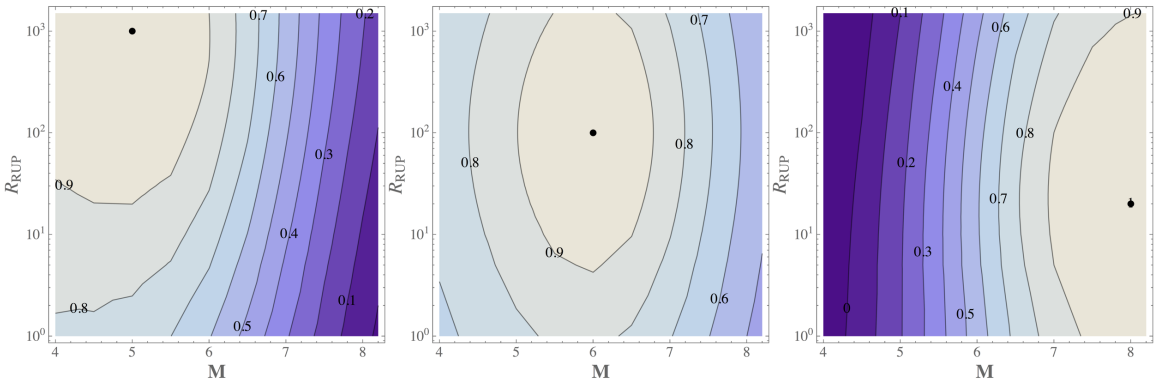


Figure 8–19 Histogram of estimates of 18 GMMs, for one scenario (left) and two scenarios (right).

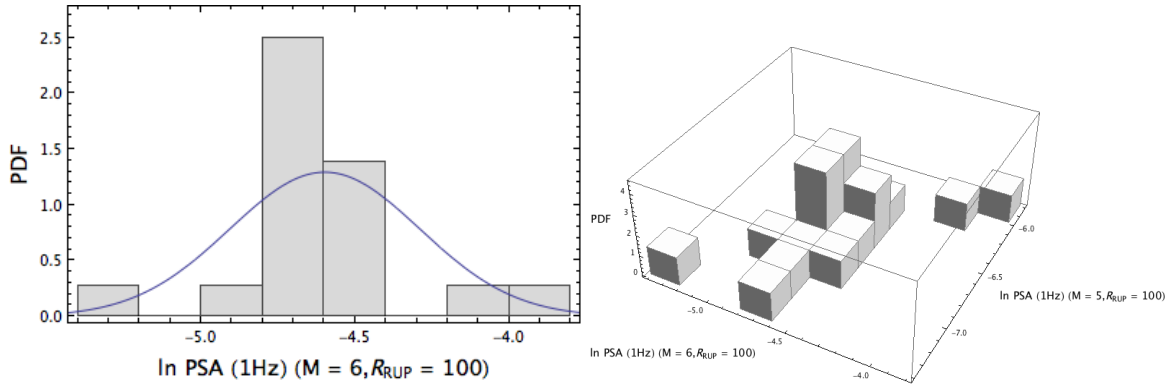


Figure 8–20 Magnitude and distance scaling for 18 GMMs used in this example.

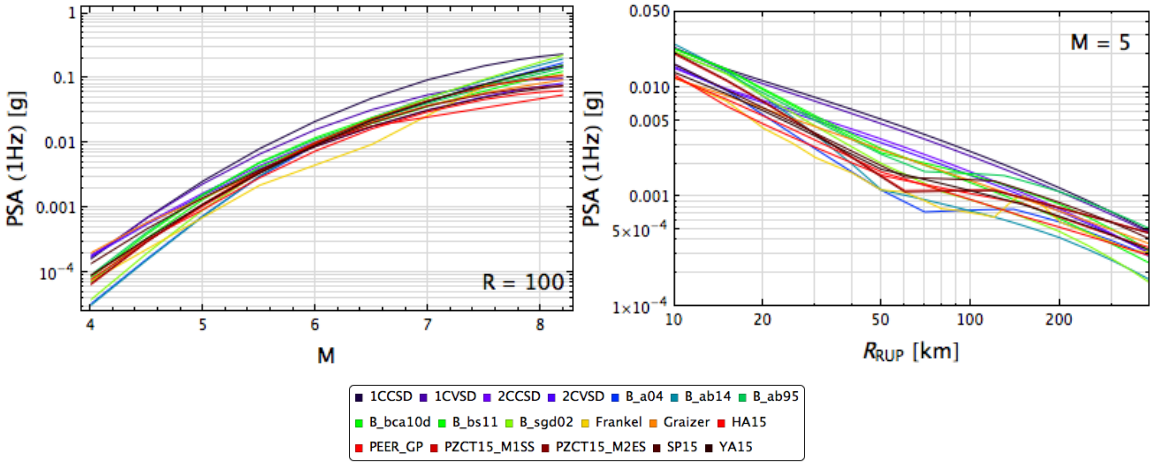


Figure 8–21 Pictures of rotated rubber ducks.

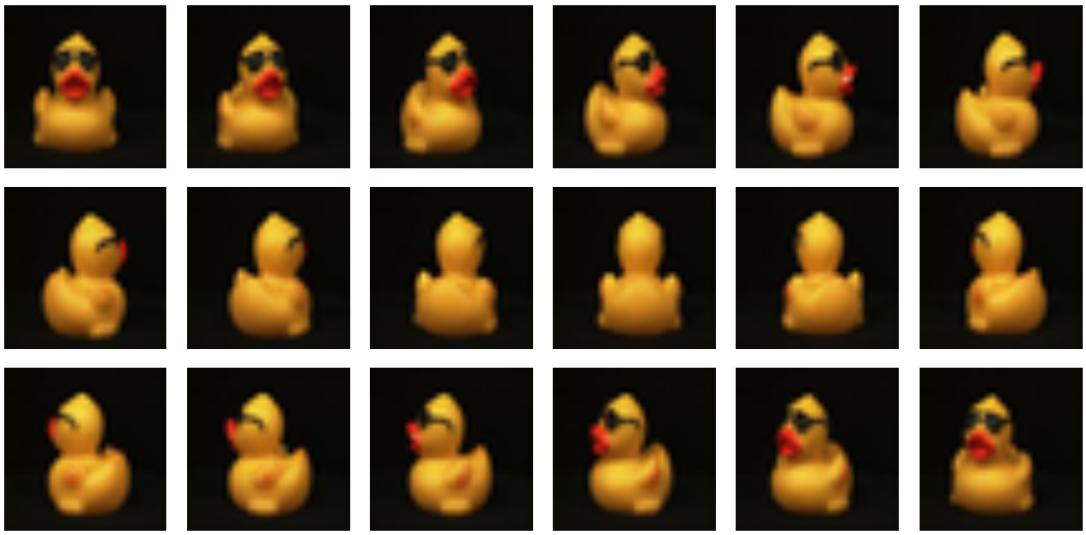


Figure 8–22 Left: Sammon map of “rubber duck" space, with insets of pictures; Right: same map, but with rotation angle of each picture at the respective coordinate.

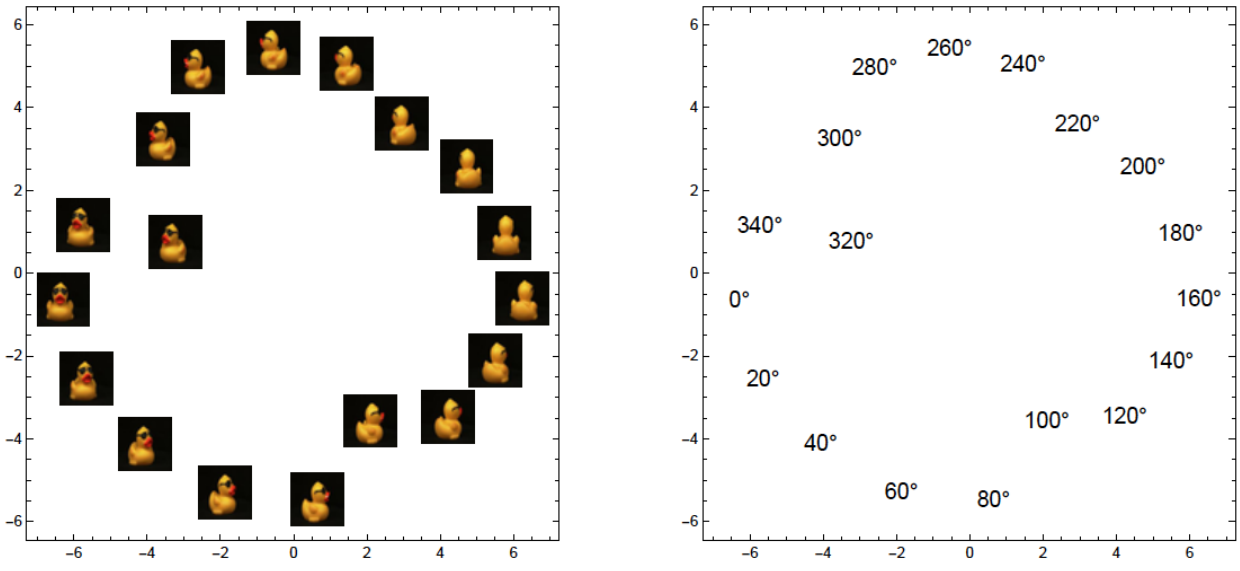


Figure 8–23 Differences between two GMMs at different M-distance values, which go into the calculation of the (high-dimensional) distance $\overline{\Delta_{GMij}}$ between GMMs i and j.

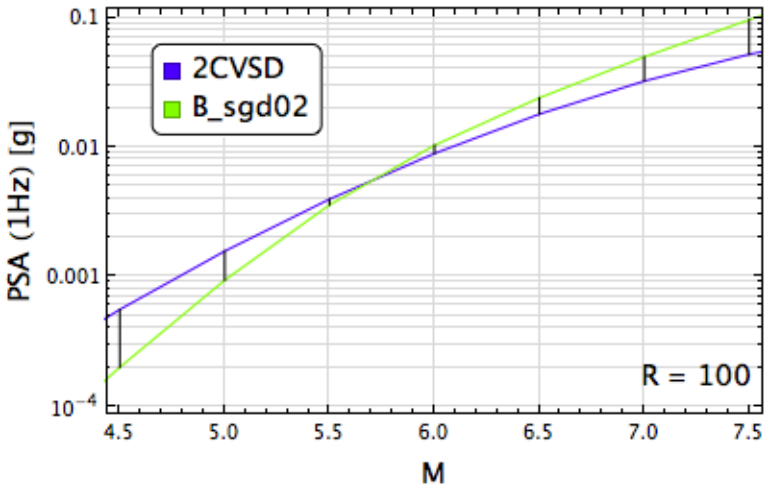


Figure 8–24 Difference in scaling of reference models that help the interpretation of the Sammon's maps.

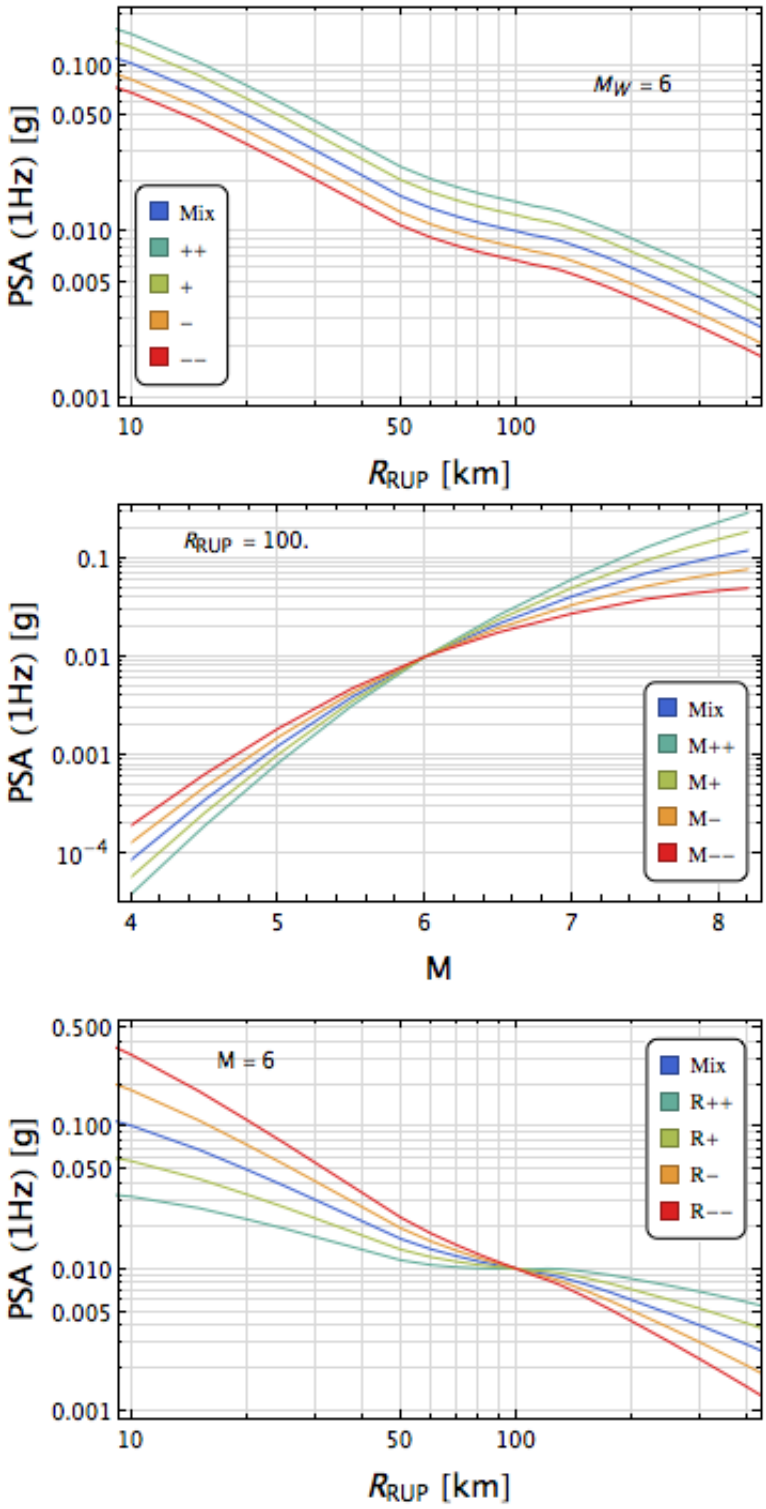


Figure 8–25 Sample 1 Hz Sammon's map for 18 GMMs, together with reference models. GMM-distances in high dimensions are calculated using the L_2-distance.

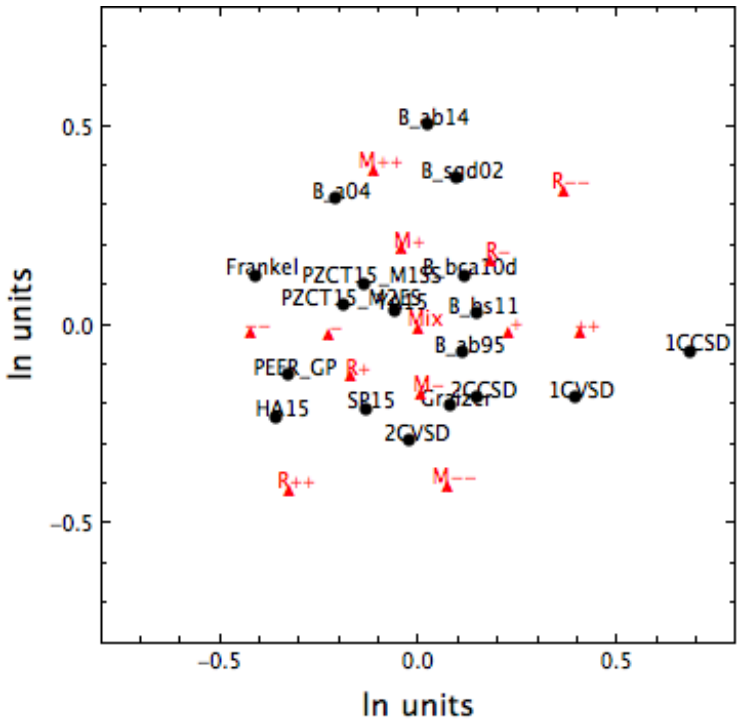


Figure 8–26 Scaling of 10 sampled models with individual seed GMMs as mean function, for three different seed GMMs. The mean is plotted as a dashed black line.

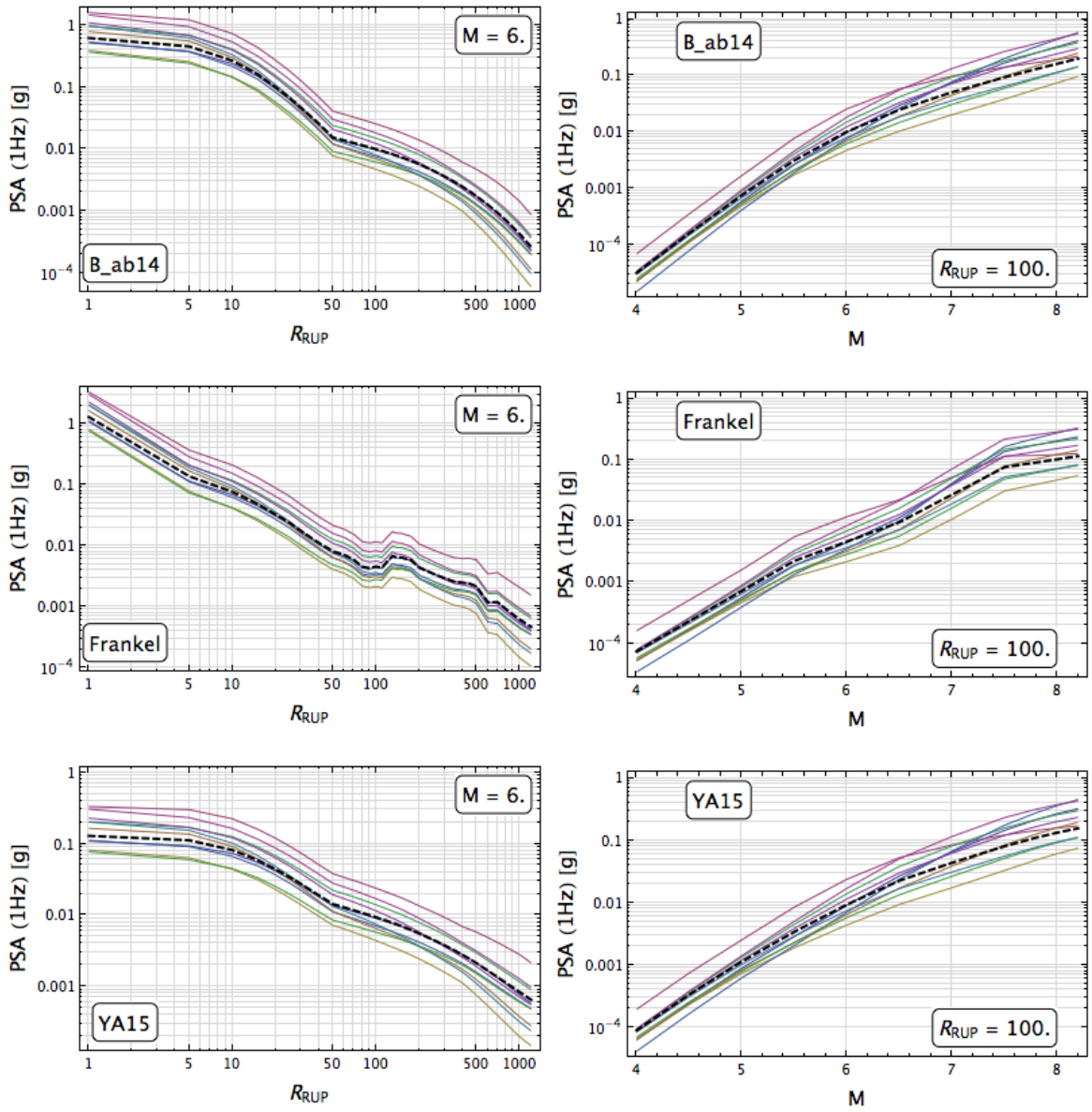


Figure 8–27 Sample 1 Hz Sammon's map for the assignment of initial weights, based on their proximity in Sammon's map space.

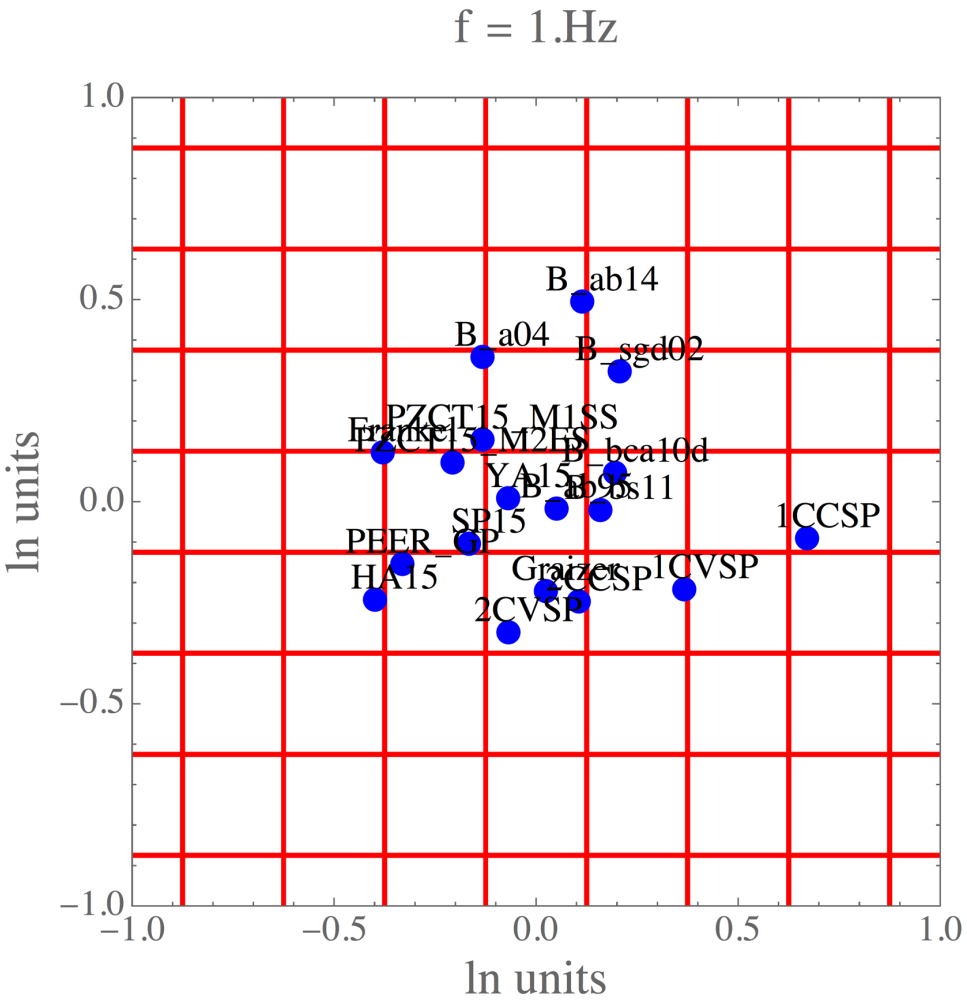


Figure 8–28(a) 1 Hz hazard sensitivity results to initial seed weight for the Manchester site.

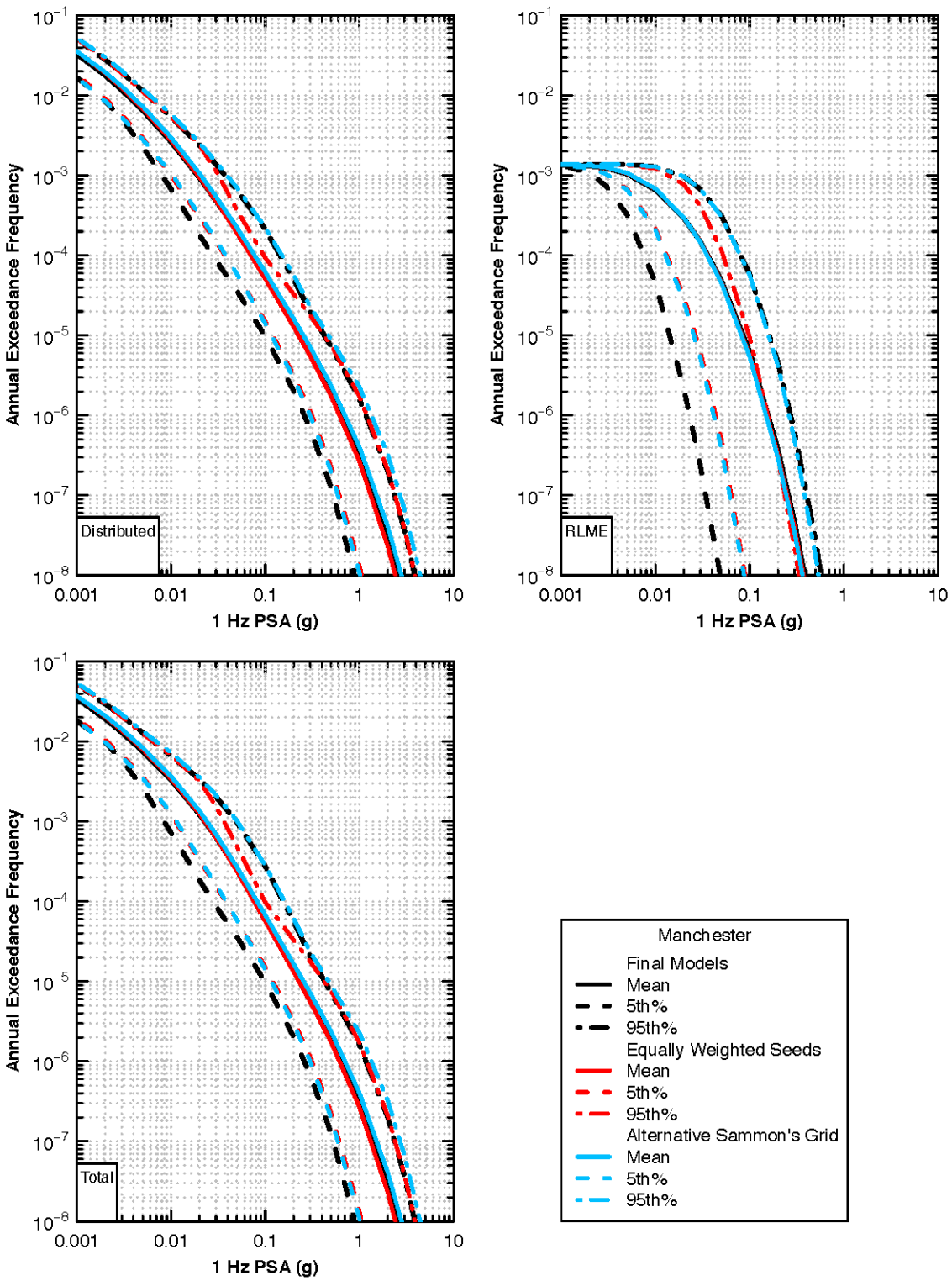


Figure 8–28(b) 10 Hz hazard sensitivity results to initial seed weight for the Manchester site.

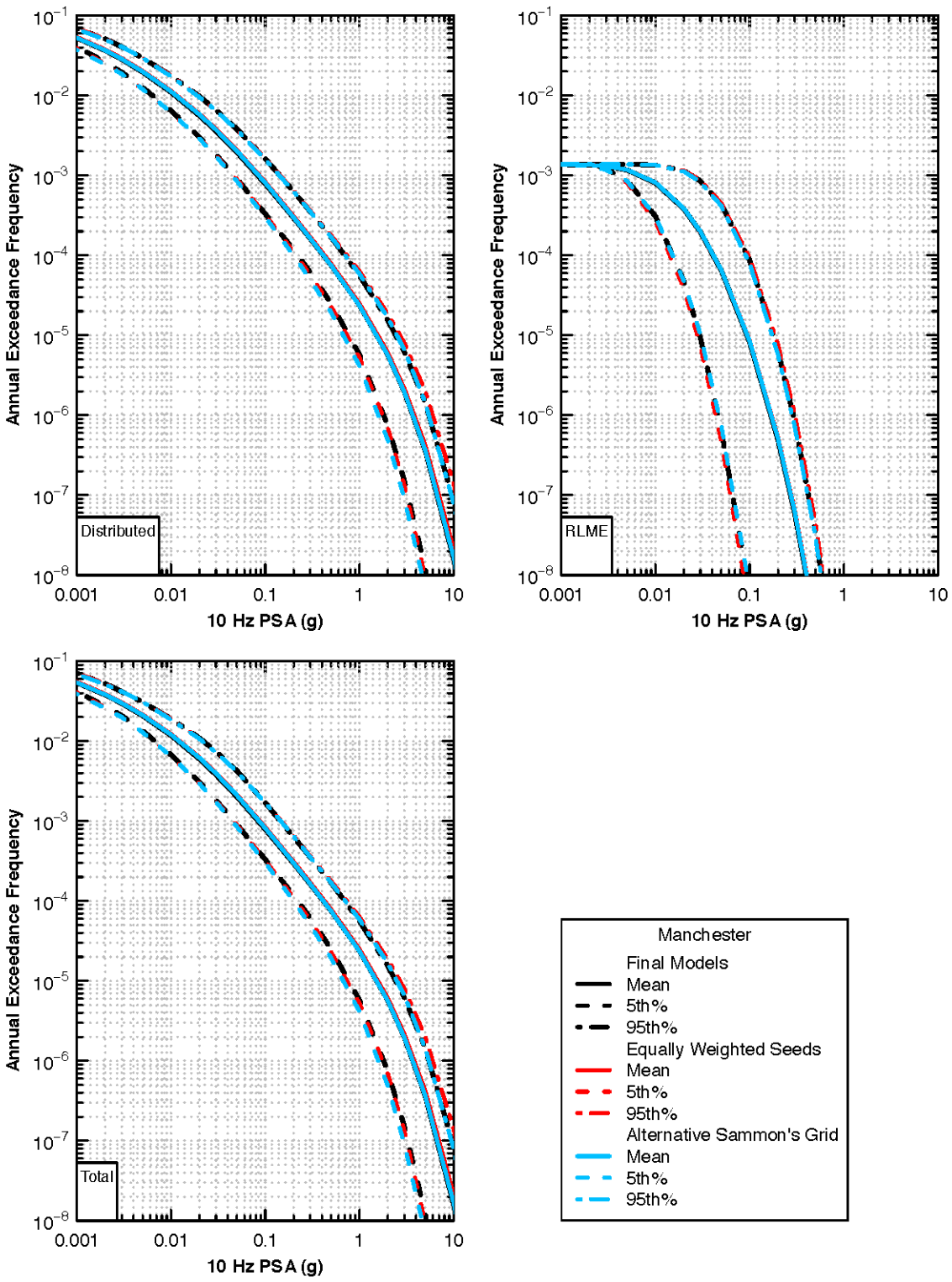


Figure 8–29 Scaling of 100 sampled models with individual seed GMMs as mean function, for three different seed GMMs. The mean is plotted as a dashed black line.

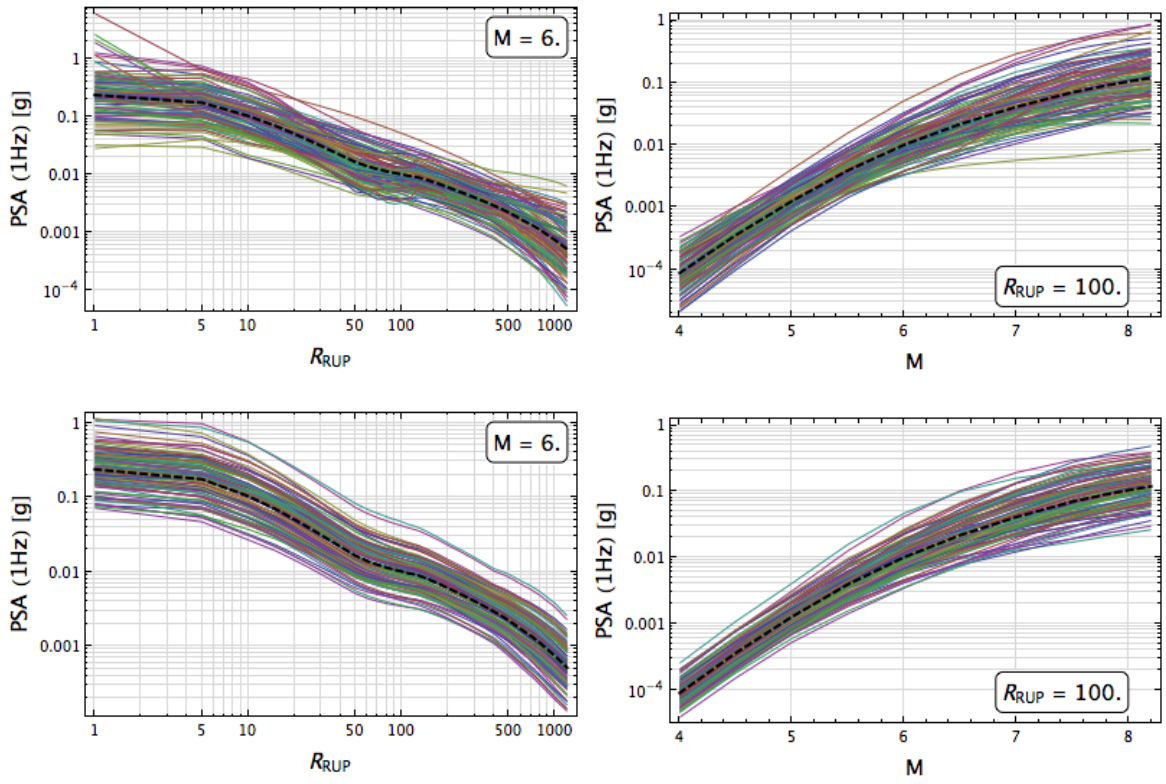


Figure 8–30 Minimum distance slopes (top) and magnitude ratios (bottom) from seed GMMs; PGA values are plotted at 200 Hz.

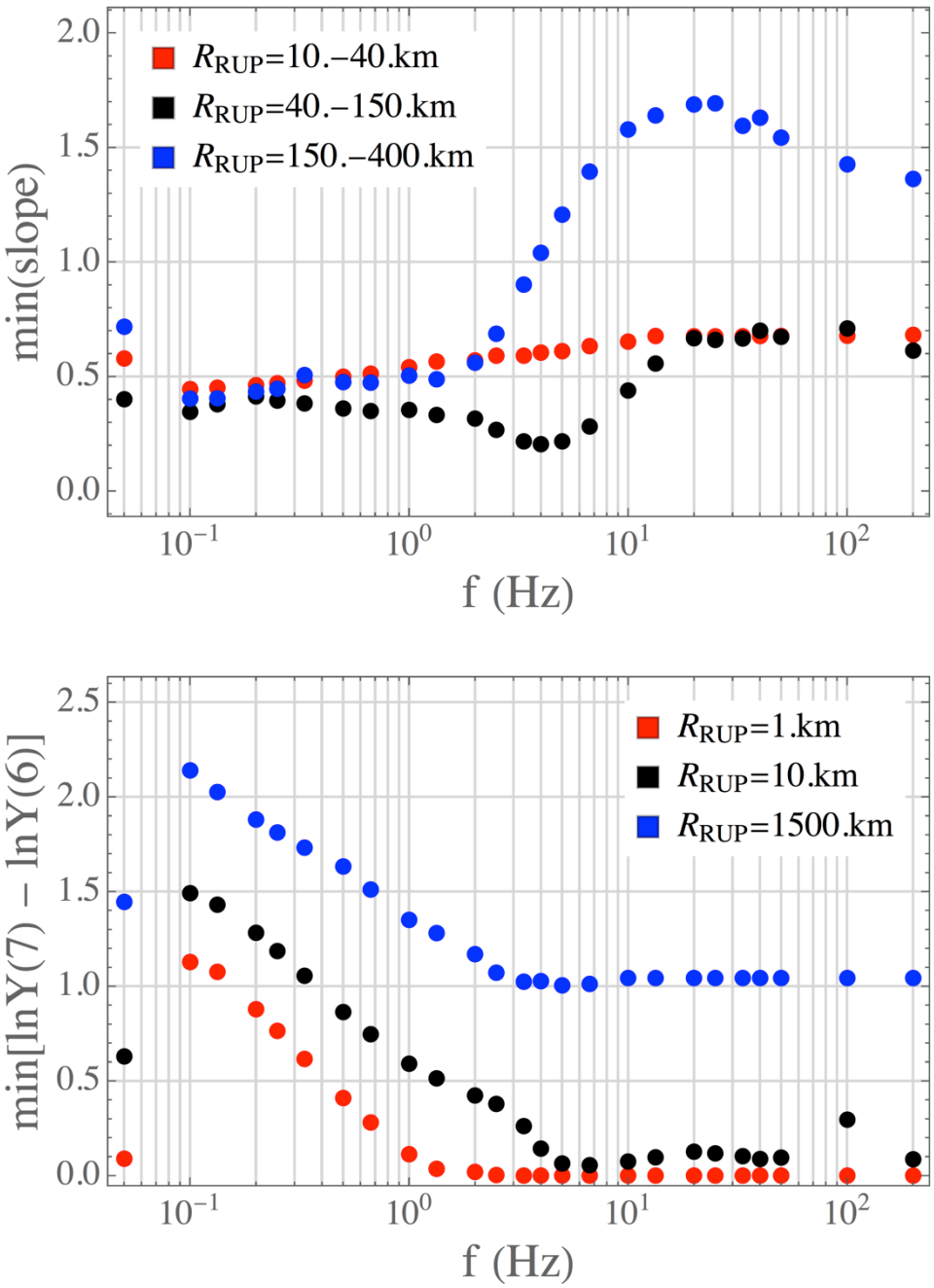


Figure 8–31 Sammon's maps space covered by 1,000 (left) and 5,000 models (right) in blue, relative to the 10,000 selected models (grey) for the 1 Hz case.

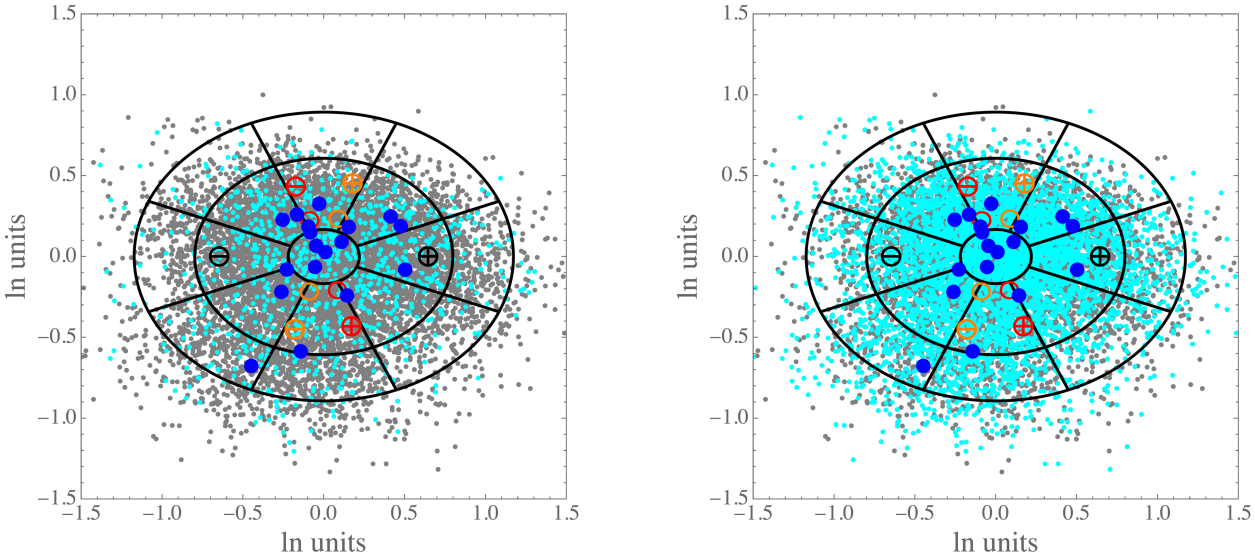


Figure 8–32 Number of rejected models due to failing the constraints on either magnitude scaling or distance slopes, for the generation of 10,000 models; PGA values are plotted at 200 Hz.

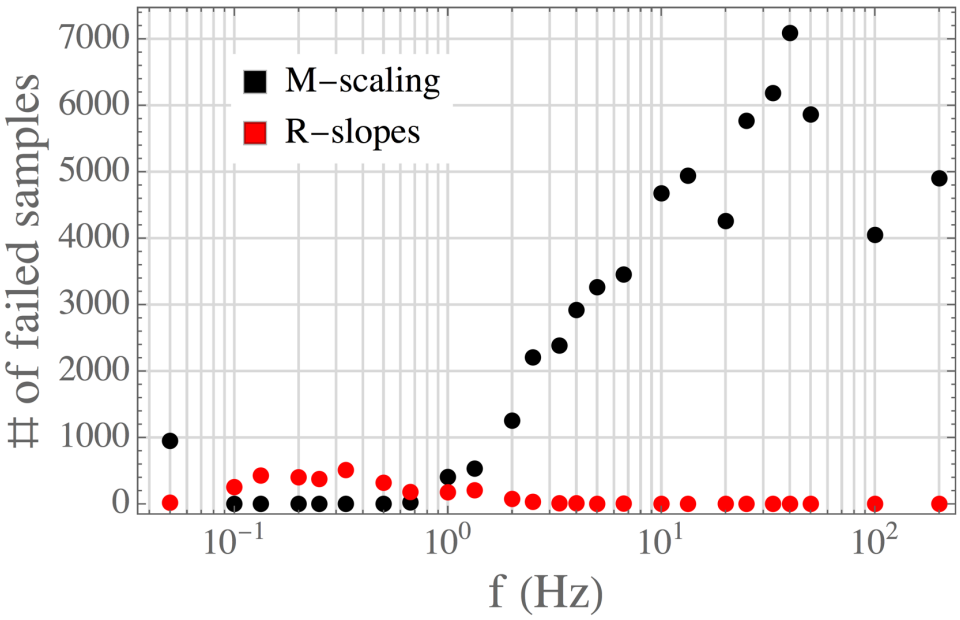


Figure 8–33 (a) Number of samples using the different seed GMMs as mean function for all NGA-East GMIMs.

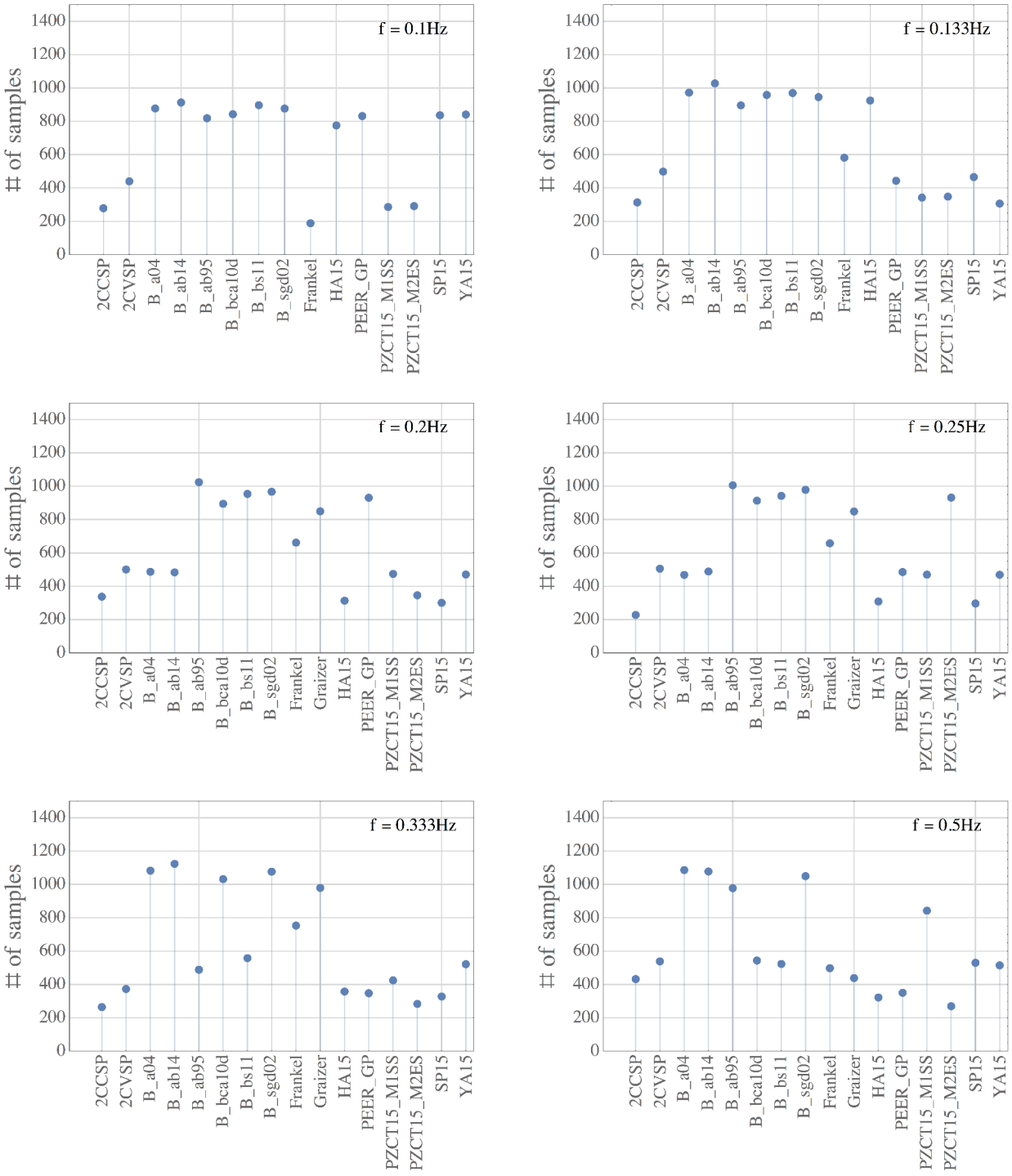


Figure 8–33(b) Number of samples using the different seed GMMs as mean function for all NGA-East GMIMs.

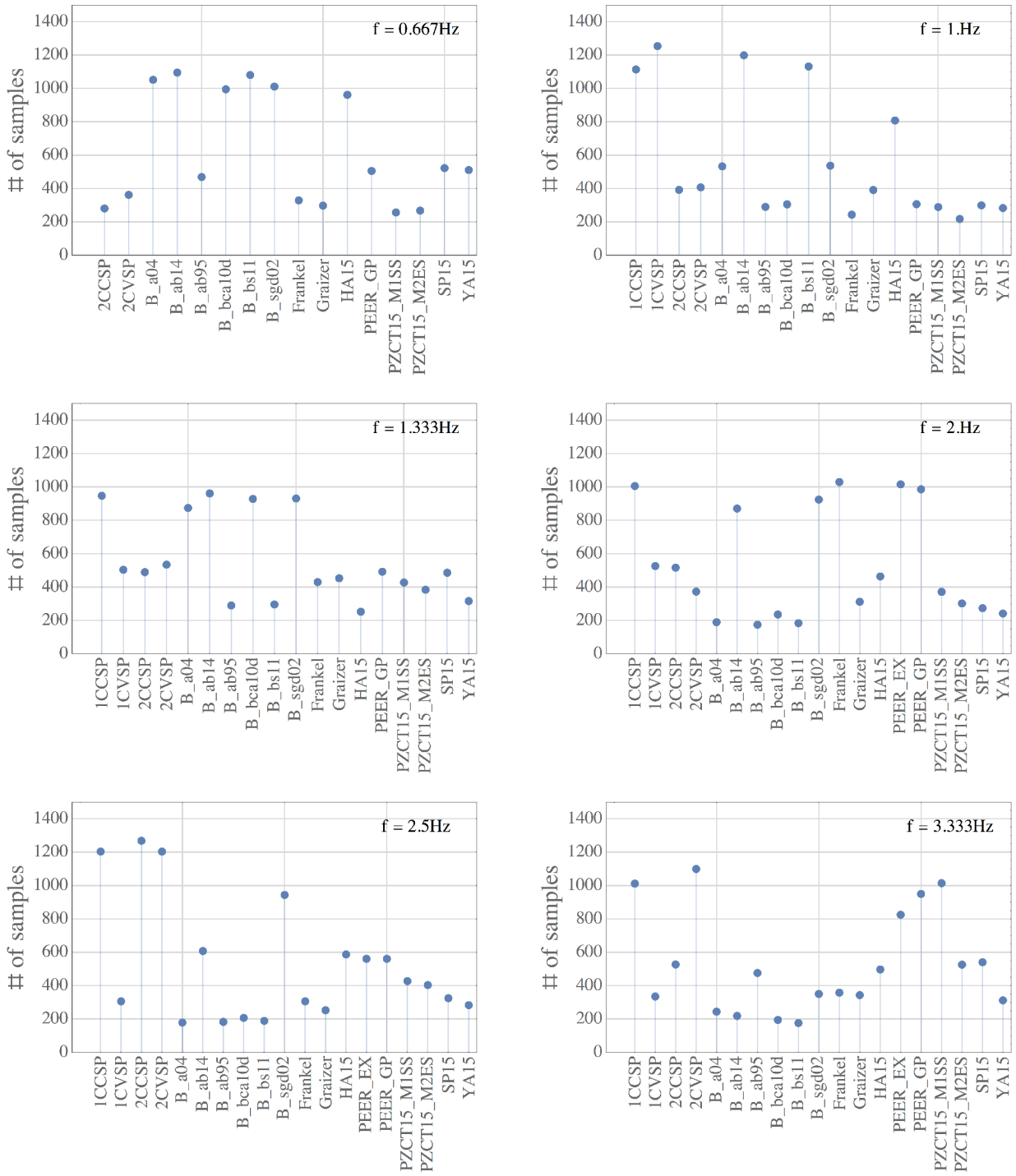


Figure 8–33(c) Number of samples using the different seed GMMs as mean function for all NGA-East GMIMs.

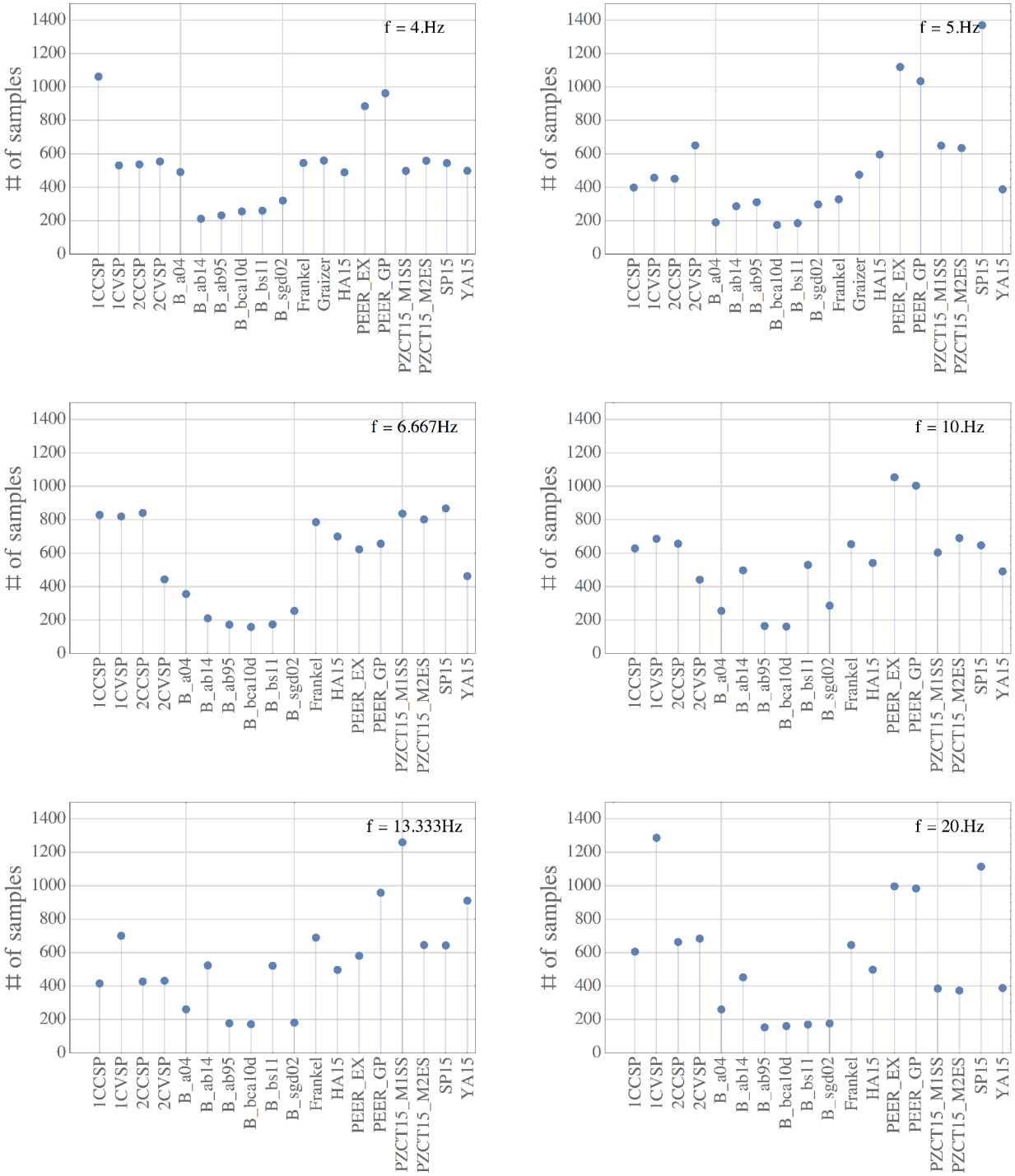


Figure 8–33(d) Number of samples using the different seed GMMs as mean function for all NGA-East GMIMs.

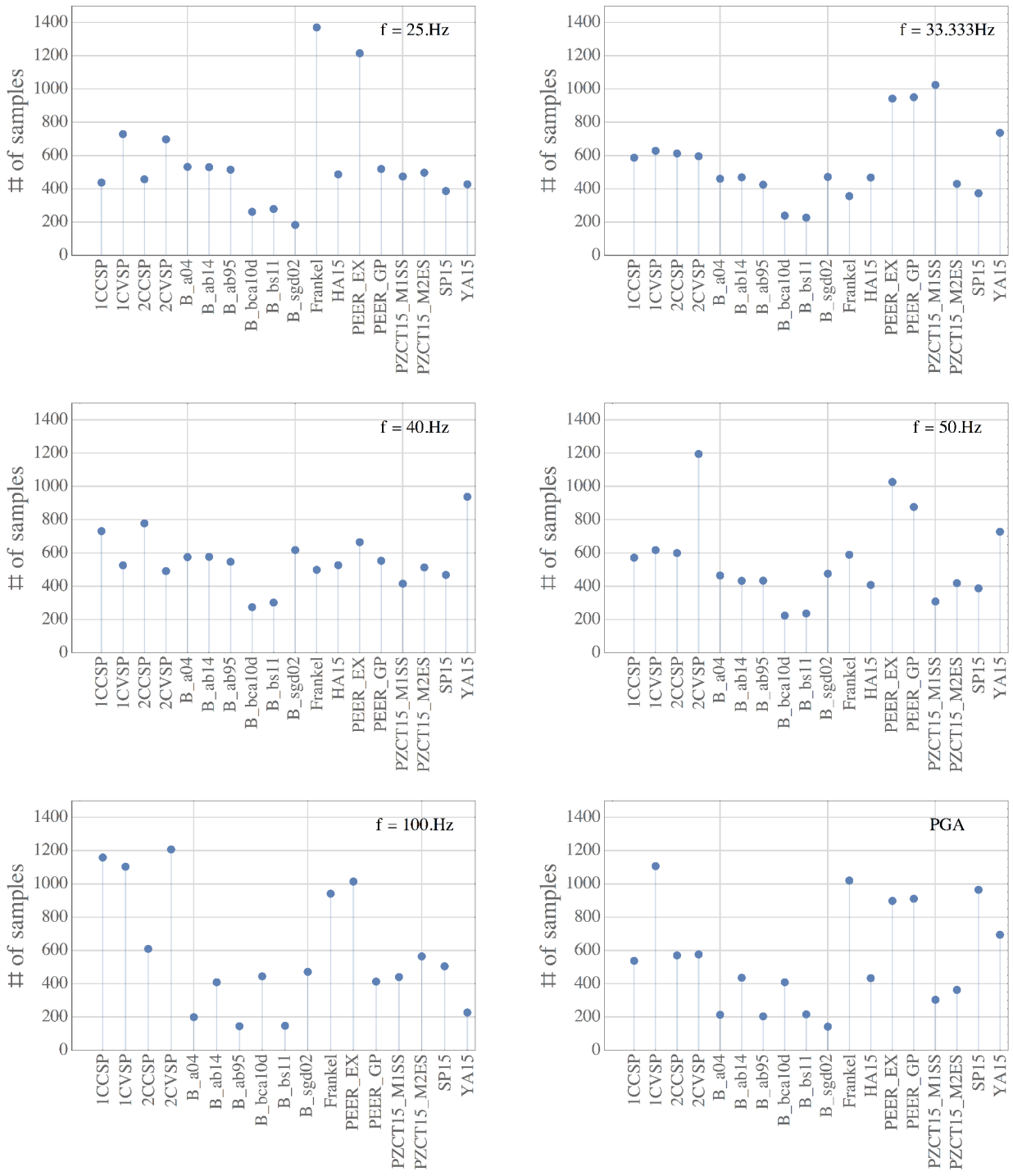


Figure 8–33(e) Number of samples using the different seed GMMs as mean function for all NGA-East GMIMs.

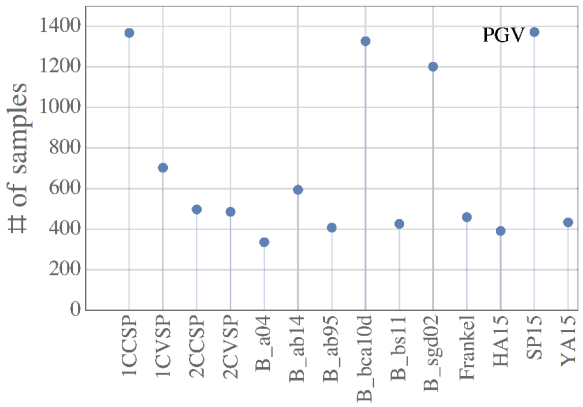


Figure 8–34 Distance and magnitude scaling for 200 sampled GMMs.

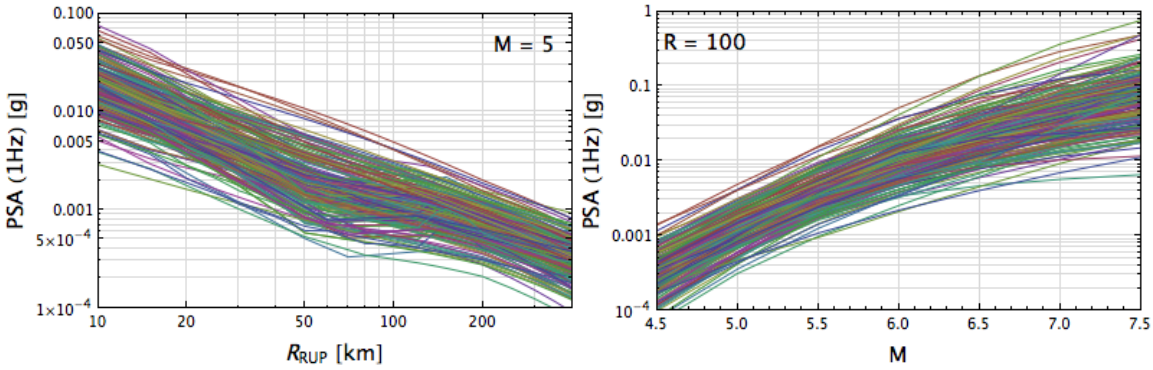


Figure 8–35 Three-dimensional ground-motion space, NGA-East correlation model, for 5000 sampled models (gray points) and the 18 seed GMMs (red points). The models form a thin cloud (with non-zero thickness).

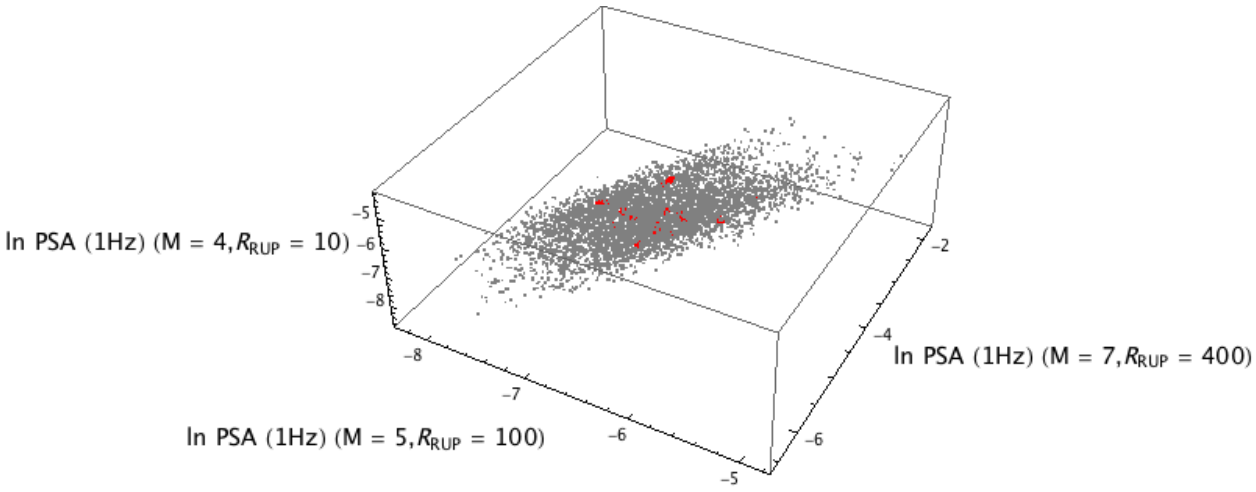


Figure 8–36 Three-dimensional ground-motion space, no correlation, for 5000 sampled models (gray points) and the 18 seed GMMs (red points). The models form a sphere.

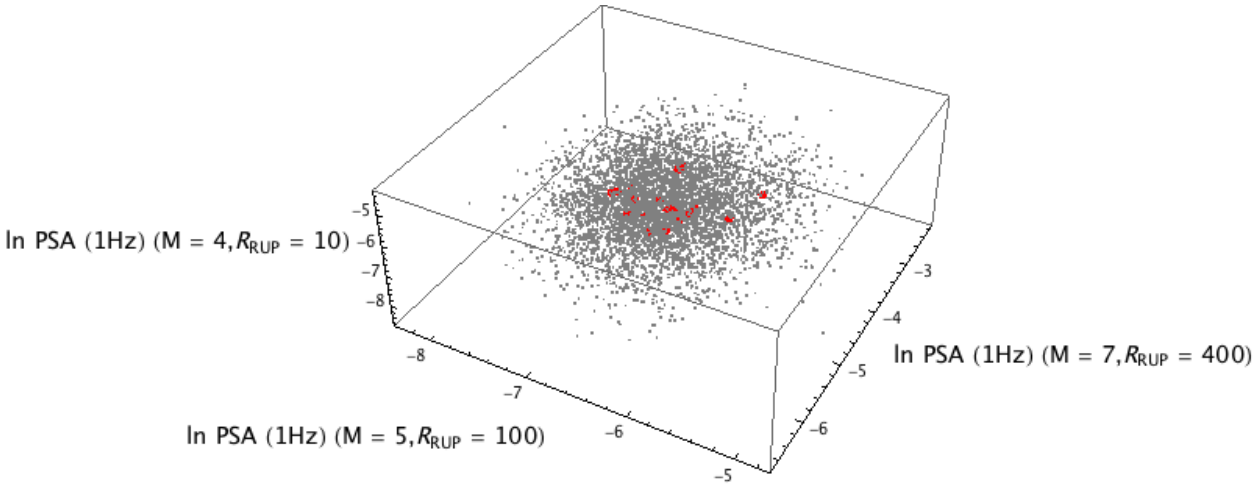


Figure 8–37 Three-dimensional ground-motion space, full correlation, for 5000 sampled models (gray points) and the 18 seed GMMs (red points). The models reside on sets of lines.

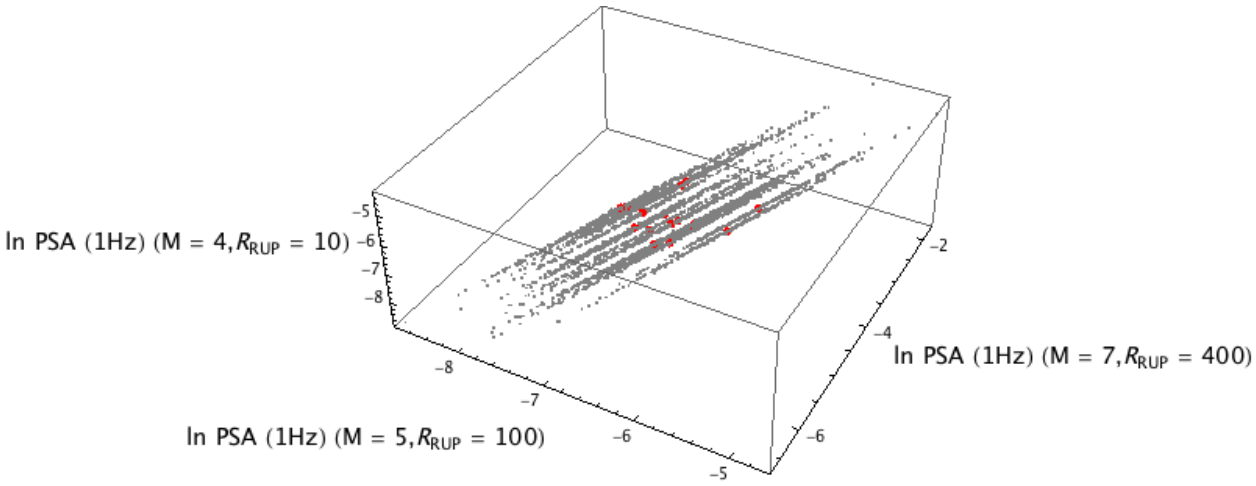


Figure 8–38 Sammon's maps for four different frequencies and 10,000 sampled models (gray points). Bue dots represent the seed GMMs and the orange and red symbol represent the magnitude and distance scaling reference models described in the text.

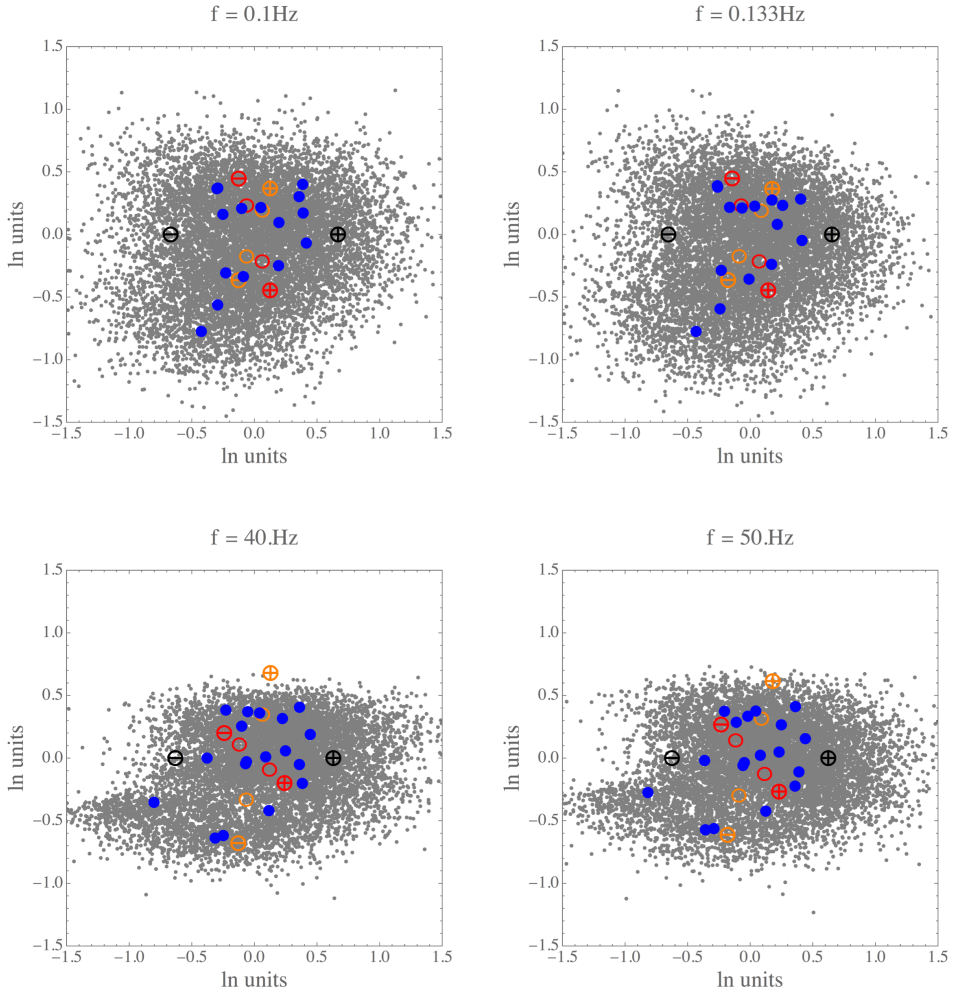


Figure 8–39 Variance of magnitude and distance scaling of the seed GMMs. The magnitude scaling is approximated by the difference between estimates of the seed models at M = 7 and M = 6 for R_{RUP} = 10 km. The distance scaling is approximated by the difference in between estimates at R_{RUP} = 10 km and R_{RUP} = 40 km for M = 6.

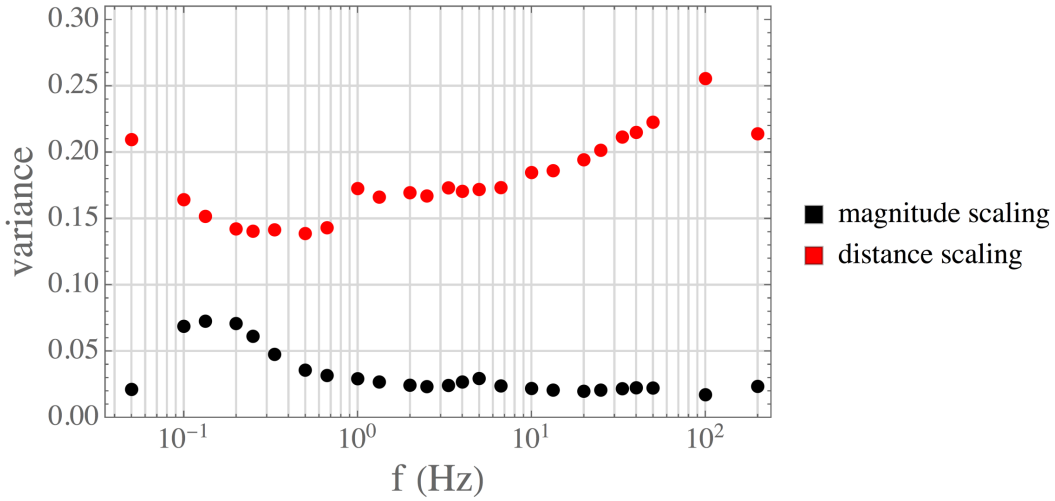


Figure 8–40 Sammon's maps for two different frequencies and 10,000 sampled models (gray points). The 95% range defined by the TI team is a black ellipse.

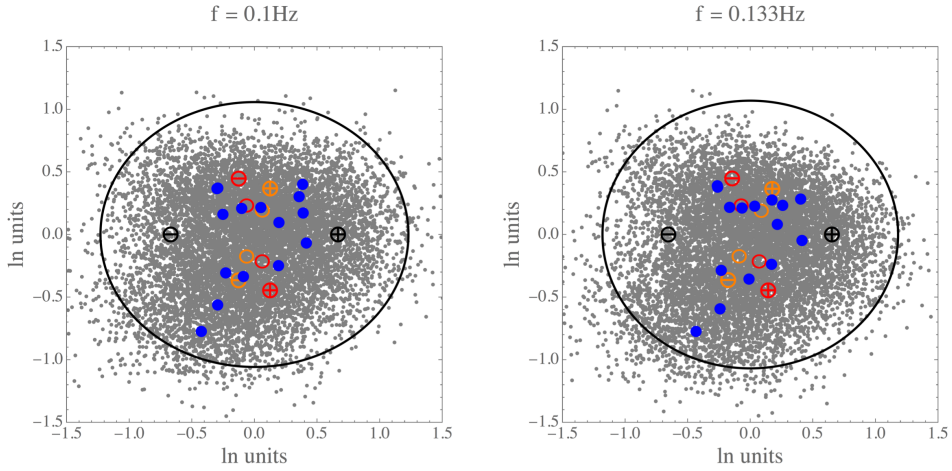


Figure 8–41 Sammon's maps for two different frequencies and 10,000 sampled models (gray points). The partition of the ground-motion space defined by the TI team are shown as black cells.

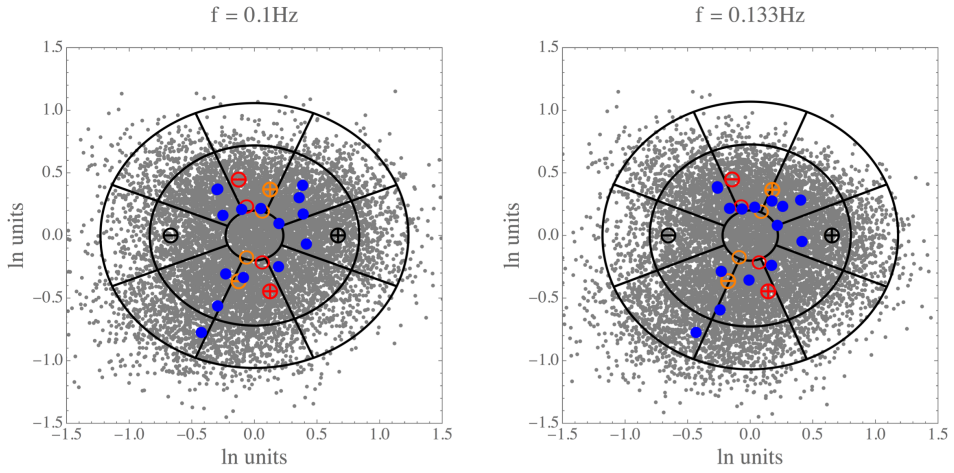


Figure 8–42 Model indices of different cells (left) and fraction of area of different cells with respect to the center cell (right).

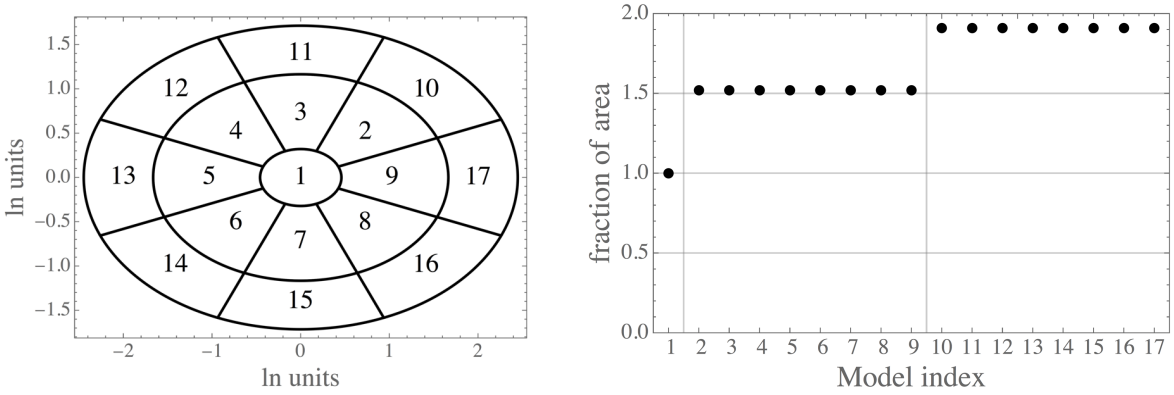


Figure 8–43 Distance and magnitude scaling of selected models for f = 1 Hz.

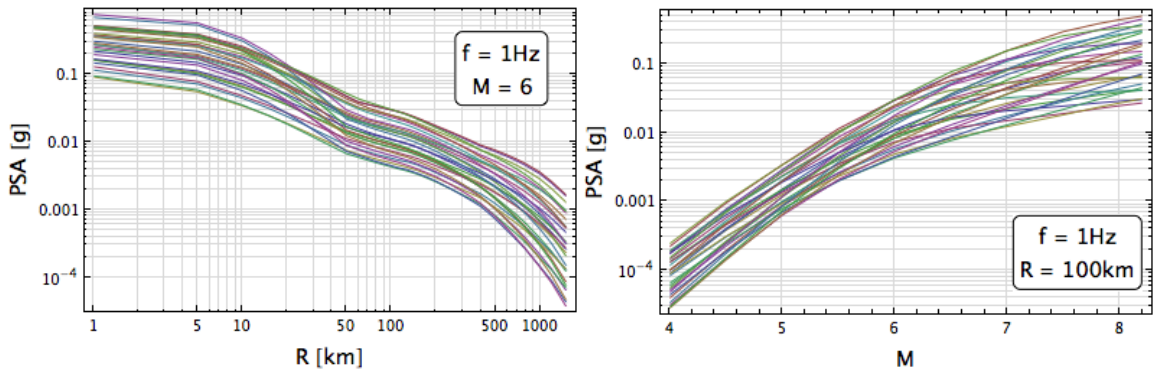


Figure 8–44 Spectra of 17 selected models for a single (M, R) scenario.

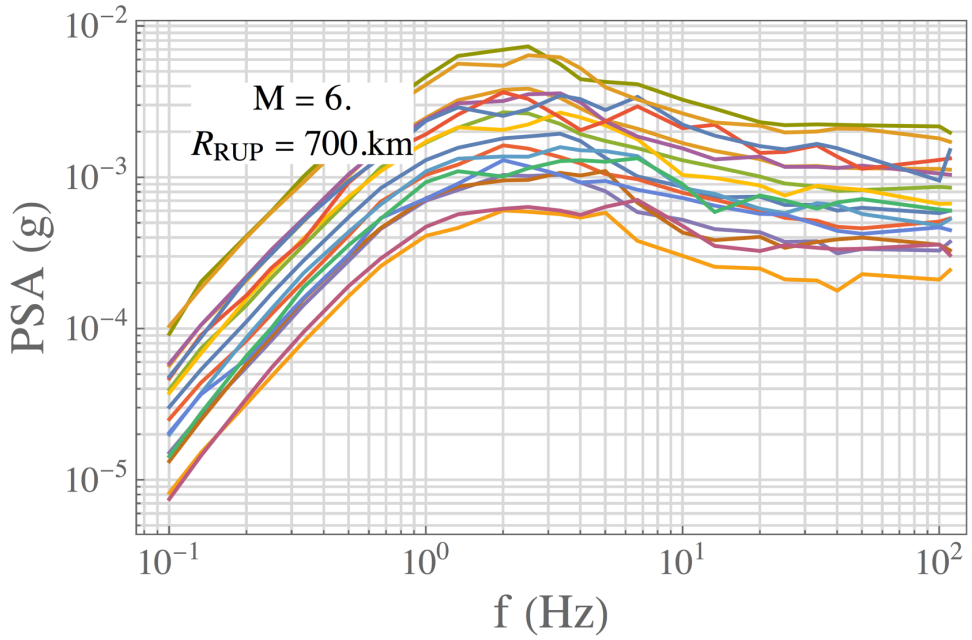


Figure 8–45 Three-dimensional ground-motion space for seeds and samples at two frequencies (1 and 10 Hz).

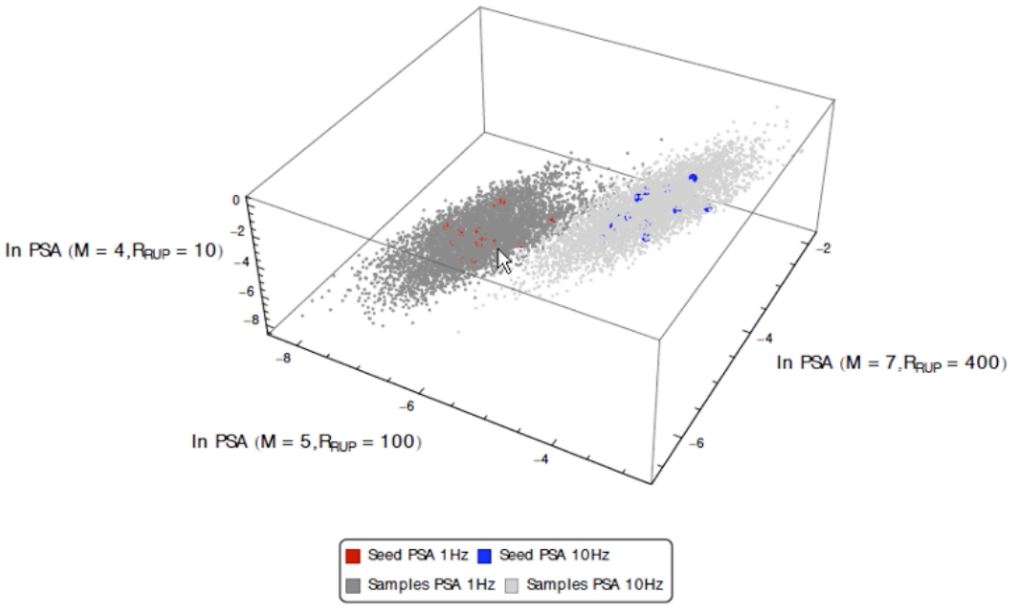


Figure 8–46 Frequency associated with PGA for spectral smoothing (see Appendix E.7).

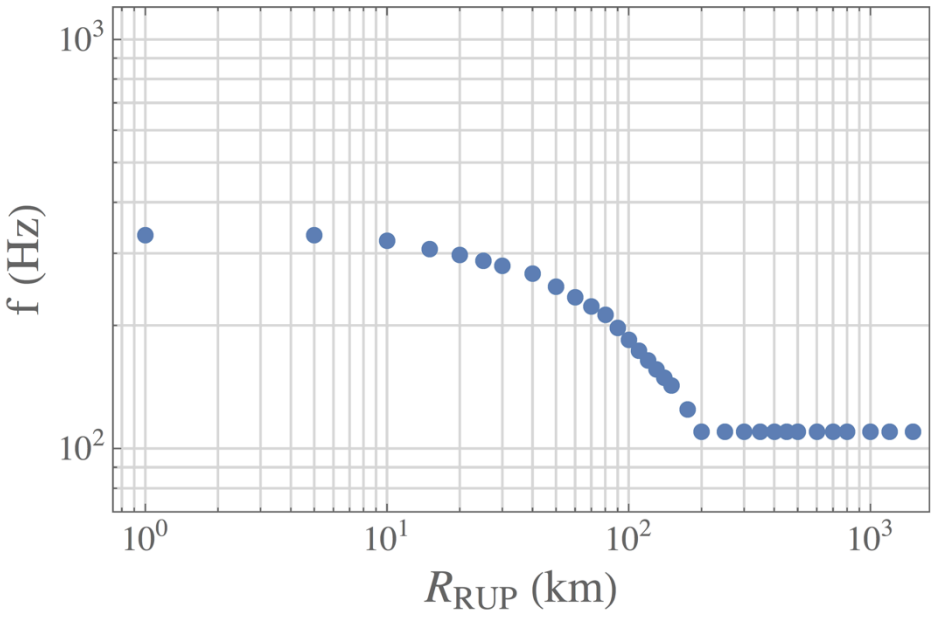


Figure 8–47 Example of spectrum smoothing (blue is raw, red is smoothed) for an individual model and a single (M, R) scenario.

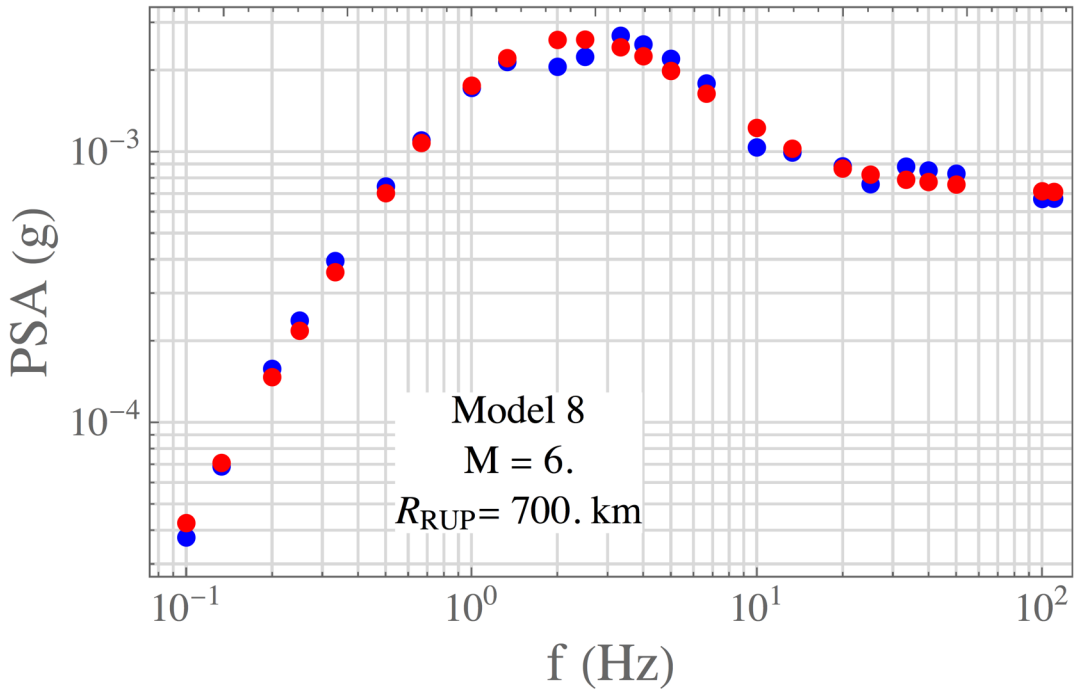


Figure 8–48 **Smoothed spectra of 17 selected models for a single (M, R) scenario (same models as in Figure 8–43).**

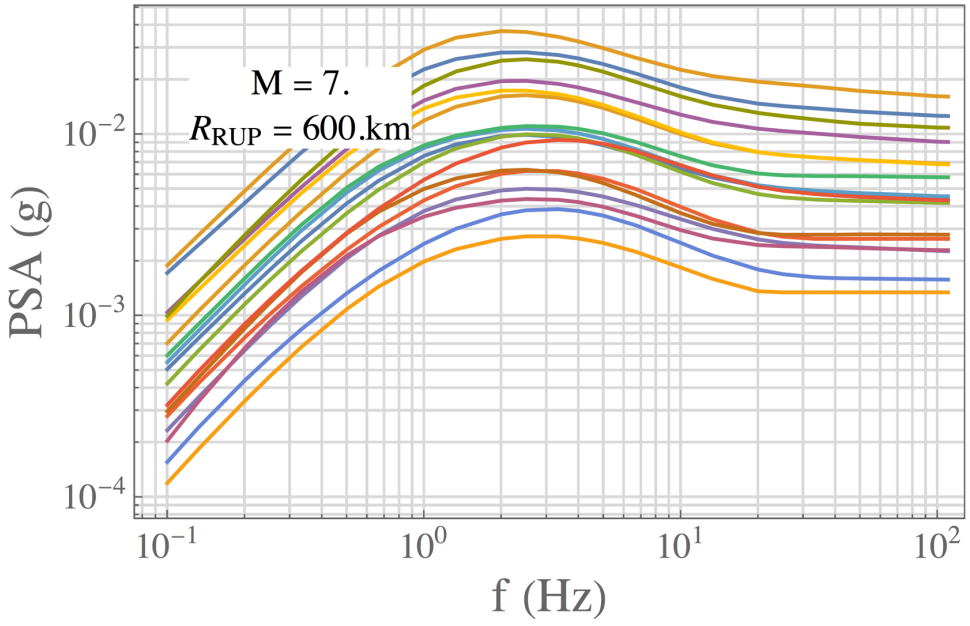


Figure 8–49 Sample 1 Hz hazard curves comparing results from the as-is and smoothed spectra, Manchester site. Weights from Chapter 9 were used for both cases.

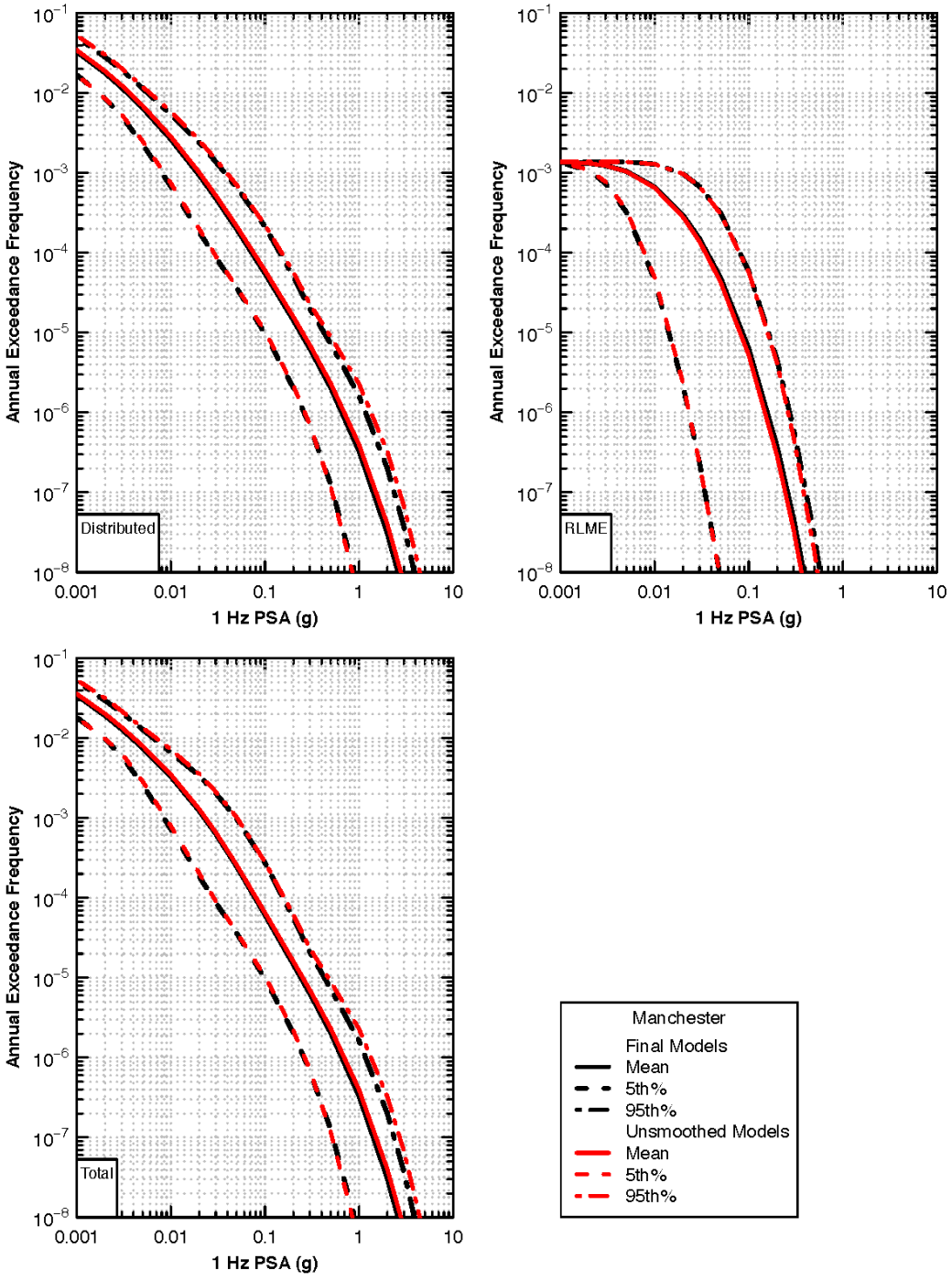
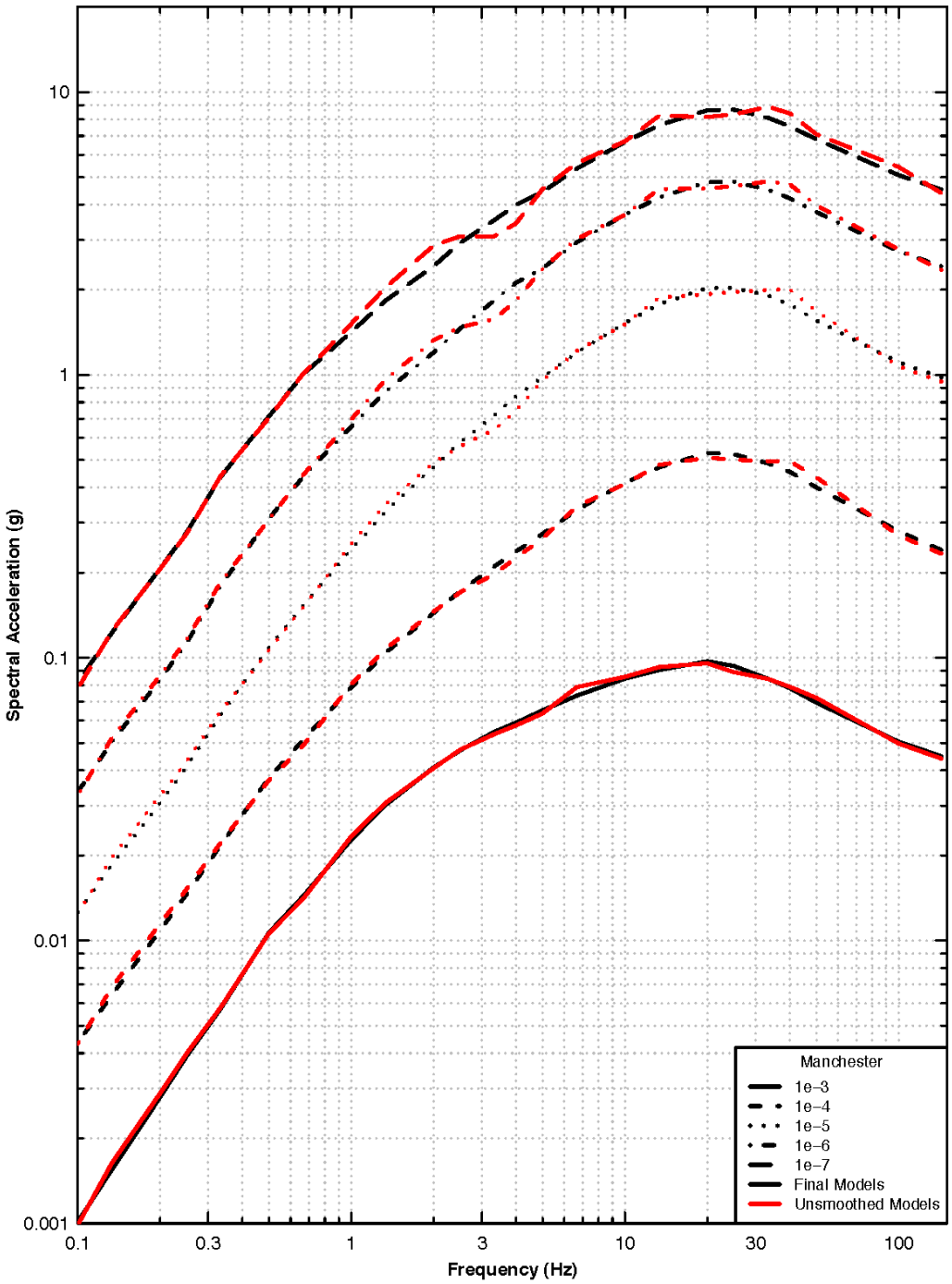


Figure 8–50 UHRS comparing results from the as-is and smoothed spectra for various hazard levels, Manchester site. Weights from Chapter 9 were used for both cases.



9. Logic Tree for Median Ground Motions

The current chapter describes Step 5 of the NGA-East approach to quantifying the epistemic uncertainty in median ground motions (Figure 6–9). The five steps are repeated here for convenience:

1. Develop a suite of seed ground-motion models (GMMs)
2. Develop parameters for continuous distributions of GMMs
3. Visualize the ground-motion space and sample GMMs
4. Re-discretize the ground-motion space
5. Assign weights

The general principles were introduced in Chapter 6. Step 1 was covered in Chapter 7, and Steps 2–4 were elaborated on in Chapter 8. Step 5 is detailed in the current chapter.

9.1 Logic-Tree Structure – Median Ground-Motion Models

The median ground-motion logic tree consists of 17 branches, each corresponding to the different GMMs developed in Chapter 8. The logic-tree structure is therefore very simple, with a weight assigned to each branch. Recall that the models were developed through the use of Sammon's maps (Sammon 1969), one at-a-time for each ground-motion intensity measure (GMIM). However, because there was a total of 25 GMIMs (a fairly large number selected to represent a smooth spectrum) and because the Sammon's maps were rotated and flipped to capture the same type of trend in each cell (ground-motion values, magnitude, and distance scaling), some of the correlation between frequencies was retained. To be consistent with the process (as is explained further below), the weights vary with frequency and for peak ground acceleration (PGA) and peak ground velocity (PGV), and are also based on Sammon's maps. The Technical Integrator (TI) team identified several different methods to set the weights for the different GMIMs. These methods were evaluated, and the TI team gave weights to each method (assigned "weights on weights"). The current chapter addresses the weight assignment for the 17 median models.

9.2 Weight Assignment Approaches

In Chapter 8, we defined P(**Y**) as the joint distribution of median ground motions at different (**M**, R_{RUP}) scenarios, for a given GMIM. This P(**Y**) distribution is therefore the full description of the center, body, and range of epistemic uncertainty associated with median ground motions. In Chapter 8, the distribution P(**Y**) was discretized into 17 distinct models. These 17 models were developed to each represent a cell, and are, therefore, a discrete representation of the center, body, and range associated with median ground motions. The 17 models need to be weighted, with the weights for each model reflecting its contribution to the center, body, and range of the epistemic uncertainty associated with median ground motions.

The TI team's approach to weighting the 17 different median GMMs was based on two considerations:

- The weighting should reflect the distribution P(**Y**).

- GMMs that better fit the observed ground-motion data should receive higher weight.

The first point reflects the view that the distribution P(**Y**), estimated from the seed GMMs, is a proper description of the epistemic uncertainty associated with median GMM estimates, and the weights for each model should reflect its likelihood with respect to that distribution.

The second point considers that available, observed ground-motion data provide information about the likelihood of the models. The 17 GMMs span a range of possible median predictions at each (**M**, R_{RUP}) scenario, so not all of them are unbiased for all scenarios. GMMs that fit the observed ground-motion data well should receive higher weights. There are different ways to evaluate a given GMM with respect to data; several approaches are described in Section 9.2.2.

Each weighting approach gives a set of weights for the 17 median GMMs. The TI team assessed the relative merits of the different weighting approaches. The justification for the weights assigned by the TI team for the different approaches is given in Section 9.4. All approaches take into account that each selected GMM is representative of an area on the map, as described by the cells in Figure 8–42. Therefore, the weight of a GMM representing an individual cell should reflect the contribution of that cell to the ground-motion space.

9.2.1 Weights Reflecting the Median Ground-Motion Distribution

As was described in Chapter 8, the range of epistemic uncertainty is defined—at each frequency—by an ellipse in Sammon's map space that contains most of the 10,000 sampled GMMs. Figure 9–1 shows an example Sammon's maps (1969) with the cell definition and the seed and sampled GMMs. Maps for all the NGA-East GMIMs were presented in Chapter 8. Two weights based on the distribution of sampled GMMs can be computed as described below. Both weights are based on the assumption that the distribution of points on the Sammon's map is a representation of the high-dimensional distribution P(**Y**).

9.2.1.1 Weight Based on the Fitted Two-Dimensional Probability Density Distribution

In the following section, the fitted 2D (elliptical) distribution is denoted by $P_{Fitted-PDF}(\mathbf{x})$, where $\mathbf{x}$ **X** is the random variable describing the location on the Sammon's map. The weight for each cell is proportional to the density of $P_{Fitted-PDF}(\mathbf{x})$ inside each cell k

$$w_k(\text{Fitted}-\text{PDF}) \propto \int_{A_k} P_{Fitted-PDF}(\mathbf{x})d\mathbf{x} \qquad (9\text{–}1)$$

where A_K is the area of the kth cell. The weights are symmetric along the two axes of the ellipse, as is expected from the fitted distribution. The weights from this approach would be constant for each cell in each ring. By definition, weights defined with this approach follow a normal distribution and are symmetric around the ellipse.

9.2.1.2 Weight Based on the Number of Sampled GMMs in a Cell

Another way to account for the distribution of sampled GMMs is to compute the weight directly from the number of models inside each cell

$$w_k\left(N_{samples}\right) \propto N_k \tag{9–2}$$

If the distribution of sampled GMMs was perfectly symmetric, the weights from Equations (9–1) and (9–2) would be identical (e.g., as was the case in the one-dimensional example in Chapter 6 and Figure 6–8). However, the weights based on Equation (9–2) take into account any asymmetry in the distribution of points on the map (Figure 9–1). Such asymmetry can be due to the characteristics of the seed GMMs themselves and/or to the application of the physicality criteria in screening the sampled GMMs, as described above. For example, Figure 8–32 showed that the number of rejected samples is different for the different seed GMMs depending on how close the seed GMM is to the physicality criteria. This can translate into asymmetry in the two-dimensional distribution of sampled points on the map.

9.2.1.3 Discussion

Figure 9–2 compares the weights based on the two different approaches for PSA at a frequency of 1 Hz. The two sets of weights are similar, with the weights from the number of points in each cell oscillating around the weights from the (theoretical) fitted distribution. However, the TI team preferred to use the number of points to assign weights from the distribution. This approach is reflective with the distribution of models in the Sammon's space and allows uneven weights around the ellipse, reflecting the effect of the physicality criteria applied by the TI team. Recall that we used a mixture distribution in the model development described in Chapter 8 (assume multivariate normal distribution, combined those statistics with mean-centered shapes from the original seed GMMs, and then applied physicality criteria, screening some of the samples out in the process).

9.2.2 Weights Reflecting the Fit to Data

The second approach to weighting takes into account that a model that fits the recorded ground-motion data better should receive a higher weight. Therefore, weights are calculated as an expectation over the area of the cell

$$w_k \propto \int_{A_k} f(\mathbf{x},D)P\left(\mathbf{x}\right)d\mathbf{x} \tag{9–3}$$

where f(**x**,D) is a function that depends on the data **D**. The TI team approach was to base weights on the mean residual with respect to data amplitudes, as well as the likelihood. To approximate the weights as calculated by Equation (9–3), the residuals and likelihood of data are calculated for each of the 10,000 sampled models. The data are a subset of the NGA-East database, with the constraint of limiting records to within R_{RUP} = 400 km (to avoid data truncation issues discussed in Section 7.6). The dataset consists of 468 records from 21 events (see Table 9–1; electronic appendix E.8.1 provides the uncorrected data in .csv format and electronic appendix E.8.2 provides the same corrected data as Table 9–1, but in the more practical .cvs format), including tectonic and potentially induced events (PIEs), all from source and station region 2 (Dreiling et al. 2014) within MCR. The records span **M** = 4.03 to 5.85 and R_{RUP} = 4.2 to 400 km, with their distribution as shown in Figure 9–3. The V_{S30} values range from 312 m/sec to 2000 m/sec. The data now have the same site condition as the seed GMMs: V_{S30} = 3000 m/sec and κ = 0.006 sec. The recorded data were corrected to that site condition using the B_avg1

method (Boore 2015) described in Section 7.6, where it was shown that different site correction methods produced slightly different adjustment factors. The B_avg1 method was selected among the available site correction methods as it resulted in the least bias, on average, relative to all the seed GMMs, across the frequency band of interest (1–10 Hz). The data-based weights are only computed for that 1–10 Hz frequency range.

For each of the 10,000 sampled models, residuals with respect to the data are calculated. The residuals $R = Y_{obs} - Y_{pred}$ (where both Y_{obs} and Y_{pred} are the natural log of the GMIM) that are calculated for each observed ground-motion data point relative to each model i (where i indexes the 10,000 sampled models), and the residuals are fit to a function

$$R = c_0 + \delta B + \delta W \tag{9–4}$$

where δB and δW are the between-event and within-event residual, respectively, and c_0 is the mean offset of the residuals. The parameters of Equation (9–4) are calculated using the random-effects algorithm of Abrahamson and Youngs (1992). The values of the between-event standard deviation τ and the within-event standard deviation ϕ are fixed to the average values of the NGA-East model (Sections 11.2 and 11.5). For each sampled model, the estimated likelihood—Equation 7 of Abrahamson and Youngs (1992)—and the mean offset c_0 are retained. Then, weights are calculated with

$$w_k(res) \propto A_k \frac{1}{\mu_k(c_0) + \delta} \tag{9–5}$$

$$w_k(L) \propto \frac{1}{N_k} A_k \sum_{i=1}^{N_k} L_i \tag{9–6}$$

where A_k is the area on the map for the k^{th} cell, $\mu_k(c_0) = \sum_{i=1}^{N_k} |c_{0i}|$ is the mean absolute offset of all models inside the k^{th} cell, and L_i is the likelihood of the ith model within the kth cell. Parameter δ serves to penalize the weights based on residuals for cells with a mean offset of (close to) zero since a cell with $\mu_k(c_0) = 0$ would otherwise have a weight of infinity. It is set to a value δ = 0.0075, which was used in the SWUS project (GeoPentech 2015). Figure 9–4 shows weights calculated according to Equation (9–5) using different values of δ, and one can see that the influence of δ is minimal.

Figures 9–5 and 9–6 show the residual and likelihood maps for all the relevant GMIMs (PSA frequencies from 1 to 10 Hz; see Section 7.6 for explanation). Figure 9–7 shows the weights calculated from Equation 9-5 and Equation 9-6 for the same GMIMs, for the 17 final GMMs.

The contour plots shown in Figures 9–5 and 9–6 are difficult to interpret. This is because the underlying Sammon's maps are a representation of high-dimensional ground-motion space that corresponds to the whole (**M**, R_{RUP}) range of NGA-East, while the data range is rather limited (cf. Figure 9–3). Hence, models that are close on the Sammon's map are on average close for a magnitude range 4–8 and a distance range 1–1500 km, while the residuals and likelihood are only calculated over a range of **M** = 4–6 and R_{RUP} =10–400 km. This leads to the appearance of noisy contour plots. Figure 9–8 shows the distance scaling for the 100 sampled models that

have the lowest absolute c_0 offset for **M** = 5. These models, although close together in a distance range where there are ground-motion observations, they diverge at very large and small distances. Since these large and small distances are used in the generation of the Sammon's maps, these models are not close on the map, leading to the noisy or spotty appearance of the residual plots in Figure 9–5. Figure 9–9 shows contour plots of residuals and likelihood for Sammon's maps that are calculated using only scenarios corresponding to the range of observed data, i.e. **M** = 4–6 and R_{RUP} =10–400 km. As expected, these contour plots are less noisy.

There are noticeable trends in the contour plots, in particular, the contour plots of the residuals for the maps based on the scenarios corresponding to the data range. These trends probably correspond to magnitude and/or distance scaling that roughly align with the scaling inherent in the data. For the purpose of this section, however, these trends (or their orientation) do not matter. Models that are more concordant with the data should receive more weights, which they do if the weights are calculated according to Equations (9–5) and (9–6).

9.2.3 Weights Combining the Distribution and Fit to Data

The third approach to assigning weights to the individual models is a combination of the previous two approaches, and takes into account both the GMM distribution and the fit to the data. It is based on calculating the posterior probability of each sampled model given the data. To calculate this probability, the value of the distribution $P_{Fitted-PDF}(\mathbf{x}_i)$ is multiplied with the likelihood L_i for the i^{th} sample model m_i:

$$P(m_i|D) = P_{Fitted-PDF}(\mathbf{x}_i)\, L_i \tag{9–7}$$

Weights for each cell can then be calculated by averaging the values of the posterior distribution over the cell k

$$w_k(post) \propto \frac{1}{N_k} A_k \sum_{i=1}^{N_k} P(m_i \mid D) \tag{9–8}$$

Figure 9–10 shows example weights as calculated from Equation (9–8). Some of the weights are zero.

9.3 Evaluation of Weight Assignment Approaches

9.3.1 Weights Reflecting the Median Ground-Motion Distribution

As mentioned earlier, the ellipse in the Sammon's map spans the epistemic uncertainty in the median GMM estimates. It is defined based on the dispersion of sampled GMMs along the horizontal and vertical axes. Hence, the two-dimensional distribution represented by the ellipse corresponds to an ideal two-dimensional normal distribution. On the other hand, the Sammon's maps show an uneven distribution of points, with some parts of the maps showing a higher density of points than others. Thus, some cells, even on the same ring, contain fewer points than others. There are several factors that come into play. On the one hand, the physicality criteria screen out models. The screened out models fail due to either their magnitude scaling (being too flat or having negative magnitude scaling) or because of a distance scaling that is too

steep. Hence, regions on the map that correspond to a steep distance scaling or a flat magnitude scaling have more samples removed, leading to an uneven distribution of points on the map. Furthermore, the ground-motion distribution P(**Y**) is a mixture distribution, with the seed models as means of the mixture components. In general, the TI team wanted the weights to reflect the unevenness of the distribution of points on the Sammon's map. For example, models that are close to being rejected due to the physicality criteria should receive a lower weight. The exact scatter of the density of the points do not matter; what is important is that the weights retain them. The w_k(Fitted–PDF) are not sensitive to the density of the points on the map, but the $w_k(N_{Samples})$ are sensitive to density. Because the TI team wanted the weighting scheme to account for this new distribution of ground motions, we could not assume that the pdf based on multivariate normality applies for all frequencies and all cases. Although Figure 9–2 shows this may be a negligible issue, it need not be the case for all the sensitivity studies; using the number of points for weight allows a systematic reproducibility of the approach without the need to assume that the normal multivariate conditions apply. Because of this, the TI team preferred to give full weight to the weighting scheme based on $w_k(N_{Samples})$.

9.3.2 Weights Reflecting the Fit to Data

The weights based on residuals to the data, w_k(res), represent an assessment of how well a model (or the samples inside cell k) fits the ground-motion data on average. The likelihood based weights, w_k(L), also take the variability of the data into account. A model that predicts the mean of the observed data for all (**M**, R_{RUP})-scenarios has a mean residual of zero but may have a large residual standard deviation depending on how closely the model tracks the data trends. Hence. it is possible to have a magnitude and/or distance scaling that is not in agreement with the data but still have a low average residual; such a model will have a large residual standard deviation. Since the between-event standard deviation and the within-event standard deviation in the calculation of the likelihood are fixed, these models have a lower likelihood. Thus the weights based on the likelihood provide more information on the agreement in magnitude/distance scaling between data and models in the (**M**, R_{RUP}) range where there are ground-motion observations.

Figure 9–8, which displays the distance scaling of the 100 sampled models with the lowest average offset to the data, shows that none of these 100 models exhibit a scaling that is unreasonable. Thus, the TI team concluded that residual weights provide some information about the fit to data. The likelihood-based weights typically assign large weights to a small number of models. Even though these models are preferable from a theoretical standpoint, as argued above, the TI team felt that heavily weighting such models would lead to a distribution of GMMs that was too narrow because of the narrow range of available data. Hence, the TI team assigned equal weights to these two residual-based approaches.

9.3.3 Weights Combining the Distribution and Fit to the Data

The weights based on the posterior distribution given the data, w_k(post), are a combination of the "prior" distribution of the GMMs, $P_{Fitted-PDF}(\mathbf{x}_i)$, and the fit to the data via the likelihood. However, as seen from Figures 9–2 and 9–4, the weights based solely on these approaches favor a relatively small number of models. Hence, their combination in terms of a posterior

weight would enforce a stronger weight on the center model (Figure 9–10). The TI team decided that this was not justified since this would reduce the epistemic uncertainty for scenarios that are hazard-relevant but are not represented within the ground-motion data. Specifically, the posterior distribution leads to a concentration of the model predictions that is driven by small-magnitude data; the TI team does not believe that this concentration should be permitted to extend to large magnitudes, for which there are no corresponding data constraints.

In addition, the prior distribution is based on the fit of a normal distribution to the distribution of points on the Sammon's map, which does not take into account the asymmetry of the distribution as explained in Section 9.2.1. Therefore, the TI team disregarded weights based on the posterior probability of the models.

9.4 Selected Weights

Figures 9–2 and 9–7 show the four types of weights presented above. The TI team's task was to assign weights to these different weighting approaches. Ideally, for a region with ample empirical data coverage in terms of magnitude, distance, and bandwidth, higher importance would be assigned to weights based on goodness-of-fit to data. However, this is not the case for the NGA-East database. As shown in Figure 9–3, the magnitude/distance distribution and limited bandwidth of the data used for calculating the weights is far from ideal. Thus, both w_k(res) and w_k(L) reflect an assessment of the models only over this limited range of magnitude, distance, and frequency. An objection was also raised that using recorded data to define weights may be double-counting the impact of data because, first, the recorded data are used to develop the seed GMMs and then to weight the final GMMs.

This objection was taken into consideration by the TI team as most GMM developers used some record-based constraints to scale the magnitude and distance behavior of their model. However, the various teams who developed candidate GMMs all used different subsets of data and different site effects models to correct those data to the reference-rock condition of V_{S30} = 3000 m/sec and κ = 0.006 sec. There is considerable epistemic uncertainty associated with this correction, which is not captured in our approach. Technically, the TI team could have used more than one site correction method, but then there would have been multiple datasets for calculating weights, which would mean that the site correction methods would have to be weighted as well. It would be possible to assess the uncertainty in V_{S30} based on the information provided in the NGA-East database; however, the uncertainty to the site response itself, given V_{S30} is more difficult to quantify, especially in CENA where large impedance contrasts are often present, thus rendering the use of V_{S30} inappropriate. Based on these considerations, the TI team decided to give a low weight to the data based approaches, and assign the bulk of the weight based on the $w_k(N_{Samples})$. The final weights are:

- Within 1-10 Hz bandwidth:
  - 80% for $w_k(N_{Samples})$, 10 % for w_k(res), and 10% for w_k (L).
- Everywhere else:
  - 100% to $w_k(N_{Samples})$

In addition, because of the limited bandwidth for which reliable data are available, the TI team felt that assigning a total of 20% weight based on fit-to-data was reasonable: this weight gives more influence to those models that behaved consistently with the data yet does not skew weights too much. Figure 9–11 shows the total weights for all the GMIMs separately for each frequency. Figure 9–12 shows the total weights for each model versus frequency. The two figures show the same information in different formats. There are variations in weights for a given cell with frequency, but the variations with frequency are relatively smooth and stable. The total weights for each GMIM are provided in Table 9–2. We provide results from hazard sensitivity analyses to different weighting schemes in Appendix E.8.

9.5 Evaluation and Discussion of Weighted Ground-Motion Models

In this section, we present our discussions and the evaluation of the final models organized around suites of plots.

9.5.1 Cumulative Ground-Motion Distributions

Figure 9–13 shows plots of the cumulative distribution function of the NGA-East final GMMs, based on the total weights. These are calculated in the following way: for a (**M**, R_{RUP})-scenario, the predictions of the 17 NGA-East final GMMs are sorted, and for each subsequent prediction the cumulative weight of all those GMMs—given a prediction that is lower—is calculated. The cumulative distribution of the NGA-East GMMs is compared with that of the seed GMMs based on the initial weights used in the sampling process (see Section 8.3). The cumulative distribution functions are shown here for two scenarios; the general trend indicates that the NGA-East GMMs achieve the goal of broadening the distribution of median predictions at large magnitudes where uncertainty is largest. Comparisons for more (**M**, R_{RUP}) scenarios and frequencies are included in Appendix E.9.

9.5.2 Fractiles of Median Ground Motions for Suites of Scenarios

We computed fractiles from the GMM predictions starting from the cumulative distributions described above. Figure 9–14 shows a suite of 10^{th}, 50^{th}, and 90^{th} fractiles GMM predictions against magnitude, plotted for binned distances along with site-corrected observed ground-motion data. The top plots are for distances of 50–100 km with the predictions computed at 70 km and the bottom plots are for distances 150–250 km with the predictions computed at 200 km. In the range where there are ground-motion observations, the distribution of the observations is wider than the fractiles, which is expected since we only show median models (no aleatory variability included). For large magnitudes and large distances, the range of the fractiles increases, which is desired, since there are no data constraints and thus the epistemic uncertainty should increase.

For a few scenarios, the range spanned by the NGA-East GMMs is narrower than that from the seeds. This is visible in Figure 9–13 (top) at the lower range of ground motions and in some of the figures compiled in Appendix E.9.1 (such as Figures 5196, 5211, and 5221). This observation only applies to a limited set of scenarios, but it is nonetheless important to address the issue and to make sure these results do not inadvertently affect the final hazard predictions. Two main factors are at play to cause these trends.

The first factor is due to the project team's decision to deliver smooth models with magnitude, distance and frequency. The smoothing of models across those three variables (**M**, R_{RUP} and *f*) all at once caused some narrowing of the range for certain selected scenarios. The changes due to smoothing are most pronounced in the case of smoothing across frequency. As with most GMMs developed to date, the seed and the NGA-East GMMs were developed one frequency at-a-time using different modeling processes. The NGA-East GMMs were smoothed through a curve-fitting process (Section 8.6) that effectively removed large deviations from the expected spectral shape trends in the models, causing a slight narrowing of the ground-motion range in limited cases, relative to the unsmoothed case.

The second factor contributing to a narrower range is applicable to very close-in distances (mostly up to 10 km). This apparent narrowing of the range is partly due to the conversion of the distance metric from R_{JB} to R_{RUP} for a suite of seeds (Section 7.7.2). In addition, several models, including those requiring the R_{RUP} conversion, had to be extrapolated to zero R_{RUP} distance (Section 7.7.3.) for numerical reasons (so that hazard codes could run the models). This was the case for most of the seed models, and it especially affected at small magnitudes. Those extrapolations were reviewed and approved by the seed developers, but the TI team was fully aware that these extrapolations were only approximations, with some of them causing undesired artifacts (see Figures 7–64 to 7–76). In some cases, the close-distance conversion and extrapolations led to narrow magnitude scaling for some seeds (e.g., B_a04 in Figure 7–65) and to wide magnitude scaling for others (e.g., 1CCSP in Figure 7–66). These issues were addressed through the TI team process for developing median models, notably by the inclusion of physicality rules and through the averaging approach described in Sections 8.3 and 8.5. The NGA-East GMMs do not exhibit the close-in artifacts (see scenario figures in Appendix E.9.2, Figures 851–1125.).

9.5.3 Hazard Sensitivity Results

To further investigate the issue described above, and to evaluate whether the conversion and extrapolation led to a narrowing of the ground motions range, we completed hazard sensitivity studies. For these analyses in hazard space, the seeds were used as provided in the PEER report (2015a) appendices (before they were converted and extrapolated by the TI team for use in the project and as document in PEER report 2015b). We used the as-provided distance metric (R_{JB} or R_{RUP}) for the seeds and computed the hazard using the appropriate distance metric. At distances less that the minimum distance supplied, the ground motions were not assumed to increase. If the distances supplied did not extend to 1000 km, the data were extrapolated using the log(R) slope fit to the last 100 km of values provided. These assumptions were required, once again, to allow the hazard codes to run. In other words, if the original seed used R_{JB}, the hazard was computed using R_{JB} distances.

Plots on Figure 9–15 to 9–17 show hazard sensitivity results computed at three sites using those original seeds with their initial weights (Section 8.3.1). Hazard results are presented for 1 Hz and 10 Hz for the Manchester site (Figure 9–15), for Central Illinois (Figure 9–16), and for Savannah (Figure 9–17). Parts (a) and (b) of these three figures show that, overall, the range of ground motions from the NGA-East GMMs (defined as the width spanned by the 5^{th} and 95^{th} percentile curves) is wider than that from the seed GMMs, especially at the mean Annual

Frequency of Exceedance (AFE) of interest (10^{-4} to 10^{-5}). Part (c) of the three figures show ratios of mean AFE from the seed GMMs relative to the final NGA-East GMMs. Mean AFE ratios of one show that the mean hazard is the same for a given ground-motion level, ratios less than one correspond to a seed hazard that is lower, and ratios above one correspond to a seed hazard that is larger, all relative to the final NGA-East GMMs. Results show the same global trends for all three sites: the ratios tend to be significantly smaller than one for the RLME sources, sometimes slightly higher than one for the distributed source and usually lower than or close to one for all the sources combined.

The TI team concluded that, overall, the objective of expanding the epistemic uncertainty from the seeds was achieved for most of the model application range, and that even though some scenarios may provide a narrower epistemic range over limited bands of magnitude and distance, these effects do no inappropriately reduce the epistemic range in hazard at the AFEs of interest.

9.5.4 Weighted Mean Spectra for a Suite of Scenarios

We have generated weighted spectra plots that combined the 17 final NGA-East GMMs using their respective weights into a single spectrum. Figures 9–18 shows examples of such plots for two selected scenarios. The top figure is representative of the typical smooth appearance of the weighted spectra. At larger distances however (Figures 9–18, bottom), the spectrum takes a more jagged appearance for frequencies beyond 10 Hz. When PSA values are less than PGA and the frequency is larger than 10 Hz, PSA should be set equal to PGA to be consistent with the definition of pseudo-absolute acceleration and to remove this artifact for application. We have observed this trend only for large distances ($\geq$ 800 km) for all the magnitudes. At these large distances, the small inconsistencies at high frequencies have no hazard significance. Weighted spectra plots have been generated for all the scenarios and are provided in Appendix E.9.1.

9.5.5 Achieved Variance for All Scenarios

Figure 9–19 and 9–20 show the seed variance (left) and achieved weighted variance (right) of the NGA-East final GMMs for all magnitudes and distances. Figure 9–19 shows results using a linear distance scale and Figure 9–20 shows the same results using a logarithmic distance scale to improve readability. The final achieved variances shown in those figures satisfied the TI team that the target variance model developed in Section 8.1 was appropriate. The final variances show that we achieved the desired objective of representing the very significant epistemic uncertainty that exists at large magnitudes while preserving the epistemic uncertainty from the seed GMMs at large distances. Figure 9–20 shows the impact of distance metric conversion and extrapolation discussed above: there is larger variance in the seed GMMs at very close distances that the NGA-East GMMs do not replicate, as intended and discussed above. The nonlinearity involved in capturing all the scenarios at once in the Sammon's maps introduced some expected variations in magnitude and distance space relative to the target model, as shown in Figures 9–19 and 9–20. Nevertheless, the final suite of median GMMs combined together with their weights reproduces the most important general trends from both

the seeds (variability at large distances) and the target variance model (variability at large magnitude).

9.6 References

Abrahamson, N.A., and R.R. Youngs. 1992. “A Stable Algorithm for Regression Analysis using the Random Effects Model.” Bull. Seismol. Soc. Am., 82(1): 505–510.

Boore, D.M. (2015). “Adjusting Ground-Motion Intensity Measures to a Reference Site for which V_{S30} = 3000 m/sec.” PEER Report No. 2015/06, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Dreiling, J., M.P. Isken, W.D. Mooney, M.C. Chapman, and R.W. Godbee. 2014. “NGA-East Regionalization Report: Comparison of Four Crustal Regions within Central and Eastern North America using Waveform Modeling and 5%-Damped Pseudo-Spectral Acceleration Response.” PEER Report No. 2014/15, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

GeoPentech. 2015. “Southwestern United States Ground Motion Characterization SSHAC Level 3 – Technical Report Rev.2.” March 2015.

Goulet, C.A., T. Kishida. T.D. Ancheta, C.H. Cramer, R.B. Darragh, W.J. Silva, Y.M.A. Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. 2014. “PEER NGA-East Database.” PEER Report No. 2014/17, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

PEER. 2015a. “NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

PEER. 2015b. NGA-East: Adjustments Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/08, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Sammon, J.W. 1969. “A Nonlinear Mapping for Data Structure Analysis.” IEEE Trans. Computers, C-18: 401–409.

Table 9–1(a) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. 2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
43	6.79E-03	2.36E-02	3.70E-02	6.03E-02	6.99E-02	7.73E-02	1.02E-01	1.57E-01	1.71E-01
49	2.19E-02	4.35E-02	7.99E-02	8.34E-02	8.98E-02	1.58E-01	1.64E-01	1.66E-01	1.57E-01
51	1.65E-02	2.74E-02	5.94E-02	8.86E-02	1.88E-01	1.66E-01	1.39E-01	1.97E-01	1.70E-01
52	1.61E-02	1.70E-02	2.68E-02	4.23E-02	7.95E-02	8.04E-02	9.36E-02	1.03E-01	9.18E-02
53	2.41E-02	3.77E-02	6.54E-02	7.83E-02	6.67E-02	8.04E-02	7.40E-02	1.22E-01	1.09E-01
54	3.48E-03	7.24E-03	2.09E-02	3.46E-02	4.07E-02	4.87E-02	8.60E-02	7.52E-02	1.23E-01
56	5.29E-03	1.18E-02	3.41E-02	3.98E-02	8.19E-02	9.88E-02	1.81E-01	2.63E-01	2.66E-01
58	2.12E-02	3.90E-02	4.18E-02	5.35E-02	8.56E-02	9.19E-02	1.13E-01	1.81E-01	8.84E-02
98	NA	NA	NA	NA	NA	7.57E-04	1.27E-03	1.75E-03	2.17E-03
99	NA	NA	NA	1.69E-03	2.56E-03	3.84E-03	5.65E-03	7.52E-03	2.06E-02
100	NA	NA	7.72E-03	1.11E-02	1.60E-02	2.48E-02	4.47E-02	5.57E-02	4.28E-02
101	NA	NA	NA	NA	8.44E-04	7.56E-04	1.05E-03	1.35E-03	1.92E-03
102	NA	NA	NA	NA	NA	1.37E-02	1.88E-02	2.99E-02	7.16E-02
103	NA	7.24E-04	1.52E-03	2.69E-03	4.88E-03	6.85E-03	9.64E-03	1.74E-02	2.94E-02
125	NA	2.66E-04	6.79E-04	8.48E-04	1.46E-03	2.76E-03	2.12E-03	2.99E-03	4.31E-03
126	NA	NA	1.63E-03	2.70E-03	4.48E-03	5.87E-03	5.70E-03	5.02E-03	4.87E-03
127	NA	NA	2.18E-03	4.79E-03	1.00E-02	1.49E-02	1.24E-02	1.14E-02	1.43E-02
129	NA	NA	1.99E-03	2.80E-03	3.79E-03	2.75E-03	2.72E-03	3.45E-03	5.28E-03
130	NA	2.27E-04	5.44E-04	8.43E-04	1.85E-03	2.25E-03	1.77E-03	2.07E-03	2.94E-03
136	4.58E-05	6.50E-05	2.10E-04	4.55E-04	4.80E-04	4.64E-04	6.36E-04	6.95E-04	5.99E-04
147	NA	1.71E-03	3.24E-03	5.11E-03	1.28E-02	2.02E-02	3.17E-02	3.01E-02	4.18E-02
150	NA	NA	NA	NA	1.79E-03	3.45E-03	3.94E-03	1.23E-02	2.32E-02
181	NA	NA	1.94E-04	1.57E-04	2.86E-04	2.81E-04	3.89E-04	4.02E-04	3.26E-04
182	NA	NA	NA	2.53E-03	3.62E-03	6.66E-03	6.38E-03	6.99E-03	2.76E-03
183	NA	NA	8.67E-05	1.05E-04	1.27E-04	1.77E-04	2.65E-04	2.80E-04	2.48E-04
184	NA	NA	NA	NA	2.70E-04	3.54E-04	6.40E-04	8.03E-04	7.71E-04
185	NA	3.09E-04	1.78E-04	2.54E-04	3.31E-04	4.38E-04	5.65E-04	8.24E-04	6.72E-04
201	6.39E-05	1.14E-04	1.47E-04	2.47E-04	2.86E-04	3.86E-04	5.43E-04	6.02E-04	NA
235	5.65E-05	1.05E-04	2.22E-04	2.61E-04	5.10E-04	9.02E-04	6.61E-04	5.96E-04	5.41E-04
240	9.95E-05	2.25E-04	2.96E-04	4.03E-04	6.87E-04	5.78E-04	5.64E-04	4.94E-04	4.45E-04
241	4.65E-05	6.99E-05	3.61E-04	3.21E-04	4.76E-04	4.94E-04	5.39E-04	9.02E-04	6.60E-04
249	1.80E-04	2.94E-04	4.94E-04	7.78E-04	1.24E-03	1.66E-03	1.67E-03	1.97E-03	2.68E-03
399	4.97E-05	1.47E-04	2.35E-04	3.49E-04	2.56E-04	2.29E-04	NA	NA	NA
537	9.33E-04	1.13E-03	1.16E-03	1.84E-03	2.39E-03	3.04E-03	2.03E-03	3.20E-03	1.87E-03
546	4.76E-04	6.73E-04	6.93E-04	1.17E-03	2.45E-03	3.70E-03	7.21E-03	4.96E-03	5.13E-03
550	2.76E-04	3.61E-04	4.79E-04	6.65E-04	1.24E-03	1.87E-03	2.21E-03	2.66E-03	4.74E-03
552	NA	NA	NA	NA	4.11E-03	3.94E-03	5.18E-03	4.42E-03	4.76E-03
555	NA	NA	NA	NA	NA	NA	NA	6.68E-03	6.78E-03
556	NA	NA	NA	NA	4.99E-03	1.07E-02	3.58E-02	3.01E-02	1.48E-02
572	2.27E-04	4.25E-04	7.86E-04	1.23E-03	1.74E-03	3.25E-03	2.96E-03	2.13E-03	2.85E-03
574	1.98E-03	4.54E-03	1.68E-02	1.01E-02	1.02E-02	1.51E-02	1.59E-02	1.05E-02	1.33E-02
581	1.81E-04	3.16E-04	8.89E-04	1.51E-03	3.03E-03	3.79E-03	3.84E-03	5.89E-03	1.01E-02
646	1.55E-03	2.33E-03	6.71E-03	1.35E-02	1.19E-02	1.08E-02	1.49E-02	3.11E-02	2.50E-02
741	2.43E-04	4.07E-04	6.20E-04	6.94E-04	7.50E-04	1.06E-03	1.76E-03	2.74E-03	NA
751	2.30E-04	4.35E-04	1.01E-03	2.03E-03	2.64E-03	1.74E-03	2.08E-03	1.71E-03	NA
754	1.62E-04	2.67E-04	5.76E-04	1.17E-03	1.41E-03	2.31E-03	2.62E-03	2.73E-03	NA
755	8.70E-05	1.93E-04	3.79E-04	3.15E-04	3.18E-04	3.18E-04	3.70E-04	4.48E-04	NA
801	9.90E-05	2.62E-04	4.64E-04	6.78E-04	1.48E-03	1.48E-03	1.41E-03	1.37E-03	NA
809	NA	NA	3.70E-04	7.17E-04	1.71E-03	2.70E-03	8.25E-03	5.84E-03	4.52E-03
1479	2.96E-05	5.53E-05	1.05E-04	1.30E-04	1.99E-04	2.38E-04	2.17E-04	1.48E-04	NA
1480	NA	1.47E-04	4.98E-04	6.10E-04	1.69E-03	1.01E-03	8.56E-04	8.44E-04	9.95E-04
1498	NA	NA	NA	3.90E-04	8.80E-04	1.41E-03	2.82E-03	3.79E-03	1.59E-03
1505	NA	5.73E-05	1.00E-04	1.93E-04	2.30E-04	3.25E-04	3.77E-04	4.71E-04	9.06E-04
1525	8.56E-05	1.61E-04	4.49E-04	9.04E-04	8.17E-04	9.08E-04	9.27E-04	1.33E-03	9.98E-04

Table 9–1(b) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 Hz	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
1679	4.34E-04	8.24E-04	2.39E-03	3.50E-03	6.30E-03	8.52E-03	9.60E-03	2.25E-02	2.20E-02
1680	NA	NA	NA	3.41E-03	6.26E-03	8.49E-03	9.49E-03	2.21E-02	2.17E-02
1681	8.57E-04	1.60E-03	4.31E-03	7.07E-03	1.17E-02	1.68E-02	2.51E-02	3.57E-02	5.02E-02
1682	NA	NA	NA	NA	NA	1.60E-02	2.49E-02	3.56E-02	5.06E-02
1683	NA	NA	7.92E-02	1.31E-01	1.86E-01	2.13E-01	2.90E-01	3.83E-01	2.35E-01
1684	8.69E-04	1.61E-03	3.47E-03	6.87E-03	1.51E-02	1.59E-02	2.54E-02	1.44E-02	2.26E-02
1685	NA	NA	NA	NA	NA	1.47E-02	2.42E-02	1.44E-02	2.21E-02
1686	NA	5.19E-03	1.02E-02	1.81E-02	3.62E-02	4.30E-02	3.91E-02	5.98E-02	8.59E-02
1687	NA	NA	NA	NA	3.32E-02	4.47E-02	4.60E-02	6.58E-02	8.66E-02
1688	2.80E-03	4.60E-03	4.52E-03	6.01E-03	9.60E-03	1.33E-02	1.60E-02	2.28E-02	3.24E-02
1689	NA	NA	NA	NA	9.49E-03	1.33E-02	1.59E-02	2.29E-02	3.29E-02
1715	6.74E-05	1.21E-04	2.07E-04	3.31E-04	5.39E-04	5.91E-04	6.94E-04	1.05E-03	1.10E-03
1721	NA	NA	NA	6.69E-03	1.19E-02	1.56E-02	2.55E-02	2.10E-02	3.04E-02
1723	8.75E-04	1.39E-03	3.09E-03	4.69E-03	9.10E-03	1.37E-02	1.89E-02	2.89E-02	2.38E-02
1724	NA	NA	3.08E-03	4.88E-03	9.57E-03	1.42E-02	2.11E-02	3.15E-02	3.38E-02
1728	8.50E-05	1.95E-04	3.62E-04	5.04E-04	7.68E-04	7.70E-04	1.18E-03	9.74E-04	1.04E-03
1731	1.05E-04	2.32E-04	4.72E-04	5.53E-04	8.76E-04	9.67E-04	9.13E-04	9.46E-04	7.96E-04
1745	NA	1.36E-03	2.19E-03	2.98E-03	5.55E-03	6.68E-03	6.67E-03	7.50E-03	1.03E-02
1746	NA	NA	NA	NA	2.90E-02	3.97E-02	5.11E-02	5.69E-02	9.55E-02
1747	NA	NA	4.69E-03	7.97E-03	1.26E-02	1.61E-02	3.13E-02	3.19E-02	4.98E-02
1752	NA	NA	NA	NA	5.84E-03	7.57E-03	1.35E-02	2.38E-02	3.18E-02
1759	NA	NA	1.68E-02	3.21E-02	6.16E-02	7.73E-02	1.38E-01	3.44E-01	2.69E-01
1761	NA	NA	NA	NA	NA	2.48E-03	2.81E-03	4.55E-03	1.15E-02
1771	1.62E-03	2.71E-03	4.63E-03	8.60E-03	1.21E-02	1.73E-02	2.72E-02	6.10E-02	4.44E-02
1772	NA	3.22E-03	8.83E-03	1.21E-02	2.50E-02	4.48E-02	7.74E-02	1.69E-01	1.97E-01
2569	3.46E-04	1.09E-03	2.97E-03	2.37E-03	1.36E-03	1.65E-03	1.55E-03	1.26E-03	NA
2576	3.97E-03	7.56E-03	5.76E-03	4.54E-03	5.10E-03	6.41E-03	8.08E-03	7.00E-03	NA
2577	4.80E-04	6.30E-04	2.09E-03	1.67E-03	1.78E-03	1.98E-03	3.27E-03	3.73E-03	NA
2610	7.33E-04	1.74E-03	3.86E-03	6.55E-03	1.16E-02	1.30E-02	1.97E-02	2.45E-02	1.61E-02
2614	9.22E-04	1.90E-03	5.14E-03	1.68E-02	3.87E-02	2.40E-02	1.69E-02	1.92E-02	1.93E-02
2615	6.53E-04	1.20E-03	3.64E-03	5.11E-03	1.08E-02	1.37E-02	2.10E-02	1.94E-02	9.59E-03
2621	1.73E-03	3.54E-03	1.16E-02	2.38E-02	4.15E-02	5.48E-02	4.98E-02	4.02E-02	2.22E-02
2622	1.49E-03	3.52E-03	1.24E-02	1.03E-02	1.09E-02	2.20E-02	2.88E-02	1.72E-02	2.29E-02
2623	4.12E-04	5.74E-04	1.23E-03	1.66E-03	2.59E-03	2.44E-03	1.97E-03	2.99E-03	NA
2628	1.73E-03	3.02E-03	1.04E-02	8.65E-03	1.53E-02	1.96E-02	2.07E-02	2.52E-02	3.86E-02
2634	9.29E-04	1.93E-03	1.88E-03	2.91E-03	3.86E-03	5.01E-03	6.78E-03	1.23E-02	1.74E-02
2643	1.78E-03	5.31E-03	1.10E-02	2.23E-02	4.81E-02	4.59E-02	5.61E-02	3.87E-02	6.37E-02
2648	1.90E-03	2.79E-03	4.85E-03	3.61E-03	5.48E-03	5.95E-03	5.85E-03	6.72E-03	5.09E-03
2649	3.72E-04	4.55E-04	8.30E-04	2.49E-03	2.95E-03	3.04E-03	3.02E-03	2.00E-03	NA
2651	1.02E-03	1.96E-03	2.81E-03	3.41E-03	7.85E-03	1.30E-02	2.34E-02	2.84E-02	1.43E-02
2654	7.46E-04	2.18E-03	8.08E-03	7.33E-03	8.90E-03	1.37E-02	1.04E-02	1.15E-02	7.32E-03
2655	2.38E-04	4.58E-04	1.15E-03	1.28E-03	1.90E-03	3.07E-03	4.18E-03	5.84E-03	3.66E-03
2656	2.34E-04	4.34E-04	1.07E-03	1.20E-03	1.78E-03	2.74E-03	3.78E-03	5.14E-03	3.08E-03
2660	1.60E-03	2.99E-03	6.70E-03	8.70E-03	9.50E-03	1.44E-02	1.89E-02	3.42E-02	5.84E-02
2661	NA	1.92E-02	4.58E-02	7.09E-02	9.29E-02	7.30E-02	1.39E-01	3.12E-01	5.16E-01
2715	1.20E-04	2.27E-04	4.10E-04	5.65E-04	8.54E-04	9.16E-04	9.00E-04	7.33E-04	NA
2719	3.27E-04	5.16E-04	1.12E-03	9.06E-04	1.38E-03	1.47E-03	1.63E-03	1.14E-03	NA
2720	1.34E-04	1.70E-04	1.95E-04	1.93E-04	4.06E-04	3.83E-04	6.45E-04	9.48E-04	NA
2723	2.96E-04	4.74E-04	8.47E-04	1.27E-03	1.66E-03	1.91E-03	1.82E-03	NA	NA
2724	1.95E-04	3.73E-04	1.12E-03	2.99E-03	7.44E-03	9.71E-03	7.98E-03	1.05E-02	8.04E-03
2726	5.24E-04	1.03E-03	2.82E-03	6.83E-03	3.01E-02	2.06E-02	1.44E-02	1.05E-02	1.43E-02
2727	NA	4.89E-04	1.30E-03	1.64E-03	4.97E-03	9.12E-03	2.23E-02	1.03E-02	6.47E-03
2730	6.01E-04	1.02E-03	3.35E-03	3.25E-03	9.48E-03	1.82E-02	2.62E-02	2.17E-02	1.09E-02

Table 9–1(c) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
2731	6.98E-04	1.27E-03	4.13E-03	7.61E-03	3.34E-03	5.53E-03	7.90E-03	7.97E-03	8.88E-03
2732	8.76E-05	1.67E-04	4.09E-04	9.11E-04	1.48E-03	2.27E-03	1.67E-03	1.88E-03	NA
2737	5.89E-04	7.42E-04	1.67E-03	1.97E-03	3.23E-03	3.98E-03	4.82E-03	7.61E-03	1.21E-02
2742	2.46E-04	4.42E-04	8.76E-04	1.52E-03	3.38E-03	3.74E-03	3.82E-03	4.67E-03	1.36E-02
2749	1.93E-03	3.16E-03	6.52E-03	7.84E-03	2.16E-02	3.32E-02	4.04E-02	5.02E-02	4.27E-02
2750	1.94E-03	3.21E-03	6.62E-03	7.99E-03	2.17E-02	3.24E-02	4.07E-02	5.21E-02	NA
2755	5.20E-04	7.49E-04	1.31E-03	1.18E-03	1.59E-03	2.80E-03	2.42E-03	2.45E-03	2.20E-03
2756	5.28E-05	1.03E-04	1.80E-04	2.89E-04	4.71E-04	7.29E-04	7.78E-04	6.43E-04	NA
2761	4.38E-04	5.44E-04	1.76E-03	2.49E-03	4.80E-03	9.30E-03	7.92E-03	7.30E-03	5.34E-03
2762	NA	NA	NA	NA	NA	9.08E-03	7.77E-03	6.92E-03	4.80E-03
2763	7.94E-05	1.28E-04	2.44E-04	3.65E-04	7.51E-04	9.97E-04	9.38E-04	1.31E-03	NA
2764	NA	NA	2.45E-04	3.98E-04	8.00E-04	1.09E-03	1.28E-03	1.74E-03	1.21E-03
2767	7.03E-04	9.63E-04	1.69E-03	1.94E-03	3.39E-03	3.83E-03	7.15E-03	1.09E-02	1.86E-02
2769	5.47E-03	1.03E-02	1.97E-02	2.52E-02	3.40E-02	5.49E-02	8.64E-02	2.34E-01	2.51E-01
2787	1.16E-03	1.05E-03	1.41E-03	2.39E-03	6.16E-03	5.38E-03	6.24E-03	6.37E-03	5.06E-03
2805	2.53E-05	7.33E-05	1.73E-04	1.72E-04	2.04E-04	2.10E-04	2.08E-04	1.45E-04	NA
2811	1.65E-04	4.08E-04	8.12E-04	8.29E-04	1.49E-03	1.57E-03	1.72E-03	1.22E-03	NA
2812	2.96E-05	4.70E-05	1.45E-04	1.48E-04	2.27E-04	2.98E-04	3.69E-04	7.50E-04	NA
2816	6.01E-05	1.49E-04	2.94E-04	4.45E-04	6.77E-04	6.58E-04	7.56E-04	4.92E-04	NA
2817	1.08E-04	2.18E-04	7.71E-04	8.69E-04	2.25E-03	2.58E-03	3.83E-03	3.52E-03	2.56E-03
2819	1.85E-04	3.82E-04	1.07E-03	2.78E-03	9.10E-03	5.70E-03	4.30E-03	3.96E-03	4.81E-03
2820	NA	NA	4.18E-04	9.35E-04	2.14E-03	2.25E-03	4.72E-03	4.53E-03	2.35E-03
2824	NA	3.45E-04	9.53E-04	1.44E-03	2.69E-03	2.98E-03	6.03E-03	4.20E-03	2.97E-03
2825	1.44E-04	4.15E-04	1.35E-03	1.36E-03	1.44E-03	2.28E-03	3.37E-03	3.73E-03	6.22E-03
2829	2.26E-04	4.30E-04	1.03E-03	1.62E-03	2.46E-03	4.14E-03	6.01E-03	8.92E-03	1.59E-02
2833	NA	NA	2.57E-04	3.97E-04	6.54E-04	9.31E-04	1.35E-03	1.83E-03	4.30E-03
2841	3.86E-04	7.17E-04	1.90E-03	4.95E-03	8.49E-03	8.87E-03	1.04E-02	1.25E-02	1.33E-02
2842	NA	NA	NA	4.93E-03	8.41E-03	8.77E-03	1.03E-02	1.20E-02	1.24E-02
2845	NA	1.97E-04	6.04E-04	4.45E-04	5.07E-04	8.68E-04	6.62E-04	7.29E-04	5.57E-04
2846	2.77E-05	3.82E-05	1.10E-04	2.21E-04	3.45E-04	4.46E-04	5.30E-04	3.91E-04	NA
2851	7.38E-05	1.58E-04	4.43E-04	7.91E-04	1.02E-03	1.68E-03	1.79E-03	2.18E-03	1.80E-03
2852	NA	NA	NA	NA	NA	1.62E-03	1.72E-03	2.09E-03	1.63E-03
2853	2.23E-05	4.37E-05	2.05E-04	1.75E-04	2.53E-04	3.32E-04	4.31E-04	NA	NA
2854	NA	NA	NA	1.73E-04	2.67E-04	3.60E-04	5.13E-04	5.54E-04	3.79E-04
2856	2.14E-04	3.73E-04	7.82E-04	1.13E-03	2.17E-03	3.49E-03	4.65E-03	9.66E-03	NA
2857	7.60E-04	1.45E-03	3.26E-03	4.32E-03	9.18E-03	1.21E-02	1.84E-02	4.22E-02	5.98E-02
2884	1.04E-04	1.47E-04	3.68E-04	5.28E-04	1.22E-03	1.06E-03	1.16E-03	1.27E-03	1.01E-03
3491	2.80E-04	3.76E-04	5.73E-04	8.01E-04	1.18E-03	1.31E-03	2.55E-03	2.74E-03	6.97E-03
3492	3.24E-04	3.96E-04	4.89E-04	6.18E-04	1.12E-03	2.00E-03	3.63E-03	3.79E-03	8.76E-03
3493	1.86E-04	2.48E-04	3.80E-04	6.11E-04	9.36E-04	1.23E-03	2.29E-03	5.60E-03	9.78E-03
3494	1.37E-04	1.79E-04	3.16E-04	5.91E-04	1.07E-03	1.15E-03	1.38E-03	2.13E-03	2.60E-03
3495	3.53E-04	4.50E-04	6.92E-04	1.00E-03	2.20E-03	3.61E-03	5.33E-03	9.11E-03	7.50E-03
3496	2.24E-04	2.49E-04	3.44E-04	7.14E-04	1.16E-03	1.47E-03	1.84E-03	3.50E-03	3.76E-03
3531	4.24E-05	6.22E-05	1.25E-04	1.35E-04	1.72E-04	1.48E-04	1.61E-04	1.36E-04	1.28E-04
3532	2.75E-05	6.42E-05	9.69E-05	2.07E-04	2.53E-04	6.24E-04	3.62E-04	2.71E-04	2.06E-04
3533	1.53E-05	2.36E-05	3.43E-05	5.08E-05	8.96E-05	9.99E-05	7.57E-05	7.46E-05	5.29E-05
3688	4.82E-05	8.22E-05	9.57E-05	1.26E-04	1.69E-04	1.66E-04	1.42E-04	1.04E-04	8.02E-05
3689	3.59E-05	6.14E-05	9.56E-05	1.28E-04	1.47E-04	1.19E-04	1.00E-04	1.01E-04	7.47E-05
3690	2.70E-05	6.65E-05	7.40E-05	1.33E-04	2.25E-04	3.63E-04	3.25E-04	2.66E-04	1.85E-04
3700	7.42E-05	1.34E-04	2.66E-04	3.36E-04	4.51E-04	6.34E-04	1.14E-03	1.00E-03	5.13E-04
3708	8.35E-05	1.79E-04	1.38E-04	1.45E-04	1.93E-04	2.58E-04	1.91E-04	1.46E-04	1.08E-04
3709	4.08E-05	5.39E-05	3.31E-04	3.61E-04	2.99E-04	3.27E-04	4.66E-04	2.77E-04	2.75E-04
3710	1.11E-04	1.41E-04	1.91E-04	3.76E-04	5.12E-04	6.29E-04	5.44E-04	4.79E-04	4.76E-04
3712	3.75E-05	6.24E-05	1.29E-04	2.10E-04	4.42E-04	4.50E-04	5.83E-04	1.08E-03	1.14E-03
3720	7.31E-05	1.14E-04	1.39E-04	1.95E-04	2.45E-04	2.57E-04	2.58E-04	1.96E-04	2.37E-04

Table 9–1(d) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
3721	1.36E-04	3.83E-04	3.55E-04	3.71E-04	4.34E-04	4.53E-04	4.39E-04	NA	NA
3722	1.15E-04	1.37E-04	4.19E-04	6.03E-04	1.14E-03	1.62E-03	3.56E-03	3.05E-03	1.32E-03
3723	1.29E-04	1.48E-04	2.81E-04	3.65E-04	5.16E-04	8.52E-04	9.22E-04	1.19E-03	1.18E-03
3730	2.06E-05	3.89E-05	5.54E-05	9.84E-05	1.22E-04	1.29E-04	1.66E-04	1.61E-04	1.44E-04
3731	1.03E-04	2.06E-04	3.43E-04	7.43E-04	9.34E-04	6.92E-04	5.81E-04	3.87E-04	NA
3732	5.39E-05	8.96E-05	1.28E-04	2.35E-04	2.77E-04	3.86E-04	3.78E-04	6.26E-04	4.96E-04
3733	1.36E-04	2.06E-04	3.63E-04	6.56E-04	1.39E-03	2.08E-03	2.21E-03	1.65E-03	2.44E-03
3734	1.54E-04	2.88E-04	3.99E-04	5.59E-04	8.13E-04	1.44E-03	1.97E-03	1.91E-03	NA
3741	4.32E-05	9.37E-05	1.02E-04	1.52E-04	1.56E-04	2.00E-04	1.78E-04	1.57E-04	1.93E-04
3742	2.42E-05	5.17E-05	1.16E-04	1.55E-04	2.44E-04	4.35E-04	4.83E-04	3.25E-04	2.92E-04
3743	3.83E-05	8.35E-05	1.31E-04	2.56E-04	2.96E-04	4.01E-04	3.39E-04	3.35E-04	5.56E-04
3744	4.23E-05	1.57E-04	2.98E-04	3.62E-04	4.92E-04	5.84E-04	5.08E-04	NA	NA
3745	1.56E-04	2.04E-04	3.39E-04	4.91E-04	6.07E-04	6.14E-04	6.31E-04	NA	NA
3753	3.20E-05	4.16E-05	6.15E-05	8.48E-05	8.24E-05	9.29E-05	7.23E-05	8.40E-05	8.99E-05
3754	3.81E-05	6.53E-05	8.46E-05	8.24E-05	1.21E-04	1.47E-04	1.84E-04	NA	NA
3755	5.59E-05	1.60E-04	2.73E-04	4.31E-04	8.07E-04	8.44E-04	1.04E-03	8.87E-04	NA
3756	1.53E-04	3.70E-04	6.06E-04	8.25E-04	1.05E-03	1.31E-03	1.66E-03	2.04E-03	1.37E-03
3757	1.94E-04	2.48E-04	4.59E-04	5.62E-04	7.49E-04	8.53E-04	9.28E-04	5.76E-04	6.54E-04
3765	2.49E-05	4.65E-05	1.14E-04	1.55E-04	1.36E-04	1.59E-04	1.68E-04	2.01E-04	NA
3767	2.01E-04	2.92E-04	7.89E-04	9.99E-04	1.50E-03	1.15E-03	2.00E-03	1.09E-03	7.03E-04
3768	1.17E-04	3.00E-04	8.15E-04	9.44E-04	1.37E-03	1.50E-03	1.29E-03	1.09E-03	8.35E-04
3776	2.33E-05	5.24E-05	7.33E-05	1.33E-04	1.89E-04	2.27E-04	3.14E-04	3.43E-04	NA
3777	4.40E-05	1.05E-04	2.02E-04	1.74E-04	2.81E-04	3.11E-04	5.59E-04	3.14E-04	NA
3778	5.14E-05	8.05E-05	2.28E-04	2.80E-04	5.39E-04	5.45E-04	5.16E-04	4.60E-04	2.51E-04
3779	7.92E-05	1.97E-04	3.26E-04	4.26E-04	4.55E-04	4.06E-04	4.21E-04	5.17E-04	4.38E-04
3780	2.96E-05	5.14E-05	7.90E-05	1.09E-04	2.33E-04	2.72E-04	2.10E-04	2.57E-04	3.59E-04
3788	2.66E-05	2.79E-05	5.63E-05	1.05E-04	1.52E-04	1.24E-04	9.81E-05	NA	NA
3791	2.09E-05	3.22E-05	4.30E-05	6.05E-05	1.50E-04	1.35E-04	1.67E-04	1.64E-04	NA
3982	2.23E-03	3.79E-03	1.36E-02	1.66E-02	9.58E-03	1.29E-02	1.27E-02	1.16E-02	1.64E-02
3984	1.76E-03	2.62E-03	3.66E-03	4.28E-03	5.42E-03	6.36E-03	1.06E-02	1.26E-02	1.75E-02
3987	1.35E-03	2.35E-03	4.83E-03	6.72E-03	5.00E-03	6.64E-03	4.86E-03	3.81E-03	3.49E-03
3990	1.48E-03	1.49E-03	2.32E-03	3.03E-03	3.24E-03	3.38E-03	4.76E-03	5.46E-03	5.40E-03
4000	1.02E-03	1.46E-03	6.12E-03	5.92E-03	1.03E-02	1.52E-02	1.52E-02	8.97E-03	9.42E-03
4011	1.44E-03	2.46E-03	3.43E-03	2.94E-03	5.15E-03	5.11E-03	6.45E-03	8.29E-03	6.98E-03
4013	6.87E-04	1.30E-03	2.36E-03	3.30E-03	3.03E-03	3.40E-03	3.31E-03	3.94E-03	3.25E-03
4014	4.56E-04	7.89E-04	1.46E-03	2.24E-03	3.87E-03	5.05E-03	7.28E-03	1.78E-02	1.28E-02
4017	2.27E-04	2.79E-04	6.75E-04	7.48E-04	1.44E-03	1.71E-03	1.84E-03	1.70E-03	1.02E-03
4019	7.11E-04	8.06E-04	1.41E-03	1.73E-03	2.15E-03	2.68E-03	2.41E-03	2.37E-03	2.19E-03
4021	NA	NA	NA	NA	4.50E-03	7.33E-03	9.97E-03	7.69E-03	1.31E-02
4022	NA	NA	NA	NA	3.01E-03	5.90E-03	7.07E-03	6.23E-03	5.28E-03
4023	NA	NA	NA	NA	NA	NA	NA	5.04E-03	7.02E-03
4027	2.77E-03	3.69E-03	8.84E-03	1.42E-02	2.64E-02	3.06E-02	2.82E-02	3.60E-02	3.42E-02
4028	4.48E-03	1.51E-02	9.78E-03	1.10E-02	2.24E-02	3.12E-02	1.85E-02	3.98E-02	NA
4031	NA	1.38E-02	3.73E-02	2.94E-02	2.86E-02	2.91E-02	3.63E-02	5.37E-02	5.79E-02
4032	5.85E-03	1.10E-02	1.97E-02	2.65E-02	3.47E-02	3.20E-02	5.07E-02	6.21E-02	6.91E-02
4033	5.53E-03	9.94E-03	1.46E-02	1.63E-02	2.11E-02	2.24E-02	4.12E-02	5.79E-02	6.08E-02
4034	NA	4.63E-02	6.70E-02	3.55E-02	5.02E-02	8.96E-02	5.92E-02	7.33E-02	5.45E-02
4037	NA	1.03E-02	1.94E-02	2.61E-02	3.39E-02	3.17E-02	5.01E-02	6.14E-02	6.87E-02
4038	1.50E-03	2.14E-03	3.48E-03	4.57E-03	6.57E-03	6.71E-03	7.12E-03	6.59E-03	5.53E-03
4039	5.97E-04	1.32E-03	3.22E-03	6.29E-03	6.71E-03	9.58E-03	1.23E-02	1.47E-02	1.71E-02
4041	1.69E-03	4.41E-03	1.80E-02	1.15E-02	8.66E-03	1.22E-02	1.39E-02	9.31E-03	7.74E-03
4043	9.13E-04	1.59E-03	2.61E-03	6.11E-03	6.76E-03	6.83E-03	9.70E-03	1.27E-02	1.48E-02
4046	2.21E-03	4.26E-03	6.68E-03	6.70E-03	7.02E-03	5.26E-03	7.19E-03	6.12E-03	5.61E-03
4058	4.79E-04	6.97E-04	1.11E-03	1.25E-03	1.04E-03	1.31E-03	1.46E-03	1.50E-03	1.24E-03

Table 9–1(e) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
4060	2.11E-03	2.08E-03	4.17E-03	3.35E-03	4.53E-03	3.85E-03	4.74E-03	6.56E-03	6.56E-03
4086	1.26E-04	2.26E-04	3.69E-04	4.91E-04	7.12E-04	5.88E-04	6.43E-04	6.70E-04	5.88E-04
4114	4.24E-04	8.61E-04	7.89E-04	1.10E-03	1.10E-03	1.54E-03	2.49E-03	1.96E-03	2.22E-03
4121	1.46E-04	2.75E-04	4.01E-04	2.97E-04	3.62E-04	4.60E-04	4.28E-04	3.44E-04	2.37E-04
4564	3.19E-04	5.18E-04	1.06E-03	1.19E-03	2.17E-03	2.19E-03	2.49E-03	2.72E-03	2.94E-03
4568	3.40E-04	5.51E-04	1.17E-03	1.80E-03	2.86E-03	1.96E-03	2.50E-03	2.21E-03	1.87E-03
4671	NA	NA	5.17E-04	1.16E-03	1.78E-03	2.28E-03	2.76E-03	4.11E-03	5.70E-03
4672	NA	NA	NA	4.79E-04	6.90E-04	1.38E-03	1.93E-03	1.40E-03	2.27E-03
4673	NA	NA	NA	1.39E-03	1.90E-03	2.76E-03	4.78E-03	7.86E-03	5.43E-03
4731	6.53E-04	1.38E-03	3.35E-03	4.82E-03	8.82E-03	1.03E-02	9.95E-03	8.92E-03	1.22E-02
4732	1.06E-03	2.04E-03	5.13E-03	7.84E-03	1.25E-02	1.59E-02	1.34E-02	1.10E-02	1.58E-02
4733	3.37E-04	6.96E-04	1.70E-03	2.65E-03	4.67E-03	5.29E-03	6.19E-03	6.16E-03	1.18E-02
4734	4.25E-04	1.01E-03	1.74E-03	2.06E-03	4.16E-03	5.29E-03	6.44E-03	8.39E-03	1.90E-02
4735	7.60E-04	1.50E-03	3.87E-03	5.76E-03	8.79E-03	1.16E-02	1.25E-02	7.98E-03	1.11E-02
4736	9.02E-04	1.77E-03	3.88E-03	6.27E-03	9.72E-03	1.11E-02	8.69E-03	1.05E-02	1.42E-02
4737	7.94E-04	1.72E-03	4.46E-03	7.68E-03	1.73E-02	2.39E-02	2.29E-02	1.67E-02	1.47E-02
4738	5.80E-04	1.22E-03	2.53E-03	3.90E-03	5.09E-03	6.13E-03	6.27E-03	7.95E-03	1.02E-02
4739	3.56E-04	7.60E-04	2.58E-03	5.17E-03	6.95E-03	7.42E-03	6.55E-03	6.51E-03	1.30E-02
4789	9.23E-05	1.38E-04	1.55E-04	1.88E-04	2.28E-04	2.55E-04	3.54E-04	2.72E-04	2.61E-04
4790	1.15E-04	1.89E-04	3.82E-04	6.01E-04	6.92E-04	7.77E-04	9.96E-04	1.10E-03	7.20E-04
4791	2.04E-04	3.48E-04	9.03E-04	1.16E-03	2.21E-03	2.47E-03	2.67E-03	2.65E-03	2.02E-03
4792	6.33E-05	1.07E-04	2.13E-04	4.17E-04	5.19E-04	8.31E-04	8.89E-04	9.13E-04	5.31E-04
4802	9.59E-05	2.61E-04	5.47E-04	5.71E-04	6.64E-04	6.56E-04	6.12E-04	8.78E-04	6.14E-04
4803	1.14E-04	1.30E-04	2.76E-04	4.03E-04	5.31E-04	7.63E-04	9.64E-04	7.93E-04	8.44E-04
4872	6.13E-05	1.21E-04	2.99E-04	4.27E-04	5.48E-04	6.42E-04	8.32E-04	1.20E-03	9.93E-04
5057	8.13E-05	1.62E-04	2.35E-04	3.80E-04	3.42E-04	2.77E-04	2.49E-04	2.27E-04	NA
5058	3.49E-05	5.60E-05	1.39E-04	1.38E-04	1.75E-04	2.15E-04	1.49E-04	1.61E-04	1.31E-04
5059	9.42E-05	1.93E-04	3.15E-04	3.71E-04	5.93E-04	8.67E-04	1.01E-03	9.13E-04	4.88E-04
5060	7.44E-05	1.57E-04	2.88E-04	4.65E-04	7.05E-04	4.99E-04	4.46E-04	4.51E-04	NA
5061	1.50E-04	2.78E-04	5.62E-04	7.28E-04	1.17E-03	1.08E-03	1.12E-03	1.36E-03	7.28E-04
5062	2.17E-04	4.43E-04	1.00E-03	9.67E-04	1.16E-03	1.31E-03	1.23E-03	7.65E-04	6.50E-04
5068	7.31E-05	1.44E-04	3.72E-04	3.51E-04	2.76E-04	2.45E-04	2.59E-04	2.21E-04	2.15E-04
5071	1.65E-04	2.91E-04	4.72E-04	4.42E-04	5.72E-04	6.43E-04	9.18E-04	1.13E-03	7.66E-04
5072	9.38E-05	2.14E-04	3.49E-04	4.55E-04	6.44E-04	7.48E-04	1.21E-03	1.31E-03	1.12E-03
5073	1.19E-04	2.18E-04	4.57E-04	6.88E-04	1.06E-03	1.49E-03	2.26E-03	1.31E-03	8.79E-04
5074	1.55E-04	2.45E-04	5.96E-04	6.90E-04	1.39E-03	2.00E-03	2.14E-03	1.43E-03	1.03E-03
5075	9.02E-05	2.54E-04	4.28E-04	6.86E-04	9.99E-04	8.25E-04	8.66E-04	6.65E-04	4.63E-04
5084	2.68E-04	4.02E-04	8.13E-04	6.58E-04	5.23E-04	5.10E-04	3.95E-04	NA	NA
5085	1.33E-04	2.94E-04	1.31E-03	1.43E-03	1.02E-03	1.03E-03	1.27E-03	9.78E-04	1.09E-03
5086	1.98E-04	4.45E-04	8.90E-04	9.49E-04	1.19E-03	1.26E-03	1.21E-03	9.81E-04	NA
5088	1.41E-04	2.11E-04	3.60E-04	7.33E-04	1.14E-03	1.47E-03	1.99E-03	2.84E-03	NA
5089	6.31E-04	1.41E-03	3.77E-03	6.45E-03	7.05E-03	7.71E-03	6.56E-03	3.59E-03	2.85E-03
5090	1.63E-04	4.45E-04	8.80E-04	1.20E-03	1.30E-03	2.10E-03	3.41E-03	2.61E-03	NA
5091	8.97E-05	1.86E-04	4.26E-04	5.41E-04	7.36E-04	9.51E-04	1.09E-03	1.38E-03	1.41E-03
5093	2.50E-04	4.37E-04	9.62E-04	2.07E-03	2.77E-03	3.69E-03	5.79E-03	4.83E-03	3.06E-03
5096	1.71E-04	2.06E-04	2.81E-04	3.68E-04	4.37E-04	3.84E-04	2.71E-04	NA	NA
5098	4.76E-04	6.21E-04	9.07E-04	9.97E-04	9.11E-04	8.42E-04	7.63E-04	6.93E-04	9.91E-04
5099	2.24E-04	5.25E-04	8.20E-04	1.13E-03	1.18E-03	1.16E-03	1.68E-03	2.62E-03	3.59E-03
5100	1.18E-04	2.42E-04	5.67E-04	1.23E-03	2.11E-03	3.24E-03	5.82E-03	6.69E-03	3.93E-03
5101	9.07E-05	1.96E-04	4.18E-04	9.37E-04	1.77E-03	1.70E-03	2.69E-03	3.62E-03	2.98E-03
5102	3.44E-04	9.18E-04	2.09E-03	2.93E-03	4.20E-03	5.49E-03	7.73E-03	6.21E-03	NA
5103	3.33E-04	4.26E-04	1.21E-03	1.63E-03	2.14E-03	3.32E-03	3.08E-03	2.61E-03	1.67E-03
5104	1.52E-04	2.57E-04	4.38E-04	6.77E-04	9.83E-04	1.76E-03	2.16E-03	2.90E-03	2.23E-03
5105	8.99E-05	1.83E-04	3.40E-04	4.62E-04	6.79E-04	1.03E-03	1.47E-03	1.70E-03	NA

Table 9–1(f) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
5108	1.64E-04	2.84E-04	3.86E-04	5.27E-04	4.21E-04	3.38E-04	2.79E-04	2.82E-04	2.98E-04
5109	7.21E-05	1.30E-04	2.55E-04	3.98E-04	4.21E-04	3.28E-04	4.05E-04	6.60E-04	5.79E-04
5110	4.59E-04	6.60E-04	1.00E-03	2.39E-03	2.87E-03	2.80E-03	2.10E-03	1.38E-03	9.81E-04
5111	4.31E-04	8.89E-04	1.58E-03	1.67E-03	1.46E-03	1.26E-03	1.38E-03	1.88E-03	2.14E-03
5112	2.36E-04	5.61E-04	9.81E-04	1.51E-03	3.45E-03	3.99E-03	4.82E-03	5.09E-03	8.49E-03
5113	1.43E-04	2.68E-04	4.38E-04	5.90E-04	8.68E-04	1.19E-03	1.24E-03	1.61E-03	2.50E-03
5114	4.58E-04	8.41E-04	1.90E-03	2.90E-03	4.71E-03	6.68E-03	7.32E-03	9.80E-03	NA
5115	4.44E-04	5.99E-04	1.60E-03	2.86E-03	2.84E-03	3.27E-03	4.45E-03	6.52E-03	6.00E-03
5116	8.15E-05	1.46E-04	2.18E-04	3.61E-04	6.28E-04	8.24E-04	1.33E-03	1.10E-03	6.99E-04
5117	7.84E-05	1.44E-04	4.44E-04	7.72E-04	1.13E-03	1.60E-03	1.41E-03	1.05E-03	1.15E-03
5119	7.49E-05	1.21E-04	1.05E-04	1.05E-04	1.46E-04	2.28E-04	1.80E-04	1.92E-04	1.39E-04
5120	2.39E-04	6.04E-04	6.20E-04	7.73E-04	7.69E-04	1.02E-03	7.71E-04	7.69E-04	8.57E-04
5121	1.96E-04	2.81E-04	6.96E-04	1.13E-03	1.48E-03	2.01E-03	2.17E-03	1.55E-03	1.47E-03
5122	1.16E-04	2.05E-04	4.54E-04	7.95E-04	1.82E-03	1.57E-03	1.30E-03	1.17E-03	NA
5123	1.19E-04	2.12E-04	3.77E-04	4.82E-04	6.92E-04	9.35E-04	1.16E-03	1.16E-03	1.27E-03
5124	6.66E-04	1.33E-03	2.74E-03	3.52E-03	3.59E-03	3.60E-03	3.61E-03	4.16E-03	7.04E-03
5125	6.38E-04	1.26E-03	3.36E-03	5.69E-03	9.32E-03	1.10E-02	9.96E-03	6.24E-03	NA
5126	3.47E-04	6.60E-04	1.43E-03	1.83E-03	2.82E-03	3.52E-03	4.08E-03	5.23E-03	NA
5127	2.12E-04	3.92E-04	9.96E-04	1.36E-03	2.70E-03	4.56E-03	4.51E-03	2.91E-03	NA
5128	1.89E-04	3.84E-04	3.58E-04	6.34E-04	7.06E-04	6.79E-04	8.51E-04	7.63E-04	8.11E-04
5129	7.46E-05	1.28E-04	2.29E-04	3.19E-04	3.47E-04	2.95E-04	3.64E-04	5.46E-04	NA
5132	4.32E-05	6.63E-05	1.11E-04	1.57E-04	2.35E-04	2.25E-04	1.80E-04	2.85E-04	NA
5133	7.25E-05	1.14E-04	2.65E-04	3.11E-04	3.63E-04	5.01E-04	7.83E-04	NA	NA
5134	1.22E-04	2.56E-04	6.02E-04	1.08E-03	1.68E-03	2.08E-03	3.36E-03	3.99E-03	2.20E-03
5135	1.58E-04	2.36E-04	5.46E-04	7.30E-04	1.11E-03	1.99E-03	3.93E-03	6.17E-03	4.72E-03
5136	4.44E-04	9.11E-04	1.56E-03	2.44E-03	2.92E-03	3.06E-03	6.25E-03	4.55E-03	3.87E-03
5137	1.61E-04	3.30E-04	4.88E-04	5.59E-04	1.05E-03	1.53E-03	2.26E-03	2.77E-03	4.67E-03
5138	4.42E-04	9.00E-04	1.84E-03	2.67E-03	5.19E-03	6.27E-03	6.57E-03	5.09E-03	NA
5139	3.60E-04	5.64E-04	8.61E-04	1.08E-03	1.04E-03	1.21E-03	1.75E-03	1.39E-03	7.76E-04
5140	9.37E-05	1.50E-04	2.51E-04	3.71E-04	4.59E-04	5.09E-04	6.50E-04	5.25E-04	3.26E-04
5143	9.20E-05	1.56E-04	2.83E-04	4.84E-04	5.64E-04	5.90E-04	3.93E-04	4.07E-04	3.75E-04
5144	7.11E-05	1.16E-04	2.64E-04	3.27E-04	3.60E-04	4.54E-04	4.45E-04	8.71E-04	8.95E-04
5146	1.80E-04	3.21E-04	1.00E-03	1.26E-03	1.74E-03	2.40E-03	3.25E-03	2.77E-03	NA
5147	2.89E-04	5.21E-04	1.51E-03	2.29E-03	3.68E-03	4.81E-03	8.29E-03	8.44E-03	5.32E-03
5148	1.37E-04	2.08E-04	4.94E-04	6.45E-04	7.38E-04	1.01E-03	1.20E-03	1.86E-03	1.29E-03
5156	6.47E-05	1.17E-04	1.98E-04	3.20E-04	5.51E-04	7.66E-04	1.07E-03	1.82E-03	NA
5157	2.16E-04	4.88E-04	9.29E-04	8.43E-04	8.81E-04	1.27E-03	1.98E-03	1.74E-03	1.09E-03
5158	4.22E-04	8.89E-04	1.37E-03	1.54E-03	3.25E-03	3.03E-03	3.37E-03	4.24E-03	NA
5159	3.97E-04	5.79E-04	1.77E-03	3.46E-03	7.91E-03	9.27E-03	9.54E-03	1.05E-02	5.48E-03
5160	9.95E-05	2.21E-04	5.32E-04	7.58E-04	1.01E-03	1.45E-03	1.62E-03	1.96E-03	2.52E-03
5168	5.18E-05	8.69E-05	2.14E-04	2.29E-04	3.30E-04	2.64E-04	1.85E-04	1.73E-04	1.48E-04
5188	6.28E-05	1.25E-04	2.93E-04	4.19E-04	4.69E-04	5.37E-04	7.57E-04	5.06E-04	NA
5196	4.89E-05	1.23E-04	4.03E-04	5.43E-04	7.74E-04	1.16E-03	1.60E-03	1.69E-03	2.13E-03
6932	6.35E-04	1.00E-03	2.01E-03	2.43E-03	4.27E-03	8.49E-03	7.56E-03	1.12E-02	1.66E-02
6933	5.10E-04	8.46E-04	7.95E-04	1.21E-03	2.40E-03	2.58E-03	3.53E-03	7.79E-03	1.63E-02
6934	7.65E-03	9.69E-03	1.20E-02	1.43E-02	1.48E-02	2.88E-02	2.86E-02	5.84E-02	5.18E-02
7010	NA	NA	NA	NA	5.81E-04	5.87E-04	6.21E-04	5.72E-04	1.19E-03
7018	2.06E-04	2.60E-04	3.46E-04	2.94E-04	6.97E-04	8.54E-04	1.09E-03	1.41E-03	4.05E-03
7019	NA	NA	NA	NA	NA	8.31E-04	1.11E-03	1.48E-03	4.84E-03
7062	6.40E-03	1.12E-02	2.30E-02	3.26E-02	2.82E-02	3.08E-02	6.38E-02	9.24E-02	1.58E-01

Table 9–1(g) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
7063	NA	1.29E-02	3.60E-02	4.76E-02	8.77E-02	1.55E-01	2.34E-01	2.23E-01	3.22E-01
7360	6.61E-05	9.61E-05	1.49E-04	2.07E-04	4.81E-04	4.89E-04	5.42E-04	4.66E-04	NA
7361	5.79E-05	8.94E-05	1.49E-04	2.57E-04	3.51E-04	3.82E-04	4.56E-04	5.79E-04	7.01E-04
7362	9.54E-05	1.29E-04	2.05E-04	2.58E-04	4.54E-04	5.41E-04	6.59E-04	7.77E-04	7.76E-04
7372	1.54E-04	2.36E-04	4.76E-04	4.69E-04	1.12E-03	1.41E-03	1.07E-03	7.16E-04	NA
7373	1.44E-04	2.97E-04	6.80E-04	7.48E-04	7.99E-04	8.90E-04	5.97E-04	4.27E-04	4.42E-04
7374	5.81E-05	8.38E-05	1.37E-04	2.27E-04	4.39E-04	5.51E-04	4.24E-04	3.62E-04	2.32E-04
7375	7.28E-05	9.95E-05	1.92E-04	2.21E-04	3.20E-04	5.33E-04	5.32E-04	4.63E-04	4.49E-04
7376	2.89E-04	3.01E-04	2.82E-04	3.38E-04	6.00E-04	6.56E-04	7.14E-04	8.32E-04	1.19E-03
7388	3.33E-04	5.58E-04	3.19E-04	4.20E-04	6.11E-04	4.54E-04	6.98E-04	8.20E-04	5.08E-04
7389	9.59E-05	1.89E-04	3.26E-04	6.65E-04	1.63E-03	1.06E-03	9.85E-04	1.16E-03	9.83E-04
7390	2.34E-04	2.92E-04	3.23E-04	3.49E-04	5.59E-04	5.69E-04	5.69E-04	8.29E-04	1.48E-03
7391	8.80E-05	1.13E-04	2.06E-04	2.72E-04	5.56E-04	9.85E-04	1.46E-03	2.14E-03	1.87E-03
7392	2.25E-04	3.84E-04	6.96E-04	8.40E-04	1.79E-03	1.81E-03	1.86E-03	1.31E-03	1.32E-03
7394	1.90E-04	2.08E-04	3.04E-04	6.65E-04	8.29E-04	7.08E-04	9.52E-04	1.15E-03	8.20E-04
7402	3.41E-04	3.59E-04	3.28E-04	4.50E-04	5.27E-04	6.32E-04	1.14E-03	9.45E-04	6.41E-04
7403	1.98E-04	2.46E-04	3.69E-04	6.69E-04	1.16E-03	8.20E-04	1.56E-03	1.79E-03	1.04E-03
7404	2.04E-04	4.70E-04	5.28E-04	1.03E-03	3.23E-03	3.22E-03	3.10E-03	4.46E-03	3.59E-03
7405	3.74E-04	9.67E-04	7.48E-04	9.36E-04	2.16E-03	2.82E-03	3.18E-03	2.95E-03	4.44E-03
7406	4.99E-04	9.51E-04	2.76E-03	5.94E-03	2.10E-02	2.16E-02	1.02E-02	6.18E-03	4.68E-03
7411	1.26E-04	1.46E-04	4.37E-04	6.92E-04	1.07E-03	1.47E-03	1.81E-03	1.55E-03	9.85E-04
7412	1.16E-04	1.78E-04	3.40E-04	6.52E-04	1.06E-03	8.93E-04	1.13E-03	1.54E-03	8.19E-04
7413	4.71E-04	9.34E-04	1.90E-03	2.34E-03	4.66E-03	3.58E-03	4.38E-03	4.87E-03	4.29E-03
7414	5.51E-04	1.04E-03	1.21E-03	2.18E-03	4.15E-03	3.87E-03	7.55E-03	6.79E-03	5.95E-03
7420	1.32E-04	1.66E-04	4.01E-04	5.55E-04	5.76E-04	5.90E-04	5.47E-04	7.49E-04	6.52E-04
7421	1.11E-04	2.22E-04	4.32E-04	8.36E-04	1.61E-03	1.14E-03	1.34E-03	2.89E-03	1.26E-03
7422	1.73E-04	3.22E-04	1.24E-03	1.27E-03	2.32E-03	1.95E-03	2.16E-03	2.85E-03	3.61E-03
7423	1.28E-04	1.48E-04	3.78E-04	5.18E-04	7.13E-04	7.69E-04	1.13E-03	1.34E-03	1.76E-03
7430	6.67E-05	1.11E-04	2.70E-04	6.00E-04	5.70E-04	6.54E-04	4.68E-04	4.68E-04	3.94E-04
7431	2.15E-04	2.91E-04	4.54E-04	7.57E-04	7.28E-04	5.40E-04	4.56E-04	4.22E-04	3.98E-04
7432	9.55E-05	1.46E-04	4.13E-04	4.64E-04	7.35E-04	5.80E-04	4.51E-04	4.91E-04	3.62E-04
7433	1.32E-04	2.01E-04	6.48E-04	8.50E-04	9.67E-04	1.25E-03	1.17E-03	1.40E-03	1.03E-03
7434	2.67E-04	7.52E-04	1.15E-03	1.89E-03	3.90E-03	5.87E-03	4.57E-03	5.51E-03	8.56E-03
7486	2.50E-04	6.58E-04	1.77E-03	2.54E-03	2.38E-03	3.19E-03	4.70E-03	3.06E-03	2.49E-03
9457	2.25E-04	3.39E-04	3.81E-04	6.71E-04	9.85E-04	1.25E-03	1.71E-03	3.20E-03	1.19E-02
9467	NA	NA	5.16E-05	7.52E-05	1.25E-04	1.94E-04	2.55E-04	3.53E-04	7.39E-04
9470	NA	NA	NA	NA	1.98E-04	3.42E-04	5.17E-04	5.90E-04	9.21E-04
9567	1.12E-04	2.08E-04	2.47E-04	4.16E-04	4.56E-04	4.82E-04	3.90E-04	3.10E-04	2.32E-04
9568	2.31E-03	2.92E-03	4.85E-03	6.45E-03	6.61E-03	9.37E-03	7.62E-03	7.01E-03	5.27E-03
9873	2.34E-04	3.66E-04	3.92E-04	4.79E-04	4.63E-04	4.55E-04	5.15E-04	6.53E-04	4.99E-04
9874	3.56E-04	4.49E-04	6.93E-04	6.52E-04	6.33E-04	7.45E-04	8.51E-04	7.23E-04	6.01E-04
9875	2.45E-04	5.98E-04	8.51E-04	1.08E-03	9.16E-04	8.97E-04	8.73E-04	8.01E-04	7.73E-04
9884	1.19E-04	2.94E-04	2.88E-04	4.68E-04	5.34E-04	5.94E-04	5.94E-04	9.90E-04	9.14E-04
9885	2.65E-04	4.24E-04	7.01E-04	1.06E-03	8.47E-04	1.18E-03	1.30E-03	1.45E-03	9.65E-04
9886	7.21E-04	9.82E-04	2.19E-03	2.34E-03	1.85E-03	1.39E-03	1.40E-03	1.38E-03	1.25E-03
9887	2.03E-04	4.17E-04	4.27E-04	5.34E-04	5.77E-04	5.78E-04	5.22E-04	5.89E-04	1.05E-03
9888	1.21E-04	2.23E-04	3.55E-04	5.67E-04	7.91E-04	8.03E-04	1.01E-03	9.58E-04	1.02E-03
9895	2.65E-04	4.23E-04	4.39E-04	7.65E-04	1.38E-03	1.54E-03	1.89E-03	3.70E-03	2.80E-03
9896	8.91E-04	1.05E-03	9.36E-04	1.32E-03	2.20E-03	2.74E-03	2.38E-03	2.59E-03	1.57E-03
9897	6.29E-04	1.68E-03	2.14E-03	1.98E-03	2.12E-03	2.44E-03	2.95E-03	2.70E-03	2.28E-03
9898	1.99E-04	4.58E-04	6.57E-04	9.76E-04	8.91E-04	8.38E-04	1.01E-03	1.17E-03	1.03E-03
9899	1.02E-04	1.59E-04	2.14E-04	3.89E-04	4.59E-04	3.61E-04	4.93E-04	7.32E-04	4.37E-04
9900	1.33E-04	1.96E-04	1.78E-04	2.75E-04	3.62E-04	4.46E-04	4.81E-04	6.01E-04	6.49E-04
9910	4.73E-04	7.53E-04	1.86E-03	3.29E-03	4.30E-03	3.16E-03	4.91E-03	9.36E-03	6.47E-03
9911	6.84E-04	1.47E-03	2.42E-03	4.80E-03	4.77E-03	7.41E-03	5.58E-03	3.82E-03	3.52E-03
9912	5.71E-04	8.40E-04	9.88E-04	1.09E-03	1.07E-03	1.83E-03	2.57E-03	3.17E-03	4.36E-03
9913	2.28E-04	6.87E-04	1.22E-03	2.33E-03	2.24E-03	2.41E-03	3.20E-03	5.44E-03	7.35E-03
9914	2.69E-04	5.34E-04	3.99E-04	4.38E-04	5.30E-04	6.10E-04	1.42E-03	1.65E-03	2.36E-03

Table 9–1(h) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
9915	6.05E-05	1.14E-04	1.27E-04	1.64E-04	2.52E-04	3.69E-04	4.38E-04	8.98E-04	1.05E-03
9921	3.09E-04	4.58E-04	6.99E-04	1.20E-03	1.94E-03	3.05E-03	3.08E-03	5.52E-03	5.01E-03
9922	1.64E-03	2.40E-03	2.10E-03	1.72E-03	2.18E-03	2.18E-03	2.10E-03	1.92E-03	2.63E-03
9923	5.00E-04	6.65E-04	1.05E-03	1.50E-03	2.42E-03	3.51E-03	4.02E-03	4.87E-03	5.29E-03
9924	4.48E-04	6.62E-04	8.97E-04	1.15E-03	2.43E-03	2.20E-03	3.67E-03	4.35E-03	5.65E-03
9925	3.25E-04	7.05E-04	8.63E-04	1.10E-03	1.45E-03	2.85E-03	4.99E-03	7.48E-03	8.04E-03
9926	1.82E-04	3.37E-04	5.95E-04	1.28E-03	2.39E-03	2.49E-03	5.54E-03	4.67E-03	6.38E-03
9927	1.29E-04	2.12E-04	3.18E-04	4.44E-04	5.33E-04	7.58E-04	6.96E-04	1.07E-03	2.28E-03
9928	1.44E-04	2.81E-04	6.02E-04	1.36E-03	3.70E-03	3.41E-03	2.71E-03	1.53E-03	1.03E-03
9935	3.18E-03	3.12E-03	6.42E-03	8.96E-03	8.00E-03	8.97E-03	8.60E-03	1.36E-02	1.94E-02
9936	3.44E-04	4.10E-04	7.09E-04	9.97E-04	1.59E-03	1.99E-03	3.65E-03	4.40E-03	4.81E-03
9937	2.13E-04	3.43E-04	7.26E-04	1.01E-03	3.12E-03	3.67E-03	3.24E-03	3.11E-03	2.39E-03
9938	5.74E-04	6.24E-04	1.52E-03	2.58E-03	4.95E-03	5.86E-03	4.81E-03	3.80E-03	3.82E-03
9939	3.65E-04	5.51E-04	6.99E-04	1.14E-03	1.71E-03	2.21E-03	1.75E-03	1.84E-03	1.42E-03
9940	1.43E-04	2.16E-04	4.29E-04	4.49E-04	5.41E-04	5.18E-04	6.09E-04	5.08E-04	3.91E-04
9946	1.63E-03	3.05E-03	3.22E-03	2.84E-03	4.47E-03	3.84E-03	4.22E-03	4.76E-03	4.06E-03
9947	1.79E-03	2.65E-03	1.89E-03	3.41E-03	3.86E-03	2.51E-03	3.64E-03	4.97E-03	5.15E-03
9948	1.19E-03	2.20E-03	3.06E-03	5.85E-03	1.44E-02	1.42E-02	8.95E-03	1.23E-02	5.37E-03
9949	1.04E-03	1.23E-03	1.44E-03	3.05E-03	2.63E-03	3.43E-03	2.01E-03	1.79E-03	2.22E-03
9950	2.01E-04	3.56E-04	3.52E-04	6.03E-04	6.11E-04	9.88E-04	8.94E-04	5.45E-04	4.44E-04
9958	3.21E-04	5.31E-04	8.44E-04	1.53E-03	1.65E-03	2.48E-03	2.50E-03	2.81E-03	3.13E-03
9959	3.39E-04	4.36E-04	7.14E-04	1.41E-03	2.08E-03	2.15E-03	2.48E-03	3.16E-03	2.78E-03
9960	5.63E-04	9.96E-04	1.44E-03	2.28E-03	1.83E-03	1.87E-03	1.87E-03	2.01E-03	1.19E-03
9961	4.08E-04	4.85E-04	8.18E-04	9.00E-04	1.12E-03	1.09E-03	7.21E-04	9.63E-04	8.34E-04
9962	2.76E-04	2.82E-04	4.29E-04	4.78E-04	5.34E-04	5.49E-04	6.04E-04	4.94E-04	3.35E-04
9963	3.17E-04	3.78E-04	4.45E-04	6.40E-04	7.78E-04	8.82E-04	8.88E-04	9.67E-04	8.37E-04
9969	4.86E-04	6.11E-04	9.93E-04	1.24E-03	1.34E-03	9.79E-04	9.49E-04	8.96E-04	7.13E-04
10016	3.23E-04	4.53E-04	5.27E-04	6.81E-04	6.91E-04	6.98E-04	5.15E-04	3.15E-04	2.87E-04
10017	2.94E-04	3.57E-04	3.03E-04	2.81E-04	2.70E-04	3.13E-04	2.87E-04	2.00E-04	NA
10023	2.98E-04	3.16E-04	4.74E-04	5.91E-04	6.46E-04	6.99E-04	4.85E-04	5.48E-04	5.28E-04
10030	3.75E-04	6.52E-04	1.06E-03	1.51E-03	1.19E-03	1.51E-03	2.09E-03	1.89E-03	1.38E-03
10031	3.74E-04	6.49E-04	1.05E-03	1.51E-03	1.18E-03	1.51E-03	2.08E-03	1.88E-03	1.32E-03
10042	1.58E-03	1.81E-03	2.64E-03	3.02E-03	4.94E-03	5.58E-03	9.09E-03	1.05E-02	2.55E-02
10054	1.13E-02	1.05E-02	9.43E-03	1.07E-02	1.12E-02	1.58E-02	2.80E-02	3.11E-02	5.13E-02
10055	1.85E-02	1.48E-02	1.17E-02	1.34E-02	1.32E-02	1.50E-02	2.55E-02	3.42E-02	6.84E-02
10056	8.79E-03	7.46E-03	8.65E-03	1.54E-02	1.77E-02	2.18E-02	2.80E-02	2.06E-02	3.36E-02
10057	5.22E-03	6.03E-03	1.20E-02	1.72E-02	2.30E-02	2.43E-02	2.03E-02	1.99E-02	3.15E-02
10058	9.65E-03	1.12E-02	1.21E-02	1.68E-02	1.56E-02	2.91E-02	4.36E-02	4.24E-02	7.71E-02
10427	NA	NA	1.07E-03	1.14E-03	1.77E-03	1.93E-03	2.17E-03	2.48E-03	NA
10428	6.19E-04	9.26E-04	1.99E-03	2.28E-03	2.05E-03	2.98E-03	2.95E-03	2.18E-03	1.75E-03
10429	7.56E-04	7.58E-04	2.57E-03	3.27E-03	2.86E-03	2.48E-03	2.65E-03	2.82E-03	2.40E-03
10438	NA	NA	1.06E-03	1.06E-03	2.02E-03	2.56E-03	2.79E-03	3.08E-03	3.19E-03
10439	6.17E-04	8.73E-04	1.54E-03	2.83E-03	2.21E-03	2.69E-03	4.47E-03	6.42E-03	3.39E-03
10440	1.55E-03	1.75E-03	3.61E-03	4.51E-03	4.07E-03	3.44E-03	3.61E-03	4.39E-03	3.12E-03
10441	7.10E-04	1.24E-03	2.22E-03	2.20E-03	2.17E-03	2.70E-03	1.79E-03	NA	NA
10442	5.86E-04	1.16E-03	1.58E-03	1.89E-03	2.35E-03	3.50E-03	2.74E-03	3.25E-03	3.66E-03
10450	NA	1.95E-03	2.64E-03	3.59E-03	4.10E-03	8.50E-03	1.29E-02	1.35E-02	1.08E-02
10451	1.08E-03	1.73E-03	2.98E-03	3.35E-03	6.74E-03	7.13E-03	9.06E-03	8.93E-03	4.54E-03
10452	1.02E-03	3.30E-03	5.96E-03	5.93E-03	5.91E-03	6.89E-03	9.03E-03	7.91E-03	NA
10453	9.09E-04	1.20E-03	3.19E-03	4.10E-03	2.67E-03	3.42E-03	3.17E-03	3.26E-03	3.16E-03
10454	7.32E-04	6.90E-04	1.44E-03	2.20E-03	1.50E-03	2.04E-03	1.68E-03	1.62E-03	1.37E-03
10455	9.20E-04	8.10E-04	9.10E-04	1.42E-03	1.63E-03	2.00E-03	1.81E-03	1.87E-03	NA
10466	1.98E-03	2.38E-03	4.57E-03	6.76E-03	8.57E-03	1.16E-02	1.52E-02	2.38E-02	NA
10467	2.77E-03	2.43E-03	5.57E-03	6.40E-03	1.03E-02	1.06E-02	1.52E-02	1.57E-02	NA

Table 9–1(i) NGA-East site-corrected PSA used for data-based weights. Additional information on each record can be found in Goulet et al. (2014) through the record sequence number (RSN).

RSN	1 HZ	1.33 Hz	2 Hz	2.5 Hz	3.33 Hz	4 Hz	5 Hz	6.67 Hz	10 Hz
10468	1.61E-03	1.37E-03	3.21E-03	3.22E-03	5.17E-03	1.07E-02	1.13E-02	1.11E-02	NA
10469	9.15E-04	1.90E-03	3.65E-03	6.78E-03	5.91E-03	7.55E-03	7.41E-03	1.54E-02	1.44E-02
10470	8.64E-04	2.37E-03	2.62E-03	2.75E-03	2.83E-03	3.48E-03	NA	NA	NA
10471	3.49E-04	6.05E-04	9.67E-04	9.52E-04	1.12E-03	1.64E-03	NA	NA	NA
10477	NA	5.59E-03	3.99E-03	3.46E-03	4.46E-03	5.01E-03	6.80E-03	6.79E-03	NA
10478	2.26E-03	4.00E-03	6.16E-03	8.87E-03	6.59E-03	1.27E-02	1.55E-02	2.55E-02	2.17E-02
10479	1.58E-03	2.96E-03	5.44E-03	6.44E-03	9.16E-03	1.29E-02	1.08E-02	1.79E-02	1.40E-02
10480	1.22E-03	2.24E-03	2.36E-03	2.53E-03	2.17E-03	2.63E-03	2.88E-03	4.00E-03	7.17E-03
10481	NA	1.60E-03	4.61E-03	7.03E-03	1.16E-02	1.05E-02	8.84E-03	5.02E-03	3.31E-03
10488	1.35E-03	4.41E-03	6.70E-03	7.79E-03	2.10E-02	1.60E-02	1.30E-02	1.18E-02	1.19E-02
10489	1.12E-03	2.99E-03	3.97E-03	3.96E-03	5.43E-03	5.93E-03	7.74E-03	7.15E-03	3.99E-03
10490	5.68E-04	8.98E-04	1.85E-03	2.19E-03	1.58E-03	1.57E-03	2.38E-03	1.67E-03	1.23E-03
10496	1.59E-03	2.83E-03	4.18E-03	3.87E-03	3.42E-03	3.55E-03	4.31E-03	5.29E-03	6.96E-03
10497	4.17E-04	6.90E-04	1.36E-03	1.15E-03	1.15E-03	1.61E-03	1.67E-03	1.57E-03	1.61E-03
10505	1.59E-03	2.21E-03	3.50E-03	4.08E-03	5.21E-03	7.00E-03	9.35E-03	8.74E-03	NA
10506	3.16E-03	3.37E-03	2.82E-03	4.01E-03	6.06E-03	9.28E-03	7.61E-03	7.43E-03	5.53E-03
10507	1.24E-03	1.57E-03	1.91E-03	1.69E-03	1.74E-03	3.19E-03	3.55E-03	4.13E-03	2.84E-03
10508	5.96E-04	7.61E-04	7.56E-04	9.62E-04	1.18E-03	1.53E-03	2.42E-03	2.21E-03	1.47E-03
10509	6.76E-04	7.59E-04	1.08E-03	1.39E-03	1.69E-03	2.51E-03	3.07E-03	4.56E-03	3.39E-03
10564	5.71E-04	9.21E-04	1.60E-03	1.68E-03	2.38E-03	2.84E-03	2.28E-03	1.03E-03	1.03E-03
10565	5.00E-04	6.75E-04	9.22E-04	7.08E-04	8.72E-04	1.36E-03	1.05E-03	6.01E-04	7.47E-04
10570	4.77E-04	6.54E-04	7.74E-04	9.77E-04	1.13E-03	1.54E-03	1.91E-03	2.18E-03	2.32E-03
10580	4.59E-03	6.70E-03	1.32E-02	1.18E-02	8.44E-03	7.91E-03	7.44E-03	1.27E-02	7.60E-03
10581	2.51E-03	4.05E-03	7.87E-03	7.56E-03	6.60E-03	6.34E-03	6.63E-03	9.58E-03	7.71E-03
12006	NA	NA	1.51E-02	1.74E-02	1.71E-02	1.40E-02	1.93E-02	2.26E-02	2.06E-02
12007	NA	7.45E-03	1.29E-02	1.03E-02	2.10E-02	2.16E-02	2.66E-02	3.48E-02	2.70E-02
12008	NA	NA	1.39E-02	1.45E-02	2.21E-02	2.44E-02	3.28E-02	4.31E-02	3.48E-02
12009	NA	NA	1.05E-02	1.44E-02	2.11E-02	2.33E-02	3.10E-02	3.23E-02	4.30E-02
12010	NA	NA	1.09E-02	1.49E-02	1.62E-02	2.17E-02	2.49E-02	2.96E-02	4.12E-02
12011	NA	NA	1.20E-02	1.23E-02	2.06E-02	2.27E-02	3.16E-02	2.73E-02	3.72E-02
12038	1.05E-04	2.24E-04	4.57E-04	5.37E-04	5.35E-04	4.97E-04	5.69E-04	9.07E-04	8.07E-04
12039	5.94E-05	8.18E-05	1.33E-04	2.93E-04	3.82E-04	6.67E-04	6.54E-04	9.06E-04	1.14E-03
12040	6.80E-05	9.85E-05	2.26E-04	3.24E-04	5.53E-04	6.67E-04	1.13E-03	1.99E-03	2.19E-03
12041	9.75E-05	1.78E-04	3.18E-04	4.54E-04	8.47E-04	1.45E-03	2.13E-03	2.04E-03	2.00E-03
12042	7.99E-05	1.28E-04	1.75E-04	2.36E-04	5.24E-04	7.89E-04	1.19E-03	1.65E-03	2.08E-03
12043	6.14E-05	9.82E-05	1.31E-04	1.51E-04	3.65E-04	4.74E-04	7.98E-04	1.09E-03	1.65E-03

Table 9–2(a) Total weights for the 17 models, for all the GMIMs (oscillator frequencies, f, in Hertz).

	f=0.1	f=0.133	f=0.2	f=0.25	f=0.333	f=0.5	f=0.667	f=1.	f=1.333	f=2.	f=2.5	f=3.333
Model 1	0.0955	0.0941	0.103	0.0994	0.0941	0.0945	0.1032	0.0998	0.1116	0.1044	0.1009	0.1013
Model 2	0.0833	0.093	0.0846	0.0904	0.0617	0.0897	0.0706	0.0749	0.0721	0.0852	0.0841	0.0683
Model 3	0.0837	0.079	0.0914	0.0935	0.0709	0.0783	0.0683	0.0684	0.0568	0.0844	0.0675	0.0732
Model 4	0.0904	0.0787	0.1071	0.1056	0.1037	0.0978	0.097	0.0922	0.086	0.0639	0.0785	0.0824
Model 5	0.0666	0.0617	0.0638	0.0673	0.0701	0.0679	0.0903	0.0917	0.0947	0.0953	0.0885	0.0733
Model 6	0.0914	0.0898	0.0658	0.065	0.0847	0.0717	0.0884	0.0885	0.0889	0.082	0.077	0.0692
Model 7	0.0969	0.0993	0.0828	0.0776	0.099	0.0842	0.0941	0.0878	0.0893	0.0787	0.0889	0.1082
Model 8	0.0778	0.0822	0.0873	0.0844	0.0878	0.0922	0.0869	0.0794	0.0956	0.0849	0.0839	0.1023
Model 9	0.0924	0.0991	0.1056	0.1114	0.09	0.1017	0.0743	0.0841	0.0808	0.0827	0.0918	0.0899
Model 10	0.0204	0.0111	0.0047	0.0087	0.004	0.0127	0.0121	0.0116	0.0189	0.0384	0.0225	0.0171
Model 11	0.0086	0.0072	0.0075	0.0077	0.0052	0.0058	0.0077	0.0096	0.012	0.0155	0.0155	0.0212
Model 12	0.0233	0.0224	0.0438	0.0375	0.0347	0.0233	0.0223	0.0278	0.0242	0.0201	0.0199	0.0359
Model 13	0.0196	0.0181	0.0183	0.0185	0.0245	0.0153	0.0287	0.0392	0.0243	0.0241	0.023	0.0182
Model 14	0.0516	0.0562	0.0395	0.036	0.053	0.0478	0.047	0.0416	0.0435	0.0411	0.0372	0.0214
Model 15	0.0464	0.0514	0.0416	0.043	0.0577	0.0548	0.0545	0.0463	0.0469	0.045	0.0518	0.0355
Model 16	0.0202	0.0267	0.0263	0.0239	0.0352	0.0328	0.0318	0.028	0.0255	0.0239	0.0353	0.0532
Model 17	0.0319	0.03	0.0269	0.0301	0.0237	0.0295	0.0228	0.0291	0.0289	0.0304	0.0337	0.0294

Table 9–2(b) Total weights for the 17 models, for all the GMIMs (oscillator frequencies, f, in Hertz).

	f=4.	f=5.	f=6.667	f=10.	f=13.333	f=20	f=25.	f=33.333	f=40	f=50.	f=100	PGA	PGV
Model 1	0.0921	0.0737	0.0683	0.1047	0.1068	0.0998	0.1069	0.1078	0.0987	0.0949	0.0935	0.1009	0.0976
Model 2	0.0585	0.0994	0.153	0.1175	0.1311	0.1315	0.1256	0.1316	0.1453	0.1176	0.1462	0.1606	0.0678
Model 3	0.0632	0.0892	0.0863	0.0723	0.0697	0.0965	0.088	0.0883	0.0996	0.0985	0.123	0.1151	0.0738
Model 4	0.0739	0.0691	0.0834	0.0676	0.0651	0.0686	0.068	0.0673	0.0653	0.0704	0.0981	0.097	0.0756
Model 5	0.0731	0.0456	0.0342	0.0677	0.0735	0.054	0.0579	0.0512	0.0396	0.0407	0.0472	0.0548	0.0702
Model 6	0.0965	0.1095	0.096	0.0553	0.0519	0.0559	0.06	0.0509	0.062	0.0666	0.033	0.0376	0.0916
Model 7	0.1198	0.102	0.0894	0.0725	0.0917	0.0649	0.0586	0.0627	0.0609	0.0643	0.0522	0.0507	0.098
Model 8	0.1123	0.0876	0.055	0.0642	0.0506	0.0743	0.0784	0.0727	0.0838	0.0984	0.0629	0.0497	0.1054
Model 9	0.0774	0.0859	0.086	0.1075	0.0938	0.1136	0.1221	0.1205	0.1057	0.1064	0.1092	0.0986	0.0956
Model 10	0.0123	0.0281	0.0212	0.0254	0.019	0.0374	0.0298	0.0245	0.0278	0.0246	0.0372	0.0372	0.0108
Model 11	0.0185	0.0214	0.0056	0.0088	0.0008	0.0191	0.0087	0.0016	0.003	0.0147	0.0123	0.01	0.0197
Model 12	0.0344	0.0293	0.0139	0.0175	0.0048	0.0178	0.0139	0.0055	0.0059	0.0174	0.0271	0.0167	0.0274
Model 13	0.0191	0.0176	0.0109	0.0158	0.0126	0.0117	0.0082	0.0057	0.0045	0.0098	0.0076	0.0119	0.0117
Model 14	0.0208	0.0366	0.063	0.082	0.098	0.0406	0.0473	0.0703	0.0648	0.0512	0.0368	0.0436	0.0257
Model 15	0.0418	0.0368	0.0688	0.0649	0.08	0.043	0.0549	0.0758	0.0626	0.0466	0.0418	0.0504	0.0365
Model 16	0.0606	0.0328	0.0196	0.0179	0.0226	0.0268	0.0272	0.0236	0.0345	0.0419	0.0266	0.0282	0.0567
Model 17	0.0257	0.0354	0.0454	0.0384	0.028	0.0445	0.0445	0.04	0.036	0.036	0.0453	0.037	0.0359

Figure 9–1 Sampled GMMs distribution on Sammon's maps for 1 Hz (see Section 8.3 for details and plots for all frequencies).

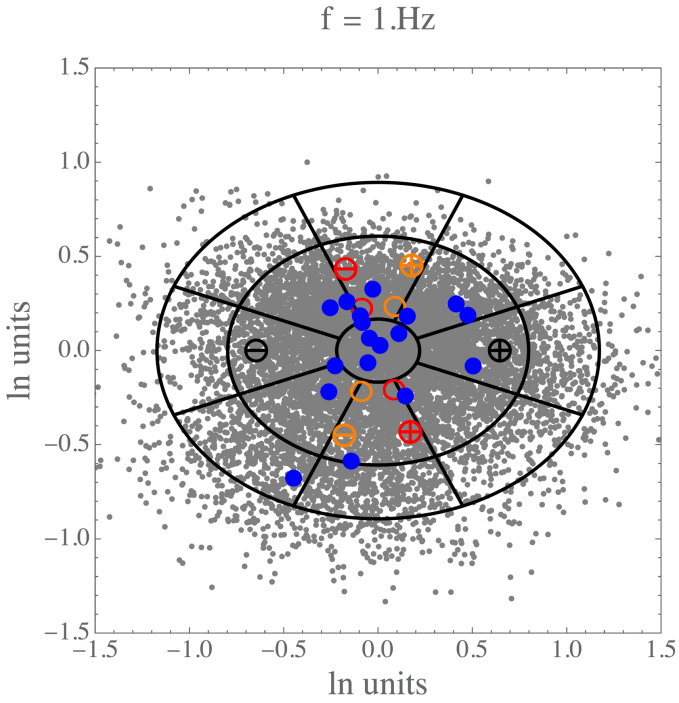


Figure 9–2 **Comparison of weights w_k (Fitted—PDF) in black and w_k ($N_{Samples}$) in red for the 1 Hz case.**

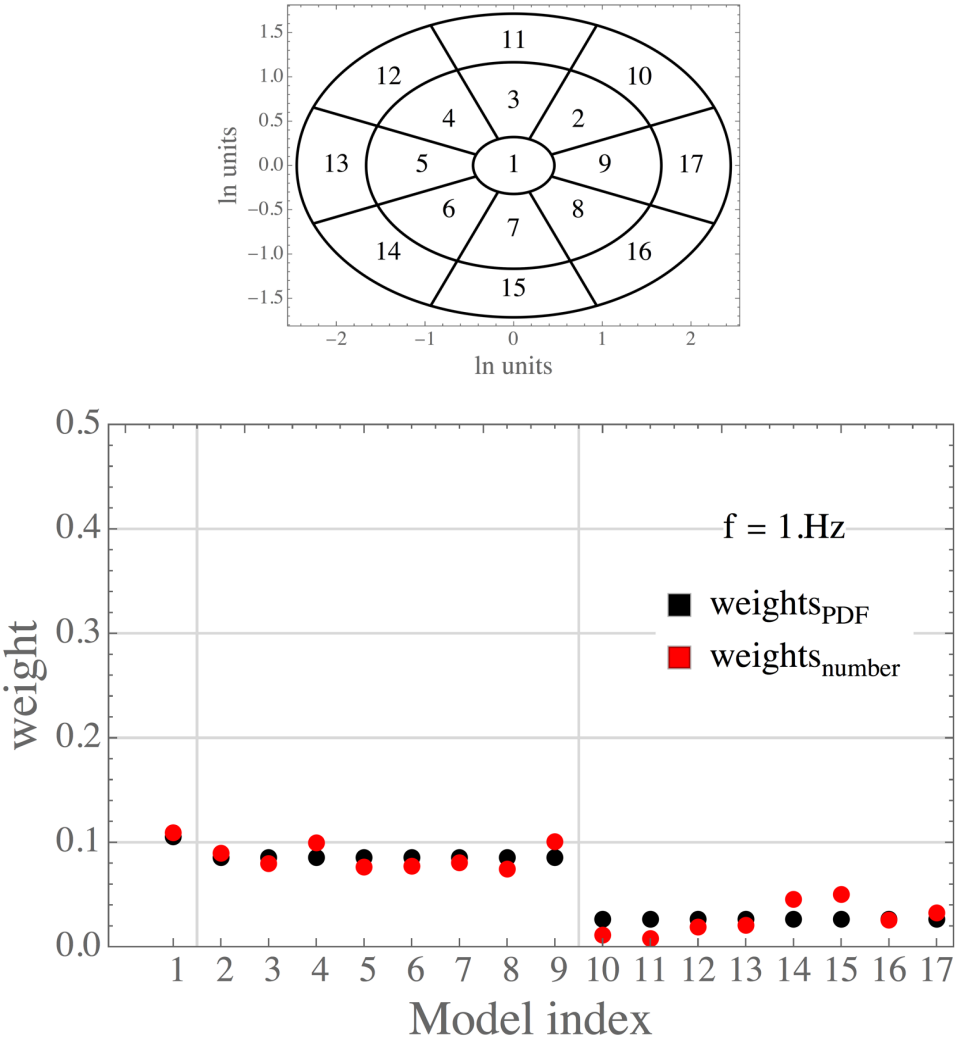


Figure 9–3 Magnitude and distance distribution of NGA-East data used to calculate residuals.

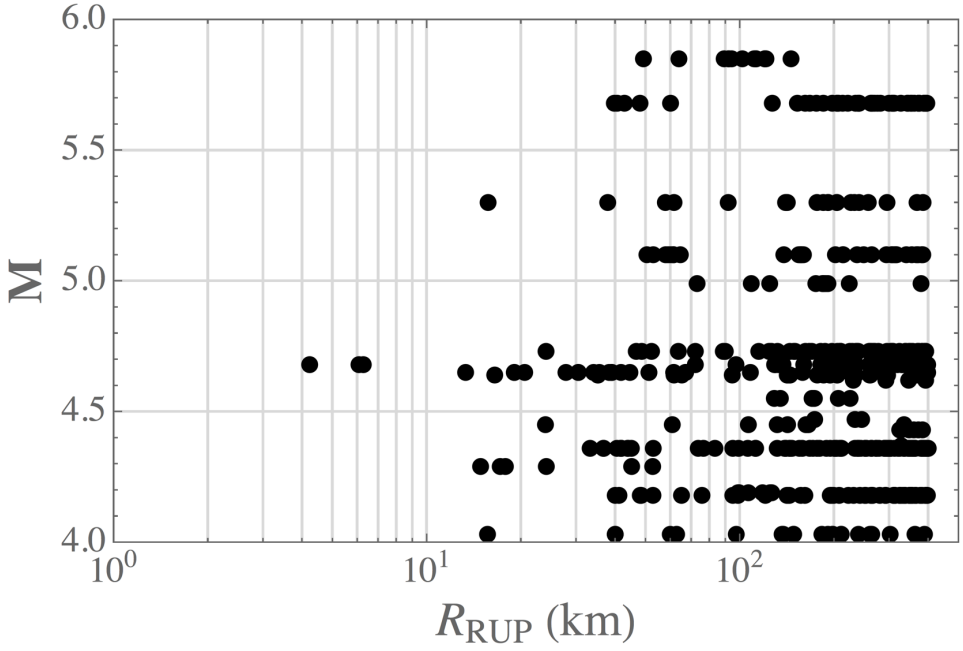


Figure 9–4 Weights based on residuals for different values of delta.

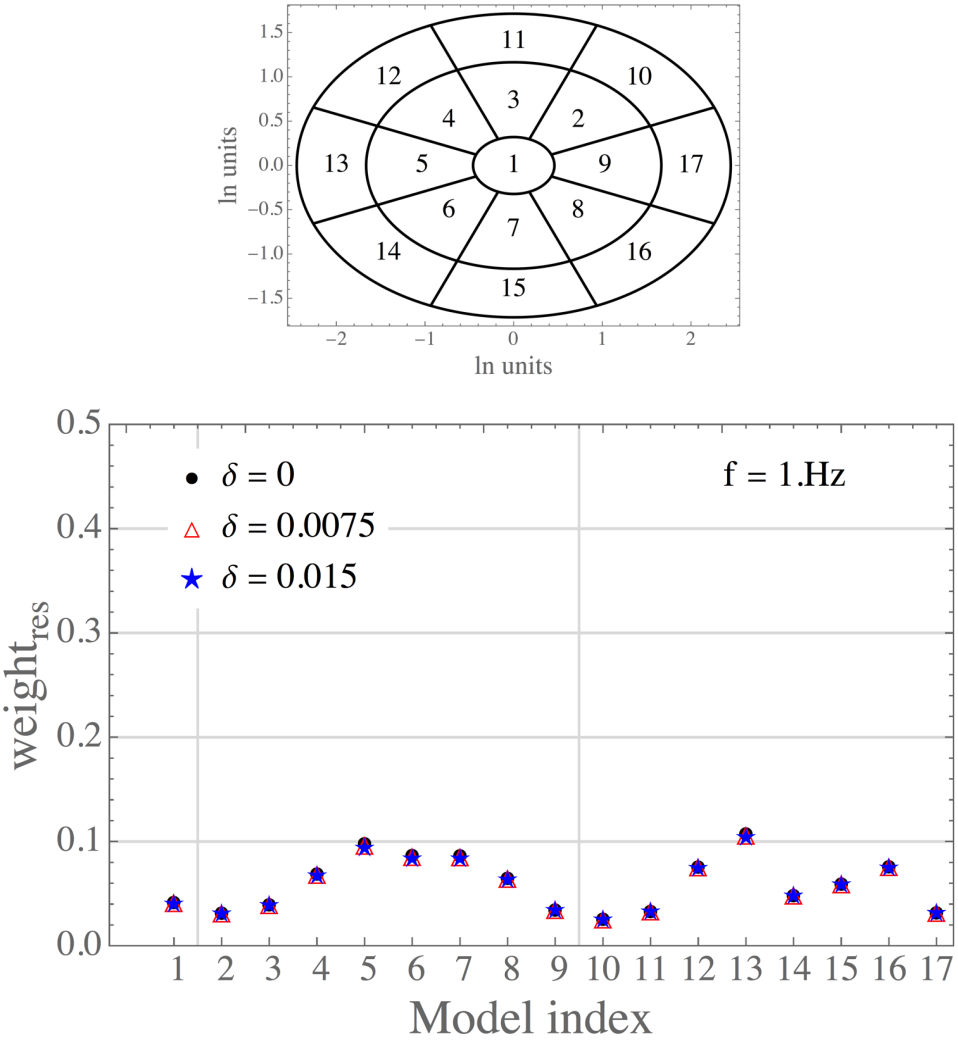


Figure 9–5(a) Contour plots of mean offset residuals for frequencies 1–4 Hz.

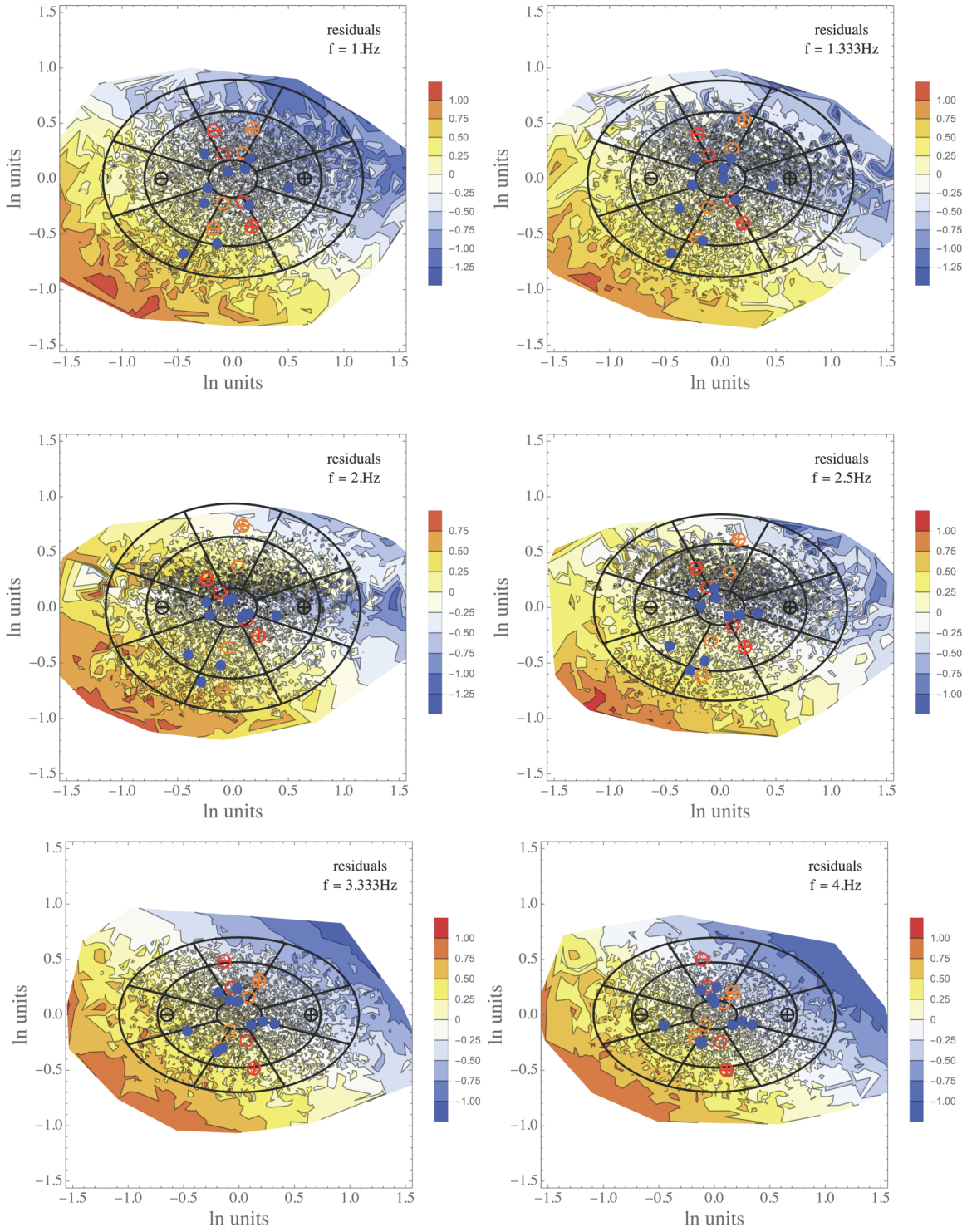


Figure 9–5(b) Contour plots of mean offset residuals for frequencies 5–10 Hz.

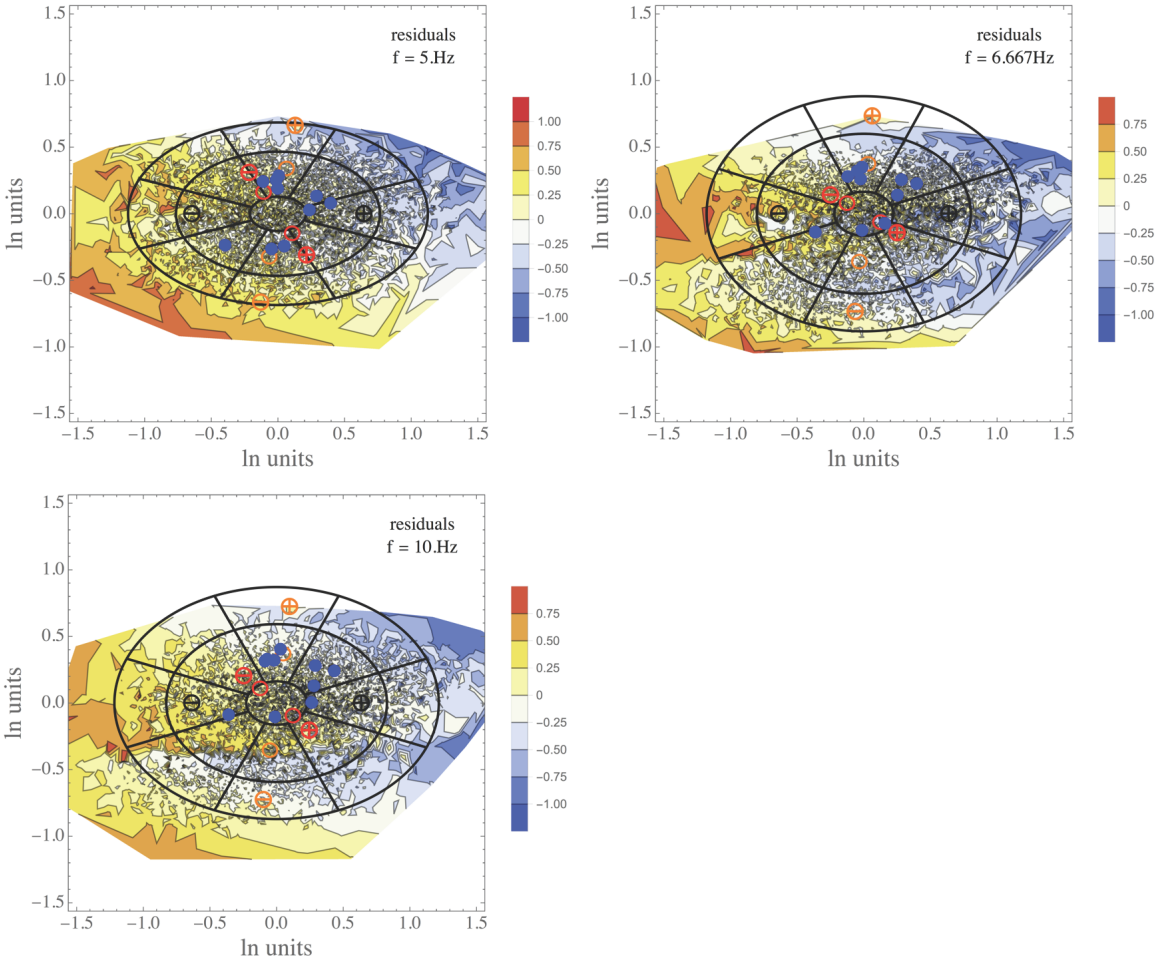


Figure 9–6(a) Contour plots of log-likelihood for frequencies 1–4 Hz.

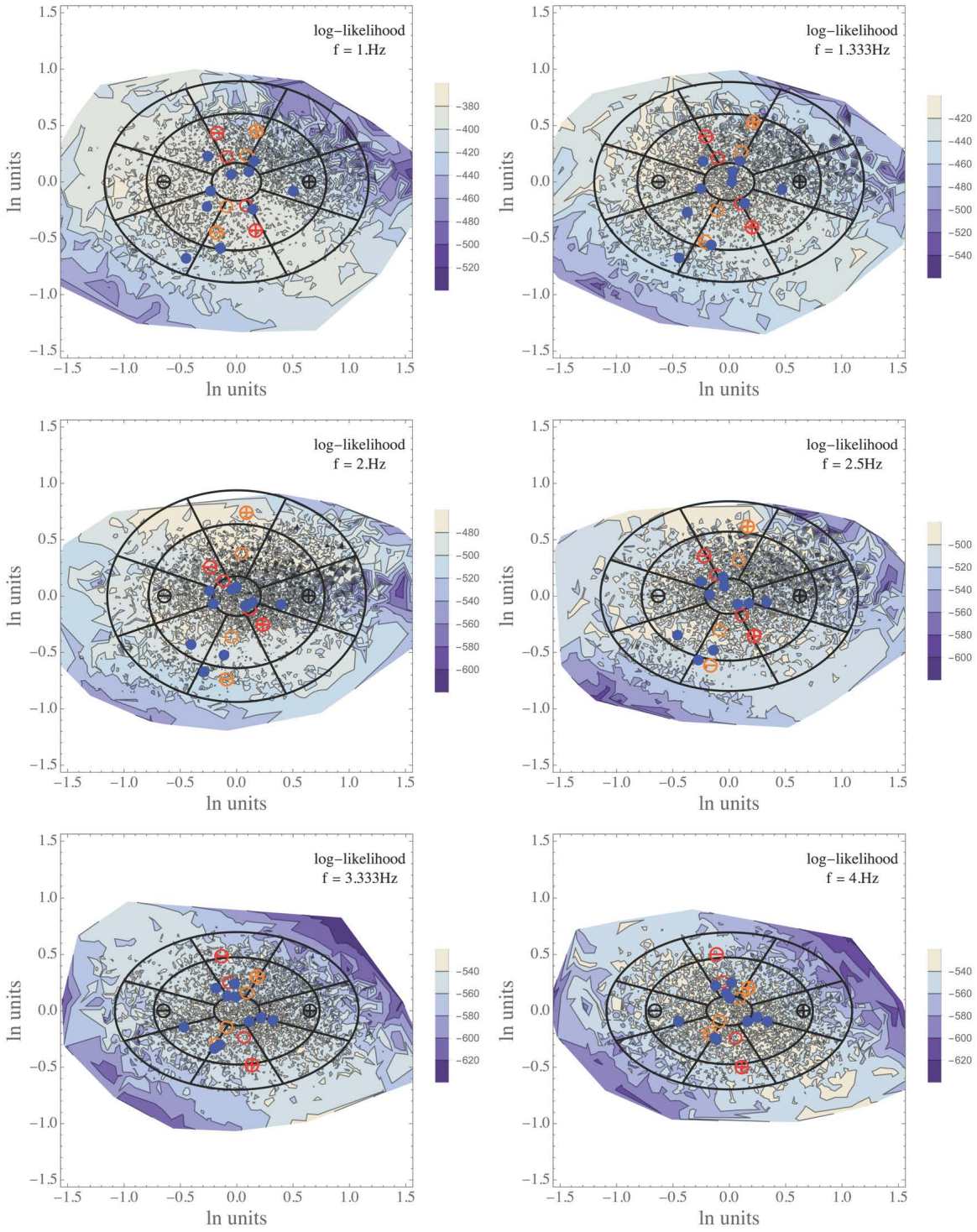


Figure 9–6(b) Contour plots of log-likelihod for frequencies 5–10 Hz.

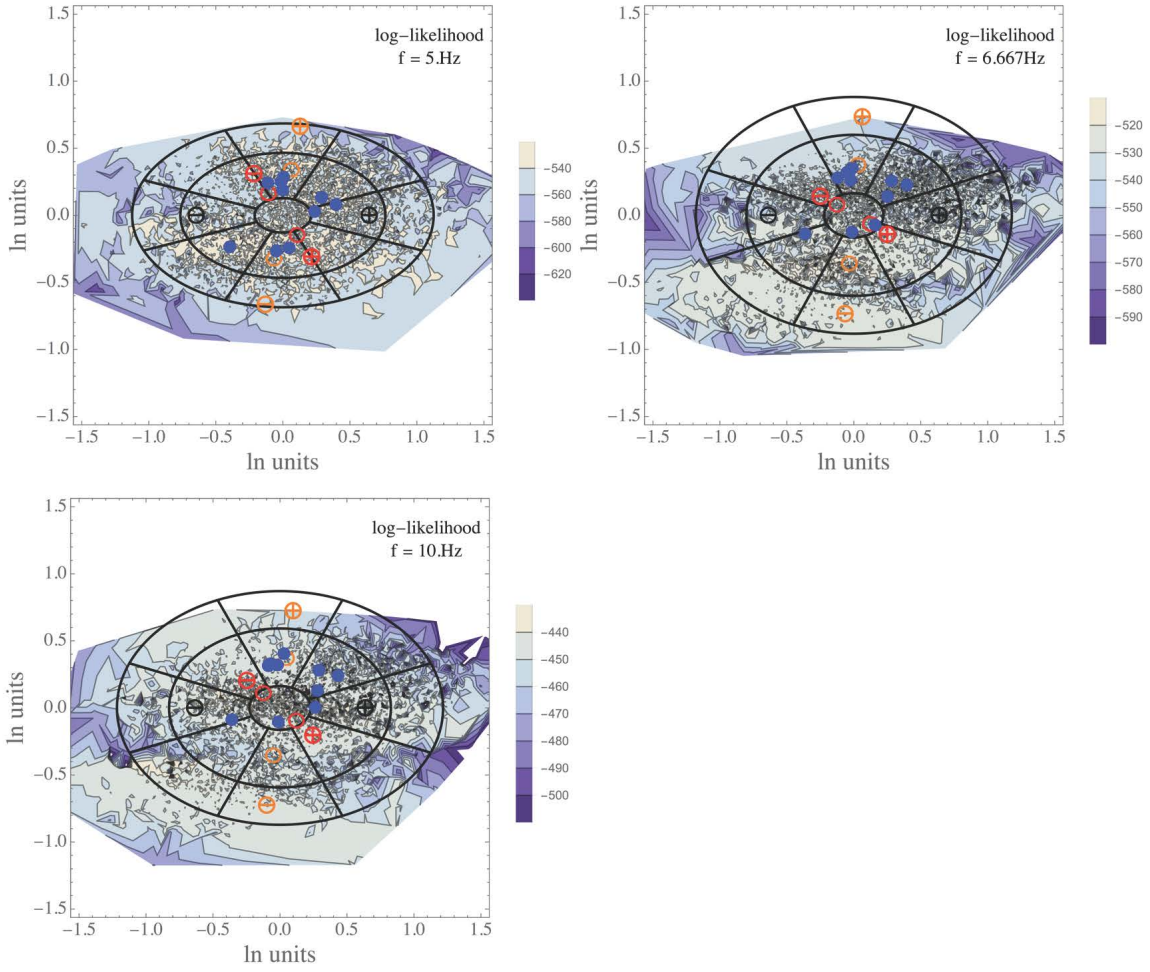


Figure 9–7(a) Weights based on residuals (black) and likelihood (red) for 1–4 Hz.

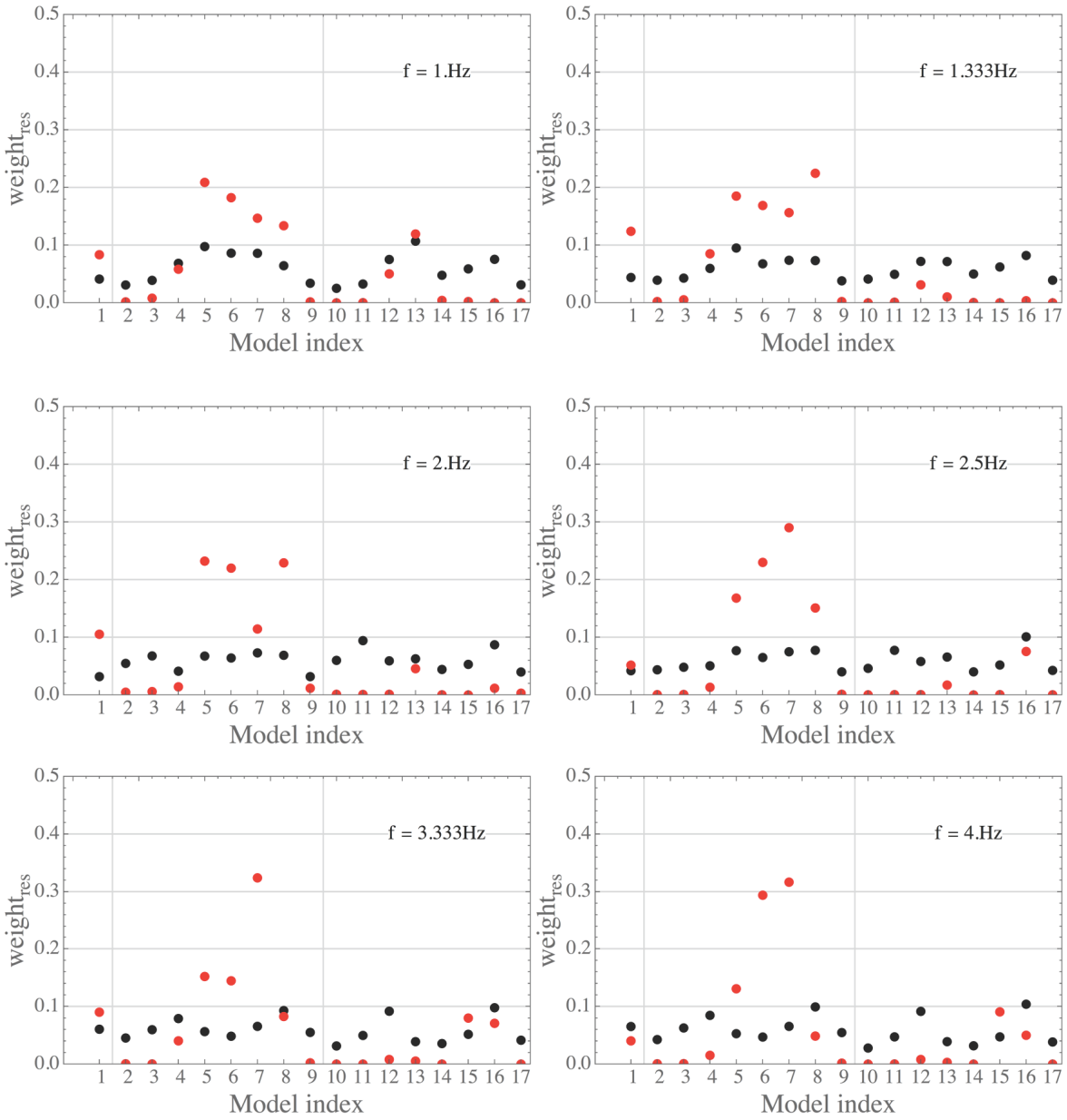


Figure 9–7(b) Weights based on residuals (black) and likelihood (red) for 5–10 Hz.

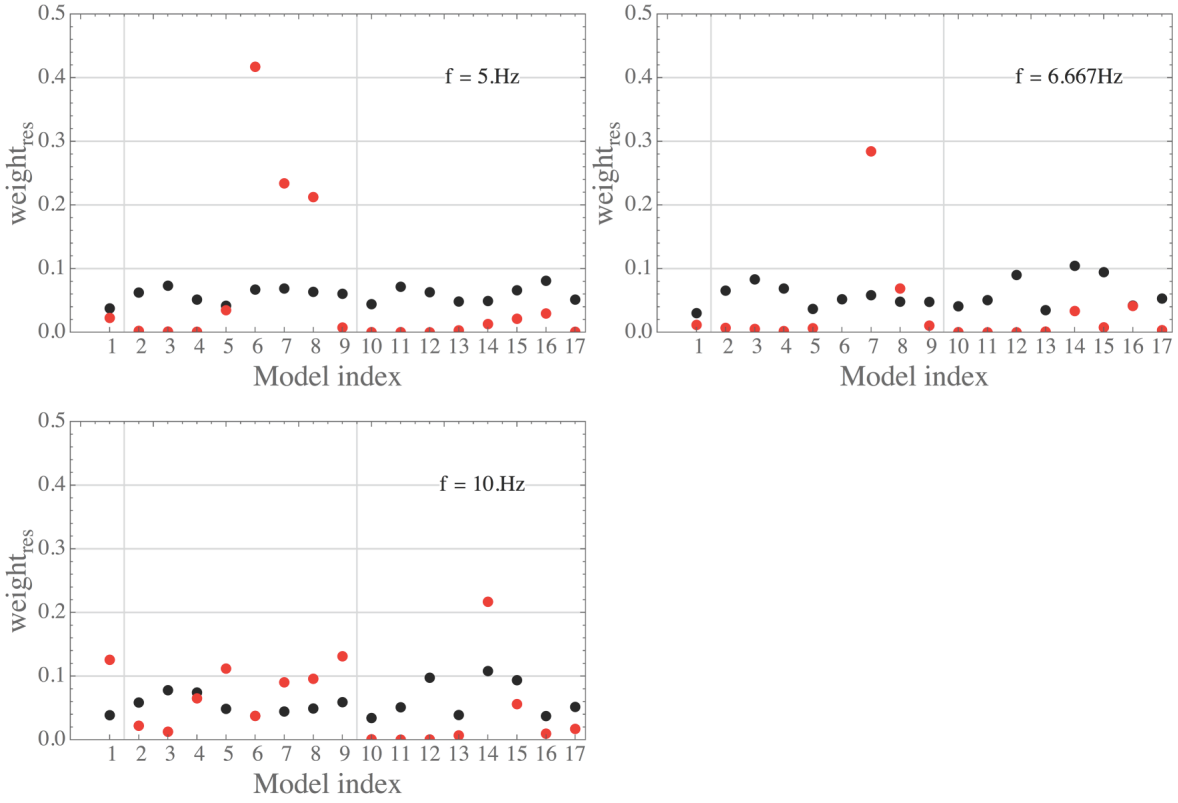


Figure 9–8 **Distance scaling of the 100 sampled models (out of 10,000) with the lowest residual (mean offset) to the data.**

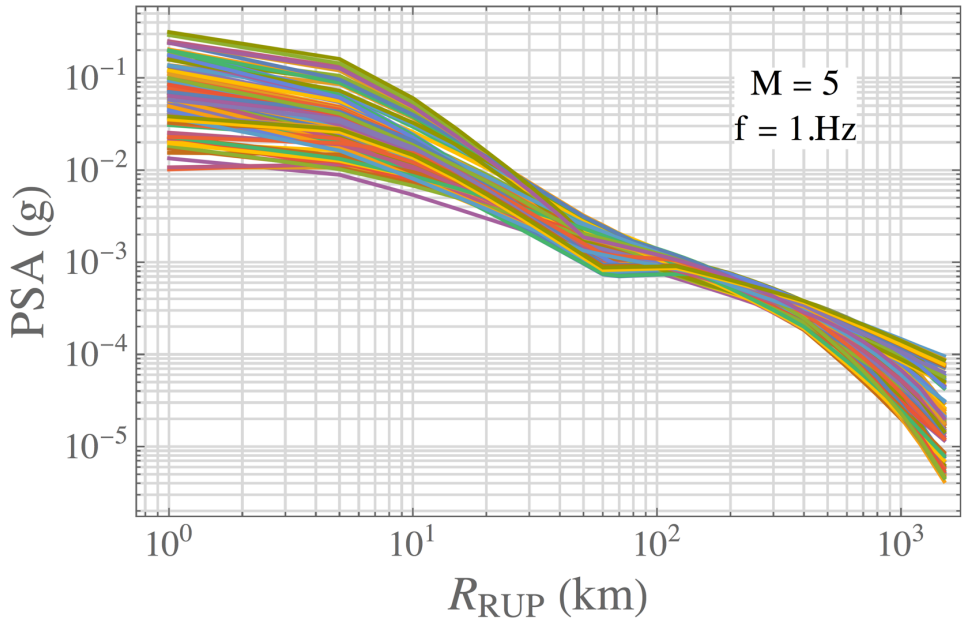


Figure 9–9 **Contour plots of mean residuals (left) and likelihood (right) for Sammon's maps calculated based on scenarios that cover only the range of the data (M = 4-6, RRUP = 10–400 km).**

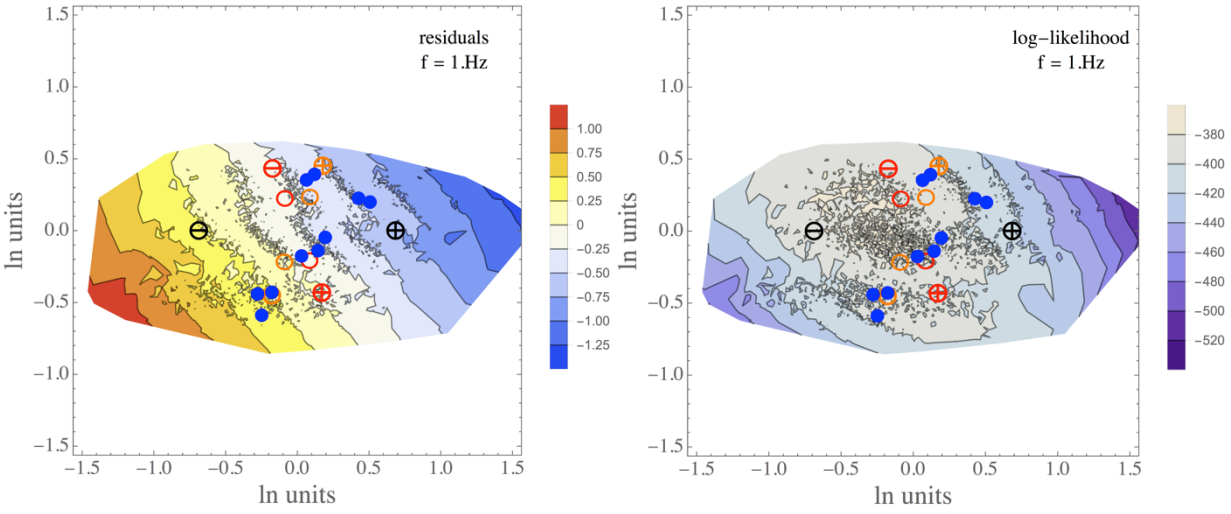


Figure 9–10 Weights based on the posterior distribution for the 1 Hz case.

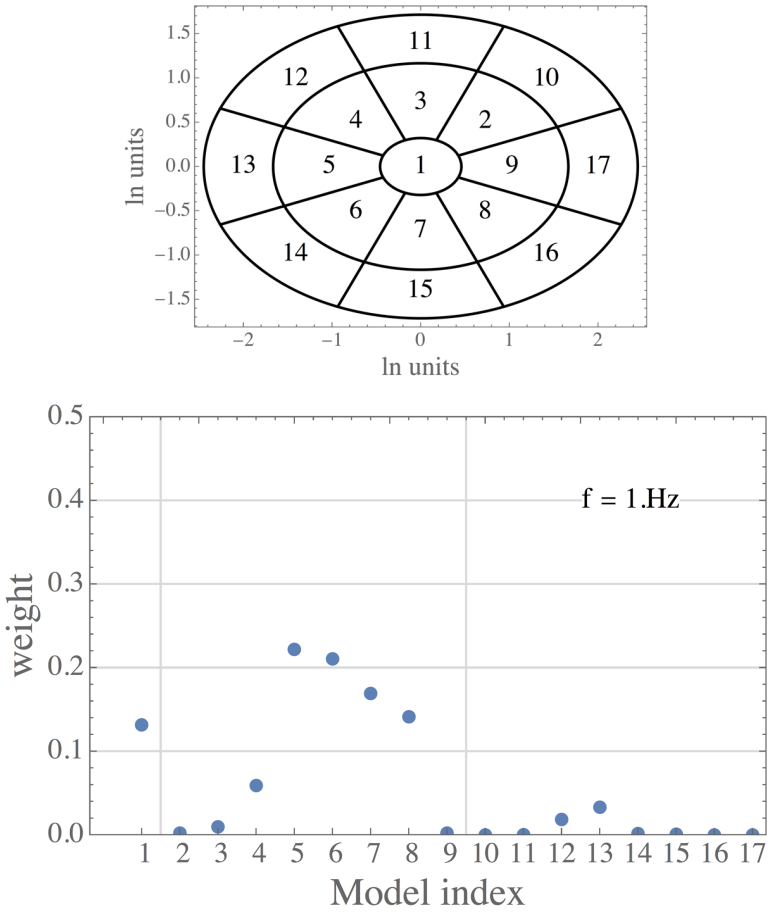


Figure 9–11(a) Total weights for frequencies 0.1–0.5 Hz.

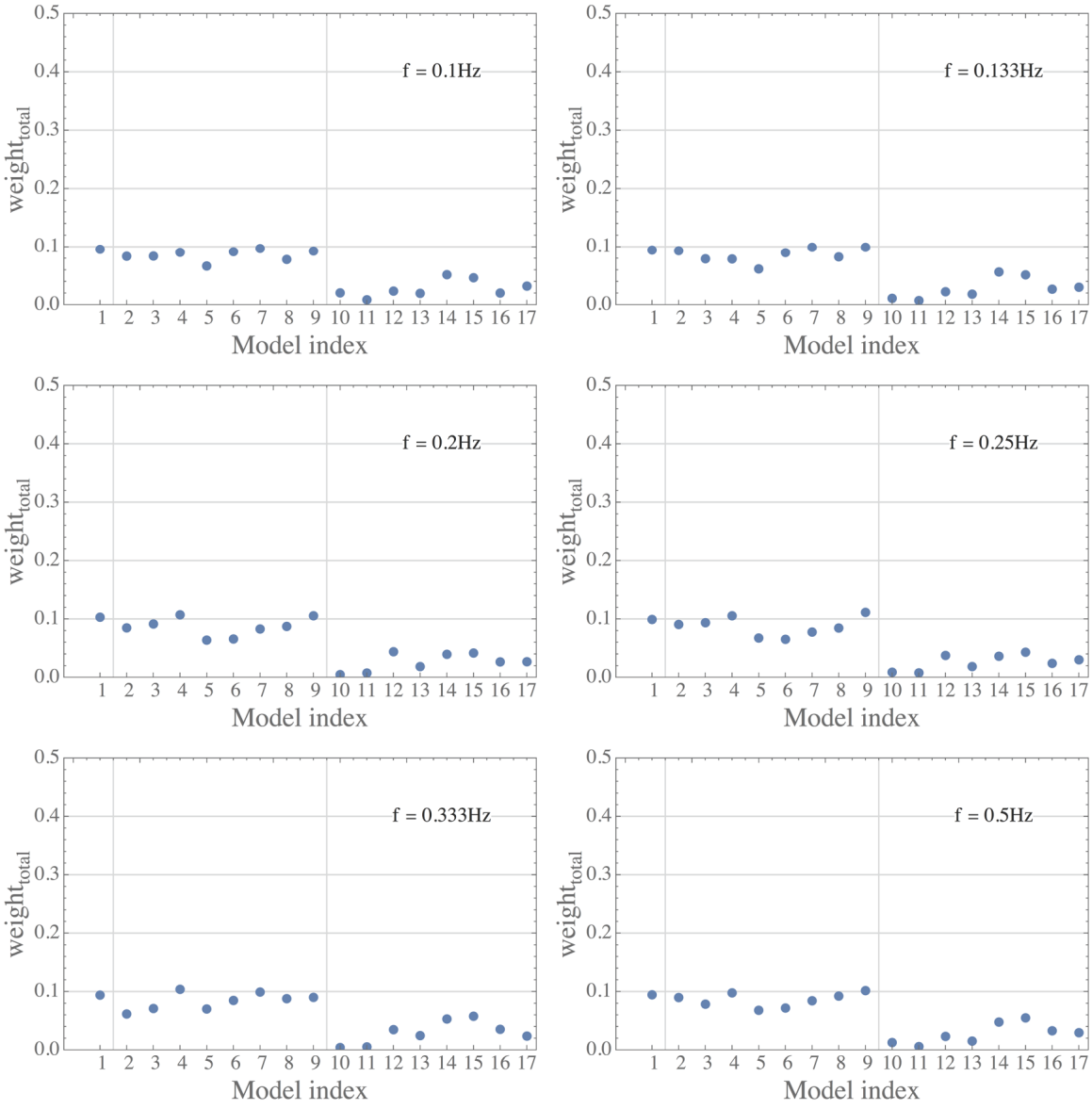


Figure 9–11(b) Total weights for frequencies 0.667–3.333 Hz.

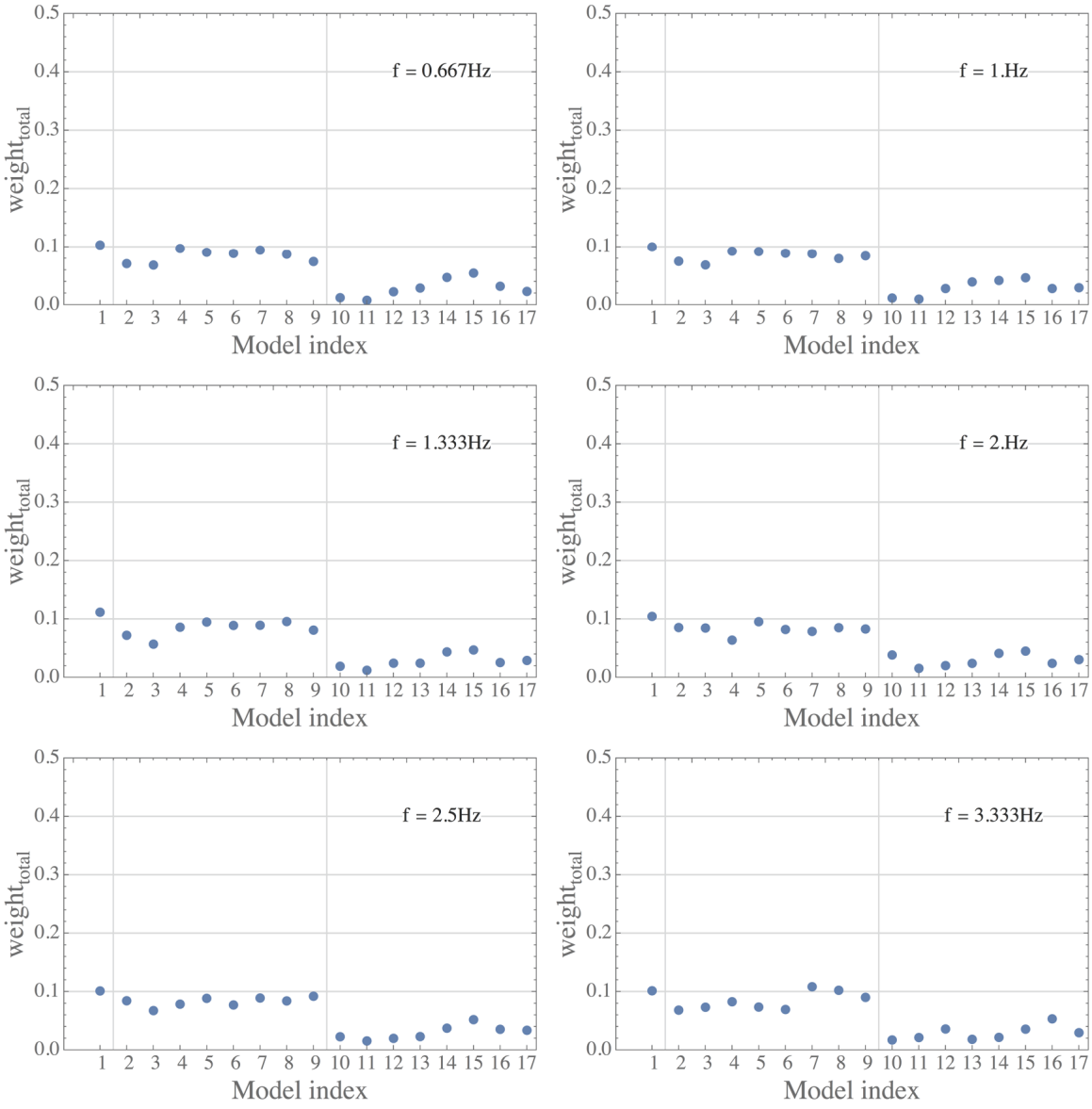


Figure 9–11(c) Total weights for frequencies 4–25 Hz.

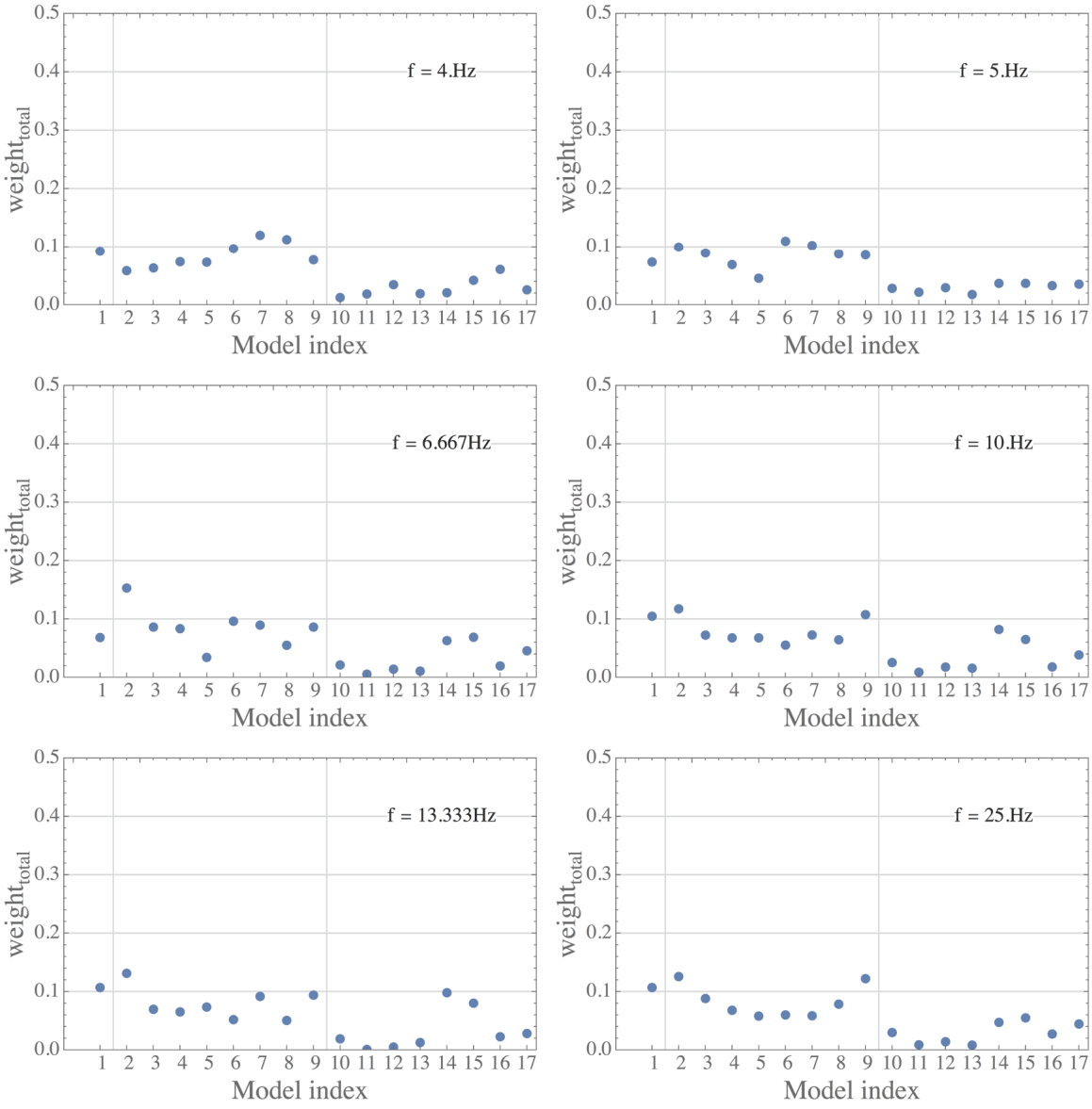


Figure 9–11(d) Total weights for frequencies 50–100 Hz and PGA.

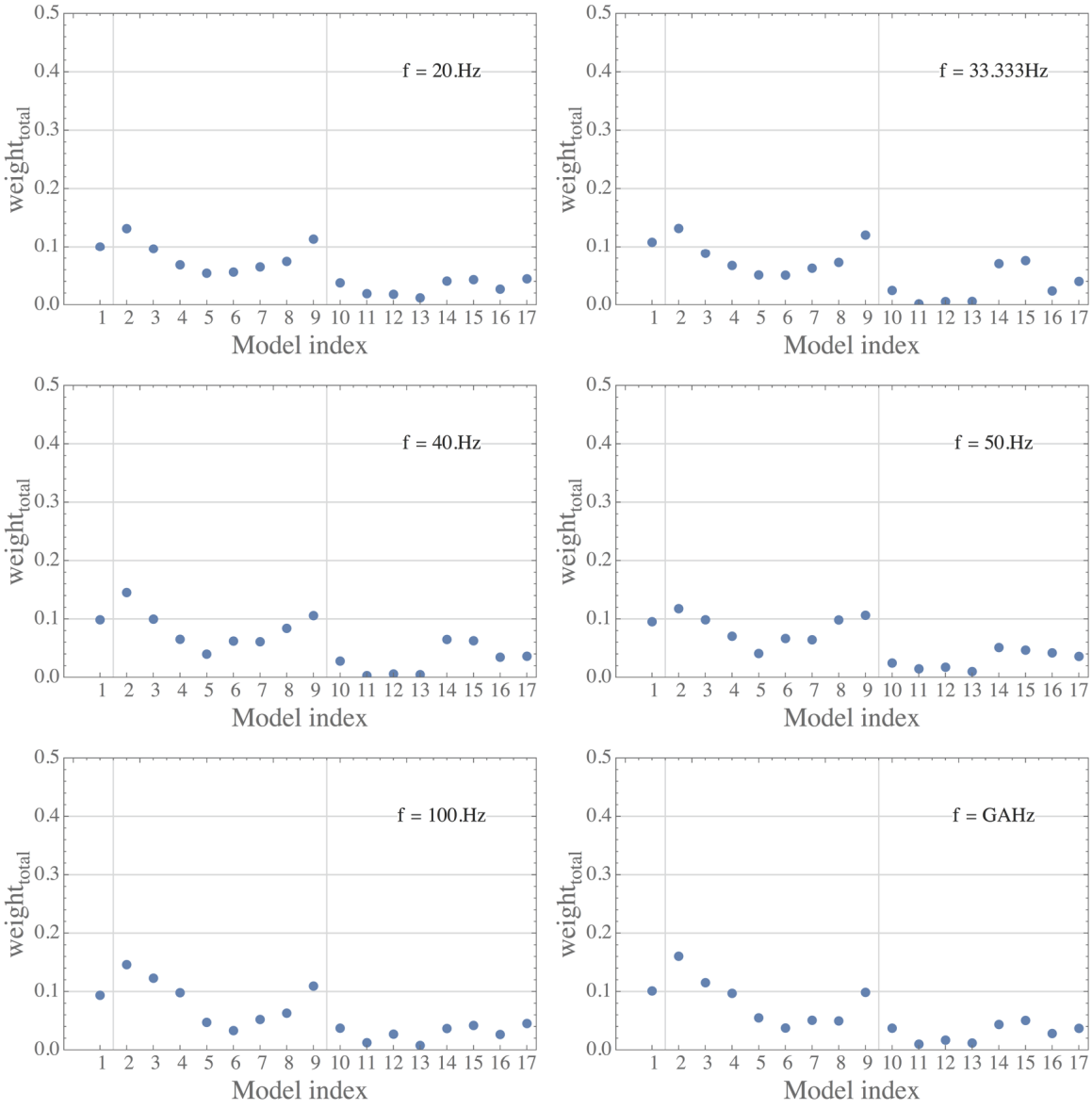


Figure 9–11(e) Total weights for PGV.

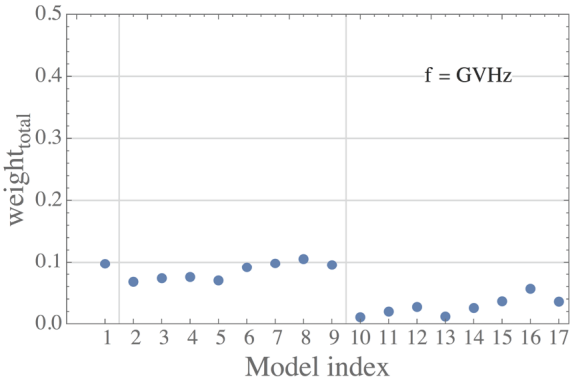


Figure 9–12(a) Total weights against frequency for the final GMMs (1-6).

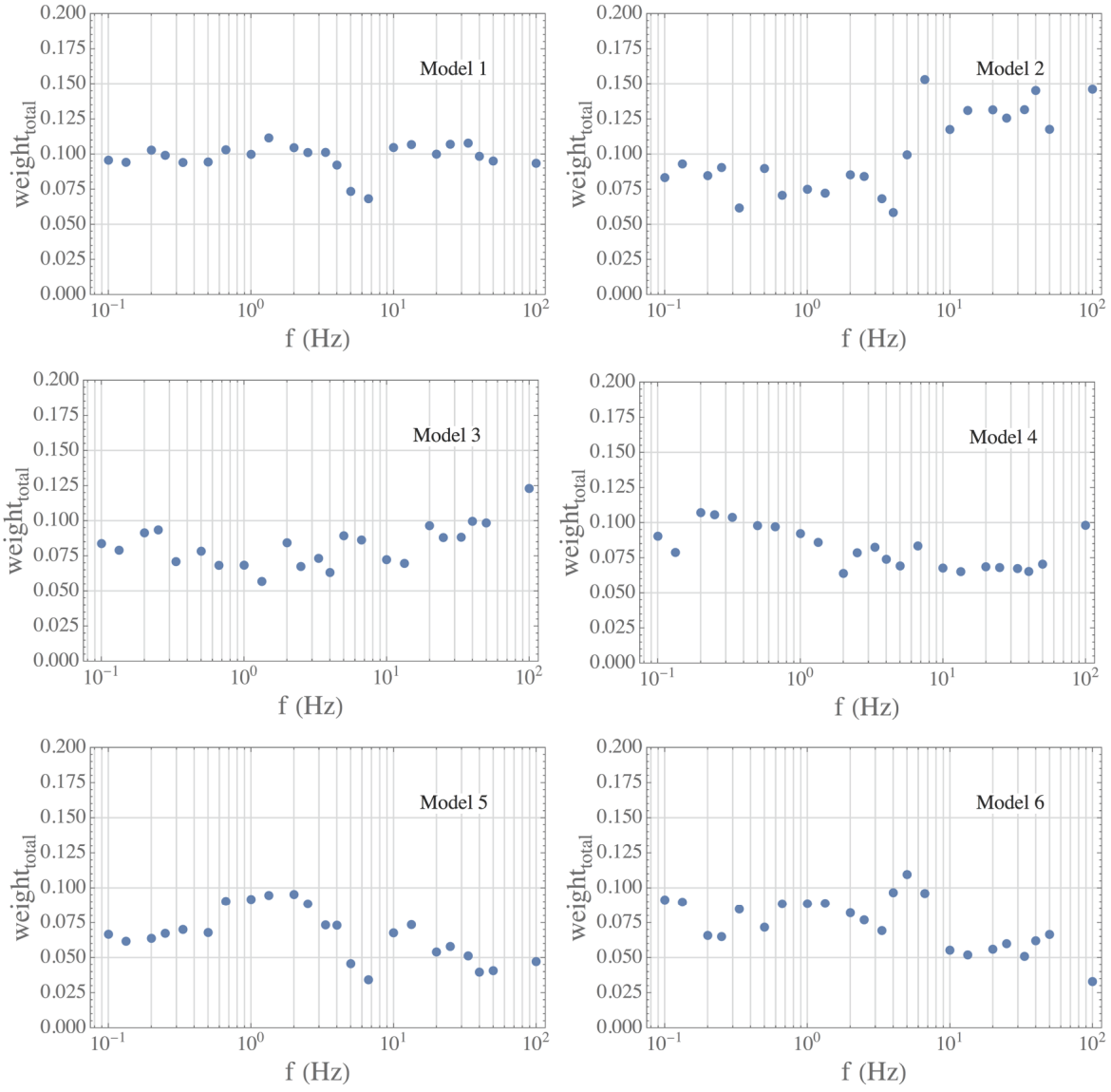


Figure 9–12(b) Total weights against frequency for the final GMMs (7-12).

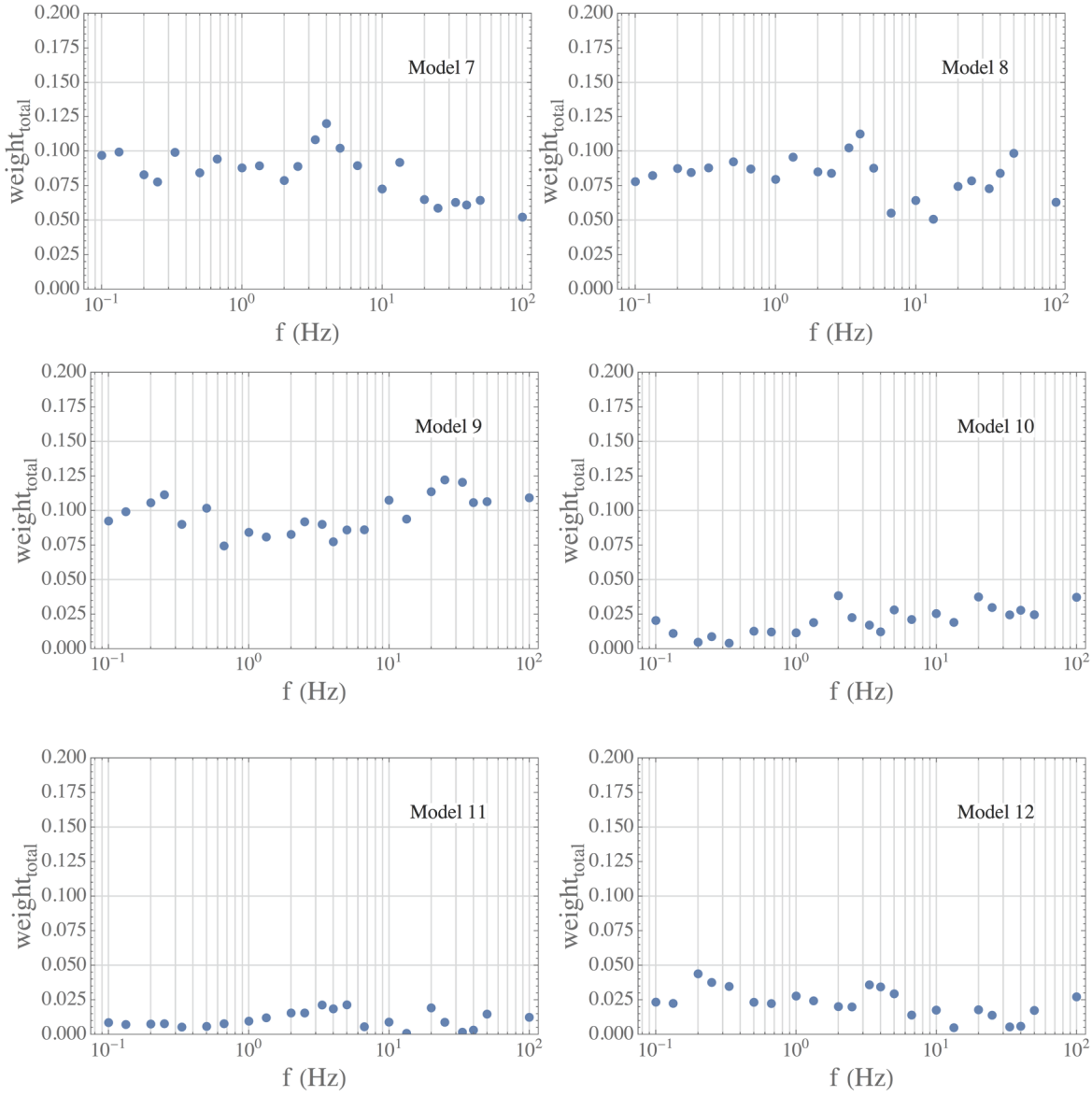


Figure 9–12(c) Total weights against frequency for the final GMMs (13-17).

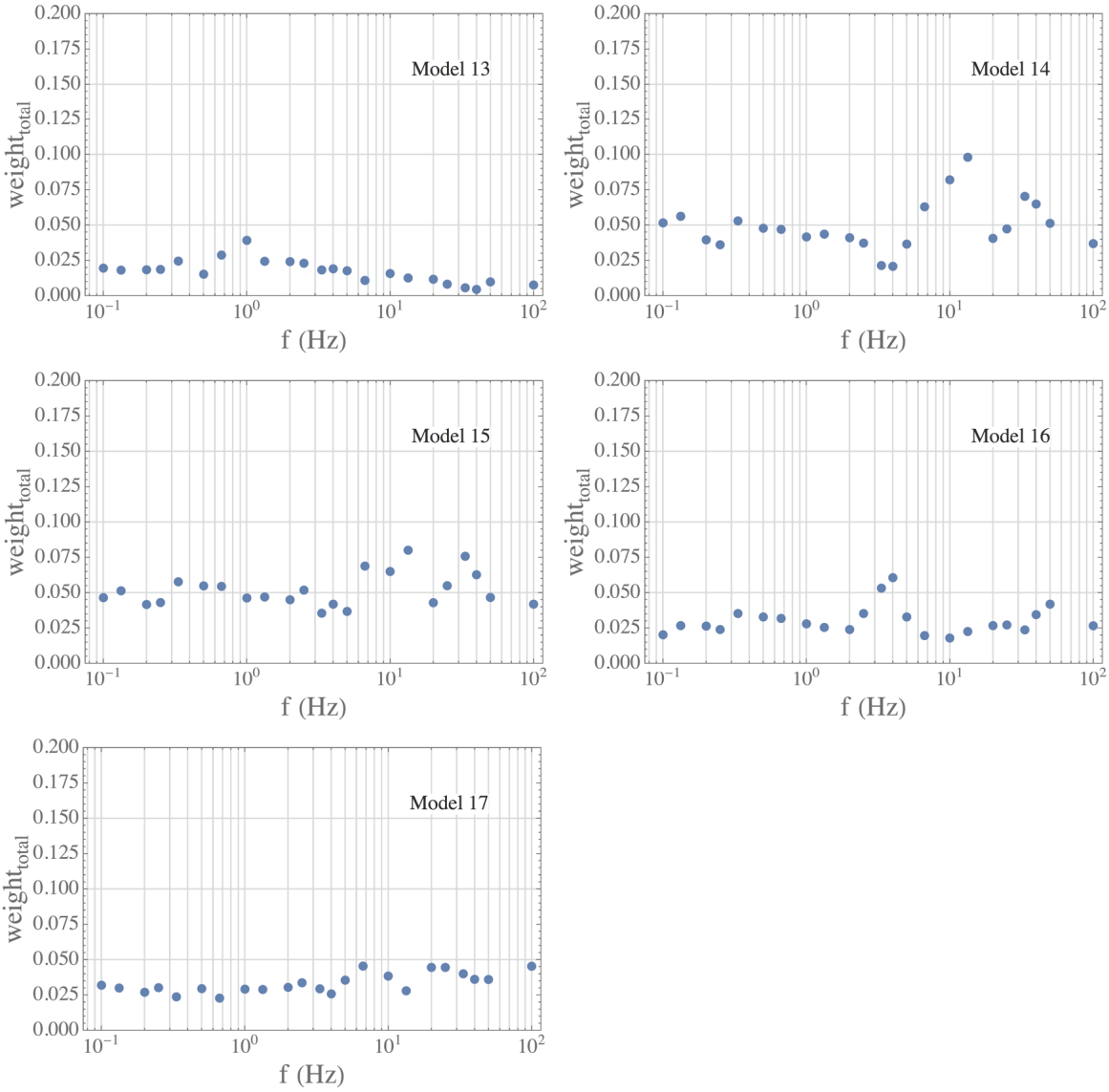


Figure 9–13 Plots of the cumulative distribution function of the NGA-East models based on total weights (black) and the seed models (blue).

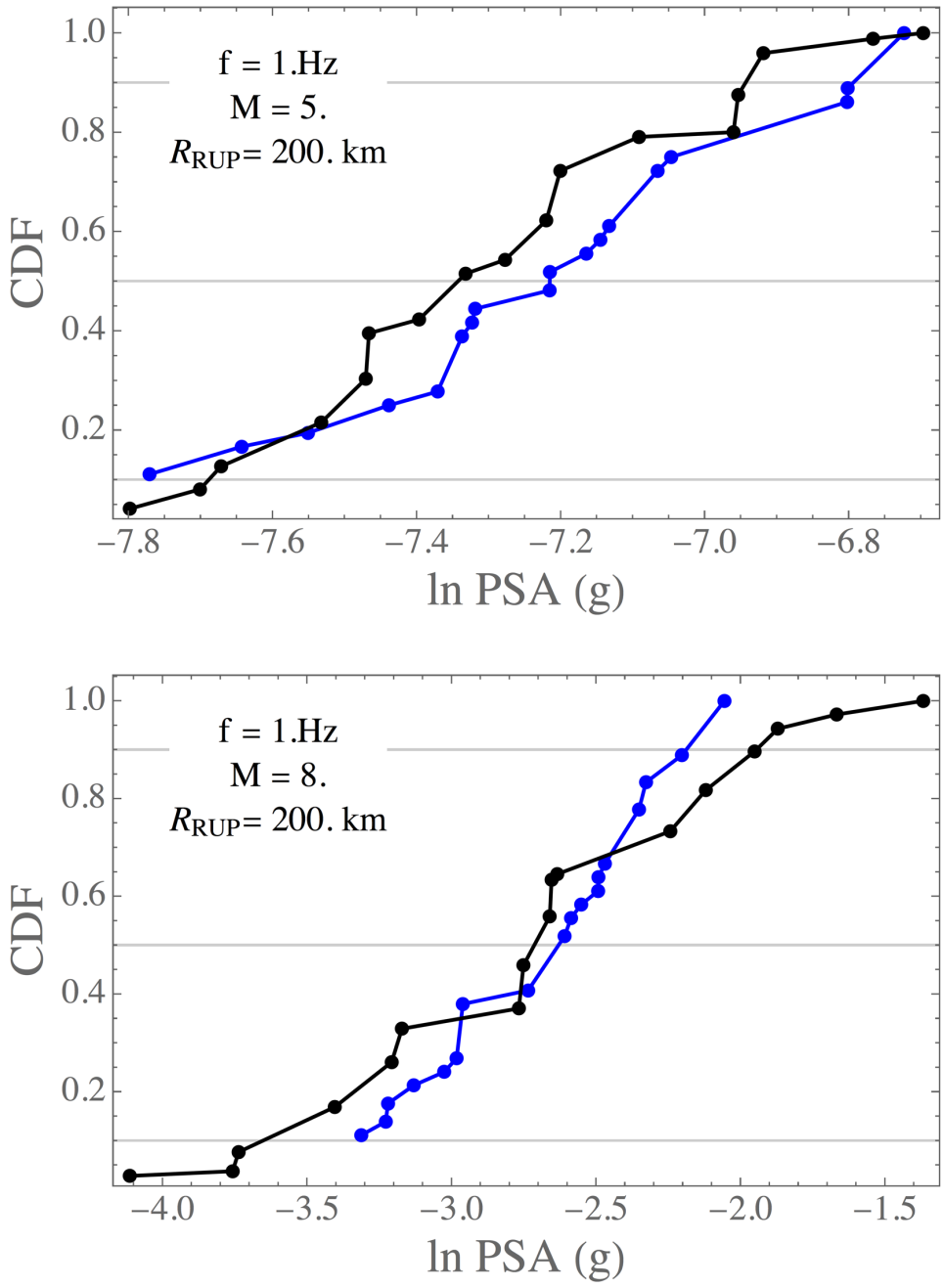


Figure 9–14(a) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 1 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

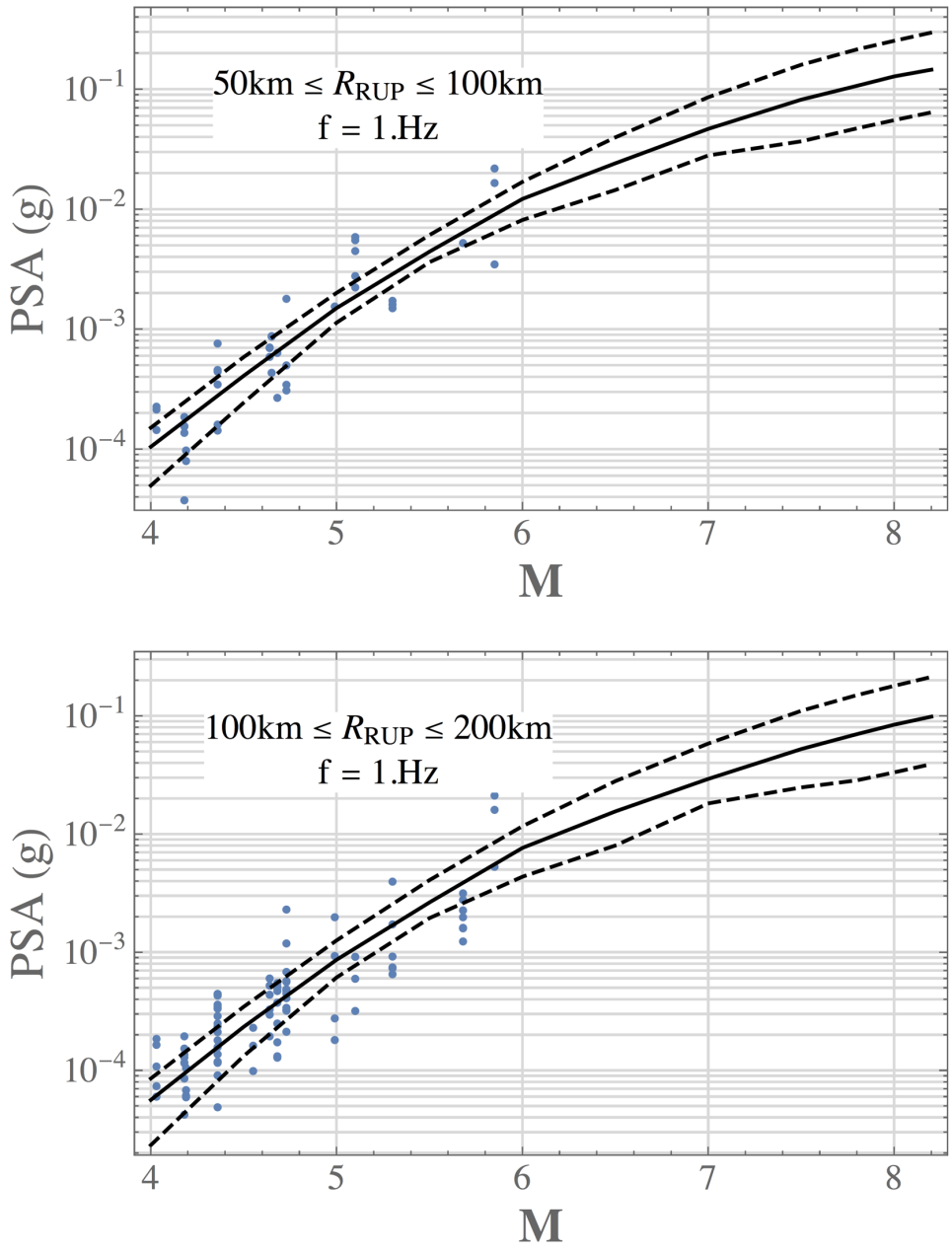


Figure 9–14(b) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 1.333 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

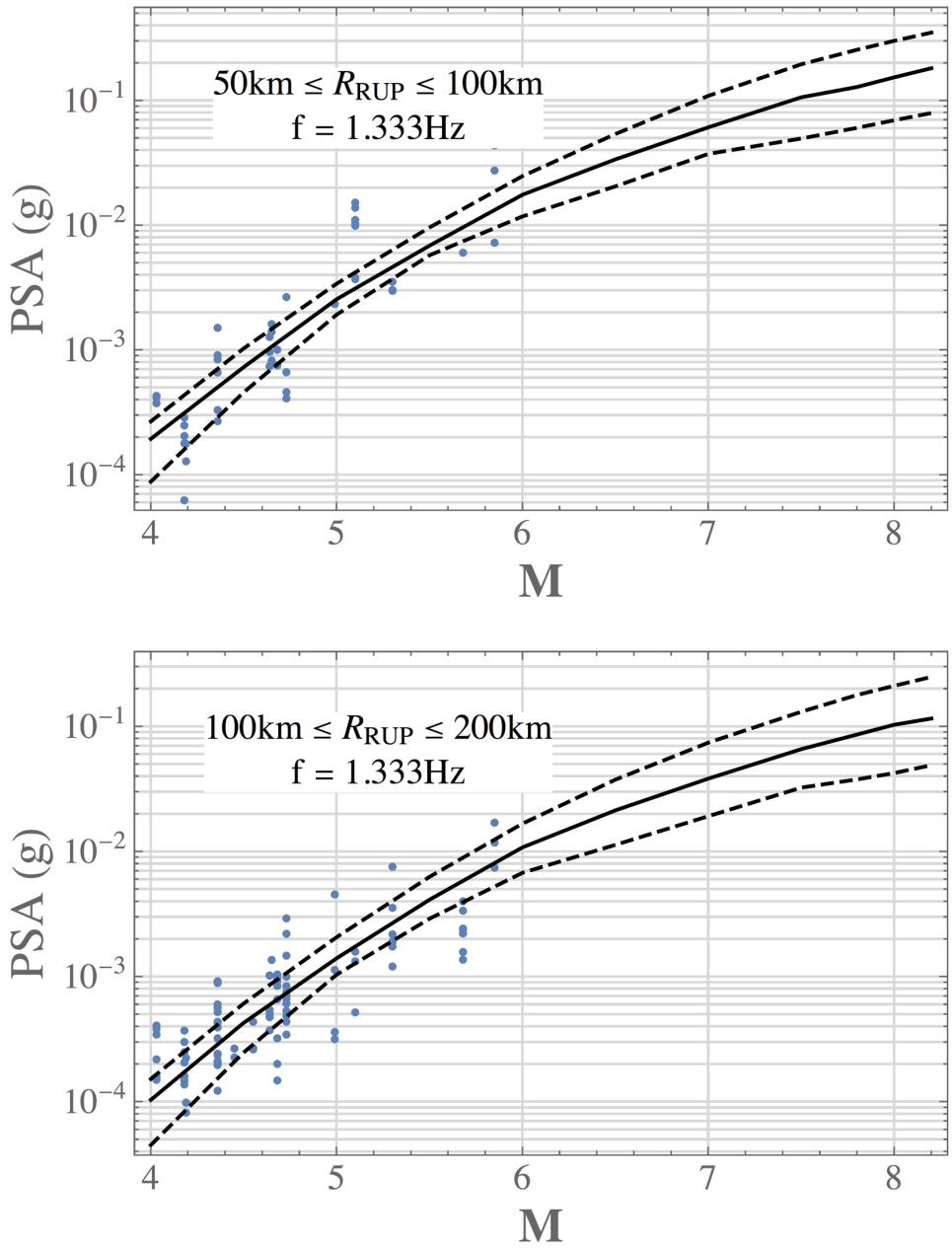


Figure 9–14(c) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 2 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

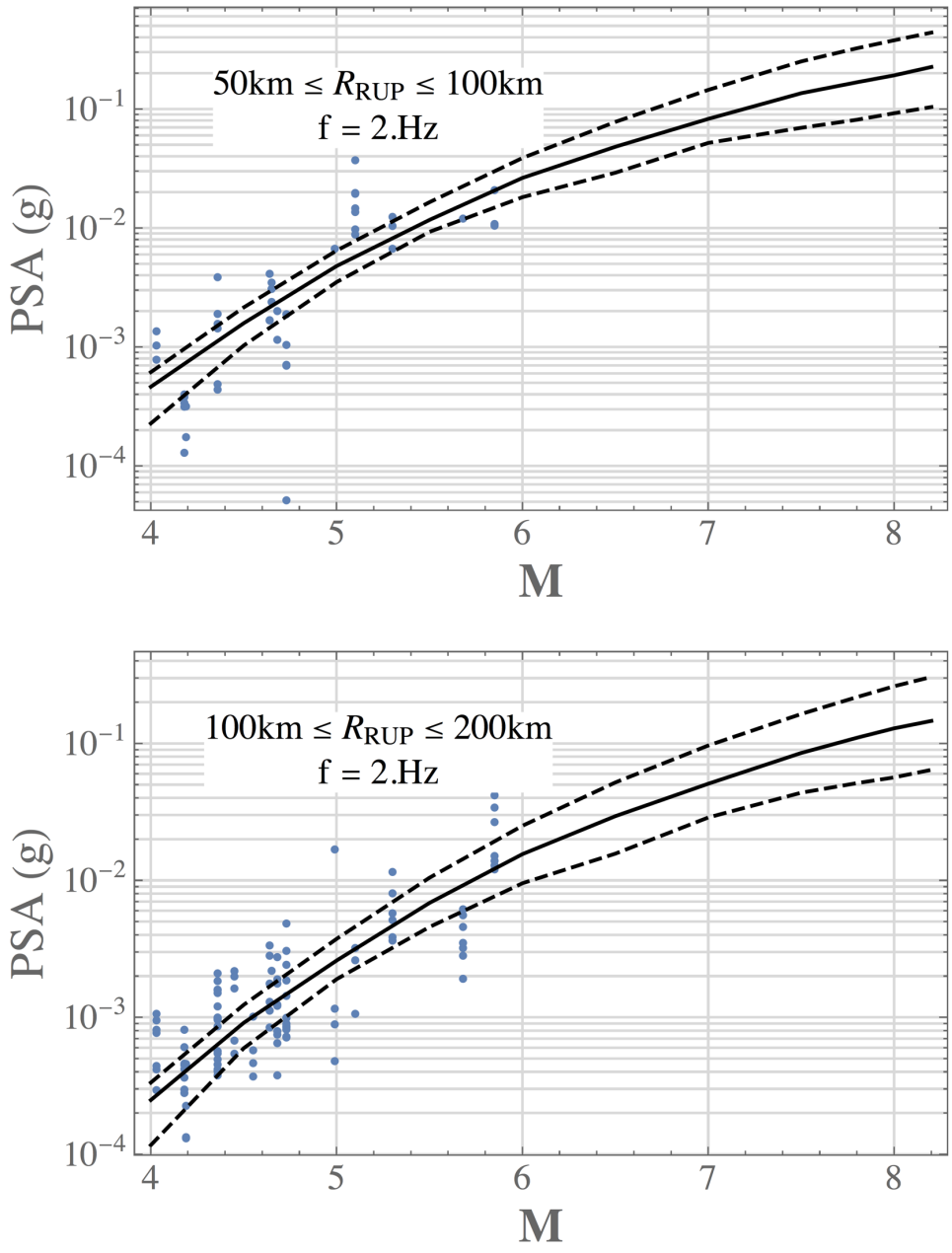


Figure 9–14(d) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 2.5 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

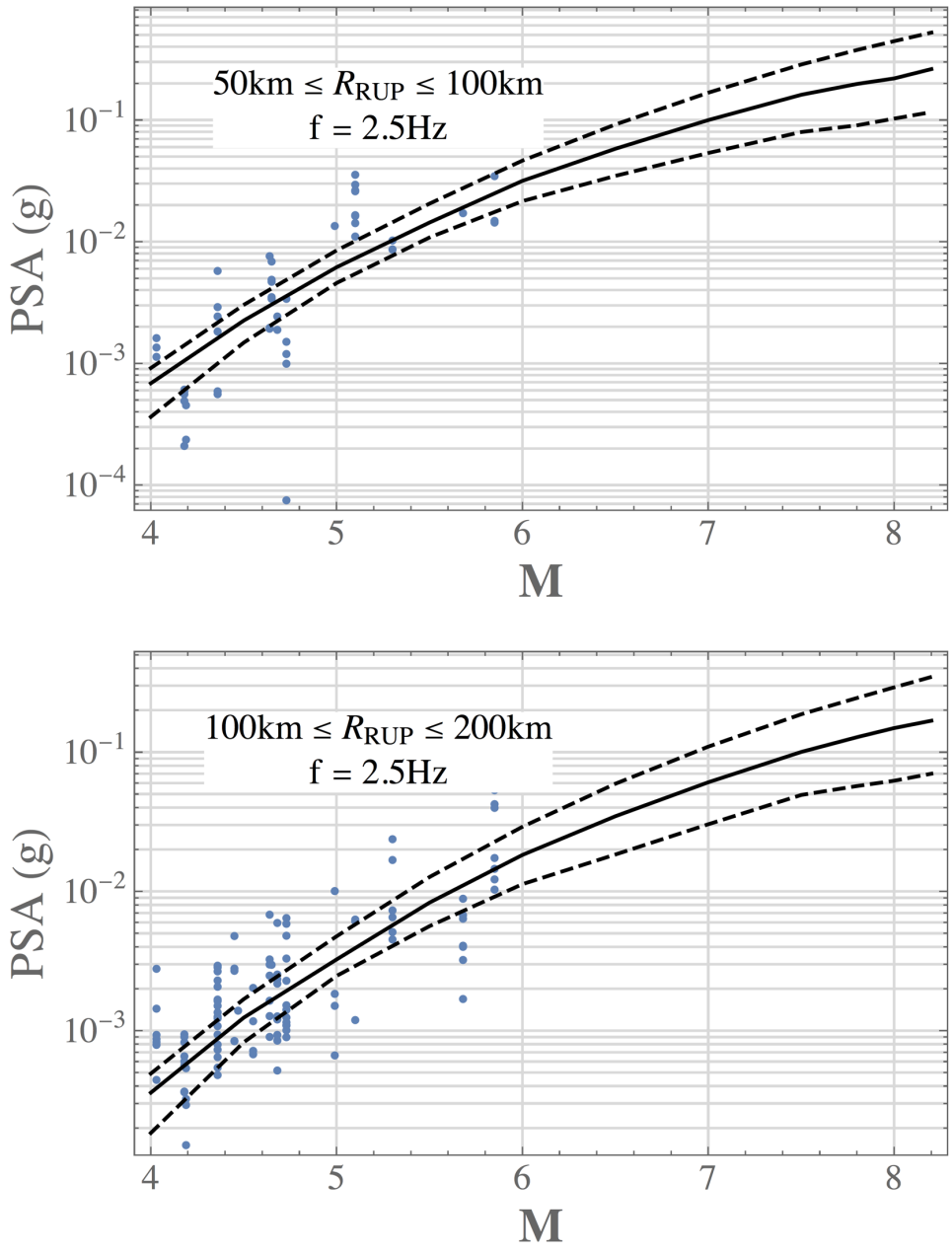


Figure 9–14(e) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 3.333 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

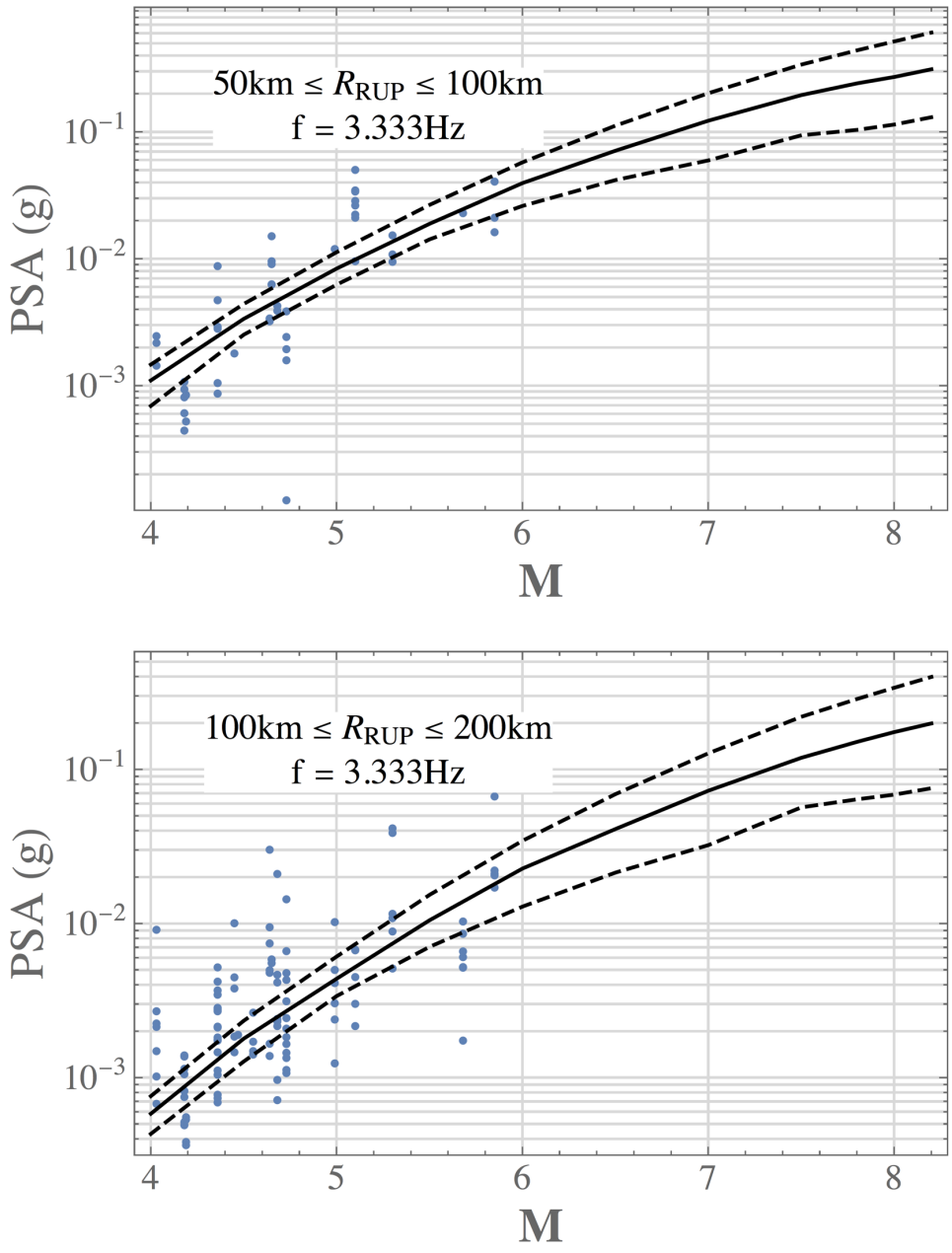


Figure 9–14(f) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 4 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

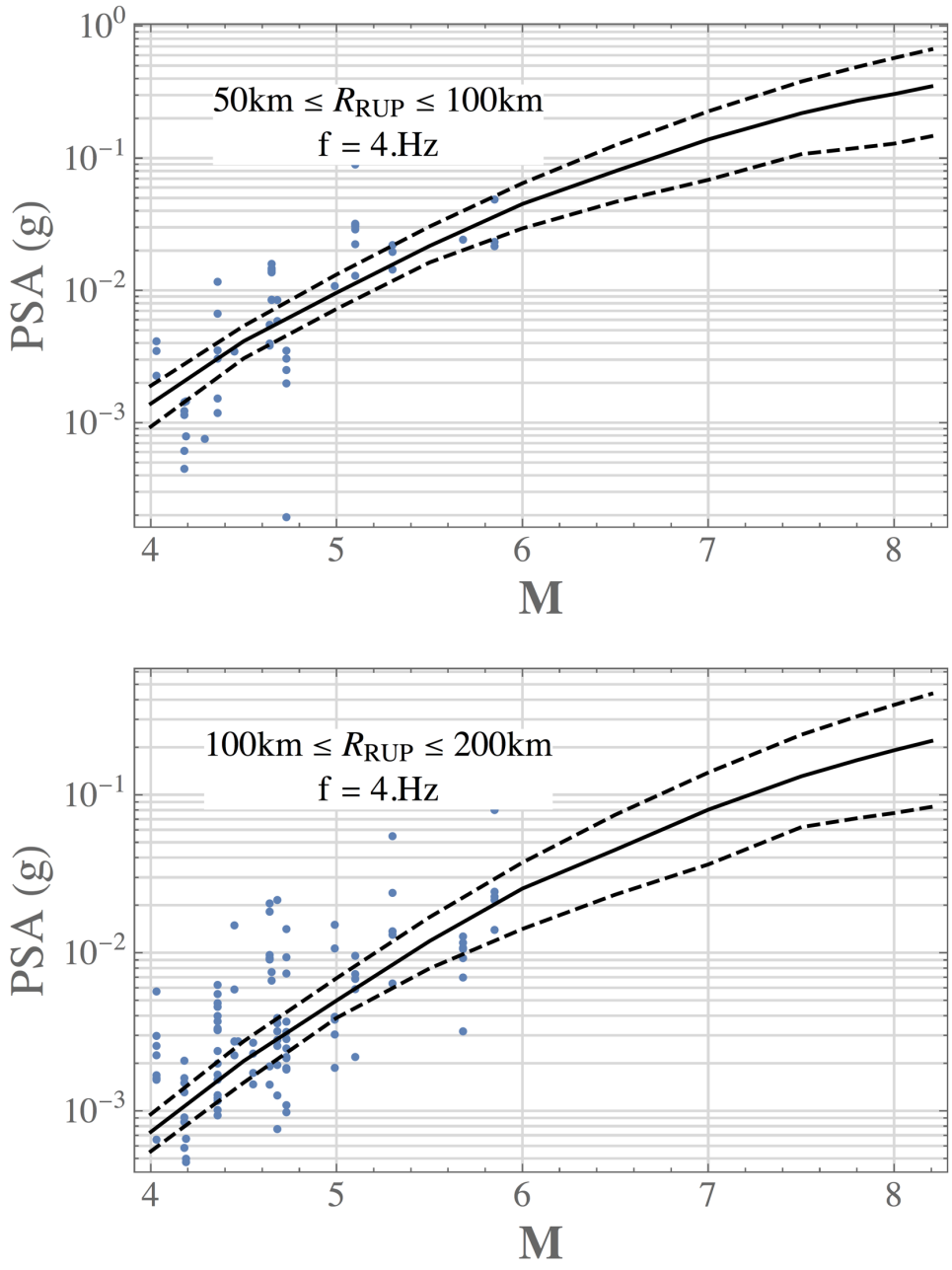


Figure 9–14(g) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 5 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

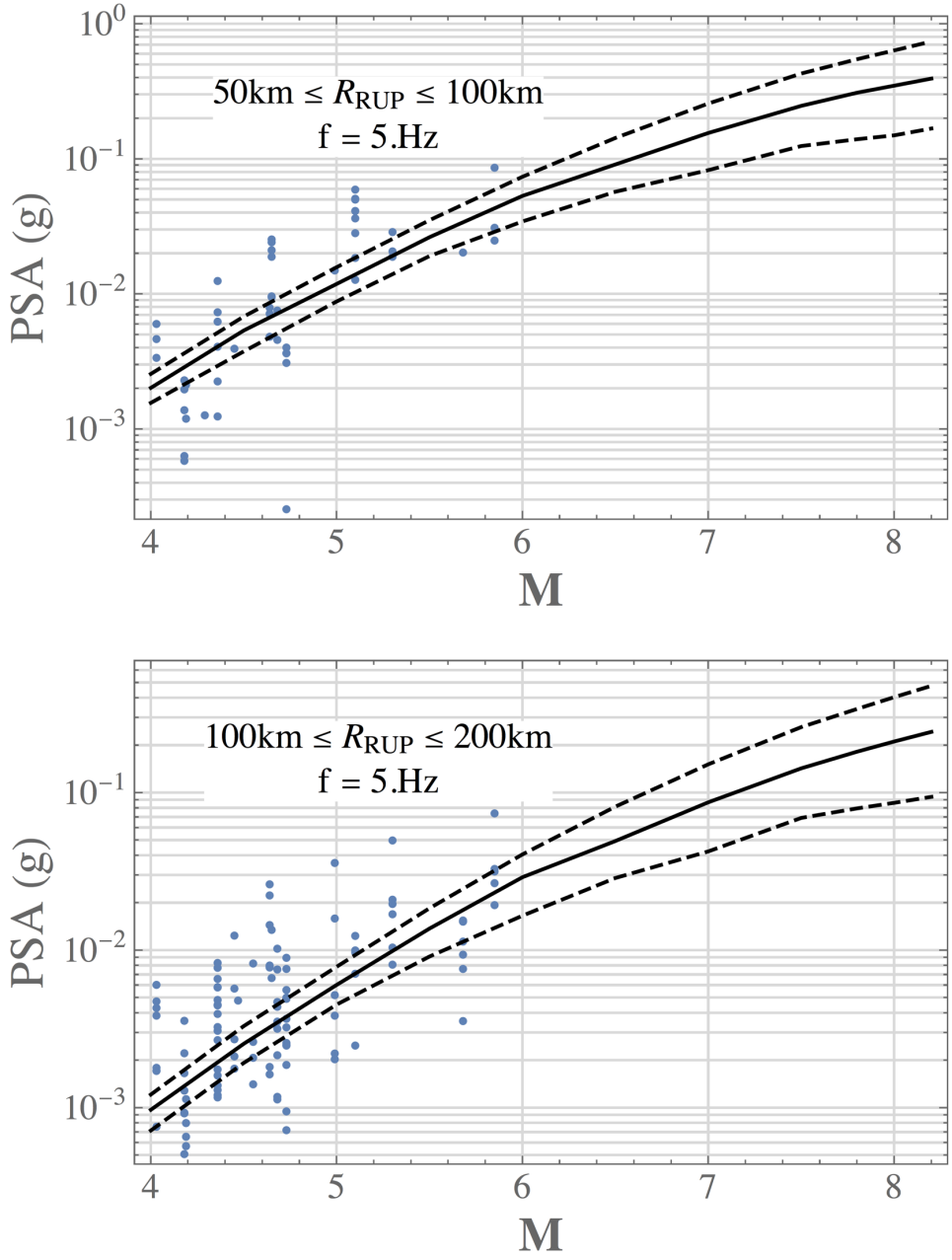


Figure 9–14(h) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 6.667 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

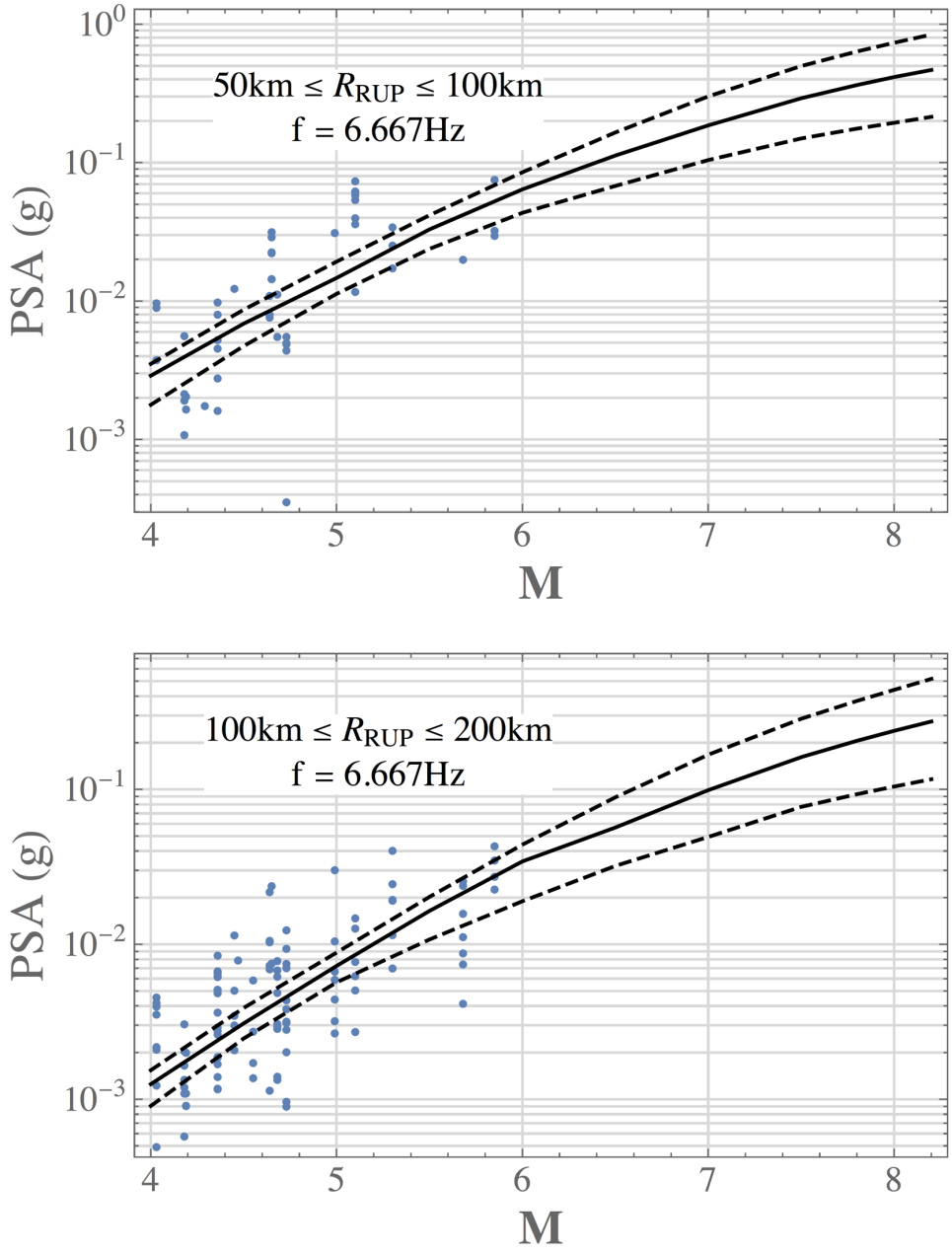


Figure 9–14(i) Magnitude and amplitude scaling of 10th, 50th, and 90th fractiles of the NGA-East final GMM distribution against observed ground-motion data (aggregated in one distance bin) for 10 Hz. The NGA-East final GMM predictions are calculated at 70 km (top) and 150 km (bottom).

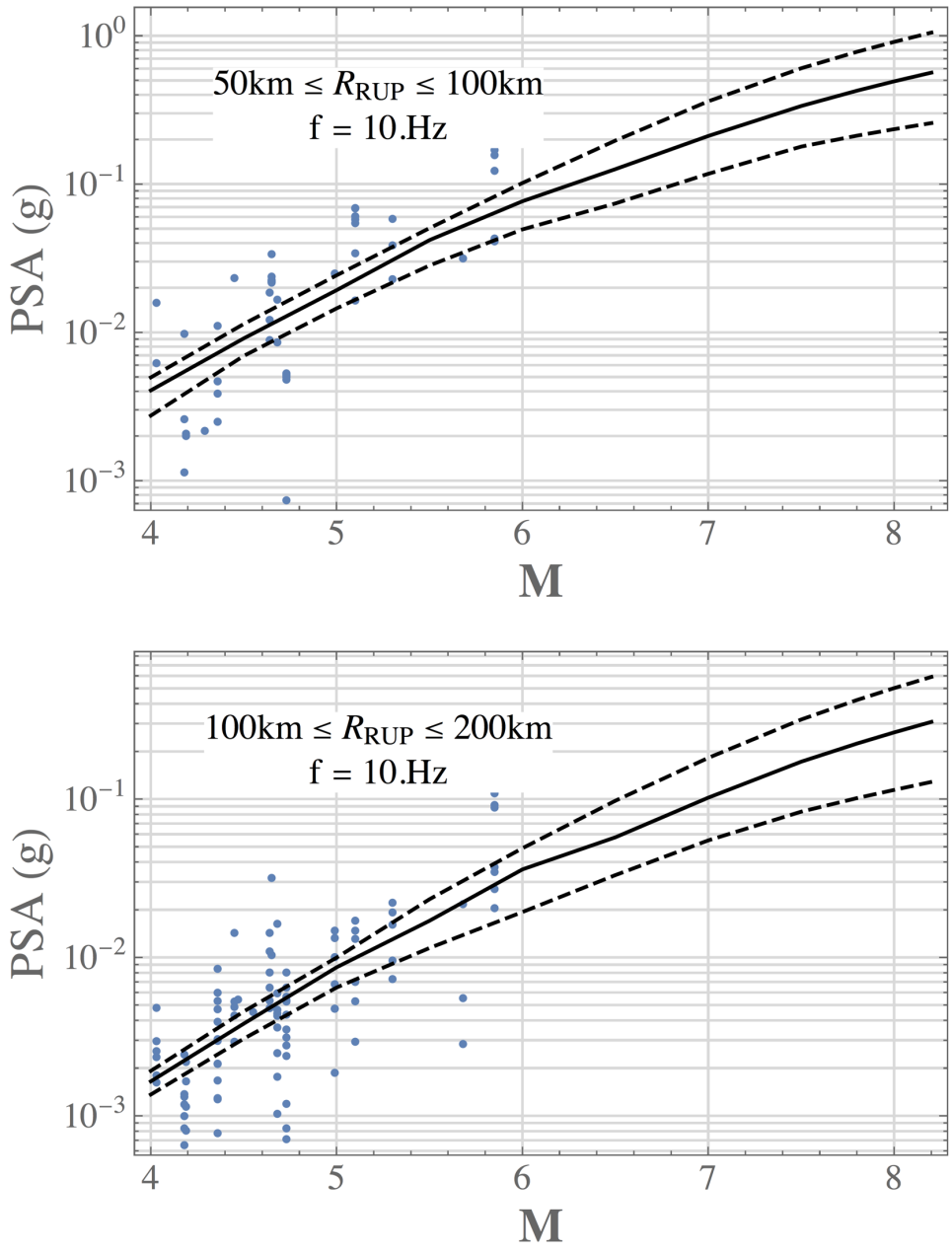


Figure 9–15(a) Hazard curves from original weighted seed GMMs using as-is distance measures (R_{JB} or R_{RUP}) and from the final NGA-East GMMs at Manchester, 1 Hz.

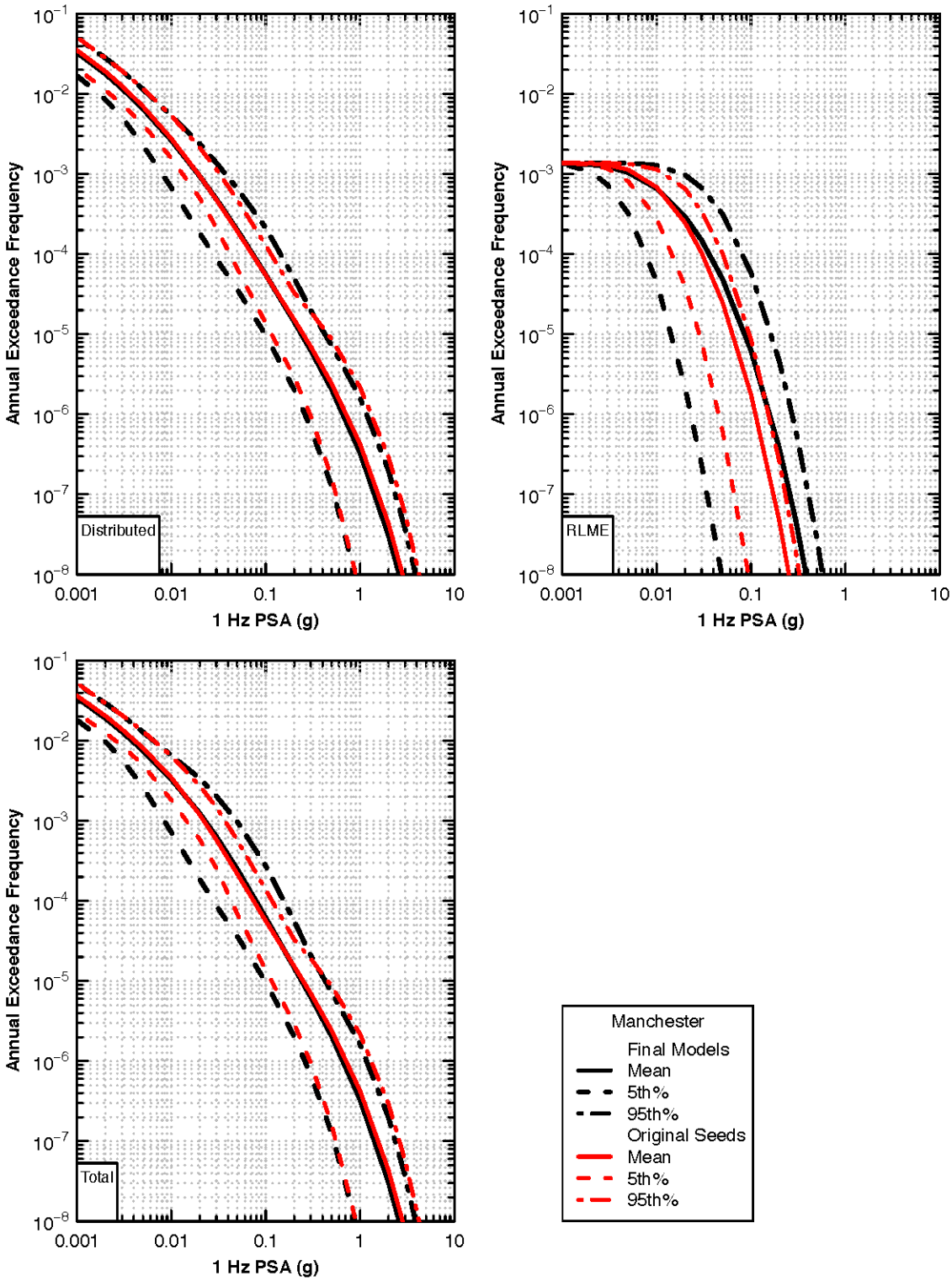


Figure 9–15(b) Hazard curves from original weighted seed GMMs using as-is distance measures (R_{JB} or R_{RUP}) and from the final NGA-East GMMs at Manchester, 10 Hz.

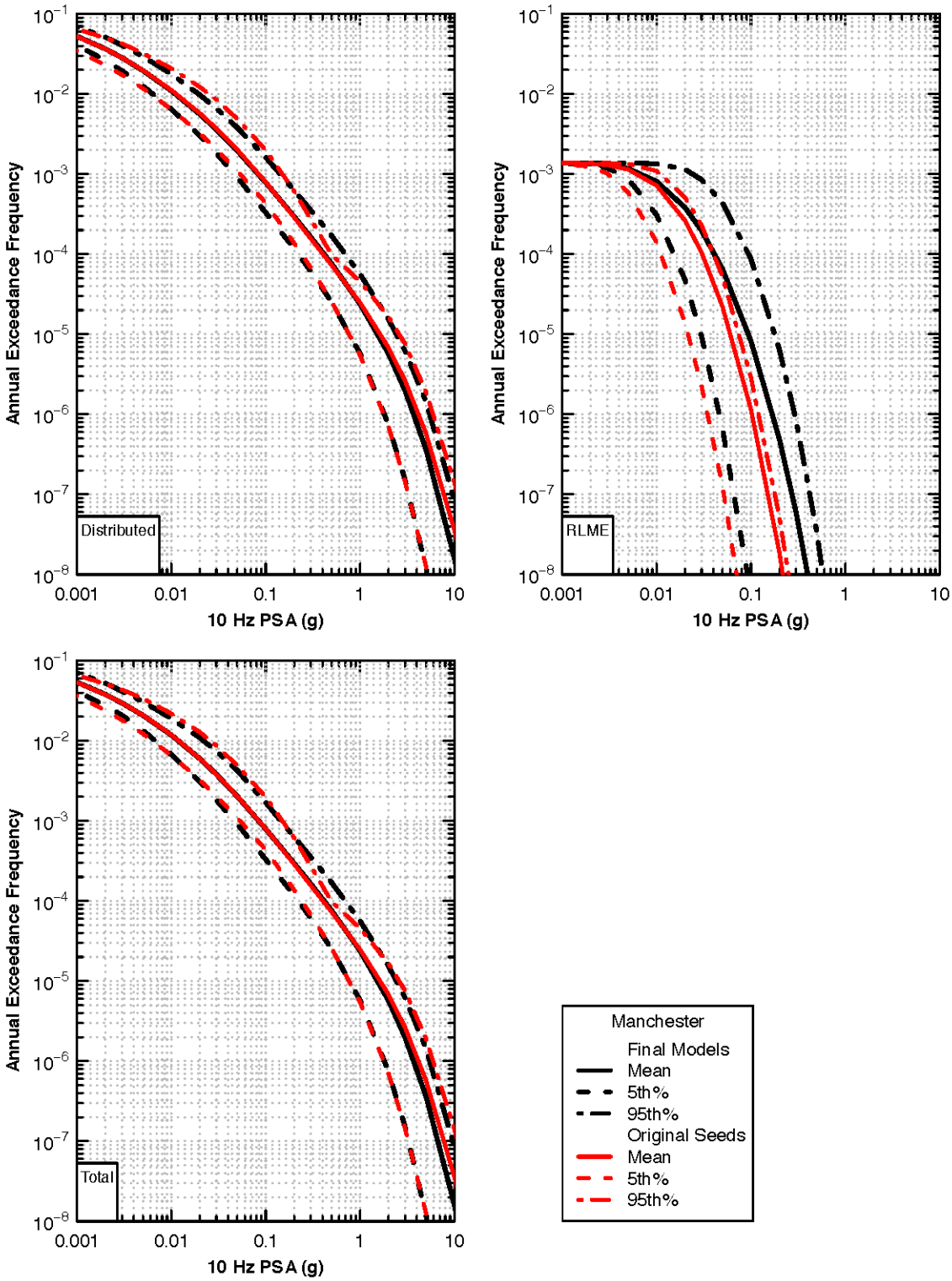
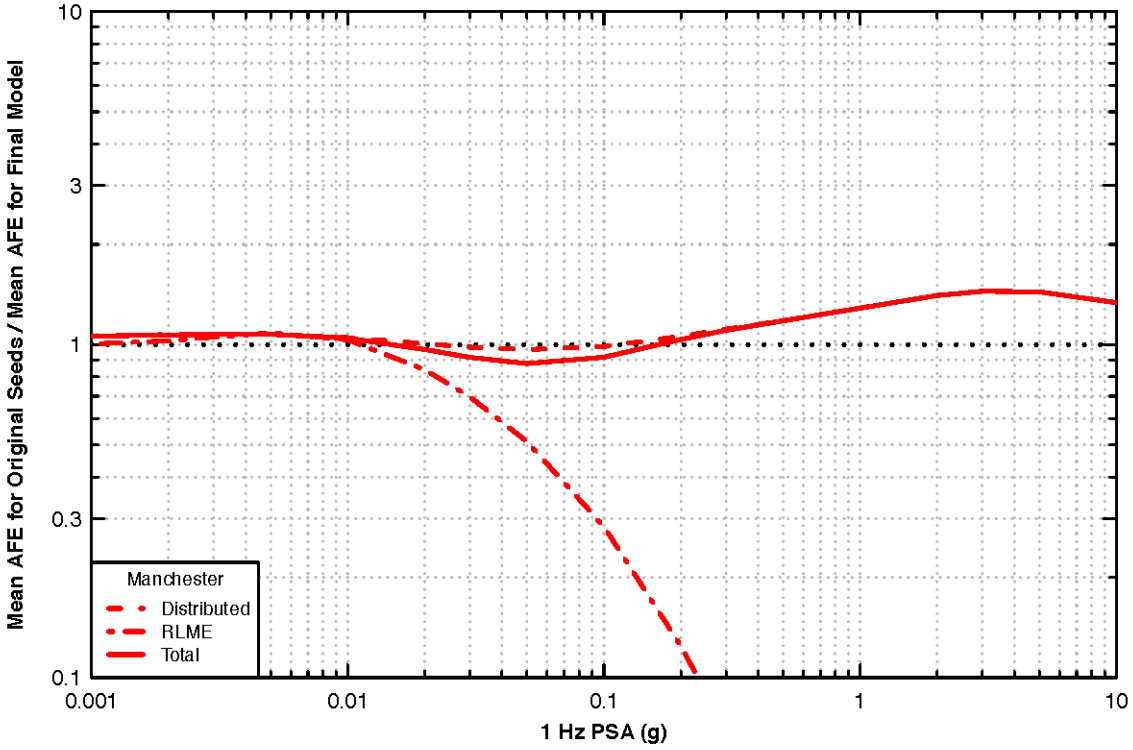


Figure 9–15(c) Hazard ratio from original weighted seed GMMs seed relative to that from the final NGA-East GMMs at Manchester (1 Hz and 10 Hz).



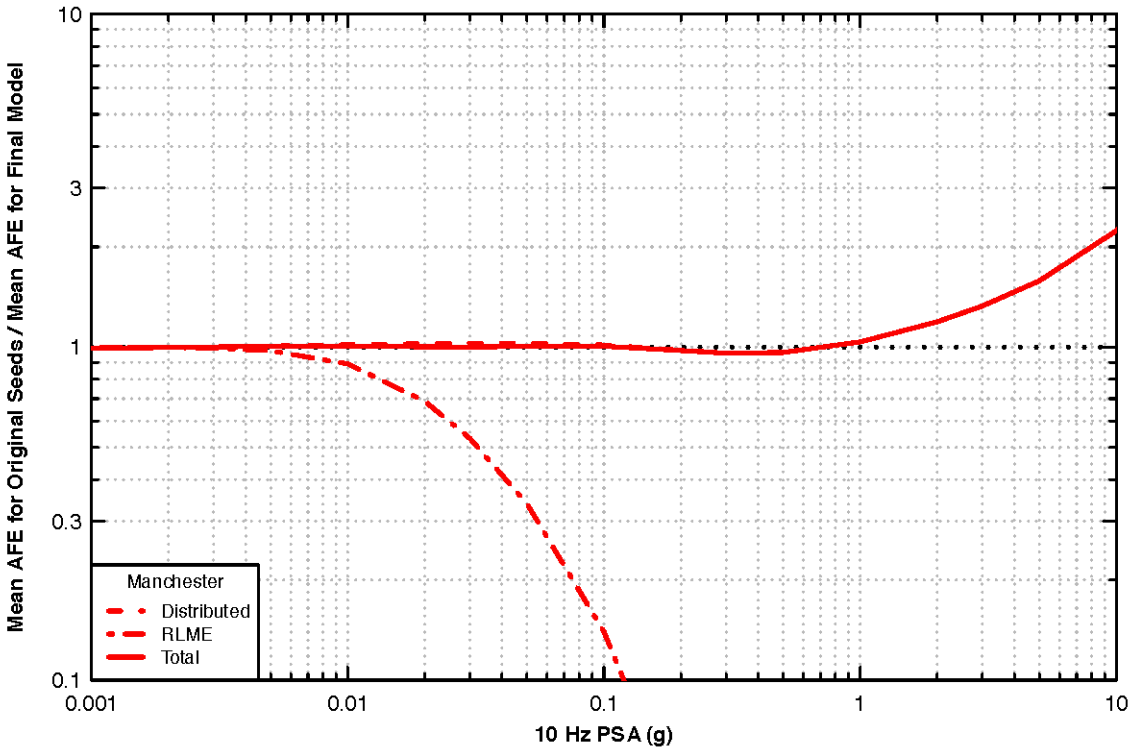


Figure 9–16(a) Hazard curves from original weighted seed GMMs using as-is distance measures (R_{JB} or R_{RUP}) and from the final NGA-East GMMs at Central Illinois, 1 Hz.

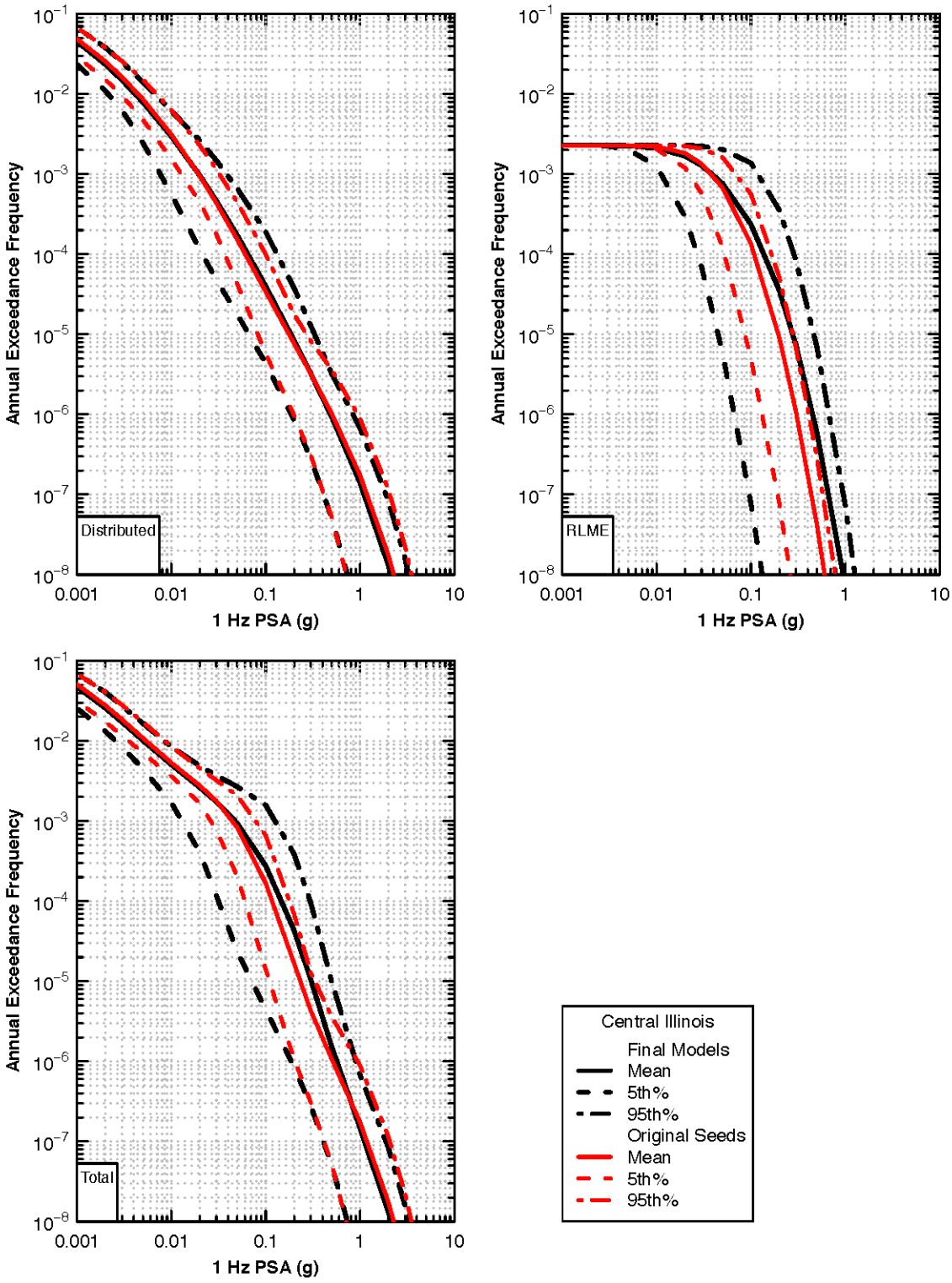


Figure 9–16(b) Hazard curves from original weighted seed GMMs using as-is distance measures (R_{JB} or R_{RUP}) and from the final NGA-East GMMs at Central Illinois, 10 Hz.

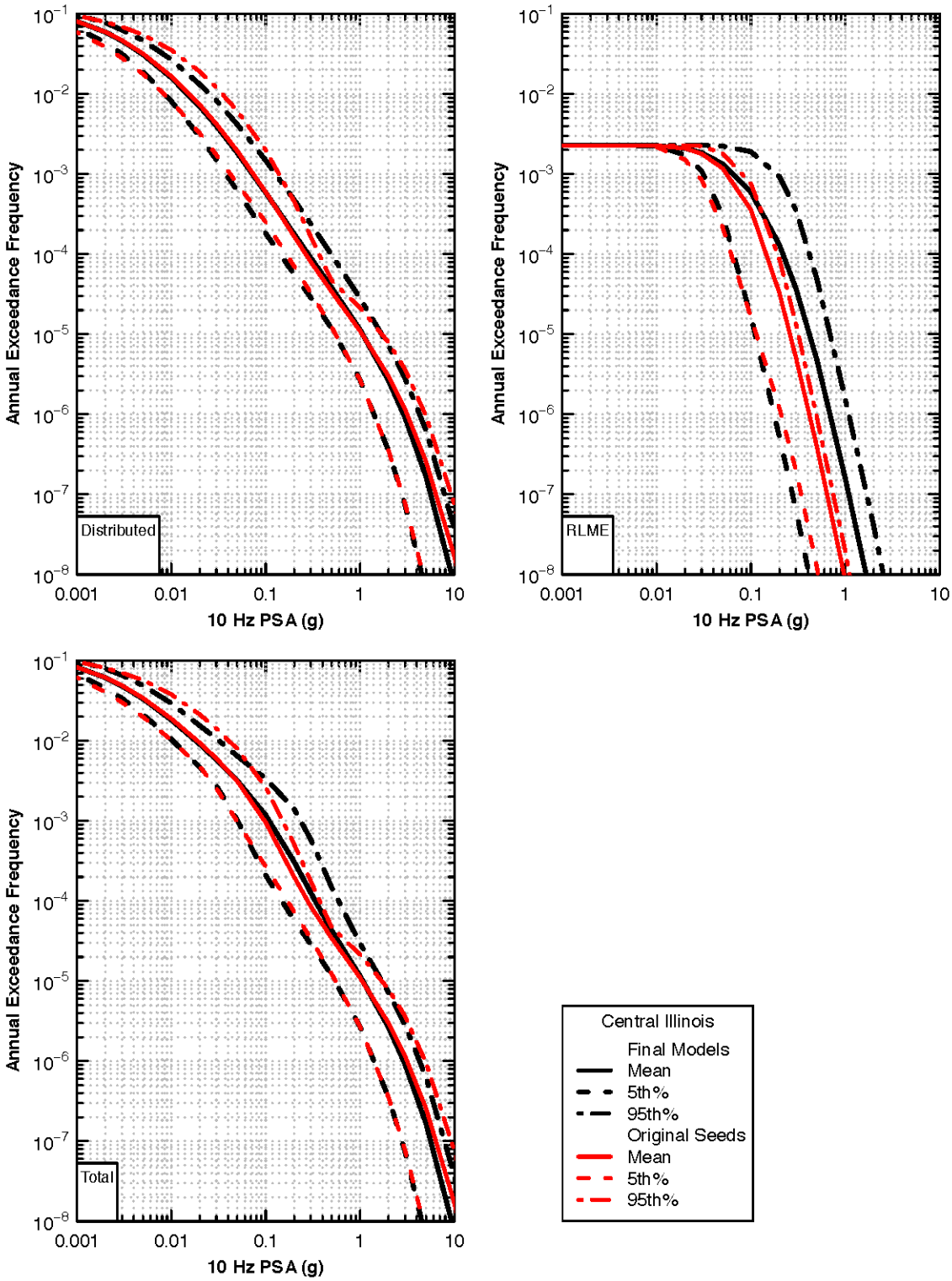
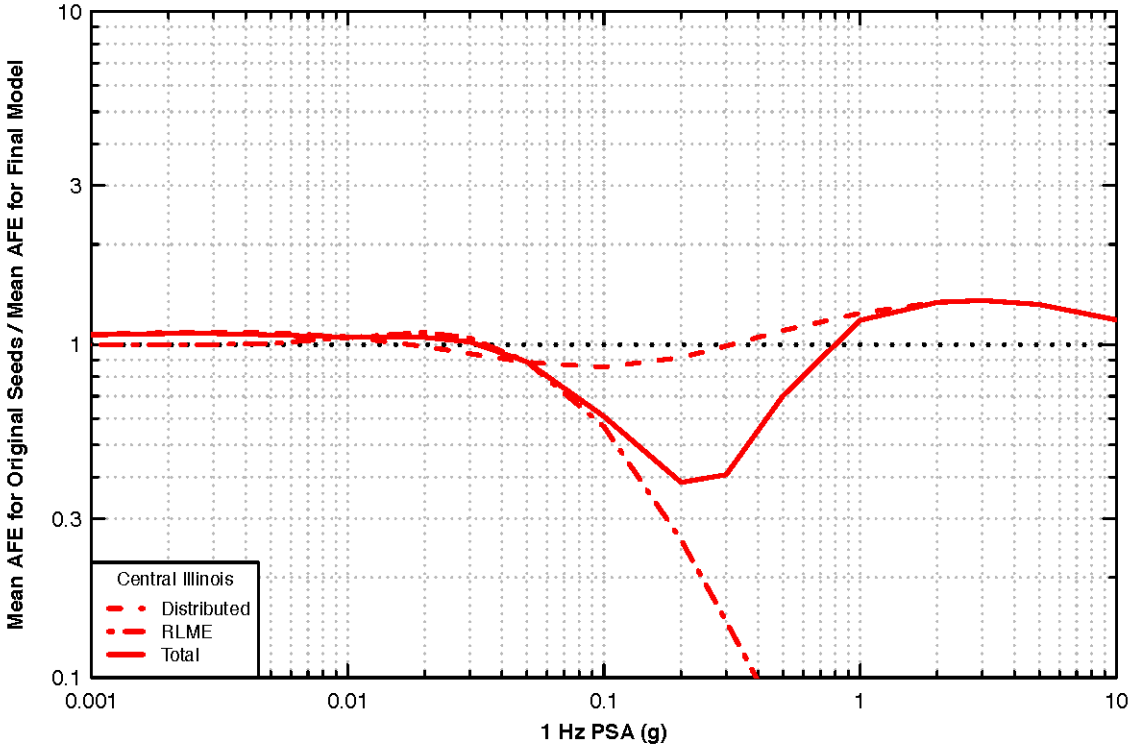


Figure 9–16(c) Hazard ratio from original weighted seed GMMs seed relative to that from the final NGA-East GMMs at Central Illinois (1 Hz and 10 Hz).



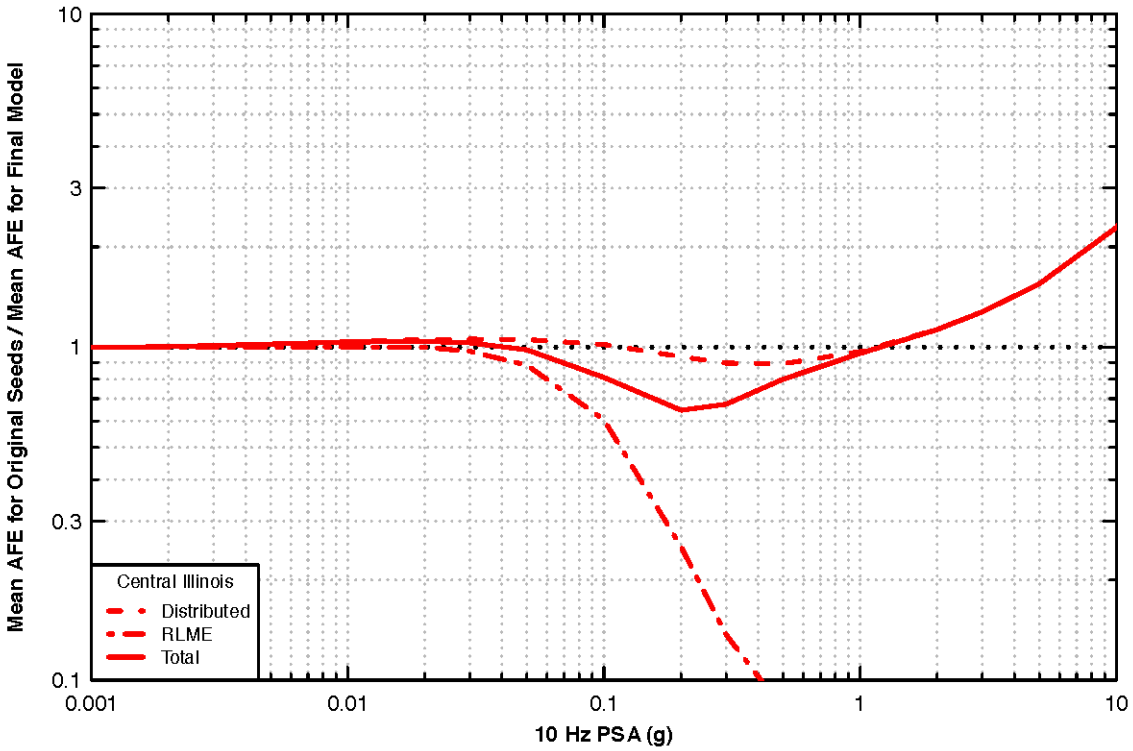


Figure 9–17(a) Hazard curves from original weighted seed GMMs using as-is distance measures (R_{JB} or R_{RUP}) and from the final NGA-East GMMs at Savannah, 1 Hz.

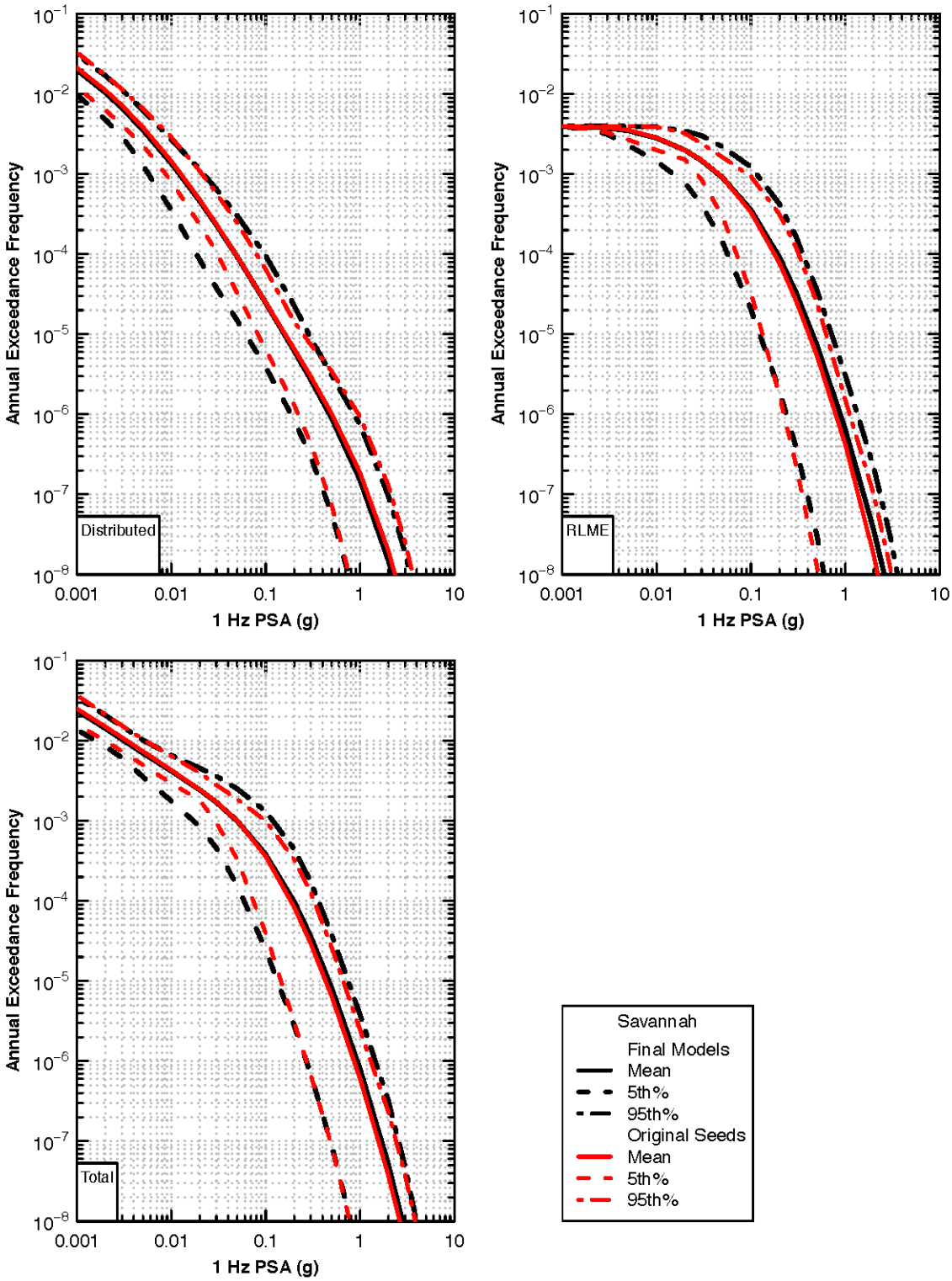


Figure 9–17(b) Hazard curves from original weighted seed GMMs using as-is distance measures (R_{JB} or R_{RUP}) and from the final NGA-East GMMs at Savannah, 10 Hz.

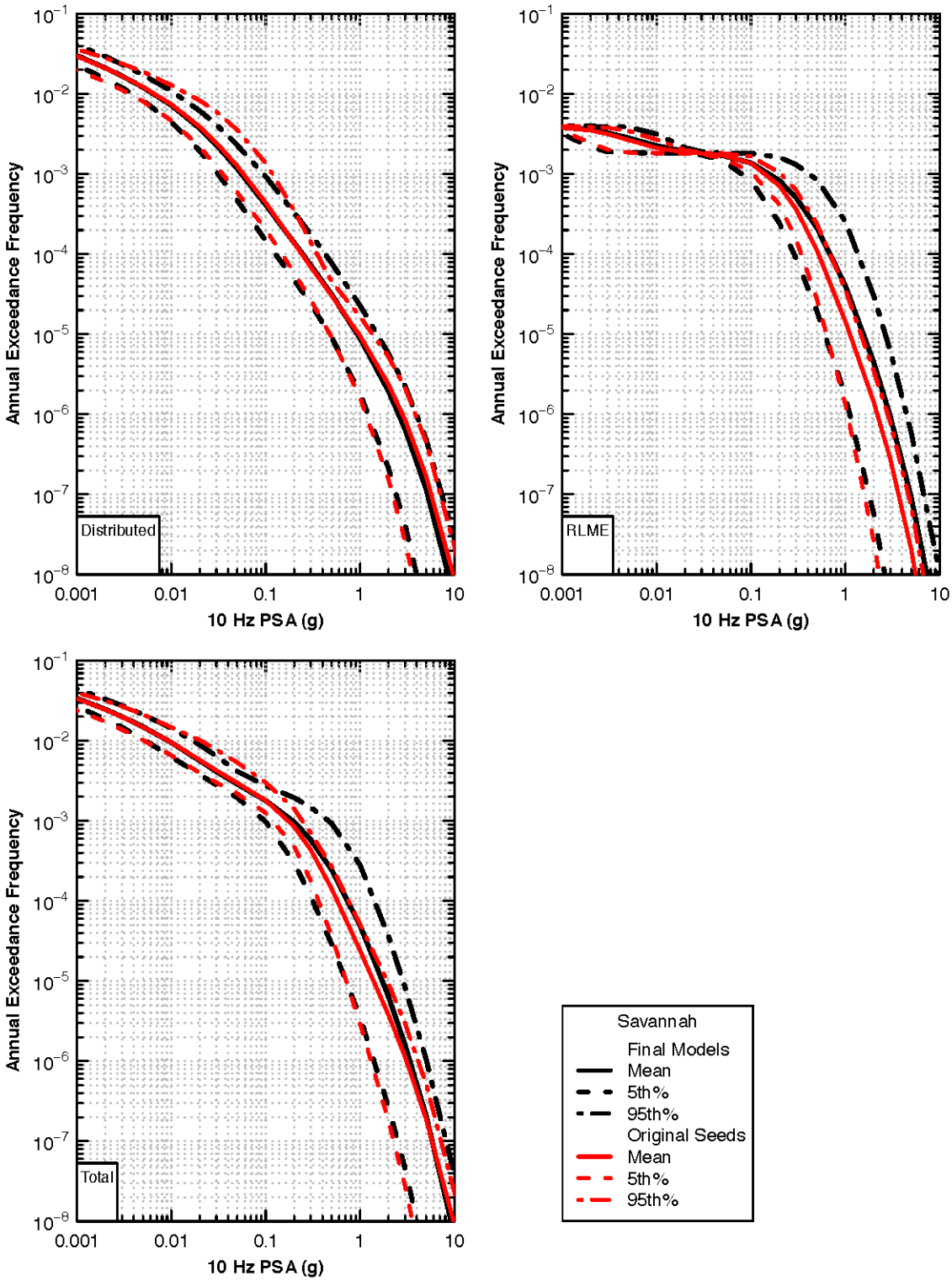
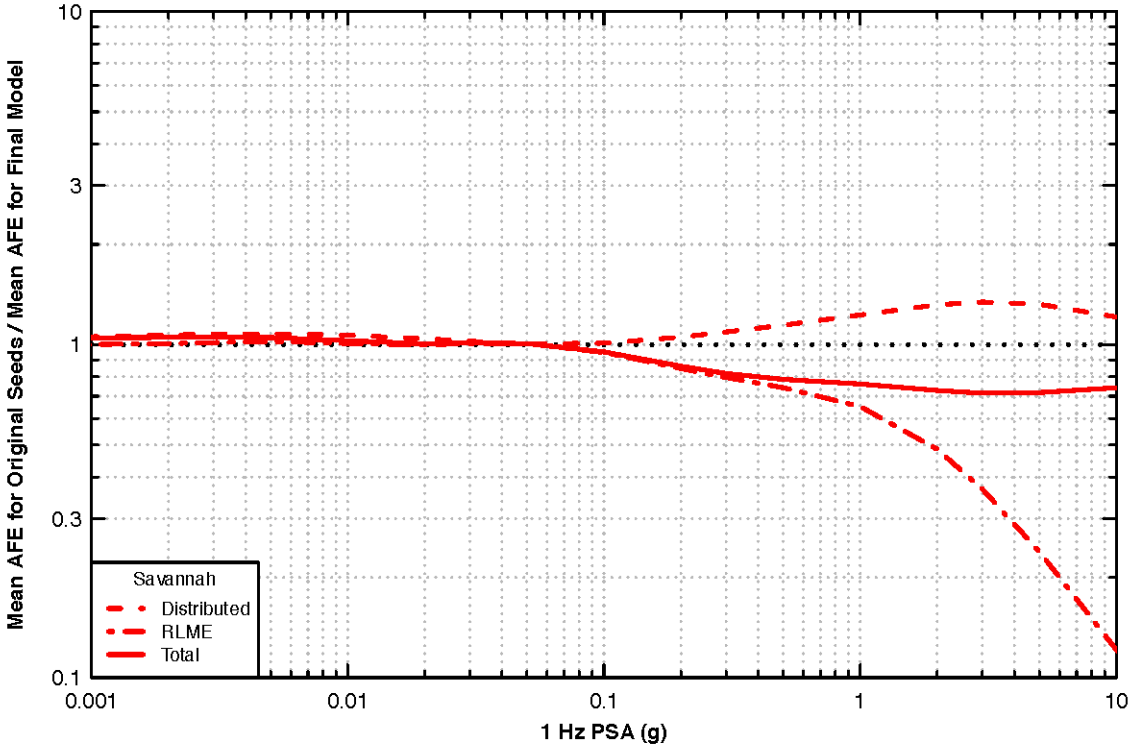


Figure 9–17(c) Hazard ratio from original weighted seed GMMs seed relative to that from the final NGA-East GMMs at Savannah (1 Hz and 10 Hz).



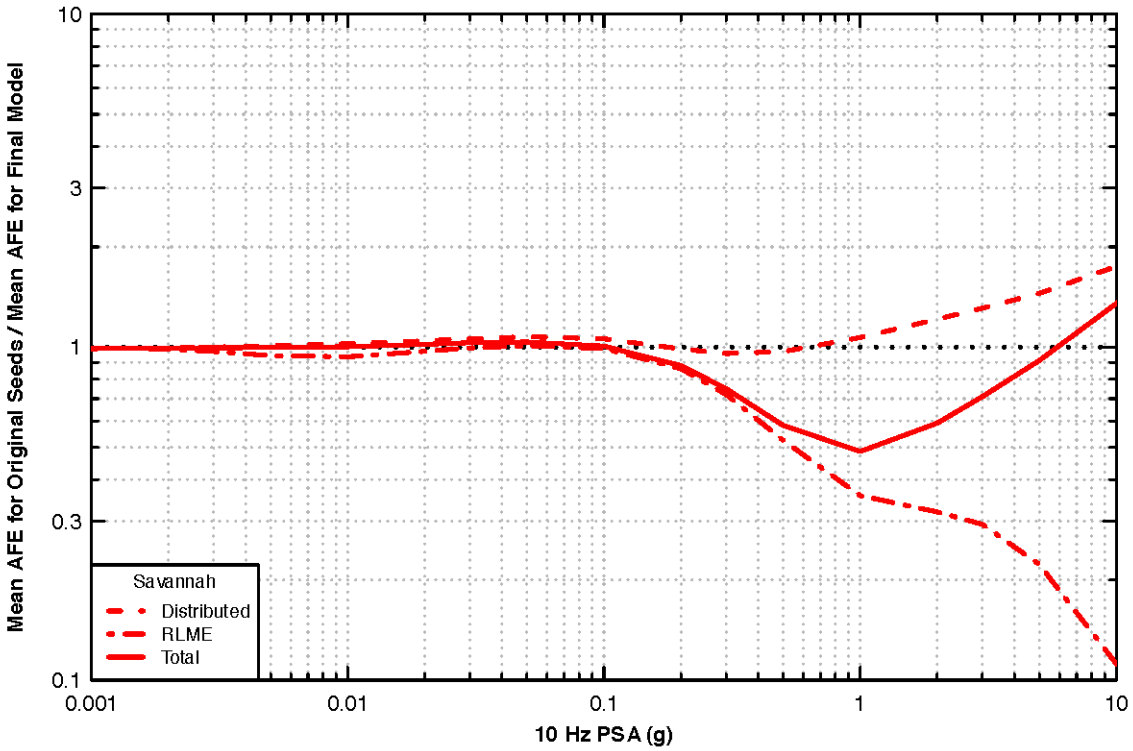


Figure 9–18 Weighted mean spectra from the 17 final NGA-East GMMs for two scenarios: M = 7.0, R_{RUP} = 50 km (top) and M = 7.0, R_{RUP} = 1000 km (bottom).

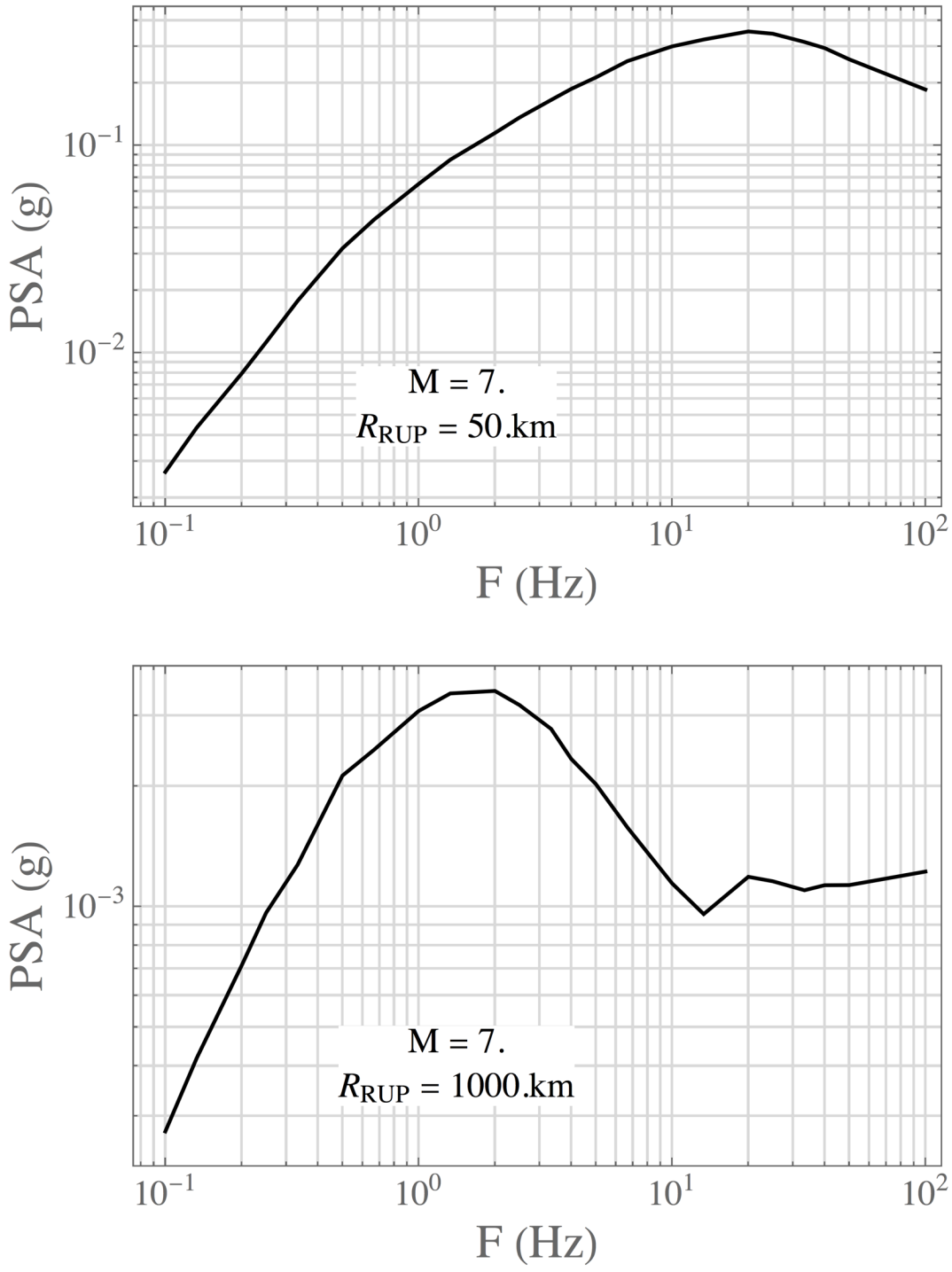


Figure 9–19(a) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.1 and 0.1333 Hz. Linear distance scale.

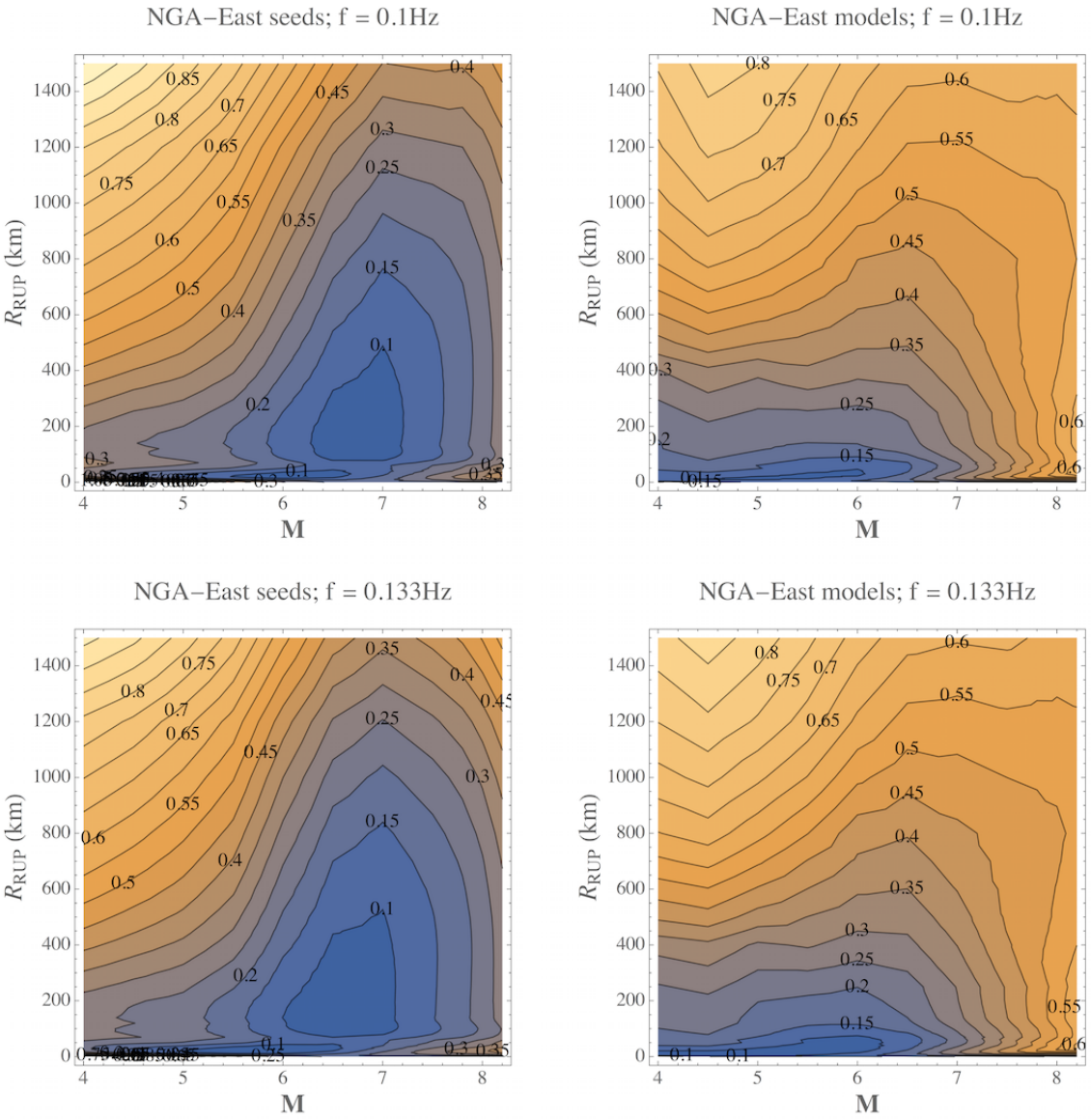


Figure 9–19(b) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.2 and 0.25 Hz. Linear distance scale.

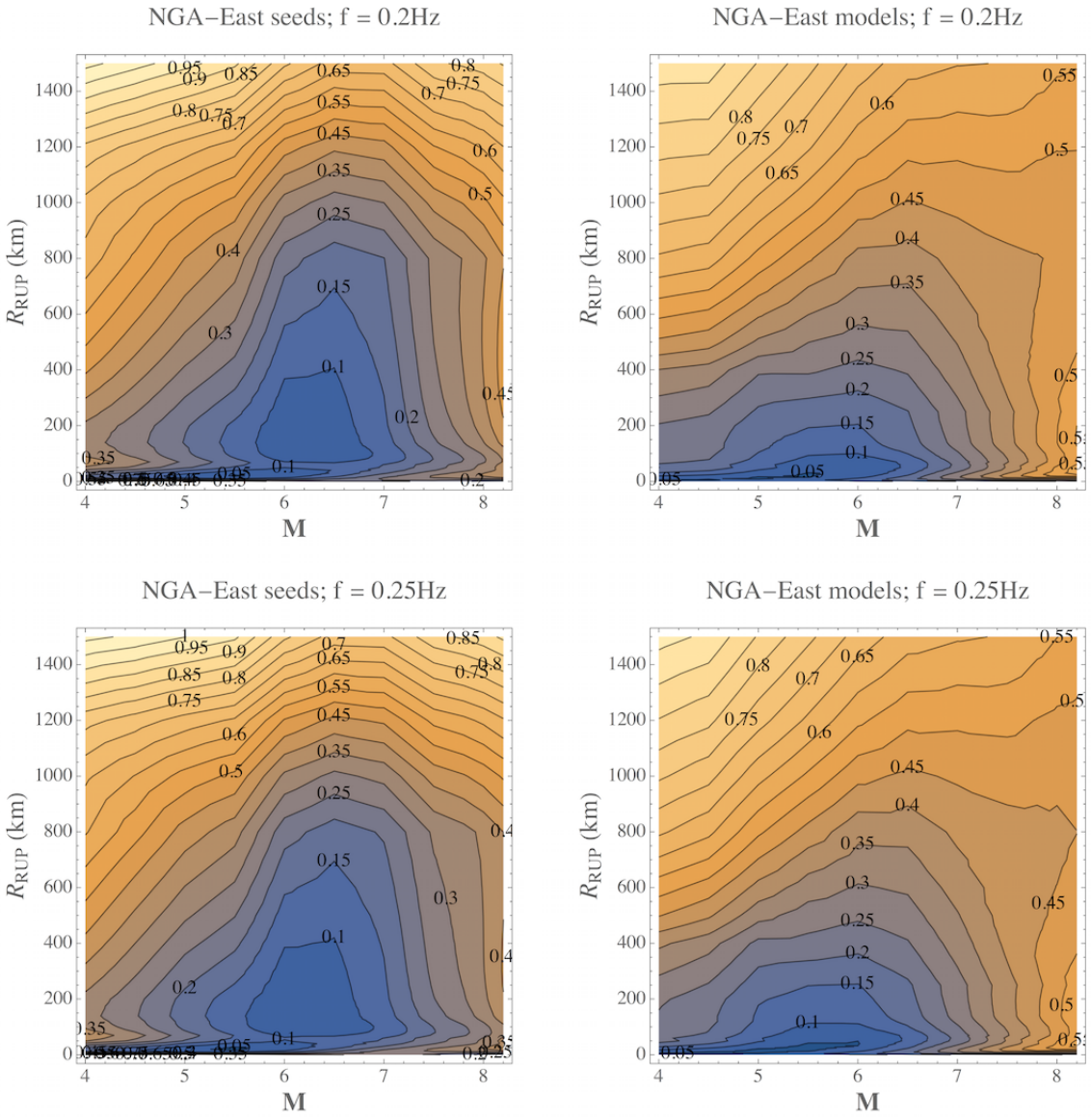


Figure 9–19(c) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.333 and 0.5 Hz. Linear distance scale.

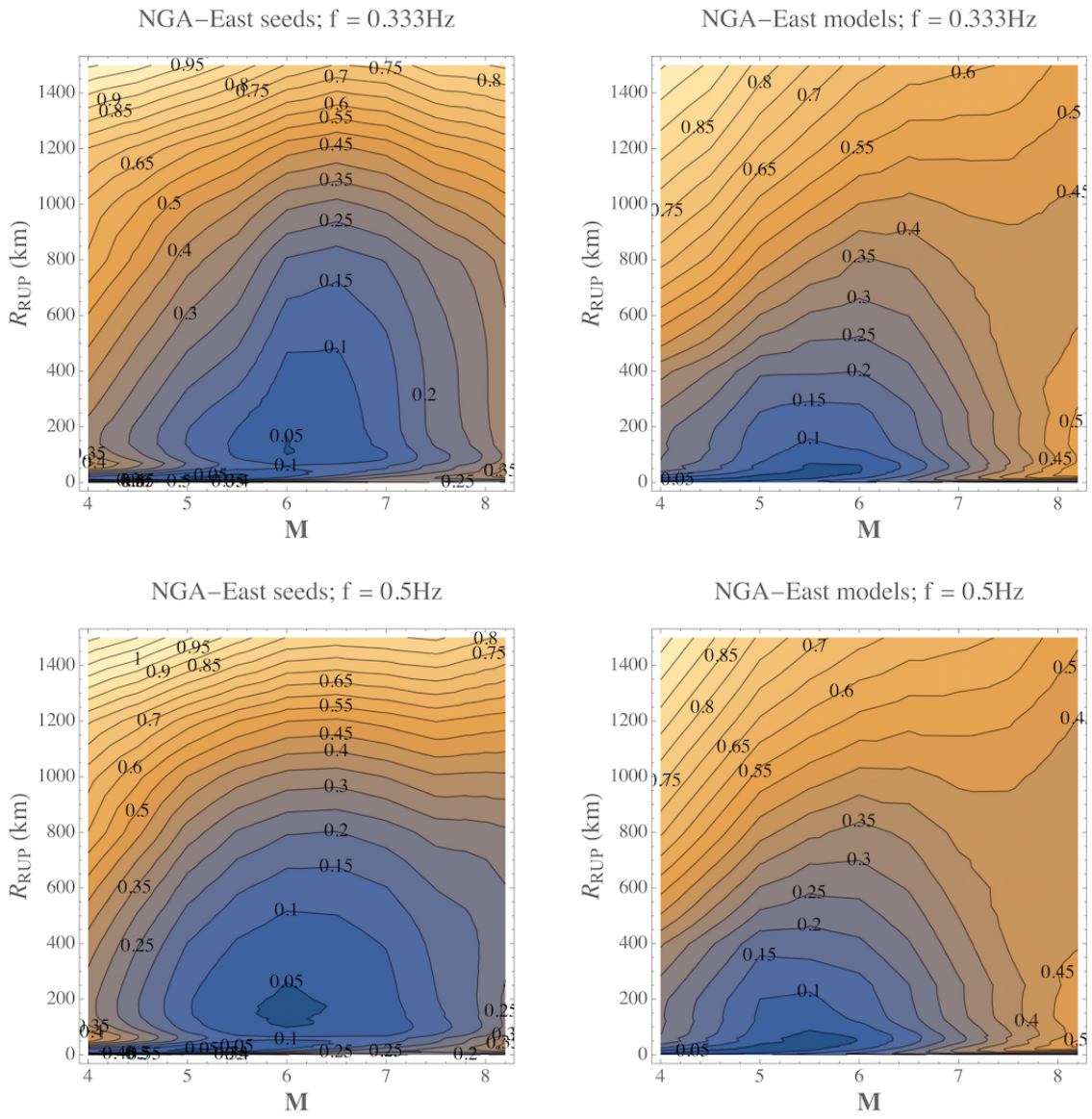


Figure 9–19(d) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.667 and 1 Hz. Linear distance scale.

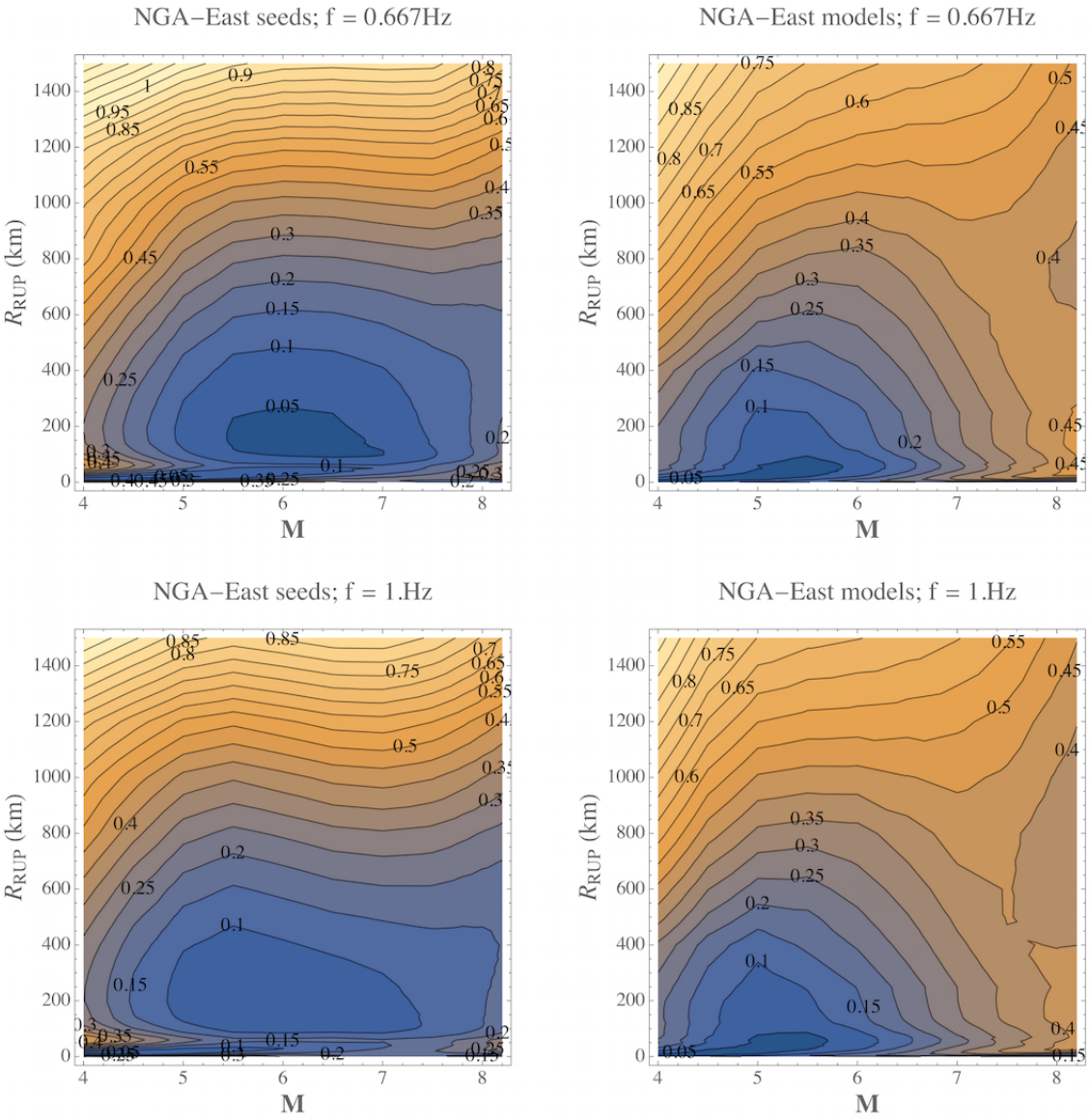


Figure 9–19(e) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 1.333 and 2 Hz. Linear distance scale.

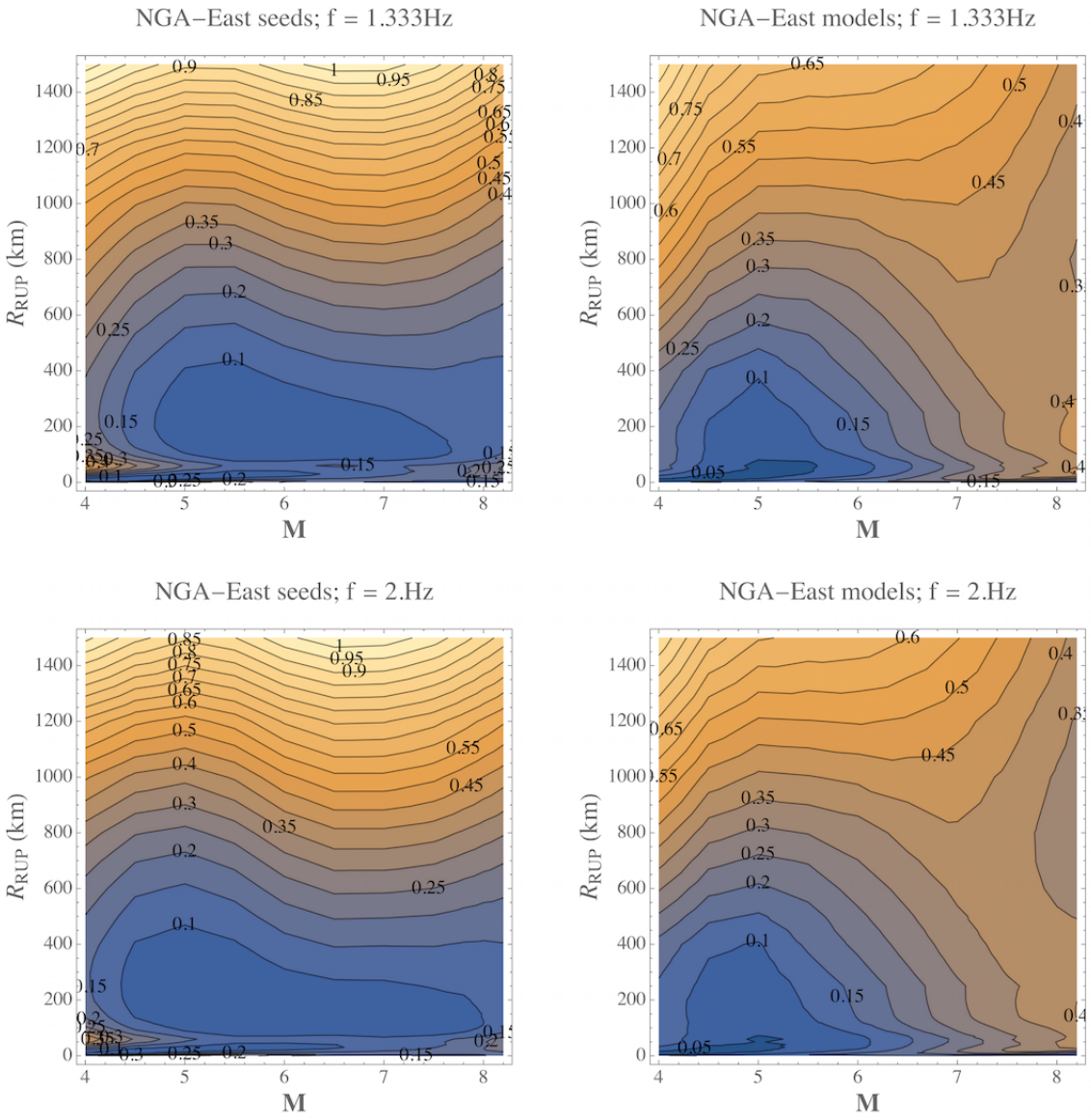


Figure 9–19(f) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 2.5 and 3.333 Hz. Linear distance scale.

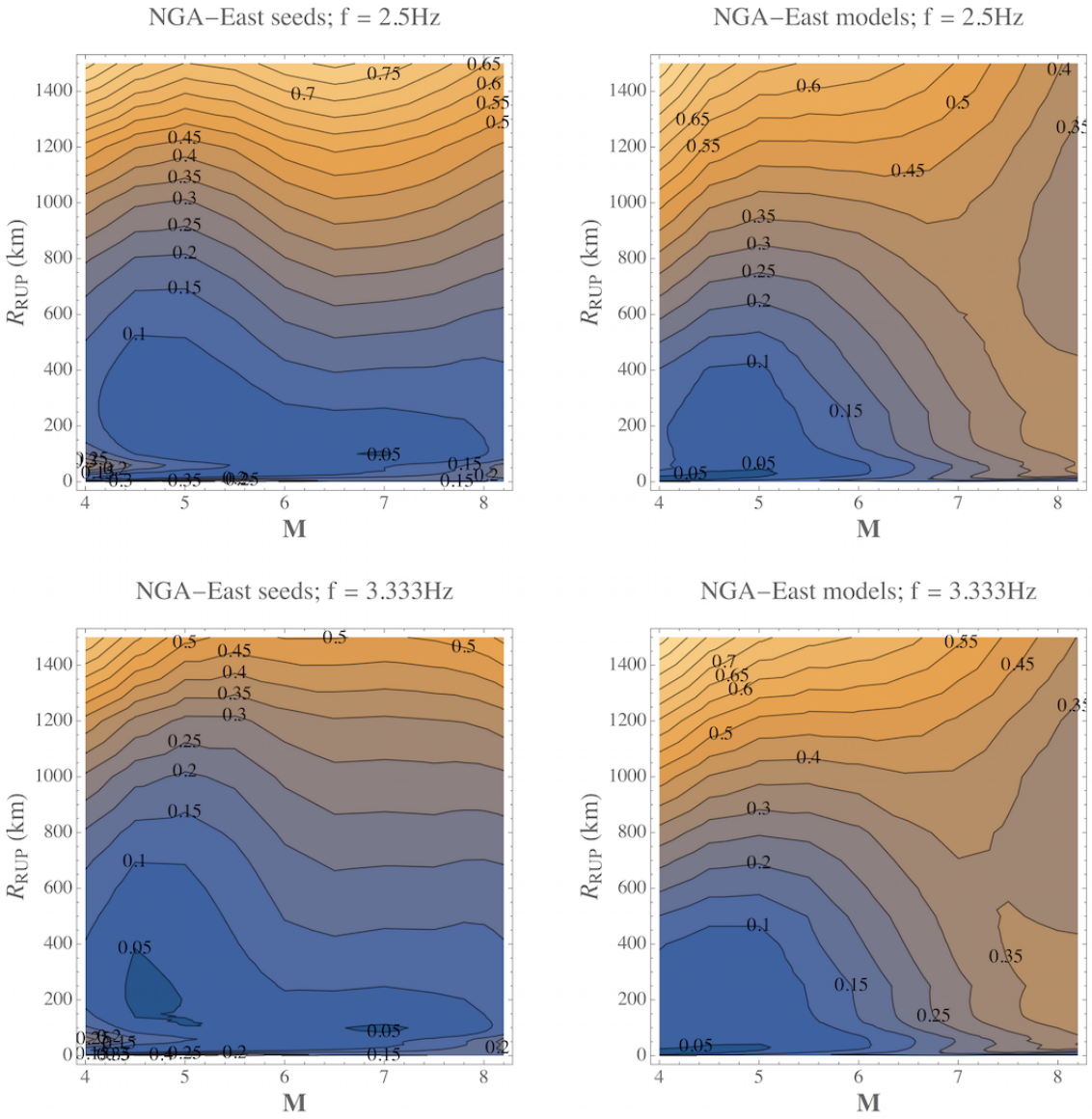


Figure 9–19(g) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 4 and 5 Hz. Linear distance scale.

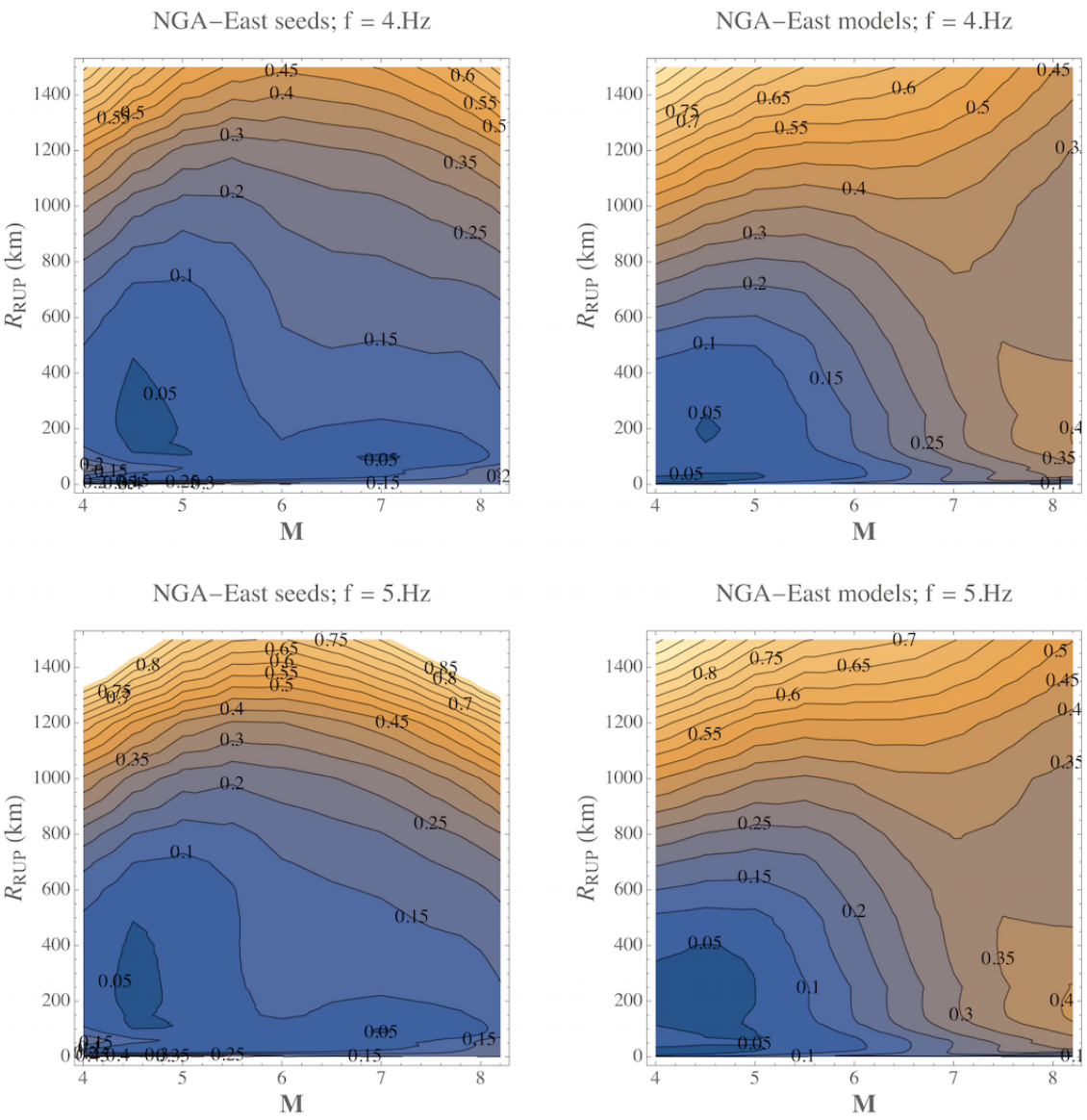


Figure 9–19(h) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 6.667 and 10 Hz. Linear distance scale.

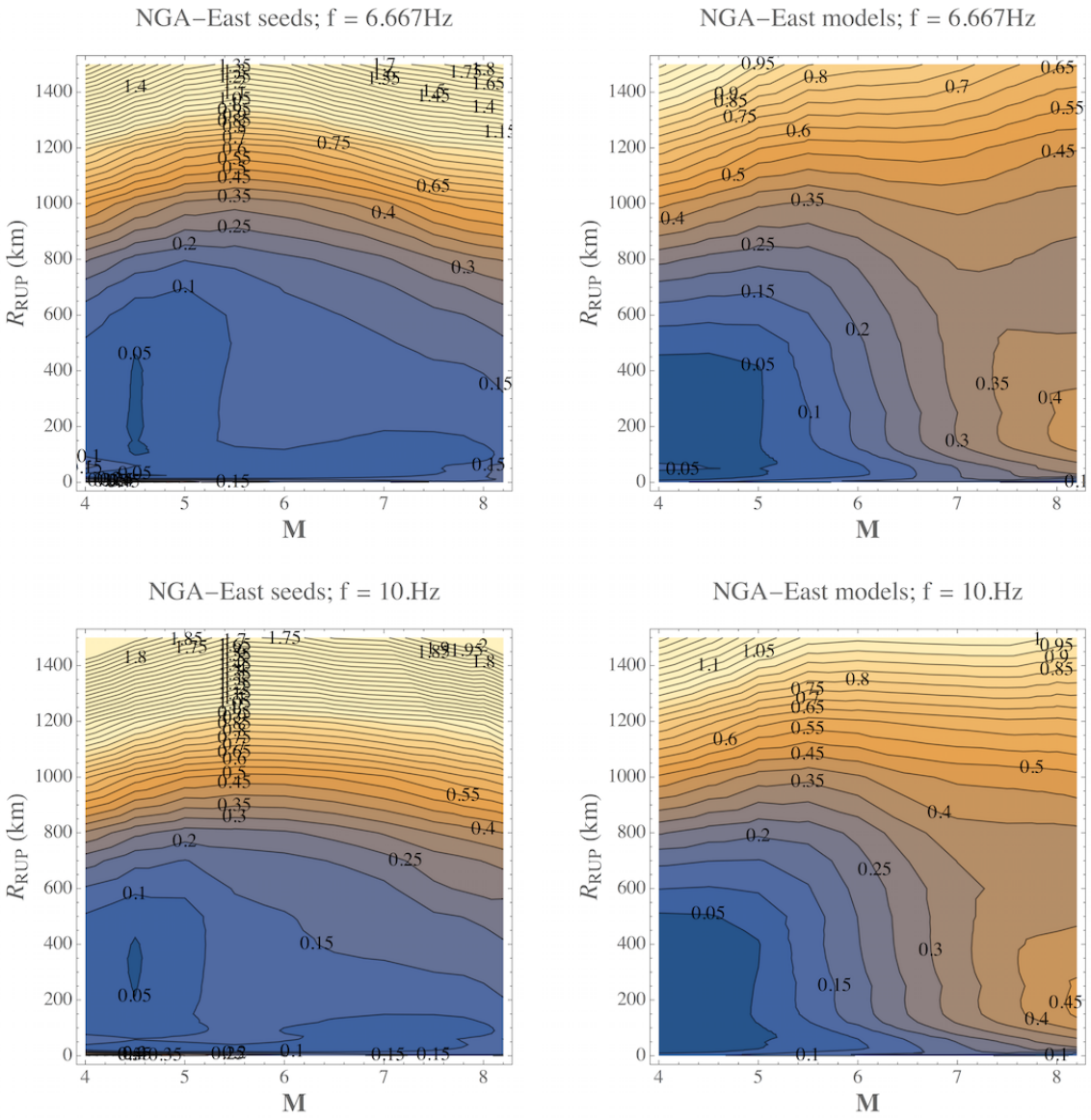


Figure 9–19(i) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 13.333 and 20 Hz. Linear distance scale.

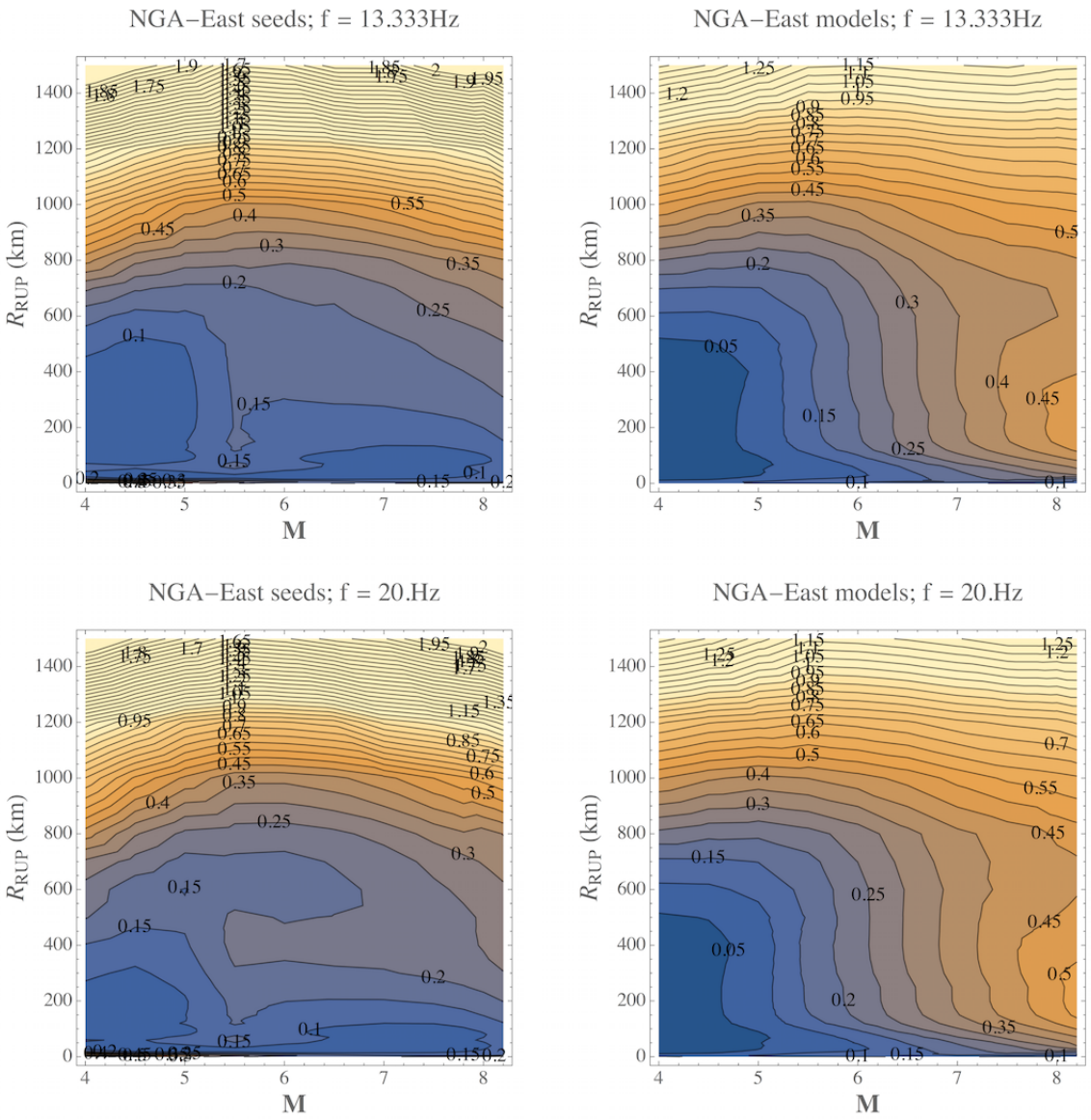


Figure 9–19(j) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 25 and 33.333 Hz. Linear distance scale.

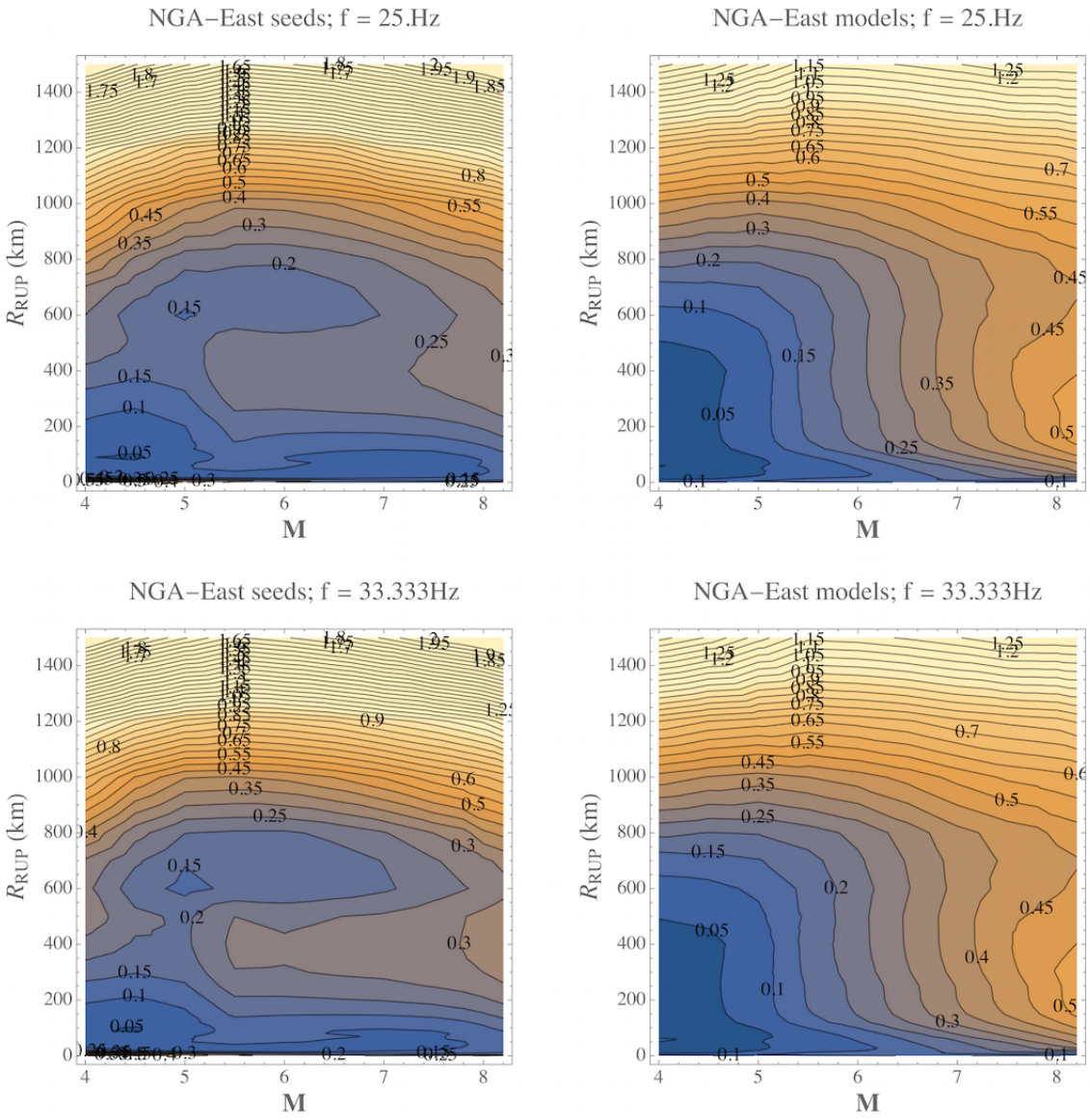


Figure 9–19(k) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 40 and 50 Hz. Linear distance scale.

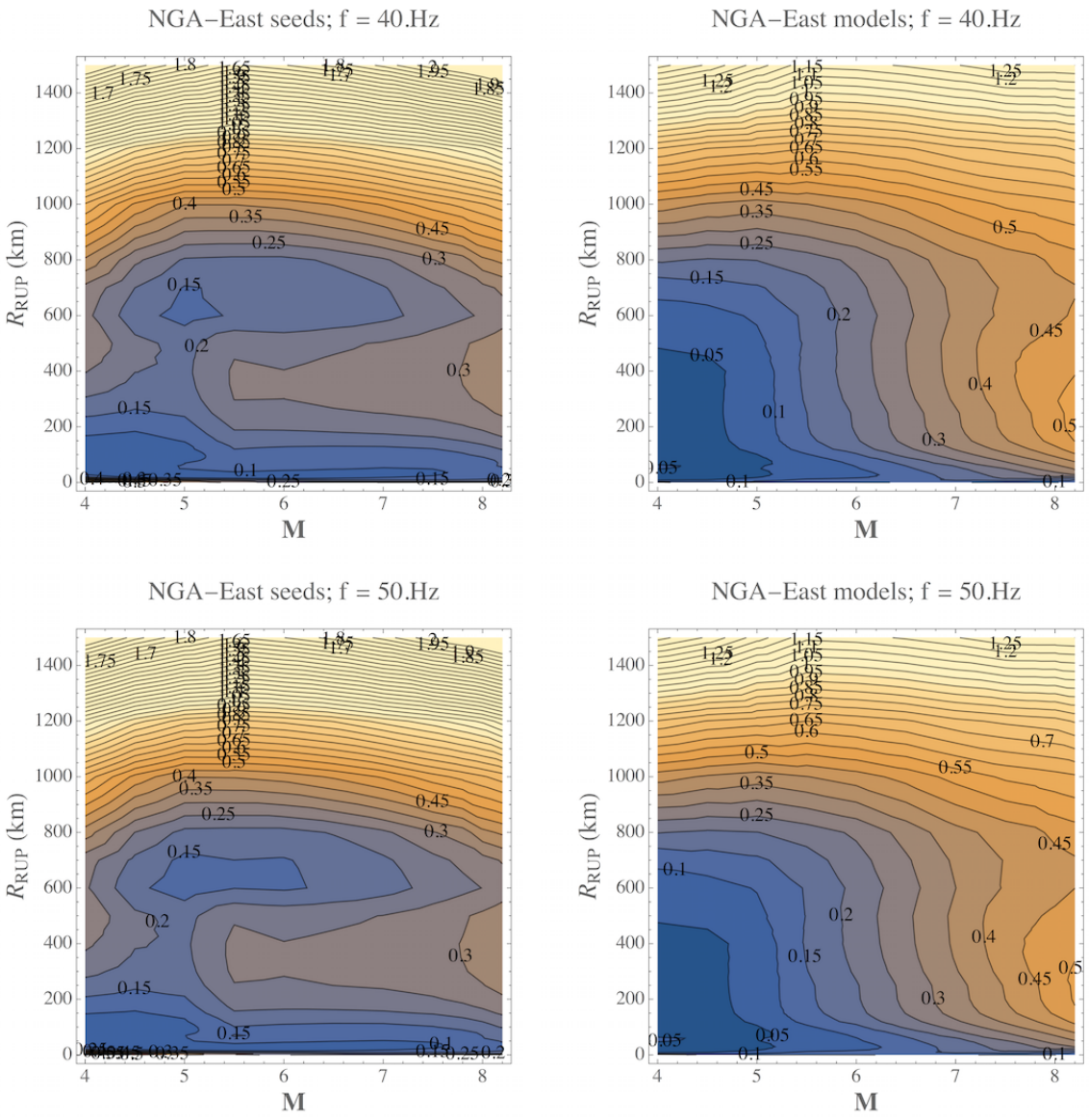


Figure 9–19(l) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 100 Hz and PGA. Linear distance scale.

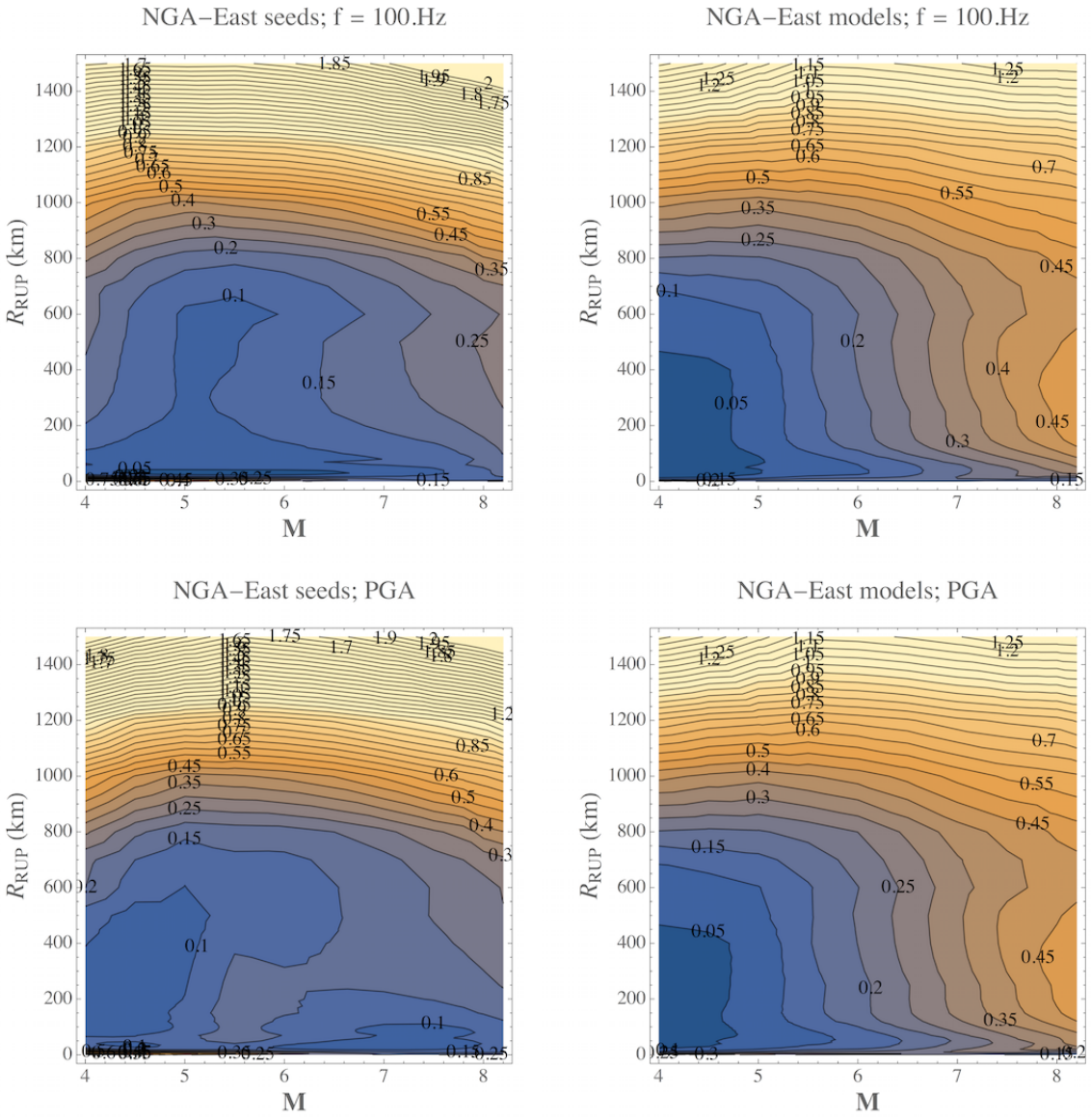


Figure 9–19(m) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for PGV. Linear distance scale.

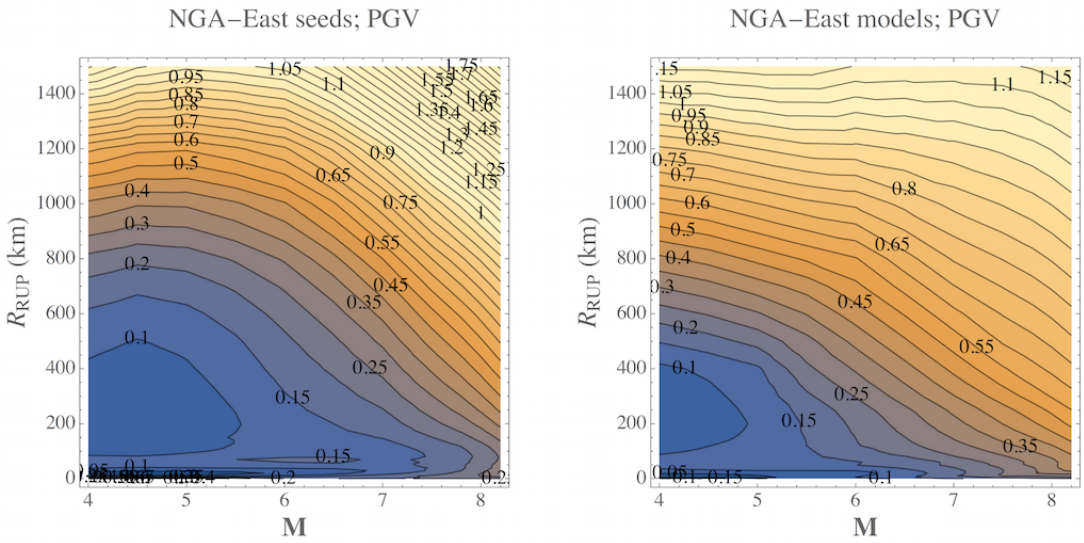


Figure 9–20(a) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.1 and 0.1333 Hz. Logarithmic distance scale.

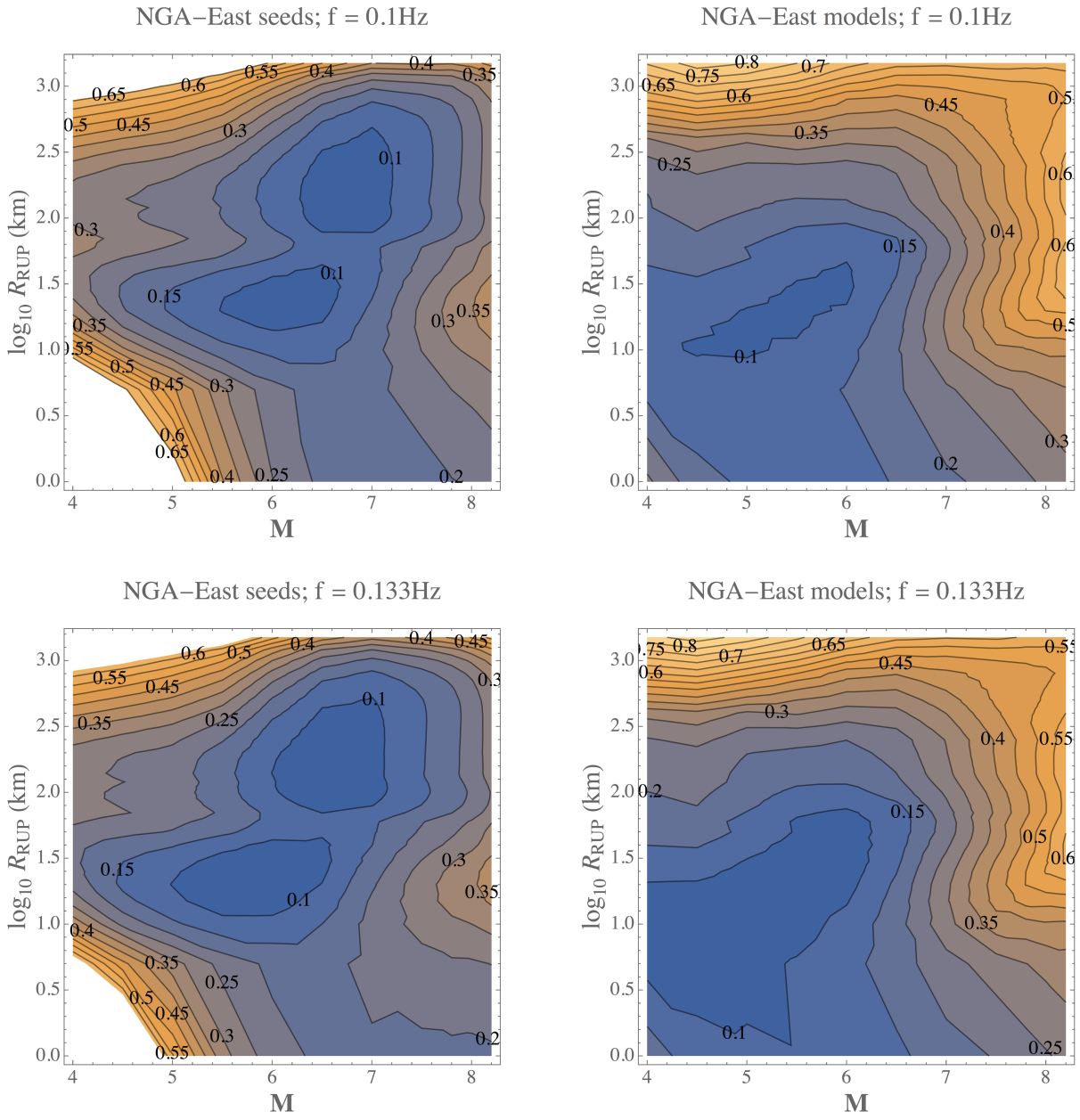


Figure 9–20(b) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.2 and 0.25 Hz. Logarithmic distance scale.

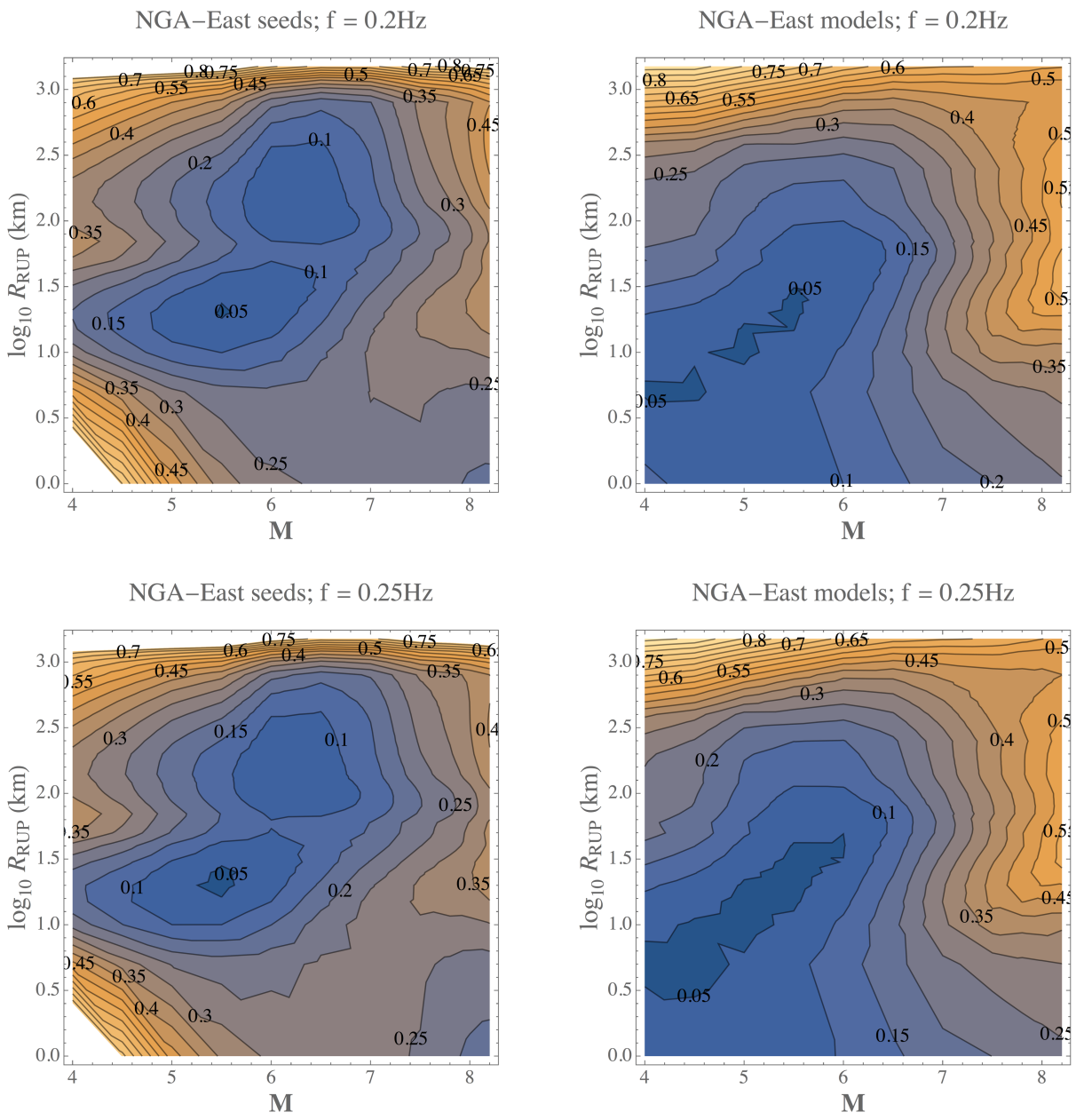


Figure 9–20(c) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.333 and 0.5 Hz. Logarithmic distance scale.

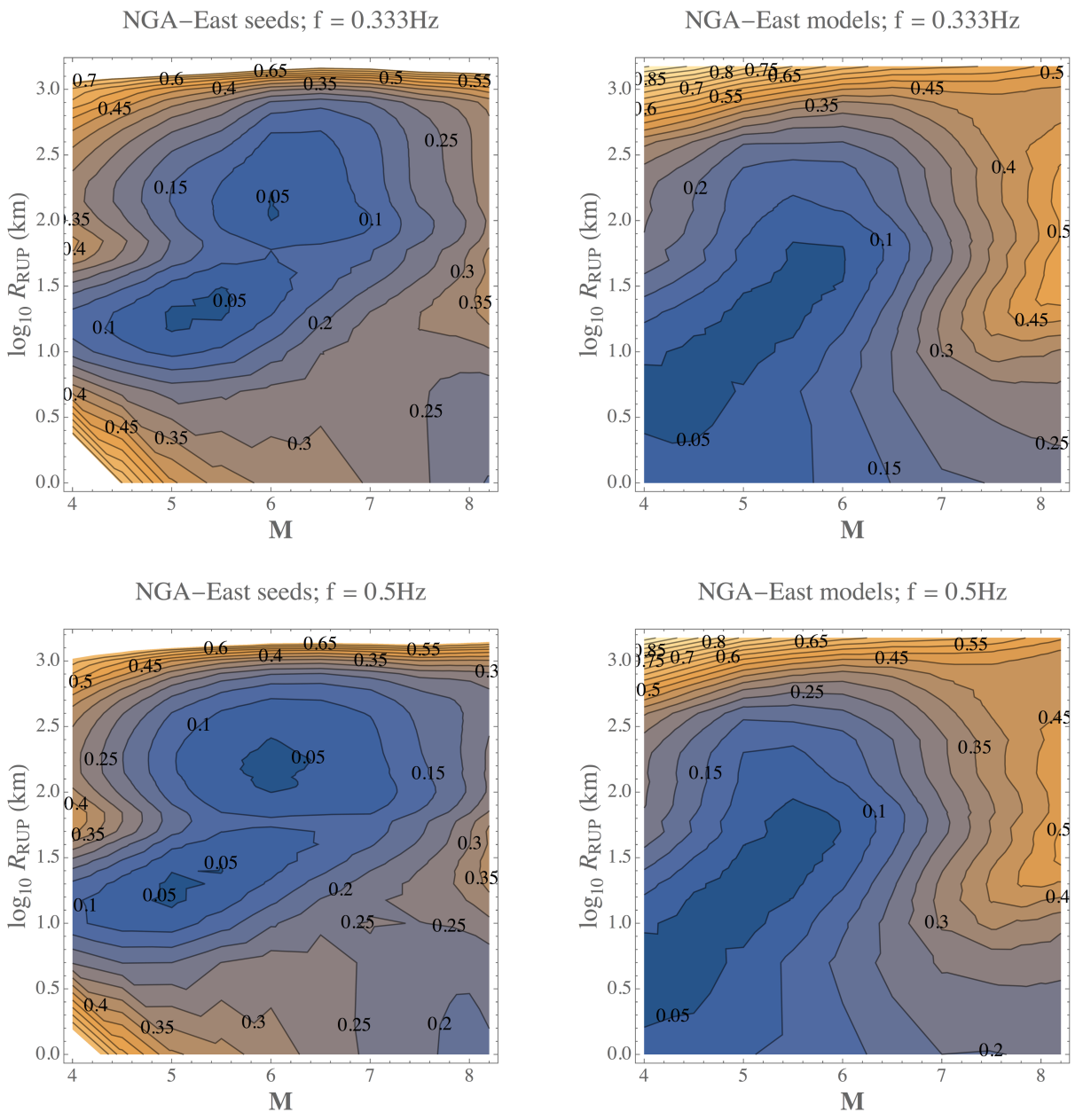


Figure 9–20(d) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 0.667 and 1 Hz. Logarithmic distance scale.

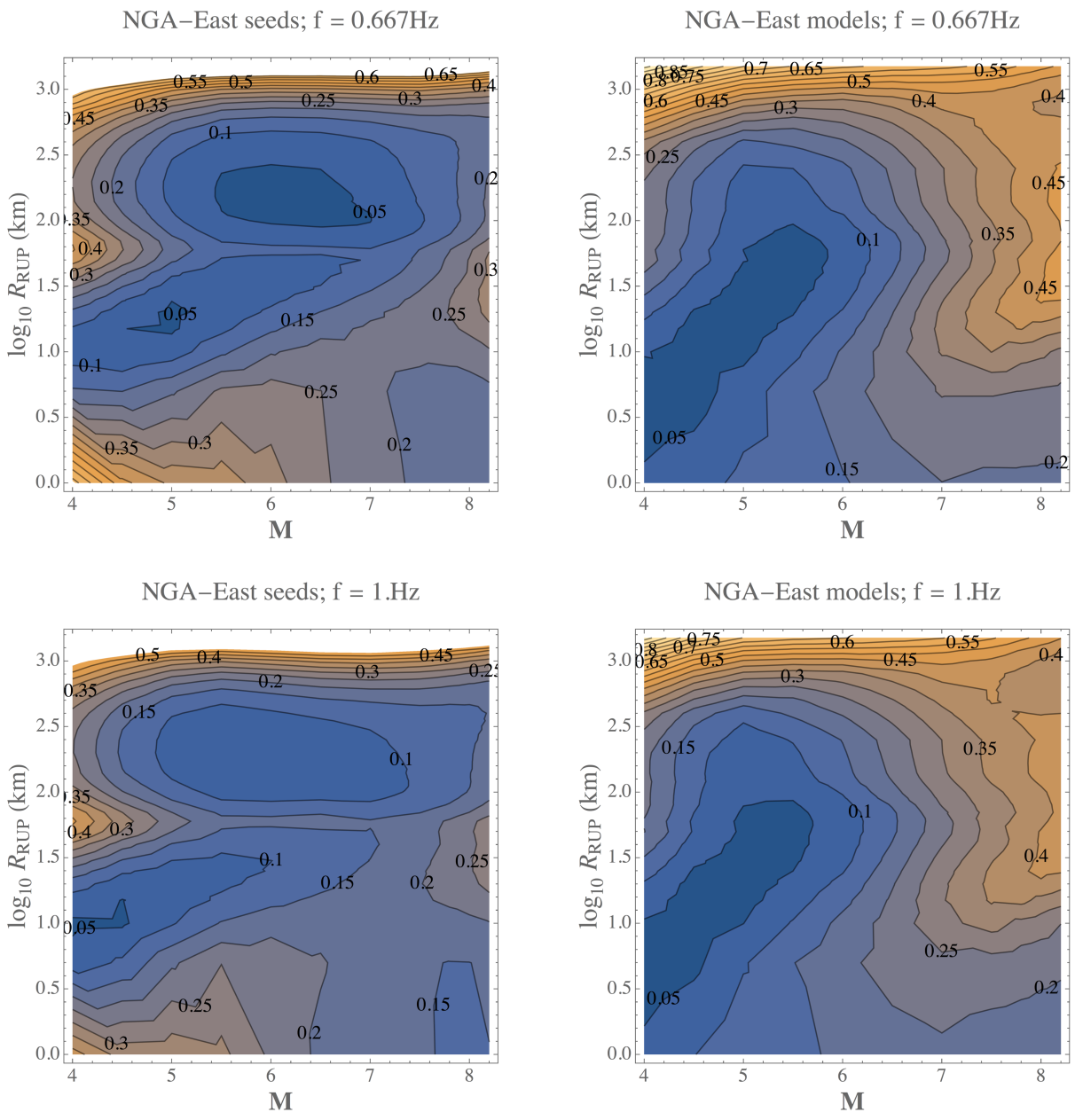


Figure 9–20(e) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 1.333 and 2 Hz. Logarithmic distance scale.

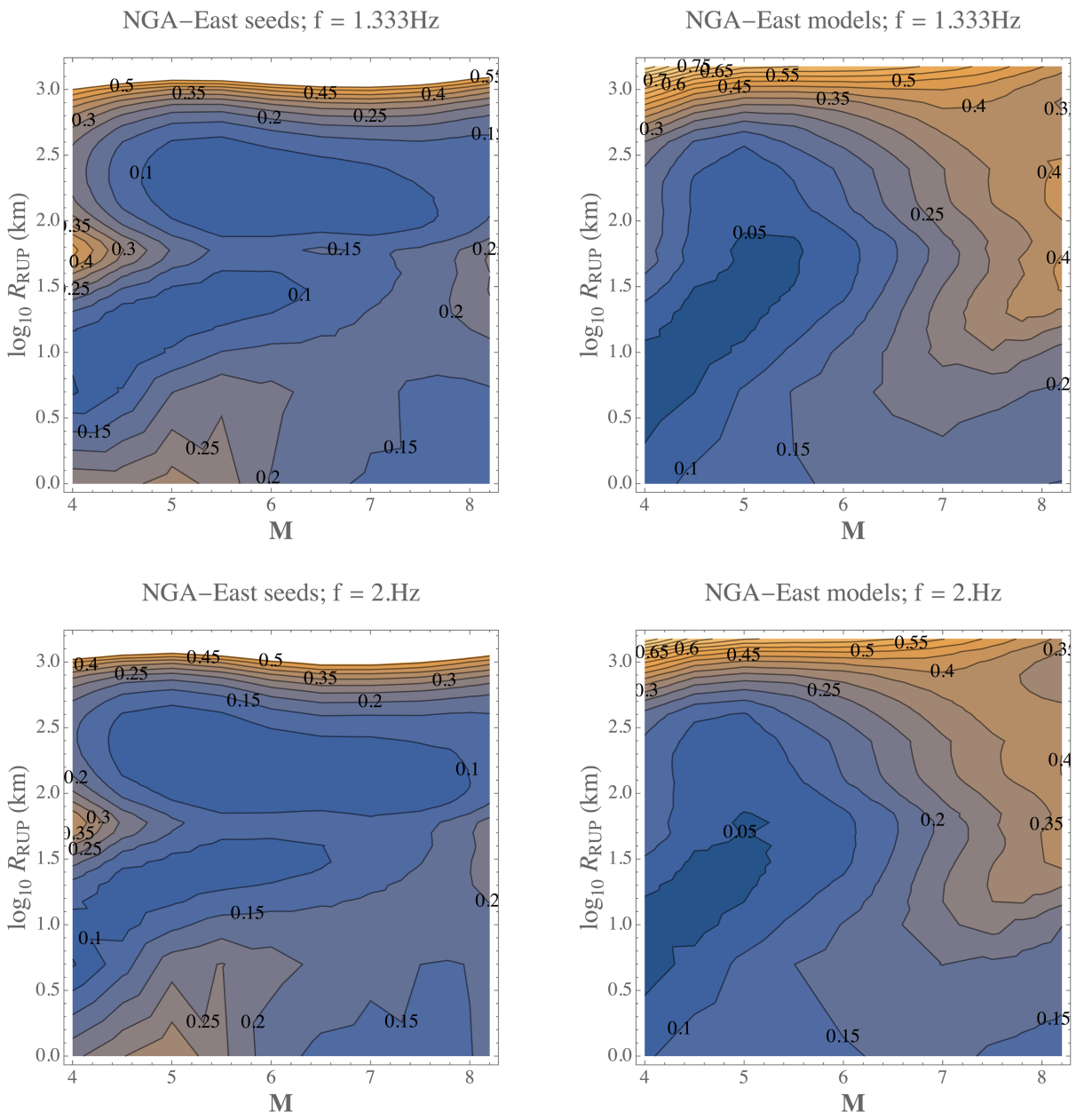


Figure 9–20(f) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 2.5 and 3.333 Hz. Logarithmic distance scale.

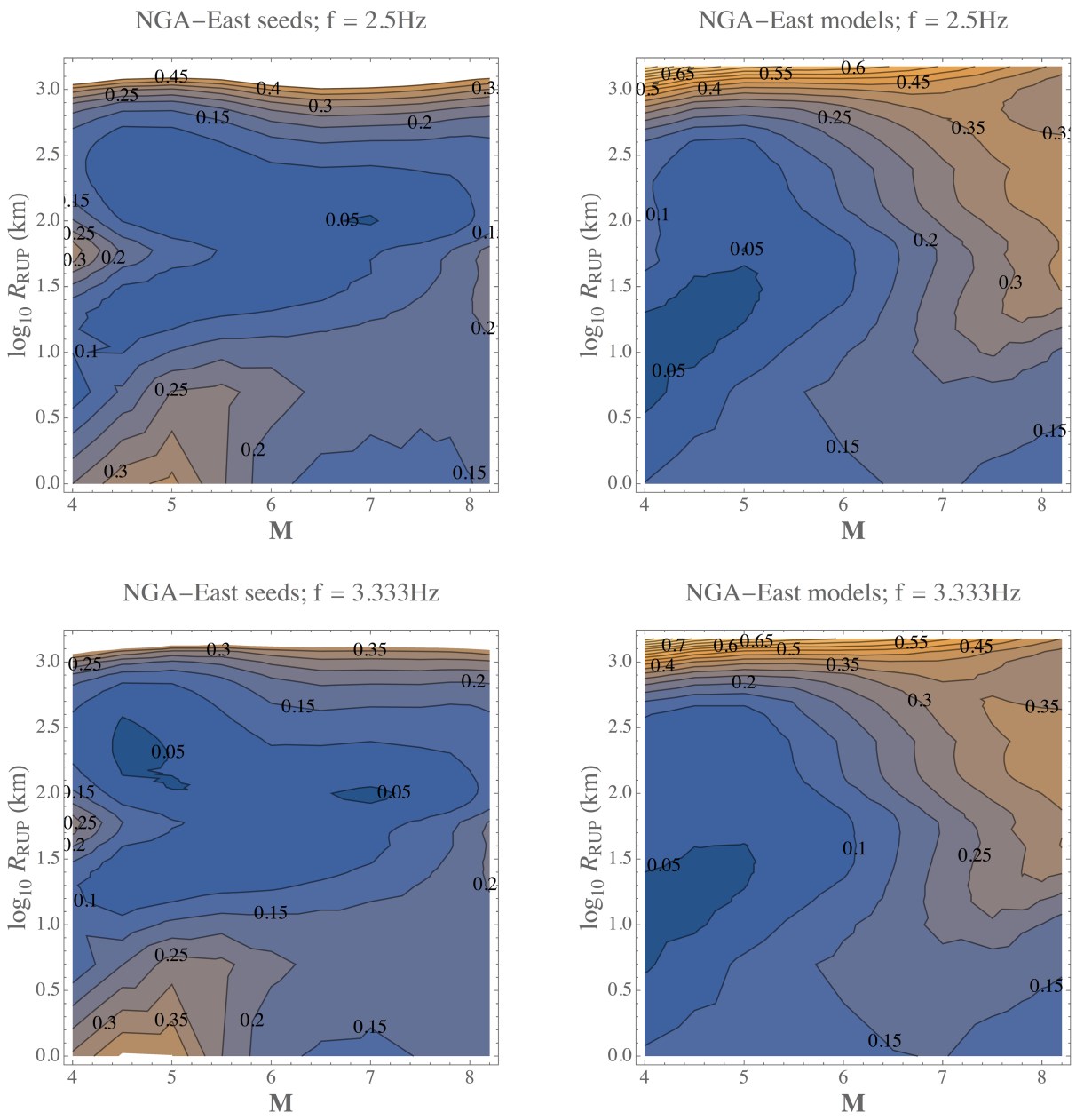


Figure 9–20(g) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 4 and 5 Hz. Logarithmic distance scale.

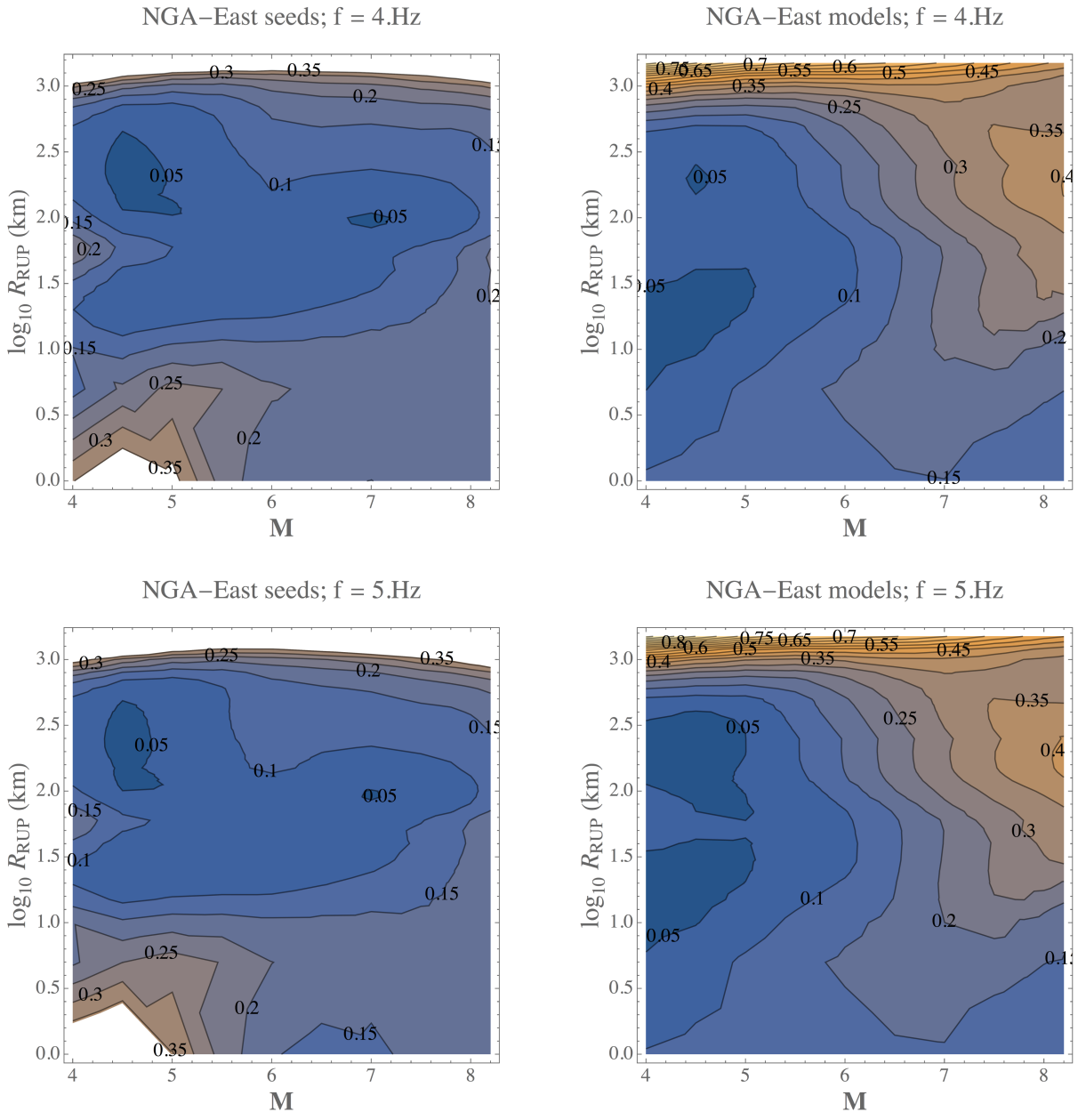


Figure 9–20(h) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 6.667 and 10 Hz. Logarithmic distance scale.

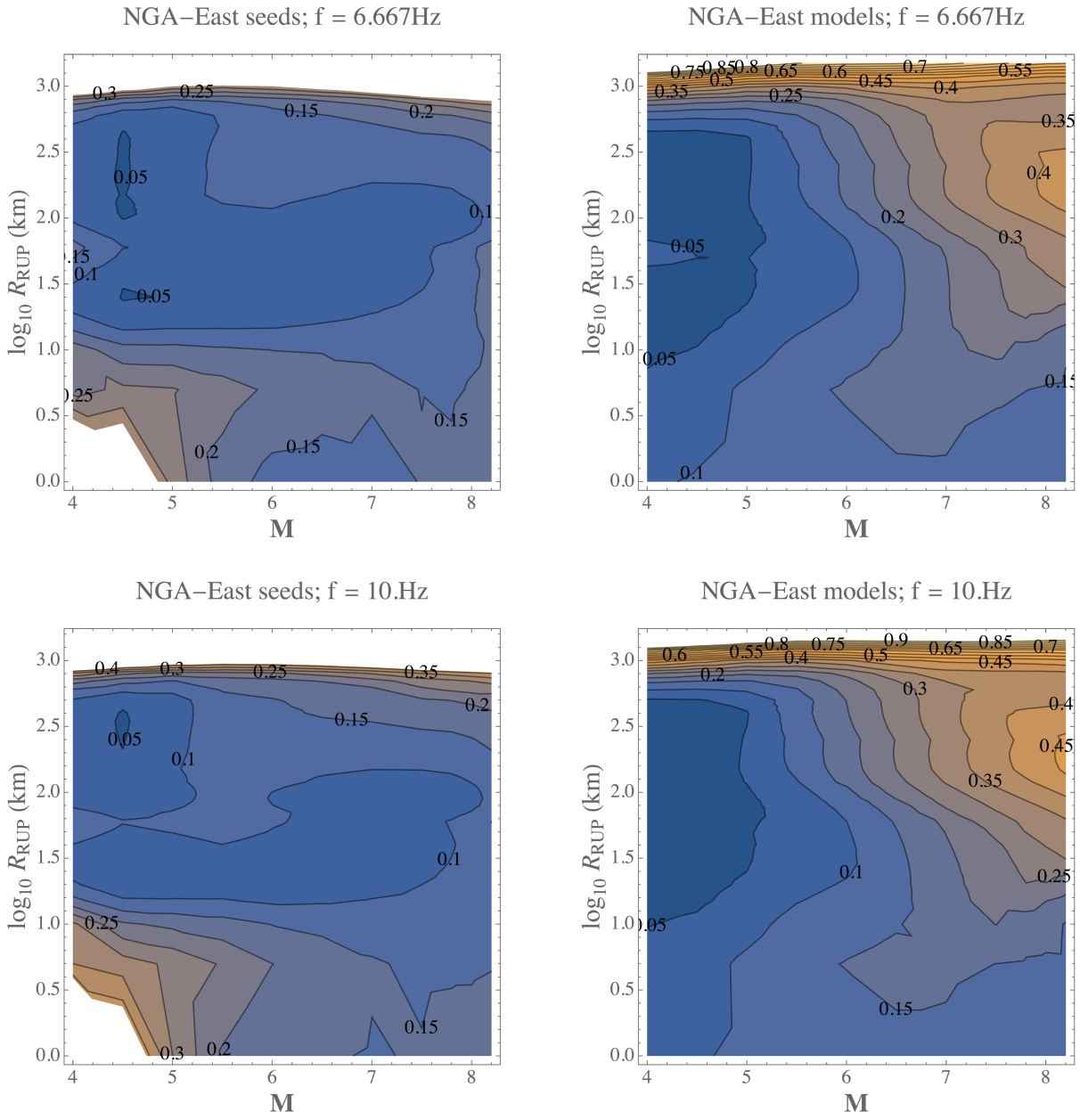


Figure 9–20(i) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 13.333 and 20 Hz. Logarithmic distance scale.

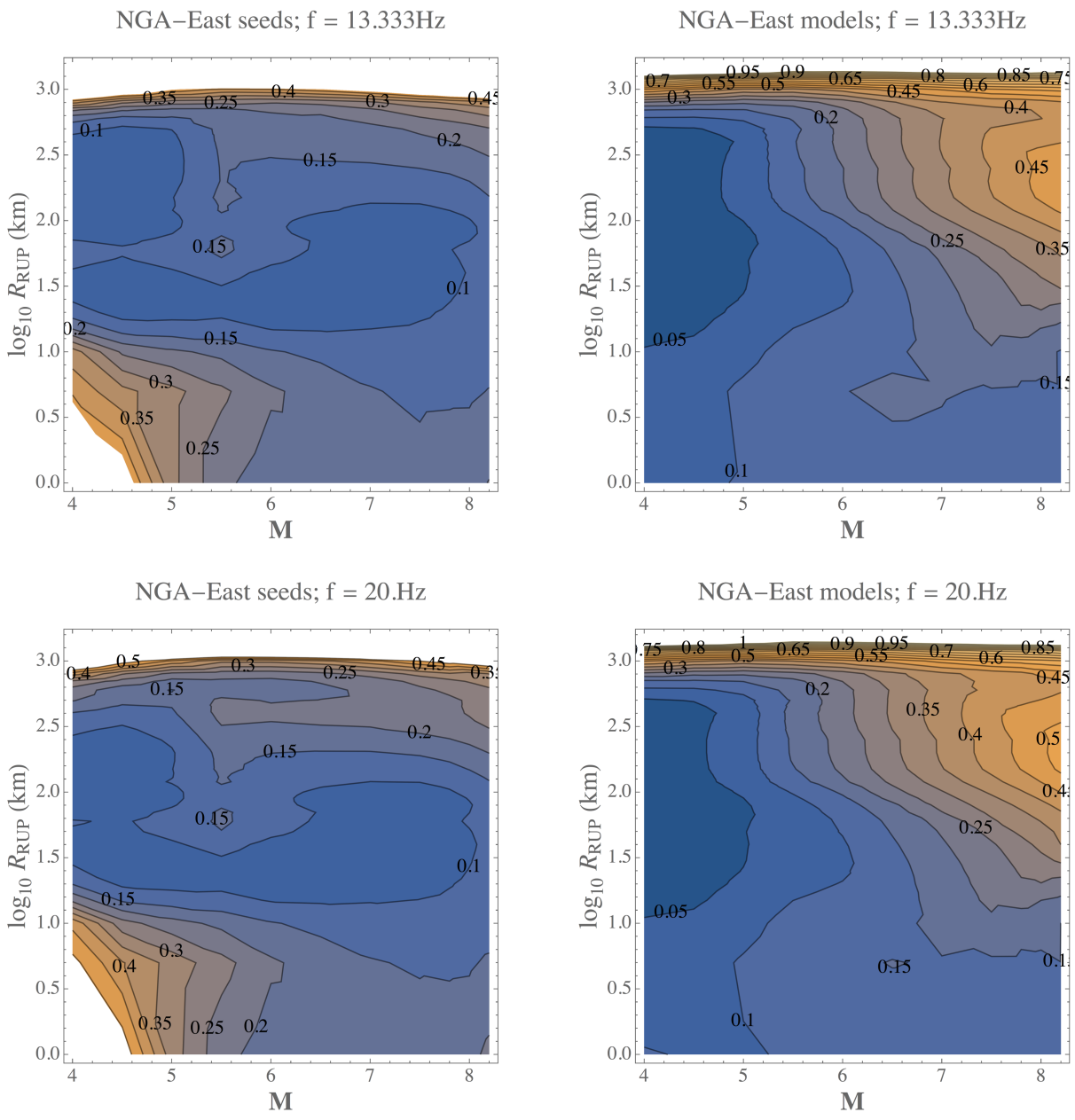


Figure 9–20(j) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 25 and 33.333 Hz. Logarithmic distance scale.

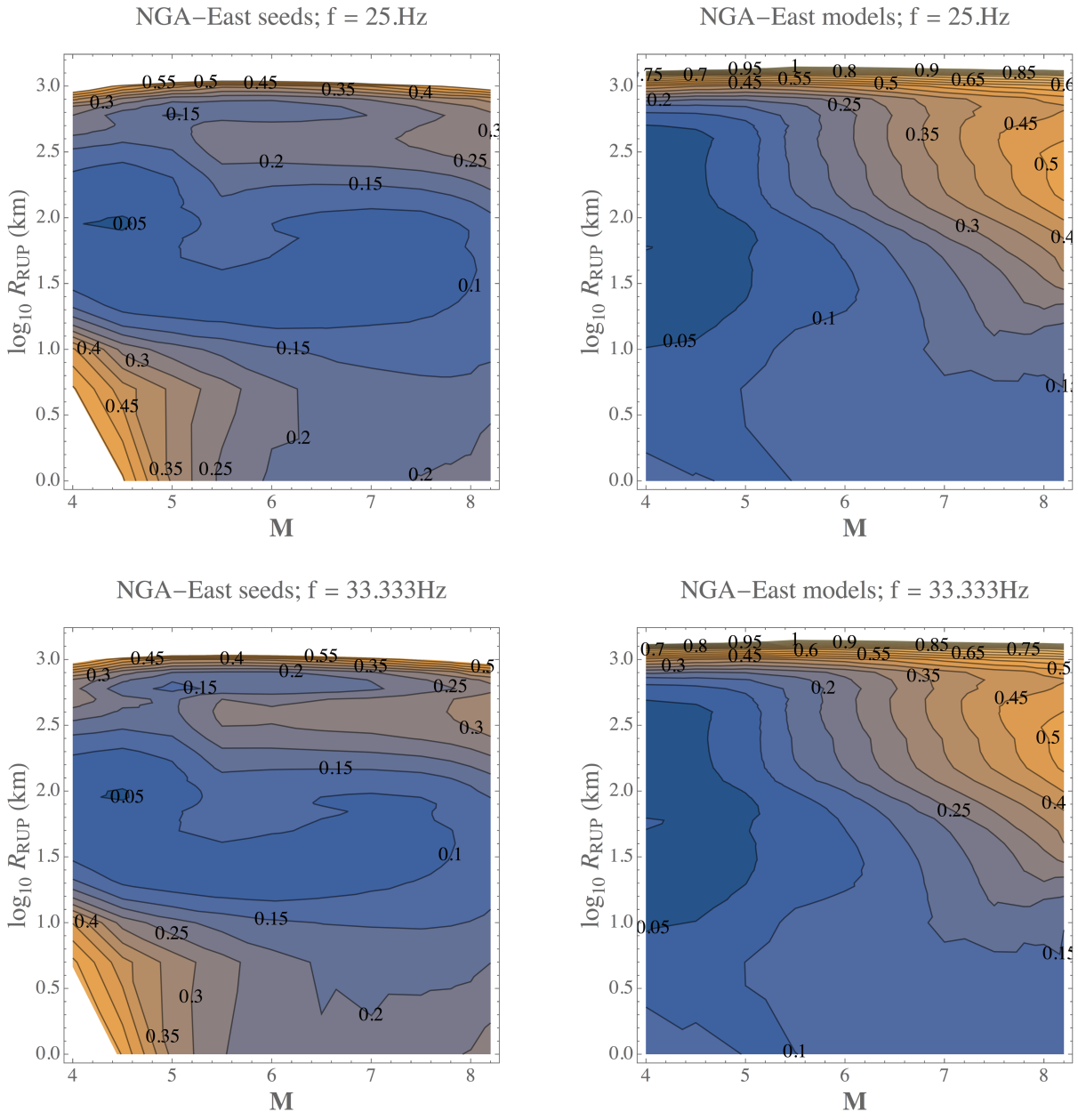


Figure 9–20(k) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 40 and 50 Hz. Logarithmic distance scale.

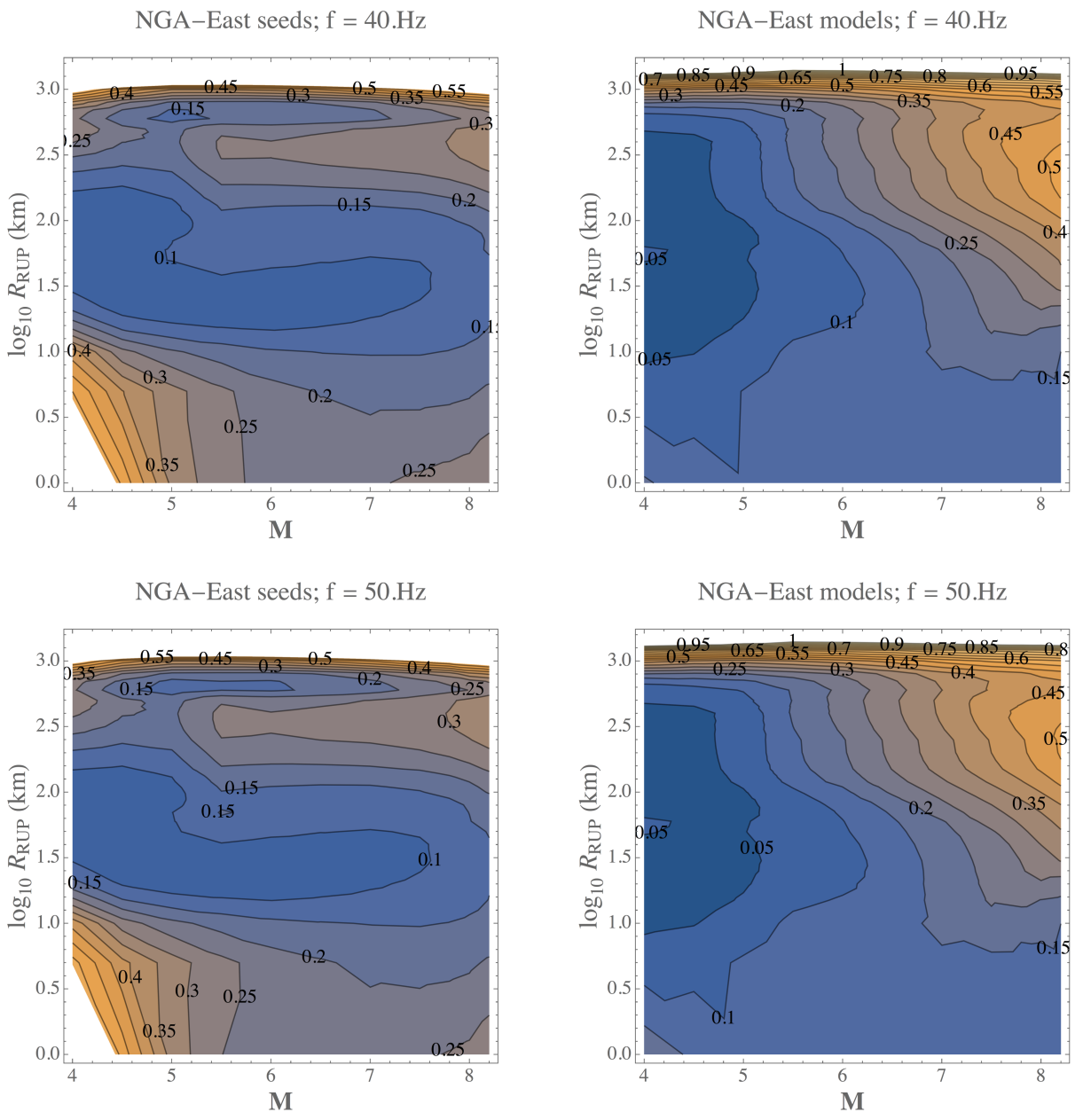


Figure 9–20(l) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for 100 Hz and PGA. Logarithmic distance scale.

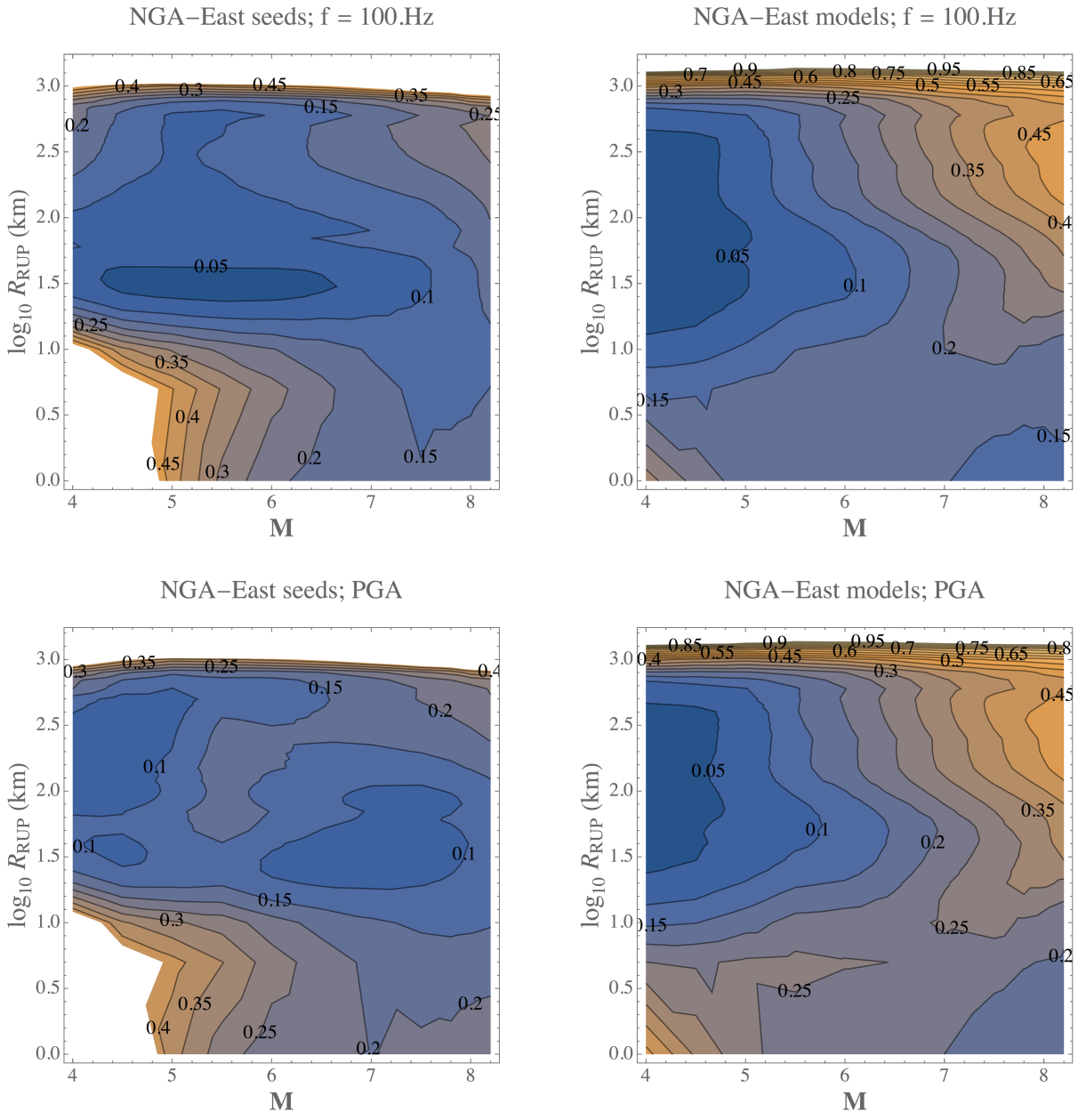
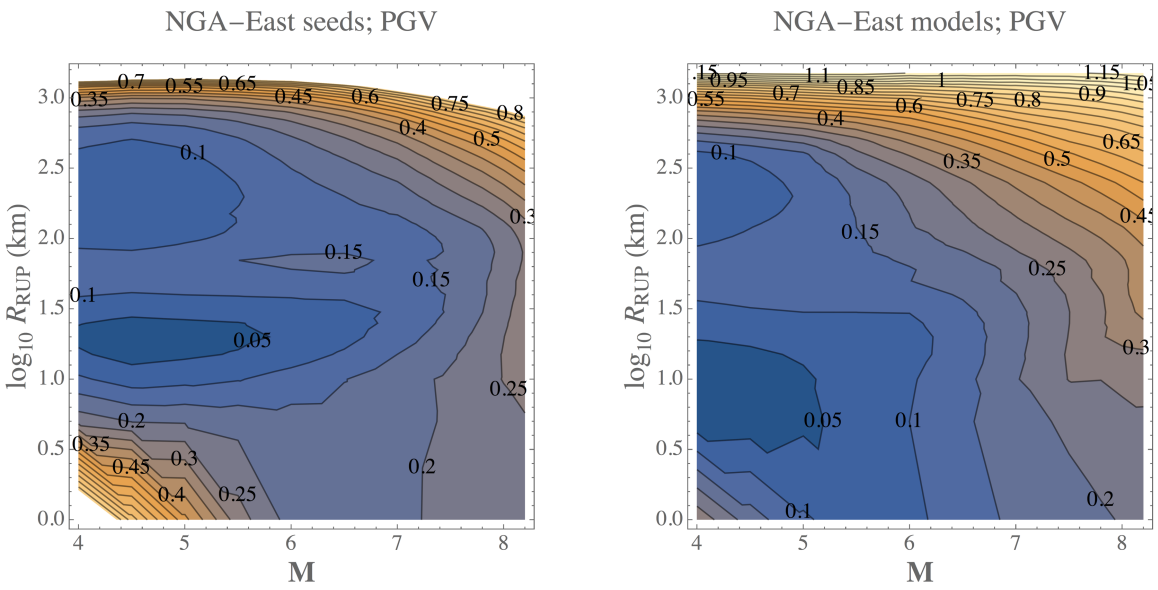


Figure 9–20(m) Original weighted seed GMMs variance (left) and achieved weighted variance of the NGA-East final GMMs (right) for PGV. Logarithmic distance scale.



10. Candidate Proponent Standard Deviation Models

10.1 Characterization of Aleatory Variability

In probabilistic seismic hazard analysis (PSHA), it is necessary to capture the full distribution (median and aleatory variability) of ground motions that may be generated by any given earthquake scenario. Ground-motion models (GMMs) discussed in Chapters 8 and 9 characterize the median ground motions for Central and Eastern North America (CENA). This chapter reviews the components of aleatory ground-motion variability, provides the framework used to develop standard deviation models for CENA, and summarizes the candidate standard deviation models evaluated for CENA. The evaluation of the candidate standard deviation models and the related weights assignment are described in Chapter 11.

10.1.1 Overview of the Components of Ground-Motion Variability

The components of ground-motion variability are described in Al Atik et al. (2010) and a brief review of the components and notations relevant to this study is provided here. Table 10–1 summarizes the adopted notation following Al Atik et al. (2010) for the components of the ground-motion residuals and their respective standard deviations. The total ground-motion residual (Δ_{es}) represents the difference in natural log (LN) units between the observed ground motion and the median predicted by GMMs. Total residuals can be separated into between-event residuals (δB_e) and within-event residuals (δW_{es}):

$$\Delta_{es} = \delta B_e + \delta W_{es} \tag{10–1}$$

where subscripts ‘e’ and ‘s’ refer to earthquake and station, respectively. As illustrated in Figure 1 of Al Atik et al. (2010), the between-event residual represents the average shift of the observed ground motion from an individual earthquake from the median ground motion predicted by the GMM. The within-event residual is the misfit between an individual observation of the earthquake at a particular station from the earthquake-specific median prediction of the GMM. The between-event and within-event residuals have standard deviations denoted as τ and ϕ, respectively, and are assumed to be uncorrelated. Regression analyses performed using different datasets as part of the NGA-East study proved this assumption to be valid. As a result, the total standard deviation σ can be written as:

$$\sigma = \sqrt{\phi^2 + \tau^2} \tag{10–2}$$

The total standard deviation σ is also referred to as ergodic standard deviation. Under the ergodic assumption (Anderson and Brune 1999), the ground-motion variability at a site is assumed to be equal to the ground-motion variability observed in a global dataset.

The within-event residual can in turn be separated into the site-to-site term $\delta S2S_s$ and the single-station within-event residual δWS_{es}:

$$\delta W_{es} = \delta S2S_s + \delta WS_{es} \tag{10–3}$$

where $\delta S2S_s$ represents the systematic deviation of the ground motion at site 's' from the median event-corrected ground motion predicted by the GMM. δWS_{es} is the site-corrected and event-corrected residual referred to as single-station within-event residual. The site-to-site term and the single-station within-event residuals have standard deviations denoted as ϕ_{S2S} and ϕ_{SS}, respectively, and are assumed to be uncorrelated; this assumption was validated based on the results of the regression analyses performed as part of this study. The total single-station standard deviation (single-station sigma, σ_{SS})—sometimes referred to as the partially non-ergodic sigma—and the ergodic standard deviation (σ)—can be written as:

$$\sigma_{SS} = \sqrt{\phi_{SS}^2 + \tau^2} \tag{10–4}$$

$$\sigma = \sqrt{\phi_{S2S}^2 + \phi_{SS}^2 + \tau^2} \tag{10–5}$$

10.1.2 Use of Ergodic and Single-Station Sigma

Probabilistic seismic hazard analysis (PSHA) studies typically aim to evaluate the hazard at a single site. This requires characterizing the ground motion (median and variability) at the particular site due to the future occurrence of different earthquakes. Since repeated recordings at the site of interest are usually not available, the ergodic assumption (Anderson and Brune 1999) is adopted, whereby the ground-motion variability over different sites and source regions is assumed applicable to an individual site. In such cases, the use of the ergodic assumption is valid as the PSHA is typically concerned with rare events with the likelihood of repeated occurrences being very low.

More recently, the availability of repeated recordings at individual stations in ground-motion databases allowed for the estimation of the site-to-site variability and for removing it from the ground-motion aleatory variability. This led to a reduced aleatory variability as observed in previous studies [e.g., Chen and Tsai (2002); Atkinson (2006); Morikawa et al. (2008); Lin et al. (2011); Rodriguez-Marek et al. (2011); Chen and Faccioli (2013); Rodriguez-Marek et al. (2013); and Luzi et al. (2014)]. The application of the partially non-ergodic approach, however, requires adjusting the median GMMs from representing average site conditions in the dataset used to develop the models to become site specific. Site-specific adjustments of the median GMMs can be done using empirical ground-motion recordings when available at or near the site of interest or as part of the more commonly used analytical site response studies. The epistemic uncertainty resulting from the limited knowledge of the site-specific parameters and adjustment factors needs to be properly accounted for as part of the partially non-ergodic approach.

In addition to the reduction in aleatory variability, the use of single-station sigma has multiple advantages. It allows for proper identification of the various components of ground-motion variability where some of the apparent randomness can be transformed to epistemic uncertainty. With the acquisition of additional data, this epistemic uncertainty can, in theory, be removed or decreased. Moreover, the use of single-station sigma has the advantage of avoiding double counting of the uncertainty in cases where the ergodic sigma is used and a site-specific response study performed. In such cases, the double counting of the uncertainty is the result of the site-to-site variability being part of the total ergodic sigma, and the uncertainty in the site amplification factors are also captured as part of the site response study.

In summary, some requirements need to be met for the application of the partially non-ergodic assumption. If these requirements are not met, ergodic sigma should be used. The use of single-station sigma in PSHA requires the estimation of the median site term ($\delta S2S_s$) and its associated epistemic uncertainty. This is typically done through site response analyses or, if available, by analysis of ground-motion data recorded at the site. When ground-motion recordings are used to estimate the site term, epistemic uncertainty arises from the limited number of recordings available. As the number of recordings increases, the epistemic uncertainty on the value of the site term typically decreases. On the other hand, site response analyses have inherent uncertainties both in the input parameters and the modeling processes, and these lead to epistemic uncertainty in the estimated site term. Another requirement for the application of partially non-ergodic sigma is that the epistemic uncertainty in the single-station within-event standard deviation be estimated and accounted for in the ground-motion logic tree. This requirement arises from observations that there is variability in the standard deviation of site- and event-corrected residuals at a single station ($\phi_{SS,S}$) compared to the average ϕ_{SS} estimated over all the stations in the database (Rodriguez-Marek et al. 2013). This uncertainty could be due to site-specific features such as topography or subsurface layering.

10.1.3 Framework for the Characterization of Aleatory Variability for CENA

The general framework for NGA-East consists of decoupling the median GMMs from the aleatory variability models (Section 6.4). Such approach is standard practice in ground-motion modeling [e.g., Renault et al. (2010), EPRI (2013), Rodriguez-Marek et al. (2014), Coppersmith et al. (2014), and GeoPentech (2015)]. This approach assumes that the developed aleatory variability models are applicable to the range of median GMMs that were developed to capture the epistemic uncertainty in the median ground-motion predictions. To explore the potential correlation between the median predictions and the aleatory variability, an exercise was undertaken using the ground-motion prediction equation (GMPE) developed by Kuehn and Abrahamson (2017) based on the Abrahamson et al. (2014) dataset whereby all the model coefficients, event terms, and site terms were simultaneously estimated using Bayesian inference via Markov–Monte Carlo sampling. A total of 800 samples were then drawn from the posterior distribution of the GMPE parameters (Kuehn, personal communication). Using these samples, the uncertainty in each parameter as well as the joint distribution of parameters could be evaluated. The uncertainty in the median ground-motion predictions for two earthquake scenarios (magnitude 6.0, distance of 20 km and magnitude 8.0, distance of 200 km; strike–slip events) at a site with V_{S30} of 760 m/sec was evaluated against the uncertainty in the standard deviations. Figures 10–1, 10–2, and 10–3 show the 800 median predictions for the magnitude 6.0 and distance 20-km scenario plotted against the posterior samples of the standard deviations for τ, ϕ_{SS}, and ϕ_{S2S}, respectively, for PGA. These plots show no correlation between the value of the standard deviation and that of the median prediction. In other words, the higher median predictions do not require smaller standard deviations and vice versa. These results are consistent with those observed for the magnitude 8.0 and distance 200-km scenario. A similar analysis was also performed using a GMPE developed based on European data for a range of spectral periods showing consistent results with those presented here using the results from the Kuehn and Abrahamson (2017) model at PGA (Kuehn, personal communication). Although this analysis was not specifically undertaken using the CENA dataset, it does indicate that the

aleatory variability model can be applied to the entire range of median GMMs developed for CENA.

Models for the components of ergodic sigma and single-station sigma were developed for the ground motion in CENA and are presented in this chapter. These models are intended for application to a magnitude range of 4.0 to 8.2, distance up to 1500 km, reference site conditions with average shear-wave velocity in the top 30 m (V_{S30}) of 3000 m/sec, and for peak ground acceleration (PGA), peak ground velocity (PGV), and spectral periods of 0.01 to 10 sec (frequencies of 0.1 to 100 Hz). The evaluation of the candidate aleatory variability models is presented in Chapter 11.

The adopted approach for the development of aleatory variability models for CENA consists first of a review of the existing standard deviation models applicable to CENA. New models for the individual components of ground-motion variability (τ, ϕ_{SS}, and ϕ_{S2S}) were developed using the CENA ground-motion dataset collected and processed as part of the NGA-East project. Additionally, this study makes use of ground-motion datasets in other regions—such as the Next Generation Attenuation Relationships for the western U.S. (NGA-West2) dataset (Ancheta et al. 2014) and the Japanese dataset (Dawood 2014) —to gain insight on the behavior of the ground-motion variability in magnitude, distance, and frequency ranges that are not well populated in the NGA-East database (Goulet et al. 2014).

Dividing the variability into its components allows a better understanding of the sources of ground-motion variability. Moreover, such a breakdown allows the use of single-station sigma for sites where site-specific analyses are performed or ground-motion recordings are available, thus avoiding double-counting of the site-to-site variability. Another advantage of this approach is that the value of the single-station within-event standard deviation (ϕ_{SS}) has been observed to be relatively constant across different regions and tectonic environments (Rodriguez-Marek et al. 2013). The observed lack of regional dependence in ϕ_{SS} allows the use of global datasets for estimating CENA ϕ_{SS}, thus bypassing the data limitations in the NGA-East dataset. As a result, logic trees were developed for CENA τ, ϕ_{SS}, and ϕ_{S2S}. These variability components were then combined into ergodic sigma and single-station sigma logic trees.

We note here that the seed GMM developers were not tasked with developing aleatory variability models for CENA as a result of the decision to decouple the median GMMs from the aleatory variability models discussed above. Moreover, most of the seed GMMs are based on non-empirical approaches (e.g., point-source stochastic simulations, finite-fault simulations, hybrid empirical approach, etc.), which are not well-calibrated for developing aleatory variability models. Although a few seed GMM developers did provide accompanying aleatory variability for their models, these aleatory variabilities were not considered for the development of the CENA aleatory variability models for the reasons discussed here. The aleatory variability models provided by the seed GMM developers are summarized in Section 10.2.3.

10.2 Existing Standard Deviation Models

10.2.1 Existing Ergodic Sigma (σ) Models

Previous CENA ground-motion studies developed ergodic standard deviation models. A comprehensive review of the existing ergodic standard deviation models for CENA can be found in Al Atik (2015). A review of the most recent models is presented here.

The EPRI (2013) study adopted the conclusion of the EPRI (2006) study that the aleatory variability of ground motions in Central and Eastern U.S. (CEUS) is similar to that in Active Tectonic Regions (ATRs). Therefore, EPRI (2013) based their aleatory variability model on the average of preliminary aleatory variability values of four NGA-West2 models [not including the Idriss (2014) GMM]. The EPRI (2013) aleatory variability model has magnitude breaks in the magnitude dependence at **M** at 5.0, 6.0, and 7.0. Similar to the conclusion reached by EPRI in 2006, the 2013 EPRI study increased τ by 0.03 LN units to adjust the values derived for ATRs for application in CEUS. In addition, EPRI (2013) concluded that the observed reduction in aleatory variability of ground motions in ATRs may not be applicable to CEUS due the greater high-frequency energy content of ground motions in CEUS than in ATRs. As a result, the values of τ and ϕ between 10 and 40 Hz were set equal to the values at 10 Hz to account for the increase in high-frequency content of the CEUS ground motions. The favored EPRI model (2013) (weight 0.6) shows no increase in the aleatory variability at small values of R_{JB}, while its alternative model included an additional aleatory component with a maximum of 0.16 for R_{JB} values less than 10 km.

Recently, Atkinson (2013) evaluated empirical ground-motion data and also concluded that aleatory variability in CENA ground motions should be similar to that in ATRs. Atkinson and Adams (2013) have implemented this concept in proposed ground-motion models for use in updating the seismic hazard maps for Canada.

10.2.2 Existing Single-Station Sigma (σ_{SS}) Models

The development and use of single-station standard deviation models has not been common practice in CENA primarily due to data limitations and the insufficient number of repeated recordings at individual stations. Atkinson (2013) used ground-motion data from small-to-moderate magnitude earthquakes recorded at six stations in the Charlevoix Region to evaluate the aleatory variability in Eastern North America (ENA) ground motions. Atkinson (2013) estimated single-station sigma to be in the range of 0.23–0.28 in log10 units (0.53–0.64 in LN units), and concluded that single-station sigma in ENA is similar to that observed in California. Since single-station within-event standard deviation (ϕ_{SS}) models have been observed to be stable across regions and tectonic environment, it is worthwhile to summarize the existing recent ϕ_{SS} models outside of CENA developed from large datasets.

1. Rodriguez-Marek et al. (2013) developed ϕ_{SS} models derived from a global dataset with **M** ≥ 4.5 and rupture distance R_{RUP} ≤ 200 km compiled as part of the PEGASOS Refinement Project (PRP) (Renault et al. 2010). The PRP dataset consisted of residuals of ground-motion data from California, Taiwan, Switzerland, Japan, and Turkey. Rodriguez-Marek et al. (2013) observed that the value of ϕ_{SS} appears to be

largely region-independent, with an average of 0.45 (natural log units) across all periods. Rodriguez-Marek et al. (2013) also observed that ϕ_{SS} shows a magnitude-dependent trend with values decreasing from a maximum at **M** = 5.0 to a minimum at **M** = 7.0 and a distance-dependence for small magnitude events. As a result, three candidate models were adopted in the PRP (referred to as PRP models): constant ϕ_{SS} model (homoscedastic), distance-dependent ϕ_{SS} model, and magnitude- and distance-dependent ϕ_{SS} model.

2. The Thyspunt Nuclear Siting Project (TNSP) in South Africa (Rodriguez-Marek et al. 2014) developed ϕ_{SS} models based on the PRP data with **M** ≥ 5.0. The TNSP adopted two ϕ_{SS} models: a homoscedastic model with ϕ_{SS} = 0.45 (LN units) and a magnitude-dependent model.
3. The Hanford Probabilistic Seismic Hazard Analysis Project, referred to herein as the Hanford Project (Coppersmith et al. 2014), developed ϕ_{SS} models for crustal and subduction earthquakes. For crustal earthquakes, the Hanford Project used the PRP data with **M** ≥ 5.0 to develop a magnitude-dependent ϕ_{SS} model.
4. The Southwestern United States Ground Motion Characterization SSHAC Level 3 Project (SWUS) (GeoPentech 2015) used a global dataset consisting of the NGA-West2 data, which was supplemented with the Lin et al. (2011) data from Taiwan [as well as the European dataset of Akkar et al. (2014)], to develop ϕ_{SS} models for the Diablo Canyon Power Plant (DCPP) and the Palo Verde Nuclear Generating Station (PVNGS) sites. Five ϕ_{SS} models were developed for the SWUS project, focusing on magnitude and distance ranges of importance to the hazard at DCPP and PVNGS. Two ϕ_{SS} models were developed using the global dataset (NGA-West2 and additional Taiwanese data). These global ϕ_{SS} models are homoscedastic and were derived using data with **M** ≥ 5.0 and R_{RUP} less than 50 km, and with **M** ≥ 5.5 and R_{RUP} between 200 and 400 km, respectively. Two magnitude-dependent ϕ_{SS} models were derived using California NGA-West2 ground-motion data with R_{RUP} less than 50 km. These models have magnitude breaks at **M** = 5.0 and 7.0, and **M** = 5.0 and 5.5, respectively. A homoscedastic ϕ_{SS} model was developed using the European dataset of Akkar et al. (2014) with **M** ≥ 5.0 and distance less than 50 km. The ϕ_{SS} models developed for the SWUS project were generally comparable to other ϕ_{SS} models particularly those developed over similar magnitude and distance ranges.

10.2.3 Aleatory Variability for Seed GMMs

As discussed in Section 10.1.3, the seed GMM developers were not tasked with providing aleatory variability models to accompany their median GMMs. This is based on the early decision made by the Technical Integrator (TI) team to decouple the median GMMs from the aleatory variability models. Additionally, the majority of the CENA seed GMMs are based on non-empirical methods, which were not necessarily well calibrated for evaluating aleatory variability. As a result, only five GMM developers provided estimates of aleatory variability along with their median models. These aleatory variability values documented in the PEER report (2015) were developed for verification purposes of the median models and were not intended for use in forward applications. For instance, the seed GMM developers did not examine trends

of their aleatory variability with parameters such as magnitude, distance, V_{S30}, etc. In some cases, the individual components of the aleatory variability (e.g., τ, ϕ_{SS}, and ϕ_{S2S}) were not provided. As a result, the TI team did not consider the aleatory variability provided by some of the seed GMM developers as candidate models in the development of the CENA aleatory variability models. A brief overview of the aleatory variability provided by some of the seed GMM developers is summarized here for the sake of completeness.

A total of five seed GMM developers provided estimates of the aleatory variability accompanying their median GMMs as documented in the PEER report (2015). Estimates of standard deviations were provided by: Darragh et al. (DASG), Al Noman and Cramer (ANC15), Graizer, Hassani and Atkinson (HA15), and Hollenback et al. (PEER models). We note that other seed GMM developers [e.g., Pezeshk et al. (PZCT) and Shahjouei and Pezeshk (SP15)] did evaluate aleatory variability, but their estimates were not provided to the TI team as this was considered outside of their intended scope of developing median GMMs for CENA (PEER 2015). We note that the tabular values of the aleatory variability of the seed GMMs were not always provided and, in some cases, only plots of standard deviations were documented (PEER 2015).

1. The DASG Team developed four GMMs based on the point-source stochastic simulation method. They presented estimates of aleatory variability for each of the GMMs [refer to Table 3.5 and Figure 3.14 of the PEER report (2015)]. Their total aleatory variability is based on summing the variances associated with parameter variations, point-source modeling of past earthquakes, and regression fit. Their total sigma values tend to range from 0.7 to 0.8 (LN units) at high frequency to 1.3 (LN units) at low frequency.
2. The ANC15 Team developed their median GMM based on the traditional pseudo-spectral acceleration-based (PSA-based) empirical approach. Their total standard deviation values are listed in Table 8.1 of the PEER report (2015) and range in values between 0.9 and 1.2 (LN units). Similarly, the Graizer median GMM was developed using a PSA-based empirical approach. The aleatory variability of this model is based on the standard error of the residuals and is presented in Figure 9.25 of the PEER report (2015), ranging from around 0.8 to 1.0 (LN units).
3. The HA15 Team developed their median GMM based on the referenced empirical approach of Atkinson (2008). Their within-event and between-event standard deviations are listed in Table 10.1 of the PEER report (2015). Their total standard deviations range in values between 0.7 and 0.9 (LN units). The authors discussed the reasons behind their relatively large standard deviation values, which they attributed to the wide ranges of site conditions and distances and to the predominance of small magnitudes in the CENA dataset.
4. The PEER median GMM was derived based on the empirical approach using the Fourier amplitude spectral (FAS) values of the CENA dataset. Figure 11.2 of the PEER report (2015) shows the components τ, ϕ_{SS}, ϕ_{S2S}, and ϕ of the aleatory variability for this FAS-based model for frequencies between 0.6 and 10 Hz. This figure shows that τ ranges between 0.38 to 0.5-0.6 (LN units) at the edges of the frequency range used. ϕ_{SS} decreases from around 0.5 (LN units) at f = 0.6 Hz to

around 0.38 at f = 10 Hz. Figure 11.2 of the PEER report (2015) shows that ϕ_{S2S} is on the order of 0.52 to 0.68 (LN units) and ϕ is relatively large (on the order of 0.66 to 0.76 LN units).

10.3 Candidate Standard Deviation Models for CENA

Candidate standard deviation models (τ, ϕ_{SS}, and ϕ_{S2S}) were selected and evaluated for CENA using the NGA-East dataset, as well as models developed based on other datasets such as the NGA-West2 project. Al Atik (2015) analyzed the components of ground-motion variability using the NGA-East dataset and describes in detail the development of the candidate models for τ, ϕ_{SS}, and ϕ_{S2S} for CENA. We summarize here the datasets used in the evaluation of the ground-motion variability as well as the candidate models for τ, ϕ_{SS}, and ϕ_{S2S} (best estimate and standard deviation of the corresponding variance). Updates made to the original ϕ_{S2S} model presented in Al Atik (2015) are discussed in this chapter. An evaluation of the regionalization of residuals in CENA is presented in Appendix F.1.

10.3.1 Datasets

10.3.1.1 CENA

10.3.1.1.1 Within-Event and Between-Event Residuals Analysis

The CENA dataset compiled and processed by the NGA-East project was used to evaluate ground-motion residuals and components τ, ϕ_{SS}, and ϕ_{S2S} of ground-motion variability in CENA. Recordings were excluded from the analysis if they fit one or more of the following criteria:

1. Recordings flagged for having data issues known to impact ground motion; these recordings were only excluded at the particular frequencies that were affected by issues such as observed microseismic noise or being outside of the filter useable frequency range;
2. Recordings flagged for having residuals with respect to Atkinson and Boore (2006 and 2011) that fall outside of +/-4sigma at all of PGA, PGV, and pseudo-spectral acceleration (PSA) (0.05 sec); these recordings were considered outliers and potentially impacted by undetected data issues;
3. Recordings flagged for having the station, earthquake, or both located outside of the CENA regions defined in Chapter 4 [recordings with path region = -999 as shown in Figure C.8 of Goulet et al. (2014)] with the exception of the three recordings from the Nahanni earthquake; these recordings were kept because they were the only events in the CENA flatfile with magnitudes greater than 6.0;
4. Recordings flagged for having the station, earthquake, or both located in the Gulf Coast Region (GCR) mapped in Figure 4–6; ground-motion recordings from the GCR were treated in a separate analysis as discussed in Section 11.9;
5. Recordings with R_{RUP} distance greater than 500 km; these recordings were excluded from the analysis for being potentially unreliable at such large distances recorded mostly from small magnitude earthquakes; and

6. Earthquakes with fewer than three recordings each. The minimum number of recordings per earthquake is discussed in Section 4.2 of Al Atik (2015). The minimum of three recordings per earthquake criterion was adopted to ensure a large enough dataset as well as reliable estimates of event terms.

The CENA ground-motion data that passed the criteria listed above were used in the within-event and between-event analysis. The magnitude and distance distribution of the resulting CENA dataset is shown in Figure 10–4 at a frequency f = 4 Hz. Figure 10–4 shows that CENA earthquakes in the dataset range in magnitude from 2.57 to 6.76, and consist of tectonic earthquakes and potentially induced events (PIEs). Table 10–2 lists the number of PIEs and tectonic recordings, earthquakes, and stations in the dataset at f = 4 Hz. The magnitude and hypocentral depth distribution of the CENA earthquakes in the dataset is shown in Figure 10–5, with PIEs having shallower average depth than the tectonic events. The CENA stations used in the within-event and between-event residuals analysis have assigned V_{S30} values ranging from 144 to 2000 m/sec. Figure 10–6 shows the number of stations in different V_{S30} bins at f = 4 Hz. We note that a relatively small number of stations in the dataset (53 stations at f = 4 Hz) had measured V_{S30} values, and other proxies were used to assign V_{S30} at the station locations where measurements were unavailable.

Figure 10–7 shows the total number of recordings versus frequency, and indicates that frequencies outside 0.67 to 13.33 Hz suffer from a significant reduction in the total number of recordings due to limitations on the useable frequency bandwidth of the recordings and therefore cannot be reliably used to evaluate ground-motion variability. The ratio of the number of recordings at each frequency outside of 0.67–13.33 Hz to that at 4 Hz is less than 60%. We note that the useable frequency bandwidth for each recording is dictated by the corner frequencies of the low-pass and high-pass filters applied as discussed in Section 5.2.2.4. Table 10–3 lists the number of PIEs and tectonic recordings, earthquakes, and stations in the dataset at f = 25 Hz. Compared to the numbers presented in Table 10–2 for f = 4 Hz, we observe a significant drop in the number of recordings at f = 25 Hz compared to 4 Hz where the number of recordings at f = 25 Hz is 28% that at f = 4 Hz. Similarly, a sharp reduction in the number of stations is observed at f = 25 Hz compared to that at f = 4 Hz. Figure 10–8 compares the magnitude and distance distribution of the tectonic earthquakes in the CENA dataset at f = 4 and 25 Hz. This figure indicates that the magnitude and distance range for f = 4 Hz is better populated by ground-motion data than that for f = 25 Hz, particularly for **M** > 4.0.

As a result of the frequency bandwidth limitations discussed above, the TI team decided to consider the frequency range of 1 to 10 Hz, where the ratio of the number of recordings with respect to that at f = 4 Hz is greater than 80% (reduction in the number of recordings is within 20%), to be reliable for evaluating ground-motion variability and thus for subsequently developing aleatory variability models using the CENA dataset. We note that the results of the ground-motion residuals analysis using the CENA data are presented in the figures of this chapter as well as Chapter 11 for the entire frequency range, although only the 1 to 10 Hz range is considered reliable. The reliable results are indicated by the solid vertical lines in the figures at 1 and 10 Hz.

Aleatory variability of ground motion is best evaluated using residuals of a ground-motion model that fits the available empirical data. The CENA ground-motion data that passed the criteria

listed above were used to derive a new median GMM for the sole purpose of analyzing ground-motion variability (within-event and between-event residuals analysis). This model is not intended for use in median ground-motion predictions and is not necessarily well constrained for extrapolation outside of the limited magnitude and distance range of the CENA dataset. The model for the median ground-motion predictions is described in Al Atik (2015). The mixed-effects algorithm described in Abrahamson and Youngs (1992) was used for the regression analysis and the separation of the total residuals into between-event and within-event residuals. The model for the median ground motion has the following functional form:

$$\ln(y) = c_1 + c_2\mathbf{M} + c_3\mathbf{M}^2 + c_4 \ln\left(\sqrt{R_{RUP}^2 + c_6^2}\right) + c_7\left(\sqrt{R_{RUP}^2 + c_6^2}\right) + c_8 \ln(V_{S30}) + c_9\left(\frac{H_{dep}}{20}\right) \text{for } R_{RUP} < 50 \text{ km}$$

$$\ln(y) = c_1 + c_2\mathbf{M} + c_3\mathbf{M}^2 + (c_4 - c_{4h})\ln\left(\sqrt{50^2 + c_6^2}\right) + c_{4h} \ln\left(\sqrt{R_{rup}^2 + c_6^2}\right) + c_7\left(\sqrt{R_{rup}^2 + c_6^2}\right) + c_8 \ln(V_{S30}) + c_9\left(\frac{H_{dep}}{20}\right)$$

$$\text{for } R_{RUP} \geq 50 \text{ km} \qquad (10\text{–}6)$$

where H_{dep} is the hypocentral depth (km). The geometrical spreading term hinges at an R_{RUP} distance of 50 km and its coefficient for distances greater than 50 km (c_{4h}) is period-dependent determined based on SMSIM simulations (Boore 2005). Coefficients of geometrical spreading within 50 km and anelastic attenuation, c_4 and c_7, distinguish between PIE and tectonic events, and have different values for earthquakes located in the Oklahoma-Arkansas region (mostly consisting of PIEs). Coefficient c_6, commonly referred to as "fictitious depth," was fixed to the same form used in the Abrahamson et al. (2014) (ASK14) model:

$$c_6 = \begin{cases} 1 & \text{for } \mathbf{M} \leq 4.0 \\ 3.5\mathbf{M} - 13 & \text{for } 4.0 < \mathbf{M} \leq 5.0 \\ 4.5 & \text{for } \mathbf{M} > 5.0 \end{cases} \qquad (10\text{–}7)$$

Coefficients c_1, c_2, c_3, c_4, c_7, c_8, and c_9 were determined from the regression, and residuals plots were given in Al Atik (2015). The coefficients of the GMM derived for the purpose of analyzing ground-motion variability are listed in Table 10–4.

10.3.1.1.2 Single-Station Residuals Analysis

The single-station analysis was performed using a subset of the CENA dataset used for the within-event and between-event residuals analysis and discussed in Section 10.3.1.1. This subset consisted of stations that recorded a minimum of three earthquakes each. The minimum number of recordings per station was discussed in Sections 4.4 and 4.5 of Al Atik (2015). Similar to the number of recordings per earthquake, a minimum of three recordings per station was applied to ensure a stable estimate of site terms as well as a large enough dataset. Figure 10–9 shows the magnitude and distance distribution of the recordings, and Table 10–5 summarizes the number of recordings, earthquakes, and stations used in this analysis. Figure 10–10 shows the number of recordings versus frequency used in the single-station analysis. Similar to the observations made on Figure 10–7 (number of recordings versus frequency for the dataset used in the within-event and between-event residuals analysis), Figure 10–10 indicates that frequencies outside the range of 0.5 to 13.33 Hz suffer from a large reduction in

the dataset (more than 60%) and cannot be reliably used to evaluate ϕ_{SS} and ϕ_{S2S} due to the useable frequency limitations of the recordings (refer to Section 5.2.2.4 on filtering). Moreover, the TI team decided to adopt the frequency range of 1 to 10 Hz, where the reduction in the number of recordings is within 20%, for the purpose of evaluating the site terms and single-station within-event residuals and developing models for ϕ_{SS} and ϕ_{S2S} using the CENA data.

The mixed-effects regression (Abrahamson and Youngs 1992) was used to separate the within-event residuals into site terms and single-station within-event residuals. Similar to the approach outlined in Abrahamson and Youngs (1992), the maximum likelihood solution for the random effects, $\delta S2S_s$, can be written as:

$$\delta S2S_s = \frac{\phi_{S2S}^2 \sum_{e=1}^{N_s} \delta w_{es}}{\phi_{SS}^2 + N_s \phi_{S2S}^2} \qquad (10\text{–}8)$$

An algorithm similar to the one outlined in Abrahamson and Youngs (1992) for estimating the event terms and the within-event residuals was adopted for the estimation the site terms and single-station within-event residuals, ϕ_{SS}, and ϕ_{S2S}.

10.3.1.2 Western United States (WUS)

The NGA-East study relies on comparisons of the components of ground-motion variability in CENA to those observed in western U.S. (WUS) to inform the extrapolation of aleatory variability in CENA to magnitude, distance, and frequency ranges not well covered in the CENA dataset. Four sets of NGA-West2 GMM within-event and between-event residuals were used by the TI team for this analysis: Abrahamson et al. (2014) (ASK14), Boore et al. (2014) (BSSA14), Campbell and Bozorgnia (2014) (CB14), and Chiou and Youngs (2014) (CY14). Idriss (2014) residuals were not used because within-event and between-event residuals were not available for this GMM. The NGA-West2 dataset consists mostly of recordings from California as well as some recordings from other regions such as Taiwan, Japan, China, etc. The data distributions of the four sets of NGA-West2 residuals were discussed in the Earthquake Spectra Special Issue on the NGA-West2 Project.

The NGA-West2 single-station analysis was performed by the TI team using stations that recorded a minimum of three earthquakes each. The mixed-effects algorithm (Abrahamson and Youngs 1992) was used to separate the within-event residuals obtained from the NGA-West2 developers into site terms and single-station within-event residuals. Plots of the magnitude-distance distribution, V_{S30} histograms, and number of recordings histograms of the subsets of the NGA-West2 data are provided in Al Atik (2015). The four NGA-West2 datasets consist of recordings with **M** = 3.0 to 8.0 and distances up to 400 km. The CB14 Team performed their regression on ground-motion data with distance less than 80 km and then applied an additional distance scaling term for distance greater than 80 km. Most of the recording stations have V_{S30} in the range of 200 to 600 m/sec. A summary of the number of recordings, earthquakes, and stations is listed in Table 10–6 for the four GMMs.

10.3.1.3 Japan

The NGA-East study made use of the single-station analysis performed on Japanese crustal ground-motion data by Dawood (2014) to support the development of the aleatory variability

model for CENA, particularly the site-to-site standard deviation ϕ_{S2S}. Insights obtained from the Japanese dataset were useful because of the general similarity in the site conditions between Japan and CENA (shallow soil cover overlying rock). The availability of ground-motion recordings at the surface and deep in the borehole in the Japanese dataset helped in the extrapolation of the ϕ_{S2S} for CENA for application to hard rock conditions with V_{S30} of 3000 m/sec.

A detailed description of the Japanese dataset and the development of a ground-motion model for active crustal earthquakes in Japan are detailed in Dawood (2014). A brief summary of the dataset relevant to the NGA-East study is presented herein. The Japanese dataset consists of ground-motion recordings from active crustal earthquakes recorded on the KiK-net stations. The dataset is comprised of 13,735 six-component (three at the ground surface level and three deep within the borehole) ground-motion recordings from 679 active crustal earthquakes recorded at 643 stations. The V_{S30} values calculated for the KiK-net stations are based on seismic velocity profiles from downhole PS logging. Dawood (2014) used a GMPE functional form adopted from ASK14 for their regression analysis. Their ϕ_{S2S} results for the surface and the borehole levels were used for comparisons in the NGA-East study. Moreover, the within-event residuals were provided by the authors and used in the NGA-East study to perform single-station regressions and evaluate trends in ϕ_{S2S} for different aspects of the dataset.

Figure 4.47 of Al Atik (2015) shows the magnitude and distance distribution of the Japanese dataset. It consists of ground-motion recordings with **M** range of 4.0 to 7.0 and distance R_{RUP} of up to about 350 km. The number of stations histogram for the different V_{S30} bins is shown in Figure 10–11. The average V_{S30} for the surface recordings is 499 m/sec. The Japanese dataset provides borehole depth and the shear-wave velocity (V_S) at the borehole level, which is denoted $V_{S,Zhole}$. Figures 10–12 and 10–13 show the histograms of the number of stations versus bins of borehole depth and $V_{S,Zhole}$, respectively. The majority of the Japanese stations have a borehole depth of 98 to 150 m with 31 stations having a borehole depth greater than 500 m.

10.3.2 Tau (τ) Models

The CENA dataset covers a limited magnitude range and does not allow a reliable extrapolation of τ for magnitudes greater than 5.5. Moreover, CENA τ values at frequencies outside of 1 to 10 Hz range are not reliable due to the frequency bandwidth limitations of the recordings (refer to the discussion of the CENA dataset in Section 10.3.1.1.1). As a result, the TI team considered three candidate models for τ:

- Global τ model based on the average of the four proposed NGA-West2 τ models (ASK14, BSSA14, CB14, and CY14)
- CENA constant τ model (homoscedastic)
- CENA magnitude-dependent τ model

The derivation of the three candidate τ models [mean and standard deviation, SD(τ^2)] is described in Al Atik (2015) and is summarized below.

10.3.2.1 Global Tau (τ) Model

The global τ model is based on the average of τ^2 for the four NGA-West2 models (ASK14, BSSA14, CB14, and CY14). These models were chosen because they were derived from a large and uniformly-processed global dataset and are applicable to a large magnitude range (**M** =3.0 to 8.0 or 8.5). All four NGA-West2 τ models are magnitude-dependent, and all models except ASK14 are also period-dependent.

Figure 10–14 shows the four NGA-West2 τ models and their average at frequencies of 1 and 100 Hz. The proposed global model is also shown on the plots. It follows closely the average τ and has four magnitude breaks at **M** = 4.5, 5.0, 5.5, and 6.5. The global τ model has the following functional form:

$$\tau = \begin{cases} \tau_1 & \text{for } \mathbf{M} \le 4.5 \\ \tau_1 + (\tau_2 - \tau_1) * \dfrac{(\mathbf{M} - 4.5)}{0.5} & \text{for } 4.5 < \mathbf{M} \le 5.0 \\ \tau_2 + (\tau_3 - \tau_2) * \dfrac{(\mathbf{M} - 5.0)}{0.5} & \text{for } 5.0 < \mathbf{M} \le 5.5 \\ \tau_3 + (\tau_4 - \tau_3) * \dfrac{(\mathbf{M} - 5.5)}{1.0} & \text{for } 5.5 < \mathbf{M} \le 6.5 \\ \tau_4 & \text{for } \mathbf{M} > 6.5 \end{cases} \quad (10\text{–}9)$$

Figure 10–15 shows the derived model coefficients as a function of frequency, and indicates that the average τ at the magnitude breaks is not constant with frequency but fluctuates around constant values. An upward bump can be observed in the average τ at a frequency of around 10–20 Hz. This bump can also be seen in some of the underlying models (BSSA14 and CB14), while the rest of the models smoothed through it (CY14 and ASK14). The origin of this bump has been investigated through point-source simulations as part of the Hanford Project (Coppersmith et al. 2014). A first set of 250 point-source simulations were conducted with an average stress drop of 50 bars and a logarithmic standard deviation of 0.5. Ground motions were computed at 50 sites per earthquake using a log-normal distribution of kappa with median of 0.035 sec and a standard deviation of 0.3 (LN units). The WUS amplification factors were used with a frequency-independent site factor that is log-normally distributed around zero with a standard deviation of 0.4 (LN units). The resulting ϕ and τ values are shown in Figure 10–16 and show a bump in ϕ but not in τ.

A second set of simulations were conducted allowing for correlation between earthquakes and kappa values. This correlation would result if there are regional kappa differences resulting in earthquakes sampling particular ranges of kappa values. The uncertainty in kappa was divided between a median value for each earthquake and a within-earthquake distribution, thus preserving the total variance of kappa; all other parameters were kept the same. The resulting standard deviations are presented in Figure 10–17, and show that the correlation between earthquakes and kappa values result in the bump occurring now in both ϕ and τ. Based on this analysis, the TI team concluded that the bump observed in τ at around 10 to 20 Hz is likely to be

an artifact of kappa in the NGA-West2 dataset; thus it was smoothed through and constant τ versus frequency was adopted, as shown with the dashed lines in Figure 10–15.

The standard deviation of τ^2 $\left[\mathrm{SD}(\tau^2)\right]$ consists of two components: within-model $\left[\mathrm{SD}(\tau_W^2)\right]$ and between-model variability $\left[\mathrm{SD}(\tau_B^2)\right]$, as shown below:

$$\mathrm{SD}(\tau^2) = \sqrt{[\mathrm{SD}(\tau_W^2)]^2 + [\mathrm{SD}(\tau_B^2)]^2} \qquad (10\text{–}10)$$

The between-model variability is the standard deviation of τ^2 for the four underlying GMMs. The within-model variability calculated as part of the regressions conducted for the CY14 model (R. Youngs, *personal communication*) was used here and represents the statistical uncertainty in their τ^2 estimates. The between-model, within-model, and total variability in τ^2 is shown in Figure 10–18 at the magnitude breaks of **M** = 4.5, 5.0, 5.5, and 6.5, respectively. The standard deviations were smoothed with a constant across frequencies similar to the mean model. Figure 10–18 shows that the total variability in τ^2 is largest at **M** = 4.5 and 5.0, and decreases as **M** increases from **M** = 5.5 to **M** = 6.5.

10.3.2.2 CENA Constant Tau (τ) Model

CENA tectonic data with **M** greater than 3.0 were used to construct a τ model. Since CENA data are limited in magnitude range to a maximum **M** of about 5.5, two alternative models (constant and magnitude-dependent) were evaluated using the CENA data to address the uncertainty in the extrapolation of τ to magnitudes larger than 5.5.

Figure 10–19 shows the magnitude-independent CENA τ values as a function of frequency obtained from the mixed-effects regression (Abrahamson and Youngs 1992). Figure 10–19 shows that the CENA τ values appear to be constant in the frequency range of 1 to 10 Hz. Outside of this frequency range, a different trend of τ versus frequency can be observed. As discussed in Section 10.3.1.1.1, CENA τ values are not considered reliable outside the frequency range of 1 to 10 Hz due to the frequency bandwidth limitations of the data. Therefore, the observed trend of τ versus frequency outside of 1 to 10 Hz is not necessarily representative of a true trend and is affected by data limitations. As a result, τ values obtained outside of the frequency range of 1 to 10 Hz were not considered in the τ model development. CENA τ values were averaged in the frequency range of 1 to 10 Hz to obtain the proposed CENA constant τ model (τ = 0.37 LN units), which is magnitude-independent and period-independent.

Figure 10–20 shows the standard deviation in τ^2 obtained from the mixed-effects regression and represents the statistical uncertainty in τ^2. The standard deviation in τ^2 is calculated as outlined in Searle (1971) (page 474, part d). These standard deviation values were averaged between 1 and 10 Hz to obtain the proposed variability in τ^2. We note that the resulting variability in τ^2 is smaller than that observed for the global τ^2 model, which is based on a bigger dataset. Recall that the variability in the global τ^2 model consists of two components: between-model variability resulting from the use of four NGA-West2 τ models and within-model variability based on the CY14 calculated statistical uncertainty in τ^2 (Youngs, personal communication). The within-model uncertainty in the global τ^2 model is generally larger than the between-model uncertainty,

particularly for **M** less than 6.5. Because a lower statistical uncertainty for the CENA τ^2 model compared to the global τ^2 model is difficult to justify given the smaller CENA dataset, the TI team decided not to use the standard deviation of the constant τ^2 model computed using the CENA data. Instead, the variability in the global τ^2 model at **M** = 5.0 was adopted for the variability in the constant CENA model.

10.3.2.3 CENA Magnitude-Dependent Tau (τ) Model

Studies of between-event variability based on large datasets that cover a wide magnitude range generally note a magnitude-dependent trend in τ, whereby τ decreases as **M** increases and reaches a constant value at **M** = 6 to 7.5 (ex. NGA-West2 models). As a result, a CENA τ model was developed to incorporate the magnitude-dependence observed in the global τ models. This model is derived using CENA tectonic data with magnitude greater than or equal to 3.0. The model has the following form:

$$\tau = \begin{cases} \tau_1 & \text{for } \mathbf{M} \le 5.0 \\ \tau_1 + (\tau_2 - \tau_1) * \dfrac{(\mathbf{M} - 5.0)}{0.5} & \text{for } 5.0 < \mathbf{M} \le 5.5 \\ \tau_2 + (\tau_3 - \tau_2) * (\mathbf{M} - 5.5) & \text{for } 5.5 < \mathbf{M} \le 6.5 \\ \tau_3 & \text{for } \mathbf{M} > 6.5 \end{cases} \qquad (10\text{–}11)$$

where the ratios τ_2/τ_1 and τ_3/τ_1 were constrained based on the global τ model; CENA data were therefore used to solve for τ_1. Figure 10–21 shows the values of τ_1 obtained from the mixed effects regression as a function of frequency. The average of τ_1 values in the frequency range of 1 to 10 Hz was used to smooth τ_1 versus frequency. Figure 10–21 also shows the resulting τ_2 and τ_3 as functions of frequency. Figure 10–22 shows the $SD(\tau_1^2)$ obtained from the regression. These values were again smoothed with a constant equal to the average of $SD(\tau_1^2)$ between frequencies of 1 and 10 Hz. Figure 10–22 also shows the SD(τ^2) for the global model at the magnitude breaks **M** = 4.5, 5.0, 5.5, and 6.5, and indicates that the variability in the global τ^2 model at all magnitudes is larger than that for the CENA magnitude-dependent model, which is based on a smaller dataset. As a result, values of SD(τ^2) for CENA at **M** = 5.0, 5.5, and 6.5 were adopted from the global model.

10.3.3 PhiSS (ϕ_{SS}) Models

The shortcomings of the CENA dataset (limited magnitude and frequency ranges) drove the TI team to consider three candidate ϕ_{SS} models:

- Global ϕ_{SS} model based on the average of the four NGA-West2 ϕ_{SS} (ASK14, BSSA14, CB14, and CY14)
- CENA constant model (homoscedastic)
- CENA magnitude-dependent model

The derivation of the three candidate ϕ_{SS} models [mean and standard deviation, SD(ϕ_{SS}^2)] is described in Al Atik (2015) and summarized below. In general, the uncertainty in the derived ϕ_{SS} models has up to three components: (1) station-to-station variability in ϕ_{SS}, (2) statistical uncertainty in the ϕ_{SS} estimates, and (3) errors in the proposed model fit to the data. We note that, while the station-to-station variability in ϕ_{SS} is always included in estimating the uncertainty in the ϕ_{SS} models, components (2) and (3) of the uncertainty are not always explicitly included in estimates of SD(ϕ_{SS}^2) depending on their applicability to the specific models. A description of the uncertainty in the ϕ_{SS} models is given below specific to each model. Finally, the error in ϕ_{SS} resulting from smoothing the model coefficients is not included in the calculation of the uncertainty in the ϕ_{SS} models.

10.3.3.1 Global PhiSS (ϕ_{SS}) Model

For each of the four NGA-West2 GMMs, single-station within-event residuals were binned by magnitude in bins of 0.5 magnitude unit width, and ϕ_{SS} was calculated in each bin. Figure 10–23 shows the resulting four NGA-West2 ϕ_{SS} values as a function of magnitude at frequencies of 0.33, 1, 10, and 100 Hz. Plots of ϕ_{SS} versus magnitude at PGV and other frequencies can be found in Al Atik (2015). A clear trend of ϕ_{SS} with magnitude can be observed at high frequencies (short periods). At frequencies less than 1 Hz (periods greater than 1 sec), the magnitude dependence becomes weaker and ϕ_{SS} becomes magnitude-independent.

Values of ϕ_{SS}^2 in each magnitude bin were averaged for the four NGA-West2 models to obtain the average WUS ϕ_{SS} as a function of magnitude. A weighted linear fit, with weights based on the standard deviation of ϕ_{SS} in each magnitude bin, was applied at each period to the average ϕ_{SS} values as a function of magnitude, with magnitude breaks at **M** = 5.0 and 6.5. Figure 10–23 shows the average ϕ_{SS} values and the fit results at frequencies of 0.33, 1, 10, and 100 Hz. The global ϕ_{SS} model has the following form:

$$\phi_{SS} = \begin{cases} a & \text{for } \mathbf{M} \le 5.0 \\ a + (\mathbf{M} - 5.0) * \dfrac{(b - a)}{1.5} & \text{for } 5.0 < \mathbf{M} \le 6.5 \\ b & \text{for } \mathbf{M} > 6.5 \end{cases} \qquad (10\text{–}12)$$

Figure 10–24 shows the derived coefficients ‘a’ and ‘b’ versus frequency. These coefficients were smoothed while preserving their trend as a function of frequency.

The uncertainty in the global ϕ_{SS} model has two components: station-to-station variability in ϕ_{SS} and errors in the proposed model fit (Equation 10–12) to the data. The statistical uncertainty in the ϕ_{SS} estimates was used to derive the weighted linear fits of the mean ϕ_{SS} model versus magnitude. The station-to-station variability in ϕ_{SS} can be estimated across all sites in the dataset. It is a measure of variability in ϕ_{SS} from one site to another due to factors such as azimuthal dependency and topographic effects, as well as other unknown factors. The station-to-station variability in ϕ_{SS} was analyzed using the ASK14 dataset with **M** larger than or equal to 4.0 as part of the SWUS project (GeoPentech 2015). Their results have been adopted in the NGA-East project and their analysis is summarized here.

The estimate of ϕ_{SS} at an individual station is denoted as $\phi_{SS,S}$. The standard deviation of $\phi_{SS,S}$, SD($\phi_{SS,S}$), was estimated at all periods using the ASK14 dataset; however, the empirical estimates of $\phi_{SS,S}$ and their standard deviation are affected by sampling error that decreases as the number of recordings per site increases. To quantify the sampling error, a statistical exercise was undertaken whereby a large set of single-station within-event residuals was simulated per station for the same number of stations as in the ASK14 dataset such that the coefficient of variation, CV, of $\phi_{SS,S}$ [CV = SD($\phi_{SS,S}$) / mean($\phi_{SS,S}$)] is zero (all stations have the same $\phi_{SS,S}$ values). Assuming a normal distribution, $\phi_{SS,S}$ values were then computed at each station using multiple realizations of the dataset (different number of recordings per station), and the resulting CV of $\phi_{SS,S}$ was calculated for each realization. Figure 10–25 shows the CV values for the different number of realizations per site compared to the CV of $\phi_{SS,S}$ from the global dataset with different minimum numbers of recordings per station used in the regression for PGA and periods of 0.1 and 1.0 sec. The curve in Figure 10–25 represents the effect of pure sampling error on the estimates of the SD($\phi_{SS,S}$), and indicates that for large numbers of recordings per station, the sampling error decreases and approaches zero.

The statistical exercise was then repeated with different CV values (0.05, 0.10, 0.15) assigned for the simulations of the large dataset. The resulting CV for different realizations of the data are shown in Figure 10–26 where the CV of $\phi_{SS,S}$ for the empirical data fall between 0.10 and 0.15. A CV value of 0.12 was therefore adopted for calculating the station-to-station variability in ϕ_{SS}. The selected value of CV($\phi_{SS,S}$) = 0.12 was obtained assuming a homoscedastic ϕ_{SS}. However, it is reasonable to adopt this value for the chosen heteroskedastic model. It can be easily shown that for normally distributed residuals and for small values of CV($\phi_{SS,S}$), the CV of the variance (ϕ_{SS}^2) is twice the CV of the standard deviation (ϕ_{SS}); hence, the standard deviation of ϕ_{SS}^2, SD(ϕ_{SS}^2), can be computed as:

$$\mathrm{SD}\left(\phi_{SS}^2\right) = 0.24\phi_{SS}^2 \qquad (10\text{–}13)$$

Figures 10–27 and 10–28 show the components of the SD(ϕ_{SS}^2) for coefficients ‘a’ and ‘b’ of the global ϕ_{SS} model. For each coefficient, the total standard deviation of the variance, SD(ϕ_{SS}^2), consists of the site-to-site variability of ϕ_{SS}^2 calculated as shown in Equation (10–13), as well as the standard error of the coefficient squared estimated from the weighted linear fit to the ϕ_{SS} values versus magnitude. The total SD(ϕ_{SS}^2) is also shown in the figures.

10.3.3.2 CENA Constant PhiSS (ϕ_{SS}) Model

The dependence of ϕ_{SS} obtained from the analysis of CENA data on magnitude, distance, event type (tectonic/PIE), etc. has been discussed in Al Atik (2015). In light of the observations presented in Al Atik (2015), a subset of the CENA dataset discussed in 10.3.1.1.2 consisting of CENA tectonic data with minimum **M** of 3.0 and maximum R_{RUP} distance of 300 km was used to construct ϕ_{SS} models. The distance cut-off of 300 km was adopted in order to limit the slightly-smaller ϕ_{SS} values at large distances from biasing the average ϕ_{SS} (Al Atik 2015). Since CENA data are limited in magnitude range to a maximum **M** of about 5.5, two alternative models (constant and magnitude-dependent) were developed using the CENA data to address the uncertainty in the extrapolation of ϕ_{SS} to magnitudes larger than about 5.5. This section presents

the CENA constant (homoscedastic) ϕ_{SS} model while the magnitude-dependent ϕ_{SS} model is presented in Section 10.3.3.3.

Figure 10–29 shows the constant CENA ϕ_{SS} values as a function of frequency obtained from the mixed-effects regression. Because CENA data suffer from frequency bandwidth limitations (refer to Section 10.3.1.1), CENA ϕ_{SS} values were averaged in the frequency range of 1 to 10 Hz to obtain the proposed CENA constant ϕ_{SS} model (ϕ_{SS} = 0.51 LN units), which is magnitude-independent and period-independent.

The uncertainty in the CENA constant ϕ_{SS} model has two components: station-to-station variability in ϕ_{SS} and statistical uncertainty in the ϕ_{SS} estimates. The errors resulting from the smoothing of the model (CENA ϕ_{SS} values averaged in the frequency range of 1 to 10 Hz) are small and were not included in the estimate of the uncertainty in the model. The station-to-station variability in ϕ_{SS} can only be estimated for large numbers of recordings per station (minimum of 10 or more recordings per station). The CENA dataset does not have enough stations with number of recordings greater than 10 to estimate the station-to-station variability in a manner similar to that performed for the WUS data. As a result, the station-to-station variability of ϕ_{SS} obtained from the WUS data is adopted as a more robust estimate for CENA. By definition, the estimation of the station-to-station variability in ϕ_{SS} should not be affected by small datasets or by varied site conditions that impact ϕ_{S2S} but not ϕ_{SS}; therefore, the station-to-station variability of ϕ_{SS} obtained from the WUS data is considered reasonable to use for CENA. The impact of the limited dataset on the variability in ϕ_{SS} for CENA is reflected in the statistical uncertainty in the ϕ_{SS} estimates. Figure 10–30 shows the standard deviation in ϕ_{SS}^2, which includes the site-to-site variability obtained using CV(ϕ_{SS}^2) = 2*0.12 and the statistical uncertainty in the ϕ_{SS}^2 estimates obtained from the regression. The total standard deviation values were averaged between 1 and 10 Hz to obtain the proposed variability in ϕ_{SS}^2.

10.3.3.3 CENA Magnitude-Dependent PhiSS (ϕ_{SS}) Model

Previous single-station within-event standard deviation studies based on large global datasets that cover a wide magnitude range observed a magnitude-dependent trend of ϕ_{SS} whereby ϕ_{SS} decreases as **M** increases, reaching a constant value at **M** ≥ 6 (e.g., PRP, TNSP, Hanford, and SWUS projects). As a result, a CENA ϕ_{SS} model was developed to incorporate the magnitude-dependence observed in the global ϕ_{SS} model. Similar to the CENA constant ϕ_{SS} model, this model was derived using the subset of the CENA dataset discussed in Section 10.3.1.1.2 and consists of CENA tectonic data with magnitude greater than or equal to 3.0 and R_{RUP} distance less than or equal to 300 km. The model has the following form:

$$\phi_{SS} = \begin{cases} a & \text{for } \mathbf{M} \le 5.0 \\ a + (\mathbf{M} - 5.0) * \dfrac{(b - a)}{1.5} & \text{for } 5.0 < \mathbf{M} \le 6.5 \\ b & \text{for } \mathbf{M} > 6.5 \end{cases} \qquad (10\text{–}14)$$

where the ratios of b/a were constrained to those of the global ϕ_{SS} model; therefore, CENA data were used to solve for coefficient ‘a’. Figure 10–31 shows coefficients ‘a’ and ‘b’ as a function of

frequency. The average of ‘a’ values in the frequency range of 1 to 10 Hz was used as a constant value of ‘a’, which was then applied across all frequencies.

Similar to the CENA constant ϕ_{SS} model, the uncertainty in the CENA magnitude-dependent ϕ_{SS} model has two components: station-to-station variability in ϕ_{SS} and statistical uncertainty in the ϕ_{SS} estimates. Figure 10–32 shows the statistical uncertainty and the site-to-site variability of ϕ_{SS}^2 at **M** = 5.0 (a^2). The values of the statistical uncertainty were again smoothed using the average between frequencies of 1 and 10 Hz. The total SD(a^2) is also shown on the plot. Similarly, Figure 10–33 shows the components of the total variability in ϕ_{SS}^2 at **M** = 6.5 (b^2).

10.3.4 PhiS2S (ϕ_{S2S}) Model

Sensitivity analyses were performed to evaluate the impact of different parameters—such as number of recordings per station, V_{S30} ranges, assigned V_{S30} codes, measured versus inferred V_{S30}, tectonic events versus PIEs, regional dependence, etc.—on the estimated site terms and ϕ_{S2S} values for CENA. Observations from these sensitivity analyses, documented in Al Atik (2015) (Sections 4.5, 4.6, and 5.4) and summarized in this section, guided the development of the ϕ_{S2S} model for CENA. Recently, additional analyses on ϕ_{S2S} were performed to investigate issues such as the dependence of ϕ_{S2S} on magnitude and the extrapolation of ϕ_{S2S} to hard-rock conditions. These analyses were important to the update of the ϕ_{S2S} model developed for CENA and are documented in this section.

The CENA ground-motion dataset described in Section 10.3.1.1 was used to estimate site terms and evaluate ϕ_{S2S}. Appendix F.1 presents the results of the evaluation of regional differences in the average site terms and ϕ_{S2S} values for CENA. The average site term for Region 3 is biased negative between frequencies of 1 and 10 Hz, while the average of the event terms in Region 3 is near zero. The negative bias in the site terms for Region 3 can be attributed to site-response effects as well as potential regional differences in the Q attenuation for Region 3. Given the limited data available in Region 3, regional Q differences and their potential impact on the site terms cannot be reliably quantified.

Figure 10–34 compares ϕ_{S2S} for PIEs versus tectonic events and indicates that ϕ_{S2S} values for PIEs are smaller than those for tectonic events. The difference in ϕ_{S2S} between PIEs and tectonic events was investigated and documented in Al Atik (2015). Figure 10–35 shows the spatial distribution of the stations used in the analysis of the PIEs and tectonic events at f = 4 Hz. Figure 10–35 indicates that the stations that recorded PIEs are located in a much narrower geographic region than those that recorded tectonic events. Based on Figure 10–35, the TI team concluded that the reduced ϕ_{S2S} values for PIEs compared to tectonic events could be due to the clustering of the stations that recorded PIEs and the presence of more similar geologic conditions than for the stations that recorded tectonic events spanning a much larger area.

The sensitivity of using PIEs and tectonic events versus only tectonic events on the resulting ϕ_{S2S} values was further explored. Two single-station sigma analyses were performed using data from tectonic events only and data from both tectonic events and PIEs, respectively. An F-test of equality of variance was conducted to test the null hypothesis of equal variance of site terms using both PIEs and tectonic events versus using tectonic-only events (results of two different

mixed-effects regression analyses): the null hypothesis (H_0), the alternative hypothesis (H_a), and the test statistic F are written as:

$$H_0 : \phi^2_{S2S}\left(\text{PIE} + \text{Tectonic}\right) = \phi^2_{S2S}\left(\text{Tectonic}\right) \quad (10\text{–}15)$$

$$H_a : \phi^2_{S2S}\left(\text{PIE} + \text{Tectonic}\right) \neq \phi^2_{S2S}\left(\text{Tectonic}\right) \quad (10\text{–}16)$$

$$F = \phi^2_{S2S}\left(\text{PIE} + \text{Tectonic}\right) \Big/ \phi^2_{S2S}\left(\text{Tectonic}\right) \quad (10\text{–}17)$$

Under the null hypothesis, F has an F-distribution with numerator degrees-of-freedom of N1-1 and denominator degrees-of-freedom of N2-1, where N1 and N2 are the sample sizes of the two datasets. Table 10–7 shows the p-values of the F-test at frequencies of 1 to 10 Hz. Values in red indicate cases where the p-values are less than 0.05, and the null hypothesis is rejected at a 5% significance level. Table 10–7 shows that for 5% significance level, the equality of ϕ^2_{S2S} values obtained using all CENA events versus tectonic events only cannot be rejected for most frequencies between 1 to 10 Hz.

The impact of the limited number of stations in the CENA dataset with measured V_{S30} and the errors in the assigned V_{S30} values on the resulting ϕ_{S2S} values was evaluated. Figure 10–36 presents a comparison of ϕ_{S2S} calculated using stations with measured versus inferred V_{S30} and indicates that ϕ_{S2S} values are generally comparable for stations with measured versus inferred V_{S30}.

Based on the observations presented above, values of ϕ_{S2S} obtained from the mixed-effects regression analysis of the CENA within-event residuals discussed in Section 10.3.1.1 consisting of PIEs and tectonic events, excluding data from the GCR, and having a minimum of three recordings per station were used to develop a CENA ϕ_{S2S} model at PGV and for frequencies between 1 and 10 Hz. Both tectonic events and PIEs were used in order to maximize the number of stations in the analysis (144 versus 275 stations for the tectonic events versus both tectonic events and PIEs at f = 4 Hz, respectively). We note that, comparing the ϕ_{S2S} values for the tectonic events versus all events in the frequency range of 1 to 10 Hz in Figure 10–34, the largest difference in ϕ_{S2S} is on the order of 16%.

Figure 10–37 compares the CENA ϕ_{S2S} values to those for the NGA-West2 GMMs as well as for Japanese data (Dawood 2014). These comparisons indicate that CENA ϕ_{S2S} values are comparable to the Japanese ϕ_{S2S} both in terms of amplitude as well as general spectral shape. Al Atik (2015) described additional analyses performed to investigate the impact of other factors such as regression approach and regional trade-offs between τ and ϕ_{S2S} on the ϕ_{S2S} results, and concluded that these factors are unlikely to have controlled the CENA ϕ_{S2S} results. The similarity of ϕ_{S2S} for CENA and Japan could be result of general similarities in the site conditions between the two regions, which consists of relatively shallow soil cover over hard rock. Relatively large ϕ_{S2S} values can be observed for both CENA and Japan in contrast with smaller ϕ_{S2S} for NGA-West2.

Based on the discussion presented in this section, the TI team decided to develop a region-independent ϕ_{S2S} model derived from all the available CENA data. Models of ϕ_{S2S} from regions

other than CENA (WUS or Japan) were not directly adopted because the variability in the site terms is not generally constant across regions. Based on the similarity of ϕ_{S2S} for CENA and Japan, the scaling of the Japanese ϕ_{S2S} with frequency was used to extrapolate the CENA ϕ_{S2S} values for frequencies outside of 1 to 10 Hz. Figure 10–38 shows the CENA ϕ_{S2S} values versus frequency between 1 and 10 Hz, and the extrapolated values outside of this frequency range. Figure 10–39 shows the variability of ϕ_{S2S}^2, SD(ϕ_{S2S}^2), obtained from the regression analysis for CENA and compared to the corresponding values for Japan. Similar to the mean ϕ_{S2S} values, the CENA SD(ϕ_{S2S}^2) were used at PGV and for frequencies between 1 and 10 Hz. Outside of this frequency range, the SD(ϕ_{S2S}^2) values were extrapolated based on the scaling of SD(ϕ_{S2S}^2) with frequency for Japan.

10.3.4.1 Magnitude Dependence of PhiS2S (ϕ_{S2S})

Because the NGA-East dataset is mostly composed of small magnitude earthquakes, the TI team performed further analyses to investigate the potential bias of ϕ_{S2S} due to the limited small magnitude range of the dataset. Ground-motion recordings used to evaluate ϕ_{S2S} for CENA range in magnitude between 2.57 and 5.8, and cannot give insight on the magnitude-dependence of ϕ_{S2S}; therefore, other datasets that cover a wide magnitude range, such as the NGA-West2 and the Japanese datasets, were used for this purpose. Figures 10–40 to 10–43 show ϕ_{S2S} calculated using all recordings versus subsets of the recordings with magnitude smaller than or greater than 5.0 for NGA-West2 ASK14, BSSA14, CB14, and CY14, respectively. For each NGA-West2 model, the ϕ_{S2S} values presented in Figures 10–40 to 10–43 were obtained from three different mixed-effects regression analyses using the three different magnitude ranges of the corresponding dataset. Figures 10–40 to 10–43 show a significant difference in ϕ_{S2S} values obtained from the small versus large magnitude subsets of the NGA-West2 data, indicating a clear bias to large ϕ_{S2S} values as a result of limiting the dataset to small magnitudes. We note that for CB14 (Figure 10–42), a bigger reduction in ϕ_{S2S} is observed for the large magnitude subset of the data compared to the other 3 NGA-West2 models, which is likely due to the limited distance range to a maximum of 80 km in the CB14 dataset.

The TI team performed a similar analysis using the Japanese within-event residuals. Figure 10–44 presents the ϕ_{S2S} values obtained from the subsets of the Japanese dataset of surface recordings with **M** 4.0 to 5.0 and **M** >=5.0. A reduction in ϕ_{S2S} is also observed for the large magnitude subset of the data compared to the smaller magnitude dataset. This reduction is not as pronounced for the Japanese dataset as for NGA-West and is likely due to the absence of magnitudes less than of 4.0 in the Japanese dataset. Figure 10–45 presents a similar comparison of the magnitude dependence of ϕ_{S2S} using the Japanese borehole within-event residuals. Similar to the surface ground-motion data, ϕ_{S2S} for small magnitude borehole data is larger than that for large magnitude data. The difference in ϕ_{S2S} between small and large magnitude subsets of the data appears to be slightly less pronounced for the borehole data compared to the surface data.

Stafford et al. (2017) examined the linear site response factors in relation to the magnitude and distance of the scenario earthquake. While it is commonly assumed that linear site response is independent of the earthquake scenario, Stafford et al. (2017) found that for short spectral

periods, the linear site response factors were dependent on the magnitude and distance of the earthquake scenario. This impact was most pronounced for small magnitude scenarios. The authors explain this dependence as the result of the frequency bandwidth of the Fourier amplitude spectrum that contributes to the response spectrum at short spectral periods. The dependence of ϕ_{S2S} on magnitude presented in Figures 10–40 to 10–45 could be related to the same observations made in Stafford et al. (2017). We note that Stafford et al. (2017) found that the dependence of the linear site response on the scenario earthquake diminishes for stiffer site conditions, which is consistent with the observation that the magnitude dependence of ϕ_{S2S} is less pronounced for the borehole Japanese data compared to the corresponding surface data.

Figure 10–46 presents a comparison of ϕ_{S2S} obtained from small magnitude subsets of the NGA-West2 and the Japanese surface datasets compared to ϕ_{S2S} for CENA. While ϕ_{S2S} for CENA and Japan are different than those for NGA-West2 in terms of both amplitude and frequency content (refer to Figure 10–37), Figure 10–46 shows more similarities between ϕ_{S2S} for CENA and NGA-West2 for the small-magnitude subsets of the data. Based on the observations made in this section, the TI team concluded that CENA ϕ_{S2S} is likely to be biased to large values as a result of the small magnitude CENA dataset (**M** = 2.57 to 5.8) used in the analysis. Correction of the CENA ϕ_{S2S} values based on factors derived from datasets that cover wide magnitude ranges was considered necessary.

10.3.4.2 PhiS2S (ϕ_{S2S}) for Hard-Rock Conditions

In order to develop a CENA ϕ_{S2S} model applicable to hard-rock conditions with V_{S30} of 3000 m/sec, the NGA-East TI team conducted analyses to examine the dependence of ϕ_{S2S} on different site conditions. We note that these analyses were limited by the small number of recording stations on hard rock in available datasets (e.g., CENA and NGA-West2 datasets). As a result, the TI team made use of comparisons of ϕ_{S2S} between surface and borehole recordings from Japan to guide the correction of CENA ϕ_{S2S} for application to hard rock conditions.

Figure 10–47 presents a comparison of the CENA ϕ_{S2S} (median, 5th, and 95th percentiles) extrapolated, as explained earlier on in this section, outside of the 1–10 Hz frequency range based on the Japanese surface ϕ_{S2S} values compared to CENA ϕ_{S2S} for stations with V_{S30} greater than or equal to 1500 m/sec. We note that, out of a total of 275 stations available for the CENA analysis at f = 4 Hz, there are only 42 stations with V_{S30} greater than or equal to 1500 m/sec. Moreover, these stations are primarily located in Canada with mostly inferred V_{S30} values and are likely not representative of ϕ_{S2S} for rock sites across the entire CENA. Figure 10–47 indicates that ϕ_{S2S} for rock sites in CENA is generally smaller than that observed for the entire dataset comprised mostly of soil sites. Moreover, the peak ϕ_{S2S} for the rock sites occurs around 15 Hz compared to a peak of around 10 Hz for the entire dataset. This is consistent with the shift in frequency content observed for rock sites with smaller kappa values.

A similar analysis was attempted to compare ϕ_{S2S} obtained from stations with V_{S30} greater than 1000 m/sec to those obtained from all stations for NGA-West2 ASk14, BSSA14, CB14, and CY14. Despite the limited subset of stations with V_{S30} greater than or equal to 1000 m/sec (e.g., 14 stations out of the total of 1227 stations for ASK14), this analysis indicates smaller ϕ_{S2S} for rock conditions compared to the softer sites as shown in Figure 10–48 for ASK14.

Figure 10–49 presents a comparison of ϕ_{S2S} from the surface versus borehole recordings of the Japanese dataset (Dawood 2014). We note that the surface recordings have an average V_{S30} of 499 m/sec while the borehole recordings have an average $V_{S,Zhole}$ of 1659 m/sec. Figure 10–49 indicates a significant difference in ϕ_{S2S} in terms of both amplitude and frequency content between the surface and borehole data. The borehole data recorded largely on rock (refer to Figure 10–12) show smaller ϕ_{S2S} values compared to their counterpart recorded on soil sites at the surface. In order to remove the softer sites from the Japanese borehole dataset, the analysis was repeated for the subset of the Japanese borehole stations with $V_{S,Zhole} \geq 1500$ m/sec. We note that $V_{S,Zhole}$ is not expected to be significantly smaller than V_{S30} below the borehole depth horizon due to generally large V_S values encountered at the borehole level. Figure 10–50 compares the ϕ_{S2S} for the Japanese borehole data with $V_{S,Zhole} \geq 1500$ m/sec to corresponding ϕ_{S2S} for small versus large magnitude subsets of the data. For rock site conditions, Figure 10–50 indicates that ϕ_{S2S} for **M** ≥ 5.0 is smaller than that calculated from the smaller magnitude subset of the data, which is consistent with observations presented in Section 10.3.4.1.

To correct the CENA ϕ_{S2S} values for application to hard-rock site conditions with V_{S30} of 3000 m/sec, the TI team adopted a set of correction factors obtained as the ratio of ϕ_{S2S} from the Japanese borehole dataset with $V_{S,Zhole} \geq 1500$ m/sec and **M** ≥ 5.0 to ϕ_{S2S} obtained from the Japanese surface data. These correction factors were multiplied by the extrapolated CENA ϕ_{S2S} values shown in Figure 10–38 to simultaneously correct for the site conditions and the small magnitude effects in the CENA dataset. Figure 10–51 presents the derived ϕ_{S2S} correction factors. The TI team acknowledges the limitation of using the Japanese borehole dataset in that the site response in the borehole data is impacted by the soil column above the depth of the recording instrument. Despite this limitation, Japanese borehole data were used to derive ϕ_{S2S} correction factors for CENA due to lack of surface recordings on hard rock sites in other available datasets and because of similarities in the spectral shape of ϕ_{S2S} for the rock sites in CENA to those for the Japanese borehole dataset (both peak around 15–20 Hz).

Figure 10–52 presents the mean CENA ϕ_{S2S} values corrected for the small magnitude bias and for application to hard rock conditions. These corrected CENA ϕ_{S2S} values were smoothed versus frequency while maintaining the spectral shape of the Japanese ϕ_{S2S} values for the borehole data with $V_{S,Zhole} \geq 1500$ m/sec and **M** ≥ 5.0 shown in Figure 10–50. Figure 10–53 presents the smoothed CENA mean ϕ_{S2S} model. The standard deviation of the CENA ϕ_{S2S} model consists of the standard deviations of the CENA ϕ_{S2S} values and those of the Japanese borehole ϕ_{S2S} values for the subset of data with $V_{S,Zhole} \geq 1500$ m/sec and **M** ≥ 5.0 combined using the square root of the sum of squares (SRSS). Figure 10–54 shows the standard deviation of the CENA ϕ_{S2S} model and indicates that the standard deviation of the Japanese borehole ϕ_{S2S} values is small compared to that of the CENA ϕ_{S2S} values. The total standard deviation of the CENA ϕ_{S2S} model was smoothed versus frequency to preserve its general spectral shape.

The value of ϕ_{S2S} for PGV is not available for the Japanese dataset. To correct the CENA ϕ_{S2S} at PGV for hard rock conditions and small magnitude effects, the correlation of PGV with other spectral periods was evaluated using the Japanese ground-motion data from the surface and

borehole recordings. This analysis indicated that, for the surface recordings, PGV is best correlated with the spectral accelerations at the frequency of 2 Hz (period of 0.5 sec). For the borehole data, PGV is best correlated with frequencies of 0.5 to 1 Hz (periods of 1 to 2 sec). As a result, the average ϕ_{S2S} ratio for frequencies between 0.5 and 2 Hz was used to correct the CENA ϕ_{S2S} at PGV.

10.4 Summary

This chapter presented the TI team's approach for the development and evaluation of aleatory variability models for CENA. Accordingly, existing and newly-developed models for the components of the ground-motion variability (τ, ϕ_{SS}, and ϕ_{S2S}) were presented and discussed. The evaluation of these candidate models is the subject of Chapter 11.

10.5 References

Abrahamson, N.A., and R.R. Youngs. 1992. "A Stable Algorithm for Regression Analyses using the Random Effects Model." Bull. Seismol. Soc. Am., 82(1): 505–510.

Abrahamson, N A., W.J. Silva, and R. Kamai. 2014. "Summary of the ASK14 Ground-Motion Relation for Active Crustal Regions." Earthq. Spectra, 30(3): 1025–1055.

Akkar, S., M.A. Sandikkaya, and J.J. Bommer. 2014. "Empirical Ground-Motion Models for Point- and Extended-Source Crustal Earthquake Scenarios in Europe and the Middle East." Bull. Seismol. Soc. Am., 12(1): 359–387, DOI 10.1007/s10518-013-9461-4.

Al Atik, L. 2015. "NGA-East: Ground Motion Standard Deviation Models for Central and Eastern North America." PEER Report No. PEER 2015/07, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Al Atik, L., N.A. Abrahamson, J.J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn. 2010. "The Variability of Ground-Motion Prediction Models and Its Components." Seismol. Res. Lett., 81(5): 794−801.

Ancheta, T.D., R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S.-J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida, and J.L. Donahue. 2014. "NGA-West2 Database." Earthq. Spectra, 30: 989–1005.

Anderson, J.G., and J.N. Brune. 1999. Probabilistic Seismic Hazard Assessment without the Ergodic Assumption." Seismol. Res. Lett., 70: 19−28.

Atkinson, G.M. 2006. "Single-Station Sigma." Bull. Seismol. Soc. Am., 96(2): 446–455.

Atkinson, G.M. 2008. "Ground Motion Prediction for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty." Bull. Seismol. Soc. Am., 98(3): 1304–1318.

Atkinson, G.M. 2013. "Empirical Evaluation of Aleatory and Epistemic Uncertainty in Eastern Ground Motions." Seismol. Res. Lett., 84: 130−138.

Atkinson, G.M., and J. Adams. 2013. Ground Motion Prediction Equations for Application To the 2015 National Seismic Hazard Maps of Canada." Canadian J. Civil Eng., 40: 988–998.

Atkinson, G.M., and D.M. Boore. 2006. "Earthquake Ground-Motion Prediction Equations for Eastern North America." Bull. Seismol. Soc. Am., 96: 2181–2205.

Atkinson, G.M., and D.M. Boore. 2011. "Modifications to Existing Ground-Motion Prediction Equations in Light of New Data." Bull. Seismol. Soc. Am., 101(3): 1121–1135.

Boore D.M. 2005. "SMSIM---Fortran Programs for Simulating Ground Motions From Earthquakes: Version 2.3---A Revision of OFR 96-80-A." U.S. Geological Survey Open-File Report, U. S. Geological Survey Open-File Report 00-509, revised 15 August 2005, 55 pgs. Available from the online publications link on http://quake.usgs.gov/~boore.

Boore, D.M., J.P. Stewart, E. Seyhan, and G.M. Atkinson. 2014. "NGA-West2 Equations for Predicting Response Spectral Accelerations for Shallow Crustal Earthquakes." Earthq. Spectra, 30(3): 1057–1085.

Campbell, K.W., and Y. Bozorgnia. 2014. "NGA-West2 Campbell-Bozorgnia Ground Motion Model for the Horizontal Components Of PGA, PGV, and 5%-Damped Elastic Pseudo-Acceleration Response Spectra for Periods Ranging from 0.01 to 10 sec." Earthq. Spectra, 30(3): 1087–1115.

Chen, L., and E. Faccioli. 2013. "Single-Station Standard Deviation Analysis of 2010–2012 Strong-Motion Data from the Canterbury Region." NZ Bull. Earthq. Eng., 11: 1617−1632.

Chen, Y-H., and C-C.P. Tsai. 2002. "A New Method for Estimation of the Attenuation Relationship with Variance Components." Bull. Seismol. Soc. Am., 92(5): 1984−1991.

Chiou, B.S.-J., and R.R. Youngs. 2014. "Update of the Chiou and Youngs NGA Ground Motion Model for Average Horizontal Component of Peak Ground Motion and Response Spectra." Earthq. Spectra. 30(3): 1117–1153.

Coppersmith, K.J., J.J. Bommer,, K. Hanson,, R. Coppersmith,, J.R. Unruh,, L. Wolf,, R.R. Youngs, L. Al Atik, A. Rodriguez-Marek, G. Toro, and V. Montaldo-Falero. 2014. "Hanford Sitewide Probabilistic Seismic Hazard Analysis." PNNL-23361, Pacific Northwest National Laboratory, Richland, WA.

Dawood, H.M. 2014. Partitioning Uncertainty for Non-Ergodic Probabilistic Seismic Hazard Analyses. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA.

EPRI. 2006. "Program on Technology Innovation: Truncation of the Lognormal Distribution and Value of the Standard Deviation for Ground Motion Models in the Central and Eastern United States." Electric Power Research Institute, EPRI Report 1013105, Technical Update, Palo Alto, CA.

EPRI. 2013. “Ground-Motion Model (GMM) Review Project.” Electrical Power Research Institute, EPRI Report 3002000717, Palo Alto, CA.

GeoPentech. 2015. “Southwestern United States Ground Motion Characterization SSHAC Level 3 – Technical Report Rev.2.” March 2015.

Goulet, C.A., T. Kishida. T.D. Ancheta, C.H. Cramer, R.B. Darragh, W.J. Silva, Y.M.A. Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. 2014. “PEER NGA-East Database.” PEER Report No. 2014/17, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Idriss, I.M. 2014. “An NGA-West2 Empirical Model for Estimating the Horizontal Spectral Values Generated by Shallow Crustal Earthquakes.” Earthq. Spectra, 30(3): 1155–1177.

Kuehn, N., personal communication.

Kuehn, N., and N. Abrahamson. 2017. “The Effect of Uncertainty in Predictor Variables on the Estimation of Ground-Motion Prediction Equations.” Bull. Seismol. Soc. Am., 108(1): 358–370.

Lin, P.-S., B.S.-J. Chiou, N.A. Abrahamson, M. Walling,, C.-T. Lee, and C.-T. Cheng. 2011. “Repeatable Source, Site, and Path Effects on the Standard Deviation for Empirical Ground-Motion Prediction Models.” Bull. Seismol. Soc. Am., 101(5): 2281–2295.

Luzi, L., D. Bindi, R. Puglia, F. Pacor, and A. Oth. 2014. “Single-Station Sigma for Italian Strong-Motion Stations.” Bull. Seismol. Soc. Am., 104(1): 467–483.

Morikawa, N., T. Kanno, A. Narita, H. Fujiwara, T. Okumura, Y. Fukushima, and A. Guerpinar. 2008. “Strong Motion Uncertainty Determined from Observed Records by Dense Network in Japan.” J. Seismology, 12(4): 529–546.

PEER. 2015. “NGA-East: Median Ground-Motion Models for the Central and Eastern North Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Renault, P., S. Heuberger, and N.A. Abrahamson. 2010. “PEGASOS Refinement Project: An Improved PSHA for Swiss Nuclear Power Plants.” Proceedings, 14th European Conference of Earthquake Engineering, Paper ID 991, Ohrid, Republic of Macedonia.

Rodriguez-Marek, A., F. Cotton, N.A. Abrahamson, S. Akkar, L. Al Atik, B. Edwards, G.A. Montalva, and H.M. Dawood. 2013. “A Model for Single-Station Standard Deviation using Data from Various Tectonic Regions.” Bull. Seismol. Soc. Am., 103: 3149–3163.

Rodriguez-Marek, A., G.A. Montalva, F. Cotton, and F. Bonilla. 2011. “Analysis of Single-Station Standard Deviation using the KiK-net Data.” Bull. Seismol. Soc. Am., 101: 1242–1258.

Rodriguez-Marek, A., E.M. Rathje, J.J. Bommer,, F. Scherbaum, and P. Stafford. 2014. “Application of Single-Station Sigma and Site-Response Characterization in a Probabilistic Seismic-Hazard Analysis for a New Nuclear Site.” Bull. Seismol. Soc. Am., 104: 1601–1619.

Searle, S.R. 1971. Linear Models. New York: Wiley.

Stafford, P.J., A. Rodriguez-Marek, B. Edwards, P. Kruiver, and J.J. Bommer. 2017. “Scenario Dependence of Linear Site Effect Factors for Short-Period Response Spectral Ordinates.” Bull. Seismol. Soc. Am., 107: (6): 2859–2872.

Youngs, R.R., personal communication.

Table 10–1 Summary of terminology for residual components and standard deviations (SD denotes standard deviation).

Residual Component	Residual Notation	Standard Deviation Component	Standard Deviation Notation
Total residual	Δ_{es} $\Delta_{es} = \delta B_e + \delta W_{es}$	Total or ergodic standard deviation	σ $\sigma = \sqrt{\phi^2 + \tau^2}$
Between-event residual (event term)	δB_e	Between-event (inter-event) standard deviation (Tau)	$\tau = SD(\delta B_e)$
Within-event residual (intra-event residual; event-corrected residual)	δW_{es} $\delta W_{es} = \delta S2S_s + \delta WS_{es}$	Within-event (intra-event) standard deviation (Phi)	$\phi = SD(\delta W_{es})$
Site-to-site residual (site term)	$\delta S2S_s$	Site-to-site variability	$\phi_{S2S} = SD(\delta S2S_s)$
Single-station within-event residual (site- and event-corrected residual)	δWS_{es}	Single-station within-event standard deviation (single-station Phi)	$\phi_{ss} = SD(\delta WS_{es})$

Table 10–2 Number of CENA recordings, earthquakes, and stations at f = 4 Hz used in the within-event and between-event residuals analysis.

	Number of Recordings	Number of Earthquakes	Number of Stations
Tectonic	1122	53	345
PIE	720	9	181
Total	1,842	62	447

Table 10–3 Number of CENA recordings, earthquakes, and stations at f = 25 Hz used in the within-event and between-event residuals analysis.

	Number of Recordings	Number of Earthquakes	Number of Stations
Tectonic	473	44	157
PIE	50	7	17
Total	523	51	169

Table 10–4 Coefficients of the CENA GMM derived for the purpose of analyzing ground-motion variability. Note that this GMM is not intended for use in median ground-motion predictions. Coefficients outside of the 1 to 10 Hz frequency range are not considered reliable due to limited data.

Period (sec)	Frequency (Hz)	c_1	c_2	c_3	c_4 PIE	c_4 Tectonic	c_7 PIE	c_7 Tectonic	c_8	c_9	c_{4h}
0.03	33.33	-6.5508	2.7840	-0.1481	-1.3019	-0.9216	0.0025	-0.0035	-0.4039	0.4730	-1.6936
0.04	25	-5.2438	2.6658	-0.1365	-1.3831	-0.9704	0.0013	-0.0044	-0.4939	0.2070	-1.5017
0.05	20	-6.1741	2.9721	-0.1663	-1.2884	-0.9103	0.0001	-0.0048	-0.4947	0.2171	-1.3005
0.075	13.33	-8.8509	3.6489	-0.2250	-1.3273	-1.0858	-0.0047	-0.0061	-0.3135	0.4594	-0.9301
0.1	10	-9.7947	4.0749	-0.2643	-1.2283	-1.0585	-0.0055	-0.0064	-0.3594	0.3906	-0.7481
0.15	6.67	-11.2908	4.5891	-0.3040	-1.2024	-1.0980	-0.0052	-0.0057	-0.3940	0.3168	-0.6321
0.2	5	-11.7502	4.6742	-0.3019	-1.1113	-1.0433	-0.0047	-0.0049	-0.4759	0.3110	-0.6067
0.25	4	-12.6926	4.9079	-0.3191	-1.0804	-1.0248	-0.0041	-0.0043	-0.5025	0.3058	-0.5979
0.3	3.33	-13.6485	5.0042	-0.3217	-1.0760	-1.0312	-0.0033	-0.0037	-0.4663	0.2892	-0.5947
0.4	2.5	-14.4784	5.1658	-0.3283	-1.1811	-1.1335	-0.0023	-0.0031	-0.4535	0.2463	-0.5945
0.5	2	-14.4444	5.0034	-0.3000	-1.1901	-1.1385	-0.0018	-0.0027	-0.4930	0.1931	-0.5972
0.75	1.33	-15.3583	4.7925	-0.2620	-1.1749	-1.1624	-0.0012	-0.0022	-0.4429	0.1434	-0.6086
1	1	-15.2826	4.4269	-0.2163	-1.1828	-1.1652	-0.0010	-0.0019	-0.4297	0.0176	-0.6261
1.5	0.67	-15.8332	4.3013	-0.1887	-1.1132	-1.1122	-0.0010	-0.0014	-0.4966	-0.0821	-0.6622
2	0.5	-18.3595	5.0045	-0.2550	-1.1292	-1.1431	-0.0007	-0.0006	-0.4600	-0.1737	-0.6909
3	0.33	-18.3934	4.4972	-0.1942	-0.9834	-1.0214	-0.0012	-0.0003	-0.4542	-0.4234	-0.7306
4	0.25	-24.2103	6.4351	-0.3789	-0.8941	-1.0169	-0.0011	-0.0004	-0.4275	-0.4002	-0.7531
5	0.2	-25.5952	6.9103	-0.4250	-0.8302	-0.9370	-0.0008	-0.0004	-0.5040	-0.4257	-0.7677
7.5	0.13	-18.0242	3.2094	-0.0347	-1.1059	-1.0431	-0.0002	-0.0009	-0.3343	-0.7069	-0.7865
10	0.1	-18.2366	2.5760	0.0401	-1.0378	-1.1393	-0.0004	0.0002	-0.2478	-0.4496	-0.7961

Table 10–5 Number of CENA recordings, earthquakes and stations at f = 4 Hz used in the single-station analysis.

	Number of Recordings	Number of Earthquakes	Number of Stations
Tectonic	927	51	189
PIE	691	9	161
Total	1618	60	275

Table 10–6 Number of recordings, earthquakes, and stations in the NGA-West2 datasets used in the single-station analysis.

	Number of Recordings	Number of Earthquakes	Number of Stations
ASK14	13,020	297	1,227
BSSA14	15,466	377	1,344
CB14	5,285	244	535
CY14	9,197	269	798

Table 10–7 P-values from the F-test of equality of ϕ_{S2S}^2 using all CENA events and tectonic events only. Values in red show cases where the null hypothesis(1) is rejected at 5% significance level.

Period (sec)	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.75	1.00
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00
All events vs tectonic only	0.316	0.249	0.158	0.094	0.058	0.035	0.080	0.070	0.048

(1) Null and alternate hypotheses are: $H_0 : \phi_{S2S}^2(\text{PIE} + \text{Tectonic}) = \phi_{S2S}^2(\text{Tectonic})$ and $H_a : \phi_{S2S}^2(\text{PIE} + \text{Tectonic}) \neq \phi_{S2S}^2(\text{Tectonic})$, respectively.

Figure 10–1 Distribution of the posterior samples of Tau (τ) versus the median ground-motion predictions for a M = 6.0 and distance of 20-km earthquake scenario obtained using the Kuhen and Abrahamson (2017) model (Kuehn, personal communication).

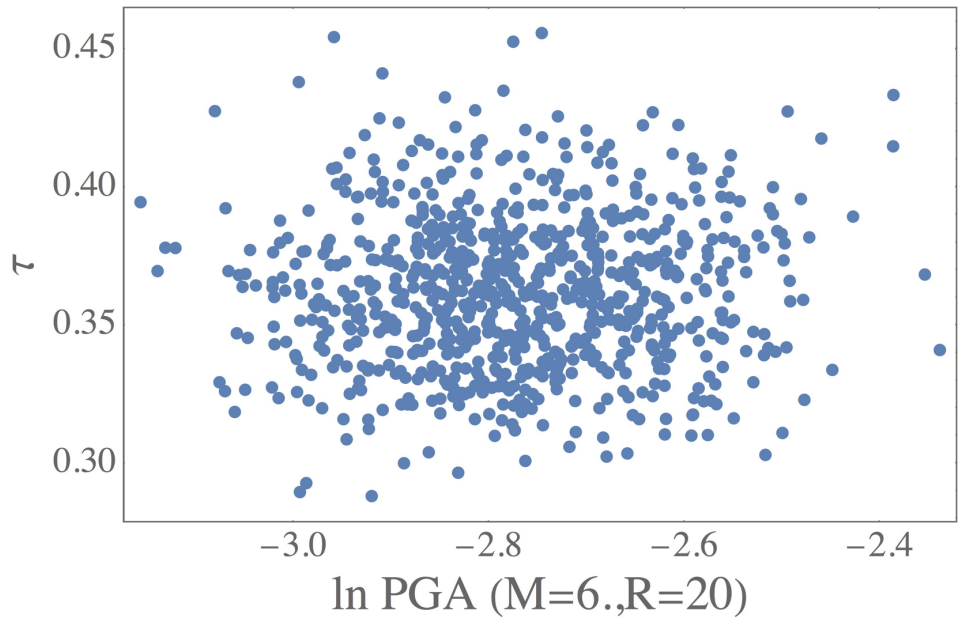


Figure 10–2 Distribution of the posterior samples of PhiSS (ϕ_{SS}) versus the median ground-motion predictions for a M = 6.0 and distance of 20-km earthquake scenario obtained using the Kuhen and Abrahamson (2017) model (Kuehn, personal communication).

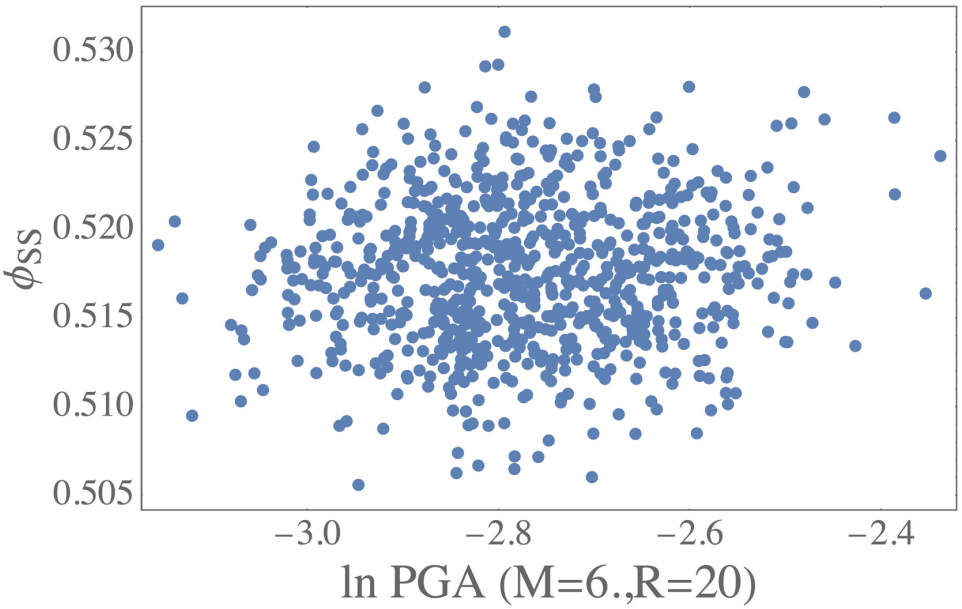


Figure 10–3 Distribution of the posterior samples of PhiS2S (ϕ_{S2S}) versus the median ground-motion predictions for a M = 6.0 and distance of 20-km earthquake scenario obtained using the Kuhen and Abrahamson (2017) model (Kuehn, personal communication).

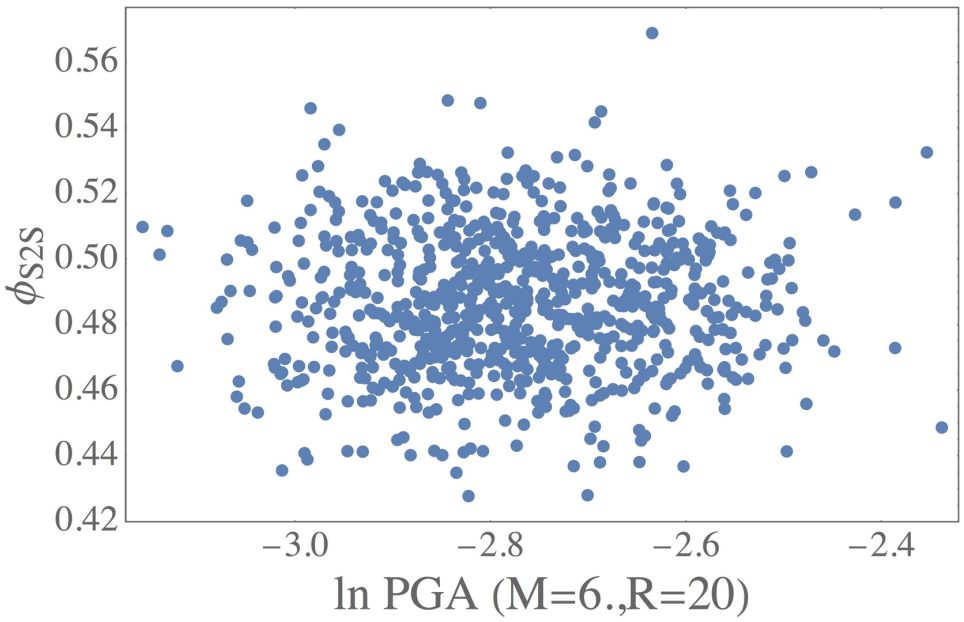


Figure 10–4 Magnitude and distance distribution of the CENA ground-motion data used in the between-event and within-event residuals analysis at f = 4 Hz.

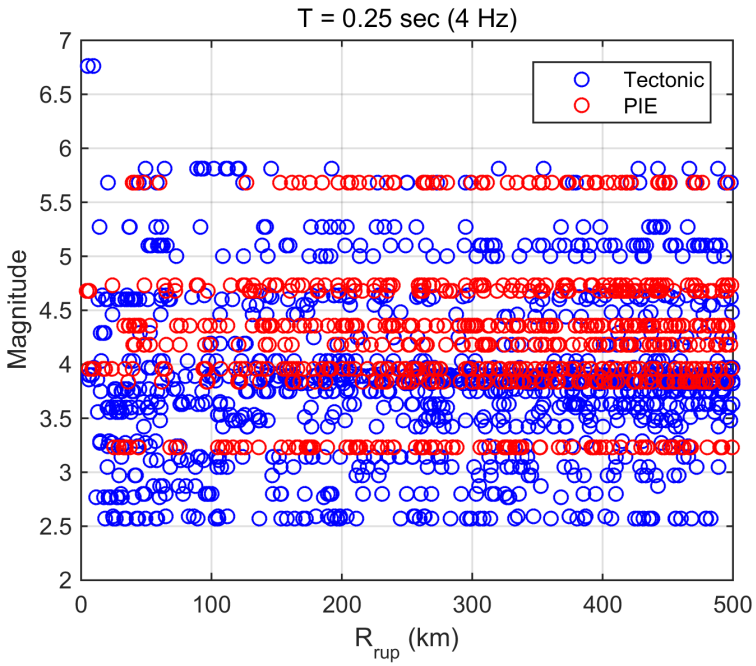


Figure 10–5 Magnitude and hypocentral depth distribution of the CENA earthquakes used in the between-event and within-event residuals analysis at f = 4 Hz.

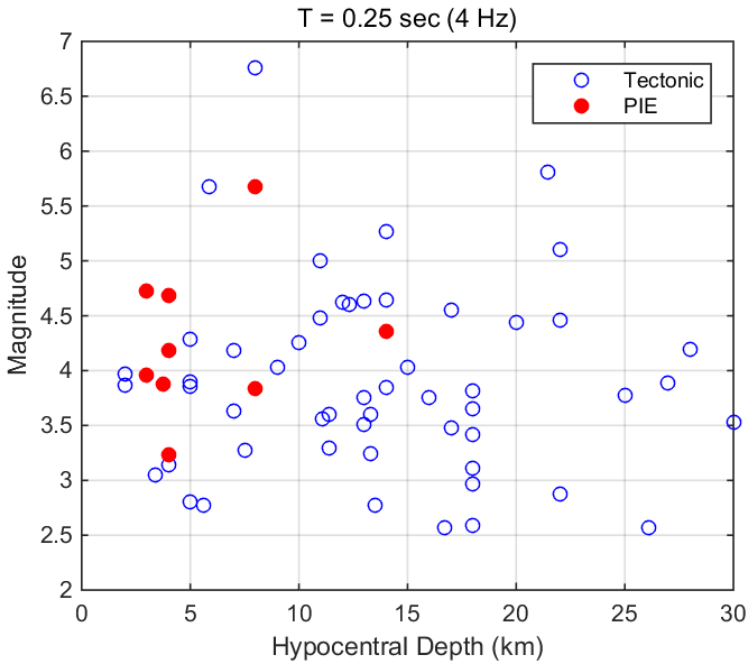


Figure 10–6 Histogram of the number of stations in the V_{S30} bins for the CENA ground-motion data used in the between-event and within-event residuals analysis at f = 4 Hz.

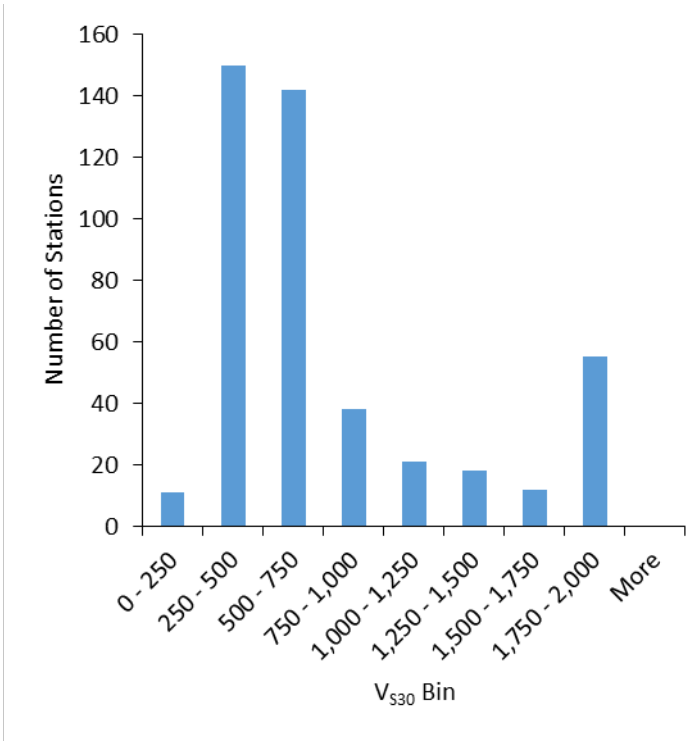


Figure 10–7 Number of recordings versus frequency for the CENA dataset used in the between-event and within-event residuals analysis.

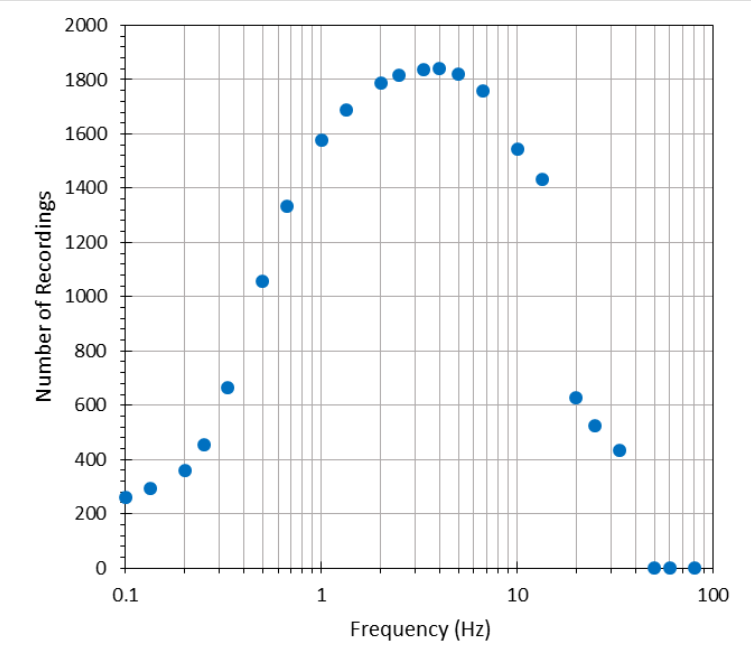


Figure 10–8 Comparison of the magnitude and distance distribution of the CENA tectonic earthquakes used in the between-event and within-event residuals analysis at f = 4 and 25 Hz.

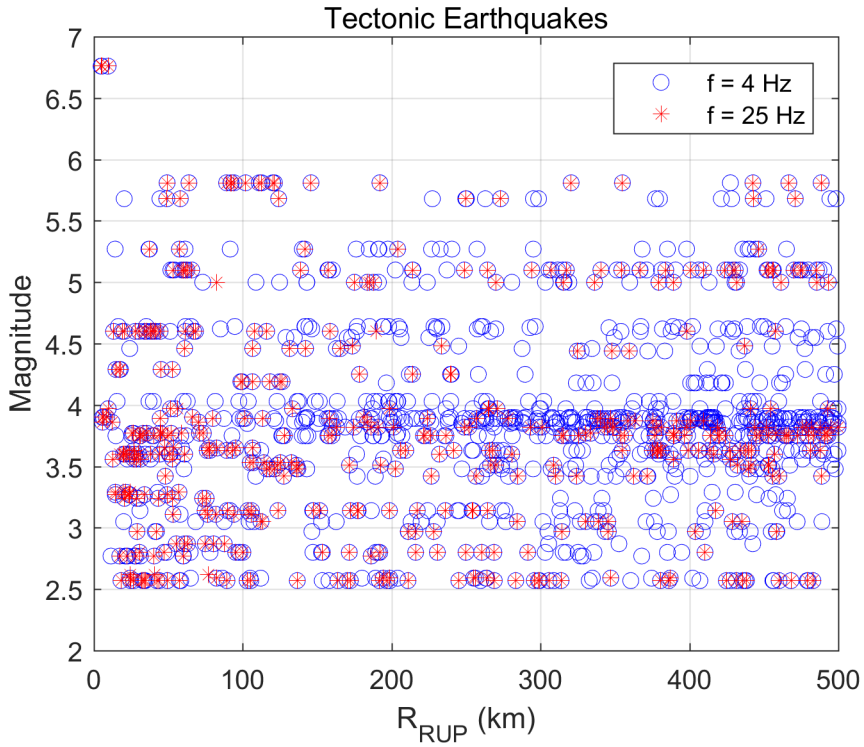


Figure 10–9 Magnitude and distance distribution of the CENA ground-motion data used in the single-station analysis at f = 4 Hz.

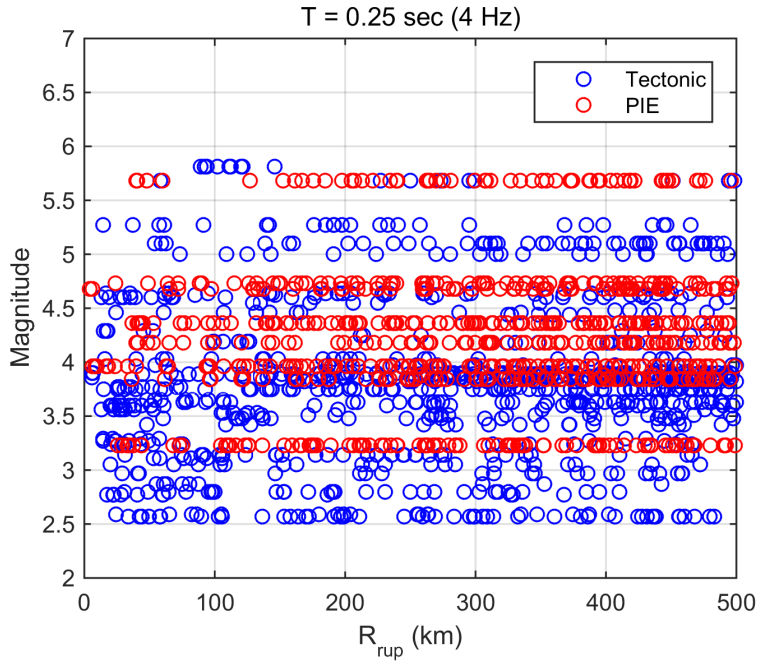


Figure 10–10 Number of recordings versus frequency for the CENA dataset used in the single-station analysis.

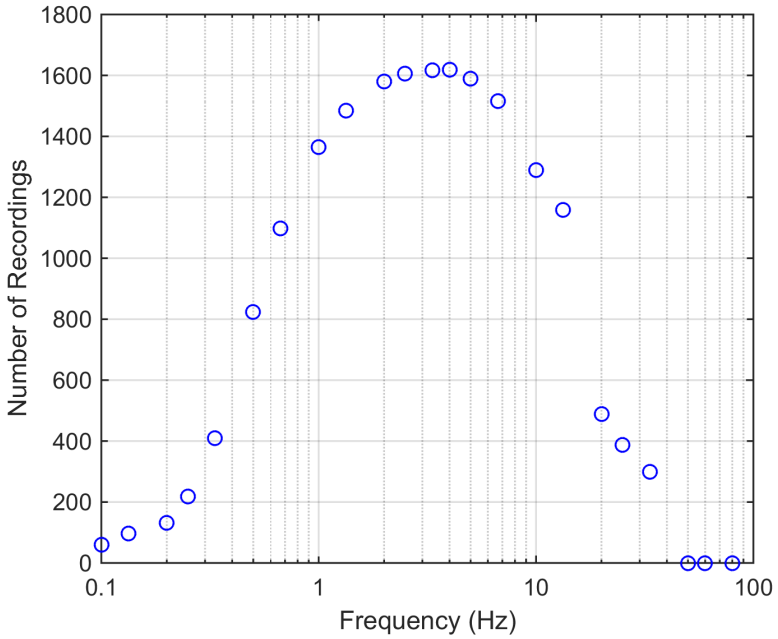


Figure 10–11 Histogram of the number of stations in the V_{S30} bins for the Japanese ground-motion dataset.

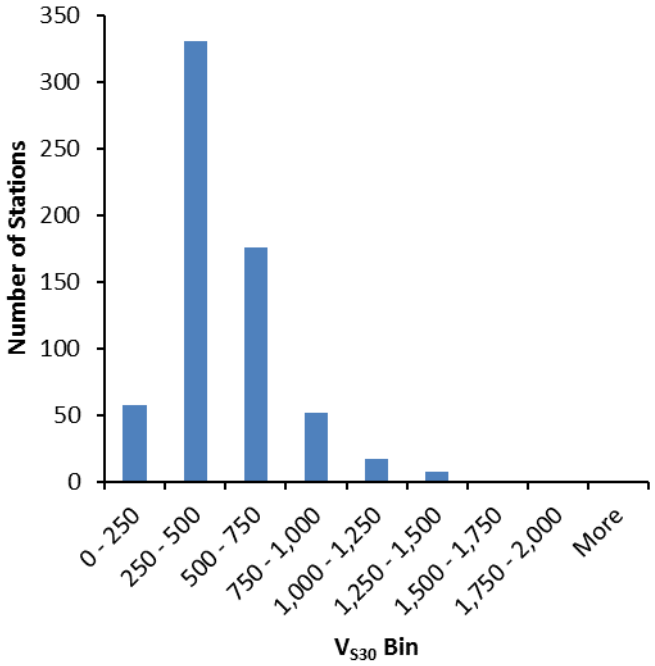


Figure 10–12 Histogram of the number of stations in the $V_{S,Zhole}$ bins for the Japanese ground-motion dataset.

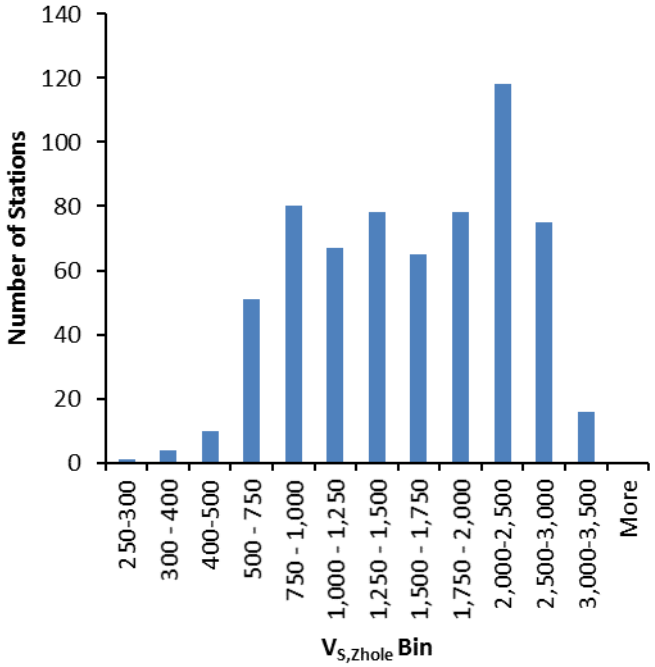


Figure 10–13 Histogram of the number of stations in the borehole depth bins for the Japanese ground-motion dataset.

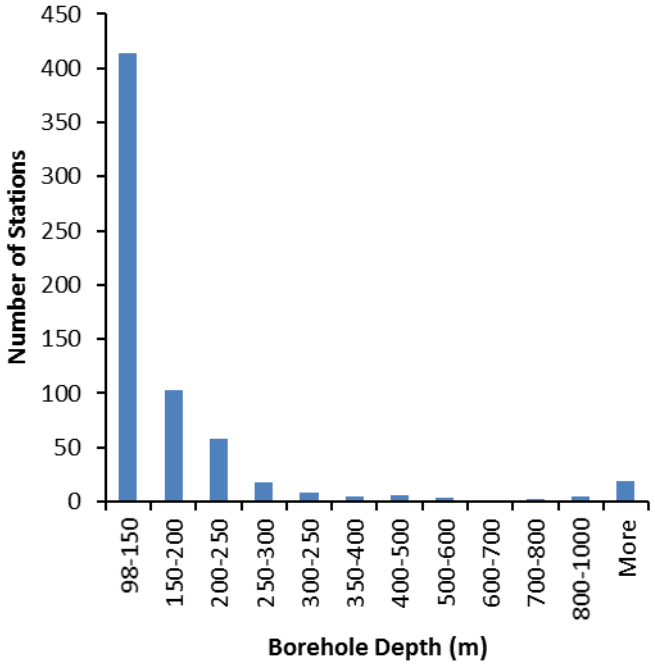


Figure 10–14 Global Tau (τ) versus magnitude at f = 1 and 100 Hz.

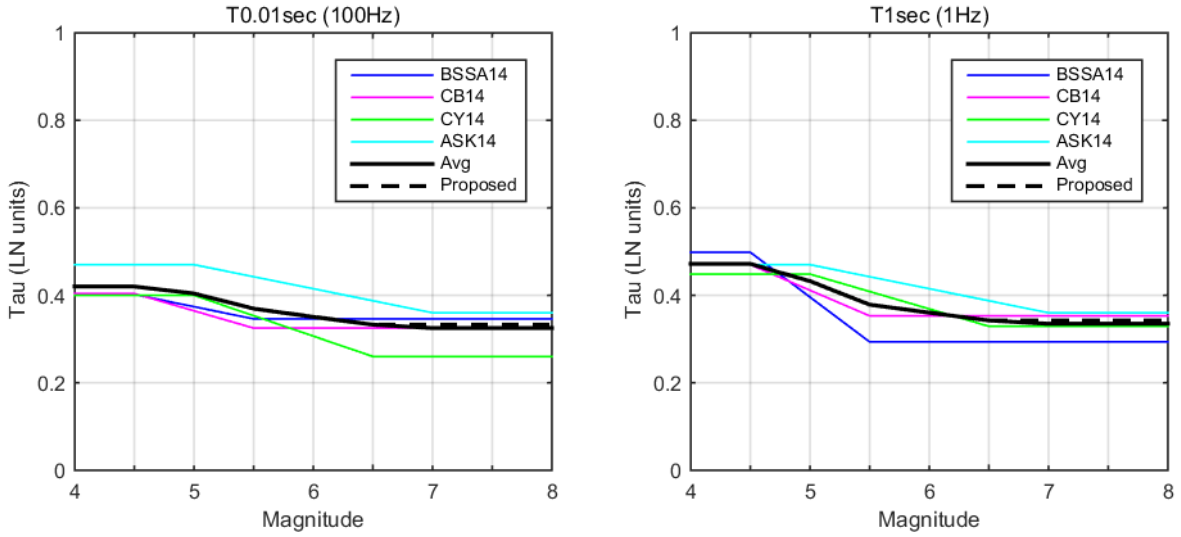


Figure 10–15 Coefficients of the global Tau (τ) model versus frequency as derived (solid lines) and smoothed (dashed lines).

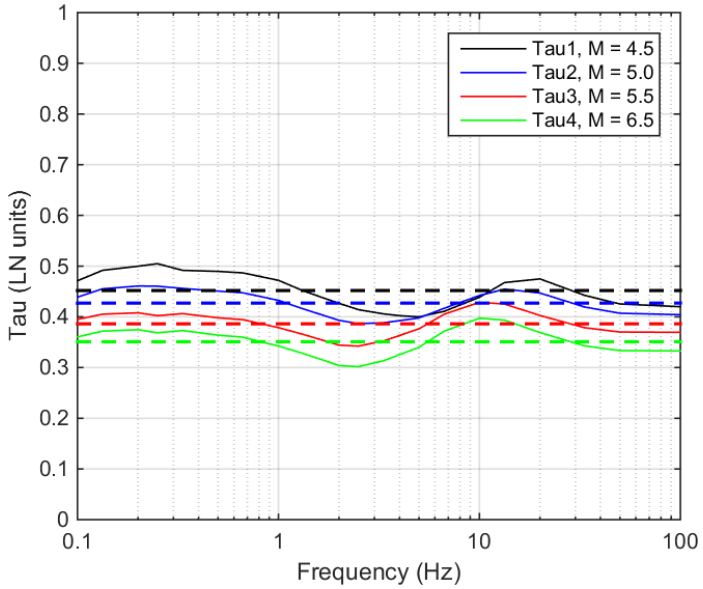


Figure 10–16 Standard deviations computed using point-source stochastic simulations with random kappa values (Coppersmith et al. 2014).

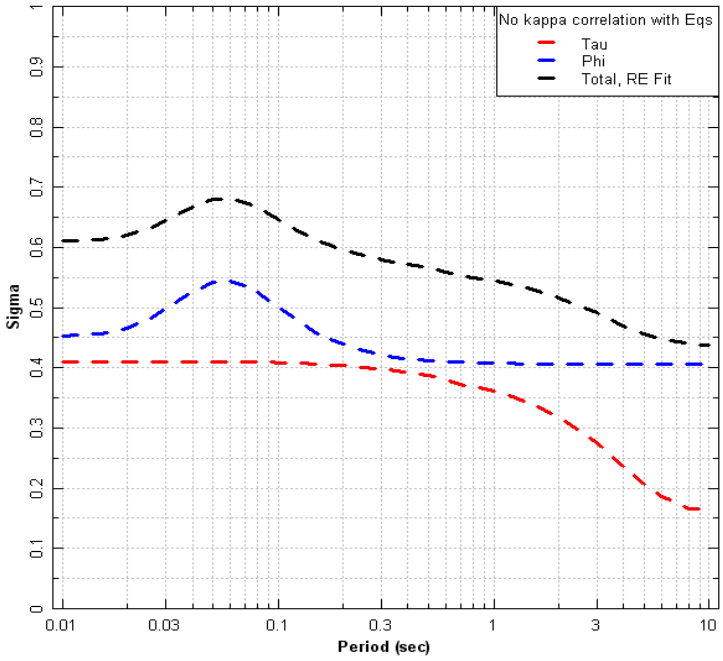


Figure 10–17 Standard deviations computed using point-source stochastic simulations with correlated kappa values (Coppersmith et al. 2014).

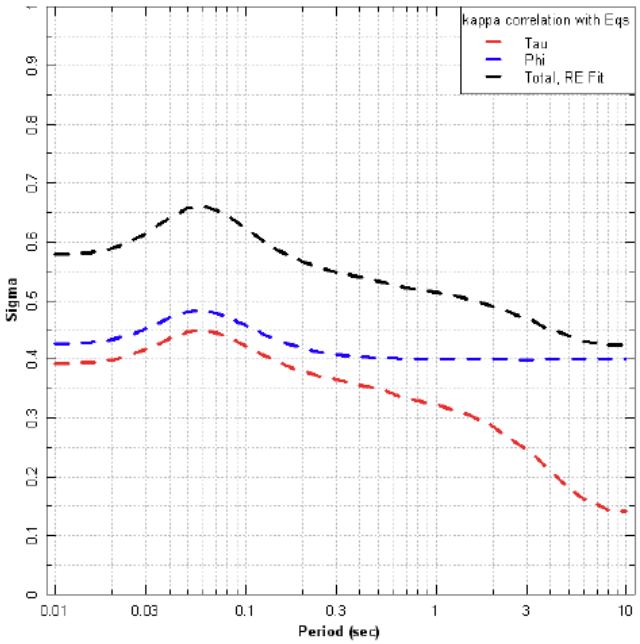


Figure 10–18 Between-model, within-model, and total variability of the global Tau^2 (τ^2) model at the magnitude breaks of M = 4.5, 5.0, 5.5, and 6.5.

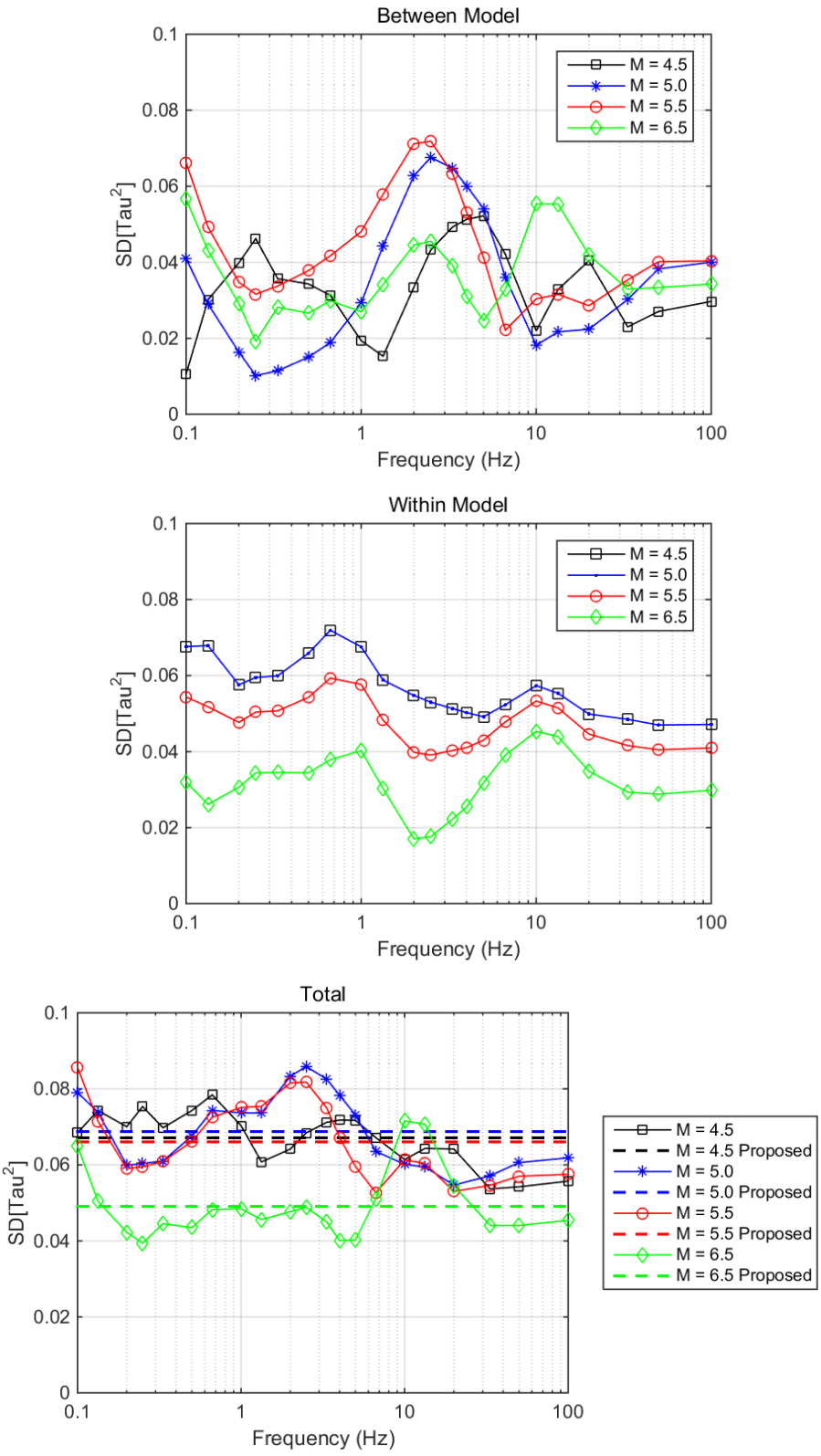


Figure 10–19 CENA constant Tau (τ) model (best estimate). Error bars represent one standard error.

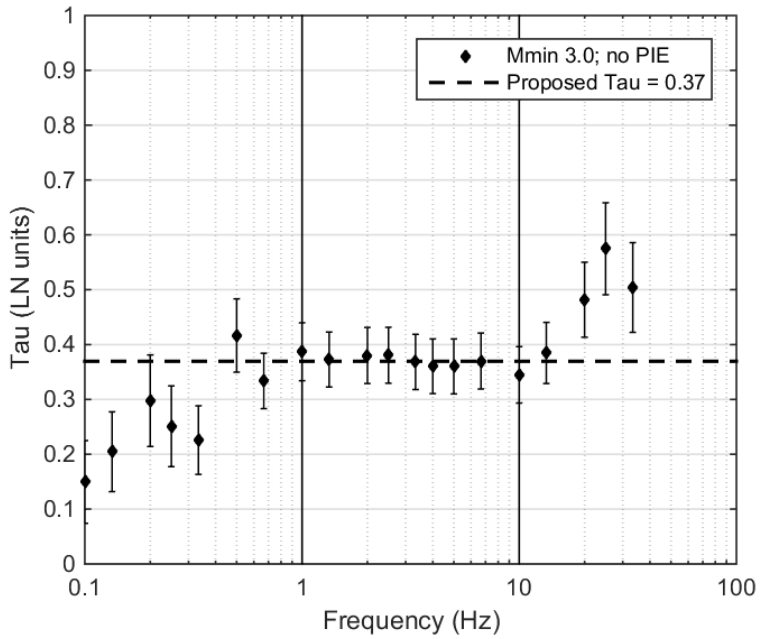


Figure 10–20 Variability in CENA constant Tau (τ) model.

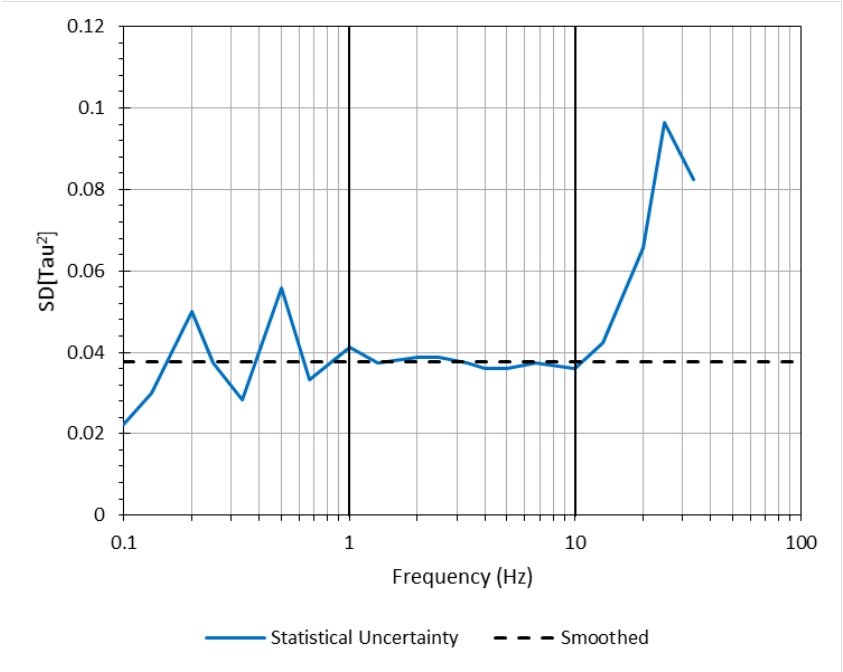


Figure 10–21 Coefficients of the CENA magnitude-dependent Tau (τ) model. Error bars represent one standard error.

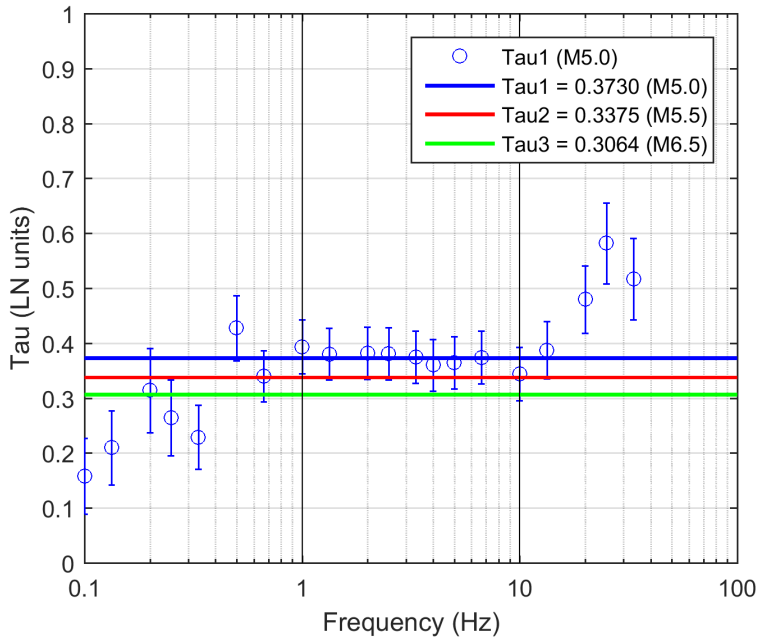


Figure 10–22 Variability in the CENA magnitude-dependent Tau (τ) model compared to the global model.

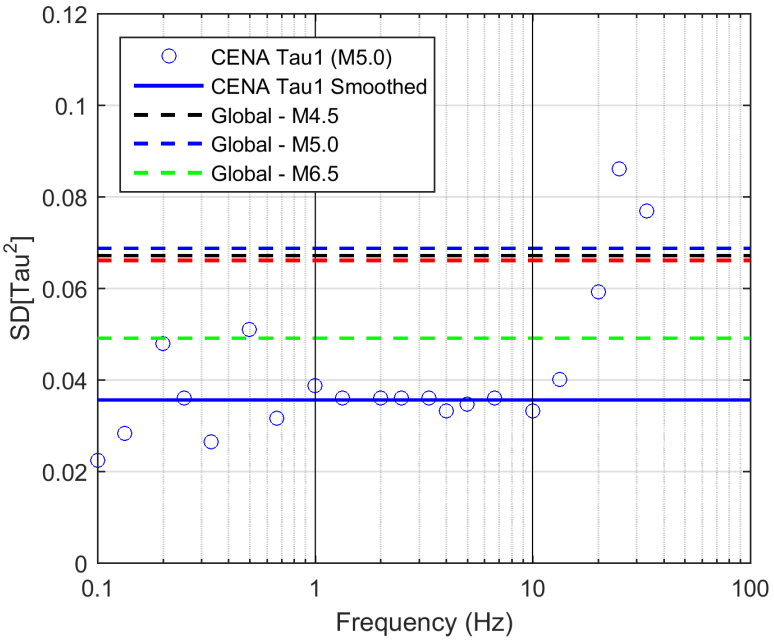


Figure 10–23 Global PhiSS (ϕ_{SS}) versus magnitude at f = 100, 10, 1, and 0.33 Hz. Error bars represent one standard error.

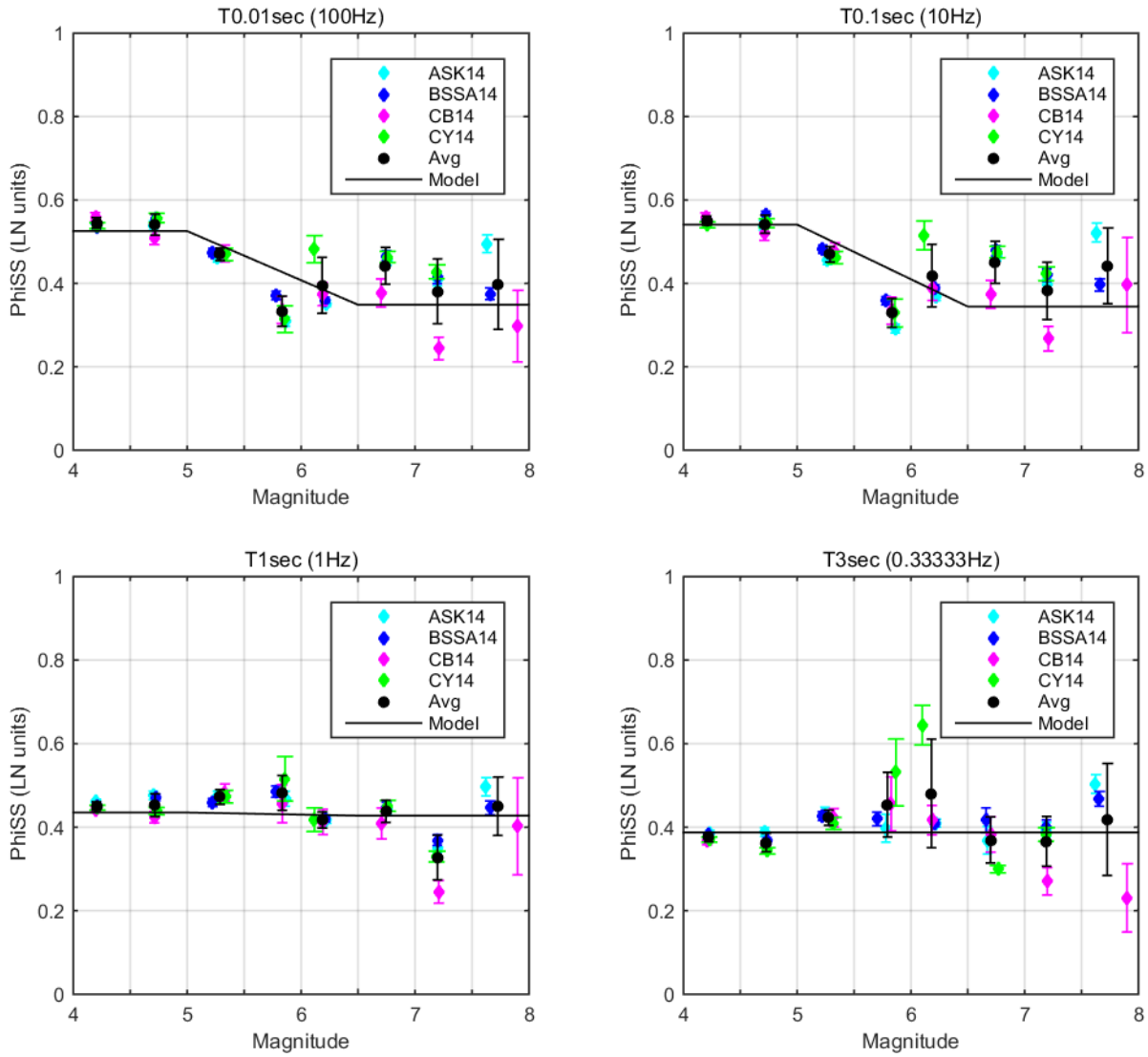


Figure 10–24 Coefficients of the global PhiSS (ϕ_{SS}) model versus frequency.

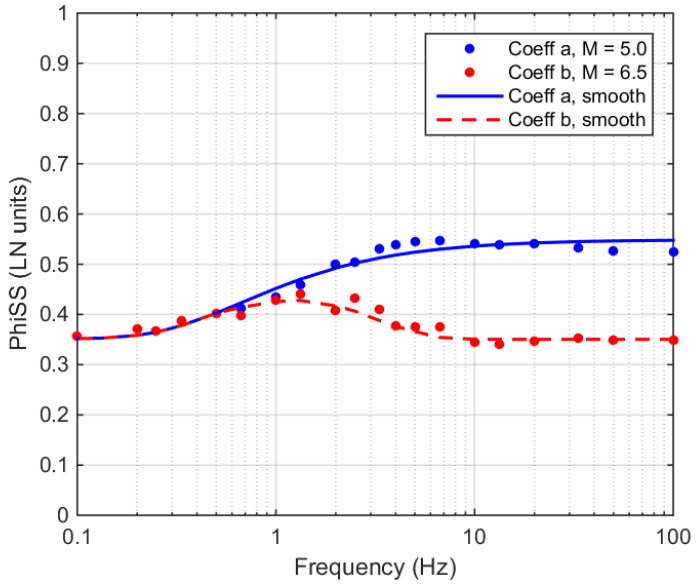


Figure 10–25 Coefficient of variation (CV) of PhiSS ($\phi_{SS,S}$) at the stations in the ASK14 dataset for different number of recordings per station (N). The blue line represents the CV from a simulated ground-motion dataset where PhiSS at all stations are equal.

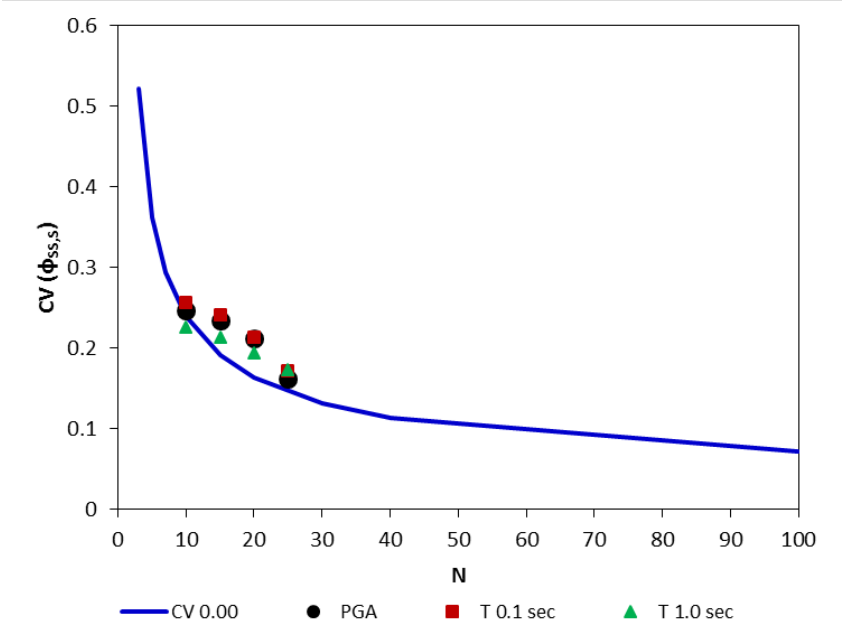


Figure 10–26 Coefficient of variation (CV) of PhiSS ($\phi_{ss,s}$) for the ASK14 dataset for different number of recordings per station (N). The blue lines represent the CV of the realizations (N) of simulated datasets with different assigned CV.

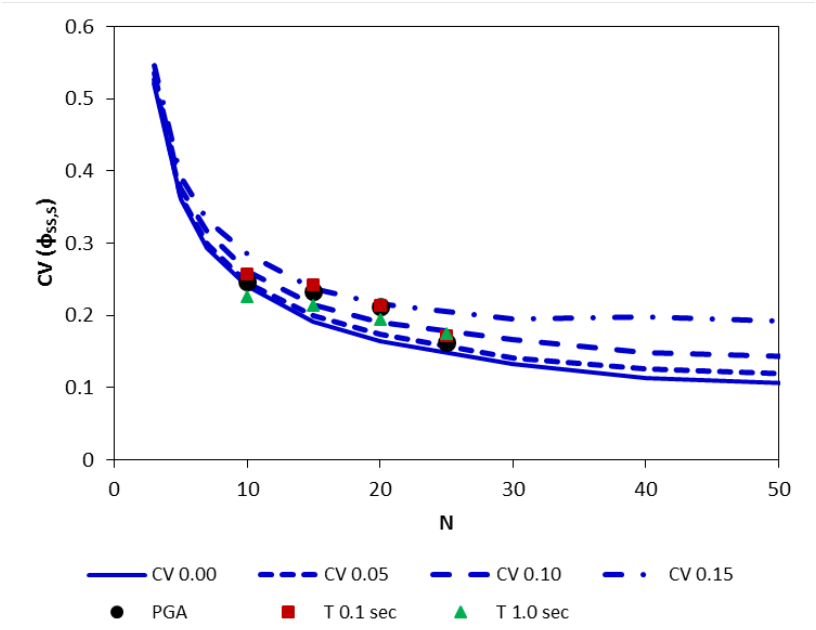


Figure 10–27 Variability in the global PhiSS (ϕ_{ss}) model at M = 5.0 (coefficient 'a').

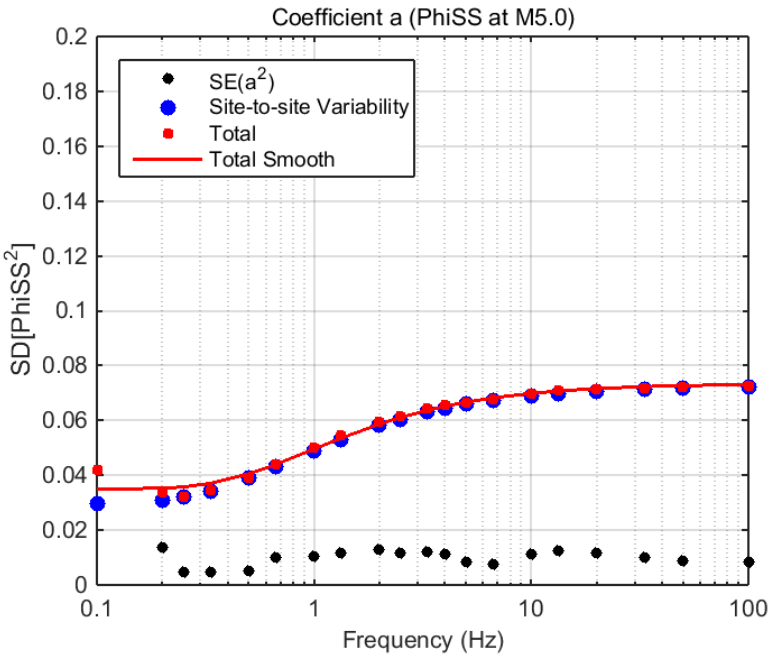


Figure 10–28 Variability in the global PhiSS (ϕ_{SS}) model at M = 6.5 (coefficient 'b').

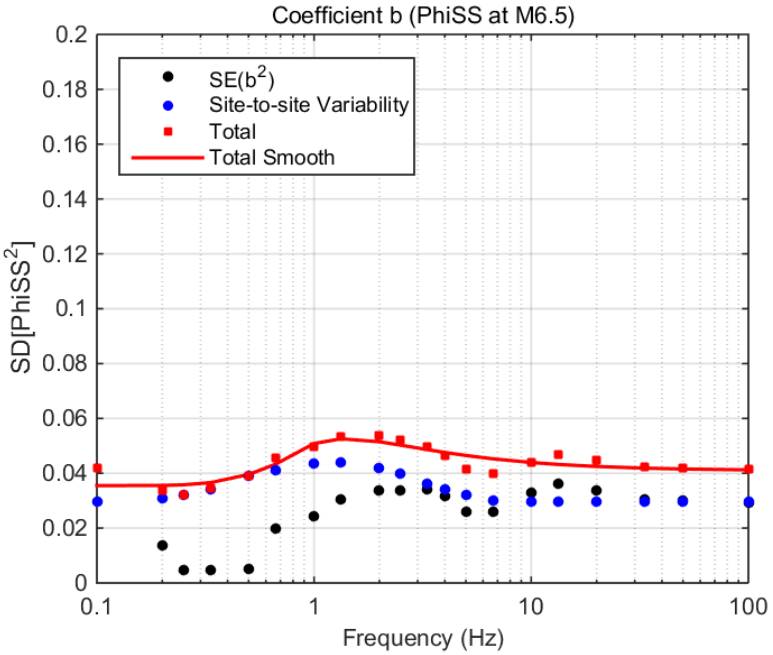


Figure 10–29 CENA constant PhiSS (ϕ_{SS}) model (best estimate). Error bars represent one standard error.

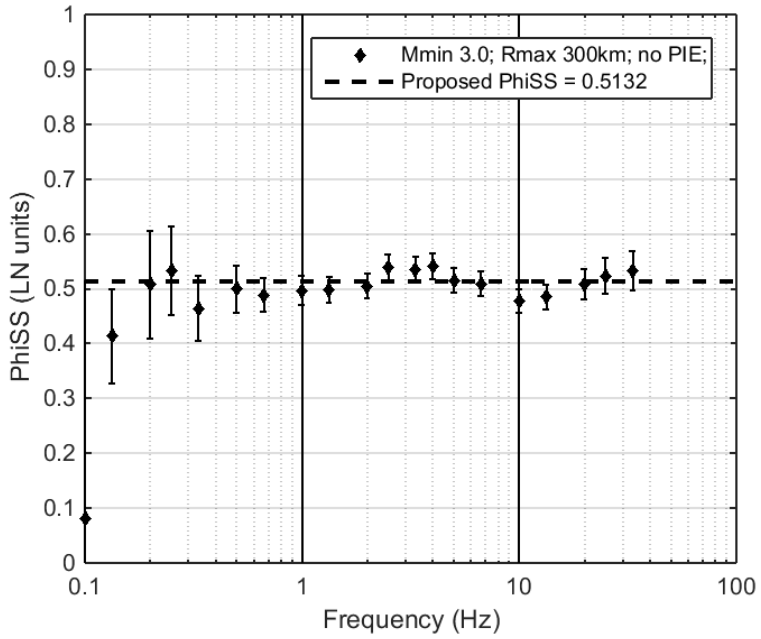


Figure 10–30 Variability in CENA constant PhiSS (ϕ_{SS}) model.

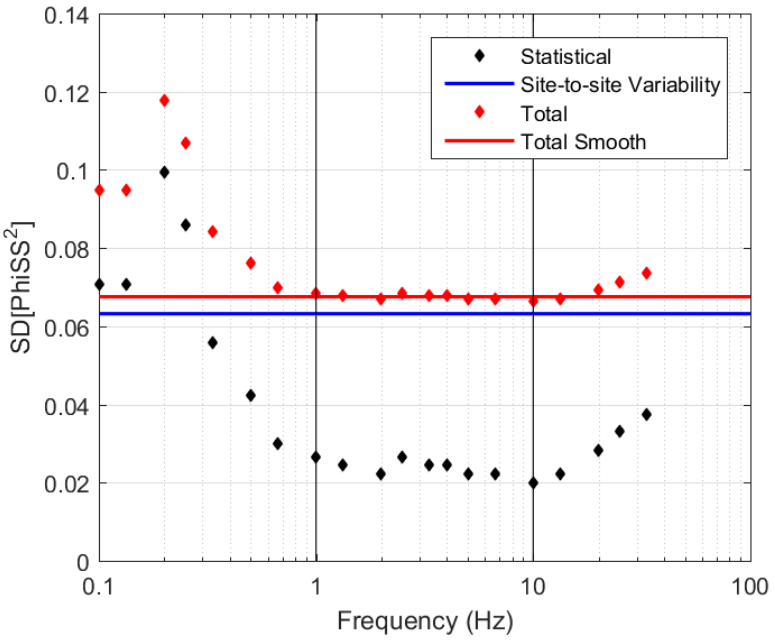


Figure 10–31 Coefficients of the CENA magnitude-dependent PhiSS (ϕ_{SS}) model. Error bars represent one standard error.

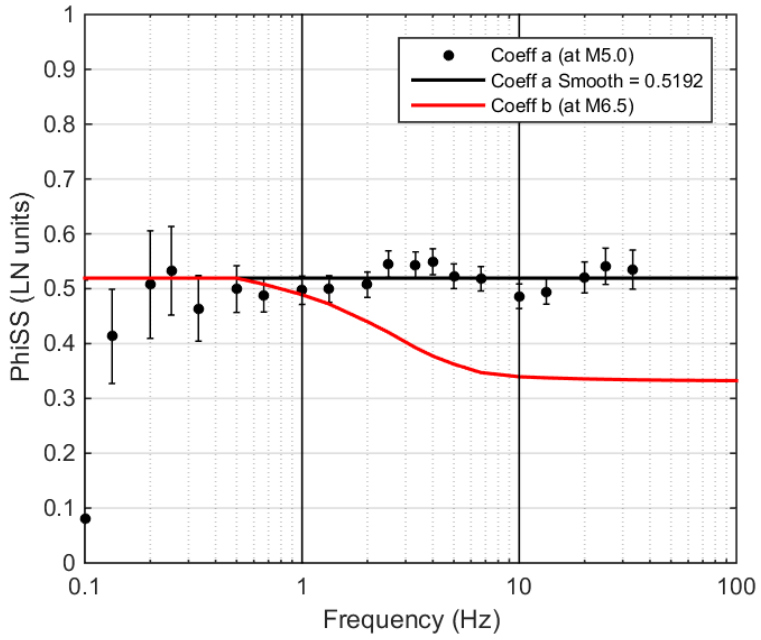


Figure 10–32 Variability in the CENA magnitude-dependent PhiSS (ϕ_{SS}) model at M = 5.0 (coefficient 'a').

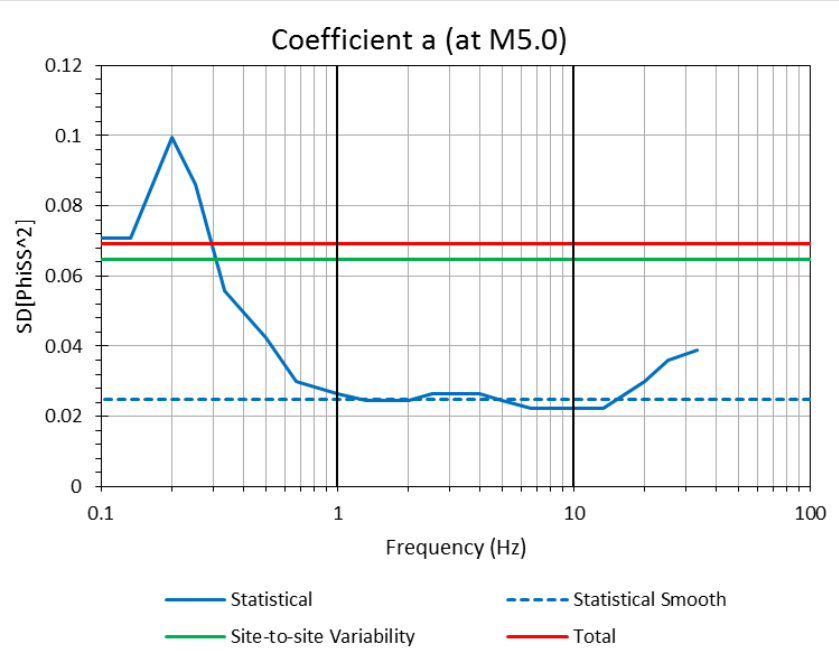


Figure 10–33 Variability in the CENA magnitude-dependent PhiSS (ϕ_{SS}) model at M = 6.5 (coefficient 'b').

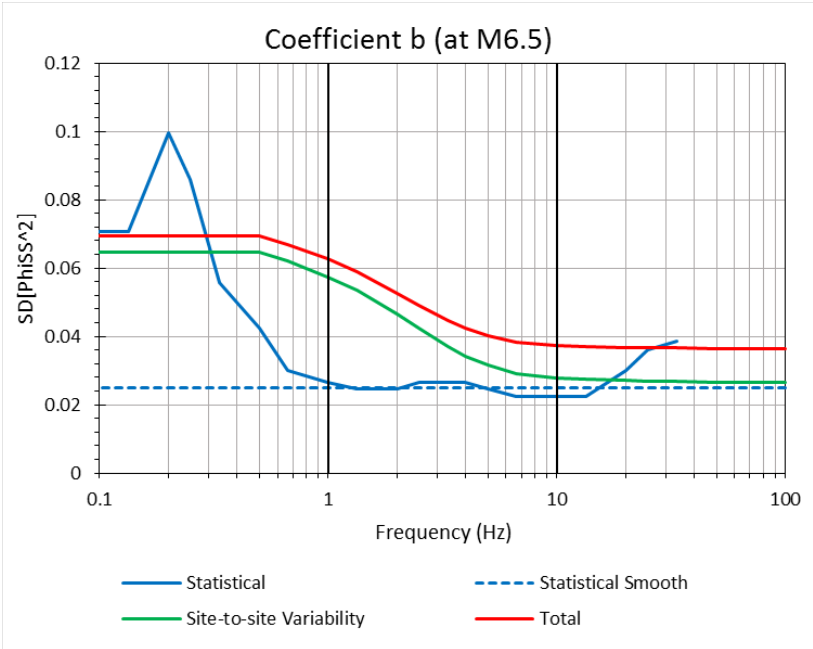


Figure 10–34 Comparison of CENA PhiS2S (ϕ_{S2S}) for PIE and tectonic events. Error bars represent one standard error.

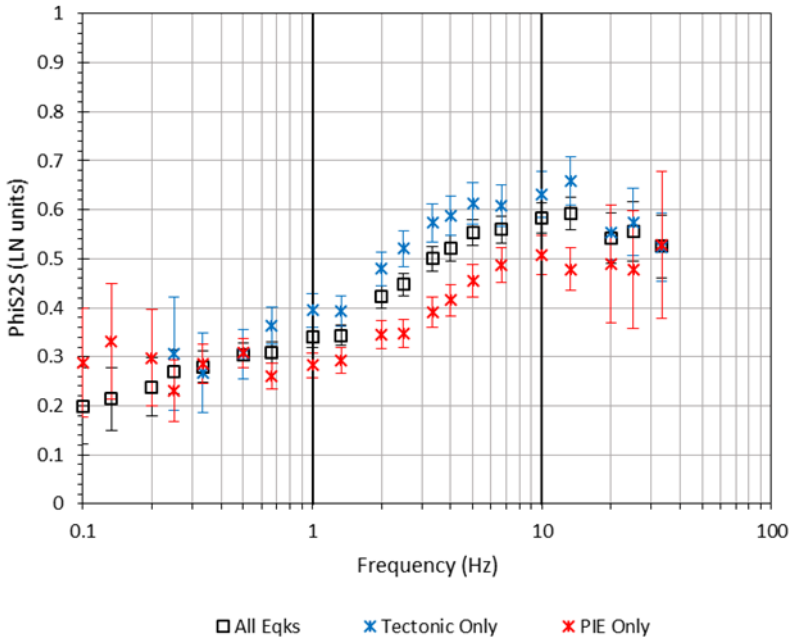


Figure 10–35 Location of the stations that recorded PIEs (red) and tectonic events (blue) for f = 4 Hz.

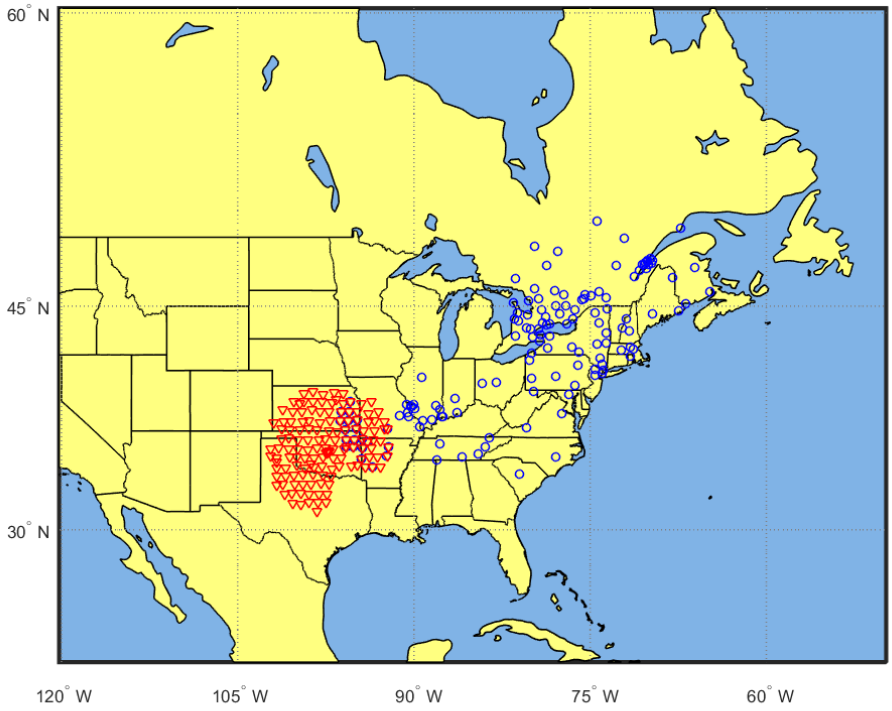


Figure 10–36 Comparison of PhiS2S (ϕ_{S2S}) for CENA stations with measured versus inferred V_{S30}. Error bars represent one standard error.

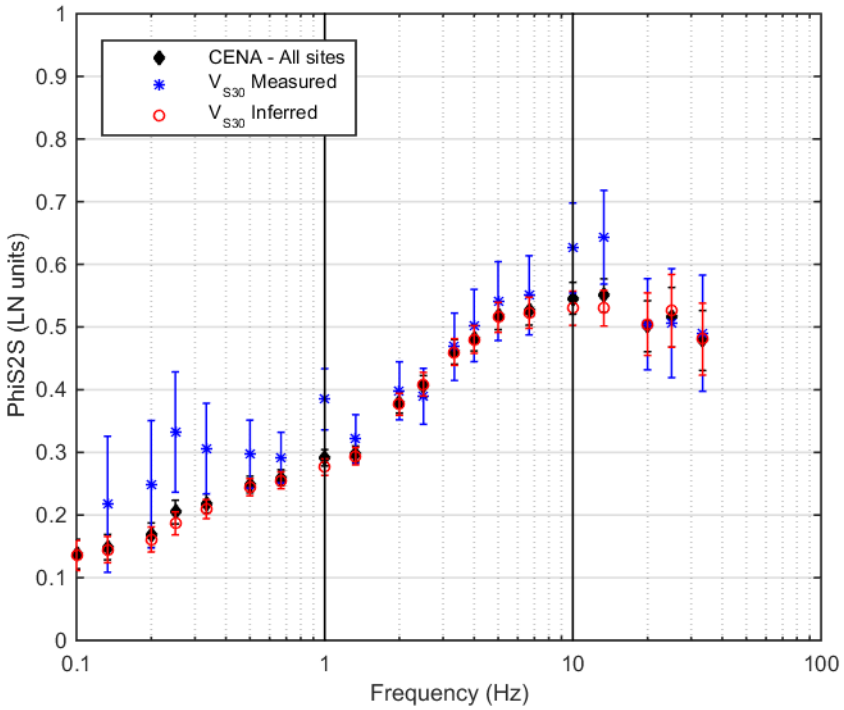


Figure 10–37 Comparison of PhiS2S (ϕ_{S2S}) for CENA, NGA-West2, and Japan. Error bars represent one standard error.

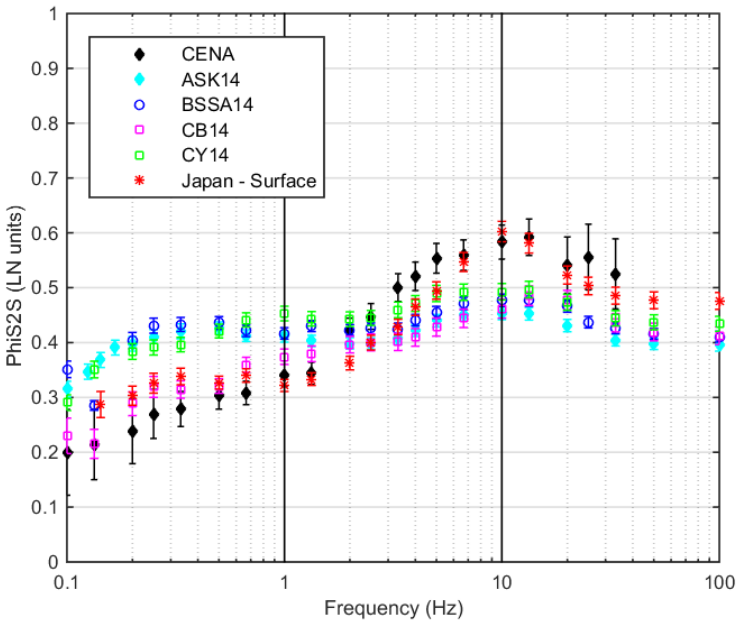


Figure 10–38 PhiS2S (ϕ_{S2S}) for CENA extrapolated outside of the 1-10 Hz frequency range.

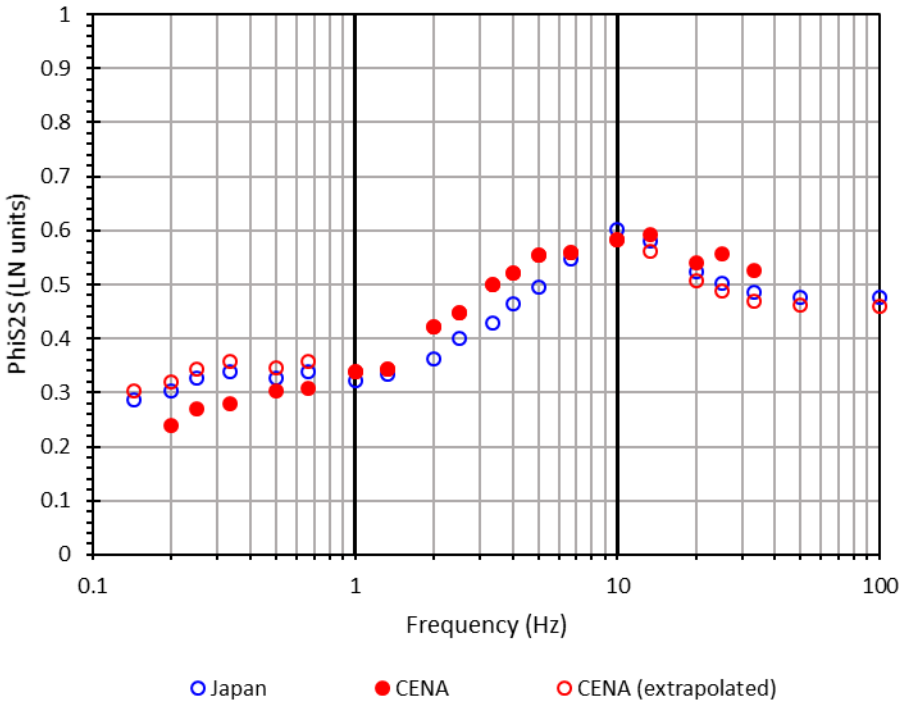


Figure 10–39 Variability of CENA PhiS2S (ϕ_{S2S}) extrapolated outside of the 1–10 Hz frequency range.

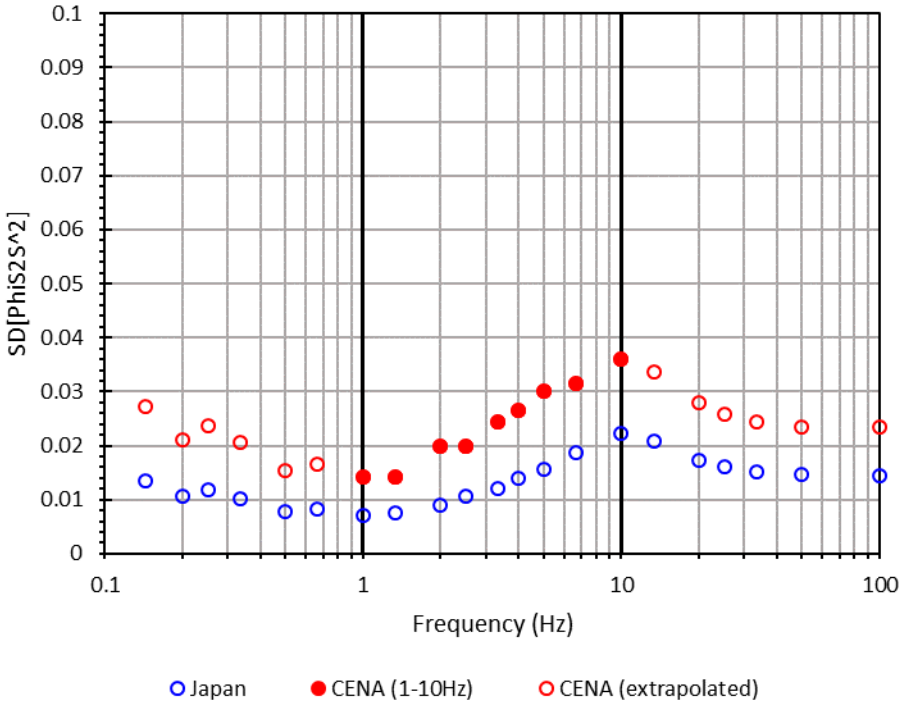


Figure 10–40 Magnitude dependence of PhiS2S (ϕ_{S2S}) observed using the ASK14 dataset. Error bars represent one standard error.

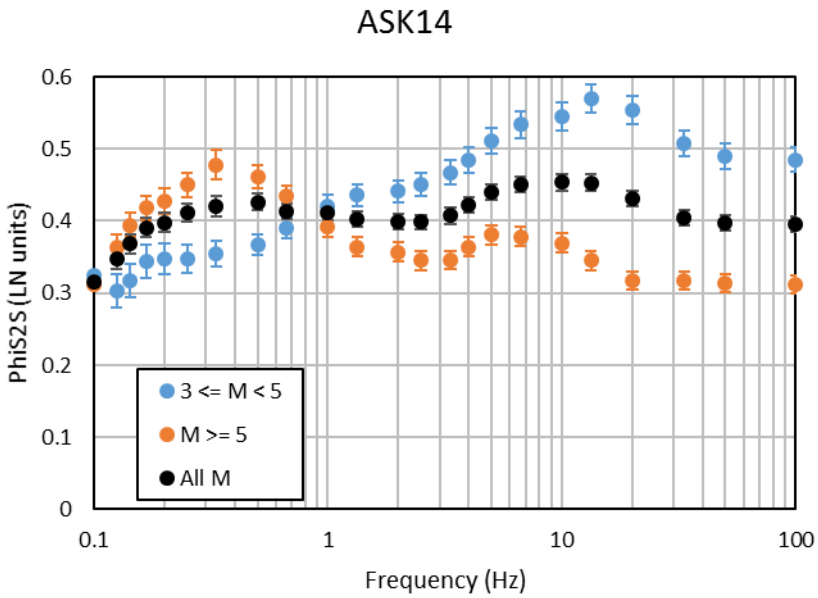


Figure 10–41 Magnitude dependence of PhiS2S (ϕ_{S2S}) observed using the BSSA14 dataset. Error bars represent one standard error.

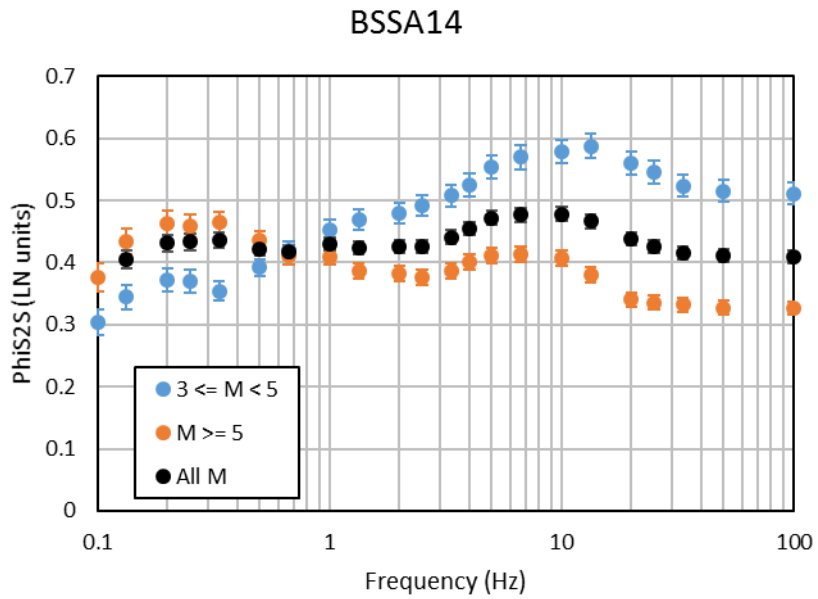


Figure 10–42 Magnitude dependence of PhiS2S (ϕ_{S2S}) observed using the CB14 dataset. Error bars represent one standard error.

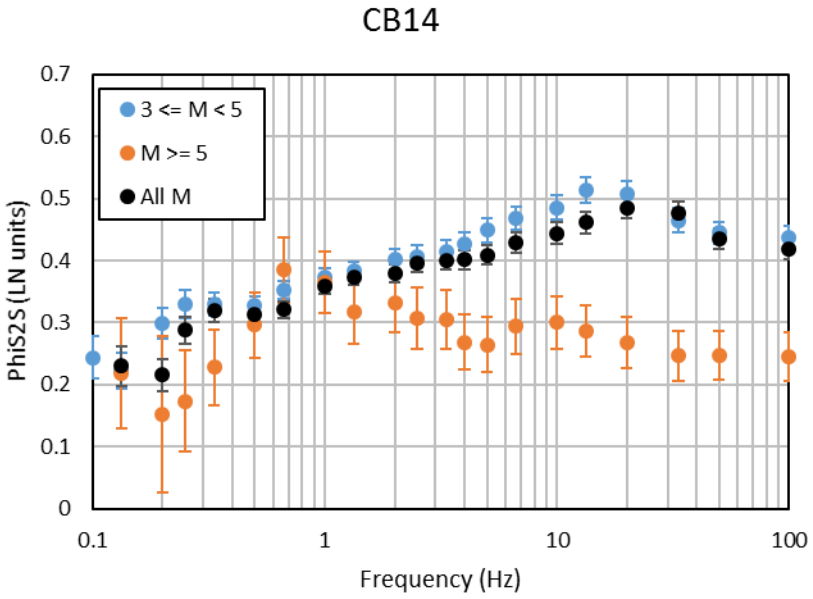


Figure 10–43 Magnitude dependence of PhiS2S (ϕ_{S2S}) observed using the CY14 dataset. Error bars represent one standard error.

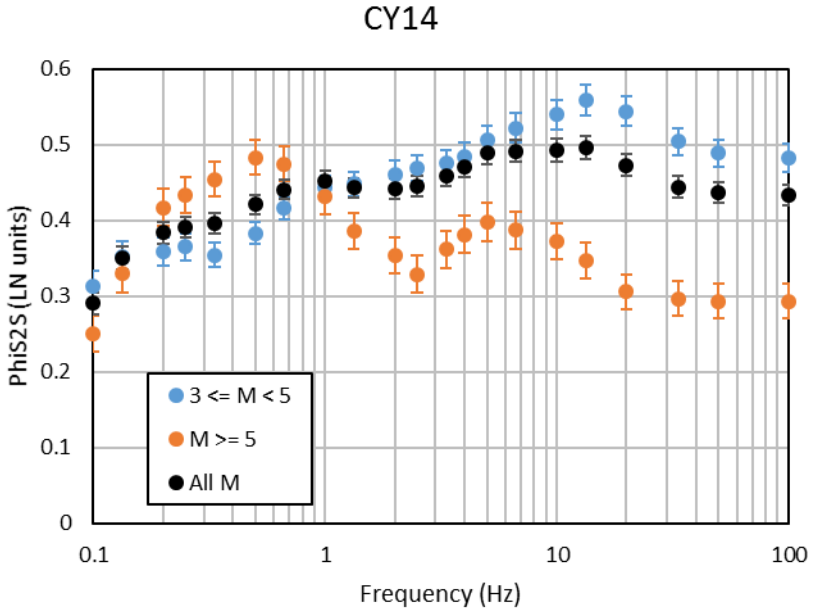


Figure 10–44 Magnitude dependence of PhiS2S (ϕ_{S2S}) observed using the Japanese surface dataset. Error bars represent one standard error.

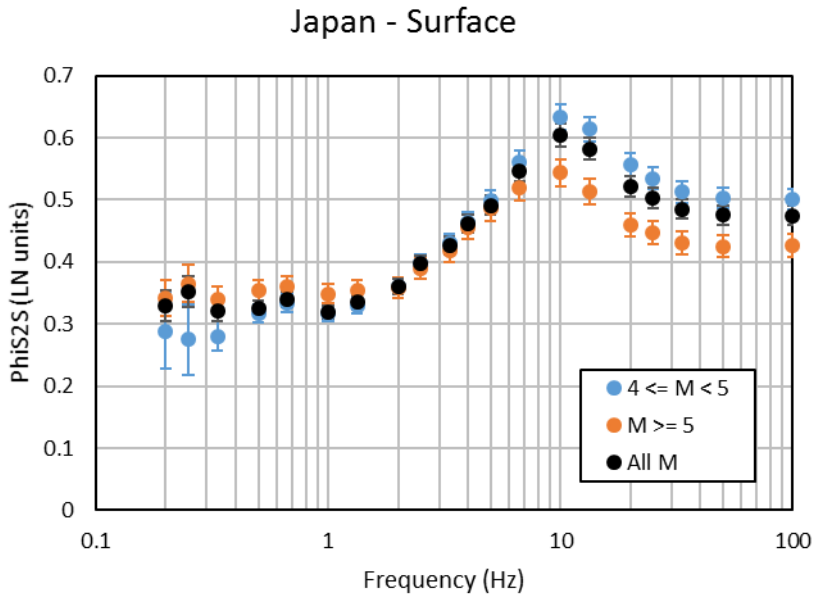


Figure 10–45 Magnitude dependence of PhiS2S (ϕ_{S2S}) observed using the Japanese borehole dataset. Error bars represent one standard error.

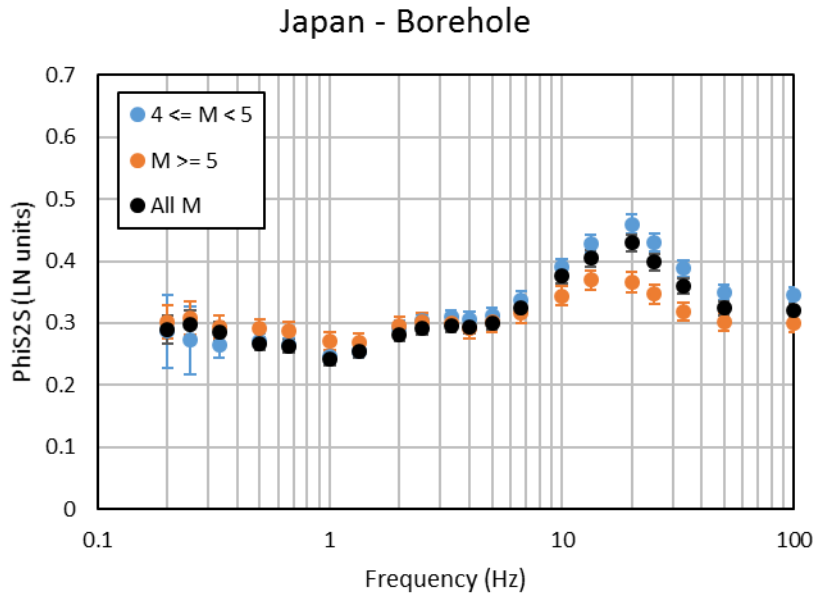


Figure 10–46 Comparison of PhiS2S (ϕ_{S2S}) for CENA and the small-magnitude subsets of the NGA-West2 and the Japanese datasets. Error bars represent one standard error.

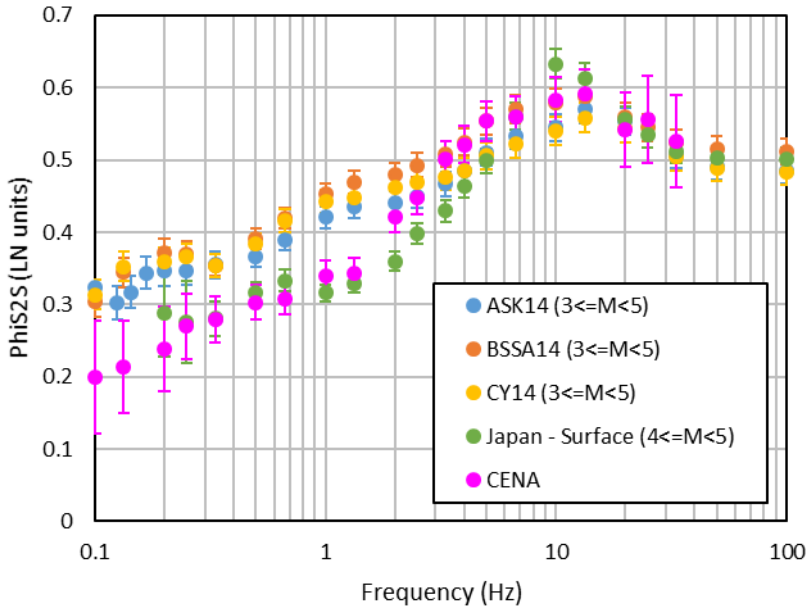


Figure 10–47 Comparison of PhiS2S (ϕ_{S2S}) derived from the entire CENA dataset (solid and dashed black lines corresponding to median, 5th, and 95th percentile) to CENA ϕ_{S2S} for rock conditions. Error bars represent one standard error.

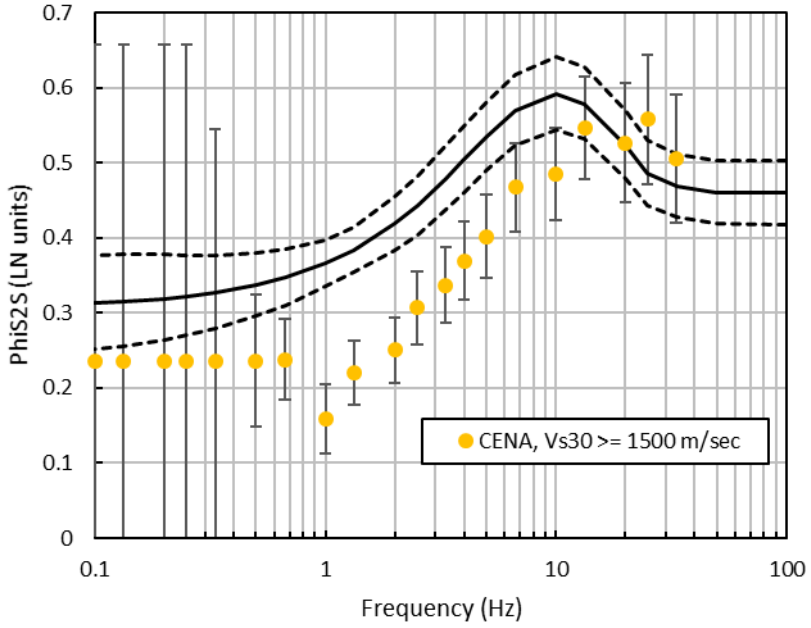


Figure 10–48 Comparison of PhiS2S (ϕ_{S2S}) for ASK14 derived from the entire dataset to Pϕ_{S2S} for rock versus soil sites . Error bars represent one standard error.

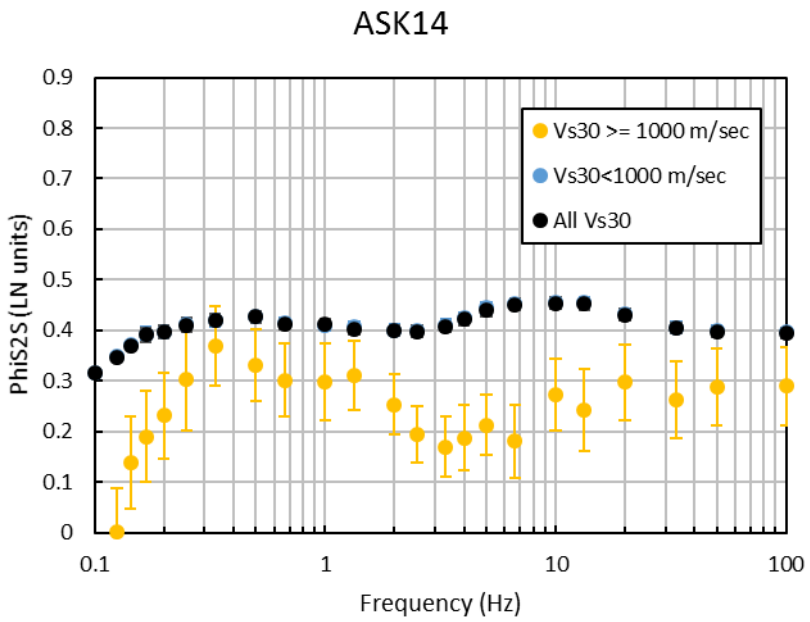


Figure 10–49 Comparison of PhiS2S (ϕ_{S2S}) for the Japanese surface and borehole datasets.

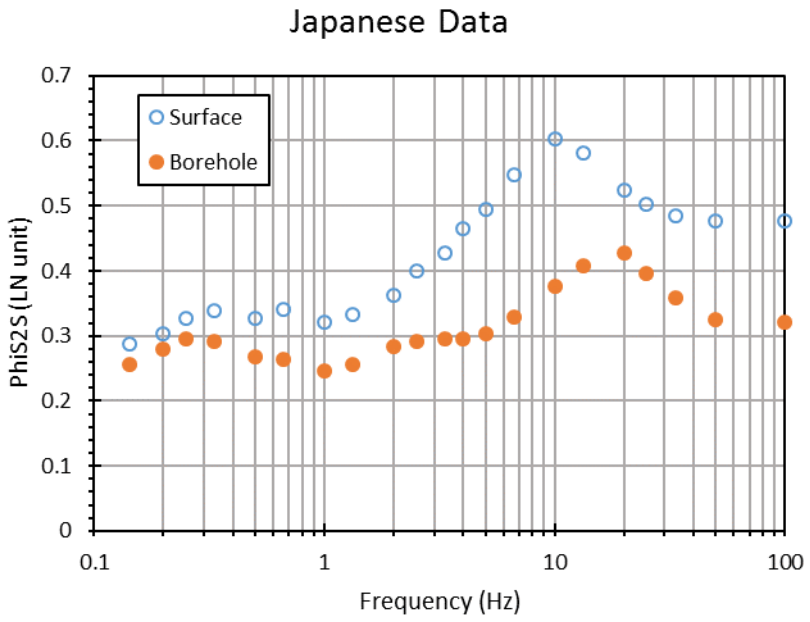


Figure 10–50 Comparison of PhiS2S (ϕ_{S2S}) for the Japanese borehole dataset with $V_S \geq 1500$ for subsets of the data with small and large magnitude. Error bars represent one standard error.

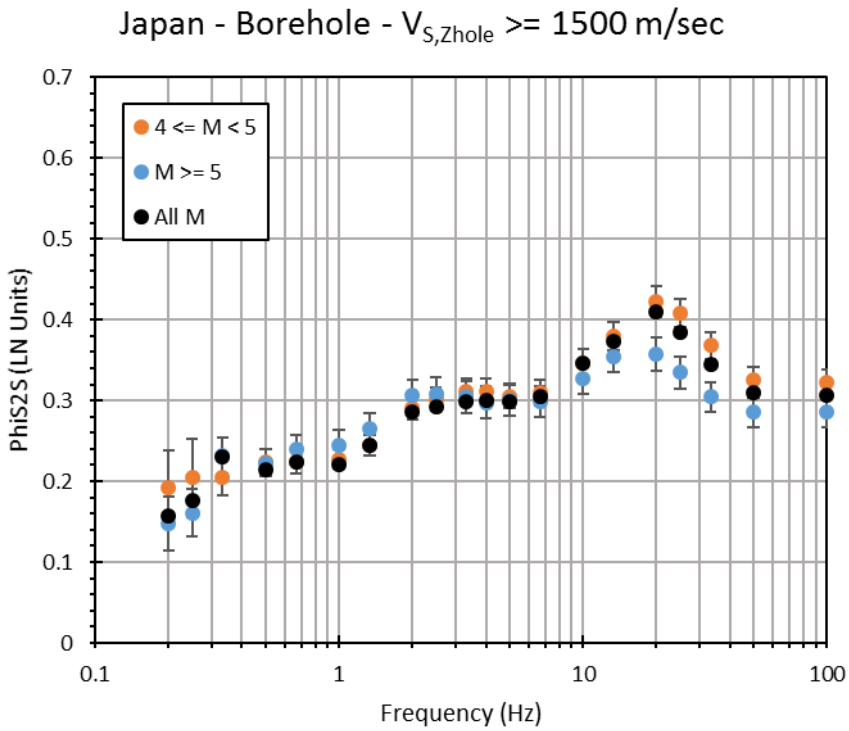


Figure 10–51 PhiS2S (ϕ_{S2S}) correction factors derived using the Japanese data.

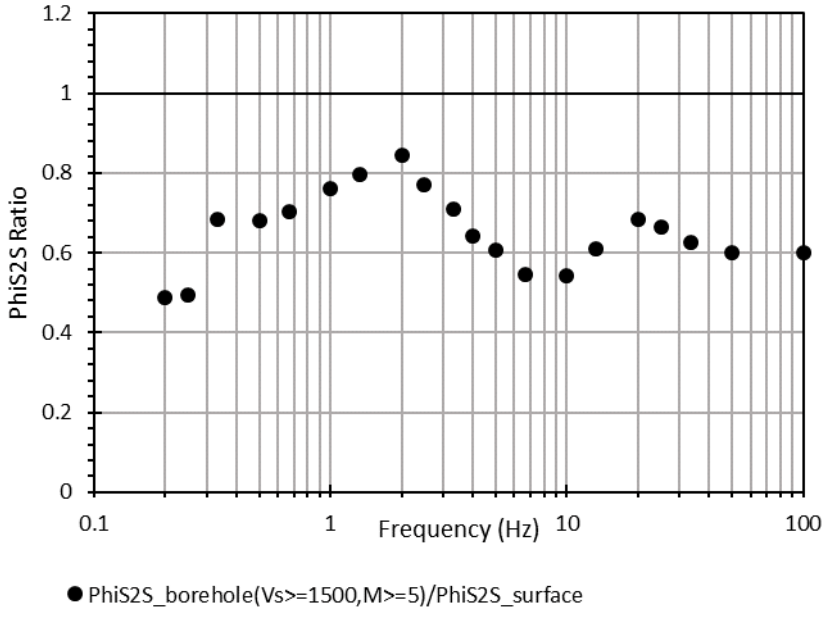


Figure 10–52 **CENA PhiS2S (ϕ_{S2S}) corrected for hard rock conditions and small magnitude effects.**

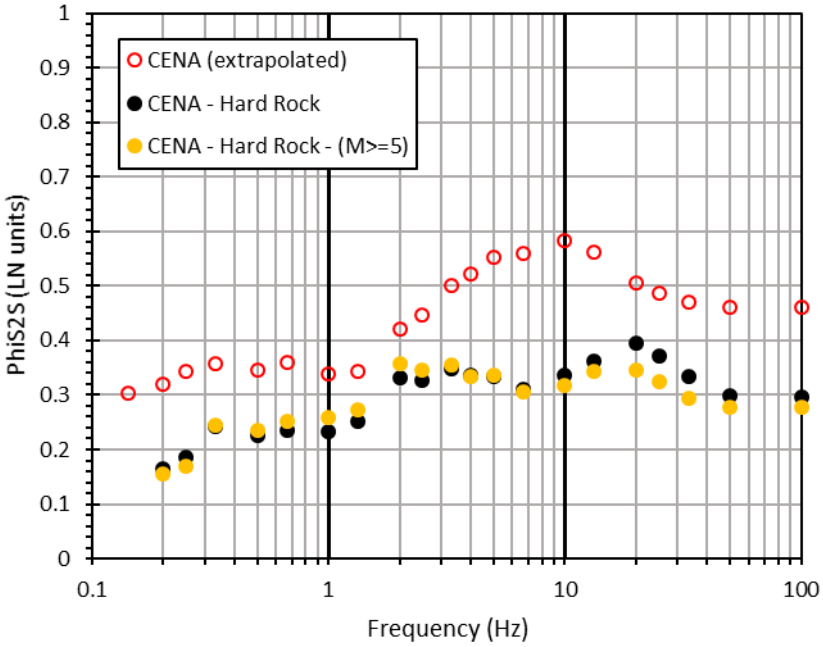


Figure 10–53 **Smoothed CENA mean PhiS2S (ϕ_{S2S}) model.**

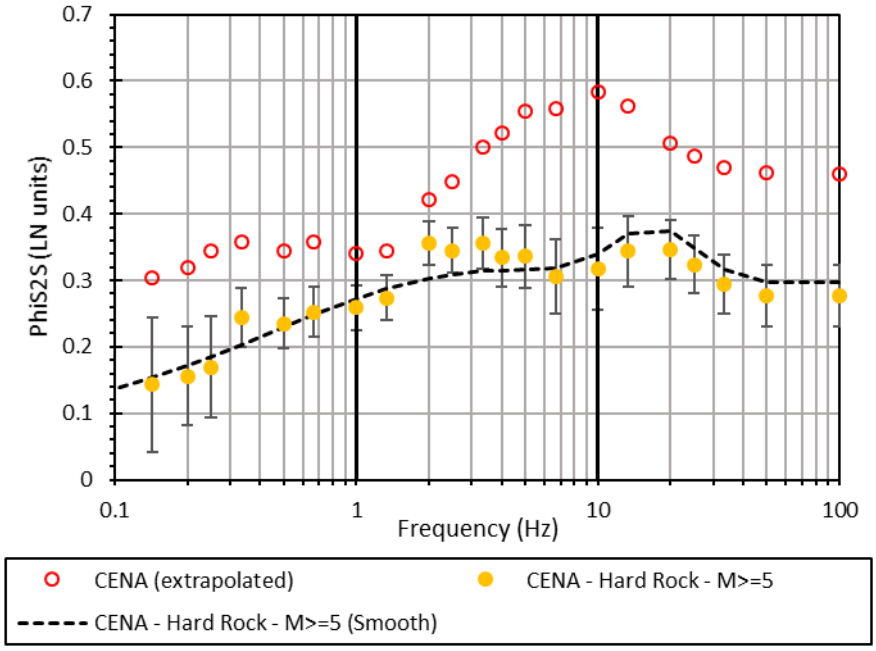
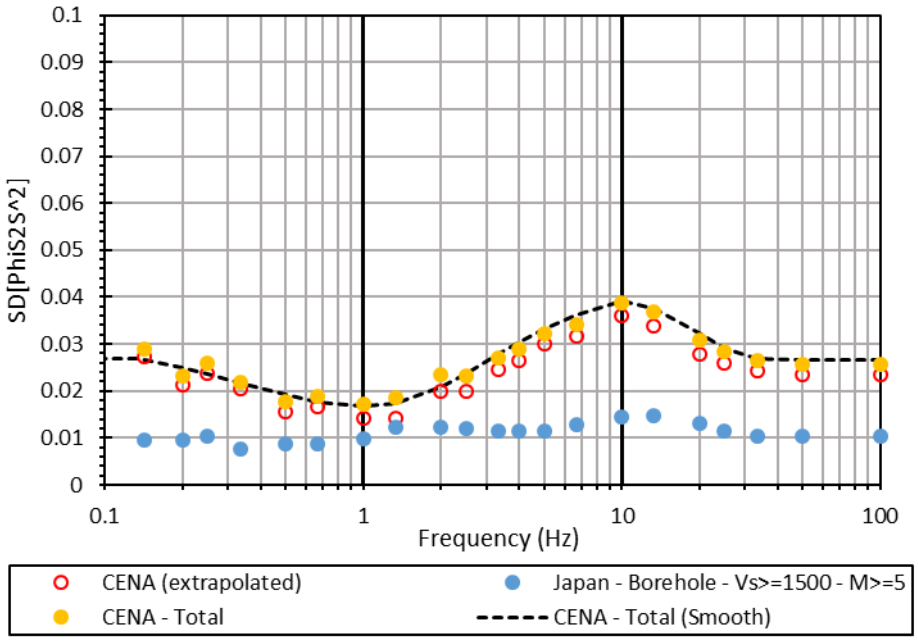


Figure 10–54 Standard deviation of CENA PhiS2S (ϕ_{S2S}) model.



11. Standard Deviation Characterization

11.1 Introduction

This chapter presents the evaluation of the candidate models for the between-event variability, τ, the single-station within-event variability, ϕ_{SS}, and the site-to-site variability, ϕ_{S2S}, presented in Chapter 10. These candidate models were developed based on evaluating the components of ground-motion variability in Central and Eastern North America (CENA) using the NGA-East dataset (Goulet et al. 2014) as well as in other regions using the Japanese dataset (Dawood 2014), and the global dataset developed for the Next Generation Attenuation Relationships for the western U.S. (NGA-West2) project (Ancheta et al. 2014).

Logic trees for CENA were developed for each of τ, ϕ_{SS}, and ϕ_{S2S}, and the assigned weights are discussed herein. Logic trees for τ and ϕ_{SS} were combined to create the single-station sigma (σ_{SS}) logic tree, which was then collapsed into three branches for efficiency in the computation of the hazard analysis. Similarly, τ, ϕ_{SS}, and ϕ_{S2S} logic trees were combined to create the logic tree for ergodic sigma (σ), which was finally collapsed into three branches.

The chi-square distribution describes the distribution of the sum of squares of independent standard normal random variables. As a result, we describe the uncertainty distributions for τ, ϕ_{SS}, and ϕ_{S2S} with continuous scaled chi-square distributions (Ang and Tang 2007). These continuous distributions are then represented by three discrete points selected at the 5^{th}, 50^{th}, and 95^{th} percentiles, and are referred to as low, central, and high branches with weights of 0.185, 0.63, and 0.185, respectively (Keefer and Bodily 1983). The calculation of the scaling parameter and the number of degrees-of-freedom of the chi-square distributions as well as the calculation of the central, high, and low branches are described in Al Atik (2015) for the different components of ground-motion variability.

The CENA models for the individual components of the ground-motion variability (τ, ϕ_{SS}, and ϕ_{S2S}) as well as the models for σ_{SS} and σ are compared in this chapter to existing aleatory variability models developed based on other datasets or used in other projects. These existing aleatory variability models were discussed in Section 10.2 and are summarized in Table 11−1. Table 11−1 summarizes the basic features (magnitude dependence, period dependence, and datasets used) of the CENA candidate aleatory variability models as well as existing models developed as part of other projects such as NGA-West2, Southwestern United States Ground Motion Characterization Project (SWUS) (GeoPentech 2015), Hanford Project, EPRI (2013) study, Atkinson and Adams (2013) etc.

As discussed in Sections 10.1.3 and 10.2.3, the general approach for NGA-East has been to decouple the development of median ground-motion models (GMMs) from aleatory variability models. As a result, the seed GMM developers were not tasked with providing aleatory variability models. Nevertheless, a few of them did provide companion sigma values to their median models. The aleatory variability of the seed GMMs are summarized in Section 10.2.3 and are not used for the development of CENA aleatory variability models.

11.2 Logic Tree for Tau (τ)

11.2.1 Elements of the Logic Tree

Figure 11–1 presents the logic tree structure for τ. It consists of two levels: the candidate models and their statistical uncertainty. Three candidate τ models were developed (Section 10.3.2) for CENA: global model, CENA constant model, and CENA magnitude-dependent model. Figure 11–2 presents a comparison of the three candidate τ models as a function of magnitude. All three models are period-independent. The solid lines show the median branches for each model, and the dashed lines show the 5th and 95th percentile branches for each model calculated assuming that τ^2 follows a scaled chi-square distribution (Al Atik 2015, Section 5.1.1) with mean and standard deviations presented in Section 10.3.2.

Figure 11–2 shows that CENA constant and magnitude-dependent τ models agree for **M** < 5.0. Both of the CENA τ models are smaller than the global model for **M** < 5.0. At larger magnitudes, the CENA magnitude-dependent model follows the same trend with magnitude as the global model. The CENA constant model agrees with the global model at **M** > 6.0.

11.2.2 Evaluation and Weights

In the evaluation of the three candidate τ models, the Technical Integrator (TI) Team chose to give higher weight to the global model over the two CENA models. This is due to three factors:

- The global model is built using a large uniformly processed global dataset, while the CENA dataset used to build the τ models (**M** larger than 3.0 and tectonic events only) is significantly smaller.
- The CENA dataset is limited to **M** < 5.5; therefore, it does not extrapolate reliably to large magnitudes. In contrast, the global model is based on data from a wide magnitude range and is reliable at large magnitudes.
- The CENA dataset is limited to the frequency range of 1 to 10 Hz (refer to Section 10.3.1.1) while the global model is reliable for the entire frequency range of 0.1 to 100 Hz. Beyond the frequency range of 1 to 10 Hz, the CENA models are extrapolated with a constant (Sections 10.3.2.2 and 10.3.2.3).

The statistical significance of the observed difference in Figure 11–2 between global and CENA τ for **M** < 5.0 (magnitude range where CENA data is most abundant) was evaluated by the TI team. Specifically, an F-test of equality of variance was conducted to test against the null hypothesis of equal variance of event terms for CENA and each of the NGA-West2 GMMs: Abrahamson et al. (2014) (ASK14), Boore et al. (2014) (BSSA14), Campbell and Bozorgnia (2014) (CB14), and Chiou and Youngs (2014) (CY14). The null hypothesis (H_0), the alternative hypothesis (H_a), and the test statistic F are defined as:

$$H_0 : \tau^2_{CENA} = \tau^2_{NGA-West2} \quad (11\text{–}1)$$

$$H_a : \tau^2_{CENA} \neq \tau^2_{NGA-West2} \quad (11\text{–}2)$$

$$F = \tau^2_{CENA} / \tau^2_{NGA-West2} \quad (11\text{–}3)$$

Under the null hypothesis, F has an F-distribution with numerator degrees-of-freedom of N1-1 and denominator degrees-of-freedom of N2-1, where N1 and N2 are the sample sizes of the two datasets. F, N1, and N2 are needed for the test. For CENA, τ values obtained from the mixed-effects regression with magnitude-dependent τ are used, and N1 is the number of events used in the regression. Note that the magnitude-dependent and the constant CENA models both yield similar τ values for **M** < 5.0. For the NGA-West2 models, two options were used to define τ and N for each of the four GMMs:

- Option 1 consisted of calculating τ for each of the NGA-West2 models based on the event terms of the model with **M** between 3.0 and 5.0, and N is the number of events in this magnitude range; however, this slightly underestimates the τ models proposed by the NGA-West2 developers.
- Option 2 consisted of averaging the proposed τ values for **M** between 3.0 and 5.0 (note that some of the NGA-West2 models are not constant in this range). N is the number of events used in each model with **M** between 3.0 and 5.0.

Tables 11–2 and 11–3 show the p-values of the F-test at peak ground velocity (PGV) and frequencies of 1 to 10 Hz with the NGA-West2 τ calculated using the two options discussed above. Shown in red are the cases where the p-values are less than 0.05, and the null hypothesis is rejected at a 5% significance levels. At most frequencies tested, Tables 11–2 and 11–3 show that the equality of τ^2 between CENA and NGA-West2 for **M** less than or equal to 5.0 cannot be rejected at 5% significance level. Based on the results of the F-test of equality of variance of the event terms (τ^2) between CENA and NGA-West2, the TI team concluded that the observed difference in τ between the CENA models and the global model for **M** < 5.0 are not statistically significant. This is likely due to the limited number of events in the CENA dataset compared to the NGA-West2 dataset with **M** < 5.0 (e.g., a total of 39 CENA tectonic events with **M** < 5.0 compared to 219 events in the ASK14 dataset).

Based on the conclusions from the F-test of equality of τ^2 between CENA and NGA-West2, and the issues outlined in the bullet points at the beginning of this section (i.e., the limited reliability of the CENA models at large magnitude and for frequencies outside of 1 to 10 Hz), the TI team decided to give the CENA τ models zero weights and to fully adopt the global τ model. Values of τ for the global model can be calculated using Equation (11-4) below discussed in Section 10.3.2.1, and the model coefficients at peak ground acceleration (PGA), PGV, and frequencies of 0.1 to 100 Hz are presented in Table 11–4 for the central, high, and low branches:

$$\tau = \begin{cases} \tau_1 & \text{for } \mathbf{M} \le 4.5 \\ \tau_1 + (\tau_2 - \tau_1) * \dfrac{(\mathbf{M} - 4.5)}{0.5} & \text{for } 4.5 < \mathbf{M} \le 5.0 \\ \tau_2 + (\tau_3 - \tau_2) * \dfrac{(\mathbf{M} - 5.0)}{0.5} & \text{for } 5.0 < \mathbf{M} \le 5.5 \\ \tau_3 + (\tau_4 - \tau_3) * \dfrac{(\mathbf{M} - 5.5)}{1.0} & \text{for } 5.5 < \mathbf{M} \le 6.5 \\ \tau_4 & \text{for } \mathbf{M} > 6.5 \end{cases} \quad (11\text{-}4)$$

We note that the coefficients of the global τ model at PGA are equal to those at f = 100 Hz.

11.2.3 Comparison to Existing Models

This section presents a comparison of the global τ model adopted for CENA to other existing τ models such as those derived for the Southwestern United States Ground Motion Characterization SSHAC Level 3 Project (SWUS) (GeoPentech 2015) and the Hanford project (Coppersmith et al. 2014). Similarities and differences between the global τ model and existing models are examined and discussed. We note that a comparison of the global τ model to the NGA-West2 τ models is not presented since the global τ model is based on the NGA-West2 models. The basic features of the existing τ models presented in this section and the datasets used to derive them are presented in Table 11–1.

Figure 11–3 compares the global τ model for CENA to the SWUS (GeoPentech 2015) τ model (central, high, and low branches). Both the global and SWUS models are magnitude-dependent and period-independent and are shown as a function of magnitude. The SWUS τ model is based on the average of τ models for the four NGA-West2 GMMs and for Zhao et al. (2006), which has a magnitude-independent τ. At small magnitudes (**M** = 4.0 to 5.0), the SWUS τ model is lower than the global model and has larger uncertainty (i.e., the spread between the 5th and 95th percentiles) due to the inclusion of the Zhao et al. (2006) τ model in SWUS, which is smaller than the four NGA-West2 models at small magnitudes. At larger magnitudes, the SWUS and the global τ models are comparable.

Figure 11–4 compares the global τ model for CENA to the Hanford τ model (Coppersmith et al. 2014) as a function of frequency for **M** = 5.0, 6.0, and 7.0. The Hanford τ model is based on the average of the four NGA-West2 τ models but adopted different magnitude breaks and smoothing with period (period-dependent τ) than the global model. Figure 11–4 shows that the Hanford and the global τ models are generally comparable in terms of their median values and uncertainty ranges. The Hanford τ model is generally larger than the global τ model at low frequencies and smaller than the global τ model at high frequencies. This is the result of the frequency-smoothing function that the Hanford model applied as opposed to the frequency-independence adopted for the global τ model discussed in Section 10.3.2.1.

11.3 Logic Tree for PhiSS (ϕ_{SS})

11.3.1 Elements of the Logic Tree

Figure 11–5 presents the logic tree structure for ϕ_{SS}, which mirrors the structure of the model for τ. Like the τ model, we developed three candidate ϕ_{SS} models for CENA (Section 10.3.3): global model, CENA constant model, and CENA magnitude-dependent model. Figure 11–6 compares the three candidate ϕ_{SS} models as a function of magnitude at PGV and f = 0.1, 0.2, 1, 10, and 100 Hz. The solid lines show the median branches for each model, and the dashed lines show the 5th and 95th percentile branches calculated assuming ϕ_{SS}^2 follows a scaled chi-square distribution with mean and standard deviations presented in Section 10.3.3. Both the global and the CENA magnitude-dependent ϕ_{SS} models are period-dependent, while the constant CENA

model is period-independent and magnitude-independent. Figure 11–7 compares the three candidate models as a function of frequency at **M** = 5.0, 6.0, and 7.0.

Figure 11–6 shows that for periods less than 1 sec, both CENA ϕ_{SS} models agree for **M** < 5.0. For longer periods, all three models become magnitude-independent with the two CENA ϕ_{SS} models having comparable values. Moreover, the CENA magnitude-dependent and the global ϕ_{SS} models are comparable at periods less than 1 sec. At longer periods, the CENA magnitude-dependent ϕ_{SS} become larger than the global ϕ_{SS} because the CENA magnitude-dependent coefficient at **M** = 5.0 is constant, while the global coefficient at **M** = 5.0 decreases as period increases. Figure 11–7 shows that for **M** = 5.0, all three models have comparable ϕ_{SS} values at high frequencies. As magnitude increases, the two magnitude-dependent models stay comparable at high frequencies while the constant CENA model has larger ϕ_{SS} values.

11.3.2 Evaluation and Weights

Analyses of ground-motion data covering a wide magnitude range indicate a magnitude-dependent trend for ϕ_{SS} (e.g., NGA-West2, SWUS, Hanford, etc.). Because available ground-motion data are mostly recorded on soil or soft-rock sites, the TI team evaluated the magnitude-dependence of ϕ_{SS} for hard-rock sites relevant to the target site conditions in CENA. To this end, the within-event residuals of the Japanese crustal ground-motion data recorded at the surface and the borehole level (Rodriguez-Marek, personal communication) and described in Section 10.3.1.3 were used in this evaluation. The borehole dataset was important for this exercise because of its abundance of recordings on hard-rock conditions (refer to Figure 10–12) that is not available in other datasets. Figure 11–8 shows ϕ_{SS} for the Japanese surface and borehole data with $V_S \geq 1500$ m/sec for two subsets of the data: **M** < 5.0 and **M** ≥ 5.0. Figure 11–8 clearly indicates a dependence of ϕ_{SS} on magnitude for rock site conditions similar to the trend observed with the surface data (ϕ_{SS} for **M** < 5.0 is greater than that for **M** ≥ 5.0 at high frequencies). The TI team concluded that the magnitude-dependence of ϕ_{SS} observed for soft sites can be extended to hard-rock conditions. This is likely the result of path effects being stronger and more variable for small magnitudes, resulting in larger ϕ_{SS} values compared to those values obtained for large magnitudes for both soil and rock sites. We note that, for each magnitude subset of the data, Figure 11–8 shows a difference in ϕ_{SS} between the surface and the borehole data. The TI team opted not to correct the candidate ϕ_{SS} models derived in Section 10.3.3 for hard-rock conditions based on the difference in ϕ_{SS} observed for the Japanese data between surface and borehole conditions. This decision was based on the fact that the difference in ϕ_{SS} between surface and borehole levels is relatively small (on the order of 5 to 6% at high frequencies), contrary to the much bigger difference observed in ϕ_{S2S} between soft- and hard-rock site conditions (Section 10.3.4.2).

In the evaluation of the three candidate ϕ_{SS} models, the TI team decided to give a higher weight to the global model over the CENA models for the same reasons as were given for the τ model, i.e., the limitations of the CENA models in frequency range to 1–10 Hz and the extrapolation to large magnitudes (see Section 11.2.2). Similar to the approach adopted for the evaluation of the

τ models, the statistical significance of the difference between global and CENA ϕ_{SS} for **M** between 3.0 and 5.0 was evaluated by the TI team using the same F-test methodology as for τ.

Table 11–5 shows the p-values of the F-test at PGV and for frequencies between 1 to 10 Hz. Values in red indicate cases where the p-values are less than 0.05, and the null hypothesis is rejected at a 5% significance level. Table 11–5 shows that for 5% significance level, ϕ_{SS} values for CENA and NGA-West2 models cannot be assumed equal at PGV and spectral periods of 0.1, 0.4, 0.75, and 1 sec. Since the F-test did not show that the variances (ϕ_{SS}^2) can be assumed to be equal between CENA and NGA-West2 at most of the frequencies tested and at 5% significance level, the TI team concluded that the difference in ϕ_{SS} between the CENA models and the global model are not insignificant at **M** < 5. As a result, the TI team decided to adopt all three candidate ϕ_{SS} models in the logic tree.

Because the global ϕ_{SS} model is more reliable than the CENA models at large magnitude and over the entire frequency range of 0.1 to 100 Hz as discussed in Section 11.2.2, the TI team gave the global model a higher weight compared to the CENA models. Weights of 0.8, 0.1, and 0.1 were assigned to the global, CENA magnitude-dependent, and CENA constant ϕ_{SS} models, respectively. The values of ϕ_{SS} for the global model can be calculated using Equation (11-5) below (discussed in Section 10.3.3.1), where coefficients 'a' and 'b' are listed in Table 11–6 for the central, high, and low branches of the model:

$$\phi_{SS} = \begin{cases} a & \text{for } \mathbf{M} \le 5.0 \\ a + (\mathbf{M} - 5.0) * \dfrac{(b-a)}{1.5} & \text{for } 5.0 < \mathbf{M} \le 6.5 \\ b & \text{for } \mathbf{M} > 6.5 \end{cases} \qquad (11\text{-}5)$$

Table 11–7 shows the CENA constant ϕ_{SS} values for the central, high, and low branches. The CENA magnitude-dependent ϕ_{SS} model has the same functional form as the global ϕ_{SS} model (refer to Section 10.3.3.3). The values of ϕ_{SS} for the CENA magnitude-dependent model can therefore be calculated using Equation (11–5) above where coefficients 'a' and 'b' are given in Table 11–8 for the central, high, and low branches. The coefficients of the ϕ_{SS} models at PGA in Tables 11–6 to 11–8 are equal to their corresponding values at f = 100 Hz as a result of the constant trend of the models at high frequencies shown in Figure 11–7.

11.3.3 Comparison to Existing Models

This section compares the ϕ_{SS} models for CENA to existing ϕ_{SS} models developed for other studies such as the Hanford project and the SWUS project. We note that the ϕ_{SS} models for CENA are not compared to the NGA-West2 ϕ_{SS} because the global ϕ_{SS} model is derived based on NGA-West2 ϕ_{SS} and such comparison would be redundant.

Figure 11–9 compares the three ϕ_{SS} models for CENA to the Hanford ϕ_{SS} model (central, high, and low branches) at spectral periods of 0.01 and 1 sec. Recall that the Hanford model is based on the PEGASOS Refinement Project (PRP) (Renault et al. 2010) data and is magnitude-dependent. Figure 11–9 shows that the Hanford ϕ_{SS} model is generally similar to the global ϕ_{SS}

model. The uncertainty range of the Hanford model is smaller than that of the CENA models because the Hanford model used a coefficient of variation of ϕ_{SS} equal to 0.1 in calculating the site-to-site variability while the NGA-East study used a coefficient of variation of 0.12 (Al Atik 2015)

Figure 11−10 compares the three ϕ_{SS} models for CENA to one of the SWUS ϕ_{SS} models at spectral periods of 0.01 and 1 sec. As discussed in Section 10.2 and summarized in Table 11−1, SWUS developed several ϕ_{SS} models based on the controlling sources (magnitude and distance) in their hazard results. The SWUS model presented in Figure 11−10 is magnitude-dependent with magnitude breaks at **M** = 5.0 and 7.0, and is derived based on the California NGA-West2 data with **M** greater than or equal to 5.0 and distance up to 50 km. Figure 11−10 shows that the SWUS ϕ_{SS} model is similar to the global model. The uncertainty in the SWUS model is smaller than that of the global model because the SWUS model did not include the uncertainty in the magnitude-dependent model fit to the data. This uncertainty was negligible for SWUS model because the ϕ_{SS} values versus magnitude were similar for the four NGA-West2 models in the magnitude and distance ranges used to build the model.

11.4 Logic Tree for PhiS2S (ϕ_{S2S})

Figure 11−11 presents the logic tree for ϕ_{S2S}. The ϕ_{S2S} model developed using CENA data and discussed in Section 10.3.4 was adopted. The model was extrapolated for frequencies outside of the reliable 1 to 10 Hz frequency range for the CENA data using the scaling versus frequency of ϕ_{S2S} for the Japanese data (Dawood 2014) based on the similarity of ϕ_{S2S} for CENA and Japan. Moreover, the ϕ_{S2S} model was corrected for the small-magnitude bias and for application to hard-rock conditions using correction factors derived based on ϕ_{S2S} for the Japanese data at the surface and the borehole levels (refer to Sections 10.3.4.1 and 10.3.4.2). Table 11−9 presents the values of the homoscedastic ϕ_{S2S} model for CENA, whereby the central, high, and low branches are obtained by assuming that ϕ_{S2S}^2 follows a scaled chi-square distribution with mean and standard deviation values presented in Section 10.3.4. Figure 11−12 presents the central, high, and low values of the CENA ϕ_{S2S} model as a function of frequency. The values of ϕ_{S2S} at PGA are equal to those at f = 100 Hz based on the fact that ϕ_{S2S} become constant at high frequencies as shown in Figure 11−12.

11.5 Logic Tree for Phi (ϕ)

The models for CENA ϕ are derived by combining the models for ϕ_{SS} and ϕ_{S2S} presented in Sections 11.3 and 11.4, respectively. The results of the mixed-effects regression performed using the CENA data to separate the within-event residuals into site terms and single-station within-event residuals (described in Section 10.3.1.1) show very weak negative correlation to no correlation between the uncertainties in the estimation of ϕ_{SS}^2 and ϕ_{S2S}^2. Therefore, the TI team assumes ϕ_{SS}^2 and ϕ_{S2S}^2 to be uncorrelated. Given this assumption, ϕ^2 and its standard deviation, SD(ϕ^2), are written as:

$$\phi^2 = \phi_{SS}^2 + \phi_{S2S}^2 \tag{11–6}$$

$$SD\left(\phi^2\right) = \sqrt{\left[SD\left(\phi_{SS}^2\right)\right]^2 + \left[SD\left(\phi_{S2S}^2\right)\right]^2} \tag{11–7}$$

where the mean and standard deviations of the ϕ_{SS}^2 and ϕ_{S2S}^2 are discussed in Sections 10.3.3 and 10.3.4, respectively. As a result, the logic tree for CENA ϕ is shown in Figure 11–13. The global ϕ model is derived by combining the CENA ϕ_{S2S} model and the global ϕ_{SS} model. The CENA constant ϕ model is derived by combining the CENA ϕ_{S2S} model and the CENA constant ϕ_{SS} model. The CENA magnitude-dependent ϕ model is the result of combining the CENA ϕ_{S2S} model and the CENA magnitude-dependent ϕ_{SS} model. The magnitude-dependent ϕ models (global and CENA) have the following form:

$$\phi = \begin{cases} a & \text{for } \mathbf{M} \le 5.0 \\ a + (\mathbf{M} - 5.0) * \dfrac{(b-a)}{1.5} & \text{for } 5.0 < \mathbf{M} \le 6.5 \\ b & \text{for } \mathbf{M} > 6.5 \end{cases} \tag{11–8}$$

where coefficients ‘a’ and ‘b’ for the central, high, and low branches of the global, and the CENA magnitude-dependent model are given in Tables 11–10 and 11–11, respectively. The central, high, and low values of coefficients ‘a’ and ‘b’ were obtained by calculating the mean and standard deviations of ϕ^2 at the magnitude breaks of 5.0 and 6.5 according to Equations (11–6) and (11–7) and then representing the continuous chi-square distribution of ϕ^2 by a three-point distribution as described in Section 11.1.The values of the CENA constant ϕ model are given in Table 11–12.

Figure 11–14 shows the three ϕ models for CENA as a function of magnitude: CENA constant model, CENA magnitude-dependent model, and global model at PGV and spectral periods of 0.01, 0.1, 0.5, 1, and 5 sec (frequencies of 100, 10, 2, 1, and 0.2 Hz). The solid lines show the median branches for each model, and the dashed lines show the 5th and 95th percentile branches for each model calculated assuming ϕ^2 follows a scaled chi-square distribution with mean and standard deviations calculated as discussed above. Similar to the observations made for the ϕ_{SS} models, Figure 11–14 shows that the CENA constant and magnitude-dependent models are comparable for M less than 5.0. For larger magnitudes, the CENA constant model is larger than the two magnitude-dependent models. For periods longer than 1 sec, all models are magnitude-independent. Figure 11–15 shows the three ϕ models as a function of frequency for **M** = 5.0, 6.0, and 7.0. For **M** = 5.0, the two CENA models have comparable ϕ values. For **M** greater than 5.0, the two CENA models are comparable at low frequencies while the two magnitude-dependent models are comparable at high frequencies.

Figure 11–16 compares the three ϕ models as a function of frequency for CENA to the NGA-West2 ϕ models for **M** = 5.0, 6.0, and 7.0. For the global ϕ model and the NGA-West2 ϕ models, the main difference lies in the difference in ϕ_{S2S} between CENA for hard-rock conditions and WUS for soil and soft-rock sites. Therefore, Figure 11–16 shows that NGA-West2 ϕ is generally larger than that of the global CENA ϕ, particularly at low frequencies. This difference is largely

due to the correction of the CENA ϕ_{S2S} for the small-magnitude bias and for hard-rock conditions.

11.6 Logic Tree for Single-Station Sigma (σ_{SS})

The models for CENA σ_{SS} are derived by combining the models for ϕ_{SS} and τ whereby the mean σ_{SS}^2 and SD(σ_{SS}^2) were derived as follows:

$$\sigma_{SS}^2 = \phi_{SS}^2 + \tau^2 \tag{11–9}$$

$$SD\left(\sigma_{SS}^2\right) = \sqrt{\left[SD\left(\phi_{SS}^2\right)\right]^2 + \left[SD\left(\tau^2\right)\right]^2} \tag{11–10}$$

The equations above assume that the event terms and the single-station within-event residuals are uncorrelated, and that the uncertainties in the estimates of ϕ_{SS}^2 and τ^2 are uncorrelated, which is justified based on the regression results (Section 10.3). A single τ model based on global τ was adopted for CENA while the logic tree for CENA ϕ_{SS} consists of three models (global, CENA constant, and CENA magnitude-dependent). The mean and standard deviations of the ϕ_{SS}^2 and τ^2 are discussed in Sections 10.3.3 and 10.3.2, respectively. The logic tree for CENA σ_{SS} is shown in Figure 11–17. The global σ_{SS} model is derived by combining the global ϕ_{SS} model and the global τ model. The CENA model-1 σ_{SS} is derived by combining the CENA constant ϕ_{SS} model and the global τ model. The CENA model-2 σ_{SS} is the result of combining the CENA magnitude-dependent ϕ_{SS} model and the global τ model. Because the global τ model is magnitude-dependent with four magnitude breaks, all three σ_{SS} models are also magnitude-dependent and have the following form:

$$\sigma_{ss} = \begin{cases} \sigma_{ss_1} & \text{for } \mathbf{M} \le 4.5 \\ \sigma_{ss_1} + \left(\sigma_{ss_2} - \sigma_{ss_1}\right) * \dfrac{\left(\mathbf{M} - 4.5\right)}{0.5} & 4.5 < \mathbf{M} <= 5.0 \\ \sigma_{ss_2} + \left(\sigma_{ss_3} - \sigma_{ss_2}\right) * \dfrac{\left(\mathbf{M} - 5.0\right)}{0.5} & \text{for } 5.0 < \mathbf{M} <= 5.5 \\ \sigma_{ss_3} + \left(\sigma_{ss_4} - \sigma_{ss_3}\right) * \dfrac{\left(\mathbf{M} - 5.5\right)}{1.0} & \text{for } 5.5 < \mathbf{M} <= 6.5 \\ \sigma_{ss_4} & \text{for } \mathbf{M} > 6.5 \end{cases} \tag{11–11}$$

We note that the TI team decided to keep all four magnitude breakpoints in the σ_{SS} models. This is due to the simple implementation of the linear σ_{SS} models despite the multiple breakpoints.

The four coefficients for the central, high, and low branches are given in Tables 11–13, 11–14, and 11–15 for the global model, CENA model-1, and CENA model-2, respectively. Figure 11–18 shows the three σ_{SS} models for CENA as a function of magnitude at PGV and spectral periods of 0.01, 0.1, 0.5, 1, and 5 sec (frequencies of 100, 10, 2, 1, and 0.2 Hz). The solid lines show the median branches for each model, and the dashed lines show the 5th and 95th percentile branches calculated assuming σ_{SS}^2 follows a scaled chi-square distribution. Figure 11–19 shows the three σ_{SS} models as a function of frequency for **M** = 5.0, 6.0, and 7.0. Figures 11–18 and 11–19 indicate that at **M** = 5.0; therefore, the two CENA models are comparable. At larger

magnitudes, the two CENA models are comparable only at low frequencies. At high frequencies, the CENA model-2 σ_{SS} agrees with the global σ_{SS} and is smaller than the σ_{SS} values of CENA model-1. This is the result of the CENA constant ϕ_{SS} model (component of CENA model-1) being larger than the magnitude-dependent ϕ_{SS} models at large magnitudes.

11.6.1 Composite Single-Station Sigma (σ_{SS}) Model

The nine single-station sigma branches discussed above represent the discrete approximations of the three continuous distributions of single-station sigma models. To reduce the computational burden in hazard analyses resulting from the use of nine alternative σ_{SS} values, the three continuous distributions of the single-station sigma models along with their weights are represented by a composite σ_{SS} model. The composite σ_{SS} model is then discretized into a three-point distribution

The cumulative density functions (CDFs) were developed for the three continuous σ_{SS} models at the magnitude breakpoints in the relationship between σ_{SS} and magnitude (**M** = 4.5, 5.0, 5.5, and 6.5) using the mean and standard deviation of σ_{SS}^2 and assuming that σ_{SS}^2 follows a scaled chi-square distribution. Figures 11–20 and 11–21 show examples of the CDFs at **M** = 4.5, 5.0, 5.5, and 6.5 at spectral periods of 0.01 and 1 sec (f = 100 and 1 Hz). The three CDFs were then combined using the weights assigned to the σ_{SS} models to develop a weighted composite CDF. The resulting composite continuous distribution is then represented by three discrete points selected at the 5th, 50th, and 95th percentiles. The composite distributions for σ_{SS} show only a small degree of skewness, and the nominal weights of 0.185, 0.63, and 0.185 on the 5th, 50th, and 95th percentiles, respectively, produce values for the mean and standard deviation of σ_{SS} that adequately represent those computed from the continuous distribution. Tables 11–16 and 11–17 compare the mean and standard deviation for σ_{SS} computed from the continuous and discrete distributions for **M** = 4.5, 5.0, 5.5, and 6.5 at PGV and spectral periods of 0.01, 0.05, 0.1, 0.25, 0.5, 1, 3, 5, and 10 sec. As indicated in the tables, the means are mostly identical to four digits, and the standard deviations from the discrete model are slightly greater by less than 4%.

The composite σ_{SS} model is represented as a function of magnitude with the same functional form [Equation (11–11)] presented above. Table 11–18 presents the coefficients of the three branches (central, high, and low with weights of 0.63, 0.185, and 0.185, respectively) of the composite single-station sigma model for CENA. Figure 11–22 presents a comparison of the composite CENA σ_{SS} model to the Hanford σ_{SS} model (central, high, and low branches) versus frequency for **M** = 5.0, 6.0, and 7.0, and indicates that the Hanford model falls within the range of the CENA σ_{SS} model.

Similarly, Figure 11–23 compares the composite CENA σ_{SS} model to the SWUS σ_{SS} model (central, high, and low branches) versus frequency for **M** = 5.0, 6.0, and 7.0. Figure 11–23 indicates that the CENA and SWUS σ_{SS} models are comparable for **M** = 6.0 and 7.0. For **M** = 5.0, CENA σ_{SS} is larger than that for SWUS, particularly at high frequencies. This is the result of larger CENA τ and ϕ_{SS} at **M** = 5.0 as shown in Figures 11–3 and 11–10, respectively. We note

that the smallest magnitude used for the SWUS model was **M** = 5.0 while the CENA model uses magnitudes as small as **M** = 3.0.

11.7 Logic Tree for Total Ergodic Sigma (σ)

The models for ergodic σ for CENA are derived by combining the models for ϕ_{SS}, ϕ_{S2S}, and τ, whereby the mean σ^2 and SD(σ^2) are derived as follows:

$$\sigma^2 = \sigma_{SS}^2 + \sigma_{S2S}^2 + \tau^2 \tag{11–12}$$

$$SD\left(\sigma^2\right) = \sqrt{\left[SD\left(\phi_{SS}^2\right)\right]^2 + \left[SD\left(\phi_{S2S}^2\right)\right]^2 + \left[SD\left(\tau^2\right)\right]^2} \tag{11–13}$$

The equations above assume that the ground-motion residual components are uncorrelated and that the uncertainties in the estimates of their variances are also uncorrelated, which is justified based on the regression results. The mean and standard deviations of ϕ_{SS}^2, ϕ_{S2S}^2, and τ^2 are discussed in Sections 10.3.3, 10.3.4, and 10.3.2, respectively. The logic tree for ergodic σ for CENA is shown in Figure 11–24. The global σ model is derived by combining the global ϕ_{SS} and τ models with the CENA ϕ_{S2S} model. The CENA model-1 σ is derived by combining the CENA constant ϕ_{SS} model with the CENA ϕ_{S2S} model and the global τ model. The CENA model-2 σ is the result of combining the CENA magnitude-dependent ϕ_{SS} model with the CENA ϕ_{S2S} model and the global τ model. Because the global τ model is magnitude-dependent with four magnitude breaks, all three σ models are also magnitude-dependent and have the following form:

$$\sigma = \begin{cases} \sigma_1 & \text{for } \mathbf{M} \le 4.5 \\ \sigma_1 + \left(\sigma_2 - \sigma_1\right) * \dfrac{\left(\mathbf{M} - 4.5\right)}{0.5} & \text{for } 4.5 < \mathbf{M} <= 5.0 \\ \sigma_2 + \left(\sigma_3 - \sigma_2\right) * \dfrac{\left(\mathbf{M} - 5.0\right)}{0.5} & \text{for } 5.0 < \mathbf{M} <= 5.5 \\ \sigma_3 + \left(\sigma_4 - \sigma_3\right) * \dfrac{\left(\mathbf{M} - 5.5\right)}{1.0} & \text{for } 5.5 < \mathbf{M} <= 6.5 \\ \sigma_4 & \text{for } \mathbf{M} > 6.5 \end{cases} \tag{11–14}$$

The coefficients for the central, high, and low branches are given in Tables 11–19, 11–20, and 11–21 for the global model, CENA model-1, and CENA model-2, respectively. Figure 11–25 shows the three σ models for CENA as a function of magnitude at PGV and spectral periods of 0.01, 0.1, 0.5, 1, and 5 sec. The solid lines show the median branches for each model, and the dashed lines show the 5th and 95th percentile branches calculated assuming σ^2 follows a scaled chi-square distribution. Figure 11–26 shows the three σ models for CENA as a function of frequency for **M** = 5.0, 6.0, and 7.0. Figures 11–25 and 11–26 indicate that for **M** = 5.0, the two CENA models are comparable. At frequencies greater than 2 Hz, all three σ models are similar for **M** = 5.0, with the global model resulting in lower σ values at low frequencies. At **M** = 6.0 and 7.0, the two CENA models are only similar at low frequencies. At high frequencies, the CENA model-2 σ is comparable to the global σ.

11.7.1 Composite Total Ergodic Sigma (σ) Model

The nine ergodic σ branches discussed above represent the discrete approximations of the three continuous distributions of the ergodic σ models. To reduce the computational burden in hazard analyses resulting from the use of nine alternative σ values, the three continuous distributions of the σ models along with their weights are represented by a composite σ model, which is then discretized by three alternative values.

The CDFs were developed for the three continuous σ models at the magnitude breakpoints in the relationship between σ and magnitude (**M** = 4.5, 5.0, 5.5, and 6.5) using the mean and standard deviation of σ^2, and assuming that σ^2 follows a scaled chi-square distribution. Figures 11–27 and 11–28 show examples of the CDFs at **M** = 4.5, 5.0, 5.5, and 6.5 at spectral periods of 0.01, 0.1, and 1 sec. The three CDFs were then combined using the weights assigned to the σ models to develop a weighted composite CDF. The resulting composite continuous distribution is then represented by three discrete points selected at the 5th, 50th, and 95th percentiles. The distributions for σ show only a small degree of skewness, and the nominal weights of 0.185, 0.63, and 0.185 on the 5th, 50th, and 95th percentiles, respectively, produce values for the mean and standard deviation of σ that adequately represent those computed from the continuous distribution. Tables 11–22 and 11–23 compare the mean and standard deviation for σ computed from the continuous and discrete distributions for **M** = 4.5, 5.0, 5.5, and 6.5 at PGV and spectral periods of 0.01, 0.05, 0.1, 0.25, 0.5, 1, 3, 5, and 10 sec. As indicated in the tables, the means are mostly identical to four digits, and the standard deviations from the discrete model are slightly greater by less than 3%.

The composite σ model is represented as a function of magnitude with the same functional form presented above [Equation (11–14)]. Table 11–24 presents the coefficients of the three branches (central, high, and low with weights of 0.63, 0.185, and 0.185, respectively) of the composite ergodic σ model for CENA. Figure 11–29 presents a comparison of the composite ergodic σ model for CENA to the NGA-West2 σ models as a function of frequency for **M** = 5.0, 6.0, and 7.0. Figure 11–29 shows that the composite σ model is generally lower than the NGA-West2 σ models, particularly at low frequencies. This is mainly due to the correction of the ϕ_{S2S} values for CENA to remove the small-magnitude bias and for application to hard-rock conditions, resulting in ϕ_{S2S} values for CENA generally smaller than those for NGA-West2 on softer sites.

Figure 11–30 presents a comparison of the composite ergodic σ model for CENA to the EPRI (2013) σ model for CEUS and the Atkinson and Adams (2013) σ model used in the seismic hazard maps for Canada for **M** = 5.0, 6.0, and 7.0. Figure 11–30 shows that the CENA σ model is generally smaller than the EPRI model (2013) model, which was based on the NGA-West2 σ models, particularly at **M** = 6.0 and 7.0. On the other hand, the CENA σ model is larger than the Atkinson and Adams (2013) model, particularly at high frequencies. This difference between the Atkinson and Adams (2013) model and the CENA model largely reflects greater variability in the range of site conditions and path effects that are included in the CENA model. This is likely due to the fact that the CENA dataset is comprised of ground-motion data from a large geographic region with more variable site and path effects while the Atkinson and Adams (2013) model is

based on a smaller and more uniform dataset in terms of sites and paths (ground-motion data for rock sites in the Charlevoix region). The use of a more uniform dataset, particularly in terms of site conditions, leads to lower ground-motion aleatory variability.

11.8 Application of Standard Deviation Models to Potentially Induced Events

The CENA ϕ_{S2S} model presented in Sections 10.3.4 and 11.4 is based on both potentially induced events (PIEs) and tectonic events in the NGA-East dataset. The models for τ and ϕ_{SS} developed for CENA are based on ground-motion residuals from tectonic events only (Sections 10.3.2 and 10.3.3). The reasons for including PIEs in the dataset used for deriving the CENA ϕ_{S2S} model are mainly related to maximizing the dataset size and the similarities in ϕ_{S2S} between tectonic events plus PIEs and only tectonic events, as discussed in Section 10.3.4. The CENA dataset includes a large number of recordings from PIEs with magnitude greater than 3.0. This section investigates applying the derived CENA standard deviation models for τ and ϕ_{SS} to ground motions from PIEs in CENA.

Figure 11–31 shows a comparison of the τ values obtained from the mixed effects regressions using all earthquakes in the CENA dataset with minimum **M** = 3.0 (54 earthquakes at f = 4 Hz) as well as PIEs only (a total of nine events at f = 4 Hz), and tectonic only events with minimum **M** = 3.0 (45 earthquakes at f = 4 Hz). Figure 11–31 shows that τ for PIEs is smaller than that for tectonic earthquakes in the frequency range of 1 to 10 Hz. The statistical significance of the difference between PIE and tectonic τ for magnitude greater than or equal to 3.0 was evaluated. An F-test of equality of variance was conducted to test against the null hypothesis of equal variance of event terms for PIE and tectonic events. Table 11–25 shows the p-values of the F-test at frequencies of 1 to 10 Hz; the relatively large p-values indicate that the equality of τ^2 for PIE and tectonic events with M larger than or equal to 3.0 cannot be rejected at 5% significance level. As a result, the TI team concluded that, with the current small dataset, the τ model developed for tectonic events in CENA (global τ model) is applicable to PIEs in CENA.

Figure 11–32 compares ϕ_{SS} obtained from the mixed effects regressions using all earthquakes in the CENA dataset with minimum **M** of 3.0 and maximum R_{RUP} of 300 km as documented in Section 10.3.3.2 (708 recordings at f = 4 Hz) as well as PIEs only (315 recordings at f = 4 Hz), and tectonic-only events (393 recordings at f = 4 Hz) within the same magnitude and distance range. Stations with a minimum of three recordings within the magnitude and distance range of interest were used in the regression. Figure 11–32 indicates that ϕ_{SS} values are comparable between tectonic events and PIEs in the frequency range of 1 to 10 Hz. Table 11–26 shows the p-values of the F-test of equality of variances of the single-station within-event residuals for PIEs and tectonic events. Table 11–26 indicates that the equality of ϕ_{SS}^2 for PIEs and tectonic events with **M** larger than or equal to 3.0 and R_{RUP} distance of less than or equal to 300 km cannot be rejected at 5% significance level. As a result, we concluded that the ϕ_{SS} models developed for tectonic events in CENA are applicable to PIEs in CENA. Hazard sensitivity results to the final standard deviation logic tree are presented in Appendix F.2.

11.9 Application of Standard Deviation Models to Gulf Coast Region

Ground-motion data from the Gulf Coast Region (GCR) were not used in developing the standard deviation models for CENA (Section 10.3.1). The GCR was treated separately in the NGA-East project, and adjustment factors to the NGA-East median GMMs (PEER 2015a) were developed to make them applicable to the GCR. This section analyzes the standard deviations of the residuals of the Gulf Coast data with respect to the two PEER-developed GMMs for the GCR and evaluates the applicability of the developed standard deviation models for the rest of CENA to the GCR. The median GMMs for the GCR are described in PEER (2015b) and in Chapter 13, and are different than those presented in Chapter 9 of this report for the rest of CENA.

The Gulf Coast data used to develop the median GMMs was documented in PEER (2015b) and consist of eight tectonic events and one PIE with **M** between 3.4 and 4.7. We note that the PIE event with **M** = 4.7 has a large negative event term at short periods compared to the tectonic events. Figure 11–33 compares the global τ model for **M** = 5.0 to the τ values for the tectonic events in the GCR; the CENA constant τ model is also included in the plot. Figure 11–33 indicates that the τ values for the GCR are within the range of the CENA constant model except for frequencies of 3 to 4 Hz. The uncertainty in the estimates of τ for the GCR is largely due to the small number of events available for the analysis. Table 11–27 presents the p-values from the F-test of equality of variance of the event terms for the GCR and the tectonic events in the rest of CENA with **M** = 3.0 to 5.0. Table 11–27 indicates that the equality of τ^2 for the GCR and the rest of CENA cannot be rejected for most of the frequencies between 1 and 10 Hz for a 5% significance level. Based on these p-values and since the available data from the GCR are insufficient to reliably develop a GCR-specific τ model, The TI team considers the global τ model to be applicable to the GCR.

Figure 11–34 compares the global ϕ_{SS} model for **M** = 5.0 and the CENA constant ϕ_{SS} model to the ϕ_{SS} values for the GCR using tectonic data and consisting of only 132 recordings at f = 4 Hz. The datasets used to derive the global and CENA constant ϕ_{SS} models were discussed in Chapter 10. Figure 11–34 shows that the ϕ_{SS} values for the GCR are generally within the range of the ϕ_{SS} models adopted for CENA. Table 11–28 shows the p-values of the F-test of equality of ϕ_{SS}^2 for the GCR and the rest of CENA using tectonic data with **M** = 3.0 to 5.0 and maximum R_{RUP} distance of 400 km (comparable **M** and R_{RUP} ranges for the GCR and the rest of CENA). Table 11–28 indicates that the equality of ϕ_{SS}^2 cannot be rejected at a 5% significance level when tested period by period. As a result, the TI team considers the CENA ϕ_{SS} models applicable to the GCR.

Figure 11–35 compares the CENA ϕ_{S2S} model to the ϕ_{S2S} values obtained for the GCR. We note that the CENA ϕ_{S2S} model shown in Figure 11–35 is the derived CENA model before the application of the adjustment factors to correct for the small-magnitude bias and for application to hard-rock conditions. The uncorrected CENA ϕ_{S2S} model was used for the Gulf Coast comparisons because it is more compatible with the Gulf Coast data. Figure 11–35 indicates that the ϕ_{S2S} values for the GCR are lower than those for the rest of CENA and have large error bars. We note that the GCR dataset used to estimate ϕ_{S2S} consisted of a total of 20 stations

compared to 275 stations used for the CENA ϕ_{S2S} model at f = 4 Hz. Table 11−29 presents the p-values of the F-test of equality of ϕ_{S2S}^2 for the GCR and the rest of CENA. Values in red indicate cases where the equality of ϕ_{S2S}^2 is rejected at a 5% significance level at frequencies of 6.67 and 10 Hz. For the rest of the frequencies between 1 and 6.67 Hz, the equality of variances is not rejected at 5% significance level. Given that the equality of ϕ_{S2S}^2 cannot be rejected at most of the frequencies between 1 and 10 Hz and because the small dataset for the GCR does not allow the development of a reliable GCR-specific ϕ_{S2S} model, the TI team adopts the CENA ϕ_{S2S} model for application to the GCR. Moreover, the CENA ϕ_{S2S} model adjusted for the small-magnitude bias and designated for application to hard-rock conditions should be applied for the GCR.

Based on the discussions presented in this section, the TI team considers the models developed for CENA for τ (global τ model), ϕ_{SS} (global and CENA constant and magnitude-dependent ϕ_{SS} models), and ϕ_{S2S} (CENA ϕ_{S2S} model) to be applicable for the GCR. As a result, the TI team considers the CENA single-station sigma and total ergodic sigma models presented in Sections 11.6 and 11.7, respectively, applicable to the GCR.

11.10 Implementation of the Mixture Model

In probabilistic seismic hazard analysis (PSHA), it is typically assumed that the distribution of ground motions follows a lognormal distribution defined by the median and standard deviation of the lognormal of the spectral accelerations. Recent studies involving large datasets—such as the NGA-West2 data (GeoPentech 2015), a Japanese dataset of crustal earthquakes, and a global subduction dataset (Coppersmith et al. 2014) —observed that at large deviations from the mean (i.e., the tails of the distribution), ground motions deviate from the lognormal distribution.

The SWUS project analyzed the within-event residuals, between-event residuals, and single station within-event residuals of the ASK14 and CY14 subsets of the NGA-West2 data. The residuals were shown in terms of quantile plots (Q-Q plots) to compare the observed data density at various values of epsilon (residuals normalized by the corresponding standard deviation) to the assumed normal distribution. The SWUS project observed a deviation from normality for the ASK14 data for a wide range of oscillator periods and for epsilon values larger than 2. This deviation from normality was observed in the analysis of the within-event residuals and single-station within-event residuals, but not for the between-event residuals. Similar observations were made using the CY14 residuals. The SWUS project modeled the heavy tail of the aleatory variability distribution using a mixture model. In general, a mixture model is a composite distribution summing multiple normal distributions with different means and standard deviations. For the SWUS project, the derived mixture model consisted of an equally weighted mixture of two normal distributions with means of zero and with standard deviations of 0.8 and 1.2 times the standard deviation of the within-event residuals. The SWUS project used two alternatives to represent the shape of the distribution of ground-motion residuals: a traditional lognormal distribution with a weight of 0.2 and a mixture of two normal distributions with weight of 0.8.

For the SWUS project, the mixture model was highly weighted because of the observed evidence of heavy tails in the tested within-event residuals and the good fit of the mixture model to the distribution of the within-event residuals. The SWUS project maintained the traditional lognormal distribution with a small assigned weight because this distribution is most widely used in current practice. The mixture model developed based on the SWUS analysis of the within-event residuals was applied to the single-station and ergodic standard deviations. GeoPentech (2015) noted that the application of the mixture model to σ_{SS} and σ may be conservative given that the between-event residuals did not demonstrate the same deviation from normality as the within-event and single-station within-event residuals. The application of the mixture model to σ_{SS} and σ was considered appropriate in the SWUS project because the values of ϕ_{SS} and ϕ are larger than those of τ, which makes the tail of the total residual distribution dominated by ϕ_{SS} and ϕ for the single-station sigma and ergodic sigma, respectively. Moreover, the smaller sample size for τ compared to the sample sizes for ϕ_{SS} and ϕ (i.e., a smaller number of events than recordings in datasets) limited the ability to evaluate the shape of the tails in the distribution of event terms.

The Hanford project (Coppersmith et al. 2014) evaluated the deviation of the tails of the ground-motion residuals from normality using: ASK14 residuals, CY14 residuals, KiK-net dataset of crustal earthquakes (Dawood et al. 2016), and a global dataset of subduction earthquakes compiled for the Hanford project (Coppersmith et al. 2014). For the ASK14, CY14, and KiK-net residuals, the within-event residuals and the single-station within-event residuals were observed to deviate from normality at epsilons greater than 2 while the between-event residuals did not show such heavy tails. For the subduction dataset, both within-event and between-event residuals deviated from normality at epsilon values around 2 and 3.5, respectively. Similar to the SWUS project, the Hanford project adopted the lognormal distribution with a weight of 0.2 and the mixture model with a weight of 0.8 to represent the shape of ground-motion distribution. The mixture model for the Hanford project was calibrated using the ASK14 residuals and consisted of an equally weighted mixture of two normal distributions with means of zero and with standard deviations of 0.8 and 1.2 times σ_{SS}.

The limited CENA dataset consisting primarily of small-to-moderate magnitude earthquakes does not allow the evaluation of the distribution of the different components of ground-motion residuals in the extreme tails. Moreover, because ground-motion residuals were observed to deviate from normality in the extreme tails in more than one dataset and for different tectonic environments (e.g., the ASK14, CY14, KiK-net dataset of crustal earthquakes, and a global subduction dataset), the TI team decided to adopt the logic tree used in the SWUS and the Hanford projects to represent the shape of the distribution of ground-motion residuals. We note that both ASK14 and CY14 residuals, which showed a deviation of the distribution of within-event and single-station within-event residuals from normality for large ground motions, were used in developing the global τ model with an assigned weight of 1.0 as well as the global ϕ_{SS} model with an assigned weight of 0.8 in the NGA-East study. Therefore, the mixture model developed in the SWUS and the Hanford projects is considered applicable to the NGA-East project.

The final sigma logic tree applicable to both single-station sigma and ergodic sigma is shown in Figure 11–36. The TI team adopted the observations made in the SWUS and the Hanford projects that the mixture of two normal distributions better fits the distribution of ground-motion residuals, particularly for large ground motions. We adopted two alternatives to represent the shape of ground-motion residuals: traditional lognormal distribution with a weight of 0.2 and a mixture of two normal distributions with weight of 0.8. For the single-station sigma, the mixture model consists of an equally weighted mixture of two normal distributions, with one component having 0.8 times the normal σ_{SS} and the other component having 1.2 times the normal σ_{SS}. For the total ergodic sigma, the mixture model consists of an equally weighted mixture with one component having 0.8 times the normal σ and one component having 1.2 times the normal σ. For the mixture model, the conditional probability of the ground motion, Sa, exceeding a ground-motion level, z, is given by:

$$P(Sa > z) = w_{Mix1}\left\{1-\Phi\left[\frac{\ln(z)-\ln(\mu)}{\sigma_{Mix1}}\right]\right\} + w_{Mix2}\left\{1-\Phi\left[\frac{\ln(z)-\ln(\mu)}{\sigma_{Mix2}}\right]\right\} \qquad (11\text{–}15)$$

where Φ is the standard normal cumulative distribution function, w_{Mix1} and w_{Mix2} are the weights of the two normal distributions (both equal to 0.5 in this case), μ is the median ground motion, and σ_{Mix1} and σ_{Mix2} are the standard deviations of the two normal distributions. For the case of ergodic sigma, σ_{Mix1} and σ_{Mix2} are equal to 1.2σ and 0.8σ, respectively, with σ calculated as described in Section 11.7.1. For the single-station sigma, σ_{Mix1} and σ_{Mix2} are equal to $1.2\sigma_{SS}$ and $0.8\sigma_{SS}$, respectively, with σ_{SS} calculated as described in Section 11.6.1.

Appendix F.2 shows the results of hazard sensitivities to the three branches (low, central, and high) of the composite single-station sigma model as well as to the distribution of the single-station sigma model. Hazard sensitivities were performed for the Central Illinois and Savannah test sites, and hazard curves are shown for PGA and frequencies of 0.1 to 50 Hz. The hazard plots in Appendix F.2 show small to negligible difference in the hazard due to the distribution of the single-station sigma model (traditional lognormal versus mixture model). Despite the small to negligible observed impact of the distribution of ground-motion residuals (normal versus mixture model) on the hazard results presented in Appendix F.2, the TI team considers the logic tree for the distribution of ground-motion residuals warranted given the evidence of deviation from normal distribution at large ground motions observed in multiple datasets and the relatively simple application of the mixture model.

11.11 References

Abrahamson, N A., W.J. Silva, and R. Kamai. 2014. “Summary of the ASK14 Ground-Motion Relation for Active Crustal Regions.” Earthq. Spectra, 30(3): 1025–1055.

Al Atik, L. 2015. “NGA-East: Ground Motion Standard Deviation Models for Central and Eastern North America.” PEER Report No. PEER 2015/07, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Ancheta, T.D., R.B. Darragh, J.P. Stewart, E. Seyhan, W.J. Silva, B.S.-J. Chiou, K.E. Wooddell, R.W. Graves, A.R. Kottke, D.M. Boore, T. Kishida, and J.L. Donahue. 2014. “NGA-West2 Database.” Earthq. Spectra, 30: 989–1005.

Ang A.H.-S., and W.H. Tang. (2007). Probability Concepts In Engineering: Emphasis on Applications to Civil and Environmental Engineering (v.1). John Wiley and Sons, Chichester U.K.

Atkinson, G.M., and J. Adams. 2013. “Ground Motion Prediction Equations for Application to the 2015 National Seismic Hazard Maps of Canada.” Canadian J. Civil Eng., 40: 988–998.

Boore, D.M., J.P. Stewart, E. Seyhan, and G.M. Atkinson. 2014. “NGA-West2 Equations For Predicting Response Spectral Accelerations for Shallow Crustal Earthquakes.” Earthq. Spectra, 30(3): 1057–1085.

Campbell, K.W., and Y. Bozorgnia. 2014. “NGA-West2 Campbell-Bozorgnia Ground Motion Model for the Horizontal Components of PGA, PGV, and 5%-Damped Elastic Pseudo-Acceleration Response Spectra for Periods Ranging from 0.01 to 10 sec.” Earthq. Spectra, 30(3): 1087–1115.

Chiou, B.S.-J., and R.R. Youngs. 2014. “Update of the Chiou and Youngs NGA Ground Motion Model for Average Horizontal Component of Peak Ground Motion and Response Spectra.” Earthq. Spectra. 30(3): 1117–1153.

Coppersmith, K.J., J.J. Bommer,, K. Hanson,, R. Coppersmith,, J.R. Unruh,, L. Wolf,, R.R. Youngs, L. Al Atik, A. Rodriguez-Marek, G. Toro, and V. Montaldo-Falero. 2014. “Hanford Sitewide Probabilistic Seismic Hazard Analysis,” PNNL-23361, Pacific Northwest National Laboratory, Richland, WA.

Dawood, H.M. 2014. Partitioning uncertainty for non-ergodic probabilistic seismic hazard analyses. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Dawood H.M., A. Rodriguez-Marek, J. Bayless, C. Goulet, and E. Thompson. 2016. “A Flatfile for the KiK-Net Database Processed using an Automated Protocol.” Earthq. Spectra, 32(2): 1281–1302.

EPRI. 2013. “Ground-Motion Model (GMM) Review Project.” Electrical Power Research Institute, EPRI Report 3002000717, Palo Alto, CA.

GeoPentech. 2015. “Southwestern United States Ground Motion Characterization SSHAC Level 3 – Technical Report Rev.2.” March 2015.

Goulet, C.A., T. Kishida. T.D. Ancheta, C.H. Cramer, R.B. Darragh, W.J. Silva, Y.M.A. Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. 2014. “PEER NGA-East Database.” PEER Report No. 2014/17, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Keefer, D.L., and S.E. Bodily. 1983. “Three-Point Approximations for Continuous Random Variables.” Management Sci., 29: 595–609.

PEER. 2015a. “NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

PEER. 2015b. “NGA-East: Adjustments to Median Ground-Motion Models for Central and Eastern North America.” PEER Report No. 2015/08, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Renault, P., S. Heuberger, and N.A. Abrahamson. 2010. “PEGASOS Refinement Project: An Improved PSHA for Swiss Nuclear Power Plants.” Proceedings, 14th European Conference of Earthquake Engineering, Paper ID 991, Ohrid, Republic of Macedonia.

Rodriguez-Marek, A., personal communication.

Zhao J.X., J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H.K. Thio, P.G. Somerville, and Y. Fukushima. 2006. “Attenuation Relations of Strong Ground Motion in Japan using Site Classification based on Predominant Period.” Bull. Seismol. Soc. Am., 96(3): 898–913.

Table 11–1 Summary of basic features and datasets used in the development of the candidate CENA aleatory variability models and other exisiting models.

	τ	ϕ_{SS}	ϕ_{S2S}	ϕ
NGA-East	- Global Model: **M**- dependent; T-independent; based on NGA-West2 τ models - CENA constant Model: **M**- and T-independent; NGA-East dataset - CENA **M**-dependent Model: **M**-dependent; T-independent; NGA-East dataset; **M**-dependence based on NGA-West2 models	- Global Model: **M**- and T-dependent; NGA-West2 dataset - CENA constant Model: **M**- and T-independent; NGA-East dataset - CENA **M**-dependent Model: **M**-dependent; T-dependent; NGA-East dataset; **M**-dependence based on NGA-West2 models	- CENA Model: T-dependent, NGA-East dataset; extrapolation outside of 1 to 10 Hz based on Japanese dataset	Based on combining ϕ_{SS} and ϕ_{S2S} models; NGA-East and NGA-West2 datasets
EPRI (2013)	Based on NGA-West2 τ models; average NGA-West2 τ increased by 0.03; τ between 10 and 40 Hz set equal to the value at 10 Hz			Based on NGA-West2 ϕ models; ϕ between 10 and 40 Hz set equal to the value at 10 Hz
Atkinson and Adams (2013)	Based on previous studies such as NGA-West1 and Atkinson (2011) and general observations that assigned aleatory variability should be less than indicated by regressions statistics			Based on previous studies such as NGA-West1 and Atkinson (2011) and general observations that assigned aleatory variability should be less than indicated by regressions statistics
ASK14	**M**-dependent and T-independent; based on NGA-West2 dataset			**M**- and T-dependent; based on NGA-West2 dataset
BSSA14	**M**- and T-dependent; based on NGA-West2 dataset			**M**- and T-dependent; based on NGA-West2 dataset
CB14	**M**- and T-dependent; based on NGA-West2 dataset			**M**- and T-dependent; based on NGA-West2 dataset
CY14	**M**- and T-dependent; based on NGA-West2 dataset			**M**- and T-dependent; based on NGA-West2 dataset
Southwestern U.S. Ground Motion Characterization Project	**M**-dependent; T-independent; NGA-West2 dataset and Zhao et al. (2006) dataset	Multiple models; **M**-dependent and **M**-independent models; T-dependent and T-independent models; based on European dataset, Taiwanese dataset, and NGA-West2 dataset (California only and global dataset)		
Hanford Project (custal models)	**M**- and T-dependent model based on NGA-West2 τ models	**M**- and T-dependent model based on the global dataset compiled for the PEGASOS Refinement project		

Table 11–2 P-values from the F-test of equality of τ^2 for CENA and the NGA-West2 GMMs. NGA-West2 τ is calculated based on event terms (option 1). Values in red show cases where the null hypothesis(1) is rejected at 5% significance level.

Period (sec)	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.75	1.00	PGV
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00	
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

(2) Null and alternate hypotheses are: $H_0 : \tau^2_{CENA} = \tau^2_{NGA-West2}$ and $H_a : \tau^2_{CENA} \neq \tau^2_{NGA-West2}$, respectively.

Table 11–3 P-values from the F-test of equality of τ^2 for CENA and the NGA-West2 GMMs. NGA-West2 τ is calculated based on the proposed models (option 2). Values in red show cases where the null hypothesis(1) is rejected at 5% significance level.

Period (sec)	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.75	1.00	PGV
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00	
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

(1) Null and alternate hypotheses are: $H_0 : \tau^2_{CENA} = \tau^2_{NGA-West2}$ and $H_a : \tau^2_{CENA} \neq \tau^2_{NGA-West2}$, respectively.

Table 11–4 Coefficients of the global Tau (τ) model.

		τ_1	τ_2	τ_3	τ_4
Central	f = 0.1 to 100 Hz	0.4436	0.4169	0.3736	0.3415
	PGA	0.4436	0.4169	0.3736	0.3415
	PGV	0.3633	0.3532	0.3340	0.3136
Low	f = 0.1 to 100 Hz	0.3280	0.2928	0.2439	0.2343
	PGA	0.3280	0.2928	0.2439	0.2343
	PGV	0.2488	0.2370	0.2278	0.2081
High	f = 0.1 to 100 Hz	0.5706	0.5551	0.5214	0.4618
	PGA	0.5706	0.5551	0.5214	0.4618
	PGV	0.4919	0.4845	0.4535	0.4333

Table 11–5 P-values from the F-test of equality of ϕ_{SS}^2 for CENA and the NGA-West2 GMMs. Values in red show cases where the null hypothesis(1) is rejected at 5% significance level.

Period (sec)	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.75	1.00	PGV
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00	
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

(1) Null and alternate hypotheses are: $H_0 : \phi^2_{SS_{CENA}} = \phi^2_{SS_{NGA-West2}}$ and $H_a : \phi^2_{SS_{CENA}} \neq \phi^2_{SS_{NGA-West2}}$, respectively

.

Table 11–6 Coefficients of the global PhiSS (ϕ_{SS}) model.

Period (sec)	Frequency (Hz)	Central		Global ϕ_{SS} Model			
				High		Low	
		a	b	a	b	a	b
0.01	100	0.5423	0.3439	0.6553	0.4446	0.4367	0.2525
0.02	50	0.5410	0.3438	0.6537	0.4452	0.4357	0.2518
0.03	33.33	0.5397	0.3437	0.6521	0.4459	0.4347	0.2510
0.04	25	0.5382	0.3436	0.6503	0.4466	0.4334	0.2503
0.05	20	0.5371	0.3435	0.6489	0.4473	0.4326	0.2496
0.075	13.33	0.5339	0.3433	0.6450	0.4489	0.4301	0.2478
0.1	10	0.5308	0.3431	0.6412	0.4505	0.4277	0.2461
0.15	6.67	0.5247	0.3466	0.6338	0.4561	0.4229	0.2478
0.2	5	0.5189	0.3585	0.6266	0.4673	0.4182	0.2600
0.25	4	0.5132	0.3694	0.6196	0.4776	0.4137	0.2712
0.3	3.33	0.5077	0.3808	0.6129	0.4879	0.4093	0.2831
0.4	2.5	0.4973	0.4004	0.6002	0.5057	0.4010	0.3037
0.5	2	0.4875	0.4109	0.5884	0.5161	0.3932	0.3142
0.75	1.33	0.4658	0.4218	0.5622	0.5264	0.3757	0.3253
1	1	0.4475	0.4201	0.5403	0.5217	0.3607	0.3263
1.5	0.67	0.4188	0.4097	0.5068	0.4985	0.3367	0.3271
2	0.5	0.3984	0.3986	0.4836	0.4818	0.3189	0.3208
3	0.33	0.3733	0.3734	0.4565	0.4556	0.2958	0.2969
4	0.25	0.3604	0.3604	0.4436	0.4437	0.2832	0.2831
5	0.2	0.3538	0.3537	0.4374	0.4381	0.2764	0.2757
7.5	0.13	0.3482	0.3481	0.4325	0.4337	0.2703	0.2691
10	0.1	0.3472	0.3471	0.4317	0.4329	0.2692	0.2679
PGA		0.5423	0.3439	0.6553	0.4446	0.4367	0.2525
PGV		0.4985	0.3548	0.6010	0.4296	0.4027	0.2850

Table 11–7 Coefficients of the CENA constant PhiSS (ϕ_{SS}) model.

Period (sec)	Frequency (Hz)	CENA Constant ϕ_{SS} Model		
		Central	High	Low
0.01	100	0.5076	0.6192	0.4037
0.02	50	0.5076	0.6192	0.4037
0.03	33.33	0.5076	0.6192	0.4037
0.04	25	0.5076	0.6192	0.4037
0.05	20	0.5076	0.6192	0.4037
0.075	13.33	0.5076	0.6192	0.4037
0.1	10	0.5076	0.6192	0.4037
0.15	6.67	0.5076	0.6192	0.4037
0.2	5	0.5076	0.6192	0.4037
0.25	4	0.5076	0.6192	0.4037
0.3	3.33	0.5076	0.6192	0.4037
0.4	2.5	0.5076	0.6192	0.4037
0.5	2	0.5076	0.6192	0.4037
0.75	1.33	0.5076	0.6192	0.4037
1	1	0.5076	0.6192	0.4037
1.5	0.67	0.5076	0.6192	0.4037
2	0.5	0.5076	0.6192	0.4037
3	0.33	0.5076	0.6192	0.4037
4	0.25	0.5076	0.6192	0.4037
5	0.2	0.5076	0.6192	0.4037
7.5	0.13	0.5076	0.6192	0.4037
10	0.1	0.5076	0.6192	0.4037
PGA		0.5076	0.6192	0.4037
PGV		0.5461	0.6502	0.4483

Table 11–8 Coefficients of the CENA magnitude-dependent PhiSS (ϕ_{SS}) model.

Period (sec)	Frequency (Hz)	CENA Magnitude-Dependent ϕ_{SS} Model					
		Central		High		Low	
		a	b	a	b	a	b
0.01	100	0.5135	0.3263	0.6267	0.4198	0.4081	0.2412
0.02	50	0.5135	0.3271	0.6267	0.4206	0.4081	0.2420
0.03	33.33	0.5135	0.3279	0.6267	0.4215	0.4081	0.2427
0.04	25	0.5135	0.3288	0.6267	0.4224	0.4081	0.2436
0.05	20	0.5135	0.3296	0.6267	0.4231	0.4081	0.2443
0.075	13.33	0.5135	0.3316	0.6267	0.4252	0.4081	0.2463
0.1	10	0.5135	0.3336	0.6267	0.4272	0.4081	0.2482
0.15	6.67	0.5135	0.3413	0.6267	0.4351	0.4081	0.2555
0.2	5	0.5135	0.3569	0.6267	0.4514	0.4081	0.2702
0.25	4	0.5135	0.3717	0.6267	0.4671	0.4081	0.2839
0.3	3.33	0.5135	0.3870	0.6267	0.4837	0.4081	0.2979
0.4	2.5	0.5135	0.4150	0.6267	0.5145	0.4081	0.3230
0.5	2	0.5135	0.4344	0.6267	0.5362	0.4081	0.3401
0.75	1.33	0.5135	0.4665	0.6267	0.5726	0.4081	0.3680
1	1	0.5135	0.4836	0.6267	0.5922	0.4081	0.3827
1.5	0.67	0.5135	0.5026	0.6267	0.6141	0.4081	0.3988
2	0.5	0.5135	0.5135	0.6267	0.6267	0.4081	0.4081
3	0.33	0.5135	0.5135	0.6267	0.6267	0.4081	0.4081
4	0.25	0.5135	0.5135	0.6267	0.6267	0.4081	0.4081
5	0.2	0.5135	0.5135	0.6267	0.6267	0.4081	0.4081
7.5	0.13	0.5135	0.5135	0.6267	0.6267	0.4081	0.4081
10	0.1	0.5135	0.5135	0.6267	0.6267	0.4081	0.4081
PGA		0.5135	0.3263	0.6267	0.4198	0.4081	0.2412
PGV		0.5575	0.3957	0.6789	0.4950	0.4445	0.3041

Table 11–9 CENA PhiS2S (ϕ_{S2S}) model.

Period (sec)	Frequency (Hz)	CENA ϕ_{S2S}		
		Central	High	Low
0.01	100	0.2936	0.3697	0.2237
0.02	50	0.2935	0.3696	0.2235
0.03	33.33	0.3122	0.3843	0.2454
0.04	25	0.3456	0.4161	0.2797
0.05	20	0.3708	0.4440	0.3022
0.075	13.33	0.3653	0.4512	0.2857
0.1	10	0.3335	0.4318	0.2442
0.15	6.67	0.3110	0.4089	0.2227
0.2	5	0.3099	0.3994	0.2285
0.25	4	0.3103	0.3925	0.2349
0.3	3.33	0.3097	0.3854	0.2398
0.4	2.5	0.3063	0.3711	0.2458
0.5	2	0.3010	0.3591	0.2464
0.75	1.33	0.2857	0.3365	0.2378
1	1	0.2707	0.3229	0.2216
1.5	0.67	0.2452	0.3053	0.1896
2	0.5	0.2246	0.2968	0.1596
3	0.33	0.1937	0.2872	0.1140
4	0.25	0.1706	0.2828	0.0808
5	0.2	0.1518	0.2808	0.0560
7.5	0.13	0.1164	0.2769	0.0212
10	0.1	0.0934	0.2700	0.0088
PGA		0.2936	0.3697	0.2237
PGV		0.3104	0.3740	0.2510

Table 11–10 Coefficients of the global Phi (ϕ) model.

Period (sec)	Frequency (Hz)	Global ϕ Model					
		Central		High		Low	
		a	b	a	b	a	b
0.01	100	0.6194	0.4560	0.7245	0.5463	0.5199	0.3714
0.02	50	0.6182	0.4559	0.7230	0.5466	0.5190	0.3708
0.03	33.33	0.6261	0.4680	0.7292	0.5571	0.5283	0.3844
0.04	25	0.6422	0.4910	0.7433	0.5784	0.5462	0.4086
0.05	20	0.6554	0.5092	0.7557	0.5970	0.5599	0.4261
0.075	13.33	0.6501	0.5055	0.7531	0.6003	0.5523	0.4163
0.1	10	0.6306	0.4834	0.7369	0.5854	0.5300	0.3881
0.15	6.67	0.6137	0.4708	0.7192	0.5742	0.5140	0.3745
0.2	5	0.6077	0.4784	0.7102	0.5784	0.5108	0.3851
0.25	4	0.6027	0.4865	0.7026	0.5839	0.5082	0.3954
0.3	3.33	0.5975	0.4945	0.6949	0.5896	0.5051	0.4053
0.4	2.5	0.5864	0.5071	0.6798	0.5989	0.4976	0.4206
0.5	2	0.5751	0.5121	0.6656	0.6030	0.4890	0.4264
0.75	1.33	0.5483	0.5118	0.6342	0.6024	0.4667	0.4263
1	1	0.5249	0.5021	0.6083	0.5914	0.4456	0.4177
1.5	0.67	0.4875	0.4798	0.5688	0.5614	0.4106	0.4025
2	0.5	0.4601	0.4603	0.5418	0.5404	0.3831	0.3846
3	0.33	0.4248	0.4249	0.5097	0.5089	0.3453	0.3461
4	0.25	0.4047	0.4047	0.4935	0.4935	0.3220	0.3220
5	0.2	0.3926	0.3926	0.4850	0.4855	0.3071	0.3066
7.5	0.13	0.3781	0.3780	0.4759	0.4768	0.2882	0.2873
10	0.1	0.3720	0.3718	0.4714	0.4722	0.2809	0.2799
PGA		0.6194	0.4560	0.7245	0.5463	0.5199	0.3714
PGV		0.5896	0.4738	0.6823	0.5434	0.5016	0.4074

Table 11–11 Coefficients of the CENA magnitude-dependent Phi (ϕ) model.

Period (sec)	Frequency (Hz)	CENA Magnitude-Dependent ϕ Model					
		Central		High		Low	
		a	b	a	b	a	b
0.01	100	0.5944	0.4426	0.6990	0.5279	0.4956	0.3625
0.02	50	0.5943	0.4431	0.6990	0.5285	0.4955	0.3629
0.03	33.33	0.6037	0.4562	0.7068	0.5395	0.5062	0.3778
0.04	25	0.6217	0.4804	0.7229	0.5619	0.5258	0.4033
0.05	20	0.6363	0.4994	0.7369	0.5815	0.5407	0.4216
0.075	13.33	0.6335	0.4971	0.7376	0.5857	0.5349	0.4134
0.1	10	0.6162	0.4761	0.7244	0.5711	0.5140	0.3870
0.15	6.67	0.6042	0.4662	0.7127	0.5609	0.5019	0.3776
0.2	5	0.6033	0.4767	0.7099	0.5681	0.5027	0.3908
0.25	4	0.6031	0.4878	0.7082	0.5772	0.5039	0.4036
0.3	3.33	0.6026	0.4989	0.7064	0.5873	0.5044	0.4156
0.4	2.5	0.6004	0.5185	0.7025	0.6067	0.5037	0.4351
0.5	2	0.5974	0.5309	0.6989	0.6203	0.5015	0.4464
0.75	1.33	0.5896	0.5492	0.6911	0.6429	0.4937	0.4605
1	1	0.5824	0.5562	0.6850	0.6538	0.4855	0.4640
1.5	0.67	0.5711	0.5613	0.6761	0.6645	0.4724	0.4642
2	0.5	0.5630	0.5630	0.6702	0.6702	0.4624	0.4624
3	0.33	0.5525	0.5525	0.6628	0.6628	0.4491	0.4491
4	0.25	0.5459	0.5459	0.6584	0.6584	0.4408	0.4408
5	0.2	0.5415	0.5415	0.6556	0.6556	0.4350	0.4350
7.5	0.13	0.5349	0.5349	0.6515	0.6515	0.4264	0.4264
10	0.1	0.5313	0.5313	0.6486	0.6486	0.4221	0.4221
PGA		0.5944	0.4426	0.6990	0.5279	0.4956	0.3625
PGV		0.6405	0.5058	0.7504	0.5926	0.5366	0.4237

Table 11–12 Coefficients of the CENA constant Phi (ϕ) model.

Period (sec)	Frequency (Hz)	CENA Constant ϕ Model		
		Central	High	Low
0.01	100	0.5893	0.6925	0.4918
0.02	50	0.5892	0.6924	0.4917
0.03	33.33	0.5987	0.7003	0.5025
0.04	25	0.6168	0.7166	0.5222
0.05	20	0.6315	0.7307	0.5372
0.075	13.33	0.6287	0.7315	0.5312
0.1	10	0.6112	0.7182	0.5102
0.15	6.67	0.5992	0.7064	0.4981
0.2	5	0.5982	0.7035	0.4988
0.25	4	0.5981	0.7018	0.5001
0.3	3.33	0.5975	0.7000	0.5006
0.4	2.5	0.5953	0.6960	0.5000
0.5	2	0.5923	0.6923	0.4978
0.75	1.33	0.5844	0.6843	0.4899
1	1	0.5772	0.6782	0.4818
1.5	0.67	0.5658	0.6692	0.4685
2	0.5	0.5576	0.6632	0.4584
3	0.33	0.5470	0.6558	0.4450
4	0.25	0.5404	0.6514	0.4365
5	0.2	0.5359	0.6486	0.4307
7.5	0.13	0.5292	0.6444	0.4219
10	0.1	0.5255	0.6416	0.4176
PGA		0.5893	0.6925	0.4918
PGV		0.6303	0.7255	0.5396

Table 11–13 Coefficients of the global single-station sigma (σ_{SS}) model.

Period (sec)	Frequency (Hz)	Global σ_{SS} Model											
		Central				High				Low			
		σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}
0.01	100	0.7054	0.6895	0.6122	0.4903	0.8232	0.8114	0.7322	0.6000	0.5939	0.5744	0.4998	0.3884
0.02	50	0.7044	0.6884	0.6115	0.4903	0.8220	0.8102	0.7315	0.6003	0.5930	0.5735	0.4990	0.3880
0.03	33.33	0.7034	0.6874	0.6109	0.4902	0.8208	0.8090	0.7309	0.6006	0.5922	0.5726	0.4983	0.3877
0.04	25	0.7023	0.6862	0.6101	0.4902	0.8196	0.8077	0.7301	0.6009	0.5912	0.5716	0.4975	0.3873
0.05	20	0.7014	0.6853	0.6095	0.4902	0.8185	0.8067	0.7296	0.6013	0.5905	0.5708	0.4969	0.3870
0.075	13.33	0.6989	0.6828	0.6078	0.4901	0.8157	0.8039	0.7280	0.6020	0.5884	0.5687	0.4951	0.3862
0.1	10	0.6965	0.6804	0.6061	0.4900	0.8130	0.8011	0.7265	0.6028	0.5863	0.5665	0.4934	0.3854
0.15	6.67	0.6919	0.6756	0.6039	0.4926	0.8077	0.7957	0.7244	0.6063	0.5822	0.5623	0.4911	0.3871
0.2	5	0.6874	0.6710	0.6038	0.5010	0.8026	0.7906	0.7243	0.6142	0.5783	0.5582	0.4910	0.3959
0.25	4	0.6831	0.6666	0.6036	0.5089	0.7977	0.7856	0.7241	0.6216	0.5745	0.5543	0.4908	0.4041
0.3	3.33	0.6789	0.6623	0.6036	0.5172	0.7930	0.7809	0.7241	0.6292	0.5709	0.5505	0.4908	0.4128
0.4	2.5	0.6710	0.6542	0.6031	0.5317	0.7843	0.7720	0.7236	0.6425	0.5639	0.5432	0.4902	0.4282
0.5	2	0.6638	0.6468	0.6006	0.5398	0.7762	0.7638	0.7213	0.6503	0.5574	0.5364	0.4877	0.4364
0.75	1.33	0.6478	0.6304	0.5920	0.5480	0.7588	0.7462	0.7133	0.6581	0.5429	0.5213	0.4787	0.4450
1	1	0.6346	0.6167	0.5820	0.5467	0.7447	0.7319	0.7040	0.6551	0.5305	0.5084	0.4682	0.4451
1.5	0.67	0.6145	0.5960	0.5646	0.5383	0.7238	0.7108	0.6878	0.6402	0.5113	0.4883	0.4498	0.4425
2	0.5	0.6005	0.5816	0.5515	0.5296	0.7099	0.6966	0.6758	0.6294	0.4975	0.4738	0.4361	0.4358
3	0.33	0.5840	0.5645	0.5334	0.5108	0.6940	0.6805	0.6592	0.6112	0.4805	0.4560	0.4169	0.4166
4	0.25	0.5758	0.5559	0.5244	0.5013	0.6865	0.6729	0.6510	0.6029	0.4718	0.4468	0.4074	0.4063
5	0.2	0.5716	0.5517	0.5198	0.4966	0.6828	0.6692	0.6469	0.5989	0.4673	0.4421	0.4026	0.4009
7.5	0.13	0.5682	0.5481	0.5161	0.4926	0.6799	0.6662	0.6435	0.5957	0.4635	0.4381	0.3986	0.3963
10	0.1	0.5676	0.5474	0.5154	0.4919	0.6794	0.6656	0.6428	0.5952	0.4628	0.4373	0.3978	0.3955
PGA		0.7054	0.6895	0.6122	0.4903	0.8232	0.8114	0.7322	0.6000	0.5939	0.5744	0.4998	0.3884
PGV		0.6221	0.6164	0.5654	0.4785	0.7333	0.7284	0.6725	0.5747	0.5171	0.5109	0.4648	0.3883

Table 11–14 Coefficients of the CENA single-station sigma (σ_{SS}) model-1.

Period (sec)	Frequency (Hz)	CENA σ_{SS} Model-1											
		Central				High				Low			
		σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}
0.01	100	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.02	50	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.03	33.33	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.04	25	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.05	20	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.075	13.33	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.1	10	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.15	6.67	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.2	5	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.25	4	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.3	3.33	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.4	2.5	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.5	2	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
0.75	1.33	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
1	1	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
1.5	0.67	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
2	0.5	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
3	0.33	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
4	0.25	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
5	0.2	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
7.5	0.13	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
10	0.1	0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
PGA		0.6790	0.6624	0.6367	0.6168	0.7965	0.7843	0.7611	0.7302	0.5680	0.5476	0.5200	0.5100
PGV		0.6609	0.6556	0.6448	0.6345	0.7722	0.7675	0.7524	0.7418	0.5556	0.5498	0.5429	0.5330

Table 11–15 Coefficients of the CENA single-station sigma (σ_{SS}) model-2.

Period (sec)	Frequency (Hz)	CENA σ_{SS} Model-2											
		Central				High				Low			
		σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}
0.01	100	0.6835	0.6670	0.5922	0.4777	0.8017	0.7897	0.7182	0.5850	0.5718	0.5515	0.4746	0.3780
0.02	50	0.6835	0.6670	0.5924	0.4783	0.8017	0.7897	0.7184	0.5855	0.5718	0.5515	0.4748	0.3786
0.03	33.33	0.6835	0.6670	0.5926	0.4788	0.8017	0.7897	0.7186	0.5860	0.5718	0.5515	0.4750	0.3792
0.04	25	0.6835	0.6670	0.5928	0.4795	0.8017	0.7897	0.7188	0.5866	0.5718	0.5515	0.4753	0.3798
0.05	20	0.6835	0.6670	0.5930	0.4799	0.8017	0.7897	0.7190	0.5871	0.5718	0.5515	0.4755	0.3803
0.075	13.33	0.6835	0.6670	0.5935	0.4813	0.8017	0.7897	0.7195	0.5884	0.5718	0.5515	0.4760	0.3818
0.1	10	0.6835	0.6670	0.5940	0.4827	0.8017	0.7897	0.7199	0.5898	0.5718	0.5515	0.4765	0.3832
0.15	6.67	0.6835	0.6670	0.5960	0.4881	0.8017	0.7897	0.7218	0.5950	0.5718	0.5515	0.4786	0.3887
0.2	5	0.6835	0.6670	0.6000	0.4992	0.8017	0.7897	0.7256	0.6058	0.5718	0.5515	0.4827	0.3999
0.25	4	0.6835	0.6670	0.6038	0.5100	0.8017	0.7897	0.7292	0.6165	0.5718	0.5515	0.4867	0.4105
0.3	3.33	0.6835	0.6670	0.6078	0.5214	0.8017	0.7897	0.7330	0.6280	0.5718	0.5515	0.4908	0.4217
0.4	2.5	0.6835	0.6670	0.6152	0.5427	0.8017	0.7897	0.7400	0.6500	0.5718	0.5515	0.4984	0.4420
0.5	2	0.6835	0.6670	0.6204	0.5577	0.8017	0.7897	0.7449	0.6660	0.5718	0.5515	0.5037	0.4562
0.75	1.33	0.6835	0.6670	0.6289	0.5833	0.8017	0.7897	0.7531	0.6936	0.5718	0.5515	0.5125	0.4797
1	1	0.6835	0.6670	0.6335	0.5971	0.8017	0.7897	0.7575	0.7087	0.5718	0.5515	0.5173	0.4921
1.5	0.67	0.6835	0.6670	0.6387	0.6127	0.8017	0.7897	0.7624	0.7261	0.5718	0.5515	0.5226	0.5060
2	0.5	0.6835	0.6670	0.6416	0.6217	0.8017	0.7897	0.7652	0.7362	0.5718	0.5515	0.5256	0.5139
3	0.33	0.6835	0.6670	0.6416	0.6217	0.8017	0.7897	0.7652	0.7362	0.5718	0.5515	0.5256	0.5139
4	0.25	0.6835	0.6670	0.6416	0.6217	0.8017	0.7897	0.7652	0.7362	0.5718	0.5515	0.5256	0.5139
5	0.2	0.6835	0.6670	0.6416	0.6217	0.8017	0.7897	0.7652	0.7362	0.5718	0.5515	0.5256	0.5139
7.5	0.13	0.6835	0.6670	0.6416	0.6217	0.8017	0.7897	0.7652	0.7362	0.5718	0.5515	0.5256	0.5139
10	0.1	0.6835	0.6670	0.6416	0.6217	0.8017	0.7897	0.7652	0.7362	0.5718	0.5515	0.5256	0.5139
PGA		0.6835	0.6670	0.5922	0.4777	0.8017	0.7897	0.7182	0.5850	0.5718	0.5515	0.4746	0.3780
PGV		0.6708	0.6656	0.6096	0.5104	0.7934	0.7888	0.7220	0.6170	0.5554	0.5496	0.5038	0.4109

Table 11–16 Mean of single-station sigma (σ_{SS}) computed from the continuous and discrete representation of the composite distributions.

Period (sec)	Frequency (Hz)	M 4.5		M 5.0		M 5.5		M 6.5	
		Continuous	Discrete	Continuous	Discrete	Continuous	Discrete	Continuous	Discrete
0.01	100	0.7017	0.7017	0.6858	0.6858	0.6141	0.6141	0.5031	0.5035
0.05	20	0.6985	0.6985	0.6825	0.6825	0.6120	0.6120	0.5032	0.5036
0.1	10	0.6946	0.6946	0.6785	0.6785	0.6094	0.6094	0.5034	0.5037
0.25	4	0.6838	0.6838	0.6675	0.6675	0.6084	0.6084	0.5212	0.5213
0.5	2	0.6684	0.6684	0.6516	0.6516	0.6076	0.6076	0.5506	0.5506
1	1	0.6450	0.6450	0.6276	0.6276	0.5941	0.5941	0.5600	0.5600
3	0.33	0.6046	0.6047	0.5859	0.5860	0.5562	0.5562	0.5336	0.5339
5	0.2	0.5948	0.5949	0.5757	0.5758	0.5454	0.5454	0.5224	0.5226
10	0.1	0.5916	0.5917	0.5724	0.5724	0.5419	0.5418	0.5186	0.5189
PGV		0.6320	0.6320	0.6265	0.6265	0.5790	0.5790	0.4984	0.4999

Table 11–17 Standard deviation of single-station sigma (σ_{SS}) computed from the continuous and discrete representation of the composite distributions.

Period (sec)	Frequency (Hz)	M 4.5		M 5.0		M 5.5		M 6.5	
		Continuous	Discrete	Continuous	Discrete	Continuous	Discrete	Continuous	Discrete
0.01	100	0.0703	0.0704	0.0727	0.0728	0.0720	0.0720	0.0751	0.0757
0.05	20	0.0699	0.0699	0.0723	0.0723	0.0720	0.0721	0.0756	0.0762
0.1	10	0.0693	0.0694	0.0718	0.0718	0.0721	0.0722	0.0762	0.0768
0.25	4	0.0682	0.0683	0.0707	0.0708	0.0721	0.0722	0.0734	0.0736
0.5	2	0.0675	0.0676	0.0701	0.0702	0.0725	0.0725	0.0691	0.0691
1	1	0.0686	0.0687	0.0714	0.0715	0.0750	0.0751	0.0689	0.0689
3	0.33	0.0765	0.0769	0.0799	0.0802	0.0848	0.0851	0.0748	0.0758
5	0.2	0.0795	0.0802	0.0830	0.0836	0.0880	0.0884	0.0788	0.0803
10	0.1	0.0807	0.0814	0.0842	0.0849	0.0891	0.0896	0.0803	0.0819
PGV		0.6320	0.0687	0.0688	0.0691	0.0692	0.0686	0.0687	0.0745

Table 11–18 Coefficients of the composite single-station sigma (σ_{SS}) model for CENA.

Period (sec)	Frequency (Hz)	Composite σ_{SS} Model											
		Central				High				Low			
		σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}	σ_{SS1}	σ_{SS2}	σ_{SS3}	σ_{SS4}
0.01	100	0.7006	0.6846	0.6127	0.4973	0.8193	0.8075	0.7348	0.6377	0.5879	0.5682	0.4981	0.3901
0.02	50	0.6998	0.6838	0.6122	0.4974	0.8183	0.8064	0.7343	0.6379	0.5873	0.5676	0.4976	0.3899
0.03	33.33	0.6990	0.6829	0.6116	0.4974	0.8173	0.8054	0.7338	0.6382	0.5867	0.5669	0.4970	0.3897
0.04	25	0.6981	0.6820	0.6110	0.4975	0.8162	0.8043	0.7333	0.6384	0.5860	0.5662	0.4964	0.3895
0.05	20	0.6974	0.6813	0.6105	0.4975	0.8153	0.8034	0.7329	0.6385	0.5855	0.5657	0.4960	0.3893
0.075	13.33	0.6955	0.6793	0.6092	0.4977	0.8129	0.8010	0.7317	0.6390	0.5840	0.5641	0.4946	0.3887
0.1	10	0.6935	0.6773	0.6079	0.4978	0.8106	0.7986	0.7306	0.6395	0.5825	0.5626	0.4933	0.3882
0.15	6.67	0.6898	0.6735	0.6063	0.5007	0.8061	0.7941	0.7292	0.6413	0.5795	0.5595	0.4917	0.3904
0.2	5	0.6862	0.6698	0.6066	0.5092	0.8019	0.7899	0.7295	0.6453	0.5765	0.5564	0.4921	0.3994
0.25	4	0.6827	0.6662	0.6068	0.5171	0.7980	0.7859	0.7297	0.6492	0.5736	0.5533	0.4924	0.4079
0.3	3.33	0.6794	0.6628	0.6072	0.5254	0.7943	0.7822	0.7300	0.6537	0.5707	0.5503	0.4928	0.4168
0.4	2.5	0.6730	0.6563	0.6075	0.5399	0.7876	0.7753	0.7305	0.6623	0.5650	0.5444	0.4931	0.4326
0.5	2	0.6671	0.6502	0.6060	0.5482	0.7816	0.7693	0.7295	0.6681	0.5595	0.5387	0.4913	0.4411
0.75	1.33	0.6541	0.6369	0.5999	0.5577	0.7697	0.7572	0.7256	0.6762	0.5467	0.5253	0.4837	0.4504
1	1	0.6432	0.6256	0.5921	0.5576	0.7611	0.7484	0.7209	0.6772	0.5354	0.5135	0.4741	0.4508
1.5	0.67	0.6262	0.6082	0.5781	0.5510	0.7502	0.7373	0.7137	0.6743	0.5171	0.4944	0.4566	0.4482
2	0.5	0.6143	0.5959	0.5674	0.5436	0.7444	0.7313	0.7093	0.6738	0.5037	0.4803	0.4433	0.4416
3	0.33	0.5998	0.5811	0.5517	0.5267	0.7391	0.7258	0.7033	0.6698	0.4870	0.4628	0.4243	0.4226
4	0.25	0.5925	0.5736	0.5438	0.5181	0.7372	0.7237	0.7009	0.6687	0.4784	0.4538	0.4148	0.4124
5	0.2	0.5889	0.5698	0.5397	0.5138	0.7364	0.7228	0.6999	0.6682	0.4740	0.4491	0.4101	0.4071
7.5	0.13	0.5858	0.5666	0.5364	0.5102	0.7358	0.7222	0.6990	0.6679	0.4702	0.4451	0.4061	0.4025
10	0.1	0.5853	0.5661	0.5357	0.5095	0.7357	0.7221	0.6989	0.6679	0.4695	0.4444	0.4053	0.4017
PGA		0.7006	0.6846	0.6127	0.4973	0.8193	0.8075	0.7348	0.6377	0.5879	0.5682	0.4981	0.3901
PGV		0.6303	0.6247	0.5765	0.4896	0.7480	0.7431	0.6960	0.6419	0.5220	0.5159	0.4703	0.3928

Table 11–19 Coefficients of the global ergodic total sigma (σ) model.

Period (sec)	Frequency (Hz)	Global σ Model											
		Central				High				Low			
		σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4
0.01	100	0.7662	0.7516	0.6815	0.5745	0.8783	0.8671	0.7939	0.6757	0.6593	0.6417	0.5750	0.4789
0.02	50	0.7652	0.7506	0.6808	0.5744	0.8772	0.8660	0.7933	0.6759	0.6585	0.6408	0.5743	0.4786
0.03	33.33	0.7715	0.7571	0.6884	0.5841	0.8823	0.8712	0.7996	0.6842	0.6658	0.6484	0.5829	0.4894
0.04	25	0.7847	0.7705	0.7036	0.6026	0.8940	0.8830	0.8131	0.7011	0.6803	0.6632	0.5995	0.5092
0.05	20	0.7955	0.7815	0.7160	0.6176	0.9041	0.8932	0.8248	0.7159	0.6917	0.6749	0.6124	0.5243
0.075	13.33	0.7912	0.7771	0.7121	0.6147	0.9015	0.8906	0.8236	0.7173	0.6858	0.6689	0.6063	0.5175
0.1	10	0.7753	0.7609	0.6954	0.5968	0.8880	0.8769	0.8102	0.7043	0.6679	0.6505	0.5866	0.4954
0.15	6.67	0.7616	0.7469	0.6828	0.5866	0.8740	0.8628	0.7981	0.6952	0.6545	0.6368	0.5737	0.4844
0.2	5	0.7567	0.7419	0.6819	0.5927	0.8674	0.8562	0.7959	0.6991	0.6512	0.6333	0.5740	0.4924
0.25	4	0.7526	0.7378	0.6815	0.5992	0.8619	0.8506	0.7944	0.7038	0.6485	0.6305	0.5746	0.5005
0.3	3.33	0.7484	0.7334	0.6810	0.6057	0.8564	0.8450	0.7930	0.7087	0.6454	0.6273	0.5749	0.5083
0.4	2.5	0.7394	0.7243	0.6785	0.6160	0.8455	0.8340	0.7894	0.7167	0.6383	0.6199	0.5735	0.5205
0.5	2	0.7304	0.7151	0.6738	0.6201	0.8352	0.8236	0.7843	0.7201	0.6304	0.6118	0.5691	0.5252
0.75	1.33	0.7094	0.6936	0.6591	0.6198	0.8126	0.8008	0.7700	0.7197	0.6110	0.5917	0.5541	0.5250
1	1	0.6913	0.6751	0.6436	0.6117	0.7943	0.7822	0.7559	0.7111	0.5932	0.5734	0.5376	0.5175
1.5	0.67	0.6632	0.6462	0.6175	0.5933	0.7668	0.7544	0.7325	0.6890	0.5648	0.5439	0.5093	0.5026
2	0.5	0.6433	0.6257	0.5980	0.5776	0.7483	0.7356	0.7157	0.6732	0.5437	0.5220	0.4877	0.4871
3	0.33	0.6185	0.6001	0.5711	0.5498	0.7264	0.7134	0.6928	0.6488	0.5165	0.4937	0.4576	0.4565
4	0.25	0.6049	0.5861	0.5563	0.5344	0.7151	0.7020	0.6808	0.6366	0.5009	0.4775	0.4406	0.4384
5	0.2	0.5969	0.5779	0.5477	0.5254	0.7090	0.6957	0.6741	0.6300	0.4916	0.4677	0.4304	0.4274
7.5	0.13	0.5877	0.5683	0.5375	0.5148	0.7022	0.6888	0.6668	0.6227	0.4802	0.4558	0.4180	0.4140
10	0.1	0.5838	0.5642	0.5332	0.5103	0.6991	0.6856	0.6634	0.6192	0.4757	0.4511	0.4130	0.4088
PGA		0.7662	0.7516	0.6815	0.5745	0.8783	0.8671	0.7939	0.6757	0.6593	0.6417	0.5750	0.4789
PGV		0.6971	0.6921	0.6471	0.5726	0.8002	0.7956	0.7453	0.6599	0.5990	0.5936	0.5536	0.4896

Table 11–20 Coefficients of the CENA ergodic total sigma (σ) model-1.

Period (sec)	Frequency (Hz)	CENA σ Model-1											
		Central				High				Low			
		σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4
0.01	100	0.7419	0.7269	0.7036	0.6855	0.8534	0.8420	0.8204	0.7925	0.6358	0.6175	0.5930	0.5838
0.02	50	0.7419	0.7268	0.7035	0.6854	0.8534	0.8419	0.8204	0.7924	0.6358	0.6175	0.5929	0.5838
0.03	33.33	0.7494	0.7345	0.7115	0.6936	0.8598	0.8485	0.8270	0.7993	0.6442	0.6262	0.6019	0.5930
0.04	25	0.7640	0.7493	0.7268	0.7093	0.8729	0.8617	0.8406	0.8135	0.6600	0.6424	0.6187	0.6099
0.05	20	0.7759	0.7615	0.7393	0.7220	0.8842	0.8731	0.8523	0.8258	0.6723	0.6550	0.6318	0.6231
0.075	13.33	0.7737	0.7592	0.7370	0.7197	0.8842	0.8731	0.8524	0.8260	0.6682	0.6508	0.6274	0.6184
0.1	10	0.7596	0.7449	0.7222	0.7045	0.8730	0.8617	0.8407	0.8140	0.6517	0.6339	0.6099	0.6005
0.15	6.67	0.7500	0.7350	0.7120	0.6941	0.8636	0.8523	0.8310	0.8039	0.6418	0.6237	0.5993	0.5899
0.2	5	0.7491	0.7342	0.7112	0.6933	0.8617	0.8503	0.8290	0.8016	0.6420	0.6239	0.5996	0.5903
0.25	4	0.7490	0.7340	0.7110	0.6931	0.8606	0.8493	0.8279	0.8004	0.6427	0.6246	0.6003	0.5911
0.3	3.33	0.7485	0.7335	0.7105	0.6926	0.8594	0.8480	0.8266	0.7989	0.6429	0.6248	0.6005	0.5914
0.4	2.5	0.7467	0.7317	0.7086	0.6906	0.8566	0.8452	0.8237	0.7958	0.6420	0.6238	0.5995	0.5906
0.5	2	0.7443	0.7293	0.7061	0.6881	0.8539	0.8424	0.8208	0.7927	0.6400	0.6218	0.5974	0.5885
0.75	1.33	0.7380	0.7228	0.6995	0.6813	0.8477	0.8362	0.8144	0.7860	0.6336	0.6152	0.5905	0.5817
1	1	0.7323	0.7170	0.6934	0.6751	0.8428	0.8312	0.8093	0.7807	0.6272	0.6086	0.5837	0.5748
1.5	0.67	0.7234	0.7079	0.6840	0.6654	0.8354	0.8238	0.8017	0.7727	0.6170	0.5981	0.5728	0.5636
2	0.5	0.7171	0.7015	0.6773	0.6585	0.8305	0.8188	0.7966	0.7675	0.6094	0.5904	0.5647	0.5553
3	0.33	0.7089	0.6930	0.6686	0.6496	0.8243	0.8125	0.7901	0.7608	0.5995	0.5802	0.5541	0.5444
4	0.25	0.7039	0.6879	0.6632	0.6441	0.8206	0.8087	0.7863	0.7569	0.5934	0.5738	0.5474	0.5375
5	0.2	0.7005	0.6844	0.6596	0.6403	0.8182	0.8063	0.7838	0.7543	0.5891	0.5695	0.5429	0.5328
7.5	0.13	0.6955	0.6793	0.6543	0.6348	0.8146	0.8027	0.7801	0.7506	0.5828	0.5630	0.5361	0.5258
10	0.1	0.6927	0.6764	0.6513	0.6318	0.8123	0.8004	0.7777	0.7481	0.5797	0.5597	0.5327	0.5223
PGA		0.7419	0.7269	0.7036	0.6855	0.8534	0.8420	0.8204	0.7925	0.6358	0.6175	0.5930	0.5838
PGV		0.7320	0.7272	0.7174	0.7082	0.8358	0.8315	0.8178	0.8081	0.6329	0.6277	0.6215	0.6128

Table 11–21 Coefficients of the CENA ergodic total sigma (σ) model-2.

Period (sec)	Frequency (Hz)	CENA σ Model-2											
		Central				High				Low			
		σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4
0.01	100	0.7461	0.7311	0.6636	0.5637	0.8583	0.8469	0.7809	0.6627	0.6392	0.6211	0.5530	0.4702
0.02	50	0.7460	0.7310	0.6638	0.5641	0.8582	0.8469	0.7810	0.6631	0.6392	0.6210	0.5531	0.4706
0.03	33.33	0.7535	0.7386	0.6724	0.5745	0.8646	0.8533	0.7882	0.6718	0.6476	0.6296	0.5629	0.4823
0.04	25	0.7680	0.7534	0.6888	0.5938	0.8777	0.8665	0.8026	0.6894	0.6633	0.6458	0.5810	0.5032
0.05	20	0.7798	0.7655	0.7021	0.6094	0.8889	0.8779	0.8151	0.7046	0.6756	0.6584	0.5949	0.5190
0.075	13.33	0.7776	0.7632	0.7001	0.6076	0.8889	0.8778	0.8157	0.7067	0.6714	0.6542	0.5906	0.5137
0.1	10	0.7636	0.7490	0.6849	0.5907	0.8777	0.8665	0.8040	0.6941	0.6551	0.6374	0.5725	0.4930
0.15	6.67	0.7540	0.7392	0.6759	0.5827	0.8684	0.8571	0.7955	0.6862	0.6452	0.6273	0.5631	0.4851
0.2	5	0.7532	0.7384	0.6786	0.5911	0.8665	0.8552	0.7967	0.6924	0.6454	0.6274	0.5669	0.4953
0.25	4	0.7531	0.7382	0.6818	0.6000	0.8654	0.8541	0.7988	0.6998	0.6461	0.6281	0.5712	0.5055
0.3	3.33	0.7526	0.7377	0.6848	0.6092	0.8642	0.8529	0.8009	0.7080	0.6462	0.6283	0.5750	0.5154
0.4	2.5	0.7508	0.7359	0.6894	0.6253	0.8615	0.8501	0.8043	0.7236	0.6453	0.6273	0.5806	0.5320
0.5	2	0.7484	0.7335	0.6914	0.6357	0.8587	0.8474	0.8059	0.7344	0.6433	0.6252	0.5830	0.5419
0.75	1.33	0.7422	0.7271	0.6924	0.6511	0.8526	0.8412	0.8070	0.7522	0.6370	0.6187	0.5839	0.5550
1	1	0.7365	0.7213	0.6905	0.6571	0.8477	0.8362	0.8059	0.7607	0.6306	0.6122	0.5812	0.5588
1.5	0.67	0.7277	0.7122	0.6858	0.6616	0.8404	0.8288	0.8029	0.7688	0.6205	0.6017	0.5751	0.5600
2	0.5	0.7214	0.7058	0.6819	0.6631	0.8355	0.8239	0.8005	0.7731	0.6130	0.5940	0.5699	0.5590
3	0.33	0.7132	0.6975	0.6733	0.6542	0.8293	0.8176	0.7941	0.7665	0.6031	0.5839	0.5593	0.5481
4	0.25	0.7082	0.6923	0.6679	0.6488	0.8256	0.8139	0.7903	0.7626	0.5970	0.5776	0.5527	0.5413
5	0.2	0.7048	0.6889	0.6644	0.6451	0.8232	0.8114	0.7878	0.7601	0.5928	0.5733	0.5482	0.5366
7.5	0.13	0.6999	0.6838	0.6590	0.6396	0.8197	0.8078	0.7841	0.7563	0.5866	0.5669	0.5415	0.5297
10	0.1	0.6971	0.6810	0.6561	0.6366	0.8174	0.8055	0.7817	0.7538	0.5834	0.5636	0.5381	0.5262
PGA		0.7461	0.7311	0.6636	0.5637	0.8583	0.8469	0.7809	0.6627	0.6392	0.6211	0.5530	0.4702
PGV		0.7410	0.7363	0.6861	0.5997	0.8550	0.8507	0.7899	0.6961	0.6327	0.6276	0.5873	0.5084

Table 11–22 Mean of sigma (σ) computed from the continuous and discrete representation of the composite distributions.

Period (sec)	Frequency (Hz)	M 4.5		M 5.0		M 5.5		M 6.5	
		Continuous	Discrete	Continuous	Discrete	Continuous	Discrete	Continuous	Discrete
0.01	100	0.7627	0.7627	0.7481	0.7481	0.6830	0.6830	0.5855	0.5858
0.05	20	0.7928	0.7928	0.7788	0.7788	0.7179	0.7179	0.6281	0.6283
0.1	10	0.7736	0.7736	0.7592	0.7592	0.6981	0.6981	0.6080	0.6082
0.25	4	0.7533	0.7533	0.7385	0.7385	0.6856	0.6856	0.6097	0.6098
0.5	2	0.7345	0.7345	0.7193	0.7193	0.6799	0.6799	0.6294	0.6294
1	1	0.7008	0.7008	0.6849	0.6849	0.6545	0.6545	0.6236	0.6236
3	0.33	0.6381	0.6382	0.6204	0.6205	0.5926	0.5926	0.5713	0.5715
5	0.2	0.6193	0.6194	0.6010	0.6011	0.5722	0.5721	0.5501	0.5503
10	0.1	0.6073	0.6074	0.5886	0.5887	0.5591	0.5590	0.5364	0.5366
PGV		0.7059	0.7059	0.7010	0.7010	0.6589	0.6589	0.5897	0.5908

Table 11–23 Standard deviation of sigma (σ) computed from the continuous and discrete representation of the composite distributions.

Period (sec)	Frequency (Hz)	M 4.5		M 5.0		M 5.5		M 6.5	
		Continuous	Discrete	Continuous	Discrete	Continuous	Discrete	Continuous	Discrete
0.01	100	0.0671	0.0672	0.0691	0.0692	0.0677	0.0677	0.0689	0.0693
0.05	20	0.0650	0.0650	0.0668	0.0668	0.0656	0.0656	0.0665	0.0668
0.1	10	0.0673	0.0673	0.0692	0.0692	0.0690	0.0691	0.0713	0.0716
0.25	4	0.0652	0.0653	0.0672	0.0673	0.0679	0.0679	0.0678	0.0680
0.5	2	0.0632	0.0632	0.0653	0.0653	0.0667	0.0668	0.0629	0.0629
1	1	0.0644	0.0645	0.0668	0.0668	0.0695	0.0696	0.0635	0.0635
3	0.33	0.0746	0.0750	0.0776	0.0779	0.0819	0.0821	0.0727	0.0735
5	0.2	0.0791	0.0796	0.0823	0.0828	0.0868	0.0872	0.0784	0.0795
10	0.1	0.0817	0.0822	0.0851	0.0856	0.0897	0.0901	0.0817	0.0829
PGV		0.0639	0.0639	0.0642	0.0642	0.0631	0.0631	0.0669	0.0686

Table 11–24 Coefficients of the composite ergodic sigma (σ) model for CENA.

Period (sec)	Frequency (Hz)	Composite σ Model											
		Central				High				Low			
		σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4
0.01	100	0.7618	0.7471	0.6819	0.5809	0.8747	0.8635	0.7963	0.7076	0.6539	0.6361	0.5736	0.4807
0.02	50	0.7610	0.7463	0.6814	0.5808	0.8737	0.8625	0.7958	0.7078	0.6533	0.6355	0.5731	0.4805
0.03	33.33	0.7676	0.7530	0.6891	0.5905	0.8791	0.8679	0.8022	0.7157	0.6610	0.6434	0.5819	0.4913
0.04	25	0.7810	0.7667	0.7044	0.6090	0.8909	0.8799	0.8158	0.7315	0.6758	0.6586	0.5987	0.5113
0.05	20	0.7920	0.7779	0.7169	0.6239	0.9011	0.8902	0.8275	0.7452	0.6874	0.6705	0.6118	0.5264
0.075	13.33	0.7881	0.7739	0.7133	0.6214	0.8990	0.8880	0.8267	0.7457	0.6820	0.6650	0.6060	0.5200
0.1	10	0.7726	0.7581	0.6969	0.6039	0.8858	0.8747	0.8137	0.7328	0.6645	0.6471	0.5867	0.4982
0.15	6.67	0.7597	0.7450	0.6850	0.5941	0.8725	0.8613	0.8023	0.7233	0.6521	0.6343	0.5743	0.4876
0.2	5	0.7556	0.7408	0.6844	0.6001	0.8668	0.8555	0.8004	0.7246	0.6496	0.6317	0.5750	0.4958
0.25	4	0.7523	0.7374	0.6844	0.6065	0.8621	0.8508	0.7994	0.7270	0.6476	0.6296	0.5761	0.5040
0.3	3.33	0.7488	0.7338	0.6842	0.6129	0.8575	0.8461	0.7983	0.7295	0.6452	0.6271	0.5767	0.5119
0.4	2.5	0.7412	0.7261	0.6825	0.6231	0.8485	0.8370	0.7956	0.7340	0.6392	0.6210	0.5760	0.5244
0.5	2	0.7335	0.7182	0.6786	0.6274	0.8402	0.8286	0.7918	0.7360	0.6323	0.6138	0.5723	0.5293
0.75	1.33	0.7151	0.6995	0.6661	0.6283	0.8228	0.8110	0.7814	0.7361	0.6145	0.5954	0.5586	0.5299
1	1	0.6992	0.6832	0.6528	0.6215	0.8096	0.7976	0.7717	0.7314	0.5977	0.5780	0.5430	0.5227
1.5	0.67	0.6742	0.6576	0.6299	0.6050	0.7916	0.7792	0.7568	0.7203	0.5703	0.5496	0.5156	0.5080
2	0.5	0.6562	0.6392	0.6127	0.5907	0.7807	0.7681	0.7471	0.7139	0.5497	0.5282	0.4945	0.4926
3	0.33	0.6337	0.6160	0.5885	0.5651	0.7687	0.7558	0.7343	0.7025	0.5229	0.5004	0.4648	0.4624
4	0.25	0.6212	0.6032	0.5749	0.5509	0.7627	0.7496	0.7276	0.6967	0.5076	0.4844	0.4480	0.4446
5	0.2	0.6139	0.5956	0.5670	0.5426	0.7591	0.7460	0.7238	0.6932	0.4983	0.4747	0.4379	0.4337
7.5	0.13	0.6053	0.5867	0.5576	0.5327	0.7545	0.7412	0.7188	0.6883	0.4871	0.4630	0.4256	0.4205
10	0.1	0.6016	0.5829	0.5535	0.5285	0.7519	0.7386	0.7160	0.6855	0.4826	0.4583	0.4207	0.4153
PGA		0.7618	0.7471	0.6819	0.5809	0.8747	0.8635	0.7963	0.7076	0.6539	0.6361	0.5736	0.4807
PGV		0.7045	0.6995	0.6569	0.5825	0.8134	0.8090	0.7661	0.7163	0.6034	0.5980	0.5587	0.4936

Table 11–25 P-values from the F-test(1) of equality of τ^2 for PIE and the tectonic events for CENA.

Period (sec)	0.100	0.150	0.200	0.250	0.300	0.400	0.500	0.750	1.00
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00
Tectonic vs PIE	0.72	0.73	0.63	0.80	0.85	0.87	0.56	0.48	0.28

(1) Null and alternate hypotheses are: $H_0 : \tau^2_{Tectonic} = \tau^2_{PIE}$ and $H_a : \tau^2_{Tectonic} \neq \tau^2_{PIE}$, respectively.

Table 11–26 P-values from the F-test(1) of equality of ϕ_{SS}^2 for PIE and the tectonic events for CENA.

Period (sec)	0.100	0.150	0.200	0.250	0.300	0.400	0.500	0.750	1.00
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00
Tectonic vs PIE	0.54	0.71	0.94	0.61	0.62	0.83	0.64	0.61	0.06

(1) Null and alternate hypotheses are: $H_0 : \phi^2_{SS_{Tectonic}} = \phi^2_{SS_{PIE}}$ and $H_a : \Phi^2_{SS_{Tectonic}} \neq \Phi^2_{SS_{PIE}}$, respectively.

5. respectively.

Table 11–27 P-values from the F-test(1) of equality of τ^2 for tectonic events in the Gulf Coast Region and the rest of CENA. Values in red show cases where the null hypothesis(1) is rejected at 5% significance level.

Period (sec)	T0.10	T0.15	T0.20	T0.25	T0.30	T0.40	T0.50	T0.75	T1.00
Frequency (Hz)	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

(1) Null and alternate hypotheses are: $H_0 : \tau^2_{CENA} = \tau^2_{Gulf}$ and $H_a : \tau^2_{CENA} \neq \tau^2_{Gulf}$, respectively.

Table 11-28 P-values from the F-test(1) of equality of ϕ_{SS}^2 for tectonic events in the Gulf Coast Region and the rest of CENA.

Period (sec)	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.75	1.00
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00
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

(1) Null and alternate hypotheses are: $H_0 : \phi^2_{SS_{CENA}} = \phi^2_{SS_{Gulf}}$ and $H_a : \Phi^2_{SS_{CENA}} \neq \Phi^2_{SS_{Gulf}}$, respectively.

Table 11–29 P-values from the F-test(1) of equality of ϕ_{S2S}^2 for the Gulf Coast Region and the rest of CENA. Values in red show cases where the null hypothesis(1) is rejected at 5% significance level.

Period (sec)	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.75	1.00
Frequency (Hz)	10.00	6.67	5.00	4.00	3.33	2.50	2.00	1.33	1.00
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

(1) Null and alternate hypotheses are: $H_0 : \phi^2_{S2S_{CENA}} = \phi^2_{S2S_{Gulf}}$ and $H_a : \phi^2_{S2S_{CENA}} \neq \tau^2_{S2S_{Gulf}}$, respectively.

Figure 11–1 Logic tree for Tau (τ). "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively.

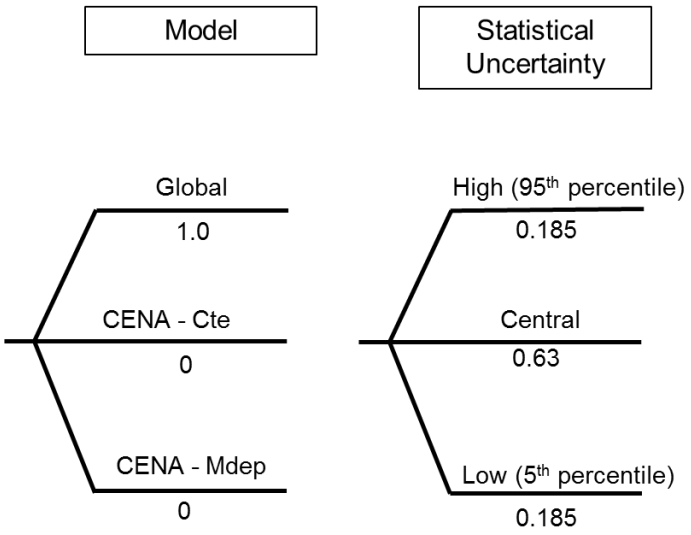


Figure 11–2 Comparison of candidate Tau (τ) models for CENA. "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively. Solid lines show the central branches. Dashed lines show the low and high branches.

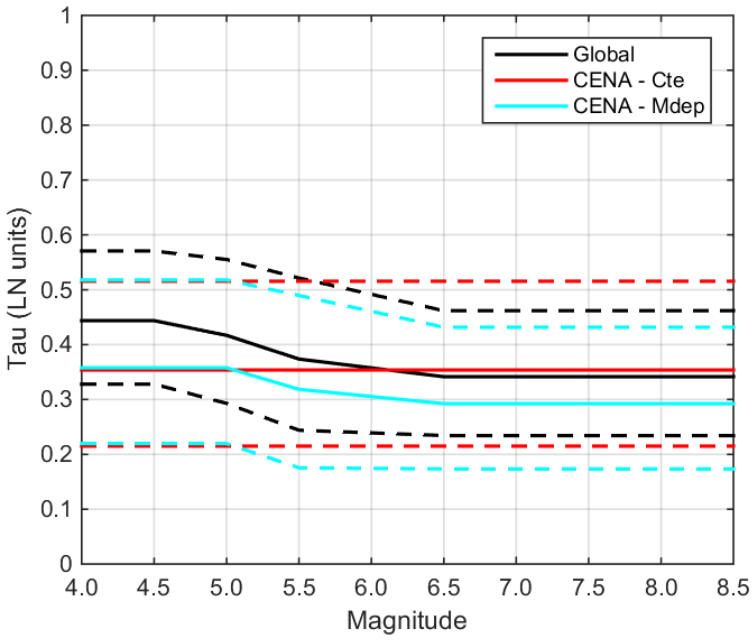


Figure 11–3 Comparison of global Tau (τ) model for CENA to SWUS τ model. Solid lines show the central branches. Dashed lines show the low and high branches.

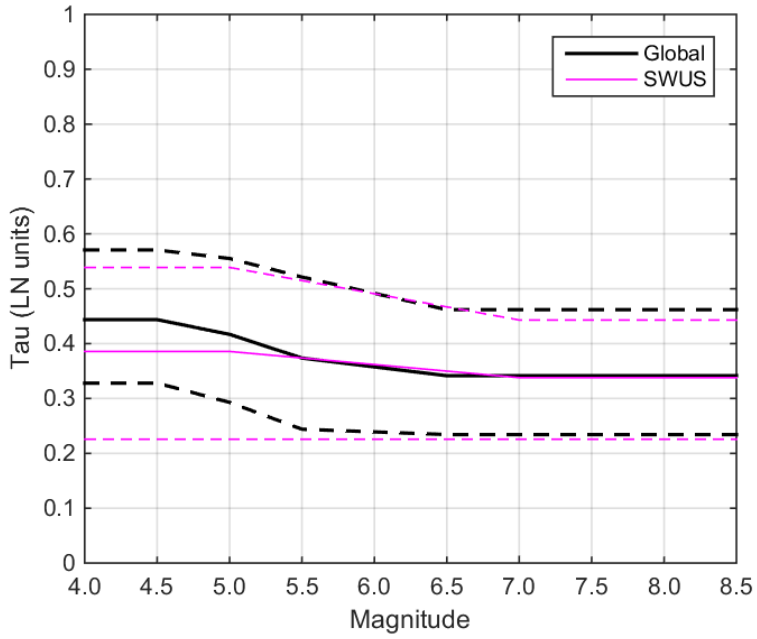


Figure 11–4 Comparison of global Tau (τ) model for CENA to the Hanford τ model at M 5.0, 6.0, and 7.0. Solid lines show the central branches. Dashed lines show the low and high branches.

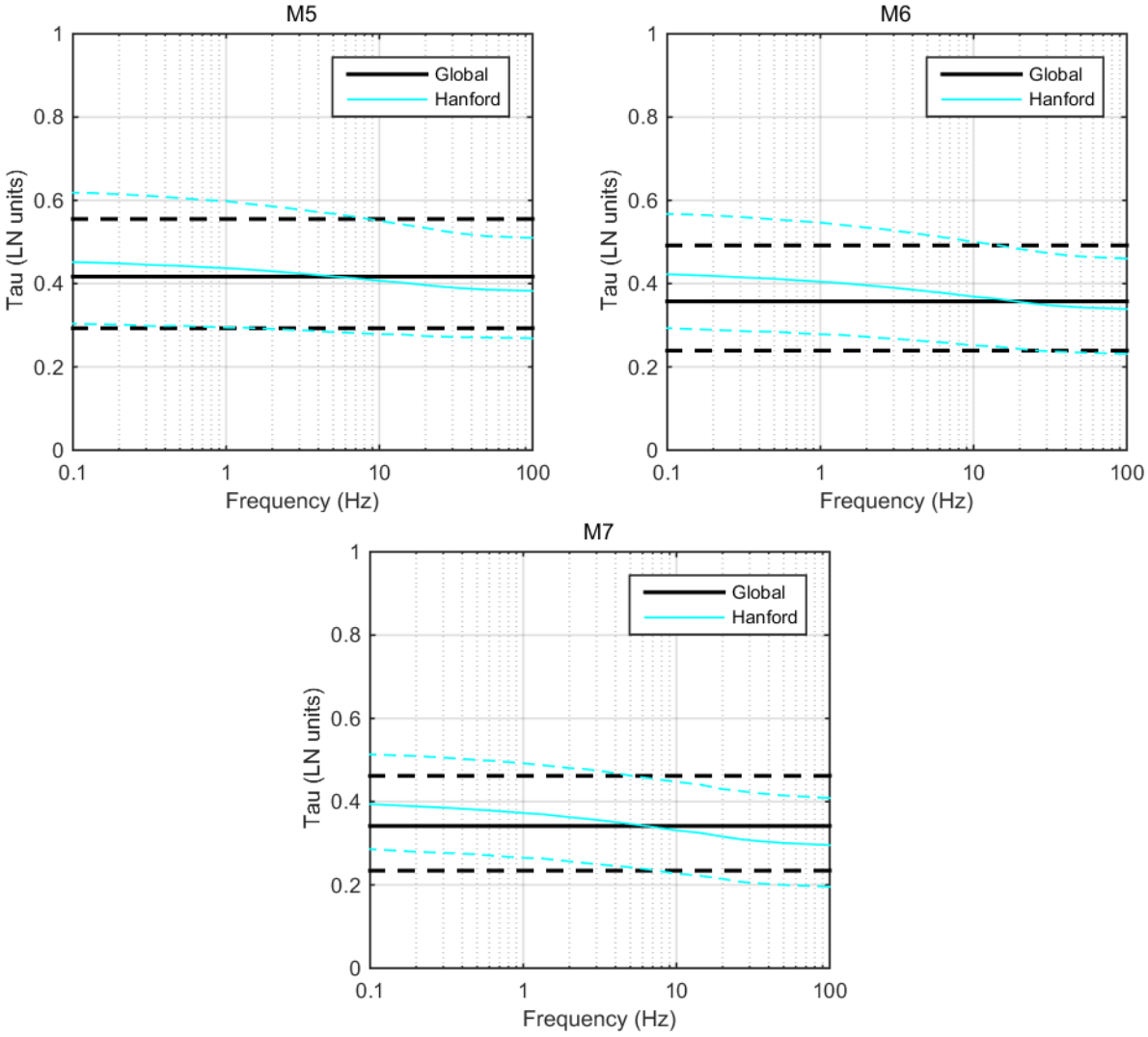


Figure 11–5 Logic tree for PhiSS (ϕ_{SS}). "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively.

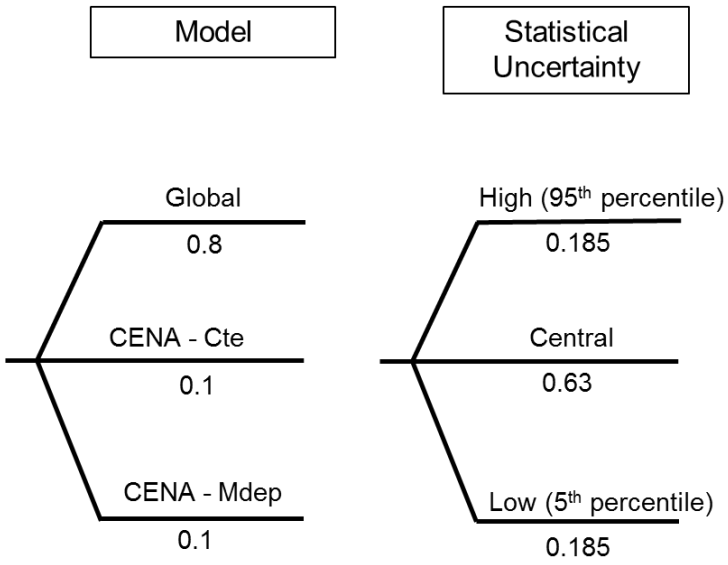


Figure 11–6 Comparison of candidate PhiSS (ϕ_{SS}) models for CENA versus magnitude at PGV, and f = 100, 10, 1, 0.2, and 0.1 Hz. "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively. Solid lines show the central branches. Dashed lines show the low and high branches.

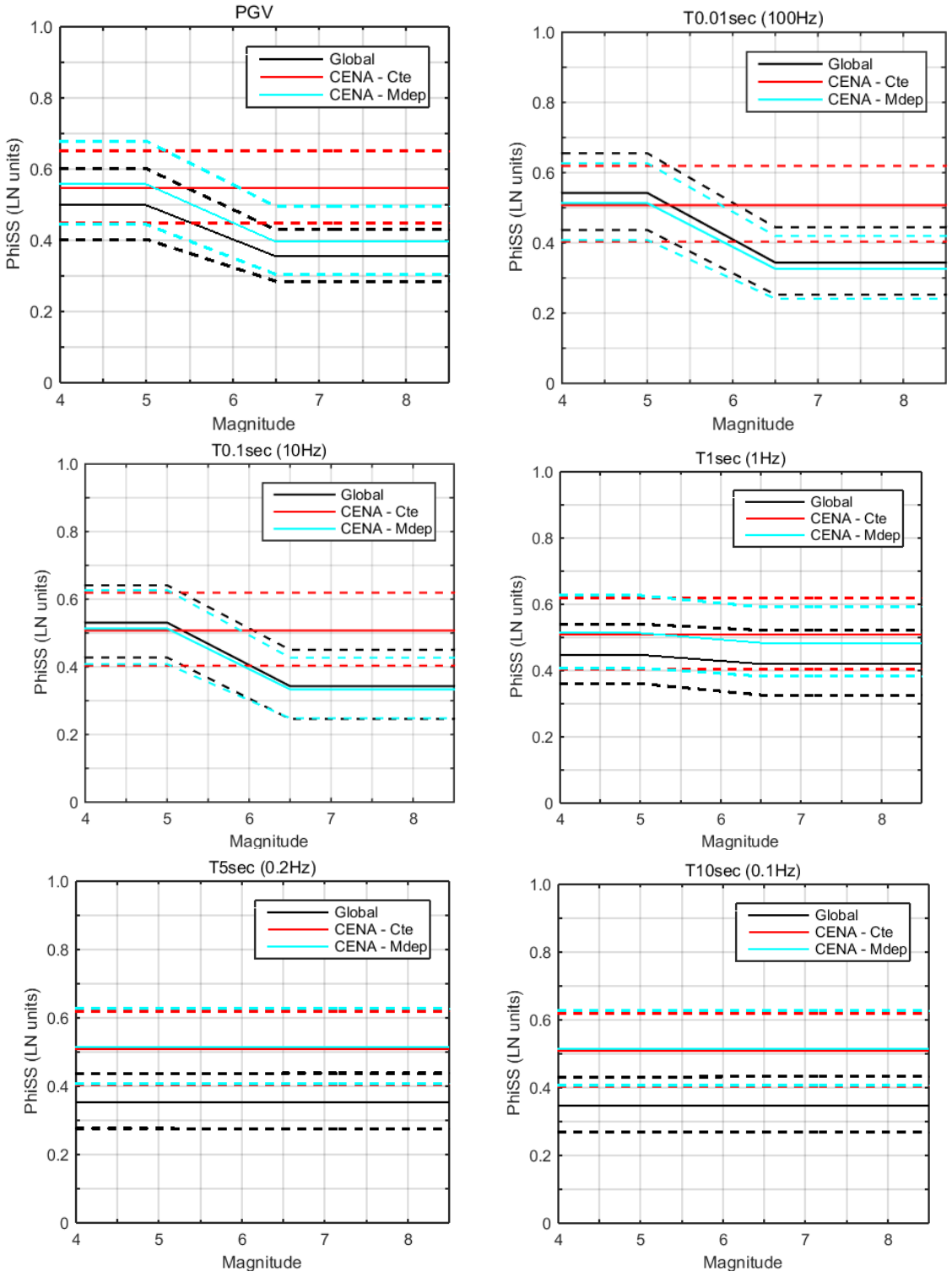


Figure 11–7 Comparison of candidate PhiSS (ϕ_{ss}) models for CENA versus frequency for M = 5.0, 6.0, and 7.0. "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively. Solid lines show the central branches. Dashed lines show the low and high branches.

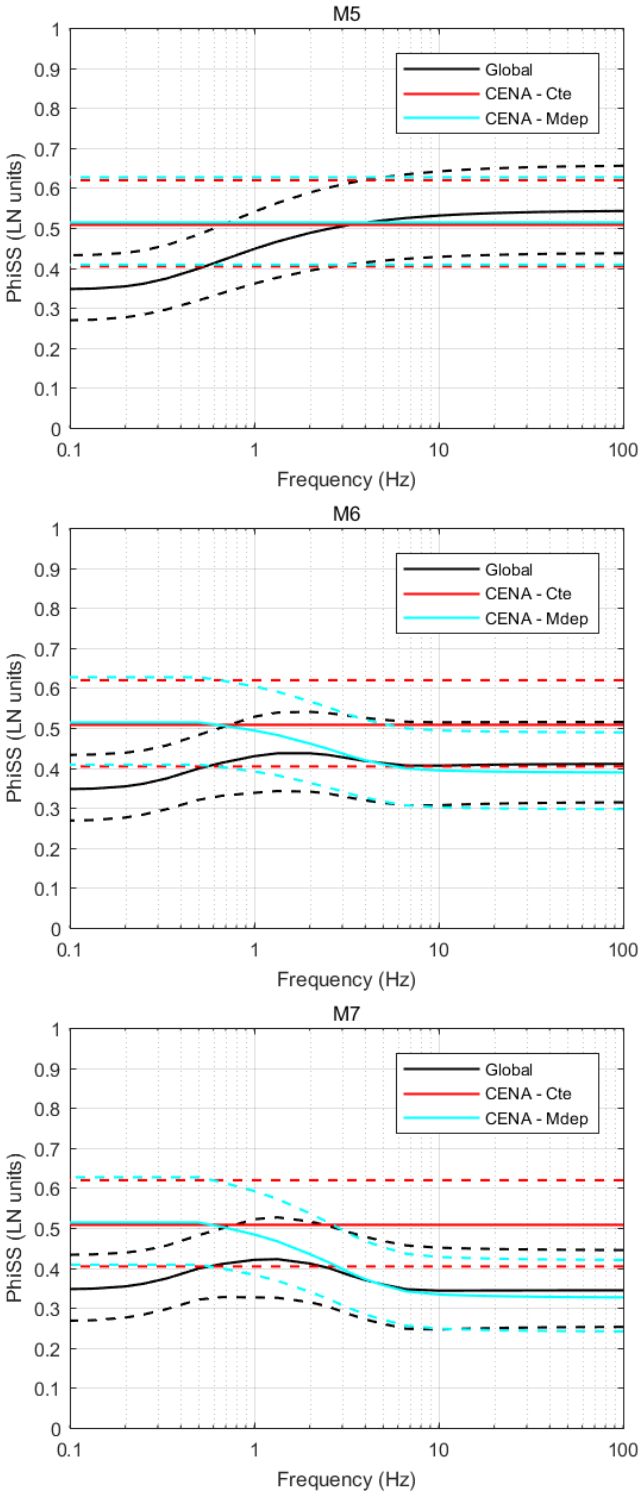


Figure 11–8 Magnitude-dependence of PhiSS (ϕ_{SS}) for the Japanese dataset. Error bars represent one standard error.

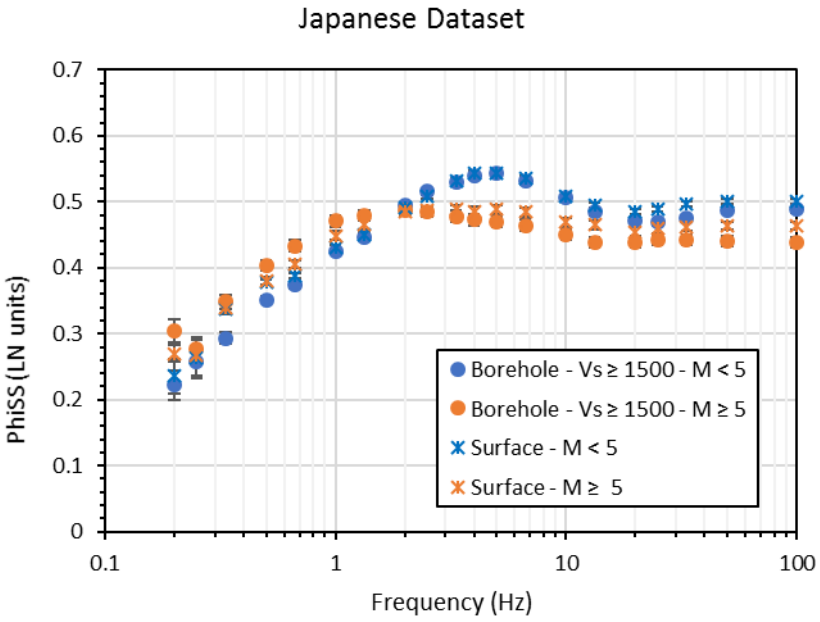


Figure 11–9 Comparison of the three PhiSS (ϕ_{SS}) models for CENA to the Hanford model at f = 100 and 1 Hz. "CENA - Mdep" refers to the magnitude-dependent CENA model. Solid lines show the central branches. Dashed lines show the low and high branches.

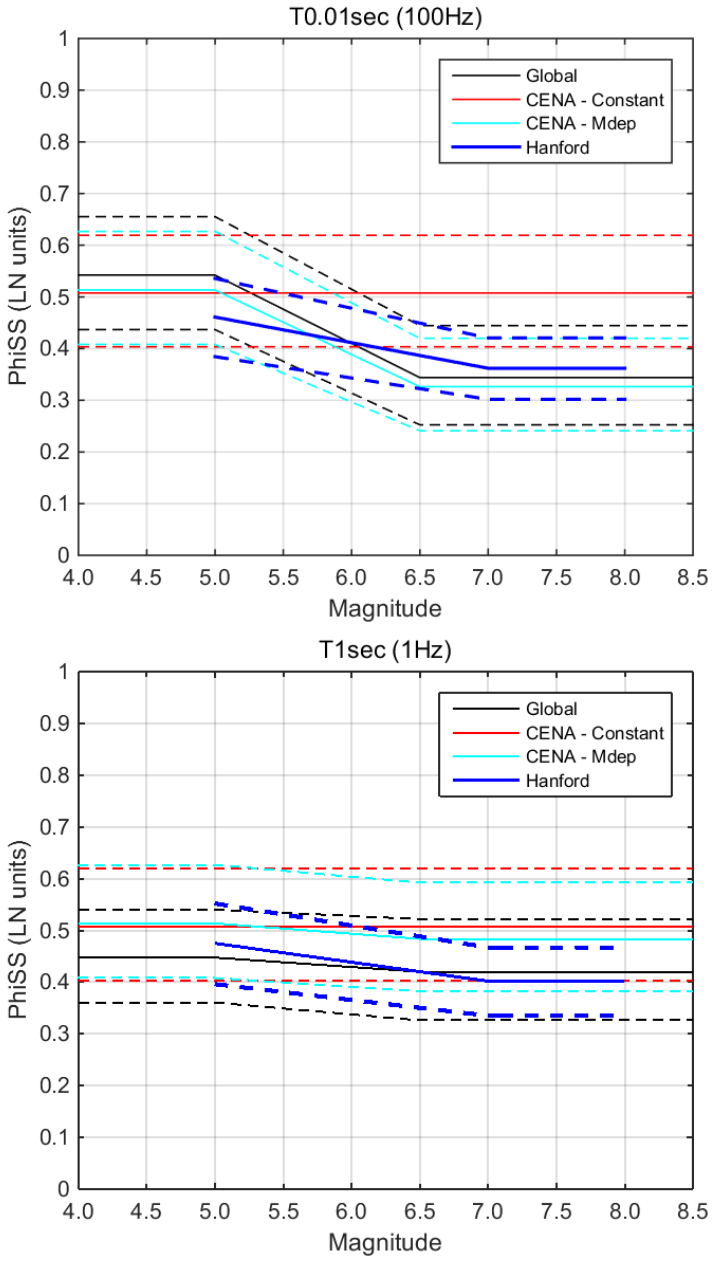


Figure 11–10 Comparison of the three PhiSS (ϕ_{SS}) models for CENA to the SWUS model at f = 100 and 1 Hz. "CENA - Mdep" refers to the magnitude-dependent CENA model. Solid lines show the central branches. Dashed lines show the low and high branches.

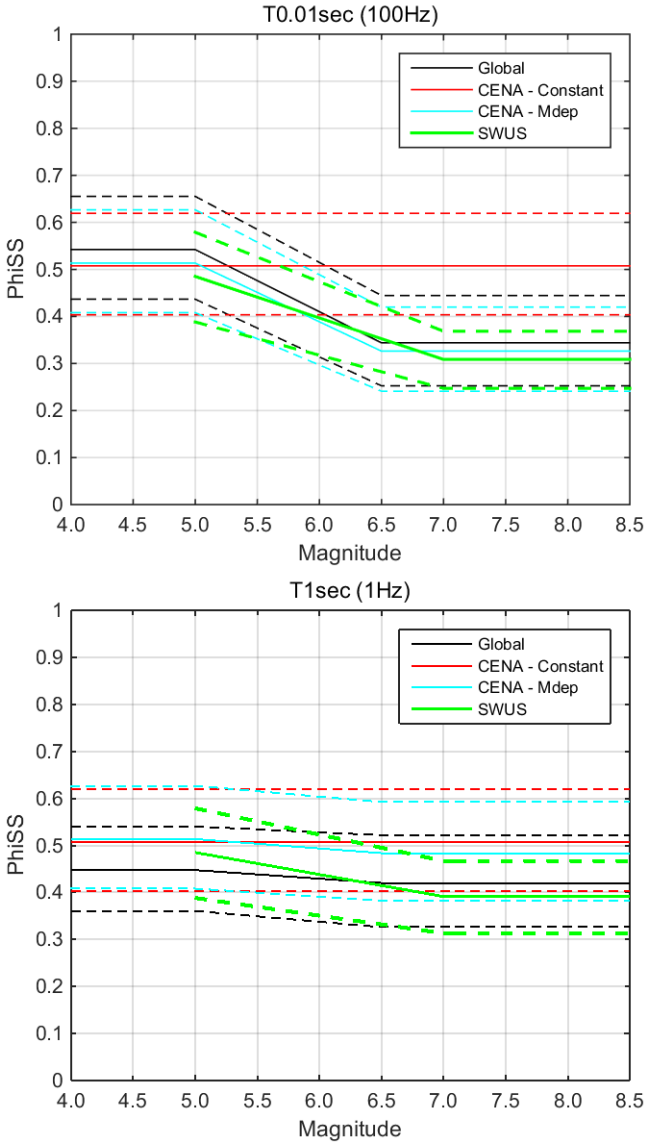


Figure 11–11 Logic tree for PhiS2S (ϕ_{S2S}).

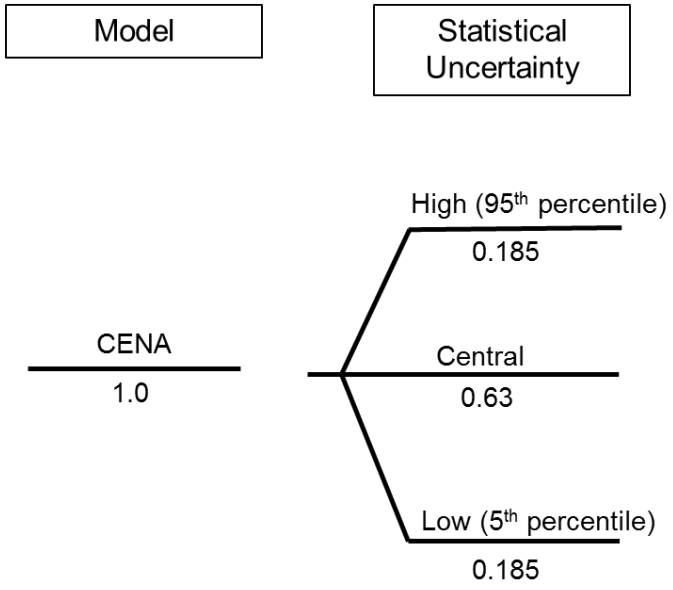


Figure 11–12 Central (solid line), high, and low branches (dashed lines) of the CENA PhiS2S (ϕ_{S2S}) model.

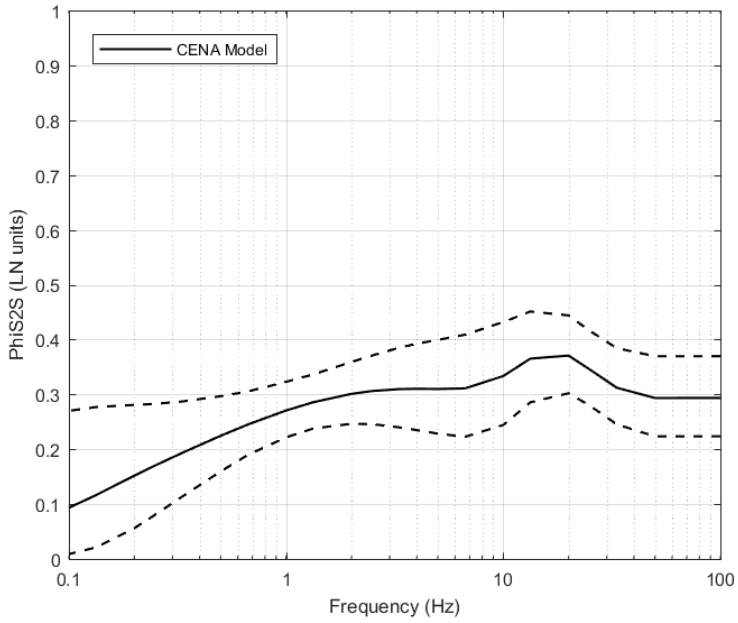
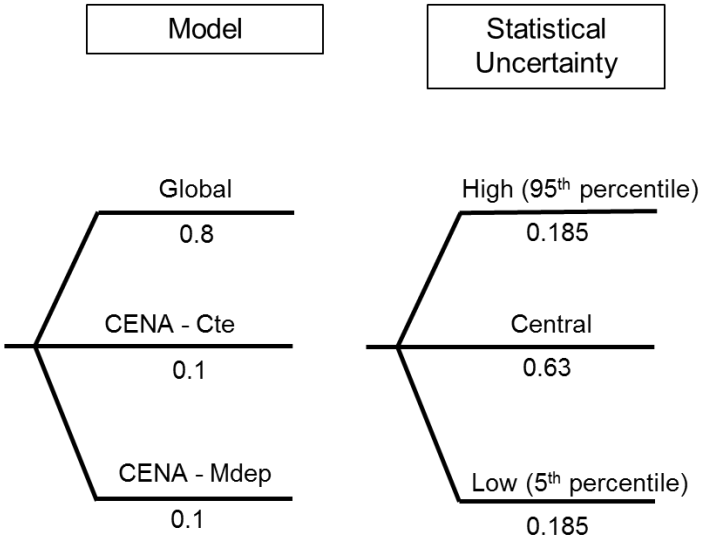


Figure 11–13 Logic tree for Phi (ϕ)[1]. "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively.



(1) All three models for Phi (ϕ) utilize the same CENA PhiS2S (ϕ_{S2S}) model.

Figure 11–14 Comparison of CENA Phi (ϕ) models versus magnitude at PGV, and f = 100, 10, 2, 1, and 0.2 Hz. "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively. Solid lines show the central branches. Dashed lines show the low and high branches.

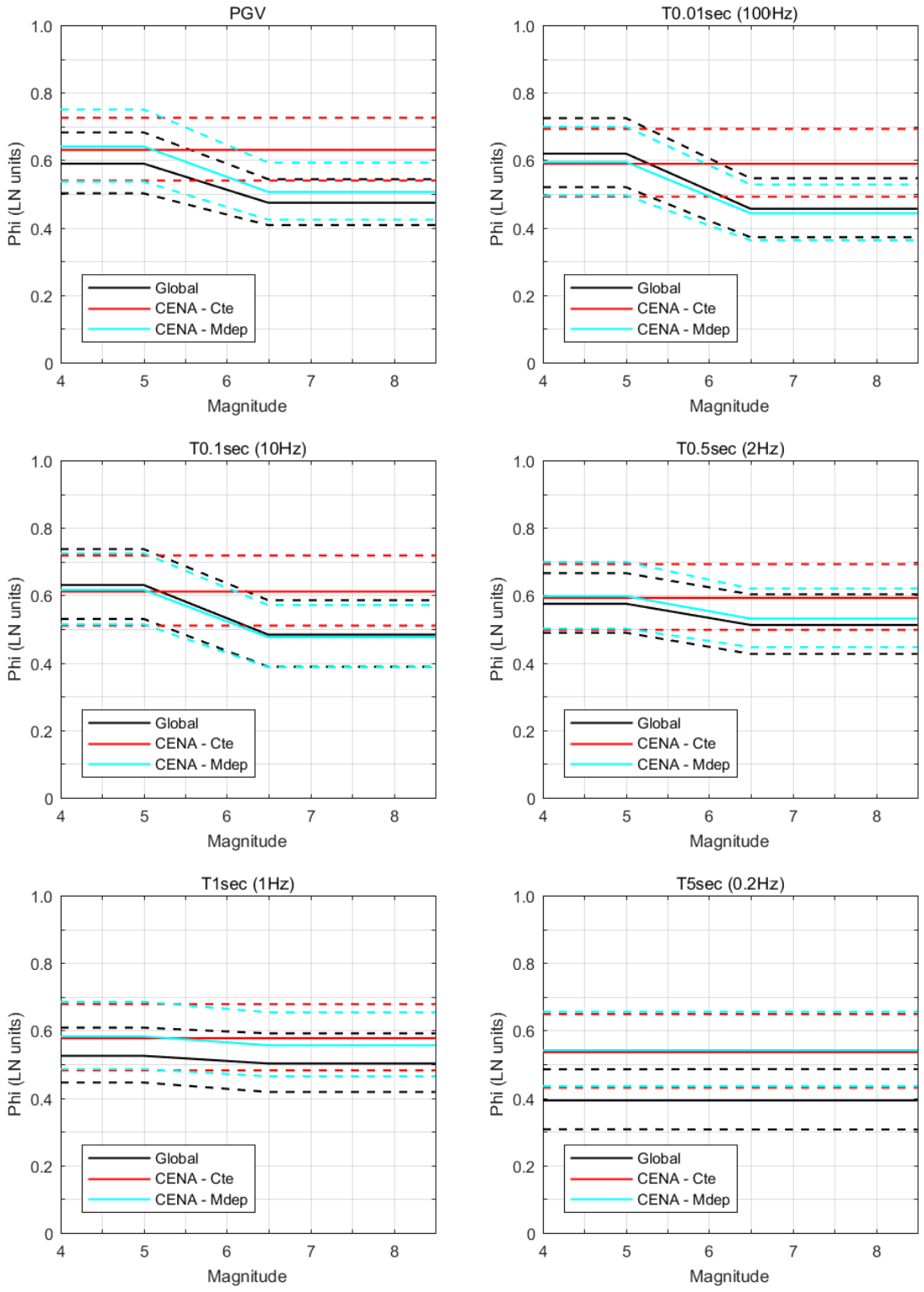


Figure 11–15 Comparison of CENA Phi (ϕ) models versus frequency for M = 5.0, 6.0, and 7.0. "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively. Solid lines show the central branches. Dashed lines show the low and high branches.

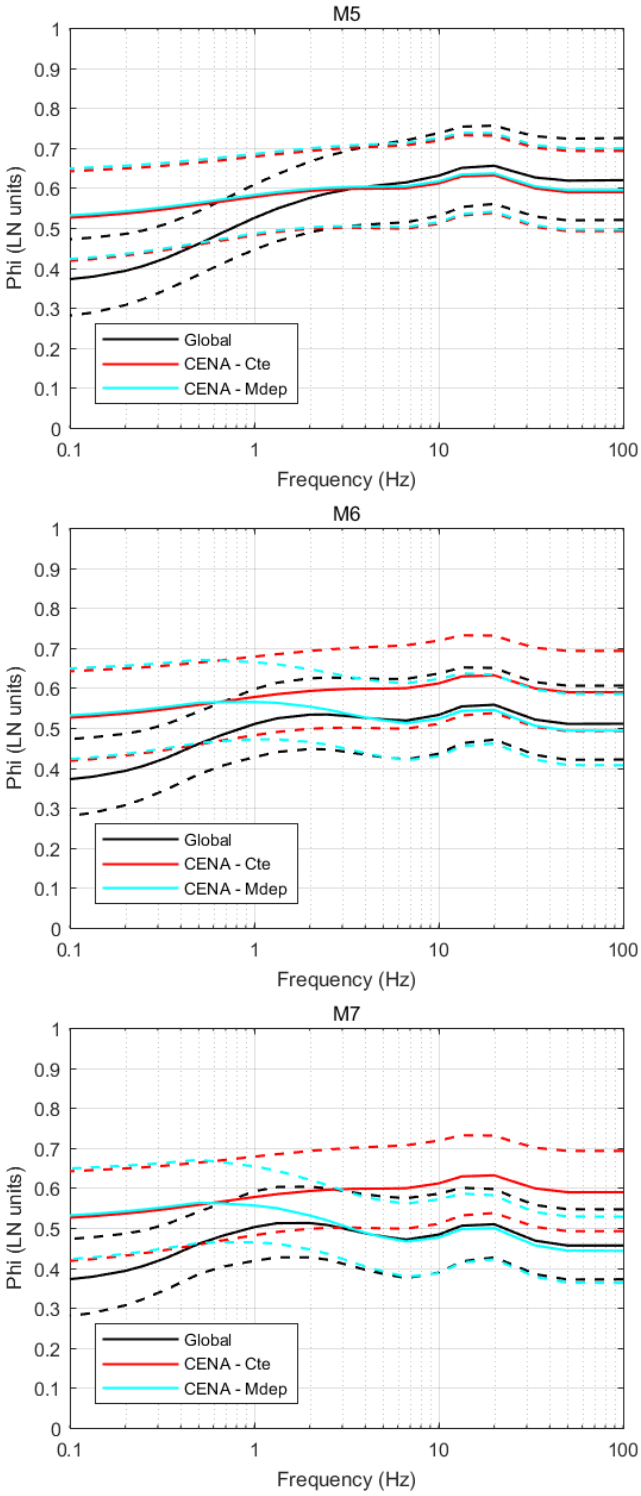


Figure 11–16 Comparison of CENA and NGA-West2 Phi (ϕ) models versus frequency for M = 5.0, 6.0, and 7.0. "CENA - Cte" and "CENA - Mdep" refer to the constant and magnitude-dependent CENA models, respectively. Solid lines show the central branches. Dashed lines show the low and high branches.

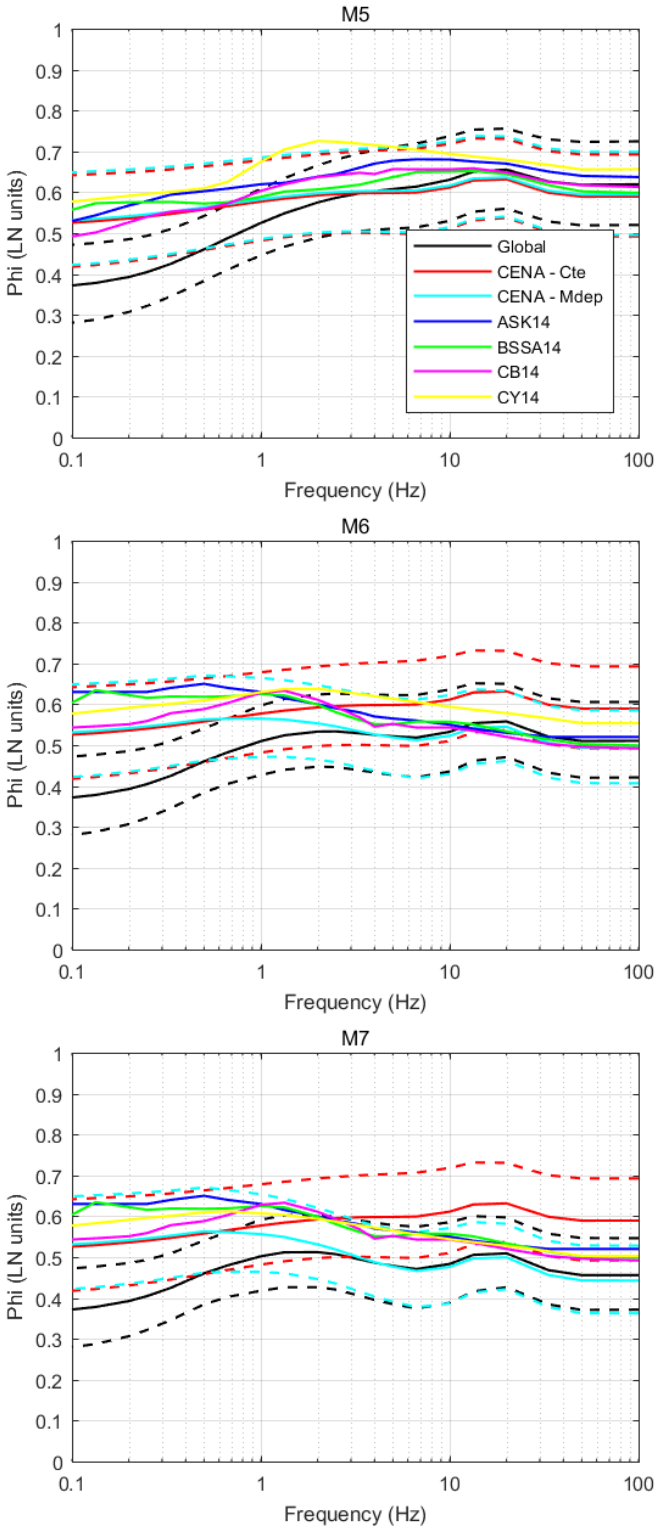
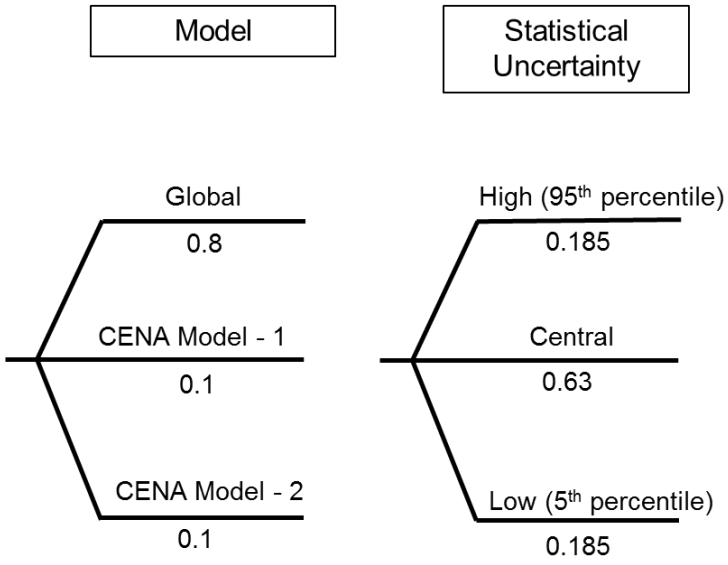


Figure 11–17 Logic tree for single-station sigma (σ_{SS})[(1)].



(1) Global Model is composed of global Tau (τ) and global PhiSS (ϕ_{SS}) models.

6. CENA Model-1 is composed of global Tau (τ) and CENA constant PhiSS (ϕ_{SS}) models.
7. CENA Model-2 is composed of global Tau (τ) and CENA magnitude-dependent PhiSS (ϕ_{SS}) models.

Figure 11–18 Comparison of CENA single-station sigma (σ_{SS}) models versus magnitude at PGV, and f = 100, 10, 2, 1, and 0.2 Hz. Solid lines show the central branches. Dashed lines show the low and high branches.

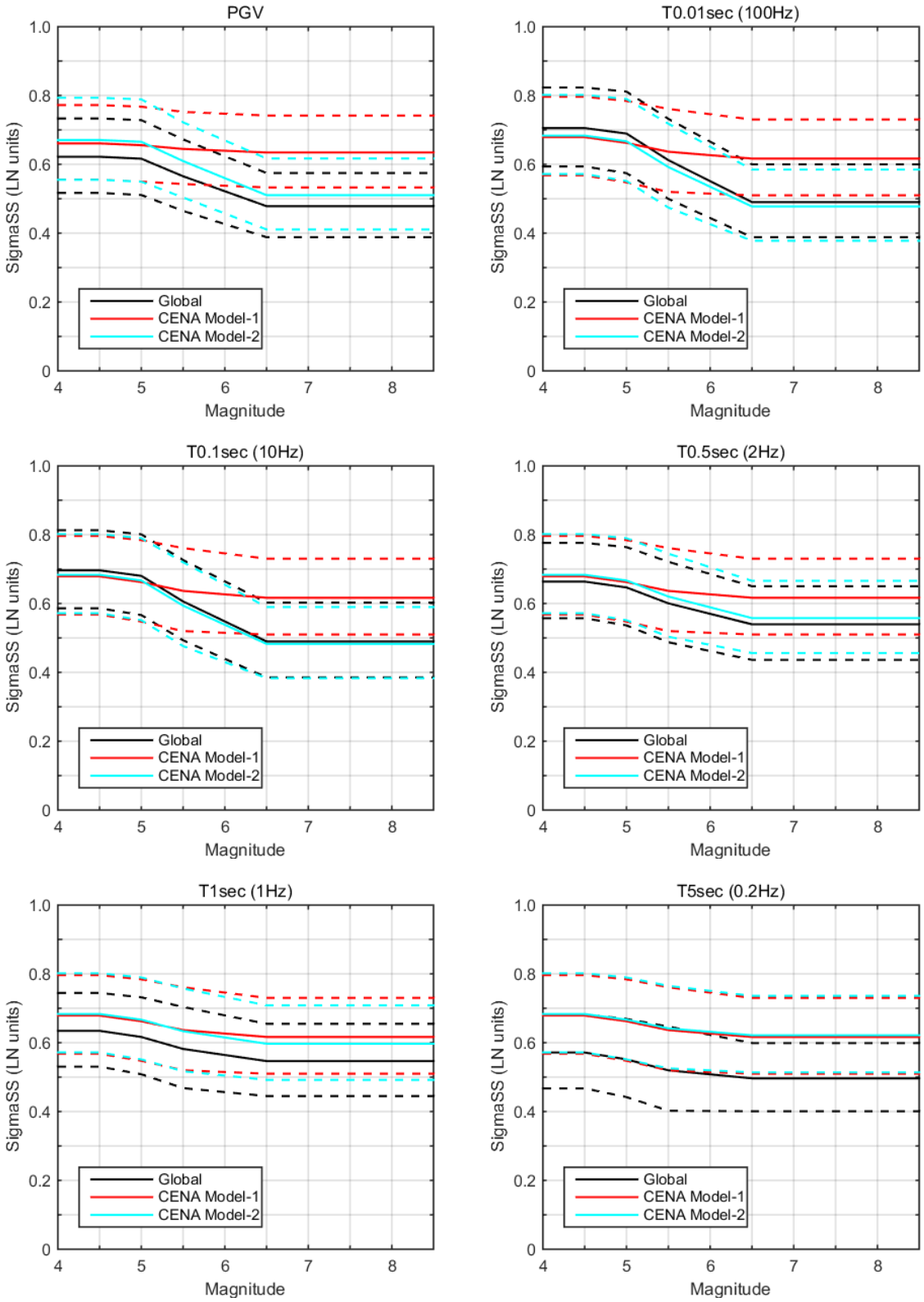


Figure 11–19 Comparison of CENA single-station sigma (σ_{SS}) models versus frequency for M = 5.0, 6.0, and 7.0. Solid lines show the central branches. Dashed lines show the low and high branches.

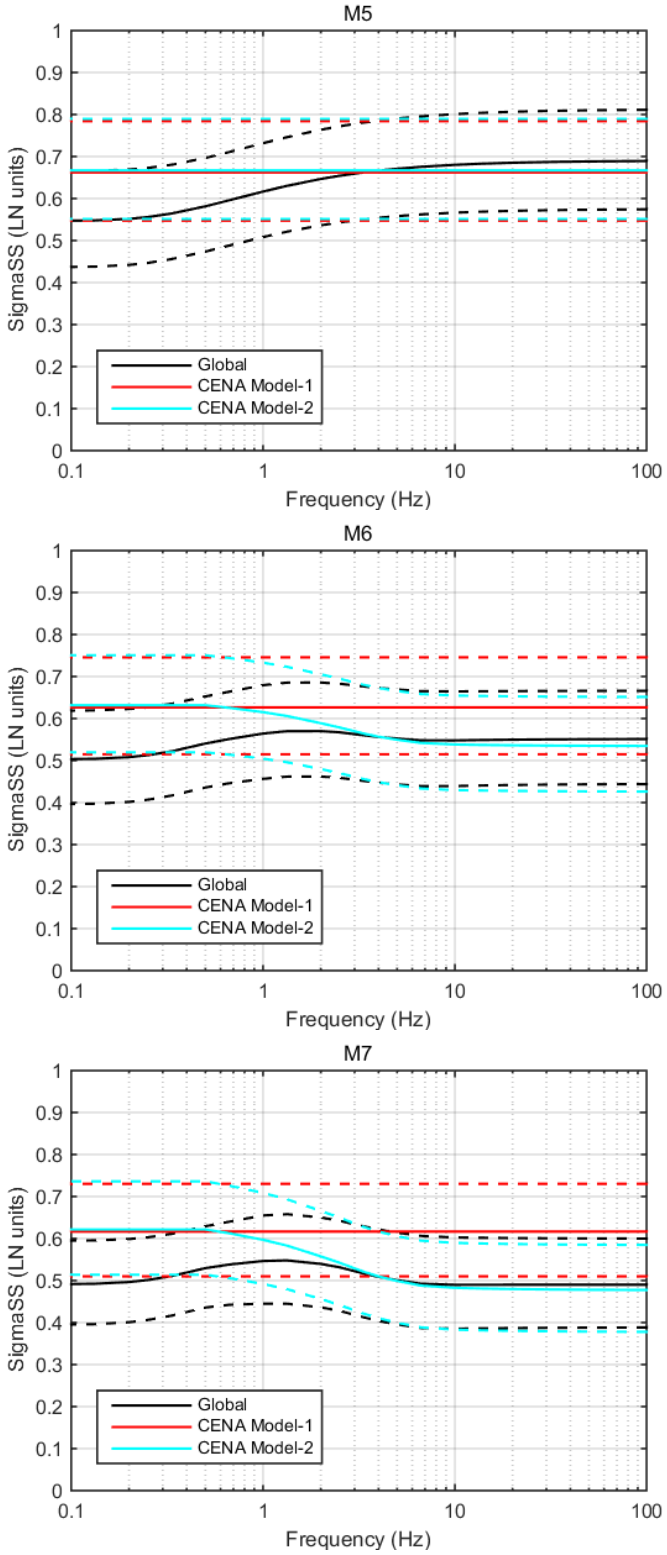


Figure 11–20 CDFs of the three single-station sigma (σ_{SS}) models and their composite for M = 4.5, 5.0, 5.5, and 6.5 at T = 0.01 sec (f = 100 Hz).

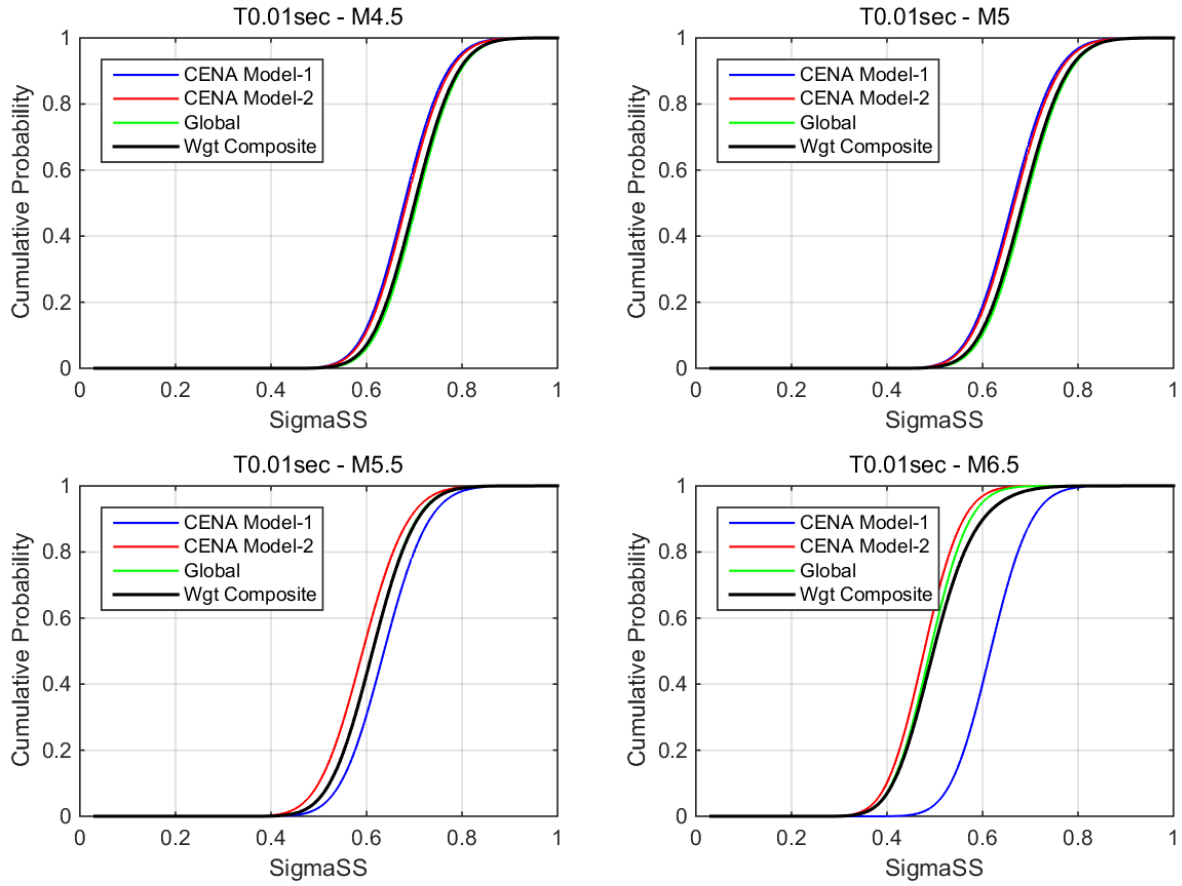


Figure 11–21 DFs of the three single-station sigma (σ_{SS}) models and their composite for M = 4.5, 5.0, 5.5, and 6.5 at T = 1 sec (f = 1 Hz).

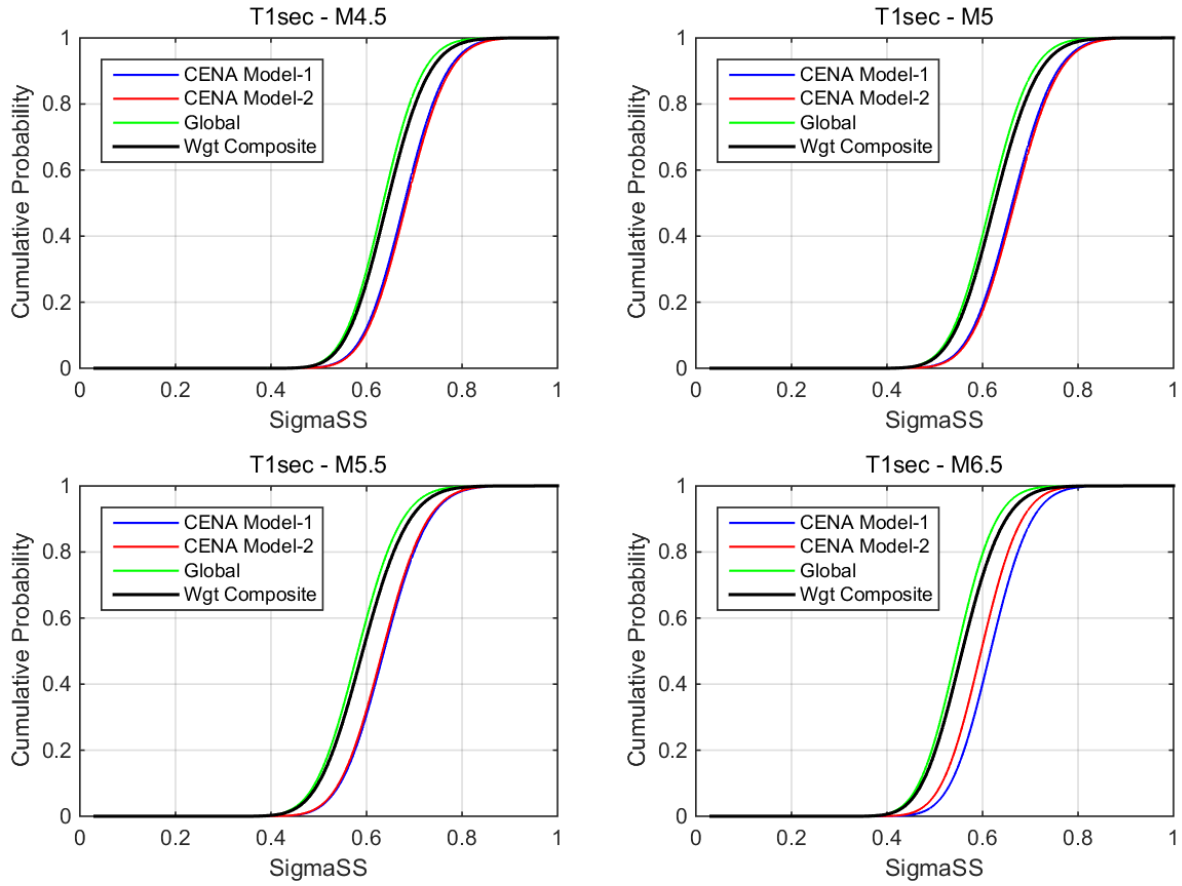


Figure 11–22 Comparison of single-station sigma (σ_{SS}) models for CENA and Hanford versus frequency for M = 5.0, 6.0, and 7.0. Solid lines show the central branches. Dashed lines show the low and high branches.

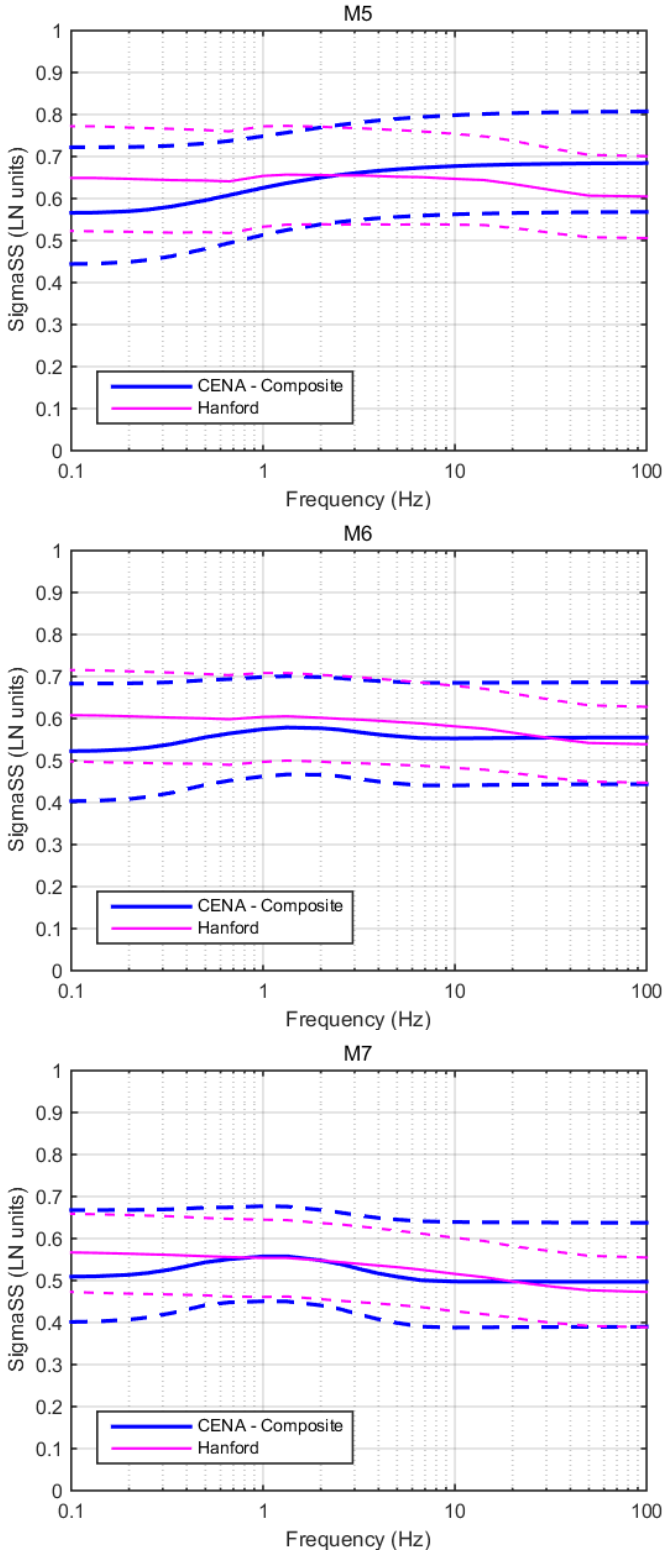


Figure 11–23 Comparison of single-station sigma (σ_{SS}) models for CENA and SWUS versus frequency for M = 5.0, 6.0, and 7.0. Solid lines show the central branches. Dashed lines show the low and high branches.

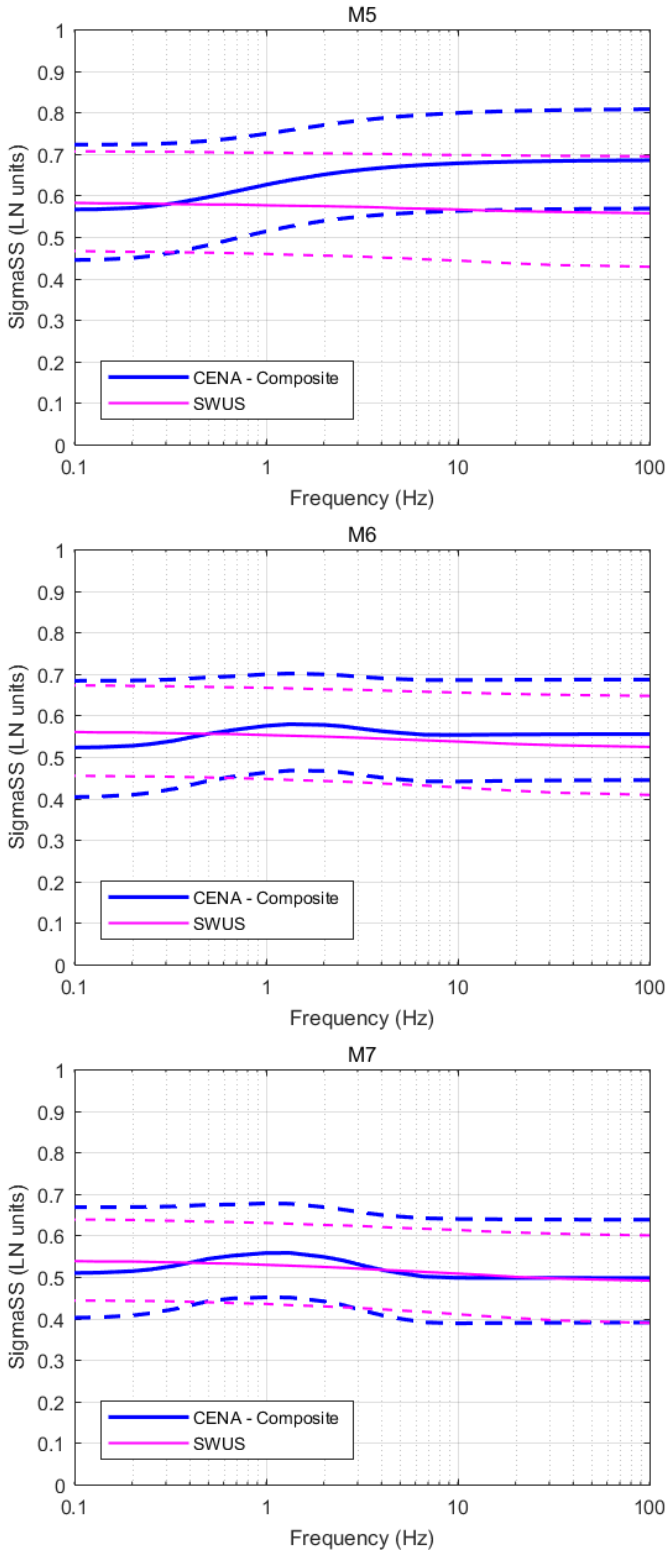
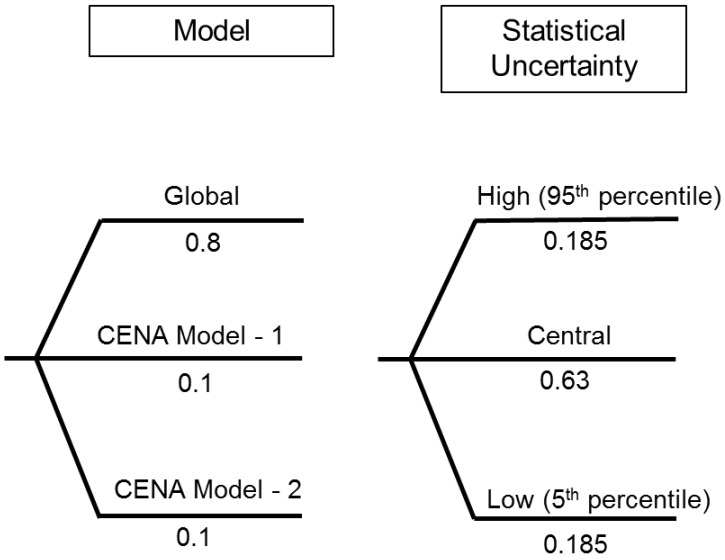


Figure 11–24 Logic tree for total ergodic sigma (σ)[(1)].



(1) Global Model is composed of global Tau (τ), global PhiSS (ϕ_{SS}), and CENA PhiS2S (ϕ_{S2S}) models.

8. CENA Model-1 is composed of global Tau (τ), CENA constant PhiSS (ϕ_{SS}), and CENA PhiS2S (ϕ_{S2S}) models.
9. CENA Model-2 is composed of global Tau (τ), CENA magnitude-dependent PhiSS (ϕ_{SS}), and CENA PhiS2S (ϕ_{S2S}) models.

Figure 11–25 Comparison of ergodic total sigma (σ) models for CENA versus magnitude at PGV and f = 100, 10, 2, 1, and 0.2 Hz. Solid lines show the central branches. Dashed lines show the low and high branches.

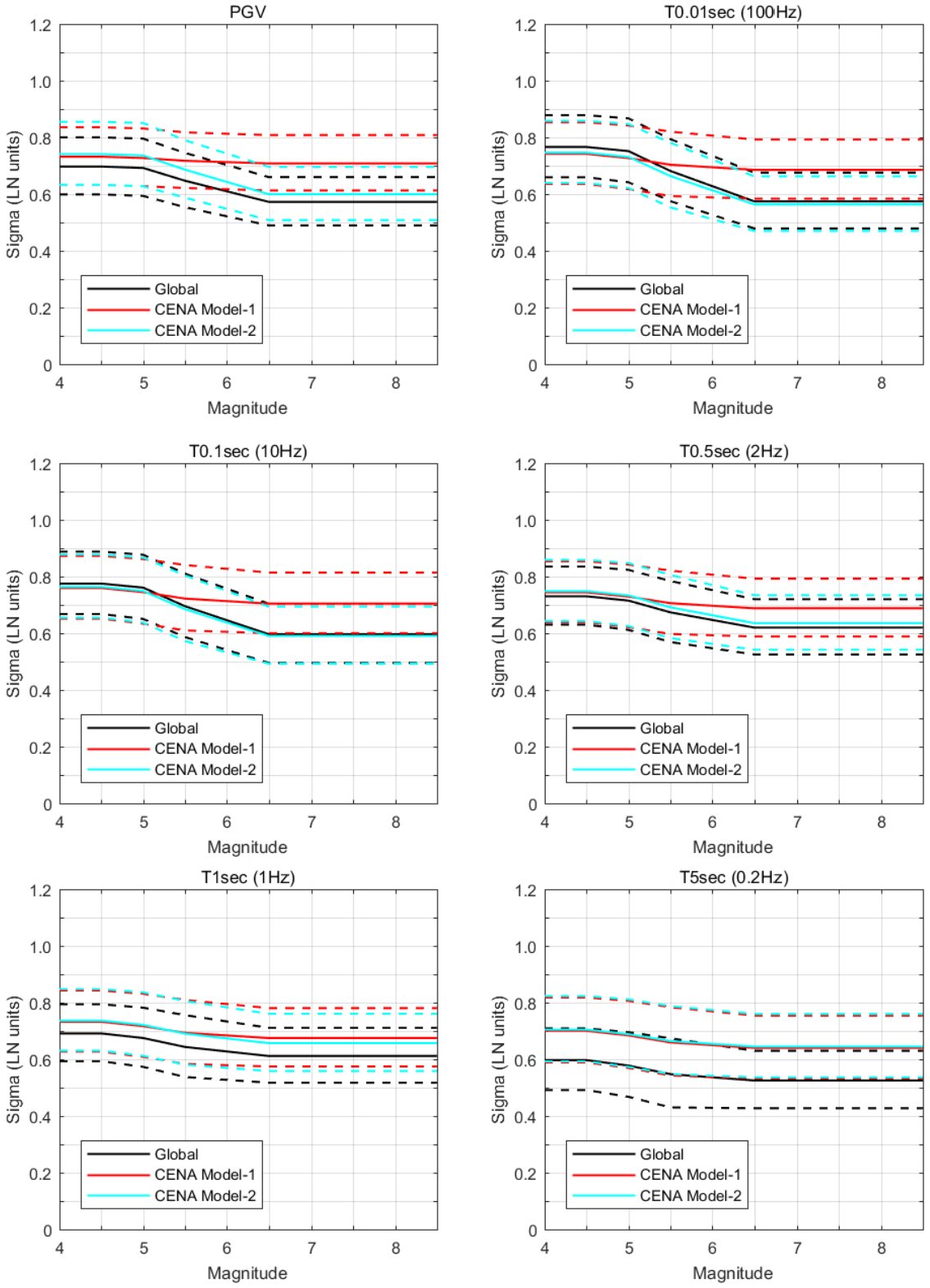


Figure 11–26 Comparison of ergodic total sigma (σ) models for CENA versus frequency for M = 5.0, 6.0, and 7.0. Solid lines show the central branches. Dashed lines show the low and high branches.

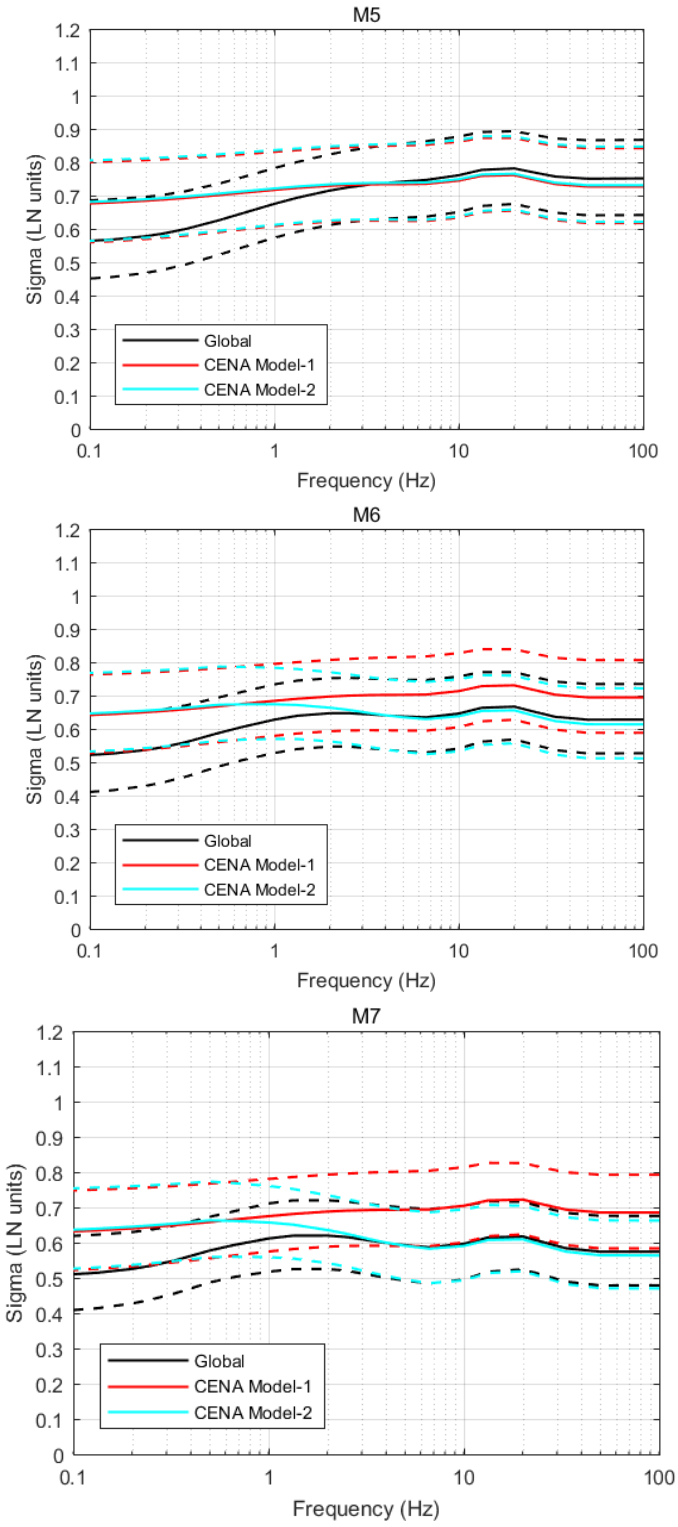


Figure 11–27 CDFs of the three ergodic total sigma (σ) models and their composite for M = 4.5, 5.0, 5.5, and 6.5 at T = 0.01 sec (f = 100 Hz).

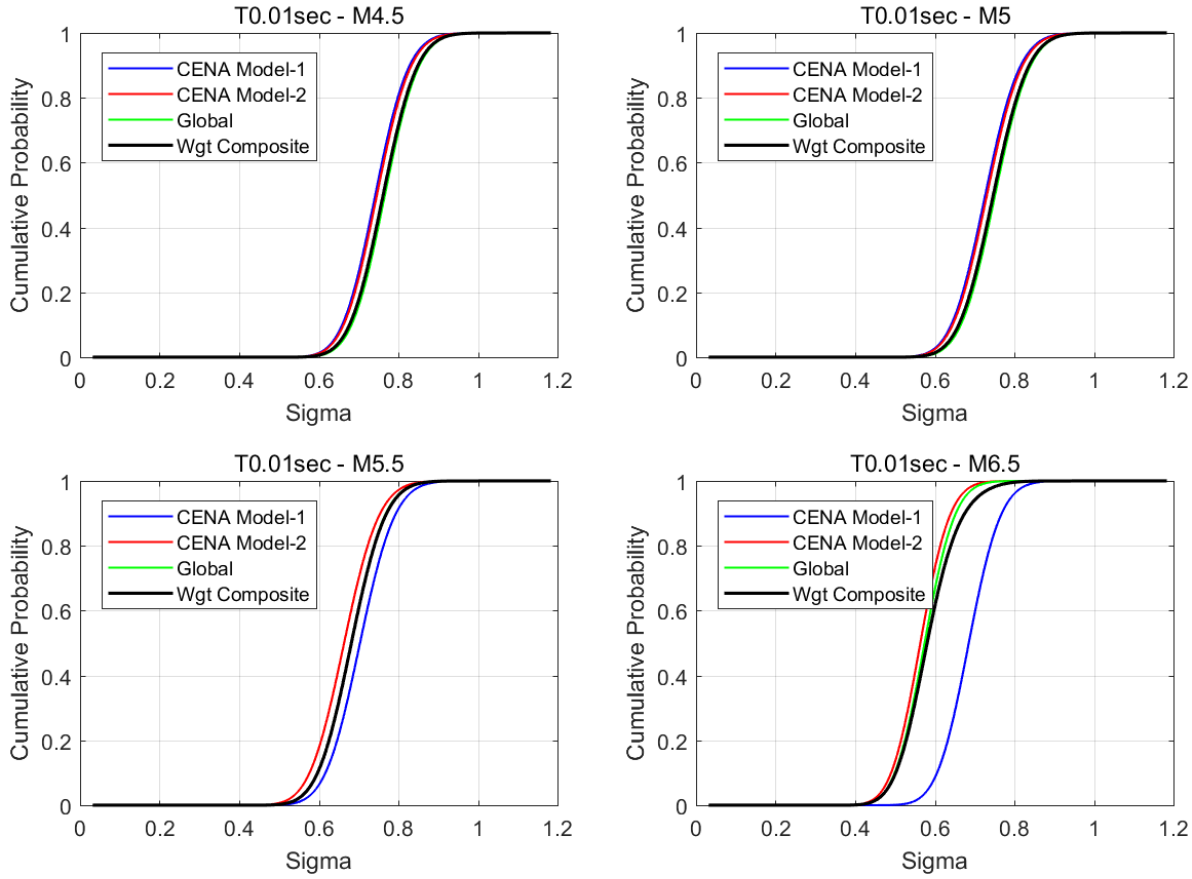


Figure 11–28 CDFs of the three ergodic total sigma (σ) models and their composite for M = 4.5, 5.0, 5.5, and 6.5 at T = 1 sec (f = 1 Hz).

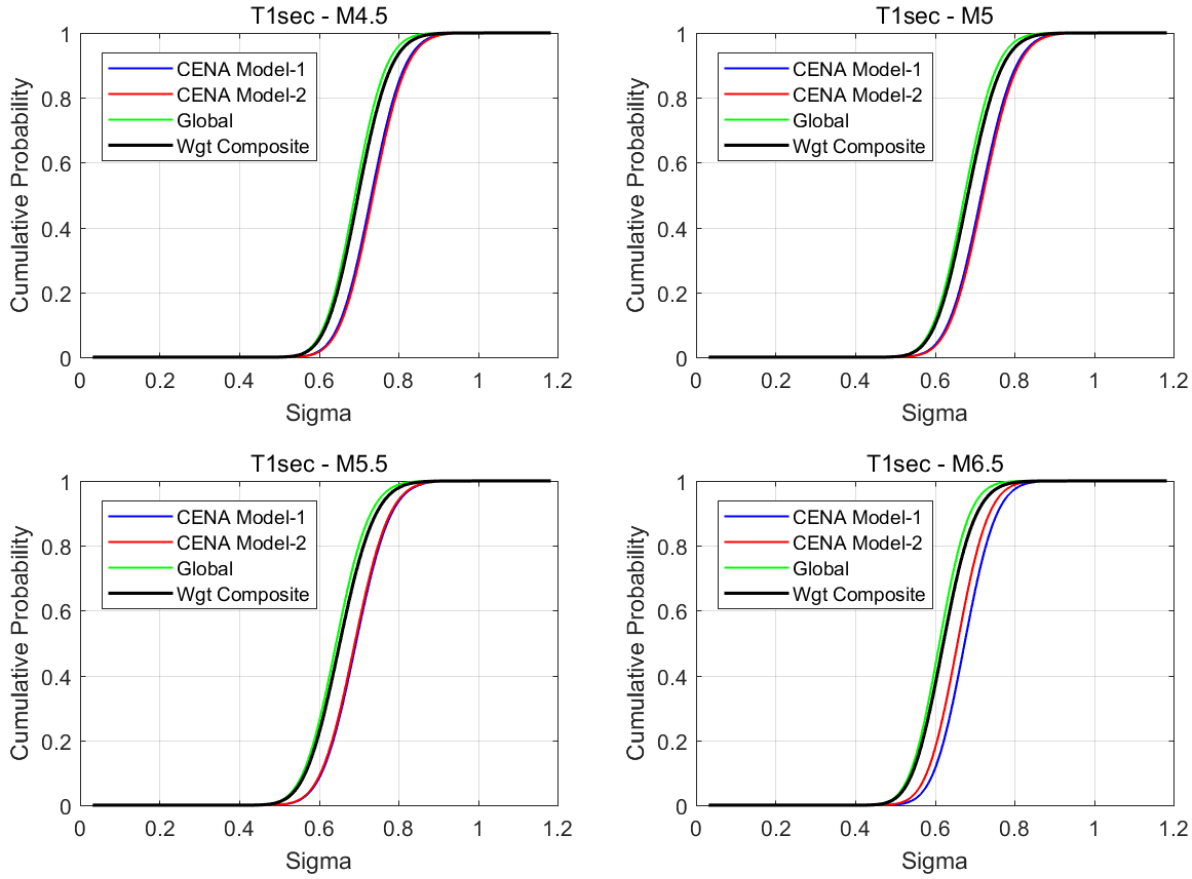


Figure 11–29 Comparison of composite ergodic total sigma (σ) model for CENA and NGA-West2 models versus frequency for M = 5.0, 6.0, and 7.0. The solid line shows the central branch. Dashed lines show the low and high branches.

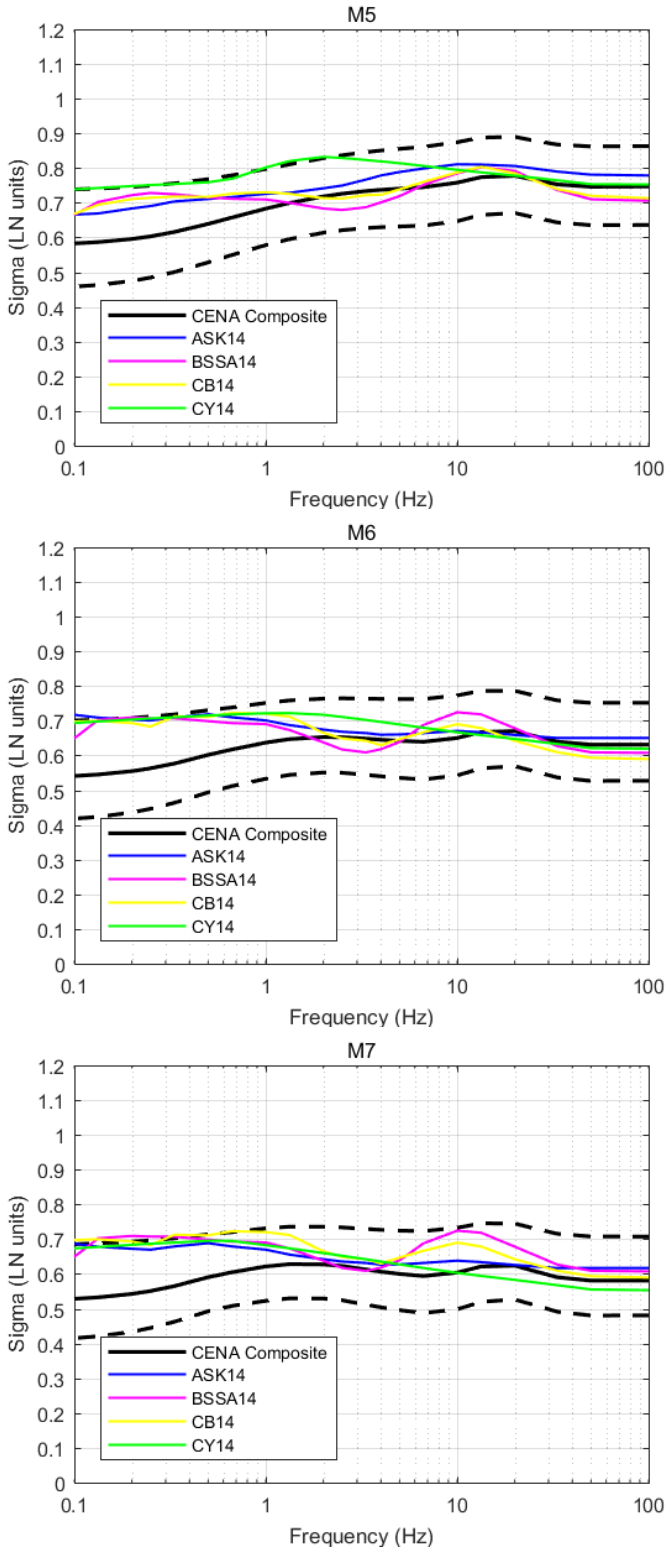


Figure 11–30 Comparison of composite ergodic total sigma (σ) model for CENA to the EPRI 2013 and Atkinson and Adams (2013) σ models versus frequency for M = 5.0, 6.0, and 7.0. The solid line shows the central branch. Dashed lines show the low and high branches.

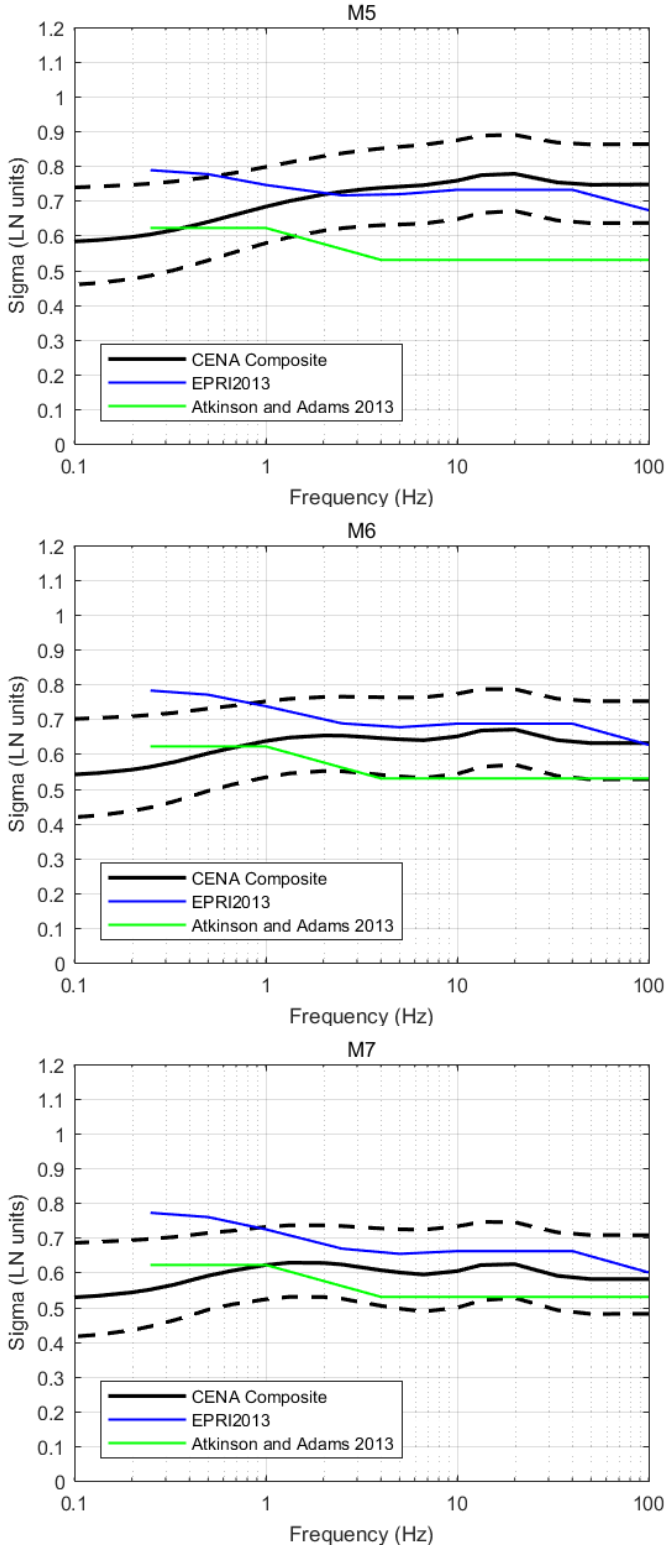


Figure 11–31 Comparison of Tau (τ) values for all CENA earthquakes, PIE only and tectonic only events. Only earthquakes with minimum M of 3.0 were used. Error bars represent one standard error.

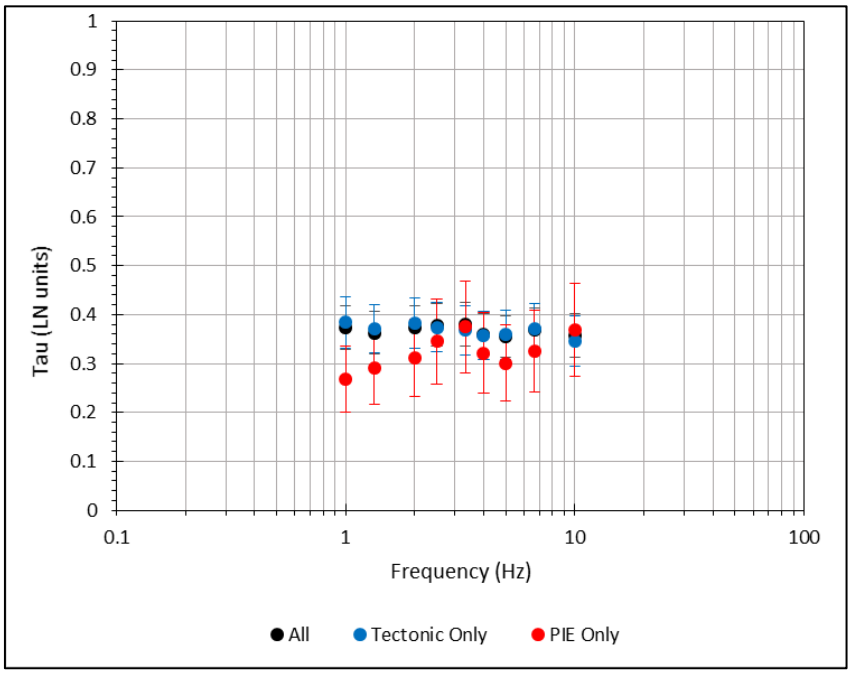


Figure 11–32 Comparison of PhiSS (ϕ_{SS}) values for all CENA earthquakes, PIE only and tectonic only events. Only earthquakes with minimum M of 3.0 were used. Error bars represent one standard error.

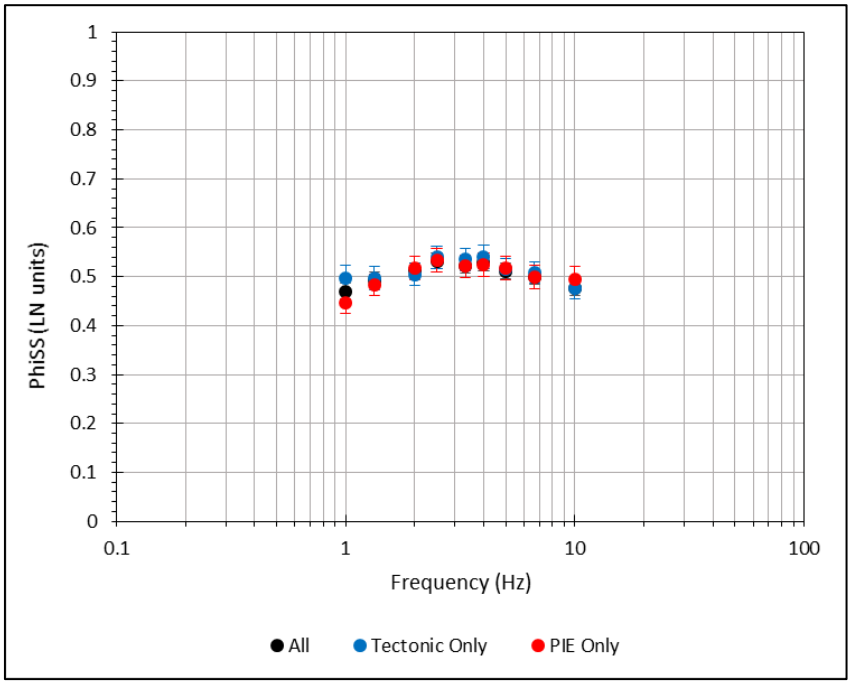


Figure 11–33 Comparison of CENA Tau (τ) models to τ values for the Gulf Coast Region. "CENA Cte" refers to the constant CENA model. Solid lines show the central branches. Dashed lines show the low and high branches. Error bars represent one standard error.

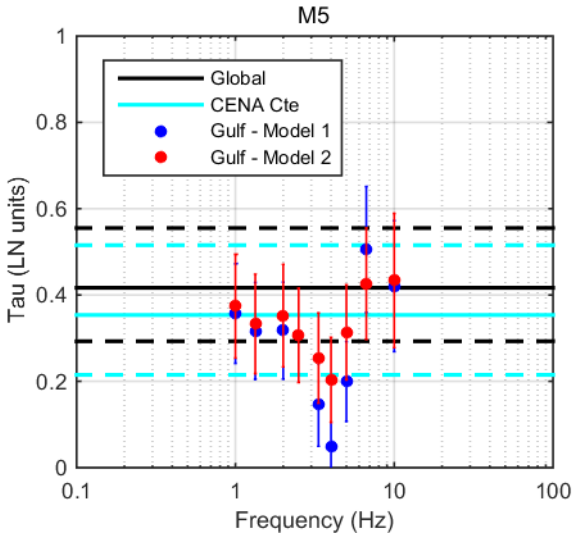


Figure.11–34 Comparison of CENA PhiSS (ϕ_{SS}) models to ϕ_{SS} values for the Gulf Coast region. "CENA Cte" refers to the constant CENA model. Solid lines show the central branches. Dashed lines show the low and high branches. Error bars represent one standard error.

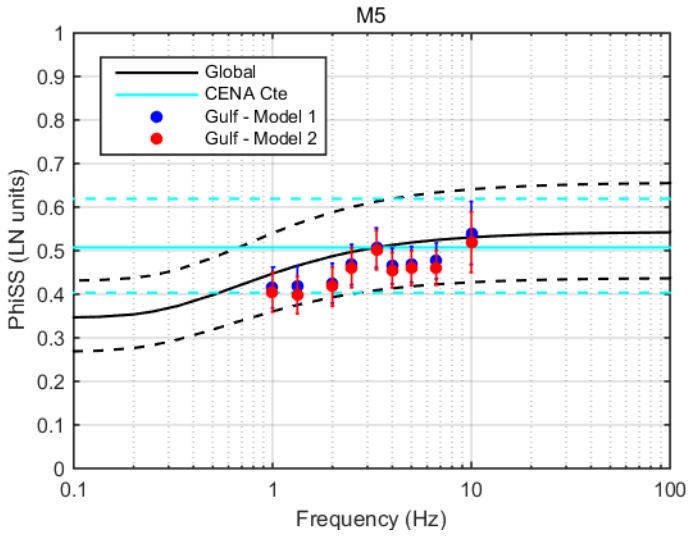


Figure 11–35 Comparison of CENA PhiS2S (ϕ_{S2S}) model unadjusted for small-magnitude bias and hard rock conditions to ϕ_{S2S} values for the Gulf Coast Region. The solid line shows the central branch. Dashed lines show the low and high branches. Error bars represent one standard error.

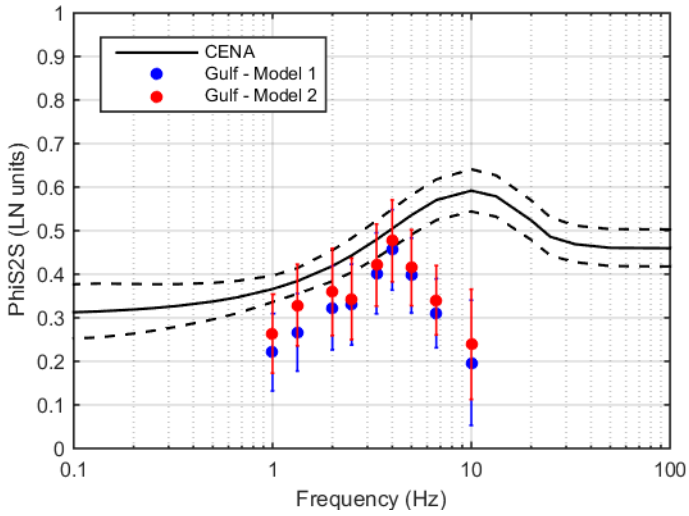
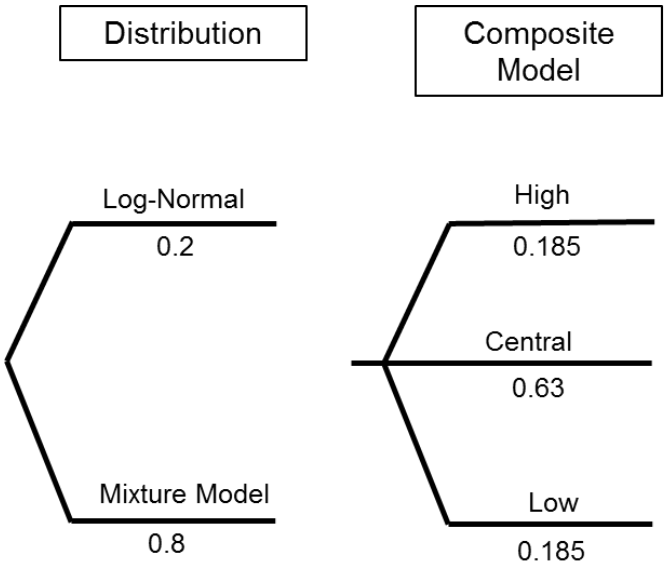


Figure 11–36 Logic tree structure applicable to single-station sigma (σ_{SS}) and ergodic sigma (σ).



12. Median Model: Hazard Results and Comparisons

12.1 Introduction and Outline of Hazard Comparisons

The current state-of-practice for conducting probabilistic seismic hazard analyses (PSHA) for nuclear facilities in the U.S. consists of using the Central and Eastern U.S. Seismic Source Characterization (CEUS SSC) model (EPRI/DOE/NRC 2012) along with the EPRI (2013) ground-motion characterization (GMC) model. The EPRI (2013) GMC was developed to fill the "time gap" between the previous generation of the EPRI GMC model and the incoming NGA-East GMC. Therefore, it is relevant to present PSHA results of NGA-East GMC as compared to those of the EPRI (2013) GMC.

The NGA-East approach for quantifying the ground-motion epistemic uncertainty is relatively new and innovative. The approach is conceptually similar to the approach used for the South Western U.S. (SWUS) utilities project (GeoPentech 2015); however, the methodology has evolved in the NGA-East project. The approach relies on mathematical and statistical modeling not commonly used in the field of seismic ground motions; thus, the NGA-East participatory peer review panel (PPRP) requested a partial demonstration of the approach described in Chapter 8 but using the EPRI (2013) as the basis for the seeds GMMs. The goal of this exercise was not to show that the NGA-East and EPRI (2013) approaches lead to similar results, but to understand the source of the differences, if any, that *may be due to the process itself*.

This chapter presents a series of PSHA results computed using the mean CEUS SSC model (Chapter 3). The focus of the demonstration is on the median ground-motion models (GMMs) and their epistemic uncertainty. As such, all the calculations were performed using NGA-East's base single-station sigma model, central branch (Chapter 11). This eliminates any potential differences due to standard deviations. A suite of hazard curves was developed and compared for the following three cases:

- **Base Model**: NGA-East base case with the 17 final models and weights defined in Chapters 8 and 9.
- **EPRI Seeds**: the individual EPRI (2013) GMMs are used directly as seed GMMs in the NGA-East approach, and the complete NGA-East process is applied (Chapter 8 and 9) to generate 17 weighted GMMs.
- **EPRI (2013):** median EPRI (2013) GMC model with cluster-based epistemic uncertainty.

Section 12.2 summarizes key results from the second bullet item above and Section 12.3 presents hazard results for all three cases.

12.2 Using EPRI (2013) GMMs as Seeds for the NGA-East Process

As elaborated in the previous chapters of this report, there are five steps in the NGA-East approach to quantify the epistemic uncertainty of median ground motion:

1. Develop a suite of seed GMMs
2. Develop parameters for continuous distributions of GMMs

3. Visualize the ground-motion space and sample GMMs
4. Re-discretize the ground-motion space
5. Assign weights

For the current computation demonstration, Step 1, which requires an evaluation of the available median GMMs, is completely skipped, and the ten original EPRI (2013) GMMs are used directly (Table 12–1). Steps 2–5 are then applied exactly as they were for the NGA-East project, and using the same "rules," i.e., the same correlation model was used to generate a continuous distribution of GMMs from which the 10,000 GMMs were sampled (Section 8.1). The same physicality rules were also applied to the sampled models (Section 8.1.5).

Note that the EPRI GMMs are only defined for seven ground-motion intensity measures (GMIMs): pseudo-spectral acceleration (PSA) at six oscillator frequencies (0.5, 1, 2.5, 5, 10, and 25 Hz), and peak ground acceleration (PGA).

Figure 12–1 shows the rejection rate of sampled GMMs using the EPRI seeds compared to the NGA-East seed GMMs. The trends are generally similar, with more models rejected for both cases due to the magnitude-scaling criterion, and the rate of rejection is generally comparable between the two sets of seeds for the distance-scaling criterion. These rates depend how closely the seed models are to the criterion to start with; and from the modeled correlation, which dictates how wide the continuous distribution of GMMs is. The large rejections of models sampled from the NGA-East seeds is due to the relatively low slopes of scaling with magnitude at close distance (see Section 9.5 for a discussion of close distance issues). Because the magnitude scaling slopes from the NGA-East seeds tend to be lower (Figure 12–2), models sampled from these seeds sometimes have a null or negative slope and are rejected based on the physicality criteria.

Using the same scenarios as for the NGA-East, the sampled models are used to generate Sammon's (1969) maps (as described in Section 8.3). The ground-motion space shown on the Sammon's maps is then defined using the same scaling rules and re-discretized into 17 cells. Figure 12–3 compares maps from NGA-East and the current demonstration computations for three GMIMs. Recall from Chapter 8 that although direct trends are difficult to extract from Sammon's maps, two key points can be made to help the interpretation:

- Due to the manner in which all the maps are rotated and aligned, the ground-motion values vary from small to large as from the left to right on the map. If a single GMM was scaled up and down, it would show as a series of points along the x-axis.
- Trends due to different magnitude and distance scaling affect the y-axis or vertical direction. Different scaling trends actually act along some diagonal axes.

Figure 12–3 shows that the two sets of maps lead to different results in terms of sampled GMMs. The weights are then computed for the EPRI-based models following the approach described in Chapter 9. As for the NGA-East models, most of the weight associated to each cell's mean model is derived from the number of models in that cell. Maps of residuals and likelihood are shown in Figures 12–4 and 12–5 for the case using EPRI seed models. The total weights are shown in Figure 12–6.

12.3 Discussion of Trends in Sammon's Map Space

The EPRI set include half the number of seed models and does not provide a smooth distribution of sampled GMMs between the seeds (Figure 12–3). The differences between the two cases are completely due to the choice of seed models and their modeled correlation across the various (**M**, R_{RUP}) scenarios used to compute the maps. In general, the maps based on the NGA-East seed GMMs tend to be smoother and filling the ellipse in a more continuous fashion. This suggests that the NGA-East seeds (and their overall distribution) tend to sample more variation in scaling than the EPRI seeds, as the blue dots show on both sets of maps in Figure 12–3. There is some clustering around the NGA-East seeds as well, but all the cells are populated, contrary to the case for the EPRI seeds.

The clustering is the result of: (1) the lower magnitude scaling variability relative to NGA-East (the red dots representing the magnitude reference models are close together on the map); (2) the larger distance scaling variability (the orange dots representing the distance scaling reference models are spaced out on the maps) relative to the NGA-East case; and (3) the imposed covariance used, which was defined for the NGA-East seeds (Figure 8–11). We expand on these points below.

The magnitude and distance scaling reference models provide insight into the clustering trends observed at certain frequencies. For example, Figure 12–3 shows that for the EPRI case at 1 Hz, the reference magnitude models (red) all sit close together while the distance reference models (orange) span a wider Sammon distance. At 10 Hz, the magnitude and distance reference models for EPRI span comparable distances and cover a wider range in the Sammon map space. The clustering in the EPRI Sammon maps is also marked at 1 Hz but not at 10 Hz.

To explain these trends, we use the differences in the slopes of magnitude and distance scaling from the two sets of seeds (NGA-East and EPRI), using the 1Hz case for the illustration. The mean and the standard deviation of differences in ground-motion predictions at **M** 6.5 and **M** 6 are calculated as proxies for the magnitude slopes. This provides an indication of the range in different magnitude slopes among the models. Since the difference can depend on the distance values where the difference is calculated, this process is repeated for all scenario distances used for the calculation of the Sammon's maps, leading to the production of Figure 12–7 (top). To get an indication of the range in distance slopes between the seed models, the standard deviations of differences in predictions at 1000 km and 1500 km are calculated, this time for all scenario magnitudes used in the calculation of the Sammon's maps. Figure 12–7 show that on average the range in magnitude slopes is larger for the NGA-East seeds than for the EPRI seeds, while it is the other way around for the distance slopes. Hence, the difference in distance slopes dominate the differences between models for EPRI compared to NGA-East. This is expressed in the 1 Hz Sammon's maps by the clustering of the reference-magnitude models (red dots) and the relatively large spacing of the reference-distance models (orange dots). This is also expressed in the models sampled from those seeds (grey dots).

The scaling of the ellipse following the assumption of a bivariate normal distribution with the imposed covariance developed for NGA-East seeds (Chapter 8 discussion and Figure 8–11) does not seem appropriate for the EPRI case, especially for the 1 and 100 Hz PSA. Recall from Chapter 8 and 9 that the imposed covariance was tuned to achieve a target (desired)

covariance defined by the TI team. During the sampling process that covariance is combined with that originating from the original seeds into a model designed to achieve the target covariance. To obtain a more even distribution of EPRI samples in Sammon's map space, we would have had to increase the imposed covariance to compensate for the different covariance of the EPRI seeds with magnitude and distance discussed above. However, the combination of covariances is not a linear process because we apply physicality criteria in the sampling process, so it's not possible to simply solve for the variance to be added. The TI team has not performed any tuning when applying the NGA-East process to the EPRI seeds. This illustrates that the NGA-East approach described in this report is indeed applicable to other sets of seeds, but that it is not a black box to be blindly applied. This is highlighted in Figure 6–18 whereby the purple diamond shapes indicate key steps that require expert input from the TI team (this step is called "Model and define variance and correlation" under Step 2). Indeed, the covariance model required to obtain a close to mutually exclusive, completely exhaustive (MECE) distribution of ground-motion models depends on the starting seed models, their number, range, and inherent covariance.

12.4 Hazard Results and Observations

The three cases described in Section 12–1 are compared in hazard space for different GMIMs. The hazard curves are calculated for two demonstration sites: Central Illinois and Savannah (Figures 3–3 and 3–4). The Savannah site is in close proximity to the Charleston Repeatable Large Magnitude Earthquake (RLME) zone, while the Central Illinois site is far from its closest RLME (the Wabash Valley zone). Those attributes make the two sites representative of most sites located in Central and Eastern North America (CENA).

Figures 12–8 to 12–14 show hazard curves for the following GMIMs, respectively: 0.5, 1, 2.5, 5, 10, and 25 Hz PSA, and PGA. On these figures, the left panel shows hazard curves for the Savannah site and the right panel shows hazard curves for the Central Illinois site. Each panel shows three plots. The top left plot shows hazard curves for the distributed sources, the right plots shows hazard curves for RLMEs, and the bottom plot shows the hazard from both types of sources. Hazard curves for the three cases described in Section 12.1 are presented with the "Final Models" in black, the "EPRI Seeds" in red and the "EPRI (2013)" in blue. For each case, the mean hazard and the 5^{th} and 95^{th} percentiles are also shown. Figures 12–15 and 12–16 show mean hazard ratio plots for all the frequencies shown in Figures 12–8 to 12–14 for Central Illinois and Savannah, respectively.

General observations on ground motions are as follows (focusing on annual exceedance frequencies, annual exceedance frequencies (AEFs), smaller than 10^{-3}, which are most relevant for design of critical facilities):

Mean hazard:

- 0.5–25 Hz: the EPRI (2013) motions tend to be smaller than for the EPRI Seeds, with differences becoming smaller with increasing oscillator frequency. The means from the Final Models (NGA-East) tend to show the largest ground motions.

Lower-bound (5th percentile):

- 0.5–10 Hz: EPRI Seeds motions < EPRI (2013) < Final Models motions.
- 25 Hz: EPRI Seeds motions are smaller than the EPRI (2013) and Final Models motions, which tend to be similar, but only at low AFEs (less than 10^{-5}).
- PGA: EPRI Seeds motions < final models motions < EPRI (2013) motions

Upper-bound (95th percentile):

- All GMIMs: Final Models motions *tend* to be larger than the EPRI Seeds motions, and both are larger than the EPRI (2013) motions.

Additional trends:

- For all GMIMs, the 95th percentile motions from EPRI (2013) tend to be equal or smaller than the mean Final Models motions. As the frequency increases, these differences become visible only at very low AEFs.
- The range of ground motions captured by the Final Models and the EPRI Seeds is very similar; overall, the EPRI Seeds ground motions are shifted to lower values relative to the Final Models motions. The range from EPRI (2013) tends to be narrower than for the other two models and also spans lower ground motions.

12.5 Discussion

The hazard results summarized in Section 12.3 provide insight into contributions from: (1) different seed GMMs and (2) different processes for quantifying the epistemic uncertainty in median ground motions. Recall that all three sets of hazard curves were computed using the same standard deviation model. The difference between the Final Models (NGA-East) and the EPRI Seeds is only the seed GMMs, while the difference between EPRI Seeds and EPRI (2013) is only in the process.

Observations from above show that the same process applied to different seeds (NGA-East and EPRI) lead to a generally similar range in ground motions. The absolute level is different and can be explained by differences in seeds. It was also shown that the Sammon's maps were different due to major differences in seeds, and by the smaller number of seeds in the case of EPRI. The weights obtained from both sets of seed GMMs were generally comparable, for the exception of cells not populated by models as described above.

It is difficult to fully explore the differences in process between EPRI (2013) and EPRI Seeds. The quantification of epistemic uncertainty requires that a process be designed and evaluated as a whole. However, the results presented here are informative, and show that the NGA-East approach tends to provide a broader range of ground motions and larger estimates depending on the GMIM and AFE. It is interesting to note than the mean ground motions from the case of using EPRI Seeds are generally between the mean ground motions for the EPRI (2013) and the base-model cases.

12.6 Reference

Atkinson, G.M. 2008. “Ground Motion Prediction for Eastern North America from a Referenced Empirical Approach: Implications for Epistemic Uncertainty.” Bull. Seismol. Soc. Am., 98(3): 1304–1318.

Atkinson, G.M., and D.M. Boore. 2006. Earthquake Ground-Motion Prediction Equations for Eastern North America.” Bull. Seismol. Soc. Am. 96, 2181–2205. (also see the erratum published in Vol. 97, No. 3, p. 1032).

Atkinson, G.M., and D.M. Boore. 2011. “Modification to Existing Ground-Motion Prediction Equations in Light of New Data.” Bull. Seismol. Soc. Am., 101(3): 1121–1135.

EPRI. 2013. “EPRI 2004-2006 ground-motion model (GMM), Final Report Project 3002000717.” June 2013, Electric Power Research Institute, Palo Alto, CA.

EPRI/USDOE/USNRC. 2012. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Frankel, A., C. Mueller, T. Barnhard, D. Perkins, E.V. Leyendecker, N. Dickman, S. Hanson, and M. Hopper. 1996. “National Seismic-Hazard Maps: Documentation June 1996.” U.S. Geological Survey, Open-File Report 96-532.

GeoPentech. 2015. “Southwestern United States Ground Motion Characterization SSHAC Level 3 – Technical Report Rev.2.” March 2015.

Pezeshk, S., A. Zandieh, and B. Tavakoli. 2011. “Hybrid Empirical Ground-Motion Prediction Equations for Eastern North America using NGA Models and Updated Seismological Parameters.” Bull. Seismol. Soc. Am., 101(4): 1859–1870.

Sammon, J.W. 1969. A Nonlinear Mapping for Data Structure Analysis. *IEEE Trans. Comp.*, C-18:401–409.

Silva, W.J., N. Gregor, and R.B. Darragh. 2002. “Development of Regional Hard Rock Attenuation Relations for Central And Eastern North America.” Pacific Engineering and Analysis Report, El Cerrito, CA.

Somerville, P.G., N. Collins, N.A. Abrahamson, R.W. Graves, and C. Saikia. 2001. “Ground Motion Attenuation Relations for the Central and Eastern United States.” Report to U.S. Geological Survey, NEHRP Award No. 99HQGR0098, 36 pgs.

Toro, G.R., N.A. Abrahamson, and J.F. Schneider. 1997. A Model of Strong Ground Motions from Earthquakes in Central and Eastern North America: Best Estimates and Uncertainties. Seismol. Res. Lett., 68(1): 41–57.

Table 12–1 EPRI (2013) review project GMMs.

Cluster	Model types	Reference
1	Single-corner Brune source	Silva et al. (2002), SCCS-Sat Silva et al. (2002), SCVS Toro et al. (1997) Frankel et al. (1996)
2	Complex/empirical source; ~R^{-1} geometrical spreading R < 70 km	Silva et al. (2002), DC-Sat A08′: Atkinson (2008) with modifications from Atkinson and Boore (2011), and EPRI
3	Complex/empirical source; ~$R^{-1.3}$ geometrical spreading R < 70 km	AB06′: Atkinson and Boore (2006) with modifications from Atkinson and Boore (2011) Pezeshk et al. (2011)
4	Finite-source; Full waveform Green's functions	Somerville et al. (2001), two models for rifted and non-rifted regions

Figure 12–1 Number of rejected sampled GMMs from EPRI seeds based on NGA-East criteria, for a suite of frequencies (100 Hz is used to represent PGA).

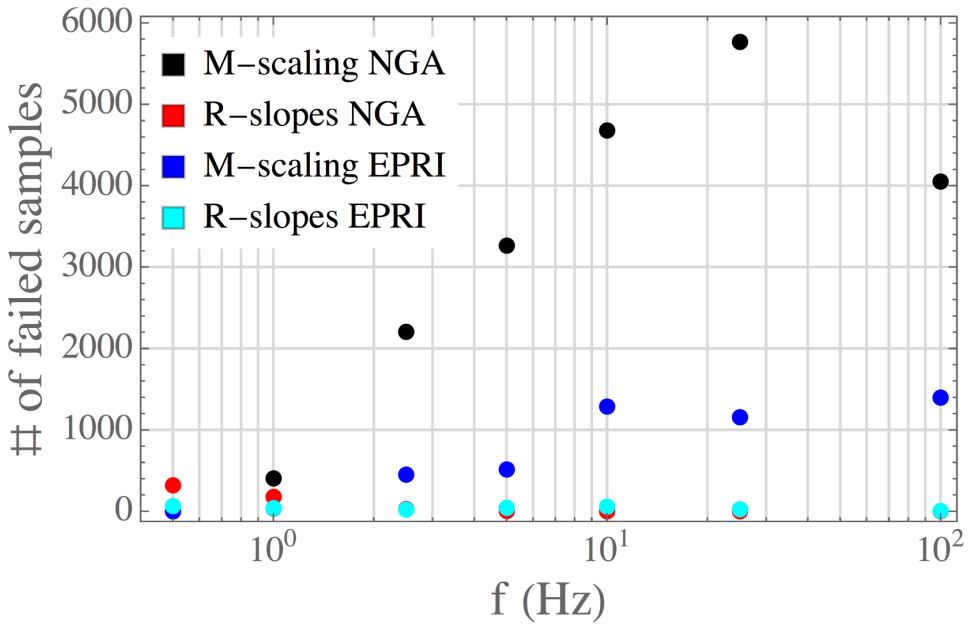


Figure 12–2 Magnitude scaling slopes for close-in distance (R_{RUP} of 5 km) for the NGA-East and EPRI seeds. The NGA-East seeds tend to have low scaling with magnitude.

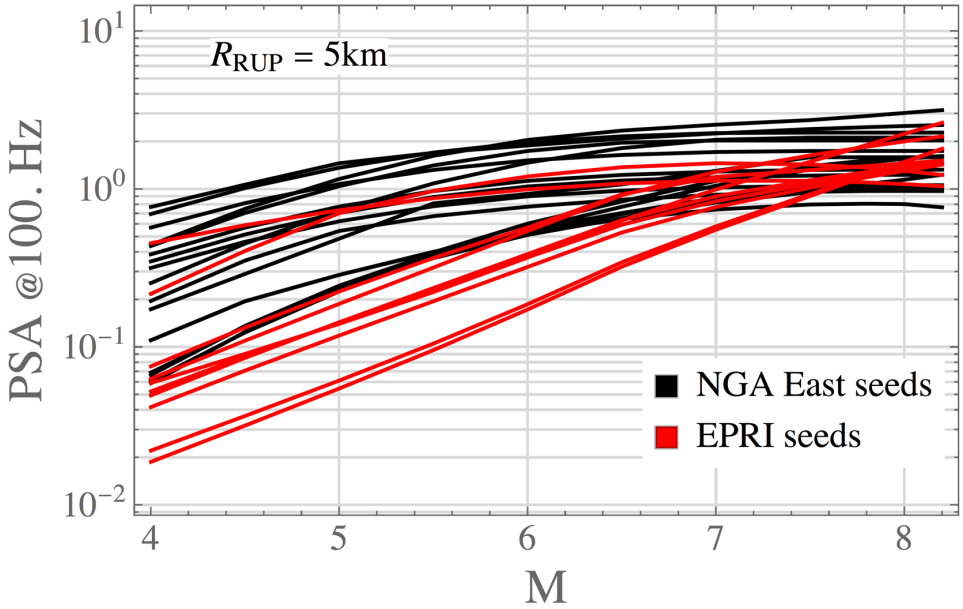


Figure 12–3 Comparison of Sammon's maps from the NGA-East case (left column) to the test case for which EPRI (2013) GMMs are used as seeds (right column); red symbols for magnitude scaling and orange symbols for distance scaling.

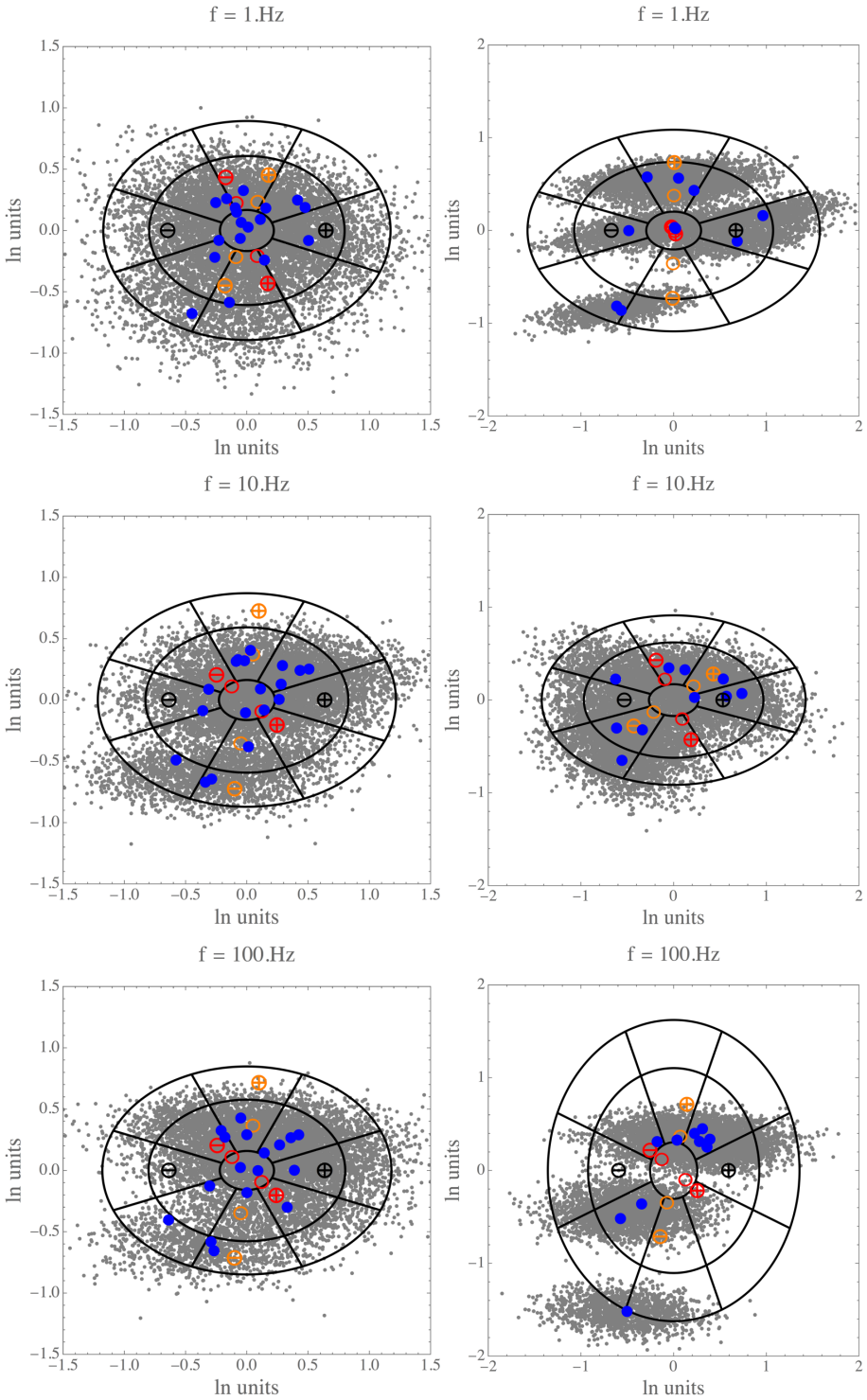


Figure 12–4 Contour plots of mean mean residuals for the sampled models based on the EPRI (2013) seed GMMs. This figure is the equivalent of Figure 9–5, developed for the NGA-East seed models.

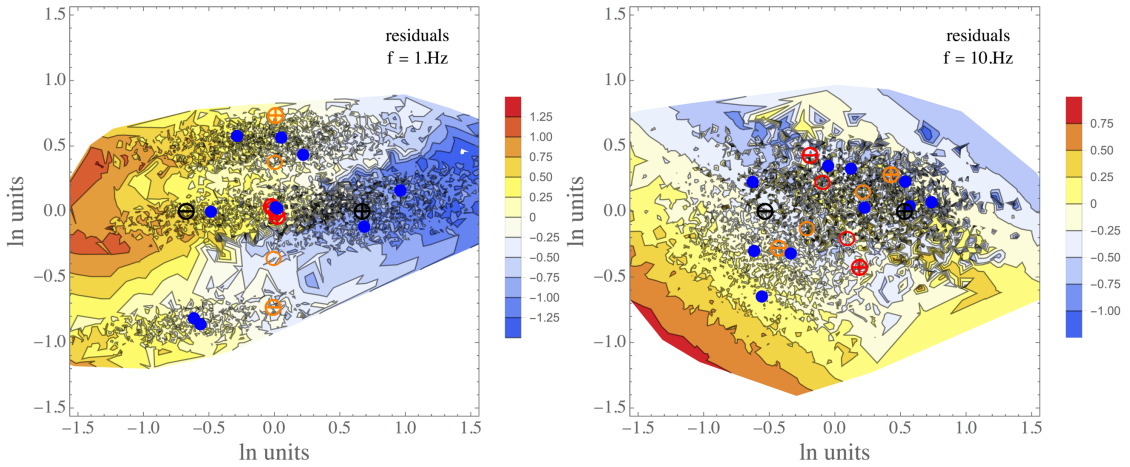


Figure 12–5 This figure is the equivalent of Figure 9–6, developed for the NGA-East seed models.

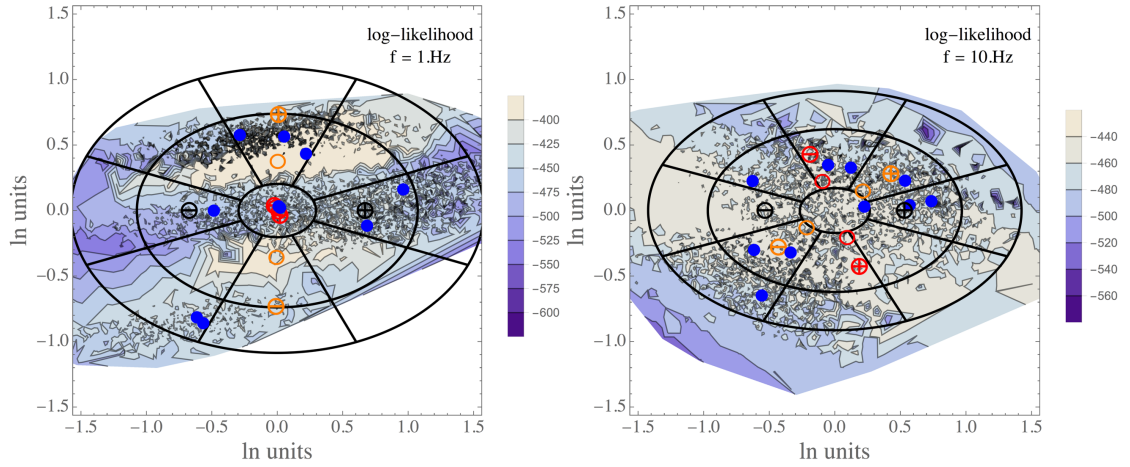


Figure 12–6 Total weights for the sampled models based on the EPRI (2013) seed GMMs. This figure is the equivalent of Figure 9–7, developed for the NGA-East seed models.

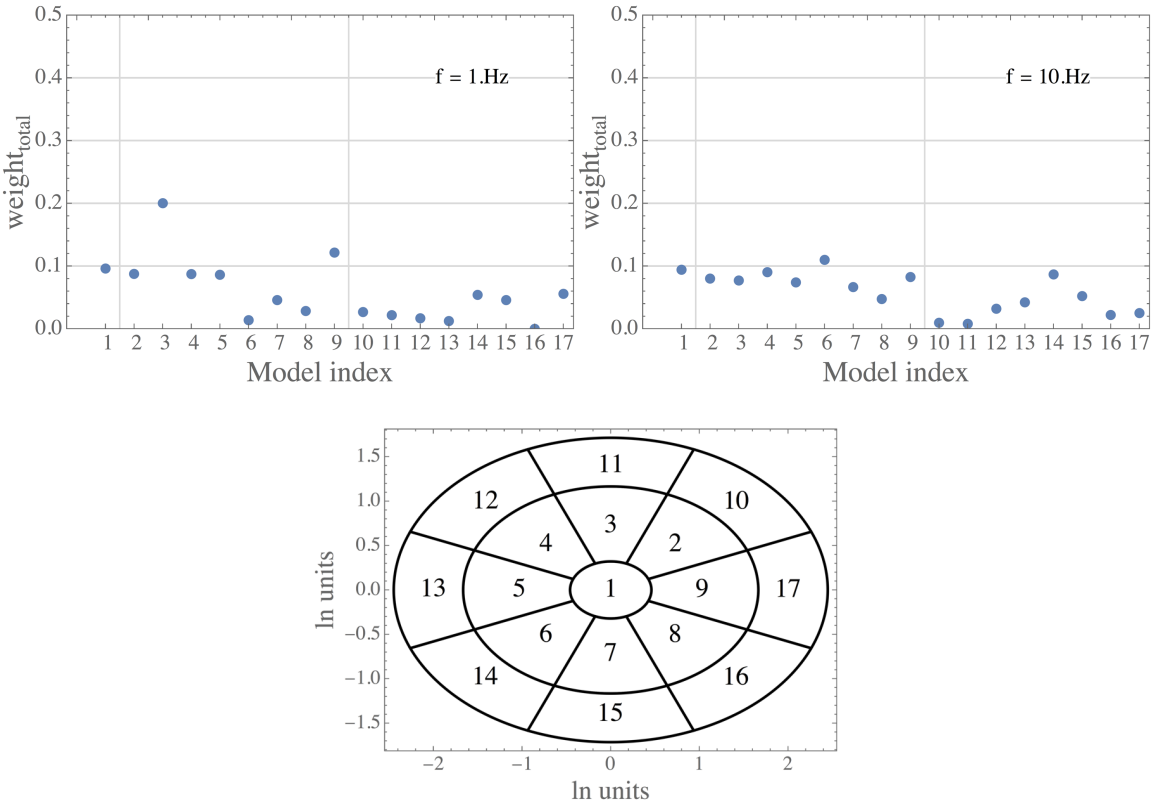


Figure 12–7 Standard deviation of ground-motion proxy for (top) magnitude and (bottom) distance scaling for the Sammon's maps scenario for NGA-East and EPRI seeds.

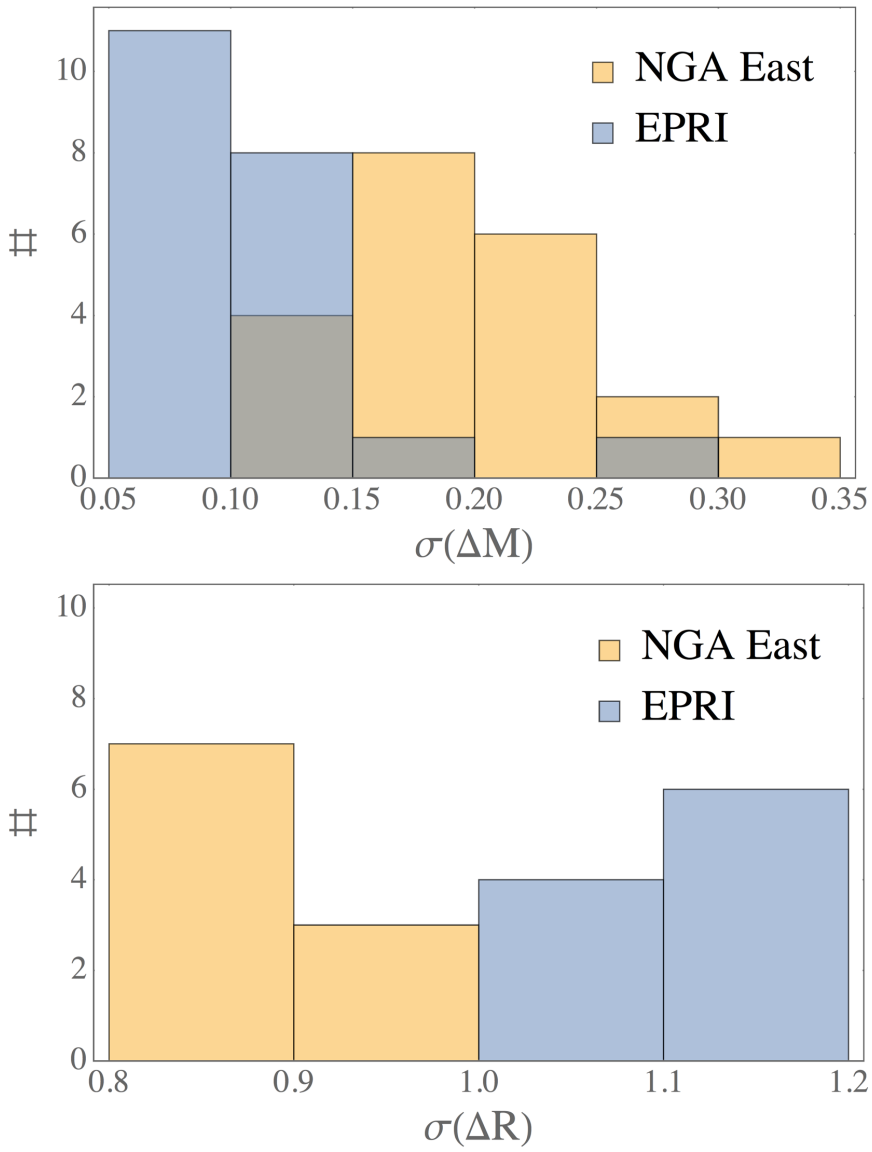


Figure 12–8 Comparison of hazard curves from variants of EPRI (2013) median GMMs to the NGA-East base case GMMs, for 0.5 Hz PSA at the Savannah site (left) and the Central Illinois site (right).

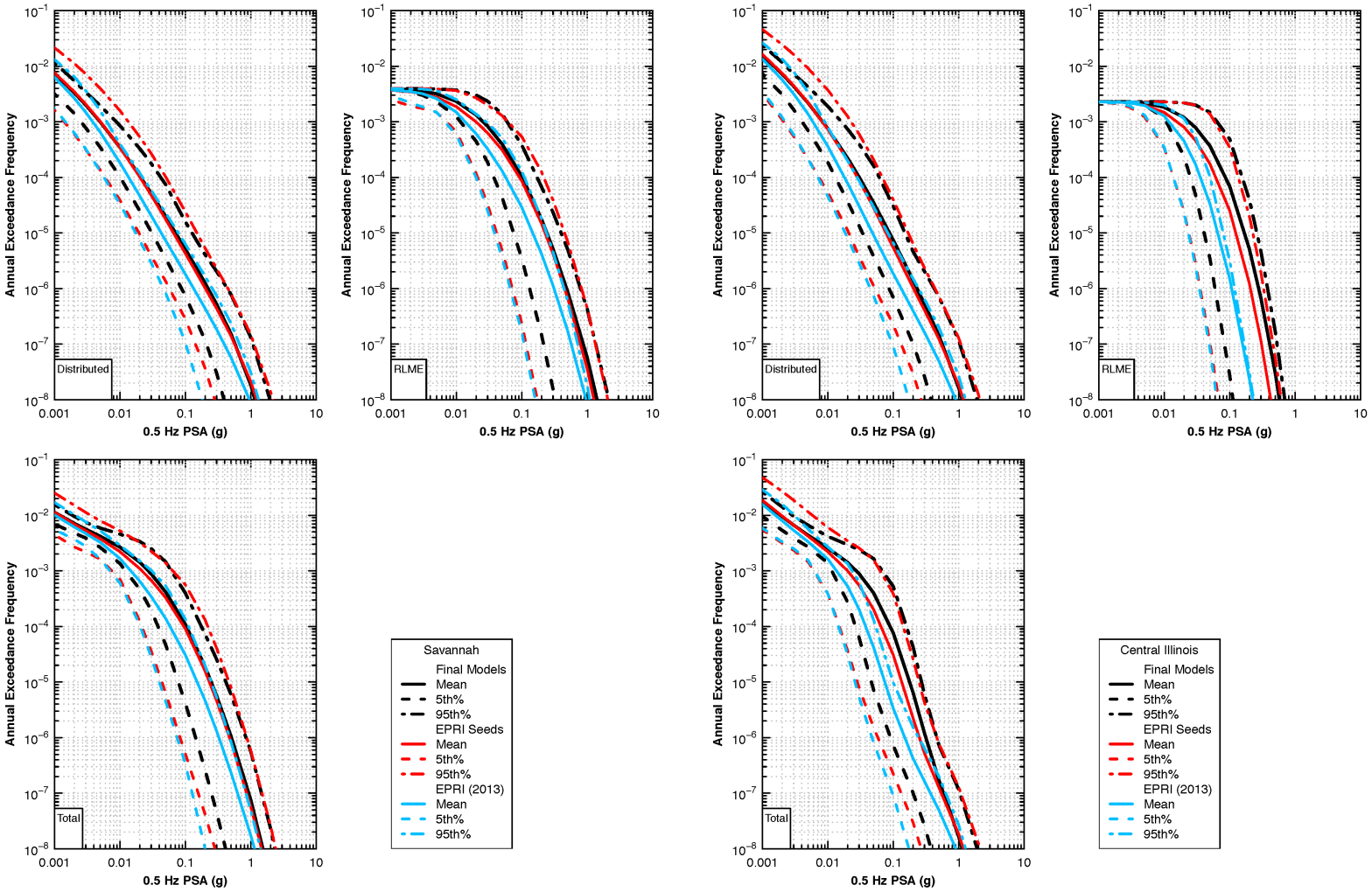


Figure 12–9 Comparison of hazard curves from variants of EPRI (2013) median GMMs to the NGA-East base case GMMs, for 1 Hz PSA at the Savannah site (left) and the Central Illinois site (right).

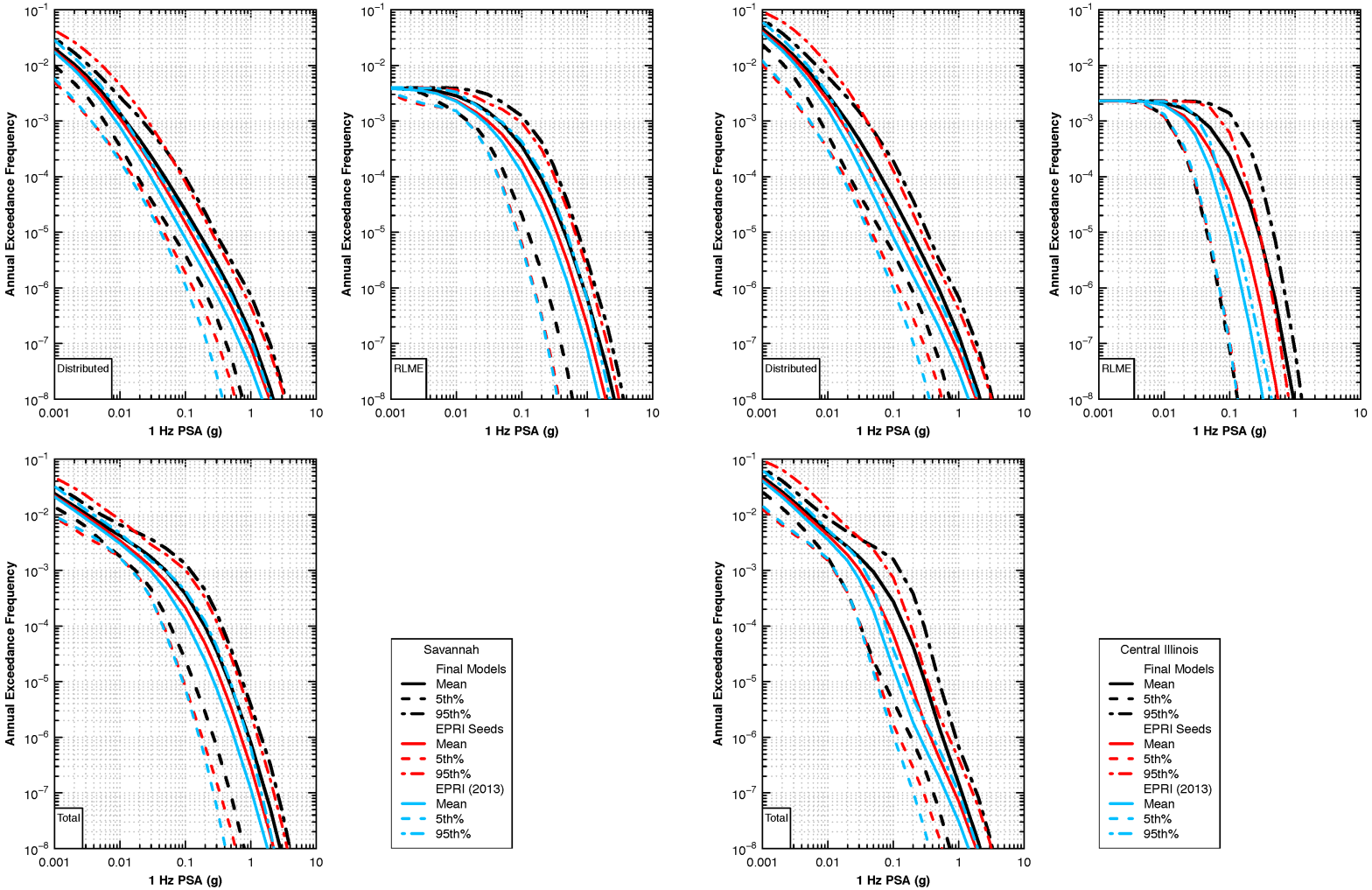


Figure 12–10 Comparison of hazard curves from variants of EPRI (2013) median GMMs to the NGA-East base case GMMs, for 2.5 Hz PSA at the Savannah site (left) and the Central Illinois site (right).

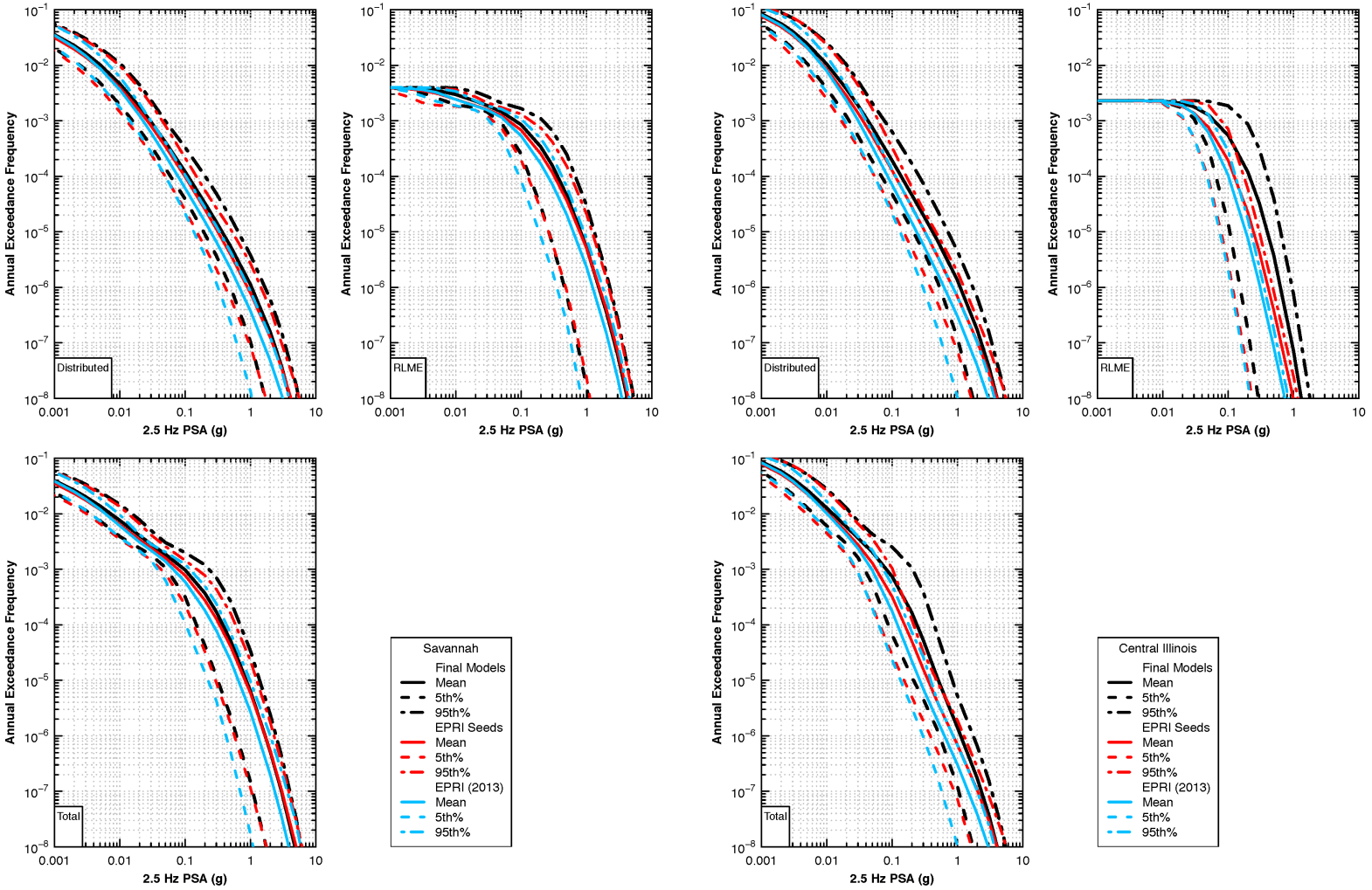


Figure 12–11 Comparison of hazard curves from variants of EPRI (2013) median GMMs to the NGA-East base case GMMs, for 5 Hz PSA at the Savannah site (left) and the Central Illinois site (right).

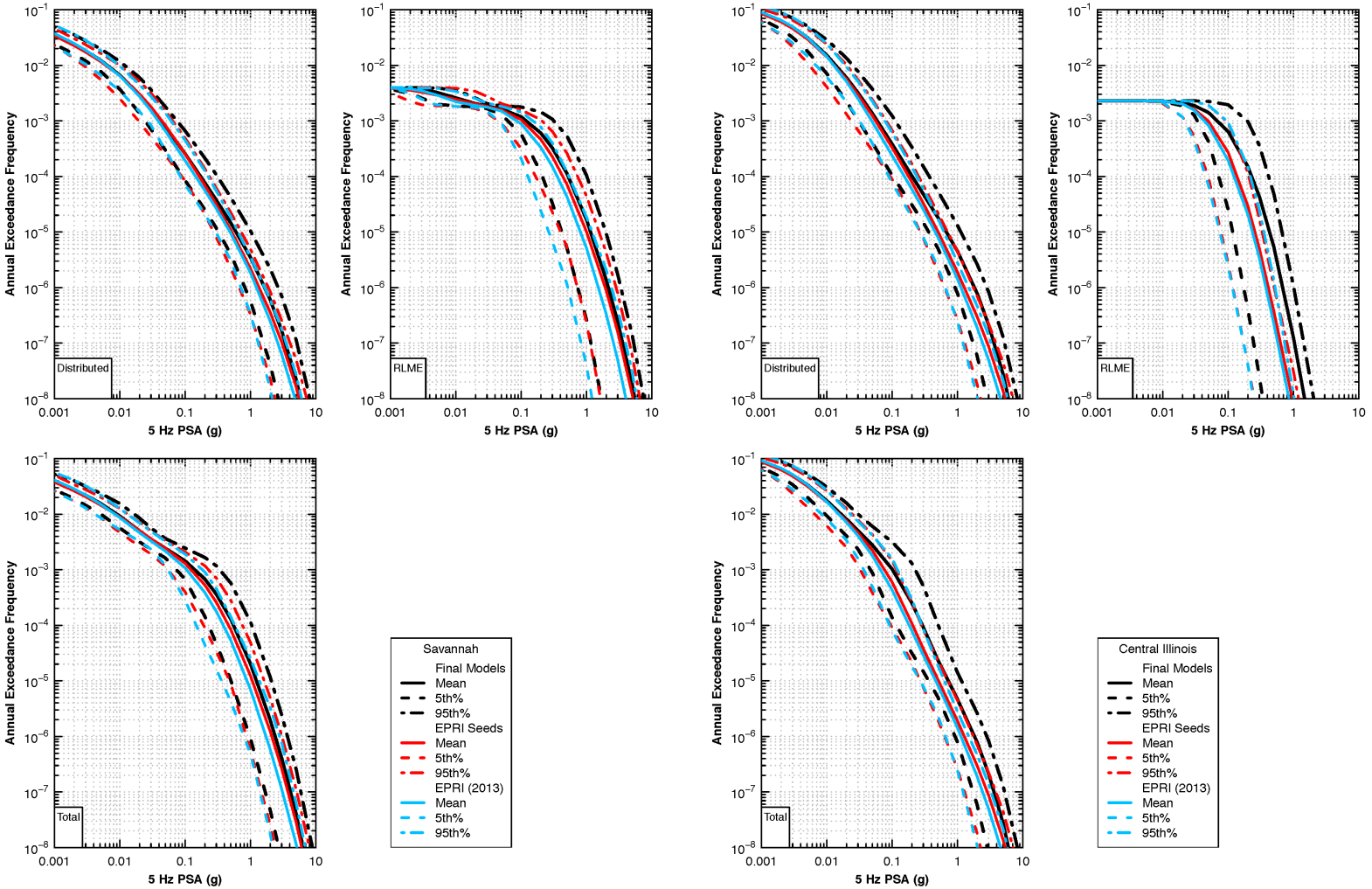


Figure 12–12 Comparison of hazard curves from variants of EPRI (2013) median GMMs to the NGA-East base case GMMs, for 10 Hz PSA at the Savannah site (left) and the Central Illinois site (right).

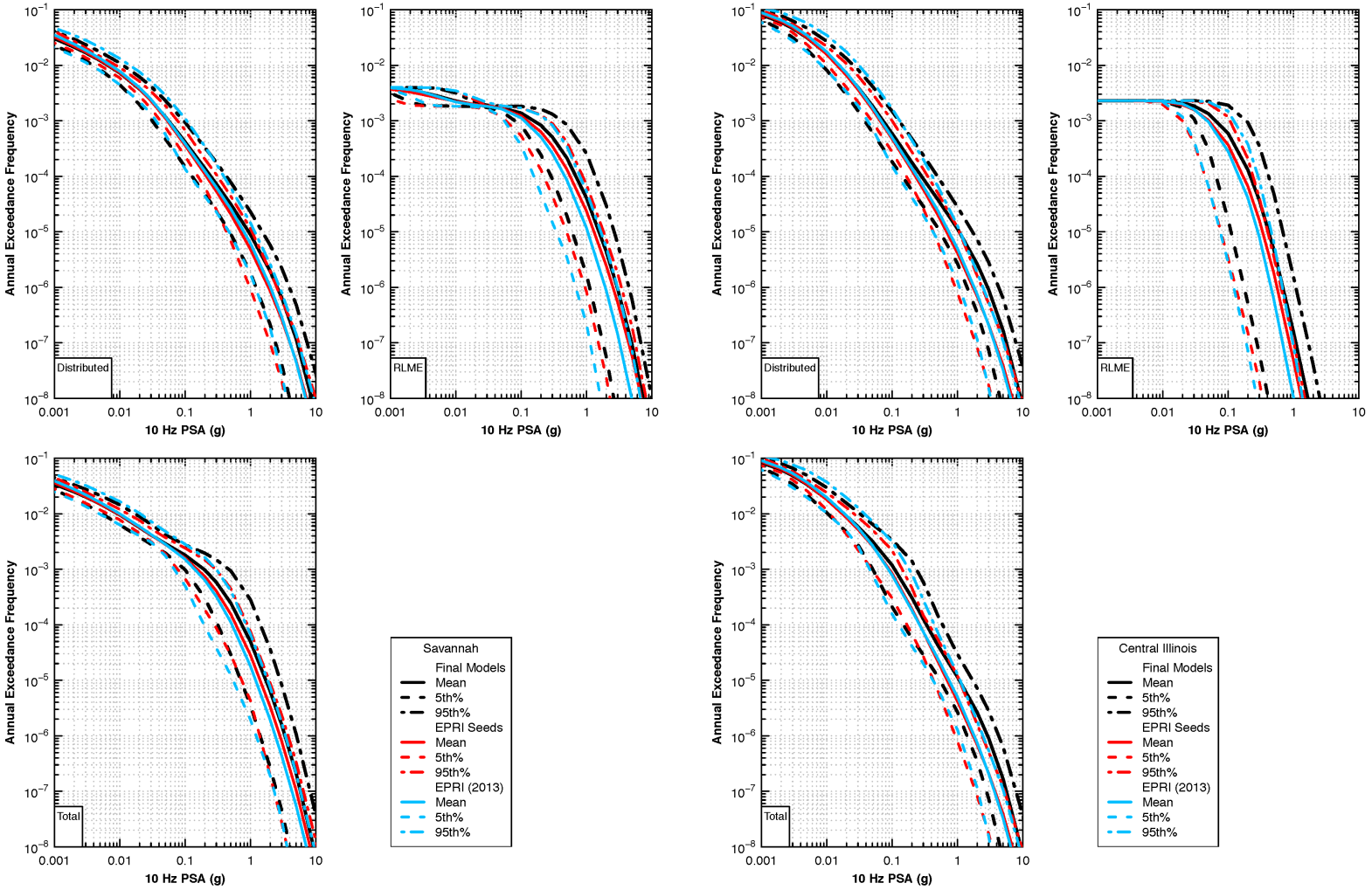


Figure 12–13 Comparison of hazard curves from variants of EPRI (2013) median GMMs to the NGA-East base case GMMs, for 25 Hz PSA at the Savannah site (left) and the Central Illinois site (right).

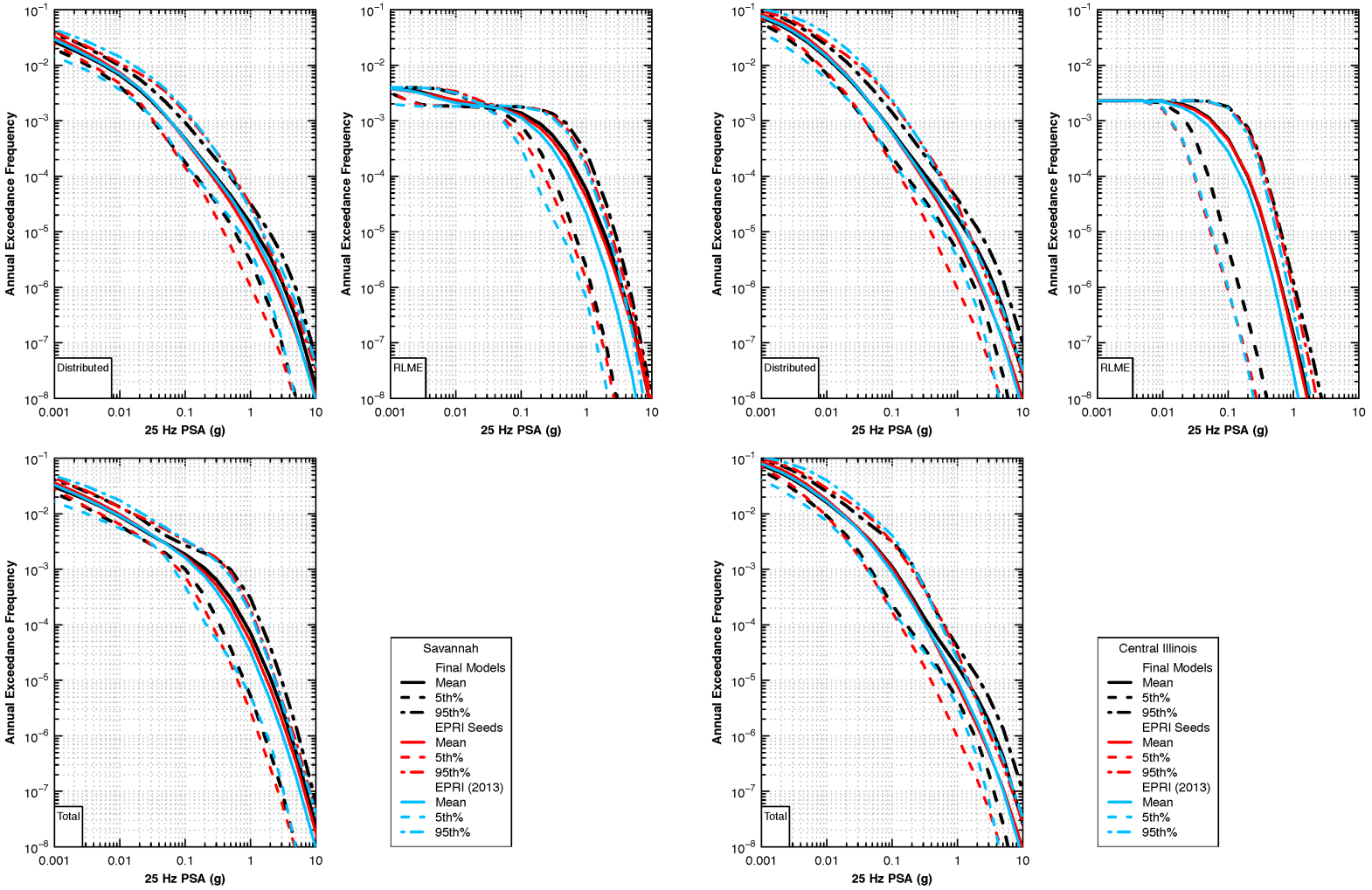


Figure 12–14 Comparison of hazard curves from variants of EPRI (2013) median GMMs to the NGA-East base case GMMs, for PGA at the Savannah site (left) and the Central Illinois site (right).

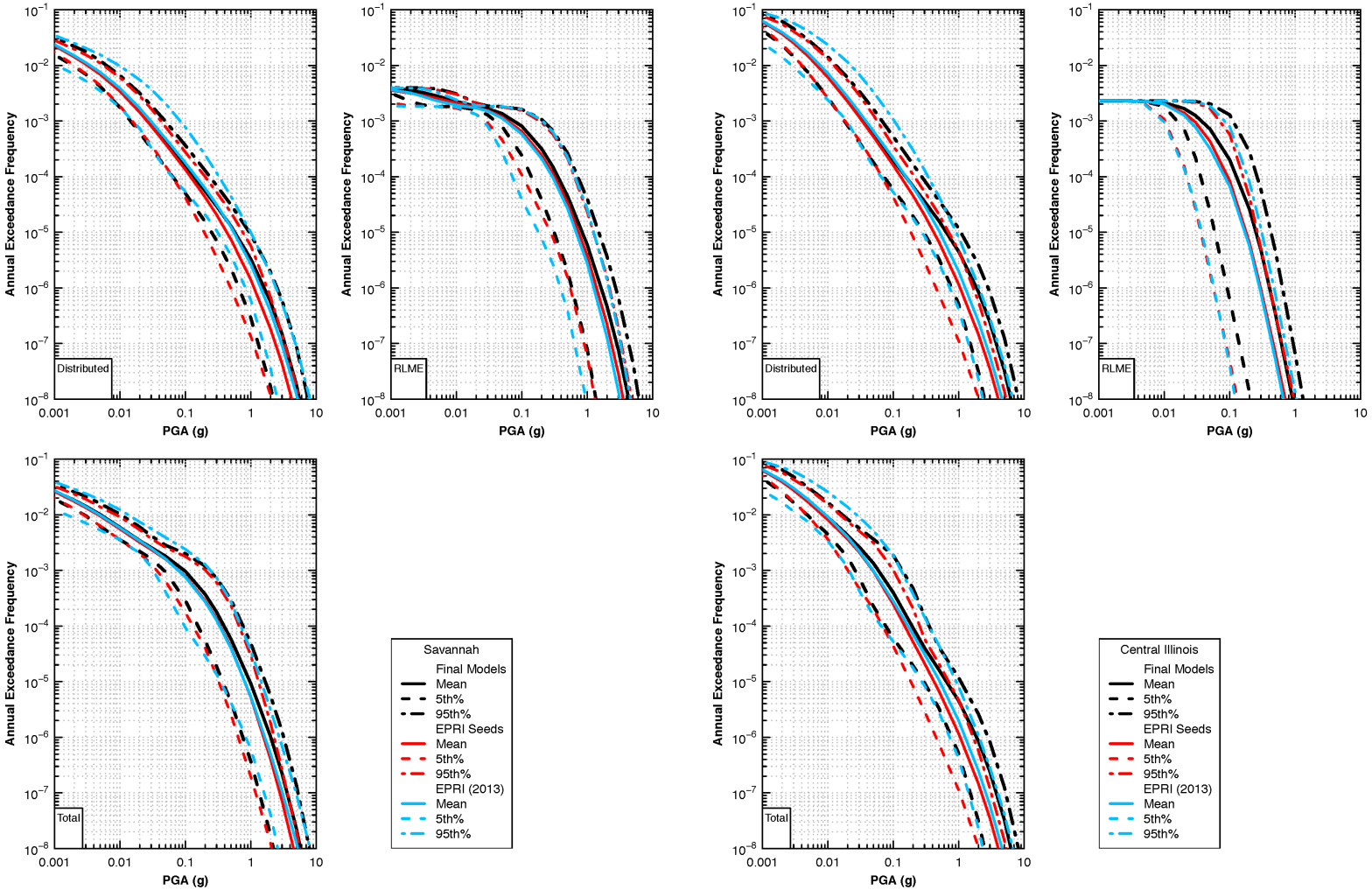


Figure 12–15(a) Hazard ratio plots of NGA-East relative to EPRI (2013) median GMMs for Central Illinois, frequencies of 0.5–5 Hz.

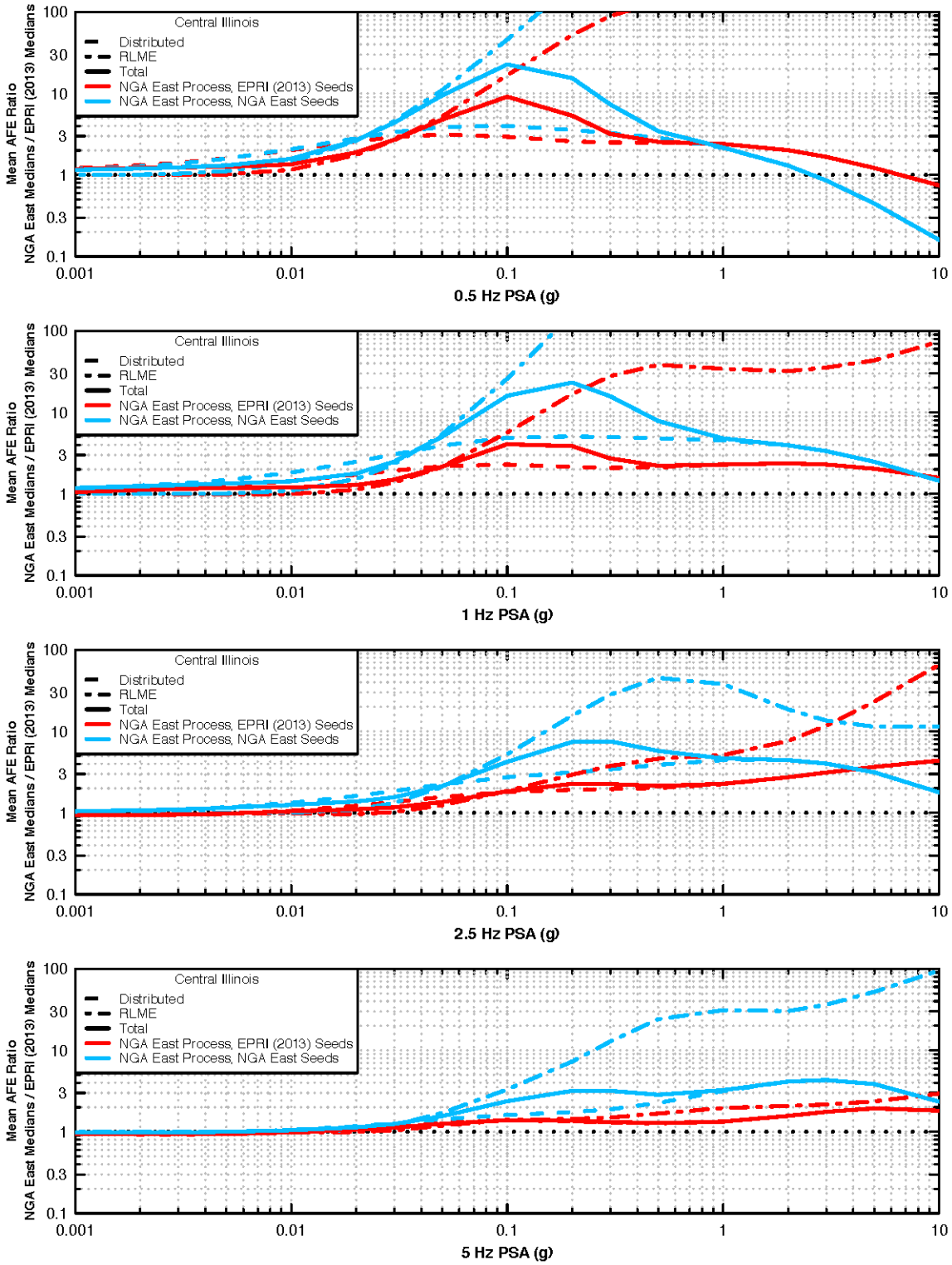


Figure 12–15(b) Hazard ratio plots of NGA-East relative to EPRI (2013) median GMMs for Central Illinois, frequencies of 10–25 Hz and PGA.

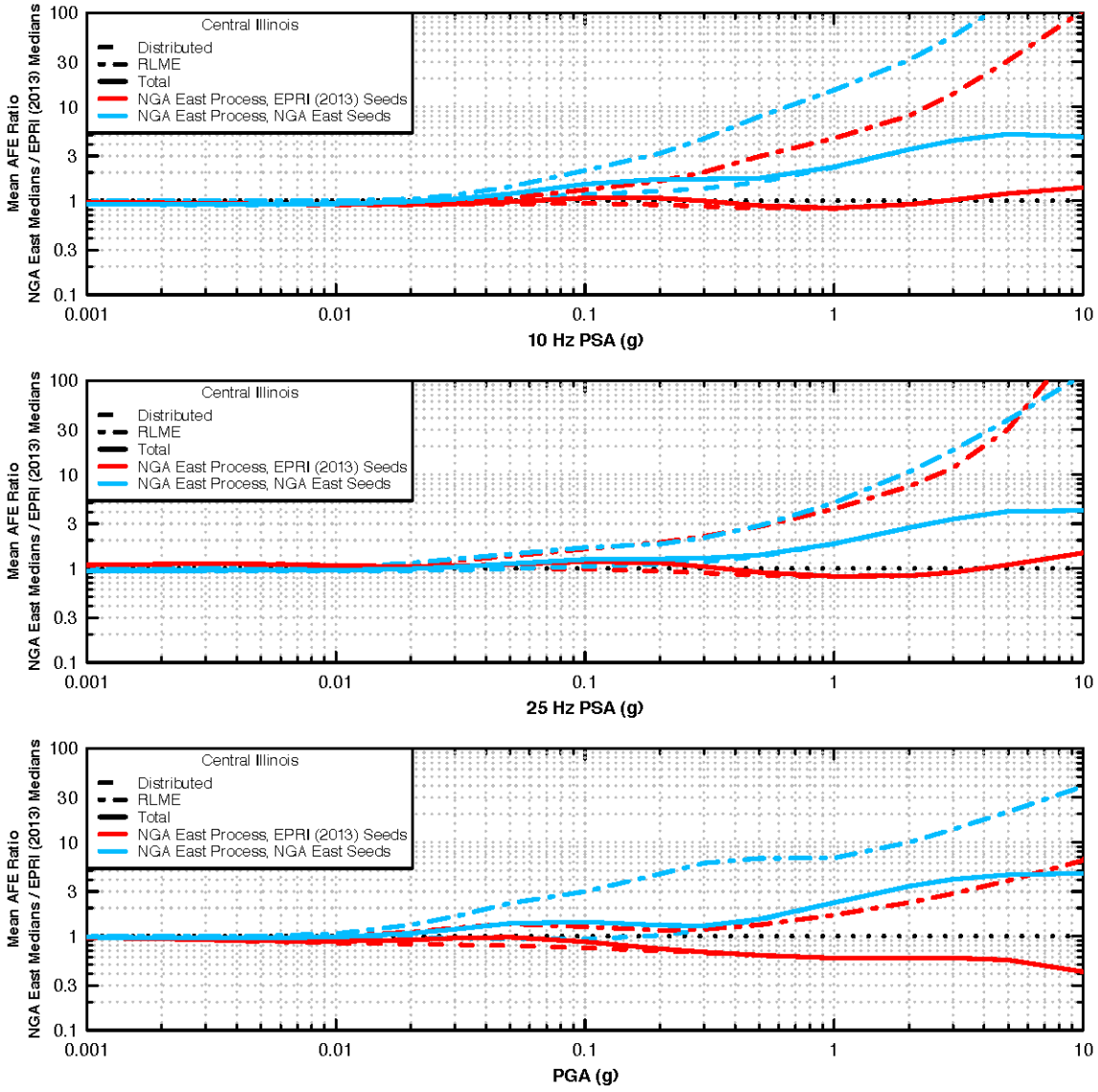


Figure 12–16(a) Hazard ratio plots of NGA-East relative to EPRI (2013) median GMMs for Savannah, frequencies of 0.5–5 Hz.

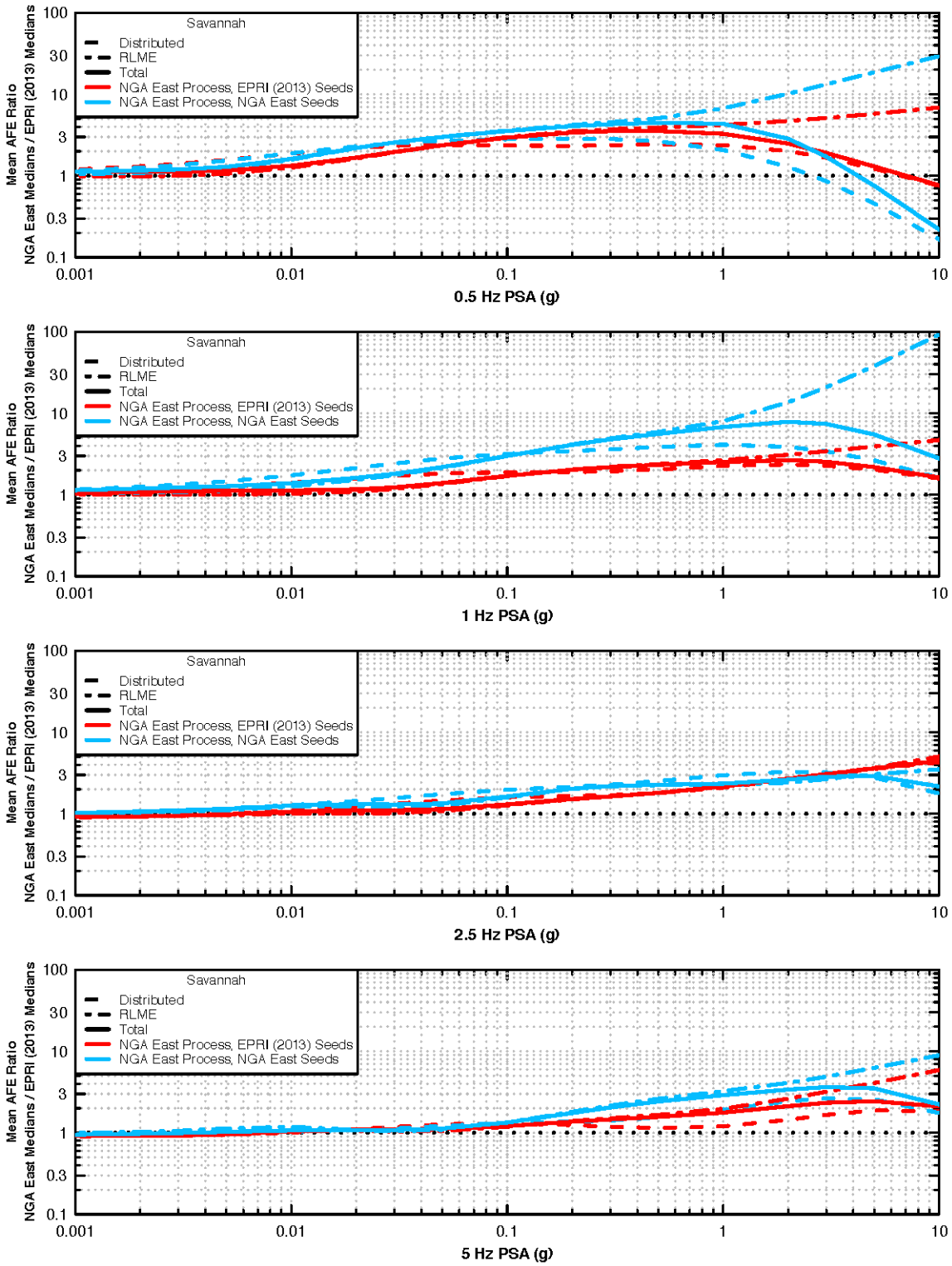
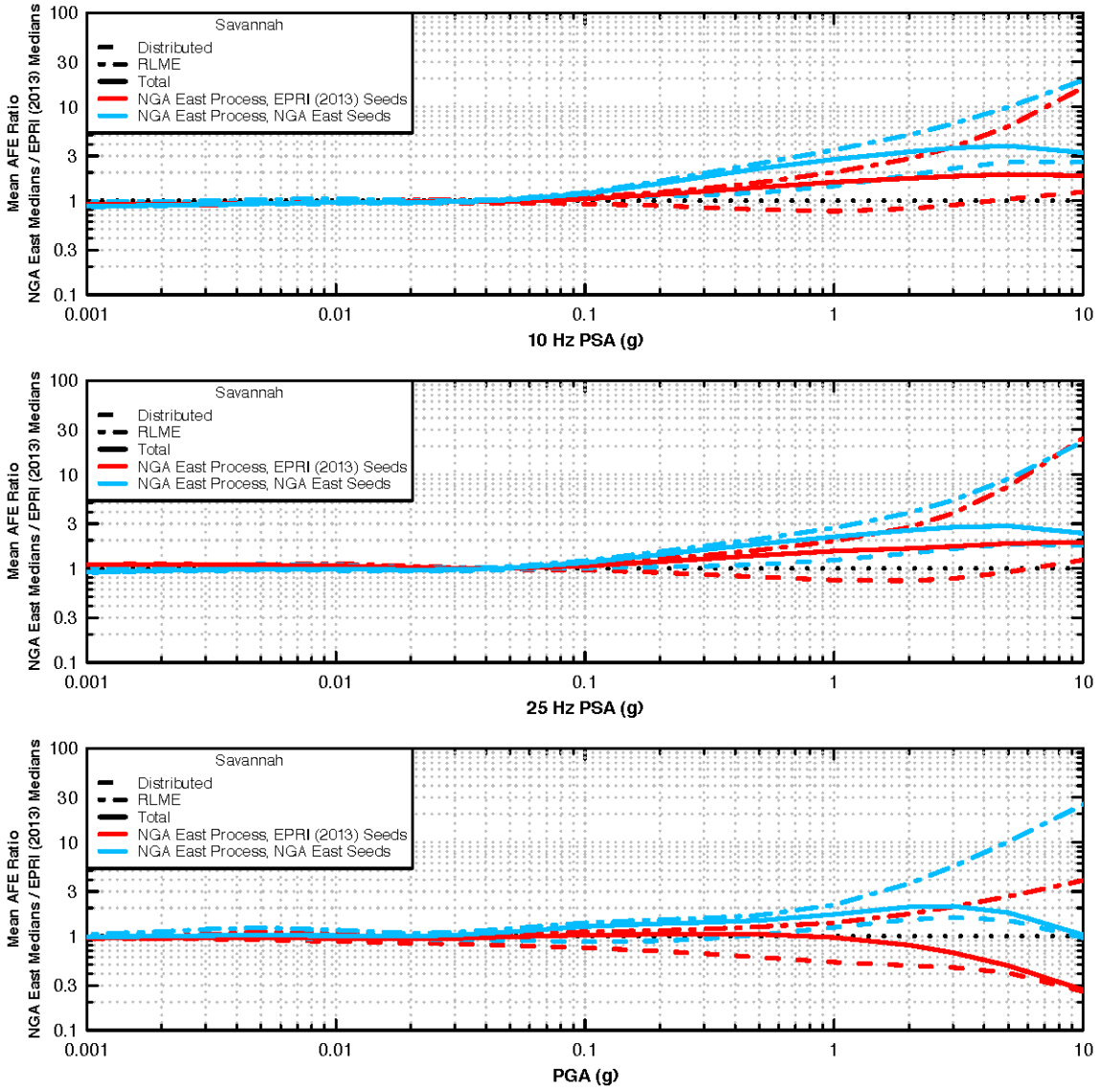


Figure 12–16(b) Hazard ratio plots of NGA-East relative to EPRI (2013) median GMMs for Savannah, frequencies of 10–25 Hz and PGA.



13. Adjustment for Gulf Coast Region, Source Depth, and Hanging Wall Effects

13.1 Purpose

This chapter summarizes the adjustments applied to the median Ground-Motion Models (GMMs) developed as part of the Next Generation Attenuation for Central and Eastern North America (CENA) project (NGA-East). Three types of adjustments are described below: (1) adjustment for the Gulf Coast Region (GCR) attenuation conditions, (2) adjustments for source-depth effects; and (3) adjustment for hanging-wall effects. The median GMMs used as seeds to develop the models presented in Chapter 8 did not formally address these specific issues, and the NGA-East Project undertook the development of three types of adjustments to address them in PEER [2015b]. The evaluation and integration of these effects into the NGA East ground-motion characterization (GMC) model is described in this chapter. Sections 13.2 and 13.3 borrow material from PEER (2015b) and Section 13.4 borrows from GeoPentech (2015).

13.2 Gulf Coast Adjustment Models

13.2.1 Introduction and Motivation

All the GMMs described in Chapter 7 were developed for the Mid-Continent Region (MCR), see Figure 4–4. There are very limited data in the GCR, making the development of stand-alone GMMs impractical. Therefore, the NGA-East median models do not account for differences in crustal structures and Q estimates for the GCR, which tend to lead to a stronger rate of attenuation. The NGA-East Project developed two new models to provide adjustments to the existing GMMs for applicability to the GCR. With these two models, which can be applied to any of the existing candidate and final GMMs, it is possible to cover the complete CENA region. The model development is described in detail in PEER (2015b); a summary of the candidate models and their evaluation are provided in the following sub-sections.

13.2.2 Candidate Models for Regional Adjustments

Two groups built on their candidate median GMM work to develop Gulf Coast adjustment models. The first group was composed of R. B. Darragh, N. A. Abrahamson, W. J. Silva, and N. Gregor (DASG), who developed the median DASG GMMs documented in Chapter 3 of PEER (2015a). The second group consisted of J. Hollenback, N. Kuehn, C. A. Goulet, and N. A. Abrahamson, who developed the PEER GMMs documented in Chapter 11 of PEER (2015a).

The model development was based on the assumption that the Gulf Coast boundary was as described in Dreiling et al. (2014) and summarized in Section 4.2.4. The data available in the NGA-East database (Goulet et al. 2014]) with both the earthquake source and the site located in the Gulf Coast are relatively sparse. Figure 13–1 shows the location of the earthquakes and travel paths used for the assessment, and Figure 13–2 summarizes the magnitude and distance ranges covered in the NGA-East database for the GCR. Both DASG and PEER used this dataset in their model development [see PEER 2015b for details on data selection from each group]. The regionalization assumption from Dreiling et al. (2014) implies that at least some of

the data includes paths sampling the Mississippi Embayment Region (Figure 13–1) but not path sampling the Florida area. The inclusion or exclusion of Florida within the GCR is addressed as an epistemic uncertainty of the regionalization from two models: Dreiling et al. (2014) and Al Noman and Cramer (2016).

The two groups adopted slightly different approaches, which were consistent with their original candidate median GMM development. The PEER group used their empirically-based Fourier amplitude spectrum (FAS) functional form, which they recalibrated with the GCR data through a residual analysis procedure. DASG use their theoretical-based approach based on the point-source stochastic model, constraining the point-source parameters by data inversions from the GCR. DASG then ran ground-motion simulations for a wide range of magnitude and distance (**M** from 4.5 to 8.5, an RJB from 1 to 1000 km) for both regions and computed PSA ratios. Both groups considered alternative formulations to model ground motions from the GCR from which they selected a preferred model. Once their respective GCR models were developed, both groups computed ratio of response spectra for various **M** and R_{RUP} conditions and defined models as a distance-dependent ratio of PSA of the form:

$$f(\theta,R) = \frac{PSA_{GRC}}{PSA_{MCR}} \qquad (13\text{–}1)$$

where PSA_{GCR} and PSA_{MCR} are the PSA values from the GCR and MCR, respectively, and θ is the set of predictive parameters. The groups did not observe a magnitude dependence in the scaling, mostly due to the lack of data available to constrain such a model. Both groups used the events shown in Figure 13–1 for which the magnitude range is narrowly limited to 3.35–4.71. This range was not sufficient to capture magnitude dependence with the PEER approach. For DASG, the narrow magnitude range spanned by the GCR data led to inversion differences only in Q_0. Therefore, although magnitude was considered in the simulations, only a distance dependence was observed and modeled.

This factor scales the median PSA values from the NGA-East GMM for MCR to obtain GCR values. The models are shown in Figure 13–3 for a subset of frequencies (Note: the PEER model is frequency independent.). Both models show ratios essentially equal to unity for roughly the first 100 km: the DASG ratio decays very slowly from distance zero, and the PEER model is exactly 1 until 100 km. The unit ratio close in and the monotonic decrease in ratios with distance implies that only differences in attenuation are captured by the models. This may appear counter-intuitive given the deep sediment deposits present in the GCR. However, the ratios were computed for the same hard-rock conditions as the candidate medians GMMs (V_{S30} =3000 m/sec and kappa (κ = 0.006 sec), thus neglecting expected site effects from the geological conditions. It is expected that defining an alternate set of reference site conditions for the Gulf Coast, which would involve additional work on kappa, would be more appropriate in capturing regional differences in ground motions for a wide range of site conditions. Tasks could include extending the work from Chapman and Conn (2016) to shear waves. Estimation of kappa and site effects for sites other than the reference rock site is not part of the NGA-East Project scope.

The PEER model is given by the relationship:

$$\ln\left(\frac{PSA_{GRC}}{PSA_{MCR}}\right) = -0.00221 \times \max\left(0, R_{JB\_GCR} - 100\right) \qquad (13\text{–}2)$$

and the DASG model is given by the relationship:

$$\ln\left(\frac{PSA_{GRC}}{PSA_{MCR}}\right) = \gamma(f) R_{JB\_GCR} \qquad (13\text{–}3)$$

where R_{JB_GCR} is the length of the horizontal travel path within the GCR. The values of $\gamma(f)$ are listed in Table 13–1. The PEER model was developed for R_{RUP} but at distances of 100 km and greater; R_{RUP} and R_{JB} are essentially equal.

13.2.3 Evaluation and Selection of Regional Adjustment Models

The two models show substantial differences (Figure 13–3), especially at low frequencies and large distances. At 0.5 Hz and 1000 km, there is approximately a factor-of-five difference between the two adjustment models. This large difference in the models is partly a result of the difference in approach each group took in developing their adjustment models and partly due to the lack of quality data for the GCR in the NGA-East database. With this lack of data, a large range in median adjustments appears appropriate for application.

Also shown on Figure 13–3 are values of the ratio PSA_{GCR}/PSA_{MCR} based on EPRI (2013) and Cramer and Al Noman (2016). The values for EPRI (2013) are computed using the MCR and GCR ground-motion prediction equations (GMPEs) for the central model of each of the four EPRI (2013) clusters for **M** = 5, 6, 7, and 8 earthquakes. The EPRI [2013] GMPEs for the GCR were developed by applying differences in Q between the MCR and GCR to the GMPEs for the MCR. As such, they were derived in a similar manner to the DASG GCR adjustment model. At low frequencies, the EPRI (2013) ratios are similar to the DASG adjustment factors but become more variable at higher frequencies, generally lying between the PEER and DASG adjustment factors. Cramer and Al Noman (2016) only provide a GMPE for the GCR derived from empirical data contained in the NGA-East database and recorded by the U.S. Transportable Array (TA), and supplemented by recordings from the 2010 **M** 7.1 Darfield, New Zealand ,earthquake obtained at distances less than 60 km.

In order to assess the implied PSA_{GCR}/PSA_{MCR} ratios, the rate of attenuation for distances greater than 100 km computed using the Cramer and Al Noman [2016] GMPE was compared to the rate of attenuation derived from the central NGA-East GMPE developed in Chapter 9. The rate of attenuation for distances greater than 100 km was used because it is at these distances where the GCR effects become evident (Figure 13–3). With the exception of 0.5 Hz, the approximate ratios derived for the Cramer and Al Noman (2016) are generally similar to the PEER adjustment factors for frequencies of 1 to 5 Hz and are similar to the DASG adjustment factors for 10 Hz and PGA (Cramer and Al Noman do not provide a GMPE for 25 Hz).

As discussed in Section 4.2.4, Dreiling et al. (2014) concluded that differences in crustal structure are a more important factor than differences in Q for producing different attenuation rates in the GCR and MCR. Therefore, the PEER model, which is based on empirical ground-

motion data, is more likely to capture these effects than the DASG model, which is based solely on numerical simulations for the effects of differences in Q. The approximate GCR adjustments derived from the Cramer and Al Noman (2016) empirical GMPE are generally more similar to the PEER model than to the DASG model. On this basis, the Technical Integrator (TI) team favors the PEER model over the DASG model. This assessment is tempered by the fact that the PEER model is based on somewhat limited data in terms of bandwidth and number of recordings, although the assessments of Q in the GCR used to develop the DASG model are also affected to some extent by limited data. Because of its direct reliance on empirical observations, the PEER model is assessed to be twice as likely to represent GCR adjustments than the DASG model and is assigned a weight of 0.67 compared to 0.33 for the DASG model.

13.2.4 Application Recommendations

The two GCR models given by Equations (13–2) and (13–3) provide an adjustment factor to apply to the predictions from the MCR models developed in Section 9. The adjustment is based on the length of the travel path within the GCR. As indicated on Figure 13–3, the GCR adjustments only have a significant effect on ground motions when the length of the travel path within the GCR exceeds 100 km. Note that this is equivalent to using distance proportion weighting to combine MCR and GCR estimates of ln(PSA) using the full travel path.

Alternative interpretations of the boundary of the GCR were discussed in Chapter 4. These different interpretations are illustrated on Figure 13–4. Using crustal structure, Dreiling et al. (2014) defined a GCR that combined the region previously defined by EPRI [1993] (see Figure 4–2) with the upper Mississippi Embayment and extending the GCR to encompass the area around New Madrid. Dreiling et al. (2014) also included all of the Florida peninsula as part of the GCR. EPRI (1993) did not address the southern portion of Florida (see Figure 4–1), and it was left out of the GCR as shown in EPRI (2004). EPRI (2013) refined the boundaries of the Dreiling et al. (2014) GCR to conform to the boundaries of CEUS SSC (EPRI/DOE/NRC 2012) seismotectonic source zones on the basis that the two boundaries are similar and both are based on crustal structure differences across the CEUS. Note that Dreiling et al. (2014) considered only a very limited amount of crustal velocity data from sites outside of the EPRI (2013) GCR in developing their crustal profiles. Therefore, the conclusions about differences in attenuation reached by Dreiling et al. (2014) can be considered applicable to the EPRI (2013) GCR. As there are some practical advantages in probabilistic seismic hazard analysis (PSHA) calculation to having the GCR boundary match source zonation boundaries, the EPRI (2013) GCR is adopted by the TI team to represent the Dreiling et al. (2014) GCR regionalization for the purpose of applying GCR ground-motion adjustments.

Cramer and Al Noman (2016) argued for a more limited GCR of greater attenuation rate based on analysis of Earthscope TA ground-motion recordings. Cramer and Al Noman's boundaries are also shown on Figure 13–4. Two boundaries are proposed. The short period boundary is based on the analysis of 5-Hz motions and represents the assessed boundary based on earthquakes located both outside of the GCR and within the GCR. Cramer and Al Noman also found that the same boundary applies to 1-Hz motions for earthquakes occurring within or near the GCR. The Cramer and Al Noman long-period boundary shown on Figure 13–4 is based on evaluation of 1-Hz motions from earthquakes located outside of the GCR. Cramer and Al

Noman (2016) indicate that their boundary locations are resolvable within the spacing of the TA stations, approximately 70 km.

The alternative GCR boundaries have two primary effects on hazard assessment. The first is on assessment of the ground motions in southern Arkansas, Louisiana, and eastern Texas from large earthquakes occurring in the New Madrid seismic zone. These earthquakes have been shown to be the major source of hazard for low-frequency ground motions in this region [e.g., Figure 8.2–3f (EPRI 2013)]. The EPRI [2013]/Dreiling et al. (2014) assessment would specify that the entire travel path from these earthquakes would occur within the GCR, while the Cramer and Al Noman(2016) long-period boundary would indicate that much if not most of the travel path is in the MCR. The Cramer and Al Noman (2016) short-period boundary places the upper Mississippi Embayment in the MCR, and their long period boundary limits the GCR to southeastern Texas, the southern two-thirds of Louisiana, and southwestern Mississippi.. Cramer and Al Noman (2016) suggest that the differences in location between the short-period and long-period boundaries may relate to differences in crustal structure within the GCR, but they do not point to specific tectonic features that would indicate differences in structure.

Figures 7.1-5 and 7.1-6 of EPRI/DOE/NRC (2012) provide maps of magnetic anomalies and residual isostatic gravity, respectively, for the CEUS. The magnetic anomalies map shows some structure at the boundary between the CEUS SSC ECC-GC and GHEX seismotectonic zones, but this boundary is located south of Cramer and Al Noman's long-period boundary by a distance approximately equal to the TA station spacing. The residual isostatic gravity map indicates a transition from areas of higher gravity to areas of lower gravity at the approximate location of Cramer and Al Noman's long-period boundary. Note that the PEER GCR model is based on analysis of recordings for which most of the travel paths cross Cramer and Al Noman's long-period boundary.

The second location where the differences in the proposed GCR boundary would have a significant impact is on the assessment of ground motions in the Florida peninsula from large earthquakes occurring in the vicinity of Charleston, South Carolina. The EPRI (2013)/Dreiling et al. (2014) assessment would specify that much of the travel path would be within the GCR, while the Cramer and Al Noman [2016] would specify that the entire travel path would be within the MCR. Cramer and Al Noman [2016] base their assessment of the Florida peninsula on TA recordings of the 2012 **M** 4.2 Whitesburg, Kentucky, earthquake and to a lesser extent on limited recordings from the 2014 **M** 5.0 North Cuba earthquake. They concluded that the ground-motion attenuation with distance for these earthquakes is similar to the MCR. They do indicate a "step down" in the recordings near the Georgia/Florida border for the short-period (5-Hz) motions from the Whitesburg earthquake. Examination of their Figure 14 also suggests that one could interpret the data to indicate a change in slope at this point. The northern portion of the Florida peninsula and the adjacent portions of southern Georgia consist of the Suwannee terrane, a portion of the African continent that was accreted to the north [e.g., Thomas et al. (1989)]. EPRI/DOE/NRC (2012) included this region as part of the ECC-GC extended crust seismotectonic source, although they indicate that the crust of the Suwannee terrane is relatively unextended compared to the portions of the ECC-GC to the west. Note that Dreiling et al. (2014) did not use velocity profiles from the Florida peninsula in developing their GCR crustal model, as indicated by their Figure 3.13. The analyses performed for the PEER model also did

not include data from Florida (see Figure 13–1), and Chapman and Conn's (2016) assessment of Q in the GCR also did not include ray paths in Florida. An alternative interpretation was reached by Florida Power and Light (2016), who concluded that the characteristics of ground-motion attenuation in the southern Florida peninsula and regions north of Cuba are more similar to the GCR GMPEs developed by EPRI (2004) than to those for the MCR.

The differences between the EPRI (2013)/Dreiling et al. (2014) and Cramer and Al Noman (2016) boundaries in other areas are not considered significant because they are generally on the order of the resolution of the Cramer and Al Noman (2016) boundaries, which is represented by the spacing of the TA stations (approximately 70 km).

The TI team recommends the use of two alternative GCR/MCR boundaries to represent the current state of epistemic uncertainty in applying the GCR adjustment models presented in Section 13.2.2. These two alternatives are shown on Figure 13–5. GCR_{LARGE} corresponds to the region defined by EPRI (2013), which approximates Region 1 of Dreiling et al. (2014) and is consistent with the CEUS SSC source zonation. The alternative is GCR_{SMALL}, which uses the Cramer and Al Noman (2016) boundaries to limit the extent of the GCR in Florida and the southern Gulf Coast. The TI team selected the Cramer and Al Noman (2016) long-period boundary because it addresses wave propagation of those seismic waves that will have the greatest impact on hazard assessment in the Gulf Coast. For example, it better captures the low-frequency motions from large earthquakes occurring on the RLME sources in the New Madrid region. The GCR_{LARGE} zonation is given slightly larger weight (0.6) compared to the GCR_{SMALL} interpretation (weight 0.4) because the GCR adjustment models developed above are based on data from travel paths that cross the GCR_{SMALL} boundary, and there is no clear tectonic interpretation for the differences in attenuation observed by Cramer and Al Noman (2016). The coordinates of the two GCR zones are provided in Table 13–2.

13.3 Adjustment for Source Depth Effects

The issue of source-depth effects was discussed in several NGA-East workshops and working meetings. Various researchers presented data-based evidence of this important contribution to ground motions in CENA. A summary of the importance of depth effects in CENA is presented in Section 3.1 and 3.2 of PEER (2015b). The conclusions from these sections indicate that there is significant depth dependence to the stress parameter (as used in the point-source model for the study) for small events at depths less than about 10 km. This includes almost all potentially-induced events (PIEs) and some tectonic earthquakes contained in the NGA-East database (Goulet et al. 2014). As Figure 13–6 shows, most tectonic events occur at depths greater than 10 km, even those with small magnitudes. For this reason, although the effect of the depth variation on ground motions is substantial, it appears to be primarily of importance for PIEs, and there is no empirical constraint to verify this assumption given the lack of large magnitude data in CENA. In addition, of the ten developers who provided ground motions for the NGA-East project, only two [Yenier and Atkinson (YA15) and PEER] included a focal-depth dependence of the motions as a fundamental part of their methodology. The motions provided by those developers were for nominal focal depths, however, and did not include a focal-depth as a parameter (such as hypocentral depth Z_{HYP} or the depth to the top of rupture Z_{TOR}). Frankel has shown that there is a significant depth dependence for longer period motions, which he

attributes to the excitation of surface waves for shallow sources. But short-period motions in his model were obtained from stochastic model simulations, which use a depth-independent stress parameter and therefore have no depth dependence. Therefore, NGA-East Project considered depth models from other regions to guide the model development for depth effects over the full magnitude range of expected events in CENA. The model was developed with the intent to be used with any of the seed or final median GMMs.

The model is parameterized in term of Z_{TOR}—a depth measure that is a less computationally intensive parameter than Z_{HYP}—to integrate into PSHA computations and has been shown to provide a good representation of the effect of earthquake depth on ground motions in active tectonic regions.

The following sub-sections provide a summary of some of the work detailed in PEER [2015b] as well as new material not previously documented to address depth modeling. Four primary pieces are required to perform depth adjustments:

1. Depth scaling model ($f_{Z_{TOR,M}}$): quantifies the scaling (amplification or de-amplification) due to depth as a function of magnitude and frequency (Section 13.3.1)
2. Depth centering model ($f_{Z_{TOR,Z}}$): provides the mean depths of which the current GMMs models are representative of (or centered around) in their current state (Section 13.3.2).
3. Evaluation of proposed models (Section 13.3.3) and recommendations for their application (Section 13.3.4)
4. Source depth model: provides the depth statistics for sources from the CEUS SSC (Section 13.3.5).

The proposed source-depth adjustment model for the NGA-East median GMMs is given by

$$f_{Z_{TOR}} = f_{Z_{TOR},M} \times \Delta Z_{TOR} \quad (13\text{–}4)$$

where $f_{Z_{TOR},M}$ is the magnitude-dependent source-depth scaling factor, and ΔZ_{TOR} is the depth deviation from the centered Z_{TOR} value. The term $f_{Z_{TOR}}$ is added to the natural log of the ground motion value obtained from the NGA-East median GMM.

13.3.1 Candidate Model for Depth Effects Scaling ($f_{Z_{TOR,M}}$)

The candidate model for adding source-depth scaling to the NGA-East median GMMs was first documented in PEER (2015b). The formulation of the model is based on three existing models developed for NGA-West2 (Chiou and Youngs 2014; Campbell and Bozorgnia 2014; and Abrahamson, Silva and, Kamai 2014), and from the “PEER” seed GMM (Hollenback et al. Chapter 11 in 2015a).

The shape of the three NGA-West2 source-depth scaling models with frequency and magnitude guides the proposed source-depth adjustment model. Additionally, the proposed model is constrained by the source-depth scaling implied by the PEER seed GMM.

The magnitude-dependent source-depth scaling factor, $f_{Z_{TOR,M}}$ is given by

$$f_{Z_{TOR,M}} = \begin{cases} b_1 & \text{for } M \le 5.0 \\ b_1 + b_2 \dfrac{M-5}{1.5} & \text{for } 5.0 < M \le 6.5 \\ b_1 + b_2 & \text{for } M > 6.5 \end{cases} \quad (13\text{–}5)$$

where b_1 and b_2 are frequency-dependent coefficients listed in Table 13–3. The coefficients b_1 and b_2 were set by averaging the three NGA-West2 models over the range of low to intermediate frequencies and by following the trend of the PEER model at high frequencies. The threshold of high and low frequency was magnitude-dependent, reflecting the magnitude-dependence trends from two of the NGA-West2 models. The functional form was selected to offer a simple, smooth transition between the source-depth scaling factor for **M** 5 and **M** 6.5. Figure 13–7 shows the average of the NGA-West2 models for the full frequency range and for **M** 5 and **M** 6.5. Figure 13–8 shows the proposed model (referred to as PEER hybrid) source-depth scale factors plotted against frequency for a range of **M**.

For all magnitudes, the average of the NGA-West2 source-depth scaling factors was taken for the proposed model between 0.33 and 5 Hz. For simplicity, the source-depth scale factors are flat at all frequencies above 5 Hz and below 0.33 Hz to avoid adopting a more complicated shape that is not informed by NGA-East data. Above 5.0 Hz the proposed model was held constant at the 5-Hz value. This constraint was imposed because 5.0 Hz is where the average of the NGA-West2 models reaches the level of scale factor from the PEER model at high frequencies. At 0.25 Hz and below, the proposed model was set at the NGA-West2 average for all the frequencies below 0.25 Hz, allowing the model to go to negative values. The same rule was applied to all magnitudes. Table 13–3 provides the model coefficients for Equation (13–5).

13.3.2 Candidate Models for Depth Effects Centering (ΔZ_{TOR})

The mechanism for centering the source-depth adjustment model is required for its implementation in PSHA. A result of selecting Z_{TOR} as the source-depth metric is that large magnitude events will have shallow Z_{TOR}. In other words, large magnitude events will tend to rupture most if not all of the seismogenic thickness of the crust and thus have Z_{TOR} values that are close to 0 km. If the source-depth adjustment model is centered on Z_{TOR} values not compatible with the magnitude and data used in the median GMM development, then the resulting ground motions could be adjusted in an inappropriate way. Therefore, a magnitude-dependent centering is recommended for application. The form of the recommended centered Z_{TOR} model, ΔZ_{TOR}, is adopted from Chou and Youngs (2014), given by:

$$\Delta Z_{TOR} = Z_{TOR} - E\left(Z_{TOR}\right) \quad (13\text{–}6)$$

where $E(Z_{TOR})$ is the magnitude-dependent centering depth (magnitude-dependent expected Z_{TOR}), and Z_{TOR} is the depth-to-top-of-rupture of the earthquake source, as derived from a SSC model.

13.3.2.1 CEUS SSC-Based Magnitude Dependence of $E(Z_{TOR})$

A relationship to predict the mean Z_{TOR} as a function of magnitude was developed [Youngs, personal communication, 2015] based on: rupture geometries from NUREG-2115 Chapter 5,

earthquake source depths from the NGA-East database, a distribution of hypocentral depth ratios from Chiou and Youngs (2008), and the rupture-area relationship from Somerville (2014). This relationship was originally developed to convert median GMMs that used R_{JB} to R_{RUP}. The development of this model was as follows

- An average hypocentral depth of 10 km is assumed for CENA. This assumption is based on the average of hypocentral depths in the NGA-East flatfile, 12 km, and the mode of the hypocentral depth distribution published in Chapter 5 of NUREG-2115, 8 km.
- A dip angle for events in CENA is assumed to be 75° for strike–slip and 45° for reverse earthquakes. This is based on Table 5.4-1 of NUREG-2115, which states that the default characteristics of CENA earthquakes are a mixture of strike-slip and reverse at a ratio of 2 to 1 with median dip angles of 75° and 45°, respectively.
- The location of the hypocenter and down dip on the rupture plane is defined based on the distribution of the fractional depth from the top of rupture of the hypocenter given in Appendix B of Chiou and Youngs (2008). The mean of this distribution is used and is 0.6375 for strike–slip and 0.628 for reverse earthquakes.
- The thickness of the seismogenic crust in CENA is assumed to be 17 km. This is the central branch of the of the seismogenic thickness logic tree in NUREG-2115.
- Rupture area as a function of magnitude is assumed to follow the Somerville (2014) relationship. The rupture aspect ratio is assumed to be 1:1 (from Table 5.4-1 of NUREG-2115).
- Using a hypocenter depth of 10 km, the specified dips, average hypocenter depth ratios, rupture area relationship, and aspect ratio, the nominal depth to the top and bottom of ruptures is calculated a range of magnitude from **M** = 4.0–8.0.

The results of this model are summarized in Table 13‒4 and plotted in Figure 13‒9 along with $E(Z_{TORi})$, the mean Z_{TOR} used for centering in the Chiou and Youngs (2014) source-depth scaling model for active tectonic regions (NGA-West2). Based on the assumptions listed above, Z_{TOR} reaches 0 before the bottom of the rupture reaches the full seismogenic thickness of the crust. Therefore, for large-magnitude events the mean value of Z_{TOR} reaches 0. The two relationships plotted in Figure 13‒9 have significantly different shapes. For strike–slip events both relationships predict Z_{TOR} = 0 at similar magnitudes (~6.7) but have differing predictions below this magnitude. For reverse earthquakes, the relationships have similar predictions in the **M** = 6.0–6.7 range but differ outside that range. The mechanism-independent average is referred to as the CEUS SSC interim average model in the following section.

13.3.2.2 $E(Z_{TOR})$ Model Based on Implied Depths from Seed GMMs

The NGA-East TI team was concerned that the depth-centering model described above may not be consistent with the GMM development and set out to develop a model based on the depths implied from the seed GMMs. To this effect, the TI team worked closely with the median seed

model developers to better understand the depths at which their models are most closely centered. First, the TI team compiled the depths assumed in modeling and those associated to the events used in the model development for all the candidate models. The depths were originally compiled during the NGA-East database development (Goulet et al. 2014) as part of the source database and included in the NGA-East flatfile. For this assessment, documentation provided by each group of modelers in PEER [2015a] was used. Second, the TI team worked with the candidate seed GMM developers to obtain guidance on the depth distribution most relevant to their model, for the full magnitude range covered by the median GMMs. In most cases, the modelers recommended using the shape of the CEUS SSC interim average model described above and adjusting it to where most of the data used was centered. Table 13–5 and Figure 13–10 summarize that information. An average model was computed assuming the same weight for each modeling team. The direct average was modified to force surface ruptures for events of **M** 7.5. The justification for this modification at large magnitude is provided in Section 13.3.4. Table 13–6 compiles the depth centering model developed through this process. Figure 13–11 shows the two candidate models available for evaluation.

13.3.3 Evaluation and Selection of Depth Effects Scaling ($f_{Z_{TOR,M}}$) and Centering (ΔZ_{TOR})

The TI team evaluated the depth-scaling model developed for CENA (PEER hybrid depth scaling model) and felt it was the most appropriate to use with the NGA-East median GMMs. This model presents the advantage of borrowing from data-rich regions for the scaling shape and is constrained by the NGA-East dataset in the range for which data are available. Scaling for **M** < 5 should follow the **M** = 5 model and scaling for **M** > 6.5 should follow the **M** = 6.5 model.

The GMM-based centering model was selected by the TI team and is recommended for use with the depth-scaling model. This model is most consistent with the depth distributions implied by the seed median GMMs as a whole. Because the final NGA-East median GMMs are based on the statistics of the original models and not on weights assigned to any seed model in particular, using a global average appears to be the most appropriate approach to recommend. A smoothed version of this model is shown on Figure 13–11 and is listed in Table 13–6. Values for intermediate magnitudes should be obtained by linear interpolation.

13.3.4 Application Recommendations

The selected depth-scaling model provided by Equation 13–4 adjusts the predictions from the median GMMs developed in Chapter 9 as a function of magnitude, frequency, and the difference between the Z_{TOR} of the specific earthquake and its expected value $E(Z_{TOR})$, ΔZ_{TOR}. The model for $E(Z_{TOR})$, provided in Table 13–6, represents a median estimate of the centering values for the seed models. Because of the variability in these centering values (Table 13–5), the median models developed in Chapter 9 are more appropriately considered as centered on a range of Z_{TOR} values rather than a single expected value. In order to capture this effect, Equation 13–6 was modified to allow Z_{TOR} values within 2 km of $E(Z_{TOR})$ to be considered as having ΔZ_{TOR} equal to zero. In addition, the results presented in PEER (2015b) indicate that

ground motions appear to not increase in CENA as the depths become greater than 10 km. After combining these two effects, the recommended relationship for ΔZ_{TOR} is thus given by

$$\Delta Z_{TOR} = \begin{cases} Z_{TOR} - E[Z_{TOR}] + 2 & \text{for } Z_{TOR} < E[Z_{TOR}] - 2 \\ 0 & \text{for } E[Z_{TOR}] - 2 \leq Z_{TOR} \leq E[Z_{TOR}] + 2 \\ \min\{10, Z_{TOR}\} - E[Z_{TOR}] - 2 & \text{for } Z_{TOR} > E[Z_{TOR}] + 2 \end{cases} \quad (13\text{–}7)$$

Application of the depth-effects model described in PSHA requires a representation of the distribution of Z_{TOR} for the seismic sources in CENA. Section 13.6 presents a generalized model that can be used to apply the depth adjustments given by Equations (13–5) and (13–7). If more detailed local information is available, it can be used in lieu of these generalized distributions.

13.3.5 Generalized Z_{TOR} Distribution for CEUS SSC Sources

The EPRI/DOE/NRC (2012) report presents information on the focal depth distribution of earthquakes within the boundaries of the CEUS SSC model (Section 5.4.4) but does not provide explicit distributions for Z_{TOR}. Accordingly, the NGA-East project developed the following generalized model for Z_{TOR} for application with the CEUS SSC model.

13.3.5.1 Approach

Because the number of earthquakes in the CEUS for which finite-fault models have been developed is very small, Z_{TOR} distributions were developed from earthquake focal-depth distributions. The approach builds upon the simulation process developed in Appendix B of Chiou and Youngs (2008). Distributions for Z_{TOR} for a specific earthquake magnitude, style of faulting, rupture dip, and seismogenic crustal thickness were developed through a simulation process as follows.

The rupture area for a given magnitude is assessed using the magnitude-rupture area relationship developed by Somerville (2014) for the CEUS. The relationship for aspect ratio as a function of magnitude and style of faulting developed was then used to compute the rupture width and corresponding depth range for the rupture [equal to rupture width times sin(dip)].

The empirical distribution for earthquake focal depths was first truncated at the seismogenic crustal thickness assigned to a seismic source and then renormalized. Then an earthquake focal depth was sampled from the renormalized distribution. For the sampled focal depth, the corresponding value of Z_{TOR} was determined by sampling from an empirical distribution for location of hypocenter within the depth range of ruptures. Chiou and Youngs (2008) defined the hypocentral depth ratio (HDR) as the depth of the hypocenter below the top of the rupture plane divided by the total depth extent of the rupture. A value of HDR was drawn from the appropriate distribution and used to define the top and bottom depths, Z_{TOR} and Z_{BOR}, respectively, for the simulated rupture above and below the simulated hypocentral depth. If the resulting values of Z_{TOR} and Z_{BOR} fell within the range of allowable depths—zero to the assumed seismogenic crustal thickness—then the simulation was accepted and the value of Z_{TOR} retained. If the values of Z_{TOR} or Z_{BOR} fell outside of the allowable range, then the simulation was rejected. The simulation process was repeated until a large number of acceptable solutions were obtained. The resulting values of Z_{TOR} form the required distribution for the specified magnitude, dip, style of faulting, and seismogenic crustal thickness.

The above process was then repeated for the range of magnitudes, styles of faulting, rupture dips, and seismogenic crustal thicknesses contained within the CEUS SSC model. The result was a series of tables containing Z_{TOR} distributions that can be used to implement the NGA East ground-motion model in association with the CEUS SSC model. The following sections describe the analyses performed to develop these distributions.

13.3.5.2 Focal Depth Distributions for CEUS Earthquakes

Figure 13−12 shows the locations of earthquakes in the CEUS SSC earthquake classified as having good quality focal depths. This catalog of earthquakes was further reviewed to remove earthquakes identified as reservoir induced by EPRI (2015). Note that the CEUS SSC earthquake catalog contains earthquakes up to a date of December 31, 2008, and thus does not include data from the recent high level of induced seismic activity in Oklahoma.

Initial evaluations of differences in depth distributions across the CEUS SSC study region indicated that the earthquakes could be grouped into five regions corresponding to various combinations of the CEUS SSC seismotectonic source zones as shown on Figure 13–13:

- A DEEP region consisting of the SLR and GMH zones
- An Extended region consisting of the AHEX, NAP, ECC-AM, ECC-GC, GHEX, and RR/RCG zones
- The PEZ-W zone
- An Interior region consisting of the MIDC-D and OKA zones
- The IBEB zone.

Potential differences in the focal depth distributions for these five regions was then evaluated using Student-t tests for differences in the mean depth in combination with Kolmogorov-Smirnov tests for differences in the cumulative distribution functions [e.g., Benjamin and Cornell (1972)]. The results of these tests indicated that the focal depth distribution for DEEP region consistently could be considered different from that of all of the other regions for a range of magnitudes, and that for earthquakes larger than E(M)3.5, the differences in the focal depth distributions between the other four regions is not statistically significant. Consequently, two generalized focal depth distributions were developed using earthquakes of E(M) 3.5 and larger, one for the DEEP region and one for the remaining portion of the CEUS study region. The cumulative distributions for focal depth were smoothed to remove irregularities resulting from the limited sample size. Figure 13−14 shows the two cumulative distribution functions and the smoothed distributions are listed in Table 13−7.

13.3.5.3 Earthquake Rupture Parameters

The following earthquake rupture parameters were used to develop example Z_{TOR} distributions. Rupture areas (RA) in km^2 were assessed using the relationship developed by Somerville (2014):

$$\log(RA) = -4.25 + M, \sigma\left[\log(RA)\right] = 0.2 \qquad (13\text{–}8)$$

The rupture aspect ratio (AR) models of Chiou and Youngs (2008) were updated using the finite-fault model database developed as part of the NGA-West2 project (Bozorgnia et al. 2014). Figure 13−15 shows the aspect ratio data and the fitted models. The models were fit to earthquakes of magnitude **M** ≤ 7 to minimize the effects of limits on rupture width. The resulting models are given by Equation (13−9).

$$\ln(RA) = \begin{cases} \max(0, -3.814 + 0.666M, & \sigma_{\ln(AR)} = 0.285 \quad \text{for normal faulting} \\ \max(0, -4.254 + 0.785M & \sigma_{\ln(AR)} = 0.395 \quad \text{for strike} - \text{slip faulting} \\ \max(0, -1.931 + 0.349M & \sigma_{\ln(AR)} = 0.414 \quad \text{for reverse faulting} \end{cases} \tag{13–9}$$

The remaining parameter is HDR (see Section 13.3.6.1), defined by the relationship:

$$HDR = \frac{Z_{Hypocenter} - Z_{TOR}}{Z_{BOR} - Z_{TOR}} \tag{13–10}$$

Data from the finite-fault database of NGA-West2 was used to develop relationships for normal, strike–slip, and reverse faulting earthquakes. Figure 13−16 shows the binned data and the smoothed model developed for the three rupture types. Table 13−8 lists the resulting CDFs for HDR used to develop Z_{TOR} distributions.

13.3.5.4 Example Z_{TOR} Distributions.

The focal depth distributions (Section 13.3.5.2) and the rupture parameter relationships (Section 13.3.5.3) were used to develop example Z_{TOR} distributions using the simulation process described in Section 13.3.5.1. The example simulations were performed without considering variability in the rupture parameters. Figure 13−17 shows Z_{TOR} distributions for vertical strike–slip faulting, and Figure 13−18 shows Z_{TOR} distributions for 45° dipping reverse faulting. The distributions for the 30-km thick crust were developed using the focal-depth distribution for the DEEP region, and the distributions for the 15-km thick crust were developed using the focal-depth distribution developed for the general CEUS study region. In each of the plots, the largest magnitude listed in the legend always ruptures to the surface (Z_{TOR} = 0 with frequency 1.0).

13.3.6 Consideration of Other Depth Effects

Additional depth considerations evaluated by the TI team are presented in Appendix G. More specifically, the TI team evaluated the importance of ground-motion effects due to Rg waves (Appendix G.1) and concluded that there was not enough evidence of systematic effects to motivate a model development for the large CENA region. In addition, the TI team evaluated the development of a depth effects model based on the simulations from Frankel. However, the simulations did not cover a wide enough range of conditions for such a development. Additional considerations and discussions are provided in Appendix G.

13.3.7 Example Application of the Depth Effects Model

Tables 13−9a and 13−9b present example calculations of the depth effect for an **M** 6 earthquake using the Z_{TOR} distribution for strike–slip and reverse earthquakes, respectively, in a 15-km thick crust shown at the top of Figures 13−17 and 13−18. The first column contains values of Z_{TOR} and the second column values of the cumulative probability (aleatory) of the Z_{TOR}

value. The third column contains incremental values of Z_{TOR} set at the midpoint of 0.5 km increments and the fourth column contains the corresponding probability masses obtained by differencing the cumulative probabilities in column 2. The fifth column contains the values of ΔZ_{TOR} computed using Equation (13−7). The value of $f_{Z_{TOR},M}$ for a **M** 6 earthquake and PGA is 0.04375 using Equation (13−5) and Table 13−3. The last column contains the values of $f_{Z_{TOR}}$ for each depth increment computed using Equation (13−4). The weighted average values are -0.061and -0.045 for strike–slip and reverse earthquakes, respectively. These correspond to scale factors [$\exp(\overline{f_{Z_{TOR}}})$] of 0.941 and 0.955, respectively on the median ground motions produced by **M** 6 earthquakes.

13.4 Hanging-Wall Effects Models

13.4.1 Introduction and Motivation

Hanging-wall (HW) effects for earthquake ground motions were first proposed by Somerville and Abrahamson (1995) based on observations from a limited number of California earthquakes. Since that time, HW effects have become increasingly incorporated into empirically-based GMPEs. It was the case for example for the recently published NGA-West2 relationships. These relationships model HW effects as an additive term to the log of ground motion that predicts an increase in ground motions at sites located above dipping earthquake ruptures compared to what would be expected at sites located at the same rupture distance but on the footwall side of the rupture. Based on a combination of empirical observation and numerical modeling, HW effects are expected to increase with decreasing rupture dip. The modeled effect decreases as the distance from the rupture increases. In the past, it has been suggested that the use of the Joyner–Boore distance metric implicitly accounts for HW effects, although discussions in GeoPentech (2015) indicated that is it likely true only for moderately dipping ruptures (45 to 60 degrees) and that it does not capture the range off effects predicted for a wide range of dips and rupture depths by numerical modeling and GMPEs that incorporate an explicit hanging wall representation.

The primary application of HW effects in ground-motion estimation has been to cases where a known fault is explicitly modeled in the hazard assessment. This condition is expected to be relatively rare in CENA as few fault-specific sources have been identified. In the CEUS SSC model, only some of the RLME sources contain mapped faults and the rest are represented by pseudo-faults with specified orientations. The impact on hazard of applying HW effects to distributed seismicity sources is investigated in Section 14.1.

13.4.2 Summary of Available Candidate Models

The most complete set of HW models consist of those included in the Abrahamson et al. (2014), Campbell and Bozorgnia (2014), and Chiou and Youngs (2014) NGA-West2 GMPEs. These models were developed using a combination of empirical data from the limited number of earthquakes with HW site recordings and numerical simulations to model the effects of dip, depth, and rupture size through the magnitude term and rupture width W. The SSHAC Level 3 SWUS project (GeoPentech 2015) reviewed these models and developed a composite model to

represent the epistemic uncertainty in assessing HW effects. The SWUS composite model is given by the expression

$$f_{HW} = \begin{cases} C_1 \cos(dip) \times \left[C_2 + (1-C_2)\tanh\left(\frac{C_3 R_X}{W\cos(dip)}\right)\right] \times [1 + C_4(M-7)] \times \\ \qquad \left[1 - \frac{R_{JB}}{R_{RUP}+0.1}\right] \times \max\left\{0, 1 - \frac{Z_{TOR}}{12\ km}\right\} & \text{for } R_X \geq 0 \\ 0 & \text{for } R_X < 0 \end{cases} \qquad (13\text{–}11)$$

The first line of Equation 13−11 models the HW effect for sites above the rupture ($R_{JB} = 0$) in terms of fault dip, rupture width (W), magnitude, and the distance metric R_X representing the horizontal distance to the site from the surface projection of the rupture measured perpendicular to the rupture strike, with positive values of R_X on the hanging wall side and negative values on the footwall site. The second line contains two tapers, one to model the decay of the HW effect with increasing distance from the rupture (increasing R_{JB}), and one to model the decay of the HW effect with increasing rupture depth. Table 13−10 lists the parameters of the SWUS HW model. The coefficients for frequencies of 40 and 25 Hz were obtained by interpolation of the values given in Table 13−10 in with respect to log(frequency). GeoPentech (2015) did not provide coefficients for PGV. Review of the NGA West2 GMMs from which the GeoPentech (2015) HW model was derived indicate that the PGV coefficients are intermediate between periods of 2 and 3 seconds. Therefore, the PGV coefficients listed in Table 13−10 are computed as the average of the values for those two periods. Epistemic uncertainty in the HW effect is modeled by five equally weighted alternative values for coefficient C_1; for sites on the footwall site of the rupture ($R_x < 0$) $f_{HW} = 0$. The term f_{HW} is to be added to the natural log of median ground motions.

Figure 13−19 shows how the amplitude of f_{HW} for PGA varies with ground-motion frequency, magnitude, and dip for sites located on the HW side of ruptures with dips of 30, 45, and 60°. These results were computed for earthquakes occurring in a crust with a seismogenic thickness of 17 km. For sites directly above the rupture ($R_{JB} = 0$), f_{HW} increases with increasing R_X to a peak at a point over the downdip edge of the rupture. Then, as distance increases beyond the downdip edge, f_{HW} decreases with increasing R_X as R_{JB} begins to increase from 0. At distances of about 50 km from the rupture, f_{HW} becomes negligible.

13.4.2.1 Evaluation and Selection of HW Effects Model

Because the SWUS HW model represents a SSHAC Level 3 evaluation of HW effects and because no more recent models have been proposed, the SWUS HW model is selected as the best representation available.

13.4.2.2 Implementation of HW Effects Model

As indicated above, epistemic uncertainty in $\mathbf{f}_{HW}$ is modeled by five equally likely values of coefficient C_1 listed in Table 13−10. Table 13−11 lists the assignments of the five models to the 17 median ground motion models. The assignments were made such that each HW model has approximately equal weight based on the weights assigned to the median models averaged over the frequency range of 1 Hz to PGA and the number of occurrences of each model is approximately equal (3 to 4 occurrences of each).

13.4.3 Application Recommendations

The SWUS HW effects model is an additive model meant to be applied to GMPEs centered on footwall or neutral site ground-motion estimates. The developers of the Pezeshk et al. and the Shahjouei and Pezeshk candidate GMPEs explicitly state that they did not include HW effects. Many of the other candidate GMPEs are based on point-source simulations, which would not explicitly address HW effects. The Hassani and Atkinson referenced empirical GMPE uses RJB, which nominally accounts for HW effects, but the model is converted to R_{RUP} for use, which likely negates any incorporation of HW effects in the resulting seed model. Frankel included thrust faults in his finite-fault simulations, so some HW data are included in developing the Frankel candidate model, but HW effects are not explicitly addressed in its development. On the whole, it appears to be reasonable to consider that the median models developed in Chapter 9 can be considered approximately footwall or neutral site centered. Thus, the SWUS HW model is considered to be appropriate to apply to the median NGA East models. It is directly applicable to cases where there is a nearby identified dipping active fault (e.g., for sites in close proximity to the Reelfoot thrust). The use of the HW model in computing hazard from distributed seismicity sources is evaluated in Section 14.1.

13.5 Summary

Three adjustments to the median GMPEs were developed in this section, adjustments for greater attenuation rates in the Gulf Coast Region, adjustments for the depth of ruptures, and adjustments for sites located on the HW side of dipping faults. The effect of these adjustments on seismic hazard computations is evaluated in Section 14.1.

13.6 References

Abrahamson, N.A., W.J. Silva, and R. Kamai. 2014. “Summary of the ASK14 Ground-Motion Relation for Active Crustal Regions.” Earthq. Spectra, 30(3): 1025–1055.

Al Noman, M.N., and C.H. Cramer. 2015. “Empirical Ground-Motion Prediction Equations for Eastern North America (Chapter 8),” in NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region, Pacific Earthquake Engineering Research Center, PEER Report No. 2015/04, University of California, Berkeley, CA., pp. 193–212.

Benjamin, J.R., and C.A. Cornell. 1970. Probability, Statistics, and Decisions for Civil Engineers, McGraw-Hill, New York, 684 pgs.

Bozorgnia, Y., N.A. Abrahamson, L., Al Atik, T.D Ancheta, G.M. Atkinson, J.W. Baker, A. Baltay, D.M. Boore, K.W. Campbell, B.S.-J Chiou, R.B. Darragh, S. Day, J. Donahue, R.W. Graves, N. Gregor, T. Hanks, I.M. Idriss, R. Kamai, T. Kishida, A. Kottke, S.A. Mahin, S. Rezaeian, B. Rowshandel, E. Seyhan, S. Shahi, T. Shantz, W.J. Silva, P. Spudich, J.P. Stewart, J. Watson-Lamprey, K. Wooddell, and R.R. Youngs. 2014. “NGA-West2 Research Project.” Earthq. Spectra, 30(3): 973–987.

Campbell, K.W., and Y. Bozorgnia. 2014. “NGA-West2 Ground Motion Model for the Average Horizontal Components of PGA, PGV, and 5% Damped Linear Acceleration Response Spectra.” Earthq. Spectra 30(3): 1087–1115.

Chapman, M., and A. Conn. 2016. “A Model of Lg Propagation in the Gulf Coast Plain of the Southern United States.” Bull. Seismol. Soc. Am., 106(2): 349–363.

Chiou, B.S.-J., and R.R. Youngs. 2008. “NGA Model for Average Horizontal Component of Peak Ground Motion and Response Spectra.” PEER Report No. 2008/09, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Chiou, B.S-J., and R.R. Youngs. 2014. “Update of the Chiou and Youngs NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra.” Earthq. Spectra, 30(3): 1117–1153.

Cramer, C.H., and N. Al Noman. 2016. “Improving Regional Ground Motion Attenuation Boundaries and Models in the CEUS and Developing a Gulf Coast Empirical GMPE using Earthscope US Array Data for Use In the National Seismic Hazards Mapping Project.” US Geological Survey External Grants Program Final Technical Report, Award Number G14AP00049, August, 52 pgs.

Dreiling, J., M.P. Isken, W.D. Mooney, M.C. Chapman, and R.W. Godbee. 2014. “NGA-East Regionalization Report: Comparison of Four Crustal Regions Within Central and Eastern North America Using Waveform Modeling And 5%-Damped Pseudo-Spectral Acceleration Response.” PEER Report No. 2014/15, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

EPRI. 1993. “Guidelines for Determining Design Basis Ground Motions. Volume 1: Method and Guidelines for Estimating Earthquake Ground Motion in Eastern North America.” Electric Power Research Institute, Technical Report, Palo Alto, CA.

EPRI. 2004. “CEUS Ground Motion Project Final Report.” Technical Report 1009684. Electric Power Research Institute, Palo Alto, CA.

EPRI. 2013. “EPRI 2004-2006 Ground-Motion Model (GMM), Final Report Project 3002000717.” June 2013, Electric Power Research Institute, Palo Alto, CA.

EPRI. 2015. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities: Review for Reservoir-Induced Seismicity (RIS) in the Southeast and Earthquakes in South Carolina Near the 1886 Charleston Earthquake.” Final Report Project 3002005288. September 2015, Electric Power Research Institute, Palo Alto, CA.

EPRI/USDOE/USNRC. 2012. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Florida Power and Light, Inc. 2016. Turkey Point Units 6&7 COLA Section 2.5.2, Rev 8. https://www.nrc.gov/docs/ML1625/ML16250A316.pdf

GeoPentech. 2015, Southwestern United States Ground Motion Characterization SSHAC Level 3-Technical Report, Rev. 2, March 2015.

Goulet, C.A., T. Kishida. T.D. Ancheta, C.H. Cramer, R.B. Darragh, W.J. Silva, Y.M.A. Hashash, J. Harmon, J.P. Stewart, K.E. Wooddell, and R.R. Youngs. 2014. “PEER NGA-East Database.” PEER Report No. 2014/17, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

PEER. 2015a. “NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

PEER. 2015b. “NGA-East: Adjustments to Median Ground-Motion Models for Central and Eastern North America.” PEER Report No. 2015/08, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Somerville, P.G. 2014. “Scaling Relations between Seismic Moment and Rupture Area of Earthquakes in Stable Continental Regions.” PEER Report No. 2014/14, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Somerville, P.G., and N.A. Abrahamson. 1995. “Prediction of Ground Motions for Thrust Earthquakes.” Report to Calif. Strong Motion Instrumentation Program, Sacramento, CA.

Thomas, W.A., T.M. Chowns, D.L. Daniels, T.L. Neathery, Lynn Glover III, and R.J. Gleason. 1989. The Subsurface Appalachians beneath the Atlantic and Gulf Coastal Plains, Chapter 10 of The Geology of North America, v. F-2, The Appalachian-Ouachita Orogen in the United States, The Geological Society of America.

Young, R.R. 2015. Personal communication.

Table 13–1 Coefficients for DASG GCR adjustment model [Equation (13-3)].

Ground-motion frequency (Hz)	γ(f)
0.1	-2.67E-04
0.133	-2.92E-04
0.2	-3.35E-04
0.25	-3.65E-04
0.333	-4.07E-04
0.5	-4.85E-04
0.667	-5.40E-04
1	-6.42E-04
1.333	-7.20E-04
2	-8.42E-04
2.5	-9.02E-04
3.333	-9.71E-04
4	-1.02E-03
5	-1.06E-03
6.667	-1.09E-03
10	-1.11E-03
13.333	-1.10E-03
20	-1.06E-03
25	-1.04E-03
33.333	-1.01E-03
40	-9.97E-04
50	-9.74E-04
100	-9.20E-04
PGA	-9.12E-04
PGV	-4.95E-04

Table 13–2(a) Coordinates of GCR/MCR boundaries for the GCR_{LARGE} regions.

Longitude (deg)	Latitude (deg)	Longitude (deg)	Latitude (deg)	Longitude (deg)	Latitude (deg)
-88.19936	31.67075	-88.43882	26.98429	-95.62877	33.77953
-87.82138	31.51676	-88.84571	27.01285	-95.2367	33.9735
-87.27002	31.51676	-89.4025	27.01998	-94.8839	34.0924
-85.48914	31.78623	-90.13061	27.00571	-94.5325	34.1942
-84.16609	31.98643	-90.88728	26.94146	-94.2431	34.2876
-82.53244	31.5576	-91.78672	26.82011	-93.9273	34.3915
-81.65436	31.19003	-92.67902	26.67734	-93.9063	34.4202
-80.91922	31.08793	-93.64984	26.60596	-92.8443	34.8499
-79.95816	31.01809	-94.28515	26.54171	-92.5124	34.9785
-79.05916	31.3506	-95.04182	26.46319	-92.1973	34.9049
-78.37686	31.62734	-95.45585	26.36325	-91.31839	35.65223
-78.29184	31.41121	-95.79135	26.19907	-90.4552	36.65989
-77.87075	31.10429	-96.06261	26.06344	-89.87543	37.17804
-77.47918	31.08481	-96.20538	25.97778	-89.45512	37.50415
-77.04531	31.26995	-96.38384	25.82074	-89.34583	37.84227
-77.06245	30.91646	-96.57657	25.58517	-89.00005	37.79535
-77.02677	30.56914	-96.75503	25.24253	-88.68714	37.75289
-76.99081	30.40436	-96.83318	25.03421	-88.42463	37.70362
-76.67064	30.28973	-96.85497	24.79281	-88.25009	37.67544
-76.27519	30.29119	-96.85497	24.48586	-87.89861	36.87449
-76.11884	30.24617	-96.84329	24.06827	-87.98561	36.47963
-76.04288	30.16769	-96.82642	23.46507	-90.31618	34.71705
-76.06076	30.01383	-96.79786	23.02963	-90.86069	34.02538
-76.11884	29.88315	-96.8012	22.98298	-91.0589	33.7567
-76.19145	29.82506	-97.32704	22.93454	-91.017	33.7009
-76.43316	29.69203	-97.37848	22.93546	-90.889	33.5592
-76.6173	29.45565	-97.5178	22.93513	-90.6863	33.4072
-76.65646	29.17741	-97.9082	22.97051	-90.4851	33.2785
-76.69728	29.02375	-97.93124	23.23314	-90.2056	33.0773
-76.56296	28.30955	-98.04224	23.85473	-89.8993	32.8367
-76.47972	27.54843	-98.35303	24.56511	-89.4511	32.475
-76.45592	27.16787	-98.77482	24.9869	-89.1971	32.2692
-76.52729	26.64459	-99.44081	25.78608	-89.065	32.2118
-76.84446	26.1704	-99.6628	26.49647	-88.81263	32.04757
-77.06608	25.42433	-99.95139	27.00705	-88.56027	31.88334
-77.25548	25.03625	-100.19559	27.51764	-88.37273	31.74138
-77.44579	24.72619	-100.59518	28.13922	-88.19936	31.67075
-77.86154	24.16199	-100.90597	28.42782		
-78.18235	23.76264	-101.15017	28.87181		
-78.6496	23.19974	-101.17236	29.1826		
-78.92136	22.89686	-100.1903	29.2302		

Longitude (deg)	Latitude (deg)	Longitude (deg)	Latitude (deg)	Longitude (deg)	Latitude (deg)
-79.39371	23.04914	-99.3228	29.1917		
-79.88964	23.19339	-98.7361	29.2406		
-80.72599	23.34348	-98.4462	29.1686		
-81.52701	23.41376	-98.2451	29.1525		
-82.02143	23.41284	-98.1339	29.1789		
-82.63713	23.36029	-97.9322	29.2738		
-83.37672	23.20048	-97.8784	29.3154		
-83.59716	23.13088	-97.7309	29.4551		
-83.91662	23.15992	-97.4717	29.821		
-84.27968	23.30516	-97.417	30.0247		
-84.57007	23.56651	-97.4331	30.467		
-84.78736	23.79594	-97.4081	30.531		
-85.12478	24.20635	-97.2297	30.7476		
-85.3663	25.00719	-97.1225	30.9378		
-85.53156	25.3377	-96.9971	31.244		
-85.74766	25.65549	-96.7996	32.0152		
-86.04031	25.99774	-96.731	32.4496		
-86.13312	26.10627	-96.7201	32.9506		
-86.55428	26.46319	-96.6645	33.0794		
-86.9112	26.67734	-96.5015	33.3864		
-87.53224	26.84866	-96.12066	33.57197		

Table 13–2(b) Coordinates of GCR/MCR boundaries for the GCR_{SMALL} regions.

Longitude (deg)	Latitude (deg)	Longitude (deg)	Latitude (deg)
-87.82138	31.51676	-100.90597	28.42782
-87.27002	31.51676	-101.15017	28.87181
-85.48914	31.78623	-101.17236	29.1826
-84.2172	31.3474	-100.1903	29.2302
-83.6302	30.7145	-99.3228	29.1917
-82.9666	30.1212	-98.7361	29.2406
-82.5348	29.3689	-98.4462	29.1686
-82.2291	28.267	-98.2451	29.1525
-83	28	-98.1339	29.1789
-83.37672	23.20048	-97.057	29.481
-83.59716	23.13088	-96.3155	30.0799
-83.91662	23.15992	-95.3106	31.3567
-84.27968	23.30516	-94.556	31.3331
-84.57007	23.56651	-93.8896	31.4167
-84.78736	23.79594	-92.9188	32.0227
-85.12478	24.20635	-92.1521	32.0617
-85.3663	25.00719	-91.4827	31.8803
-85.53156	25.3377	-90.6856	32.0422
-85.74766	25.65549	-89.8958	32.0322
-86.04031	25.99774	-87.82138	31.51676
-86.13312	26.10627		
-86.55428	26.46319		
-86.9112	26.67734		
-87.53224	26.84866		
-88.43882	26.98429		
-88.84571	27.01285		
-89.4025	27.01998		
-90.13061	27.00571		
-90.88728	26.94146		
-91.78672	26.82011		
-92.67902	26.67734		
-93.64984	26.60596		
-94.28515	26.54171		
-95.04182	26.46319		
-95.45585	26.36325		
-95.79135	26.19907		
-96.06261	26.06344		
-96.20538	25.97778		
-96.38384	25.82074		
-96.57657	25.58517		
-96.75503	25.24253		

Longitude (deg)	Latitude (deg)	Longitude (deg)	Latitude (deg)
-96.83318	25.03421		
-96.85497	24.79281		
-96.85497	24.48586		
-96.84329	24.06827		
-96.82642	23.46507		
-96.79786	23.02963		
-96.8012	22.98298		
-97.32704	22.93454		
-97.37848	22.93546		
-97.5178	22.93513		
-97.9082	22.97051		
-97.93124	23.23314		
-98.04224	23.85473		
-98.35303	24.56511		
-98.77482	24.9869		
-99.44081	25.78608		
-99.6628	26.49647		
-99.95139	27.00705		
-100.19559	27.51764		
-100.59518	28.13922		

Table 13–3 Coefficients for source-depth scaling factor model [Equation (13-5)].

Frequency/GMIM	b_1	b_2
0.10	-0.01653	0.0173
0.13	-0.01653	0.0173
0.20	-0.01653	0.0173
0.25	-0.01653	0.0173
0.33	-0.01351	0.02875
0.50	-0.00607	0.03055
0.67	-0.00111	0.03174
1.00	0.00479	0.02996
1.33	0.01003	0.02646
2.00	0.02279	0.0161
2.50	0.03069	0.00821
3.33	0.03958	-0.00069
4.00	0.04658	-0.00768
5.00	0.05346	-0.01457
5.88	0.05346	-0.01457
6.67	0.05346	-0.01457
8.33	0.05346	-0.01457
≥10.00	0.05346	-0.01457

Table 13–4 Magnitude dependence of Z_{TOR} from the CEUS SSC (interim model).

M	Z_{TOR} strike–slip (km)	Z_{TOR} reverse (km)	Z_{TOR} average (km)
4	9.54	9.67	9.60
4.1	9.48	9.63	9.55
4.2	9.42	9.58	9.50
4.3	9.35	9.53	9.44
4.4	9.27	9.47	9.37
4.5	9.18	9.41	9.29
4.6	9.08	9.34	9.21
4.7	8.97	9.25	9.11
4.8	8.84	9.16	9.00
4.9	8.70	9.06	8.88
5	8.54	8.95	8.74
5.1	8.36	8.82	8.59
5.2	8.16	8.67	8.42
5.3	7.94	8.51	8.22
5.4	7.69	8.33	8.01
5.5	7.40	8.13	7.77
5.6	7.09	7.90	7.49
5.7	6.73	7.64	7.19
5.8	6.33	7.35	6.84
5.9	5.88	7.03	6.46
6	5.38	6.67	6.03
6.1	4.82	6.26	5.54
6.2	4.19	5.81	5.00
6.3	3.48	5.30	4.39
6.4	2.68	4.72	3.70
6.5	1.79	4.08	2.93
6.6	0.79	3.36	2.07
6.7	0.00	2.55	1.27
6.8	0.00	1.64	0.82
6.9	0.00	0.61	0.31
7	0.00	0.00	0.00
7.1	0.00	0.00	0.00
7.2	0.00	0.00	0.00
7.3	0.00	0.00	0.00
7.4	0.00	0.00	0.00
7.5	0.00	0.00	0.00
7.6	0.00	0.00	0.00
7.7	0.00	0.00	0.00
7.8	0.00	0.00	0.00
7.9	0.00	0.00	0.00
8	0.00	0.00	0.00
8.1	0.00	0.00	0.00
8.2	0.00	0.00	0.00

Table 13–5 Magnitude dependence of Z_{TOR} recommended by candidate GMM developers.

Model / M	4	4.5	5	5.5	6	6.5	7	7.5	8	8.2	Data notes and modelers recommendation
Boore (all)	15.3	14.8	14	11.3	8.8	4.3	0	0	0	0	Median Z_{TOR} for 9 events between **M** 4 and **M** 5.5 was 12 km. Recommended use of CEUS SSC interim average adjusted Z_{TOR} =14 km for **M** = 5.
DASG (all)	13.3	12.9	12.1	10.8	8.4	4.1	0	0	0	0	Assumed Z_{TOR} for simulations to be average 8 km with a standard deviation of ln(0.6). Use CEUS SSC interim average, adjusted to Z_{TOR} =13.3 km at **M** = 3.9.
YA	9.44	9.19	8.83	8.31	7.56	6.48	4.91	2.64	0	0	Z_{TOR} = max(10-0.5*0.6*(10^(-1.01+0.32***M**)), 0)
PZCT (all)	11.7	11.3	10.7	9.5	7.3	3.6	0	0	0	0	Used all non-Gulf, non-PIE events **M** > 3 for calibration. Use CEUS SSC interim average adjusted to Z_{TOR} =11.7 km at **M** = 3.9.
Frankel	5	5	5	5	5	5	5	1	1	1	Use depths from simulations directly: Z_{TOR} =5 km, for **M** < 7.5, and Z_{TOR} = 1 km for **M** ≥ 7.5,
SP	9.6	9.29	4	4	4	3	3	2.5	2.5	2.5	Based on simulation depths (Table 7.1 in PEER 2015a), use CEUS SSC interim average Z_{TOR} for **M** < 5.
Grazier	8.9	8.6	8.1	7.2	5.6	2.7	0	0	0	0	Used 48 events with **M** > 3.75, R_{RUP}<1000 km and recommended the use of CEUS SSC interim average adjusted for Z_{TOR} = 8.9 km at **M** = 4.7.
HA	9.44	9.19	8.83	8.31	7.56	6.48	4.91	2.64	0	0	Used an average of focal depth of 8 km; recommended to use the same Z_{TOR} distribution as YA15.
PEER	9.6	9.29	8.74	7.77	6.03	2.93	0	0	0	0	Used the CEUS SSC interim average directly.
Average	**10.23**	**9.95**	**8.92**	**8.02**	**6.7**	**4.29**	**1.98**	**0.98**	**0.39**	**0.39**	

Table 13–6 Smoothed median seed GMM-based magnitude dependence of $E(Z_{TOR})$.

M	$E(Z_{TOR})$(km)
4	10.0
4.5	10.0
5	8.9
5.5	8.0
6.0	6.7
6.5	4.3
7	2.0
7.5	0.0
8	0.0
8.2	0.0

Table 13–7 Smoothed focal depth distributions for CEUS earthquakes.

Focal Depth (km)	Smoothed Cumulative Distribution	
	DEEP Region	**Rest of CEUS Region**
0	0	0
0.5	0	0
1.5	0.0045	0.0123
2.5	0.0135	0.0362
3.5	0.0269	0.0748
4.5	0.0449	0.1245
5.5	0.0673	0.1865
6.5	0.0942	0.2681
7.5	0.1256	0.3595
8.5	0.1615	0.4601
9.5	0.2019	0.5442
10.5	0.2468	0.6092
11.5	0.2962	0.6650
12.5	0.3500	0.7129
13.5	0.4100	0.7552
14.5	0.4700	0.7920
15.5	0.5300	0.8252
16.5	0.5900	0.8454
17.5	0.6429	0.8643
18.5	0.6922	0.8821
19.5	0.7378	0.8986
20.5	0.7797	0.9138
21.5	0.8180	0.9278
22.5	0.8526	0.9406
23.5	0.8836	0.9522
24.5	0.9109	0.9625
25.5	0.9346	0.9715
26.5	0.9547	0.9794
27.5	0.9710	0.9860
28.5	0.9838	0.9913
29.5	0.9928	0.9955
30.5	0.9982	0.9984
31.5	1	1

Table 13–8 Smoothed HDR distributions developed from NGA-West2 database.

HDR	Cumulative distribution function for:		
	Normal	**Strike-Slip**	**Reverse**
0	0	0	0
0.1	0	0.025	0.03
0.2	0	0.05	0.08
0.3	0	0.09	0.14
0.4	0	0.15	0.25
0.5	0	0.24	0.36
0.6	0.05	0.35	0.47
0.7	0.1	0.53	0.58
0.8	0.25	0.72	0.69
0.9	0.5	0.86	0.83
1.0	1	1	1

Table 13–9(a) Example PGA depth scaling for a M 6 strike–slip earthquake [$E(Z_{TOR}) = 6.7$ km, $f_{Ztor,M} = 0.04375$].

Z_{TOR} (km)	Cumulative Probability	Z_{TOR} Increment (km)	Probability	ΔZ_{TOR} (km)	$f_{Z_{TOR}}$
0	0	0.25	0.0805	-4.45	-0.1947
0.5	0.0805	0.75	0.0491	-3.95	-0.1728
1	0.1296	1.25	0.0546	-3.45	-0.1509
1.5	0.1842	1.75	0.0597	-2.95	-0.1291
2	0.2439	2.25	0.0640	-2.45	-0.1072
2.5	0.3079	2.75	0.0667	-1.95	-0.0853
3	0.3746	3.25	0.0685	-1.45	-0.0634
3.5	0.4431	3.75	0.0681	-0.95	-0.0416
4	0.5112	4.25	0.0651	-0.45	-0.0197
4.5	0.5763	4.75	0.0618	0	0
5	0.6381	5.25	0.0582	0	0
5.5	0.6963	5.75	0.0538	0	0
6	0.7501	6.25	0.0499	0	0
6.5	0.8000	6.75	0.0461	0	0
7	0.8461	7.25	0.0424	0	0
7.5	0.8885	7.75	0.0394	0	0
8	0.9279	8.25	0.0365	0	0
8.5	0.9644	8.75	0.0337	0.05	0.0022
9	0.9981	9.25	0.0019	0.55	0.0241
9.5	1				

Table 13–9(b) Example PGA depth scaling for a M 6 reverse earthquake [$E(Z_{TOR}) = 6.7$ km, $f_{Ztor,M} = 0.04375$].

Z_{TOR} (km)	Cumulative Probability	Z_{TOR} Increment (km)	Probability	ΔZ_{TOR} (km)	$f_{Z_{TOR}}$
0	0	0.25	0.0576	-4.45	-0.1947
0.5	0.0576	0.75	0.0362	-3.95	-0.1728
1	0.0938	1.25	0.0415	-3.45	-0.1509
1.5	0.1353	1.75	0.0464	-2.95	-0.1291
2	0.1817	2.25	0.0519	-2.45	-0.1072
2.5	0.2336	2.75	0.0565	-1.95	-0.0853
3	0.2901	3.25	0.0607	-1.45	-0.0634
3.5	0.3508	3.75	0.0631	-0.95	-0.0416
4	0.4139	4.25	0.0640	-0.45	-0.0197
4.5	0.4779	4.75	0.0631	0	0
5	0.5410	5.25	0.0609	0	0
5.5	0.6019	5.75	0.0584	0	0
6	0.6603	6.25	0.0549	0	0
6.5	0.7152	6.75	0.0506	0	0
7	0.7658	7.25	0.0467	0	0
7.5	0.8125	7.75	0.0423	0	0
8	0.8548	8.25	0.0388	0	0
8.5	0.8936	8.75	0.0356	0.05	0.0022
9	0.9292	9.25	0.0335	0.55	0.0241
9.5	0.9627	9.75	0.0308	1.05	0.0459
10	0.9935	10.25	0.0065	1.55	0.0678
10.5	1				

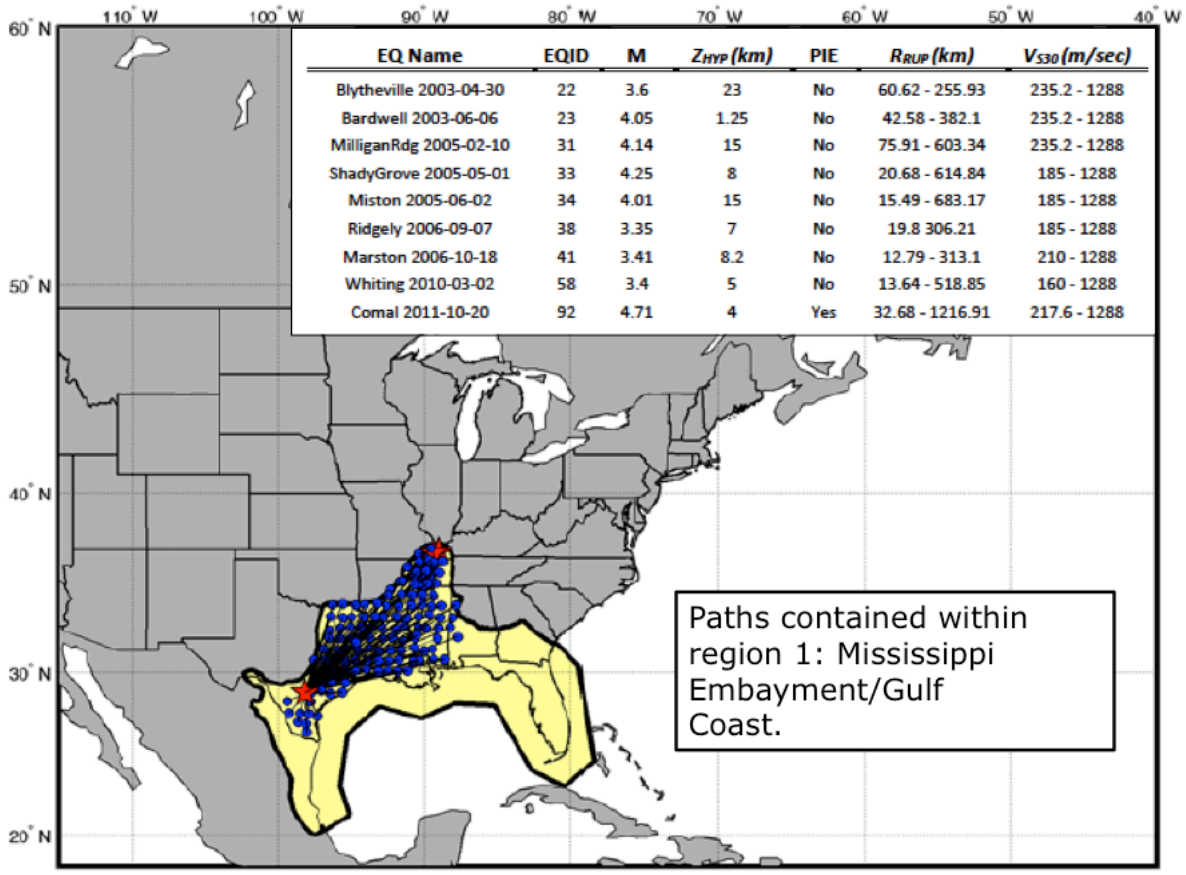
Table 13–10 Parameters of the SWUS HW effects model (Equation 13-11).

Period (sec)	Model-dependent C_1 coefficients					Coefficients held constant for all five models		
	Model HW1	**Model HW 2**	**Model HW 3**	**Model HW 4**	**Model HW 5**	**C_2**	**C_3**	**C_4**
0.01	0.868	0.982	1.038	1.095	1.209	0.2160	2.0289	0.1675
0.02	0.867	0.987	1.046	1.106	1.226	0.2172	2.0260	0.1666
0.03	0.856	0.997	1.067	1.138	1.278	0.2178	2.0163	0.1670
0.05	0.840	1.027	1.121	1.215	1.402	0.2199	1.9870	0.1699
0.075	0.857	1.041	1.133	1.226	1.410	0.2218	1.9906	0.1817
0.1	0.848	1.040	1.135	1.231	1.422	0.2213	1.9974	0.1717
0.15	0.868	1.009	1.080	1.150	1.292	0.2169	2.0162	0.1814
0.2	0.850	1.005	1.082	1.160	1.315	0.2131	1.9746	0.1834
0.25	0.868	0.985	1.044	1.102	1.219	0.1988	1.9931	0.1767
0.3	0.839	0.974	1.041	1.108	1.242	0.2019	2.0179	0.1658
0.4	0.780	0.934	1.011	1.089	1.243	0.2090	2.0249	0.1624
0.5	0.741	0.902	0.982	1.063	1.223	0.2053	2.0041	0.1719
0.75	0.613	0.869	0.997	1.125	1.380	0.1713	1.8697	0.1866
1	0.621	0.788	0.872	0.955	1.123	0.1571	1.8526	0.3143
1.5	0.506	0.662	0.740	0.818	0.974	0.1559	1.8336	0.3195
2	0.391	0.537	0.609	0.682	0.828	0.1559	1.7996	0.3246
3	0.128	0.245	0.304	0.362	0.480	0.1616	1.6740	0.3314
4	0	0.034	0.088	0.138	0.231	0.1616	1.6740	0.3314
5	0	0	0	0	0.040	0.1616	1.6740	0.3314
7.5	0	0	0	0	0	0.1616	1.6740	0.3314
10	0	0	0	0	0	0.1616	1.6740	0.3314

Table 13–11 Assignment of HW models to NGA east medians.

NGA East Median Model	HW C_1 Coefficient Table 13-10
1	3
2	5
3	1
4	3
5	4
6	4
7	1
8	2
9	2
10	5
11	5
12	2
13	4
14	5
15	4
16	3
17	1

Figure 13–1 List of earthquake events used in the Gulf Coast adjustment models. Red stars show the events locations, blue dots show station locations and the blue lines show the sampled wave propagation paths.



EQ Name	EQID	M	Z_{HYP} (km)	PIE	R_{RUP} (km)	V_{S30} (m/sec)
Blytheville 2003-04-30	22	3.6	23	No	60.62 - 255.93	235.2 - 1288
Bardwell 2003-06-06	23	4.05	1.25	No	42.58 - 382.1	235.2 - 1288
MilliganRdg 2005-02-10	31	4.14	15	No	75.91 - 603.34	235.2 - 1288
ShadyGrove 2005-05-01	33	4.25	8	No	20.68 - 614.84	185 - 1288
Miston 2005-06-02	34	4.01	15	No	15.49 - 683.17	185 - 1288
Ridgely 2006-09-07	38	3.35	7	No	19.8 306.21	185 - 1288
Marston 2006-10-18	41	3.41	8.2	No	12.79 - 313.1	210 - 1288
Whiting 2010-03-02	58	3.4	5	No	13.64 - 518.85	160 - 1288
Comal 2011-10-20	92	4.71	4	Yes	32.68 - 1216.91	217.6 - 1288

Figure 13–2 Magnitude (M) and rupture distance (R_{RUP}) ranges for records in the NGA-East database from Path Region 1 (i.e., both the earthquake source and the site are located in the GCR).

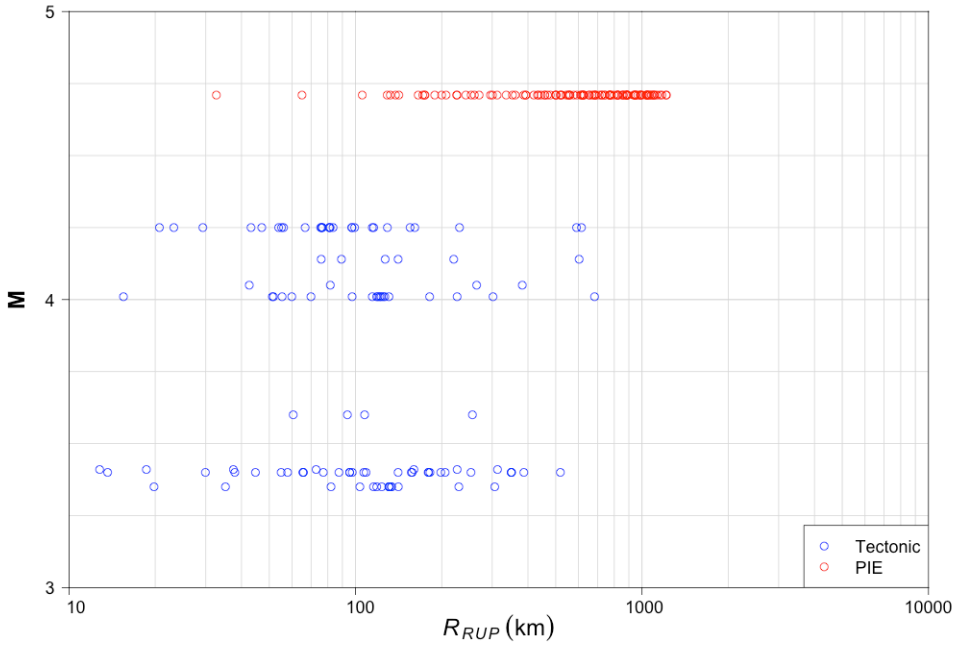


Figure 13–3 Comparison of Gulf Coast adjustment ratios between DASG modeland the PEER modelplotted against distance for PSA at frequencies 0.5, 1.0, 2.5, 5.0, 10, and 25 Hz, and PGA.

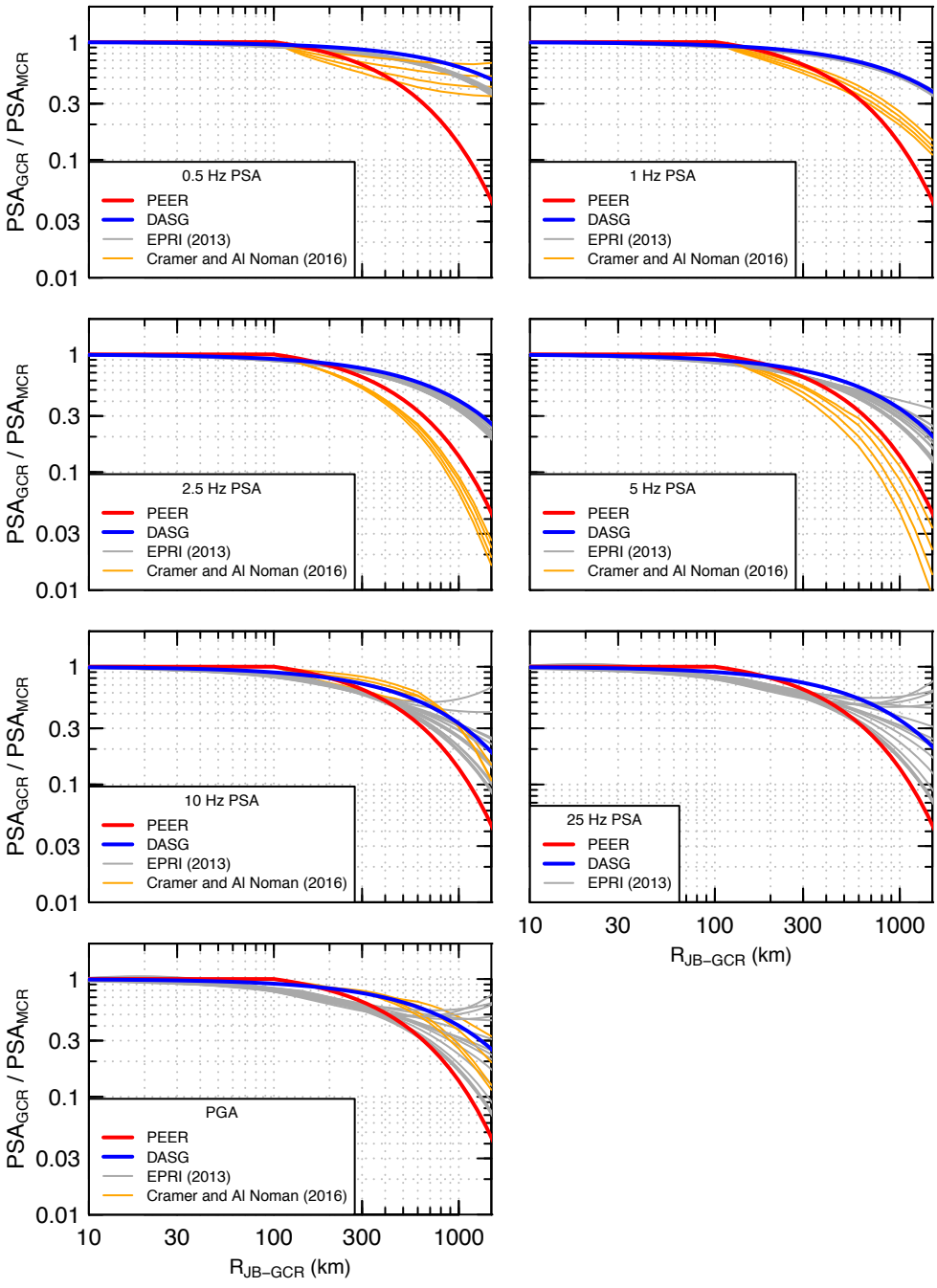


Figure 13–4 Alternative proposed boundaries for GCR.

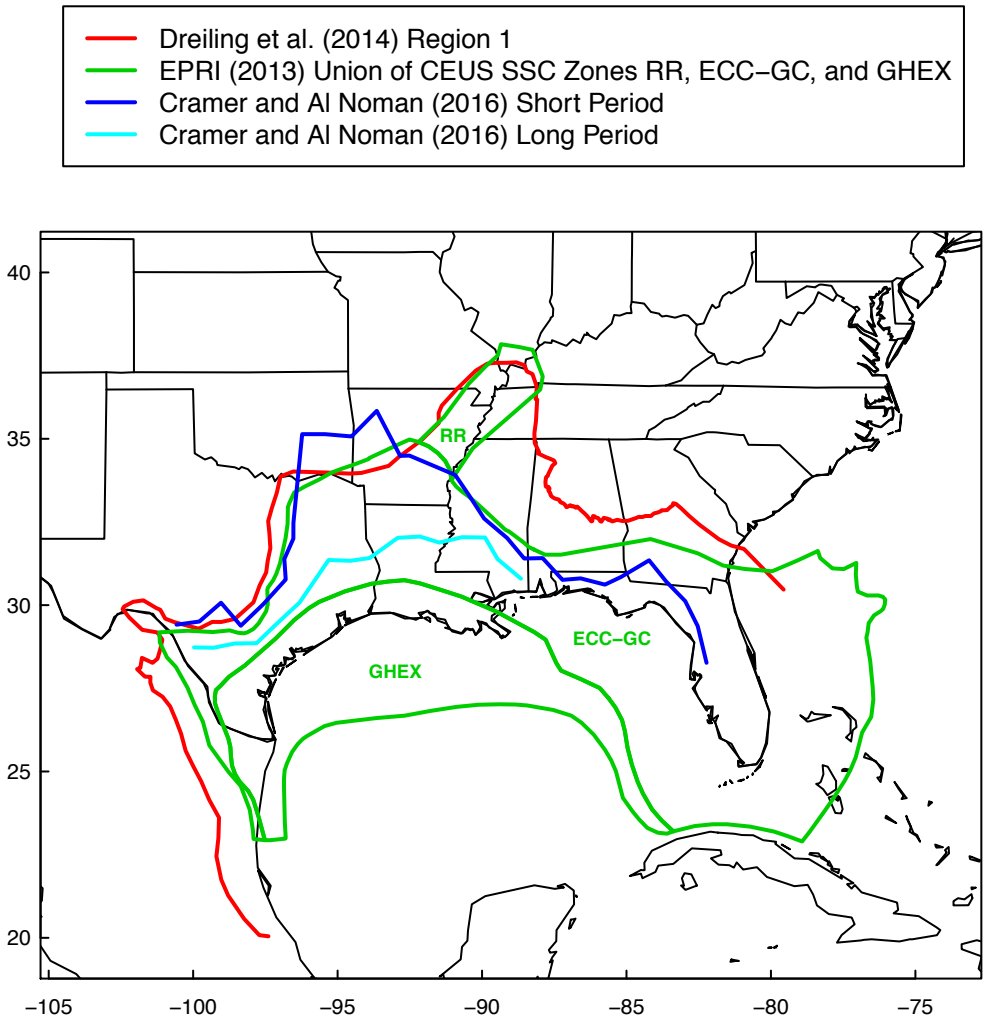


Figure 13–5 Alternative GCR zonations for applying GCR adjustments.

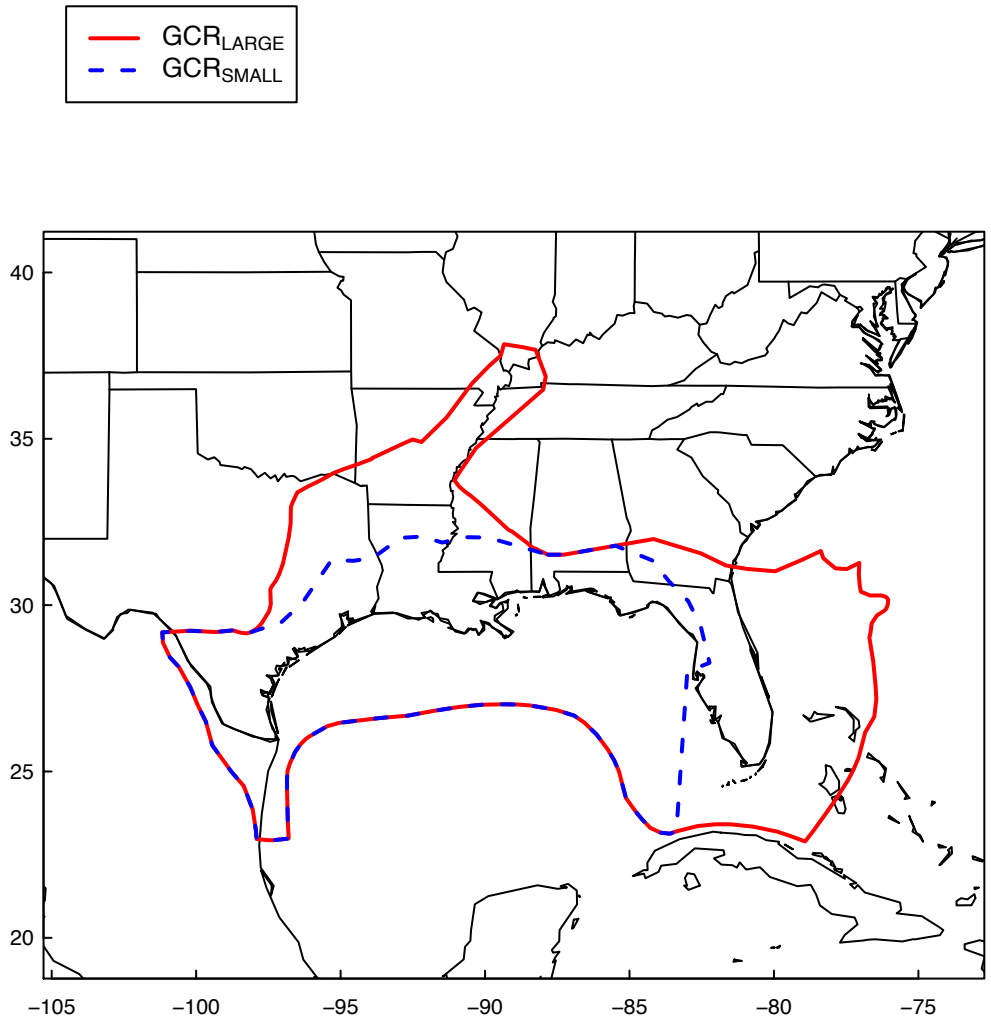


Figure 13–6 Focal depth—magnitude scatterplot for events in the NGA-East database flatfile (figure originally published in PEER 2015b).

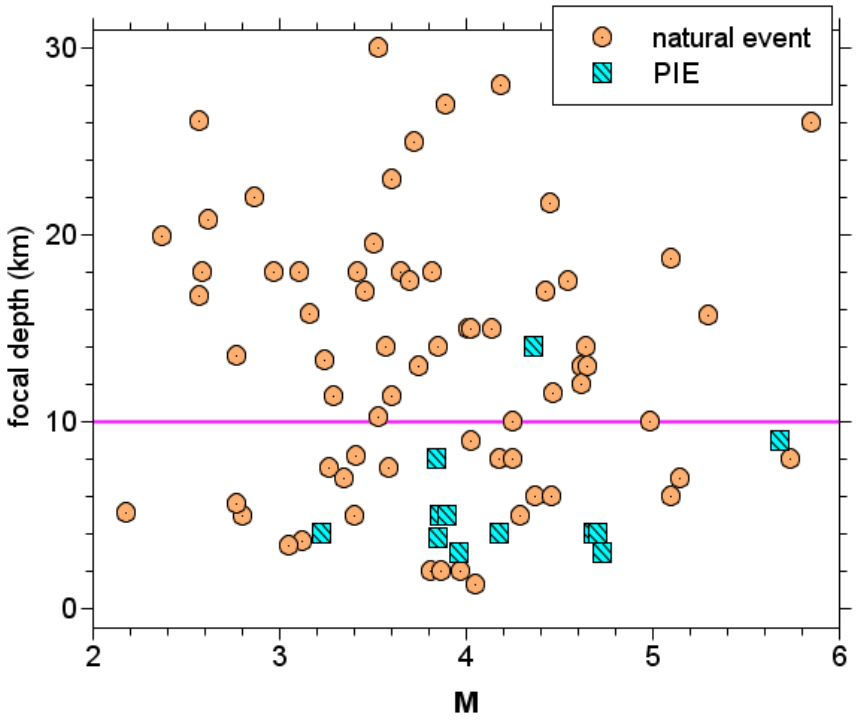


Figure 13–7 Source-depth model scaling factors for the average of the three NGA-West2 models (cyan. The NGA-West2 models and the implied source-depth scaling factor from the PEER NGA-East model are included for comparison: CY14 in dark green, ASK14 in blue and, CB14 in black, and PEER NGA-East in magenta (top M = 5.0 and bottom M = 6.5).

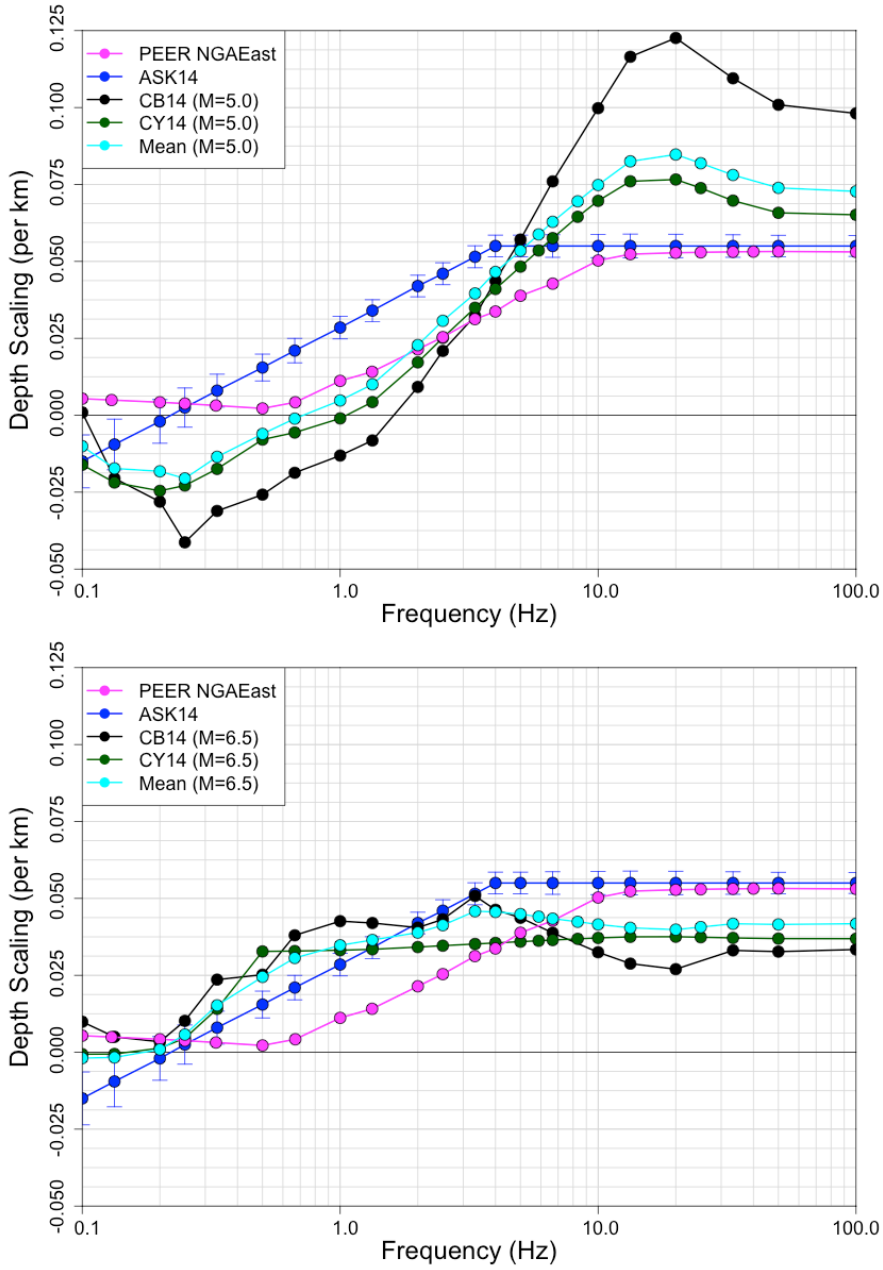


Figure 13–8 PEER hybrid source-depth model scale factors for M = 5.0, 5.5, 6.0, and 6.5.

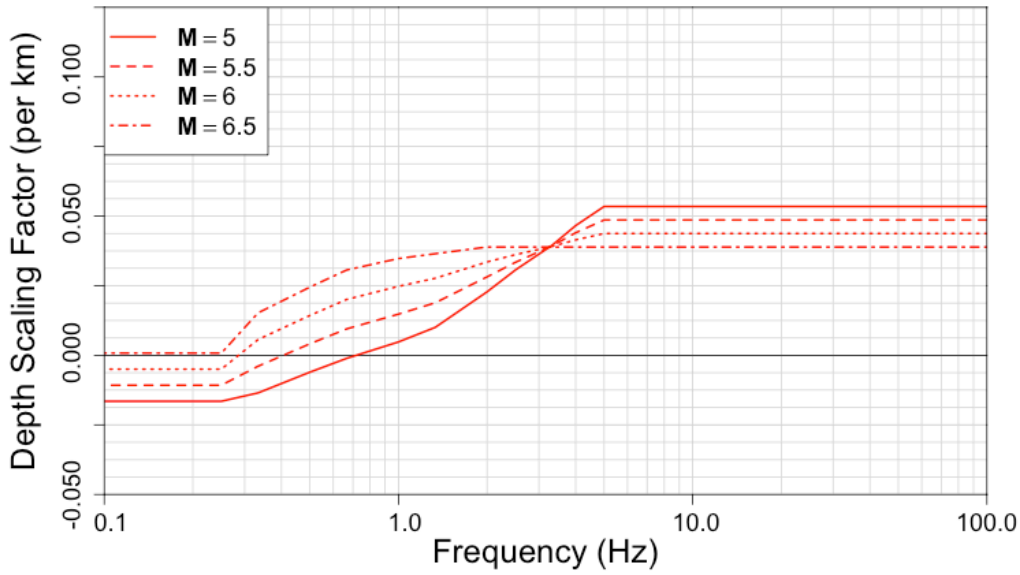


Figure 13–9 Expected values of Z_{TOR} from NGA-West2 [Chiou and Youngs (2014)] and for the CEUS region.

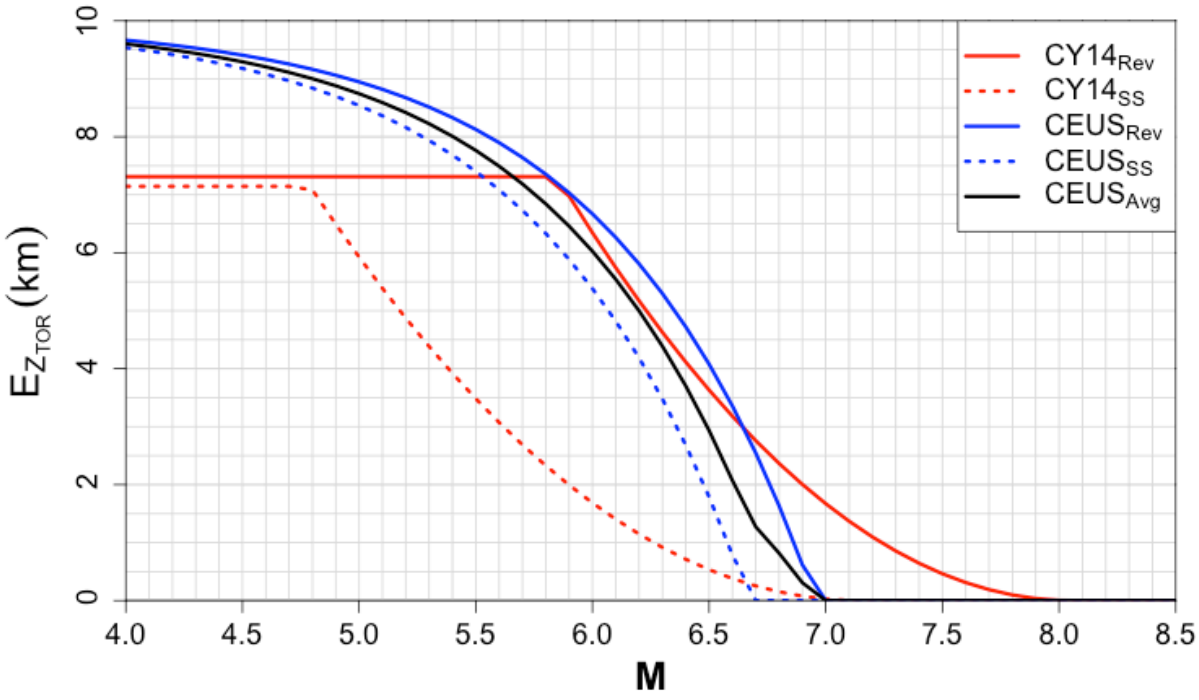


Figure 13–10 Seed GMM-based expected values of Z_{TOR} (as detailed in Table 13–3). Red diamonds show the source depth distribution for the NGA-East database.

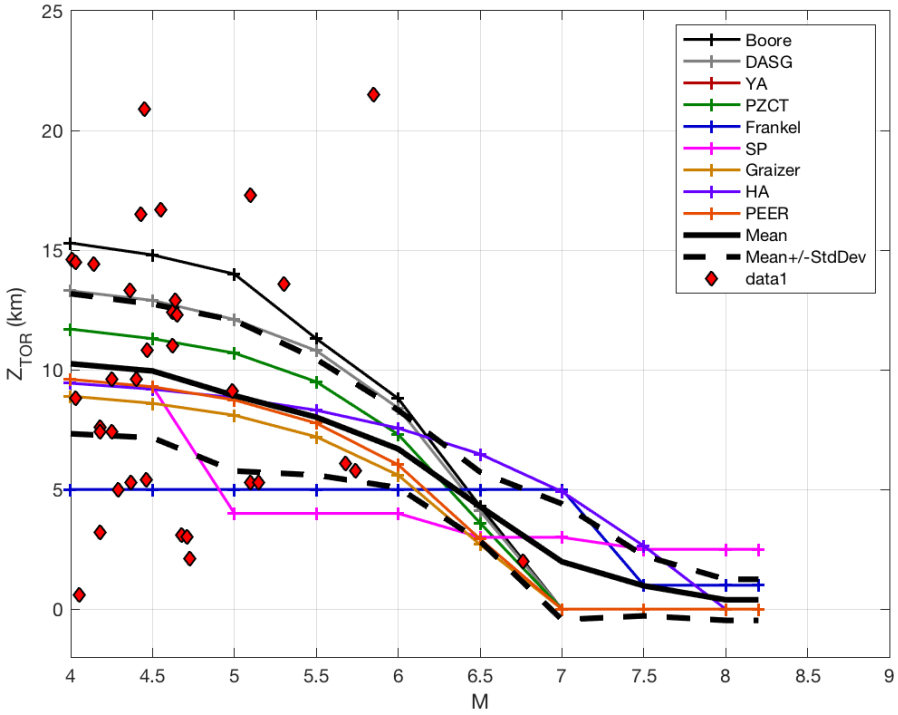


Figure 13–11 Illustration of two candidate models for depth centering Z_{TOR} (from Tables 13–1 and 13–3).

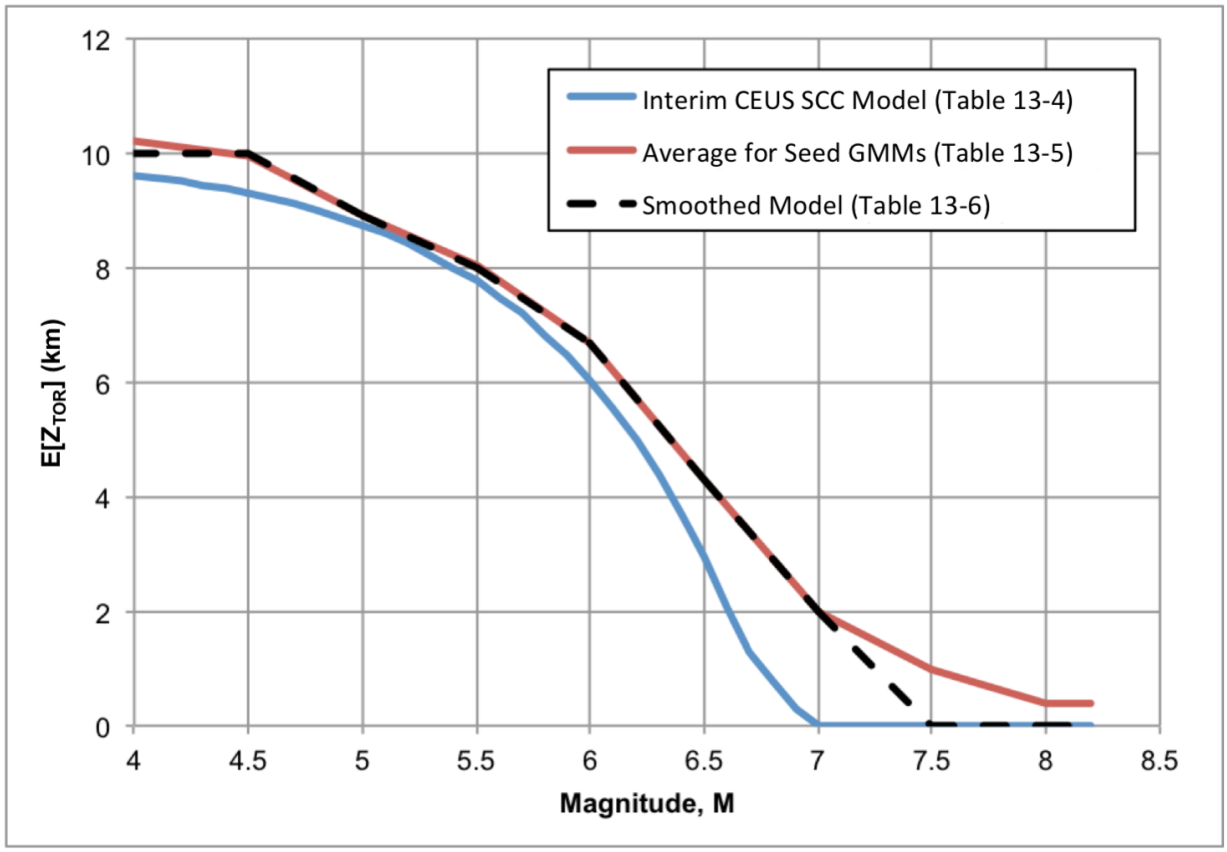


Figure 13–12 Location of earthquakes with better defined focal depths in the CEUS SSC earthquake catalog (EPRI/DOE/NRC, 2012).

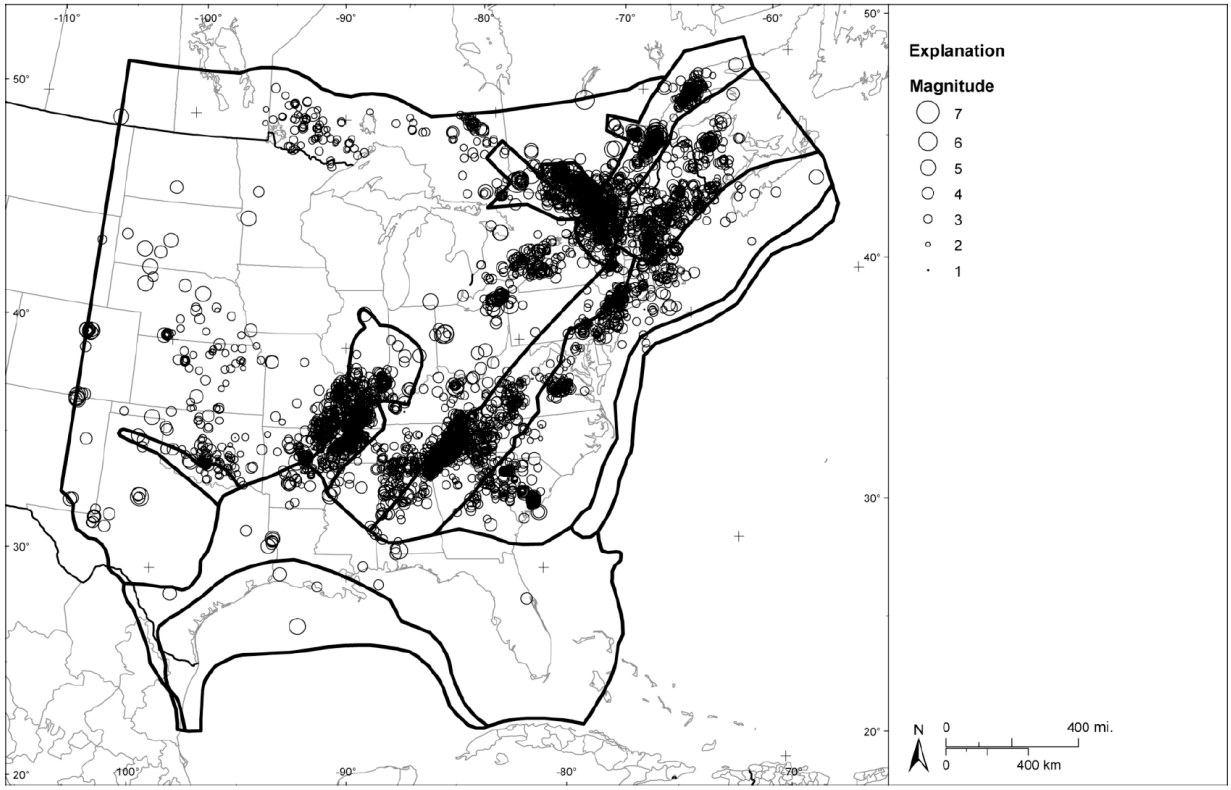


Figure 13–13 CEUS SSC Seismotectonic Source Zones (EPRI/DOE/NRC, 2012).

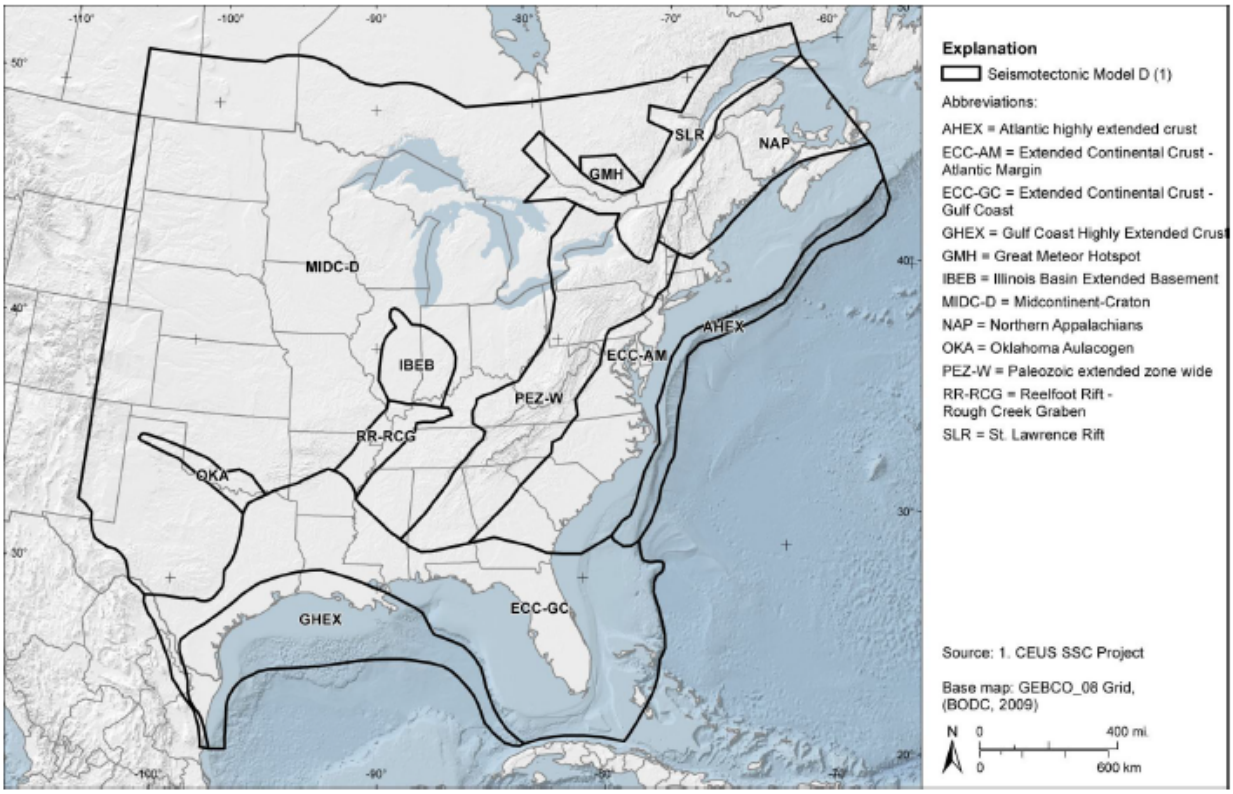


Figure 13–14 Focal depth distributions for the CEUS SSC model.

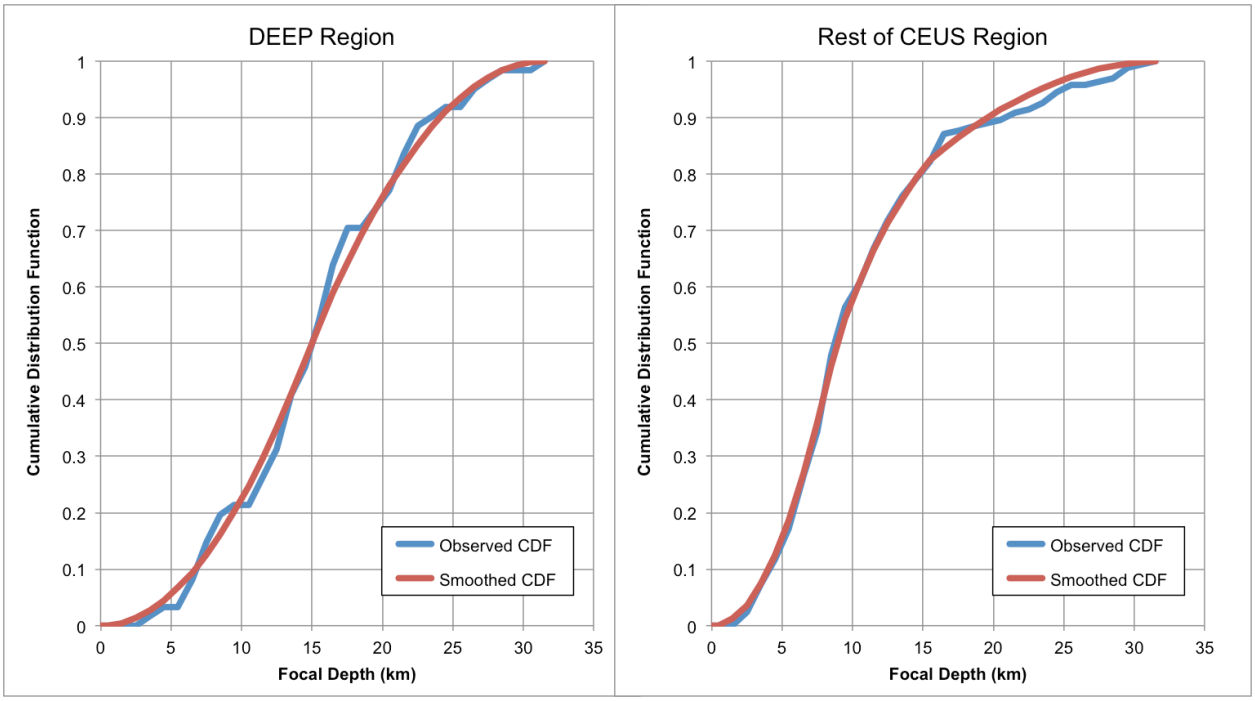


Figure 13–15 Rupture area aspect ratio models developed from the NGA-West2 finite-fault database.

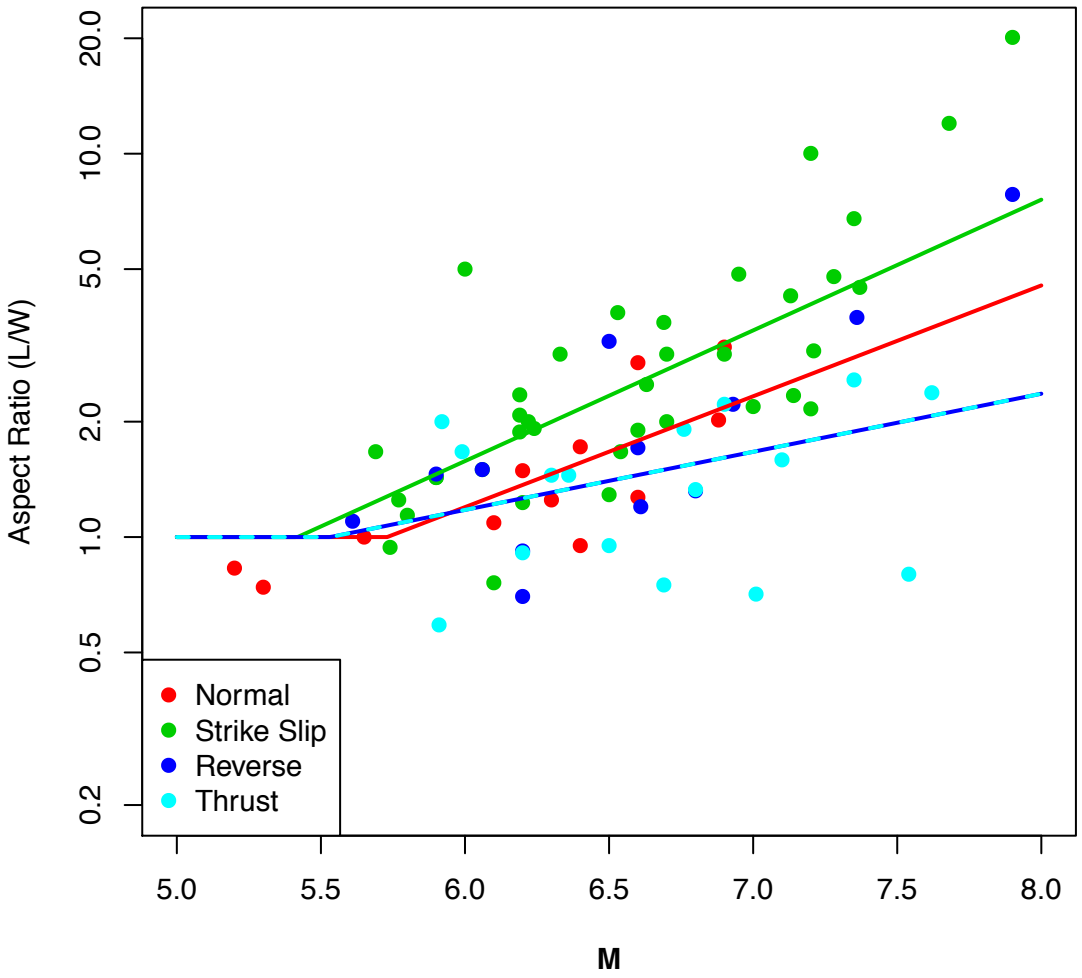


Figure 13–16 HDR models developed from the NGA-West2 finite-fault database.

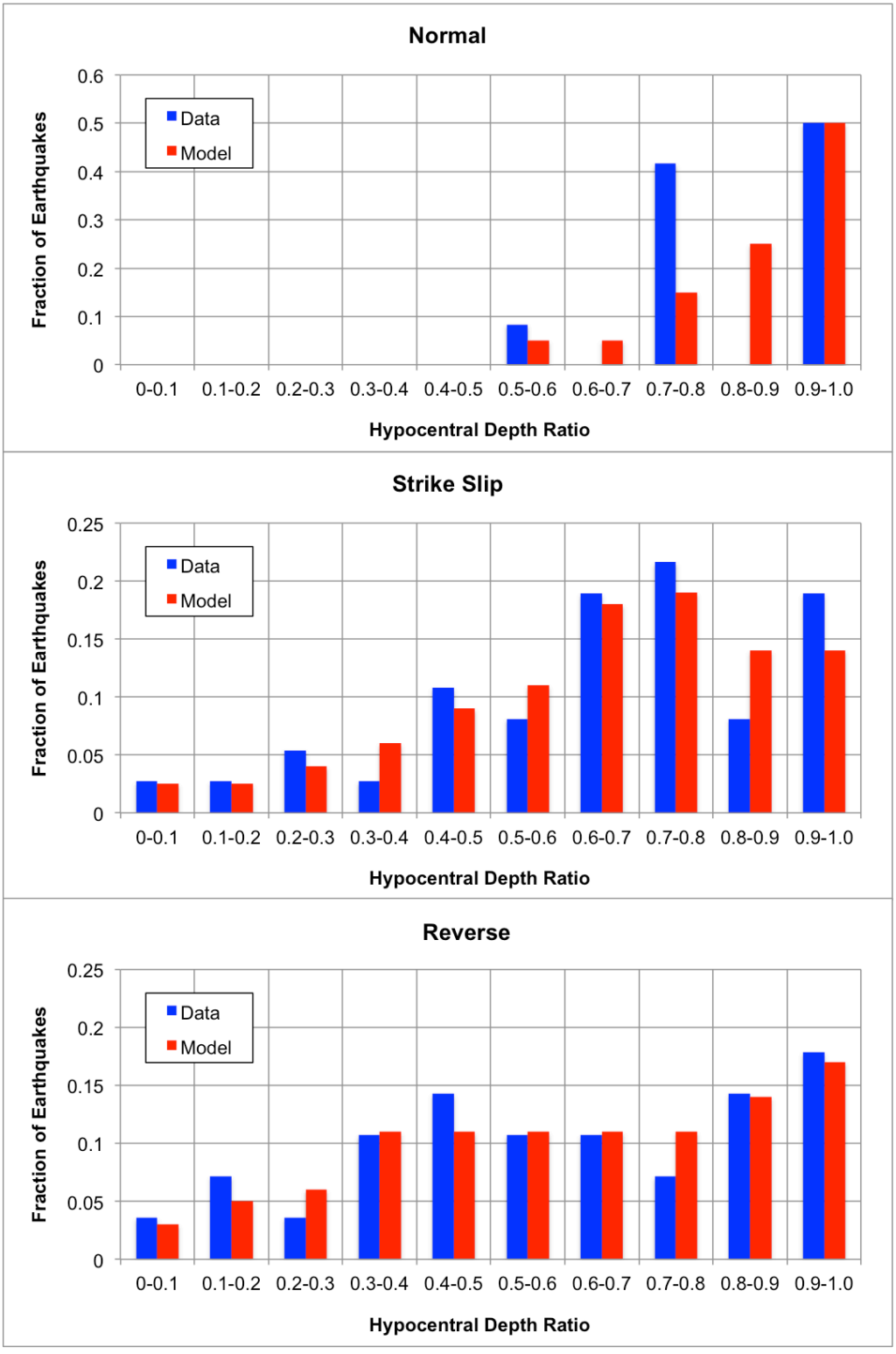
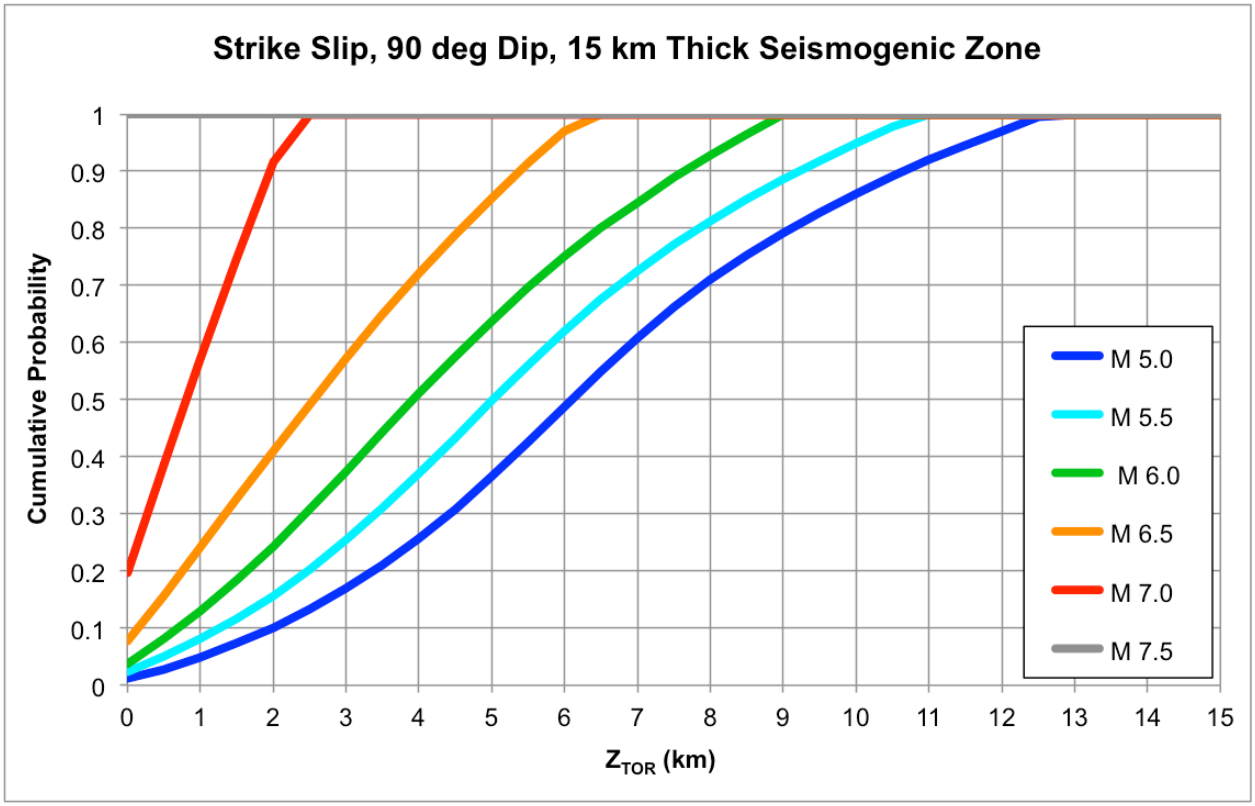


Figure 13–17 Example Z_{TOR} distributions for strike–slip faulting earthquakes.



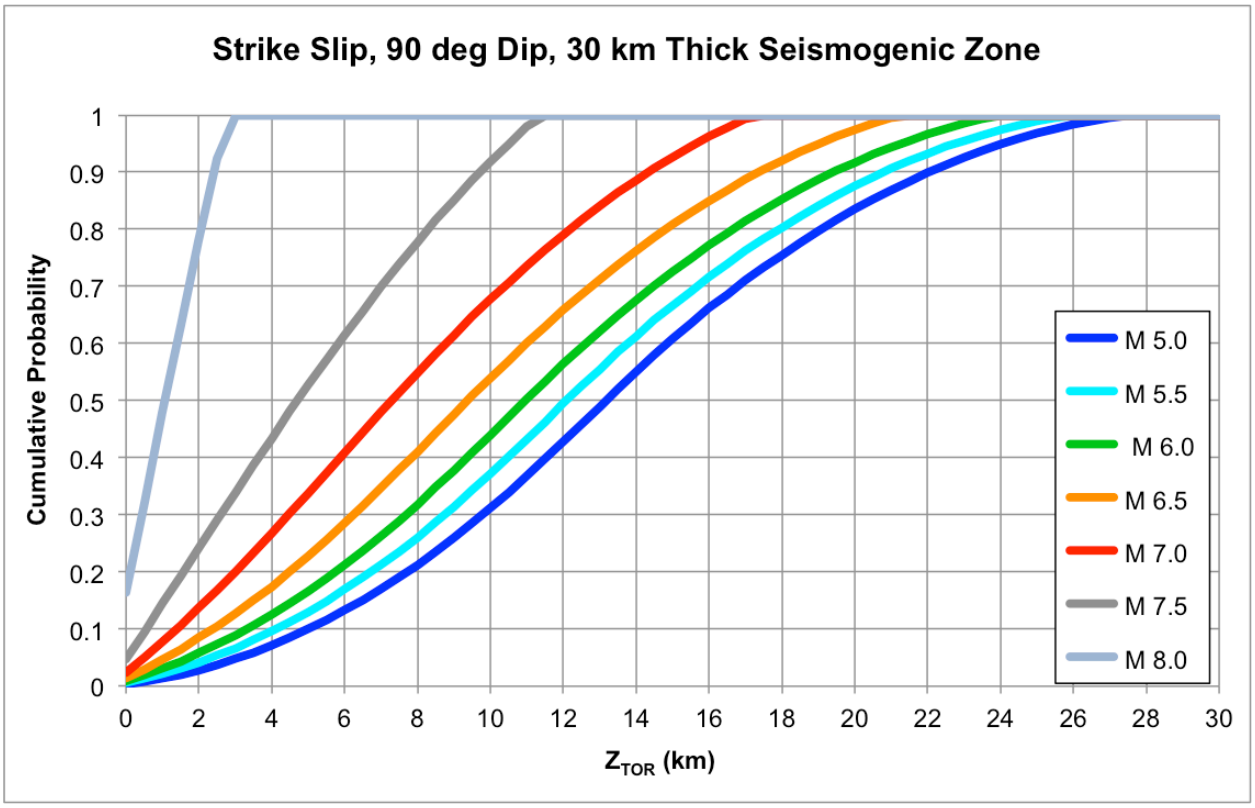
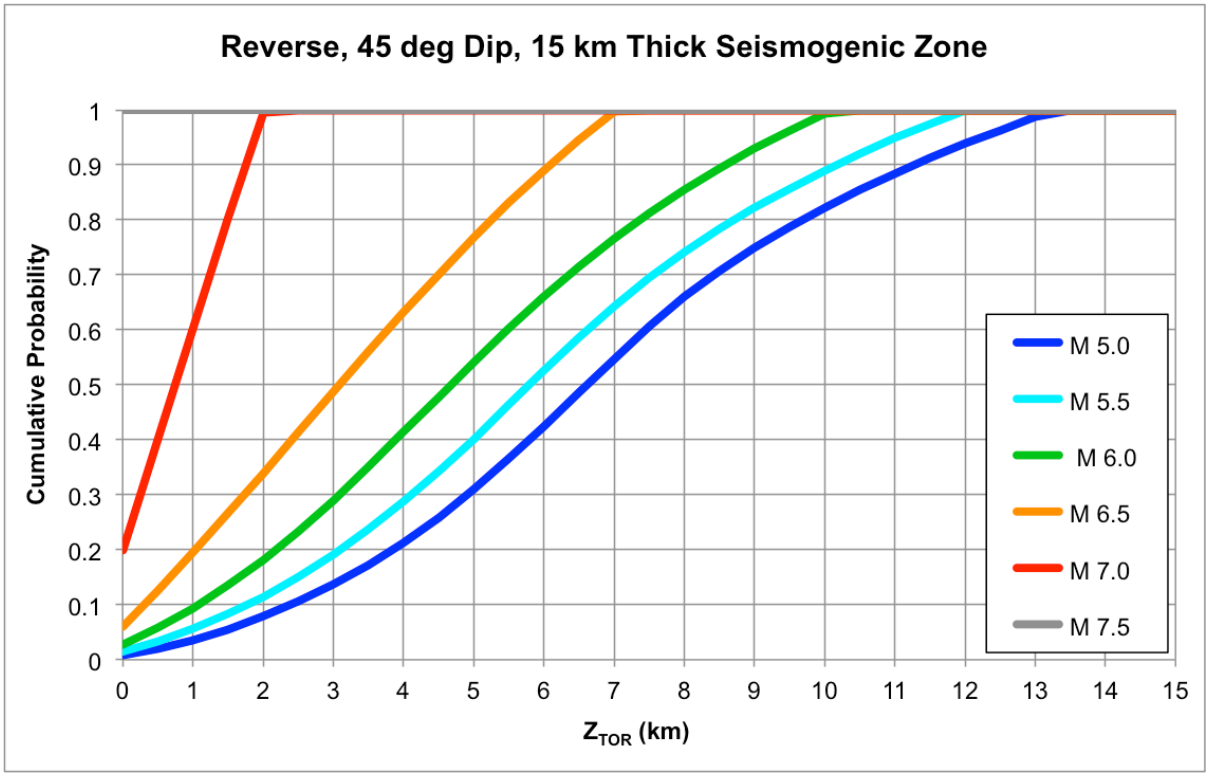


Figure 13–18 **Example Z_{TOR} distributions for reverse faulting earthquakes.**



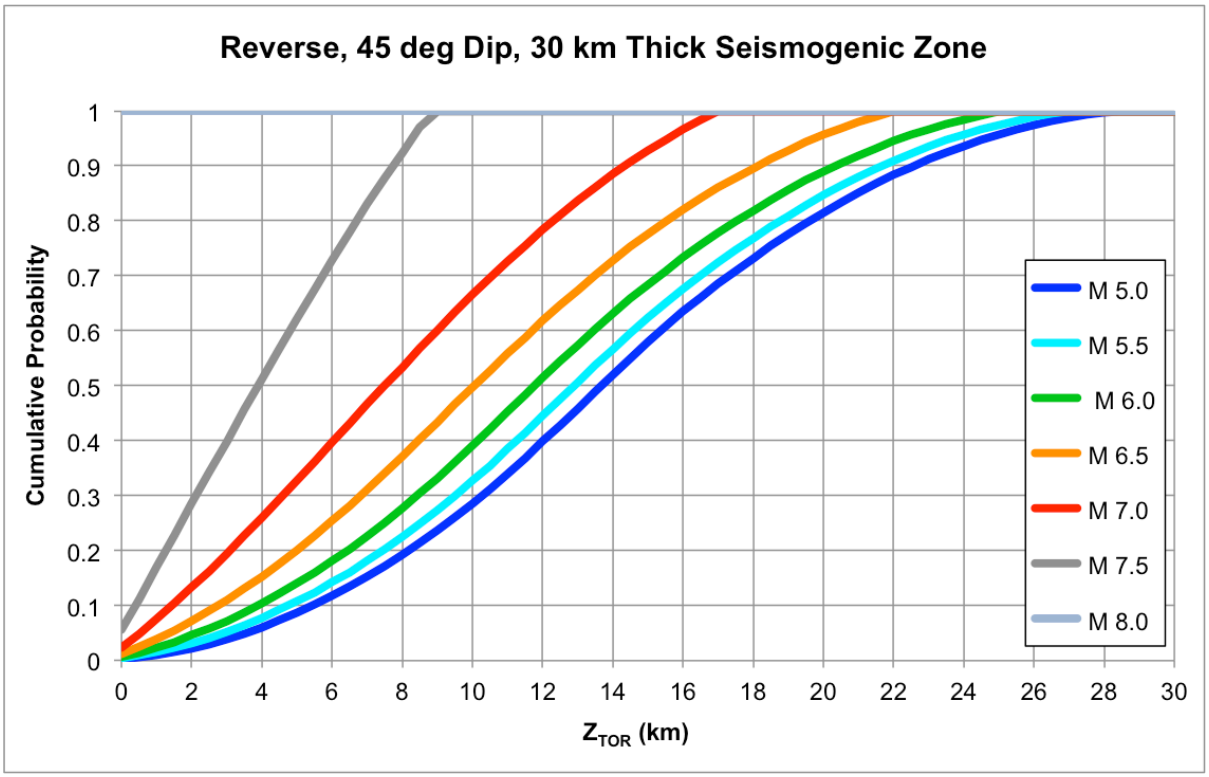
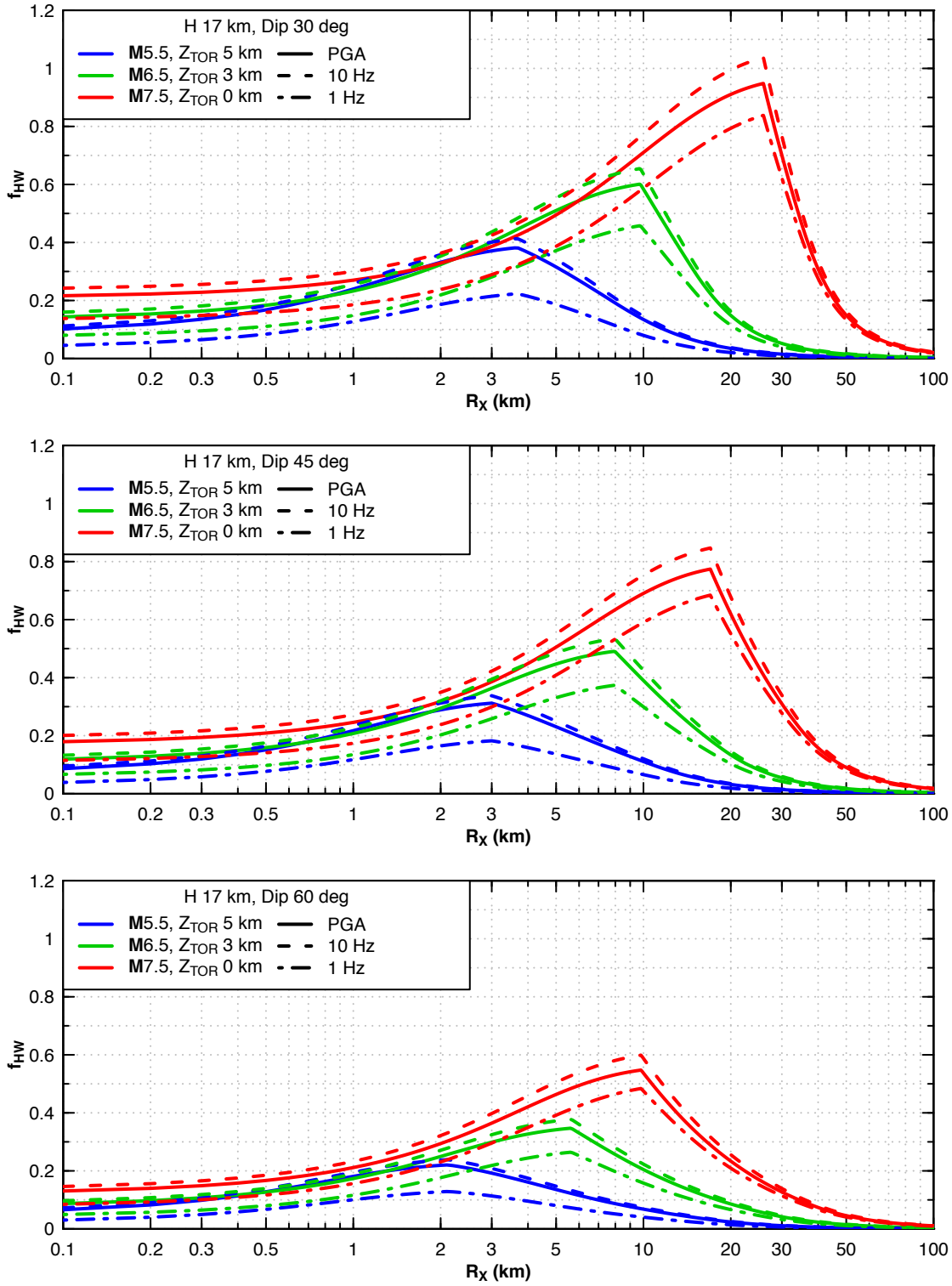


Figure 13–19 Example f_{HW} variations as a function of frequency, magnitude, and fault dip.



14. Full Model Implementation

Chapter 13 presents the development of additional components of the median ground motion model to address the effect of greater attenuation rates within the Gulf Coast region, the effect of source depth, and the inclusion of hanging wall effects for dipping ruptures. Chapter 11 presents the development of the aleatory variability model. This chapter presents the effects of implementation of these components on seismic hazard results and provides guidance on their use in probabilistic seismic hazard analyses (PSHA). In addition, PSHA calculations at the seven demonstration sites using the full NGA-East ground-motion model (GMM) are compared to PSHA results obtained using the full EPRI (2013) GMM. These comparisons provide an indication of the expected impact of the NGA-East GMM on PSHA results in Central and Eastern North America (CENA). Hazard computations are done with the mean representation of the Central and Eastern U.S. (CEUS) seismic source characterization (SSC) model (EPRI/DOE/NRC 2012) such that the uncertainty in the hazard represents only epistemic uncertainty in the GMMs. Finally, documentation on what constitutes the reference-site condition is provided.

14.1 Implementation of the Adjustments to the NGA-East GMM

Chapter 13 presents adjustments to the median GMMs to account for three effects, differences in attenuation in the Gulf Coast Region (GCR), the effect of source depth, and hanging-wall (HW) effects for dipping faults. The effects of these adjustments on seismic hazard calculations are presented below. In addition, the effect of the updated aleatory variability model on hazard calculations is presented.

14.1.1 Adjustment for Gulf Coast Region

Section 13.2 presents the model for GCR adjustments to the median GMMs. The model consists of two alternative adjustment factors presented in Section 13.2.2, and two alternative definitions of the boundary of the GCR, presented in Section 13.2.4. The adjustment models are the PEER model [Equation (13–2)], also referred to as the PEER hybrid model) weighted 0.67 and the DASG model [Equation (13–3)] weighted 0.33. The two definitions of the GCR boundary are shown on Figure 14–1. The larger region is weighted 0.6, and the smaller region is weighted 0.4. Seismic hazard calculations were performed for the two demonstration sites that lie within the GCR—Houston and Jackson (shown in Figure 14–1)—using the full logic tree for median motions developed in Chapter 9, the full logic tree for the aleatory variability model developed in Chapter 11, and the full logic tree for the GCR adjustments described above. The use of the GCR adjustments defined by Equations (13–2) and (13–3) requires calculation of the horizontal path length within the GCR. For the repeated large magnitude event (RLME) sources, this distance was computed by obtaining the point where the straight-line path from the closest point on the source to the site intersects the GCR boundary. For the distributed seismicity sources, this distance was computed by obtaining the point where the straight-line path from the closest point in each ¼ x ¼ degree or ½ x ½ degree cell that intersects the GCR boundary. The hazard computations were performed without including HW effects.

Figures 14–2 through 14–5 show the resulting seismic hazard curves for 1 Hz and 10 Hz motions at the two sites. The solid black curves on each panel show the mean hazard computed from the combined median, aleatory variability, and GCR adjustment models, and the dashed and dashed dot curves show the 5th percentile and 95th percentile hazard, respectively, computed from the combined models. As was the case for the hazard calculations presented in Chapter 12, these results show the distribution in hazard resulting from only the epistemic uncertainty in the GMM characterization. Also shown on Figures 14–2 through 14–5 by the colored curves is the mean hazard conditional on giving full weight to either the PEER or the DASG GCR adjustment models.

Figures 14–6 and 14–7 show the ratio of mean hazard at Houston and Jackson, respectively, for hazard computed using the PEER model to that computed using the DASG model. For distributed seismicity sources, the alternative GCR adjustment models have a noticeable effect on the hazard at Houston only at low ground-motion levels. Once the ground-motion level reaches about 0.01g, the mean distance of earthquakes contributing to the hazard becomes less than about 200 km. As shown on Figure 13–3, the differences between the two GCR adjustment models are small for distances less than 200 km. The effect of the alternative GCR adjustment models on the hazard from the RLME sources at Houston is much more pronounced. This is due to the fact that the primary RLME sources contribution to the hazard at Houston are located in the New Madrid region at a large distance from the site such that the path length within the GCR is generally over 600 km, and at these distances there are large differences in the two GCR adjustment models (Figure 13–3).

At Jackson, the effect of the alternative GCR adjustment models on the hazard from the distributed seismicity sources is slightly less that that shown for Houston. This occurs because there is a significant contribution from the distributed seismicity sources outside of the GCR and the distance to the GCR boundary is small. The effect of the alternative GCR adjustment models on the hazard at Jackson from RLME sources is very small. This is because the path length within the GCR from these sources is generally less than about 100 km.

Figures 14–8 through 14–11 repeat the combined model hazard results shown on Figures 14–2 through 14–5, but these figures now show the sensitivity of the mean hazard for 1 Hz and 10 Hz motions at the two sites to the alternative GCR boundaries shown on Figure 14–1. Again, each panel of the plots shows the mean, 5th percentile, and 95th percentile hazard computed using the full ground-motion characterization (GMC) logic tree. The colored curves show the mean hazard conditional on giving full weight to each of the two alternative GCR boundaries, using the weighted GCR adjustment models.

Figures 14–12 and 14–13 show the ratio of mean hazard at Houston and Jackson, respectively, for hazard computed using the large GCR to that computed using the small GCR. For distributed seismicity sources, the alternative GCR boundaries have a greater effect on the hazard at Houston than the alternative GCR adjustment models as they affect the length of the travel path within the GCR for much of the source zones contributing to the low ground-motion level hazard. Once the ground-motion level reaches about 0.1g, earthquakes that lie within about 100 km from the site contribute nearly all of the hazard from the distributed sources, and the alternative boundaries do not affect the GCR travel path length. The effect of the alternative GCR adjustment models on the hazard from the RLME sources at Houston is, again, large

because the alternative boundaries produce a large difference in the travel path length within the GCR (Figure 13–3).

The effect of the alternative GCR boundaries on the hazard at Jackson from distributed sources is similar to that at Houston in that it is somewhat greater than the effect of the alternative GCR adjustment models for the same reason. The alternative GCR boundaries have essentially no effect on the hazard from the RLME sources because the path lengths within the large GCR are generally less than 100 km and the site lies outside the small GCR. As shown on Figure 13–3, at distances less than about 100 km, the GCR adjustments are essentially unity.

14.1.2 Effect of Source Depth

Section 13–2 presents a model for adjusting the median ground motions as a function of the depth to top of rupture, Z_{TOR}. The model was implemented in seismic hazard calculations for the demonstration sites in the following manner. For each seismic source, the process described in Section 13.2.6 was used to generate a Z_{TOR} distribution using the rupture geometry aleatory distributions defined in Tables 5.4-1 and 5.4-2 of EPRI/DOE/NRC (2012). Distributions were computed for each of the alternative seismogenic crustal thicknesses and then a weighted average Z_{TOR} distribution was produced. This distribution was used in conjunction with the depth effects model to produce adjustments to the median ground motions from earthquakes at each grid point in the distributed seismicity sources and from earthquakes associated with each RLME in HW effects.

Figures 14–14 through 14–23 compare the resulting seismic hazard curves for 1 Hz and 10 Hz motions at the Savannah, Central Illinois, Manchester, Chattanooga, and Topeka demonstration sites to seismic hazard curves computed without applying the depth-effects model. Each plot shows the mean, 5th percentile, and 95th percentile hazard curves based on the full NGA-East GMM. As was done for the comparisons shown in Chapter 12, the seismic hazard calculations were computed using the mean of the SSC model. Figures 14–25 through 14–28 show the ratio of the mean hazard computed without depth effects to that computed with depth effects for the five sites.

The comparisons for 10-Hz hazard indicate that incorporation of the depth effects model produces slightly lower hazard from the distributed seismicity sources and essentially the same hazard from the RLME sources as obtained without consideration of the depth effect. The one exception to this is the hazard from RLME sources at Manchester where inclusion of the depth effect produces slightly higher hazard (Figures 14–18, and 14–26). As shown on the figures, the 1-Hz hazard is essentially the same with and without the depth effect for both distributed seismicity and RLME sources. Again, the 1-Hz hazard from RLME sources at Manchester is slightly higher with the depth effect than without.

The reason for the small impact of the depth-effects model on hazard is illustrated on Figure 14–29. Shown are the depth effects, expressed as $\exp(f_{Z_{TOR}})$, on median ground motions as a function of magnitude for PGA and PSA at frequencies of 0.1, 1, and ≥ 10 Hz. For motions of 10 Hz and higher, the depth-effects model produces on average about a 5 to 7% reduction in motions for magnitudes of **M** 6 or less in the majority of the CEUS and less than about 3% adjustment on average for larger magnitudes. The average adjustments decrease with

decreasing frequency below 10 Hz. Magnitudes of **M** 5 to 6 are large contributors to high-frequency hazard from the distributed sources, but are not a large contributor to low-frequency hazard and do not contribute from the RLME sources. The RLME hazard at Manchester is from the Charlevoix RLME, which is located in the DEEP focal depth distribution zone. In this zone, the larger earthquakes have a deeper Z_{TOR} distribution such that there is an average increase in the median ground motions from the depth-effects model.

Also shown on Figure 14–29 are the ranges in depth effects as a function of magnitude for the three frequencies. These ranges reflect the limits of the effects that occur at the extremes of the Z_{TOR} distributions.

14.1.3 Hanging-Wall Effects

The impact of including HW effects on the computed hazard was tested at four locations, Chattanooga, Manchester, Central Illinois, and Topeka. These four sites were used for the sensitivity analyses as they display a range in nearby seismic activity. Among the demonstration sites, Chattanooga and Manchester have the highest levels of local distributed seismicity and Central Illinois and Topeka have the lowest outside of the GCR. At each site, the mean hazard was computed for PGA and 10 Hz PSA to span the range of frequencies where the largest HW effects occur (Table 13–10). The calculations were performed the full logic tree for ground motions, including the 17 alternative median models and the six alternative aleatory variability models. Table 13–11 lists the assignments of the five alternative HW models to the 17 median ground motion models. Earthquake ruptures were simulated using the rupture geometry aleatory distributions defined in Table 5.4–2 of EPRI/DOE/NRC (2012) and the implementation of the Z_{TOR} distributions as described above in Section 14.1.2. For each simulated rupture, values of R_{RUP}, R_{JB}, and R_X were computed, and Equation (13–11) was used to compute the factor f_{HW} that is added to the median value for the natural log of ground motions. HW effects were added to the hazard computed for the distributed seismicity sources only as the RLME sources are located at distances greater than 100 km from each site.

Figure 14–30 compares the total hazard curves computed with and without HW effects at the four sites. Figure 14–31 shows the ratios of mean AEF computed with HW effects to that computes without HW effects. The top two plots show the hazard ratios plotted against ground motion level and the bottom two plots show the hazard ratios plotted against the mean AEF computed without HW effects. At each of the sites, the hazard computed using HW effects begins to increase above that computed without HW effects at ground motion levels corresponding to AEF of about 10^{-4}. At a mean AEF of 10^{-5}, the increase in AEF is about 5 to 10 percent. For a given level of mean AEF, the increase in hazard is larger at sites with higher levels of local seismicity (e.g. Chattanooga and Manchester) than at sites with lower levels of local seismicity. At very low levels of AEF, the increase in hazard is in the range of a factor of 1.5 to 2.

For frequencies less than 10 Hz, the HW effect will decrease at these sites both because the magnitude of the effect decreases as frequency decrease and because there is an increasing contribution to hazard from the RLME sources, which are located at distances greater than 100 km from the sites such that there is no HW effect on the motions produced by the RLMEs.

14.1.4 Partially Non-Ergodic Aleatory Variability

Chapter 11 developed a partially non-ergodic (single-station) aleatory variability model for use with the median models to compute hazard for the reference site condition. A partially non-ergodic aleatory variability model is recommended for computing reference hard-rock hazard because the current standard of practice is to account for epistemic uncertainty in characterizing the site amplification from reference hard rock to appropriate control point levels [e.g., EPRI (2013) and Rodriguez-Marek et al. (2014)]. The use of single-station sigma avoids double counting of the uncertainty in cases where the ergodic sigma is used and a site-specific response study performed. It should be noted that the EPRI (2013) GMM uses a fully ergodic aleatory variability.

Figure 14–32 compares the EPRI (2013) fully ergodic sigma to the single-station sigma developed in Chapter 13. The central estimate of the single-station sigma is lower than the EPRI (2013) ergodic sigma at all frequencies except 100 Hz (PGA). For frequencies above 1 to 2 Hz, the uncertainty distribution for the single-station model encompasses the ergodic EPRI (2013) values. At frequencies below 1 Hz, the EPRI (2013) ergodic values are above the range of the single-station sigma estimates.

Sensitivity hazard calculations were performed to demonstrate the effect of the two alternative aleatory variability models. Calculations were performed for the Central Illinois and Savannah sites using the EPRI (2013) median models combined with the full logic tree for the single-station sigma model developed in Chapter 11, and combined with the EPRI (2013) full logic tree for aleatory variability. Figures 14–33 through 14–39 compare the hazard at the two sites using the two aleatory variability models. Figures 14–40 and 14–41 show the ratio of the mean hazard computed using the single-station sigma model to that computed using the EPRI (2013) ergodic model. As is expected, use of the lower single-station sigma produces lower hazard in general, especially at low frequencies where the differences are greater. The mean hazard ratios for the RLMEs at high ground-motion levels show an increase because these results are produced by the extreme tails of the ground-motion distributions. The epistemic uncertainty distribution for the single-station sigma model includes larger sigma values than the EPRI (2013) ergodic model, and includes the mixture model that allows for heavier tails in the ground-motion distribution.

14.2 Comparison with EPRI (2013)

Seismic hazard calculations were performed to compare hazard results computed using the full NGA-East GMM to those obtained using the EPRI (2013) GMM. The calculations were performed for the seven CEUS SSC demonstration sites shown on Figure 3–1 and for the seven ground-motion measures of the EPRI (2013) GMM: PSA at frequencies of 0.5, 1, 2.5, 5, 10, and 25 Hz, and PGA. The calculations were performed using the rupture geometry aleatory distributions defined in Tables 5.4-1 and 5.4-2 of EPRI/DOE/NRC (2012) and the implementation of the Z_{TOR} distributions as described above in Section 14.1.2. For the EPRI (2013) model, the same rupture geometry distributions were used with R_{JB} distances calculated for each rupture geometry rather than using the point source adjustment factors given in EPRI (2013). This provides a more direct comparison of the differences in the hazard produced by the two models. The hazard was computed without incorporating the HW model developed in

Section 13.4. The hazard was again computed using the mean representation of the CEUS SSC model such that the uncertainty in the hazard represents only epistemic uncertainty in the GMMs. The hazard was computed using the full epistemic distribution for partially non-ergodic aleatory variability for the NGA-East GMM and the full epistemic distribution for ergodic aleatory variability for the EPRI (2013) GMM. The comparison was performed in this manner because the ergodic EPRI (2013) aleatory variability is currently used to compute reference hard-rock hazard, and it is recommended that the partially non-ergodic aleatory variability model developed in Chapter 11 be used to compute hazard for the reference hard rock condition for the NGA-East GMM.

The seismic hazard comparisons are provided on a set of eight figures for each of the seven demonstration sites. The first seven show comparisons for the seven EPRI (2013) frequencies. Each figure contains four plots with the black curves (indicating results obtained using the NGA-East GMM) and the red curves [indicating results obtained using the EPRI (2013) GMM]. The upper-left plot compares the mean hazard contributed by the distributed seismicity sources and by the RLME sources as well as the total mean hazard. The remaining three plots on the figures compare the center (50^{th} percentile), body (16^{th} and 84^{th} percentiles), and range (5^{th} and 95^{th} percentiles) of hazard results using the two GMMs. The final plot in the series shows the ratio of the mean hazard computed using the two GMMs for the seven spectral periods.

14.2.1 Comparisons at Individual Sites

Figures 14–42 through 14–49 show the hazard comparisons for the Savannah site. At low frequencies, the NGA-East GMM produces higher hazard for both the RLME and distributed seismicity sources, and produces a broader epistemic uncertainty distribution than the EPRI (2013) GMM. At very large ground motions and low AFEs, the hazard results produced by the two GMMs converge. This is likely due to the much larger low-frequency ergodic aleatory variability for the EPRI (2013) GMM compared to the single-station sigma for the NGA-East GMM (Figure 14–32). The differences between the mean hazard produced by the two models become less at higher frequencies, with the NGA-East GMM generally producing larger hazard and a slightly broader range in epistemic uncertainty at most AFEs. At PGA, the EPRI (2013) GMM produces slightly higher hazard for the distributed seismic sources, but lower hazard from the RLME, such that the total is lower than produced by the NGA-East GMM. The differences in hazard shown in these figures are less than shown in Chapter 12, reflecting the effect of the lower single-station sigma used in the NGA-East GMM, as illustrated in Section 14.1.4.

Figures 14–50 through 14–57 show the hazard comparisons for the Central Illinois site. At low frequencies the differences between the hazard produced by the NGA-East and EPRI (2013) GMMs is larger than at Savannah. The RLMEs are at a greater distance from the Central Illinois site than from the Savannah site, indicating that the difference between the NGA-East and EPRI (2013) ground-motion predictions at large magnitudes increases with increasing distance. At higher frequencies, the difference between the mean hazard produced by the two models again decreases, with the two GMMs producing similar mean hazard at frequencies above 5 Hz. As was the case for the Savannah site, the NGA-East GMM generally produces a slightly broader range in epistemic uncertainty at most AFEs. At PGA, the EPRI (2013) GMM produces slightly higher hazard for the distributed seismic sources, but lower hazard from the RLME, such that

the total is lower than produced by the NGA-East GMM. The differences in hazard shown in these figures are less than shown in Chapter 12, reflecting the effect of the lower single-station sigma used in the NGA-East GMM.

Figures 14–58 through 14–65 show the hazard comparisons for the Manchester site. The differences between the hazard produced by the NGA-East and EPRI (2013) GMMs is smaller than at Savannah or Central Illinois. This is likely due to the fact that the RLME contribution is much smaller at this site. At low frequencies, the NGA-East GMM again produces higher hazard, but the difference is small compared to the other sites because the magnitudes contributing to the hazard are lower where the two GMMs have smaller differences than for large, distant earthquakes. Comparing Figure 14–58 for Manchester with Figure 14–50 for Central Illinois indicates that the relative differences in 0.5-Hz mean hazard for the distributed seismicity and for the RLME sources are similar at the two sites, being larger for the RLME sources than for the distributed seismicity sources. At Central Illinois, the RLME hazard dominates, resulting in larger differences in total mean hazard than at Manchester, where the distributed seismicity sources have a larger contribution than the RLME source. The EPRI (2013) GMM produces slightly higher mean hazard than the NGA-East GMM at high frequencies. Again, the breath of the hazard uncertainty distributions is in general greater for the NGA-East GMM that for the EPRI (2013) GMM.

Figures 14–66 through 14–73 show the hazard comparisons for the Chattanooga site. The differences in hazard at low frequencies are similar to those for Central Illinois as the low-frequency hazard at the two sites is dominated by distant RLMEs. The differences in hazard at high frequencies are similar to the other sites, being intermediate between that observed at Central Illinois and that observed at Manchester. Both Manchester and Chattanooga are in areas of higher distributed seismicity locally compared to Central Illinois, but there is a larger RLME contribution to the high-frequency hazard at Chattanooga than at Manchester.

Figures 14–74 through 14–81 show the hazard comparisons for the Topeka site. The differences in hazard at low frequencies are similar to those for Central Illinois and Chattanooga as the low frequency hazard at the three sites is dominated by distant RLMEs. The differences in hazard at high frequencies are again similar to the other sites, being very similar to those observed for the Central Illinois site as both are located in areas of lower local seismicity.

Figures 14–82 through 14–89 show the hazard comparisons for the Houston site. The low-frequency hazard differences are larger at Houston than at Topeka. At both sites, the low-frequency hazard is dominated by the RLME contributions, but at Houston there is the added difference in the differences in the characterization of the GCR between the NGA-East and EPRI (2013) GMMs. As shown on Figure 13–3, there is a larger range in the GCR adjustment factors in the NGA-East GMM compared to the EPRI (2013) GMM, and the NGA-East GMM includes alternative GCR boundaries that affect the path length to which the GCR adjustments are applied. At high frequencies, the hazard differences between the NGA-East and EPRI (2013) GMMs are similar to those at other sites. As the high-frequency hazard is largely contributed by earthquakes occurring at distances less than 200 km, the GCR adjustments are small (Figure 13–3) and have a limited uncertainty range.

Figures 14–90 through 14–97 show the hazard comparisons for the Jackson site. Hazard differences are similar to the other sites. The differences at higher frequencies are similar to those at Savannah because there is a large contribution to the high-frequency hazard from RLME sources. In addition, the Jackson site is potentially located outside of the GCR (weight 0.4), which contributes to higher hazard resulting from the NGA-East GMM from more distant sources than for the EPRI (2013) GMM.

Figure 14–98 presents a comparison of the mean hazard at the seven demonstration sites in terms of the ratio of UHRS for reference rock conditions at AFEs of 10^{-4} and 10^{-5}. For frequencies of 10 Hz and higher, the NGA-East GMM produces ground motions that are about 90–120% of those obtained using the EPRI (2013) GMM. At lower frequencies, the ratio increases to a peak at 1 Hz, where the NGA-East GMM produces ground motions that are approximately 140–210% of those produced by the EPRI (2013) model. The differences in hazard are similar for AFEs of 10^{-4} and 10^{-5}.

14.2.2 Summary of NGA-East and EPRI (2013) GMM Comparisons

The differences in hazard produced by the NGA-East GMM compared to that produced by the EPRI (2013) GMM are attributed primarily to differences in the median models counterbalanced to some extent by differences in the aleatory variability models. As indicated on Figure 14–98, the hazard differences are greatest at low frequencies (≤1 Hz). The differences are also large where the controlling earthquakes are larger and more distant, as seen by comparing the UHRS ratios for Central Illinois and Topeka to those for Savannah. Figures 12–8, 12–9, 12–15(a), and 12–16(a) indicate that the NGA-East median models produce higher low frequency hazard from RLME sources than the EPRI (2013) medians, with the difference increasing with magnitude and distance (e.g., Savannah versus Central Illinois). Thus, the low frequency UHRS ratios are smallest for Manchester because the RLME has a relatively lower contribution to the hazard compared to the other sites. The results on Figures 12–8, 12–9, 12–15(a), and 12–16(a) also indicate that the difference in the hazard produces by the alternative sets of medians is due partly to the differences in the seed models and partly due to the differences in the process used to model epistemic uncertainty. The differences in hazard shown on Figures 14–42 and 14–43 for Savannah and 14–50 and 14–51 for Central Illinois are smaller than those shown on Figures 12–8 and 12–9 for differences just due to different median models, reflecting the impact of using the lower values of single-station sigma compared to the EPRI (2013) ergodic sigma at frequencies ≤1 Hz (see Figure 14–32).

The comparisons shown on Figures 12–12, 12–13, 12–14, 12–5(b), and 12–16(b) for higher frequencies (≥10 Hz) indicate that at low ground-motion levels the median models for the two GMMs produce similar hazard from the distributed seismicity sources; however, the difference increases as the ground-motion level increases. The comparisons also indicate that the differences at larger ground motions are likely due to differences in the seed models, as the hazard results obtained by applying the NGA-East process to the EPRI (2013) seeds (the red curves) are close to those obtained using the EPRI (2013) medians (light blue curves). The comparisons shown on Figures 14–46 through 14–48 for Savannah and Figures 14–54 through 14–56 for Central Illinois indicate somewhat the opposite trend, with the hazard from distributed seismicity sources computed using the EPRI (2013) GMM being slightly higher than that

computed using the NGA-East GMM at high ground-motion levels. This change in behavior is attributed to the use of the lower single-station sigma values in the NGA-East GMM as well as the incorporation of depth effects, which affect the results at high ground-motion levels (e.g., Figures 14–24 and 14–25).

In summary, the NGA-East GMM produces similar high frequency (≥ 5 Hz) mean hazard and higher low-frequency (< 5 Hz) mean hazard compared with the EPRI GMM. At high frequencies, the NGA-East medians produce higher hazard at large ground motions, which would represent larger medians at close distances. The effect of the difference in medians is countered somewhat by the depth-effects model and by the use of single-station sigma for calculation of reference site hazard. The process used by NGA-East leads to slightly broader epistemic uncertainty for higher-frequency motions, but such differences in the mean hazard that occur are attributed more to the differences in the seed models. At low frequencies, the NGA-East medians produce substantially higher mean hazard than the EPRI (2013) medians, with the difference increasing with the magnitude and distance of the contributing earthquakes. These differences are attributed both to differences in the seed models and to differences in the process. The NGA-East process produces significantly broader epistemic uncertainty distributions for low-frequency motions, with the range of the results increasing as the magnitude and distance of the contributing earthquakes increases. At low frequencies, the use of single-station sigma again works to counteract the differences in the median models.

14.3 Reference Site Conditions

As indicated in Section 1.2.3, the target reference site conditions for the NGA-East GMM are a shear-wave velocity Vs of 3000 ± 300 m/sec and a kappa (κ) of 0.006 ± 0.001 sec [see Hashash et al. (2014) and Campbell et al. (2014)]. The assigned uncertainty in Vs for the reference profile top layer Vs was based on producing a ±5% variation in crustal amplification. The uncertainty range for κ is based on the epistemic uncertainty recommended by Campbell et al. (2014) when uncertainty in source and path parameters are considered, which can be argued is the case for the NGA-East median models taking into account the multiple seed models used in their development.

Table 14–1 summarizes the properties used by the developers in the development of the seed GMMs. Figure 14–99 compares the crustal Vs profiles associated with the seed GMMS to the target reference Vs value. Hashash et al. (2014) do not provide an overall crustal profile. Their Figure 6 indicates that the reference velocity is relatively constant to a depth of about 2 km and that depth is used is this comparison plot. Hashash et al. (2014) do indicate a mean velocity gradient for the reference rock of 2 m/s/m. However, it is unlikely that this gradient should be applied over a large depth range, as it would lead to values well outside of those the report.

Hassani and Atkinson (2015), Graizer (2015), Shahjouei and Pezeshk (2015), and Yenier and Atkinson (2015) use the empirical crustal amplification factors developed by Siddiqqi and Atkinson (2002) as reported by Atkinson and Boore (2006). No specific Vs profile is associated with these factors other than that the surface Vs is ≥ 2,000 m/s. Atkinson and Boore (2006) discuss the expected Vs values as a function of depth and their values are used to construct the profile shown by the heavy gray line on Figure 14–99. Hassani and Atkinson (2015) and Yenier

and Atkinson (2015) assume that there the difference in crustal amplification between a surface Vs of 2,000 m/s and 3,000 m/s is small enough to neglect and this range is used to define the gray shaded area on Figure 14–99. Graizer (2015) also indicates that the crustal amplification for hard rock sites is similar to that reported by Atkinson and Boore (2006) and specifies a surface Vs of 2,800 m/s, which is within the gray shaded area on Figure 14–99. Boore and Thompson (2015) develop the hard rock Vs profile shown on Figure 14–99 with a surface Vs of 3,000 m/s. The resulting amplification factors were used in the development of the seed GMMs by Boore (2015), Pezeshk et al. (2015), Shahjouei and Pezeshk (2015), and PEER (2015). Alternative Vs profiles were developed by Darragh et al. (2015), Frankel (2015), and Shahjouei and Pezeshk (2015).

As indicated on Figure 14–99, the resulting range of surface Vs associated with the seed GMMs is broader than the range suggested by Hashash et al. [2014] for CEUS reference hard rock. The effect of this range in Vs profiles on the seed GMMs is illustrated on Figure 14–100 by comparing the resulting crustal amplification functions. The top plot shows the crustal amplification functions, and the bottom plot shows the amplifications multiplied by the kappa operator. For the Darragh et al. and Frankel seed models, the amplification functions were computed using the specified velocity models and the quarter-wavelength method implemented in Boore [2005]. For the Frankel and Shahjouei and Pezeshk seeds, separate amplification functions were computed for low-frequency and high-frequency motions, and then spliced together over the frequency range of 1 to 2 Hz to approximate the approach used to produce the simulated ground motions. For the low-frequency case, the source Vs and density were selected from the second crustal layer to reflect the use of an extended source. Note also that the Graizer model was not used in developing the NGA-East GMM for frequencies above 5 Hz.

The results presented on Figure 14–100 indicate that there is approximately a 25% range in crustal amplification at high frequencies across the seed models. This range contributes in a small part to the epistemic uncertainty in the median models. In light of these differences, the TI team recommends that the NGA-East GMM should be considered applicable to sites whose crustal amplification in combination with kappa fall within the range shown on the lower portion of Figure 14–99. Although this range is broader than suggested by Hashash et al. [2014], it is consistent with how the seed models were developed.

14.4 Application Recommendations

Based on the results presented in Section 14.1, the following guidance in applying the NGA-East GMM is provided.

The hazard ratios on Figures 14–6 and 14–12 show that at sites like Houston, which are located well within the GCR, the alternative GCR adjustment models and GCR boundaries have appreciable impact on low-frequency hazard. This is due to the fact that a major contribution to the low-frequency hazard is from large distant earthquakes where there is a large path length within the GCR. The effect of the alternative models is less at higher frequencies because the contributing earthquakes are typically at closer distances such that the differences between the two GCR adjustment models are smaller, and the GCR path lengths for the majority of contributing earthquakes are similar in length within the two GCR boundaries. The hazard ratios

on Figures 14–7 and 14–13 show that for sites like Jackson located near the GCR boundary the alternative GCR adjustment models and the alternative GCR boundaries have only a small effect on the hazard at both high and low frequencies. This is due to the short path lengths within the GCR for most contributing earthquakes. In summary, the alternative GCR adjustment models and regions are important only for sites located well within the GCR where travel paths within the GCR are much larger than 100 km.

The hazard ratios on Figures 14–24 through 14–28 show that, in general, inclusion of the depth-effects model has a small effect on hazard. The largest effect is for large amplitude, high-frequency motions. These are typically the result of very nearby moderate magnitude earthquakes occurring at shallower depths. For these earthquakes, inclusion of the depth-effects model produces slightly lower hazard as the shallow ruptures are predicted to have lower median motions. Use of the depth-effects model has almost no effect on the hazard from the RLME sources and low-frequency hazard in general because these earthquakes are forecast to occur over a narrow range of Z_{TOR} values. The exception to this is shown on Figure 14–26, where exclusion of the depth-effects model produces lower RLME hazard. The contributing RLME is Charlevoix, which is in a region of thicker seismogenic crust, allowing large earthquakes to occur deeper. In summary, inclusion of the depth-effects model is important primarily for sites where there is a hazard contribution from earthquakes in the Saint Lawrence Rift (SLR) and Great Meteor Hotspot (GMH) distributed seismicity source zones and/or the Charlevoix RLME source.

The hazard comparisons and hazard ratios on Figures 14–30 and 14–31 show that inclusion of the HW effect for distributed seismicity sources becomes important at mean AEF values below 10^{-5}, with the effect increasing as AEF decreases. This occurs because as the hazard increases, it is being contributed to by nearby earthquakes, and, generally, there is an important fraction of those events that are forecast to occur such that the site is located on the hanging wall of a dipping rupture. The 5 to 10 percent increase in the 10^{-5} AEF hazard shown on Figure 14–31 translates into a 2 to 5 percent increase in 10^{-5} UHRS, and a 2 to 4 percent increase in the Ground Motion Response Spectrum (GMRS) computed using the performance-based approach outlined in the US Nuclear Regulatory Commission's Regulatory Guide 1.208. These are considered minimal differences.

In summary, the HW model should be incorporated into calculations for hazard from distributed seismicity sources when there is interest in AEF < 10^{-5}. The HW model should also be used where the site is located in the vicinity of the hanging wall of a dipping RLME source (e.g., the Reelfoot thrust portion of the NMF RLME). Figure 13–19 shows the distance range for significant impact of the HW model.

14.5 Summary and Concluding Remarks

Under the NGA-East program, a comprehensive GMC for CENA was developed. Like other successful NGA programs, the project started as a multidisciplinary multi-researcher science-based project. Under the NGA-East program, numerous experts in the field were brought together, while maintaining their full intellectual independence. The Acknowledgements Section of this report provides evidence of such inclusiveness. Early after the onset of the project, it was

converted to a formal SSHAC Level 3 project. A TI team was formed to perform the evaluation and integration, and a Participatory Peer Review Panel (PPRP) was assembled to support the formal review that is required by the SSHAC process. The project hosted several SSHAC Workshops as well as numerous public meetings on various topics. The Working Groups created at the beginning of the Project (Chapter 2) were integrated into the framework and all of their participants played official roles such as Resources Experts and Proponent Experts at one point or another during the project. Additional experts were convened as necessary throughout the project. The project was successful at fostering additional PPRP interactions by encouraging the PPRP to engage during the workshops with their questions. This was especially beneficial as the project included new technologies and approaches that required increased communications. In addition, the formal communication stream expected in SSHAC Level 3 projects was maintained outside of the workshops with official written communications between the PPRP and the TI team, through the Project Manager. The process was extremely valuable, and the PPRP provided comprehensive and detailed comments and feedback on all aspects of the project, which supported the development of better documentation.

Several innovative technologies were developed and implemented in an attempt to increase the transparency and repeatability of the GMC building process, and to capture the center, body, and range of the technically defensible interpretations (CBR of TDI) as required by SSHAC Level 3 process. The NGA-East process and products include several important features, which are briefly highlighted below. The NGA-East project:

1. Developed a large public comprehensive ground-motion database for CENA. Early versions of the database were shared with the EPRI (2013) project as available, but NGA-East made use of the complete and final dataset, which included a complete flatfile of metadata, the seismograms and tables detailing the earthquake sources and recording station. The database was used in the development of candidate models. The database was also routinely used in other components of NGA-East such as checking the quantification of the epistemic uncertainty, among others.
2. Performed a complete reassessment of regionalization and the development of updated reference site conditions for CENA.
3. Coordinated the development of an unprecedented number of candidate GMMs covering a wide range of approaches. The development was coordinated by the TI team and included extensive interaction among all the model developers throughout the process.
4. Fostered a greater use of extended source (finite-fault models) simulations in the development on candidate GMMs.
5. Developed and implemented a new extensive technical approach to quantify and capture epistemic uncertainty aimed at encouraging transparency and repeatability of results. Through this project, the evaluation and integration focused on capturing the epistemic uncertainty in ground-motion values not in modeling approaches. This involved using and expanding the original candidate GMM set to define and capture the CBR of TDI of median ground motions. Through such an elaborated process, epistemic uncertainty could be quantified more objectively than before, and with a process that is repeatable.

6. Developed a new approach for modeling the aleatory variability, implemented independently from the median GMMs. The development made extensive use of the CENA database but also borrowed data from other parts of the worlds when relevant, and led to an integrated suite of models capturing the epistemic uncertainty in aleatory variability.
7. Developed the aleatory variability model for appropriate use of reference site conditions, favoring the partition of aleatory and epistemic uncertainty in site response through the use of single-station sigma.
8. Included and extensive and continuous review by PPRP members who are all specialized in ground motion characterization. The review process increased the quality and clarity of the outcome of NGA-East.
9. Fostered intellectual deliberations at various levels of interactions among numerous experts and organizations.

Finally, we learned several lessons during the NGA-East Project, most of them at the science level. One very important lesson stands out at the organization level. As indicated earlier, the project started as a multidisciplinary multi-researcher science-based project and before the science was completed, the project was converted to a SSHAC Level 3 project. This approach was costly, both financially and in terms of time. For future projects, we recommend completing the science part of the project, followed by a SSHAC Level 3 process for the evaluation and integration of the models developed in the science phase.

14.6 References

Atkinson, G.M., and D.M. Boore. 2006. Earthquake Ground-Motion Prediction Equations for Eastern North America." Bull. Seismol. Soc. Am. 96, 2181–2205. (also see the erratum published in Vol. 97, No. 3, p. 1032).

Boore, D.M. 2005. "SMSIM – Fortran Programs for Simulating Ground Motions from Earthquakes: Version 2.3 – A Revision of OFR 96-80-A." U.S. Geological Survey Open-File Report, 00-509, revised 15 August 2005, 55 pgs.

Boore, D.M. 2015. "Adjusting Ground-Motion Intensity Measures to a Reference Site for which VS30 3000 m/s." PEER Report No. 2015/06, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Boore, D.M., and E.M. Thompson. 2015. "Revisions to Some Parameters used in Stochastic-Method Simulations of Ground Motion." Bull. Seismol. Soc. Am., 105: 1029–1041.

Campbell, K.W., Y.M.A. Hashash B. Kim, A.R. Kottke, E.M. Rathje, W.J. Silva, and J.P. Stewart. 2014. "Reference-Rock Site Conditions for Central and Eastern North America: Part II – Attenuation Definition." PEER Report No. 2014/12, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

EPRI. 2013. "Ground-Motion Model (GMM) Review Project." Electrical Power Research Institute, EPRI Report 3002000717, Palo Alto, CA.

EPRI. 2013b. “Seismic Evaluation Guidance: Screening, Prioritization and Implementation Details (SPID) for the Resolution of Fukushima Near-Term Task Force Recommendation 2.1: Seismic.” Technical Report 1025287, Electric Power Research Institute, Palo Alto, California, February.

EPRI/USDOE/USNRC. 2012. “Central and Eastern United States Seismic Source Characterization for Nuclear Facilities.” U.S. Nuclear Regulatory Commission Report, NUREG-2115; EPRI Report 1021097, 6 Volumes; DOE Report# DOE/NE-0140, Washington, D.C.

Frankel, A., C. Mueller, T. Barnhard, D. Perkins, E.V. Leyendecker, N. Dickman, S. Hanson, M. Hopper. 1996. “National Seismic Hazard Maps—Documentation June 1996.” U.S. Geological Survey Open-File Report 96-532, Menlo Park, CA, 110 pgs.

Hashash, Y.M.A., A.R. Kottke, J.P. Stewart, K.W. Campbell, B. Kim, E.M. Rathje, and W.J. Silva. 2014. “Reference Rock Site Condition for Central and Eastern North America.” Bull. Seismol. Soc. Am., 104: 684–701.

PEER. 2015. “NGA-East: Median Ground-Motion Models for the Central and Eastern North America Region.” PEER Report No. 2015/04, Pacific Earthquake Engineering Research Center, University of California, Berkeley, CA.

Rodriguez-Marek, A., E.M. Rathje, J.J. Bommer, F. Scherbaum, and P. Stafford. 2014. “Application of Single-Station Sigma and Site-Response Characterization in a Probabilistic Seismic-Hazard Analysis for a New Nuclear Site.” Bull. Seismol. Soc. Am., 104: 1601–1619.

Siddiqqi, J., and G.M. Atkinson. 2002. “Ground Motion Amplification at Rock Sites across Canada, as Determined from the Horizontal-To-Vertical Component Ratio.” Bull. Seismol. Soc. Am., 92: 877–884.

Table 14–1 Crustal properties used in developing seed GMPEs.

Seed model set	Crustal amplification	Kappa (s)	Source Vs, ρ (km/sec, g/cc)
Boore	Specified amplification function, Table 2.3 of PEER (2015), surface Vs 3 km/sec (Boore and Thompson 2015)	0.006	3.7, 2.8
Darragh et al.	Specified crustal Vs profile, Table 3.2 of PEER (2015), surface Vs 3 km/sec	0.006	3.52, 2.71
Yenier and Atkinson	Atkinson and Boore (2006) referencing Siddiqqi and Atkinson (2002), surface Vs ≥ 2 km/sec	0.005	3.7, 2.8
Pezeshk et al.	Specified amplification function, Table 5.1 of PEER (2015), surface Vs 3 km/sec	0.006	3.7, 2.8
Frankel	Specified crustal model for low frequencies (LF), surface Vs 2.6 km/sec Specified amplification for high frequencies (HF) Frankel et al. (1996), surface Vs 2.8 km/sec	0.0033 LF 0.006 HF	3.4, 2.5 LF 3.6, 2.8 HF
Shahjouei and Pezeshk	Specified crustal model for LF, surface Vs 2.83 km/sec Two alternatives for HF: Atkinson and Boore (2006) and Boore and Thompson (2015)	0.005 HF	3.7, 2.8 HF
Graizer	Atkinson and Boore (2006), surface Vs specified as 2.8 km/sec	0.006	3.7, 2.8
Hassani and Atkinson	Atkinson and Boore (2006), surface Vs ≥ 2 km/sec	0.005	3.7, 2.8
PEER	Amplification specified as Boore (2015), equivalent to Boore and Thompson (2015), surface Vs 3 km/sec	0.006	3.7, 2.8

Figure 14–1 Alternative Gulf Coast Regions and Location of Houston and Jackson Demonstration Sites.

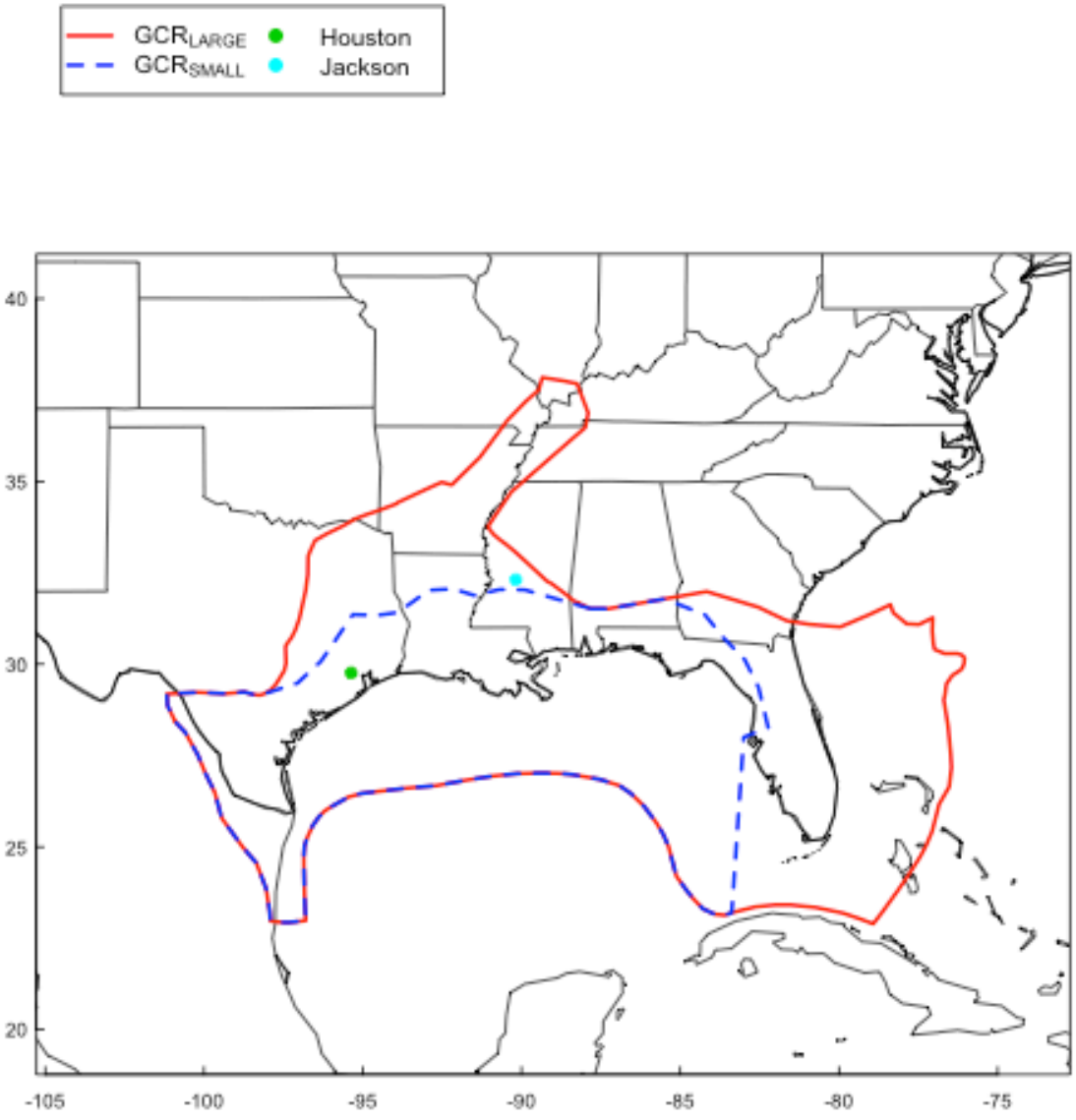


Figure 14–2 Sensitivity of 1 Hz hazard at Houston to alternative GCR adjustment models.

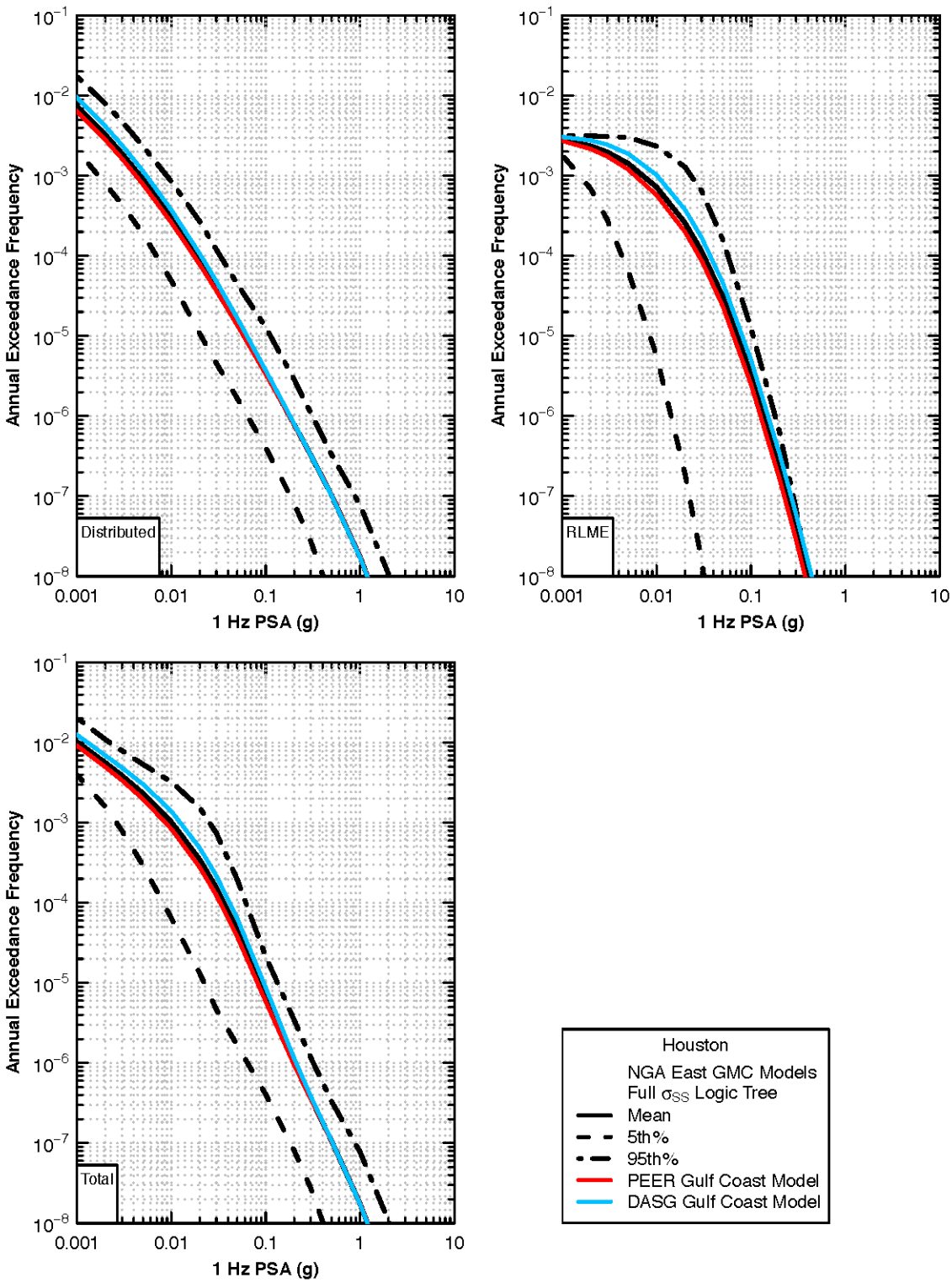


Figure 14–3 Sensitivity of 10 Hz hazard at Houston to alternative GCR adjustment models.

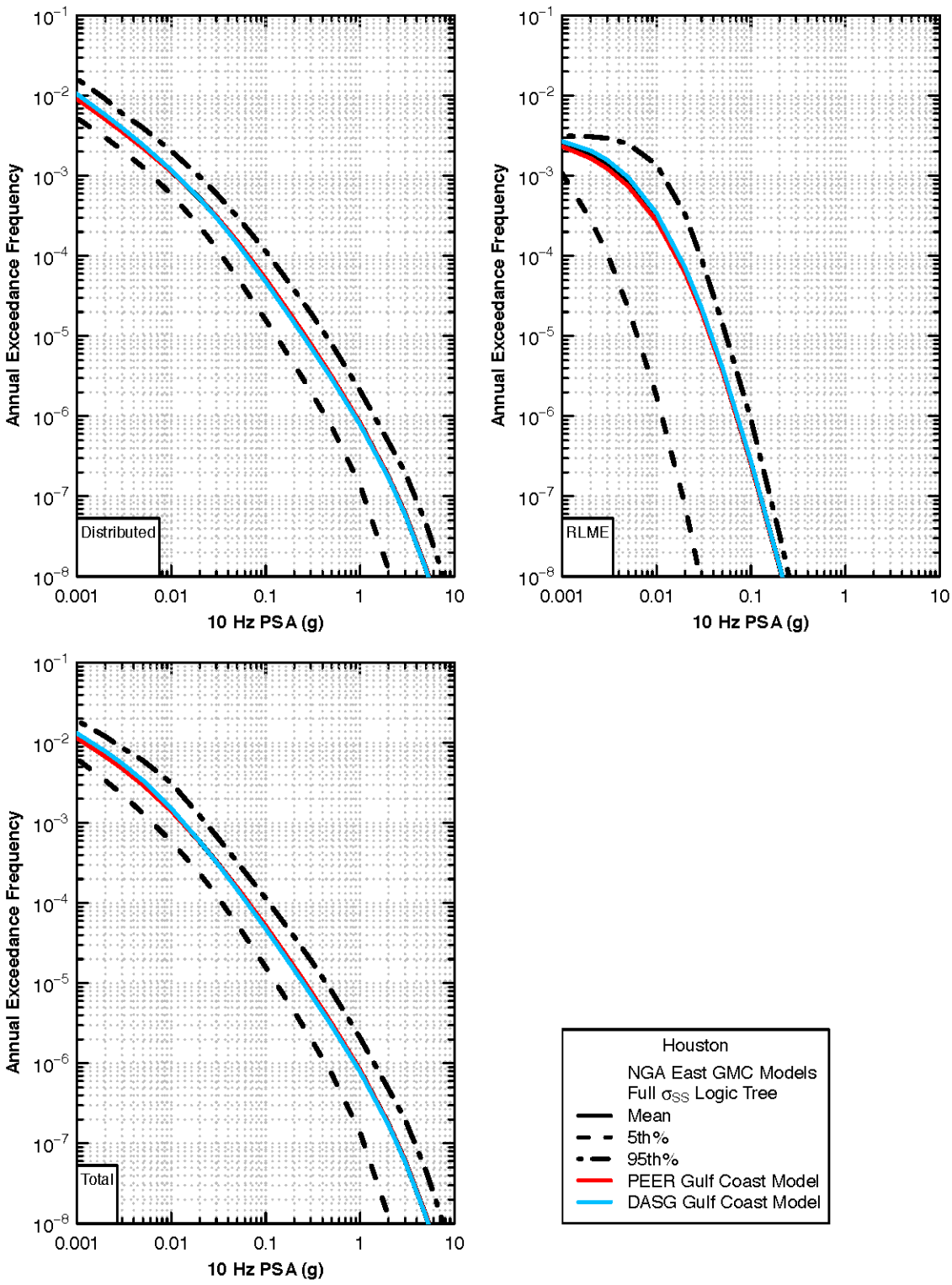


Figure 14–4 Sensitivity of 1 Hz hazard at Jackson to alternative GCR adjustment models.

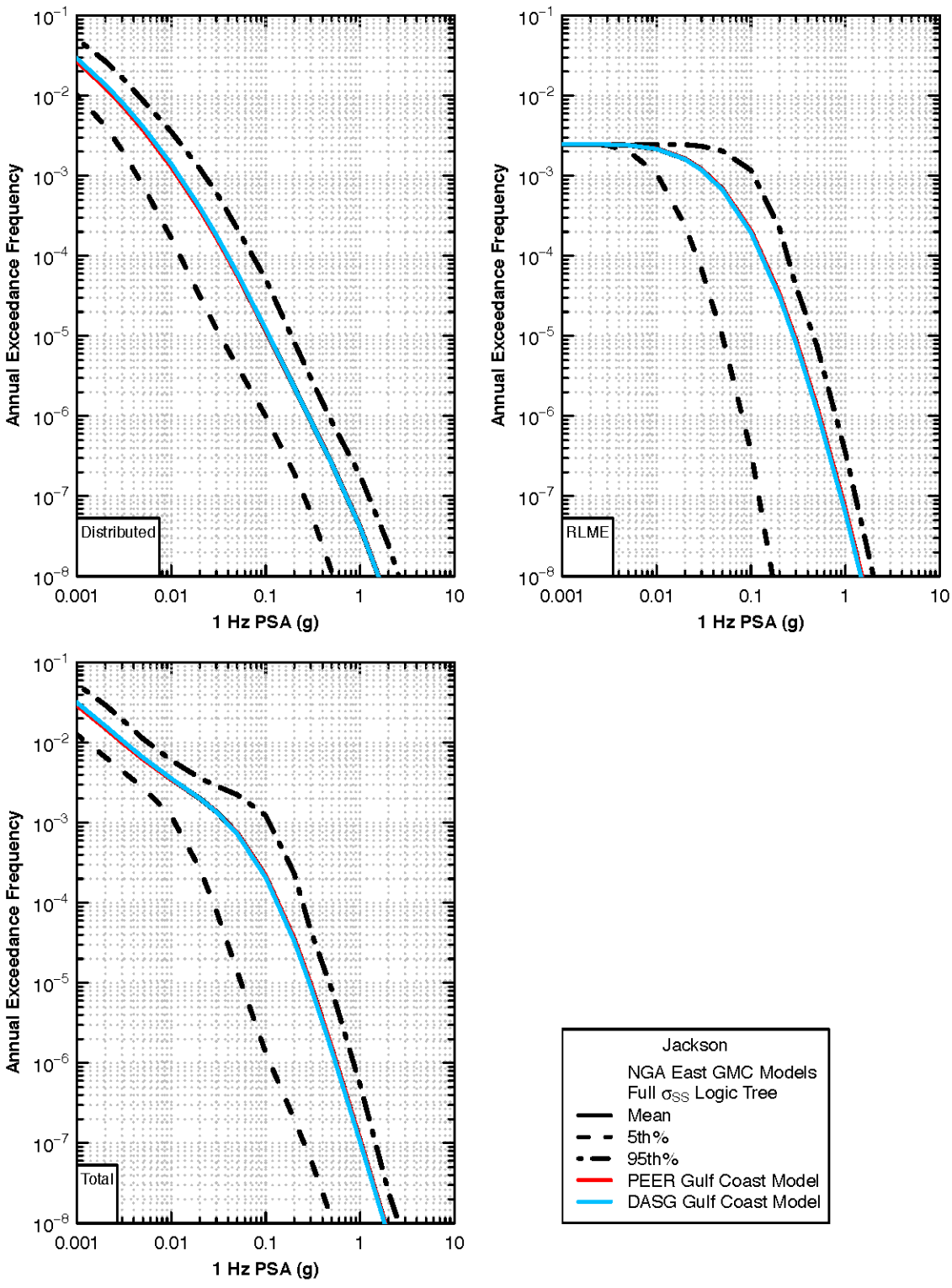


Figure 14–5 Sensitivity of 10 Hz hazard at Jackson to alternative GCR adjustment models.

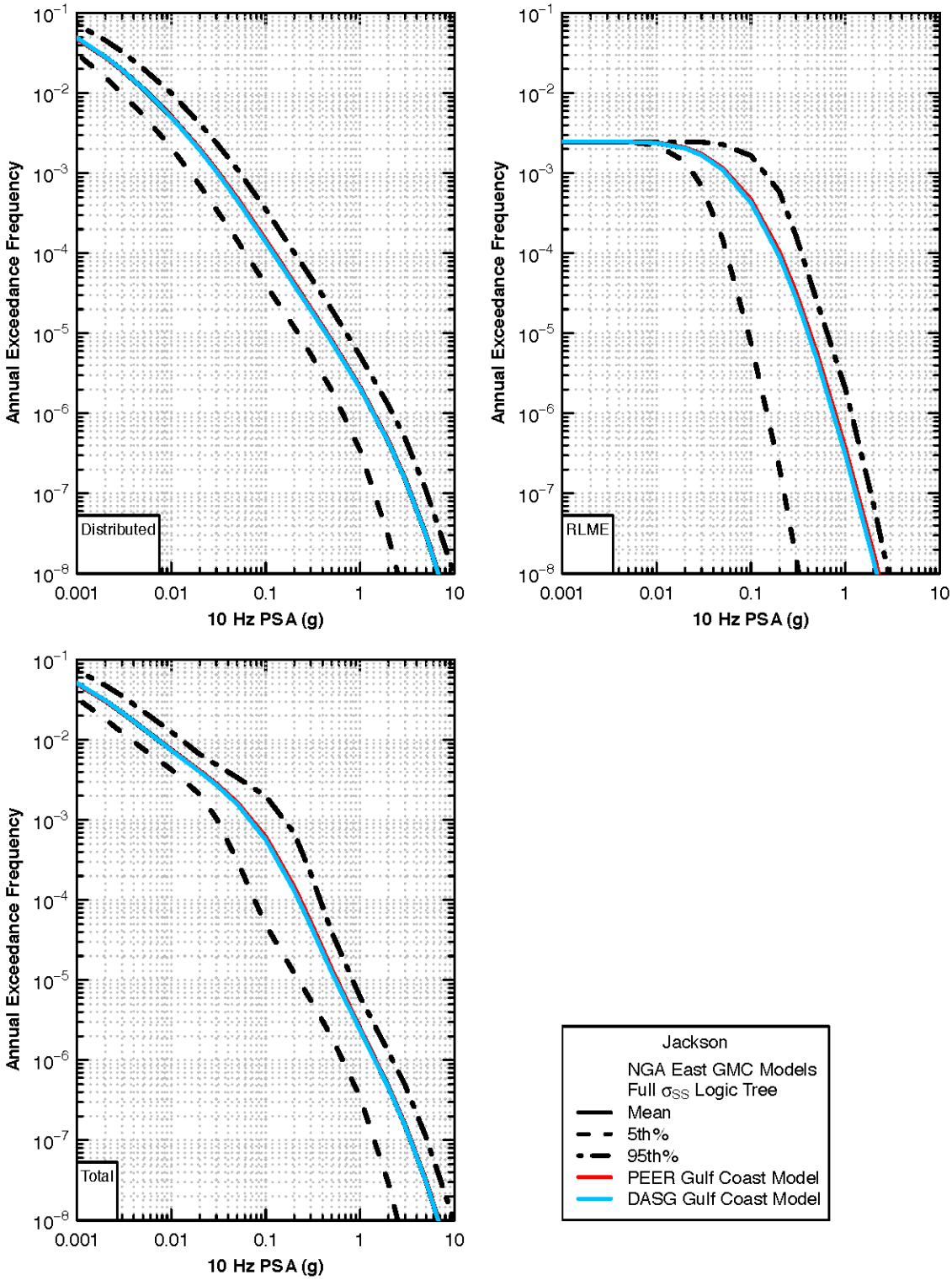


Figure 14–6 Ratio of mean hazard at Houston to alternative GCR adjustment models.

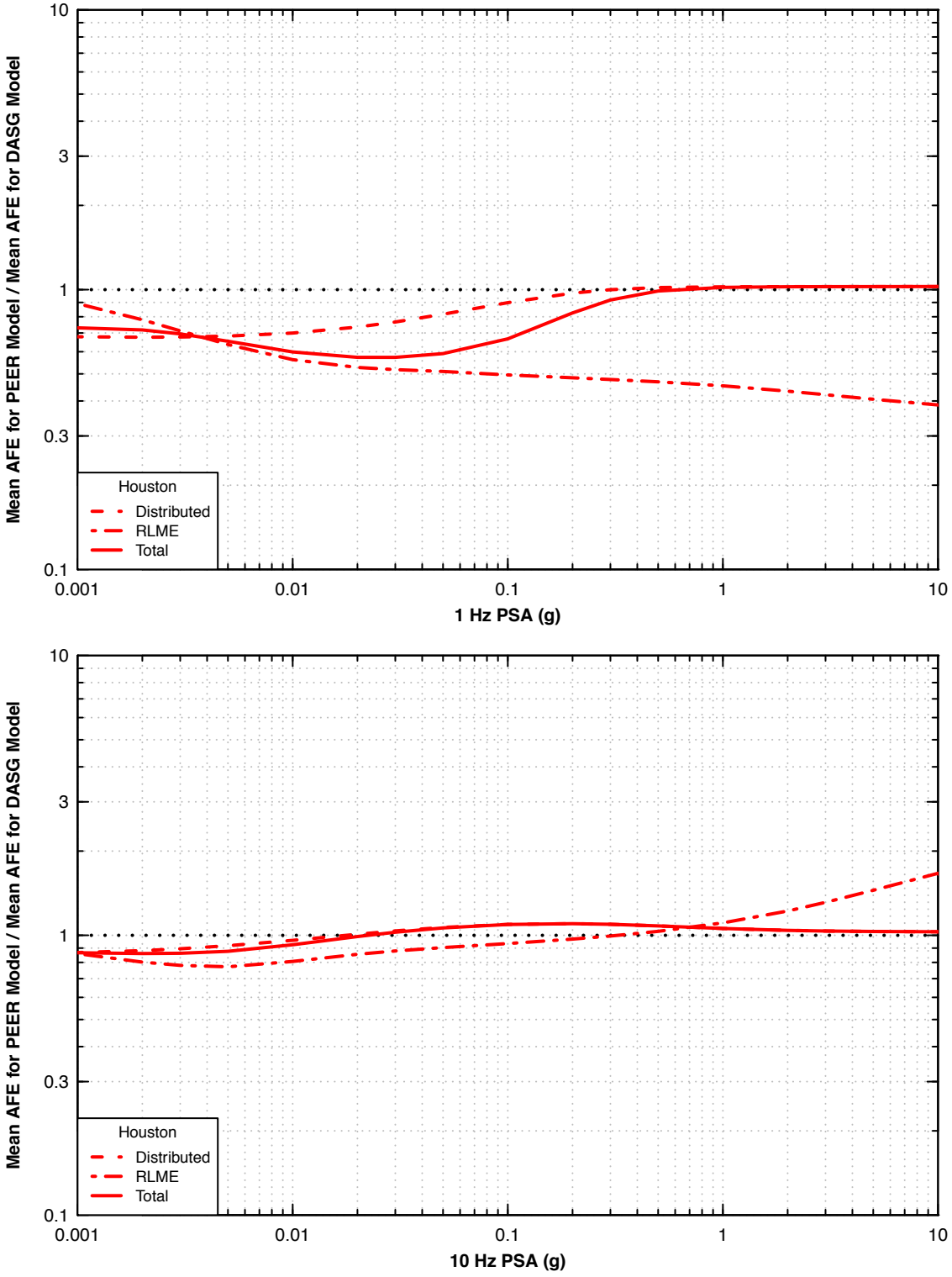


Figure 14–7 Ratio of mean hazard at Jackson to alternative GCR adjustment models.

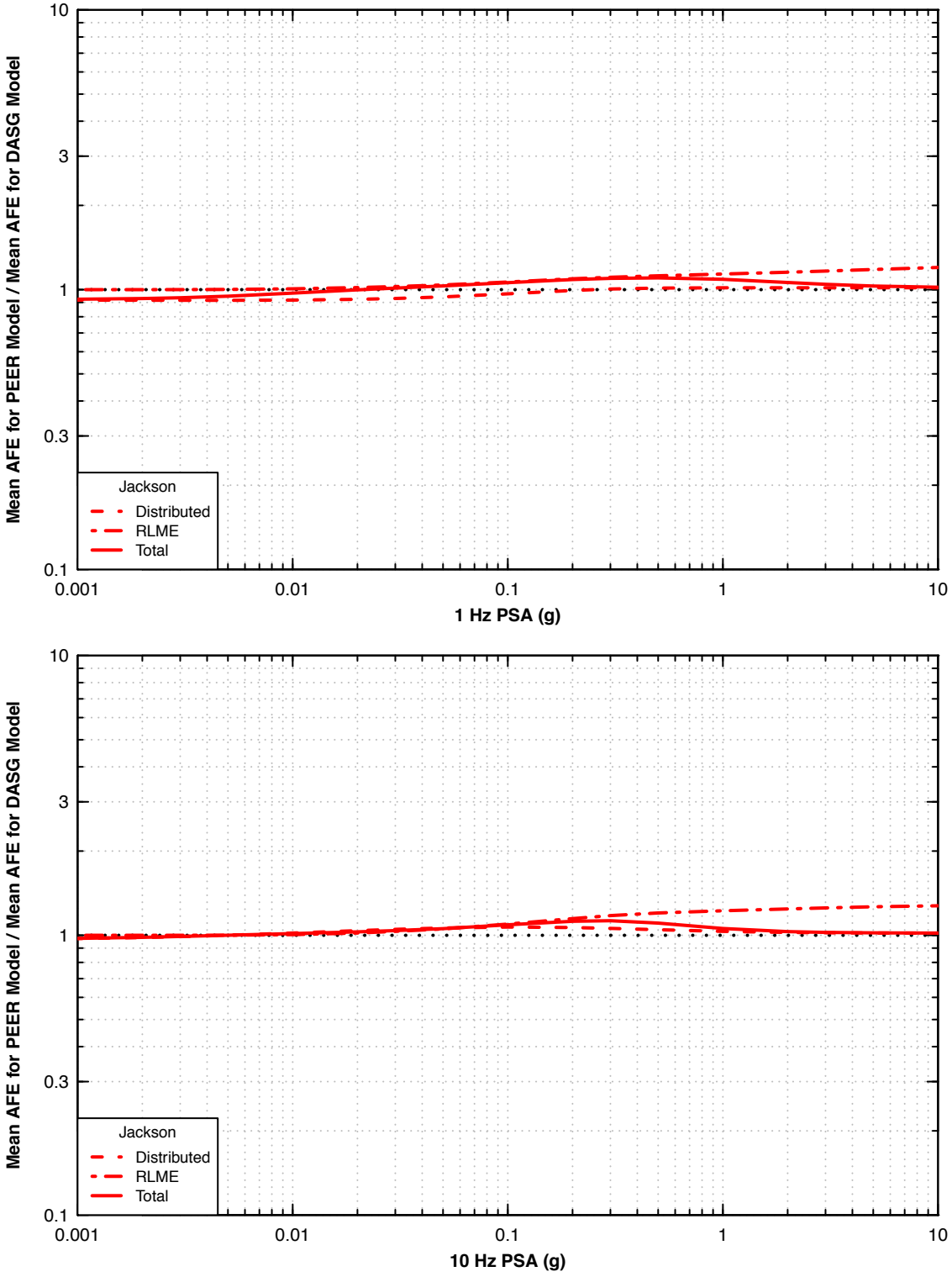


Figure 14–8 Sensitivity of 1 Hz hazard at Houston to alternative GCR boundaries.

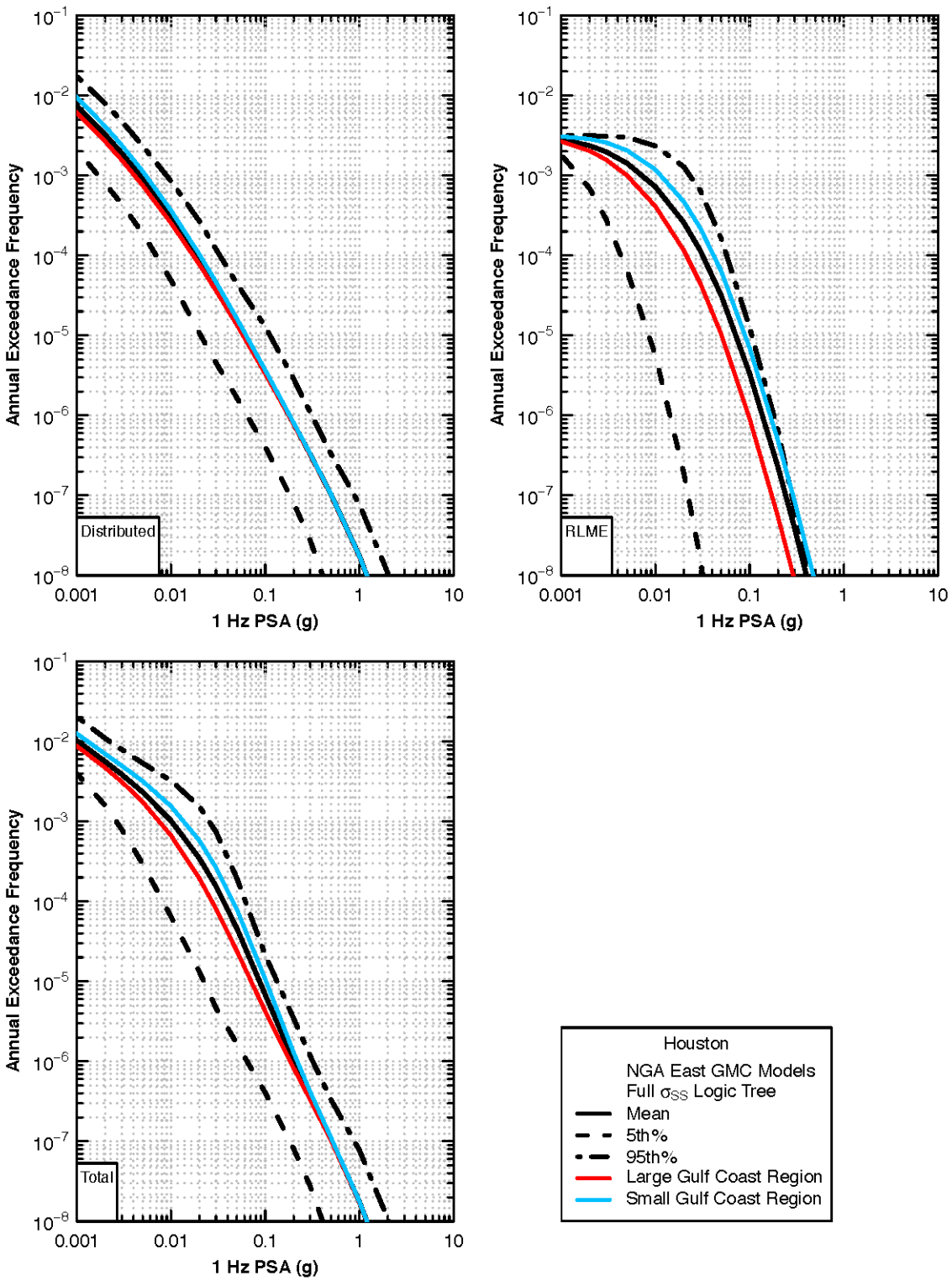


Figure 14–9 Sensitivity of 10 Hz hazard at Houston to alternative GCR boundaries.

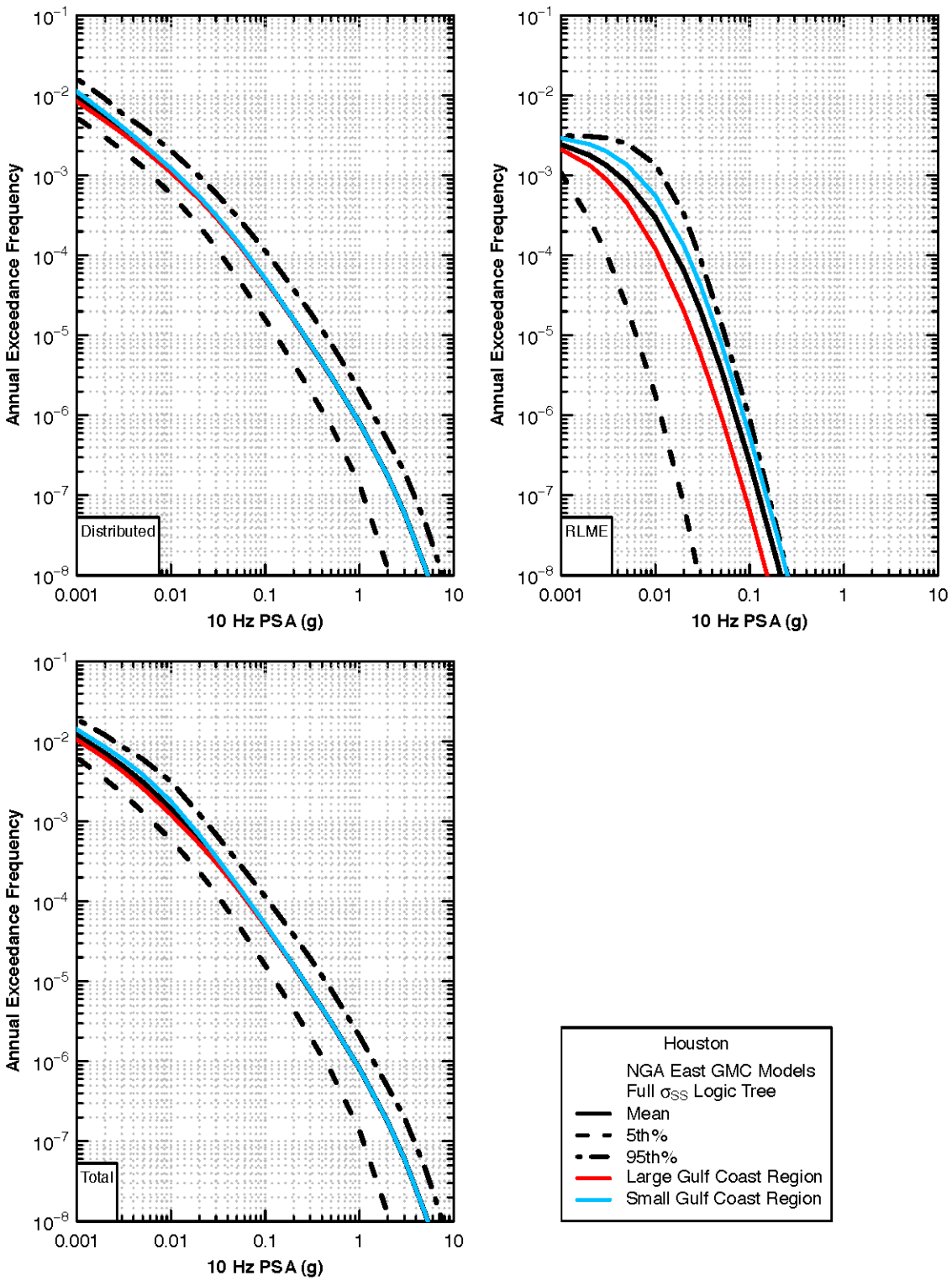


Figure 14–10 Sensitivity of 1 Hz hazard at Jackson to alternative GCR boundaries.

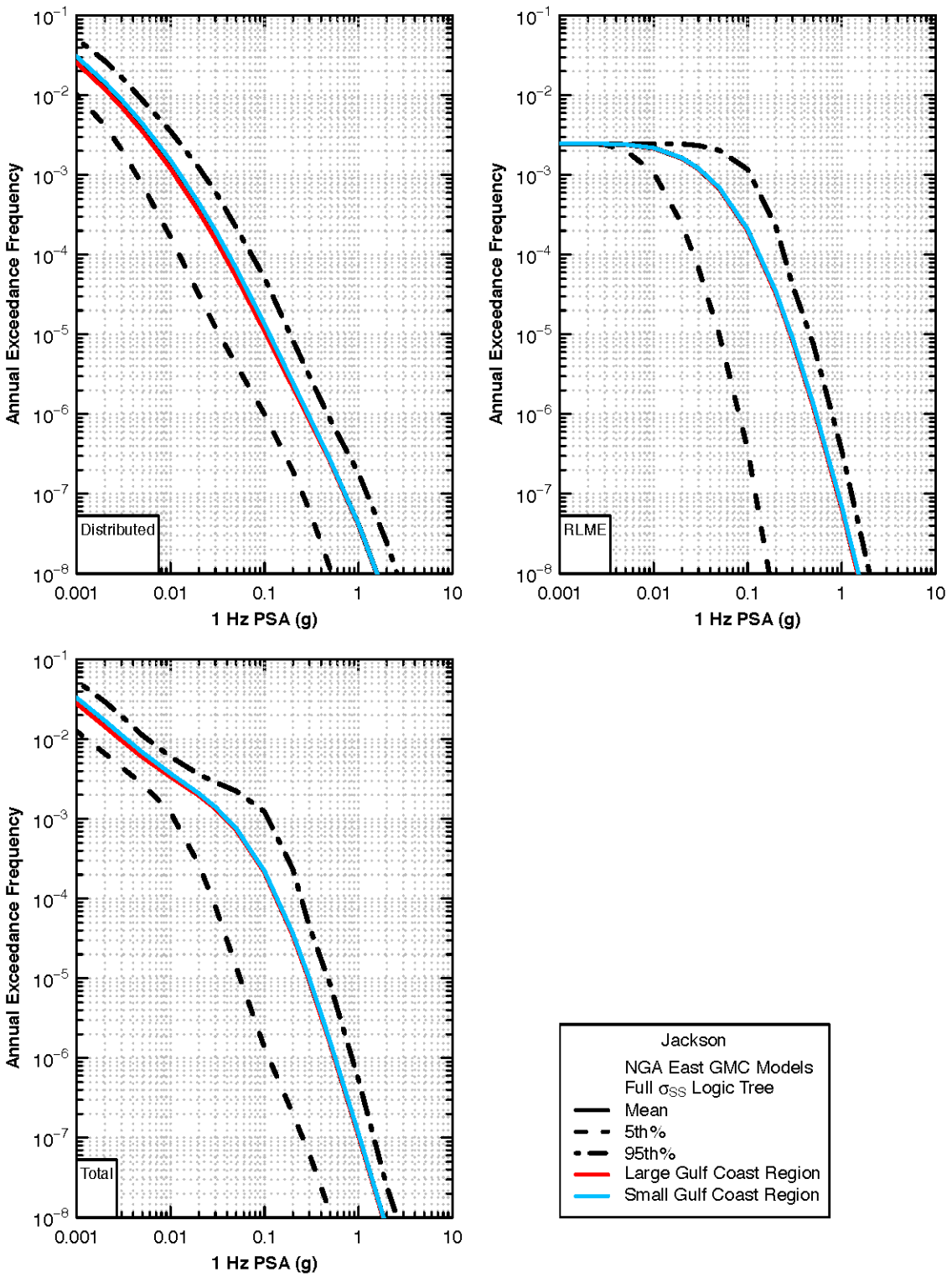


Figure 14–11 Sensitivity of 10 Hz hazard at Jackson to alternative GCR boundaries.

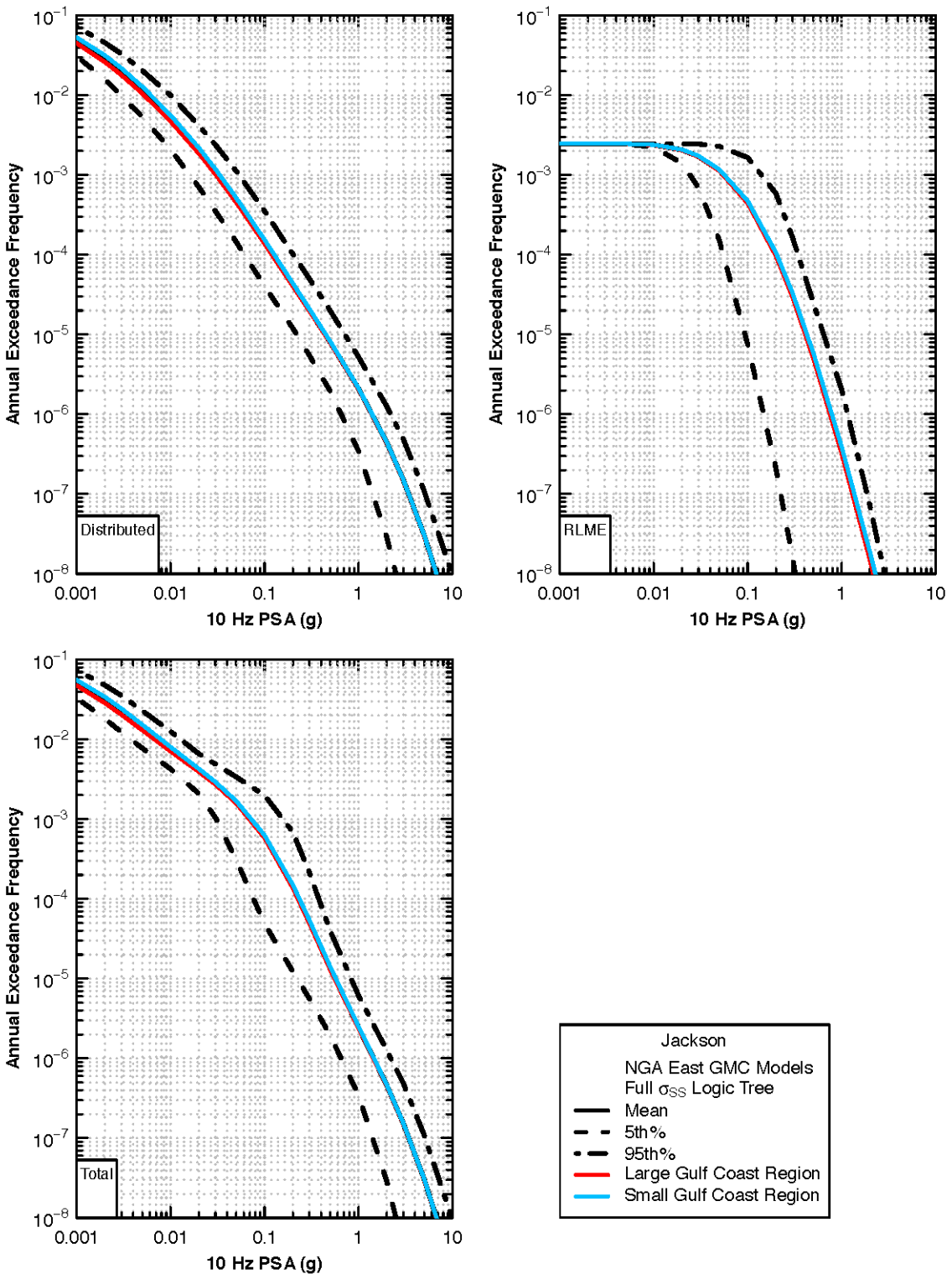


Figure 14–12 Ratio of mean hazard at Houston to alternative GCR boundaries.

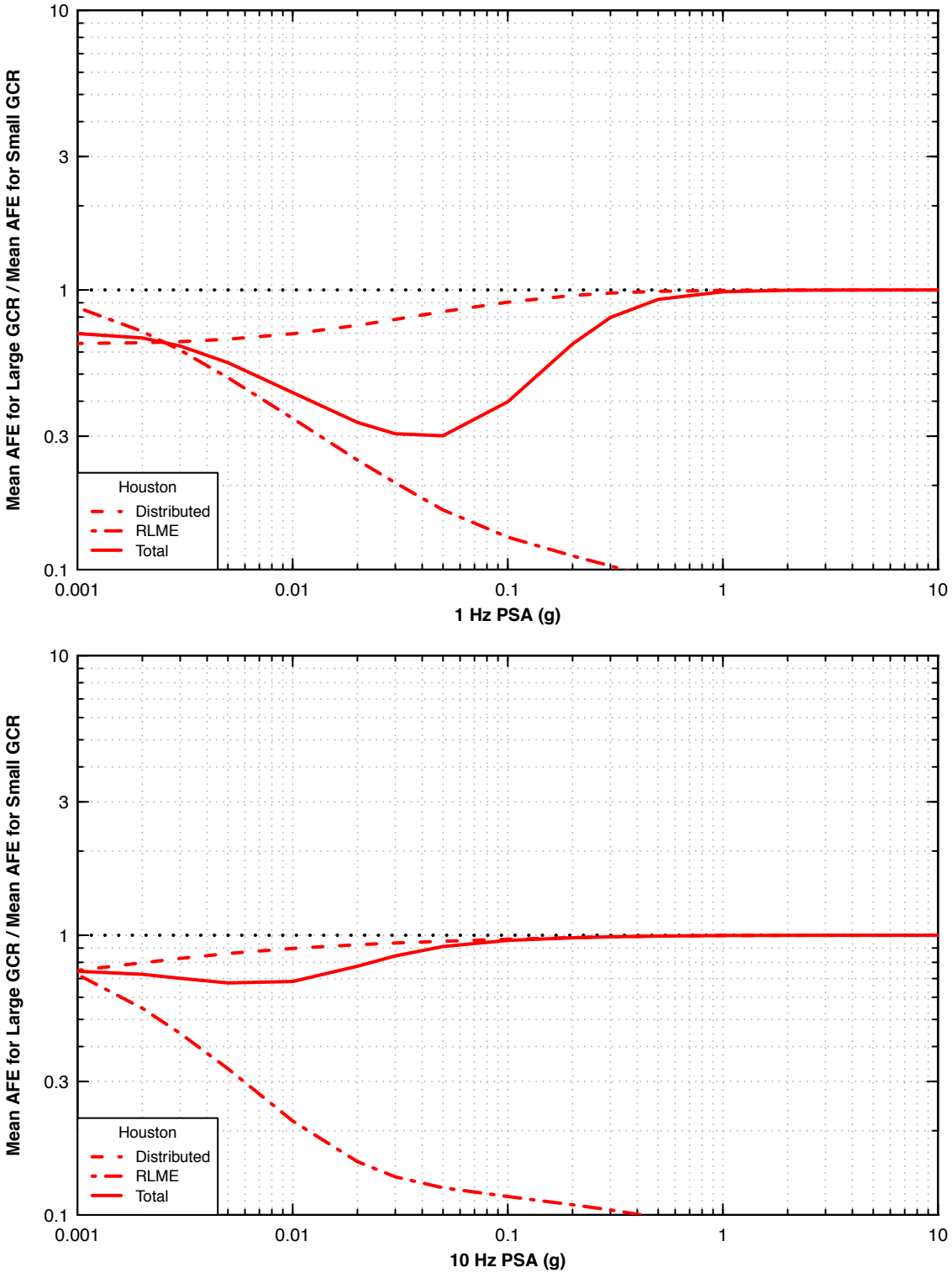


Figure 14–13 Ratio of mean hazard at Jackson to alternative GCR boundaries.

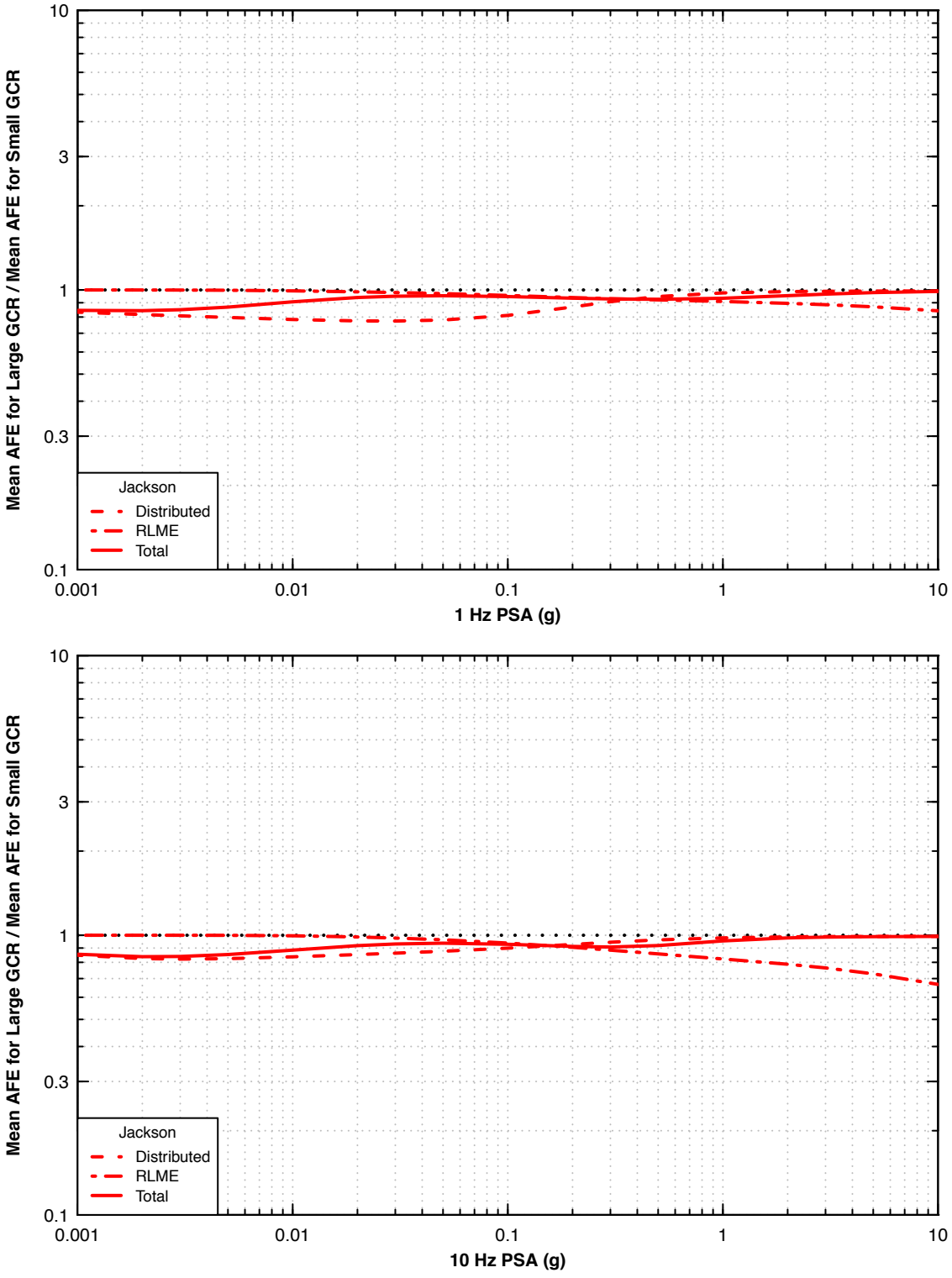


Figure 14–14 Comparison of 1 Hz hazard at Savannah computed with and without depth-effects adjustments.

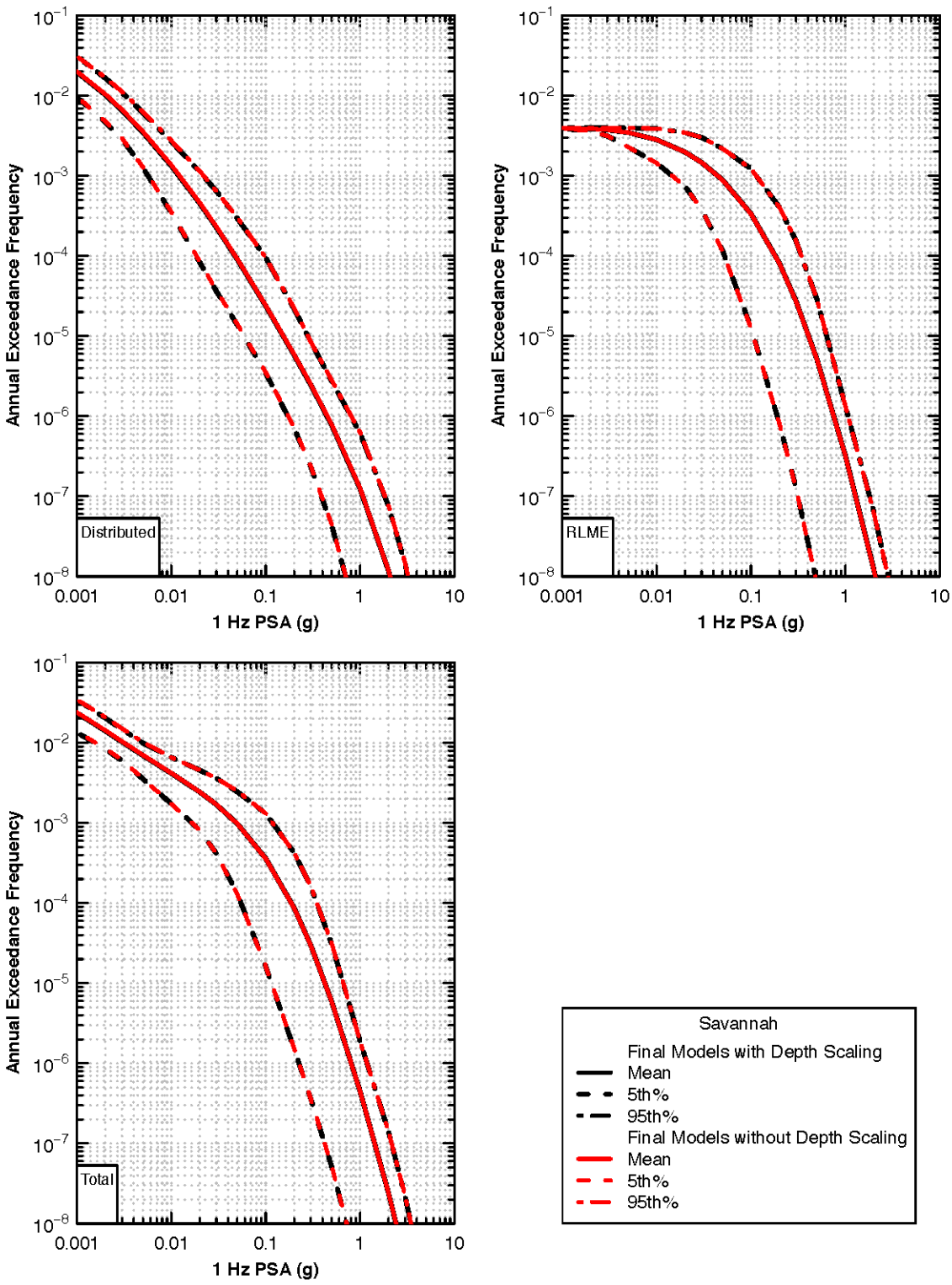


Figure 14–15 Comparison of 10 Hz hazard at Savannah computed with and without depth-effects adjustments.

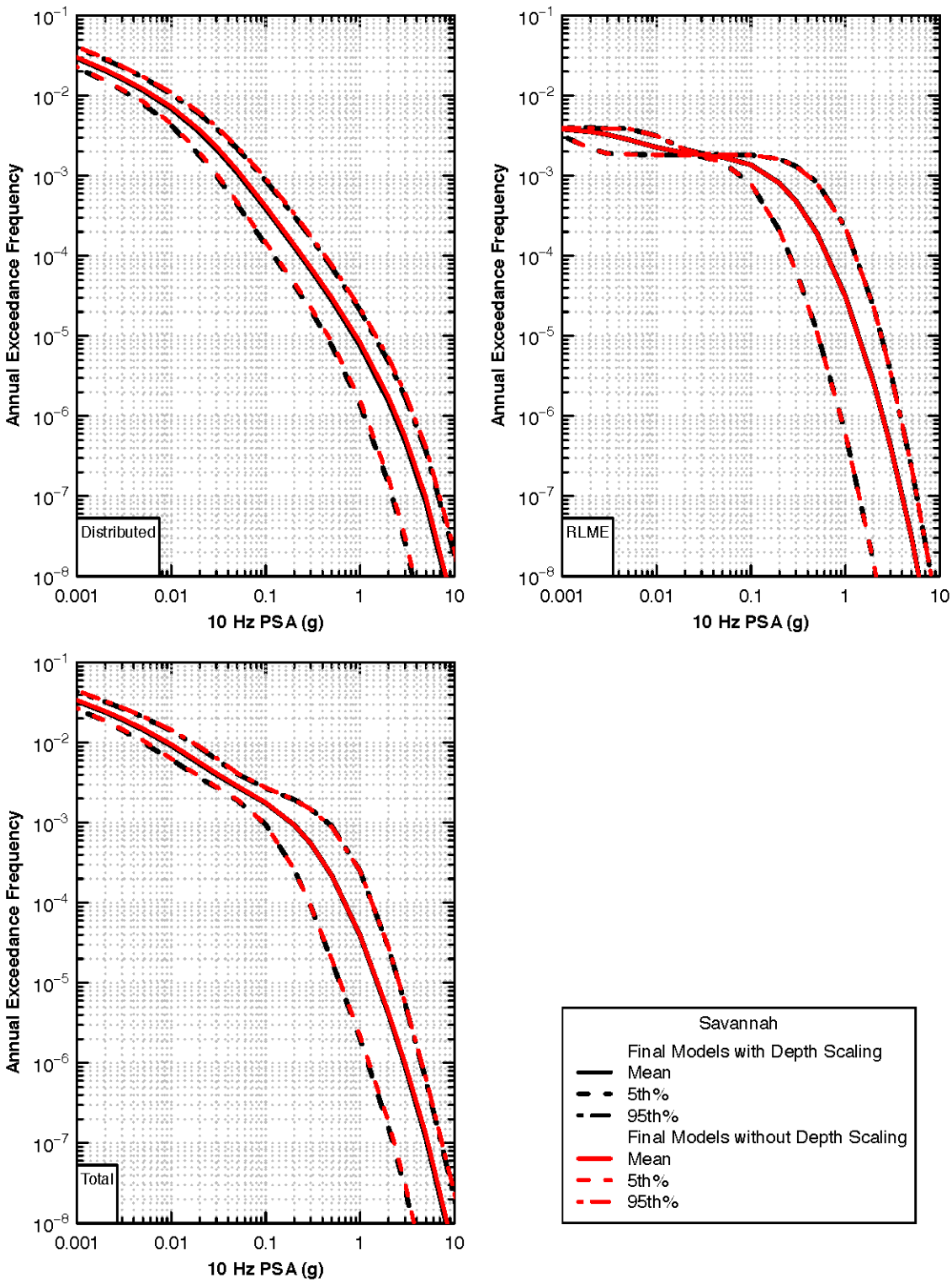


Figure 14–16 Comparison of 1 Hz hazard at Central Illinois computed with and without depth-effects adjustments.

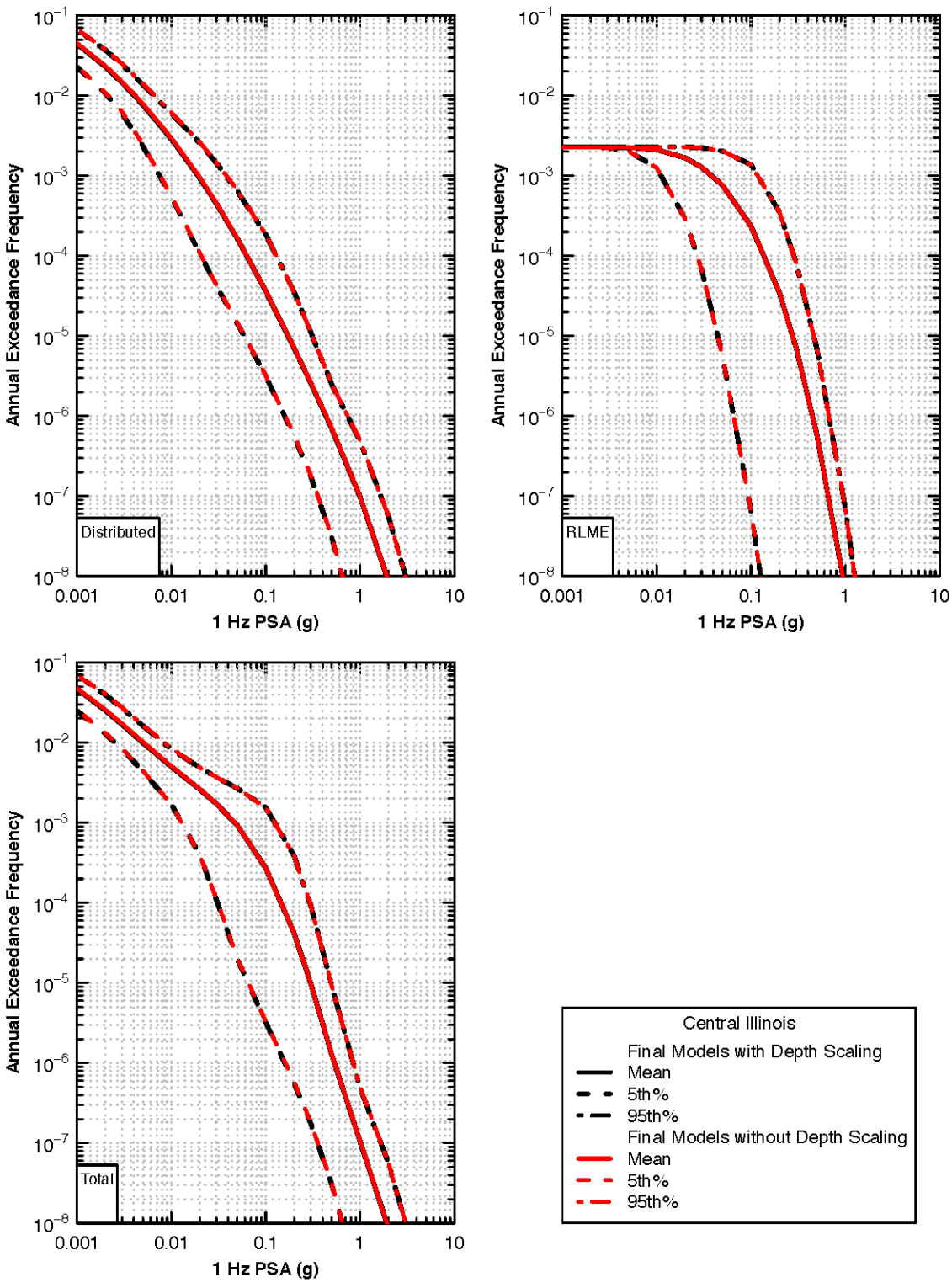


Figure 14–17 Comparison of 10 Hz hazard at Central Illinois computed with and without depth-effects adjustments.

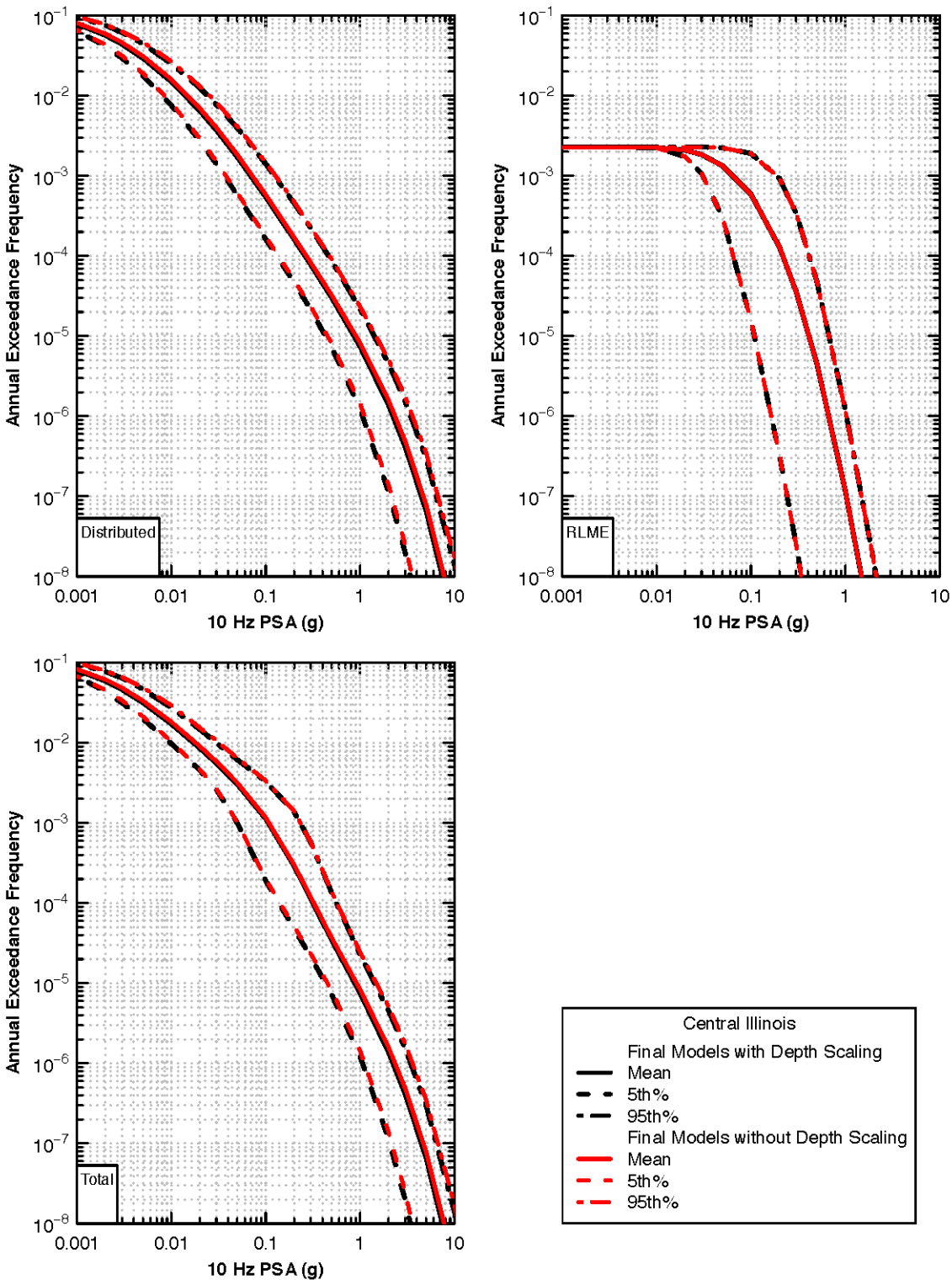


Figure 14–18 Comparison of 1 Hz hazard at Manchester computed with and without depth-effects adjustments.

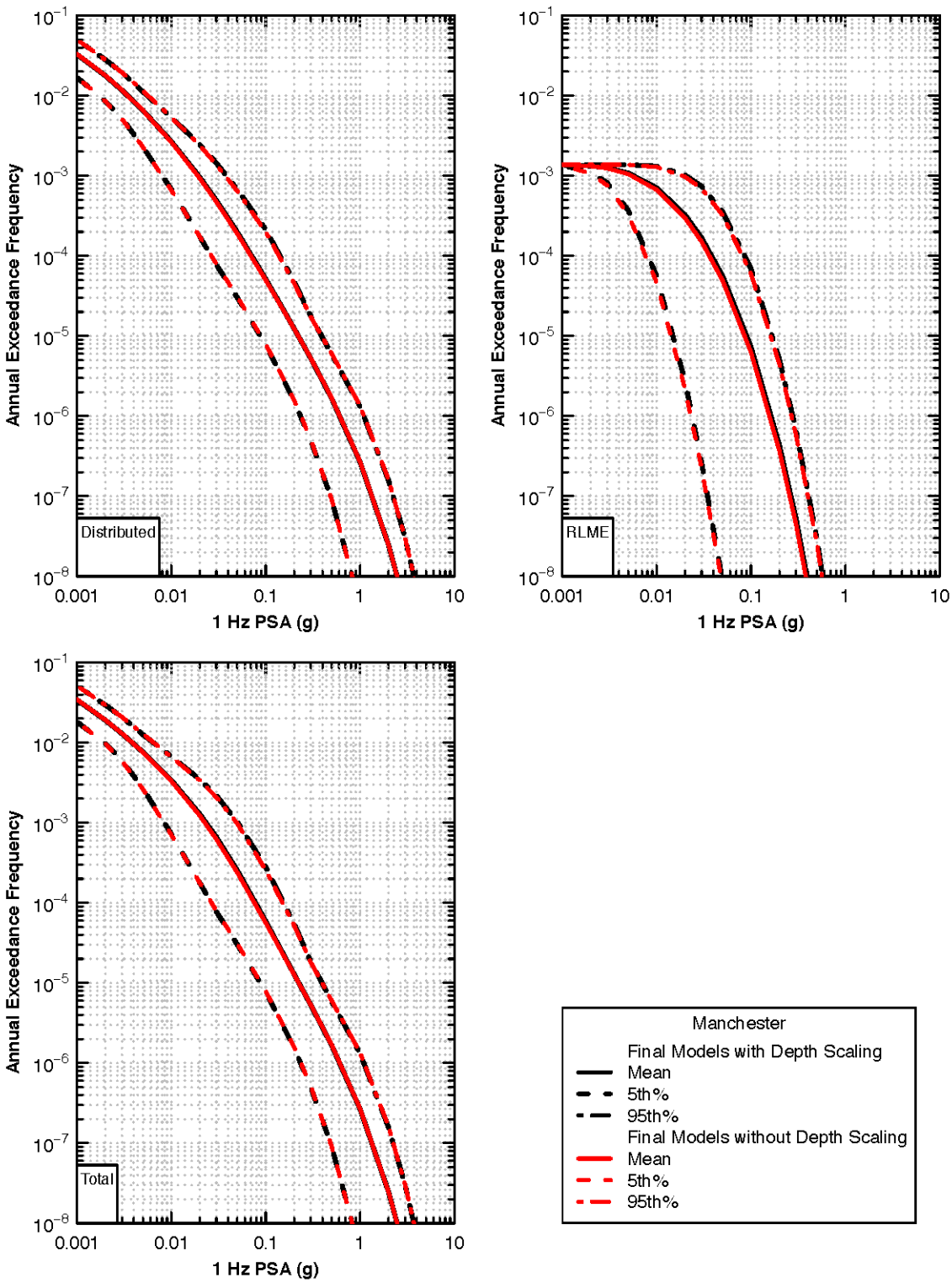


Figure 14–19 Comparison of 10 Hz hazard at Manchester computed with and without depth-effects adjustments.

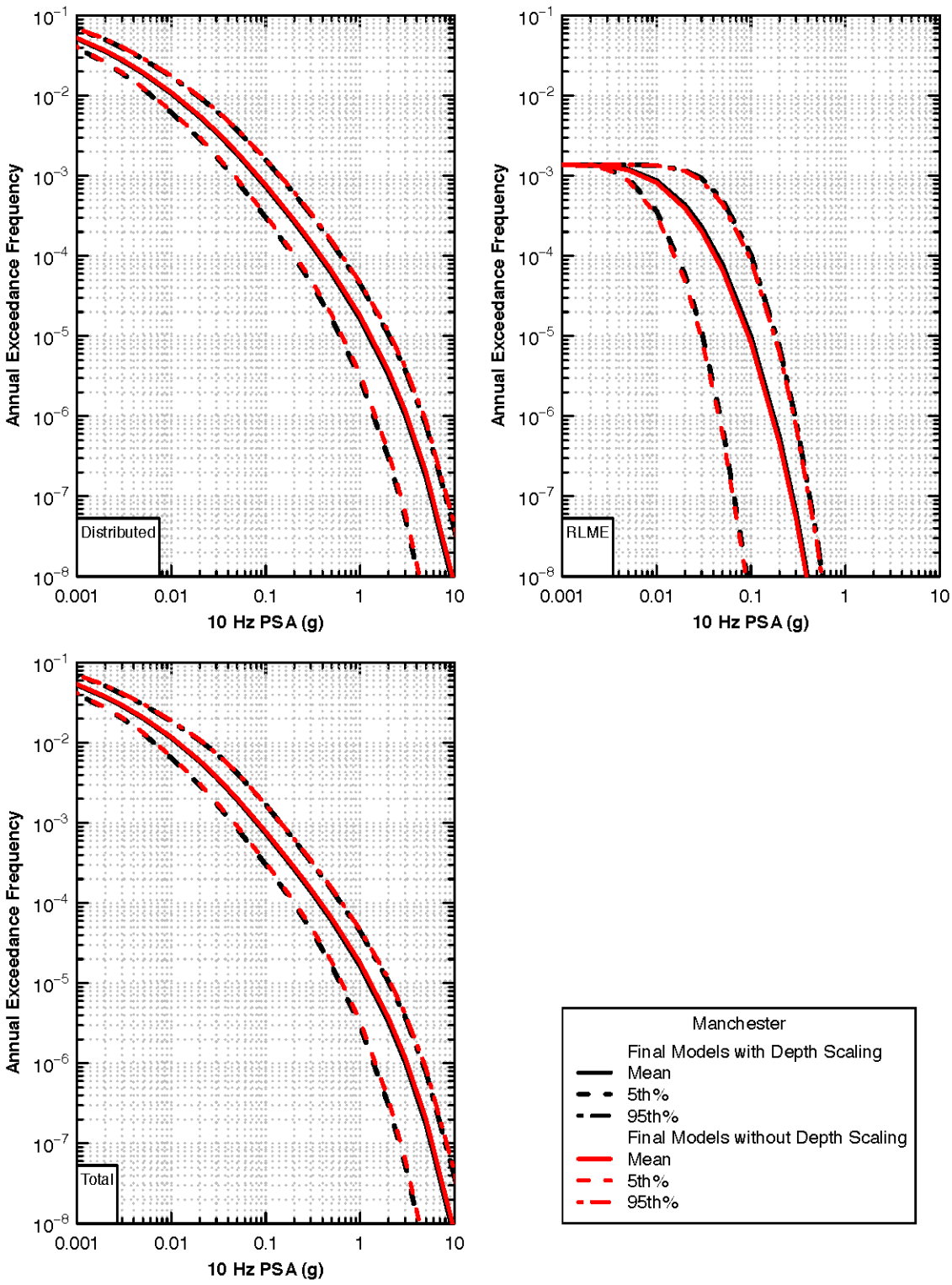


Figure 14–20 Comparison of 1 Hz hazard at Chattanooga computed with and without depth-effects adjustments.

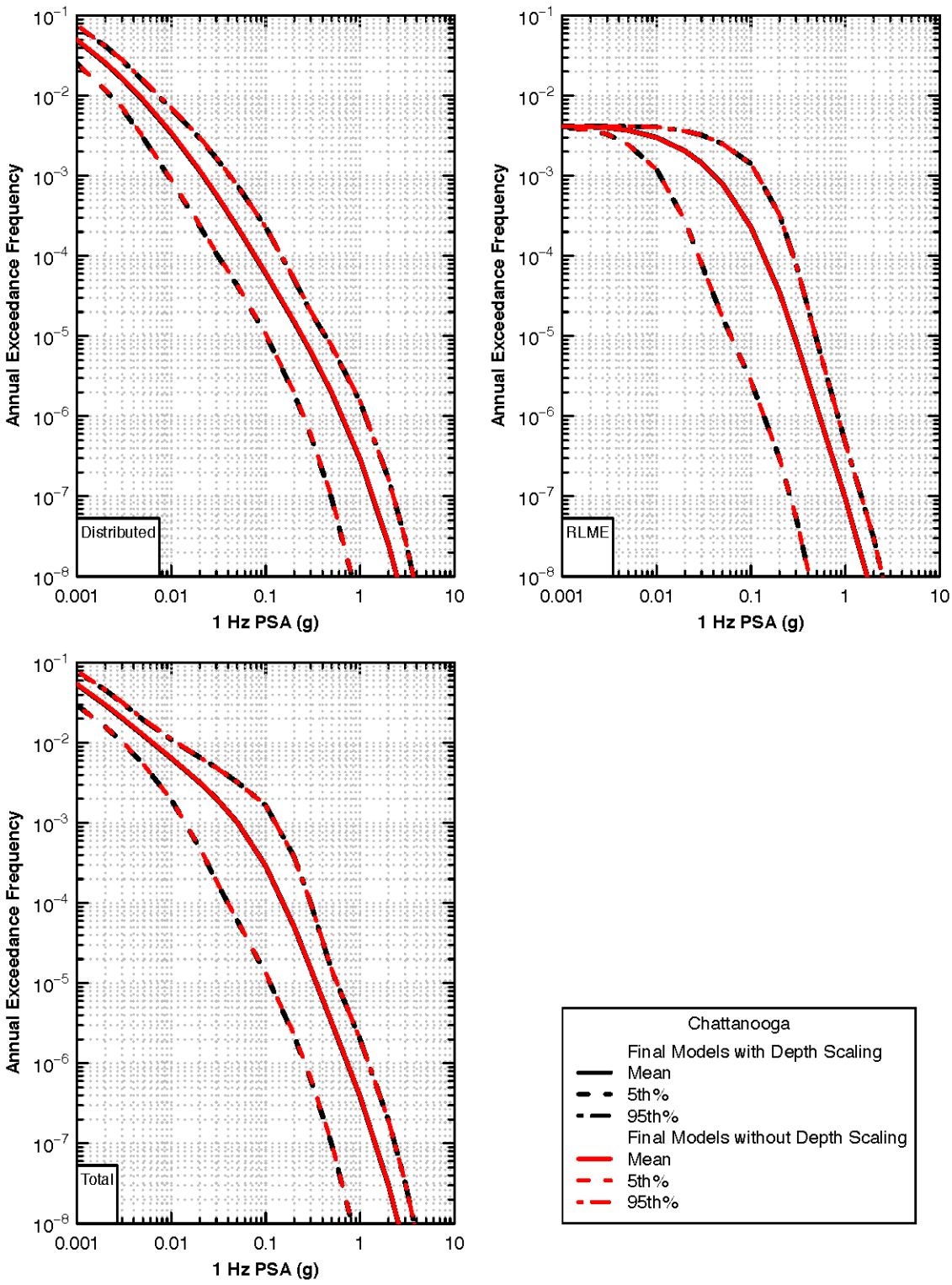


Figure 14–21 Comparison of 10 Hz hazard at Chattanooga computed with and without depth-effects adjustments.

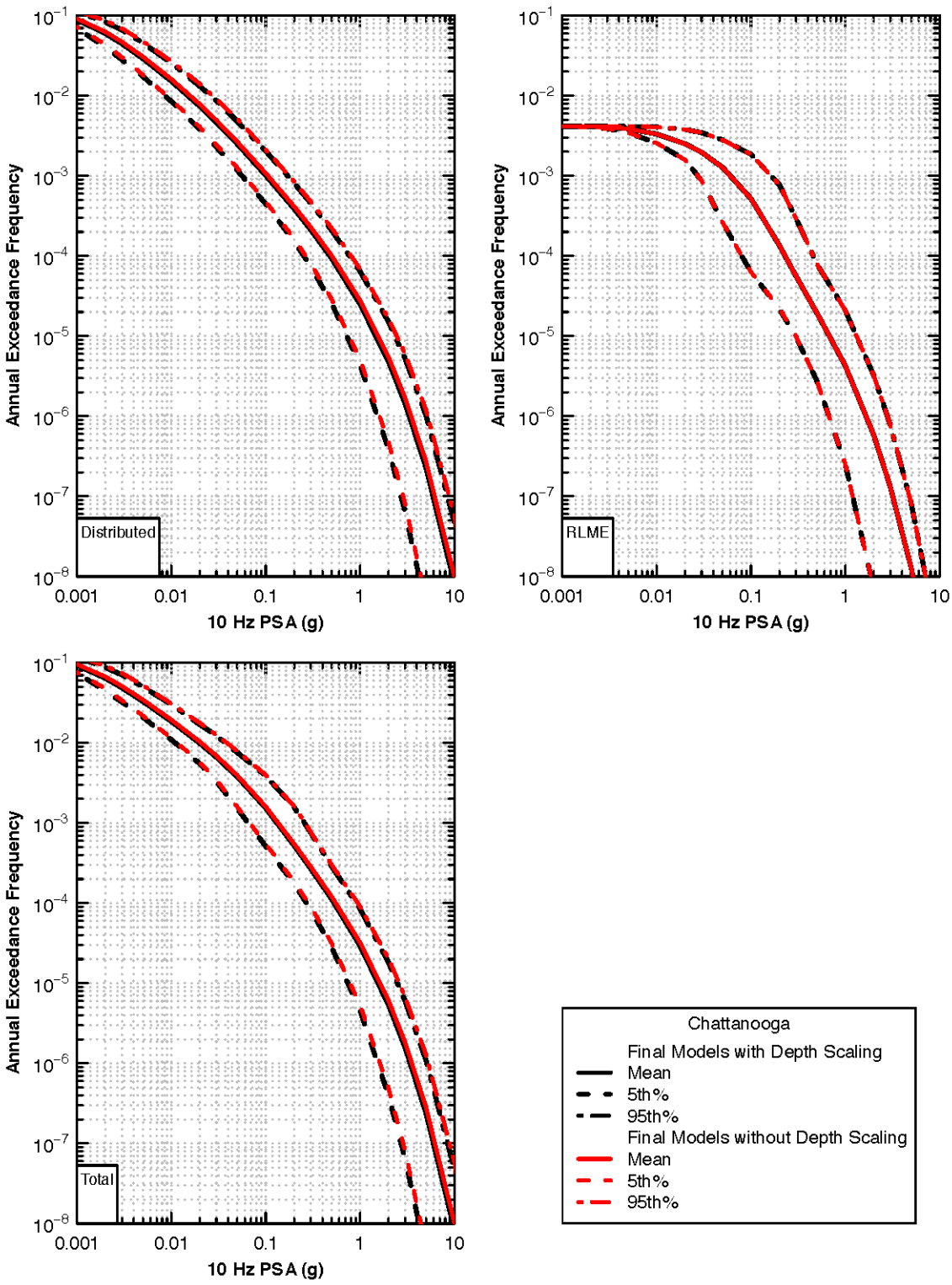


Figure 14–22 Comparison of 1 Hz hazard at Topeka computed with and without depth-effects adjustments.

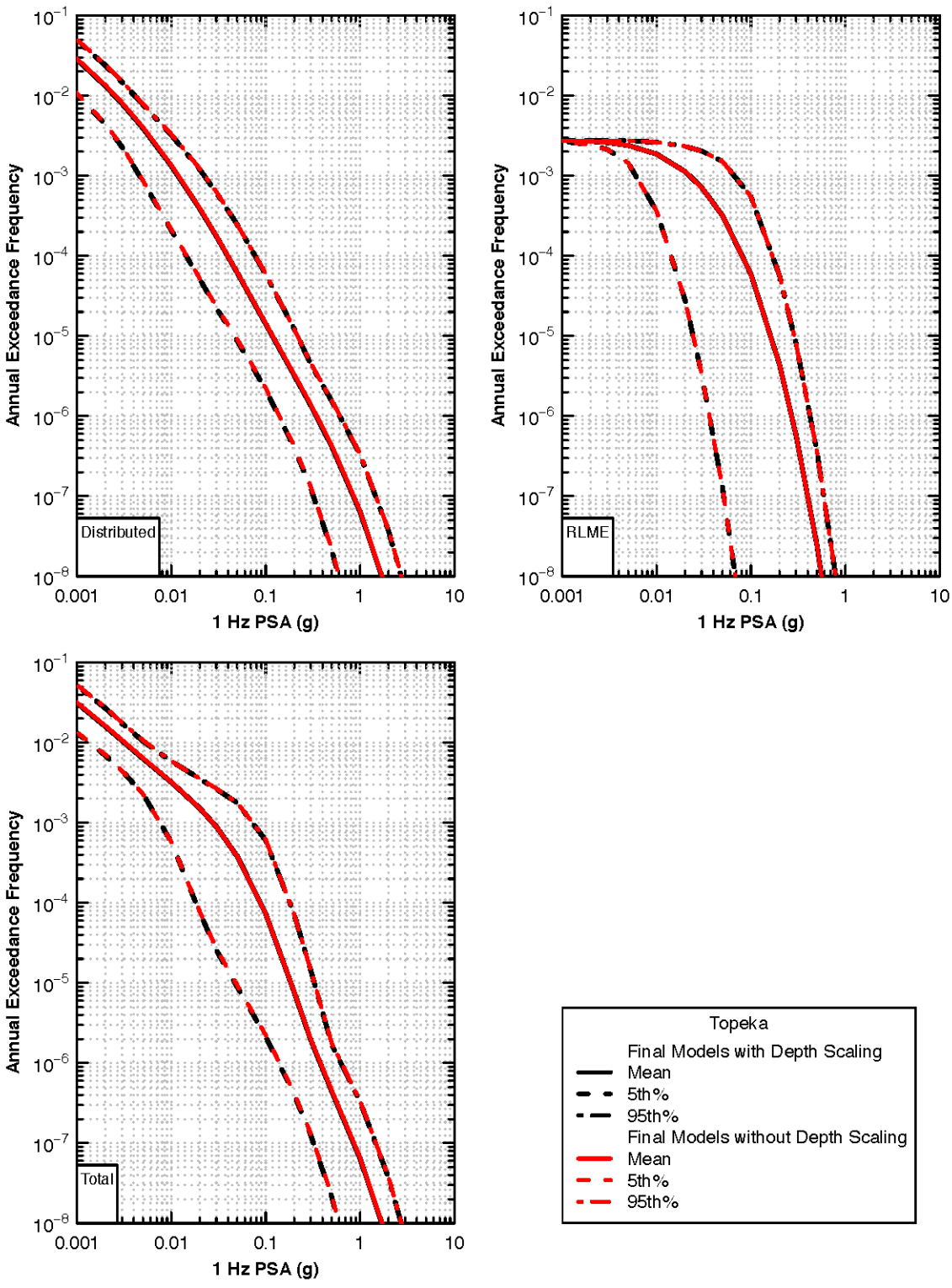


Figure 14–23 Comparison of 10 Hz hazard at Topeka computed with and without depth-effects adjustments.

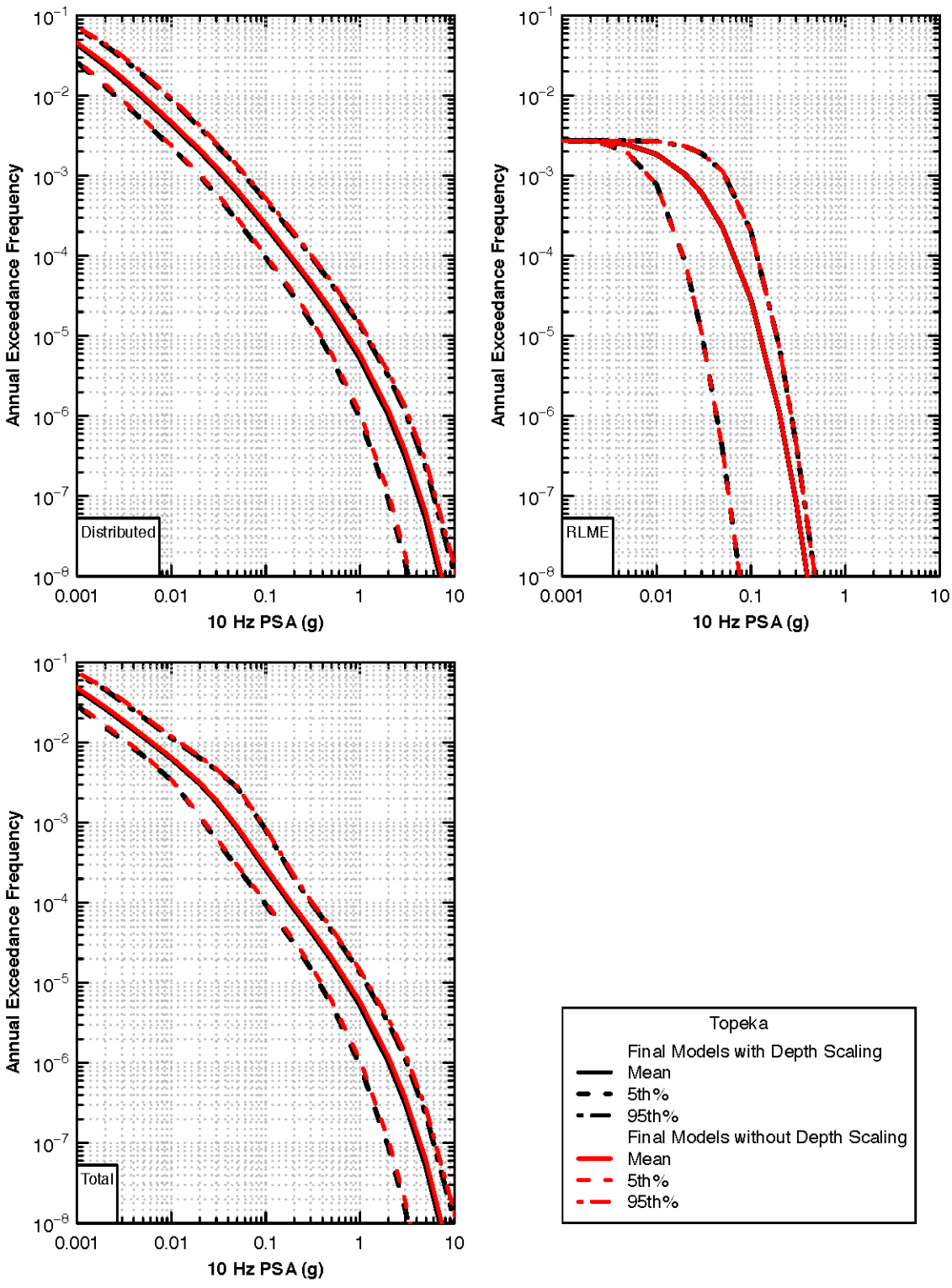


Figure 14–24 Ratio of mean hazard at Savannah computed without depth-effects adjustments to that computed with depth effects.

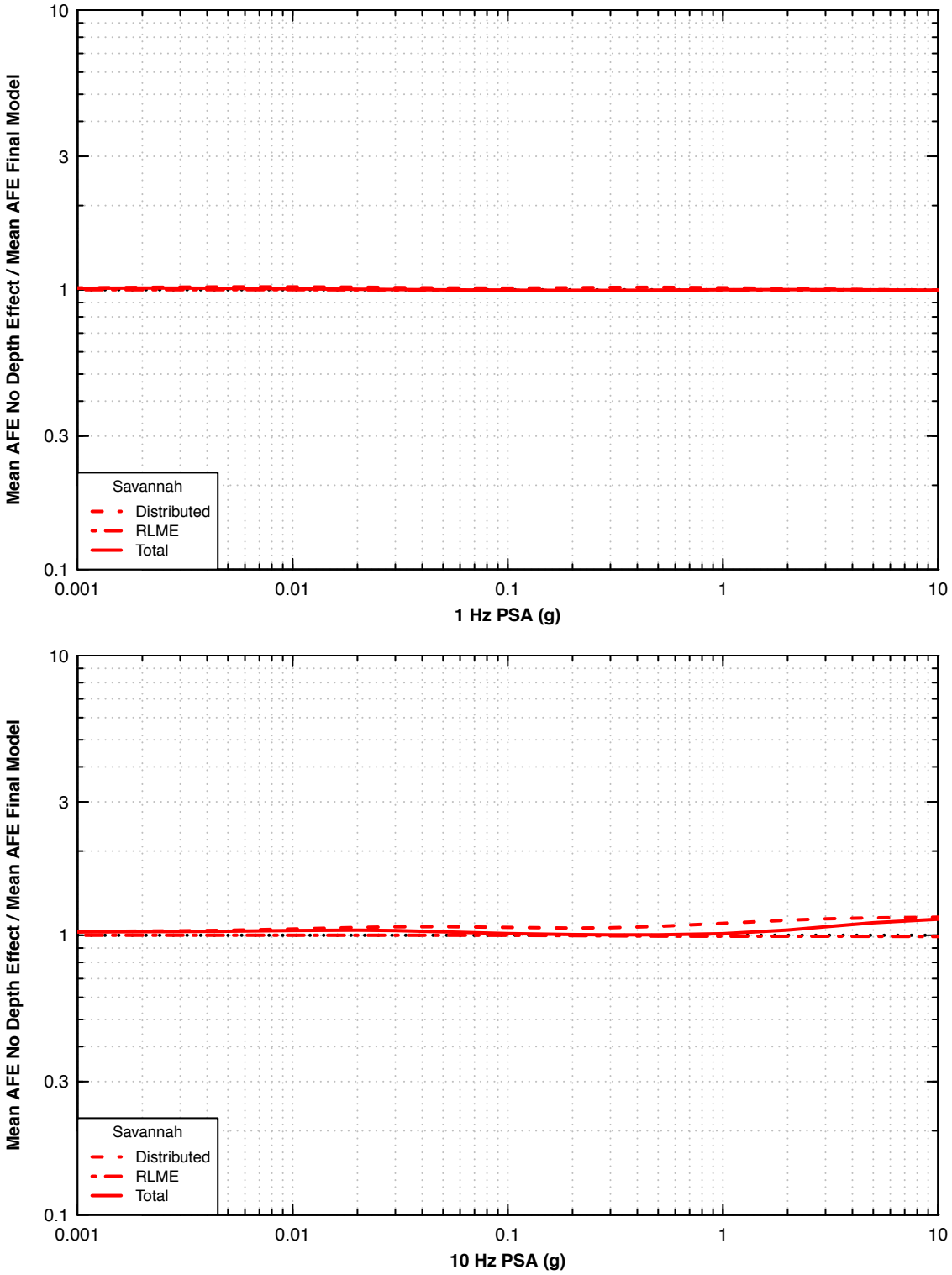


Figure 14–25 Ratio of mean hazard at Central Illinois computed without depth-effects adjustments to that computed with depth effects.

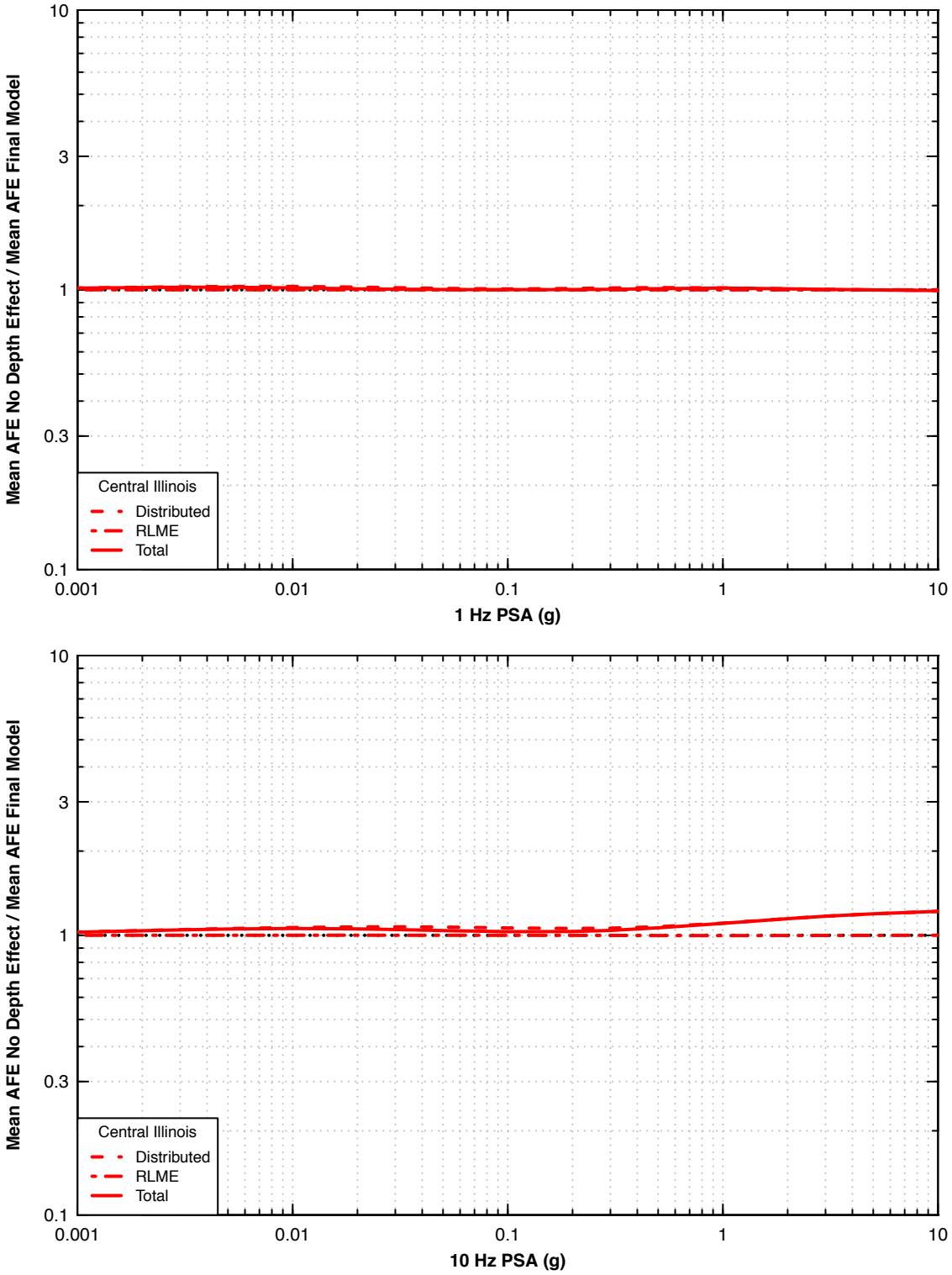


Figure 14–26 Ratio of mean hazard at Manchester computed without depth-effects adjustments to that computed with depth effects.

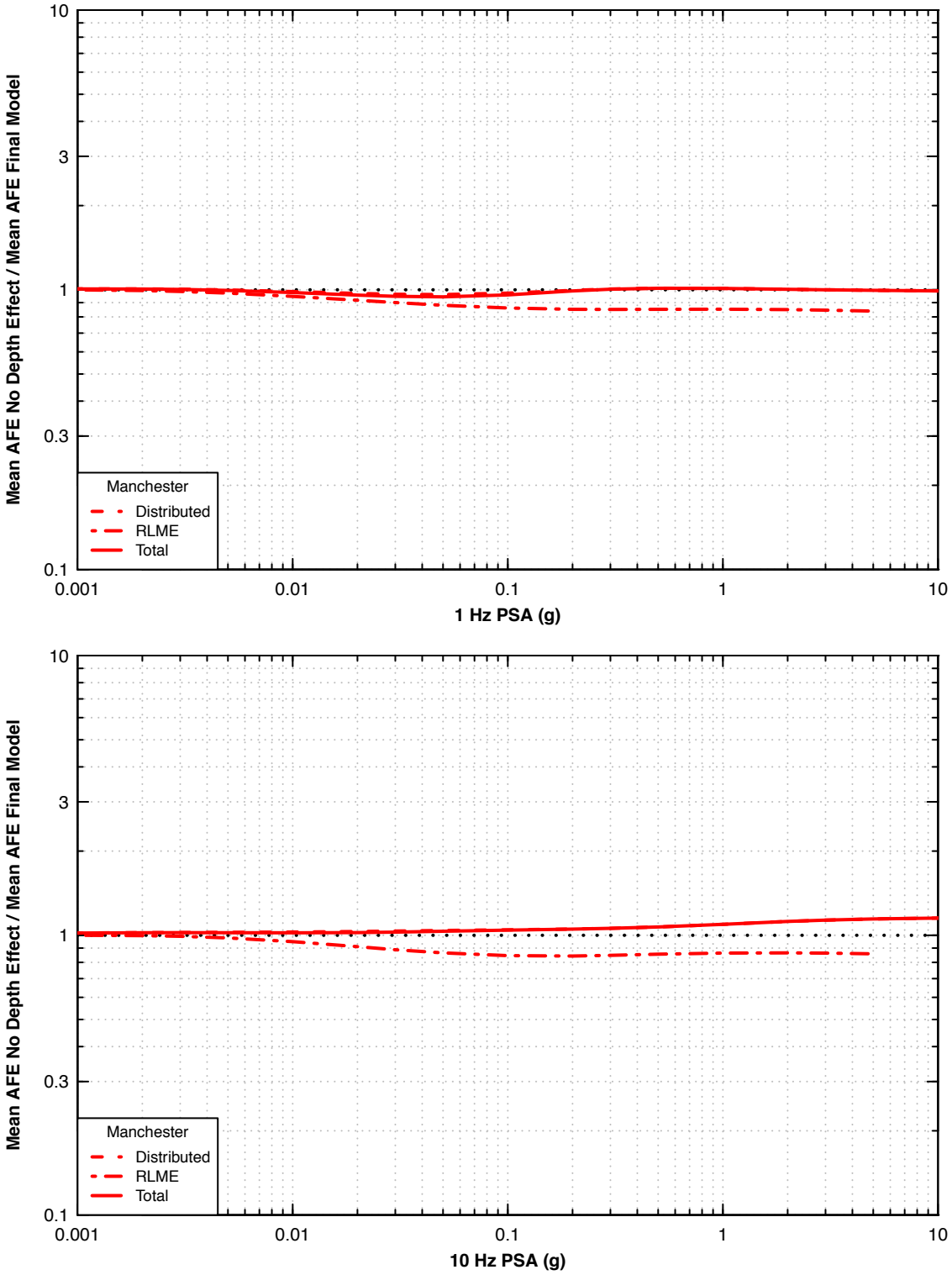


Figure 14–27 Ratio of mean hazard at Chattanooga computed without depth-effects adjustments to that computed with depth effects.

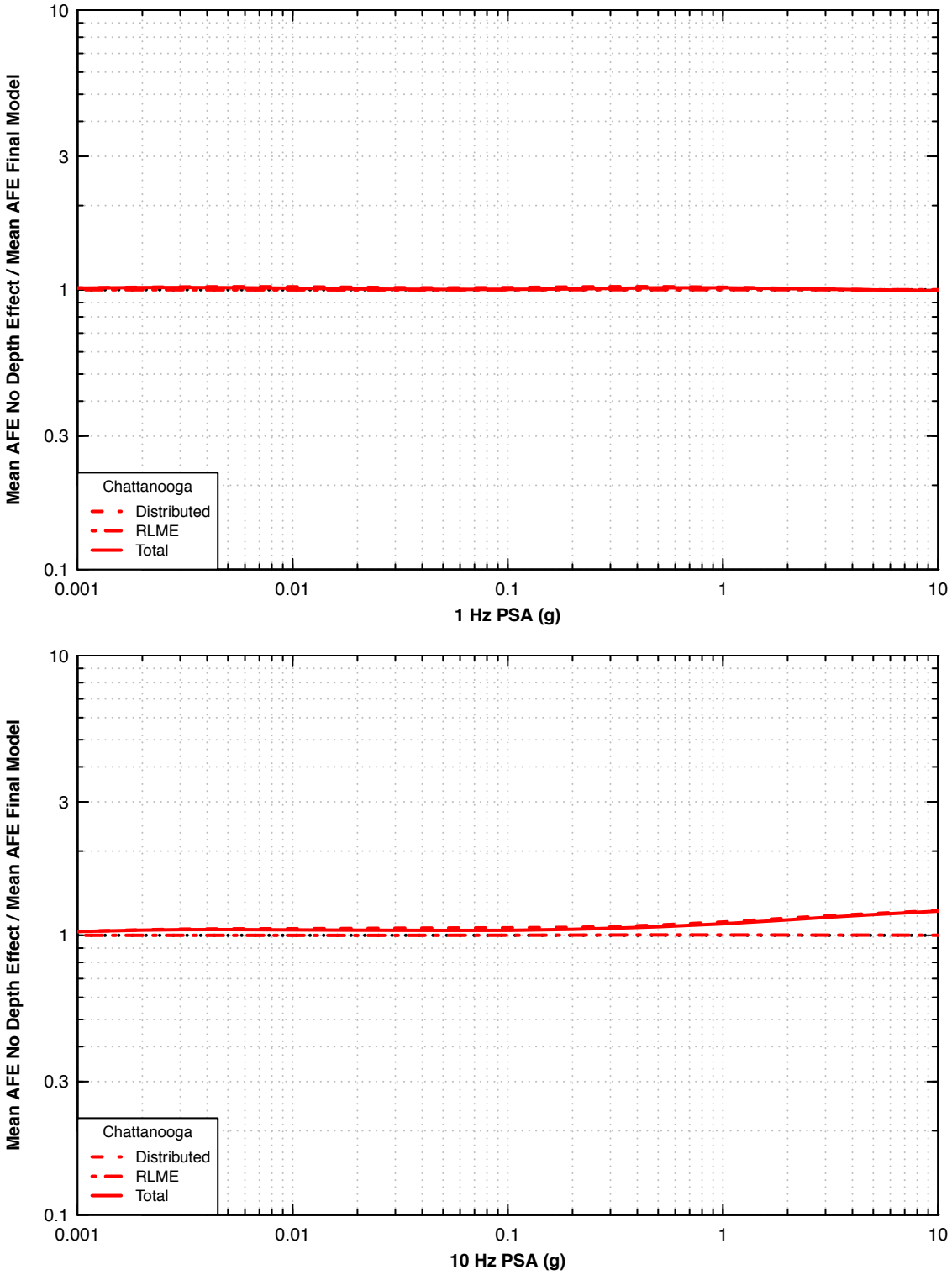


Figure 14–28 Ratio of mean hazard at Topeka computed without depth-effects adjustments to that computed with depth effects.

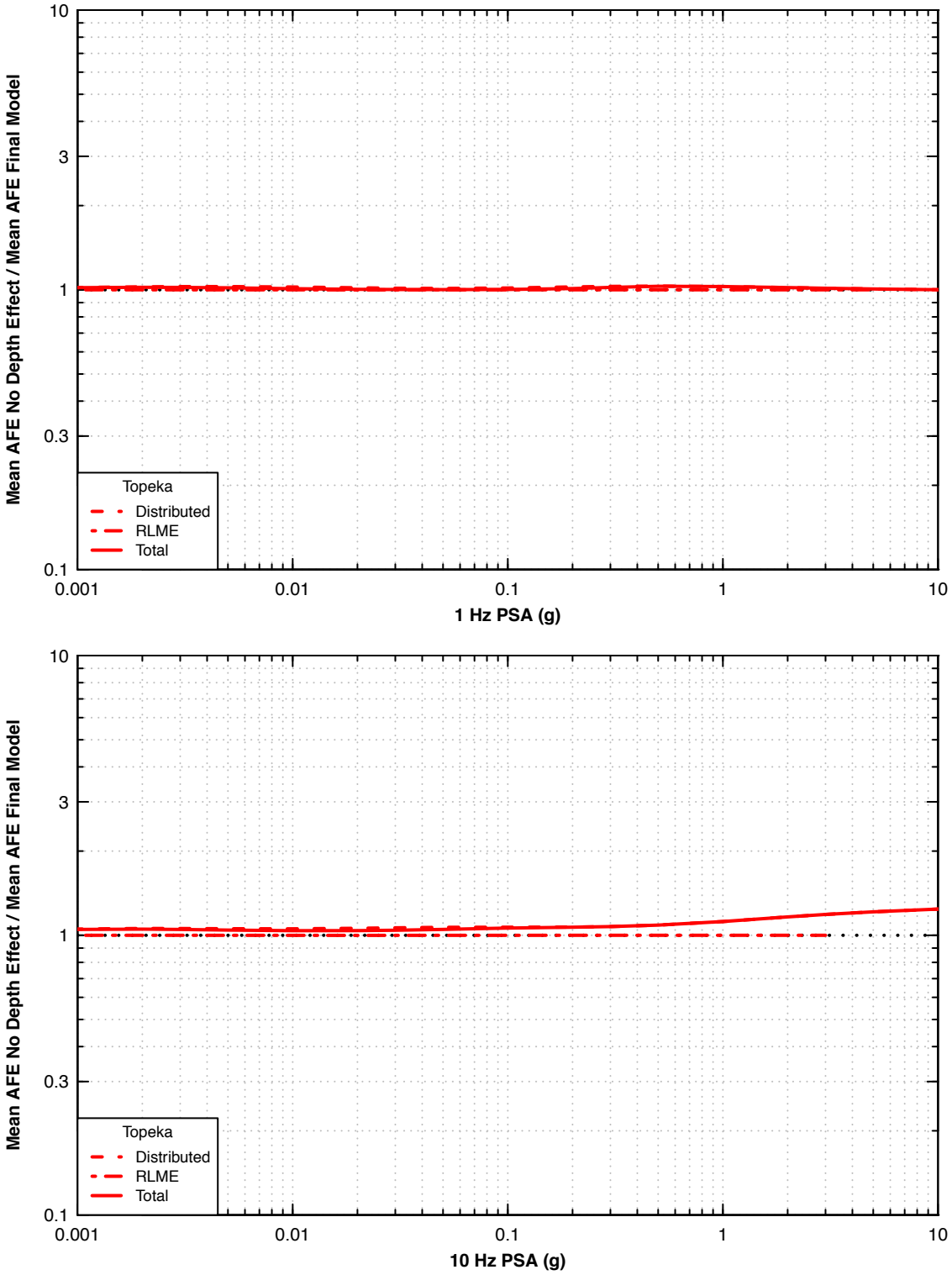


Figure 14–29 Average depth effects adjustments as a function of magnitude for the majority of the CEUS (top) and the DEEP (SLR) zone (bottom).

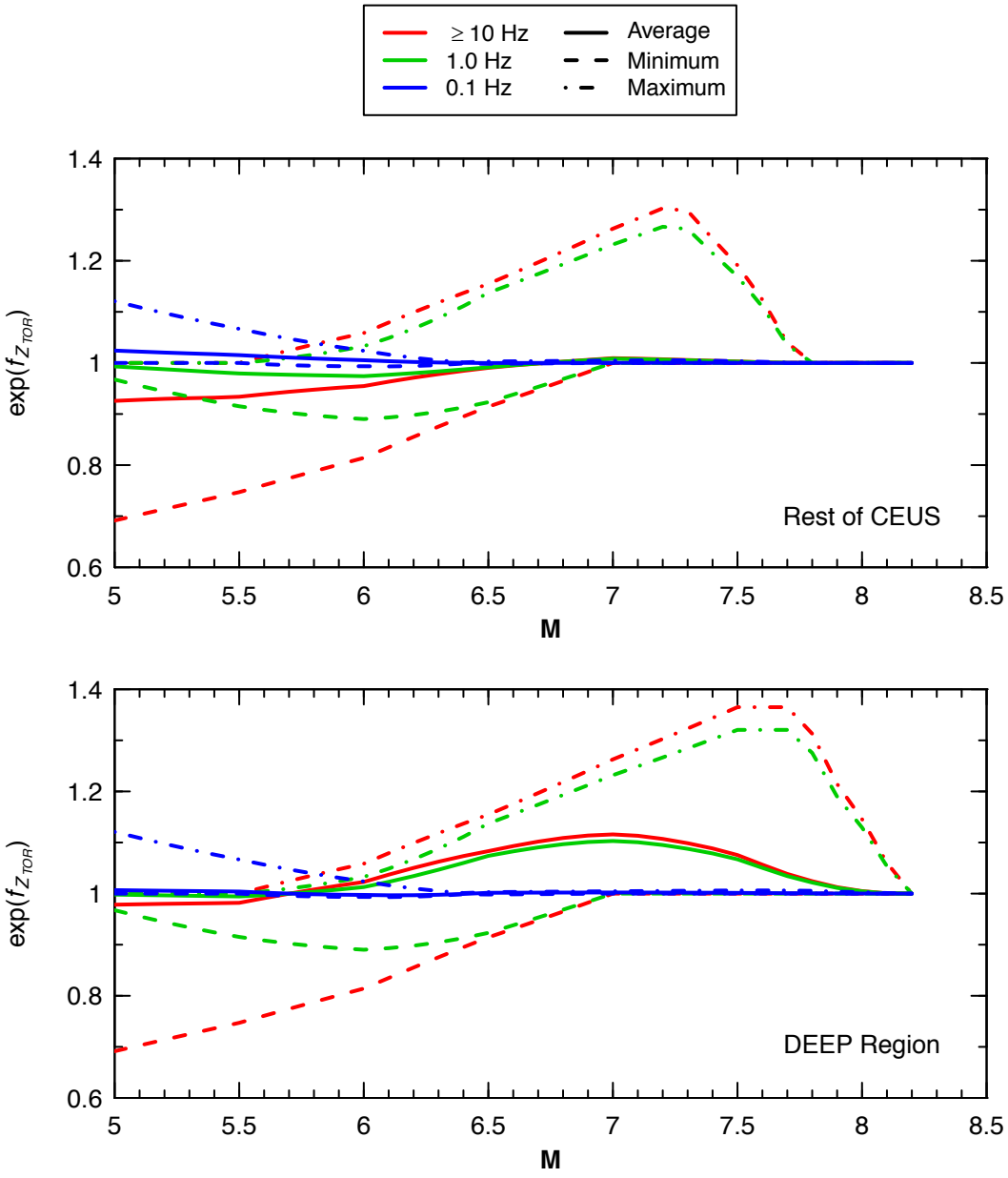


Figure 14–30 Comparison of hazard from host distributed seismicity source zone computed with and without the hanging wall effects model.

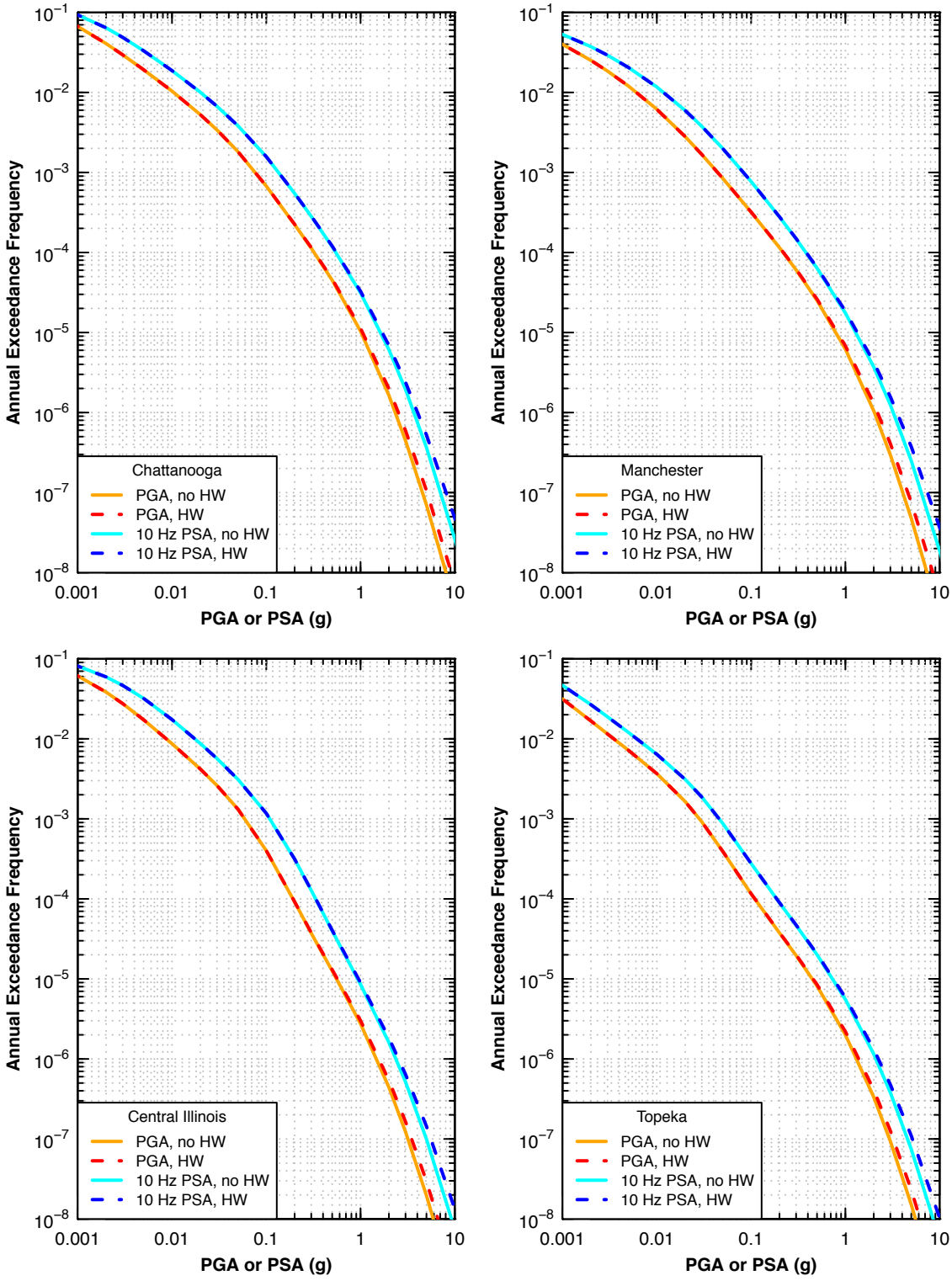


Figure 14–31 Ratio of hazard from host distributed seismicity source zone computed with and without the hanging wall effects model.

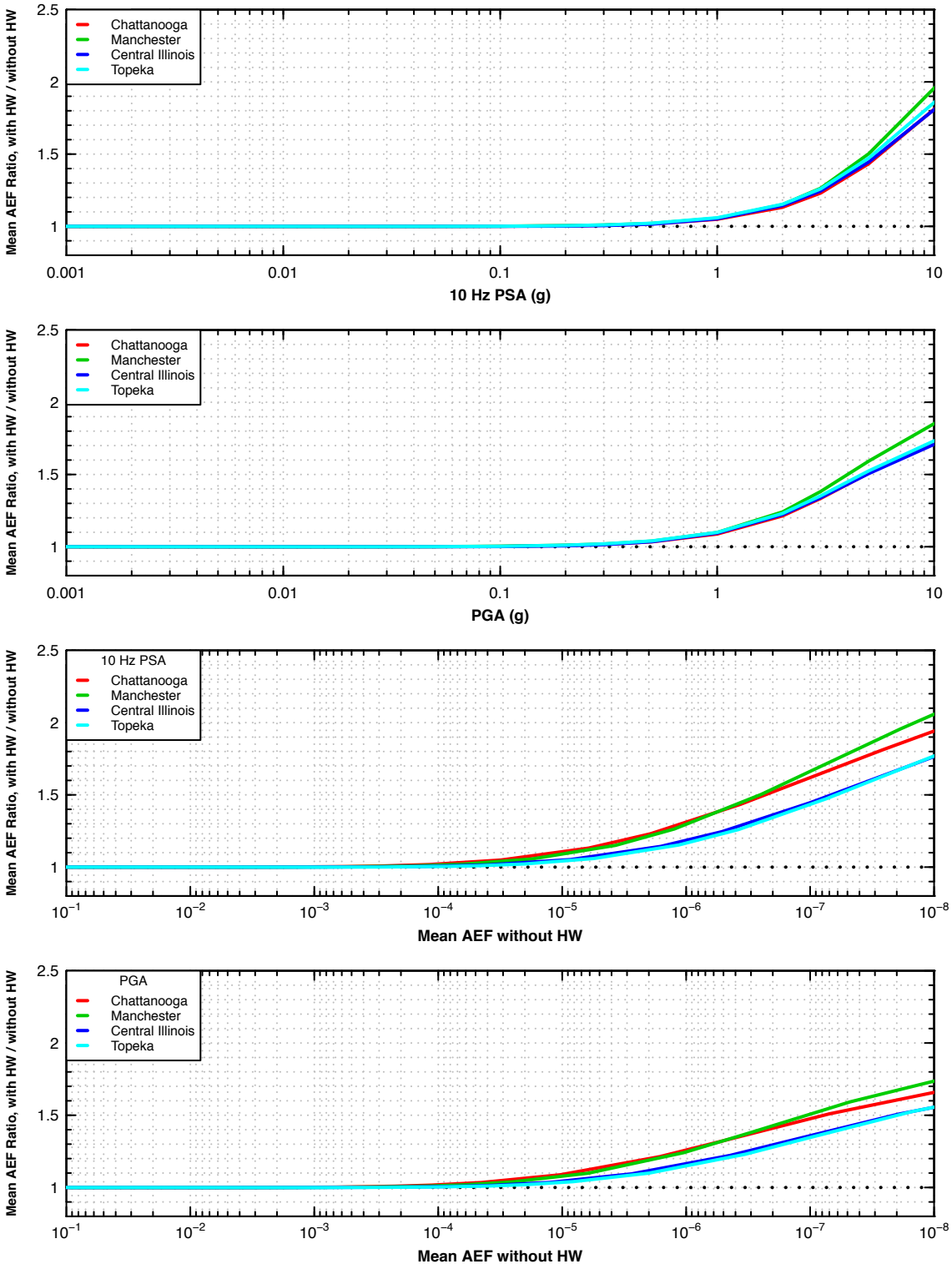


Figure 14–32 Comparison of the NGA-East single-station sigma model and the EPRI [2013] ergodic sigma model.

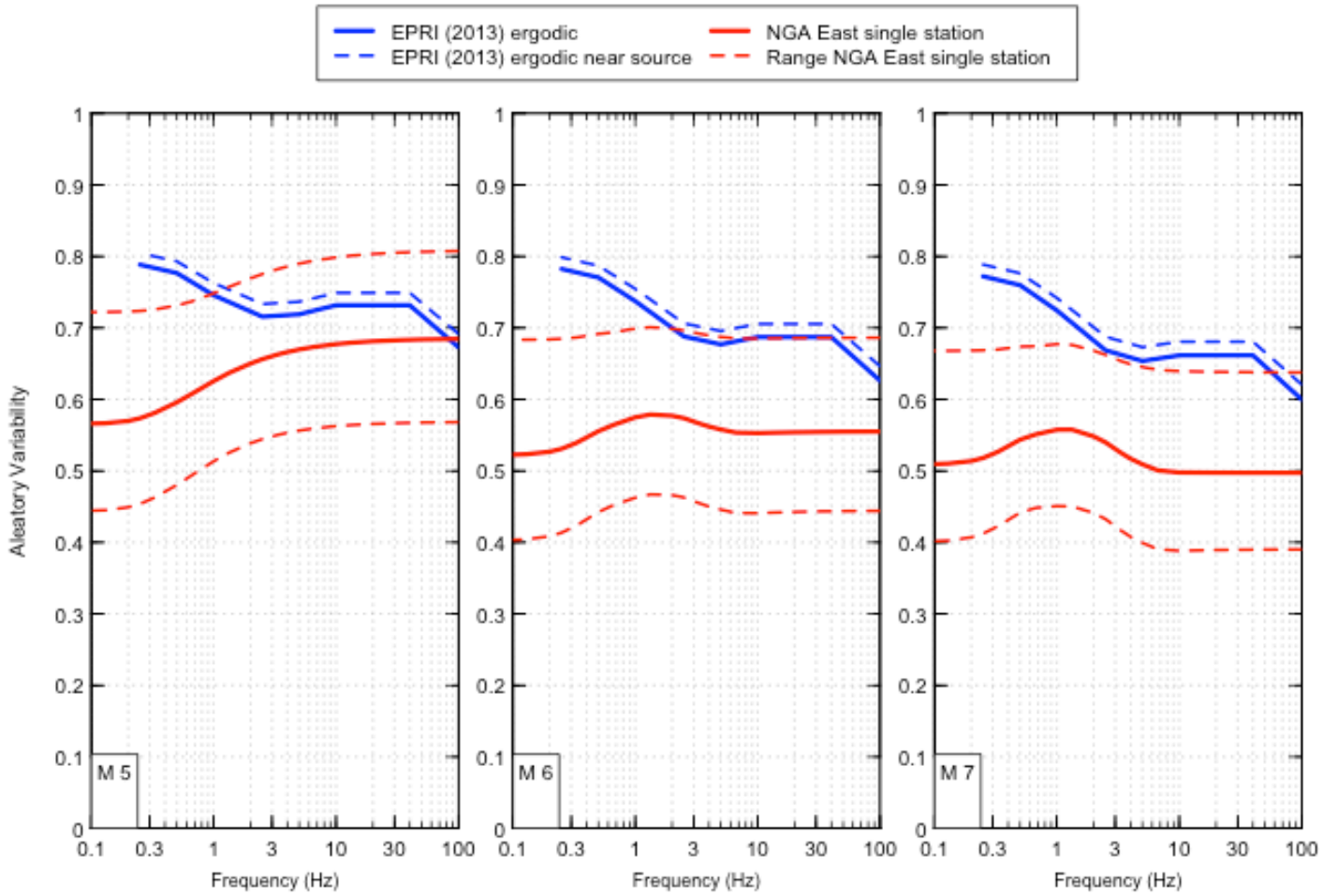


Figure 14–33 Comparison of 0.5 Hz PSA hazard at Savannah (left) and Central Illinois (right) computed using the NGA-East single-station sigma model and the EPRI [2013] ergodic sigma model, both using the EPRI (2013) median models.

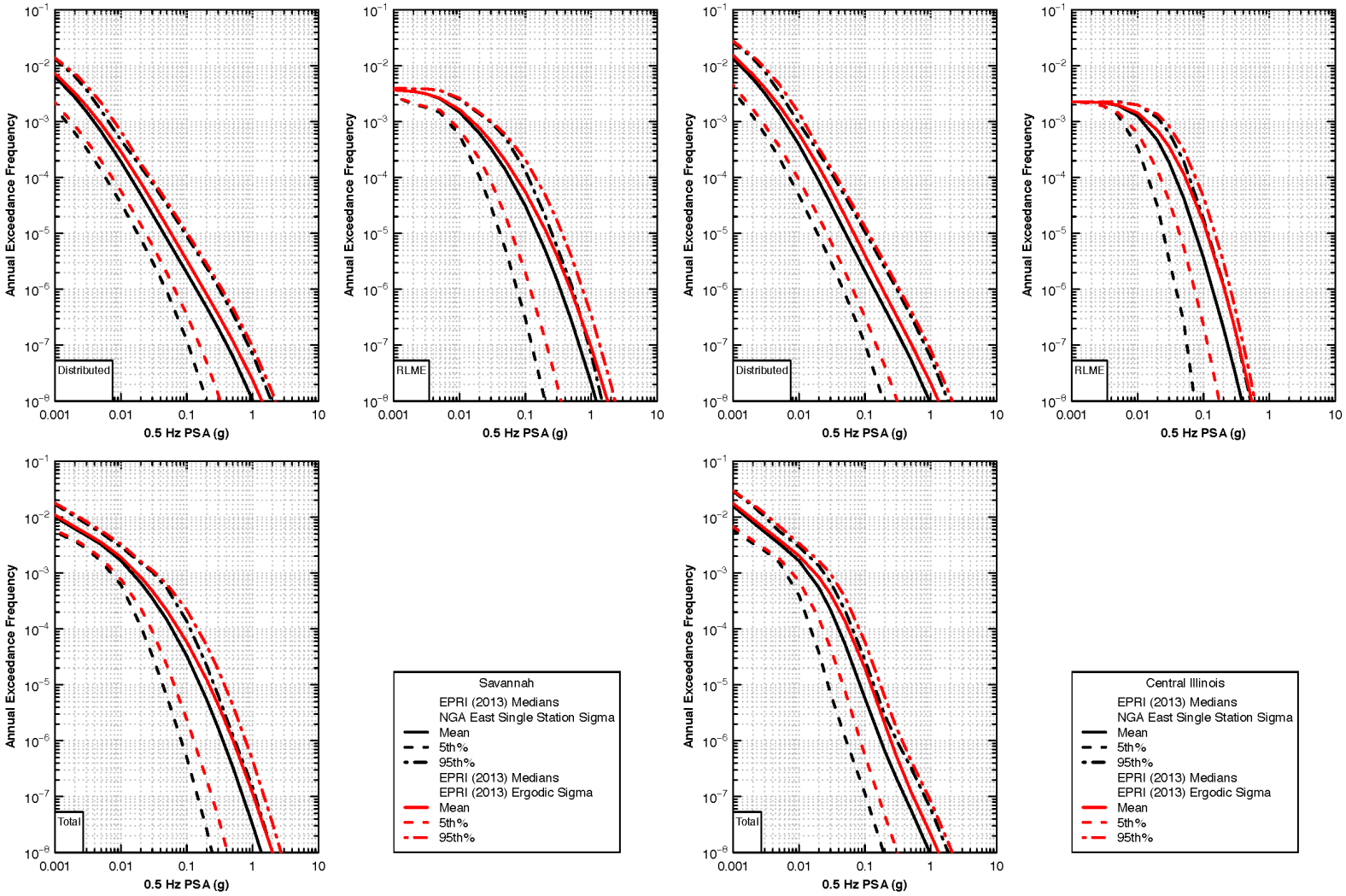


Figure 14–34 Comparison of 1 Hz PSA hazard at Savannah (left) and Central Illinois (right) computed using the NGA-East single-station sigma model and the EPRI (2013) ergodic sigma model, both using the EPRI (2013) median models.

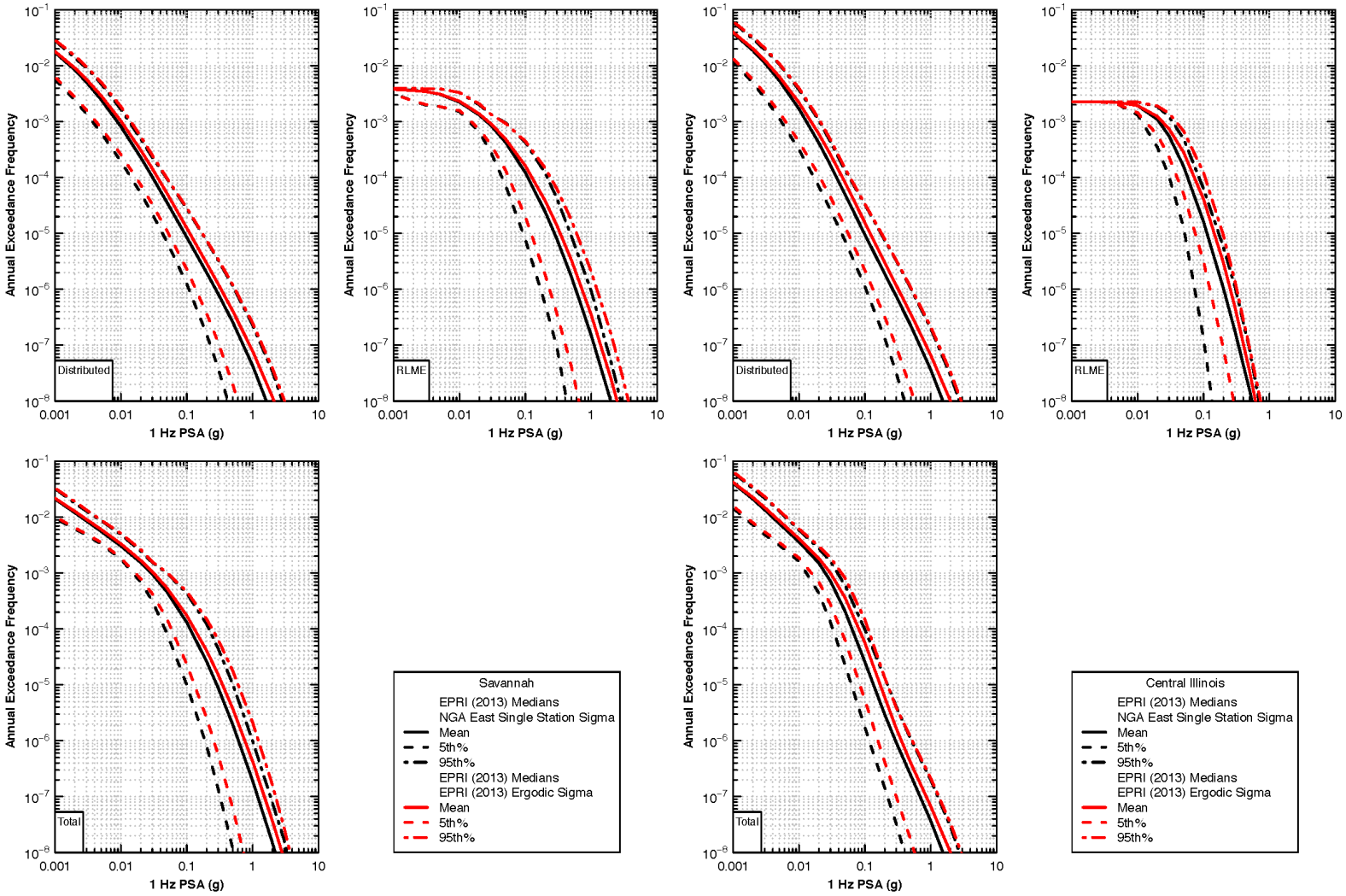


Figure 14–35 Comparison of 2.5 Hz PSA hazard at Savannah (left) and Central Illinois (right) computed using the NGA-East single-station sigma model and the EPRI (2013) ergodic sigma model, both using the EPRI (2013) median models.

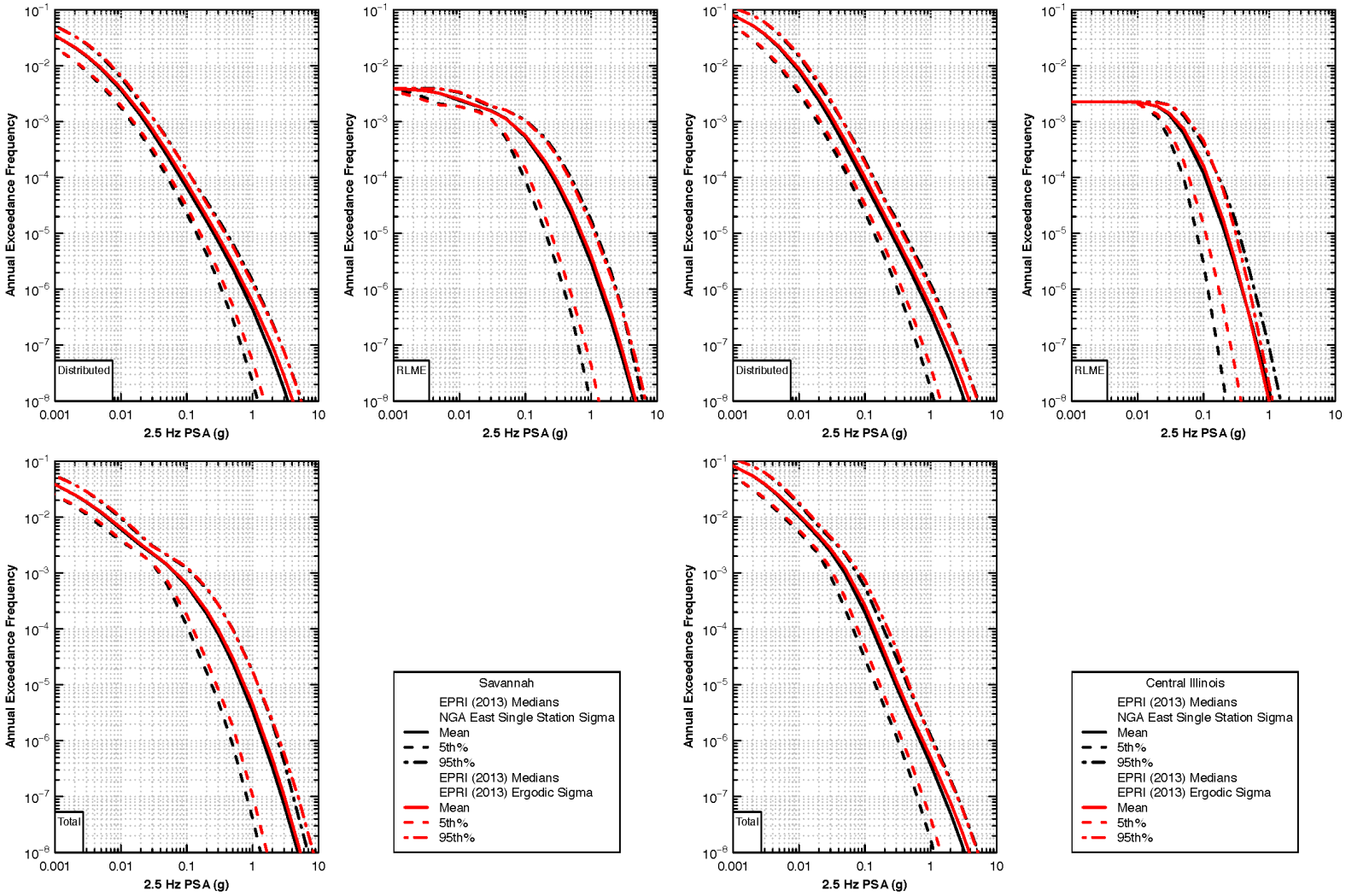


Figure 14–36 Comparison of 5 Hz PSA hazard at Savannah (left) and Central Illinois (right) computed using the NGA-East single-station sigma model and the EPRI (2013) ergodic sigma model, both using the EPRI (2013) median models.

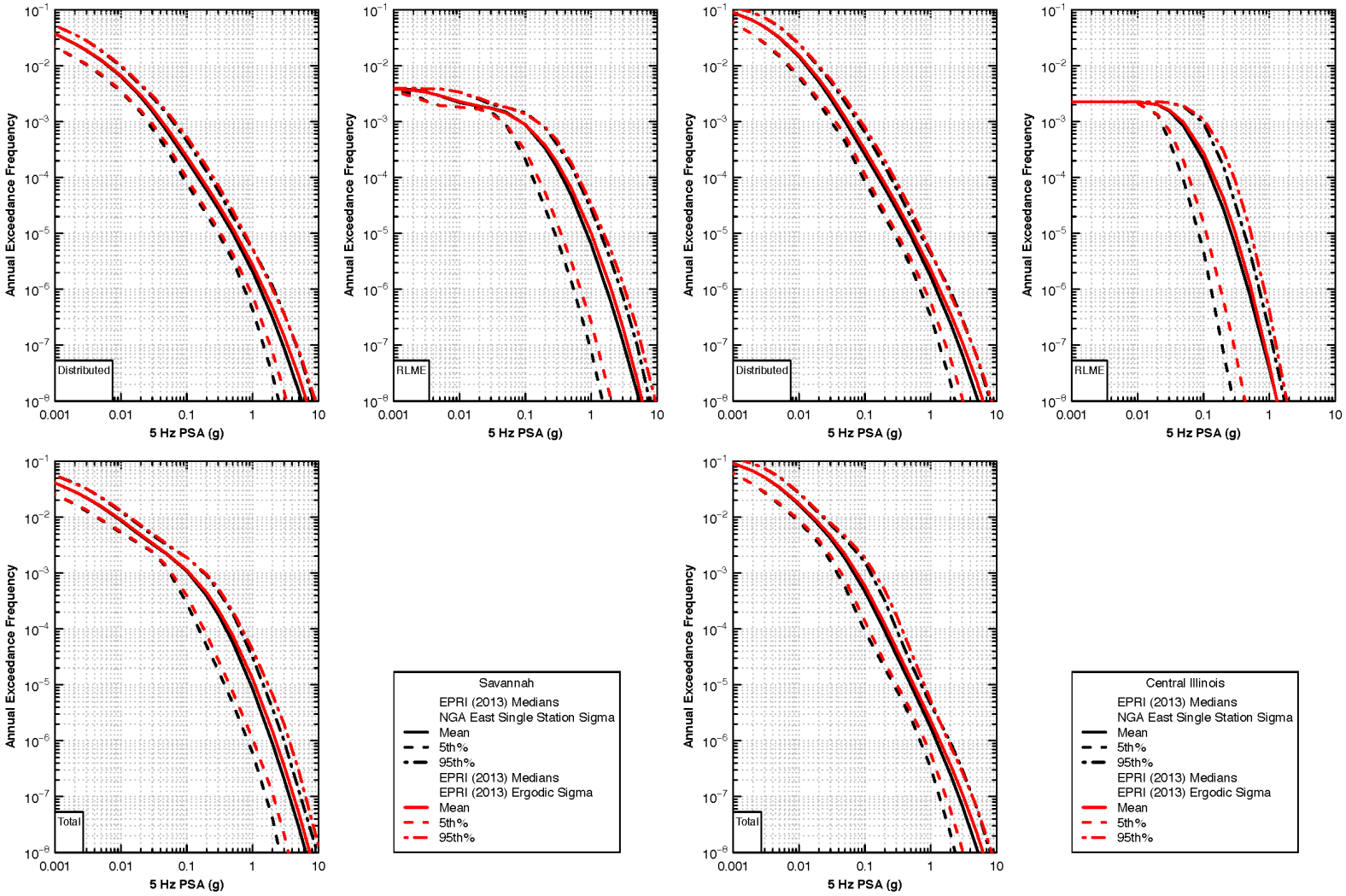


Figure 14–37 Comparison of 10 Hz PSA hazard at Savannah (left) and Central Illinois (right) computed using the NGA-East single-station sigma model and the EPRI (2013) ergodic sigma model, both using the EPRI (2013) median models.

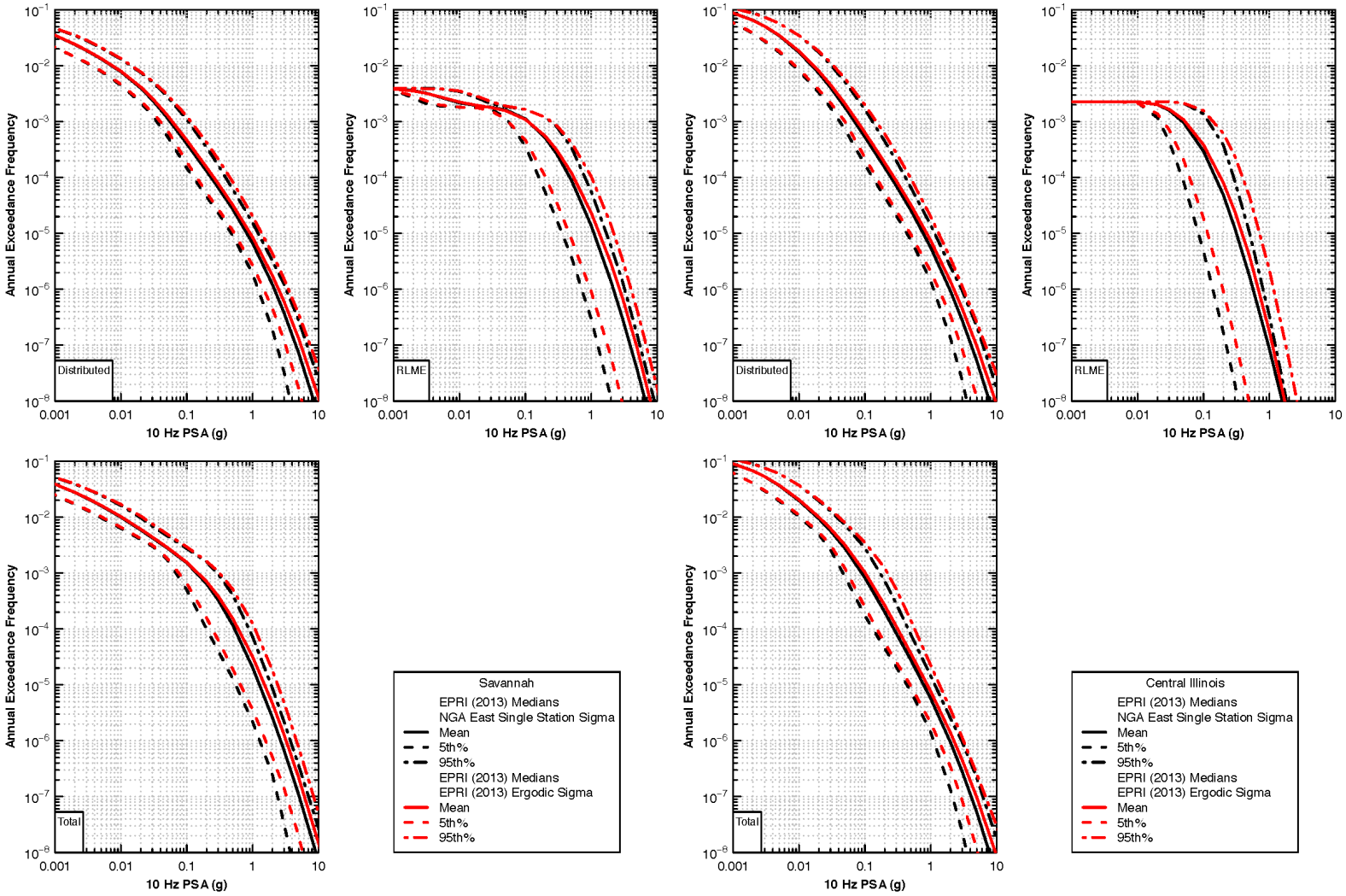


Figure 14–38 Comparison of 25 Hz PSA hazard at Savannah (left) and Central Illinois (right) computed using the NGA-East single-station sigma model and the EPRI (2013) ergodic sigma model, both using the EPRI (2013) median models.

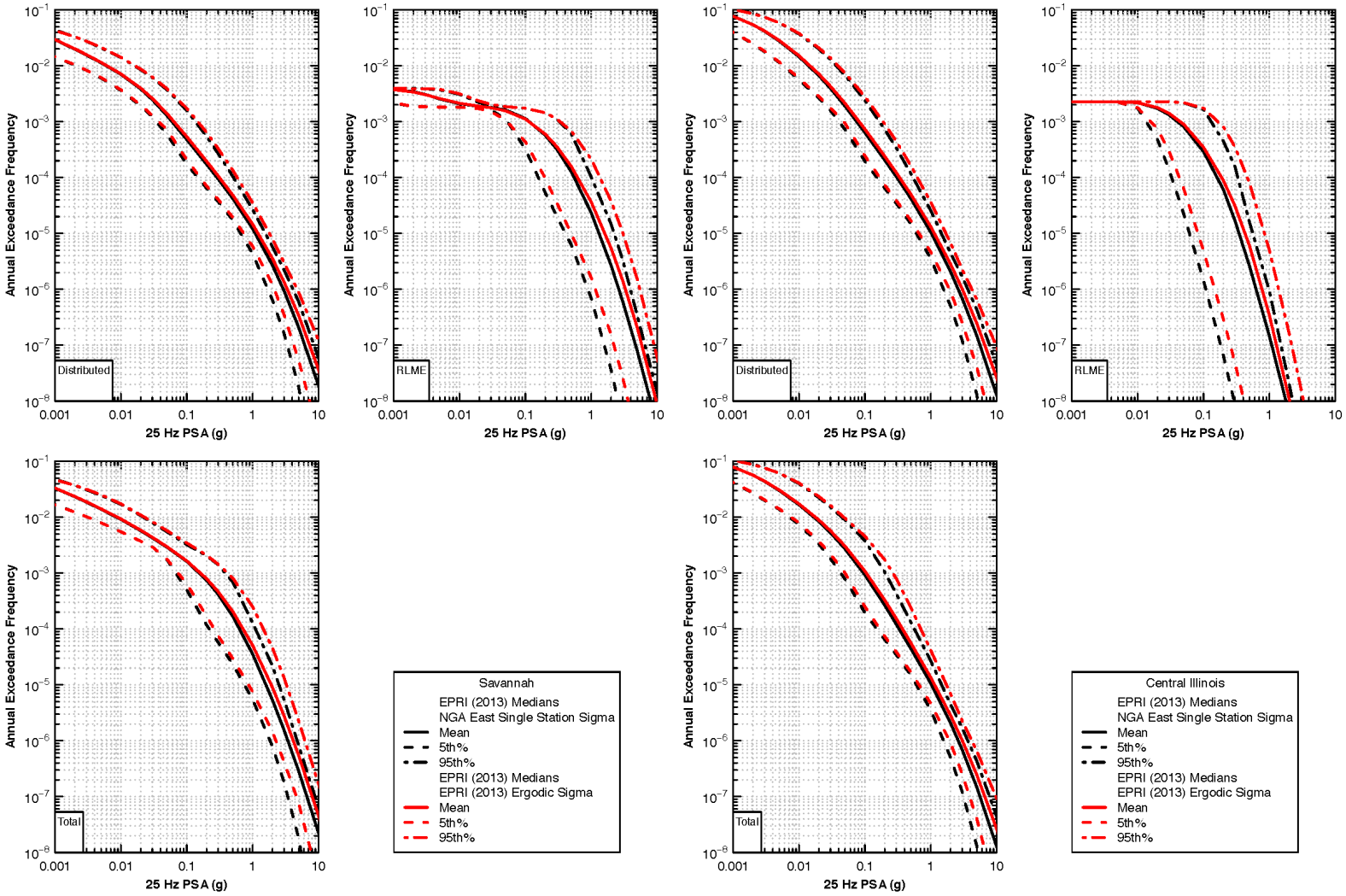


Figure 14–39 Comparison of PGA hazard at Savannah (left) and Central Illinois (right) computed using the NGA-East single-station sigma model and the EPRI (2013) ergodic sigma model, both using the EPRI (2013) median models.

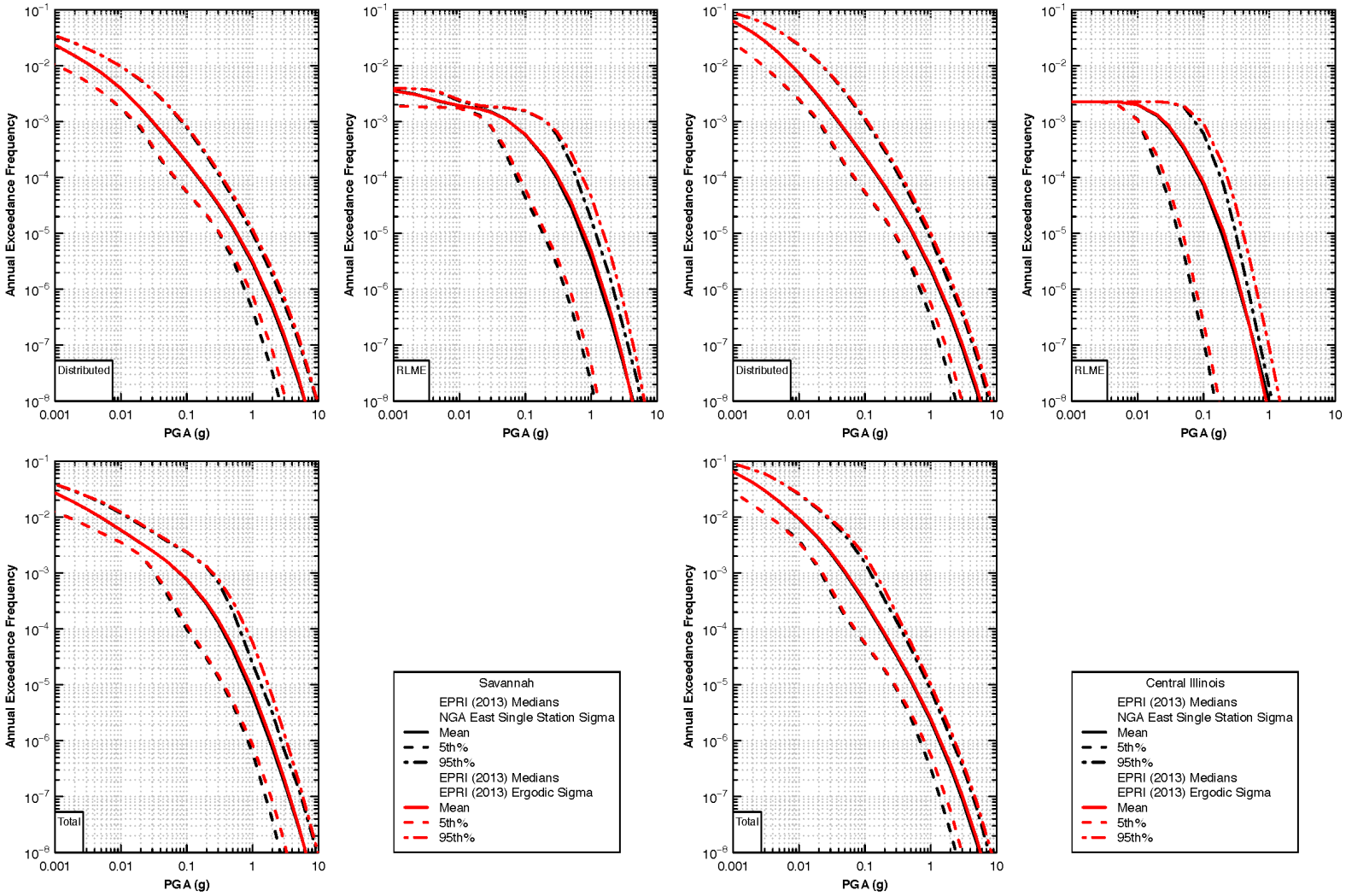


Figure 14–40(a) Ratio of mean hazard at Savannah computed using the NGA-East single-station sigma model to that computed using the EPRI (2013) ergodic sigma model.

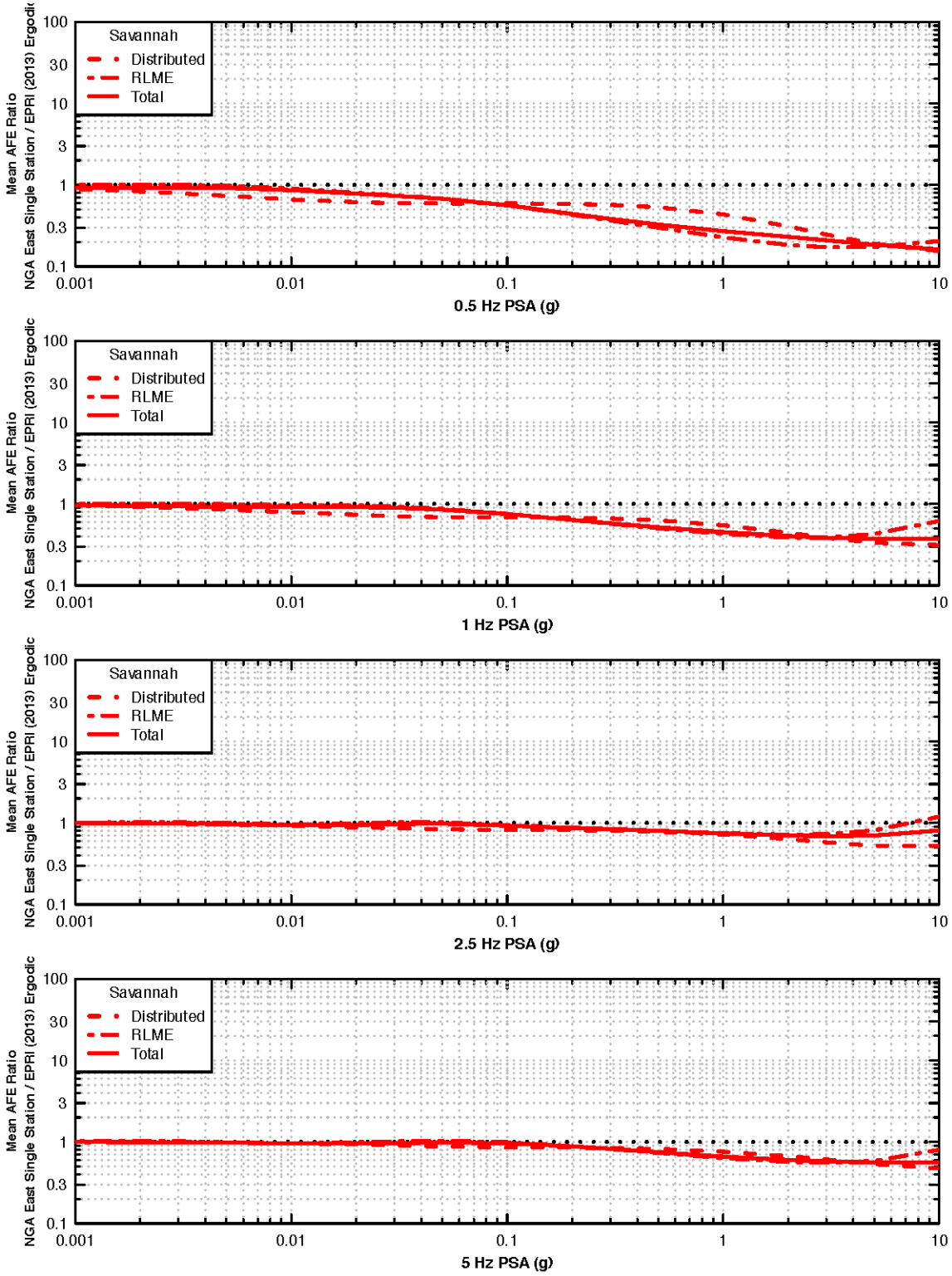


Figure 14–40(b) Ratio of mean hazard at Savannah computed using the NGA-East single-station sigma model to that computed using the EPRI (2013) ergodic sigma model.

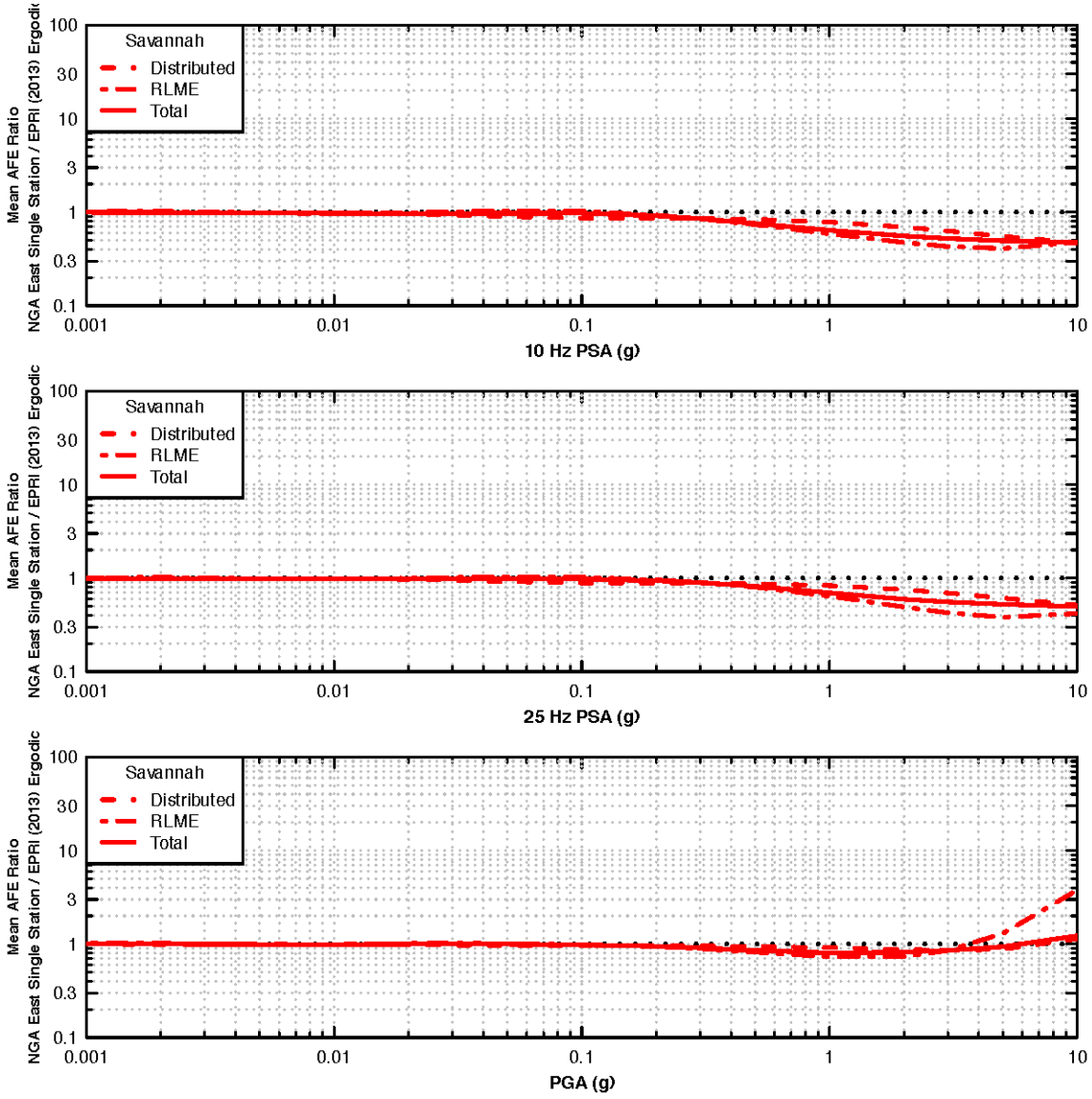


Figure 14–41(a) Ratio of mean hazard at Central Illinois computed using the NGA-East single-station sigma model to that computed using the EPRI (2013) ergodic sigma model.

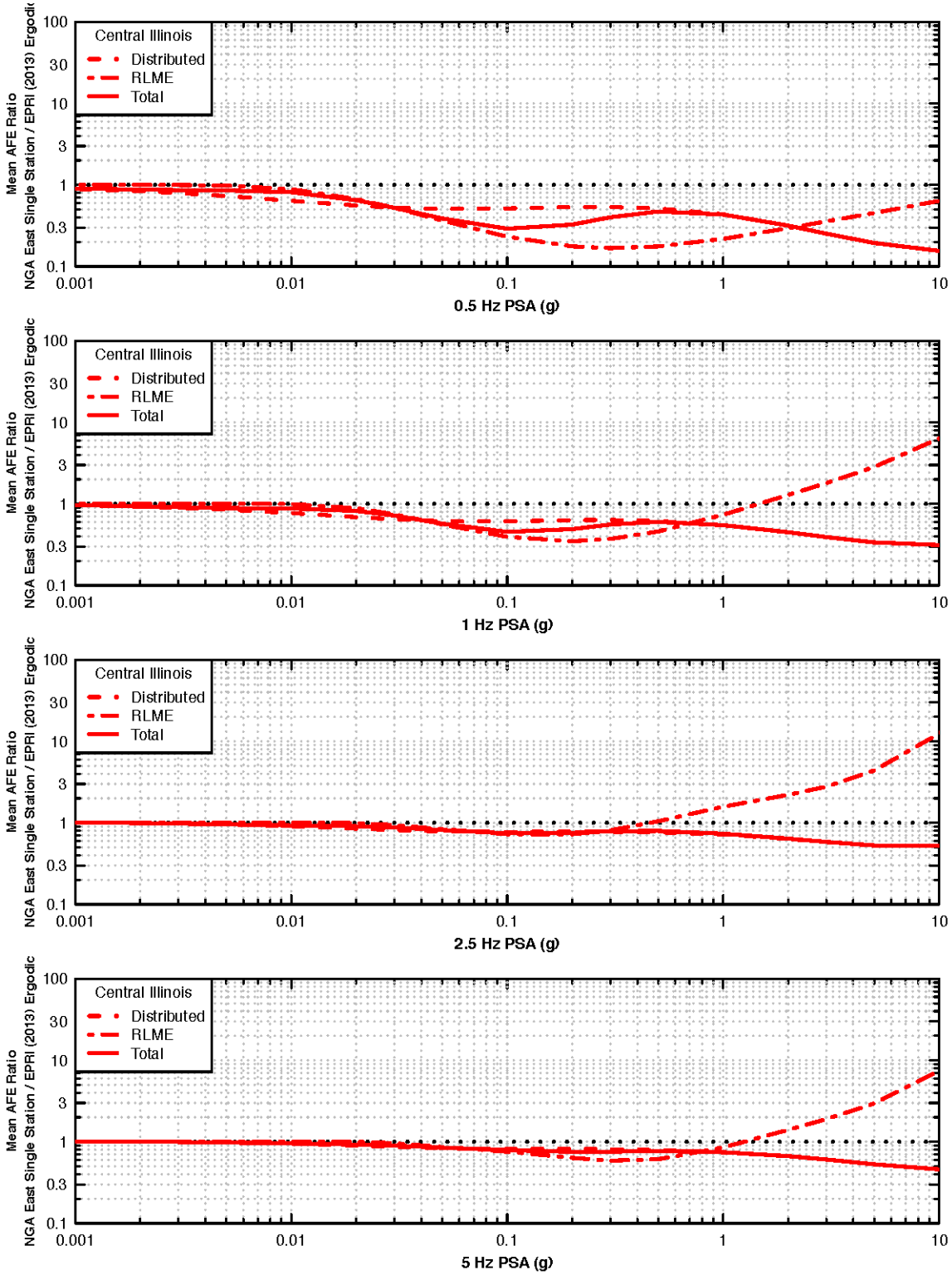


Figure 14–41(b) Ratio of mean hazard at Central Illinois computed using the NGA-East single-station sigma model to that computed using the EPRI (2013) ergodic sigma model.

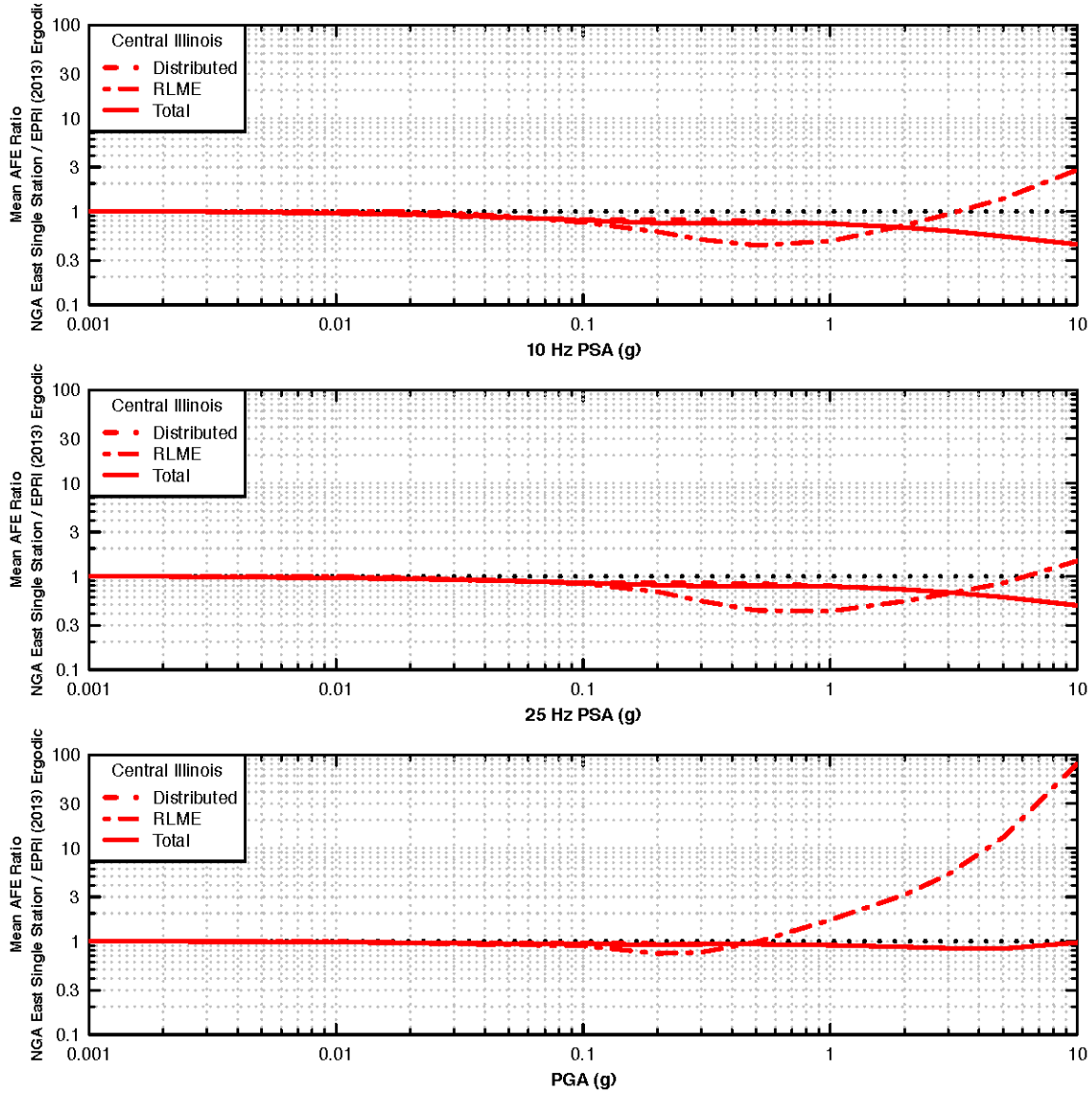


Figure 14–42 Comparison of 0.5 Hz PSA hazard at Savannah computed using the complete NGA-East and EPRI (2013) GMMs.

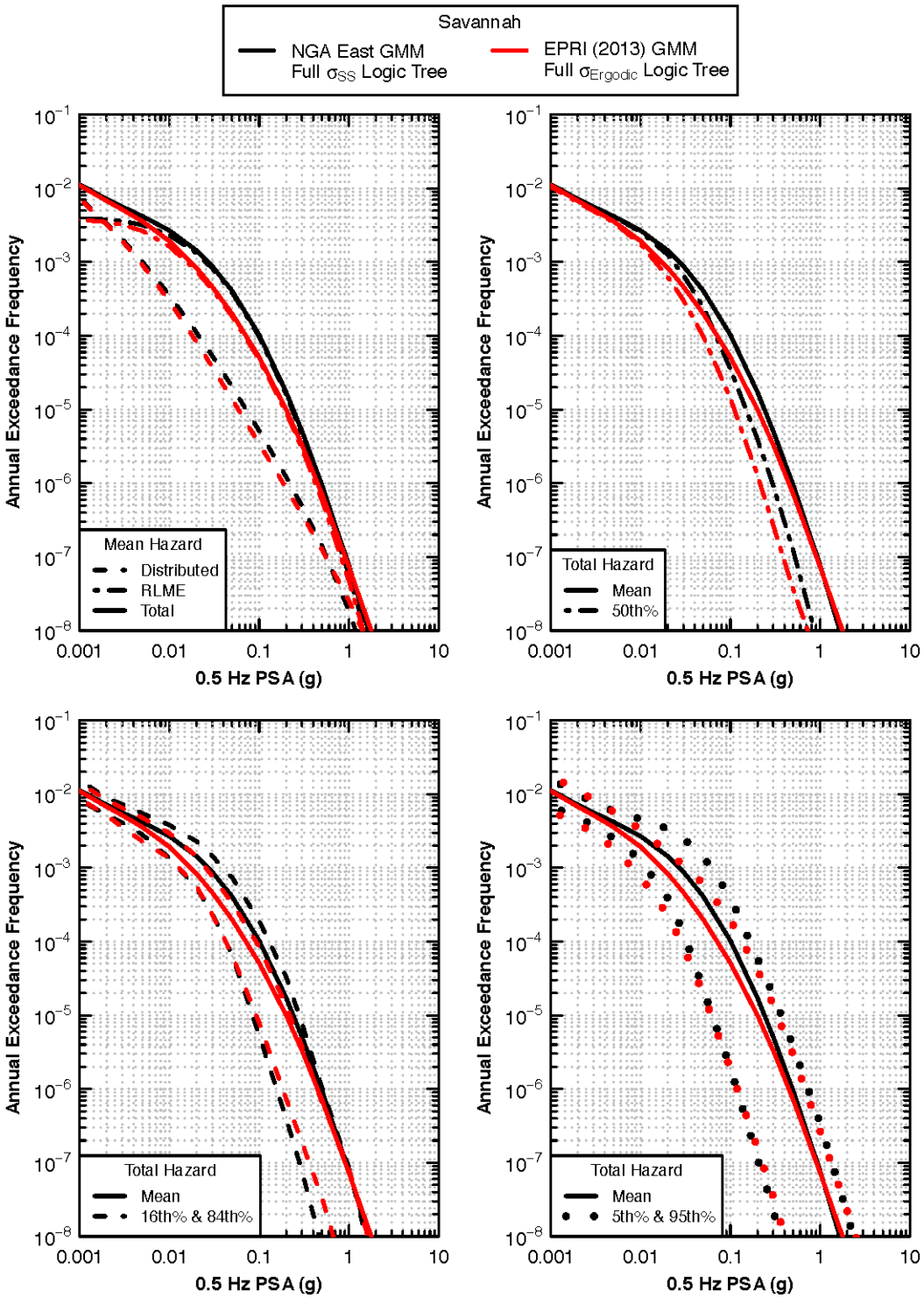


Figure 14–43 Comparison of 1 Hz PSA hazard at Savannah computed using the complete NGA-East and EPRI (2013) GMMs.

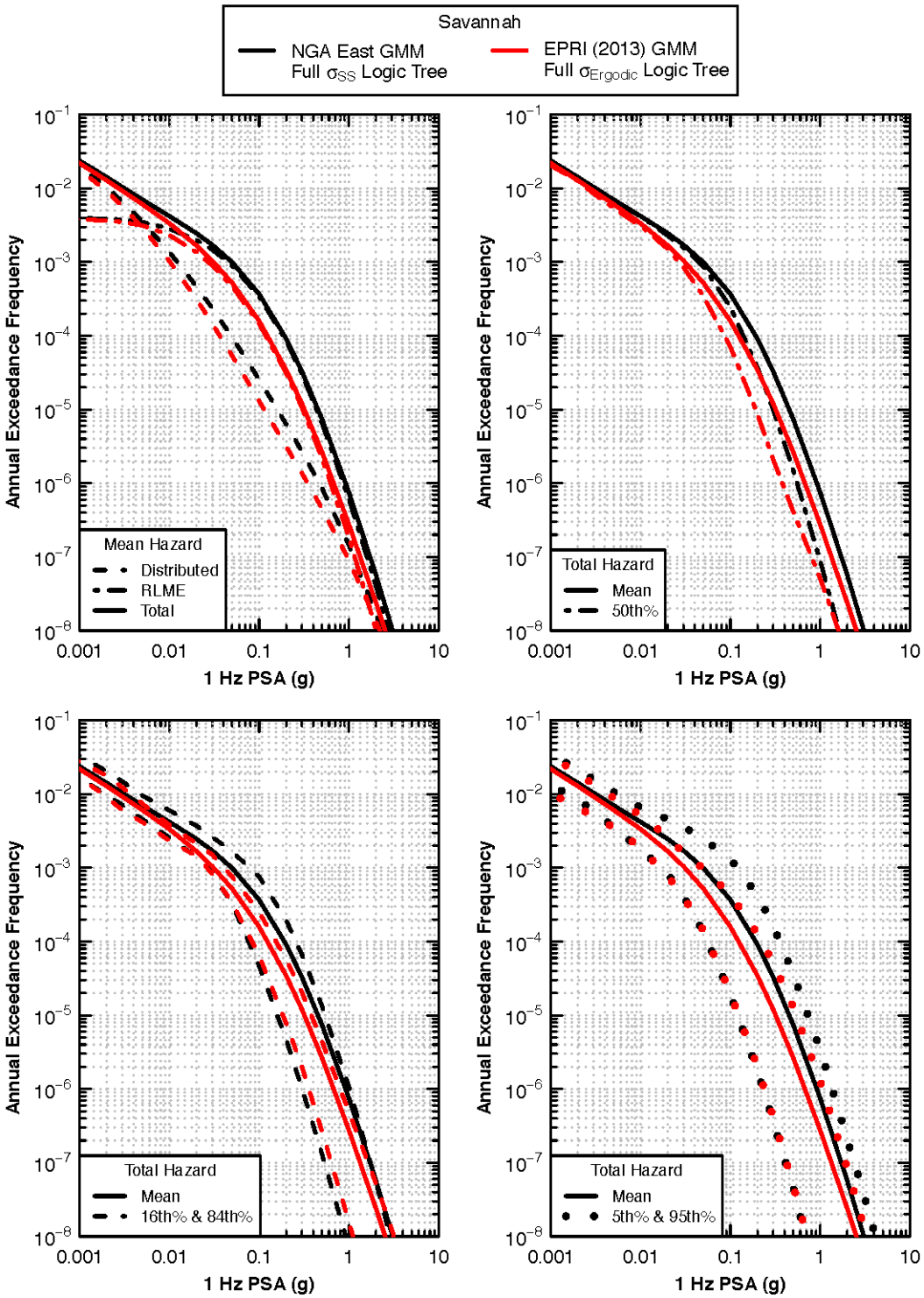


Figure 14–44 Comparison of 2.5 Hz PSA hazard at Savannah computed using the complete NGA-East and EPRI (2013) GMMs.

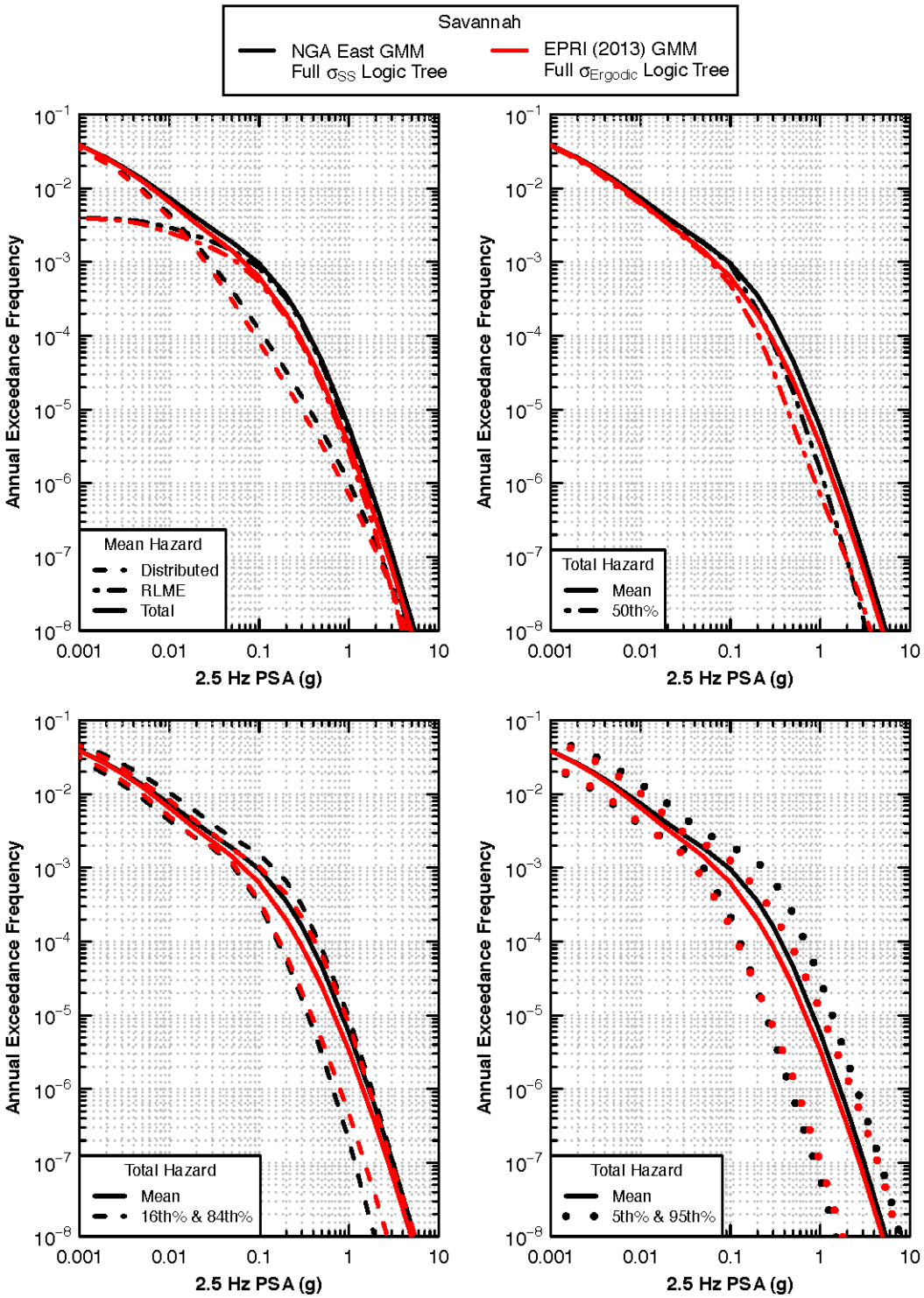


Figure 14–45 Comparison of 5 Hz PSA hazard at Savannah computed using the complete NGA-East and EPRI (2013) GMMs.

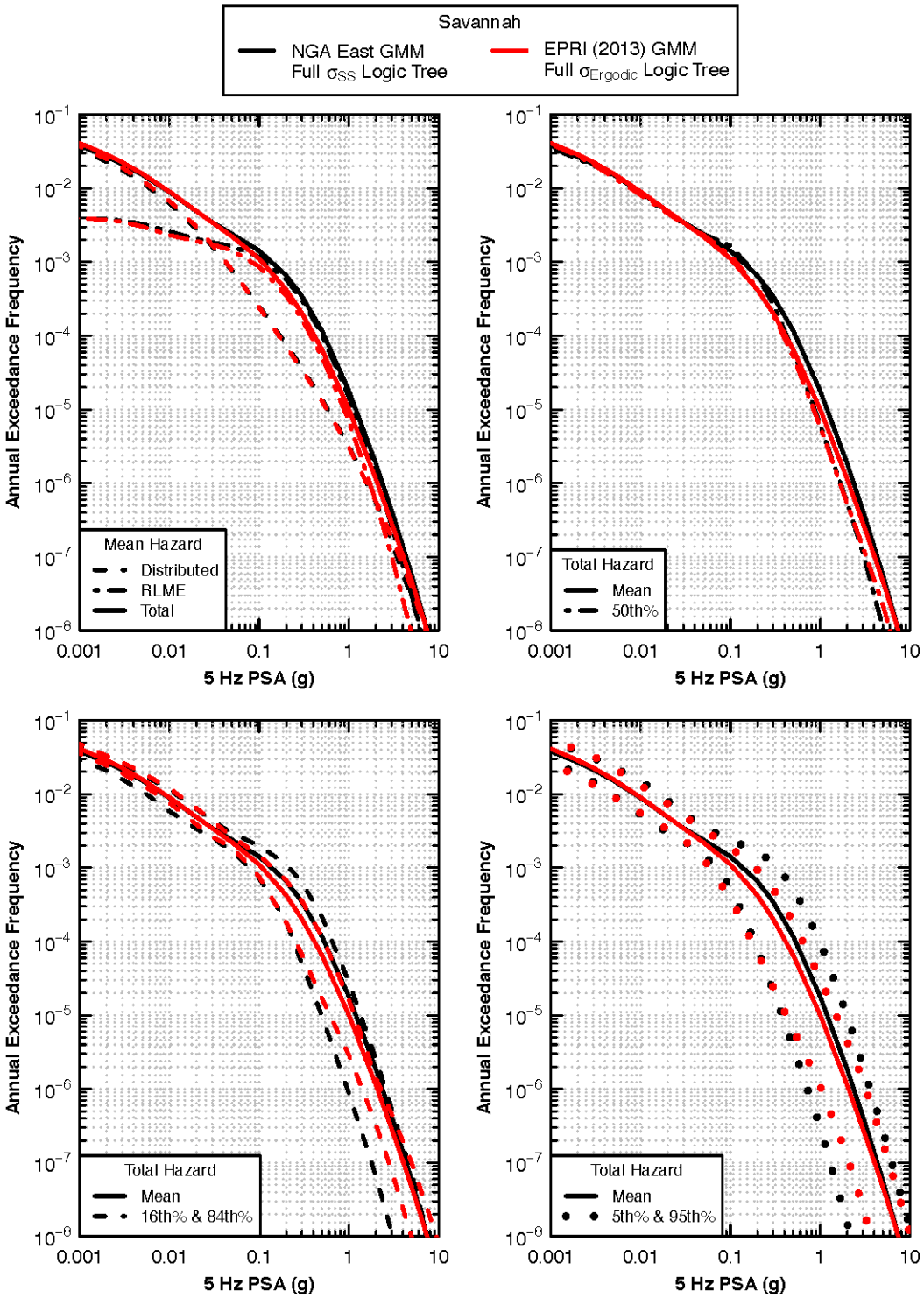


Figure 14–46 Comparison of 10 Hz PSA hazard at Savannah computed using the complete NGA-East and EPRI (2013) GMMs.

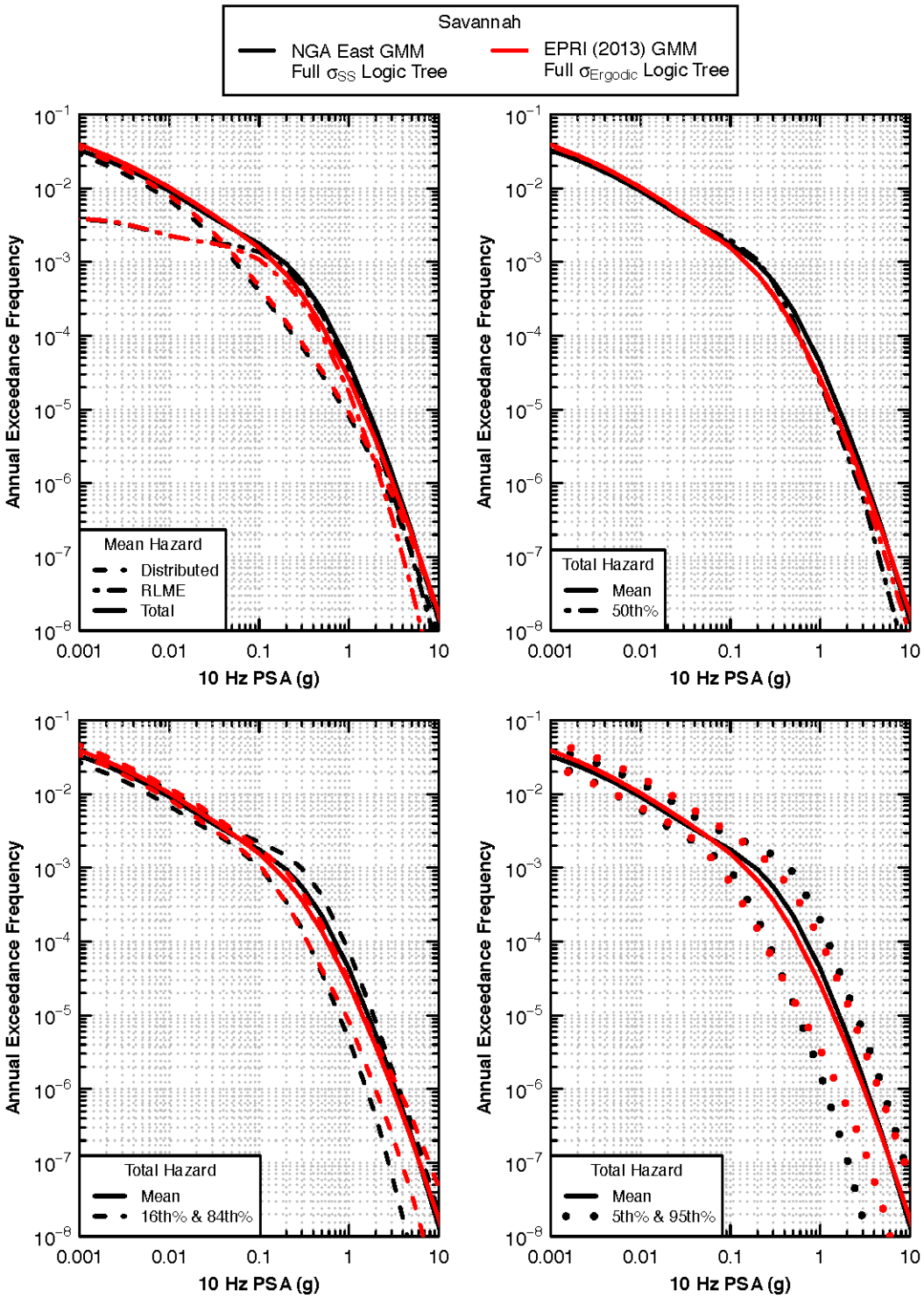


Figure 14–47 Comparison of 25 Hz PSA hazard at Savannah computed using the complete NGA-East and EPRI (2013) GMMs.

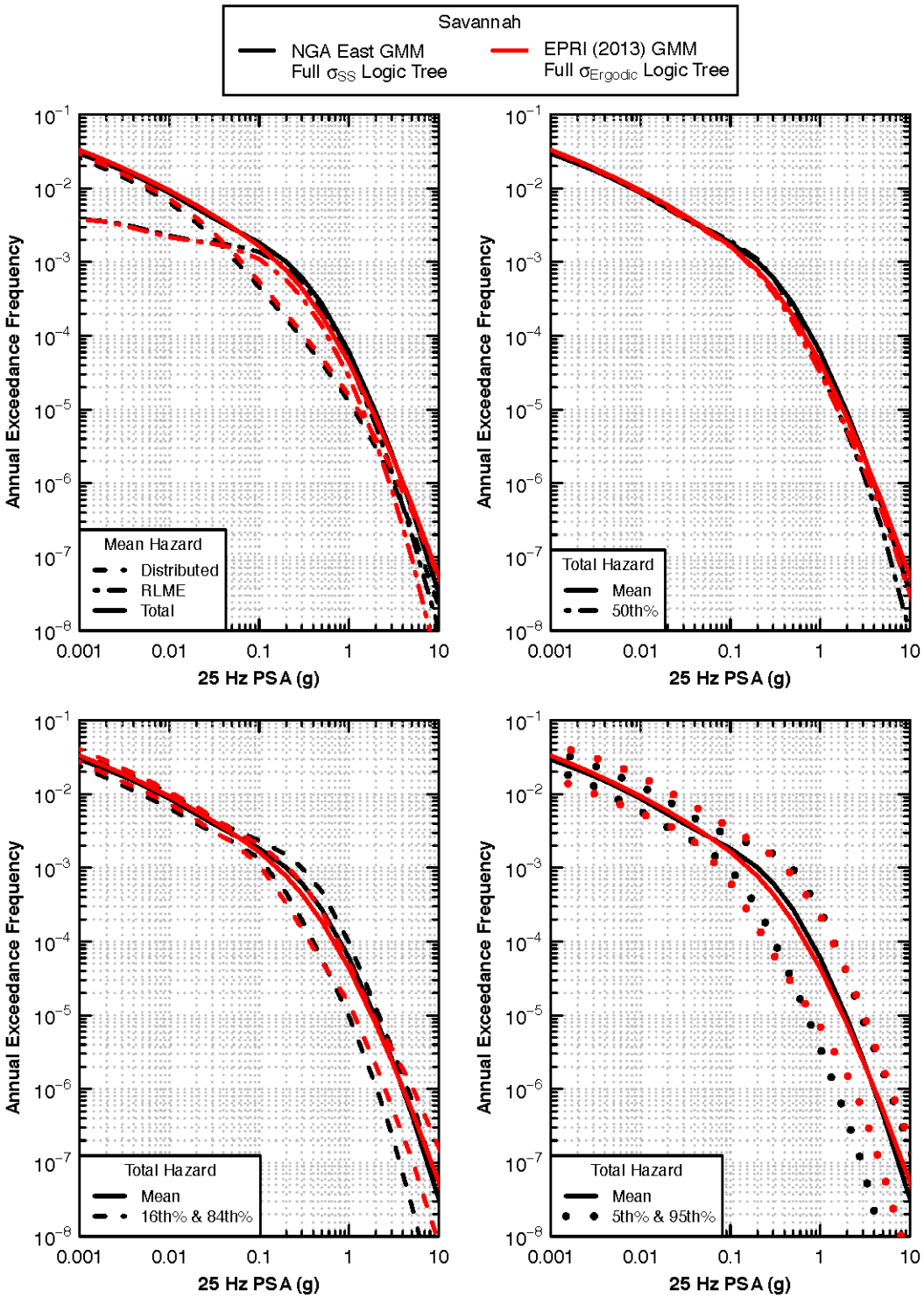


Figure 14–48 Comparison of PGA hazard at Savannah computed using the complete NGA-East and EPRI (2013) GMMs.

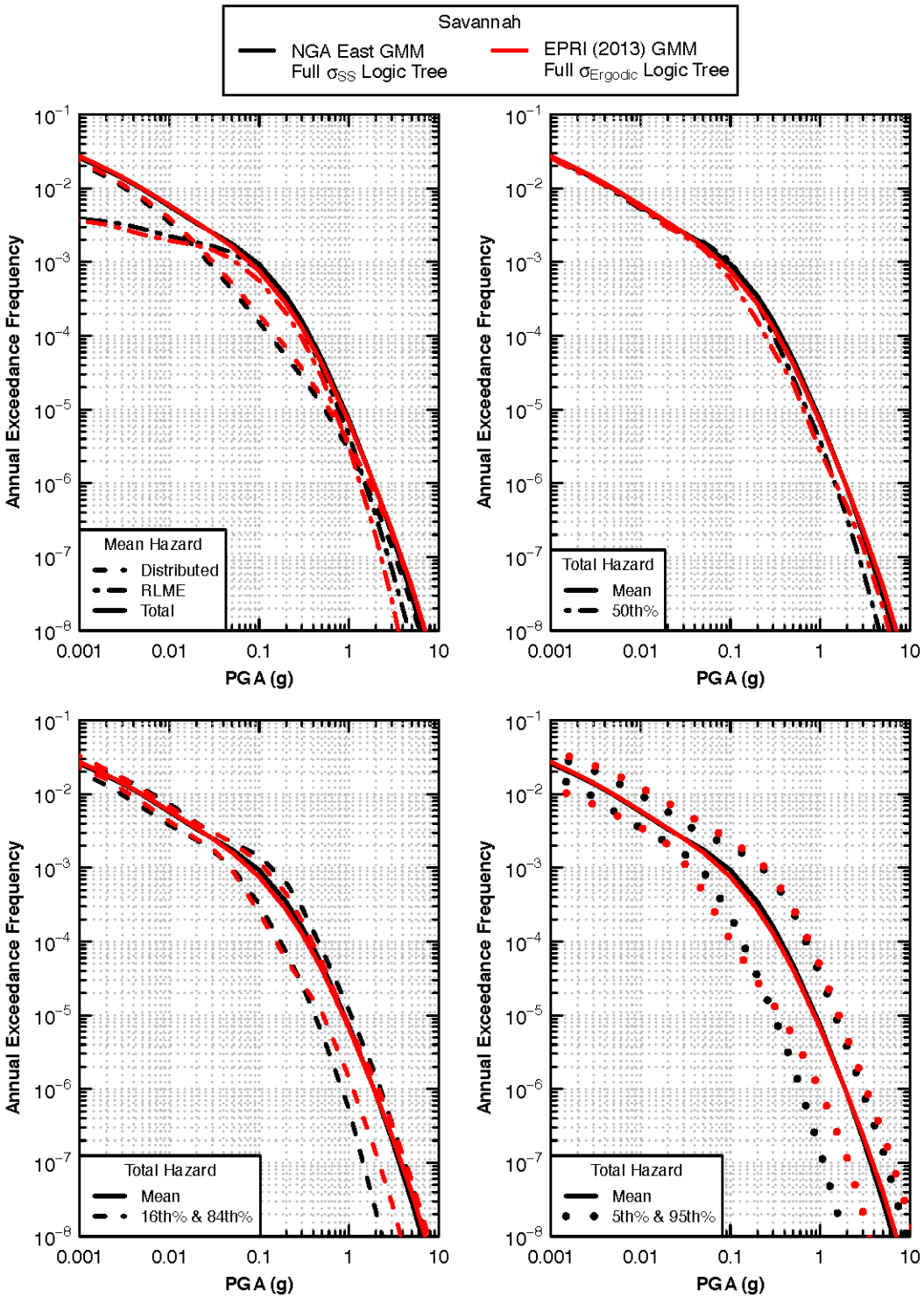


Figure 14–49(a) Ratio of mean hazard at Savannah computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

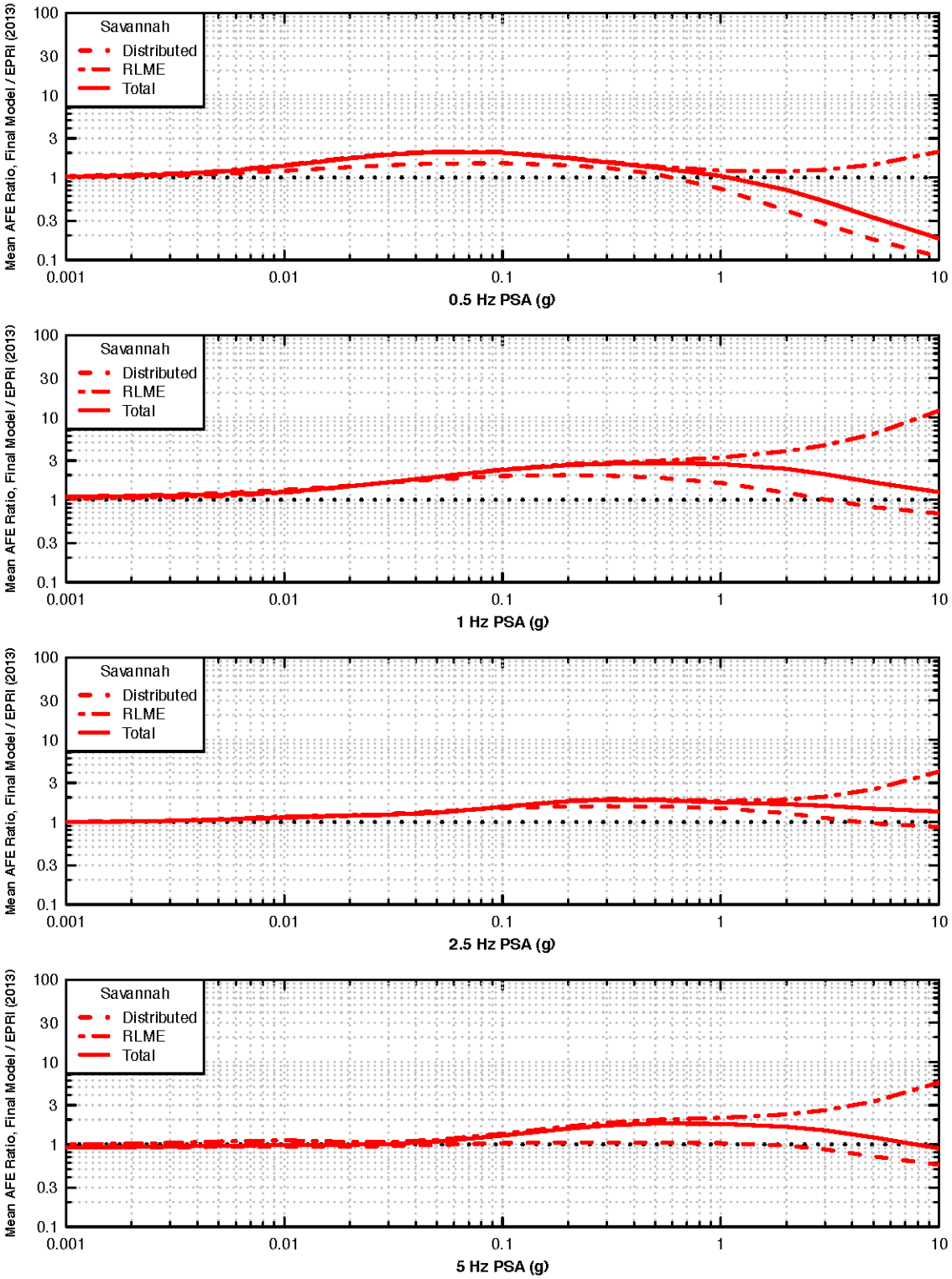


Figure 14–49(b) Ratio of mean hazard at Savannah computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

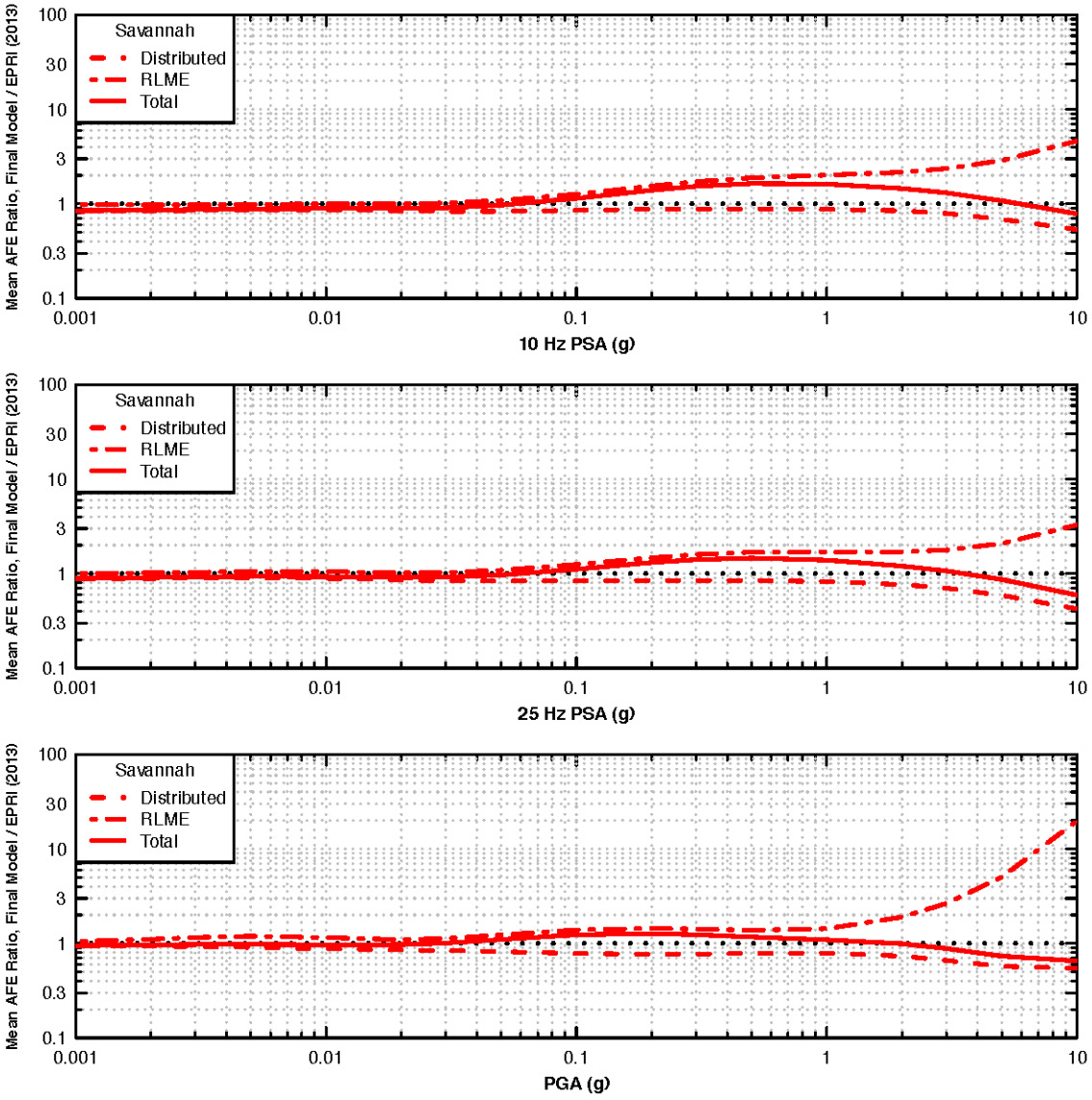


Figure 14–50 Comparison of 0.5 Hz PSA hazard at Central Illinois computed using the complete NGA-East and EPRI (2013) GMMs.

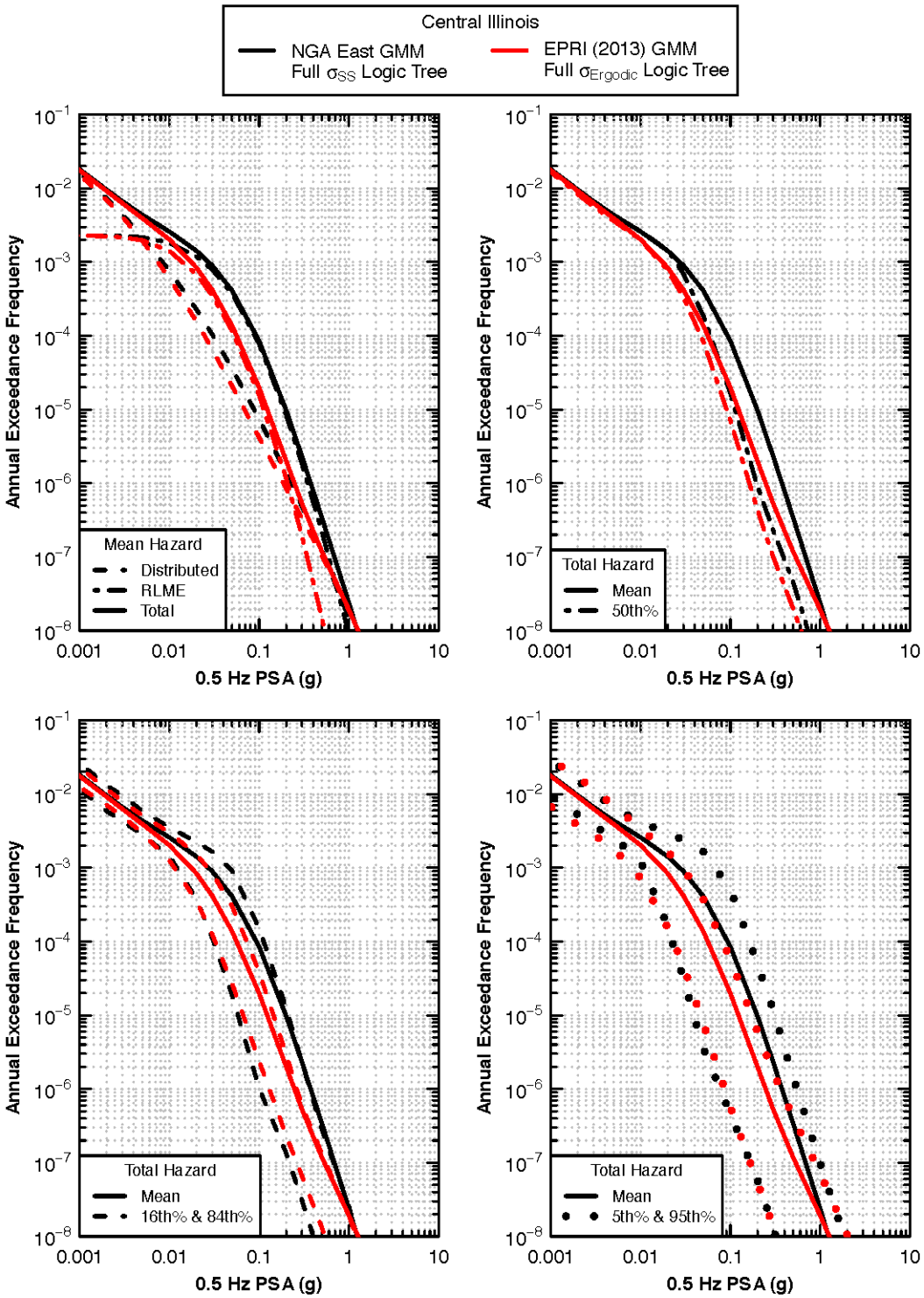


Figure 14–51 Comparison of 1 Hz PSA hazard at Central Illinois computed using the complete NGA-East and EPRI (2013) GMMs.

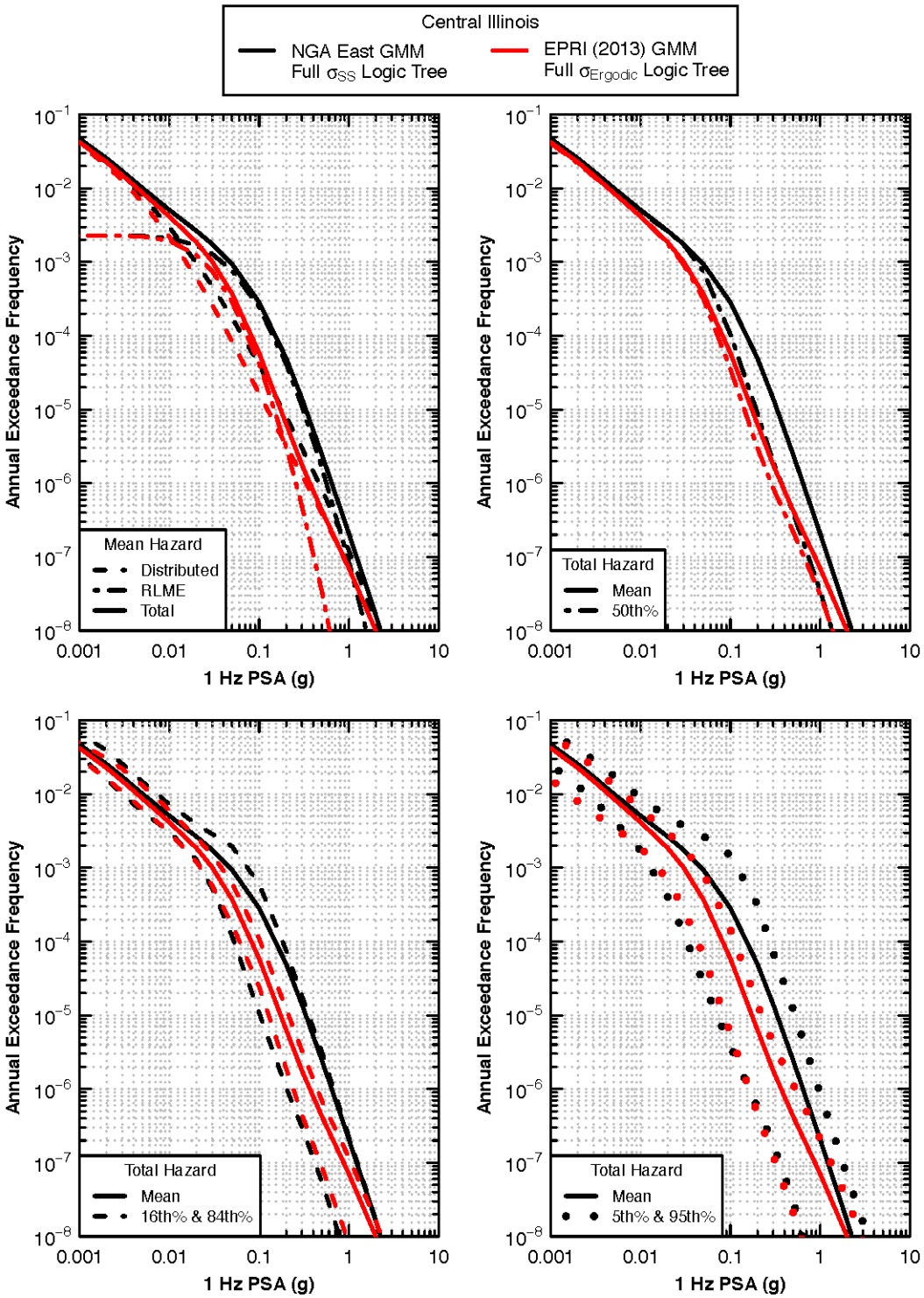


Figure 14–52 Comparison of 2.5 Hz PSA hazard at Central Illinois computed using the complete NGA-East and EPRI (2013) GMMs.

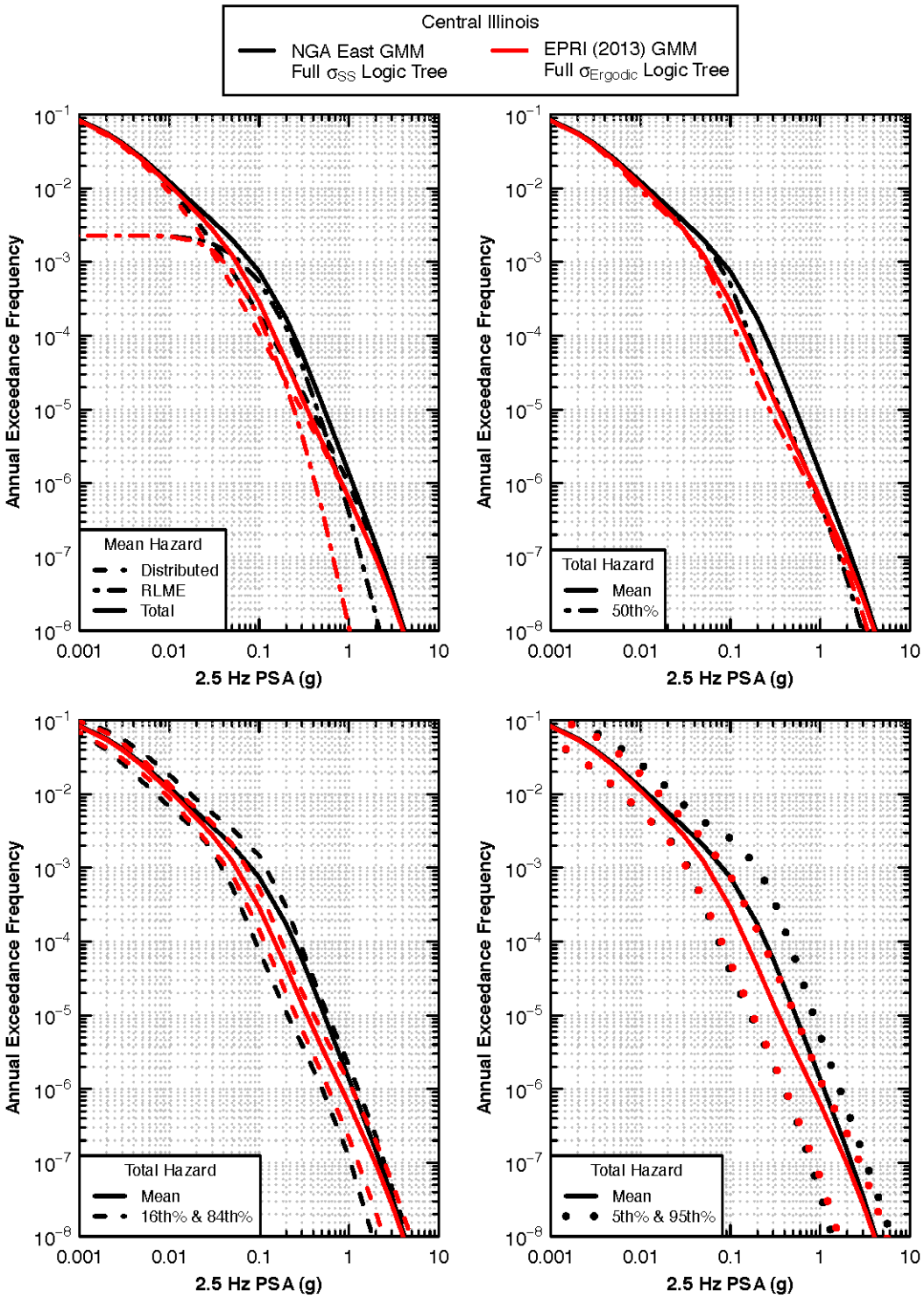


Figure 14–53 Comparison of 5 Hz PSA hazard at Central Illinois computed using the complete NGA-East and EPRI (2013) GMMs.

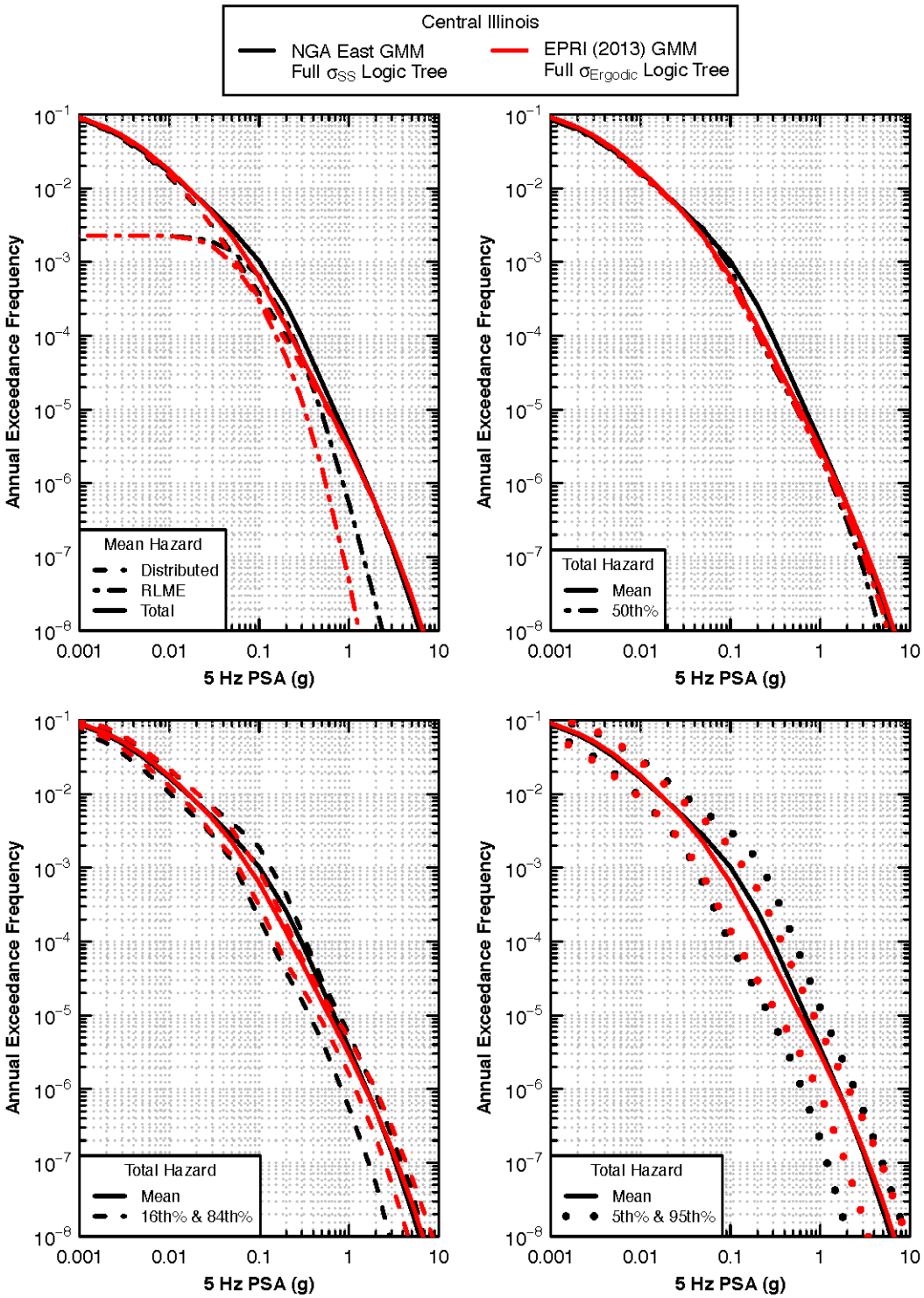


Figure 14–54 Comparison of 10 Hz PSA hazard at Central Illinois computed using the complete NGA-East and EPRI (2013) GMMs.

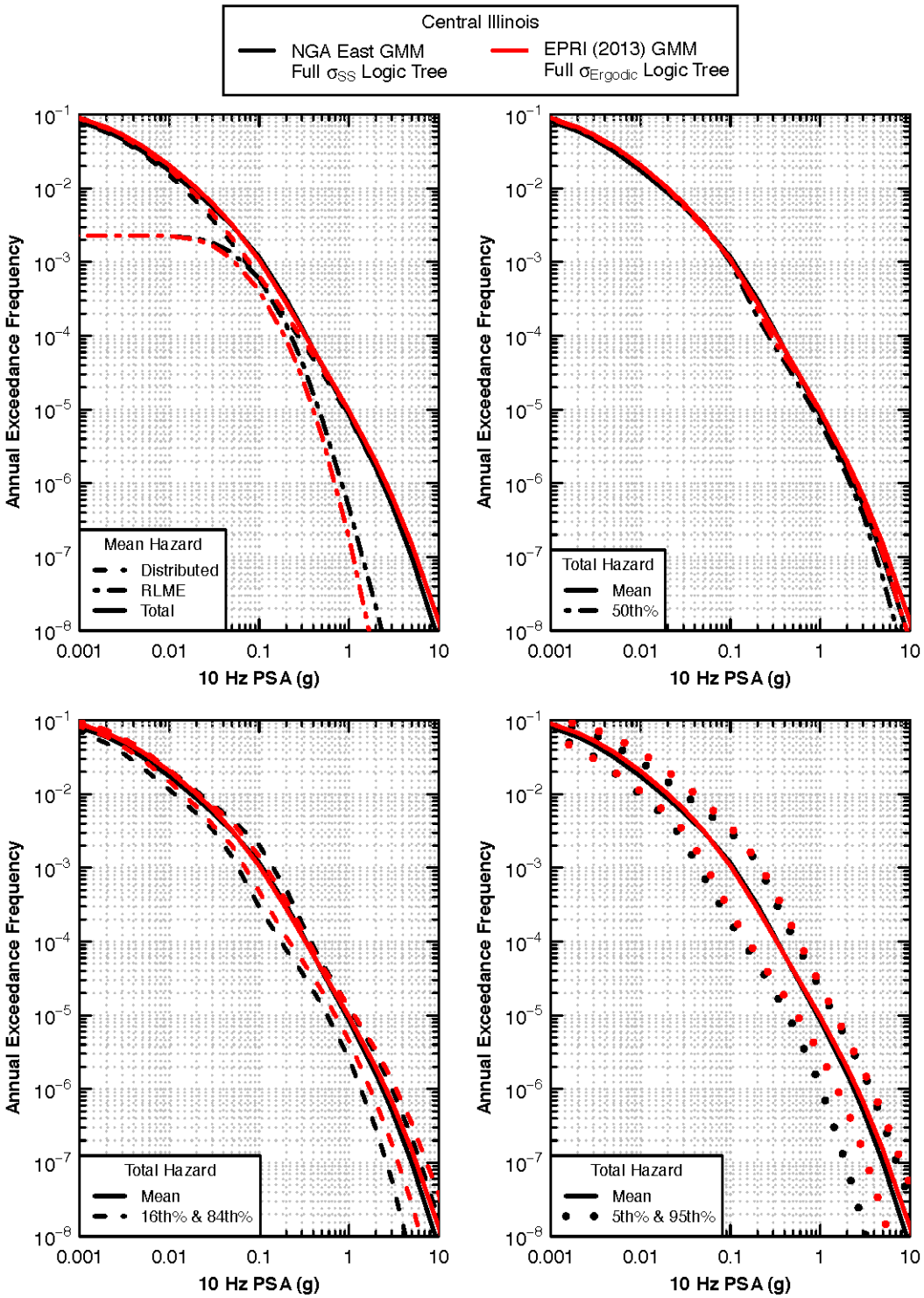


Figure 14–55 Comparison of 25 Hz PSA hazard at Central Illinois computed using the complete NGA-East and EPRI (2013) GMMs.

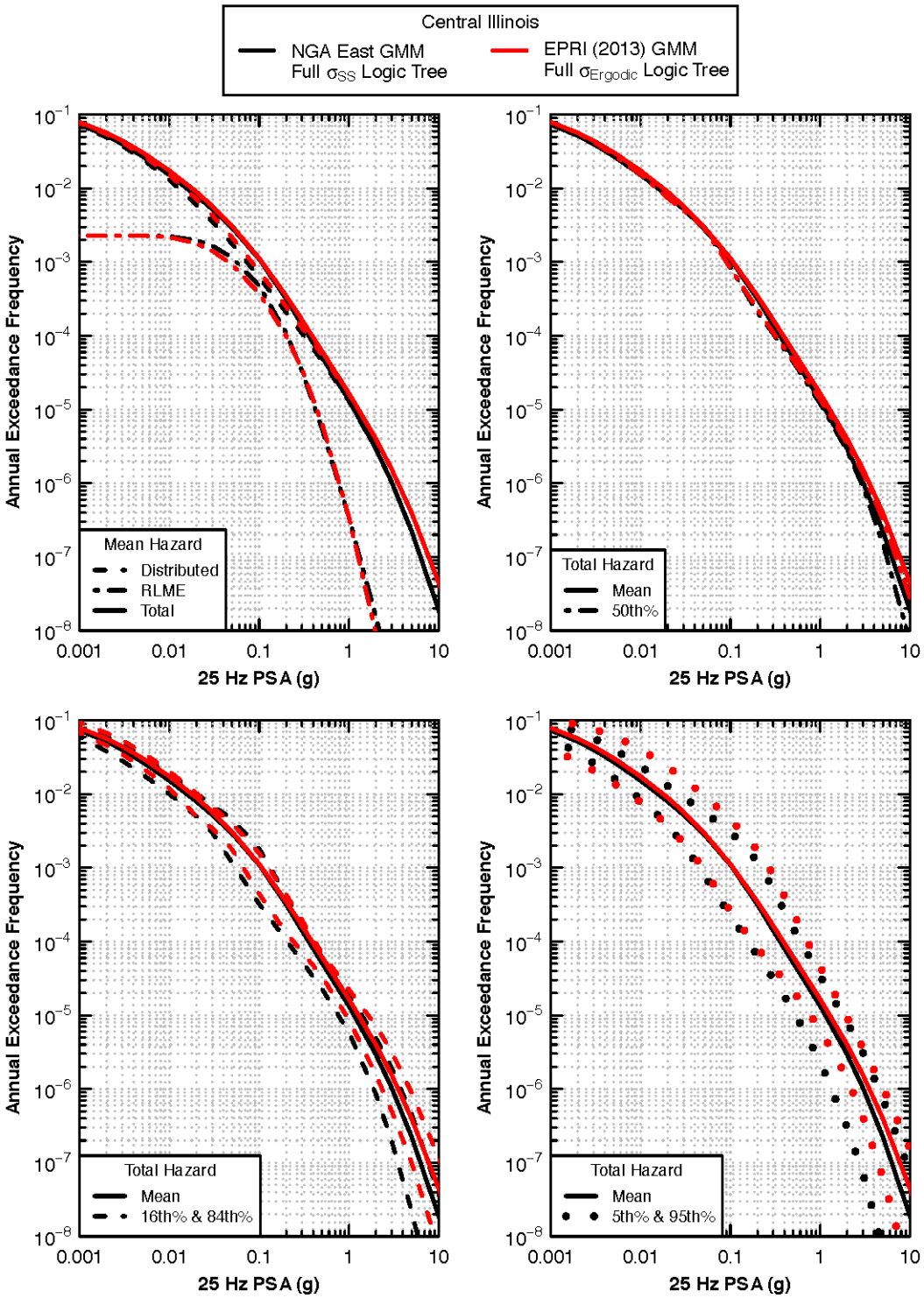


Figure 14–56 Comparison of PGA hazard at Central Illinois computed using the complete NGA-East and EPRI (2013) GMMs.

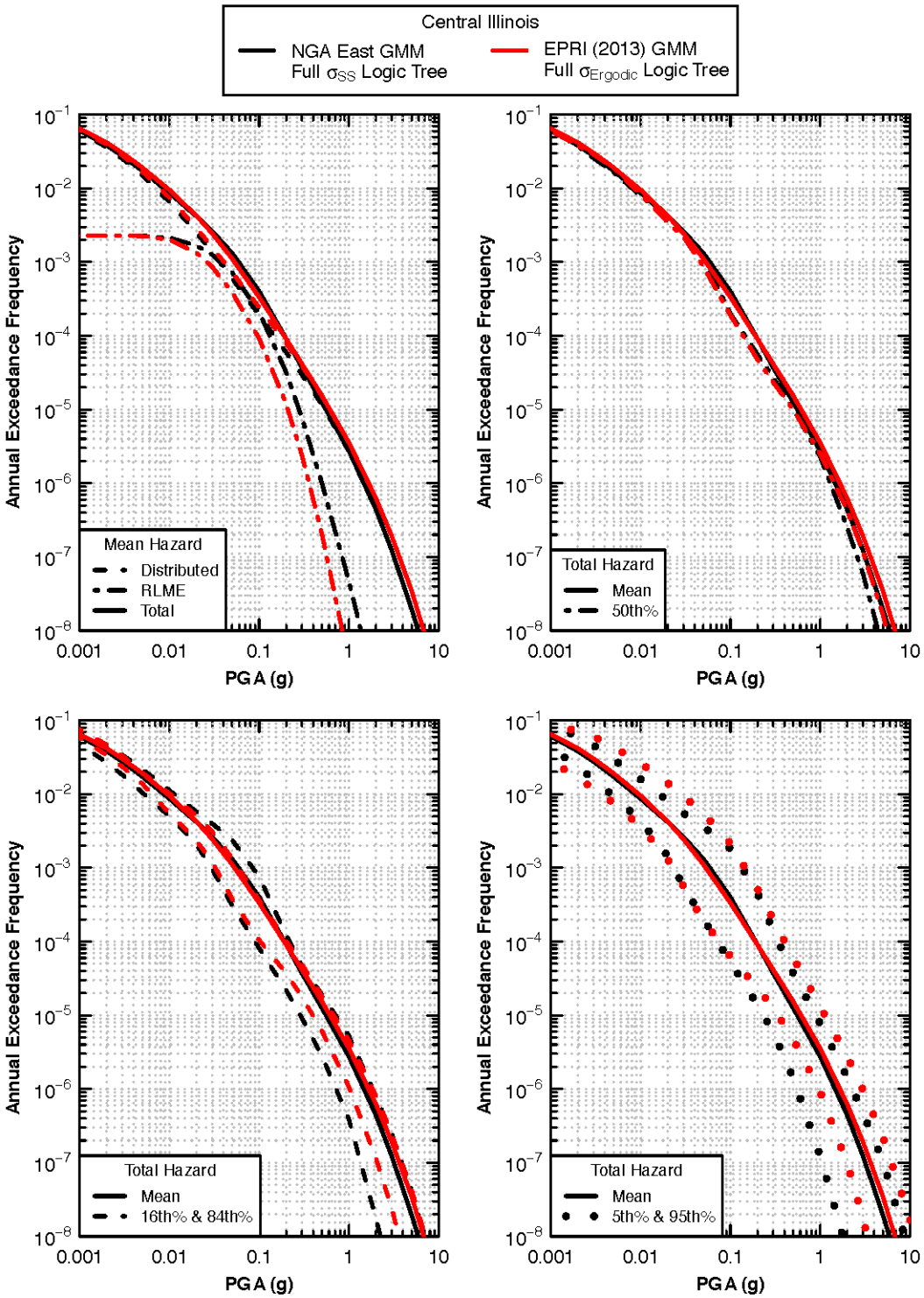


Figure 14–57(a) Ratio of mean hazard at Central Illinois computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

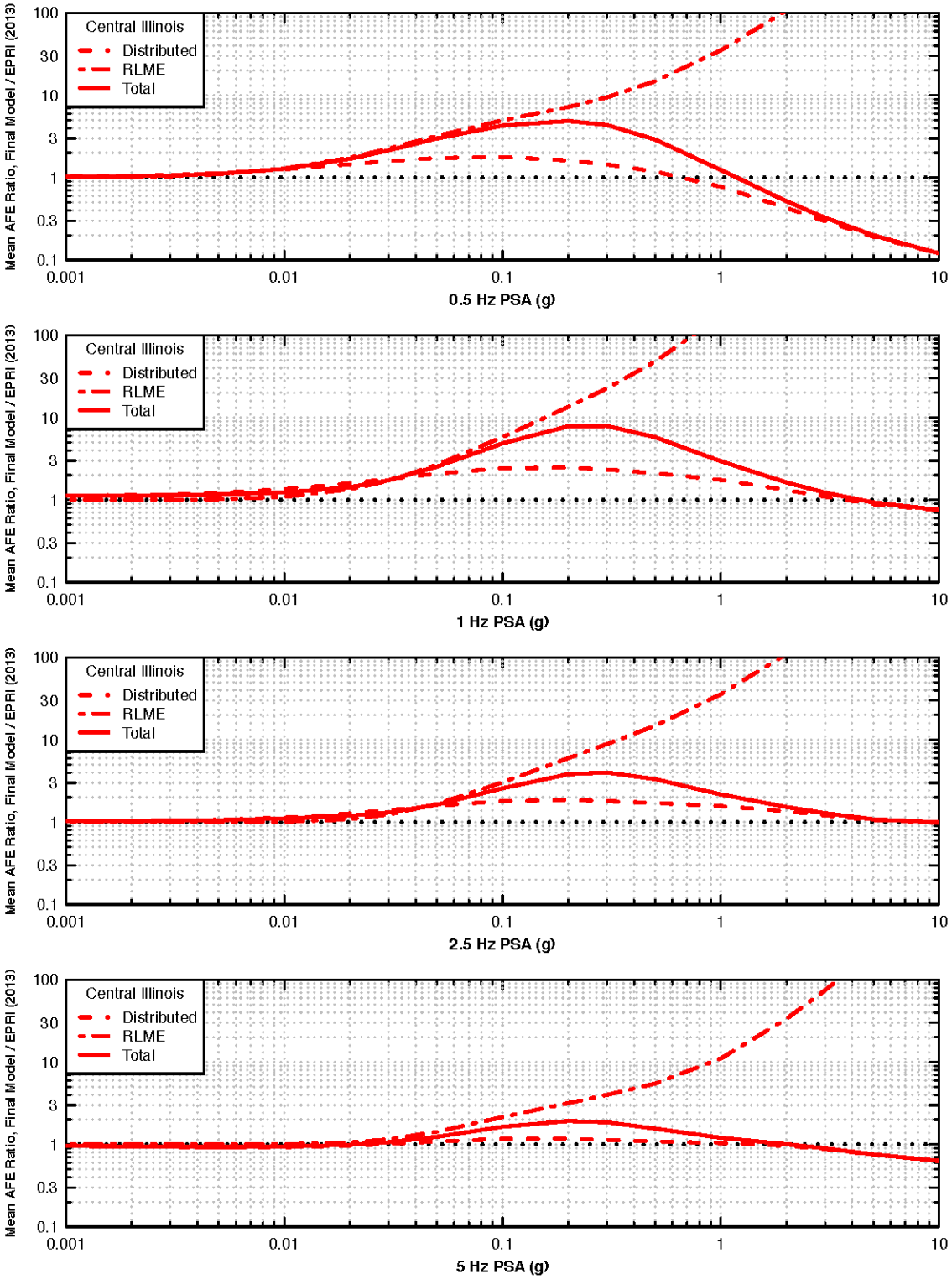


Figure 14–57(b) Ratio of mean hazard at Central Illinois computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

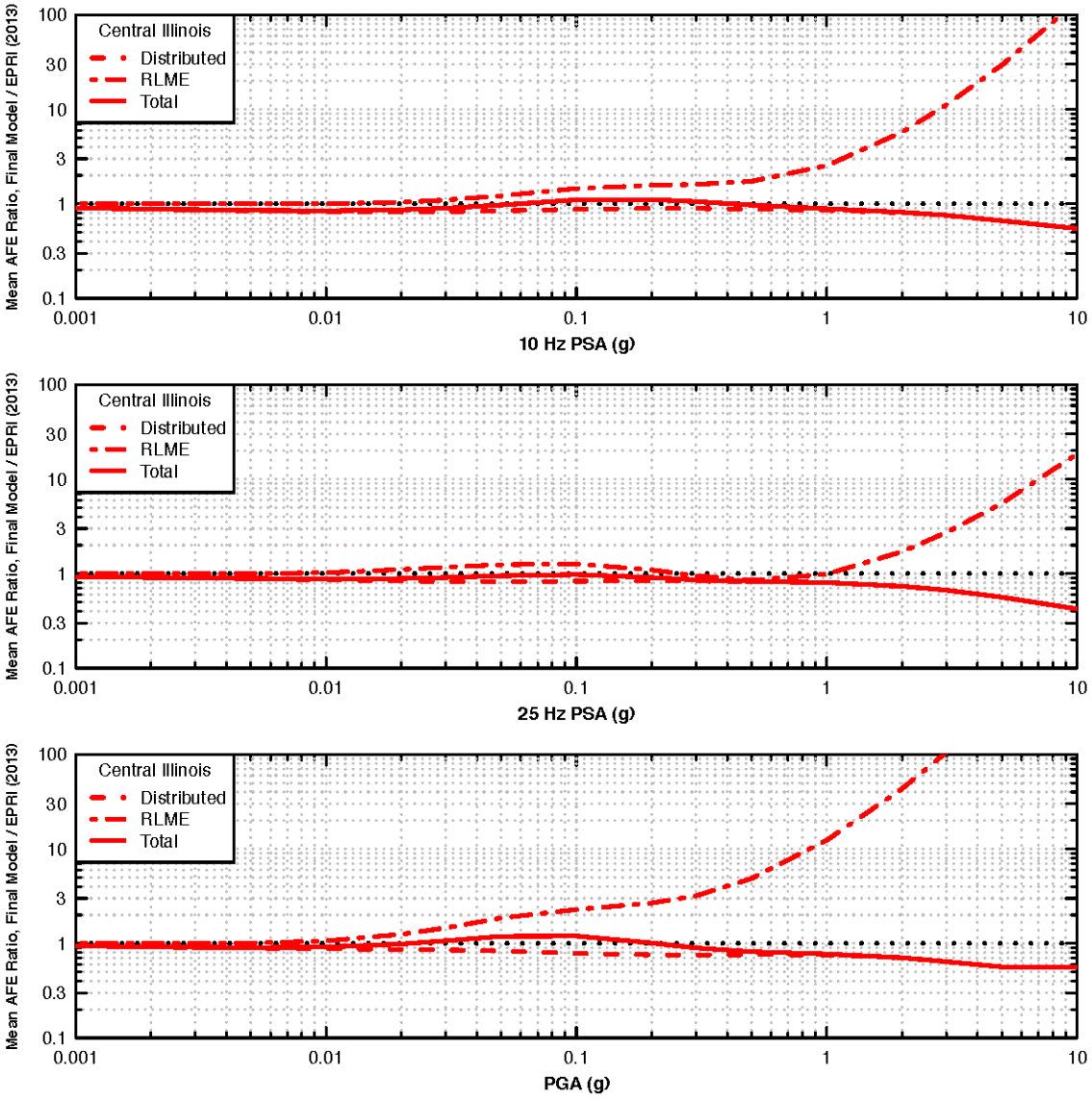


Figure 14–58 Comparison of 0.5 Hz PSA hazard at Manchester computed using the complete NGA-East and EPRI (2013) GMMs.

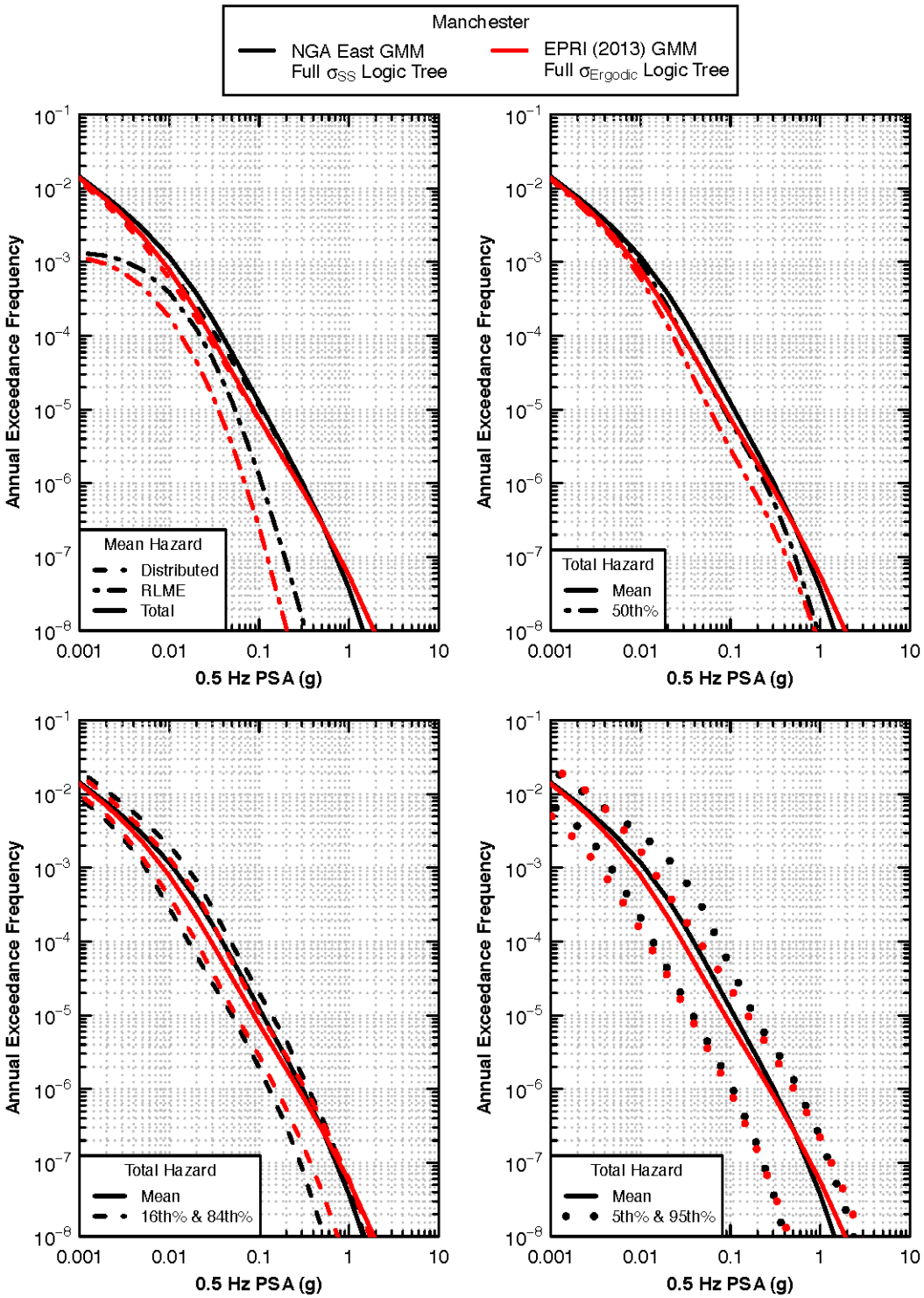


Figure 14–59 Comparison of 1 Hz PSA hazard at Manchester computed using the complete NGA-East and EPRI (2013) GMMs.

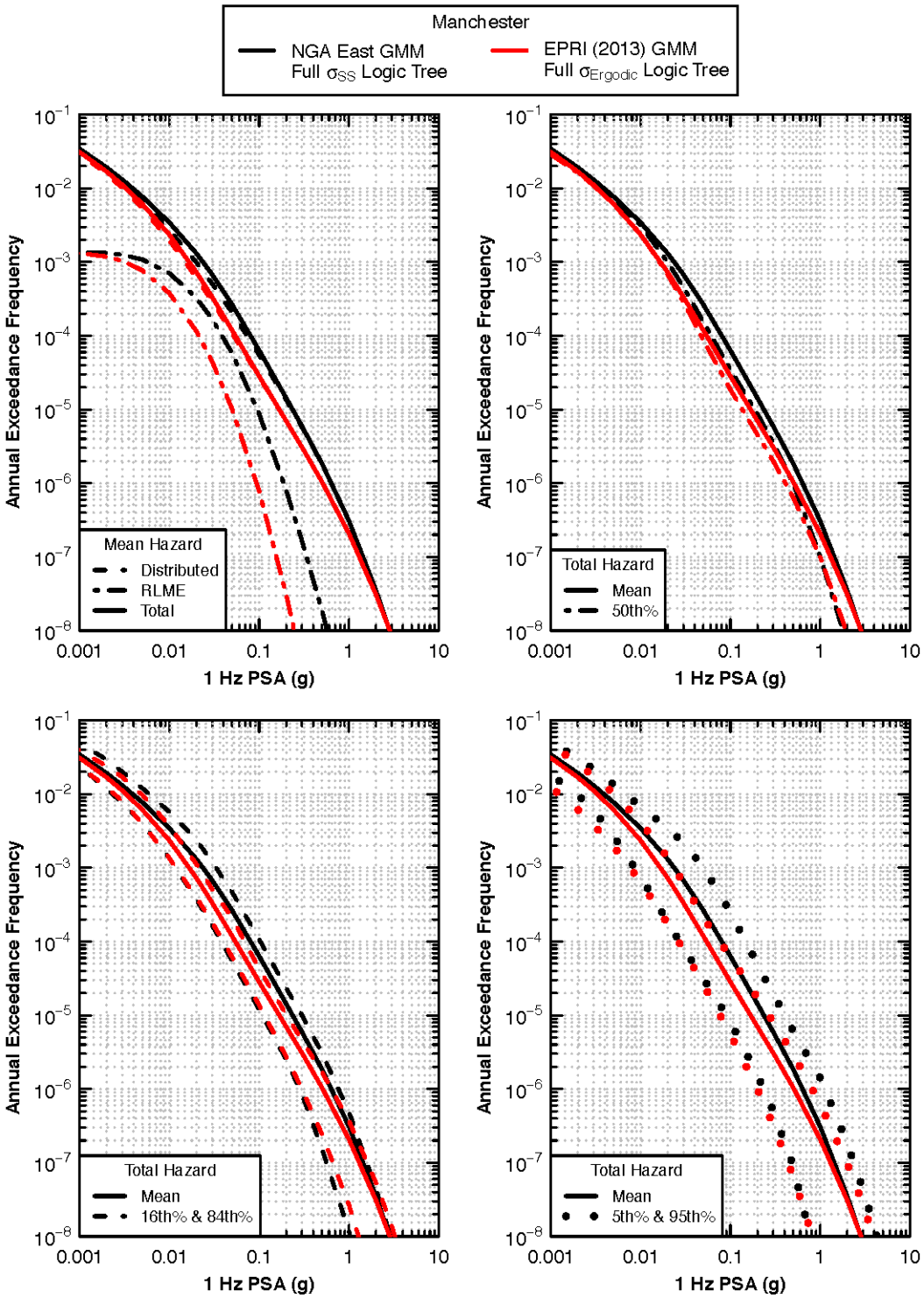


Figure 14–60 Comparison of 2.5 Hz PSA hazard at Manchester computed using the complete NGA-East and EPRI (2013) GMMs.

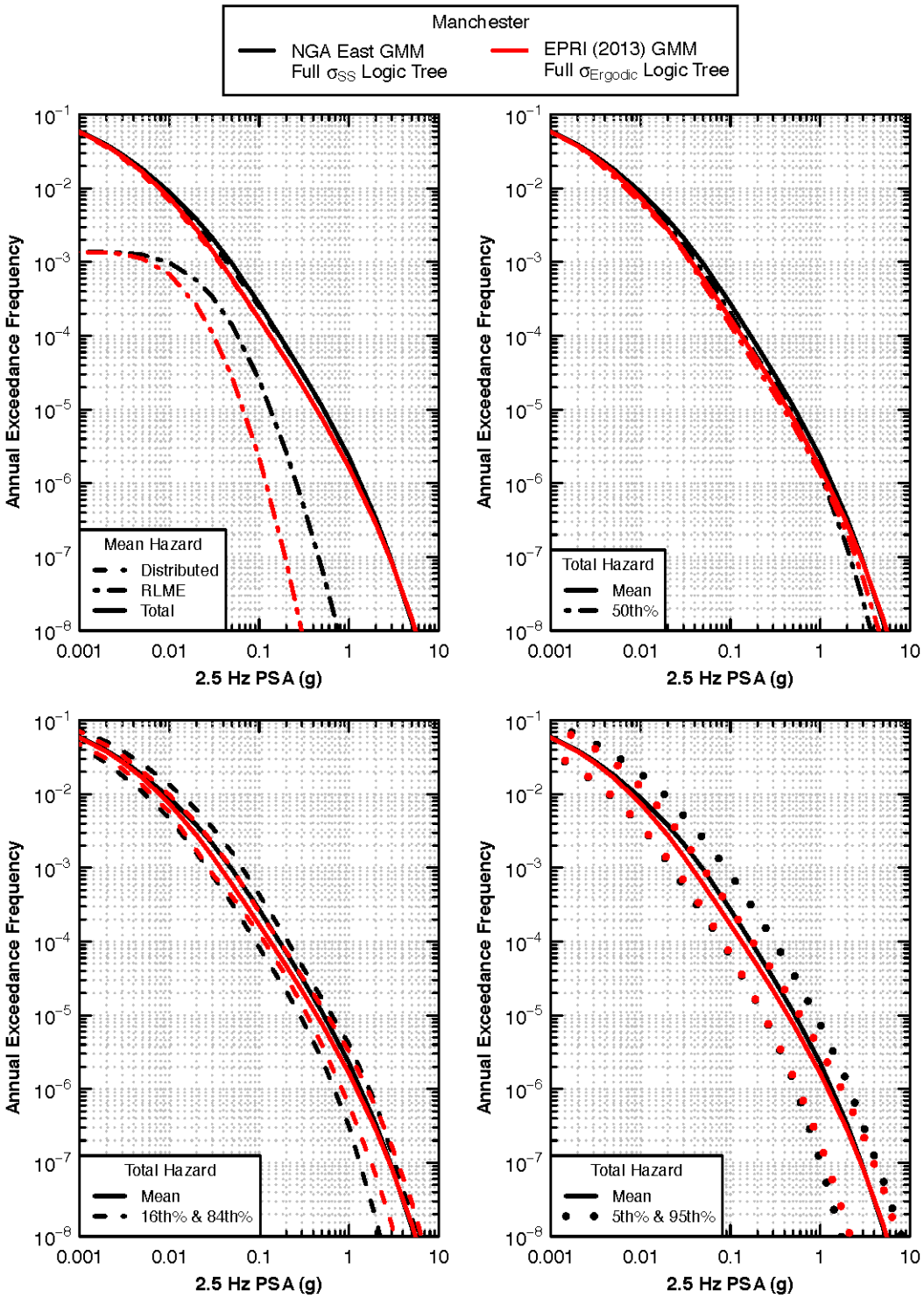


Figure 14–61 Comparison of 5 Hz PSA hazard at Manchester computed using the complete NGA-East and EPRI (2013) GMMs.

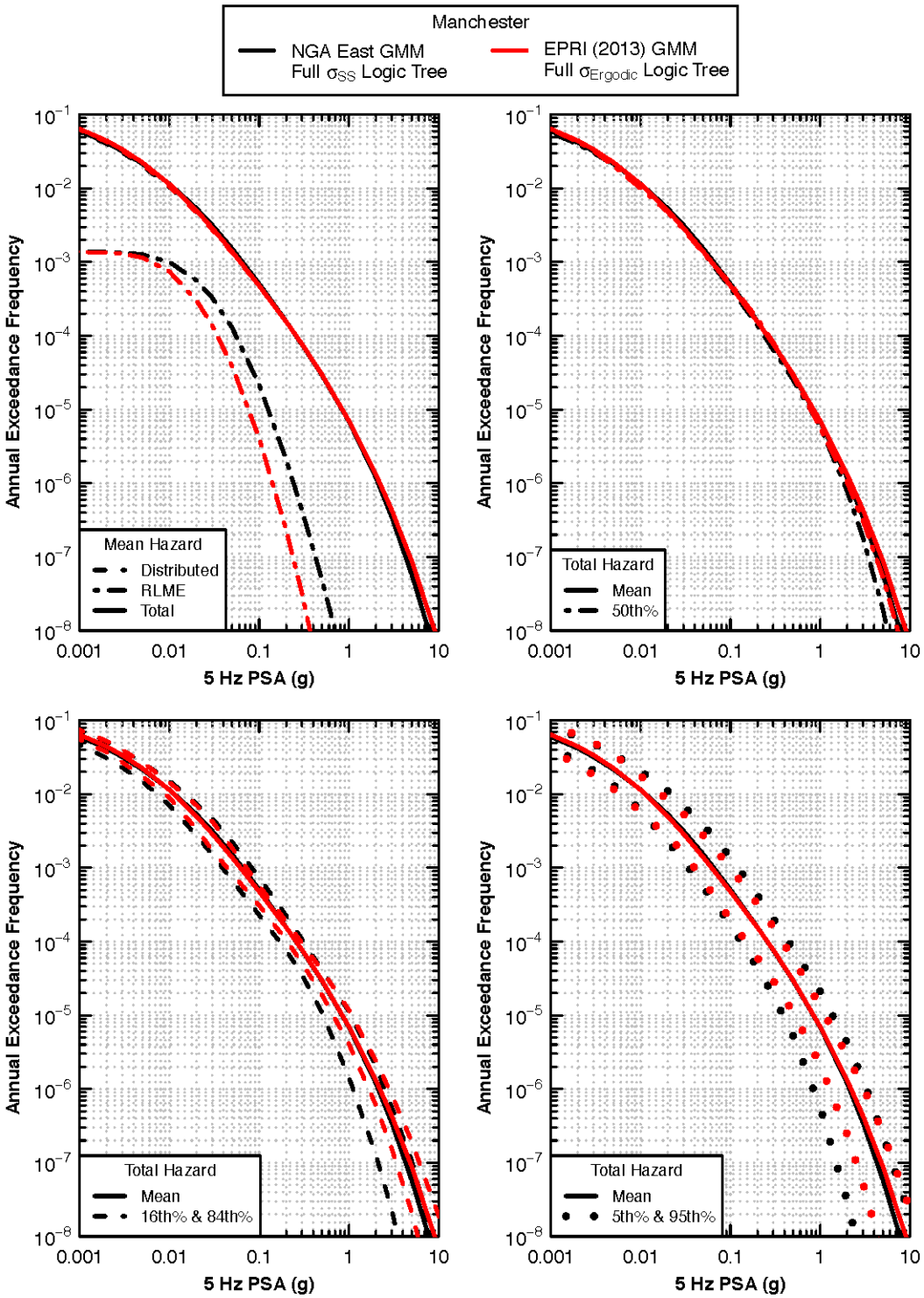


Figure 14–62 Comparison of 10 Hz PSA hazard at Manchester computed using the complete NGA-East and EPRI (2013) GMMs.

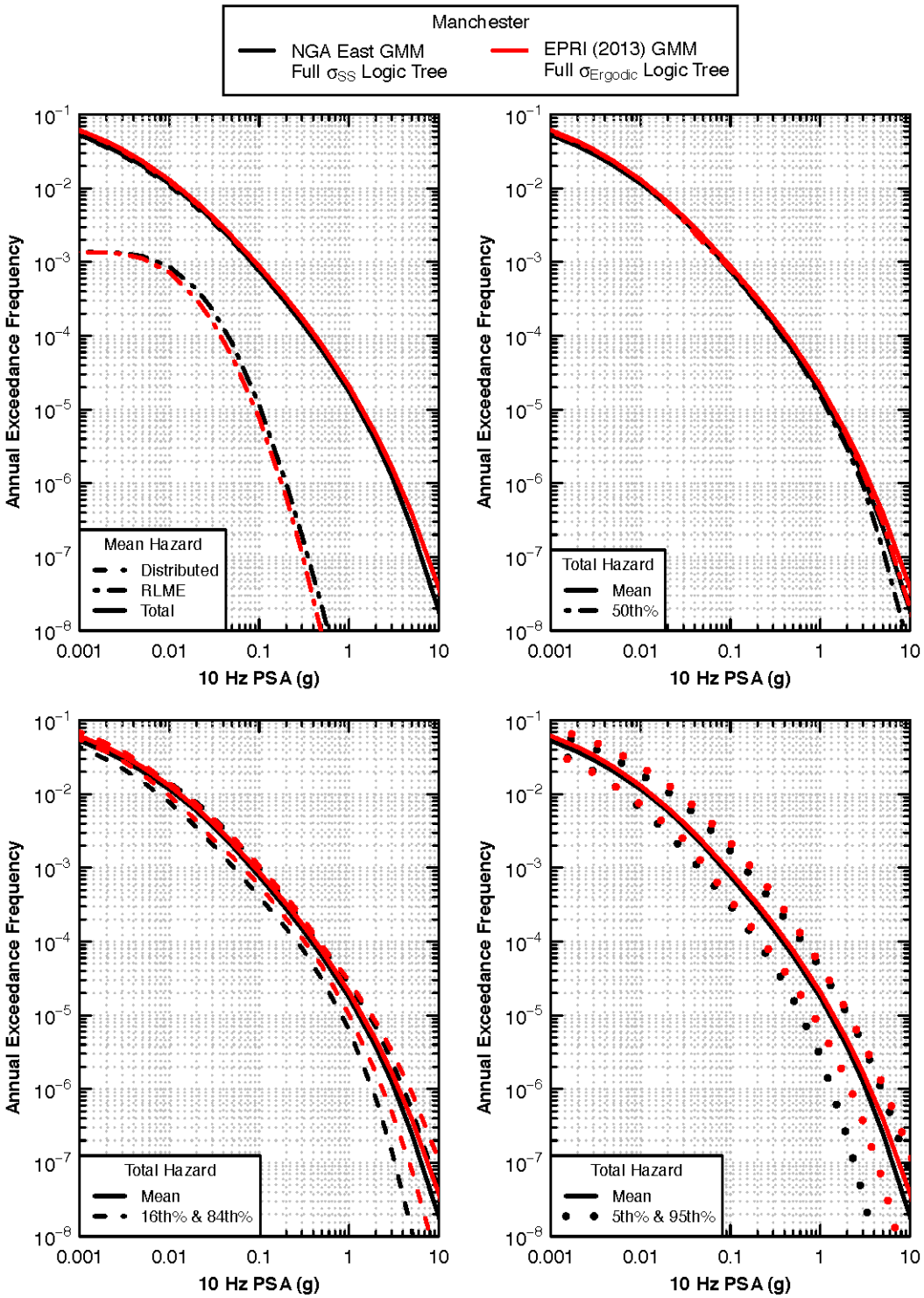


Figure 14–63 Comparison of 25 Hz PSA hazard at Manchester computed using the complete NGA-East and EPRI (2013) GMMs.

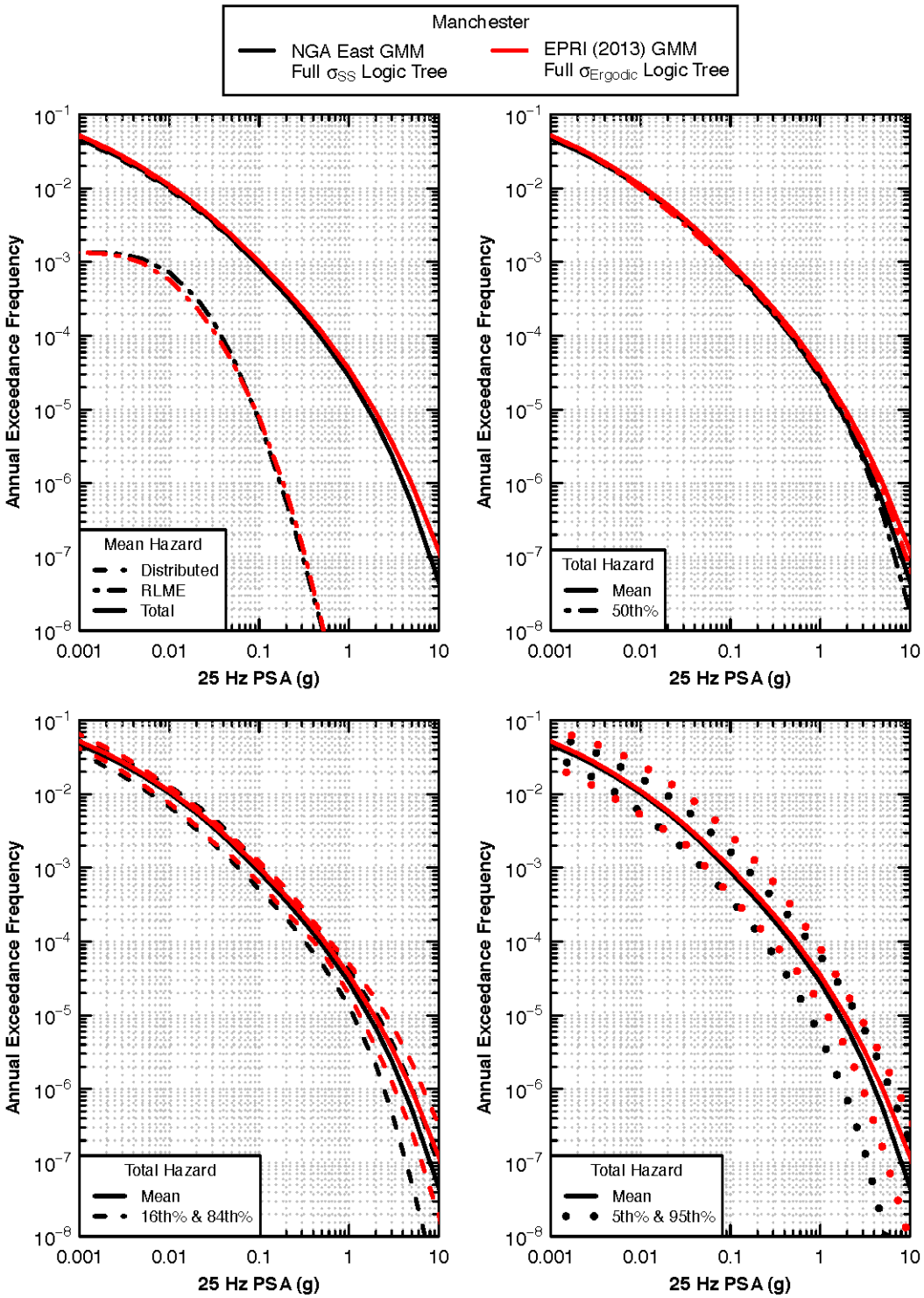


Figure 14–64 Comparison of PGA hazard at Manchester computed using the complete NGA-East and EPRI (2013) GMMs.

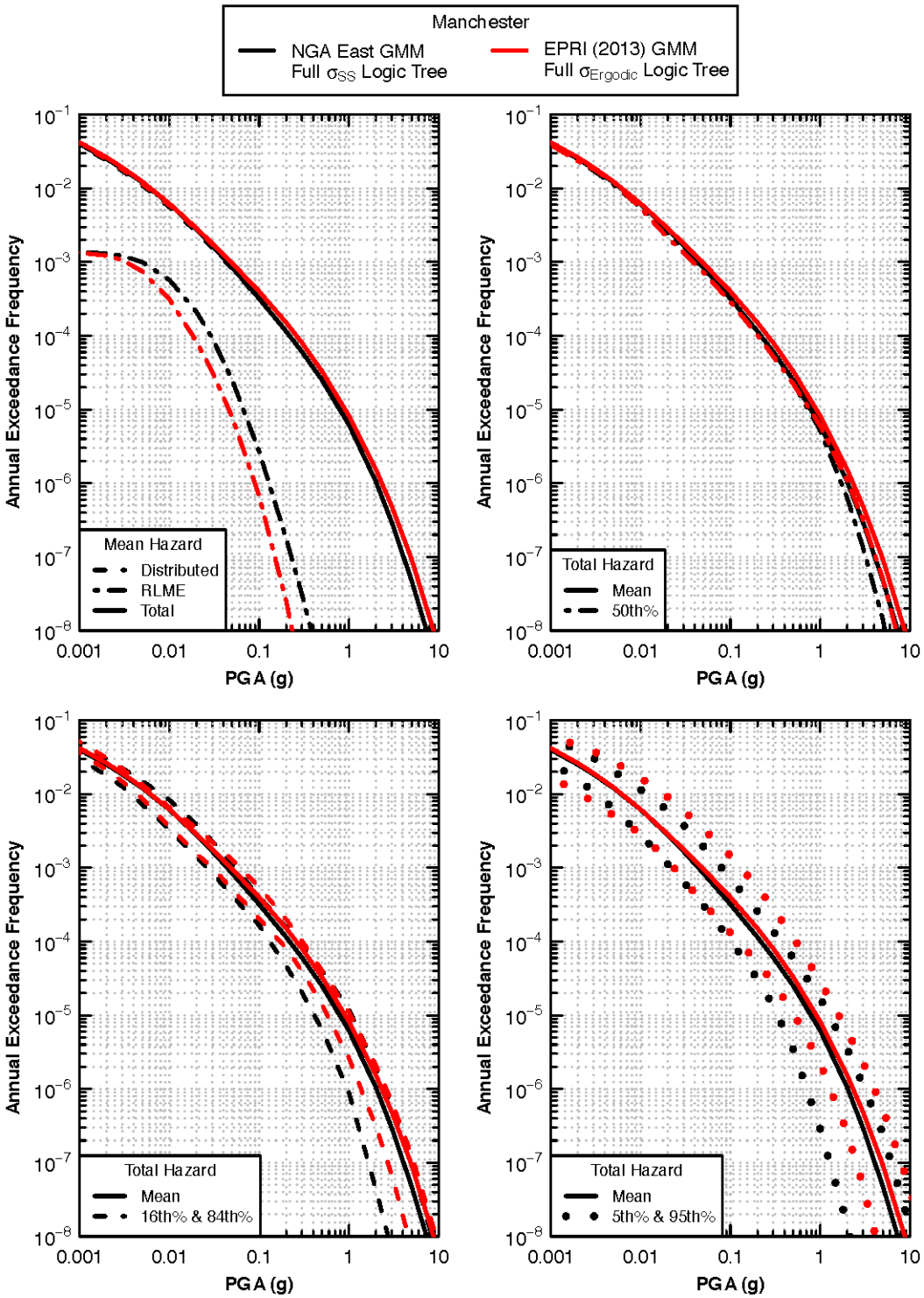


Figure 14–65(a) Ratio of mean hazard at Manchester computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

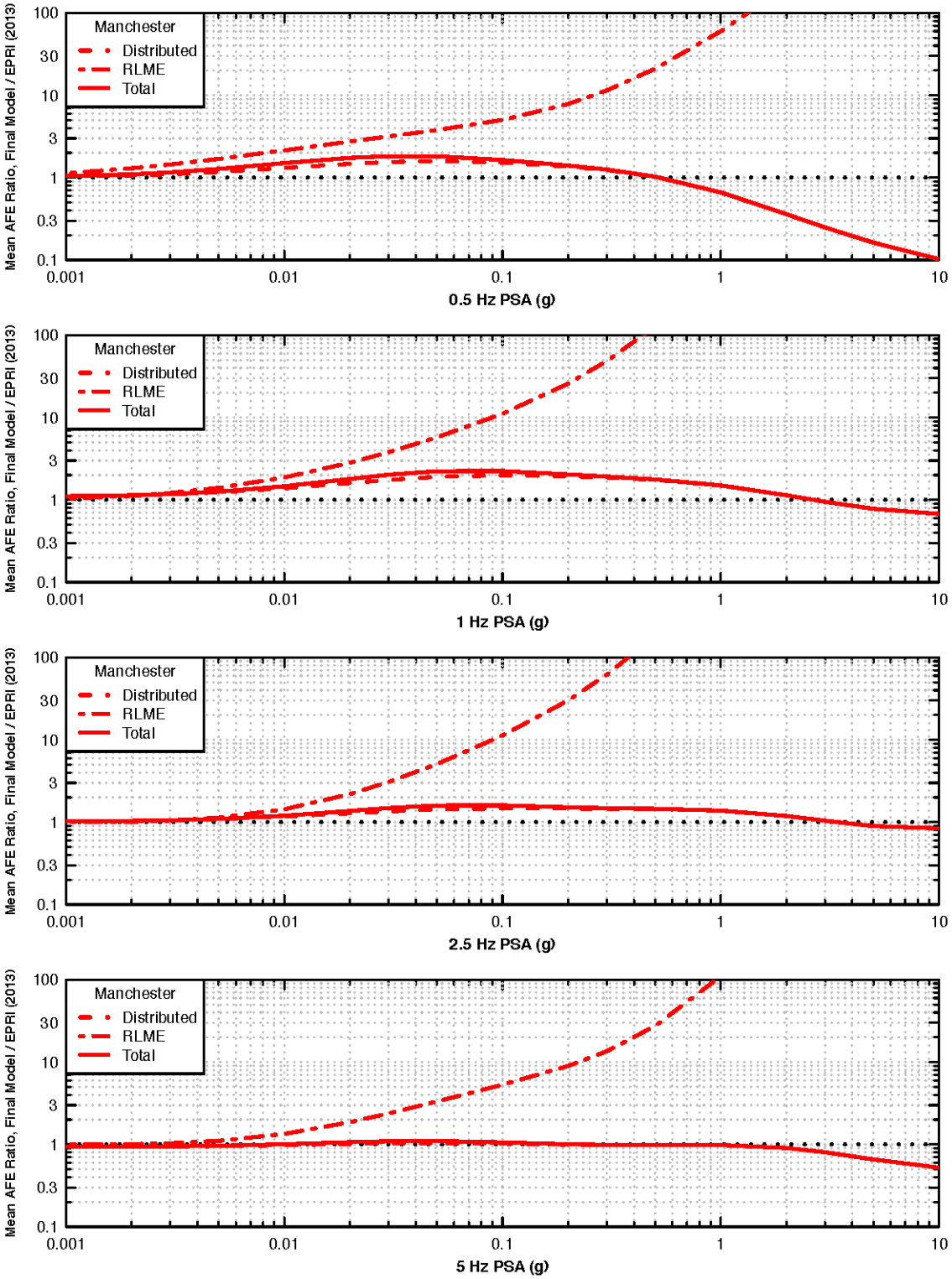


Figure 14–65(b) Ratio of mean hazard at Manchester computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

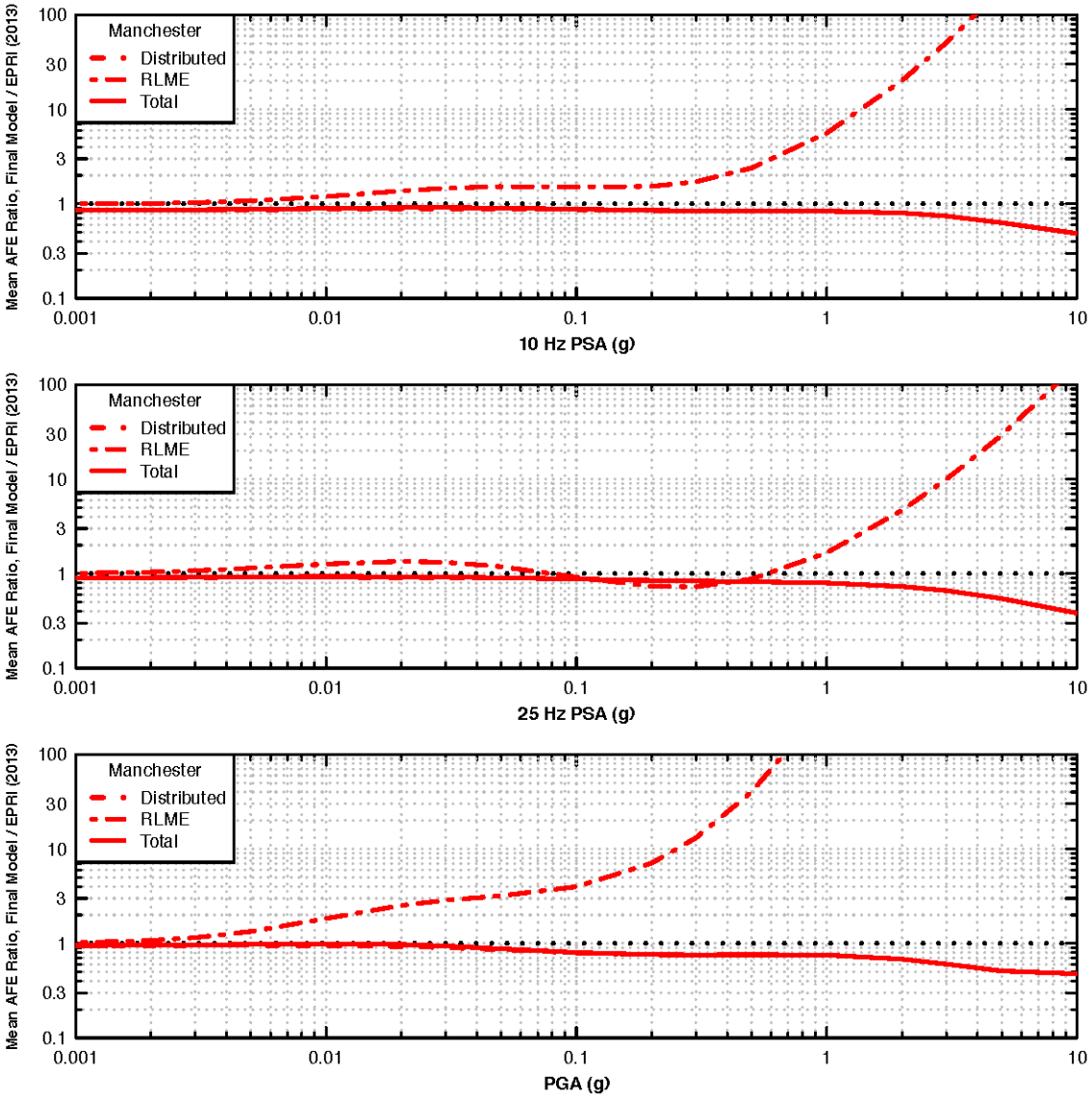


Figure 14–66 Comparison of 0.5 Hz PSA hazard at Chattanooga computed using the complete NGA-East and EPRI (2013) GMMs.

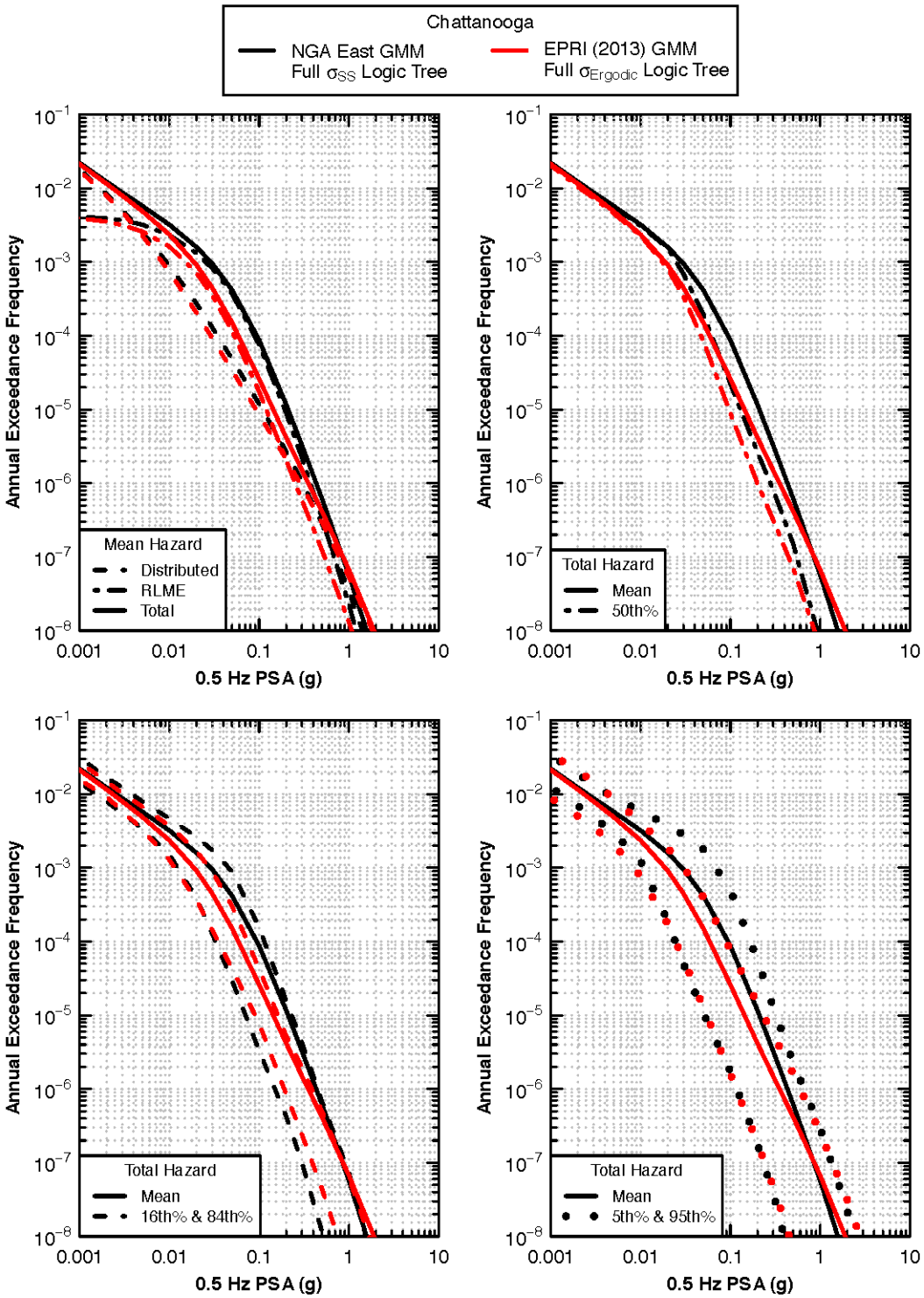


Figure 14–67 Comparison of 1 Hz PSA hazard at Chattanooga computed using the complete NGA-East and EPRI (2013) GMMs.

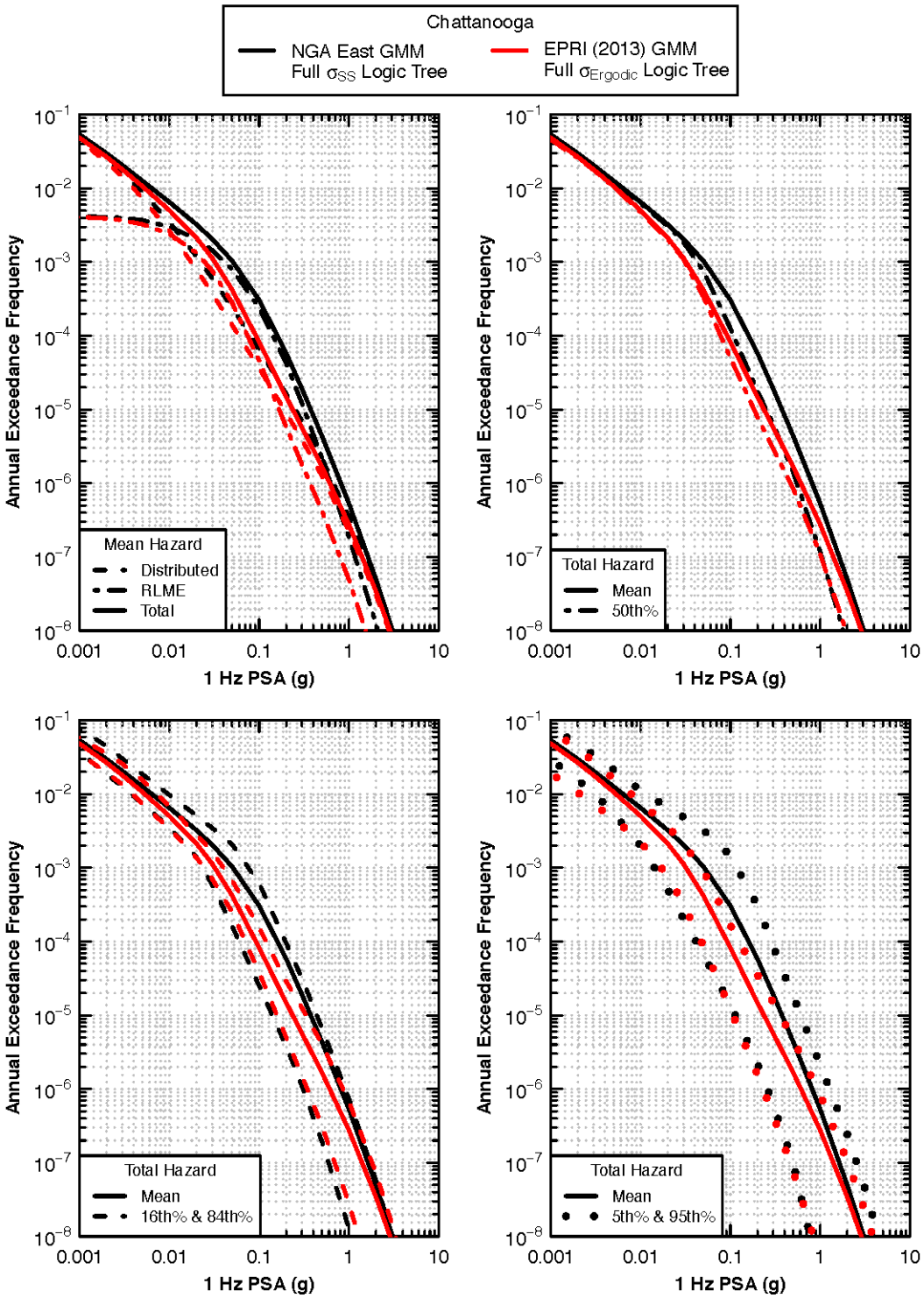


Figure 14–68 Comparison of 2.5 Hz PSA hazard at Chattanooga computed using the complete NGA-East and EPRI (2013) GMMs.

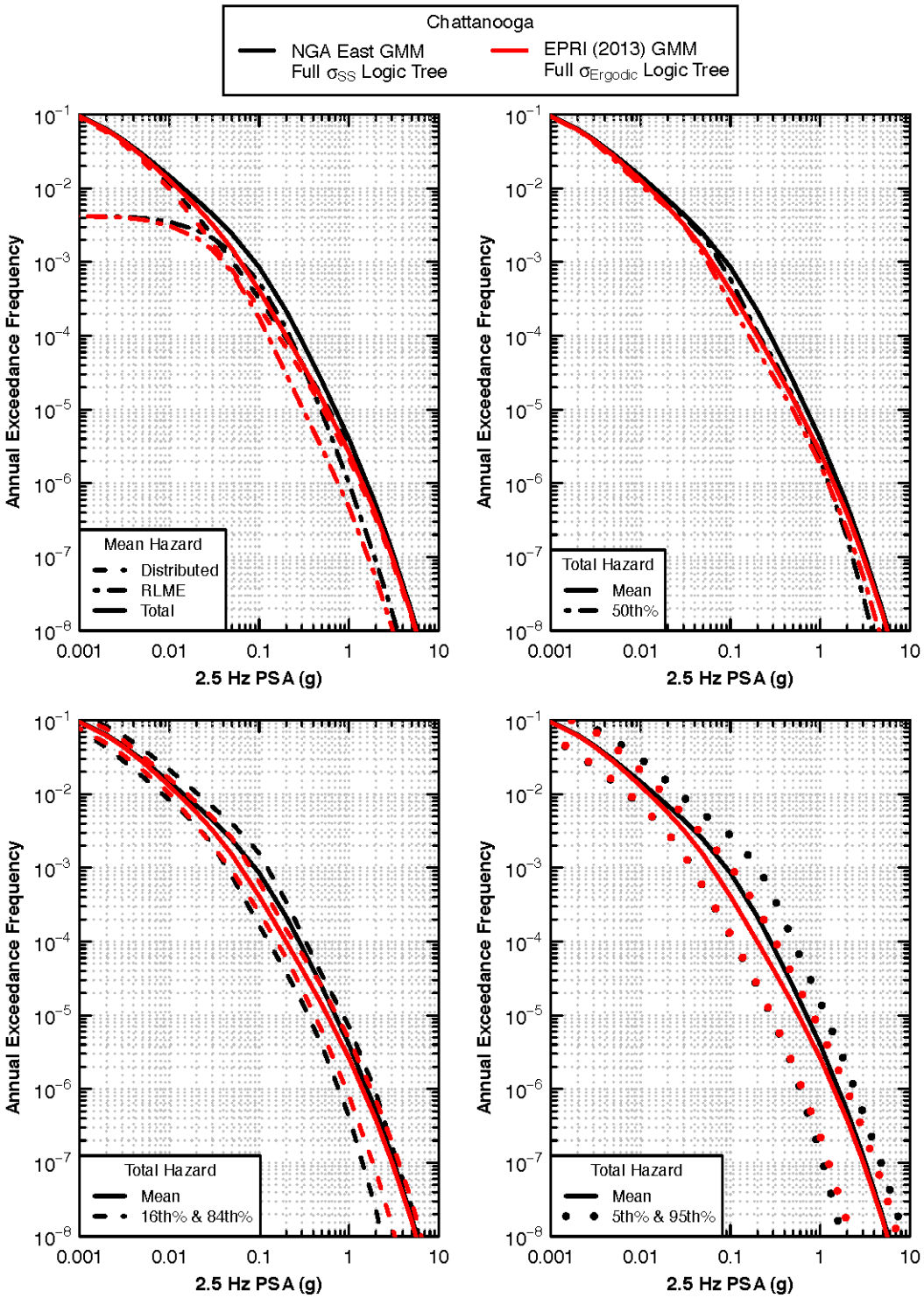


Figure 14–69 Comparison of 5 Hz PSA hazard at Chattanooga computed using the complete NGA-East and EPRI (2013) GMMs.

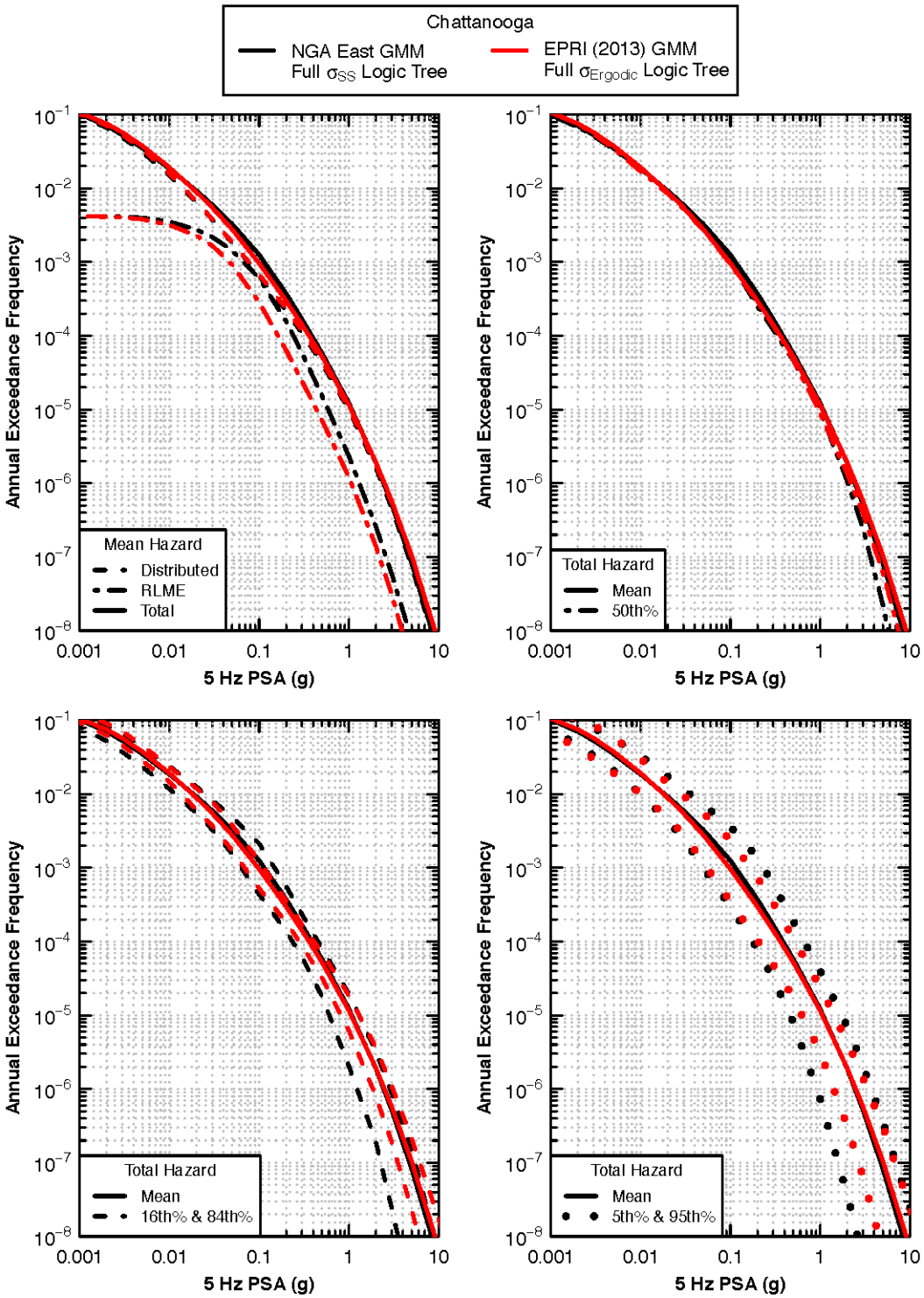


Figure 14–70 Comparison of 10 Hz PSA hazard at Chattanooga computed using the complete NGA-East and EPRI (2013) GMMs.

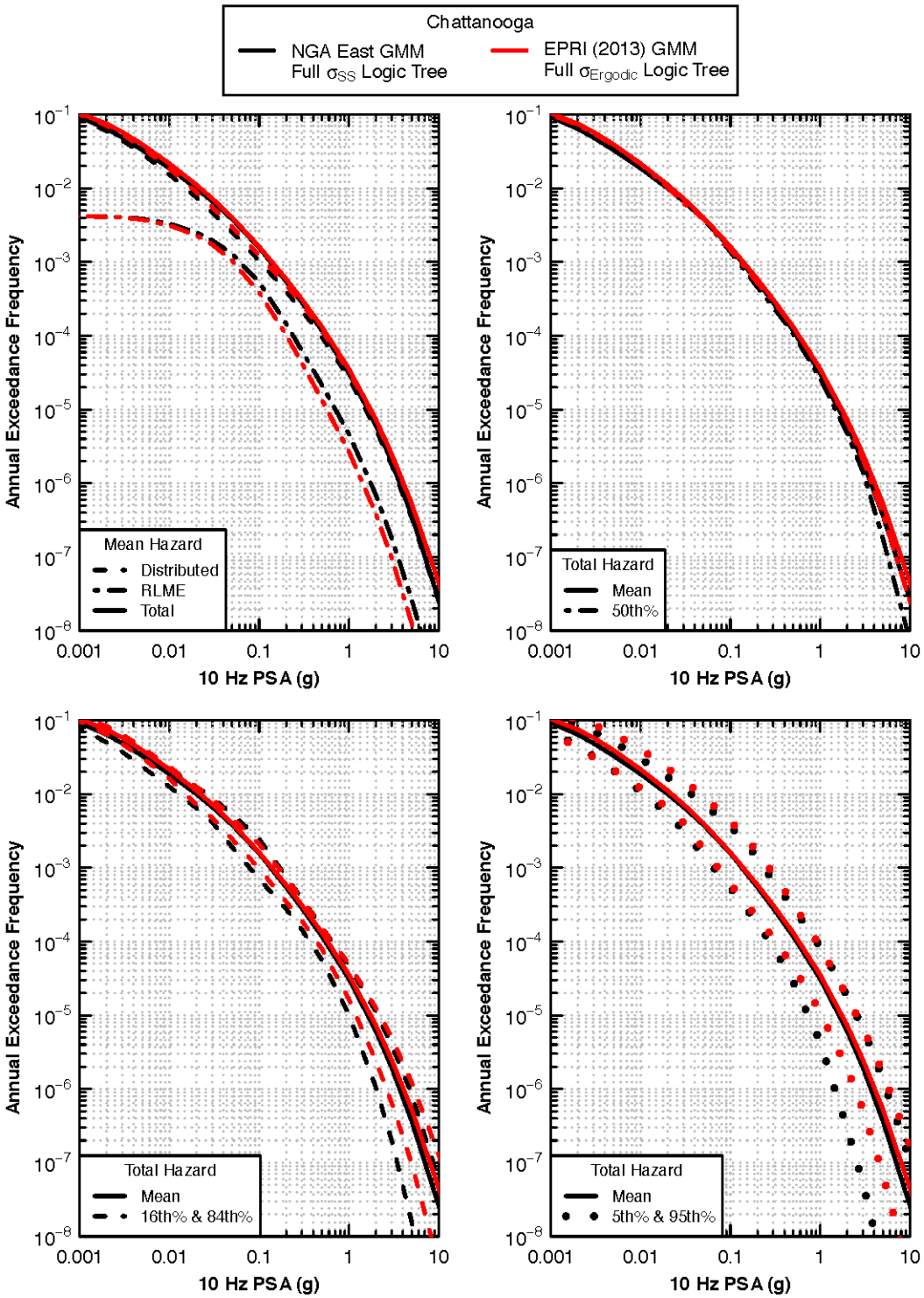


Figure 14–71 Comparison of 25 Hz PSA hazard at Chattanooga computed using the complete NGA-East and EPRI (2013) GMMs.

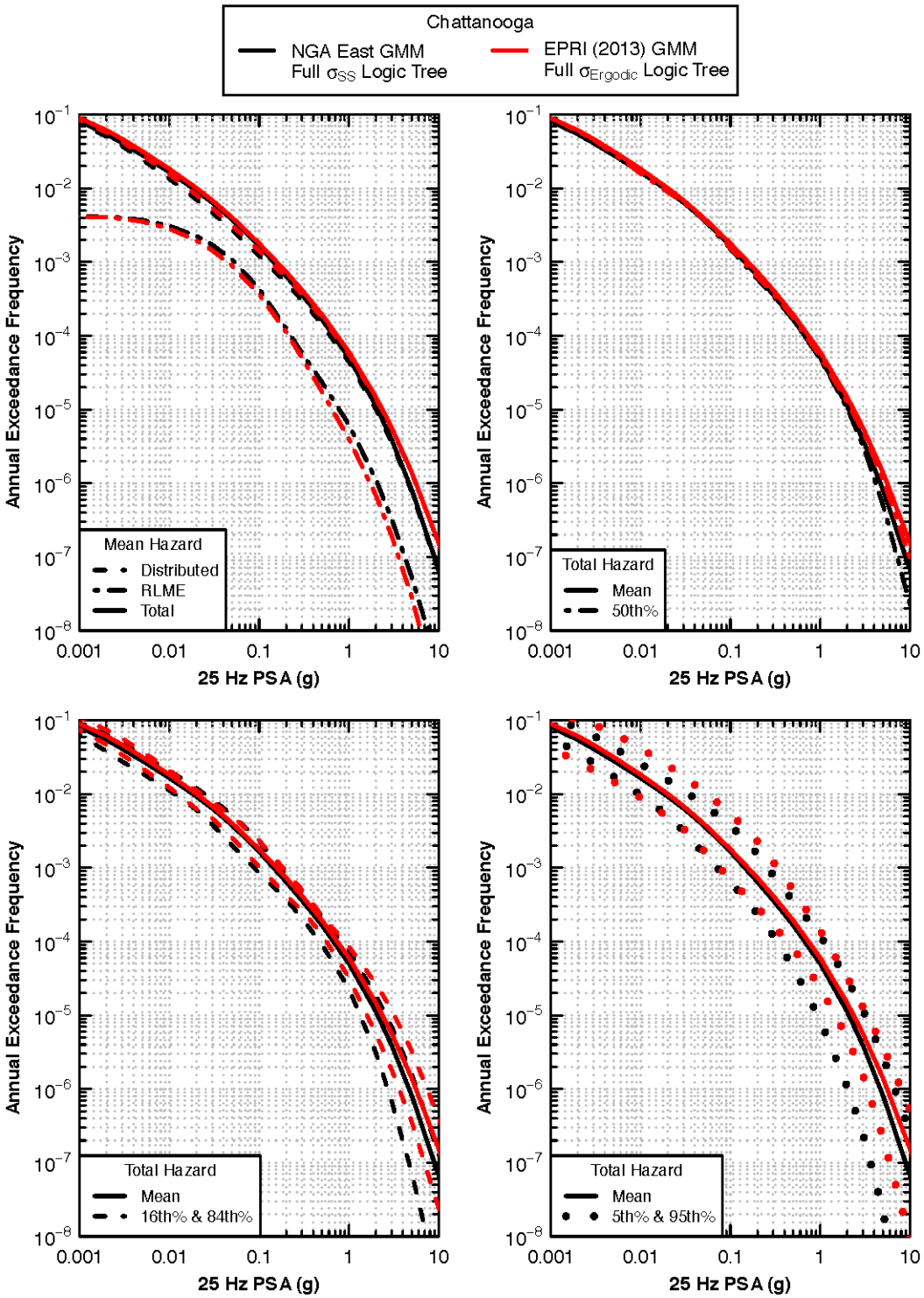


Figure 14–72 Comparison of PGA hazard at Chattanooga computed using the complete NGA-East and EPRI (2013) GMMs.

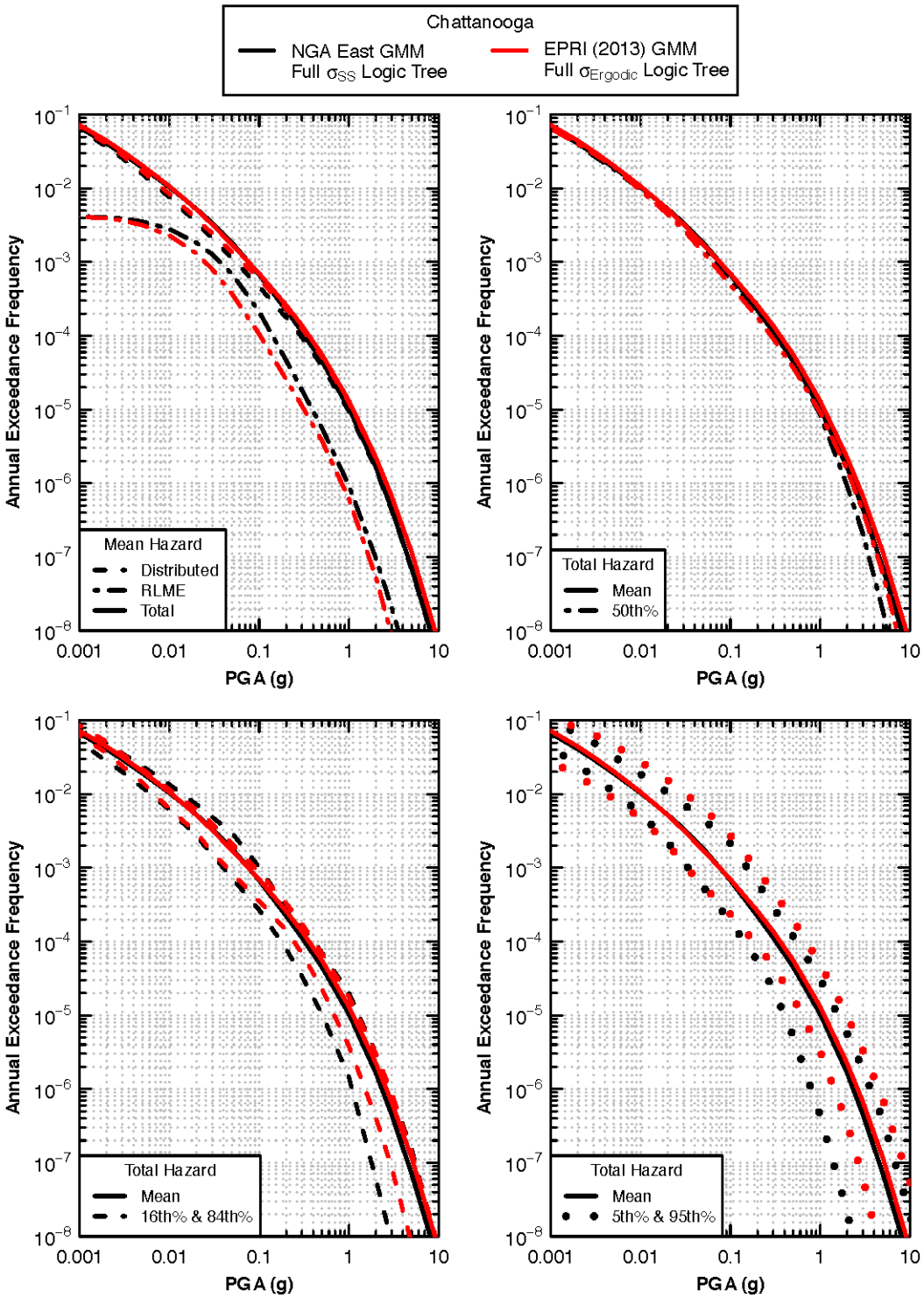


Figure 14–73(a) Ratio of mean hazard at Chattanooga computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

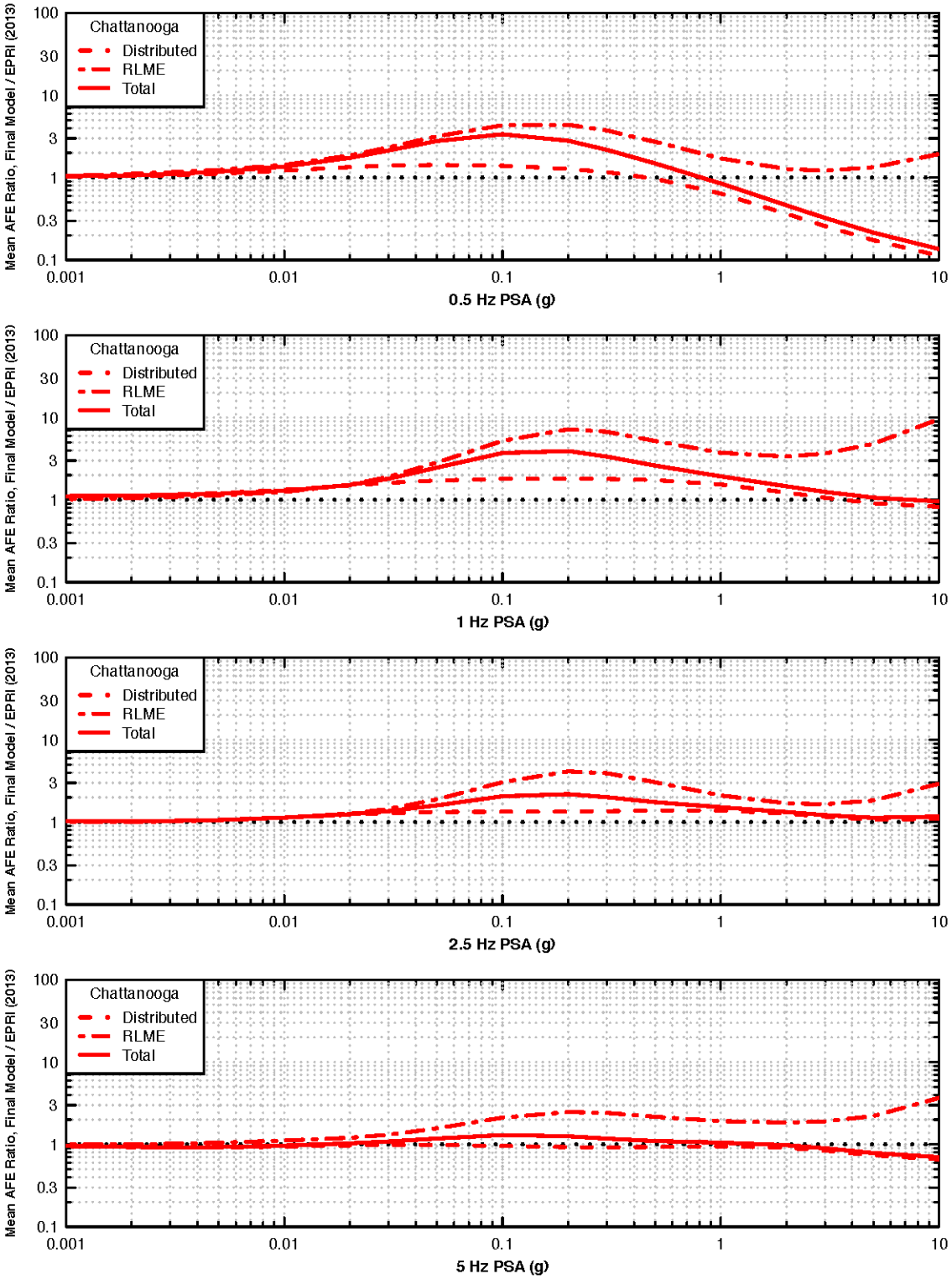


Figure 14–73(b) Ratio of mean hazard at Chattanooga computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

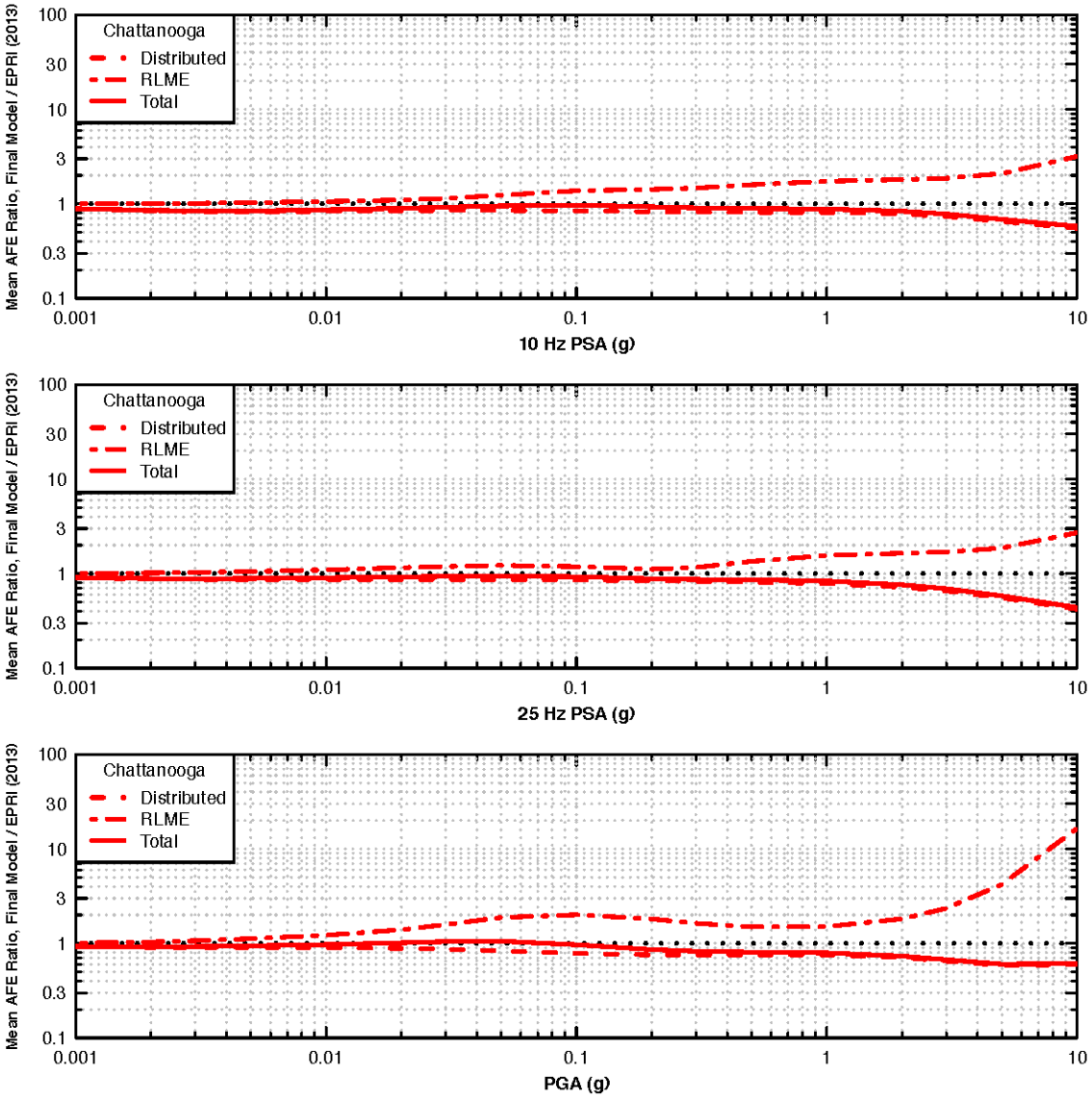


Figure 14–74 Comparison of 0.5 Hz PSA hazard at Topeka computed using the complete NGA-East and EPRI (2013) GMMs.

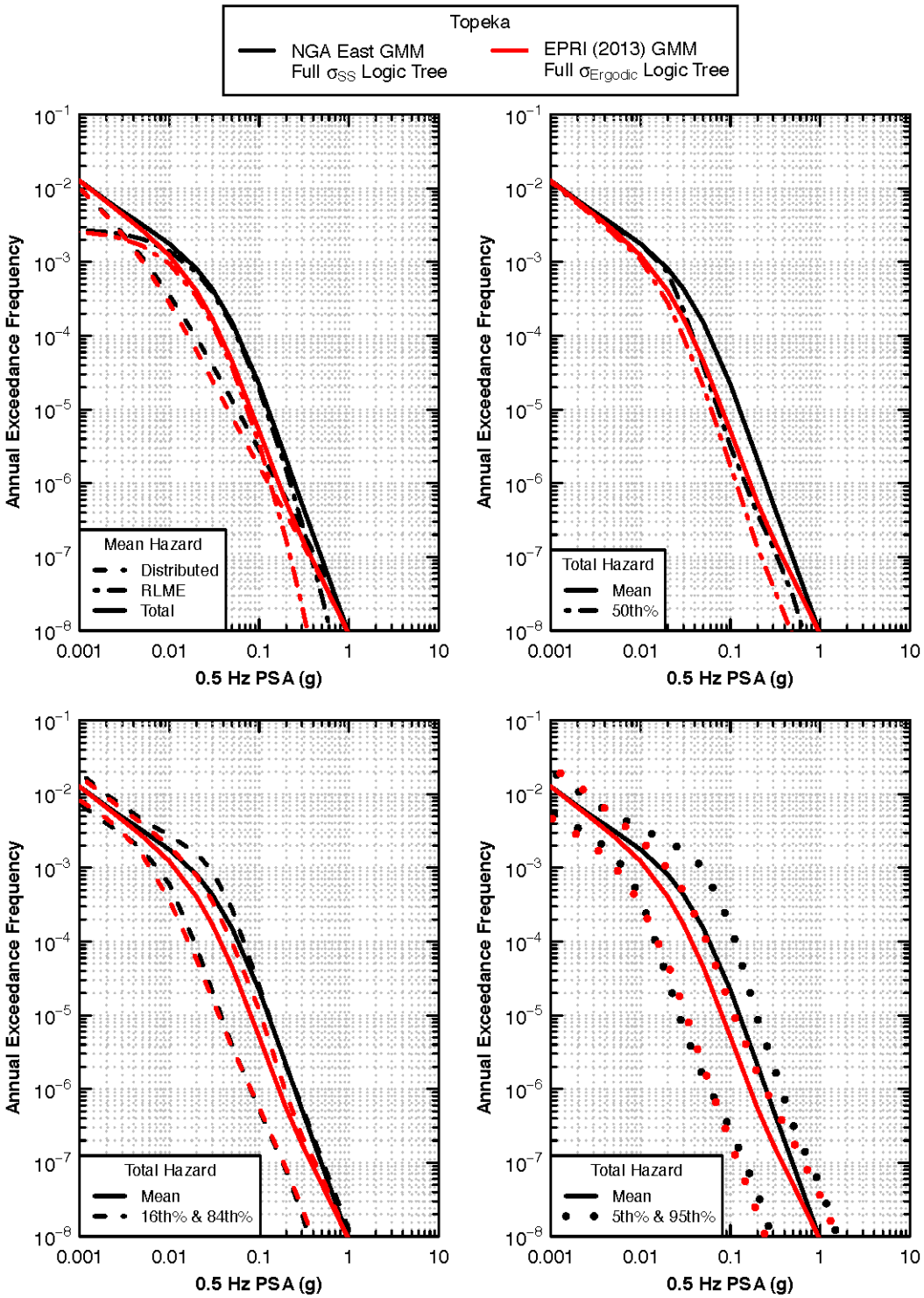


Figure 14–75 Comparison of 1 Hz PSA hazard at Topeka computed using the complete NGA-East and EPRI (2013) GMMs.

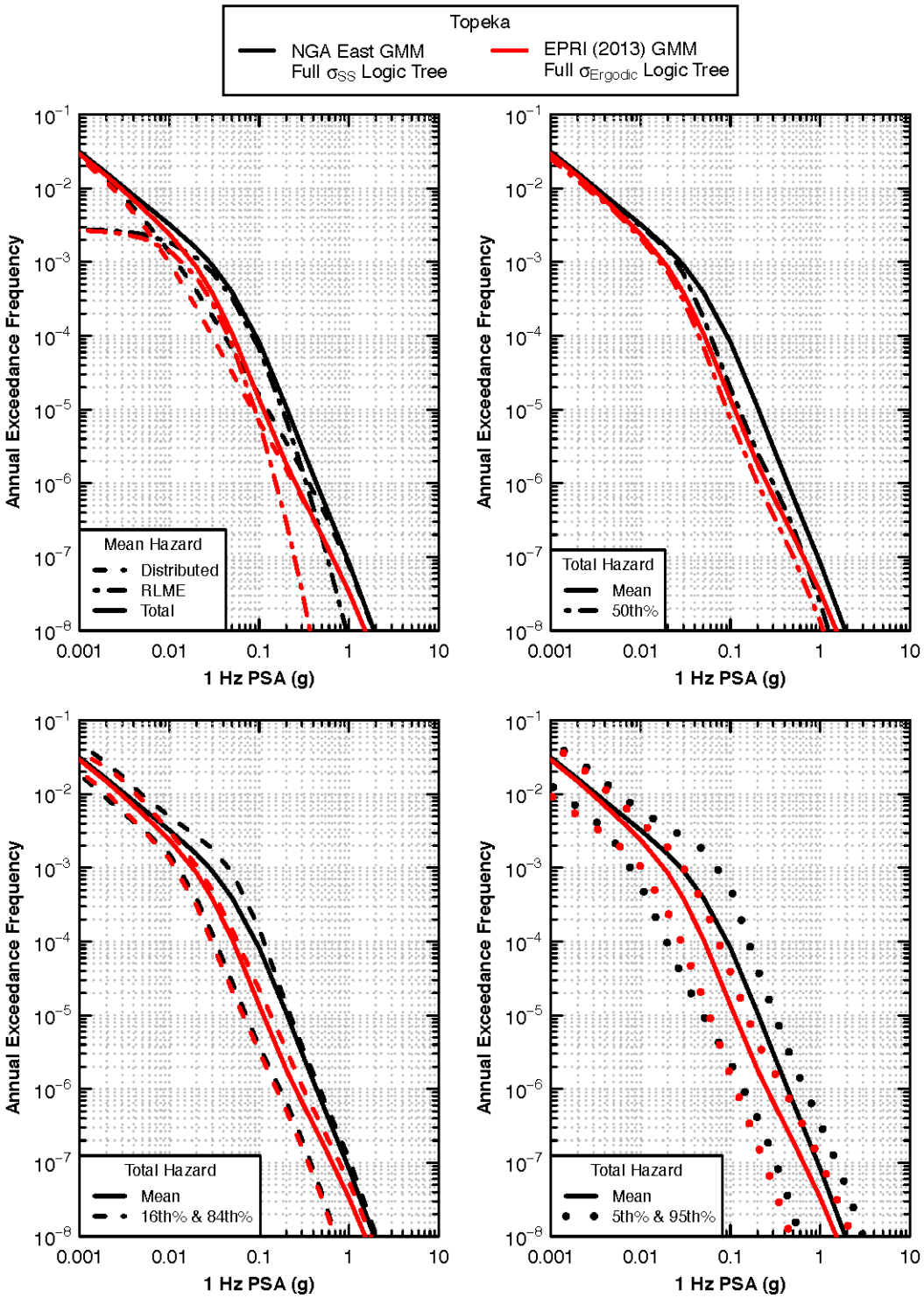


Figure 14–76 Comparison of 2.5 Hz PSA hazard at Topeka computed using the complete NGA-East and EPRI (2013) GMMs.

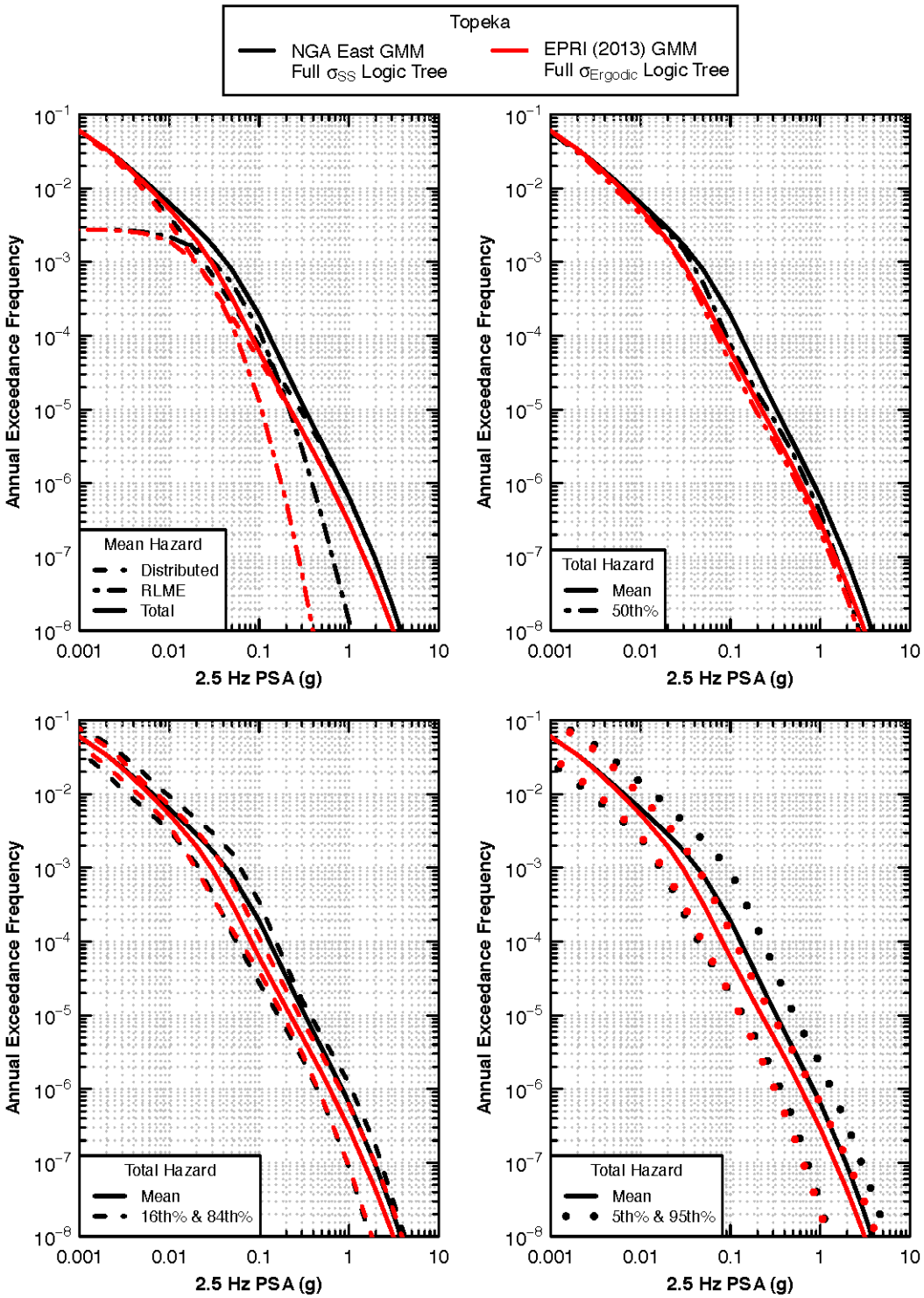


Figure 14–77 Comparison of 5 Hz PSA hazard at Topeka computed using the complete NGA-East and EPRI (2013) GMMs.

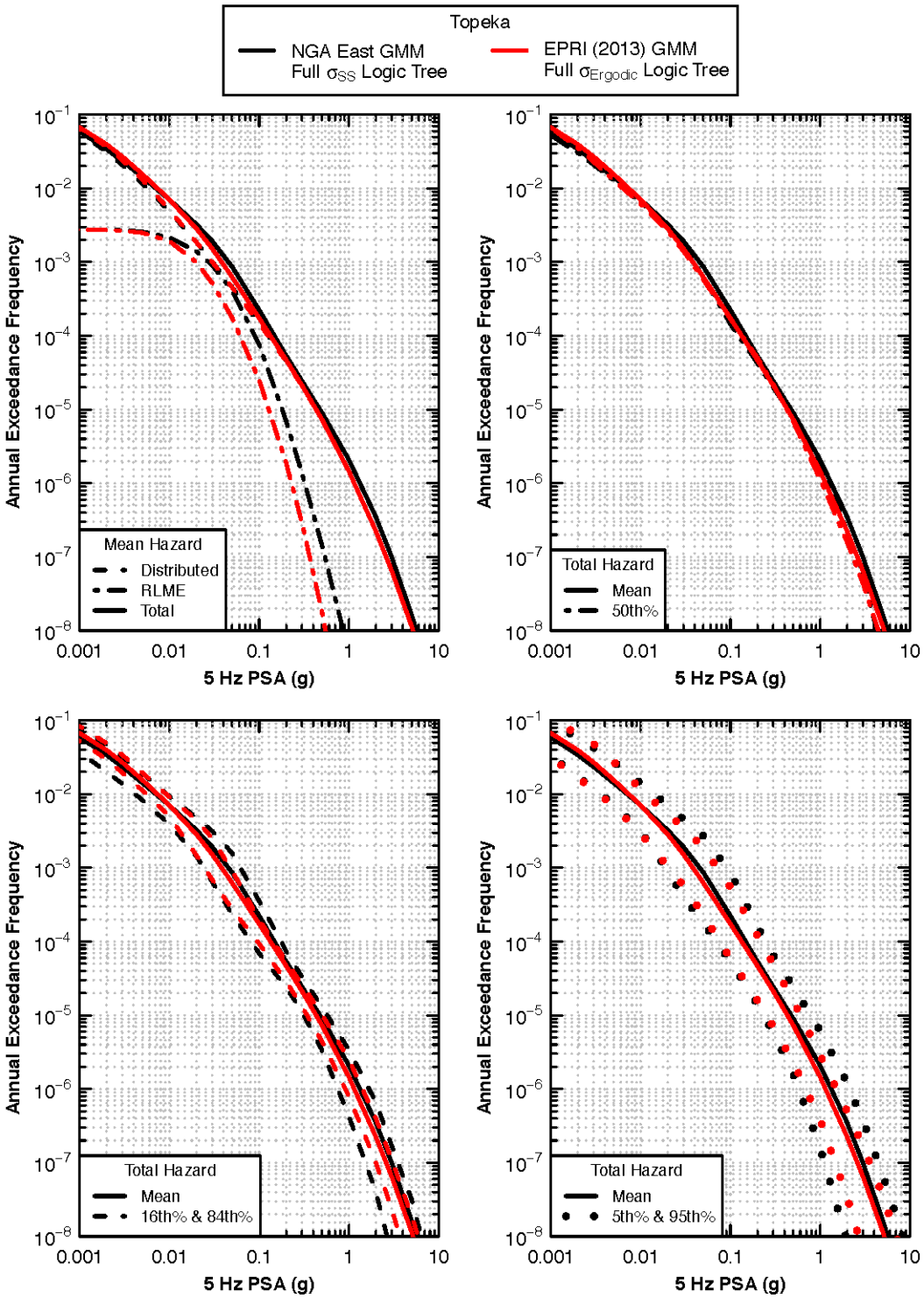


Figure 14–78 Comparison of 10 Hz PSA hazard at Topeka computed using the complete NGA-East and EPRI (2013) GMMs.

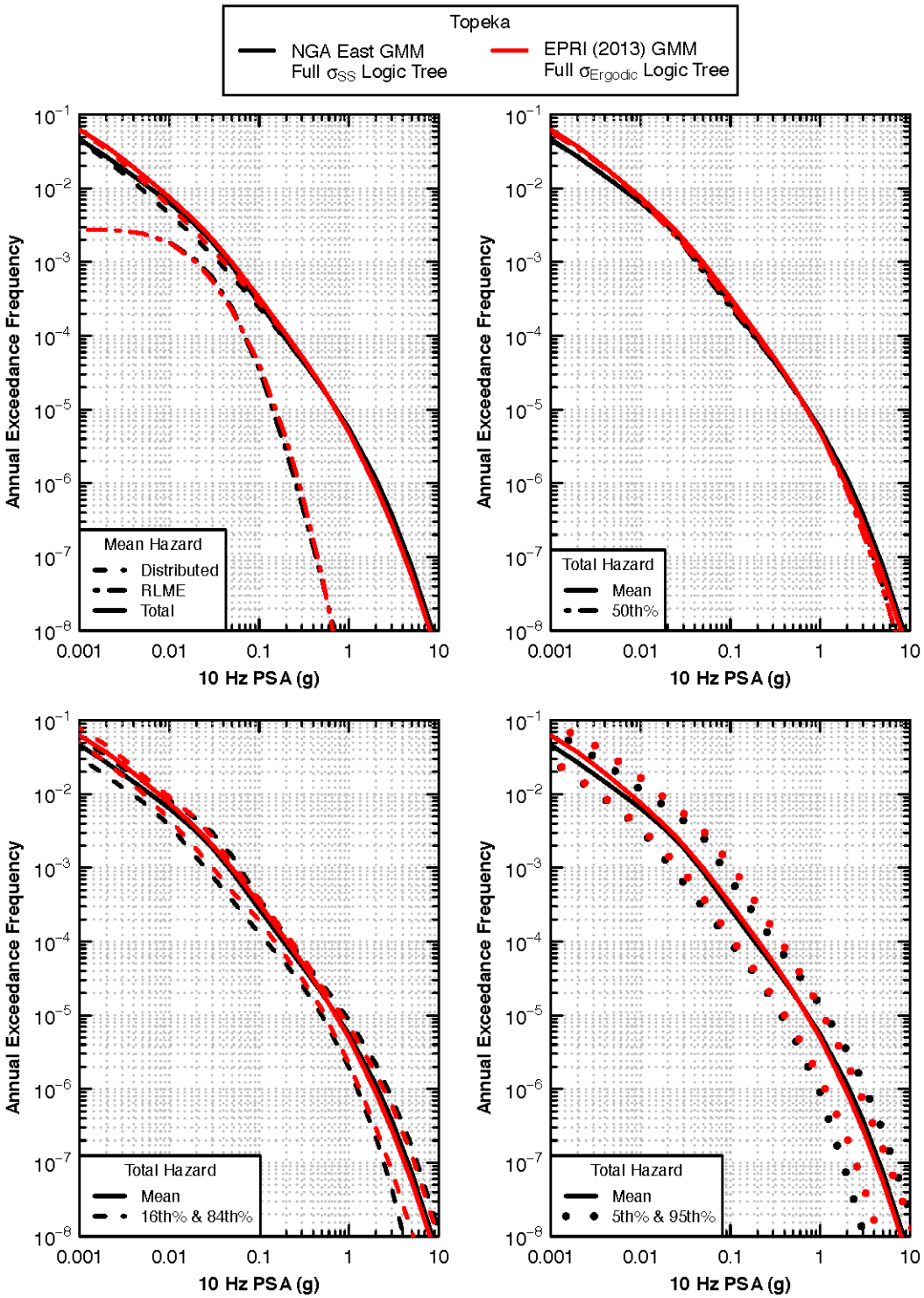


Figure 14–79 Comparison of 25 Hz PSA hazard at Topeka computed using the complete NGA-East and EPRI (2013) GMMs.

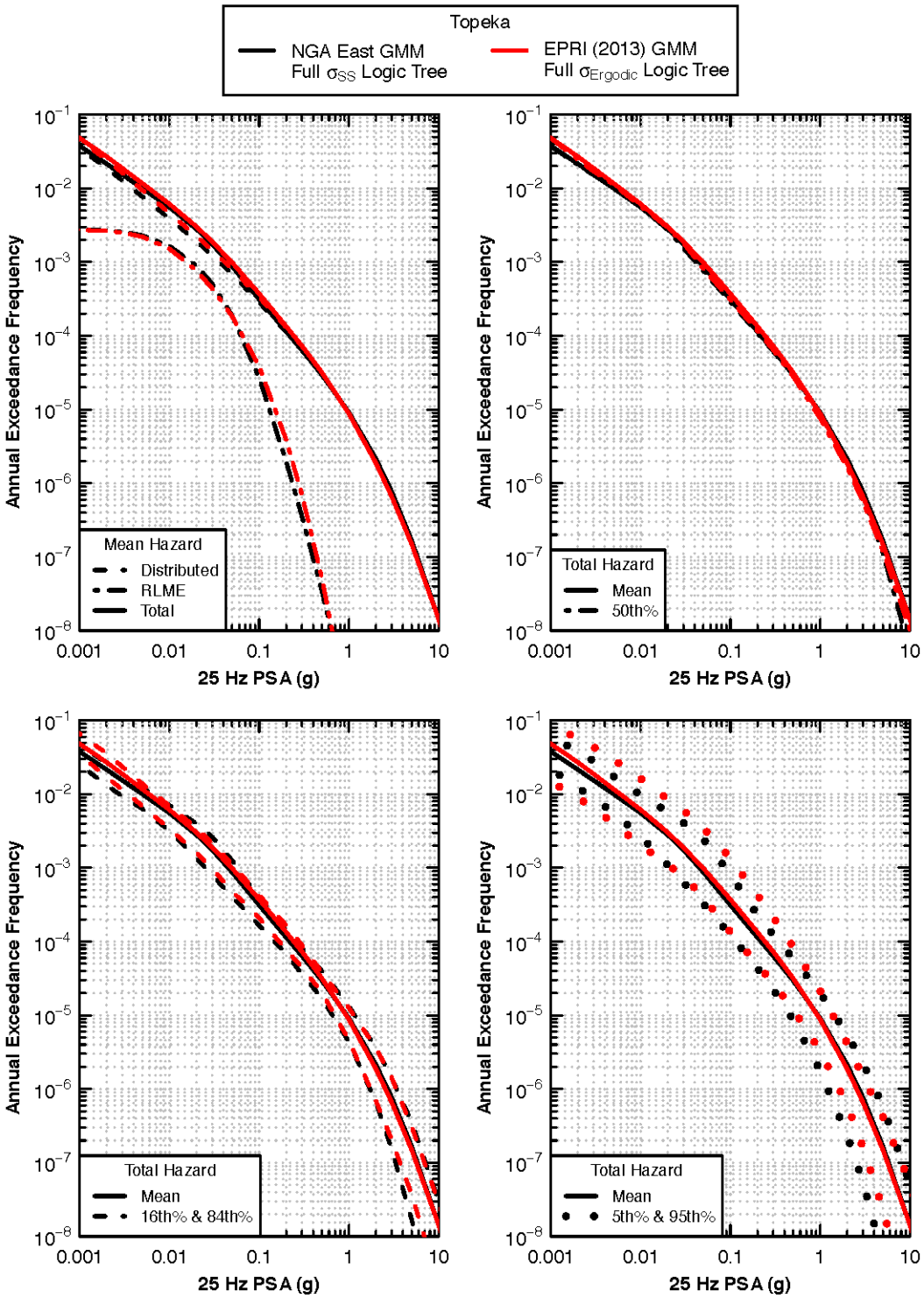


Figure 14–80 Comparison of PGA hazard at Topeka computed using the complete NGA-East and EPRI (2013) GMMs.

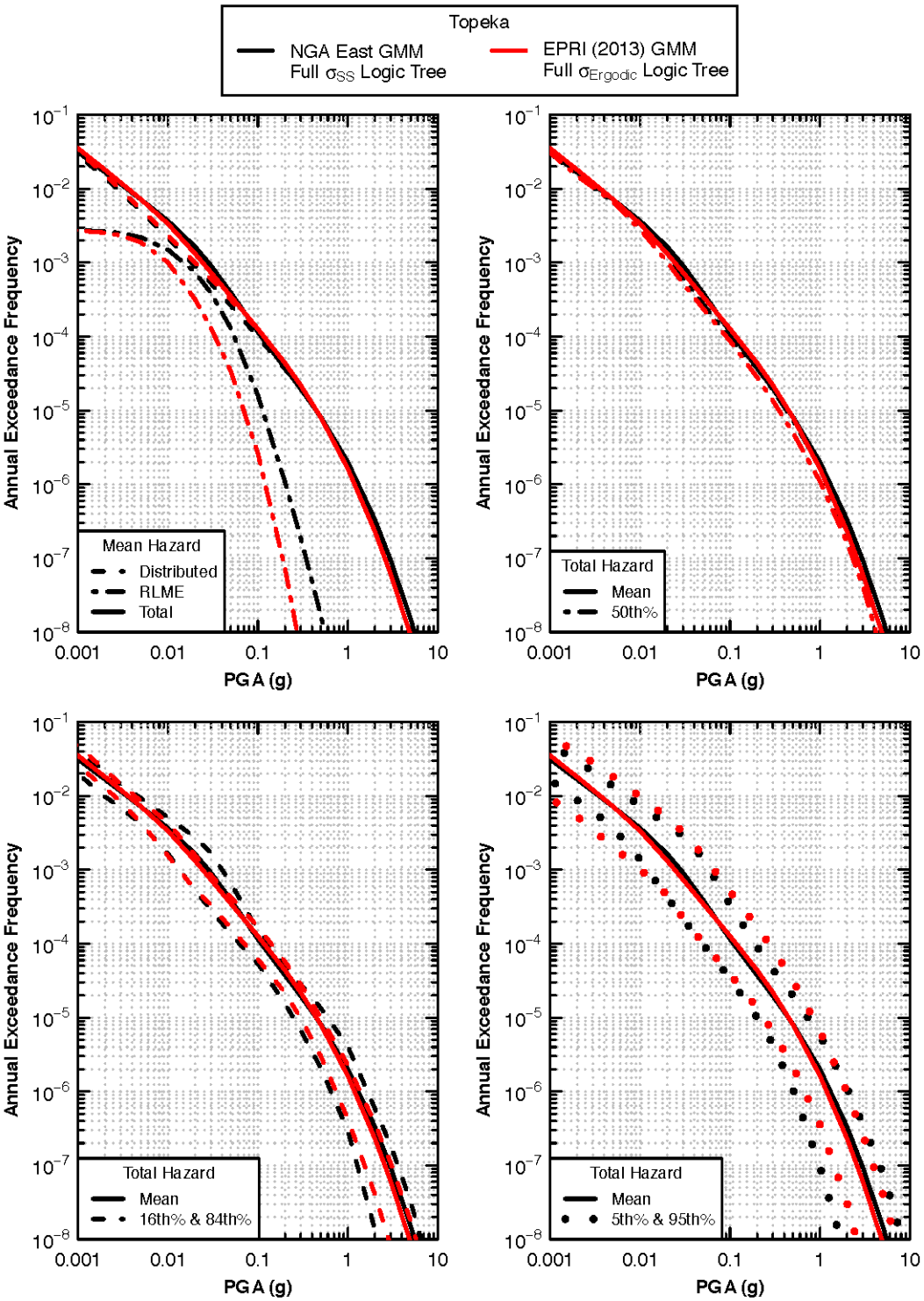


Figure 14–81(a) Ratio of mean hazard at Topeka computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

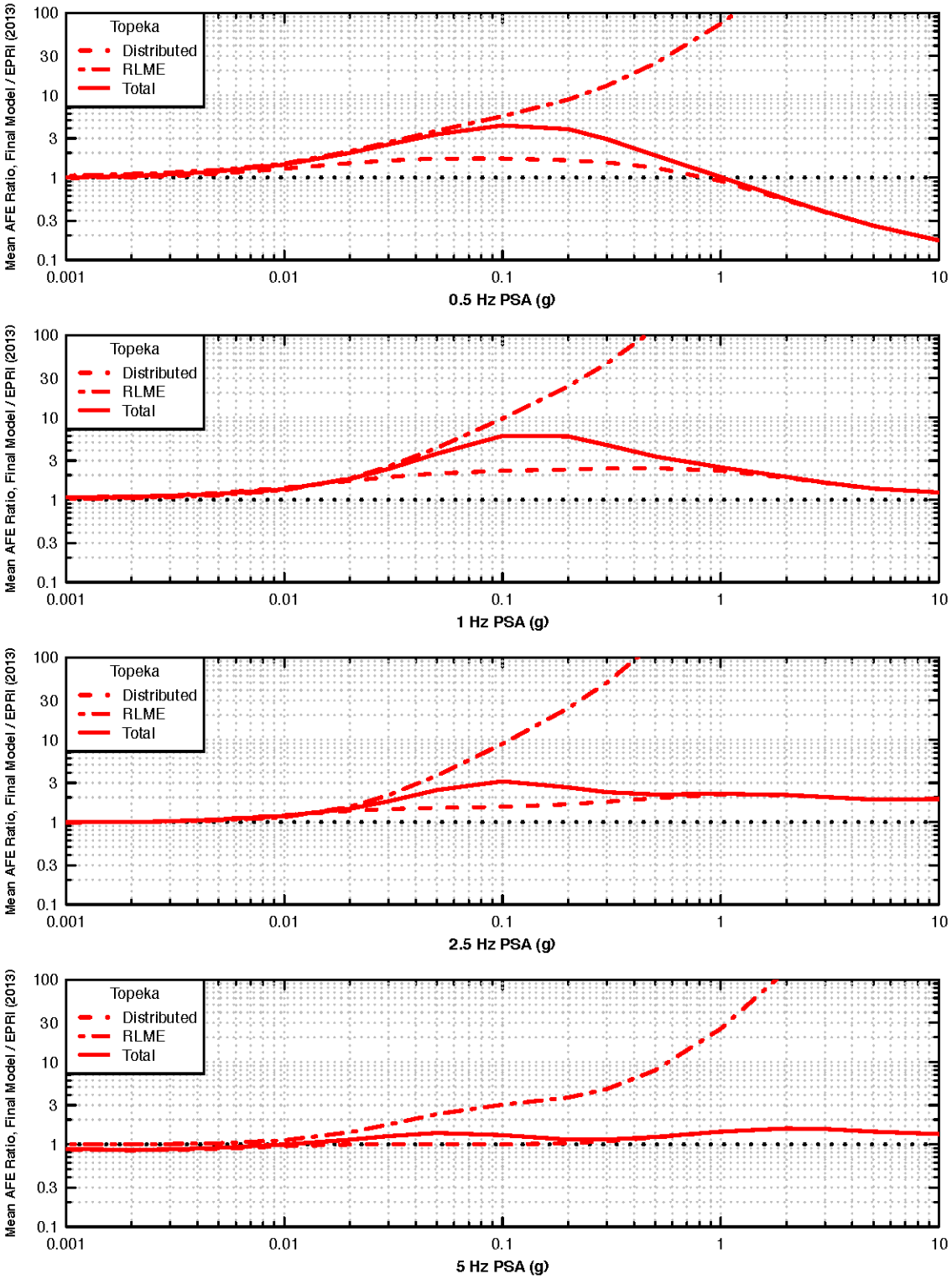


Figure 14–81(b) Ratio of mean hazard at Topeka computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

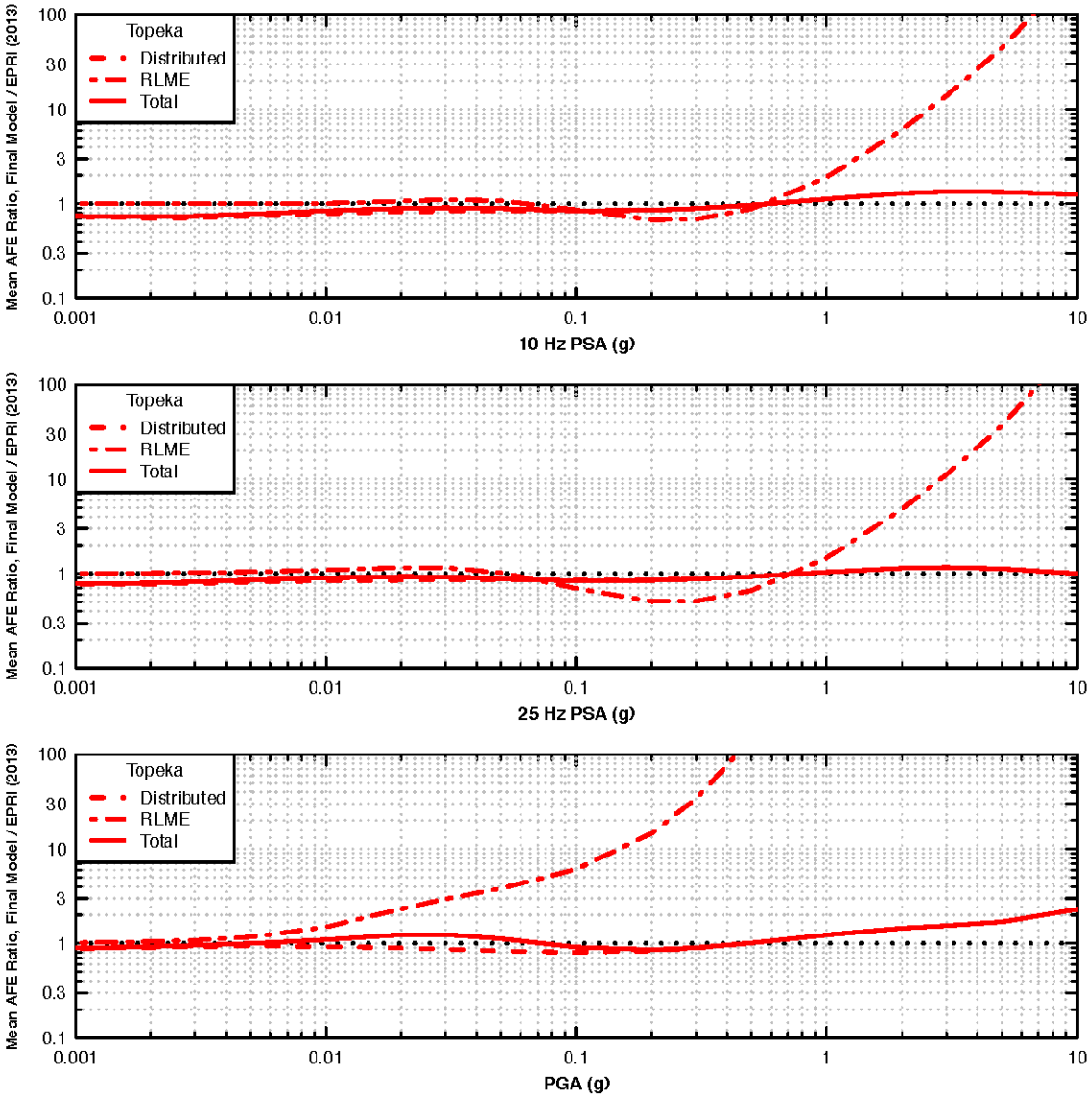


Figure 14–82 Comparison of 0.5 Hz PSA hazard at Houston computed using the complete NGA-East and EPRI (2013) GMMs.

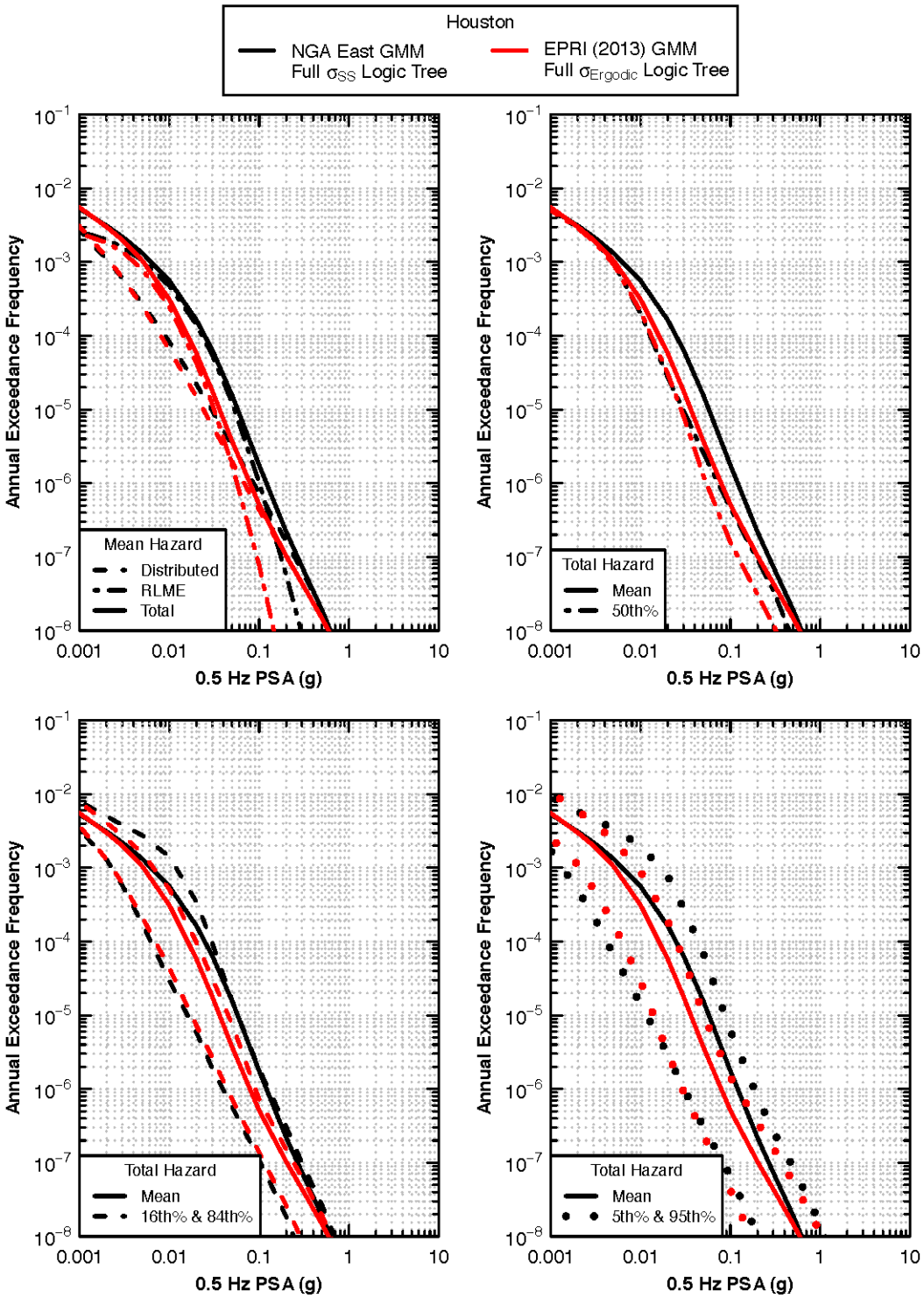


Figure 14–83 Comparison of 1 Hz PSA hazard at Houston computed using the complete NGA-East and EPRI (2013) GMMs.

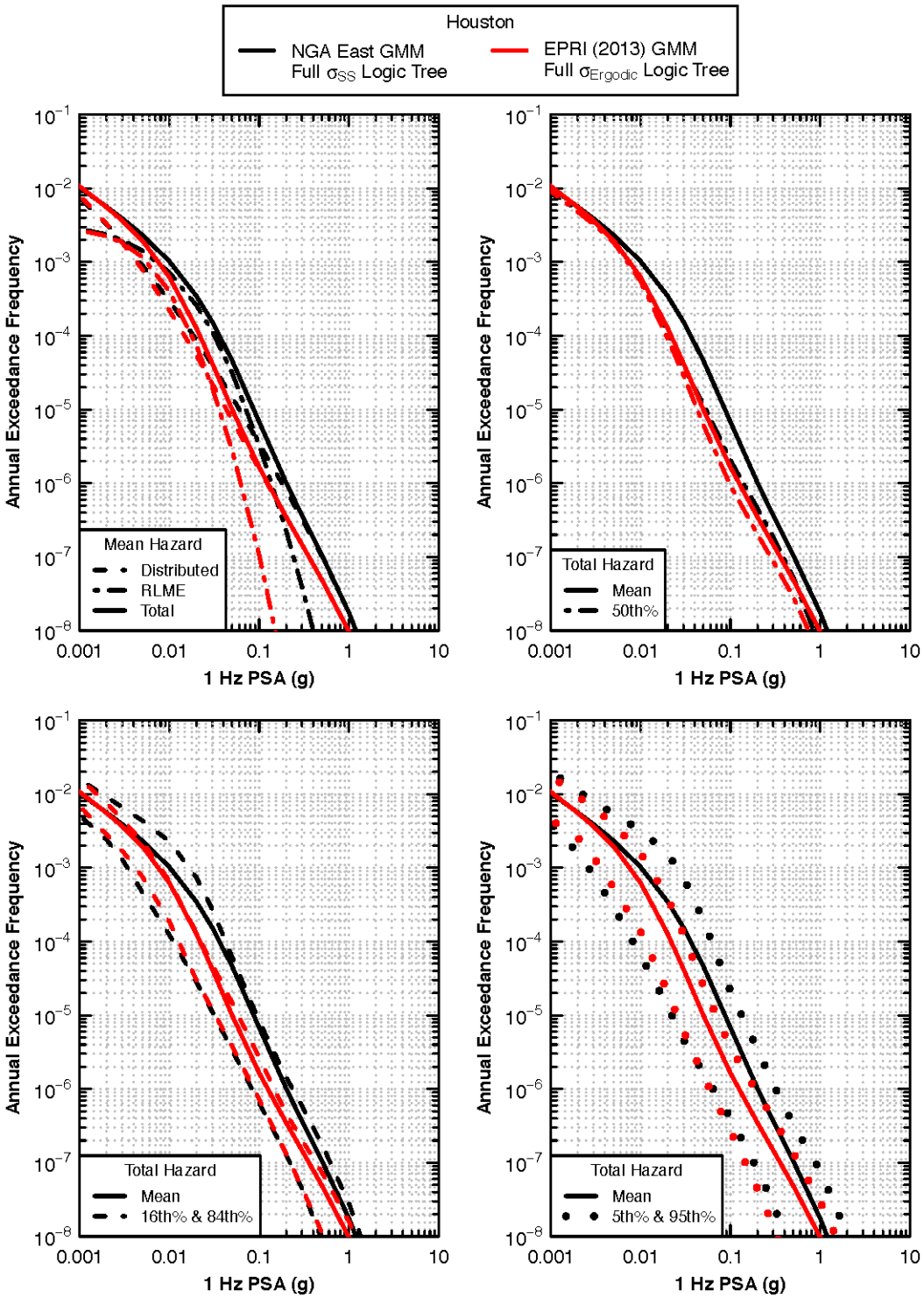


Figure 14–84 Comparison of 2.5 Hz PSA hazard at Houston computed using the complete NGA-East and EPRI (2013) GMMs.

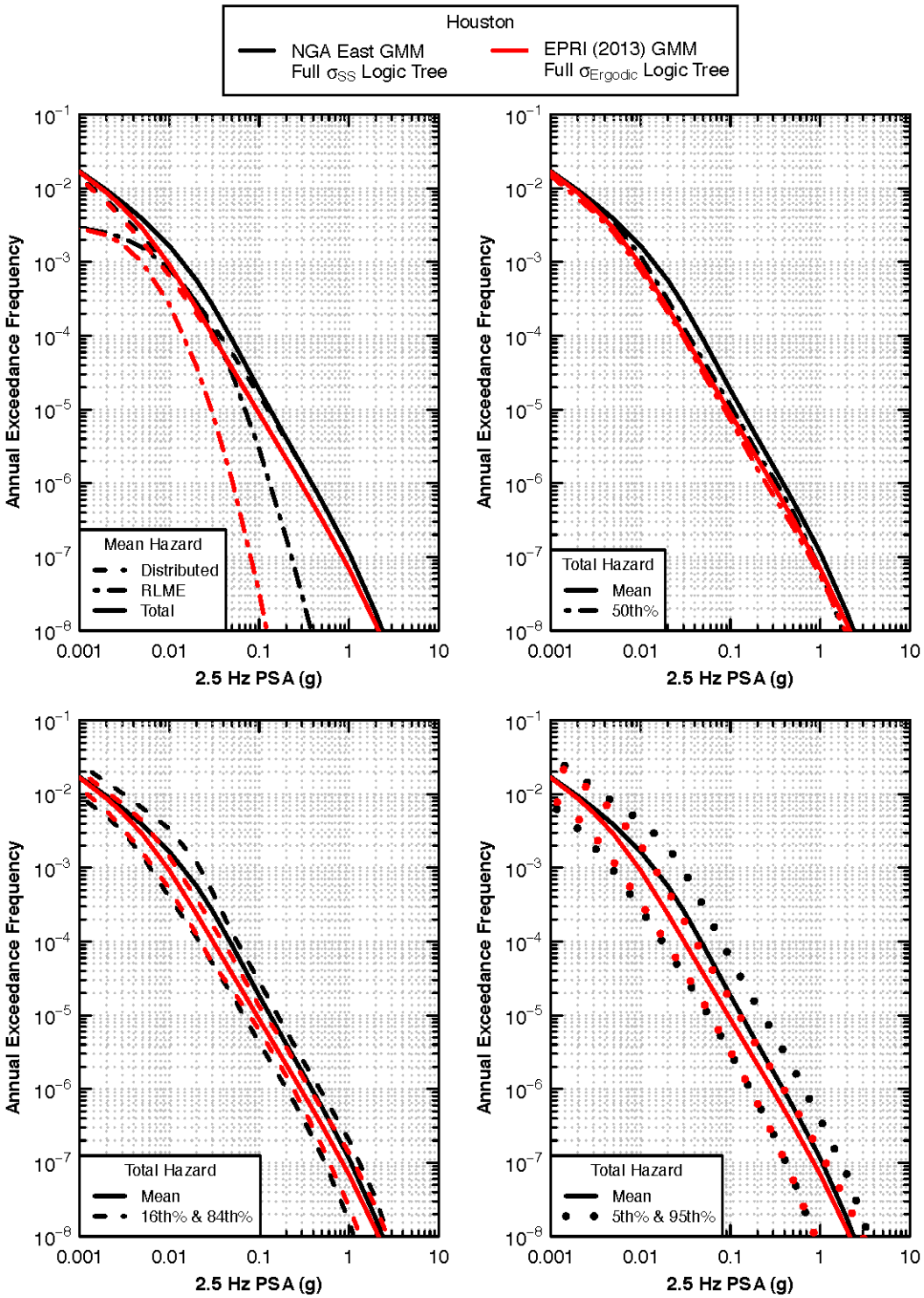


Figure 14–85 Comparison of 5 Hz PSA hazard at Houston computed using the complete NGA-East and EPRI (2013) GMMs.

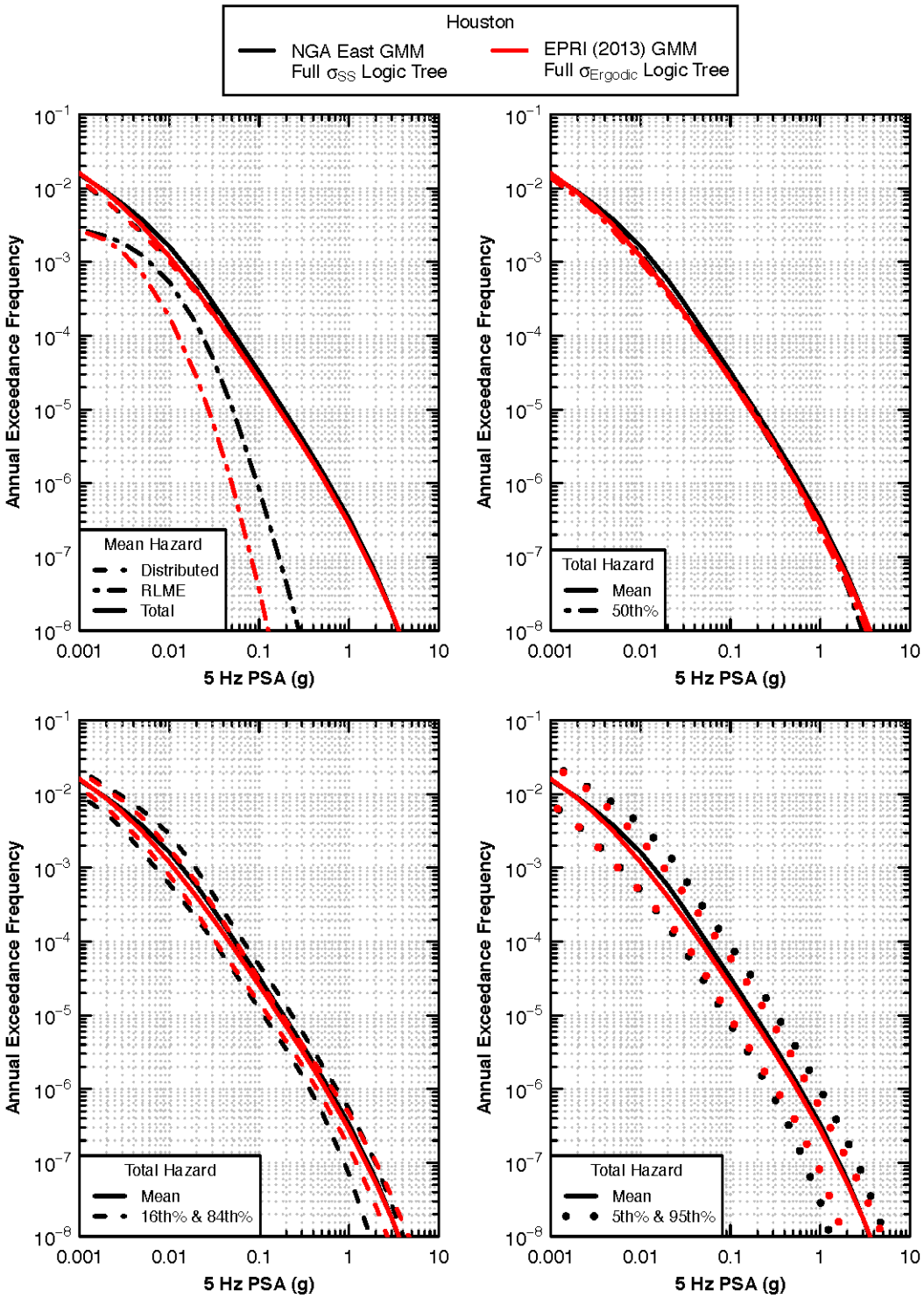


Figure 14–86 Comparison of 10 Hz PSA hazard at Houston computed using the complete NGA-East and EPRI (2013) GMMs.

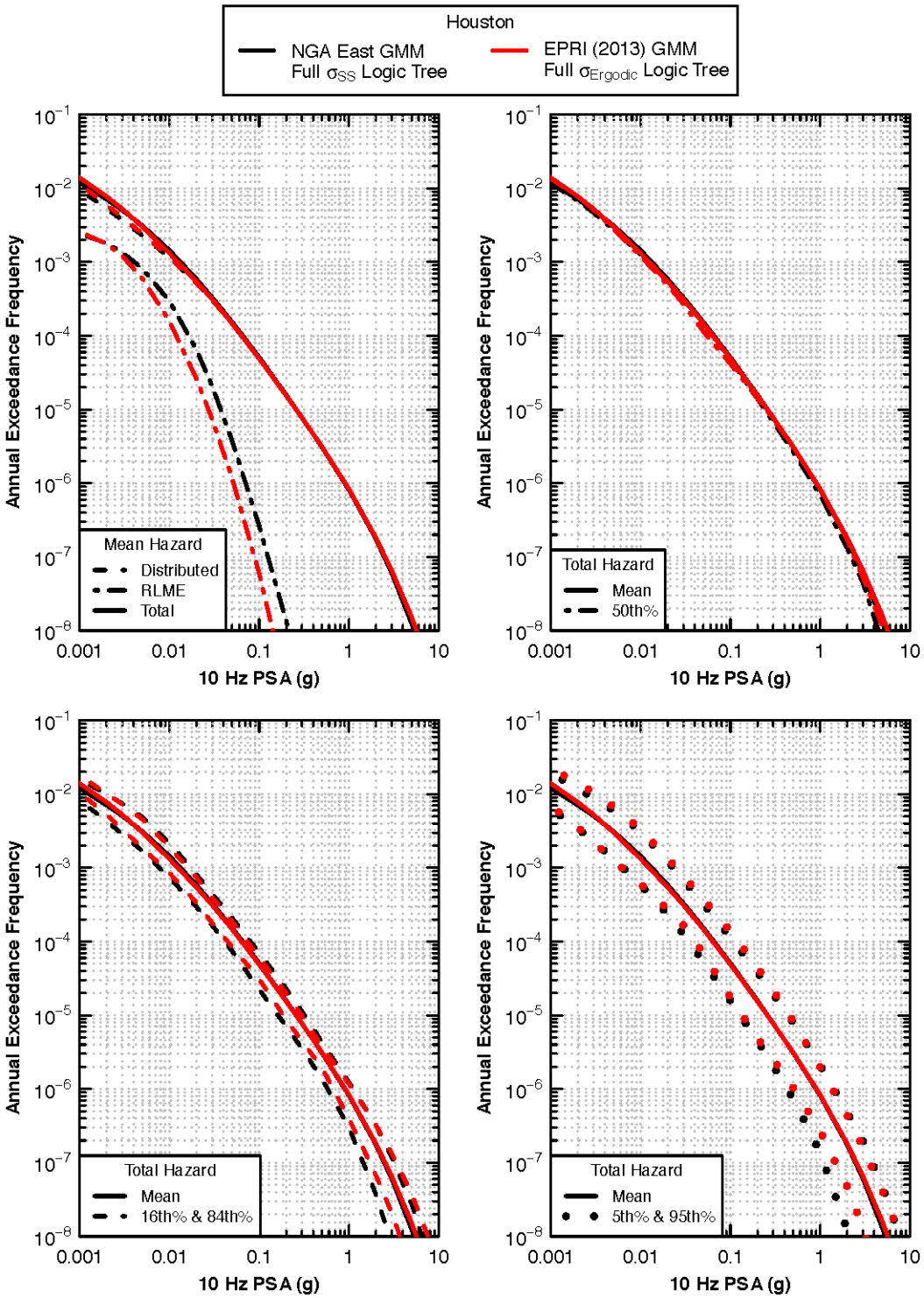


Figure 14–87 Comparison of 25 Hz PSA hazard at Houston computed using the complete NGA-East and EPRI (2013) GMMs.

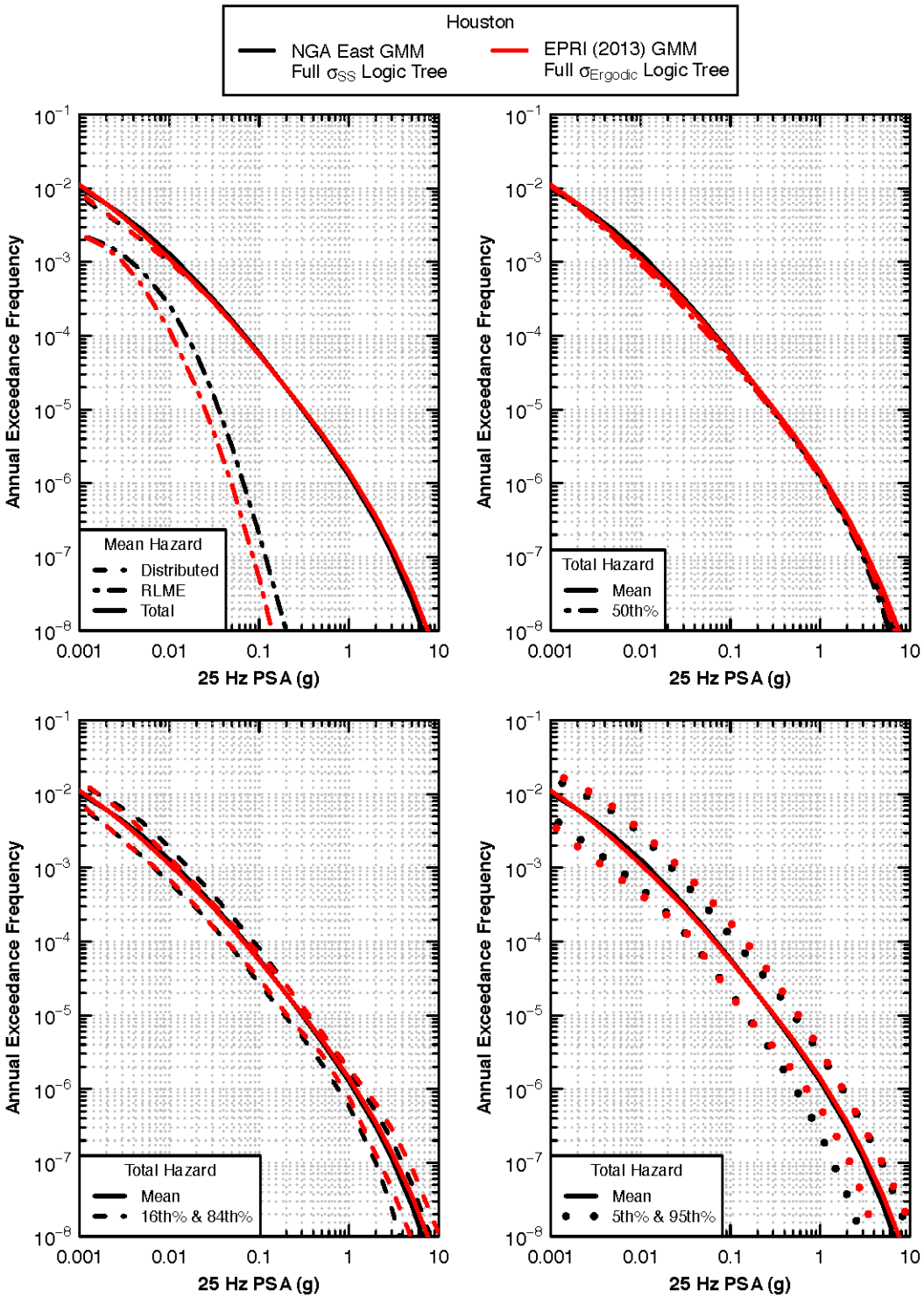


Figure 14–88 Comparison of PGA hazard at Houston computed using the complete NGA-East and EPRI (2013) GMMs.

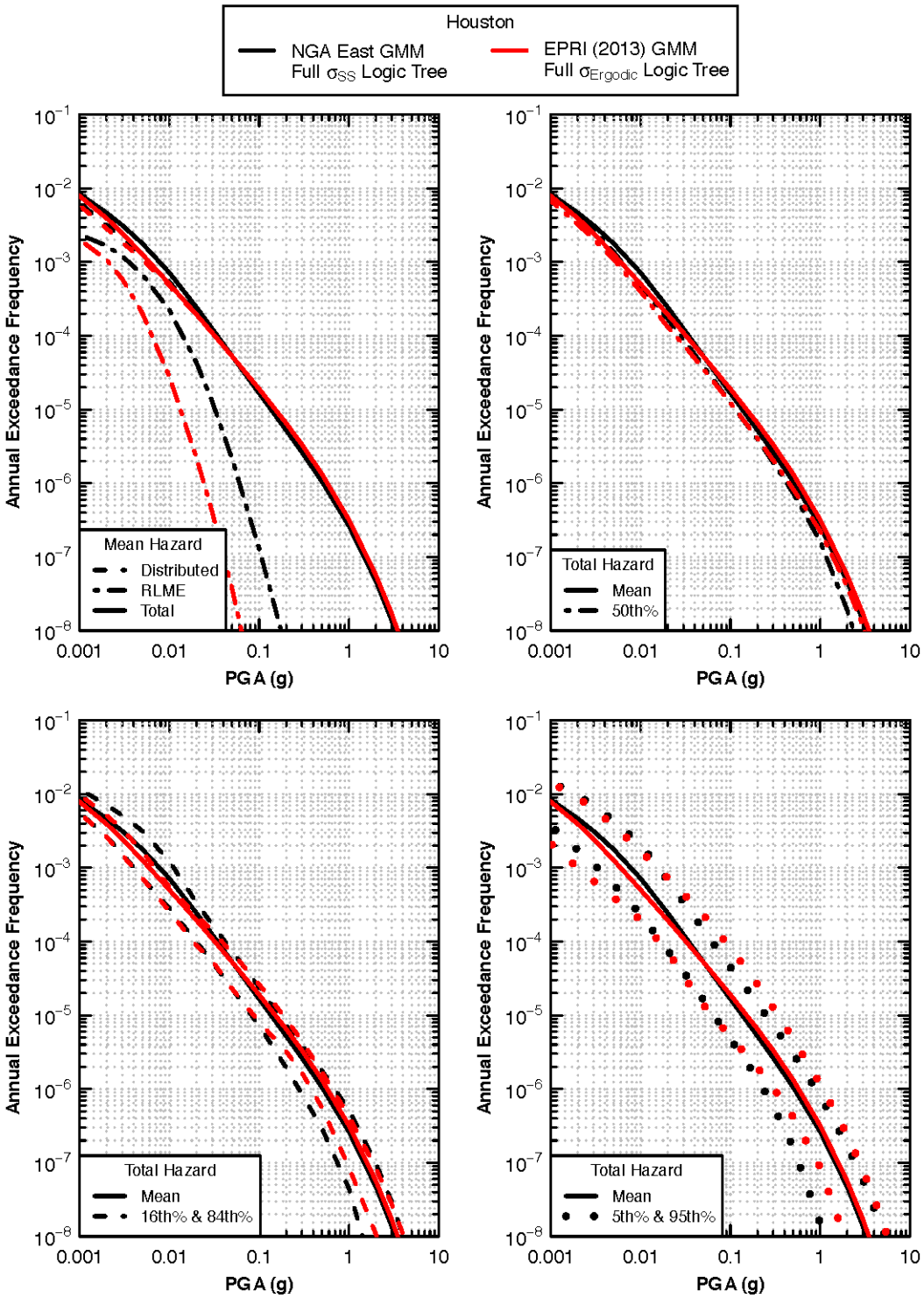


Figure 14–89(a) Ratio of mean hazard at Houston computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

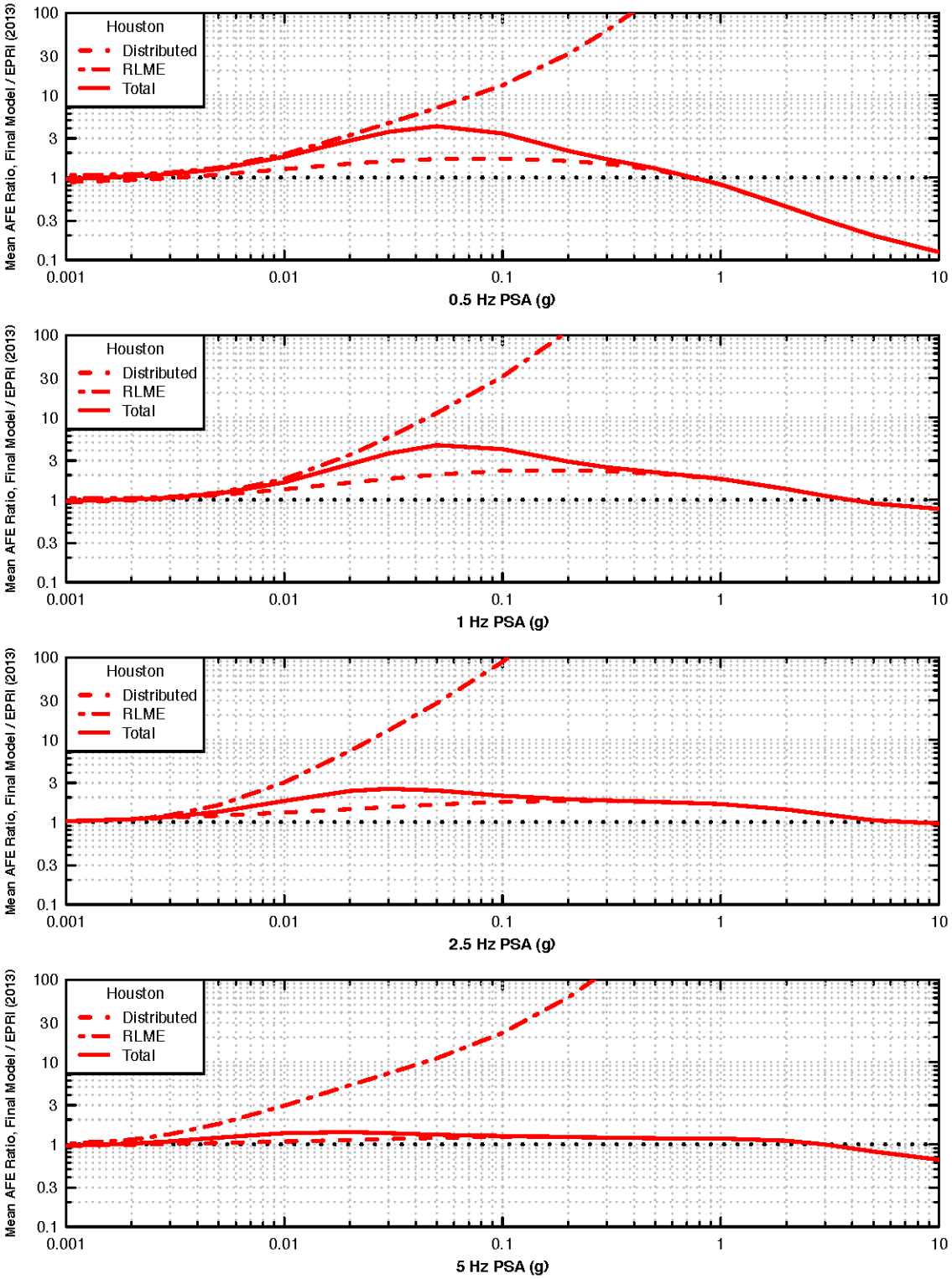


Figure 14–89(b) Ratio of mean hazard at Houston computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

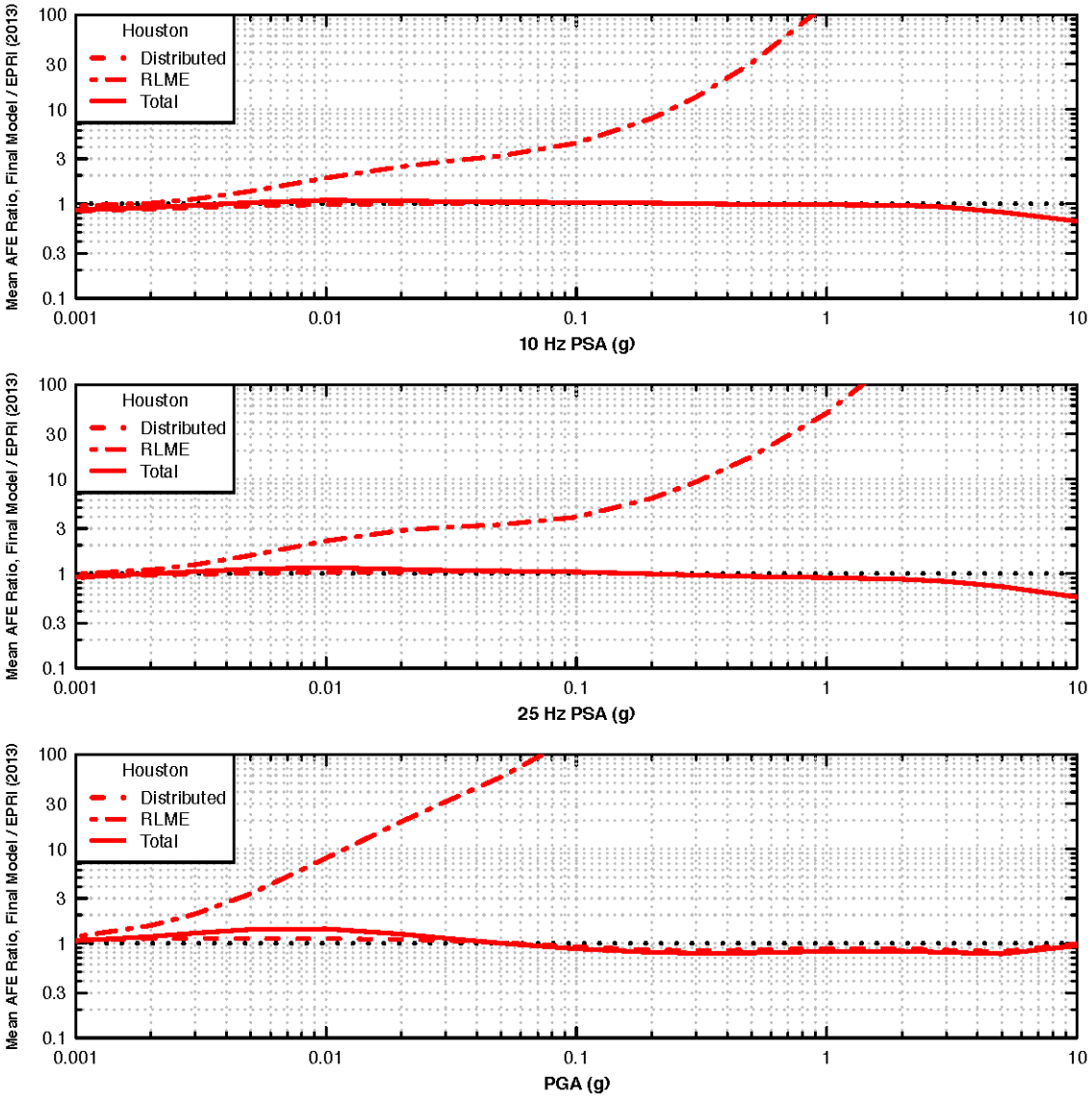


Figure 14–90 Comparison of 0.5 Hz PSA hazard at Jackson computed using the complete NGA-East and EPRI (2013) GMMs.

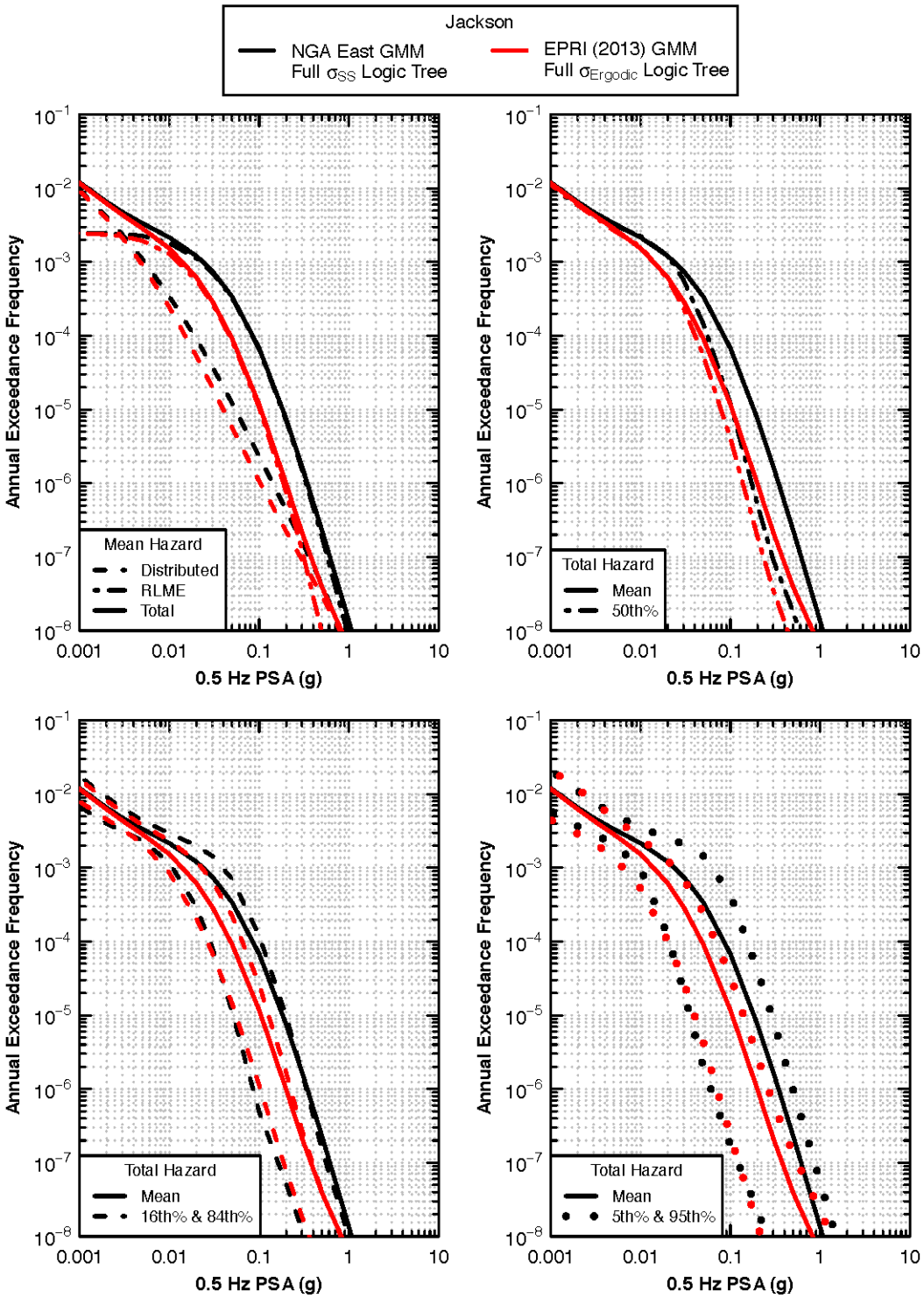


Figure 14–91 Comparison of 1 Hz PSA hazard at Jackson computed using the complete NGA-East and EPRI (2013) GMMs.

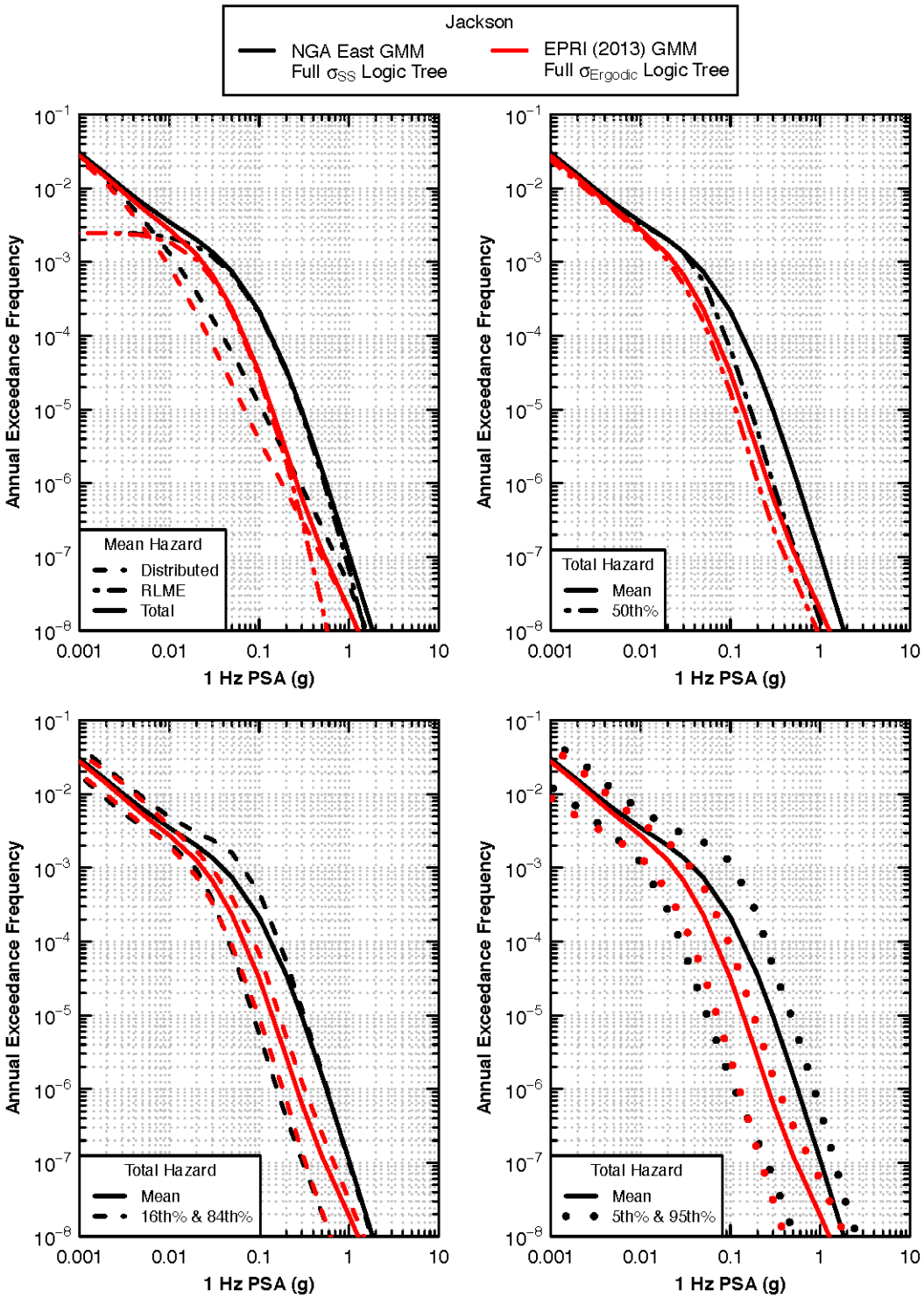


Figure 14–92 Comparison of 2.5 Hz PSA hazard at Jackson computed using the complete NGA-East and EPRI (2013) GMMs.

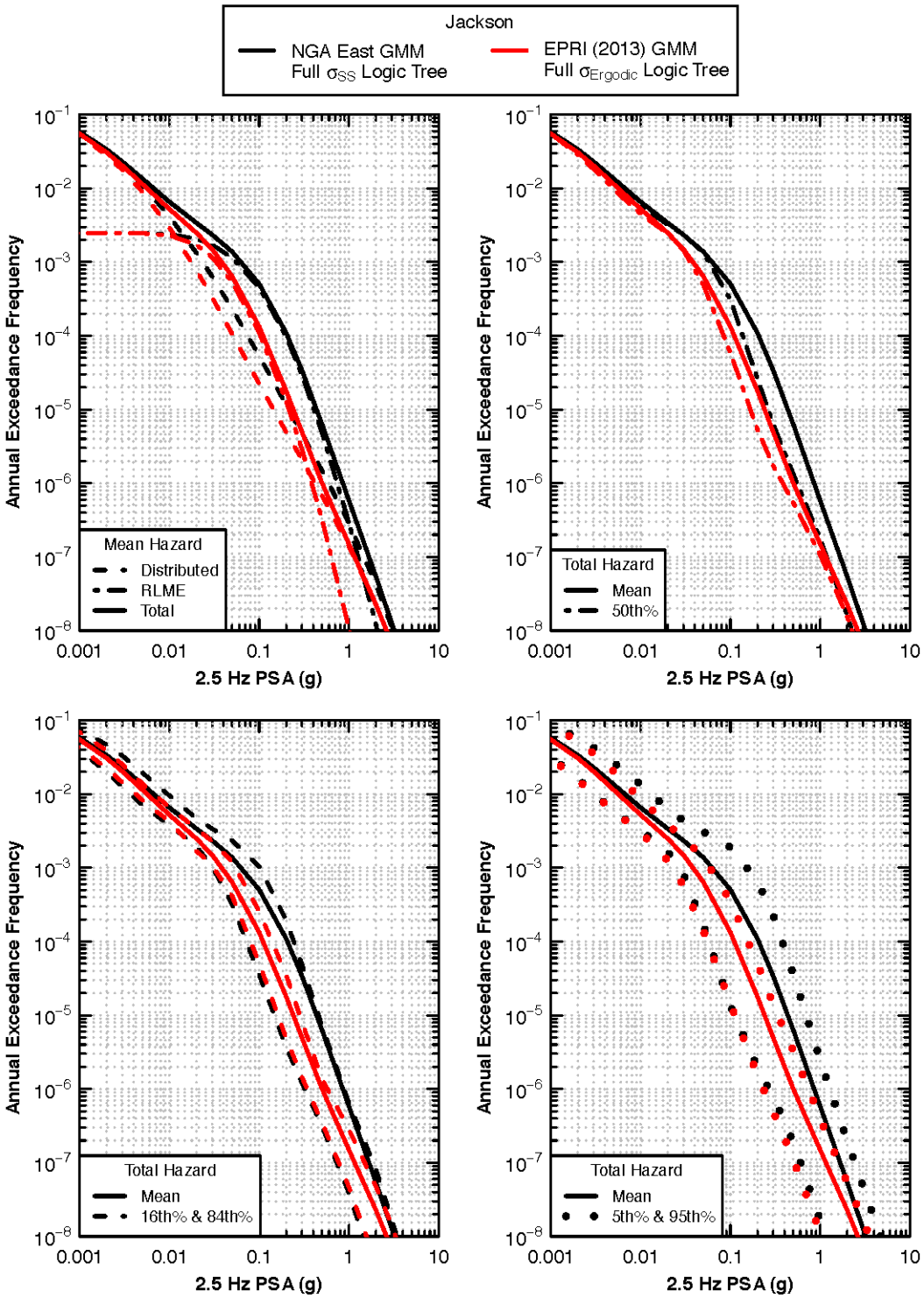


Figure 14–93 Comparison of 5 Hz PSA hazard at Jackson computed using the complete NGA-East and EPRI (2013) GMMs.

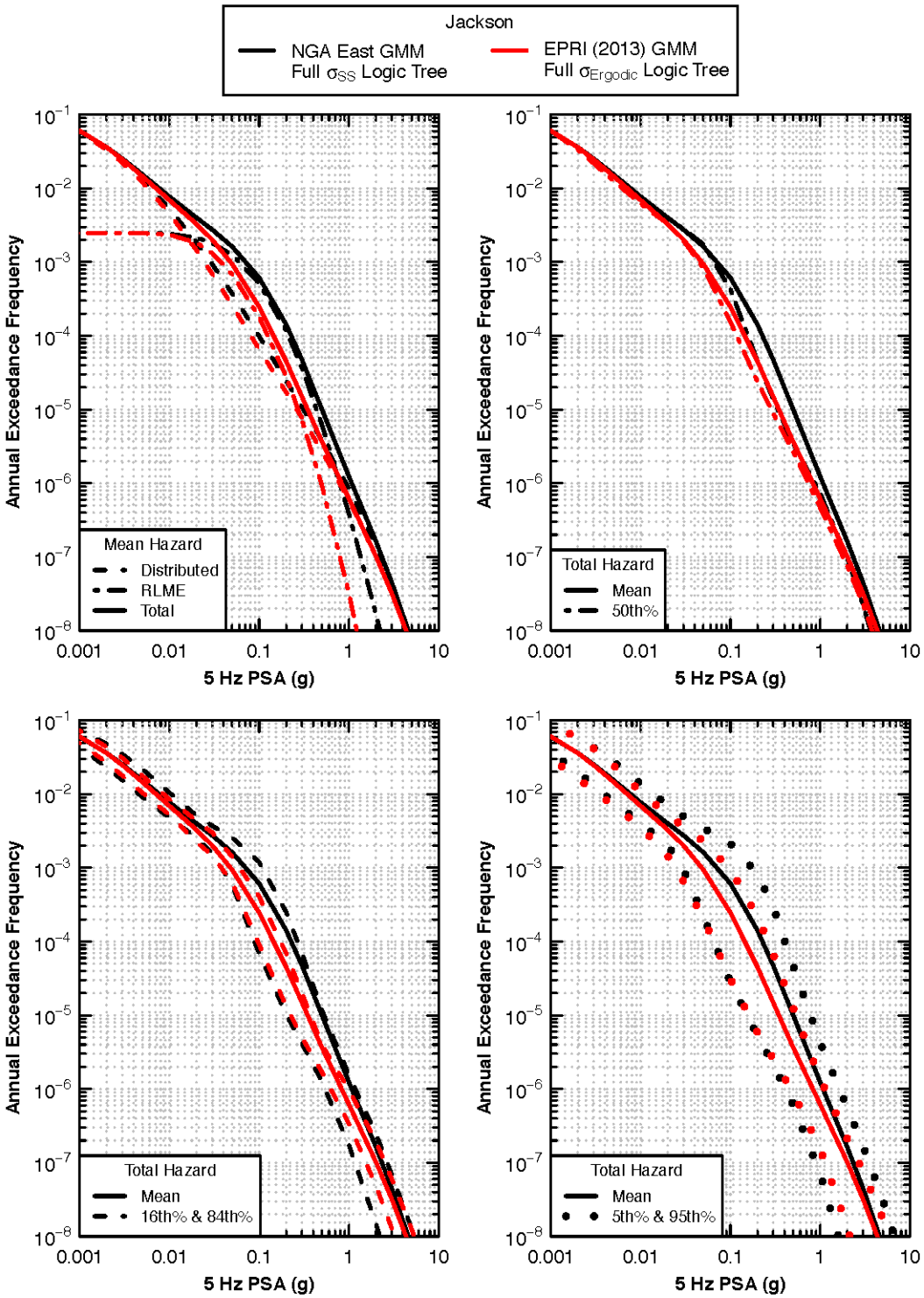


Figure 14–94 Comparison of 10 Hz PSA hazard at Jackson computed using the complete NGA-East and EPRI (2013) GMMs.

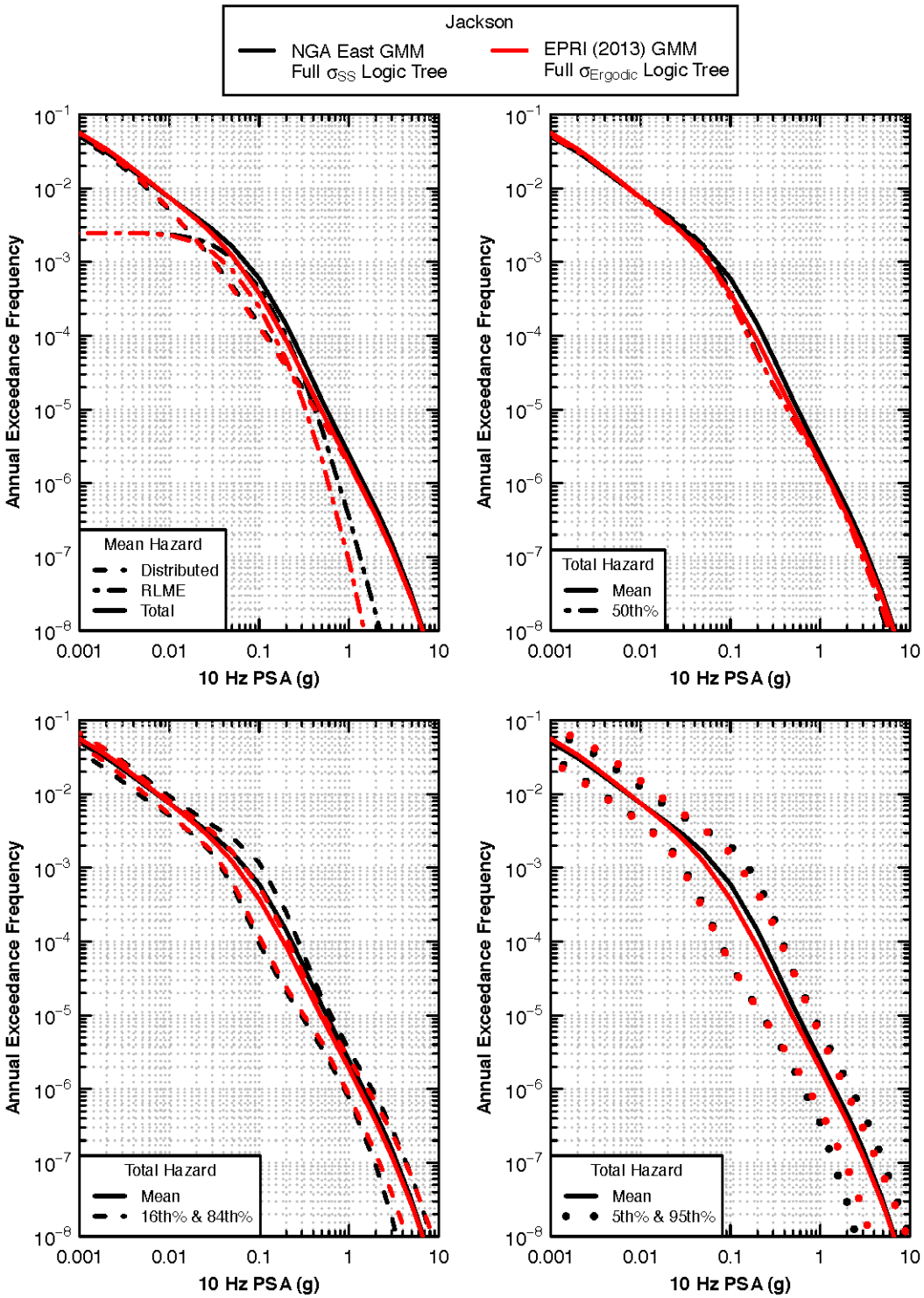


Figure 14–95 Comparison of 25 Hz PSA hazard at Jackson computed using the complete NGA-East and EPRI (2013) GMMs.

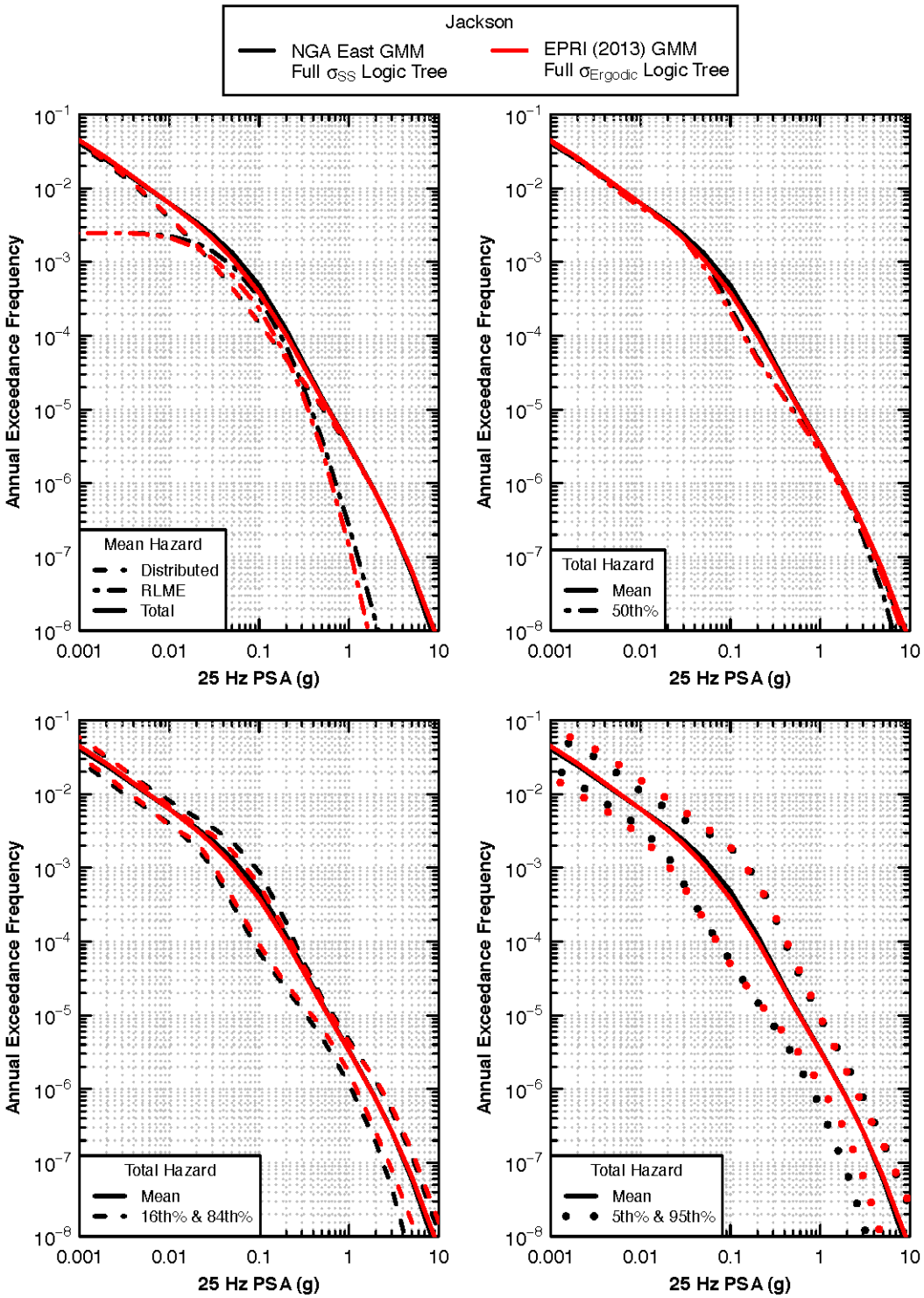


Figure 14–96 Comparison of PGA hazard at Jackson computed using the complete NGA-East and EPRI (2013) GMMs.

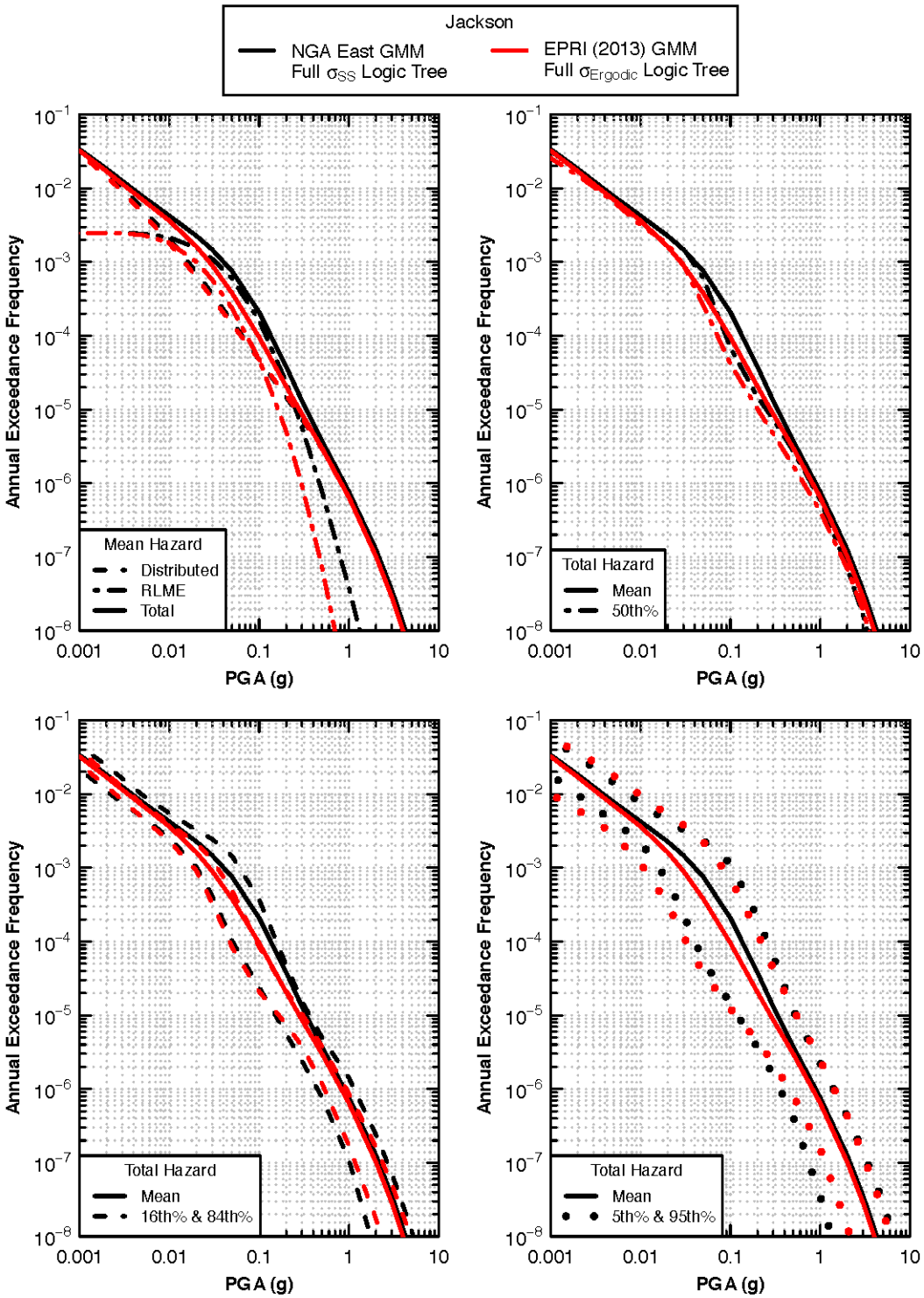


Figure 14–97(a) Ratio of mean hazard at Jackson computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

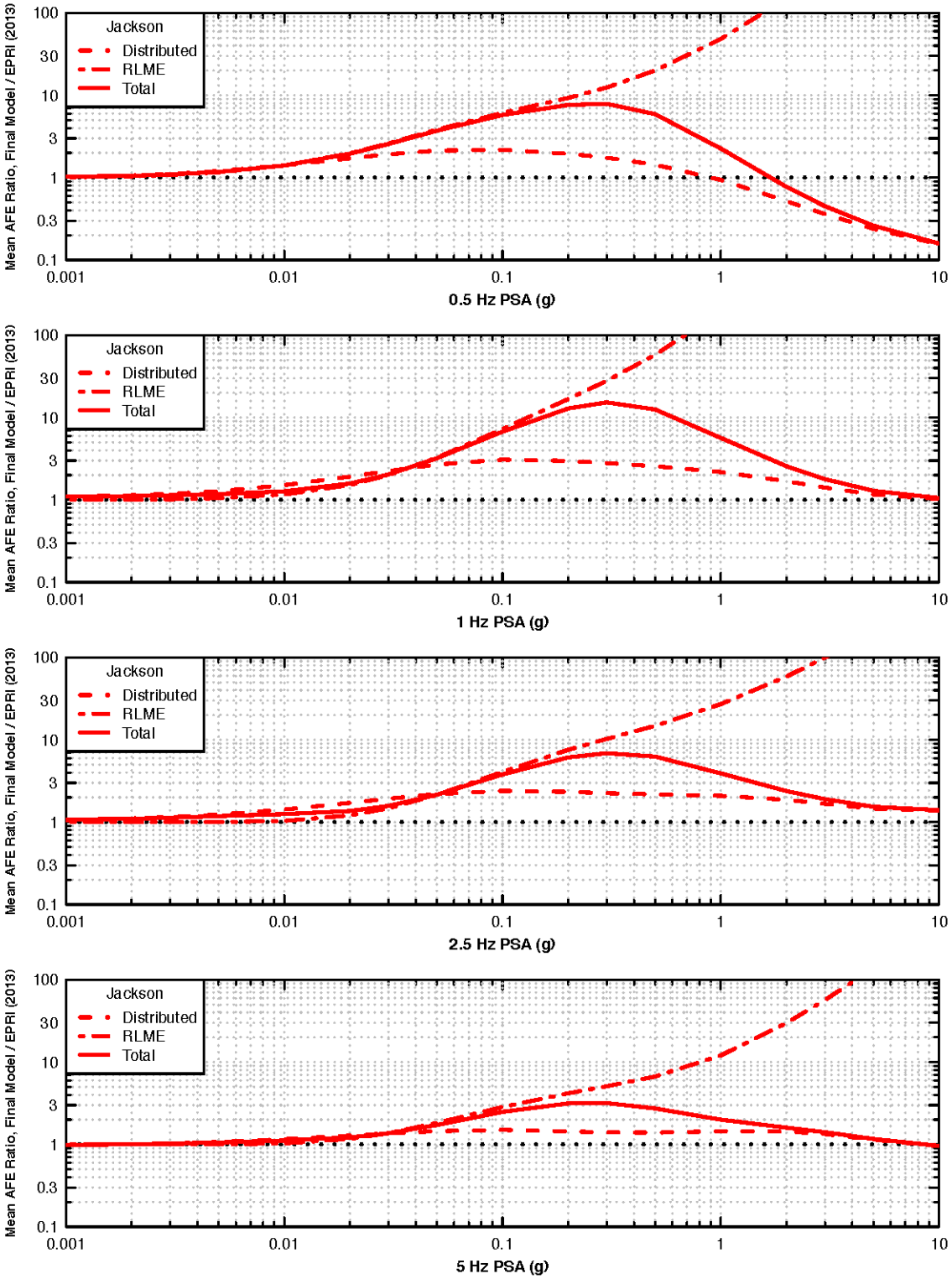


Figure 14–97(b) Ratio of mean hazard at Jackson computed using the complete NGA-East GMM to that computed using the EPRI (2013) GMM.

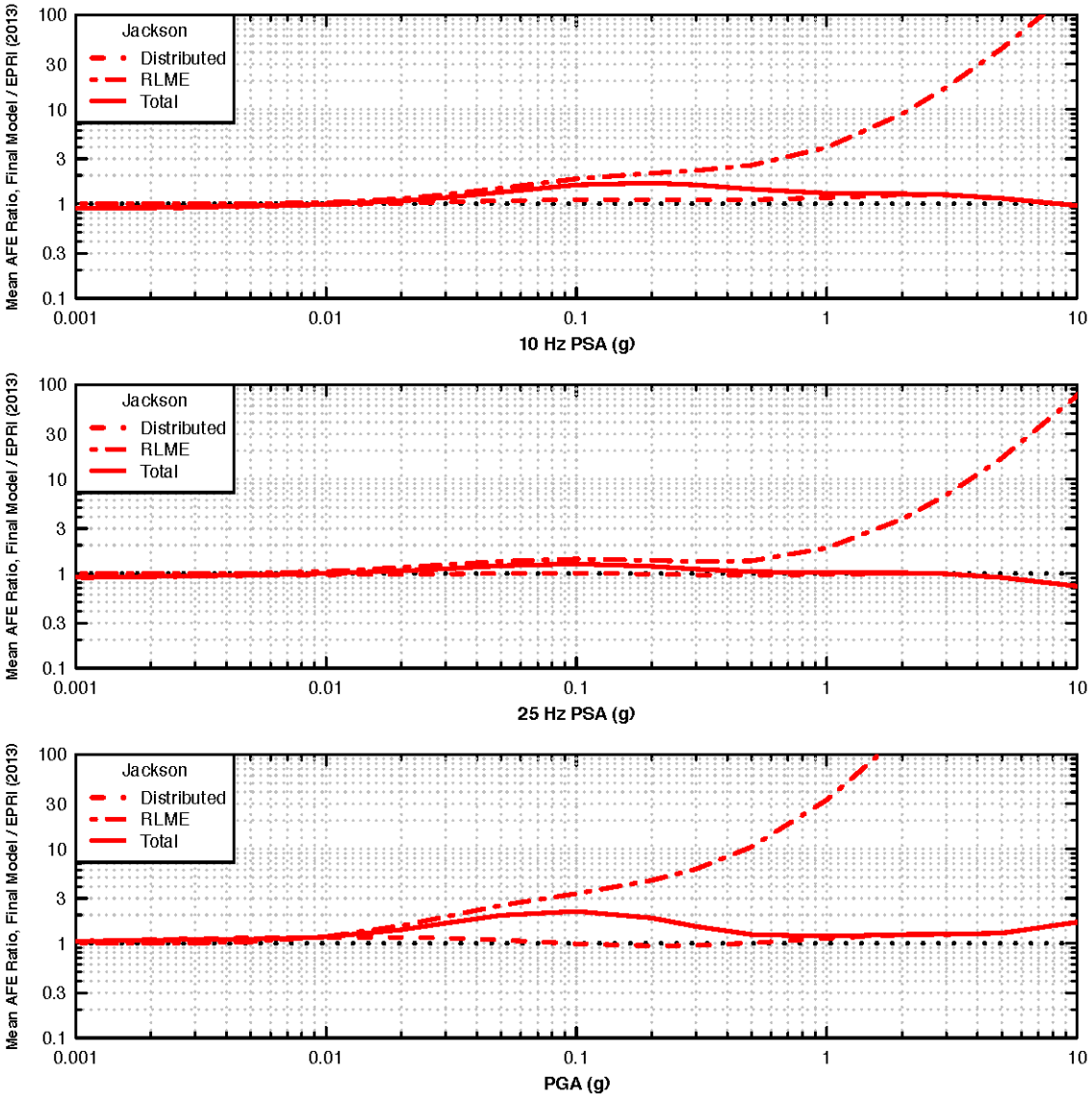


Figure 14–98 Ratio of UHRS computed using the NGA-East and EPRI (2013) GMMs.

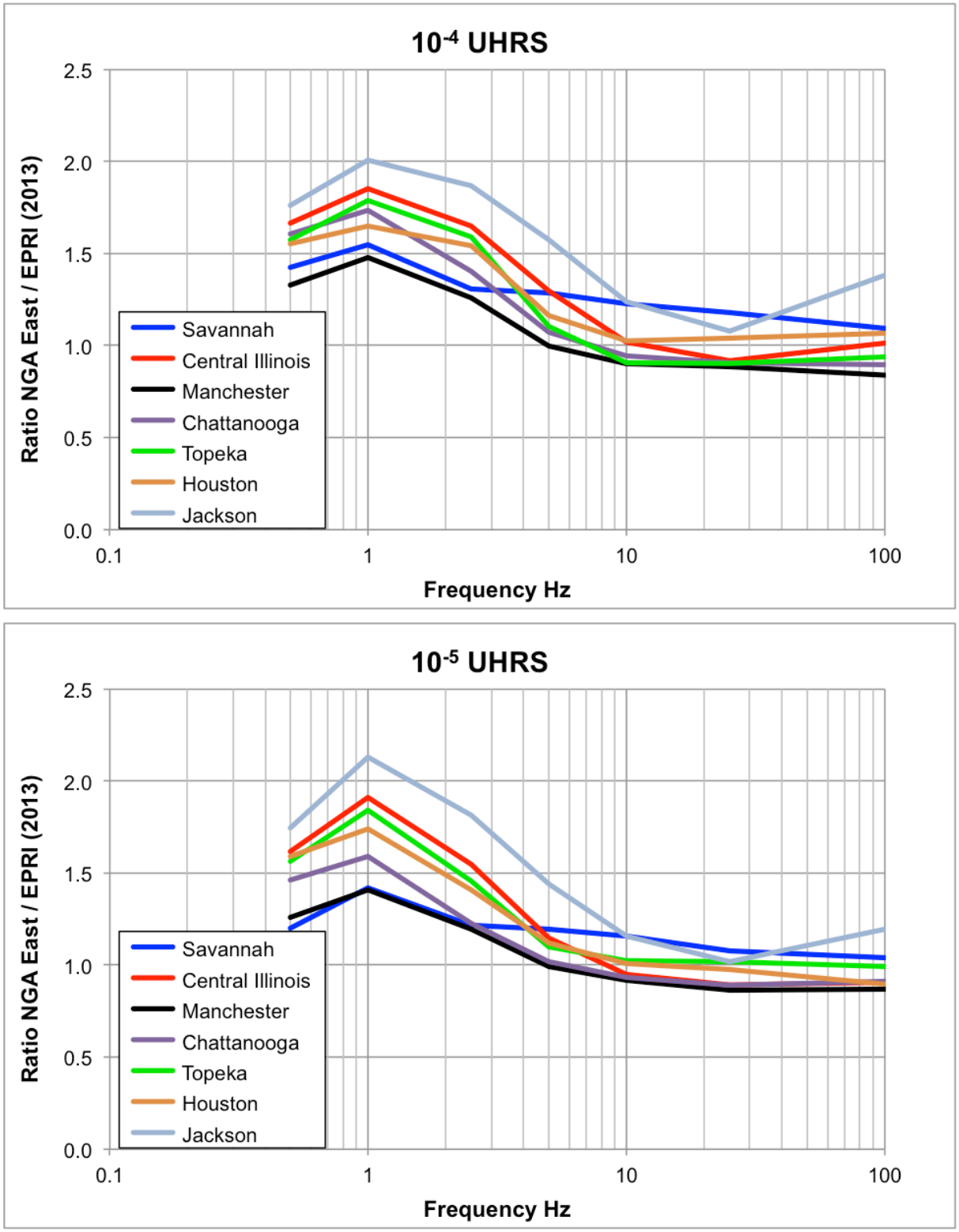


Figure 14–99 Comparison of the reference hard rock shear wave velocity recommended by Hashash et al. (2014) to the Vs profiles used in the development of the seed GMMs.

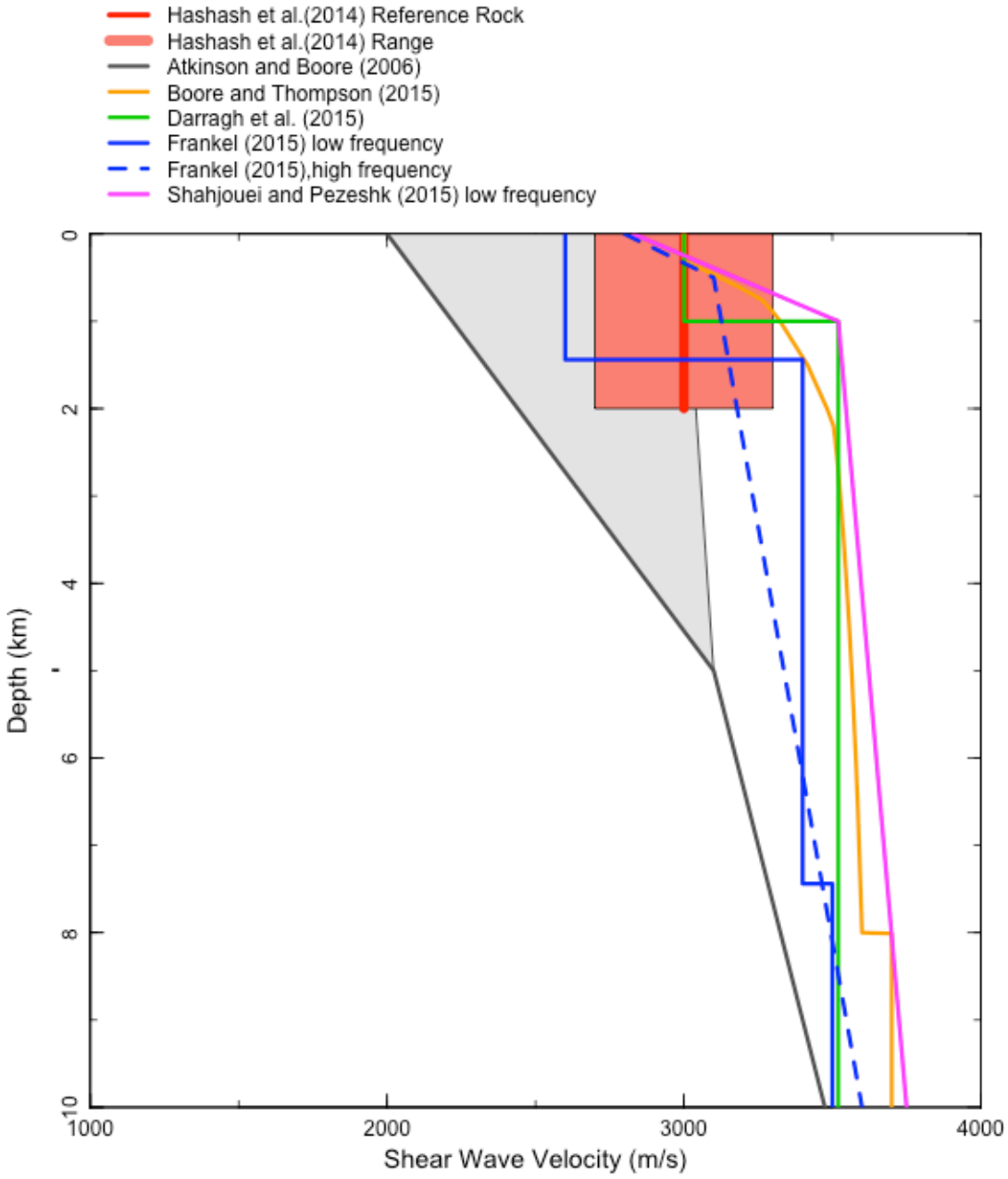
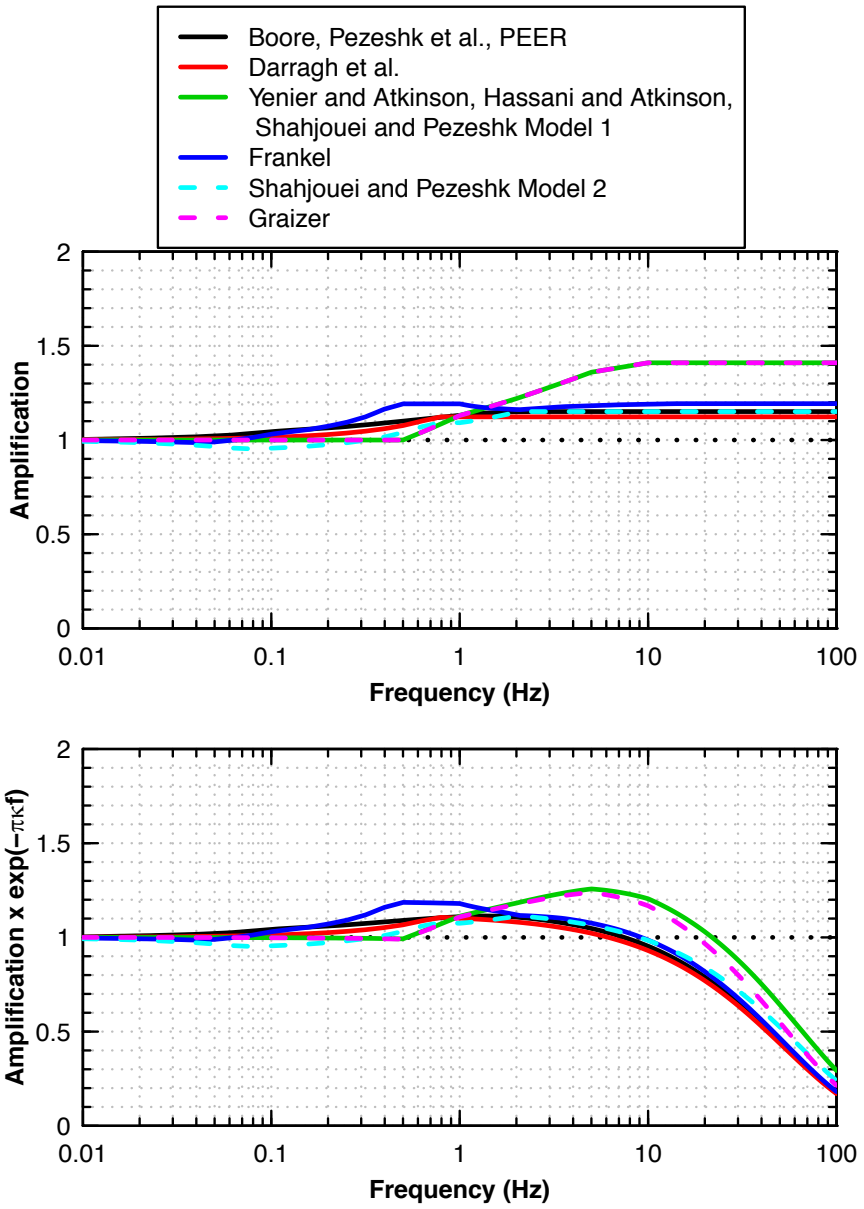


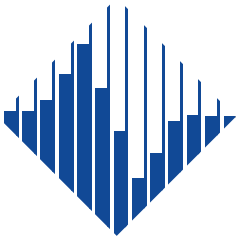
Figure 14–100 Site amplification functions incorporated in the NGA seed models.



The Pacific Earthquake Engineering Research Center (PEER) is a multi-institutional research and education center with headquarters at the University of California, Berkeley. Investigators from over 20 universities, several consulting companies, and researchers at various state and federal government agencies contribute to research programs focused on performance-based earthquake engineering.

These research programs aim to identify and reduce the risks from major earthquakes to life safety and to the economy by including research in a wide variety of disciplines including structural and geotechnical engineering, geology/seismology, lifelines, transportation, architecture, economics, risk management, and public policy.

PEER is supported by federal, state, local, and regional agencies, together with industry partners.



PEER Core Institutions

University of California, Berkeley (Lead Institution)
California Institute of Technology
Oregon State University
Stanford University
University of California, Davis
University of California, Irvine
University of California, Los Angeles
University of California, San Diego
University of Nevada, Reno
University of Southern California
University of Washington

PEER reports can be ordered at https://peer.berkeley.edu/peer-reports or by contacting

Pacific Earthquake Engineering Research Center
University of California, Berkeley
325 Davis Hall, Mail Code 1792
Berkeley, CA 94720-1792
Tel: 510-642-3437
Email: peer_center@berkeley.edu

ISSN 1547-0587X