A comparison of RMS-to-peak factors using simulated and recorded earthquakes

Eric M. Thompson[1] and David M. Boore[2]

9 January, 2011

1 INTRODUCTION

This report has two goals: first, to summarize and clarify the role of RMS-to-peak conversion factors in random vibration theory (RVT), and second, to compare two different RMS-to-peak factors using simulated and recorded ground motions. For the second task, we look for trends in the differences between RMS-to-peak factors in terms of the 5% damped pseudo spectral acceleration (PSA). We look for trends as a function of the oscillator period (T) of a single-degree-of-freedom (SDOF) system, the earthquake magnitude (**M**), and the epicentral distance (R).

The differences in the PSA that result from the two RMS-to-peak factors that we address in this report are illustrated in Figure 1. Subfigures (a) and (b) show the mean PSA from 100 ground-motion simulations of an **M** 6 earthquake at distances of 5 km and 50.7 km, respectively. The simulation method is described in section 2.1 and the RVT methods to compute the PSA from the ground acceleration are described in section 2.2. Subfigures (c) and (d) show the residuals, defined as the ratio of the two different PSA curves that are predicted using spectral moments calculated from simulated ground motions, along with the RMS-to-peak ratio from RVT; we use two versions of this ratio: the curve labeled B03 is the ratio described by Boore (2003) and the curve labeled DK is the ratio described by Der Kiureghian (1979; 1983); these two ratios are described in detail later. The PSA curves are generated from the same set of 100 simulated time series and the only difference is RMS-to-peak factors. Thus, the differences in PSA can only be

[1] Tufts University, Medford, MA, USA; eric.thompson@tufts.edu
[2] U. S. Geological Survey, Menlo Park, CA, USA

attributed to the different RMS-to-peak methods. We see from subfigures (a) and (b) that the largest differences in an absolute sense tend to occur near the maximum PSA. But looking at the PSA ratios, we see that important differences also occur longer periods.

We include both simulated and recorded ground motions in our analysis because each has an advantage over the other. The benefit of simulated time histories is that we can generate as many records as necessary for the values of T, **M**, and R of interest. Though the simulated records exhibit the proper frequency content and duration for a given **M** and R, the possibility remains that some other characteristic of earthquake ground motion will affect the RMS-to-peak factor (such as whether or not the time series are clustered). Thus, the appropriateness of conclusions based on simulations may be questioned. To address this issue, we also compare the results from different RMS-to-peak factors to recorded ground motions. The major drawback to this approach is that we are limited by the available data, and thus some earthquake parameters are under represented. In order to achieve a reasonable sample size for recorded data, we must choose data bins that are defined for a range of magnitudes and distances. We use records that have been collected and processed for the Next Generation of Attenuation (NGA) West-2 project.

The sample size (i.e., the number of recordings used to compare theoretical and empirical values) is a fundamental issue in this analysis. The RVT predictions should be accurate on average, but any individual time series may deviate significantly from the average response. For a given **M** and R, we compute 100 synthetic simulations. For recorded ground motions, we bin the records for a range of **M** and R. The magnitudes were chosen to span the range that is of primary interest to the PEGASOS project (**M** 4.5 to 7.5; Renault, pers. comm., 2010). The goal when choosing the distance bins was to achieve a minimum number of records of approximately 100 records in each bin while keeping the distance ranges constant across the different magnitude classifications (it was not possible to achieve this goal for all bins).

We begin by reviewing RVT and the methods for simulating earthquake time series. Our analysis begins with the stochastic simulation of ground motions. From these data we estimate the “true” (though simulated) PSA and compare the results to those obtained for two different RMS-to-peak factors. Then we repeat this analysis replacing

the simulated time series with recorded ground motions and analyze the trends in terms of the difference between the predicted and observed peaks, as well as the difference between the two different methods of predicting the peak motions.

2 METHODS

2.1 Stochastic Simulation of Ground Motions

For this report, we use the Boore (1983) method to generate ground motion simulations, as implemented in the program SMSIM (Boore, 2005). This method begins by generating a time series of band-limited random white Gaussian noise with zero expected mean. The variance is constrained such that the spectral amplitude is unity on average. The amplitude spectrum of the noise is multiplied by the target amplitude spectrum, which is then transformed back into the time domain. The target spectrum is defined from an earthquake source model, such as the omega-squared model (Aki, 1967). In this report, we use the Atkinson and Silva (2000) 2-corner source model and western generic rock amplifications (the parameter file for a sample run is given in Appendix 2). We do not expect that the results regarding the RMS-to-peak factors are sensitive to the assumed source model (it is beyond the scope of this project to test this assumption). Following Boatwright and Choy (1992), the lower corner frequency (f_a) of the 2-corner source model determines the source duration:

$$D_s = 0.5\, f_a^{-1}. \tag{1}$$

The total duration of the ground motion time domain window is

$$D_{gm} = D_s + 0.05R\ , \tag{2}$$

where $0.05R$ is the component of the duration due to path effects (Atkinson and Silva, 2000). The exponential window function used by SMSIM is

$$w(t) = a(t/t_\eta)^b \exp[-c(t/t_\eta)]\ , \tag{3}$$

where t_η is $2D_{gm}$, and

$$b = -\varepsilon \ln \eta / [1 + \varepsilon(\ln \varepsilon - 1)] \ ,$$
$$c = b / \varepsilon D_{gm} \ , \text{ and}$$
$$a = [e / \varepsilon D_{gm}]^b \ ,$$

where $\varepsilon = 0.2$ and $\eta = 0.05$ are based on the results from Saragoni and Hart (1974).

2.2 Random Vibration Theory

RVT uses the moments of the squared amplitude spectrum to relate the peak (x_{max}) to the RMS (x_{rms}) of a random time series $x(t)$. The *i*th spectral moment is defined as

$$m_i = 2\int_0^\infty (2\pi f)^i \left| X(f) \right|^2 df \ , \qquad (4)$$

where *f* is frequency and *X* is the amplitude spectrum of $x(t)$, which can be defined from a source and path model (e.g., Atkinson and Silva, 2000), or estimated from a recorded or simulated time series. Note that x_{rms} can be computed from the zeroth spectral moment

$$x_{rms} = \sqrt{m_0 / D_{rms}} \ , \qquad (5)$$

where D_{rms} is defined differently when $x(t)$ is a ground motion and when $x(t)$ is the response of a SDOF system. When $x(t)$ is a ground motion, D_{rms} is simply equal to D_{gm} (equation 2). When $x(t)$ is the response of a SDOF system, we use the Boore and Joyner (1984) method to compute D_{rms} from D_{gm} and *T*.

For simulated ground motions, D_{gm} can be computed from the source model and epicentral distance (equation 2); for the recorded NGA time series, however, D_{gm} is not known *a priori* and it must be computed from the observed time series. Thus, we estimate

D_{gm} from $\int x^2(t)dt$ (i.e., the Husid plot). To be consistent in our treatment of the simulated and recorded data, we also estimate D_{gm} from the Husid plot for the simulated data. The duration is computed from the times where the Husid curve reaches minimum and maximum values as a percent of the maximum of the Husid curve. The minimum and maximum percentages must be defined differently for the SMSIM simulations and recorded data. The reason for this is that SMSIM does not include surface waves, unlike the recorded data. Following Bommer *et al.* (2009), we use $D_{95\text{-}5}$, which is the time between 5% and 95% of the Husid curve, to estimate the duration of the full wave field for simulated time series. In contrast, we use $D_{75\text{-}5}$, to estimate only the duration of the body waves from time series of observed records that may contain surface waves (Bommer *et al.*, 2009).

We define the RMS-to-peak factor as

$$p = \frac{x_{\max}}{x_{rms}} \ . \tag{6}$$

We also must distinguish between two different methods of estimating p that have resulted in distinctly different estimates of PSA (see Renault, 2010; and Figure 1 of this report). We refer to the method developed by Boore (1983) and updated by Boore (2003) as p_{B03}. We refer to the method described by Der Kiureghian (1979) and updated by Der Kiureghian (1983) as p_{DK}. We also use a few modifications to p_{DK} described by Asfura (2008).

Boore (2003) provided an update to the equation for p_{B03} given by Boore (1983). Both equations are based on equation 6.8 in Cartwright and Longuet-Higgins (1956). The update by Boore (2003) removes the integrable singularity in equation 6.8 by applying the change of variable $z = \sqrt{\theta}$, yielding

$$p_{B03} = 2\int_0^\infty \left(1 - [1 - \xi\exp(-z^2)]^{Ne}\right)dz \ , \tag{7}$$

where $\xi = m_2(m_0 m_4)^{-1/2}$, and N_e is the number of extrema of $x(t)$, given by

$$N_e = \max\left(\frac{1}{\pi}\sqrt{m_4/m_2}D_{rms},\ 1.002\right). \quad (8)$$

Unlike Boore (1983; 2003), Der Kiureghian (1983) makes use of a shape factor (δ) to compute p_{DK} (denoted q in Der Kiureghian, 1979). It is a dimensionless ratio of spectral moments:

$$\delta = \sqrt{1 - \frac{m_1^2}{m_0 m_2}}\ . \quad (9)$$

This shape factor has a value between 0 and 1 and is a measure of the dispersion of $X(f)$ about its centroid. According to Der Kiureghian (1979), Vanmarcke (1975) showed that δ is a measure of the dependence between zero crossings; δ is inversely proportional to the cluster size of the zero crossings of $x(t)$. Thus, small values of δ indicate that the assumption of independence of zero crossings is violated (i.e., the Poisson model assumption). In this report, we employ an empirical modification of δ described by Asfura (2008): $\delta' = \max(\delta, \zeta)$, where ζ is the fraction damping of the SDOF system. It is beyond the scope of this report to assess the appropriateness and the effects of this modification. We apply it here because we believe it was used to produce the results discussed in Renault (2010) and resolving the differences between the contractor's results that were identified in Renault (2010) is the primary goal of this report.

In contrast to equation 7, Der Kiureghian (1983) defines the RMS-to-peak factor with the asymptotic form of the Cartwright and Longuet-Higgins (1956) integral

$$p_{DK} = \sqrt{2\ln N_z' D_{rms}} + 0.5772/\sqrt{2\ln N_z' D_{rms}}\ , \quad (10)$$

where N_z' is the number of zero crossings, which is computed from the number of zero crossings assuming independence (N_z) using the following empirical relationship:

$$N_z' = \begin{cases} \max[2.1, 2(\delta')N_z] & 0.00 < \delta' \le 0.10 \\ [1.63(\delta')^{0.45} - 0.38]N_z & 0.10 < \delta' \le 0.69 \\ N_z & 0.69 < \delta' \le 1.00 \end{cases}, \quad (11)$$

where

$$N_z = \frac{1}{\pi}\sqrt{m_2 / m_0}\, D_{rms} \ . \quad (12)$$

Similar to the minimum value applied to δ above, Asfura (2008) applies an empirical minimum value to N_z' of 2.1.

2.3 Definition of Residuals

The goal of this report is to analyze the difference between the two different RVT methods of estimating the peak motion from the RMS motion (p_{B03} and p_{DK}). The primary variable of interest for engineering design is PSA, so we define the residual as the ratio of the observed PSA to the predicted PSA:

$$\mathrm{PSA}^{rr}(T) = \mathrm{PSA}^{o}(T) / \mathrm{PSA}^{p}(T) \ . \quad (13)$$

Both PSA^{o} and PSA^{p} are computed from the relative oscillator displacement time series [$u(t)$], which is computed from either a simulated or a recorded acceleration time series [$a(t)$]. PSA generally follows a lognormal distribution, which is why we define the residual as a ratio in equation 13. A ratio also has the advantage that it is not sensitive the large range of PSA values across the period range of interest for a single SDOF response. We have found that for the smaller values of PSA, the arithmetic difference may appear negligible while PSA^{rr} may be as large as 2, indicating that the predicted values are half of the observed values. However, the arithmetic residual may be important for some applications, so we define

$$\mathrm{PSA}^{ar}(T) = \mathrm{PSA}^{o}(T) - \mathrm{PSA}^{p}(T) \ . \quad (14)$$

Note that the superscript *rr* in equation 13 indicates that it is the "residual ratio", while *ar* in equation 14 indicates that it is the "arithmetic residual".

To compute $\mathrm{PSA}^o(T)$, we first compute the relative displacement response spectra as

$$SD(T) = |u(t)|_{\max} \ , \tag{15}$$

from which $\mathrm{PSA}^o(T)$ is defined as

$$PSA^o(T) = (2\pi/T)^2 SD(T) \ , \tag{16}$$

and $u(t)$ is the SDOF response as a function of time for a specified period and damping (5% in this report).

The key difference between PSA^o and PSA^p is in the computation of *SD*. As indicated by equation 15, PSA^o uses the actual maximum of the absolute value of $u(t)$. In defining PSA^p, our goal is to isolate the effects of the RMS-to-peak factor (*p*; equation 6). Thus, we compute the u_{rms} from $u(t)$ with equation 5, from which we estimate *SD* with the two different RMS-to-peak factors:

$$SD_{B03}(T) = u_{rms} p_{B03} \ , \text{ and} \tag{17}$$

$$SD_{DK}(T) = u_{rms} p_{DK} \ , \tag{18}$$

which yields two different PSA^p:

$$PSA^P_{B03}(T) = (2\pi/T)^2 SD_{B03}(T) \ , \text{ and} \tag{18}$$

$$PSA^P_{DK}(T) = (2\pi/T)^2 SD_{DK}(T) \ . \tag{19}$$

The RMS and peak factors make use of spectral moments computed from Fourier spectra of the windowed portion of the oscillator time series, using equation 4. The windowed

oscillator time series begins at the time where the Husid plot reaches 5% of its maximum and has duration D_{rms} (as defined by Boore and Joyner, 1984). Since there are two different PSA^{p} which can be used to compute PSA^{rr} and PSA^{ar}, there are four different PSA residual definitions: PSA^{rr}_{B03}, PSA^{ar}_{B03}, PSA^{rr}_{DK}, and PSA^{ar}_{DK}. For each residual, we report the geometric mean at a given period. For the simulations, we generate 100 samples, while the number of records varies for each magnitude-distance bin for the recorded data. In the Results section, we focus on PSA^{rr}, but we include analogous plots for PSA^{ar} in the Appendix 1 for comparison.

3 RESULTS

We are primarily concerned with earthquakes near **M** 6 because deaggregation indicates that this is the magnitude range that is most important for the hazard in Switzerland (Renault, pers. comm., 2010). For this reason we use stochastically simulated records from an **M** 6 earthquake and recorded earthquakes from **M** 5.5 to 6.5. To investigate the magnitude dependence of the results, we also present results for stochastically simulated records from **M** 5 and **M** 7 earthquakes, and analogously, records from recorded ground motions from events within **M** 4.5-5.5 and **M** 6.5-7.5 bins. The events for each of these bins are described in Tables 1-3, which also give the number of records that we use from each event.

The accuracy of the time-domain simulations is an important issue to address before we analyze the different RMS-to-peak factors. This is a consistency check between the SMSIM time-domain simulations and the target spectrum that these simulations are trying to match. Any mismatch that we observe here cannot be attributed to the different RMS-to-peak factors.

We compare PSA^{p}_{B03} to the target PSA from the SMSIM ground-motion simulations to the target PSA computed from the Atkinson and Silva (2000) model for an **M** 6 earthquake in Figure 2. The left column of plots is for a distance of 5 km and the right column of plots is for a distance of 50.7 km. The key difference between the two curves in subfigures (a) and (b) is that PSA^{p}_{B03} is computed from 100 time domain (TD)

simulations, while the target PSA (labeled RV) are based solely on the amplitude spectra predicted by the Atkinson and Silva (2000) model. Subfigures (c) and (d) plot the arithmetic difference between the RV and TD curves, whereas subfigures (e) and (f) give the ratio of the TD curve to the RV curve. It is interesting to note that different conclusions about the difference between the TD and RV curves will be reached in some cases if the primary concern is the arithmetic difference (subfigures c and d) or the ratio (subfigures e and f). As an example, compare the two different residuals at 3 sec for the R=50.7 km plots (subfigures d and f). The arithmetic difference is negligible, while the residual ratio shows a significant difference, indicating that the TD PSA is 5-10% larger than the target PSA.

3.1 Simulated Time Series

In this section, we assume that the "true" or "observed" peaks are those in the SMSIM time-domain simulations. We plot PSA^{rr}_{B03} and PSA^{rr}_{DK} for simulated **M** 5, 6 and 7 earthquakes at six different distances in Figures 3, 4, and 5, respectively (analogous figures can be found in Appendix A1 for PSA^{ar}_{B03} and PSA^{ar}_{DK}). The distances for the subfigures were chosen to be approximately the mid-point of the range of the distance bins for the NGA data. These results indicate that there is negligible distance dependence of PSA^{rr} for the SMSIM simulated time series.

Values of PSA^{rr} greater than one indicate that the predicted peaks are smaller than the observed. Thus, we see that p_{DK} consistently predicts smaller peaks than p_{B03} (consistent with Figure 1 and the results described by Renault, 2010). Additionally, we note that both the p_{DK} method and the p_{B03} method give nearly identical results for PGA, both of which consistently under-predict the observed value by 5-10%.

Figures 3-5 also indicates that the residuals exhibit a negative slope with respect to period. The predicted peaks of both methods are generally within 10% of the observed peaks for periods between 0.04 and 3 sec. The p_{B03} method tends to be more accurate at short periods (0.04-0.2 sec) whereas the p_{DK} method tends to be more accurate at long periods (1-3 sec). For intermediate periods, the observed peaks tend to fall between the two RMS-to-peak methods.

3.2 NGA Records

In this section, we assume that the "true" or "observed" peaks are from the NGA West-2 records. Figures 6-8 are similar to Figures 3-5 except that PSA^{o} is computed from time series from the NGA West-2 database. Additional differences are: (1) we include earthquakes with magnitudes ranging from +/-0.5 of the magnitudes in Figures 6-8, and (2) each subfigure includes a range of Joyner-Boore distances (R_{JB}). For those recordings for which R_{JB} has not been included in the NGA West-2 flatfile, we use the epicentral distance (this was only needed for earthquakes in the **M** 4.5-5.5 bin). The number of ground motions included in each bin is reported in the title of each subfigure (analogous figures can be found in Appendix 1 for PSA^{ar}_{B03} and PSA^{ar}_{DK}). We do not use the vertical component records, and we process the two horizontal components as independent observations.

Generally, we see similar trends in the PSA^{rr} for the NGA data as we saw in the simulations. An unexpected difference is that the two different RMS-to-peak methods give more similar results for the NGA data than for the simulated data. We expected that the two methods would be more similar for the simulated data because there is no reason to expect that the zero crossings are clustered. In the absence clustering, the two methods should give similar results because the fundamental difference between the methods is the use of the shape factor δ which is meant to account for clustering in the time series.

The distance-dependence of PSA^{rr} is another difference between the results from simulated records in Figures 3-5 and the NGA records in Figures 6-8. While PSA^{rr} appears not to be influenced by distance for the simulated data, the PSA^{rr} do exhibit trends with distance for the NGA records. The PSA^{rr} values tend to increase with R_{JB}; the PSA^{rr} in subfigure (a) of Figures 6-8 are centered around unity, whereas PSA^{rr} are generally greater than one for almost all periods in subfigures (d) through (f) of Figures 6-8 (distances greater than 40 km). Since the p_{B03} method generally predicts larger values than the p_{DK} method, the p_{B03} method tends to give better estimates at longer distances. This observation should be qualified by noting that these results are sensitive to the estimate of D_{gm}. For recorded data we use $D_{75\text{-}5}$ to estimate D_{gm}, which attempts to remove the influence of surface waves. The influence of surface waves tends to increase

with distance, and thus the estimation of D_{gm} is likely to be less accurate as distance increases. This may be the underlying cause of the distance trends noted above.

We also wish to directly address trends in the differences between the Boore (1983; 2003) method and the Der Kiureghian (1979; 1983) method (ignoring the "true" values). Thus, we plot the ratio of PSA^{p}_{B03} to PSA^{p}_{DK} for the NGA data in Figure 9. Figure 9 (a) gives the ratio for **M** 4.5-5.5, (b) gives the ratio for **M** 5.5-6.5, and (c) gives the ratios for **M** 6.5-7.5. The data in each magnitude bin is also categorized by the same distance bins as in Figures 6-8. We see very similar trends across the different magnitude bins: the ratios are near unity at 0.04 and 3 sec, and increase to a maximum of 1.05-1.10 near 0.2 to 0.3 sec. The maximum of the ratio decreases slightly as magnitude increases. The maximum B03/DK ratio in the $120 < R_{JB} < 200$ km is smaller and located at a larger period than the other distance bins for the **M** 5.5-6.5 and **M** 6.5-7.5 data. The $120 < R_{JB} < 200$ km curve for **M** 4.5-5.5 only has 18 samples and so it should not be considered reliable.

4 SUMMARY

This report investigates the consequences of the differences between the Boore (1983; 2003) method of computing peak motions from RMS motions and the Der Kiureghian (1979; 1983) method. The key difference is that the Der Kiureghian method uses a shape factor that is based on the spectral moments to account for clustering of the zero-crossings of the time series.

We compare the peaks computed with the Der Kiureghian method and the Boore method using both simulated and recorded time series. The comparisons of the two approaches to simulated motions (Figures 3-5) indicate that the Boore method may be more accurate at short periods (0.04-0.2 sec) while the Der Kiureghian method may be more accurate for long periods (1-3 sec). For intermediate periods, the Boore method tends to over-predict the peak motions by 5-10% and the Der Kiureghian method tends to under-predict the peak motions by 5-10%. As indicated by Figure 2, differences this small are near the limit of the accuracy of the RVT approach, as indicated by the differences between the RVT time domain PSA and the target PSA.

One could question the appropriateness of the conclusions based on simulated data because the clustering of the SMSIM time series may not be representative of real earthquake motions. For this reason we also studied recordings of ground motion. The comparisons to the NGA records (Figures 6-8) indicate that the differences between the two methods are smaller than for the simulated data (e.g., Figure 9 vs Figure 1). It is not clear which method is more accurate because the overall magnitude of the residuals is sensitive to the duration of shaking, which is not known with precision. The trends in the observed data, however, are similar to the trends in the simulated data: the Boore method consistently predicts larger peaks than the Der Kiureghian method. The differences approach unit at small periods (0.04 sec) and larger periods (3 sec) with a maximum of 5-10% at approximately 0.2 to 0.3 sec. The trend is relatively consistent across the magnitude-distance bins that we analyzed in this report for both simulated and recorded motions: **M** 4.5-7.5, and R_{JB} between 0 and 200 km.

ACKNOWLEDGEMENTS

We thank Norm Abrahamson for valuable discussions that informed and motivated this report. We also thank Philippe Renault for providing documentation of the results and methods provided by the different PEGASOS contractors and Technical Note QA-TN-1118 that summarizes the findings RVT reference test case verification. The Technical Note provided the motivation for this report.

REFERENCES

Aki, K. (1967). Scaling law of seismic spectrum, *J. Geophys. Res.* 72, 1217-1231.

Asfura, A. (2008). Appendix A: Random Vibration Theory used in SHAKE program, letter to Dr. Klügel provided to the authors by Swissnuclear on 10 December, 2010.

Atkinson, G. M., and W. Silva (2000). Stochastic modeling of California ground motions, *Bull. Seism. Soc. Am.* 90, 255-274.

Boatwright, J., and G. L. Choy (1992). Acceleration source spectra anticipated for large earthquakes in northeastern North America, *Bull. Seism. Soc. Am.* 82, 660-682.

Bommer, J. J., P.J. Stafford, and J. E. Alarcón (2009). Empirical equations for the prediction of the significant, bracketed, and uniform duration of earthquake ground motion, *Bull. Seism. Soc. Am.* 99, 3217-3233.

Boore, D. M. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, *Bull. Seism. Soc. Am.* 73, 1865-1894.

Boore, D. M. (2003). Simulation of ground motion using the stochastic method, *Pure Appl. Geophys.* 160, 635-676.

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.

Boore, D. M., and W. B. Joyner (1984). A note on the use of random vibration theory to predict peak amplitudes of transient signals, *Bull. Seism. Soc. Am.* 74, 2035-2039.

Cartwright, D. E. and M. S. Longuet-Higgins (1956). The statistical distribution of the maxima of a random function, *Proc. R. Soc. London* 237, 212–232.

Der Kiureghian, A. (1979). On response of structures to stationary excitation, UCB/EERIC-79/32.

Der Kiureghian, A. (1983). Lecture notes for short course on structural reliability: theory and applications, Berkeley, March 23-25.

Renault, P. (2010). RVT reference test case verification summary, Technical Note QA-TN-1118, PEGASOS Refinement Project.

Saragoni, G. R. and Hart, G. C. (1974). Simulation of artificial earthquakes, *Earthq. Eng. Struct. Dyn.* 2, 249–267.

Vanmarcke, E. H. (1975). Distribution of first-passage time for normal stationary random processes, *Journal of Applied Mechanics-Transactions of the ASME* 42, 215-220.

TABLES

Table 1. NGA West-2 events included in the **M** 4.5-5.5 bin.

EQID	Earthquake Name	Year	Magnitude	Number of records
0163	Anza-02	2001	4.92	147
0161	Big Bear-02	2001	4.53	86
0170	Big Bear City	2003	4.92	73
0077	Coalinga-02	1983	5.09	41
0079	Coalinga-04	1983	5.18	22
0166	Gilroy	2002	4.90	72
0051	Imperial Valley-07	1979	5.01	33
0029	Lytle Creek	1970	5.33	27
0150	Northridge-05	1994	5.13	18
0151	Northridge-06	1994	5.28	126
0039	Oroville-03	1975	4.70	27
0114	Whittier Narrows-02	1987	5.27	142
0160	Yountville	2000	5.00	58

Table 2. NGA West-2 events included in the **M** 5.5-6.5 bin.

EQID	Earthquake Name	Year	Magnitude	Number of records
0126	Big Bear-01	1992	6.46	82
0171	Chi-Chi, Taiwan-02	1999	5.90	906
0174	Chi-Chi, Taiwan-05	1999	6.20	984
0076	Coalinga-01	1983	6.36	94
0274	L'Aquila, Italy	2009	6.30	138
0101	N. Palm Springs	1986	6.06	66
0179	Parkfield-02, CA	2004	6.00	180
0113	Whittier Narrows-01	1987	5.99	235

Table 3. NGA West-2 events included in the **M** 6.5-7.5 bin.

EQID	Earthquake Name	Year	Magnitude	Number of records
278	Chuetsu-oki	2007	6.80	1848
158	Hector Mine	1999	7.13	253
279	Iwate	2008	6.90	1101
127	Northridge-01	1994	6.69	305
280	Sierra El Mayor	2010	7.20	480

Figures

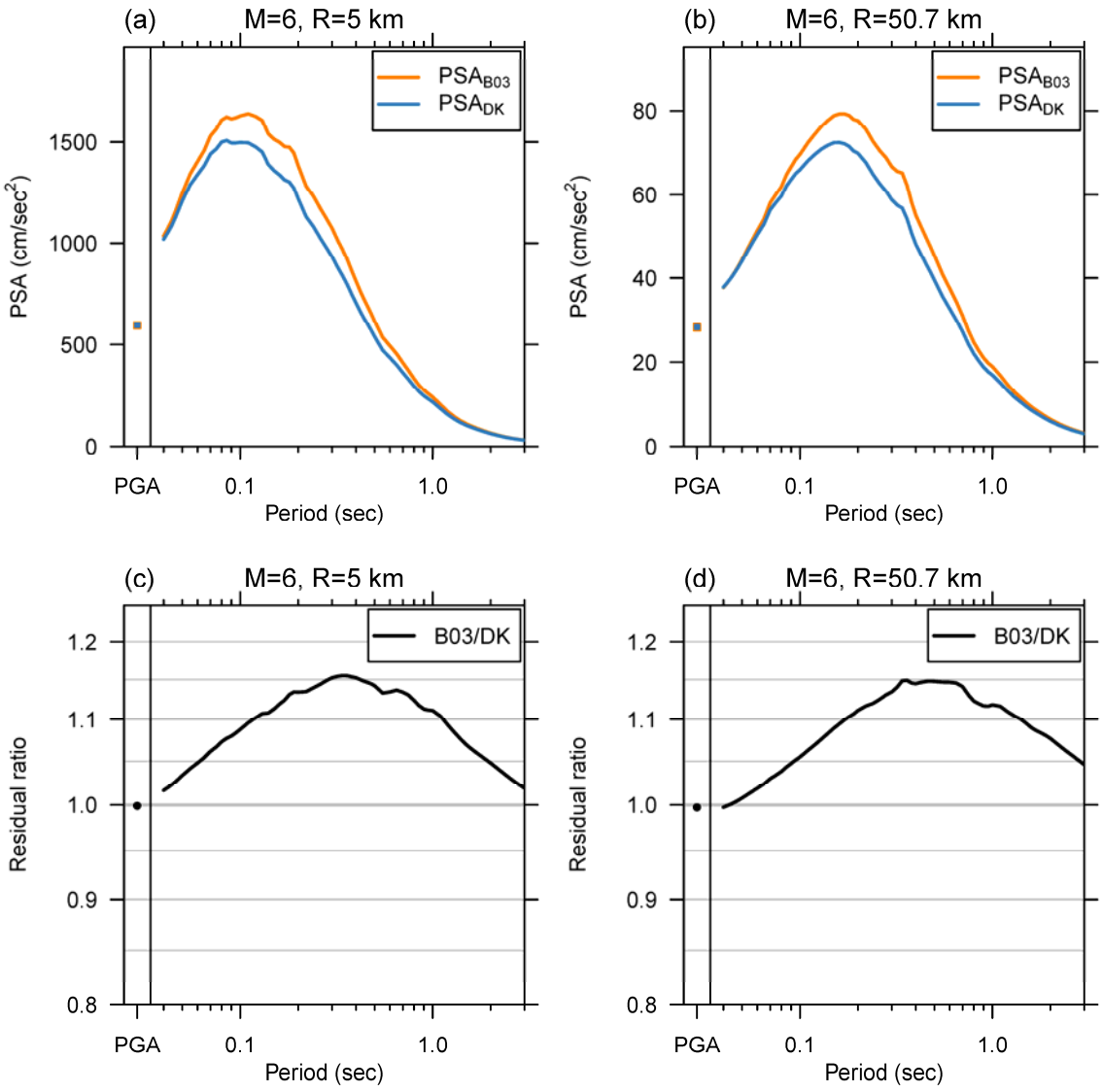

Figure 1. PSA from simulated ground motions from an **M** 6 earthquake at distances of (a) 5 km and (b) 50.7 km. The corresponding ratios for the two different RMS-to-peak factors are given in (c) and (d).

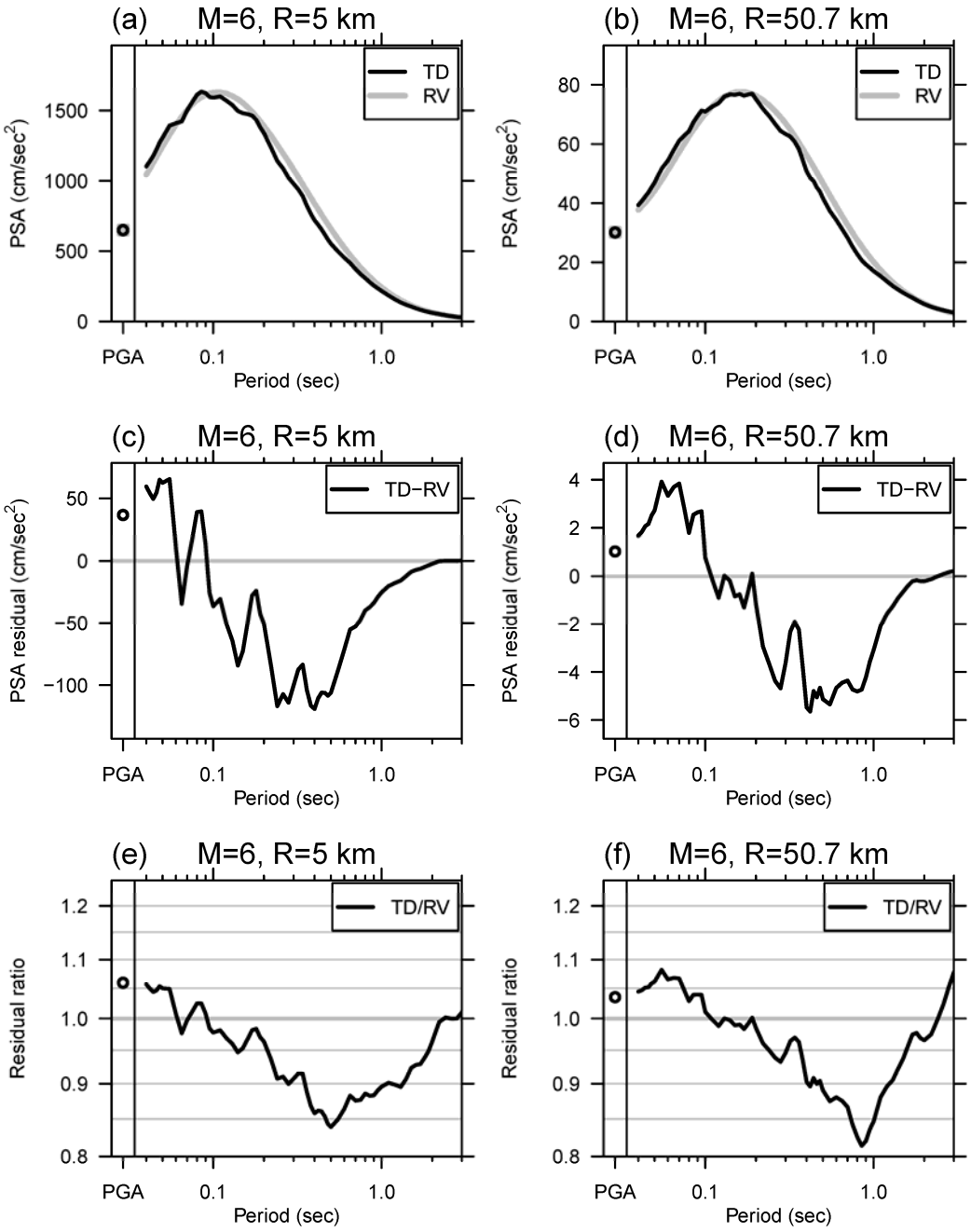

Figure 2. Comparison of the target RVT PSA (labeled RV) and the PSA_{B03}^{p} computed as the geometric mean of 100 time domain (TD) simulations for an **M** 6 earthquake: (a), (c) and (e) are the PSA, the arithmetic residual between the RV and TD curves, and ratio of the TD curve to the RV curve at a distance of 5 km, respectively; (b), (d), and (f) are analogous figures at a distance of 50.7 km.

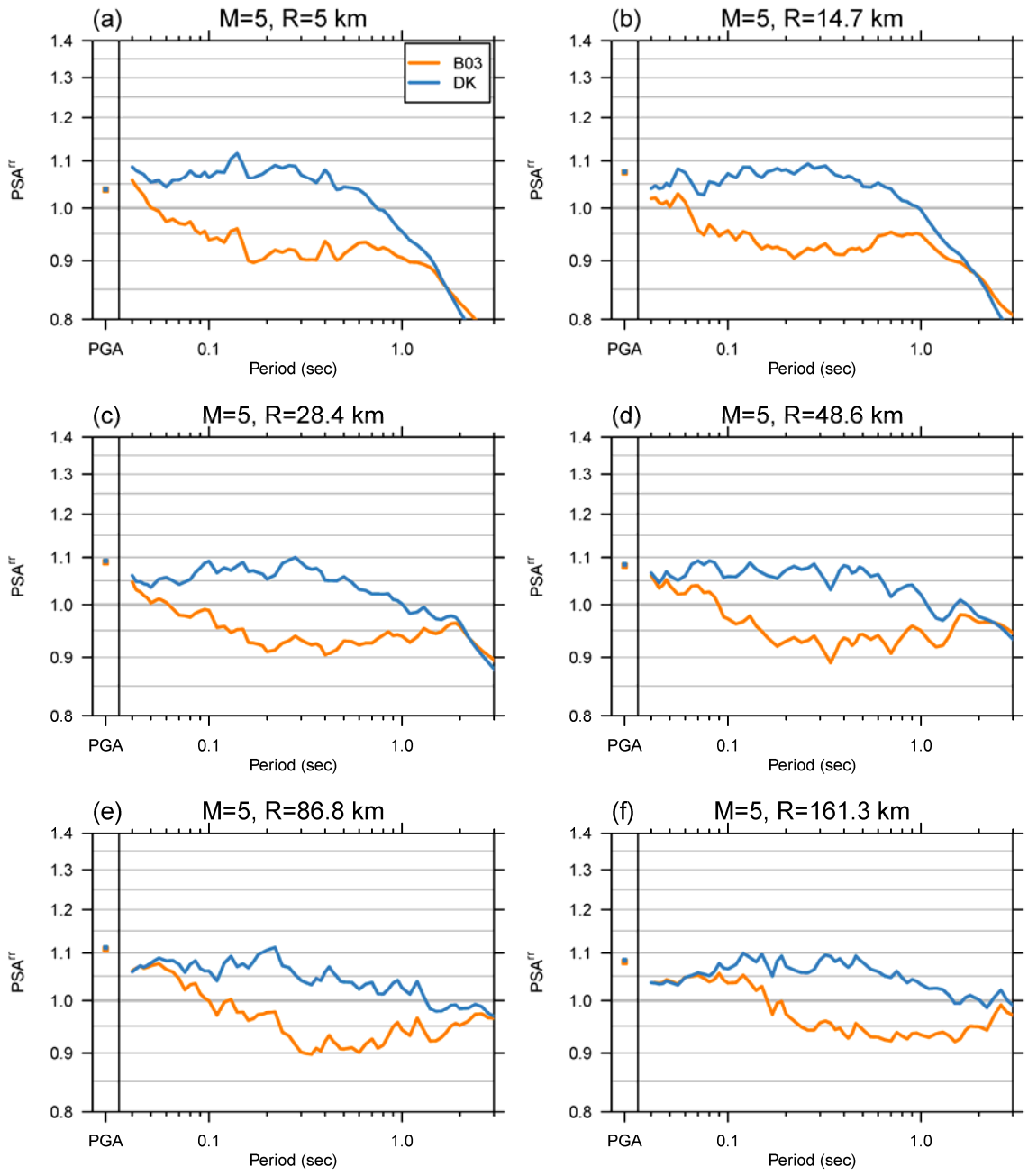

Figure 3. PSA^{rr}_{B03} and PSA^{rr}_{DK} for a simulated **M** 5 earthquake at six different distances: (a) 5 km, (b) 14.7 km, (c) 28.4 km, (d) 48.6 km, (e) 86.8 km, and (f) 161.3 km.

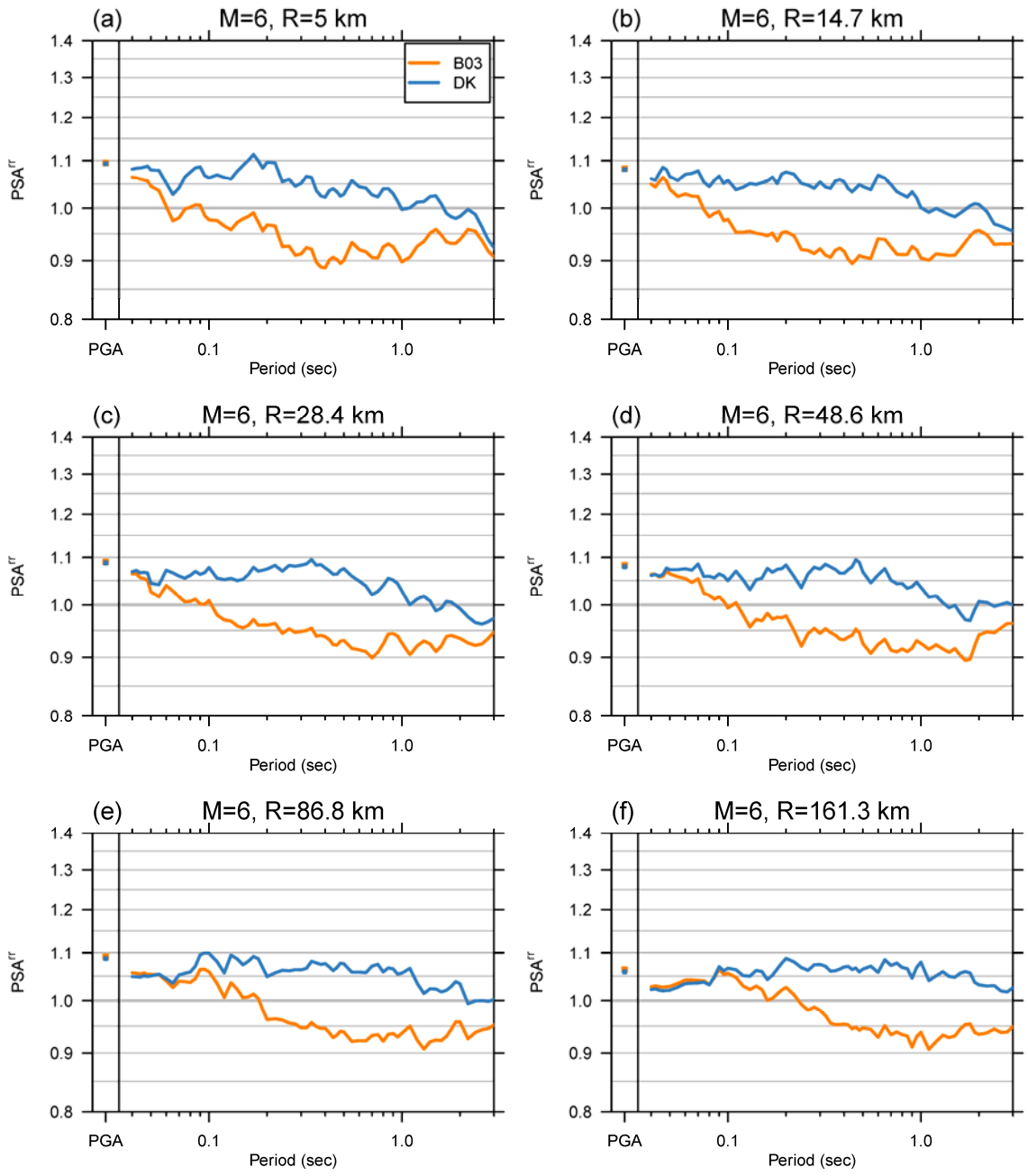

Figure 4. PSA^{rr}_{B03} and PSA^{rr}_{DK} for a simulated **M** 6 earthquake at six different distances: (a) 5 km, (b) 14.7 km, (c) 28.4 km, (d) 48.6 km, (e) 86.8 km, and (f) 161.3 km.

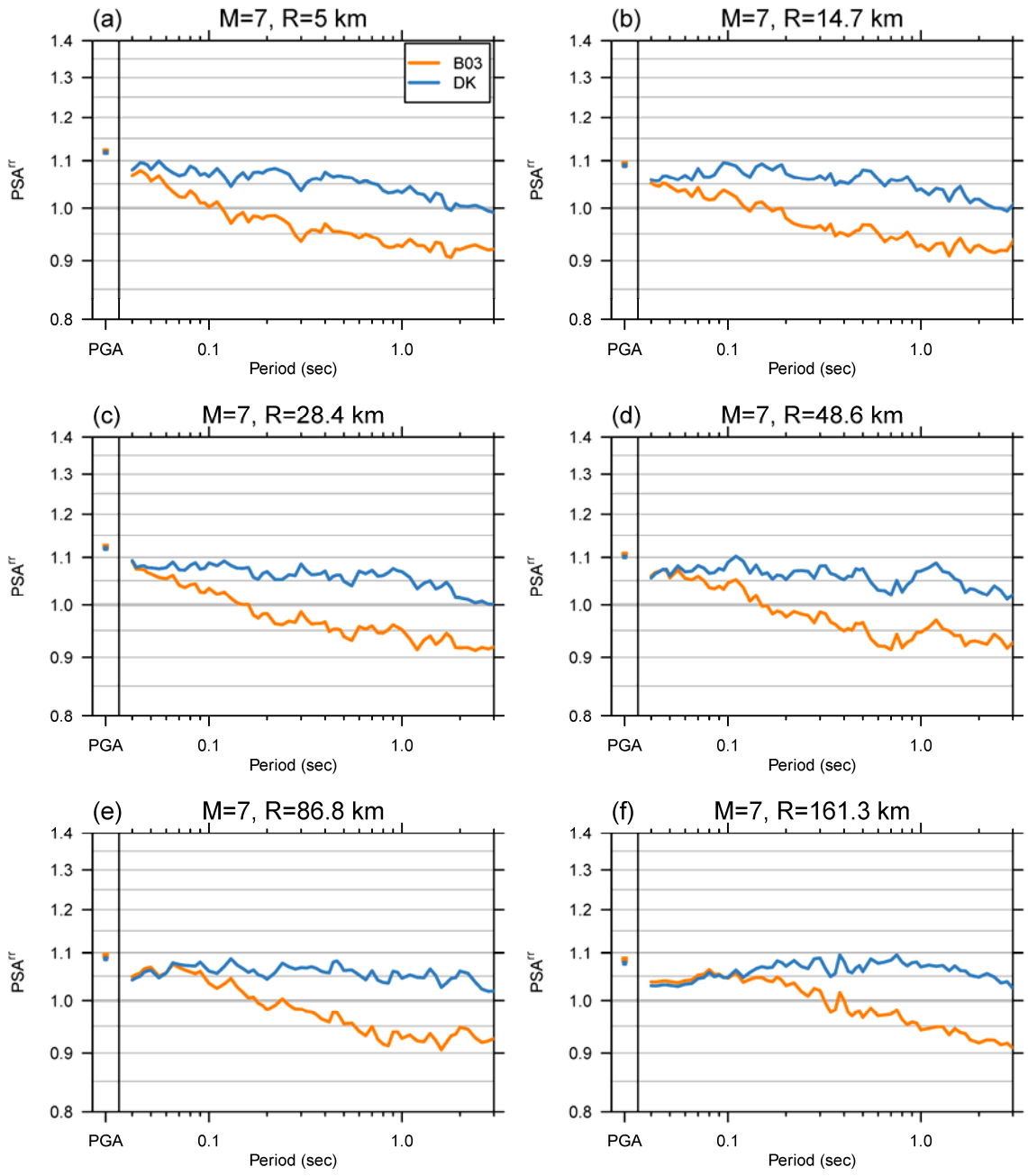

Figure 5. PSA^{rr}_{B03} and PSA^{rr}_{DK} for a simulated **M** 7 earthquake at six different distances: (a) 5 km, (b) 14.7 km, (c) 28.4 km, (d) 48.6 km, (e) 86.8 km, and (f) 161.3 km.

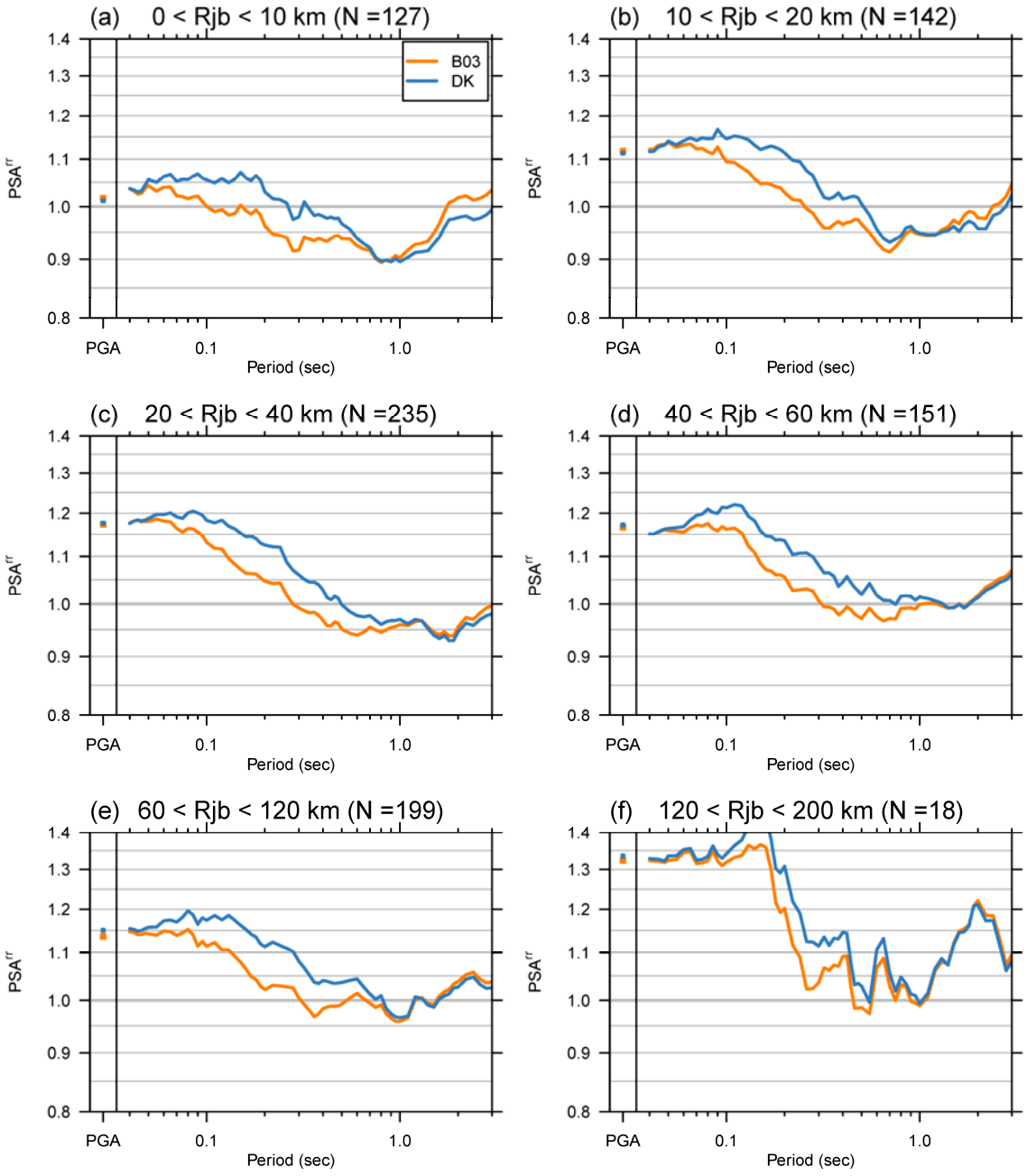

Figure 6. PSA^{rr}_{B03} and PSA^{rr}_{DK} for observed data from earthquakes from **M** 4.5 to 5.5 at six different distances: (a) $0 < R_{JB} < 10$ km, (b) $10 < R_{JB} < 20$ km, (c) $20 < R_{JB} < 40$ km, (d) $40 < R_{JB} < 60$ km, (e) $60 < R_{JB} < 120$ km, and (f) $120 < R_{JB} < 200$ km.

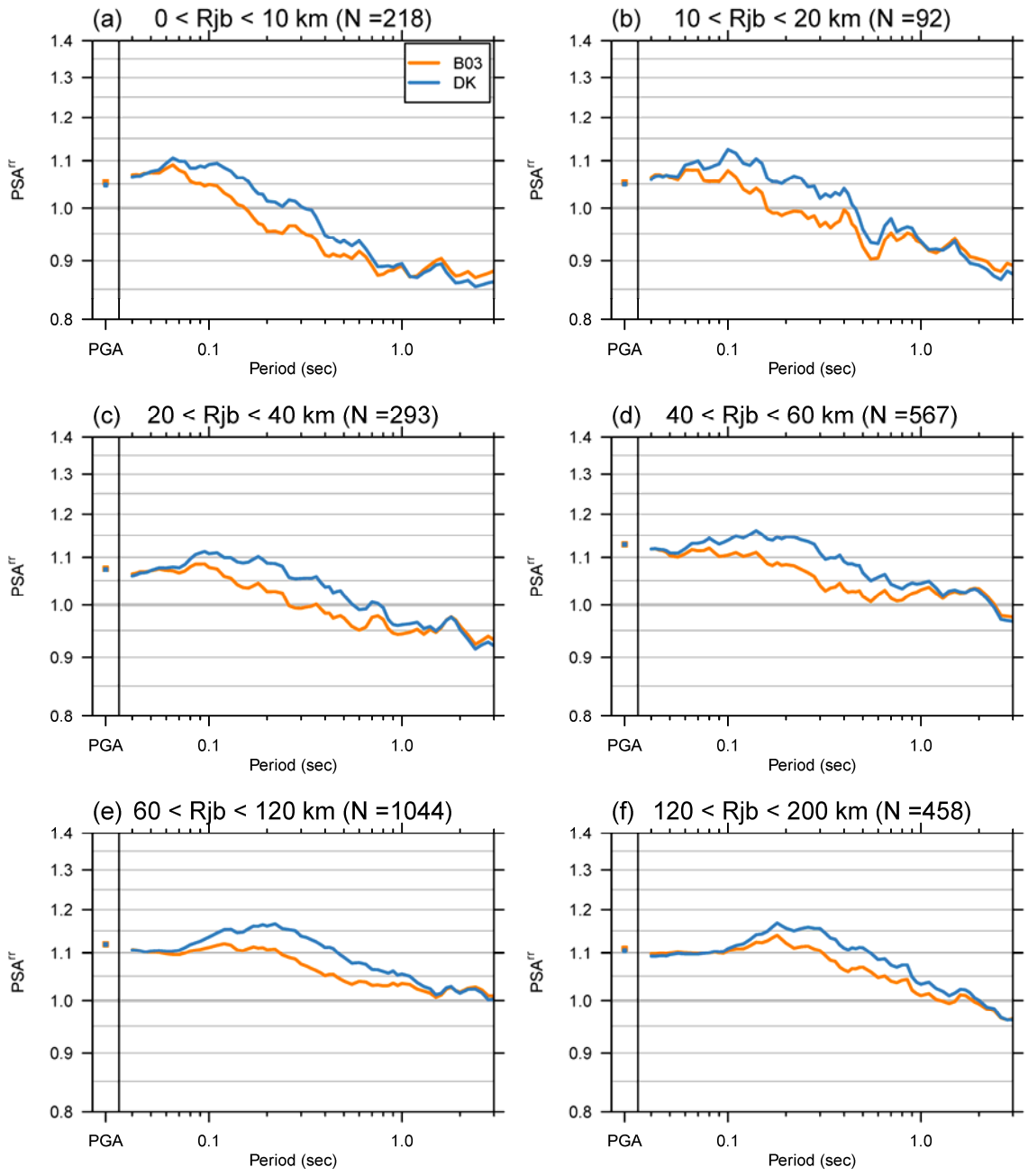

Figure 7. PSA^{rr}_{B03} and PSA^{rr}_{DK} for observed data from earthquakes from **M** 5.5 to 6.5 at six different distances: (a) $0 < R_{JB} < 10$ km, (b) $10 < R_{JB} < 20$ km, (c) $20 < R_{JB} < 40$ km, (d) $40 < R_{JB} < 60$ km, (e) $60 < R_{JB} < 120$ km, and (f) $120 < R_{JB} < 200$ km.

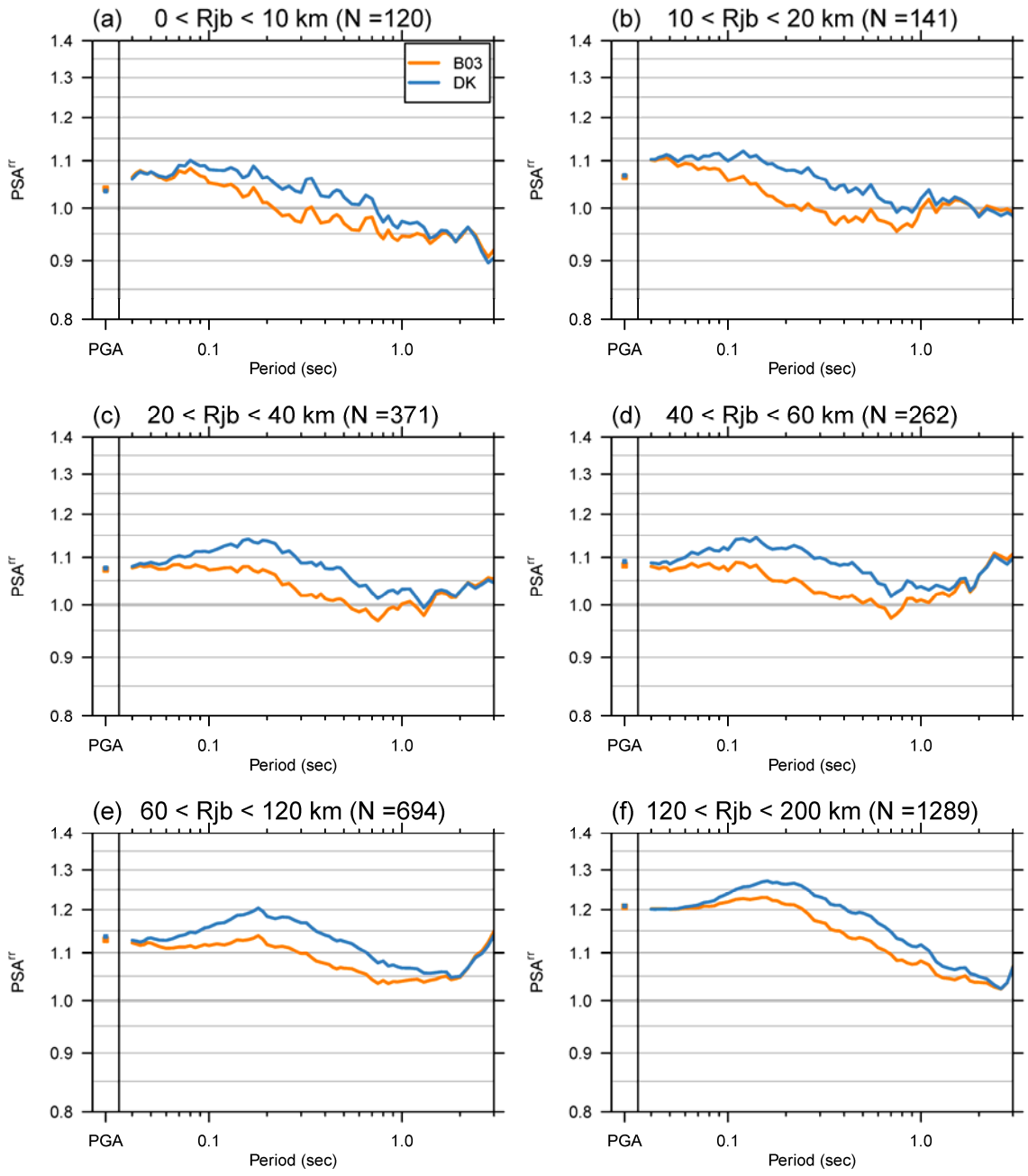

Figure 8. PSA^{rr}_{B03} and PSA^{rr}_{DK} for observed data from earthquakes from **M** 6.5 to 7.5 at six different distances: (a) $0 < R_{JB} < 10$ km, (b) $10 < R_{JB} < 20$ km, (c) $20 < R_{JB} < 40$ km, (d) $40 < R_{JB} < 60$ km, (e) $60 < R_{JB} < 120$ km, and (f) $120 < R_{JB} < 200$ km.

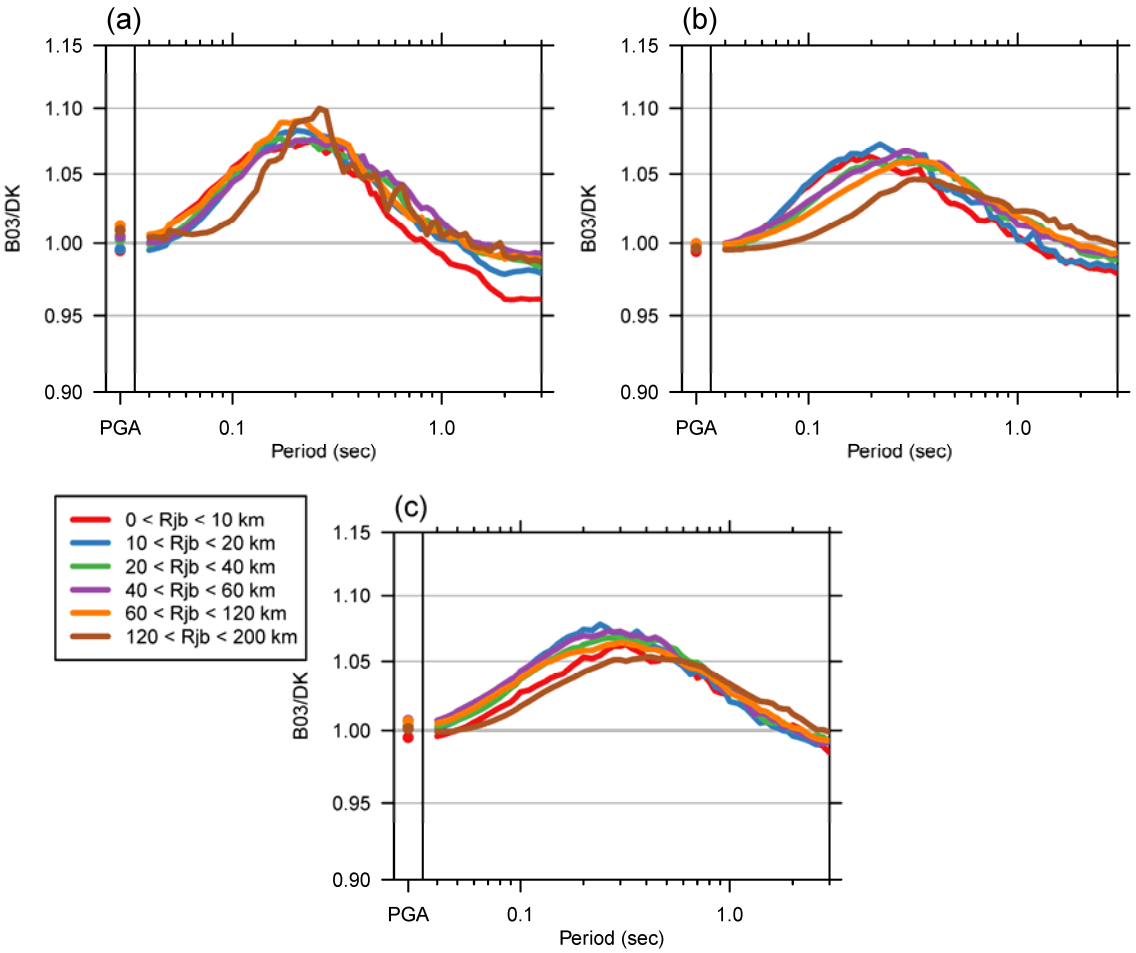

Figure 9. Ratio of PSA^{p}_{B03} to PSA^{p}_{DK} for the NGA data for (a) **M** 4.5-5.5, (b) **M** 5.5-6.5, and (c) **M** 6.5-7.5.

APPENDIX 1: Arithmetic residuals for **M** 5, **M** 6, and **M** 7 events.

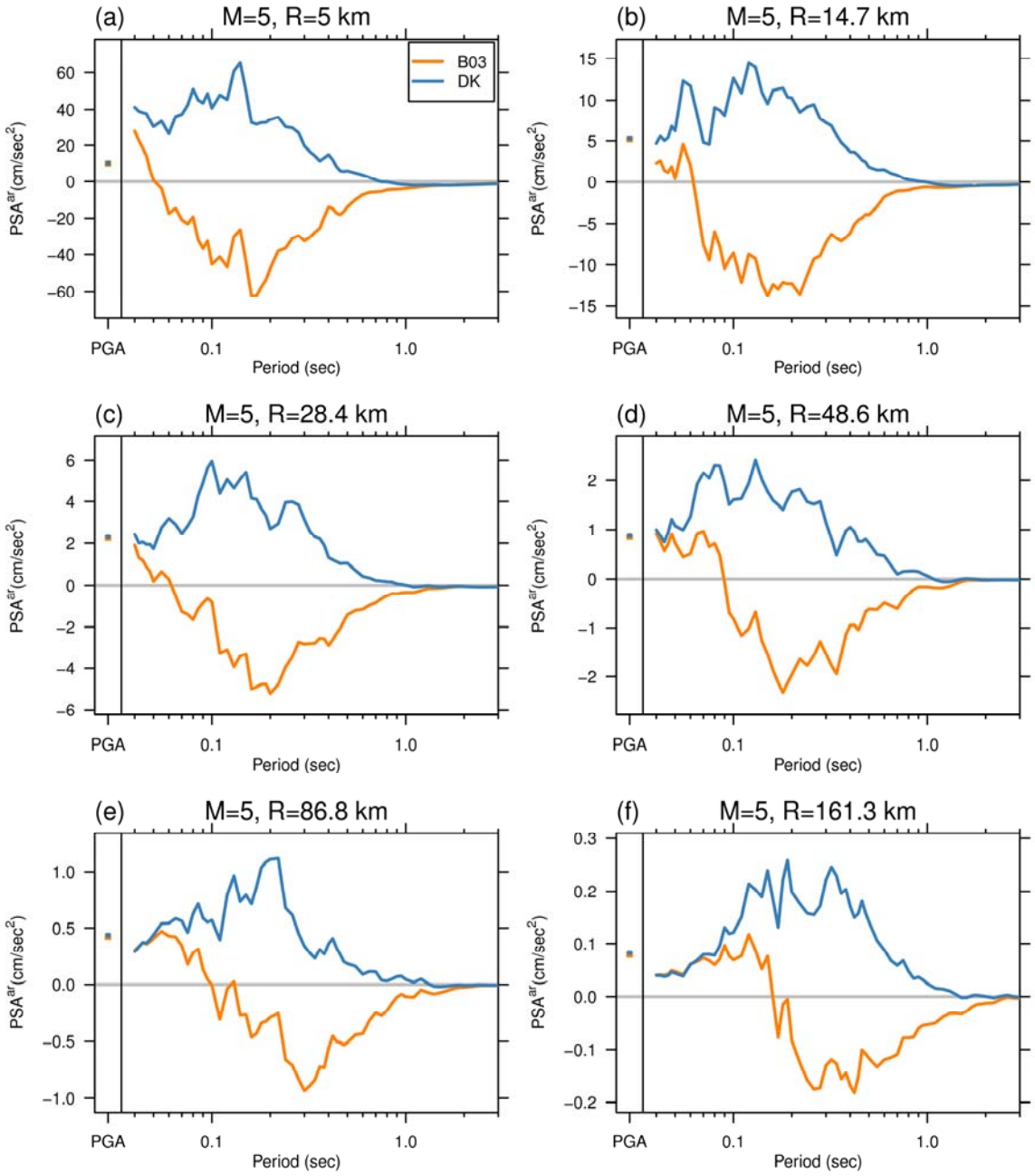

Figure A1. PSA^{ar} for 100 simulated time series for an **M** 5 earthquake at six different distances: (a) 5 km, (b) 14.7 km, (c) 28.4 km, (d) 48.6 km, (e) 86.8 km, and (f) 161.3 km.

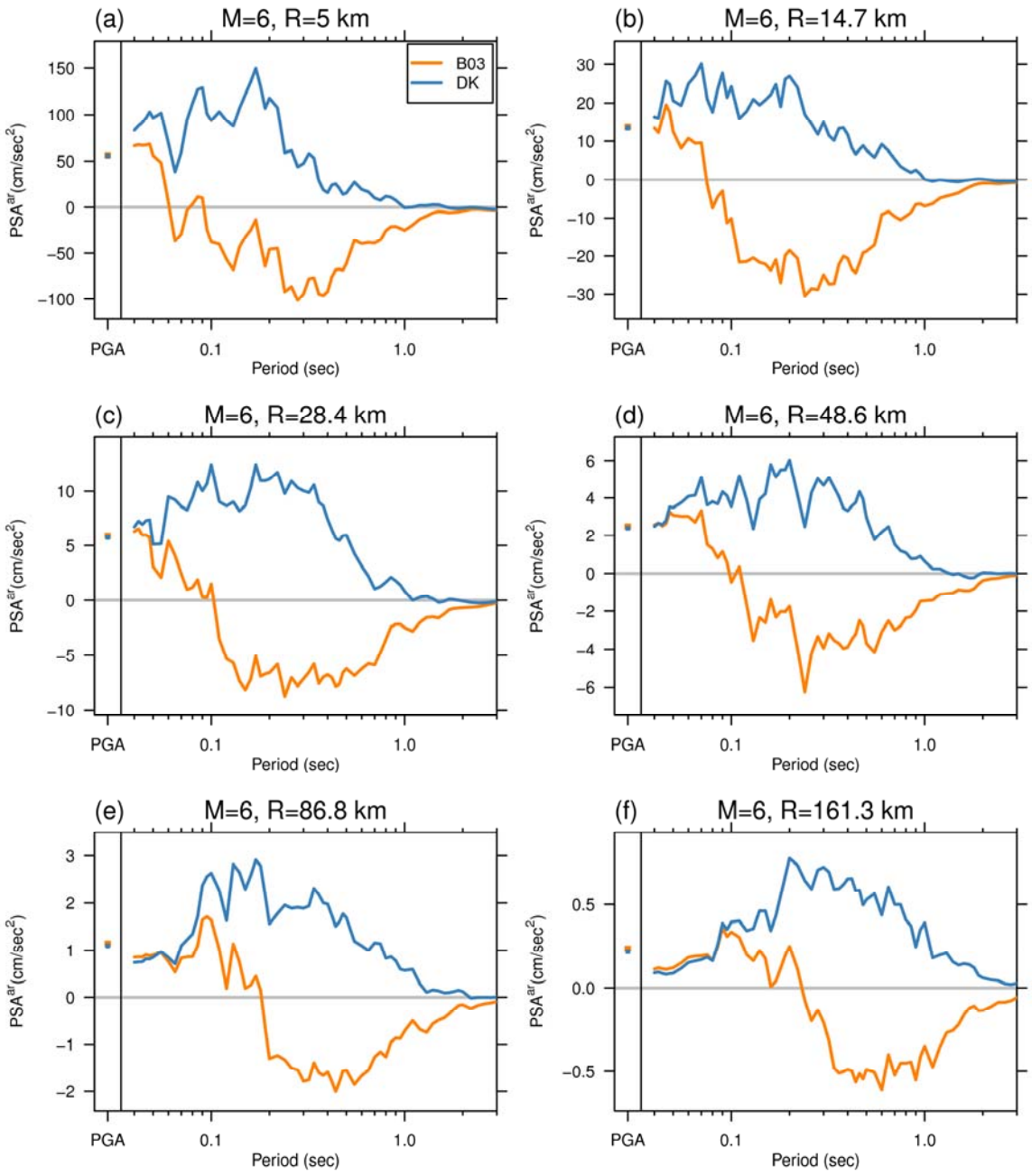

Figure A2. PSA^{ar} for 100 simulated time series for an **M** 6 earthquake at six different distances: (a) 5 km, (b) 14.7 km, (c) 28.4 km, (d) 48.6 km, (e) 86.8 km, and (f) 161.3 km.

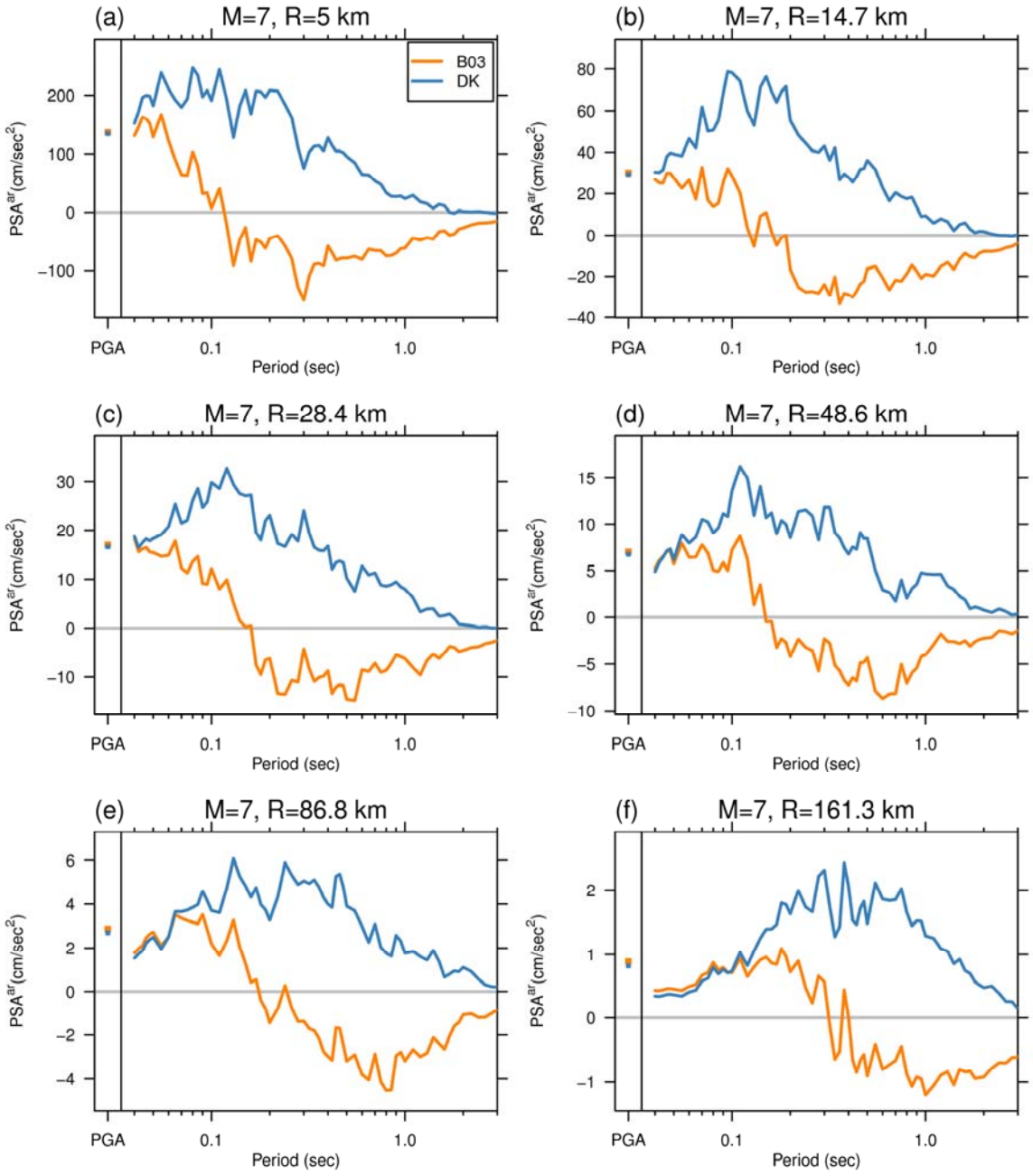

Figure A3. PSA^{ar} for 100 simulated time series for an **M** 7 earthquake at six different distances: (a) 5 km, (b) 14.7 km, (c) 28.4 km, (d) 48.6 km, (e) 86.8 km, and (f) 161.3 km.

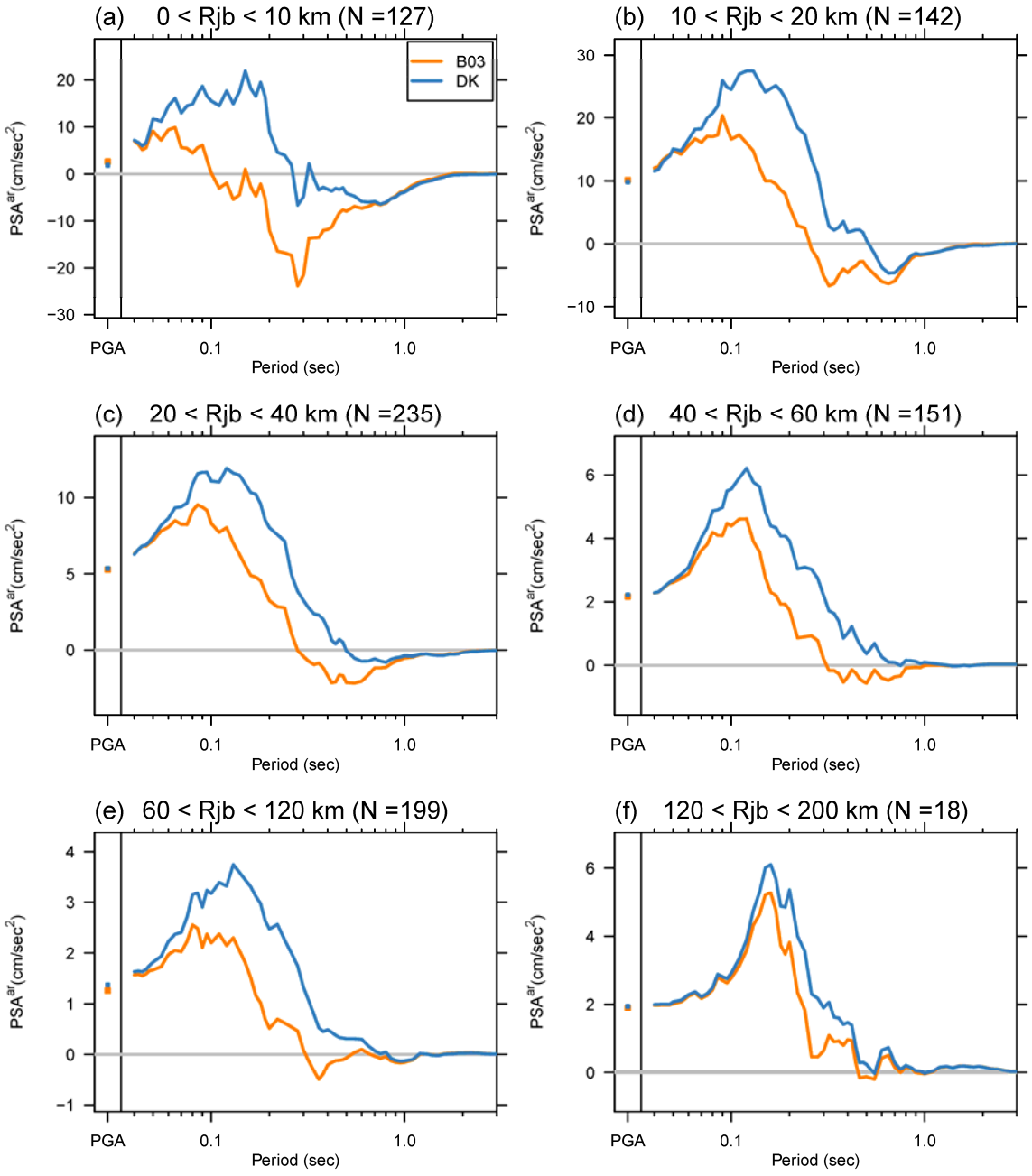

Figure A4. PSA^{ar} for NGA data for **M** 4.5-5.5 events at six different distances: (a) 0 < R_{JB} < 10 km, (b) 10 < R_{JB} < 20 km, (c) 20 < R_{JB} < 40 km, (d) 40 < R_{JB} < 60 km, (e) 60 < R_{JB} < 120 km, and (f) 120 < R_{JB} < 200 km.

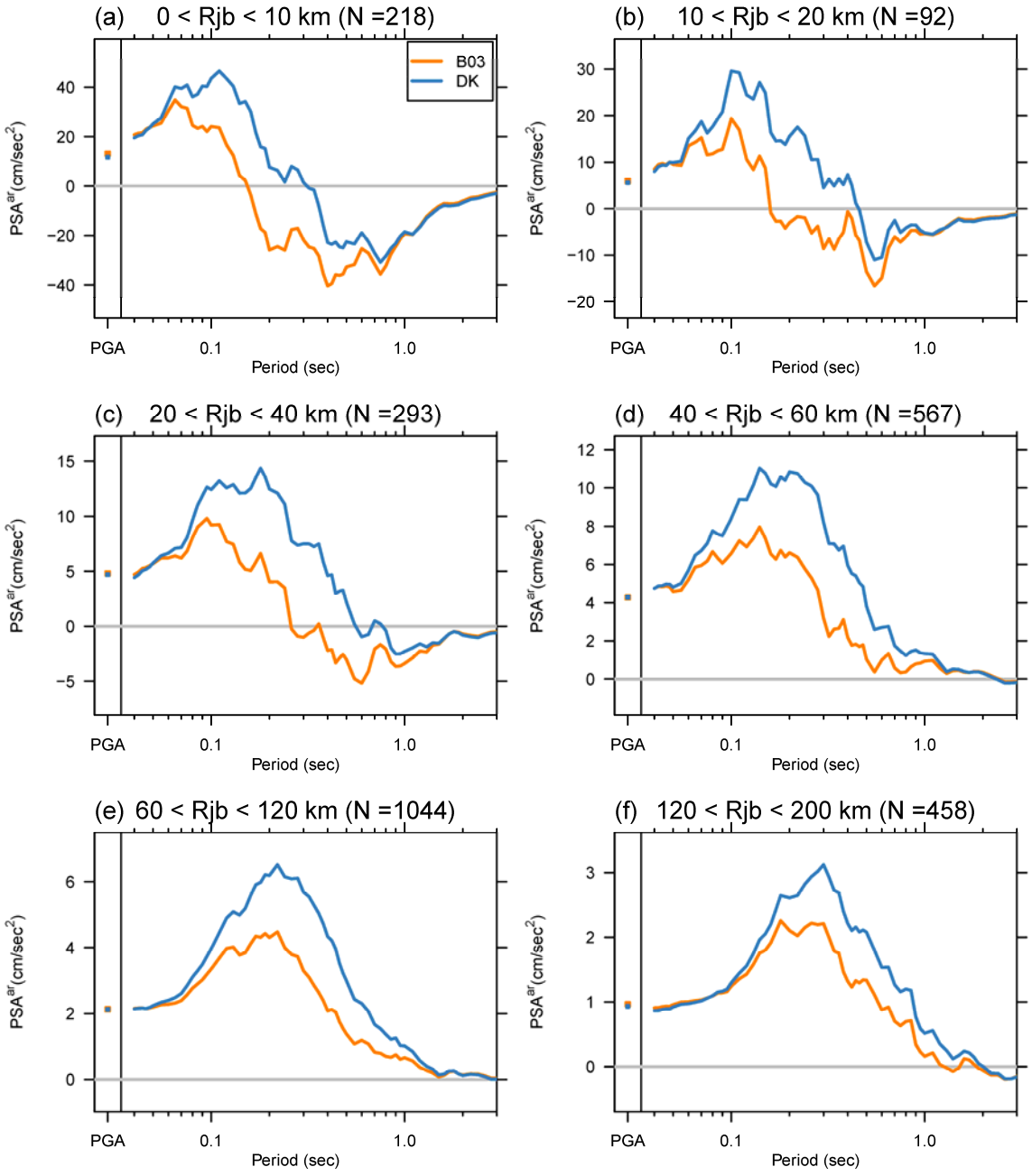

Figure A5. PSA^{ar} for NGA data for **M** 5.5-6.5 events at six different distances: (a) 0 < R_{JB} < 10 km, (b) 10 < R_{JB} < 20 km, (c) 20 < R_{JB} < 40 km, (d) 40 < R_{JB} < 60 km, (e) 60 < R_{JB} < 120 km, and (f) 120 < R_{JB} < 200 km.

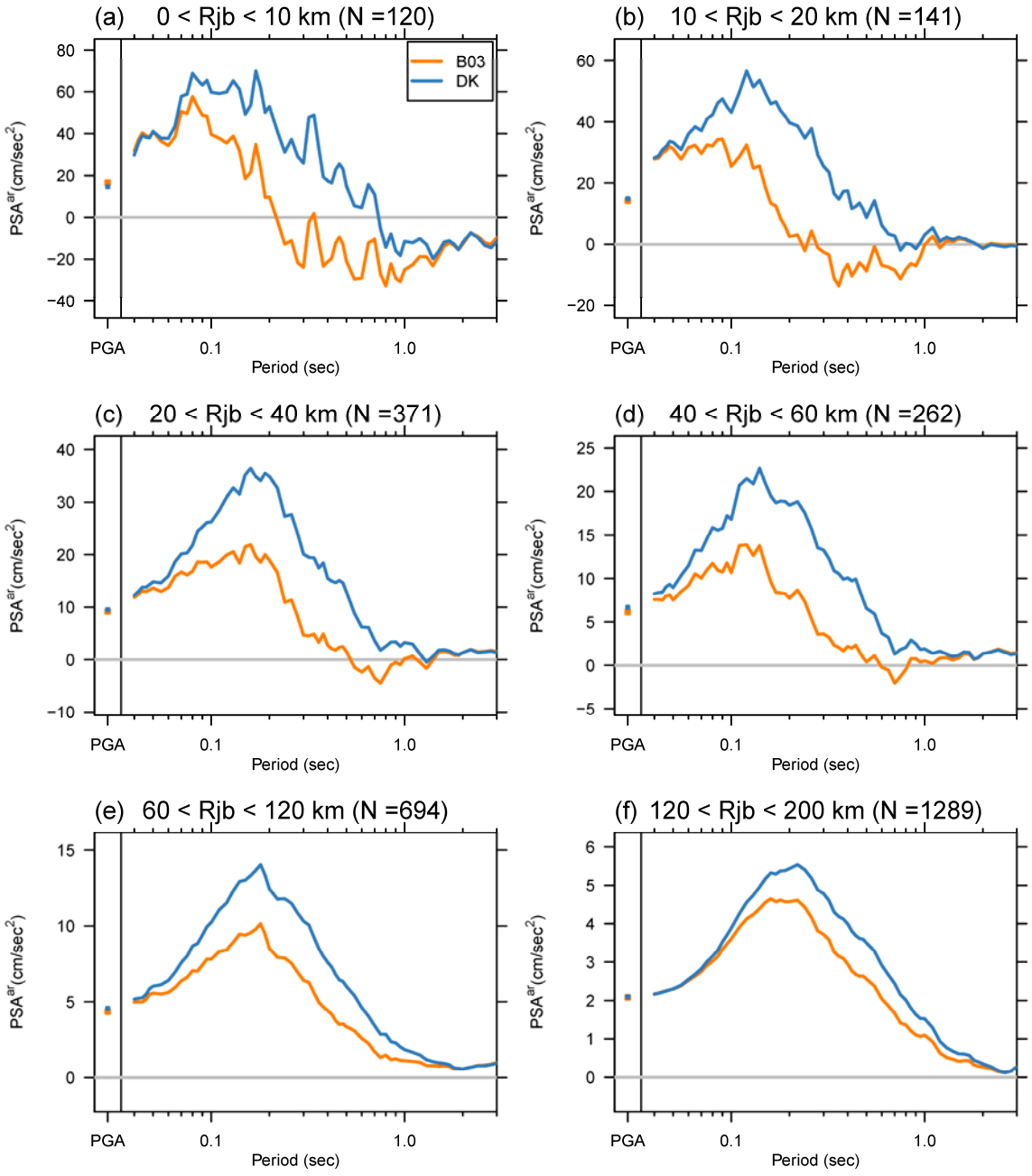

Figure A1.6. PSA^{ar} for NGA data for **M** 6.5-7.5 events at six different distances: (a) 0 < R_{JB} < 10 km, (b) 10 < R_{JB} < 20 km, (c) 20 < R_{JB} < 40 km, (d) 40 < R_{JB} < 60 km, (e) 60 < R_{JB} < 120 km, and (f) 120 < R_{JB} < 200 km.

APPENDIX 2: Sample SMSIM input parameter file.

```
!Revision of program involving a change in the parameter file on this date:
   12/16/09
!Title
Params for AS2000 (Atkinson and Silva, BSSA 90, 255--274) source, applied to WNA
!rho, beta, prtitn, radpat, fs:
    2.8 3.5 0.707 0.55 2.0
!spectral shape: source number, pf_a, pd_a, pf_b, pd_b
! where source number means:
!  1 = 1-corner (S = 1/(1+(f/fc)**pf_a)**pd_a)
!  2 = Joyner (BSSA 74, 1167--1188)
!  3 = Atkinson (BSSA 83, 1778--1798; see also Atkinson & Boore, BSSA 85,
!      17--30)
!  4 = Atkinson & Silva (BSSA 87, 97--113)
!  5 = Haddon 1996 (approximate spectra in Fig. 10 of
!          Haddon's paper in BSSA 86, 1300--1313;
!          see also Atkinson & Boore, BSSA 88, 917--934)
!  6 = AB98-California (Atkinson & Boore BSSA 88, 917--934)
!  7 = Boatwright & Choy (this is the functional form used by
!                         Boore & Atkinson, BSSA 79, 1736--1761, p. 1761)
!  8 = Joyner (his ENA two-corner model, done for the SSHAC elicitation
!      workshop)
!  9 = Atkinson & Silva (BSSA 90, 255--274)
! 10 = Atkinson (2005 model),
! 11 = Generalized two corner model
!      (S = [1/(1+(f/fa)**pf_a)**pd_a]*[1/(1+(f/fb)**pf_b)**pd_b])
! pf_a, pd_a, pf_b, pd_a are used for source numbers 1 and 11, usually
! subject to the constraint pf_a*pd_a + pf_b*pd_b = 2 for an omega-squared
! spectrum. The usual single-corner frequency model uses
! pf_a=2.0,pd_a=1.0; the Butterworth filter shape is given by
! pf_a=4.0,pd_a=0.5.  pf_b and pd_b are only used by source 11, but dummy
! values must be included for all sources.
     9 2.0 1.0 0.0 0.0
!spectral scaling: stressc, dlsdm, fbdfa, amagc, c1_fa, c2_fa, amagc4fa
! (stress=stressc*10.0**(dlsdm*(amag-amagc))
! (fbdfa, amagc for Joyner model, usually 4.0, 7.0)
! c1_fa, c2_fa are the coefficients relating log fa to M in
! source 11, as given by the equation log fa = c1_fa + c2_fa*(M-amagc4fa).
! fb for source 11 is given such that the high-frequency spectral level
! equals that for a single corner frequency model with a stress parameter
! given by stress=stressc*10.0**(dlsdm*(amag-amagc).
! See Tables 2 and 3 in Boore (2003) for various source descriptions
! (Note: the parameters in the line below are not used for most of the
! sources, for which the spectrum is determined by fixed relations between
! corner frequency and seismic moment, but placeholders are still needed)
    100.0 0.0 4.0 7.0  0.0 0.0 0.0
!iflag_h_eff, c1_log10_h_eff, c2_log10_h_eff
!  If iflag = 1, compute an effective depth as
!  h_eff = 10.0**(c1_log10_h_eff + c2_log10_h_eff*amag) and modify the closest
!  distance by this depth: rmod = sqrt(r^2+h_eff^2)); use rmod in
!  the calculations
!  1 -0.05 0.15 ! Atkinson and Silva (2000) values
   0  0.0 0.0
!gsprd: r_ref, nsegs, (rlow(i), a_s, b_s, m_s(i))  (Usually set r_ref = 1.0 km)
    1.0
    2
      1.0 -1.0 0.0 6.5
     40.0 -0.5 0.0 6.5
!q: fr1, Qr1, s1, ft1, ft2, fr2, qr2, s2, c_q
    1.0 180 0.45 1.0 1.0 1.0 180 0.45 3.5
!source duration: weights of 1/fa, 1/fb
    0.5 0.0
!path duration: nknots, (rdur(i), dur(i), slope of last segment
    1
      0.0 0.0
     0.05
```

```
!crustal amplification, from the source to the site (note that this can include
! local site amplification): namps, (famp(i), amp(i))
    11
       0.01                1.00
       0.09                1.10
       0.16                1.18
       0.51                1.42
       0.84                1.58
       1.25                1.74
       2.26                2.06
       3.17                2.25
       6.05                2.58
      16.6                 3.13
      61.2                 4.00
!site diminution parameters: fmax, kappa, dkappadmag, amagkref
! (NOTE: fmax=0.0 or kappa=0.0 => fmax or kappa are not used.  I included this
!  to prevent the inadvertent use of both fmax and kappa to control the diminution
!  of high-frequency motion (it would be very unusual to use both parameters
!  together.  Also note that if do not want to use kappa, dkappadmag must also
! be set to 0.0).
    0.0 0.03 0.0 0.0
!low-cut filter parameters: fcut, nslope (=4, 8, 12, etc)
    0.04 4
!rv params: zup, eps_int (int acc), amp_cutoff (for fup), osc_crrctn(1=b&j;2=l&p)
    10.0 0.00001  0.001 1
!window params: idxwnd(0=box,1=exp), tapr(<1), eps_w, eta_w, f_tb2te, f_te_xtnd
    1 0.05 0.2 0.05 2.0 2.0
!timing stuff: dur_fctr, dt, tshift, seed, nsims, iran_type (0=normal;1=uniform)
    1.3  0.005 20.0 123.0 100 0
```