# Appendix A: MATLAB and HAZ45 Model Implementations

The repository <https://github.com/jay14bay/USGS_G22AP00199> contains the following MATLAB files:

* Bea24\_Example.m: a script which calculates the directivity effect (median and standard deviation) for an example scenario and produces a map, making use of the functions below.
* GC2.m: Calculates the GC2 coordinates. This function is a conversion to MATLAB of Brian Chiou's R functions (pers. comm.)
* Bea24.m: a function which implements the Bea24 directivity model.

The repository <https://github.com/jay14bay/USGS_G22AP00199> contains the following files for implementation of the model into HAZ45.2:

* Directivity\_bea24.f: a fortran subroutine which implements the directivity model.
* Modified versions of the following HAZ45.2 programs and subroutines: Directivity.f, Haz\_main2.f, gc2.f, cldist.f, and declare1.h
* ChangeLog-7Feb2024.txt: a list of the required changes from HAZ45.2 required to implement the model.

# Appendix B: Directivity Centering

This appendix provides detailed descriptions of the two directivity centering concepts: neutrality and centering, followed by an evaluation of the neutrality of the NGA-W2 database.

**Concept 1: Neutrality of the Directivity Condition**

The ‘neutrality’ refers to the bias of the average directivity condition contained within the empirical data as compared to the average directivity condition for uniformly distributed sites and hypocenters. Any models developed from the empirical data will reflect this bias. Therefore, this concept can apply to a GM database or to an empirical GMM derived from such a database. If the bias is zero, the directivity condition is neutral or ‘centered’ for future earthquakes. The neutrality issue applies to all GMMs, whether that GMM explicitly accounts for rupture directivity effects or not.

The neutrality of the directivity condition is a sampling bias issue. For a given earthquake, the bias will depend on the number of and locations of recording stations, including their distribution of source-site azimuths, and on the finite-fault properties of the rupture, including the rupture geometry and earthquake hypocenter location. For a GM database (consisting of recordings from a large set of earthquakes and recording stations), the neutrality condition depends on the collection of source-site recording pairs and can vary with distance. At a given distance, a database can be biased towards a forward directivity condition (characterized by larger than average long-period ground-motion amplitude with shorter durations), a backwards directivity condition (characterized by smaller than average long-period ground-motion amplitude with longer durations), or a neutral condition.

As an example, consider a hypothetical ground-motion database consisting of many well-recorded earthquakes, one of which is the vertical strike-slip earthquake shown in Figure B1a. In Earthquake #1, all five of the recording stations are at locations where forward directivity effects are expected to some degree, and therefore larger than average long-period ground-motion amplitudes are expected for this distance, corresponding to positive total residuals. This is due to their source-site azimuths and distances, and to the hypocenter location, which results in rupture propagation to the north towards the sites. As a result, the ground-motion recordings from Earthquake #1 will have a forward directivity bias. The **M**6.19 1984 Morgan Hill earthquake is an example of a similar situation from the NGA-W2 database (Ancheta et al., 2014; Figure B1b). The near-fault sampling of data in NGA-W1 (Chiou et al., 2008) was biased towards sites with forward directivity (Donahue et al., 2019; D19 hereafter).

|  |  |
| --- | --- |
| A graph of a number of miles  Description automatically generated with medium confidence | A graph of a number of points  Description automatically generated with medium confidence |

**Figure B1.** (a) Hypothetical Earthquake #1, with fault trace (black), hypocenter (red star), and recording stations (blue triangles) reflecting a directivity condition that is biased toward forward directivity. (b) A map of the 1984 **M**6.19 Morgan Hill earthquake rupture trace (black) and recording stations within 50 km with usable period up to 5 sec (blue triangles).

A hypothetical database made up of only situations like Earthquake #1 would have a strong forward directivity bias, and a GMM derived from this database would have a forward directivity bias in the median. In a database of real earthquakes, the distribution of source-site azimuths will approach a uniform distribution as the sample of earthquakes and recording stations increases; however, the distribution of source-site azimuths may be skewed at a given distance, especially short distances and large magnitudes where recorded databases are still relatively sparse. Additionally, the sampling of the locations of hypocenters in the recorded database with respect to the recording stations may not reflect a directivity neutral condition.

Because of these factors, the neutrality of directivity condition needs to be checked for recorded databases and empirical GMMs. If the bias is zero at a given distance range, the directivity condition is neutral for that GMM. If the bias is non-zero at a given distance range, the directivity condition in the GMM is also biased. This needs to be checked before a directivity adjustment model can be derived from the same database (and subsequently applied to the GMM), because a directivity model based upon a biased directivity condition will also be biased.

Similar concepts apply to the neutrality condition of the aleatory variability. The range of modeled directivity effects from the distribution of stations in the data set is quantified by their standard deviation of the directivity effects. This standard deviation can be compared with the standard deviation of the directivity effect from uniformly distributed stations with the full range of azimuths to determine if the aleatory variability of the directivity condition is neutral. This method for checking the neutrality of the directivity condition is applied to the NGA-W2 database below.

**Concept 2: Centering of a Directivity Model**

The second concept related to the term ‘centering’ applies to a DM which is designed to adjust the median and standard deviation of a GMM without directivity. A DM is centered if it does not change the average magnitude and distance scaling of the GMM.

A racetrack is a set of sites uniformly distributed over all azimuths around the finite-fault rupture and with equal source-site distance. For a given earthquake scenario and source-site distance, the mean of the directivity parameter on a racetrack is zero if the DM is centered. When the centered DM is applied to the GMM, the average distance scaling, which does not traditionally vary with azimuth, is unchanged. Similarly, the overall magnitude scaling is unchanged because the mean directivity adjustment over all distances and azimuths for the scenario is also zero.

In other words, a centered DM predicts azimuthally varying median adjustments, but the mean of these is zero over all azimuths at a given distance. This is necessary because the magnitude and distance scaling of the GMM are centered with respect to the recorded data used to develop the GMM. Applying a non-centered DM implies that the magnitude and/or distance scaling of the GMM were biased.

The Chiou and Youngs (2014; CY14) GMM is the only NGA-W2 model to explicitly incorporate directivity effects. CY14 uses the centered directivity predictor . The approach taken by CY14 is to center on its mean, defined as the mean value over a suite of sites located at the same distance to an earthquake over all azimuths (sites on a racetrack). Because mean is specific to an earthquake, CY14 fit a parametric approximation of mean to each NGA-W2 earthquake. The approach of centering the directivity predictor on its mean is also taken by Rowshandel (2013; Chapter 3 of Spudich et al. 2013) and Rowshandel (2018). Centering of the model described in this article is discussed further below.

**Relationship between GMM Neutrality and DM Centering**

For any GMM without directivity, the average directivity effect in the observed dataset is implicitly included in the median. The GMM neutrality addresses the reference directivity condition corresponding to that median. The CY14 and Bea20 DMs were developed primarily from NGA-W2 ground-motion data, either as part of the GMM regression (CY14) or from GMM residuals (Bea20). Therefore, these DMs also implicitly include the average directivity effect in the dataset, to the extent that the models are based on the NGA-W2 data.

However, the centering term of a DM is not necessarily the same as the GMM neutrality term (the average directivity bias). Due to data limitations, Bea20 grouped recordings from all distances (with source-to-site distances less than about 80km) rather than developing a centered model at discrete rupture distances. Bea20 was based on the relationship between the residuals and directivity parameters; these parameters include the source-site azimuth and the rupture propagation distance between the hypocenter and the site, but not the rupture distance. This resulted in a model which was centered with respect to the data used to create it, but, in forward application is not centered for any given distance when considering the full range of source-site azimuths.

To address this shortcoming, the Bea24 directivity predictor is centered by removing the centering term at a given rupture distance and for a given scenario. The centering term is the mean of the un-centered directivity predictor over a suite of uniformly distributed sites located at the same distance to an earthquake over all azimuths (sites on a racetrack). The centering of Bea24 is described in *Rupture Directivity Adjustment Model*, and in Supplemental Appendix C.

**Neutrality of the NGA-W2 GMMs**

D19 examined the distribution of directivity parameter DPP (Chiou and Youngs, 2014) for NGA-W2 events with **M**>6.5 and sites with Rrup < 40 km. For events with at least 10 recordings per earthquake, D19 binned the recordings by distance and found mean DPP values ranging from 0.002 to 0.13. D19 found the mean of all DPP values in the NGA-W2 database, regardless of distance, is 0.011. D19 concluded that, on average, the NGA-W2 data reflect a directivity-neutral condition using DPP.

This section evaluates the neutrality of the mean and standard deviation of the directivity condition of the NGA-W2 GMMs in more detail than D19. In this evaluation, the centered Bea24 and CY14 rupture directivity models are used. The Bea24 model predicts rupture directivity adjustment in natural log units of amplification. At a given rupture distance, the mean of over all source-to-site azimuths is zero because the Bea24 model is centered. This evaluation uses recordings from three NGA-W2 GMMs: Abrahamson et al. (2014), Boore et al. (2014), and Campbell and Bozorgnia (2014). These three GMMs were selected because these GMMs were used to develop the directivity model described in Section 4. Chiou and Youngs (2014) was not included in the model development because it contains directivity adjustments.

The CY14 directivity adjustment (natural log units of amplification) is and is based on the centered directivity predictor . For the analysis here, Jennifer Donahue (personal communication) provided a spreadsheet of values calculated by Brian Chiou for each recording in the NGA-W2 database. The values were converted to using Equation 7 of CY14, which applies a constant, magnitude taper, distance taper, and period dependence to .

**Evaluation Method**

We select the set of earthquakes listed in Table B-1 to evaluate the neutrality of NGA-W2. The earthquakes need to be large enough in magnitude to have a finite-fault model for calculating the CY14 and Bea24 directivity adjustments. Because directivity models are less established for reverse and normal style of faulting earthquakes, the focus here is on strike-slip and select oblique style of faulting earthquakes. The procedure used to evaluate the neutrality of the directivity condition is:

1. For each earthquake listed in Table B-1, identify NGA-W2 recordings with Rrup < 100 km and which were used by the following GMMs: ASK14, BSSA14, and CB14. At rupture distances larger than 100 km, both the Bea24 and CY14 directivity models have no effect. Calculate and for these earthquakes and recording locations. Note that the and are unrelated to GMM residuals. These depend only on the DM and the parameters describing the earthquake sources (rupture geometry, magnitude, hypocenter location) and stations (rupture distance and location of the station with respect to the finite-fault rupture).
2. For each earthquake listed in Table B-1, calculate on a densely sampled grid of station locations (0.5 km grid spacing). Figure B-1 shows a map of at T=3 sec for the Landers earthquake at the dense grid of stations (contours; step 2) and at the recording stations (circles; step 1).
3. For each earthquake listed in Table B-1, group the NGA-W2 recording stations into rupture distance bins and calculate the sample mean of and ( and ) and sample standard deviations ( and ) within each bin. Table B-1 lists and for each earthquake and for two distance bins: Rrup less than 20 km and Rrup less than 40 km. Evaluate the neutrality of the directivity condition for individual earthquakes.
4. For the complete set of earthquakes listed in Table B-1, group the gridded stations into rupture distance bins and calculate the population standard deviations () within each bin. The population mean of is equal to zero for the gridded stations because the Bea24 model is centered.
5. Evaluate the neutrality of the directivity condition for the NGA-W2 recordings. For a given rupture distance bin, a completely neutral directivity condition will have equal to zero and will have equal to . This procedure is repeated for different spectral periods (T).

|  |
| --- |
| A comparison of a diagram  Description automatically generated with medium confidence |

**Figure B1.** A map of the Landers earthquake showing at T=3 sec for at all locations (contours), at the locations with ground-motion recordings (circles), the fault trace (black), and hypocenter (star).

**Evaluation of Individual Earthquakes**

The values presented here, including Table B-1, are for T=3 seconds. The Landers earthquake (Figure B1) has = -0.018 from four stations within 20 km, and = -0.045 from 13 stations within 40 km. These values represent an approximately neutral directivity condition, as expected because of the good azimuthal coverage of the recording stations. Of the 22 events, 14 have < 0.1, representing less than about 10% difference from a neutral mean directivity condition for their respective recording station locations.

The Morgan Hill earthquake has = 0.247 from 17 stations within 40 km, representing a strong forward-directivity bias. The 2010 Darfield and 1999 Duzce earthquakes were relatively well-recorded and also have strong forward directivity biases ( = 0.237 and 0.185, respectively) for stations within 40 km. The value = 0.247 represents a 28% difference from a neutral mean directivity condition. The Manjiil earthquake has the strongest forward directivity bias; however, this is determined from a single available recording station within 40 km.

The Parkfield and Wenchuan earthquakes have the strongest backward-directivity biases of the events listed in Table B1 with = -0.045 and -0.046 for stations within 40 km, respectively. These backward-directivity biases, at only a few percent different from neutral, are weaker than the strongest forward-directivity cases (Morgan Hill, Darfield, and Duzce). This is reflected in the mean of from all 22 earthquakes, giving equal weight to each earthquake, which is 0.092 (9.6%) for stations within 20 km and 0.076 (7.9%) for stations within 40 km.

**Table B1.** Evaluation the neutrality of the directivity condition of three NGA-W2 GMMs, T=3 sec.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Earthquake Name** | **M** | **NGA-W2 EQID** | **No. of recordings with Rrup < 20 km** | **R < 20 km** | **No. of recordings with Rrup < 40 km** | **R < 40 km** |
| 1979 Coyote Lake | 5.74 | 48 | 7 | -0.018 | 9 | -0.041 |
| 1979 Imperial Valley | 6.53 | 50 | 23 | 0.097 | 31 | 0.109 |
| 1984 Morgan Hill | 6.19 | 90 | 9 | 0.31 | 17 | 0.247 |
| 1986 North Palm Springs | 6.06 | 101 | 7 | 0.072 | 9 | 0.063 |
| 1986 Chalfant Valley | 6.19 | 103 | 3 | -0.006 | 9 | -0.081 |
| 1987 Whittier Narrows | 5.99 | 113 | 18 | 0.011 | 58 | 0.013 |
| 1987 Superstition Hills | 6.54 | 116 | 6 | 0.12 | 8 | 0.092 |
| 1989 Loma Prieta | 6.93 | 118 | 17 | 0.147 | 37 | 0.108 |
| 1992 Landers | 7.28 | 125 | 4 | -0.018 | 13 | -0.045 |
| 1995 Kobe | 6.9 | 129 | 9 | 0.137 | 16 | 0.067 |
| 1999 Kocaeli | 7.35 | 136 | 4 | 0.014 | 6 | -0.015 |
| 1999 Duzce | 7.14 | 138 | 11 | 0.236 | 14 | 0.237 |
| 1990 Manjil | 7.37 | 144 | 1 | 0.365 | 1 | 0.365 |
| 1999 Hector Mine | 7.13 | 158 | 1 | 0.29 | 2 | 0.199 |
| 2002 Denali | 7.9 | 169 | 1 | -0.016 | 1 | -0.016 |
| 2000 Tottori | 6.6 | 176 | 10 | 0.002 | 21 | 0.042 |
| 2003 Bam | 6.5 | 178 | 1 | 0.285 | 1 | 0.285 |
| 2004 Parkfield | 6 | 179 | 54 | -0.059 | 57 | -0.045 |
| 2008 Wenchuan | 7.9 | 277 | 6 | -0.053 | 14 | -0.046 |
| 2010 El Mayor-Cucapah | 7.2 | 280 | 8 | -0.036 | 22 | -0.038 |
| 2010 Darfield | 7 | 281 | 15 | 0.179 | 29 | 0.185 |
| 2011 Christchurch | 6.2 | 346 | 17 | -0.029 | 22 | -0.006 |
|  | | | Mean: | 0.092 | Mean: | 0.076 |

**Neutrality of the Mean, All Events**

Table B2 lists results for the 22 earthquakes evaluated by combining all the recording stations from each earthquake into distance bins, for T=3 seconds. Figure B2 shows and (sample means of the median directivity effect at the NGA-W2 stations) in distance bins along with their 95% confidence intervals. When the 95% confidence intervals do not overlap zero, the sample means have a statistically significant difference from zero at the 95% confidence level. This figure shows that for most distance bins, the sample means of the median directivity adjustments from both models are not statistically significant from zero, indicating a neutral directivity condition. Exceptions are the distance bins 3-5 km, 10-15 km, and 40-60 km, where there is a small (statistically significant) bias towards the forward-directivity condition. There are no distance bins with a statistically significant bias towards backwards directivity. The Bea24 model has generally larger sample means than the CY14 model.

The bottom of Table B2 lists the statistics for distance bins of 0-20 km, 0-40 km, and 0-100 km. All three bins have a small bias towards forward directivity (both models) with the largest of these for 0-20 km bin ( = 0.051 ln units; 5.2% increase from neutral). All sites with Rrup < 40 km have = 0.049 ln units (5.0%), and sites with Rrup < 100 km have = 0.039 ln units (4.0%).

**Table B2.** Results by distance bin for the 22 earthquakes listed in Table B1.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Rrup bin (km) | No. of NGA-W2 stations in distance bin | Sample means and standard deviations; from NGA-W2 stations | | | | Population standard deviation; from gridded stations |
| Bea24:   (ln) | Bea24:   (ln) | CY14:   (ln) | CY14:   (ln) | Bea24:  (ln) |
| 0 Rrup < 1 | 12 | 0.071 | 0.243 | 0.036 | 0.198 | 0.173 |
| 1 Rrup < 3 | 26 | 0.042 | 0.230 | -0.003 | 0.130 | 0.192 |
| 3 Rrup < 5 | 29 | 0.098 | 0.190 | 0.023 | 0.091 | 0.209 |
| 5 Rrup < 10 | 62 | -0.024 | 0.253 | -0.017 | 0.119 | 0.220 |
| 10 Rrup < 15 | 45 | 0.131 | 0.255 | 0.054 | 0.179 | 0.221 |
| 15 Rrup < 20 | 58 | 0.047 | 0.210 | 0.010 | 0.182 | 0.219 |
| 20 Rrup < 30 | 99 | 0.047 | 0.209 | 0.012 | 0.140 | 0.215 |
| 30 Rrup < 40 | 66 | 0.041 | 0.192 | 0.000 | 0.159 | 0.211 |
| 40 Rrup < 60 | 110 | 0.056 | 0.183 | 0.005 | 0.117 | 0.189 |
| 60 Rrup < 80 | 102 | 0.014 | 0.114 | -0.003 | 0.025 | 0.096 |
| 80 Rrup < 100 | 76 | 0.001 | 0.013 | 0.000 | 0.000 | 0.000 |
|  | | | | | | |
| Rrup 20 | 232 | 0.051 | 0.237 | 0.013 | 0.153 | 0.216 |
| Rrup 40 | 397 | 0.049 | 0.223 | 0.010 | 0.151 | 0.214 |
| Rrup 100 | 685 | 0.039 | 0.191 | 0.006 | 0.124 | 0.145 |

|  |
| --- |
| A graph with numbers and lines  Description automatically generated |

**Figure B2.** Distance bins of and (sample means of the directivity effect at the NGA-W2 stations) with their 95% confidence intervals. Confidence interval bar distance widths indicate the extent of the distance bins.

|  |
| --- |
| A graph with numbers and symbols  Description automatically generated with medium confidence |

**Figure B3.** Distance bins of (sample standard deviation of the directivity effect at the NGA-W2 stations) with 95% confidence intervals, and (population standard deviation of the directivity effect at the gridded stations). Confidence interval bar distance widths indicate the extent of the distance bins.

**Neutrality of the Standard Deviation, All Events**

A completely neutral directivity condition for the standard deviation will have (sample standard deviation of the directivity effect at the NGA-W2 stations) equal to (population standard deviation of the directivity effect at the gridded stations).

Figure B3 shows and with 95% confidence intervals for T=3 seconds and for the Bea24 model. The CY14 model predictions (based on ) were provided only at the NGA-W2 recording stations, not the gridded stations. As a result, Table B2 provides the sample standard deviations for CY14, but there is no comparison made for this model.

All distance bins with Rrup < 60 km have a population standard deviation contained within the sample standard deviation confidence intervals. At distances shorter than 3 km, where the recorded data is very limited, the confidence intervals are wide and appears biased high. The distance bins with Rrup > 60 have a very small positive bias in .

**Conclusions on Neutrality of the NGA-W2 GMMs**

Table B3 summarizes the NGA-W2 directivity condition of the mean using Bea24. At long periods, there is a small bias towards the forward directivity condition. The largest bias, of approximately 6%, is for 10 seconds period and for sites with Rrup 20 km. When the distance range for this evaluation is reduced to include only sites with Rrup 10 km, the forward directivity bias is smaller for all spectral periods; this is consistent with the finer distance bins used in Supplemental Appendix B, where the Rrup 10-15 km bin has the largest . The bias towards the forward directivity condition generally decreases with decreasing spectral period. As the distance range of sites to consider is expanded from Rrup 20 to Rrup 40 km and Rrup 100 km, the forward directivity bias also decreases.

**Table B3.** Sample mean of by distance bin and period for the 22 earthquakes evaluated in Supplemental Appendix B.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Rrup bin  (km) | T=1 sec (ln) | T=3 sec  (ln) | T=5 sec  (ln) | T=7.5 sec  (ln) | T=10 sec  (ln) |
| Rrup 10 | -0.005 | 0.026 | 0.035 | 0.031 | 0.041 |
| Rrup 20 | 0.009 | 0.051 | 0.059 | 0.061 | 0.065 |
| Rrup 40 | 0.013 | 0.049 | 0.058 | 0.054 | 0.055 |
| Rrup 100 | 0.010 | 0.039 | 0.046 | 0.040 | 0.039 |

For distance bins of 3-5 km, 10-15 km, and 40-60 km, there is a small bias (statistically significant at the 95% confidence level) towards the forward-directivity condition, as shown in Supplemental Appendix B. For all other distance bins evaluated, the sample means of the directivity adjustments models are not statistically significant from zero, indicating a neutral directivity condition.

For comparison, D19 binned the NGA-W2 database recordings by distance and found mean DPP values ranging from 0.002 (Rrup < 40) to 0.13 (Rrup < 20), with mean of all DPP values in the NGA-W2 database, regardless of distance, of 0.011. For reference, at 3 seconds period, a value of DPP = 0.01 corresponds to approximately = 0.002 natural log units of amplification (0.2%), DPP = 0.1 to approximately = 0.025 ln units (2.5%). D19 concluded that, on average, the NGA-W2 data reflect a directivity-neutral condition for the median of ground-motions using DPP.

In summary, the mean directivity condition of the NGA-W2 recordings (and GMMs derived from them) have a small bias towards forward directivity. That bias varies with distance and spectral period; at its largest, it is approximately 5% at long spectral periods and close distances. The forward-directivity bias is smaller at shorter spectral periods and larger distances. To the extent that Bea24 is based on the NGA-W2 recordings, we have assumed that the bias in the mean directivity condition, over all distances and spectral periods, is small enough to ignore. This conclusion is consistent with D19.

The standard deviations of the directivity effect at the NGA-W2 recording stations are not inconsistent with the standard deviations of the directivity effect at all possible station locations. Therefore, to the extent that the standard deviation of the NGA-W2 GMM models contain rupture directivity effects, these models can be considered to reflect a directivity-neutral condition of the standard deviation of directivity effects.

# Appendix C: Calculation of

Equations 3a through 3d (repeated below) require the calculation of the directivity predictor centering term at a given rupture distance,, defined as the mean value over a suite of sites located at the same distance to an earthquake over all azimuths (sites on a racetrack). The value of is specific to a scenario with given hypocenter location, rupture dimensions, and rupture distance, .

There are several ways the calculation of can be approached for a given racetrack:

1. Closed form (analytic) solution. Solve for the mean value of using the Fundamental Theorem of Calculus (Mean Value Theorem). The mean value is the definite integral of the continuous function .
2. Numerical method. Solve for the mean value of numerically.
3. Functional form method. Create a model for which approximates.

Al Atik et al., (2023) used the functional form method for their statewide hazard analysis with (Chiou and Youngs, 2014). Al Atik et al., (2023) created a database of scenario earthquakes with a range of rupture dimensions and hypocenter locations and used a functional form to fit a model for given the hypocenter location, rupture distance, and rupture length. This approach is straightforward to apply but may be subject to the largest errors out of the three methods.

The closed form solution is preferable to the other methods because it will be exact and efficient to calculate. We attempted to solve the integrals defined below, first by hand, and then with integral solver software available online, but the solutions were extremely complex and seemed unnecessary. We found it was most straightforward to use the numerical method and approximately solve the integrals by splitting the racetrack into four regions, as shown below.

**Regions 1 and 2: between the fault ends**

The figure below shows the map view of a scenario vertical strike-slip rupture (red) with coordinate system origin at the hypocenter (star). The along-strike dimension is x and the strike perpendicular dimension is y. Region 1 is between the origin and the end of the rupture in the along strike direction (). Region 2 is between the origin and the end of the rupture in the anti-strike direction ().

A red line with black text

Description automatically generated

Considering the racetrack at rupture distance R, and setting the rake angle to zero, the following result from Equations 3a through 3d for the along-strike direction:

Substituting and integrating , we have for Region 1:

For the anti-strike direction (Region 2), equivalently:

**Regions 3 and 4: off the fault ends**

Off the ends of the fault, is constant and decreases with increasing because the racetrack curves towards the along-strike direction. Region 3 has and Region 4 has (not shown).

A red line with black text

Description automatically generated

For Region 3, the following result from Equations 3a through 3d:

(circle formula)

Substituting and integrating, we have:

For the anti-strike direction (Region 4), equivalently:

**All Regions**

Using the Mean Value Theorem, is the definite integral of the continuous function divided by the interval length, or equivalently:

If solved, this would represent the closed form solution. Only one side of the fault is included in this calculation because of symmetry. As a result, when this method is applied to curved ruptures the solution for is an approximation.

In the Matlab implementation of the model provided with this article, the numerical method is used. In this method, the integrands of , , , and are calculated numerically for sites using the spacing = 0.1 km, and the mean value for sites in all four regions is computed.

# Appendix D: Median Model Development

The median model is developed following a similar approach taken as in Bea20. The approach utilizes within-event residuals and fits a parametric model relating the Bea24 directivity predictor () to these residuals to model the directivity effect () and its period-dependence. Residuals are for the RotD50 (Boore et al., 2010) horizontal component of 5% damped spectral acceleration, calculated from three NGA-W2 GMMs (Abrahamson et al., 2014; Boore et al., 2014; Campbell and Bozorgnia, 2014). Two databases are used: one from the suite of simulations described in Bea20, and the second from NGA-W2 recordings of strike-slip earthquakes with a finite fault model and at least 5 recordings.

The steps taken to develop the median model are performed on both databases (simulated and recorded) separately. The steps are:

For each event:

1. Calculate the value Bea24 centered directivity predictor () at each site where there is data (recorded or simulated).
2. For each spectral period, use the residuals to fit the free parameters from Equation 2: and . The limiting upper and lower bound of is , and represents the slope of the relationship between and .

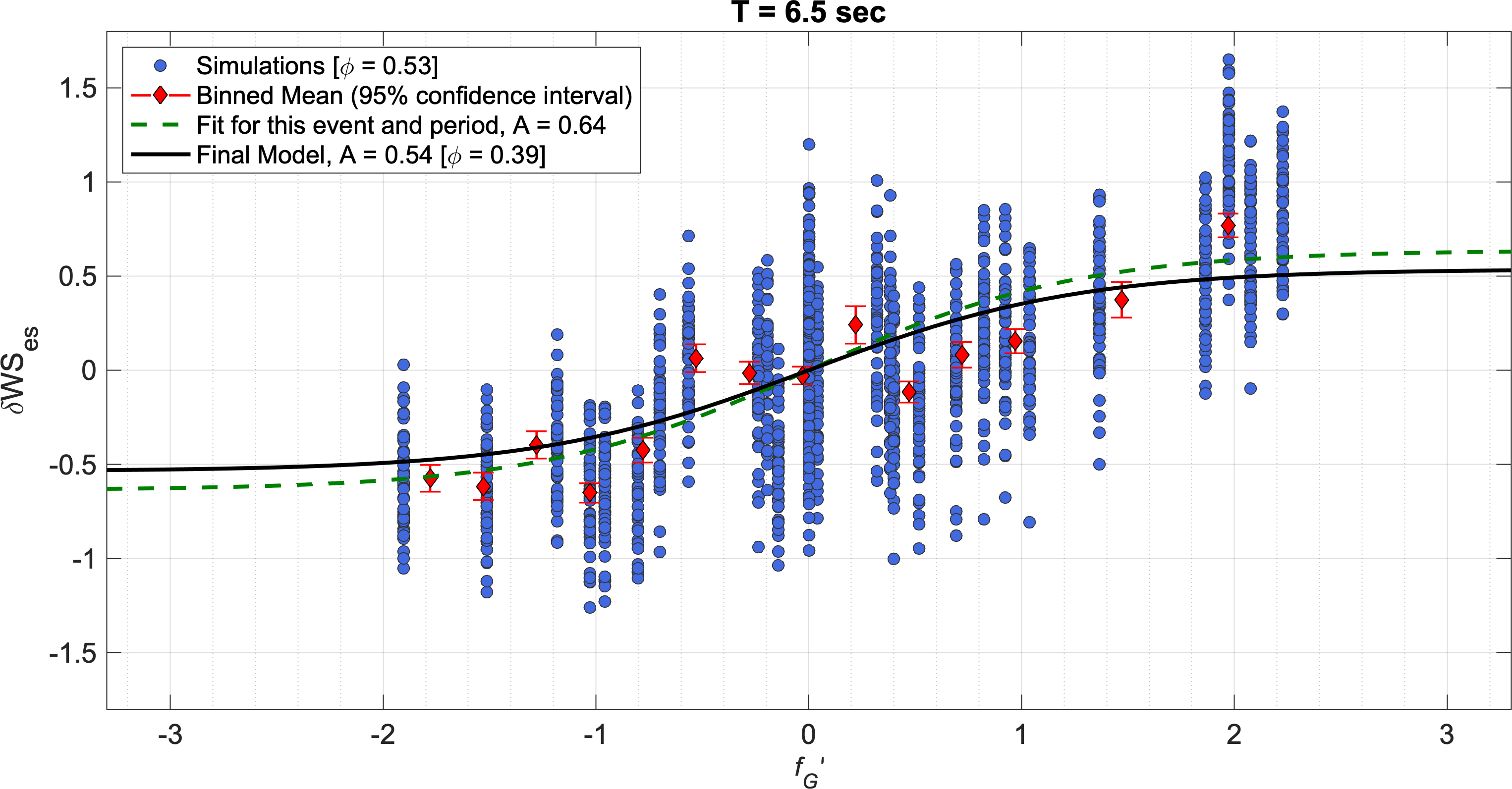
Considering all events:

1. Evaluate the magnitude and period dependence of and of the peak value of ().
2. Evaluate and model the magnitude dependence of the period corresponding to (; Equation 4b)
3. Model with a Gaussian function of period (Equation 4a). This function has maximum value and is centered on period . The standard deviation (width parameter) of the Gaussian function is .
4. Perform a nonlinear least squares regression using the database residuals to derive the period-independent model coefficients and . These coefficients are estimated by a joint regression with combined data of all periods. Because there is potential for tradeoff in the coefficients (in this case, arriving at the same result with two sets of values for and ), in the final version we fixed . The value of was determined to be 1.58 based on step 3 in the above process.

The final step (regression) was performed for both datasets independently to inform the rupture directivity modeling. We found that the simulations, which have significantly more near-fault stations and better azimuthal coverage than the data, demonstrate stronger scaling with the directivity parameters and over a broader period range. The NGA-W2 data generally demonstrate weaker scaling over a narrower period range. As a result, directivity model coefficients are provided based on both datasets. The models derived from both datasets have the same functional form and only differ in their coefficients.

**Simulations**

Figure D1 provides an example of Step 2 for the Landers earthquake simulation using the Graves and Pitarka (2014) simulation method; the simulations are described in Bea20. This figure shows within-event residuals (; blue) versus at T=6.5 sec. Binned means with 95% confidence interval for small intervals of are shown in red. The model for fit to the simulation residuals shown (at this period) is given by the dashed green line. The final Bea24 model (Model 1; resulting from Step 7) is shown in black. By applying the final Bea24 model to these residuals, the within-event residual standard deviation reduces from 0.53 to 0.39 natural log units. As described in the above approach outline, Step 2 is repeated for each event and spectral period.



**Figure D1.** Summary of Step 2 for the Landers earthquake simulation, at T=6.5 sec.

Figure D2 summarizes the results of Step 3 for the simulated database, where each line color or marker represents one of the 8 simulated events. Figure D3 (upper panel) shows Step 4, and the lower panel shows Step 5. The standard deviation (width parameter) of the Gaussian function is 0.38.

Figure D4 shows within-event residuals () versus at period (this period varies by event) from complete set of the simulations. Binned means with 95% confidence interval for small intervals of are shown in red. The Bea24 median model (result of Step 6) is shown in black, with the value of determined from the regression. Use of the logistic function saturates for extreme values of ; this causes the flattening of the black curve.

A collage of graphs

Description automatically generated

A collage of graphs

Description automatically generated

**Figure D2**. Summary of Step 3 for the simulated dataset.

A collage of graphs

Description automatically generated

A collage of graphs

Description automatically generated

**Figure D3.** Summary of Steps 4-5 for the simulated dataset. The colored lines in the bottom panel correspond to the legend in Figure D2.

A graph showing a graph of a graph

Description automatically generated with medium confidence

**Figure D4.** Within-event residuals () versus at period (this period varies by event) from the simulations. Binned means with 95% confidence interval for small intervals of are shown in red. The Bea24 median model from the simulations (Model 1) is shown in black.

**NGA-W2**

Figure D5 summarizes Steps 3-5 as applied to the NGA-W2 data. Figure D6 shows within-event residuals () versus at period (this period varies by event) from the recorded data. Binned means with 95% confidence interval for small intervals of are shown in red. The Bea24 median model from the recorded data (Model 2; result of Step 5) is shown in the solid black, and the model from simulations (Model 1) is given by the dashed line.

**Simulated and Recorded events**

Figure D7 summarizes the model approach for the combined dataset, where simulations are in grey and recorded events are in blue. Panel a shows the period dependence of (Step 3), panel b shows the Gaussian function of period used to model (Step 5), panel c shows the evaluation for magnitude dependence of (Step 3), and panel d shows the magnitude dependence of (Step 4).

A collage of graphs

Description automatically generated

**Figure D5.** Summary of Steps 3-5 for the recorded dataset.

A graph with blue and red dots

Description automatically generated

**Figure D6.** Within-event residuals () versus at period (this period varies by event) from the recorded dataset. Binned means with 95% confidence interval for small intervals of are shown in red. The Bea24 median model from the recorded data (Model 2) is shown in the solid black, and the model from simulations (Model 1) is given by the dashed line.

A close-up of a graph

Description automatically generated

**Figure D7.** Summary of Steps 3-5 for the simulated and recorded events.

# Appendix E: Aleatory Variability Model Development

**Part 1: NGA-W2 Data**

Using the final model coefficients, the within-event residuals of the NGA-W2 data are adjusted for median directivity effects. We calculate before and after this adjustment for the NGA-W2 events used to develop the median model. In these calculations, stations are included only if they are within the distance range for which the model predicts directivity adjustments (i.e. for rupture distances less than 80 km). At both stages, the within-event residuals are inspected versus rupture distance and the centered directivity predictor, , as shown in Figure E1, where the red circles are the residuals after applying the directivity median adjustment. The purpose of this evaluation is to confirm that the distance scaling is not adversely affected, and to visualize the effect of the directivity model, which reduces the residuals for positive values of , and increases them for negative . In the case of the Landers earthquake (T=7.5 sec), application of the model reduces as indicated in the bottom panel of Figure E1.

A screenshot of a graph

Description automatically generated

**Figure E1.** The effect of applying the directivity adjustment model on Abrahamson et al. (2014) within-event residuals for the Landers earthquake, considering sites within 80 km rupture distance, at T=7.5 sec. Blue and green circles are the residuals before the adjustment, and smaller red circles are after.

The period dependence of from all events is shown in Figure 2 of the main article, and the values are listed in Table 2.

**Part 2: Simulations**

The procedure for the NGA-W2 data was followed for the simulations used to develop the median model. The period dependence of from all simulated events is shown in Figure 2 of the main article, and the values are listed in Table 2.

# Appendix F: Examples

This appendix provides a series of example applications of the Bea24 model.

**Example 1**

This scenario (borrowed from the test scenarios in Table 1.2 of Spudich et al., 2013) is a **M**7.2 vertical strike-slip rupture with rake angle of 180 degrees, and with length of 80 km and down-dip width of 15 km. The hypocenter (red star) is located 10 km from the southern end of the rupture, at 10 km depth. The origin for the GC2 calculation required by Bea24 is the rupture surface trace ordinate of the up-dip projection of the hypocenter, which is the same as the epicenter for a vertically dipping rupture. Figure F1 shows the spatial fields of median amplification (exponential of ) at T=3 sec spectral period, for Model 1 (simulation-based, left) and Model 2 (NGA-W2 data-based, right). The spatial patterns of amplification are the same between the models because both models have the same functional form and only differ in their coefficients.

|  |
| --- |
| A comparison of a model  Description automatically generated with medium confidence |

**Figure F1.** The spatial fields of amplification for the Example 1 scenario rupture, T=3 sec.

**Example 2**

This example is based on the 2002 Denali, Alaska earthquake using the rupture model adopted by NGA-West2. This rupture model has **M**7.9 and is composed of three strands. Strand 1, which contains the hypocenter, is 45.1 km in length and 24 km width and is shallowly dipping (dip=32 deg). Strand 2 is 213.7 km long and 15 km wide, with 80-degree dip. Finally, Strand 3 is 67.9 km long and 15 km wide, with 90-degree dip. A representative rake of 171 degrees is assumed based on the larger rupture areas of strands 2 and 3 compared to strand 1. The origin for the GC2 calculation required by Bea24 is the rupture surface trace ordinate of the up-dip projection of the hypocenter (white circle). Figure F2 shows the spatial field of Bea24 median amplification (exponential of ) at T=5 sec spectral period, for Model 1 (simulation-based).

|  |
| --- |
| A map of a map  Description automatically generated |

**Figure F2.** The spatial fields of amplification for the Example 2 (Denali earthquake) rupture, T=5 sec.

**Example 3**

This example is for the **M**7.8 scenario earthquake representing the 1906 San Francisco event. The rupture is defined as in Aagaard et al., (2009). The rupture is 478 km in length and has a representative rake angle of zero degrees. Figure F3 shows the spatial field of Bea24 (Model 1) median amplification (exponential of ) at T=7 sec spectral period.

|  |
| --- |
| A map of a meteorite  Description automatically generated with medium confidence |

**Figure F3.** The spatial fields of amplification for the Example 3 (1906 SF earthquake) rupture, T=7 sec.

**Example 4**

In this example, the model is applied to a hypothetical scenario earthquake. This strike-slip rupture is set up with two disconnected strands, each vertically dipping and with 22 km width. The first strand has strike of 0 degrees and 40 km length, with the hypocenter located 30 km from the southern end. The second has 45-degree strike and 28.28 km length. There is a gap of about 14 km between these strands, which is 9 km larger than the “maximum jump distance” allowed in the plausibility filters used to define UCERF3 ruptures (Field et al., 2013; Appendix T). This example demonstrates the flexibility of the model to accommodate ruptures with gaps between strands. Using the Leonard (2010) relationship for rupture area and magnitude, the scenario is prescribed **M**7.15. Figure F4 shows the spatial field of Bea24 (Model 1) median amplification (exponential of ) at T=3 sec spectral period.

|  |
| --- |
|  |

**Figure F4.** The spatial fields of amplification for the Example 4 rupture, T=3 sec.

**Example 5**

Figure F5 shows Bea24 (Model 1) median amplification maps for four vertical strike-slip scenarios, each north-south striking, and with hypocenter placed at the center of the rupture plane. The maps are shown at the spectral period corresponding to the scenario peak amplification (). The parameters used to define these scenarios are listed in Table F1. With increasing magnitude, and the maximum distance in the distance taper also increase.

|  |
| --- |
|  |

**Figure F5.** The mapped median directivity adjustment (amplification) for the four strike-slip scenarios described in the text.

**Table F1.** Parameters of the scenarios used to illustrate the model behavior with magnitude.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number | M | Length (km) | Width (km) | (sec) |
| Strike-slip 1 | 6.0 | 12.6 | 8.12 | 1.9 |
| Strike-slip 2 | 6.5 | 25.1 | 12.8 | 3.0 |
| Strike-slip 3 | 7.0 | 50.2 | 20.4 | 4.7 |
| Strike-slip 4 | 8.0 | 465.0 | 22.0 | 12.0 (10.0 sec shown) |

**Example 6**

Figure F6 shows T=3 sec Bea24 (Model 1) median amplification maps for four vertical strike-slip scenarios, each north-south striking and with hypocenters placed at the center of the rupture plane, with increasing (0, 5, 10, and 15 km).

|  |
| --- |
| A group of images of different colors  Description automatically generated with medium confidence |

**Figure F6.** The mapped median directivity adjustment (amplification) at T=3 sec for four strike-slip scenarios with increasing .

**Example 7**

This example applies Bea24 to a series of vertical strike-slip scenario earthquakes with increasing magnitude. In each scenario, the Bea24 (Model 1) is applied with unknown hypocenter locations modeled as described in *Model Implementation: Deterministic*. The Melgar and Hayes (2019) hypocenter distribution model is used to assign such that such that . The maps in Figures F7 through F10 show the spatial fields of and , which are calculated using Equations 8 and 9, respectively.

|  |
| --- |
|  |

**Figure F7**. Application of the Bea24 directivity model (Model 1) with unknown hypocenter location to an **M**6.0 scenario earthquake with 13 km rupture length, at 3 seconds spectral period.   
Left: Contours of (ln units). Right: Contours of (ln units).

|  |
| --- |
| A comparison of a flower with a blue background  Description automatically generated with medium confidence |

**Figure F8.** Application of the Bea24 directivity model (Model 1) with unknown hypocenter location to an **M**6.6 scenario earthquake with 30 km rupture length, at 3 seconds spectral period.   
Left: Contours of (ln units). Right: Contours of (ln units).

|  |
| --- |
|  |

**Figure F9.** Application of the Bea24 directivity model (Model 1) with unknown hypocenter location to an **M**7.2 scenario earthquake with 80 km rupture length, at 3 seconds spectral period.   
Left: Contours of (ln units). Right: Contours of (ln units).

|  |
| --- |
|  |

**Figure F10.** Application of the Bea24 directivity model (Model 1) with unknown hypocenter location to an **M**7.8 scenario earthquake with 400 km rupture length, at 10 seconds spectral period.   
Left: Contours of (ln units). Right: Contours of (ln units).