A High-frequency Distance Metric in Ground-Motion Prediction Equations Based on Seismic Array Back-Projections

Tian Feng[1], Lingsen Meng[1*]
1. Earth, Planetary and Space Sciences, University of California, Los Angeles, CA 90095, USA.

*meng@ess.ucla.edu, 1805 Geology Building
University of California, Los Angeles, Phone: (310) 825-1885

Key Points:

1. We propose a new frequency-dependent distance metric R_{hf} for use in ground-motion prediction equations (GMPEs).
2. R_{hf} improves the accuracy of high-frequency (0.5 – 4 Hz) ground-motion prediction for the megathrust earthquakes
3. GMPEs utilizing R_{hf} can potentially provide rapid ground-motion intensity estimation for megathrust earthquakes.

Abstract

Typical ground-motion prediction equations (GMPEs) measure source-to-site distances relative to the closest point on the rupture plane (R_{rup}). However, for megathrust earthquakes (Mw > 8), the over-simplification of the earthquake source characteristics in distance metrics results in significant bias. Recent studies suggest that the high-frequency (HF) and low-frequency energy tend to emanate from different portions of the megathrusts. This phenomenon motivates an alternative distance metric based on the array Back-Projection (BP) imaging technique that effectively captures regions releasing HF energy. Herein, we define a HF distance metric (R_{hf}) as the distance from the site to the high-frequency radiation zone. We study five Mw > 7.2 megathrust earthquakes in Japan and Chile and find that R_{hf} outperforms R_{rup} in predicting the ground shaking intensity between 0.5 - 4 Hz. We consider R_{hf} as a complementary measure to conventional GMPE distance metrics and a more accurate ground-motion predictor in many cases.

1. Introduction

Recent advances in earthquake source imaging of major subduction earthquakes highlight the frequency-dependent and depth-varying seismic radiation along the plate interfaces. Low-frequency (LF) energy mainly emanates from the shallower portion of the megathrusts while dominant high-frequency (HF) energy often radiates from the deeper portion of the megathrust (Lay et al., 2012; Yao et al., 2013). In the 2011 Mw 9.0 Tohoku-Oki earthquake, the along-dip separation between fault segments that radiate HF and LF seismic waves is as large as about 100 km (Meng et al., 2011; Yao et al., 2011). These frequency-dependent source behaviors have significant impacts on the spatial pattern of the near-field ground shaking and therefore should be considered in ground-motion predictions, which are crucial to both earthquake source sciences and civil engineering applications (Anderson et al., 2007; Bozorgnia et al., 2014). The standard approach to predict the expected seismic shaking intensity is the ground-motion prediction equations (GMPEs). A GMPE provides estimates of ground-motion intensity measures based on regression analysis of the source, path, and site effects (Atkinson et al., 2007; Gregor et al, 2014; Abrahamson et al., 2014; Boore et al., 1994, 1997 and 2014, etc). GMPEs are widely used in Probabilistic Seismic Hazard Analysis (PSHA) due to their simplicity, as it does not require local 3D velocity models, previously observed strong motion records, or additional input such as empirical Green's functions.

The source-to-site distance is an important parameter of the GMPEs that represents the path effect of ground motion during an earthquake (Fig. S1). The classical GMPEs use point-source measures relative to the epicentral and hypocentral distances (R_{epi} and R_{hyp}, respectively). Considering the finite dimensions of large earthquakes, point-source distance metrics were replaced by extended-source distance metrics. Two widely used metrics are the rupture distance R_{rup} (Douglas, 2003), representing the closest distance (orthogonal) to the fault plane, and the Joyner-Boore distance R_{JB} (Joyner and Boore, 1981), the shortest distance to the surface projection of the fault plane. A variant of R_{JB}, the equivalent hypocentral distance EHD, measures the shortest distance to the area of principle moment release on the fault (Ohno et al., 1993; Si and Midorikawa, 2000). Many current GMPEs based on these frequency-independent distance metrics rely on numerous modifications of distance to accurately model near-field ground motions. Because these modifications (e.g. additional distance definitions, saturation terms) often result in complicate functional forms without seismological basis, Thompson and Baltay (2018) advocate for the use of mean rupture distance R_p, which is frequency-dependent and in a simpler form. Frequency-independent distance metrics perform reasonably well in predicting the shaking intensity of small earthquakes, but the frequency dependence of source distances can become more significant for large subduction-zone earthquakes (Meng et al., 2011; Lay et al., 2012; Yao et al., 2013). In the case of the Tohoku earthquake, GMPEs using frequency-independent distance metrics produce substantial biases in the estimated level of shaking. Stewart et al. (2013) compare the performances of four different GMPEs (Atkinson and Boore, 2007; Abrahamson et al., 2016; Zhao et al., 2006; Si and Midorikawa, 2000), all of which underestimate the near-field (< 100 km) ground-motion intensity of the Tohoku earthquake at short periods and overestimate shaking at long periods. To address this bias, a statistical event term is applied after residual analysis is performed, comparing GMPEs to observed ground-motion intensity (Stewart et al., 2013). This term is positive at short periods and zero or negative at long periods. The physical basis of this term is unknown, and it is not clear how the event term can be generalized to other megathrust earthquakes.

Here, based on the observations that the short-and-long-period seismic radiations are generated from different portions of the megathrust, we attributed this event term to the frequency-dependent source-to-site distances. In the following sections, we propose a new frequency-dependent distance metric R_{hf} based on Back-Projection Imaging and compare it with R_{rup} in five large megathrust events: the 2011 Mw 9.0 Tohoku earthquake, the 2003 Mw 8.0 Tokachi earthquake, the 2005 Mw 7.2 off-Miyagi earthquake, the 2014 Mw 8.1 Iquique earthquake, and the 2015 Mw 8.3 Illapel earthquake. In all cases, R_{hf} outperforms R_{rup} in the HF band (0.5 – 2 Hz) by showing little or no bias in ground-motion estimations.

2. High-frequency Distance Metric

2.1. Back-Projection Imaging (BP)

The frequency-dependent radiation patterns, particularly at high-frequencies (HF), can be effectively captured with the Back-Projection (BP) approach. BP is an earthquake-rupture imaging technique utilizing the coherent teleseismic P wavefield based on seismic array processing (Ishii et al., 2005; Kiser and Ishii, 2017; Kruger and Ohrnberger, 2005; Meng et al., 2011). Back-tracking of seismic waves recorded by dense arrays allows BP to determine the spatiotemporal properties of the rupture (length, direction, speed, segmentation) (Fig. S2). Over recent decades, development of large-scale dense seismic networks has enabled BP imaging of the rupture process of major large earthquakes.

The advantages of BP are three-fold. First, in contrast to conventional finite-fault modeling, BP is an imaging procedure that does not involve the solution of an inverse problem. Thus, it does not suffer from being ill-posed or underdetermined, and does not rely on restrictive parameterizations of the rupture kinematics and fault geometry, or additional smoothing (Kiser and Ishii, 2017). Because fewer assumptions and choices are involved, BP results tend to be more robust and are more mutually consistent between different research groups than finite fault models (Kiser and Ishii, 2017). Second, BP algorithms (such as beamforming) are simple and do not require heavy computations, such as large-scale matrix inversions or wave-propagation simulations. Therefore, BP is easily automated (e.g. IRIS Back-Projection products) and well suited for implementation in real-time for the purpose of rapid seismic hazard assessment or earthquake/tsunami early warning (Meng et al., 2014; An and Meng, 2016). Third, since BP does not attempt to deterministically fit seismic waves and relies only on the coherent phase of seismic array signals, it does not require the detailed knowledge of the Green's functions. This benefit allows BP to be applied to HF wave-field ($f \geq 1$ Hz), which is generally thought to be too noisy and uncertain to be modeled deterministically.

In comparison, classic source inversions requiring the deterministic Green's function are limited to relatively low frequencies ($f < 0.1$ Hz), for which the crustal structure is reasonably approximated. The complex rupture patterns that emerge in dynamic ruptures require finer details of the source process that are apparent in high frequencies ($f > 1$ Hz), which are hard to retrieve through using conventional source inversions. Furthermore, the ability of BP to image HF sources is crucial to improve the prediction of strong ground motion, since the typical frequency range of intense shaking is between 1-10 Hz (e.g., Anderson et al., 2007; Sokolov and Chernov, 1998). The natural frequency of building response falls in this particular frequency band: a rule of thumb is that the building period equals the number of stories divided by 10, so the 1-10 Hz band is roughly responsible for the shaking of 1-10 story residential and commercial buildings.

2.2. High-frequency Radiator (HFR) and High-frequency Distance (R_{hf})

To systematically evaluate the performances of the frequency-dependent distance metrics in ground-motion predictions, we studied five large earthquakes in Japan and Chile (Table. S1). These earthquakes are chosen because they occurred after 2003, when strong motion data are available for systematic strong-motion analysis, and because their GCMT solutions indicate typical thrust mechanism (Fig. S3). In our study, we adopt the hypocenters and magnitudes determined by Japan Meteorological Agency and Centro Sismólogico Nacional (CSN). We perform BP analysis of the vertical velocity seismograms recorded by broadband station in North America for the five earthquakes at 4 representative frequencies (0.5 Hz, 1 Hz, 2 Hz, and 4 Hz). The data are available through the IRIS data management center. We chose stations at teleseismic distances between 40º to 85º because of simplicity and the high coherency of the direct P arrivals. We adopt the Multiple Signal Classification (MUSIC) Back-Projection method with the reference window strategy which provides superior resolution and less artifact than conventional beamforming (Meng et al., 2011; Meng et al., 2012). We first filter the seismograms aligned by initial P arrivals and then determine the HF sources (large energy bursts) as the peak locations of short-period energy released in each time frame with an interval of 1 s. We name the location of these energy bursts as high-frequency radiator (HFR) since 0.5 – 4 Hz is relatively high in the ground-motion spectrum of earthquakes.

The spatial distribution of HFRs at four frequencies show clear along-dip patterns, with higher frequency HFRs closer to the mainland (Fig. 1a). Assuming the coseismic rupture occurring on the megathrust, we estimate the depths of HFRs by projecting HFRs onto the pre-defined plate interface (SLAB 1.0, Hayes et al., 2012). We then define the high-frequency distance (R_{hf}) as the shortest distance from a given site to all HFRs during the earthquake, analogous to the distance metric R_{rup}. Since the locations of HFRs can be evaluated with BP performed at different frequency bands, the distance metric R_{hf} (f) is also frequency-dependent. To avoid the contamination of the coda wave, only the HFRs within the earthquake duration are used to determine R_{hf}. The event duration is determined according to the average SCARDEC moment-release functions (Vallée and Douet, 2016), and we consider the earthquake over when the moment release function falls below 1% of the peak moment rate.

2.3. Effective Period and Uncertainty of R_{hf}

It is worth noting that the effective period of R_{hf} metric is limited from 0.1 s to 10 s. In this study, we choose 0.25 to 2 s, where teleseismic BP performs best. When the period is longer than 10 s, the BP result is less reliable because of the swimming artifact, a systematic drift resulting from signal non-stationarity (Meng et al., 2012). The swimming artifact produces apparent energy bursts migrating in the direction of the array as time lapses because of the tradeoff between source origin time and distances (Koper et al., 2011). The swimming artifact is most prominent in low-frequency beamforming Back-Projection and can lead to apparent frequency-dependent rupture behavior (Meng et al., 2012). On the other hand, when period is lower than 0.1 s, the waveforms are too noisy and incoherent to achieve the reliable BP results. To extend the valid frequency range of GMPEs, we could combine R_{hf} with other distance metrics (e.g. to use R_{hypo} when studying longer period ground motion).

Recent studies show that the spatial uncertainty of BP mainly comes from the travel time error due to 3D path effects (Kiser and Ishii, 2013; Meng et al., 2016; Fan and Shearer, 2017; Chen et al., 2018). This uncertainty is generated by applying a uniform the hypocenter time correction across the entire rupture area without considering the structural heterogeneity of the earth's interior. The uncertainty can be inferred by comparing the BP-imaged aftershcoks with their catalog locations. Fan and Shearer (2017) analyze the BPs of 46 M5-M7 earthquakes in Japan and determine that the median location error is about 25 km. Similarly, Chen et al. (2018) study two deep-earthquakes (M > 5) near the 2015 Mw 7.9 Bonin earthquake and estimate the spatial error around 15km. Besides, Meng et al. (2018) perform the slowness-enhanced BP based on the aftershocks of 2015 Mw 8.3 Illapel earthquake and reduce the spatial error from 24 km to 8 km. Here, we consider HFR location error and therefore R_{hf} error are on the order of 20 km, which is reasonably small comparing to the dimension of megathrust events (M > 7) and the gaps between HF and LF sources.

2.4. Comparison between R_{hf} and R_{rup}

Traditionally, when calculating R_{rup}, the preferred finite fault model is often trimmed by removing low-slip (< 10% of the peak slip) subfaults. Then, R_{rup} is measured as the closest distance to the edge of trimmed fault area (> 10% of peak slip, blue pattern in Fig. 1a). This procedure allows R_{rup} to represent the source-receiver distance from the principle slip area and is a routine processing in the PEER NGA-West2 project (Ancheta et al., 2012; Stewart et al., 2012). Because different data sources (e.g. teleseismic, strong motion, GPS, InSAR, tsunami) are used alone or jointly in the finite fault model inversions, the edge of trimmed fault areas could be different under the same criteria. For the Iquique and Illapel earthquakes, the edge is spatially correlated with higher frequency HFRs (Fig. 1a). Thus, we can deduce that frequency-independent R_{rup} is close to R_{hf} at high frequency while smaller than R_{hf} at low frequency. This difference between R_{rup} and R_{hf} increase from HF to LF. However, for the Tokachi, off-Miyagi, and Tohoku earthquake, the edge is closer to lower frequency HFRs. Accordingly, R_{hf} is close to R_{rup} at low frequency and decreases with the frequency.

3. Ground-Motion Predictions

3.1. Strong-motion Data

We collect the strong motion data (horizontal accelerograms) recorded on the Japanese K-Net (Kyoshin network) and Chilean CSN stations, and then compare them with the GMPE predictions. Owing to the densely distributed K-NET stations, over 300 recordings of the ground-motion observations are available to study GMPEs for each Japan event. For the two Chile events, 11 and 48 recordings are used for analysis respectively. We use 5% damping pseudo-spectra acceleration (PSa), the maximum acceleration response of equivalent single degree of freedom systems, as our ground-motion intensity measure, extracted directly from the seismograms. PSa is commonly used in GMPEs, designed to approximate first-order building response. We use the Newmark Method for linear systems to compute the PSa in our study (Chopra, 2001).

3.2. Ground-Motion Prediction Equations (GMPEs)

As an example of the utility of representing the high-frequency source generation, we replace the distance metric in existing GMPE with R_{hf} and compare the performances of the GMPE with different distance metrics. We adopted the BC2016 GMPE designed for subduction earthquakes with magnitudes up to Mw 9.0 (Abrahamson et al., 2016). Even though some other GMPEs use subduction zone data, they are not suitable for large earthquakes in the subduction zone, i.e. they don't consider the forearc/backarc effect, which Stewart et al. (2013) shows is essential for the 2011 Tohoku earthquake. The original form of the GMPE uses rupture distance R_{rup}, the shortest distance from the site to the fault plane. The original form of BC2016 for the plate-interface earthquakes is:

$$ln(PSa) = \delta C_1 + F_M(M) + F_D\left(R_{rup}, M\right) + F_{FABA}\left(R_{rup}\right) + F_S(Vs30, PGA_{1000}) \qquad (1)$$

Where PSa is 5% pseudo-spectra acceleration, M is moment magnitude, $Vs30$ is the average shear velocity of the top 30 meters, and PGA_{1000} is the reference median PGA for $Vs30 = 1000\, m/s$. F_M , F_D , F_{FABA} , and F_S represent the magnitude scaling, distance scaling, forearc/backarc scaling, and site attenuation functions. The event term δC_1 , suggested by Stewart et al. (2013), is estimated from the residual analysis of the 2010 Mw 8.8 Maule and 2011 Mw 9.0 Tohoku earthquakes. In our analysis, we replace R_{rup} with R_{hf}. To focus on the effect of the distance metric, we also drop the event and site amplification functions. The GMPE is simplified as:

$$ln(PSa) = F_M(M) + F_D(R_{rup}, M) + F_{FABA}(R_{rup}) \quad (2)$$

To compare the accuracy of different distance metrics in GMPEs, we compute the residual R_i between the observed and model-predicted PSa.

$$R_i = \ln(PSa_i) - \ln(\widehat{PSa_i}) \quad (3)$$

Where PSa_i is the PSa measured from recording i, $\widehat{PSa_i}$ is the corresponding value predicted by the GMPE. This residual is later used for the correction of site attenuation, specifically in reference to the scaling of ground motions with $Vs30$ (Stewart et al., 2013). We then utilize the Root-Mean-Squared Error (RMSE) to represent the average logarithmic residuals between predictions with the data:

$$RMSE = \sqrt{\sum_{i=1}^{n} R_i^2/n} \quad (4)$$

Note that n is the total number of recordings.

4. Results

4.1. The 2011 Mw 9.0 Tohoku Earthquake

First, we select six representative stations in different regions across Japan (Fig. 3; S3) and study the PSa predictions (Fig. 2a). In all six stations, we find that R_{rup} underestimates the ground-motion intensity at the short periods (T < 1 s). On the other hand, the prediction based on R_{hf} fits the observation very well in a broad spectral range (0.1 < T < 10 s). Note that R_{hf} still underestimates the shaking at MYG0081, possibly due to the local site effect (Fig. 2a).

Next, we compare performances of two distance metrics R_{rup} and R_{hf} at all K-net stations, using BC 2016 (Fig. 2b), shown for two periods of T = 2 s and T = 0.5 s. For each case, we combine the forearc and backarc predictions to obtain a single RMSE. We find that the R_{hf}'s RMSE for both periods remains a small value of about 0.74, i.e. the predicted PSa is on average 2.15 times either larger or smaller than the observations. On the other hand, while the RMSE of R_{rup} at T = 2 s is reasonably small at 0.745, it increases dramatically to 1.38 at T = 0.25 s, which renders R_{rup} a poor distance metric for high-frequency. Our results indicate that R_{hf} is a more robust and accurate metric for this event. However, even though R_{hf} performs significantly better than R_{rup} in the near-field (R < 100 km), there is still bias in R_{hf}, overestimating at T= 0.25 s while underestimating at T= 2 s (Fig. 2b). We then add two more comparisons in the periods of T = 1 s and T = 0.5 s (Fig. S4, Fig. S5). It indicates that while RMSE of R_{hf} remains stable in both frequency bands, the RMSE of R_{rup} increases with frequency (Fig. S4). The RMSEs of R_{hf} and R_{rup} are both reasonably small at T = 2 s. We consider that the prediction by R_{rup} is invalid for T < 0.5 s for the Tohoku earthquake.

The advantage of R_{hf} can be also verified with the site attenuation analysis (Text. S1, Fig. S6, Fig. S7). We focus on the rate $c(T)$ at which site attenuation term is changing with respect to logarithmic value of $Vs30$. $c(T)$, in other words, is the slope between residual $R(T)$ and $\ln(Vs30)$ (Fig. S7). In the Tohoku earthquake, the variation of $c(T)$ (Stewart et al., 2013) with the period T is much larger than the previous study in active crustal regions (ACR; Choi & Stewart, 2005) when using R_{rup}. However, when using R_{hf}, $c(T)$ is much more consistent with that inferred from ACR data (Fig. S7). This agreement between the Tohoku earthquake case and ACR study indicates that R_{hf} is a better approximation of the distance term. A more detailed discussion of the methodology can be found in the supporting information (Barani et al., 2010; Borcherdt and Glassmoyer, 1994; Borcherdt 1994; Choi and Stewart, 2005; Stewart et at., 2013).

4.2. Analysis of All Five Megathrust Events

To further understand the performance of R_{hf} in other earthquakes, we analysis four additional megathrust earthquakes in Japan and Chile in the same manner as the Tohoku earthquake (Table. S1): the 2003 Mw 8.0 Tokachi earthquake (Koketsu et al., 2005; Nozu and Irikura, 2008), the 2005 Mw 7.2 off-Miyagi earthquake (Shao and Ji, 2005), 2014 Mw 8.1 Iquique earthquake (Wei, 2014), and the 2015 Mw 8.3 Illapel earthquake (Okuwaki et al., 2017) (Fig. S8-S15). Combining the results of all five earthquakes, we compare the RMSE of R_{rup} and R_{hf} as a function of frequency (Fig. 1b). For the three Japanese earthquakes (left panel), the mean RMSE of R_{hf} remains around 0.8 for all frequency bands, while that of R_{rup} increases rapidly with frequency, from 0.8 (T = 2 s) to 1.2 (T = 0.25 s). This can be explained by our previous analysis (see section 2.4): R_{rup} is close to R_{hf} at low frequency and the difference increases with frequency. For the two Chilean earthquakes (right panel), the mean RMSE of R_{rup} shows the opposite trend, decreasing rapidly with the frequency, from 1.3 (T = 2 s) to 0.7 (T = 0.25 s) with the corresponding explanation that R_{rup} is close to R_{hf} at high frequency and the difference decreases with frequency. Our results suggest that frequency-independent R_{rup} only work well in one single frequency and this frequency depends on the fault trimming process and the earthquake itself. Thus, R_{hf} outperforms R_{rup} despite the trend of the errors are different for Japanese and Chilean earthquakes within the overall HF band (0.5 - 2 Hz).

5. Discussion

Here, we demonstrate that modeling the locations of HF generating areas from BP gives a new distance metric, R_{hf}, that more accurately predicts strong ground motion at distance as compared to the traditionally used metric of R_{rup}. To understand the role of HFRs in generating high-frequency ground motions, we compare HFRs at the period of 0.25 s with the strong motion generation areas (SMGAs) and the coseismic slip of the 2011 Tohoku Event (Fig. 3) and Tokachi Event (Fig. S16). The SMGA is defined as the source areas that are responsible for the large amplitude of high-frequency near-field strong-motions (Miyake et al., 2003). SMGAs are typically interpreted as the source of large slip velocities or large stress release (~10 MPa). The locations of the SMGAs are estimated based on the onset times of the wave-packets at near-field stations (Kurahashi & Irikura, 2011). Another terminology, asperity, in the context of kinematic source models, is a region on the fault surface that produces large coseismic slip relative to the average slip. These asperities often locate at shallower depths in the dip-slip faults of crustal earthquakes (Somerville et al., 1999). SMGAs are often associated with the asperities in the corresponding source models by assuming that large slip velocity is often proportional to large static slip (Miyake et al., 2003). However, in the case of the Tohoku earthquake, the identified SMGAs are more consistent with the HFRs inferred from teleseismic BP rather than the large shallow asperities (Ye et al., 2013; Stewart et al., 2013). This agreement between SMGAs and HFRs confirms the common origin of HF bursts (~ 10 Hz) observed at teleseismic and local distances. In comparison, the region with large slip velocity is not necessarily the region of large slip (Fig. 3).

One possible scenario to produce large slip velocity with small slip is the short rise times. Average slip velocities equal the total slip divided by the rise time. When the rise time is small, the slip velocity of subfaults can be large even with small final slip. This hypothesis is supported by the observation that the HFRs are colocated with down-dip subfaults of short rise times (< 10 s) and moderate localized slip (~ 5 m) in Tohoku earthquake (Wei et al., 2012). The generation of these HFRs can be due to the presence of frictional heterogeneities at the seismogenic zone (Madariaga, 1982; Meng et al., 2011). The conditionally stable background is dotted with plenty of patchy, small-scale regions, which produce coherent short-period radiations when broken (Lay et al., 2012).

In future work, we will continue to test and improve the performances of R_{hf} in GMPEs and develop new attenuation relationship based on R_{hf}. The GMPEs using R_{hf} can be used to guide the interpolation of sparse data when producing the USGS ShakeMap. When a finite fault model is unavailable immediately after an earthquake, a quick, robust prediction of ground-motion intensity using R_{hf} can be developed based on near-real-time automatic BP. One can determine the HFRs directly from IRIS Data Management Center's automated BP results for global M > 6.5 earthquakes, based on coherent P-wave energy from regional arrays (North America, Europe, and Australia) and Global Seismographis Network (Trabant et al., 2012). The duration and magnitude of an earthquake can also be estimated from BP based on an empirical equation (Wang et al., 2017). Thus, we have all the necessary information to estimate ground-motion intensity. An important aspect of the GMPE is to estimate the ground motion intensity for the future earthquake scenarios, which requires the assumptions of source locations. For Probabilistic Seismic Hazard Analysis (PSHA) purposes, R_{hf} can't be simply estimated using BP as it is only available after the earthquake. These deep HFRs are the closest to the landward strong-motion stations and therefore dominate the determinations of R_{hf}. The mechanism of repeating earthquakes and HFRs are both interpreted as the repeated failures of small patches driven by either slow-slip or dynamic ruptures. Thus, these small patches in brittle-ductile transition could be the potential locations to generate coseismic high-frequency energy in megathrust earthquakes. The spatial consistency between HFRs during the 2011 Tohoku earthquake (Fig. 3) and the 2003 Tokachi earthquake (Fig. S16) and pre-seismic repeating earthquakes is an evidence for such hypothesis (Meng et al., 2011; Igarashi et al., 2001; Uchida et al., 2016; Lay et al., 2012). Next, we will explore which repeating earthquakes could be used as proxies for HFRs of potential future ruptures for PSHA.

6. Conclusion

Recent studies, exploring the frequency-dependent and depth-varying seismic radiation at plate interfaces from several major megathrust earthquakes motivate us to consider frequency-dependent source effect for GMPEs. The corresponding high-frequency distance metric R_{hf} is based on high-frequency radiators, which are quickly and robustly located by Back-Projection (BP) Imaging technique without any prior knowledge of rupture kinematics. We demonstrate that R_{hf} performs better in GMPEs over the traditional frequency-independent metric R_{rup} in several recent large megathrust earthquakes based on the analysis of the 2011 Tohoku earthquake, the 2003 Tokachi earthquake, the 2005 off-Miyagi earthquake, the 2014 Iquique eathquake, and the 2015 Illapel earthquake. R_{rup} shows significant bias in the prediction of 5% pseudo-spectral acceleration at high frequencies (0.5 – 4 Hz), which is close to the natural frequency of low-rise buildings. However, the distance metric R_{hf} based on BP shows little or no bias in ground-motion estimations. R_{hf} therefore potentially has the ability to improve the accuracy of GMPE predictions. GMPEs utilizing on R_{hf} can potentially provide real-time ground-motion intensity estimations for large earthquakes based on automated BP.

Acknowledgement

This research was supported by NSF Geophyiscs EAR-1723192 and by the Leon and Joann V.C. Knopoff Foundation. We appreciate Sean Ahdi and Jonathan Stewart for valuable discussions. The strong-motion from K-Net (http://www.kyoshin.bosai.go.jp/) and Centro Sismólogico Nacional (CSN) (http://www.sismologia.cl/) are used to validate the prediction of ground-motion intensity. The waveform from Incorporated Research Institutions for Seismology (IRIS) (http://ds.iris.edu/) are utilized for the Back-Projection Imaging analysis. The background topography and bathymetry used in our figures are provided by the National Centers for Environmental Information of NOAA (https://www.ngdc.noaa.gov/mgg/global/etopo1sources.html). The Vs30 data for K-Net stations are provided by David Boore in his personal website (http://www.daveboore.com/data_online.html). The authors are also grateful for Annemarie Baltay and one anonymous reviewer for their evaluations of the paper.

References

Abrahamson, N. A., Silva, W. J., & Kamai, R. (2014). Summary of the ASK14 ground motion relation for active crustal regions. *Earthquake Spectra*, *30*(3), 1025-1055.

Abrahamson, N., Gregor, N., & Addo, K. (2016). BC Hydro ground motion prediction equations for subduction earthquakes. *Earthquake Spectra*, *32*(1), 23-44.

Atkinson, G. M., & Boore, D. M. (2007). *Boore-Atkinson NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters*. Pacific Earthquake Engineering Research Center.

An, C., & Meng, L. (2016). Application of array backprojection to tsunami prediction and early warning. *Geophysical Research Letters*, *43*(8), 3677-3685.

Ancheta, T. D., Bozorgnia, Y., Darragh, R., Silva, W. J., Chiou, B., Stewart, J. P., ... & Idriss, I. M. (2012). PEER NGA-West2 database: A database of ground motions recorded in shallow crustal earthquakes in active tectonic regions. In *Proceedings, 15th World Conference on Earthquake Engineering*.

Anderson, J. G. (2007). Physical processes that control strong ground motion.

Barani, S., De Ferrari, R., Ferretti, G., & Spallarossa, D. (2010). Calibration of soil amplification factors for Real-time ground-motion scenarios in Italy.

Boore, D. M., Joyner, W. B., & Fumal, T. E. (1994). *Estimation of response spectra and peak accelerations from western North American earthquakes: an interim report* (No. USGS-OFR-93-509). US Department of the Interior, US Geological Survey.

Boore, D. M., Joyner, W. B., & Fumal, T. E. (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work. *Seismological research letters*, *68*(1), 128-153.

Boore, D. M., Stewart, J. P., Seyhan, E., & Atkinson, G. M. (2014). NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthquake Spectra*, *30*(3), 1057-1085.

Borcherdt, R. D. (1994). Estimates of site-dependent response spectra for design (methodology and justification). *Earthquake spectra*, *10*(4), 617-653.

Borcherdt, R. D., & Glassmoyer, G. (1994). INFLUENCES OF LOCAL GEOLOGY ON STRONG AND WEAK GROUND MOTIONS RECORDED IN THE SAN FRANCISCO BAY REGION. *US geological survey professional paper*, 77.

Bozorgnia, Y., Abrahamson, N. A., Atik, L. A., Ancheta, T. D., Atkinson, G. M., Baker, J. W., ... & Darragh, R. (2014). NGA-West2 research project. *Earthquake Spectra*, *30*(3), 973-987.

Chen, Y., Meng, L., Zhang, A., & Wen, L. (2018). Source complexity of the 2015 Mw 7.9 Bonin earthquake. *Geochemistry, Geophysics, Geosystems*.

Choi, Y., & Stewart, J. P. (2005). Nonlinear site amplification as function of 30 m shear wave velocity. *Earthquake Spectra*, *21*(1), 1-30.

Chopra, A. K. (2001). Dynamics of structures: Theory and applications.

Douglas, J. (2003). Earthquake ground motion estimation using strong-motion records: a review 524 of equations for the estimation of peak ground acceleration and response spectral 525 ordinates, *Earth-Science Reviews* 61, 43–104. 526

Fan, W., & Shearer, P. M. (2017). Investigation of backprojection uncertainties with M6 earthquakes. *Journal of Geophysical Research: Solid Earth*, *122*(10), 7966-7986.

Graves, R. W., & Pitarka, A. (2010). Broadband ground-motion simulation using a hybrid approach. *Bulletin of the Seismological Society of America*, *100*(5A), 2095-2123.

Gregor, N., N.A. Abrahamson, G.M. Atkinson, D.M. Boore, Y. Bozorgniz, K. Campbell, B. Chiou (2014). Comparison of NGA-West2 GMPEs, Eq. Spectra, 30(3), 1179-1197.

Hayes, G. P., Wald, D. J., & Johnson, R. L. (2012). Slab1. 0: A three-dimensional model of global subduction zone geometries. *Journal of Geophysical Research: Solid Earth*, *117*(B1).

Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2005). Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature*, *435*(7044), 933.

Joyner, W.B. and D.M. Boore (1981). Peak horizontal acceleration and velocity from strong 560 motion records including records from the 1979 Imperial Valley, California, earthquake, 561 *Bull. Seismol. Soc. Am.* 71, 2011–2038

Kiser, E., & Ishii, M. (2013). Hidden aftershocks of the 2011 Mw 9.0 Tohoku, Japan earthquake imaged with the backprojection method. *Journal of Geophysical Research: Solid Earth*, *118*(10), 5564-5576.

Kiser, E., & Ishii, M. (2017). Back-Projection Imaging of Earthquakes. *Annual Review of Earth and Planetary Sciences*, *45*, 271-299.

Koketsu, K., Hatayama, K., Furumura, T., Ikegami, Y., & Akiyama, S. (2005). Damaging long-period ground motions from the 2003 Mw 8.3 Tokachi-oki, Japan earthquake. *Seismological Research Letters*, *76*(1), 67-73.

Koper, K. D., Hutko, A. R., Lay, T., Ammon, C. J., & Kanamori, H. (2011). Frequency-dependent rupture process of the 2011 M w 9.0 Tohoku Earthquake: Comparison of short-period P wave backprojection images and broadband seismic rupture models. *Earth, planets and space*, *63*(7), 16.

Krüger, F., & Ohrnberger, M. (2005). Tracking the rupture of the M w= 9.3 Sumatra earthquake over 1,150 km at teleseismic distance. *Nature*, *435*(7044), 937.

Kurahashi, S., & Irikura, K. (2011). Source model for generating strong ground motions during the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, planets and space*, *63*(7), 11.

Lay, T., Kanamori, H., Ammon, C. J., Koper, K. D., Hutko, A. R., Ye, L., ... & Rushing, T. M. (2012). Depth-varying rupture properties of subduction zone megathrust faults. *Journal of Geophysical Research: Solid Earth*, *117*(B4).

Madariaga, R. (1983). High frequency radiation from dynamic earthquake. *Ann. Geophys.*, *1*, 17.

Meng, L., Inbal, A., & Ampuero, J. P. (2011). A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake. *Geophysical Research Letters*, *38*(7).

Meng, L., Ampuero, J. P., Luo, Y., Wu, W., & Ni, S. (2012). Mitigating artifacts in back-projection source imaging with implications for frequency-dependent properties of the Tohoku-Oki earthquake. *Earth, planets and space*, *64*(12), 5.

Meng, L., Allen, R. M., & Ampuero, J. P. (2014). Application of seismic array processing to earthquake early warning. *Bulletin of the Seismological Society of America*, *104*(5), 2553-2561.

Meng, L., Zhang, A., & Yagi, Y. (2016). Improving back projection imaging with a novel physics-based aftershock calibration approach: A case study of the 2015 Gorkha earthquake. *Geophysical Research Letters*, *43*(2), 628-636.

Miyake, H., Iwata, T., & Irikura, K. (2003). Source characterization for broadband ground-motion simulation: Kinematic heterogeneous source model and strong motion generation area. *Bulletin of the Seismological Society of America*, *93*(6), 2531-2545.

Nozu, A., & Irikura, K. (2008). Strong-motion generation areas of a great subduction-zone earthquake: waveform inversion with empirical Green's functions for the 2003 Tokachi-oki earthquake. *Bulletin of the Seismological Society of America, 98*(1), 180-197.

Ohno, S., Ohta, T., Ikeura, T., & Takemura, M. (1993). Revision of attenuation formula considering the effect of fault size to evaluate strong motion spectra in near field. *Tectonophysics*, *218*(1-3), 69-81.

Okuwaki, R., Yagi, Y., Aránguiz, R., González, J., & González, G. (2017). Rupture process during the 2015 Illapel, Chile earthquake: Zigzag-along-dip rupture episodes. In *The Chile-2015 (Illapel) Earthquake and Tsunami* (pp. 23-32). Birkhäuser, Cham.

Shao and Ji (UCSB, Honshu 2005).Preliminary Result of the Aug 16, 2005 Mw 7.19 honshu Earthquake,http://www.geol.ucsb.edu/faculty/ji/big_earthquakes/2005/08/smooth/honshu.html,last accessed August 22, 2013.

Si, H., & Midorikawa, S. (2000). New attenuation relations for peak ground acceleration and velocity considering effects of fault type and site condition. In *Proceedings of twelfth world conference on earthquake engineering*.

Sokolov, V. Y., & Chernov, Y. K. (1998). On the correlation of seismic intensity with Fourier amplitude spectra. *Earthquake spectra*, *14*(4), 679-694.

Somerville, P., Irikura, K., Graves, R., Sawada, S., Wald, D., Abrahamson, N., ... & Kowada, A. (1999). Characterizing crustal earthquake slip models for the prediction of strong ground motion. *Seismological Research Letters*, *70*(1), 59-80.

Stewart, J. P., Lanzo, G., Pagliaroli, A., Scasserra, G., Di Capua, G., Peppoloni, S., ... & Gregor, N. (2012). Ground motion recordings from the Mw 6.3 2009 L'Aquila earthquake in Italy and their engineering implications. *Earthquake Spectra*, *28*(1), 317-345.

Stewart, J. P., Midorikawa, S., Graves, R. W., Khodaverdi, K., Kishida, T., Miura, H., ... & Campbell, K. W. (2013). Implications of the Mw9. 0 Tohoku-Oki Earthquake for Ground Motion Scaling with Source, Path, and Site Parameters. *Earthquake Spectra*, *29*(s1), S1-S21.

Thompson, E. M., & Baltay, A. S. (2018). The Case for Mean Rupture Distance in Ground-Motion Estimation. *Bulletin of the Seismological Society of America*.

Trabant, C., Hutko, A. R., Bahavar, M., Karstens, R., Ahern, T., & Aster, R. (2012). Data products at the IRIS DMC: Stepping stones for research and other applications. *Seismological Research Letters*, *83*(5), 846-854.

Uchida, N., Iinuma, T., Nadeau, R. M., Bürgmann, R., & Hino, R. (2016). Periodic slow slip triggers megathrust zone earthquakes in northeastern Japan. *Science*, *351*(6272), 488-492.

Vallée, M., & Douet, V. (2016). A new database of source time functions (STFs) extracted from the SCARDEC method. *Physics of the Earth and Planetary Interiors*, *257*, 149-157.

Wei, S., Graves, R., Helmberger, D., Avouac, J. P., & Jiang, J. (2012). Sources of shaking and flooding during the Tohoku-Oki earthquake: A mixture of rupture styles. *Earth and Planetary Science Letters*, *333*, 91-100.

Wei, S. 2014. Apr./0.1/2014 (Mw 8.1), Iquique, Chile. Source Models of Large Earthquakes. http://www.tectonics.caltech.edu/slip_history/2014_chile/index.html, last accessed May 18, 2014.

Wang, D., Kawakatsu, H., Zhuang, J., Mori, J., Maeda, T., Tsuruoka, H., & Zhao, X. (2017). Automated determination of magnitude and source length of large earthquakes using backprojection and P wave amplitudes. *Geophysical Research Letters*, *44*(11), 5447-5456.

Yao, H., Shearer, P. M., & Gerstoft, P. (2013). Compressive sensing of frequency-dependent seismic radiation from subduction zone megathrust ruptures. *Proceedings of the National Academy of Sciences*, *110*(12), 4512-4517.

Yao, H., Gerstoft, P., Shearer, P. M., & Mecklenbräuker, C. (2011). Compressive sensing of the Tohoku-Oki Mw 9.0 earthquake: Frequency-dependent rupture modes. *Geophysical Research Letters*, *38*(20).

Ye, L., Lay, T., & Kanamori, H. (2013). Ground shaking and seismic source spectra for large earthquakes around the megathrust fault offshore of northeastern Honshu, Japan. *Bulletin of the Seismological Society of America*, *103*(2B), 1221-1241.

Zhao, J. X., Zhang, J., Asano, A., Ohno, Y., Oouchi, T., Takahashi, T., ... & Fukushima, Y. (2006). Attenuation relations of strong ground motion in Japan using site classification based on predominant period. *Bulletin of the Seismological Society of America*, *96*(3), 898-913.

Figures and Captions

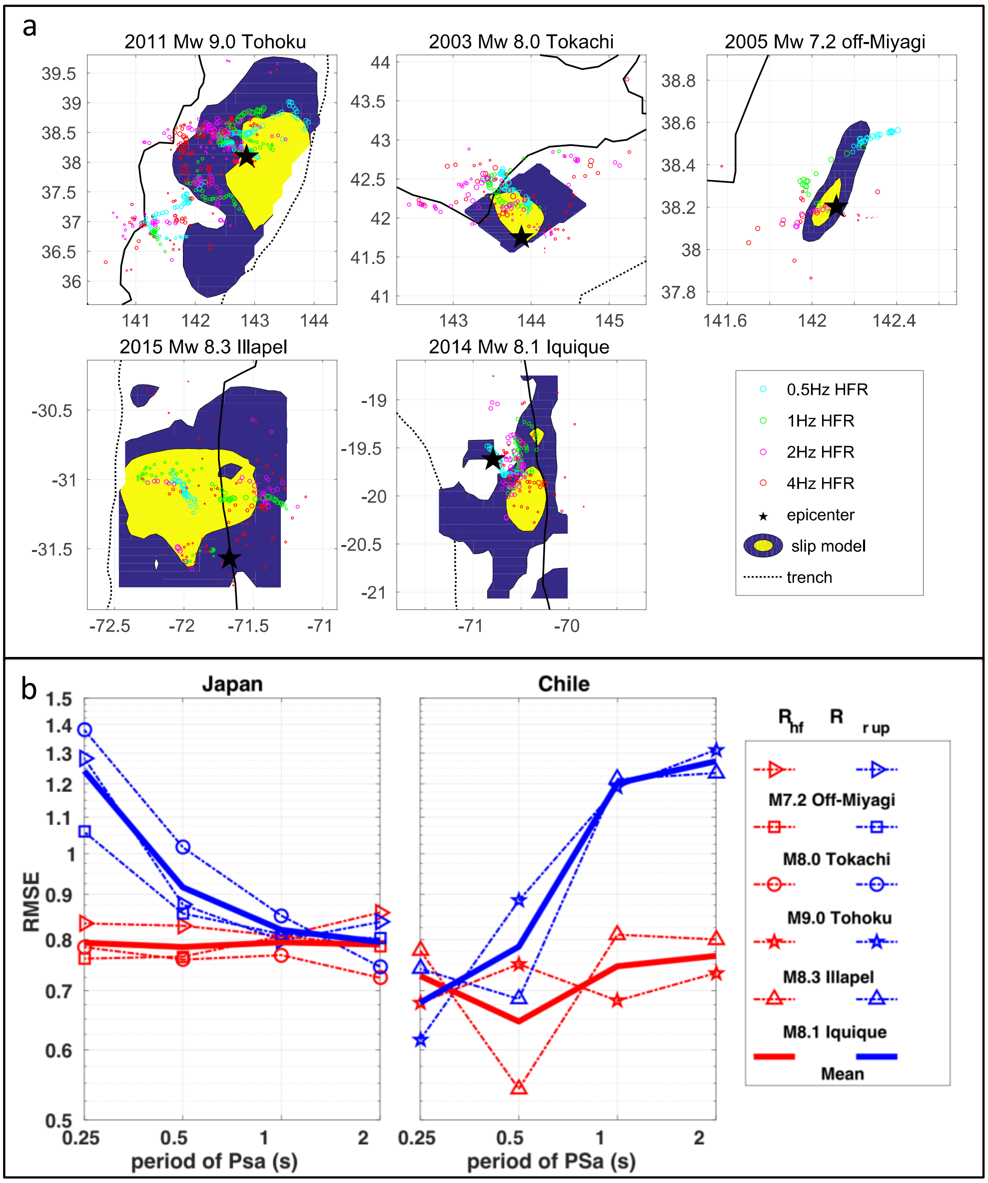

Figure 1. (a) Back-Projection results of five large earthquakes in Japan and Chile: the 2011 Mw 9.0 Tohoku earthquake, the 2003 Mw 8.0 Tokachi earthquake, the 2005 Mw 7.2 off-Miyagi earthquake, the 2014 Mw 8.1 Iquique earthquake, and the 2015 Mw 8.3 Illapel earthquake. Black star marks the hypocenters of the five earthquakes. Color open circles represent HFRs in four frequency bands. The yellow and blue patches represent areas with coseimic slip larger than 50% and 10% of peak slip respectively. (b) Comparison of the performance between R_{rup} and R_{hf} in five earthquakes. Blue dashed lines represent R_{rup}, while red lines represent R_{hf}. The R_{rup}'s mean RMSE value of three Japanese events (solid blue line) decreases with period while R_{hf}'s mean RMSE (solid red line) changes little (left panel). The R_{rup}'s mean RMSE value of three Chilean events increases with period while R_{hf}'s mean RMSE remains small (right panel).

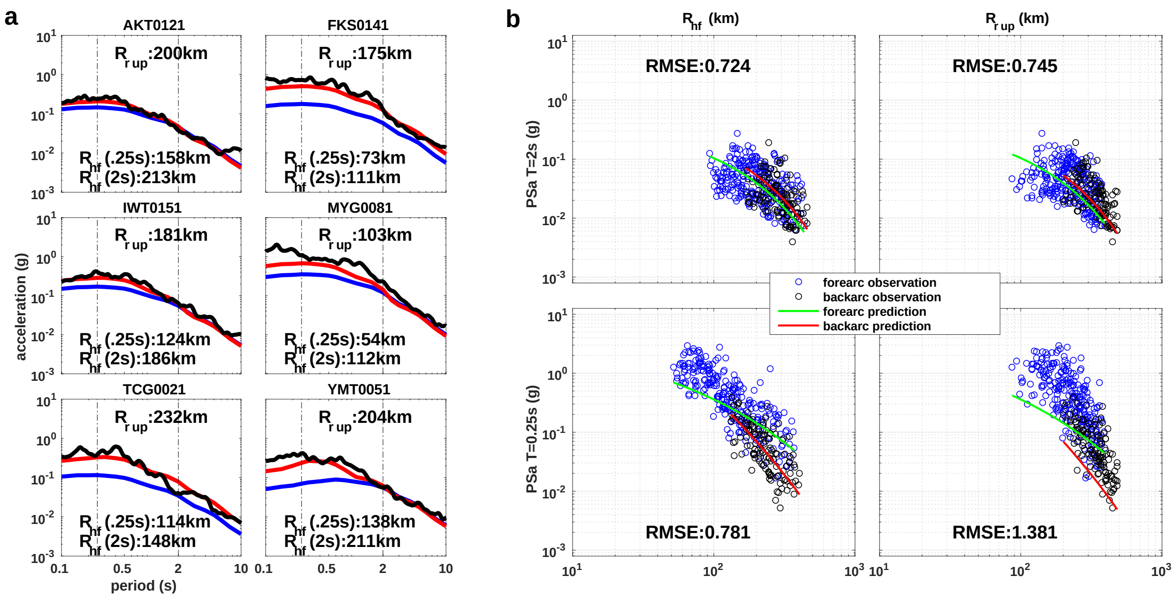

Figure 2. (a) Comparisons of PSa (g) predicted with distance metrics R_{hf} and R_{rup} of 2011 Tohoku earthquake (Mw 9.0) at 6 selected K-Net stations. The black, blue, and red lines denote the observation, the prediction based on R_{rup} and the prediction based on R_{hf} respectively. (b) Comparisons of PSa in different frequencies using R_{hf} and R_{rup} for the 2011 Tohoku earthquake. The x-axis marks distance in km, and the y-axis denotes PSa in gravitational acceleration g. The green and red traces denote the forearc and backarc PSa values predicted with GMPE BC2016. The blue and black open circles are the observed PSa values at forearc and backarc K-net stations, respectively.

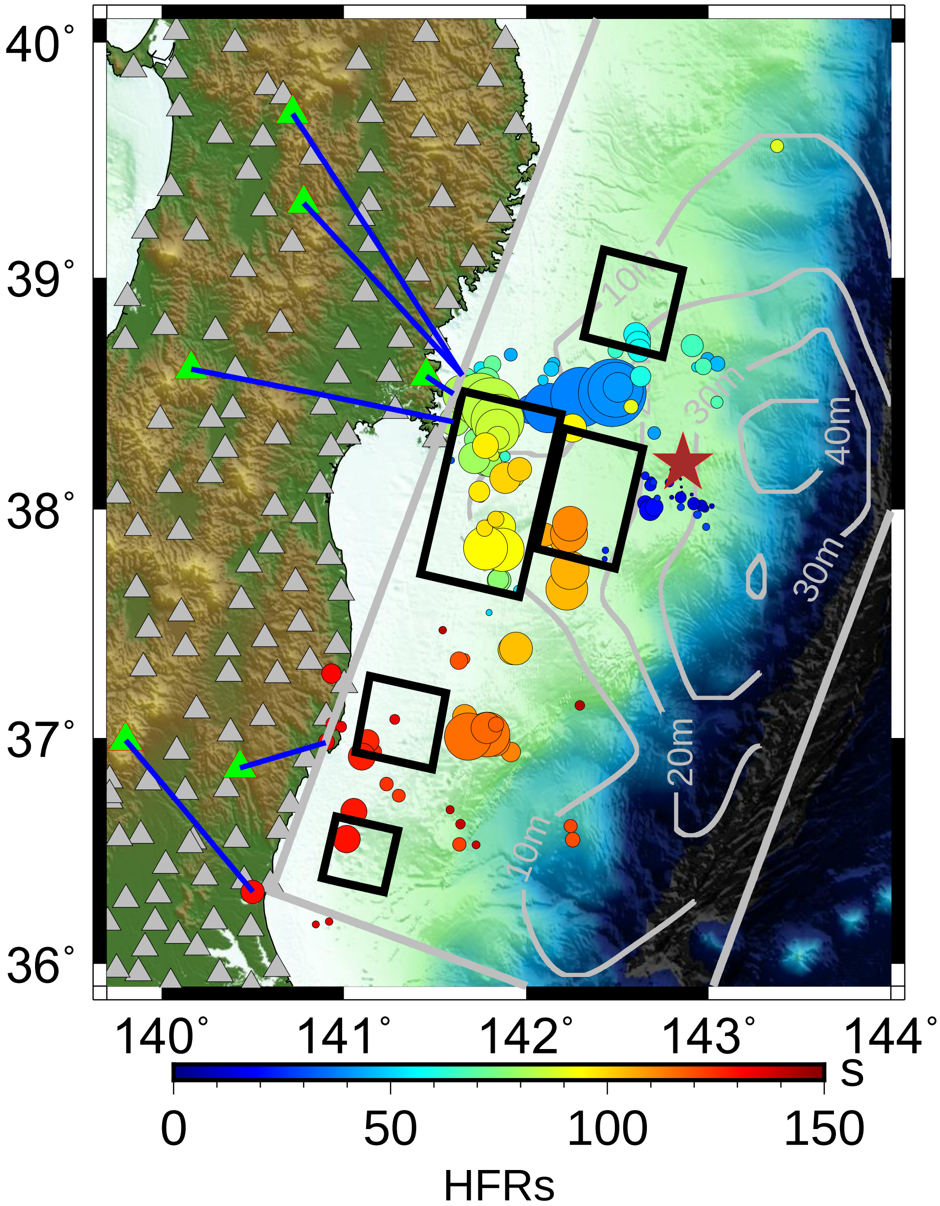

Figure 3. Finite fault model (Wei et al., 2012) of 2011 Mw 9.0 Tohoku earthquake (gray rectangle), strong motion generation areas from Kurahashi and Irikura (2011; black rectangles), the 4 Hz high-frequency radiator locations (color spheres), and deep pre-seismic repeating earthquakes (Uchida et al., 2016) (purple open squares). Contours indicate coseismic slip (m) and the red star is the JMA epicenter of the main shock. Blue lines denote the surface projection of R_{hf} at six selected stations (green triangles).