Assessing the impact of regional geology on the ground motion model variability at the Kashiwazaki-Kariwa Nuclear Power Plant (Japan) via physics-based numerical simulation

David Castro-Cruz, Filippo Gatti, Fernando Lopez-caballero

Version of Record: https://www.sciencedirect.com/science/article/pii/S0267726121003699
Manuscript_4fc5dbeba9fa0fe211aec382866046b8

Assessing the impact of regional geology on the ground motion model variability at the Kashiwazaki-Kariwa Nuclear Power Plant (Japan) via physics-based numerical simulation.

David Castro-Cruz[a], Filippo Gatti[a] and Fernando Lopez-Caballero[a]

[a]*Université Paris-Saclay- CNRS - CentraleSupélec*
Laboratoire de Mécanique des Sols, Structures et Matériaux (MSSMat) UMR CNRS 8579
8-10 rue Joliot Curie 91190, Gif-sur-Yvette, France

ARTICLE INFO

Keywords:
Physics-Based earthquake Simulation
2007 M_W6.6 Niigata earthquake
Kashiwazaki-Kariwa Nuclear Power Plant
Site-effects
Ground Motion Prediction Equations

ABSTRACT

This study investigates the sensitivity of physics-based earthquake prediction to the choice of the regional geological model. The generally poor information on the mechanical properties of the Earth's crust induces large uncertainty margins on the synthetic seismic response. Therefore, uncertainty quantification must be associated with high-fidelity numerical simulation, especially when verging on broad-band simulations (i.e., numerical theoretical accuracy higher than the conventional 1 Hz limit). In this paper, the synthetic seismic response of the Niigata region (Japan), during the Chūetsu offshore earthquake (M_W6.6) serves as representative case study. Three plausible three-dimensional (3-D) geological structures are compared for two point-wise aftershock numerical simulations, duly validated. The framework of comparison targets the Ground Motion Prediction Equation (GMPE) for pseudo-spectral acceleration at the surface. The impact of the geology on the variability of between-event and the within-event residuals is inferred from synthetic simulations and related to common parameters representing site-specific conditions.

1. Introduction

With the increasing broader access to high-quality seismic databases, empirical relationships to predict ground motion intensity flourished, cast into Ground Motion Prediction Equations (GMPEs, whose omni-comprehensive compendium can be found at `http://www.gmpe.org.uk/`, see Douglas, 2021), of the general form:

$$\log\left(Y_{es}\right) = f\left(\{X\}_i;\{\theta\}_j\right) + \Delta_{es} \tag{1}$$

where Y_{es} is the generic intensity measure (IM) for earthquake e at station s. The pseudo-spectral acceleration Sa (approximating the maximum absolute acceleration of the single degree of freedom system with 5% damping), the Peak Ground Acceleration (PGA) and the Peak Ground Velocity (PGV) are among the most commonly used IMs. The functional form f embodies the geometric mean IM prediction (log value). $\{X\}_i$ embraces the source characteristics (usually earthquake magnitude, type of earthquake, focal mechanism), the notion of source-to-site distance (hypocentral, epicentral and Joyner-Boore among others), the site parameters (usually the site class based on the average shear-wave velocity of the first 30 m below the ground surface, noted as $V_{S,30}$). $\{\theta\}_j$ is the set of coefficients obtained via non-linear regression on large seismic databases.
The residual Δ_{es} represents the shift between the geometric mean prediction f and observations. Δ_{es} is usually modeled as a centered normal random variable with distribution $\mathcal{N}(0,\sigma)$ (Atik et al., 2010; Rodriguez-Marek et al., 2013). Moreover, Δ_{es} is modeled as the sum of the *between-event variability* δB_e (the earthquake-wise station-average deviation of the observed ground motion from the geometric mean f) and of the *within-event variability* δW_e (the station-wise misfit between the e observed earthquake ground motion and the sum $f+\delta B_e$, Rodriguez-Marek et al., 2013).

$$\Delta_{es} = \delta B_e + \delta W_{es} \tag{2}$$

*Corresponding author
ORCID(s): 0000-0002-2404-8653 (D. Castro-Cruz); 0000-0001-7174-4048 (F. Gatti); 0000-0002-1010-1230 (F. Lopez-Caballero)

δB_e and δW_{es} are both modeled as centered normal random variable with distributions $\mathcal{N}(0, \tau)$ and $\mathcal{N}(0, \phi)$ respectively.

GMPEs usually benefit from the ever-increasing access to seismic data and meta-data. More complex functional f can be calibrated by adding new parameters θ_j to the recipe in order to take into account site-specific and source-specific features. However, current seismic hazard analysis struggles in reducing the uncertainty Δ_{es} and the associated standard deviation σ_t (especially for earthquakes with high return periods, Strasser et al., 2009). Some authors achieved this goal by removing the so-called *ergodicity assumption* (see, for instance, Atik et al., 2010), i.e. including site- and scenario-dependent parameters. However, this approach suffers from (1) a lack of observed records in some regions to sufficiently constrain the regression process and (2) poor generalization, i.e., the impossibility of extrapolating reliable predictions in general contexts (Abrahamson and Hollenback, 2012; Chen and Tsai, 2002). Those drawbacks hinder the establishment of high-fidelity site-/scenario-specific GMPEs. A very appealing alternative is provided by high-fidelity earthquake simulations, duly validated against past recorded earthquakes (Bradley, 2018). This approach sacrifices the GMPE handy regression framework to explicitly tackle the modeling uncertainties at stake, namely: (1) the definition of a reasonable and realistic rupture scenario on active faults is required, (2) the 3-D crustal structure, (3) the surficial site conditions (related to the so-called *site effects*). For those reasons, waveform inversion and array back-projection are employed to detect and characterize fault asperities that generated the strong ground motion (Aoi et al., 2008; Honda and Aoi, 2009). GMPEs and numerical simulations are being used together for Probabilistic Seismic Hazard Assessment (PSHA). For instance, Milner et al. (2021) managed by updating non-ergodic PSHA for Southern California via direct numerical simulation.

The identification of the geology in the region of interest (bedrock and softer sediments) is either achieved by interpretation and analysis of borehole logs in the surroundings (Fantoni and Franciosi, 2010), or by seismic tomography (Lee et al., 2014; Jian and Fanhua, 2009). Finally, laboratory tests are necessary when including non-linear soil rheology (see the PRENOLIN international benchmark on 1-D numerical wave-propagation in Régnier et al., 2016). Despite the computational performance achieved by the numerical tools, the mentioned uncertainties prevent the complete characterization of broad-band ground shaking scenarios, especially at short periods (i.e. $T < 0.5$ s, Paolucci et al., 2018). In recent years, physics-based numerical simulation improved by merging its fidelity at long periods (usually for natural periods $T > 0.2$ s) with data-driven ground motion models, conditioned by the outcome of the numerical analysis (Gatti and Clouteau, 2020; Castro-Cruz et al., 2021; Jayalakshmi et al., 2021). Several comparisons (Olsen et al., 2000; Smerzini et al., 2011) showed large differences between 1-D 2-D and 3-D simulations, establishing that 3-D simulations can better capture the source footprint and the impact of the geological structures. However, the power of 3-D models is sometimes limited by the only available 1-D velocity profiles around specific areas.

In this work, high-fidelity 3-D numerical simulations are exploited to clarify their sensitivity to the geological model considered in a relatively low-frequency range (0-3 Hz). The uncertainty quantification is inferred from the reproduced variability of the ground motions for three different yet plausible geological models. The analysis follows the context of GMPE predictions, given the large number of synthetic time-histories extracted from the numerical simulation at several geographical locations in the region of interest. This work focuses on a real case study, the Niigata prefecture (central-west Japan), which hosts the largest nuclear power plant in the world, located in Kashiwazaki-Kariwa (KKNPP, see Figure 1).
This region has been struck by several significant earthquakes, with the remarkable M_W 6.6 Niigata Chuetsu-Oki earthquake of 2007 (NCOEQ2007), which did cause the KKNPP shut-down. Gatti et al. (2018a) showed the importance of considering the complex 3-D shallow underground structure to approach the recorded seismic response, compared to traditional layered geology. They confirmed an improved fit of the recorded time-histories at KKNPP, when the complex *folded* geology below KKNPP was included in their numerical model. Gatti et al. (2018c) employed the technique named `ANN2BB` (Paolucci et al., 2018) to enrich this synthetic wave-motion at high-frequency and successfully perform a Soil-Structure Interaction (SSI) study of the Unit 7 reactor building at KKNPP. These studies highlighted the difficulty of constructing and calibrating a regional digital twin, including the uncertainty of the available geological information (usually 2-D geological maps) and the uncertainty related to the numerical discretization.

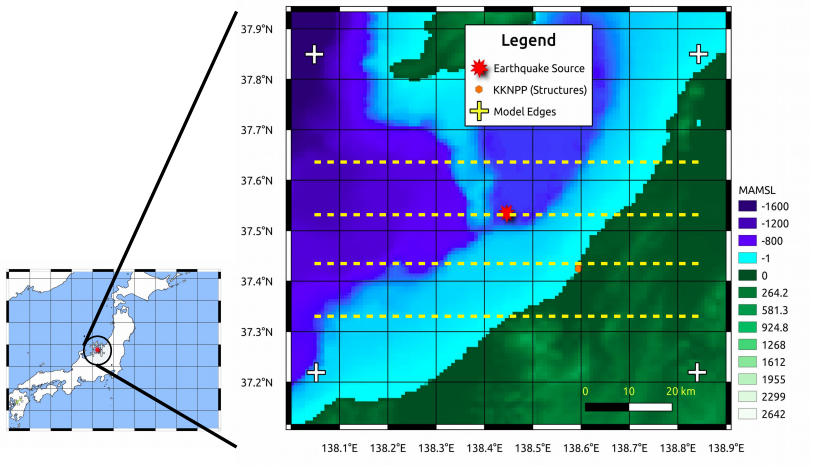

Figure 1: Digital Elevation Model of the region considered in the analysis, including the Niigata basin and the surroundings of the Kashiwazaki-Kariwa Nuclear Power Plant (KKNPP). Yellow dashed horizontal lines represent the traces of the vertical clips shown in Figure 4.

2. Data

3-D numerical models require detailed information on the site geological configuration and source characteristics, in order to accurately reproduce the complex 3-D wave field. Ideally, high-quality information is required at three different scales: the regional scale (strictly connected to the long-period part of the rendered ground motion), the site-scale (at which most of the non-linear multi-path effects take place), and the borehole scale (necessary when investigating complex SSI effects). Nevertheless, these ideal conditions are rarely met, given the tremendous effort needed to collect seismological observations, to perform geophysical and geological campaigns, to invert and to exploit those data. More realistically, 1-D geological profiles are provided for the seismological region (characteristic dimension $\approx$ 1 km). 2-D or 3-D geological structures are also available for some sites of interest, such as at the KKNPP (Watanabe et al., 2009).

2.1. Available 1-D velocity structures

Figure 2 compares the available 1-D vertical profiles proposed in the literature for the region surrounding KKNPP. Specifically, `DA&2010-1` and `DA&2010-2` were proposed by Ducellier and Aochi (2010), in their numerical simulations of the NCOEQ2007 mainshock. `Aochi2013` was proposed by Aochi et al. (2013), for their finite difference analysis of the NCOEQ2007 mainshock and aftershock sequence. `Cirella2008` velocity values were employed by Cirella et al. (2008) to perform a waveform inversion of NCOEQ2007 source mechanism. Finally, the profiles at stations `NIG004`, `NIG016`, `NIG026` belong to K-NET network and `NIGH12` to KiK-Net network (NIED, 2019; Aoi et al., 2004). Those crustal models are reliable at a very low-frequency range (f< 0.5 Hz).

2.2. Other geological models

Figure 3 displays the cross-section SC of a local *folded* geology striking at N 145° E and located underneath KKNPP. Watanabe et al. (2009) employed it in a 2-D finite element earthquake analysis. The folded structure was constructed from a boring and seismic reflection survey (Kobayashi et al., 1995). The SC cross-section spans approximately a region 7.6 km wide and 4.8 km. Seven strata are reported, including the seismic bedrock (see Table 1).
The simulation results rendered by Watanabe et al. (2009) showed good agreement with the observed strong motion records at Unit 1, explaining the strong ground motion spatial variability within the nuclear facility (due precisely to the first and second boundaries). Those findings are confirmed by Tokumitsu et al. (2009). Tsuda et al. (2011) clarified the role of this folded structure by performing a 3-D finite difference earthquake simulation of the NCOEQ2007 mainshock and aftershock sequence. The 3-D velocity structure they proposed is the combination of a *broad model*, proposed by Japan Nuclear Energy Safety Organization (JNES, see Kamae, 2016). Although accurate for the regional wave field, JNES did not include the folding structure. Therefore, Tsuda et al. (2011) included a 3-D version of the folding

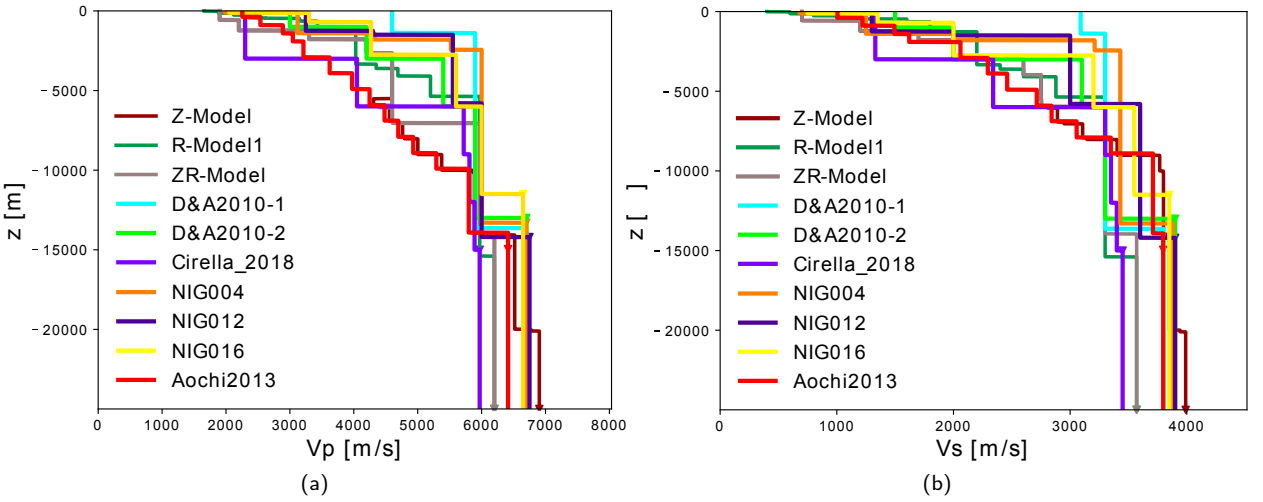

Figure 2: 1-D velocity models of the different boreholes from the region around KKNPP. The triangles are located at the maximum borehole depths. a) Compressional wave speed V_P profile; b) Shear wave speed V_S profile.

Table 1
Geological properties of the folding structure underneath KKNPP. V_P and V_S are the compressional- and shear-wave velocities, respectively. The ** indicates the interface chosen to plug the folding structure into the original 1-D *Aochi2013* profile, granting a smooth transition from one model to another.

Layer	V_S [m/s]	V_P [m/s]	ρ [kg/m^3]
Nishiyama	700	1900	1700
Shiiya	1200	2200	2100
Upper Teradomari	1700	3300	2300
Lower Teradomari	2000	4200	2400
Nanaya	2000	4600	2500
**Green tuff	2600	5200	2600
Seismic bedrock	2600->2750	5200	2600

structure, based on the previous studies of Tokumitsu et al. (2009), which consists of 7 cross-sections. This *local folding model* was interpolated to reconnect to the external *broad model*. Using the horizontal-vertical (H/V) ratio technique (Nakamura, 1989), Tsuda et al. (2011) matched the synthetic H/V with those of the observed records for some sites on the KKNPP, obtaining a good agreement of the predominant frequencies around 0.3 Hz. The numerical simulation for the NCOEQ2007 showed good agreement with the recorded time-histories at KKNPP, in the range 0-4 Hz.

Finally, Gatti et al. (2018a) validated a similar 3-D geological model by comparing the synthetic records with the recorded shakings from two aftershocks (M_{JMA}4.4 and M_{JMA}4.2) of the NCOEQ2007 (NIED DMC; Obara et al., 2005; Okada et al., 2004), confirming that a simple layered 1-D model (`Aochi2013` proposed by Aochi et al. (2013)) cannot reproduce the ground motion amplification accurately at Unit 1 and the focalization effect due to the folding structure.

2.3. Proposed 3-D geological models

In this work, three geological models are considered, obtained from the interpolation of the available 1-D and 2-D geological information described in the previous subsection, integrated with the geological model of reference, provided by the Geological Survey of Japan for the Niigata area (Sekiguchi et al., 2009). Available information was alternatively combined in each model, aiming at reproducing the regional and site path effects. Namely, the three models are referred to as:

1. **Z-Model**: this virtual geology corresponds to the model proposed by Gatti et al. (2018a) and Gatti et al. (2018c)

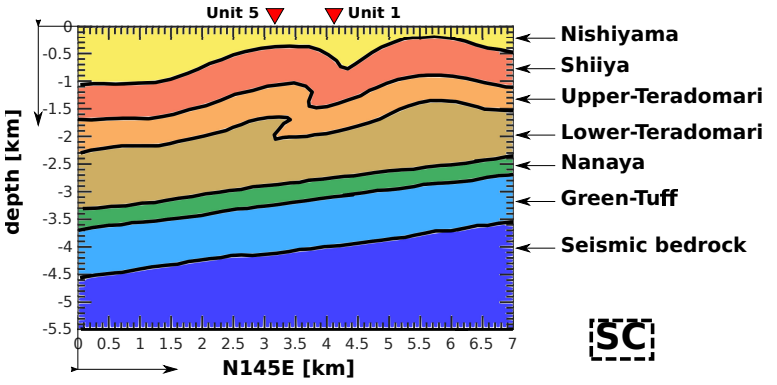

Figure 3: Cross-section (SC) of the folded geology below KKNPP (Unit 1 and Unit 5) proposed by Watanabe et al. (2009); Tsuda et al. (2011). Layer names are reported in Table 1. Reprinted from Gatti et al. (2018a).

to predict the structural response of the KKNPP Unit 7 reactor building. It blends two geological structures: a regional sub-horizontally layered geology inspired by Aochi et al. (2013) (labeled as `Aochi2013` in Figure 2) and a model obtained from the extrusion of SC cross-section in Figure 3. The two geological structures are joined by linear interpolation. The `Aochi2013` model was calibrated to reproduce the regional wave field in the 0-0.5 Hz frequency range. The hybrid model, including the folding, was instead verified in the 0-5 Hz frequency range, for two NCOEQ2007 aftershocks, at several locations within the KKNPP site (the borehole arrays deployed within the nuclear site) and at several KiK-Net and K-NET stations.

2. **R-Model**: this model was proposed by the Geological Survey of Japan (GSJ, Sekiguchi et al., 2009). It is an improvement of the previous NIED model (National Research Institute for Earth Science and Disaster Prevention) proposed by Fujiwara et al. (2006) for the Niigata area, based on several seismological observations. The depth of layer boundaries is provided on a grid mesh of approximately 0.5×0.5 km. The model contains 50 layers above the Moho (Mohorovicic discontinuity). Aochi et al. (2013) tested and compared this geological structure with two other 3-D models of the Niigata basin and reproduced the NCOEQ2007 mainshock time-histories at several KiK-Net (NIGH11) and K-NET stations (NIG019 and NIG017) in the frequency range 0.1-0.5 Hz. The authors specified that the synthetics at near-field soil sites were quite realistic, but the model needed improvements and a finer shallow structure.
3. **ZR-Model**: this synthetic 3-D configuration is a combination of the previous two, i.e., the folding structure (Z-model) is included in the regional structure of the Niigata basin (corresponding to the R-model). This hybrid geological structure shares the deepest layers (> 6 km) with the R-Model and the surficial ones with the Z-Model. The main reason why this geology was considered to coincide with the purposes of Tsuda et al. (2011) to include the prediction of both the regional wave field (interacting with the Niigata basin-like geology) and the influence of the local geology on it.

The left and right columns in Figure 4 show four West-East cross-sections of the Z-Model and the R-Model, respectively. In Appendix A, the detailed 3-D cuts of the three geological structures in the Niigata region are reported. It is important to note that the Z-model is relatively regular for depths below 5 km (layered velocity structure proposed by Aochi et al., 2013), whereas the R-model shows 3-D sharp basin-like structures reaching a depth of 10 km.

3. Physics-based numerical simulation

3.1. 3-D Numerical Model

To assess the uncertainty related to the geological and geophysical models proposed in the previous section, the constructed numerical model covers 64 km $\times$ 64 km around the earthquake position (see Figure 1). The model characteristics are summarized in Table 2. All three models exploit the very same hexahedral mesh, neglecting the regional topography. Since different geologies characterize the three models, their accuracy varies from model to model. However, all the models are accurate enough to study the influence of the geology for periods $T > 0.5s$.

Earthquake simulations were performed using `SEM3D` (CEA and CentraleSupélec and IPGP and CNRS, 2017), a high-

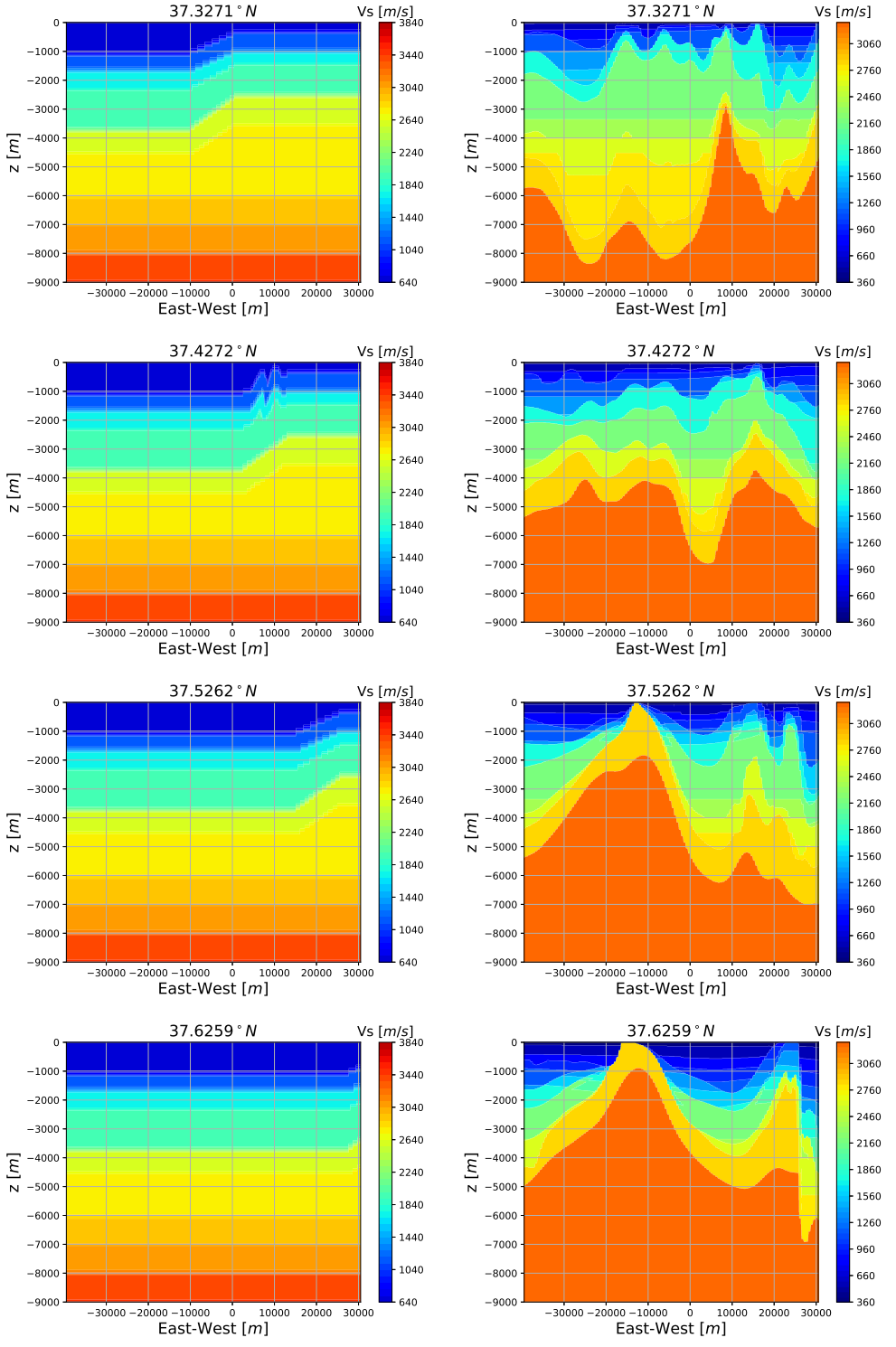

Figure 4: West-East cross-sections at different latitudes (reference system: Tokyo CS VIII EPSG:30168 system), showing the V_S contours. Figure 1 shows the position of the cross-sections in the map. Left column belong to the Z-Model and right column to the R-Model.

Table 2
Parameters of the template numerical model employed in this paper. n_{el} and n_{DOF} represent the number of hexahedral elements and degrees of freedom (DOF) respectively. ΔL_{min} is the minimal element characteristic size, N_{GLL} is the number of Gauss-Lobatto-Legendre integration points per element size, f_{max} is the maximum frequency that the model can theoretically propagate.

Model	n_{el} [1]	n_{DOF} [1]	ΔL_{min} [m]	N_{GLL} [1]	Vs_{min} [m/s]	f_{max} [Hz]
Z	$\approx 4.02 \cdot 10^6$	$\approx 1.50 \cdot 10^9$	139	5 ×5×5	700	5
R/ZR	-	-	-	-	400	2.1

performance software, implementing the spectral element method (SEM, Komatitsch and Vilotte, 1998; Faccioli et al., 1997). This software is based on `RegSEM` code (Cupillard et al., 2012; Festa and Vilotte, 2005). The SEM represents a high-order version of the finite element method (FEM), and it is well known for the relatively easy parallelization algorithm associated (domain decomposition over distributed memory computing cores, using MPI protocol Göddeke et al., 2014). Those aspects make the 3-D simulation an attractive tool for seismic hazard analysis (Olsen et al., 1995; Bradley, 2018). The mesh has higher refinement at shallow layers to preserve the numerical accuracy where slowness values are larger. The maximum theoretical frequency propagated can be estimated as:

$$f_{max} = \frac{V_{S,min} \cdot N_{GLL}}{5 \cdot \Delta L} \tag{3}$$

with $V_{S,min}$ being the minimum shear wave velocity in the domain, ΔL the corresponding element characteristic size and N_{GLL} the number of Gauss-Lobatto-Legendre integration points featuring the mesh element (for the SEM, N_{GLL} is usually greater than 5).

Given the intricate geologies at stake and due to the adopted integration rule over 3-D hexahedra, a *not-honoring* approach is adopted to associate the spatially distributed geological properties to the mesh point locations (Casarotti et al., 2008). In other words, the geological discontinuities are not directly meshed, but rather *interpolated* over the GLL grid. This approach has the advantage of simplifying the inclusion of surfaces of complex geometry, yet inducing some spurious amplification due to spatial interpolation. The influence on the earthquake numerical predictions presented in this study was duly checked and it can be neglected, due to the rather low-frequency range considered. Finally, Gatti et al. (2018a) verified the limited influence of surface topography on the seismic response at KKNPP, which was not therefore included. Appendix B provides some insights on the computational resources employed for each simulation.

As far as the source is concerned, in order to assess the uncertainty related to the geological model, two earthquake point sources were adopted: (1) EQ1, corresponding to the M_{JMA}4.4 (Japan Meteorological Agency) aftershock of the NCOEQ2007 sequence (employed by other authors (Gatti et al., 2018a)) and (2) EQ2, corresponding to the NCOEQ2007 mainshock hypocenter location. Table 3 summarizes the focal mechanism and location of the two sources. These parameters represent the small aftershocks employed to validate the numerical model by Gatti et al. (2018a) against the observed recordings. Statistical analysis on the source's uncertainty, such as by Shang and TkalÄĐiÄĞ (2020), are out of the scope of this paper. For the sake of simplicity, the mainshock source (EQ2) was

Table 3
Summary of the aftershock parameters employed in this analysis. $(\phi_S; \lambda; \delta)$ represent the strike, rake and dip angles, respectively, estimated by F-NET Centroid Moment Tensor solution (Kubo et al., 2002). τ_R represents the source rise-time.

Event	M_{JMA}	M_0[Nm]	$(\phi_S; \lambda; \delta)$ [°]	d [km]	τ_R [s]
EQ1 (07/16/07-21:08)	4.4	$5.21 \cdot 10^{15}$	187; 70; 54	11	0.113
EQ2 (07/16/07-01:13)	6.8	$7.72 \cdot 10^{18}$	30; 78; 38	17	8.015

considered as a point source to be coherent with the idea of this paper of studying the effect of epistemic uncertainty on geological and geophysical information. The full moment tensor components were provided by F-NET. More accurate analysis requires an extended source model, whose rupture path highly affected the mainshock pulse-like seismic response at the site (near-source effects shown by Aoi et al., 2008; Gatti et al., 2018a). However, the following sections

show the interest in considering a mainshock source due to its Source Time Function (STF) featured by large rise-time τ_R compared to EQ1 (see Table 3).

3.2. Synthetic seismic response of the Niigata region

Before performing statistical analyses on the post-processed results (see Section 4), a validation task is presented hereafter. In doing so, synthetic time-histories are compared to recorded signals for EQ1, an aftershock widely recorded by the down hole arrays and surface seismometers located within KKNPP (see Figure 5). The site is instrumented with two recording networks, including four borehole arrays (Gatti et al., 2017), with KSH in Figure 5 representing a free-field control point for this analysis (Gatti et al., 2018a).

Given the accuracy of the numerical models, all the synthetic and recorded time-histories were band-pass filtered in

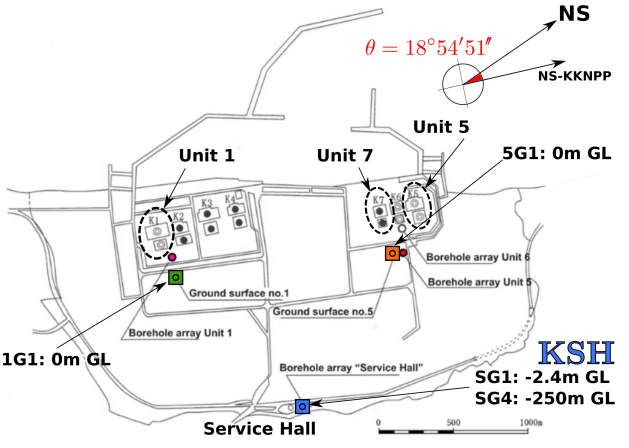

Figure 5: Map of the KKNPP site. The locations of downhole arrays of seismometers are reported: KSH represents the Service Hall array (four devices, SG1-SG4, reaching depth of -250 m G.L.), which entirely recorded the NCOEQ2007.

the 0.1-2.8 Hz frequency band (4^{th} order Butterworth filter). Moreover, to quantify the predictive capabilities of the numerical model, the criterion proposed by Kristekova et al. (2009) was adopted, which proposes a time-frequency metrics in order to quantify the Goodness of Fit (GoF) between synthetic and recorded signals. Kristekova's criterion defines two indices that quantify the misfit in time and frequency domains: Envelope Goodness (EG) and Phase Goodness (PG). Figure 6 summarizes the EG and PG GoF values obtained for three EQ1 runs at the KKNPP, each one featuring a different geological model among the investigated ones (R-, Z-, ZR-model respectively). The comparison at other stations outside the KKNPP is out of the scope of this paper, being those stations located too far away from the nuclear site. The synthetic and recorded time-histories at KSH, employed for the validation exercise, are reported in Appendix C, along with the time-frequency GoF values (see Figure 21). Overall, the GoFs are satisfactory in both Envelope (≈ 5) and Phase (≈ 6) for the three virtual geologies, with slight differences among them. The Z-model, as already verified by Gatti et al. (2018b), appears as the most adapted to predict the KKNPP seismic response, being primarily influenced by the layered folding striking underneath KKNPP. However, Figure 6 proves that the three models are all equally plausible in this relatively low-frequency range, thus suggesting the need for an accurate uncertainty quantification study.

4. Uncertainty quantification on the synthetic ground motion

In this section, the uncertainty of the numerical earthquake prediction related to the chosen digital geological model is assessed, exploiting EQ2 source. The outcomes of each simulation, obtained by plugging in the three different geological models rendered in Figure 4, are compared in terms of trends and residual values, following the GMPE framework (Rodriguez-Marek et al., 2013). Compared to GMPE, numerical earthquake simulations yield a more accurate prediction of site-specific seismic response, which is generally difficult to cast into non-linear regressions, given their intrinsic uniqueness. Moreover, in this paper, up to 261120 virtual stations were deployed across a square grid, allowing a significant statistical analysis in the post-processing phase.

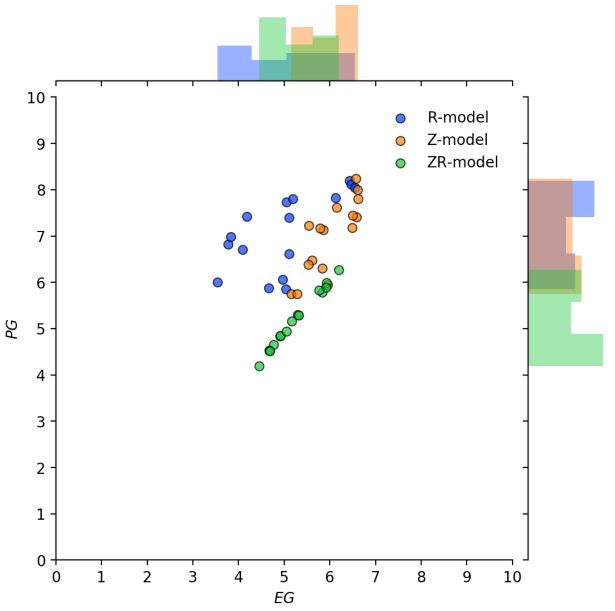

Figure 6: Average time-frequency Envelope GoF (EG) and Phase GoF (PG) (proposed by Kristekova et al., 2009) for recording stations at KKNPP, during EQ1 (aftershock). Synthetic and recorded time-histories were band-pass filtered in the 0.1-2.8 Hz frequency range. GoFs were computed between each numerical model (R-model (blue), Z-model (orange) and ZR-model (green) and recorded signals. The histograms collecting EG and PG GoFs are reported on each axis.

4.1. Sensitivity of physics-based simulations to geological structure

For engineering purposes, the pseudo-spectral acceleration Sa (5% damping) is usually adopted as an insightful IM characterizing the local seismic response. Figure 7 shows the contour values of $Sa(T = 1s)$ for each model (Figure 7a, 7c and 7e), along with the respective contour maps of the thickness of the layer with $Vs < 1500$ m/s (Figure 7b, 7d, 7f). The difference between the synthetic seismic response obtained with R-Model (Niigata basin) and the one obtained with the Z-model (layered regional geology) is evident, which leads to infer that this discrepancy relies on the thickness of surficial soil layers. Compared to the Z-Model and the ZR-Model, which share the same horizontal layering close to the surface, the R-model induce a much more heterogeneous Sa distribution.

Z-model seemingly better fits the recorded response (see Figure 6), whereas including the Niigata basin geology into the earthquake simulation (R-model) over-amplifies the response at shorter periods. This statement is supported by Figure 8a, which shows the response spectra Sa for the three geological models considered. It is observed that Z-Model (layered regional geology and local folding structure) provides the lowest Sa values overall. On the other hand, Figure 8b reports the spectral ratio Sa_*/Sa_{**} between two alternative synthetic responses, $*$ and $**$, respectively. The spectral ratio assesses the relative amplification - at several natural periods - of two seismic responses predicted by considering two different geological models. The influence of the geological structures interacting with the impinging wave field is highlighted according to their characteristic dimensions (e.g., layer thickness) relative to the incident wavelengths. In the period range, $0.5 \text{ s} < T < 5$ s, the ratios range between 0.5 and 1.5, with Sa_{ZR} and Sa_Z exhibiting the far-most comparable response in the range 1.5-3.3 s, where the significant impact of the local folding geology is evident on the synthetic ground motion at KKNPP.

The simulated seismic response was compared to the Abrahamson et al. (2016) GMPE, suitable for the seismic context at stake. Figure 9 displays this comparison, for Sa values at $T = 1$ s (all stations in Figure 9a and geometric means in Figure 9c) and $T = 3$ s (all stations in Figure 9b and geometric means in Figure 9d) along the source-to-site distance R. $\bar{Sa}$ represents the geometric mean response spectra (García-Fernández et al., 2019) expressed as:

$$\bar{Sa}(T) = \frac{1}{N}\sum_{s=1}^{N} \log Sa_s(T) \tag{4}$$

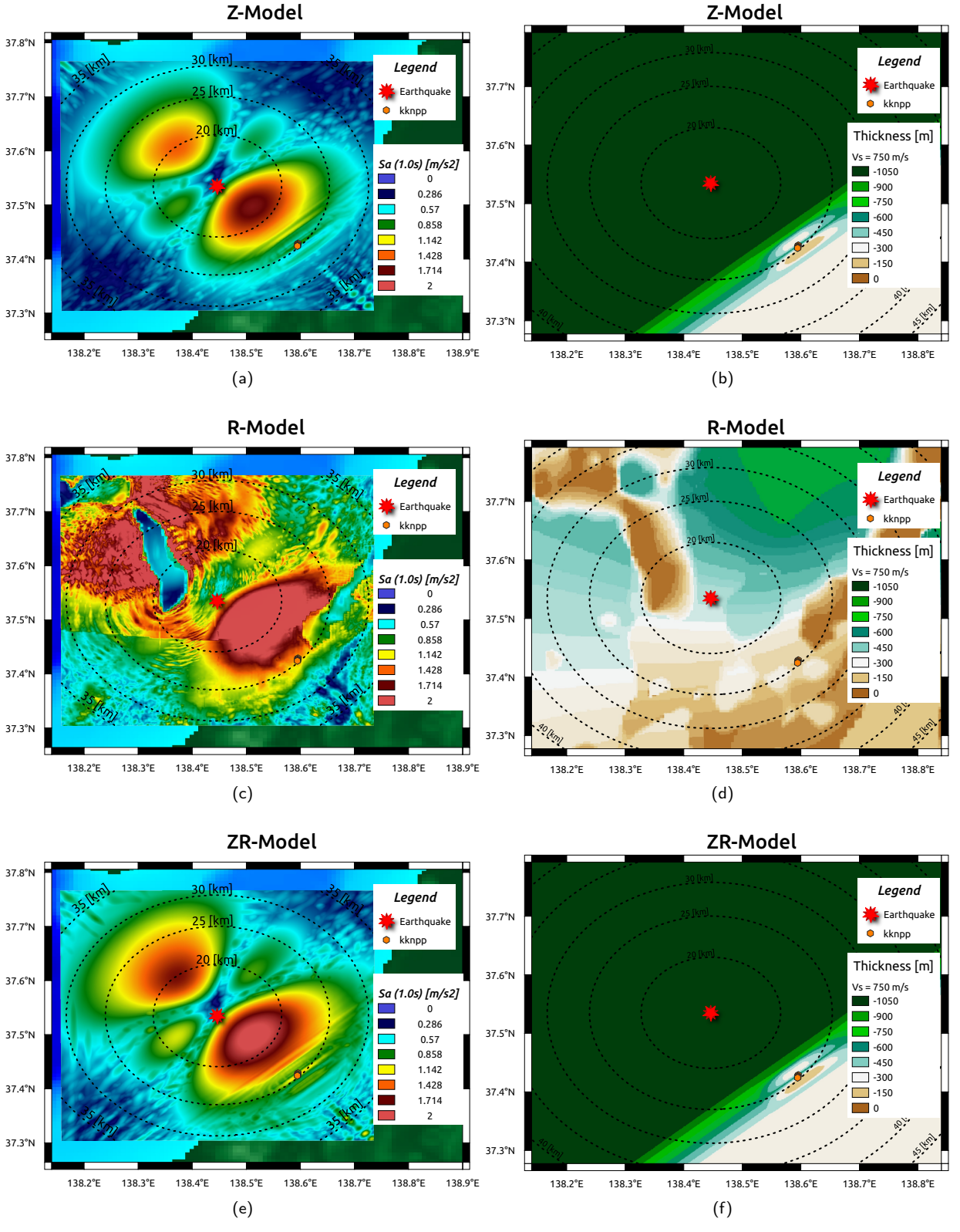

Figure 7: Contour plot of Sa values at $T = 1$ s (5% critical damping) for the three different models: (a) Z-model, (c) R-model, (e) ZR-model. Contour maps of the thickness of the geological layers with $V_S < 1500$ m/s for the three models: (b) Z-model, (d) R-model, (f) ZR-model.

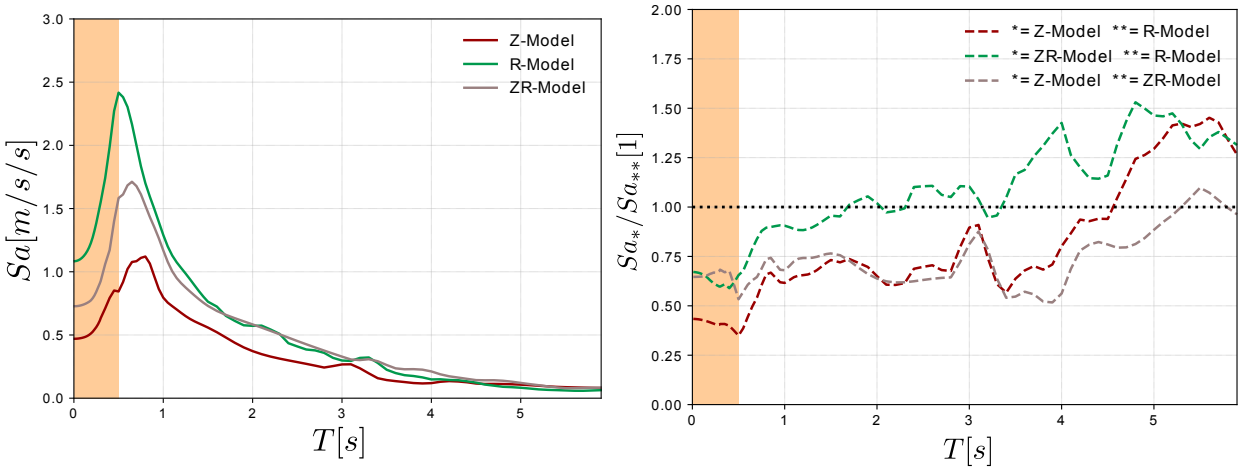

Figure 8: Comparison of the response spectra Sa at the KKNPP location (KSH - 37.4° N, 138.6° E) for the three geological models considered. (a) measures the Sa geometric average of the horizontal components. (b) measures the ratio Sa_*/Sa_{**} between two alternative synthetic response $*$ and $**$, corresponding to different geological models. Orange area covers the frequency band highly affected by the applied filter ($T < 1/f_{max}$).

with Sa_s being the response spectra at the station s. For each model, the synthetic Sa values (both at $T = 1$ s and $T = 3$ s) are quite dispersed, compared to the $\pm\sigma$ (i.e. plus/minus one standard deviation) margins indicated by the Abrahamson et al. (2016) GMPE, for a reference $V_{S,30} = 400$ m/s (approximately corresponding to the lowest V_S value in the R-model, see Figure 4). However, Sa values computed at the KKNPP (KSH-SG1 station in Figure 5, referenced by a star in Figure 9a and Figure 9b) for R-model are higher than $\bar{Sa}$, at the same distance from the source. On the contrary, despite the fact that the smallest V_S value in Z-model is 700 m/s (the Nishiyama layer in Table 1) the corresponding Sa value computed at the KKNPP is in very good agreement with the $\bar{Sa}$ prediction of the Abrahamson et al. (2016) GMPE. This inconsistency can be explained by the fact that the GMPEs generally lack an explicit description of 3-D site effects, that cannot be effectively condensed into the $V_{S,30}$ value.

It is interesting to notice that trends in Figure 9c and Figure 9d can be adopted as site-specific and scenario-specific GMPE predictions. A stable decaying trend is observed within the 0-35 km distance range, where the number of stations is statistically representative. The thinner solid lines in Figure 9c and Figure 9d represent the deviation from the station-average Sa value at a fixed distance to the source. The R-model and the Z-Model present the overall largest and smallest deviation. Compared to horizontally layered geology (featuring the Z-model), the R-Model embraces the intricate basin-like structures of the Niigata region, arguably inducing an incoherent ground motion across the region of interest. The ZR-model and Z-model show comparable trends, seemingly due to the folding structure they share, located underneath the KKNPP, which focuses the wave field in that region, compared to the coherent wave field in the rest of the region due to the sub-horizontally layered geology (as observed and numerically reproduced by Tsuda et al., 2011; Gatti et al., 2018c, , among others, see Figure 7).

To summarize, the R-model has an intricate geology, with softer sediments close to the surface, which amplify the ground motion in the surroundings. The outcome from the Z-model is systematically lower than the others at all distances from the source (see Figure 9), although representing a valid geological model to predict the seismic response at the KKNPP. The ZR-model represents a trade-off between the Z-model and R-model.

In order to further highlight the correlation between different geological model, Figure 10 displays the correlation coefficient ρ of the Sa values simulated by each numerical model, all sharing the same earthquake sources. Figure 10 shows the high correlation between the Z-Model and the ZR-Model ($\rho_{Z,ZR} \geq 0.6$), for both $Sa(T = 1$ s) and $Sa(T = 3$ s). This is seemingly related to the folding structure underneath KKNPP, shared by both Z- and ZR-model models and that impacts the correlation at short distances (i.e., around 22-24 km, where the near-field effects are predominant). The correlation coefficients $\rho_{Z,R}$ and $\rho_{R,ZR}$ range around a value of ≈ 0.5 for $Sa(T = 1$ s), evolving in the

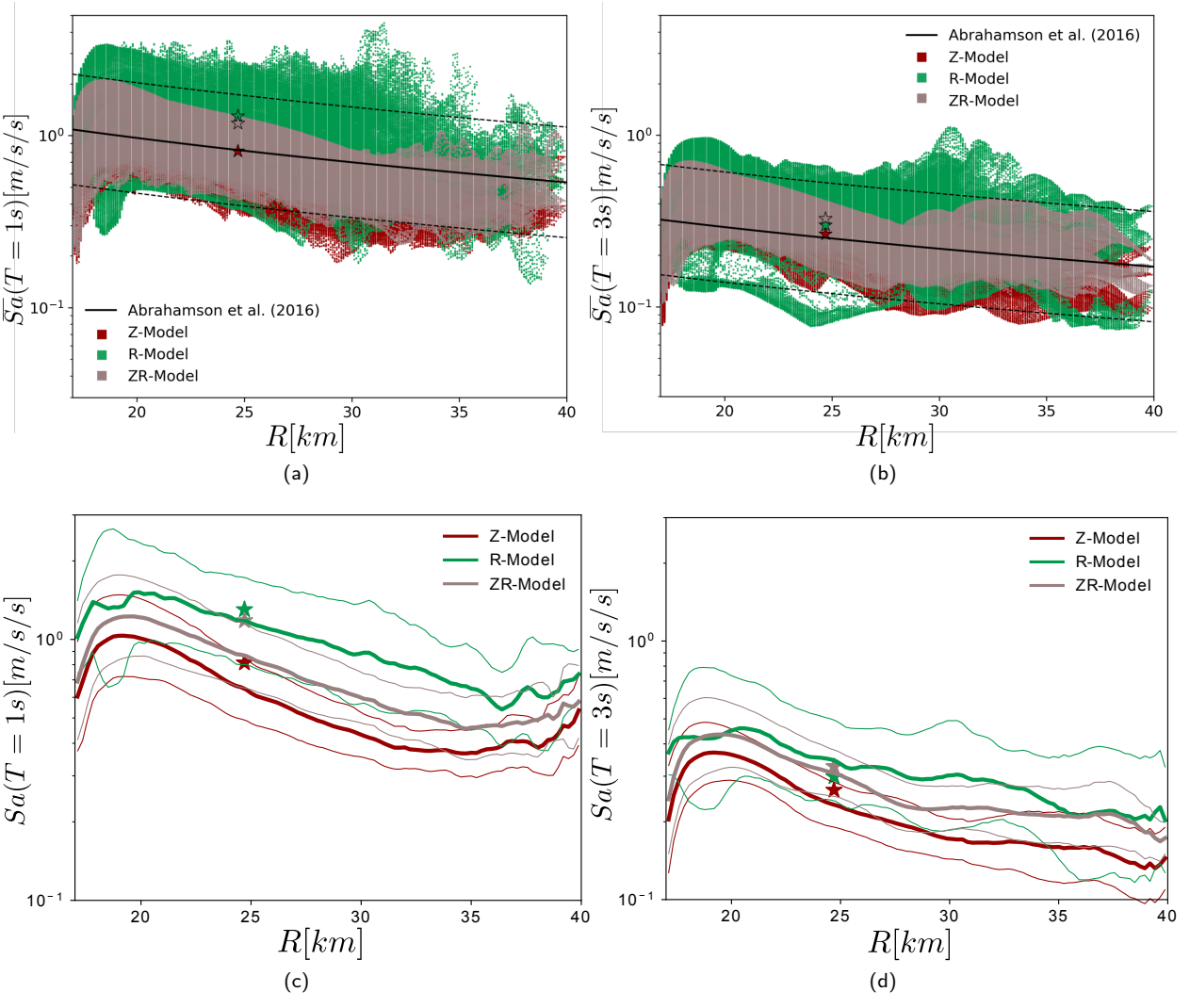

Figure 9: Sa values obtained at surface for the three physics-based simulations at: (a) and (c) $T = 1$ s; (b) and (d) $T = 3$ s. (a-b) show the synthetic response at each station, compared with the GMPE proposed by Abrahamson et al. (2016) (computed for a reference $V_{S,30} = 400$ m/s. The geometric mean $\bar{Sa}$, for the three physics-based simulations, is depicted with a solid thick line in (c-d), whereas the respective $\pm\sigma$ margins are marked by thinner dashed lines. Stars indicate the free-field response at KSH station, located within the KKNPP.

same way along with the source-to-site distance, but they consistently drop below 0.4 at longer period (see Figure 10b). A possible explanation comes from the impact of the deep geological structure of the Niigata basin (R-model) on the long-period ground motion prediction ($Sa(T= 3$ s)), which is remarkably different from the one estimated with the layered configuration in the Z-model. However, the influence of the complex yet deep geological structure seemingly vanishes at shorter periods (e.g., at $T = 1$ s), where Sa values are mainly affected by the folding structure located within the first 5 km below the surface.

Hereafter, the synthetic seismic response estimated for the three geological models is analyzed in order to estimate the residual of the synthetic strong ground motion prediction.

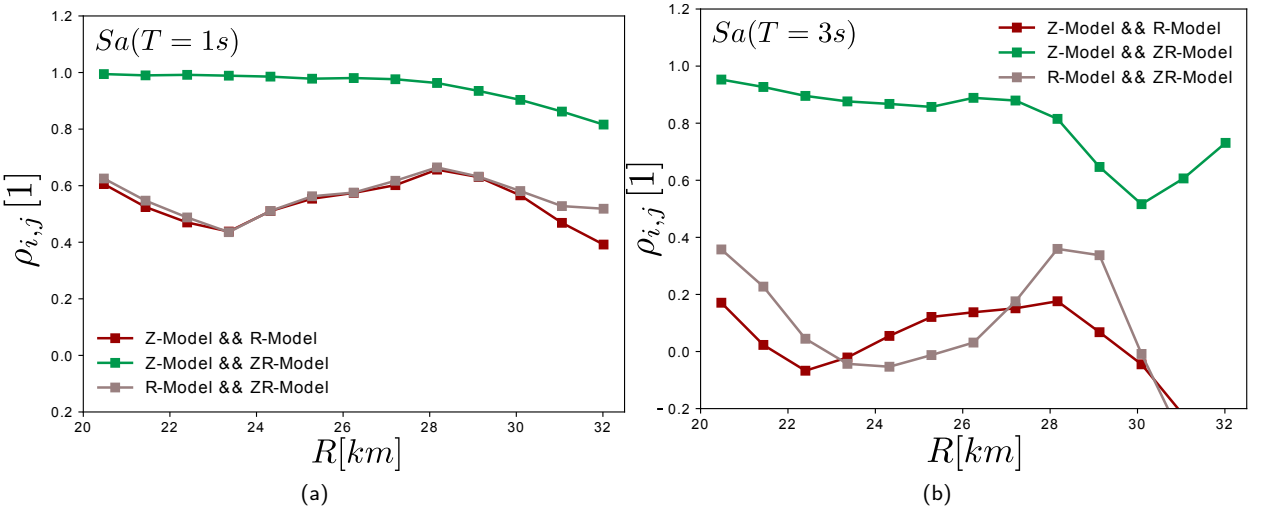

Figure 10: Correlation coefficient $\rho_{i,j}$ of the Sa values, computed between the i^{th} and j^{th} numerical models as a function of the source-to-site distance R. (a) Correlation coefficient of $Sa(T= 1$ s); (b) correlation coefficient of $Sa(T= 3$ s).

4.2. Estimation of within-event variability

Several studies (e.g., Atkinson, 2011; García-Fernández et al., 2019) estimated the standard deviation ϕ of the within-event residual δW_{es} at a certain source-to-site distance as:

$$\phi(T) = \sqrt{\frac{1}{N-1}\sum_{s=1}^{N}\left[\log\left(\frac{Sa_s(T)}{\bar{S}a(T)}\right)\right]^2} \tag{5}$$

Figure 11 displays the estimation of ϕ for the present study, at different source-to-site distances and for $T = 1$ s and $T = 3$ s. For Z-model and ZR-model and for both natural periods, Figure 11 depicts a rather constant trend along

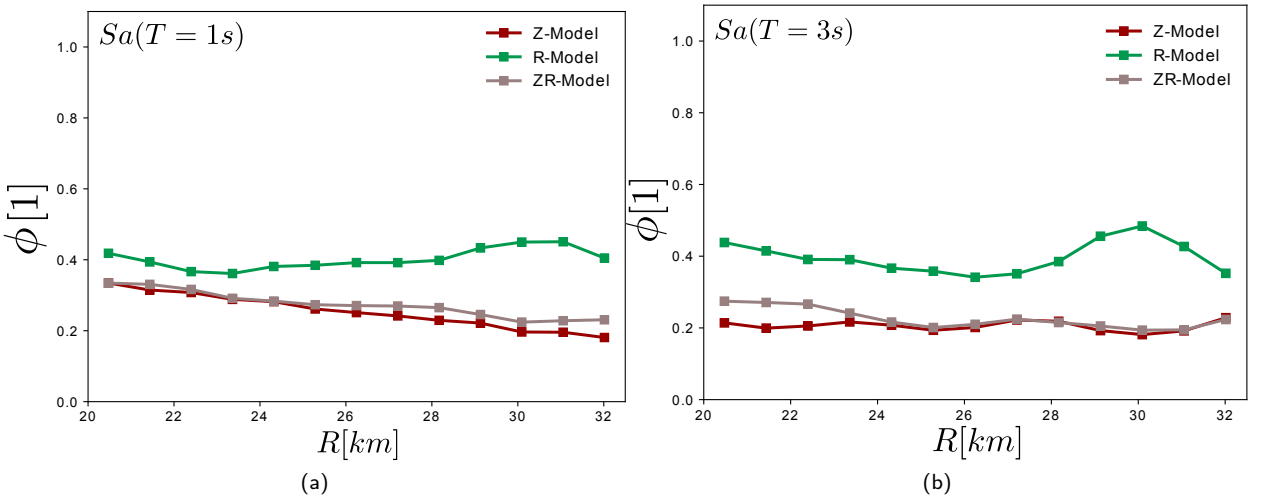

Figure 11: Standard deviation ϕ of the within-event residual δW_{es} computed on Sa with respect to the source to site distance R, at natural period $T = 1$ s (a) and $T = 3$ s (b).

with the source-to-site distance. This constant trend is insignificant for $T = 3$ s. Since those models share the same shallow horizontal layering (including the folding), ϕ is mainly related to shallow geology, indicating once again its major impact compared to deep geology. Exploiting the rare advantage of having many stations close to the source for a single event, Figure 11 proves that ϕ is approximately independent of the distance, even in the case of R-Model where there is a complex layering (non-horizontal). In many cases, the value of ϕ is considered to be independent on the distance to the source (Atkinson, 2011). However, recent studies proposed a functional dependence of ϕ on the distance from the source (see Boore et al., 2014), although this variation occurs at distances larger than 80 km.

4.3. Influence of the site effects on the within-event variability

To explain the variability of the within-event residuals (δW_{es}), its sensitivity with the respect to the local site conditions is hereafter analyzed. In doing so, δW_{es} is related to several parameters such as: (1) the harmonic-mean shear velocity in the first d m below the surface ($V_{S,d}$) and (2) the thickness of the equivalent basin-like structure (H_v). Those parameters are used in the recipes proposed by Eurocode 8 (EC8, CEN, 2004) and they are currently being discussed in the framework of the EC8 site categorization update (Paolucci et al., 2021). $V_{S,d}$ is often computed for $d = 30$ m and $d = 100$ m. The thickness of the equivalent basin-like structure is instead defined based on the shear-wave velocity value of the *engineering bedrock*, such as H_{800}, corresponding to the depth at which the *engineering bedrock* is found, i.e. the geological layer at which the shear velocity ranges around 800 m/s (according to Ansal and Tönük, 2007). In this study, H_{1500} is also considered, in order to assess the role of deeper geological layers (such as the Shiiya and Upper-Teradomari layers in the folding structure, see Figure 3). Table 4 shows the correlation between the R-Model parameters for each mesh point location at surface. As expected, $V_{S,30}$ and $V_{S,100}$ are well-correlated with $\rho = 0.98$. Similarly, the thicknesses H_{800} and H_{1500} are correlated with $\rho = 0.78$. $V_{S,30}$ and $V_{S,100}$ and H_{800} and H_{1500} are poorly correlated and they can be used as complementary parameters to predict δW_{es} better and reduce ϕ. Figure 12 shows

Table 4
Correlation matrix of the four analyzed site parameters for the R-model: $V_{S,30}$, $V_{S,100}$, H_{800} and H_{1500}.

	$V_{S,30}$	$V_{S,100}$	H_{800}	H_{1500}
$V_{S,30}$	1.0	0.98	0.40	0.40
$V_{S,100}$		1.00	0.50	0.46
H_{800}			1.00	0.78
H_{1500}	sym			1.00

δW_{es} values as a function of the four parameters $V_{S,30}$, $V_{S,100}$, H_{800} and H_{1500} respectively. Positive δW_{es} indicate that the synthetic ground motion at the station of interest is larger than the geometric mean prediction. Negative δW_{es} values, on the contrary, point out that the synthetic IM at the station underestimate the geometric mean prediction. Adopting the R-model, the δW_{es} estimation (for both $T = 1$ s and $T = 3$ s) is positive but it rapidly decays with $V_{S,30}$ and $V_{S,100}$ values below 800 m/s, whereas it remains approximately constant and close to -1 for larger shear-wave velocity values. Numerical simulations seemingly overestimate the $\bar{S}a$ values across softer sediments, highlighting the presence of 3-D site effects that the GMPEs can barely picture. Moreover, the trend in Figure 12 are in agreement with the EC8 classification and with the GMPE guidelines (e.g., Abrahamson and Silva, 1993; Lee et al., 1995) that usually classify a site according to its $V_{S,30}$, being a *soil deposit* for $V_{S,30} < 800$ m/s and engineering bedrock for $V_{S,30} \geq 800$ m/s. In addition, this result supports the choice of H_{800} to condense the basin-like site effects.
On the contrary, no clear trend is observed for the estimated δW_{es} at $T = 1$ s neither with respect to H_{800} nor with the respect to H_{1500} (see Figure 12). At longer period (i.e., $T = 3$ s), as suggested by decaying δW_{es} trend approaching small H_{800} or H_{1500}, the residual δW_{es} diminishes for lower H_{1500}, highlighting the influence of harder and deeper rock layers on the site effect prediction.

Figure 13 shows the absolute correlation coefficient $|\rho|$ between δW_{es} and each parameter $V_{S,30}$, $V_{S,100}$, H_{800} and H_{1500}, at $T = 1$ s (blue histogram) and $T = 3$ s (orange histogram). In this figure, the same correlation is computed for the Z-model for the sake of comparison. $|\rho|$ with respect to H_{800} and H_{1500} almost doubles at long period ($T = 3$ s) highlighting the importance of site effects on the residual estimation. In the case of Z-Model (Figure 13b), there is no correlation between δW_{es} and $V_{S,30}$ and poor correlation with $V_{S,100}$, due to the horizontal layering and the coarse geological model. Both parameters $V_{S,30}$ and $V_{S,100}$ are uninformative at this scale and resolution. However, H_{800} and H_{1500} represent complementary site-characterization parameters to $V_{S,30}$.

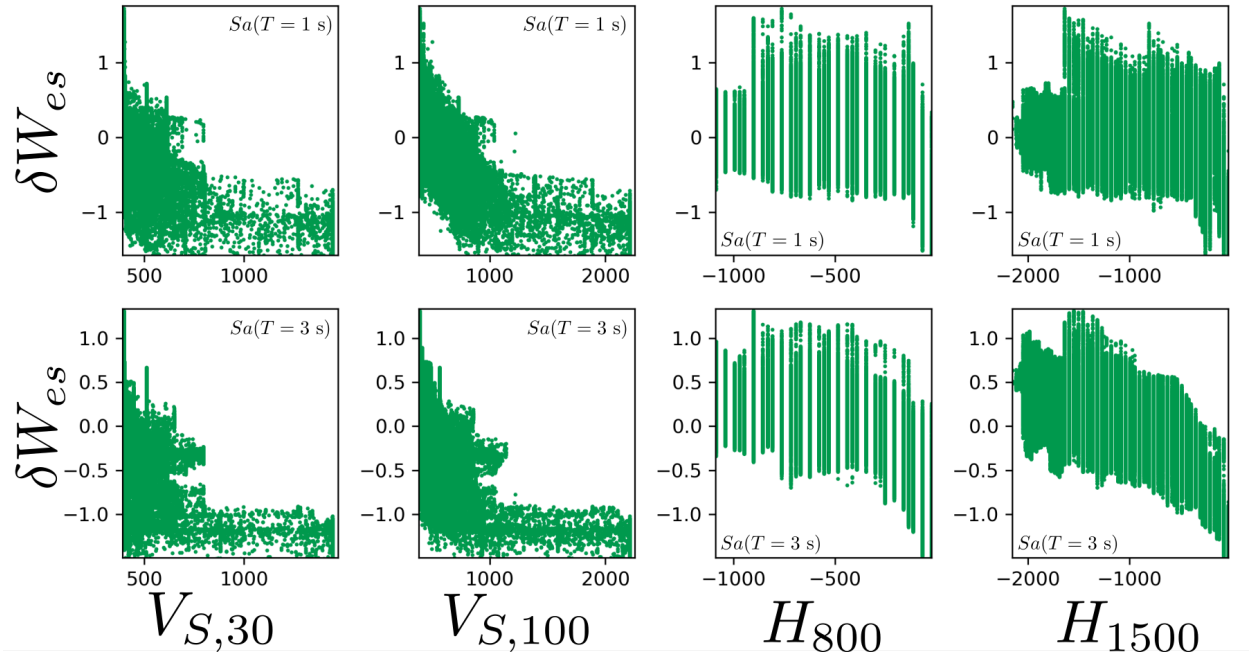

Figure 12: Within-event residual δW_{es} computed for the R-model, as function of $V_{S,30}$, $V_{S,100}$, H_{800} and H_{1500}. Top panels: δW_{es} computed for $Sa(T = 1\ \mathrm{s})$. Bottom panels: δW_{es} computed for $Sa(T = 3\ \mathrm{s})$.

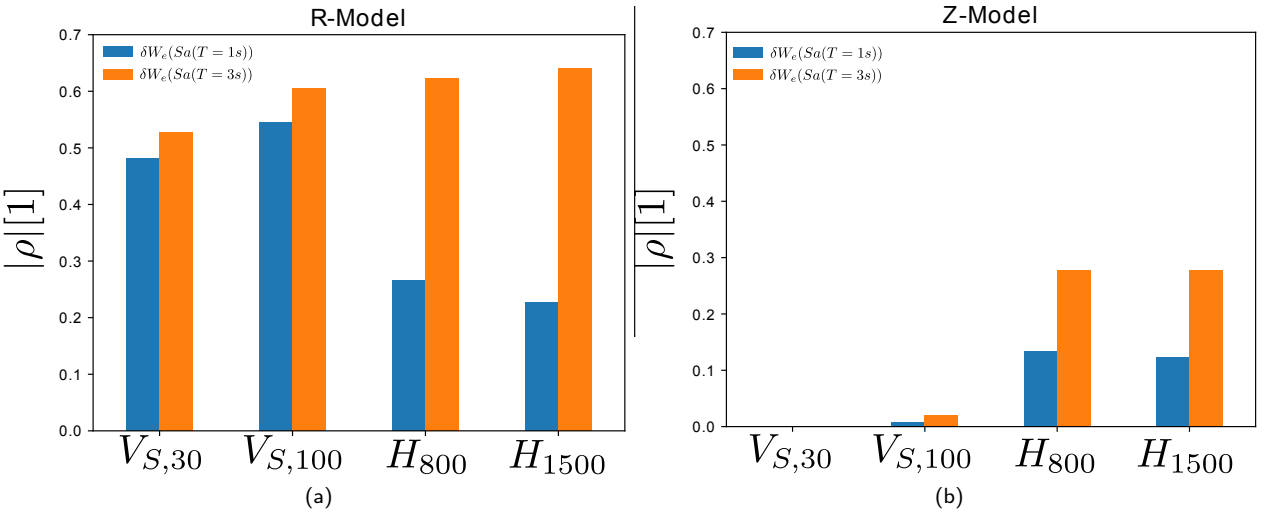

Figure 13: Absolute correlation coefficient $|\rho|$ of δW_{es} with site parameters. R-Model (a) and Z-Model horizontal layered case (b).

Finally, the intensity of the ground motion is affected by the pattern of the source. Figure 7 showed that Sa spatial variability at surface depends on the station back-azimuth. Figure 14 confirms this assumption, displaying the value of ϕ by clustering each station based on its back-azimuth.

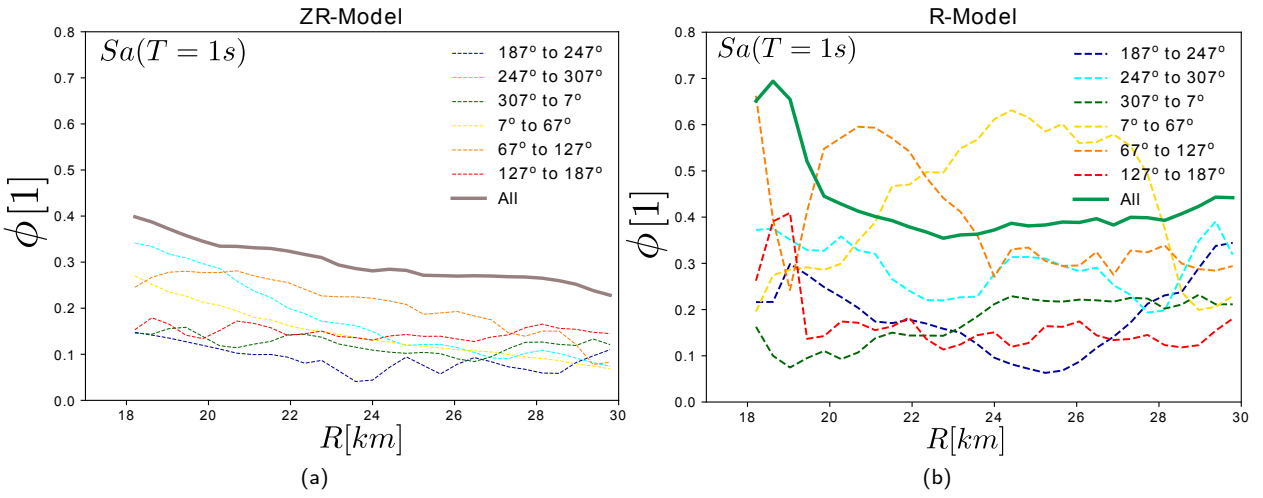

Figure 14: Standard deviation ϕ of the within-event residual δW_{es} as a function of the source-to-site distance R. ϕ values at each station are clustered based on the station back-azimuth. (a) ZR-Model; (b) R-Model.

For the ZR-model, featured by a horizontally layered regional geology (see Figure 14a), the ϕ clustering based on the back-azimuth highlights a maximum difference of up to 40% between clusters. When considering the R-Model (see Figure 14b), those discrepancies between cluster are less evident. For back-azimuths between 7° and 127°, the high ϕ values are seemingly related to the complex geology in this zone. This is congruent with several studies (such as Ripperger et al., 2008; Vyas et al., 2016) that also proved the variation of ϕ as a function of the source pattern effects. In this study, point sources are considered, although the back-azimuth dependency could be higher for extended-fault scenarios due to directivity effects.

4.4. Estimation of between-event variability

As far as δB_e is concerned, the extensive number of monitoring points allows to draw solid statistics in a large source-to-site distance range. In this study, the quota of epistemic uncertainty linked to the source mechanism is not investigated. Therefore, δB_e can be indirectly estimated by disentangling it from the within-event variability. Figure 15 shows the estimated δB_e for both $Sa(T= 1$ s) and $Sa(T= 3$ s). However, the fact that the three cases considered herein display the same source pattern and location affects its estimation (Atik et al., 2010). Concerning the variability of the between-event residual τ, one easily observes that the variance of the total residual σ^2 is the sum of the variances of between-event and within-event residuals, which are modeled as independent random variables. Therefore, the between-event variance is estimated as:

$$\tau^2 = \sigma^2 - \phi^2 \tag{6}$$

with ϕ^2 estimated in Section 4.2 from the synthetic time-histories. Therefore, the first step is to compute the model-to-model deviation, inferred directly from the Sa_*/Sa_{**} ratio, according to the following expression:

$$\Delta_{Sa_*/Sa_{**}} = \Delta_{Sa_*} - \Delta_{Sa_{**}} \tag{7}$$

where Δ_{Sa}, for any model and period, is considered as distributed as centered log-normal random variable. With this assumption, the variance of $\Delta_{Sa_*/Sa_{**}}$, namely $\sigma^2_{Sa_*/Sa_{**}}$ is computed as:

$$\sigma^2_{Sa_*/Sa_{**}} = \sigma^2_{Sa_*} + \sigma^2_{Sa_{**}} - 2 \cdot \rho_{Sa_*,Sa_{**}} \cdot \sigma_{Sa_*} \cdot \sigma_{Sa_{**}} \tag{8}$$

with σ^2_{Sa} corresponding to the variance of the total residual of each model and $\rho_{Sa_*,Sa_{**}}$ being the correlation coefficient between each combination of models, displayed in Figure 10. The evolution of $\sigma_{Sa_*/Sa_{**}}$ for each model duos, as a function of the source-to-site distance, is displayed in Figure 16. The standard deviation of the $\sigma_{Sa_Z/Sa_{ZR}}$ ratio

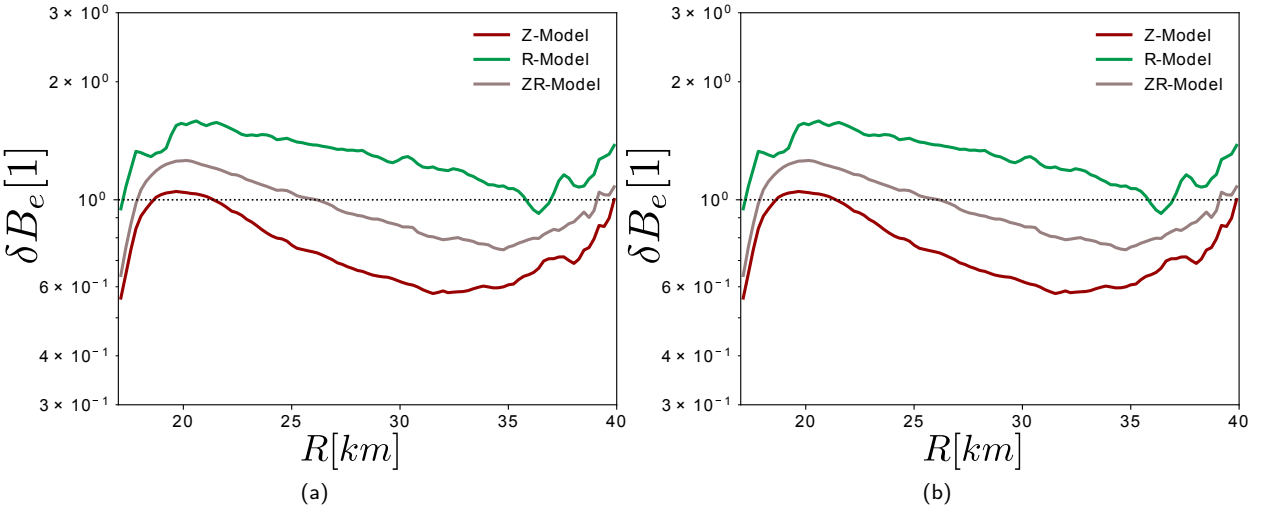

Figure 15: Between-event residuals for (a) $Sa(T= 1$ s) and (b) $Sa(T= 3$ s) as functions of the source-to-site distance R.

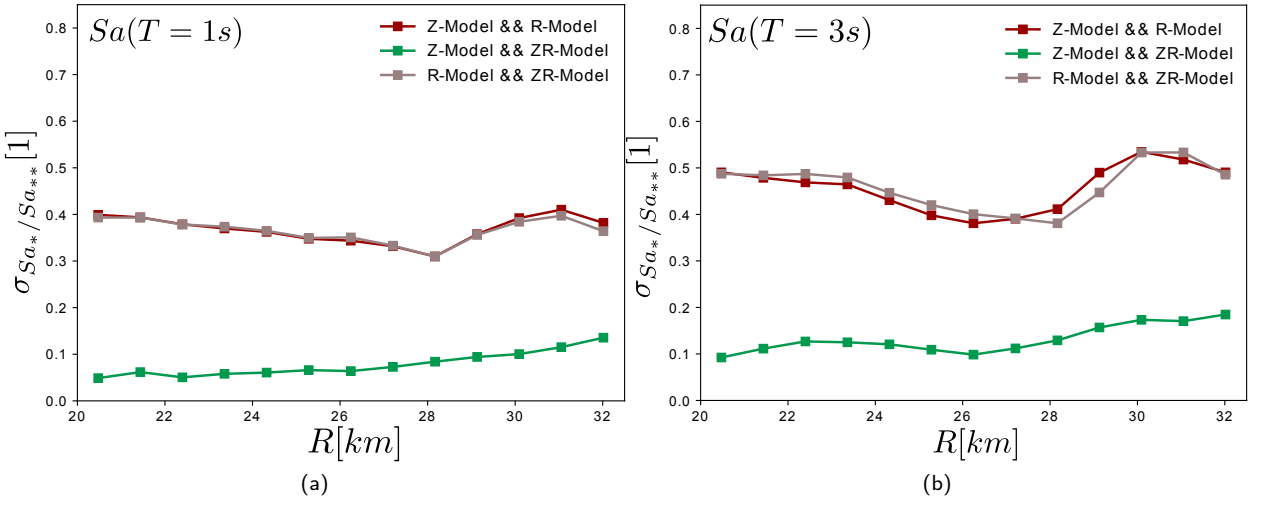

Figure 16: Standard deviation $\sigma_{Sa_*/Sa_{**}}$ computed along source-to-site distance R for each model duos (*,**) and for (a) $T = 1$ s and (b) $T = 3$ s.

is very low. This aspect, along with the high correlation between both models (see Figure 10), indicates that, even if they do not share similar geologies at depth, the two models are almost interchangeable for this very earthquake.

The second step to estimate the variability of δB_e consists into assembling a system of the three equations in the form of Equation (8), considering σ_{Sa_R}, σ_{Sa_Z} and $\sigma_{Sa_{ZR}}$ as unknowns. The system is iteratively solved by exploiting the outcome of the three earthquake simulations. Finally, τ is computed by exploiting Equation (6). Figure 17 depicts the τ estimation along the source-to-site distance (not taking into account the variability that could come from the model or the parameterization of the source event). Figure 17 shows that the three τ curves are slowly varying along the source-to-site distance. Despite some minor fluctuation between 20 and 24 km, τ is practically constant for all the three models at $T = 3$ s. Moreover, the curves referring to the ZR-Model and the Z-Model are very similar for both $T = 1$ s (slowly decaying values with increasing distance from the source) and $T = 3$ s (lower deviation, constant

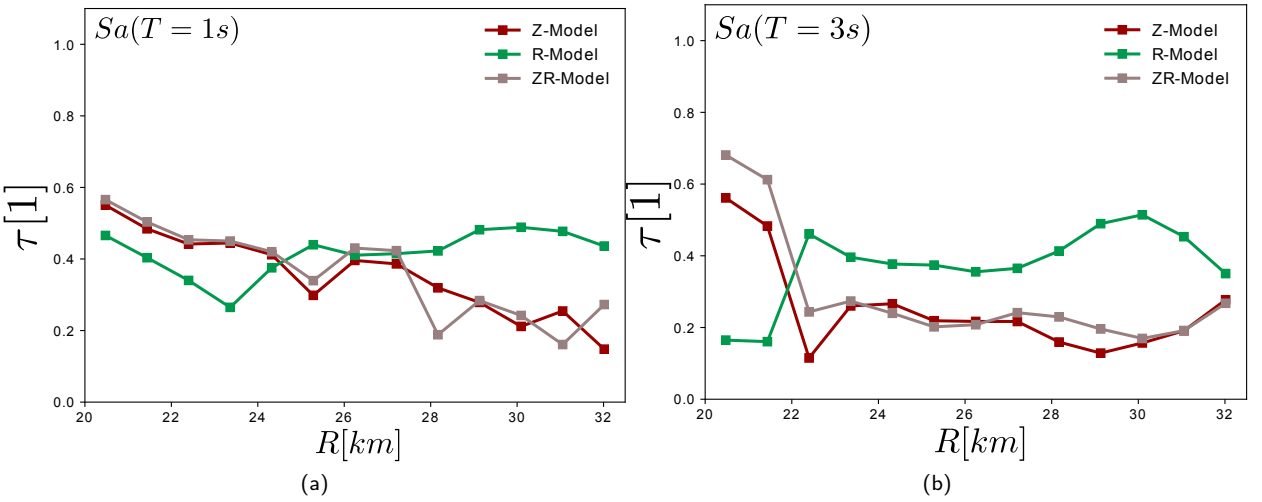

Figure 17: Standard deviation τ of the between-event residual δB_e computed along source-to-site distance R, for $T = 1$ s (a) and $T = 3$ s (b).

along the source-to-site distance). The deviation associated to the R-model is almost insensitive to the natural period ($\tau \approx = 0.45$) except for its drop to lower values at short distances and long period, arguably due to the rather constant geological properties (compared to the layered medium in Z-model) at the surface depth.

5. Conclusions

This study compares the synthetic outcomes of three physics-based numerical models, each featuring a different yet plausible geological structure for the region at stake. A realistic test case (the 2007 Niigata earthquake) is studied, so to prove the sensitivity of high-fidelity earthquake simulators to input geological information. The comparison proved that complex geological features (the Niigata regional basin plugged into the R-Model), compared to horizontally layered deep geology and folding geology close to the surface (Z-model and ZR-Model) might induce less than 10% difference on the surface seismic response (in terms of Sa values), leading to the possibility of simplifying the geological medium to reduce the computational costs). Moreover, physics-based simulations were employed to study the geology's influence on the spatial variability of the ground shaking.

Considering the 261120 virtual stations covering the region of interest, the paper adopted the GMPEs formulation in Equation (1) to infer synthetic site-specific ground motion prediction variability. From the analysis, we evinced that ϕ (i.e. the within-event residual standard deviation) mostly depends on shallow layers properties, increasing with the geological complexity. Furthermore, the two models sharing the surficial site conditions but different ones at deep layers (Z-Model and ZR-Model) show a remarkable difference in the estimated between-event residual δB_e by an almost constant factor (with source-to-site distance). This suggests that records from different tectonic regions but similar site conditions (basin-like structures) can be integrated into a site-specific GMPE, by applying a correction factor that considers the diverse deep geology.

The impact of complex layering on τ (i.e. the between-event residual standard deviation) is reflected in the R-model results, where higher τ is obtained than in simpler geology models. Shallow geology is proven as a significant factor in determining the average trend and variability of the synthetic seismic site response. At long period ($T = 3$ s), the between-event variability τ is seemingly highly affected by the choice of the geological model, in a source-to-site distance range of 0-25 km.

The outcome of the numerical exercise suggests the possibility of disentangling the contribution of shallow and deep geological structures on the overall synthetic seismic prediction at the site.

Finally, we unveiled a relation between the within-event residuals (δW_{es}) and the properties of the sedimentary layers, such as basin thickness and harmonic shear-wave velocity. Those two parameters are seemingly equally correlated to

the residual δW_{es}. In particular, the depth of the *engineering bedrock* is a rather easily accessible information, inferred from regional geological maps or obtained by non-invasive geophysical tests. Therefore, either H_{800} or H_{1500} are suitable to reduce the variability ϕ of the Sa GMPE prediction, especially at long period. In this study, the influence of the source pattern and the back-azimuth on both ϕ and τ variability is assessed.

For the present case study, the results of this numerical analysis are congruent with several previous studies on real records. This shows that the numerical simulation from a well-calibrated model can be integrated with real records to create synthetic GMPEs, or complete the database of a region to produce more accurate GMPE. The previous results were found with a frequency band 0-2.1 Hz and not including the topographical effects. It is possible that at higher frequencies, the local conditions are more relevant, especially in the value of ϕ. However, the results found in this simulation are still valid.

Acknowledgements

The research reported in this paper has been supported in part by the SEISM Paris Saclay Research Institute. This work, within the SINAPS@ project, benefited from French state funding managed by the National Research Agency under program RNSR Future Investments bearing Reference No. ANR-11-RSNR-0022-04, from the FY 2018 Postdoctoral Fellowship issued by the Japanese Society for Promotion of Science (JSPS) International Fellowships for Research in Japan, hosted by Prof. H. Kawase at Disaster Prevention Research Institute (DPRI) and by the digital geological data provided by Prof. H. Sekiguchi, as well as from valuable strong motion records during the 2007 Niigataken Chuetsu-Oki earthquake disseminated by TEPCO (Tokyo Electric Power Company).

A. Details of the geological models

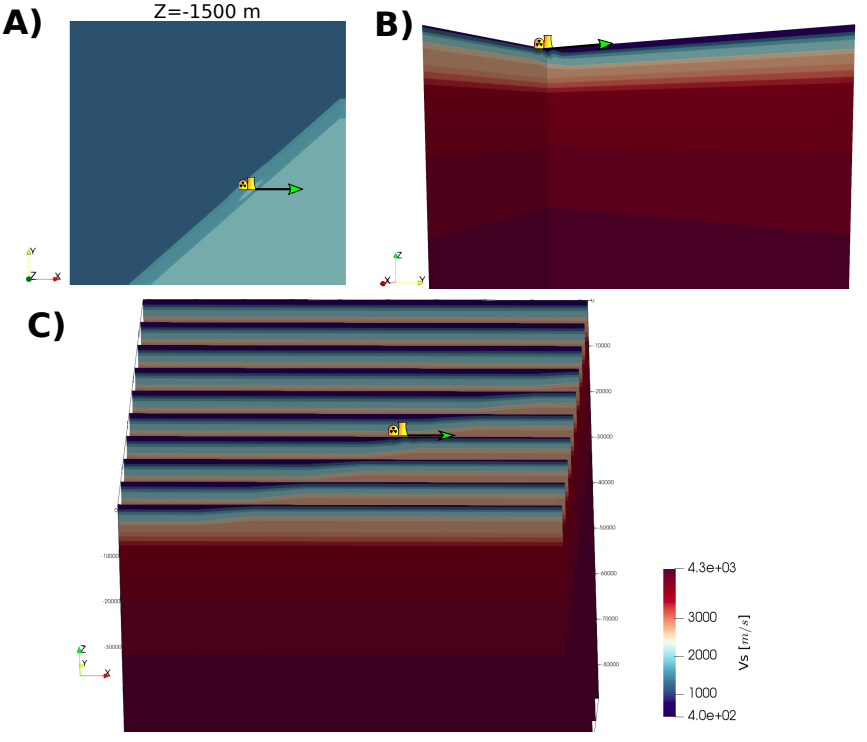

Figure 18: View-clips of the Z-Model. A) Superior view at z = -1500 m. B) Longitudinal cut by the South-North direction (y-Axis). C) different slips by the East-West (x-axis) direction.

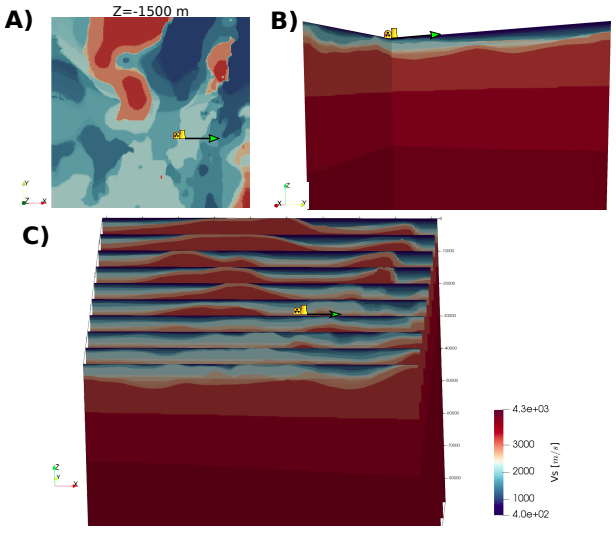

Figure 19: View of the R-Model marking the KKNPP comples with the nuclear plant symbol. A) Superior view at z = -1500 m. B) Longitudinal cut by the South-North direction (y-Axis). C) different slips by the East-West (x-axis) direction.

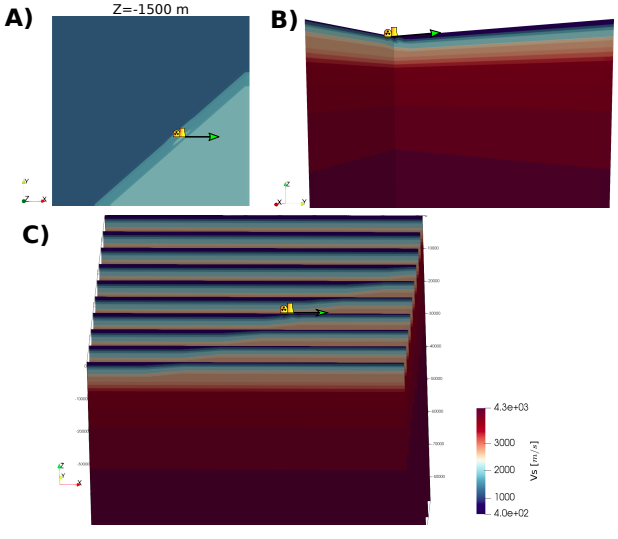

Figure 20: View of the ZR-Model marking the KKNPP buildings with the nuclear plant symbol. A) Superior view at z =1500 m. B) Longitudinal cut by the South-North direction (y-Axis). C) different slips by the East-West (x-axis) direction.

B. Performance of the numerical analysis

This work was granted access to the HPC resources of CINES under the allocation 2018-A0040410444, 2019-A0060410444, 2020-A0070411083 and 2020-A0080410444, made by GENCI (Grand équipement national de calcul intensif). Computations were also performed using HPC resources allocated by the Mésocentre Moulon, the super-computer of CentraleSupélec and École Normale Supérieure de Paris Saclay, Paris Saclay University.

Given the scalability properties of SEM3D (https://www.researchgate.net/publication/349254101_SEM3D-High-res seismic_wave_propagation_modelling_from_the_fault_to_the_structure_for_realistic_earthquake_

scenarios_GENCI_Allocation_A0080410444/stats), each run (30 s earthquake simulation, on the template model) took approximately 80 minutes (≈ 960 hours CPU-time) on 720 MPI cores Intel Xeon Gold 6230 20C @2.1 GHz Cascade Lake.

C. Numerical model validation

In Figure 21, the synthetic and recorded accelerograms and Kristekova's GoF at KSH (SG4) are compared, considering each numerical model (R-model, Z-model and ZR-model respectively). As expected, the synthetic ground motion rendered at the KKNPP differs remarkably from model to model. Moreover, Z-Model displays a shorter *coda*-wave compared to the R-Model, where the scattering effect caused by the more intricate geological structure increases the effective propagation path and rendering large-duration time-histories.

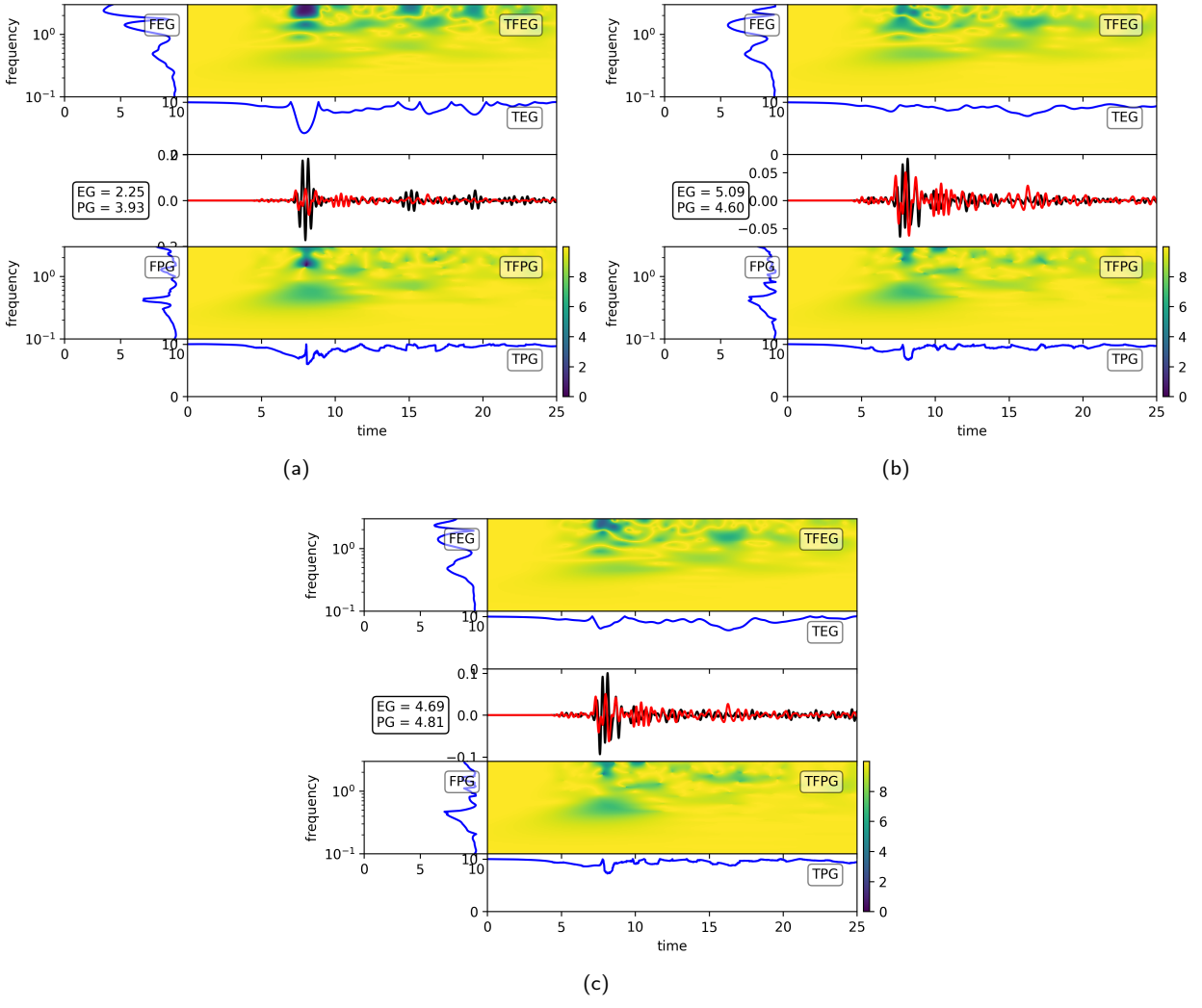

Figure 21: Synthetic acceleration time-histories (EW component, filtered between 0-2.1 Hz) at KSH location (device SG4, G.L. -250 m) for the three models for EQ1. Red time-histories refer to earthquake recording (TEPCO, 2007), whereas black time-histories refer to the outcome of the numerical simulation. (a) R-model; (b) Z-model; (c) ZR-model.

References

Abrahamson, N., Gregor, N., Addo, K., 2016. BC Hydro ground motion prediction equations for subduction earthquakes. Earthquake Spectra 32, 23–44. doi:10.1193/051712EQS188MR.

Abrahamson, N., Hollenback, J., 2012. Application of single-station sigma ground motion prediction equations in practice, in: Proceedings of the 15th world conference on earthquake engineering, Lisbon, Portugal, paper.

Abrahamson, N., Silva, W., 1993. Attenuation of long period strong ground motions. ASME-PUBLICATIONS-PVP 256, 187–187.

Ansal, A., Tönük, G., 2007. Source and site factors in microzonation, in: Earthquake geotechnical engineering. Springer, pp. 73–92. doi:10.1007/978-1-4020-5893-6_4.

Aochi, H., Ducellier, A., Dupros, F., Delatre, M., Ulrich, T., de Martin, F., Yoshimi, M., 2013. Finite difference simulations of seismic wave propagation for the 2007 Mw 6.6 Niigata-ken Chuetsu-Oki earthquake: Validity of models and reliable input ground motion in the near field. Pure and Applied Geophysics 170, 43–64. URL: https://hal-brgm.archives-ouvertes.fr/hal-00980238, doi:10.1007/s00024-011-0429-5. springer Verlag (Germany).

Aoi, S., Kunugi, T., Fujiwara, H., 2004. Strong-motion seismograph network operated by NIED: K-NET and KiK-net. Journal of Japan association for earthquake engineering 4, 65–74.

Aoi, S., Sekiguchi, H., Morikawa, N., Kunugi, T., 2008. Source process of the 2007 Niigata-ken Chuetsu-oki earthquake derived from near-fault strong motion data. Earth, Planets and Space 60, 1131–1135. URL: http://dx.doi.org/10.1186/BF03353149, doi:10.1186/BF03353149.

Atik, L.A., Abrahamson, N., Bommer, J.J., Scherbaum, F., Cotton, F., Kuehn, N., 2010. The Variability of Ground-Motion Prediction Models and Its Components. Seismological Research Letters 81, 794–801. URL: https://doi.org/10.1785/gssrl.81.5.794, doi:10.1785/gssrl.81.5.794, arXiv:https://pubs.geoscienceworld.org/ssa/srl/article-pdf/81/5/794/2762386/794.pdf.

Atkinson, G.M., 2011. An empirical perspective on uncertainty in earthquake ground motion prediction. Canadian Journal of Civil Engineering 38, 1002–1015. doi:10.1139/l10-120.

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, 1057–1085. doi:10.1193/072114EQS116M.

Bradley, B.A., 2018. On-going challenges in physics-based ground motion prediction and insights from the 2010-2011 Canterbury and 2016 Kaikoura, New Zealand earthquakes. Soil Dynamics and Earthquake Engineering , –URL: https://www.sciencedirect.com/science/article/pii/S0267726117308631, doi:https://doi.org/10.1016/j.soildyn.2018.04.042.

Casarotti, E., Stupazzini, M., Lee, S., Komatitsch, D., Piersanti, a., Tromp, J., 2008. Cubit and seismic wave propagation based upon the spectral-element method: An Advanced Unstructured Mesher for Complex 3D Geological Media, in: Proceedings of the 16th International Meshing Roundtable, pp. 579–597.

Castro-Cruz, D., Gatti, F., Lopez-Caballero, F., 2021. High-fidelity broad-band prediction of regional seismic response: a hybrid coupling of Physics-Based synthetic simulation and Empirical Green's functions. Natural Hazards URL: https://link.springer.com/article/10.1007%2Fs11069-021-04766-x, doi:10.1007/s11069-021-04766-x.

CEA and CentraleSupélec and IPGP and CNRS, 2017. SEM3D Ver 2017.04 Registered at French Agency for Protection of Programs (Dépôt APP).

CEN, 2004. Eurocode 8: design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings. Technical Report. URL: https://eurocodes.jrc.ec.europa.eu/showpage.php?id=138.

Chen, Y.H., Tsai, C.C.P., 2002. A new method for estimation of the attenuation relationship with variance components. Bulletin of the Seismological Society of America 92, 1984–1991. doi:10.1785/0120010205.

Cirella, A., Piatanesi, A., Tinti, E., Cocco, M., 2008. Rupture process of the 2007 niigata-ken chuetsu-oki earthquake by non-linear joint inversion of strong motion and gps data. Geophysical Research Letters 35.

Cupillard, P., Delavaud, E., Burgos, G., Festa, G., Vilotte, J.P., Capdeville, Y., Montagner, J.P., 2012. RegSEM: a versatile code based on the spectral element method to compute seismic wave propagation at the regional scale. Geophysical Journal International 188, 1203–1220.

Douglas, J., 2021. Ground motion prediction equations 1964–2021. Technical Report. arXiv:http://www.gmpe.org.uk/gmpereport2014.pdf.

Ducellier, A., Aochi, H., 2010. Numerical simulation of the mw 6.6 niigata, japan, earthqhake: Reliable input ground motion for engineering purpose, in: 14th European Conference on Earthquake Engineering, pp. 8–p.

Faccioli, E., Maggio, F., Paolucci, R., Quarteroni, A., 1997. 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method. Journal of Seismology 1, 237–251. doi:10.1023/A:1009758820546.

Fantoni, R., Franciosi, R., 2010. Tectono-sedimentary setting of the Po Plain and Adriatic foreland. Rendiconti Lincei 21, 197–209. doi:10.1007/s12210-010-0102-4.

Festa, G., Vilotte, J.P., 2005. The Newmark scheme as velocity–stress time-staggering: an efficient PML implementation for spectral element simulations of elastodynamics. Geophysical Journal International 161, 789–812.

Fujiwara, H., Kawai, S., Aoi, S., Senna, S., Ooi, M., Matsuyama, H., Iwamoto, K., Suzuki, H., Hayakawa, Y., 2006. A subsurface structure modeling of whole of Japan for strong-motion evaluation, in: 12th Japan Earthquake Engineering Symposium, pp. 1466–1469.

García-Fernández, M., Gehl, P., Jiménez, M.J., DâĂŹAyala, D., 2019. Modelling Pan-European ground motions for seismic hazard applications. Bulletin of Earthquake Engineering 17, 2821–2840. doi:10.1007/s10518-019-00605-4.

Gatti, F., Clouteau, D., 2020. Towards blending Physics-Based numerical simulations and seismic databases using Generative Adversarial Network. Computer Methods in Applied Mechanics and Engineering 372, 113421. URL: http://www.sciencedirect.com/science/article/pii/S004578252030606X, doi:10.1016/j.cma.2020.113421.

Gatti, F., Lopez-Caballero, F., Clouteau, D., Paolucci, R., 2018a. On the effect of the 3-D regional geology on the seismic design of critical structures: the case of the Kashiwazaki-Kariwa Nuclear Power Plant. Geophysical Journal International 213, 1073–1092. doi:10.1093/gji/ggy027. doi:10.1093/gji/ggy027.

Gatti, F., Lopez-Caballero, F., Paolucci, R., Clouteau, D., 2018b. Near-source effects and non-linear site response at Kashiwazaki-Kariwa Nuclear Power Plant, in the 2007 Chuetsu-Oki earthquake: evidence from surface and downhole records and 1D numerical simulations. Bulletin of

Earthquake Engineering 16, 1105–1135. URL: https://doi.org/10.1007/s10518-017-0255-y, doi:10.1007/s10518-017-0255-y.
Gatti, F., Paludo, L.D.C., Svay, A., Cottereau, R., Clouteau, D., et al., 2017. Investigation of the earthquake ground motion coherence in heterogeneous non-linear soil deposits. Procedia engineering 199, 2354–2359. doi:10.1016/j.proeng.2017.09.232.
Gatti, F., Touhami, S., Lopez-Caballero, F., Paolucci, R., Clouteau, D., Fernandes, V.A., Kham, M., Voldoire, F., 2018c. Broad-band 3-D earthquake simulation at nuclear site by an all-embracing source-to-structure approach. Soil Dynamics and Earthquake Engineering 115, 263–280. URL: http://www.sciencedirect.com/science/article/pii/S0267726118303890, doi:10.1016/j.soildyn.2018.08.028.
Göddeke, D., Komatitsch, D., Möller, M., 2014. Finite and Spectral Element Methods on Unstructured Grids for Flow and Wave Propagation Problems, in: Numerical Computations with GPUs. Springer, pp. 183–206.
Honda, R., Aoi, S., 2009. Array back-projection imaging of the 2007 Niigataken Chuetsu-Oki earthquake striking the world's largest nuclear power plant. Bulletin of the Seismological Society of America 99, 141–147. doi:10.1785/0120080062.
Jayalakshmi, S., Dhanya, J., Raghukanth, S., Mai, P.M., 2021. Hybrid broadband ground motion simulations in the indo-gangetic basin for great himalayan earthquake scenarios. Bulletin of Earthquake Engineering , 1–30doi:10.1007/s10518-021-01094-0.
Jian, W., Fanhua, L., 2009. Prediction of oil-bearing single sandbody by 3d geological modeling combined with seismic inversion. Petroleum Exploration and Development 36, 623–627. URL: https://www.sciencedirect.com/science/article/pii/S1876380409601506, doi:https://doi.org/10.1016/S1876-3804(09)60150-6.
Kamae, K. (Ed.), 2016. Earthquakes, Tsunamis and Nuclear Risks. Springer Japan, Tokyo. URL: http://link.springer.com/10.1007/978-4-431-55822-4, doi:10.1007/978-4-431-55822-4.
Kobayashi, I., Tateishi, M., Yoshimura, T., Ueda, T., Kato, T., 1995. Geology of the Kashiwazaki District. Geological Survey of Japan.
Komatitsch, D., Vilotte, J.P., 1998. The Spectral Element Method: An Efficient Tool to Simulate the Seismic Response of 2D and 3D Geological Structures. Bulletin of the Seismological Society of America 88, 368–392.
Kristekova, M., Kristek, J., Moczo, P., 2009. Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of time signals. Geophysical Journal International 178, 813–825. doi:10.1111/j.1365-246X.2009.04177.x.
Kubo, A., Fukuyama, E., Kawai, H., Nonomura, K., 2002. NIED seismic moment tensor catalogue for regional earthquakes around Japan: quality test and application. Tectonophysics 356, 23 – 48. URL: http://www.sciencedirect.com/science/article/pii/S004019510200375X, doi:10.1016/S0040-1951(02)00375-X. seismic Source Mechanism through Moment Tensors.
Lee, E., Chen, P., Jordan, T.H., Maechling, P.B., Denolle, M.A.M., Beroza, G.C., 2014. Full 3D tomography for crustal structure in Southern California based on the scattering integral and the adjoint wavefield methods. Journal of Geophysical Research: Solid Earth 119, 6421–6451. URL: http:https://doi.org/10.1002/2014JB011346, doi:10.1002/2014JB011346.
Lee, V., Trifunac, M., Todorovska, M., Novikova, E., 1995. Empirical equations describing attenuation of peaks of strong ground motion, in terms of magnitude, distance, path effects and site conditions. Terms of Magnitude, Distance, Path Effects and Site Conditions , 95–02.
Milner, K.R., Shaw, B.E., Goulet, C.A., RichardsâĂŘDinger, K.B., Callaghan, S., Jordan, T.H., Dieterich, J.H., Field, E.H., 2021. Toward PhysicsâĂŘBased Nonergodic PSHA: A Prototype Fully Deterministic Seismic Hazard Model for Southern California. Bulletin of the Seismological Society of America 111, 898–915. doi:10.1785/0120200216, arXiv:https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/111/2/898/5258622/bssa-2020216.1.pdf.
Nakamura, Y., 1989. A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Railway Technical Research Institute, Quarterly Reports 30. URL: http://worldcat.org/oclc/3127232.
NIED, N., 2019. K-net, kik-net, national research institute for earth science and disaster resilience doi:10.17598/NIED.0004.
NIED DMC, D., . Nied hi-net, national research institute for earth science and disaster resilience doi:10.17598/NIED.0003.
Obara, K., Kasahara, K., Hori, S., Okada, Y., 2005. A densely distributed high-sensitivity seismograph network in Japan:Hi-net by National Research Institute for Earth Science and DisasterPrevention. Review of Scientific Instruments 76, 021301. URL: https://doi.org/10.1063/1.1854197, doi:10.1063/1.1854197, arXiv:https://doi.org/10.1063/1.1854197.
Okada, Y., Kasahara, K., Hori, S., Obara, K., Sekiguchi, S., Fujiwara, H., Yamamoto, A., 2004. Recent progress of seismic observation networks in Japan Hi-net, F-net, K-NET and KiK-net. Earth, Planets and Space 56, xv–xxviii. doi:10.1186/BF03353076.
Olsen, K.B., Nigbor, R., Konno, T., 2000. 3D Viscoelastic Wave Propagation in the Upper Borrego Valley, California, Constrained by Borehole and Surface Data. Bulletin of the Seismological Society of America 90, 134–150. URL: https://doi.org/10.1785/0119990052, doi:10.1785/0119990052, arXiv:https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/90/1/134/2710460/134.pdf.
Olsen, K.B., Pechmann, J.C., Schuster, G.T., 1995. Simulation of 3D elastic wave propagation in the Salt Lake Basin. Bulletin of the Seismological Society of America 85, 1688–1710. arXiv:https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/85/6/1688/2708590/BSSA0850061688.pdf.
Paolucci, R., Aimar, M., Ciancimino, A., Dotti, M., Foti, S., Lanzano, G., Mattevi, P., Pacor, F., Vanini, M., 2021. Checking the site categorization criteria and amplification factors of the 2021 draft of Eurocode 8 Part 1âĂŞ1. Bulletin of Earthquake Engineering URL: https://link.springer.com/10.1007/s10518-021-01118-9, doi:10.1007/s10518-021-01118-9.
Paolucci, R., Gatti, F., Infantino, M., Ozcebe, A.G., Smerzini, C., Stupazzini, M., 2018. Broad-band ground motions from 3D physics-based numerical simulations using Artificial Neural Networks. Bulletin of the Seismological Society of America 108, 1272–1286. doi:10.1785/0120170293.
Régnier, J., Bonilla, L.F., Bard, P.Y., Bertrand, E., Hollender, F., Kawase, H., Sicilia, D., Arduino, P., Amorosi, A., Asimaki, D., Boldini, D., Chen, L., Chiaradonna, A., De Martin, F., Ebrille, M., Elgamal, A., Falcone, G., Foerster, E., Foti, S., Garini, E., Gazetas, G., Gélis, C., Ghofrani, A., Giannakou, A., Gingery, J., Glinsky, N., Harmon, J., Hashash, Y., Iai, S., Jeremić, B., Kramer, S., Kontoe, S., Kristek, J., Lanzo, G., di Lernia, A., Lopez-Caballero, F., Marot, M., McAllister, G., Diego M., E., Moczo, P., Montoya-Noguera, S., Musgrove, M., Nieto-Ferro, A., Pagliaroli, A., Pisanó, F., Richterova, A., Sajana, S., Santisi d'Avila, M., Shi, J., Silvestri, F., Taiebat, M., Tropeano, G., Verrucci, L., Watanabe, K., 2016. International Benchmark on Numerical Simulations for 1D, Nonlinear Site Response (PRENOLIN): Verification Phase Based on Canonical Cases. Bulletin of the Seismological Society of America 106, 2112–2135. URL: http://www.bssaonline.org/content/106/5/2112.abstract, doi:10.1785/0120150284, arXiv:http://www.bssaonline.org/content/106/5/2112.full.pdf+html.

Ripperger, J., Mai, P.M., Ampuero, J.P., 2008. Variability of Near-Field Ground Motion from Dynamic Earthquake Rupture Simulations. Bulletin of the Seismological Society of America 98, 1207–1228. URL: https://doi.org/10.1785/0120070076, doi:10.1785/0120070076, arXiv:https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/98/3/1207/3670862/1207.pdf.

Rodriguez-Marek, A., Cotton, F., Abrahamson, N.A., Akkar, S., Al Atik, L., Edwards, B., Montalva, G.A., Dawood, H.M., 2013. A Model for Single-Station Standard Deviation Using Data from Various Tectonic Regions. Bulletin of the Seismological Society of America 103, 3149–3163. URL: https://doi.org/10.1785/0120130030, doi:10.1785/0120130030.

Sekiguchi, H., Yoshimi, M., Horikawa, H., Yoshida, K., Suzuki, H., Matsuyama, H., Morino, M., Takizawa, F., Ying, L., 2009. 3d subsurface structure model of the niigata sedimentary basin. Geological Survey of JapanâĂŤAIST annual report on active fault and paleoearthquake researches 9, 175–259.

Shang, X., TkalÄĐiÄĞ, H., 2020. Point-source inversion of small and moderate earthquakes from p-wave polarities and p/s amplitude ratios within a hierarchical bayesian framework: Implications for the geysers earthquakes. Journal of Geophysical Research: Solid Earth 125, e2019JB018492. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019JB018492, doi:https://doi.org/10.1029/2019JB018492, arXiv:https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019JB018492.

Smerzini, C., Paolucci, R., Stupazzini, M., 2011. Comparison of 3D, 2D and 1D numerical approaches to predict long period earthquake ground motion in the Gubbio plain, Central Italy. Bulletin of Earthquake Engineering 9, 2007–2029. doi:10.1007/s10518-011-9289-8.

Strasser, F.O., Abrahamson, N.A., Bommer, J.J., 2009. Sigma: Issues, Insights, and Challenges. Seismological Research Letters 80, 40–56. URL: https://doi.org/10.1785/gssrl.80.1.40, doi:10.1785/gssrl.80.1.40, arXiv:https://pubs.geoscienceworld.org/ssa/srl/article-pdf/80/1/40/2760588/040.pdf.

TEPCO, 2007. The data analysis recorded at the Kashiwazaki Kariwa Nuclear Power Plant during the 2007 Niigata-ken Chuetsu-oki earthquake. Technical Report. The Tokyo Electric Power Company, Inc. URL: http://www.tepco.co.jp/cc/press/betu07_j/images/070730d.pdf. in Japanese.

Tokumitsu, R., Kikuchi, M., Nishimura, I., Shiba, Y., Tanaka, S., 2009. Analysis of the strong motion records obtained from the 2007 Niigataken Chuetsu-Oki earthquake and determination of the design basis ground motions at the Kashiwazaki Kariwa Nuclear Power Plant. Part 1. Outline of the strong motion records and estimation of factors in large amplification. Technical Report.

Tsuda, K., Hayakawa, T., Uetake, T., Hikima, K., Tokimitsu, R., Nagumo, H., Shiba, Y., 2011. Modeling 3D Velocity Structure in the Fault Region of the 2007 Niigataken Chuetu-Oki Earthquake with Folding Structure, in: 4th IASPEI/IAEE International Symposium-Effects of Surface Geology on Seismic Motion, pp. 1–11.

Vyas, J.C., Mai, P.M., Galis, M., 2016. Distance and Azimuthal Dependence of GroundâĂŘMotion Variability for Unilateral StrikeâĂŘSlip Ruptures. Bulletin of the Seismological Society of America 106, 1584–1599. URL: https://doi.org/10.1785/0120150298, doi:10.1785/0120150298, arXiv:https://pubs.geoscienceworld.org/ssa/bssa/article-pdf/106/4/1584/2646281/1584.pdf.

Watanabe, T., Moroi, T., Nagano, M., Tokumitsu, R., Kikuchi, M., Nishimura, I., 2009. Analysis of the strong motion records obtained from the 2007 Niigataken Chuetsu-Oki earthquake and determination of the design basis ground motions at the Kashiwazaki Kariwa Nuclear Power Plant. Part 2. Difference of site amplification based on the 2D FEM analysis of the folded structure. Technical Report.