Earth, Planets and Space

FULL PAPER Open Access

Predicted results of weak and strong ground motions at the target site of the blind prediction exercise as steps-2 and -3, Report for the experiments for the 6th international symposium on effects of surface geology on seismic motion

Seiji Tsuno[1*], Fumiaki Nagashima[2], Hiroshi Kawase[2], Hiroaki Yamanaka[3] and Shinichi Matsushima[2]

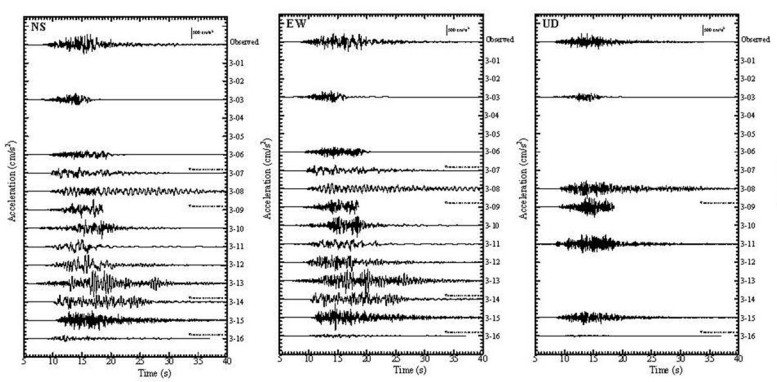
Abstract

In this study, we compared observations and predictions submitted by participants for blind prediction experiments for ground motions using aftershocks, foreshock, and mainshock of the 2016 Kumamoto earthquake sequence, Japan, to improve our understanding of the quality of state-of-the-art methods on the reproducibility of the effects of surface geology on seismic motions. In the blind predictions, 1D, 2D, and 3D methods, Green's function methods, spectral ratio approaches, and other approaches were applied. As for PGA/PGV, acceleration/velocity duration, Fourier spectrum, pseudo-velocity response spectrum, and site amplification factors, the observed values are mostly within the range of average $\pm\sigma$ of all the predictions in the case of weak and strong ground motions. The results of the mean absolute percentage errors for these indices show that the applied methods can predict weak and strong ground motions for the three components in the range of one-half to twice the observations. The average goodness-of-fit (GOF) scores for weak and strong ground motions indicate either a very good fit (6.5–8) or a good fit (4.5–6.5) for the three components. Finally, examples of the categorized methods are quite limited; however, results indicate that the predictions by all the categorized methods can adequately reproduce weak and strong ground motions within either a very good or good fit. Although we could not find a significant difference in the results from the categorized methods, scores by the 2D and 3D methods in the frequency range of 0.5–1 and 1–2 Hz for all the blind predictions are higher than the scores by the other methods. The GOF score for the part after the S-wave by the 2D and 3D methods is higher than that by the 1D method. This supports that the predictions by the 2D and 3D methods due to the accounting of the proper geometry could reproduce the basin-induced and/or basin-transduced surface waves excited by the basin-edge effect more than the 1D method using the earthquake record observed at the reference site.

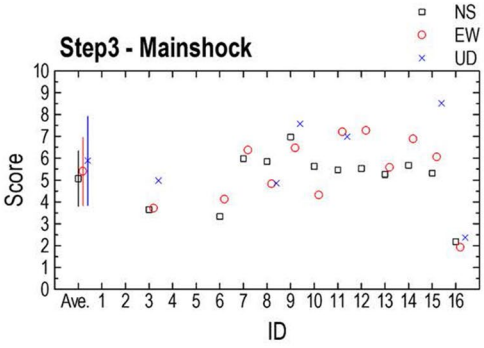
Keywords Weak motion, Strong motion, Blind prediction, Validation, MAPE, GOF, Basin-edge effect, 2016 Kumamoto earthquake, Kumamoto Plain

*Correspondence:
Seiji Tsuno
tsuno.seiji.75@rtri.or.jp
Full list of author information is available at the end of the article

Graphical Abstract



Accelerations at the target site for the 2016 Kumamoto earthquake (*Mj* 7.3) by the 12 predictions including all the methods



GOF scores for the 2016 Kumamoto earthquake (Score: 8–10 excellent fit, 6.5–8 very good fit, 4.5–6.5 good fit, 3.5–4.5 fair fit, and 0–3.5 poor fit)

Introduction

To improve our understanding of the quality of state-of-the-art methods for the reproducibility of the effects of surface geology on seismic motion, we performed a blind prediction exercise at the 6th IASPEI/IAEE International Symposium: Effects of Surface Geology on Seismic Motion (ESG6). In the ESG6, based on the theme "Progress of ESG research during the last three decades: How accurately can we predict site amplification?", the ESG6 local organizing committee prepared blind prediction exercises consisting of three steps (step-1: identification of the subsurface structure at the target site, step-2: simulation of weak ground motions observed at the target site, step-3: simulation of strong ground motions observed at the target site) (Chimoto et al. 2023; Matsuhima et al. 2023). The experiments, with their aims, of past ESG and ESG6 are well-described in Matsushima et al. (2023).

As the target site, we selected a seismic station in the Kumamoto Plain (Ishizaka et al. 1995) that recorded strong ground motions during the 2016 Kumamoto earthquake sequence in Japan, which has not been yet published. In blind predictions of step-2 (BP2) and step-3 (BP3), the aftershocks, foreshock (*Mj* 6.5), and mainshock (*Mj* 7.3) of the 2016 Kumamoto earthquake sequence (e.g., Asano and Iwata 2016) were selected as the target earthquakes to predict weak and strong ground motions at the Kumamoto test site. Earthquake ground motions were categorized into weak and strong ground motions with observed PGAs of less than 50 cm/s^2 and over 50 cm/s^2, respectively.

In this study, we performed quantitative validation analyses using the mean absolute percentage error (MAPE) (Morley et al. 2018) and goodness-of-fit (GOF) (Anderson 2004) between observations and predictions for weak and strong motions using the aftershocks, foreshock, and mainshock of the 2016 Kumamoto earthquake sequence.

Outline of blind predictions and submitted waveforms

Outline of blind predictions

The blind predictions of step-2 (BP2) and step-3 (BP3) were simulations of the weak and strong ground motions observed at the Kumamoto test site, respectively (Matsushima et al. 2023). The earthquake records observed at the reference site for the target earthquakes (BP2: *Mj* 5.9 occurred on April 16, 2016 at 3:03 JST, BP3: foreshock of *Mj* 6.5 occurred on April 14, 2016 at 21:26 JST, and mainshock of *Mj* 7.3 occurred on April 16, 2016 at 1:25 JST) were released. However, the earthquake records observed at the test site for the target earthquakes were blindfolded. Earthquake ground motions for the target earthquakes at the test site, predicted by simulation techniques such as the 1D method, GMPE, empirical Green's function method, and 2D/3D simulations were requested.

Strong ground motions during the 2016 Kumamoto earthquake sequence and weak ground motions during the aftershocks were recorded at a reference site (KU.KMP1). However, the records following the arrival of wave packets with the maximum amplitude for the mainshock were lost because of technical issues with the backup battery (Tsuno et al. 2017). The test site at which the strong ground motion station (KUMA) has been installed and maintained by the Kyushu Railway Company is located on the Quaternary layers in the northern part of the Kumamoto Plain, Japan (Ishizaka et al. 1995; Chimoto et al. 2023; Matsushima et al. 2023). Inversely, the

reference site in which the strong ground motion station (KU.KMP1) has been installed and maintained by the Institute of Seismology and Volcanology, Faculty of Science, Kyushu University, is located on Mt. Kinbo, where andesite is widely distributed (Hoshizumi et al. 2004). The locations of the KUMA and KU.KMP1 are shown in Fig. 1. The accelerations of the target earthquakes of BP2 and BP3 at KU.KMP1 and KUMA are shown in Figs. 2 and 3, respectively. Information on the foreshock, mainshock (BP3), and aftershock (BP2) of the 2016 Kumamoto earthquake sequence with the epicentral distance, fault distance, observed PGA, and PGV at KU.KMP1 and KUMA are shown in Table 1. Please note that the waveforms in Fig. 3 and the information in Table 1 were released to the participants only after the blind prediction results were collected.

Matsushima et al. (2023) reported that the surface geology of KUMA consists of alluvium forming the flood plain of the Shirakawa River and a modern landfill, as borehole surveys to a depth of 39 m conducted at the site indicated. The preferred model was simply integrated with both the models of PS logging data (Matsushima et al. 2023) for a shallow profile down to the layer with an S-wave velocity (Vs) of 290 m/s and an array microtremor survey (Senna et al. 2018) for a deep profile in and around this area. In particular, gravel with Vs 95 m/s in the 1st layer of the subsurface structure was confirmed by boring and PS logging. We provided the results of the in situ measurements and laboratory tests, the preferred 1D velocity model by the ESG6 local organizing committee (Matsushima et al. 2023), information from the Japan integrated velocity structure model (JIVSM: Koketsu et al. 2009, 2012), and a geological map of the Kumamoto test site by the National Institute of Advanced Industrial Science and Technology (AIST). For example, laboratory tests were conducted on five samples obtained at different

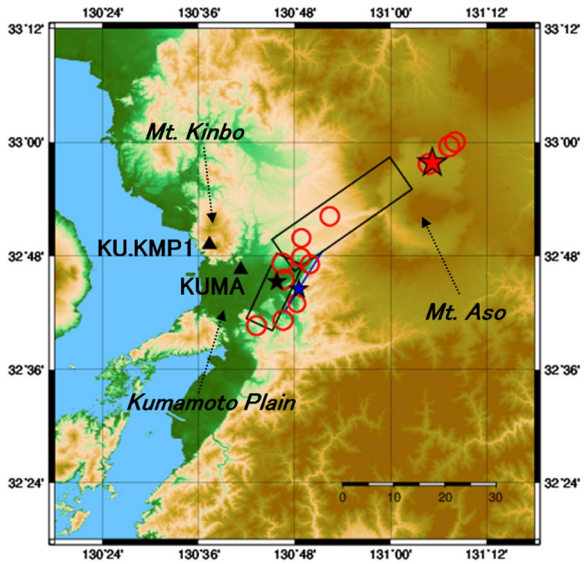


Fig. 1 Locations of strong ground motion stations and epicenters of earthquakes provided in the blind predictions (BP2 and BP3). Circles denote epicenters of the aftershocks of the 2016 Kumamoto earthquake sequence. A red star denotes an epicenter of the target earthquake in BP2. Black stars denote epicenters of the foreshock and mainshock of the 2016 Kumamoto earthquake sequence that are the target earthquakes in BP3. Source fault planes of the foreshock and mainshock by Asano and Iwata (2016) are projected

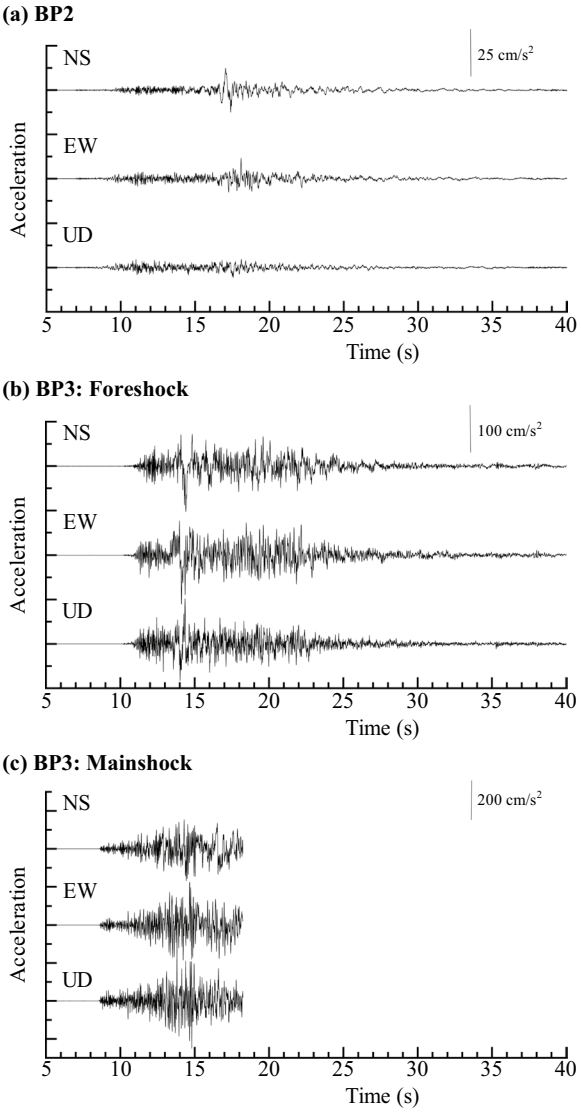


Fig. 2 **a** Accelerations of the three components at KU.KMP1 for the largest aftershock of *Mj* 5.9. Waveforms start on April 16, 2016, at 03:03:10 JST. **b** Accelerations of the three components at KU.KMP1 for the foreshock of *Mj* 6.5. Waveforms start on April 14, 2016, at 21:26:30 JST. **c** Accelerations of the three components at KU.KMP1 for the mainshock of *Mj* 7.3 before technical issues with the back-up battery. Waveforms start on April 16, 2016 at 01:25:00 JST

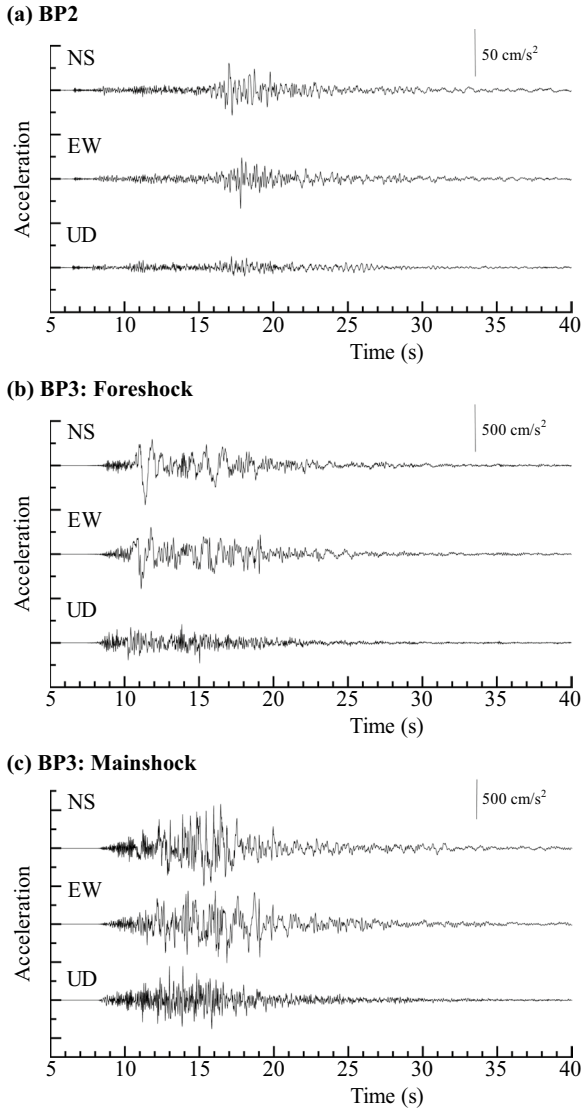


Fig. 3 **a** Accelerations of the three components at KUMA for the largest aftershock of *Mj* 5.9. **b** Accelerations of the three components at KUMA for the foreshock of *Mj* 6.5. **c** Accelerations of the three components at KUMA for the mainshock of *Mj* 7.3

depths at KUMA based on the Japanese Industrial Standard (JIS) and the Japanese Geotechnical Society Standard (2015, 2017). The particle size, specimen conditions, and test results of the cyclic triaxial test such as G-γ and h-γ relation of the samples were opened to all participants before the exercises (Matsushima et al. 2023).

We simply required to simulate earthquake ground motions for the target earthquakes at KUMA, by any simulation techniques, such as the 1D method, GMPE, empirical Green's function method, and 2D/3D simulation. As the minimum requirement of the results, acceleration data for the horizontal component and the Fourier spectrum in the reliable frequency range were requested, along with the starting time and duration of the simulation. Submission of two- or three-component data was optional but preferred. In the cases of theoretical methods and non-linear analysis, we requested a description of the structural model, a non-linear parameter used, and an estimated shear strain.

Methods applied by participants

In the blind prediction of step-2 (BP2), 18 predictions were submitted by 15 teams from five countries, as shown in Table 2. Ten teams from BP2 participated in BP1 (Chimoto et al. 2023). As methods applied by the participants, there were the 1D theoretical simulation methods in 4 predictions, 2D theoretical simulation methods in 2 predictions, empirical/stochastic Green's function methods in 4 predictions, spectral ratio approaches in 3 predictions, 3D theoretical simulation method in 1 prediction, and other approaches in 5 predictions. The nonlinearity of soft soil for the shallow subsurface structure was considered in three predictions, including the equivalent–linear method. Ten predictions were obtained by considering the 1D subsurface structure at KUMA. Most participants used their own models in BP1 or models estimated using earthquake records. Only three participants used the preferred model offered by the ESG6 local organizing Committee. In contrast, eight predictions were obtained without any information on the subsurface structures, in which spectral ratio approaches or empirical/stochastic Green's function methods were applied. The reliable frequency ranges are mostly 0.1 to 10 Hz among the 18 predictions in BP2.

In the blind prediction of step-3 (BP3), 16 predictions were submitted by 14 teams from five countries, as shown in Table 3. Eight teams from BP3 participated in BP1. As the methods applied by the participants, there were the 1D theoretical simulation methods in 8 predictions, 2D theoretical simulation methods in 2 predictions, empirical/stochastic Green's function methods in 3 predictions, spectral ratio approaches in 2 predictions, 3D theoretical simulation method in 1 prediction, and other approaches in 3 predictions. The nonlinearity of soft soil for the shallow subsurface structure was considered in the 15 predictions, leaving out one prediction. Thirteen predictions were obtained by considering the 1D subsurface structure at KUMA. Seven participants used their own models in BP1 or models estimated using earthquake records. Six participants used the preferred model and PS logging offered by the ESG6 local organizing committee. However, three predictions were obtained without any information on the subsurface structures, in which the spectral ratio approaches and the empirical Green's function method were applied. The reliable frequency ranges are mostly

Table 1 Information of the foreshock, mainshock (BP3), and aftershock (BP2) of the 2016 Kumamoto earthquake sequence with the epicentral distance, fault distance, observed PGA and PGV, at KU.KMP1 and KUMA

Target earthquake	Origin Time (JST)	Lat	Long	Depth (km)	*Mj*	KU.KMP1							KUMA						
						Epicentral Dis. (km)	Fault Dis. (km)	PGA (cm/s^2)			PGV (cm/s)			Epicentral Dis. (km)	Fault Dis. (km)	PGA (cm/s^2)			PGV (cm/s)		
								NS	EW	UD	NS	EW	UD			NS	EW	UD	NS	EW	UD
Foreshock	2016-04-14 21:26:34.43	32.74	130.81	11.39	6.5	19.45	19.70	102.4	112.8	100.2	9.4	11.8	4.3	11.92	12.09	441.5	388.8	227.0	69.2	46.4	12.9
Mainshock	2016-04-16 01:25:05.47	32.75	130.76	12.45	7.3	15.00	16.78	168.2	231.5	249.1	20.8	25.0	8.0	7.41	10.47	574.7	509.1	444.0	66.8	52.2	18.5
Aftershock	2016-04-16 03:03:10.78	32.96	131.09	6.89	5.9	46.26	–	12.8	11.5	5.7	1.0	0.7	0.4	42.76	—	30.6	33.2	12.6	1.97	1.89	0.63

PGVs at Kuma were obtained by velocities in the frequency range of 0.1 to 10 Hz

Table 2 Summary of methods applied in the blind prediction of step-2

Team	Prediction	Methods	Nonlinearity	Subsurface structure	Reliable frequency range (Hz)
1	1	2D-FDM	Yes	1D (Preferred model) + 2D (JIVSM)	0.1–10
	2	2D-FDM	Yes	1D (Preferred model) + 2D (JIVSM)	0.1–10
2	3	Equivalent–linear (1D)	Yes	1D (Own model in BP1)	0.1–20
3	4	1D	No	1D (Own model in BP1)	0.1–25
4	5	Hybrid (1D + 3D-FDM)	No	1D (Own model in BP1) + 3D (JIVSM)	0.1–20
5	6	Stochastic Green's Function	No	1D (Own estimated model)	0.5–10
6	7	Stochastic Green's Function	No	—	0.2–10
7	8	Stochastic approach	No	1D (Own model in BP1)	0.25–25
8	9	Empirical Green's Function	No	—	0.1–10
9	10	Generalized Standard Spectral ratio	No	—	0.5–20
	11	Standard Spectral ratio	No	—	0.5–30
10	12	Semi-empirical approach	No	—	0.2–10
11	13	Deffuse Field concept	No	1D (Own estimated model)	0.1–20
12	14	Site effect substitution	No	—	0.1–10
13	15	Physics-based simulation	No	1D (Own estimated model)	0.01–50
14	16	1D	No	1D (Preferred model)	0.1–10
15	17	Empirical Green's Function	No	—	0.1–20
	18	Spectral ratio	No	—	0.1–20

Table 3 Summary of methods applied in the blind prediction of step-3

Team	Prediction	Methods	Earthquake	Nonlinearity	Subsurface structure	Reliable frequency range
1	1	2D-FDM	Foreshock	Yes	1D (Preferred model) + 2D (JIVSM)	0.1–10 Hz
	2	2D-FDM	Foreshock	Yes	1D (Preferred model) + 2D (JIVSM)	0.1–10 Hz
2	3	Empirical non-linear transfer function	Foreshock/mainshock	Yes	–	Whole frequency range
3	4	Equivalent–linear (1D)	Foreshock	Yes	1D (Own model in BP1)	0.1–20 Hz
	5	Nonlinear (1D)	Foreshock	Yes	1D (Own model in BP1)	0.1–30 Hz
4	6	Nonlinear (1D)	Foreshock/mainshock	Yes	1D (Own model in BP1)	0.1–25 Hz
5	7	Stochastic approach	Foreshock/mainshock	Yes	1D (Own model in BP1)	0.25–25 Hz
6	8	Hybrid (Nonlinear 1D + EGF + DWN)	Foreshock/mainshock	Yes	1D (PS logging and own estimated model)	0.05–10 Hz
7	9	Empirical non-linear transfer function	Foreshock/mainshock	Yes	–	Whole frequency range
8	10	Empirical Green's Function	Foreshock/mainshock	No	–	0.2–10 Hz
9	11	Effective stress analysis (1D)	Foreshock/mainshock	Yes	1D (PS logging)	0.1–10 Hz
10	12	Hybrid (1D + 3D-FDM)	Foreshock/mainshock	Yes	1D (Own model in BP1) + 3D (JIVSM)	0.1–20 Hz
11	13	Equivalent–linear (1D)	Mainshock	Yes	1D (Preferred model)	up to 20 Hz
12	14	Hybrid (Stochastic Green's Function + Equivalent–linear 1D	Foreshock/mainshock	Yes	1D (Preferred model)	0.5–10 Hz
13	15	Deffuse Field concept	Foreshock/mainshock	Yes	1D (Own estimated model)	0.1–20 Hz
14	16	Physics-based simulation	Foreshock/mainshock	Yes	1D (Own estimated model)	0.01–50 Hz

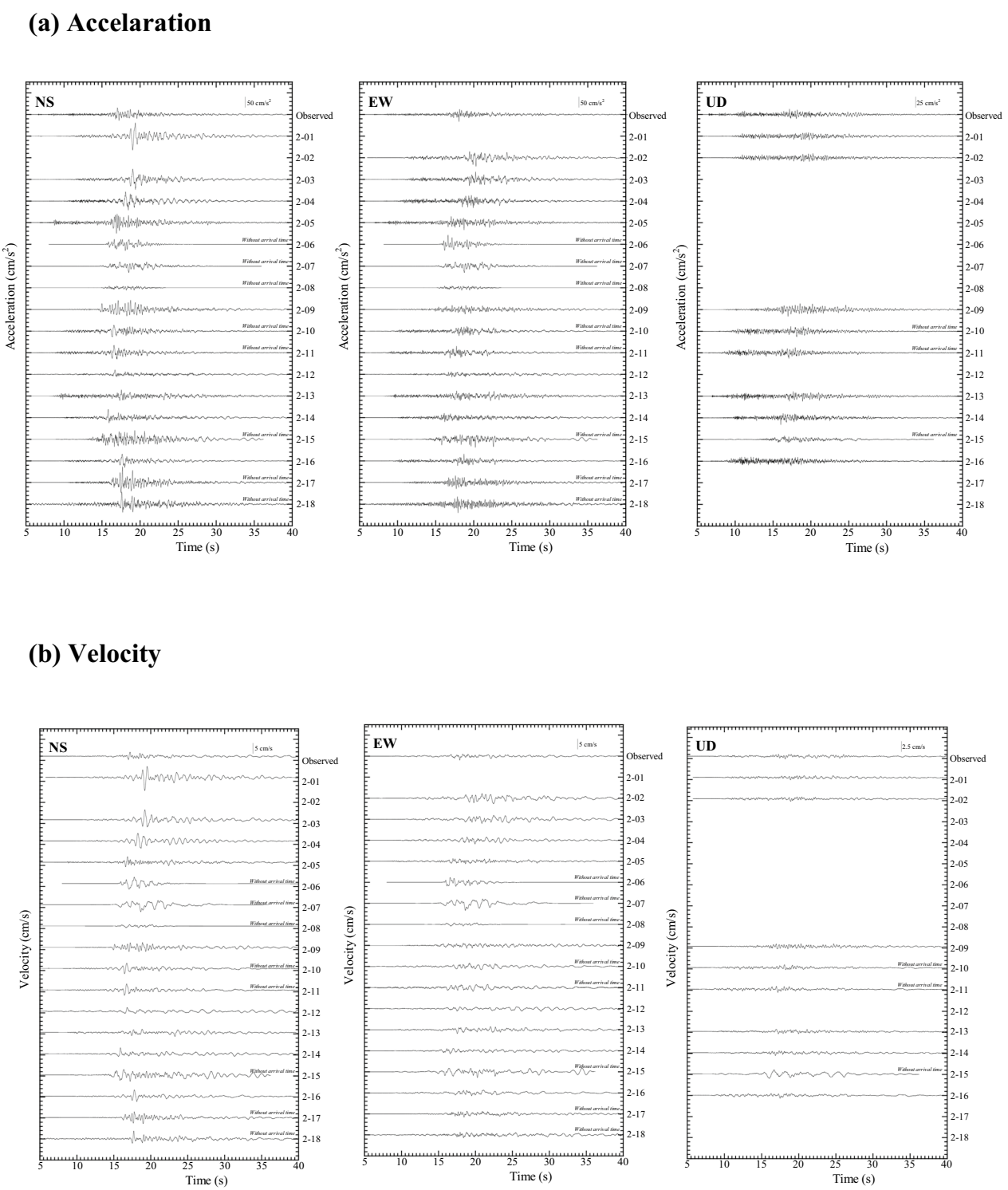


Fig. 4 **a** Accelerations of the three components at KUMA for the largest aftershock of *Mj* 5.9. Acceleration observed is also shown. Waveforms start on April 16, 2016 at 03:03:10 JST. **b** Velocities of the three components in the frequency range of 0.1 to 10 Hz at KUMA by integrating the accelerations. Velocity observed is also shown

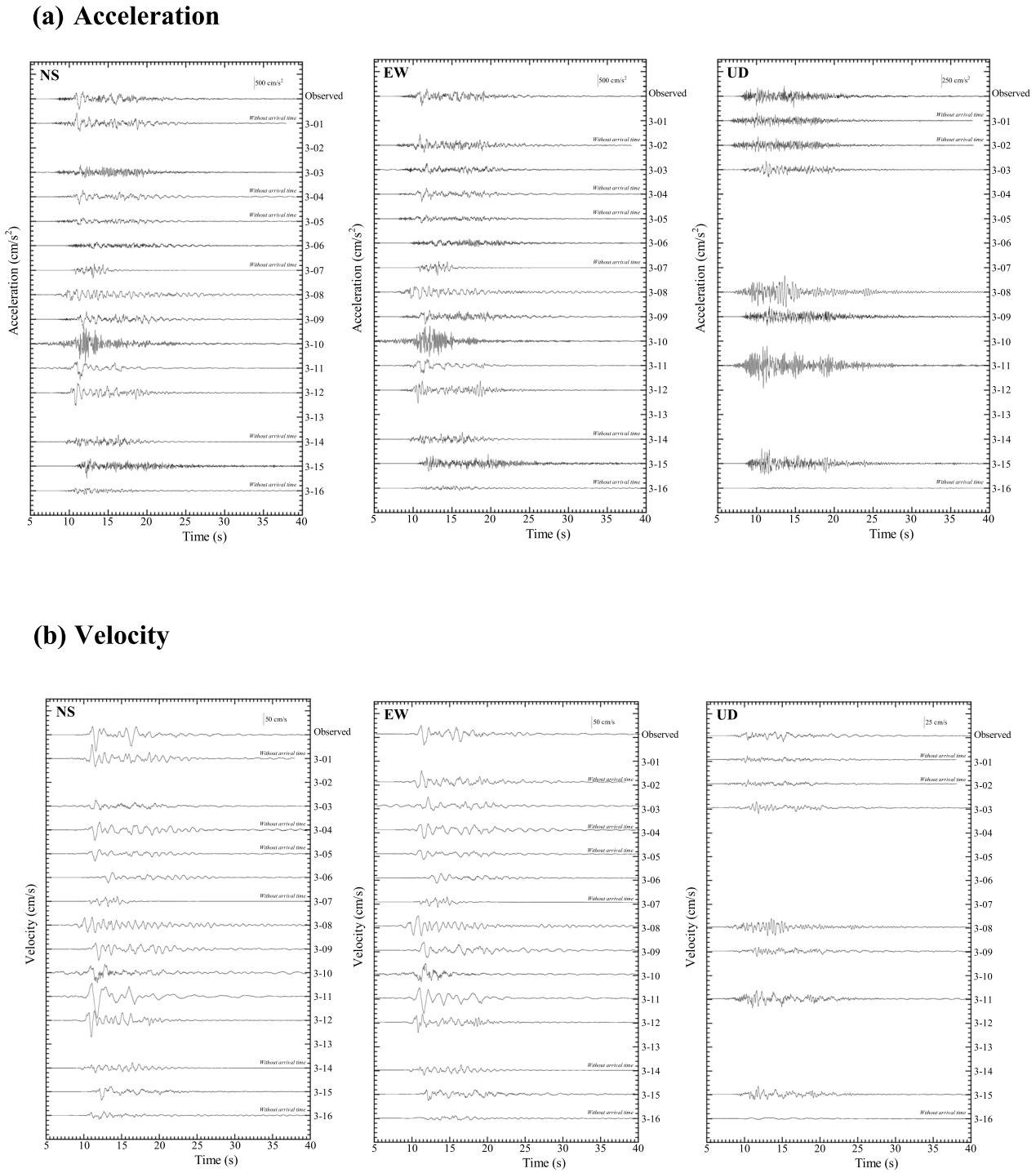


Fig. 5 **a** Accelerations of the three components at KUMA for the foreshock of *Mj* 6.5. Acceleration observed is also shown. Waveforms start on April 14, 2016 at 21:26:30 JST. **b** Velocities of the three components in the frequency range of 0.1 to 10 Hz at KUMA by integrating the accelerations. Velocity observed is also shown

0.1 to 10 Hz among the 16 predictions in BP3. As for the 2D/3D simulation in BP2 and BP3, the Japan integrated velocity structure model (JIVSM) (Koketsu et al. 2009, 2012) was used in the Kumamoto Plain, replacing the structural model around KUMA with the preferred model by the ESG6 local organizing committee (Matsushima et al. 2023) and the own model in BP1, respectively. For the 2D simulation, the earthquake record at KU.KMP1 was

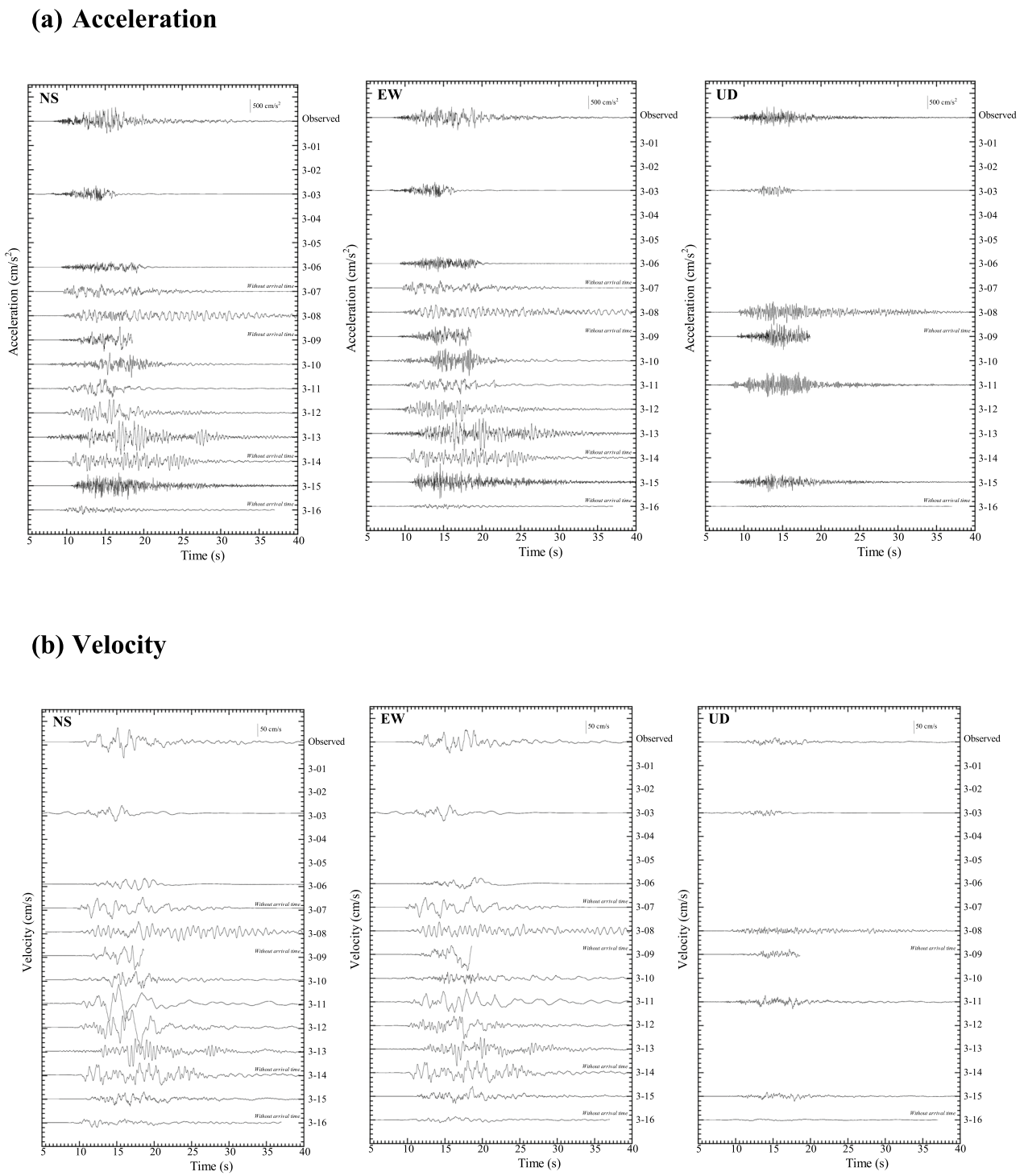


Fig. 6 **a** Accelerations of the three components at KUMA for the mainshock of *Mj* 7.3. Acceleration observed is also shown. Waveforms start on April 16, 2016 at 01:25:00 JST. **b** Velocities of the three components in the frequency range of 0.1 to 10 Hz at KUMA by integrating the accelerations. Velocity observed is also shown

used as an input ground motion at the bedrock. On the other hand, for the 3D simulation, the finite fault models of the foreshock and mainshock including slip velocities estimated by the kinematic waveform inversion (Asano and Iwata 2016) were used. Note that there is basically no information on the input motions, because there was no

request for the input motions of the participants in BP2 and BP3.

Submitted waveforms

The accelerations of the three components at KUMA for the largest aftershock of *Mj* 5.9 (the target earthquake of BP2) by the 18 predictions, foreshock of *Mj* 6.5 (the target earthquake of BP3) by the 15 predictions, and mainshock of *Mj* 7.3 (the target earthquake of BP3) by the 12 predictions are shown with the observed accelerations in Figs. 4, 5, and 6a, respectively. In these figures, the waveforms started on April 16, 2016 at 03:03:10 JST, April 14, 2016 at 21:26:30 JST, and April 16, 2016 at 01:25:00 JST. Several predictions did not fix arrival times. Several teams predicted the acceleration of the UD component. In Figs. 4, 5, and 6b, the velocities of the three components in the frequency range of 0.1 to 10 Hz at KUMA by integrating the accelerations are shown with the observed velocity.

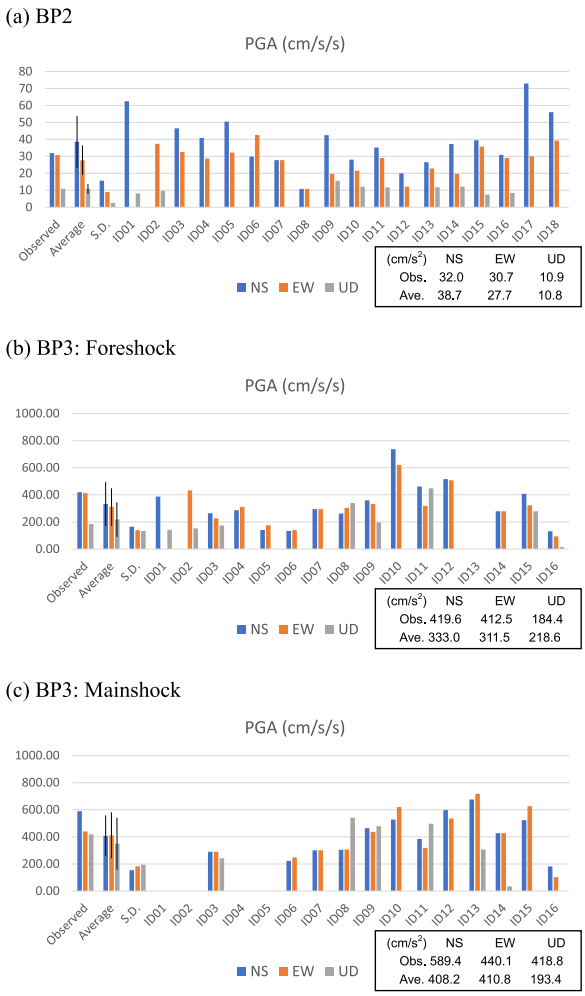


Fig. 7 PGAs of the three components in the frequency range of 0.1 to 10 Hz at KUMA for predictions with the average and the standard deviation. A bar indicates the average ± the deviation. PGA observed is also shown. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

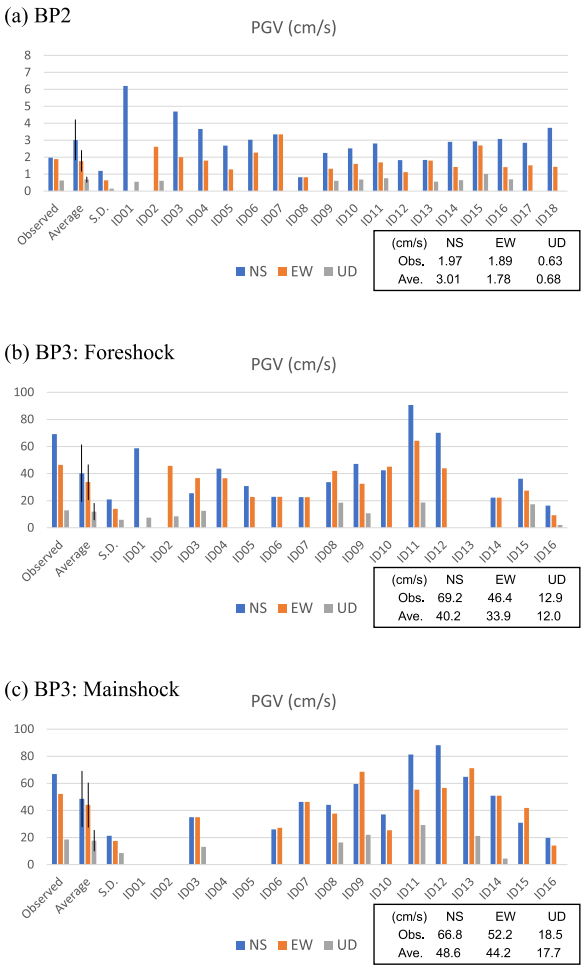


Fig. 8 PGVs of the three components in the frequency range of 0.1 to 10 Hz at KUMA for predictions with the average and standard deviation. A bar indicates the average ± the deviation. PGV observed is also shown. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

Results

We summarize the results of the submitted waveforms in terms of peak ground acceleration (PGA), peak ground velocity (PGV), Fourier spectrum, velocity response spectrum with 5% damping, duration, and site amplification factor without any interpretation or discussion in this section. We calculated the arithmetic average and standard deviation for the PGA, PGV, and duration for comparison on a linear scale. We also calculated the geometric average and standard deviation of the Fourier spectrum, velocity response spectrum, and site amplification factor on a logarithmic scale.

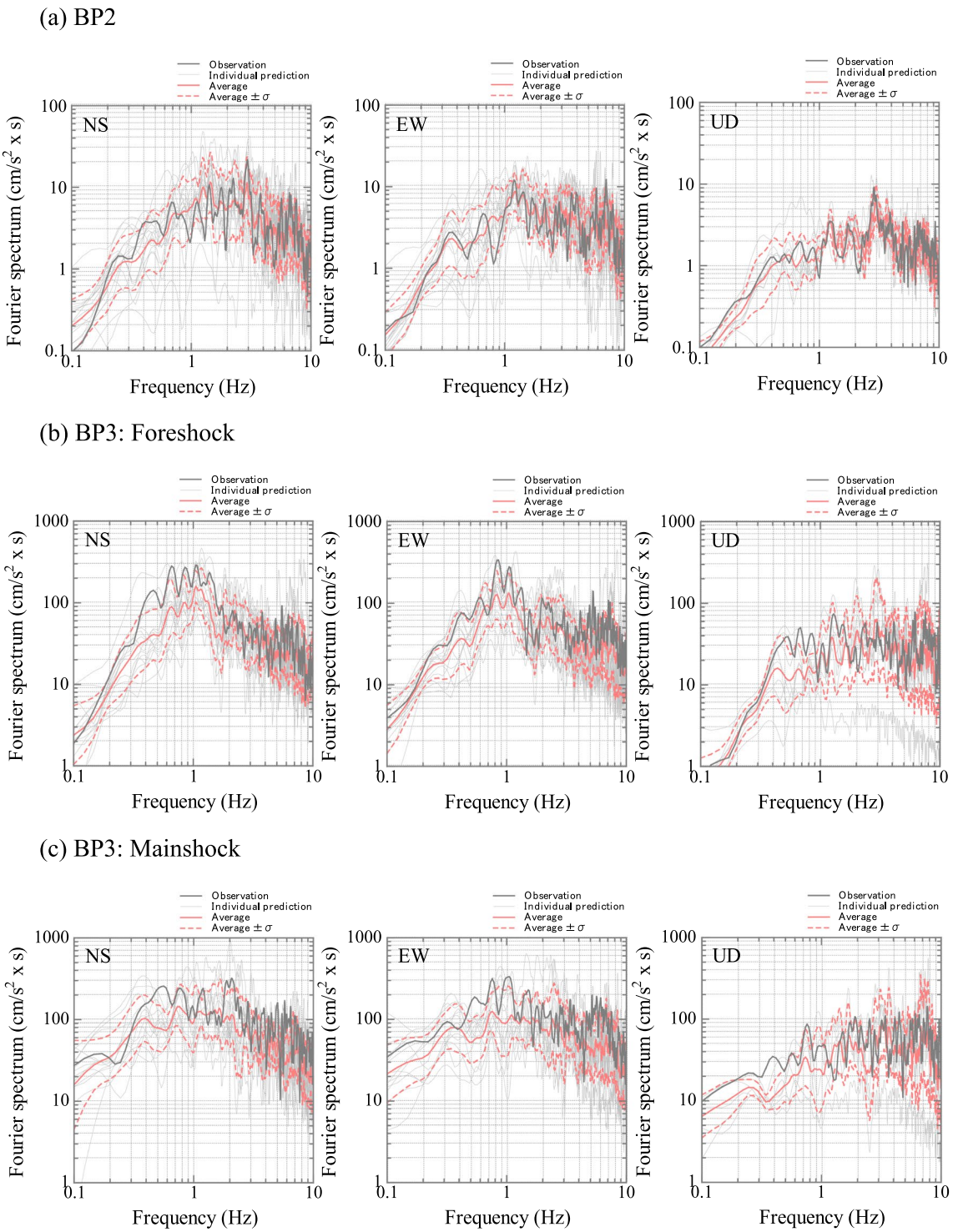


Fig. 9 Fourier spectra of the three components at KUMA for predictions with the average and the standard deviation. Fourier spectrum observed is also shown. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

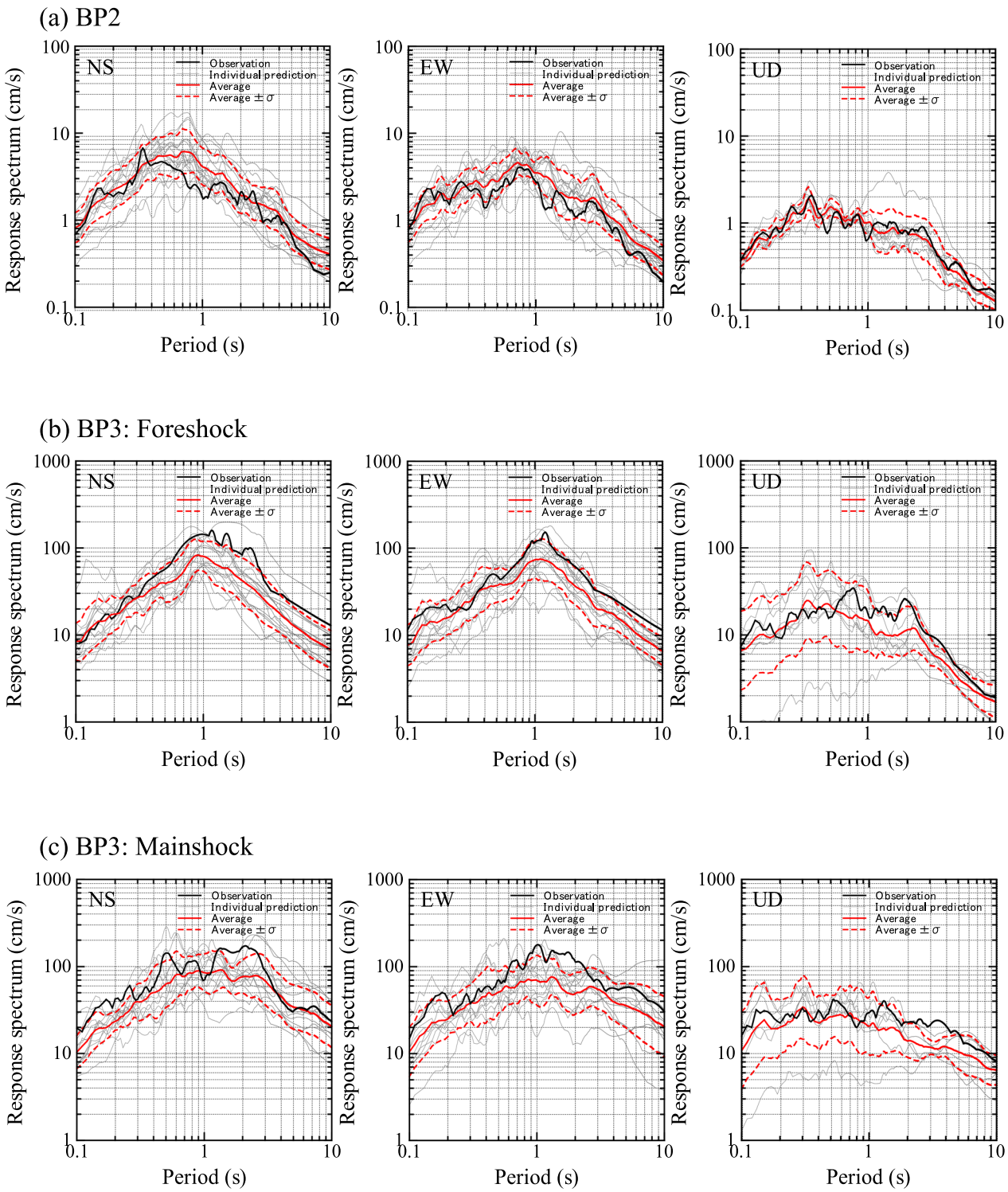


Fig. 10 Pseudo-velocity response spectra ($h = 0.05$) of the three components at KUMA for predictions with the average and the standard deviation. Pseudo-velocity response spectrum ($h = 0.05$) observed is also shown. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

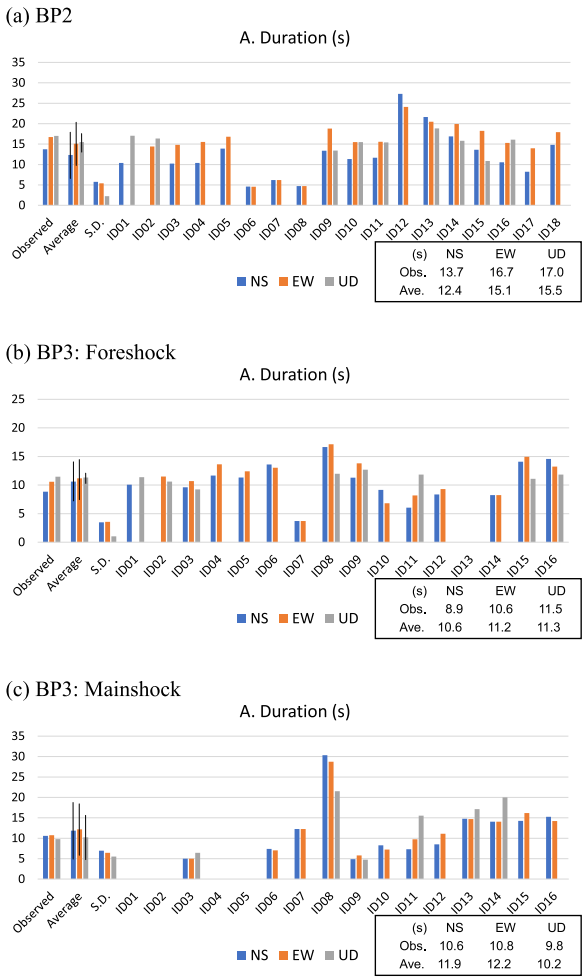


Fig. 11 Acceleration durations of the three components in the frequency range of 0.1 to 10 Hz at KUMA for predictions with the average and the standard deviation. A bar indicates the average ± the deviation. Acceleration duration observed is also shown. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

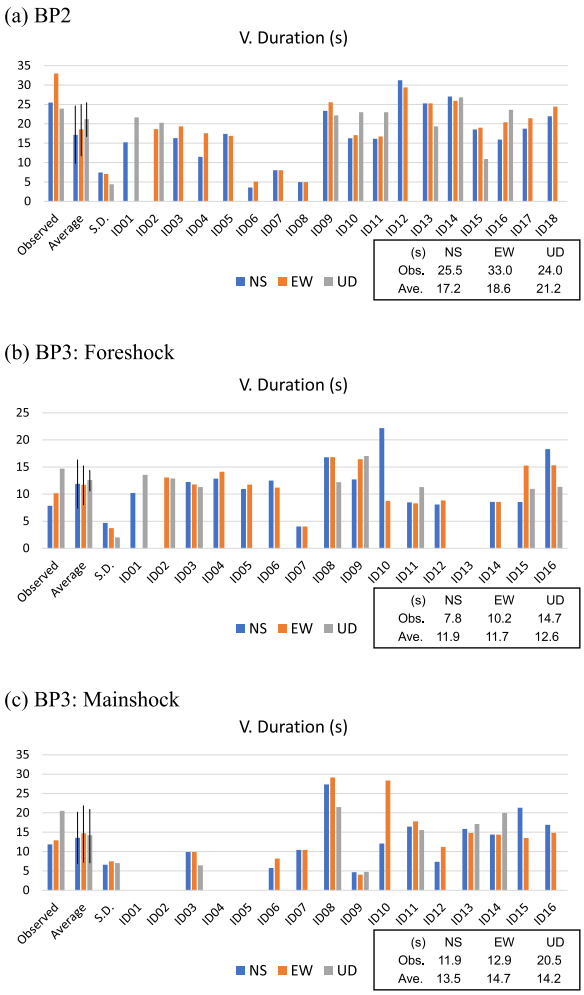


Fig. 12 Velocity durations of the three components in the frequency range of 0.1 to 10 Hz at KUMA for predictions with the average and standard deviation. A bar indicates the average ± the deviation. Acceleration duration observed is also shown. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

PGA and PGV

PGAs and PGVs of the three components in the frequency range of 0.1 to 10 Hz at KUMA for the largest aftershock of *Mj* 5.9 (BP2), foreshock of *Mj* 6.5 (BP3: foreshock), and mainshock of *Mj* 7.3 (BP3: mainshock) are shown with the observed PGAs and PGVs in Figs. 7 and 8, respectively. The averages and standard deviations σ of the PGAs and PGVs are also shown.

Fourier spectrum

Fourier spectra of the three components at KUMA are shown with the observation in Fig. 9, using the accelerations of 40.96 s with the Parzen's window of 0.1 Hz. The averages and standard deviations σ of the predictions are also shown. The variability of the individual Fourier spectra for the UD component in BP2 is clearly smaller than that of the individual Fourier spectra for the NS and EW components, particularly at frequencies higher than 1 Hz.

Velocity response spectrum with 5% damping

Pseudo-velocity response spectra of the three components at KUMA are shown with the observation in Fig. 10. The averages and standard deviations σ of the predictions are also shown. The variability of the individual pseudo-velocity response spectra for the UD component in BP2 is clearly smaller than that of the individual pseudo-velocity response spectra for the NS and EW components, particularly at the period less than 1 s, with the same tendency as the Fourier spectrum.

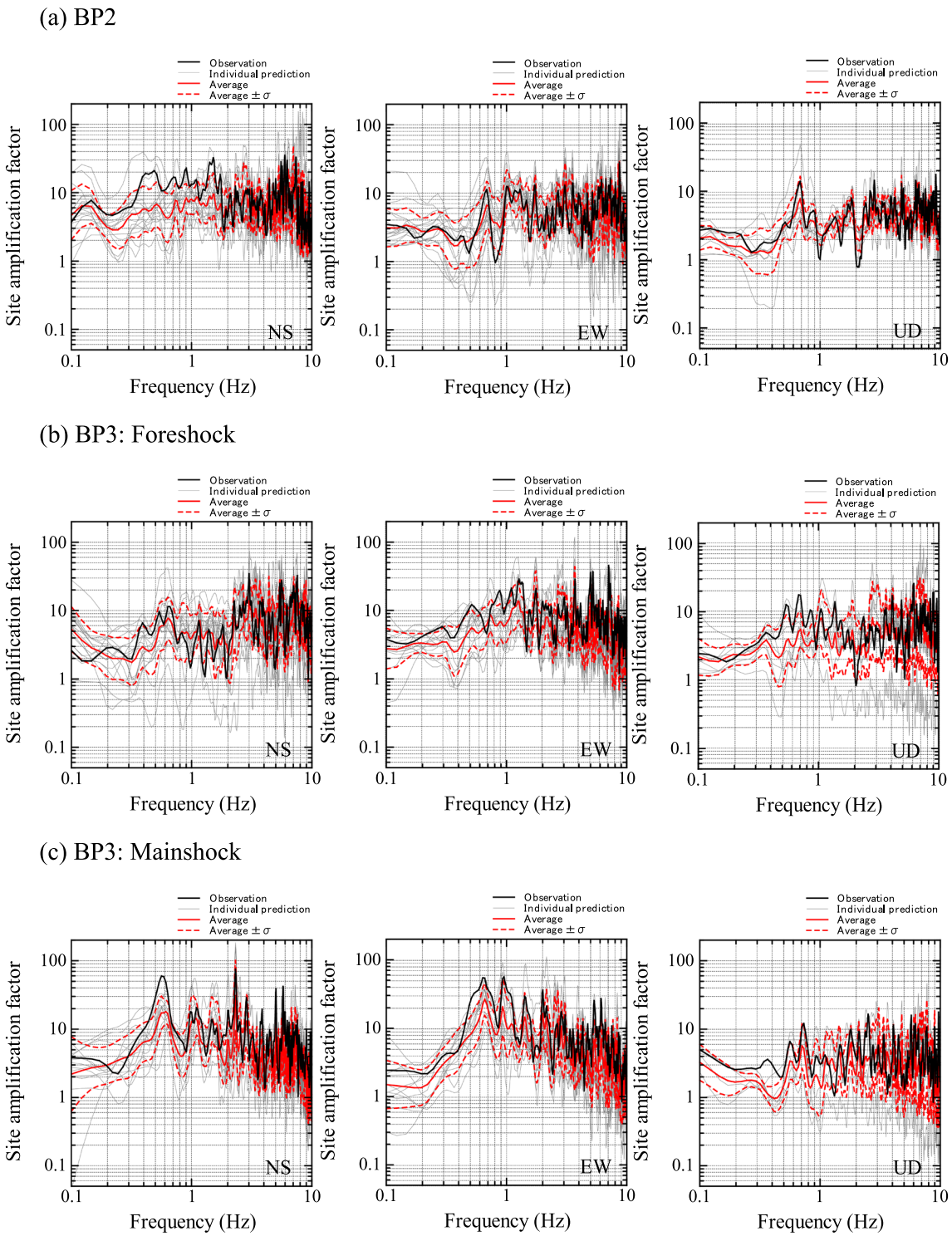


Fig. 13 Site amplification factors of the three components at KUMA from the bedrock for predictions with the average and the standard deviation. Site amplification factor observed is also shown. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

Duration

As acceleration duration of the earthquake ground motions, we adopted the Arias intensity (Arias 1970) that acceleration duration should be defined as the 5–95% duration of the squared acceleration integral, so the Arias intensity is also called Arias integral (Trifunac and Brady 1975). As velocity duration, we adopted the energy integral proposed by Anderson (2004) with the same process as that for the acceleration duration.

The acceleration and velocity durations of the three components in the frequency range of 0.1 to 10 Hz at KUMA are shown with the observation in Figs. 11 and 12, respectively. The averages and standard deviations σ of the predictions are shown. In general, the duration depends on the magnitude of the earthquakes. In these cases, however, the durations of the largest aftershock of *Mj* 5.9 are longer than those of the foreshock of *Mj* 6.5 and the mainshock of *Mj* 7.3, in the observations as well as in the predictions. This might be because the epicentral distance of the largest aftershock to KUMA was the longest among the earthquakes.

Site amplification factors

We estimated the site amplification factors at KUMA, that is, the Fourier spectrum of each prediction at KUMA divided by the input motion of 1/2 Fourier spectrum of the earthquake record at KU.KMP1. The empirical site amplification factor was estimated using earthquake records from KUMA and KU.KMP1. The site amplification factors of the three components at KUMA are shown in Fig. 13 with the averages and standard deviations σ of the predictions. In the cases of BP2 and BP3, the observations are mostly in the range of the average ± σ of the predictions for the three components, as similar to the PGA, PGV, Fourier spectrum, velocity response spectrum with 5% damping, and durations.

Validations between observations and predictions

Mean absolute percentage error

As an index of validation between observations and predictions (e.g., Morley et al. 2018), we first adopted the mean absolute percentage error (MAPE). We applied the MAPE with the geometric mean procedure to the individual prediction results of the PGAs, PGVs, acceleration durations, and velocity durations in BP2, BP3 (foreshock), and BP3 (mainshock) in Fig. 14, using Eq. (1). P_i and O indicate the individual prediction and observation values, respectively. The number of predictions is n. In the case of averaging the Fourier spectrum and the pseudo-velocity response in the frequency range of 0.1 to 10 Hz, Eq. (2) was applied. a represents number of frequencies:

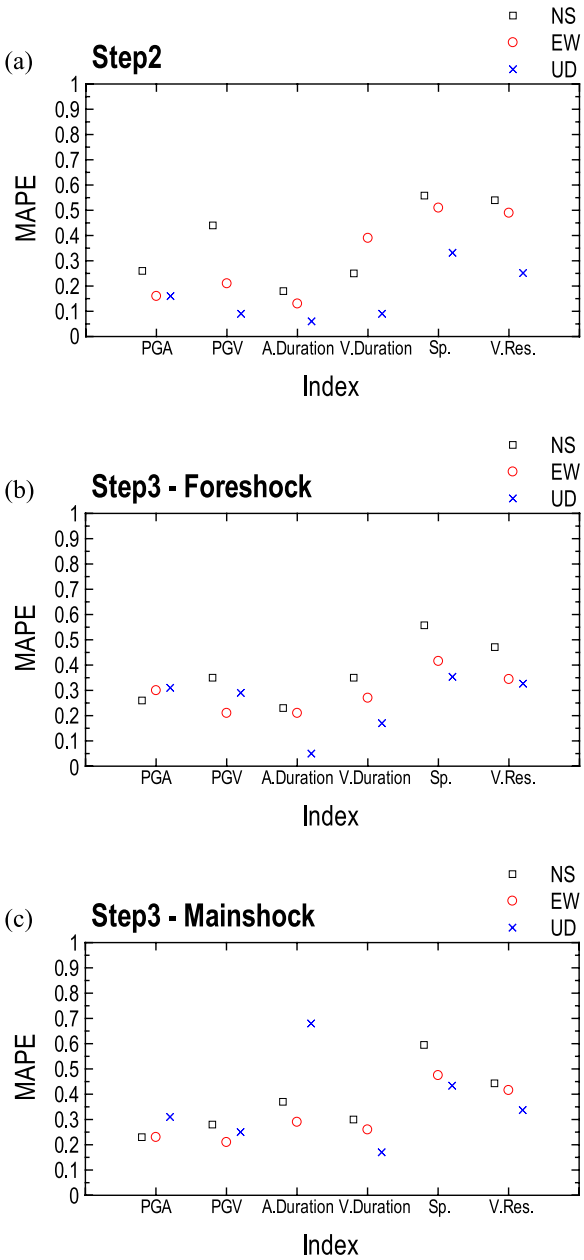


Fig. 14 MAPE for the PGAs, PGVs, acceleration duration, velocity duration, and Fourier spectra and the pseudo-velocity responses averaged in the frequency range of 0.1 to 10 Hz. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

$$\mathrm{MAPE} = \left(\prod_{i=1}^{n} \left| \frac{P_i - O}{O} \right| \right)^{1/n} \tag{1}$$

$$\mathrm{MAPE}infreq = \left(\prod_{i=1}^{n} \left(\frac{1}{a} \sum_{j=1}^{a} \left| \frac{P_i(f_j) - O(f_j)}{O(f_j)} \right| \right) \right)^{1/n} \tag{2}$$

In BP2, the MAPE for the PGAs, PGVs, acceleration durations, and velocity durations are 0.13–0.44 in the NS and EW components and 0.06–0.16 in the UD component. The ranges of the MAPE for the pseudo-velocity responses are approximately 1/2–2 times. However, those MAPE in the three components are 0.05–0.35 and 0.17–0.37 in BP3 (foreshock) and BP3 (mainshock), respectively. The ranges of the MAPE for the pseudo-velocity responses are almost 1/2–2 times in the BP3 (foreshock) and BP3 (mainshock). The MAPE for the predictions results in the UD component in BP2 is less than 0.1, indicating high accurate forecasting, as shown by Lewis (1982). The MAPE for other prediction results in BP2 and BP3 are 0.2–0.5 as reasonable forecasting shown by Lewis (1982).

The results of the validations by the MAPE between the observations and predictions in the blind prediction for the 2016 Kumamoto earthquake sequence in ESG6 are better than those in the blind prediction for the 1995 Kobe earthquake in ESG2 (Kawase and Iwata 1999, 2000). As for the PGA and PGV, the results from the blind predictions in ESG2 and ESG6 had the same tendency that the observations are mostly in the range of the average ± σ of the predictions in the three components. As for the velocity response spectra, on the other hand, the predictions submitted in ESG6 are rather well-constrained than those in ESG2.

Goodness-of-fit

The goodness-of-fit (GOF) can be used to investigate the credibility of simulated waveforms for engineering applications (Anderson 2004; Kristeková et al. 2006, 2009; Olsen and Mayhew 2010). We applied the GOF proposed by Anderson (2004) to the predicted weak and strong ground motions. The GOF in this study is composed of nine criteria with different characteristics, which have a scale from 0 to 10 for the Arias duration (C1), energy duration (C2), Arias intensity (C3), energy integral (C4), peak acceleration (C5), peak velocity (C6), peak displacement (C7), response spectra (C8), and Fourier spectra (C9). We excluded cross-correlation (C10) from the ten items of Anderson (2004), because the arrival time was not estimated in several predictions.

We calculated the scores of the GOF criterion in six frequency ranges (0.1–0.5, 0.5–1, 1–2, 2–5, 5–10, and 0.1–10 Hz), that are important for evaluating earthquake ground motions in engineering fields. Finally, the GOF score between the observations and predictions was obtained by averaging the GOF scores calculated for the six different frequency ranges. The GOF scores between the observations and predictions in BP2, BP3 (foreshock), and BP3 (mainshock) are shown in Fig. 15. Olsen and Mayhew (2010) proposed the following classification of the GOF score by Anderson (2004): 8–10 excellent fit, 6.5–8 very good fit, 4.5–6.5 good fit, 3.5–4.5 fair fit, and 0–3.5 poor fit. The GOF scores in BP2 are the highest among those of the three blind predictions, despite having the largest number of participants. The average GOF scores in BP2, BP3 (foreshock), and BP3 (mainshock) are over 4.5 for three components indicating either a very good fit or a good fit. These results indicate that techniques to simulate earthquake ground motions can adequately predict weak and strong ground motions in

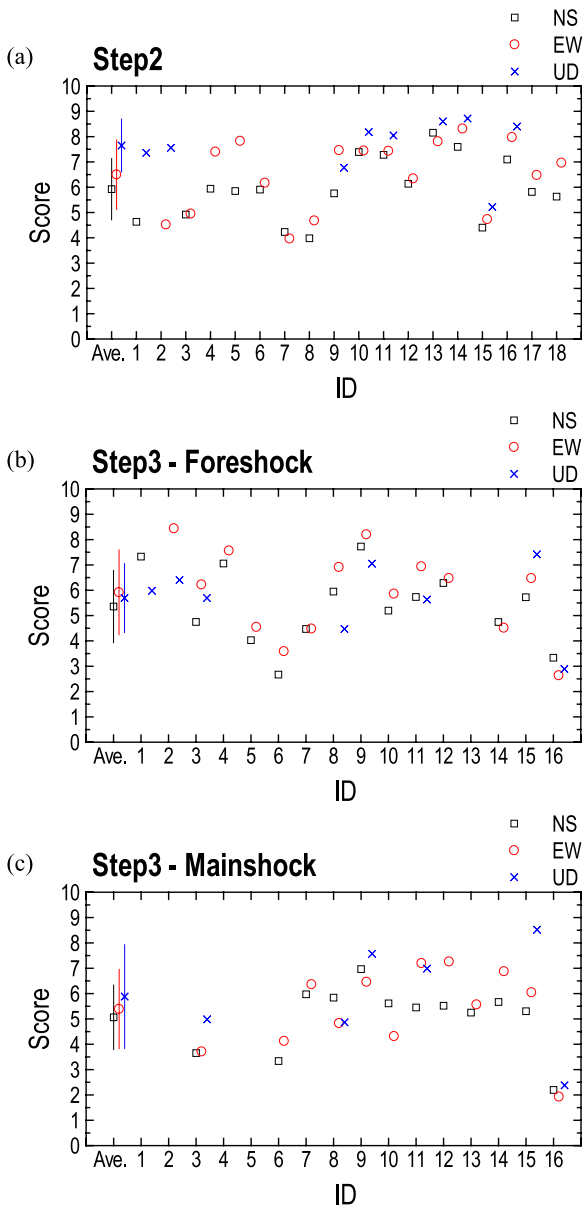


Fig. 15 GOF scores by Anderson (2004) between the observations and predictions. A bar indicates the average ± the deviation. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

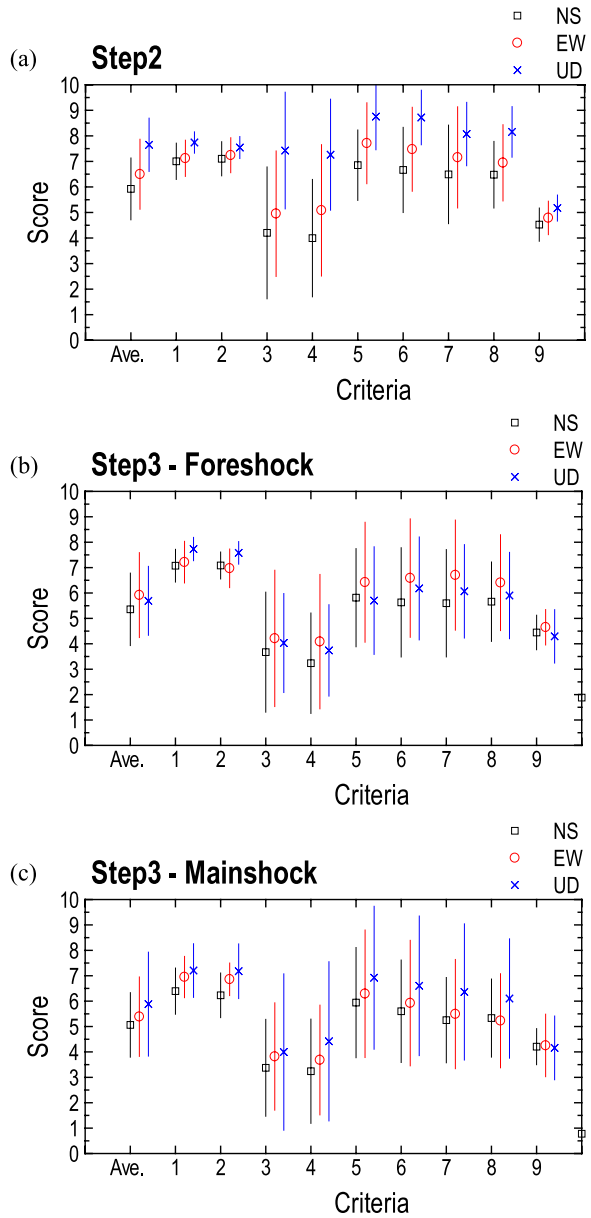


Fig. 16 GOF scores on the individual criteria by Anderson (2004) between the observations and predictions. A bar indicates the average ± the deviation. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

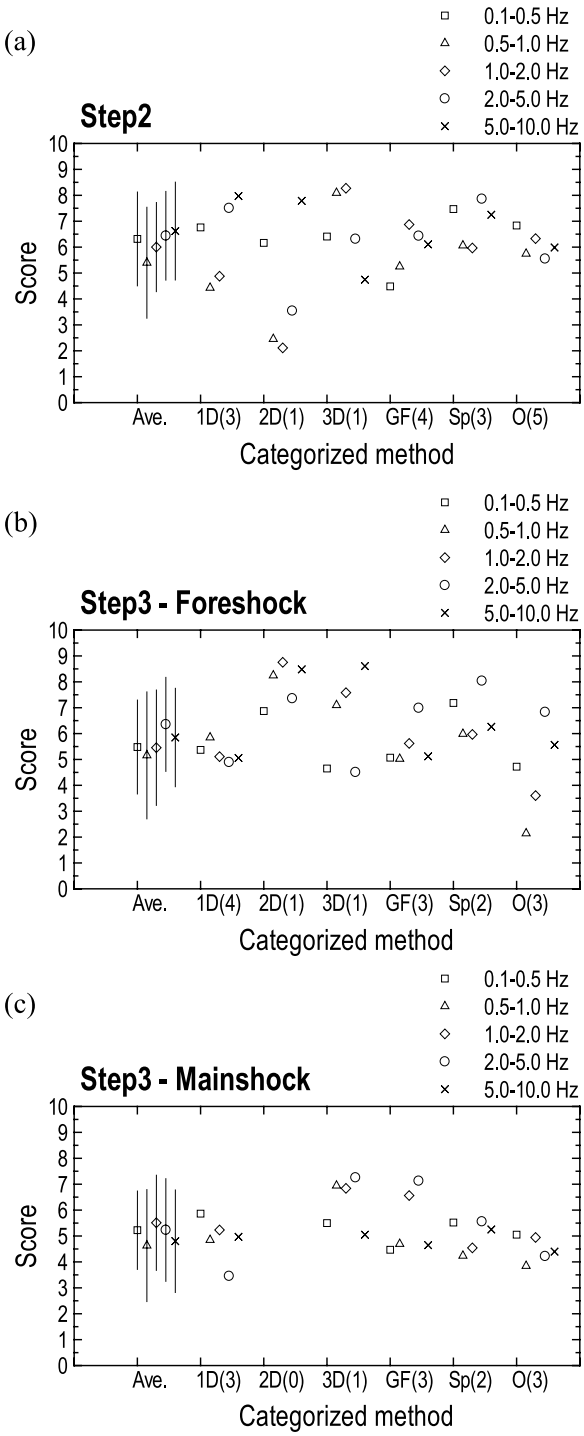


Fig. 17 GOF scores in the different frequency ranges between the observations and predictions by the categorized methods. 1D, 2D, 3D, GF, Sp, and O mean 1D method, 2D method, 3D method, Green's function method, spectral ratio approach, and other approaches. The number in the brackets indicates the number of predictions. **a** BP2, **b** BP3: foreshock, **c** BP3: mainshock

engineering applications when the parameters at the site are well-provided (Matsushima et al. 2023). The GOF scores on the individual criteria (C01–C09) by Anderson (2004) between the observations and predictions in BP2, BP3 (foreshock), and BP3 (mainshock) are shown in Fig. 16. The criteria corresponding to the energy contained in the seismic waves (C3 and C4) are relatively lower than the other criteria corresponding to the maximum value of the seismic waves (C5, C6, and C7). It is assumed that most results using the earthquake record

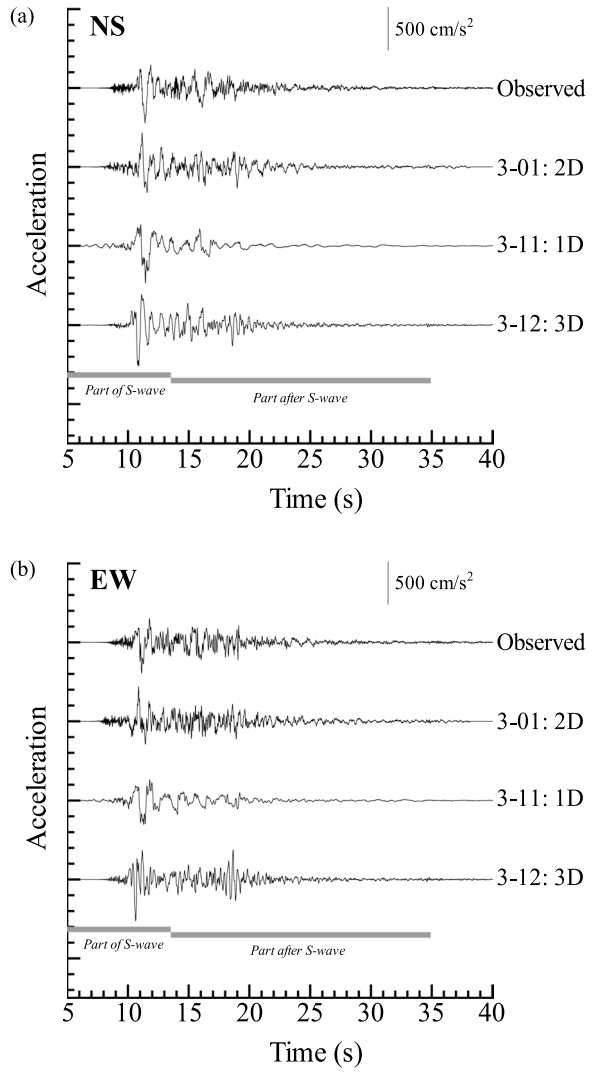


Fig. 18 Accelerations of the predictions by 3-01 (2D method), 3-11 (1D method), and 3-12 (3D method) with the observation in BP3 (foreshock) without the process of filtering. Shadow lines indicate the part of the S-wave (Part 1) and the part after the S-wave (Part 2). **a** NS component, **b** EW component

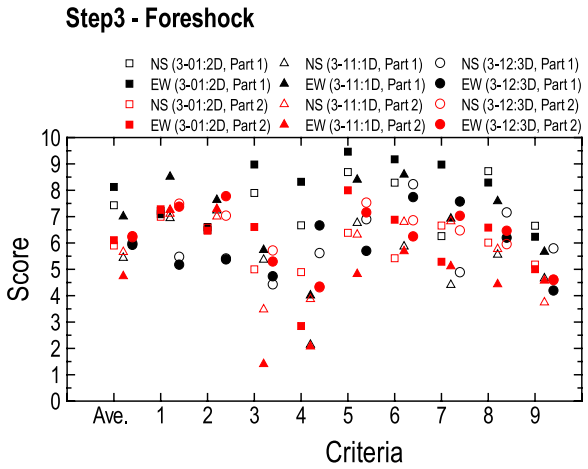


Fig. 19 GOF scores on individual criteria by Anderson (2004) between the observations and predictions by 3-01 (2D method) in the square, 3-11 (1D method) in the triangle, and 3-12 (3D method) in the diamond. NS and EW components showed in the open and the closed marks, respectively. Black and white indicate results using the part of the S-wave (Part 1) and the part after the S-wave (Part 2), respectively

observed at the reference site (KU.KMP1) could not properly reproduce the basin-induced surface waves and/or basin-transduced surface waves (Kawase 1996).

Discussion

2D and 3D methods accounting the basin-edge effect

To investigate the validity of the simulated waveforms against the observed waveforms among the different methods, the GOF scores were averaged using the categorized methods (1D, 2D, 3D, Green's function, spectral ratio, and others). Note that we could not clearly categorize the applied methods because of the combination of the multiple methods (Tables 2 and 3). The GOF scores in different frequency ranges (0.1–0.5, 0.5–1, 1–2, 2–5, and 5–10 Hz) between the observations and predictions by the categorized methods in BP2, BP3 (foreshock), and BP3 (mainshock) are shown in Fig. 17. We did not estimate the standard deviation for the individual categorized methods because of the limited number of applications. Although we could not find a significant difference in the results from categorized methods, it has the tendency that the scores by the 3D method in the frequency range of 0.5 to 1 and 1 to 2 Hz for all the blind predictions seem to be higher than the scores by the other methods. The scores by the 2D method in the frequency range of 0.5 to 1 and 1 to 2 Hz are the highest among the results in BP3: foreshock. The methods which can consider the irregular shape of subsurface structures may take advantage in predicting the earthquake ground motions in the frequency range of 0.5 to 1 and 1 to 2 Hz in these cases. Figure 18 shows the accelerations of the predictions by 3-01 (2D method), 3-11 (1D method), and 3-12 (3D method) with the observation in BP3 (foreshock) without the process of filtering. As a representative case of the 1D method, the result of 3-11 was selected owing to having the average GOF score among the 1D method. We found that the amplitudes of the part after the S-wave predicted by 3-01 (2D method) and 3-12 (3D method) agree compared with those predicted by 3-11 (1D method) for the observations. Figure 19 shows the GOF scores on individual criteria by Anderson (2004) between the observations and predictions for the part of the S-wave and the part after the S-wave. In the result, the GOF scores in the total for the part after the S-wave predicted by 3-01 (2D method) and 3-12 (3D method) are

Table 4 GOF scores by Anderson (2004) between the observations and predictions by the 1D method using the preferred model and its own model

(a) BP3: foreshock			
	Score	**Individual**	
Preferred model	6.34	6.34	(1)
Own model	4.91	7.31, 4.29, 3.13	(3)
(b) BP3: mainshock			
	Score	**Individual**	
Preferred model	5.87	6.32, 5.41	(2)
Own model	3.73	3.73	(1)

The number in the brackets indicates the number of predictions

in agreement compared with that predicted by 3-11 (1D method) for the observations, similar to the amplitudes. In particular, the GOF scores in the total for the part of the S-wave predicted by 3-12 (3D method) were the worst; however, the GOF scores in total for the part after the S-wave was the best. These results indicate that the predictions by the 2D and 3D methods due to the accounting of the proper geometry could reproduce basin-induced surface waves and/or basin-transduced surface waves excited by the basin-edge effect (Kawase 1996) better than the 1D method using the earthquake record observed at the reference site (KU.KMP1). Although the number of examples of the categorized methods is quite limited, these results lead us to conclude that the predictions by all categorized methods can adequately reproduce strong ground motions as well as weak ground motions within a very good fit and a good fit for the GOF.

Finally, we investigated the differences in the results of the 1D method using the preferred model and its own model in BP3. The GOF scores by Anderson (2004) between the observations and predictions by the 1D method are shown in Table 4. The number of examples here is limited; however, the results of the 1D method using the preferred model appear to be better than those using its own model in this study.

Comparing to the ground motion prediction equation

We verified the accuracy of the predictions for the foreshock and mainshock of the 2016 Kumamoto earthquake, by comparing them with a ground motion prediction equation (GMPE). Figure 20 shows the ground motion prediction equations for the foreshock and mainshock of the 2016 Kumamoto earthquake with the PGAs of the observation and the average of the predictions at KUMA without filtering. The GMPEs for the foreshock and mainshock of the 2016 Kumamoto earthquake were estimated according to Si and Midorikawa (1999), which explains the strong ground motions during the mainshock of the 2016 Kumamoto earthquake (Suzuki et al. 2017). The fault distances at KUMA for both earthquakes are listed in Table 1. Note that the GMPEs (Si and Midorikawa 1999) were estimated using earthquake records by 21 earthquakes in the fault types of inter-plate, intra-plate, and crustal on the surface in the entire Japan.

The observations and the averages of the predictions for the foreshock and the mainshock of the 2016 Kumamoto earthquake are within the range of average $\pm\sigma$ of these GMPEs. Particularly in the mainshock of the 2016 Kumamoto earthquake, the observation and average of the predictions are close to the average of the GMPE. The standard errors in the GMPE for PGA proposed by Si and Midorikawa (1999) were 0.27 and 0.25 against all data and the data with fault distance less than 100 km or equal, respectively. On the other hand, the standard errors for the predictions for PGA in BP3 (foreshock) and BP3 (mainshock) are 0.2 and 0.17 smaller than those by the GMPE, respectively. This suggests that the predictions for the foreshock and mainshock of the 2016 Kumamoto earthquake in BP3 reproduce the observations better than the GMPE estimations.

Conclusions

To improve our understanding of the quality of state-of-the-art methods on the reproducibility of the effects of surface geology on seismic motions, we performed validation analyses of MAPE and GOF between the observations and predictions for weak ground motions (BP2) and strong ground motions (BP3) using the aftershocks, foreshock, and mainshock of the 2016 Kumamoto earthquake sequence, Japan.

As for the PGA/PGV, acceleration/velocity duration, Fourier spectrum, and pseudo-velocity response spectrum, the observed values are mostly within the range of average $\pm\sigma$ of all the predictions in the case of weak and

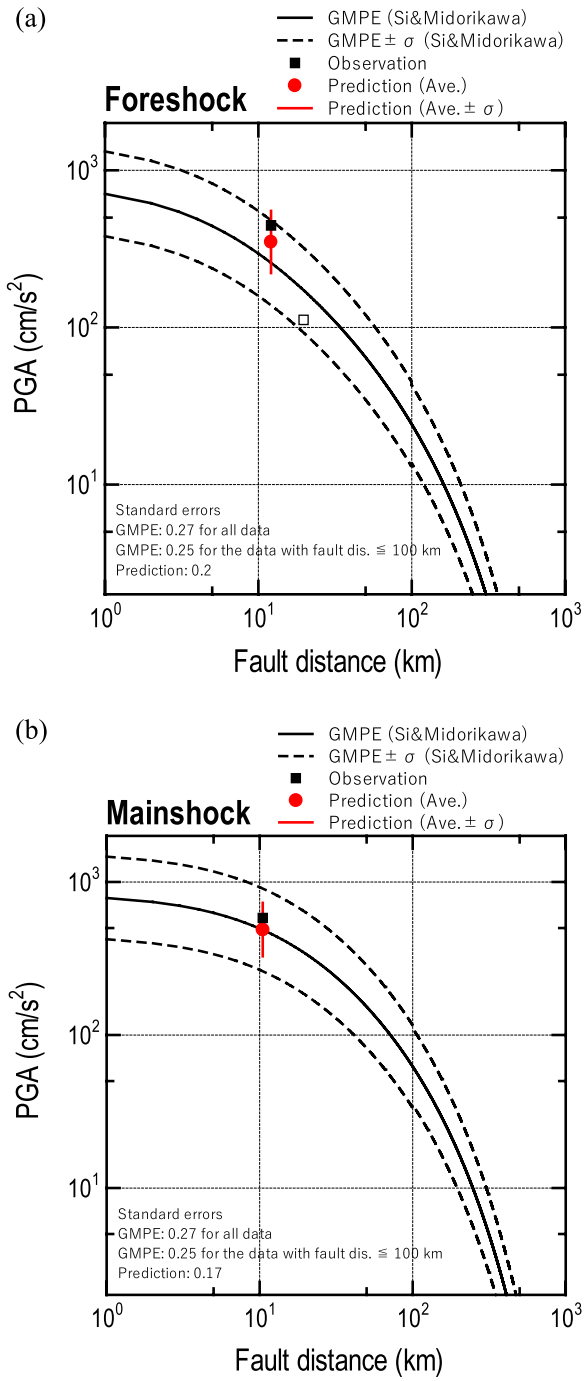


Fig. 20 Ground motion prediction equations for the foreshock and mainshock of the 2016 Kumamoto earthquake with the PGAs of the observation and the average of the predictions at KUMA without filtering. Bar indicates the average of the predictions ± deviation. The observation at KU.KMP1 for the foreshock is shown by an open square in the figure. **a** BP3: foreshock, **b** mainshock

strong ground motions. The relationship of the site amplification factors between the observations and predictions at the Kumamoto test site to the reference site also has the same tendency. The results of the MAPE for these indices show that the applied methods can predict weak and strong ground motions for the three components in the range of one-half to twice the observations. Moreover, the average GOF scores for weak and strong ground motions indicate either a very good fit (6.5–8) or a good fit (4.5–6.5) for the three components. Finally, examples of the categorized methods (1D, 2D, 3D, Green's function, spectral ratio, and others) are quite limited; however, results indicate that the predictions by all the categorized methods can adequately reproduce weak and strong ground motions within either a very good fit or good fit for GOF. Although we could not find a significant difference in the results from the categorized methods, GOF scores by the 2D and 3D methods in the frequency range of 0.5–1 and 1–2 Hz for all the blind predictions are higher than the scores by the other methods. The GOF score for the part after the S-wave by the 2D and 3D methods is higher than that by the 1D method for the observation. This supports that the predictions by the 2D and 3D methods due to the accounting of the proper geometry could reproduce the basin-induced surface waves and/or basin-transduced surface waves excited by the basin-edge effect (Kawase 1996) more than the 1D method using the earthquake record observed at the reference site (KU.KMP1).

To enhance the validity of earthquake ground motion predictions, both qualitatively and quantitatively for engineering applications, it is important to continue experiments such as blind predictions of earthquake ground motions.

Acknowledgements

We thank the Kyushu Railway Company and Institute of Seismology and Volcanology, Faculty of Science, Kyushu University for providing us with the earthquake records of the 2016 Kumamoto earthquake sequence. We thank Mr. Sugama, a graduate student at the Tokyo Institute of Technology, for helping with the data analysis. The participants of blind prediction Step-2 (BP2) without honorifics were Mercerat, Mezher, Bordoni, Milana, Lucarelli, Di Giulio, Hailemicael, Vassallo, Foti, Comina, Cosentini, Sappa, Srihari, Hayakawa, Honda, Sato, Kosaka, Nagasaka, Nagashima, Kawase, Ito, Nakano, Nozu, Yamanaka, Chimoto, Torimoto, Huang, Kuo, Karimzadeh, Askan, and Asten. The participants of blind prediction Step-3 (BP3) without honorifics were Castro-Cruz, Régnier, Bertrand, Derras, Bard, Bordoni, Milana, Lucarelli, Di Giulio, Hailemicael, Vassallo, Foti, Comina, Cosentini, Sappa, Srihari, Kosaka, Nagasaka, Nozu, Hayakawa, Yamanaka, Chimoto, Saifuddin, Nagashima, Kawase, Ito, Nakano, Huang, Kuo, Karimzadeh, Askan, and Asten. We thank all the participants for their contributions to BP2 and BP3 of ESG6. The authors also acknowledge the support and cooperation provided to conduct the surveys and all members of the ESG6 local organizing committee as well as the members of the research committee on the strong motion evaluation of JAEE. The concept building of BPs for ESG6 preparation and management was conducted by working groups (WGs) for the construction of the preferred velocity model (WG-PM) and for the analysis and selection of earthquake records (WG-GM) of the ESG6 local organizing committee. The members of the WG-GM were Prof. Hiroshi Kawase, Dr. Seiji Tsuno, and Dr. Takashi Hayakawa of Shimizu Corporation; Dr. Kazuhiro Kaneda of Chiba Institute of Technology; Dr. Tomonori Ikeura of Kajima Corporation; Prof. Tomotaka Iwata of Kyoto University; and Dr. Shinako Noguchi of the Association for the Development of Earthquake Prediction. We

thank two anonymous reviewers and the guest editor, Prof. Cotton Fabrice, for their helpful comments and suggestions on improving the manuscript.

Author contributions
ST and FN analyzed the earthquake records. ST drafted the manuscript. All the authors have revised the manuscript accordingly.

Funding
This study was partially supported by "Core-to-Core Collaborative Research between Earthquake Research Institute, The University of Tokyo and Disaster Prevention Research Institute, Kyoto University."

Availability data and materials
We provide digital data on the weak and strong ground motions released in these blind predictions (BP2 and BP3) on the website of the Japanese Association for Earthquake Engineering (JAEE), which was the main sponsor of ESG6. The data sets generated and/or analyzed during the current study are not publicly available but are available from the corresponding author upon reasonable request.

Declarations

Competing interests
The authors declare no competing financial interests.

Author details
[1]Railway Technical Research Institute, 2-8-38 Hikari-cho, Kokubunji-shi, Tokyo 185-8540, Japan. [2]Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan. [3]Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa 227-8503, Japan.

Received: 1 October 2022 Accepted: 19 August 2023
Published online: 29 August 2023

References

Advanced Industrial Science and Technology (AIST) Seamless digital geological map of Japan (1:200,000). https://gbank.gsj.jp/seamless/index_en.html. Accessed 17 Aug 2023.
Anderson JG (2004) Quantitative measure of the goodness-of-fit of synthetic seismograms. 13th World Conf. on Earthquake Engineering, 243.
Arias A (1970) A measure of earthquake intensity. In: Hansen R (ed) Seismic design of nuclear power plants. MIT Press, Cambridge, pp 438–483
Asano K, Iwata T (2016) Source rupture processes of the foreshock and mainshock in the 2016 Kumamoto earthquake sequence estimated from the kinematic waveform inversion of strong motion data. Earth Planets Space 68:147. https://doi.org/10.1186/s40623-016-0519-9
Chimoto K, Yamanaka H, Tsuno S, Matsushima S (2023) Predicted results of the velocity structure at the target site of the blind prediction exercise from microtremors and surface wave method as Step-1, Report for the experiments for the 6th international symposium on effects of surface geology on seismic motion. Earth Planets Space 75:79. https://doi.org/10.1186/s40623-023-01842-3
Hoshizumi H, Ozaki M, Miyazaki K, Matsuura H, Toshimitsu S, Uto K, Uchiumi S, Komazawa M, Hiroshima T, Sudo S (2004) Geological map of Japan 1:200,000. Geological Survey of Japan, AIST, Kumamoto
Ishizaka S, Iwasaki Y, Hase Y, Watanabe K, Iwauchi A, Taziri M (1995) Subsidence and sediments of the last interglacial epoch in the Kumamoto Plain, Japan. Quat Res 34:335–344 (**in Japanese with English abstract**)
Kawase H (1996) The cause of the damage belt in Kobe: "The Basin-edge effect", Constructive interference of the direct S-wave with the basin-induced diffracted/Rayleigh waves. Seismo Res Lett 67(5):25–34. https://doi.org/10.1785/gssrl.67.5.25
Kawase H, Iwata T (1999) A report on submitted results of the simultaneous simulation for Kobe. In: Irikura K, Kudo K, Okada H, Sasatani T (eds) The effects of surface geology on seismic motion. Balkema, Rotterdam, pp 1311–1337
Kawase H, Iwata T (2000) Simultaneous simulation for Kobe: What we have learned at ESG98. 12th World Conf. on Earthquake Engineering, 2684
Koketsu K, Miyake H, Afnimar TY (2009) A proposal for a standard procedure of modeling 3-D velocity structures and its application to the Tokyo metropolitan area, Japan. Tectonophysics 472:290–300. https://doi.org/10.1016/j.tecto.2008.05.037
Koketsu K, Miyake H, Suzuki H (2012) Japan integrated velocity structure model version 1. 15th World Conf on Earthquake Engineering, 1773
Kristeková M, Kristek J, Moczo P, Day SM (2006) Misfit criteria for quantitative comparison of seismograms. Bull Seismol Soc Am 96(5):1836–1850. https://doi.org/10.1785/0120060012
Kristeková M, Kristek J, Moczo P (2009) Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of time signals. Geophys J Int 178(2):813–825. https://doi.org/10.1111/j.1365-246X.2009.04177.x
Lewis CD (1982) Industrial and business forecasting methods. Butterworths Scientific, London
Matsushima S, Yamanaka H, Tsuno S, Chimoto K, Suzuki H, Kawase H (2023) Investigation of the subsurface structure at the target site in Kumamoto, Japan and the distributed data of the blind prediction exercise, Report for the experiments for "The 6th International Symposium on Effects of Surface Geology on Seismic Motion", submitted to this special volume
Morley SK, Brito TV, Welling DT (2018) Measures of model performance based on the log accuracy ratio. Space Weather 16:69–88. https://doi.org/10.1002/2017SW001669
Olsen KM, Mayhew JE (2010) Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 Mw 5.4 Chino Hills, California, Earthquake. Seismo Res Lett 81(5):715–723. https://doi.org/10.1785/gssrl.81.5.715
Senna S, Wakai A, Suzuki H, Yatagai A, Matsuyama H, Fujiwara H (2018) Modeling of the subsurface structure from the seismic bedrock to the ground surface for a broadband strong motion evaluation in Kumamoto plain. J Disaster Res 13(5):917–927. https://doi.org/10.20965/jdr.2018.p0917
Si H, Midorikawa S (1999) New attenuation relationships for peak ground acceleration and velocity considering effects of fault type and site condition. J Struct Constr Eng 523:63–70 (**in Japanese with English abstract**)
Suzuki W, Aoi S, Kunugi T, Kubo H, Morikawa N, Nakamura H, Kimura T, Fujiwara H (2017) Strong motions observed by K-NET and KiK-net during the 2016 Kumamoto earthquake sequence. Earth Planets Space 69:19. https://doi.org/10.1186/s40623-017-0604-8
The Japanese Geotechnical Society (2015) Japanese geotechnical society standards: laboratory testing standards of geomaterials (Vol. 1)
The Japanese Geotechnical Society (2017) Japanese geotechnical society standards: laboratory testing standards of geomaterials (Vol. 2)
Trifunac MC, Brady AG (1975) A study of the duration of strong earthquake ground motion. Bull Seismol Soc Am 65:581–626
Tsuno S, Korenaga M, Okamoto K, Yamanaka H, Chimoto K, Matsushima T (2017) Local site effects in Kumamoto City revealed by the 2016 Kumamoto earthquake. Earth Planets Space 69:37. https://doi.org/10.1186/s40623-017-0622-6

Publisher's Note