17th World Conference on Earthquake Engineering, 17WCEE
Sendai, Japan - September 13th to 18th 2020

GROUND STRUCTURE MODELING FOR PHYSICAL LOGGING DATA WITH SPARSITY REGULARIZATION

T. Kurita[(1)], Y. Shingaki[(2)], and M. Yoshimi[(3)]

[(1)] *Tokyo Electric Power Services Co., Ltd., kurita@tepsco.co.jp*
[(2)] *Tokyo Electric Power Services Co., Ltd, shingaki@tepsco.co.jp*
[(3)] *National Institute of Advanced Industrial Science and Technology, yoshimi.m@aist.go.jp*

Abstract

Sparse estimation is utilized as an effective modeling technique in the fields of machine learning and image processing. In this study, sparse modeling was applied to seismic engineering to build a rational modeling approach and it's the effectiveness of this approach was examined. Fused lasso, which is one sparse modeling method, was applied to model the ground structure using physical logging data.

The fused lasso is a kind of generalized lasso proposed by Tibshirani and Taylor (2011). The 1D fused lasso formulation is a simple case of the generalized lasso. In the formulation, L_1 norm with respect to unknown parameters is incorporated into the general least squares problem. L_1 norm is controlled by the regularization parameter.

Sparse modeling was applied to model the ground structures using the density and PS logging results of the subsurface ground obtained in the heavily damaged areas of the 2016 Kumamoto earthquake. Fused lasso solutions corresponding to fluctuated values of the regularization parameter were obtained. As the value of the regularization parameter increases, the degree of freedom of the estimated model decreases.

Since the seismic amplification characteristics of the subsurface ground depend on the ground structure model, theoretical values expected from the estimated models were compared with earthquake and microtremor observation records to confirm the validity of abovementioned approaches. The observation records were obtained from the temporal earthquake observation conducted after the 2016 Kumamoto earthquake. The theoretical values of the surface waves calculated from the estimated model were compared with the H/V spectra of the observation records. They correspond well, thus confirming the validity of the estimation model. Then, an attempt was made to select appropriate models based on the information criterion.

Finally, the effect of modeling on the amplification characteristics of the surface ground was examined using the obtained sparse solutions. The difference between the original models based on the physical logging data and the optimum models selected from the sparse solutions increased as the amplitude level of the input acceleration increased.

Keywords: sparse modeling; fused lasso; physical logging; ground structure; amplification characteristics

1. Introduction

Analysis of the seismic response of subsurface ground requires the in-depth physical properties of the ground at the target site. Thus, boring surveys are often performed. There are two types of PS logging: downhole PS logging and suspension PS logging. Since the measurement mechanism of the two is different, the result is naturally different. The former oscillates on the ground surface and receives signals in the borehole, while the latter oscillates in the borehole and receives the signal in the same borehole. In downhole PS logging, the velocity structure for each geological layer is obtained. On the other hand, suspension PS logging can produce more detailed velocity distribution, so it can identify local velocity changes in a geological layer or the existence of thin layers. In general, downhole PS logging tends to yield the average velocity structure of the ground, and suspension PS logging tends to yield local velocity structure. The analysis of the seismic response of subsurface ground using the one-dimensional wave propagation theory assumes that the target site is the horizontally layered ground, so it is important that the velocity distribution obtained from the PS logging results should be sufficiently continued in the horizontal direction.

When response analysis of subsurface ground is performed using high resolution data obtained by suspension PS logging and density logging, it is common to construct a ground structure model for numerical analysis. Since this modeling is usually based on empirical knowledge, it tends to be arbitrary. Therefore, in this study, we tried to pursue an approach to modeling using a methodology that is as objective as possible. The method we focused on is a fused lasso, a kind of sparse modeling. Sparse modeling is a technique that has been developed as an information processing technology, and suggests that the essence of a thing can be found by making complicated matters sparse. Fused lasso is a lasso with a constraint term added so that the coefficients of adjacent explanatory variables have the same value, and can be applied to matrix data in which the variables are ordered [1]. Lasso (least absolute shrinkage and selection operator) is an L_1 regularization method proposed by Tibshirani [2] for the purpose of dimensional compression (prediction using as few variables as possible).

The authors modeled the ground structure by applying the fused lasso to physical logging data in Mashiki town, which was affected by the 2016 Kumamoto earthquake, and appropriate models were selected based on the temporal seismic observation data and information criterion [3]. In this study, we tried the model selection based on another information criterion. Then, the effect of modeling on the amplification characteristics of the surface ground was examined using the obtained sparse solutions.

2. Methodology and Data

2.1 Theory of fused lasso

The fused lasso is a kind of generalized lasso proposed by Tibshirani and Taylor (2011) [4]. The 1D fused lasso solution represented by $\hat{\boldsymbol{\beta}}$ is obtained using the following formula,

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \frac{1}{2}\sum_{i=1}^{n}\left(y_i - \beta_i\right)^2 + \lambda\sum_{i=1}^{n-1}\left|\beta_{i+1} - \beta_i\right| \tag{1}$$

where y_i is an observation variable, β_i is an unknown model parameter, and $\lambda \geq 0$ is a regularization parameter. In Eq. (1), L_1 norm with respect to β_i is incorporated into the general least squares problem.

2.2 Summary of physical logging data

In this study, we used the results of suspension PS logging in Mashiki town performed by Yoshimi et al. [5, 6] and the results of density logging performed in the same area by Shingaki et al. [7] after the 2016 Kumamoto earthquake. The PS logging was performed at three points from GS-MSK-1 to GS-MSK-3, and density logging was performed only at GS-MSK-1. The subsurface ground of the survey sites is mainly composed of volcanic deposits.

The data of the suspension PS logging were obtained in increments of 50 cm in the depth direction. On the other hand, since the density logging data was obtained at intervals of 6 cm, the averaged value was obtained for each 50 cm section so as to match the PS logging.

2.3 Temporal seismic observation

Temporal seismic observation was conducted on the ground surface at the three sites mentioned above [3]. The observation period was from October 2016 to January 2017. Acceleration records of 20 earthquakes are obtained as principal events. In addition to the aftershocks of the Kumamoto earthquake, records of earthquakes that occurred in remote areas such as the earthquake on 21 October in eastern Tottori prefecture (Mw 6.2) have been obtained. In this study, we focused on this far field seismogram because the surface waves are expected to be dominant in the coda part of the ground motions. Furthermore, since the data was recorded as a continuous record, microtremors between events can also be used.

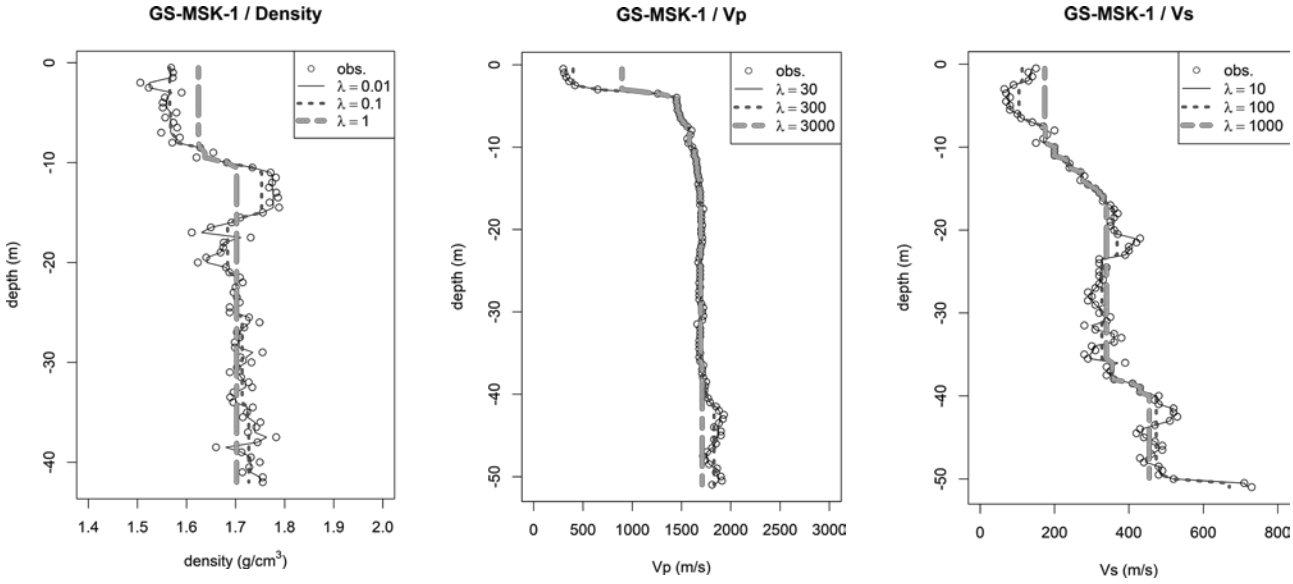
3. Fused Lasso Solutions

31 Modeling of ground structure by fused lasso

Fig.1 shows the examples of ground structure models for GS-MSK-1 obtained in the previous study [3]. The models are the fused lasso solutions for varying regularization parameters (λ). Table 1 shows the combinations of regularization parameters. As the regularization parameter is increased, a highly sparse ground structure model is obtained.

Table 1 – Values of the regularization parameter (λ) in each modeling case

modeling	λ		
	density	Vp	Vs
case 1	0.01	30	10
case 2	0.1	300	100
case 3	1	3000	1000



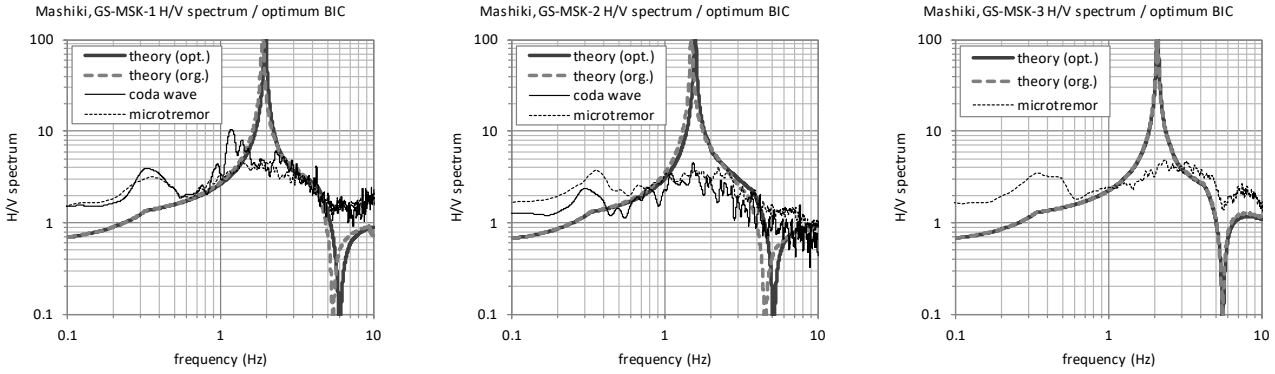
(1) density (2) P wave velocity (3) S wave velocity

Fig.1 – Modeling of ground structure by fused lasso [3]

3.2 Verification of sparse models by observation record reproducibility

In the previous study, several ground structure models were obtained depending on the value of the regularization parameter. However, what is actually needed is a model that can adequately express the amplification characteristics of the subsurface ground. Therefore, a comparison against the observation data obtained from the temporary earthquake observation was performed. We focused on the H/V spectra of the coda part of the seismic waves and microtremors. Though the H/V spectrum based on surface wave theory does not directly represent the amplification characteristics of the subsurface ground, it is used in inverse analysis of the ground structure because it is closely related to the ground structure.

Fig. 2 shows a comparison of the H/V spectra between the observation records and the theoretical values calculated from the ground structure model. The observed values indicate the geometric mean of multiple H/V spectra. The theoretical values were calculated based on the theory of Arai and Tokimatsu [8] using the ground structure model obtained by fused lasso, and only the fundamental mode of the surface wave considering both Rayleigh waves and Love waves was used. In the theoretical H/V spectra, “org.” means the calculated value using the original physical logging data. On the other hand, “opt.” indicates the similar value based on the optimum ground structure model obtained from the following investigation.



(1) GS-MSK-1 (2) GS-MSK-2 (3) GS-MSK-3

Fig.2 – Comparison of H/V spectra between observed data and theoritical model obtained by fused lasso

3.3 Appropriate model selection based on the information criterion

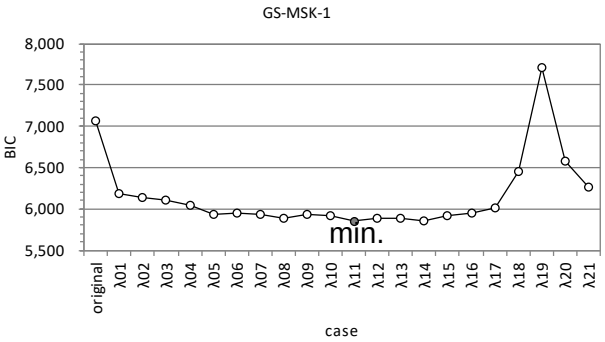
As shown in the previous studies, multiple sparse ground structure models can be obtained by the fused lasso. In addition, it seems that there is an appropriate ground structure model which can explain the observation record well. In this paper, we tried to select an appropriate model based on the information criterion.

BIC (Bayesian Information Criterion) [9] was used as a criterion for model selection. The idea of BIC is that the model with the highest probability of being a true model is better. BIC is defined as follows:

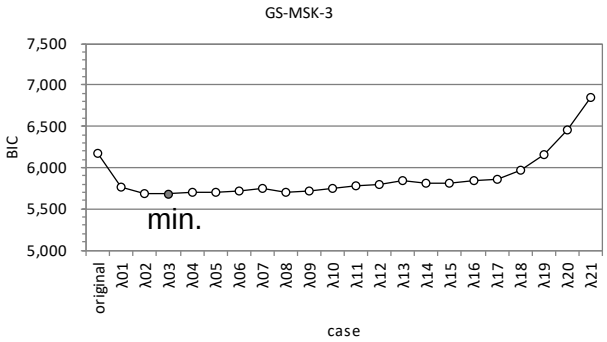
$$\mathrm{BIC} = -2 \cdot MLL + m \cdot \ln(n) \tag{2}$$

Here, MLL : maximum logarithmic likelihood, m : number of independent variables (degree of freedom), and n : number of observed data. The frequency range in which the residual sum of squares was calculated was 0.5 Hz to 7.0 Hz. Since the value of BIC indicates the poor fitness of the model, the model with the smallest BIC is set as the optimum model.

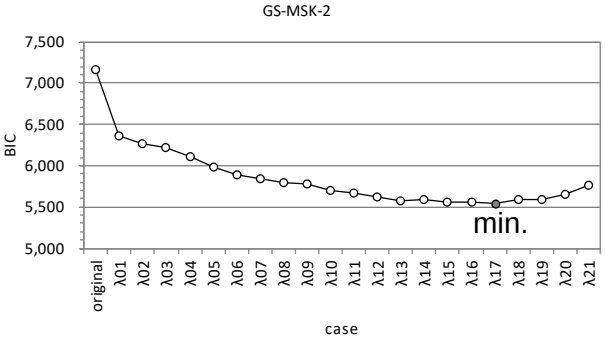
Fig. 3 shows the transition of the BIC for each estimation model obtained by changing the regularization parameter. Symbols that show the minimum value in the figure are filled with gray. From the figure, at GS-MSK-3, BIC shows the minimum value (λ03) when the regularization parameter is increased by two steps from the minimum regularization parameter (λ 01). On the other hand, at GS-MSK-1 and GS-MSK-2, BIC is lowest when the value of λ is relatively large. These models were selected as the optimum models based on BIC.



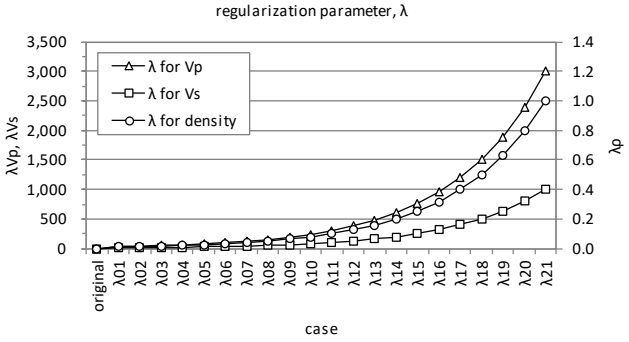
(1) GS-MSK-1



(3) GS-MSK-3

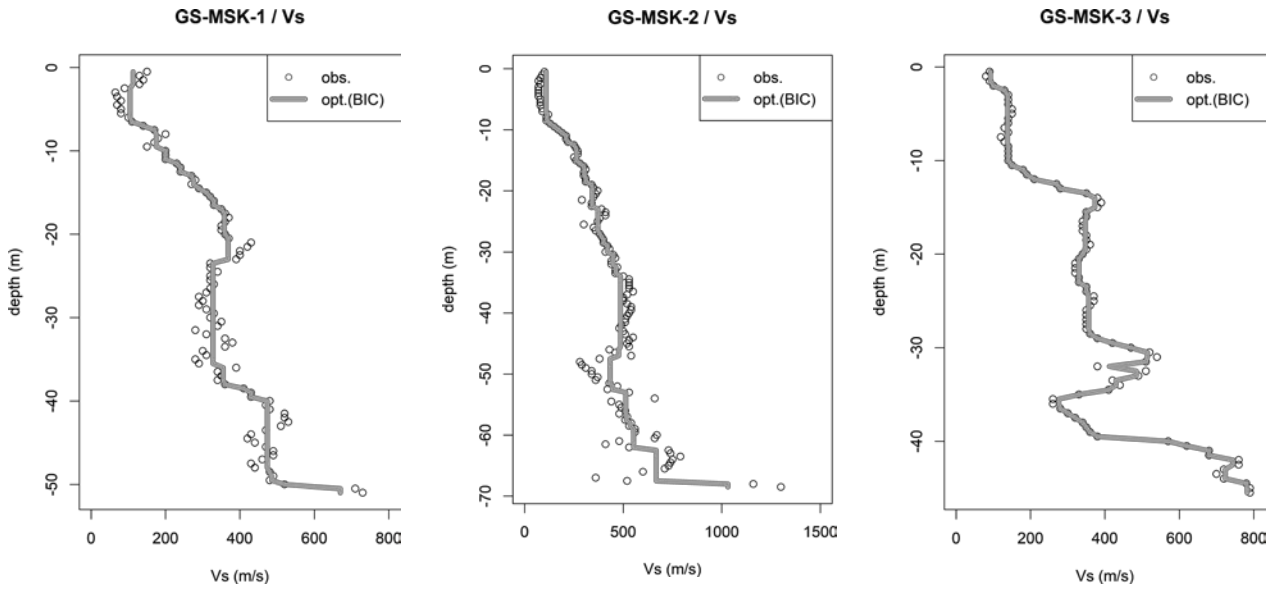


(2) GS-MSK-2



(4) regularization paremeter

Fig.3 – Transition of BIC varying with regularization parameter



(1) GS-MSK-1 (2) GS-MSK-2 (3) GS-MSK-3

Fig.4 – Optimum Vs structure models selected based on the BIC

4. Effect of Model Selection on Nonlinear Amplification Characteristics of SH Waves

The SH wave amplification characteristics of the optimum ground structure model selected based on the BIC were investigated in consideration of the nonlinearity of the soil properties.

4.1 Analysis conditions

The SH wave that was pulled back from the EW component of seismic record at the ground surface of KiK-net Mashiki (KMMH16) observed during the 2016 Kumamoto Earthquake to the stiff layer of GL-101m (S wave velocity:1389.4m/s) was set as the reference input seismic wave. At this time, the optimum ground structure model proposed by Kurita (2016) [10] was used as the ground structure model for the earthquake response analysis. Fig. 5 shows the time history of the reference input seismic wave.

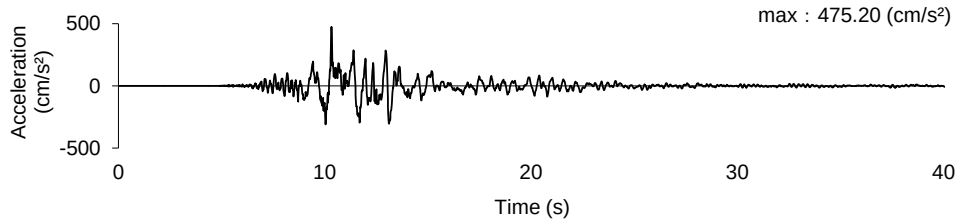


Fig.5 – Time history of the reference input seismic wave

Fig.6 shows the dynamic deformation characteristics for the equivalent linear seismic response analysis used in the following examination. The dynamic deformation characteristics were prepared based on the soil samples collected in the aforementioned boring survey.

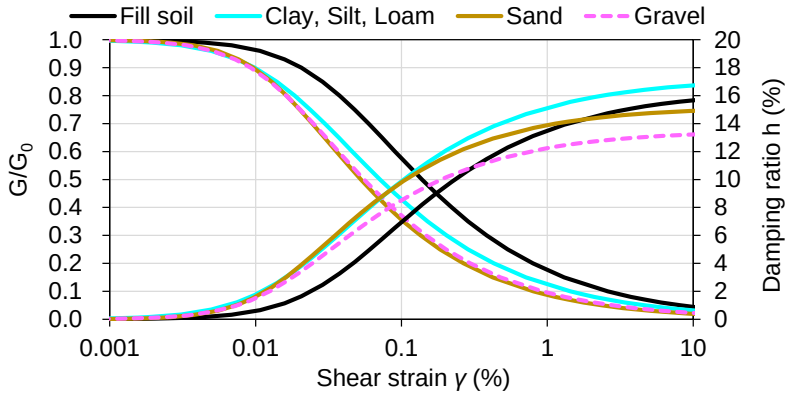


Fig.6 – Dynamic deformation characteristics of borehole test site

The seismic response analysis for the ground structure models was conducted using the equivalent linear analysis based on the one-dimensional wave propagation theory.

4.2 Comparison of seismic response analysis results according to input ground motion intensity

Input ground motions with five different amplitude levels adjusted from the reference input seismic wave (1.5, 1.0, 0.5, 0.2, and 0.1 times exaggerations from the reference) are applied for the seismic response analysis.

i) 1.5 times the reference wave: 712.8 cm/s^2

ii) Reference wave: 475.2 cm/s^2

iii) 0.5 times the reference wave: 237.6 cm/s^2

iv) 0.2 times the reference wave: 95.0 cm/s^2

v) 0.1 times the reference wave: 47.5 cm/s^2

The differences at GS-MSK-2 between the results of the equivalent linear analysis using the original model and the optimum model were obtained in proportions as shown in Fig. 7. The original model refers to the ground structure model based on the physical logging data. At the same time, the optimum model refers to the ground structure obtained from the model selection based on the BIC. In this figure, the result of the maximum shear strain is clearly smaller than others. There is a tendency for the maximum acceleration on the ground surface to generally increase as the input ground motion increases. The other maximums hit the ceiling at increased acceleration levels.

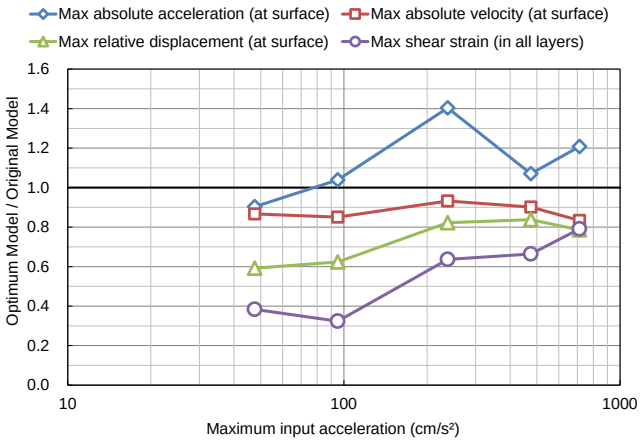


Fig.7 – Ratio of various maximum value (optimum model / original model) at GS-MSK-2

Seismic response analysis results for different input ground motion intensities at GS-MSK-2 are shown in Fig. 8 and Fig. 9. In Fig. 8, the equivalent S wave velocity and equivalent damping ratio mean that converged values were obtained from the iteration calculation in the seismic response analysis. Both figures show a comparison between the optimum model and the original model. As the input acceleration increases, the difference in the calculation results between the models increases. From Fig. 8(4) and Fig. 9, the maximum acceleration on the ground surface is larger in the optimum model than in the original model. This is because the impedance ratios of the latter near GL-35m and shallower are larger than that of the former. Fig. 8(6) shows that the maximum displacement and the maximum shear strain near the ground surface are larger in the original model than in the optimum model. This is because the former impedance ratio near GL-10m is larger than the latter.

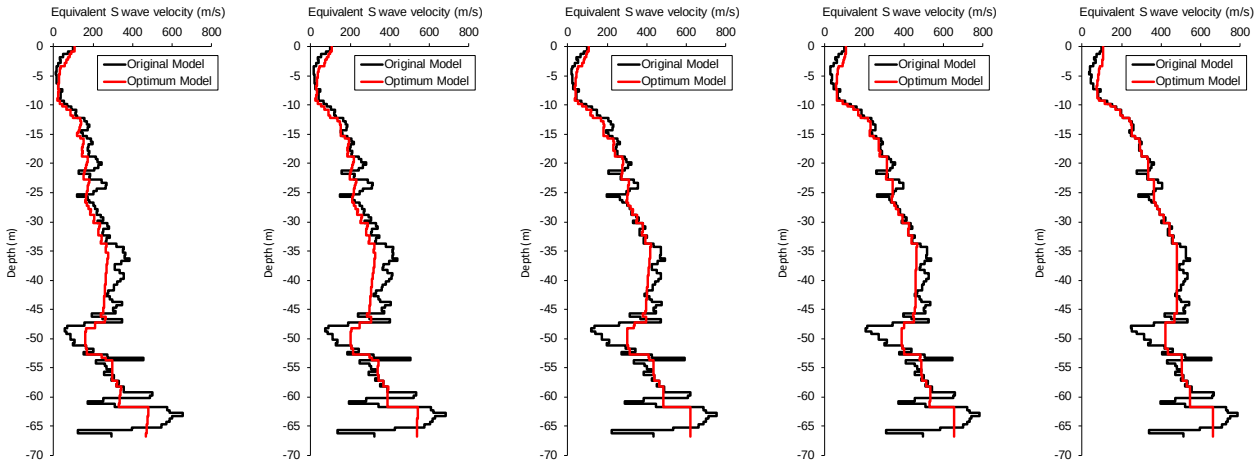
In order to quantitatively express the difference between the two models at each site, an evaluation was performed using the following index. The nRMSEs (normalized Root Mean Square Errors) of the calculation results of the original model and the optimum model were determined using the following equation:

$$\mathrm{nRMSE} = \frac{1}{\sigma_{org}} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left\{ x_{opt}(i) - x_{org}(i) \right\}^2} \tag{3}$$

$$\sigma_{org} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left\{ x_{org}(i) - \overline{x}_{org} \right\}^2}$$

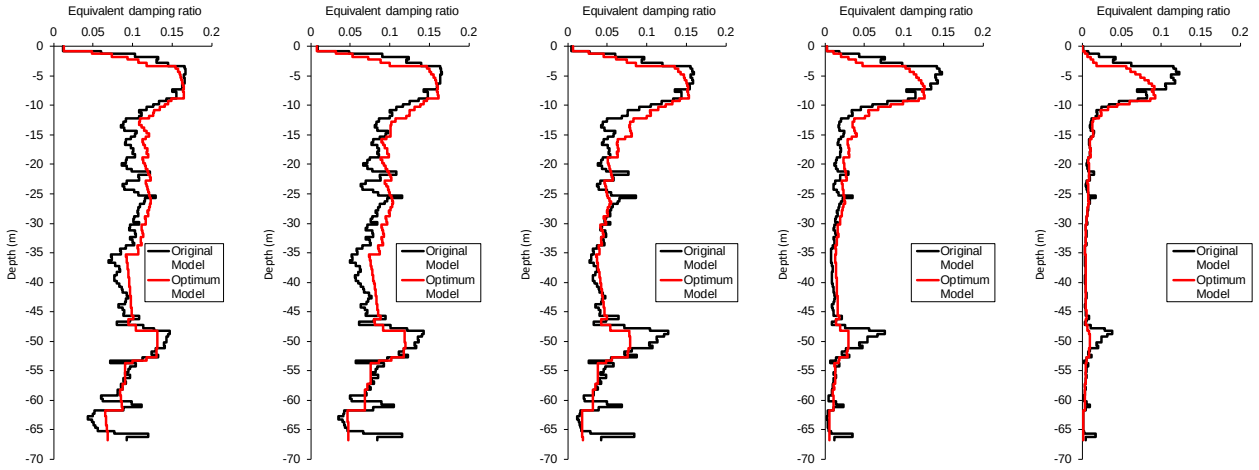
$$\overline{x}_{org} = \frac{1}{n} \sum_{i=1}^{n} x_{org}(i)$$

Here, $x_{opt}(i)$: results of seismic response analysis with optimum ground structure model, $x_{org}(i)$: results of seismic response analysis with original physical logging data, n: number of data, respectively.



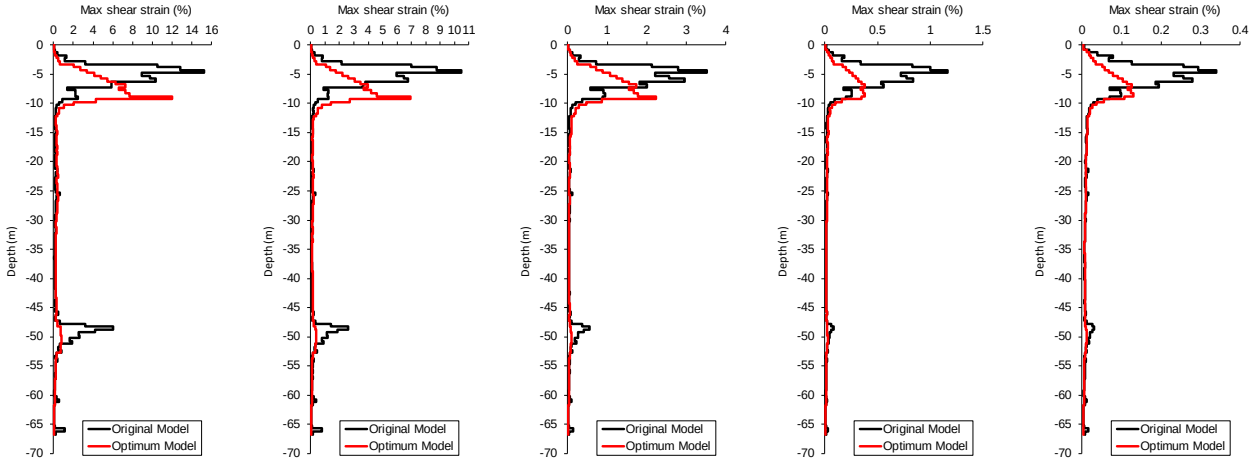
input × 1.5 (712.8cm/s²) input × 1.0 (475.2cm/s²) input × 0.5 (237.6cm/s²) input × 0.2 (95.04cm/s²) input × 0.1 (47.52cm/s²)

(1) Distribution of equivalent S wave velocity in the depth direction



input × 1.5 (712.8cm/s²) input × 1.0 (475.2cm/s²) input × 0.5 (237.6cm/s²) input × 0.2 (95.04cm/s²) input × 0.1 (47.52cm/s²)

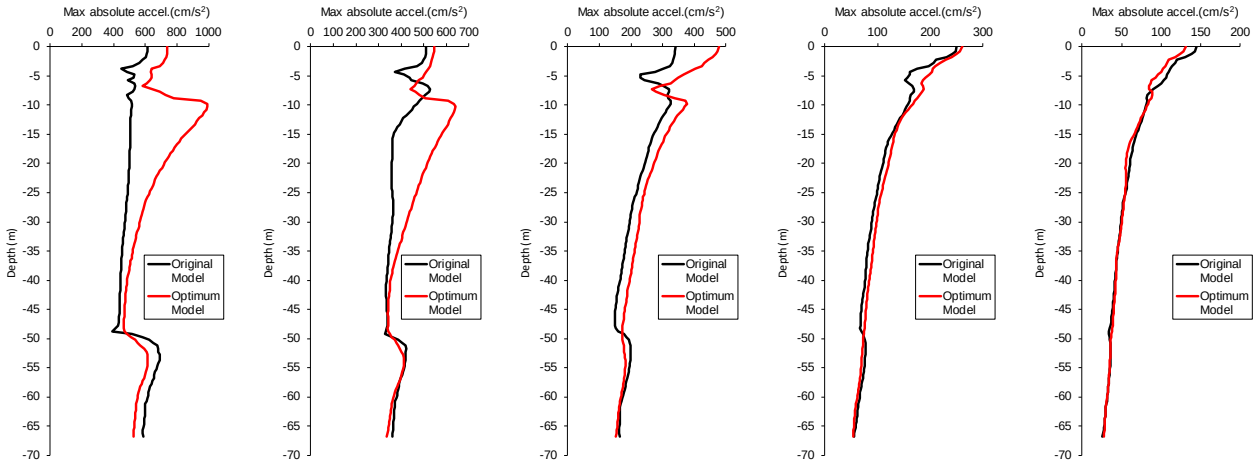
(2) Distribution of equivalent damping ratio in the depth direction



input × 1.5 (712.8cm/s²) input × 1.0 (475.2cm/s²) input × 0.5 (237.6cm/s²) input × 0.2 (95.04cm/s²) input × 0.1 (47.52cm/s²)

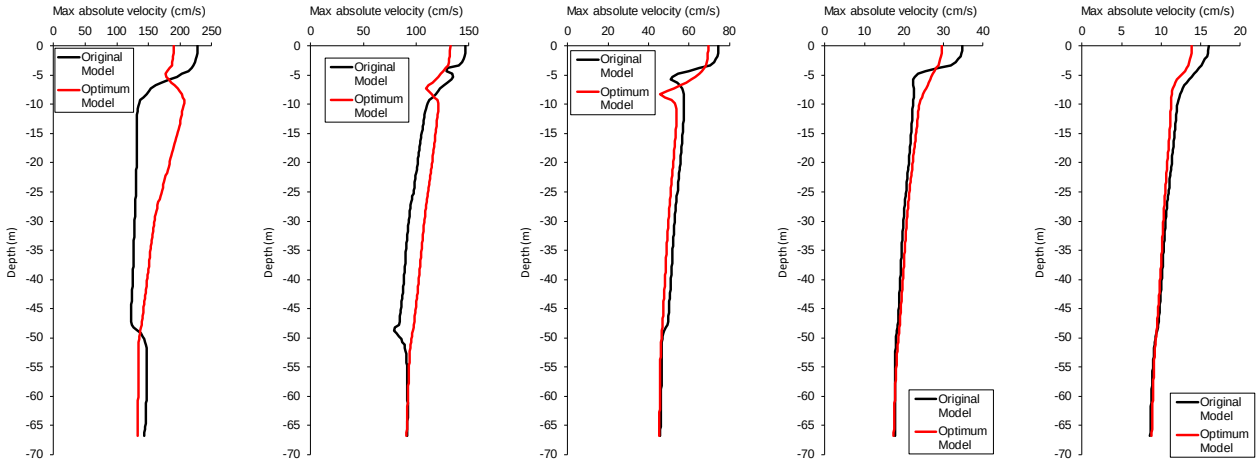
(3) Distribution of maximum shear strain in the depth directio

Fig.8 – Seismic response analysis results for different input ground motion intensities at GS-MSK-2 (1)



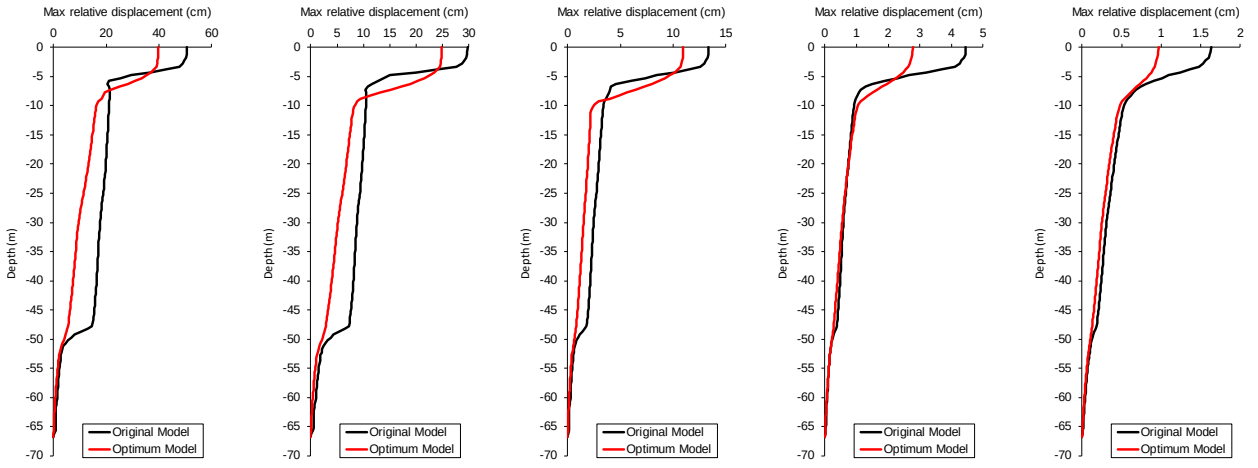
input × 1.5 (712.8cm/s²) input × 1.0 (475.2cm/s²) input × 0.5 (237.6cm/s²) input × 0.2 (95.04cm/s²) input × 0.1 (47.52cm/s²)

(4) Distribution of maximum absolute acceleration in the depth direction



input × 1.5 (712.8cm/s²) input × 1.0 (475.2cm/s²) input × 0.5 (237.6cm/s²) input × 0.2 (95.04cm/s²) input × 0.1 (47.52cm/s²)

(5) Distribution of maximum absolute velocity in the depth direction



input × 1.5 (712.8cm/s²) input × 1.0 (475.2cm/s²) input × 0.5 (237.6cm/s²) input × 0.2 (95.04cm/s²) input × 0.1 (47.52cm/s²)

(6) Distribution of maximum relative displacement in the depth direction

Fig.8 – Seismic response analysis results for different input ground motion intensities at GS-MSK-2 (2)

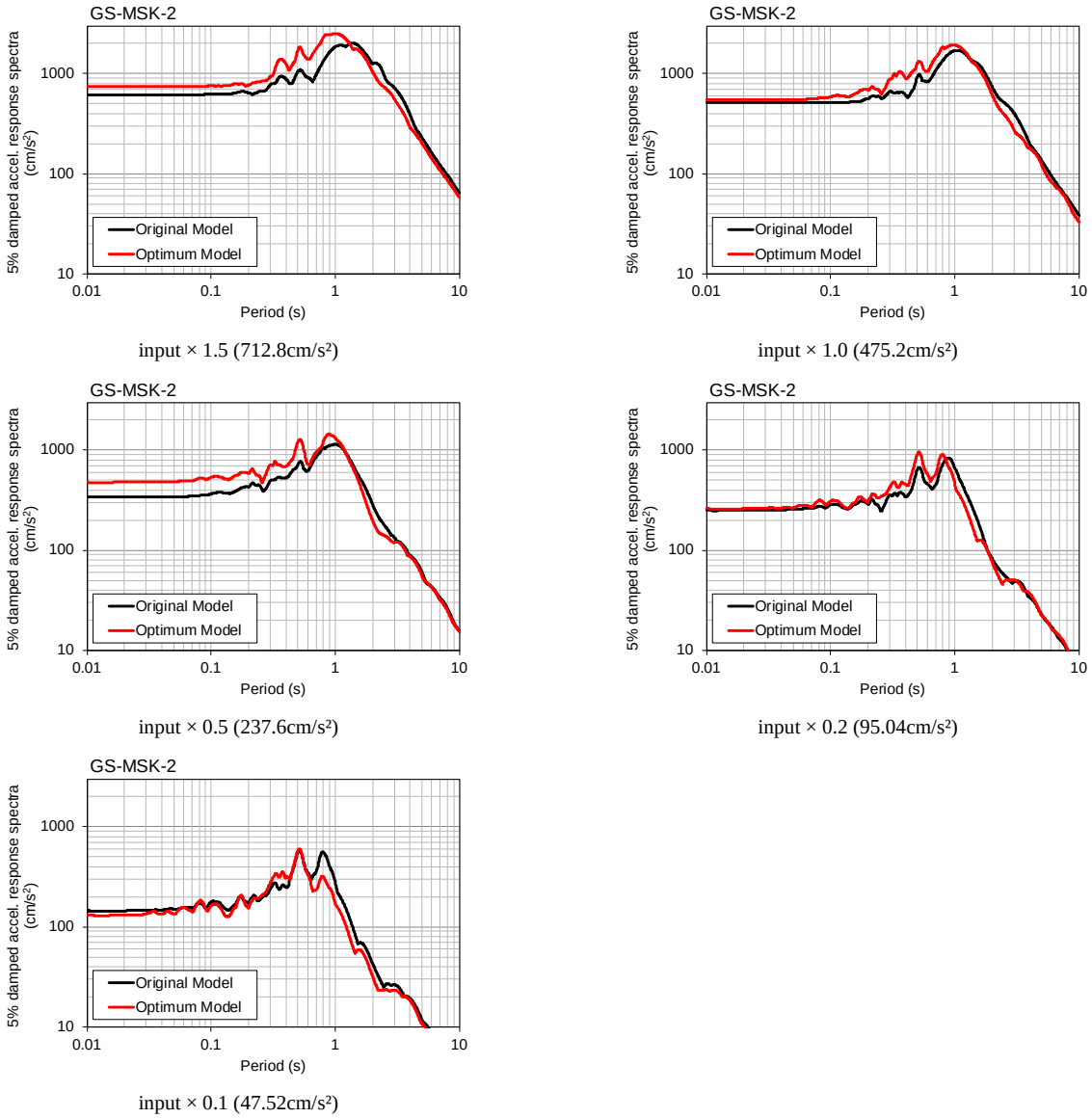
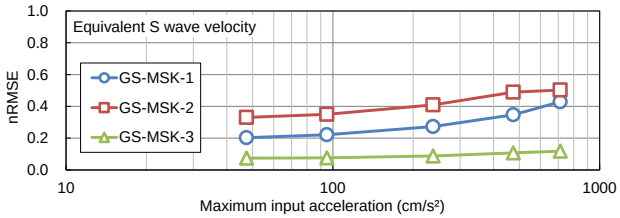


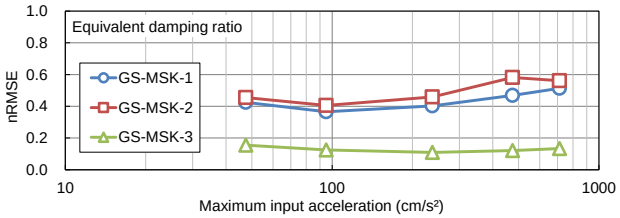
Fig.9 – Acceleration response spectra at ground surface for different input ground motion intensities at GS-MSK-2

The calculated values used to obtain nRMSE are output as follows. In the case of acceleration response spectra, a period between 0.01 seconds and 10 seconds was set as an interval of 0.01 seconds. For other quantities, the distance from the ground surface to the basement stiff layer was in 0.5 m increments.

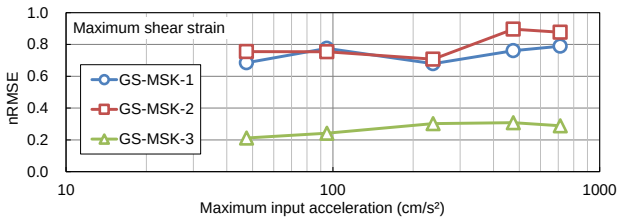
nRMSEs for different input ground motion intensities are displayed in Fig.10. The figure shows that the difference between the models is small at GS-MSK-3. This is because the optimal model at GS-MSK-3 is similar to the original model (see Fig. 4 (3)). On the other hand, the nRMSEs of GS-MSK-1 and GS-MSK-2 are almost the same, but the nRMSEs in GS-MSK-2 are often the largest. Also, as the input ground motion increases, the difference between the models (i.e. nRMSE) tends to increase. In addition, for all physical quantities, the difference between models tends to increase as the input ground motion increases. In the intercomparison between the physical quantities, the differences between the models of the maximum absolute acceleration is larger than the others.



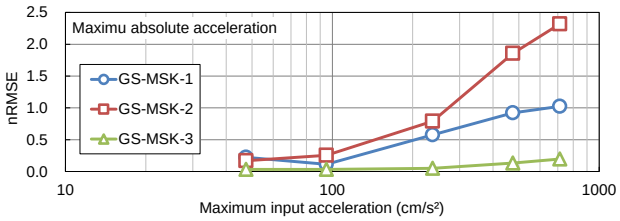
(1) Equivalent S wave velocity



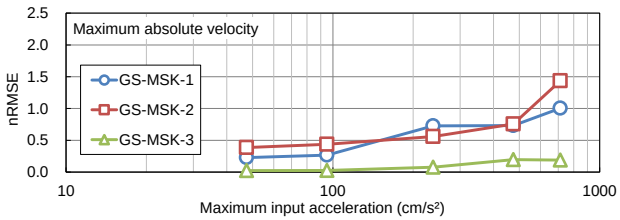
(2) Equivalent damping ratio



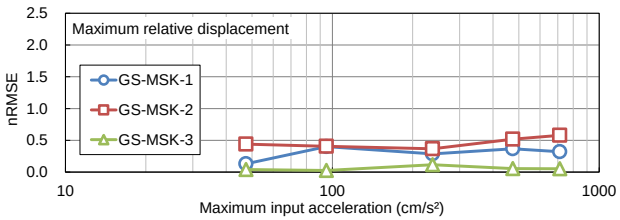
(3) Maximum shear strain



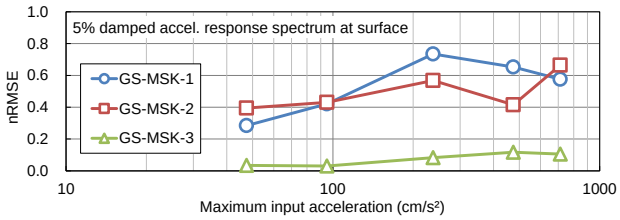
(4) Maximum absolute acceleration



(5) Maximum absolute velocity



(6) Maximum relative displacement



(7) 5% damped acceleration response spectra

Fig.10 – nRMSE for different input ground motion intensities

5. Conclusions

In this study, the ground structures were modelled by applying the fused lasso to the physical logging data, and the appropriate model selection was examined based on temporally seismic observation data. The findings obtained in this study are summarized below.

i. Using the fused lasso, ground structure models with reduced degree of freedoms were obtained from physical logging data. We also confirmed that simplified models can be obtained by increasing the value of the regularization parameter.

ii. It was confirmed that the H/V spectra theoretically calculated from the physical logging data in the affected area in Mashiki town corresponds well with the H/V spectra obtained from the coda wave of teleseismic events and the microtremor observation records.

iii. Using the degree of freedom of the fused lasso solution model and the residual sum of squares between the model theoretical value and the observation records, we tried to select appropriate models using the

Bayesian information criterion (BIC). A model that was relatively close to the original logging data was selected at GS-MSK-3, and a slightly different model was selected at GS-MSK-1 and GS-MSK-2.

iv. When the optimum ground structure model selected by BIC was applied to calculate the amplification characteristics of the SH wave by the equivalent linear analysis method, the difference between the original models based on the physical logging data and the optimum models selected from the sparse solutions increased as the amplitude level of the input acceleration increased.

6. Acknowledgements

We used the Japan Meteorological Agency Earthquake Monthly Report (catalog edition) for the hypocenter information. Also, the acceleration time history data of the strong-motion seismograph network (NIED KiK-net) installed by the National Research Institute for Earth Science and Disaster Resilience, doi:10.17598/NIED.0004 was employed. We would like to express our gratitude to all those involved.

7. References

[1] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight: Sparsity and smoothness via the fused lasso, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 67, No. 1, pp. 91-108. 2005.

[2] R. Tibshirani: Regression Shrinkage and Selection via the lasso, *Journal of the Royal Statistical Society: Series B (Methodlogical)*, Vol. 58, No.1, pp. 267-288, 1996.

[3] T. Kurita, Y. Shingaki and M. Yoshimi: Fused lasso solutions for physical logging data modeling, *Proceedings of the 39th JSCE Earthquake Engineering Symposium*, JSCE, B12-1394, 2019. (in Japanese)

[4] R.J. Tibshirani and J. Taylor: The solution path of the generalized lasso, *The Annals of Statistics*, 39, 3, 1335-1371, 2011.

[5] M. Yoshimi, Y. Hata, H. Goto, T. Hosoya, S. Morita, and T. Tokumaru: Borehole exploration in heavily damaged area of the 2016 Kumamoto earthquake, Mashiki town, Kumamoto, *Proceedings of the Japanese Society for Active Fault Studies 2016 Annual Meeting*, p. 17, 2016. (in Japanese)

[6] M. Yoshimi, H. Goto, Y. Hata, N. Yoshida: Nonlinear site response at the worst-hit area of the 2016 Kumamoto earthquakes in the Mashiki Town, Kumamoto, Japan, DPRI Annual Meeting 2017, Kyoto University, No. A05, 2017. (in Japanese)

[7] Y. Shingaki, M. Yoshimi, H. Goto, T. Kurita, K. Sato, T. Hosoya, Y. Arai and S. Morita: Physical and dynamic properties of the volcanic ash soil in the heavily damaged site of the 2016 Kumamoto earthquake, *Journal of JSCE A1*, 73, 3, 552-559, 2017. (in Japanese)

[8] H. Arai and K. Tokimatsu: S-wave velocity profiling by inversion of Microtremor H/V spectrum, *Bulletin of the Seismological Society of America*, Vol. 94, No. 1, pp. 53–63, 2004.

[9] G. E. Schwarz: Estimating the dimension of a model, *The Annals of Statistics*, 6, 2, 461-464, 1978.

[10] T. Kurita: Nonlinearity amplification of subsurface ground at KiK-net Mashiki site during the 2016 Kumamoto earthquake, *Journal of JSCE A1*, 73,4, pp. I_74-I_82, 2017. (in Japanese)

