

Identification of spatial distribution for ground stiffness by data fusion of geotechnical and geophysical data

Shin-ichi Nishimura, Toshifumi Shibata

Faculty of Environmental, Life and Natural Science and Technology, Okayama University, Japan, theg1786@okayama-u.ac.jp

ABSTRACT: The surface wave method (SWM) and the screw weight sounding (SWS) are employed as a geophysical exploration method and a geotechnical test, respectively to identify the spatial distribution of the stiffness of an earth-fill dam in the present study. The ensemble Kalman filter (EnKF) is used as a data assimilation technique to fuse two types of data. It can estimate the spatial distribution of the Young's modulus as the stiffness of a geo-structure by assimilating the travel time to the first arrival of the surface waves. By the ensemble data assimilation, the measured data from the SWM is applied to simultaneously estimate the Young's modulus and evaluate the uncertainties. The SWS results are employed as the prior information to generate the initial ensemble through the sequential Gaussian simulation (sGs). Proposed method has been applied to the actual data of the SWM and the SWS measured at an earth-fill dam site. Consequently, it has been clarified the proposed approach could identify the appropriate spatial distribution of Young's modulus.

KEYWORDS: surface wave method; screw weight sounding; data fusion; data assimilation; ensemble Kalman filter.

1 INTRODUCTION

Although boring and standard penetration tests are usually conducted to examine the inside of geo-structures, the interval of the test points must be sparse due to low efficiency and the possibility of damage to the structures. Consequently, weak locations may be overlooked. To deal with these shortcomings, the screw weight sounding test (SWS) (JGS 2015) is employed in this study as the static sounding method for obtaining detailed information for the ground stiffness on the inside of an earth-fill and the underneath ground. The SWS is advantageous in that it makes short-interval exams possible, which are highly efficient, and it causes little damage to the structures. Additionally, the surface wave method (SWM), which is one of the geophysical survey methods, is employed to obtain information for the ground stiffness of wider area, since it is a much more efficient approach. To examine wide area of an earth-fill efficiently and accurately, two methods, namely, the SWS and the SWM, are theoretically fused in this study.

Although the sounding results are reliable to estimate the stiffness of the ground, only the information of the point estimate is available. While, the geophysical exploration can present the spatially averaged information of the stiffness. In this study, two kinds of the test results are fused by the data assimilation technique, and the technique is applied to the identify the spatial distribution, namely random field of the stiffness. The ensemble Kalman filter (EnKF) is employed as a data assimilation technique. It can estimate the spatial distribution of the Young's

modulus as the stiffness of the ground by assimilating the travel time to the first arrival of the surface waves (Ren *et al.* 2022).

For the SWM, almost traditional inversion processes lead to a deterministic spatial distribution of the S-wave velocity (Jongmans and Demanet 1993). That means the uncertainty of the estimation cannot be evaluated. While, by the ensemble data assimilation, the measured data from the SWM is applied to simultaneously estimate the expected valued of the parameters and the uncertainties. The SWS results are used as the prior information to generate the initial ensembles of the ground stiffness through the sequential Gaussian simulation (sGs) (Deutsch and Journel 1992).

In this study, the proposed approach is applied for the test results of the SWS and the SWM conducted at an earth-fill site. Finally, it is verified that the identification of the parameter field, namely, the spatial distribution of Young's modulus is possible, and the uncertainties of the identified parameters can be reduced by the assimilation.

2 SURFACE WAVE METHOD

The measurements obtained by the SWM are utilized for the assimilation. According to the study of Hayashi *et al.* (2001), the geophones in this study are arranged at 2 m intervals, and the shot points are placed between the geophones and at the ends of them, so that the number of shot points is 1 more than the geophones. During the in-situ tests, a sledgehammer is used to generate surface waves by striking it against the shot points, and the shot-

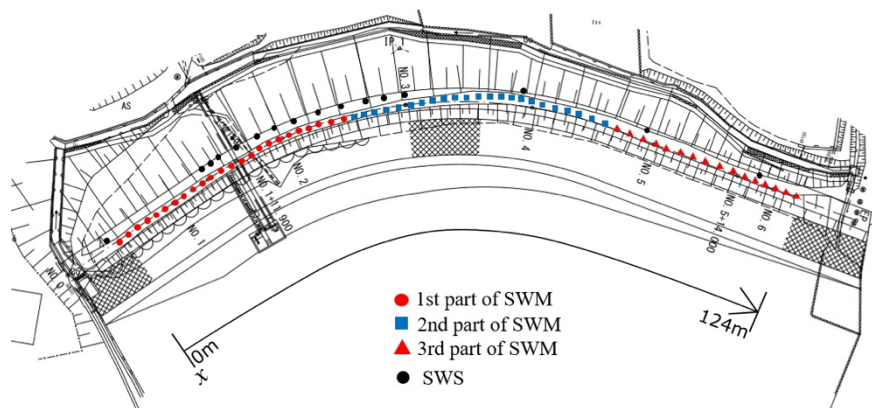


Figure 1. Plan view of SWS and SWM test locations.

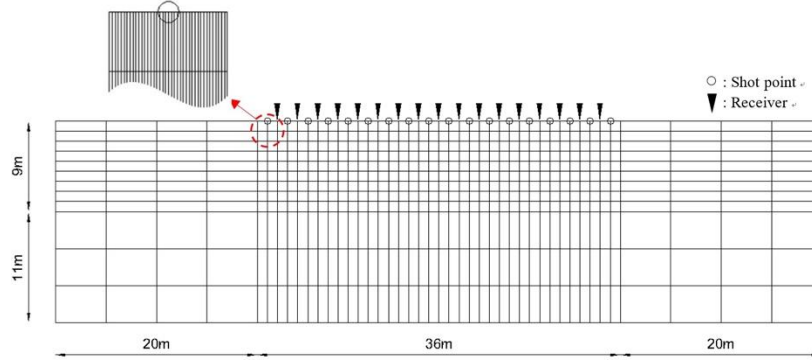


Figure 2. Example of mesh design of FEM model.

gather obtained by recording the waveforms is used as the measurement data for the corresponding times.

In the data assimilation to analyze the surface wave, the finite element method (FEM) is employed. The FEM allows the boundary conditions of different models to be taken into consideration, and can introduce the heterogeneity of the ground into the models easily. The first arrival time of the surface waves is used as the measurement in this study, which allows for a model with relatively large grids to be used, compared with the case that the frequency of the wave is used. As the noise in realizations can be eliminated with appropriate boundary conditions, the first arrival time of the analysis can be determined automatically.

3 SCREW WIGHT SOUNDING TEST

The spatial distribution of the stiffness parameters of decrepit earth-fill dams is discussed, and an identification method for the distribution is proposed. Generally, the identification of spatial correlation of the soil parameters is very difficult, since the number of the sampling points is limited to model the spatial variability. The difficulty can be solved by the sounding tests, since they can be conducted at widely spread points. Although the strength of earth-fill dams is generally predicted from standard penetration test (SPT) N -values, screw weight sounding (SWS) test is employed in this research as a simpler method to obtain the spatial distribution of the N -values.

The relationship between the SPT N -value and the SWS N -values is modelled (Inada 1960);

$$N_{SWS} = 0.067N_{SW} + 0.002W_{SW} \quad (1)$$

where N_{SWS} is the N -value derived from SWS, N_{SW} is the number of half rotations and W_{SW} is the total weight of the loads (N).

Next, the N -values are required to be transformed into Young's modulus. A roughly suitable model which makes the measurements of initial spatial distribution of the Young's modulus get close to the real measurements by the data assimilation could be chosen for convenience. The transformation model used in this study is

$$E = cN \quad (2)$$

where c is a constant transformation coefficient. It is common for c to be 700 or 2800 (e.g. for borehole load tests or horizontal load tests).

4 IDENTIFICATION OF RANDOM FIELD

4.1 Identification method

The N -value distribution derived from the SWS tests is spatially interpolated with the Monte Carlo method (MCM). The random field theory has been used to depict the spatial variability of soil

properties. Nishimura *et al.* (2016) employed the geostatistical method with data from cone penetration tests to create a reasonable N -value field. The spatial distribution of the ground in this section is created by the analogous method. The data from the SWS tests, taken from the same site as the SWM tests, are used to establish a statistical model for the N -value field (the spatial distribution of SWS N -value).

Figure 1 shows the plan view and the SWS test points of the studied site of the earth-fill dam located in Okayama. The determined trend function of $\log N$ is:

$$m = 0.912 - 0.004x - 0.028z + 0.00003x^2 + 0.013z^2 + 0.0019xz \quad (3)$$

The covariance function is presented by Equation (4). The horizontal correlation distance is identified as 15.7 m, while the vertical correlation distance is 1.24 m respectively.

$$C = [C_{ij}] = N_e \sigma^2 \exp\left(-\frac{|x_i - x_j|}{15.7} - \frac{|z_i - z_j|}{1.24}\right)$$

$$\sigma = 0.383$$

$$N_e = 1 \quad \left(\begin{array}{l} |x_i - x_j| = 0, |z_i - z_j| = 0 \\ |x_i - x_j| \neq 0, |z_i - z_j| \neq 0 \end{array}\right) \quad (4)$$

$$N_e = 0.593 \times 0.690 \quad \left(\begin{array}{l} |x_i - x_j| \neq 0, |z_i - z_j| = 0 \\ |x_i - x_j| = 0, |z_i - z_j| \neq 0 \end{array}\right)$$

$$N_e = 0.593 \quad \left(\begin{array}{l} |x_i - x_j| \neq 0, |z_i - z_j| = 0 \\ |x_i - x_j| = 0, |z_i - z_j| \neq 0 \end{array}\right)$$

$$N_e = 0.690 \quad \left(\begin{array}{l} |x_i - x_j| \neq 0, |z_i - z_j| \neq 0 \\ |x_i - x_j| = 0, |z_i - z_j| \neq 0 \end{array}\right)$$

Based on this statistical model, the random fields of N -values are generated by the sGs and converted to Young's modulus fields.

4.2 Ensemble Kalman filter

The inversion process of the EnKF consists of the evolution step and the update step. In the evolution step, a group of numerical models (the ensemble) with different initial states are used to forecast the evolution of a real system to calculate those measurements. Then, the states of the numerical models can be calibrated by the difference between the measurements of the numerical models and those of the real system in the update step. The EnKF is successful in many high-dimensional, nonlinear, and non-Gaussian application (Katzfuss, et al. 2016). In the present study, the finite element method (FEM) is used to solve the numerical models and predict the measurements.

The updating scheme of the EnKF is presented by:

$$\hat{x}_t^{(i)} = \tilde{x}_{t|t-1}^{(i)} + K_t \left(y_t - \tilde{y}_t^{(i)} \right), \quad i = 1, \dots, N \quad (5)$$

where $\tilde{x}_{t|t-1}^{(i)}$ is a set of forecast state vectors, which consists of all the state variables that needed to be estimated at time step t –

1 and $\hat{x}_{t|t}^{(i)}$ is a set of updated state vectors at time step t . i is the ensemble number, and N is the number of ensemble members. K_t is so called Kalman gain at the time step t . The state vectors here actually refer to the joint state vectors, which include the m parameters to be estimated (Young's modulus) and the n states to be assimilated (first arrival time of the surface waves). y_t is the measurement vector (n -dimensional), which consists of all the measurements at time step t , and $\hat{y}_t^{(i)}$ is the measurement vector of the i th ensemble.

According to Equation (5), not only Young's moduli, but also the first arrival time of the surface waves, are calibrated in the updating process. The updated parameters are propagated, and the state is recalculated from the simulation, which is presented by the function f in Equation (6), namely, the FEM, and at is the first arrival at each time step.

$$\mathbf{a}_t = f(\mathbf{E}_t) \quad (6)$$

In the updating process, the EnKF may lead to unrealistic geological parameters, for instance, a negative Young's modulus. Consequently, the logarithm (base 10) of Young's modulus is updated, and then all the updated Young's moduli which will be used in the simulation are positive. Accordingly, an attenuation factor β , greater than zero and less than one, is applied to reduce the error covariance matrix, which is determined by the measurements and the increments in measurements. A rational size of β depends on the numerical model (i.e., Equation (6)). It can be determined by trial and error so that the increments in the state vectors in each updating process are sufficiently small. The joint state variable being updated is:

$$\mathbf{x}_t = [\log \mathbf{E}_t, \beta \mathbf{a}_t] \quad (7)$$

The attenuation factor β is set to be 0.005 in this study.

5 APPLICATION RESULT FOR STUDIED SITE

5.1 Analytical model

In this section, a model to simulate SWM tests for the studied site, is utilized. The FEM model is shown in Figure 2. Corresponding to the arrangement of the geophones, there are 17 measurement points. There are 18 shot points in total, and the shot points are struck sequentially. Each hit is regarded as a time step, in which the EnKF should be updated. The height of the grid used for the analysis is 1 m, with a total height of 9 m. Within a radius of 1 m of a shot point, the length of the grid is set to be very small (0.05 m) in the horizontal direction, and the rest of the grid is 1 m in length. It is necessary to resort to this mesh division for removing the parasitic oscillation in the waves.

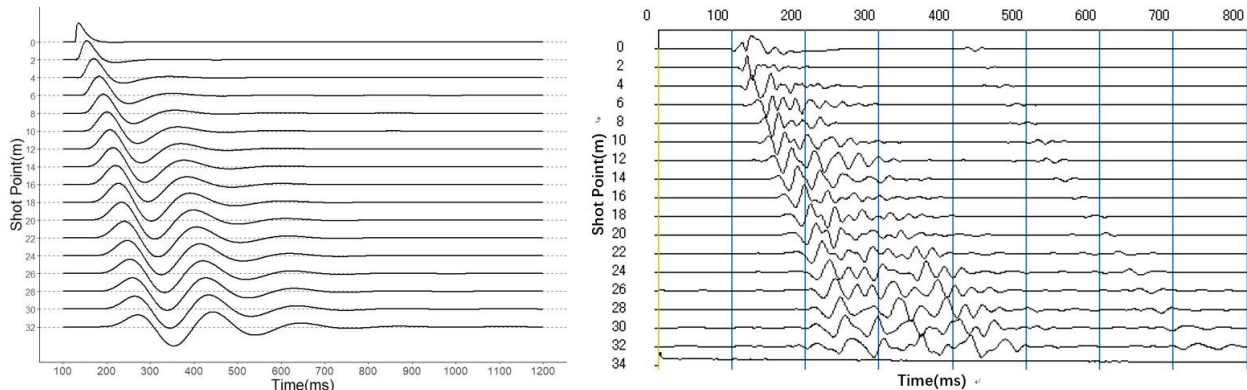


Figure 3. Example of simulation of surface wave propagation (left) and real measurement (right)

As such, the grid of the FEM model is rebuilt at each time step; however, a static grid is necessary for applying the EnKF updating. Therefore, the grid created for the sGs during the formation of the initial ensemble is used in the following updating, and the parameters of the elements for the FEM model in the next time step are determined by applying the updated grid. The Young's modulus of every element in the FEM model equals the average Young's modulus of every element in the updated grid, which falls within the search square centered at the location of each element. The horizontal length of the search square is 1 m in this experiment.

In a homogeneous medium, shear wave velocity V_s is given by:

$$V_s = \sqrt{\frac{E}{2(1+\nu)\rho}} \quad (8)$$

where ν is Poisson's ratio, E is Young's modulus, and ρ is the mass density. Compared with Young's modulus, which may change in a large range, the ranges of Poisson's ratio (0.3 ~ 0.4 for sandy soil) and mass density (19 kN/m³) for most of the sandy soil are relatively small. Therefore, the geological parameters, other than Young's modulus, are assumed to be known in order to limit the degrees of freedom of the unknowns. Poisson's ratio is set to be 0.35, and the unit weight is set to be 19 kN/m³. The time increment for one step of the FEM is 0.001 s.

Figure 3 gives an example of a simulation of the surface wave propagation. It can be seen that the characteristics of the waveform roughly match those of a real surface wave, but the calculated frequency of the wave is lower than the measured surface wave due to the grid size. However, the first arrival time of the waves is less influenced by the grid size or the frequency.

5.2 Analytical result

The SWM tests were implemented on the studied site, consisting of three parts as shown in Figure 1. The length of the first two parts is 46 m, and each part is arranged with 24 geophones and 25 shot points. The length of the third part is 32 m. For the relationship between the sites of the SWS tests and those of the SWM tests, the first part of the waveform data obtained from the SWM tests corresponds to the section from 0 to 46 m (1st part) in Figure 1. The second part corresponds to the section from 46 to 92 m (2nd part), and the third part corresponds to the section from 92 to 124 m (3rd part). The FEM model presented in Figure 2 corresponds to the 3rd part. The plan view of the SWM tests is shown in Figure 1 and an example of the waveform shown in Figure 3 corresponds to the 3rd part. To improve the accuracy of the estimates, the maximum penetration depth in the SWS tests is utilized to distinguish the dam and the bedground, which is treated as prior information for the EnKF. The portion below this maximum penetration depth is considered to be hard bedground.

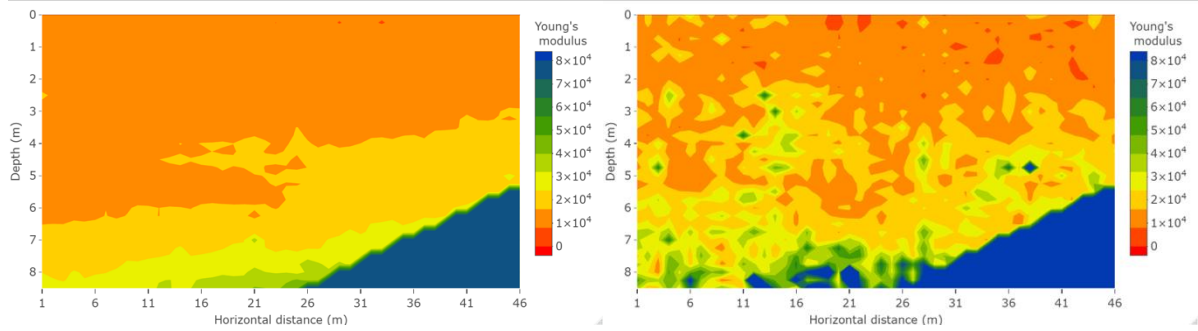


Figure 4. Mean of Young's moduli in 1st part at initial step (left) and final step (right).

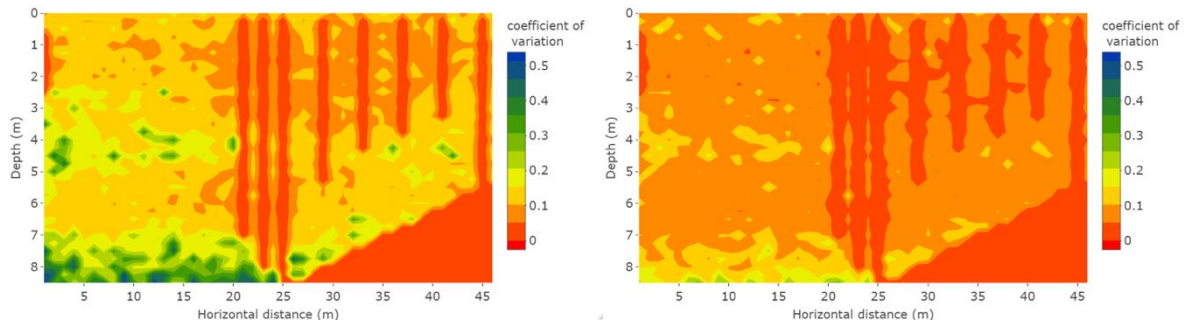


Figure 5. Coefficient of variation of Young's moduli in 1st part at initial step (left) and final step (right)

In this section, the Young's modulus of the bedground part is fixed to be five times the average of the dam part. The transformation coefficient c is assumed to equal 2800 in presented simulations.

Figures 4 show the spatial distributions of expected value of the Young's moduli updated by the EnKF. Only the Young's moduli at the initial and final time steps are shown in the figures. The large blue portions in Figures 4 represent the bedground. Not only the initial ensemble and the measurements, which are assimilated, but also prior information on the stratum, affect the estimated spatial distribution very strongly.

Figure 5 shows the spatial distribution of the coefficients of variation. The spatial distribution of the coefficients of variation represents the uncertainties of Young's modulus. As such, the probability distribution of the Young's modulus on each grid cell is described by the ensemble spread and the ensemble mean. Aside from the bedrock, there are some places in Figure 5, where the coefficient of variation is 0. These places are actually the locations of the SWS tests, and the parameters on these corresponding grid cells are treated as true values when random fields are generated using sGs. As for the other regions, it can be seen that the uncertainties decrease substantially after the assimilation of the EnKF.

6 CONCLUSIONS

- (1) The proposed method was applied for the measurement data from SWM tests conducted on an earth-fill dam located in Okayama, Japan. As prior information, the data from the SWS tests were used to distinguish the bedground and the dam, and to introduce the spatial variability of the Young's moduli.
- (2) The first arrival time of artificially excited surface waves was assimilated to estimate the Young's modulus of an earth-fill dam. The spatial distribution of the Young's moduli could be identified.
- (3) In the present study, it has been found that the EnKF is able to evaluate the uncertainties of the parameters. With the assimilation, the coefficient of variation was seen to have substantially decreased. The ability of the EnKF to reduce the uncertainties of the parameters has been verified.

7 ACKNOWLEDGEMENTS

This work was partly supported by JSPS KAKENHI Grant Numbers 20H00442 and 24H00534. Authors greatly appreciate the contribution of Dr. Yuxiang Ren for this research project.

8 REFERENCES

- Deutsch, C. V. and Journel, A. G. 1992. Geostatistical Software Library and User's Guide, Oxford University Press.
- Hayashi, K., Suzuki, H., and Sito, H. 2001. Surface wave method using artificial sources. Development and application to civil engineering investigations. *Oyo Tech Rep.* 21, 9-39.
- Inada, M. 1960. Usage of Swedish Weight Sounding Results. Tsuchi-to-Kiso, *Journal of JSSMGE* 8(1), 13-18. (in Japanese)
- Japanese Geotechnical Society. 2015. JGS Standards, Geotechnical and Geoenvironmental Investigation Methods, *JIS A 1221:2020*. JGS.
- Jongmans, D. and Demanet, D. 1993. "The importance of surface waves in vibration study and the use of Rayleigh waves for estimating the dynamic characteristics of soils. *Eng Geol* 34(1-2): 105-113. Doi.org/10.1016/0013-7952(93)90046-f
- Katzfuss, M., Stroud, JR., and Wikle, CK. 2016. "Understanding the ensemble kalman filter." *Am Stat.* 70(4), 350-357. <https://doi.org/10.1080/00031305.2016.1141709>
- Nishimura, S., Shibata, T., and Shuku, T. 2016. "Diagnosis of earth-fill dams by synthesized approach of sounding and surface wave method", *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, 10(4), 312-319, DOI: 10.1080/17499518.2016.1197406
- Ren, Y., Nishimura, S., Shibata, T. and Shuku, T.: Data assimilation for surface wave method by ensemble Kalman filter with random field modeling 2022, *Int J Numer Anal Methods Geomech.* 46(15), 2944-2961. DOI: 10.1002/nag.3435