Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

doi: 10.3850/978-981-18-5182-7_17-007-cd

ISGSR 2022

Inverse Process of Multichannel Analysis of Surface Wave by Using Ensemble Kalman Filter

Yuxiang Ren[1] and Shinichi Nishimura[2] and Toshifumi Shibata[3] and Takayuki Shuku[4]

[1]Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan.
E-mail: py3g9cvm@s.okayama-u.ac.jp
[2]Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan.
E-mail: theg1786@okayama-u.ac.jp
[3]Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan.
E-mail: tshibata@cc.okayama-u.ac.jp
[4]Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan.
E-mail: shuku@cc.okayama-u.ac.jp

Abstract: The Young's modulus is a common and critical soil property. It is difficult to infer the spatial distribution of Young's modulus on a large ground, owing to the uncertainty. In this study, a method using ensemble Kalman filter (EnKF) to solve the inverse problem in the multichannel analysis of surface wave and estimate the spatial distribution of Young's modulus which includes quantified uncertainty is presented. The statistic model derived from other investigation data could be integrated into the inverse process to increase the accuracy of estimate by using sequential Gaussian simulation (sGs) to generate a reasonable initial ensemble. The practical effectiveness of this framework is verified by numerical experiments using synthetic and realistic data from an earth-fill dam. The impact of the initial ensemble and the two update schemes (stochastic and deterministic updates) is discussed through the experiments as well.

Keywords: data assimilation; sequential Gaussian simulation; surface wave method; ensemble Kalman filter; inverse problem.

1 Introduction

The surface wave method (SWM) is a prevalent in-situ investigation. Compared with other destructive tests like the standard penetration test (SPT), the SWM can be implemented quickly and conveniently. Because the velocity of the surface wave is highly relevant to the velocity of the shear wave, many inversion processes of SWM lead to a spatial distribution of the shear wave velocity. The shear wave velocity, which relates to the stiffness of the medium, is an important property of subsoil. But it is hard for commonly used inversion processes like non-linear least squares to quantify uncertainties of its result which is a single estimate in a position. The ensemble data assimilation built on top of Monte Carlo methods is used in this study as an inversion process in the SWM to evaluate the stiffness with quantified uncertainty.

Ensemble data assimilation methods such as ensemble Kalman filter (EnKF) (Evensen 1994) and particle filter utilize samples to approximate the probability distribution of the model states. The EnKF can be used to deal with high-dimensional and nonlinear problems. The fact that the distribution of state is assumed to be Gaussian in EnKF gives rise to a low computational cost. The number of required samples in EnKF is far less than the number in the particle filter. Therefore, the EnKF and its variants are almost the only way to approximate many realistic and complex systems. Furthermore, compared with the elastic half-space hypothesis in traditional methods, in EnKF, boundary conditions of systems could be taken into consideration with ease through the numerical model which underlies the assimilation. There are two types of update schemes for the EnKF: stochastic updates introduced by Evensen and deterministic updates, e.g., the ensemble adjustment Kalman filter (EAKF) introduced by Anderson (2001).

The ensemble method is widely used in meteorology and oceanography and is gradually being applied in mechanical engineering, petroleum engineering, and geotechnical engineering. By assimilating the observations, the EnKF can modify the poorly known parameters and lead to best-guess estimates. There should be a correlation between the observation data and the parameters to be updated. For different models, appropriate observation data must be carefully chosen. Although the EnKF is used in many other disciplines as described above, few studies applied the EnKF to infer the stiffness of subsoil with the data of the surface wave. It is necessary to evaluate the validity of the method in this situation. Therefore, this work is a proof of concept for the new inversion process of SWM based on EnKF.

2 Methodology

The update step of inversion process is the focus of EnKF. The deterministic methods which belong to the family of square root filters are more accurate than stochastic methods for very small ensemble sizes. For the limited sample size in this work, the deterministic method that is also named ensemble adjustment Kalman filter is implemented. Another two-step algorithm of this deterministic method, which is based on an assumption that the prior distribution is Gaussian, is more efficient than addressing the covariance matrix directly. The two-step algorithm is used in this work.

This algorithm first updates the observations of ensemble members (samples). The k-th observation $y_k^{(i)}$ of a member is derived from the corresponding state vector of ensemble member $x^{(i)}$, where $i = 1, \dots, N$. Then, the observation increments, Δy_k, for all ensemble members should be calculated from updated observation $y_{u,k}$ and prior observation y_k. Assume that the observation y is Gaussian, i.e., $y \sim N(\bar{y}, \sigma^2)$, where $\bar{y}$ is mean and σ^2 is variance. Similarly, the observation from instrument is also Gaussian and the variance is σ_o^2. Then, the variance of updated observation σ_u^2 is given by

$$\sigma_u^2 = \left[(\sigma_k^2)^{-1} + \left(\sigma_{o,k}^2\right)^{-1}\right]^{-1} \tag{1}$$

and the mean of updated observation is given by

$$\overline{y_u} = \sigma_u^2 \left[(\sigma_k^2)^{-1}\overline{y_k} + \left(\sigma_{o,k}^2\right)^{-1} y_{o,k}\right] \tag{2}$$

Then, for each ensemble members the Δy_k for k-th observation can be calculated using

$$\Delta y_k = \sqrt{\frac{\sigma_u^2}{\sigma_k^2}}(y_k - \bar{y}_k) + \bar{y}_u - y_k \tag{3}$$

It should be noted that the Δy_k is a function of the variance of real-world observations which is very difficult for some tests to determine. So, there is a numerical experiment to study the influence of σ_o^2 in the following section. Once the increments of the observations are calculated, the increments of the state variables can be calculated by linear regression.

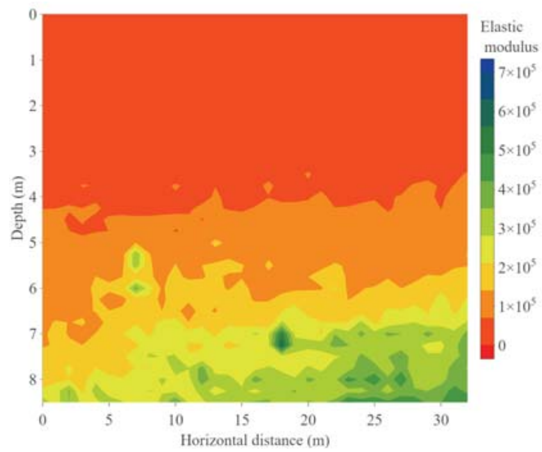

Figure 1. Reference spatial distribution of elastic modulus

3 Initial distribution

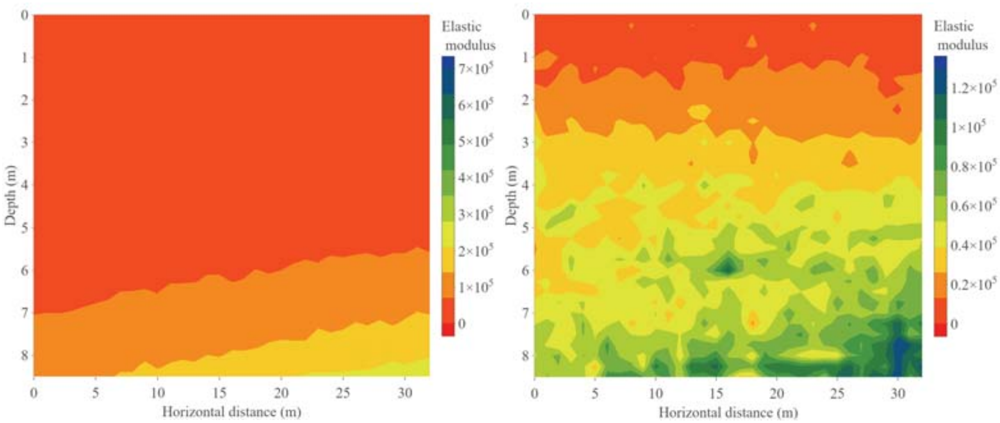

Figure 2. The initial state (left) and updated state (right).

For EnKF, the ensemble of initial states should be determined appropriately. Quite often, the initial states for all grid points are interpolated by observations at time point 0. in this section, the initial states are generated from the data of N-value of Swedish weight sounding test (SWS), which was taken on an earth-fill dam, by sequential Gaussian simulation(sGs). The realizations of sGs are treated as the samples of EnKF. The statistic model for sGs consist of a mean function and a covariance function suggested by minimizing Akaike information criterion estimation (NISHIMURA 2011) and semi-variogram, respectively. The mean function tells the trend of spatial distribution and the covariance function tells the correlation.

4 Numerical experiment

The effectiveness of the method is examined by numerical experiments. The proposed initial ensemble is based on the measurements from SWS tests implemented on an earth-fill dam located in Okayama, Japan. The numerical experiments are conducted as follows. First, the initial spatial distribution of elastic modulus is computed as described above to construct models. Then, the propagation of the surface wave is simulated by FEM as observation data. One of the random fields generated by the sGs is treated as the reference spatial distribution and its observations are treated as real-world observations for experiments. The rest of the samples which treated as ensemble members are biased to make the mean of elastic modulus different from the reference. The result of assimilation is evaluated by the residual sum of squares (RSS).

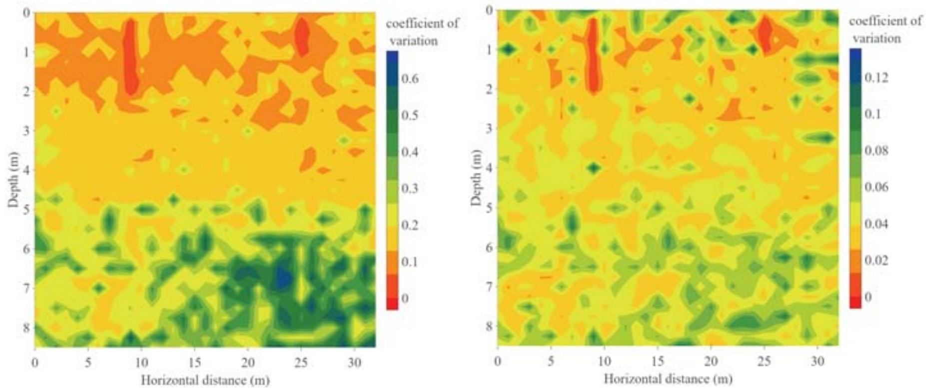

Figure 3. Coefficient of variation before (left) and after update (right).

The uncertainty is quantified by the coefficient of variation. The observation increments depend on the variance of real-world observations σ_o^2. In this work, it is difficult to give an exact σ_o^2 of the data acquired using geophones, because the σ_o^2 depends on both the error of instrument and background noise.

In the first case, the error of the instrument is assumed to be very small, i.e., $\sigma_o^2 = 0.1$. The reference spatial distribution of elastic modulus is shown in Figure 1. The prior spatial distribution (i.e., the initial condition) and updated one (updated by all observations) are shown in Figure 2. It should be noted that the spatial distributions here are ensemble mean. In this case, the RSS quickly and significantly reduces after the first update process. The observations here used to compute RSS are not the updated observations which are always closer to the reference than the prior observations, but the forecast observations derived from updated elastic modulus for next time or the initial condition for the first time. Therefore, the scheme can be considered efficient. Although it is difficult to adjust the ensemble to perfectly match the reference, the scheme can still significantly improve the accuracy. The

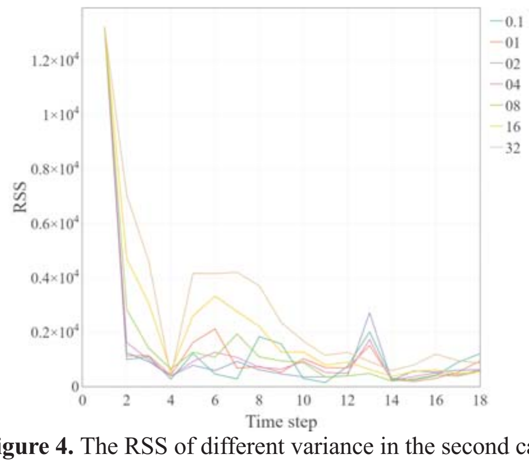

Figure 4. The RSS of different variance in the second case

Figure 3 shows the reduction of coefficient of variation for elastic modulus after data assimilation. This indicates the uncertainty has also been reduced significantly.

In the second case, instead of using fixed variance, the σ_o^2 is 0.1, 1, 2, 4, 8, 16 and 32 times of the ensemble variance, respectively, i.e., $\alpha = \sigma_o^2/\sigma^2 = 0.1,1,2,4,8,16,32$. The result of the assimilation experiment is shown in Figure 4. It can be seen that the larger the σ_o^2, the slower the rate of RSS reduction. Even there is no noise in reference observations which are calculated from the numerical model and the σ_o^2 should have been zero, the small α (i.e., 0.1 or 1) brings about a result that is not better than the others at the final step. When $\alpha = 32$, The RSS is greater than other RSSs in almost step. This indicates that the σ_o^2 should not be too large either. In practice, for a time-variant system, the excessive σ_o^2 will reduce the efficiency of the update process and increase the error of forecast observation in every step. For a time-invariant system in this work (the reference elastic modulus is constant), the relatively great σ_o^2 could still induce low RSS if there are enough observations for a sufficient quantity of updates.

5 Summary

In this study, the first arrival time of artificially excited surface waves was assimilated to estimate the elastic modulus of an earth-fill dam model. The initial ensemble is generated by sGs. The theoretical and practical effectiveness of this scheme is verified by numerical experiments. The difference between an updated parameter and reference and the uncertainty are both reduced by data assimilation. The reproducibility of observations is verified by the reduction of the RSS.

References

Anderson, J. L. (2001) An ensemble adjustment Kalman filter for data assimilation, *Monthly Weather Review*, vol. 129, no. 12, pp. 2884–2903, Dec.

Evensen, G. (1994). Sequential data assimilation with a nonlin-ear quasi-geostrophic model using monte carlo methods to forecast error statistics, *Journal of Geophysical Re-search*, vol. 99, no. C5, p. 10 143.

NISHIMURA, S. etc. (2011). Prediction of spatial distribution for n-values in earth-fill embankments. *Journal of Japan Society of Civil Engineers*, Ser. C (Geosphere Engineering) Volume 67 Issue 2 Pages 252-263.