Inversion for Dispersion Curve of Rayleigh Wave by Ensemble Kalman Filter

Okayama University, student member, ○Yuxiang Ren
Okayama University, regular member, Shinichi Nishimura
Okayama University, regular member, Toshifumi Shibata
Okayama University, regular member, Takayuki Shuku

1. Introduction

When propagating in inhomogeneous media, surface waves of different wavelengths will result in different phase velocities, a property known as dispersion. The surface wave method utilizes this characteristic of Rayleigh waves propagating in soil to calculate the shear wave velocity. Shear wave velocity is an important parameter in geotechnical engineering, and has a significant correlation with parameters such as N value and elastic modulus, which can be used to evaluate the stiffness of soil. After obtaining the waveform data of Rayleigh waves through geophones, the dispersion curve can be calculated in various ways. Inverting the dispersion curve can estimate the near-surface structure.

However, this type of inversion has strong non-uniqueness. In engineering practice, non-linear least squares methods are often used to estimate the soil structure. For example, Hayashi uses the Jacobian matrix to analyze the relationship between wavelength and detection depth. Considering convergence and robustness issues, this method ignored the non-uniqueness of the solution. Some studies also use Monte Carlo methods to handle inversion problems and quantitatively evaluate its uncertainty. However, this method requires many samples, and when we use it for two or even three-dimensional problems, it may lead to excessive computational complexity.

The ensemble Kalman filter (EnKF) introduced by Evensen utilizes samples to approximate the probability distribution of the state variables (here the physical properties of soil). 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. Therefore, the EnKF and its variants are almost the only way to approximate many realistic and complex systems. There are two types of update schemes for the EnKF: stochastic update and deterministic update[1)] which is used in this study. This work is a proof of concept for the new inversion process for the dispersion curve of Rayleigh wave.

2. Methodology

The inversion process of this study is as follows:

1.Establish dozens of random initial soil models.

2.Solve the forward problem for each model to obtain the corresponding dispersion curve.

3.Take the phase velocity corresponding to certain frequencies in the dispersion curve as the observed value, calculate the misfit between the observed values of each sample and the true model, and use the EnKF to adjust the sample models to obtain the estimates and their uncertainties.

The reliability of the forward problem directly affects the accuracy of the estimate. In order to apply this method to complex structures and different boundary conditions, such as earth-fill dams and slopes, finite element method (FEM) is used to simulate the propagation of Rayleigh waves, and then the phase-shift method is used to solve the dispersion curve. For reference, the tau-p transform is also used to solve the forward problem.

As for the inversion problem, A two-step algorithm of deterministic update of EnKF is used in this study. This algorithm first updates the observations of ensemble members (samples). Suppose that $\boldsymbol{x}$ is a unobserved state vector which can be used to describe the discretized ground. For example, $\boldsymbol{x} = \left(v_s^1, v_p^1, \cdots, v_s^n, v_p^n\right)$ where v_s and v_p are the velocity of S-wave and P-wave respectively. The time evolution model is $x_{t+1} = f(x_t) + v_t$, where v_t is Gaussian noise. The k-th observation $y_k^{(i)}$ of the i-th member is derived from the corresponding state vector of ensemble member $\boldsymbol{x}^{(i)}$ by forward model (i.e., $\boldsymbol{y} = Hx + w$, where $\boldsymbol{w}$ is Gaussian noise). All observations are assumed to be Gaussian. The updated variance $\sigma_{u,k}^2$ of the k-th observation is

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

where σ_k^2 is real observation and $\sigma_{p,k}^2$ is a priori variance. Then the mean of updated k-th observation is given by

$$\bar{y}_{u,k} = \sigma_u^2 \left[(\sigma_k^2)^{-1} y_k + \left(\sigma_{p,k}^2\right)^{-1} \bar{y}_{p,k}\right] \tag{2}$$

where y_k is real observation and $\bar{y}_{p,k}$ is mean of a priori observation. The updated k-th observation $y_{u,k}$ can be calculated using

$$y_{u,k} = \sqrt{\frac{\sigma_u^2}{\sigma_k^2}} \left(y_{p,k} - \bar{y}_{p,k}\right) + \bar{y}_{u,k} \tag{3}$$

where $y_{p,k}$ is a prior single observation.

The second step is to update the state for a member by global least squares fit. For the j-th state in i-th member, the increment due to k-th observation is given by

$$\Delta x_{k,j}^{(i)} = \hat{\beta}\left(\boldsymbol{y}_k, \boldsymbol{x}_j\right) \Delta y_k^{(i)} \tag{4}$$

$$\hat{\beta}(\boldsymbol{a}, \boldsymbol{b}) = \frac{\sum_{n=1}^{N} (a_n - \bar{a}) b_n}{\sum_{n=1}^{N} (a_n - \bar{a})^2} \tag{5}$$

where $\boldsymbol{y}_k$ and $\boldsymbol{x}_j$ are vectors which consist of all sample observations and states, respectively.

3. Numerical experiment

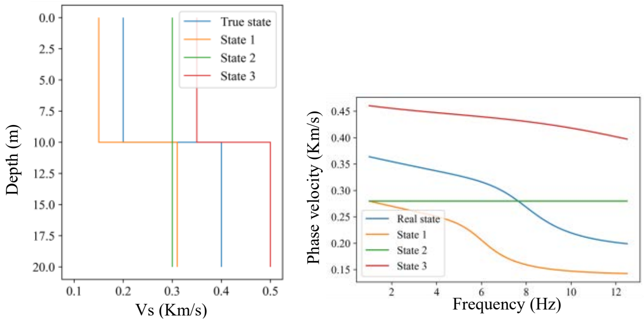
The numerical experiments in this study use a relatively simple model, i.e., a two-layered one-dimensional soil structure.

3.1. Using tau-p transform for forward problem

The tau-p transform introduced by McMechan and Yedlin (1981) can extract phase velocities dispersion from Rayleigh waves, and this method has been implemented in Computer Programs in Seismology (CPS) provided by Herrmann (2013).

In this study, we used part of the CPS code to forward calculate dispersion curves from soil models.

Figure 1(a) shows the true state of the geological and geotechnical structure, as well as different initial models. Except for the true model, the other initial states represent the mean of 100 samples. The variance is 0.025 times the mean, which represents the uncertainty of the initial parameters and should be carefully determined. If the uncertainty is too high, the adjustment range of the parameters may be too large during data assimilation, which can cause errors in nonlinear problems. Figure 1(b) shows the corresponding dispersion curves of these structures, taking the phase velocities at 3Hz and 20Hz as observations, which correspond to the Rayleigh wave velocities of the second (deep) and first (shallow) layers, respectively.



(a) Soil Models (b) Dispersion Curves

Fig.1 Different soil models and corresponding dispersion curves.

Figure 2 shows the model after data assimilation. It can be seen from Figure 2 that in this experiment, there is no requirement for an initial model that is reasonably accurate to a certain extent. The mean errors of the first and second layers are 5% and 4.2%, respectively, which are generally considered acceptable.

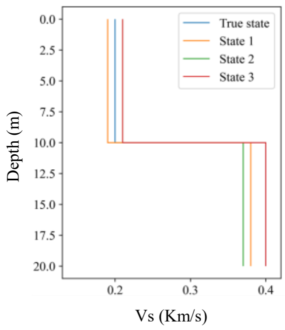


Fig. 2 Updated states.

3.2. Using FEM for forward problem

In order to handle complex structures in future studies, we attempted to assess the applicability of this method to FEM. We incorporated a viscoelastic boundary into the FEM model to absorb reflected waves. After simulating the waveforms using FEM, we transformed each waveform into the frequency-domain using Fourier transformation.

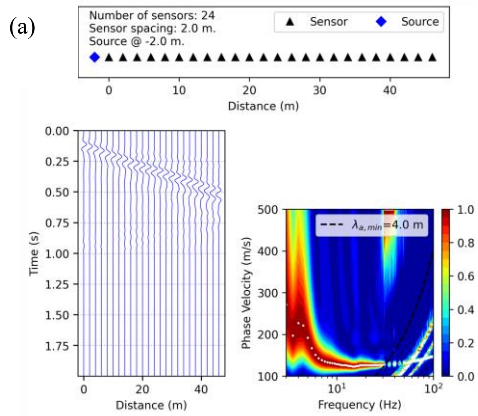
$$U(x,w) = \int u(x,t)e^{iwt}dt \quad (6)$$

where u is the waveform and x represents the offset of geophones. The dispersion curve for a given phase velocity is determined by a peak of $V(w,c)$:

$$V(w,c) = \int e^{\frac{iwx}{c}} U(x,w)/|U(x,w)|dx \quad (7)$$

where c is the phase velocity.

The SWprocess tool developed by Vantassel (2022) was used to process waveforms and calculate dispersion curves. Figure 3 shows the forward simulation result of a two-layer model, with shear wave velocity of 132 m/s and 241 m/s in the first and second layers, respectively. A Ricker wavelet was used as the source in this experiment. Figure 4 shows the inversion result. Due to the high computational cost of finite element method, only 16 samples were used. After data assimilation, the errors in the first and second layers were 8.6% and 0.7%, respectively, indicating that the inversion was effective in this experiment.



(b) Waveforms (c) Dispersion curve

Fig.3 Simulated waveforms and its dispersion curve. (a) shows the arrangement of geophones and source. (b) shows synthetic waveforms. (c) can be obtained by applying Fourier transformation to every trace in (b) and sum spectrums after applying offset-dependent phase shift.

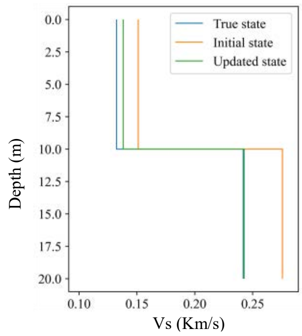


Fig.4 Inversion result for simulated waveforms.

4. Summary

This study used EnKF to invert the data of Rayleigh waves. A relatively simple one-dimensional model was used to verify the effectiveness of this method. The method is not sensitive to the initial model selection and has shown sufficient effectiveness for different forward models. It can be expected that this method will be used for more complex soil models or combined with other observation data (e.g., first arrival of P wave) for joint inversion.

References

(1) J. L. Anderson, "An ensemble adjustment Kalman filer for data assimilation," Monthly Weather Review, vol. 19, no. 12, pp. 2884–2903, Dec. 2001.