Joint inversion of receiver function and surface wave dispersion based on unscented Kalman Methodology

LongLong Wang

INSTITUTE OF GEOLOGY AND GEOPHYSICS, CHINESE ACADEMY OF SCIENCES

Key Points:
- Joint Inversion
- Unscented Kalman Inversion
- Receiver Function Inversion
- Surface Wave Inversion

Corresponding author: Wang Long Long. wanglonglong@mail.iggcas.ac.cn
Abstract
The Bayesian methodology is widely used for inversion analysis and uncertainty analysis. In this work, we focus on the joint inversion of receiver function and surface wave dispersion. Based on unscented Kalman methodology, we present a novel joint inversion framework (Innocent unscented Kalman Inversion, IUKI), which assumes the thickness and S-wave velocity at every artificial layer are both Gaussian random variables and can adaptively adjust their means and variances until convergence. This approach is derivative-free and can efficiently provide uncertainty estimations of models with noisy data. Furthermore, our method can also be easily extended as a generalized joint inversion framework of multimodal geophysical data. With comprehensive experiments, the proposed framework demonstrates superior performance in terms of accuracy and stability.

1 Introduction
The crustal and upper mantle velocity models are of great importance for understanding continental tectonics, thermal and compositional structures of the lithosphere (Shen et al., 2013). Because of the sensitivity of P-wave receiver functions to the strong gradients in elastic properties, there are many studies to use receiver function to imagine and inverse the S-wave velocity structure. But it is well known that RF can only result in a relative shear velocity, to better constrain the absolute shear velocity structure, surface wave dispersion has been added for joint inversion. Linearized inversions and Non-linear inversion based on Markov Chain Monte Carlo have been proposed successively to solve this problem.

In this paper, we introduce the Kalman inversion approach as the framework of joint inversion. Kalman inversion is derived from Kalman filter, which is first published to describe a recursive solution to the discrete-data linear filtering problem in 1960 (Kalman, 1960). Since that time, Kalman filters have been the subject of expansive research and application, particularly in the area of Signal Processing and Dynamic Systems. Kalman methodology can also be used for inversion from the optimization perspective when artificial dynamics based on state augmentation are constructed (Welch & Bishop, n.d.). Considering the essence of derivative-free and the connection with generalized Levenberg-Marquardt algorithm, Kalman inversion has been widely used as an optimization method for parameter estimation.

In the Kalman method and its variants, a vital operation performed is the propagation of a Gaussian random variable (GRV) through the system dynamics (Wan & Van Der Merwe, 2000). The unscented Kalman methods (UKF and UKI) use a deterministic sampling technique known as the unscented transformation (UT) (S. J. Julier & Uhlmann, n.d.; S. Julier, 2002) to solve this problem. By picking the sigma points deterministically and applying a quadrature rule, it provides an effective tool in signal processing, multi-sensor fusion, and parameter learning problems (Huang & Huang, 2021; Huang et al., n.d.; Wan & Van Der Merwe, 2000; Cao & Huang, 2021).

The remainder of the paper is organized as follows. In Section 2, we introduce our joint inversion framework based on unscented Kalman methodology. Section 3 is the test with synthetic examples and real data. Conclusions are offered in Section 4.

2 Method
We use the Thomson-Haskell matrix method to compute the teleseismic receiver function and SWD data (Herrmann, 2013). To get a better understanding of this joint inversion algorithm, we follow the work from Daniel Z. Huang and derive the UKI from Gaussian Approximation Algorithm (GAA) in subsection 2.1. Subsection 2.2 shows how to employ unscented transformation to well approximate Gaussian expectations and co-
variance which results in the UKI algorithm. In Subsection 2.3 we introduce the IUKI, which is a new joint inversion framework of multimodal geophysical.

2.1 Gaussian Approximation Algorithm

The GAA, originally in geostatistics, also known as Gaussian process regression, is a method of interpolation based on the Gaussian process that maps Gaussian to Gaussian and gives insight into the Kalman methodology (Huang et al., n.d.). The algorithm proceeds with two steps. In the first step we assume that \( \mu_n \approx N(\mu_n, C_n) \) and with a linear transformation \( \hat{\mu}_{n+1} \) is also Gaussian:

\[
\hat{m}_{n+1} = \mathbb{E}[\theta_{n+1}|Y_n] = m_n \quad \hat{C}_{n+1} = \text{Cov}[\theta_{n+1}|Y_n] = C_n + \Sigma_w. \tag{2.1}
\]

Then we can get the joint distribution of \( \{\theta_{n+1}, y_{n+1}\} | Y_n \):

\[
\mathcal{N}
\begin{bmatrix}
\hat{m}_{n+1} \\
\hat{y}_{n+1} \\
\hat{C}_{n+1}^\theta \\
\hat{C}_{n+1}^{\theta y} \\
\hat{C}_{n+1}^{yy}
\end{bmatrix}
\begin{bmatrix}
\hat{m}_{n+1} \\
\hat{y}_{n+1} \\
\hat{C}_{n+1}^\theta \\
\hat{C}_{n+1}^{\theta y} \\
\hat{C}_{n+1}^{yy}
\end{bmatrix}
\right), \tag{2.2}
\]

where

\[
\hat{y}_{n+1} = \mathbb{E}[y_{n+1}|Y_n] = \mathbb{E}[G(\theta_{n+1})|Y_n],
\]

\[
\hat{C}_{n+1}^{\theta y} = \text{Cov}[\theta_{n+1}, y_{n+1}|Y_n] = \text{Cov}[\theta_{n+1}, G(\theta_{n+1})|Y_n], \tag{2.3}
\]

\[
\hat{C}_{n+1}^\theta = \text{Cov}[\theta_{n+1}|Y_n] = \text{Cov}[G(\theta_{n+1})|Y_n] + \Sigma_v.
\]

Computing the conditional distribution of joint distribution to find \( \theta_{n+1}|Y_{n+1} \):

\[
m_{n+1} = \hat{m}_{n+1} + \hat{C}_{n+1}^{\theta y} (\hat{C}_{n+1}^{yy})^{-1} (y_{n+1} - \hat{y}_{n+1}),
\]

\[
C_{n+1} = \hat{C}_{n+1} - \hat{C}_{n+1}^{\theta y} (\hat{C}_{n+1}^{yy})^{-1} \hat{C}_{n+1}^{\theta y} T. \tag{2.4}
\]

When assuming all observations \( y_n \) are identical to \( y (Y_n = y) \) and cycle from 2.1 to 2.4, the Kalman inversion will be established to solve the inverse problem.

2.2 UKI

Recall that the vital operation performed in the kalman methodology is the propagation of a Gaussian random variable. Specifically speaking is how to approximates the integral in 2.3. Unscented kalman methodology approximates the integrals by deterministic quadrature rules (Huang & Huang, 2021).

**Definition 2.1** (Modified Unscented Transform). Let denote Gaussian random variable \( \theta \sim \mathcal{N}(m, C) \in \mathbb{R}^N, 2N\theta + 1 \) symmetric sigma points are chosen deterministically:

\[
\theta^0 = m \quad \theta^j = m + c_j [\sqrt{C}]_j \quad \theta^{j+N} = m - c_j [\sqrt{C}]_j \quad (1 \leq j \leq N_\theta), \tag{2.5}
\]

where \([\sqrt{C}]_j\) is the jth column of the Cholesky factor of C. The quadrature rule approximates the mean and covariance of the transformed variable \( G_i(\theta) \) as follows,

\[
\mathbb{E}[G_i(\theta) \approx G_i(\theta^0)] \quad \text{Cov}[G_1(\theta), G_2(\theta)] \approx \sum_{j=1}^{2N_\theta} W_j^c (G_1(\theta^j) - \mathbb{E} G_1(\theta)) (G_2(\theta^j) - \mathbb{E} G_2(\theta))^T \tag{2.6}
\]

Here these constant weights are:

\[
c_1 = c_2 \cdots = c_{N_\theta} = \sqrt{N_\theta + \lambda} \quad W_j^c = \frac{1}{2(N_\theta + \lambda)} (j = 1, \cdots, 2N_\theta),
\]

\[
\lambda = a^2 (N_\theta + \kappa) - N_\theta \quad \kappa = 0 \quad a = \min\left\{\sqrt{\frac{4}{N_\theta + \kappa}}, 1\right\}. \tag{2.7}
\]
The parameters are chosen as:
\[ \Sigma_v = 2\Sigma_\eta \text{ and } \Sigma_w = C_n, \] (2.8)
which guarantees that the algorithm can have a good estimate of the posterior probability through the converged mean and covariance (Huang et al., 2022).

2.3 IUKI

Inversion problems can be set as follows:
\[ y = \mathcal{G}(\theta) + \eta, \] (2.9)
where \( y \in \mathbb{R}^{N_y}, \theta \in \mathbb{R}^{N_\theta} \) and the map \( \mathbb{R}^{N_\theta} \rightarrow \mathbb{R}^{N_y} \) is a observational noise and assumed to drawn from a Gaussian with distribution \( \mathcal{N}(0, \Sigma_\eta) \). For the joint inversion, we divided them to the two parts and set:
\[
\begin{align*}
\mathbf{y} &= \begin{bmatrix} \alpha * y_{rf}, y_{swd} \end{bmatrix} \\
\mathbf{G} &= \begin{bmatrix} \alpha * \mathcal{G}_{rf}, \mathcal{G}_{swd} \end{bmatrix} \\
\mathbf{\eta} &= \begin{bmatrix} \alpha * \eta_{rf}, \eta_{swd} \end{bmatrix} \\
\mathbf{\theta} &= \begin{bmatrix} \theta_{vs}, \theta_{thickness} \end{bmatrix},
\end{align*}
\] (2.10)
where \( \alpha \) is selected according to the relatively data quality. We use the squared error cost function as the objective function:
\[
\Phi(\theta) = \frac{1}{2} \left\| \Sigma_\nu^{-\frac{1}{2}} (y - \mathcal{G}(\theta)) \right\|^2. \] (2.11)
For a joint inversion of receiver function and surface wave dispersion, the objective function is defined by the linear combination of misfits of the weighted receiver functions \( \Phi_{RF} \) and the \( \Phi_{SWD} \), using the L2 norm, thus takes the form:
\[
\Phi(\theta) = \alpha \Phi_{RF} + \Phi_{SWD}, \] (2.12)
where \( \alpha \) is selected according to the relative data quality, relative data length between RF and SWD.

3 Results

3.1 Synthetic Example

The synthetic RF and SWD are calculated with the method described in section 2. Refer to Bodin et al. (2012), SWD noise are generated by a diagonal covariance matrix \( \sigma_{true}^{SWD} = 0.012 \) and \( r_{true}^{SWD} = 0 \) while RF noise are generated submit an exponential correlation law with values \( \sigma_{true}^{RF} = 0.005 \) and \( r_{true}^{RF} = 0.92 \).

In this section, we assume that the initial model has 25 layers and the first three layers thickness means are all 1.5km, the fourth layer to the seventh layer thickness means is 2km and the remaining layer thickness mean are all 3km.

We use three known velocity models to evaluate our algorithm, the three models are: i) A 8 horizontal layers with a low S-wave velocity layer in the crust and a strong velocity increase at the Moho. The model is modified by a 6 layers model refer to Bodin et al. (2012); ii) GAr1 model which can be viewed as a smoother version than model i; iii) GAr3 model which has a slight difference in the upper crust compared with model ii.
Figure 1. Iterated unscented Kalman inversion of the simulated RF and SWD. (a) Posterior probability distribution for Vs at each depth. (b) Simulated RF data with the Gaussian random noise and the inversion result. (c) Simulated SWD data with Gaussian random noise and the inversion result.

Figure 2. Iterated unscented Kalman inversion. (a) The synthetic true velocity model is plotted as a red line, the color from light blue to dark blue shows the evolution of the final model through each iteration. (b) The optimization error $\frac{1}{2} \Sigma^{-\frac{1}{2}} (y - \mathcal{G}(\theta))^2$ at each UKI iteration for the joint inversion.
Figure 3. Separate and joint inversion results for the synthetic data. (a) SWD inversion. (b) RF inversion. (c) Joint inversion. (d)(e)(f) are separately the optimization error of (a)(b)(c).

Figure 4. Joint inversion with different initial model for Model ii)

Figure 5. Joint inversion with different initial model for Model iii)
Figure 6. Joint inversion of field data for station KIGAM. (a) The evolution of the final model through each iteration. (b) The posterior probability of layer velocity model. (c) Receiver function measurements. The red curve shows the RF calculated with the model that best fits the observed RF. (d) Fundamental model Rayleigh wave phase velocity dispersion measurements. The red curve shows the dispersion curve obtained with the model that best fits these data.

From Figure 1 and Figure 2, we can see that the algorithm can adaptively adjust the parameter (i.e., velocity and thickness) and fits the artificial observations well.

The result of joint inversion has been compared with the result of separate RF inversion and SWD inversion in Figure 3. Overall, the joint inversion gets the best parameter estimate compared with separate inversion of surface wave and receiver function.

The aforementioned experiments are all set in the same initial model, which has 25 layers. The impact of different initial models on the parameter estimate can also be tested (Figure 4, Figure 5). These numerical experiments demonstrate that discontinuities with sharp speed changes can be effectively recovered with a properly initial model. and different initial models differ in their ability to present details.

3.2 Field Case

Our algorithm is further tested using a real data set from a KIGAM station at Seoul National University. This data set is available from the web page: http://www.eas.slu.edu/eqc/eqc_cps/TUTORIAL/STRUCT/index.html. In this data set, the dispersion measurements and receiver functions at station KIGAM have been obtained and performed well (Herrmann, 2013). Recall that our goal is to illustrate that our methodology can get the model well that best fits observed SWD and observed RF, so we select one receiver function with a Gaussian factor $a=2.5$ and SWD with periods from 5-90s. The initial model thickness parameter is set to refer to the Figure 6a, which is similar to section 3.1 but adds more layers.

From Figure 6a can see that the iterative process has been cycled only 10 times and our algorithm can simultaneously adjust the thickness and S-wave velocity at each iteration. The results and their sensitivities are shown in Figure 6b. As the result shows, The Moho is apparent which has a strong discontinuity lies 37–42 km underground. Another strong discontinuity layer arises in the depth around 100-110km, we suspect that this is the LAB layer below the station. By the two layers, the structure has been divided into 3 parts. The first part (0-30km) has a low-velocity uppermost structure which
may relate to unconsolidated sediments or weathered exposed rocks. With a jump layer, the velocity structure tends to a value of about 3.8 km/s. The second part is relatively constant with a mean Vs = 4.2 km/s. The third part has a relatively high velocity of about 4.8 km/s. The Figure 6c, 6d are separately the inversion results of RF and SWD and our inversion model fit their individual data sets well.

4 Conclusions

The purpose of the current study was to introduce IUKI as an effective inversion framework that can perform joint inversion of RF and SWD. By performing in three synthetic data sets and a real data set, this inversion framework shows the outstanding ability for at least four reasons: 1. Derivative-free and non-intrusive. 2. Generally requires $O(20)$ iterations. 3. Naturally support joint inversion of RF and SWD. 4. Based on Bayesian methodology that can evaluate the uncertainty of the solution by posterior distribution efficiently.

Acknowledgments

References

Bodin, T., Sambridge, M., Tkalčič, H., Arroucau, P., Gallagher, K., & Rawlinson, N. (2012, February). Transdimensional inversion of receiver functions and surface wave dispersion: TRANSDIMENSIONAL INVERSION OF RF AND SWD. Journal of Geophysical Research: Solid Earth, 117(B2), n/a–n/a. Retrieved 2021-10-15, from http://doi.wiley.com/10.1029/2011JB008560 doi: 10.1029/2011JB008560

Cao, S., & Huang, D. Z. (2021, June). Bayesian Calibration for Large-Scale Fluid Structure Interaction Problems Under Embedded/Immersed Boundary Framework. arXiv:2105.09497 [cs, math]. Retrieved 2021-12-10, from http://arxiv.org/abs/2105.09497 (arXiv: 2105.09497)

Herrmann, R. (2013). Computer programs in seismology: An evolving tool for instruction and research.

doi: 10.1785/0220110096

Huang, D. Z., & Huang, J. (2021, April). Unscented Kalman Inversion: Efficient Gaussian Approximation to the Posterior Distribution. arXiv:2103.00277 [cs, math]. Retrieved 2021-12-10, from http://arxiv.org/abs/2103.00277 (arXiv: 2103.00277)

Huang, D. Z., Schneider, T., & Stuart, A. M. (n.d.). Unscated Kalman Inversion. 49.

Huang, D. Z., Schneider, T., & Stuart, A. M. (2022, January). Iterated Kalman Methodology For Inverse Problems. arXiv:2102.01580 [cs, math]. Retrieved 2022-02-06, from http://arxiv.org/abs/2102.01580 (arXiv: 2102.01580)

Julier, S. (2002). The scaled unscented transformation. 4555–4559 vol.6. Retrieved 2021-12-29, from http://ieeexplore.ieee.org/document/1025369/ (00000) doi: 10.1109/ACC.2002.1025369

Julier, S. J., & Uhlmann, J. K. (n.d.). A New Extension of the Kalman Filter to Nonlinear Systems. 12. doi: 10.1117/12.280797

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems.

Shen, W., Ritzwoller, M. H., Schulte-Pelkum, V., & Lin, F.-C. (2013, February). Joint inversion of surface wave dispersion and receiver functions: a Bayesian Monte-Carlo approach. Geophysical Journal International, 192(2), 807–836. Retrieved 2022-01-21, from http://academic.oup.com/gji/article/192/2/807/580799/Joint-inversion-of-surface-wave-dispersion-and doi: 10.1093/gji/ggs050
Wan, E., & Van Der Merwe, R. (2000). The unscented Kalman filter for nonlinear estimation. 153–158. Retrieved 2022-01-26, from http://ieeexplore.ieee.org/document/882463/ doi: 10.1109/ASSPCC.2000.882463

Welch, G., & Bishop, G. (n.d.). An introduction to the kalman filter. 16.