

Elastic full waveform inversion results uncertainty analysis: a comparison between the model uncertainty given by conventional FWI and machine learning methods

Tianze Zhang¹, Scott Keating², Kristopher Innanen¹

1 University of Calgary, 2 Swiss Federal Institute of Technology in Zurich

Summary

Uncertainty analysis is an important aspect of quantifying the results of the inversion problem. In this report, we compare the uncertainty analysis given by two methods for full waveform inversion. The first method is by using the approximation of the inverse Hessian to perform the uncertainty analysis, as the approximation of the inverse Hessian is closely related to the posterior model covariance matrix. The second method is based on a machine learning-based method, which uses the Bayesian neural network (BNN) to generate elastic models and then performs the inversion. In the BNN, each trainable weight is represented as a Gaussian distributed probability distribution function (pdf). When BNN is well trained, we can forward calculate the BNN several times and perform the statistic analysis for the prediction results and give the uncertainty analysis for the generated models. Our numerical results suggested that both methods can generate promising inversion results and reasonable uncertainty quantification when compared with the true model errors.

Theory

Full-waveform inversion (FWI) addresses the geophysical inverse problem of estimating subsurface model parameters from observed waveform data. In most geophysical applications, FWI is introduced as an iterative, local optimization problem that attempts to minimize the least-squares residuals between observed and synthetic data. Mathematically, the inverse problem is ill-posed, leading to a non-uniqueness of the solutions. It remains challenging to solve inverse problems practically due to limitations in data acquisition, measurement uncertainties and the non-uniqueness of the solution (Tarantola, 1984; Lailly, 1983).

Estimations of the resolution or uncertainty in seismic inversions have a long history in geophysics and can be analyzed with mathematical tools such as the posterior covariance matrix. The posterior covariance matrix is closely related to the inverse Hessian (Fichtner and Trampert, 2011; Zhu et al., 2016). However, for practical problems with millions of parameters, it is unfeasible to store such vast matrices. With least-squares QR factorization Zhang and McMechan (1995) modify classic inversion algorithms with least-squares QR factorization to handle large-scale inverse problems. The spatial resolution lengths with a Gaussian approximation to the resolution matrix were introduced by An (2012). Trampert et al. (2013) sample the tomographic models for resolution lengths with random probing and analyze the direction-dependent resolution lengths of waveform tomography by Fichtner and Leeuwen (2015) autocorrelating the randomly sampled Hessian. Rawlinson et al. (2014) also gives a detailed explanation for the uncertainty estimation for the seismic inversion problem.

Randomized singular-value decomposition (SVD) also attracted attention to geophysicists, Halko et al. (2011), with the development of matrix probing theories in applied mathematics. The Bayesian inference workflow for waveform tomography is formulated by Bui-Thanh et al. (2013)

by deriving an approximation to the posterior covariance matrix by decomposing the data-misfit Hessian into eigenvalues and eigenvectors with randomized SVD. Zhu et al. (2016) improve the efficiency of the Hessian computation by exploiting point-spread function (PSF) tests. More recent ensemble-based approaches have been introduced to tomography problems by Jordan (2015) with the utilization of the Kalman Filter (KF) theory (Kalman, 1960; Evensen, 1994). Liu and Peter (2019) used the Square-root variable metric-based elastic FWI to quantify the uncertainties for V_p and V_s . The theory part of this report is referred to the paper of Liu and Peter (2019) to provide a recent review of uncertainty estimations. In this abstract, we will first use the inverse Hessian approximation to give the uncertainty quantification for the two parameters of elastic FWI. As the inverse Hessian approximation in this paper is calculated by the Quasi-Newton method, which requires starting point given manually, we discuss how this initial guessing of the inverse Hessian could influence the final uncertainty quantification. Next, we briefly introduce the BNN-based EIFWI, which uses the BNN to generate elastic models. Uncertainty analysis can be given by calculating the well-trained neural network several times to perform the statistical analysis of the generated elastic models. Then, we will compare the uncertainty quantification given by these two methods.

Results

In this section, we will use part of the full Marmousi model as the V_p and V_s models to perform the inversion. The size of the model is 100×200 . The grid length we use here is $dx = dz = 20$. The inversions are all carried out in the time domain. We assume that all the sources are well known, and we use Ricker's wavelet as the source, with the main frequency 10Hz. All the sources and receivers are located on the model's surface, with a shot interval of 600m and a receiver interval of 20m. We use the Wolfe condition to calculate the step (Nocedal and Wright, 1999).

Figure 1 (a) and (b) are the true V_p , V_s models respectively. Figure 1 (c) and (d) are inversion results for V_p and V_s after 500 iterations after using the BFGS method. We can see that the inversion is successful. Most of the structures of the elastic model have been correctly updated, especially for the anomaly located in the center of the models. The deeper part of the model is less updated, and due to that, we have limited acquisition illumination. If the acquisition system is cross-well, which means that we have shots on one side of the model and receivers on the other side of the model, we could have better illumination for the deeper part of the model. Figures 1 (e) and (f) illustrate the uncertainty analysis of the FWI for the V_p , and V_s . Figures 1 (e) and (f) are the standard deviation of the posterior sampling for parameter V_p , and V_s using equation 21. Figure 1 (g) and (h) are the absolute model errors for the V_p , V_s respectively, which is the difference between the true models and MAP models. Our assumption for a successful uncertainty quantification is that the uncertainty should match well with the absolute model error. From the comparison between the third and the fourth rows, we can see that the errors of the center anomaly align well with the standard deviation uncertainty indicating the correct quantification of the uncertainty for the shallower part of the elastic model. However, the model errors for the deeper parts of the model are poorly reflected in the standard deviation plots. This may be because the inverse approximation Hessian is constructed with the gradients and the model updates of the FWI. If the model updates and the gradients have small updates in the deeper part of the model, then the uncertainty quantification according to such a Hessian also has few information uncertainty updates for the deeper part of the model. A reasonable guess for the uncertainty for the deeper part of the model should be larger than the shallower part.

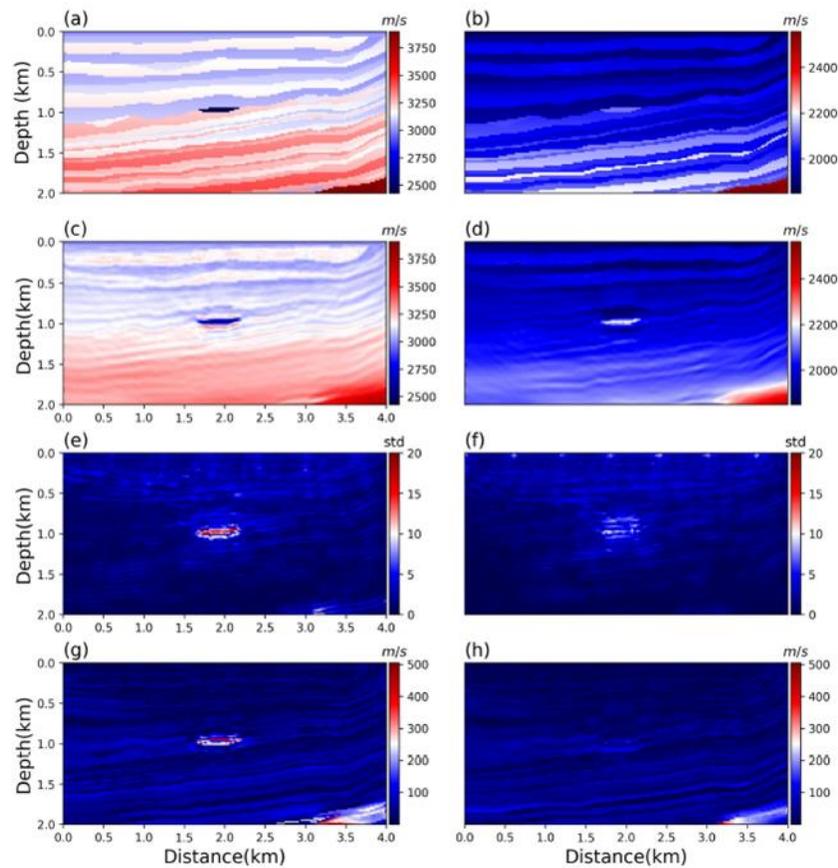


FIG. 1. FWI uncertainty analysis. (a) true Vp. (b) true Vs. (c) Vp inversion result. (d) Vs inversion result (e) Vp standard deviation. (f) Vs standard deviation. (e) Absolute Vp model error. (f) Absolute Vs model error.

We perform the forward calculation using the well-trained BNN 1000 times and obtain a set of the 1000 models for all the Vp, Vs, and ρ models. As the weights are drawn from the well-trained posterior probability distribution function, each forward calculation gives relatively different velocity models. Then we calculate the mean and the standard deviation of the velocity models to give the uncertainty analysis of these velocities generated with BNN. If the velocity model agrees with each other at a certain location of the model, the uncertainty for this point is low (and vice-versa). We will compare the standard deviation of the prediction sets for the elastic models with the absolute model error. If the standard deviations match well with the model error, then we consider that this is a valid uncertainty quantification. The third row of Figure 2 illustrates the standard deviation, and the last row demonstrates the absolute model uncertainty. We can see that the standard deviations estimated by the BNN match well with the absolute model error. We can see that most of the prediction errors are positioned in the center anomaly of the model and the deeper layers, and these model errors are all correctly reflected on the standard deviation plots. For instance, the largest model error for Vp in Figure 2 (g) is located below the anomaly, and we can also observe that the standard deviation below the anomaly is large. We can also clearly see more noise is presented on these model errors, especially for the deeper part of the model, indicating that we should have high uncertainty in these areas, indicating the successful

uncertainty quantification of the FWI. Both the uncertainties given by the inverse approximation Hessian and the BNN have the ability to characterize the inverse results for the model uncertainty of FWI. BNN method has a good ability to characterize overall uncertainty of the elastic model, but noise can be observed in the MAP model and the uncertainty quantification model. The inverse Hessian approximation method is not noise-free but has some limitations on estimating the uncertainty of the model where it has weakly illuminated, and a modification of the initial guessing could help to release this issue.

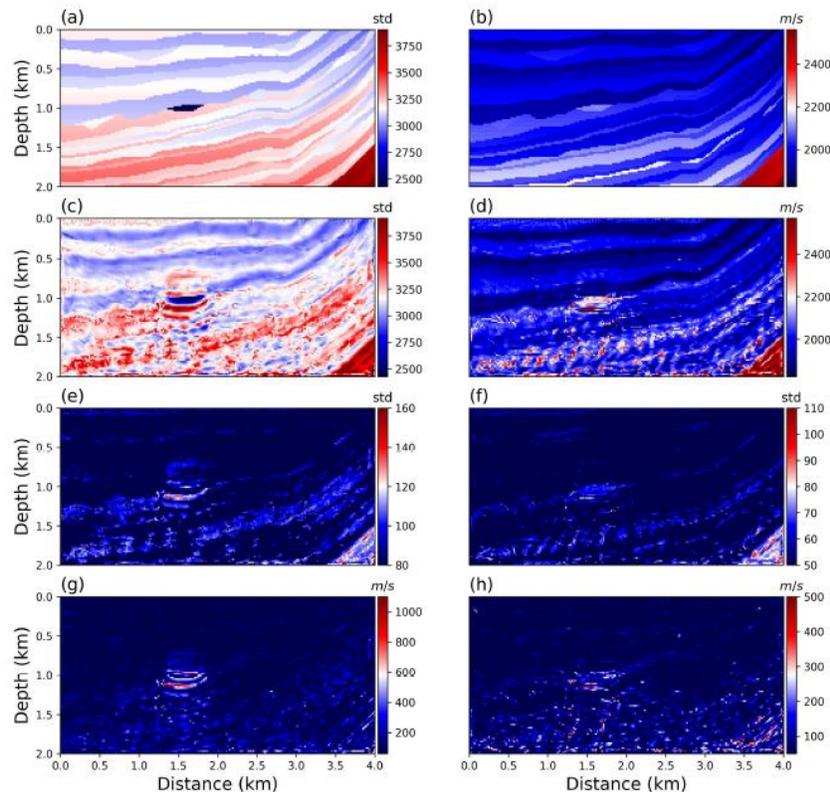


FIG3. FWI uncertainty analysis with given by BNN. (a) true V_p . (b) true V_s . (c) V_p BNN FWI result. (d) V_s BNN result (e) V_p BNN FWI standard deviation. (f) V_s BNN FWI standard deviation. (g) Absolute V_p model error. (h) Absolute V_s model error.

We perform the forward calculation using the well-trained BNN 1000 times and obtain a set of the 1000 models for all the V_p , V_s , and ρ models. As the weights are drawn from the well-trained posterior probability distribution function, each forward calculation gives relatively different velocity models. Then we calculate the mean and the standard deviation of the velocity models to give the uncertainty analysis of these velocities generated with BNN. If the velocity model agrees with each other at a certain location of the model, the uncertainty for this point is low (and vice-versa). We will compare the standard deviation of the prediction sets for the elastic models with the absolute model error. If the standard deviations match well with the model error, then we consider that this is a valid uncertainty quantification. The third row of Figure 3 illustrates the standard deviation, and the last row demonstrates the absolute model uncertainty. We can see that the standard deviations estimated by the BNN match well with the absolute model error. We

can see that most of the prediction errors are positioned in the center anomaly of the model and the deeper layers, and these model errors are all correctly reflected on the standard deviation plots. For instance, the largest model error for V_p in Figure 3 (g) is located below the anomaly, and we can also observe that the standard deviation below the anomaly is large. We can also clearly see more noise is presented on these model errors, especially for the deeper part of the model, indicating that we should have high uncertainty in these areas, indicating the successful uncertainty quantification of the FWI.

Conclusions

Both the uncertainties given by the inverse approximation Hessian and the BNN have the ability to characterize the inverse results for the model uncertainty of FWI. BNN method have good ability to characterize overall uncertainty of the elastic model, but noise can be observed in the MAP model and the uncertainty quantification model. Inverse Hessian approximation method is noise free, but have some limitation on estimating the uncertainty of the model where it have weekly illuminated, and a modification of the initial guessing could help to release this issue. We compare the uncertainty analysis given by the inverse Hessian approximation method in conventional FWI with the uncertainty analysis given by the BNN. Both of the method have successfully captured the main characterization of model error, though with a difference in uncertainty patterns. In the uncertainty analysis, using the inverse Hessian approximation method, we improve the uncertainty analysis by changing the initialization of the inverse Hessian. The deeper part of the model has better uncertainty analysis by utilization of such a method.

Acknowledgements

We thank the sponsors of CREWES for their continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19. The first author is also supported by the Chine Scholarship Council (CSC).

References

An, M., 2012, A simple method for determining the spatial resolution of a general inverse problem: *Geophysical Journal International*, 191, No. 2, 849–864.

Bui-Thanh, T., Ghattas, O., Martin, J., and Stadler, G., 2013, A computational framework for infinite dimensional bayesian inverse problems part i: The linearized case, with application to global seismic inversion: *SIAM Journal on Scientific Computing*, 35, No. 6, A2494–A2523.

Chen, K., and Sacchi, M. D., 2020, Time-domain elastic gauss–newton full-waveform inversion: a matrix-free approach: *Geophysical Journal International*, 223, No. 2, 1007–1039

Evensen, G., 1994, Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics: *Journal of Geophysical Research: Oceans*, 99, No. C5, 10,143–10,162.

Fichtner, A., and Leeuwen, T. v., 2015, Resolution analysis by random probing: *Journal of Geophysical Research: Solid Earth*, 120, No. 8, 5549–5573.

Fichtner, A., and Trampert, J., 2011, Hessian kernels of seismic data functionals based upon adjoint techniques: *Geophysical Journal International*, 185, No. 2, 775–798.

Fletcher, R., 2013, *Practical methods of optimization*: John Wiley & Sons.

Fletcher, R., and Powell, M. J., 1963, A rapidly convergent descent method for minimization: *The computer journal*, 6, No. 2, 163–168.