

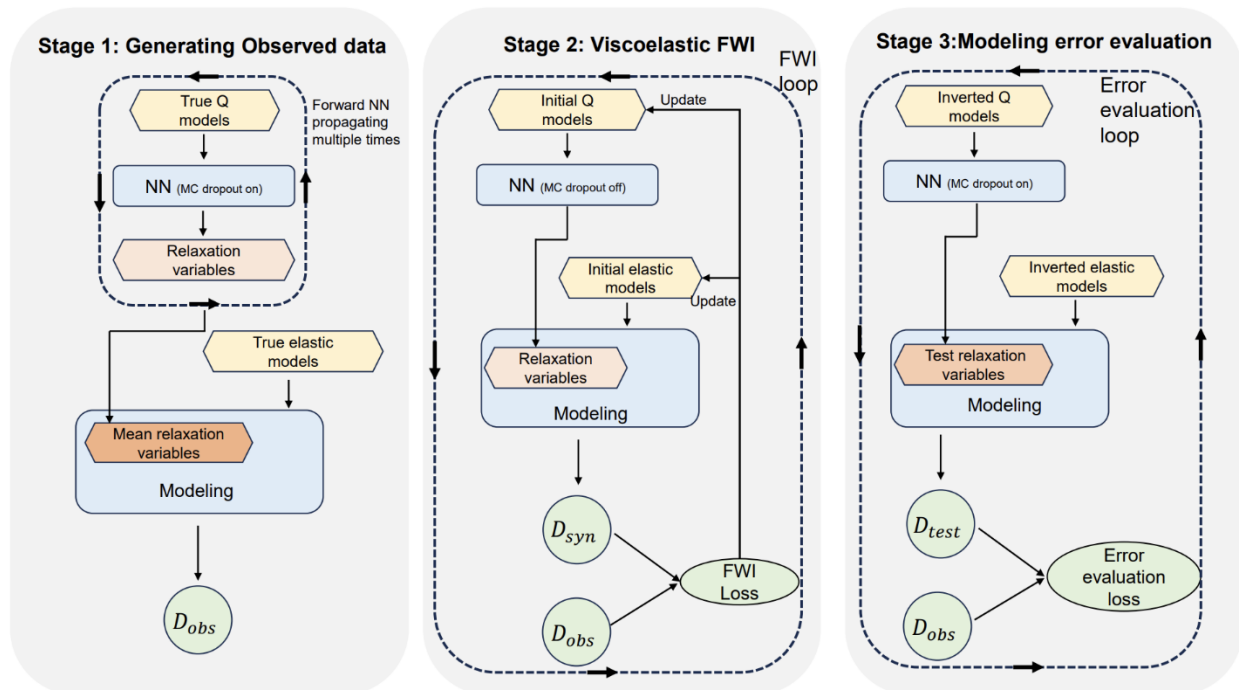
## A specific type of modeling error evaluation for viscoelastic full waveform inversion

Tianze Zhang, Kristopher Innanen, Daniel Trad  
CREWES project, University of Calgary

### Summary

We propose a novel method where a Multi-Layer Perceptron (MLP) is pre-trained to learn the mapping from  $Q$  values to relaxation variables. This trained MLP is then integrated into a Recurrent Neural Network (RNN)-based GSLs viscoelastic FWI framework, creating a complete computational graph. This graph connects  $Q$  models, relaxation variables, and elastic models directly to the synthetic data, thereby enabling the direct inversion of the  $Q$  model. Additionally, we employ the Monte Carlo dropout technique within the neural network to quantify the uncertainty associated with the MLP's learning process for mapping  $Q$  values to relaxation variables. In our assumption of a constant  $Q$  model, this uncertainty quantification reflects the relaxation variables' limited capacity to accurately represent a constant  $Q$  model. The potential of extending this approach to variable  $Q$  models is straightforward. The impact of this limited capacity, which is effectively a modeling error, on the forward modeling data is also examined.

### Workflow



**Fig1.** This study follows an overall numerical testing diagram comprising three stages.

We perform viscoelastic FWI and attempt to quantify the modeling error arising from the limited capability of the relaxation variables to represent a constant  $Q$  model. Our numerical test comprises three stages, and the overall inversion diagram is depicted in Figure 1. In the first stage, I generate the observed data assuming that  $Q$  is frequency-independent within the

bandwidth of interest. During the observed data generation, I forward propagate the well-trained neural network to produce a set of relaxation variables using the true  $Q$  model by enabling the Monte Carlo (MC) dropout in the neural network. As shown in propagation of the well-trained neural network allows the mean relaxation variable to provide a more accurate evaluation of the constant  $Q$  model. I use the true elastic model and the mean relaxation variables to compute the observed data.

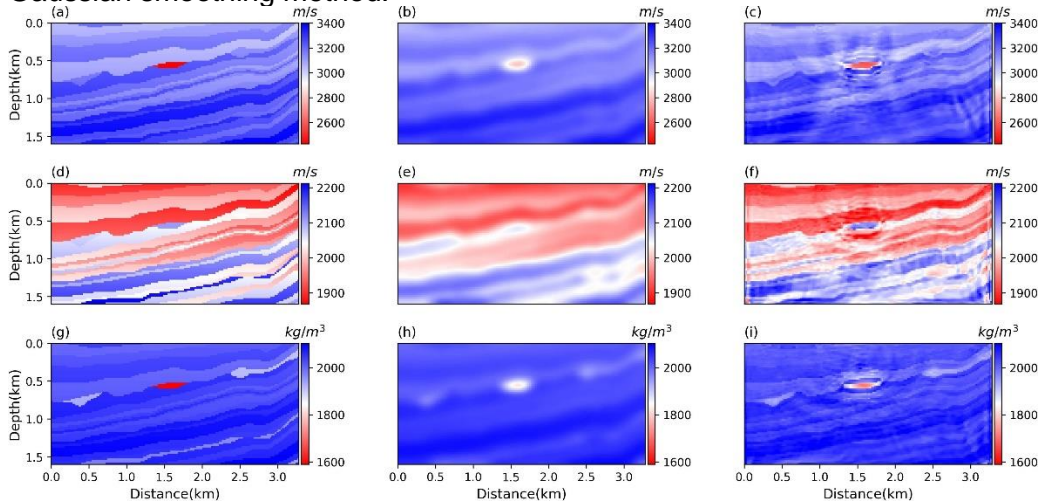
In stage 2, We calculate the viscoelastic FWI. In this study, we use the recurrent neural network work-based inversion strategy, which uses the automatic differential method to update the elastic models and the  $Q$  models directly. During the second stage, we do not enable the Monte Carlo (MC) dropout in the neural network and use a single realization of the relaxation variables from the neural network for inversion. Consequently, the relaxation variables are generated differently from stage 1. If we were to enable the MC dropout, it would increase the nonlinearity between the data and the  $Q$  model. Since Full Waveform Inversion (FWI) is already a highly nonlinear inversion method, introducing additional nonlinearity would render the inversion unstable. In contrast to stage 1, where we forward calculated the relaxation variables multiple times, we compute only one realization of the relaxation variables in stage 2 to reduce the computational cost. As a result, the forward modeling used in inversion and data generation differs, ensuring that we do not commit the inverse crime in our numerical test.

In the third stage, we aim to assess the modeling error arising from the limited ability of a single realization of the relaxation field to represent the constant  $Q$  model. We achieve this by "turning on" the Monte Carlo (MC) dropout switch of the neural network and letting the neural network generate different relaxation variables and evaluate the corresponding synthetic data. In more detail, We first use the inverted  $Q$  model as input to realize the relaxation variables, referred to as the test relaxation variables. Using the inverted elastic models and the test relaxation variables, we obtain the synthetic data  $D_{test}$  and compute the misfit between  $D_{test}$  and  $D_{obs}$ , which we refer to as the error evaluation loss. If the error evaluation loss is equivalent to or within the acceptable range of the final FWI loss, then  $D_{test}$  is considered synthetic data that differs from the final FWI synthetic data  $D_{test}$  but does not affect the final FWI loss due to the variation of the relaxation variables. Such  $D_{test}$  will be collected into a dataset for quantifying the modeling error. Subsequently, we allow the neural network to generate another realization of the test relaxation variable and evaluate a new  $D_{test}$ . By repeating this process several times, we obtain a dataset of  $D_{test}$  values that differ solely due to the variation of the relaxation variables but have misfit values close to the final FWI loss. we consider the standard deviation of the collected  $D_{test}$  dataset as the modeling errors. These errors arise from the inability of a single realization of the relaxation variable to represent the constant  $Q$  model.

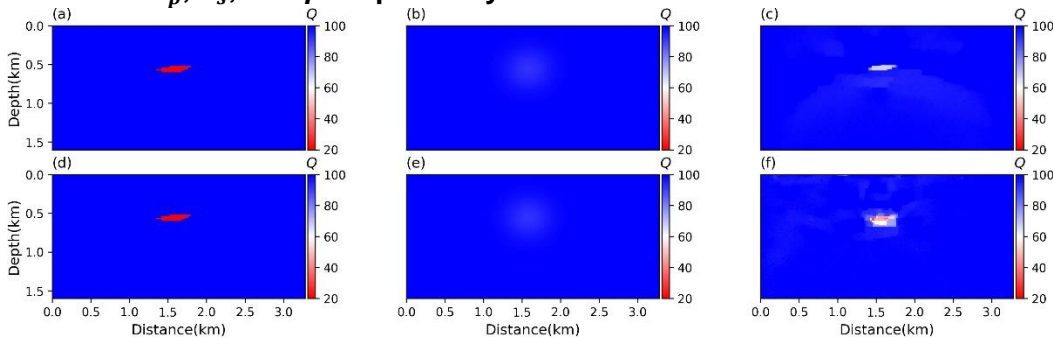
## Results, and observationsd

In this section, we calculate viscoelastic FWI on a portion of the Marmousi model, where a reservoir is situated in the upper center of the model. The true models for  $V_p$ ,  $V_s$ , and  $\rho$  are shown in Figure 2 (a), (d), and (g), respectively. The true attenuation models for  $Q_p$ , and  $Q_s$  are depicted in Figure 3 (a) and (d), respectively. The model dimensions are  $80 \times 160$ , with a grid spacing of  $dx = dz = 20$  m. I use  $L = 2$  layers of the SLS mechanism and forward propagate the relaxation variables generated with the post-trained neural network 1000 times to obtain the mean relaxation variables for generating the observed data. The source wavelets are Ricker wavelets

with a central frequency of 13 Hz, and the reference frequency used here is 40 Hz. The maximum recording time is 1.5 s with a time interval of 0.002 s. The initial models for  $V_p$ ,  $V_s$ , and  $\rho$  are plotted in Figures 2 (b), (e), and (h), respectively. The initial models for the attenuation models are plotted in figures 3 (b), and (e), respectively. The initial models are obtained by smoothing the true models with the Gaussian smoothing method.

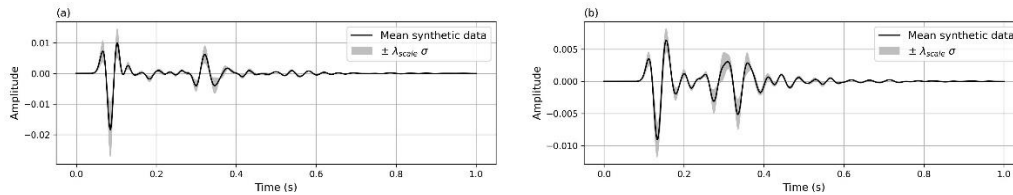


**Fig2.** Part of the Marmousi model  $V_p$ ,  $V_s$ , and  $\rho$  viscoelastic FWI. (a), (d) and (g) are the true models for  $V_p$ ,  $V_s$ , and  $\rho$ . (b), (e) and (h) are the initial models. (c), (f) and (i) are the inversion results  $V_p$ ,  $V_s$ , and  $\rho$  respectively.

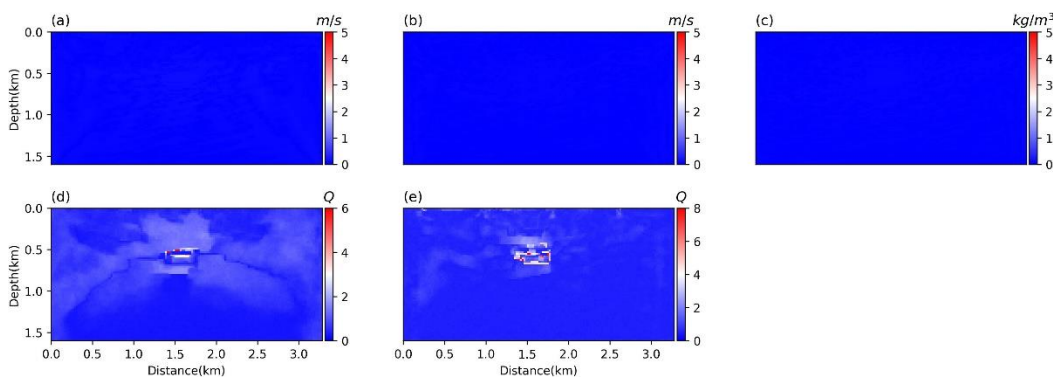


**Fig3.** Part of the Marmousi model  $Q_p$ ,  $Q_s$  viscoelastic FWI. (a) and (d) are the true models for  $Q_p$ , and  $Q_s$ , respectively. (b), (e) are the initial models for  $Q_p$ , and  $Q_s$ , respectively. (c) and (f) are the inversion results for  $Q_p$ , and  $Q_s$ , respectively.

I use the total variation regularization in the objective function to mitigate artifacts caused by cross-talk between parameters. I employ the Adam algorithm for optimization and set the maximum iteration number to 300. The relaxation variables during the inversion are obtained with a single forward realization of the neural network. The inversion results for the elastic models  $V_p$ ,  $V_s$ , and  $\rho$  are shown in Figures 2 (c), (f), and (i), respectively. The inversion results for the attenuation models  $Q_p$ , and  $Q_s$  are displayed in Figures 3 (c) and (f), respectively. Compared to the true models, I obtain promising inversion results for both the elastic and attenuation models. The inversion results for  $Q_s$  also exhibit some difficulty in accurately describing the shape of the central reservoir.



**Fig4. Modeling error evaluation of the Marmousi model for 2 synthetic seismic records. Synthetic records at the last iteration of FWI and its variation caused by the modeling error evaluation.**



**Fig5. Standard deviation models of 87 updated models for  $V_p$  (a),  $V_s$  (b), and  $\rho$  (c),  $Q_p$  (d), and  $Q_s$  (e). These Standard deviation models represent how the inability of relaxation variables to quantify a constant  $Q$  model will influence the FWI results.**

Next, we will quantify the modeling error, corresponding to the demonstration in stage 3 of Figure 1. The MC dropout rate we set here is 0.43, and we will calculate the loop of the modeling error evaluation in stage 3 for 100 iterations. We set the upper and lower loss bound as the  $\pm 2\%$  in the neighborhood of the final FWI loss. we evaluate 100 relaxation variables, and 87 of them have data misfit values which are within the acceptable range. Thus, 87  $D_{test}$  will be collected to evaluate the modeling error. The standard deviation of the model collected are regarded as the modeling error evaluated. I plot two traces of the records in Figure 4. In Figure 4, the black lines are the mean synthetic records for the last iteration of FWI, and the gray area represents the change of the range of the signal due to the change of the relaxation variables, quantified by the scaled standard deviation.

We obtain 87 updated models for  $V_p$ ,  $V_s$ ,  $\rho$ ,  $Q_p$ , and  $Q_s$ . We plot the standard deviation of these updated models for  $V_p$ ,  $V_s$ ,  $\rho$ ,  $Q_p$ , and  $Q_s$  in Figure 5 (a), (b), (c), (d), and (e), respectively. These figures represent the impact of the relaxation variables' limited ability to quantify a constant  $Q$  model on the FWI results for the elastic Marmousi model section. We do not observe significant variation in the updates of the elastic model at the observable scale. In contrast, the variations in the updates for the attenuation models are readily apparent. This again suggests that such an inability primarily affects the evaluation of the attenuation models.

## Conclusion

In this report, we train a neural network to map the attenuation factor  $Q$  to the relaxation times for different numbers of the SLS mechanism. We observe that the mean  $Q$  spectrum calculated using the relaxation variables generated by the well-trained neural network, can more accurately represent the constant  $Q$  model within the seismic frequency bandwidth. The well-trained network can map any  $Q$  value to its corresponding relaxation variable, enabling us to directly update the  $Q$  model when incorporating the network into the RNN-based inversion framework.

The observed data are generated using the mean relaxation variables obtained from multiple forward propagations of the well-trained neural network, providing a better approximation of the constant  $Q$  model. During the inversion, we calculate only one realization of the relaxation variables to maintain the same computational cost. This means that the forward modeling used to obtain the observed data and during the inversion calculation are different, releasing the inverse crime problem that we are facing in the synthetic inversion tests.

After the inversion, by enabling the Monte Carlo dropout method in the neural network, we can evaluate the modeling error caused by the relaxation variables' insufficient ability to model the constant  $Q$  model. The overall scale of the modeling error for the seismic records is small, but it can influence the amplitude and phase of the records. This modeling error has a greater impact on the evaluation of the attenuation models than on the elastic models.

## Acknowledgements

The sponsors of CREWES are gratefully thanked for continued support. This work was funded by CREWES industrial sponsors, NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19.

## References

- [1] Day, S. M., and Minster, J. B., 1984, Numerical simulation of attenuated wavefields using a padé approximant method: *Geophysical Journal International*, 78, No. 1, 105–118.
- [2] Emmerich, H., and Korn, M., 1987, Incorporation of attenuation into time-domain computations of seismic wave fields: *Geophysics*, 52, No. 9, 1252–1264.
- [3] Fabien-Ouellet, G., Gloaguen, E., and Giroux, B., 2017, Time domain viscoelastic full waveform inversion: *Geophysical Journal International*, 209, No. 3, 1718–1734.
- [3] Fan, N., Zhao, L.-F., Xie, X.-B., Ge, Z., and Yao, Z.-X., 2016, Two-dimensional time-domain finite-difference modeling for viscoelastic seismic wave propagation: *Geophysical Journal International*, 206, No. 3, 1539–1551.
- [4] Fichtner, A., and van Driel, M., 2014, Models and fréchet kernels for frequency-(in) dependent  $q$ : *Geophysical Journal International*, 198, No. 3, 1878–1889.