geoconvention
Virtual Event
September 21-23
2020
CSPG • CSEG • CWLS • GAC • MAC • IAH

Theory based machine learning viscoelastic full waveform inversion based on recurrent neural network

Tianze Zhang, Kristopher Innanen, Jian Sun, Daniel Trad
CREWES, Department of Geoscience, University of Calgary

Summary

In this study, we use a recurrent neural network (RNN) to achieve viscoelastic full waveform inversion. The RNN is a typical type of neural network that consists of several RNN cells. In this study, each RNN cell is designed according to the stress velocity viscoelastic wave equation. With the Automatic Differential engine built in the machine learning library, the exact gradient for the trainable parameters, the velocity models and attenuation models, would be given based on the computational graph. Both the simple and complex model numerical inversion tests prove that the inversion based on this theory-guided recurrent neural network can give accurate inversion results. We also test the inversion with different objective function. L1, L2 and Huber objective function are tested. We conclude that all the objective function inversion can give promising results.

Introduction

Full waveform inversion (FWI) could be considered as a powerful method based on data fitting to invert velocity models. During full waveform inversion, we would first generate synthetic wavefields by using the initial models. The forward propagating wavefields would be calculated, and shotrecords would also be recorded at the same time to form the synthetic shotrecords. The zero-lag correlation between the backpropagation wavefields and the forward propagation wavefields would be calculated as the gradients to update the models. Sometimes, the Hessian matrix would also be calculated to tackle the crosstalk problem. In recent years, adding attenuation into FWI had been studied by more researchers by using the viscoelastic wave equation based on the velocity-stress formulation. Fabien-Ouellet et al. (2017) explored the use of OpenGL to develop a portable code that can take advantage of the many parallel processor architectures now available and presented a program for 2D and 3D viscoelastic FWI in the time domain. Yang et al. (2016) studied 3-D multiparameter full waveform inversion (FWI) in viscoelastic media based on the generalized Maxwell/Zener body including the arbitrary number of attenuation mechanisms. Groos et al. Groos et al. (2012) explored which degree viscoelastic modeling is relevant during a full waveform inversion of shallow seismic surface waves, and concluded that if we use Q factors that are too far away from the Q factor of the observed data the inversion result becomes worse than ignoring attenuation. Trinh et al. (2018) used the SEAM Phase II Foothill dataset to simultaneously invert the P and S wave speed.

Jian et al. (2019) used the recurrent neural network to achieve the scalar wave full waveform inversion. Each of the RNN cells is designed according to the scalar wave equation, which forms the theory-based machine learning seismic data inversion method. In this paper, based on their idea, we introduce the recurrent neural network (RNN) to achieve viscoelastic FWI. In this study, the cells in RNN are designed according to the viscoelastic wave equation. Exact gradients for the trainable parameters would be given by the Automatic Differential engine built in the machine

learning framework. By using an optimization method and the step length for each model, we can update the models and reduce the misfit between the observe data and synthetic data.

Theory and method

1. Forward modeling

The 2-D first order viscoelastic wave equation can be derived from the momentum conservation equation and the viscoelastic constitutive relationship. Equation (1) shows the viscoelastic wave equation written in the stress-velocity form. σ_{xx} σ_{zz} and σ_{xy} are the stress tensors. v_x and v_y are the velocity fields in horizontal and vertical directions respectively. r_{xx}, r_{yy} and r_{xy} are the memory variables. τ_σ is the relaxation time which we use for both the P-waves and S-waves (Robertsson et al. (1994)). τ_ε^P and τ_ε^S define the attenuation level of the media. Π is the relaxation modulus corresponding to P-waves analogous to $\lambda + 2\mu$ in the elastic case, where λ and μ are the Lame constants. Following Robertsson et al (1994), we have the viscoelastic wave equations:

$$
\begin{aligned}
&\frac{\partial \boldsymbol{v_x}}{\partial t} = \frac{1}{\rho}\left(\frac{\partial \boldsymbol{\sigma_{xx}}}{\partial x} + \frac{\partial \boldsymbol{\sigma_{xy}}}{\partial y}\right)\\
&\frac{\partial \boldsymbol{v_y}}{\partial t} = \frac{1}{\rho}\left(\frac{\partial \boldsymbol{\sigma_{xy}}}{\partial x} + \frac{\partial \boldsymbol{\sigma_{yy}}}{\partial y}\right)\\
&\frac{\partial \boldsymbol{\sigma_{xx}}}{\partial t} = \pi\frac{\tau_\varepsilon^P}{\tau_\sigma}\left(\frac{\partial \boldsymbol{v_x}}{\partial x} + \frac{\partial \boldsymbol{v_y}}{\partial y}\right) - 2\mu\frac{\tau_\varepsilon^S}{\tau_\sigma}\frac{\partial \boldsymbol{v_y}}{\partial y} + r_{xx}\\
&\frac{\partial \boldsymbol{\sigma_{yy}}}{\partial t} = \pi\frac{\tau_\varepsilon^P}{\tau_\sigma}\left(\frac{\partial \boldsymbol{v_x}}{\partial x} + \frac{\partial \boldsymbol{v_y}}{\partial y}\right) - 2\mu\frac{\tau_\varepsilon^S}{\tau_\sigma}\frac{\partial \boldsymbol{v_x}}{\partial x} + r_{yy}\\
&\frac{\partial \boldsymbol{\sigma_{xy}}}{\partial t} = \mu\frac{\tau_\varepsilon^S}{\tau_\sigma}\left(\frac{\partial \boldsymbol{v_x}}{\partial y} + \frac{\partial \boldsymbol{v_y}}{\partial x}\right) + r_{xy}\\
&\frac{\partial \boldsymbol{r_{xy}}}{\partial t} = -\frac{1}{\tau_\sigma}\left(r_{xy} + \mu\left(\frac{\tau_\varepsilon^S}{\tau_\sigma} - 1\right)\left(\frac{\partial \boldsymbol{v_x}}{\partial y} + \frac{\partial \boldsymbol{v_y}}{\partial x}\right)\right)\\
&\frac{\partial \boldsymbol{r_{xx}}}{\partial t} = -\frac{1}{\tau_\sigma}\left(r_{xx} + \pi\left(\frac{\tau_\varepsilon^P}{\tau_\sigma} - 1\right)\left(\frac{\partial \boldsymbol{v_x}}{\partial x} + \frac{\partial \boldsymbol{v_y}}{\partial y}\right) - 2\mu\left(\frac{\tau_\varepsilon^S}{\tau_\sigma} - 1\right)\frac{\partial \boldsymbol{v_y}}{\partial y}\right.\\
&\left.\frac{\partial \boldsymbol{r_{yy}}}{\partial t} = -\frac{1}{\tau_\sigma}\left(r_{yy} + \pi\left(\frac{\tau_\varepsilon^P}{\tau_\sigma} - 1\right)\left(\frac{\partial \boldsymbol{v_x}}{\partial x} + \frac{\partial \boldsymbol{v_y}}{\partial y}\right) - 2\mu\left(\frac{\tau_\varepsilon^S}{\tau_\sigma} - 1\right)\frac{\partial \boldsymbol{v_x}}{\partial x}\right.\right.
\end{aligned}
\tag{1}
$$

Equations (1) show the viscoelastic wave equation we use in this study. The partial derivatives in time and space are approximated by finite-difference of order 2 on a staggered grid, in which velocities are updated at integer time steps Δt. The stress fields and memory variables are updated at half-time steps. In this paper, the particle derivative is calculated according to the image convolution. We create kernels to calculate partial derivatives in different directions according to the staggered grid method. The kernels would scan the wavefields and calculate image convolutions to get the partial derivatives.

2. Automatic differential

By using the machine learning library, like TensorFlow or Pytorch. We do not need to calculate the gradients all by ourselves; the gradients can be calculated by using the Automatic Differential engine built-in these machine learning frameworks. During the forward propagation, every

mathematical operation is recorded to form the Dynamic Computational Graph. The exact gradients are calculated according to this Dynamic Computational Graph by using the backpropagation method as we just explained.

Synthetic test

In this section, we will perform some numerical tests to examine the efficiency of the proposed methods. We also test the inversion with different objective functions. We tested three objective functions, which are the L1 norm. L2 norm and Huber objective function. From the definition for the loss functions, we can see that L2 objective function is the least-squares norm, and L1 objective function is the least-absolute values norm. The Huber objective function is the combination of the L1 norm and the L2 norm. The l2 norm is the misfit we use in the traditional FWI.

$$loss_{L_1} = |x - y| \tag{2}$$

$$loss_{L_2} = \frac{1}{2}|x - y|^2 \tag{3}$$

$$loss_{Huber} = \begin{cases} 0.5|x - y|^2, if\ |x - y| < 1 \\ |x - y| - 0.5, \ otherwise \end{cases} \tag{3}$$

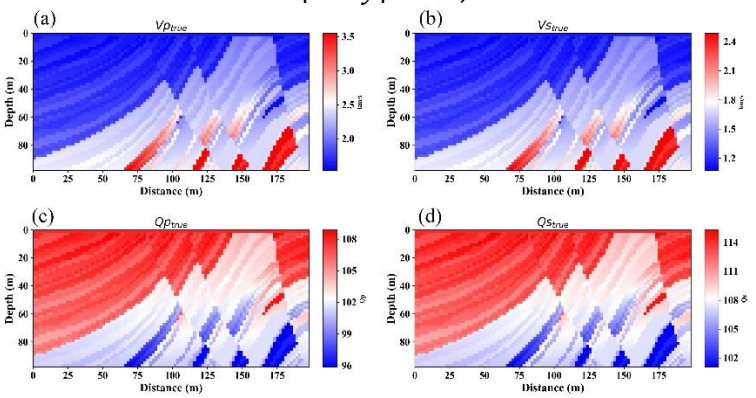

Figure 1. (a), True Vp model, (b) True Vs model, (c) True Qp model, (d) True Qs model

In this numerical test, we use part of the Marmousi model to test the efficiency of this inversion method. The source of the wavelet is the Ricker's wavelet with a main frequency of 35Hz. The size of the model is 50×100 grid points. The grid length of the model is dz = dx = 2m. Seven shots are evenly distributed on the top of the model. Every grid point at the top of the model is located a receiver. The total receiving time for the shot records is3s with a time step of 0:004s. We use 2nd order in time and 2nd order in space staggered grid method to simulate the synthetic data. Figure 1 (a), (b), (c) and (d) are the true models for Vp, Vs, Qp and Qs. Figure 1 (a), (e), (i), (m) are the initial models for Vp, Vs, Qp and Qs respectively.

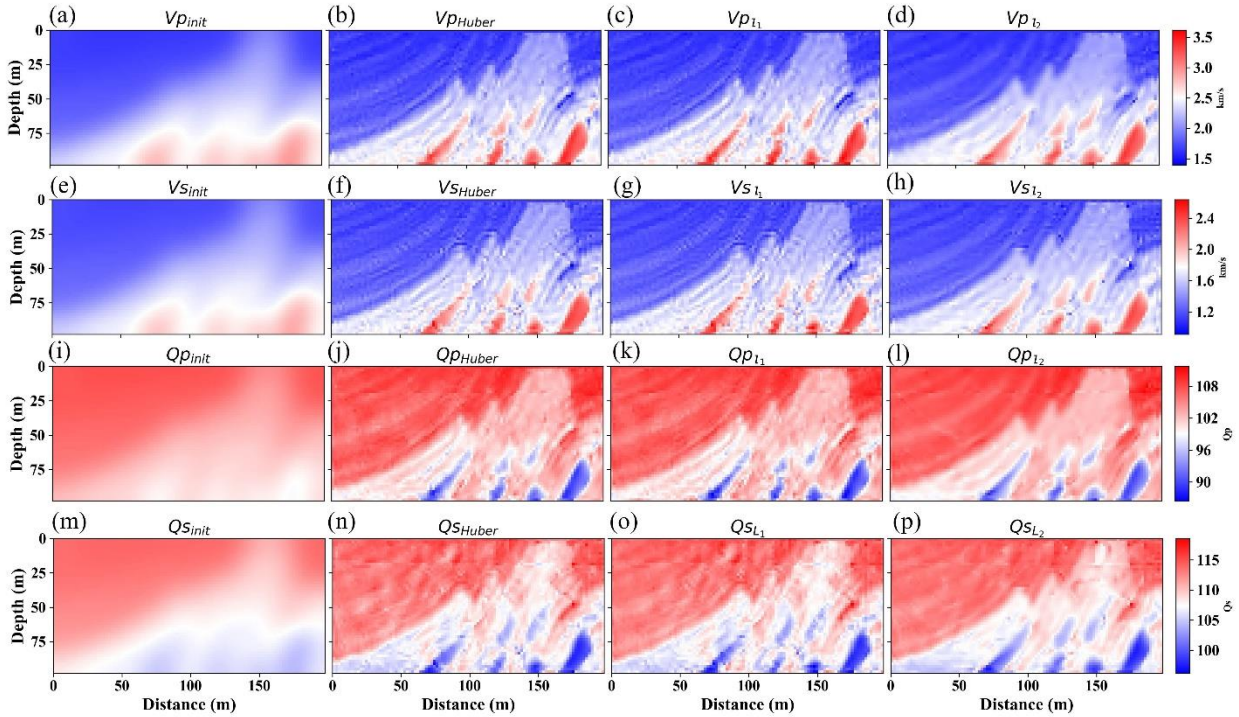

Figure 2. Inversion results for different objective function

Figure 2 shows the inversion results for different objective functions. From the inversion results we can see that all three objective function can give the right inversion results also the very high wavenumber components of the model are better reconstructed by using the Huber and l1 objective function. On the one hand, the bandwidth-limited problem may cause this inaccuracy problem for the very high component of the models. On the other hand, this may also due to the nature of using different objective functions to perform inversion, which means that the l1 and Huber function would be more robust objective functions for the full waveform inversion. From Figure 2 we can also see that the high attenuation part of the model has been correctly updated.

Conclusion

In this study, based on the viscoelastic wave equation, we build the viscoelastic RNN cell and performed the viscoelastic full waveform inversion, which forms the theory-based machine learning full waveform inversion. We also test the L1, L2 and Huber objective function to perform RNN based FWI. All, the three objective functions have given us the right inversion results, however, the L1 and Huber norm have the better ability to reconstruct the high wavenumber component of the models.

Acknowledgment

We thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors, and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. We also thank the support from the China Scholarship Council (CSC).

References

Belahi, T., Fuji, N., and Singh, S., 2015, Elastic versus viscoelastic full waveform inversion of near-offset and wide-angle data in the presence of attenuation, in 77th EAGE Conference and Exhibition 2015.
Belahi, T., Singh, S., and Fuji, N., 2016, Viscoelastic full waveform inversion of sea bottom long offset seismic data in presence of attenuation, in 78th EAGE Conference and Exhibition 2016.
Bohlen, T., 2002, Parallel 3-d viscoelastic finite difference seismic modelling: Computers & Geosciences, 28, No. 8, 887–899.
Brossier, R., Operto, S., and Virieux, J., 2010, Which data residual norm for robust elastic frequency-domain full waveform inversion?: Geophysics, 75, No. 3, R37–R46.
Day, S., and Minster, J., 1984, Numerical simulation of wavefields using a pade approximation method: Geophys. JR astr. Soc, 78–105.
Djikpéssé, H. A., and Tarantola, A., 1999, Multiparameter l1 norm waveform fitting: Interpretation of gulf of mexico reflection seismograms: Geophysics, 64, No. 4, 1023–1035.
Emmerich, H., and Korn, M., 1987, Incorporation of attenuation into time-domain computations of seismic wave fields: Geophysics, 52, No. 9, 1252–1264.
Fabien-Ouellet, G., Gloaguen, E., and Giroux, B., 2016, The adjoint state method for the viscoelastic wave equation in the velocity-stress formulation, in 78th EAGE Conference and Exhibition 2016.
Fabien-Ouellet, G., Gloaguen, E., and Giroux, B., 2017, Time-domain seismic modeling in viscoelastic media for full waveform inversion on heterogeneous computing platforms with opencl: Computers & Geosciences, 100, 142–155.
Groos, L., Schäfer, M., Forbriger, T., and Bohlen, T., 2012, On the significance of viscoelasticity in a 2d full waveform inversion of shallow seismic surface waves, in 74th EAGE Conference and Exhibition incorporating EUROPEC 2012.
Lin, Y., andWu, Y., 2018, Inversionnet: A real-time and accurate full waveform inversion with convolutional neural network: The Journal of the Acoustical Society of America, 144, No. 3, 1683–1683.
Liu, H.-P., Anderson, D. L., and Kanamori, H., 1976, Velocity dispersion due to anelasticity; implications for seismology and mantle composition: Geophysical Journal International, 47, No. 1, 41–58.
Mosser, L., Dubrule, O., and Blunt, M., 2018, Stochastic seismic waveform inversion using generative adversarial networks as a geological prior, in First EAGE/PESGB Workshop Machine Learning.
Pyun, S., Son, W., and Shin, C., 2009, Frequency-domain waveform inversion using an l 1-norm objective function: Exploration Geophysics, 40, No. 2, 227–232.
Robertsson, J. O., Blanch, J. O., and Symes, W. W., 1994, Viscoelastic finite-difference modeling: Geophysics, 59, No. 9, 1444–1456.
Sun, H., and Demanet, L., 2018, Low frequency extrapolation with deep learning, in SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2011–2015.
Trinh, P.-T., Brossier, R., Métivier, L., and Virieux, J., 2018, Data-oriented strategy and v p/v s modelconstraint for simultaneous v p and v s reconstruction in 3d visco-elastic fwi: Application to the seam ii foothill dataset, in SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 1213–1217.
Yang, F., and Ma, J., 2019, Deep-learning inversion: a next generation seismic velocity-model building method: Geophysics, 84, No. 4, 1–133.
Yang, P., Brossier, R., Métivier, L., and Virieux, J., 2016, A review on the systematic formulation of 3-d multiparameter full waveform inversion in viscoelastic medium: Geophysical Journal International, 207, No. 1, 129–149