Integrated networks for viscoelastic FWI: mapping from Q to relaxation variables and quantifying modelling error

Tianze Zhang, Kristopher A. Innanen, and Daniel Trad, University of Calgary*

SUMMARY

In viscoelastic full waveform inversion (FWI), which is typically based on the generalized standard linear solid (GSLS) model, the quality factor (Q) is not inverted for directly, but through an intermediate transformation into relaxation variables. This adds to the inverse problem for elastic and attenuation models a complex conversion back to Q. We formulate a method integrating a pre-trained multi-layer perceptron (MLP) to do the Q-relaxation variable mapping into a Recurrent Neural Network (RNN)-based FWI framework, which directly inverts for Q and elastic parameters using multiple SLS layers. The formulation allows straightforward assessment and mitigation of mapping uncertainty through Monte Carlo dropout. For instance, the approach quantifies the uncertainty associated with the approximate representation of a constant Q model with relaxation time variables. The influence of such uncertainty on the forward modelling synthetic data and the inversion results are both analyzed. Our approach is in principle extendable to frequency-variant Q values. We illustrate several inversion examples to demonstrate the significance of this type of viscoelastic modelling error analysis for inversion.

INTRODUCTION

Mechanical wave propagation through real earth media involves attenuation, a process influenced by geometric spreading, anelastic attenuation, losses during transmission, mode conversion, scattering, and refraction (Samec and Blangy, 1992; Krebes, 2019). In contrast to perfectly elastic media, where wave energy is conserved, the inherently anelastic nature of earth materials results in the progressive alteration of the amplitude and phase of the seismic wave. The quality factor Q, a dimensionless measure, quantifies attenuation as the ratio of mean energy retained in the medium against the energy dissipated in a single cycle (Fan et al., 2016). Attenuation phenomena critically influence both the amplitude and phase of seismic recordings, affecting key data modeling practices, including reverse time migration (RTM) and full-waveform inversion (FWI) (Fathalian et al., 2020).

The generalized standard linear solid (GSLS) viscoelastic modeling method utilizes relaxation time variables to define the viscoelastic constitutive relationship. Day and Minster (1984) investigated the integration of elasticity into 2-D time-domain modeling via a Padé approximation, leading to the proposal of the generalized standard linear solid (GSLS) method by Emmerich and Korn (1987). This technique, grounded in rheological principles, leverages the finite difference method for improved computational efficiency. Robertsson et al. (1994) developed a staggered grid finite difference approach, tailored to model viscoelastic wavefield using a single layer of standard linear solid (SLS). In this context, the quality factor Q is translated into stress and strain relaxation times essential for forward modeling. Each SLS layer is characterized by distinct stress and strain relaxation times, making the accurate determination of these times from the respective Q value a critical aspect of viscoelastic wavefield simulation.

Viscoelastic full waveform inversion (FWI) is recognized for its potential to refine seismic imaging and facilitate applications such as reservoir characterization and CO_2 monitoring (e.g., Fabien-Ouellet et al., 2017; Keating and Innanen, 2019, 2020; Gao et al., 2023; Mirzanejad et al., 2022; Moradi and Innanen, 2015; Pan et al., 2023), striving to achieve high-resolution inversion results for both elastic and attenuation models. Despite its promise, the implementation of viscoelastic FWI faces several technical hurdles. For instance, practical inversion for attenuation models relies on a simplistic single-layer SLS model, which has the potential of inadequately representing the frequency dependent variations of the quality factor Q. This is a problematic example of FWI modeling error.

In this study,we explore the use of neural networks for quantifying this type of forward modeling uncertainty in viscoelastic FWI. A critical and ongoing area of research is the evaluation of uncertainty in neural network predictions. Our focus is on employing the Monte Carlo (MC) dropout method to approximate the posterior uncertainty (Gal and Ghahramani, 2016; Abdar et al., 2021) of the predicted relaxation time variables after training. The structure of this abstract is organized as follows: First, we detail the process of training a neural network to transform input attenuation level Q values into relaxation time variables, adhering to the standard linear solid model under the assumption of a constant Q. To mitigate over-fitting, dropout is employed during training, which is subsequently used post-training to approximate posterior uncertainty. We then ascertain the posterior uncertainty of the relaxation variables for the constant Q model by applying dropout to the neural network post-training and evaluating its outputs multiple times. Following this, we calculate viscoelastic FWI using a Recurrent Neural Network (RNN) (Zhang et al., 2021), ensuring no inverse crime is committed, and assess the modeling error arising from the relaxation time variables' limited capability to accurately represent the constant Q model.

THEORY

Training data are synthetically generated, adhering to the constant Q assumption utilized in this study. This assumption implies that the Q spectrum remains constant Kjartansson (1979) across the frequencies within the exploration seismic bandwidth. Accordingly, random Q values, maintained as constant across discrete angular frequencies ω_i, are generated to serve as labeled data for training purposes. A total of 2000 random Q values are produced, with 1000 being integer values ranging

from 1 to 1000, and the remaining 1000 values derived from a Uniform distribution bounded between 1 and 1000. These random Q values, despite their randomness, are all treated as constants with respect to frequencies and are denoted as $Q_{\text{label}}(\omega)$.

In our formulation, the relaxation variables generated by the neural network are expressed as follows:

$$\mathscr{N}_{\text{mlp}}(Q;\boldsymbol{\theta}) = \tau_{\varepsilon_{\text{nn}_1}}, \tau_{\sigma_{\text{nn}_1}}, \tau_{\varepsilon_{\text{nn}_2}}, \tau_{\sigma_{\text{nn}_2}}, \cdots, \tau_{\varepsilon_{\text{nn}_L}}, \tau_{\sigma_{\text{nn}_L}}, \quad (1)$$

where $\mathscr{N}_{\text{mlp}}(Q;\boldsymbol{\theta})$ denotes the neural network function mapping the Q value to the relaxation times, parameterized by $\boldsymbol{\theta}$, taking Q value as input, where $\boldsymbol{\theta}$ is the weight vector containing all the trainable parameters. Here, $\tau_{\varepsilon_{\text{nn}_l}}$ and $\tau_{\sigma_{\text{nn}_l}}$ represent the strain and stress relaxation times, respectively, for the l^{th} standard linear solid mechanism generated with the neural network. The neural network employed in this study is a multilayer perceptron (MLP) network, recognized for its simplicity, comprising fully connected layers, activation functions, and the dropout method. Taking the MLP network for training the 2 SLS relaxation time variable as an example, this network will accept a Q value as input ($\mathbb{R}^1$) and outputs relaxation time variables ($\mathbb{R}^4$), specifically $\tau_{\varepsilon_{\text{nn}_1}}$, $\tau_{\varepsilon_{\text{nn}_2}}$, $\tau_{\sigma_{\text{nn}_1}}$, and $\tau_{\sigma_{\text{nn}_2}}$ for the two SLS layers. We utilize Tanh() activation functions for the hidden layers and a Sigmoid() activation function for the output layer, ensuring the relaxation time variables are positive.

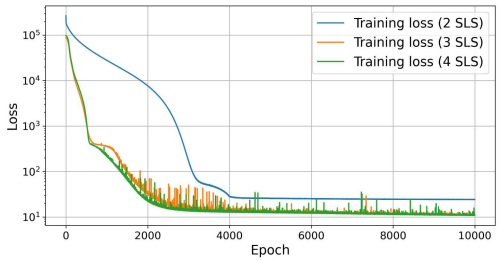

Figure 1: Variation of training loss values in generating relaxation variables with an MLP under the constant Q Model. $L = 2$ SLS (blue line), $L = 3$ SLS (yellow line), $L = 4$ SLS (green line).

Frequency-dependent Q modeled with the relaxation variables generated by the neural network $\mathscr{N}_{\text{mlp}}$ is denoted as Q_{nn}, which has the formulation of:

$$Q_{nn}(\omega, \mathscr{N}_{\text{mlp}}(Q;\boldsymbol{\theta})) = \frac{1 - L + \sum_{l=1}^{L} \frac{1+\omega^2 \tau_{\varepsilon_{\text{nn}_l}} \tau_{\sigma_{\text{nn}_l}}}{1+\omega^2 \tau_{\sigma_{\text{nn}_l}}^2}}{\sum_{l=1}^{L} \frac{\omega(\tau_{\varepsilon_{\text{nn}_l}} - \tau_{\sigma_{\text{nn}_l}})}{1+\omega^2 \tau_{\sigma_{\text{nn}_l}}^2}}. \quad (2)$$

The objective function for neural network training is defined as:

$$\theta^* = \underset{\boldsymbol{\theta}}{\text{argmin}} \sqrt{\sum_{i=1}^{I} \left[\frac{1}{Q_{nn}\left(\omega_i, \mathscr{N}_{\text{mlp}}(Q;\boldsymbol{\theta})\right)} - \frac{1}{Q_{\text{label}}(\omega_i)} \right]^2}, \quad (3)$$

where I is the maximum number of discrete angular frequencies. A well-trained network provides a continuous mapping function relating the Q values to the relaxation variables, implying that the neural network can map any Q value within the range of interest to its corresponding relaxation variables on the real axis. The Adam algorithm, with a learning rate of 10^{-4}, is used as the optimizer. The stopping condition is set to a maximum training epoch of 10000, with a dropout rate of 0.01 employed to prevent overfitting during the training phase. The objective function value during the training stage for SLS layers 2, 3, and 4, with respect to the training epoch, are plotted in Figure 1. It is observed that as the number of epochs increases, the objective function value decreases accordingly; around 4000 iterations, the objective function values decrease to a level where the rate of change is minimal, indicating the convergence of the training process.

In Figure 2(a), we plot the predicted $Q_{nn}(\omega_i, \mathscr{N}_{\text{mlp}}(Q;\boldsymbol{\theta}))$ for Q=30. Using the distribution of the relaxation variables, generated with the post-trained $\mathscr{N}_{\text{mlp}}$, with a dropout rate 0.5, will lead us to a series of synthetic Q spectrum plotted as gray lines in Figure 2. The area spanned by the gray lines can be seen as the uncertainty quantification of the synthetic Q spectrum. The synthetic Q spectrum, calculated from the mean generated relaxation time variables plotted as the yellow line. From Figure 2, we can see that the mean predicted line matches the constant Q line with an acceptable error. Similar observations can be seen from the predictions of other Q values. In Figure 2(b), (c), and (d), we illustrate the comparisons for Q values of 80, 115, and 170, respectively. We can also see that the synthetic Q spectrum provided by the mean relaxation time variable results could better represent the constant Q spectrum within the expiration seismic bandwidth of interest.

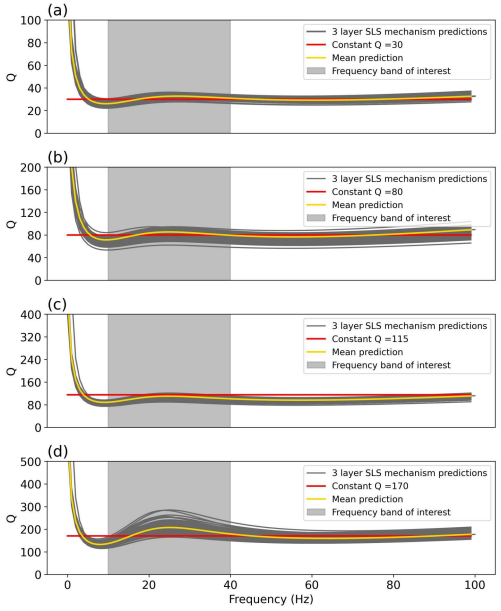

Figure 2: Prediction of constant Q models using a three-layer Standard Linear Solid (SLS) mechanism, with the stress and strain relaxation times generated from a well-trained neural network employing a 0.5 dropout rate. The Q spectrum, represented by the yellow line, is calculated from the mean relaxation time variables generated by the neural network

NUMERICAL INVERSION TEST

Inversion overview

To better elucidate the impact of modeling errors on inversion

results, we calculate viscoelastic FWI without committing inverse crime numerically. We will use different numerical modeling methods to calculate the observed and synthetic data. Specifically, the observed data, D_{obs}, are calculated by using the mean relaxation time variables generated by the $\mathcal{N}_{\text{mlp}}$ (3 SLS) 1,000 times. During inversion, we utilize a single realization of the post-trained $\mathcal{N}_{\text{mlp}}$ (2 SLS) to calculate synthetic data. Thus, the numerical methods to calculate D_{obs}, and D_{syn} are different. We use the GSLS viscoelastic wave equation derived according to Fan et al. (2016) with all the relaxation time variables generated with the post-trained neural network.

Next, our goal is to evaluate the modeling error due to a single realization's inability to accurately represent the constant Q model. This is achieved by applying Monte Carlo (MC) dropout in $\mathcal{N}_{\text{mlp}}$ (2 SLS) to produce various relaxation variables and corresponding synthetic data D_{test}. Specifically, we use the inverted Q model to generate test relaxation variables and calculate the misfit between D_{test} and observed data D_{obs}, termed error evaluation loss. If this loss is within the acceptable range of the final FWI loss, D_{test} is considered valid for modeling error quantification. Repeating this process yields a dataset of D_{test} values, with their standard deviation representing the modeling errors. These errors, stemming from the relaxation variables' limited representation of the constant Q model, inform updates to the elastic and attenuation models, aiding in analyzing the impact of modeling errors on model updates.

Marmousi model inversion results

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 ρ are shown in Figure 3 (a), (d), and (g), respectively. The true attenuation models for Q_p and Q_s are depicted in Figure 3 (j) and (m), respectively. The model dimensions are 80×160, with a grid spacing of $d_x = d_z = 20\,\text{m}$. We use $L = 3$ layers of the SLS mechanism and forward propagate the relaxation variables 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 for modeling used here is 20 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 ρ are plotted in Figures 3 (b), (e), and (h). 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. The receivers are located on the model's surface, left and right-most well-log, with an interval of 20m, for good data illumination.

We 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 post-trained neural network (2 SLS). The inversion results for test 3, the elastic models V_p, V_s, and ρ are shown in Figures 3 (c), (f), and (i), respectively. The inversion results for the attenuation models Q_p and Q_s, are displayed in Figures 3 (l) and (o). Compared to the true models, we obtain promising inversion results for both the elastic and attenuation models. Acceptable model errors are located in the lower right corner of the elastic models. The inversion results for Q_s also exhibit some difficulty in accurately describing the shape of the central reservoir.

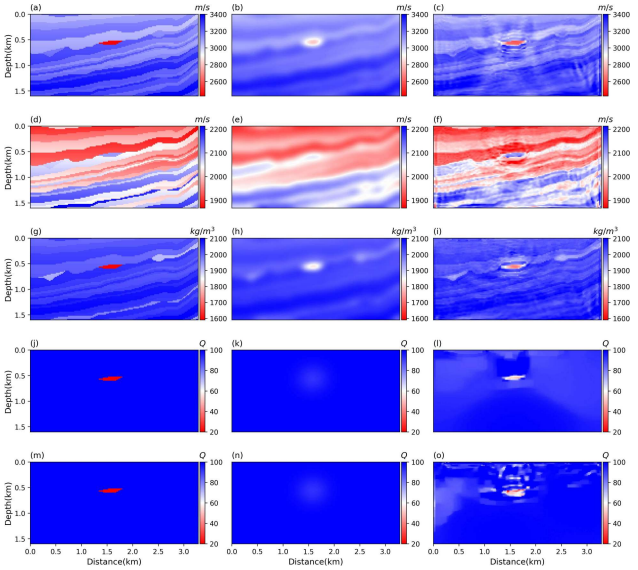

Figure 3: Part of the Marmousi model V_p, V_s, and ρ viscoelastic FWI. The models for V_p, V_s, and density ρ are from top to bottom in rows. From left to right, in columns are the true models, initial models, inversion results for test 1, inversion results for test 2, and inversion results for test 3, respectively.

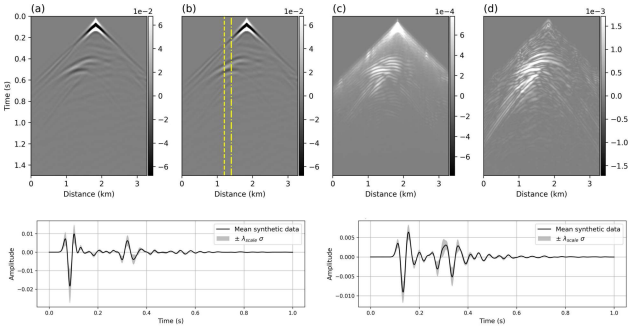

Figure 4: Modeling error evaluation for part of the Marmousi model inversion test. (a) The observed record at 1000 m of the model. (b) Synthetic data from the final iteration of FWI. (c) The standard deviation record of the collected D_{test}, which is regarded as the modeling error. (d) Absolute error between (a) and (b). Lower panel: examples of the mean and scaled standard deviation of two traces calculated from the collected D_{test} data sets.

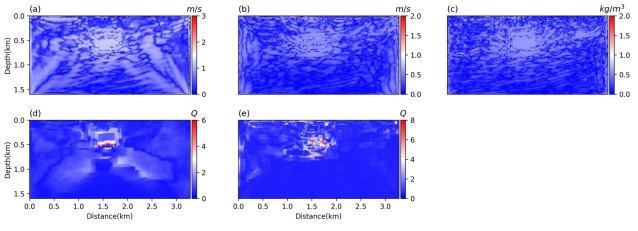

Figure 5: Standard deviation models of 87 updated models for V_p (a), V_s (b), ρ (c), Q_p (d), and Q_s (e). These Standard deviation models represent how the inability of relaxation variables to quantify a constant Q model will influence the FWI results.

Next, we will quantify the modeling error. The MC dropout rate we set here is 0.5, and we will calculate the loop of the

modeling error evaluation 100 times. 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 have data misfit values within the acceptable range. Thus, 87 D_{test} will be collected to evaluate the modeling error. The standard deviation of the 87 D_{test} is regarded as the modeling error evaluated.

In the first row of Figure 4, we display the modeling error evaluated for the shot located at 1000m of the model. We plot the observed data, synthetic data at the final iteration of FWI, the modeling error evaluated, and the absolute error between the observed data and the final synthetic data in Figure 4 (a), (b), (c), and (d), respectively. We can see that the modeling error evaluated in Figure 4(c) has reflected the main feature of the data error in Figure 4(d), especially for the reflection waves arrived at around $0.4s$ of the records. We should not expect the modeling error evaluated to have the ability to capture all the features of the absolute data error, as Figure 4 (d) includes all the errors in inversion. To further observe how the modeling error influences the D_{test}, we plot two traces of the records in Figure 4, where the trace location is marked in Figure 4(b). In the bottom panels of Figure 4, the black lines are the 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 can see that the evaluated modeling error influences both the amplitude and the phase of the synthetic records.

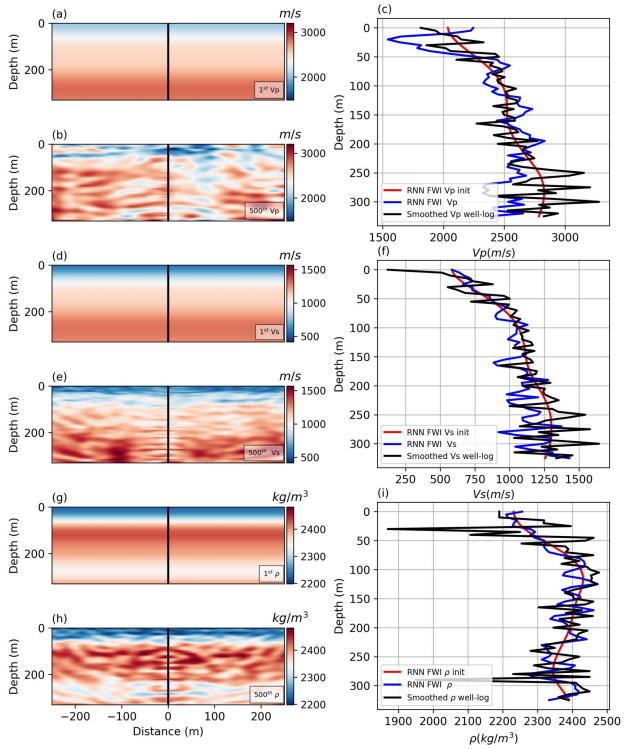

Figure 6: 2022 Snowflake data V_p , V_s, ρ line4 results (2 SLS). (a), (b) and (c) are the initial model, inversion result, and well-log comparison for V_p. (d), (e) and (f) are the initial model, inversion result, and well-log comparison for V_s. (g), (h) and (i) are the initial model, inversion result, and well-log comparison for ρ.

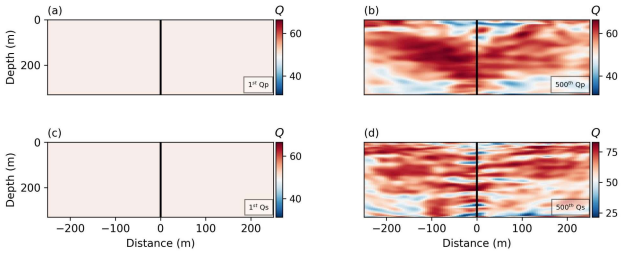

Figure 7: 2022 Snowflake data Q_p , Q_s line4 results (2 SLS). (a), and (b) are the initial model, inversion result for Q_p. (c) and (d) are the initial model, inversion results for Q_s.

Next, we use each of the collected 87 D_{test} datasets to evaluate modeling error. The results are plotted in Figure 5. These Figures represent the impact of the relaxation variable's limited ability to quantify a constant Q model on the FWI results for the elastic Marmousi model section.

Field data inversion results

To substantiate the efficacy of integrating networks within the viscoelastic FWI framework, we calculate inversion analysis using the line4 field data from the 2022 VSP Snowflake dataset (Hall et al., 2019). The study area is located near Brooks, Alberta, focusing on innovative approaches for monitoring carbon dioxide sequestration within the Basal Belly River Sandstone formation, situated approximately 300 meters below the surface. As of the end of 2023, approximately 85 tonnes of CO_2 had been injected into this formation. The inversion results for V_p, V_s, and density (ρ) models are illustrated in Figure 6. Initial models were derived by applying a smoothing filter to well-log data centered within the study model. These inversion outcomes are juxtaposed with the smoothed well-log data, as depicted in Figures 6 (c), (f), and (i). The comparison reveals a promising correlation between the inversion results and the well-log data. The inversion results for Qp, and Q_s are plotted in Figure 7.

CONCLUSION

In this study, we develop a neural network to correlate the attenuation factor Q with relaxation times for various SLS mechanisms. Our results show that the Q spectrum, obtained from the mean relaxation variables of a well-trained network, effectively represents the constant Q model within the seismic frequency range. This network facilitates direct Q model updates in an RNN-based inversion framework. Additionally, we explore the impact of modeling errors on viscoelastic FWI, focusing on errors from inadequate representation of the Q spectrum by relaxation time variables. Through numerical and field data inversion tests, we highlight the significant our methodology's ability to calculate viscoelastic FWI.

ACKNOWLEDGMENTS

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