Uncertainty estimation in seismic inversion using Invertible Neural Networks

Arnab Dhara[1], **Ke Wang**[2], **Jinsong Chen**[2], **Hermes Malcotti**[2], **Reynaldo Cardona**[2]

[1]Institute for Geophysics, The University of Texas at Austin
[2]Chevron Technical Center
adhara@utexas.edu, Ke.Wang@chevron.com, JChen@chevron.com, HMalcotti@chevron.com, RCardona@chevron.com

Abstract

The main goal of seismic inversion is to predict subsurface rock properties from seismic data. Seismic inversion algorithms have served as a viable reservoir characterization tool for decades . Neural network based inversion techniques has become widely popular in recent years. Compared to traditional inversion methods, neural networks provide a high-resolution estimation of rock properties in a reasonable time. But they give a single solution, which is a "best" estimate or the most likely value of the property of interest. However, seismic inversion is a one-to- many problem. Thus the inverse problem is ill-posed and non- unique. Several probabilistic algorithms have been developed over the years to tackle the issue of non-uniqueness. Probabilistic inversion algorithms, for example the Markov Chain Monte Carlo (MCMC) methods, give a probability distribution of the model parameters of interest. Thus we get uncertainty estimates along with the most likely value. But the solution from MCMC methods are generated at a significant computational cost. To overcome such issues, we investigate the use of Invertible Neural Networks (INNs) for uncertainty quantification. INNs can solve probabilistic inverse problems and can provide approximations to complicated posterior distributions. Tests on our model demonstrate that the developed method can accurately predict elastic properties well, and can quantify uncertainty. We also compare our uncertainty es- timates with those found using MCMC methods and find them to be consistent.

Introduction

Geoscientists use well and seismic data to develop subsurface models for carbon sequestration, natural resources exploration and scientific studies for evolution of the earth. Seismic inversion is an important tool that enables us to combine the well data with seismic data to generate models of subsurface properties. Pre-stack seismic data can be converted to elastic subsurface properties using Amplitude-vs-Offset inversion. Several deterministic and probabilistic algorithms (Grana et al. 2022) have been developed throughout the decades for Amplitude-vs-Offset (AVO) inversion. Recently, several authors have proposed different flavours of Deep Neural Networks (DNN) for the problem of AVO inversion. Neural networks based inversion methods can successfully generate a high resolution estimate of the model parameters by learning a mapping between data and model parameters. But seismic inverse problems have nonunique solutions due to data noise, nonlinearity of the physical parameters and data, and the fundamental physics of the problem. Therefore, uncertainties in AVO inversion need to be quantified to interpret inversion results. However, most neural network based methods trained using a supervised approach do not provide reliable uncertainty estimates. Probabilistic inversion methods like Markov Chain Monte Carlo (MCMC) based sampling methods can be used to derive useful posterior probability statistics because they can generate a set of samples from the posterior probability density function of the elastic model parameters. The probability density function (pdf) can describe the nonuniqueness of the parameters. But the solution from MCMC methods are generated at a significant computational expense since MCMC methods require several forward modeling runs to generate samples.This motivated us to use Invertible Neural Networks (INNs) to solve the probabilistic AVO inverse problem. Ardizzone et al. (2018) first used invertible neural networks to solve the inverse problem and demonstrated that the method can provide accurate approximation to multimodal posterior distribution. Unlike the conventional neural network which learns only the inverse mapping from data to model parameters, INNs provide bijective mapping between data (input) and models (output). Since the network also learns the forward mapping between input and output, it uses the information loss during the forward mapping to improve its prediction. Moreover, because of the additional latent variables introduced by the network on the output side it is possible for the network to learn the full posterior distribution. Subsequently, INNs have been applied to solve several inverse problems, for example, Zhang and Curtis (2021). However, the success of such applications has been limited to low dimensional problems. In this work, we investigate the use of Invertible Neural Networks for two kinds of high-dimensional AVO inversion problems : post-stack inversion and pre-stack inversion. The goal of post stack inversion is to get an estimate of P-impedance. While prestack inversion is more complex and generates estimates of P-impedance, S-impedance and density. The INN framework is not designed to estimate uncertainty due to data noise. We overcome this issue by including noise as an additional parameter in training. Our examples demonstrate that once an INN is trained, it can produce posterior distribution much faster compared to MCMC based methods. We also compare our INN results with that of MCMC and find them to be consistent.

Methodology

Problem

Our aim here is to solve the AVO inverse problem for elastic properties, which are the model parameters, $\mathbf{m}$ based on the observed data, i.e., seismic measurements $\mathbf{d_{obs}}$. The forward model G defines the relation between the model parameters and the data.

$$\mathbf{d} = G(\mathbf{m}) + \mathbf{e_d}, \tag{1}$$

where $\mathbf{e_d}$ is the modeling error.
For the forward modeling G, we use the following equation to generate reflectivity in case of post-stack inversion

$$R_{pp}(t) = \frac{d\, ln\, z_p(t)}{dt} \tag{2}$$

where $z_p(t)$ is the P-impedance as a function of time. For the pre-stack case, we use the (Zoeppritz 1919) equation to calculate the reflectivity from elastic model properties and then we generate seismic data at a particular time t as a convolution of the reflectivity $\mathbf{R_{pp}}(t, \theta)$ and the wavelet $s(t, \theta)$.
Since the inverse problem is non-unique, an infinite number of models can produce acceptable solutions. So, we use the Bayesian approach which gives a measure of uncertainty in the estimation of our model parameters, by generating a posterior probability distribution (PPD) of model parameters ($\mathbf{m}$). This PPD is estimated by generating samples. The predominant methodology for sampling from such a probability density is Markov chain Monte Carlo (MCMC) (Chen and Hoversten 2012).

Invertible Neural Networks

Invertible neural networks have several properties that make them useful for probabilistic inverse problem. They have bijective mappings between input and output. We build our INN using a set of GLOW coupling blocks (Kingma and Dhariwal 2018).The basic unit of GLOW block is a reversible block consisting of two complementary affine coupling layers. A GLOW coupling block consists of *actnorm* step, followed by an invertible 1 × 1 convolution layer and affine transformation layer. Given an input vector $\mathbf{m}$, the affine transformation layer splits it into two halves $\mathbf{m_1}$ and $\mathbf{m_2}$, which are transformed by an affine function with coefficients $exp(s_i)$ and t_i $(i \in 1, 2)$, using element-wise multiplication ($\odot$) and addition:

$$\mathbf{v_1} = \mathbf{m_1} \odot exp(s_2(\mathbf{m_2})) + t_2(\mathbf{m_2}) \tag{3}$$

$$\mathbf{v_2} = \mathbf{m_2} \odot exp(s_1(\mathbf{v_1})) + t_1(\mathbf{v_1}) \tag{4}$$

In our implementation, the mappings s_i and t_i could be a succession of two fully connected layers or two convolutional layers (depending on the position of the network) with Sigmoid activations. Our deep invertible network is composed of a sequence of four reversible blocks with invertible flatten and reshape layers in between. To increase model capacity, we insert permutation layers between the reversible blocks, which shuffle the elements of the subsequent layer's input in a randomized, but fixed way.

INN based probabilistic AVO inversion

Our AVO inverse problem has a non-unique solution. To account for uncertainties in the solution, the $\mathbf{d}$ vector is augmented with additional latent variables $\mathbf{z}$ (Ardizzone et al. 2018). AVO inversion is a one-to-many process. The latent variables make the pair comprising the augmented data $(\mathbf{d}, \mathbf{z})$ and model $\mathbf{m}$ and unique. So, the network can now map a unique model parameter $\mathbf{m}$ to a unique pair $(\mathbf{d}, \mathbf{z})$ of measurements and latent variables. The output of the network in the forward direction is given by $f(\mathbf{m}; \theta)_d$. Our network is trained to approximate the forward modeling. After training, $f(\mathbf{m}; \theta)_d \approx G(\mathbf{m})$ where subscript $\mathbf{d}$ represents the data estimate from the network, and the latent variable $\mathbf{z}$ predicted by the network is constrained to be a normal distribution. We run the network in the reverse direction with our specific measurement $\mathbf{d_{obs}}$ and latent variable $\mathbf{z}$ selected randomly from the same normal distribution to obtain the model parameters corresponding to the data:

$$\mathbf{m} = f^{-1}(\mathbf{d_{obs}}, \mathbf{z}, \theta), \mathbf{z} \approx \mathcal{N}(0, \mathbf{I}) \tag{5}$$

We take many samples of $(\mathbf{d_{obs}}, \mathbf{z})$ with a constant $\mathbf{d_{obs}}$ and multiple samples of $\mathbf{z}$ generated from normal distribution. The trained network then gives the posterior distribution $p(\mathbf{m}|\mathbf{d_{obs}})$ by transforming the distribution prior $p(\mathbf{z})$. With the posterior distribution, we can characterize the uncertainty in the estimation of the model parameters. The original INN only accounted for innate uncertainties in parameter estimation because of the physics. To account for uncertainty because of data noise, we follow Zhang and Curtis (2021) and treat the noise as additional model parameter. Our INN architecture is shown in Figure 1. Just like conventional inversion, the low frequency is also given as input to the network. To apply the INN to our AVO inversion problem, we go through the following steps:

- The first step is training data set generation. To generate a training dataset, we use the statistics from the well logs and generate pseudo well logs using the statistics. From the pseudo well logs we generate seismic data using a known wavelet.
- To train the network, we use our generated pseudo well logs and the seismic data. The network training is bidirectional. Suppose the network output distribution in the forward direction is $q(\mathbf{d}, \mathbf{z}; \theta)$ and the output of the network in the forward direction is $f(\mathbf{m}, \epsilon; \theta)$, (ϵ is the noise realization and θ are the network weights) the training loss in the forward direction can be defined as:

$$L_{fwd} = \sum_{i=1}^{N} ||\mathbf{d}^i_{obs} - f(\mathbf{m}^i, \epsilon^{\mathbf{i}}; \theta)|| + \\ MMD[q(\mathbf{d}^i, \mathbf{z}^i; \theta), p(\mathbf{d}^i)p(\mathbf{z}^i)] \tag{6}$$

  where each model vector corresponds to single data vector and single latent vector. N is the net batch numbers. $p(\mathbf{d}^i)$ is the prior distribution of data in the ith batch. $p(\mathbf{z}^i)$ is typically a Normal distribution. Maximum Mean Discrepancy (MMD) which is a measure of dissimilarity of two distributions. We also include two loss function

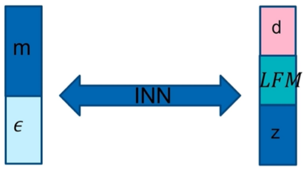

Figure 1: INN maps between $\mathbf{m}$ and $\mathbf{d}$ conditioned by Low Frequency Model (LFM). ϵ is the data noise realization.

on the input side to facilitate convergence.

$$L_{rev} = MMD[q(\mathbf{m}^i, \epsilon^i; \theta), p(\mathbf{m}^i)p(\epsilon^i)] - \gamma \log(p[(\mathbf{m}^i, \epsilon^i) = f^{-1}(\mathbf{d}^i, \mathbf{z}^i; \theta)]|\det J_{f^{-1}}(\mathbf{d}^i, \mathbf{z}^i; \theta)|) \quad (7)$$

where the first term matches the input distribution predicted by the network acting in the reverse direction to the true distribution of the model parameters and the second term is the likelihood term which matches the predicted model parameters to the true model parameters. We noticed that adding the likelihood term increases the rate of convergence. The total loss is a combination of both forward and backward loss

$$L_{total} = L_{fwd} + L_{rev} \quad (8)$$

Experiments and Results

Uncertainty estimation in post stack inversion Our experiments are based on model from Kuito field in Angola. We start with post-stack inversion. Our goal is to estimate P-impedance given our post-stack seismic data. To generate training data, we use a low frequency P-impedance model and perturb it using a gaussian distribution. The seismic data was generated with a wavelet of peak frequency 25 Hz. We used 10000 traces for training. 9500 traces were kept for training and 500 was used for testing. Finally we apply the trained network on our synthetic model. We first tested a case where the seismic had low noise with a Signal-to-Noise (S/N) ratio of 30. We also looked at the case when the seismic had a S/N ratio of 1. Figure 2 compares the result of the mean of all realizations (orange) from invertible neural networks to the true P-impedance (blue). The LFM model is shown in green and the lower and upper bounds of the 95% confidence interval shown in cyan. We compare the uncertainty bounds of the data with low noise to the one with high noise. The uncertainty bounds of P-impedance inversion is higher for the data with noise. From the histogram of the resulting set of samples of $\mathbf{m}$ at 550 ms time stamp, we notice that when noise affects the data, the posterior distribution is wider with values at the tails of the distribution.

Uncertainty estimation in pre-stack inversion Next, we implemented our neural network on the problem of pre-stack inversion. We consider prestack data ranging from 3^o to 30^o with 6 equally placed angles. To train the network, we use Gaussian mixture model with three components and fit it to the well-log data at a location. Using the Gaussian mixture model parameters, we create 16000 pseudo well logs for P-impedance, Vp/Vs and Density. From the pseudo well-log, we generated pre-stack data. frequency to generate pre-stack data. We train the network using 15500 samples and keep 500 samples for testing. Now we apply the trained network to estimate P-impedance, velocity ratio and density for our 2D model. The seismic data and the results of the inversion are shown in figure 3a and 3b respectively. The top figure shows the true P-impedance, velocity ratio and density while the bottom plots show the mean P-impedance, velocity ratio and density samples from our invertible neural network. We plot the traces at CDP 800 in figure 4.

Comparison to MCMC We compare marginal pdfs by histogramming samples from INN based inversion and MCMC based approach for the same problem. The MCMC approach is discussed in detail in Chen and Hoversten (2012). We look at the samples from a window length of 10 ms at two time stamps (Figure 3). The posterior distribution from the two methods looks similar. The MCMC method took 848 seconds whereas the run time for the INN was 0.5 seconds to generate results for one location.

Conclusion

We investigated the use of INNs to solve AVO inverse problems in this project. We also looked at the posterior pdfs given by the network and compared the results to posterior pdfs estimated by MCMC. The INN and MCMC based inversion show similar posterior distribution. After training, INNs can predict posterior pdfs in less than a second thus making them an attractive tool for uncertainty estimation.

Acknowledgements

We thank Chevron Technical Center, a division of Chevron U.S.A. Inc. for the permission to publish this work.

References

Ardizzone, L.; Kruse, J.; Wirkert, S.; Rahner, D.; Pellegrini, E. W.; Klessen, R. S.; Maier-Hein, L.; Rother, C.; and Köthe, U. 2018. Analyzing inverse problems with invertible neural networks. *arXiv preprint arXiv:1808.04730*.

Chen, J.; and Hoversten, G. M. 2012. Joint inversion of marine seismic AVA and CSEM data using statistical rock-physics models and Markov random fields. *Geophysics*, 77(1): R65–R80.

Grana, D.; Azevedo, L.; de Figueiredo, L.; Connolly, P.; and Mukerji, T. 2022. Probabilistic inversion of seismic data for reservoir petrophysical characterization: Review and examples. *Geophysics*, 87(5): M199–M216.

Kingma, D. P.; and Dhariwal, P. 2018. Glow: Generative flow with invertible 1x1 convolutions. *Advances in neural information processing systems*, 31.

Zhang, X.; and Curtis, A. 2021. Bayesian geophysical inversion using invertible neural networks. *Journal of Geophysical Research: Solid Earth*, 126(7): e2021JB022320.

Zoeppritz, K. 1919. Erdbebenwellen vii. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1919: 57–65.

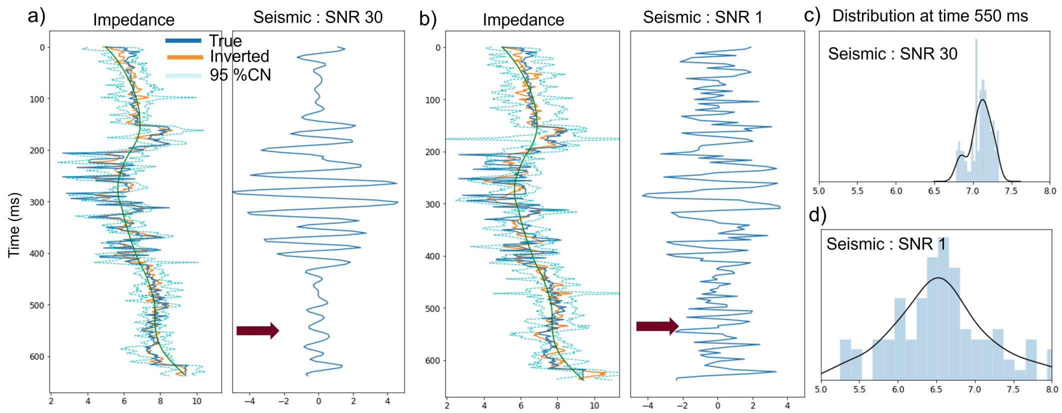

Figure 2: Poststack inversion : a) SNR 30 b) SNR 1. Posterior distribution at 550 ms with c) SNR 30 d) SNR 1

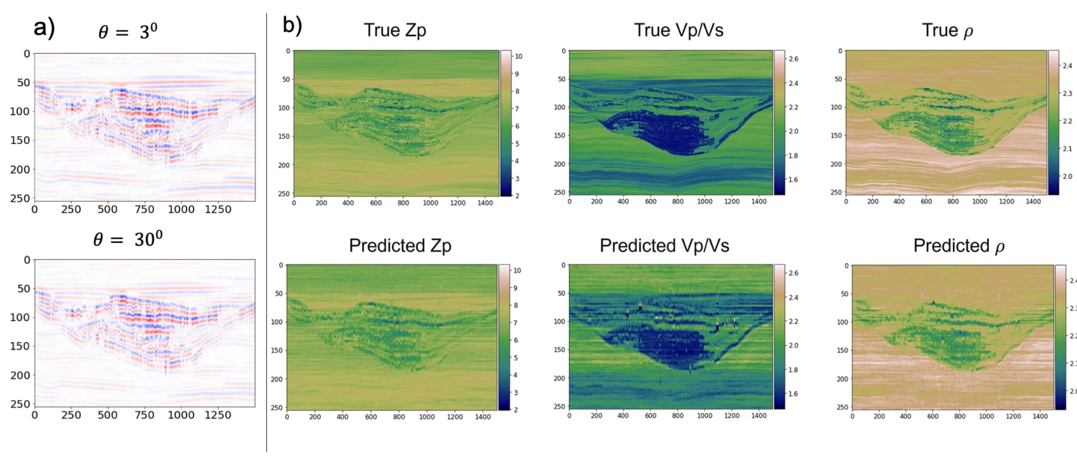

Figure 3: a) Prestack data b) Prestack inversion result: mean of realizations compared against true properties

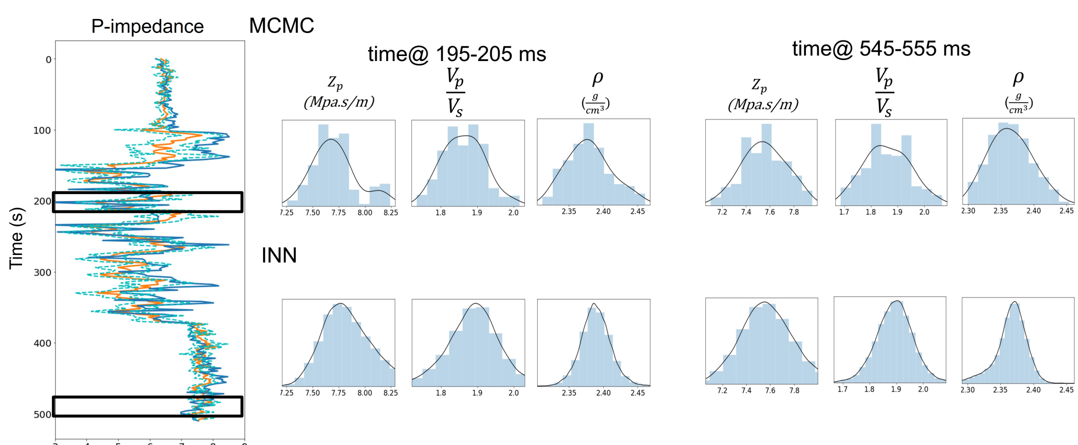

Figure 4: Prestack inversion : Posterior distribution comparison at two locations