

# Joint Inversion of Geophysical Data for Geologic Carbon Sequestration Monitoring: A Differentiable Physics-Informed Deep Learning Model

Mingliang Liu<sup>1</sup>, Dario Grana<sup>2</sup>, and Tapan Mukerji<sup>1</sup>

<sup>1</sup>Stanford University

<sup>2</sup>University of Wyoming

November 24, 2022

## Abstract

Geophysical monitoring of geologic carbon sequestration is critical for risk assessment during and after carbon dioxide (CO<sub>2</sub>) injection. Integration of multiple geophysical measurements is a promising approach to achieve high-resolution reservoir monitoring. However, joint inversion of large geophysical data is challenging due to high computational costs and difficulties in effectively incorporating measurements from different sources and with different resolutions. This study develops a differentiable physics model for large-scale joint inverse problems with reparameterization of model variables by deep neural networks and implementation of a differentiable programming approach of the forward model. The main novelty is the use of automatic differentiation and parallel computing for efficient multiphysics data assimilation. The application to the Sleipner benchmark model demonstrates that the proposed method is effective in estimation of reservoir properties from seismic and resistivity data and shows promising results for CO<sub>2</sub> storage monitoring.

# **Joint Inversion of Geophysical Data for Geologic Carbon Sequestration Monitoring: A Differentiable Physics-Informed Deep Learning Model**

**Mingliang Liu<sup>1</sup>, Dario Grana<sup>2</sup> and Tapan Mukerji<sup>1</sup>**

<sup>1</sup> Department of Energy Resources Engineering, Stanford University, USA.

<sup>2</sup> Department of Geology and Geophysics, University of Wyoming, USA.

Corresponding author: Mingliang Liu ([mliu9@stanford.edu](mailto:mliu9@stanford.edu))

## **Key Points:**

- A novel inverse model is developed by combining differentiable physics and deep neural networks
- The developed model provides an accurate and efficient approach for the joint inversion of geophysical data from different sources
- The inverted models accurately characterize subsurface properties and structures and identify the migration of CO<sub>2</sub> plume

## ABSTRACT

Geophysical monitoring of geologic carbon sequestration is critical for risk assessment during and after carbon dioxide (CO<sub>2</sub>) injection. Integration of multiple geophysical measurements is a promising approach to achieve high-resolution reservoir monitoring. However, joint inversion of large geophysical data is challenging due to high computational costs and difficulties in effectively incorporating measurements from different sources and with different resolutions. This study develops a differentiable physics model for large-scale joint inverse problems with reparameterization of model variables by deep neural networks and implementation of a differentiable programming approach of the forward model. The main novelty is the use of automatic differentiation and parallel computing for efficient multiphysics data assimilation. The application to the Sleipner benchmark model demonstrates that the proposed method is effective in estimation of reservoir properties from seismic and resistivity data and shows promising results for CO<sub>2</sub> storage monitoring.

## Plain Language Summary

This study develops a complete inversion model for the joint quantification and interpretation of geophysical measurements from different sources for geologic carbon sequestration monitoring. By combining deep neural networks for model reparameterization and differentiable programming for inverse modeling, the developed approach accurately characterizes subsurface reservoirs, it identifies the migration of CO<sub>2</sub> plume, and it quantifies global parameters that are uncertain in the forward models. One of the major advantages of this method is that all components in the model are seamlessly integrated and updated simultaneously. Moreover, the model can be easily deployed to high-performance computing platforms, thereby providing a computationally efficient approach for large geophysical data. Therefore, the developed model illustrates promising results for geophysical subsurface monitoring.

## 1. Introduction

Geologic CO<sub>2</sub> sequestration (GCS) is one of the main mitigation strategies to prevent carbon dioxide (CO<sub>2</sub>) from entering the atmosphere by capturing and injecting it into subsurface geological formations, such as depleted oil reservoirs and deep saline aquifers (Metz et al. 2005; Aminu et al. 2017). To ensure safe and long-term CO<sub>2</sub> storage, geophysical monitoring surveys are periodically acquired to surveil the behavior of the injected CO<sub>2</sub> and make informed decisions for storage management (Davis et al., 2019).

Time-lapse (4-D) seismic data is one of the monitoring techniques widely deployed in GCS projects, where a series of seismic surveys are sequentially acquired to quantify changes in reservoir properties (e.g., elastic properties, fluid pressure and saturation) during and after CO<sub>2</sub> injection (Chadwick et al., 2010; Caspari et al., 2011; Rubino et al., 2011; Grude et al., 2013; Egorov et al., 2017; Glubokovskikh et al., 2020; Li and Li, 2021). However, in some cases, depending on the rock stiffness, the time-lapse seismic response might be insensitive to fluid changes. (Gasperikova et al., 2022). To address this issue, non-seismic monitoring techniques have been investigated and successfully applied to GCS. For example, interferometric synthetic-aperture radar (InSAR) is an effective tool for measuring reservoir surface displacements and to infer reservoir pressure changes (Tang et al., 2022). Time-lapse gravity data are sensitive to bulk density, hence mass balance changes due to CO<sub>2</sub> injection (Bonneville et al., 2022). Electromagnetic (EM) and electrical techniques are used to estimate CO<sub>2</sub> saturation because of the distinct resistivity differences between brine- and CO<sub>2</sub>-saturated rocks (Bhuyian et al., 2012). The spatial resolution of non-seismic surveys, however, is generally lower than the resolution of seismic surveys. Joint time-lapse monitoring of seismic and non-seismic measurements is then a potential approach to achieve high-resolution reservoir monitoring and

66 reduce the uncertainty. This approach has been used in hydrocarbon exploration but it is  
67 relatively new for GCS monitoring. Eliasson et al. (2014) investigated the integration of  
68 controlled-source electromagnetic (CSEM) inversion with full waveform seismic inversion for  
69 CO<sub>2</sub> quantification. Dupuy et al. (2021) developed a Bayesian joint inversion method for CO<sub>2</sub>  
70 storage monitoring with uncertainty quantification. Huang (2022) comprehensively reviewed the  
71 advantages and limitations of different geophysical techniques and their applications in long-  
72 term GCS monitoring.

73 The main challenge of joint geophysical monitoring is how to efficiently assimilate  
74 geophysical measurements from different sources. In general, joint geophysical inversion  
75 methods can be categorized into sequential or simultaneous schemes (Tveit et al., 2020). The  
76 sequential scheme inverts geophysical measurements in separate steps with the resulting model  
77 from one measurement as initial model or constraint for the next step, while the simultaneous  
78 scheme inverts all measurements in a single step. In the sequential scheme, measurements with  
79 lower resolution are generally inverted first, such that low frequency models are used as initial  
80 models for the inversion of higher resolution datasets. In the simultaneous scheme, model  
81 parameters are coupled with structural or rock-physics constraints and the forward operator  
82 includes multi-physics models. For this reason, simultaneous schemes are preferred to sequential  
83 schemes despite the higher computational complexity.

84 Recently, scientific machine learning (SciML) is getting increasing attention in the  
85 geophysics community, especially for its ability to address the computational challenges in large-  
86 scale geophysical inverse problems. These approaches mainly involve physics-informed neural  
87 networks (PINNs) (Raissi et al., 2019; Rasht-Behesht et al., 2022) and differentiable physics  
88 (Hernández et al., 2022). PINNs embed the governing physics laws (i.e., partial differential

equations) into deep neural networks (DNNs) as prior information to overcome the issue of limited data, bridging the gap between the data scarcity of inverse problems and data requirements of DNNs. PINNs are a mesh-free method that is theoretically advantageous for high-dimensional problems. However, PINNs still face difficulties of training and convergence in practice due to the lack of information about connectivity between neighboring points (or nodes). Differentiable physics provides an alternative approach for physics modeling by integrating differentiable programming with classical numerical methods (e.g., finite difference method and finite element method). Differentiable programming is a new programming paradigm in which all numerical operations are implemented by differentiable building blocks. Such differentiable mechanism enables solving the forward and inverse problems via automatic differentiation and can be easily combined with DNNs. Differentiable programming has been applied in geophysics, such as geostatistical seismic inversion (Liu and Grana, 2019), full waveform inversion (Zhu et al., 2022) and subsurface flow problems (Li et al., 2020).

To address the computational challenges in geophysical CO<sub>2</sub> monitoring, we propose a differentiable physics model that assimilates time-lapse seismic and resistivity data to predict porosity and the evolution of CO<sub>2</sub> saturation over time. In this approach, large geological models are represented by a DNN with fewer parameters than the actual model and geophysical forward operators are implemented using the differentiable programming framework. We also use a Bayesian neural network (BNN) to estimate global parameters that are uncertain in the forward physics model. By adopting automatic differentiation and graphics processing unit (GPU), model parameters are automatically updated in a computationally efficient manner to minimize the mismatch between predicted and true measurements.

## 2. Methods

In this study, we focus on joint inversion of time-lapse pre-stack seismic and resistivity data, with the goal of predicting porosity and the evolution of CO<sub>2</sub> saturation over time. Seismic data are linked to the reservoir properties (i.e., porosity and fluid saturations) through rock physics models, e.g., the unconsolidated sand model (Dvorkin and Nur, 1996) and convolution-based seismic wave propagation model (Yilmaz, 2001), whereas resistivity is linked to porosity and saturation through empirical equations, e.g., Archie's law (Archie, 1942). Rock physics parameters, such as coordination number and critical porosity in the unconsolidated sand model and the cementation and saturation exponents in Archie's law, are often uncertain in practical applications; hence, we assume that those rock physics parameters are unknown global model parameters. The proposed differentiable physics inversion is illustrated in Figure 1a. The global model parameters (coordination number, critical porosity, and cementation and saturation exponents) and reservoir model spatial variables (porosity and CO<sub>2</sub> saturations) are re-parameterized by a BNN and a convolutional neural network (CNN), respectively. Then, time-lapse seismic and resistivity data are predicted from the output model parameters of the neural networks through the forward models implemented by differentiable programming. With the technique of automatic differentiation, the network parameters are iteratively updated to minimize the loss function. After training, the inverted global parameters and reservoir properties can be obtained as the outputs of the trained BNN and CNN, respectively.





gray circles ( $\mathbf{z}_g$  and  $\mathbf{z}_p$ ) represent constant inputs (latent vectors) that are non-trainable. Other symbols and notations are defined in Section 2.

## 2.1 Forward Modeling

The forward operator for the prediction of the seismic response includes a rock physics model that maps petrophysical properties (porosity and water saturation) to elastic properties (P- and S-wave velocity and density) and a convolution-based seismic wave propagation model that computes seismic travel time and reflection amplitudes. The mathematical-physical equation of the forward seismic model is expressed symbolically as follows

$$\mathbf{d}_{\text{seis}} = \mathbf{f}_{\text{seis}}(\mathbf{f}_{\text{rpm}}(\mathbf{m}_g, \mathbf{m}_p)) + \epsilon_{\text{seis}}, \quad (1)$$

where  $\mathbf{d}_{\text{seis}}$  represents the predicted seismic data,  $\mathbf{m}_g$  represents the global model parameters (coordination number, and critical porosity),  $\mathbf{m}_p$  represents the spatially-varying petrophysical properties (porosity and water saturation),  $\mathbf{f}_{\text{rpm}}$  is the rock physics model for the prediction of elastic properties  $\mathbf{m}_e$ ,  $\mathbf{f}_{\text{seis}}$  is the seismic wave propagation model, and  $\epsilon_{\text{seis}}$  represents the measurement errors. The rock physics model adopted in this study is the unconsolidated sand model (Dvorkin and Nur, 1996). The detailed descriptions of the unconsolidated sand model and the convolution-based seismic wave propagation model can be found in Appendix A and B, respectively.

The forward operator for resistivity is Archie's law (Archie, 1942), that calculates the electrical resistivity of reservoir rocks  $\mathbf{d}_{\text{res}}$  as a function of  $\mathbf{m}_g$  and  $\mathbf{m}_p$

$$\mathbf{d}_{\text{res}} = \mathbf{f}_{\text{res}}(\mathbf{m}_g, \mathbf{m}_p) = \phi^{-m} R_w S_w^{-n} + \epsilon_{\text{res}}, \quad (2)$$

where  $\mathbf{f}_{\text{res}}$  is the electrical rock physics model (Archie's law), the cementation  $m$  and saturation exponent  $n$  are two empirical parameters,  $\mathbf{m}_p$  includes porosity  $\phi$  and water saturation  $S_w$ ,

while  $\epsilon_{\text{res}}$  is the associated error. Here we assume that CO<sub>2</sub> saturation  $\mathbf{S}_{\text{CO}_2} = 1 - \mathbf{S}_w$  and the resistivity of the brine  $R_w$  is known and constant.

In the following example, we introduce measurement error ( $\epsilon_{\text{seis}}$  and  $\epsilon_{\text{res}}$ ) to the “true” seismic ( $\mathbf{d}_{\text{seis}}$ ) and resistivity ( $\mathbf{d}_{\text{res}}$ ) measurements to avoid the “inversion crime” and discard the error terms when predicting data ( $\tilde{\mathbf{d}}_{\text{seis}}$  and  $\tilde{\mathbf{d}}_{\text{res}}$ ) from inverted models.

## 2.2 Model Reparameterization

Two neural networks ( $\mathcal{N}_g$  and  $\mathcal{N}_p$ ) are used for the reparameterization of the global  $\mathbf{m}_g$  and petrophysical parameters  $\mathbf{m}_p$ :

$$\mathbf{m}_g = \mathcal{N}_g(\mathbf{z}_g; \boldsymbol{\theta}_g), \quad (3)$$

$$\mathbf{m}_p = \mathcal{N}_p(\mathbf{z}_p; \boldsymbol{\theta}_p), \quad (4)$$

where  $\mathbf{z}$  represents the input latent vectors of the neural networks that are sampled from standard Gaussian distribution and are non-trainable, and  $\boldsymbol{\theta}$  represents the trainable parameters (network weights) of the neural networks. The goal is to estimate  $\boldsymbol{\theta}_g$  and  $\boldsymbol{\theta}_p$  instead of  $\mathbf{m}_g$  and  $\mathbf{m}_p$ . Model reparameterization with neural networks provides a convenient way to control the range of model parameters, for example  $\mathbf{S}_{\text{CO}_2}$  and  $\mathbf{S}_w$  within  $[0, 1]$ , which is helpful for the stability of inversion.

Specifically,  $\mathcal{N}_g$  is a BNN that takes a 4-dimensional latent vector  $\mathbf{z}_g$  as input and outputs a 4-dimensional vector (whose elements correspond to coordination number, critical porosity, cementation  $m$  and saturation exponent  $n$ , respectively) after the transformation of three fully-connected layers. As illustrated in Figure 2, BNN is a probabilistic model where unlike standard neural networks the weights  $\boldsymbol{\theta}$  are not fixed but follow a distribution  $p(\boldsymbol{\theta})$  (Dürr et al., 2020) and thus it allows quantifying the uncertainty of the global parameters. The goal of BNN is to find

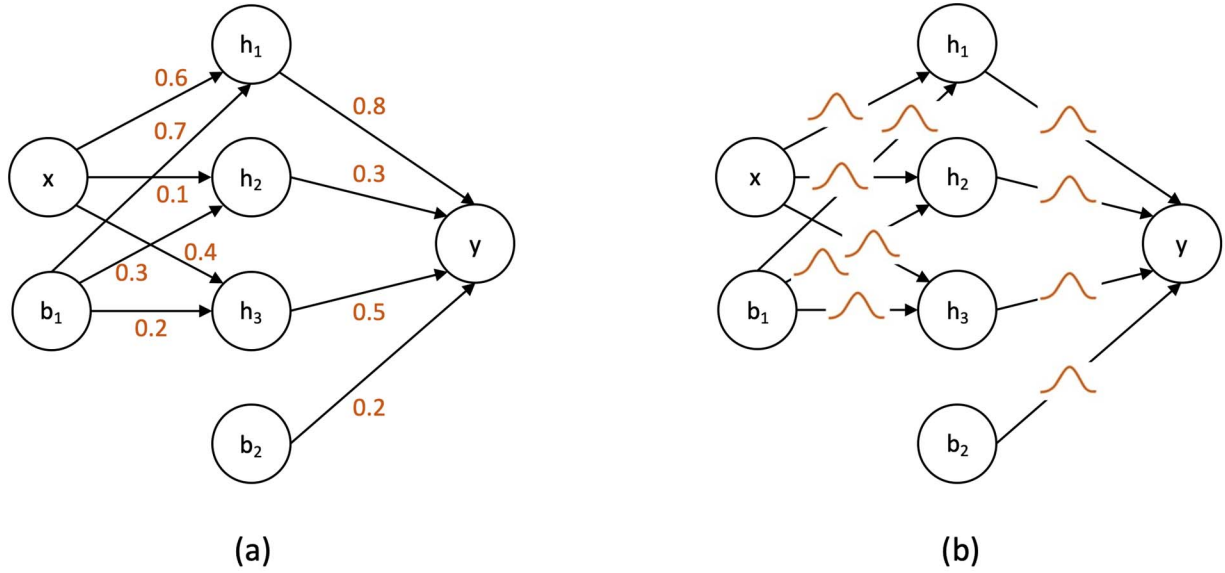
181 the posterior distribution  $p(\boldsymbol{\theta}|\mathcal{D})$  that best fit the data  $\mathcal{D}$ . The distribution of network weights  
182 allows us to quantify the prediction uncertainty. The solution for the posterior distribution  
183  $p(\boldsymbol{\theta}|\mathcal{D})$  is usually not tractable analytically due to the complexity of BNN. We therefore  
184 approximate the posterior distribution with a variational distribution  $q_{\lambda}(\boldsymbol{\theta})$  with variational  
185 parameters  $\lambda$  according to the variational inference approach. In practice, the variational  
186 distribution is generally assumed to be a Gaussian distribution with parameters  $\lambda = (\mu, \sigma)$ , i.e.,  
187 mean  $\mu$  and variance  $\sigma$ . The objective of variational inference is to estimate the variational  
188 parameters  $\lambda$  so that  $q_{\lambda}(\boldsymbol{\theta})$  accurately approximates  $p(\boldsymbol{\theta}|\mathcal{D})$ . Mathematically, this objective is  
189 achieved by minimizing the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951)  
190 between the variational and true distribution

$$191 \quad \text{KL}[q_{\lambda}(\boldsymbol{\theta})||p(\boldsymbol{\theta}|\mathcal{D})] = \int q_{\lambda}(\boldsymbol{\theta}) \log \frac{q_{\lambda}(\boldsymbol{\theta})}{p(\boldsymbol{\theta}|\mathcal{D})} d\boldsymbol{\theta}. \quad (5)$$

192 Therefore, the corresponding optimization problem associated to BNNs is expressed as  
193 determining the optimal variational parameters  $\lambda^*$  as

$$194 \quad \lambda^* = \underset{\lambda}{\operatorname{argmin}} \{ \text{KL}[q_{\lambda}(\boldsymbol{\theta})||p(\boldsymbol{\theta}|\mathcal{D})] - \mathbb{E}_{\boldsymbol{\theta} \sim q_{\lambda}} [\log p(\mathcal{D}|\boldsymbol{\theta})] \}, \quad (6)$$

195 where the first term is the KL divergence and the second (negative) term is the average negative  
196 log-likelihood that measures the data misfit.



**Figure 2.** (a) Standard neural network; (b) Bayesian neural network.

The network  $\mathcal{N}_p$  is a CNN that takes another latent vector  $\mathbf{z}_p$  of size  $n_c \times n_x \times n_y \times n_t$  as input and outputs a  $(N_\tau + 1) \times N_x \times N_y \times N_t$  ( $N_\tau$  is the number of monitoring years,  $N_x$ ,  $N_y$  and  $N_t$  are the number of grids along the  $x$ -,  $y$ - and  $z$ - direction) tensor whose elements correspond to the porosity and CO<sub>2</sub> saturations of  $N_\tau$  years. In the above,  $N_\tau$  is the number of monitoring surveys,  $N_x$ ,  $N_y$  and  $N_t$  are the number of grids along the  $x$ -,  $y$ - and  $z$ - direction, and the output has a dimension of  $(N_\tau + 1)$  as there are  $N_\tau$  surveys and one porosity model that is assumed to be constant over the time of the surveys. In this study, the latent vector is increased by eight times after three up-sampling layers (i.e.,  $N_x = 8n_x$ ,  $N_y = 8n_y$ ,  $N_t = 8n_t$ ), while  $n_c$  is the number of channels in the CNN input layer. The parameters of CNN are listed in Table 1.

**Table 1.** Parameters of each layer in the CNN model.

Layer	Kernel/Scale Factor	Activation function	Output size
1. Input	-	-	64×15×15×28

2.	Conv3D	3×3×3	ReLU	64×15×15×28
3.	Upsample	2	-	64×30×30×56
4.	Conv3D	3×3×3	ReLU	32×30×30×56
5.	Upsample	2	-	32×60×60×112
6.	Conv3D	3×3×3	ReLU	16×60×60×112
7.	Upsample	2	-	16×120×120×224
8.	Conv3D	3×3×3	Sigmoid	4×120×120×224

---

### 2.3 Loss Function

The definition of an appropriate objective function (or cost function) is crucial for large-scale geophysical inverse problems. In this study, the objective function is defined as follows

$$J_{\text{loss}}(\boldsymbol{\theta}_g, \boldsymbol{\theta}_p) = \alpha_1 \frac{1}{N_{\text{seis}}} \|\mathbf{d}_{\text{seis}} - \tilde{\mathbf{d}}_{\text{seis}}\|_2^2 + \alpha_2 \frac{1}{N_{\text{res}}} \|\mathbf{d}_{\text{res}} - \tilde{\mathbf{d}}_{\text{res}}\|_2^2 + \alpha_3 \text{KL}[q(\boldsymbol{\theta}_g) \| p(\boldsymbol{\theta}_g)] + \alpha_4 \frac{1}{N_{\text{CO}_2}} \mathbf{S}_{\text{CO}_2}^T (1 - \mathbf{S}_{\text{CO}_2}), \quad (7)$$

where  $N_{\text{seis}}$ ,  $N_{\text{res}}$  and  $N_{\text{CO}_2}$  are the number of elements of the seismic data, resistivity data and  $\text{CO}_2$  saturation model, respectively;  $\alpha_i$  ( $i = 1, \dots, 4$ ) are the weights of the loss terms. The first two terms in Equation 7 are mean square errors (MSE) that measure the mismatch between “true” and predicted data. The third term is the KL divergence of the Bayesian neural network  $\mathcal{N}_g(\mathbf{z}_g; \boldsymbol{\theta}_g)$  that measures the difference between the variational distribution  $q(\boldsymbol{\theta}_g)$  and the true distribution  $p(\boldsymbol{\theta}_g)$  of the BNN weights. The last term is a regularization term to constrain the model values of  $\text{CO}_2$  saturation in the interval  $[0,1]$  that is necessary for multimodal variables like saturations (Vo and Durlofsky, 2014).

### 2.4 Differentiable Programming

The inversion workflow is implemented according to the differentiable programming framework. Differentiable programming provides an efficient way to jointly estimate model parameters (i.e., global parameters, spatial distribution of porosity, and time-dependent spatial distribution of CO<sub>2</sub> saturation) where all variables are simultaneously updated.

Differentiable programming organizes a program as a Data Flow Graph (DFG). The DFG is a directed acyclic graph in which graph nodes represent mathematical operations (e.g., addition and multiplication) or variables (tensors) and edges correspond to the flow of intermediate values between the nodes (Martin and Estrin, 1967). Figure 1b shows the DFG of the inversion model in this study. Different from traditional programming paradigms that specify explicit instructions to the computer, constructing DFG is like building blocks in which the program is defined by the input-output data and the associated transformations composed of differentiable blocks. The mechanism of DFG makes the program very flexible and scalable to modify and incorporate prior knowledge.

The network parameters  $\theta_g$  and  $\theta_p$  in the DFG are learned from the data by minimizing the loss function in Equation 7

$$[\tilde{\theta}_g, \tilde{\theta}_p] = \arg \min_{\theta_g, \theta_p} J_{\text{loss}}(\theta_g, \theta_p). \quad (8)$$

With the technique of automatic differentiation (Rall, 1981), the gradients of the loss function with respect to  $\theta_g$  and  $\theta_p$  are automatically computed according to the chain rule

$$\frac{\partial J_{\text{loss}}}{\partial \theta_p} = \left( \frac{\partial J_{\text{loss}}}{\partial \mathbf{f}_{\text{seis}}} \frac{\partial \mathbf{f}_{\text{seis}}}{\partial \mathbf{f}_{\text{rpm}}} \frac{\partial \mathbf{f}_{\text{rpm}}}{\partial \mathcal{N}_p} + \frac{\partial J_{\text{loss}}}{\partial \mathbf{f}_{\text{res}}} \frac{\partial \mathbf{f}_{\text{res}}}{\partial \mathcal{N}_p} \right) \frac{\partial \mathcal{N}_p}{\partial \theta_p}, \quad (9)$$

$$\frac{\partial J_{\text{loss}}}{\partial \theta_g} = \frac{\partial J_{\text{loss}}}{\partial \mathbf{f}_{\text{seis}}} \frac{\partial \mathbf{f}_{\text{seis}}}{\partial \mathbf{f}_{\text{rpm}}} \frac{\partial \mathbf{f}_{\text{rpm}}}{\partial \mathcal{N}_g} \frac{\partial \mathcal{N}_g}{\partial \theta_g} + \frac{\partial J_{\text{loss}}}{\partial \mathbf{f}_{\text{res}}} \frac{\partial \mathbf{f}_{\text{res}}}{\partial \mathcal{N}_g} \frac{\partial \mathcal{N}_g}{\partial \theta_g}. \quad (10)$$

The optimal network parameters  $\tilde{\theta}_g$  and  $\tilde{\theta}_p$  are then obtained by the gradient-based

backpropagation method. Here, we use the adaptive moment estimation algorithm (Adam) (Kingma and Ba, 2014), which is computationally efficient for high-dimensional problems.

### 3. Application

We demonstrate the proposed model for joint geophysical monitoring of CO<sub>2</sub> storage with an synthetic model and data modified from the Sleipner 2019 Benchmark Model (<https://co2datashare.org/dataset/sleipner-2019-benchmark-model>). The Sleipner field located in the Norwegian Sea is the first commercial GCS plant and is the longest ongoing CO<sub>2</sub> storage project in the world. The Utsira Formation is the target storage unit, consisting of nine sandstone layers that are vertically separated by eight thin shale layers with low-permeability and a caprock layer. Porosity and permeability of each layer are assumed to be constant in the original benchmark dataset. To make the GCS model more realistic, porosity and permeability of each layer of the reservoir model used in this study are taken to be spatially variable and are geostatistically simulated by sequential Gaussian simulation (SGS) (Deutsch and Journel, 1998) to mimic the spatial heterogeneity of subsurface rocks. Porosity and permeability are assumed to be constant in time. The reference “true” models of porosity and permeability are shown in Figure 3a and 3b, respectively. The GCS model consists of 64×118×263 grid cells. The lateral resolution of the model is 50 m by 50 m and the vertical resolution is about 2 m for the sandstone layers, 0.5 m for the shale layers and 5 m for the caprock. The model parameters such as fluid properties and well configurations are provided in the benchmark dataset. The multiphase flow behavior of CO<sub>2</sub> flow is simulated from the year 1996 to 2010 using the Open Porous Media Flow reservoir simulator (Rasmussen et al., 2021). The detailed parameters of the reservoir are listed in Table 2.



270 **Table 2.** Reservoir model parameters of the example modified from the Sleipner benchmark model.

Layer	Formation	Porosity mean	Porosity range	Permeability mean (mD)	Permeability range (mD)
1	Caprock	0.016	[0.002, 0.036]	5.93e-4	[0, 0.0058]
2	Sandstone	0.17	[0.018, 0.40]	734.44	[0.031, 11969]
3	Shale	0.017	[0.002, 0.04]	7.36e-4	[0, 0.0115]
4	Sandstone	0.18	[0.018, 0.35]	629.81	[0.028, 5172.8]
5	Shale	0.017	[0.002, 0.036]	5.55e-4	[0, 0.0059]
6	Sandstone	0.17	[0.021, 0.35]	559.58	[0.048, 5317.1]
7	Shale	0.018	[0.002, 0.033]	6.40e-4	[0, 0.0041]
8	Sandstone	0.20	[0.027, 0.35]	810.07	[0.12, 5439.1]
9	Shale	0.022	[0.003, 0.034]	0.0010	[0, 0.0048]
10	Sandstone	0.21	[0.024, 0.38]	1065.6	[0.079, 9202.0]
11	Shale	0.021	[0.003, 0.039]	9.75e-4	[0, 0.0097]
12	Sandstone	0.20	[0.019, 0.39]	822.73	[0.037, 9599.8]
13	Shale	0.019	[0.002, 0.037]	7.87e-4	[0, 0.0078]
14	Sandstone	0.20	[0.02, 0.40]	910.26	[0.043, 11269]
15	Shale	0.019	[0.003, 0.036]	8.93e-4	[0, 0.0067]
16	Sandstone	0.19	[0.019, 0.35]	771.31	[0.033, 5152.0]
17	Shale	0.018	[0.003, 0.032]	6.68e-4	[0, 0.0035]
18	Sandstone	0.18	[0.019, 0.35]	701.91	[0.035, 5604.5]

271

272 In this application, we aim to quantify porosity and CO<sub>2</sub> plumes at years 2000, 2003 and

273 2006 from the base seismic and resistivity surveys at year 1996 (before CO<sub>2</sub> injection) and

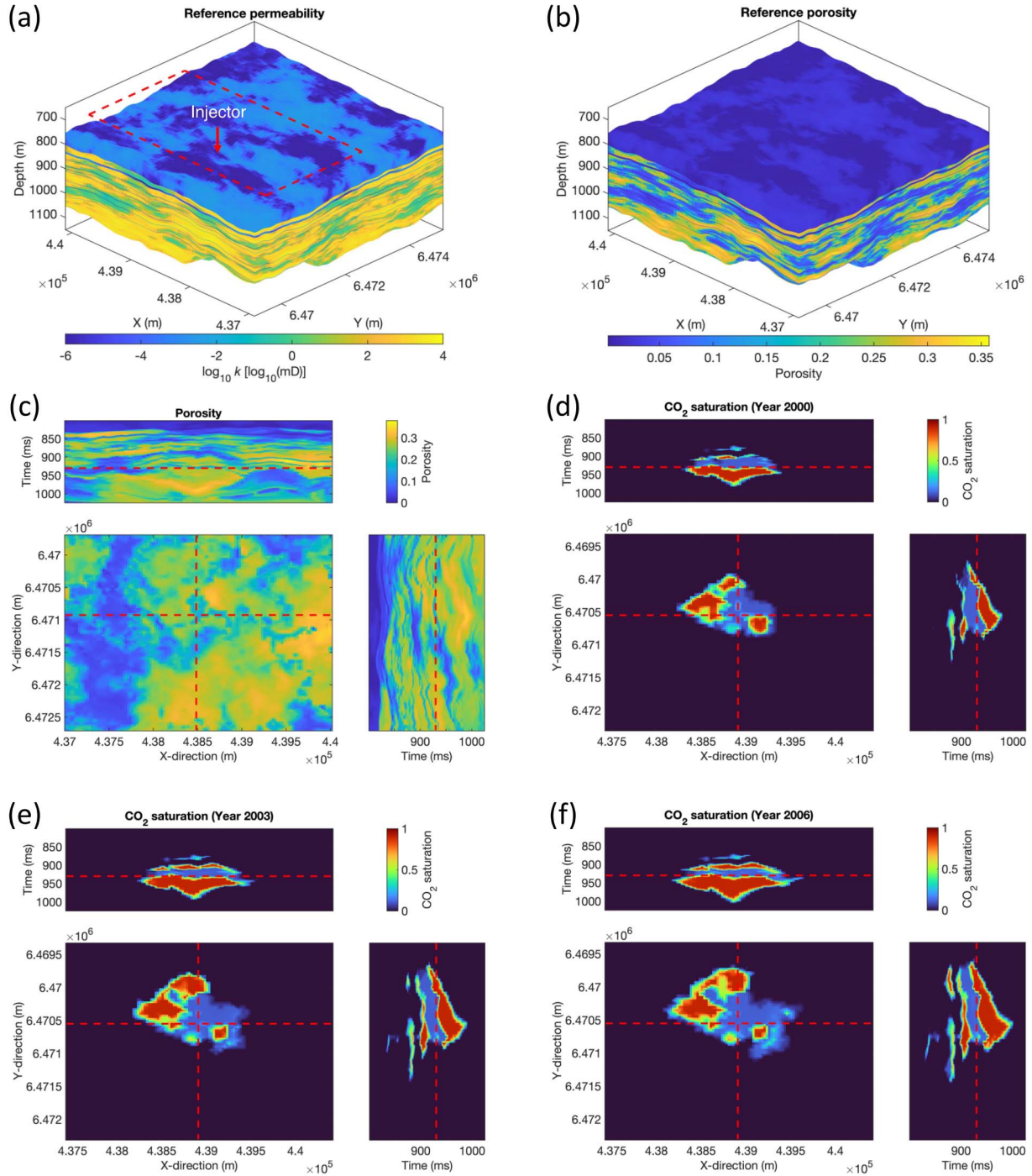
274 monitor seismic and resistivity surveys of the three years. The area of interest is a sub-region of

275 the reservoir model around the injector well (dashed red rectangle in Figure 4a). The synthetic

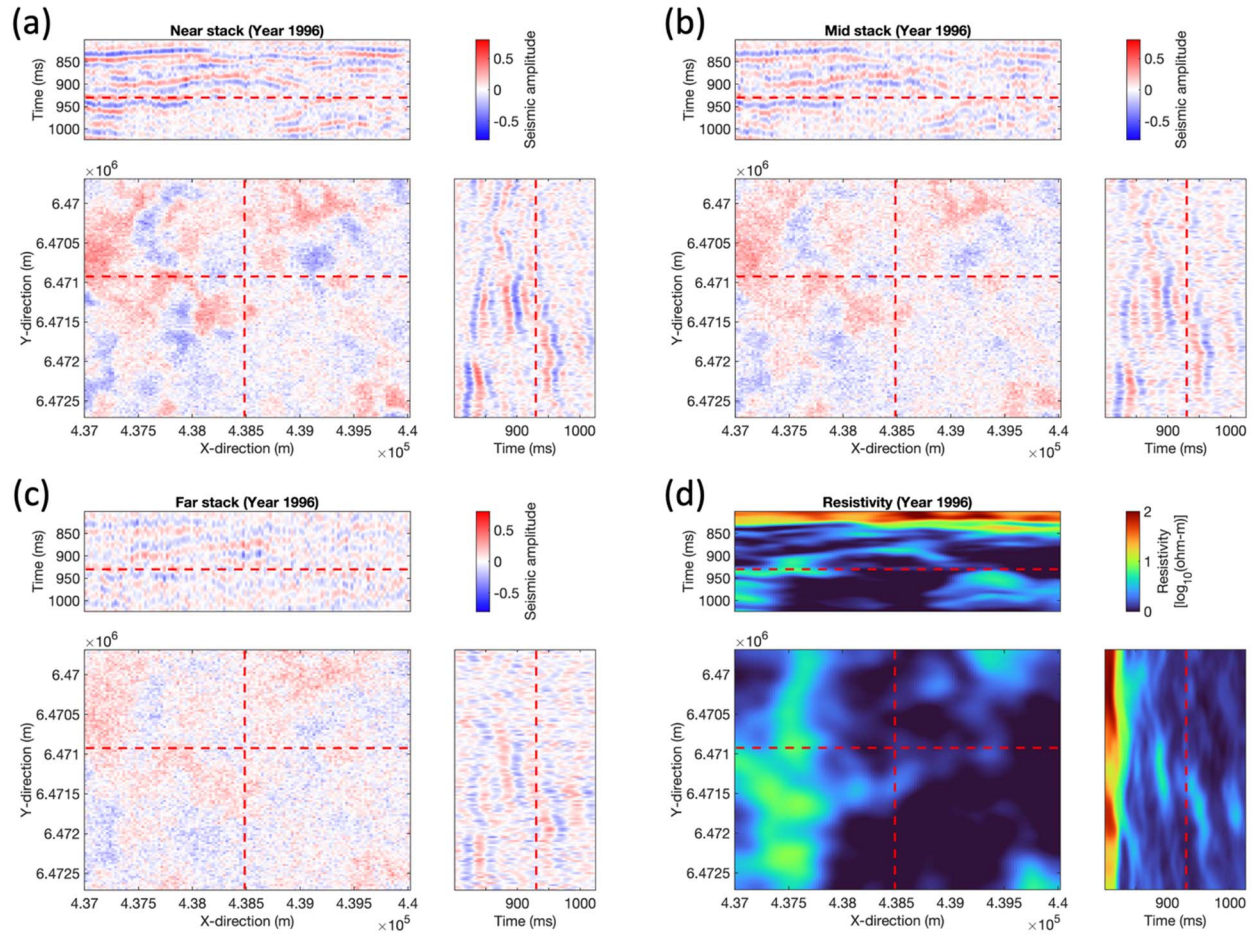
276 data in the area of interest are simulated in time domain and consist of 120×120×224 grid cells

277 with lateral resolution of 25 m (downscaled by scatter interpolation in 3-D) and time interval of 1

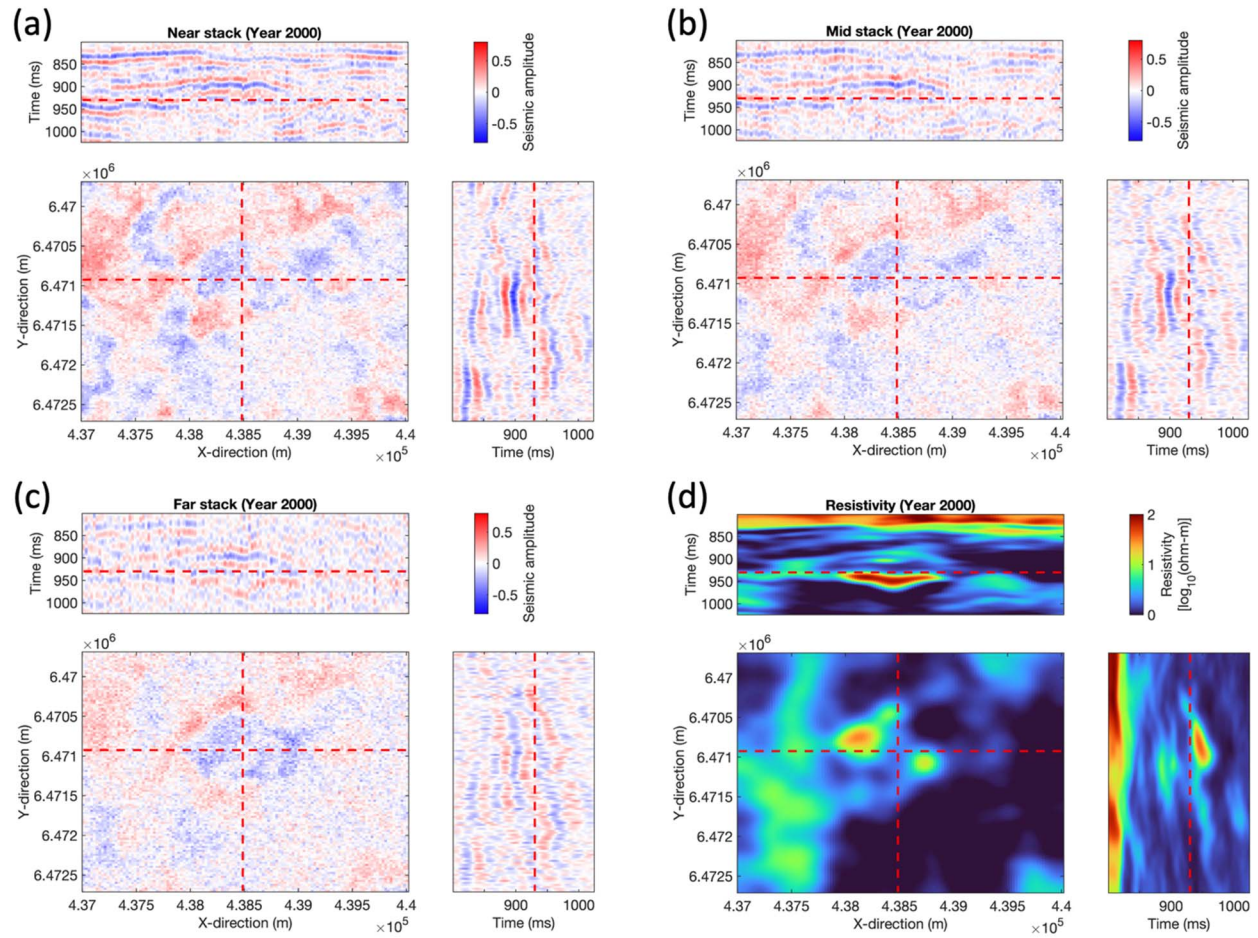
ms. Porosity and CO<sub>2</sub> saturations in the area of interest are shown in Figure 3c – f. The reference “true” seismic and resistivity responses predicted by the forward model (Section 2.1) at year 1996, 2000, 2003 and 2006 are shown in Figure 4 – 7. Observations of year 1996 are base surveys that are acquired before CO<sub>2</sub> injection, while observations of the latter three years are monitor surveys aiming to characterize CO<sub>2</sub> displacement. The pre-stack seismic data include near, mid and far stacks corresponding to incidence angles of 12, 24 and 36 degrees, respectively. The dominant frequencies of wavelets for near, mid and far stacks are 30, 25 and 20 Hz, respectively. We introduce measurement error to the seismic data by adding Gaussian noise with zero mean and 0.02 standard deviation to the reflection coefficients and then convolve them with wavelets. For the resistivity, we add Gaussian noise with zero mean and 0.5 standard deviation in the common logarithm scale and then smooth it with a Gaussian filter with standard deviation of 3 pixels. This represents for example, resistivity that would be available from an inversion of CSEM data. The implementation details and the parameters of the reference models and geophysical observations are included in the code.



**Figure 3.** Reservoir models of porosity (a) (dashed red rectangle is the area of interest) and permeability (b) for CO<sub>2</sub> flow simulation. Porosity (c) and CO<sub>2</sub> saturations at year 2000 (d), 2003 (e) and 2006 (f) in the area of interest. The injector well is located at the intersection of the dashed red lines.

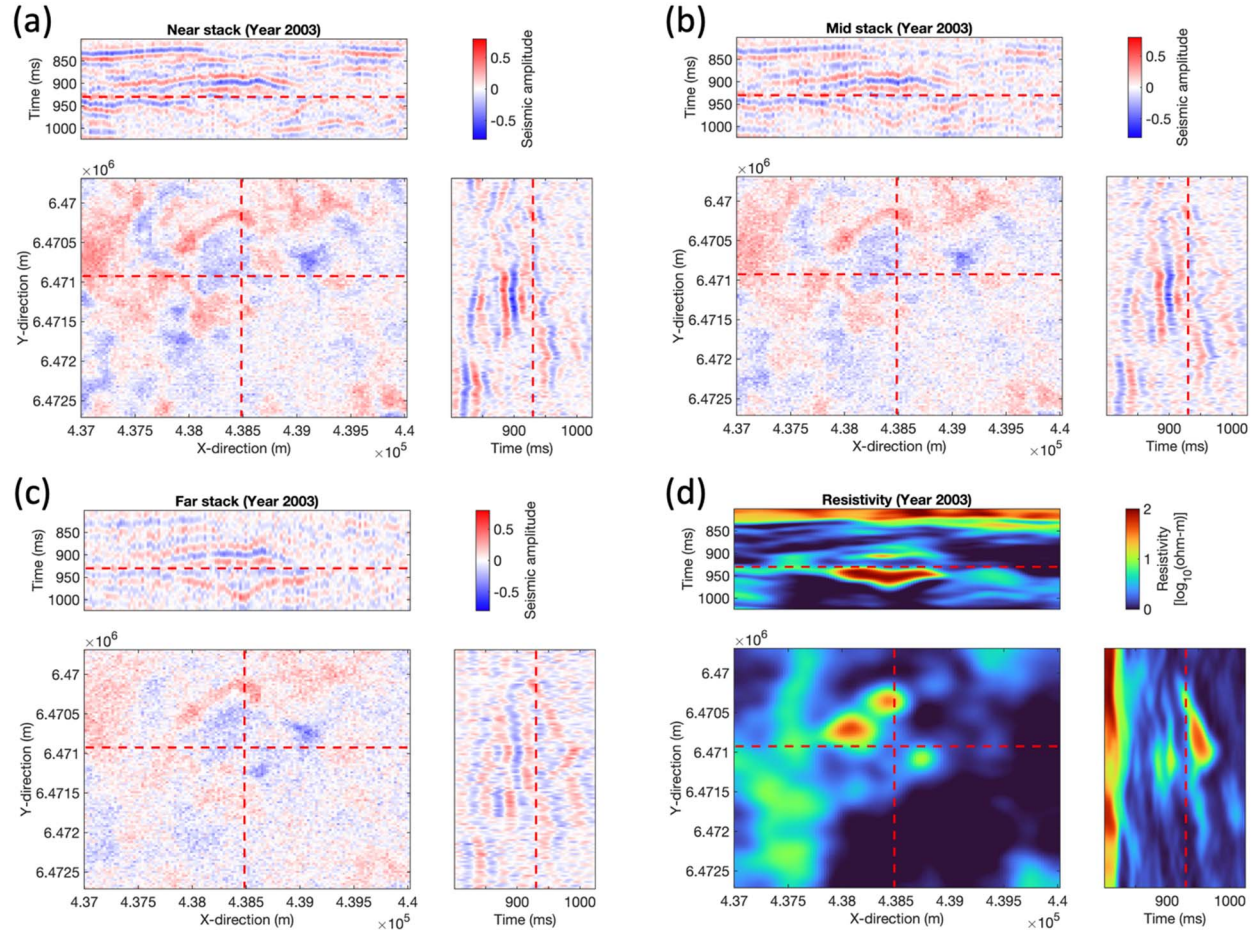


**Figure 4.** Observations of year 1996 (base survey): (a) near seismic stack; (b) mid seismic stack; (c) far seismic stack; (d) resistivity data.

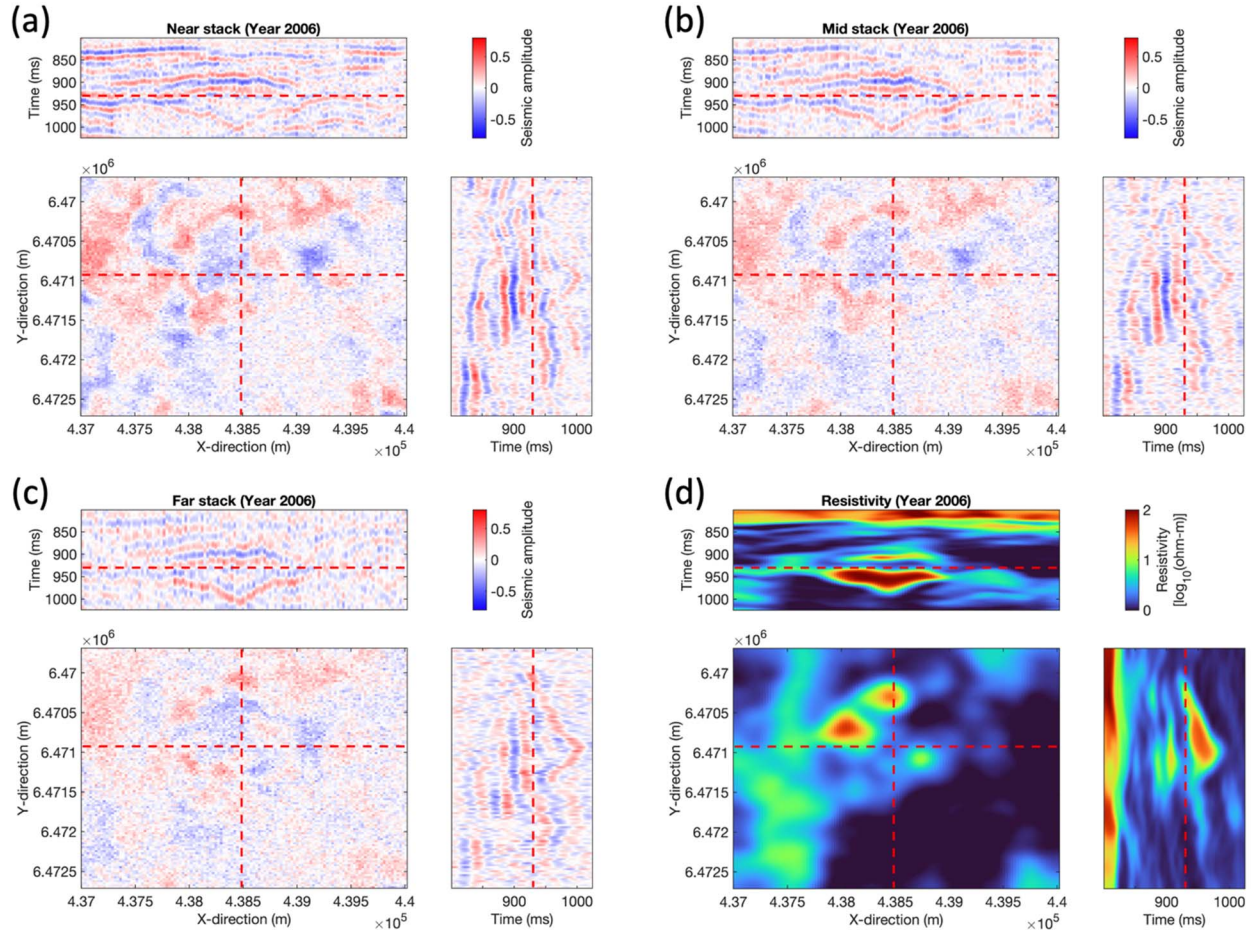


**Figure 5.** Observations of year 2000 (monitor survey): (a) near seismic stack; (b) mid seismic stack; (c) far seismic stack; (d) resistivity data.





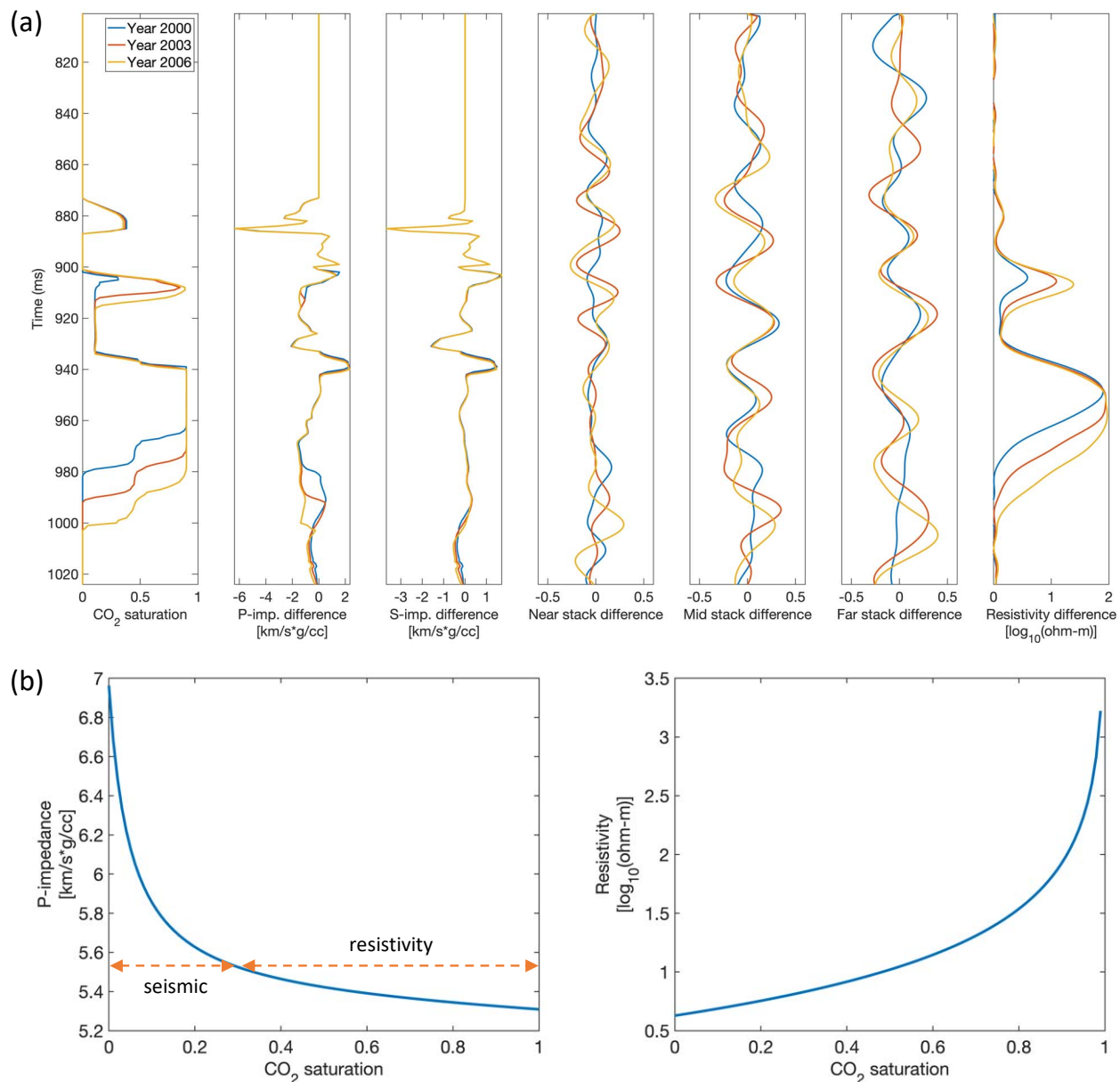
**Figure 6.** Observations of year 2003 (monitor survey): (a) near seismic stack; (b) mid seismic stack; (c) far seismic stack; (d) resistivity data.



**Figure 7.** Observations of year 2006 (monitor survey): (a) near seismic stack; (b) mid seismic stack; (c) far seismic stack; (d) resistivity data.

Figure 8 shows the seismic and resistivity responses as function of the change in  $\text{CO}_2$  saturation and the sensitivity of P-impedance and resistivity to  $\text{CO}_2$  saturation. In this case, it will be hard to detect  $\text{CO}_2$  saturation changes above 0.3 using time-lapse seismic data. Inversion of seismic data alone can predict the porosity model and roughly delineate the  $\text{CO}_2$  plumes, but the  $\text{CO}_2$  saturations are obviously underestimated. Instead, time-lapse resistivity data are highly sensitive to changes in fluid volumes but the spatial resolution of the resistivity measurements is limited and the coupling of porosity and the  $\text{CO}_2$  saturation in Equation 2 makes the inversion of

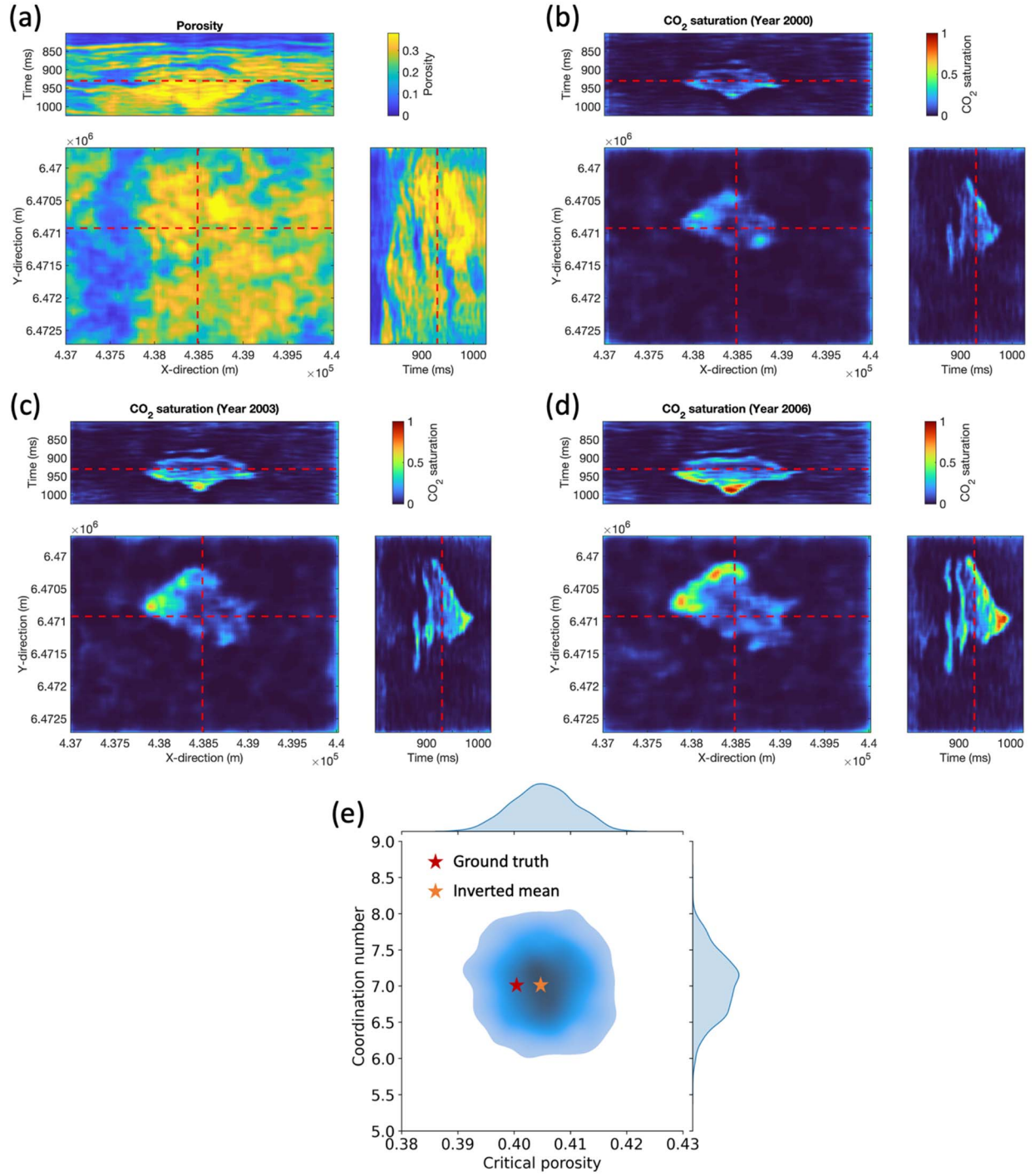
316 resistivity data by itself highly uncertain. The inverted results using only one type of data, i.e.,  
 317 seismic or resistivity data alone, are shown in Figure 9 and 10.



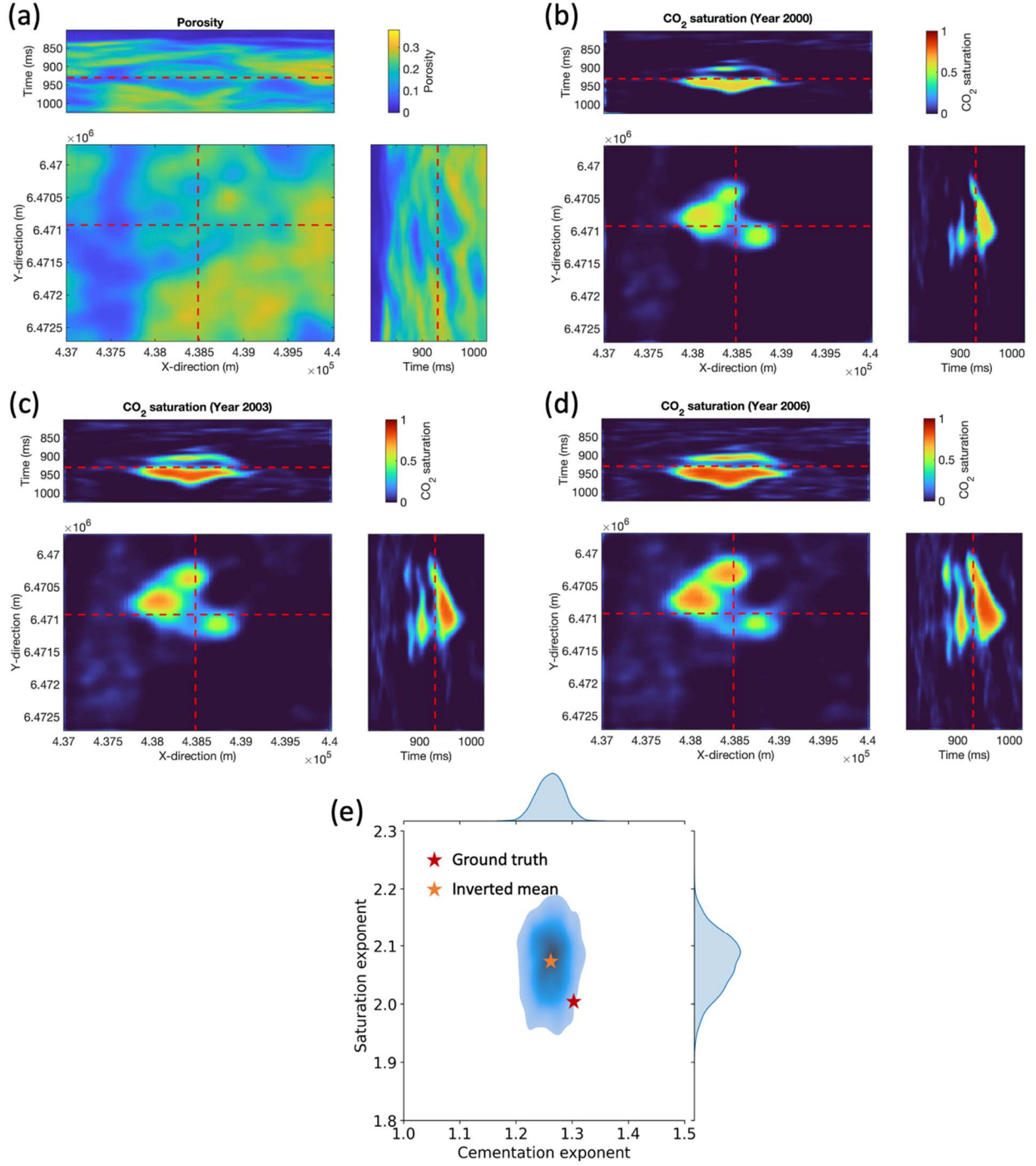
318  
 319 **Figure 8.** (a) Geophysical responses to the change in CO<sub>2</sub> saturation and (b) sensitivity of P-impedance  
 320 and resistivity to CO<sub>2</sub> saturation.

321



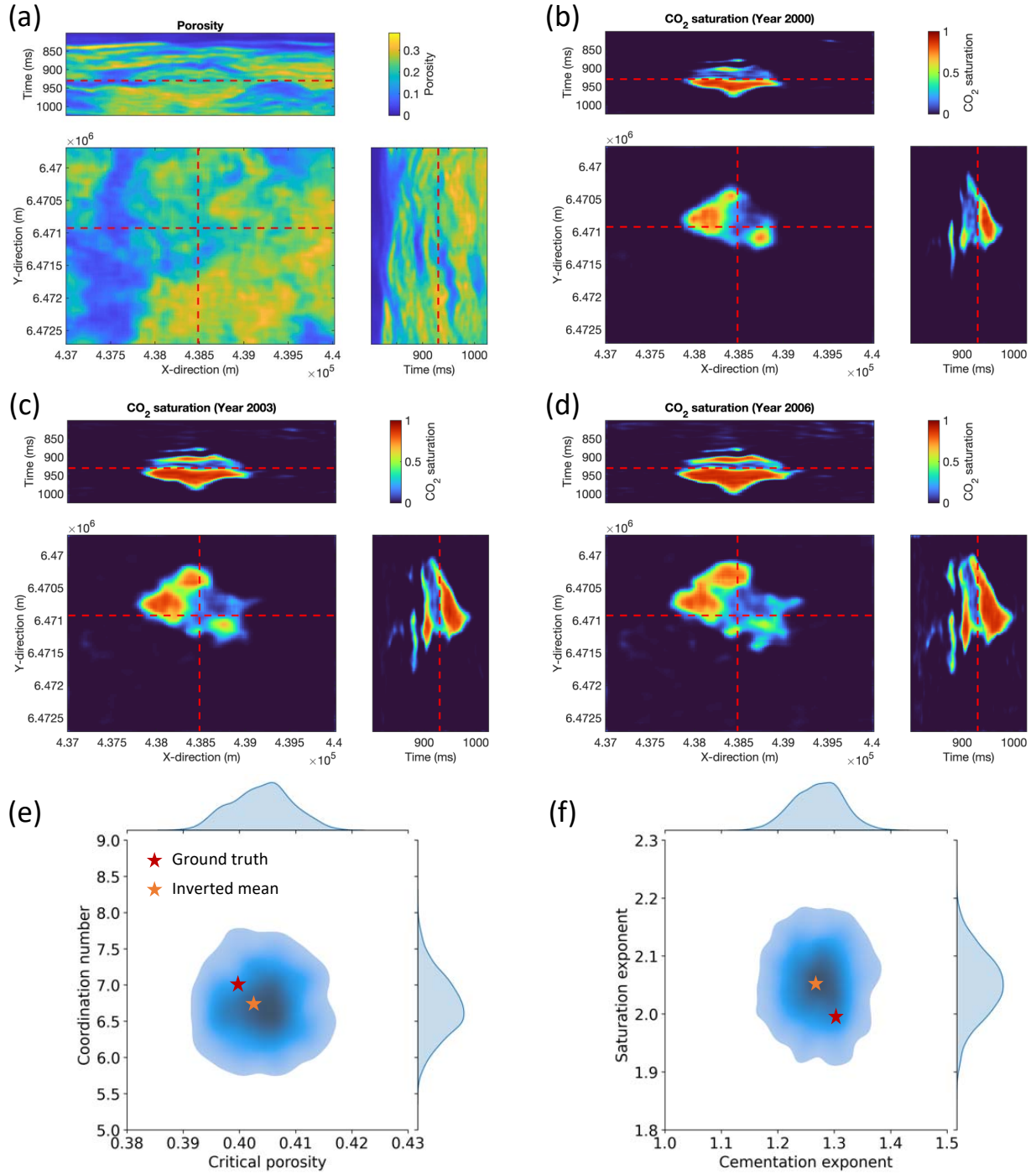


**Figure 9.** Inverted results with only seismic data: (a) porosity; (b) CO<sub>2</sub> saturation at year 2000; (c) CO<sub>2</sub> saturation at year 2003; (d) CO<sub>2</sub> saturation at year 2006; (e) critical porosity and coordination number obtained from the trained BNN.

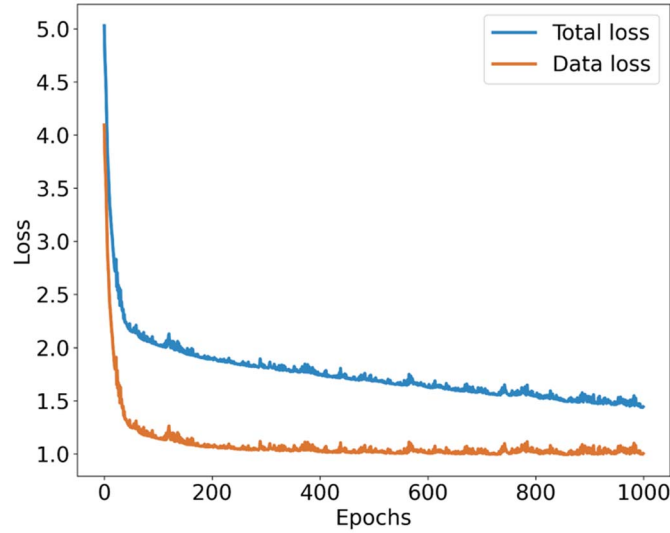


**Figure 10.** Inverted results with only resistivity data: (a) porosity; (b) CO<sub>2</sub> saturation at year 2000; (c) CO<sub>2</sub> saturation at year 2003; (d) CO<sub>2</sub> saturation at year 2006; (e) cementation and saturation exponents obtained from the trained BNN.

The proposed differentiable physics model provides an efficient way to jointly invert time-lapse seismic and resistivity data for the simultaneous prediction of porosity and CO<sub>2</sub> saturations. Figure 11a – d shows the inverted models of porosity and CO<sub>2</sub> saturations at years 2000, 2003 and 2006. Compared to the reference models in Figure 3, the predicted porosity and CO<sub>2</sub> saturations accurately match the true model. The  $R^2$  score of porosity is 0.92 and of CO<sub>2</sub> saturations are 0.90, 0.92 and 0.92 for the three years under examination. Figure 11e and 11f show the global model parameters, i.e., critical porosity, coordination number, cementation and saturation exponent, with associated uncertainties predicted by the BNN. The mean values of the predicted global parameters are close to the true parameters used to generate the synthetic data. Unlike traditional trace-by-trace inversion algorithms, the proposed model predicts geologically consistent and spatially continuous models, thanks to the simultaneous inversion in 4-D and the prior information encoded in the CNN for model re-parameterization (Zhu et al., 2022). The total loss defined by Equation 7 and the MSE between the predicted and true data over iterations are shown in Figure 12. The other advantage of the proposed method is the computational efficiency. The inversion takes approximately 2.2 hours with an Intel Xeon E5-1650 CPU and the computing time is further reduced to 15 minutes with a Nvidia GTX 1080 GPU. Overall, the proposed inversion model shows great promises for joint geophysical monitoring for CO<sub>2</sub> storage and could be applied to any large scale geophysical inverse problem with noisy data from multiple sources.



**Figure 11.** Joint inversion results of the proposed model: (a) porosity; (b) CO<sub>2</sub> saturation at year 2000; (c) CO<sub>2</sub> saturation at year 2003; (d) CO<sub>2</sub> saturation at year 2006; (e) posterior distribution of critical porosity and coordination number; and (f) cementation and saturation exponents, obtained from the trained BNN.



**Figure 12.** Loss function values over iterations.

#### 4. Discussion

The differentiable physics model provides a promising tool for joint inversion of large geophysical data for GCS monitoring. The synthetic example focuses on the inversion of pre-stack seismic and resistivity data based on the seismic wave propagation and elastic-electrical rock physics models and provides an accurate validation of the approach. The proposed model can be also extended to the inversion and assimilation of other data types and other geophysical operators, including, for example, seismic angle gathers with full waveform inversion or amplitude and phase measurements of electromagnetic fields with electromagnetic wave propagation models by implementing the forward operators according to the differentiable programming approach. According to the proposed model, the gradients of the forward operators can be efficiently evaluated by automatic differentiation and the computational effort can be deployed to high-performance computing platforms (e.g., GPUs). One potential limitation of the proposed model is the high demand for computing resources in cases with extremely large

geological models and geophysical data (up to billions of grids), especially when involving computationally intensive forward modeling. A possible solution might include the integration of PINNs and deep-learning surrogate models into the inversion model.

## 5. Conclusion

A novel inversion model has been developed for the joint inversion of large-scale geophysical data from different sources and applied to geologic carbon sequestration monitoring. The proposed model is built upon the technique of differentiable physics with the integration of deep neural networks for model re-parameterization. Thanks to automatic differentiation and parallel computing, the inversion is computationally efficient, and all parameters are updated simultaneously to minimize the loss function through the backpropagation algorithm. Moreover, the model can also predict global model parameters that are uncertain in the forward model. The presented approach has been validated on a synthetic example modified from the Sleipner benchmark model. The inverted models accurately reproduced the actual reference model, including porosity and time dependent CO<sub>2</sub> saturation. We conclude that the proposed model provides an efficient and accurate tool to assimilate multiphysics data for subsurface monitoring.

## Appendix A: Unconsolidated sand model

In the unconsolidated sand model (Dvorkin and Nur 1996), the matrix bulk and shear moduli ( $K_{\text{mat}}$  and  $\mu_{\text{mat}}$ ) are computed by the Voigt-Reuss-Hill average as function of the mineral fractions (e.g., clay and quartz volumes), while the moduli of dry rock ( $K_{\text{HM}}$  and  $\mu_{\text{HM}}$ ) at the critical porosity  $\phi_0$  are computed using the Hertz-Mindlin equations (Mindlin, 1949)

$$K_{\text{HM}} = \sqrt[3]{\frac{P_e[n(1-\phi_0)\mu_{\text{mat}}]^2}{18[\pi(1-\nu)]^2}}, \quad (\text{A-1})$$

$$\mu_{\text{HM}} = \frac{5-4\nu}{5(2-\nu)} \sqrt[3]{\frac{3P_e[n(1-\phi_0)\mu_{\text{mat}}]^2}{2[\pi(1-\nu)]^2}}, \quad (\text{A-2})$$

where  $P_e$  is the effective pressure,  $\nu$  is the grain Poisson's ratio and  $n$  is the coordination number (i.e., average number of contacts per grain). Then, the moduli of the dry rock ( $K_{\text{dry}}$  and  $\mu_{\text{dry}}$ ) with porosity in the range  $[0, \phi_0]$  are obtained by interpolating two end members (matrix and dry rock moduli at critical porosity) using the modified Hashin-Shtrikman lower bounds (Hashin and Shtrikman 1963)

$$K_{\text{dry}} = \left( \frac{\frac{\phi}{\phi_0}}{K_{\text{HM}} + \frac{4}{3}\mu_{\text{HM}}} + \frac{1-\frac{\phi}{\phi_0}}{K_{\text{mat}} + \frac{4}{3}\mu_{\text{HM}}} \right)^{-1} - \frac{4}{3}\mu_{\text{HM}}, \quad (\text{A-3})$$

$$\mu_{\text{dry}} = \left( \frac{\frac{\phi}{\phi_0}}{\mu_{\text{HM}} + \frac{1}{6}\xi\mu_{\text{HM}}} + \frac{1-\frac{\phi}{\phi_0}}{\mu_{\text{mat}} + \frac{1}{6}\xi\mu_{\text{HM}}} \right)^{-1} - \frac{1}{6}\xi\mu_{\text{HM}}, \quad (\text{A-4})$$

with

$$\xi = \frac{9K_{\text{HM}} + 8\mu_{\text{HM}}}{K_{\text{HM}} + 2\mu_{\text{HM}}}. \quad (\text{A-5})$$

For the fluid saturated rock, the bulk and shear moduli ( $K_{\text{sat}}$  and  $\mu_{\text{sat}}$ ) are computed using fluid substitution with Gassmann's equations (Gassmann 1951)

$$K_{\text{sat}} = K_{\text{dry}} + \frac{\left(1 - \frac{K_{\text{dry}}}{K_{\text{mat}}}\right)^2}{\frac{\phi}{K_{\text{fl}}} + \frac{1-\phi}{K_{\text{mat}}} - \frac{K_{\text{dry}}}{K_{\text{mat}}^2}}, \quad (\text{A-6})$$

$$\mu_{\text{sat}} = \mu_{\text{dry}}. \quad (\text{A-7})$$

where  $K_{\text{fl}}$  is the bulk modulus of the fluid phase and it depends on the water and  $\text{CO}_2$  saturations and their corresponding bulk moduli. Here the bulk moduli of brine water and  $\text{CO}_2$  are 3.06 and 0.10 GPa, respectively; the density of brine water and  $\text{CO}_2$  are 1.08 and 0.72 g/cc, respectively (Davis et al., 2019). Finally, the P- and S-wave velocities can be computed from the moduli by

413 definition

$$414 \quad V_p = \sqrt{\frac{K_{\text{sat}} + \frac{4}{3}\mu_{\text{sat}}}{\rho_{\text{sat}}}}, \quad (\text{A-8})$$

$$415 \quad V_s = \sqrt{\frac{\mu_{\text{sat}}}{\rho_{\text{sat}}}}, \quad (\text{A-9})$$

416 where  $\rho_{\text{sat}}$  is the density of the saturated rock which is weighted average of the matrix density

417  $\rho_{\text{mat}}$  and fluid density  $\rho_{\text{fl}}$

$$418 \quad \rho_{\text{sat}} = (1 - \phi)\rho_{\text{mat}} + \phi\rho_{\text{fl}}. \quad (\text{A-10})$$

419

## 420 **Appendix B: Convolution-based seismic model**

421 In the convolution-based seismic wave propagation model (Yilmaz, 2001; Grana et al., 2021),  
422 the seismic response  $\mathbf{d}_{\text{seis}}(x, t, \theta)$  is computed by convolving the source wavelet  $\mathbf{w}(\theta)$  with a  
423 series of reflectivity coefficients  $\mathbf{R}(x, t, \theta)$  as

$$424 \quad \mathbf{d}_{\text{seis}}(x, t, \theta) = \mathbf{w}(\theta) * \mathbf{R}(x, t, \theta), \quad (\text{B-1})$$

425 where  $x$  is the spatial location,  $t$  is the two-way vertical seismic travel time and  $\theta$  is the angle of  
426 incidence. For pre-stack seismic data, the reflectivity coefficients  $\mathbf{R}(x, t, \theta)$  are obtained from the  
427 elastic parameters (P-wave velocity  $V_p$ , S-wave velocity  $V_s$ , and density  $\rho$ ) by the Knott-  
428 Zoeppritz equations (Aki and Richards, 2002)



$$\begin{aligned}
& \begin{bmatrix} R_{PP}(\theta_1) \\ R_{PS}(\theta_1) \\ T_{PP}(\theta_1) \\ T_{PS}(\theta_1) \end{bmatrix} = \begin{bmatrix} -\sin \theta_1 & -\cos \varphi_1 & \sin \theta_2 & \cos \varphi_2 \\ \cos \theta_1 & -\sin \varphi_1 & \cos \theta_2 & -\sin \varphi_2 \\ \sin 2\theta_1 & \frac{V_{P1}}{V_{S1}} \cos 2\varphi_1 & \frac{\rho_2 V_{S2}^2 V_{P1}}{\rho_1 V_{S1}^2 V_{P2}} \sin 2\theta_2 & \frac{\rho_2 V_{S2} V_{P1}}{\rho_1 V_{S1}^2} \cos 2\varphi_2 \\ -\cos 2\theta_1 & \frac{V_{S1}}{V_{P1}} \sin 2\varphi_1 & \frac{\rho_2 V_{P2}}{\rho_1 V_{P1}} \cos 2\theta_2 & -\frac{\rho_2 V_{S2}}{\rho_1 V_{P1}} \sin 2\varphi_2 \end{bmatrix}^{-1} \begin{bmatrix} \sin \theta_1 \\ \cos \theta_1 \\ \sin 2\theta_1 \\ \cos 2\theta_1 \end{bmatrix}, \quad (\text{B-2})
\end{aligned}$$

where  $\theta_1$  is the angle of incidence,  $\theta_2$  is the angle of the transmitted P-wave,  $\varphi_1$  is the angle of reflected S-wave, and  $\varphi_2$  is the angle of the transmitted S-wave;  $R_{PP}(\theta_1)$ ,  $R_{PS}(\theta_1)$ ,  $T_{PP}(\theta_1)$  and  $T_{PS}(\theta_1)$  are the reflected P-, reflected S-, transmitted P-, and transmitted S-wave amplitude coefficients, respectively; the subscripts 1 and 2 represent elastic parameters of the upper and lower layers, respectively. In this study, we only consider reflected P-wave amplitudes (i.e.,  $R_{PP}$ ).

## Data Availability Statement

The code and data are freely available on the GitHub repository (<https://github.com/theanswer003/JointInvDP>) (Liu, 2022).

## Acknowledgements

We acknowledge the sponsors of the Stanford Center for Earth Resources Forecasting (SCERF) and support from Prof. Steve Graham, the Dean of the Stanford School of Earth, Energy and Environmental Sciences. The authors also would like to thank Equinor for sharing the Sleipner benchmark model.

447   **References**

- 448   • Aki, K., & Richards, P. G. (2002). Quantitative Seismology: *University Science Books, 2nd*  
449       *ed.*
- 450   • Aminu, M.D., Nabavi, S.A., Rochelle, C.A., & Manovic, V. (2017). A review of  
451       developments in carbon dioxide storage. *Applied Energy, 208*, pp.1389-1419.
- 452   • Archie, G.E. (1942). The electrical resistivity log as an aid in determining some reservoir  
453       characteristics. *Transactions of the AIME, 146*(01), pp.54-62.
- 454   • Bhuyian, A.H., Landrø, M., & Johansen, S.E. (2012). 3D CSEM modeling and time-lapse  
455       sensitivity analysis for subsurface CO<sub>2</sub> storage. *Geophysics, 77*(5), pp.E343-E355.
- 456   • Bonneville, A., Black, A.J., Hare, J.L., Kelley, M.E., Place, M., & Gupta, N. (2021). Time-  
457       lapse borehole gravity imaging of CO<sub>2</sub> injection and withdrawal in a closed carbonate reef.  
458       *Geophysics, 86*(6), pp.G113-G132.
- 459   • Caspari, E., Müller, T.M., & Gurevich, B. (2011). Time-lapse sonic logs reveal patchy CO<sub>2</sub>  
460       saturation in-situ. *Geophysical Research Letters, 38*(13).
- 461   • Chadwick, A., Williams, G., Delepine, N., Clochard, V., Labat, K., Sturton, S., Buddensiek,  
462       M.L., Dillen, M., Nickel, M., Lima, A.L., & Arts, R. (2010). Quantitative analysis of time-  
463       lapse seismic monitoring data at the Sleipner CO<sub>2</sub> storage operation. *The Leading Edge,*  
464       *29*(2), pp.170-177.
- 465   • Davis, T.L., Landrø, M., & Wilson, M. (2019). Geophysics and geosequestration. *Cambridge*  
466       *University Press.*
- 467   • Deutsch, C.V., & Journal, A.G. (1998). GSLIB: Geostatistical software library and user's  
468       guide. *New York: Oxford University Press.*
- 469   • Dupuy, B., Romdhane, A., Nordmann, P.L., Eliasson, P., & Park, J. (2021). Bayesian rock-

physics inversion: Application to CO<sub>2</sub> storage monitoring. *Geophysics*, 86(4), pp.M101-M122.

- Dürr, O., Sick, B., & Murina, E. (2020). Probabilistic Deep Learning: With Python, Keras and TensorFlow Probability. *Manning Publications*.
- Dvorkin, J., & Nur, A. (1996). Elasticity of high-porosity sandstones: Theory for two North Sea data sets. *Geophysics*, 61(5), pp.1363-1370.
- Egorov, A., Pevzner, R., Bóna, A., Glubokovskikh, S., Puzyrev, V., Tertyshnikov, K., & Gurevich, B. (2017). Time-lapse full waveform inversion of vertical seismic profile data: Workflow and application to the CO<sub>2</sub>CRC Otway project. *Geophysical Research Letters*, 44(14), pp.7211-7218.
- Eliasson, P., Romdhane, A., Jordan, M., & Querendez, E. (2014). A synthetic Sleipner study of CO<sub>2</sub> quantification using controlled source electro-magnetics and full waveform inversion. *Energy Procedia*, 63, pp.4249-4263.
- Gasperikova, E., Appriou, D., Bonneville, A., Feng, Z., Huang, L., Gao, K., Yang, X., & Daley, T. (2022). Sensitivity of geophysical techniques for monitoring secondary CO<sub>2</sub> storage plumes. *International Journal of Greenhouse Gas Control*, 114, p.103585.
- Gassmann, F. (1951). Elastic waves through a packing of spheres. *Geophysics* 16(4), 673–685.
- Glubokovskikh, S., Pevzner, R., Gunning, J., Dance, T., Shulakova, V., Popik, D., Popik, S., Bagheri, M., & Gurevich, B. (2020). How well can time-lapse seismic characterize a small CO<sub>2</sub> leakage into a saline aquifer: CO<sub>2</sub>CRC Otway 2C experiment (Victoria, Australia). *International Journal of Greenhouse Gas Control*, 92, p.102854.
- Grana, D., Mukerji, T., & Doyen, P. (2021). Seismic reservoir modeling: Theory, examples,

and algorithms. *John Wiley & Sons*.

- Grude, S., Landrø, M., & Osdal, B. (2013). Time-lapse pressure–saturation discrimination for CO<sub>2</sub> storage at the Snøhvit field. *International Journal of Greenhouse Gas Control*, 19, pp.369-378.
- Hernández, A., Millerioux, G., & Amigó, J.M. (2022). Differentiable programming: Generalization, characterization and limitations of deep learning. *arXiv preprint arXiv:2205.06898*.
- Hashin, Z., & Shtrikman, S. (1963). A variational approach to the theory of the elastic behaviour of multiphase materials. *Journal of the Mechanics and Physics of Solids*, 11(2), 127–140.
- Huang, L. (2022). Geophysical Monitoring for Geologic Carbon Storage. *Wiley*.
- Kingma, D.P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kullback, S., & Leibler, R.A. (1951). On information and sufficiency. *Ann Math Stat* 22:79–86.
- Li, B., & Li, Y.E. (2021). Neural Network-Based CO<sub>2</sub> Interpretation From 4D Sleipner Seismic Images. *Journal of Geophysical Research: Solid Earth*, 126(12), p.e2021JB022524.
- Li, D., Xu, K., Harris, J.M., & Darve, E. (2020). Coupled time-lapse full-waveform inversion for subsurface flow problems using intrusive automatic differentiation. *Water Resources Research*, 56(8), p.e2019WR027032.
- Liu, M., & Grana, D. (2019). Accelerating geostatistical seismic inversion using TensorFlow: A heterogeneous distributed deep learning framework. *Computers & Geosciences*, 124, pp.37-45.

- 516 • Liu, M. (2022). Joint inversion of geophysical data for geologic carbon sequestration  
517 monitoring: a differentiable physics-informed deep learning model. [Source code].  
518 <https://github.com/theanswer003/JointInvDP>
- 519 • Martin, D., & Estrin, G. (1967). Models of computations and systems—evaluation of vertex  
520 probabilities in graph models of computations. *Journal of the ACM (JACM)*, 14(2), pp.281-  
521 299.
- 522 • Metz, B., Davidson, O., & De Coninck, H. eds. (2005). Carbon dioxide capture and storage:  
523 special report of the intergovernmental panel on climate change. *Cambridge University Press*.
- 524 • Mindlin, R.D. (1949). Compliance of elastic bodies in contact. *J. Appl. Mech., ASME* 16,  
525 259–268.
- 526 • Raissi, M., Perdikaris, P., & Karniadakis, G.E. (2019). Physics-informed neural networks: A  
527 deep learning framework for solving forward and inverse problems involving nonlinear  
528 partial differential equations. *Journal of Computational physics*, 378, pp.686-707.
- 529 • Rall, L.B. (1981). Automatic differentiation: Techniques and applications. *Berlin, Heidelberg:*  
530 *Springer Berlin Heidelberg*.
- 531 • Rasht-Behesht, M., Huber, C., Shukla, K., & Karniadakis, G.E. (2022). Physics-Informed  
532 Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions. *Journal of*  
533 *Geophysical Research: Solid Earth*, 127(5), p.e2021JB023120.
- 534 • Rasmussen, A.F., Sandve, T.H., Bao, K., Lauser, A., Hove, J., Skaflestad, B., Klöfkorn, R.,  
535 Blatt, M., Rustad, A.B., Sævareid, O., & Lie, K.A. (2021). The open porous media flow  
536 reservoir simulator. *Computers & Mathematics with Applications*, 81, pp.159-185.
- 537 • Rubino, J.G., Velis, D.R., & Sacchi, M.D. (2011). Numerical analysis of wave-induced fluid  
538 flow effects on seismic data: Application to monitoring of CO<sub>2</sub> storage at the Sleipner field.

- 539      *Journal of Geophysical Research: Solid Earth*, 116(B3).
- 540      • Tang, H., Fu, P., Jo, H., Jiang, S., Sherman, C.S., Hamon, F., Azzolina, N.A., & Morris, J.P.  
541      (2022). Deep Learning-Accelerated 3D Carbon Storage Reservoir Pressure Forecasting  
542      Based on Data Assimilation Using Surface Displacement from InSAR. *arXiv preprint*  
543      arXiv:2201.08543.
- 544      • Tveit, S., Mannseth, T., Park, J., Sauvin, G., & Agersborg, R. (2020). Combining CSEM or  
545      gravity inversion with seismic AVO inversion, with application to monitoring of large-scale  
546      CO<sub>2</sub> injection. *Computational Geosciences*, 24(3), pp.1201-1220.
- 547      • Vo, H.X., & Durlofsky, L.J. (2014). A new differentiable parameterization based on principal  
548      component analysis for the low-dimensional representation of complex geological models.  
549      *Mathematical Geosciences*, 46(7), pp.775-813.
- 550      • Yilmaz, Ö. (2001). Seismic data analysis (Vol. 1, p. 1809). *Tulsa: Society of exploration*  
551      *geophysicists*.
- 552      • Zhu, W., Xu, K., Darve, E., Biondi, B., & Beroza, G.C. (2022). Integrating deep neural  
553      networks with full-waveform inversion: Reparameterization, regularization, and uncertainty  
554      quantification. *Geophysics*, 87(1), pp.R93-R109.