Title

Deep learning multiphysics network for imaging CO2 saturation and estimating uncertainty in geological carbon storage

Journal

Geophysical Prospecting, 72(1)

ISSN

0016-8025

Authors

Um, Evan Schankee
Alumbaugh, David
Commer, Michael
et al.

Publication Date

2024

DOI

10.1111/1365-2478.13257

Peer reviewed

Title: Deep-Learning Multiphysics Network for Imaging CO_2 Saturation and Estimating Uncertainty in Geological Carbon Storage

Short Running Title: DL Multiphysics Network for Imaging CO_2 Saturation

1st revision: June 1, 2022

2nd revision: July, 15, 2022

3rd revision: July, 26, 2022

List of Authors:

Evan Schankee Um, Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 94720, esum@lbl.gov

David Alumbaugh, Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 94720, dlalumbaugh@lbl.gov

Michael Commer†, Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 94720, mcommer@lbl.gov

Shihang Feng†, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA, 87545, shihang@lanl.gov

Erika Gasperikova†, Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, 94720, egasperikova@lbl.gov

Yaoguo Li†, Department of Geophysics, Colorado School of Mines, Golden, CO, USA, 80401, ygli@mines.edu

Youzuo Lin†, Earth and Environment Division, Los Alamos National Laboratory, Los Alamos, NM, USA, 87545, ylin@lanl.gov

Savini Samarasinghe†, Department of Geophysics, Colorado School of Mines, Golden, CO, USA, 80401, savini.ms@gmail.com

†These authors have contributed equally to this work.

Conflicts of Interest:

The authors declare no conflicts of interest in preparing this manuscript.

Abstract

Multiphysics inversion exploits different types of geophysical data that often complement each other and aims to improve overall imaging resolution and reduce uncertainties in geophysical interpretation. Despite the advantages, traditional multiphysics inversion is challenging because it requires a large amount of computational time and intensive human interactions for preprocessing data and finding trade-off parameters. These issues make it nearly impossible for traditional multiphysics inversion to be applied as a real-time monitoring tool for geological carbon storage. In this paper, we present a deep-learning (DL) multiphysics network for imaging CO_2 saturation in real time. The multiphysics network consists of three encoders for analyzing seismic, electromagnetic, and gravity data, and shares one decoder for combining imaging capabilities of the different geophysical data for better predicting CO_2 saturation. The network is trained on pairs of CO_2 label models and multiphysics data so that it can directly image CO_2 saturation. We use the bootstrap aggregating method to enhance the imaging accuracy and estimate uncertainties associated with CO_2 saturation images. Using realistic CO_2 label models and multiphysics data derived from the Kimberlina CO_2 storage model, we evaluate the performance of the DL multiphysics network and compare their imaging results to those from the DL single-physics networks. Our modeling experiments show that the DL multiphysics network for seismic, electromagnetic, and gravity data not only improves the imaging accuracy but also reduces uncertainties associated with CO_2 saturation images. Our results also suggest that the DL multiphysics network for the non-seismic data (i.e., electromagnetic and gravity) can be used as an effective low-cost monitoring tool in between regular seismic monitoring.

Keywords: Full waveform, Electromagnetics, Gravity, Inversion and Monitoring

Introduction

Geological carbon storage (GCS) is a viable option for reducing CO_2 emission into the atmosphere (Metz et al., 2005; Benson and Cole, 2008; Davis et al., 2019; Ringrose, 2020). A large amount of CO_2 is captured from fossil-fuel power stations and other major industrial CO_2 sources and is injected into depleted reservoirs or saline aquifers. The injection of CO_2 into a reservoir changes geomechanical, geochemical, hydrological states inside and around the reservoir and can threaten the seal integrity of GCS (e.g., Rutqvist, 2012; Zoback and Gorelick, 2012; Jenkins et al., 2015; Harbert et al., 2016). For safe and efficient GCS operations, it is important to accurately track the movement of CO_2 plumes inside and detect CO_2 leak from a reservoir in real time.

Geophysical methods provide the possibility of cost-effective long-term monitoring for GCS. Various types of geophysical data are often acquired and interpreted together as they are sensitive to different geophysical properties, and the methods also exhibit different scales of resolving power. For example, the seismic method can provide high-resolution subsurface images and is highly sensitive to changes in CO_2 saturations when the saturation is relatively low (Vasco et al., 2014). Thus, seismic imaging serves as a great tool for delineating the boundaries of CO_2 plumes in detail (e.g., Lazaratos and Marion, 1997; Arts et al., 2003; Chadwick et al., 2010; Ajo-Franklin et al., 2013; Queißer and Singh, 2013; Li et al., 2021). In contrast, electromagnetic measurements are sensitive to changes in saturation at relatively higher concentrations of CO_2 (Gasperikova and Hoversten, 2006). Therefore, the electromagnetic method is better suited to characterizing higher concentration portions of CO_2 plumes as well as recovering higher saturation values. Gravity data are sensitive to the full range of CO_2 saturation and can also be used for estimating the overall distribution of density changes caused by CO_2 injection (e.g., Eiken et al., 2008; Alnes et al., 2011; Gasperikova and Li, 2021; Yang et al., 2022). By using several geophysical data types together, a multiphysics inversion approach aims to improve the overall resolution of GCS imaging and decreases uncertainties in geophysical interpretation.

Traditional multiphysics inversion is not trivial and tends to be quite challenging from multiple standpoints. First and foremost, multiphysics inversion needs a significant amount of computer resources and time to repeatedly complete the forward modelling that is required for each of the geophysical data types (Commer and Newman, 2008; Shin and Cha, 2009; Virieux and Operto, 2009; Fichtner, 2010; Um et al., 2014). Multiphysics inversion can also be highly non-linear, and thus its success often relies on multiple trial and error as the nonlinearity can lead the inversion algorithm to get 'stuck' at local minima. Finding proper data weighting and trade-off parameters inside a multiphysics objective function also heavily relies on intensive human interactions. Even without these aspects of computational complexity and non-linearity related difficulties, the reservoir properties we are interested in for monitoring in a GCS project (e.g., CO_2 saturation) are empirically rather than theoretically related to the geophysical properties that the measurements are sensing. Thus, the conversion from one to the other tends to pose additional levels of uncertainty to the problem. Because of these issues, traditional multiphysics inversion is nearly impossible to be applied as a real-time monitoring tool for GCS.

Recently, deep learning (DL) imaging has drawn attention in computational geophysics as it overcomes some of the main drawbacks that traditional inversion exhibits (Araya-Polo et al., 2018; Yang and Ma, 2019; Wu and Lin, 2019; Zhang and Alkhalifah, 2019; Puzyrev, 2019; Colombo et al., 2020; Zhang and Lin, 2020, Kaur et al., 2021; Li and Yang, 2021; Um et al., 2022; Yang et al., 2022). A deep neural network is trained such that it can learn complex non-linear correlations between earth models and corresponding geophysical data. Therefore, once fully trained, the network can instantaneously predict an earth model from newly acquired geophysical monitoring data. The prediction can be completed in real or near-real time.

Nonetheless, using DL imaging does not remove the computational challenges of the multiphysics CO_2 monitoring problem. For example, in order to generate a realistic set of DL training models and data, one needs to simulate a number of CO_2 flow models and their geophysical counterparts by solving their governing partial differential equations (Zeng et al., 2021; Um et al., 2022). The flow simulation cost may be reduced by simply inserting many different shapes of CO_2 bodies into a background model (Puzyrev, 2019; Yang et al., 2022). In either case, a large number of geophysical forward modeling tasks should be completed. However, the flow models and associated synthetic geophysical data required for training the neural network can be generated prior to the geophysical monitoring data being acquired. This opens a possibility that DL multiphysics imaging can monitor GCS processes in near-real time with little human interaction and bias.

Based on the successful numerical modeling studies of DL imaging on single geophysical data as listed above, it is natural to extend DL from single to multiphysics data. In recent years, DL multiphysics imaging has been applied to onshore and offshore geophysics problems. For example, Oh et al. (2020) demonstrate a cooperative DL imaging network for marine controlled-source electromagnetic data and seismic information (i.e., seismic salt-top boundaries) for enhancing salt delineation. Sun et al. (2020) proposes a set of deep neural network architectures for marine seismic and electromagnetic data for salt reconstruction. Hu et al. (2021) present a DL enhanced joint imaging framework for crosswell seismic and DC resistivity data. Guo et al. (2020) use deep residual convolutional neural networks to assist multiphysics inversion of seismic and magnetotelluric data and demonstrate that the DL-assisted multiphysics inversion can predict an earth model with lower data misfit than single-physics inversions.

In this paper, we present DL multiphysics networks for imaging CO_2 plumes and estimating image uncertainty. The network architectures are designed to exploit seismic, electromagnetic, and gravity data separately as well as together. To directly image CO_2 saturation rather than geophysical proxy properties (e.g., P wave velocity, electrical resistivity, or density), the network is trained on pairs of CO_2 saturation label models and associated geophysical monitoring data (Um et al., 2022). The CO_2 saturation label models are generated by simulating CO_2 injection and flow over a 3D CO_2 storage model that was constructed based on real geologic and hydrogeologic data. The training multiphysics data are generated by solving various governing partial differential equations with geophysical earth models derived from rock-physics conversion of the CO_2 flow models. We utilize the imaging networks in an ensemble learning framework for further improving overall CO_2 imaging accuracy and estimating uncertainties associated with CO_2 images. In order to evaluate the performance of the DL multiphysics networks for CO_2 monitoring, we perform both DL single-physics and multiphysics imaging and systematically compare imaging results and uncertainty estimates.

Deep-Learning Multiphysics Imaging Network

Figure 1 shows a DL multiphysics imaging network architecture that we have assembled for estimating CO_2 saturation. This neural network can be considered a modified version of the U-Net (Ronneberger et al., 2015) that was originally developed for medical imaging segmentation. While the original U-Net has one encoder and one decoder, our multiphysics network consists of three encoders and one decoder. Each encoder takes one type of geophysical data and repeatedly applies convolution operation, batch normalization, rectified linear activation and max-pooling operations. The convolution operation has multiple convolution filters and produces feature maps that highlight regions of the input layer that resembles each filter. Using the convolution operations, the network analyzes spatial hierarchies of data. The batch normalization subtracts the mean from the data and divides it by its standard deviation. This operation helps the gradient propagate effectively through the deep neural network. We use rectified linear activation (ReLU), which allows a better gradient propagation. Dropout layers randomly set some input values to zero, preventing the network from overfitting the training data. The max pooling operation down-samples the feature maps by taking the maximum input value in each kernel and reduces large feature maps to smaller summary maps. Note that we do not implement any low-pass filter before the max pooling operation because a network like the U-Net using multiple convolution layers is known to learn an anti-aliasing filter (Ribeiro and Schon, 2021). Each decoder repeats the series of the operations described above four times. In order to constrain an output CO_2 saturation between 0 and 1, a sigmoid activation function is used in the final layer.

The three encoders share one decoder such that the network can combine imaging capabilities of three different types of geophysical data for better predicting CO_2 saturation. To share the common decoder, the outputs from the three different encoders are concatenated before they are fed into the next convolution layer. Similar to the encoders, the decoder consists of four repetitions. Inverse convolution is followed by the convolution, the batch normalization and the activation operation. To improve information flow through the deep neural network architecture, the decoder concatenates its feature maps with those from the encoders. For details on the operations mentioned above, the reader is referred to Goodfellow et al. (2016), Chollet (2017), and James et al. (2021).

As a default, the DL multiphysics network is designed to support three different types of input data (i.e., seismic, electromagnetic and gravity data) but can easily incorporate different available types of data. For example, when only one type of geophysical data is available, the encoder for that type of data is built but the other two encoders are not. In this case (**Figure** 2), the DL network reduces to the classic U-Net architecture (Ronneberger et al., 2015) and can be used as a DL single-physics network for imaging CO_2 saturation (Um et al., 2022). The details about their implementation will be described later.

To improve the overall accuracy of CO_2 images, we use the DL multiphysics network in an ensemble learning framework called the bootstrap aggregating or bagging method (Friedman et al., 2006; James et al., 2021). The bagging method uses bootstrapping as a sampling method and create many unique 'sub-data sets' out of the original multiphysics data sets. Then, the method independently trains a number of DL networks using the bootstrapped training data sets. Although the bagging method requires large computational cost for training many networks, all training tasks are independent from each other. Thus, multiple networks can be trained simultaneously on a modern graphic processing unit (GPU) cluster. For imaging CO_2, all networks make a prediction. The bagging method calculates an average value and standard deviation of CO_2 saturation within each cell in the imaging domain. By using the average of all predictions, the bagging method can yield more accurate estimates than a single strong predictor, reduce variance in noisy data and mitigate the chance of overfitting. The standard deviation also provides a measure of the uncertainty associated with the DL multiphysics network. In this work, we use both DL single- and multi-physics imaging networks and systematically compare their results in terms of accuracy and uncertainty.

CO_2 Models and Multiphysics Data

Traditional inversion repeatedly solves geophysical governing equations (e.g., seismic wave equation, electromagnetic diffusion equations and gravity potential equation) and updates a geophysical earth model such that the differences between measured and predicted geophysical data can be reduced during the inversion. Accordingly, the governing physics is directly embedded into the inversion. In contrast, the DL imaging networks (**Figures** 1 and 2) do not involve solving the geophysical governing equations. To embed the governing physics into the DL multiphysics networks, we generate CO_2 label models and their associated geophysical data using 3D flow and 2D and 3D geophysical forward modeling algorithms along with rock-physics relationships to convert reservoir properties to geophysical models. Then, the multiphysics imaging networks are trained on these models and data to learn the governing physics.

To train and evaluate our imaging networks for CO_2 monitoring, we generate a set of seismic, electromagnetic and gravity modeling data for the Kimberlina 1.2 CO_2 storage and flow model (Zhou and Birkholzer, 2011, Wainwright et al., 2013). The Kimberlina model was developed for understanding a commercial-scale CO_2 storage candidate site in the Southern San Joaquin Basin of California, 30 km northwest of Bakersfield, CA, USA. For realistic evaluation of various geophysical techniques for monitoring CO_2 plumes, the Kimberlina model was created based on geological and hydrological data acquired from a number of wells in the region. CO_2 injection and flow are simulated using TOUGH2

(Pruess, 1999), a modeling software for nonisothermal flows of multicomponent, multiphase fluids in porous and fractured media.

The Kimberlina 1.2 model consists of 300 different CO_2 injection and flow realizations. Each realization is based on a unique set of reservoir and flow modeling parameters created stochastically. For detailed description about the 3D Kimberlina 1.2 model, the reader is referred to Mansoor et al. (2018). In this study, we choose a single case from them. The single 3D realization consists of 35 snapshots of 3D simulated hydrological properties from 0 to 200 years. CO_2 is injected into a sandstone reservoir at 2750m in depth. The 3D simulation starts with CO_2 injection at a constant rate of 5 million tons per year for the first 50 years and covers a post-injection period of the remaining 150 years. Note that in the first 100 years, the 3D snapshots are taken at fine time intervals (Year 0, 1, 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100) because CO_2 plumes are rapidly formed and move. After the first 100 years, the CO_2 plumes move slowly. Thus, the snapshot interval gradually increases (i.e., 110, 120, 130, 140, 150, 175, 200 years). The first column of **Figure** 3 shows cross-sectional views of CO_2 saturation models at selected time intervals.

There are three layers of CO_2 that represent the three different high porosity zones inside the reservoir. The hydrogeological and reservoir properties (e.g., dissolved solids, temperature, formation porosity, fluid saturation, bulk modulus, shear modulus, density and others) that the Kimberlina model includes are converted to seismic velocity, resistivity and density models using a conversion workflow described in Wang et al. (2018), Yang et al. (2019) and Alumbaugh et al. (2021). **Figure** 3 shows P-wave velocity, electrical resistivity, and density difference models at selected time intervals. For more details on the conversion steps involved in generating the geophysical models, the reader is referred to the references mentioned above.

Here, we briefly describe the numerical simulation of multiphysics data over the Kimberlina geophysical models. The synthetic surface seismic data are generated by solving the acoustic wave equation (Alford et al., 1974; Moczo et al., 2007). To make the computational cost tractable, we slice a 3D Kimberlina velocity model at intervals of 100m and create 53 2D velocity models (**Figure** 4a). As a result, the 35 3D Kimberlina velocity models produce 1855 (=53 2D velocity models × 35 snapshots) 2D velocity models. Each 2D model has six pointwise surface pressure sources at an interval of 1 km (i.e., y=0, 1, 2, 3, 4, 5 km, and z=0 km). We use a Ricker wavelet with a peak frequency of 25 Hz as a source waveform. Seismic full waveform data are sampled using 601 surface geophones ranging from y=-2 km to 4 km at an interval of 10 m. The data that share a common source are gathered together (i.e., a common shot gather) and used as an input for the DL imaging network.

For electromagnetic (EM) monitoring of CO_2 plumes, we utilize a borehole-to-surface EM configuration where a vertical electric dipole source is placed near the injection depth below the two observation wells (Figure 4b) and operate at a frequency range from 0.1 to 8 Hz (i.e., 0.1, 0.3, 0.6, 0.8, 1.0, 3.0, 6.0, and 8.0 Hz) and the electric fields are measured on the surface. Note we have chosen to simulate this borehole-to-surface EM data acquisition configuration as it was noted by Gasperikova et al. (2022) to provide high sensitivity to injected CO_2 at depth. For simplicity, the effects of steel-cased injection and observation wells on EM measurements are ignored here. Note that diffusive EM modeling is computationally lighter than seismic wave modeling because EM modeling meshes can be coarser than in the seismic case. Here we use a 3D finite-difference EM diffusion algorithm (Commer and Newman, 2008) for simulating 2D inline borehole-to-surface EM responses over the 3D resistivity models (i.e., pseudo-2D data, **Figure** 4b) rather than slicing the 3D models. Horizontal surface electric fields are recorded using 31 receiver stations from -2 to 4 km in the y direction at an interval of 200 m. A common shot gather configuration is used to collect and prepare EM data for the DL network.

Gravity data are also simulated along the same surface survey lines used for the EM as well as within the two observation wells from 1500 m to 2500 m in depth. In practice, gravity data contain a significant component of time-invariant signals. The signals include the background rock density and the topography effects. Therefore, CO_2 monitoring uses the time-lapse difference gravity data because they are only sensitive to the density changes associated with CO_2 plumes. The three orthogonal components of the time-lapse gravity anomaly data are simulated using the 3D gravity modeling algorithm developed by Rim and Li (2015). **Figure** 5 shows some samples of the multiphysics data used for this paper.

The three synthetic geophysical data sets have different sampling densities. Hence, using down-sampling and linear interpolation, we map the data on a common data array, (x, 512, 256) where x is the number of sources and/or data components. For example, due to the memory and computational cost issues, each seismic shot-gather (e.g., Figure 5a) is down-sampled to an array of (512, 256) where the row and column are time and geophone positions. Because one velocity model has six source positions as described earlier, six down-sampled shot gathers are combined together, form an array of (6, 512, 256) and are fed into the DL imaging networks. EM data are linearly interpolated and mapped onto an array of (512, 256) where the row and column are frequency and sensor positions, respectively. As one conductivity model has two borehole source positions, EM sensors measure three data components (two horizontal electric fields and one horizontal magnetic field), and each EM data consists of real and imaginary components, EM data for one conductivity model are packed into (12, 512, 256). The time-lapse gravity data are also linearly interpolated and mapped onto an array of (512, 256) where the row and column are depth and sensor positions. As done in the other two data, the three components of the gravity data are combined together and packed into (3, 512, 256).

Implementing and Training Deep Neural Network

The multiphysics imaging networks have been implemented in Python using TensorFlow (Abadi et al., 2015) and Keras libraries (Chollet, 2018). These libraries include a range of DL functions and optimization tools and allow us to rapidly implement and evaluate the imaging capability of the proposed networks. In this work, we implement three network architectures having a different number of encoders ranging from one to three. The network with one encoder is used for inverting a single type of geophysical data (i.e., either seismic, EM or gravity data), whereas the network with multiple encoders is used for simultaneously inverting multiple types of data. Despites the differences in the number of encoders, the three networks share common building blocks as shown in **Figures** 1 and 2. Thus, their implementations are also similar to each other, facilitating the development of the single and multiphysics imaging networks. For example, **Figure** 6 compares the core implementation between single physics and multiphysics imaging networks. Both shares the same encoder function. Their decoder functions are slightly different from each other to concatenate a different number of encoder outputs.

Once the DL networks are implemented, the CO_2 label models and multiphysics data are split into three different sets: test, training and validation sets. As shown in **Figure** 3, the CO2 plumes are rapidly formed in the early time (e.g., 0-50 years) but change little in the late time (e.g., 100-200 years). We select the data and model for year 20 for the test sets because we observe a sufficient amount of CO_2 inside the reservoir after 20 years of CO_2 injection. Because we slice the 3D test CO_2 model (i.e., year 20) at intervals of 100m and create 53 2D label models as mentioned earlier, the test dataset has 53 different 2D label models. The 53 models can be thought of as results from 2D CO_2 flow simulation at different times. The test data are used only for evaluating the prediction accuracy of the DL imaging networks after the training phase is completed. During the evaluation phase, the test data are contaminated with Gaussian noise with zero mean and standard deviation of 10%. In short, our test dataset consists of 53 CO_2 label models and multiphysics data, and the remaining 1802 pairs of multiphysics data and CO_2 label models are used as the training and validation dataset as will be discussed below.

The bootstrap aggregating (bagging) method (Friedman et al., 2006; James et al., 2021) is an ensemble learning method and utilizes the bootstrap method for generating a number of training data sets. During the bootstrapping phase, new training data are randomly sampled from the original training set assuming a uniform distribution. Because of the nature of the replacement process, some data can be drawn more than once and some are never employed. On average, each bootstrapped training data set includes about two-thirds of the original training data that are employed. The remaining one-third of the data not drawn from the original data set are used as a validation set. As a result, using this bagging method each DL imaging network is trained with its own unique training and validation data. These two data sets are used differently during the training phase. The training data set is directly fed to the network. In contrast, the validation set is not directly fed to the network for training but instead used for estimating a prediction error when a new data set is used during the prediction phase. Loss values (i.e., misfits) on the training and validation data sets are called the training loss and the validation loss, respectively. The training phase ends when the validation loss no longer decreases.

Training individual networks for the bagging method is an embarrassingly parallel problem. In our work, we use eight NVIDIA Tesla P100-PCIE GPUs and simultaneously train eight networks for each bagging method (i.e., three DL single-physics imaging networks for seismic, EM and gravity and four DL multiphysics imaging networks for seismic-EM, seismic-EM-gravity, EM-gravity, and seismic-gravity). On average, it took about 2 hours to complete training one multiphysics imaging network. In contrast, the single-physics imaging network was trained in about 1 hour due to its relatively small size. Once fully trained, both single-physics and multiphysics imaging network can predict CO_2 saturation and its uncertainty in a few seconds on the GPUs, enabling us to monitoring GCS in real time. Based on our experience, it would take a few weeks or months to invert this kind of multiphysics data through conventional joint inversion experiments. Nonetheless, it does not mean that the computational cost of the DL inversion is significantly lower than the conventional inversion as discussed earlier in this paper. In contrast to the conventional conversion, the DL inversion requires preparing a large set of realistic label models and geophysical data before field data are acquired. In other words, the DL inversion pays most of the computational cost up front, whereas the conventional inversion performs its major computation after field data are measured.

We chose a mean squared error (MSE) as a loss function for our training. **Figure** 7 shows some examples of training history of the seven DL imaging networks. The Adam optimizer (Kingma and Ba, 2015) is used with a batch size of 32 and a learning rate of 10^{-4}. Note that once training is completed, the multiphysics networks (**Figure** 7b) show smaller final validation loss values than the single-physics networks (**Figure** 7a), indicating that the use of multiphysics data can improve the prediction accuracy. In general, using a large number of networks for the bagging method does not result in overfitting (James et el., 2021). In practice, however, the number of networks used for the bagging method is limited due to computational costs, and needs to be determined based on the characteristics of the data and the estimator (i.e., the DL imaging network). After trial and error using different numbers of the networks varying from 10 to 400, we have found that mean and standard deviation values of the recovered CO_2 images change little and settle down when more or less 70-80 predictions are made and used together.

In this work, we want to tweak the number of the networks so that the method can predict CO_2 saturation with the smallest possible variance without critically increasing the computational cost. To safely ensure the convergence of mean and standard deviation values, we use 100 independent networks in the bagging method and predict 100 CO_2 saturation models. After the prediction phase is done, we have a distribution of CO_2 saturation at each cell, and calculate a mean CO_2 saturation and its standard deviation. The mean CO_2 saturation model is used as a final image and the standard deviation as a measure of uncertainty. Once the training phase is completed, the DL imaging and statistical analysis are completed in real time without human interactions.

DL Multiphysics Imaging Experiments

Before we evaluate the performance of the DL multiphysics networks, we first perform the DL single-physics imaging (i.e., seismic, EM and gravity), and compare the imaging results for the single data types (the 2nd to 4th columns of **Figure** 8) to the true CO_2 saturation test models (the 1st column of **Figure** 8). The DL seismic network recovers the CO_2 plumes fairly well. For example, the network clearly recovers the three layers of CO_2 plumes that represent three high porosity sand layers, and the lateral extent of the recovered plumes is close to that of the true model. This DL seismic imaging also recovers a high concentration of CO_2 saturation near the injection point as shown in **Figure** 8f. Note that a rock physics model used for converting the Kimberlina model into the P-wave velocity model is based on the average of the upper and lower Hashin-Shtrikman bounds (Yang et al., 2019; Mavko et al., 2020), Thus, the seismic training data are sensitive to a broader range of CO_2 saturation even at high CO_2 saturation levels. However, in practice, if a rock physical relationship between P-wave velocity and CO_2 saturation is close to the lower Hashin-Shtrikman bound, the sensitivity of the seismic imaging to the high concentration of CO_2 saturation would diminish (Kim et al., 2010; Davis et al., 2019, Gasperikova and Li, 2021). The non-seismic methods such as EM and gravity methods can help to fill the gap.

The 3rd column of **Figure** 8 shows the DL EM imaging results. EM also delineates the CO_2 plumes very well. The high CO_2 saturation near the injection is recovered as expected, and the three layers of high saturation of CO_2 are clearly recovered. The lateral extent of the recovered CO_2 plume is recovered reasonably well but less accurately compared to the seismic images due to the fact that borehole-to-surface EM mainly illuminates a triangular region defined by the borehole sources and the surface receiver array. The reader is referred to Um et al. (2020) for a discussion of the region of sensitivity of the borehole-to-surface EM measurement configuration.

The DL gravity imaging also clearly detects the presence of the top CO_2 plume layer and recovers its lateral extent fairly well. However, it does not recover the lower two layers. This relatively poor result can be inferred from the training history plots (**Figure** 7a) where the final validation loss of the DL gravity network is an order of magnitude larger than that of the other two single-physics imaging networks. This is mainly due to the fact that the gravity data are all recorded above the top CO_2 layer and are not sensitive enough to the fine structures that the CO_2 label models have, so the DL imaging network cannot be trained for fully recovering such details and yields relatively high validation loss values. The overall imaging sensitivity and resolution of the gravity-generated images presented here are comparable to those found in DL geophysics literature (e.g., Yang et al., 2022). Note that the gravity imaging does not compete against seismic or EM in terms of resolution when used in a multiphysics interpretation on imaging mode. Instead, it provides unique sensitivity to the density change due to fluid substitution and complements the other two methods.

Figure 9 shows the calculated uncertainty estimates in terms of the standard deviation. In all three cases, non-zero uncertainty values are found inside and around the recovered plumes. The EM images shows a slightly higher magnitude and greater lateral spread of the non-zero standard deviation values compared to the seismic generated images of CO_2. In contrast, the uncertainty estimates of the gravity-generated images look counterintuitive. The absence of the 2nd and 3rd CO_2 layer in the gravity images is not correlated with abnormally high standard deviation values (**Figures** 9c, 9f and 9i). This is because all 100 gravity-generated images used for the bagging method equally fail to recover the 2nd and 3rd layers and thus the standard deviation does not reflect imaging errors associated with the absence of the two layers. Rather, the standard deviation merely measures dispersion of the 100 recovered CO_2 images at each cell. In other words, the ensemble method neither indicates nor overcomes the limitations of the gravity method. Instead, based on geophysical principles and sensitivity studies, one must carefully choose one or multiple geophysical methods for monitoring a given GCS site. Despite the limitation of the gravity method as a single geophysical imaging tool, we demonstrate that the DL multiphysics imaging network including the gravity component can best improve prediction accuracy and reduce uncertainty as shown below.

Figure 10 compares four multiphysics imaging results (i.e., seismic-EM-gravity, seismic-EM, EM-gravity, and seismic-gravity). For this particular synthetic data set based on the Kimberlina reservoir model and associated rock-physics conversions to geophysical properties, the DL multiphysics imaging does not significantly improve the overall imaging accuracy compared with the DL seismic imaging because the seismic rock physics transform that was used both for the test and training data are sensitive to a broad range of CO_2 saturations as discussed earlier. Thus, the DL seismic imaging alone can predict CO_2 saturation fairly well. For example, at x=0 km, the DL seismic-EM-gravity image (**Figure** 10f) is nearly identical to the image from the DL seismic network (**Figure** 8f). However, at the edge of the plume (i.e., x=1 km), the DL multiphysics image (**Figure** 10j) better recovers the lateral extent of the plumes than the DL seismic image (**Figure** 8j). The DL seismic-gravity image (**Figure** 10l) also shows the improved lateral resolution. Accordingly, it is reasonable to infer that this improvement results from the gravity data because the DL seismic-EM-generated CO_2 image (**Figure** 10i) does not show such enhanced lateral resolution. It is also worth mentioning that the EM-gravity-generated images (**Figures** 10c, 10g and 10k) are nearly comparable to those from the DL seismic network (**Figures** 8b, 8f and 8j), suggesting that the DL multiphysics imaging network for the non-seismic data can serve as a cost-effective tool for long-term monitoring of a GCS site.

Last, we compare the uncertainties associated with the DL multiphysics images (**Figure** 11) to those from the DL single-physics imaging networks (**Figure** 9). In general, the multiphysics imaging networks clearly reduce the magnitude of standard deviation values as well as their lateral spread. For example, the DL seismic-EM-gravity images show the smallest magnitude and lateral spread of the non-zero standard deviation values than any single-physics-generated image. The DL seismic-gravity images also approach a similar level of uncertainty. The DL EM-gravity images show slightly higher magnitude but smaller lateral spread of the non-zero standard deviation values compared with the seismic-generated images, demonstrating that the DL multiphysics network for inexpensive non-seismic data could effectively complement the seismic inversion.

Conclusions

We have developed a novel DL multiphysics network for imaging CO_2 plumes and estimating image uncertainty in real time. The DL multiphysics network consists of three encoders and one decoder. The three encoders separately analyze three different geophysical data (i.e., seismic, EM, gravity) but share one decoder to combine imaging capabilities of the different geophysical data for improving the prediction accuracy. The network is trained on pairs of CO_2 label models and multiphysics data so that CO_2 saturation is directly imaged. Using a bootstrap aggregating method, we train a number of the DL imaging networks simultaneously and use these results to calculate a mean CO_2 saturation as well as a standard deviation.

Using CO_2 label models and multiphysics data that are based on the realistic Kimberlina GCS flow models, we systematically compare the DL multiphysics imaging results with DL single physics imaging results. We demonstrated that the DL multiphysics network for seismic, EM, and gravity data not only improves the prediction accuracy but also reduces uncertainties associated with CO_2 saturation images. Our numerical modeling studies also showed that the DL multiphysics network for EM and gravity data produces CO_2 images nearly comparable to those from the DL seismic imaging network, suggesting that the DL imaging network for the non-seismic data would be an effective low-cost monitoring tool in between regular seismic monitoring.

In this work, we have evaluated the imaging and uncertainty analysis capabilities of the DL inversion network using numerical modeling data contaminated with Gaussian noise. However, real field data always include a range of unknown noise and measurement errors that cannot be easily duplicated in numerical modeling experiments such as those we have demonstrated here. We have also assumed that rock-physics relationships between reservoir properties and geophysical models are exactly known to us. However, in real field data, the relationships may neither be exactly known nor fully understood. Accordingly, these aspects will increase the level of uncertainty beyond those demonstrated in this work. Our future plans involve applying the DL imaging network on GCS field data and evaluating its robustness and limitations in real-world situations.

Acknowledgements

This work was completed as part of the Science-informed Machine learning to Accelerate Real Time decision making for Carbon Storage (SMART-CS) Initiative (edx.netl.doe.gov/SMART). Support for this initiative was provided by the U.S. Department of Energy's (DOE) Office of Fossil Energy's Carbon Storage Research program through the National Energy Technology Laboratory (NETL). The authors wish to acknowledge Mark McKoy (NETL, Carbon Storage Technology Manager), Darin Damiani (DOE Office of Fossil Energy, Carbon Storage Program Manager), and Mark Ackiewicz (DOE Office of Fossil Energy, Director, Division of Carbon Capture and Storage Research and Development), for programmatic guidance, direction and support. The numerical modeling experiments presented in this work were carried out using high performance computing facilities provided by the Geosciences Measurement Facility (GMF) at Lawrence Berkeley National Laboratory and NETL.

Data Availability Statement

At the point of submitting this work to Geophysical Prospecting, the models and data used here are available only to registered users of U.S. Department of Energy's NETL EDX system, https://edx.netl.doe.gov/. The models and data will be transferred over to a publically available location on EDX in late 2022.

Figure Captions

Figure 1. The DL multiphysics network architecture for imaging CO_2 saturation. The size of the feature maps is shown at the top of the layers. The number of channels is shown at the bottom of the layers.

Figure 2. The single-physics network for imaging CO_2 saturation. (Um et al., 2022).

Figure 3. Cross-sectional (y-z plane) views of the Kimberlina CO_2 saturation models (the 1st column) and their corresponding P-wave velocity models (the 2^{nd} column), the electrical resistivity models (the 3^{rd} column) and the density difference models (the 4^{th} column).

Figure 4. The geophysical survey configurations for monitoring CO_2 plumes. (a) The surface seismic configuration. The 53 red lines indicate surface seismic survey lines. (b) the borehole-to-surface EM configuration. On the cross-sectional view (y=0 km), the two vertical red arrows indicate borehole electric dipole sources. On the map view (z=2.9 km), the red lines indicate the surface electric field survey lines. The same surface lines and boreholes are used for the gravity data generation.

Figure 5. Examples of Kimberlina multiphysics data for training the DL single-physics and multiphysics networks. (a) Surface seismic modeling data. (b) Borehole-to-surface EM modeling data. (c) Surface and borehole gravity data.

Figure 6. Comparison of the core implementation between (a) the single-physics imaging network and (b) the multiphysics imaging network. Both shares the same encoder function.

Figure 7. Training and validation loss plots for (a) the DL single-physics imaging networks and (b) the DL multiphysics imaging networks. Seis, EM and GRV stand for seismic, electromagnetic and gravity, respectively. TL and VL stand for training loss and validation loss, respectively.

Figure 8. Comparison of the true Kimberlina CO_2 saturation model (the 1st column) and the mean values of CO_2 saturation images recovered from the three different DL single-physics imaging methodologies (the 2^{nd} to 4^{th} columns).

Figure 9. Cross-sectional views of uncertainties associated with the CO_2 saturation images from the three DL single-physics networks.

Figure 10. Cross-sectional views of the mean values of CO_2 saturation images recovered from the four different DL multiphysics networks.

Figure 11. Cross-sectional views of uncertainties associated with the CO_2 saturation images from the four DL multiphysics networks.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S. (2015) TensorFlow: Large-scale machine learning on heterogeneous systems.

Ajo-Franklin, J.B., Peterson, J., Doetsch, J. and Daley, T.M. (2013) High-resolution characterization of a CO_2 plume using crosswell seismic tomography: Cranfield, MS, USA. International Journal of Greenhouse Gas Control, 18, 497-509.

Alford, R.M., Kelly, K.R. and Boore, D.M. (1974) Accuracy of finite-difference modeling of the acoustic wave equation. Geophysics, 39(6), 834-842.

Alnes, H., Eiken, O., Nooner, S., Sasagawa, G., Stenvold, T. and Zumberge, M. (2011) Results from Sleipner gravity monitoring: Updated density and temperature distribution of the CO2 plume. Energy Procedia, 4, pp.5504-5511.

Alumbaugh, D., Commer, M., Crandall, D., Gasperikova, E., Harbert, W., Li, Y., Lin, Y., Samarasinghe, S., and Yang, X. (2021) Development of a Multi-Scale Synthetic Data Set for the Testing of Subsurface CO2 Storage Monitoring Strategies. AGU Meeting.

Araya-Polo, M., Jennings, J., Adler, A. and Dahlke, T. (2018) Deep-learning tomography. The Leading Edge, 37(1), 58-66.

Arts, R., Eiken, O., Chadwick, A., Zweigel, P., Van der Meer, L. and Zinszner, B. (2003) Monitoring of CO_2 injected at Sleipner using time lapse seismic data. Greenhouse Gas Control Technologies-6th International Conference, 347-352.

Benson, S.M. and Cole, D.R. (2008) CO_2 sequestration in deep sedimentary formations. Elements, 4(5), 325-331.

Chadwick, A., Williams, G., Delepine, N., Clochard, V., Labat, K., Sturton, S., Buddensiek, M.L., Dillen, M., Nickel, M., Lima, A.L. and Arts, R. (2010) Quantitative analysis of time-lapse seismic monitoring data at the Sleipner CO_2 storage operation. The Leading Edge, 29(2), 170-177.

Chollet, F. (2017) Deep learning with Python. Simon and Schuster.

Chollet, F. (2018) Keras: The python deep learning library. Astrophysics Source Code Library.

Colombo, D., Li, W., Sandoval-Curiel, E. and McNeice, G.W. (2020) Deep-learning electromagnetic monitoring coupled to fluid flow simulators. Geophysics, 85(4), WA1-WA12.

Commer, M. and Newman, G.A. (2008) New advances in three-dimensional controlled-source electromagnetic inversion. Geophysical Journal International, 172(2), 513-535.

Davis, T.L., Landrø, M. and Wilson, M. eds. (2019) Geophysics and Geosequestration. Cambridge University Press.

Eiken, O., Stenvold, T., Zumberge, M., Alnes, H. and Sasagawa, G. (2008) Gravimetric monitoring of gas production from the Troll field. Geophysics, 73(6), WA149-WA154.

Fichtner, A., 2010. Full seismic waveform modelling and inversion. Springer Science & Business Media.

Friedman, J., Hastie, T. and Tibshirani, R. (2006) The elements of statistical learning. Springer.

Gasperikova, E. and Hoversten, G.M., 2006. A feasibility study of nonseismic geophysical methods for monitoring geologic CO_2 sequestration. The Leading Edge, 25(10), 1282-1288.

Gasperikova, E. and Li, Y. (2021) Time-lapse electromagnetic and gravity methods in carbon storage monitoring. The Leading Edge, 40(6), 442-446.

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_2 storage plumes, Int. J. Greenhouse Gas Control, 114, 103585.

Goodfellow, I., Bengio, Y. and Courville, A. (2016) Deep learning. MIT press.

Guo, R., Yao, H.M., Li, M., Ng, M.K.P., Jiang, L. and Abubakar, A. (2020) Joint inversion of audio-magnetotelluric and seismic travel time data with deep learning constraint. IEEE Transactions on Geoscience and Remote Sensing, 59(9), 7982-7995.

Harbert, W., Daley, T.M., Bromhal, G., Sullivan, C. and Huang, L. (2016) Progress in monitoring strategies for risk reduction in geologic CO_2 storage. International Journal of Greenhouse Gas Control, 51, 260-275.

Hu, Y., Wei, X., Wu, X., Sun, J., Chen, J., Chen, J. and Huang, Y. (2021) Deep learning-enhanced multiphysics joint inversion. In First International Meeting for Applied Geoscience & Energy, 1721-1725, Society of Exploration Geophysicists.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2021) An introduction to statistical learning. Springer.

Jenkins, C., Chadwick, A. and Hovorka, S.D. (2015) The state of the art in monitoring and verification—ten years on. International Journal of Greenhouse Gas Control, 40, 312-349.

Kaur, H., Zhong, Z., Sun, A. and Fomel, S. (2021) Time-lapse seismic data inversion for estimating reservoir parametersusing deep learning. Interpretation, 10(1), 1-91.

Kim, J., Xue, Z. and Matsuoka, T. (2010) Experimental study on CO_2 monitoring and saturation with combined P-wave velocity and resistivity. In International Oil and Gas Conference and Exhibition in China. OnePetro.

Kingma, D.P. and Ba, J. (2015) A method for stochastic optimization. In: 3rd International Conference on Learning Representations. Conference Track Proceedings.

Lazaratos, S.K. and Marion, B.P. (1997) Crosswell seismic imaging of reservoir changes caused by CO_2 injection. The Leading Edge, 16(9), 1300-1308.

Li, D., Peng, S., Huang, X., Guo, Y., Lu, Y. and Cui, X. (2021) Time-lapse full waveform inversion based on curvelet transform: Case study of CO_2 storage monitoring. International Journal of Greenhouse Gas Control, 110, 103417.

Li, Y. and Yang, D. (2021) Electrical imaging of hydraulic fracturing fluid using steel-cased wells and a deep-learning method. Geophysics, 86(4), E315-E332.

Mansoor, K., T. Buscheck, X. Yang, S. Carroll, X. Chen. 2018, LLNL Kimberlina 1.2 NUFT Simulations, https://edx.netl.doe.gov/dataset/llnl-kimberlina-1-2-nuft-simulations-june-2018, DOI: 10.18141/1603336

Mavko, G., Mukerji, T. and Dvorkin, J. (2020) The rock physics handbook. Cambridge university press.

Metz, B., Davidson, O., De Coninck, H.C., Loos, M. and Meyer, L. (2005) IPCC special report on carbon dioxide capture and storage. Cambridge University Press.

Moczo P., J.O. Robertsson and L. Eisner (2007) The finite-difference time-domain method for modeling of seismic wave propagation: Advances in geophysics, 48, 421-516.

Oh, S., Noh, K., Seol, S.J. and Byun, J. (2020) Cooperative deep learning inversion of controlled-source electromagnetic data for salt delineation. Geophysics, 85(4), E121-E137.

Pruess, K., Oldenburg, C.M. and Moridis, G.J. (1999) TOUGH2 user's guide version 2 (No. LBNL-43134). Lawrence Berkeley National Lab.

Puzyrev, V. (2019) Deep learning electromagnetic inversion with convolutional neural networks. Geophysical Journal International, 218(2), 817-832.

Queißer, M. and Singh, S.C. (2013) Full waveform inversion in the time lapse mode applied to CO_2 storage at Sleipner. Geophysical prospecting, 61(3), 537-555.

Rim, H. and Li, Y. (2015) Advantages of borehole vector gravity in density imaging. Geophysics, 80(1), G1-G13.

Ringrose, P. (2020) How to Store CO_2 Underground: insights from early-mover CCS Projects. Springer.

Ronneberger, O., Fischer, P. and Brox, T. (2015) U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention, 234-241.

Rutqvist, J. (2012) The geomechanics of CO_2 storage in deep sedimentary formations. Geotechnical and Geological Engineering, 30(3), 525-551.

Shin, C. and Cha, Y. H. (2009) Waveform inversion in the Laplace—Fourier domain. Geophysical Journal International, 177(3), 1067-1079.

Sun, Y., Denel, B., Daril, N., Evano, L., Williamson, P. and Araya-Polo, M. (2020) Deep learning joint inversion of seismic and electromagnetic data for salt reconstruction. In SEG Technical Program Expanded Abstracts, 550-554, Society of Exploration Geophysicists.

Um, E.S., Commer, M. and Newman, G.A., 2014. A strategy for coupled 3D imaging of large-scale seismic and electromagnetic data sets: Application to subsalt imaging. Geophysics, 79(3), ID1-ID13.

Um, E.S., Kim, J. and Wilt, M. (2020) 3D borehole-to-surface and surface electromagnetic modeling and inversion in the presence of steel infrastructure. Geophysics, 85(5), E139-E152.

Um, E.S., Alumbaugh, D., Lin, Y. and Feng, S. (2022) Real time Deep Learning Inversion of Seismic Full Waveform Data for CO_2 Saturation and Uncertainty in Geological Carbon Storage Monitoring, Geophysical Prospecting, https://doi.org/10.1111/1365-2478.13197

Vasco, D.W., Daley, T.M. and Bakulin, A. (2014) Utilizing the onset of time-lapse changes: A robust basis for reservoir monitoring and characterization. Geophysical Journal International, 197(1), 542-556.

Virieux, J. and Operto, S. (2009) An overview of full-waveform inversion in exploration geophysics. Geophysics, 74(6), WCC1-WCC26.

Wainwright, H.M., Finsterle, S., Zhou, Q. and Birkholzer, J.T. (2013) Modeling the performance of large-scale CO_2 storage systems: a comparison of different sensitivity analysis methods. International Journal of Greenhouse Gas Control, 17, 189-205.

Wang, Z., Harbert, W.P., Dilmore, R.M. and Huang, L. (2018) Modeling of time-lapse seismic monitoring using CO2 leakage simulations for a model CO_2 storage site with realistic geology: Application in assessment of early leak-detection capabilities. International Journal of Greenhouse Gas Control, 76, 39-52.

Wu, Y. and Lin, Y. (2019) InversionNet: An efficient and accurate data-driven full waveform inversion. IEEE Transactions on Computational Imaging, 6, 419-433.

Yang, F. and Ma, J. (2019) Deep-learning inversion: A next-generation seismic velocity model building method. Geophysics, 84(4), R583-R599.

Yang, X., Buscheck, T.A., Mansoor, K., Wang, Z., Gao, K., Huang, L., Appriou, D. and Carroll, S.A. (2019) Assessment of geophysical monitoring methods for detection of brine and CO_2 leakage in drinking water aquifers. International Journal of Greenhouse Gas Control, 90, 102803.

Yang, X., Chen, X. and Smith, M.M. (2022) Deep learning inversion of gravity data for detection of CO_2 plumes in overlying aquifers. Journal of Applied Geophysics, 196, p.104507.

Zeng, Q., Feng, S., Wohlberg, B. and Lin, Y. (2021) Inversionnet3d: Efficient and scalable learning for 3d full waveform inversion. IEEE Transactions on Geoscience and Remote Sensing.

Zhang, Z.D. and Alkhalifah, T. (2019) Regularized elastic full-waveform inversion using deep learning. Geophysics, 84(5), R741-R751.

Zhang, Z. and Lin, Y. (2020) Data-driven seismic waveform inversion: A study on the robustness and generalization. IEEE Transactions on Geoscience and Remote sensing, 58(10), 6900-6913.

Zhou, Q. and Birkholzer, J.T. (2011) On scale and magnitude of pressure build-up induced by large-scale geologic storage of CO_2. Greenhouse Gases: Science and Technology, 1(1), 11-20.

Zoback, M.D. and Gorelick, S.M., 2012. Earthquake triggering and large-scale geologic storage of carbon dioxide. Proceedings of the National Academy of Sciences, 109(26), pp.10164-10168.

Figures

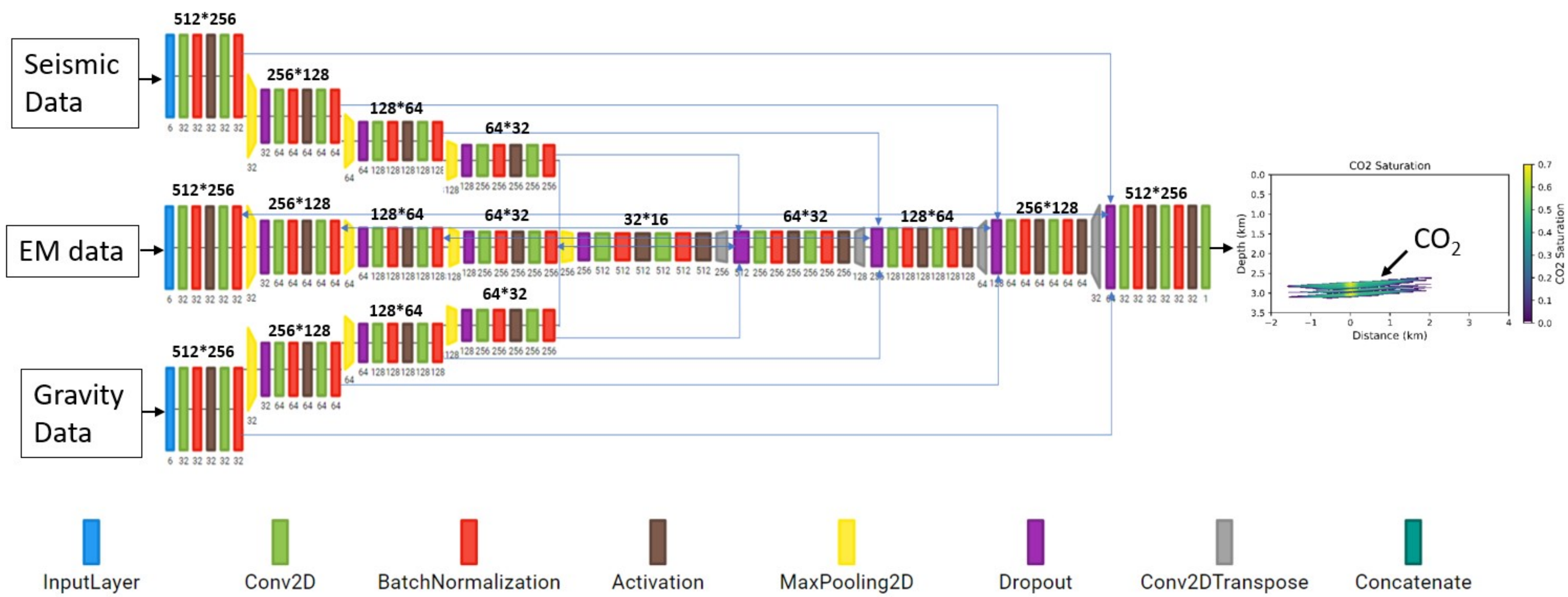


Figure 1. The DL multiphysics network architecture for imaging CO_2 saturation. The size of the feature maps is shown at the top of the layers. The number of channels is shown at the bottom of the layers.

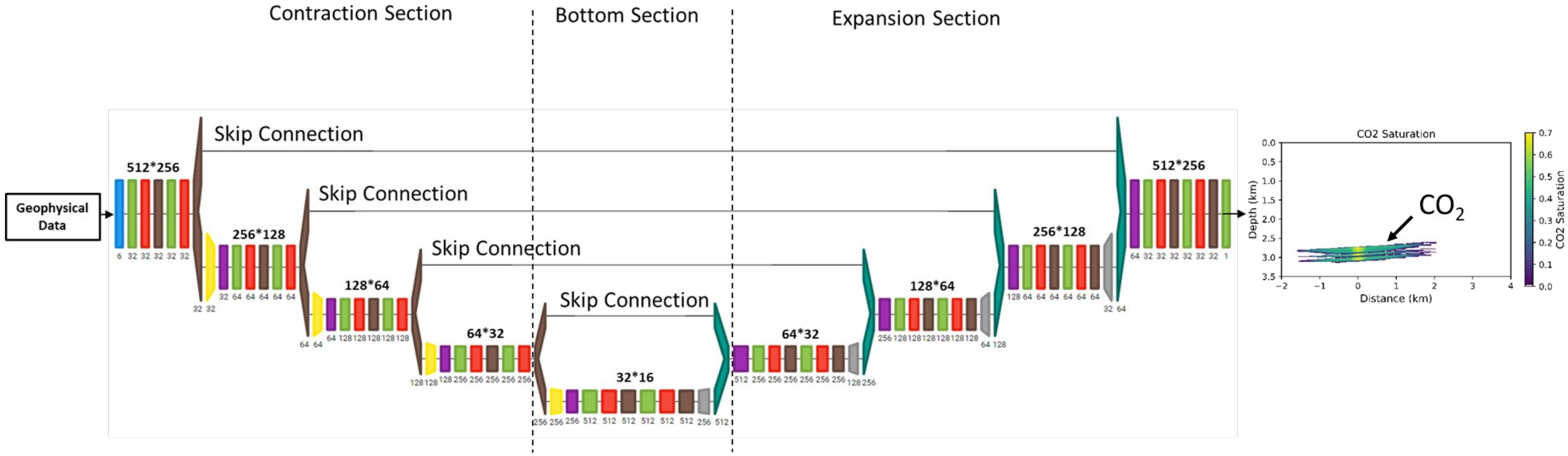


Figure 2. The single-physics network for imaging CO_2 saturation. (Um et al., 2022).

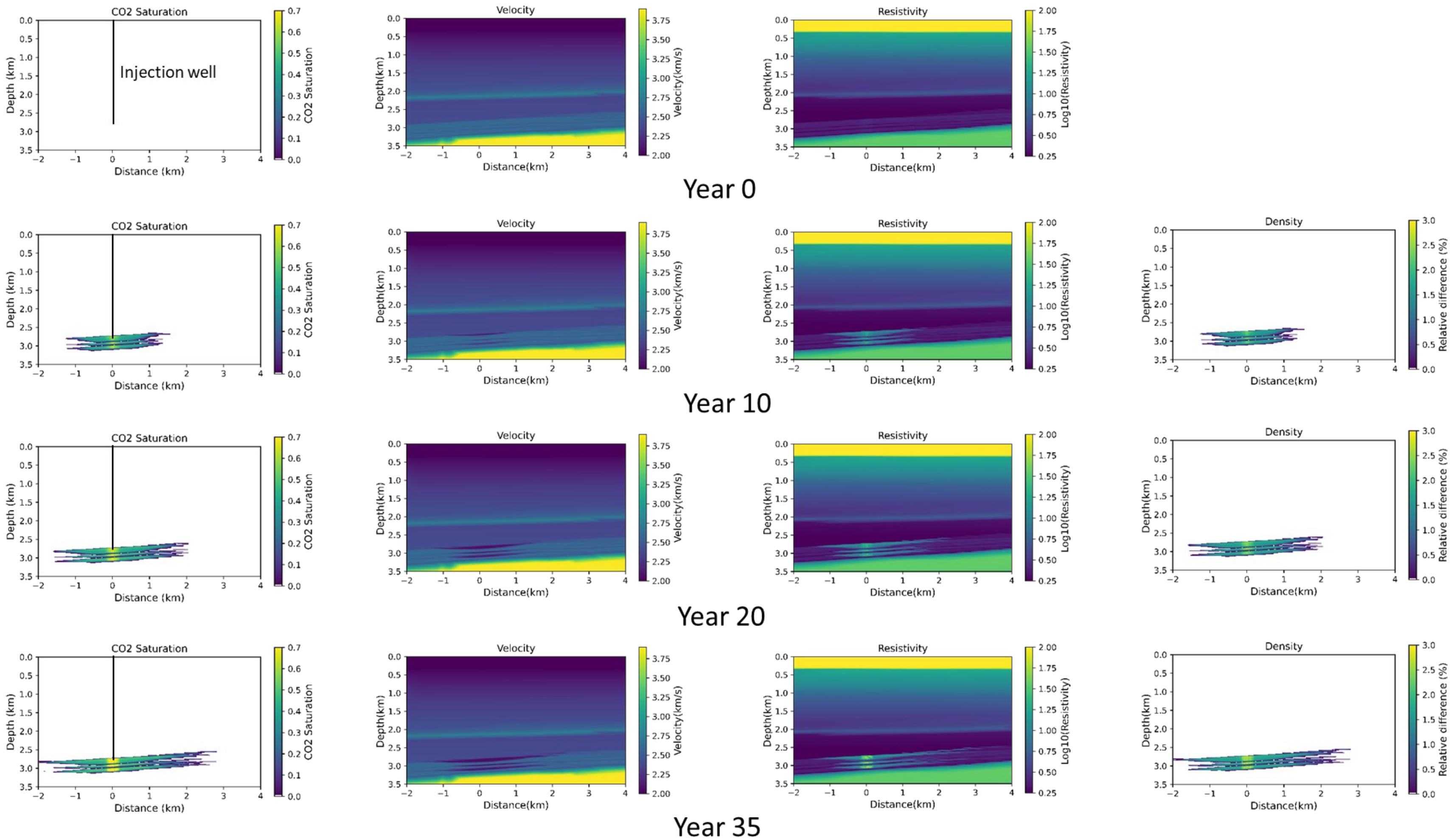


Figure 3. Cross-sectional (y-z plane) views of the Kimberlina CO_2 saturation models (the 1st column) and their corresponding P-wave velocity models (the 2nd column), the electrical resistivity models (the 3rd column) and the density difference models (the 4th column).

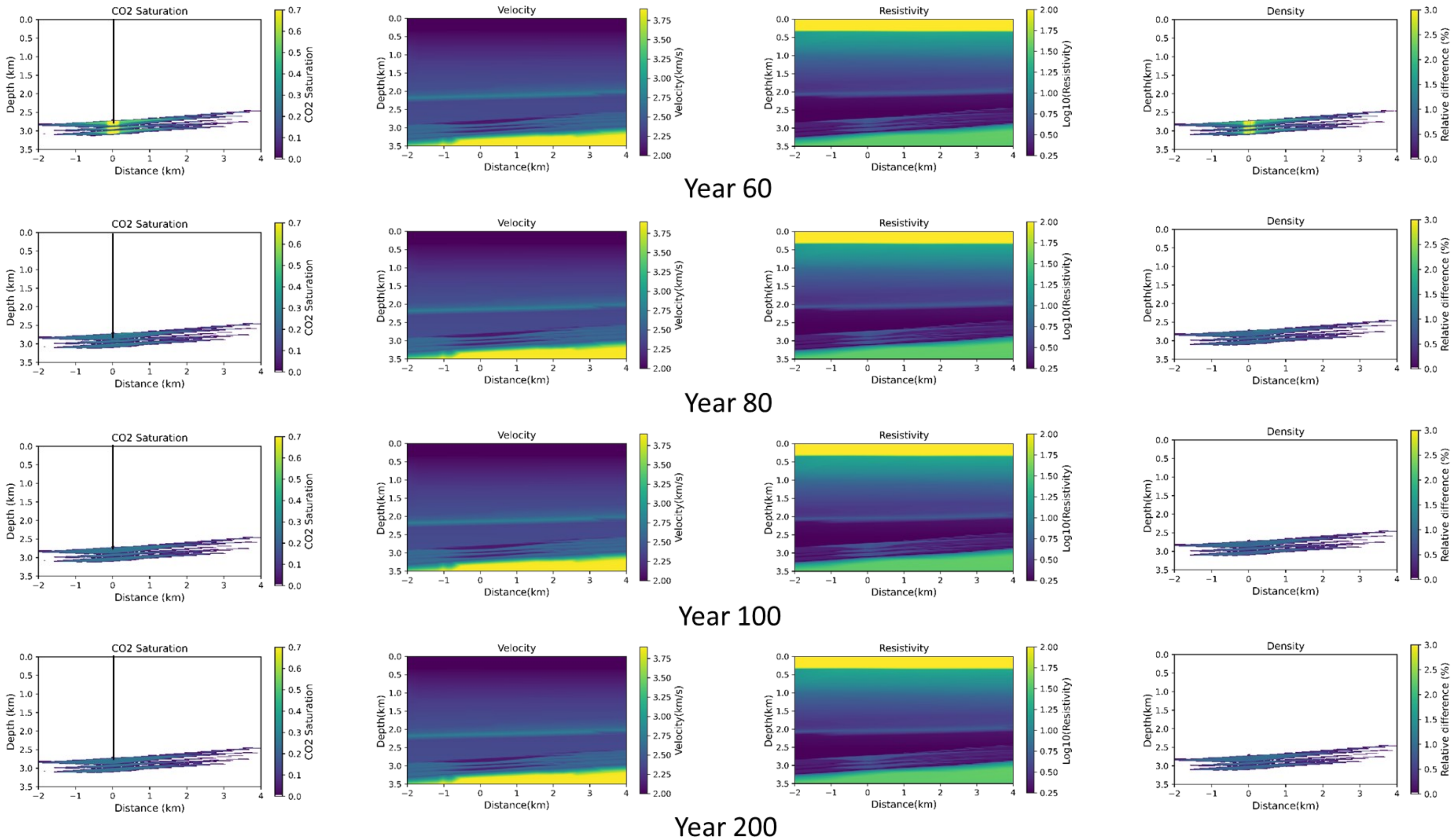
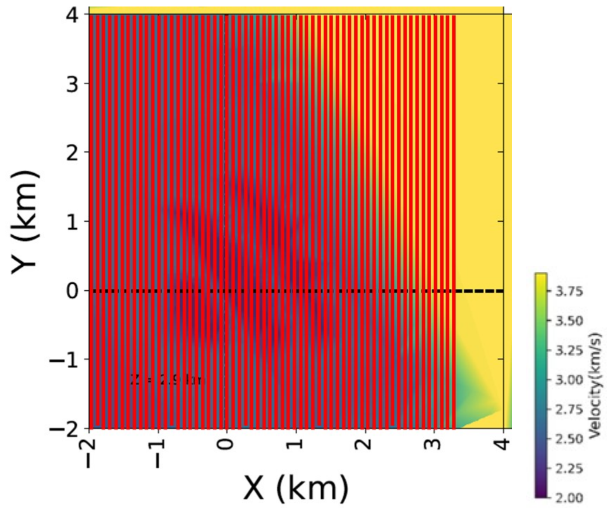
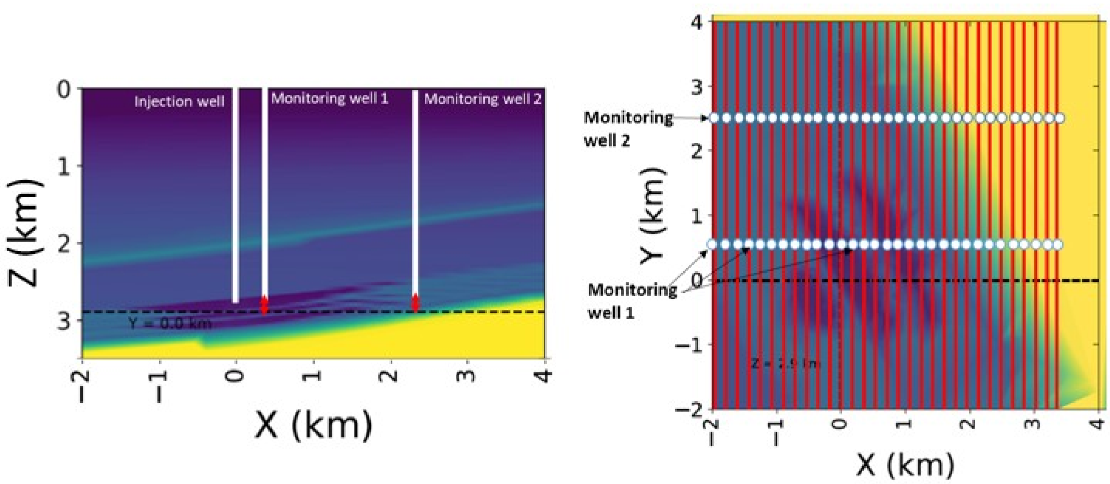


Figure 3. Continued.

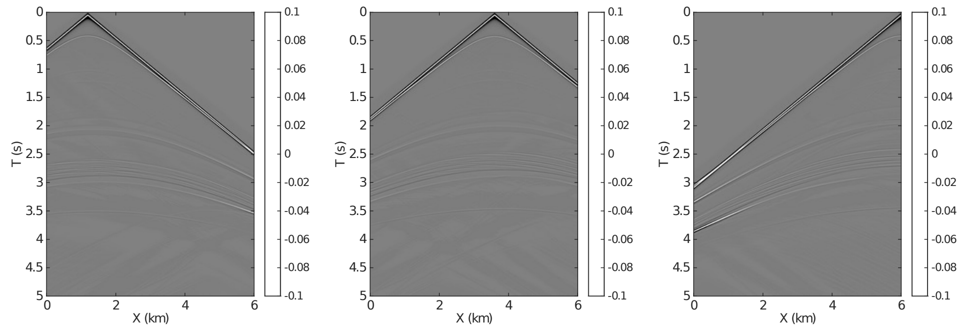


(a)

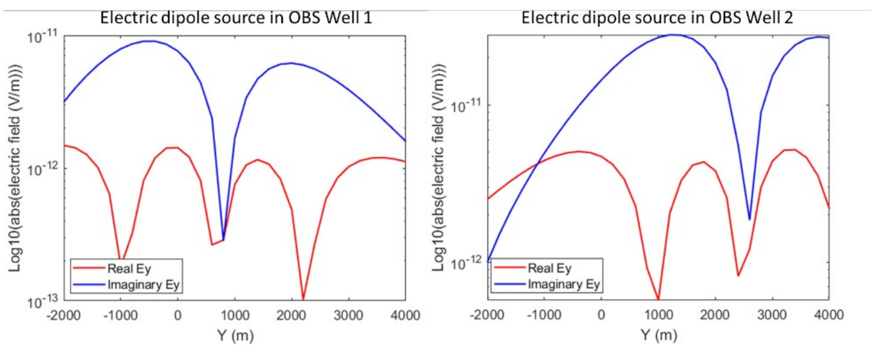


(b)

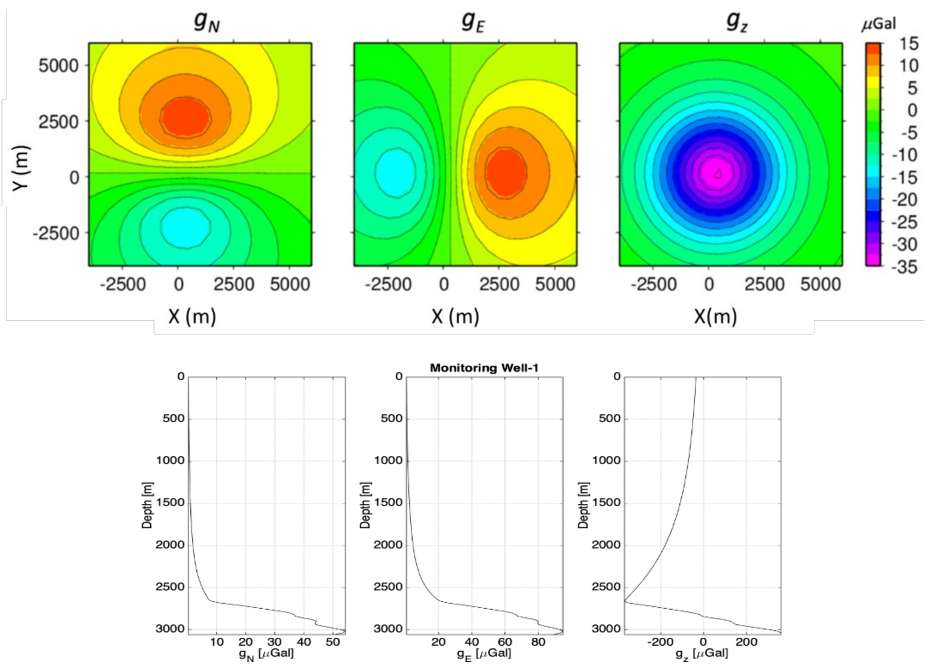
Figure 4. The geophysical survey configurations for monitoring CO_2 plumes. (a) The surface seismic configuration. The 53 red lines indicate surface seismic survey lines. (b) the borehole-to-surface EM configuration. On the cross-sectional view (y=0 km), the two vertical red arrows indicate borehole electric dipole sources. On the map view (z=2.9 km), the red lines indicate the surface electric field survey lines. The same surface lines and boreholes are used for the gravity data generation.



(a)



(b)



(c)

Figure 5. Examples of Kimberlina multiphysics data for training the DL single-physics and multiphysics networks. (a) Surface seismic modeling data. (b) Borehole-to-surface EM modeling data. (c) Surface and borehole gravity data.

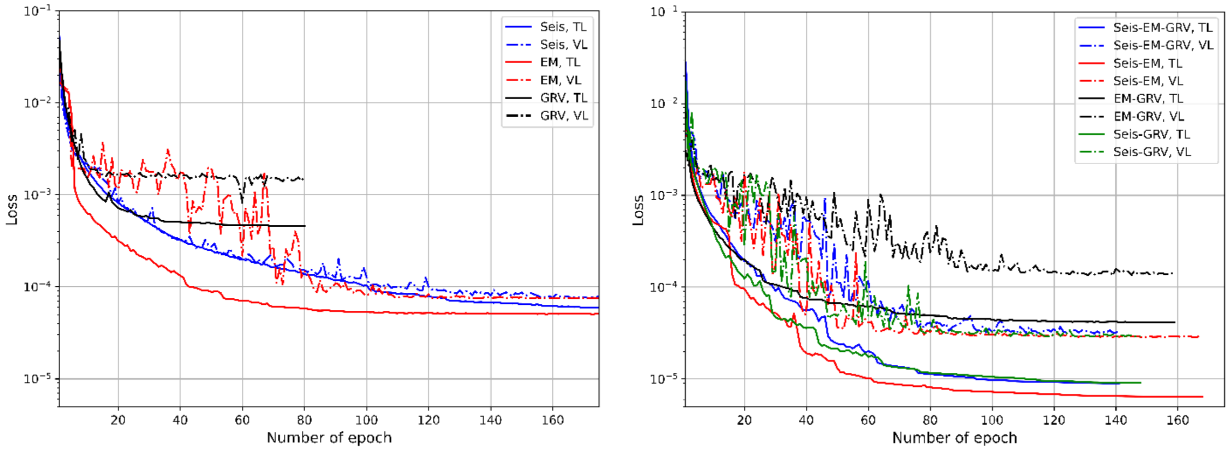
```
def conv_block(input, num_filters):
    x = Conv2D(num_filters, 3, padding="same")(input)
    x = BatchNormalization()(x)
    x = Activation("relu")(x)
    x = Conv2D(num_filters, 3, padding="same")(x)
    x = BatchNormalization()(x)
    x = Activation("relu")(x)
    return x

def encoder_block(input, num_filters):
    x = conv_block(input, num_filters)
    p = MaxPool2D((2, 2))(x)
    p = Dropout (dropout_value)(p)
    return x, p

def decoder_block(input, sf1, num_filters):
    x = Conv2DTranspose(num_filters, (2, 2), strides=2, padding="same")(input)
    x = Concatenate()([x, sf1])
    x = Dropout (dropout_value)(x)
    x = conv_block(x, num_filters)
    return x

def build_net(input_shape_seis):
    input_seis = Input(input_shape_seis)
    s1, p1 = encoder_block(input_seis, 32)
    s2, p2 = encoder_block(p1, 64)
    s3, p3 = encoder_block(p2, 128)
    s4, p4 = encoder_block(p3, 256)

    b1 = conv_block(p4, 512)

    d1 = decoder_block(b1, s4, 256)
    d2 = decoder_block(d1, s3, 128)
    d3 = decoder_block(d2, s2, 64)
    d4 = decoder_block(d3, s1, 32)

    outputs = Conv2D(1, 1, padding="same", activation="sigmoid")(d4)
    model = Model(input_seis, outputs)
    return model
```

(a)

```
def decoder_block2(input, sf1, sf3, num_filters):
    x1 = Conv2DTranspose(num_filters, (2, 2), strides=2, padding="same")(input)
    x1 = Concatenate()([x1, sf1])
    x1 = Concatenate()([x1, sf3])
    x = Dropout (dropout_value)(x1)
    x = conv_block(x, num_filters)
    return x

def build_net(input_shape_seis, input_shape_em):
    input_seis = Input(input_shape_seis)
    input_em = Input(input_shape_em)

    s1b, p1b = encoder_block(input_seis, 32)
    s2b, p2b = encoder_block(p1b, 64)
    s3b, p3b = encoder_block(p2b, 128)
    s4b, p4b = encoder_block(p3b, 256)

    s1c, p1c = encoder_block(input_em, 32)
    s2c, p2c = encoder_block(p1c, 64)
    s3c, p3c = encoder_block(p2c, 128)
    s4c, p4c = encoder_block(p3c, 256)

          p4=Concatenate()([p4b, p4c])
    b1 = conv_block(p4, 512)

    d1 = decoder_block2(b1, s4b, s4c, 256)
    d2 = decoder_block2(d1, s3b, s3c, 128)
    d3 = decoder_block2(d2, s2b, s2c, 64)
    d4 = decoder_block2(d3, s1b, s1c, 32)

    outputs = Conv2D(1, 1, padding="same", activation="sigmoid")(d4)
    model = Model([input_seismic, input_em], outputs)
    return model
```

(b)

Figure 6. Comparison of the core implementation between (a) the single-physics imaging network and (b) the multiphysics imaging network. Both shares the same encoder function.



(a) (b)

Figure 7. Training and validation loss plots for (a) the DL single-physics imaging networks and (b) the DL multiphysics imaging networks. Seis, EM and GRV stand for seismic, electromagnetic and gravity, respectively. TL and VL stand for training loss and validation loss, respectively.

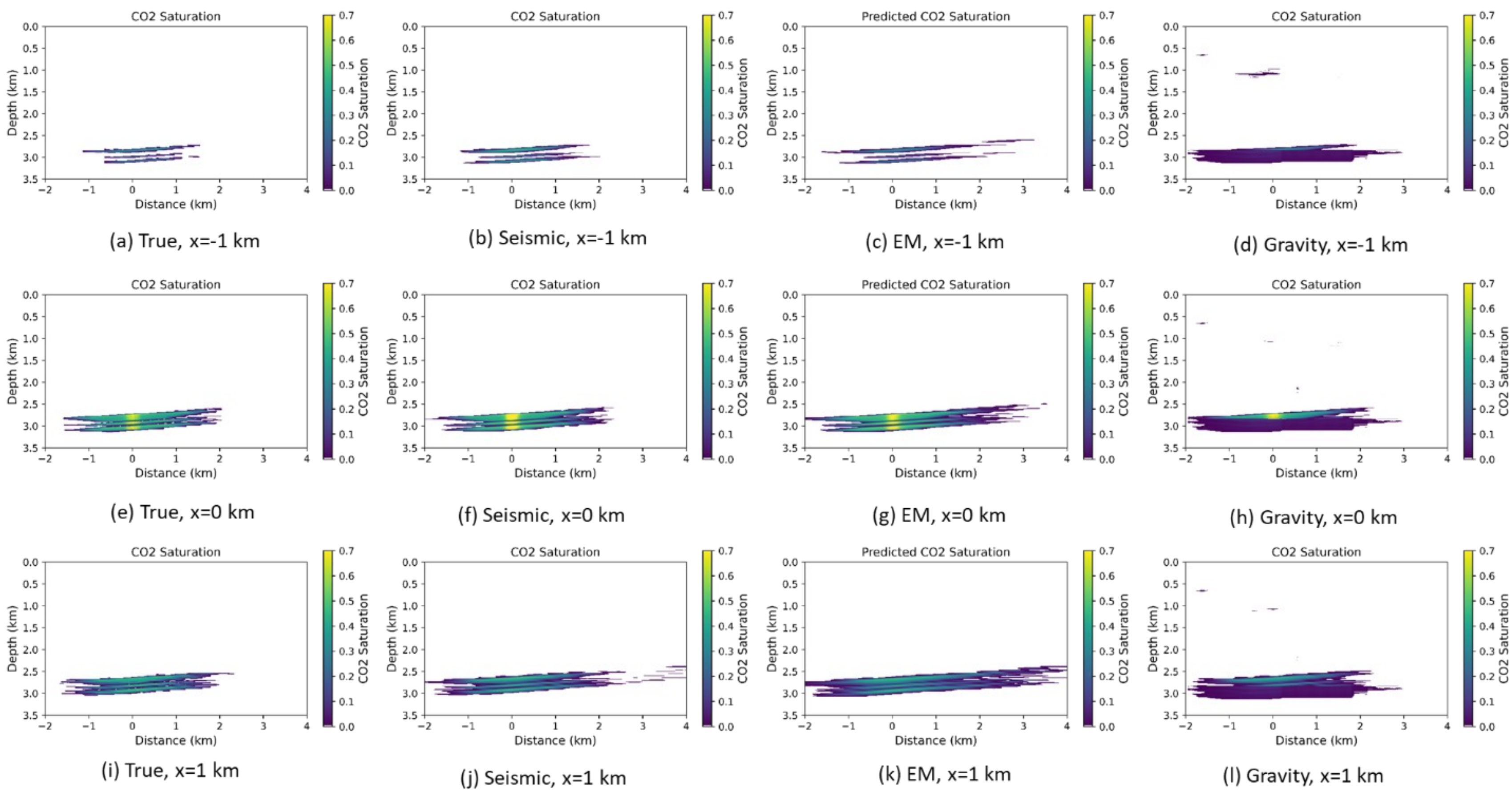


Figure 8. Comparison of the true Kimberlina CO_2 saturation model (the 1st column) and the mean values of CO_2 saturation images recovered from the three different DL single-physics imaging methodologies (the 2nd to 4th columns).

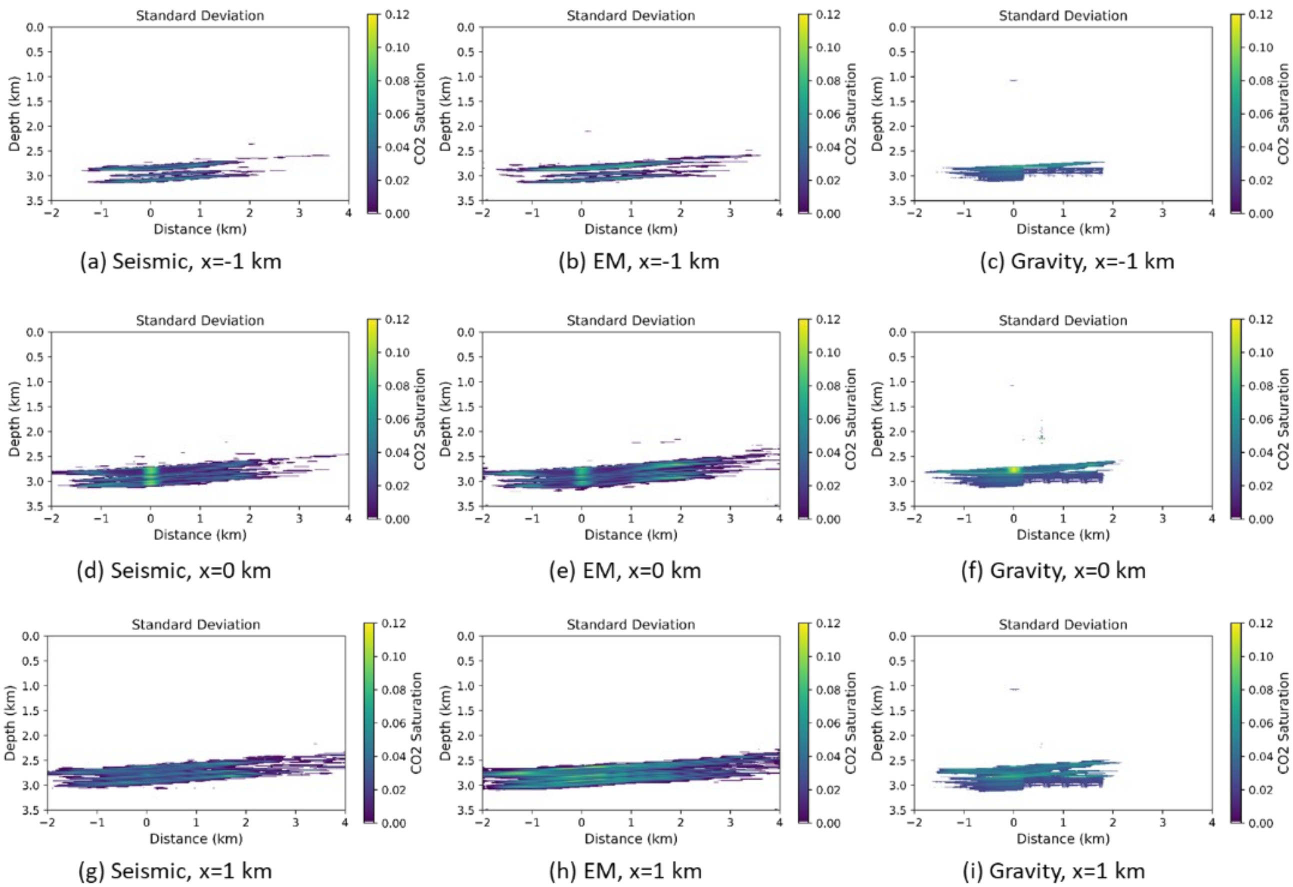


Figure 9. Cross-sectional views of uncertainties associated with the CO_2 saturation images from the three DL single-physics networks.

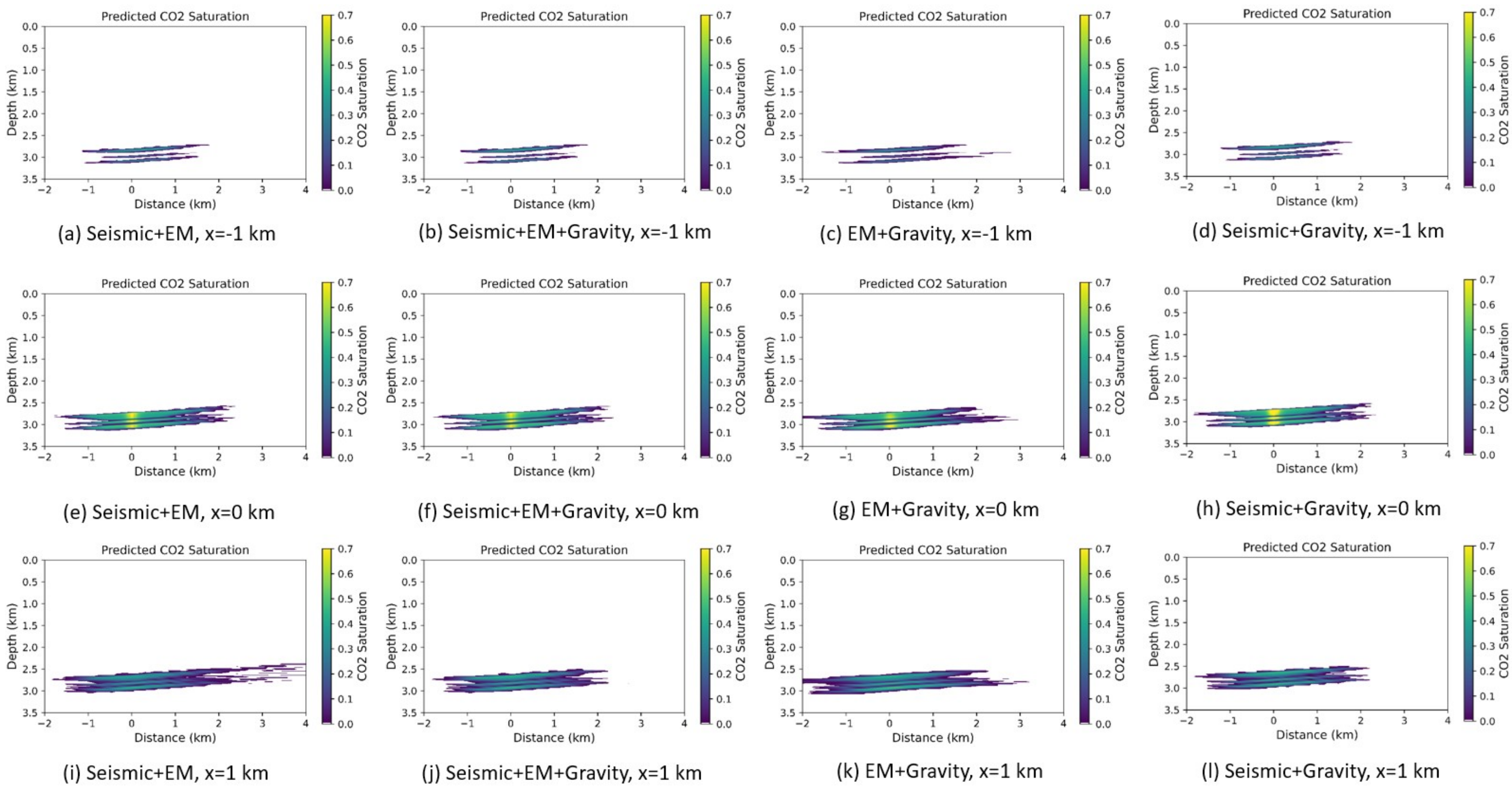


Figure 10. Cross-sectional views of the mean values of CO_2 saturation images recovered from the four DL multiphysics networks.

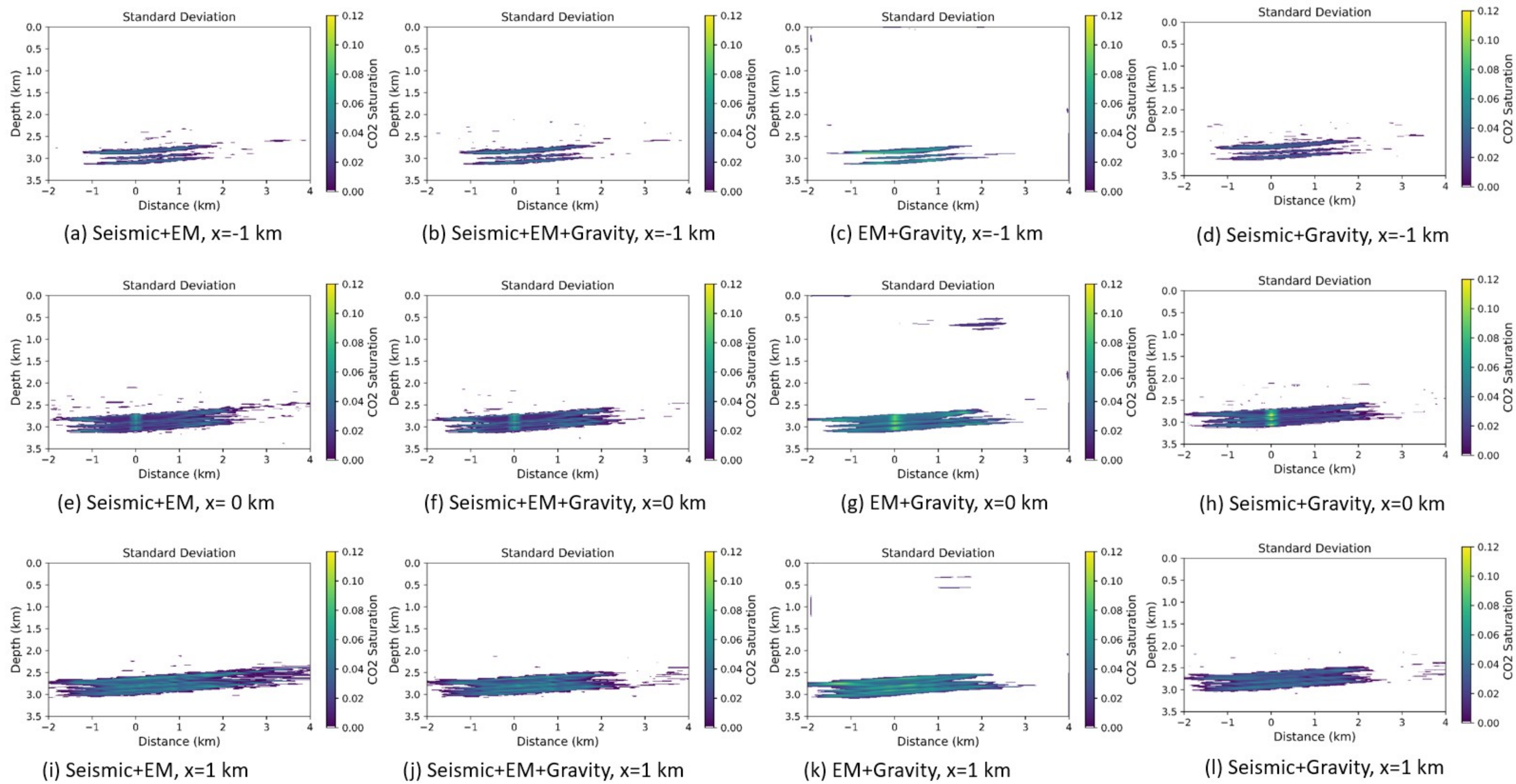


Figure 11. Cross-sectional views of uncertainties associated with the CO_2 saturation images from the four DL multiphysics networks.