

PNNL-34107

NRAP-Open-IAM Multisegmented Wellbore Reduced-Order Model

Improvement and Quality Assurance

March 2023

Seunghwan Baek
Diana H Bacon
Nicolas J Huerta

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, **makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical
Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: <http://www.ntis.gov>

NRAP-Open-IAM Multisegmented Wellbore Reduced-Order Model

Improvement and Quality Assurance

March 2023

Seunghwan Baek
Diana H Bacon
Nicolas J Huerta

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Abstract

The multisegmented wellbore model (MSW) semi-analytically estimates the amount of CO₂ and brine leakage from a leaking legacy well by segmenting it into intervals to simulate site-specific stratigraphic and hydrogeologic properties. The model is a component of the National Risk Assessment Partnership Open-Source Integrated Assessment Model (NRAP-Open-IAM), which was developed to perform risk assessment for geologic CO₂ storage.

The new wellbore leakage model, which uses deep learning networks for a caprock segment, was developed to enhance the analytical MSW. The model was trained and validated using a synthetic data set of Subsurface Transport Over Multiple Phases (STOMP) multiphase flow simulations from various geological, well attribute, and operational conditions to ensure its quality. The results demonstrate that the model is more accurate than the existing model in predicting the transport of two-phase fluids (brine and injected CO₂) through the well. This report provides a detailed explanation of the model development and quality assurance.

Acknowledgments

This work was completed as part of the U.S. Department of Energy's (DOE's) National Risk Assessment Partnership project, supported by the DOE Office of Fossil Energy's Cross-Cutting Research Program. The work was performed by Pacific Northwest National Laboratory under DOE contract number DE-AC05-76RL01830.

Acronyms and Abbreviations

DOE	U.S. Department of Energy
GCS	geological carbon storage
MSE	mean squared error
MSW	multisegmented wellbore
NRAP	National Risk Assessment Partnership
NRAP-Open-IAM	NRAP Open-Source Integrated Assessment Model
ROM	reduced-order model
STOMP	Subsurface Transport Over Multiple Phases
STOMP-CO ₂ e	STOMP with the CO ₂ operational mode

Contents

Abstract.....	iii
Acknowledgments.....	iv
Acronyms and Abbreviations.....	v
1.0 Introduction	1
2.0 Methodology.....	3
2.1 Data Generation	3
2.2 Data Processing	6
2.3 Model Training	7
3.0 Results and Discussion	9
3.1 Numerical Simulation Input Parameter Distributions	9
3.2 Model Validation and Quality Assurance.....	10
4.0 Conclusion	16
5.0 References.....	17
Appendix A – STOMP Input File for a Benchmark Problem	A.1
Appendix B – Machine Learning Model for Fluid Properties	B.1
Appendix C – Detailed Input Information for NRAP-Open-IAM Implementation.....	C.1

Figures

Figure 1.	Diagram of deep-learning-assisted multisegmented wellbore ROM development and quality assurance.....	3
Figure 2.	Numerical reservoir model.....	4
Figure 3.	CO ₂ injection rate profile.....	4
Figure 4.	Prediction workflow for CO ₂ leakage rate, CO ₂ saturation, and brine leakage rate.....	7
Figure 5.	Histograms of input parameters of the numerical simulation: (a) reservoir permeability, (b) reservoir porosity, (c) wellbore permeability, (d) wellbore length, (e) wellbore bottom depth, (f) wellbore radius, (g) temperature gradient, (h) salinity, and (i) injection rate.....	9
Figure 6.	Comparison of the histogram of wellbore top depth: (a) present study and (b) Baek et al. (2023).....	10
Figure 7.	Learning curves of the developed models: (a) CO ₂ classification, (b) CO ₂ regression, (c) brine classification, (d) brine large leakage rate regression, (e) brine small leakage rate regression, (f) brine negative leakage rate regression, and (g) saturation regression.....	13
Figure 8.	Error comparison of regression models: left column – present study, right column – analytical MSW (Baek et al. 2021). (a, b) CO ₂ , (c, d) brine large	

leakage, (e, f) brine small leakage, (g) brine negative leakage, and (h) histogram of absolute errors for brine negative leakage. 14

Figure 9. Comparison between model in NRAP-Open-IAM (red) and the standalone model (blue). Numerical modeling result (i.e., true data) is also given in yellow..... 15

Tables

Table 1. Numerical model parameters for the benchmark problem.5

Table 2. Testing results of CO₂ leakage classification model. 11

Table 3. Testing results of brine leakage classification model. 11

Table 4. Mean squared error (MSE) of leakage rate regression models. 11

Table 5. Mean squared error (MSE) of saturation regression model. 12

1.0 Introduction

To ensure safe geologic carbon storage (GCS), it is essential to have quantitative estimates of a site's long-term performance, taking into account subsurface hydrologic and geochemical processes, CO₂ injection, and associated uncertainties. An effective modeling approach is needed that can couple component models efficiently and allow for rapid computation (Pawar et al. 2015). In 2017, as part of the National Risk Assessment Partnership (NRAP) tool development effort, a set of 10 tools was released to estimate the risks associated with GCS (Vasykivska et al. 2021). The toolset, called the NRAP Open-source Integrated Assessment Model (NRAP-Open-IAM), considers uncertainties in storage reservoirs, leakage scenarios, and shallow groundwater impacts effectively through a stochastic framework. The set of component models used for rapid probabilistic risk assessments comprises analytical models or physics-based reduced-order models (ROMs) (Vasykivska et al. 2022).

At a CO₂ sequestration site, wells are commonly viewed as high-risk pathways for fluid leakage due to defects resulting from poor well completion or subsequent damage caused by chemical reactions or thermal or mechanical stresses (Gasda et al. 2008). To estimate the impacts of potential leakage from existing well penetrations to the reservoir, injection well, and monitoring wells during and after GCS operations, NRAP-Open-IAM incorporates several wellbore model components, including the open wellbore model (Bacon et al. 2021), cemented wellbore model (Jordan et al. 2015), and multisegmented wellbore (MSW) model (Baek et al. 2021). These models estimate fluid leakage along the wellbore from a storage reservoir, where CO₂ is injected, to overlying aquifers or the atmosphere.

The MSW was originally proposed in Nordbotten et al. (2009) for risk assessment of field-scale GCS applications with multiple overlying aquifers (Celia et al. 2011). The model analyzes the leakage of CO₂ and brine through legacy wells, based on a 1D multiphase version of Darcy's law. The model assumes that the leak occurs along a compromised wellbore, such as discrete pathways created by residual drilling fluid, gas channels formed during primary cementing, and micro-annuli created by pressure and temperature cycling, and the leak path is composed of alternating well segments, each with a length and an effective permeability value for the flow path.

This segmented leak path allows for approximating vertical heterogeneity and considering varying permeability formations into which the fluid may leak (e.g., the primary seal or underground source of drinking water). However, the original model does not account for varying fluid properties over depth (Baek et al. 2021; Viswanathan et al. 2008), which can be significant for CO₂, given its drastic changes in properties (e.g., density and viscosity) with depth, pressure, and temperature conditions. Additionally, the original analytical approach models lateral leakage in aquifer layers using a simple empirical method, which can be inaccurate when the system becomes more complex (Baek et al. 2021).

The improved MSW ROM addresses these limitations and enhances model accuracy for more reliable risk assessments of GCS operations. Baek et al. (2023) used deep learning to develop a model for caprock segment wellbore leakage. They generated synthetic data sets using a full physics numerical reservoir simulator that considered varying fluid properties over depth. Their work included various geologic conditions, well attributes, and operation conditions to cover various scenarios used in the literature. The present work improves these efforts further by including data sets for shallower (less than 1,000 m deep) aquifers. To build on this approach, we developed a new deep learning model and confirmed its correct implementation into NRAP-

Open-IAM. This work provides readers with a better understanding of the deep-learning-assisted MSW model and its quality assurance.

2.0 Methodology

This section outlines the process of developing and evaluating a deep-learning-assisted wellbore model, as shown in Figure 1. The Subsurface Transport Over Multiple Phases (STOMP) code, a multi-phase flow simulator developed by Pacific Northwest National Laboratory (White et al. 2012), was used to generate a comprehensive data set covering a wide range of subsurface pressure and temperature conditions, well characteristics, and operations based on various literature sources (Appendix A). The input and output data from the simulations were processed and used to train and test the deep learning models, which were optimized using an automated hyperparameter optimization process. The models included multiple classification and regression models for CO₂ and brine leakage and CO₂ saturation. Separate testing data sets were used for model validation, and the performance of the new model was compared with the analytical MSW model. This method is similar to that of Baek et al. (2023).

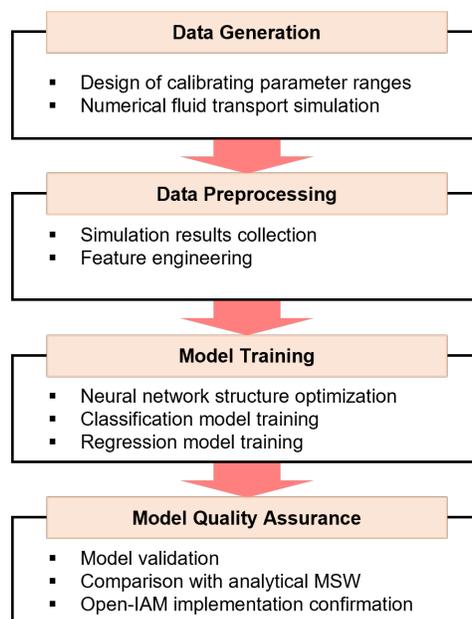


Figure 1. Diagram of deep-learning-assisted multisegmented wellbore ROM development and quality assurance.

2.1 Data Generation

STOMP with the CO₂ operational mode (STOMP-CO₂e) was used to create data for developing the deep learning model. A 2D axisymmetric reservoir model was employed, as shown in Figure 2. The model consists of a storage reservoir layer, a caprock sealing layer, and a leaky well. The model covers a lateral extent of 100 km from the well and is divided into 100 cells of logarithmically increasing size, from 0.075 to 11,527 m. The storage reservoir has five 5-m-thick cells; only the far-left column of cells in the caprock layer was activated, and the remaining part was considered impermeable and incompressible to save computational costs. The height of each cell in the wellbore is approximately 1 m, and the well length varies across realizations.

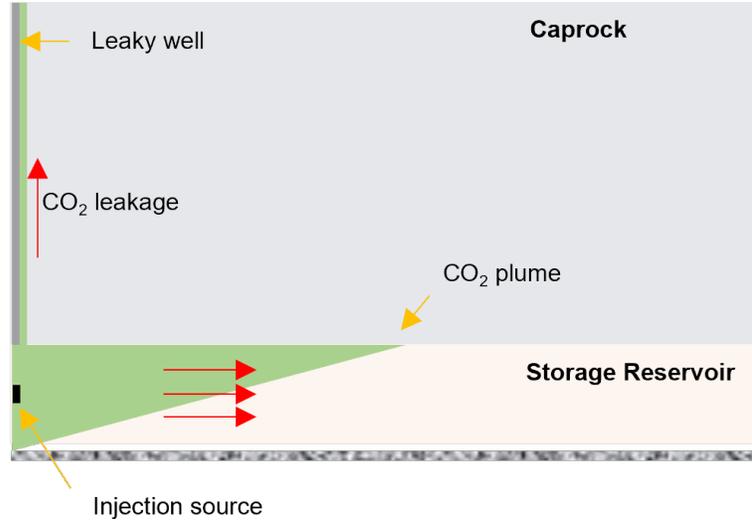


Figure 2. Numerical reservoir model.

For the deep learning model, dynamic data at the bottom of the well or the top layer of the storage reservoir is used as input data, and data at the top of the wellbore is used as output data. CO₂ is injected directly below the leaky well, and its injection profile is shown in Figure 3.

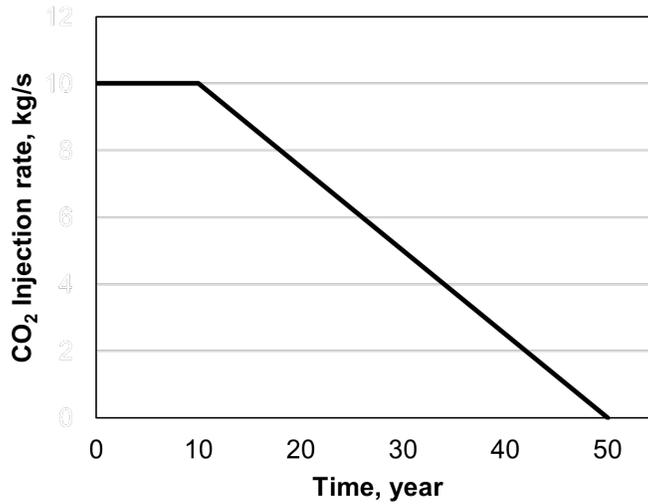


Figure 3. CO₂ injection rate profile.

Leakage along the well is assumed to follow multiphase Darcy flow as

$$q_{\alpha} = -\frac{k_{rel,\alpha}k_{well}}{\mu_{\alpha}}\left(\frac{\partial p_{\alpha}}{\partial z} + \rho_{\alpha}g\right) \tag{1}$$

where, for a fluid phase α (CO₂ or brine), q_{α} is the volumetric flux along the well, $k_{rel,\alpha}$ is the relative permeability of the well, k_{well} is the absolute permeability of the well, μ_{α} is viscosity, ∂p_{α} is the change in pressure across an increment in the vertical direction (i.e., ∂z), ρ_{α} is the fluid density, and g is the gravitational constant, 9.8 kg/m². To simplify modeling of complex fluid flow along the wellbore, a bulk effective permeability approach was used despite the fact that the

leakage path is not truly porous media and the Darcy equation [Eq. (1)] may not fully describe its physics (Carey 2018).

Hydrostatic pressure gradient (i.e., 9.8 MPa/km) is assumed, and a constant pressure boundary condition (i.e., Dirichlet) is applied for the right and top boundaries. The bottom of the model is sealed, leading to no vertical flow communication.

Linear relative permeability curves (i.e., x-type) were used for the entire domain for two-phase fluid flow, and capillary pressure was neglected. The model was initially saturated with brine while the residual brine saturation – the fraction of irreducible volume occupied by brine – was set to zero. Rock pore volume compressibility was set to 1.0×10^{-12} 1/Pa, and it is assumed that its properties are isotropic and homogeneous. Brine properties (i.e., density, viscosity) were calculated using the American Society of Mechanical Engineers steam table formulations (Meyer et al. 1993), and CO₂ properties were computed using the equation of state for CO₂ developed by Span and Wagner (1996) and the formulation of Fenghour et al. (1998). The properties are calculated as a function of pressure, temperature, and salinity internally during simulation. Initial salinity varies over realization, and no gradient of salinity was considered. All processes were considered isothermal, and mutual dissolution between different phases was neglected.

Nine parameters were varied, with minimum and maximum bounds, to generate a set of representative transient and steady state reservoir behavior and leakage scenarios based on the literature (Appendix A) with varying CO₂ injection scenarios (Table 1) – the constant injection rate for the first 10 years was varied. Five-thousand realizations of combinations of the parameters were sampled using uniform distributions. This study intends to expand the range of the top depth of the wellbore or caprock layer up to a shallower depth than used in the previous study, and the wellbore bottom depth and length were controlled.

Of the 5,000 STOMP simulation results, 4,053 cases were collected and used for deep learning model development; the remaining 947 simulations were removed because they failed due to convergence issues or took longer than 72 hours of simulation time on a single core.

Table 1. Numerical model parameters for the benchmark problem.

Parameter	Units	Min	Max
Reservoir permeability	m ²	1×10^{-15}	1×10^{-11}
Reservoir porosity	-	0.05	0.45
Wellbore permeability	m ²	1×10^{-16}	1×10^{-12}
Wellbore length	m	30	2,000
Wellbore bottom depth	m	776	4,970
Wellbore effective radius ^(a)	m	0.01	0.05
Temperature gradient	C°/km	18	32
Salinity	g/kg	1	25
CO ₂ injection rate	Kg/s	1.8	108

(a) The cell size is fixed at 0.075 m for the wellbore, but the porosity varies. The effective radius of the wellbore is calculated based on the cell size and porosity.

2.2 Data Processing

Based on the input parameters of the simulation (Table 1), 33 feature parameters were selected, consisting of:

- 5 well attributes: bottom depth, top depth, length, permeability, radius.
- 11 fluid properties: bottom and top well CO₂ density, bottom and top well brine density, bottom and top well CO₂ viscosity, bottom and top well brine viscosity, bottom well CO₂ saturation, bottom well brine saturation, salinity.
- 4 pressure-related parameters: bottom well pressure, top well pressure, pressure difference between top and bottom of the wellbore, bottom well pressure change with respect to the initial bottom well pressure. For top well pressure, the hydrostatic pressure was used.
- 3 temperature-related parameters: bottom well temperature, top well temperature, temperature gradient.
- 10 composite parameters: CO₂ driving force ratio [Eq. (2)], brine driving force ratio [Eq. (2)], CO₂ conductivity at the top and bottom wellbore [Eq. (3)], brine conductivity at the top and bottom wellbore [Eq. (3)], ratio of CO₂ conductivity at the top wellbore to that of the bottom wellbore [Eq. (4)], ratio of brine conductivity at the top wellbore to that of the bottom wellbore [Eq. (4)], kinematic viscosity ratio at the top and bottom wellbore [Eq. (5)].

The last 10 parameters are defined as follows:

$$\Gamma_{ph} = \frac{(P_{well,bot} - P_{well,top})}{0.5 \cdot (\rho_{ph,bot} + \rho_{ph,top})gh} \quad (2)$$

where Γ , defined as a driving force ratio, is a ratio of pressure force to the gravitational force; $P_{well,bot}$ and $P_{well,top}$ are pressure (Pa) at the bottom and top of the well, respectively; $\rho_{ph,bot}$ and $\rho_{ph,top}$ are CO₂ density (kg/m³) at the bottom and top of the well, respectively. Subscript ph is *CO₂* or *brine*; g is gravity constant (kg/m³); and h is well length (m).

$$\eta_{ph_2,i} = \frac{k_{well}\rho_{ph_2,i}g}{\mu_{ph,i}} \quad (3)$$

$$\kappa = \eta_{CO_2,top} / \eta_{CO_2,bot} = \left(\frac{\rho_{CO_2,top}}{\mu_{CO_2,top}} / \frac{\rho_{CO_2,bot}}{\mu_{CO_2,bot}} \right) \quad (4)$$

where $\eta_{ph,i}$ is phase conductivity (m/s), defined after hydraulic conductivity (Nordbotten and Celia 2012). Subscript ph is *CO₂* or *brine* and i is *bot* or *top*, indicating bottom or top of the well; k_{well} is well permeability (m²); $\mu_{ph,i}$ is phase viscosity (Pa-sec); and the κ conductivity ratio is a ratio of phase conductivity at the top of the well to that at the bottom of the well.

$$v_i = v_{brine,i}/v_{CO_2,i} = \left(\frac{\mu_{brine,i}}{\rho_{brine,i}} / \frac{\mu_{CO_2,i}}{\rho_{CO_2,i}} \right) \tag{5}$$

where v_i is a ratio of brine kinematic viscosity to CO₂ kinematic viscosity. Subscript i is *bot* or *top*, indicating the bottom or top of the well. $\rho_{brine,i}$ and $\mu_{brine,i}$ are brine density (kg/m³) and brine viscosity (Pa-sec), respectively.

The target data set for the deep learning model consists of the flow rates of CO₂ and brine, and CO₂ saturation at the top of the well for every year, resulting in a total of 199,551 data points for each. Less than 1% of the CO₂ saturation at the bottom wellbore and less than 10⁻⁷ kg/s of CO₂ were considered zero or no leakage, respectively. The data set was randomly divided into three subgroups for training, validation, and testing with a ratio of 0.7:0.2:0.1. The feature data set includes leaking rates of CO₂, well permeability, CO₂ conductivity, and kinematic viscosity ratio at the top and bottom wellbore, all of which were transformed into log-scale to improve modeling results (Chen et al. 2019). The feature data was also normalized between zero and one using the MinMaxScaler function from the scikit-learn Python library (v.1.0.1) (Pedregosa et al. 2011).

2.3 Model Training

This section describes the development of multiple sub-deep learning models to predict CO₂ and brine leakage rates in a wellbore. To predict CO₂ leakage, a classification model is used first to determine if there is any leakage through the leaky wellbore across the caprock layer. If there is leakage, then regression models are used to predict the CO₂ leakage rate and saturation at the top of the well. For brine leakage, a classification model is used to predict the flow rate, which can be large, small, or negative. The pressure gradient required for brine to leak upward is higher compared to CO₂, and in some cases, the gravity force can cause the brine to flow downward, resulting in a negative flow rate (Baek et al. 2023). The corresponding regression model predicts the brine leakage rate once the type of leakage is identified.

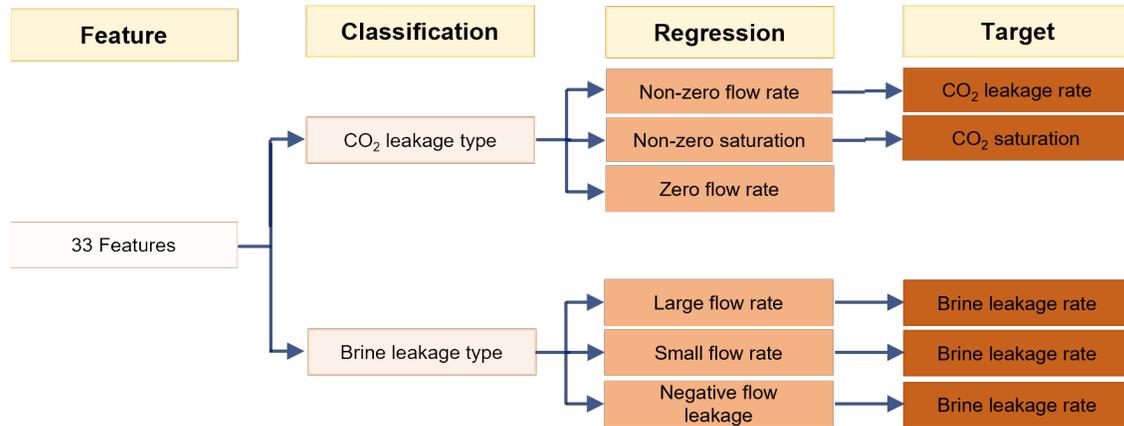


Figure 4. Prediction workflow for CO₂ leakage rate, CO₂ saturation, and brine leakage rate.

The developed models' neural networks were optimized using the Python library AutoKeras (v.1.0.16) (Jin et al. 2019), which uses Bayesian optimization to tune hyperparameters such as the number of dense layers, activation functions, and dropout values for deep learning model structures. To optimize the models, 100 neural network structures were searched with a batch

size that varied between 256 and 1,024. The models were trained for 150 epochs with early stopping if there was no improvement after 50 epochs. The data set was shuffled after every epoch.

Binary cross-entropy, categorical cross-entropy, and mean squared error (MSE) were used as accuracy metrics for binary classification, multi classification, and regression, respectively, which are calculated as

$$H_p = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (6)$$

where H_p is binary cross-entropy, y is the label (1 or 0 for binary classification), $p(y)$ is the predicted probability that the label y is true, and N is the number of data points considered.

$$CE = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K y_{i,k} \log p_{i,k} \quad (7)$$

where CE is categorical cross-entropy, which expands Eq. (6) for multiple labels, K .

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_{i,true} - y_{i,pred})^2 \quad (8)$$

where MSE is mean squared error, $y_{i,true}$ is the reference value, $y_{i,pred}$ is the value predicted by the regression model, and N is the number of data points considered.

Precision, recall, and F1 score were given to measure the quality of the classification models:

$$P = \frac{T_p}{T_p + F_p} \quad (9)$$

where P is precision, T_p is the number of true positives, and F_p is the number of false positives.

$$R = \frac{T_p}{T_p + F_n} \quad (10)$$

where R is recall, T_p is the number of true positives, and F_n is the number of false negatives.

$$F1 = 2 \frac{P \times R}{P + R} \quad (11)$$

where $F1$ is the F1 score, which is defined as the harmonic mean of precision and recall.

3.0 Results and Discussion

This section presents ranges of the numerical simulation input parameters used to develop the new model and predict its performance. Comparison with the analytical MSW model (Baek et al. 2021) shows improvement made in the new model.

3.1 Numerical Simulation Input Parameter Distributions

The necessary data for deep learning model development was generated using the numerical reservoir simulations. For the simulations, the nine parameters given in Figure 5 were sampled, and based on the samples and simulation results, the aforementioned 33 features were prepared. Figure 5 shows the distributions of the simulation input parameters for the cases that ran successfully and were collected for the wellbore model development.

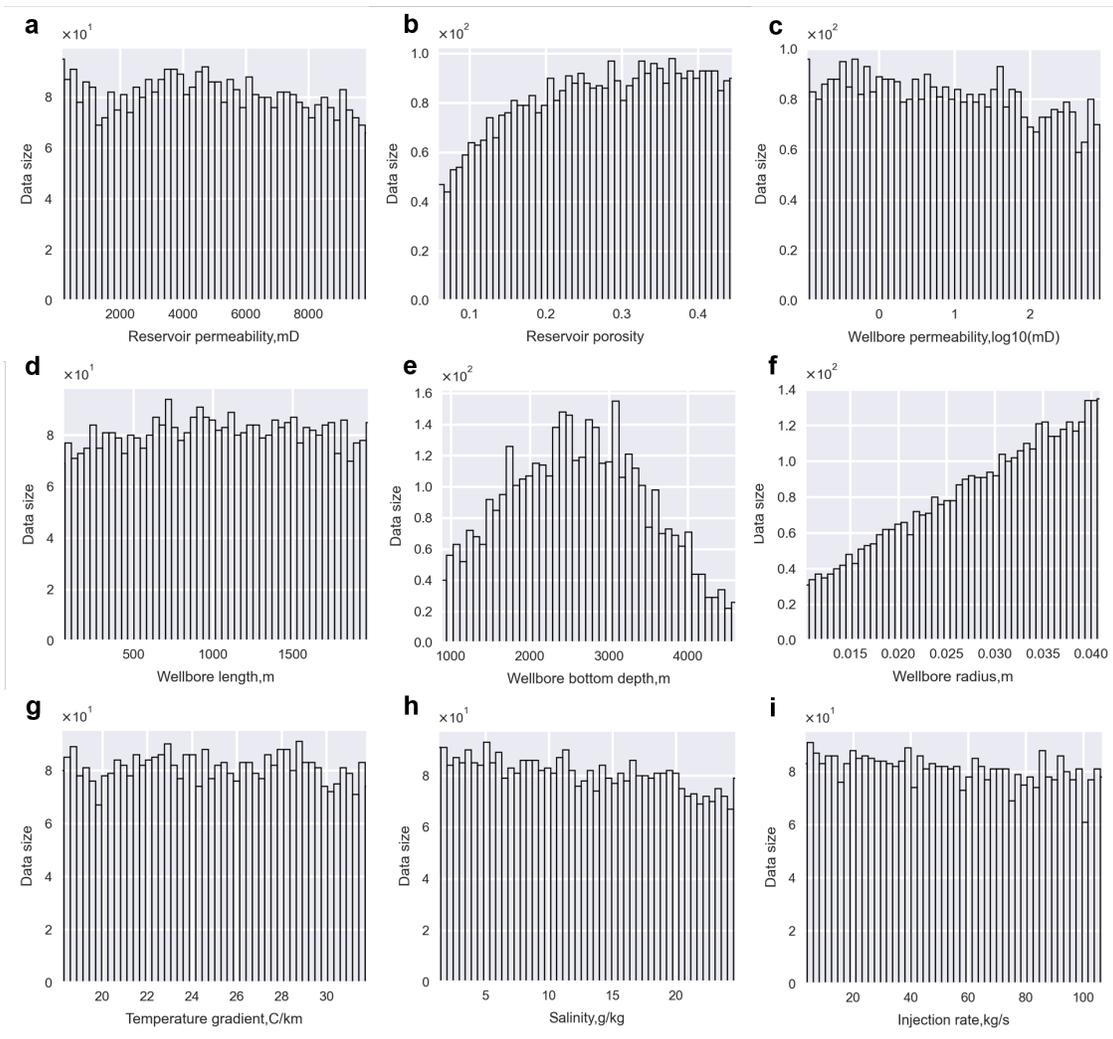


Figure 5. Histograms of input parameters of the numerical simulation: (a) reservoir permeability, (b) reservoir porosity, (c) wellbore permeability, (d) wellbore length, (e) wellbore bottom depth, (f) wellbore radius, (g) temperature gradient, (h) salinity, and (i) injection rate.

Input parameters were sampled uniformly as intended except for wellbore bottom depth and wellbore radius. Filtering and additional calculation led to unproportional distribution of these two parameters. Wellbore bottom depth was initially sampled uniformly, and the cases where the top depth of the wellbore – which is calculated based on wellbore bottom depth and wellbore length – was shallower than 800 m were removed. For wellbore radius, porosity of the cells for wellbore, not the radius, was sampled for easier implementation in the numerical simulation, and the effective wellbore radius, r_w , was calculated as

$$r_w = \sqrt{\frac{(0.075)^2 \phi}{\pi}} \quad (12)$$

where ϕ is porosity and 0.075 m is the dimension size of the square-shaped cells of the wellbore. This conversion caused the imbalance as shown in Figure 5f.

The purpose of this study is to expand the applicable ranges of the previous model (Baek et al. 2023) to include a shallower wellbore top depth for improved applicability for shallow underground sources of drinking water. As seen as in Figure 6, while Baek et al. (2023) did not include the wellbore top depth shallower than 1,000 m (Figure 6b), the present study (Figure 6a) includes data between 25 and 1,000 m. Wellbore top depth was not sampled but was calculated using wellbore bottom depth and wellbore length. More frequent simulation failures were observed as the top depth is shallower, and this is probably because of increased non-linearity due to the phase changes of CO₂.

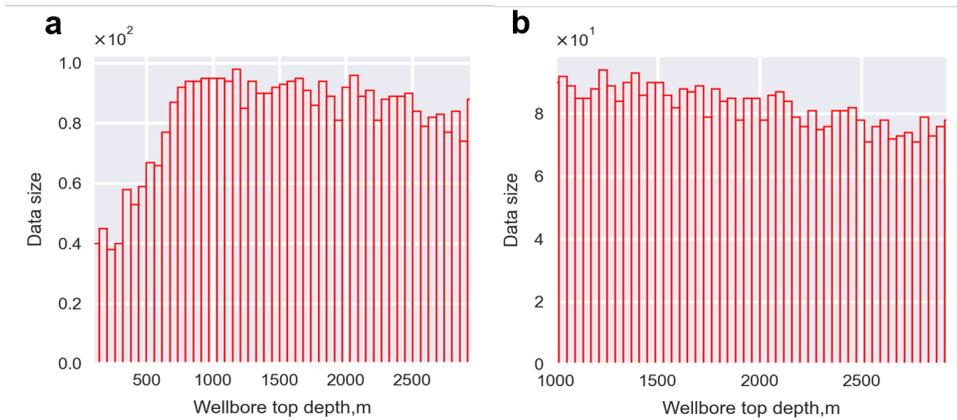


Figure 6. Comparison of the histogram of wellbore top depth: (a) present study and (b) Baek et al. (2023).

3.2 Model Validation and Quality Assurance

A total of seven deep learning models were developed in this study: two classification models for CO₂ and brine and five regression models for CO₂ leakage, CO₂ saturation, brine large leakage, brine small leakage, and brine negative leakage. This session discusses how the models were validated for quality assurance.

The CO₂ leakage classification model predicts whether there is leakage at the top of the wellbore, which is binary classification – true or false – using binary cross-entropy as a loss metric [Eq. (6)]. Ten percent of the data set, so-called testing data, was not included during

model training and instead was used for validation. Table 2 summarizes the model performance with precision [Eq. (9)], recall [Eq. (10)] and F1 score [Eq. (11)]. For no leakage and leakage cases, the numbers of the data set used were 11,834 and 8,122, respectively, and overall the performance is good.

Table 2. Testing results of CO₂ leakage classification model.

Target	Precision	Recall	F1 score	Data size
No leakage	0.99	0.99	0.99	11,834
Leakage	0.99	0.99	0.99	8,122

The brine leakage classification model predicts three classes: negative leakage, small leakage, and large leakage at the top of the wellbore, which is multi-classification using categorical cross entropy [Eq. (8)] as a loss metric. Precision [Eq. (9)], recall [Eq. (10)], and F1 score [Eq. (11)] are summarized in Table 3, which shows that overall model performance is good.

Table 3. Testing results of brine leakage classification model.

Target	Precision	Recall	F1-score	Data size
Negative leakage rate	0.93	0.93	0.93	1,241
Small leakage rate	0.99	0.96	0.98	6,058
Large leakage rate	0.97	0.99	0.98	12,657

For leakage rate regression models, MSEs [Eq. (8)] for training, validation, and testing are provided in Table 4. The MSEs for training and validation are at the 150th epoch. Negative brine leakage rate was multiplied by -1 to calculate and show it in the same format as the others. Testing results from the unused data set for modeling training show small MSEs comparable with those of training and validation.

Table 4. Mean squared error (MSE) of leakage rate regression models.

Data Set	CO ₂		Brine	
	Leakage ^(a) (log ₁₀ (kg/s)) ^{0.5}	Negative ^(b) (log ₁₀ (-kg/s)) ^{0.5}	Small ^(c) (log ₁₀ (kg/s)) ^{0.5}	Large ^(d) (log ₁₀ (kg/s)) ^{0.5}
Training	1.3216e-04	2.9932e-04	3.3945e-04	5.4337e-04
Validation	2.0539e-04	6.5208e-04	9.8773e-04	7.1911e-04
Testing	1.9513e-04	6.8580e-04	6.1010e-04	6.9066e-04

- (a) The values of data for training, validation, and testing are 56,282, 16,078, and 8,041, respectively.
 (b) The values of data for training, validation, and testing are 8,829, 2,521, and 1,262, respectively.
 (c) The values of data for training, validation, and testing are 42,603, 12,172, and 6,087, respectively.
 (d) The values of data for training, validation, and testing are 88,257, 25,212, and 12,608, respectively.

For saturation regression model, Table 5 shows MSEs [Eq. (8)] for training, validation, and testing. MSEs at the 150th epoch are shown for training and validation. MSE results for the testing data are small and are comparable with those of training and validation.

Table 5. Mean squared error (MSE) of saturation regression model.

Data Set	CO ₂ saturation ^(a)
Training	0.0012
Validation	0.0013
Testing	0.0013

(a) The values of data for training, validation, and testing are 56,282, 16,078 and 8,041, respectively.

To understand the robustness of the models, the learning curves were recorded during model training and are shown for all models in Figure 7. Loss metrics for binary classification, multi classification, and regression are binary cross-entropy [Eq. (6)], categorical cross-entropy and MSE [Eq. (7)], respectively. Figure 7 shows that models are neither over-fitted nor under-fitted.

In Figure 8, the regression performance of the new wellbore leakage model and the analytical MSW for CO₂ leakage were compared using testing data and a unit plot. Non-leakage data was not included in the comparison. The dot indicates the leakage flow rate data at each time from various time series data, which were generated from simulations with different realizations.

Figure 8 allows quick visual inspection of how well the regression models predict. When the dots are well-aligned along the diagonal line, the model performance is better. The new deep learning models show good performance in the left column of Figure 8, and comparison to the results of the analytical MSW on the right column in Figure 8 clearly shows how much the prediction accuracy was improved in the new model. On the other hand, the prediction by the analytical MSW is poorer with brine leakage rate, and it turns out that it does not have a capability to predict the negative leakage rate (Figure 8d). So, the absolute errors between the true values from the numerical simulation data and the predicted values from the analytical MSW are displayed using a histogram instead of the unit plot.

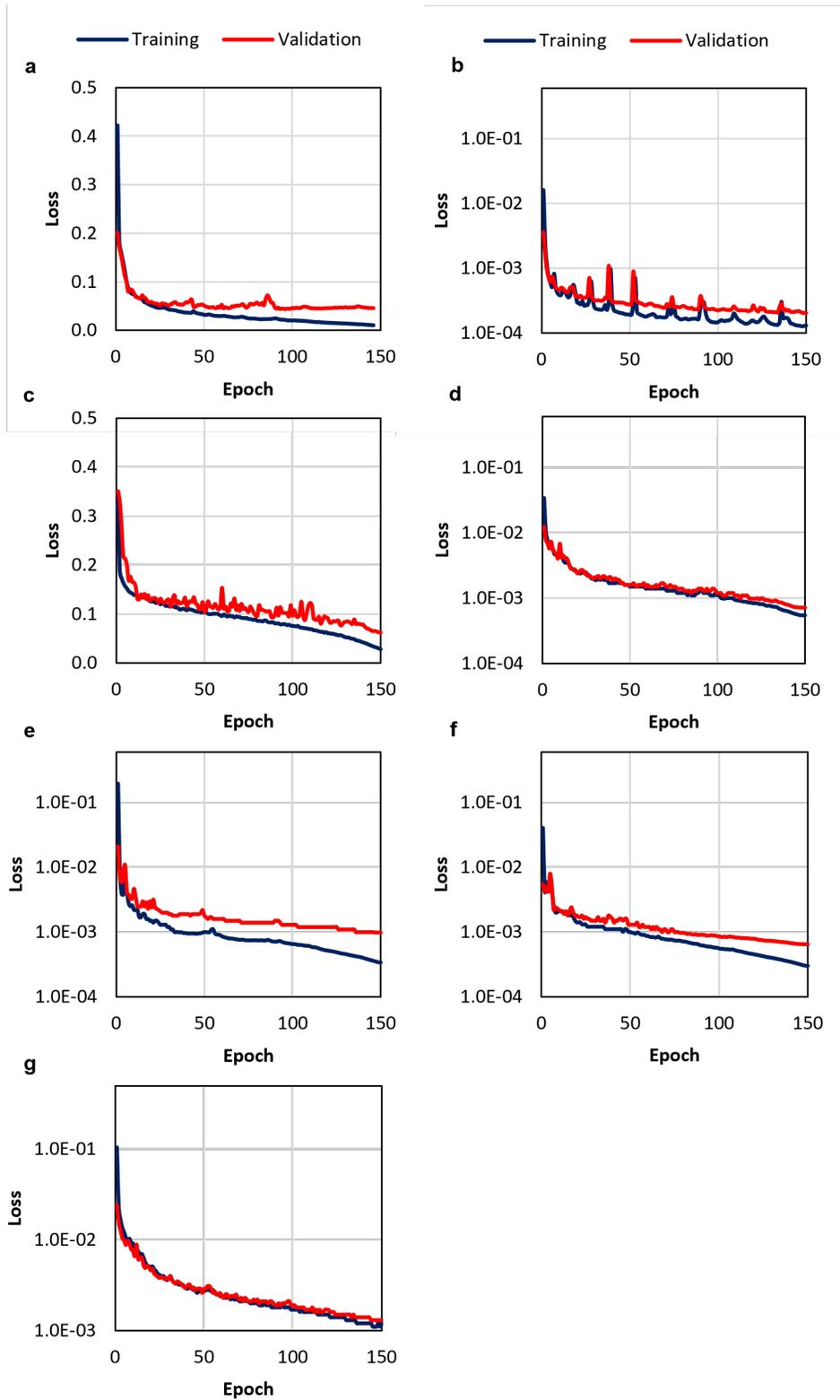


Figure 7. Learning curves of the developed models: (a) CO₂ classification, (b) CO₂ regression, (c) brine classification, (d) brine large leakage rate regression, (e) brine small leakage rate regression, (f) brine negative leakage rate regression, and (g) saturation regression.

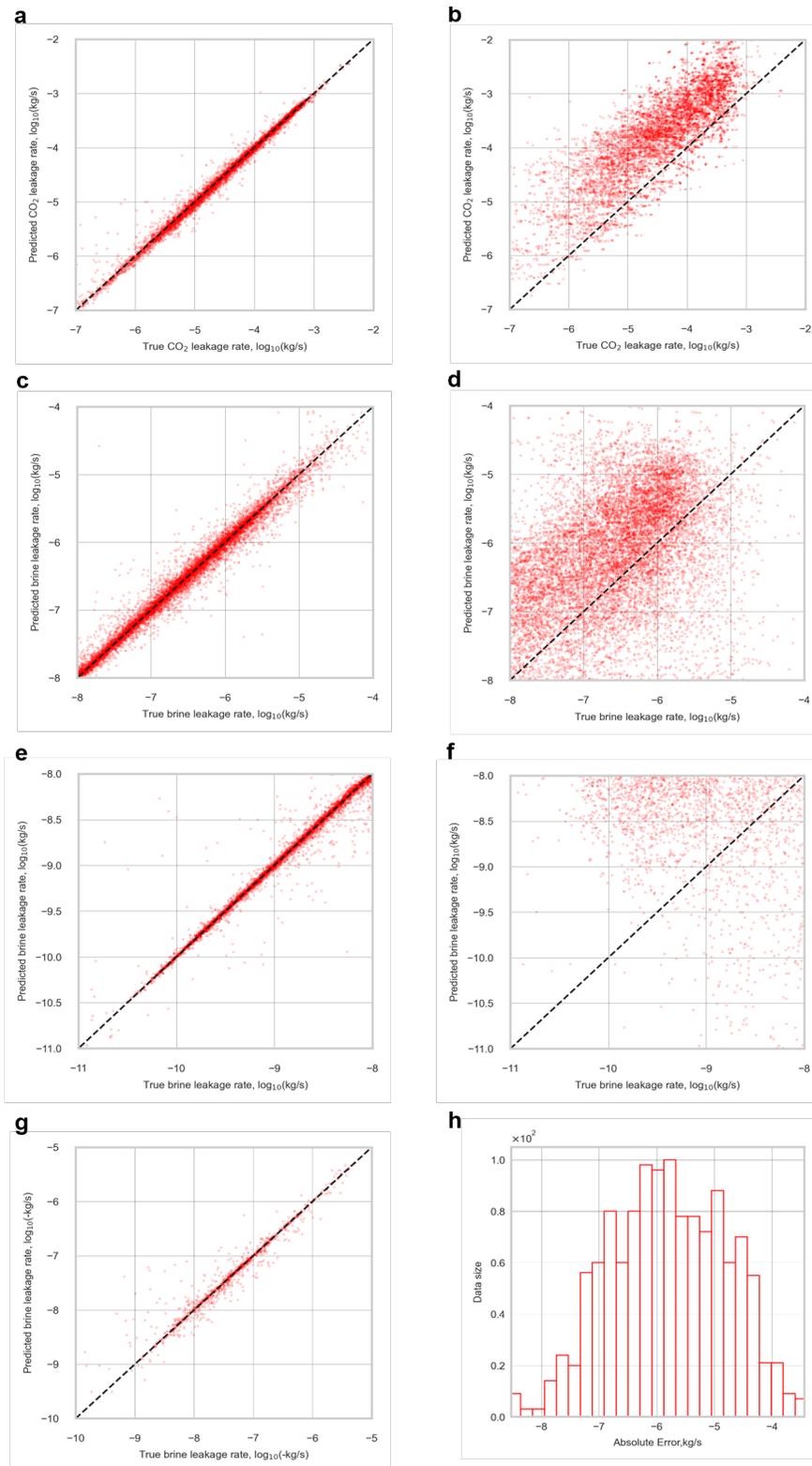


Figure 8. Error comparison of regression models: left column – present study, right column – analytical MSW (Baek et al. 2021). (a, b) CO₂, (c, d) brine large leakage, (e, f) brine small leakage, (g) brine negative leakage, and (h) histogram of absolute errors for brine negative leakage.

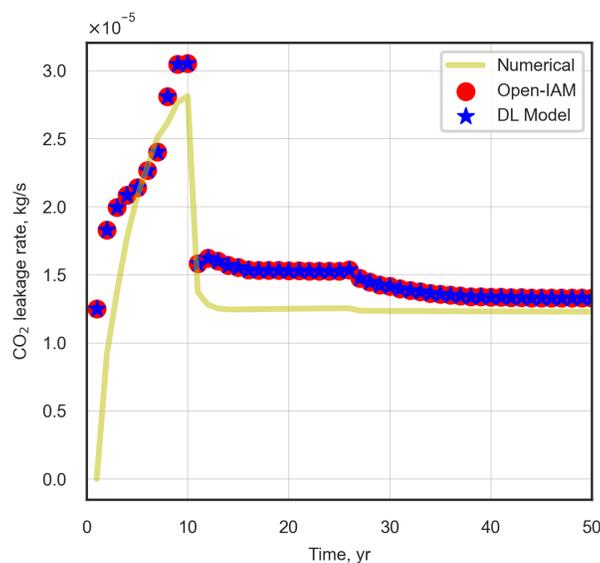


Figure 9. Comparison between model in NRAP-Open-IAM (red) and the standalone model (blue). Numerical modeling result (i.e., true data) is also given in yellow.

The developed wellbore model is implemented into the existing MSW model in NRAP-Open-IAM. The new model calculates the leakage rates of CO₂ and brine from the storage reservoir to the aquifer immediately above the reservoir along the leaky well path and CO₂ saturation at the bottom of the aquifer immediately above the reservoir, while other parts use the existing approach of the analytical MSW model (Baek et al. 2021). Deep learning models for the other parts will be developed in future. NRAP-Open-IAM provides the option to use either the analytical approach only, as in Baek et al. (2021), or a hybrid approach integrating both the deep learning model and analytical approach.

Figure 9 compares the CO₂ leakage rate calculation by both the wellbore model implemented into NRAP-Open-IAM and a standalone model (i.e., DL Model in figure), which was developed separately before implementing into NRAP-Open-IAM. The true value calculated by the numerical simulation (yellow line in Figure 9) is also provided. The NRAP-Open-IAM model uses dynamic pressure, CO₂ saturation, and salinity from the separate numerical reservoir model. For fluid property calculation, separate machine learning models were developed for both CO₂ and brine, which predict density and viscosity as a function of pressure and temperature for CO₂ and pressure, temperature, and salinity for brine, respectively (Appendix B). The comparison in Figure 9 confirms that computational implementation is done correctly. The detailed dynamic and static input data is included in Appendix C.

4.0 Conclusion

This report presents the improvement of an MSW ROM in NRAP-Open-IAM (Baek et al. 2021). A deep learning model was created to predict the leakage rates of CO₂ and brine, and CO₂ saturation, at the top well for a bottom caprock segment using a large synthetic data set of STOMP multiphase flow simulations. The new model expands the applicable range of the top depth of the wellbore or bottom depths of the overlying aquifer from 1,000 m to 25 m, improving on the work by Baek et al. (2023).

The wellbore ROM was effective in capturing the underlying complexity in multiphase flows along leaky paths for a wide range of geological conditions, well attributes, and operation conditions. The developed model demonstrated good predictive performance and significant improvement over the analytical MSW ROM (Baek et al. 2021). The new model has been correctly implemented into NRAP-Open-IAM for a bottom caprock segment and can be useful for practical risk assessment in GCS operations when coupled with other component models.

The results suggest that the new model could be used to simulate leakage through more complex leaky paths with multiple intervals and different leak levels by linking multiple segments of the model. This approach would allow for the simulation of more realistic wellbore leakage risks with greater flexibility in system design. This work represents the first step toward developing the individual segment model.

5.0 References

- Bacon DH, L Pan, and CM Oldenburg. 2021. *NRAP-Open-IAM: Open Wellbore Component v2.0*. PNNL-31543. Pacific Northwest National Laboratory: Richland, WA. <https://doi.org/10.2172/1825929>
- Baek S, DH Bacon, and NJ Huerta. 2021. *NRAP-Open-IAM Multisegmented Wellbore Reduced-Order Model*. PNNL-32364. Pacific Northwest National Laboratory: Richland, WA. <https://doi.org/10.2172/1840652>
- Baek S, DH Bacon, and NJ Huerta. 2023. “Enabling site-specific well leakage risk estimation during geologic carbon sequestration using a modular deep-learning-based wellbore leakage model.” *International Journal of Greenhouse Gas Control* 126:103903. <https://doi.org/10.1016/j.ijggc.2023.103903>.
- Birkholzer JT and Q Zhou. 2009. “Basin-scale hydrogeologic impacts of CO₂ storage: Capacity and regulatory implications.” *International Journal of Greenhouse Gas Control* 3(6):745-756. <https://doi.org/10.1016/j.ijggc.2009.07.002>.
- Carey JW. 2018. *Probability Distributions for Effective Permeability of Potentially Leaking Wells at CO₂ Sequestration Sites*. NRAP-TRS-III-021-2017. NRAP Technical Report Series; U.S. Department of Energy, National Energy Technology Laboratory: Morgantown, WV. <https://doi.org/10.2172/1438142>
- Celia MA, JM Nordbotten, B Court, M Dobossy, and S Bachu. 2011. “Field-scale application of a semi-analytical model for estimation of CO₂ and brine leakage along old wells.” *International Journal of Greenhouse Gas Control* 5(2):257-269. <https://doi.org/10.1016/j.ijggc.2010.10.005>.
- Chen B, DR Harp, RJ Pawar, PH Stauffer, HS Viswanathan, and RS Middleton. 2019. “Frankenstein’s ROMster: Avoiding pitfalls of reduced-order model development.” *International Journal of Greenhouse Gas Control* 93. <https://doi.org/10.1016/j.ijggc.2019.102892>.
- Doughty C. 2010. “Investigation of CO₂ Plume Behavior for a Large-Scale Pilot Test of Geologic Carbon Storage in a Saline Formation.” *Transport in Porous Media* 82(1):49-76. <https://doi.org/10.1007/s11242-009-9396-z>.
- Fenghour A, WA Wakeham, and V Vesovic. 1998. “The Viscosity of Carbon Dioxide.” *Journal of Physical and Chemical Reference Data* 27(1):31-44. <https://doi.org/10.1063/1.556013>.
- Gasda SE, MA Celia, and JM Nordbotten. 2008. “Upslope plume migration and implications for geological CO₂ sequestration in deep, saline aquifers.” *The IES Journal Part A: Civil & Structural Engineering* 1(1):2-16. <https://doi.org/10.1080/19373260701620154>.
- Jin H, Q Song, and X Hu. 2019. *Auto-keras: An efficient neural architecture search system*. ACM.
- Jordan AB, PH Stauffer, D Harp, JW Carey, and RJ Pawar. 2015. “A response surface model to predict CO₂ and brine leakage along cemented wellbores.” *International Journal of Greenhouse Gas Control* 33:27-39. <https://doi.org/10.1016/j.ijggc.2014.12.002>.
- Jordan P and C Doughty. 2009. “Sensitivity of CO₂ migration estimation on reservoir temperature and pressure uncertainty.” *Energy Procedia* 1(1):2825-2832. <https://doi.org/10.1016/j.egypro.2009.02.055>.
- Kang M and RB Jackson. 2016. “Salinity of deep groundwater in California: Water quantity, quality, and protection.” *Proceedings of the National Academy of Sciences* 113(28):7768-7773. <https://doi.org/10.1073/pnas.1600400113>.
- Meyer CA, RB McClintock, GJ Silvestri, and RC Spencer. 1993. *ASME steam tables*. The American Society of Mechanical Engineers, New York.

- Nguyen BN, Z Hou, DH Bacon, GV Last, and MD White. 2017. "Analysis of a Complex Faulted CO₂ Reservoir Using a Three-dimensional Hydro-geochemical-Mechanical Approach." *Energy Procedia* 114:3496-3506. <https://doi.org/10.1016/j.egypro.2017.03.1479>.
- Nordbotten JM and MA Celia. 2012. *Geological Storage of CO₂: Modeling Approaches for Large-Scale Simulation*. Wiley, Hoboken, NJ.
- Nordbotten JM, D Kavetski, MA Celia, and S Bachu. 2009. "Model for CO₂ Leakage Including Multiple Geological Layers and Multiple Leaky Wells." *Environmental Science & Technology* 43(3):743-749. <https://doi.org/10.1021/es801135v>.
- Onishi T, MC Nguyen, JW Carey, B Will, W Zaluski, DW Bowen, BC Devault, A Duguid, Q Zhou, SH Fairweather, LH Spangler, and PH Stauffer. 2019. "Potential CO₂ and brine leakage through wellbore pathways for geologic CO₂ sequestration using the National Risk Assessment Partnership tools: Application to the Big Sky Regional Partnership." *International Journal of Greenhouse Gas Control* 81:44-65. <https://doi.org/10.1016/j.ijggc.2018.12.002>.
- Pawar RJ, GS Bromhal, JW Carey, W Foxall, A Korre, PS Ringrose, O Tucker, MN Watson, and JA White. 2015. "Recent advances in risk assessment and risk management of geologic CO₂ storage." *International Journal of Greenhouse Gas Control* 40:292-311. <https://doi.org/10.1016/j.ijggc.2015.06.014>.
- Pedregosa F, G Varoquaux, A Gramfort, V Michel, B Thirion, O Grisel, M Blondel, P Prettenhofer, R Weiss, V Dubourg, J Vanderplas, A Passos, D Cournapeau, M Brucher, M Perrot, and É Duchesnay. 2011. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research* 12:2825–2830.
- Span R and W Wagner. 1996. "A New Equation of State for Carbon Dioxide Covering the Fluid Region from the Triple - Point Temperature to 1100 K at Pressures up to 800 MPa." *Journal of Physical and Chemical Reference Data* 25(6):1509-1596. <https://doi.org/10.1063/1.555991>.
- USGS. 2013. *National assessment of geologic carbon dioxide storage resources: data*. Reston, VA. <https://doi.org/10.3133/ds774>
- Vasylykivska V, R Dilmore, G Lackey, Y Zhang, S King, D Bacon, B Chen, K Mansoor, and D Harp. 2021. "NRAP-open-IAM: A flexible open-source integrated-assessment-model for geologic carbon storage risk assessment and management." *Environmental Modelling & Software* 143:105114. <https://doi.org/10.1016/j.envsoft.2021.105114>.
- Vasylykivska V, S King, D Bacon, D Harp, E Lindner, S Baek, E Keating, Y Yang, Y Zhang, B Chen, and K Mansoor 2022. NRAP-Open-IAM User's Guide, Release alpha 2.6.1-22.09.30, 2/8/2023.
- Viswanathan HS, RJ Pawar, PH Stauffer, JP Kaszuba, JW Carey, SC Olsen, GN Keating, D Kavetski, and GD Guthrie. 2008. "Development of a Hybrid Process and System Model for the Assessment of Wellbore Leakage at a Geologic CO₂ Sequestration Site." *Environmental Science & Technology* 42(19):7280-7286. <https://doi.org/10.1021/es800417x>.
- White MD, DH Bacon, BP McGrail, DJ Watson, SK White, and ZF Zhang. 2012. *STOMP Subsurface Transport Over Multiple Phases: STOMP-CO₂ and STOMP-CO₂e Guide: Version 1.0*. PNNL-21268. Pacific Northwest National Laboratory: Richland, WA. <https://doi.org/10.2172/1059044>
- William Carey J, R Svec, R Grigg, J Zhang, and W Crow. 2010. "Experimental investigation of wellbore integrity and CO₂-brine flow along the casing-cement microannulus." *International Journal of Greenhouse Gas Control* 4(2):272-282. <https://doi.org/10.1016/j.ijggc.2009.09.018>.
- Yonkofski C, G Tartakovsky, N Huerta, and A Wentworth. 2019. "Risk-based monitoring designs for detecting CO₂ leakage through abandoned wellbores: An application of NRAP's

- WLAT and DREAM tools.” *International Journal of Greenhouse Gas Control* 91:102807. <https://doi.org/10.1016/j.ijggc.2019.102807>.
- Zhang L, R Dilmore, N Huerta, Y Soong, V Vasyukivska, A Namhata, Y Wang, and X Li. 2018. “Application of a new reduced-complexity assessment tool to estimate CO₂ and brine leakage from reservoir and above-zone monitoring interval (AZMI) through an abandoned well under geologic carbon storage conditions.” *Greenhouse Gases: Science and Technology* 8(5):839-853. <https://doi.org/10.1002/ghg.1813>.

Appendix A – STOMP Input File for a Benchmark Problem

Table A.1. Reference for parameter ranges of GCS systems.

Parameters	References
Wellbore permeability	(Birkholzer and Zhou 2009; Celia et al. 2011; Doughty 2010; Nguyen et al. 2017; William Carey et al. 2010; Yonkofski et al. 2019; Zhang et al. 2018)
Temperature gradient	(Doughty 2010; Yonkofski et al. 2019)
Wellbore length	(Celia et al. 2011; Doughty 2010; Onishi et al. 2019; Yonkofski et al. 2019)
Wellbore bottom depth	(Kang and Jackson 2016; USGS 2013)
Brine salinity	(Jordan and Doughty 2009; Kang and Jackson 2016)

Appendix B – Machine Learning Model for Fluid Properties

Random forest regression models were developed with 359,191 data points for CO₂ and 8,999 data points for brine. The data for CO₂ were collected from numerical reservoir simulations ran for this study while the data for brine was generated using (Meyer et al. 1993). Figure B.1 shows the unit plots for the testing data sets (39,911 data points for CO₂ and 1,000 for brine), and the regression model are accurate. Table B.1 and Table B.2 include input parameter ranges for the models. These models are also implemented into MSW of NRAP-Open-IAM.

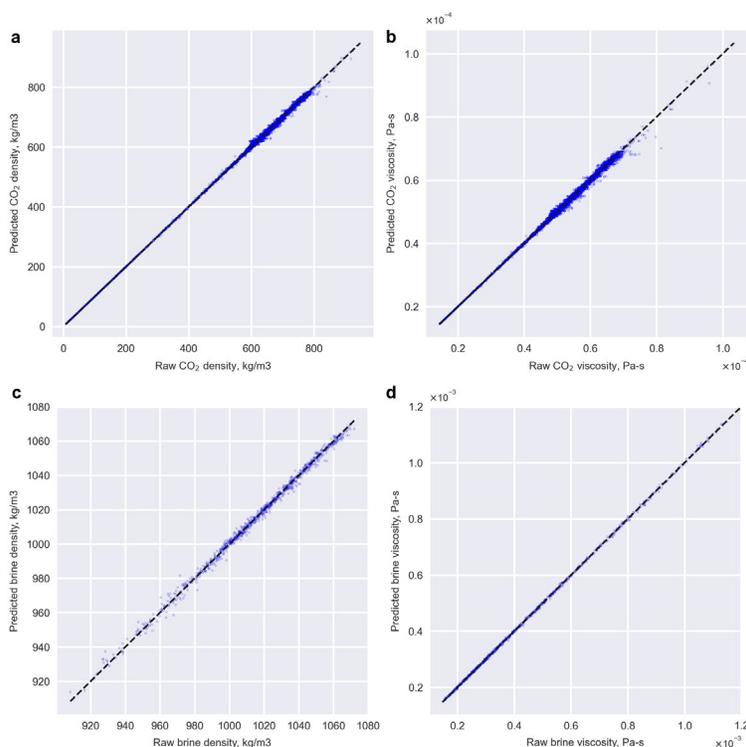


Figure B.1. Validation of the machine learning models for fluid properties: (a) CO₂ density, (b) CO₂ viscosity, (c) brine density, (d) brine viscosity.

Table B.1. Applicable ranges of the CO₂ fluid model.

Parameter	Unit	Min	Max
Temperature	°C	15.9	163.5
Pressure	MPa	0.488	79.309

Table B.2. Applicable ranges of the brine fluid model.

Parameter	Unit	Min	Max
Temperature	°C	15.9	234.3
Pressure	MPa	0.662	123.362
Salinity	kg/kg	0.0	0.010

Appendix C – Detailed Input Information for NRAP-Open-IAM Implementation

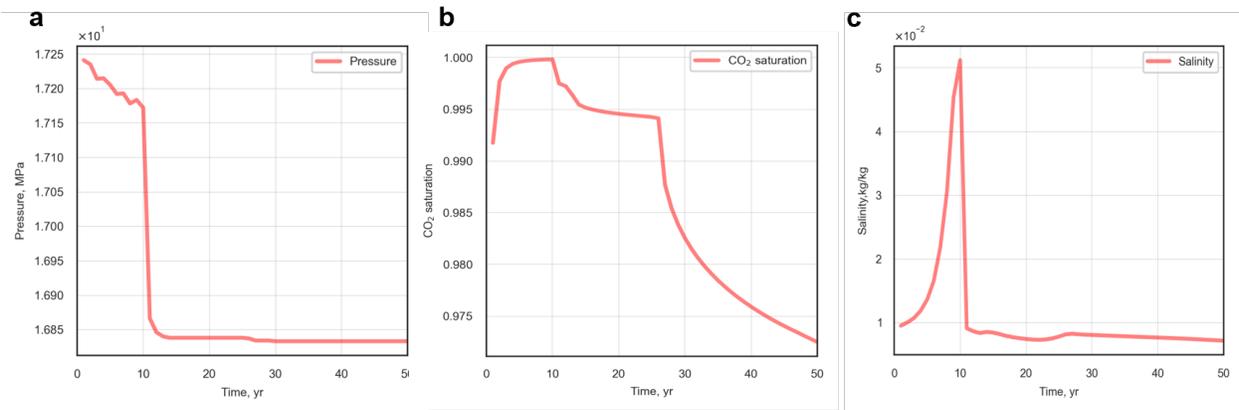


Figure C.1. Dynamic input data at the bottom of the leaky well for NRAP-Open-IAM: (a) pressure, (b) CO₂ saturation, and (c) salinity.

Table C.1. Static input parameters for NRAP-Open-IAM.

NRAP-Open-IAM input parameter	Units	Values
numberOfShaleLayers	-	3
logWell1Perm	log ₁₀ (m ²)	-13.728
logWell2Perm	log ₁₀ (m ²)	-100.0
logWell3Perm	log ₁₀ (m ²)	-100.0
logAqu1Perm	log ₁₀ (m ²)	-100.0
logAqu2Perm	log ₁₀ (m ²)	-100.0
shale1Thickness	m	80.3
shale2Thickness	m	30.0
shale3Thickness	m	1539.9
aquifer1Thickness	m	30.0
aquifer2Thickness	m	30.0
brineResSatAquifer1	-	0.0
wellRadius	m	0.03412
tempGrad	C°/km	24.164

*Default values were used for not given input parameters here.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov