# Optimal monitoring design for uncertainty quantification during geologic $CO_2$ sequestration: A machine learning approach

Misael M. Morales[a,b,*], Mohamed Mehana[a] and Bailian Chen[a]

[a] Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87544, USA
[b] Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX 78712, USA

ARTICLE INFO

*Keywords*:
Geologic carbon sequestration
Monitoring design optimization
Machine learning
Reduced-order modeling
Data assimilation
Uncertainty quantification

ABSTRACT

Geologic $CO_2$ sequestration (GCS) projects have large uncertainties in geologic properties, and require optimal monitoring designs for risk assessment and management. An effective monitoring design is crucial to ensure the safe and permanent geologic storage of $CO_2$. Optimal monitoring design involve an optimal placement of monitoring wells, and optimal monitoring measurement data (pressure, $CO_2$ saturation, temperature, etc.). We have developed a filtering-based data assimilation approach to design an optimal GCS monitoring strategy for well placement and monitoring data design. To efficiently solve the optimization problem and reduce computational costs, Artificial Neural Networks are used to develop computationally efficient reduced-order models based on full-physics numerical simulations of $CO_2$ injection in saline aquifers. We demonstrate our approach in two scenarios of $CO_2$ leakage through legacy or abandoned wellbores where an optimal monitoring strategy are devised to reduce the uncertainty in cumulative $CO_2$ leakage in the GCS site. The optimal monitoring design resulted in an uncertainty reduction in the cumulative leakage of $CO_2$ of approximately 73% and 62% in each case, respectively. The proposed approach is efficient in developing monitoring designs under geologic uncertainty and enables safe geologic carbon sequestration operations.

## 1. Introduction

Geologic $CO_2$ sequestration (GCS) has emerged as an important technology to reduce anthropogenic greenhouse gas emissions to the atmosphere [1–8]. This has become increasingly popular worldwide due to the need to meet international climate protection agreements [9, 10]. Different types of underground formations have been proposed to store $CO_2$ emissions including oil and gas reservoirs, coal beds and seams, and deep saline aquifers [11]. The main concern in GCS projects is potential leakage of the $CO_2$ through leakage pathways, such as improperly abandoned wells, faults, and fractures [1, 12–15]. Such risks can pose a major threat to overlying resources (e.g., groundwater resources, oil and gas reservoirs, etc.) and human health [16, 17]. Monitoring and verifying $CO_2$ behavior within the subsurface reservoir are crucial for detecting potential leakage, assessing storage capacity, and evaluating environmental impacts [18–20].

To ensure safe and efficient operations in a large-scale GCS site, risk management techniques are used to minimize and mitigate potential risks during $CO_2$ injection and post-injection periods [21–25]. Monitoring is thus an important aspect of GCS risk management, and one of the main goals of the Department of Energy (DOE) Office of Fossil Energy National Risk Assessment Partnership (NRAP) [26]. For this goal, several monitoring techniques have been developed, including near surface $CO_2$ flux and tracer measurements [27, 28], groundwater chemistry monitoring [29, 30], seismic surveying [28, 31–33], and pressure monitoring [34–38].

Optimal sensor placement and monitoring design play a critical role in achieving accurate and efficient monitoring in GCS projects. Depending on the reservoir properties and heterogeneity, the placement of monitoring wells can provide a more accurate measurement of the injected $CO_2$ plume and help mitigate potential leakage risks [39–42]. In common GCS operations, each injection well is paired with one monitoring well, though large-scale projects often incorporate a larger number of monitoring wells [43–45]. Moreover, the selection of monitoring measurement plays an important role in reducing uncertainties and quantifying risks in GCS operations [30, 46–49]. Therefore, it is crucial to define an optimal monitoring strategy in terms of both well placement and monitoring measurement type.

*Corresponding author
misaelmorales@lanl.gov (M.M. Morales); ~~mzm@lanl.gov (M. Mehana)~~
ORCID(s): 0000-0001-6923-1032 (M.M. Morales); 0000-0003-2472-2879 (M. Mehana); 0000-0003-3655-8340 (B. Chen)

Recent advancement in monitoring systems such as smart or intelligent wells are capable of providing large amounts of data in terms of volume, velocity, variety, value, and veracity [50–52]. Classical techniques in data processing and forecasting are sometimes hindered by big data, therefore machine learning provides a promising approach to enhance data-driven subsurface energy resource systems [53–57]. By analyzing extensive data sets, machine learning algorithms can uncover complex latent patterns and relationships that may not be discernible through traditional methods [58–65]. Machine learning approaches, when combined with reduced-order modeling (ROM) techniques, enable efficient and accurate prediction of key parameters [66–72], including pressure distribution, $CO_2$ plume migration, and reservoir behavior [73–77]. These insights facilitate the optimization of sensor placement and monitoring strategies, enabling better decision making and forecasting in GCS projects.

Accurately quantifying uncertainties is vital for the reliability of predictions and optimizing monitoring design under uncertain conditions [15, 39, 78–84]. Uncertainty quantification is particularly important in GCS due to inherent complexities and variabilities associated with subsurface conditions, fluid flow, and measurement errors [46, 85–87]. Several approaches for history matching or data assimilation have been applied to subsurface flow and transport, including Markov Chain Monte Carlo (MCMC) [61, 81, 82, 88, 89], randomized maximum likelihood (RML) [90], filter-based or rejection sampling (RS) [91–94], ensemble Kalman filtering (EnKF) [95–99] and ensemble smoother with multiple data assimilation (ES-MDA) [46, 100–105]. Filter-based approaches provide a robust framework for characterizing uncertainties associated with reservoir properties, operating conditions, and measurement errors, and with reduced complexity and cost compared to previously-mentioned techniques. Leveraging data assimilation techniques allows for informed risk assessment, ensuring the safety and efficiency of GCS projects.

Numerous research endeavors have been dedicated to addressing monitoring design, sensor placement, and uncertainty quantification in GCS. Previous studies have explored various modeling techniques, simulation frameworks, and optimization algorithms to enhance monitoring strategies and improve forecasting.

Efforts have been made to select the optimal monitoring measurements for GCS projects. Yonkofski et al. [49] use a simulated annealing (SA) global optimization approach to obtain the optimal monitoring measurement design in a GCS project. Their objective is to minimize the estimated time to first detection (ETFD) by iteratively mutating potential monitoring designs. Oladyshkin et al. [106] propose a polynomial chaos expansion (PCE) and bootstrap filtering approach for assimilating pressure data into reservoir models and quantifying the uncertainty reduction in $CO_2$ leakage rate at a GCS site. Liu and Grana [107] propose a deep convolutional autoencoder as a ROM strategy to assimilate seismic monitoring data in GCS. Their method requires HFS to obtain $CO_2$ saturation plume predictions from an ensemble of prior models, which is then used to calculate the seismic response. The autoencoder is used to project the observed monitoring measurements into latent space where ES-MDA is used to update the model parameters and quantify the uncertainty in predictions.

Similar efforts have been made in the area of optimal monitoring well placement. Sun et al. [47] propose an approach to optimize monitoring well location based on pressure measurements for GCS under geologic uncertainty. Using binary integer programming problem (BIPP) formulation, they effectively select optimal monitoring locations for homogeneous and fluvial heterogeneous reservoirs. However, their method requires a large number of forward simulations, which can be computationally costly and time consuming. Sun and Durlofsky [39] use a data-space inversion (DSI) approach to optimize the monitoring well locations in a GCS project with a genetic algorithm (GA) global optimization. Using principal component analysis (PCA) as a model reduction strategy, they reduce the uncertainty in $CO_2$ saturation plume using a RML approach. In this approach, posterior geological models are not generated in the DSI method, which is different from traditional ensemble-based data assimilation approaches.

Besides optimal well placement and monitoring measurement selection, several research studies have been conducted to quantify the uncertainty in GCS projects. Jia et al. [85] propose a Bayesian model average and Monte Carlo simulation to quantify parameter uncertainty based on a PCE-ROM. However, Monte Carlo strategies require a very large number of realizations and can be extremely computationally inefficient. Chen et al. [46] propose a risk assessment approach using ES-MDA with geometric inflation factors (ES-MDA-GEO) to quantify the uncertainty monitoring data and calibrate the prior uncertain geologic models. Their work leverages continuous data assimilation as new monitoring data becomes available in GCS projects to improve the underlying model and reduce uncertainties. Mehana et al. [108] provide a ROM-based approach to quantify wellbore leakage from depleted reservoirs in $CO_2$-EOR operations. They compare the performance of different machine learning-based ROMs for prediction of cumulative leakage and quantify the uncertainty using Monte Carlo simulations. Pawar et al. [109] provide a robust framework for quantitative risk assessment of leakage in GCS. Utilizing the NRAP-open-IAM (Integrated Assessment Model)

tool, they are able to quantify the leakage risk through legacy or abandoned wells in large-scale GCS projects. This framework can then be used to support permit applications for GCS projects.

In this paper, we build upon the work of Chen et al. [110] to systematically design an optimal monitoring placement and measurement strategy for large-scale GCS beyond naive monitoring well placement and monitoring design. Chen et al. [110] developed a robust framework for uncertainty reduction in cumulative $CO_2$ using Multivariate Adaptive Regression Splines (MARS) [111]. Using a filter-based data assimilation process, they quantify the uncertainty reduction in cumulative $CO_2$ leakage. However, their work assumes a predetermined, uninformed placement of the monitoring well and monitoring measurement type, relying solely on engineering judgement and fixed monitoring configurations.

We propose a method for optimal GCS monitoring design based on well placement optimization and monitoring measurement selection. We develop an artificial neural network ROM to predict cumulative $CO_2$ leakage from a prior ensemble of uncertain model parameters, and implement a filter-based data assimilation approach to select the most informative monitoring well location and measurement type in order to reduce uncertainties and $CO_2$ leakage risks. The structure of this paper is as follows: Section 2 present our methodology, Section 3 presents the results of our approach for two synthetic cases, and Section 4 summarizes our findings, discusses their implications, and outlines potential avenues for future research in the field of GCS.

## 2. Methodology

In this section we will discuss the approaches for uncertainty quantification, ROM development, ROM training and performance, and optimal monitoring workflow design.

### 2.1. Uncertainty Quantification

The goal of this study is to evaluate the value of data in GCS monitoring design. The value of data is quantified by the amount of uncertainty that is reduced in the cumulative $CO_2$ leakage, $M_c$, over the duration of a GCS project. The prior probability density function (PDF) of the cumulative $CO_2$ leakage is denoted as $P(M_c)$. In this study, prior refers to the probability distribution before a monitoring program is implemented. The distribution of potential monitoring data that could be measured at the monitoring wells is denoted as $D = [d_1, d_2, \dots, d_{n_d}]$, where $\{d_i\}_{i=1}^{n_d}$ are the individual monitoring data points obtained if a monitoring design were implemented in a particular leakage scenario, and $n_d$ is the total number of monitoring data points in $D$. In this study, monitoring data is sampled monthly, and can represent pressure, $CO_2$ saturation, or temperature values at the monitoring well. Thus, we denote $D^j$ as the $j^{th}$ realization of $D$. For each $D^j$, we obtain a posterior PDF denoted by $P(M_c|D^j)$, which can be calculated using a data assimilation procedure as the cumulative $CO_2$ leakage, $M_c$, for a given monitoring design data $D^j$. The objective is to quantify the value of information (VOI) estimated from a distribution of potential monitoring design, allowing us to choose an optimal monitoring well placement and monitoring measurement type to minimize the uncertainty in potential leakage scenarios. Following Chen et al. [81, 110] and Le and Reynolds [112], the VOI is quantified by the uncertainty reduction in the objective function. We denote the amount of uncertainty in cumulative $CO_2$ leakage distribution $P(M_c)$ as $U[P(M_c)]$, defined as:

$$U[P(M_c)] = P_{90}[P(M_c)] - P_{10}[P(M_c)], \tag{1}$$

where $P_{10}[\bullet]$ is the $10^{th}$ percentile of a distribution and $P_{90}[\bullet]$ is the $90^{th}$. The distribution of cumulative $CO_2$ leakage can be attributed to the uncertainty in model parameters, in this case the number of and the vertical transmissibility of potential leaky pathways, $k_v^{\ell}$, and the reservoir permeability multiplier, $k_R$. Therefore, selecting a monitoring design that reduces the uncertainty in $M_c$ ensures that the monitoring design will function effectively under multiple possible potential leakage scenarios.

The expected posterior uncertainty distribution in $M_c$ given $D$ is given by:

$$E_d[U[P(M_c|D]] = \frac{1}{\ell_d}\sum_{j=1}^{\ell_d} U[P(M_c|D^j)], \tag{2}$$

where $E_d$ is the expectation with respect to all realizations of D, and $l_d$ is the number of data realizations. The expected uncertainty reduction, $U_R$, as a result of data acquisition from a potential monitoring design is given by the difference

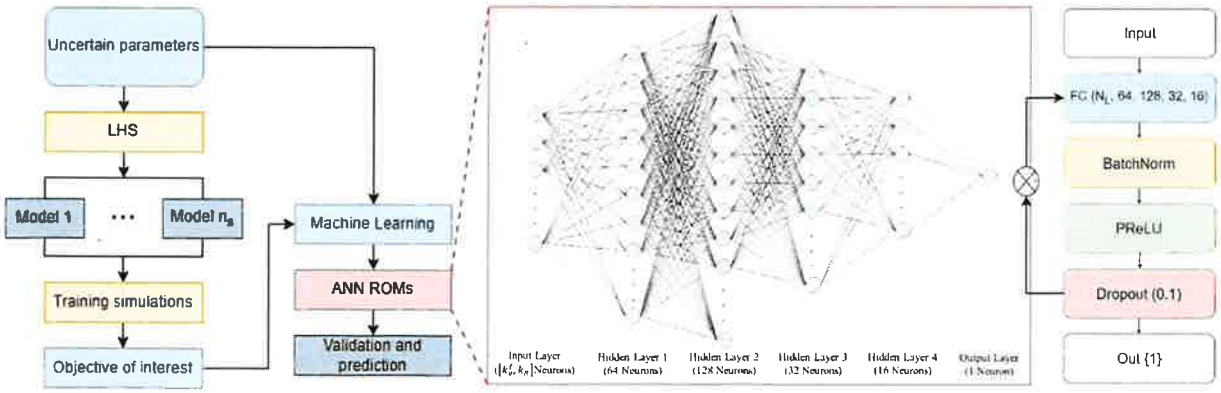



**Figure 1:** Workflow diagram for machine learning-based ROM development.

between the prior uncertainty and the expected posterior uncertainty in cumulative $CO_2$ leakage, as defined by:

$$U_R = U[P(M_c)] - E_d[U[P(M_c|D)]] \tag{3}$$

By selecting the optimal monitoring well placement and monitoring measurement type, the uncertainty reduction, $U_R$, quantifies the effectiveness of the particular GCS monitoring design, where the higher the uncertainty reduction the higher the VOI in the monitoring data obtained in the monitoring design.

## 2.2. Reduced-Order Model Development

Given the computational cost of traditional filter-based data assimilation, a reduced-order model is developed in this study. The workflow for the ROM development is illustrated in Fig. 1. This section provides a summary of the main steps in the ROM development workflow:

Step 1. *Experimental design*: Given a set of uncertain parameters $k_U^{\ell_d}$ and $k_R$, we generate $n_s$ training samples using Latin Hypercube Sampling (LHS) [113, 114].

Step 2. *Forward simulations*: Physics-based HFS of $CO_2$ injection and post-injection migration is performed with each of the $n_s$ training samples using the Finite Element Heat and Mass Transfer (FEHM) simulator [115].

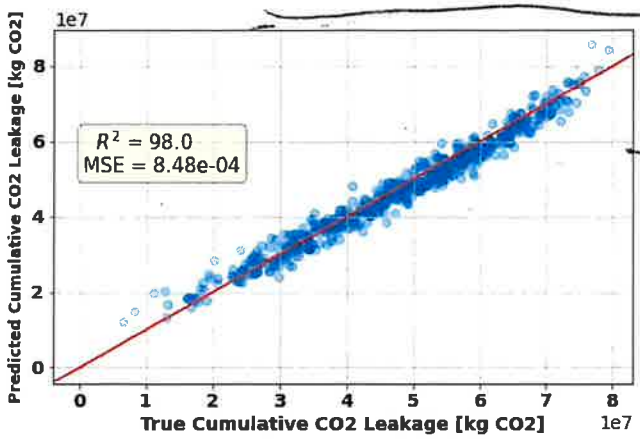
Step 3. *Collect training data*: For each training realization, the set of uncertain parameters, monitoring data, and cumulative $CO_2$ leakage are collected. In Fig. 1, we see that the uncertain parameters are inputs for the ROM training and the objectives of interest (cumulative $CO_2$ leakage and monitoring data) are the corresponding outputs.

Step 4. *Train ROMs for the objectives of interest*: A reduced-order model is used to map the relationship between the training parameters inputs and outputs. We build an ensemble of ROMs, one for each objective of interest, namely the cumulative $CO_2$ leakage ($M_c$) and the simulated monitoring data ($D$) at each specified time-step. A fully-connected artificial neural network (ANN) is implemented to build the ROMs. Fig. 1 shows the architecture of the ANN.

Step 5. *Validate the ROMs against the HFS*: Using 10-fold cross-validation [116], we test the predictions from the ROMs against the HFS results in order to perform hyper-parameter tuning and obtain robust ROMs that can be used for further predictions.

Using the Python TensorFlow/Keras packages [117, 118], we develop a fully-connected ANN architecture to build the ROMs. Each ANN consists of four hidden layers with sizes 64, 128, 32, and 16, respectively, with a total number of parameters equal to 14,705. A kernel regularizer is applied with the $\ell_1$-norm, and dropout of 10% is used on each hidden layer. The activation function is the parametric rectified linear unit (PReLU), which learns the negative slope for each batch in each epoch. The Adam optimizer [119] is used with a mean squared error (MSE) loss function. Training is performed on an NVIDIA RTX A6000 GPU in about 2 minutes for each ROM using 10-fold cross-validation. The average validation MSE is approximately $8.5 \times 10^{-4}$ and the correlation coefficient ($R^2$) is approximately 0.98. The truth vs. prediction performance for a set of 500 realizations of uncertain parameters is shown in Fig. 2.

(2) This will add a bit more weight to your paper
I remember you showed me one plot for the comparison before ?




(1) Could you please compare ANN based ROM with MARS ?
→ This is one improvement compared to previous work.

**Figure 2:** Cumulative $CO_2$ leakage prediction from ANN ROM vs true cumulative $CO_2$ leakage.

### 2.3. Workflow for Optimal Monitoring Design

In this section we present a filtering and ROM based workflow for optimal monitoring design of GCS. The workflow diagram is shown in Fig. 3. The main steps for the optimal monitoring design workflow are summarized below.

Step 1. *Develop ROMs for the objective function, $M_c$, and predict monitoring data, $D$:* A detailed description of the ROM development workflow were presented in the previous section. We build one ROM for each monitoring data point, $d_i$, in each data vector $D^j$. The vector of predicted monitoring data is denoted as $O(m) = [O_1(m), O_2(m), \ldots, O_{n_d}(m)]^T$, where $m$ is the vector of uncertain model input parameters, namely $k_v^\ell$ and $k_R$. The ROMs are used to replace FEHM physics-based simulations and to predict the objectives of interest for a set of new input parameters not in the training data.

Step 2. *Generate an ensemble of realizations of monitoring data, $D$:* Initially, $l_d$ realizations are sampled from the prior PDF of $m$, and are denoted as $\{\tilde{m}^j\}_{j=1}^{\ell_d}$. The corresponding monitoring data, $\tilde{d}_{obs}^j$, for each $\tilde{m}^j$ are given by:

$$\tilde{d}_{obs}^j = O(\tilde{m}^j) + e^j \tag{4}$$

where $O(\tilde{m}^j)$ is the ROM prediction for $n_d$ monitoring data points, and $e^j$ denotes the $j^{th}$ realization of measurement errors which follow a Gaussian distribution.

Step 3. *Generate Monte Carlo samples, and calculate prior uncertainty:* A large number (50,000) Monte Carlo samples are generated from the prior distribution of $m$, and denotes as $\{\hat{m}^k\}_{k=1}^{\ell_{MC}}$. The Monte Carlo samples are used to calculate the prior PDF and the amount of uncertainty in the prior can be computed using Eq. (1).

Step 4. *Filter the Monte Carlo samples, and compute expected posterior uncertainty:* Using a filtering-based method [94], also known as rejection sampling, we construct a posterior distribution of m conditional to each $\tilde{d}_{obs}^j$. First, using the Monte Carlo samples, $\hat{m}^k$, generated in Step 3, we simulate the corresponding monitoring data $\hat{d}^k$ with the ROMs generated in Step 1, such that $\hat{d}^k = O(\hat{m}^k)$. Here, $\hat{d}^k$ represents a realization from the distribution of potential monitoring data sets that capture potential $CO_2$ leakage scenarios given the uncertain input parameters $k_v^\ell$ and $k_R$. The data assimilation error is defined as the maximum absolute error (MAE) as follows:

$$MAE(d_{obs}^j) = \max_{1 \le i \le n_d} |\tilde{d}_{obs,i}^j - \hat{d}_i^k| \tag{5}$$

Given a threshold value $\tau$, the $\hat{m}^k$ sample is accepted as a legitimate realization of the posterior distribution according to the following acceptance probability:

$$P_{acc}(\hat{m}^k) = \begin{cases} 1, & \text{if } MAE < \tau \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

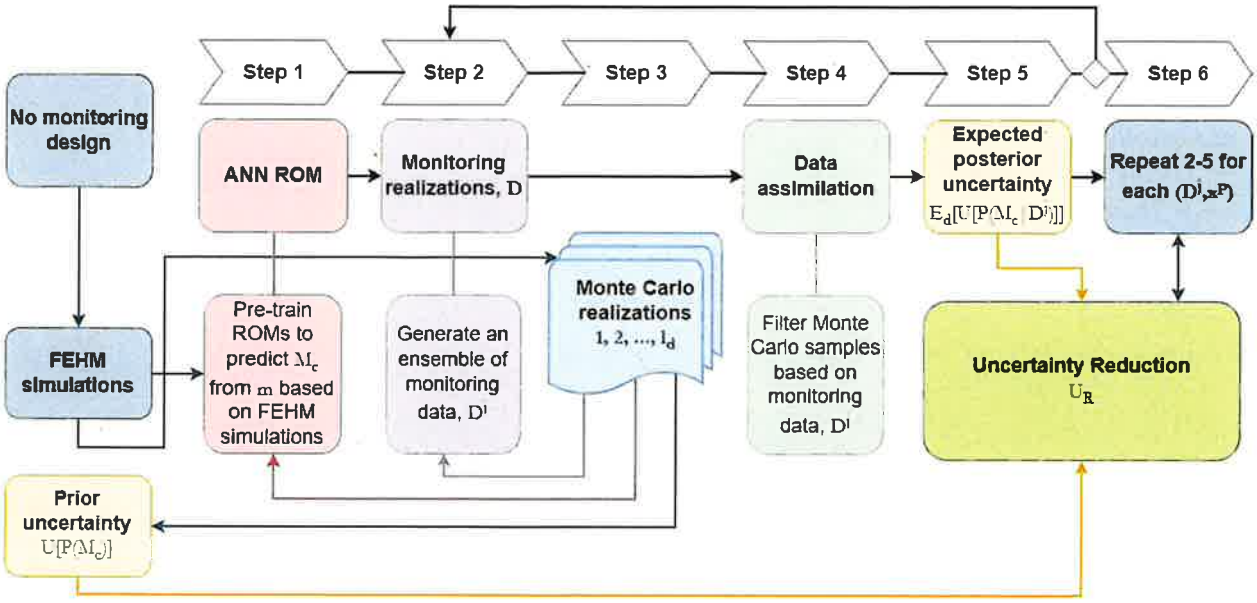



**Figure 3:** Workflow diagram for optimal monitoring design. First, calculate the prior uncertainty with no monitoring design based on Monte Carlo samples ~~of the FEHM simulations~~. Then use the pre-trained ROMs to predict $CO_2$ cumulative leakage using Monte Carlo samples and perform data assimilation of monitoring measurement $D^j$ for monitoring well placement $x^p$ to compute expected posterior uncertainty. The uncertainty reduction for case $(D^j, x^p)$ is given by $U^{x_p}[P(M_c)] - E_d[U^{x_p}[P(M_c|D^j]]$, and we repeat for each possible $D^j$ and $x^p$.

The threshold value, $\tau$, is chosen based on engineering judgement and takes into consideration the measurement and modeling errors. Therefore, $\hat{m}^k$ is accepted if it is deemed sufficiently consistent with the true monitoring data realization. Every Monte Carlo sample is evaluated using Eq. (6) and the accepted samples constitute the posterior distribution of m conditional to the monitoring data realization $\tilde{d}^j_{obs}$ such that $\ell_d$ posterior samples of $m$ are obtained. The expected posterior uncertainty is calculated using Eq. (2).

Step 5. *Calculate the expected amount of uncertainty reduction $U_R$.* The expected amount of uncertainty reduction, $U_R$, is calculated by comparing the uncertainty in the prior distribution and the expected value of the uncertainty in the posterior distribution using Eq. (3).

Step 6. *Monitoring well placement optimization.* We repeat Steps 2-5 for every possible monitoring well location in the GCS area of review (AOR), conditional to the data for each possible measurement type, $D^j$. In order to accelerate the optimization procedure, we coarsen the simulation grid into a $4 \times 4$ subgrid, meaning there are 16 possible monitoring well locations. We calculate the expected amount of uncertainty reduction for each monitoring data type, $D^j$, for each possible monitoring well location $\{x^p\}_{p=1}^{16}$, and obtain the monitoring design that maximally reduces the uncertainty in cumulative $CO_2$ leakage (maximally reducing the uncertainty is equivalent to minimizing the negative expected uncertainty reduction), as shown in Eq. (7)

$$x_p^* = \min_{1 \le p \le 16} -U_R^{x_p} \quad (7)$$

This results in an exhaustive search in the subgrid to obtain the optimal well location, $x_p^*$, that yields the highest uncertainty reduction, defined by $U_R^{x_p}$ as follows:

$$U_R^{x_p} = U^{x_p}[P(M_c)] - E_d[U^{x_p}[P(M_c|D^j]] \quad (8)$$

With this optimal monitoring design workflow, the expected uncertainty reduction in cumulative $CO_2$ leakage for each potential monitoring measurement and each potential monitoring well location can be computed, and the optimal monitoring design that reduces the uncertainty in the simulated amount of $CO_2$ leakage is obtained.

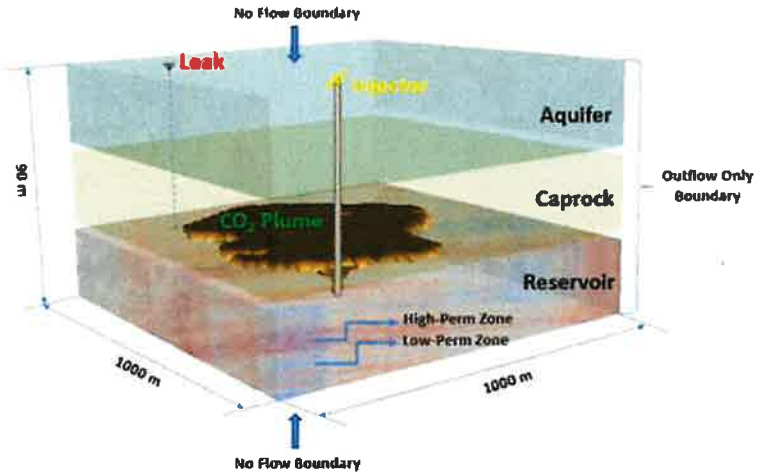



**Figure 4:** Schematic of the base model, in which a storage reservoir and aquifer are separated by a caprock. At the center is a $CO_2$ injection well. The vertical axis is exaggerated 7 times.

## 2.4. Model Description

We implement the optimal monitoring design workflow on a synthetic GCS model consisting of a heterogeneous storage reservoir, a homogeneous caprock layer and a homogeneous aquifer, as shown in the schematic of the base model in Fig. 4. The thickness of each of the three layers is 30 $m$, and the model is 1 $km$ wide in the horizontal dimensions. The depth from ground surface to the top of the model is 1000 $m$. A $CO_2$ injection well is placed at the center of the reservoir and multiple potential leakage pathways traverse the caprock, where $CO_2$ could potentially leak into the aquifer. Note that only one possible leakage pathway is shown in Fig. 4, while we have considered several scenarios with multiple potential leakage pathways. The caprock and aquifer layers have a homogeneous permeability distribution equal to $1\times10^{-1}$ $m^2$ and $1\times10^{-13}$ $m^2$, respectively. The storage reservoir has a heterogeneous permeability distribution, as shown in Fig. 5. The base model is generated using a spherical variogram model [120, 121] with major and minor correlation lengths of 680 $m$ and 280 $m$, respectively, with a major direction of 45° from the positive $x$-axis.

The mean of the permeability field is $1\times10^{-13}$ $m^2$. For each realization, we assume that the reservoir permeability is uncertain, and to honor this uncertainty we use a permeability multiplier, $k_R$, to multiply the aforementioned base permeability distribution. The lower and upper bounds for the multiplier $k_R$ and the potential leaky pathways $k_v^\ell$ are shown in Table 1.

**Table 1**
Uncertain parameters and their lower and upper bounds.

| Uncertain parameters | Symbol | Lower bound | Upper bound | Unit |
|---|---|---|---|---|
| Reservoir permeability multiplier | $k_R$ | 0.5 | 2 | – |
| Permeability of leaky pathway(s) | $k_v^\ell$ | -19<br>0.001 | -14<br>10 | $log_{10}$ $[m^2]$<br>$mD$ |

A numerical mesh for the reservoir simulation is made using the grid generation toolkit *LaGriT* [122]. The numerical mesh has 51 grid nodes in both the $x$- and $y$-directions, and 31 grid nodes in the $z$-direction. The distance between each grid node in the x- and y-directions is 20 $m$, and in the $z$-direction it is 3 m. The total number of grid nodes used in the simulation is 80,631, with 26,010 grid nodes in the reservoir and caprock, respectively, and 28,611 grid nodes in the aquifer. FEHM is used for 3D multi-phase flow simulations [115]. The boundary conditions of the reservoir are defined as Dirichlet boundaries, allowing $CO_2$ to flow out but not in, and water pressure above hydrostatic. The top and bottom boundary conditions of the simulation model are no-flow boundaries. The thermal conditions of the model are initialized using a geothermal gradient of 0.03°C/m with a temperature of 20°C at the top. Pressure gradients are initialized at $9.81\times10^{-3}$ MPa/m with a pressure of 0.2 MPa along the top. In this study, $CO_2$ is constantly injected in a five-year period, monitored monthly, with a constant injection rate of 0.1 million metric tons/year.