CCUS: 4011841

De-risking Saline Aquifer-Type CO_2 Storage Resources Via Machine Learning-based Reservoir Modelling. Case Study, Bunter Sandstone Formation, Southern North Sea

Edwin Tillero*[1], Jose L. Mogollon[1], Francisco Tillero[1], 1. Movus Energy Solution.

Abstract

This work aims to develop a surrogate model (machine learning-based) to expedite the assessment of geologic CO_2 storage (GCS) sites by predicting the CO_2 containment efficiency of saline aquifers. De-risking as much as possible geologic CO_2 storage resources is an absolute priority to help with the global goal of reaching net-zero by mid-century. Geologic CO_2 storage numerical modelling (GCSNM) is a comprehensive technique to understand the long-term containment of CO_2, but it's a time-consuming multi-parametrial process, with relative high cost, especially when groups of geologic sites must be evaluated.

To overcome before mentioned drawbacks, machine learning-based reservoir modelling arises as a cheap, quick, and computationally efficient tool for assessing multiples storage sites. In this paper, a case study is presented. A neural network-based geologic CO_2 storage (NN-GCS) model was built and fed with a seven-parameter subset as input (CO_2 residual saturation, horizontal permeability, vertical to horizontal permeability ratio, porosity, brine salinity, flow rate, and elapsed time) and 4-parameter subset as output (Residual, Solubility, and Structural Trapping Index along with CO_2 injected volume). Such dataset was built from hundreds of experimental design-based numerical realizations derived from a synthetic aquifer numerical model of the dome-like shape Bunter Sandstone Closure 36 aquifer in Southern North Sea (SNS), UK Continental Shelf (UKCS). The NN-GCS model architecture was designed in Python and used root mean square error (RMSE) and coefficient of determination (R_2) as best-fit indicators of the NN-GCS performance. In addition, it was optimized regarding numbers of nodes (40) and layers (3), showing accuracies, for instance R_2 for training and testing with 96% and 95% precision respectively. Finally, a field application of the NN-GCS model was performed based on basic geologic information from a group of Bunter Closures of interest in SNS basin to assure the feasibility to extend its application to another dome-shape deep saline aquifers in UKCS.

Results showed, at the end of 100-years injection case, a Structural, Residual, and Solubility Trapping Index averaging 83%, 11%, and 6% respectively. The variation coefficient averaging 5% indicate properly

predicted trapping indices because all the aquifers have similar structures (dome-like shape) and relatively same properties. In addition, CO_2 injected volume predictions were ranging from 397 to 456 million tons/reservoir, totaling 2.1 giga ton (Gt) of potential storage capacity which represents 70% of total theoretical volumetric capacity. These results show the great impact of accelerating geologic CO_2 storage sites (closure saline aquifers) assessment by implementing a ML-based modelling to de-risk and classify groups of saline aquifers as ready-to-be-considered potentially feasible CO_2 storage sites.

Introduction

CO_2 capture and storage (CCS) surges as an order-of-the-day solution in capturing million tons of CO_2 –generated from power generation, industrial processes, and other sources– and containing such volume in underground geologic formations. In that sense, This CO_2 emission reduction technology is aligned with 2050 net-zero emission goal.

Regarding CO_2 storage resources in UK Continental Shelf (UKCS), its potential is vast mainly in deep saline aquifers, which has been assessed approximately at 68,000 million tons (Mt) or 68 giga tons (Gt), representing over 85% of the national total CO_2 storage resource [1].

However, several factors –like site geology and business model– challenge the technical feasibility and commercial development of CCS projects. Regarding geology-related factors, geologic trap characteristics and reservoir properties along with fluids properties constitute most fundamentals aspects to consider for CO_2 storage site modelling stage and its management. Accordingly, geologic CO_2 storage modelling (GCSM) is the first step in CO_2 storage site appraisal. The GCSM may go from too simplistic analytical model to fast approximate data-driven techniques (surrogate models) and to complex reservoir numerical simulation which power and cost increase accordingly [2].

The GCSNM presents some similarities with oil & gas reservoir numerical modelling processes which go from a subsurface geology and reservoir data collection, the building of a geo-grid static model, to the building of a compositional numerical reservoir model. However, it may differ when geochemical and geomechanical effects of CO_2 injection must be considered. In addition, modelling larger area usually must be carried out to capture the long-term migration of CO_2 plume, longer time scales to monitor post-injection CO_2 movement, and brine management, among others.

Although GCSNM is considered a comprehensive technique to understand the CO_2 behavior in geologic store sites, questions arise about how often the GCSNM can be used when it is applied to multiple sites and for wide variation ranges of aquifer parameters to be sensitized. In addition, how much time the typical timeframe of CO_2 storage appraisal process can be reduced when it has been shown to be time-consuming. In the other hand, the urgency to reduce CO_2 emissions requires more storage sites been classified as ready-to-inject, which demands quicker but efficient GCSM process.

Proxy (surrogate) reservoir models are considered an alternative to GCSNM by utilizing plenty numerical reservoir data and generating approximate results with significant reduction in computational cost and at the same time honoring reservoir variability when referring to several geologic storage sites assessment. Among such proxy techniques are machine learning (ML) approach, particularly artificial neural networks (ANN), which have proven to be efficient tools to relate inputs and outputs from a high-dimensionality processes such as GCS [3]. This is what this work is about and its application for GCS modelling in saline aquifers lacking numerical modelling.

The ANN model developed is fed by a parametrical dataset resulting from implementing an experimental design-based numerical realizations process. The experimental design is performed via an optimization tool. Input to the ANN-GCS model are reservoir characteristics and elapsed times, and as output there would be indicators related to CO_2 trapping efficiency and cumulative volume of injected CO_2.

The numerical aquifer base model was derived from an existing saline aquifer numerical model called Bunter Closure 36 which is one of multiple dome-like shape geological structures (anticlines) located in the UK Continental Shelf (UKCS), more specifically in the Southern North Sea (SNS) area [4]. It is worth mentioning that SNS presents a considerable number of saline aquifer anticlines which lack numerical models. In addition, SNS is close to an important group of Southern East onshore's CO_2 emitters which makes SNS area a strategic place with transport and storage conditions necessary to deploy CCS projects [5]. Therefore, therein lies the study's contribution which will help to speed up the assessment of such group of saline aquifers and to the best of my knowledge, this is the first adoption of ML technique as predictive tool to evaluate GCS performance for deep saline aquifers in Bunter Sandstone formation in SNS area for UKCS.

1. Theory

Guidelines have been discussed to characterize underground geological storage sites by numerical modelling [6]. In general, GCSNM has most of elements of an oil & gas reservoir numerical modelling. A multidisciplinary approach is needed to reconstruct geologic site architecture (static modelling) by describing its geometry, boundaries, and quantifying reservoir rock properties. In addition, multi-phase flow, geochemical, and geomechanical phenomena are modelled by a numerical model (dynamic modelling) from which injection scenarios, development plan, optimization, and risk assessment can be derived. A comprehensive GCSNM provides estimates of storage volume, injection capacity, geological containment assurance for long-term, and leakage risk assessment.

Numerical simulation of geologic CO_2 storage (GCS)

The CO_2 geo-sequestration numerical modelling require a high-fidelity geological model an involves the solution of 1) transport of reactive multi-components multiphase flow (Darcy's Law), 2) equations for thermodynamic equilibrium between gas and aqueous phase (Equations of State), and 3) mineral dissolution/precipitation reaction equations. General governing equation for CO_2 storage numerical dynamic modelling is described as follows, assuming phase and thermal equilibrium [7]:

$$\underbrace{\sum_{j=CO_2,w} \Delta \mathrm{T} \frac{k_{ij}}{\mu_{ij}} \rho_j\, y_{ij}\, \Delta \mathrm{P}_j}_{convection} + \underbrace{\sum_{j=CO_2,w} \Delta \mathrm{AD}_{ij} \Delta \rho_j\, y_{ij}}_{difussion,\ dispersion} + \underbrace{r_i}_{reaction} + \underbrace{q_i}_{injection} = \underbrace{\frac{V}{\Delta \mathrm{t}}\left(N_i^{n+1} - N_i^{n}\right)}_{accumulation} \qquad (Eq. 1.1)$$

The convection term expresses the multi-phase flow of components caused by a delta pressure (ΔP) and explained by Darcy's law (considers k_r, μ, ρ). The second term represents the rate of diffusion and dispersion of CO_2 in the liquid phase (solubility). Source term for CO_2 injection is given by injection and reaction due to precipitation and dissolution mechanisms derived from rock-fluids interactions. The right-side term expresses accumulated molar volume of components in the control volume which is function of saturation of components and porosity. Gas density (ρ_g) can be estimated by appropriate EOS. Gas viscosity (μ_g), water density (ρ_w) and water viscosity (μ_w) in the aqueous phase are calculated with proper correlation.
Regarding an efficient design of dynamic models, a coarser grid model is usually required due to the initial large size and finely gridding static model. Coarser grid is used to improve the computational efficiency of the dynamic model. This is crucial as the model will be run for hundreds or thousands of years into the future.

GCS performance assessment for deep saline aquifer

The performance quantification of underground CO_2 containment is one of the main requirements to identify potential storage site. Therefore, identification of acting underground CO_2 trapping mechanisms plays a significant role in assessing the CO_2 storage performance. CO_2 storage performance may vary depending on location (sedimentary basin), reservoir characteristics, and timescales. Structural, residual gas, solubility, and mineral CO_2 trapping mechanisms have been identified as main ways for trapping CO_2

in saline aquifers [8]. Most of trapping mechanisms have been described in reported literature as trapping indices, as follows:

$$\text{Residual Gas Trapping index RTI(t)} = \frac{\text{Total mass of } CO_2 \text{ trapped as residual gas at time t}}{\text{Total mass of } CO_2 \text{ injected at time t}} \quad (Eq. 1.2)$$

$$\text{Solubility Trapping index STI(t)} = \frac{\text{Total mass of } CO_2 \text{ soluble in brine at time t}}{\text{Total mass of } CO_2 \text{ injected at time t}} \quad (Eq. 1.3)$$

A trapping index defined as Structural Trapping Index (StTI) is introduced in this work and it is defined as:

$$\text{Structural Trapping index StTI(t)} = \frac{\text{Total mass of } CO_2 \text{ trapped by buoyancy at time t}}{\text{Total mass of } CO_2 \text{ injected at time t}} \quad (Eq. 1.4)$$

GCS numerical modelling (GCSNM) in UK

There are several studies reported in literature about characterizing geologic CO_2 storage sites in UK which go from analogues, analytical, to large-scale and detailed UCS numerical modelling [1,9,10]. The most comprehensive characterization study of GCS sites in UK was carried out by ETI (Energy Technologies Institute) [1]. This project identified several CO_2 storage sites and carried out detailed modelling to selected sites given their potential contribution to pivot commercial-scale carbon capture and storage projects for power and industry emitters, and to de-risk these stores for future storage developers. Among selected sites were Bunter Closure 36, Viking Gas Field, Captain Aquifer, Hamilton Gas Field, and Forties 5 Aquifer. Special attention has been given to water-bearing Bunter Sandstone formation (deep saline aquifer) in the Southern North Sea (SNS) area, UK Continental Shelf (UKCS). SNS area, in addition to contain large closure structures (anticlines), it is close to an important group of UK's CO_2 stationary emitters. Estimation of up to 20 $MtCO_2$ per year for a 50-years injection timeframe was obtained by numerical dynamic modelling of NE part of Southern North Sea [10]. More details about Bunter 36 aquifer modelling results are given in methodology section since it was selected as base numerical model for this study.

Although GCSNM is a comprehensive technique to understand the long-term geological CO_2 storage behavior, it is a multi-parametrial and multi-scale process and it requires quantity, quality, and variety of data. In addition, characterizing geological storage can take years, especially less-studied saline aquifer sites so, it has shown to be time-consuming and with relatively high cost. For instance, BEIS (Business, Energy, and Industrial Strategy, BEIS) considers that the site characterization (static and dynamic modelling) of a UCS project may take between 1 to 4 years [1].

Machine learning approach for GCSNM

Proxy models (e.g. machine learning) have been applied to model lifecycle of CO_2 geo-sequestration creating simplified approaches of reservoir responses in place of the reservoir simulators [3]. Among machine learning (ML) techniques are artificial neural networks (ANN) [2,11,12].

In general, ANN consists of an input layer, one or more hidden layers, an output layer, and transfer functions with the goal of finding nonlinear relationship among variables. In the context of GCS modelling by a neural network architecture, mathematical formulation is as follows:

$$y_m = f_0\left[\sum_{j=1}^{m} w_{jm} f_h \sum_{i=1}^{n}\left(w_{ij}x_i + b_j\right) + b_m\right] \quad (Eq. 1.5)$$

Where x_i is input parameters, y_m is output variables, both from numerical realizations dataset. w_{ij} are input and hidden layers link connections' weight from i to j node, b_j and b_m are bias for the hidden and output layers respectively. f_0 and f_h are transfer nonlinear functions for hidden and output layers; m and n are number of hidden nodes and input variables respectively.
This work will develop an ANN-based model to predict trapping efficiencies and CO_2 injected volume to surrogate GCSNM process to a group of deep saline aquifers as an early-stage appraisal of the potential of a UCS site and as an alternative to geologic sites assessment lacking numerical models.

2. Methodology

The ANN-based GCS modelling (ANN-GCS) workflow proposed in this work is illustrated in Figure 2.1

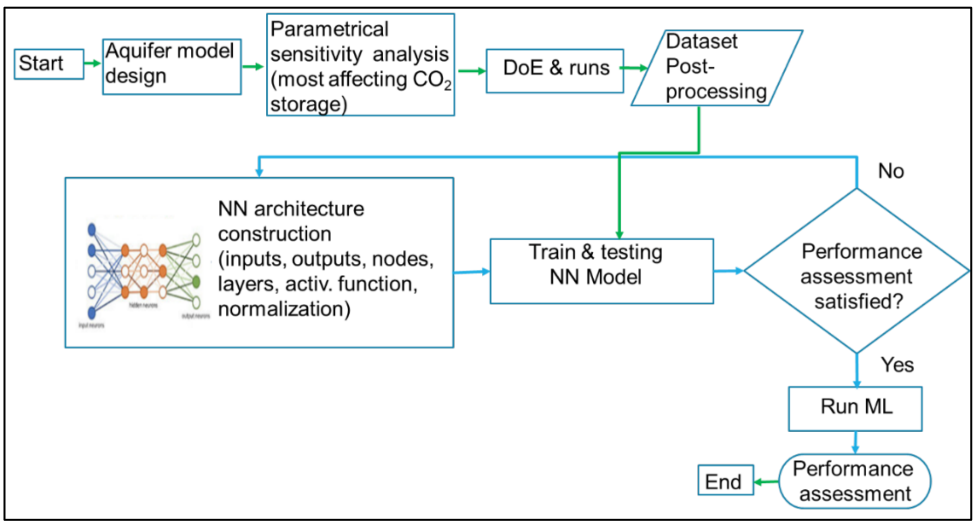

Figure 2.1 Workflow for GCS dynamic modelling by Artificial Neural Network

The methodology applied to study case is described as follows:

Step 1: GCS numerical aquifer base model design. One relevant aspect of an aquifer numerical base model to be used is its flexibility in honor geological typology, reservoir properties variations (heterogeneities), and operating conditions. As a result, a synthetic aquifer model that matches aquifers variant is guaranteed. The Bunter 36 saline aquifer numerical model was chosen from the select site inventory of ETI (Energy Technologies Institute) which is an open license data base [1]. Bunter sandstone is a geologic formation which domains in the Southern North Sea (SNS) area. The SNS is characterized by large anticlines which represent a strategic national geologic storage resource. In addition, it is close to an important group of UK's CO_2 stationary emitters (fossil fuel-based power stations) and is located near future shore terminal Barmston Beach [5].
Bunter Closure 36 aquifer is a dome-like structure, and it is 1200m below sea level and 200m thick. ETI's study concluded that competent caprocks (impermeable strata) exist above (Rot Halite Formation) and below (Bunter Shale) and likelihood of CO_2 leakage risk due to geological failure of such caprocks is low [6]. On the other hand, based on a geochemical modelling undertaken in the same ETI study, the impact of CO_2 injection with time over mineralogical changes is low due to a low rate of reaction in the reservoir given the quartz-dominated mineralogy and low temperature would not reach equilibrium even after 10,000 years, hence it suggests negligible impact on the injection timescale to be defined in this study [6].

Therefore, geomechanical and geochemical phenomena will not be modelled in our study, hence numerical complexity will be reduced.

The original grid of Bunter Closure 36 is 124x134x41 (total grid number of 681,256) with four injection wells run in 1 hour. In this study, the model was upscaled to 84x94x15 size (total grid number of 118,440) to optimize the run time to around quarter hour. It is worth mentioning that the upscaling has been weighted to ensure that a representative value of properties has been captured at coarser model. The top and bottom boundaries were set as closed, with no flow, while boundary conditions of four sides were set as open because non geological features at boundaries are described in its geomodel, hence it is considered infinite acting aquifer.

Figure 2.2 illustrates a version of the original numerical dynamic model of Bunter Closure 36 aquifer from ETI's study which served as base model to generate an aquifer synthetic dynamic model. CO_2 trapping indices were previously forecasted by ETI's study under injection scenario of 56 years, resulting in 73% of the CO_2 injected mass is structurally trapped (StTI), 22% residually trapped (RTI), and 5% dissolvably trapped (STI) respectively at the end of period.

Table 2.1 shows aquifer parameters and constraints of the model. ETI's study forecasted for 40-years and 56-years of injection period, totally 280 and 391 $MtCO_2$ respectively (dynamic storage capacity). Those results will be used as reference to validate the ANN model.

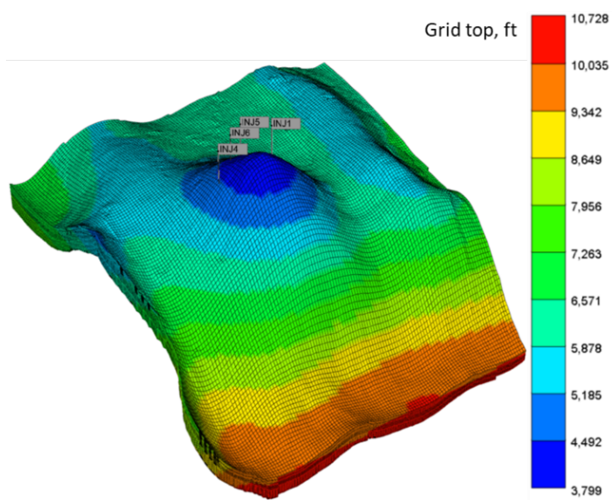

Figure 2.2. Upscaled numerical dynamic model of Bunter Closure 36 aquifer

Parameter	Value	Parameter	Value
Porosity (fraction)	0.2	Reservoir pressure @ 1,170 m, bar	119
Horizontal permeability (miliDarcies, mD)	210	Reservoir temperature, °C	45
Thickness (m)	200	Rock fracture pressure @ 1,170 m, bar	197
Reference depth (m)	1170	Max. well bottomhole pressure @ 1,170 m,	148
Salinity (ppm)	250,000	CO2 injection rate per well (Mton/year)	1.75
Residual gas saturation (fraction)	0.3	Number of injection wells	4
Vertical to horizontal permeability ratio,	0.35	Start of injection, year	2027
		Injection evaluation period(s), years	40 and 56

Table 2.1. Properties and constraints of Bunter Closure 36 aquifer model.

Step 2: Output parameters selection for CO_2 storage performance assessment. To estimate the effectiveness of CO_2 trapping in the aquifer, output parameters related to CO_2 trapping efficiency will be used in this study (equations 1.2 to 1.4). In addition, injected CO_2 volume is a parameter included, which is related to trapping indices which are fractions of total CO_2 injected volume.

Step 3: Input parameters selection and sensitivity analysis. Ranges of CO_2 storage affecting factors were studied and selected from previous works, such ETI's study, and based on data from UKCS' aquifers. A sensitivity analysis for Bunter Sandstone 36 aquifer model to identify the most influential factors affecting CO_2 sequestration capacity was carried out in ETI's project [6]. During sensitivity analysis a randomized values selection process of each influential factor was performed. ETI study's parametric sensitivity analysis showed that storage capacity is negatively affected by wells design factors such as closer wells placement, greater number of injection wells, and higher injection rate, all of them significantly increasing reservoir pressure near rock fracture pressure. Such drawback effects related to wells design were addressed and optimized during ETI's study. Similarly, reservoir properties also adversely affected storage capacity such as smaller aquifer size, lower permeability, lower vertical to horizontal permeability ratio, and lower residual gas saturation. In summary, to honor aquifer model parametrial variability, most affecting factor were porosity (φ), permeability (k), vertical to horizontal permeability ratio (kv/kh), salinity, residual gas saturation (S_{gr}), and injection rate [6].
Table 2.2 reflects range of influential factors (input parameters) used during design of experiment to generate numerical realizations which were later used to create dataset for training the ANN model.

Input variables	minimum	most likely	maximum
Porosity (fraction)	0.03	0.2	0.4
Permeability (mD)	0.1	200	2300
Salinity (ppm)	1,000	200,000	500,000
Residual gas saturation (fraction)	0.05	0.3	0.5
Kv/Kh (fraction)	0.01	0.36	0.8
CO_2 total injection rate (MT/year)	1.5	7.0	10.0

Table 2.2. Ranges of influential factors upon storage performance.

Step 4. Design of experiment (DoE) and data post-processing. Hundreds, even thousands of numerical experiments may be required to feed an ANN model. In this work, experiments were launched in the compositional simulator CMG-GEM and automated by an optimizer software (CMG-CMOST) considering the range of input variables. A design of experiment (DoE) or sampling technique was required to synthesize experiments/simulation data. This DoE process was performed based on Monte Carlo's approach. To analyze the effect of extended injection period on cumulative volume of CO_2, the design of experiment was carried out for longer injection period (100 years) in comparison to 56 years of maximum injection period defined by ETI's study [6].

Step 5. ANN architecture design, train, and testing. A supervised machine learning (ML) model was developed in this study based on ANN approach. The ANN algorithm was fed with dataset from DoE cases (step 4). Once the dataset was cleaned and adequate at format to feed the proxy model, it was randomly divided in training and testing, splitting data with a starting point of 80% and 20% respectively. The ANN model architecture was designed in Python programming language and the library Keras with Tensor Flow was used. The model in this study was based on the perceptron approach devised from human neural neurons. The rectified linear unit (ReLU) function is the most used activation function because of its lower run time than other functions. Finally, the ANN model predicted RTI, STI, StTI, and CO_2 injected volume.
Optimization of numbers of neurons and hidden layers.
The definition of number of neurons and hidden layer is an important process during the artificial neural network design because it impacts its performance for modelling. To find optimal number of layers and nodes, a trial-and-error process is performed based on different sizes of both layer and node units added

one a at time until best performance is obtained. Around 40 neurons and 3 layers were considered as the optimal ANN architecture.

ANN-GCS performance assessment. Expected performance of ANN model was defined by adopting two error functions such as the coefficient of determination (R^2) and the root mean squared error (RMSE). R^2 magnitudes the variance from dependent variable, being R_2 near 1 indicative of best performance and vice versa. Equation is shown as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{i_{act}} + y_{i_{pred}}\right)}{\sum_{i=1}^{n}\left(y_{i_{act}} + \bar{y}_{i_{pred}}\right)} \qquad (Eq. 2.1)$$

where $y_{i_{act}}$ is the actual value, $y_{i_{pred}}$ is the predicted value, and $\bar{y}_{i_{pred}}$ is the average of predicted values. Regarding RMSE, it expresses the spread of predicted errors giving an idea about how close or far is the data from the regression model so, the lower RMSE the more accurate performance and vice versa. The equation is:

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}\left(y_{i_{act}} - y_{i_{pred}}\right)^2\right]^{0.5} \qquad (Eq. 2.2)$$

Step 6. ANN-GCS model field application. In this task, basic input data was used from similar geologic sites in SNS area to evaluate stability of new ANN model and to ensure the applicability of the predictive model to similar CO_2 storage sites lacking numerical models.
Field application of the developed ANN-GCS model was conducted using basic parameters from the CO2Stored® public UK's database by selecting some CO_2 potential aquifer storage sites, specifically a group of Bunter Sandstone aquifers [14]. In addition, following premises were considered:

- Present geologic similarities (anticline closures) like the base aquifer model used to train the NN model.
- Migration and leakage through the sealing caprock is not considered to be major risk.
- They are located relatively close to the biggest concentration of carbon emitters in the UK.
- The same number of injection wells are expected from comparative development concept (4 injection wells).
- Injection period of 100 years.

Four (4) Bunter Sandstone closure aquifers were selected from the CO2Stored® database to predict their trapping indices and CO_2 cumulative injected volume, which values have not been before predicted because such aquifers lacking numerical models. The same input parameters used to train the ANN model were also used for field application of the new ANN-GCS model, as shown in Table 2.3. Bunter Closure aquifer 3, 9, 35, and 40 were selected from CO2Stored® database repository. Worth noting that Bunter Closure 36, which was the aquifer base model, is just presented in the table as reference.

Site Description	Porosity, frac	Permeability, mD	Kv/Kh ratio	Residual Gas Saturation	Thickness, m	Salinity, ppm	Total Injection rate (P50) Mt/yr
Bunter Closure 3	0.21	350	0.3	0.3	240	180000	5
Bunter Closure 9	0.21	350	0.3	0.3	300	180000	15
Bunter Closure 35	0.26	400	0.3	0.3	245	180000	10
Bunter Closure 36 **(as reference)**	**0.2**	**200**	**0.35**	**0.3**	**220**	**200000**	**7**
Bunter Closure 40	0.2	271	0.3	0.3	230	180000	2

Table 2.3. Basic data of selected Bunter Sandstone aquifers (CO2Stored® database). Field application.

3. Results

Approximately 700 numerical realizations were generated and 690 selected during DoE process and post-processing of data. Figure 3.1 illustrates trapping indices profiles and CO_2 cumulative injected volume vs. time from numerical realizations (light blue curves). In addition, the original aquifer base case is shown (black curve). Elapsed times comprising 690 samples each were defined. Finally, 8,280 samples for a total of 12 elapsed times (1, 5, 10, 20, 40, 70, 120, 140, 170, 220, 240, 272 years) from year 2028 to 2300 were post-processed to build the final dataset feeding the ANN model to predict of trapping indices and CO_2 injected volume over time.

During DoE process, defined input parameters to be used in ANN model are randomly generated. Among parameters collected from numerical realizations were porosity (φ), k, kv/kh, residual CO_2 saturation S_{gr}, injection rate, and elapsed time. In addition, trapping indices (RTI, STI, and StTI) and CO_2 injected volume were collected as output to ANN model. Collected data will be used to feed the ANN model during training, testing, and validation process.

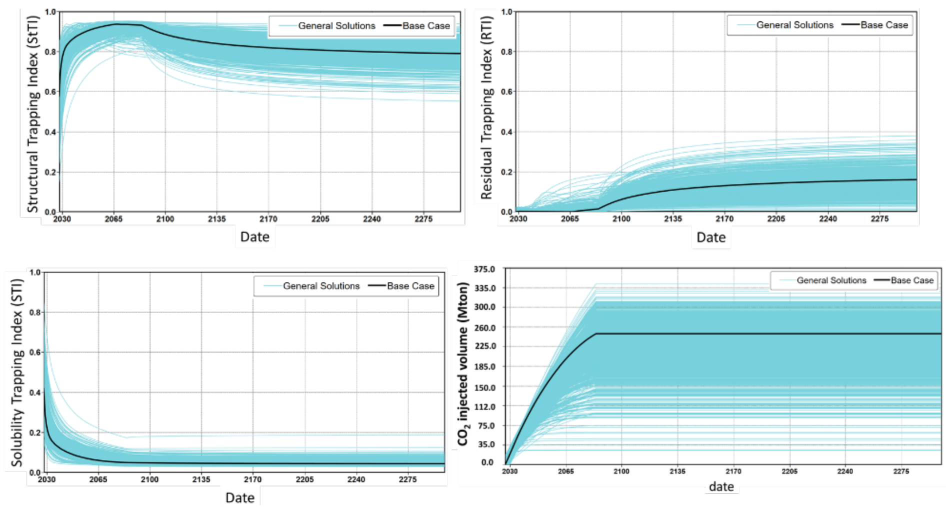

Figure 3.1. Trapping indices and injected volume profile distribution from DoE process.

Figure 3.2 illustrates the range of CO_2 cumulative injected volume from the DoE process expressed in a histogram. Considering 95% of numerical experiments, the average cumulative volume for 100-years of injection was 329 Mt (ranging from 115 to 560 Mt) respectively.

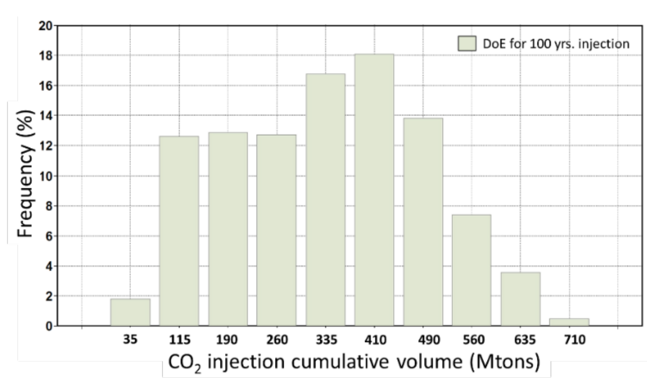

Figure 3.2. CO_2 cumulative injected volume histogram from DoE processes.

Regarding ANN model, an optimal performance was achieved for RMSE 0.0034 and R^2 0.998. In the other hand, the comparison between training data and predicted data with testing data and respective predicted data resulted in R^2 of 0.96 and 0.95 respectively, and RMSE of 0.0033 and 0.004 respectively.
Figure 3.3 shows trapping indices and CO_2 cumulative volume profiles for 12 elapsed times (from 1 to 272 years) which were predicted by the ANN model to each one of the Bunter Closure sites for 100-years injection period. A gradual reduction or steady increase of respective trapping indices is observed which behaviors may be delayed for longer injection time until their stabilization over time in comparison with original 56-year injection period case. On the other hand, structural trapping (StTI) is the most significant trapping mechanism during and after injection stops, which trend declines gradually until stabilizes over time. Moreover, the solubility process stabilizes, and residual trapping steadily increases over time. Worthy to mention that CO_2 cumulative injected volume variation is strongly related to aquifer reservoir properties and constrains related to operating conditions (injection rate and pressure limit).

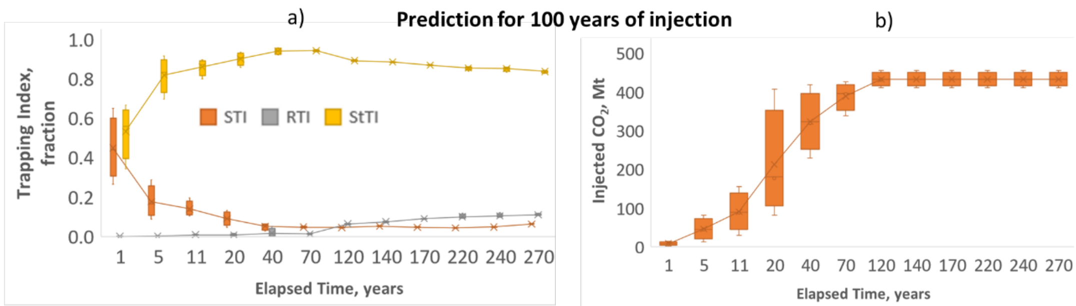

Figure 3.3. a) Trapping indices, b) CO_2 injected volume predictions by the GCS-ANN model for 100-years injection.

Table 3.1 summarizes average trapping indices at the end of year 2300 for the 100-years injection case predicted by ANN model and they are compared with indices from the original 56-years injection case. It appears that the later the injection stops the greater the delay in trapping processes stability.

Injection Scenario	Average Solubility Trapping Index, % @ end period (year2300)	Average Residual Trapping Index, % @ end period (year2300)	Average Structural Trapping Index, % @ end period (year2300)
56 years	5	18	77
100 years	6	11	83

Table 3.1. Average trapping indices comparison at end-period.

On the other hand, CO_2 injected volume (dynamic storage capacity) of selected sites were predicted for a 100-year injection period and are shown in Table 3.2. Results indicate that about 70% of theoretical volumetric capacity is reached at 100-year injection period.
Based on these results, selected aquifer sites from field validation process may be considered preliminarily de-risked by using the new ANN-GCS model and it can be said that up to 2.1 Gt of CO_2 may be accommodated and classified as ready-to-be-considered as potentially feasible storage sites in UK.

Aquifer Site*	Theoretical Volumetric Capacity * (Mton)	end-period CO2 injected volume. 100-yrs injection (Mton)
Bunter Closure 3	409	432
Bunter Closure 9	1691	432
Bunter Closure 35	554	456
Bunter Closure 36	232 (350 ***)	397 (398 **)
Bunter Closure 40		411
TOTAL	**3,004**	**2,128**

* Taken from CO2Stored® database

** from numerical simulation

*** Volumetric capacity from static model

Table 3.2. CO_2 injected volume for 100-years injection case vs. theoretical volumetric capacity.

Finally, an execution time comparison between the traditional numerical reservoir modelling and the approach of combining both numerical and ML model (ANN model) is shown in Table 3.3, from which just the reservoir numerical simulation job for 5 similar sites could take 1 year using the same hardware used in this study, and the combination of both technologies (ML + numerical simulation) could take 4 months, hence 8 months were saved which may represent 66% of time and cost saving considering a mostly labor-based cost model.

Approach	Aquifer base model design	DoE, 1000 realization	ML Design	5 similar sites modeling	Total
Numerical	2.0	-	-	~ 10.0	12.0
Numerical + Machine Learning	2.0	1.0	1.0	minutes	4.0

A Laptop 11th Gen Intel® Core i7-11800H 16 GB processor was used in this study.

Table 3.3. Execution time (months) comparison between both numerical and numerical +ML approaches

Discussion

One thing to highlight from the results is that stabilization of every trapping index is lagged as injection period is active. Once injection stops, it starts an increase in residual trapping and gradual reduction in structural trapping mechanism as time goes by.

Regarding cumulative injected volume during simulation period, the most rapid pace in cumulative volume trend occurs for the first decades (~50 years), period after which well injection rates drop drastically as reservoir pressure is closer to fracture pressure limit. Therefore, greater CO_2 cumulative volume occurs as reservoir pressure is maintained far from caprock fracture pressure.

Reaching the theoretical volumetric capacity will depend on geologic properties, operating conditions, and injection time. Hence, larger closures will require pressure management-based brine production to reach theoretical storage capacity.

Conclusions

• The developed ML-based model using ANN showed 95% performance and predicting time variations of trapping indices and CO_2 cumulative injected volume (dynamic storage capacity) for dome-like shape deep saline storage aquifers in Bunter sandstone formation at Southern North Sea area in UK. It is expected that this approach can address the lack of numerical models for similar aquifers in the area.

• For the reservoirs used in this study, the model saved ~8 months of labor hours, resulting in ~ 66% cost reduction. And demonstrated that up to 2100 CO_2 Mt can be storage over at least 100 years of injection period.

• This type of model is a useful tool to quickly predict with good approximation the CO_2 storage performance in saline aquifers and other types of geological storage sites at an early stage of a UCS project just by using basic geologic data.

References

[1] Progressing Development of the UK's Strategic Carbon Dioxide Storage Resource. *Energy Technologies Institute (ETI)*. 2016.

[2] Mohaghegh S., et al. Grid-Based Surrogate Reservoir Modeling (SRM) for Fast Track Analysis of Numerical Reservoir Simulation Models at the Grid block Level. *SPE Western Regional Meeting*. California, USA. 2012. Paper SPE-153844.

[3] Misra S. Machine Learning Tools for Fossil and Geothermal Energy Production and Carbon Geo-Sequestration. *Circular Economy and Sustainability.* Springer Nature, 2021.

[4] Industrial Carbon Dioxide Emissions and Carbon Dioxide Storage Potential in the UK. *British Geological Survey (BGS)*. DTI 2006. Report No. COAL R308, DTI/Pub, URN 06/2027.

[5] A picture of CO_2 Storage in the UK. Learnings from the ETI's UKSAP and derived projects. *Energy Technologies Institute (ETI)*. 2016. https://www.eti.co.uk

[6] Strategic UK CCS Storage Appraisal Project. D10: WP5A - Bunter Storage Development Plan. 10113ETIS-Rep-13-03, 2016. Energy Technologies Institute (ETI). 2016.

[7] Nghiem L., et al. Modelling CO_2 Storage in Aquifer with a Fully-Coupled Geochemical EOS Compositional Simulator. *SPE/DOE Fourteenth Symposium on Improved Oil Recovery*, 2004, Tulsa, Oklahoma, USA. Paper SPE-89474.

[8] Nghiem L., et al. Simulation of CO_2 Storage in Saline Aquifers. *SPE Reservoir Characterization and Simulation Conference*, Abu Dhabi, UAE, 2009. Paper SPE-125848.

[9] Holloway S., et al. The Potential for Aquifer Disposal of Carbon Dioxide in the UK. *Energy Conversion and Management,* 1993. Volume 34, No. 9-11, pages 925-932.

[10] Noy D., et al. Modeling large-scale carbon dioxide injection into the Bunter Sandstone in the UK Southern North Sea. *International Journal of Greenhouse Gas Control*. 9 (2012), 220-233.

[11] Song Y., et al. Application of an artificial neural network in predicting the effectiveness of trapping mechanisms on CO_2 sequestration in saline aquifers. *International Journal of Greenhouse Gas Control*, Elsevier, 2020. 98 103042. https://doi.org/10.1016/j.ijggc.2020.103042

[12] Vo Thanh H., et al. Application of machine learning to predict CO_2 trapping performance in deep saline aquifers. *Energy*, Elsevier, 2022. 239 122457. https://doi.org/10.1016/j.energy.2021.122457

[13] Schuetter J., et al. Building Statistical Proxy Models for CO_2 Geologic Sequestration. *Energy Procedia 63* (2014) 3702-3714.

[14] CO2Stored®. http://www.co2stored.co.uk