PAPER • **OPEN ACCESS**

Machine learning for underground gas storage with cushion CO_2 using data from reservoir simulation

To cite this article: J O Helland *et al* 2023 *IOP Conf. Ser.: Mater. Sci. Eng.* **1294** 012058

Machine learning for underground gas storage with cushion CO_2 using data from reservoir simulation

J O Helland[1,*], H A Friis[1], M Assadi[1] & S Nagy[2]

[1] University of Stavanger, Stavanger, Norway
[2] AGH University of Science and Technology, Kraków, Poland

E-mail: *johan.o.helland@uis.no, jhel@norceresearch.no

Abstract. Underground natural gas storage (UNGS) is a means to store energy temporarily for later recovery and use. In such storage operations, carbon dioxide (CO_2) can be injected as cushion gas to improve the operating efficiency of the working gas and then be permanently stored in the same reservoir. A potential obstacle for widespread use of this technology is that the mixing of the different gases can lead to undesired CO_2 production. Herein, we use a two-component flow model to simulate injection and withdrawal periods of methane (CH_4) in idealized reservoirs containing CO_2. First, we simulate cases with a single well for both CH_4 injection and production. From 1200 simulations with systematic variation of reservoir temperature, porosity, permeability, height, and injection time, we find that the reservoir height and permeability have the most significant impact on the production time until the well stream reaches 1% mole fraction of CO_2. In another set of simulations, we investigate the impact of well spacing in seasonal gas storage scenarios with separate wells for CH_4 injection and production, while CO_2 injection occurs from a third well. Based on the simulated data we construct artificial neural networks (ANNs) that describe the relations between the varied input parameters and the production time of CH_4, well-block mole fraction and pressure. We conclude that trained and validated ANN models are useful tools to optimize important parameters for UNGS operations, including well positioning, with the aim at maximizing the amounts of delivered gas.

1. Introduction

Natural gases, e.g., methane (CH_4), are regarded as one of the cleanest energy resources among the fossil fuels [1]. In underground natural gas storage (UNGS) the gas is stored temporarily in salt caverns, reservoirs, or aquifers, before it is produced upon demand (e.g., for electrical heating in winter seasons) and to maintain a steady energy supply [1, 2]. Storage operations in an underground reservoir can benefit from using a cushion gas to improve the operating efficiency of the working gas by contributing with pressure support [2]. The compressible properties of carbon dioxide (CO_2) are an advantage for a cushion gas as it allows for large storage capacities of natural gas in the reservoir, while its expansion to lower pressure aids the gas production [3]. Another potential advantage is that injected CO_2 can be stored permanently in the reservoir after it served as a cushion gas for UNGS. However, in such storage strategies a disadvantage is that the mixing of gases in the reservoir can lead to undesired CO_2 production during continued seasons of injections and withdrawals.

In the literature, reservoir simulation studies of UNGS with cushion gas focus on demonstrating the advantageous density changes of CO_2 near its critical point in idealized reservoirs [3]; the mechanical responses to pressure build-ups in idealized, yet inclined reservoirs

[4]; and simulations of working gas quality caused by mixing with cushion gas [5]. While the entry of data-driven methods has proven useful in research on CO_2 storage [6], there are to date few data-driven methodologies demonstrated for UNGS with cushion CO_2 in reservoirs or aquifers. In the literature on UNGS, most data-driven approaches are concerned with describing the relations between gas deliverability (that is, the amount of gas that can be withdrawn per day [2]), the flowing bottom hole pressure (BHP) in the well, and reservoir pressure [7, 8]. On the other hand, Mann III & Ayala [9] developed an artificial neural network (ANN) model aimed at determining the optimum design of a storage facility, but their model did not include information on which gas was considered as cushion gas, nor the complete reservoir conditions, both of which are important for the behaviour of the gases and their mixtures.

Herein, we perform systematic reservoir simulations of UNGS at reservoir conditions with cushion CO_2 in idealized reservoir geometries, inspired by Oldenburg [3] and Ma et al. [4]. The set of simulations generates data that we use to develop ANN models that learn parameter relationships for seasonal storage of CH_4 with cushion CO_2. Our objective is twofold: First, we show that ANN models can predict production time until the CO_2 mole fraction exceeds 1% in the well block (that is, the grid block where the bottom hole of the well is located). Secondly, for seasonal storage with fixed injection and production periods, we develop ANN models that can describe the CO_2 mole fraction in the well block after each seasonal storage cycle. The prediction of these parameters under different well configurations and injection/production schedules contribute to the optimization of UNGS operation, with the aim at minimizing costly and undesired CO_2 production.

2. Methods

Reservoir simulation was used to generate data for seasonal CH_4 storage scenarios in idealized 2D reservoirs initially containing CO_2. Using the simulation results, artificial neural network (ANN) models were developed to describe relations between input parameters and calculated parameters, such as production time, well-block mole fraction and well-block pressure.

2.1. Reservoir models

For the reservoir simulations we use Matlab Reservoir Simulation Toolbox (MRST) [10] and its *Compositional* module [11]. MRST is free, open-source software written in Matlab and the MRST module *Compositional* includes some well-known solution strategies for compositional reservoir simulations. The Peng-Robinson equation of state (EoS) and Lohrenz-Bray-Clark correlation for viscosity [12], with industry standard corrections that account for mixing behaviour of gases, are default settings and are also used in the simulations carried out in this work. The compositional module of MRST has previously been benchmarked against the commercial reservoir simulator Eclipse 300, as well as AD-GPRS (the Stanford University research simulator) [11].

We use a two-component model with gravity to simulate various seasonal CH_4 storage scenarios in idealized, vertically aligned, rectangular 2D reservoirs, initially containing CO_2. 2D setups were used in this study, because 2D simulations are less computationally costly than 3D simulations, and hence it allows us to explore a more comprehensive set of parameter variations with simulations in shorter time. The aim here is thus not to develop ANN models that describe storage in a real reservoir, but rather to build confidence in ANN models as useful tools to describe relations between influential parameters in UNGS with cushion CO_2. All simulations assume the initial reservoir pressure is 6 MPa. During injection the pressure rises above the critical point where CO_2 exists as a supercritical phase. Hence, the simulations account for conditions where advantageous, large density changes of CO_2 occur with pressure [3]. The reservoirs have fixed length 1000 m, and we always assume vertical permeability is 10% of horizontal permeability, $K_v = 0.1K_h$, which is a reasonable assumption in non-fractured,

laminated rock [13]. All simulations use vertical wells with radius 10 cm, while all other reservoir boundaries were sealed. Injection and production rates are specified at standard surface conditions (temperature 273.15 K and 1 atm pressure). For this setup, we carried out simulations with different grid block resolutions and maximal time steps and found that grid blocks of size $5 \times 1 \times 1$ m^3 (assuming depth of the 2D reservoir is 1 m) and maximal time step of 3 days were suitable for the investigations. These results were also validated against the commercial simulator GEM [14].

2.2. Artificial neural network models

To tailor the setup of feed-forward neural networks with backpropagation, data generated from the reservoir simulations were utilized. Due to its flexibility in design, it was decided to use *TensorFlow* and the high-level *Keras* library installed in a *Python* environment for the ANN model development. Keras is an open-source library that provides a Python interface for neural networks, and it acts as an interface for the TensorFlow library. Within this code framework we have high flexibility in determining the structure of the ANN and in training of the model.

When developing the optimal choice of the ANN structure, we started with a small number of hidden layers, and a small number of neurons in each layer. Then the networks were trained by systematically increasing the number of neurons and layers until no significant improvement in the results was observed. For each network the choice of activation functions, optimization method, and specifications of the learning rate were altered. We also investigated the deactivation of a fraction of neurons in each layer (called "dropout fraction") and the use of regularization weights in every layer based on a metric (for example L_2-norm) that will add a penalty to the total loss function. These latter functionalities are useful means to avoid underfitting and overfitting in ANN models.

Based on this experimentation, ANN models with two hidden layers and a relatively large number of neurons were developed. The model specifications also include the activation functions recified linear unit (*ReLu*) and tanh, whereas the use of dropout fractions did not improve the results. The ANN training also made use of a slightly declining learning rate and L_2-regularization. As optimization method we used Adam optimizer while mean squared error was used both as metric in the training and as loss function.

3. Results

We begin by describing two scenarios of UNGS with cushion CO_2 and present results from the corresponding reservoir simulations. These scenarios are outlined in Section 3.1.1 (Case 1) and Section 3.1.2 (Case 2). Then we proceed with presenting the corresponding ANN models for these cases and their performance.

3.1. Reservoir simulations

3.1.1. Case 1: One-well CH_4 injection & withdrawal cycle. Case 1 has one well in the upper left corner of the reservoir (3 m below the top seal, see Figure 1) which is used to simulate CH_4 injection followed by production with a rate of 0.015 m^3/s. Production ceases when well-block CO_2 mole fraction exceeds 1%, so that production time is different from injection time. The impact of reservoir thickness (vertical distance) H, horizontal permeability K_h, porosity ϕ, temperature T, and injection time t_{inj}, were studied by varying these parameters independently in different simulations. Figure 1 shows the distribution of CH_4 at a time during injection and production from one such simulation. The lighter gas, CH_4, tends to stay above CO_2 while a mixing zone of both gas components develops in between. Figure 2 shows the evolution of well-block pressure and well-block CH_4 mole fraction over time for this simulation. The first part of the production exhibits a time interval in which the well primarily produces CH_4, whereas at later times the well-block CH_4 mole fraction declines, leading to significant amounts of undesired

CO_2 production. We also note that the well-block pressure increases steadily during injection, but in the beginning of the production stage it decreases rapidly to enable flow toward the well from distant zones in the reservoir.

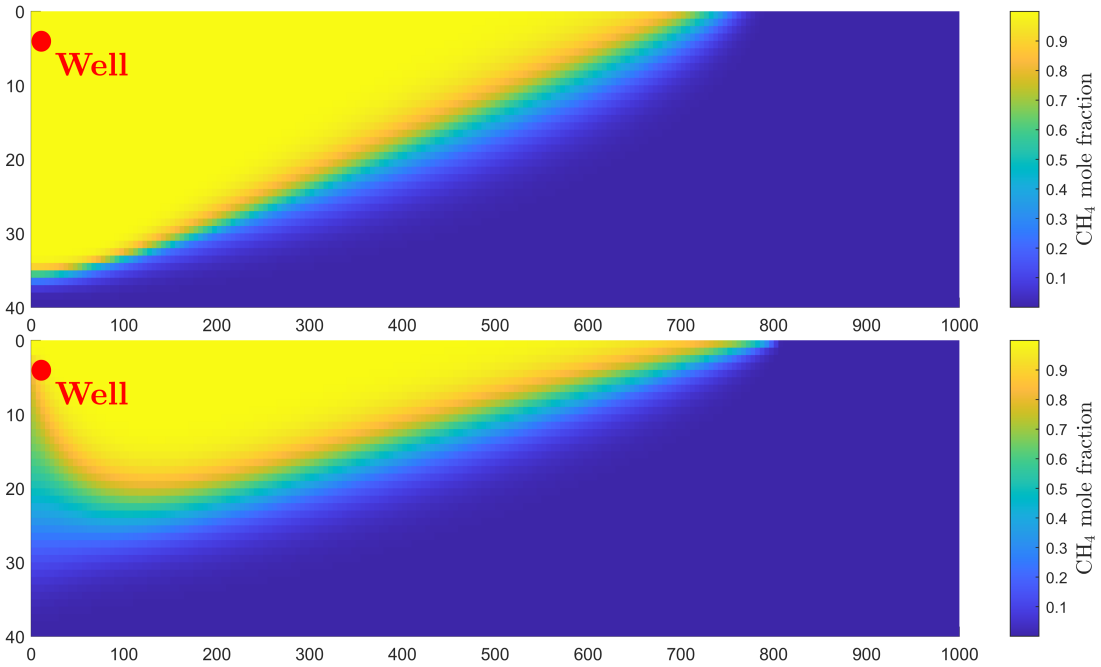

Figure 1. CO_2 and CH_4 distribution at 90 days (*top*) and 240 days (*bottom*) from a typical Case 1 simulation with $T = 323.15$ K, $H = 40$ m, $\phi = 0.2$, $K_h = 1000$ mD, and $t_{inj} = 180$ days.

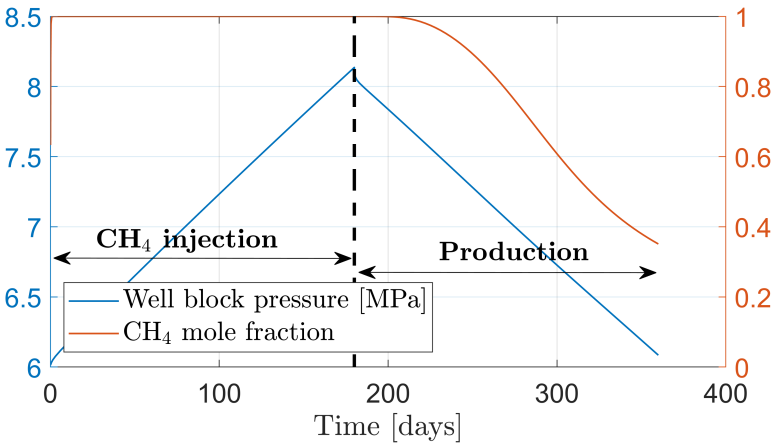

Figure 2. Pressure and CH_4 mole fraction in well block over time from a typical Case 1 simulation with $T = 323.15$ K, $H = 40$ m, $\phi = 0.2$, $K_h = 1000$ mD, and $t_{inj} = 180$ days.

Table 1 shows the range and resolution of the parameters varied independently in the simulations. Thus, a total number of 1200 ($= 4 \times 5 \times 4 \times 5 \times 3$) simulations were carried out with different parameter combinations. To identify the most influential parameters on the simulated well-block pressure and well-block mole fraction, a simulation with midrange parameters $T = 328.15$ K, $H = 60$ m, $\phi = 0.155$, $K_h = 700$ mD, and $t_{inj} = 180$ days was conducted. The midrange parameter simulation was then compared with new simulations where the midrange value of one parameter at a time was substituted with its maximum and

Table 1. List of parameter values varied independently in the MRST simulations.

Temperature T [K]	Reservoir thickness H [m]	Porosity ϕ [frac.]	Horizontal permeability K_h [mD]	Injection time t_{inj} [days]
313.15	30	0.11	100	150
323.15	40	0.14	300	180
333.15	50	0.17	500	210
343.15	70	0.20	900	–
–	90	–	1300	–

minimum value. Among the varied parameters, the results show that the reservoir thickness H and horizontal permeability K_h have the most significant impact on the well-block mole fraction, particularly in the interval between the minimum and midrange values, see Figure 3. Consequently, additional simulations were carried out with variation of H and K_h in this interval, so that the step size for the variation of these parameters is uneven, see Table 1.

From the set of simulations, we saved to file the five varied input parameters in addition to the maximum pressure, average reservoir pressure and average CH_4 mole fraction before and after production, well-block pressure after production, and production time (i.e., the time until well-block CO_2 mole fraction exceeds 1%).

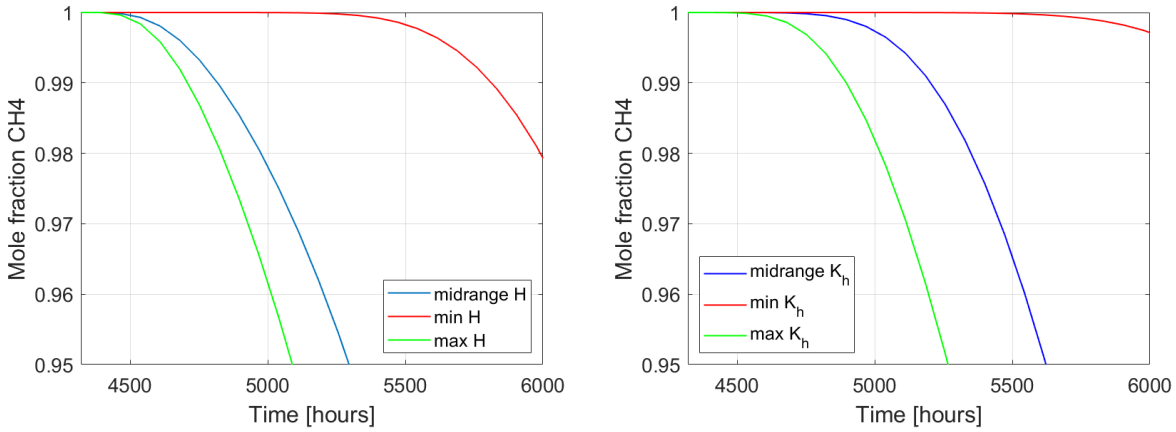

Figure 3. Well-block mole fraction of CH_4 during production from a Case 1 simulation with midrange parameter values compared to Case 1 simulations with maximum and minimum values of reservoir thickness H (*left*) and horizontal permeability K_h (*right*).

3.1.2. Case 2: Seasonal CH_4 injection & withdrawal cycles with continuous CO_2 injection using different wells. Case 2 explores the behaviour in seasonal storage cycles when CH_4 injection and production occur from two different wells, while continuous CO_2 injection occurs from a third well. This is a scenario for using the reservoir both as a facility for seasonal gas storage with cushion CO_2 and as a site for permanent CO_2 storage. In this case we set $H = 40$ m, $K_h = 1000$ mD, $\phi = 0.2$, and $T = 323.15$ K. The location of the CO_2 injection well is 37 m below the top seal at left boundary, the location of the CH_4 injection well is 3 m below the top seal at a distance $x = d_W$ from the left boundary, while the location of the production well is at the right boundary, 3 m below the top seal. This setup is illustrated in Figure 4 (top). The

schedule in the simulations is as follows: (1) initial CH_4 injection with rate 0.025 m^3/s for 360 days, (2) 180 days with all wells closed which allows some gravitational segregation, and (3) three cycles of CH_4 injection (180 days) and withdrawal (180 days) with rate 0.015 m^3/s while continuous CO_2 injection occurs from the third well. We studied the impact of CO_2 injection rate ($Q_{CO2} = 0$, 0.005 and 0.010 m^3/s) and CH_4 injection well position ($d_W = 200$, 400, 600 and 800 m) by varying these parameters independently, which constitutes 12 ($= 3 \times 4$) simulations.

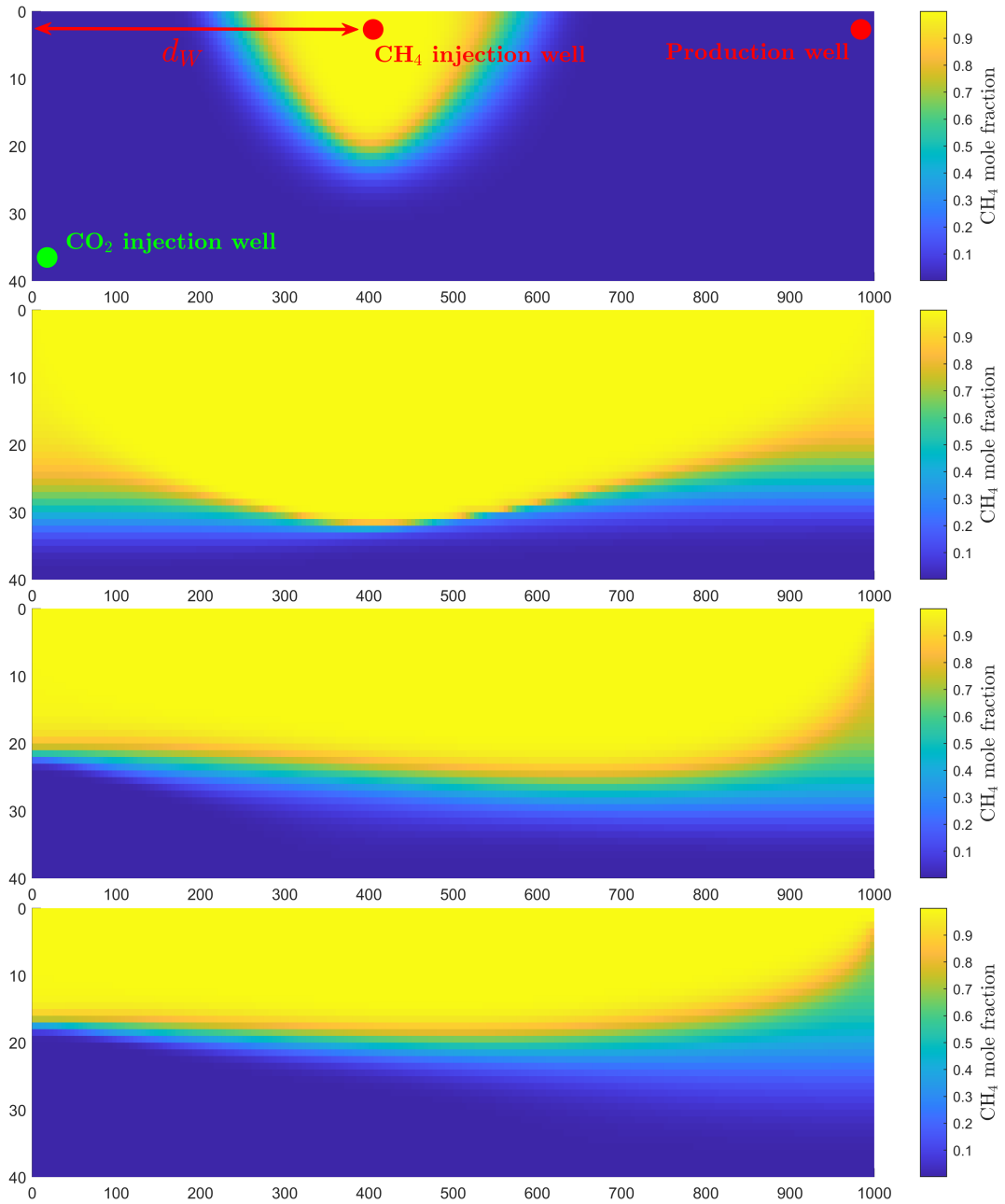

Figure 4. CH_4 distribution for Case 2 with $Q_{CO2} = 0.01$ m^3/s and $d_W = 400$ m at four time slots (ordered from top to bottom): during initial CH_4 injection (30 days), after the initial CH_4 injection (360 days), after the first seasonal production (900 days), and after the third seasonal production (1620 days).

Figure 4 shows images of the CH_4 distribution in the reservoir at four times, for the case with $Q_{CO2} = 0.010$ m^3/s and $d_W = 400$ m. After the initial CH_4 injection, the most significant CH_4/CO_2 mixing zones take place at the flanks. During CO_2 injection, the mixing zone at the left boundary diminishes while the mixing zone at the right boundary, where production occurs, expands slightly. Figure 5 shows the impact of Q_{CO2} on pressure and mole fraction in the production well-block, and the impact of injection well position d_W on the mole fraction in the production well-block. The impact of well position d_W on the pressure in the production well-block is insignificant and is not shown here. The results show, as expected, that a continuous CO_2 injection leads to an undesired pressure build-up over time which gets worse with injection rate and subsequent seasonal injection/withdrawal cycles. A significant increase in pressure could damage the reservoir rock, alter its permeability [4], and potentially lead to gas leakage. With respect to mole fraction, an increased CO_2 injection rate has a marginal positive effect in the first cycle, as it results in lower production of CO_2, while the subsequent seasonal cycles show more severe negative effects as the amount of CO_2 production increased significantly. Further, the simulations with varying well position shows that a decreased spacing between CH_4 injection and production wells leads to the least CO_2 production during seasonal cycles. Hence, the optimal setting is to use the same well for production and injection.

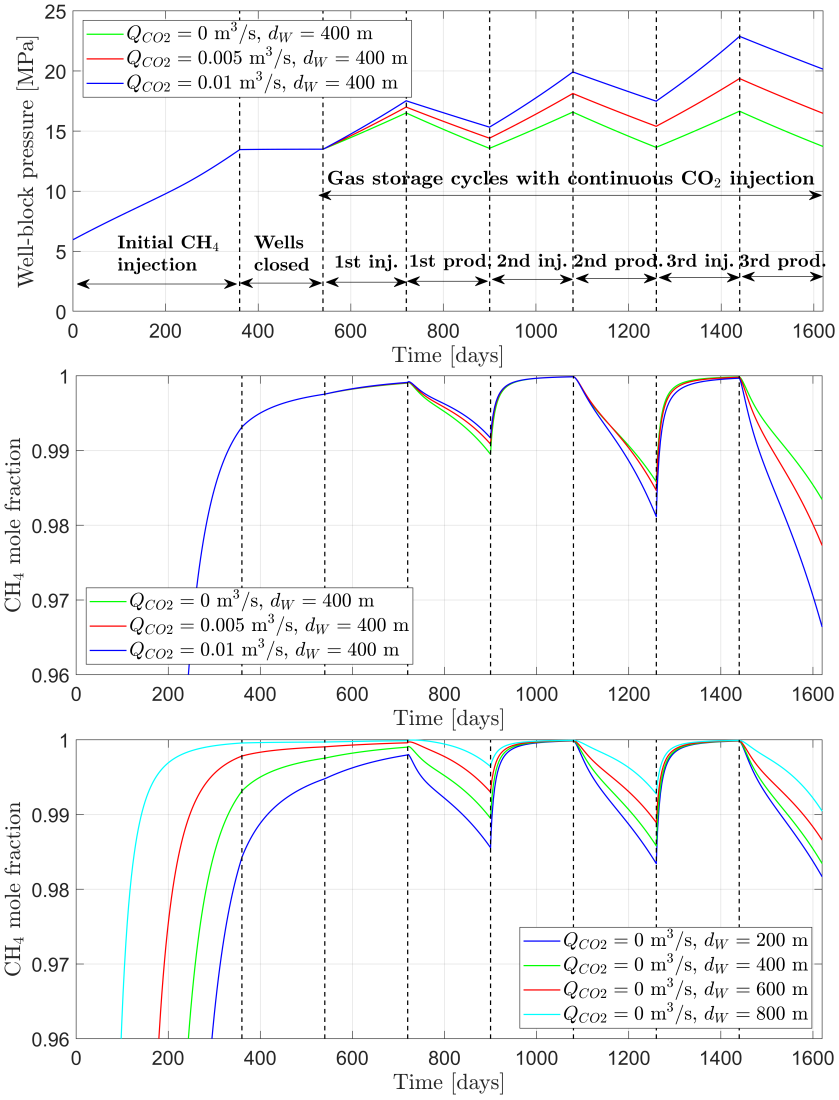

Figure 5. Well-block pressure (*top*) and well-block CH_4 mole fraction (*middle*) over time for Case 2 simulations with different CO_2 injection rates Q_{CO2} and fixed CH_4 injection-well position $d_W = 400$ m. *Bottom:* Well-block CH_4 mole fraction over time for Case 2 simulations with $Q_{CO2} = 0$ m^3/s and different d_w. Vertical lines (*black, dashed*) indicate the different injection/production schedules.

3.2. ANN results

ANN models for Case 1 and Case 2 were developed separately. The ANN model for Case 1 aims at predicting the production time until well-block CO_2 mole fraction exceeds 1%, while for Case 2 where the production time is fixed, the ANN model aims at predicting the CO_2 mole fraction in the production well-block after each seasonal withdrawal.

3.2.1. ANN model for Case 1. The ANN model for Case 1 takes as input the five varied parameters from the simulations (see Table 1). Before training the ANN model, it is crucial to normalize all the data to the range $[0, 1]$. Normalization is important for stability of the training, as the data varies significantly in magnitude between the parameters. We also shuffle the 1200 data sets from the simulations randomly, with possibility for replicating the shuffle if desired. After the shuffle, the first 75 data sets were taken as the validation set, the next 75 data sets as testing set, and the remainder (1050 data sets) as the training set. In the optimized ANN model structure for Case 1, the following number of neurons were obtained: 5 (input layer) – 42 (hidden layer) – 15 (hidden layer) – 1 (output). We used *ReLu* as activation function, a slightly declining learning rate, and L_2-regularizer penalty coefficient of 10^{-4}. The model was trained for 200 epochs. Figure 6 shows the ANN model's excellent performance on the training data, validation data, and test data sets.

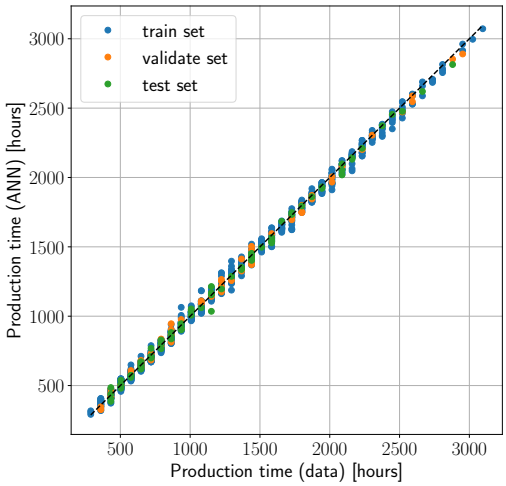

Figure 6. ANN calculation of production time until the well-block mole fraction exceeds 1% CO_2 versus corresponding data sets from training, validation, and testing.

3.2.2. ANN model for Case 2. For Case 2 only the CO_2 injection rate and the CH_4 injection-well position were varied. In addition, each simulation provided results for three seasonal injections and productions. Thus, for the Case 2 ANN model the injection rate (Q_{CO2}), well position (d_W), and the seasonal cycle number (1, 2 or 3) were considered as inputs. These yield 36 ($= 3 \times 4 \times 3$) data sets, all of which were used for training the ANN model. The output is the CO_2 mole fraction in the production well-block.

In the optimized ANN model structure for Case 2, the following number of neurons were obtained: 3 (input layer) – 21 (hidden layer) – 5 (hidden layer) – 1 (output). We used *ReLu*, tanh and *ReLu* as activation functions in the hidden layers and the output layer, a slightly declining learning rate, and L_2-regularizer penalty coefficient of 10^{-3}. This model was trained for 600 epochs. Figure 7 (left) shows the ANN model's excellent performance on the training data. To evaluate its performance on other data, we performed two new MRST simulations for different combinations of Q_{CO2} and d_W: $Q_{CO2} = 0.0075$ m^3/s, $d_W = 300$ m, and $Q_{CO2} = 0.0025$ m^3/s,

$d_W = 700$ m. Figure 7 (right) shows that the ANN model predicts the mole fraction at the end of each seasonal cycle very well for both simulations.

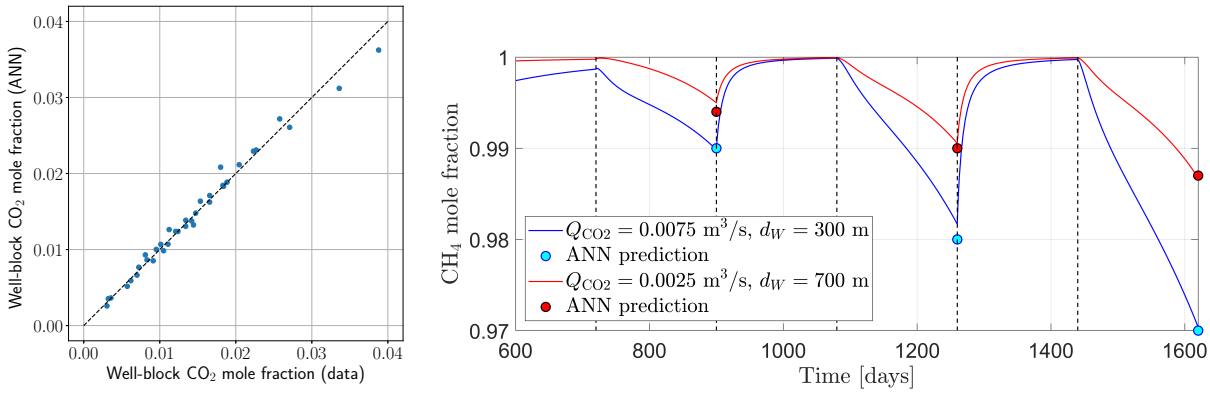

Figure 7. *Left:* ANN calculation of well-block CO_2 mole fraction for Case 2 versus corresponding data used in the training. *Right:* Performance of ANN model for well-block mole fraction when it is applied to two new simulations of Case 2.

Finally, the ANN model for Case 2 was expanded slightly by also taking pressure in the production well-block as output. For this problem, the optimal ANN structure included more neurons in each layer: 3 (input layer) – 27 (hidden layer) – 15 (hidden layer) – 2 (output). This time the L_2-regularizer penalty coefficient was set to 10^{-4}. Otherwise, the ANN specifications remained unchanged. This model was also trained for 600 epochs.

Figure 8 shows the ANN model's excellent performance on the training data set with respect to describing simultaneously the well-block CO_2 mole fraction and the well-block pressure accurately. The ANN model was also compared against the two new MRST simulations for validation. Figure 9 shows an excellent agreement between these simulations and the ANN model with respect to both pressure and mole fraction.

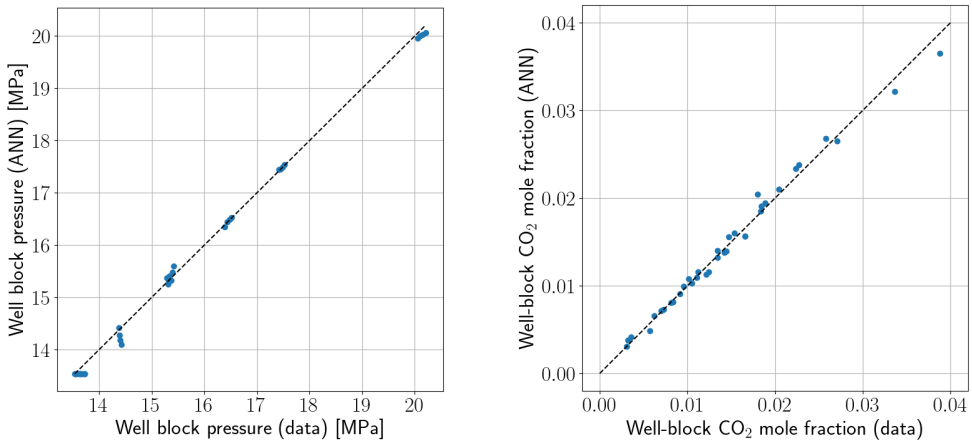

Figure 8. Simultaneous ANN calculation of well-block pressure (*left*) and CO_2 mole fraction (*right*) versus the data used in the training (from Case 2).

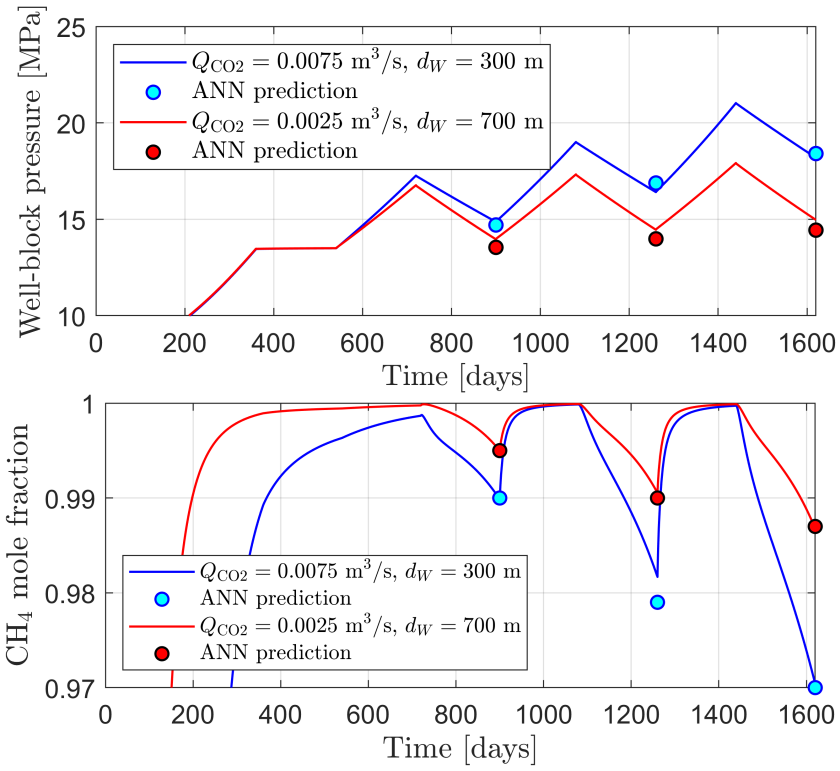

Figure 9. Performance of the ANN model for Case 2 that calculates both well-block CO_2 mole fraction and well-block pressure simultaneously. ANN predictions of well-block pressure (*top*) and well-block mole fraction (*bottom*) compared with two new test simulations of Case 2.

For the ANN model of Case 2 which takes both well-block mole fraction and well-block pressure as outputs, the RMSEs are 1.6×10^{-3} for mole fraction and 0.32 MPa for pressure when compared against the new test simulation with $Q_{CO2} = 0.0075$ m^3/s and $d_W = 300$ m, while the RMSEs are 3.37×10^{-4} (mole fraction) and 0.48 MPa (pressure) with respect to the other test simulation with $Q_{CO2} = 0.0025$ m^3/s and $d_W = 700$ m. In comparison, the ANN model that only takes well-block mole fraction as output obtains RSMEs of 9.95×10^{-4} and 6.6×10^{-4} for the first and second test simulation, respectively. Hence, accuracy is maintained in the ANN model when increasing the number of outputs, albeit at the expense of increasing the number of neurons in the hidden layers in the ANN. This shows promise in expanding the data-driven models with higher numbers of inputs and outputs, accompanied by more data for training and validation, to increase their predictive capability and utility value for seasonal UNGS with cushion CO_2.

4. Conclusions

Reservoir simulation and ANN models were used to explore seasonal cycles in UNGS with cushion CO_2. Our conclusions are as follows:

- To avoid CO_2 production, it is paramount to have a significant amount of CH_4 present in the reservoir before seasonal injection & withdrawal cycles begin.
- Decreased spacing between CH_4 injection & production wells leads to the least CO_2 production during seasonal cycles. Hence, the optimal setting is to use the same well for production and injection.
- Continuous CO_2 injection during UNGS has a marginal positive effect in the first cycle, while later cycles bear disadvantages with increased CO_2 production and undesired pressure build-up that worsen with increased CO_2 injection rate.
- Trained ANN models capture the behaviour of critical parameters that describes undesired

CO_2 production during UNGS. Hence, ANN models are useful tools to optimize UNGS operations.

Future work should supplement the data generated here with simulations of other schedules for CO_2 injection in conjunction with seasonal CH_4 injection and production. It will also be needed to expand the ANN models with other output parameters. For example, with the addition of well-block pressure and average reservoir pressure, the ANN model could contribute to establishing their relations to well deliverability [7, 8]. It is also desirable to explore deep-learning methods to describe the CH_4/CO_2 distribution in the reservoir, in particular the extent of mixing zones, potentially by utilizing methods developed for CO_2 storage that describe the migration of CO_2 plumes in sandstone formations [6].

Acknowledgments
This research was funded by EEA and Norway Grants, operated by The National Center for Research and Development, grant number NOR/POLNORCCS/AGaStor/0008/2019-00.

References
[1] International Energy Agency (IEA) Fuel & technologies, gas Available online URL `https://www.iea.org/fuels-and-technologies/gas`
[2] US Energy Information Administration (EIA) The basics of underground natural gas storage. natural gas report November 16, 2015. Available online URL `https://www.eia.gov/naturalgas/storage/basics/`
[3] Oldenburg C M 2003 *Energy & Fuels* **17** 240–246
[4] Ma J, Li Q, Kempka T and Kühn M 2019 *Energy & Fuels* **33** 6527–6541
[5] Sadeghi S and Sedaee B 2022 *Journal of Energy Storage* **46** 103885
[6] Chu A K, Benson S M and Wen G 2023 *Energies* **16** 246
[7] Ali A 2021 *Energy* **229** 120648
[8] Thanh H V, Zamanvad A, Safaei-Farouji M and Ashraf U 2022 *Renewable Energy* **200** 169–184
[9] Mann III A W and Ayala L F 2009 *Int. J. Modelling and Simulation* **29** 214–223
[10] Lie K A 2019 *An introduction to reservoir simulation using MATLAB/GNU Octave: User guide to the MATLAB Reservoir Simulation Toolbox* (Cambridge University Press)
[11] Lie K A and Møyner O (eds) 2021 *Advanced Modeling with the MATLAB Reservoir Simulation Toolbox* (Cambridge University Press) chap 8. Compositional simulation with the AD-OO framework, pp 324–374
[12] Lohrenz J, Bray B G and Clark C R 1964 *Journal of Petroleum Technology* 1171–1176
[13] Fanchi J R 2010 *Integrated Reservoir Asset Management* ed Fanchi J R (Boston: Gulf Professional Publishing) pp 49–69 ISBN 978-0-12-382088-4
[14] CMG 2015 CMG GEM User's Guide Computer Modeling Group Ltd., Calgary, Alberta, Canada