Data-driven Proxy at Hydraulic Fracture Cluster Level: A Technique for Efficient CO₂-Enhanced Gas Recovery and Storage Assessment in Shale Reservoir

Amirmasoud Kalantari-Dahaghi*, Shahab Mohaghegh, Soodabeh Esmaili
West Virginia University, United States

E-mail addresses: akalanta@mix.wvu.edu, Amirmasoud.KalantariDahaghi@CRC.com (A. Kalantari-Dahaghi),
Shahab.Mohaghegh@mail.wvu.edu (S. Mohaghegh),
Soodabeh.Esmaili@CRC.com, semaili@mix.wvu.edu (S. Esmaili).

Abstract

The continuing development of the organic-rich and extremely low permeability shale reservoirs in the United States has the potential to positively impact the future of carbon storage. Due to the unique characteristics of shale reservoirs, not only can CO₂ be safely stored, it also can be preferentially adsorbed and displace methane, leading to enhanced gas recovery.

However, CO₂ storage in depleted or nearly depleted shale formations is not completely risk free. Thus, prior to making the economic commitment to a full-field CO₂ sequestration project, a systematic analysis of the complete set of variables must be considered in the planning of a shale-CO₂ storage initiative. Numerical modeling and simulation is a robust tool that can provide an insight into how the system may operate in order to further understand the feasibility and assist in the design and operation of such a project, and to predict changes that may occur.

In order to perform a comprehensive uncertainty analysis, a large number of simulation runs are required. Designing and running simulation cases to model enhanced gas recovery and storage in shale by applying the Explicit Hydraulic Fracture modeling technique (EHF) is long and laborious, and its implementation is computationally expensive.

In this paper, a data-driven approach with pattern recognition algorithms is used to develop a new generation of a shale proxy model at the hydraulic fracture cluster level, as a replica of a reservoir simulation model. For more accurate analysis, instead of commonly used mechanistic models, a history-matched hydraulic fractured Marcellus shale pad with multiple stages/clusters is used as a base case to perform the analysis. The detailed procedure for development of the data-driven proxy model is explained and the model is validated using blind simulation runs. The developed data-driven proxy model is capable of accurately reproducing the calculated CO₂ injection, CO₂/CH₄ production profiles, and CO₂ breakthrough time from the numerical simulation model, for each cluster/stage and horizontal lateral. Joint use of the deterministic reservoir model with the data-driven proxy model can serve as a novel screening and optimization tool for the techno-economic evaluation of the CO₂-Enhanced Gas Recovery (EGR) and Storage process in shale systems.

Introduction

Over the past century, the burning of fossil fuels has increased greenhouse gas (GHG) emissions, leading to increasing concern regarding its link to global warming. CO₂, in particular, is an efficient heat-trapping gas generated from fossil fuel combustion and is responsible for more than 70% of the greenhouse effect among the other greenhouse gases (Kruger and Franklin, 2006; Zhang et al., 2009). However, fossil fuels are expected to remain a main element for supplying the world's energy in the near future, so the challenge is to find ways to reduce CO₂ emissions into the atmosphere from burning fossil fuels (Bachu, 2007).

Capturing and storing CO₂ in oil and gas reservoirs is one option for reducing greenhouse gas emissions in the atmosphere. Additionally, it has been demonstrated that CO₂ can be used for commercial-scale CO₂-EOR. Advantages of hydrocarbon reservoirs as storage reservoirs include limited exploration costs, existing effective traps and seals, known reservoir properties (porosity, permeability, pressure, temperature, overall storage capacity), and existing equipment on the surface and in the subsurface that can be reused (Meer, 2005; Aydin, 2010).

With the relatively recent realization that thermally-matured organic rich shales are prolific reservoirs of natural gas, these unconventional resources are now being evaluated for CO₂ storage. In shale gas reservoirs, natural gas exists as free gas in the pores and open/partially open natural fractures and also as adsorbed phase on clay and kerogen surfaces. Organic black shale reservoirs may behave similarly to coalbed methane by having high affinity to adsorb CO₂ and displace methane at a ratio of two to one, or more. If this is the case, organic shales may serve as an excellent sink for CO₂ and have the added benefit of serving to enhance natural gas production (Nuttall *et al.*, 2005, Schepers *et al.*, 2009, Eshakalak 2015).

As an example, the continuing development of the low-permeability, continuous, and organic-rich Marcellus Shale has the potential to be a good candidate for future of Carbon Capture & Sequestration (CCS) in the Appalachian Basin. It can act both as a seal for injection of CO₂ into sandstone and limestone formations below the Marcellus and by itself as a storage reservoir for captured CO₂ (U.S. DOE 2010).

Reservoir simulation is a groundbreaking tool for detailed analysis of production from shale and CO₂ Enhanced Gas Recovery and Storage process (CO₂-EGR&S). The base simulation model, which is properly calibrated with available production history at primary recovery phase, should be used to model the reservoir response during CO₂ injection process.

The integration of natural fracture networks and hydraulic fractures with the desorption-diffusion phenomena in the shale matrix is a challenging task in the simulation of shale gas reservoirs. Explicit Hydraulic Fracture (EHF) and Stimulated Reservoir Volume (SRV) techniques are the two most commonly used techniques for the simulation of shale gas production. EHF modeling attempts to predict the impact of hydraulic fracturing at the cluster/stage level by incorporating hydraulic fracture characteristics in the simulation model. Model setup for the EHF technique is long and laborious and its implementation is computationally expensive, so much that it becomes impractical to model and history match beyond a single pad.

Furthermore, to model the CO₂ Enhanced Gas Recovery and Storage process (CO₂-EGR&S), a compositional simulator should be utilized, and this significantly increases computation time. Therefore, the perennial challenge in modeling primary production and the CO₂-EGR&S process is to strike a balance between explicit representation of reservoir complexity and long simulation run time for multiple realizations. Proxy models can be developed and used to overcome the aforementioned problem to address the under-utilization of numerical simulation models.

Traditional vs. Data-driven Proxy Model

The proxy models are mathematically driven functions that substitute large numerical simulation cases by replicating simulation model output and providing fast-approximated solutions to be used efficiently in the development planning, uncertainty analysis, and operational design optimization.

The most frequently used proxy models in the oil and gas industry are reduced order models and response surfaces that reduces simulation run-time by approximating the problem and /or the solution space (Kalantari *et al.*, 2012). Response surfaces are categorized as statistical-based proxy models that require a large number of simulation runs to facilitate optimization and uncertainty analysis.

As stated by Mohaghegh (2014), there are two main problems associated with statistics, especially when applied to problems with well-defined physics behind them: a) the issue of “correlation vs. causality,” and b) imposing a pre-defined functional form such as linear, polynomial, exponential, etc., to the data being analyzed or modeled. This approach will fail when data representing the nature of a given complex problem does not lend itself to a pre-determined functional form and changes behavior multiple times.

Unlike traditional proxy models, Data-driven analytics (DDA) takes a different approach to developing proxy models. In this approach, unlike reduced order models, physics and space-time resolution are not reduced and instead of using pre-defined functional forms that are more frequently used to develop response surfaces, a series of machine learning algorithms that conform to the system theory are used for training, with ultimate goal of more accurately modeling the intricacy of a developed shale-CO₂-EGR&S numerical reservoir simulation model.

1
2
3
4 An ensemble of multiple, interconnected, adaptive neuro-fuzzy systems (Nauck, 1997; Jang, 1995) create the core
5 for the development of these models. An artificial neural network is an information processing system with certain
6 performance features analogous to biological neural networks.
7

8
9 In this system, neurons are clustered into different layers, including input, output, and hidden layers. The number of
10 parameters in the dataset defines the number of neurons in the input layer. One or multiple output(s) can be defined
11 during a neuro-model development. One important step toward building a proxy model with machine learning is
12 feature extraction by including the hidden layer(s) and defining the hidden neurons in that layer, which increases the
13 dimensionality that accommodates classification and pattern recognition.
14

15 Neural networks can be divided into two main classes based on training algorithms, namely unsupervised and
16 supervised. In the supervised training mode, both input and output in the form of a spatio-temporal database are
17 presented to the neural network to permit learning on a feedback basis. Different clustering algorithms can be used
18 for partitioning the cases into three main groups: Training, Calibration and Validation.
19

20 During the training process, the weights between the processing elements are adjusted. Memorization and over-
21 training are two main issues that must be avoided during the development of a data-driven proxy models. The
22 calibration portion is used for that purpose to examine the trained neural networks generalization capabilities. The
23 proxy model development workflow is completed, once the verification dataset is used to test the predictive
24 capability of the model.
25

26 **Reservoir Simulation Model**

27

28 In order to make a more realistic model, information from 77 Marcellus shale gas wells in Southwestern
29 Pennsylvania – with a total number of 652 stages of hydraulic fracture and 1,893 clusters in an area of about 53,241
30 acres – was used to develop a simulation model. A representative pad (called WVU pad) – with six horizontal
31 laterals in terms of reservoir and hydraulic fracture characteristics in the study area – was selected to perform the
32 simulation.
33

34 A compositional, matrix-discretized, dual-porosity simulation model consisting of 200,000 grid blocks was
35 constructed. Hydraulic fractures were included into the three-layered simulation model explicitly and the resulting
36 model had nine simulation layers in the refined regions. The simulation model was used to do history matching and
37 methane production forecasting.
38

39
40 Three years of production from six laterals in the WVU pad were successfully history matched and the production
41 was forecasted for 90 years to generate a base case for further CO₂-Enhanced Gas recovery and storage (Shale CO₂-
42 EGR&S) study. Figure 1 shows the entire study area, with 77 horizontal wells and the location of the target WVU
43 pad for CO₂ injection.
44

45 Simulation of the CO₂ injection process in shale was started by defining and running a large number of injection
46 scenarios to recommend a proper set of injection realizations for enhancing gas recovery and CO₂ storage. To
47 optimize the process, an objective function was defined to maximize methane recovery, delay CO₂ breakthrough
48 time, and maximize the volume of stored carbon dioxide in the reservoir.
49

50 As stated before, designing and running the simulation cases to model EGR and Storage in the shale by applying the
51 Explicit Hydraulic Fracture modeling technique (EHF), is long and laborious, and its implementation is
52 computationally expensive, thus requiring the development of a proxy model as a an alternative for fast track
53 uncertainty analysis.
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1. Shale Simulation model (3D) and the location of the targeted pad

Data-driven Proxy Model Development

Spatio-Temporal Database Generation- The most important step in the development of any data-centric reservoir model is the assimilation of a comprehensive spatio-temporal database by considering the uncertainty domain and operational limitations, which forms the foundation of a data-driven shale proxy model for the CO₂-EGR&S process. The success of the data-driven proxy model depends greatly on the degree to which the training dataset represents the fluid flow and storage behavior of the shale reservoir.

In this study, the CO₂ injection plan included injecting through four patterns with different matrix, fracture characteristics, and spacing between injector and producer. From a practical point of view, the methane production from the prospective injection well should reach an economic limit before it can be converted to an injector. In order to meet that criteria, when the production well has drained 75% of its recoverable reserve (based on 100 years EUR), CO₂ injection starts and continues until the end of t=100 years.

Five simulation runs were defined for each case (total of 20 runs) to cover the desired range of reservoir characteristics and different operational conditions. The generated comprehensive spatio-temporal database was used to teach and validate a multilayer feed-forward, back-propagation neural network to accurately mimic the reservoir simulator behavior; to re-generate CH₄ production and CO₂ injection profiles; and also to predict CO₂ breakthrough time and CO₂ production profile for each cluster of hydraulic fracture in the producer and injector(s), as well as corresponding laterals (Figure 2). (The summation of gas rates from all clusters generates a production/injection profile for the corresponding horizontal lateral.)

In order to take the validation one step further, the developed data-driven proxies were validated with a set of completely blind simulation runs that were not used during the training process. The detailed workflow for the data-driven shale proxy model development for the CO₂-EGR&S process is shown in Figure 3.

Figure 2. Illustration of generated production profile by Data-driven proxy model at hydraulic fracture cluster level

Figure 3. CO₂-EGR&S Data-driven Shale proxy model development workflow

Figure 4 illustrates the production injection patterns in four cases. For each pattern (or case), five simulation runs are defined to cover the desired range of reservoir characteristics (i.e. Matrix, natural fracture and hydraulic fracture properties as well as different Langmuir isotherms) and different operational constraints (Flowing bottom-hole pressure and Bottom-hole injection pressure).

In the first three proposed simulation cases, the average distance between producer and injector varies from 2,020 to 900 ft. In Case 4, the hydraulic fractures for both production and injection laterals are overlapping and therefore, early breakthrough can be expected for this pattern. This case is included in the spatio-temporal database to make sure that the developed data-driven proxy model is robust enough to predict production and injection performance in any situation.

Neural networks training, calibration, and validation - The representative spatio-temporal database includes 116,000 pairs of input-output that are used for the training and validation process. The inputs are static data (e.g., reservoir and hydraulic fracture characteristics, sorption features, etc.) and operational constraint, and the outputs are CH₄ production and CO₂ injection profiles, and also CO₂ breakthrough time and CO₂ production profiles for each cluster of hydraulic fracture in the production and injection well(s).

Figure 4. CO₂-EGR&S Process Simulation Cases (Producer and Injector patterns)

In order to take into account the impact of different grid blocks' properties on each cluster, a production/injection "Tier system" was defined. Three different tiers were defined and a property for each tier was calculated by averaging the properties of all the grid blocks in the corresponding tier. Moreover, in order to teach the NNs, the interference effect between the clusters, they were divided into four classes based on their relative location respect to the other offset ones.

Table 1 lists the required information needed for building the training and validation dataset to be used for developing the proxy model.

Table 1. Required information for CO₂-EGR&S Data-driven shale proxy models development

Matrix porosity [0.054-0.125]	Matrix permeability [0.0001-0.0008(md)]	Natural fracture porosity [0.01-0.035]	Natural fracture permeability [0.001-0.004 (md)]	Sigma factor [0.005-0.08]
Hydraulic fracture height [100-125 ft]	Hydraulic fracture length [200-1100 ft]	Hydraulic fracture conductivity [0.1-5.4 (md-ft)]	Net pay thickness [113-128(ft)]	CH ₄ -Langmuir volume [55-91(scft/ton)]
CH ₄ -Langmuir pressure [600-790 psi]	CO ₂ -Langmuir volume [70 -120 (scf/ton)]	CO ₂ -Langmuir pressure [400-580 psi]	CH ₄ -Diffusion coefficient [0.2-4(ft ² /day)]	CO ₂ -Diffusion coefficient [1-20 (ft ² /day)]
Bottom-hole Injection Pressure[1680-3360 psi]	Flowing Bottom-hole Pressure[820-130 psi]			

The back propagation technique was used to perform the training process. Different clustering algorithms can be used for partitioning the cases into three main groups: Training, Calibration and Validation. During the training process, the weights between the processing elements were properly adjusted.

Memorization and over-fitting are two main issues that must be avoided during the development of data-driven proxy models. The calibration portion was used for that purpose to examine the trained neural networks generalization capabilities. The proxy model development workflow was completed, once the neural network reached the best calibration and the trained model passed the final predictive competence test.

Results and discussion

In this section, the results of four data-driven CO₂-EGR&S proxy models at hydraulic fracture cluster level are presented. The data-driven proxies are developed to predict: a) CH₄ production rate, b) CO₂ injection rate, c) CO₂ breakthrough time, and d) CO₂ production rate for each hydraulic fracture clusters, as wells as corresponding laterals. The summation of the rates from all clusters generates a production/injection profile for the corresponding horizontal lateral.

Methane (CH₄) Production Profile Prediction

The first data-driven CO₂-EGR&S proxy model was developed to re-generate an annual methane production rate for each cluster of hydraulic fractures for 100 years, which had previously been generated by the commercial reservoir simulator. Twenty simulation runs were designed and run based on four Injector/Producer patterns to generate 1,160 unique methane production profiles with different petrophysical data, natural and hydraulic fracture characteristics, sorption features (sorption time, diffusion coefficients, and Langmuir isotherms), and operational constraints.

Before starting the training and validation process, the fuzzy pattern recognition technique was used to determine the relative importance and impact of the input parameters on CH₄ production. According to this analysis, the most influential parameters on methane production are duration of production, operational constraints, producer/injector patterns, and relative location of clusters. The second most influential parameters are those that control the reserve, such as isotherms, net pay thickness, natural fracture porosity, and matrix porosity.

Some other parameters – such as hydraulic fracture components, natural fracture properties that control the accessibility of gas to the clusters – can be considered as a third group of parameters that has a great influence on production. They are essential to the productivity of shale gas wells at the beginning of production, but for long-term production performance, operational constraints and parameters that control the reserve play a more important role.

To continue with the training and validation procedure, the neural network architecture was designed to have one intermediate hidden layer with 55 hidden neurons that were selected based on the number of data records available, and the number of input parameters selected in each training process. In the training process, the data set was partitioned into three separate segments by using the intelligent partitioning technique. The data was partitioned with 70% training fraction, 20% for calibration, and 10 % for verification.

The training, calibration, and verification results for annual methane production rate (Mscf/year) (from left to right) are shown in Figure 5. In these plots, the x-axis corresponds to the neural network predicted gas rate, and the y-axis shows the simulated gas rate by the commercial reservoir simulator.

The result, with an R² of more than 0.99 in all steps, shows the successful development of a data-driven shale proxy model for CH₄ production profile prediction.

Figure 6 and Figure 7 show some examples of the comparison of reservoir simulation output for annual methane production rate during the CO₂-EGR&S process, with the predicted one by the data-driven proxy model for some of the clusters and some of the laterals for all four production and injection patterns/cases.

Figure 5. The results of training, calibration and verification sets (From left to right respectively) - CH₄ Production (Mscf/year)

Figure 6. Comparison of CH₄ production profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the clusters (Cases1, 2 and 3)

In all the plots blue dots represent the annual methane production rate generated by the commercial reservoir simulator and the solid red line is the result of the data-driven CO₂-EGR&S proxy model. The results are self-descriptive enough to show the capability of the data-driven proxy model in predicting the CH₄ production profile at hydraulic fracture cluster, and lateral level for different production/injection patterns.

In all four cases, WVU2-1 (the left lateral in Figure 6 and Figure 7) continue methane production for 100 years while the CH₄ production from WVU2-2(Case2), WVU3-2(Case2), WVU3-3(Case3), and WVU3-1(Case4) is stopped after producing the 75% of their accessible gas and then converted to an CO₂ injection well. A boost in the methane production profile in WVU2-1 well can be observed when the displaced methane by CO₂ (during the counter diffusion process) reaches the producing well after 8-20 years, depends on the distance, reservoir characteristics, and bottom-hole injection pressure.

Figure 7. Comparison of CH_4 production profile (Mscf/year) from simulator and CO_2 -EGR&S Data-driven proxy model for some of the laterals (Cases 1, 2, 3 and 4)

CO₂ Injection Profile Prediction

The second data-driven CO_2 -EGR&S proxy model was developed to mimic the numerical simulation behavior and re-generate the annul CO_2 injection rate (Mscf/year) from the injection startup date until $t=100$ years. The same procedure for building the first proxy (CH_4 production) is followed here as well.

The representative database with 116,000 pairs of input-output is used to identify the critical parameters and their degree of influence on CO₂ injectivity in the shale formation. Starting time of injection, operational constraints and the inj./prod. Configuration and relative location of clusters are top ranked parameters during CO₂ injection process and CO₂ /CH₄ Langmuir isotherms, fracture porosity, net pay and fracture half-length are the next top influential parameters for long-term CO₂ injection (44 to 58 years).

Having the hydraulic fracture conductivity as low ranked parameter does not necessarily diminish their importance during the CO₂ injection process. It should be noted that, hydraulic fracture properties are critical components to allow CO₂ injection in Nano-Darcy permeability shale to be initiated. When the injection is started and last for many years, then the other parameters come into the picture and show their contribution for long time injection practice.

By having the better understanding of the key parameters during CO₂ injection process, the spatio-temporal database is partitioned with 70% training, and 30% for calibration and verification. The designed neural network for predicting the CO₂ injection profile has one hidden layers with 58 hidden neurons.

The cross plots for predicted values of CO₂ injection rate (Mscf/year) by the neural network vs. numerical simulation results for training, calibration and verification steps (from left to right) are shown in

Figure 8. In these plots, x-axis corresponds to the neural network predicted CO₂ injection rate (Mscf/year) and the y-axis shows the simulated CO₂ injection rate by the commercial reservoir simulator.

The data driven CO₂-EGR&S proxy model is successfully developed to predict CO₂ injection rate (Mscf/year) with the R² of more than 0.99 in all steps (Training, Calibration and verification).

Figure 9 and Figure 10 show some examples of the comparison of reservoir simulation output for annual CO₂ injection rate during CO₂-EGR&S process with the predicted one by the data-driven proxy model for some of the clusters as well as some of the laterals for all four production and injection patterns/cases.

In all the plots blue dots represent the CO₂ injection rate (Mscf/year) generated by Eclipse and the solid red line is the result of the data-driven CO₂-EGR&S proxy model. The results show that the data-driven proxy model is predicted the CO₂ injection profile at hydraulic fracture cluster, and lateral level for different production/injection patterns with high accuracy.

Figure 8.The results of training, calibration and verification sets (From left to right respectively) – CO₂ Injection rate (Mscf/year)

Figure 9.Comparison of CO₂ injection profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the clusters (Cases1, 2, 3 and 4)

Figure 10. Comparison of CO₂ injection profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the laterals (Cases 1, 2, 3 and 4)

CO₂ Breakthrough Time and Production Profile Predication

The last data-driven CO₂-EGR&S proxy model was developed to re-generate the CO₂ production profile (Mscf/year) from the offset producing well. CO₂ breakthrough (BT) occurrence was dependent on reservoir characteristics, hydraulic fracture characteristics, sorption features, and bottomhole injection pressure.

The same procedure for building the first and second proxies (CH₄ production and CO₂ injection date prediction) was followed here as well.

In order to identify the key parameters that affect CO₂ breakthrough, key performance analysis (KPI) was performed. According to this analysis, injector/producer pattern, relative location of clusters, CO₂ breakthrough time indicator – which was calculated using the CO₂ data-drive proxy for CO₂ breakthrough time prediction – and bottom-hole injection pressure are essential parameters that control the amount of CO₂ production from the offset production well. The other important parameters that have an impact on the amount of CO₂ breakthrough include Langmuir isotherms for both CO₂ and CH₄, natural fracture permeability, and reserve-related parameters such as matrix and fracture porosity and pay thickness. As a result, designed parameters (i.e. well trajectory, hydraulic fracture placement, and operational constraint) should be optimized before starting the CO₂-Enhanced gas recovery process.

The representative database with 116, 000 pairs of input-output were used to train a multilayer feed-forward back-propagation neural network. The data was partitioned with 70% training and 30% validation. The designed neural network had one hidden layer with 52 hidden neurons.

The cross plots for predicted and simulated values of CO₂ production rate (Mscf/year) for training, calibration, and verification steps (from left to right) are shown in Figure 11. In these plots, the x-axis corresponds to the neural network predicted CO₂ production rate (Mscf/year), and the y-axis shows the simulated CO₂ production rate by the reservoir simulator (Eclipse).

The data-driven CO₂-EGR&S proxy model was successfully developed to predict the CO₂ production rate (Mscf/year) with an R2 of more than 0.99 for training, calibration, and verification.

Figure 12 and Figure 13 show some examples of the comparison of reservoir simulation output (blue dots) for CO₂ production rate (Mscf/year) during the CO₂-EGR&S process with the predicted one by the data-driven proxy model (red solid line) for some of the clusters and some of the laterals for all four production and injection patterns/cases. The plots clearly show that the developed data-driven CO₂-EGR&S proxy model is capable of re-generating the CO₂ production rate with acceptable accuracy.

Figure 11. The results of training, calibration and verification sets (From left to right respectively) – CO₂ Production (Mscf/year)

Figure 12. Comparison of CO₂ production profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the hydraulic fracture clusters (Cases1, 2, 3 and 4)

Figure 13. Comparison of CO₂ production profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the laterals (Cases 1, 2, 3 and 4)

Data-driven CO₂-EGR&S Proxy Model Validation by Blind Cases

Validating/testing the developed data-driven proxy model was the final phase of the workflow to examine the capability of the model in predictive mode. In order to do that, one of the cases (Case 2), including five simulation runs, was completely removed from the training and initial validation process. The same input parameters that were used for developing the four data-driven CO₂-EGR proxy models (CH₄ Production, CO₂ Injection rates, CO₂ Breakthrough time, and CO₂ production rate prediction) were used for developing a new proxy model based on 15 simulation runs. Upon completion of the training and validation of the new proxy models, the inputs (i.e., static properties and operational constraints) for the five simulation runs (with 58 clusters of hydraulic fracture per run) were introduced to the proxy models.

The corresponding outputs – including CH₄ Production, CO₂ Injection rates (Mscf/year), CO₂ Breakthrough time (day), and CO₂ production rate (Mscf/year) – were generated, and some of the results are illustrated in Figure 14 (CH₄ production rate),

Figure 15 (CO₂ injection rate), and Figure 16 (CO₂ production rate). In all the plots, the blue dots are the simulation outputs and the red solid lines are the proxy model results. As can be seen in Figure 14, the data-driven CO₂-

EGR&S proxy model predicted the methane production rates for both production and the prospective injector with acceptable accuracy. Since hydraulic fracture length and number of clusters for WVU2-1 does not change for all 15 runs, therefore the neural network could capture the production behavior of this well better than WVU3-2.

Figure 14. Comparison of CH₄ production profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the laterals –Blind Case

Figure 15 and Figure 16 show the comparison of the CO₂ injection and CO₂ production rates from the offset producer calculated by Eclipse with the profiles predicted by the proxy model for some of the runs, which were completely blind to the neural network. According to these figures, the Data-driven proxy model was able to mimic the simulation behavior with good enough accuracy.

Figure 15. Comparison of CO₂ injection profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the laterals –Blind Case

Figure 16. Comparison of CO₂ production profile (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the laterals –Blind Case

The second method used to validate/test the capability of the developed data-driven proxy model was to design a completely new horizontal well with unique reservoir and hydraulic fracture characteristics for each cluster and with new operational constraints, completely different (but in the uncertainty range) from the initially designed 20 simulation runs for training the neural network. Therefore, a new synthetic well (WVU-2013) was drilled and completed with 25 clusters of hydraulic fractures.

Figure 17. Schematic of new synthetic well for validation of the proxy models

Figure 18. Comparison of CH₄, CO₂ production profiles (Mscf/year) and CO₂ Injection rate (Mscf/year) from simulator and CO₂-EGR&S Data-driven proxy model for some of the laterals –Blind Case

Figure 17 shows the horizontal well trajectory with orange color and its relative position to the other offset laterals in the WVU pad. The simulation outputs were compared with the data-driven proxy model results and are shown in Figure 18. In this figure, the top figures are the comparison of simulation and proxy model results for the methane production for the producing and the prospective injector wells. The bottom left plot shows CO₂ production and the bottom right figure shows the CO₂ injection profile comparisons.

Finally, the developed data-driven CO₂-EGR&S proxy models can be used as an efficient and fast uncertainty quantification tool since thousands of proxy models runs can be made quickly to fulfill any uncertainty analysis, techno-economic evaluation, and optimization purposes.

Summary and Conclusions

In this paper, data-driven CO₂-EGR&S proxy models for shale formation were developed based on pattern recognition capabilities of Artificial Intelligence and validated by completely blind simulation runs to reproduce the injection and production profiles for the CO₂-Enhanced gas recovery and storage process. This technique provides the ability to perform fast, detailed uncertainty and optimization analysis, instead of using a numerical simulator for which the model set-up and implementation are laborious and computationally expensive.

Additionally, key performance analysis was performed using the fuzzy pattern recognition technique, and the most influential parameters that control CH₄ production rate, CO₂ injection rate, CO₂ breakthrough time, and CO₂ production rate were identified as important to consider before designing any CO₂ injection process for the purpose of enhanced gas recovery and storage.

Acknowledgements

The authors would like to acknowledge the U.S. Department of Energy (DOE) - NETL-RUA for financially supporting this project. The authors also would like to acknowledge Intelligent Solutions, Inc. (ISI) and Schlumberger for providing the ISMA and Eclipse/ Petrel package.

References

- Aydin, G.; Karakurt, I.; Aydiner, K. "Evaluation of geologic storage options of CO₂: Applicability, cost, storage, capacity and safety," *Energy Policy*. 2010, 38, 5072-5080.
- Bachu, S.; Bonijoly, D.; Bradshaw, J.; Burruss, R.; Holloway, S.; Christensen, N.; Mathiassen, O. "CO₂ storage capacity estimation: Methodology and gaps," *International Journal Of Greenhouse Gas Control*. 2007: 1, 430-443.
- Cipolla, C.L. et al: "Modeling Well performance in Shale-Gas Reservoirs," SPE 125532, SPE/WAGE Reservoir Characterization, and Simulation Conference held in Abu Dhabi, UAE, Oct. 19 – 21, 2009.
- Eshkalak, M. and Aybar, U. (2015) Carbon Dioxide Storage and Sequestration in Unconventional Shale Reservoirs. *Journal of Geoscience and Environment Protection*, 3, 7-15. doi: 10.4236/gep.2015.31002.
- Kalantari-Dahaghi, A.; Esmaili, S.; Mohaghegh, S. "Fast Track Analysis of Shale Numerical Models," SPE 162699, Canadian Unconventional Resources Conference held in Calgary, Alberta, Canada, 30 October–1 November 2012.
- Kalantari-Dahaghi, A.; Mohaghegh, S.D. "Numerical Simulation and Multiple Realizations for Sensitivity Study of Shale Gas Reservoirs," SPE 141058, SPE Production and Operations Symposium, Oklahoma City, Oklahoma, USA, 2011.
- Kalantari-Dahaghi, A. "Numerical Simulation and Modeling of Enhanced Gas Recovery and CO₂ Sequestration in Shale Gas Reservoirs-A Feasibility Study," SPE 139701, SPE International Conference on CO₂ Capture, Storage, and Utilization, New Orleans, LA, USA, 2010.
- Khazani, Y., Mohaghegh S.D. "Intelligent Production Modeling Using Fullfield Pattern Recognition," 2011 SPE Reservoir, Vol.14, No. 6, pp.735-749.
- Krueger, D.; Franklin, P. "The Methane to Markets Partnership: Opportunities for coal mine methane utilization," in Mutmanský, J.; Ramani, R. 11th US/North American Mine Ventilation Symposium, Penn State Univ, University Pk, PA, June 05-07, 2006; 11th U.S./ North American Mine Ventilation Symposium 2006 Book Series: Proceedings and Monographs in Engineering, Water and Earth Sciences, 2006, pp. 3-8.
- Lewis, R. et al. "New Evaluation Techniques for Gas Shale Reservoirs," Schlumberger paper for 2004 Reservoir Symposium.
- Meer, V.B. "Carbon dioxide storage in natural gas reservoirs," *Oil and Gas Science and Technology – Review IFP60*. 2005, 3, pp. 527-536.
- Mohaghegh, S.; Abdull, F. "Production Management Decision Analysis Using AI-Based Proxy Modeling of Reservoir Simulations – A Look-Back Case Study," SPE 170664, SPE Annual Technical Conference and Exhibition held in Amsterdam, The Netherlands, 27–29 October 2014.
- Mohaghegh, S.D. "Reservoir simulation and modeling based on artificial intelligence and data mining (AI&DM)," *Journal of Natural Gas Science and Engineering*, Elsevier, 3 (2011) p.697-705.
- Mongalvy, V.; Chaput, E.; Agarwal, S.; and Lu, L. "A New Methodology for Shale Reservoir Performance," paper SPE 144154, prepared for presentation at the 2011 Americas Unconventional Gas Conference, Woodlands, TX, 14-16 June.

Nuttal, B.; Eble, C.; Drahovzal, J.; Bustin, R. "Analysis of Devonian Black Shales in Kentucky for potential carbon dioxide sequestration and enhanced natural gas production," Kentucky Geological Survey, 2005, DE-FC26 02NT41442.

Rubin, B. "Accurate Simulation of Non-Darcy Flow in Stimulated Fracture Shale Reservoirs," SPE 132293, prepared for presentation at the Western Regional Conference, Anaheim, CA, May 27 – 29, 2010.

Rubin, B. "Accurate Simulation of Non-Darcy Flow in Stimulated Fractured Shale Reservoirs," Paper SPE 132093, Western Regional Meeting, Anaheim, CA, May 27-29, 2010.

Schepers, K.; Nuttal, B.; Oudinot, A.; Gonzalez, R. "Reservoir modeling and simulation of the Devonian gas shale of Eastern Kentucky for enhanced gas recovery and CO₂ storage," SPE12662, 2009.

U.S. Energy Information Administration (EIA) "U.S. Crude Oil, Natural Gas, and Natural Gas Liquids Reserves," U.S. Department of Energy, Washington, DC (November 30, 2010).

Wang, J. and Liu, Y. "Well Performance Modeling in Eagle Ford Shale Oil Reservoir," SPE 144427, Americas Unconventional Gas Conference, Woodlands, TX, 14-16 June 2011.

Zhang, W; Li, Y; Xu, T; Cheng, H; Zheng, Y; and Xiong, P. "Long-term variations of CO₂ trapped in different mechanisms in deep saline formations: A case study of the Songliao Basin, China," International Journal Of Greenhouse Gas Control. 2009, 3, 161-180.

APPENDIX A: Comparison of CH₄, CO₂ production and CO₂ injection rate generated by Data-driven CO₂-EGR&S proxy models with the numerical simulation results

In all the plots, blue dots represent the rates generated by Eclipse and the solid red line is the result of the data-driven CO₂-EGR&S proxy model. There are four plots in each figure to compare the data-driven proxy model results for CH₄ Production (top) CO₂ production (bottom left), and CO₂ injection (bottom right) with numerical simulation results.

Figure 19. Comparison of CH₄, CO₂ Production rates (Mscf/year) and CO₂ injection rate from simulator and CO₂-EGR&S Data-driven proxy model for producer/injector pairs of laterals (Cases1, Run1)

Figure 20. Comparison of CH₄, CO₂ Production rates (Mscf/year) and CO₂ injection rate from simulator and CO₂-EGR&S Data-driven proxy model for producer/injector pairs of laterals (Cases2, Run7)

Figure 21. Comparison of CH₄, CO₂ Production rates (Mscf/year) and CO₂ injection rate from simulator and CO₂-EGR&S Data-driven proxy model for producer/injector pairs of laterals (Cases3, Run13)

Figure 22. Comparison of CH₄, CO₂ Production rates (Mscf/year) and CO₂ injection rate from simulator and CO₂-EGR&S Data-driven proxy model for producer/injector pairs of laterals (Cases4, Run20)