

Estimation of Adsorption Capacity of CO₂, CH₄, and their Binary Mixtures in Quidam Shale using LSSVM: Application in CO₂ Enhanced Shale Gas Recovery and CO₂ Storage

Amin Bemani¹, Alireza Baghban^{2,*}, Amir H. Mohammadi^{3,4}, Pål Østebø Andersen^{5,*}

¹Department of Petroleum Engineering, Petroleum University of Technology (PUT), Ahwaz, Iran

²Department of Chemical Engineering, Amirkabir University of Technology, Mahshahr Campus, Mahshahr, Iran

³ Institut de Recherche en Genie Chimique et Petrolier (IRGCP), Paris Cedex, France

⁴ Discipline of Chemical Engineering, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa

⁵ Department of Energy Resources, University of Stavanger, Norway and the National IOR Centre of Norway, University of Stavanger, Norway

Abstract - Carbon dioxide enhanced shale gas recovery depends strongly on adsorption properties of carbon dioxide and methane. In this work, Least Squares Support Vector Machine (LSSVM) optimized by Particle Swarm Optimization, has been proposed to learn and then predict adsorption capacity of methane and carbon dioxide from pure and binary gas mixtures in Jurassic shale samples from the Qaidam Basin in China based on input parameters pressure, temperature, gas composition and TOC. A literature dataset of 348 points was applied to train and validate the model. The predicted values were compared with the experimental data by statistical and graphical approaches. The coefficients of determination of carbon dioxide adsorption were calculated to 0.9990 and 0.9982 for training and validation datasets, respectively. For CH₄ the numbers are 0.9980 and 0.9966. The model was extrapolating reasonable trends beyond measurement ranges. More extensive datasets are needed to properly parameterize the role of shale properties.

Keywords – Adsorption capacity of CO₂ and CH₄; Carbon storage in shales; Multicomponent gas isotherm; LSSVM; PSO.

1. Introduction

* Corresponding authors email addresses: alireza_baghban@alumni.ut.ac.ir (A. Baghban) and pal.andersen@uis.no (P. Andersen)

Due to economic development and population growth, worldwide demand for energy is increasing [1]. Shale gas reservoirs have become one of the dominant unconventional energy sources with vast resources particularly in Canada, China and the United States. They are associated with a high storage capacity particularly due to significant adsorption [2-7]. In the stated countries, the study and development of shale gas reservoirs have become the focus of energy industries and producing the adsorbed gas is a central topic [8, 9]. At high pressure and temperature conditions, carbon dioxide has higher adsorption capacity than methane especially in the micropore volume fraction [10]. In recent investigations, it was shown that five carbon dioxide molecules could take the place of one methane molecule. Hence, carbon dioxide injection into organic-rich gas shales has benefits such as enhancing adsorbed methane recovery, effective sequestration of carbon dioxide and mitigating global warming caused by carbon dioxide emissions[11]. Gas shale thus is an interesting alternative for carbon dioxide sequestration [12, 13]. Currently, this approach has not yet been commercialized, although numerous efforts have been made to investigate its feasibility [14-16]. Key investigated topics include transport properties of gas mixtures in shale, adsorption mechanisms and solid-gas molecular interactions. The parameters affecting carbon dioxide enhanced methane recovery and co-adsorption of a mixture of carbon dioxide and methane still are investigated. A recent review of experimental results indicating trends on the adsorption capacity of CO₂ and CH₄ in shales is reported by Klewiah et al. [7]. The focus of this study will be on the correlation and prediction of adsorption of CH₄-CO₂ mixtures in shales.

The first studies of the shale gas adsorption were mainly about the mechanism of shale gas storage [17, 18]. Furthermore, numerous investigations on the relation between adsorption and effective parameters of shale such as clay mineral content, organic matter, and pore structure were carried out. The previous investigations expressed that there is an approximately linear relationship

between total organic carbon content (denoted TOC) of shale and methane adsorption ability [18, 19]. The thermal maturity and organic matter type can also strongly affect methane adsorption capacity and the rate of adsorption [20, 21]. Clay mineral content and their structure also have clear impacts on adsorption ability of shale [22, 23]. Cheng et al. expressed that larger surface areas can enhance the methane adsorption in shale [24]. Bai et al. proposed a linear relationship between BET surface area and methane adsorption capacity [25]. Ross et al. concluded that as the micropore volume and organic carbon content increase, the methane adsorption capacity increases [6]. Weniger and coworkers developed correlations to estimate adsorption of carbon dioxide and methane for shales and coals in Brazil [26]. Chareonsuppanimit et al. investigated methane, carbon dioxide and nitrogen adsorption in shale. Their results shows the higher amount of adsorption for carbon dioxide [10]. Kang et al. studied the carbon dioxide storage capacity in shales in terms of pressure by considering adsorption and compressibility effects [5].

Artificial intelligence approaches have demonstrated extensive applications in interpretation and prediction [27-29]. In this work we seek to determine the relations controlling adsorption capacity of CO₂ and CH₄ in shale under various conditions. For doing so, we apply the Least Squares Support Vector Machine (LSSVM) algorithm as a predictive tool to forecast adsorption of methane and carbon dioxide binary mixtures adsorption in shale. This method is considered accurate and user-friendly. Optimal training parameters for the model were found using Particle Swarm Optimization (PSO). The algorithms are implemented in Matlab. Furthermore, sensitivity analysis carried out to identify the most influential of the input parameters on gas adsorption capacity.

2. Methodology

2.1. Aims and workflow

The main purpose of the present research is the development of an LSSVM algorithm which optimized by the PSO algorithm can be used to predict adsorption capacity of carbon dioxide and methane when these gases are exposed to shale as pure component or binary mixtures. Based on the available dataset the input parameters are pressure, temperature, gas mole fraction of CO₂ and TOC of shale as shown in [Figure 1](#). Effectively this provides a multicomponent adsorption isotherm that does not assume a particular relation between sorption and pressure (which is the most frequent measurement while holding other parameters fixed) and also integrates the effect of the other stated parameters. The input data are normalized before implementing in LSSVM-PSO and after that outputs are realized and altered to adsorption capacity. The workflow of our approach is summarized in [Figure 2](#) while more details are explained in the following sections.

2.2. Data Gathering

The consistency and efficient performance of any proposed model extensively depend on the quality of the dataset the model was trained from. A comprehensive dataset of experimental measurements generated by [30] was applied in this work; which includes 348 measurements each for carbon dioxide or methane adsorption capacity (measured in cm³/g) on Jurassic shale from the Qaidam Basin in China for different combinations of the four (input) parameters pressure, temperature, carbon dioxide mole fraction and TOC. The ranges of input variables are reported in [Table 1](#). It is worth mentioning that Luo et al. [30] used van der Waals equation of state to determine absorbed gas density, a parameter difficult to measure experimentally, which was further used to calculate the excess adsorbed gas from the experimental data [30]. In order to

prepare and test the machine learning models, the collected data were randomly divided into two groups, respectively for training and validation. Note that the terms ‘validation’ and ‘testing’ will be used interchangeably.

2.3. Identification of outliers

The utilized data points in preparation of prediction model can influence the predictive ability of proposed model. In analysis of large datasets, identifying outlier data is important to assess their representativeness. To this end, Leverage analysis was employed where the so-called Hat matrix is defined as follows [31, 32]:

$$H = X(X^T X)^{-1}X^T \quad (1)$$

In the aforementioned equation, X represents the $m \times n$ design matrix with m and n standing for the number of experimental data points (348 in this case) and parameters of the predicting model (4 in this case), respectively. The hat values for each experimental point can be given by the main diagonal of the Hat matrix. For visualization of outliers, a William's plot is made where normalized residuals are plotted on the vertical axis against hat values on the x-axis. As indicated in [Figure 3](#), three outlier points were pointed out by standard residuals being outside the range -3 to 3. A warning leverage value (H^*) was also included in the William's plot defined by:

$$H^* = 3(n + 1)/m \quad (2)$$

No points had hat values above this threshold and hence were not given too much leverage in the data analysis. Hence, although three points appeared different from the rest, they or others did not have a strong influence on the model calibration compared to other data points. For simplicity all data were included.

2.4. Least square support vector machine (LSSVM)

The Support Vector Machine (SVM) approach was introduced by Vapnik et al. [33-35] as a method for classification. It transfers a set of input variables (used to predict an output) via non-linear functions to a multidimensional space where the points are separated by decision surfaces. This method was further applied to regression of data, then termed Support Vector Regression, where an objective function should be minimized under inequality constraints describing how far the estimated outputs could deviate from the true value [36]. Suykens and Vandewalle [37] made the Least Square Support Vector Machine (LSSVM) algorithm as a simpler form of SVM by reformulating the problem in terms of linear, rather than non-linear, constraints on the function error terms. The LSSVM approach is outlined in the following based on the notation and derivations in [38, 39]. A minimization problem is then formulated as:

$$\min_{w,e} J(w, e) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \gamma \sum_{i=1}^n e_i^2 \quad (3)$$

where $J \in R$ is an objective function which should be minimized for a given weight vector \mathbf{w} , T denotes the transpose operation, e_i ($i = 1, \dots, n$) is the difference between output estimate $f(x_i) \in R$ and the true output $y_i \in R$ (the output is the adsorption capacity in this case) and $\gamma \in R$ is a so-called regularization parameter (defining which term is more important to minimize). n is the number of measurements. The estimated output of data point i is defined by the weighted combination of the elements in a nonlinear function $\phi(x_i)$ and a constant bias parameter b . This function maps the input space (in our case a four dimensional vector with values of pressure, temperature, CO2 fraction and TOC) to a higher dimension with same order as w . Note that ϕ is a useful temporary variable for the method derivation, but is not explicitly calculated, hence its

dimension need not be specified. The true output y_i is by definition equal to the estimate $f(x_i)$ plus the error e_i :

$$y_i = f(x_i) + e_i = \mathbf{w}^T \boldsymbol{\phi}(x_i) + b + e_i, \quad (i = 1, \dots, n) \quad (4)$$

This formulation of function estimation as an optimization (minimization) problem with linear constraints is characteristic for the LSSVM method. The problem can be reformulated to finding the saddle point of the Lagrange function:

$$L(\mathbf{w}, b, \mathbf{e}, \boldsymbol{\alpha}) = J(\mathbf{w}, \mathbf{e}) - \sum_{i=1}^n \alpha_i (\mathbf{w}^T \boldsymbol{\phi}(x_i) + b + e_i - y_i) \quad (5)$$

where α_i are the so-called Lagrange multipliers and its associated terms correspond to the incorporation of the equality constraints in (4). The saddle point is found by setting partial derivatives of L with respect to the parameters $\mathbf{w}, b, e_i, \alpha_i$ to zero which implies that:

$$\mathbf{w} = \sum_{i=1}^n \alpha_i \boldsymbol{\phi}(x_i) \quad (6)$$

$$\sum_{i=1}^n \alpha_i = 0 \quad (7)$$

$$\alpha_i = \gamma e_i, \quad (i = 1, \dots, n) \quad (8)$$

$$\mathbf{w}^T \boldsymbol{\phi}(x_i) + b + e_i - y_i = 0, \quad (i = 1, \dots, n) \quad (9)$$

Eliminating \mathbf{e} and \mathbf{w} from this linear set of equations results in a reduced set of equations where the non-linear function $\boldsymbol{\phi}$ only appears in terms of the following form:

$$\boldsymbol{\Omega}_{kj} = (\boldsymbol{\phi}(x_k))^T \boldsymbol{\phi}(x_j) = \psi(x_k, x_j), \quad (k, j = 1, \dots, n) \quad (10)$$

The set of equations to be solved can be expressed by:

$$\begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & A \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ y \end{bmatrix}, \quad A = \Omega + \gamma^{-1}I \quad (11)$$

$\psi(x_k, x_j)$ is a kernel function which must be defined with some restrictions. With this function we effectively replace the need to define and compute ϕ . A is symmetric and positive semi-definite which guarantees the existence of its inverse. For a defined kernel function $\psi(x_k, x_j)$ the optimal values of b and the α vector can be computed:

$$b = \frac{\mathbf{1}^T A^{-1} y}{\mathbf{1}^T A^{-1} \mathbf{1}}, \quad \alpha = A^{-1}(y - b\mathbf{1}) \quad (12)$$

Using that the weight vector w is given by (4) we can update the output estimate $f(x)$ in (4) for a general input vector as follows:

$$f(x) = \mathbf{w}^T \phi(x) + b = \sum_{i=1}^n \alpha_i (\phi(x_i))^T \phi(x) + b = \sum_{i=1}^n \alpha_i \psi(x, x_i) + b \quad (13)$$

One of the most popular choices of kernel functions applied in LSSVM, and also applied here, is the (Gaussian) Radial Basis function (see [36] for other choices):

$$\psi(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \quad (14)$$

Here σ^2 indicates the squared bandwidth and $\|x_i - x_j\|$ denotes the magnitude of the vector difference $x_i - x_j$. σ^2 and the regularization coefficient γ are both given parameters for an application of the LSSVM algorithm. To further optimize the performance of the LSSVM algorithm in terms of matching given data and prediction of remaining data; these two parameters were optimized using Particle Swarm Optimization as described in the following section.

2.5. Particle swarm optimization (PSO)

The assigned parameters γ, σ^2 in the structure of the LSSVM algorithm can have extensive impacts on the performance quality for prediction so it is required to determine these parameters carefully. In the literature, different heuristic and analytical methods exist to lead predicting tools to their best structure. Among analytical approaches, gradient-based manners can be utilized to forecast these parameters, while in heuristic approaches, evolutionary algorithms like simulated annealing, particle swarm optimization (PSO), and genetic algorithms are utilized. In this study, PSO was applied. This algorithm has great ability to improve estimates of local and global optima [40, 41] and is simple to implement.

In the PSO optimization algorithm a number of candidate solutions ($r_k = [\gamma_k, \sigma_k^2]$), each called a particle, are initially randomly distributed in the search space and given initial velocities v_k also randomly. The entire set of ‘particles’ is called the swarm. The solution estimate of particle k corresponds to its position r_k . At a given iteration the quality of the solution parameters for each particle is calculated; This was quantified by the Root Mean Square Error (RMSE) evaluated over the training data when the optimized LSSVM algorithm applied the ‘particle’ position values of γ and σ^2 . The best solution position for every individual particle (throughout its movements / iterations) and the best solution position of all the particles (the swarm) are kept track of and denoted p_k and p_g , respectively. New velocities v_k^{new} are then calculated for each particle as a function of the previous velocity v_k^{old} , and based on how far the particle is from its historic best position p_i and from the swarm’s historic best position p_g [42-47]:

$$v_k^{new} = \omega v_i^{old} + U(0, C_1)(p_i - r_i) + U(0, C_2)(p_g - r_i) \quad (15)$$

$$r_i^{new} = r_i^{old} + V_i^{new} \quad (16)$$

Here C_1 and C_2 are known as the acceleration constants, stating how quickly the particle will be steered towards the two currently best positions. U is the uniform distribution function and assigns a random value between 0 and the indicated acceleration constant. It has the role of applying stochastic weight in the particle velocity updating. ω is the so-called inertia weight and is used for controlling the search velocity balancing between global exploration (large steps to cover the search space) and local optimization (small steps near the optimum). This parameter is reduced with increasing iteration number as calculated below:

$$\omega^{new} = \omega^{old} f_{damp} \quad (17)$$

Here ω^{new} and ω^{old} indicate the inertia weight at the new and old iteration step, respectively, and f_{damp} is the damping factor at which ω is reduced at every iteration. Particle position updates may in some cases fall outside of the defined search space. This is solved by randomly assigning a new position to an ‘existing’ particle i to a new location in the search space as follows:

$$x_{ji}^{new} = (x_{j,max} - x_{j,min})U_j(0,1) + x_{j,min}, \quad (j = 1, \dots, 4) \quad (18)$$

Here x_{max} and x_{min} indicate the vectors with maximum and minimum values of the parameter range of the search space.

Briefly summarized, the optimization process can be described in the following steps:

- (1) First assign initial values of positions and velocities, after that, the inertia weight, learning parameters and maximum number of iteration should be selected.
- (2) The fitness of particles should be computed.
- (3) The position and speed of particles should be updated.

(4) The stop condition will be checked. If the assumed condition is obtained, the loop is terminated, otherwise, the loop will be continued.

2.6. Assessment of matching quality

Some statistical approaches described in the following formulations were utilized to evaluate the explanatory power (quality of match) between the observed data y_i and the simulated data $y_i^{est} = f(x_i)$.

1. Coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y_i^{est})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (19)$$

2. Mean relative error (MRE):

$$MRE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i^{est} - y_i|}{y_i} \quad (20)$$

3. Root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y_i^{est})^2} \quad (21)$$

4. Standard deviation of the error (STD_{err}):

$$STD_{err} = \left(\frac{1}{N-1} \sum_{i=1}^N (e_i - \bar{e}) \right)^{0.5} \quad (22)$$

In the above formulas, \bar{y} denotes the mean of observed data and \bar{e} the mean error between estimated and observed data.

3. Results and discussion

In this section results are presented for CO₂ adsorption as function of the four input parameters temperature, pressure, CO₂ composition and TOC. The analysis for CH₄ adsorption is completely similar and the results are summarized in the Appendix.

3.1. Parameter relevance for prediction

The impact of the four input parameters on CO₂ adsorption were first assessed by the Pearson correlation coefficient r_j , also termed the relevancy factor. r_j indicates the impact of input parameter j in the x vector on the output y and is defined as follows [48, 49]:

$$r_j = \frac{\sum_{i=1}^n (x_{j,i} - \bar{x}_j)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_{j,i} - \bar{x}_j)^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (j = 1, \dots, 4) \quad (23)$$

The factor r_j is in the range of -1 to 1. Higher absolute value expresses stronger correlation, and a positive or negative sign indicates whether the dependent parameter y increases or decreases with variations in x_j , respectively. The determined relevancy factors for each input parameter for impact on CO₂ adsorption are shown in [Figure 4](#). The figure expresses that all four parameters contribute to affect adsorption capacity, although the TOC content is less significant compared to the others ($r \sim 0.2$). Pressure dependence appears most significant ($r \sim 0.7$), followed by CO₂ composition ($r \sim 0.4$) and temperature ($r \sim -0.4$). In accordance with findings in the literature, adsorption is negatively correlated only with temperature. It should be noted that the weak correlation to TOC could be due to a lack of measurements of this parameter. The reported dataset were classified into three shale types where only one value of TOC was assigned to each shale type. Although this value can vary from sample to sample and usually is considered strongly linked to adsorption capacity the statistical variation of this parameter was not reported in the dataset. We also note that

the relevancy factor calculation assumes a linear dependence between the parameters, which is generally not the case. Hence, a higher correlation should be expected when non-linear functional dependence is introduced via the LSSVM algorithm.

3.2. Model calibration and validation

After PSO optimization, the LSSVM model had been parameterized and calibrated to the training dataset. The parameters applied in the optimized LSSVM model (particularly including the band width σ^2 and regularization coefficient γ) are listed in **Table 2**, while details related to the PSO algorithm (such as the number of particles, number of iterations, acceleration constants C_1, C_2 , inertia weight ω and damping factor f_{damp}) are listed in **Table 3**. By applying the latter parameters in the PSO algorithm the former parameters were obtained for the LSSVM algorithm, defining the optimum LSSVM structure for describing the data. A graphical comparison between forecasted and actual adsorption values of CO₂ is shown in **Figure 5**. Both the training dataset and validation dataset are included. As it can be observed in this demonstration, the forecasted and actual adsorption values cover each other with high accuracy. A cross plot of actual and model calculated data are shown for both the training dataset and the validation dataset in **Figure 6**. Great agreement exists between calculated and actual values of carbon dioxide adsorption capacity values for both datasets with correlation coefficients R^2 of 0.999 and 0.998 for the training and validation datasets, respectively. Similarly, **Figure 7** demonstrates the relative deviation of forecasted adsorption compared to actual values $\frac{y_i^{est} - y_i}{y_i}$. The relative errors are mainly limited to within ~10%, for adsorption capacity from 1 to 12 cm³/g, but higher deviation is seen at lower adsorption

capacities. More uniform deviations could be obtained by minimizing the relative (rather than absolute) errors.

After the visual comparison, a quantitative comparison can be useful. A summary of the aforementioned statistical indices (STD, RMSE, MRE and R^2 defined in (19) to (22)) is given in **Table 4** regarding the ability of the model to reproduce the experimental adsorption values for carbon dioxide adsorption. The indices are calculated for the training, validation and total datasets. These indices are for the training dataset set; $R^2=0.999$, $MRE=3.77\%$, $MSE=0.0092$, and $STD=0.0758$ and for the validation set; $R^2=0.9982$, $MRE=7.38\%$, $MSE=0.0206$, and $STD=0.0994$. Although the performance naturally is better for the trained set, both sets have very high and comparable performance and especially the model is able to predict data in the validation dataset.

Figure 8 illustrates the adsorption capacity in presence of pure CO_2 gas composition for different values of pressure and temperature and for the three shale types. The model and experimental data are in excellent agreement. Moreover, CO_2 adsorption capacity in a binary gas system as function of pressure and CO_2 mole fraction is shown in **Figure 9** for shale 1, shale 2, and shale 3 at constant temperature conditions of 308 K. Also here good agreement between model and experiments is demonstrated, although higher deviations can be seen in this case, with differences up to $0.5 \text{ cm}^3/\text{g}$ for some points. Furthermore, the overall trends of adsorption capacity of CO_2 in terms of gas composition, pressure and temperature show reasonable variations based on the sensitivity analysis.

3.3. Prediction

A significant strength of the LSSVM model is that it can be calibrated to match large datasets of non-linearly correlated data and then be used to predict behavior under new conditions. For demonstration, the LSSVM model is hence used for prediction of CO₂ adsorption in binary gas systems vs pressure at two temperatures of 318 and 338 K as shown in [Figure 10](#) and [Figure 11](#). This is demonstrated for all three shale types. Smooth trends are obtained where increased pressure and increased CO₂ fractions both lead to higher adsorption of CO₂.

4. Conclusions

In this paper the machine learning algorithm Least Squares Support Vector Machine (LSSVM) was applied, optimized by Particle Swarm Optimization (PSO), to correlate CO₂ adsorption in shale from pure or binary CO₂-CH₄ gas mixtures as function of pressure, temperature, CO₂ gas fraction and Total Organic Carbon (TOC). A total of 348 data points (combinations of these four inputs and the adsorption capacity) were used with 80 % for calibration and the remaining 20 % for validation. The same methodology was applied to CH₄ adsorption. The obtained capacity values were compared with the experimental values for both the training and validation datasets by graphical and statistical methods. Coefficients of determination of $R^2 \sim 0.998$ between experimental and calculated adsorption were obtained for both datasets. The applied dataset did not contain a sufficiently detailed reporting of TOC to reliably estimate the impact of this parameter. We also note that this model was parameterized against a dataset from the Qaidam Basin in China and generalization to other shales is uncertain.

Appendix

The proposed algorithm has demonstrated an ability to match and predict gas adsorption capacity for CO₂ as discussed in the main text of the paper. For the interested reader, in this Appendix the analysis is repeated for the corresponding experimental results of CH₄ from the same experimental dataset. The parameters used to optimize and calibrate the LSSVM and PSO algorithms are listed in **Table 2** and **Table 3**, respectively. **Figure A1** compares model performance against each data point for the training and test dataset, while **Figure A2** directly compares modeled vs predicted adsorption and **Figure A3** shows the relative deviation. A quantitative summary of the LSSVM-PSO model performance for CH₄ adsorption is given in **Table 4**. **Figure A4** presents CH₄ modeled and experimental adsorption vs pressure at different temperatures and shale types when pure CH₄ is exposed to shale and **Figure A5** presents CH₄ adsorption vs pressure at different gas compositions and shale types for the same temperature of 308 K. The model is used to extrapolate the adsorption behavior to mixed gas composition at higher temperatures in **Figure A6** (318 K) and **Figure A7** (338 K).

Acknowledgments

Andersen acknowledges the Research Council of Norway and the industry partners, ConocoPhillips Skandinavia AS, Aker BP ASA, Vår Energi AS, Equinor ASA, Neptune Energy Norge AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, and Wintershall DEA, of The National IOR Centre of Norway for support.

References

- [1] B. Dhinesh, M. Annamalai, A study on performance, combustion and emission behaviour of diesel engine powered by novel nano nerium oleander biofuel, *Journal of cleaner production*, 196 (2018) 74-83.
- [2] G.R. Chalmers, R.M. Bustin, Lower Cretaceous gas shales in northeastern British Columbia, Part I: geological controls on methane sorption capacity, *Bulletin of Canadian petroleum geology*, 56 (2008) 1-21.

- [3] D.J. Ross, R.M. Bustin, Characterizing the shale gas resource potential of Devonian–Mississippian strata in the Western Canada sedimentary basin: Application of an integrated formation evaluation, *AAPG bulletin*, 92 (2008) 87-125.
- [4] B.C. Nuttal, C. Eble, R.M. Bustin, J.A. Drahovzal, Analysis of Devonian black shales in Kentucky for potential carbon dioxide sequestration and enhanced natural gas production, in: *Greenhouse Gas Control Technologies 7*, Elsevier, 2005, pp. 2225-2228.
- [5] S.M. Kang, E. Fathi, R.J. Ambrose, I.Y. Akkutlu, R.F. Sigal, Carbon dioxide storage capacity of organic-rich shales, *Spe Journal*, 16 (2011) 842-855.
- [6] D.J. Ross, R.M. Bustin, The importance of shale composition and pore structure upon gas storage potential of shale gas reservoirs, *Marine and Petroleum Geology*, 26 (2009) 916-927.
- [7] I. Klewiah, D.S. Berawala, H.C.A. Walker, P.Ø. Andersen, P.H. Nadeau, Review of experimental sorption studies of CO₂ and CH₄ in shales, *Journal of Natural Gas Science and Engineering*, 73 (2020) 103045.
- [8] C.C. Iferobia, M. Ahmad, A review on the experimental techniques and applications in the geomechanical evaluation of shale gas reservoirs, *Journal of Natural Gas Science and Engineering*, (2019) 103090.
- [9] Y. Zhang, D. Li, X. Sun, P. Li, New theoretical model to calculate the apparent permeability of shale gas in the real state, *Journal of Natural Gas Science and Engineering*, 72 (2019) 103012.
- [10] P. Chareonsuppanimit, S.A. Mohammad, R.L. Robinson Jr, K.A. Gasem, High-pressure adsorption of gases on shales: Measurements and modeling, *International Journal of Coal Geology*, 95 (2012) 34-46.
- [11] F. Liu, K. Ellett, Y. Xiao, J.A. Rupp, Assessing the feasibility of CO₂ storage in the New Albany Shale (Devonian–Mississippian) with potential enhanced gas recovery using reservoir simulation, *International Journal of Greenhouse Gas Control*, 17 (2013) 111-126.
- [12] F.M. Orr, Onshore geologic storage of CO₂, *Science*, 325 (2009) 1656-1658.
- [13] F. Luo, R.-N. Xu, P.-X. Jiang, Numerical investigation of the influence of vertical permeability heterogeneity in stratified formation and of injection/production well perforation placement on CO₂ geological storage with enhanced CH₄ recovery, *Applied energy*, 102 (2013) 1314-1323.
- [14] K.C. Schepers, B.C. Nuttall, A.Y. Oudinot, R.J. Gonzalez, Reservoir modeling and simulation of the Devonian gas shale of eastern Kentucky for enhanced gas recovery and CO₂ storage, in: *SPE International Conference on CO₂ Capture, Storage, and Utilization*, Society of Petroleum Engineers, 2009.
- [15] H. Sun, J. Yao, S.-h. Gao, D.-y. Fan, C.-c. Wang, Z.-x. Sun, Numerical study of CO₂ enhanced natural gas recovery and sequestration in shale gas reservoirs, *International Journal of Greenhouse Gas Control*, 19 (2013) 406-419.
- [16] M. Godec, G. Koperina, R. Petrusak, A. Oudinot, Potential for enhanced gas recovery and CO₂ storage in the Marcellus Shale in the Eastern United States, *International Journal of Coal Geology*, 118 (2013) 95-104.
- [17] E.C. Smith, A practical approach to evaluating shale hydrocarbon potential, in: *Second Eastern Gas Shales Symposium*, 1978, pp. 73.
- [18] X.-C. Lu, F.-C. Li, A.T. Watson, Adsorption measurements in Devonian shales, *Fuel*, 74 (1995) 599-603.
- [19] S. Wang, Z. Song, T. Cao, X. Song, The methane sorption capacity of Paleozoic shales from the Sichuan Basin, China, *Marine and Petroleum Geology*, 44 (2013) 112-119.

- [20] D.M. Jarvie, R.J. Hill, T.E. Ruble, R.M. Pollastro, Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment, *AAPG bulletin*, 91 (2007) 475-499.
- [21] T. Zhang, G.S. Ellis, S.C. Ruppel, K. Milliken, R. Yang, Effect of organic-matter type and thermal maturity on methane adsorption in shale-gas systems, *Organic geochemistry*, 47 (2012) 120-131.
- [22] R. Aringhieri, Nanoporosity characteristics of some natural clay minerals and soils, *Clays and clay minerals*, 52 (2004) 700-704.
- [23] M. Gasparik, A. Ghanizadeh, P. Bertier, Y. Gensterblum, S. Bouw, B.M. Krooss, High-pressure methane sorption isotherms of black shales from the Netherlands, *Energy & fuels*, 26 (2012) 4995-5004.
- [24] A.-L. Cheng, W.-L. Huang, Selective adsorption of hydrocarbon gases on clays and organic matter, *Organic geochemistry*, 35 (2004) 413-423.
- [25] D. Bai, X. Zhang, G. Chen, W. Wang, Replacement mechanism of methane hydrate with carbon dioxide from microsecond molecular dynamics simulations, *Energy & Environmental Science*, 5 (2012) 7033-7041.
- [26] P. Weniger, W. Kalkreuth, A. Busch, B.M. Krooss, High-pressure methane and carbon dioxide sorption on coal and shale samples from the Paraná Basin, Brazil, *International Journal of Coal Geology*, 84 (2010) 190-205.
- [27] M. Arabloo, S. Rafiee-Taghanaki, SVM modeling of the constant volume depletion (CVD) behavior of gas condensate reservoirs, *Journal of Natural Gas Science and Engineering*, 21 (2014) 1148-1155.
- [28] A. Tatar, A. Shokrollahi, M. Mesbah, S. Rashid, M. Arabloo, A. Bahadori, Implementing radial basis function networks for modeling CO₂-reservoir oil minimum miscibility pressure, *Journal of natural gas science and engineering*, 15 (2013) 82-92.
- [29] H.A. Zamani, S. Rafiee-Taghanaki, M. Karimi, M. Arabloo, A. Dadashi, Implementing ANFIS for prediction of reservoir oil solution gas-oil ratio, *Journal of Natural Gas Science and Engineering*, 25 (2015) 325-334.
- [30] X. Luo, S. Wang, Z. Wang, Z. Jing, M. Lv, Z. Zhai, T. Han, Adsorption of methane, carbon dioxide and their binary mixtures on Jurassic shale from the Qaidam Basin in China, *International Journal of Coal Geology*, 150 (2015) 210-223.
- [31] P.J. Rousseeuw, A.M. Leroy, *Robust regression and outlier detection*, John Wiley & sons, 2005.
- [32] A.H. Mohammadi, F. Gharagheizi, A. Eslamimanesh, D. Richon, Evaluation of experimental data for wax and diamondoids solubility in gaseous systems, *Chemical engineering science*, 81 (2012) 1-7.
- [33] V. Vapnik, *The Nature of Statistical Learning Theory*. 6. [MJ New York, Springer-Verlag, 1 (1995) 995.
- [34] V. Vapnik, V. Vapnik, *Statistical learning theory*, in, Wiley New York, 1998.
- [35] V. Vapnik, *The support vector method of function estimation*, in: *Nonlinear Modeling*, Springer, 1998, pp. 55-85.
- [36] H. Drucker, C.J. Burges, L. Kaufman, A.J. Smola, V. Vapnik, *Support vector regression machines*, in: *Advances in neural information processing systems*, 1997, pp. 155-161.
- [37] J.A. Suykens, J. Vandewalle, *Least squares support vector machine classifiers*, *Neural processing letters*, 9 (1999) 293-300.
- [38] S.J. AK, *Least squares support vector machines*, World Scientific, 2002.

- [39] Z. Guo, G. Bai, Application of least squares support vector machine for regression to reliability analysis, *Chinese Journal of Aeronautics*, 22 (2009) 160-166.
- [40] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, Ieee, 1995, pp. 39-43.
- [41] O. Castillo, *Type-2 fuzzy logic in intelligent control applications*, Springer, 2012.
- [42] J.E. Onwunalu, L.J. Durlofsky, Application of a particle swarm optimization algorithm for determining optimum well location and type, *Computational Geosciences*, 14 (2010) 183-198.
- [43] A. Sharma, G. Onwubolu, Hybrid particle swarm optimization and GMDH system, in: *Hybrid self-organizing modeling systems*, Springer, 2009, pp. 193-231.
- [44] M.-Y. Chen, A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering, *Information Sciences*, 220 (2013) 180-195.
- [45] J.-S. Chiou, S.-H. Tsai, M.-T. Liu, A PSO-based adaptive fuzzy PID-controllers, *Simulation Modelling Practice and Theory*, 26 (2012) 49-59.
- [46] Y. Shi, R. Eberhart, A modified particle swarm optimizer, in: *1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360)*, IEEE, 1998, pp. 69-73.
- [47] R. Poli, J. Kennedy, T. Blackwell, Particle swarm optimization, *Swarm intelligence*, 1 (2007) 33-57.
- [48] A. Baghban, M. Kahani, M.A. Nazari, M.H. Ahmadi, W.-M. Yan, Sensitivity analysis and application of machine learning methods to predict the heat transfer performance of CNT/water nanofluid flows through coils, *International Journal of Heat and Mass Transfer*, 128 (2019) 825-835.
- [49] F. Zarei, M.R. Rahimi, R. Razavi, A. Baghban, Insight into the experimental and modeling study of process intensification for post-combustion CO₂ capture by rotating packed bed, *Journal of cleaner production*, 211 (2019) 953-961.