

Correlating Stochastically Distributed Reservoir Heterogeneities with Steam-Assisted Gravity Drainage Production

Cui Wang¹, Zhiwei Ma¹, Juliana Y. Leung^{1,*} and Stefan D. Zanon²

¹ School of Mining and Petroleum Engineering, Department of Civil and Environmental Engineering, University of Alberta, Alberta, Edmonton T6G 2W2 - Canada

² Nexen Energy ULC, 801 - 7th Avenue S.W, Alberta, Calgary T2P 3P7 - Canada
e-mail: juliana2@ualberta.ca

*Corresponding author

Abstract — Application of big data analytics in reservoir engineering has gained wide attention in recent years. However, designing practical data-driven models for correlating petrophysical measurements and Steam-Assisted Gravity Drainage (SAGD) production profiles using actual field data remains difficult. Parameterization of the complex reservoir heterogeneities in these reservoirs is not trivial. In this study, a set of attributes pertinent to characterizing stochastic distributions of shales and lean zones is formulated and used for correlating against a number of production performance measures. A comprehensive investigation of the heterogeneous distribution (continuity, size, proportions, permeability, location, orientation and saturation) of shale barriers and lean zones is presented. First, a series of two-dimensional SAGD models based on typical Athabasca oil reservoir properties and operating conditions are constructed. Geostatistical techniques are applied to stochastically model shale barriers, which are imbedded in a region of degraded rock properties referred to as Low-Quality Sand or LQS, among a background of clean sand. Parameters including correlation lengths, orientation, proportions and permeability anisotropy of the different rock facies are varied. Within each facies, spatial variations in water saturation are modeled probabilistically. In contrast to many previous simulation studies, representative multiphase flow functions and capillarity models are assigned in accordance to individual facies. A set of input attributes based on facies proportions and dimensionless correlation lengths are formulated. Next, to facilitate the assessment of different scenarios, production performance is quantified by numerous dimensionless output attributes defined from recovery factor and steam-to-oil ratio profiles. An additional dimensionless indicator is implemented to capture the production time during which the instantaneous steam-to-oil ratio has exceeded a particular economic threshold. Finally, results of the sensitivity analysis are employed as training and testing datasets in a series of neural network models to correlate the pertinent system attributes and the production performance measures. These models are also used to assess the consequences of ignoring lateral variation of heterogeneities when extracting petrophysical (log) data from vertical delineation wells alone. An important contribution of this work is that it proposes a set of input attributes for correlating reservoir heterogeneity introduced by shale barriers and lean zones to SAGD production performance. It demonstrates that these input attributes, which can be extracted from petrophysical logs, are highly correlated with the ensuing recovery response and heat loss. This work also exemplifies the feasibility and utility of data-driven models in correlating SAGD performance. Furthermore, the proposed set of system variables and modeling approach can be applied directly in field-data analysis and scale-up study of experimental models to assist field-operation design and evaluation.

INTRODUCTION

Steam-Assisted Gravity Drainage (SAGD) is considered a proven technology for heavy-oil/bitumen recovery over the past decades (Ipek *et al.*, 2008). In SAGD, steam is injected near the bottom of the pay zone and rises as the steam chamber expands; the heated oil has lower viscosity and tends to drain along the chamber edge towards to the producer located about 5–10 m below the injector (Butler and Mcnab, 1981; Butler, 1985). Its performance, however, can be significantly affected by reservoir heterogeneities (Richardson *et al.*, 1978; Webb *et al.*, 2005). An effective design of the SAGD process requires a detailed understanding of the complex effects of reservoir heterogeneities introduced by shale barriers and varying water saturation (*i.e.*, lean zones) (Pooladi-Darvish and Mattar, 2002; Baker *et al.*, 2008; Xu *et al.*, 2014a).

Richardson *et al.* (1978) developed a mathematical model to study the effects of shale barriers on SAGD performance. The impacts of shale barriers were realized by a reduction in vertical permeability. The model illustrated that both the size and spatial distribution of shale barriers could play an essential role on SAGD production. In the experiments by Joshi and Threlkeld (1985), a box-shaped sand pack was filled with 20–30 mesh (0.84–0.58 mm in diameter) Ottawa sand, where shale barriers were represented by a series of 6 mm thick plastic sheets. Although the final SAGD oil recovery was reduced after a certain time, the initial oil production rate, however, was higher because the injected steam was diverted by the shale barrier to reach regions located above the shale barrier. Yang and Butler (1992) constructed similar models with 2-mm and 3-mm glass beads, while a thin shale layer was represented by a 0.4 cm thick, reinforced, phenolic resin divider. It was concluded that short shale barriers did not affect the recovery performance significantly. These observations were corroborated by numerical simulation studies presented by Chen *et al.* (2008), in which stochastic realizations of the shale distribution were constructed to model varying proportion and continuity of shale barriers. The shale barrier was modeled as a facies with low vertical permeability. Their results showed that thin and laterally-extensive shale barriers posed adverse impacts on steam chamber development, while short shale had only minor effects.

Another origin of reservoir heterogeneity stems from spatially-varying water distribution. Lean zones with high water saturation could act as thief zones causing severe heat loss (Xu *et al.*, 2014a). Highly-permeable lean zones could also promote lateral spreading of the steam chamber. Law *et al.* (2003) studied the impacts of confining/non-confining top water zone with numerical simulations. Inefficient oil production and steam-to-oil ratio were observed as steam chamber expanded into the top water zone. Masih *et al.* (2012) conducted a similar study to assess the influences on

heat transfer and production due to bottom water coning. They concluded that bottom water coning would reduce the oil mobility near the producer, resulting in a lower oil rate and a higher steam-to-oil ratio. A sensitivity analysis was presented in Ricardo (2013), in which the thickness of top/bottom water zones and the pressure of an overlying gas cap were varied. Results of these studies highlighted the adverse impacts on SAGD efficiency due to heat loss through these heterogeneous features. However, quantifying the impacts of randomly-distributed water saturation (as often inferred from well log measurements) is lacking. Xu *et al.* (2014b) presented a simulation study, where the facies model was constructed stochastically, but the water saturation was assumed to be constant within each facies.

The aforementioned studies have provided many useful insights regarding the modeling of shale barriers/lean zones and their impacts on short-/long-term recovery performance; however, a set of input parameters for correlating reservoir heterogeneity introduced by shale barriers and lean zones to SAGD production performance are rarely defined. Another important consideration is the ability to formulate such parameters from well log or other petrophysical data alone, as it would facilitate the construction of data-driven or statistical models for SAGD production analysis.

A dimensionless Shale Indicator (*SI*) was implemented by Amirian *et al.* (2015) to parameterize the location and size of the shale barrier that is in the closest proximity to the well pair. A recent modeling study by Wang and Leung (2015) included a systematic investigation of the heterogeneous distribution of shale barriers and lean zones (top/bottom water layers) in SAGD process. A few drawbacks of their study include: (1) constant porosity and permeability were assigned to each facies; (2) distribution of lean zones within the pay zone was ignored; (3) shales were modeled as deterministic, uniform rectangular barriers distributed in horizontal direction; and (4) the size of the Low-Quality Sand (LQS) is limited to be one-grid block surrounding the shale barrier. Nonetheless, their models have incorporated a LQS that exists as a transition zone between the clean sand and shale barriers, depicting a more realistic representation of the geology commonly observed in the McMurray Formation (Smith *et al.*, 2009; Hubbard *et al.*, 2011). A set of parameters were derived following a sensitivity analysis of the heterogeneous distribution. Artificial Neural Networks (ANN) were employed to correlate those parameters to a number of SAGD performance indicators. ANN has also been used to analyze heavy oil recovery in a number of previous works (Queipo *et al.*, 2002; Ahmadloo *et al.*, 2010; Karambeigi *et al.*, 2011; Popa *et al.*, 2011; Zerafat *et al.*, 2011; Popa and Patel, 2012; Amirian *et al.*, 2015; Ma *et al.*, 2016). In particular, ANN was implemented in Ma *et al.* (2015) recently to construct a series of data-driven models from an actual SAGD field dataset.

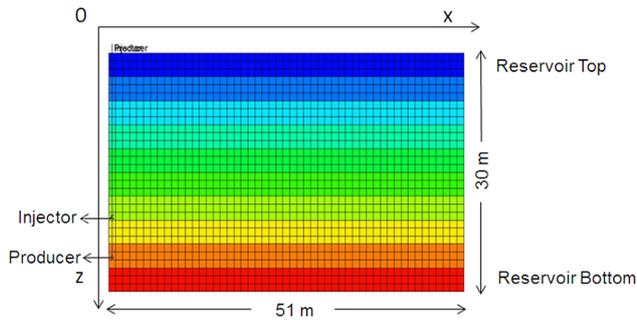


Figure 1

Configuration and set-up for one half of a symmetric 2-D (X-Z plane) SAGD simulation. Reservoir length ($L=51$ m) and thickness ($H=30$ m) are indicated.

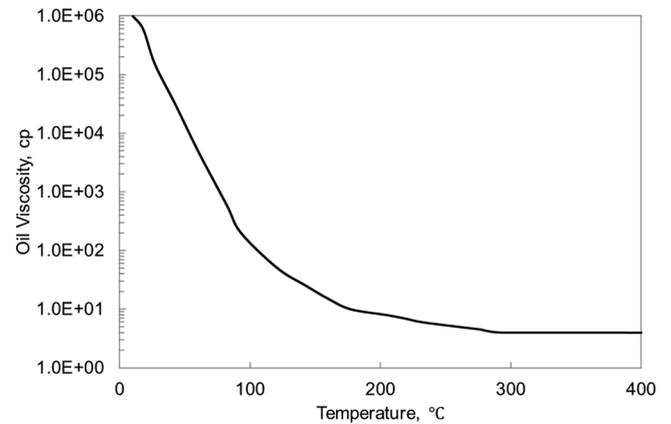


Figure 2

Oil viscosity profile as a function of temperature.

A major drawback of these previous studies is that the proposed parameters capture only limited characteristics of the nearest heterogeneous features; information regarding the overall shale distribution is not represented. The primary objectives of this study are to (1) extend the reservoir heterogeneity description to include stochastically distributed shale barriers and lean zones; (2) formulate and correlate a set of attributes pertinent to characterizing realistic distribution of shale barriers/LQS and lean zones to production performance measures; and (3) demonstrate the potential of data-driven models for correlating production time-series with reservoir heterogeneities. First, a series of 2D SAGD models based on typical Athabasca oil reservoir properties are constructed. Nested sequential indicator simulation is applied to model the heterogeneous distribution (continuity, size, proportions, permeability, location, orientation and saturation) of shale barriers, LQS and clean sand distribution. Next, a set of attributes are identified to represent the characteristics of reservoir heterogeneities. Finally, ANN modeling is used to correlate the pertinent system attributes and production performance measures.

This paper is organized as follows: first, detailed description of methodology and simulation models is presented; next, a set of attributes are identified to describe the reservoir heterogeneities; finally, ANN is used to study the relationships between these inputs and SAGD performance.

1 METHODOLOGY

1.1 Reservoir Model Description

A commercial thermal compositional simulator (Computer Modeling Group, 2013) is used to construct a 2-D (X-Z) numerical model, as shown in Figure 1. Assuming symmetry around the well pair, only one half of the reservoir is modeled. Petrophysical properties, such as the porosity, permeability, initial oil saturation and net pay thickness, for the base case have been taken from representative values of

several pads in the Christina Lake project. Data including well logs and production data are extracted from the public domain (Ma *et al.*, 2015). The pay-zone thickness is 30 m. The model is 51 m and 30 m in the x - and z -direction, respectively, with $\Delta x = \Delta z = 1$ m. This corresponds to a well spacing of approximately 100 m. A well length of 900 m is oriented along the y -direction. The injector-producer pair, which is 5 m apart, is located at 9.5 m from the bottom of the pay zone. The locations of the well pair are also shown in Figure 1. The model set-up depicts a confined drainage pattern, similar to many previous works such as Chen *et al.* (2008). The viscosity of the *in-situ* oil is 600 000 cp at the initial temperature of 18 °C. Figure 2 shows the oil viscosity profile with temperature. For all simulations, 100% quality steam is injected at 1900 kPa continuously for a total simulation time (t_s) of 20 years. A pre-heating period of 2 months is modeled.

1.2 Facies and Rock Property Modeling

A number of conceptual models have been proposed to describe the middle unit of the McMurray Formation. One commonly-adopted model consists of two particular facies: Cross-Stratified Sandstone (CSS) and the Inclined Heterolithic Strata or IHS (Hassanpour *et al.*, 2013). Porosity, permeability and oil saturation are high within the CSS, while characterization of IHS is more complex. It consists of many inclined sets of alternating sandstone and mudstone. These IHS sets can be described as either sandy or muddy, depending on the amount of mud content. The IHS deposits may contain centimeter-scale features that extend over kilometers. Object-based modeling techniques (Hassanpour *et al.*, 2013) have been proposed in recent years to capture the complexities of these features. However, this type of modeling is quite complicated and time-consuming, and it often requires calibration with outcrops.

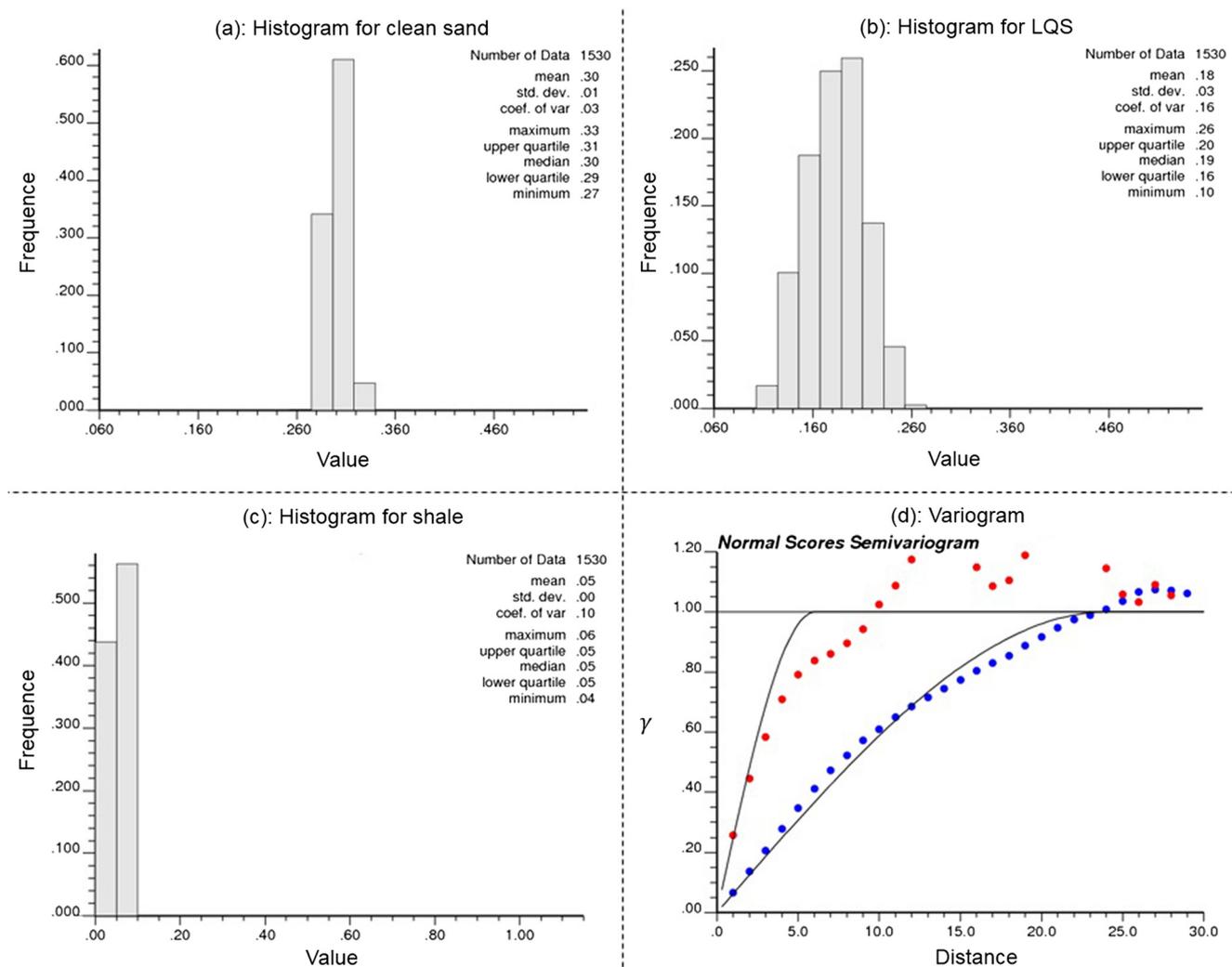


Figure 3

Modeling of facies and porosity distribution: (a-c) histograms of porosity distribution in clean sand, LQS and shale facies; (d) semi-variograms along the direction of maximum anisotropy for the facies modeling.

In this work, a simpler and more widely-used covariance-based technique is implemented to model the distribution of these heterogeneous features. Three different rock facies are modeled: clean sand, shale barrier and LQS. The LQS represents a transition zone between the clean sand and the shale barrier; it also depicts a mixture of muddy and sandy IHS at the larger scale. The LQS is also important for modeling water saturation variability among the reservoir; the presence of clay materials would result in an increase in capillarity and higher initial water saturation.

To model the facies distribution, nested sequential indicator simulation (Deutsch, 2002) is applied to stochastically model shale barriers, which are imbedded in a region of LQS, among a background of clean sand. Within each facies, porosity values are populated using sequential Gaussian simulation (Deutsch and Journel, 1998). LQS properties are modeled to depict a continuous gradation between shale

barrier and clean sand. The histograms of porosity for the three individual facies are shown in Figure 3a-c, while the semi-variogram (γ) for the facies modeling is presented in Figure 3d. Multiple realizations of facies and porosity distribution are modeled.

Same value of thermal conductivity is assigned for all three rock types, as suggested by Hampton *et al.* (2013), while different fluid thermal conductivities are assigned (Yang and Butler, 1992). These values and other reservoir properties are shown in Table 1. These parameters are comparable to those presented in the experimental studies of Yang and Butler (1992) and simulation studies of Chen *et al.* (2008). Chen *et al.* (2008) assumed that the presence of shale in sand reduces the vertical permeability dramatically but has no effect on the horizontal permeability. Following their suggestion, a constant horizontal permeability of $5D$ is applied to all three facies in this study, and factors of 10^{-8} – 10^{-4} and 0.2–0.8 is applied to

TABLE 1
Reservoir and fluid properties.

Parameters	Value	Unit
Reservoir depth	200	m
Reservoir pressure	1500	kPa
Well length	900	m
Horizontal permeability	5	D
Reservoir temperature	18	°C
Oil viscosity (18 °C)	6 00 000	cp
Rock compressibility	9.6×10^{-6}	kPa^{-1}
Formation heat capacity	2350	$\text{kJ/m}^3 \cdot \text{K}$
Rock thermal conductivity	6.6×10^5	$\text{J/m} \cdot \text{d} \cdot \text{°C}$
Oil thermal conductivity	1.15×10^4	$\text{J/m} \cdot \text{d} \cdot \text{°C}$
Water thermal conductivity	5.35×10^4	$\text{J/m} \cdot \text{d} \cdot \text{°C}$
Gas thermal conductivity	1.4×10^2	$\text{J/m} \cdot \text{d} \cdot \text{°C}$

the k_v/k_h in shale barriers and clean sand, respectively, which are similar to the values provided in [Chen et al. \(2008\)](#) and [Dang et al. \(2010\)](#). The continuous variation in k_v/k_h within each facies is modeled by sequential Gaussian simulation. The variation in k_v/k_h ratio in the LQS is modeled as the arithmetic average between the clean sand and shale barrier.

1.3 Saturation Modeling

Archie model ([Archie, 1942](#)) is used to model the varying saturation distribution within the reservoir as a function of local porosity and formation resistivity:

$$S_w = \left(\frac{aR_w}{\phi^m R_t} \right)^{1/n}, \quad (1)$$

where a , m = cementation factor; n = saturation exponent; ϕ = porosity in fraction; R_t = true or formation resistivity, and R_w = water resistivity. [McCoy and Grieves \(1997\)](#) demonstrated that values a and m do not vary significantly for different facies; however, n is observed to be the highest for clean sand and the lowest for shale barrier. The values of R_w and R_t can be derived from the well log measurements. Values of these parameters, as summarized in [Table 2](#), are assigned based on the representative trends for different rock groups presented in [Palacky \(1987\)](#). The formation resistivity is assumed to be constant within the facies. An example of facies, porosity and water-saturation distributions is shown in [Figure 4](#). In these models, S_w in sand or LQS is typically ≤ 0.2 (ranging between 0.15 and 0.3), while S_w in shale is close to 1. These ranges are in agreement with log data from the field.

TABLE 2
Parameters in archie equation.

Parameter	Clean sand	LQS	Shale	Unit
a	1	1	1	–
m	2	2	2	–
n	2	2.5	3	–
R_t	500	300	100	$\Omega \cdot \text{m}$
R_w	0.2	0.2	0.2	$\Omega \cdot \text{m}$

1.4 Multiphase Flow Functions

Relative permeability relationships of oil-water and gas-liquid systems for all three facies are adopted from [Wang and Leung \(2015\)](#), in which it was concluded that capillary pressure could not be ignored, especially when shale barriers are extensively distributed and present in regions between the injection and production wells.

1.5 Performance Ranking

The following dimensionless indicator based on oil Recovery Factor (RF) and Cumulative Steam-to-Oil Ratio ($CSOR$) is implemented:

$$R = \frac{RF}{CSOR}, \quad (2)$$

where $RF = COP/OOIP$ = Cumulative Oil Production/Original Oil In Place and $CSOR$ = Cumulative Steam-to-Oil Ratio. Another measure based on the concept of Discounted Barrel (DB) of oil is also implemented ([Wang and Leung, 2015](#)). It takes into account the economic impact of steam consumption and variation in fluid properties. In the simplest term, it represents the net energy that can be obtained from the process.

Another important variable in describing the efficiency of steam injection is t_{iSOR} , which is defined as the duration over which the monthly average Steam-to-Oil Ratio ($iSOR$) exceeds 5 (a commonly-accepted upper limit for typical SAGD wells). For a given value of COP, low values of t_{iSOR} and $CSOR$ would correspond to higher steam injection efficiency. A dimensionless form defined as $t_{DISOR} = t_{iSOR}/t_s$ is considered as an output attribute in the ANN modeling. In this work, $t_s = 7304$ days (20 years).

1.6 Neural Network Modeling

The ANN is an artificial intelligence technique useful for approximating the complex nonlinear relationship between a set of input and target variables. A schematic of

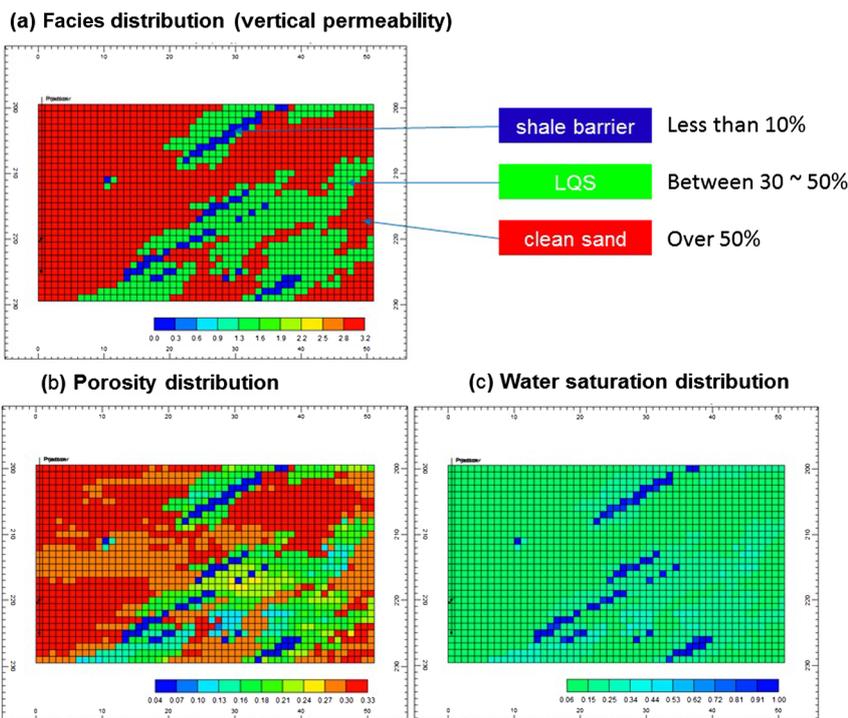


Figure 4
Facies, porosity and water-saturation distributions for a randomly selected realization.

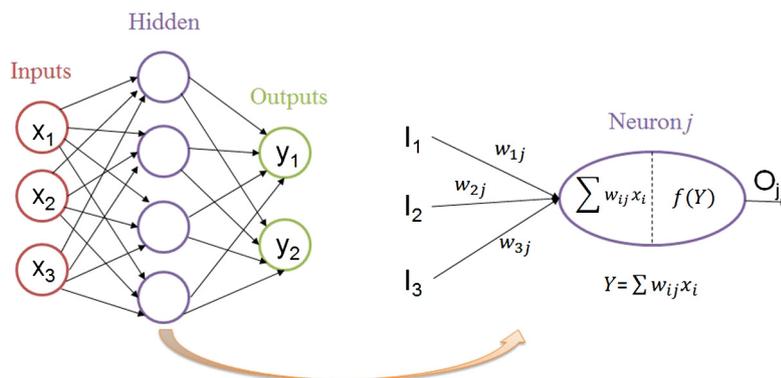


Figure 5
An artificial neural network is an interconnected group of nodes. The circle represents an artificial neuron and the arrow represents a connection between two neurons.

various elements of an ANN consisting of only one hidden layer is illustrated in Figure 5. The network consists of an input layer, an output layer and a number of hidden layers. A series of neurons are assigned in the hidden layers. Weighted summation of the signals from the input layer is calculated according to Equation (3):

$$Y_j = w_0 + \sum_{i=1}^n w_{ij} x_i, \quad (3)$$

where Y_j is weighted sum of input signals at node j ; w_0 is threshold (bias); w_{ij} is the weight associated with the connection between node j and the input node i ; x_i is the signal of input node i ; n is number of input nodes. An activation function is applied to the weighted sum according to Equation (4) to compute the output signal from node j , which is considered as the input signal to the next layer.

$$f(Y) = \frac{1}{(1 + e^{-Y})}. \quad (4)$$

Equations (3) and (4) are applied repeatedly until the final output layer is reached and value for the target variable is computed. Network parameters (*i.e.*, weights and biases associated with each connection) are determined in a supervised learning procedure (Francis, 2001). For example, in a Back-Propagation Neural Network (BPNN), the error is back-propagated from the output layer to train the unknown network parameters in a gradient-descent minimization algorithm (Bishop, 1995), during which the mismatch between network predictions and known values of the target variables is minimized with a set of training data consisting of known input and output attributes. Due to the large disparity in scales of different data sources, normalization or standardization procedures are performed on all input and output attributes (Francis, 2001). The training dataset is also used to design the optimal network configuration. Readers may refer to various references for additional information about ANN (Zupan 1994; Shahab, 1995; Shahab, 2000, Al-Fattah and Startzman, 2001; Weiss et al., 2002). In addition, Principal Component Analysis (PCA) is performed to reduce the dimensionality of the original dataset (X) through a linear projection onto a lower-dimensional subspace. PCA is also useful for removing correlation among the inputs of the original dataset, while retaining much of its information. A mean-adjusted dataset (Z) is attained by subtracting the mean of each variable in X . Next, the covariance between two variables is calculated to eliminate bias due to large disparity in mean values.

$$COV(X_j, X_k) = \frac{\sum_{i=1}^M Z_{ij}Z_{ik}}{M-1}, \quad (5)$$

where X_j and X_k represent two particular variables in X ; M denotes the total number of samples in X . The calculation in Equation (5) is repeated for all pairs of variables to compute the covariance matrix, which is subjected to eigenvalue decomposition. Individual eigenvalue signifies the contribution of the variance from the corresponding eigenvector to the total variance of the original data. The eigenvectors with highest eigenvalues, or the Principal Components (PC), can be obtained by sorting the eigenvalues in decreasing order. Principal Scores (PS), which are regarded as new inputs attributes in subsequent ANN modeling, are computed using Equation (6).

$$PS = PC \times Z^T. \quad (6)$$

1.7 Grid Sensitivity Analysis

In order to select a cell size that balances both accuracy (numerical dispersion) and computational efficiency, three different cell sizes ($\Delta x \times \Delta y$) are tested: 1 m \times 1 m (coarse),

0.5m \times 0.5m (medium), 0.25m \times 0.25m (fine), respectively. The average initial water saturation over the entire reservoir is 0.35, and the proportions of shale barriers, LQS and sand are 20%, 40% and 40% respectively. Only minor differences (less than 0.1%) in COP are observed among the three cases after 20 years of production. The run-time for the 0.25m \times 0.25m model is 34 and 2154 times those of the 0.5m \times 0.5m model and the 1m \times 1m model, respectively on a 3.4 GHz, 16 GB of RAM, Intel[®] Core[™] i7-2600 CPU. Therefore, a cell size of 1m \times 1m is selected in this study.

2 CASE STUDY

A suite of stochastic realizations of rock property distribution is constructed by varying facies proportions, vertical permeability, correlation range and orientation (*i.e.*, continuity and direction of anisotropy) of shale barriers and LQS. Water saturation and lean zone distributions are modeled as described in the previous section. The reservoir is assumed to be composed of primarily clean sand, with the combined proportions of shale barriers and LQS to be less than 50%. In particular, LQS proportion is approximately 30–50%, while the shale barrier proportion is less than 10%, as LQS depicts a gradation of properties between shale barriers and clean sand.

In order to verify the negative impacts of lean zones, two particular scenarios are compared: (1) lean zones are distributed stochastically, and (2) mean saturation in the clean sand is assigned everywhere. It is clear that the steam chamber advances more quickly in the reservoir with stochastically distributed lean zones. Although a higher RF is observed, the corresponding $CSOR$ is also increased; as a result, both R and DB decrease, while t_{ISOR} increases. Lean zones may accelerate the expansion of steam chamber by providing alternative pathway for water drainage and serving as easy conduits for steam entry, as steam does not need to mobilize the bitumen in order to gain penetration into the reservoir. However, additional energy is consumed to vaporize the water in the lean zones. This observation illustrates the importance of capturing the effects of S_w variations in the analysis.

A number of parameters related to these heterogeneous features are introduced and explained in Figure 6. Here, the correlation lengths of the LQS and shale barriers along the maximum direction of anisotropy (L_{LQS} and L_{sh}) are normalized against the total reservoir length in the x -direction (L): $L_{DLQS} = L_{LQS}/L$ and $L_{Dsh} = L_{sh}/L$, respectively. Similarly, the correlation lengths along the minimum direction of anisotropy (H_{LQS} and H_{sh}) are normalized against the total reservoir thickness (H): $H_{DLQS} = H_{LQS}/H$, and $H_{Dsh} = H_{sh}/H$. Since the physical thickness of a shale barrier is generally smaller than the resolution of the

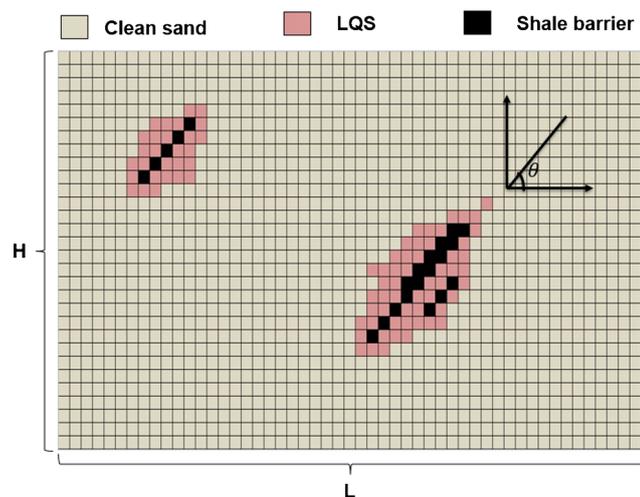


Figure 6

Schematic illustrating the definitions of various parameters employed in this study.

numerical model, H_{sh} is assumed to be the size of a single grid block of 1 m; as a result, H_{Dsh} is also a constant. Orientation of these heterogeneous features is represented by the azimuth angle or θ , which is measured anti-clockwise away from the x -direction. θ_{LQS} in the LQS is assumed to be the same as θ_{sh} in the shale. The rationale is that LQS serves to model the transitioning property between a shale barrier and the background clean sand; it essentially represents a shaly region near a shale barrier. It would be reasonable to assume that the geological processes contributing to the formation of both shale barriers and shaly regions to be similar; hence the orientation should be the same. A range of value between 0 and 30° has been used for θ_{sh} . Prior to ANN modeling, these azimuth angles are normalized to dimensionless forms as: $\theta_{DLQS} = \theta_{LQS}/360^\circ$ and $\theta_{Dsh} = \theta_{sh}/360^\circ$. These parameters have sufficiently captured the essential information extracted from the semi-variogram models. In addition, the dimensionless shale indicator $SI_D = V/d_{min}wH$ is considered, where d_{min} refers to the closest distance between a shale barrier and the producer, and w is the width of the reservoir. The product $d_{min}wH$ describes the volume of rock between the vertical plane at the heel of the well pair and the corresponding shale barrier; SI_D essentially represents the ratio of the shale-barrier volume to the rock volume that exists between the well-pair and the shale barrier. A thick shale barrier located close to the well pair can be reflected by a high SI value. Other parameters including average k_v/k_h and average water saturation in reservoir S_w are computed. All these variables are considered as input attributes for the subsequent ANN modeling, and their ranges are summarized in Table 3.

A dataset consisting of a total of 249 cases with stochastically distributed facies, rock properties and lean zones is assembled for ANN modeling. The total dataset (m)

TABLE 3

Range of dimensionless input attributes for ANN modeling.

Variable	Min	Max
LQS proportion	0.2484	0.4346
L_{DLQS}	0.2941	0.3922
H_{DLQS}	0.1667	0.3333
Shale proportion	0.0307	0.0804
L_{Dsh}	0.0980	0.1961
$\theta_{DLQS} = \theta_{Dsh}$	0.0000	0.0833
k_v/k_h in Clean Sand	0.2052	0.7971
k_v/k_h in Shale	5.31E-07	9.96E-05
S_w	0.1348	0.1981
SI_D	4.09E-05	5.14E-04

is partitioned into two parts: (1) m_1 samples are designated for training and validation of the BPNN model; an n -fold cross-validation is implemented to identify the optimal network architecture (Ma et al., 2015); and (2) the remaining ($m_2 = m - m_1$) samples are assigned for final testing in a prediction mode using the previously trained network parameters. In this study, the entire dataset is subdivided into 209 cases for training and validating, with the remaining 40 cases for testing.

ANN modeling is implemented in Matlab (R2009b) to correlate the non-linear relationships between the extracted input attributes and multiple SAGD performance indicators (R, DB and t_{DISOR}). A total of 10 dimensionless input attributes including LQS proportion, L_{DLQS} and H_{DLQS} , shale proportion, L_{Dsh} , θ_{Dsh} , SI_D , k_v/k_h in clean sand and shale, and S_w are adopted in this work.

3 RESULTS AND DISCUSSION

Cross-plots between actual target values and the network predictions for the training datasets are shown in Figure 7. The predictive quality of the ANN models is quantified by the coefficient of determination (R^2) and Mean Squared Error (MSE) between the target and predicted values. The corresponding results are compared in Tables 4 and 5. Overall, reasonable agreement is observed between target and predicted values. The results demonstrate that the parameterization scheme is useful for capturing reservoir heterogeneities pertinent to production performance, as evidenced by the high R^2 and low MSE values corresponding to all three outputs for both training and testing datasets. The results for the testing dataset are also shown in Figure 7. It appears that model predictability for t_{DISOR} is not as good

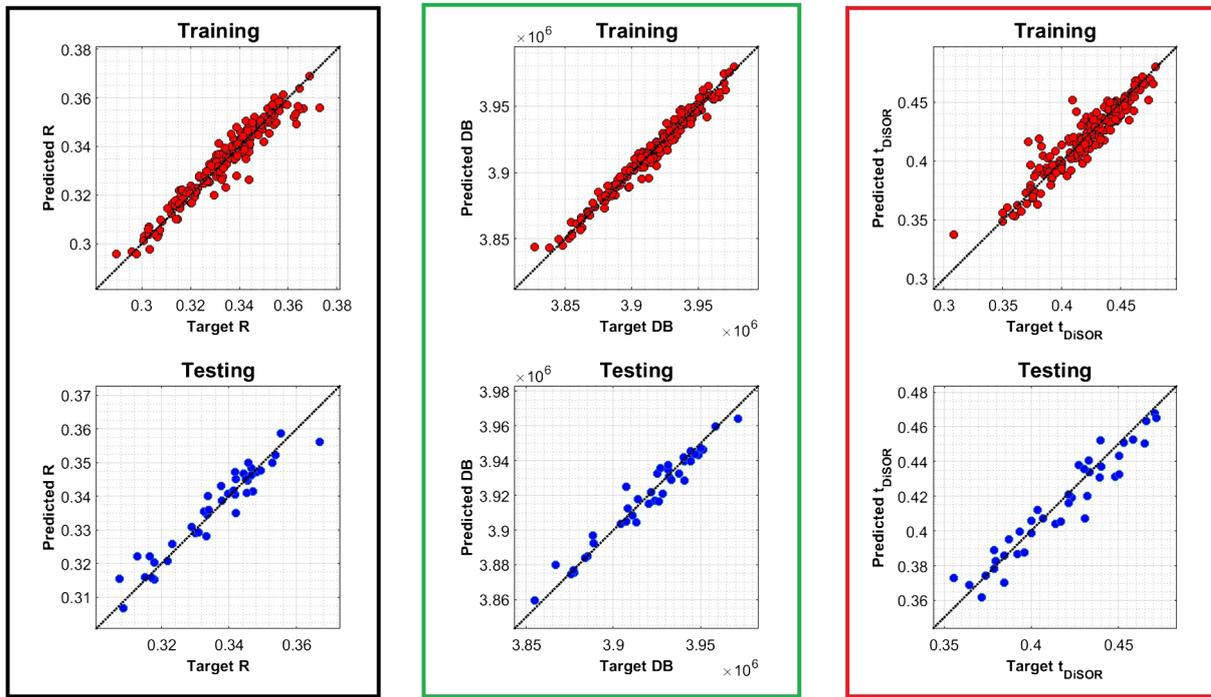


Figure 7

Cross plot of actual flow simulation target values of R (in black frame), DB (in green frame) and t_{DISOR} (in red frame) against network predictions using all 10 original input attributes: training dataset (top) and testing dataset (bottom).

TABLE 4
Coefficient of determination (R^2) for ANN modeling.

Output attributes	Base scenario				Ignoring lateral heterogeneity	
	Original 10 input attributes		6 Principal scores		Original 7 input attributes	
	Training	Testing	Training	Testing	Training	Testing
R	0.9281	0.9178	0.9256	0.9241	0.8875	0.7882
DB	0.9717	0.9510	0.9703	0.9614	0.9573	0.9317
t_{DISOR}	0.8777	0.9140	0.8818	0.8857	0.8474	0.9087

as for the other two output variables, as minor deviation between target and predicted values is observed. Considering that some of the input parameters are correlated (e.g., S_w is related to facies proportions), it is postulated that removal of data redundancy may improve the model robustness and predictive capability. Therefore, PCA is performed next to remove correlation among the input parameters in the original dataset.

Figure 8 shows the variance contribution of each principal component in decreasing order. The variances of the first six components appear to dominate (accounting for >99.6% of the total variance), suggesting that the remaining four components can be neglected. Therefore, the modeling has been repeated after reducing the dimensionality of the

original input space to 6. The application of PCA has reduced the number of input attributes to 6 principal scores. The results are summarized in Figure 9 and Tables 4 and 5. The ANN performance with only 6 PCA-transformed inputs is comparable to that obtained with all 10 original input variables. The results would suggest that there is some redundancy among the data; for instance, SI_D and S_w are dependent on the other input variables, as they are computed based on realizations generated from the other variables. This dependency can be removed via PCA, and reducing the dimension does not compromise the model prediction. Another important remark is that with a relative small set of input variables (10 in this case), which have been selected with care to capture much information about the underlying

TABLE 5
Mean Squared Error (MSE) for ANN modeling.

Output attributes	Base scenario				Ignoring lateral heterogeneity	
	Original 10 input attributes		6 Principal scores		Original 7 input attributes	
	Training	Testing	Training	Testing	Training	Testing
R	1.74E-05	1.57E-05	1.80E-05	1.44E-05	2.72E-05	4.03E-05
DB	2.39E + 07	3.65E + 07	2.51E + 07	2.88E + 07	3.61E + 07	5.09E + 07
t_{DisOR}	1.03E-04	8.61E-05	9.99E-05	1.14E-04	1.29E-04	9.15E-05

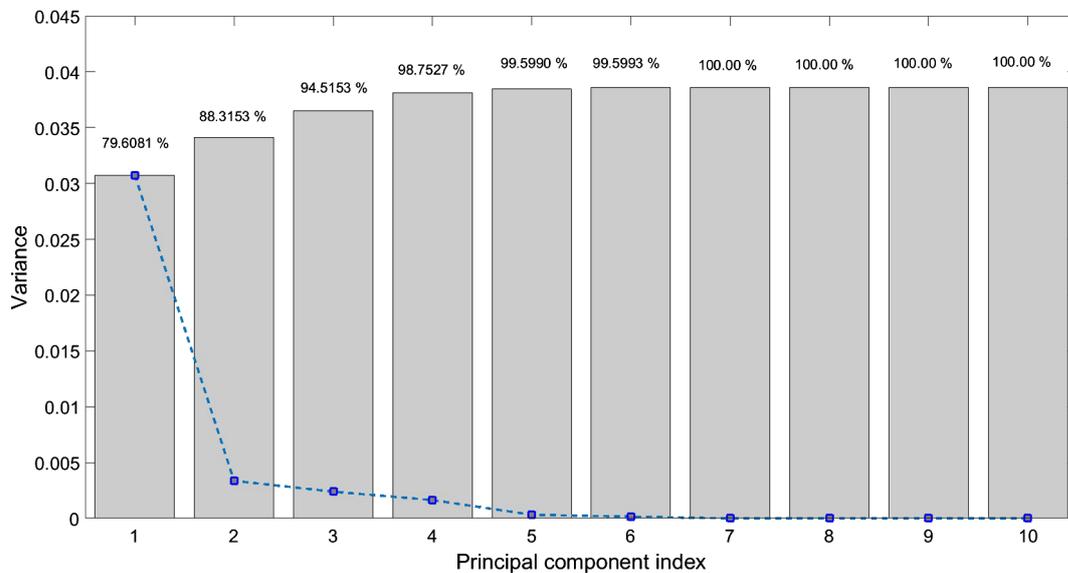


Figure 8

Principal component analysis of the original input dataset.

heterogeneity and to minimize internal data redundancy, it is reasonable to expect that the removal of internal data redundancy would not lead to a significant improvement in model training and prediction; nevertheless, it is an important step in enhancing the robustness of data-driven models. Given the wide range of heterogeneity in the dataset, the results support the validity of the proposed set of parameters for correlating reservoir heterogeneity introduced by shale barriers and lean zones to SAGD production performance. It is highly possible that a group of reservoir models with similar characteristics would exhibit similar production behavior; future work should incorporate techniques such as cluster analysis to identify these internal data groupings.

Due to the randomness often associated with stochastically-distributed properties in heterogeneous reservoirs, parameterization of these properties for correlating against SAGD production performance is not well defined in the literature. In contrast to other recovery technologies, SAGD

production is highly dependent on effective steam chamber growth, it is particularly important to capture heterogeneous features that have significant impact on steam chamber development.

It is proposed that the same parameterization strategy can be adopted when analyzing actual field data. However, in most practical scenarios, petrophysical logs are available only at vertical delineation wells (Ma *et al.*, 2015), rendering conditioning data for heterogeneity modeling to be available at just a few sampled locations. To explore the potential implications of this limitation, the modeling workflow is repeated, where only information along the vertical direction of wellbore is extracted to calculate the input attributes. This procedure aims to replicate a practical field dataset. Some modifications have to be made in the input attribute formulation. For example, parameters including θ_{Dsh} , L_{DLQS} and L_{Dsh} are omitted, as they are difficult to be inferred from well log data alone (additional geological studies and

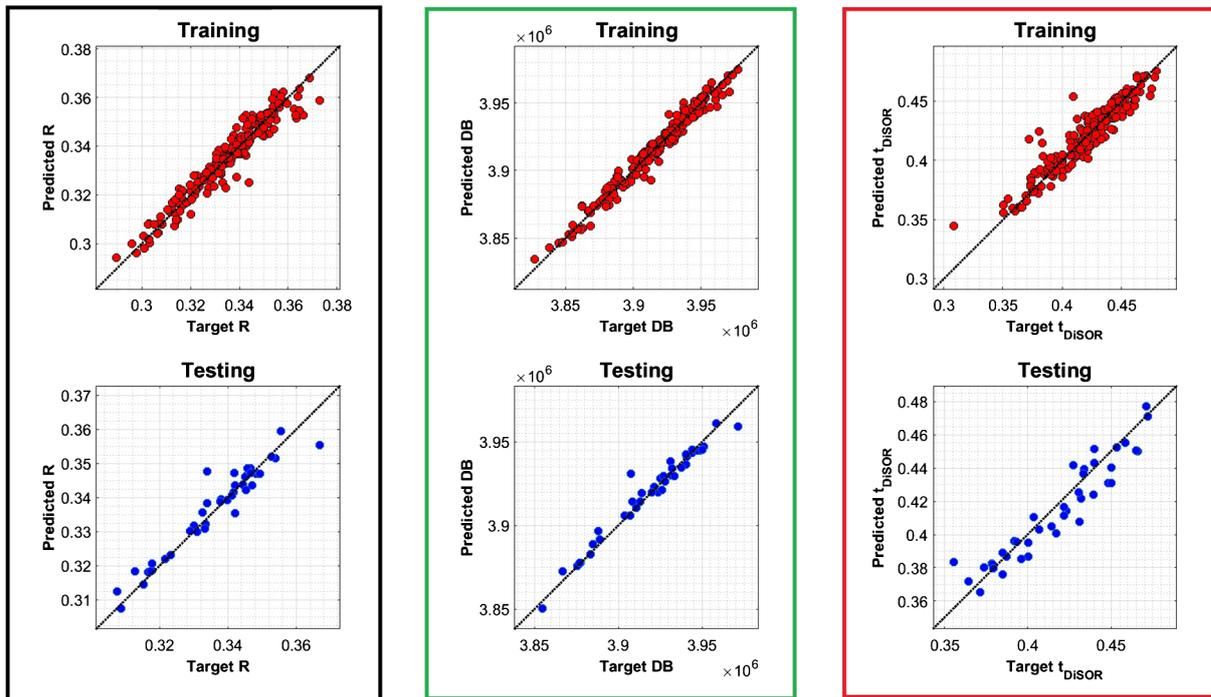


Figure 9

Cross plot of actual flow simulation target values of R (in black frame), DB (in green frame) and t_{DISOR} (in red frame) against network predictions using 6 principal scores as input attributes: training dataset (top) and testing dataset (bottom).

understanding of the depositional environment are needed to constrain these parameters). With a reduced input attribute space and less redundancy among data, no PCA is applied. Therefore, a total of 7 input variables and 3 outputs are employed. It is obvious that the extracted input variables would capture only a portion of the information related to the reservoir heterogeneity. The objectives of this analysis are to: (1) assess the performance of ANN models when only limited petrophysical information is available; and to (2) illustrate the potential implications of ignoring inter-well heterogeneities. To facilitate reliable comparison with the previous results, the same 249 data samples of the training and testing are used here, and the results are presented in Figure 10. Though more scattering can be observed in comparison to Figures 7 and 9, most data points follow the 45° line. As shown in Tables 4 and 5, although the prediction accuracy has dropped, comparing to the ANN models constructed in the previous section, a reasonably high coefficient of determination (approximately 0.8) can still be achieved. Simulation studies and field experience have confirmed that SAGD performance is more sensitive to the variability in lateral extent, rather than thickness, of the shale barriers, it is expected that not capturing the effects of θ_{Dsh} , L_{DLQS} and L_{Dsh} would, to a certain extent, compromise the model fidelity. The results show that the reduced set of input parameters is still highly correlated to SAGD performance, highlighting the robustness of the proposed parameterization

scheme. If data, either dynamic (e.g., 4D seismic and temperature measurements) or static (regional geologic mapping) is available to estimate the lateral distribution of the shale barriers, it can be readily incorporated to constrain the 3 missing input variables.

CONCLUSION

A set of 2D models capturing a range of heterogeneous features commonly encountered in SAGD reservoirs are constructed. The petrophysical properties are representative of a typical reservoir in the Athabasca oil sands. Continuity, size, proportions, permeability, location, orientation and saturation of shale barriers and a surrounding region of degraded rock properties referred (LQS) are modeled stochastically. Capillarity and relative permeability effects, which have been ignored in many previous simulation studies, are incorporated in the shales to model bypassed oil trapping. Despite these aspects of heterogeneities have been studied to various extents in the past, the models used in this paper have integrated all these effects into a single comprehensive dataset.

The main objective of this study is to identify a set of parameters for correlating the effects of heterogeneities in lean zones and shales to SAGD performance. Ten dimensionless input attributes including LQS proportion, L_{DLQS} and

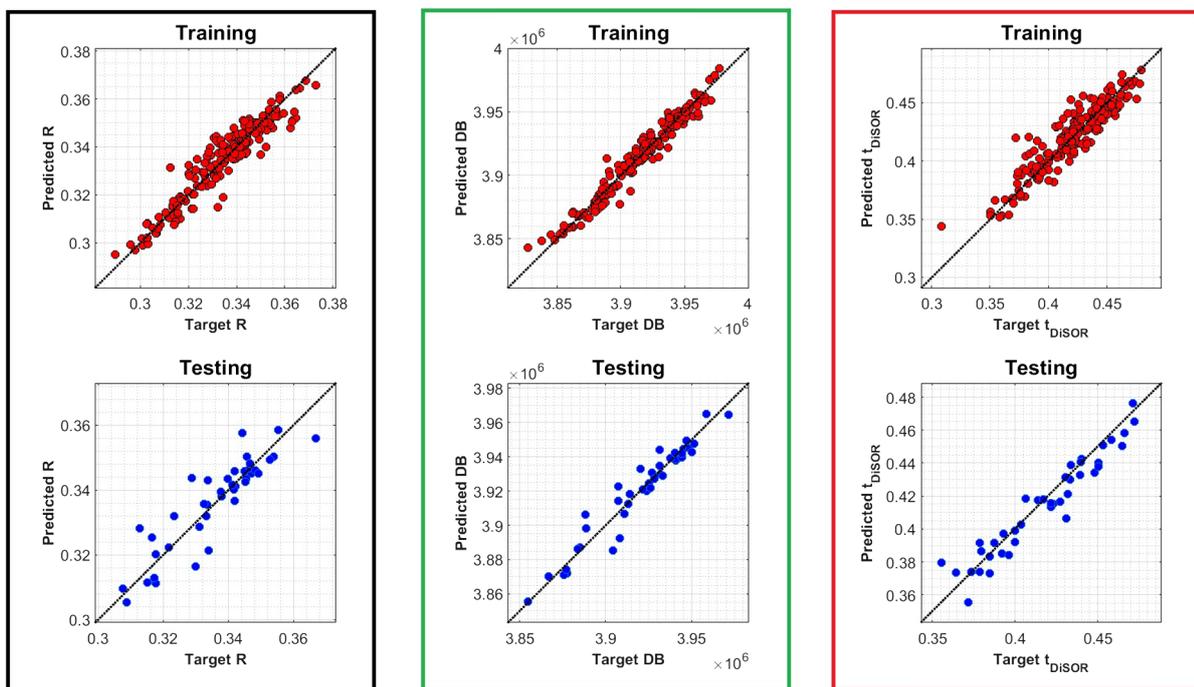


Figure 10

Cross plot of actual flow simulation target values of R (in black frame), DB (in green frame) and t_{DISOR} (in red frame) against network predictions for the case where inter-well heterogeneities are ignored by using all 7 original input attributes: training dataset (top) and testing dataset (bottom).

H_{DLQS} , shale proportion, L_{Dsh} , θ_{Dsh} , SI_D , k_v/k_h in clean sand and shale, and S_w are formulated to capture the heterogeneities due to proportions, location, orientation, continuity and saturation of different facies. Three production performance indicators (R , DB and t_{DISOR}) that take into account both recovery factor (RF) and cumulative steam injection efficiency ($CSOR$) are employed to facilitate the comparison of production performance. An artificial neural network involving the ten input attributes and three production performance indicators is constructed. It is observed that removal of internal data redundancy (or correlation among input parameters) with techniques such as PCA is helpful for improving model robustness and accuracy.

Modeling of heterogeneous reservoirs is often facilitated by random variables and multiple realizations; however, parameterization of stochastically distributed properties for the purpose of production correlation is not well understood, particularly in SAGD reservoirs. Therefore, an important contribution of this work is that it explores and proposes a set of dimensionless variables that are demonstrated to correlate reasonably well with a number of production performance indicators.

A particular case study is presented to assess the utility of these parameters and the predictability of the ensuing ANN models, when only limited information is available along the vertical direction above a given well pair. The results, though not as good as those where inter-well heterogeneities

are accounted for, demonstrate promising potential in the application with practical field dataset typically consisting of only petrophysical logs. This work also highlights the feasibility and utility of data-driven models in correlating SAGD performance. Future work should model the effects of gas cap. It should also incorporate time series data such as the oil rate and instantaneous SOR as the ranking criteria.

NOMENCLATURE

Symbols

a_w	constant, dimensionless
a_o	constant, dimensionless
c_o	constant, dimensionless
c_w	constant, dimensionless
C_w	specific heat for water, J/kg/°C
d	distance between shale barrier and producer, m
$f(Y)$	activation function
H	reservoir thickness, m
H_{sh}	thickness of shale barrier, m
H_{LQS}	thickness of LQS, m
H_{Dsh}	dimensionless thickness of shale barrier
H_{DLQS}	dimensionless thickness of LQS
k_v/k_h	vertical-to-horizontal permeability ratio, mD/mD
L	reservoir length, m
L_{sh}	length of shale barrier, m

L_{LQS}	length of LQS, m
L_{Dsh}	dimensionless length of shale barrier
L_{DLQS}	dimensionless length of LQS
m	cementation factor, number of samples in a dataset
m_1	number of samples for training and validation
m_2	number of samples for testing
m_w	water mass in kg per m ³
M	number of samples in X
n	number of input nodes or saturation exponent in equation (1)
Q_o	energy of produced oil per cubic meter, J/m ³
Q_w	energy required to generate one cubic meter of steam, J/m ³
R	ranking indicator
R^2	coefficient of determination
R_t	true or formation resistivity, Ω -m
R_w	water resistivity, Ω -m
S_w	water saturation, m ³ /m ³
S_w	average water saturation in reservoir, m ³ /m ³
t_{iSOR}	duration over which the monthly average steam-to-oil ratio exceeds a threshold, day
t_{DiSOR}	dimensionless form of t_{iSOR}
t_s	simulation time, year
T_s	surface temperature, °C
V	volume of shale barrier, m ³
V_w	vaporization heat for water, kJ/kg
w	width of reservoir, m
w_0	bias
w_{ij}	weight associated with the connection between nodes i and j
x_i	signal from input node i
Y_j	weighted sum of input signals
ρ_o	oil density, kg/m ³
γ	semi-variogram
ϕ	porosity in fraction
θ	orientation, °
θ_{sh}	orientation of shale barrier, °
θ_{LQS}	orientation of LQS, °
θ_{Dsh}	dimensionless orientation of shale barrier
θ_{DLQS}	dimensionless orientation of LQS

Acronyms

ANN	Artificial Neural Network
BPNN	Back-Propagation Neural Network
COP	Cumulative Oil Production
CSS	Cross-Stratified Sandstone
CSOR	Cumulative Steam-to-Oil Ratio
DB	Discounted Barrel
IHS	Inclined Heterolithic Strata
iSOR	monthly average Steam-to-Oil Ratio
LQS	Low Quality Sand
MSE	Mean Squared Error

OOIP	Original Oil In Place
PC	Principal Component
PCA	Principal Component Analysis
PS	Principal Score
RF	Recovery Factor
SAGD	Steam-Assisted Gravity Drainage
SI _D	Dimensionless Shale Indicator
SOR	Steam-to-Oil Ratio
X	original data vector
Z	mean-adjusted data vector

ACKNOWLEDGMENTS

This research was supported by Nexen Energy ULC under the Collaborative Research and Development Grant program administered by *Natural Sciences and Engineering Research Council of Canada (NSERC)*. The authors would also like to thank the University of Alberta for granting access to the Numerical and Statistical Server. Academic licenses for STARS are provided by the *Computer Modelling Group (CMG) Ltd.*

REFERENCES

- Ahmadloo F., Asghari K., Renouf G. (2010) A new diagnostic tool for performance evaluation of heavy oil water floods: Case study of western canadian heavy oil reservoirs, *SPE Western Regional Meeting*, Anaheim, California, USA.
- Al-Fattah S.M., Startzman R.A. (2001) Predicting natural-gas production using artificial neural network, *SPE Hydrocarbon Economics and Evaluation Symposium*, Dallas, Texas, USA.
- Amirian E., Leung J.Y., Zanon S., Dzurman P. (2015) Integrated cluster analysis and artificial neural network modeling for steam-assisted gravity drainage performance prediction in heterogeneous reservoirs, *Expert Syst. Appl.* **42**, 723–740.
- Archie G.E. (1942) The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics. *SPE-942054-G Trans. AIME* **146**.
- Baker R.O., Fong C., Li T., Bowes C., Toews M. (2008) Practical considerations of reservoir heterogeneities on SAGD Projects, *International Thermal Operations and Heavy Oil Symposium*, October 20–23, Calgary, Alberta, Canada.
- Bishop C.M. (1995) *Neural networks for pattern recognition*, Clarendon Press, Oxford.
- Butler R.M. (1985) A new approach to the modeling of steam-assisted gravity drainage, *J. Can. Pet. Technol.* **24**, 03, 42–51.
- Butler R.M., McNab G.S. (1981) Theoretical studies on the gravity drainage of heavy oil during *in-situ* steam heating, *Can. J. Chem. Eng.* **59**, 04, 455–460.
- Chen Q., Gerritsen M.G., Kovscek A.R. (2008) Effects of reservoir heterogeneities on the steam-assisted gravity-drainage process, *SPE Reserv. Eval. Eng.* **11**, 05, 921–932.
- Computer Modeling Group. 2013. STARS: Advanced Processes & Thermal Reservoir Simulator User's Guide (Version 2013), Computer Modeling Group Limited, Calgary, Alberta, Canada.

- Dang C.T.Q., Nguyen N.T.B., Bae W., Nguyen H.X., Tu T., Chung T. (2010) Investigation of SAGD recovery process in complex reservoir, *Paper Presented at Asia Pacific Oil and Gas Conference and Exhibition*, 18–20 October, Brisbane, Queensland, Australia.
- Deutsch C.V. (2002) *Geostatistics reservoir modeling*, Oxford University Press, Oxford, UK.
- Deutsch C.V., Journel A.G. (1998) *GSLIB. Geostatistical software library and user's guide*, 2nd ed, Oxford University Press, New York.
- Francis L. (2001) The basics of neural networks demystified, *Contingencies*, **11**, 12, 56–61.
- Hampton T., Kumar D., Azom P., Srinivasan S. (2013) Analysis of impact of thermal and permeability heterogeneity on SAGD performance using a semi-analytical approach, *SPE Heavy Oil Conference*, June 11–13, Calgary, Alberta, Canada.
- Hassanpour M.M., Pyrcz M.J., Deutsch C.V. (2013) Improved geostatistical models of inclined heterolithic strata for McMurray formation, *AAPG Bull.* **97**, 7, 1209–1224.
- Hubbard S.M., Smith D.G., Nielsen H., Leckie D.A., Fustic M., Spencer R.J., Bloom L. (2011) Seismic geomorphology and sedimentology of a tidally influenced river deposit, lower cretaceous Athabasca oil sands, Alberta, Canada, *Am. Assoc. Pet. Geol.* **95**, 07, 1123–1145.
- Ipek G., Frauenfeld T., Yuan J.Y. (2008) Numerical study of shale issues in SAGD, *Canadian International Petroleum Conference*, 17–19 June, Calgary, Alberta, Canada.
- Joshi S.D., Threlkeld C.B. (1985) Laboratory studies of thermally aided gravity drainage using horizontal wells, *ASOTRA J. Res.* **2**, 1, 11.
- Karambeigi M.S., Zabihi R., Hekmat Z. (2011) Neuro-simulation modeling of chemical flooding, *J. Pet. Sci. Eng.* **78**, 2, 208–219.
- Law D.H.-S., Nasr T.N., Goog W.K. (2003) Field-Scale numerical simulation of sagd process with top-water thief zone, *J. Can. Pet. Technol.* **42**, 08, 32–38.
- Ma Z., Leung J., Zanon S., Dzumann P. (2015) Practical implementation of knowledge-based approaches for SAGD production analysis, *Expert Syst. Appl.* **42**, 21, 7326–7343.
- Ma Z., Leung Y., Zanon S. (2016) Integration of artificial intelligence and production data analysis for shale heterogeneity characterization in SAGD reservoirs, *SPE Canada Heavy Oil Technical Conference*, Calgary, Canada.
- Masih S., Ma K., Sanchez J., Patino F., Boida L. (2012) The effect of bottom water coning and its monitoring for optimization in SAGD, *SPE Heavy Oil Conference*, 12–4 June, Calgary, Alberta, Canada.
- MATLAB 2009. Version 7.9.0 (R2009b), Natick, Massachusetts, The MathWorks Inc.
- McCoy D.D., Grieves W.A. (1997) Use of resistivity logs to calculate water saturation at Prudhoe bay, *SPE Reserv. Eng.* **12**, 1, 45–51.
- Palacky G.J. (1987) Clay mapping using electromagnetic methods, *First Break* **5**, 8, 295–306.
- Pooladi-Darvish M., Mattar L. (2002) SAGD operations in the presence of overlying gas cap and water layer – effect of shale layers, *J. Can. Pet. Technol.* **41**, 06, 40–51.
- Popa S., Patel A. (2012) Neural networks for production curve pattern recognition applied to cyclic steam optimization in diatomite reservoirs, *SPE Western Regional Meeting*, March 21–23, Bakersfield, CA, USA.
- Popa A.S., Cassidy S.D., Mercer M. (2011) A Data Mining Approach to Unlock Potential from an Old Heavy Oil Field, *SPE Western North American Regional Meeting*, May 7–11, Anchorage, Alaska, USA.
- Queipo N.V., Goicochea J.V., Pintos S. (2002) Surrogate modeling-based optimization of SAGD process, *J. Pet. Sci. Eng.* **35**, 1–2, 83–93.
- Ricardo M. (2013) Simulation sensitivity study and design parameters optimization of SAGD process, *SPE Heavy Oil Conference*, June 11–13, Calgary, Alberta, Canada.
- Richardson J.G., Harris D.G., Rossen R.H., VanHee G. (1978) The effect of small, discontinuous shales on oil recovery, *J. Pet. Technol.* **30**, 11, 1531–1537.
- Shahab M. (1995) Neural network: What it can do for petroleum engineers, *SPE J.* **47**, 01, 42–42.
- Shahab M. (2000) Virtual-Intelligence applications in petroleum engineering: Part 1-artificial neural networks, *J. Pet. Technol.* **52**, 09, 64–73.
- Smith D.G., Hubbard S.M., Leckie D.A., Fustic M. (2009) Counter point bar deposits: Lithofacies and reservoir significance in the meandering modern peace river and ancient McMurray formation, Alberta, Canada, *Int. Assoc. Sedim.* **56**, 1655–1669.
- Wang C., Leung J. (2015) Characterizing the effects of lean zones and shale distribution in steam-assisted-gravity-drainage recovery performance, *SPE Reserv. Eval. Eng.* **18**, 3, 329–345.
- Webb A.C., Schroder-Adams C.J., Pedersen P.K. (2005). Regional subsurface correlations of Albian sequences north of the Peace river in NE British Columbia: northward extent of Sandstones of the Falher and Notikewin members along the eastern flank of the foredeep, *Bull. Can. Pet. Geol.* **53**, 2, 168–188.
- Weiss W.W., Balch R.S., Stubbs B.A. (2002) How artificial intelligence methods can forecast oil production, *SPE/DOE Improved Oil Recovery Symposium*, April 13–17, Tulsa, Oklahoma, USA.
- Xu J.Z., Chen Z.X., Cao J.L., Li R. (2014a) Numerical study of the effects of lean zones on SAGD performance in periodically heterogeneous media, *SPE Heavy Oil Conference*, June 10–12, Calgary, Alberta, Canada.
- Xu J.Z., Chen Z.X., Yu Y.G., Cao J.L. (2014b) Numerical thermal simulation and optimization of hybrid CSS/SAGD process in long lake with lean zones, *SPE Heavy Oil Conference*, June 10–12, Calgary, Alberta, Canada.
- Yang G., Butler R.M. (1992) Effects of reservoir heterogeneities on heavy oil recovery by steam-assisted gravity drainage, *J. Can. Pet. Technol.* **31**, 08, 37–43.
- Zerafat M.M., Ayatollahi S., Mehranbod N., Barzegari D. (2011) Bayesian network analysis as a tool for efficient EOR screening, *SPE Enhanced Oil Recovery Conference*, July 19–21, Kuala Lumpur, Malaysia.
- Zupan J. (1994) Introduction to artificial neural network (ANN) methods: What they are and how to use them, *Acta Chim. Slov.* **41**, 3, 327–352.

Manuscript submitted in 1 November 2016

Manuscript accepted in 11 December 2017

Published online in April 2018