Expert Systems with Applications 42 (2015) 723–740

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Integrated cluster analysis and artificial neural network modeling for steam-assisted gravity drainage performance prediction in heterogeneous reservoirs

Ehsan Amirian [a], Juliana Y. Leung [a,*], Stefan Zanon [b], Peter Dzurman [b]

[a] School of Mining and Petroleum Engineering, Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Alberta T6G 2W2, Canada
[b] Nexen Energy ULC, 801 – 7th Avenue S.W., Calgary, Alberta T2P 3P7, Canada

ARTICLE INFO

Article history:
Available online 30 August 2014

Keywords:
Thermal recovery
Heavy oil
Data-driven
Artificial neural networks
Principal components
Cluster analysis

ABSTRACT

Evaluation of steam-assisted gravity drainage (SAGD) performance that involves detailed compositional simulations is usually deterministic, cumbersome, expensive (manpower and time consuming), and not quite suitable for practical decision making and forecasting, particularly when dealing with high-dimensional data space consisting of large number of operational and geological parameters. Data-driven modeling techniques, which entail comprehensive data analysis and implementation of machine learning methods for system forecast, provide an attractive alternative.

In this paper, artificial neural network (ANN) is employed to predict SAGD production in heterogeneous reservoirs, an important application that is lacking in existing literature. Numerical flow simulations are performed to construct a training data set consists of various attributes describing characteristics associated with reservoir heterogeneities and other relevant operating parameters. Empirical Arps decline parameters are tested successfully for parameterization of cumulative production profile and considered as outputs of the ANN models. Sensitivity studies on network configurations are also investigated. Principal components analysis (PCA) is performed to reduce the dimensionality of the input vector, improve prediction quality, and limit over-fitting. In a case study, reservoirs with distinct heterogeneity distributions are fed to the model. It is shown that robustness and accuracy of the prediction capability are greatly enhanced when cluster analysis are performed to identify internal data structures and groupings prior to ANN modeling. Both deterministic and fuzzy-based clustering techniques are compared, and separate ANN model is constructed for each cluster. The model is then tested using a validation data set (cases that have not been used during the training stage).

The proposed approach can be integrated directly into most existing reservoir management routines. In addition, incorporating techniques for dimensionality reduction and clustering with ANN demonstrates the viability of this approach for analyzing large field data set. Given that quantitative ranking of operating areas, robust forecasting, and optimization of heavy oil recovery processes are major challenges faced by the industry, the proposed research highlights the significant potential of applying effective data-driven modeling approaches in analyzing other solvent-additive steam injection projects.

1. Introduction

The world's total reserves of heavy oil, tar sands, and bitumen are estimated to be approximately 6 trillion barrels of oil in place (Das 1998), with about 1.8 trillion bbl of bitumen located in Alberta, Canada (Energy Resources Conservation Board, 2012). The main challenge for heavy oil recovery is its high fluid viscosity (e.g., in-situ fluid viscosity of the Athabasca bitumen is over 10^6 cp), rendering conventional non-thermal recovery technology uneconomical. SAGD (Butler & Stephens, 1981) is commonly adopted for commercial production of bitumen in Alberta. Two horizontal wells, which are approximately a few meters apart, are placed near the bottom of the target formation. Pressurized steam is injected into the upper injector well and propagates both vertically and laterally in the formation, reducing the viscosity of the bitumen. The heated crude oil then drains by gravitational force along the steam chamber edge into the lower producer.

* Corresponding author. Tel.: +1 (780) 492 3338.
E-mail addresses: amirian@ualberta.ca (E. Amirian), juliana2@ualberta.ca (J.Y. Leung), stefan.zanon@nexencnoocltd.com (S. Zanon), peter.dzurman@nexencnoocltd.com (P. Dzurman).

http://dx.doi.org/10.1016/j.eswa.2014.08.034
0957-4174/

Nomenclature

Symbols

b	decline exponent
D_i	decline rate, 1/day
d_{ij}	Euclidean distance between ith data and jth cluster
d_k^p	modeled value at time step k
$d_{obs,k}$	observed value at time step k
d_{sh}	distance of shaly layers to the injection well, m
d_{sh_avg}	average distance of shaly layers to the injection well, m
d_{sh_min}	distance of the closest shaly layer to the injection well, m
f	total squared error function
$f(Y)$	activation function
h_{sh}	thickness of shaly layer, m
h_{sh_avg}	average thickness of shaly layers, m
J	main objective function
k	permeability, md
m	fuzziness index
N_k	the total number measurement time steps
N_{sh}	number of shaly layers in each realization
P_{inj}	bottom hole injection pressure, psi
q_i	initial oil rate, m^3/day
Q_p	cumulative oil production, m^3
t	elapsed time from start of production, day
V_{DP}	Dykstra–Parson coefficient
Var(d_{sh})	variance of the distances of shaly layers to the injection well
Var(h_{sh})	variance of the thickness of shaly layers
x_i	signal from input node i
Y_j	weighted sum of input signals
w_0	bias
w_{ij}	weight associated with the connection between nodes i and j

Greek letters

ϕ	porosity, %
ϕ_{avg}	average porosity, %
σ	standard deviation of normal distribution
μ	mean of normal distribution
μ_h	center of cluster h
μ_{ij}	the membership of ith data to jth cluster center
v_j	jth cluster center

Acronyms

ANN	artificial neural network
BPNN	backpropagation neural network
EM	expectation-maximization
EOR	enhanced oil recovery
FCM	fuzzy c-means
NN	neural network
PC	principal component
PCA	principal component analysis
PS	principal score
SAGD	steam assisted gravity drainage
SI	shale indicator
RF	recovery factor, %
RMSE	Root Mean Square Error

Present workflow for evaluation of steam-assisted gravity drainage (SAGD) performance typically entails construction of static 2D or 3D geologic models of reservoir properties, which are subjected to numerical compositional simulations for flow and recovery predictions. This process provides only approximate solutions to recovery responses, as numerous simplifications and assumptions must be invoked. The modeling process itself is also quite labor-intensive and time-consuming, particularly when dealing with high-dimensional data space consisting of large number of operational and geological parameters, rendering it not quite suitable for practical decision-making and forecasting involving large number of realizations of reservoir properties or development scenarios. Data-driven modeling techniques, which entail comprehensive data analysis and implementation of machine learning methods for system forecast, provide an attractive alternative.

Artificial neural network (ANN) is a virtual intelligence method used to identify or approximate a complex non-linear relationship between input and target variables. A data set consists of both input (or predicting) and target variables is used to train the network, and the unknown parameters of the network (weights and biases) are estimated in an inverse problem procedure where the mismatch between the network output and the known values of the target variables is minimized. History and evolution of ANN technique including all associated issues and behaviors are presented in Suh (2012).

A wide variety of neural network applications can be found in petroleum engineering (Mohaghegh, 2002; Saputelli, Malki, Canelon, & Nikolaou, 2002; Stundner, 2001), particularly in the areas of: classification (Stundner 2001), reservoir characterization or property prediction (Aminian, Ameri, Oyerokun, & Thomas, 2003; El-Sebakhy et al., 2012; Irani & Nasimi, 2011; Raeesi, Moradzadeh, Ardejani, & Rahimi, 2012; Tang, Toomey, & Meddaugh, 2011), proxy for recovery performance prediction (Awoleke & Lane, 2011; Lechner & Zangl, 2005), history matching (Ramagulam, Ertekin, & Flemings, 2007), and design or optimization of production operations and well trajectory (Artun, Ertekin, Watson, & Miller, 2012; Luis, Ayala, & Ertekin, 2007; Malallah & Sami Nashawi, 2005; Oberwinkler, Ruthammer, Zangl, & Economides, 2004; Stoisits, 1999; Yeten & Durlofsky, 2003; Zangl, Graf, & Al-Kinani, 2006). In particular, neural networks have been utilized in recent years to perform EOR (enhanced oil recovery) screening (Karambeigi, Zabihi, & Hekmat, 2011; Parada & Ertekin, 2012; Zerafat, Ayatollahi, Mehranbod, & Barzegari, 2011); to characterize reservoir properties in unconventional plays (Holdaway, 2012); and to evaluate performance of CO_2 sequestration process (Mohammadpoor, Qazvini Firouz, & Torabi, 2012).

Neural network has also been employed in the subject of heavy oil recovery; for example, it has been used as a proxy model to forecast SAGD performance from operational parameters (Queipo, Goicochea, & Pintos, 2002) and to analyze production characteristics of cyclic steam injection process (Popa, Cassidy, & Mercer, 2011; Popa & Patel, 2012) in homogeneous reservoirs; however, its application in analysis of heterogeneous SAGD reservoirs is lacking. In a conventional SAGD analysis workflow, petrophysical measurements and well logs are used to extract various predicting variables descriptive of the underlying heterogeneity; log traces, which commonly encompass measurements at approximately 0.1 m-resolution over a few hundred meters of formation thickness, can be further considered as conditioning data in the construction of geologic models for flow prediction. In data-driven modeling applications, quantifying and parameterizing relevant heterogeneity descriptions as input attributes, however, remain challenging. To that end, Ahmadloo, Asghari, and Renouf (2010) applied three