Cognitive Data-Driven Proxy Modeling for Performance Forecasting of Water-flooding Process

Ehsan Amirian[*] and Zhangxing John Chen

University of Calgary, Calgary, Alberta, Canada

[*]**Corresponding author:** Ehsan Amirian, University of Calgary, Calgary, Alberta, Canada, Tel: 5877078489; Fax: 5877078489; E-mail: eamirian@ucalgary.ca

Abstract

Assessment of diverse operational constraints and risk appraisal associated with reservoir heterogeneities are essential foundation of production optimization and oil field development scenarios. Water-flooding performance evaluation that comprises comprehensive numerical simulations is typically cumbersome in terms of time and money, which is not reasonably appropriate for practical decision making and future performance forecasting. Cognitive data-driven proxy modeling practices, which incorporate data-mining techniques and machine learning concepts, offer a fascinating substitute for explicit models of the underlying process that can be instantaneously reassessed, especially for extremely nonlinear system forecasts. In this paper, an exploratory data analysis is applied to create a comprehensive data set from Water-flooding actual field data, which entails different characteristics labeling reservoir heterogeneities and other pertinent operational constraints. Artificial neural network (ANN) is applied as a cognitive data-driven proxy modeling effort to predict Water-flooding production in heterogeneous reservoirs. This study presents the great potential of cognitive data-driven proxy modeling techniques for practical applications and as a feasible add on for investigating a huge quantity of real field data efficiently. In addition, the suggested methodology can be incorporated directly into most present reservoir development decision making routines.

Keywords: Data-driven; Cognitive; Artificial neural network; Water-flooding

Nomenclature

Symbols

$f(Y)$: Activation function; k: Permeability, mdk_{avg}: Average permeability; mdS_{ro}: Residual oil saturation; S_{rw}: Residual water saturation; V_{DP}: Dykstra-Parson coefficient; x_i: Signal from input node; iY_j: Weighted sum of input signals; $w0$: bias; w_{ij}: Weight associated with the connection between nodes i and j.

Greek letters

φ = Porosity, (%); φ_{avg} = Average porosity (%).

Acronyms

ANN=Artificial Neural Network; BPNN=Back-Propagation Neural Network; MLP=Multilayer Perceptron; NN=Neural Network; SAGD=Steam Assisted Gravity Drainage; SL= Single Layer; RF=Recovery Factor (%).

Introduction

Water flooding or water injection is one of the most vital techniques for enhancing oil recovery. Injected water into a reservoir aims to provide pressure support to the reservoir which is called voidage replacement, and to displace or drive oil from the reservoir to production wells. Water flooding performance assessment has been widely investigated in both experimental [1-5] and detailed numerical simulation [6-10] frameworks. Numerical modeling of water flooding recovery performance can be carried out with traditional simulators.

The current flow simulators require a huge number of input parameters such as initial saturation and pressure distributions, porosity, permeability, multi-phase flow functions, and well parameters. Inference of these input parameters is time-consuming, while accurate measurements are often not readily available. Furthermore, many assumptions associated with the process physics are often invoked for a numerical solution. Given the extremely nonlinear relationships between input variables and output objective functions (e.g., an oil production profile), the computational time is also extremely high.

Therefore, there has been an increased drive and interest to integrate data-driven proxy approaches for modeling the recovery response of Enhanced Oil Recovery (EOR) processes. A data-driven proxy modeling approach provides a practical alternative for assessment of diverse operational constraints and risk appraisal associated with reservoir heterogeneities, which are essential foundation of production optimization and oil field development plans. High-dimensional data including enormous geological and operational parameters can be processed for efficient and fast decision-making.

Cognitive data-driven proxy modeling employs machine-learning techniques to construct induced models that explain the behavior of the underlying physical process. Detailed data analysis characterizing the desired process is the foundation of this data-driven modeling approach. Common techniques utilized in cognitive data-driven proxy modeling include data-mining and statistical methods, artificial and computational intelligence tools, and fuzzy logic concepts. Currently, these approaches have progressed outstandingly beyond the ones

applied in traditional empirical regression. These techniques are employed for big data classification, predictive modeling, recreating extremely nonlinear relationships, and constructing rule-based expert systems.

The common themes of data-driven modeling have been developed with integration and contribution from various interdisciplinary specializations involving artificial and computational intelligence, machine learning and pattern recognition, data-mining alongside with the statistical data analysis, knowledge discovery in databases, and soft computing. A data-driven modeling layout is based on the assumption that a primary process has generated a database of observed cases, expert experience and knowledge. As demonstrated in Figure 1, the ultimate goal of cognitive data-driven proxy modeling is to fuse these multiple information sources to present a representative model for the primary process. If the presented model approximation is acceptable for that process, it can be employed to address the other questions regarding the properties of the underlying process [11].
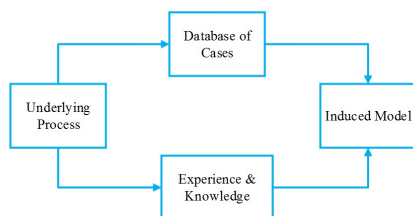

Figure 1: Illustration of data-driven modeling approach [11].

In this study, an artificial intelligence (AI) technique, called artificial neural network (ANN), is used to identify or approximate a complex nonlinear relationship connecting pertinent input variables to the desired target objective functions. Artificial neural network (ANN) employs a series of processing units (neurons, nodes) in the hidden layers where the weighted summation of input variables is subjected to a nonlinear transfer function. A data set consisting of both input (pertinent predicting) variables and desired target objective functions is employed to train the network.

The network unknown parameters, classically the connecting weights and biases (thresholds) that link between all the nodes, are assessed using the procedure of inverse problem theory in which the objective is to minimize the existing mismatch between the artificial neural network predicted output and the known actual values of the desired objective functions variables.

In Figure 2, a typical flow chart for artificial neural network training is illustrated. Since various parameters can be both continuous and discrete (categorical), the most universal uses of ANN include cognitive proxy modeling for function estimation and pattern classification. Brief history and advancements for the artificial neural network technique including ANN learning rules, architecture configuration, hybrid practices and convergence situations are described in [12-14].
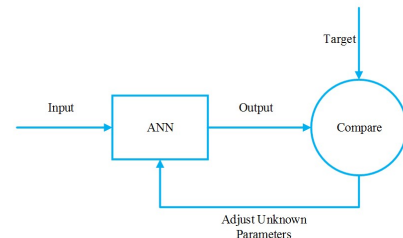

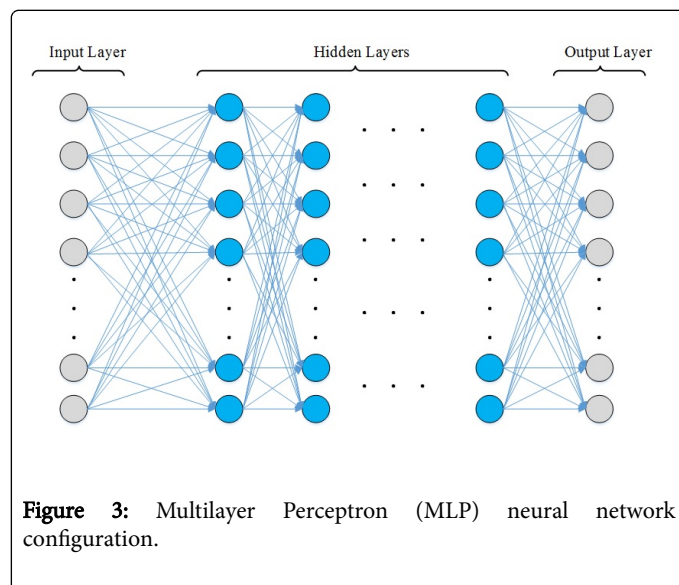**Figure 2:** General flow chart for artificial neural network training.

ANN is developed by training a network to represent the intrinsic relationships existing within the data. The idea of neural network alludes back to 1943 when neurophysiologist Warren McCulloch and mathematician Walter Pitts published their research work on how neurons might perform [15]. Any neural network is trained using a learning algorithm and training data set. In general, there are two types of neural network learning algorithms classification: unsupervised learning and supervised learning. The unsupervised learning is used to find hidden structure in unlabeled data. The objective is to categorize or discover features or regularities in the training data. A cluster analysis is the most common use of unsupervised learning. In contrast, the supervised learning method requires that target values be provided. A training dataset is needed as the input vector and will generate the rules according to the desired output by adjusting the weights. The weights are then used for processing the inputs of a test data set. After providing the desired output to the net, the weights will be adjusted to match the model to the desired goal. The learning process iteration will be continued until the desired goal is reached.

In petroleum engineering an extensive variety of neural network applications can be found [16-19], particularly in the areas of: reservoir characterization or property prediction [20-23], classification [19], proxy for recovery performance prediction [24,25], history matching [26], and design or optimization of production operations and well trajectory [27-33]. In particular, neural networks have been utilized in recent years as a proxy model to predict heavy oil recoveries [34-39], to perform EOR (enhanced oil recovery) screening [40-42]to characterize reservoir properties in unconventional plays [43], and to evaluate performance of a CO_2 sequestration process [44]. As a data-driven proxy modeling workflow numerical reservoir simulation models are subjected to the commercial simulator to build the comprehensive training data set. This representative data set consist of pertinent input variables labelling reservoir uncertainty due to the geological heterogeneities and crucial injection/production constraints with the corresponding desired objective function output values including cumulative oil production profile and ultimate recovery factor for the underlying recovery process. Artificial neural network models are trained employing the representative training data set and eventually applied as a data-driven proxy modeling alternate to forecast the desired objective function target values including cumulative oil production profile and ultimate recovery factor during the SAGD process [38,39].

The principal objective of this research is to develop a data-driven proxy model using artificial neural network as a substitute to smartly forecast Water-flooding recovery performance in heterogeneous reservoirs. This research aims to assess the key demonstrative predicting (input) variables related to Water-flooding performance forecasting with practical application to the heterogeneous reservoirs.
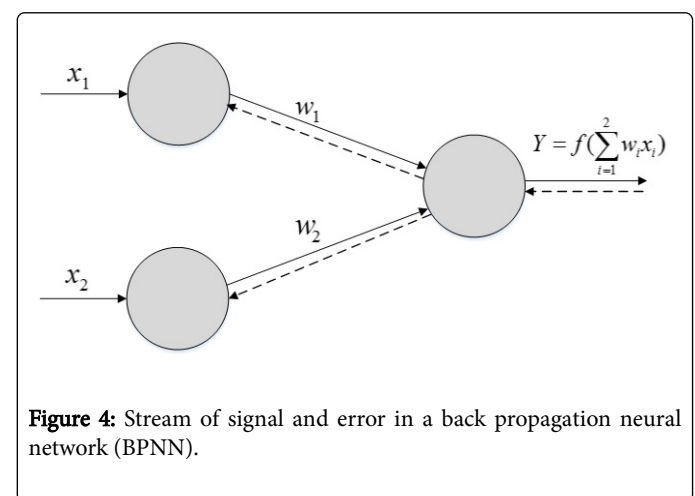
Methodology

Artificial neurons that are linked together according to specific network architecture create an artificial neural network. The two most common categories of network architectures are Single Layer Perceptron (SLP) and Multilayer Perceptron (MLP) which consists of single input and output layer with any number of hidden layers. Since the complexity, heterogeneity and nonlinearity are the main challenges when facing a comprehensive real field data set, multilayer perceptron stands as the most common neural network architecture. A general schematic of MLP neural network is shown in Figure 3. Selection of the number of hidden layers and the number of hidden processing units (neurons, nodes) within the hidden layers are of great importance during an artificial neural network architecture design. There is a transaction between accuracy and overfitting of data: A difference between artificial neural network forecasts and desired target values of output objective functions could not be minimized with an insufficient number of neurons, whereas too many neurons can result in an overfitting of network parameters (i.e., weights and bias connections).

Figure 3: Multilayer Perceptron (MLP) neural network configuration.

Researchers have established some rules of thumb to opt the number of neurons in a neural network. The number of independent (input) variables is generally much larger than the number of dependent (target) variables. Ferreira [45] proposed that the number of neurons should lie between the number of input parameters and the number of output parameters; in particular, the number of neurons should be two thirds of the number of input parameters, plus the number of output parameters, but no more than twice the number of input parameters. Haykin [46] explained that the number of free parameters (i.e., the number zf weights and bias connections) in the hidden layer should be a function of the input vector dimension and the total training data set size. In this paper, a sensitivity analysis on the network configuration is implemented; the optimal architecture is selected by comparing the error/mismatch in network prediction between different configurations.

In a feed-forward neural network, a signal is passed from an input layer of neurons (nodes) through a series of hidden layers to an output layer of neurons. The input nodes represent the independent variables that are nonlinearly related to a set of dependent or target variables, characterized by a series of neurons at the output layer. Free parameters including weights and biases specified for each connection are determined using a training data set via a supervised learning process in which a gradient-based minimization technique is utilized to minimize the mismatch between artificial neural network predictions and actual values of the desired target objective function values [47]. Figure 4 illustrates exactly how the signal flows forward (feed-forward) from the input to the output layer and also who the error is back-propagated from the output to the input layer.

Figure 4: Stream of signal and error in a back propagation neural network (BPNN).

The sum of the multiplication of the input values is added to the biases of each processing unit to get the value of Y as in Equation 1.

$$Y_j = w_0 + \sum_{i=1}^{n} w_{ij} x_i \tag{1}$$

where Yj is a weighted sum of input signals at node j; w0 is a bias value; wij is the weight linked with the connection between node j and the input node i; xi is the value of input node i; n is the number of input nodes. As shown in Equation 2, a sigmoid activation function is applied to the weighted Yj.

$$f(Y) = \frac{1}{1 + e^{-Y}} \tag{2}$$

The value calculated from Equation 2 is the output signal from node j, which can be considered as the input signal to the next layer. In this paper, a Feedforward Backpropagation Neural Network or Backpropagation Neural Network (BPNN) model as the most common algorithm for estimating the unknown network parameters (weights and biases) is employed. As a gradient-based minimization technique that utilizes a supervised learning process with feed-forward network architecture, BPNN propagates an error backwards from the output nodes to the input nodes which is shown in Figure 4. The algorithm evaluates the gradient of the error associated with the network's unknown parameters. A gradient-descent algorithm is then applied to estimate parameters that minimize the error [48].

Data assembly

Core data is frequently quoted as the "ground-truth" in petrophysical assessment of reservoir rock [49]. Cores can be physically studied and measurements made can be interpreted in terms of representative lithological characteristics. It should be noted that the cores themselves must be representative of the evaluated reservoir section. A core analysis is employed to define not only the porosity and permeability of the reservoir rock, but also to exhume the fluid saturation and grain density. All of these measurements help geologists and engineers better understand the conditions of a well and its potential productivity.

Core data based on core analyses for 700 cored wells are collected from a public domain in an oilfield located in the Alberta province, where Water-flooding has been performed during years. Given that these data contain noises and considering the missing data attributes in some locations, 61 cored wells are used as the representative data set for this data-driven modeling. A number of input/output attributes describing the reservoir properties and production characteristics including permeability, porosity, residual oil and water saturations and cumulative oil production have been analyzed for this data set. Although other information such as a log analysis and seismic data might be available for a portion of the data set, we are interested in building records that have the reliable measurements from core data.

Case study

Heterogeneity in hydrocarbon reservoirs creates a great amount of risk during recovery processes. In this case study, ANN is employed to forecast Water-flooding oil recovery performance for a series of producing wells with varying porosity and permeability values.

An ANN model is constructed involving one output/target variable cumulative oil production after five years for each well and a total of five input variables that include the mean of porosity values in each well: ϕavg, the mean of permeability values in each well: kavg, the Dykstra-Parsons coefficient: VDP, residual oil saturation: Sro and residual water saturation: Srw. The Dykstra-Parsons coefficient [50] has been employed in much research to characterize heterogeneous permeability distribution in layered reservoirs.
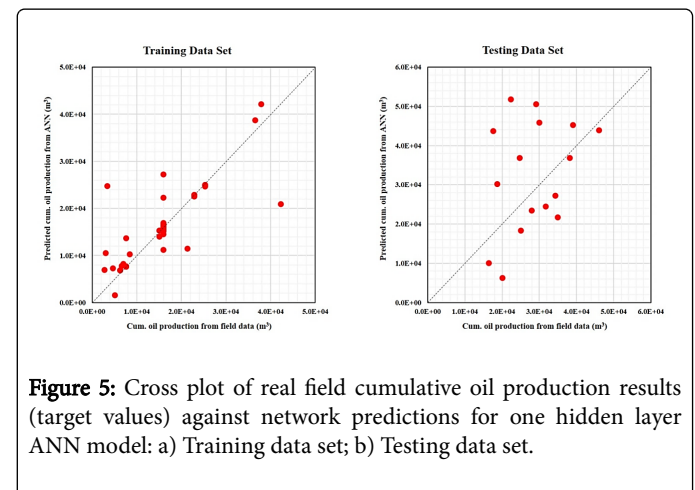
According to the immense discrepancy in scales of different databases, normalization is frequently implemented [47]. As an essential preliminary processing (pre-processing) stage for data driven proxy modeling, normalization is applied to transform all data values to shift and lie in a certain positive range, for instance (0, 1), shown in Equation 3. This step boosts the equality of the training stage by retaining inputs with large values from replacing out other inputs that are equally important with smaller values [51].

$$x_{\text{new}} = \frac{x_{\text{old}} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \quad (3)$$
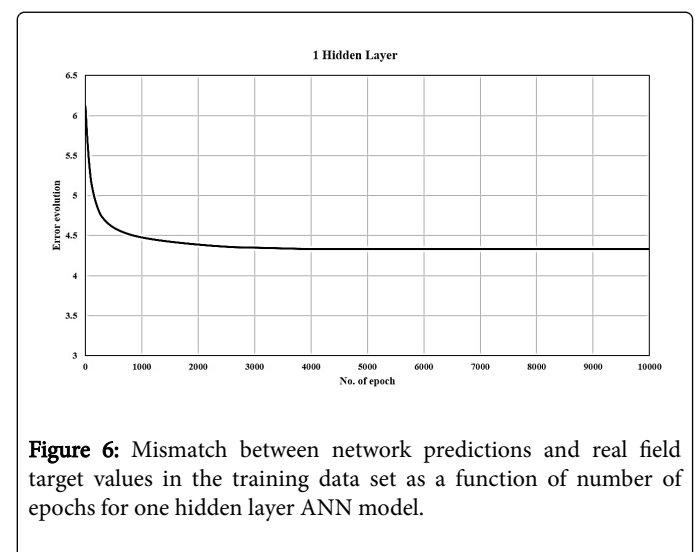
All 61 well records are subjected to a pre-processing stage and afterwards the ANN training and testing phases. In this study, 75 percent of the recorded data points are used as the training data set and the rest 25 percent are utilized as the testing data set. A sensitivity analysis of the network configuration is performed where the mismatch between network predictions and actual values of target variables after a fixed number of epochs is compared among different configurations.

Results and Discussion

The ANN model configuration is set to have one hidden layer with six hidden nodes. The predicted results for this configuration are illustrated in Figure 5. As one can see, the training stage performance after a constant number of epochs is unsatisfactory and the ANN predicted results for cumulative oil production are different from the ones from real field data (Figure 5a). This insufficiency is revealed in the testing stage where the trained network parameters are employed to predict the cumulative oil production for the testing data set after five years of production, as shown in Figure 5b.

Figure 5: Cross plot of real field cumulative oil production results (target values) against network predictions for one hidden layer ANN model: a) Training data set; b) Testing data set.
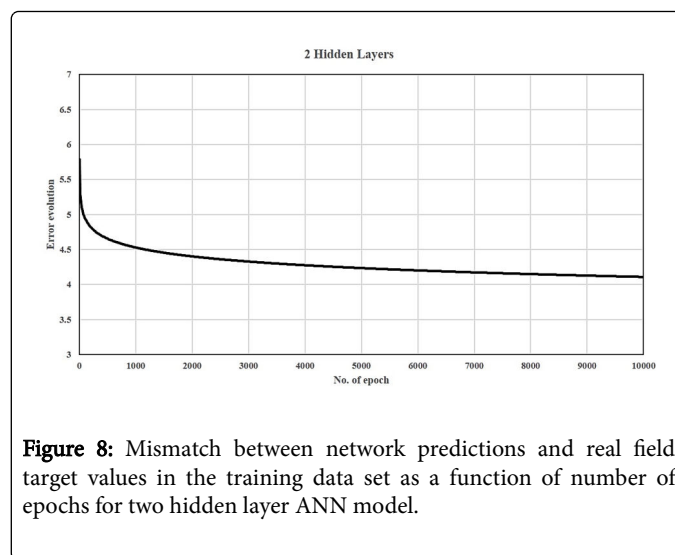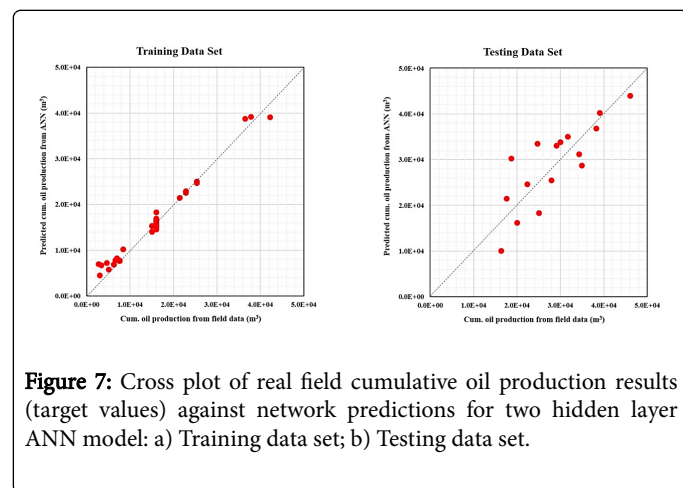
The error evolution in the training stages is plotted per number of epochs for this ANN model configuration (Figure 6). The error has been stabilized after epoch no. 4,000 and there is no decrease in the mismatch between ANN prediction and real field target values. Increasing the number of hidden nodes made no improvements in the model predictability. This fact lead us to increase the number of hidden layers to capture the nonlinearity existing among the data-driven proxy model input variables and desired output parameters.

Figure 6: Mismatch between network predictions and real field target values in the training data set as a function of number of epochs for one hidden layer ANN model.
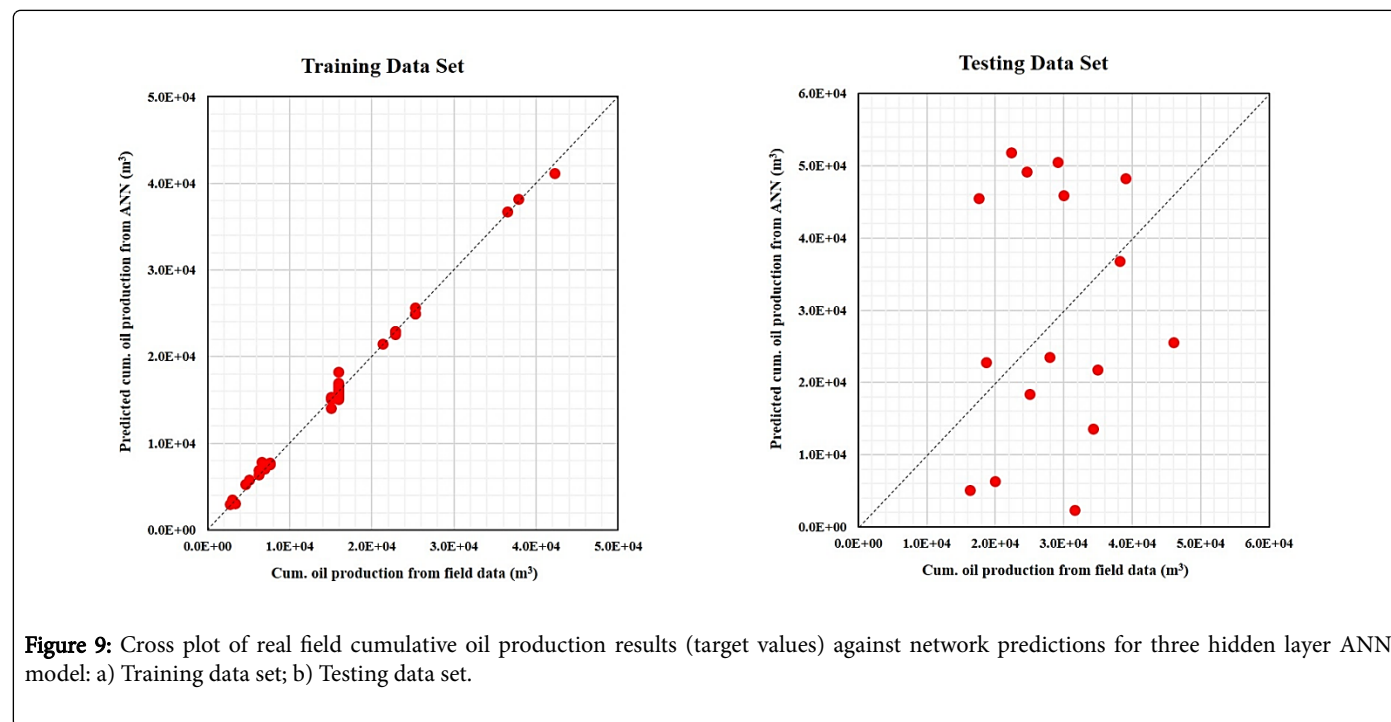
Next, two hidden layers, each including six nodes, are employed to train the proxy model and estimate the network unknown parameters. Cross-plots of the real field cumulative oil production values against proxy forecasts for training and testing data segments are

demonstrated in Figure 7. A decent match between the objective function desired actual values and data-driven proxy model predictions can be observed, as most points lie very close to the 45°line. Figure 8 presents a decline of the error/mismatch as a function of the number of epochs for this network configuration; a significant error reduction is achieved as the training progresses. The decrease in error evolution after reaching the end of an epoch limit is comparable to the one from one hidden layer ANN model. The results suggest that the chosen network configuration can be used successfully to predict recovery performance.

Figure 7: Cross plot of real field cumulative oil production results (target values) against network predictions for two hidden layer ANN model: a) Training data set; b) Testing data set.

Figure 8: Mismatch between network predictions and real field target values in the training data set as a function of number of epochs for two hidden layer ANN model.

Performance of a three hidden layer ANN model is also investigated for this case study with the same six nodes in each hidden layer. ANN performance during the training stage is fundamentally flawless as excellent matches between targets and predicted output values demonstrated in Figure 9a. The error evolution plot also certifies a good training stage where the error has been significantly decreased as shown in Figure 10. Nevertheless, once the trained data-driven proxy model is applied to the testing data set, evaluation of proxy model forecasts with the desired target objective function actual values revealed to be awfully ill-conditioned. As seen in Figure 9b, the cross-plot between ANN predicted values and real field data for cumulative oil production after five years are way far from the 45° line which indicates a drop in network predictability.
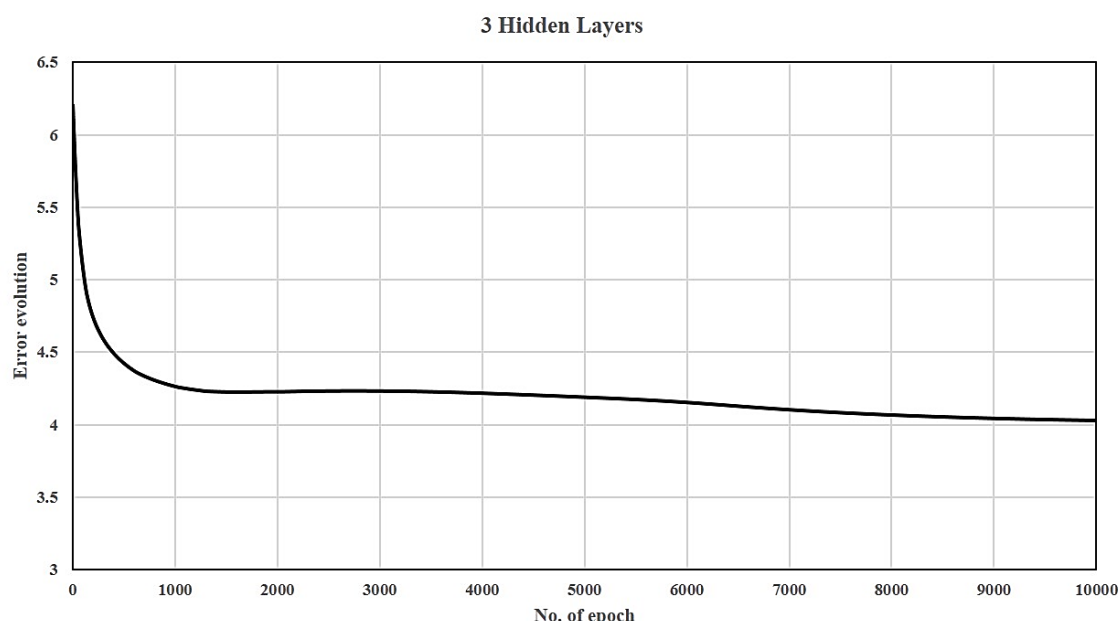

Figure 9: Cross plot of real field cumulative oil production results (target values) against network predictions for three hidden layer ANN model: a) Training data set; b) Testing data set.

Figure 10: Mismatch between network predictions and real field target values in the training data set as a function of number of epochs for three hidden layer ANN model

The other issue rather than over-fitting due to the network configuration can be the existence of hidden structures within the data set. In this case, a cluster analysis and principal components analysis can be applied in upcoming research works to detect the internal hidden patterns within the input attributes of the underlying process and make recommendations for dimensionality reduction to have independent input parameters for cognitive data-driven proxy modeling [39].

Comparison between all three network configurations implies that too many hidden layers will increase the chance of an over-fitting issue, whereas the forecasting performance of a cognitive data-driven proxy model is conceded with inadequate number of hidden processing nodes. In this study the optimum network configuration has been shown to be a two hidden layer ANN model where the network prediction is in a good match with the target values from real field data.

Conclusion

A core data analysis is employed to identify numerous parameters describing characteristics associated with reservoir heterogeneities. A comprehensive Water-flooding real field data set from a public domain is compiled. The constructed data set involves one output/target variable cumulative oil production after five years for each well and a total of five input variables that include the mean of porosity and permeability, the Dykstra-Parsons coefficient, residual oil saturation and residual water saturation.

Data-driven modeling processes such as artificial neural networks are still considered recent advancements and have not been widely adopted in most sectors of the oil sands industry. In this research,

ANN is successfully applied to predict recovery performance of the Water-flooding process.

Performance of different ANN model configurations is evaluated and the optimal network architecture is tested successfully to forecast the Water-flooding performance for the testing data set.

The proposed research has great potential to be integrated directly into most existing reservoir management and decision-making routines and applied in SAGD and other solvent-additive steam injection projects.

Future studies will incorporate data from other producing fields and integrate other data mining techniques, such as clustering to recognize internal hidden structures among a data set and a principal component analysis as a treatment to the curse of dimensionality to reduce the number of independent input parameters.

Acknowledgement

References

1. Noirot JC, Van Den Hoek PJ, Zwarts D, Bjoerndal HP, Stewart G, et al. (2003) Water injection and water flooding under fracturing conditions. Middle East Oil Show. Bahrain: Society of Petroleum Engineers.

2. Aslanyan A, Aslanyan I, Demski C, Al Touqi M (2007) Evaluation of water flooding efficiency through the static high precision temperature logging. SPWLA Middle East Regional Symposium. Abu Dhabi, UAE: Society of Petrophysicists and Well-Log Analysts.

3. Cense A, Berg S, Bakkar K, Jansen E (2011) Direct visualization of designer water flooding in model experiments. SPE Enhanced Oil

Recovery Conference. Kuala Lumpur, Malaysia: Society of Petroleum Engineers.

4. Dong Y, Dindoruk B, Ishizawa C, Lewis EJ, Kubicek T (2011) An experimental investigation of carbonated water flooding. Denver: SPE, SPE Annual Technical Conference and Exhibition.

5. Al-Attar HH, Mahmoud MY, Zekri AY, Channam MT (2013) Low salinity flooding in a selected carbonate reservoir: Experimental approach. EAGE Annual Conference & Exhibition incorporating SPE Europec. London, UK: Society of Petroleum Engineers.

6. Zhu Z, Gerritsen M, Thiele M (2010) Thermal streamline simulation for hot water-flooding. SPE Reservoir Evaluation & Engineering 13: 372-382.

7. Pajonk O, Schulze-Riegert R, Krosche M, Hassan M, Nwakile MM (2011) Ensemble-based water flooding optimization applied to mature fields. SPE Middle East Oil and Gas Show and Conference. Manama, Bahrain: Society of Petroleum Engineers.

8. Lavie G, Kamp AM (2011) An Evaluation of criteria for dynamic gridding during reservoir simulation of water flooding and solvent injection. SPE Reservoir Simulation Symposium. The Woodlands, Texas, USA: Society of Petroleum Engineers.

9. Fjelde I, Asen SM, Omekeh VA (2012) Low salinity water flooding experiments and interpretation by simulations. SPE Improved Oil Recovery Symposium. Tulsa, Oklahoma, USA: Tulsa, Oklahoma, USA.

10. Alzayer H, Sohrabi M (2013) Numerical simulation of improved heavy oil recovery by low-salinity water injection and polymer flooding. SPE Saudi Arabia Section Technical Symposium and Exhibition. Al-Khobar, Saudi Arabia: Society of Petroleum Engineers.

11. Kjaerulff UB, Madsen AL (2008) Bayesian networks and influence diagrams. Springer Science+ Business Media 200, 114.

12. Mehrotra K, Mohan CK, Ranka S (1997) Elements of artificial neural networks. Cambridge, MA: A Bradford Book, MIT Press.

13. Suh SC (2012) Practical applications of data mining. Jones & Bartlett Learning,.

14. Poulton MM (2001) Computational neural networks for geophysical data processing. Elsevier, Oxford.

15. McCulloch WS, Pitts W (1943) A logical calculus of ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5: 115-133.

16. Mohaghegh S (2000) Virtual-intelligence applications in petroleum engineering: Part 3—Fuzzy logic. Society of Petroleum Engineers (SPE).

17. Bravo C, Saputelli L, Rivas F, Pérez AG, Nikolaou M, et al. (2012) State-of-the-art application of artificial intelligence and trends in the E&P industry: A technology survey. Utrecht, The Netherlands: SPE Intelligent Energy International.

18. Saputelli L, Malki H, Canelon J, Nikolaou M (2002) A critical overview of artificial neural network applications in the context of continuous oil field optimization. San Antonio, TX. USA: SPE Annual Technical Conference and Exhibition.

19. Stundner M (2001) How data-driven modeling methods like neural networks can help to integrate different types of data into reservoir management. SPE Middle East Oil Show., Bahrain.

20. Tang H, Toomey N, Meddaugh WS (2011) Using an artificial-neural-network method to predict carbonate well log facies successfully. SPE Reservoir Evaluation Engineering 14: 35-44.

21. Raeesi MM, Moradzadeh A, Ardejani FD, Rahimi M (2012) Classification and identification of hydrocarbon reservoir lithofacies and their heterogeneity using seismic attributes, logs data and artificial neural networks. J Pet Sci Eng 82: 151-165.

22. Aminian K, Ameri S, Oyerokun A, Thomas B (2003) Prediction of flow units and permeability using artificial neural networks. Long Beach, CA, USA.: SPE Western Regional/AAPG Pacific Section Joint Meeting.

23. Hemmati-Sarapardeh A, Shokrollahi A, Tatar A, Gharagheizi F, Mohammadi AH, et al. (2014) Reservoir oil viscosity determination using a rigorous approach. Fuel 116: 39-48.

24. Awoleke OO, Lane RH (2011) Analysis of data from the barnett shale using conventional statistical and virtual intelligence techniques 14: 544-556.

25. Lechner JP, Zangl G (2005) Treating uncertainties in reservoir performance prediction with neural networks. Madrid, Spain: SPE Europec/EAGE Annual Conference.

26. Ramagulam A, Ertekin T, Flemings PB (2007) Utilization of artificial neural networks in the optimization of history matching. SPE. Buenos Aires, Argentina.

27. Stoisits RF, Crawford KD, MacAllister DJ, McCormack MD, Lawal AS, et al. (1999) Production optimization at the Kuparuk River field utilizing neural networks and genetic algorithms. Oklahoma City, OK, USA: SPE Mid-Continent Operations Symposium.

28. Luis F, Ayala H, Ertekin T (2007) Neuro-simulation analysis of pressure maintenance operations in gas condensate reservoirs. J Pet Sci Eng 58: 207-226.

29. Artun E, Ertekin T, Watson R, Miller B (2012) Designing cyclic pressure pulsing in naturally fractured reservoirs using an inverse looking recurrent neural network. Computational Geosciences 38: 68-79.

30. Yeten B, Durlofsky LJ (2003) Optimization of nonconventional well type, location, and trajectory. SPE Journal 8: 200-210.

31. Oberwinkler C, Ruthammer G, Zangl G, Economides MJ (2004) New tools for fracture design optimization. Lafayette, LA, USA: SPE International Symposium and Exhibition on Formation Damage control.

32. Malallah A, Sami Nashawi I (2005) Estimating the fracture gradient coefficient using neural networks for a field in the Middle East. J Pet Sci Eng 49: 193-211.

33. Zangl G, Graf T, Al-Kinani A (2006) Proxy modeling in production optimization. Vienna, Austria: SPE Europec/EAGE Annual Conference and Exhibition.

34. Queipo, NV, Goicochea JV, Pintos S (2002) Surrogate modeling-based optimization of SAGD Processes. J Pet Sci Eng 35: 83-93.

35. Popa A, Cassidy S, Mercer M (2011) A data mining approach to unlock potential from an old heavy oil field. Anchorage, Alaska, USA: SPE Western North American Regional Meeting.

36. Popa AS, Patel A (2012) Neural networks for production curve pattern recognition applied to cyclic steam optimization in diatomite reservoirs. Bakersfield, CA, USA: SPE Western Regional Meeting.

37. Ahmadloo F, Asghari K, Renouf G (2010) A new diagnostic tool for performance evaluation of heavy oil waterfloods: case study of western Canadian heavy oil reservoirs. Anaheim, CA, USA: SPE Western Regional Meeting.

38. Amirian E, Leung JY, Zanon S, Dzurman P (2013) Data-driven modeling approach for recovery performance prediction in SAGD operations. Calgary, Canada: SPE Heavy Oil Conference - Calgary, Alberta, Canada.

39. Amirian E, Leung JY, Zanon S, Dzurman P (2015) Integrated cluster analysis and artificial neural network modeling for steam-Assisted gravity drainage performance prediction in heterogeneous reservoirs. Expert Systems with Applications 42: 723-740.

40. Zerafat MM, Ayatollahi S, Mehranbod N, Barzegari D (2011) Bayesian network analysis as a tool for efficient EOR screening. Kuala Lumpur, Malaysia: SPE Enhanced Oil Recovery Conference.

41. Karambeigi MS, Zabihi R, Hekmat Z (2011) Neuro-simulation modeling of chemical flooding. J Pet Sci Eng 78: 208-219.

42. Parada CH, Ertekin T (2012) A new screening tool for improved oil recovery methods using artificial neural networks. Bakersfield, CA, USA: SPE Western Regional Meeting.

43. Holdaway KR (2012) Oilfield data mining workflows for robust reservoir characterization. Utrecht, The Netherlands: SPE Intelligent Energy International.

44. Mohammadpoor M, Qazvini-Firouz A, Torabi F (2012) Implementing simulation and artificial intelligence tools to optimize performance of the CO_2 sequestration in coalbed methane reservoirs. Orlando, FL, USA: Carbon management technology conference.

45. Ferreira I, Gammiero A, Llamedo M (2012). Design of a neural network model for predicting well performance after water shutoff treatments using polymer gels. Mexico City, Mexico: SPE.

46. Haykin S (2005) Neural networks, A Comprehensive Foundation (2nd edn) Prentice-Hall of India, Delhi, India.

47. Francis L (2001) The basics of neural networks demystified. Contingencies 11: 56-61.

48. Werbos PJ (1994) The roots of back-propagation. From ordered derivatives to neural networks and political forecasting. New York, NY: John Wiley & Sons Inc.

49. Skopec RA, Collee PE, Van Puymbroeck L, Shallenberger LK (1995) Downhole core preservation for improved formation evaluation. Annual Logging Symposium, 36th. Paris, France: Society of Professional Well Log Analysts.

50. Dykstra H, Parsons RL (1950) The prediction of oil recovery by water-flooding, Secondary Recovery of Oil in the United States. New York: API.

51. Al-Fattah SM, Startzman RA (2001) Predicting natural gas production using artificial neural network. SPE Hydrocarbon Economics and Evaluation Symposium. Dallas, Texas, USA: Society of Petroleum Engineers.