Society of Petroleum Engineers

SPE-181578-MS

Reservoir Characterization Using Fuzzy Kriging and Deep Learning Neural Networks

M. M. Korjani, University of Southern California; A. S. Popa, E. Grijalva, and S. Cassidy, Chevron Corporation; I. Ershaghi, University of Southern California

This paper was prepared for presentation at the SPE Annual Technical Conference and Exhibition held in Dubai, UAE, 26-28 September 2016.

Abstract

The sandy-clayed heavy oil hydrocarbon reservoirs of the Miocene formations in San Joaquin Valley, California, present relatively high complex geologic settings. These reservoirs consist in general of sandstone, silts and shale bodies, and present anywhere from six to eight degrees of inclination. Geostatistical tools were used to characterize the reservoir and develop a full earth model for the field. The data used was mainly based on the suite of resistivity logs (shallow, medium and deep), which seems to best represent the formations, while a classic kriging technique was used for modeling.

In the present research, we introduce a new approach for characterization of these heavy-oil reservoirs by using fuzzy kriging and deep learning neural networks. The new model better captures the uncertainty associated with the characteristics of the reservoirs. Once successfully trained, the system was applied to generate a 3D earth model and estimate reservoir properties at any point in the field. The new earth model successfully improves the resolution and realization of the subsurface model, as compared to the existing one developed using classic kriging. Integrating fuzzy kriging solves the limitation of smoothness of kriging estimation and reproduction of extreme values. Additional capability includes generation of synthetic logs for wells anywhere in the reservoir, with a significant application for infill drilling.

The methodology presented in this research proved to be very successful as developed for modeling alluvial deposits such as the heavy oil reservoirs in San Joaquin Valley, California. Further research should be considered to address other geologies where laterally formations vary in characteristics or where pinch outs or faulting occur.

Introduction

Kriging is one of the most popular geostatistical models for the spatial inference of quantities in unobserved locations, and to quantify the uncertainty associated with the results. It provides the best linear unbiased estimator for the distance weighted average of a variable. The kriging system depends on only data values and their covariance model; therefore, the kriging weights and variance can be predicted before recording data values. However, because the estimation variance is not a complete measure of local accuracy,

the results show smoothness effect and exact interpretation. This, in turn, generates shortcomings in the development of 3D full field earth models.

To overcome the limitations of the classic kriging, we proposed a new approach where fuzzy kriging parameters are used as inputs, along with other measured reservoir attributes, in deep learning neural networks to improve the reservoir characterization and thus the reservoir model.

Deep learning neural networks are an emerging technology getting significant attention in image recognition, language translation, and signal/voice processing. In contrast to the classic neural networks, deep learning requires large volumes of data, regarding both features and samples, for proper training.

Fuzzy kriging is a relatively new and enriched methodology, which through its mathematical model accounts for the uncertainty existing in the nature of hydrocarbon reservoirs. Fuzzy kriging weights account for spatial correlation between data points leading to an improved deep learning estimation of petrophysical characteristics of reservoirs. Large heavy oil fields with more than 5000 wells present an attractive application of this technology. Training data consisted of a sufficient suite of logs including SP, GR, SRES, MRES, DRES, NPHI, CALIPER, etc., together with added dependencies such as distance, dip, angle of nearest neighbor wells, and fuzzy kriging weights to complete a set of 100+ input attributes.

The model captured the correlation between the well geographic locations using kriging parameters as well as the well logs, and reservoir facies. Once successfully trained, the system was applied to estimate reservoir properties at any point in the formation and to generate synthetic reservoir facies, which lead ultimately to find future drilling locations and production improvement.

A subject field located in San Joaquin Valley, California was selected to pilot the proposed methodology. The reservoir consists of multiple productive sandstone formations separated by silts and shale bodies. The reservoirs present a slight inclination of about six to eight degrees and contain heavy oil. Starting in the early 1960's, steam injection was introduced as an enhanced oil recovery technology. Today, the field continues to sustain a strong production with the introduction of horizontal well drilling and intelligent steam management.

The article is structured into seven sections as follows: Section II covers the objective of the research; Section III provides an overview of kriging and fuzzy kriging; deep learning neural networks are presented in Section IV; Section V covers the methodology and results; a brief discussion and summary is covered in Section VI, and conclusions are presented in section VII.

Objective

An accurate reservoir model leads to the best decision making in the development of hydrocarbon reservoirs. However, the accuracy comes to a certain price mostly driven by the cost of data collection. This data can be seen as drilling additional wells, in the case of exploration, or running the comprehensive suite of logs to measure formation properties for every new infill drill. The latter data collection process can become unfeasible and unsustainable for mature reservoirs with hundreds or thousands of wells.

In the case of the subject field, reservoir modeling was completed using available geophysical log data from the 5,000+ wells drilled in the field. The application of geostatistical tools, classic kriging methodology, Gaussian simulation and exponential model variograms were used to estimate and populate the spatial distribution of petrophysical properties in the grid cells of the 3D full field model. The most important properties in the case of the subject field were the facies representation, effective porosity (PHIE) and permeability (k), which was obtained as a transform function from deep resistivity.

While the model served its purpose well for the last decade, it has its limitations and shortcomings. For example, the interpolation of the facies between wells shows rather rigid edges with segmented continuity. Secondly, the model is not able to capture the sand channels created by the alluvial deposits existing in the reservoir mostly due to classical kriging methodology and internal modeling to account for the reservoir inclination.

The objective of this work was to research the application of new alternatives and technologies that would improve the special distribution of the petrophysical properties and enhance the resolution of the subsurface full field model.

Kriging Technology

Over the years, multiple alternatives of kriging methodology were proposed. A listing of the most popular models includes; ordinary kriging, simple kriging, universal kriging, indicator kriging, probability kriging, disjunctive kriging, co-kriging, and fuzzy kriging.

In this section, a brief describing of simple kriging is provided to establish the foundation of spatial interpolation. Next, we will describe the fuzzy kriging, which was used in the current research.

Simple Kriging

Kriging is one of the most popular methods of spatial interpolation of data [1–3]. It is an estimator that incorporates the spatial relationship between data and the unknown as well as within data points (through variogram). The output of kriging algorithm is an estimated value for a given spatial location from the weighted linear combination of input data. The estimator is

$$\phi^*(u_0) = \lambda_0 + \sum_{i=1}^{n} \lambda_i \phi(u_i) \tag{1}$$

where to estimate the value of #* in point u_0, it requires using linear combination of 'n' inputs $\#(u_1),\ldots,\#(u_n)$ at different locations $u_1,\ldots,u_n$. Linear estimation model from four inputs is presented in Fig. 1.

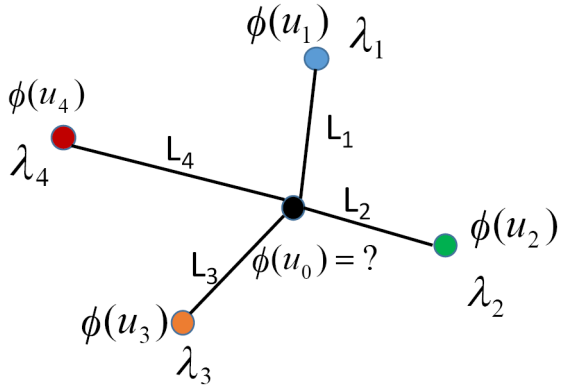


Figure 1—Kriging linear estimation for four input sources

The input weights are determined by minimizing estimation variance [4].

$$\min_{\lambda} Var[\phi^*(u_0) - \phi(u_0)] \tag{2}$$

and the estimation is unbiased

$$E[\phi^*(u_0) - \phi(u_0)] = 0 \tag{3}$$

which leads to non-negative weights $\lambda_i \geq 0$, $i=1,\ldots,n$ and

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{4}$$

The function describing the degree of spatial dependency between input sources is known as variogram. It is defined by the variance of the difference between input values at two locations. Variogram models play a critical role in the calculation of a reliable kriging estimation; however, there are lots of uncertainties in

obtaining a consistent model. Several methods to perform the calculation of variograms exist [5–8], but they still can lead to substantially different fits. The choice of the variogram type and parameters are known to influence the kriging results [9–10].

Dealing with uncertainty in a variogram needs a flexible set of solutions that can properly model different kinds of uncertainty in data. Different sources of uncertainties are: uncertainty in measurement, lack of enough measurements, and lack of knowledge about certain phenomenon. Defining variogram through fuzzy sets it is possible to account for all the uncertainties in a system. Although a variogram can be crisp (Fig. 2), a range of different curves can be fitted to the variogram data points, which can be seen as a fuzzy variogram (Fig. 3).

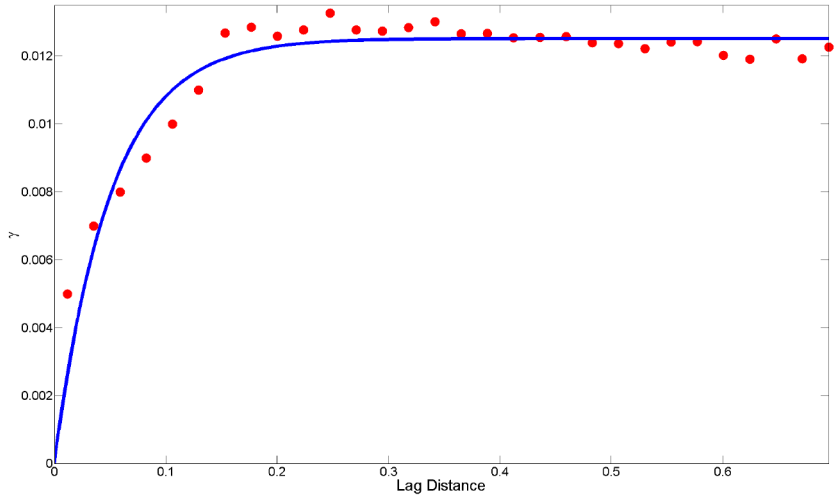


Figure 2—Simple variogram fitted to a set of data

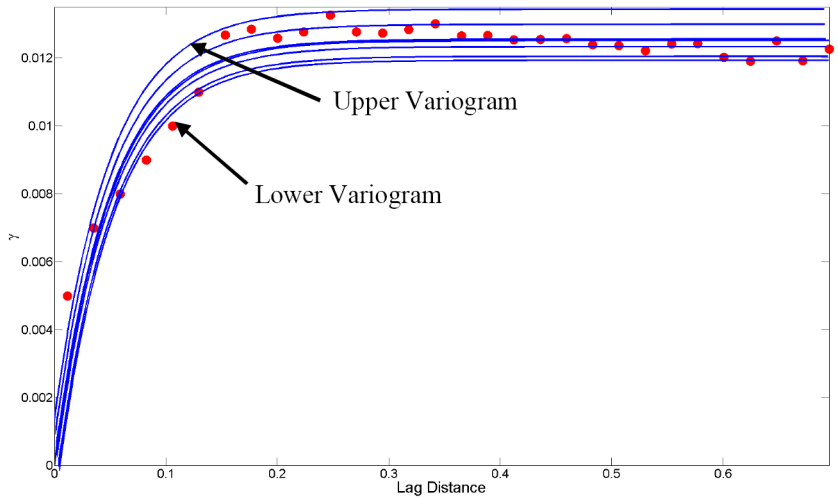


Figure 3—Different variograms fitted to a set of data can be seen as fuzzy variogram

Fuzzy Kriging

An alternative to classical kriging that accounts for larger uncertainty in the data was proposed by Diamond [Diamond at. al.], using fuzzy variograms. Fuzzy variogram can incorporate uncertainty on variogram parameters in geostatistical prediction [11–14]. Furthermore, the fuzzy variogram present the advantage of providing knowledge of possible intervals for variogram parameters since a single variogram cannot be fitted easily with a theoretical model. Preferably, the fitted model must represent the entire population. An example of a variogram as illustrated in Fig. 2 shows a case where the fit is relatively simple. The variogram

fits could vary in precision or rather imprecision if different curves are to be fitted to a set of data as shown in Fig. 3. Nevertheless, the results of different variograms fitted to a set of data are bounded between a lower and upper variograms. All the variograms in the range (between lower and upper variograms) are possible and represent different sections or areas of the field.

To convert the main kriging equation into a fuzzy function, the Extension Principle [15] is used to expand the crisp sets into fuzzy sets:

$$\tilde{\phi}^{*}(u_0) = \tilde{\lambda}_0 + \sum_{i=1}^{n} \tilde{\lambda}_i \tilde{\phi}(u_i) \tag{5}$$

where $\tilde{\lambda}_i = [\lambda_{\min}, \lambda_{\max}], i = 0,...,n$ are fuzzified λ_i that can be calculated from Fuzzy variogram. The results are fuzzy variables $\tilde{\phi}(u_i)$ bounded between [$\#_{\min}(u_i)$, $\#_{\max}(u_i)$]. Note that there are different ways to fuzzify the kriging algorithm. In this paper we incorporate imprecision of variograms using fuzzy interval kriging coefficients. The interval range calculated from fuzzy variogram shows the interval of spatial relationship between data sources. The fuzzy kriging coefficient range was determined and used as inputs in the deep learning neural network.

Deep Learning

Deep learning is a class of machine learning where many layers of information processing are performed in a hierarchical way from preprocessing to features extraction and selection. It composes multiple levels of non-linear transformations that can approximate any complex function. Each level can be seen as a separate quantity which can transform its input to its output deploying non-linear functions. The non-linearity in layers is distributed throughout the network where parameters of non-linear transformations are optimized using historical data. The learning process of a deep learning algorithm is that it extracts useful information from raw data in a sequential way using a large quantity or volume of information.

Deriving features in hierarchical order allows the deep learning model to learn complex correlations from data. The mapping between inputs and outputs is critical in complicated systems; even experts often are unable to explicitly specify information in raw data. The ability to automatically learn features will become increasingly important as the amount of data and range of applications to machine learning methods continues to grow.

Differences between current machine learning algorithms and deep learning networks consist in a depth of architecture where the number of levels of non-linear transformations is significantly higher. Each layer within the series between input and output data identifies important features and further processes them in a series of stages. Through this capability, the deep learning nets are taking away the time consuming, difficult and sometimes impossible feature selection task from the expert. Deep neural networks naturally are exposed to raw data; next, they pre-process the data, extract and select critical features for complex mapping problems and use them for prediction or classification.

Another difference between the classical neural network and deep learning is number of parameters and amount of data needed to train the networks. Due to the depth of architecture in deep learning networks, the number of parameters to estimate is much higher than a classical neural network. Therefore, to estimate the parameters of the network, a larger volume of data is needed to train the network. To speed up the training process on big data, deep learning structure typically is implemented in parallel and distributed processing.

Until recently, most machine learning and signal processing techniques had exploited shallow-structured architectures where the architectures typically contain few layer of nonlinear feature transformations. However, after the recent development of computational power, the significantly lowered cost of computing hardware, and recent advances in machine learning and signal/information processing, researchers have demonstrated successes of deep learning in diverse applications, including classification [15–17], regression

[18], image processing [19], natural language processing [20–21], text mining [22], robotics [23], and petroleum industry [24–26].

A deep learning neural network consists of many simple processing units called neurons. The input neurons gets activated through weighted connection from input data. Linear combination of weighted inputs activates a nonlinear function to generate output (Fig. 4) [13]. Also, a neuron may have a bias which shifts the net input of activation function depending on its range. The learning process of a neural network is about finding the weights and bias of each neuron that make the network illustrate the desired prediction rate.

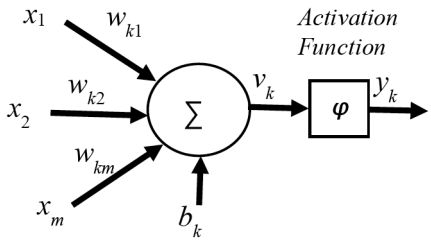


Figure 4—A neuron model

The mathematical model that describe a neuron is

$$v_k = b_k + \sum_{i=1}^{m} w_{ki} x_i \tag{6}$$

and

$$y_k = \varphi(v_k) \tag{7}$$

where x_i are input signals, w_{ki} are weights of neuron k, b_k is the bias, φ is the activation function, and y_k is the output of the neuron k. Fig. 5 shows the structure of a deep neural network constructed by organizing neurons in the form of layers with input, output, and hidden layers.

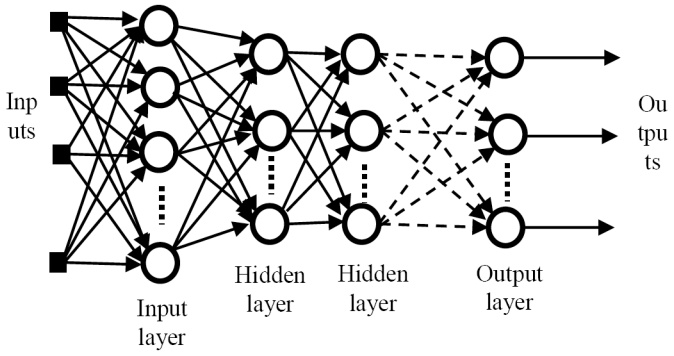


Figure 5—Deep neural network architecture

The internal parameters or weights of a deep learning neural network are determined through the learning algorithm. At each iteration in the learning process, input data propagates feedforward through the network and generate an output. Then the error between actual output and calculated output propagates back through the network to adjust the weights and biases. These steps are performed sequentially for many iterations for all input/output pairs until the desired performance defined by objective function achieves. The learning process from large volumes of data is a computationally intensive task that depends on the number of neurons in each layer, number of layers, and size of data.

Methodology and Results

The application of deep learning and fuzzy kriging involves the following steps: selection of the study field or reservoir, geophysical well logs, location of the wells, generating the reservoir variograms, determination of fuzzy kriging parameters, defining the data sets for modeling, training the model, pseudo-wells generation, and populating the 3D full field model. In this paper, we are presenting only the distribution of the facies, which are derived from the suite of resistivity logs. The other petrophysical parameters such as effective porosity (PHIE), permeability (K), and oil saturation (So) model distributions can be obtained in a similar way once the facies model was developed.

The study field consists of more than 5,000 wells drilled in relatively small five spot patterns. Despite the large number and age, most of the wells in the field have a good suite of logs including SP, GR, SRES, MRES, DRES, NPHI, CALIPER, etc. The data is stored in large databases and is easily accessible to different tools for reservoir surveillance and modeling.

Given a large number of the wells and the size of the model, we selected only a small part of the field consisting of 1202 wells to prove the methodology. We focused our study only on the productive intervals, thus selecting just the log intervals covering the target information. With this consideration, the total dataset comes to be about 4 million plus data points. We divided input/output data pairs into three data sets: *training, validation* and *testing*. About 10% of data was selected and set aside for testing the performance of the system (121 wells). The remaining 90% of data (1081 wells) was used in the learning or training process. This set was also divided into training and validation folds. The location of the wells used in the training and testing are shown in Fig. 6.

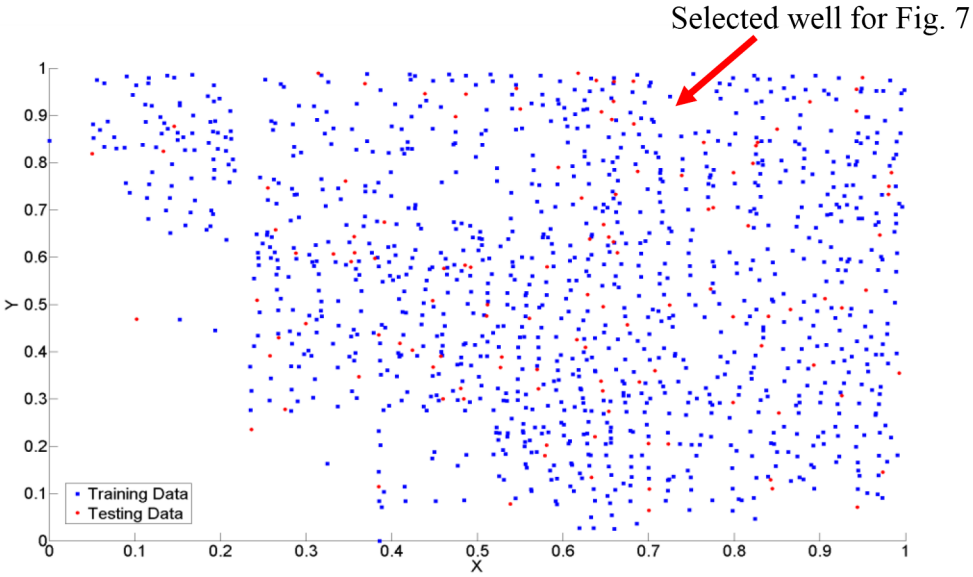


Figure 6—Location of training and testing wells

The development of the modeling dataset required some additional considerations. Since the model is intended to estimate the petrophysical properties at any given location, the information of nearest neighbor wells was included in the learning dataset. A representation of the nearby wells selection is shown as marked connections in Fig. 7.

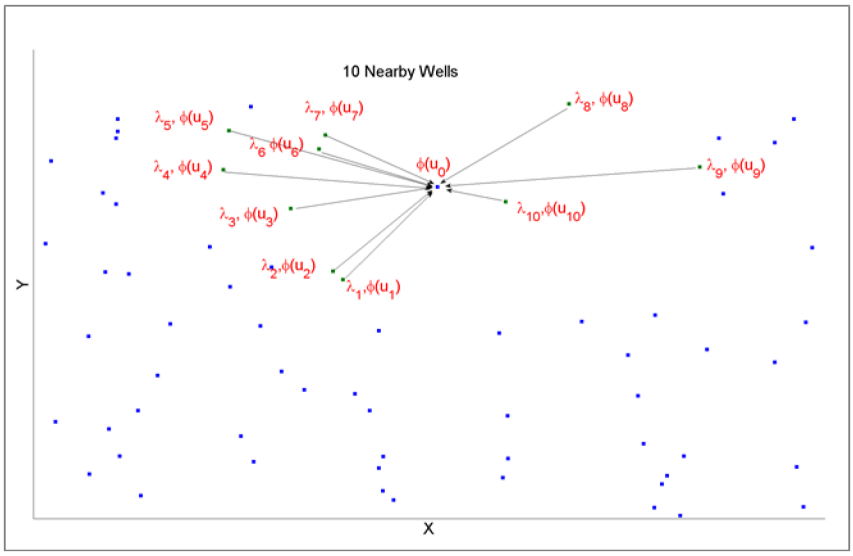


Figure 7—An example of 10 nearby wells

The pertinent information of nearby wells was included to construct a modeling data set. Consequently, the new constructed dataset consists of sufficient features, and it is suitable to predict facies everywhere in the field. Commonly, the number of nearby wells included is based on the distribution and density of the wells, as well as the size of the field.

In our study, we developed three datasets for the three models we proposed and tested. The first dataset is based on the angle and distance of nearby wells to the target well; the second dataset capture the uncertainty of the field using kriging coefficients; and the third dataset includes fuzzy kriging range. Three deep learning neural networks models were trained based on the three data sets and the results were compared with the existing earth model developed based on kriging technology. All deep learning models considered have 200 to 300 input attributes. The networks have between 3 to 6 hidden layers, and the number of hidden neurons ranges between a minimum of a 100 to a maximum 250 respectively. The training process ends whenever the validation error starts increasing to avoid overfitting, or the minimum objective function is achieved.

Two different means are presented to demonstrate the performance of the model. First, the pseudo-logs generated from the three models were compared to the measured logs from newly drilled wells. This comparison will establish the deep learning model using fuzzy kriging range, which has a better performance as contrasted to the other two models, distance-angle and simple kriging respectively. Secondly, we compared the cross-sections from the existing earth model to the fuzzy kriging deep learning. The contrast will prove a significant resolution improvement and more realistic realization of the subsurface model.

Fig. 8 and 9 shows four well examples where the measured log recorded in the well is contrasted to the three different deep learning neural network models, distance-angel, kriging and fuzzy kriging. The plots show that pseudo-logs generated by the deep learning using the fuzzy kriging range have illustrated the closest fit to the measured log, thus demonstrating that model performance is increased when the geological uncertainty of the field is captured by a range of variograms.

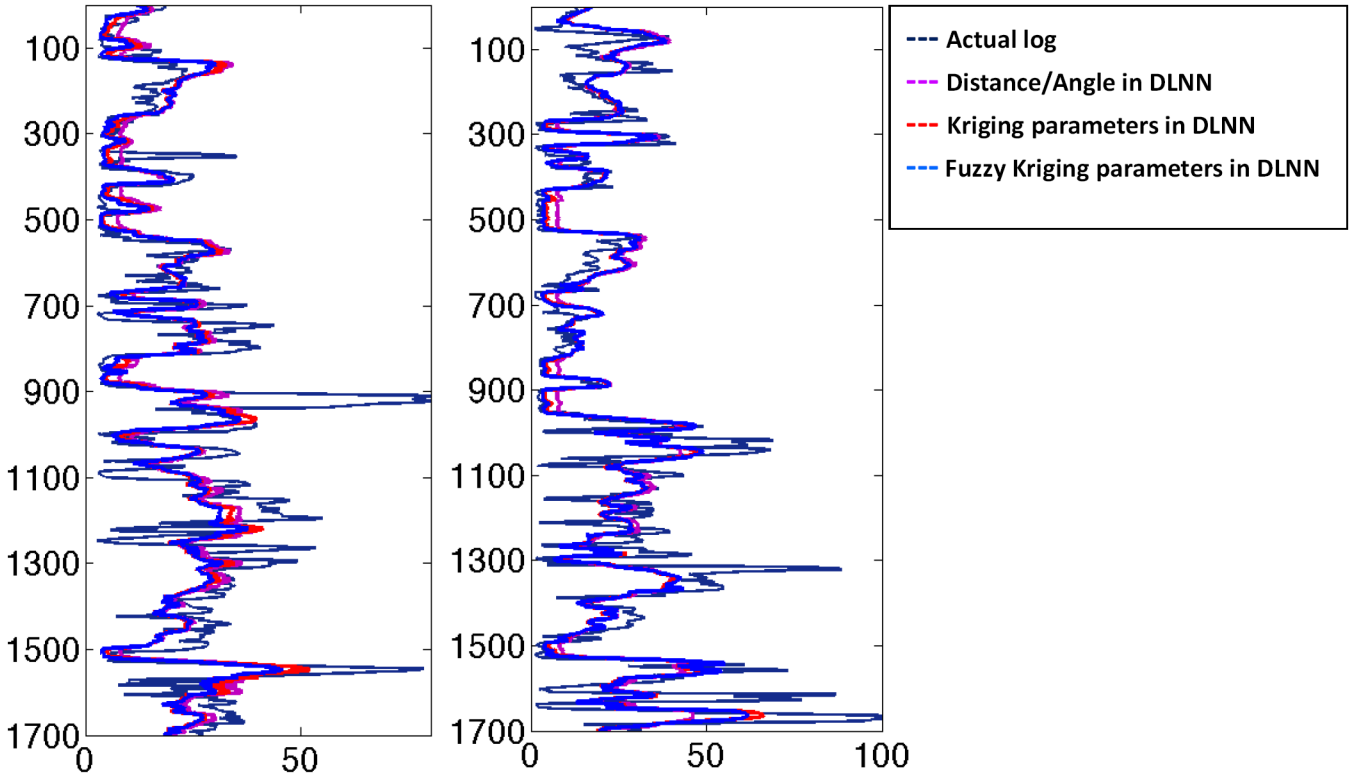


Figure 8—Comparison of measured logs vs the three pseudo-logs generated using deep learning: Distance/Angle, Kriging, and Fuzzy Kriging (Wells WH-221 and WH-767)

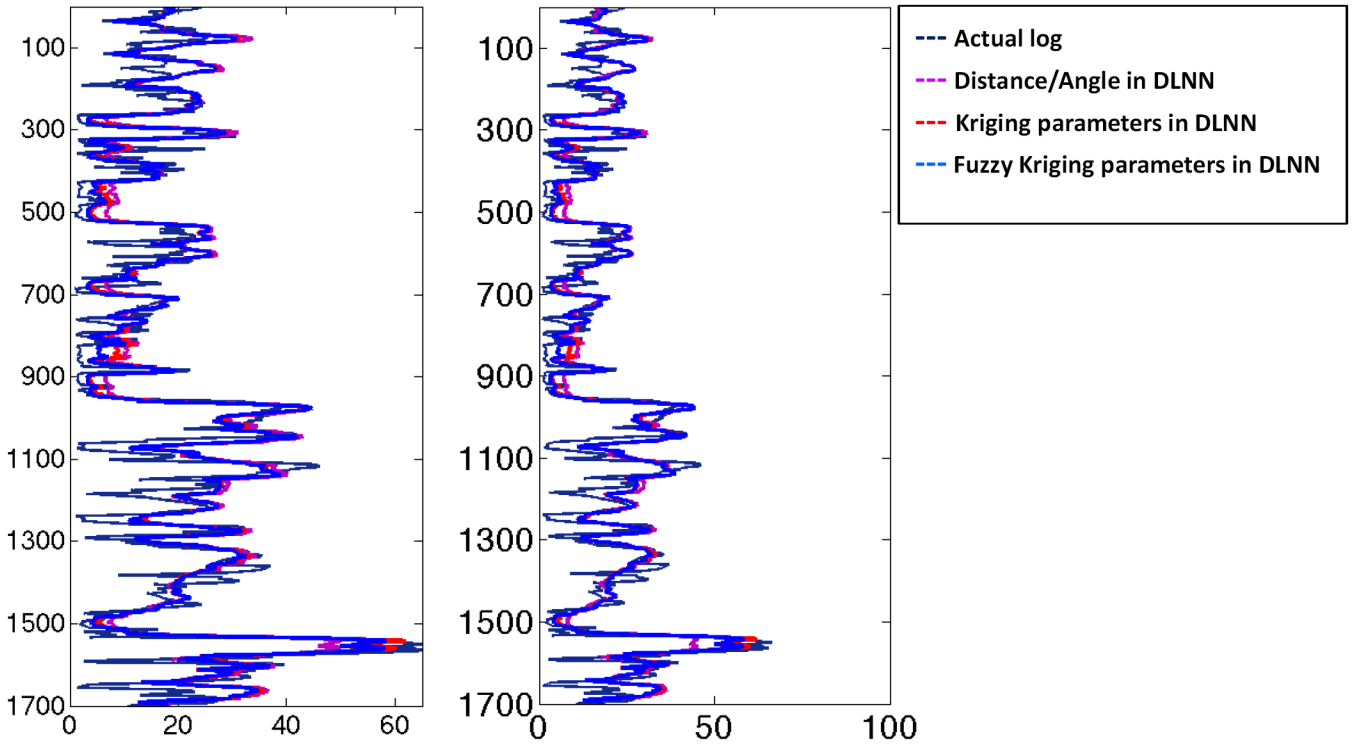


Figure 9—Comparison of measured logs vs the three pseudo-logs generated using deep learning: Distance/Angle, Kriging, and Fuzzy Kriging (Wells WH-1221 and WH-1637)

Fig. 10 contrasts only the recorded well log with the deep learning neural network model using fuzzy kriging range. In both cases, the pseudo-logs show a very strong correlation to the measured traces, even though the wells are located in the opposite parts of the field.

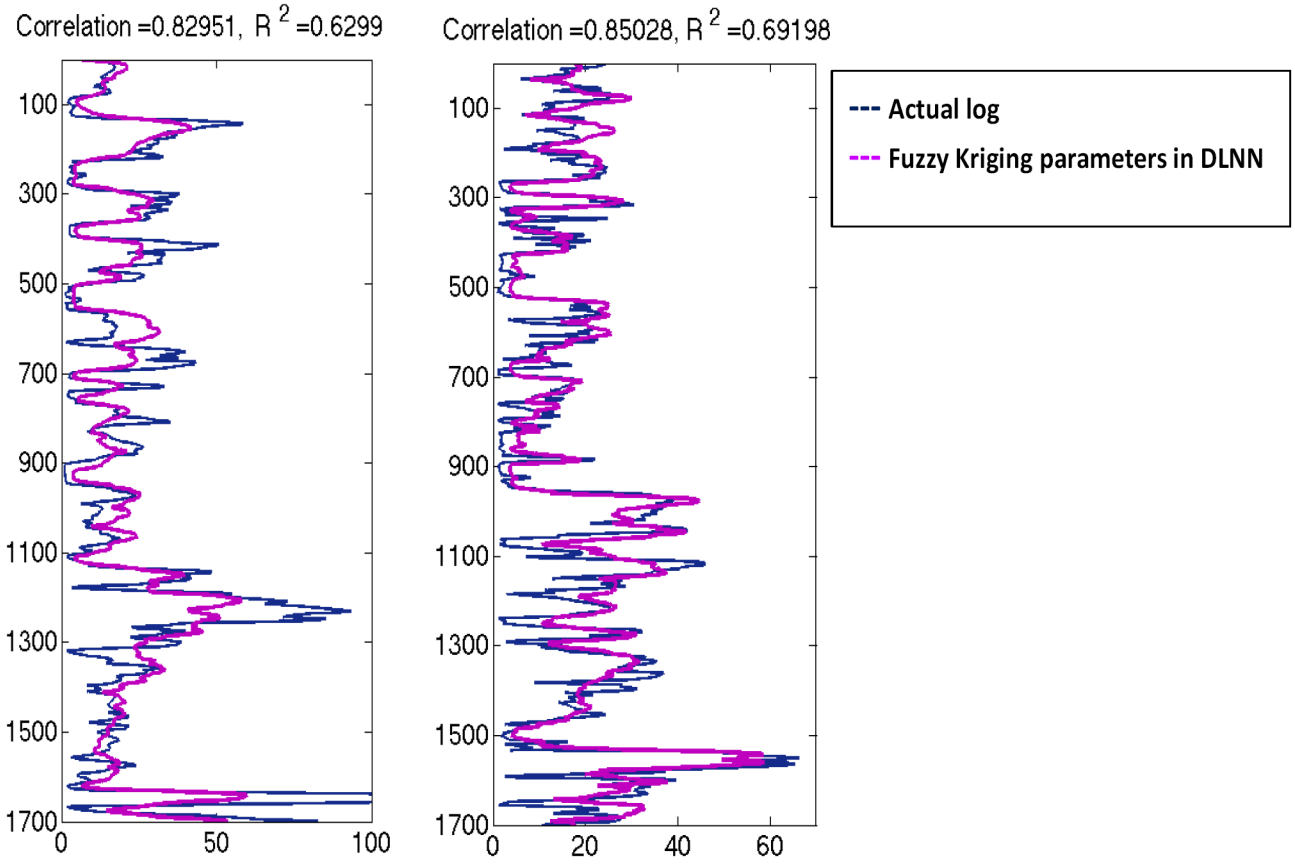


Figure 10—Actual vs. predicted DRES logs for new wells using fuzzy kriging and deep learning

Table I shows a comparison of the average correlation coefficient and R-square for testing wells. To obtain the average and standard deviation of different method we run the model on every testing well and find the correlations and R-square between actual output and predicted output. The data confirms that deep learning with fuzzy kriging provides the highest correlation coefficient and R-square. It implies that fuzzy kriging better captures the uncertainty between geological locations of wells and provides better model performance.

Table I—Correlation of Deep learning neural network with different spatial estimator

	Distance	Kriging	Fuzzy Kriging
Average correlation	0.6578	0.6702	0.6899
STD correlation	0.1152	0.1206	0.1191

The second test of the model is to contrast cross-sections from the existing earth model to the deep learning model using fuzzy krigging. A subsection of the field was selected (Fig. 11) together with its coresponding existing grid cell base earth model. The deep learning model was applied to the existing earth model and each grid cell was populated with the corresponding facies (resistivity).

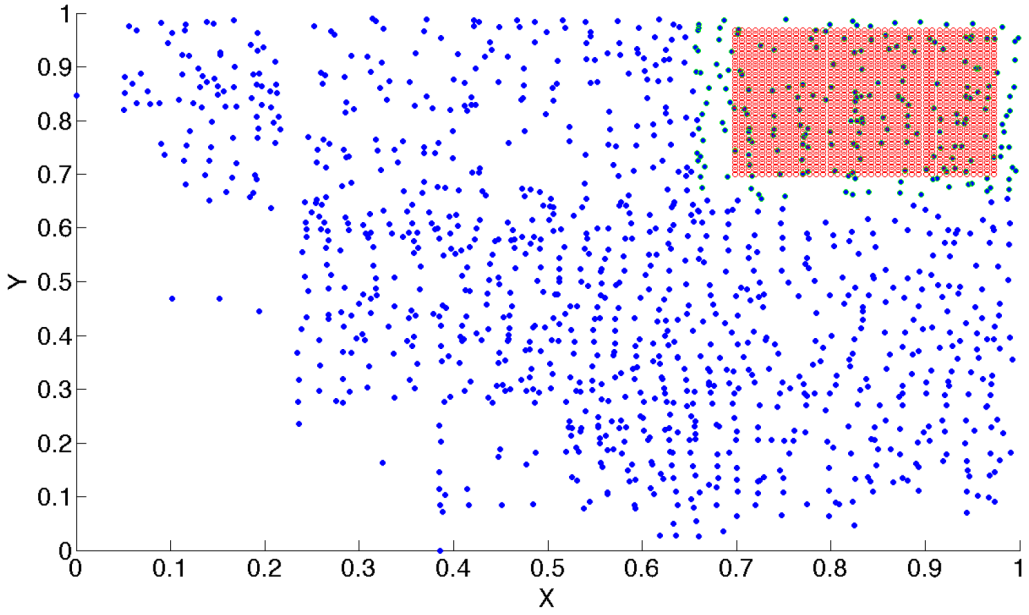


Figure 11—Subsection of the field to generate facies

Figures 12 and 13 show two vertical cross sections of the sub-grid field taken at the same location. Fig. 12 illustrates in comparison an N-E cross section of the model. The left cross section is taken from the current existing kriging model, while the right one from the deep learning neural (DLNN) net model. One can observe a significant difference regarding the continuity of formations with the DLNN model showing a natural trend for the depositional environment as compared to classical kriging algorithm. The sharp sand facies changes shown in the classical kriging algorithm appear less realistic and do not match the natural rends of the sands and silts as seen in the outcrop. In contrast, deep learning using the 'γ' variance from the kriging algorithm seems to better capture the natural geologic trends and reduces the reservoir characterization uncertainty of the field. This method takes advantage of the natural dip of the well data values among various wells. It more naturally reproduces the sand/silt contacts and facies distribution. The DLNN model represents a more natural depositional environment. The high ohm-m sand lobes in the DLNN model are more naturally distributed and a clearer picture of the fluvial channels is represented.

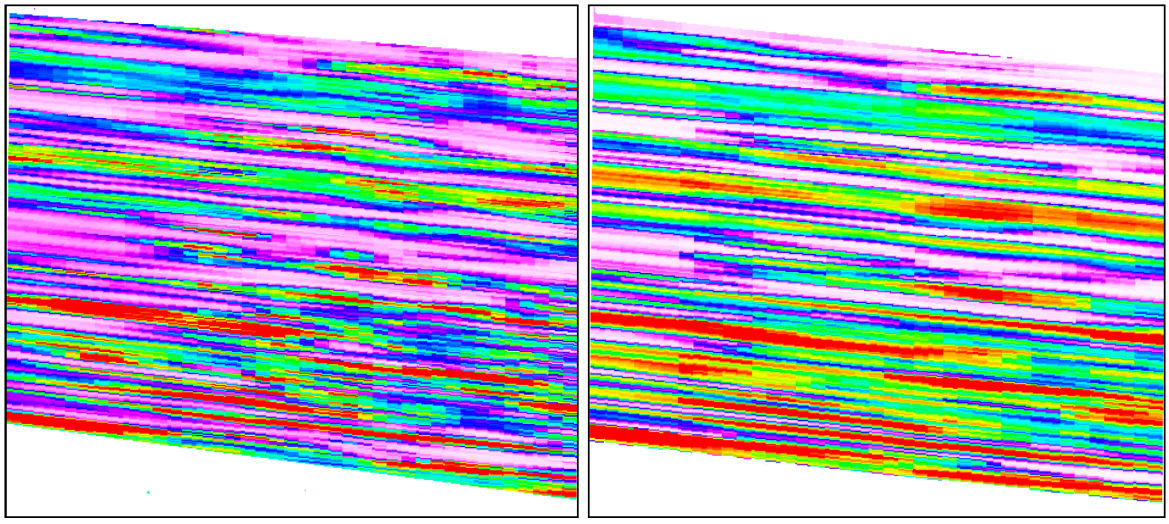
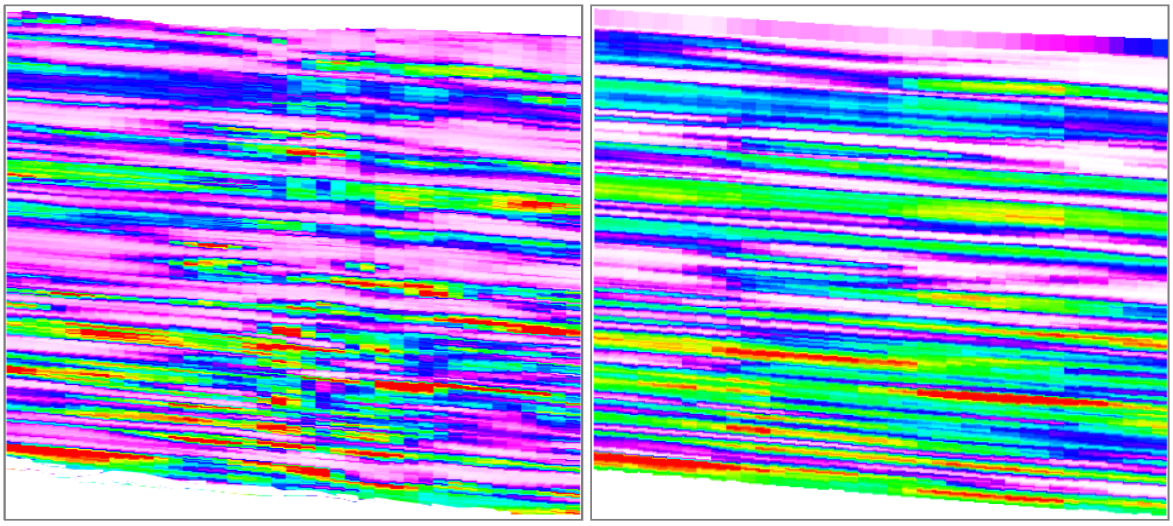


Figure 12—N - E Horizontal cross section (a) kriging estimate (b) deep neural network estimate



(a) (b)

Figure 13—S – W Horizontal cross section (a) kriging estimate (b) deep neural network estimate

Fig 13 illustrates a second example where two S - W crosssections are compared. Same observations as discussed above apply.

Figures 14 and 15 show two horizontal crosssections of the subgrid field taken at different depths. In both figures, the left picture illustrates the existing kriging earth model, while the right picture the deep learning neural net. One can observe the natural trend of the channel which corresponds to the fluvial/ alluvial depositional environment as shown in the DLNN model. In constrast, the existing kriging model seems to show more of a "piecemeal" interpolatiton, with dotted kriging artifacts (property modeling errors) that distort the property distribution and are unable to capture the natural channel trends. The DLNN model is able to reproduce channel trends that are natural and reflective of seismic data. The DLNN model does not allow for unrealistic sand/silt facies distribution and erroneous juxtapositions of fascies. As a result the DLNN model will better allow the project team to recognize the connectedness of the sand bodies in the reservoir and better delineate the project areas. Greater understanding of the reservoir connectedness makes more efficient use of steam for planned and executed steamflood projects. It also increases the likelyhood of success in our committed recovery of hydrocarbons.

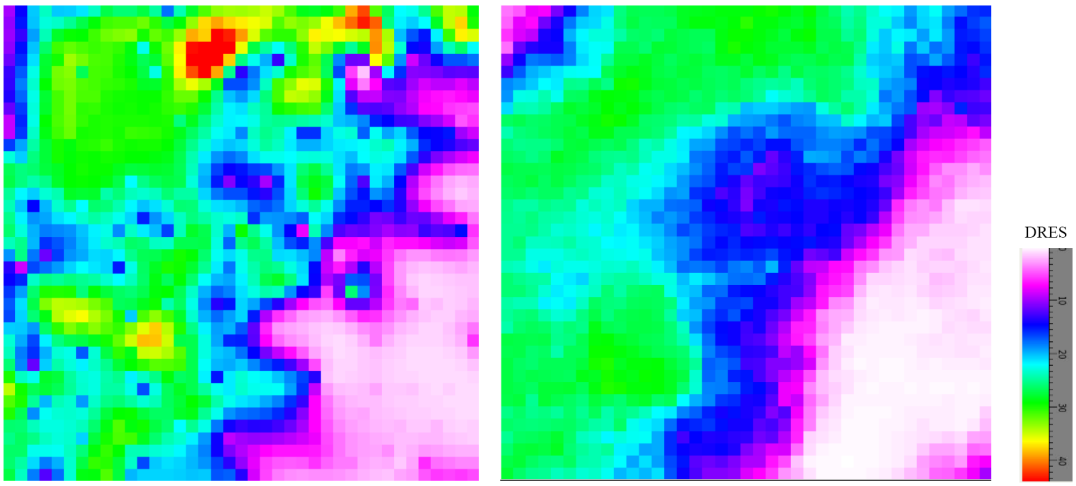


Figure 14—Horizontal cross section (a) kriging estimate (b) deep neural network estimate

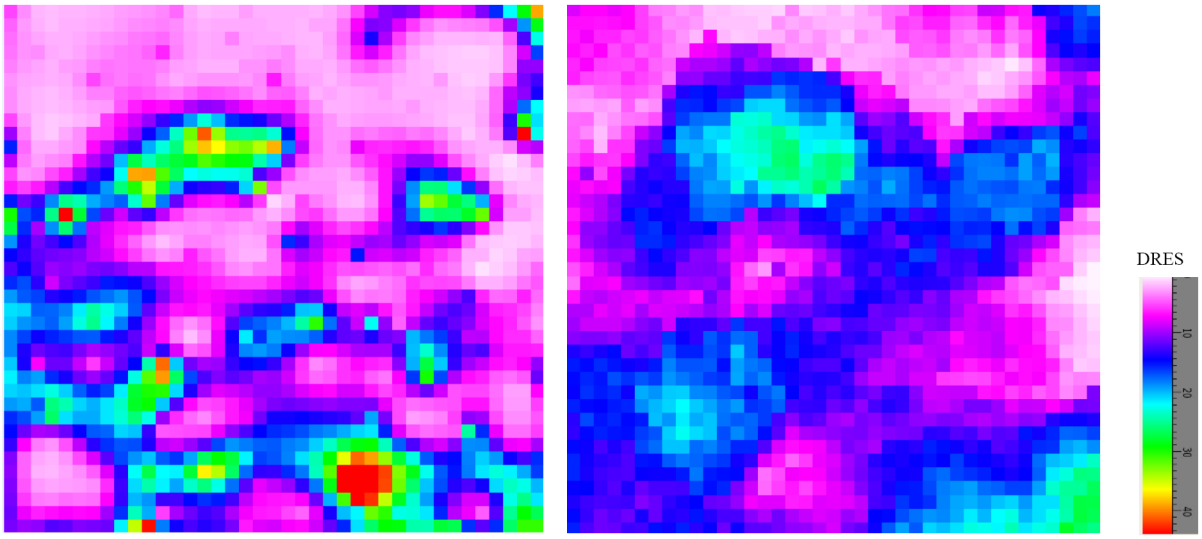


Figure 15—Horizontal cross section (a) kriging estimate (b) deep neural network estimate

Discussion

The methodology presented performed well in this fluvial/alluvial reservoir. More research should be conducted to investigate if this methodology can be applied in other depositional environments where lateral formation variability has different characteristics. Examples would be environments where pinch outs cause a disappearance of certain features or where faulting affects the facies continuity.

An additional deep learning neural network model was further developed and used to predict other reservoir properties like oil saturation and steam. The predicted relationship was based on the natural association seen in the well data, and not interpolated independent of one another and bootstrapped together incongruously as typically done in models.

Developing such a new approach and model is not an easy task. Moreover, an adoption of such models seems to be even harder to achieve.

Conclusion

We present a new research that uses fuzzy variogram as input in deep learning neural network models to capture the geological uncertainty and considerably enhance the generation of synthetic logs at any location in the field. We applied our methodology to a heavy oil field with a large number of wells, which has the reservoir consisting of alluvial fans. Furthermore, the application of deep learning neural network assisted by fuzzy kriging or fuzzy variograms proved to be a powerful technology for improved reservoir characterization when large volumes of data are available.

There is a considerable business value in having an accurate full field model for a reservoir. This new modeling method delivers a more natural prediction of sand/silt facies, it does not juxtapose sand and silt unnaturally, displays the natural continuity of reservoir connectivity, and provides a clearer picture of the reservoir architecture. Furthermore, it helps to more accurately define steam containment of existing and planned steamfloods; and improve project hydrocarbon volume prediction. Better modeling of the reservoir not only affects the volume of oil but the movement of steam and the drainage of the mobilized oil. Most important, an accurate and realistic model aids in well placement optimization of both vertical and horizontal well planning and execution.

Acknowledgement

The authors would like to thank Chevron North America Exploration and Production for allowing the publication of this work. Additional gratitude is expressed to Jeff Woolford and Robin Fleming for their input and guidance.

References

1. A.G. Jornel and Ch.J. Huijbregts, *Mining Geostatistics*, Academic Press, London, 1978.
2. B.D. Ripley, *Spatial Statistics*, Wiley, New York, 1981.
3. N. A. C. Cressie, "The Origins of Kriging," *Mathematical Geology*, v. **22**, pp 239–252, 1990.
4. F. P. Agterberg, *Geomathematics, Mathematical Background and Geo-Science Applications*, Elsevier Scientific Publishing Company, Amsterdam, 1974.
5. N. Cressie and D. Hawkins, "Robust Estimation of the Variogram," *Mathematical Geology*, v. **12**, p. 115–125, 1980.
6. P. A. Dowd, "The Variogram and Kriging: Robust and Resistant Estimators," *Geostatistics for Natural Resources Characterization*, Amsterdam, p. 91–106, 1984.
7. A. G. Joumel, "New Distance Measures: The Route Toward Truly Non-Gaussian Geostatistics," *Mathematical Geology*, v. **20**, p. 459–475, 1988.
8. M. R. Srivastava, and H. M. Parker, "*Relative Variograms and Robust Spatial Continuity Measures*," Kluwer Academic Publishers, p. 295–308, 1989.
9. P. I. Brooker, "A Parametric Study of Robustness of Kriging Variance as a Function of Range and Relative Nugget Effect for a Spherical Semivariogram," *Mathematical Geology*, vol. **18**, pp. 477–488, 1986.
10. M. Armstrong, "Improving the Estimation and Modelling of the Variogram," *Geostatistics for Natural Resources Characterization*, Amsterdam, pp.1–19, 1984.
11. A. Bardossy, I. Bogardi, and W. E. Kelly, "Kriging with Imprecise (Fuzzy) Variograms. I: Theory," *Mathematical Geology*, Vol. **22**, No. 1, 1990.
12. P. Diamond, "Fuzzy Kriging," *Fuzzy Sets and Systems*, vol. **33**, pp. 315–332, 1989.
13. A. Consonni, R. Iantosca, and P. Ruffo, "Interval and Fuzzy Kriging Techniques Applied to Geological and Geophysical Variables," *Soft Computing for Reservoir Characterization and Modeling*, Physica-Verlag HD, pp. 73–103, 2002.
14. A. L. Zadeh, "Fuzzy sets." *Information and control*, vol. **8**.3, pp. 338–353, 1965.
15. H. Lee, P. Pham, Y. Largman, and A. Y. Ng, "Unsupervised feature learning for audio classification using convolutional deep belief networks," *In Advances in neural information processing systems*, pp. 1096–1104. 2009.
16. R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, "Convolutional-recursive deep learning for 3d object classification," *In Advances in Neural Information Processing Systems*, pp. 665–673. 2012.
17. R. Salakhutdinov and G. E. Hinton, "Using deep belief nets to learn covariance kernels for Gaussian processes," in *Advances in Neural Information Processing Systems*, pp. 1249–1256, Cambridge, MA: MIT Press, 2008.
18. D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," *In Computer Vision and Pattern Recognition (CVPR)*, 2012 IEEE Conference on, pp. 3642–3649. IEEE, 2012.
19. R. Collobert, and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," In 25th international conference on Machine learning, pp. 160–167. ACM, 2008.

20. T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," In *Advances in neural information processing systems*, pp. 3111–3119. 2013.
21. X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," In International Conference on Artificial Intelligence and Statistics, pp. 315–323. 2011.
22. I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *The International Journal of Robotics Research* **34**, no. 4–5: 705–724, 2015.
23. A. Aggarwal, and S. Agarwal, "ANN Powered Virtual Well Testing," Offshore Technology Conference, doi:10.2118/174871-MS, 2014.
24. S. P. Ketineni, T. Ertekin, K. Anbarci, and T. Sneed, "*Structuring an Integrative Approach for Field Development Planning Using Artificial Intelligence and its Application to an Offshore Oilfield*." Society of Petroleum Engineers. doi:10.2118/174871-MS, September **28**, 2015.
25. Haykin, Simon S., Simon S. Haykin, Simon S. Haykin, and Simon S. Haykin. *Neural networks and learning machines*. Vol. **3**. Upper Saddle River: Pearson Education, 2009.
26. M. Korjani, A. Popa, E. Grijalva, S. Cassidy, and I. Ershaghi, "*A New Approach to Reservoir Characterization Using Deep Learning Neural Networks*," Society of Petroleum Engineers, doi:10.2118/180359-MS, 2016.