SPE 162699

Fast Track Analysis of Shale Numerical Models

A. Kalantari-Dahaghi ,SPE, S. Esmaili, SPE, West Virginia University, S.D. Mohaghegh, SPE, Intelligent Solution Inc. & West Virginia University

This paper was prepared for presentation at the SPE Canadian Unconventional Resources Conference held in Calgary, Alberta, Canada, 30 October–1 November 2012.

Abstract

Latest advances in shale gas reservoir simulation and modeling have made it possible to optimize and enhance the production from organic rich shale gas reservoirs.

Reservoir simulator is no longer used with a simple description of the complex shale gas reservoirs, but with multiple, equally probable realizations to allow risk assessment. Nevertheless, the perennial challenge in shale reservoir modeling is to strike a balance between explicit representation of reservoir complexity and long simulation run time for multiple realizations.

Focus of this study is on the development, calibration and validation of a Shale Surrogate Reservoir Model (AI-based proxy model) that represents a series of complex shale numerical simulation models. The Shale Surrogate Reservoir Model is then used for fast track analysis of the shale numerical model. Reservoir simulation model for a generic shale gas reservoir are constructed using a popular commercial simulator that is capable of handling complex fracture network (natural and multiple stages of hydraulic fractures), different sorption types (instantaneous and time dependent one), and capturing long transient nature of the flow in shale matrix. Validation of the Shale Surrogate Reservoir model using several blind simulation runs is also presented.

Shale Surrogate Reservoir Model is a replica for the numerical simulation model with response that is measured in fractions of a second. As such, it provides the means for comprehensive and fast track analysis of the model in a relatively short period of time, allowing the reservoir engineer to scrutinize different realizations and propose development strategies.

Introduction

Numerical models capable of modeling the most important features of tight gas and gas shales are undergoing further development to include better representations of the basic physics controlling gas flow as the industry learns more. (Lee and Sidle, 2010)

Various attempts have been made to model flow in shale gas systems. Models describing detailed physical processes can be built in a numerical setting where fractures can be discretely modeled and matrix blocks are assigned to transfer gas through diffusion and desorption into the fracture blocks (e.g. Lewis, 2004; Cipolla et al. 2009, Rubin 2010; Wang 2011; Mongalvy 2011).

Many difficulties on shale gas modeling caused engineers to rely heavily on the simplest, most accessible tool: such as using a reduced physics models (Wilson et al.2012) or using the simplest production data analysis approach(e.g. Decline curve analysis) by knowing the fact that tools of traditional production data analysis have not been sufficient identifying flow behavior in shale system. At the same time, other researchers try to present some conditions of the design parameters that should be considered to build proper and more accurate shale gas simulator. Civan et al. (2011) proposed a quad-porosity approach accounting for a complicated reservoir pore structure that includes pores in the organic matter, inorganic matter; natural and hydraulic fractures with heterogeneous wettability, and different relative permeability and capillary pressure functions.

In this paper, our objective is not to compare and contrast different shale modeling and simulation technique nor detailing the strength and weakness of current simulator for shale gas modeling. Our goal is to introduce a technique that can be helpful for fast track and efficient analysis of shale simulation model. (By assuming that the shale simulation model is fully capable of modeling hydrocarbon storage and production from this type of reservoir)

Proxy Models vs. Surrogate Reservoir Model (SRM)

Proxy models-They have become popular in E&P industry over past several years. They are mathematically/statistically defined functions that replicate simulation model output for selected input parameters and provide fast-approximated solutions that substitute large numerical simulation models and are used to fulfill many different purposes.

They are used to assist in the field development planning, uncertainty analysis, optimization of operational design, and history matching. Most common proxy models are either reduced models (increase run-time speed by grossly approximating the problem) or response surfaces (grossly approximating the solution space).

Current state of building proxy models, when it comes to representing numerical reservoir simulation models, leaves much to be desired. Nevertheless, they are routinely developed and used in order to generate fast solutions to changes in the input space.

Given the fact that computational power has increased vastly in recent years, an argument can be made about the role and even necessity of developing proxy models. However, Williams et al. (2004, 2006) observe that gains in computational power "are generally used to increase the complexity of the models rather than to reduce model run time." Therefore, still there is a need in development of effective and robust proxy models.

The most prominent and commonly used approaches for developing a proxy model that can be used in reservoir simulation and modeling are as follows:

RSM-Polynomial regression models are especially suitable in cases where the problem is known to be governed predominantly by low-order effects (linear or quadratic) and the number of input variables is limited ($n \leq 10$). When defining the setup of RSMs, the only free choice is in the regressors. On the one hand, this limited number of options makes RSM easy to handle, but on the other hand, it restricts the range of possible applications.

MLS-The moving-least-squares approach makes use of regression techniques to define a global model based on a locally weighted polynomial approximation, which typically does not interpolate the observations. Accordingly, MLS models are favorable whenever RSMs cannot be established to be globally valid, but the function to be approximated is smooth enough such that in the proximity of the prediction point a polynomial relationship can be accepted.

Kriging- This type of proxy model is extremely flexible due to the large variety of feasible correlation formulations but from a computational point of view, kriging models are less favorable. For the fitting process, an n-dimensional optimization is necessary to find the correlation parameters via maximum likelihood estimation

RBF- Radial basis function approximations are comparable to kriging models. In this technique, the flexibility to adapt to many different applications is due to the variety of radial basis functions. The accuracy of the obtained approximation is affected by the opted radial basis function and its free parameter. Similar to the problem of specifying the weighting functions for MLS approximations, there is no general rule how to find the best radial basis function for a problem.

Shale Surrogate Reservoir Model (AI-Based Proxy Model) –is an Artificial intelligence-based proxy model that has the capability of reproducing highly accurate well-based (focus of this paper) and grid-based simulation responses as a function of changes to all the involved input parameters (reservoir characteristics and operational constraints) in fraction of a second. It integrates reservoir engineering and reservoir modeling with machine learning and data mining.

The approximation quality can be controlled by the selection of appropriate activation functions and efficient spatio-temporal data base generation. On the other hand, the wide range of alternatives encountered while setting up artificial neural network (ANN) approximations can pose a problem for the inexperienced user, which often leads to a categorical rejection of this method. However, if employed properly, ANNs are a suitable approximation technique for high-dimensional and highly nonlinear problems.

As presented by Mohaghegh at al. (2012), this is not the first time that a technology has been misused, consequently misjudged, and prematurely dismissed. The brief explanation provided in Zubarev's study (2009) on how the Neural Network has been used to build the proxy model presents ample reasoning on why it did not work. Without going into the details of what was wrong with the way the Neural Network was used in the aforementioned study, it suffices to say that whenever

Neural Networks have been used purely as a regression tool it has resulted in disappointing outcomes. Neural Networks should not be used merely as a regression tool, without paying attention that as part of a larger toolset, it attempts to observe, learn, and generalize. This is due to the fact that artificial intelligence and data mining (as an overarching discipline) are far more than regression tools and certain understanding of machine learning activities are required for their effective use and deployment.

Shale surrogate model, which is based on the pattern recognition capabilities of artificial intelligence and data mining, has several advantages over traditional proxy models such as response surfaces or reduced models. These advantages include:

- No need to approximate the existing simulation model to develop an AI-based proxy model
- The number of simulation runs required for the development of an AI-based proxy model is at least an order of magnitude less than traditional proxy models, and
- Beyond representing the pressure and production profiles at each well individually, AI-based proxy model can replicate, with high accuracy, the pressure and saturation changes at each grid block.

During the development process, numerical reservoir simulation model is used to teach the SRM (AI-based proxy model) the principles of fluid flow through shale porous media and the complexities of the heterogeneous reservoir represented by the geological model and its impact on the fluid production and pressure changes in the reservoir.

Unlike other proxy models that require several hundreds of simulation runs for their development, AI-based one requires only a small number of simulation runs by using a representative spatio-temporal database. This database is uniquely built for each problem based on reservoir simulation runs for training the neural network to understand the impact of changing all parameters involved in shale gas modeling (e.g. matrix and natural fracture properties, hydraulic fracture characteristics ,sorption type and flow regimes). In other words, all that you want the SRM to learn (and therefore be able to accurately reproduce) must be included in this database.

Necessity of SRM has to do with the fact that massive potentials of the existing numerical reservoir simulation models go unrealized because it takes a long time to make a single run. Numerical models that are built to simulate complex reservoirs with multi-million grid blocks require considerable run-time even on cluster of parallel CPUs.

Exhaustive and comprehensive evaluation of the solution space for designing field development strategies as well as quantification of uncertainties associated with the static model are the type of analyses that require large number of simulation runs in order to provide meaningful and usable results. When a numerical simulation model takes hours for a single run, performing such analyses become impractical and the engineers have to compromise by designing and running a much smaller number of runs in order to make decisions.

When the quantification of uncertainties associated with the geologic model is one of the objectives, SRM can prove to be valuable asset. Since a single run of the fullfield model takes several minutes or hours, uncertainty analysis becomes a painful and time-consuming process. In order for the uncertainty analysis to be meaningful, the number of geologic realizations of the reservoir must be statistically significant. This is the reason behind the fact that most of the techniques used for quantification of uncertainties call for analysis of several (sometimes hundreds of) geological realization of the reservoir. As the number of independent parameters involved in a problem increases, so does the number of realizations needed for statistical significance. (Mohaghegh.2011, 2012)

Spatio-Temporal Data Set Generation

To collect training data efficiently, the locations for sampling points have to be chosen systematically thus assuring a maximum gain in information with minimal effort. Various methods were proposed in the literature for this purpose, typically summarized under the notion of design of experiment.

It is advantageous to minimize the number of sampling points in order to reduce the experimental effort, but on the other hand, it is important to gather as much information as possible about the major characteristics of the shale system under investigation. Finally, the individual specification of the SRM (AI-based proxy model) formulation intended for the approximation must be considered when selecting adequate coordinate settings of sampling points.

There are different design of experiment techniques such as: Full Factorial designs, Fractional Factorial designs, Central composite design, Box-Behnken design, D-optimal design, Plackett –Burman design (For detail explanation of each

technique please refer to Jurecka 2007). Selecting the appropriate design of experiment method depends on the comparative, screening and response surface objectives which is summarized in Table 1(e-Handbook of Statistical method, 2012, NIST)

Table 1-Summary table for choosing an experimental design for comparative, screening, and response surface designs

Number of Factors	Comparative Objective	Screening Objective	Response Surface Objective
1	1-factor completely randomized design		
2-4	Randomized block design	Full or fractional factorial	Central composite or Box-Behnken
5 or more	Randomized block design	Fractional factorial or Plackett-Burman	Screen first to reduce number of factors

In order to generate a spatio-temporal data set in this study, two different approaches are taken. In the first approach, 54 different simulation runs are designed by having 15 parameters and 5 different values for each, which are uniformly distributed over the uncertainty domain. In each run, only one-parameter is changed at the time, which is completely linear way of looking at a non-linear system.

Since in shale gas numerical reservoir simulation, we are dealing with highly non-linear system therefore input dataset with experiments uniformly distributed over the uncertainty domain may not be sufficient for construction of an proxy model. To overcome this limitation Latin Hypercube Sampling (LHS) with a genetic algorithm is recommended and used for sampling.

Latin hypercube sampling, due to McKay et al. (1979), is a strategy for generating random sample points ensuring that all portions of the vector space is represented.

Latin hypercube sampling (LHS) (Carnell 2011, Stocki 2005, Stein 1987) was developed to generate a distribution of collections of parameter values from a multidimensional distribution. A square grid containing possible sample points is a Latin square if there is only one sample in each row and each column. A Latin hypercube is the generalization of this concept to an arbitrary number of dimensions. When sampling a function of k variables, the range of each variable is divided into n equally probable intervals. n sample points are then drawn such that a Latin Hypercube is created. Latin Hypercube sampling generates more efficient estimates of desired parameters than simple Monte Carlo sampling.

Simulation Model Description

Net pay thickness of 100 ft with five simulation layers is considered to build the simulation model using commercial reservoir simulator. Orthogonal corner point gridding method is used to divide the reservoir into 5838 lateral grid (50*50 ft) blocks (139*42).

A horizontal well with a lateral length of 3050 ft completed in a typical shale gas reservoir is considered in this study. The well is completed with eight stages of hydraulic fracturing treatment. Logarithmic local grid refinement technique is used to model the hydraulic fractures. To account for a natural fracture network, two sets of fracture are defined then the constructed DFN upscaled to generate natural fracture porosity, natural fracture permeability, and matrix-fracture transfer function (sigma) distribution.

Traditional dual porosity models assume that the matrix to fracture flow is in steady state, and thus the matrix cell can be regarded as a single cell. In shale gas reservoirs, the flow is not instantaneous and requires matrix subdivision to capture transient nature of the matrix to fracture flow. Therefore a discretized matrix model is used, which sub-divides each matrix cell into a series of nested sub-cells (3 sub-cell in our case), allowing the simulator to predict the transient behavior in shale matrix.

In order to account for near-wellbore pressure drop associated with turbulence, which causes a non-Darcy addition to the pressure drop, Forchheimer's coefficient (also known as β-factor) was included in the model.

Because physical sorption is the key process in gas shale systems, different sorption type (instant and time dependent) should be introduced to the simulation model. In our case, time dependent type of sorption is considered.

Results and Analysis

As explained before, two scenarios are defined for spatio-temporal data base generation. In the first case, linear sensitivity analysis approach is used to generate the data set, which is a very simple way to approach complex and non-linear problem.

In this case, 54 simulations runs are defined and in each run, only one-parameter is changing while the rest of parameters remained unchanged. The input is randomly sampled and uniformly distributed over the uncertainty domain. A complete list of inputs that are included in the data set is shown in Table 2.

Table 2-Main input data list for first Shale-SRM development

Matrix porosity	Matrix permeability	Rock compaction	Natural fracture porosity	Natural fracture permeability
Sigma factor	Hydraulic fracture height	Hydraulic fracture half length	Hydraulic fracture conductivity	Hydraulic fracture Spacing
Sorption type(instant or time dependent)	Longmuir volume	Longmuir pressure	Diffusion coefficient	Bottom hole pressure

In this study, a multilayer neural networks or multilayer perceptions, are used (Haykin 1999).These networks are most suitable for pattern recognition especially in non-linear problems Neural networks have one hidden layers with different number of hidden neurons that are selected based on the number of data record available and the number of input parameters selected in each training process. The training process of the neural networks is conducted using a back propagation technique (Chauvin et al.1995).In the training process, the data set is partitioned into three separate segments. This is done in order to make sure that the neural network will not be trapped in the memorization phase (Maren et al.1990)

The intelligent partitioning process allows the network to adapt to new data one it is being trained. The first segment, which includes the majority of the data, is used to train the model. In order to prevent the memorizing and overtraining effect in the neural network training process, a second segment of the data is taken for calibration that is blind to the neural network and at each step of training process, the network is tested for this set. If the updated network gives better predictions for the calibration set, it will replace the previous neural network; otherwise, the previous network is selected. Training will be continued once the error of predictions for both the calibration and training data set is satisfactory. This will be achieved only if the calibration and training partitions are showing similar statistical characteristics. Verification partition is the thirds and last segment used for the process that is kept out of training and calibration process and is used only to test the precision of the neural networks. Once a network is trained and calibrated, then the final model is applied to the verification set. If the results are satisfactory then the neural network is accepted as part of the entire prediction system. (Khazani,Mohaghegh 2012)

Key performance indicator process is performed to rank the most influential input parameters on shale gas production. In this process, not only the impact of each input parameters but also influence of a combination of different inputs on monthly gas production can be identified. Figure 1 show the comparison of sixteen different parameters influence.

Parametr	% Degreee of infulenc
K_f(md)	100
HF-length	70
Diffusion coeff.	58
HF_K(md)	55
HF_K*HF_Length	54
K_m(md)	35
Poro_F(%)	30
HF_Spacing	26
No.of Stages	26
Bottomhole press	21
Sigma factor	20
Longmuir Vol(Mscf/ton)	19
Poro_M (%)	18
HF_Height	17
Rock compaction	15
Sorption time	9

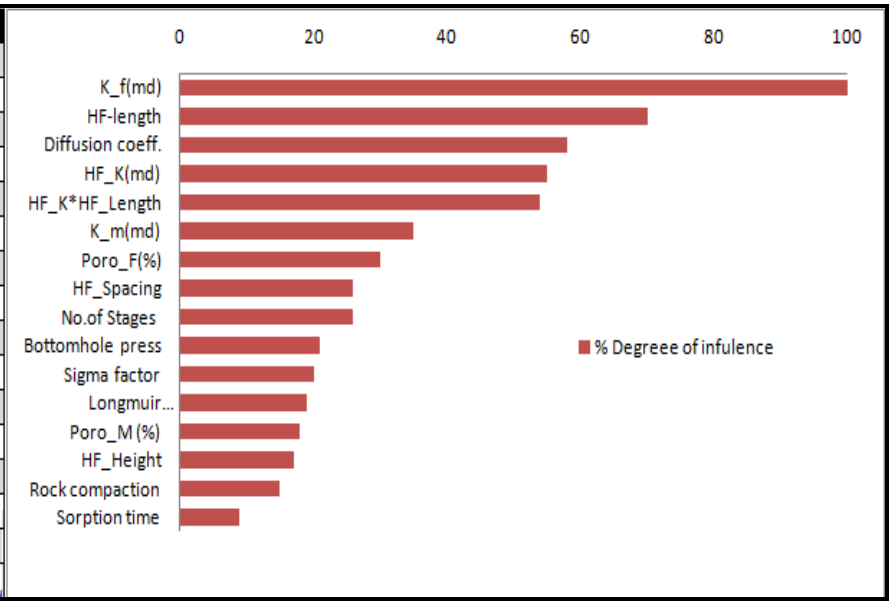

Figure 1-Comparison of the impact of different parameter on monthly gas rate of a shale well

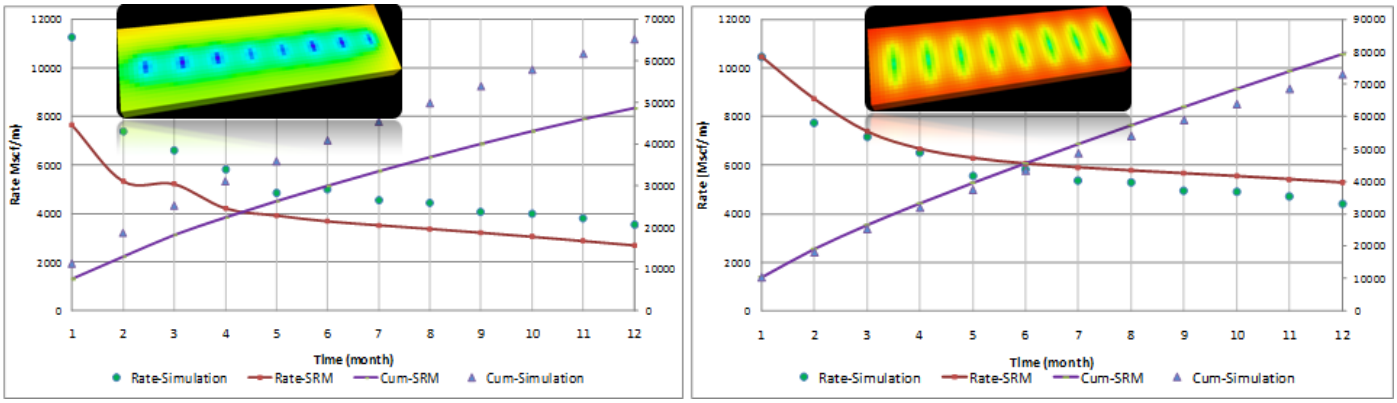

Figure 2-Monthly production rate generated by reservoir simulation and shale surrogate reservoir model

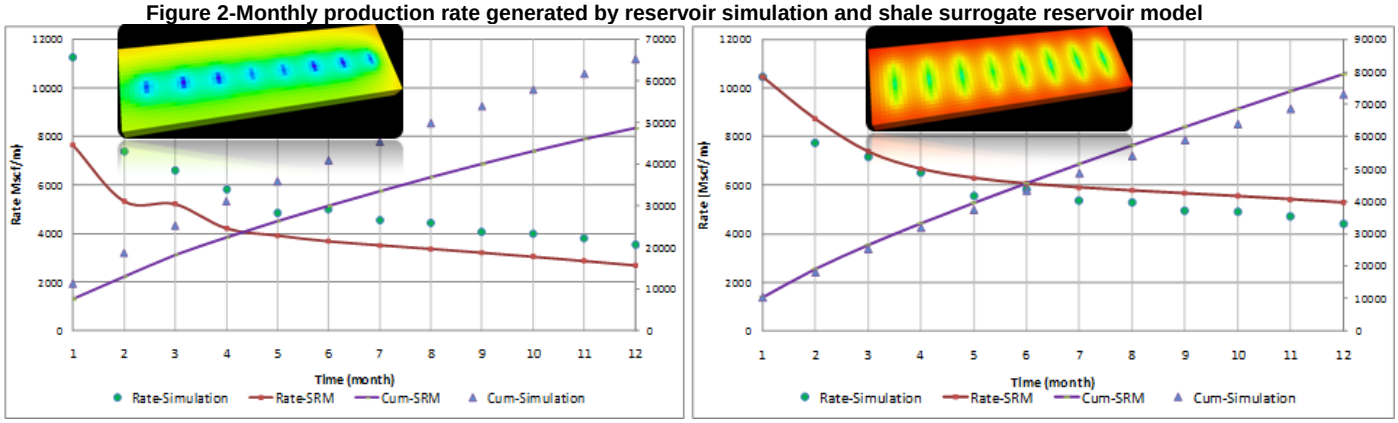

Figure 2 shows two examples of surrogate shale model results for the first scenario. In this graph, the green dots represent the monthly rates from reservoir simulation model while the red solid line shows the shale surrogate model results. The purple triangular points represent the cumulative production (reservoir simulation) while purple solid line is corresponding to cumulative production output by AI-based surrogate model.

As shown clearly in the aforementioned graph, approaching a non-linear problem linearly may not yield to a successful proxy model. Therefore, effective sampling and selecting robust design of experiment for developing spatio-temporal database is a key in successful development of AI-based proxy model.

In the second scenario, we generate a Latin Hypercube Sample (LHS) with a genetic algorithm by creating random permutations of the first *n* integers in each of *k* columns and then transforming those integers into *n* sections of a standard uniform distribution. Random values are then sampled from within each of the *n* sections. Once the sample is generated, the uniform sample from a column can be transformed to any distribution by using the quantile functions, e.g. *qnorm()*. Different columns can have different distributions.

S-optimality seeks to maximize the mean distance from each design point to all the other points in the design, so the points are as spread out as possible.

Genetic Algorithm was defined by "*Genetic_LHS*(*n, k, pop, gen, pMut*)" where:

n :The number of partitions (simulations or design points)

k :The number of replications (variables)

pop :The number of designs in the initial population

gen :The number of generations over which the algorithm is applied

pMut :The probability with which a mutation occurs in a column of the progeny

The systematic procedure for generating genetic LHS is briefly explained below and the schematic of result is shown in Figure 3. (The matrix's diagonal represents the parameters and the other Latin cubes shows the samples distribution in space based on the provided range for each parameter)

- Generate pop random Latin hypercube designs of size *n* by *k*
- Calculate the *S* optimality measure of each design
- Keep the best design in the first position and throwaway half of the rest of the population
- Take a random column out of the best matrix and place it in a random column of each of the other matrices, and take a random column out of each of the other matrices and put it in copies of the best matrix thereby causing the progeny
- For each of the progeny, cause a genetic mutation *pMut* percent of the time. The mutation is accomplished by switching two elements in a column

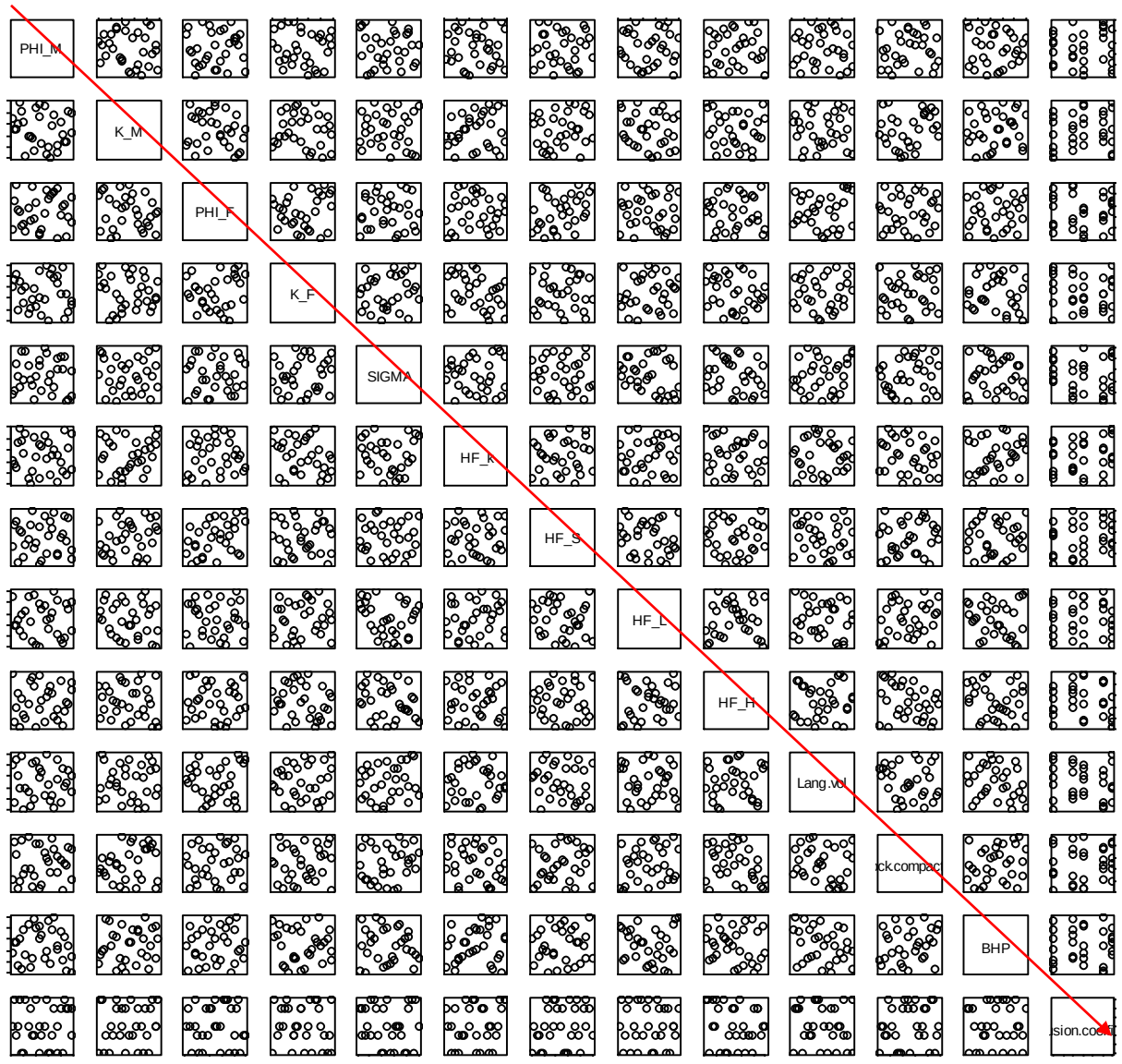

Figure 3- Latin cubes- LHS result by genetic algorithm for the parameters involve in shale modeling

In this approach, the spatio-temporal database is generated and the neural network training, calibration, and verification process is performed to develop a shale surrogate model with acceptable accuracy without simplifying and /or reducing the complexity of the problem, which is the case in development of many proxy model or response surfaces.

We come up with the optimum number of runs for generating a robust spatio-temporal data base which for our case is 31 runs. Upon completion of the training and calibration, the shale surrogate model has learned the intricate details of fluid flow in this reservoir therefore; it is capable of generalizing the simulation model's behavior and accurately replicating its results.

Two specific neural networks are trained, calibrated, and verified to predict the rate at the first month of production and the second one is developed for predicting the production rate for second to twelfth month.

After estimation of initial production rate for the first month, the production rate is modeled in a time-successive fashion. In other word, the production rate is predicated on the basis of previous production rate that is coming from the numerical simulator. The predicted initial rate by surrogate shale model is then used for predicting the second month of production rate and for the subsequent months, the same procedure is followed. This process is called cascading.

For both neural networks, data are partitioned with a 80% training fraction, 20% for calibration and verification (10% for each). The crossplots for predicted and simulated values of monthly flow rate (Mscf/m) for initial rate and second to twelfth months rate prediction by neural networks are shown in Figure 4.These plots show that the trained network also work very well for the blind data.

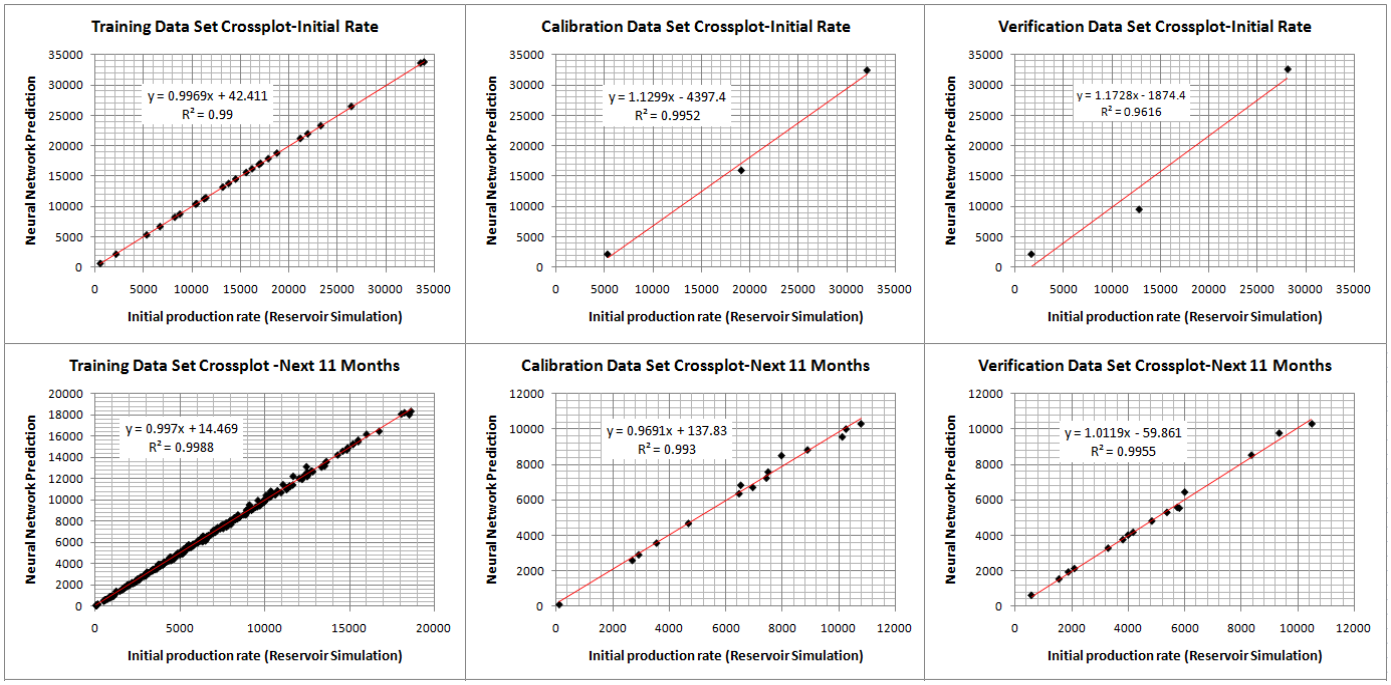

Figure 4-Initial rate (first row plots) and second to twelfth months (second row plots) rate prediction by neural networks

Example of shale surrogate model performance vs. the simulation runs for eight runs (12 months of production) that are used for training and calibration purposes is shown in Figure 5 through Figure 8. The accuracy of shale surrogate model results in replicating the results of the simulation runs are clearly demonstrated in these figures. In these graphs, the green dots represent the monthly rates from reservoir simulation model while the red solid line shows the shale surrogate model results. The purple triangular points represent the cumulative production from reservoir simulation while purple solid line is corresponding to cumulative production output by AI-based surrogate model.

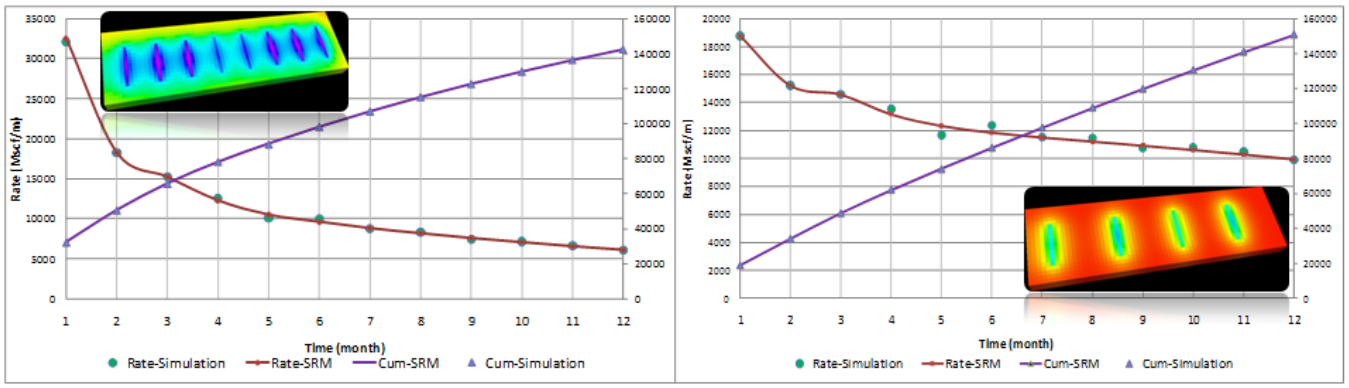

Figure 5- Comparison of shale surrogate model results with conventional simulator. (Gas rate and cum. production for 12 months)

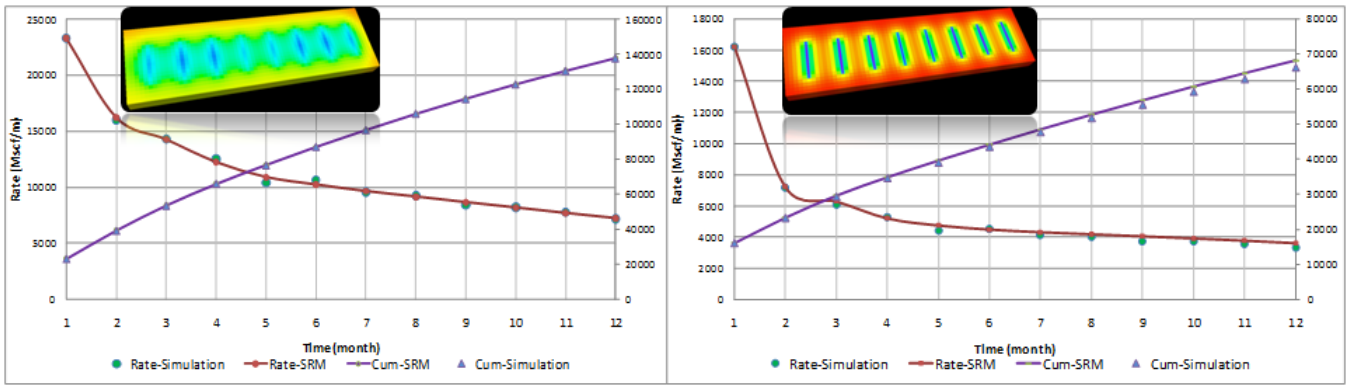

Figure 6- Comparison of shale surrogate model results with conventional simulator. (Gas rate and cum. production for 12 months)

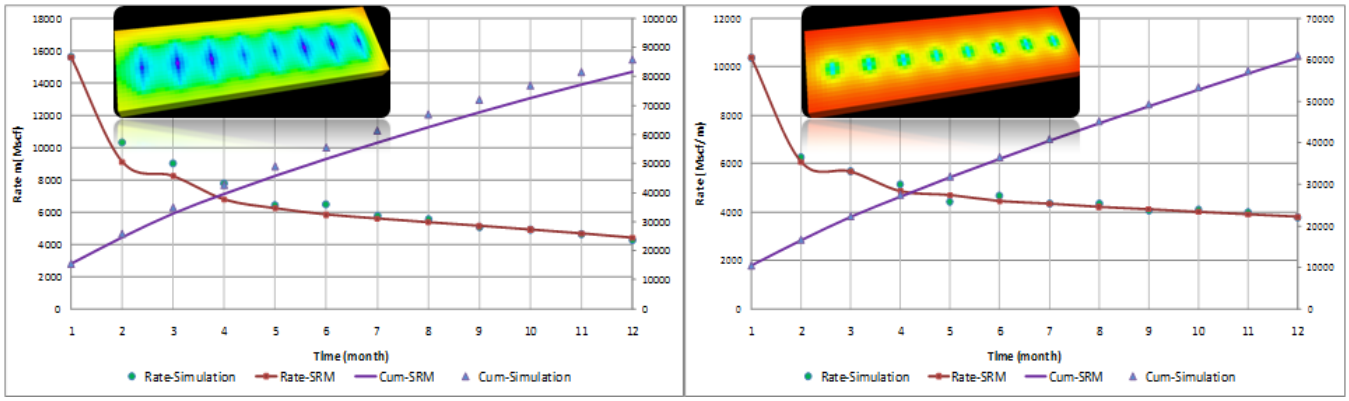

Figure 7- Comparison of shale surrogate model results with conventional simulator. (Gas rate and cum. production for 12 months)

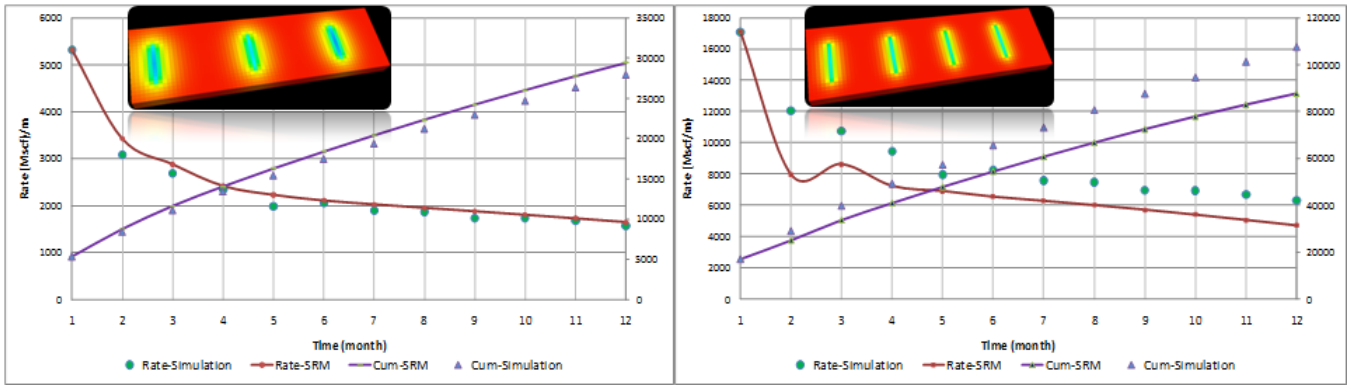

Figure 8- Comparison of shale surrogate model results with conventional simulator. (Gas rate and cum. production for 12 months)

As explained before, during shale surrogate model development some of the data are taken out from the data set for calibration, which is blind to the neural network, and at each step of training process, the network is tested for this set.

Taking validation one-step further, two simulation runs are designed in a way to be completely different (but in uncertainty range) from those 31 runs which are used for training, calibration and verification. In other word, these two runs are completely blind to neural network.

Figure **9** shows the result of Shale surrogate model performance vs. completely blind simulation runs to test the capability of AI-based proxy model in predictive mode.

Good results shows that the developed shale surrogate model can be effectively used as a tool for uncertainty quantification since thousands of shale surrogate model runs can be made in a few seconds to fulfill the requirement of the Monte Carlo simulation for uncertainty analysis.

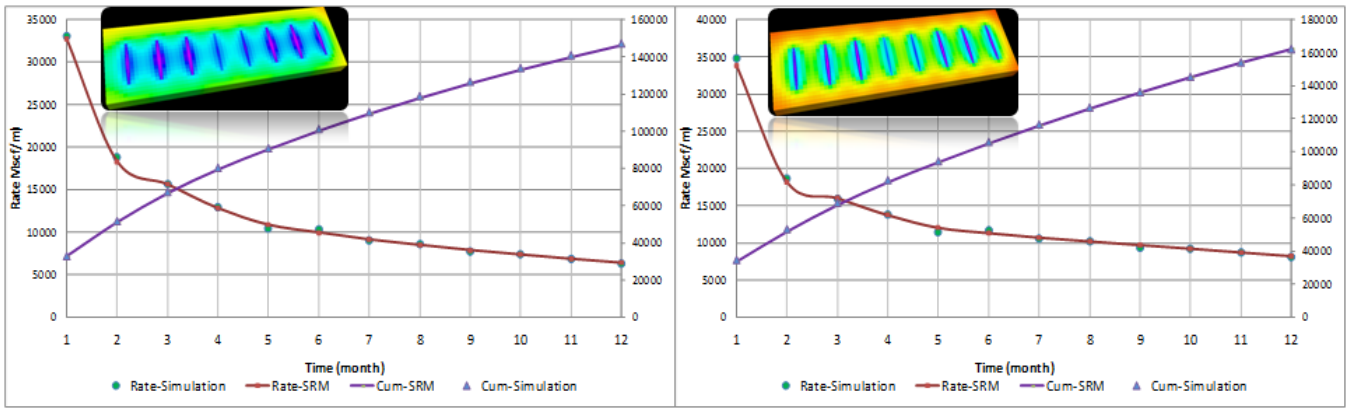

Figure 9- Shale surrogate reservoir model validation using two blind runs

Conclusion

A proof of concept study to evaluate the applicability of surrogate reservoir model in shale gas reservoir was conducted. Results of training, calibration, and validation of well-based shale surrogate reservoir model for a generic shale gas reservoir were presented. Shale surrogate model is accurate replicas of shale reservoir simulation models that run in fraction of a second and can serve as effective reservoir management tools. The validity of developed AI-based proxy model was continuously tested during the training, calibration process and also the model was validated by completely blind simulation runs successfully.

Acknowledgments

The Authors would like to thank US Department of Energy-NETL for partially funding this study (ICMI ARRA project, Subcontract No. RES1000026). We also thank and acknowledge Intelligent Solutions Inc. for providing IDEA software packages for the development of Surrogate Reservoir Models.

References

Andrade, J., Civan, F., Devegowda, D., Sigal, R.”Accurate Simulation of Shale Gas Reservoirs” SPE135564. Annual Technical Conference and Exhibition held in Florence, Italy, 19-22 September 2010.

Biswas,D.” Shale Gas Predictive Model (SGPM) – An Alternate Approach to Predict Shale Gas Production” SPE 148491, Eastern Regional Meeting held in Columbus, Ohio, USA, 17–19 August 2011.

Carnell, Rob 2011 “R-Statistical software- The Latin Hypercube Sample (lhs) Package” http://rgm2.lab.nig.ac.jp/RGM2/func.php?rd_id=lhs:lhs-package

Chauvin,Y.and Rumelhart,D.E.ed.”Backpropagation:Theory, Architectures, and Applications” London: Psychology press/Taylor & Francis Group1995

Cipolla, C.L. et al: “Modeling Well performance in Shale-Gas Reservoirs,” SPE 125532, SPE/WAGE Reservoir Characterization, and Simulation Conference held in Abu Dhabi, UAE, Oct. 19 – 21.2009

Haykin,S.”Neural Network, a Comprehensive Foundation” second edition, Upper Saddle River, NJ: Prentice Hall Inc.

Jurecka F.” Robust Design Optimization Based on Meta modeling Techniques” Technische Universität München 2007

Khazani,Y,. Mohaghegh S.D”Intelligent Production Modeling Using Fullfield Pattern Recognition” 2011 SPE Reservoir Vol.14, No.6 pp.735-749

Lee, W.J., Sidel R: "Gas Reserves Estimation in Resource Plays," paper SPE 130102 ,Unconventional Gas Conference, Pittsburgh, PA., 23-25 February 2010.

Lewis, R. et al.: “New Evaluation Techniques for Gas Shale Reservoirs,” Schlumberger paper for 2004 Reservoir Symposium.

Maren,A.J.,Harston,C.T.,and Pap,R.M" Handbook of neural computing applications "San Diego, California: Academic Press. 1990

Mohaghegh, S.D, Liu,J ,Gaskari,R., Meysami M.,Olukoko,O." Application of Well-Based Surrogate Reservoir Models (SRMs) to Two Offshore Fields in Saudi Arabia, Case Study" SPE 153845, Western North American Regional Meeting held in Bakersfield, California, USA, 19–23 March 2012.

Mohaghegh, S.D" Reservoir simulation and modeling based on artificial intelligence and data mining (AI&DM)" Journal of Natural Gas Science and Engineering, Elsevier, 3 (2011) p.697-705

Mongalvy, V., E. Chaput, S. Agarwal and L. Lu: "A New Methodology for Shale Reservoir Performance," paper SPE 144154 prepared for presentation at the 2011 Americas Unconventional Gas Conference held in the Woodlands, TX, 14-16 June.

McKay, M. D., Beckmann, R. J., and Conover, W. J. "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code" Technometrics, 21(2):239–245, 1979.

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/2012

Rubin, B: "Accurate Simulation of Non-Darcy Flow in Stimulated Fracture Shale Reservoirs," SPE 132293 prepared for presentation at the 2010 Western Regional Conference held in Anaheim, CA, May 27 - 29.

Stocki, R. "A method to improve design reliability using optimal Latin hypercube sampling" Computer Assisted Mechanics and Engineering Sciences 12, 87–105. 2005

Stein, M." Large Sample Properties of Simulations Using Latin Hypercube Sampling". Technometrics. Vol. 29, No.2, pp. 143–151,1987

Wang, J. and Y. Liu: "Well Performance Modeling in Eagle Ford Shale Oil Reservoir," SPE 144427, Americas Unconventional Gas Conference held in the Woodlands, TX, 14-16 June. 2011

Williams, G.J.J., Mansfield, M., MacDonald, D.G., and Bush, M.D. 2004. "Top-Down Reservoir Modeling" SPE 89974 presented at the 2004 SPE Annual Technical Conference and Exhibition, Houston, 26–29 September.

Williams, M.A. "Assessing Dynamic Reservoir Uncertainty: Integrating Experimental Design with Field Development Planning" 2006. SPE Distinguished Lecturer Series presentation given for Gulf Coast Section SPE, Houston, 23 March

Wilson, K. C., Durlofsky, L. J." Computational Optimization of Shale Resource Development Using Reduced-Physics Surrogate Models" SPE 152946, SPE Western Regional Meeting, 21-23 March 2012, Bakersfield, California, USA

Zubarev, D.I "Pros and Cons of Applying Proxy Models as a Substitute for Full Reservoir Simulations" SPE 124815, SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 4-7 October 2009.