Spring 4-10-2020

Subsurface Analytics: Contribution of Artificial Intelligence and Machine Learning to Reservoir Engineering, Reservoir Modeling, and Reservoir Management

Shahab D. Mohaghegh
West Virginia University, Shahab.Mohaghegh@mail.wvu.edu

PETROLEUM EXPLORATION AND DEVELOPMENT
Volume 47, Issue 2, April 2020
Online English edition of the Chinese language journal

Cite this article as: PETROL. EXPLOR. DEVELOP., 2020, 47(2): 225–228.

EDITORIAL

Subsurface analytics: Contribution of artificial intelligence and machine learning to reservoir engineering, reservoir modeling, and reservoir management

MOHAGHEGH Shahab D.*
Petroleum & Natural Gas Engineering, West Virginia University, Morgantown, WV 26506, USA

Introduction

Traditional Numerical Reservoir Simulation has been contributing to the oil and gas industry for decades. The current state of this technology is the result of decades of research and development by a large number of engineers and scientists. Starting in the late 1960s and early 1970s, advances in computer hardware along with development and adaptation of clever algorithms resulted in a paradigm shift in reservoir studies moving them from simplified analogs and analytical solution methods to more mathematically robust computational and numerical solution models.

The new computational paradigm overcame the mathematical limitations of the analytical solution methods. It introduced a more realistic solution when compared to the simple analog models such as CRMs (Capacitance-Resistance Modeling introduced to the oil industry in 1943 by W. A. Bruce)[1]. Complex, second order, non-linear partial differential equations that governs fluid flow in the porous media were solved numerically at speeds that were unthinkable just a few years before[2]. Today, the capabilities of this technology to model reservoirs is hardly contested. It is now a widely accepted technology among engineers and scientists in the oil and gas industry.

The foundation of the traditional numerical reservoir simulation technology is our current understanding of the physics of the storage and transport phenomena followed by our mathematical modeling capabilities. The complexities associated with the physics and the geology of the reservoirs being modeled determine the amount of compromise that is required during the modeling process. The application of traditional numerical reservoir simulation to unconventional resources such as shale is a good example of how much compromise is required during the modeling process. Compromises in the application of numerical reservoir simulation in unconventional resources such as shale plays appear in the form of gross assumptions and simplifications that in some cases make the entire process irrelevant, and at best, an academic exercise.

As far as the traditional numerical reservoir simulation is concerned, if our developed model does not match some of the field measurements such as hydrocarbon production or flowing bottom-hole pressure, then we modify other field measurements, such as reservoir characteristics, in order to achieve a history match. The question that has been rising in the past several years, asks: *Is it possible to accomplish this same objective with another set of techniques that minimize the amount of the required compromises or assumptions*. The objective being the incorporation of all field measurements representing the injection and production wells, reservoir (rock and fluid) characteristics, completion design and implementation, and operational constraints in order to build a robust reservoir simulation model that can accurately match the field's production history as well as forecasting its behavior. Subsurface Analytics that is defined as the application of Artificial Intelligence and Machine Learning (AI&ML) in reservoir engineering, reservoir modeling, and reservoir management is a new technology that provides a positive answer to the above question: *Yes, it is possible*.

1. Subsurface Analytics

Subsurface Analytics is a new technology that changes the way reservoir simulation and modeling is performed. Instead of starting with the construction of mathematical equations to model the physics of the fluid flow through porous media and then modification of the geological models in order to achieve history match, Subsurface Analytics that is a completely data-driven reservoir simulation and modeling technology[3] takes a completely different approach. In data-driven reservoir

Received date: 18 Feb. 2020; **Revised date:** 31 Mar. 2020.
*** Corresponding author.** E-mail: Shahab.Mohaghegh@mail.wvu.edu
https://doi.org/10.1016/S1876-3804(20)60041-6

modeling, field measurements form the foundation of the reservoir model. Using data-driven, pattern recognition technologies; the physics of the fluid flow through porous media is modeled through discovering the best, most appropriate relationships between all the measured data in a given reservoir.

In this approach in order to match the dynamic field measurements such as fluid (oil, gas, and water) productions, reservoir pressure, and water saturation, the interaction between all the field measurements such as reservoir characteristics, well placement and trajectory, completion details in space and time, operational constraints, etc. are modeled through discovery of the complex set of relationships and patterns between all the field measurements. The key characteristics of Subsurface Analytics are *(a)* No Interpretations, *(b)* No Assumptions, *(c)* No Complex Initial Geological Model, and therefore, *(d)* No Upscaling. Furthermore, it is important to note that the main series of dynamic variables that are used to build this model are measured on the surface (flow line and wellhead pressure and temperature, choke setting, as well as oil, gas and water production) while other major static (well logs, cores, seismic, etc.) and sometimes even dynamic (completions) characteristics are based on subsurface measurements. Using this combined series of surface and subsurface field measurements make this technology to be a coupled reservoir and wellbore simulation models rather than just a reservoir model.

The model history matches every individual well in the field. The "History Matching" process of the Subsurface Analytics is completely automated taking a small fraction of the time when compared with the history matching of the traditional numerical simulation models. Unlike history matching of the traditional numerical simulation models where local (well-based) modification of the model (measured reservoir characteristics, transmissibility, skin, etc.) plays a crucial role, especially for highly complex mature fields, the automatic history matching of Subsurface Analytics does not include any local modifications of the model.

Subsurface Analytics completely changes how mature fields are modeled. Deployment of this Coupled Reservoir-Wellbore Simulation (CRWS) modeling technology that is based on Artificial Intelligence and Machine Learning has a small computational footprint. The small computational footprint significantly contributes to many post modeling analyses such as single and multi-parameter sensitivity analysis, uncertainty quantifications, production and recovery optimizations, infill location optimization, injection optimization, and field development planning.

2. Subsurface Analytics vs. Traditional Numerical Reservoir Simulation

To build a reservoir simulation model based on our current understanding of the physics, we code computer programs to solve a set of mathematical formulations using numerical solution methods. During this process, we tell the computer through a set of pre-determined computer codes, exactly what to do with an amazing amount of systematic details. During this process, we can make a simple mistake and the code will fall apart. If the mistake is part of the coding language, for example in "C" or "C++", if we miss a semicolon (;), then hundreds of thousands of lines of code will collapse. If the mistake that we make is from an engineering point of view, for instance using the wrong unit system for a given variable, again, the code will give you completely non-sense answers. In other words, when you are building a numerical reservoir simulation model, *it is all about "YOU"* (and your team).

Numerical reservoir simulation and modeling has to do with your understanding of the physics of the fluid flow through porous media as well as your understanding of the geology of the field and the reservoir that you are modeling. Please do not forget that none of us has ever actually been down there and have seen a hydrocarbon reservoir that is thousands of feet under the ground. Furthermore, numerical reservoir simulation has to do with your experience of how many of such reservoirs you have modeled in the past, and how many times the models that you have developed helped the operator make correct decisions, or how many times it provided wrong information for decision making. Many times during the development of numerical reservoir simulation models, specifically for complex reservoirs, your potential success is a function of how many experts in geology, petro-physics, geo-physics, and reservoir engineering are involved in your team and how much time they spend together to come up with common agreements of what is happening thousands of feet under the ground.

Furthermore, we all know that one of the most important characteristics of building numerical reservoir simulation has to do with the deadline. The length of time that you and your team have been given to build and history match the numerical reservoir simulation model largely determines the quality of your model. If you have any doubt about this fact, ask actual reservoir modelers that are working for operating companies or service companies that have to meet deadlines for field development planning purposes. With all of these facts that were mentioned about the development of numerical reservoir simulation models, your team will take all its understandings of these details and communicates them with the computer through computer code/language. Your team will tell the computer what exactly needs to be done, systematic and with incredible amount of details.

Therefore, in one sentence, *traditional numerical reservoir simulation and modeling is all about you, and your understanding of the physics and geology*. In this technology, data is there, to serve your understanding of the physics and geology. You will be using a series of mathematical formulations that have already been developed to model the physics of fluid flow through porous media. These mathematical formulations determine what variables must be measured in order to provide a solution. Please keep in mind that any kind of biases and preconceived notions or assumptions that are made by anyone in the team of geo-scientists and reservoir engineers will influence the final solution and the final decision-making. Also, please do not forget that certain amounts and types of problem or solution simplifications are always required in order to make it possible to build numerical reservoir simulations. Therefore, you better have a team of highly educated and highly experienced geologists, petro-physicists, geo-physicists, and reservoir engineers in order to be able to trust the outcome of such exercises. These are facts, no matter how you choose to deal with them or how you end up defining them.

To build a reservoir models based on Artificial Intelligence and Machine Learning (Subsurface Analytics) the most important requirement is "Data", i.e. field measurements. The major differences between using Artificial Intelligence and Machine Learning to build a reservoir model and make decisions based on the model's outcome versus using traditional numerical reservoir simulation are the avoidance of human biases, preconceived notions, and any type of assumptions or problem/solution simplifications. When it comes to purely data-driven reservoir modeling, it is all about facts, field measurements, and data. You must let the data, the field measurements, to guide the solutions, not you. In other words, when you are building a data-driven reservoir simulation model, *it is all about "DATA"* (field measurements).

I hope this does not create confusion in people that truly and correctly believe that domain expertise (in this case reservoir engineering) is the most important characteristic of the team members that will be developing data-driven reservoir model. I have been preaching the critical importance of domain expertise in the application of AI and Machine Learning in our industry for more than 25 years [3]. However, recently it has become highly evident that when the domain experts that are in charge of developing AI-based reservoir models are not expert practitioners of Artificial Intelligence and Machine Learning and have superficial understanding of these new technologies, then they will end up developing so called "hybrid models".

To put it simplistically, the objective of the open "Machine Learning" algorithms that are used to build reservoir models are to explore and find patterns in the field measurements that must make physical sense to the scientist and engineers in the field. The patterns that the Machine Learning algorithms try to discover from large amounts of data should be able to accurately and simultaneously correlate all of the output dynamic variables with all other static and dynamic field measurements. The output of the AI-based reservoir models are oil, gas, and water production (as well as reservoir pressure and water saturation) from each well in the field, while the input to the model are a series of parameters that represent reservoir characteristics and identify how each of the wells were drilled, completed, and operated.

Obviously, this is not a simple endeavor. If you think that by learning the mathematics behind the open Machine Learning algorithms you can simply accomplish this task, then you will soon be disappointed. Apparently, such disappointments have been the main reason that some individuals and start-up companies have gone back to the mathematical formulations to get their models to make sense, and then end up calling them "hybrid models".

3. Hybrid Models

Recently, some individuals and/or start-ups have been using terminologies that to some people may sound like a new approach to modeling using Artificial Intelligence and Machine Learning. They have been using many names such as Hybrid Modeling, Physics-Based Data-Driven Modeling, Physics-based AI-models, Augmented AI Models, Data-Physics, etc. that seem to be referring to same type of thinking. It seems that all these approaches have been developed for the same following reasons:

a. Lack of ability to build a reasonable model using only Artificial Intelligence and Machine Learning, therefore using the traditional mathematical formulations in order to serve their modeling purpose,

b. Using Artificial Intelligence and Machine Learning as a marketing tool for their traditional approaches to modeling that now have a different interface,

c. Lack of ability to explain the results generated by models that are developed using Artificial Intelligence and Machine Learning,

d. Lack of ability to respond to the challenges put forward by the traditionalists in our industry,

When traditional mathematical equations are used in combination with the data-driven models that are developed by Artificial Intelligence and Machine Learning techniques, these models cannot be referred to as a new approach to engineering problem solving. These techniques are really the same as a series of techniques that have been used in our industry for decades. The fact is that in such cases, the data-driven approach becomes a purely statistical approach rather than an Artificial Intelligence and Machine Learning approach [4].

The main reason behind the fact that they cannot develop AI-based reservoir models that can justify and or explain the physics and geology is the statistical approach to problem solving. This means they are looking for correlation and use statistical regression approaches to deal with the data. This is a traditional statistical approach. It has no connection to physics and geology. That is why it cannot explain the causations behind the correlations. Therefore, in order to make their model sensible, these individuals and/or start-ups use the traditional mathematical formulations to serve their models, sometimes by generating data sometime by other approaches.

The main reasons for these approaches (if it is not all about marketing) is the lack of understanding of Artificial Intelligence and Machine Learning and how this technology can help scientists and engineers to solve physics-based problems.

We have been using traditional statistical algorithms in our industry since early 1960s (Arp's Decline Curve Analysis) and later as geo-statistics. That is why a very large number of engineers (and even academicians) around the world think that Artificial Intelligence and Machine Learning is the same as traditional statistics.

References

[1] BRUCE W A. An electrical device for analyzing oil reservoir behavior. AIME Transactions, 1943, 157: 112.

[2] AZIZ K and SETTARI A. Petroleum reservoir simulation. London: Applied Science Publishers Ltd., 1979.

[3] MOHAGHEGH S D. Data-driven reservoir modeling. USA: Society of Petroleum Engineers, ISBN: 978-1-61399-560-0, 2017.

[4] MOHAGHEGH S D. Traditional statistics vs. artificial intelligence and machine learning. (2019-11-11) [2020-1-11]. https://pubs.spe.org/en/dsde/dsde-article-detail-page/?art=6215.