Journal of Unconventional Oil and Gas Resources 15 (2016) 146–157

Contents lists available at ScienceDirect

Journal of Unconventional Oil and Gas Resources

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

Regular Articles

Determining the main drivers in hydrocarbon production from shale using advanced data-driven analytics – A case study in Marcellus shale

Shahab D. Mohaghegh

Intelligent Solutions, Inc., West Virginia University, United States

ARTICLE INFO

Article history:
Received 3 May 2016
Revised 16 July 2016
Accepted 28 July 2016
Available online 3 August 2016

Keywords:
Shale production optimization
Shale analytics
Data driven analytics
Data mining
Shale gas
Marcellus shale
Artificial intelligence

ABSTRACT

Is it the quality of the formation or the quality of the completion that determines or controls the productivity of a shale well? In this paper we attempt to address this important question. We present a case study using a fit-for-purpose approach with no attempt to generalize the final conclusions. The analysis presented in this article is based on field measurements. No assumptions are made regarding the physics of the storage and/or the transport phenomena in shale. Our objective is to let the data speak for itself.

The case study includes a large number of wells in a Marcellus shale asset in the northeast of the United States. Characteristics such as net thickness, porosity, water saturation, and TOC are used to qualitatively classify the formations surounding each well. Furthermore, wells are classified based on their productivity. We examine the hypothesis that reservoir quality has a positive correlation with the well productivity (wells completed in shale with better reservoir quality will demonstrate better productivity). The data from the field will either confirm or dispute this hypothesis.

If confirmed, then it may be concluded that completion practices have not harmed the productivity and are, in general, in harmony with the reservoir characteristics. The next step in the analysis is to determine the dominant trends in the completion and judge them as best practices. However, if and when the hypothesis is disproved (wells completed in shale with better reservoir quality will NOT demonstrate better productivity), one can and should conclude that completion practices are the main culprit for the lack of better production from better quality shale. In this case, analysis of the dominant trends in the completion practices should be regarded as identifying the practices that need to be modified.

Results of this study show that production from shale challenges many of our preconceived notions. It shows that the impact of completion practices in low quality shale are quite different from those of higher quality shale. In other words, completion practices that results in good production in low quality shale are not necessarily just as good for higher quality shale. Results of this study will clearly demonstrate that when it comes to completion practices in shale, "One-Size-fit-All" is a poor prescription.

Published by Elsevier Ltd.

1. Introduction

The conventional wisdom developed over several decades in the oil and gas industry states that better quality rocks produce more hydrocarbon. In other words there is a positive correlation between reservoir characteristics and production as depicted by the blue line in Fig. 1. Since production from shale wells is the result of significant human intervention (in the form of long laterals with a large number of hydraulic fractures), many operators started asking a question that used to be considered as the ground truth. The question is directed toward the impact of reservoir characteristics (rock quality) and its relationship with completion practices.

At the first glance it may seem that such question should be easy to answer. If the answer is not quite obvious from the operations (which one will quickly realize that it is not – please see Figs. 16 and 17 at the end of this article as examples), then we can refer to our models for the answer. The procedure should not be very complicated. In our models, we can keep the completion and hydraulic fracturing characteristics constant and change the reservoir characteristics and observe its impact on production and then answer the above question. It sounds pretty simple and straight forward, until one realizes that such models (capable of realistically addressing questions such as this) do not exist for shale.

In other words, the formulations that are currently used to model fluid flow (and therefore production) in shale,[1] does not

[1] This includes analytical and/or numerical solutions to the fluid flow equations that need to take into account the propagation of induced fracture in shale, its interaction with the natural fracture system, and many other nuances that are inherent in production from shale.

E-mail address: Shahab.Mohaghegh@mail.wvu.edu

http://dx.doi.org/10.1016/j.juogr.2016.07.004
2213-3976/Published by Elsevier Ltd.

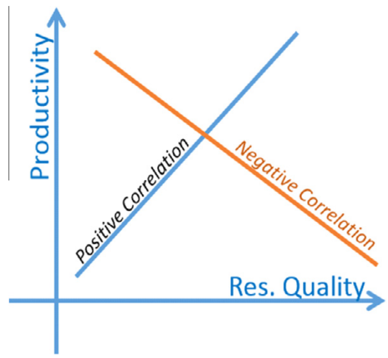

Fig. 1. Conventional Wisdom: Productivity in a well increases with reservoir quality.

really represent what is happening, and therefore, scientists and engineers cannot fully trust the results generated by these models. This is true at multiple levels, including the modeling of the storage, the transport of the fluids, and the propagation of the induced fractures. A comprehensive and critical review of the state of reservoir modeling in shale has already been published (Mohaghegh, 2013) and therefore, it will not be repeated here.

In this article, to answer the question posed earlier for a given asset (there is no claim that the results shown in this article are general in nature. We recommend similar study be applied to each field), we will only refer to actual field measurements, or as we call them "Hard Data". Hard data is defined as field measurements such as inclination, azimuth, well logs (gamma ray, density, sonic, etc.), lateral and stage lengths, number of clusters per stage, fluid type and amount, proppant type and amount, ISIP, breakdown and closure pressures, and corresponding injection rates, etc. As far as the reservoir characteristics are concerned, we use measurements such as net pay thickness, porosity, gas saturation and TOC to define rock quality. Furthermore, we use pressure corrected production as indicator of productivity. Furthermore, as part of our Advanced Data Driven Analytics technology, we introduce Supervised Fuzzy Cluster Analysis (SFCA) that is used to perform and to reach the conclusions in this study.

2. Methodology

To explain how we carried out this analysis we first need to briefly introduce two very simple ideas. The first idea is called Supervised Fuzzy Cluster Analysis (SFCA), and the second idea is the use of SFCA to classify shale qualities, in a straight forward and non-controversial manner.

Fuzzy Cluster Analysis (Bezdek, 1984) that is an implementation of Fuzzy Set Theory (Zadeh, 1965) in cluster analysis was introduced several years ago. In this study we have modified the original algorithm such that engineers and geo-scientists with domain expertise can define the location of the cluster centers (shale quality). This is a simple but very important modification to the Fuzzy Cluster Analysis algorithm[2] in order to accommodate the type of analysis that is presented here. Again, the objective of this analysis is to answer a specific question regarding the importance and the influence of reservoir quality on production in shale basins. As the reader will note, this study would not have been possible without making this modification to the classic Fuzzy Cluster Analysis algorithm.

Cluster analysis, by nature is an unsupervised process. It aims at discovering order and patterns in seemingly chaotic, hyper-dimensional data. The modification to this algorithm is based on a simple observation that allows us to impose certain domain expertise into our purely data driven analysis[3]. In other words, we attempt to address a common observation by engineers and geo-scientist when they are exposed to the data-driven analytics. Since we do know certain underlying physics regarding the shale quality, we will guide (supervise) our analysis of the data in such a way that it can identify the relative quality of the shale based on its measured reservoir characteristics. For example, if I can distinguish between "Good" and "Poor" rock qualities, I would like to learn to what degree the formation encompassing each of my wells are represented by each of these semantics.

As was mentioned in the beginning of this section, the second simple idea has to do with judging the quality of the rock (shale), based on measured parameters. Since calculation of reserves in shale still is an ongoing topic of research, in order to be on the safe side and make the results of this study acceptable by engineers and scientists of all persuasions, we will not use any formulation to calculate reserves (as a proxy for reservoir quality) in shale. Instead, we will try to identify characteristics that is acceptable by almost anyone that has any background in reserve calculation of any type of formation, including shale. The rules of distinction between "Good" and "Poor" rock qualities will be based on simple observations, such as the following, (everything else being equal):

1. Formations with higher values of Net Pay Thickness should have more hydrocarbon reserves than formations with lower values of Net Pay Thickness.
2. Formations with higher values of Porosity should have more hydrocarbon reserves than formations with lower values of Porosity.
3. Formations with higher values of Hydrocarbon Saturation should have more hydrocarbon reserves than formations with lower values of Hydrocarbon Saturation.
4. Formations with higher values of TOC should have more hydrocarbon reserves than formations with lower values of TOC.

2.1. Supervised Fuzzy Cluster Analysis (SFCA)

In conventional cluster analysis, as shown in Fig. 2, clusters are separated by crisp boundaries. In this figure the two data points that are identified by red crosses belong to cluster "A", and do not have membership in cluster "B". In this figure cluster centers are identified by brown circles. In Fig. 2 both identified data points have a membership of "1" in cluster "A" and a membership of "0" in cluster "B".

If Fig. 2 was not observable (for example instead of two, it was part of a hyper-dimensional dataset that could not be plotted for observation) and you would only be exposed to the algorithm output, then you would assume that these two points are quite similar. For example, if the cluster centers were representative of rock qualities (A = Good Shale and B = Poor Shale), both these wells were completed in "Good" quality shale. However, the reality, as presented in Fig. 2 is quite different from this interpretation.

When the idea of Fuzzy Sets is introduced and Fuzzy Cluster Analysis is used to identify order in this data, as shown in Fig. 3, the first data point (the well represented with the red cross on the left) has a membership of "0.95" in cluster "A" and a membership of "0.05" in cluster "B", while the second data point (the well represented with the red cross on the right) has a membership of "0.55" in cluster "A" (A = Good Shale) a membership of "0.45" in

[2] This is a new algorithm as part of the advanced data driven analytics algorithm developed by Intelligent Solutions, Inc. (Intelligent Solutions, 2015).

[3] As you will notice, the domain expertise we refer to here, is far from being bias, or based on assumptions, or interpretation of the data.