Oilwell Completion Regression - Literature Review

Thomas Wilson

September 9, 2018

Introduction

For centuries, oilwells have been drilled vertically and completed by perforation only. Since the invention of horizontal drilling, literature as shown clearly that various completion parameters have a significant effect on oil production.

Body

Advances in horizontal completion design are evolving. (Fulks et al. 2016)

Big data has value. (Fulks et al. 2016)

There is a trend in industry towards using statistical modeling and machine learning to optimize reservoir production. (Fulks et al. 2016)

(Holdaway, Laing, and others 2015) have demonstrated how to achieve data-driven insights in planning, design, and operation of oilwell drilling.

Marrying data-driven analytics with first principles converts raw data into actionable knowledge. (Holdaway, Laing, and others 2015)

Proxy modeling is not a new concept and had been in the industry for many years.

Data Quality

General Workflow

Data-driven workflows, models, and analysis can address a diverse array of business problems in the oil and gas industry. (Holdaway, Laing, and others 2015)

Parallel Coordinates plots are a good visual analysis for testing any input's impact on production performance. (Groulx et al. 2017)

The number of performance measures directly correlate with the number of patterns identified. (Groulx et al. 2017)

With enough statistical variability, the same approach finds patterns in all plays. (Groulx et al. 2017)

Parallel coordinates approach makes identification of thresholds and correlation windows easy. Which can be valuable input for other regression efforts. (Groulx et al. 2017)

The predictive proxy approach has a wide variety of applications to completion engineering and management. (Pankaj et al. 2018)

Drilling and lifting in addition to completion also benefit from faster decision making. (Pankaj et al. 2018)

Feature Importance

High proppant, high-fluid completion designs described by (Fulks et al. 2016) have shown success every basin.

degradable diversion improve cluster efficiency.(Fulks et al. 2016)

A 3 dimensional earth model determines the optimal lateral landing zone. (Fulks et al. 2016)

Large amounts of data from multiple wells as input to multivariate analysis determines the factors driving production. (Fulks et al. 2016)

In order of importance fracture half-length, proppant amount, zone coverage, and slurry volume are important for predicting production. (Temizel et al. 2015)

data analytics makes it possible to determine not only feature importance and significance, but also effect direction. (Temizel et al. 2015)

Simulation

Data analytics can replace fracture models built with a commercial simulators. (Temizel et al. 2015)

The approach described in (Pankaj et al. 2018) creates synthetic data from first principles.

(Pankaj et al. 2018) uses parametric explosion of parameters to create a model which predicts performance in unconventional reservoirs.

Engineers use a model to design completions and predict production in real-time. (Pankaj et al. 2018)

Proxy modeling indicates the optimal well completion much faster than full simulations, a matter of minutes instead of months.(Pankaj et al. 2018)

Model Limitations

In the absence of data proppant, fluid, and diversion are important for establishing baseline performance. (Fulks et al. 2016)

Conclusions from data driven models are specific to the model used.(Temizel et al. 2015)

Careful consideration is necessary for normalizing both performance measures and inputs. (Groulx et al. 2017)

Before discernible conclusions, patterns detected must be reviewed to refine insights. (Groulx et al. 2017)

"A calibrated model is a fundamental step to create a reliable predictive proxy." (Pankaj et al. 2018)

The rapid proliferation of sensing and data acquisition systems combined with cheap data storage is resulting in the age of "Big Data" and the most scarce resource going forward is likely to be expert time to go through, clean and label the data.(Subrahmanya et al. 2014)

This has led to the rising importance of adopting new strategies and workflows for the development of advanced analytics models.(Subrahmanya et al. 2014)

This trend is already visible in other industries and is likely to be a big challenge for the oil and gas industry as well. (Subrahmanya et al. 2014)

In this work, we explored the use of active learning to intelligently identify data points with the highest value of information and semi-supervised learning to combine the information from labeled and unlabeled sources in an optimal fashion. (Subrahmanya et al. 2014)

It was demonstrated that these advanced machine learning methods show considerable promise and can be integral tools in making the entire analysis procedure more interactive and productive. (Subrahmanya et al. 2014)

Effective adoption of these advanced machine learning techniques for interactive learning will be dependent on their integration with excellent data visualization and effective user interface design.(Subrahmanya et al. 2014)

A framework of big data analytics is developed to measure the efficacy factor of fracturing fluid ingredients in terms their effectiveness on well productivity, based on FracFocus and well production databases. (Luo 2018)

Such efficacy factors could serve as justification for chemical product approvals and denials.(Luo 2018)

As a useful demonstration, chemical ingredients that are conducive to production of wells with Marcellus shale geology are quantitatively differentiated those that negatively impact gas productivity.(Luo 2018)

The specific efficacy factors may not be directly applicable in other fields, but the framework of data analysis can be easily applied.(Luo 2018)

The data analytics framework helps providing leads in understanding the intricate chemical and physical processes of fracturing fluids.(Luo 2018)

It points the direction to the causality - those kinds of cause-effect relationships – between ingredient and productivity straight from real-world data. (Luo 2018)

Each well is a geophysical experiment. (Luo 2018)

Compared with fracturing fluid lab test that still needs to be optimized in the field, the data analytics framework learns hidden connections from field data, which is the ultimate experiment, and reduces or stops the guesswork. (Luo 2018)

The horizontal wells with multi-stage hydraulic fracturing seem to be more preferable than vertical wells with hydraulic fracturing treatment in the Field C, reservoir. Bozoev and Demidova (2016)]

The technology (Frack ports or Ball dropping process) of multi-stage hydraulic fracturing is applied in the Field C, reservoir. Bozoev and Demidova (2016)]

Analytical approaches do not take into account the interference effect. Bozoev and Demidova (2016)]

However, Economices method should be used with caution, despite the fact that the possible wells with interference effect are observed.Bozoev and Demidova (2016)]

4. The simulation model tends to be an appropriate approach to choosing the optimum completion of the wellbore.(Bozoev and Demidova 2016)

In this paper we proposed three types of data sets as input features and investigated the use of deep learning for the classification and prediction of two-phase flow, based on experimental data, obtained . (Ezzatabadipour et al. 2017)

We proposed six types of input features, and a corresponding architecture to precisely predict flow patterns. (Ezzatabadipour et al. 2017)

First, we showed that the network can learn surprisingly well as using our chosen architecture and parameters allowed us to achieve high classification accuracy.(Ezzatabadipour et al. 2017)

Second, we showed that the network can classify the different flow patterns with high efficiency. (Ezzatabadipour et al. 2017)

Finally, we achieved high precision predicting different combinations of classes. (Ezzatabadipour et al. 2017)

Our experiments indicate that a deep learning approach, has the potential to capture flow patterns, which may boost the classification performance. (Ezzatabadipour et al. 2017)

These investigations could be further improved in future studies by carrying out more exhaustive searches for the parameters in the architectures.(Ezzatabadipour et al. 2017)

The result would be improved overall performance of these systems.(Ezzatabadipour et al. 2017)

Finally, deep learning can be used to predict flow patterns using pipe characteristics, fluid properties and superficial velocities of the two-phase flows.(Ezzatabadipour et al. 2017)

Oil & Gas industry demands effective methods to complete its Industry 4.0 digital transformation.(Khodabakhsh, Ari, and Bakir 2017)

This leads them to resort to open–source big data tools.(Khodabakhsh, Ari, and Bakir 2017)

In this paper, we proposed an architecture to solve big data problems seen in oil refinery sensor data and implemented the architecture in private cloud utilizing open–source software.(Khodabakhsh, Ari, and Bakir 2017)

The results are applicable to oil drilling industry as well.(Khodabakhsh, Ari, and Bakir 2017)

We developed real–time analytical models for detecting and classifying gross errors.(Khodabakhsh, Ari, and Bakir 2017)

Our contributions are enablers of other stream mining algorithms and provisioning of "Refinery as a Service".(Khodabakhsh, Ari, and Bakir 2017)

In the future, we plan to continue and complement our current work with other machine learning and deep learning. (Khodabakhsh, Ari, and Bakir 2017)

The workflow we presented in this paper is a systematic methodology for using and comparing machine learning methods to predict production in unconventional plays using well logs and production data from previous explorations, taking into account multiple formations. (Guevara et al. 2017)

Our experimental results show that this workflow can outperform more conventional techniques, such as kriging, in particular for the case of gas production prediction. (Guevara et al. 2017)

As future work, we would like to add to our workflow the capability of extracting and integrating features from horizontal well logs, in additional to features from vertical well logs.(Guevara et al. 2017)

We also expect to be able to add well completion parameters in the models, which we expect will be very important in terms for improving the accuracy of the production predictions.(Guevara et al. 2017)

Different techniques have been used to build models that predict and estimate various geophysical properties, by means of either regression or classification tasks including but not limited to: deterministic petrophysical modeling, using shale, matrix, and fluid properties; stochastic modeling, where an approximate curve is used as input, and the reconstructed curve is the output; and different soft-computing algorithms.(Lopes and Jorge 2017)

These approaches focus on a single well or on a few wells from the same block.(Lopes and Jorge 2017)

In this work we compare the results of using Generalised Linear Models (OLS), Bayesian Regression (BRR), RANdom SAmple Consensus (RANSAC), Random Forests (RF), and Artificial Neural Networks (ANN), on the prediction of missing gaps in a single well.(Lopes and Jorge 2017)

For that purpose we chose a single well with complete logs (without gaps).(Lopes and Jorge 2017)

From these logs we generated 30 random gaps for each gap size of the first three quartiles (respectively 16, 66, and 168 points for the first, second, and third quartiles). (Lopes and Jorge 2017)

These will be used to average the models' performance, since it depends not only on the size of the gap but also on the values of the gap itself. (Lopes and Jorge 2017) The results described in the following paragraphs were obtained with vanilla implementations of the afore mentioned algorithms, provided with the Scikit-Learn

[PVG + 11] Python machine learning library. (Lopes and Jorge 2017) The distribution of the models' errors (mean absolute error) over the thirty random gaps for each gap size can be observed in Figure 3. (Lopes and Jorge 2017)
It shows clearly that the performance of the algorithms depends on the gap size and on the gap itself, given the variance displayed. (Lopes and Jorge 2017)
In general the ANNs seems to perform better more often as the gap size increases. (Lopes and Jorge 2017) For the smaller gaps the remaining methods yield better predictions. (Lopes and Jorge 2017) However, the statistical difference is not significant, and both RFs and ANNs are computationally more expensive than the remaining regressions. (Lopes and Jorge 2017) Plotted in Figure 4 one can see an example gap for each of the quartiles with the corresponding predictions for each algorithm, and the mean absolute error in the legend (eg., OLS: 0.01) (Lopes and Jorge 2017)

In this paper an overview of data mining applications in oil and gas exploration was carried out. Two major categories of geoscientific problems, which have been the focus of data mining efforts, are studied: structural geology and reservoir property-issues. Table 1 presents these categories and provides the key information deduced on them. (Nikhalat-Jahromi and Jorge 2017) There are two ways of expanding this paper: (i) There is another category of geoscientific problems in which data mining is applied. (Nikhalat-Jahromi and Jorge 2017) Here, an array of ancillary data which provide high-level information in hydrocarbon exploration are covered. (Nikhalat-Jahromi and Jorge 2017) Remote sensing (including satellite infrared, radar, and microwave surveying; and a serial magnetic, electromagnetic, and gravity surveying) is the most important topic in this category. (Nikhalat-Jahromi and Jorge 2017) Although there are cases of data mining application to geobotany prospecting and geochemical exploration as well. (Nikhalat-Jahromi and Jorge 2017) In remote sensing, swaths of earth are covered, resulting in an understanding of the geological megastructures, with the aim of narrowing down to locations with the potential for hydrocarbon. (Nikhalat-Jahromi and Jorge 2017) After that further survey ing in these places provides precise details on the subsurface formations. (Nikhalat-Jahromi and Jorge 2017) This category of problems, although less popular in academia than those two considered in this paper, is interesting to be overviewed too. (Nikhalat-Jahromi and Jorge 2017) (ii) This overview given its novel way of categorizing data mining applica tions in oil and gas exploration is worthy of being expanded to a comprehensive review. (Nikhalat-Jahromi and Jorge 2017) Such a review will be informative to both geoscientists and data miners and deepen their mutual understanding. (Nikhalat-Jahromi and Jorge 2017)

This paper has analyzed the uncertainties in an integral plume model simulating the Deepwater Horizon oil spill using Polynomial Chaos surrogates. (Wang et al. 2016)
The study quantified the impacts of uncertainties in six input parameters on the model's estimates of the trap and peel heights, and the gas mass fluxes. (Wang et al. 2016)
The input uncer- tainties were primarily caused by missing data, and the available information about the input variables was used to define reasonable uncertainty ranges for the parameters. (Wang et al. 2016)
The major computational challenges dur- ing the surrogate construction were the sampling of a large six-dimensional parameter space, and handling the model noise. (Wang et al. 2016)
Both problems were successfully addressed using BPDN to calculate the PC coefficients using only 200 model samples. (Wang et al. 2016)
Various error metrics suggest that the PC surrogate is able to predict the response surface and the statistical information with acceptable fidelity. (Wang et al. 2016)
The PC surrogates were used to estimate the PDFs of the different model outputs. (Wang et al. 2016)
Comparisons with fluo- rescence observations, a proxy for the oil accumulation at the trap height, show the observations coinciding with the mode of the trap height PDF for the LUE, thus suggesting that the latter range to be quite consist- ent with the fluorescence observations. (Wang et al. 2016)
The HUE case showed a lower probability of occurrence for the observations. (Wang et al. 2016)
Additionally, the natural gases do not dissolve completely at the trap height under any sce- nario, and therefore, the next stage numerical models must be able to handle this mixture of oil and gas in order to simulate the oil fate realistically. (Wang et al. 2016)
The sensitivity analysis of the trap height, peel height, and gas mass fluxes, indicates that the huge uncertainty range for the flow rate leads to substantial uncertainties in the ensuing quantities of interest whereas constraining the range to reasonable values shifts the dominance to the uncertainty in the largest

droplet size and to uncertainties in the entrainment coefficients. (Wang et al. 2016)
The present investigation suggests that the model outputs are insensitive to uncertainties in the flow rate provided it is known reasonably accurately. (Wang et al. 2016)
It also suggests that the next priority in observation should be aimed at estimating the largest droplet size and GOR if improvements in mass gas flux estimates are desired, whereas the priority would fall on measuring the largest droplet size and the entrainment parame- ters if improvement in trap and peel heights is desired. (Wang et al. 2016)
If direct observations of these quantities prove to be difficult, then the collection of alternate data can still be useful as it can be used in an inverse uncertainty propagation exercise [Sraj et al ., 2013] to correct the PDF of the model's input parameters. (Wang et al. 2016)
Likewise, if obser- vations prove to be difficult or impractical, an additional compromise can be made, such as resorting to simulated data from very high-resolution numerical simulations [Fabregat et al ., 2015]. (Wang et al. 2016)
(Wang et al. 2016)
This paper presents an oil well productivity computation method based on a brain-inspired cognitive architecture. (Yu et al. 2018) The architecture, consisting of two interacting sensorimotor loops, realizes prediction and other cognitive functions through the internal simulation of the interaction with the external environment. (Yu et al. 2018) In the proposed method, the IPR parameters were fitted in the inner loop. (Yu et al. 2018) The fitting results were considered in the productivity computing in the outer loop after achieving the threshold value. (Yu et al. 2018) The brain-inspired productivity computation model fully reflects the dynamic production features of oil wells, thanks to the integration of the whole spectrum of production data. (Yu et al. 2018) With a high prediction accuracy, this model can guide the design of production engineering project, well production optimization, and well performance analysis. (Yu et al. 2018)

the results presented in this study reveal that in the early hydration stages of a cement slurry, the transmission of hydrostatic pressure (in a fresh cement slurry) follows Pascal's law. (Liu et al. 2018) However, because of the sedimentation of cement particles, the hydrostatic pressure of the slurry can only be transferred by the pore solution and the hydrostatic pressure of cement slurry is reduced. (Liu et al. 2018) With an increase in the hydration time, the hydration reaction is accelerated and some hydration products are formed between the cement particles. (Liu et al. 2018) Because of the formation of such products, some of which are porous, some macro-pores are filled with hydration products to form a number of micropores, which leads to most of the free water turning into capillary water and gel water. (Liu et al. 2018) The interaction between capillary water, gel water, and micropores is stronger than that between free water and macro-pores, which leads to a reduction in the hydrostatic pressure of the cement slurry. (Liu et al. 2018) Some hydration products are formed in the pores of the cement slurry, due to which the pores are filled. (Liu et al. 2018) This also leads to a reduction in the height of the hydrostatic column and hydrostatic pressure of the cement slurry. (Liu et al. 2018) (Liu et al. 2018)

Conclusion

Bozoev, A M, and E A Demidova. 2016. "Selection of the Optimal Completion of Horizontal Wells with Multi-Stage Hydraulic Fracturing of the Low-Permeable Formation, Field c." *IOP Conference Series: Earth and Environmental Science* 33 (1): 012035. http://stacks.iop.org/1755-1315/33/i=1/a=012035.

Ezzatabadipour, Mohammadmehdi, Parth Singh, Melvin Deloyd Robinson, Pablo Guillen-Rondon, and Carlos Torres. 2017. "Deep Learning as a Tool to Predict Flow Patterns in Two-Phase Flow." *CoRR* abs/1705.07117. http://arxiv.org/abs/1705.07117.

Fulks, Robert, Simon Hughes, Ingo Geldmacher, and others. 2016. "Optimizing Unconventional Completions-an Integrated Approach." In *Abu Dhabi International Petroleum Exhibition & Conference*. Society of

Petroleum Engineers.

Groulx, Bertrand, Jim Gouveia, Don Chenery, and others. 2017. "Multivariate Analysis Using Advanced Probabilistic Techniques for Completion Design Optimization." In *SPE Unconventional Resources Conference*. Society of Petroleum Engineers.

Guevara, Jorge, Matthias Kormaksson, Bianca Zadrozny, Ligang Lu, John Tolle, Tyler Croft, Mingqi Wu, Jan Limbeck, and Detlef Hohl. 2017. "A Data-Driven Workflow for Predicting Horizontal Well Production Using Vertical Well Logs." *CoRR* abs/1705.06556. http://arxiv.org/abs/1705.06556.

Holdaway, Keith R, Moray L Laing, and others. 2015. "Drilling and Completion Optimization in Unconventional Reservoirs with Data-Driven Models." In *SPE Asia Pacific Unconventional Resources Conference and Exhibition*. Society of Petroleum Engineers.

Khodabakhsh, Athar, Ismail Ari, and Mustafa Bakir. 2017. "Cloud-Based Fault Detection and Classification for Oil & Gas Industry." *CoRR* abs/1705.04583. http://arxiv.org/abs/1705.04583.

Liu, Kaiqiang1, chengxw@swpu.edu.cnzj-656@163.com Cheng Xiaowei1, Xingguo2 Zhang, Zaoyuan2 Li, Jia1 Zhuang, and Xiaoyang2 Guo. 2018. "Relationship Between the Microstructure/Pore Structure of Oil-Well Cement and Hydrostatic Pressure." *Transport in Porous Media* 124 (2): 463–78. http://ezproxy.uhd.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=130935522&site=eds-live&scope=site.

Lopes, Rui L., and Alípio Jorge. 2017. "Mind the Gap: A Well Log Data Analysis." *CoRR* abs/1705.03669. http://arxiv.org/abs/1705.03669.

Luo, & Zhang, H. 2018. "Mining Fracfocus and Production Data for Efficacy of Fracturing Fluid Formulations." In *Society of Petroleum Engineers*. Society of Petroleum Engineers.

Nikhalat-Jahromi, Hamed, and Alípio M. Jorge. 2017. "An Overview of Data Mining Applications in Oil and Gas Exploration: Structural Geology and Reservoir Property-Issues." *CoRR* abs/1705.06345. http://arxiv.org/abs/1705.06345.

Pankaj, Piyush, Steve Geetan, Richard MacDonald, Priyavrat Shukla, Abhishek Sharma, Samir Menasria, Tobias Judd, and others. 2018. "Need for Speed: Data Analytics Coupled to Reservoir Characterization Fast Tracks Well Completion Optimization." In *SPE Canada Unconventional Resources Conference*. Society of Petroleum Engineers.

Subrahmanya, Niranjan, Peng Xu, Amr El-Bakry, Carmon Reynolds, and others. 2014. "Advanced Machine Learning Methods for Production Data Pattern Recognition." In *SPE Intelligent Energy Conference & Exhibition*. Society of Petroleum Engineers.

Temizel, C, S Purwar, A Abdullayev, K Urrutia, Aditya Tiwari, and others. 2015. "Efficient Use of Data Analytics in Optimization of Hydraulic Fracturing in Unconventional Reservoirs." In *Abu Dhabi International Petroleum Exhibition and Conference*. Society of Petroleum Engineers.

Wang, Shitao, Mohamed Iskandarani, Ashwanth Srinivasan, W. Carlisle Thacker, Justin Winokur, and Omar M. Knio. 2016. "Propagation of Uncertainty and Sensitivity Analysis in an Integral Oil-Gas Plume Model." *Journal of Geophysical Research: Oceans* 121 (5): 3488–3501. http://ezproxy.uhd.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=geh&AN=2016-081255&site=eds-live&scope=site.

Yu, Yuan, Zhang Suian, Yuan Shuqin, Wu Yanqiang, Liu Xinjia, and Wang Hongli. 2018. "Oil Well Productivity Computation Based on a Brain-Inspired Cognitive Architecture." *NeuroQuantology* 16 (5): 776–82. http://ezproxy.uhd.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=132351350&site=eds-live&scope=site.