
Artificial intelligent in upstream oil and gas industry: a review of applications, challenges and perspectives

Kenioua Abdelhamid¹, Touati Brahim Ammar² and Kenioua Laid³

¹Departement of Mechanical Engineering, University of Boumerdes, Algeria

²Schlumberger

³Department of Information and Technology, University of El-Oued, Algeria
a.kenioua@univ-boumerdes.dz

Abstract. In the last two decades, oil and gas (O&G) industries are facing several challenges and issues in different levels; from the decrease in commodity prices to the dynamic and unexpected environment. There has been a constant urge to maximize benefits and attain values from limited resources. Traditional empirical and numerical simulation techniques have failed to provide comprehensive optimized solutions in little time due to the Immense amount of data generated on daily basis with various formats, techniques and process. The proper technical analysis of this “explosion of data” is to be carried out to improve performance of O&G industries.

Artificial intelligence (AI) has found extensive usage in simplifying complex decision-making procedures in practically every competitive market field, and O&G industry is not an exception. This paper provides a comprehensive state-of-art review in the field of machine learning and artificial intelligence to solve O&G industry problems. We focus on the upstream segment as the most capital-intensive part of oil and gas and the segment of enormous uncertainties to tackle. Based on a summary of various researchers work on machine learning and AI applications, we outline the most recent trends in developing AI-based tools and identify their effects on accelerating the process in the industry. This paper discusses also the main challenges related to non-technical factors that prevent the intensive application of AI in the upstream O&G industry.

Keywords: Artificial Intelligence, Oil and gas industry, Upstream.

1 Introduction

Digital transformation has a tremendous influence on business and society. With time it has been regarded as the "fourth industrial revolution", characterized by the convergence of technologies that blur the boundaries between the physical, digital and biological realms, such as artificial intelligence, robotics and autonomous vehicles. Artificial Intelligence (AI) technology is gaining considerable attention and becomes the most important general-purpose of today [1]. Because of its rapid response speed and

robust capacity for generalization, it is quickly entering industries and creating potential of innovations and growth. AI triggered substantial changes and transformed the competition rules in media, transportation, finance and healthcare. Instead of relying on traditional and human-centered business processes, companies from these industries create value using AI solutions [2]. Advanced algorithms trained on large and useful datasets, and continuously supplied with new data drive the value creation process.

However, not only companies from digital-savvy industries are profiting from AI. Oil and gas companies are the latecomers to digitalization [3, 4], but they are also getting more and more dependent on AI solutions. Although the first applications of AI in the oil and gas industry were considered in the 1970s [5], the industry has started to look more proactively for AI application opportunities several years ago [6, 7]. It coincides with the exponential growth of AI capabilities and the industry's movement towards the Oil and Gas 4.0 concept, whose core goal is to achieve higher value utilizing advanced digital technologies [8].

As oil and gas companies are much quicker to adopt new technologies than to experiment with and change their business models [12], their AI's primary target (and other digitalization) efforts are to improve efficiency. In practice, that typically means to accelerate processes and reduce risks [5, 8].

the application of AI technology in the petroleum field, in order to organize regular relevant discussions. Based on the search result from the Onepetro platform, the number of articles on AI has increased significantly the last decade, whose main algorithms include the artificial neural network (ANN), fuzzy logic, support vector machine (SVM), hybrid intelligent system (HIS), genetic algorithm (GA), particle swarm optimization (PSO), etc. This suggests an increasing interest of the researchers in the application of artificial intelligence in the oil industry, and among all the algorithms, the ANN is the most studied one (Fig. 1).

Oil and gas deposits are often located thousands of feet below the earth's surface. The process of extracting oil and gas from these depths to the surface, and then converting it into a usable source of energy involves a large variety of operations. Fig. 1 shows different sectors in the oil and gas industry operations. Broadly, the entire process of producing oil and gas is divided into the three industry sectors, upstream sector, midstream sector and downstream sector. The upstream summarizes the subsurface (mining) part of the industry, operations in the upstream industry are focused on identifying locations below the earth's surface, which have the potential of producing oil and gas. Following the identification of a potential reserve, detailed planning of exploration, drilling oil wells, and producing oil and gas also comes under the upstream industry. Midstream stands for transportation of oil and gas, and downstream is for refinery i.e., production of fuels, lubricants, plastics, and other products.

In this paper we focus our research and discussion on the upstream sector by explaining in detail many of the upstream activities, we discuss points where AI solutions are already applied and their results. We also highlight where we expect AI to be used and what results can come out of its application.

The upstream is of particular interest as it is the most capital-intensive and important of the three segments in the oil and gas business [9]. The saying "one rock, two geol-

ogists, three opinions" tells a lot about the high uncertainties and risks oil and gas companies have to deal with. Manual handling of these enormous uncertainties and

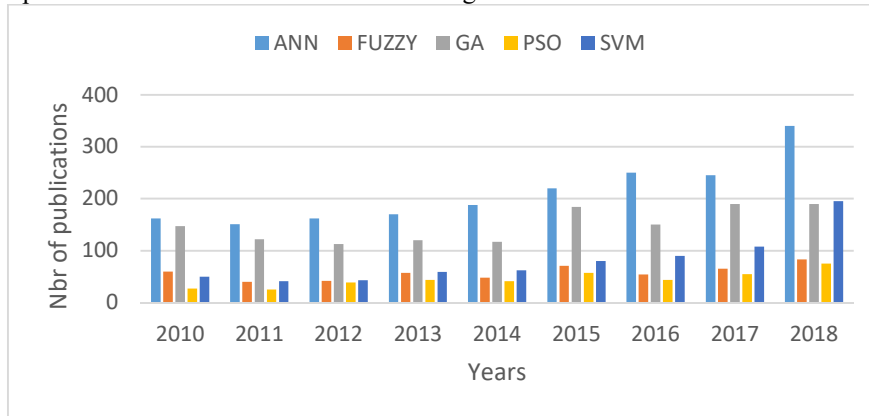


Fig. 1. AI the statistic of increasing number of articles on AI algorithms in Onepetro.

the rely to expert knowledge instead of the actual data can be very risky specially when making multibillion decisions on where and how to invest in the coming 5–20 years. However, despite the complex and uncertain nature of management problems in the sector, the single-criterion approaches have historically dominated decision-making [10]. To use existing field data to account for uncertainties associated with practitioners' subjective perception and decision-making based on experience, the first steps in using artificial intelligence and machine learning in the upstream are made, becoming increasingly popular [9]. The paper summarizes the different research works using AI to solve the problems and limitations in upstream sector. We covered the AI based researches in the whole spectrum of the upstream activities, geological assessment of the reservoirs, drilling optimization, reservoir engineering, field development, and production optimization.

The paper is organized as follow, in Section 2, a brief overview of the AI approaches and algorithms is outlined with a description of the most used algorithms in the sector. In Section 3, we discuss the use of AI based tools in the different sections of upstream sector and identifying their effect on the accelerating process and enhancing performances. We discuss in Section 4 the main challenges face the AI intensive application in the industry. Finally, we conclude with perspective of how AI can change the upstream sector and probably the whole O&G industry in the nearest future.

2 Machine learning algorithm

Machine Learning is a subset of Artificial Intelligence. ML involves teaching the machines to handle large datasets for recognizing the patterns and extracting the relevant information with an enhanced efficiency. The computer science and technology

focusing on building machines when combined with the statistical tools and inferences provides remarkable outcome as ML. As depicted in Fig. 2, ML has broadly

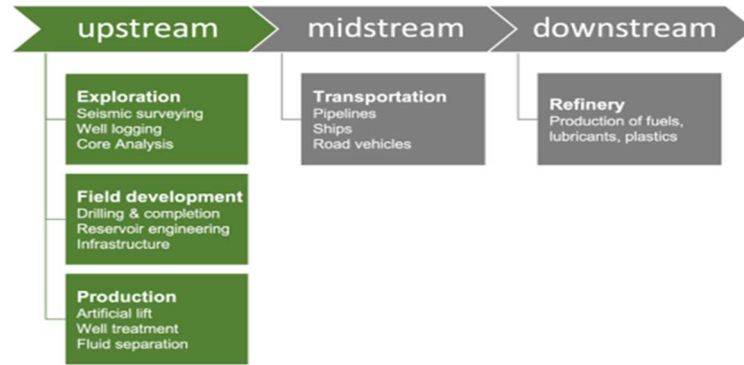


Fig. 2. Sectors of Oil and Gas industry

been classified into: Supervised learning, Unsupervised learning, Semi-supervised learning and Reinforcement learning. In supervised learning, the data-driven model is built by processing a known labeled dataset that includes desired inputs (features) and outputs (labels/responses). Supervised learning finds the model that generates the outputs based on the inputs. A physics-driven model is a theory-based mathematical mapping that relates the input and output, whereas supervised learning identifies patterns in the available dataset, learns from observations, and makes the necessary predictions based on statistical mapping of inputs and outputs. During the process of building the supervised learning model, the predictions are compared to the output and the model is improved based on a loss function. This process continues until the data-driven model achieves a high level of accuracy and performance so the loss function can be minimized. Unsupervised learning processes datasets to identify patterns, relations, and commonalities without using examples, labels, and human instruction. This ML technique organizes the data in a certain way that describes the structure, variance, density, distribution, etc. of the dataset. This might mean grouping the data into clusters or arranging it in a way that looks more organized or easy to visualize. Few examples of unsupervised learning: dimensionality reduction, data compression, manifold learning, and feature extraction. An inherent issue with unsupervised learning is to assess the efficacy and reliability of the model. A training set encompassing unlabeled and labeled datasets characterizes semi-supervised learning algorithms. This approach can yield better results in some cases compared to supervised learning techniques. Finally, using trial and error, the algorithm repeatedly self-trains in reinforcement learning. Feedback is provided to the algorithm on the causes of error but no instruction for error rectification is provided.

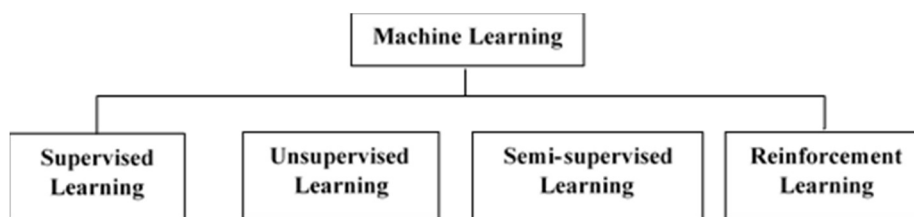


Fig. 3. Classification of ML approaches

3 Applications of ML in upstream oil and gas

This section of the paper focuses on the researches that have been conducted to implement the ML tools and techniques in various sectors of upstream oil and gas industry which have been mainly categorized into Exploration & Drilling Operations, Reservoir Engineering, and Petroleum Production System. We have reviewed over 173 works compared machine learning techniques with traditional models. Most of them showed that the use of learning algorithms provided a more-accurate prediction than the use of traditional models. The reason for that is the capability of those models to capture non-linear relationship among the variables. Fig. 2 shows the ML methods used in the studies as well as their dispatching per discipline in upstream sector. As shown in Fig. 2. a), ANN were employed 59 times followed by homogeneous ensembles (e.g. RF, GBM) which used 17 times then followed by the other techniques. In the scope of studies reviewed in this paper we can notify that ML used 78 times in drilling optimization followed by the production discipline with 41 studies then the exploration with 35 studies over the 173 works reviewed in this paper.

3.1 Exploration

Traditional methods of geological and geophysical modeling involve generating computerized representations of sub-surface observations obtained from various geological and geophysical surveys for evaluating the structural and stratigraphic description along with estimation of reservoir rock and fluid characteristics. These tasks typically consist of reservoir-scale seismic surveying, well logging, and lab core analysis.

Mahdi Rastegarnia et al. [11] utilized AI systems to extract electrofacies volumes and 3D Flow Zone Index (FZI) from large volume of 3D seismic data. Multi Resolution Graph based Clustering (MRGC) was employed to optimize the electrofacies volumes and multi attribute analysis was employed to create 3D FZI model. The electrofacies models are further improved using the Probabilistic Neural Network (PNN), and 3D FZI model was improved using Radial Basis Function (RBF) Network, Multilayer Feed Forward Network (MLFFN) and PNN. Modern pattern recognition techniques based on deep learning have started to dive into this seismic-related operation, accelerating the interpretation by a factor of 10–1000 [12]. Ke-ran Qian et al. [13] proposed SVM-based algorithm in combination with multi-scale and multi-resource information from geology, seismic, drilling and logging for sweet spot prediction along

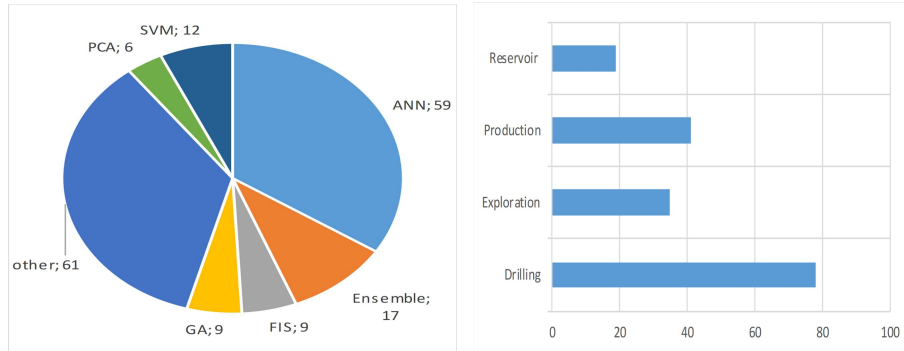


Fig. 4. a) ML methods employed in upstream studies. b) Studies used ML in upstream by sector.

and characterization of shale reservoirs with high accuracy. There is a low probability that the AI techniques will optimize the physical part (i.e., amount, cost, and placement layout of sensors) of the first seismic surveying at an asset. Still, they add value in the optimization of the secondary surveys at the same asset. The mathematics of recommender systems and interpolation capabilities of machine learning algorithms will enable proper recommendations on making the secondary surveys cheaper with a minor loss in the value of acquired information [14]. The petrophysical interpretation is a rather time-consuming process, and the result of the interpretation depends strongly on the interpreter. AI-aided technologies are the obvious way to accelerate and, maybe even more critical, to exclude the subjective part of the interpretation process. Wood DA [15] predicted the porosity, permeability and water saturation using optimized nearest-neighbor associated with data mining techniques. Meshalkin Y et al. [16] developed Robotized petrophysics workflow using machine learning and thermal profiling for automated mapping of lithotypes in unconventional. A numerical simulation model coupled with ANN and SVM classifier was developed by Kyung Jae Lee [17] to determine and classify the kerogen characteristics. Baraboshkin EE, et al. [18] used the deep learning to generate an automated well logs interpretation process for estimation rock types.

3.2 Reservoir engineering

Once the initial geological model is built, it goes to reservoir engineers. They perform upscaling and build a reservoir model from the geological model using reservoir modeling software[19, 20]. This model can estimate reservoir flows at various field development schemes which contains the plan for well drilling and well operation. The result of each of the reservoir modeling runs is a forecast of oil/gas production for forthcoming years for a particular field development scheme. Performing many runs, the reservoir engineers select the optimal field development scheme and field development plan for both green-fields and brown-fields.

Deep neural networks technique is used for acceleration of reservoir modeling. Modern surrogate reservoir models with a new computation engine based on deep neural networks has been used in [21]; this technique can compress the mathematical problem dimensionality and approximate the time derivatives promise 100–1000 times the conventional models' speedup while keeping similar functionality. Simonov et al [22] used different machine learning techniques to speed up the 3D modeling for the well in development. The upscaling process has also benefited from the machine learning algorithms to increase objectiveness and speed of the process by using deep learning algorithm trained on multiple cases of manual upscaling[23-25].

3.3 Drilling engineering

Drilling a well is a challenging task as it involves no or minimal prior information of the sub-surface characteristics and the complexity increases with increasing depth or deviation of the well from its vertical path. Moreover, occurrence of drilling problems such as lost circulation, pipe sticking, shale sloughing, dogleg severity and others add up to make the job highly crucial. Most of well drilling cost is not product cost dependent, but time dependent [26]. Therefore, one of the main goals of drilling optimization is to reduce the total time, maintaining the risks as low as possible. One way to achieve it is through selecting of optimum drilling variables prior a run (e.g. selecting a suitable drill-bit and drilling fluid type). Another approach relies on real-time analysis in order to optimize operational parameters (e.g. bit weight, rotary speed) while drilling [27]. Farough Agin et al. [28] conducted research for estimation of lost circulation during drilling operations with the application of Adaptive Neuro-Fuzzy Inference System (ANFIS). The model was trained, tested and checked to predict lost circulation volume. Ekaterina Gurina et al. [29] have demonstrated ML algorithm to detect accidental events in directional wells by comparing the real – time MWD data with the past data. The proposed model performs anomaly detection by analyzing the similarity of events using time- series comparison and gradient boosting classification. Optimization of the pre-specified key performance indicators including ROP (ft/hr) and mechanical specific energy (psi) was carried out in the research [30] using coupled end-to-end model for drilling optimization by taking into account all the parameters of interest including Weight on Bit (WOB), rotary speed, flow rate and rock strength. The researchers utilized the Random Forest (RF) algorithm for developing models for each individual parameter under consideration that were coupled using ML algorithm.

3.4 Production engineering

Producing reservoirs are attractive for AI-aided tools as well as the green fields. There are obvious machine learning applications for various pumps to implement predictive maintenance and select the optimal operation regimes concerning operational costs vs. production. Many of the pumps, including electric submersible pumps, pumps for injection wells, hydraulic fracturing, and other well treatment pumps, are equipped with a high number of sensors measuring pressures, temperatures, vibrations, flow

rates, etc. There are many examples when an entirely data-driven or a hybrid model containing physics-driven and data-driven math helps optimize the regimes, prevent unexpected failures, and save on maintenance-on-schedule [31, 32]. Li et al. [33] used a Neural Decision Tree (NDT) model for prediction of oil production by considering interconnectedness among the input variables. N. Chithra Chakra et al. [34] applied Higher Order Neural Network (HONN) to predict cumulative oil production (m³) from a conventional oil field with limited training data. It was found to be a satisfactory tool to forecast the cumulative oil production for short – term as well as long-term planning. Using electrical and frequency data as input features, Guo et al. in [32], developed an SVR workflow to forecast failures in Electrical Submersible Pump (ESP). Gupta et al., 2016, deployed a hybrid mathematical method consisting of an intelligent predictive monitoring KPI that automatically identifies imminent glitches, diagnose root causes, and prescribe corrective actions to abnormal ESP operational situations in real time raising alarms through predictive, diagnostic, and prescriptive analytics. There is an excellent opportunity to reduce the investment risks by accumulating the data from already produced well treatment jobs. Pioneering efforts on predicting the efficiency of hydraulic fracturing jobs in [35] and ML-based analysis of injectivity issues have already been made in [36].

4 Challenges and perspectives

The success of artificial intelligence critically depends on human intelligence. AI solutions have to be customized to the business context and database of a company. Thus, to actively use AI in processes and products, companies must grow in-house teams composed of data and AI specialists. These teams should be able to support development of AI infrastructure (algorithms and datasets) and, at least to customize tools that companies will later utilize in their operations. That means that oil and gas companies will become (partially) data-driven companies and, that AI specialists will become irreplaceable in supporting almost all innovation efforts in oil and gas companies in the next 10 years. AI tools need the good quality data of a suitable volume to be trained and then to work properly in the operational mode. While using smarter algorithms may help in getting better results from datasets of limited size, no manipulation can help with bad data. Thus, access to big and quality data is a crucial enabler and barrier for AI applications' successful development. Artificial intelligence is born in open and collaborative environment as a consequence of academia being a leading force in AI research for decades, almost without any business influences. This created culture of free sharing and open publishing which companies across industries (and across the globe) had to embrace as a standard to succeed in the era of AI once they joined the race. oil and gas companies should re-think strategies for collaborating and interacting with universities.

5 Conclusion

The aim of this paper is to present a review concerning the myriad of different applications and benefits of Artificial Intelligence and Machine Learning techniques in the upstream sector of oil and gas industry which cover disciplines of exploration, reservoir, drilling and production. The paper compiles the major workflows and achievements of the industry on a higher-level overview with a focus on its leverage over other traditional modelling techniques. The literature review of oil and gas industry is well-poised to take benefits of machine learning regarding their abilities of processing big data and fast computational speed. Machine learning has the potential of unequivocally changing the numerous critical actions made every day by administrators and engineers in the oil and gas sector. The future advantages of information can be achieved if appropriate techniques are used to implement different data types or structures and convert it into useful information that contributes to intelligent judgements. AI and machine learning will change the face of oil and gas industry and lead to speed up and de-risk many business processes associated to this critical and crucial industry.

References

1. Brynjolfsson E, Mitchell, T: *What can machine learning do?. Workforce implications. Science ;358(6370):1530–4 80 (2017).*
2. Iansiti M, Lakhani KR. *Competing in the age of AI. Harv Bus Rev (2020).* [Online]. Available : . Available from: <https://hbr.org/2020/01/competing-in-the-age-of-ai>
3. R. Kohli and S. Johnson: *Digital transformation in latecomer, 10(4), 141–156, (2011).*
4. Kane GC, Palmer D, Phillips AN, Kiron D, Buckley N: *Strategy, not technology, drives digital transformation. MIT Sloan Manag Rev (2015).*
5. Li H, Yu H, Cao N, Tian H, Cheng S. *Applications of artificial intelligence in oil and gas development. Arch Comput Methods Eng (2020).*
6. BCG Homepage : *Going digital is hard for oil and gas companies—but the payoff is worth it, <https://www.bcg.com/ru-ru/publications/2019/digital-value-oil-gas.aspx> Last Accessed 24/08/2021.*
7. BCG Homepage: *Big oil, big data, big value, <https://www.bcg.com/publications/2019/big-oil-data-value.aspx>. Last accessed 24/08/2021.*
8. Lu H, Guo L, Azimi M, Huang K: *Oil and gas 4.0 era: a systematic review and outlook. Comput Ind 111(6) pp:68–90, (2019).*
9. Shafiee M, Animah I, Alkali B, Baglee D: *Decision support methods and applications in the upstream oil and gas sector. J Pet Sci Eng 173 pp:1173–1186 (2019).*
10. Strantzali E, Aravossis K. *Decision making in renewable energy investments: a review. Renew Sustain Energy Rev ;55 pp:885–898 (2016).*
11. M. Rastegarnia, A. Sanati, and D. Javani, : *A comparative study of 3D FZI and electrofacies modeling using seismic attribute analysis and neural network*

- technique: A case study of Cheshmeh-Khosh Oil field in Iran,” *Petroleum*, 2(3), pp: 225–235, (2016).
12. Cunha A, Pochet A, Lopes H, Gattass M : Seismic fault detection in real data using transfer learning from a convolutional neural network pre-trained with synthetic seismic data. *Comput Geosci*: 104344 (2019).
 13. K. R. Qian, Z. L. He, X. W. Liu, and Y. Q. Chen : Intelligent prediction and integral analysis of shale oil and gas sweet spots,” *Petroleum Science*, 15(4), pp: 744–755, (2018).
 14. Portugal I, Alencar P, Cowan D : The use of machine learning algorithms in recommender systems: a systematic review. *Expert Syst Appl* 97 pp:205–227 (2018).
 15. Wood DA : Predicting porosity, permeability and water saturation applying an optimized nearest-neighbour, machine-learning and data-mining network of well-log data. *J Pet Sci Eng* (2019).
 16. Meshalkin Y, Koroteev D, Popov E, Chekhonin E, Popov Y : Robotized petrophysics: machine learning and thermal profiling for automated mapping of lithotypes in unconventional. *J Pet Sci Eng* (2018).
 17. K. J. Lee : Characterization of kerogen content and activation energy of decomposition using machine learning technologies in combination with numerical simulations of formation heating,” *Journal of Petroleum Science and Engineering*, vol. 188, (2020).
 18. Baraboshkin EE, et al. Deep convolutions for in-depth automated rock typing. *Comput Geosci* (2019) doi:10.1016/j.cageo.2019.104330.
 19. Gasda SE, Celia MA : Upscaling relative permeabilities in a structured porous medium. *Adv Water Resour* ;28(5) pp:493–506 (2005).
 20. Fanchi JR, Christiansen RL. *Introduction to petroleum engineering*. Wiley Publishing; (2016).
 21. Temirchev P, et al : Deep neural networks predicting oil movement in a development unit. *J Pet Sci Eng* (2020).
 22. M. Simonov et al., “Application of machine learning technologies for rapid 3D modelling of inflow to the well in the development system,” doi: 10.2118/191593-18rptc-ru. (2018).
 23. Barker JW, Thibeau S. A critical review of the use of pseudorelative permeabilities for upscaling. *SPE Reserv Eng* 12(2) pp:138–143 (1997).
 24. Farmer CL. Upscaling: a review. *Int J Numer Methods Fluids* ;40(1–2) pp:63–78 (2002).
 25. Pickup GE, Stephen KD, Ma J, Zhang P, Clark JD. Multi-stage upscaling: selection of suitable methods. In: *Upscaling multiphase flow in porous media: from pore to core and beyond*. Netherlands: Springer; p: 191–216 (2005).
 26. Lyons, W.C., Plisga, G.J.: *Standard Handbook of Petroleum and Natural Gas Engineering*, second ed. Gulf Professional Publishing, (2004).
 27. Payette, G.S., Spivey, B.J., Wang, L., Bailey, J.R., Sanderson, D., Kong, R., Pawson, M., Eddy, A.,. *Real-time well-site based surveillance and optimization platform for drilling: technology, basic workflows and field results*. In: *SPE/IADC Drilling Conference and Exhibition*. Society of Petroleum Engineers, Hague, The Netherlands (2017).

-
28. F. Agin, R. Khosravanian, M. Karimifard, and A. Jahanshahi, *Application of adaptive neuro-fuzzy inference system and data mining approach to predict lost circulation using DOE technique (case study: Maroon oilfield)*. Southwest Petroleum University, (2019).
 29. E. Gurina et al., "Application of machine learning to accidents detection at directional drilling," *Journal of Petroleum Science and Engineering*, vol. 184, September, (2020).
 30. C. Hegde, M. Pyrcz, H. Millwater, H. Daigle, and K. Gray, "Fully coupled end-to-end drilling optimization model using machine learning," *Journal of Petroleum Science and Engineering*, vol. 186, (2020). doi: 10.1016/j.petrol.2019.106681.
 31. J. Sneed, "Predicting ESP lifespan with machine learning," 2017, doi: 10.15530/urtec-2017-2669988.
 32. D. Guo, C. S. Raghavendra, K. T. Yao, M. Harding, A. Anvar, and A. Patel, "Data driven approach to failure prediction for electrical submersible pump systems," 2015, doi: 10.2118/174062-ms.
 33. X. Li, C. W. Chan, and H. H. Nguyen, "Application of the Neural Decision Tree approach for prediction of petroleum production," *Journal of Petroleum Science and Engineering*, vol. 104, pp. 11–16, 2013, doi: 10.1016/j.petrol.2013.03.018.
 34. N. Chithra Chakra, K. Y. Song, M. M. Gupta, and D. N. Saraf, "An innovative neural forecast of cumulative oil production from a petroleum reservoir employing higher-order neural networks (HONNs)," *Journal of Petroleum Science and Engineering*, vol. 106, pp. 18–33, 2013, doi: 10.1016/j.petrol.2013.03.004.
 35. Makhotin I, Koroteev D, Burnaev E. Gradient boosting to boost the efficiency of hydraulic fracturing. *J Pet Explor Prod Technol* 2019. doi:10.1007/s13202-019-0636-7.
 36. Orlov D, Koroteev D. Advanced analytics of self-colmatation in terrigenous oil reservoirs. *J Pet Sci Eng* 2019. doi:10.1016/j.petrol.2019.106306.