ISSN 1870-4069

Data-Driven Methodology for Candidate Well Selection and Ranking

Héctor De Paz Galicia[1], Eduardo Roldán Reyes[1], Leonid Sheremetov[2]

[1] Tecnológico Nacional de México,
Instituto Tecnológico de Orizaba, Veracruz,
Mexico

[2] Instituto Mexicano del Petróleo,
Mexico

hdepazg@ito-depi.edu.mx

Abstract. This research work proposes a new data-driven methodology that helps at different stages of wells productivity analysis to increase the production in oil and gas fields. The first two stage is aimed in developing a data-driven methodology for the identification, selection and ranking of candidate wells for productivity analysis based on the historical oilfield data. The methodology integrates several supervised and unsupervised machine learning techniques, multi-criteria decision analysis and case-based reasoning.

Keywords: Candidate-wells selection and ranking, data mining, TOPSIS, genetic algorithm, NARX neural network, CBR.

1 Introduction

Wells productivity analysis (WPA) has been seen as a main strategy for increasing production in oil and gas fields [1]. Production analysis of historical data evidence that different workovers within the same oilfield produce very non-uniformly, and up to 30% may not be producing at all due to a combination of reservoir heterogeneity, geomechanical, workover design and deployment factors. Diverse studies on WPA, underline the effectiveness of the analysis and the quality of the results, which are based on three critical points: the choice of the variables involved in the process, the candidate-wells selection (CWS) and the workover design.

WPA begins with the selection and prioritization of wells that have the best characteristics to increase the probability of success. However, the main difficulties that arise during this analysis, is the large number of wells to study, the necessity to examine the high volumes of data to identify opportunities for production of wells operating, which require considerable effort and a significant investment of resources using conventional numerical analysis methods [2]. Therefore, a data-driven methodology that streamlines these information processing operations is a time challenge.

2 Related Work

Recently many efforts have been focused on different processes of well productivity. Analyzing candidate well selection (CWS) for one of such processes, hydraulic fracturing (HF), Zoveidavianpoor and Gharibi classified the methods into conventional and advanced [3]. Unfortunately, conventional approaches to CWS are limited by the experience, expertise and preconceived notions of the specialist who is conducting the evaluation. Another limitation of these techniques is the time required to perform the analysis. Advanced methodologies represent an alternative approach to the selection and ranking of wells. Particularly, by the opportunity to examine the data and their relationships in different ways, maximizing the potential of the data. A portfolio of advanced algorithms includes analytic hierarchy process (AHP), decision trees, random forest, ANNs, linear regression and support vector machines (SVM) [4, 5]. This classification can be extended to other well productivity processes like the stimulation treatment and workovers, where gradient boosting and probabilistic expert system technique have been used [6, 7]. Although data mining and AI methods have been reported for CWS, these have only been focused on particular processes like HF, workovers, stimulation, or refracturing. An integrated data-driven methodology for WPA have not been reported in the literature so far.

3 Methodology

The proposed data-driven CWSR methodology follows the main stages of a process of the productivity analysis, which begins with the acquisition of information and ends with the proposed solution to increase wells' production:

1. Data acquisition, pre-processing, and transformation.
2. CWS for productivity analysis.
3. OPS analysis and diagnostics.
4. Optimization, selection and ranking of workover proposals.
5. Conceptualization and knowledge base.

The first stage can be considered as a typical stage of any data analytics project. Supervised and unsupervised machine learning techniques are used for data pre-processing and candidate well selection, as well as estimation of the type of intervention (1, 2). A forecasting model based on Nonlinear Autoregressive network with eXogenous inputs (NARX) is capable of predicting the response that the oil production will have, in the following three months after the treatment (3). A Multi-Criteria Decision Analysis (MCDA) method called a Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is applied for selected wells' ranking (4). The measures borrowed from information retrieval, cumulative gain and discounted cumulative gain (DCG), are applied under the premise that the first well from the list of candidates has the highest probability of success.

A genetic algorithm is proposed to optimize the weights applied in TOPSIS. Finally, a case-based reasoning (CBR) will allow for a weighted evaluation of the workover proposals based on previous experiences and best practices (5).

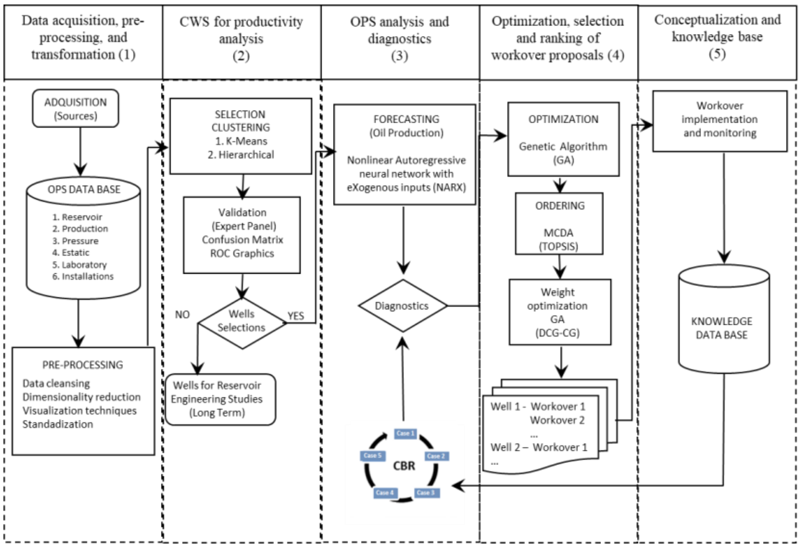

Fig. 1. Proposed methodology for the selection and ranking of wells.

4 Results and Discussion

To test the proposed methodology, a mature field located in the southeast of Mexico is considered. For the case study, 50 wells producing from the Mesozoic formation were selected with a total of 700 workovers and the dimensions of the original data sample with 38 variables and 11381 observations. The first four stages of the methodology have been developed. At the first stage, the α-trimmed mean ($\alpha = 0.2$) technique was used for the outliers' detection. Data transformation included both the definition of additional characteristics to be taken into consideration, dimensionality reduction with PCA, variables' selection and scaling. After dimensionality reduction, nine variables were selected.

The second stage deals with CWS, where both K-means and hierarchical clustering algorithms were used. In order to validate the obtained results, the candidate wells' list generated by the experts' panel was used as the reference solution. As a result of CWS, 30 wells were selected as candidates for further analysis and ranking and 20 as non-candidates with the values of F1 Score 0.95 and 0.933 for the hierarchical clustering and K-Means respectively. At the third stage, wells' models were developed for well productivity analysis based on available historical data in order to predict expected outcome of the workover for each well from the list. The forecast tool helped to identify the future event by providing the conditions (estimated variables) to carry out the diagnosis of the well. Time series forecasting techniques were used at this stage [8]. Finally, for wells' ranking, multi-objective data analysis (TOPSIS) was integrated with the evolutionary algorithm to optimize the parameters of the developed models. With the use of GA with the DCG-based fitness for the optimization of the weights, a clear correlation with experts' panel is obtained ($R^2 = 0.99$), not the case with the CG-based fitness. The differences in the ranking found by TOPSIS and the experts' panel can be attributed to the subjectivity of the experts in assigning the weights on the selected criteria; in the same way, the results of the relative proximity indicate that, for values that are very close or equal, the allocation of the ranking of the wells becomes indistinct.

5 Conclusions and Future Work

The selection and ranking of wells should be considered as a fundamental part in the methodology of well productivity. The application of the hybrid approach based on the integration of data analytics, MCDA, and evolutionary optimization to support decision making, is an alternative that offers efficient solutions to solve this problem. The approach allows to obtain acceptable results in a short time, reducing the analysis time for the CWSR by 78% in comparison with conventional methods. At the moment of writing, the experiments on productivity analysis and forecasting are in their final steps. Meanwhile, the CBR model, which will permit to capitalize knowledge by reusing past experiences, has been under development. To increase its efficiency, the proposal of a similarity measure relying on the spherical indexing algorithm is envisaged. This will reduce the adaptation effort since this task is a common issue to be solved.

References

1. Hendraningrat, L., A Salam, I.H., Mohsin, N., Wah, T.S., Rinadi, M., Sedaralit, M.: Enhancing and accelerating reservoir modelling and history matching of an offshore mature oilfield with geological and production system complexity using data-driven approach. In: Proceeding of the International Petroleum Technology Conference (2020)
2. Enciso-Paredes, E.J., de Antuñano-Muñoz, Y., Pérez-Herrera, R., Perera-Pérez L.M.: Metodología para identificar oportunidades de producción en campos maduros. Ingeniería Petrolera, 59(2), pp. 18 (2019)
3. Zoveidavianpoor, M., Abdoullatif, G.: Applications of type-2 fuzzy logic system: handling the uncertainty associated with candidate-well selection for hydraulic fracturing. Neural Computing and Applications, 27(7), pp. 1831–51 (2016)
4. Temizel, C., Purwar, S., Abdullayev, A., Urrutia, K., Tiwari, A.: Efficient use of data analytics in optimization of hydraulic fracturing in unconventional reservoirs. In: Proceeding of the Abu Dhabi International Petroleum Exhibition and Conference (2016)
5. Wang, Y., Salehi, S.: Refracture candidate selection using hybrid simulation with neural network and data analysis techniques. Journal of Petroleum Science and Engineering, 123, pp. 138–46 (2014)
6. Makhotin, I., Koroteev, D., Burnaev, E.: Gradient boosting to boost the efficiency of hydraulic fracturing. Journal of Petroleum Exploration and Production Technology, 9(3), pp. 1919–25 (2019)
7. Zangl, G., Neuhofer, L., Zabel, D., Tippel, P., Pantazescu, C.I., Krcmarik, V., Krenn, L., Hachmöller, B.: Smart and automated workover candidate selection. In: Proceeding SPE Intelligent Energy International Conference and Exhibition (2016)

8. Sheremetov, L., Cosultchi, A., Martínez-Muñoz, J., Gonzalez-Sánchez, A., Jiménez-Aquino, M.A.: Data-driven forecasting of naturally fractured reservoirs based on nonlinear autoregressive neural networks with exogenous input. Journal of Petroleum Science and Engineering, 123, pp. 106–19 (2014)