Journal of Natural Gas Science and Engineering 88 (2021) 103819

Contents lists available at ScienceDirect

Journal of Natural Gas Science and Engineering

journal homepage: http://www.elsevier.com/locate/jngse

Machine-learning predictions of the shale wells' performance

Mohamed Mehana *, Eric Guiltinan, Velimir Vesselinov, Richard Middleton, Jeffrey D. Hyman, Qinjun Kang, Hari Viswanathan

Computational Earth Science Group (EES-16), Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, USA, 87545

ARTICLE INFO

Keywords:
Machine learning
Estimated ultimate recovery
Well performance
Hydraulic fracturing

ABSTRACT

The ultra-low permeability nature of shale reservoirs leads to an extended linear flow and necessitates horizontal wells with multi-stage engineered fractures to efficiently extract hydrocarbons resources. These artificially-generated and naturally-occurring fractures form complex networks that create complex flow regimes which control oil production. These fractures are neither identical nor equally-spaced, which leads to a production profile with a masked onset of the boundary-dominated flow. The combination of the extended linear flow with the indeterminate onset of the boundary-dominated flow challenges the current deterministic analytic approaches to forecast the estimated ultimate recovery (EUR). Herein, we propose a novel machine-learning approach which overcomes these challenges and provides reliable EUR estimates based on field-wide analyses. We implement a novel unsupervised machine learning (ML) methodology, which allows for automatic identification of the optimal number of features (signals) present in the data based on non-negative matrix/tensor factorization coupled with *k*-means clustering incorporating regularization and physics constraints. In the presented analyses, the input data to the ML algorithm is the available (public) production history from the field collected at existing unconventional reservoirs. We validate our approach through hindcasting of the production data, where we achieved an excellent agreement. In addition, our approach is able to identify the poorly-performing wells, which could benefit from early refracing. Our approach provides fast and accurate estimations of the well performance without presumptions about the state of the well or the flow regime.

1. Introduction

The combination of multi-stage hydraulic fracturing with horizontal drilling has untapped the potential of shale reservoirs (Hughes, 2013; Middleton et al., 2017). Currently, shale reservoirs account for 70% of the US natural gas production and 60% of the US oil production as shown in Fig. 1 (Total primary energy supp, 1990). However, the majority of the pores in these reservoirs are in the nano-scale which necessities modifications to the conventional development plans (Bustin et al., 2008; Passey et al., 2010; Mehana and El-monier, 2016; Neil et al., 2020). In addition, shale wells usually experience extended linear flows that limit the applicability of most performance prediction techniques, which typically rely on simplified representations of the governing physical processes (Ozkan et al., 2011; Al-Rbeawiet al., 2020).

Traditionally, hydrocarbon reserves are estimated either by reservoir simulation (Lee et al., 2011), volumetric calculation (Abrahamsen et al., 1992), rate-transient analysis (Clarkson, 2013) or decline curve analysis (DCA) (Ayeni, 1989). DCA is a commonly-used approach to provide fast and accurate estimations of the estimated ultimate recovery (EUR). DCA relies on empirically providing the best fit to the production history and does not require costly data similar to other approaches (Arps, 1945). developed their well-known DCA model assuming a boundary-dominated flow and a constant bottom-hole pressure and drainage area. However, shale wells usually experience extended linear flow before boundary-dominated flow with no standard operating conditions and indefinite drainage area which violate most of the assumptions implemented for the development of Arps' equations (Arps, 1945).

Several DCA models have been specifically proposed to predict the EUR of shale reservoirs (Valko, 2009). proposed a Stretch Exponential Production Decline (SEPD) model which guarantees a bounded EUR estimates. In addition (Duong, 2010), developed a transient model which is well-suited to handle the extended linear flow observed in shale reservoirs. However, both SEPD and Duong's models tend to misestimate the EUR as the well enters the boundary-dominated flow regime (i.e. the drainage radius reaches the reservoir boundaries). In response, modified versions were proposed by switching to the hyperbolic decline

* Corresponding author.
E-mail address: mzm@lanl.gov (M. Mehana).

https://doi.org/10.1016/j.jngse.2021.103819
Received 30 October 2020; Received in revised form 15 January 2021; Accepted 17 January 2021
Available online 12 February 2021
1875-5100/Published by Elsevier B.V.

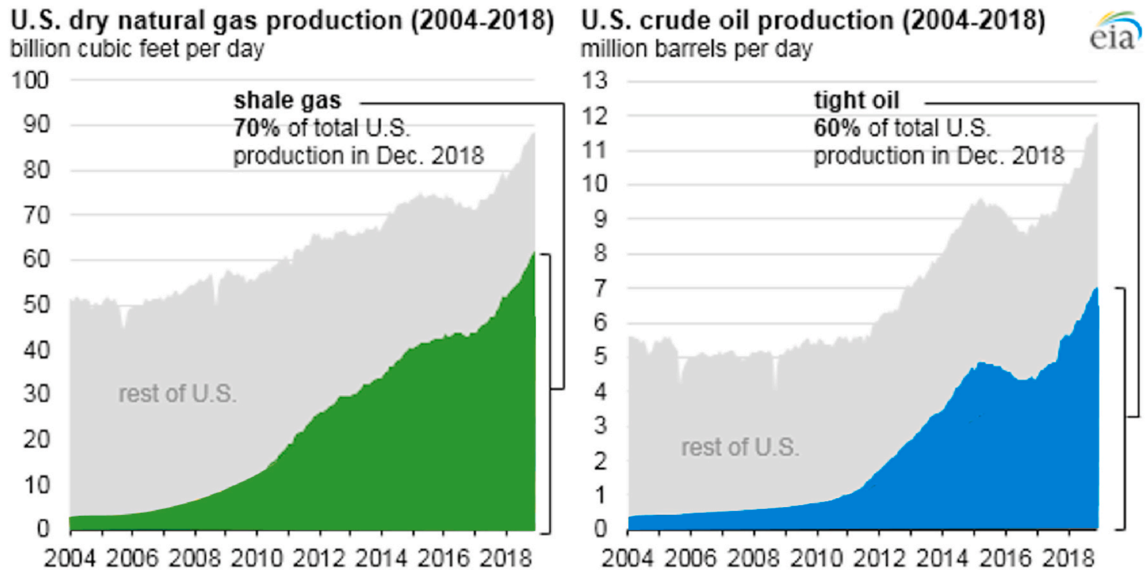

Fig. 1. Shale hydrocarbon production in the United States of America (Total primary energy supp, 1990).

model of (Arps, 1945) at late times (Joshi and Lee, 2013; Statton, 2012). Despite providing better estimates of the EUR, these methods unveiled a tuning parameter for the switch from the SEPD/Duong's models to hyperbolic decline model.

Further empirical models have been proposed to overcome challenges posed by shale characteristics (Clarkson, 2013; Mehana and Callard, 2018). Clark et al. (2011) adopted the logistic growth model to predict production decline in shale reservoirs using the carrying capacity concept, which refers to the maximum size a population can grow to, to constrain the forecast. In the same vein, Ilk et al. (2011) proposed a power-law model that uses the loss ratio, the production rate normalized by the derivative of the production rate with respect to time, to define the decline rate and is capable of modelling the linear and boundary-dominated flows. Patzek et al. (2013) proposed a field-wide approach to forecast the well performance through a scaling method. Additionally, Zhang et al. (2016) proposed the "growing drainage volume" to overcome the shortcomings of traditional DCA which does not require a switch from a transient model to a boundary-dominated flow model. In addition, Ma and Liu (2018) developed a nonlinear model integrating Arps' model and the Kernel method which outperformed the traditional Arps model.

The heterogeneity of the shale reservoirs and the complexity of the completion schemes impose uncertainty on the EUR estimate (Qin et al., 2019; Cheng et al., 2008; Al-Rbeawi, 2017). Fortunately, probabilistic approaches, such as Monte Carlo (MC), can be used to quantify this uncertainty (Macmillan and Hons, 2000; King et al., 2005; Anderson et al., 2012). Recently, Mehana et al. (2020) integrated production analysis and MC simulation to forecast the EUR of transient wells based on the performance of the mature wells. Previously, Gonzalez et al. (2012) used MC to develop a probabilistic approach to examine several models. In addition, Vonei et al. (Voneiff et al., 2014) combined the multivariate regression analysis with MC to provide a probabilistic approach to forecast the performance of 425 wells in the Montney formation where a 95% confidence match was observed.

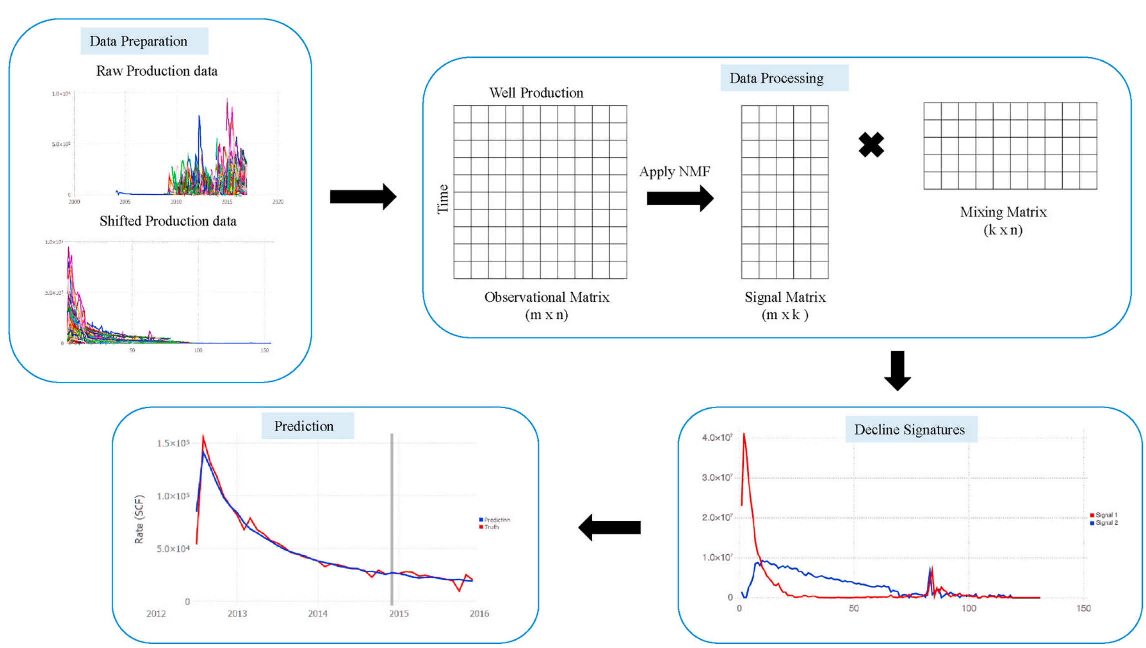

Fig. 2. Workflow of the main steps of NMFK approach.

Recently, machine learning techniques have been adopted to provide real-time predictions of the well performance (Mudunuru et al., 2020). In one of the earliest papers, Li et al. (2013) evaluated the potential of a Neural-Based Decision Tree (NDT) learning model for prediction of production where Artificial Neural Network (ANN) outperformed NDT. In addition, Chakra et al. (2013) used a high-order neural network to forecast production. On the other hand, Klie (2015) proposed surrogate models to replace the computationally-intensive numerical models to forecast the well performance. In addition, Fulford et al. (2016) integrated supervised machine learning with a well-calibrated bias to improve the estimation of uncertainty of the posterior distribution of forecasts. Furthermore,Nguyen-Le et al. (2020) developed prediction models based on simulation data that relate the early production data to EUR.

We propose a holistic machine-learning approach that does not require a priori flow regime identification. Our approach uses Non-negative Matrix Factorization coupled with k-means clustering (NMFk) to identify and learn the decline signatures from field-wide production history to forecast well performance at new and existing wells. We applied the proposed approach to forecast the well performance of a subset of wells in Eagle ford formation, Texas where the efficiency and accuracy of the approach is evaluated. The paper is organized as follows: the Methodology section presents an overview of the machine-learning technique and the Case Study; the Results section discusses the results of our work; and the Conclusions section summarizing the main findings.

2. Methodology

The production rate of a well placed in a homogeneous reservoir can be forecasted after the reservoir reaches boundary-dominated flow (a few minutes to hours in conventional reservoirs). However, shale wells usually necessitate completion schemes and possess heterogeneous reservoir characteristics including natural fractures and ultra-low permeability of the matrix, which result in a complex fracture network of natural and hydraulic fractures (Gong et al., 2019, 2020). Therefore, the production profile of a shale well is highly determined by these completion and reservoir characteristics. Our approach, Julia-based, decouples these contributions (signals) and allows for the reconstruction of the production profile and prediction of the performance of new wells. We discuss our novel NMFk algorithm in this section (Fig. 2) and contrast it with the classical Non-negative Matrix Factorization (NMF) method (Lee and Seung, 1999).

In a typical NMF problem, the observational data, X, is reconstructed by multiplication of an unknown “signal” matrix, W, and an unknown “mixing” matrix, H, i.e.,

$$X = W \times H + E, \tag{1}$$

where E is a “noise” matrix representing the NMF deconstruction error. E also denotes presence of possible noise or unbiased errors in the data X (also unknown); any biased or systematic errors in the data will be extracted as signals in the matrix W. If the NMF reconstruction is successful, it is expected that E represents white noise.

The dimension of X is $n \times m$ where m is the number of wells and n is the number of temporal production records (for example, on a daily or monthly scale). The “signal” matrix W has a size of $n\times k$ where k is the unknown number of signals (features) present in the analyzed data X. The signals in this case are temporal vectors embedded as columns of W. The ‘mixing” matrix H has a size of $k \times m$. The rows of H capture how the signals in W are mixed in each well.

The NMF equation (Eq. (1)) is solved subject to the following constraints related to W and H matrix elements:

$$W_{n,k} > 0, H_{k,m} > 0; \quad \forall \quad n, k, m. \tag{2}$$

Additional constraints that provide physics information can be implemented in NMFk as well. The non-negativity constraints lead to reconstruction of the observations (matrix X) as linear combinations of the elements of H and W that cannot cancel mutually.

The NMF algorithm solves Eq. (1) by minimizing the cost (objective) function, O, which in our case is the Frobenius norm,

$$O = \frac{1}{2}\|X - W^*H\|_F^2 \tag{3}$$

The NMF solution starts with a random guess for H and W elements. Minimizing the Frobenius norm (Eq. (3)) with non-negativity constraints (Eq. (2)) is equivalent to representing the discrepancies between the observations, X, and the reconstruction, $\times H$, as white noise.

The classical NMF requires a priori knowledge of the number of signals k present in the data X. However, k is typically unknown. To estimate the optimal k (k_{opt}), we have developed NMFk algorithm by coupling NMF with a custom semi-supervised k-means clustering. We have demonstrated that the number of signals can be estimated based on their reproducibility (fit) and robustness (Alexandrov and Vesselinov, 2014; Vesselinov et al., 2018, 2019). NMFk explores consecutively all possible numbers of signals k ranging from 1 to d ($k = 1, 2, \ldots, d$; $d < min(m, n)$, and then estimates the accuracy and robustness of obtained set of solutions with different number of signals.

Thus, NMFk performs l NMF runs. Each run is performed using a different number of signals, $k = 1, 2, \ldots, d$, with random initial guesses about the elements of W and H. After that, NMFk applies a custom semi-supervised k-means clustering to assign each of these l solutions to one of the k clusters. This customization of k-means clustering imposes a constraint which keeps the number of solutions in each cluster equal to the number of NMF runs l. For example, for the case with $k = 2$, after the execution of $l = 1,000$ NMF runs (performed with random initial guesses for the W and H matrix elements), each of the two clusters processes by our custom k-means will contain $1,000$ solutions. We have enforced the condition that the clusters have equal number of solutions, since each NMF simulation contributes equal number of solutions for each signal. During the clustering, the similarity between clustered solutions is measured using the cosine distance (also known as cosine similarity) (Alexandrov and Vesselinov, 2014; Vesselinov et al., 2018, 2019; Tan et al., 2016).

The main idea for estimating the unknown number of signals in NMFk is to use the separation between the clusters as a measure of how good a particular choice of k is as an accurate estimate of the number of unknown signals. We estimate the degree of clustering for different number of signals, and plot it as a function of k, we expect a sharp drop after we cross the k_{opt} value (Alexandrov and Vesselinov, 2014; Vesselinov et al., 2018, 2019).

To quantify this behavior, after the clustering, we compute the average silhouette width (Rousseeuw, 1987), $S(k)$, which is a measure of how well the solutions are clustered for given number of signals, k. The average silhouette width of the clusters for the NMFk solutions for different $S(k)$ values can be applied to evaluate the optimal number of signals, k_{opt}. In general, $S(k)$ declines as k increases. Theoretically, $S(k)$ varies between 1 and -1. For $k = 1$, $S(1) = 1$ since there is only one solution. Typically, $S(k)$ declines sharply (less than 0) after the optimal number of signals, k_{opt}, is reached.

In NMFk, in addition to the robustness, the reconstruction error (Eq. (3)) is used to evaluate the accuracy with which the derived solution $W \times H$ reproduces the observations X. In general, the solution accuracy increases (while the solution robustness decreases) with the increase of the number of unknown signals. Hence, the average silhouette width and Frobenius norm for each of the k cluster solutions can be used to define the optimal number of signals, k_{opt}. Specifically, k_{opt} can be select to be equal to the minimum number of signals that accurately reconstruct the data (i.e., the Frobenius norm is less than a given value or hit plateau) and the clusters of solutions are sufficiently robust (or stable, i.e., the average silhouette width S is larger than 0.8). The optimal number of signals is automatically estimated in the NMFk code. With a

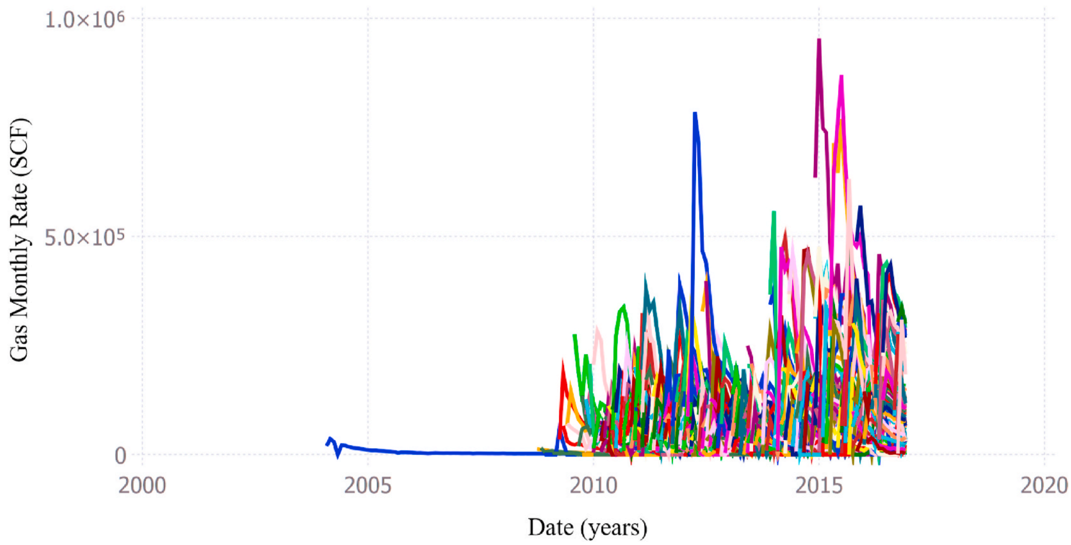

Fig. 3. The production profiles of our dataset (The majority of our wells have been completed between 2010 and 2017).

database of these signals, we predict the well performance and EUR based on the performance of the mature wells in the database.

3. Case Study data

We have selected a subset of wells from the Eagle Ford shale to showcase the capability and flexibility of our approach to forecast well performance. We got our production data from welldatabase (Welldatabase) The Eagle Ford play in Texas extends from the Mexican border through East Texas, roughly 50 miles wide and 400 miles long with an average thickness of 250 feet. It is Cretaceous in age resting between the Austin Chalk and the Buda Lime at a depth of approximately 4000 to 12, 000 feet. Our dataset includes the monthly production data from 300 gas wells. In this study, we present the results of a single-phase gas wells, however, our approach is applicable to all well and reservoir types. Unfortunately, we did not have access to the wells' geological and petrophysical parameters, which might differ among wells. However, our approach is capable of incorporating these parameters if they are available.

A typical shale production profile starts with a flow-back period where the production is dominated by the fracturing fluid, then the hydrocarbon production picks up to a peak, after that a sharp decline is observed as the well depletes the fractured media and finally a slow production decline is observed as the well depletes the ultra-low permeability matrix (Hyman et al., 2016; Mehana et al., 2019). Fig. 3 shows the monthly production profiles of our wells. The peak production rate reaches up to 0.8 MMSCF and the majority of our wells have more than 50 months of production history.

4. Results

Our framework uses the available production history to extract the decline characteristics which we use to forecast the well performance. Consequently, we expect better results as more data is used. We have designed two scenarios to showcase the capability of this framework and quantify the impact of depth of the production history on the accuracy of the results. In the first scenario (Scenario 1), we use all the production data up to 2015 to forecast one-year well performance. On the other hand, only the data up to 2012 is used in the second scenario (Scenario 2). Note that we do not take the well age into account as we extract the decline signatures.

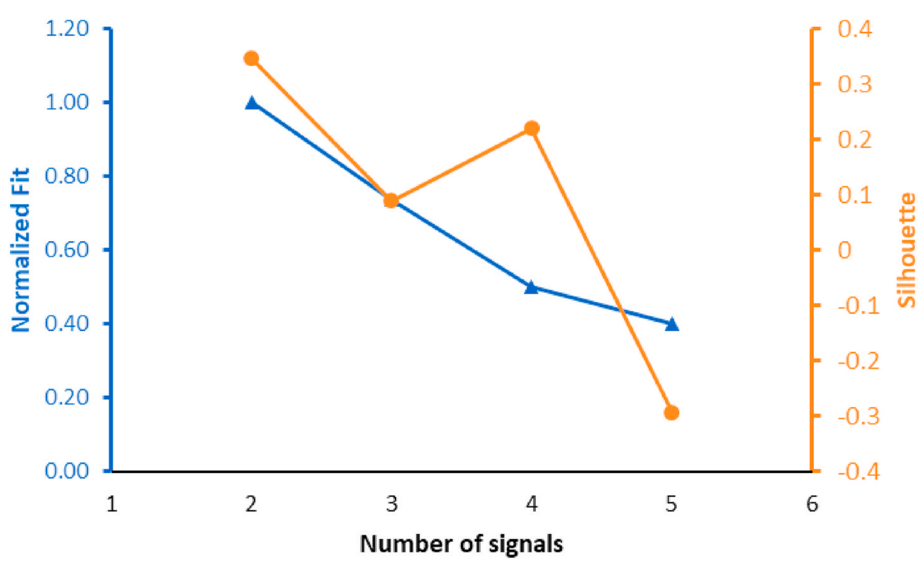

Fig. 4. The normalized fit and silhouette versus the number of signals using the dataset of the second scenario.

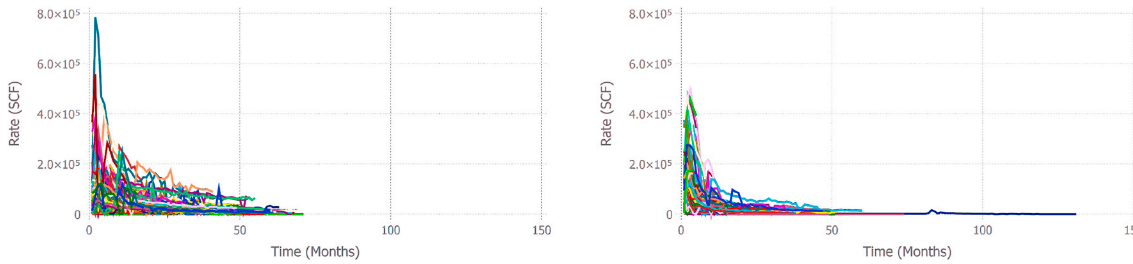

Fig. 5. The production profiles of the two classified groups. Group (A) is on the left and group (B) is on the right.

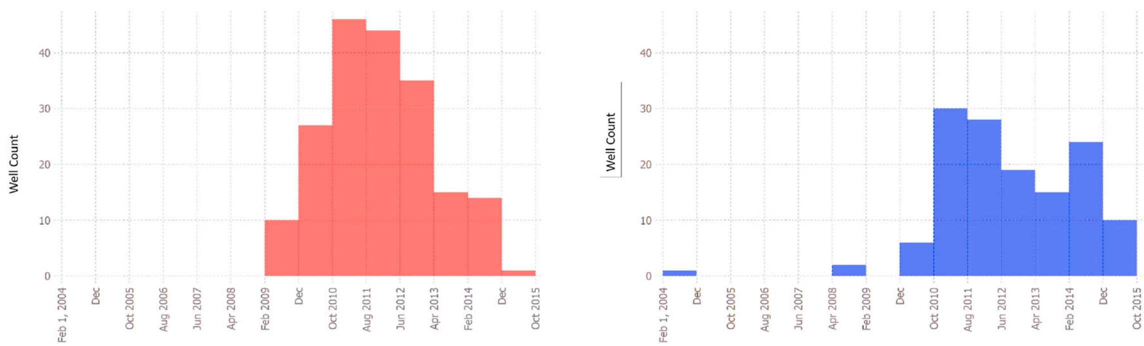

Fig. 6. The distribution of the starting date for group A and B.

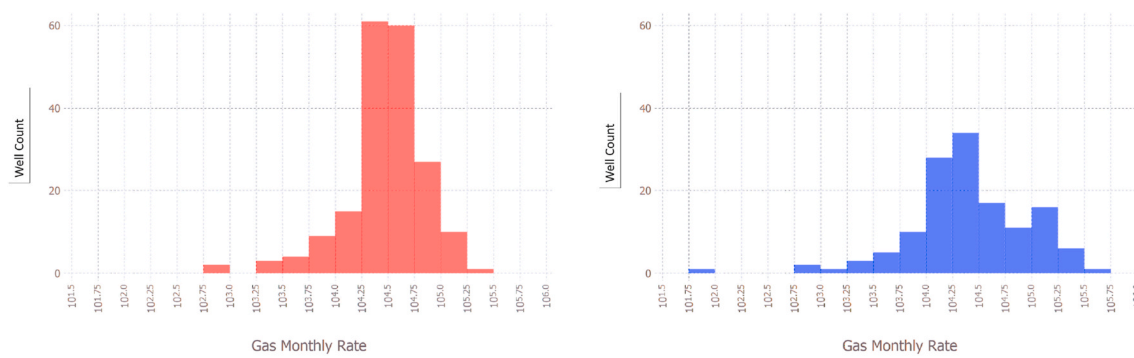

Fig. 7. The distribution of the gas monthly rate for group A and B.

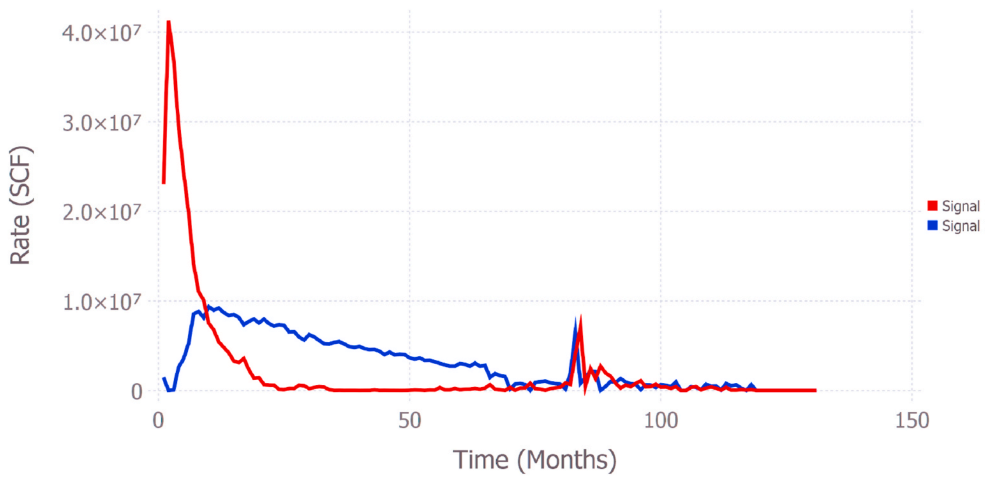

Fig. 8. The decline signals extracted from our dataset.

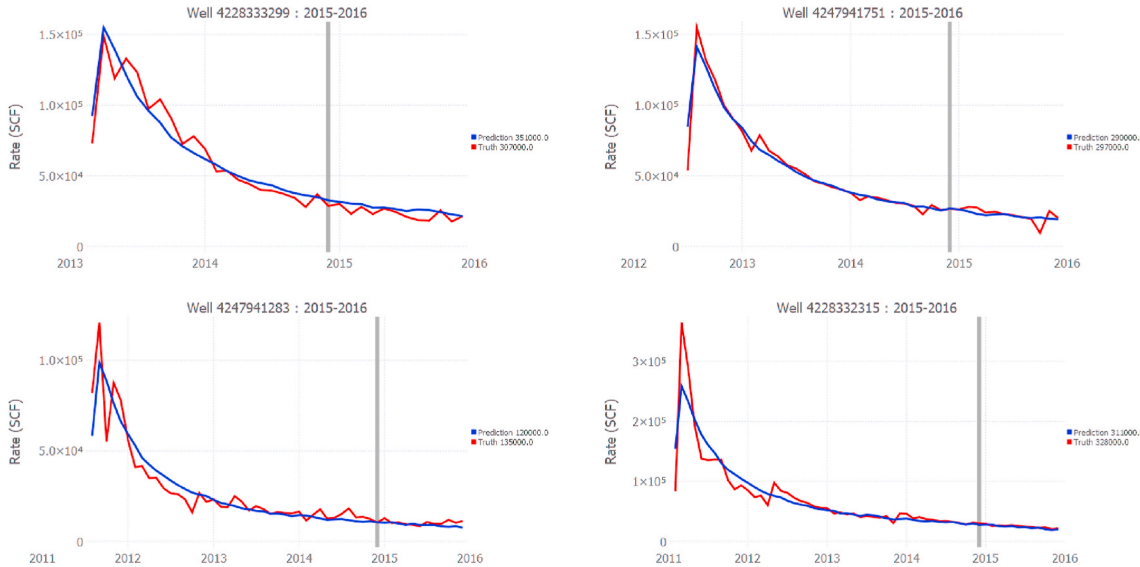

Fig. 9. Performance predictions of different wells. Note that the data before 2015 are for reference and the fit quality is estimated based on 2015–2016 predictions.

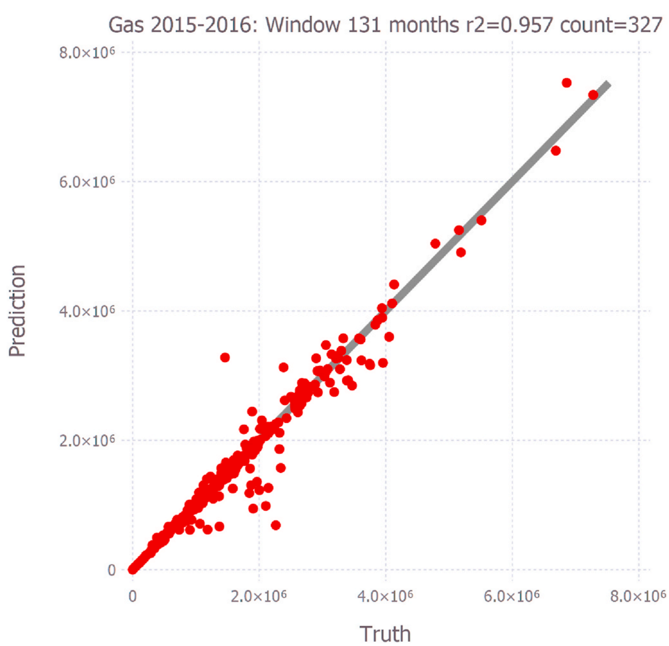

Fig. 10. Estimated vs true shale hydrocarbon production.

4.1. *Scenario 1 (longest production history)*

In this section, we used the production history up to 2015 to predict the well performance for the next year. We choose the number of signals to be two which yields the best normalized fit and silhouette values (Fig. 4). Our unsupervised machine-learning approach classified the wells into two main groups: Group A which has higher peak rates and Group B which has lower peak rates (Fig. 5). Figs. 6 and 7 present the distribution of the peak rate and the starting date for Groups A and B. We observed that Group A includes older wells with higher peak rates relative to Group B. This observation is inline with the recent literature which attributes the poor performance of the newer wells to less efficient completion schemes caused by the stress shadowing effects and parent/child well interactions (Guo et al., 2019; Liang et al., 2019; Yu et al., 2017).

With the current dataset, our framework identifies two decline signals (Fig. 8). These two signals are directly related to the two well groups discussed above: Signal 1 corresponds to Group A and Signal 2 refers to Group B. Although the actual production profile of each well would be a combination of these two signals, these signals carry physical attributes of the well. While Signal 1 exhibits rapid flow-back and decline periods, Signal 2 has slower flow-back and an almost-straight-line decline period. Thus, Signal 1 might represent properly-stimulated wells with an efficient fracture network, while Signal 2 might represent pressure-controlled wells. Both signals share a spike at late time which might correspond to the re-stimulation operations conducted in the field.

Apart from formation characteristics, the operational conditions also affect the well performance. However, without prior knowledge of these conditions, it is not possible to predict these events or adjust the forecast to reflect them. Fig. 9 compares our predictions to the actual production where we are predicting one year of performance for all the selected wells. Although these wells are in different stages of their lives, our approach accurately predicts their behavior. In contrast to the traditional techniques, there is no inherent bias in our predictions (Fig. 10). However, we cannot accommodate the changes in the operational conditions and their impact on the well performance.

4.2. *Scenario 2 (limited production history)*

In this section, we only use the production history up to 2012 to predict the future five years of production. We choose the number of signals to be four (Fig. 11). Although it does not have the best normalized fit or silhouette values, it provides the best R^2. These four unique production signals are combined to produce the actual production profiles (Fig. 12). Note that only Signal 1 describes the production profile after 30 months as the other signals phase out. Signals 1, 2 and 3 capture the flowback period; however, this period lasts longer in Signal 2 than 3 and 1. On the other hand, Signal 4 begins with a rapid decline. A clear re-stimulation signature is observed in signal 1 compared to Fig. 8.

Despite the limited production history used for predictions, our algorithm successfully forecast the well performance (Fig. 13). However, the overall comparison for all the wells has dropped from more than 0.95 R^2 in the previous case to 0.67 in the current case. It is worth noting that there is no inherit bias in our forecast (Figs. 10 and 14).

Reserve estimation is a subject of ongoing importance in the petroleum industry, particularly controlling field development-related decisions and providing valuation of corporations. Shale reservoirs are

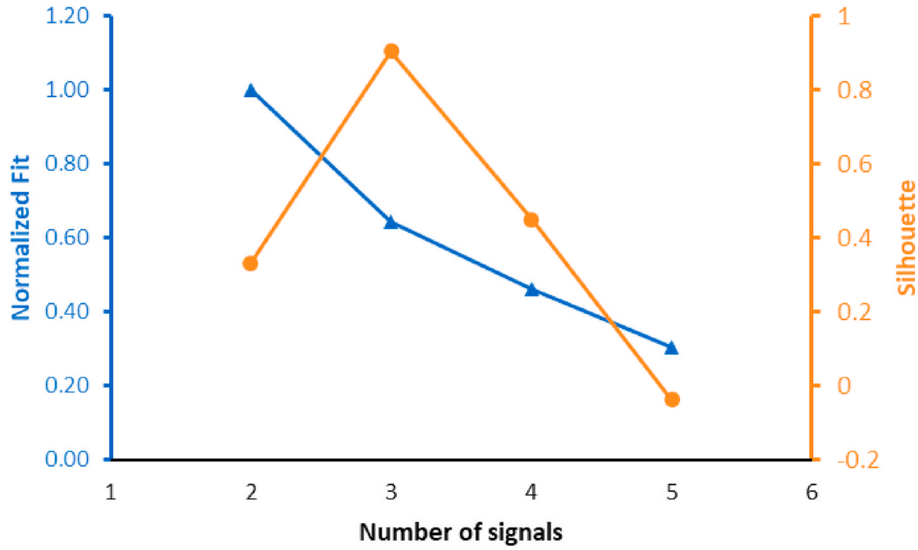

Fig. 11. The normalized fit and silhouette versus the number of signals using the dataset of the second scenario.

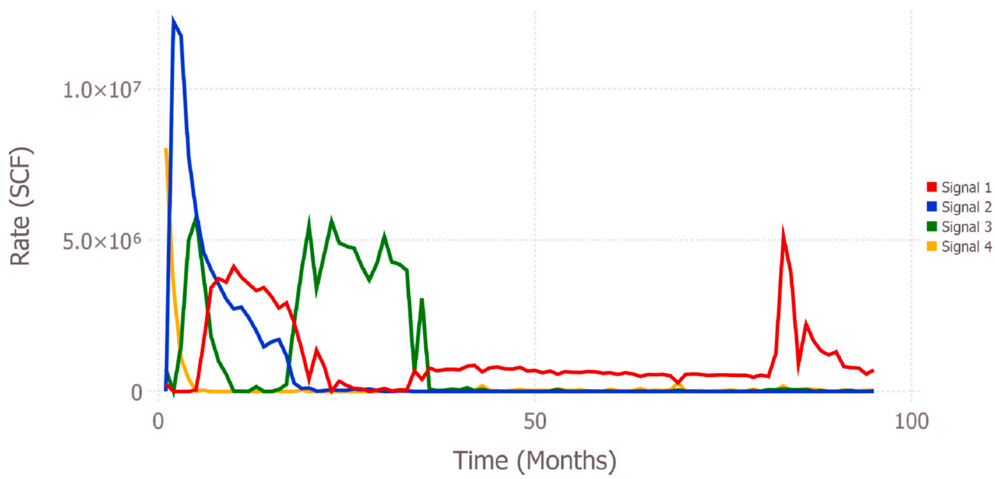

Fig. 12. The decline signals extracted from our dataset based on limited production history.

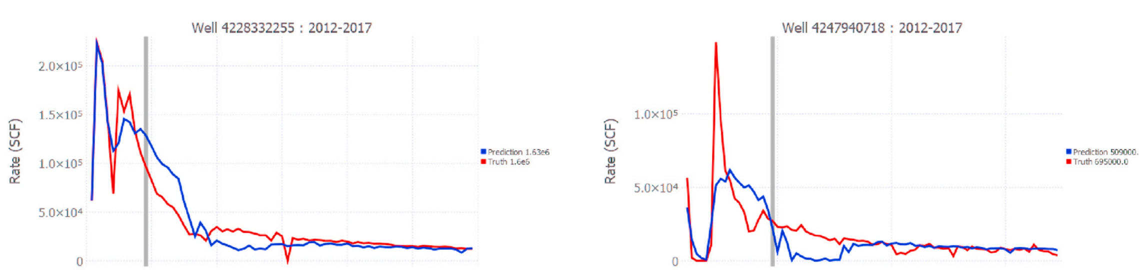

Fig. 13. Performance predictions of different wells using limited production history. Note that the data before 2012 are for reference and the fit quality is estimated based on 2012–2017 predictions.

usually completed with multistage fractures within horizontal wellbores, which results in a complex fracture network and composite flow regimes, which is a significant challenge for deterministic techniques. Recently, machine-learning techniques have revolutionized scientific research into this phenomenon, overcoming the limitations of physical models. Herein, we have developed an unsupervised machine-learning approach which was previously used to analyze reactive mixing Vesselinov et al. (2019a) and identify the contaminant sources Vesselinov et al. (2019b). We extend the capability of our approach to forecast the performance of oil & gas production well.

In general DCA is an empirical technique with assumptions that shale wells do not usually satisfy. This results in poor estimations of well performance. Because our technique utilizes a field-wide approach to forecast the performance of a single well, it does not require presumptions about the current state of the flow regime nor the operating conditions. Therefore, it enables fast and accurate estimations of the well performance. These predictions become more robust as more data is used. In essence, our approach starts by classifying the wells into groups

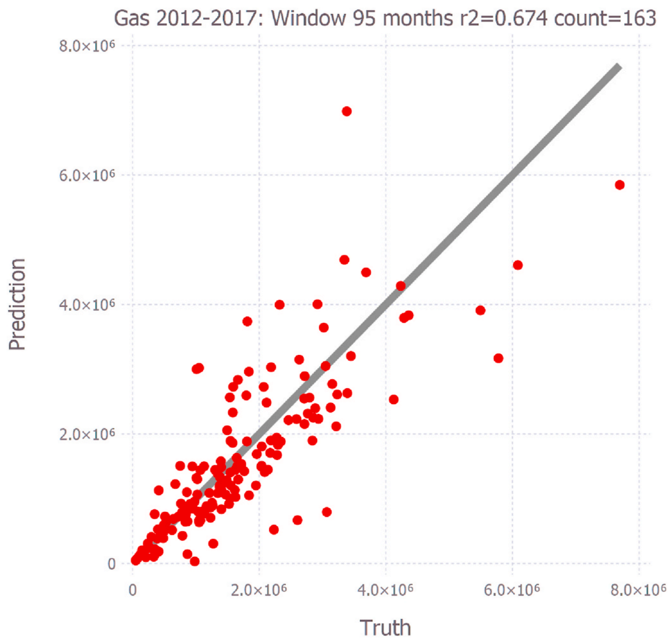

Fig. 14. Estimated vs true shale hydrocarbon production with limited production history.

based on their profile characteristics, then decouples the production profile and identifies the constituent signals. These signals could be applied to forecast the performance of newer wells. In addition, these signals could be related to well attributes or formation characteristics which enable fast identification of ideal candidates of stimulation.

5. Conclusions

We propose an efficient unsupervised machine-learning approach that allows fast and accurate real-time predictions and updates of the EUR of production wells in unconventional reservoirs. Our machine-learning framework adopts a field-wide approach that does not require flow regime analyses prior forecasting the EUR in shale reservoirs. Furthermore, it classifies the wells into categories (groups) depending on the production performance. From the well production curves, a set of unique characteristic decline signatures are extracted. These signatures can be applied to characterize the production of all the wells and used afterwards to forecast the production profile of newer wells. We found that The reliability of the predictions is highly dependent on the availability of mature wells in the analyzed reservoir dataset. We have validated our approach through hindcasting of the production data where an excellent agreement is achieved. We also have quantified the impact of the depth of the production history on the forecast quality. In addition, we found that the decline signatures usually carry information about the well attributes and the formation characteristics. For instance, we can use our approach to early identify the poorly-performing wells which could benefit from refracing simulations.

Credit roles

Conceptualization: MM, EG and VV Methodology: MM and VV Software: VV Investigation: MM, EG, VV, RM and JH Data curation: VV and MM Writing – original draft: MM Writing – review & editing: MM, EG, VV, RM, JH, QK and HV Funding acquisition: QK and VV.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded in part by the U.S. Department of Energy through Los Alamos National Laboratory's Laboratory Directed Research and Development program (LANL-LDRD).

References

Abrahamsen, P., Lia, O., Omre, H., 1992. An integrated approach to prediction of hydrocarbon in place and recoverable reserve with uncertainty measures. In: European Petroleum Computer Conference. Society of Petroleum Engineers.
Al-Rbeawi, S., 2017. How much stimulated reservoir volume and induced matrix permeability could enhance unconventional reservoir performance. J. Nat. Gas Sci. Eng. 46, 764–781.
Al-Rbeawi, S., et al., 2020. Flow-regime-based inflow-performance relationships of unconventional fractured reservoirs. SPE Prod. Oper. 35, 654–680.
Alexandrov, B.S., Vesselinov, V.V., 2014. Blind source separation for groundwater pressure analysis based on nonnegative matrix factorization. Water Resour. Res. 50, 7332–7347.
Anderson, D.M., Liang, P., Okouma Mangha, V., 2012. Probabilistic forecasting of unconventional resources using rate transient analysis: case studies. In: SPE Americas Unconventional Resources Conference. Society of Petroleum Engineers.
Arps, J.J., 1945. Analysis of decline curves. Transactions of the AIME 160, 228–247.
Ayeni, B.J., 1989. Parameter Estimation for Hyperbolic Decline Curve.
Bustin, R.M., Bustin, A.M., Cui, A., Ross, D., Pathi, V.M., 2008. Impact of shale properties on pore structure and storage characteristics. In: SPE Shale Gas Production Conference. Society of Petroleum Engineers.
Chakra, N.C., Song, K.-Y., Gupta, M.M., Saraf, D.N., 2013. An innovative neural forecast of cumulative oil production from a petroleum reservoir employing higher-order neural networks (honns). J. Petrol. Sci. Eng. 106, 18–33.
Cheng, Y., Lee, W.J., McVay, D.A., 2008. Quantification of uncertainty in reserve estimation from decline curve analysis of production data for unconventional reservoirs. J. Energy Resour. Technol. 130.
Clark, A.J., Lake, L.W., Patzek, T.W., 2011. Production forecasting with logistic growth models. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.
Clarkson, C.R., 2013. Production data analysis of unconventional gas wells: review of theory and best practices. Int. J. Coal Geol. 109, 101–146.
Duong, A.N., 2010. An unconventional rate decline approach for tight and fracture-dominated gas wells. In: Canadian Unconventional Resources and International Petroleum Conference. Society of Petroleum Engineers.
Fulford, D.S., Bowie, B., Berry, M.E., Bowen, B., Turk, D.W., 2016. Machine learning as a reliable technology for evaluating time/rate performance of unconventional wells. SPE Econ. Manag. 8, 23–39.
Gong, Y., Mehana, M., Xiong, F., Xu, F., El-Monier, I., 2019. Towards better estimations of rock mechanical properties integrating machine learning techniques for application to hydraulic fracturing. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.
Gong, Y., Mehana, M., El-Monier, I., Viswanathan, H., 2020. Proppant placement in complex fracture geometries: a computational fluid dynamics study. J. Nat. Gas Sci. Eng. 103295.
Gonzalez, R.A., Gong, X., McVay, D.A., 2012. Probabilistic decline curve analysis reliably quantifies uncertainty in shale gas reserves regardless of stage of depletion. In: SPE Eastern Regional Meeting. Society of Petroleum Engineers.
Guo, X., Wu, K., An, C., Tang, J., Killough, J., 2019. Numerical investigation of effects of subsequent parent-well injection on interwell fracturing interference using reservoir-geomechanics-fracturing modeling. SPE J. 24 (04), 1,884–1,902.
Hughes, J.D., 2013. Energy: a reality check on the shale revolution. Nature 494, 307.
Hyman, J., Jiménez-Martínez, J., Viswanathan, H.S., Carey, J.W., Porter, M.L., Rougier, E., Karra, S., Kang, Q., Frash, L., Chen, L., et al., 2016. Understanding hydraulic fracturing: a multi-scale problem. Phil. Trans. Math. Phys. Eng. Sci. 374, 20150426.
Ilk, D., Rushing, J.A., Blasingame, T.A., 2011. Integration of production analysis and rate-time analysis via parametric correlations–theoretical considerations and practical applications. In: SPE Hydraulic Fracturing Technology Conference. Society of Petroleum Engineers.
Joshi, K., Lee, W.J., 2013. Comparison of various deterministic forecasting techniques in shale gas reservoirs. In: SPE Hydraulic Fracturing Technology Conference. Society of Petroleum Engineers.
King, G.R., Lee, S., Alexandre, P., Miguel, M., Blevens, M., Pillow, M., Christie, G., 2005. Probabilistic forecasting for mature fields with significant production history: a nemba field case study. SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.
Klie, H., 2015. Physics-based and data-driven surrogates for production forecasting. In: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers.
Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. Nature 401, 788–791.
Lee, W.J., Sidle, R.E., McVay, D., 2011. Reservoir simulation: a reliable technology?. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Li, X., Chan, C., Nguyen, H., 2013. Application of the neural decision tree approach for prediction of petroleum production. J. Petrol. Sci. Eng. 104, 11–16.
Liang, B., Du, M., Yanez, P.P., 2019. Subsurface well spacing optimization in the permian basin. J. Petrol. Sci. Eng. 174, 235–243.
Ma, X., Liu, Z., 2018. Predicting the oil production using the novel multivariate nonlinear model based on arps decline model and kernel method. Neural Comput. Appl. 29, 579–591.
Macmillan, F., Hons, M., 2000. Risk, Uncertainty and Investment Decision-Making in the Upstream Oil and Gas Industry. Ph.D. thesis. Aberdeen University.
Mehana, M., Callard, J., 2018. Reserve estimation with unified production analysis. In: Unconventional Resources Technology Conference, Houston, Texas, 23-25 July 2018. Society of Exploration Geophysicists, American Association of Petroleum â€¦, pp. 691–696.
Mehana, M., El-monier, I., 2016. Shale characteristics impact on nuclear magnetic resonance (nmr) fluid typing methods and correlations. Petroleum 2, 138–147.
Mehana, M., Callard, J., Mansi, M., Gong, Y., 2019. Integrating production analysis with Monte Carlo simulation for estimated ultimate recovery eur prediction. In: SPE Eastern Regional Meeting. Society of Petroleum Engineers.
Mehana, M., Callard, J., Kang, Q., Viswanathan, H., 2020. Monte Carlo simulation and production analysis for ultimate recovery estimation of shale wells. J. Nat. Gas Sci. Eng. 83, 103584.
Middleton, R.S., Gupta, R., Hyman, J.D., Viswanathan, H.S., 2017. The shale gas revolution: barriers, sustainability, and emerging opportunities. Appl. Energy 199, 88–95.
Mudunuru, M.K., O'Malley, D., Srinivasan, S., Hyman, J.D., Sweeney, M.R., Frash, L.P., Carey, J.W., Gross, M.R., Welch, N.J., Karra, S., et al., 2020. Physics-informed Machine Learning for Real-Time Unconventional Reservoir Management, Technical Report. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).
Neil, C.W., Mehana, M., Hjelm, R.P., Hawley, M.E., Watkins, E.B., Mao, Y., Viswanathan, H., Kang, Q., Xu, H., 2020. Reduced methane recovery at high pressure due to methane trapping in shale nanopores. Communications Earth & Environment 1, 1–10.
Nguyen-Le, V., Shin, H., Little, E., 2020. Development of shale gas prediction models for long-term production and economics based on early production data in barnett reservoir. Energies 13, 424.
Ozkan, E., Brown, M.L., Raghavan, R., Kazemi, H., 2011. Comparison of fractured-horizontal-well performance in tight sand and shale reservoirs. SPE Reservoir Eval. Eng. 14, 248–259.
Passey, Q.R., Bohacs, K., Esch, W.L., Klimentidis, R., Sinha, S., 2010. From oil-prone source rock to gas-producing shale reservoir-geologic and petrophysical characterization of unconventional shale gas reservoirs. In: International Oil and Gas Conference and Exhibition in China. Society of Petroleum Engineers.
Patzek, T.W., Male, F., Marder, M., 2013. Gas production in the barnett shale obeys a simple scaling theory. Proc. Natl. Acad. Sci. Unit. States Am. 110, 19731–19736.
Qin, J., Cheng, S., He, Y., Wang, Y., Feng, D., Yang, Z., Li, D., Yu, H., 2019. Decline curve analysis of fractured horizontal wells through segmented fracture model. J. Energy Resour. Technol. 141.
Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65.
Statton, J.C., 2012. Application of the Stretched Exponential Production Decline Model to Forecast Production in Shale Gas Reservoirs. Ph.D. thesis. Texas A & M University.
Tan, P.-N., Steinbach, M., Kumar, V., 2016. Introduction to Data Mining. Pearson Education India.
Total primary energy supply (tpes) by source, world 1990-2017, https://www.iea.org/data-and-statistics,???? Accessed: 2020-01-14.
Valko, P.P., 2009. Assigning value to stimulation in the barnett shale: a simultaneous analysis of 7000 plus production hystories and well completion records. In: SPE Hydraulic Fracturing Technology Conference. Society of Petroleum Engineers.
Vesselinov, V.V., Alexandrov, B.S., O'Malley, D., 2018. Contaminant source identification using semi-supervised machine learning. J. Contam. Hydrol. 212, 134–142.
Vesselinov, V.V., Alexandrov, B.S., O'Malley, D., 2019. Nonnegative tensor factorization for contaminant source identification. J. Contam. Hydrol. 220, 66–97.
Vesselinov, V.V., Mudunuru, M.K., Karra, S., O'Malley, D., Alexandrov, B.S., 2019a. Unsupervised machine learning based on non-negative tensor factorization for analyzing reactive-mixing. J. Comput. Phys. 395, 85–104.
Vesselinov, V.V., Alexandrov, B.S., O'Malley, D., 2019b. Nonnegative tensor factorization for contaminant source identification. J. Contam. Hydrol. 220, 66–97.
Voneiff, G., Sadeghi, S., Bastian, P., Wolters, B., Jochen, J., Chow, B., Chow, K., Gatens, M., 2014. Probabilistic forecasting of horizontal well performance in unconventional reservoirs using publicly-available completion data. In: SPE Unconventional Resources Conference. Society of Petroleum Engineers.
Welldatabase, Welldatabase, ???? URL: https://welldatabase.com/.
Yu, W., Xu, Y., Weijermars, R., Wu, K., Sepehrnoori, K., 2017. Impact of well interference on shale oil production performance: a numerical model for analyzing pressure response of fracture hits with complex geometries. In: SPE Hydraulic Fracturing Technology Conference and Exhibition. Society of Petroleum Engineers.
Zhang, H., Rietz, D., Cagle, A., Cocco, M., Lee, J., 2016. Extended exponential decline curve analysis. J. Nat. Gas Sci. Eng. 36, 402–413.