

Original Paper

Production decline curve analysis of shale oil wells: A case study of Bakken, Eagle Ford and Permian



Hui-Ying Tang^{a,*}, Ge He^a, Ying-Ying Ni^a, Da Huo^b, Yu-Long Zhao^a, Liang Xue^c,
Lie-Hui Zhang^{a,**}

^a State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu, 610500, Sichuan, China

^b Stanford University, Stanford, CA, USA

^c State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China

ARTICLE INFO

Article history:

Received 21 February 2024

Received in revised form

19 July 2024

Accepted 31 July 2024

Available online 7 August 2024

Edited by Yan-Hua Sun

Keywords:

U.S. shale oil wells

Production curve

Decline curve analysis

Production prediction

ABSTRACT

The shale revolution has turned the United States from an oil importer into an oil exporter. The success of shale oil production in the U.S. has inspired many countries, including China, to begin the exploitation and development of shale oil resources. In this study, the production curves of over 30,000 shale oil wells in the Bakken, Eagle Ford (EF) and Permian are systematically analyzed to provide reference and guidance for future shale oil development. To find out the most suitable decline curve models for shale oil wells, fifteen models and a new fitting method are tested on wells with production history over 6 years. Interestingly, all basins show similar results despite of their varieties in geological conditions: stretched exponential production decline (SEPD) + Arps model provides most accurate prediction of estimated ultimate recovery (EUR) for wells with over 2 years' production, while the Arps model can be used before the two years' switch point. With the EUR calculated by decline curve analysis, we further construct simple regression models for different basins to predict the EUR quickly and early. This work helps us better understand the production of shale oil wells, as well as provide important suggestions for the choices of models for shale oil production prediction.

© 2024 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

In 2019, the shale oil production in the U.S. reached 376 million tons (U. S. Energy Information Administration, 2020), accounting for nearly 50% of its total crude oil production, which helped it become an oil and gas exporter and profoundly affecting the world energy landscape (Tsvetkova and Partridge, 2017). As a result, many countries, such as Canada, China, Russia, and Japan, start to learn from the shale development experiences in North America, leading to the rapid growth of shale oil over the world (Altawati et al., 2022). Though the shale revolution has kept on for decades in the U.S., the other countries are still in their early stages in the development of shale oil resources (Lei et al., 2023). Thus, it is necessary to analyze the shale oil wells in the U.S. with relatively long-time

production to help other countries better understand the characteristic of shale oil production.

By analyzing the production curves, researchers try to answer three questions: How the oil/gas production changes with time, what factors could affect the production and how to predict the future production accurately. Currently, most studies focus on the third question, which model can provide better prediction of production. Data-driven models (Park et al., 2021; Luo et al., 2022; Niu et al., 2022) (such as machine learning) and empirical decline curve analysis are often used for production prediction when the amount of production data is large.

For shale oil wells, most researchers choose to use data-drive models partially due to the short production history they used, when the traditional decline curve analysis fails to provide accurate prediction. In 2015, Tunstall (2015) used the U.S. shale oil quarterly dataset of 2003–2017 to verify the feasibility of their proposed nonlinear metabolism grey model combined with auto regressive integrated moving average (NMGM-ARIMA) technique in oil production prediction, which shows good accuracy in oil production

* Corresponding author.

** Corresponding author.

E-mail addresses: tanghuiying@swpu.edu.cn (H.-Y. Tang), zhangliehui@vip.163.com (L.-H. Zhang).

prediction. But the engineering data, geologic data, and decline curve analysis were not detailed presented and shown. The shale oil production data of Eagle Ford from 2006 to 2010 was used to test a bass diffusion model proposed by Wang et al. (2018) and they predicted higher daily production than the previous works. Also, the production curves of wells were not provided. Liang et al. (2019) also used the Eagle Ford data (4067 wells) to establish a multivariate linear regression model between EUR and dominant factors, fracturing length, total proppant, total vertical depth, and total organic matter, which were determined by Pearson coefficient analysis. 42 Chinese shale oil wells in Li-151 Block of the Changqing Oilfield were analyzed using hierarchical clustering analysis and principal component analysis (Wei et al., 2021) which found that the shut-in time, total fluid volume, fracturing stage number, and matrix porosity were most important factors for the shale oil production. As discussed above, different data driven models have been proposed and different conclusions on dominate factors have been obtained. The key reason for such discrepancies is the difference of data. The accuracy of statistical/machine learning models is strongly dependent on the quantity and quality of sample data. Thus, more production data is still required to further improve the model generalization performance, especially those with long production time.

Besides the above-mentioned data-driven models, decline model is also widely used in unconventional oil and gas reservoir analysis. Since the production curves of shale oil and gas wells are often characterized with early rapid decline and long tail in late stage, single decline model analysis model often fails to provide enough accurate estimation of future production. To further improve the reliability of production prediction, the combined models are proposed. Joshi and Lee (2013) believe that when the decline rate of shale gas wells is less than 5% per month, it enters the boundary dominate flow stage, at which point the Arps model could be used. For shale oil wells, when the decline rate of shale oil wells is less than 10% or 15% per month, the wells enter the boundary dominate flow stage and the turning point is approximately 24 months (Harris and Lee, 2014). Liang et al. (2020) analyzed 115 shale gas wells in the Changning Block using 18 decline models (including single and combined models), and the choice of models are suggested to be determined based on the length of production time. The coefficients in the above decline analysis models are calculated by curve fitting the whole curve, while the value of EUR is mainly determined by the late-time production. Thus, Tang et al. (2021) and Zhao et al. (2022) improved the prediction accuracy of the model by increasing the weight of the late-time production data in the fitting process of the decline curve analysis models. Some researchers choose to use the data-drive model to predict the parameters in decline curve analysis models such as Arps' (Arps, 1945) and Duong models (Zhou et al., 2023), which are then applied for shale gas and shale oil wells, respectively. In the above studies, compared with shale gas wells, the analysis of shale oil wells are limited, and the focus is primarily on the methods rather the data themselves.

To sum up, there have been many works on production decline curve analysis and data analysis, but the dataset is small (a single basin) and most are about shale gas wells. To better understand the production curves of shale oil wells as well as to find the key influence factors and proper production prediction models, we use the production data of 30,000 shale oil wells in the Bakken, Eagle Ford shale plays, and the Permian Basin in North America and carefully present and analyze the data. The distribution of engineering parameters and features of production curves are systematically demonstrated for all basins. More than 7000 wells are selected to find the most suitable decline curve analysis models for

wells with different lengths of production histories. With the calculated EUR, a multiple linear regression model between early production parameters and EUR is established for different basins to realize quick estimation of EUR of shale oil wells.

2. Introduction of shale oil basins and production data

2.1. Geological features of Bakken, Eagle Ford and Permian

Most of the shale oil in the U.S. is produced from the Bakken, Eagle Ford shale plays and Permian Basin. In this work, the oil production data from all shale plays (circled with red dashed lines in Fig. 1) will be used. Before presenting the data, a brief introduction of the shale plays is first provided.

The Bakken Formation (Sonnenberg and Pramudito, 2009; Li et al., 2015) (Fig. 1(a)), spanning across the Montana and North Dakota states, is located in the Williston Basin. The Bakken Formation, which is the target formation for shale oil production, consists primarily of sandstone, shale, and carbonate. It is located at depths of 8500 to 11,500 ft and is composed of the upper shale, middle dolomite, and lower shale layers. The Eagle Ford Formation (Alotaibi et al., 2015a; Martin et al., 2011; Walters et al., 2023; Pranesh, 2018) (Fig. 1(b)) in Texas is 50 miles wide and 400 miles long, and covers 23 counties in South Central Texas. The main production layer of the Eagle Ford Formation consists of upper layer and lower layer. The upper layer is primarily composed of calcareous shale, limestone, and quartz siltstone, while the lower layer is rich in organic materials and consists of dark shale. The main production layer of the Permian shale (Fig. 1(c)) is Wolfcamp, which is rich in organic materials and mainly composed of shale and argillaceous carbonate (Baskoro et al., 2023). The Wolfcamp Formation is further divided into 4 layers: A, B, C, and D, in which the A and B layers are the primary drilling targets for shale oil wells and are located at depths of 7300 to 10,700 ft (Bhandari et al., 2019; Nicot et al., 2020).

The geological parameters of the basins are summarized in Table 1. The pore pressure in the three shale plays is high since the pressure coefficient is over 1.2. The crude oil in Bakken and Permian is light oil with the oil gravity less than 45, while the oil gravity of Eagle Ford can reach 54.2, indicating condensate oil. Among the three shale plays, the Bakken shale play has the highest content of total organic carbon (TOC) and Permian Basin has the greatest thickness (Table 1).

The exploration and development of shale oil in the United States has begun from the Bakken Formation of the Williston Basin as early as the 1950s. Before 1985, the Williston Basin was produced by vertical wells with focusing on the Bakken and Upper Three Forks Formation, when the average daily oil rate was around 209 bbl/d. In 1987, the horizontal well was first drilled in the upper Bakken Formation of the Elkhorn Ranch Field, increasing the daily oil production to 258 bbl/d. Affected by the low oil price and unreliable EUR assessment, the shale oil development in the Williston Basin was slowed down. In 2000, with the discovery of the Elm Coulee Field, the middle Bakken Formation became the main production layer (Sonnenberg and Pramudito, 2009). After 2005, the Bakken shale oil wells can be produced economically thanks to the horizontal well and hydraulic fracturing technology. Following the success of Bakken shale, the Permian and Eagle Ford are also used for shale oil production (Martin et al., 2011) (Fig. 2). At the beginning of 2020, shale oil production accounted for approximately 66% of the total crude oil production in the United States, making it an oil exporter and significantly impact the global energy (Guo et al., 2023).

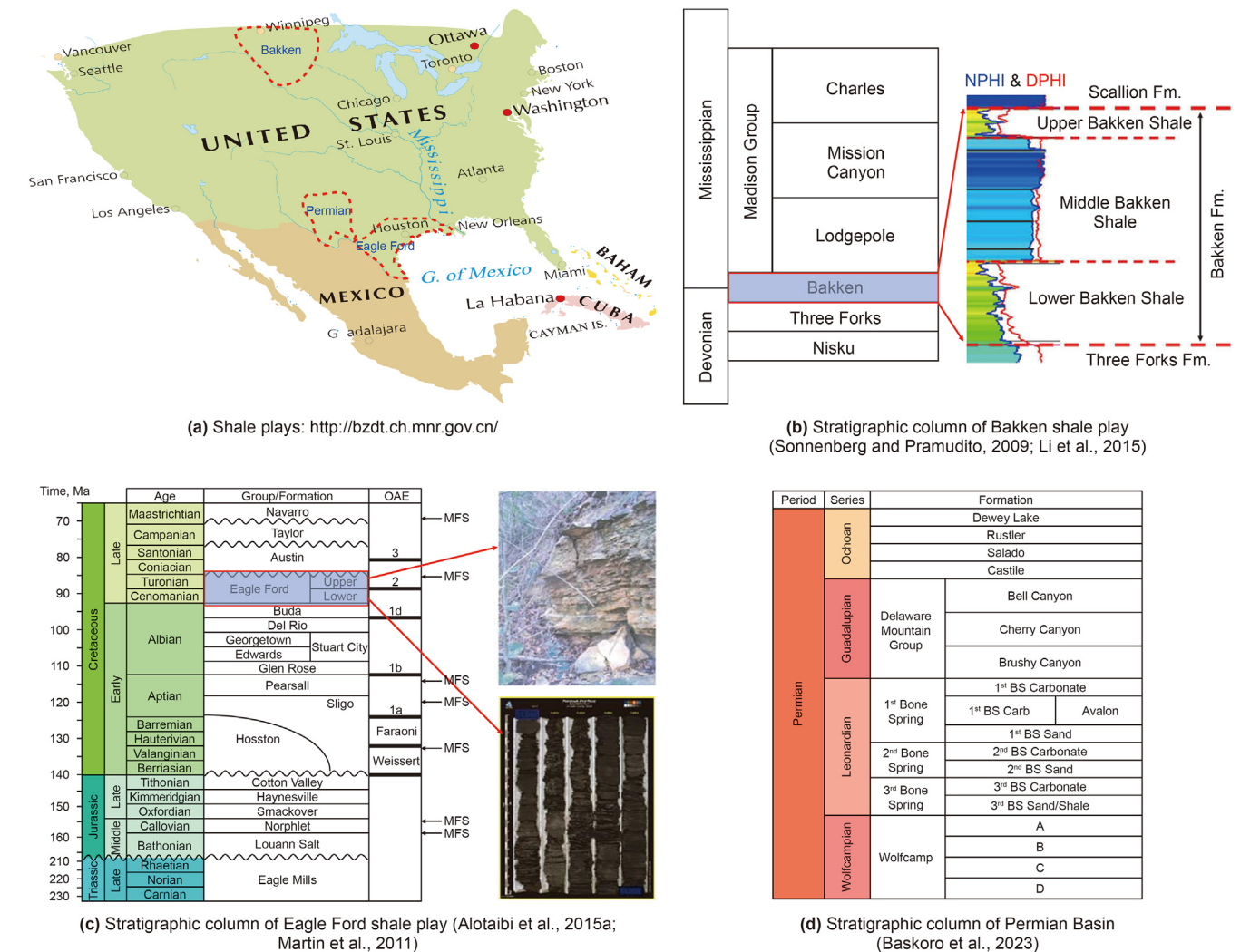


Fig. 1. The location and layers of Bakken, Eagle Ford and Permian.

Table 1			
Key parameters of Bakken, Eagle Ford and Permian (Sonnenberg and Pramudito, 2009; Li et al., 2015; Alotaibi et al., 2015a; Martin et al., 2011; Walters et al., 2023; Pranesh, 2018; Baskoro et al., 2023; Bhandari et al., 2019; Nicot et al., 2020).			
Shale play	Bakken	Eagle Ford	Permian
Depth, ft	8500–11500	4000–12000	7300–10700
Thickness, ft	10–92	70–350	131–443
Brittle mineral content, %	73.4	67.1	/
R _o , %	0.7–1.3	0.5–2.0	0.6–1.5
TOC, %	10.0–20.0	4.0–7.0	1–8
Pressure coefficient	1.3–1.6	1.3–1.8	1.5
Porosity, %	3–9	2–13	4–8
Permeability, mD	0.005–0.02	0.021–0.1	0.016–0.13
Crude oil density, g/cm ³	0.78–0.83	0.77–0.79	0.77–0.79
Oil gravity	40.4–43.4	41.0–54.2	38.0–42.5

2.2. Introduction of the dataset of the Bakken, Eagle Ford and Permian

As introduced above, the Bakken, Eagle Ford and Permian are three most important shale oil plays in the U.S. In this study, the

monthly oil, gas, water production data of over 30,000 wells, including 10,842 wells in the Bakken shale play, 12,328 wells in the Eagle Ford shale play, and 13,364 wells in the Permian Basin are analyzed. The locations of the shale oil wells in these shale plays are marked with different colors in Fig. 3(a). The number of producing wells in different years is also demonstrated in Fig. 3(b). The production of shale oil starts from 2005 in the Bakken and Eagle Ford shale plays, which means some wells have been produced for 17 years and their production trends can provide important implications for the new wells. Though the Permian Basin starts the shale oil production latest, it has the most wells. The well density is also an import parameter for development designs. We calculate the well density by summation the number of wells in a grid with area of 5.5 km × 5.5 km (Fig. 4). As shown, the well density is largest in the Permian Basin and smallest in the Bakken shale. The average density of well is 48.3 wells/30.25 km², 19.8 wells/30.25 km² and 44.6 wells/30.25 km² in the Permian, Bakken and Eagle Ford, respectively.

Here, both the total production of shale plays and the production of a standard well are presented. Eq. (1) is used to calculate the total production rate and total cumulative production, respectively:

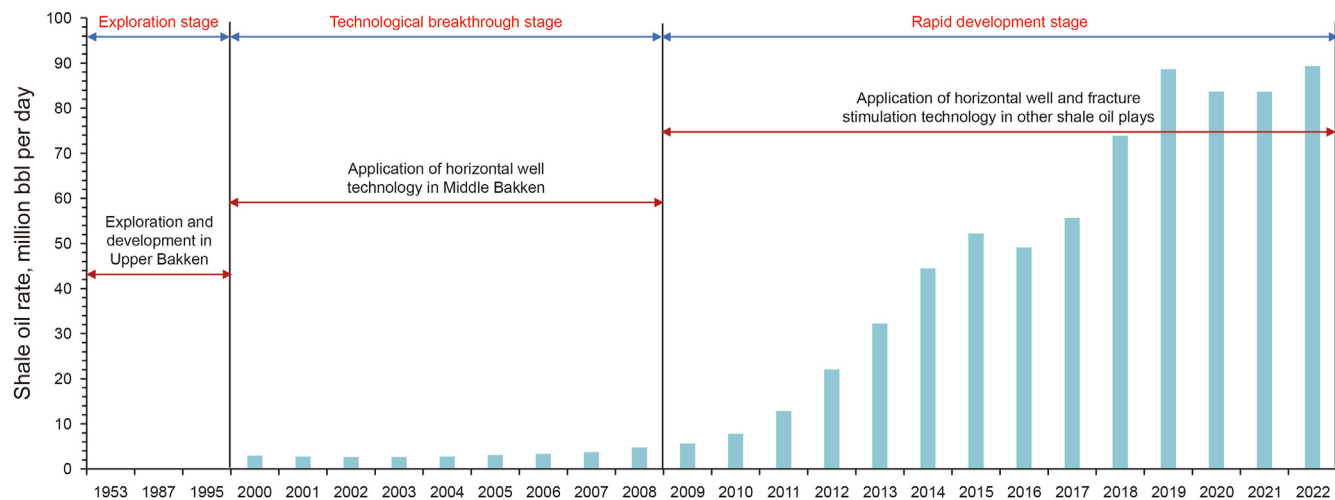


Fig. 2. The development of the United States shale oil (Guo et al., 2023).

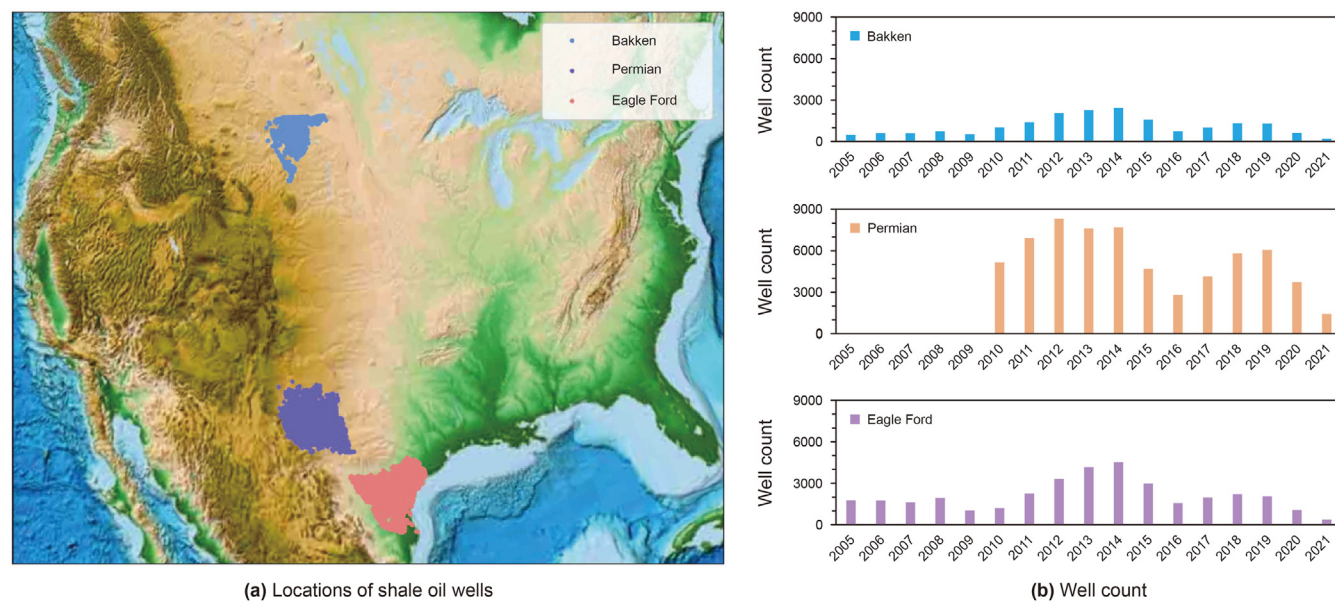


Fig. 3. Locations of shale oil wells and well count of the Bakken, Eagle Ford and Permian.

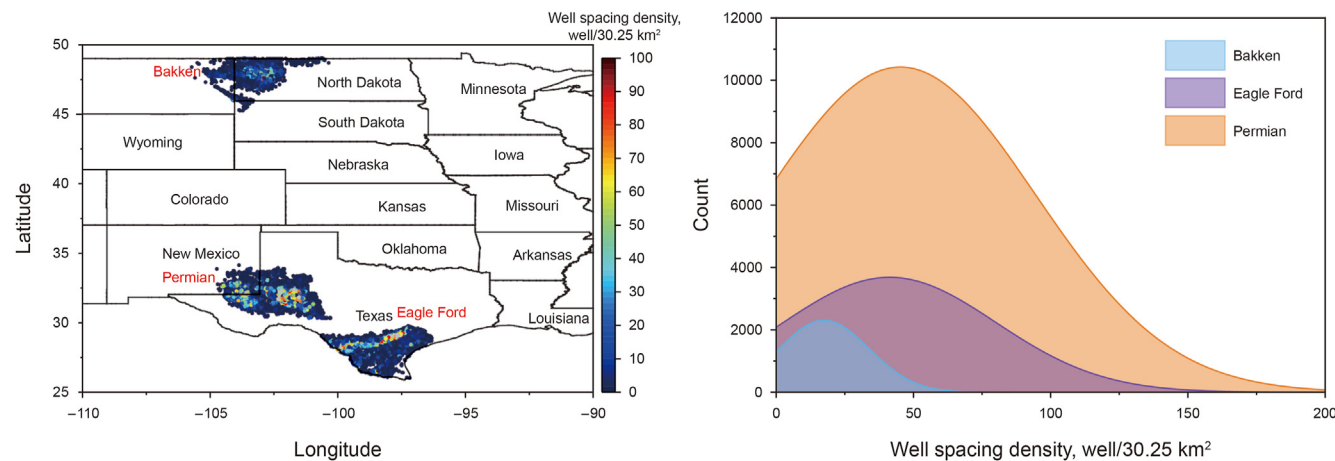


Fig. 4. Well density in the Bakken, Eagle Ford and Permian.

$$\begin{cases} q_{\text{date}} = \sum_{i=1}^n q_{i,\text{date}} \\ Q_{\text{date}} = \sum_{i=1}^n Q_{i,\text{date}} \end{cases} \quad (1)$$

where q_{date} is the total monthly oil production rate; Q_{date} is the total cumulative production; and $q_{i,\text{date}}$ is the monthly production rate of each well; n is the number of wells.

The standard well production rate and cumulative production are defined with Eq. (2) by averaging over all wells in each basin:

$$\begin{cases} q'_t = \sum_{i=1}^n q_{i,t} / n \\ Q'_t = \sum_{i=1}^n Q_{i,t} / n \end{cases}, \quad t = 1, 2, 3, \dots, m \quad (2)$$

where q'_t is the monthly production rate of a standard well at time t ; Q'_t is the cumulative production rate of a standard well at time t ; $q_{i,t}$ is the monthly production rate of each well at time t .

The production of oil, water, and gas are presented in Figs. 5–7. As can be seen from the basin production, the oil production is the largest in the Bakken shale play, while the water production and gas production are the largest in the Permian and Eagle Ford, respectively. Similar with Fig. 2, the first reduction of shale oil production starts at the second quarter of 2015 when the oil price plunge occurred, accompanied with the decrease of well count (Fig. 3(b)). The production, especially the production in the Permian Basin, starts to grow from 2017. But affected by the COVID-19, the production in all basins began to decline again since 2019 (many wells are shut in).

Like the shale gas wells (Patzek et al., 2014), the oil production curves of standard wells in all shale plays demonstrate rapid decline in the early stage and a long tail in the late stage (Fig. 5). The water produced in Permian is nearly twice as other basins (Fig. 6) and the Eagle Ford produces the most gas (Fig. 7). The large production of water (pore space blockage) and gas (high gas to oil ratio reduces oil production) often plays negative role in oil production, which could be one of the reasons that the Bakken has the highest oil production. The changes of monthly oil production rate are demonstrated in Fig. 8. As shown, the oil production decrease quite fast in the first three years that the monthly oil production rate after three years is only 18.9%, 11.1%, and 11.8% of peak oil in the Bakken, Eagle Ford, Permian shale, respectively. After ten years' production in the Permian and Eagle Ford shale, the oil rate is below 3% of the peak oil, while the Bakken wells still have 5% peak oil rate after

fifteen years' production.

Besides the production curves, some key fracturing parameters are also collected, including perforation length (Fig. 9(a)), stage count (Fig. 9(b)), proppant per perforated foot (Fig. 9(c)), fluid per perforated foot (Fig. 9(d)), and the buried depth (Fig. 9(e)). Since some parameters are not collected for certain wells, the number of points in Fig. 9 is bit different. As shown in Fig. 9(a) and (b), the perforation length and stage count (strongly correlated with the perforation length) in Bakken shale is significantly larger than others. The well length in the west and east side of Permian, the north of Eagle Ford is relatively high. By direct observation (Fig. 9(h)), the oil peak in the regions with longer perforation is greater. Though the length of perforation in Bakken is longer, more fluid and proppant are injected in the Permian and Eagle Ford shale (Fig. 9(c) and (d)). The oil peak in the regions of large fluid and proppant injection in the Permian and Eagle Ford can even exceed that in the Bakken shale wells (Fig. 9(h)). The shale wells are buried at depth of 5000–15000 ft (Fig. 9(e)) and the average depth of wells in Bakken is the largest. The variation of well depths in Permian and Eagle Ford shale is more significant than Bakken shale. Assuming the water produced in the first year all comes from the fracturing fluid, we calculated the first month flow back ratio in all shale plays, which are 5%–30%, 5%–25% and 10%–60% in the Bakken, Eagle Ford and Permian shale, respectively (Fig. 9(f)). The flow back ratio in Permian Basin is higher than others. The value of gas oil ratio (GOR) is much greater in the Eagle Ford and the GOR in Bakken is small (Fig. 9(g)). According to the definition of oil types (Alotaibi et al., 2015b) with GOR, the crude oil in the Bakken are mainly black oil and volatile oil. The crude oil in the Permian Basin are mainly black oil, volatile oil and condensate oil. The crude oil in the Eagle Ford covers all fluid types because of the high organic maturity (Table 1). The peak oil rate ranges from 0.5×10^4 to 4×10^4 bbl/month (Fig. 9(h)), where the average oil production of Bakken shale is the best among these three shale plays.

To quantitatively evaluate the impact of different parameters on production, the Pearson correlation coefficients are calculated (Fig. 10). As shown, the well length, stage count, fluid and proppant injected per perforation foot are strongly positively correlated with the peak oil in all shale plays, especially in the Permian Basin. As analyzed above, using long horizontal well or large injection can both obtain high oil production rate. Besides, the buried depth is also positively related to the oil production, which is related to the depth of the primary production layers (Fig. 1). Also, we can see that the larger the well density, the higher the peak oil. This is because that more wells will be drilled in the areas with good reservoir quality. The average first month flowback ratio in the Bakken, Eagle Ford and Permian shale are 12.4%, 6.04%, and 19.36%, respectively,

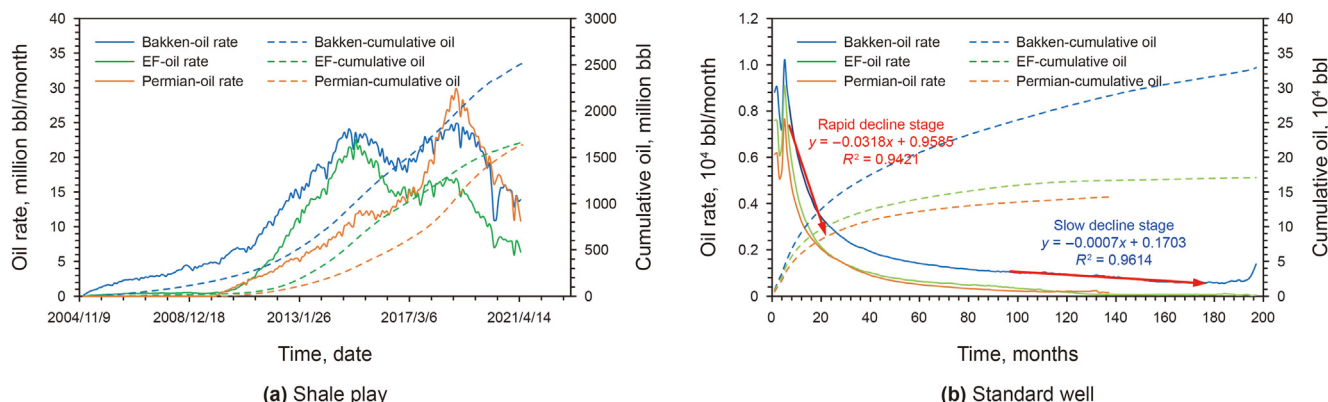


Fig. 5. Monthly oil rate and cumulative oil production for shale play and standard well. EF: Eagle Ford.

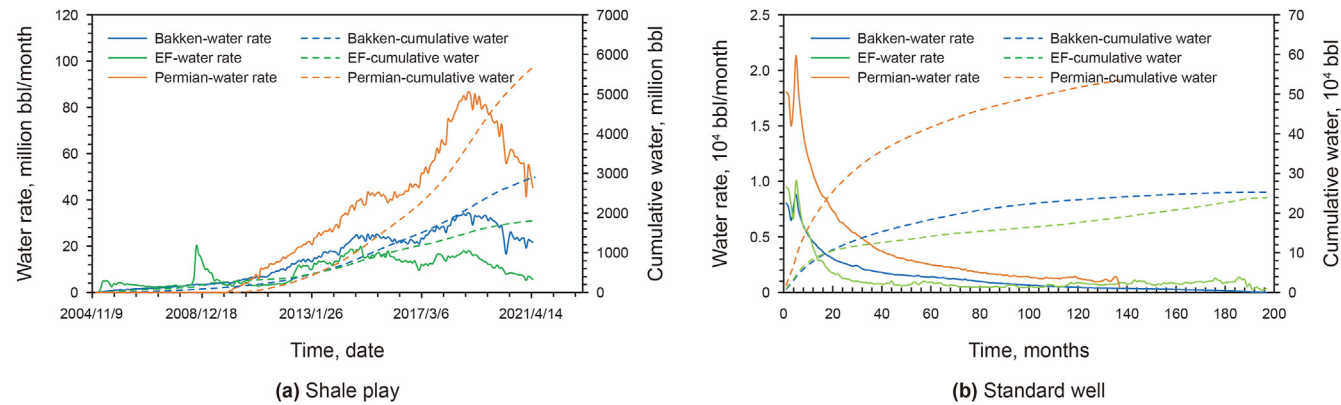


Fig. 6. Monthly water rate and cumulative water production for shale play and standard well.

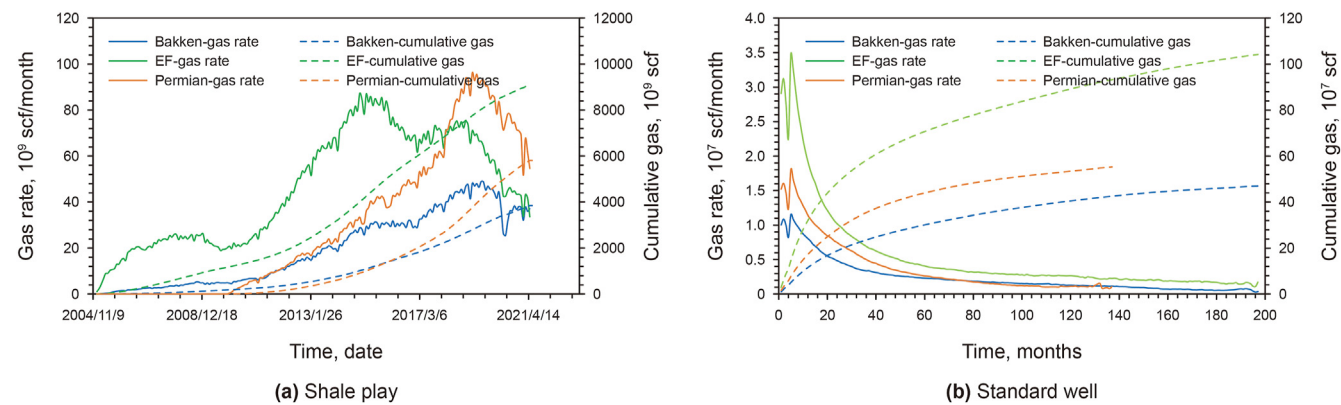


Fig. 7. Monthly gas rate and cumulative gas production for shale play and standard well.

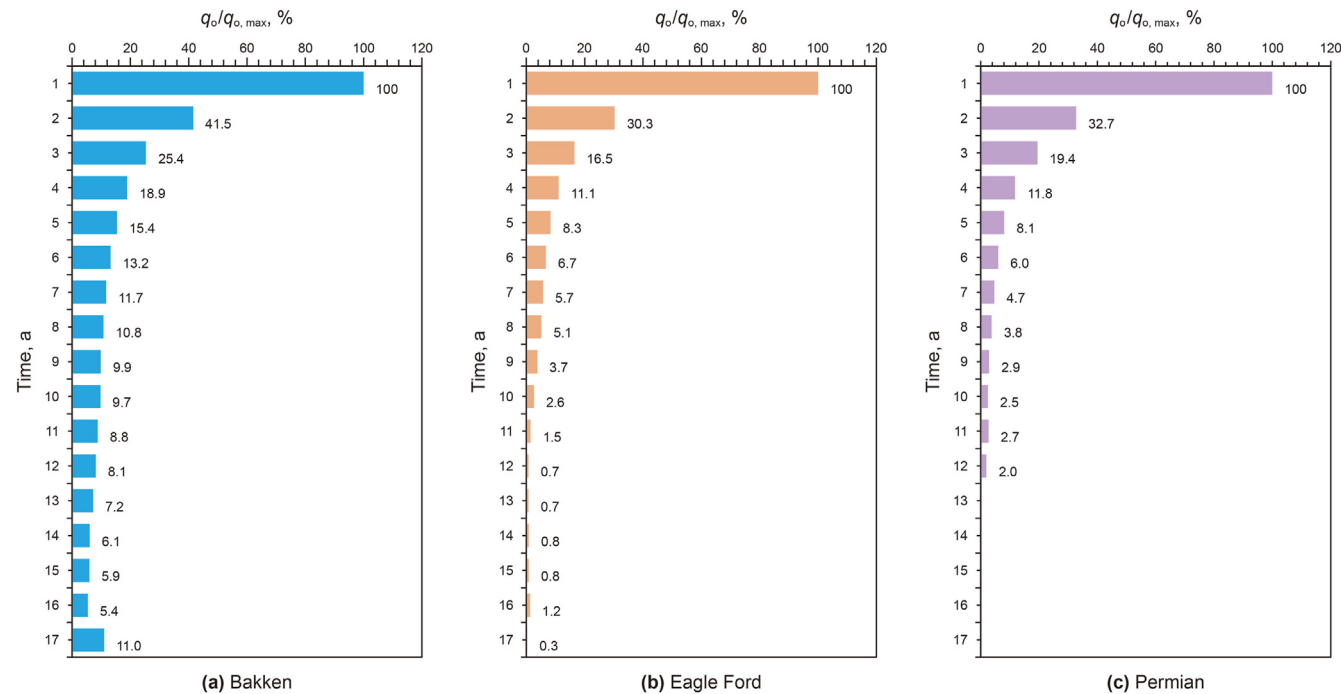


Fig. 8. Standard well monthly oil production normalized by the peak oil rate.

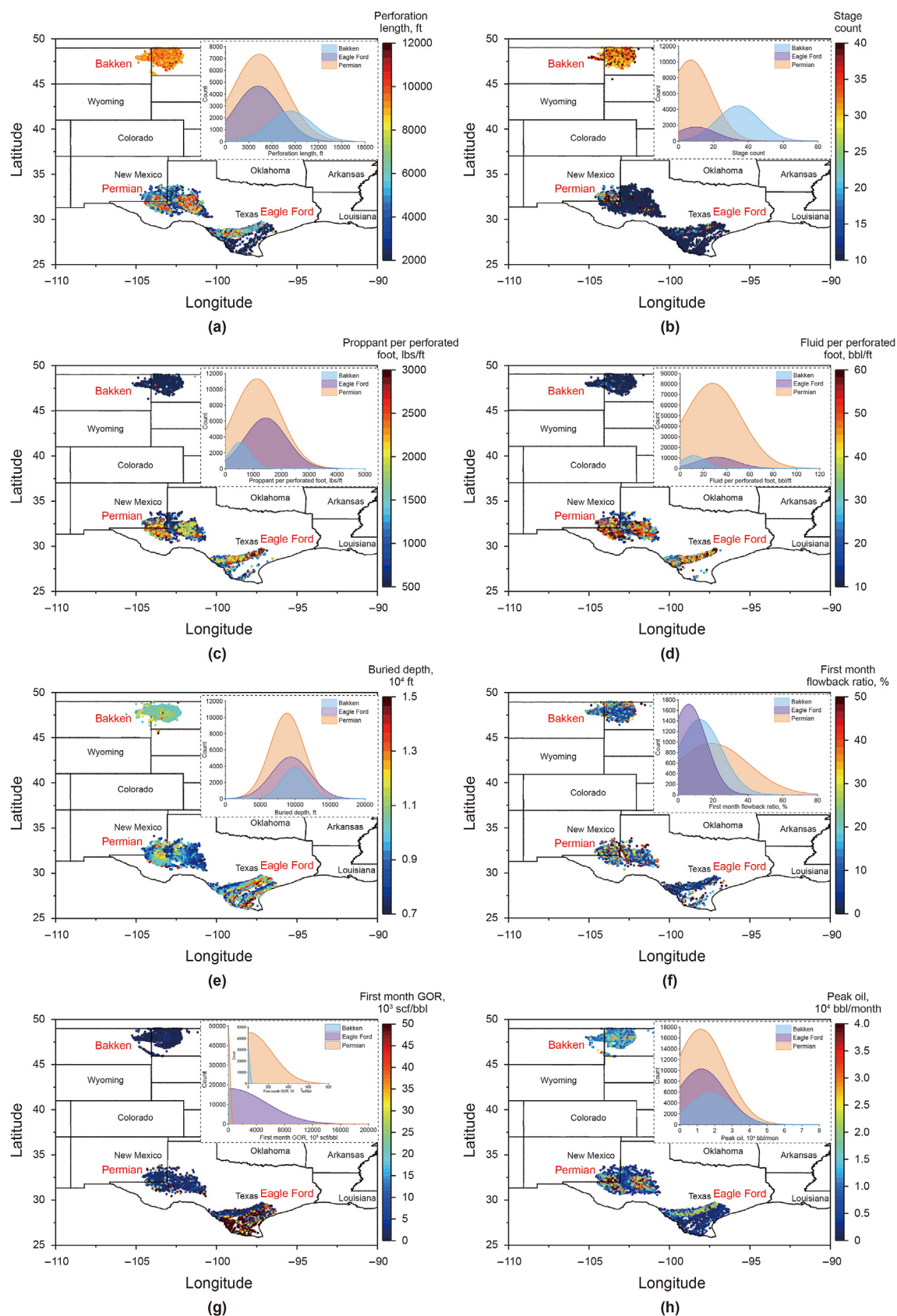


Fig. 9. Distribution of fracturing and production parameters in all shale plays.

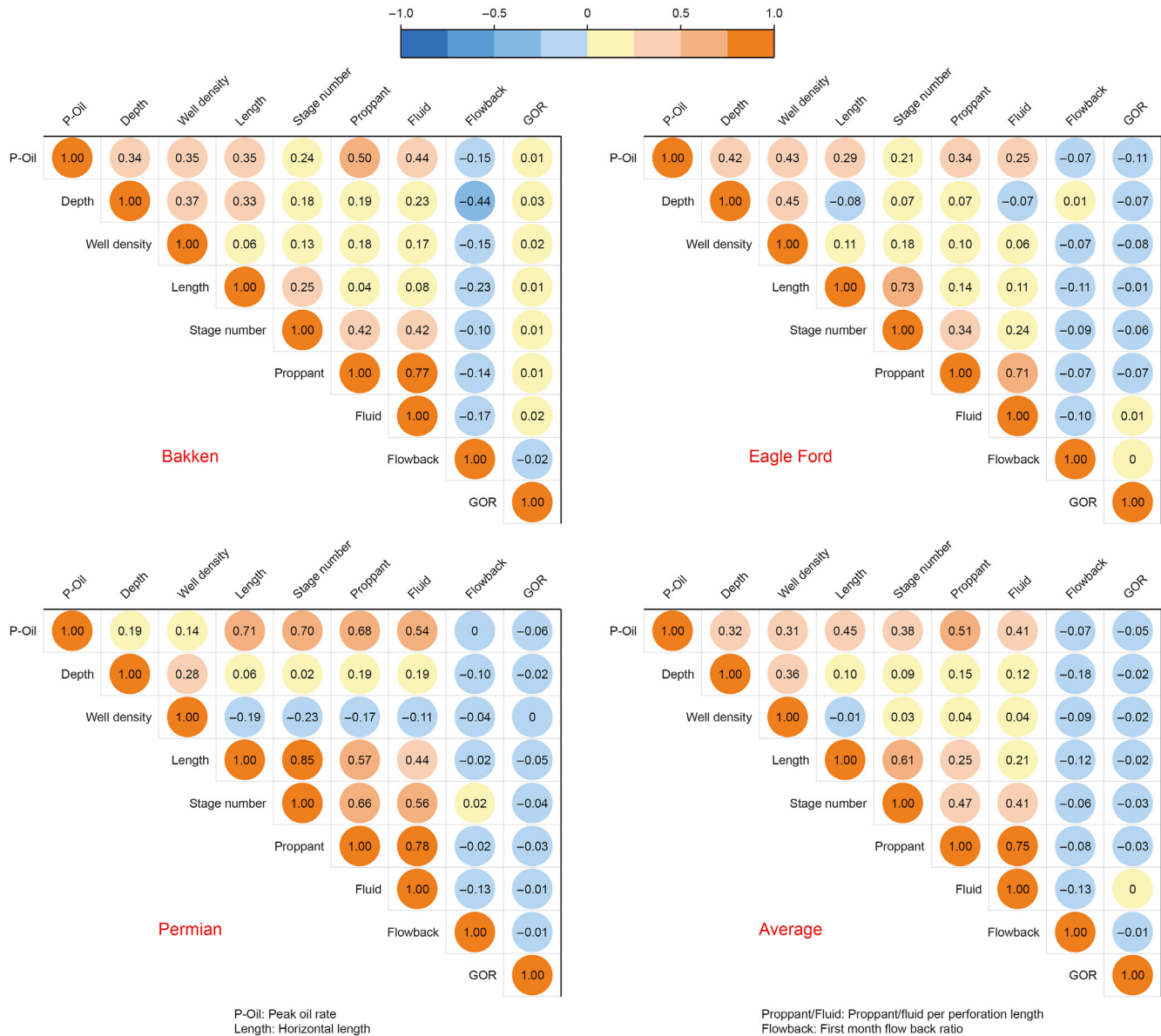


Fig. 10. Pearson correlation coefficients between different parameters.

which shows negative correlation with peak oil, similar with the observations in shale gas reservoirs (Jiang et al., 2021; Ghanbari et al., 2013). Peak oil was most negatively correlated with the first month GOR in the Eagle Ford and Permian shale, while a slight positive correlation was observed in the Bakken shale play. Among the nine parameters analyzed here, there is a significant positive correlation between the perforation length and the stage number, as well as between the proppant per perforated foot and the fluid per perforated foot. Based on the above analysis, parameters such as the stage number and fluid per perforated foot are ignored to reduce overfitting in the EUR rapid prediction model in Section 4.

3. Introduction and application of decline curve analysis models

3.1. Introduction of decline curve analysis models

Since the production data we collected only includes monthly oil, gas and water rate, the decline curve analysis models are chosen

for production analysis and prediction. Currently, the decline curve analysis models (written as decline model in the following) are divided into four types: (1) Classic Arps (1945) model and its modified models, such as Matthews–Lefkowitz (ML) model (Matthews and Lefkowitz, 1956). (2) Decline models based on the flow regimes of unconventional oil and gas wells, such as Extended КОПЫТОВ (Extended K) model (Liang et al., 2020), power law exponential model (PLE) (Ilk et al., 2008), Duong's model (Duong, 2010) and Logistic growth model considering shale gas reservoirs and development characteristics (RB-LGM) (Zhao et al., 2020). (3) Combined models, such as Duong + Arps model (Liang et al., 2020), typically take on a "two-stage" form, with the first model suitable for the linear flow stage, and the second model usually being Arps to describe the boundary dominated flow stage. In this paper, different kinds of combined models are used and compared as shown in Table 2. (4) Decline models with data-driven, such as decline curve analysis (DCA) model with long short-term memory (LSTM) neural network (DCA + LSTM) (Xue et al., 2023). Some widely used decline models are summarized in Table 2, consisting

Table 2
Some typical decline curve analysis models.

Method	Model	Basic structure	Number of fitting parameters	Model parameter annotation	Suitable flow regime
Arps model (Arps, 1945)	$q = q_i \exp(-D_i t), n = 0$	Exponential function	2	q_i : Initial production rate	Boundary dominated flow
	$q = q_i / (1 + n D_i t)^{1/n}, 0 < n < 1$	Power function	3	D_i : Initial decline rate	
	$q = q_i / (1 + D_i t), n = 1$	Power function	2	n : Loss rate	
ML model (Matthews and Lefkovits, 1956)	$q = q_i (a t / n + 1)^{-1/n}$	Power function	3	a and m : Tuning parameters	Boundary dominated flow
Extended K model (Liang et al., 2020)	$q = q_i [b / (b + t)]^2$	Power function	2	b : Fitting coefficient	
Extended Weng model (Liang et al., 2020)	$q = a \cdot t^b \cdot \exp(-t/c)$	Exponential function and power function	3	a, b , and c : Fitting coefficients	
PLE model (Ilk et al., 2008)	$q = q_i \cdot e^{-D_\infty \cdot t - \frac{D_i t^n}{n}}$	Exponential function	4	D_∞ : Decline rate when the production time approaches infinity	Linear flow
	$D = D_\infty + D_i t^{n-1}$				
SEPD model (Valko, 2009)	$q = q_i \cdot e^{-(t/\tau)^n}$	Exponential function	3	τ : The median of the characteristic number of time	
				n : Empirical constant	
Duong model (Duong, 2010)	$q = q_i \cdot t^{-m} \cdot e^{\frac{a}{1-m}(t^{1-m} - 1)}$	Exponential function and power function	3	Q : Cumulative production	
	$q/Q = a t^{-m}$			a and m : Tuning parameters	
Variable decline modified Arps (VDMA) model (Gupta et al., 2018)	$q = q_i \cdot \exp[-D_i \cdot t^{(1-a)}]$	Exponential function and power function	3	a : Decline index	
Li model (Wang et al., 2017)	$q = q_i \cdot e^{-\lambda(\ln t)^2}$	Exponential function	2	λ : Empirical coefficients	
PLE + Arps (Liang et al., 2020)	$\begin{cases} q = q_i \cdot e^{-D_\infty \cdot t - \frac{D_i t^n}{n}}, & t \leq t_1 \\ q = \frac{q_i}{(1 + n D_i t)^{1/n}}, & t > t_1 \end{cases}$	Exponential function and power function	7	t_1 : Switch point	Linear flow and boundary dominated flow
SEPD + Arps (Joshi and Lee, 2013)	$\begin{cases} q = q_i \cdot e^{-(t/\tau)^n}, & t \leq t_1 \\ q = q_i / (1 + n D_i t)^{1/n}, & t > t_1 \end{cases}$		6		
Duong + Arps (Joshi and Lee, 2013)	$\begin{cases} q = q_i \cdot t^{-m} \cdot e^{\frac{a}{1-m}(t^{1-m} - 1)}, & t \leq t_1 \\ q = q_i / (1 + n D_i t)^{1/n}, & t > t_1 \end{cases}$		6		
Arps hyp + Arps exp (Arps hyperbolic model + Arps exponential model) (Liang et al., 2020)	$\begin{cases} q = q_i / (1 + n D_i t)^{1/n}, & t \leq t_1 \\ q = q_i e^{-D_i t}, & t > t_1 \end{cases}$		5		
VDMA + Arps	$\begin{cases} q = q_i \cdot \exp[-D_i \cdot t^{(1-a)}], & t \leq t_1 \\ q = q_i / (1 + n D_i t)^{1/n}, & t > t_1 \end{cases}$		6		
Li + Arps	$\begin{cases} q = q_i \cdot e^{-\lambda(\ln t)^2}, & t \leq t_1 \\ q = q_i / (1 + n D_i t)^{1/n}, & t > t_1 \end{cases}$		5		

mainly of exponential functions and power functions, with typically three parameters for the single model.

3.2. Application of decline models for shale oil wells with different production history

The models in Table 2 are then used to predict oil production in the target shale plays. In this section, we try to find the most suitable decline curve analysis models for shale oil wells. According to our experience in shale gas wells (Tang et al., 2021) when the length of production history changes, the most suitable model could be different. So, we further consider the impact of producing time on model choice. According to the work of Harris and Lee (2014), the switch point of combined model is $t_1 = 2a$, indicating that the combined models can only be used for wells with more 2 years of production. To test the accuracy of decline models, the production history of test well should also be long enough. So, we only select the wells with production history over 6 years and show obvious decline trend for further analysis. With the above criteria, 2792 wells in Bakken shale play, 2256 wells in Eagle Ford shale play and 2192 wells in Permian Basin are selected.

The production data is divided into a train set and a validation set. The first 6, 12, 18, 24, 30, 36, 42, and 48 months' production history is used as the train set, while the remaining production data is taken as the validation set. The train set is used to fit the parameters of the decline model, and the validation set is used to compare the prediction accuracy of different models. Relative error δ (Eq. (3)) is used to evaluate the prediction accuracy.

$$\delta = \frac{Q_{\text{model}} - Q_0}{Q_0} \times 100\% \quad (3)$$

where Q_{model} is the cumulative production calculated by decline models, bbl; Q_0 is the actual cumulative production, bbl. When the δ is greater than 0, it indicates that the model prediction results are higher than the actual. When the δ is less than 0, it indicates that the model prediction results are lower than the actual. The smaller the absolute value of δ , the higher the model accuracy.

To improve the prediction accuracy of the decline model, a modified nonlinear regression algorithm is adopted in this paper. We have applied this method to Changning shale gas wells in the Sichuan Basin, and the results show that this method can effectively capture the trend of later dynamic data (Zhao et al., 2022). The steps

are as follow: (1) remove data points with production rate of 0 from shale oil well production data; (2) calculate the Euclidean distance between each production point and the initial point during the production decline stage; (3) use the Euclidean distance in Step (2) as the weight of each production point to fit the parameters of the decline model with weighted least square method. Different fitting methods are compared in Fig. 11, indicating that the modified method can improve the prediction accuracy of decline models.

From Fig. 12, we can see that the values of δ of PLE, SEPD, VDMA, KM, Li and Weng models are most negative, while the δ of combined models (except Arps hyp + Arps exp) are positive, indicating the combined models tend to over-estimate the production of shale oil. Fig. 13 shows the fitting results of 15 decline models. The discontinuity phenomenon occurs at the combined point due to the unequal rate fitting strategy. Table 3 summarizes the optimal model for wells with different production history in all shale plays. As shown, when production history is less than 24 months, the best model of three basins is Arps, with average δ value of 27.25%, 23.49%, and 26.85%, respectively. When the production history is over 24 months, the optimal combined model for all shale plays is SEPD + Arps model, while the optimal single model is different among shale plays.

Besides, we can also see from Fig. 12 that the longer the production history, the higher the accuracy of decline models. We select the model that corresponds to the minimum absolute value

of relative error and calculate the EUR of all selected wells. The calculated EUR are shown in Fig. 14(a). Similar with the values of oil peak, the EUR in Bakken shale play is the highest among all shale plays. From Table 1 and Section 2.2, we believe that the high TOC and long perforation length could be the primary reasons for the high EUR in Bakken shale play. Though the highest oil peak in Permian Basin seems a bit larger than Eagle Ford (Fig. 9(h)), the calculated EUR in Eagle Ford is greater than the Permian Basin, which is consistent with the standard well production data (Fig. 5).

As shown in Figs. 14(b) and 15(a), EUR, peak oil production and first year cumulative oil production are significantly positively correlated in log-log coordinates, while a higher correlation between peak oil production and EUR. Besides, we can see from Fig. 15(b) that the correlation of other factors with EUR, peak oil, and first-year cumulative production is consistent and it is reasonable to choose the peak oil rate as the indicator of well oil production effectiveness.

The ratio between cumulative oil production and EUR are then calculated to reveal the production potential of wells with different producing history (Fig. 16). Half of the EUR can be produced in the first 2–3 years. To produce 90% of the EUR, it would take 15, 12, and 15 years for wells in the Bakken, Eagle Ford and Permian shale plays respectively. The data in Fig. 16 can help us quickly estimate how much oil can be produced in the future and make plans of new wells to maintain the oil production in the shale plays.

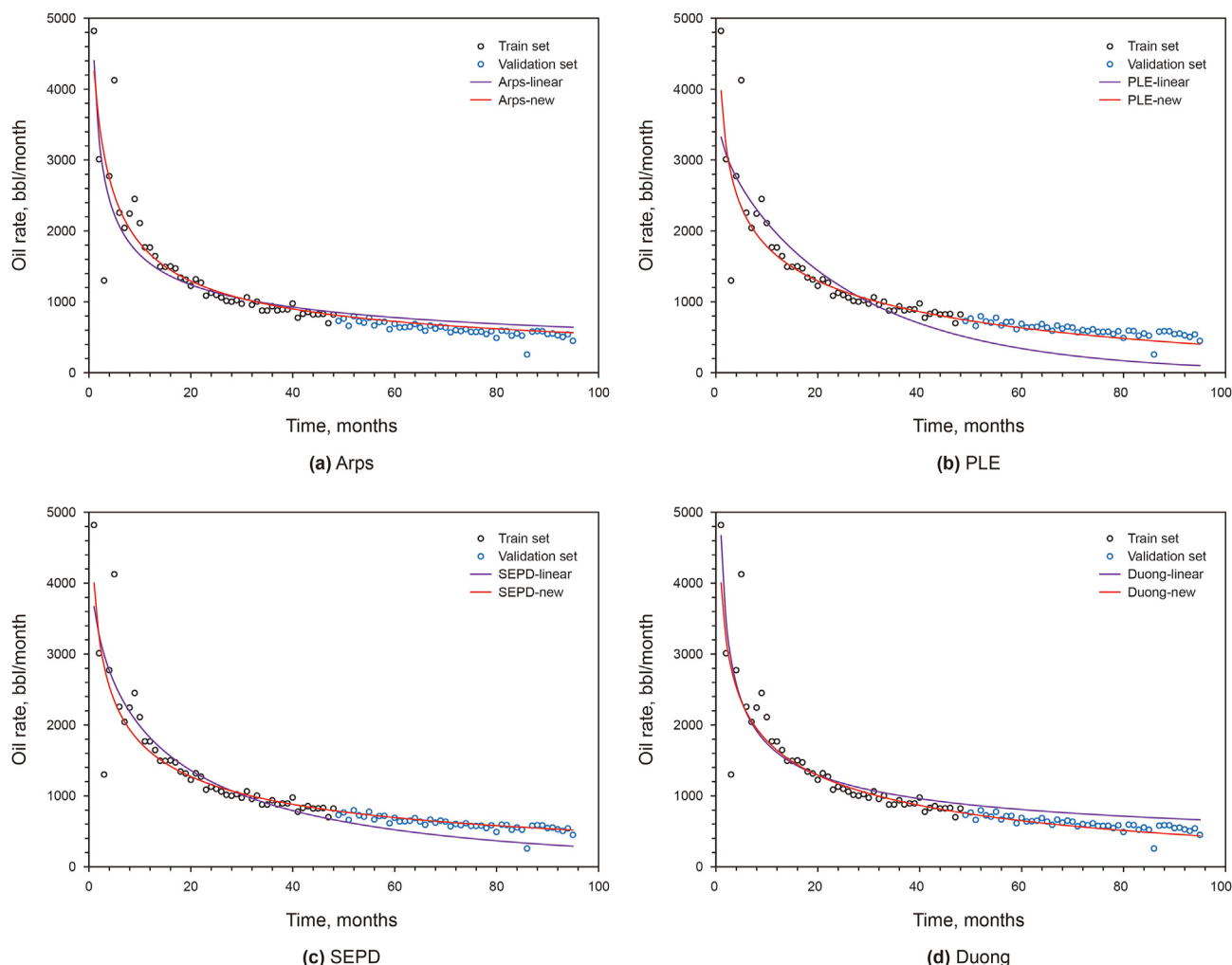


Fig. 11. Comparisons of the linearization method and the new method.

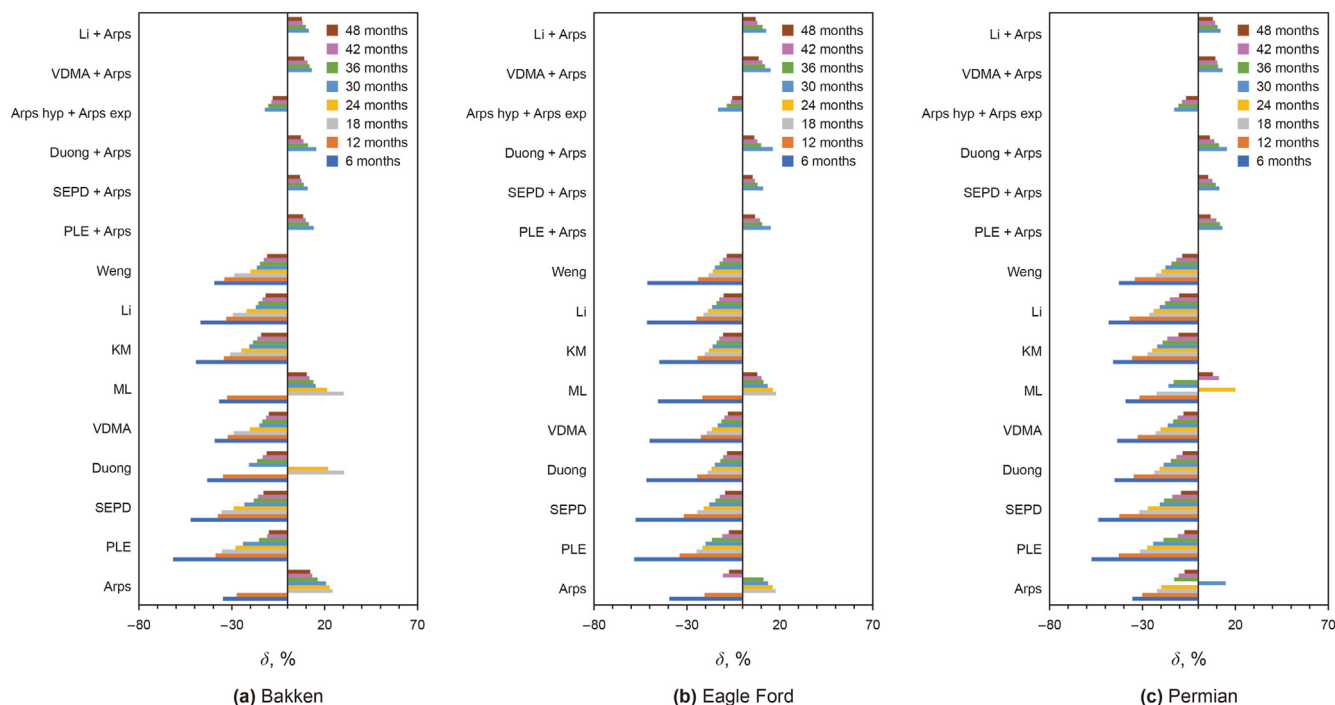


Fig. 12. The average value of relative error with different production history in different shale plays.

4. Rapid estimation of EUR at early stage

The application of decline model in Section 3.2 shows that the shorter the production history, the lower the accuracy of the EUR prediction. To achieve rapid and accurate EUR estimation with limited production data, we try to build some simple relationship between EUR and other parameters for each basin. The 'rapid' here has two meanings. First, it means that we can directly estimate EUR without fitting which is quite convenient. Second, the rapid estimation model can estimate EUR before the production of wells, which is quick that we do not need to wait the production history.

Based on the correlation analysis in Section 2.2, since some parameters are strongly mutually dependent, only the perforated length and proppant used per perforated foot are considered in the formulation. The parameters that show no obvious relationship with peak oil, such as flow back ratio and first month GOR, are also not included. Two situations are considered here, before and after oil peak has appeared. If the well hasn't begun to produce oil or the production is too short to arrive the oil peak, the perforation length and proppant injected per perforated foot are used to predict the EUR of shale oil wells. Otherwise, the peak oil is also used in the formulation. According to the linear relationship between the EUR and the peak oil rate in the double logarithmic coordinate (Fig. 14 (b)), the logarithmic transformation is applied to each parameter. The regression models between the input parameters and EUR are then established as presented in Eqs. (4) and (5),

$$\lg(\text{EUR}') = b_1 \lg(L'_f) + b_2 \lg(P') \quad (4)$$

$$\lg(\text{EUR}') = a_1 \lg(q'_{\max}) + a_2 \lg(L'_f) + a_3 \lg(P') \quad (5)$$

where EUR' is the normalized estimated ultimate reserve of shale oil well; $q'_{\max}$ is the normalized peak oil rate; L'_f is the normalized perforation length; P' is the normalized proppant per perforated foot. The values are normalized with the ranges shown in Table 4.

To verify the reliability of the model, the wells in Section 3.2 are randomly divided into train set and validation set at a ratio of 9:1, and an additional 300 wells are selected from three shale plays as test set. To note that, the additional 300 wells are not included in the wells studied in Section 3.2. The fitted parameters in Eqs. (4) and (5) are listed in Table 5. As show in Figs. 17 and 18, the average relative error for both models is less than 20%, indicating that the models could be used for EUR prediction at early production stage. If the peak oil has been arrived, the accuracy of EUR prediction can be further increased by using Eq. (5). From Tables 1 and 4, we can see that the Eagle Ford and Permian shale have similar reservoir parameters such as thickness, TOC and permeability, leading to the similar model parameters (Table 5). The models proposed here can also be used in other shale oil wells when their geological parameters are similar to the shale plays in this study to realize rapid EUR prediction.

The large amount of data in this paper can help provide better understanding of the shale oil well production. A complete comparison of fifteen decline curve models across multiple shale plays and with a large number of wells could help guide the choice of DCA models in other shale oil basins. The new fitting method proposed in our previous shale gas study is used in this study which performs better in EUR prediction over traditional fitting strategies. Thus, the rapid EUR estimation model built based on more accurate EUR prediction and more practical data has the potential to perform better over the previous methods.

5. Conclusions

In this paper, the engineering and production data of more than 30,000 shale oil wells are used, including 10,842 wells in the Bakken shale play, 12,328 wells in the Eagle Ford shale play and 13,364 wells in the Permian Basin. A systematical and detailed analysis of the production data is conducted. The impact of some key fracturing parameters on production is also investigated. Furthermore, 15 decline curve analysis models are used and

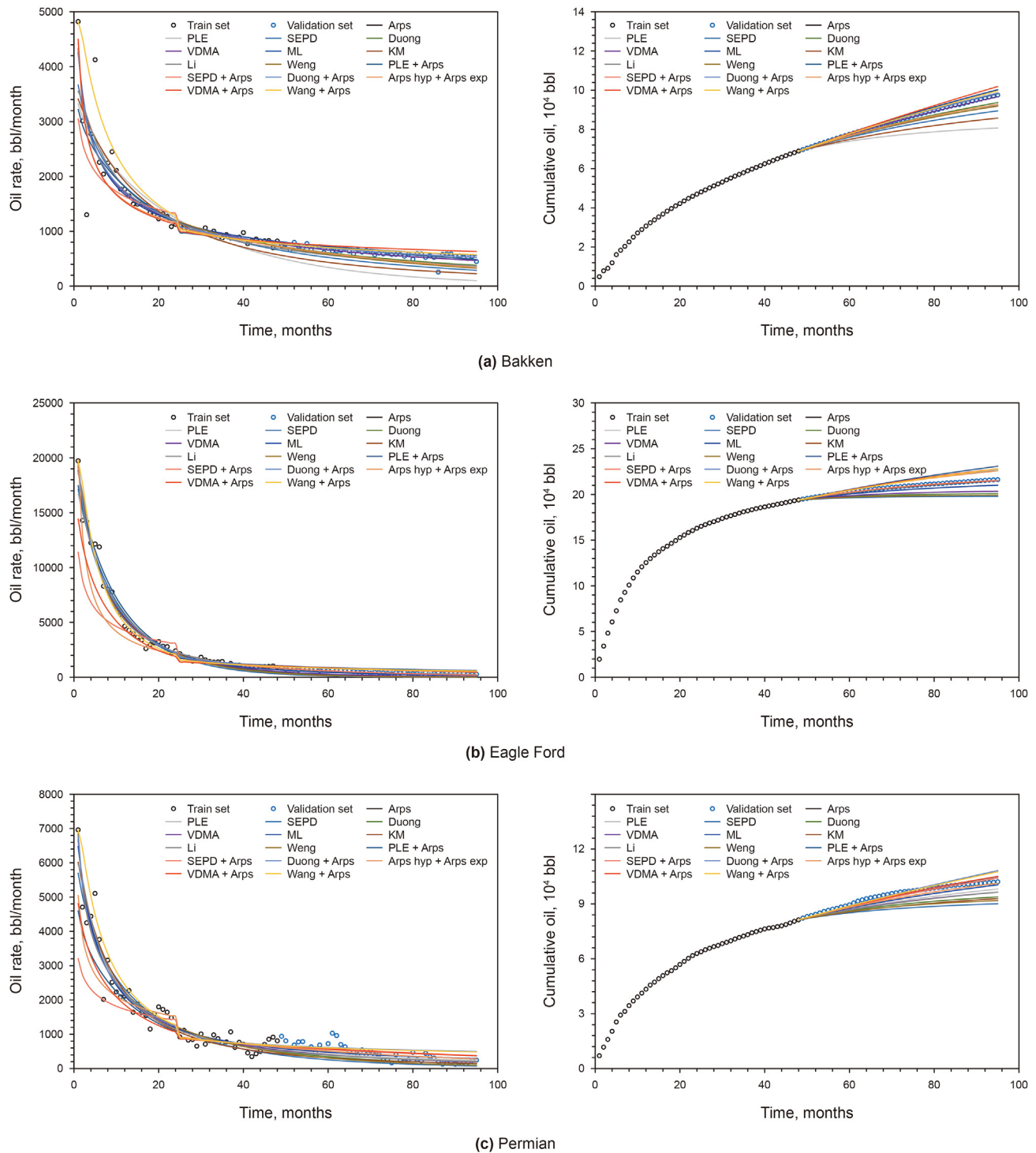


Fig. 13. The fitting results of the 15 decline models.

compared for EUR prediction based on the new fitting method. With the calculated EURs, two simple and rapid EUR prediction models are established. According to the above results and analysis, the following conclusions can be obtained.

- (1) The production data of shale oil wells from 2004 to 2021 in Bakken, Eagle Ford and Permian shale is systematically presented and analyzed in this work. The oil production in Bakken is highest and the Permian wells produce the most

water. The gas production in Eagle Ford is significantly higher than other shale plays. The oil production is featured with a rapid decline in the early stage and a long low-rate tail.

- (2) The fracturing parameters, such as perforated length, volume of proppant and fluid injection, are all positively correlated with the peak oil. The gas oil ratio and fluid flow back show slight negative impact on the oil production. According to the parameter distribution, long perforated length could be the key reason that the Bakken shale has good oil production.

Table 3
The optimal model for different production history.

Shale play	Optimal model	Production history, months							
		6	12	18	24	30	36	42	48
Bakken shale play	The optimal single model	Arps	Arps	Arps	Weng	ML	VDMA	PLE	PLE
	The optimal combined model	—	—	—	—	SEPD + Arps	SEPD + Arps	SEPD + Arps	SEPD + Arps
	The optimal model	Arps	Arps	Arps	Weng	SEPD + Arps	SEPD + Arps	SEPD + Arps	SEPD + Arps
Eagle Ford shale play	The optimal single model	Arps	Arps	Arps	Weng	VDMA	ML	VDMA	ML
	The optimal combined model	—	—	—	—	SEPD + Arps	SEPD + Arps	SEPD + Arps	SEPD + Arps
	The optimal model	Arps	Arps	Arps	Weng	SEPD + Arps	SEPD + Arps	SEPD + Arps	SEPD + Arps
Eagle Ford shale play	The optimal single model	Arps	Arps	Arps	Weng	Arps	Arps	Arps	Arps
	The optimal combined model	—	—	—	—	SEPD + Arps	SEPD + Arps	SEPD + Arps	SEPD + Arps
	The optimal model	Arps	Arps	Arps	Weng	SEPD + Arps	SEPD + Arps	SEPD + Arps	SEPD + Arps

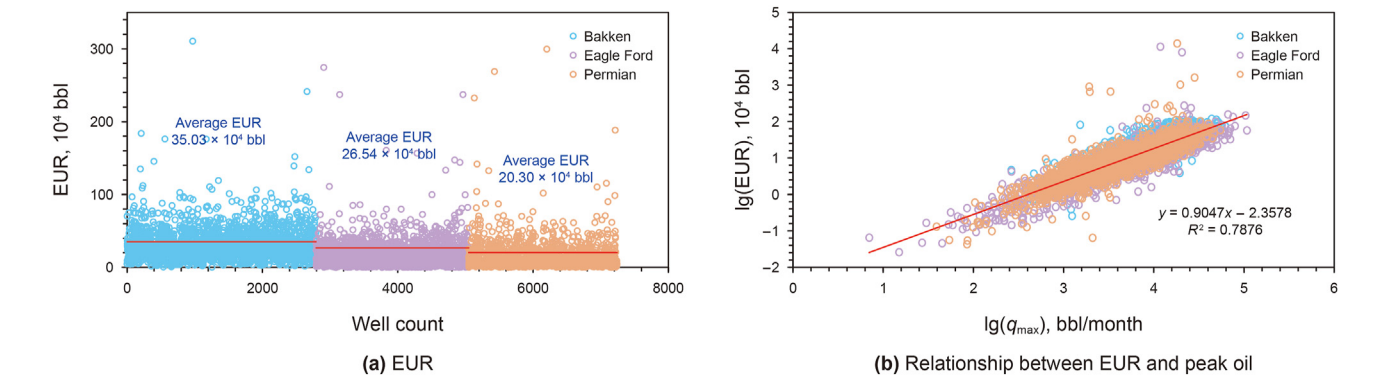


Fig. 14. EUR of shale oil wells in different basins and the relationship between EUR and peak oil.

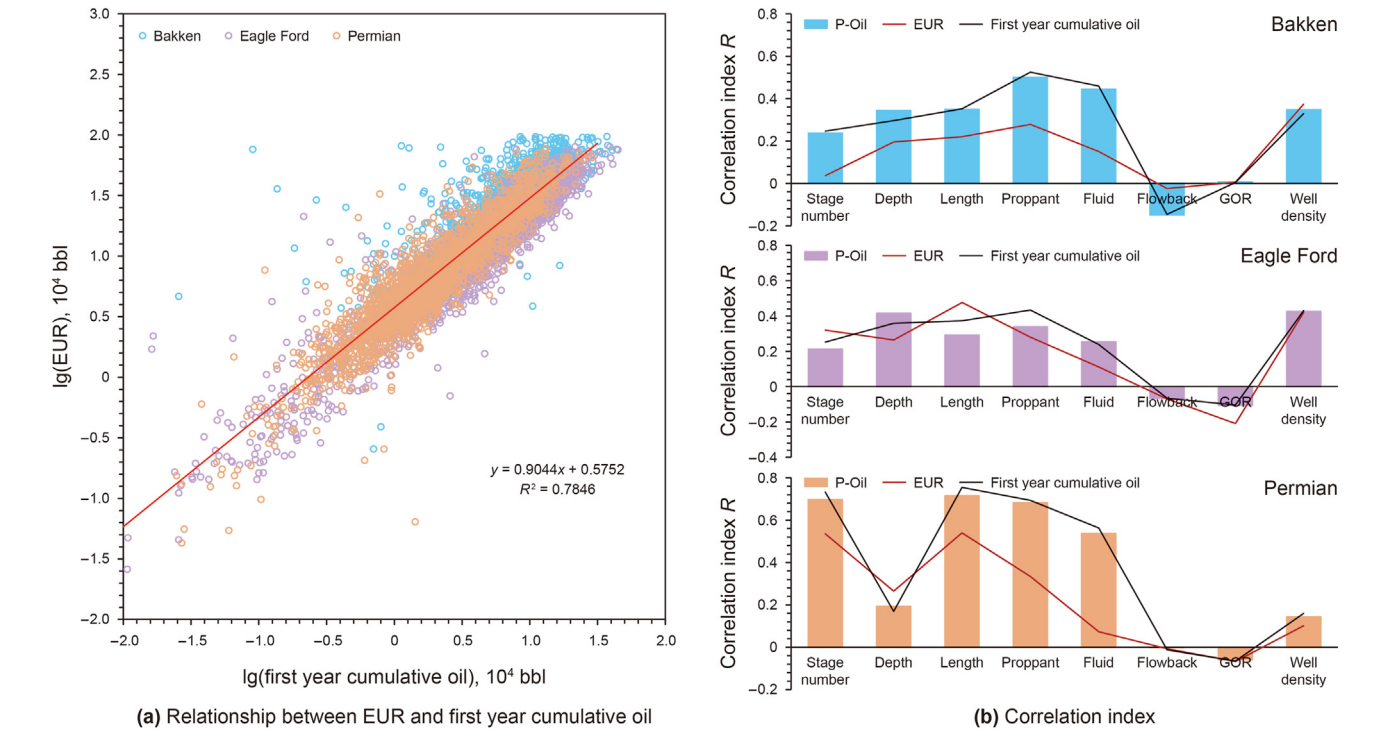


Fig. 15. The relationship between EUR and first year cumulative oil and the correlation index *R* between peak oil, EUR, first year cumulative oil production and other parameters in Bakken, Eagle Ford and Permian.

(3) Over ten decline curve models are tested on the shale oil data. The results show that the EUR predicted by the single PLE, SEPD, VDMA, KM, Li and Weng models could be smaller than the practical values, while the combined models (except Arps hyp + Arps exp) tend to overestimate the EUR. If the well has been produced more than 2 years, the SEPD + Aprs

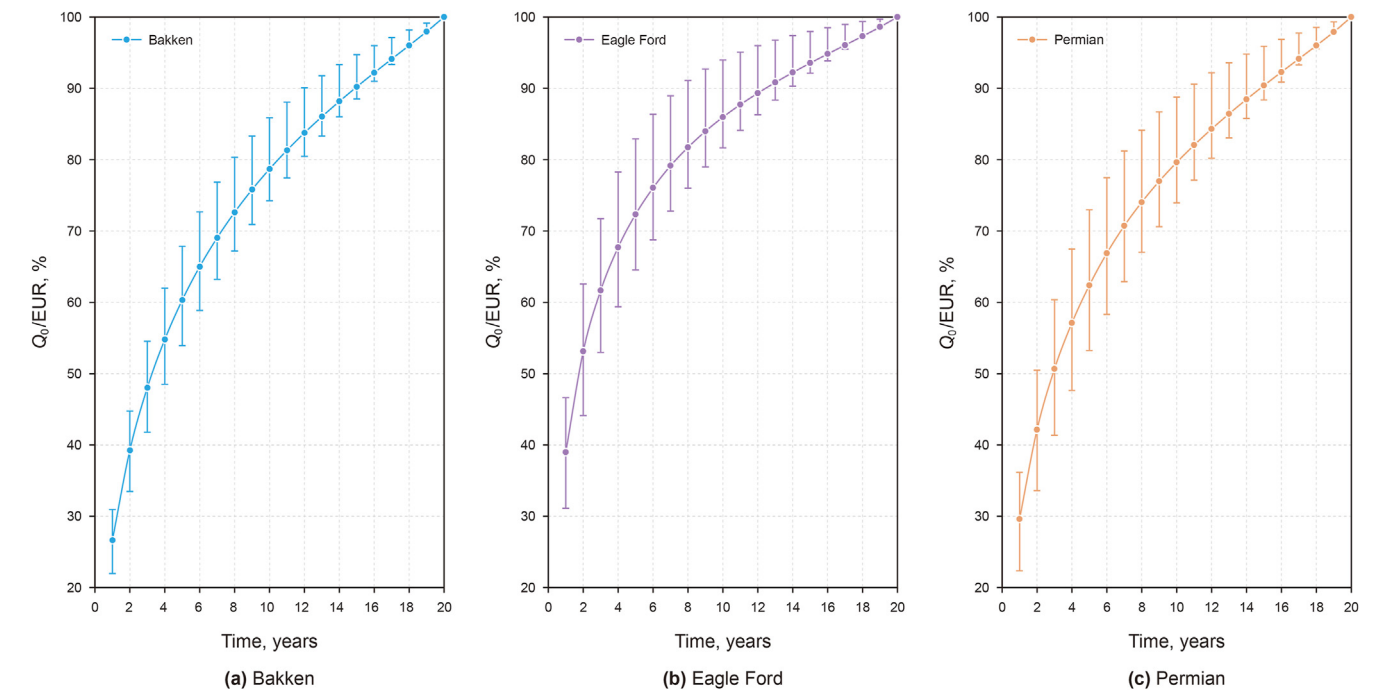


Fig. 16. The error bars of the ratio between cumulative oil production and EUR.

Table 4
The maximum and minimum values used for normalization in different shale plays.

Parameter	Bakken		Eagle Ford		Permian		Unit
	Maximum	Minimum	Maximum	Minimum	Maximum	Minimum	
lg(EUR)	1.9863	−0.5922	4.0566	−1.0318	4.1429	−1.1935	10 ⁴ bbl
lg(<i>q</i> _{max})	4.7905	2.8388	5.0320	2.0294	4.7371	1.9085	bbl/month
lg(<i>L</i> _{<i>t</i>})	4.1682	2.7612	4.0170	2.7177	4.1486	2.6990	ft
lg(<i>P</i>)	3.7327	0.8856	3.6785	1.3647	3.7049	0.9577	lbs/ft

Table 5
The parameters of EUR prediction model.

Shale play	<i>a</i> ₁	<i>a</i> ₂	<i>a</i> ₃	<i>b</i> ₁	<i>b</i> ₂
Bakken	0.6218	0.3402	0.1871	0.5935	0.5231
Eagle Ford	0.5162	0.0377	0.1366	0.3094	0.2851
Permian	0.4731	0.0383	0.1342	0.3840	0.3000

model is suggested, otherwise the Arps model could be a good choice.

- (4) Using the predicted EUR by decline curve models, we further construct a simple correlation between EUR and peak oil, perforation length, and proppant per perforated foot. The accuracy of the model could be over 85%. If the peak oil has

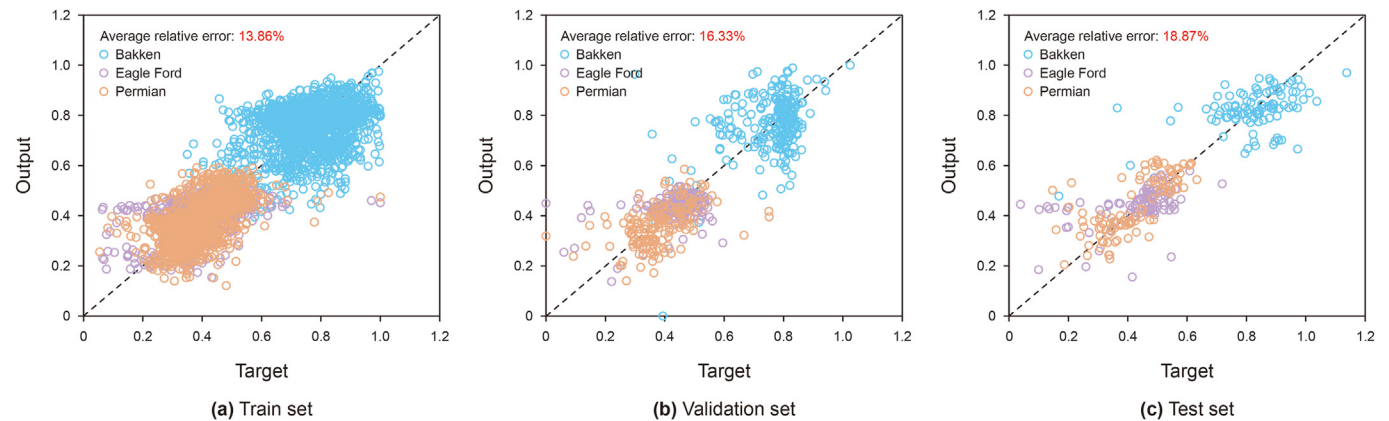


Fig. 17. Comparison between predicted and practical EURs with Eq. (4).

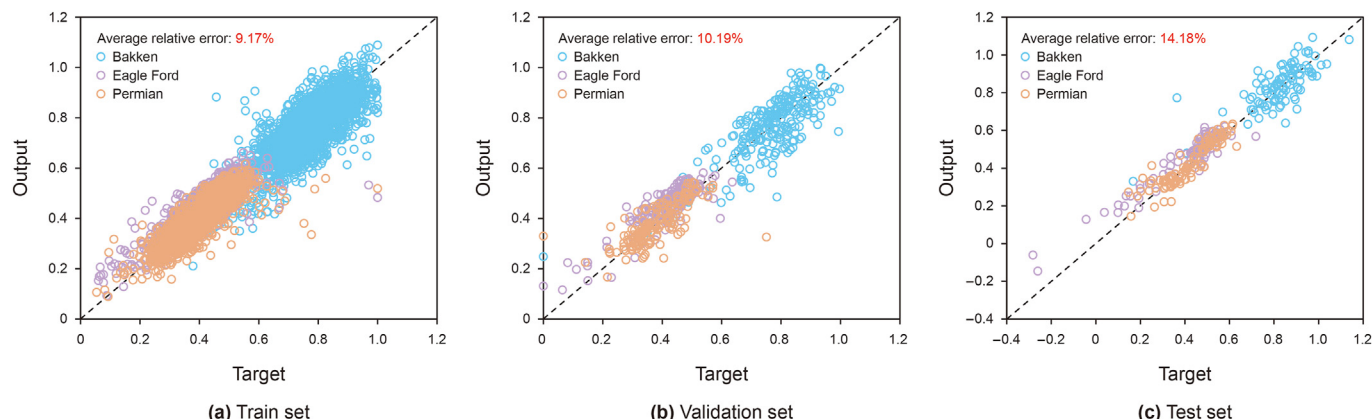


Fig. 18. Comparison between predicted and practical EURs with Eq. (5).

not been appeared, the EUR can also be estimated with the perforation length and proppant per perforated foot, which also provides EUR with 80% reliability.

CRediT authorship contribution statement

Hui-Ying Tang: Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Ge He:** Writing – original draft, Methodology, Investigation. **Ying-Ying Ni:** Writing – review & editing, Methodology, Investigation. **Da Huo:** Writing – review & editing, Investigation, Data curation. **Yu-Long Zhao:** Writing – review & editing, Investigation, Conceptualization. **Liang Xue:** Writing – review & editing, Methodology, Investigation. **Lie-Hui Zhang:** Writing – review & editing, Project administration, Funding acquisition.

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.

Acknowledgement

This work was funded by the National Natural Science Foundation of China (Grant No. 52374043) and the Key Program of National Natural Science Foundation of China (Grant No. 52234003).

References

- Alotaibi, B., Schechter, D., Wattenbarger, R.A., 2015a. Production forecast, analysis and simulation of Eagle Ford shale oil wells. In: SPE Middle East Unconventional Resources Conference and Exhibition. <https://doi.org/10.2118/SPE-172929-MS>.
- Alotaibi, B., Schechter, D., Wattenbarger, R.A., 2015b. Production forecast, analysis and simulation of Eagle Ford shale oil wells. In: SPE Middle East Unconventional Resources Conference and Exhibition. <https://doi.org/10.2118/SPE-172929-MS>.
- Altawati, F., Emadi, H., Khalil, R., Heinze, L., Menouar, H., 2022. An experimental investigation of improving Wolfcamp shale-oil recovery using liquid-N₂-assisted N₂ and/or CO₂ huff-n-puff injection technique. *Fuel* 324 (Part A), 124450. <https://doi.org/10.1016/j.fuel.2022.124450>.
- Arps, J.J., 1945. Analysis of decline curves. *Transaction of the AIME* 160 (1), 228–247. <https://doi.org/10.2118/945228-G>.
- Baskoro, A.S., Baur, F., Yu, A.Z., Grossman, E.L., 2023. Restoring source rock initial quality and quantity with kinetic-based inversion - applied to the Wolfcamp play in the Permian Delaware Basin. *Mar. Petrol. Geol.* 150 (Suppl. C), 106143. <https://doi.org/10.1016/j.marpetgeo.2023.106143>.
- Bhandari, A.R., Flemings, P.B., Ramirez, S.R., Hofmann, R., Polito, P.J., 2019. Gas and liquid permeability measurements in Wolfcamp samples. *Fuel* 236, 1026–1036. <https://doi.org/10.1016/j.fuel.2018.09.038>.
- Duong, A.N., 2010. An unconventional rate decline approach for tight and fracture-dominated gas wells. In: Canadian Unconventional Resources and International Petroleum Conference. <https://doi.org/10.2118/137748-MS>.
- Ghanbari, E., Abbasi, M.A., Dehghanpour, H., Bearinger, D., 2013. Flowback volumetric and chemical analysis for evaluating load recovery and its impact on early-time production. In: SPE Unconventional Resources Conference. <https://doi.org/10.2118/167165-MS>.
- Guo, X.S., Li, M.W., Zhao, M.Y., 2023. Shale oil development and utilization and its role in energy industry. *Bull. Chin. Acad. Sci.* 38 (1), 38–47. <https://doi.org/10.16418/j.issn.1000-3045.20221027001> (in Chinese).
- Gupta, I., Rai, C., Sondergeld, C., Devegowda, D., 2018. Variable exponential decline: modified Arps to characterize unconventional-shale production performance. *SPE Reservoir Eval. Eng.* 21 (4), 1045–1057. <https://doi.org/10.2118/194005-PA>.
- Harris, S., Lee, W.J., 2014. A study of decline curve analysis in the Elm Coulee field. In: SPE Unconventional Resources Conference. <https://doi.org/10.2118/169018-MS>.
- Ilk, D., Rushing, J.A., Perego, A.D., Blasingame, T.A., 2008. Exponential vs hyperbolic decline in tight gas sands: understanding the origin and implications for reserve estimates using Arps' decline curves. In: SPE Annual Technical Conference and Exhibition. <https://doi.org/10.2118/116731-MS>.
- Jiang, P., Wang, W.X., Li, J., Wang, F., Zhou, Y.Q., Liu, Y.L., 2021. Pressure control flowback technology for shallow shale gas wells: Taking Zhaotong National Shale Gas Demonstration Area as an example. *Nat. Gas. Ind.* 41 (A1), 186–191. <https://doi.org/10.3787/j.issn.1000-0976.2021.S1.028> (in Chinese).
- Joshi, K., Lee, J., 2013. Comparison of various deterministic forecasting techniques in shale gas reservoirs. In: SPE Hydraulic Fracturing Technology Conference. <https://doi.org/10.2118/163870-MS>.
- Lei, Q., Weng, D.W., Guan, B.S., Shi, J.F., Cai, B., He, C.M., Sun, Q., Huang, R., 2023. Shale oil and gas exploitation in China: technical comparison with US and development suggestions. *Petrol. Explor. Dev.* 50 (4), 824–831. <https://doi.org/10.11698/PED.20220714>.
- Li, H.N., Bruce, H., Matthew, D., Eric, R., 2015. Characterizing the Middle Bakken: laboratory measurement and rock typing of the Middle Bakken Formation. In: SPE/AAPG/SEG Unconventional Resources Technology Conference. <https://doi.org/10.15530/URTEC-2015-2172485>.
- Liang, H.B., Zhang, L.H., Zhao, Y.L., Zhang, B.N., Chang, C., Chen, M., Bai, M.X., 2020. Empirical methods of decline-curve analysis for shale gas reservoirs: review, evaluation, and application. *J. Nat. Gas Sci. Eng.* 83, 103531. <https://doi.org/10.1016/j.jngse.2020.103531>.
- Liang, Y., Liao, L.L., Guo, Y., 2019. A big data study: correlations between EUR and petrophysics/engineering/production parameters in shale formations by data regression and interpolation analysis. In: SPE Hydraulic Fracturing Technology Conference and Exhibition. <https://doi.org/10.2118/194381-MS>.
- Luo, S.G., Ding, C., Cheng, H.F., Zhang, B.N., 2022. Estimated ultimate recovery prediction of fractured horizontal wells in tight oil reservoirs based on deep neural networks. *Advances in Geo-Energy Research* 6 (2), 111–122. <https://doi.org/10.46690/ager.2022.02.04>.
- Martin, R., Bailly, J., Malpani, R., Lindsay, G., Atwood, W.K., 2011. Understanding production from Eagle ford–Austin chalk system. In: SPE Annual Technical Conference and Exhibition. <https://doi.org/10.2118/145117-MS>.
- Matthews, C.S., Lefkowitz, H.C., 1956. Gravity drainage performance of depletion-type reservoirs in the stripper stage. *Transactions of the AIME* 207 (12), 265–274. <https://doi.org/10.2118/665-G-P>.
- Nicot, J.P., Darvari, R., Eichhubl, P., Scanlon, B.R., Elliott, B.A., Bryndzia, L.T., Gale, J.F.W., Fall, A., 2020. Origin of low salinity, high volume produced waters in the Wolfcamp Shale (Permian), Delaware Basin, USA. *Appl. Geochem.* 122, 104771. <https://doi.org/10.1016/j.apgeochem.2020.104771>.
- Niu, W.T., Lu, J.L., Sun, Y.P., Guo, W., Liu, Y.Y., Mu, Y., 2022. Development of visual prediction model for shale gas wells production based on screening main

- controlling factors. *Energy* 250, 123812. <https://doi.org/10.1016/j.energy.2022.123812>.
- Park, J., Gupta, A.D., Singh, A., Sankaran, S., 2021. Hybrid physics and data-driven modeling for unconventional field development and its application to US onshore basin. *J. Petrol. Sci. Eng.* 206, 109008. <https://doi.org/10.1016/j.petrol.2021.109008>.
- Patzek, T., Male, F., Marder, M., 2014. A simple model of gas production from hydrofractured horizontal wells in shales. *AAPG (Am. Assoc. Pet. Geol.) Bull.* 98 (12), 2507–2529. <https://doi.org/10.1306/03241412125>.
- Pranesh, V., 2018. Subsurface CO₂ storage estimation in Bakken tight oil and Eagle Ford shale gas condensate reservoirs by retention mechanism. *Fuel* 215 (3), 580–591. <https://doi.org/10.1016/j.fuel.2017.11.049>.
- Sonnenberg, S.A., Pramudito, A., 2009. Petroleum geology of the giant Elm coulee field, Williston Basin. *AAPG (Am. Assoc. Pet. Geol.) Bull.* 93 (9), 1127–1153. <https://doi.org/10.1306/052809090006>.
- Tang, H.Y., Zhang, B.N., Liu, S., Li, H.Y., Huo, D., Wu, Y.S., 2021. A novel decline curve regression procedure for analyzing shale gas production. *J. Nat. Gas Sci. Eng.* 88, 103818. <https://doi.org/10.1016/j.jngse.2021.103818>.
- Tsvetkova, A., Partridge, M., 2017. The shale revolution and entrepreneurship: an assessment of the relationship between energy sector expansion and small business entrepreneurship in US counties. *Energy* 141, 423–434. <https://doi.org/10.1016/j.energy.2017.09.101>.
- Tunstall, T., 2015. Iterative bass model forecasts for unconventional oil production in the Eagle Ford shale. *Energy* 93 (Part 1), 580–588. <https://doi.org/10.1016/j.energy.2015.09.072>.
- U. S. Energy Information Administration, 2020. *Drilling Productivity Report: for Key Tight Oil and Shale Gas Regions*. EIA Independent Statistics & Analysis, Washington, DC.
- Valko, P.P., 2009. Assigning value to stimulation in the Barnett Shale: a simultaneous analysis of 7000 plus production histories and well completion records. In: *SPE Hydraulic Fracturing Technology Conference*. <https://doi.org/10.2118/119369-MS>.
- Walters, C.C., Gong, C.R., Sun, X., Zhang, T.W., 2023. Geochemistry of oils and condensates from the lower Eagle Ford formation, South Texas. Part 3: basin modeling. *Mar. Petrol. Geol.* 150, 106117. <https://doi.org/10.1016/j.marpetgeo.2023.106117>.
- Wang, K., Li, H.T., Wang, J.C., Jiang, B.B., Bu, C.Z., Zhang, Q., Luo, W., 2017. Predicting production and estimated ultimate recoveries for shale gas wells: a new methodology approach. *Appl. Energy* 206, 1416–1431. <https://doi.org/10.1016/j.apenergy.2017.09.119>.
- Wang, Q., Song, X., Li, R., 2018. A novel hybridization of nonlinear grey model and linear ARIMA residual correction for forecasting U.S. shale oil production. *Energy* 165 (Part B), 1320–1331. <https://doi.org/10.1016/j.energy.2018.10.032>.
- Wei, J.X., Zhang, Y., Shang, J.H., Lv, N., Liu, W.C., Wang, H.K., Ma, F.J., Zhang, Q.T., 2021. Principal factor analysis on initial productivity in shale oil development: a case study of Block Li-151 in Changqing oilfield. *Petroleum Reservoir Evaluation and Development* 11 (4), 550–558. <https://doi.org/10.13809/j.cnki.cn32-1825/te.2021.04.011> (in Chinese).
- Xue, L., Wang, J.B., Han, J.X., Yang, M.J., Mwasmwasa, M.S., Nanguka, F., 2023. Gas well performance prediction using deep learning jointly driven by decline curve analysis model and production data. *Advances in Geo-Energy Research* 8 (3), 159–169. <https://doi.org/10.46690/ager.2023.06.03>.
- Zhao, Q., Wang, H.Y., Sun, Q.P., Jiang, X.C., Yu, R.Z., Kang, L.X., Wang, X.F., 2020. A logical growth model considering the influence of shale gas reservoirs and development characteristics. *Nat. Gas. Ind. B* 7 (6), 656–663. <https://doi.org/10.1016/j.ngib.2020.05.005>.
- Zhao, Y.L., He, G., Liu, X.Y., Zhang, L.H., Wu, J.F., Chang, C., 2022. A new method for fitting empirical production decline model based on data weighting: a case study on Changning Block of the Sichuan Basin. *Nat. Gas. Ind.* 42 (11), 66–76. <https://doi.org/10.3787/j.issn.1000-0976.2022.11.007> (in Chinese).
- Zhou, Q., Lei, Z.D., Chen, Z.W., Wang, Y.H., Liu, Y.S., Xu, Z.H., Liu, Y.Q., 2023. Shale oil production predication based on an empirical model constrained CNN-LSTM. *Energy Geoscience* 5 (2), 100252. <https://doi.org/10.1016/j.engeos.2023.100252>.