Energy Engineering

Tech Science Press

https://doi.org/10.32604/ee.2026.079317

ARTICLE

Long-Term Production Prediction Method for Shale Oil Based on Bayesian Physical-Information Neural Networks

Longqiao Hu[1], Yunjin Wang[1,*], Jia Liu[1], Jiacheng Yin[1], Siyu Zhang[2], Mengyu Li[1], Qi Wu[1], Jiawei Li[1] and Fujian Zhou[3]

[1]College of Petroleum Engineering, China University of Petroleum-Beijing, Beijing, China
[2]Oil & Gas Technology Research Institute, Petrochina Changqing Oilfield Company, Xi'an, China
[3]National Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum-Beijing, Beijing, China

*Corresponding Author: Yunjin Wang. Email: wangyj@cup.edu.cn

Received: 19 January 2026; Accepted: 24 February 2026; Published: 30 August 2026

ABSTRACT: The flow mechanisms of shale oil are inherently complex, characterized by diverse occurrence states. Establishing a robust production forecasting model is essential for thoroughly evaluating well deliverability and the efficacy of reservoir stimulation. While conventional analytical techniques, numerical reservoir simulation, and production decline analysis constitute standard practice, they often struggle to balance computational efficiency with predictive fidelity. Intelligent algorithms offer superior precision in short-term forecasting following extensive data training; however, they frequently exhibit poor generalization and underfitting during long-term performance projections. Consequently, integrating domain-specific production decline theory as physical constraints represents a strategic pathway to enhance the robustness of data-driven frameworks. In this study, the ordinary differential equation (ODE) governing the Duong decline model was derived and embedded into a deep neural network (DNN) architecture. Concurrently, Bayesian optimization was employed to fine-tune both the ODE parameters and the DNN hyperparameters, culminating in the development of a Bayesian-optimized physics-informed neural network (PINN). To validate the model, production data from shale oil wells in the Jimsar Sag were analyzed to evaluate predictive performance. The results demonstrate that Bayesian optimization reduces errors by directly using cumulative oil production as the fitting target for decline model parameter optimization, while also accurately capturing data characteristics. Compared with the traditional decline analysis method, the prediction ability of the Duong decline model optimized by Bayesian is improved by about 68.5% on average. Regarding network architecture and hyperparameter tuning, Bayesian optimization leverages surrogate models and acquisition functions to systematically identify optimal configurations. While its impact on long short-term memory (LSTM) networks proved marginal, it substantially enhanced DNN and PINN performance, yielding an average improvement of roughly 53.1% for the PINN. Under physical constraints, the Bayesian-PINN outperformed standalone DNN and harmonic decline models in future long-term forecasting by approximately 61.7% and 21.4%, respectively. Notably, the predictive superiority of the PINN becomes increasingly pronounced as the forecast horizon extends. This integration of production decline laws with deep learning feature extraction provides a robust technical foundation for the digital transformation of oilfield operations.

KEYWORDS: Production forecasting; physical constraints; artificial intelligence; PINN; shale oil

1 Introduction

Shale oil constitutes a critical component of the global energy supply. In the United States, its daily output has exceeded five million barrels, representing nearly 70% of the nation's domestic crude oil production [1].

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

However, shale oil development faces challenges in resource assessment and production prediction [2,3]. Accurate production prediction is important for optimizing extraction strategies, improving resource utilization efficiency, and reducing development risks [4].

Experts and scholars have conducted extensive analyses and curve fitting based on oil and gas well production history. Arnold and Anderson [5] first conceptualized production decline and summarized the calculation methods for decline curve analysis (DCA) by observing well production characteristics. Building on this, Arps [6] synthesized prior research to propose the Arps decline model. Capable of describing various decline patterns, the Arps model remains one of the primary forecasting tools in major oilfields due to its broad applicability and computational simplicity. Denney [7] further investigated the adaptive range of the Arps model and observed poor fitting performance in low-porosity, low-permeability reservoirs; he further noted that the decline exponent, n, serves as a critical reference parameter, with fitting quality often degrading when $n < 1$. Furthermore, Rushing et al. [8] utilized field production curves to highlight the limitations of the Arps hyperbolic decline model, specifically its low accuracy in fitting transient linear flow. This limitation arises because the Arps model primarily describes production characteristics after the flow reaches a quasi-steady state. To address this, Fetkovich et al. [9] integrated analytical solutions for unsteady-state flow with the Arps model, proposing the Fetkovich type-curve analysis method. To overcome the poor performance of exponential decline models in unconventional reservoirs, Ilk et al. [10] introduced the power-law exponential (PLE) model, which is better suited for characterizing the early transient flow stage. With the widespread adoption of hydraulic fracturing, Duong [11] analyzed dynamic production data from Barnett shale wells and postulated that fractured shale reservoirs exhibit long-term fracture-dominated linear flow. Consequently, the Duong decline method was derived, positing that the contribution of matrix permeability to productivity is significantly lower than that of the fracture network. Subsequently, numerous scholars have developed a series of DCA models by integrating the Duong model with boundary-dominated flow, unsteady linear flow, and production data fitting techniques [12–14]. While these models utilize mathematical functions to fit historical data for future extrapolation, they lack robustness against complex operational conditions, such as sand plugging and inter-well interference.

Post-fracturing production data represents a typical time series; thus, production forecasting can be formulated as a time-series forecasting problem. Methodologies are generally categorized into traditional statistical approaches and deep learning algorithms [15–17]. Among traditional methods, the autoregressive model, moving average model, and autoregressive integrated moving average (ARIMA) have been successively applied to fit well production [18]. Gupta et al. [19] noted that for shale gas wells, the prediction accuracy of ARIMA is comparable to that of the Duong model, while Olominu and Sulaimon [20] suggested that ARIMA slightly outperforms Arps in medium- to long-term forecasting. Although traditional methods are mature and computationally efficient, they struggle with non-stationary data arising from complex field operations [21]. With the advancement of machine learning, recurrent neural networks (RNN), which incorporate internal memory mechanisms and feature extraction capabilities, have been deployed to capture temporal evolution in production data [22]. Studies have demonstrated RNN's ability to capture data fluctuations caused by sand plugging and well interference, achieving favorable prediction results. To address RNN's limitations in capturing long-term dependencies and the vanishing gradient problem, long short-term memory (LSTM) networks introduced forgetting, input, and output gates to process long sequences efficiently [23]. The gated recurrent unit (GRU) further optimized this architecture by using only two gates (reset and update), resulting in lower computational costs and faster training speeds [24]. Alimohammadi et al. [25] compared algorithms including LSTM, GRU, and Bi-directional RNN on shale gas data, identifying GRU as having the highest relative accuracy. Furthermore, optimization algorithms such as particle swarm optimization and Bayesian optimization have been employed to tune network structures and hyperparameters, significantly enhancing model performance [26,27]. However, while these data-driven methods excel in short-term forecasting, their performance deteriorates over long horizons due to error accumulation across multiple time steps. To mitigate this, strategies such as iterative prediction, transfer learning, physical constraints, and multi-input architectures have been applied to improve the multi-step performance of LSTM and GRU [28–30]. In summary, while intelligent time-series forecasting can capture interference from complex working conditions, it often yields suboptimal results for long-term shale oil prediction. Integrating domain knowledge as a physical constraint offers a promising pathway to enhance long-term accuracy, yet effective algorithms for fusing these physical constraints remain scarce.

In this work, we derive the ordinary differential equation (ODE) of the Duong decline model and integrate it into a deep neural network (DNN) as a physical constraint, establishing a physics-informed neural network (PINN). Bayesian optimization is employed to simultaneously optimize the parameters of the Duong ODE and the DNN architecture, resulting in a novel Bayesian-PINN dynamic production forecasting model. Finally, the predictive capability of the proposed model is evaluated using production data from shale oil wells in the Jimsar Sag.

2 Bayesian-PINN Method

The DCA employs mathematical functions to fit observed well production data and forecasts future performance by extrapolating the fitted trends. Within this framework, Duong [11] proposed a specialized model based on the observation that shale wells in the United States exhibit long-term linear flow dominated by fractures. This methodology posits that the well remains in a linear flow regime throughout its production life, characterized by a linear relationship between cumulative production and time on a double logarithmic plot. However, actual shale oil production is governed by a complex interplay of geological attributes, such as permeability and saturation, in addition to engineering factors including hydraulic fracturing techniques. Consequently, even the Duong model fails to yield precise production forecasts for these complex unconventional reservoirs.

2.1 ODE of Duong Decline Model

The artificial fracture network formed in the shale oil reservoir after large-scale fracturing has become the main flow channel for shale oil production. Constructing physical constraints based on experience accumulation can improve machine learning to identify the sequence change relationship of production. However, in the long-term production prediction of actual wells, the flow differential equations are difficult to effectively embedded into the loss function because parameters such as pressure and saturation are usually unknown. The Duong decline model demonstrates good applicability for predicting production from fractured wells in unconventional oil and gas reservoirs, which can be used to effectively constrain the accuracy of the long-term production prediction.

Based on the production characteristics of long-term linear flow in shale reservoir, Duong proposed a model for predicting production, and named it Duong decline model [11].

$$q(t) = q_i t^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}, \tag{1}$$

$$Q = \frac{q_i}{a} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}, \tag{2}$$

where $q(t)$ is the production of oil and gas wells at different times, ton/d; t is time, d; q_i is the initial production, ton/d; Q is the cumulative production at time t, ton; a and m are coefficients.

During oil and gas production, the decline rate in the production decline stage is defined as follows:

$$D(t) = -\frac{1}{q_t}\frac{dq_t}{dt}, \tag{3}$$

where $D(t)$ is the decline rate, d^{-1}; q_t is the production in a certain period of time in the decline stage, ton/d.

Based on the definition of decline rate, Eq. (1) is substituted into Eq. (3) to obtain:

$$D(t) = -\frac{1}{q_i t^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}}\frac{d}{dt}\left(q_i t^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}\right). \tag{4}$$

Assume that there are intermediate variables u and v.

$$\begin{cases} u(t) = t^{-m} \\ v(t) = e^{\frac{a}{1-m}\left(t^{1-m}-1\right)} \end{cases}. \tag{5}$$

Then its derivative is:

$$\begin{cases} u'(t) = -mt^{-m-1} \\ v'(t) = at^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)} \end{cases}. \tag{6}$$

Substituting Eq. (5) into Eq. (4) gives:

$$\begin{aligned} D(t) &= -\frac{1}{t^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}}\frac{d}{dT}\left(u(t)\,v(t)\right) \\ &= -\frac{1}{t^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}}\left(u(t)\,v'(t) + u'(t)\,v(t)\right). \end{aligned} \tag{7}$$

Derived by substituting Eqs. (5) and (6) into Eq. (7):

$$\begin{aligned} D(t) &= -\frac{1}{t^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}}\left(t^{-m} at^{-m} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)} + -mt^{-m-1} e^{\frac{a}{1-m}\left(t^{1-m}-1\right)}\right) \\ &= -\frac{1}{t^{-m}}\left(t^{-m} at^{-m} + -mt^{-m-1}\right) = mt^{-1} - at^{-m}. \end{aligned} \tag{8}$$

Eq. (3) can be modified to:

$$\frac{dq(t)}{dt} + D(t)\,q(t) = 0. \tag{9}$$

Substituting Eq. (8) into Eq. (9) gives:

$$\frac{dq(t)}{dt} + \left(mt^{-1} - at^{-m}\right)q(t) = 0. \tag{10}$$

This can be obtained from the relationship between daily oil production and cumulative oil production:

$$\frac{d^2Q}{dt^2} + \left(mt^{-1} - at^{-m}\right)\frac{dQ}{dt} = 0. \tag{11}$$

If the cumulative oil production is directly taken as the prediction target, its value span is large. Therefore, this work uses the maximum and minimum normalization method to normalize the cumulative oil production.

$$Q_d = \frac{Q - Q_{\min}}{Q_{\max} - Q_{\min}}, \tag{12}$$

where Q_d is the normalized cumulative oil production; $Q_{\min}$ is the minimum value in the cumulative oil production, t; $Q_{\max}$ is the maximum value in the cumulative oil production, ton.

This can be obtained by bringing the Eq. (12) into the Eq. (13):

$$\left(\frac{1}{Q_{\max} - Q_{\min}}\right)\left(\frac{d^2 Q_d}{dt^2} + \left(mt^{-1} - at^{-m}\right)\frac{dQ_d}{dt}\right) = 0. \tag{13}$$

2.2 PINN Based on ODE

The PINN method solves the problem of predicting the long-term production of shale oil by training a neural network model to minimize the loss function of the point [31]. The loss function of the PINN method consists of two components, which are the error of the ODE of the point and the error of the known production prediction.

Based on the ODE of the Duong decline model, this work uses square error to measure the physical loss.

$$L_p = \left\|\left(\frac{1}{Q_{\max} - Q_{\min}}\right)\left(\frac{d^2 \widehat{Q_d}}{dt^2} + \left(mt^{-1} - at^{-m}\right)\frac{d\widehat{Q_d}}{dt}\right)\right\|^2, \tag{14}$$

where L_p is the physical constraint loss; $d\widehat{Q_d}$ is the predicted value of normalized cumulative oil production.

For the initial production pattern of oil wells, the square error between the predicted and true values is used as the loss of the neural network.

$$L_d = \left\|Q_d - \widehat{Q_d}\right\|^2, \tag{15}$$

where L_d is the loss of the neural network.

The total loss function of PINN can be expressed as:

$$L = \eta L_d + L_p, \tag{16}$$

where η is the weighting factor of the equilibrium loss function, which serves to balance the data constraint and the physical constraint of the Duong decline model. A smaller value of η indicates a stronger physical constraint imposed by the decline model, whereas a larger value enhances the influence of the data constraint.

2.3 Bayesian Optimization Algorithm

The hyperparameters in the neural network structure and the weighting factor in the loss function have a large impact on the accuracy of the model. Poorly matched parameter settings will lead to the prediction results being seriously shifted from the target variable. On the contrary, appropriate hyperparameter settings can significantly improve the prediction accuracy of the model [32]. The adjustment of hyperparameters in neural networks has been studied by many scholars [33–35]. Compared with grid search and random search optimization, Bayesian optimization has the characteristics of considering the previous parameter information, fewer iterations, and still effective for non-convex problems. Therefore, using Bayesian optimization to optimize and solve the PINN algorithm will be conducive to constructing a model with good prediction performance.

Bayesian optimization is an algorithm based on the Bayesian process without using gradient optimization. It can optimize neural networks and other algorithms to achieve the optimal solution in the interval, and is an important technical means in automatic machine learning [36]. The most critical steps in Bayesian optimization are the establishment of the agent model and the construction of the acquisition function. The agent model is used to give the distribution of the unknown objective function based on the observed data, and the acquisition function gives the objective of the next optimization step. Fig. 1 shows the pseudo-code of the Bayesian algorithm: first, n design points sequence $X = \{x_1, x_2, \ldots, x_n\}$ are selected by initialization method, and the objective function value $y = \{y(x_1), \ldots, y(x_n)\}$. Then, construct an agent model for the response values based on the data sequence $\{X, y\}$ of the sampling points, and derive the posterior distribution model of the objective function. Build the acquisition function $L(x)$ based on the posterior distribution model in the above steps, and add the next design point by finding the most value of the acquisition function. If the iterative stopping condition is not satisfied, the agent model continues to be updated on the new {X, y} dataset until the end of the program.

```
Pseudocode of Bayesian optimization
Input: objective function f(x), acquisition function L(x), initial point X={x_1, x_2,…, x_n}
        and maximum number of iterations T
Dataset←{(x_i, f(x_i)) for i in 1 to n}
Training the surrogate model M(x) on the dataset
For t=1 to T Do:
    Select the next sampling point x_{n+1}=argmax_{x∈Ω} L(x)
    Evaluate the objective function y_{n+1}=f(x_{n+1})
    Update the data set X=X∪{x_{n+1}}, y=y∪{y_{n+1}}
    Training and updating agent model M(x)
End
Outfut: x*=argmax{y_n For all (x_n, y_n) in dataset}
Return x*, f(x*)
```

Figure 1: Bayesian algorithm pseudocode.

The process of Bayesian optimization includes the initialization method, agent model and acquisition function. The commonly used initialization method is the Latin hypercube design, which ensures that the design points are well populated in the design space through random number generation, thus effectively avoiding the Bayesian optimization process from falling into local optimal solutions. In this way, the global search capability of the optimization can be improved to ensure that the search process covers more potential solution space.

Most Bayesian optimization methods use Gaussian processes, which are nonparametric models consisting of multiple Gaussian random variables, as agent models. Assume that $y(x)$ is a Gaussian process that satisfies the following relation:

$$y \sim N\left(\mu \cdot I_n, \Sigma\right), \tag{17}$$

where I_n is a column vector with n-dimensional elements of 1 and μ is the expected mean value. $\sum$ is the n × n covariance matrix and the matrix elements are computed as:

$$\Sigma_{ij} = \mathrm{Cov}\left(y\left(x_i\right), y\left(x_j\right)\right) = \sigma^2 \mathcal{K}_\psi\left(x_i, x_j\right), \tag{18}$$

$$\psi = \left(\psi_1, \cdots, \psi_d\right)^T, \tag{19}$$

where $\mathcal{K}_{\psi}\left(x_i, x_j\right)$ is a pre-determined correlation function or kernel function, commonly known as the Gaussian kernel function.

$$\mathcal{K}_{\psi}\left(\mathrm{x}_i, \mathrm{x}_j\right) = \exp\left(-\sum_{s=1}^{d}\left(x_{is} - x_{js}\right)^2 \psi_s\right). \tag{20}$$

The distribution of $y(x)$ can be determined by the parameters $\theta = (\mu, \sigma^2, \psi)$.

$$l\left(\mu, \sigma^2, \psi\right) = \log f\left(y \mid \mu, \sigma^2, \psi; \mathcal{X}\right), \tag{21}$$

$$\propto -\frac{1}{2\sigma^2}\left(y - \mu \cdot I_n\right)^T R_{\psi}^{-1}\left(y - \mu \cdot I_n\right) - \frac{n}{2}\ln\left(\sigma^2\right) - \frac{1}{2}\ln\left|R_{\psi}\right|, \tag{22}$$

where R_{ψ} is the matrix of correlation coefficients for y. The maximum likelihood estimates of μ and σ^2 can be obtained by derivation of Eq. (22):

$$\mu_{\psi} = \frac{I_n^T R_{\psi}^{-1} y}{I_n^T R_{\psi}^{-1} I_n}, \sigma_{\psi}^2 = \frac{1}{n}\left(y - \mu_{\psi} \cdot I_n\right)^T R_{\psi}^{-1}\left(y - \mu_{\psi} \cdot I_n\right). \tag{23}$$

Bringing Eq. (22) into Eq. (21), $\hat{\psi} = argmax_{\psi \in D\psi} l\left(\mu_{\psi}, \sigma_{\psi}^2, \psi\right)$ can be obtained. D_{ψ} is a bounded region of the hypercube. Assume that $\hat{\mu} = \mu_{\hat{\psi}}$, $\hat{\sigma}^2 = \sigma_{\hat{\psi}}^2$, $\hat{\theta} = \left(\hat{\mu}, \hat{\sigma}^2, \hat{\psi}\right)$. $y\left(x \mid X\right)$ is the posterior distribution of $y(x)$, which is obtained according to the Bayesian formula.

$$y\left(x \mid \mathcal{X}\right) \sim N\left(\hat{y}\left(x\right), \hat{s}^2\left(x\right)\right), \tag{24}$$

where $\hat{y}\left(x\right)$ and $\hat{s}^2\left(\mathrm{x}\right)$ are given as follows:

$$\hat{y}\left(x\right) = \hat{\mu} + \rho_x^T R_{\hat{\psi}}^{-1}\left(y - \hat{\mu} \cdot 1_n\right), \tag{25}$$

$$\hat{s}^2\left(x\right) = \hat{\sigma}^2\left(1 - \rho_x^T R_{\hat{\psi}}^{-1} \rho_x\right), \tag{26}$$

where ρ_x is the correlation coefficient vector of $y(x)$ and y.

There are three commonly used acquisition functions, such as expected improvement (EI), upper confidence bound (UCB), and probability of improvement (POI). Assuming $y_{\min} = \min\left\{y\left(x_1\right), \ldots, y\left(x_n\right)\right\}$, EI calculates the expectation of the design point to improve relative to the current optimal objective function, the acquisition function is calculated as follows.

$$L\left(x\right) = \left(y_{\min} - \hat{y}\left(\mathrm{x}\right)\right)\Phi\left(\frac{y_{\min} - \hat{y}\left(\mathrm{x}\right)}{\hat{s}\left(\mathrm{x}\right)}\right) + \hat{s}\left(\mathrm{x}\right)\varphi\left(\frac{y_{\min} - \hat{y}\left(\mathrm{x}\right)}{\hat{s}\left(\mathrm{x}\right)}\right), \tag{27}$$

where Φ and φ are the density function and distribution function of the standard normal distribution, respectively.

2.4 Framework of Bayesian-PINN for Predicting Long-Term Production

The structure of the Bayesian-PINN model used in this work is shown in Fig. 2. The neural network contains input layer, hidden layer and output layer. The neural network uses a fully connected neural network NN(t), with 1 neuron in the input layer for the time t on the production prediction. The structure of the hidden layer is N × M (N is the number of hidden layers and M is the number of neurons contained in each hidden layer). There is one neuron in the output layer, and the output is the cumulative yield on the production prediction problem. The point is randomly generated in the solution domain t. The output is performed by a fully connected NN(t). In the model, the multistep gradient at each time is automatically differentiated, so that the $\frac{dQ}{dt}$ and $\frac{d^2Q}{dt^2}$ can be calculated and integrated into the loss function of the model.

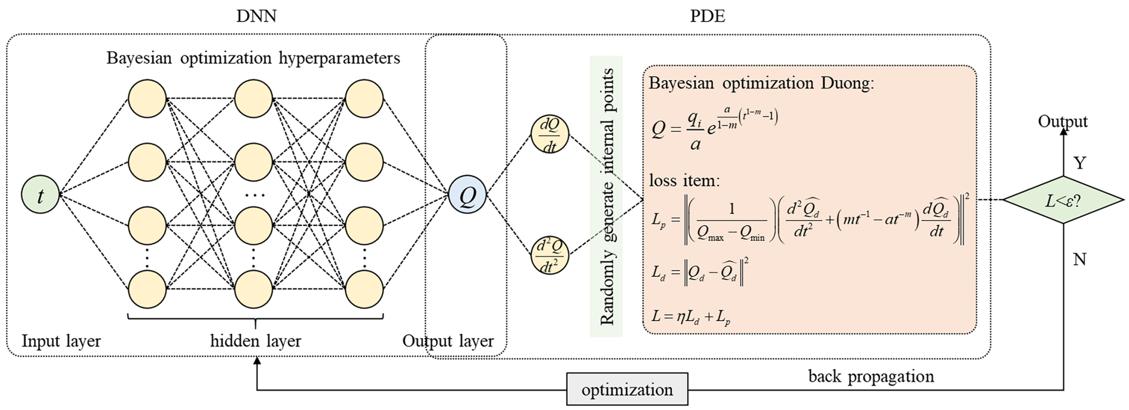


Figure 2: Network framework of Bayesian-PINN.

The loss function containing physical information is trained by the model and the loss continues to decrease until the convergence condition of the model is reached. On the contrary, the backpropagation updates the neural network parameters, and the loss continues to decrease by updating the parameters of the neural network. After the training, the normalized cumulative oil production prediction value for a long time in the future is output. Finally, the predicted value of cumulative oil production of oil wells is obtained by inverse normalization calculation of Eq. (12).

In this work the Bayesian optimization is used to optimize the learning rate λ, N, M, η, m and a in the Bayesian-PINN and the total loss function L of PINN is used as the objective function. During the iterative process, a Gaussian process is used as an agent model based on the distribution of the current data points in the loss function. The acquisition function is usually EI, through which the data points that are most likely to improve the current best solution are selected to determine the objective of the next optimization step. The distribution of the new data points in the loss function is re-observed through a Gaussian process and the model is updated. This process is iterated until the stopping condition is satisfied and the optimal parameters are finally determined.

The computational process of the Bayesian-PINN model is shown in Fig. 3. Input the production data from the training set and use Bayesian optimization to determine the parameters m and a of the Duong decline model. The internal points are randomly generated in the pre-calculated time interval, and the predicted production is calculated by forward propagation in PINN. Calculate the data loss L_d between the predicted and the actual productions. The physical loss L_p is calculated by using the production calculated by the internal points and the previously obtained optimization results m and a. Then, perform a weighted sum to calculate the total loss L. The total loss L is used to update the weights and biases in the neural network via backpropagation. Training ends when convergence condition is met. Then, the PINN is iteratively calculated by Bayesian optimization, and the output result after the maximum number of iterations is used as the optimal parameter combination of λ, η, M and N. The PINN model from the final iteration serves as the final production prediction model.

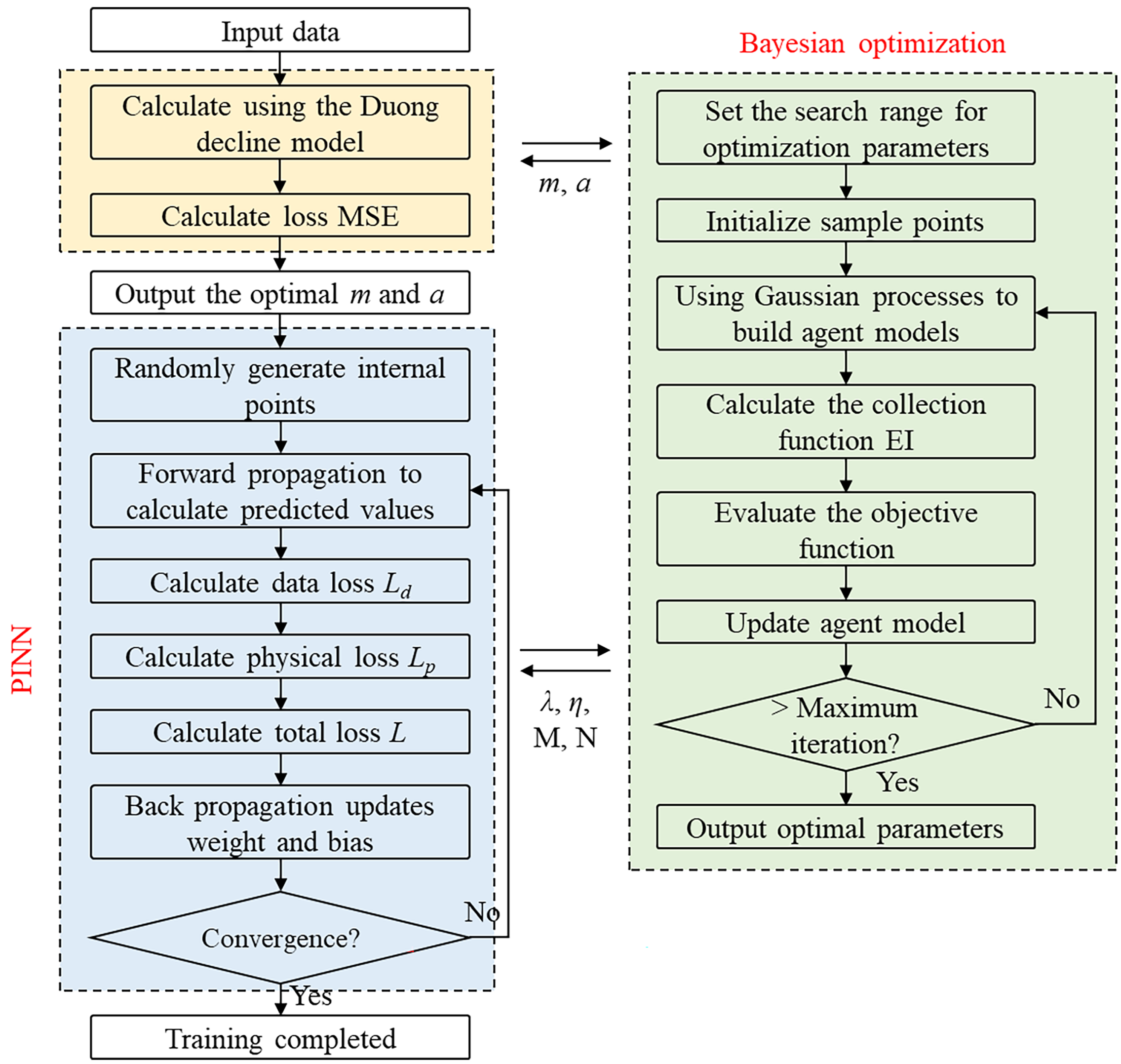


Figure 3: The calculational flowchart of Bayesian-PINN.

2.5 Model Evaluation Indicators

In this work, root mean squared error (RMSE) and mean absolute error (MAE) are selected to assess the prediction performance of long-term production.

RMSE responds to the deviation between predicted and observed values. The smaller the value, the better.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(y_i - f\left(x_i\right)\right)^2}{n}}, \tag{28}$$

where n is the total number of samples; y_i is the actual value of the ith target variable; $f(x_i)$ is the predicted value of the ith variable x_i in the model.

MAE denotes the mean of the absolute errors between the predicted and true values, again the smaller the value the better.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - f\left(x_i\right)\right|. \tag{29}$$

3 Results and Discussion

In this section, the Bayesian-PINN model is applied to forecast the long-term performance of multiple fractured horizontal wells using early-stage production data. Fig. 4 presents the production history of four representative wells in the Jimsar Sag. As observed, daily oil rates exhibit significant volatility driven by complex operational conditions, including sand plugging, inter-well fracturing interference, and manual interventions. Taking well JHW033 as a case study: the well was shut in on day 43 to accommodate fracturing operations in offset wells. Subsequently, it underwent sand washing, drilling, and fishing operations on day 305 following a sand plugging incident, a scheduled shut-in on day 359, and a workover on day 875. In contrast to the highly fluctuating daily rates, cumulative oil production profiles display smooth, monotonic trends that effectively mitigate the noise induced by these transient factors. To evaluate the long-term forecasting accuracy of the models, the production dataset is partitioned into training and testing sets using a 7:3 split. The training set is utilized to calibrate the Duong, Arps exponential, and Arps harmonic decline models, as well as the DNN, PINN, and LSTM algorithms. The testing set is strictly excluded from the training process and serves solely to validate the generalization performance of the models.

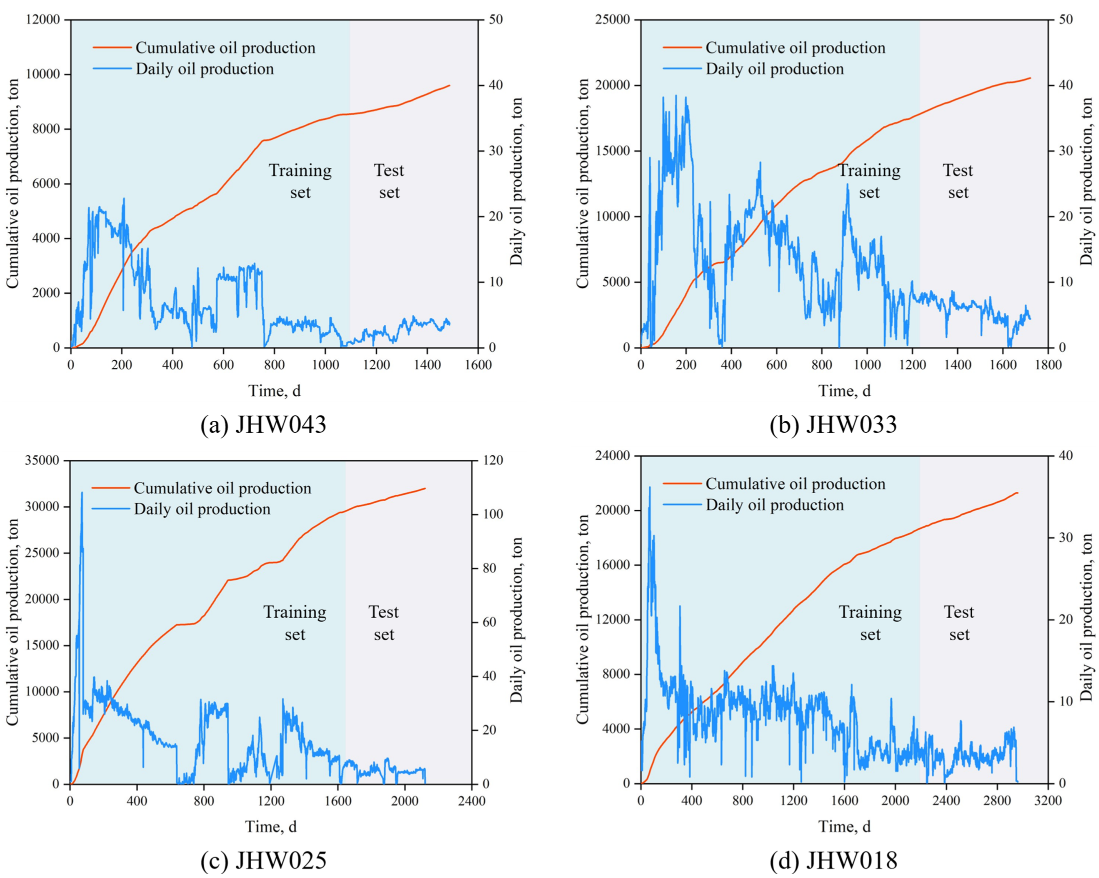


Figure 4: Production data for shale oil wells in Jimsar.

The search space of Bayesian optimization for the hyperparameters of the decline model in this work is as follows. Sliding window $w \in [2, 20]$, $M \in [1, 100]$, $N \in [1, 5]$, $\lambda \in [0.00001, 0.01]$, $m \in [0.5, 1.5]$, $a \in [0.5, 4]$, $\eta \in [0.0001, 10{,}000]$ and $D_i \in [0.00001, 0.01]$. The optimal results for the hyperparameters found by using Bayesian optimization in the training set are shown in Table 1. The m, a and D_i have certain differences due to different oil wells. Therefore, this work uses Bayesian optimization to fit the data in the training set of wells separately. To fairly evaluate each prediction method, the hyperparameters of DNN and LSTM are the same as those of PINN, and the parameters of Duong decline, exponential decline and harmonic decline are also obtained by fitting using Bayesian optimization.

Table 1: Hyperparameters at different prediction models.

Model	Hyperparameter
Duong decline	m and a are fitted by Bayesian optimization
Exponential decline	D_i is optimally fitted by Bayesian
Harmonic decline	D_i is optimally fitted by Bayesian
DNN	$\lambda = 1.8 \times 10^{-4}$, $M = 60$, $N = 2$
PINN	$\lambda = 1.8 \times 10^{-4}$, $M = 60$, $N = 2$, $\eta = 3050$, m and a are fitted by Bayesian optimization
LSTM	$\lambda = 1.8 \times 10^{-4}$, $M = 60$, $N = 2$, $w = 16$

3.1 Bayesian Optimization for DCA Model

Fig. 5 shows the optimization results of the parameters of the decline model in the training set using the traditional method in JHW043 well. In all the decline models, the daily oil production is treated as a linear feature, and the model parameters are further calculated by calculating the slopes and intercepts of the linear equations. The Duong decline model needs to logarithmically process the time, which will amplify the influence of the early production data on the linear equation during the fitting process. Most of the oil wells in Jimsar are unstable in the early period due to engineering factors, so it is necessary to deal with the unstable production in the early period. It can be seen from the figure that the Duong decline model has the characteristics of rapid decrease in the early period and slow decrease in the late period. Under the influence of logarithmic processing and unstable data, the early prediction value is slightly larger. Due to the large fluctuation of production, the fitting effect of the exponential decline model is low ($R^2 = 0.422$). The harmonic decline model sets q_i as the maximum daily oil production in this work, thus utilizing the slope of the linear equation to calculate D_i. From the results of the model calculations, the prediction results of these decline models in the early period have a large difference, and the difference in the later period is relatively small. However, it is worth noting that the traditional method of calculating the decline model has a strong human-guided factor, and the complex working conditions will amplify the influence of this human factor, so the simulation results are affected by personal experience, which may lead to a large deviation in the results.

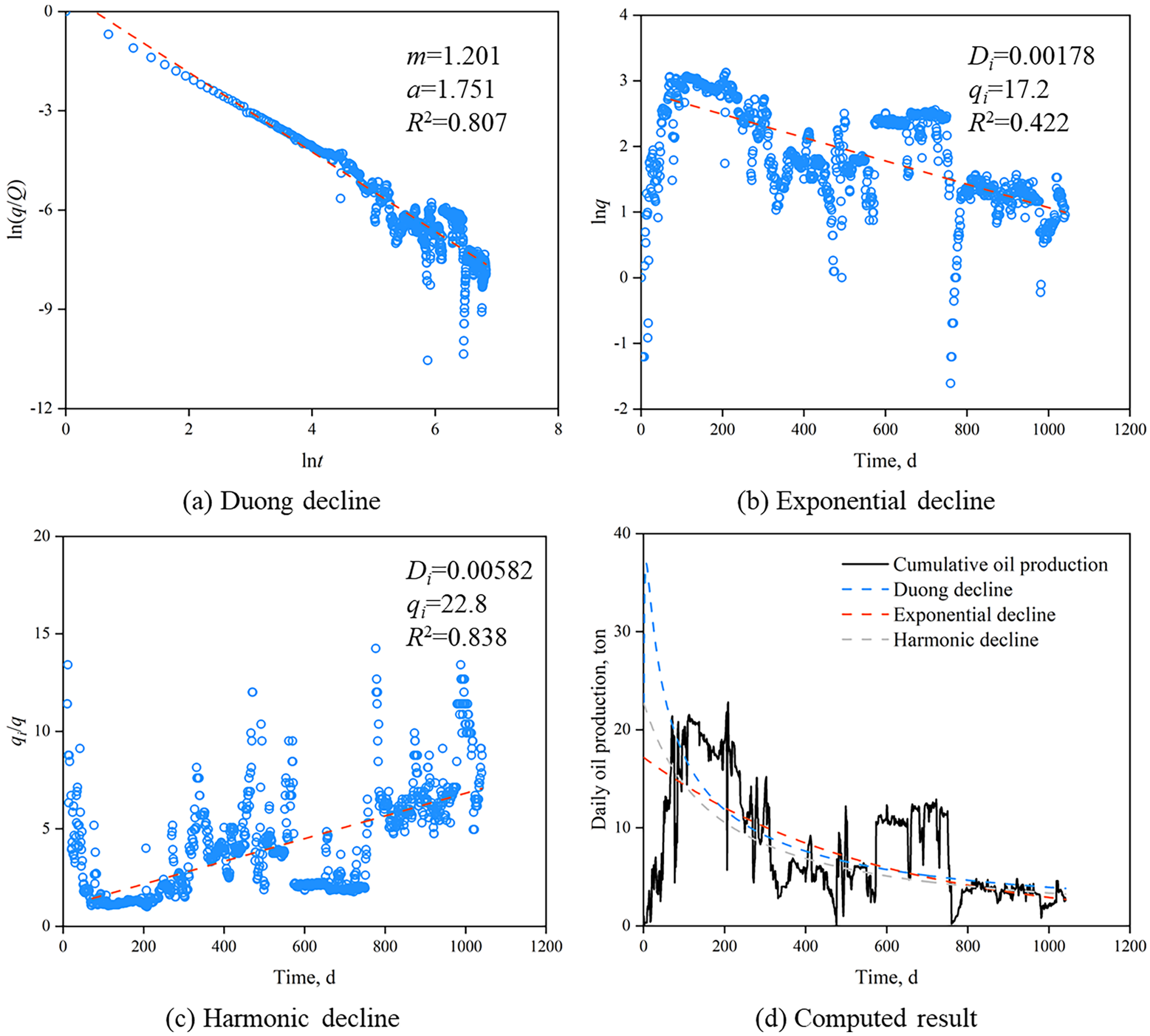


Figure 5: Fitting curves for each decline model on well JHW043.

This work compares and analyzes the prediction results of JHW043 well using traditional methods and Bayesian optimization to calculate the parameters of DCA (Fig. 6). In the traditional method, the exponential decline and harmonic decline models have relatively good prediction ability in the early period, and the exponential decline model has the best prediction effect in the late period. The prediction value of the Duong decline model is always larger than the actual production, which is caused by the large prediction value of the Duong decline model for the daily oil production in the early stage. Fig. 7 and Table 2 show the prediction performance of DCA calculated using traditional methods and Bayesian optimization. Compared with the traditional method, the prediction accuracy of the decline model using Bayesian optimization is significantly improved, and its predicted values are almost always located near the actual cumulative oil production. The most obvious improvement in the prediction performance is the Duong decline model, with an average improvement of about 68.5%, which greatly alleviates the effects of logarithmic processing and unstable data in the traditional method. The predictions of exponential and harmonic decline models are also improved by Bayesian optimization, with an average improvement of about 23.9% and 54.6%, respectively. Bayesian optimization avoids errors arising from significant fluctuations in daily oil production by directly fitting the cumulative oil production as the target. It also enables precise capture of data characteristics, thereby reducing prediction bias caused by subjective empirical differences.

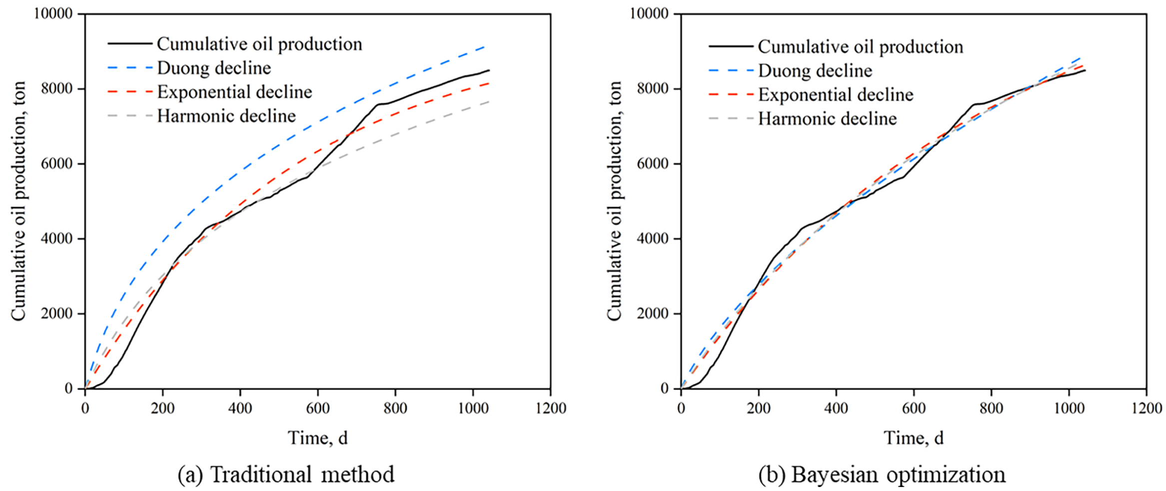


Figure 6: Prediction results of decline models at traditional methods and Bayesian optimization.

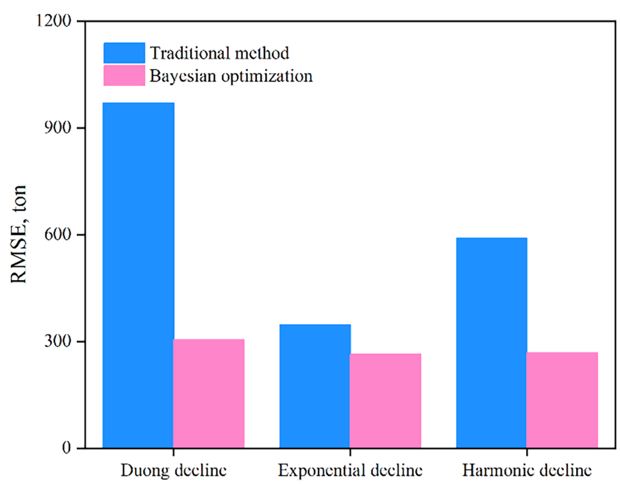


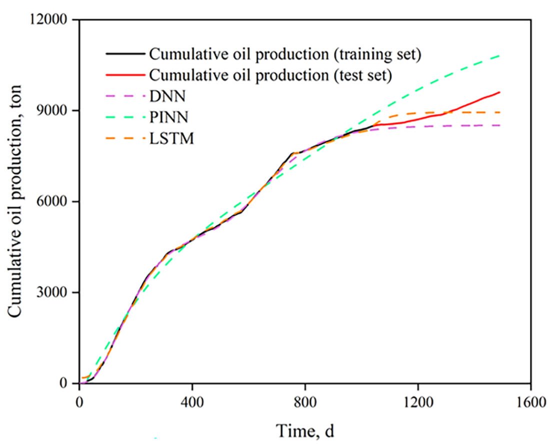
Figure 7: Fitting performance of traditional methods and Bayesian optimization.

Table 2: Prediction performance of traditional methods and Bayesian optimization in DCA models.

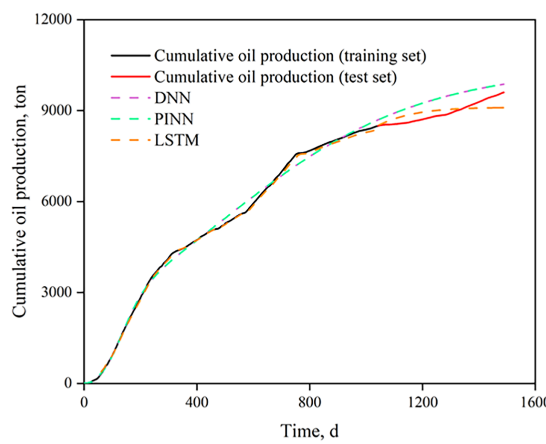
Methods	Traditional Method		Bayesian Optimization	
	RMSE, t	MAE, t	RMSE, t	MAE, t
Duong decline	970.4930	908.5421	305.4512	243.2861
Exponential decline	346.9710	306.5115	264.0001	215.3771
Harmonic decline	589.8976	464.8973	268.0386	220.4179

3.2 *Optimization Effects of Bayesian in Intelligent Prediction Models*

This section compares and analyzes the effect of Bayesian optimization on the prediction performance of intelligent models such as DNN, PINN and LSTM. The control group is the intelligent prediction model without Bayesian optimized hyperparameters. The hyperparameters of the control group are set as M = 20, N = 1, λ = 0.001, η = 10,000, and w = 10. Fig. 8 shows the prediction results of the intelligent prediction model of the JHW043 well in the control group and Bayesian optimization, respectively. The comparative analysis shows that the hyperparameters after Bayesian optimization have little influence on LSTM, and have relatively large influence on DNN and PINN. In the control group, the predicted values of DNN and PINN in the test set deviate significantly, showing conservative (predicted values lower than actual values) and radical (predicted values higher than actual values) prediction characteristics, respectively. This is due to the fact that PINN is physically constrained by ODE of Duong decline, while the prediction process of Duong decline model is relatively more radical. The growth trend of the predicted value of DNN in the test set is slower, and the predicted value is the smallest compared with other intelligent algorithms. After Bayesian optimization, the predicted values of DNN and PINN are quickly close to the actual values, and the prediction accuracies are significantly improved. The conservative DNN and radical PINN models are effectively alleviated, respectively.



(a) Control group



(b) Bayesian optimization

Figure 8: Predictions of intelligent models at control and Bayesian optimization.

Table 3 and Fig. 9 compare the predictive performance of the baseline models (control group) against those tuned via Bayesian optimization on the test set. The results indicate that Bayesian optimization universally enhances the forecasting capability of the intelligent algorithms. The PINN model exhibits the most substantial improvement, yielding an average performance gain of 53.1%, followed by the LSTM (21.1%) and the DNN (13.6%). Notably, while the DNN showed the most marginal accuracy improvement, it displayed significant variability (Fig. 8). Furthermore, the LSTM achieved the highest overall accuracy, the performance metrics of the DNN and PINN were closely aligned. However, the LSTM predictions in the test set exhibit a tendency toward slow rates of change; consequently, its reliability for extended forecast horizons warrants further investigation.

Table 3: Prediction performance of control group and Bayesian optimization.

Method	Control Group (Test Set)		Bayesian Optimization (Test Set)	
	RMSE, t	MAE, t	RMSE, t	MAE, t
DNN	539.2811	448.8641	466.0341	452.9827
PINN	1047.4829	1013.5173	491.1971	473.1497
LSTM	283.8283	235.5057	223.9413	194.9489

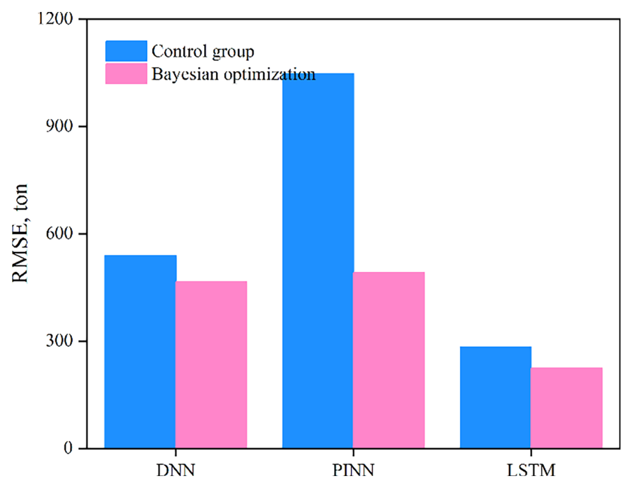


Figure 9: Prediction performance of control group and Bayesian optimization in intelligent models.

In summary, Bayesian optimization can use the agent model and the acquisition function to optimize the structure and hyperparameters in the neural network, so as to promote the intelligent algorithm to capture the data characteristics and strengthen its ability to apply to the long-term production prediction. Bayesian optimization has less impact on LSTM, but has a more obvious effect on DNN and PINN, which can help its prediction value to approach the actual value quickly, so that its prediction accuracy is significantly improved.

3.3 Long-Term Production Prediction Based on Bayesian-PINN

This section compares and studies the prediction performance of each long-term production prediction method in 4 fractured wells. Fig. 10 shows the prediction results of each prediction method for the training set and the test set. The future production predicted using the DCA methods are all larger than the actual production. In most cases (JHW043, JHW033, JHW018), there is a relationship of Duong decline > harmonic decline > exponential decline. The prediction results of Duong decline have the largest bias. LSTM has the worst performance on long-term production prediction. This is because LSTM has a better fitting effect for time-series data, but based on the concept of sliding window, as LSTM continues to predict future data, the proportion of future data in its sliding window will increase. With the gradual accumulation of errors in the iterative process, the final prediction results seriously deviate from the real results. PINN has similar prediction characteristics to DNN and LSTM in the training set, showing strong prediction ability, but the difference is that PINN and DNN have better prediction performance on future production.

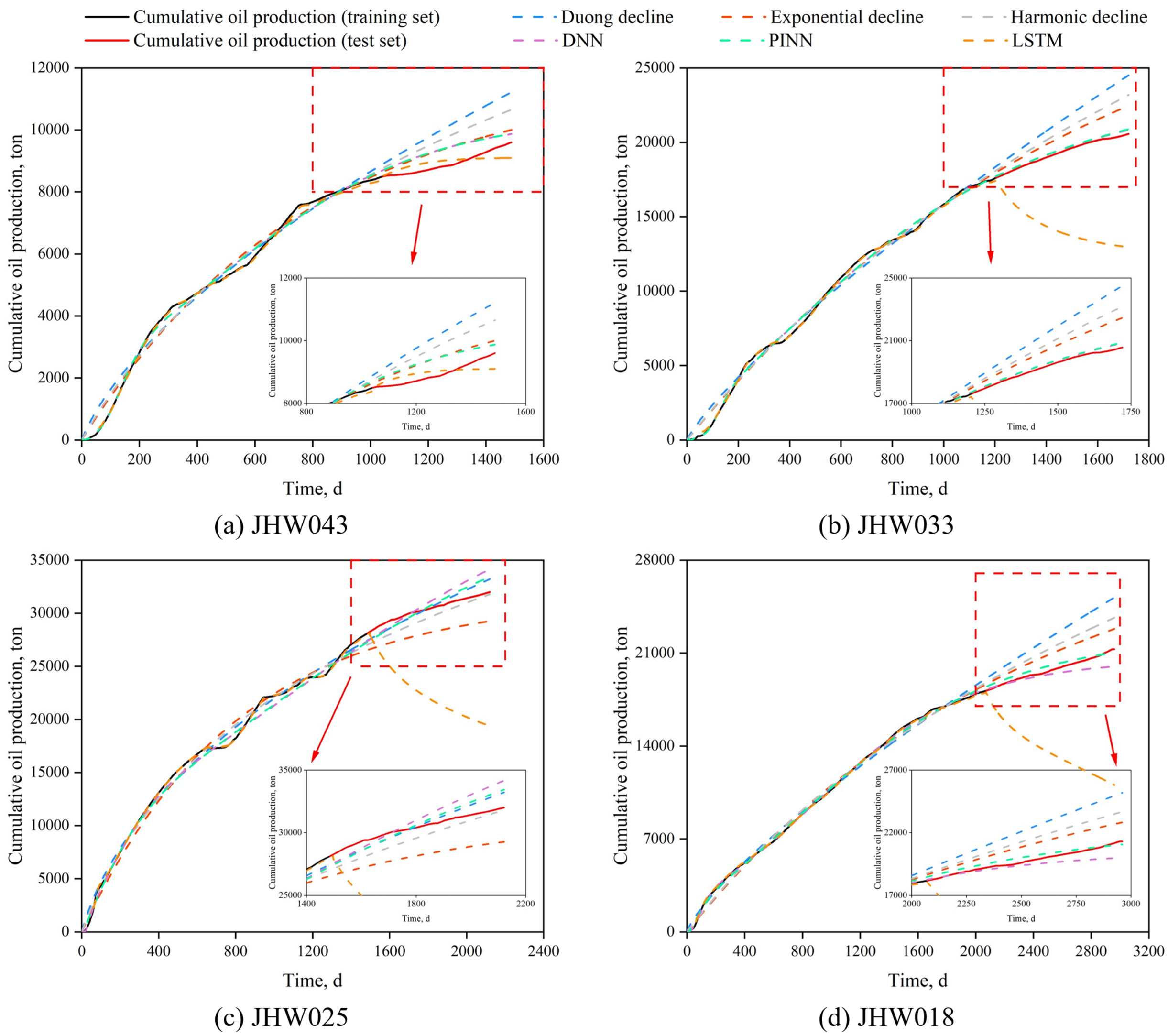
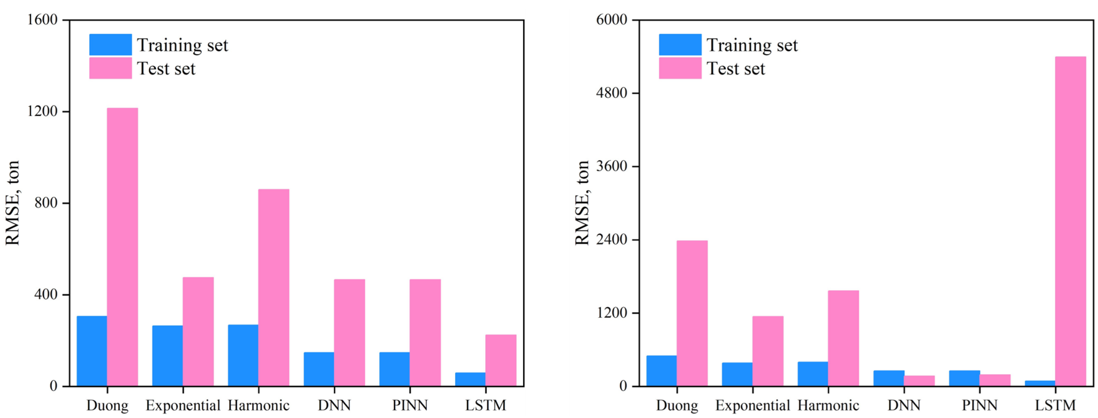


Figure 10: Prediction results of each method for long-term production (training set:test set = 7:3).

Table 4 and Fig. 11 show the prediction accuracy of each prediction method for the training and test sets. It can be seen from the results that the difference of the DCA method in the training set is small, but the prediction results for the test set are quite different. And the prediction ability in the training and test sets follows the same relationship. The algorithms that perform strongly in the training set have the same effect in the test set. The exponential decline performs best in the DCA method, and its prediction ability is 23.9% higher than Duong decline average. Intelligent algorithms such as PINN and DNN are slightly better than the traditional DCA methods in both the training and test sets. LSTM predicts the training set much better than the other algorithms, but the prediction effect on long-term yield is seriously deviated. This is because machine learning algorithms can effectively capture data features, which leads to better prediction accuracy of DNN, PINN and LSTM in the training set than DCA model. PINN with physical constraints is better than DNN, but there is no significant difference. The prediction ability is increased by about 16.1% on average, which is about 56.7% higher than that of exponential decline model. In terms of the performance of comprehensive training results, the best prediction method is PINN, which can fully utilize DNN to effectively capture data features while also integrating the physical laws of shale oil production decline.

Table 4: Production prediction accuracy of each prediction method (training set:test set = 7:3).

Well Number	Method	RMSE, t		MAE, t	
		Training Set	Test Set	Training Set	Test Set
JHW043	Duong decline	305.4512	1214.2431	243.2861	1161.4553
	Exponential decline	264.0001	475.3104	215.3771	460.9846
	Harmonic decline	268.0386	859.4906	220.4179	825.9737
	DNN	147.5052	466.0341	114.5367	452.9827
	PINN	147.5229	466.1971	114.5584	453.1497
	LSTM	58.3017	223.9413	49.3223	194.9489
JHW033	Duong decline	305.4512	1214.2431	243.2861	1161.4553
	Exponential decline	264.0001	475.3104	215.3771	460.9846
	Harmonic decline	268.0386	859.4906	220.4179	825.9737
	DNN	147.5052	466.0341	114.5367	452.9827
	PINN	147.5229	466.1971	114.5584	453.1497
	LSTM	58.3017	223.9413	49.3223	194.9489
JHW025	Duong decline	569.5076	652.1003	483.9403	575.2621
	Exponential decline	765.9256	2290.2088	650.2399	2275.7221
	Harmonic decline	575.5304	897.0031	459.7117	822.8341
	DNN	524.3917	1109.8443	422.7991	921.1825
	PINN	540.1325	784.8445	451.5947	691.3709
	LSTM	79.0214	8531.4381	60.4361	7817.3933
JHW018	Duong decline	347.0931	2630.8761	394.9455	2466.3539
	Exponential decline	321.7532	1208.7297	277.2003	1152.0331
	Harmonic decline	324.8019	1693.2146	286.0808	1595.1737
	DNN	88.6499	548.9311	75.1359	411.5229
	PINN	157.1789	295.3388	134.5394	277.9787
	LSTM	65.6941	6524.7618	54.7926	5974.8042



(a) JHW043 (b) JHW033

Figure 11: (continued)

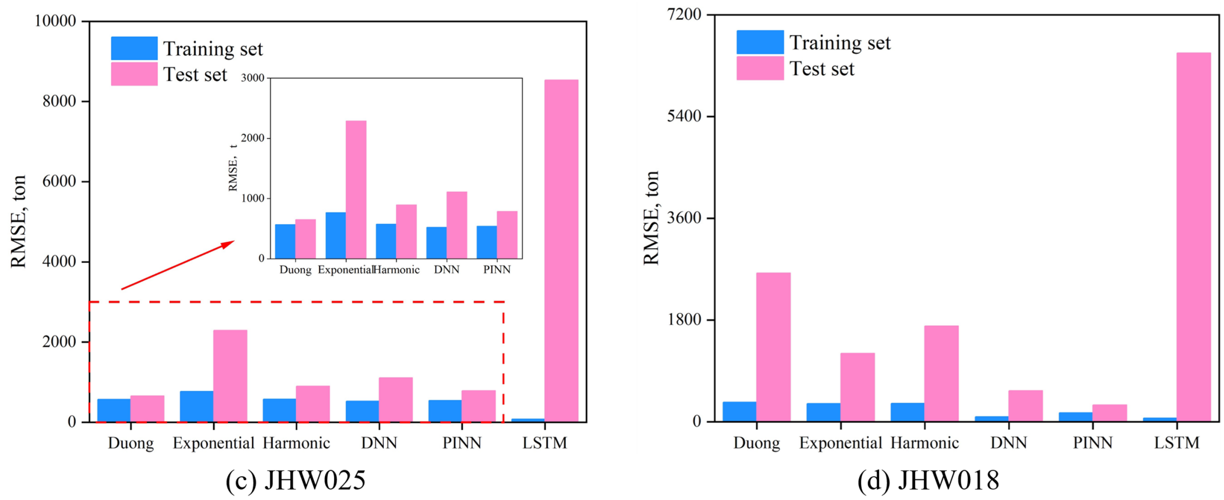


Figure 11: MSE of each forecasting method in long-term yield forecasting (training set:test set = 7:3).

In the preceding analysis, the amount of data in the training set exceeded that of the test set. However, in practical field applications, the required forecasting horizon is often significantly longer than the current production cycle, particularly when predicting the estimated ultimate recovery (EUR). To rigorously evaluate the predictive capability of each algorithm over such extended periods, this section re-partitions the production data into training and testing sets using a 4:6 split.

Fig. 12 shows the prediction ability of each prediction algorithm for long-term production when training set:test set = 4:6. As can be seen from the Figure, the features and relationships of each prediction algorithm still exist, but the prediction accuracy of each method on the test set slightly decreases as the number of samples in the test set increases. The most serious impact is the DNN algorithm, for the production data with certain fluctuations (JHW043, JHW025), its prediction ability on the test set is significantly weakened, and the prediction trend is closer to the exponential decline model. For relatively smooth production data (JHW033, JHW018), its prediction trend is closer to the harmonic decline model. This indicates that the DNN is less resistant to interference and is susceptible to the influence of production system and complex working conditions. After adding physical constraints, PINN shows stronger anti-interference ability, and its prediction ability is significantly improved for production data with fluctuating characteristics. The prediction trend of the LSTM algorithm gradually appears. With the increase of prediction time, its predicted long-term production gradually converges to a constant.

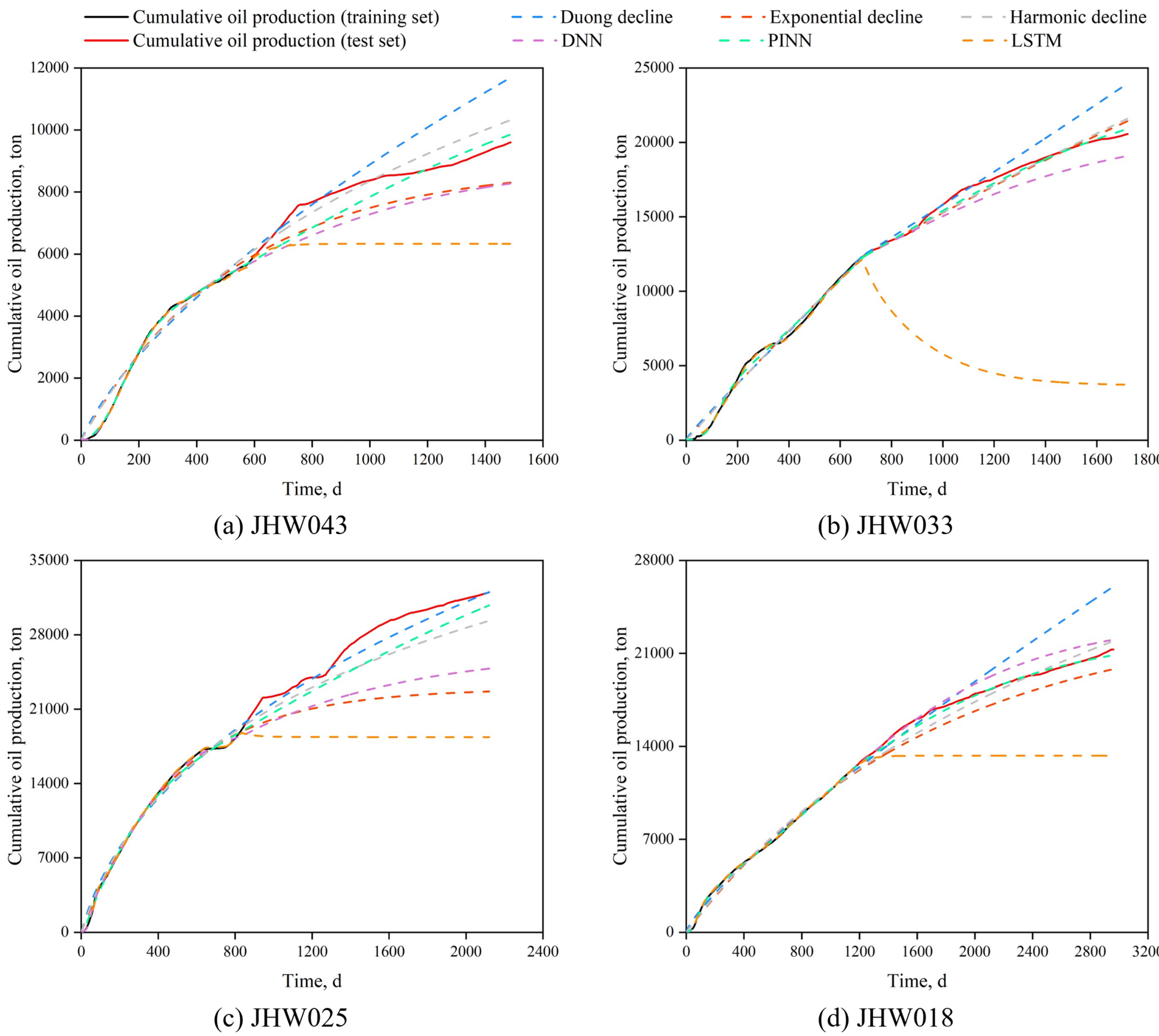
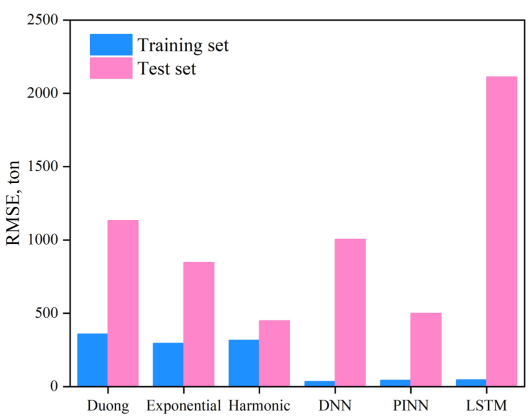


Figure 12: Prediction results of each prediction method for long-term production (training set:test set = 4:6).

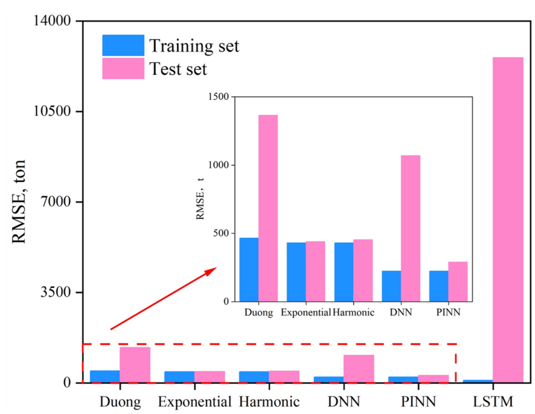
Table 5 and Fig. 13 show the prediction accuracy of each prediction method for long-term production when the ratio of training set to test set is 4:6. As the future prediction time increases, the intelligent models (DNN, PINN and LSTM) still have significantly higher prediction ability for the training set than the DCA models, but the prediction ability in the test set changes significantly. Among them, the algorithm with the best prediction ability in the DCA changes from the exponential decline to the harmonic decline, and prediction ability of the harmonic decline is about 37.9% higher than the exponential decline. The prediction ability of DNN for future long-term production is obviously weakened, and the advantages of experience-guided DCA are gradually emerging. The prediction ability of PINN on future long-term production is significantly higher than that of DNN, and the prediction ability is increased by about 61.7% on average, which is about 21.4% higher than that of harmonic decline model. The ability to predict future long-term production roughly meets the trend of PINN > DNN ≈ DCA > LSTM. In addition, the ODE and physical loss introduced by PINN generate additional consumption. When Epoch is 10,000, the training time of DNN and PINN is 22 and 55 s, respectively. Since the production prediction needs to pursue high prediction accuracy and does not require real-time calculation, these additional losses are completely acceptable.

Table 5: Production prediction accuracy of each prediction method (training set:test set = 4:6).

Well Number	Method	RMSE, t		MAE, t	
		Training Set	Test Set	Training Set	Test Set
JHW043	Duong decline	358.5726	1131.5891	301.8511	868.3157
	Exponential decline	294.8748	846.7465	235.6182	800.9106
	Harmonic decline	316.4763	448.8219	265.4743	374.1681
	DNN	34.8072	1004.9895	27.4593	980.2631
	PINN	43.4968	499.5526	31.6584	422.0396
	LSTM	46.4588	2111.9961	35.4367	1945.7245
JHW033	Duong decline	466.5636	1365.7512	360.4719	950.8223
	Exponential decline	432.1402	439.9175	350.4404	352.4601
	Harmonic decline	432.1183	454.0231	350.1881	360.0986
	DNN	224.1857	1069.6167	181.3378	933.6481
	PINN	224.1857	290.7503	181.3378	225.4184
	LSTM	99.2362	12,581.1659	86.0922	11,751.931
JHW025	Duong decline	624.3025	895.6322	553.6621	751.4469
	Exponential decline	359.9764	6264.6421	299.8001	5658.0087
	Harmonic decline	508.6534	2332.9931	441.8619	2132.8007
	DNN	199.2035	5252.7112	132.2244	4877.6322
	PINN	330.6985	2285.0052	275.2161	2170.0829
	LSTM	113.5959	9582.8215	82.4257	8841.1626
JHW018	Duong decline	217.1678	2317.0031	179.6529	1753.2359
	Exponential decline	309.3814	1194.5962	266.9365	1176.4617
	Harmonic decline	289.8119	652.0833	256.2402	584.0303
	DNN	61.8848	816.1821	53.2621	678.0561
	PINN	56.8582	267.4215	47.6295	196.4033
	LSTM	41.5592	5064.7721	37.5485	4543.4367



(a) JHW043



(b) JHW033

Figure 13: (Continued)

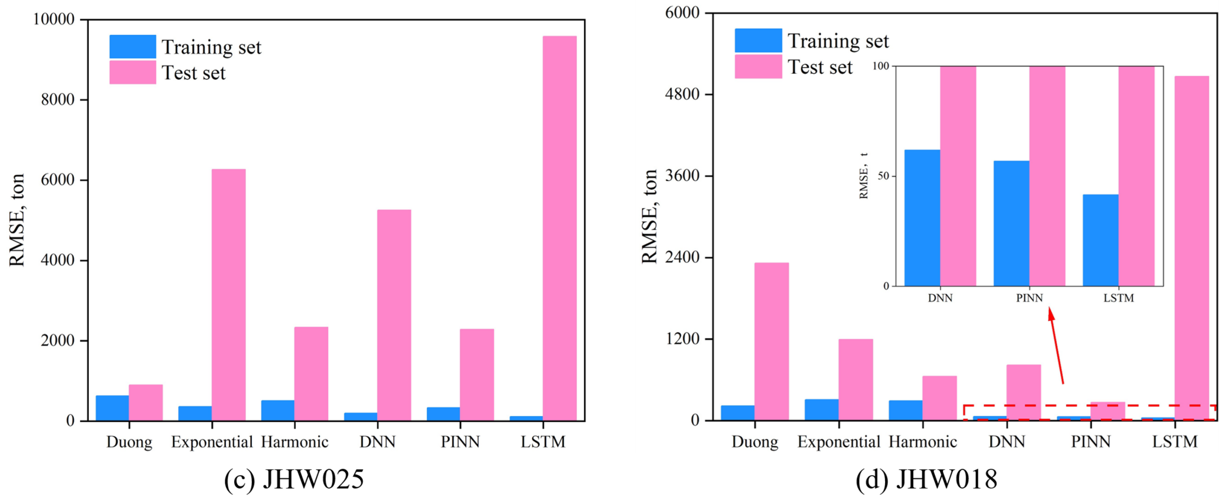


Figure 13: MSE of each method in long-term production (training set:test set = 4:6).

The EUR refers to the amount of recoverable oil and gas calculated from the beginning of production to a specified date, a certain abandonment pressure limit or the minimum daily oil and gas production of a single well as the cut-off value, which is one of the key indicators of the development potential and effect of shale oil [37]. In this section, the PINN is used as the EUR prediction tool to train the cumulative oil production data, and the cumulative oil production corresponding to a daily production rate of 0.1 t/d is used as the EUR of the well.

The results of long-term production prediction of four fractured wells using PINN are shown in Fig. 14. The calculated EUR values of JHW043, JHW033, JHW025 and JHW018 wells are 13,500, 36,864, 49,088 and 32,159 t, respectively. Studies have shown that the horizontal length of horizontal wells is an important factor affecting EUR. Under the same conditions, the longer the horizontal section is, the higher the production capacity of the well is. The horizontal length of the four fractured wells are 1389, 1532, 1200, and 1794 m, respectively, with a relatively large span. The unitization of EUR using the horizontal length can offset the effect of horizontal length on the production capacity, thus accurately describing the mapping relationship between geology, engineering parameters and production capacity.

$$\mathrm{EUR_Z} = \frac{\mathrm{EUR}}{L}, \tag{30}$$

where, $\mathrm{EUR_z}$ is the EUR per unit length, t/m; L is the horizontal length of horizontal well, m.

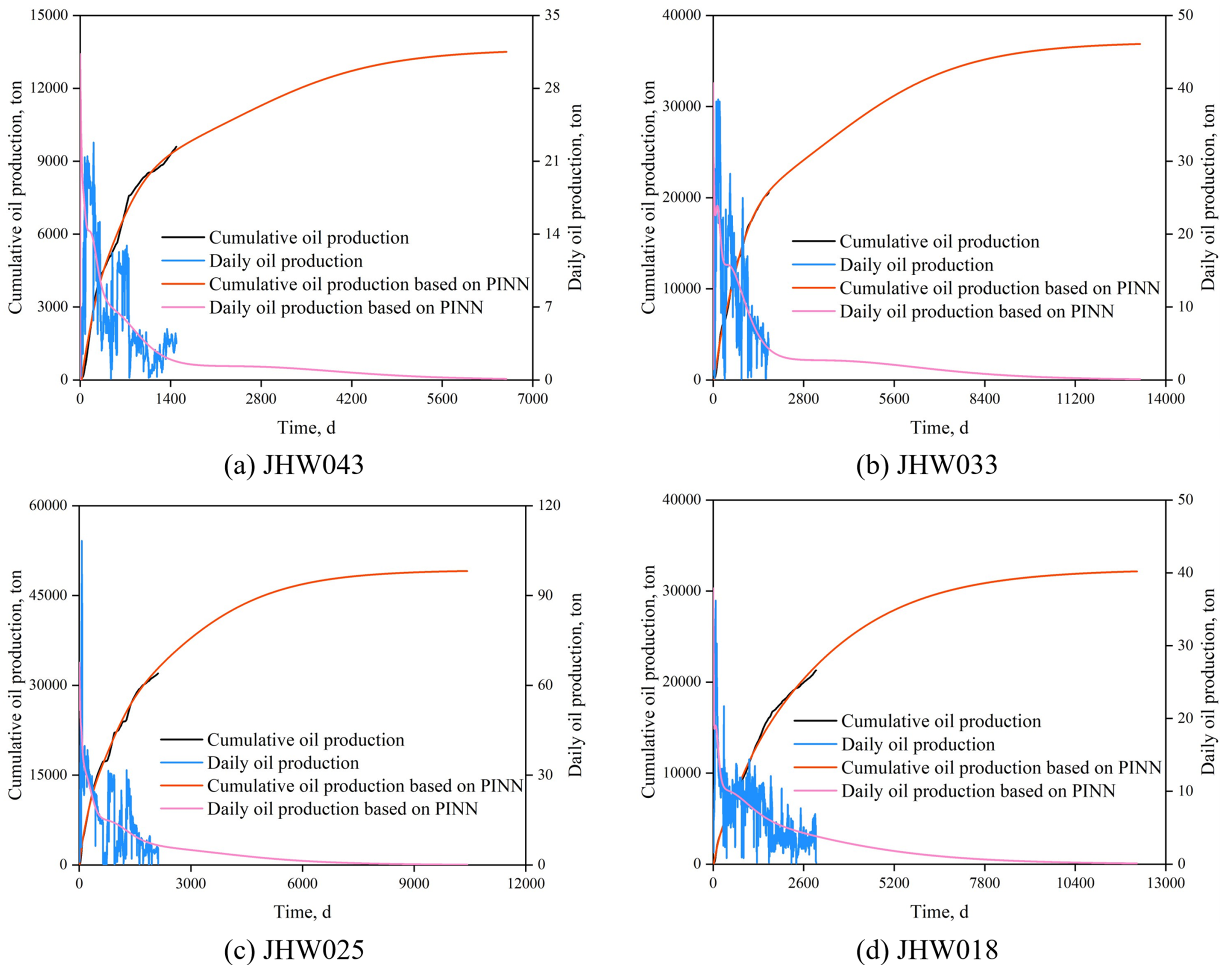
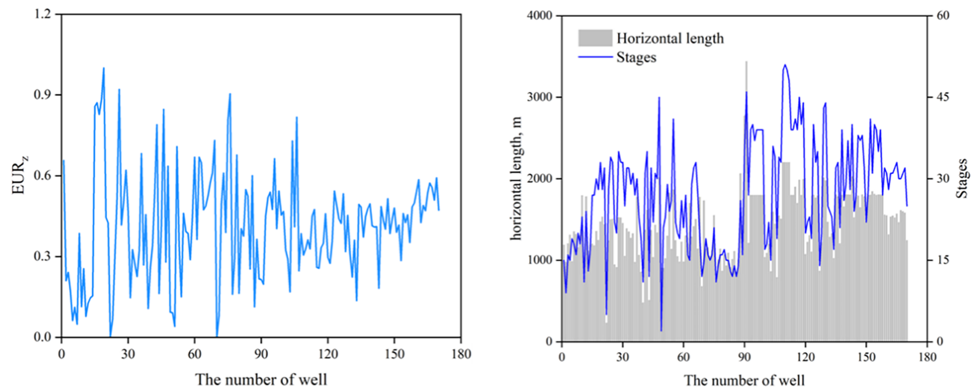


Figure 14: Predicted results of fractured well EUR using PINN.

The calculated EUR_z of JHW043, JHW033, JHW025 and JHW018 wells are 9.7, 24.1, 40.9, and 17.9 t/m, respectively. The horizontal lengths of wells JHW025 and JHW043 are almost the same, but the EUR of JHW025 is 3.6 times higher than that of JHW043. Meanwhile, the EUR of JHW018 is 58.1% higher than that of JHW043, but the EUR_z is only 45.8% higher than that of JHW043. Compared with EUR, EUR_z can more objectively describe the contribution of fracturing performance to production of shale oil, and is more suitable as the target variable of fracturing prediction and optimization model.

PINN predicts future long-term production and EUR by learning from production data of shale oil well. If the existing production data of the oil well is insufficient, the prediction accuracy of PINN after learning will be low. Research has shown that when production days fall below 200, the prediction accuracy is relatively low. Therefore, this paper sets 200 days of oil production as the threshold. If the current oil production days of an oil well are below this threshold, it is considered that declining analysis and intelligent algorithms cannot be used to calculate EUR. Due to the high confidentiality level of this important data, EUR is normalized to EUR_z for representation in this work.

Fig. 15a presents the EUR_z predicted by Bayesian-PINN for wells in the Jimsar region. During the initial development phase, well productivity was low and exhibited high volatility, reflecting limited reservoir characterization and the exploratory nature of early hydraulic fracturing techniques. Through iterative technological refinement and continuous optimization of fracturing equipment and design, subsequent operations achieved greater maturity, resulting in significantly enhanced and stabilized production performance. Fig. 15b illustrates the evolving trends in horizontal wellbore length and the number of fracturing stages for these stimulated wells. Early development was characterized by relatively short laterals and fewer stages; the first 90 wells had average values of 1243 m and 22 stages, respectively. As development progressed, these parameters increased substantially. For wells following the 90th, the average lateral length and stage count rose to 1643 m and 33, representing increases of 32.2% and 50.0%, respectively. The more rapid growth in stage count relative to lateral length indicates a trend toward reduced stage spacing, demonstrating an improved understanding of the reservoir and advances in completion strategy. The developmental trajectory of stimulation technology in this block aligns closely with the EUR trend predicted by the Bayesian-PINN model. This correlation validates the model's effectiveness for EUR forecasting in this context.



(a) EUR$_z$ predicted value (b) Horizontal length and stage

Figure 15: Sand and liquid dosage and EURz predicted value.

4 Conclusions

In this work, a new Bayesian-PINN prediction model is proposed to predict long-term production with ODE of Duong decline as physical constraint. Based on the production data of Jimsar shale oil, the optimization effect of Bayesian optimization on the prediction ability of Duong decline and PINN model is compared and analyzed. In addition, the prediction ability of Bayesian-PINN for different future time steps is analyzed. The main findings are as follows:

(1) Bayesian optimization can not only directly use the cumulative oil production as the fitting target, avoiding the problem of secondary amplification of the error, but also accurately captures the data characteristics and avoiding the problem of large differences in the prediction results caused by complex working conditions and personal experience. Compared with the traditional DCA method, the Duong decline using Bayesian optimization improves about 68.5% on average.

(2) Bayesian optimization can use the agent model and the acquisition function to find the optimal structure and hyperparameters in the neural network, which has less influence on LSTM, but has more significant improvement on DNN and PINN. The prediction performance of PINN is improved by about 53.1% on average.

(3) Under physical constraints, the prediction performance of Bayesian-PINN on future long-term production is about 61.7% higher than that of DNN model, and about 21.4% higher than that of harmonic decline. In general, the prediction ability for the future long-term production of oil wells generally meets PINN > DNN ≈ DCA > LSTM.

The Bayesian-PINN established in this work provides a framework for production prediction combined with physical constraints. In this study, Bayesian optimization is used to determine the parameters of domain knowledge, and differential equations are derived and used as soft constraints to influence the change of production. The Bayesian-PINN established in this work is suitable for the production prediction of unconventional reservoirs after fracturing. If applied to conventional reservoirs with high permeability, the use of the Arps decline model as a physical constraint may be better than the Duong decline model. Particularly in the late production stage, when the flow regime shifts to boundary-dominated flow, the Duong decline model often performs poorly in forecasting. In contrast, the Bayesian-PINN can learn dynamic characteristics from production data; learning production data from the boundary-dominated flow stage is expected to compensate for this limitation of the Duong model. Automated Machine Learning (AutoML) lacks physical interpretability, whereas Bayesian-PINN embed physical laws, making them more suitable for oil production prediction. PINN ensure physical consistency, while AutoML is preferable for large-scale, physics-agnostic applications. Furthermore, due to the influence of error accumulation, the introduction of physical constraints is an effective method to improve the prediction accuracy of intelligent algorithms for long-term production prediction. In the future, machine learning will be applied to the field of oil and gas development, and the integration of domain knowledge is still an important field for machine learning to alleviate errors and improve accuracy.

Acknowledgement: None.

Funding Statement: This research was funded by National Natural Science Foundation of China, grant numbers U23B2084 and 5241002, and Science Foundation of China University of Petroleum, Beijing, grant number 2462025XKBH032.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Longqiao Hu, Yunjin Wang, Jiacheng Yin, Siyu Zhang, Fujian Zhou; data collection: Jia Liu, Jiacheng Yin, Qi Wu; analysis and interpretation of results: Jia Liu, Mengyu Li, Qi Wu, Jiawei Li, Fujian Zhou; draft manuscript preparation: Longqiao Hu, Yunjin Wang. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. McMahon TP, Larson TE, Zhang T, Shuster M. Geologic characteristics, exploration and production progress of shale oil and gas in the United States: an overview. Pet Explor Dev. 2024;51(4):925–48. doi:10.1016/S1876-3804(24)60516-1.
2. Wang Y, Zheng L, Chen G, Kong M, Yuan L, Wang B, et al. A genetic particle swarm optimization with policy gradient for hydraulic fracturing optimization. SPE J. 2025;30(2):560–72. doi:10.2118/223956-PA.
3. Wang Y, Wang D, Li Y, Zheng L, Su H, Zhang Y, et al. Numerical simulation: diverting study and analysis on nanoparticle-viscoelastic-surfactant acid in high-temperature carbonate reservoir. Phys Fluids. 2024;36(7):076627. doi:10.1063/5.0220673.
4. Gu S, Wang J, Xue L, Tu B, Yang M, Liu Y. Deep-learning-based production decline curve analysis in the gas reservoir through sequence learning models. Comput Model Eng Sci. 2022;131(3):1579–99. doi:10.32604/cmes.2022.019435.

5. Arnold R, Anderson R. Preliminary report on the Coalinga oil district, Fresno and Kings counties, California. Washington, DC, USA: U.S. Government Printing Office; 1908. 142 p. doi:10.3133/b357.
6. Arps JJ. Analysis of decline curves. Trans AIME. 1945;160(1):228–47. doi:10.2118/945228-G.
7. Denney D. Gas-reserves estimation in resource plays. J Pet Technol. 2010;62(12):65–7. doi:10.2118/1210-0065-JPT.
8. Rushing JA, Perego AD, Sullivan RB, Blasingame TA. Estimating reserves in tight gas sands at HP/HT reservoir conditions: use and misuse of an Arps decline curve methodology. In: Proceedings of the SPE Annual Technical Conference and Exhibition; 2007 Nov 11–14; Anaheim, CA, USA. doi:10.2118/109625-MS.
9. Fetkovich MJ, Vienot ME, Bradley MD, Kiesow UG. Decline-curve analysis using type curves—case histories. SPE Form Eval. 1987;2(4):637–56. doi:10.2118/13169-PA.
10. Ilk D, Rushing JA, Perego AD, Blasingame TA. Exponential vs. hyperbolic decline in tight gas sands—understanding the origin and implications for reserve estimates using Arps' decline curves. In: Proceedings of the SPE Annual Technical Conference and Exhibition; 2008 Sep 21–24; Denver, CO, USA. doi:10.2118/116731-MS.
11. Duong AN. An unconventional rate decline approach for tight and fracture-dominated gas wells. In: Proceedings of the Canadian Unconventional Resources and International Petroleum Conference; 2010 Oct 19–21; Calgary, AB, Canada. doi:10.2118/137748-MS.
12. Aboaba A, Cheng Y. Estimation of fracture properties for a horizontal well with multiple hydraulic fractures in gas shale. In: Proceedings of the SPE Eastern Regional Meeting; 2010 Oct 12–14; Morgantown, WV, USA. doi:10.2118/138524-MS.
13. Baihly JD, Malpani R, Altman R, Lindsay G, Clayton R. Shale gas production decline trend comparison over time and basins—revisited. In: Proceedings of the SPE/AAPG/SEG Unconventional Resources Technology Conference; 2015 Jul 20–22; San Antonio, TX, USA. doi:10.15530/URTEC-2015-2172464.
14. Nobakht M, Mattar L, Moghadam S, Anderson DM. Simplified yet rigorous forecasting of tight/shale gas production in linear flow. In: Proceedings of the SPE Western Regional Meeting; 2010 May 27–29; Anaheim, CA, USA. doi:10.2118/133615-MS.
15. Wang J, Xie H, Matthai SK, Hu J, Li C. The role of natural fracture activation in hydraulic fracturing for deep unconventional geo-energy reservoir stimulation. Pet Sci. 2023;20(4):2141–64. doi:10.1016/j.petsci.2023.01.007.
16. Liu X, Sun M, Lin B, Gu S. SGP-GCN: a spatial-geological perception graph convolutional neural network for long-term petroleum production forecasting. Energy Eng. 2025;122(3):1053–72. doi:10.32604/ee.2025.060489.
17. Wang Y, Zhou F, Yin J, Wang Y, Zheng L, Huang G, et al. An integrated sequence-to-sequence framework for well production prediction. Phys Fluids. 2025;37(10):107140. doi:10.1063/5.0301508.
18. Wang Q, Song X, Li R. A novel hybridization of nonlinear grey model and linear ARIMA residual correction for forecasting U.S. shale oil production. Energy. 2018;165(6339):1320–31. doi:10.1016/j.energy.2018.10.032.
19. Gupta S, Fuehrer F, Jeyachandra BC. Production forecasting in unconventional resources using data mining and time series analysis. In: Proceedings of the SPE/CSUR Unconventional Resources Conference—Canada; 2014 Sep 30–Oct 2; Calgary, AB, Canada. doi:10.2118/171588-MS.
20. Olominu O, Sulaimon AA. Application of time series analysis to predict reservoir production performance. In: Proceedings of the SPE Nigeria Annual International Conference and Exhibition; 2014 Aug 5–7; Lagos, Nigeria. doi:10.2118/172395-MS.
21. Razak SM, Cornelio J, Cho Y, Liu H, Vaidya R, Jafarpour B. Transfer learning with recurrent neural networks for long-term production forecasting in unconventional reservoirs. SPE J. 2022;27(4):2425–42. doi:10.2118/209594-PA.
22. Mohammadian E, Mohamadi-Baghmolaei M, Azin R, Hadavimoghaddam F, Rozhenko A, Liu B. RNN-based CO_2 minimum miscibility pressure (MMP) estimation for EOR and CCUS applications. Fuel. 2024;360(12):130598. doi:10.1016/j.fuel.2023.130598.
23. Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions. Eur J Oper Res. 2018;270(2):654–69. doi:10.1016/j.ejor.2017.11.054.

24. Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); 2014 Oct 25–29; Doha, Qatar. p. 1724–34. doi:10.3115/v1/D14-1179.
25. Alimohammadi H, Rahmanifard H, Chen N. Multivariate time series modelling approach for production forecasting in unconventional resources. In: Proceedings of the SPE Annual Technical Conference and Exhibition; 2020 Oct 26–29; Virtual. doi:10.2118/201571-MS.
26. Bao A, Gildin E, Huang J, Coutinho EJ. Data-driven end-to-end production prediction of oil reservoirs by EnKF-enhanced recurrent neural networks. In: Proceedings of the SPE Latin American and Caribbean Petroleum Engineering Conference; 2020 Jul 27–31; Virtual. doi:10.2118/199005-MS.
27. Song X, Liu Y, Xue L, Wang J, Zhang J, Wang J, et al. Time-series well performance prediction based on long short-term memory (LSTM) neural network model. J Pet Sci Eng. 2020;186(2):106682. doi:10.1016/j.petrol.2019.106682.
28. Chaikine IA, Gates ID. A machine learning model for predicting multi-stage horizontal well production. J Pet Sci Eng. 2021;198(1):108133. doi:10.1016/j.petrol.2020.108133.
29. Werneck RDO, Prates R, Moura R, Gonçalves MM, Castro M, Soriano-Vargas A, et al. Data-driven deep-learning forecasting for oil production and pressure. J Pet Sci Eng. 2022;210(1):109937. doi:10.1016/j.petrol.2021.109937.
30. Yang R, Qin X, Liu W, Huang Z, Shi Y, Pang Z, et al. A physics-constrained data-driven workflow for predicting coalbed methane well production using artificial neural network. SPE J. 2022;27(3):1531–52. doi:10.2118/205903-PA.
31. Zhang X, Lu Y, Jin Y, Chen M, Zhou B. An adaptive physics-informed deep learning method for pore pressure prediction using seismic data. Pet Sci. 2024;21(2):885–902. doi:10.1016/j.petsci.2023.11.006.
32. Xin L, Rao X, Peng X, Xu Y, Chen J. Production dynamic prediction method of waterflooding reservoir based on deep convolution generative adversarial network (DC-GAN). Energy Eng. 2022;119(5):1905–22. doi:10.32604/ee.2022.019556.
33. Li B. A productivity prediction method based on artificial neural networks and particle swarm optimization for shale-gas horizontal wells. Fluid Dyn Mater Process. 2023;19(10):2729–48. doi:10.32604/fdmp.2023.029649.
34. Qian Q, Lu M, Zhong A, Yang F, He W, Li M. Production capacity prediction method of shale oil based on machine learning combination model. Energy Eng. 2024;121(8):2167–90. doi:10.32604/ee.2024.049430.
35. Wang Y, Zhou F, Su H, Zheng L, Li M, Yu F, et al. Intelligent optimization method of fracturing parameters for shale oil reservoirs in Jimsar Sag, Junggar Basin, NW China. Pet Explor Dev. 2025;52(3):830–41. doi:10.1016/S1876-3804(25)60606-9.
36. Fernandes FJD, Teixeira L, Freire AFM, Lupinacci WM. Stochastic seismic inversion and Bayesian facies classification applied to porosity modeling and igneous rock identification. Pet Sci. 2024;21(2):918–35. doi:10.1016/j.petsci.2023.11.020.
37. Nobakht M, Clarkson CR. Analysis of production data in shale gas reservoirs: rigorous corrections for fluid and flow properties. J Nat Gas Sci Eng. 2012;8(5):85–98. doi:10.1016/j.jngse.2012.02.002.