An Innovative Method for Hydraulic Fracturing Parameters Optimization to Enhance Production in Tight Oil Reservoirs

Shide Yan, **Weirong Li***, **Min Zhang**, Xi'an Shiyou University, Xi'an, China; **Zhengbo Wang**, **Zhaoxia Liu**, Research Institute of Petroleum Exploration and Development, PetroChina, Beijing, China; **Keze Lin**, China University of Petroleum (Beijing), Beijing, China; **Hongliang Yi**, Petrochina Liaohe Oilfield Company, Panjin, China

Abstract

The hydraulic fracturing technology for horizontal wells is one of the key techniques for the effective development of tight oil and gas reservoirs. Optimizing fracturing parameters can significantly enhance fracturing effectiveness, reduce development risks, and improve oil and gas production as well as economic efficiency. Rapid and accurate optimization of hydraulic fracturing construction parameters for tight oil horizontal wells has always been a challenge in reservoir development and management. This study introduces a novel workflow for optimizing fracturing parameters by combining reservoir numerical simulation and machine learning techniques. The paper first establishes a single-well numerical model using commercial simulator, and calibrates the reservoir model through matching historical production data. Eight main parameters are selected, and an initial feature dataset is generated using Monte Carlo method, while production dataset is obtained from reservoir numerical simulation. Subsequently, various machine learning algorithms are employed to construct fracture productivity models under different combinations of reservoir and hydraulic fracture parameters. The selected machine learning model with best performance is then integrated with an economic evaluation model to establish an optimization model for hydraulic fracturing parameters optimization for tight oil horizontal wells. The research indicates that the production prediction model established based on the CNN-LSTM method exhibits a high level of accuracy. The optimization model for hydraulic fracturing parameters in tight oil horizontal wells can rapidly optimize fracturing parameters. The proposed methodology in the paper has the potential to enhance horizontal well production and improve economic benefits in tight oil horizontal wells, and can also be applied to similar field development and engineering parameter optimization scenarios.

Introduction

Currently, crude oil and natural gas play crucial roles in the global energy landscape, providing abundant energy and resources for people's production and daily lives. Simultaneously, the extraction of conventional oil and gas resources is encountering increasingly formidable challenges worldwide, prompting unconventional oil and gas resources to emerge as a significant avenue for the development of the oil and gas industry (Zou et al. 2015). Unconventional oil and gas resources primarily encompass tight oil, oil sands, shale gas, and natural gas hydrates. Their extraction poses significant challenges and comes with high costs; however, they are characterized by substantial development potential and abundant resources (Rezaee 2022). Regarding tight oil, it is commonly co-produced alongside shale gas. Led by the United States, effective development of tight oil has also been achieved in Canada and Argentina, with production in 2020 reaching 25 million tons and 5.2 million tons, respectively, while the total production of tight oil and shale oil in the United States reached 350

DOI: 10.14800/IOGR.1262

Received June 2, 2023; revised September 14, 2023; accepted November 14, 2023.
*Corresponding author: weirong.li@xsyu.edu.cn

million tons (Kelsey et al. 2016). Substantial remaining oil in tight reservoirs remains untapped, prompting the utilization of fracturing techniques after depletion-based extraction to boost output and extend production cycles (Todd and Evans 2016). Given the high drilling costs and potential environmental issues associated with hydraulic fracturing, researches aimed at enhancing recovery in tight oil reservoirs have become exceptionally significant.

The hydraulic fracturing technology for horizontal wells is an advanced technique used to enhance oil and gas recovery rates. Its fundamental principle involves injecting fracturing fluid into a horizontal well under high pressure, creating fractures within rock fissures. This process enhances reservoir permeability and effective porosity, consequently increasing the recovery rate of oil and gas (Wu et al. 2012). As a result, optimizing fracturing parameters is crucial for successful hydraulic fracturing. The methods for optimizing fracturing parameters encompass traditional empirical formulas, physical simulations, statistical approaches, and machine learning techniques. These methods aim to enhance the productivity and economic efficiency of horizontal wells by optimizing parameters, such as fracturing fluid concentration, viscosity, and injection pressure.The following outlines the evolution of models and optimization of fracturing parameters: Cleary (1980) utilized experimental data and mathematical models to establish a set of design formulas for predicting fracture pressure and length during hydraulic fracturing. These formulas aimed to optimize hydraulic fracturing design, enhancing efficiency and recovery. Yang et al. (1996) introduced a method employing multivariate optimization techniques for hydraulic fracturing design. By constructing a comprehensive hydraulic fracturing model, this approach incorporated various factors, such as hydraulic fracturing parameters and reservoir properties, and conducted comprehensive parameter optimization to achieve optimal fracturing outcomes. Elrafie and Wattenbarger (1997) employed computational fluid dynamics simulations to model the hydraulic fracturing process. Sensitivity analyses were performed on fracturing parameters and well spacing. By contrasting production from horizontal and vertical wells, recommendations for optimal horizontal well and fracturing designs for the reservoir were proposed.

Dahaghi (2010) employed numerical simulation methods to analyze gas recovery and carbon dioxide sequestration processes. Different parameters' impacts on reservoir pressure, pore pressure, and saturation were explored. It was found that utilizing logarithmically spaced locally refined grids accurately simulated the volume fractured region of horizontal wells during hydraulic fracturing simulations. Cipolla et al. (2010) introduced an analysis method for reservoir properties and productivity characteristics, along with a fluid dynamics model. This enabled the holistic modeling of complex fracture networks in tight reservoirs, validated through microseismic monitoring results. Zhou et al. (2014) employed data mining techniques to assess production performance in the Marcellus shale gas region. They applied classification and regression algorithms from machine learning to process and model the data, subsequently validating their models. The study demonstrated that data-driven methods effectively predicted shale gas well production performance, offering valuable insights for optimizing production control. Schuetter et al. (2018) utilized data analysis methods to construct predictive models for unconventional shale oil and gas reservoir production. They employed multivariate linear regression and k-nearest neighbors algorithms, evaluating and optimizing their models. The results indicated that both methods were effective for predicting reservoir production in shale oil and gas formations, exhibiting high predictive accuracy across different datasets. Luo et al. (2019) utilized three machine learning methods (neural networks, decision trees, and support vector machines) to perform extensive data analysis on Bakken shale oil horizontal wells. Leveraging historical production data and multiple influencing factors such as geological attributes, fracturing parameters, and production strategies, they established predictive models. The outcomes demonstrated that machine learning models accurately forecasted well production performance and offered recommendations for production optimization. Duplyakov et al. (2020) employed machine learning techniques for optimizing hydraulic fracturing design using field data. They developed a digital database and employed various machine learning algorithms to analyze and model field data, predicting optimal fracturing design parameters and production enhancement. The effectiveness of the model was validated through field experiments, demonstrating the potential of utilizing machine learning methods for optimizing hydraulic fracturing design. Dong et al. (2022) addressed issues with traditional trial-and-error-based hydraulic fracturing parameter optimization methods by introducing a hybrid optimization approach that combines machine learning and evolutionary algorithms. This method utilized machine learning algorithms to construct a hydraulic fracturing model based on experimental data. Subsequently, evolutionary algorithms were employed to optimize critical parameters within the model, ultimately yielding the optimal combination of fracturing parameters.

However, optimizing fracturing parameters in tight oil reservoirs presents several challenges and difficulties, such as the lack of precise physical models and limited data. With the rapid advancement of data science in recent years, big data analytics methods have found extensive application in the field of oil and gas exploration and development (Zhan et al. 2019; Li et al. 2022). Simultaneously, machine learning-driven optimization of hydraulic fracturing parameters in horizontal wells necessitates substantial well group data, but this also gives rise to issues such as large data volumes and high costs. To address these challenges, this research aims to introduce a novel process for optimizing fracturing parameters. Specifically, it involves the utilization of reservoir numerical simulation and machine learning techniques to optimize hydraulic fracturing parameters in horizontal wells. By combining numerical simulation with machine learning, it becomes possible to achieve an optimized prediction of fracture morphology, production capacity, and fracturing parameters in horizontal well hydraulic fracturing. Such a comprehensive approach harnesses the strengths of both methods, establishing predictive models from extensive experimental data to further enhance the efficiency and precision of hydraulic fracturing.

The following is a detailed explanation of the structure of the article. Section 2 presents the machine learning methods employed for parameter optimization, encompassing modeling and prediction of sequential data, along with the requisites for data and model establishment, including dataset generation. The optimal machine learning model is identified in Section 3, followed by a single-factor sensitivity analysis of fracturing parameters. The practical optimization of fracturing parameters outlined in Section 4. Section 5 encompasses discussion and future prospects of this study. The conclusion is presented in Section 6.

Methodology and Workflow

The establishment of a machine learning production prediction model requires robust data support, which can be achieved through the integration of datasets from reservoir numerical simulation, fracturing simulation, and historical production fitting. However, in cases where high-quality real data is scarce or unavailable, synthetic data obtained from numerical or analytical simulations can be utilized (Kulga et al. 2017). This subsection introduces geological description of study area, tuning reservoir parameters through historical production matching, and employing reservoir numerical simulation to generate production datasets. And then, it introduces three machine learning algorithms used to forecast production in this study.

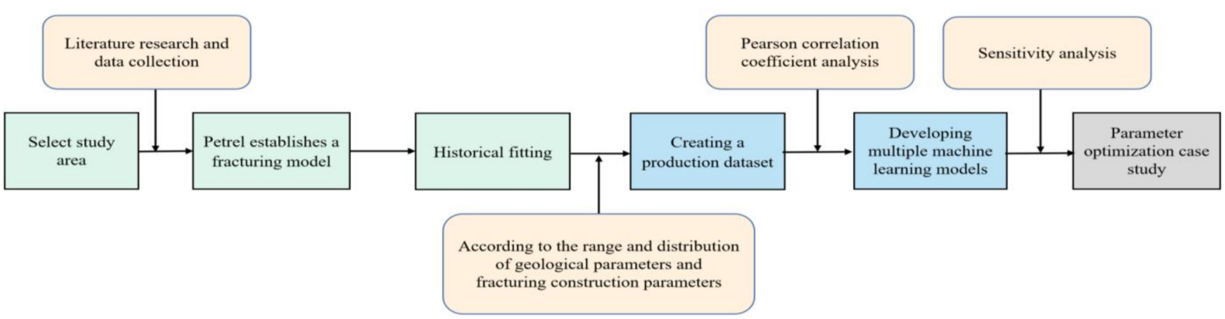


Figure 1—Structural flowchart for fracturing parameter optimization.

Figure 1 illustrates the workflow for the design of fracturing parameter optimization. It begins with the establishment of a geological model for the study area, followed by historical fitting through hydraulic fracturing. Subsequently, a machine learning-required dataset is generated based on the range and distribution of geological parameters and fracturing parameters. Three machine learning models are established and compared. Finally, the best production prediction machine learning model is selected and further integrated with an economic model to establish the fracturing parameter optimization model, which is then subjected to optimization case studies.

Research Area. The Yanchang formation, consisting of seven segments, is an important reservoir in the Pankou area of Ordos Basin. It is distributed in the Yan'an region of Shaanxi province and Shizuishan region of Ningxia, at the border between Shaanxi and Inner Mongolia. It is considered as one of the key areas for oil and gas exploration and development in this region. As shown in **Figure 2**, the study area is located in the secondary structural unit of the Yishan slope, in the southwestern part of the Ordos Basin. This region is an important reservoir for tight oil and shale gas production, with abundant potential oil and gas resources.

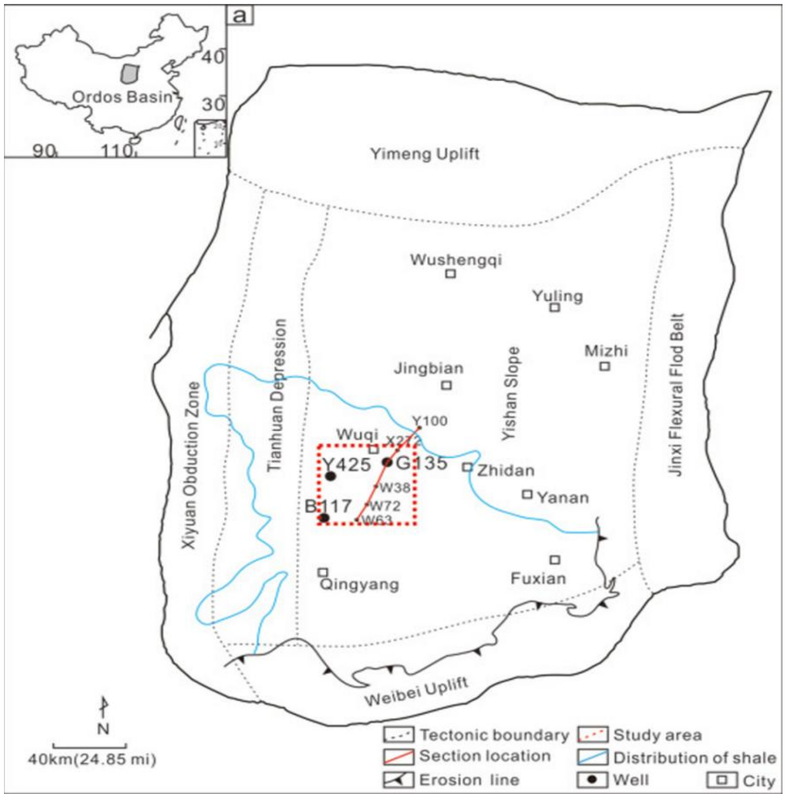


Figure 2—Simplified geological map of the Chang 7 formation in the Ordos Basin. (a) depicts the tectonic framework of the Ordos Basin (data source: Hou et al. 2023).

In the research field, the Chang 7 section of the Yanchang Formation is a crucial reservoir unit known for its abundant tight oil resources and high-quality characteristics. However, an analysis of the core physical properties data in the study area reveals that the overall physical properties of this reservoir are poor (Wang et al. 2015). The porosity distribution of the sandstone ranges from 3.07% to 18.75%, with an average value of 10.77%. The permeability distribution ranges from 0.03 to 3.23 mD, with an average value of 0.18 mD. This reservoir is characterized by low porosity and extremely low permeability, and the presence of microfractures makes it the primary pathway for oil and gas migration (Xiao et al. 2017). As a result, fractures and microfractures are the main locations for oil and gas accumulation in this reservoir. Additionally, the reservoir in the study area exhibits strong heterogeneity, with significant variations in the effectiveness of hydraulic fracturing.

Dataset Generation. Before establishing a machine learning production forecasting model, the primary task is to build a single well geological model, as it forms the basis for optimizing fracturing parameter design. In the numerical model, it is necessary to input the geological information and reservoir properties of the target well group, as well as the fracturing construction parameters, and set appropriate boundary conditions and numerical methods. By solving the model equations, it is possible to predict the fracturing effect and productivity of the target well group, which is of crucial significance for optimizing fracturing parameters.

Establishment of Numerical Model. Based on the geological model of the seven sections of the Guping well group leader, a geological model of multi-stage fractured horizontal well (MFHW) was established using the commercial simulator (Petrel). The reservoir petrophysical parameters, hydraulic fracturing parameters, and boundary conditions were determined. Multiple data sources, including seismic data, well logging data, and core data were imported into the simulator for interpretation and analysis, resulting in an accurate three-dimensional reservoir model which provides a reliable basis for optimizing fracturing parameters and predicting production rates. **Table 1** presents the reservoir and hydraulic fracturing parameters for the MFHW. By considering these reservoir and hydraulic fracturing parameters, an accurate geological model for the horizontal well was constructed to optimize fracturing design and predict production. **Figures 3** illustrate the schematic diagram of the reservoir model established. This model was constructed based on a comprehensive evaluation of various geological parameters and hydraulic fracturing parameters, aiding in the optimization of fracturing design and production rate prediction.

Table 1—Geological parameters and hydraulic fracture parameters for horizontal wells.

Parameter	Value
X length (m)	2000
Y length (m)	500
Z length (m)	18
Reservoir temperature (℃)	80
Reservoir length (m)	1200
Formation pressure (Mpa)	20
Reservoir thickness (m)	20
Porosity (%)	11.17
Permeability (mD)	0.14
Oil saturation (%)	51.53

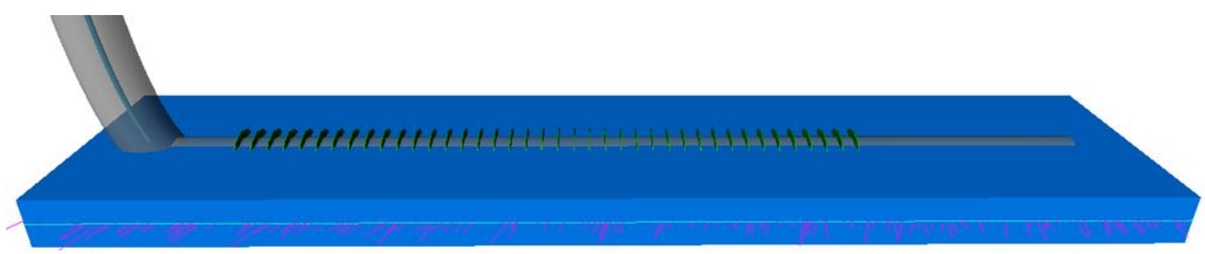


Figure 3—Conceptual illustration of a single well geological model for petrel horizontal wells.

After establishing the numerical model, the hydraulic fracture network simulation of a typical well group was conducted using the fracturing simulation software Petrel-Kinetix. During the fracturing operation, the fluid intensity and proppant intensity are two parameters that need to be balanced. Higher fluid intensity may require higher pumping pressure to extend the fractures, while higher proppant intensity can provide better fracture support but may also increase the demand for fluid pumping. Therefore, when selecting the parameters, it is necessary to balance the fluid intensity and proppant intensity according to specific requirements in order to achieve the optimal fracturing effect. The specific selected parameters are shown in **Table 2**.

Table 2—Liquid strength and Sand strength.

Parameter	Value
Fracture length (m)	283.1
Fracture height (m)	11.91
Fracture permeability (mD)	102.3
Sand proportion (%)	18.2
Single well injected fluid volume (m^3)	30477
Single-stage fluid volume (m^3)	1270
Single well sand volume (m^3)	3398.5
Single-stage sand volume (m^3)	141.6
Single-stage displacement (m^3/min)	9
Fracture spacing (m)	58.2

The natural fracture characteristics of reservoirs primarily encompass two key aspects: the storage capacity and the fluid flow properties. Taking the natural fractures in the Long 7 formation of the Ordos Basin as an example, specific data is given in **Table 3**.

Table 3—Natural fracture characteristics of Long 7 formation.

Well number	Well #18-#30
Fracture direction (°)	Formation strike: N60°W, Formation dip: 30 degrees
Fracture length (μm)	1200
Fracture width (μm)	50
Fracture description	Asphalt filled
Porosity (%)	0.014

Figure 4 illustrates the hydraulic fracturing effect of a typical well. The blue lines represent the positions of natural fractures, while the hydraulic fracturing network is depicted by varying shades of color indicating the width of the fractures, with darker shades representing wider fractures. By comparing these fractures, it is possible to assess the effectiveness of hydraulic fracturing and determine whether it has successfully expanded the fracture network, thereby increasing the permeability and productivity of the reservoir. Additionally, the overlap between the natural fractures and the hydraulic fracturing network can be analyzed to gain further insights into the coverage and effectiveness of hydraulic fracturing.

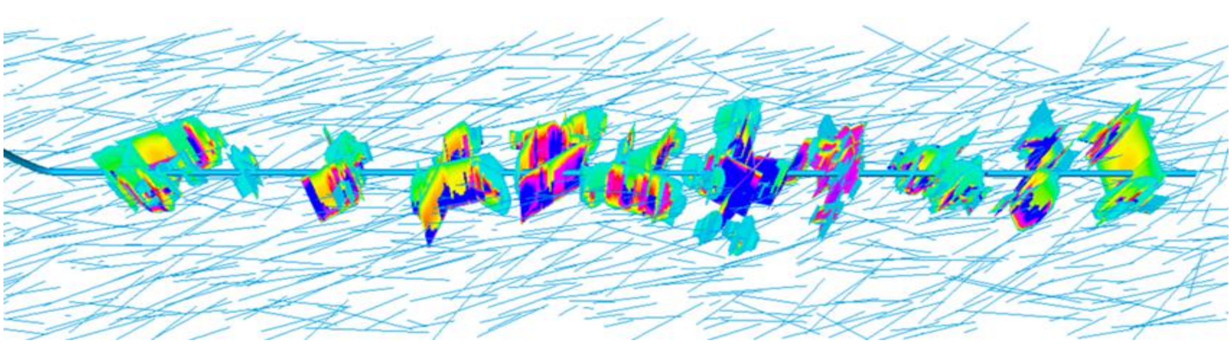


Figure 4—Hydraulic fracturing network.

Production History Match. Production history match refers to the process of comparing historical production data with numerically simulated production data in order to validate and adjust the accuracy of the numerical simulation model (Zhang and Awotunde 2016). **Figure 5** illustrates the basic workflow of history matching. The parameters are only output when the difference between the fitted values and the actual values is smaller than a specific threshold value.

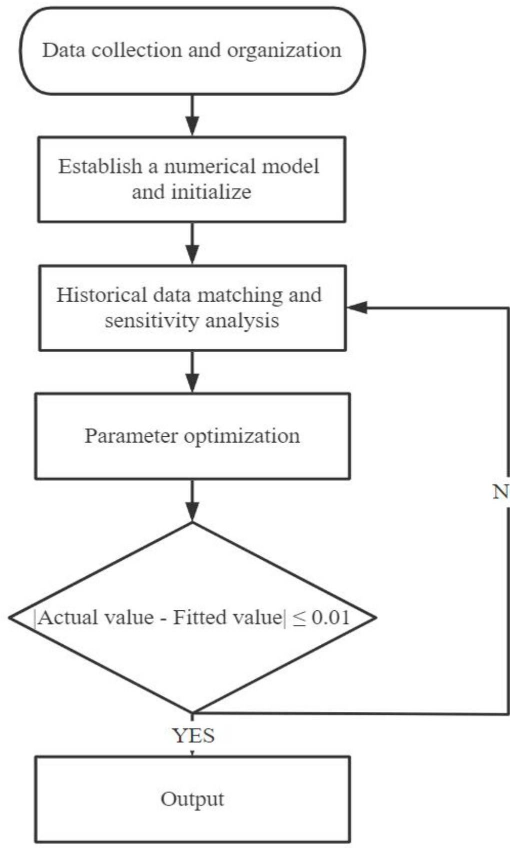


Figure 5—Workflow of production history match.

Table 4 presents the initial values and history matching results of the basic fluid and reservoir parameters for the established numerical model of this study.

Table 4—Initial and final values of parameters.

Parameter	Initial value	Final value
Permeability (mD)	0.14	0.05
Porosity (%)	11.17	11
Oil saturation (%)	51.53	55
Fracture permeability(mD)	112.6	102.3
Natural fracture permeability(mD)	0.65	0.42

Generation of Production Dataset. According to the actual conditions of the Yanchang Formation reservoir in the Pankou area of Ordos Basin, the range of various geological factors of the reservoir were determined through analysis and statistical analysis of field data, including core analysis and well logging, combined with comparison and validation using numerical models. The range for each reservoir parameter in the study area are presented in **Table 5**. **Figure 6** illustrates the distribution of reservoir parameters in typical well groups within the study area.

Table 5—Range and boundary of geological parameters.

Parameter	Minimum	Maximum
Length of reservoir section (m)	400	1600
Porosity (%)	9	12
Permeability (mD)	0.01	0.25
Oil saturation (%)	45	70

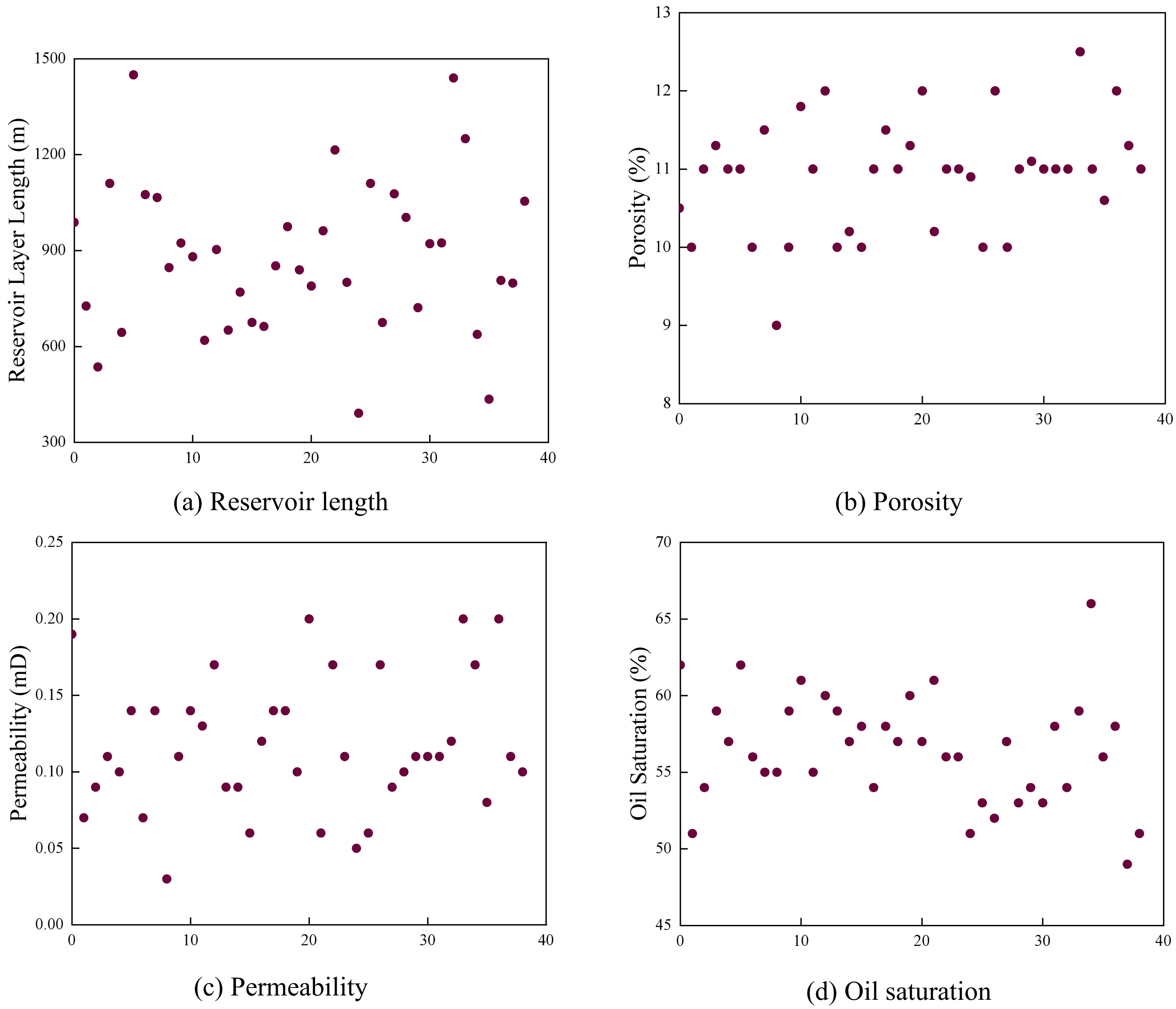
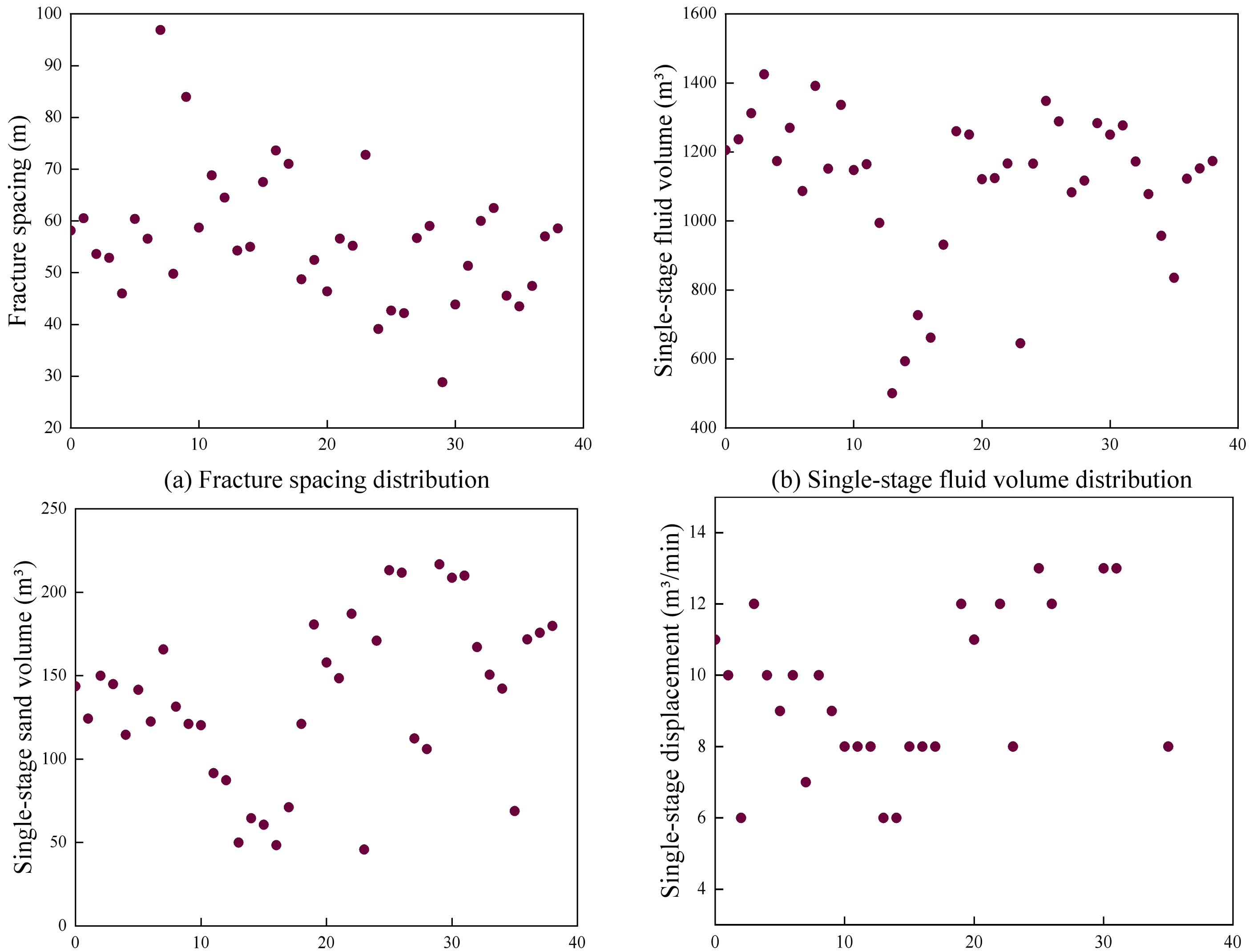


Figure 6—Distribution of reservoir parameters.

By collecting fracturing data from typical production wells in the target area and comparing them with fracturing data from similar wells with similar geological structures, well types, well depths, lithologies, etc., the range of fracturing parameters can be preliminarily determined. Additionally, referencing existing fracturing experiments and empirical data also helps to establish the range of these parameters. The range for each fracturing parameter are presented in **Table 6**. **Figure 7** presents the distribution of fracturing parameters of typical well groups in the study area. It is aware that adjustments and optimizations of the parameters should be made based on the specific conditions onsite to ensure the effectiveness.



(a) Fracture spacing distribution
(b) Single-stage fluid volume distribution
(c) Single-stage sand volume distribution
(d) Single-stage displacement distribution

Figure 7—Distribution of fracturing parameters.

Table 6—Ranges and boundary of fracturing parameters.

Parameter	Minimum	Maximum
Fracturing spacing (m)	30	100
Single-stage fluid volume (m^3)	500	1500
Single-stage sand volume (m^3)	50	250
Single-stage displacement (m^3/min)	5	15

Numerical simulations are run to generate production dataset based on the combination of reservoir and fracturing parameters in the above ranges. **Table 7** presents the distribution characteristics of the production dataset, while **Figure 8** shows the distribution of cumulative oil production. A total of 2698 sets of dynamic production data were obtained, including reservoir parameters, such as reservoir thickness, porosity, permeability, and oil saturation, as well as hydraulic fracturing parameters, such as fracture spacing, single-stage fluid volume, single-stage sand volume, and single-stage displacement. The dataset also includes monthly oil production over ten years.

Table 7—Characteristics of production dataset generated by numerical simulation.

	Interval length (m)	Porosity (%)	Perm. (mD)	Saturation (%)	Fracture spacing (m)	Fluid volume (m^3)	Sand volume (m^3)	Displacement (m^3/min)	Cum.oil production (t)
Sum	2698	2698	2698	2698	2698	2698	2698	2698	2698
Average	1244	10.99	0.15	55.0	64	1025	161	10	28033
Stand. Dev.	704	1.16	0.08	6.6	26	739	86	3	4524
Minimum	400	9.00	0.01	45.0	30	500	30	5	3156
25%	750	9.75	0.10	52.0	40	750	50	7	8024
50%	1000	10.82	0.15	57.5	60	1000	125	10	15816
75%	1250	11.35	0.20	63.0	80	1250	200	13	23578
Maximum	1600	12.00	0.25	70.0	100	1500	250	15	35916

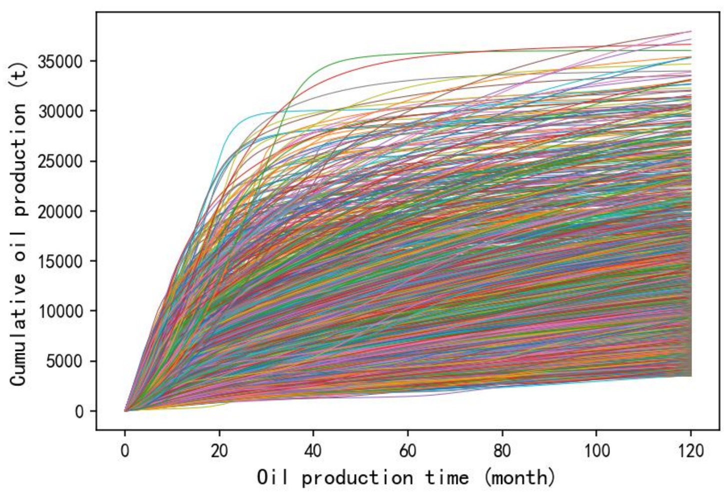


Figure 8—Distribution of cumulative oil production in the production dataset.

Machine Learning Algorithms for Production Forecasting. Machine learning refers to the process of automatically adjusting the parameters of algorithm models by learning patterns and rules from a large amount of data, with the aim of improving the accuracy of prediction and classification. The basic principle of neural networks involves training the model to map input data to output data, establishing a relationship between the two. Machine learning methods can be broadly categorized into supervised learning and unsupervised learning . Supervised learning involves training the model using known input and output data samples to predict the output for new unknown data. On the other hand, unsupervised learning aims to discover the structure and patterns within the data itself through analysis and learning without any given output samples.The following briefly introduces these three machine learning methods used for production forecasting in this study.

CNN Method. Convolutional Neural Network (CNN) is a deep learning technique used for analyzing data with grid-like structures, such as images, speech, and text. Compared to traditional neural networks, CNN can automatically learn features from input data while reducing the number of parameters, thus improving the efficiency and accuracy of the model (Albawi et al. 2017). **Figure 9** depicts the structure of a one-dimensional convolution. In CNN, the convolutional layers extract features from input images, while the pooling layers are used to reduce the size and number of parameters in the feature maps, preventing overfitting. The fully connected layers combine the output features from the convolutional and pooling layers to perform classification or regression tasks.

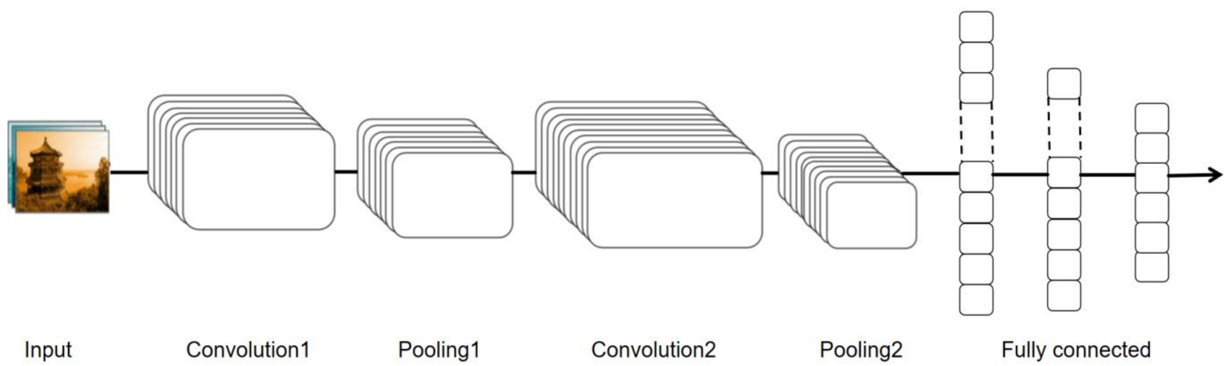


Figure 9—CNN one-dimensional convolutional structure diagram.

LSTM Method. Long Short-Term Memory (LSTM), a type of recurrent neural network (RNN) model, is designed for handling sequence data. As shown in **Figure 10**, LSTM utilizes three gates to control the flow of information: the forget gate, input gate, and output gate. These gates allow LSTM to selectively retain or forget information and produce predictions at the current time step. As a result, LSTM has been widely applied in sequence data processing tasks such as natural language processing, speech recognition, and stock prediction.

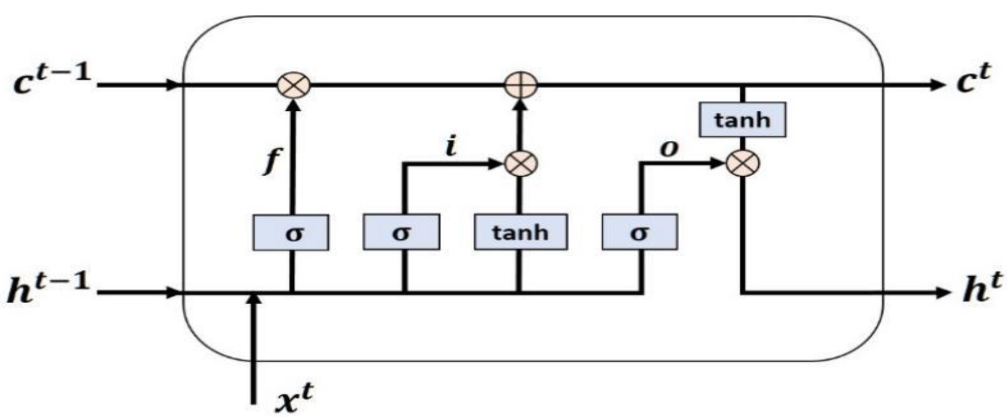


Figure 10—LSTM neural network unit structure diagram.

CNN-LSTM Method. CNN-LSTM is a complex neural network architecture that combines the characteristics of CNN and LSTM, making it suitable for modeling and predicting sequential data. As shown in **Figure 11**, it first processes the sequential data through convolutional layers to extract spatial features. Then, the output of the convolutional layers is fed into the LSTM layer, which learns the temporal dependencies in the sequence through its memory cells and gate units. Lastly, the output of the LSTM layer is passed to the output layer for prediction.

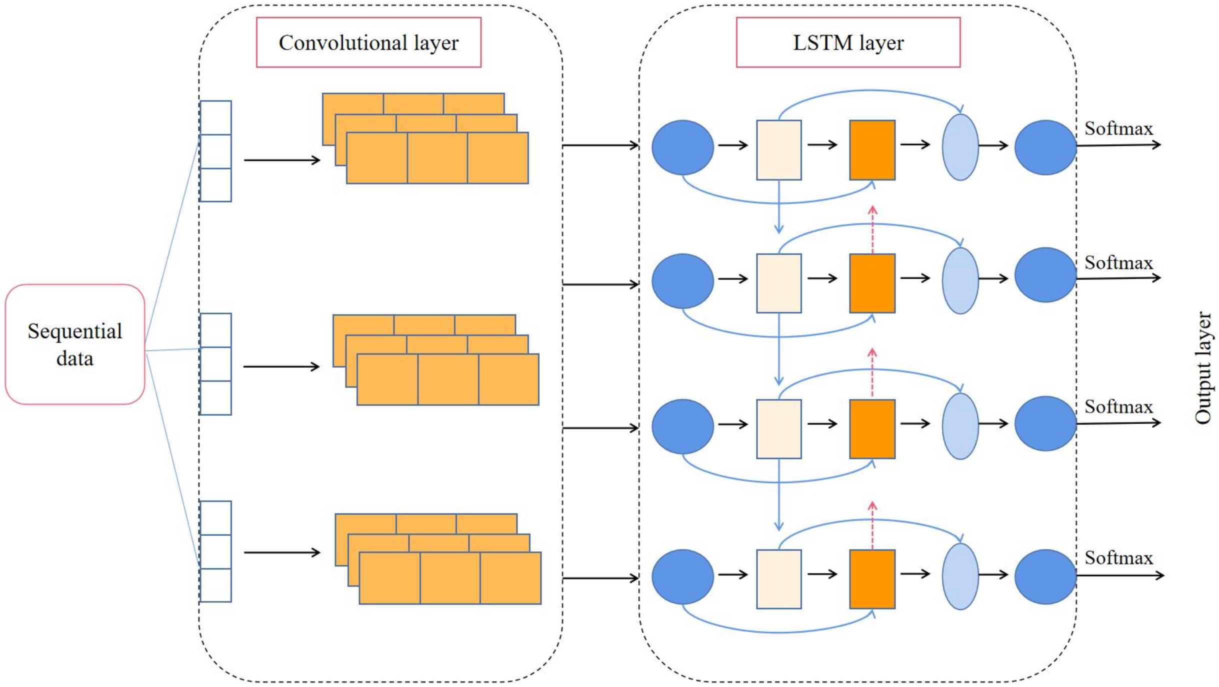


Figure 11— CNN-LSTM neural network unit structure diagram.

Analysis of Model Training Results

This section evaluates the predictive accuracy of different machine learning models through correlation analysis, selects the optimal one, followed by conducting a univariate sensitivity analysis of hydraulic fracturing parameters. Finally, it introduces the case study of fracturing parameters optimization through instance-specific investigations.

Correlation analysis. Correlation analysis involves the application of statistical methods to assess the degree of association between two or more variables. In data analysis and modeling, the utilization of Pearson correlation coefficient analysis is employed to investigate the interrelationships among variables, determining their connections, and deciding whether to incorporate these variables within the model. This yields valuable guidance for feature selection and model refinement.

The Pearson correlation coefficient is a statistical measure used to gauge the extent of linear association between two variables. Typically denoted by the symbol r, the correlation coefficient's values range from -1 to 1. A r value of 0 signifies the absence of a linear relationship between the two variables. A r value of -1 indicates a perfect negative correlation, while a r value of 1 signifies a perfect positive correlation between the variables.

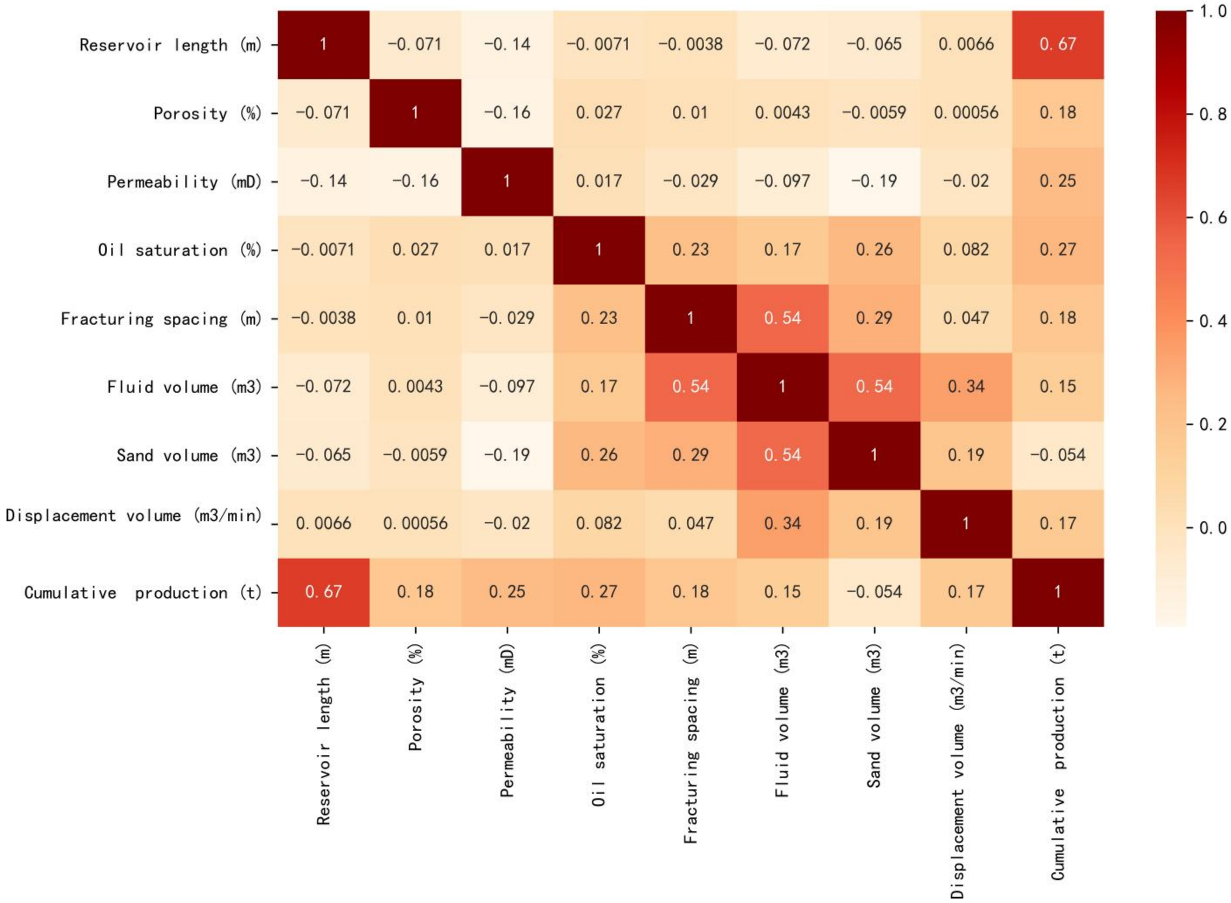


Figure 12—Pearson correlation coefficient heatmap.

Figure 12 displays the Pearson correlation coefficient heatmap between input parameters and cumulative oil production. The intensity of colors reflects the degree of correlation between them, accompanied by corresponding correlation coefficient values. Reservoir parameters exhibit notable correlations with cumulative oil production, particularly the highest correlation observed between reservoir thickness and cumulative oil production at 0.67. Subsequently, oil saturation, permeability, and porosity follow with correlations of 0.27, 0.25, and 0.18, respectively. The correlation between hydraulic fracturing parameters and cumulative oil production is relatively weak, with a coefficient of -0.18. Conversely, there is stronger correlation among hydraulic fracturing parameters, notably the highest correlation being between single-stage fluid volume and fracturing spacing , as well as single-stage sand volume, at 0.54. The magnitudes of the correlation coefficients between various features do not induce multicollinearity issues, thus ensuring the model's stability and accuracy.

Machine Learning Optimization Model. We utilized the Particle Swarm Optimization (PSO) algorithm, which is an intelligent optimization technique based on collective cooperation and global exploration, drawing inspiration from the migration and clustering behaviors observed in avian foraging processes. PSO algorithm involves adapting particle values by driving changes through the objective function. This is accomplished by dynamically comparing the optimal positions independently found by individual particles with the optimal position discovered by the entire population. We applied the PSO algorithm to optimize the hyperparameters of the CNN model, and the training outcomes are illustrated in **Figure 13**. The model exhibited a Root Mean Square Error (RMSE) of 1364.99 and a coefficient of determination (R^2) of 0.961 on the training dataset. On the testing dataset, the model achieved an RMSE of 1588.35 and an R^2 of 0.955.

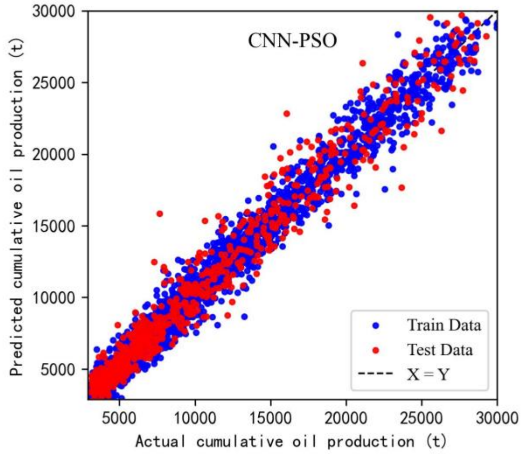


Figure 13—The performance of the CNN model.

Particle swarm optimization (PSO) was used to optimize the parameters of LSTM model, and the accuracy of the model was not improved. The training outcomes post PSO optimization are illustrated in **Figure 14**. Following optimization, the model exhibited a Root Mean Square Error (RMSE) of 1320.53 and a coefficient of determination (R^2) of 0.943 on the training dataset. On the testing dataset, the model achieved an RMSE of 1740.03 and an R^2 of 0.937.

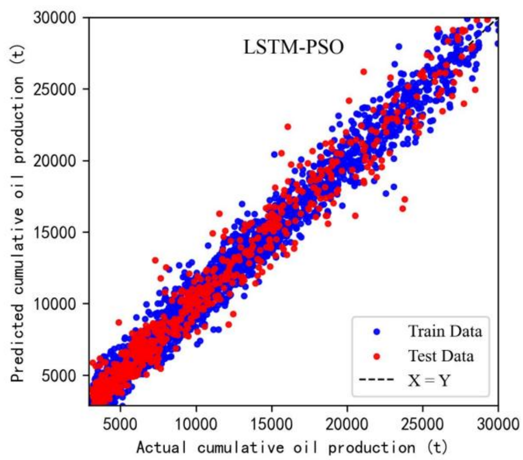


Figure 14—The performance of the LSTM model.

Through iterative optimization with the PSO algorithm, a superior combination of parameters for the CNN-LSTM model was obtained, enhancing its fitting capacity and predictive accuracy. The training outcomes are depicted in **Figure 15**. Ultimately, the optimized model achieved an RMSE of 1286.01 and an R^2 of 0.981 on the training set, and an RMSE of 1393.91 and an R^2 of 0.963 on the testing set. The iterative optimization with the PSO algorithm significantly improved the performance of the CNN-LSTM model, enhancing its precision and predictive ability on both the training and testing datasets. The optimized model is now better equipped to accurately predict the target variable and provide more reliable results.

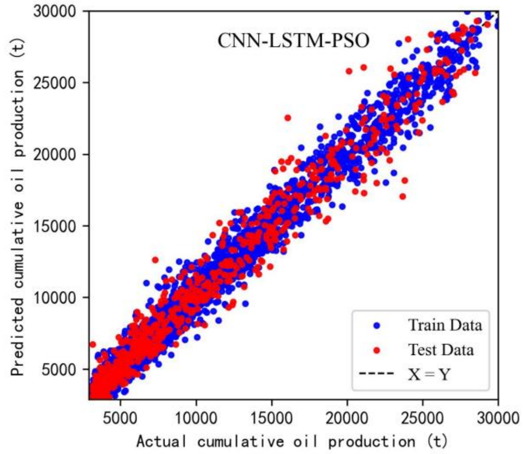


Figure 15—The performance of the CNN-LSTM model.

After optimizing the three models, distinct production prediction models have been obtained. The next step is to assess these models and select the optimal one. Model performance evaluation employs metrics, such as RMSE and R^2. On the testing dataset, a comparison is made among the three machine learning models that have undergone PSO parameter optimization, based on RMSE and R^2. **Figure 16** illustrates the comparison of RMSE and R^2 on the testing dataset for the three machine learning models following PSO parameter optimization. The results demonstrate that the CNN-LSTM-PSO model exhibits the smallest RMSE (1393.91) and the highest R^2 (0.963) on the testing dataset. Consequently, the CNN-LSTM-PSO model is chosen as the optimal production prediction model. It is important to emphasize that this conclusion is specifically applicable to the current dataset and task. If applied to different datasets or tasks, a re-evaluation of model performance is necessary to determine the optimal model choice.

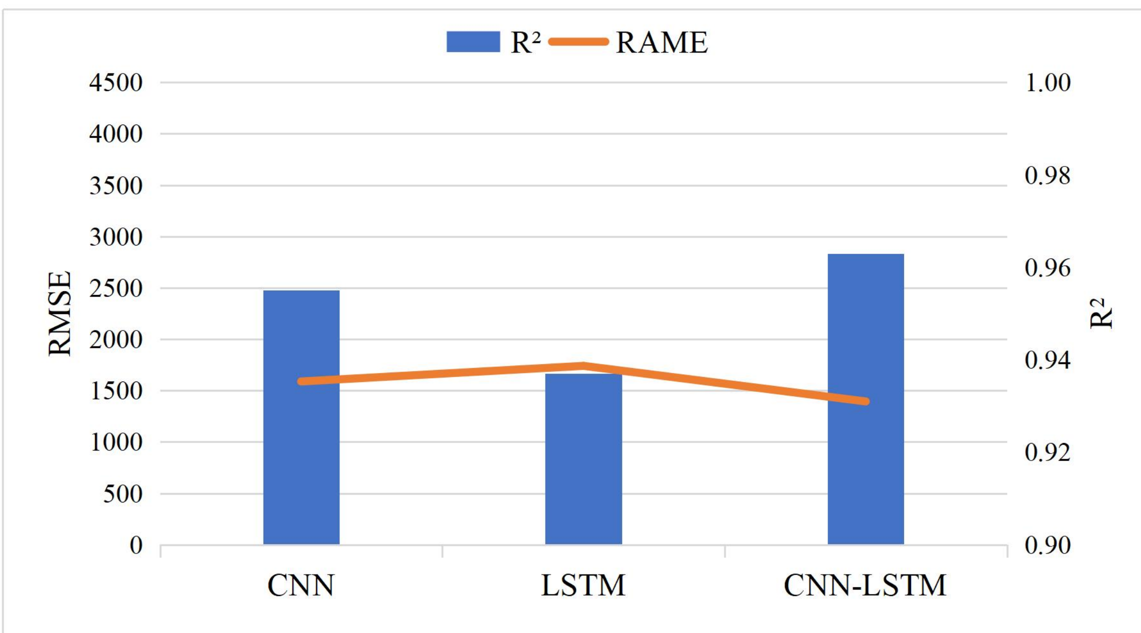


Figure 16—RMSE and R^2 comparison on the testing dataset for different machine learning models after optimization.

Sensitivity Analysis. Sensitivity analysis is a crucial method for evaluating how a model responds to variations in input parameters. It plays a significant role in optimizing model performance and enhancing decision quality. In the context of univariate sensitivity analysis, the values of each parameter are altered individually to observe the corresponding changes in Net Present Value (NPV). This aids in making more accurate decisions by understanding how the model's output reacts to parameter adjustments.

Economic Evaluation. This model comprehensively considers reservoir parameters, hydraulic fracturing parameters, and economic benefits with the aim of maximizing the economic returns of oil wells. It serves as a vital decision-making reference for oilfield development. The model's inputs encompass the reservoir parameters of the target well, such as reservoir thickness, porosity, permeability, and oil saturation. By incorporating optimized hydraulic fracturing parameters, it predicts the production of the target well. These production predictions are then applied in the economic evaluation model to calculate the NPV under the given hydraulic fracturing parameter sets. The model construction process is illustrated in **Figure 17**.

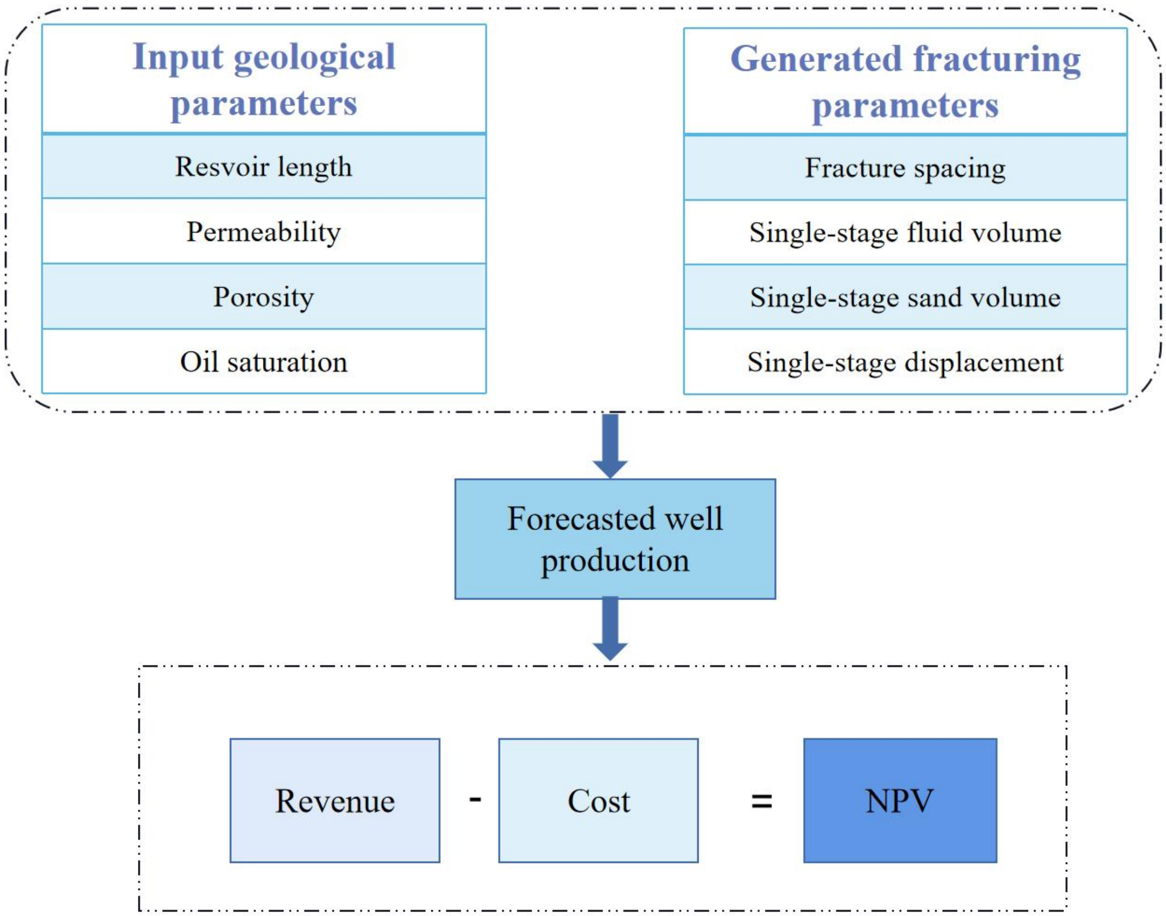


Figure 17—Structure of the economic evaluation model.

Univariate Sensitivity Analysis. During the process of univariate sensitivity analysis, the initial values of reservoir and hydraulic fracturing parameters are established. Reservoir parameters keep the same for all the cases, including reservoir length of 1200 m, porosity of 11%, permeability of 0.05 mD, and oil saturation of 55%. Maintaining other hydraulic fracturing parameters constant, while the fracturing spacing individually various and is set to be 40 m, 55 m, 70 m, 85 m, and 100 m, to conduct the sensitivity analysis and explore the impact of fracturing spacing on NPV. **Figure 18** illustrates the results. It can be observed from Figure 18 that NPV increases with fracturing spacing increasing, but there might be an optimal fracturing spacing where NPV starts to decrease after the certain threshold. In the context of conventional wells, reducing the fracturing spacing could lead to a higher number of fracturing stages, resulting in a significant increase in fluid and proppant volume per well. This, in turn, would escalate fracturing costs. The rise in costs could potentially offset the benefits gained from increased production, leading to a decline in NPV. Therefore, the optimal fracturing spacing may vary across different oilfields and scenarios, necessitating a balanced consideration of costs and benefits.

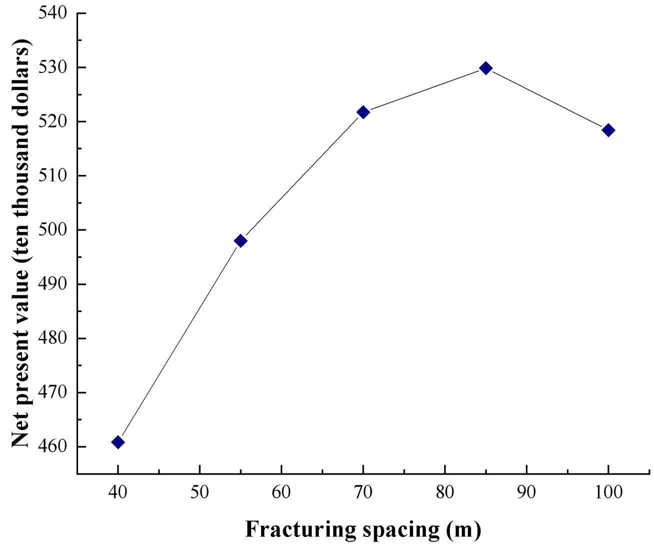


Figure 18—Sensitivity analysis of fracturing spacing on NPV.

Based on the results presented in **Figure 19**, variations in NPV under different single-stage fluid volumes (500m^3, 750m^3, 1000m^3, 1250m^3, and 1500m^3) are evident. From Figure 19, it can be observed that the NPV reaches its peak when the single-stage fluid volume reaches 1000 m^3. Single-stage fluid volume is a critical parameter in hydraulic fracturing, affecting injection rates, fracture propagation, and proppant permeation. Increasing the single-stage fluid volume often leads to higher oil well production. However, it also escalates costs and environmental impact, thereby potentially reducing NPV. Hence, the optimal single-stage fluid volume varies based on distinct well and geological conditions.

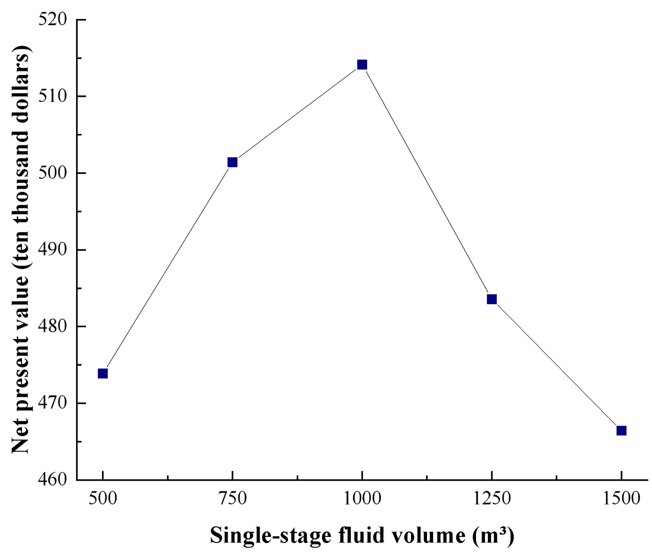


Figure 19—Sensitivity analysis of single-stage fluid volume on NPV.

By examining different single-stage sand volumes (50 m^3, 100 m^3, 150 m^3, 200 m^3, and 250 m^3), we have observed variations in NPV. It is evident from **Figure 20** that the impact of single-stage sand volume on NPV is relatively limited. However, a significant reduction in NPV becomes apparent when the single-stage sand volume is increased to 150 m^3. This implies that raising the single-stage sand volume beyond 150 m^3 may have an adverse effect on NPV.

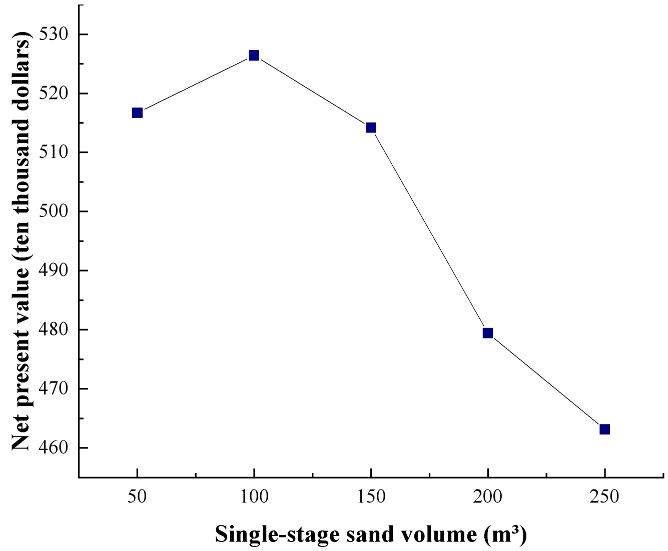


Figure 20—Sensitivity analysis of single-stage sand volume on NPV.

When considering single-stage displacement of 5 m^3/min, 7.5 m^3/min, 10 m^3/min, 12.5 m^3/min, and 15 m^3/min, the variations in NPV were observed. The changing outcomes are presented in **Figure 21**. The study reveals that within a certain range, increasing the single-stage displacement can notably enhance the NPV. Elevating the single-stage displacement leads to a corresponding increase in oil well production, with a relatively minor impact on fracturing costs. Consequently, a positive correlation exists between the single-stage displacement and oil well production. Augmenting the single-stage displacement proves beneficial in improving oil well yields, thereby augmenting the NPV.

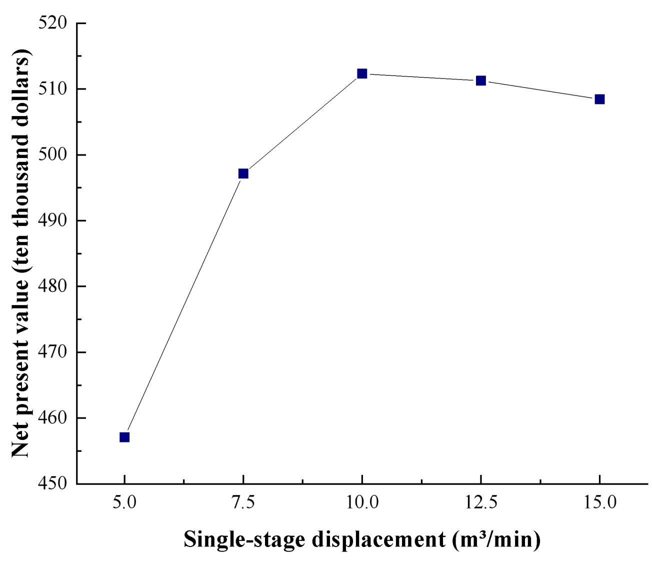


Figure 21—Sensitivity analysis of single-stage displacement on NPV.

Fracturing Parameters Optimization Example

Finally, the fracturing parameters for horizontal wells were optimized using the PSO algorithm, with the objective of maximizing NPV. Following a predefined objective function, the algorithm iteratively searched for the optimal solution. **Figure 22** illustrates the construction process of the NPV fracturing parameters optimization model for MFHWs.

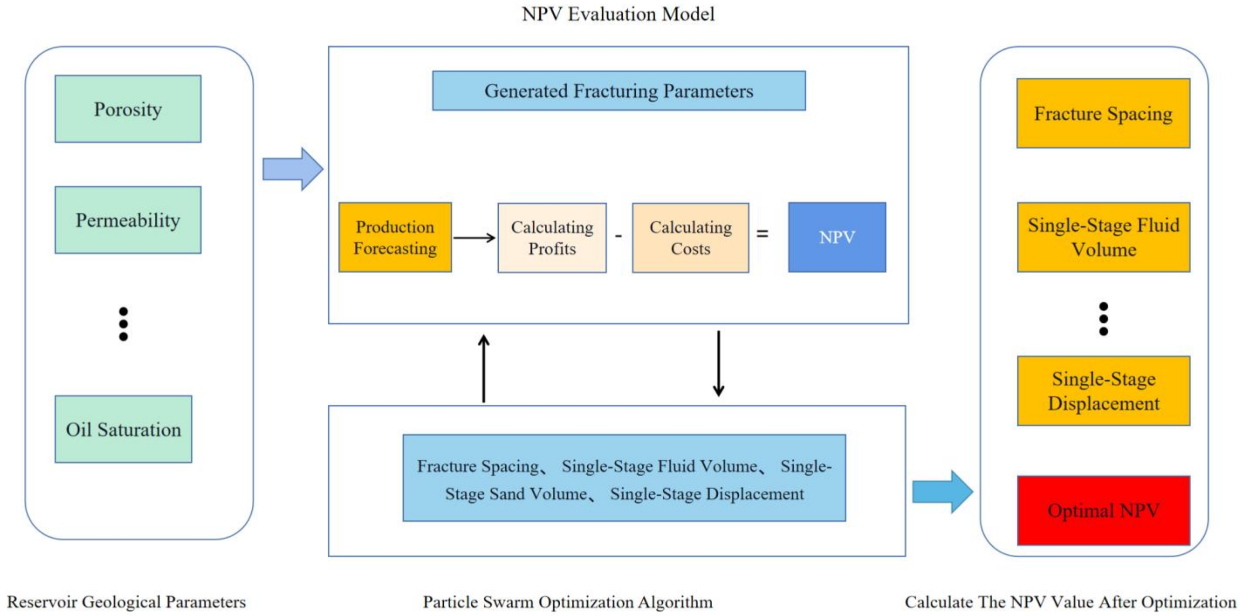


Figure 22—Structure of NPV fracturing parameters optimization model.

By analyzing on-site data, the initial values for geological and fracturing parameters for the experimental well are provided as follows: reservoir length of 1200 m, porosity of 11%, permeability of 0.05 mD, oil saturation of 55%, fracturing spacing of 58.2 m. The PSO algorithm is then utilized for optimizing the fracturing parameters, with the primary objective being to maximize the NPV. Through multiple iterations, the optimal combination of fracturing parameters is determined to achieve the maximization of economic benefits for horizontal wells. **Table 8** presents the initial values, ranges, and the final values after optimization using the PSO algorithm for the fracturing parameters.

Table 8—Fracturing parameters optimization results with NPV as the objective.

Fracturing parameters	Initial value	Range of values	Optimum value
Fracture spacing (m)	58.2	[30,100]	83.05
Single-stage fluid volume (m^3)	1000.0	[500,1500]	1125.00
Single-stage sand volume (m^3)	141.6	[50,250]	91.00
Single-stage displacement (m^3/min)	9.0	[5,15]	11.82

Figure 23 illustrates the iterative process of optimizing fracturing parameters using the PSO algorithm. The horizontal axis represents the iteration number, while the vertical axis represents the NPV. The results indicate that as the number of iterations increases, the range of NPV values gradually converges. Through continuous iterations, the NPV value stabilized at $5.84 million, successfully achieving the economic benefit target. By optimizing the fracturing parameters, it is possible to maximize well production and economic benefits, reduce unnecessary operations and resource wastage, lower costs, and attain the highest economic returns.

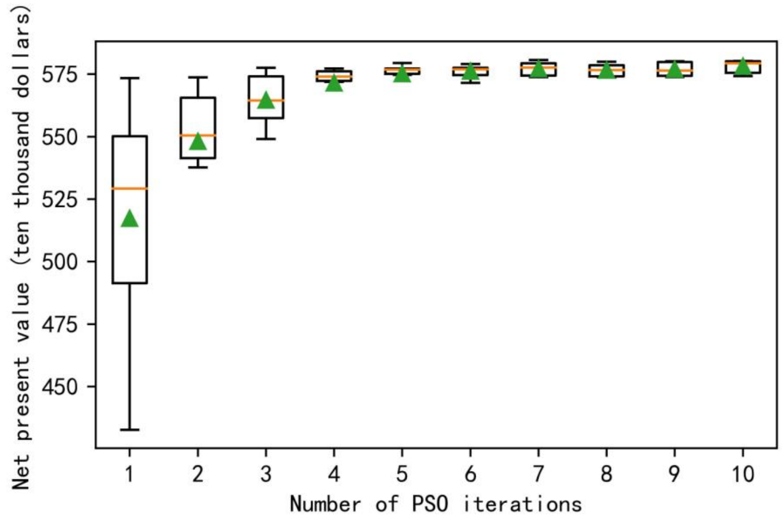


Figure 23—Iterative process of NPV for fracturing parameters optimization using PSO algorithm.

Discussion and Prospects

Fracturing parameter optimization holds significant importance in oil and gas exploration as it can maximize well productivity and economic returns. With the continuous advancement of technology, the utilization of machine learning for production forecasting and fracturing parameter optimization is expected to gain broader applicability. However, through the analysis and discussion of the empirical results, it is evident that there are several areas where the model can be enhanced. Firstly, to better accommodate diverse geological conditions and reservoir characteristics, further research is warranted to explore the response patterns of different types of reservoirs. Additionally, improving the quantity and quality of available data and adopting more sophisticated algorithms and models can enhance the precision of predictions and the efficiency of optimization. Moreover, delving into multi-objective optimization algorithms that consider multiple targets and constraints can lead to more comprehensive optimization strategies. These measures collectively contribute to the advancement of reservoir optimization and oilfield development, addressing the identified limitations in the current model.

In the future, machine learning and deep learning technologies will continue to evolve, and real-time online analysis will become a crucial trend in fracturing parameter optimization. Real-time online analysis will enable timely monitoring and collection of operational data, production data, geological data, etc., which can be directly input into optimization models for analysis and decision-making. This approach allows for immediate feedback on the current reservoir status and performance, assisting engineers in real-time adjustments and optimization of fracturing parameters to adapt to ever-changing oilfield conditions. Furthermore, the future of fracturing parameter optimization will involve a greater consideration of multidisciplinary factors. Knowledge from various disciplines such as geology, geophysics, and rock mechanics will be integrated into optimization models to build more comprehensive and holistic optimization strategies. This interdisciplinary approach will lead to the creation of solutions that take into account a broader range of influences and factors.

In summary, fracturing parameter optimization will leverage advanced technology and data analysis to establish intelligent and efficient models, providing effective support and decision-making for oilfields. The advancements in machine learning, deep learning, and real-time online analysis will enhance the intelligence and automation of optimization models, enabling them to adapt to the actual reservoir conditions and drive the efficient development of oilfields.

Conclusions

This paper integrated numerical simulation and machine learning techniques in the field of oil and gas reservoir development to establish a comprehensive workflow for optimizing hydraulic fracturing parameters. This study obtained a best machine learning-based production prediction model based on the synthetic production dataset generated from numerical simulation. The Particle Swarm Optimization (PSO) algorithm was applied to optimize fracturing parameters by maximizing well productivity and economic returns, minimizing unnecessary operations and resource waste, thus reducing costs and achieving maximum economic benefits. The main conclusions are follows.

1. The CNN-LSTM model was identified as the optimal production prediction model.
2. From Univariate Sensitivity Analysis, it is clear that increasing the single-stage fluid volume and the single-stage displacement can notably enhance the NPV; there is an optimal fracturing spacing corresponding to the highest NPV; while increasing the single-stage sand volume beyond a certain threshold may have an adverse effect on NPV.

References

Albawi, S., Mohammed, T. A., Alzawi, S. 2017. Understanding of a Convolutional Neural Network. Paper presented at the 2017 International Conference on Engineering and Technology (ICET), 21-23 August. 17615756.

Cipolla, C. L., Lolon, E. P., Erdle, J. C., et al. 2010. Reservoir Modeling in Shale-Gas Reservoirs. *SPE Reservoir Evaluation & Engineering* **13** (4): 638-653.

Cleary, M. P. 1980. Comprehensive Design Formulae for Hydraulic Fracturing. Paper presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, 21-24 September. SPE-9259-MS.

Dong, Z., Wu, L., Wang, L., et al. 2022. Optimization of Fracturing Parameters with Machine-Learning and Evolutionary Algorithm Methods. *Energies* **15** (16): 6063.

Duplyakov, V., Morozov, A., Popkov, D., et al. 2020. Practical Aspects of Hydraulic Fracturing Design Optimization Using Machine Learning on Field Data: Digital Database, Algorithms and Planning the Field Tests. Paper presented at the SPE Symposium: Hydraulic Fracturing in Russia. Experience and Prospects, Virtual, 11-13 January. SPE-205310-MS.

ElRafie, E. A. and Wattenbarger, R. A. 1997. Comprehensive Evaluation of Horizontal Wells with Transverse Hydraulic Fractures in the Upper Bahariyia Reservoirs. Paper presented at the Middle East Oil Show and Conference, Bahrain, 15-18 March. SPE-37759-MS

Hou, L., Pang, Z., Luo, X., et al. 2023. Sedimentary Features of the Gravity Flow Influenced Shale Deposit: A Case Study of Chang 7 Member in the Ordos Basin, China. *Marine and Petroleum Geology* **153** (6): 106277.

Dahaghi, A, K. 2010. Numerical Simulation and Modeling of Enhanced Gas Recovery and CO_2 Sequestration in Shale Gas Reservoirs: A Feasibility Study. Paper presented at the SPE International Conference on CO_2 Capture, Storage, and Utilization, New Orleans, Louisiana, USA, 10-12 November. SPE-139701-MS.

Kelsey, T.W., Partridge, M.D., and White, N. E. 2016. Unconventional Gas and Oil Development in the United States: Economic Experience and Policy Issues. *Applied Economic Perspectives and Policy* **38**(2): 191-214.

Kulga, B., E. Artun, and Ertekin, T. 2017. Development of a Data-Driven Forecasting Tool for Hydraulically Fractured, Horizontal Wells in Tight-Gas Sands. *Computers & Geosciences* **103** (6): 99-110.

Li, W., Wang, L., Dong, Z., et al. 2022. Reservoir Production Prediction with Optimized Artificial Neural Network and Time Series Approaches. *Journal of Petroleum Science and Engineering* **215**: 110-586.

Luo, G., Yao, T., Mariia, B., et al. 2019. Production-Strategy Insights Using Machine Learning: Application for Bakken Shale. Paper presented at the SPE Unconventional Resources Technology Conference, Houston, 23-25 July. SPE-195681-PA.

Rezaee, R. 2022. Editorial on Special Issues of Development of Unconventional Reservoirs. *Energies* **15**(7): 2617.

Schuetter, J., Srikanta, M., Zhong, M., et al. 2018. A Data-Analytics Tutorial: Building Predictive Models for Oil Production in an Unconventional Shale Reservoir. *SPE Journal* **23**(4): 1075-1089. SPE-189969-PA

Todd, H. B. and Evans, J. G. 2016. Improved Oil Recovery IOR Pilot Projects in the Bakken Formation. Paper presented at the SPE Low Perm Symposium, Denver, Colorado, USA, 5-6 May. SPE-180270-MS.

Wang, M., Zhang, S., Zhang, F., et al. 2015. Quantitative Research on Tight Oil Microscopic State of Chang 7 Member of Triassic Yanchang Formation in Ordos Basin, NW China. *Petroleum Exploration and Development* **42**(6): 827-832.

Wu, Q., Xu, Y., Wang, X., et al. 2012. Volume Fracturing Technology of Unconventional Reservoirs: Connotation, Design Optimization and Implementation. *Petroleum Exploration and Development* **39**: 377-384.

Xiao, Y., Feng, S., Wang, J., et al. 2017. Stress Sensitivity and Its Influence Factors of Tight Oil Reservoir in Chang 7 Member, Ordos Basin. *China Petroleum Exploration* **22**(5): 64-72.

Yang, Z. , Crosby, D. G., and Khurana, A. K. 1996. Multivariate Optimisation of Hydraulic Fracture Design. *The Appea Journal* **36** 1): 516–527.

Zhan, C., Sankaran, S., LeMoine, V., et al. 2019. Application of Machine Learning for Production Forecasting for Unconventional Resources. Paper presented at the Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July. URTEC-2019-47-MS.

Zhang, X. and Awotunde, A. A. 2016. Improvement of Levenberg-Marquardt Algorithm during History Fitting for Reservoir Simulation. *Petroleum Exploration and Development* **43**(5): 876–85.

Zhou, Q., Dilmore, R., Kleiter, A., et al. 2014. Evaluating Gas Production Performances in Marcellus Using Data Mining Technologies. Paper presented at the Unconventional Resources Technology Conference, Denver, Colorado, 25-27 August. URTEC-1920211-MS.

Zou, C., Zhai, G., Zhang, G., et al. 2015. Formation, Distribution, Potential and Prediction of Global Conventional and Unconventional Hydrocarbon Resources. *Petroleum Exploration and Development* **42**(1): 14-28.

Shide Yan, is a master candidate in Petroleum Engineering department at Xi'an Shiyou University. He has focused his research in areas involving reservoir simulation and enhance oil recovery.

Weirong Li, is a Professor in the Petroleum Engineering Department at Xi'an Shiyou University. His research interests include unconventional resources/reserves estimates, reservoir simulation, well testing, and production analysis. Li holds a bachelor's degree in mathematics from Northeast Petroleum University, China; a master's degree in petroleum engineering from Research Institute of Petroleum Exploration and Development, China; and a PhD degree in petroleum engineering from Texas A&M University.

Min Zhang, is a master candidate in Petroleum Engineering department at Xi'an Shiyou University. She has focused her research in areas involving reservoir simulation, chemical flooding, and enhanced oil recovery.

Zhengbo Wang, is a senior reservoir engineering in Research Institute of Petroleum Exploration and Development, PetroChina. He specializes in enhanced oil recovery. Dr. Wang holds a a PhD degree in petroleum engineering from Research Institute of Petroleum Exploration and Development, PetroChina.

ZhaoXia Liu, is a senior reservoir engineering in Research Institute of Petroleum Exploration and Development, PetroChina. She specializes in enhanced oil recovery. Dr. Liu holds a a PhD degree in petroleum engineering from Research Institute of Petroleum Exploration and Development, PetroChina.

Keze Lin, is a undergraduate student of China University of Petroleum (Beijing), Beijing.

Hongliang Yi, is a senior reservoir engineering in Liaohe Oilfield Company of PetroChina. He specializes in enhanced oil recovery.