

Original Paper

An intelligent differential optimization method for fracture parameters in long horizontal wells based on reservoir quality zoning



Nan-Lin Zhang^{a,b}, Xin-Qi Zeng^a, Zhi-Feng Luo^{a,*}, Jun-Wen Tao^c, Ping-Li Liu^a,
Mao-Han Yuan^a, Qi Zhu^d, Fu-Shen Liu^e

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

^b School of Civil Engineering and Geomatics, Southwest Petroleum University, Chengdu, 610500, Sichuan, China

^c Changqing Oil Field Company, PetroChina, Xi'an, 710016, Shaanxi, China

^d CNPC Bohai Drilling Engineering Company Limited, Tianjin, 300457, China

^e Computing Center for Geotechnical Engineering, College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, 310058, Zhejiang, China

ARTICLE INFO

Article history:

Received 3 January 2026

Received in revised form

2 March 2026

Accepted 18 May 2026

Available online 27 May 2026

Edited by Meng-Jiao Zhou

Keywords:

Shale gas

Horizontal wells

Intelligent zoning

Fracture parameter optimization

Reservoir quality index

ABSTRACT

Shale reservoirs are characterized by pronounced heterogeneity, rendering conventional homogeneous fracture spacing and experience-based design approaches inadequate for achieving efficient hydraulic fracturing. Meanwhile, numerical simulation methods suffer from high computational cost and limited global optimization capability. To address these challenges, this study proposes an intelligent optimization method for fracture parameter design. First, a reservoir quality index (RQI) is constructed to perform accurate zoning along the horizontal section, enabling targeted fracture parameter design for intervals with distinct reservoir properties. Subsequently, the Tent chaotic mapping, golden sine strategy, and sparrow search algorithm (SSA) are integrated with a backpropagation neural network to develop a TENT-SSA-BPNN model, thereby enhancing the global search capability of fracture parameters. Finally, the net present value (NPV) is adopted as the optimization objective, and the optimal fracture parameter combination is obtained under physical constraints including fracture half-length, fracture conductivity, and stage number. Results demonstrate that the proposed method outperforms conventional models in both accuracy and stability. The optimized design achieves approximately a 30% increase in production, accompanied by a 3 to 5 fold improvement in NPV. The proposed approach provides an efficient and feasible technical pathway for differentiated and precision stimulation design of shale gas horizontal wells.

© 2026 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Shale gas, as one of the most important clean energy resources, has become a key component of global energy transition and carbon reduction strategies (Zhao et al., 2025). Owing to the ultra-low porosity, ultra-low permeability, and strong multiscale heterogeneity of shale reservoirs, effective development cannot be achieved by relying solely on natural productivity (Gao et al., 2025). Consequently, horizontal wells combined with multistage hydraulic fracturing have gradually become the core technologies

for large-scale shale gas development (Du et al., 2025). Horizontal wells significantly increase the contact area between the wellbore and the reservoir (Sun et al., 2026), while multistage hydraulic fracturing creates multiple artificial fractures along the horizontal section, thereby enlarging the stimulated reservoir volume and enhancing single-well production (Chu et al., 2025). However, field practice indicates that a continuous increase in fracturing scale does not necessarily lead to a proportional improvement in production. On the one hand, increasing the number of fracturing stages and fracture size substantially raises construction costs and operational risks (Wei et al., 2025). On the other hand, when reservoir properties vary significantly along the wellbore, the adoption of uniform fracture placement and identical stimulation parameters may cause certain well sections with better reservoir quality or lower in-situ stress to develop fractures with dominant propagation capability, known as dominant hydraulic fractures

* Corresponding author.

E-mail address: lzf103429@163.com (Z.-F. Luo).

Peer review under the responsibility of China University of Petroleum (Beijing).

(Suo et al., 2026). During fracture propagation, these dominant hydraulic fractures suppress the initiation and growth of adjacent fractures through stress shadow effects, resulting in an unevenly developed fracture network concentrated in limited regions (Tang et al., 2026), which ultimately reduces stimulation efficiency (Jiang et al., 2026). Therefore, under the constraints of engineering feasibility and economic performance, rational optimization of key fracture parameters, such as fracture length, number of fracturing stages, and fracture conductivity, has become a critical scientific and engineering challenge in shale gas development.

Reservoir heterogeneity is mainly manifested by pronounced variations in reservoir properties along the horizontal wellbore. Different well sections exhibit significant differences in porosity, permeability, formation pressure, gas saturation, and natural fracture density (Yin et al., 2025). Such along-wellbore heterogeneity leads to fundamentally different responses of various well sections to hydraulic fracturing, thereby imposing higher requirements on fracture parameter optimization. Generally, for well sections with favorable reservoir properties, stimulation designs tend to focus on creating high-conductivity, relatively short fractures, aiming to alleviate near-wellbore damage and reduce pressure drop around the wellbore (Yao et al., 2023). In contrast, for well sections with poor reservoir quality, longer fractures are usually required to increase the contact area between hydraulic fractures and the reservoir (Jin, 2025). Under conditions of strong heterogeneity along long horizontal wells, if a uniform stimulation or evenly segmented design strategy is still applied, forcing well sections with different reservoir qualities to adopt identical fracture lengths and conductivities, the mismatch between fracture parameters and reservoir quality may result in suboptimal fracture conductivity, fracture length, and fracture density, thereby weakening the overall stimulation effectiveness.

Traditional studies on hydraulic fracture parameter optimization have primarily relied on reservoir numerical simulation methods. By constructing coupled wellbore–fracture–reservoir flow models, the impacts of different fracture parameter combinations on production performance can be systematically evaluated to identify optimal fracture designs. Numerous sensitivity analyses have been conducted based on equivalent fracture models (Ji et al., 2026), dual-porosity/dual-permeability models, and discrete fracture models (Zhang et al., 2021), focusing on fracture half-length, number of fractures, fracture placement patterns, and fracture conductivity. These methods possess clear physical foundations and can accurately describe the flow behavior of wellbore–fracture–reservoir systems. However, in strongly heterogeneous shale reservoirs, the interaction between natural fractures and hydraulic fractures becomes significantly enhanced, and the relationship between fracture parameters and production exhibits strong nonlinearity (Qin et al., 2023). As a result, numerical simulation approaches are computationally expensive and often fail to adequately explore high-dimensional parameter spaces during multi-parameter optimization, making them prone to convergence toward local optima (Xu et al., 2017).

With the rapid development of artificial intelligence and data-driven technologies, machine learning methods have increasingly been introduced into fracture parameter optimization studies (Zhao et al., 2022). By constructing surrogate models that relate fracture parameters to production or economic indicators and training them using historical data, parameter prediction and optimization can be performed efficiently without repeatedly invoking numerical simulations, thereby substantially improving optimization efficiency (He et al., 2025). For example, surrogate models based on neural networks, radial basis functions, or differential evolution algorithms can effectively capture the nonlinear relationships between fracture parameters, production,

and economic indicators (Fang et al., 2025; Guo et al., 2025; Huang et al., 2025; Nan et al., 2025). However, purely data-driven models may generate optimization results that violate practical fracturing constraints when physical and engineering conditions are not explicitly considered, for instance, fracture length, fracture conductivity, or number of fracturing stages exceeding feasible operational limits (Ren et al., 2023; Dong et al., 2025; Chen et al., 2026; Niu et al., 2026), thus limiting their engineering applicability and reliability. Moreover, although some algorithms (Xu et al., 2021) can incorporate prior knowledge to enhance optimization efficiency, their performance remains sensitive to the quality of initial samples and objective definitions, and they often lack a systematic integration of physical constraints and economic objectives (Gou et al., 2024). To address these issues, several studies have attempted to introduce prior knowledge or simplified physical models to constrain intelligent optimization algorithms (Niu et al., 2024). Nevertheless, most of these approaches focus on optimizing a single or a limited number of parameters and typically assume that fracture optimization is conducted at the whole-well or uniformly segmented level. As a result, they fail to fundamentally resolve the issue of inappropriate optimization unit definition under strong along-wellbore heterogeneity.

For long horizontal wells, systematic variations in reservoir quality along the wellbore imply that different well sections should correspond to different optimal fracture parameters. If strongly heterogeneous reservoirs are directly input into machine learning models for implicit learning, not only does the training difficulty increase, but the underlying physical response relationships may also become mixed within the sample space. Therefore, prior to fracture parameter optimization, it is essential to apply a rational segmentation method based on reservoir properties to divide the wellbore into multiple relatively homogeneous target sections. Such fine segmentation reconstructs the optimization problem at the physical-structural level and provides a solid foundation for subsequent multi-parameter joint optimization.

Accordingly, this study proposes a physics–data collaborative intelligent fracture parameter optimization method based on fine reservoir quality segmentation (Fig. 1). First, a reservoir quality index (RQI) is constructed to perform precise segmentation of long horizontal wells, dividing strongly heterogeneous reservoirs into multiple differentiated well section units. Then, Tent chaotic mapping, the golden sine strategy, and the sparrow search algorithm (SSA) are introduced to improve the BPNN model, thereby enhancing its global search capability in multi-constraint, high-dimensional decision spaces. Finally, with net present value (NPV) as the optimization objective, optimal fracture parameter combinations for different reservoir sections are obtained under explicit physical constraints on fracture half-length, fracture conductivity, and number of fracturing stages. Compared with optimization approaches that focus solely on production, NPV explicitly captures the trade-off between incremental production benefits and fracturing costs, making it more consistent with practical engineering decision-making. The proposed method provides an intelligent optimization pathway for horizontal well fracturing design in heterogeneous shale gas reservoirs that balances engineering feasibility and economic optimality.

2. A physics–data collaborative driven method for fracture parameter optimization in horizontal wells

Shale gas reservoirs, as a typical type of unconventional hydrocarbon reservoir, exhibit pronounced heterogeneity, manifested by strong spatial variations in lithological composition, petrophysical properties, fracture development, and gas-bearing capacity. Such complex geological characteristics render

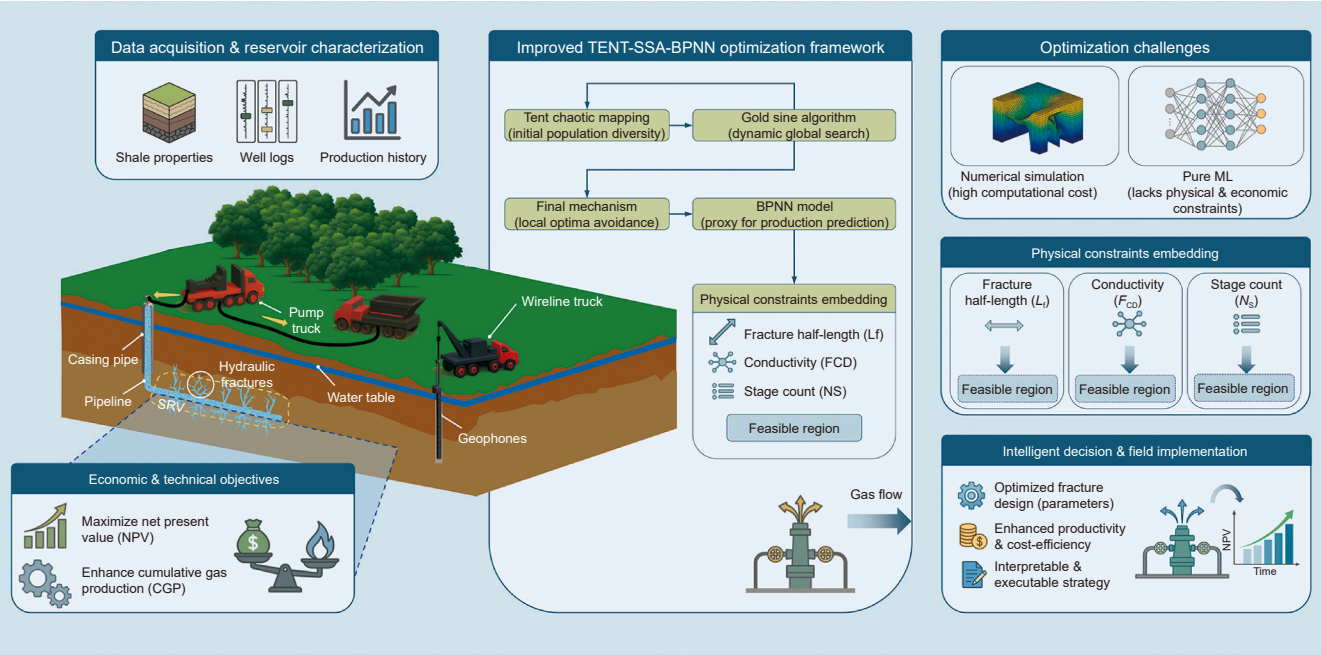


Fig. 1. A physics–data collaborative driven research framework for fracture parameter optimization.

conventional uniform or experience-based stage partitioning methods incapable of effectively matching actual stimulation requirements, often resulting in inefficient fracture energy utilization or insufficient reservoir stimulation. To improve the efficiency and accuracy of fracture parameter optimization, this study develops an intelligent optimization algorithm framework integrated with physical constraints (Fig. 2). The framework is built upon both reservoir quality index (RQI)-based zoning data and original reservoir datasets. Key controlling factors are first identified and

normalized. Subsequently, an interpolation approach combined with a median-based zoning method is employed to perform accurate segmentation of the near-wellbore reservoir along the horizontal section, enabling a precise characterization of strong reservoir heterogeneity. A backpropagation neural network (BPNN) is then constructed as the production prediction model. To further enhance optimization performance, Tent chaotic mapping and the sparrow search algorithm (SSA) are introduced to optimize network weights and strengthen the global search capability,

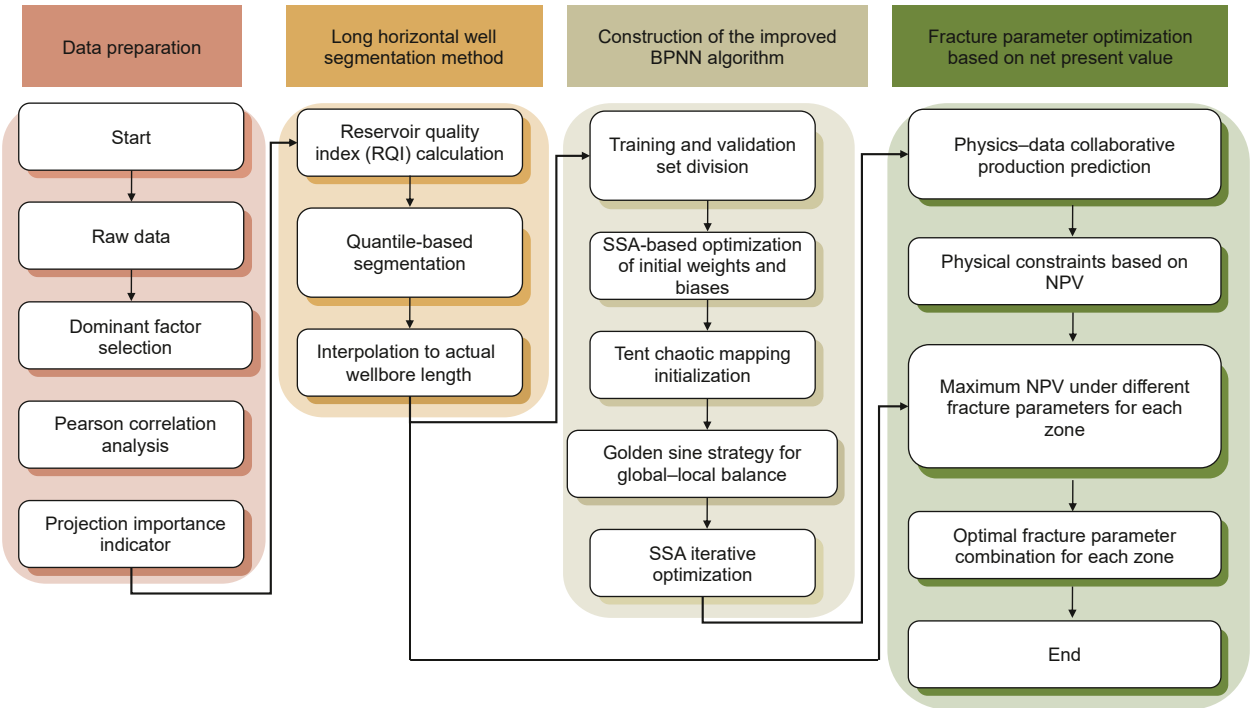


Fig. 2. Flowchart of the physics–data collaborative driven fracture parameter optimization process.

thereby achieving a coupled workflow of production prediction and fracture parameter optimization. Finally, the net present value (NPV) is adopted as the optimization objective to determine the optimal fracture parameter combination for each intelligently zoned interval under explicit physical constraints. In this manner, precision optimization of fracture parameters in horizontal wells under strongly heterogeneous reservoir conditions is achieved.

2.1. Zoning method for long horizontal well sections

The complexity of shale gas reservoirs is reflected in the spatial heterogeneity of multiple parameters, including porosity, permeability, lithology, temperature, and pressure. Variations in geological conditions, depositional environments, fracture development, as well as thermal and pressure regimes lead to pronounced spatial differences in these parameters, thereby exerting significant influence on gas storage capacity, flow behavior, and production performance of shale gas reservoirs. To identify the key controlling factors governing fine-scale zoning of horizontal wells and fracture parameter optimization, data standardization was first performed. Taking the three-year cumulative gas production after stimulation as the target variable, Pearson correlation analysis and permutation importance analysis were applied to evaluate the relationships between near-wellbore reservoir parameters along the horizontal section and post-fracturing fracture parameters. The identified dominant factors were subsequently utilized to guide the design of stage partitioning and fracture parameters for horizontal wells.

The Pearson correlation coefficient is expressed as:

$$X_j(k) = \frac{x_j(k)}{\bar{x}_j(k)} \quad (1)$$

$$r = \frac{\sum_{k=1}^n [X_j(k) - \bar{X}_j] [Y(k) - \bar{Y}]}{\sqrt{\sum_{k=1}^n [X_j(k) - \bar{X}_j]^2} \sqrt{\sum_{k=1}^n [Y(k) - \bar{Y}]^2}} \quad (2)$$

The permutation importance of the j -th feature is defined as:

$$I_{vipj} = \sqrt{\frac{m}{R_d(Y, t_1, t_2, \dots, t_h)} \sum_{k=1}^h R_d(Y, t_k) w_{kj}^2} \quad (3)$$

X denotes the standardized independent variable; the subscript j denotes the index of the independent variable, $j = 1, 2, \dots, m$, $k = 1, 2, \dots, n$; x represents the original independent variable; $\bar{x}$ denotes the mean value of the independent variable; m denotes the number of independent variables; n denotes the number of samples; r denotes the Pearson correlation coefficient; $\bar{X}$ denotes the mean of the standardized independent variable; $\bar{Y}$ denotes the mean of the standardized dependent variable; Y denotes the standardized dependent variable; I_{vipj} denotes the projection importance index of the j -th variable; Y denotes the k -th principal component; h denotes the number of principal components; $R_d(Y, t_1, t_2, \dots, t_h)$ denotes the correlation coefficient between Y and all principal components, representing the overall explanatory power of all principal components for Y ; $R_d(Y, t_k)$ denotes the correlation coefficient between Y and the single principal component t_k , representing the explanatory ability of t_k for Y .

By jointly considering reservoir parameters and fracture parameters, nine geological parameters with significant correlations were identified, as illustrated in Fig. 3. Based on these key variables, a reservoir quality index (RQI) is defined to perform

quantitative reservoir evaluation for horizontal wells, which is expressed as:

$$y_{\text{hor}} = 0.07x_f + 0.20C_f + 0.42h + 0.16k_m + 0.21\phi_m + 0.47P_{\text{res}} + 0.19S_g + 0.08L_h + 0.23N_s \quad (4)$$

where y_{hor} denotes the reservoir quality index of the horizontal well; x_f denotes the standardized fracture half-length; C_f denotes the standardized fracture conductivity; h denotes the standardized reservoir thickness; k_m denotes the standardized matrix permeability; ϕ_m denotes the standardized matrix porosity; P_{res} denotes the standardized reservoir pressure; S_g denotes the standardized gas saturation; L_h denotes the standardized horizontal section length; N_s denotes the standardized number of fracturing stages.

To evaluate the multicollinearity among the selected independent variables, the Variance Inflation Factor (VIF) method was employed. Then, the Variance Inflation Factor (VIF) is calculated as:

$$\text{VIF}_j = \frac{1}{1 - R_j^2} \quad (5)$$

In the equation, VIF_j is the variance inflation factor of variable; R_j^2 represents the linear explanatory power of the other m independent variables on results. Generally, when the VIF_j is greater than 10, it indicates that there is a high correlation among the explanatory variables and a serious degree of multicollinearity. The VIF analysis shows that all input features have variance inflation factors ranging from 1.01 to 2.80, well below the commonly adopted multicollinearity threshold of 5.0. The relatively higher VIF values for horizontal well length (2.80) and number of fracturing stages (2.33) reflect the well-known field practice that longer horizontal sections generally require more fracturing stages. This weak physical coupling does not introduce meaningful redundancy into the model inputs. The remaining features exhibit VIF values close to unity, indicating near-complete statistical independence. These findings are consistent with the Pearson correlation matrix presented in Fig. 3(b), collectively confirming that the selected feature set is free from collinearity issues. The BPNN model therefore captures genuine reservoir–fracture response behavior rather than overfitting to empirical correlations embedded in historical completion designs.

Although the RQI formulation includes fracture-related parameters (fracture half-length, fracture conductivity, and number of fracturing stages), these are not used as optimization variables in the RQI-based zoning stage. Instead, the fracture parameters used in the RQI calculation are derived from historical field completion data of offset wells, which represent the as-completed fracture configurations in each well section. The purpose of including these parameters in the RQI is to characterize the combined geological-engineering quality of each interval based on existing field data. The subsequent optimization step then treats these fracture parameters as decision variables with predefined physical constraints (Eq. (13)), independently searching for optimal values that maximize NPV for each zoned interval. Therefore, the RQI serves as a pre-optimization diagnostic index rather than a component of the optimization loop, and no methodological circularity or information leakage exists in the proposed workflow (Fig. 4).

Because shale gas wells typically experience a substantial production decline to a relatively low level after approximately three years of production, refracturing is often required. Therefore, the three-year cumulative gas production was selected as the

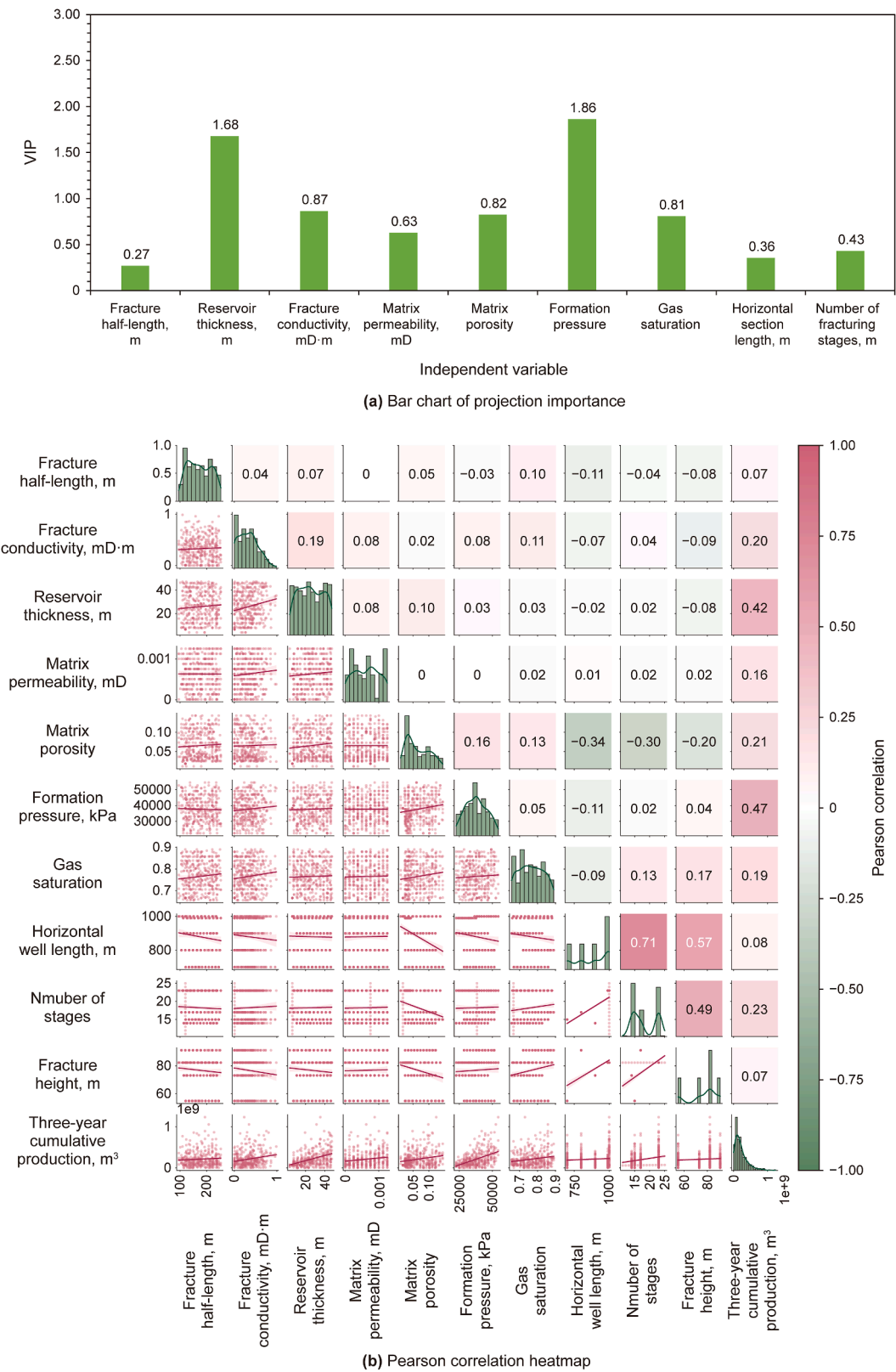


Fig. 3. Parameter importance analysis for variable selection.

dependent variable in this study. Fig. 5 illustrates the relationship between the reservoir quality index (RQI) and the post-stimulation three-year cumulative gas production for 30 horizontal wells. A clear positive correlation can be observed between the two variables. The coefficient of determination reaches $R^2 = 0.9496$, indicating that the proposed RQI can accurately characterize near-

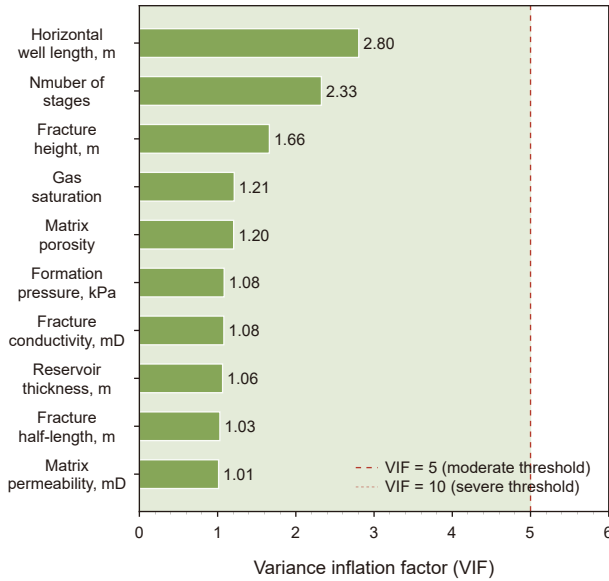


Fig. 4. Variance inflation factor (VIF) analysis.

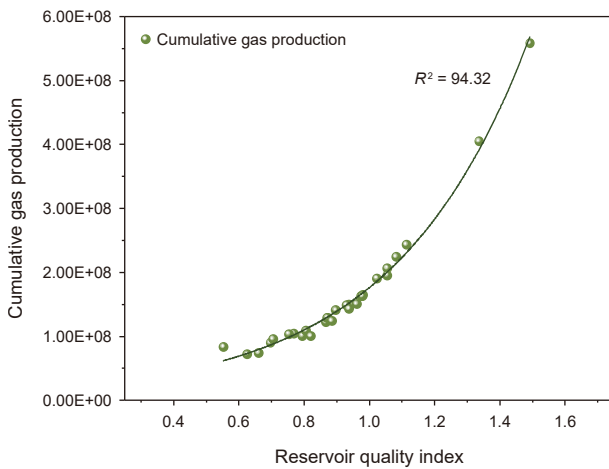


Fig. 5. Scatter plot showing the correlation between the reservoir quality index (RQI) and three-year cumulative gas production.

wellbore reservoir quality and effectively evaluate reservoir characteristics relevant to production performance.

To achieve targeted fracture parameter optimization for horizontal wells in strongly heterogeneous shale gas reservoirs, it is necessary to perform fine-scale zoning based on near-wellbore reservoir characteristics, while jointly considering post-stimulation fracture parameters and production data to enable rational fracture placement design.

In this study, a quantile-based dynamic zoning method is adopted. Taking Well X as an example, the reservoir quality index profile along the horizontal section is first calculated based on the available characteristic parameters, and a color gradient is used to represent the magnitude of the reservoir quality index (RQI). Subsequently, dynamic zoning is performed according to the following rules.

- (1) Zone boundaries are determined dynamically based on variations in the RQI values, rather than being uniformly divided along the horizontal section.

- (2) Key intervals of data variation are identified using quantiles, where the selected quantiles in this study include the minimum value, 25th percentile, 50th percentile, 75th percentile, and maximum value.

The quantile calculation is given by:

$$Q_p = \text{percentile}(y, p) \quad (6)$$

where Q_p denotes the p quantile of the dataset y , representing the value located at the p position in the ordered data sequence. The dataset y is first arranged in ascending order, and Q_p represents the value below which p of the observations fall.

- (3) To map the quantile positions to the actual horizontal well length, an interpolation-based approach is employed to ensure that each zoned interval accurately reflects the distribution characteristics of the RQI.

The interpolation equation is expressed as:

$$\text{Area}_p = \frac{Q_p - \min(y)}{\max(y) - \min(y)} L_x \quad (7)$$

where Area_p denotes the corresponding position of the p quantile along the horizontal wellbore, the location on the horizontal section length; $\max(y)$ and $\min(y)$ represent the maximum and minimum values of the RQI, respectively; L_x denotes the total length of the horizontal section, m.

- (4) Based on the interpolated zoning results, the entire horizontal section is divided into multiple zones with distinct reservoir quality characteristics, as illustrated in Fig. 6 and summarized in Table 1.

To evaluate the robustness of the proposed zoning scheme, a sensitivity analysis was conducted by varying the number of quantile divisions (3-zone, 4-zone, 5-zone, and 6-zone configurations) (see Table 2). Fig. 7 illustrates the results indicate that when fewer than four zones are used, the zoning resolution is insufficient to distinguish the local variations in reservoir quality, leading to suboptimal fracture parameter assignment. When more than five zones are adopted, the marginal improvement in NPV becomes negligible (less than 2%), while the operational complexity of implementing zone-specific fracture designs increases substantially. The five-zone configuration adopted in this study represents an optimal balance between zoning resolution and engineering practicality. Furthermore, Fig. 8 illustrates the quantile thresholds (minimum, 25th, 50th, 75th percentiles, and maximum) are statistically robust, non-parametric descriptors that do not require assumptions about the data distribution.

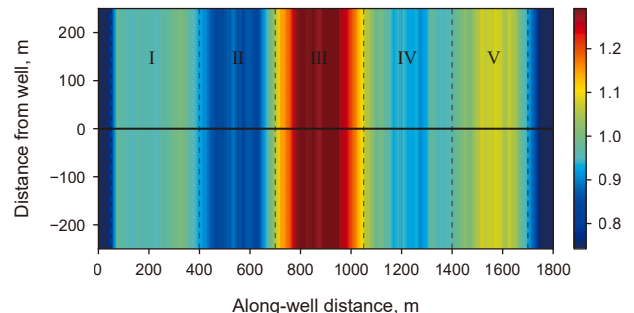


Fig. 6. Distribution of the reservoir quality index (RQI) along the horizontal wellbore.

Table 1
Statistical distribution of reservoir and fracture parameters in different zoned regions.

Parameter	I	II	III	IV	V
Fracture half-length, m	190.52	165.93	183.15	179.69	170.14
Reservoir thickness, m	27.21	25.00	26.16	26.87	25.34
Fracture conductivity, mD·m	1.14	1.07	1.26	1.25	1.12
Matrix permeability, mD	0.00061	0.00058	0.00072	0.00062	0.00063
Matrix porosity	0.034	0.034	0.081	0.073	0.069
Reservoir pressure, MPa	30.50	30.32	40.02	36.68	39.66
Gas saturation	0.75	0.78	0.76	0.77	0.75
Horizontal section length, m	400	250	350	350	300
Number of fracturing stages	23	16	23	15	17
Reservoir quality index	0.9–1.0	0.7–0.8	>1.2	0.9–1.0	1.0–1.1

Table 2
Different plans and strategies.

Scheme	Implementation strategy
Adopted	25th percentile, 50th percentile, 75th percentile, and maximum value.
Scheme B	20th percentile, 40th percentile, 60th percentile, and 80th percentile.
Scheme C	30th percentile, 55th percentile, 70th percentile, and maximum value.
Scheme D	20th percentile, 48th percentile, 73rd percentile, and maximum value.

Alternative quantile schemes (e.g., 20th/40th/60th/80th percentiles) were also tested and produced similar zoning boundaries (within 5% deviation), confirming that the results are not highly sensitive to the specific quantile choices. Although the demonstration is presented for a single well (Well X), the zoning methodology is general and can be applied to any horizontal well with available logging and completion data (see Fig. 23).

After applying the quantile-based dynamic zoning method, the average parameter values and RQI ranges for different zones are obtained. By performing reservoir quality-based zoning of the near-wellbore reservoir, more precise fracture parameter design can be implemented for each horizontal well interval, thereby improving stimulation effectiveness while maintaining cost control. Compared with traditional evenly spaced stage partitioning, which assigns identical fracture parameters to all intervals regardless of reservoir heterogeneity, the proposed data-driven zoning approach captures the spatial variability of reservoir quality along the horizontal wellbore and enables differentiated fracture designs for each zone. The superiority of this approach is further evidenced by the simulation results presented in Section 3.5 (Fig. 24), where the zoned fracture design yields a more

uniform pore-fluid pressure distribution and significantly higher cumulative gas production than the conventional uniform design.

2.2. Construction of the BPNN model

Common neural network models include the backpropagation neural network (BPNN), recurrent neural network (RNN), and convolutional neural network (CNN). In this study, the widely adopted BPNN is employed as the base predictive model. A BPNN consists of multiple neuron nodes and typically comprises three layers: an input layer, one or more hidden layers, and an output layer (Fig. 9). Data are fed into the input layer and propagated forward through the network via weighted connections between neurons. Each neuron performs a linear combination of its inputs using weights and biases, and then applies a nonlinear activation function to generate the output. Backpropagation constitutes the core mechanism of the BPNN. If the discrepancy between the predicted output and the ground truth does not satisfy a pre-defined accuracy criterion, the prediction error is propagated backward from the output layer to the hidden and input layers. The network parameters (weights and biases) are iteratively updated to minimize the error until the prescribed accuracy requirement is met (Liu et al., 2025).

2.3. Improvement of the BPNN model based on TENT-SSA

In conventional BPNN training, the initial weights and biases are typically assigned in a random manner, which often leads to slow convergence during the backpropagation process and increases the risk of the network being trapped in local optima (Su et al., 2024). To alleviate these limitations, the sparrow search algorithm (SSA) is introduced in this study to optimize the initial weights and biases of the neural network, as illustrated in Fig. 10. By performing global optimization in the parameter space prior to gradient-based learning, SSA provides high-quality initial solutions for the BPNN, thereby accelerating convergence and enhancing the stability and robustness of the training process.

The sparrow search algorithm (SSA) is a novel swarm intelligence optimization algorithm proposed in 2020, which effectively addresses common shortcomings of traditional metaheuristic algorithms, such as weak global search capability and slow convergence. The algorithm is inspired by the foraging and anti-predation behaviors of sparrow populations in nature (Xue and Shen, 2020).

During the processes of foraging and anti-predation, the sparrow population is divided into three roles: producers (discoverers), scroungers (joiners), and scouts (warners). Producers are responsible for exploring the environment, identifying food sources, and providing search directions for the population. Once a food source is located by the producers, scroungers move toward favorable positions to obtain food. Meanwhile, scouts remain vigilant to potential threats arising from predators or competition within and

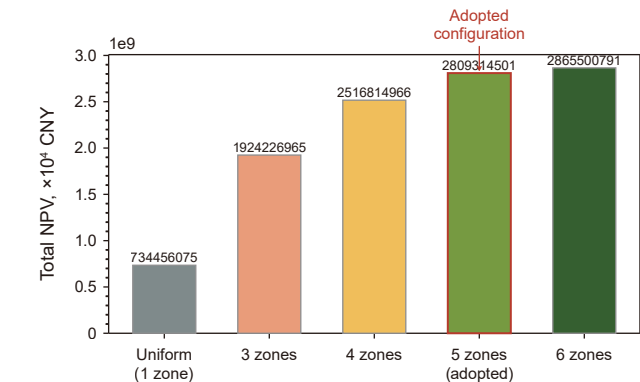


Fig. 7. Zone configuration sensitivity.

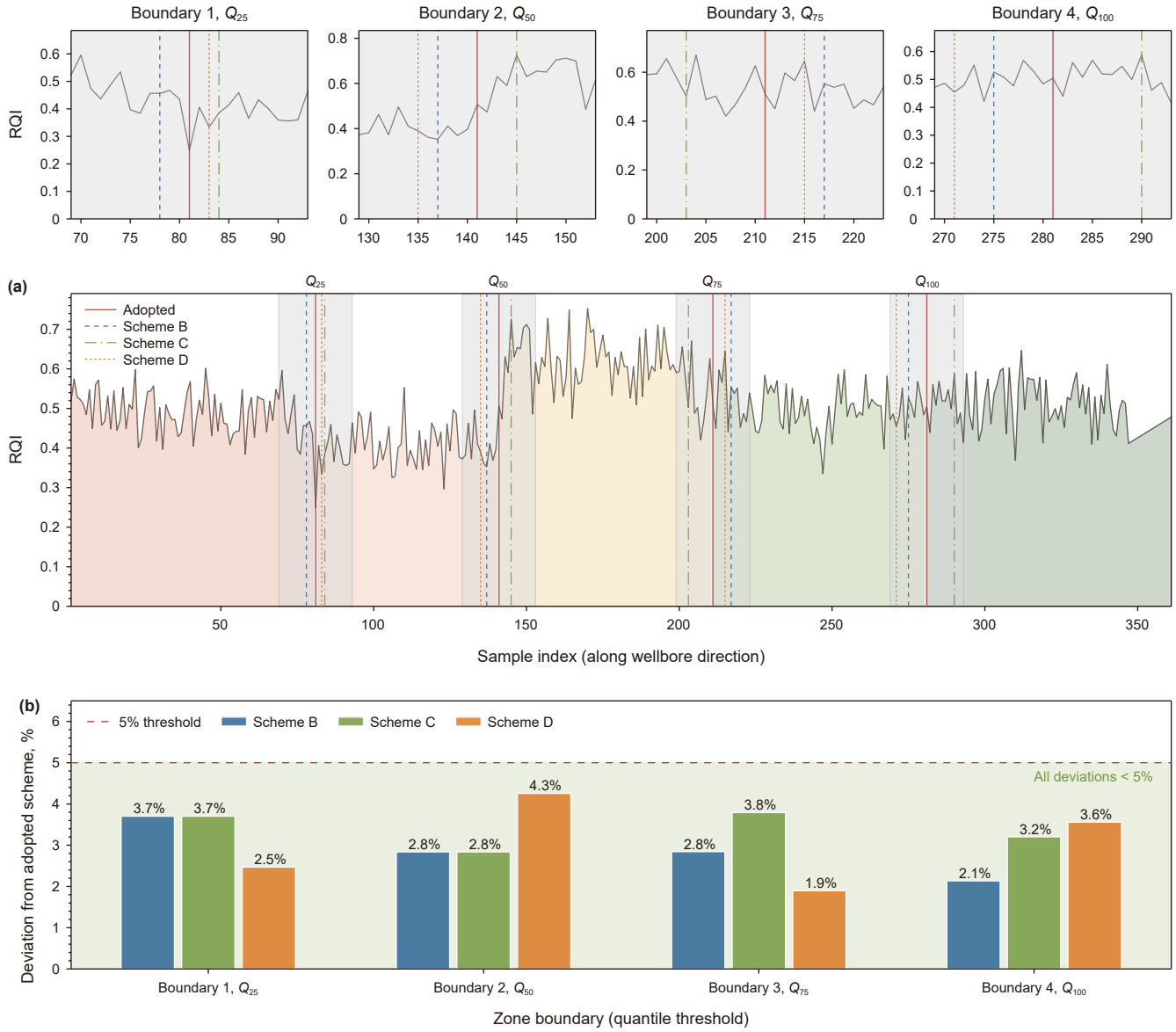


Fig. 8. Quantile threshold sensitivity.

outside the population. When a threat is detected, the population initiates anti-predation behavior to avoid danger and maintain population safety (Chen et al., 2024).

The SSA iteratively updates the positions of individuals based on a fitness function, following specific update rules associated with different roles, thereby achieving adaptive global optimization of the objective function. The detailed update mechanisms are described as follows.

- (1) The initial population is generated with n sparrows. The position of each sparrow is represented by a D -dimensional vector $\mathbf{X} = (x_1, x_2, \dots, x_D)$, and the corresponding fitness value is evaluated as $f = f(x_1, x_2, \dots, x_D)$ (Cheng et al., 2022).
- (2) A proportion of 10%–20% individuals with higher fitness values are selected as producers (discoverers), which are responsible for searching for food sources. To explore higher-quality solutions and satisfy the foraging requirements of the population, the search range of producers is larger than that of scroungers. During each iteration, the

positions of producers are updated according to Eq. (7) while seeking the global optimum (Yu et al., 2024).

$$\mathbf{x}_{ij}^{t+1} = \begin{cases} \mathbf{x}_{ij}^t \cdot \exp\left(\frac{-i}{\alpha \cdot \text{iter}_{\max}}\right), & \text{if } R_2 < ST \\ \mathbf{x}_{ij}^t + Q \cdot \mathbf{L}, & \text{if } R_2 \geq ST \end{cases} \quad (8)$$

where $\mathbf{x}_{ij}^t$ denotes the position of the i -th individual in the j -th dimension at the t -th iteration; $\mathbf{x}_{ij}^{t+1}$ denotes the updated position at the $(t+1)$ -th iteration; $\alpha [0, 1]$ is a random number; $\text{iter}_{\max}$ is the maximum number of iterations; Q is a random number following a normal distribution; $\mathbf{L}$ is a vector with all elements equal to 1; $R_2 [0, 1]$ is a warning value; $ST [0.5, 1]$ is the safety threshold.

When $R_2 < ST$, the producer is considered to be in a safe state, and the search range is expanded to enhance global exploration. When $R_2 \geq ST$, the producer is regarded as being in a dangerous state and moves toward other safe positions to avoid potential threats.

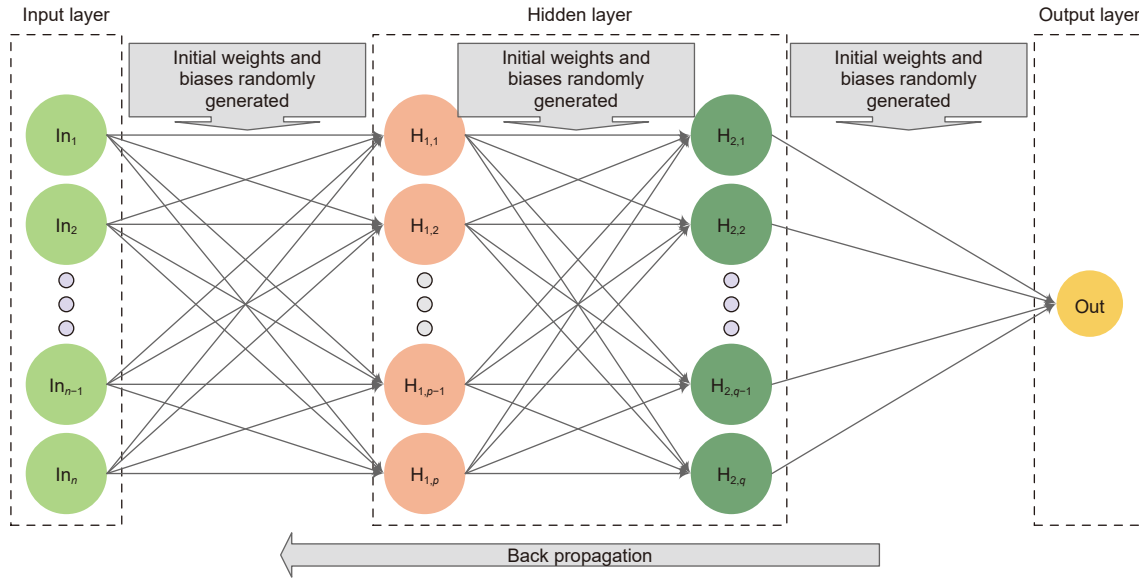


Fig. 9. Schematic diagram of the conventional backpropagation neural network (BPNN).

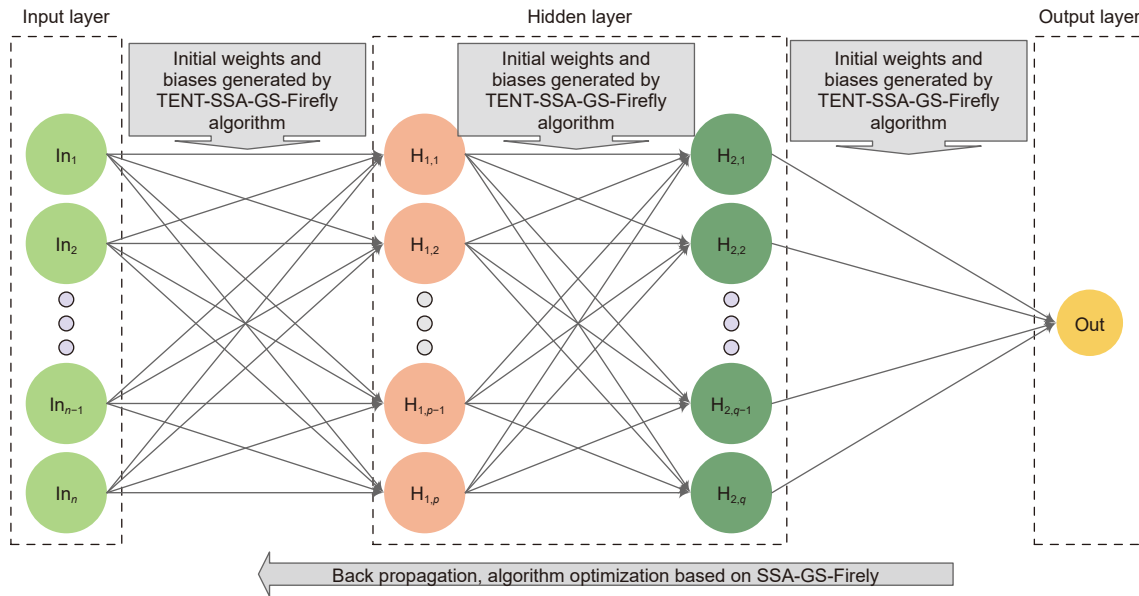


Fig. 10. Architecture of the improved TENT-SSA-BPNN model.

- (3) Scroungers determine whether they can obtain energy based on food quality and their relative positions. When a scrounger perceives that a producer has acquired higher-quality food, it moves forward to compete for the resource. If the competition is successful, the scrounger replaces the position of the producer (during this process, the numbers of producers and scroungers are dynamically adjusted). Otherwise, the position of the scrounger is updated according to Eq. (8).

The position update rule for scroungers is given by:

$$\mathbf{x}_{ij}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{\mathbf{x}_{\text{worst}}^t - \mathbf{x}_{ij}^t}{i^2}\right), & \text{if } i > \frac{n}{2} \\ \mathbf{x}_{\text{best}}^{t+1} + |\mathbf{x}_{ij}^t - \mathbf{x}_{\text{best}}^{t+1}| \cdot \mathbf{A}^+ \cdot \mathbf{L}, & \text{otherwise} \end{cases} \quad (9)$$

where $\mathbf{x}_{\text{worst}}^t$ denotes the position of the worst individual at the t -th iteration; $\mathbf{x}_{\text{best}}^{t+1}$ denotes the best position obtained by the producers at the $(t+1)$ -th iteration; $\mathbf{A}$ is a random matrix with elements of either 1 or -1, $\mathbf{A}^+ = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$ is the

Moore–Penrose pseudoinverse of $\mathbf{A}$; n denotes the number of scroungers.

When $i > \frac{n}{2}$, the corresponding sparrow exhibits relatively low fitness and fails to obtain food due to insufficient competitiveness, indicating a hungry state. In this case, the individual searches for food in other regions to better satisfy the energy requirements of the population.

- (4) A proportion of 10%–20% individuals in the population are randomly selected as scouts (warners). Scouts monitor potential danger during the search process and issue warning signals when a threat is detected. Upon receiving a warning, the population searches for new positions to forage, and the positions of scouts are updated according to Eq. (9).

$$\mathbf{X}_{ij}^{t+1} = \begin{cases} \mathbf{X}_{\text{best}}^t + \beta \cdot |\mathbf{X}_{ij}^t - \mathbf{X}_{\text{best}}^t|, & \text{if } f_i > f_g \\ \mathbf{X}_{ij}^t + K \cdot \left(\frac{|\mathbf{X}_{ij}^t - \mathbf{X}_{\text{worst}}^t|}{(f_i - f_w) + \varepsilon} \right), & \text{if } f_i = f_g \end{cases} \quad (10)$$

where β is a random number drawn from a normal distribution with mean 0 and variance 1; K is a random number uniformly distributed in $[-1, 1]$; ε is a very small constant to avoid division by zero; f_i denotes the fitness of the current individual; f_g denotes the global best fitness, f_w denotes the worst fitness in the population.

When $f_i > f_g$, the individual is considered to be located at the boundary of the population, where it is more vulnerable to attacks, and thus moves toward the best individual. When $f_i = f_g$, the individual is regarded as being within the population core but still potentially close to danger, and therefore approaches other individuals to enhance safety.

Compared with other intelligent optimization algorithms, the sparrow search algorithm (SSA) exhibits high search accuracy, fast convergence speed, and strong robustness. However, similar to many metaheuristic algorithms, when the search process approaches the global optimum, population diversity tends to decrease, which may cause the algorithm to be trapped in local optima. To alleviate this issue, a Tent chaotic mapping–based initialization strategy is introduced to replace purely random population initialization. Furthermore, the original exponential position update strategy of producers is improved by incorporating a golden sine–based position update mechanism, and the best individual is retained to enhance convergence reliability.

The Tent chaotic mapping is defined as:

$$x_{n+1} = \begin{cases} \frac{x_n}{\alpha}, & \text{if } x_n \in [0, \alpha) \\ \frac{1 - x_n}{1 - \alpha}, & \text{if } x_n \in [\alpha, 1] \end{cases} \quad (11)$$

where x_n denotes the initial system value; the subscript n represents the iteration index of the sparrow population; and α is the control parameter of the Tent map, with $\alpha \in [0, 1]$.

When applying the Tent chaotic mapping, the initial value x_n and the parameter α must not be identical; otherwise, the system degenerates into a periodic state rather than evolving into an aperiodic (chaotic) system. In this study, $\alpha = 0.501$, is adopted. Fig. 8(a) illustrates the distribution of the Tent chaotic sequence after 50,000 iterations, showing that the chaotic values are uniformly distributed within the interval $[0, 1]$. Fig. 11(b) further presents the frequency distribution of chaotic values across different subintervals, indicating an approximately uniform distribution. These characteristics demonstrate that Tent chaotic mapping can effectively

enhance population diversity and improve the global exploration capability during the initialization stage.

The rationale for selecting these two specific improvement strategies is as follows. First, the primary weakness of the standard SSA lies in insufficient population diversity during initialization, which increases the risk of premature convergence. Among various chaotic mapping techniques (e.g., Logistic, Chebyshev, and Tent maps), Tent chaotic mapping was chosen because it provides a more uniform distribution within the $[0, 1]$ interval compared with Logistic mapping, which tends to cluster near the boundary values 0 and 1. This superior uniformity, as demonstrated in Fig. 12, ensures better spatial coverage of the search domain during population initialization. Second, the standard SSA employs an exponential position update for producers, which may overshoot the optimal region in the later stages of iteration. The golden sine strategy was introduced because its sinusoidal oscillation combined with the golden section ratio provides an inherent mechanism for balancing global exploration and local exploitation, enabling more refined search behavior near the optimum without sacrificing the ability to escape local optima. Together, these two improvements complement each other: Tent mapping enhances the diversity of the initial search, while the golden sine strategy refines the subsequent iterative search process.

The golden sine strategy is an iterative optimization method based on sinusoidal functions. To balance the trade-off between global exploration and local exploitation during the SSA search process, a golden section–based sine update mechanism is introduced to improve the position update of producers. The improved producer position update rule is expressed as (Tanyildizi and Demir, 2017; Wang et al., 2024):

$$\mathbf{X}_{ij}^{t+1} = \begin{cases} \mathbf{X}_{ij}^t \cdot |\sin \gamma_1| - \gamma_2 \cdot \sin \gamma_1 \cdot |c_1 \mathbf{X}_p^t - c_2 \mathbf{X}_{ij}^t|, & \text{if } R_2 < ST \\ \mathbf{X}_{ij}^t + Q \cdot L, & \text{if } R_2 \geq ST \end{cases} \quad (12)$$

where γ_1 is a random number uniformly distributed in $[0, 2\pi]$; γ_2 is a random number uniformly distributed in $[0, \pi]$; $\mathbf{X}_p^t$ denotes the target position; c_1 and c_2 are the golden section coefficients.

2.4. Fracture parameter optimization objective based on net present value

For shale gas horizontal wells, the number of fracturing stages and fracture conductivity are two key parameters that predominantly determine the effectiveness of multistage hydraulic fracturing (Zhang et al., 2024). In general, increasing the number of hydraulic fractures can enhance shale gas production to a certain extent. However, a larger number of fractures and higher fracture conductivity inevitably increase operational complexity and construction difficulty, leading to substantially higher stimulation costs. Therefore, to achieve increased production while controlling hydraulic fracturing investment, the net present value (NPV) is adopted in this study as the objective function for fracture parameter optimization. Under different assumptions, NPV can be expressed in various mathematical forms. In general, it consists of two main components: the discounted cash inflow and the capital expenditure (Luo et al., 2024).

The mathematical expression of NPV can be written as:

$$\text{NPV} = \sum_{i=1}^N \left[\frac{g_{\text{prod},i} S_{\text{gas}} - g_{\text{prod},i} S_{\text{cost}}}{(1+r)^i} \right] - \text{FC} - C_{\text{well}} - N_f C_f \quad (13)$$

where NPV denotes the net present value (CNY); N denotes the total production period (years); $g_{\text{prod},i}$ denotes the annual gas

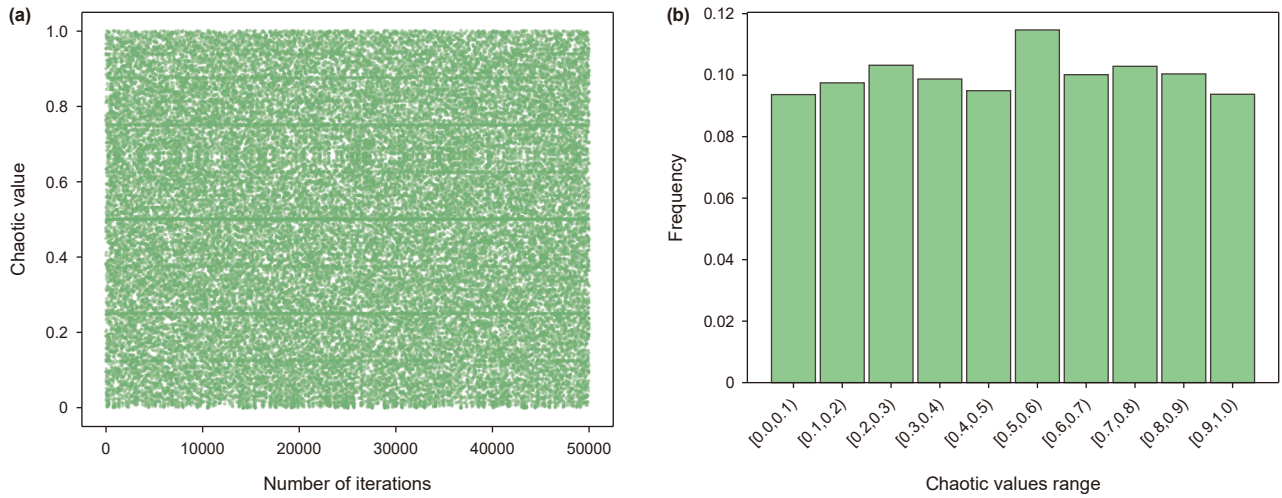


Fig. 11. Distribution characteristics of the Tent chaotic mapping. (a) Distribution of Tent chaotic mapping values. (b) Interval-wise distribution of Tent chaotic values.

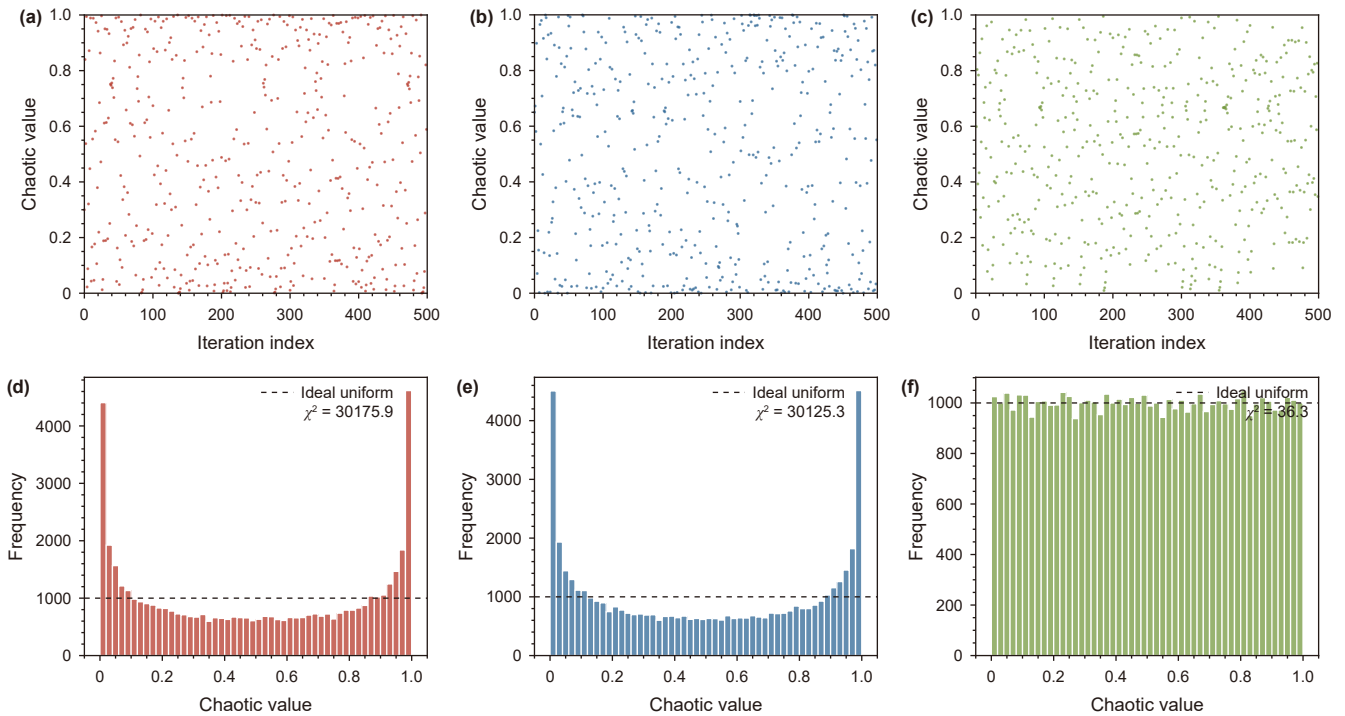


Fig. 12. Distribution characteristics of different chaotic mappings. (a)–(c) Chaotic value distributions of Logistic, Chebyshev, and Tent mappings, respectively. (d)–(f) Interval-wise frequency distributions of Logistic, Chebyshev, and Tent chaotic values, respectively.

production in the i -th year; $\$_{\text{gas}}$ denotes the unit gas price; $\$_{\text{cost}}$ denotes the unit operating cost per unit gas production; FC denotes the fixed cost per well; C_{well} denotes the drilling cost per well; N_f denotes the number of fracturing stages per well; C_f denotes the fracturing cost per fracture; r denotes the discount rate.

When maximizing the NPV, a set of physical and operational constraints must be satisfied. According to the current level of hydraulic fracturing technology and field operational feasibility, the feasible ranges of the number of fracturing stages (n_f), fracture half-length (L_f), and fracture conductivity (Con_f) are predefined. Therefore, the NPV maximization problem is solved subject to the following constraints:

$$\begin{cases} n_{f,\min} \leq n_f \leq n_{f,\max} \\ L_{f,\min} \leq L_f \leq L_{f,\max} \\ \text{Con}_{f,\min} \leq \text{Con}_f \leq \text{Con}_{f,\max} \end{cases} \quad (14)$$

where $n_{f,\min}$ and $n_{f,\max}$ denote the minimum and maximum numbers of fracturing stages, respectively; $L_{f,\min}$ and $L_{f,\max}$ denote the minimum and maximum fracture half-lengths, respectively; $\text{Con}_{f,\min}$ and $\text{Con}_{f,\max}$ denote the minimum and maximum fracture conductivities, respectively.

2.5. Implementation procedure of the improved TENT-SSA-BPNN model

The improved TENT-SSA-BPNN model employs the sparrow search algorithm (SSA) to optimize the initial weights and biases of the neural network, replacing conventional random initialization and effectively alleviating issues such as slow convergence and entrapment in local optima. Meanwhile, Tent chaotic mapping is introduced for population initialization in the SSA to enhance population diversity and further reduce the risk of premature convergence. In addition, a golden sine-based position update strategy is incorporated to improve the original exponential update mechanism of producers, thereby enhancing search efficiency and balancing global exploration and local exploitation. The overall implementation procedure of the improved TENT-SSA-BPNN model is illustrated in Fig. 13.

- (1) Data normalization. The input variables are normalized prior to model training to eliminate the influence of scale differences among features and to prevent slow convergence caused by inconsistent data magnitudes.
- (2) BPNN architecture design. The topological structure of the backpropagation neural network is determined, including the number of layers and neurons in each layer. An empirical formula is adopted to estimate the number of hidden-layer neurons, as given in Eq. (15).

$$p = \text{round}(\sqrt{q+1}) + s \quad (15)$$

where *round* denotes the rounding operator; *p* denotes the number of neurons in the hidden layer; *q* denotes the number of neurons in the input layer; *l* denotes the number of neurons in the output layer; *s* is an integer adjustment factor ranging from 1 to 10.

- (3) Initialization of the improved SSA. The parameters of the SSA are initialized, including the population size, maximum number of iterations, numbers of producers, scroungers, and scouts, as well as the warning (alert) threshold.
- (4) Population initialization and perturbation. The positions of the sparrow population are initialized using Tent chaotic mapping to enhance population diversity and reduce the risk of premature convergence. Individuals with the best fitness values are selected to form the initial population.
- (5) Fitness evaluation and position updating. The fitness of each sparrow is evaluated. Individuals ranked within the top 10%–20% in fitness are selected as producers, while 10%–20% of individuals are randomly chosen as scouts; the remaining individuals act as scroungers.

When updating producer positions, the golden sine strategy is incorporated, in which the movement step size is dynamically adjusted according to the golden section ratio to balance global exploration and local exploitation. All individuals update their positions adaptively based on the warning and safety thresholds, and individuals trapped in local optima are reinitialized.

- (6) Termination of SSA iterations. The optimal position and corresponding fitness are compared at each iteration. When the maximum number of iterations is reached, the algorithm terminates and outputs the global optimal solution; otherwise, the iteration continues.
- (7) Neural network training and prediction. The optimal solution obtained by the SSA is used as the initial weights and biases of the neural network. The BPNN is then fine-tuned using error backpropagation to improve prediction accuracy.

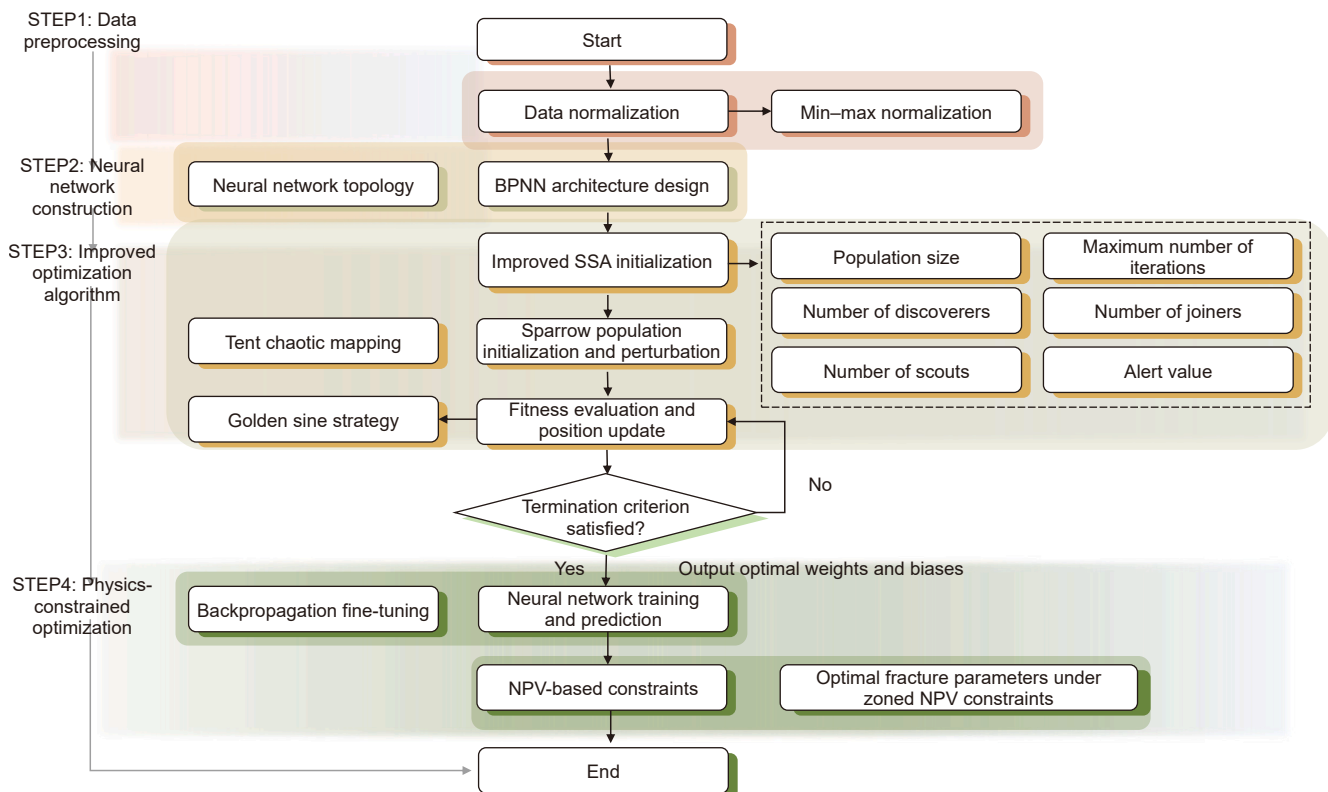


Fig. 13. Implementation flowchart of the improved TENT-SSA-BPNN model.

- (8) NPV-based optimization decision-making. Different combinations of fracture parameters are input into the trained model to predict production performance. The corresponding net present value (NPV) is calculated based on the predicted production. The fracture parameter combination that maximizes the NPV is identified as the optimal fracture design.

3. Results and discussion

3.1. Performance comparison of the improved algorithm

To verify the performance of the proposed improved algorithm, the classical Ackley function was employed to test both the original algorithm and the improved algorithm. The Ackley function is a highly nonlinear and multimodal benchmark function, characterized by a single global optimum and multiple local optima (Gharoie Ahangar et al., 2022). Owing to these properties, it is widely used to evaluate the global search capability and convergence speed of optimization algorithms. The mathematical expression of the Ackley function is given in Eq. (16) (Ihan 2016), and its three-dimensional landscape is illustrated in Fig. 14. The search domain of the function is defined as $[-32, 32]$. When $z_1, z_2 \in [-1.7904, 4.9727]$, the function attains its minimum value of 1.6505.

$$f(\mathbf{z}) = -a \cdot \exp\left(-b \cdot \sqrt{\frac{1}{d} \sum_{i=1}^d z_i^2}\right) - \exp\left(\frac{1}{d} \sum_{i=1}^d \cos(c \cdot z_i)\right) + a + \exp(1) \quad (16)$$

where $\mathbf{z}$ denotes the solution vector; z_i represents the i -th element of the vector $\mathbf{z}$; D denotes the number of decision variables; $a = 20$, $b = 0.2$, $c = 2\pi$.

Using identical parameter settings, both the conventional SSA and the improved SSA were applied to search for the minimum of the classical Ackley function within the domain $[-32, 32]$. The minimum function values obtained at different iteration numbers are presented in Fig. 15. In the conventional SSA, the population is initialized using a standard uniform random distribution, whereas the improved SSA generates the initial population based on Tent chaotic mapping. Owing to the ergodicity and pseudo-random characteristics of the Tent map, the initial population of the improved SSA exhibits superior spatial uniformity and diversity.

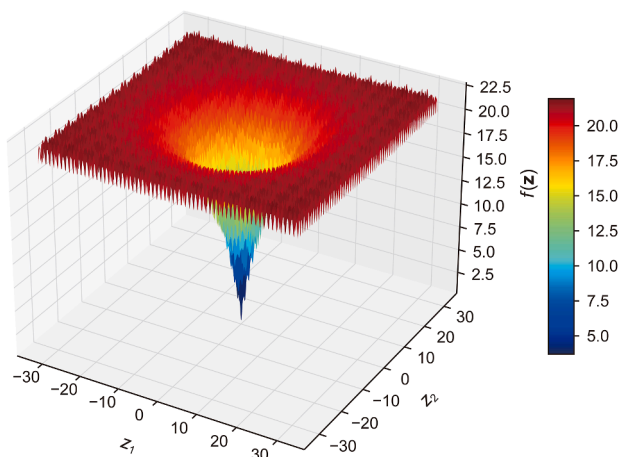


Fig. 14. Function values of the classical Ackley function with two decision variables.

This effectively avoids the population clustering phenomenon that may occur in conventional SSA due to random initialization, thereby providing higher-quality initial solutions.

As a result, the initial best solution obtained by the improved SSA is significantly better than that of the conventional SSA. The conventional SSA reaches the optimal value after approximately 20 iterations, whereas the improved SSA converges to the optimal solution within only 10 iterations. These results demonstrate that the improved algorithm can effectively avoid entrapment in local optima and substantially enhance convergence speed. Consequently, the improved SSA exhibits superior performance when solving complex, highly nonlinear optimization problems with no prior information.

To further validate the effectiveness of the improved sparrow search algorithm in optimizing the BP neural network, the proposed model was compared with the conventional BPNN, the SSA-BPNN, and the Tent-SSA-BPNN in terms of prediction performance. The dataset consisted of fracture and reservoir parameters, including fracture half-length, reservoir thickness, fracture conductivity, matrix permeability, matrix porosity, reservoir pressure, gas saturation, horizontal section length, number of fracturing stages, and three-year cumulative gas production. Five independent production prediction experiments were conducted, in which 80% of the data were used for training and the remaining 20% were used for testing.

The prediction performance of each model was evaluated using the root mean square error (RMSE), standard deviation (SD), mean absolute percentage error (MAPE), and the coefficient of determination (R^2), as shown in Figs. 16 and 17. The results indicate that the proposed model exhibits overall advantages in terms of RMSE, SD, MAPE, and R^2 . Although in certain individual cases the prediction performance of the proposed model is slightly inferior to that of the Tent-SSA-BPNN, for example, the performance of Case 4 in the Tent-SSA-BPNN is marginally better than that of Cases 2, 4, and 5 of the proposed model (Fig. 16), the overall average performance across the five cases of the proposed model is significantly superior to that of the other three models (Fig. 17).

3.2. Cross-validation of the neural network

Because the amount of complete field data available in oil and gas reservoirs is relatively limited, prediction results obtained from a single train-test split may be subject to randomness. To further evaluate the reliability of the proposed model, five-fold cross-validation was employed to assess its performance (Su et al., 2024). Specifically, the dataset was divided into five subsets of equal or nearly equal size. In each iteration, one subset was

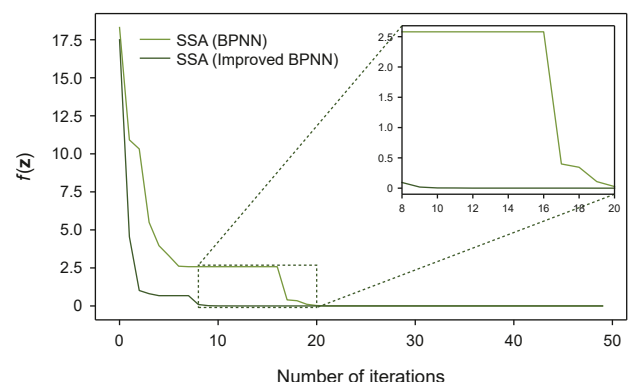


Fig. 15. Minimum function values at different iteration numbers.

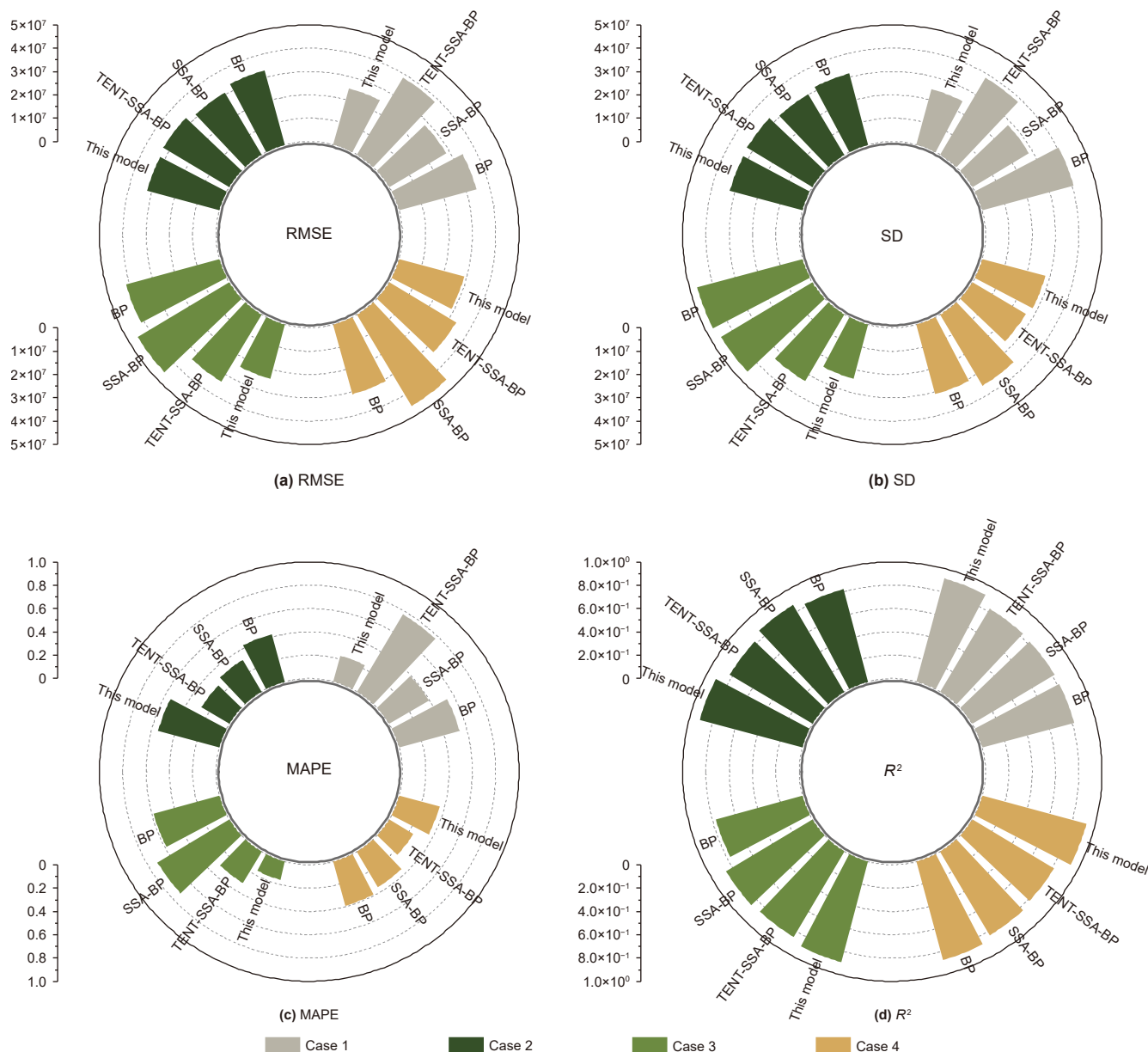


Fig. 16. Comparison of RMSE, SD, MAPE, and R^2 obtained from four prediction methods.

selected as the validation set, while the remaining four subsets were used for training. The model was trained on the training set and evaluated on the validation set. This procedure was repeated five times, with each subset serving once as the validation set.

Finally, the overall performance of the model was quantified by averaging the evaluation metrics obtained from the five validation rounds. This cross-validation strategy effectively reduces the risk of overfitting, enhances the generalization capability of the model, and provides more stable and reliable performance estimates, particularly for datasets with limited sample sizes.

As shown in Fig. 18, the training and prediction results of the optimal model demonstrate an overall prediction accuracy of 90.3%. The sample distribution approximately follows a normal distribution, indicating good data quality and reliable prediction

performance. The five-fold cross-validation results summarized in Table 3 show that the average RMSE, SD, and MAPE obtained for production prediction are 37090374, 28842064, and 31.55%, respectively. These results indicate that the prediction errors between the model outputs and the observed values remain relatively small across different training and testing datasets, confirming that the proposed model can stably and accurately predict production performance (see Fig. 19).

Moreover, the average coefficient of determination obtained from the five-fold cross-validation reaches $R^2 = 0.90$, which is close to unity. This suggests a strong correlation between the predicted and observed values, demonstrating that the model exhibits robust fitting capability and good generalization performance across different training and testing subsets.

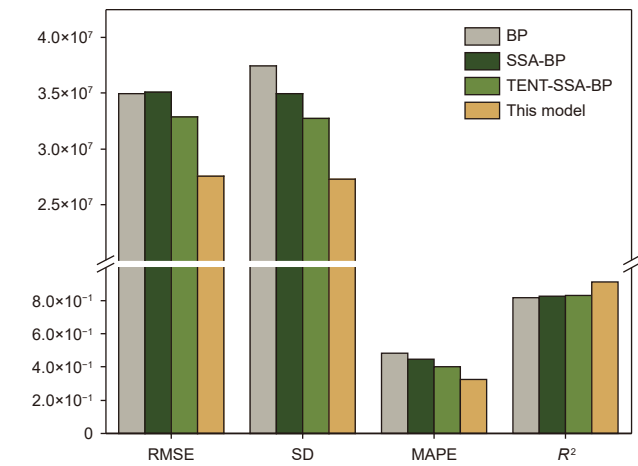


Fig. 17. Comparison of average RMSE, SD, MAPE, and R^2 over five runs for different prediction methods.

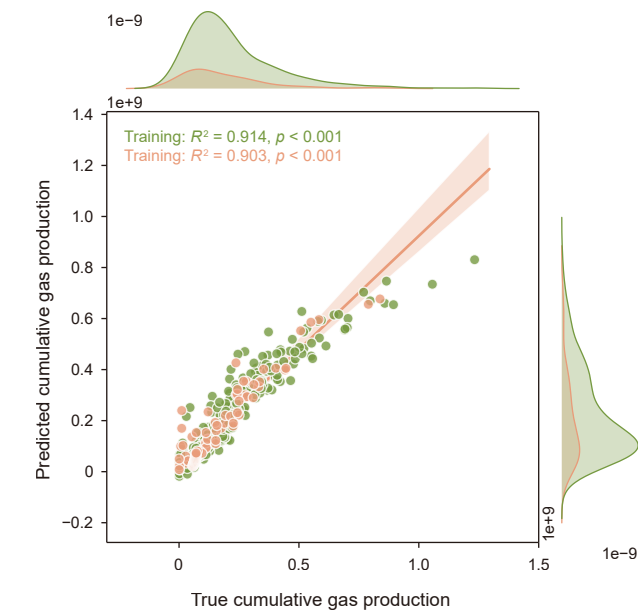


Fig. 18. Training and prediction results of the optimal model.

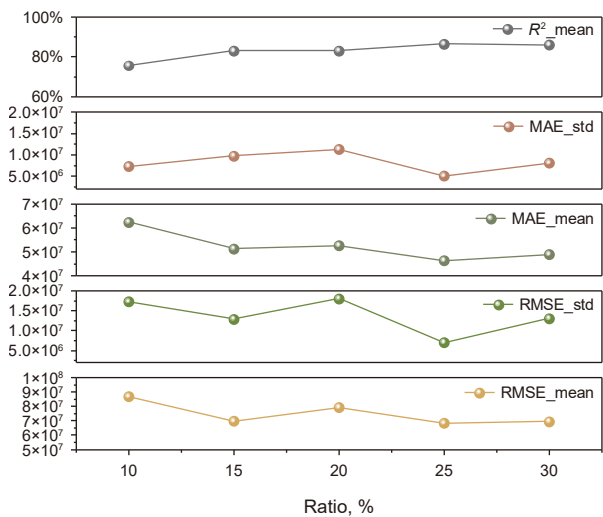


Fig. 19. Comparison of model transferability under different random subset extrapolation ratios.

3.3. Neural network transferability validation

The transferability of the proposed neural network was evaluated using a random subset extrapolation strategy. In this approach, extrapolated subsets (test sets) were constructed by randomly sampling 10%, 15%, 20%, 25%, and 30% of the entire dataset, while the remaining samples were used for model training. For each extrapolation ratio, multiple random splits were performed. In each split, data normalization, model training, and prediction were conducted independently, allowing for a robust assessment of model performance stability under distributional perturbations.

For each extrapolation ratio, the mean values of RMSE, MAE, and R^2 were computed over repeated experiments, and the corresponding standard deviations were used to quantify uncertainty. These statistics were employed to analyze the generalization behavior of the model as the size of the training dataset decreased and the test samples gradually deviated from the original data distribution.

The results of the random subset extrapolation indicate that, as the extrapolation ratio increased from 10% to 30%, the mean values of RMSE and MAE exhibited some fluctuations but remained within a comparable range, with no evident trend of systematic

Table 3
Results of five-fold cross-validation.

Fold	RMSE	SD	MAPE	R^2	Average RMSE	Average SD	Average MAPE	Average R^2
1	56708341	34857880	28.39%	0.90	37090374	28842064	31.55%	0.90
2	67090374	23765990	30.33%	0.89				
3	29398386	23430706	42.54%	0.90				
4	30208539	28835680	20.78%	0.91				
5	59747783	33320064	35.71%	0.90				

Table 4
Reservoir and wellbore parameters, and physical constraints imposed on the fracture parameters for optimization.

Reservoir thickness, m	Matrix permeability, mD	Matrix porosity	Reservoir pressure, MPa	Gas saturation	Horizontal section length, m	Fracture conductivity, mD-m	Number of fracturing stages	Fracture half-length, m
27	0.001	4.5%	35.4	0.77	800	10, 20, ..., 90, 100	10, 11, ..., 24, 25	100, 120, ..., 180, 200

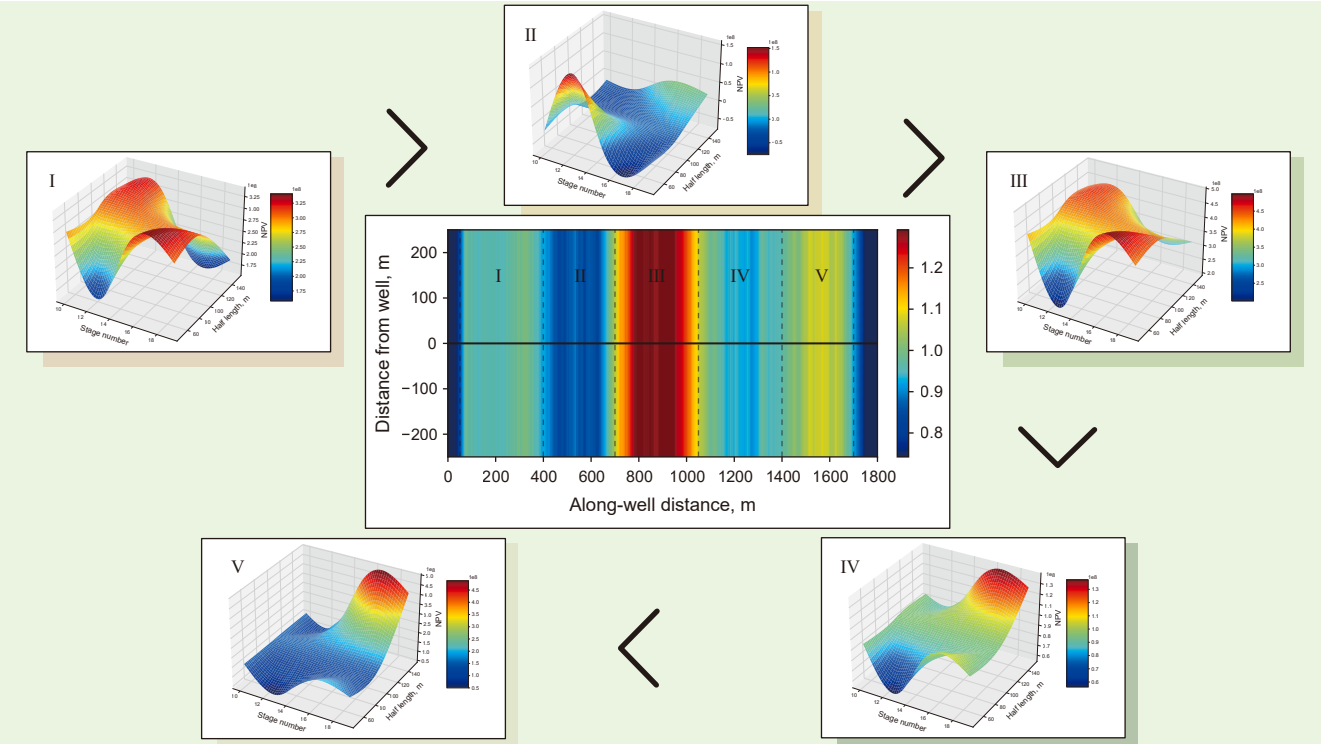


Fig. 20. Net present value (NPV) under different fracture parameter combinations.

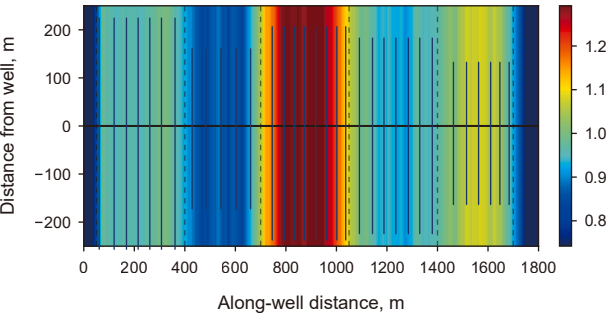


Fig. 21. Schematic illustration of fracture distribution in different reservoir zones.

degradation. Meanwhile, the relatively small standard deviations of these metrics demonstrate that the model performance is stable across different random splits and does not rely on specific training sample combinations. The R^2 values remained consistently high across all extrapolation ratios and did not decrease with increasing extrapolation proportion; in some cases, a slight improvement was even observed. This suggests that the key features learned by the model possess a certain degree of transferability and generalizability.

Overall, considering both the magnitude of error variations and the associated uncertainty ranges, the proposed neural network maintains robust and stable predictive performance under random distributional perturbations and moderate extrapolation conditions, indicating good generalization capability.

3.4. Neural network prediction results

The proposed precise segmentation method for horizontal wells in heterogeneous shale gas reservoirs, together with the

Table 5			
Optimal fracture parameters obtained for different reservoir zones.			
Zone	Fracture conductivity, mD-m	Number of fracturing stages	Fracture half-length, m
I	70	8	221
II	60	5	171
III	90	8	194.05
IV	70	7	207.2
V	80	6	163.4

improved TENT-SSA-BPNN model, was applied to optimize key fracture parameters of shale gas horizontal wells, including fracture conductivity, number of fracturing stages, and fracture half-length. The optimization objective was to maximize the net present value (NPV) for each segmented region while ensuring that the optimized fracture parameters satisfy practical engineering and physical constraints.

To achieve this objective, reservoir properties and wellbore parameters from different shale gas wells were collected and incorporated into a physics-constrained optimization framework. Based on these inputs, the trained neural network was used as a surrogate model to predict production performance under different fracture parameter combinations, and the corresponding NPV was evaluated. Table 4 summarizes the reservoir and wellbore parameters used for optimization, as well as the physical constraints imposed on the fracture parameters.

When calculating the NPV, the economic parameters were set as follows: the shale gas price was 1.275 CNY/m³ (based on the average wellhead price of shale gas in the Sichuan Basin during 2023–2025, as reported by PetroChina Southwest Oil & Gasfield Company), the operating cost was 0.32 CNY/m³ (determined from field operational expenditure records of representative shale gas

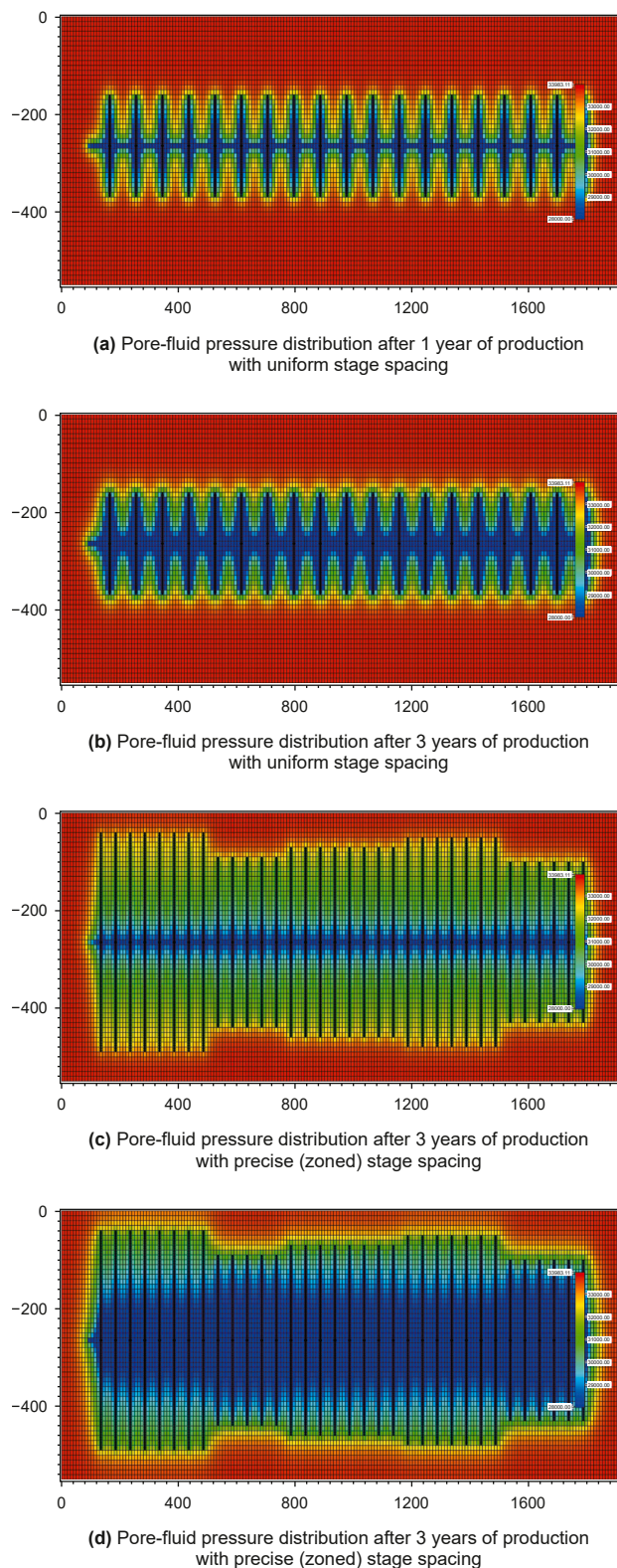


Fig. 22. Pore-fluid pressure distribution after a period of production.

wells in the study area), the fixed cost per well was 15 million CNY, and the drilling cost per well was 30 million CNY. The fracturing cost per fracture included the material costs of fracturing fluid and proppant as well as the fracturing service fee. The discount rate was set to 0.08, which is a commonly adopted value in the

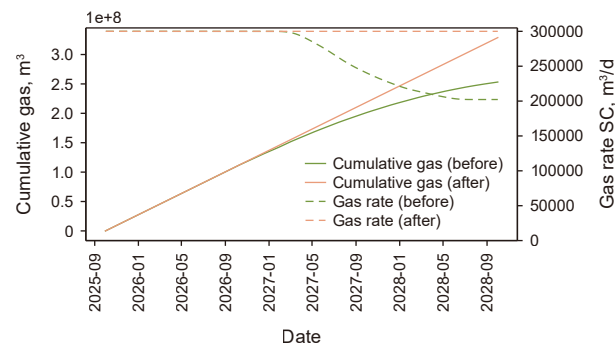


Fig. 23. Comparison of three-year cumulative gas production and daily gas production between conventional uniform fracture design and precise zoned fracture design.

economic evaluation of shale gas projects in China (National Development and Reform Commission guidelines). Using these parameters, the global maximum NPV was achieved at a fracture conductivity of 90 mD-m, 19 fracturing stages, and a fracture half-length of 100 m; therefore, this parameter set was identified as the optimal fracture design. As shown in Fig. 20, the maximum NPV for each zoned region was also determined, and the corresponding points were highlighted in red. The detailed numerical results are summarized in Table 5 (see Fig. 21).

Based on the final NPV prediction outcomes, the region-wise maximum NPV was obtained, and the corresponding fracture parameter combination was adopted to perform integrated optimization within each zone. As illustrated in the results, the proposed fracture placement design achieves a precise match to reservoir heterogeneity. Zone III, characterized by the highest reservoir quality index (RQI), was assigned a relatively higher fracture intensity (i.e., increased stage density and a more aggressive treatment design) to maximize stimulated reservoir volume and production response. Zones I and II, with comparatively poorer reservoir conditions, compensated for their inherent disadvantages by adopting reduced cluster spacing and enhanced fracture conductivity. In contrast, Zones IV and V employed a moderate-scale design to balance stimulation effectiveness and economic efficiency. Overall, this “zoning–optimization–matching” strategy enables differentiated stimulation designs across heterogeneous intervals, avoiding energy waste or insufficient reservoir contact caused by homogeneous fracture placement. Under combined economic and engineering constraints, the proposed approach effectively improves the development performance of the entire horizontal well.

3.5. Results of horizontal well fracture parameter optimization

For clarity, the baseline case used in the comparison (Fig. 22(a) and (b)) is defined as the conventional uniform fracture placement design, in which all fracturing stages are evenly distributed along the entire horizontal wellbore with identical fracture parameters: a fracture half-length of 180 m, fracture conductivity of 50 mD-m, and equal stage spacing. This uniform design represents the standard industry practice without any reservoir quality-based differentiation. In contrast, the optimized case (Fig. 22(c) and (d)) adopts the zone-specific fracture parameter combinations determined by the proposed TENT-SSA-BPNN optimization framework, as summarized in Table 5.

Based on the optimized fracture parameters obtained from the proposed method, the CMG commercial simulator (Zhang et al., 2025) was employed to simulate the reservoir pore-fluid pressure distributions after 1 year and 3 years of production under the

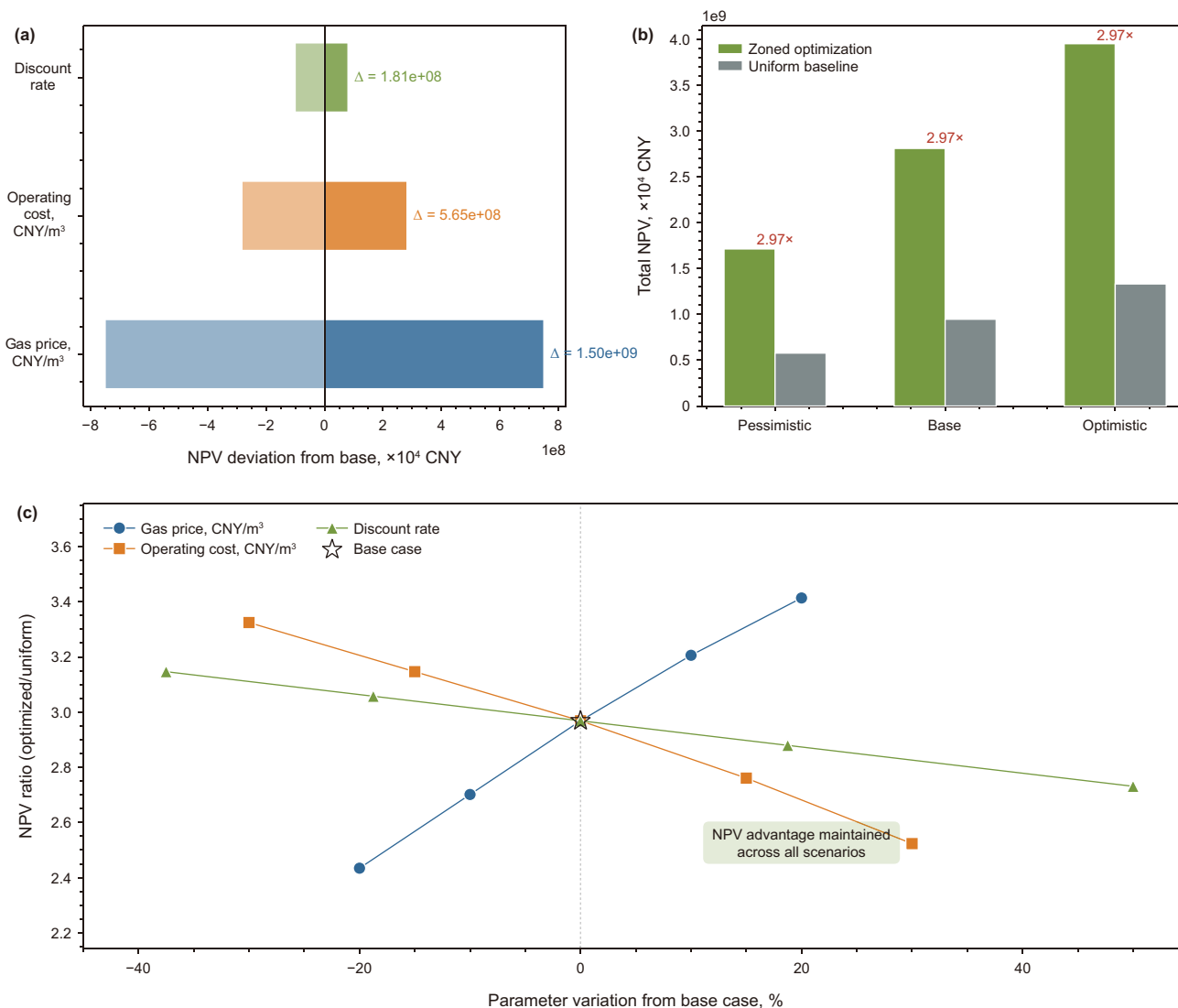


Fig. 24. NPV sensitivity analysis under varying economic parameters.

optimal fracture design, as shown in Fig. 22(a) and (b), respectively. The simulation results indicate that the pressure decline in the vicinity of fractures is pronounced, while the spatial distribution of pressure depletion remains relatively uniform. Moreover, the rate of pressure decline is well matched with the production behavior, and no excessively rapid pressure depletion leading to premature production decline is observed.

As illustrated in Fig. 22(c) and (d), after applying the proposed horizontal well fracture parameter optimization strategy, the stable production period is significantly prolonged, and no productivity exhaustion occurs during the simulated production time. The overall gas production increases by approximately 30% compared with the baseline case. In addition, the net present value (NPV) shows a substantial improvement; relative to uniform fracture placement, the NPV increases by approximately 3–5 times. These results demonstrate that the optimized fracture parameters effectively promote balanced shale gas production, significantly reducing residual gas caused by reservoir heterogeneity or inappropriate fracture design, and thereby enhancing the overall recovery efficiency of the horizontal well.

To evaluate the robustness of the reported NPV improvement under economic uncertainty, a sensitivity analysis was performed

by systematically varying the three key economic parameters in the NPV formulation. The gas price was varied within $\pm 20\%$ of the base value (1.02–1.53 CNY/m³), the operating cost was tested over a $\pm 30\%$ range (0.224–0.416 CNY/m³), and the discount rate was swept from 0.05 to 0.12. As shown in the tornado diagram (Fig. 24(a)), gas price exerts the strongest influence on absolute NPV, followed by operating cost, with the discount rate having a comparatively modest effect. The spider plot (Fig. 24(c)) further reveals that the NPV ratio between the optimized zoned design and the uniform baseline ranges from approximately 2.4 $\times$ to 3.4 $\times$ across all single-parameter variations. An increase in gas price amplifies the economic benefit of zone-specific fracture optimization, as the revenue differential between high-production and low-production zones becomes more pronounced. Conversely, rising operating costs compress the profit margin and moderately reduce the ratio, though the optimized design still retains a clear advantage. When all three parameters are simultaneously shifted to their most unfavorable levels (pessimistic scenario), the zoned design maintains an NPV approximately 2.97 times that of the uniform case (Fig. 24(b)), confirming that the relative superiority of the proposed approach is not an artifact of favorable economic assumptions. These results demonstrate that while the absolute

NPV is expectedly sensitive to market conditions, the comparative advantage of reservoir-quality-driven fracture optimization over conventional uniform placement is robust across a wide range of plausible economic scenarios.

4. Conclusions

To address the strong heterogeneity of shale gas reservoirs, this study proposes a physics–data collaborative intelligent optimization method based on fine reservoir segmentation to improve the accuracy and efficiency of fracture parameter optimization for horizontal wells. By integrating Tent chaotic mapping, the golden sine strategy, and the sparrow search algorithm (SSA), an improved TENT-SSA-BPNN model was developed and coupled with a reservoir quality index (RQI)–based segmentation scheme to achieve precise characterization of near-wellbore reservoir heterogeneity. With the net present value (NPV) as the ultimate optimization objective and under explicit physical constraints, a differentiated “segmentation–optimization–matching” design strategy was established, providing an innovative and practical solution for the development of heterogeneous shale reservoirs. The main conclusions are summarized as follows.

- (1) A fine segmentation method for horizontal wells based on the reservoir quality index was proposed, effectively addressing development challenges caused by strong reservoir heterogeneity. By quantifying key geological and engineering controlling factors and applying a dynamic quantile-based segmentation approach, the horizontal wellbore was divided into multiple reservoir quality zones (Zones I–V), replacing the conventional uniform stage spacing strategy. This segmentation method accurately captures spatial heterogeneity around the wellbore and provides a solid physical foundation for implementing zone-specific stimulation strategies.
- (2) A physics–data collaborative optimization framework based on the improved TENT-SSA-BPNN model was constructed. Tent chaotic mapping was employed to enhance population diversity, while the golden sine strategy was introduced to balance global exploration and local exploitation, significantly improving the global optimization capability in high-dimensional nonlinear search spaces. Compared with traditional numerical simulation methods and single machine learning models, the proposed framework substantially improves computational efficiency, effectively avoids local optima, and achieves high-accuracy prediction of complex nonlinear relationships between fracture parameters and production performance.
- (3) In terms of variable selection and feature engineering, Pearson correlation analysis and projection importance indicators were applied to quantitatively identify the dominant factors influencing production. By constructing a comprehensive reservoir quality index, redundant variables were eliminated, and the stability and predictive accuracy of reservoir characterization were significantly enhanced, thereby ensuring the robustness and reliability of the fine segmentation results.
- (4) A differentiated fracture parameter optimization strategy based on fine segmentation and NPV maximization was successfully implemented. For reservoir zones with different quality levels, optimal combinations of fracture conductivity, number of fracturing stages, and fracture half-length were matched accordingly. The results demonstrate that this heterogeneity-adaptive fracture design significantly outperforms uniform fracture placement in terms of

economic benefits and ultimate recovery, providing a highly innovative and practically valuable optimization scheme for efficient shale gas development.

CRediT authorship contribution statement

Nan-Lin Zhang: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xin-Qi Zeng:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Zhi-Feng Luo:** Software, Data curation. **Jun-Wen Tao:** Data curation. **Ping-Li Liu:** Validation, Supervision. **Mao-Han Yuan:** Visualization, Data curation, Conceptualization. **Qi Zhu:** Supervision. **Fu-Shen Liu:** Funding acquisition, Data curation.

Declaration of competing interest

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

References

- Chen, L., Hao, Y., Li, Q., et al., 2024. Traffic volume forecast model based on BP neural network optimized by improved sparrow search algorithm. *Harbin Gongye Daxue Xuebao/Journal of Harbin Institute of Technology* 56 (7), 94–101. <https://doi.org/10.11918/202207077>.
- Chen, J., Shi, K., Jiang, F., et al., 2026. Wettability characteristics of continental shale and machine learning-based prediction of shale wettability. *Fuel* 404 (Part C), 136356. <https://doi.org/10.1016/j.fuel.2025.136356>.
- Cheng, M., Jia, G., Fang, W., et al., 2022. SSA and BPNN based efficient situation prediction model for cyber security. In: 2022 18th International Conference on Mobility, Sensing and Networking (MSN). IEEE, pp. 809–813. <https://doi.org/10.1109/msn57253.2022.00131>.
- Chu, H.-Y., Fang, S.-D., Ren, Z.-Q., et al., 2025. An integrated multi-scale approach for transient analysis of multi-well horizontal pad with well interference in shale gas reservoirs: methodology and case study. *Pet. Sci.* 22 (3), 1155–1170. <https://doi.org/10.1016/j.petsci.2024.12.022>.
- Dong, Y., Hao, Y., Lu, D., 2025. A hybrid neural network model for shale gas production prediction considering production plans and produced water. *Eng. Appl. Artif. Intell.* 162 (Part C), 112543. <https://doi.org/10.1016/j.engappai.2025.112543>.
- Du, H., Xing, H., Feng, J., et al., 2025. Damage-integrated fracture prediction model for multi-stage fracture in deep carbonates: Application to the A Oilfield, Northern Tarim Basin. *Mar. Petrol. Geol.* 182, 107599. <https://doi.org/10.1016/j.marpetgeo.2025.107599>.
- Fang, L., Deng, Q., Yang, D., 2025. Integration of multi-domain data analysis and machine learning for wellbore flow regimes identification in shale gas reservoirs. *Energy* 332, 137123. <https://doi.org/10.1016/j.energy.2025.137123>.
- Gao, H., Cheng, P., Wu, W., et al., 2025. Gas-in-place (GIP) model and resource potential of the ultra-deep Lower Cambrian gas-rich shales in the southern Sichuan Basin, China. *Mar. Petrol. Geol.* 182, 107606. <https://doi.org/10.1016/j.marpetgeo.2025.107606>.
- Gharoie Ahangar, R., Pavur, R., Gharavi, H., 2022. A global optima search field division method for evolutionary algorithms. *J. Oper. Res. Soc.* 73 (5), 1085–1104. <https://doi.org/10.1080/01605682.2021.1890531>.
- Gou, L., Yang, Z., Min, C., et al., 2024. A novel domain adaptation method with physical constraints for shale gas production forecasting. *Appl. Energy* 371, 123673. <https://doi.org/10.1016/j.apenergy.2024.123673>.
- Guo, T., Wang, H., Chen, M., et al., 2025. Numerical simulation of fracture propagation in deflagration-hydraulic composite fracturing of unconventional reservoirs. *Petrol. Explor. Dev.* 52 (4), 1017–1028. [https://doi.org/10.1016/S1876-3804\(25\)60619-7](https://doi.org/10.1016/S1876-3804(25)60619-7).
- He, Y., Qiu, S., Qin, J., et al., 2025. Feasibility of CO₂–water alternate flooding and CO₂ storage in tight oil reservoirs with complex fracture networks based on embedded discrete fracture model. *Energy* 319, 135065. <https://doi.org/10.1016/j.energy.2025.135065>.
- Huang, L., Zhan, W., Zhao, H., et al., 2025. A novel fluid-solid coupling method for fractured reservoirs: 3D DDM-EDFM integration with proppant mechanics. *Comput. Geotech.* 181, 107127. <https://doi.org/10.1016/j.compgeo.2025.107127>.

- Ilhan, I., 2016. Mobile device based test tool for optimization algorithms. *Comput. Appl. Eng. Educ.* 24 (5), 744–754. <https://doi.org/10.1002/cae.21747>.
- Ji, D., Feng, Y., Deng, J., 2026. Modeling and parametric analysis of immobile LCM bridges for wellbore strengthening in fractured formations. *Geoenergy Sci. Eng.* 257 (Part A), 214248. <https://doi.org/10.1016/j.geoen.2025.214248>.
- Jiang, S., Qian, Y., Xu, P., et al., 2026. In-Situ CO₂ sequestration and shale gas recovery mechanisms in geologically representative organic-rich clay reservoirs. *Chem. Eng. Sci.* 321 (Part B), 122923. <https://doi.org/10.1016/j.ces.2025.122923>.
- Jin, Z.-H., 2025. Stresses in a packed fracture-filling porous bed with application in wellbore strengthening. *Eng. Fract. Mech.* 330, 111676. <https://doi.org/10.1016/j.engfractmech.2025.111676>.
- Liu, C., Liu, P., Wang, Q., et al., 2025. Stratified allocation method for water injection based on machine learning: A case study of the Bohai A oil and gas field. *Nat. Gas Ind. B* 12 (2), 207–218. <https://doi.org/10.1016/j.ngib.2025.03.005>.
- Luo, S., Zhao, Y., Xiao, H., et al., 2024. Fracture parameter optimization of tight sandstone gas reservoirs based on the hybrid optimization algorithm and deep neural network model. *Nat. Gas. Ind.* 44 (9), 140–151. <https://doi.org/10.3787/j.issn.1000-0976.2024.09.013> (in Chinese).
- Nan, T., Hu, T., Wang, Z., et al., 2025. Modeling hydro-thermal processes in fractured geothermal reservoirs using embedded discrete fracture model (EDFM) and MRST. *Adv. Water Resour.* 206, 105120. <https://doi.org/10.1016/j.advwatres.2025.105120>.
- Niu, W., Lu, J., Sun, Y., et al., 2024. Techno-economic integration evaluation in shale gas development based on ensemble learning. *Appl. Energy* 357, 122486. <https://doi.org/10.1016/j.apenergy.2023.122486>.
- Niu, W., Sun, Y., Zhang, M., et al., 2026. Unsupervised learning-driven insights into shale gas Reservoirs: Production prediction and strategic applications. *Geoenergy Sci. Eng.* 256, 214170. <https://doi.org/10.1016/j.geoen.2025.214170>.
- Qin, Z., Wu, J., Wang, C., et al., 2023. A novel calculation model, characteristics and applications of Archie's cementation exponent in dual porosity reservoirs with intersecting dual fractures. *Geoenergy Sci. Eng.* 231 (Part A), 212390. <https://doi.org/10.1016/j.geoen.2023.212390>.
- Ren, L., Wang, Z., Zhao, J., et al., 2023. Shale gas effective fracture network volume prediction and analysis based on flow back data: A case study of southern Sichuan Basin shale. *Geoenergy Sci. Eng.* 228, 211963. <https://doi.org/10.1016/j.geoen.2023.211963>.
- Su, Z.-L., Yan, W.-L., Li, H.-M., et al., 2024. Prediction of agricultural carbon emission based on improved BP neural network with optimized sparrow search algorithm. *Environ. Sci. J. Integr. Environ. Res.* 45 (12), 6818–6827. <https://doi.org/10.13227/j.hjkk.202401258> (in Chinese).
- Sun, R., Huang, D., Yao, X., et al., 2026. Effect of dynamic expansion of interface micro-annulus on casing failure mechanism during multistage fracturing. *Geoenergy Sci. Eng.* 257 (Part A), 214227. <https://doi.org/10.1016/j.geoen.2025.214227>.
- Suo, Y., Wei, X.-H., Zhao, Y.-J., et al., 2026. Simulation analysis of hydraulic fracture initiation and propagation mechanisms under mixed-mode and reservoir fracturing. *Pet. Sci.* 23 (1), 220–235. <https://doi.org/10.1016/j.petsci.2025.09.025>.
- Tang, J., Li, J., Zhang, Z., et al., 2026. Differential impacts of multi-scale natural fractures on hydraulic fracture network formation. *Earth Sci. Rev.* 272, 105315. <https://doi.org/10.1016/j.earscirev.2025.105315>.
- Tanyildizi, E., Demir, G., 2017. Golden sine algorithm: A novel math-inspired algorithm. *Adv. Electr. Comput. Eng.* 17 (2), 71–78. <https://doi.org/10.4316/aeece.2017.02010>.
- Wang, Z., Dai, H., Chen, B., et al., 2024. Effluent quality prediction of the sewage treatment based on a hybrid neural network model: Comparison and application. *J. Environ. Manag.* 351, 119900. <https://doi.org/10.1016/j.jenvman.2023.119900>.
- Wei, C., Li, H., Zhu, X., et al., 2025. Fracture parameter diagnostic method during staged multi-cluster fracturing based on distributed temperature sensing. *Petrol. Explor. Dev.* 52 (2), 496–505. [https://doi.org/10.1016/S1876-3804\(25\)60582-9](https://doi.org/10.1016/S1876-3804(25)60582-9).
- Xu, Y., Cavalcante Filho, J.S.A., Yu, W., et al., 2017. Discrete-fracture modeling of complex hydraulic-fracture geometries in reservoir simulators. *SPE Reservoir Eval. Eng.* 20 (2), 403–422. <https://doi.org/10.2118/183647-PA>.
- Xu, B., Hu, J., Hu, T., et al., 2021. Quantitative assessment of seismic risk in hydraulic fracturing areas based on rough set and Bayesian network: A case analysis of changning shale gas development block in Yibin City, Sichuan Province, China. *J. Petrol. Sci. Eng.* 200, 108226. <https://doi.org/10.1016/j.petrol.2020.108226>.
- Xue, J., Shen, B., 2020. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Systems Science and Control Engineering* 8 (1), 22–34. <https://doi.org/10.1080/21642583.2019.1708830>.
- Yao, Y., Wang, L., Wang, K., et al., 2023. A novel high-dimension shale gas reservoir hydraulic fracture network parameters optimization framework. *Geoenergy Sci. Eng.* 229, 212155. <https://doi.org/10.1016/j.geoen.2023.212155>.
- Yin, Z., Zhang, F., Wang, X., et al., 2025. Four-dimensional stress induced by hydraulic fracturing and long-term extraction for shale gas well platforms: Implications for refracturing design. *J. Rock Mech. Geotech. Eng.* <https://doi.org/10.1016/j.jrmge.2025.07.025>.
- Yu, W., Zhu, K., Yu, Y., 2024. Variable universe fuzzy PID control for active suspension system with combination of chaotic particle swarm optimization and road recognition. *IEEE Access* 12, 29113–29125. <https://doi.org/10.1109/ACCESS.2024.3368762>.
- Zhang, Q., Hou, B., Lin, B., et al., 2021. Integration of discrete fracture reconstruction and dual porosity/dual permeability models for gas production analysis in a deformable fractured shale reservoir. *J. Nat. Gas Sci. Eng.* 93, 104028. <https://doi.org/10.1016/j.jngse.2021.104028>.
- Zhang, N., Liu, F., Jiang, L., et al., 2024. Effects of phase-transition heat on fracture temperature in self-propping phase-transition fracturing technology. *Petrol. Explor. Dev.* 51 (6), 1587–1597. [https://doi.org/10.1016/S1876-3804\(25\)60562-3](https://doi.org/10.1016/S1876-3804(25)60562-3).
- Zhang, N., Cao, B., Liu, F., et al., 2025. CO₂ flooding effects and breakthrough times in low-permeability reservoirs with injection-production well patterns containing hydraulic fractures. *Nat. Gas Ind. B* 12 (3), 339–355. <https://doi.org/10.1016/j.ngib.2025.05.007>.
- Zhao, Y., Yang, H., Wu, J., et al., 2022. Choke management simulation for shale gas reservoirs with complex natural fractures using EDFM. *J. Nat. Gas Sci. Eng.* 107, 104801. <https://doi.org/10.1016/j.jngse.2022.104801>.
- Zhao, Y., Gan, P., Li, Y., et al., 2025. Simulation methods for shale gas adsorption: Fundamentals and applications at the molecular and mesoscopic scales. *Comput. Mater. Sci.* 259, 114198. <https://doi.org/10.1016/j.commatsci.2025.114198>.