A New Production Forecasting Method of the Multifractured Horizontal Wells Based on Cluster Analysis

Mingjing Lu and Zenglin Wang

1Petroleum Engineering Technology Research Institute of Shengli Oilfield, Sinopec, No. 306 West Road, Dongying, China 257067
2Postdoctoral Scientific Research Working Station of Shengli Oilfield, Sinopec, No. 306 West Road, Dongying, China 257067

Correspondence should be addressed to Mingjing Lu; lumingjing001@126.com

Received 27 November 2020; Accepted 24 September 2021; Published 27 October 2021

The seepage mechanism of multifractured horizontal wells is complex in tight reservoirs, which make that the production is very difficult to forecast. This article put forward a new way called the developed clustering analysis to forecast well production which is based on the practical production data of 10 multifractured horizontal wells. This method first uses the information analysis method to obtain the weight of the influencing factors of horizontal well production and normalizes the influencing factors of production. Second, the feature matrix is constructed by combining the weight of each factor, and the distance between the feature matrix of different production wells and the optimal feature matrix is calculated. Finally, the relationship curve between distance and production is plotted, and the production chart of the block is obtained. Taking 9 multifractured horizontal wells in the tight reservoir as an example, the production prediction chart of the block is calculated. At the same time, the production data of the 10th well are used to verify the production chart of the block. The results show that the horizontal well production has a high fitting relationship with the distance. The error between the new well production predicted by the chart and the actual production is 4.7%, which meets the requirements of the field error. The model was also used in a reservoir with 154 wells and also verified the accuracy of the model. The prediction method proposed in this paper can accurately predict the production of volume fractured horizontal wells in the experimental area and provide certain guiding significance for the development and adjustment of tight reservoirs.

1. Introduction

In recent years, with the successful exploitation of tight oil reservoirs such as Bakken shales in the United States, the development of tight oil reservoirs in China has gradually embarked on the right track on the basis of drawing on its mining experience [1, 2]. Tight oil reservoir is a typical unconventional oil and gas resources. It is necessary to implement volume fracturing technology for tight oil reservoir to reach the economic production. The reservoir structure after fracturing becomes very complex, and the seepage characteristics become more complex [3, 4]. Therefore, it is very important to establish a new method to predict production.

The dynamic prediction of oilfield development is mainly divided into two parts: the analysis of the characteristics of the dynamic change of development indexes and the prediction model of oil well productivity. Many scholars have done a lot of research on the seepage theory of fractured oil and gas reservoirs, and the seepage mode is from Darcy seepage to non-Darcy seepage. The dynamic analysis methods for studying the productivity and pressure of fractured reservoirs mainly include numerical simulation method and well test analysis method [5–7]. Numerical simulation methods are mainly continuous medium model and discrete fracture model [8]. The dual-medium model is one of the most widely used mathematical models for fluid flow coupling. Warren and Root established the ideal dual-medium model and studied the geometric characteristics and seepage process of matrix-fracture dual-medium reservoir [9]. However, the Warren-Root dual-medium model cannot consider the unsteady flow of fractured reservoirs.
In view of the problem of dual-media model, scholars proposed the discrete fracture model (DFM) for description [11, 12]. Compared with the dual medium model, DFM can describe the geometric characteristics of cracks more accurately and has higher calculation accuracy [13]. Well test analysis can effectively obtain the complex seepage law of unconventional fractured reservoir. At present, the well test interpretation model of fractured horizontal well is mainly established by point source function method, elliptical seepage theory, and linear seepage theory [14–17]. The point source analysis method was proposed by Lefkovits [14], and the point source function method was used to analyze the variation law of bottom hole production and bottom hole pressure in each layer. The point source function method can easily solve the combined reservoir problems with different boundaries, but it has great limitations for large hydraulic fracturing reservoirs [15]. Elliptic seepage theory can consider the fluid flow characteristics around complex fractures and can be used to solve the well test problem of fractured wells in large-scale hydraulic fracturing. Prats et al. [16] found that the fluid flow around the fracture was elliptical when studying the seepage law of contaminated wells after fracturing, and the elliptical equation was used to describe the fluid seepage. Well test analysis method based on elliptical seepage theory can effectively analyze the fluid flow characteristics of fractured reservoirs, but the interfracture interference cannot be considered. Subsequently, El-Banbi and Wattenburger [17] considered that linear flow was dominant in tight oil and gas reservoirs, so a multilinear seepage equation was established for well test analysis of tight oil and gas reservoirs. This method is suitable for dynamic analysis of production and pressure of horizontal wells in tight reservoirs after large-scale fracturing.

At present, the common methods for predicting horizontal well productivity include mathematical model, numerical simulation method, and empirical formula method. Mathematical model can directly express the influence of various factors of production [18]. However, it is established based on assumptions, and the formula is difficult to solve. Numerical simulation can consider many different factors and establish different models for different reservoir characteristics. Peng et al. [19] established a numerical simulation model of fine fractured horizontal well according to the geological characteristics of tight reservoir and studied the sensitivity of influencing factors of multistage fracturing productivity of horizontal well. However, for reservoirs with different characteristics, modeling takes a long time, and the prediction accuracy is affected by the model fineness. The empirical formula method is based on the existing production data, which is convenient to use. Liu et al. [20] believed that most of the oilfield production data showed that most of the decline conformed to the hyperbolic decline Arps formula, and the applicable conditions of Arps formula were relatively harsh, which required that the bottom hole pressure remained unchanged, had a constant pressure boundary, and produced with a fixed production capacity. Duan et al. [21] based on the Arps production decline method cannot be applied to the productivity prediction of early production data of gas wells. They used the power law exponential decline model to improve the deficiency that Arps can only predict the stable flow state and analyzed the production data of a shale gas fracturing horizontal well in Sichuan Basin based on the power law exponential decline method. However, due to the existence of linear flow, the conventional Arps method is not suitable for tight reservoirs [22–24]. The proposed empirical formula is more complex for tight oil reservoirs, and it has limitations due to considering the flow stage.

Some scholars proposed the application of cluster analysis to predict production of tight oil wells to avoid the complexity of the seepage mechanism in tight oil reservoirs [25]. However, this method can only classify wells based on existing data of production wells and obtain the prediction range, which cannot accurately predict single well. In this paper, we proposed an improved cluster analysis method based on previous research. This method can consider the characteristics of tight reservoirs and the characteristics of segmented multifractured horizontal wells. For single well prediction, it is not only simple to use but also the prediction error is within a reasonable range.

## 2. Improved Cluster Analysis Method to Predict Production

In cluster analysis, the affinity between samples is mainly measured by the distance between samples. There are \( K \) numbers in the sample, and the sample can be regarded as a point in the \( K \)-dimensional space. The distance between samples is the distance between \( K \)-dimensional space points. Using cluster analysis to predict productivity is to classify new wells into one category according to existing data and predict the productivity interval of new wells. The productivity prediction based on cluster analysis is easy to use. However, it can only predict wells with similar conditions to existing production wells. If the new wells are not in the classification interval, the production cannot be predicted. For new wells that can be predicted, the predicted productivity value is an interval value, and the size of the error depends on the classification of the old well. In view of the above deficiencies, this paper improves the clustering analysis method.

In this paper, according to the idea of cluster analysis, the optimal value of horizontal well parameters as the cluster center. According to the distance between the geological parameters of each horizontal well and the transformation parameters to the cluster center, the production wells are evaluated, and the relationship between the total distance of all parameters and productivity is obtained. The processing flow of this method is shown in Figure 1. It is necessary to normalize the data and consider the influence of the weight of each parameter in the process of calculation. The improved clustering method can be used to predict the productivity of new wells under any conditions, and the predicted result is a constant value with relatively small errors.

### 2.1. The Weight of Production Capacity Influence Factors

The information analysis method is based on the actual dynamic data to analyze the relative size of the influencing factors of production capacity [26]. This method has strong
consistency with gray correlation and orthogonal experimental design methods, and the calculation results are targeted for specific blocks and easy to apply. The amount of information represents the impact of parameters on productivity. The greater the amount of information, the greater the impact. According to the amount of information, a weight analysis is performed on the factors affecting productivity.

To study the actual data of a block and calculate the amount of information, a weight analysis is performed on the factors affecting productivity. The greater the amount of information, the greater the impact. According to the amount of information, a weight analysis is performed on the factors affecting productivity.

To study the actual data of a block and calculate the amount of information, a weight analysis is performed on the factors affecting productivity. The greater the amount of information, the greater the impact. According to the amount of information, a weight analysis is performed on the factors affecting productivity.

The steps for calculating the amount of information for each factor are as follows:

1. To convert the frequencies mapped in group A and group B into probability frequencies \( y_{A\delta} \) and \( y_{B\delta} \), \( \delta \) is the interval number

2. To calculate the average probability frequency of each interval \( \bar{y}_\delta \)

\[
\bar{y}_\delta = 0.1(y_{\delta-2} + 2y_{\delta-1} + 4y_\delta + 2y_{\delta+1} + y_{\delta+2}).
\] (1)

3. To calculate the average frequency ratio \( \bar{y}_{A\delta}/\bar{y}_{B\delta} \)

4. To calculate diagnostic coefficient \( Z_\delta \)

\[
Z_\delta = 10 \log \left( \frac{\bar{y}_{A\delta}}{\bar{y}_{B\delta}} \right).
\] (2)

5. To calculate the information amount of the parameter in each change interval \( I_\delta \)

\[
I_\delta = 0.5Z_\delta(\bar{y}_{A\delta} - \bar{y}_{B\delta}).
\] (3)

6. To calculate the total amount of information \( I \)

\[
I = \sum I_\delta.
\] (4)

7. To calculate the weight of the capacity influencing factors according to the amount of information

\[
\text{Weight} = \frac{I_\delta}{\sum I_\delta}.
\] (5)

2.2. Parameter Matrix Normalization. In order to eliminate the influence of different dimensions between data, the original data matrix composed of the main factors affecting the productivity of horizontal wells in the target area is normalized to make the original data distributed in the interval \([0,1]\). There are many ways to process data to make it normalized. In this paper, different parameters are divided by the maximum value \( X_i \) to achieve normalization.

\[
x = \left[ \begin{array}{cccc}
    x_{11} & x_{12} & x_{13} & \cdots & x_{1k} \\
    x_{21} & x_{22} & x_{23} & \cdots & x_{2k} \\
      \cdots & \cdots & \cdots & \cdots & \cdots \\
    x_{N1} & x_{N2} & x_{N3} & \cdots & x_{Nk}
\end{array} \right].
\] (6)

2.3. Combine Weights to Construct Feature Matrix. The weights of influencing factors of production capacity are \( a_1, a_2, \cdots, a_k \). The normalized parameter matrix and weight are multiplied to obtain the feature matrix.

\[
x' = \left[ \begin{array}{cccc}
    x_{11}a_1 & x_{12}a_2 & x_{13}a_3 & \cdots & x_{1k}a_k \\
    x_{21}a_1 & x_{22}a_2 & x_{23}a_3 & \cdots & x_{2k}a_k \\
      \cdots & \cdots & \cdots & \cdots & \cdots \\
    x_{N1}a_1 & x_{N2}a_2 & x_{N3}a_3 & \cdots & x_{Nk}a_k
\end{array} \right].
\] (7)

To analyze the influence law of each parameter standard of production capacity, we can obtain the optimal value of each column of parameters to form an optimal value matrix, as:

\[
x_{opt} = [X_{1opt} \ X_{2opt} \ X_{3opt} \ \cdots \ X_{kopt}].
\] (8)

Normalized optimal value matrix, as:

\[
x_{opt} = [x_{1opt} \ x_{2opt} \ x_{3opt} \ \cdots \ x_{kopt}].
\] (9)
Multiply the normalized optimal value matrix with the weight to obtain the optimal feature matrix, as:

\[ x'_{\text{opt}} = \begin{bmatrix} x_{1\text{opt}} \cdot a_1 & x_{2\text{opt}} \cdot a_2 & \cdots & x_{k\text{opt}} \cdot a_k \end{bmatrix}. \]  

(10)

### 2.4. Solve Feature Distance

Calculate the distance from each parameter of each well to the optimal value to form a distance matrix:

\[
\begin{bmatrix}
  d_{11} & d_{12} & d_{13} & \cdots & d_{1k} \\
  d_{21} & d_{22} & d_{23} & \cdots & d_{2k} \\
  \cdots & \cdots & \cdots & \cdots & \cdots \\
  d_{N1} & d_{N2} & d_{N3} & \cdots & d_{Nk}
\end{bmatrix},
\]

(11)

where

\[ d_{ij} = x_i - x_{\text{opt}} \cdot a_j, \quad i = 1, 2, \ldots, N; \quad j = 1, 2, \ldots, K. \]

(12)

The sum of the distance between the \(i^{th}\) well and the optimal value is

\[ d_i = |d_{i1}| + |d_{i2}| + \cdots + |d_{ik}|. \]

(13)

Draw a fitting curve between the distance and cumulative production of \(N\) wells in block A. The logarithmic curve is selected for fitting according to the production characteristics of the actual well. The characteristics of the fitting curve are as follows: when the parameters are far from the optimal value and the distance is larger, the result will be greatly improved if the parameters are slightly improved. When each index is closer to the optimal value and the distance is smaller, the index is improved, and the output change is smaller.

### 3. Application

In this part, the proposed method was used in two typical reservoirs. The first one has 10 multifractured horizontal wells, and a detailed process for production prediction was introduced. The other reservoir has 154 multifractured horizontal wells and was also introduced in this work to verify the proposed.

#### 3.1. Typical Tight Oil Reservoirs with 10 Multifractured Horizontal Wells

There are two types of sedimentary facies including delta and lacustrine for the tight reservoirs. Turbidite reservoirs formed by delta front landslides are mainly formation. The main reservoir is located between 1700 and 1900 meters, and the thickness of the sand body is 15~30 meters. This block is a typical tight oil reservoir of low porosity and low permeability. The local microfractures are developed in the reservoir. The reservoir is brittle and compressible and has been developed by hierarchical multicluster fracturing completion technology, and 10 wells are in production. The production conditions of the 10 wells are similar. The average oil production of well at the beginning of the reservoir is about 14.34 t/d. The initial production is relatively high, and the decline is slower. This block adopts advanced fracturing technology, and it induced many complex fractures after fracturing. It is a demonstration block of tight oil reservoir with volume fracturing in China. In this paper, using the actual data, combined with the geological characteristics of tight reservoir and fracturing effect to predict productivity, it has guiding significance for the future development scheme design of the tight reservoir.

The parameter data of 10 multistages and multicluster fracturing horizontal wells from tight oil reservoir have been listed in Table 1.

The information content of the seven parameters in Table 1 is calculated, and the results of the information content of each parameter are shown in Figure 2. The values are main fracture half-length (71.64), oil zone length (64.89), number of fractures (52.21), main fracture conductivity (47.20), matrix permeability (46.84), porosity (34.11), and number of fracture clusters (8.14). From the calculation results, the parameter weights are as follows: matrix permeability (14.41%), porosity (10.49%), reservoir length (10.49%), reservoir length (19.96%), half-length of main fracture (22.04%), conductivity of main fracture (14.52%), number of fractures (16.06%), and number of clusters (25.1%).

The original data of the 7 main productivity influencing factors of the 10 horizontal wells in the experimental area are normalized, and the normalized matrix is

\[
\begin{bmatrix}
  0.72 & 0.96 & 0.83 & 0.55 & 0.92 & 0.75 & 0.33 \\
  0.06 & 0.67 & 0.69 & 0.58 & 0.92 & 0.50 & 0.33 \\
  0.99 & 0.97 & 0.89 & 0.61 & 1.00 & 0.63 & 0.33 \\
  0.74 & 1.00 & 0.83 & 0.61 & 0.66 & 0.63 & 0.33 \\
  0.22 & 0.96 & 1.00 & 1.00 & 0.61 & 1.00 & 1.00 \\
  0.14 & 0.96 & 0.97 & 0.89 & 0.71 & 0.92 & 0.67 \\
  0.18 & 0.88 & 0.90 & 0.34 & 0.77 & 0.50 & 0.33 \\
  1.00 & 0.94 & 0.52 & 0.37 & 1.00 & 0.50 & 0.33 \\
  0.99 & 0.94 & 0.75 & 0.82 & 0.69 & 1.00 & 0.67 \\
  0.18 & 0.94 & 0.87 & 0.66 & 0.53 & 0.83 & 0.67
\end{bmatrix}. \]

(14)

Characteristic matrix is

\[
\begin{bmatrix}
  0.10 & 0.10 & 0.16 & 0.12 & 0.13 & 0.12 & 0.01 \\
  0.01 & 0.07 & 0.14 & 0.13 & 0.13 & 0.08 & 0.01 \\
  0.14 & 0.10 & 0.18 & 0.13 & 0.15 & 0.10 & 0.01 \\
  0.11 & 0.10 & 0.17 & 0.13 & 0.10 & 0.10 & 0.01 \\
  0.03 & 0.10 & 0.20 & 0.22 & 0.09 & 0.16 & 0.03 \\
  0.02 & 0.10 & 0.19 & 0.20 & 0.10 & 0.15 & 0.02 \\
  0.03 & 0.09 & 0.18 & 0.08 & 0.11 & 0.08 & 0.01 \\
  0.14 & 0.10 & 0.10 & 0.08 & 0.14 & 0.08 & 0.01 \\
  0.14 & 0.10 & 0.15 & 0.18 & 0.10 & 0.16 & 0.02 \\
  0.03 & 0.10 & 0.17 & 0.15 & 0.08 & 0.13 & 0.02
\end{bmatrix}. \]

(15)
The single-factor law of this block has been studied in this work. The parameters of matrix permeability, porosity, oil zone length, main fractures half-length, main fractures conductivity, and number of fractures are the larger the value, the higher the productivity, and the number of fracture clusters. However, the two clusters have the highest productivity when the number of cracks is the same.

According to the above laws, the optimal value matrix composed of the optimal values of each column of parameters is obtained:

\[
X_{opt} = \begin{bmatrix}
0.36 & 10.67 & 1506 & 380 & 301 & 48 & 2
\end{bmatrix}
\] (16)

Optimal characteristic matrix is

\[
x_{opt}' = \begin{bmatrix}
0.14 & 0.11 & 0.20 & 0.22 & 0.15 & 0.16 & 0.01
\end{bmatrix}
\] (17)

Distance matrix is

\[
d = \begin{bmatrix}
0.041 & 0.004 & 0.035 & 0.099 & 0.011 & 0.040 & 0.000 \\
0.136 & 0.034 & 0.062 & 0.093 & 0.011 & 0.080 & 0.000 \\
0.001 & 0.004 & 0.021 & 0.087 & 0.000 & 0.060 & 0.000 \\
0.038 & 0.000 & 0.034 & 0.086 & 0.049 & 0.060 & 0.000 \\
0.113 & 0.004 & 0.000 & 0.000 & 0.056 & 0.000 & 0.017 \\
0.124 & 0.004 & 0.006 & 0.024 & 0.042 & 0.013 & 0.008 \\
0.119 & 0.012 & 0.019 & 0.145 & 0.034 & 0.080 & 0.000 \\
0.000 & 0.007 & 0.096 & 0.140 & 0.000 & 0.080 & 0.000 \\
0.001 & 0.007 & 0.049 & 0.041 & 0.045 & 0.000 & 0.008 \\
0.119 & 0.007 & 0.026 & 0.075 & 0.069 & 0.027 & 0.008
\end{bmatrix}
\] (18)
The distance and cumulative production of 9 wells in the well area except well 10 are fitted.

According to the production characteristics of actual wells, the logarithmic curve is selected for fitting. When the parameters are far from the optimal value and the distance is large, the production will be greatly improved if the parameters are slightly improved. When the value is close and the distance is small, the index is improved, and the output change is small. According to formula fitted of the above, 9 horizontal wells are distances and the cumulative production. As shown in Figure 3, the cumulative production of this well is predicted by combining the parameters of well 10. The sum of the distances obtained after processing the parameters of well 10 is 0.33, and the predicted cumulative production is $1.76 \times 10^4$ m$^3$ according to the fitted formula. The prediction error is 4.7% comparing the cumulative $1.68 \times 10^4$ m$^3$. The prediction is more accurate, and the effect is better. It can be seen from the figure, the sum of distance has a great influence on the production, and the smaller the sum of distance is, the smaller the influence will be.

3.2. Shale Oil Reservoirs with 154 Multifractured Horizontal Wells. In this part, a shale oil reservoir with 154 multifractured horizontal wells was applied in the model. The shale reservoir is located in Xinjiang, China. The reservoir has been produced for more than 5 years and has enough production data and reservoir parameters. In this work, we choose 154 wells with relatively complete data. After analysis, the mainly production influence factors include porosity, oil saturation, permeability, thickness, well length, viscosity, fracture spacing, fracture number, and fracture length. The
weight of above parameters is 14.26%, 10.64%, 15.84%, 12.85%, 15.56%, 10.74%, 7.75%, 5.39%, and 6.97%, respectively. According to the above process, we can get the production prediction curve, as shown in Figure 4. To tell the truth, the production prediction has a relative error to the first reservoir with 10 wells. The largest prediction error is 25% in the 154 wells. However, most wells have a relatively good prediction accuracy.

4. Conclusion

(1) According to the analysis of the data of tight reservoirs, the weights of the factors affecting the productivity of the block are as follows: mass permeability (14.41%), porosity (10.49%), length (19.96%), main fracture half-length (22.04%), main fracture conductivity (14.52%), number of fractures (16.06%), and number of clusters (2.51%)

(2) In this paper, 9 multifractured horizontal wells are used as examples to calculate the production forecast plate of the block. The production data of the 10th well are used to verify the production chart of the block. The error between the predicted new well production and the actual production is 4.7%

(3) The improved cluster analysis method is fast and accurate for productivity prediction. When the parameters are far from the optimal value and the distance is larger, the output growth rate is greater; when the indicators are closer to the optimal value, and the distance is smaller, the output growth rate is smaller. According to the fitting curve of the actual block of the tight reservoir, it can be speculated that if the production parameters of a well are consistent with the optimal conditions, the productivity can reach $2.19 \times 10^4 \text{m}^3$

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no competing interest.

Acknowledgments

This study was supported by the Key Postdoctoral Technology Project of Shengli Oilfield, Sinopec (YKB2110).

References


