energies

MDPI

Article

Thief Zone Assessment in Sandstone Reservoirs Based on Multi-Layer Weighted Principal Component Analysis

Bin Huang [1], Rui Xu [1], Cheng Fu [1,2,*], Ying Wang [3] and Lu Wang [4]

1 College of Petroleum Engineering, Northeast Petroleum University, Daqing 163318, China; huangbin111@163.com (B.H.); xu_rui_0131@163.com (R.X.)
2 Post-Doctoral Scientific Research Station, Daqing Oilfield Company, Daqing 163413, China
3 Harold Vance Department of Petroleum Engineering, Texas A&M University, College Station, TX 77843, USA; annieyw@tamu.edu
4 China National Petroleum Corporation Exploration and Development Research Institute, Beijing 100083, China; wanglu2791@petrochina.com.cn
* Correspondence: sygcxytyb@163.com; Tel.: +86-459-650-4325

Received: 28 March 2018; Accepted: 14 May 2018; Published: 16 May 2018

Abstract: Many factors influence the evaluation process of thief zones. The evaluation index contains very complex information. How to quickly obtain effective information is the key to improve the evaluation quality for thief zones. Considering that the correlation and information redundancy among the evaluation indexes will seriously affect the evaluation results for the thief zone, based on the principal component analysis (PCA) method, this paper proposes a multi-layer weighted principal component analysis method (MLWPCA). Firstly, factor analysis is performed on the original data to obtain the plurality subsystems of the evaluation index. Then, a principal component is analyzed through the subsystems of the evaluation index PCA to obtain the principal component score. Finally, the subsystem is weighted by the factor score and the comprehensive thief zone score is obtained by combining the subsystem weight and the subsystem score. A case study on the Daqing oilfield shows the effectiveness of the method, verified by tracer tests when applying the MLWPCA method to evaluate the thief zone. The thief zone of the Daqing oilfield is obviously affected by effective thickness, coefficient of permeability variation and interwell connectivity. At present, there are 10 well developed thief zones and eight medium developed thief zones in Daqing oilfield. The accuracy of this method is 94.44%. Compared with PCA, this method has better pertinence in evaluating thief zones, and is more effective in determining the principle influencing factors.

Keywords: thief zone; multi-layer weighted; principal-component-analysis; tracer

1. Introduction

During long-term water flooding, some geologic factors, such as sand production and clay erosion, and a large number of production factors, such as injection pressure and high recovery rate, will contribute greatly to the heterogeneity of formation structures, which may lead to the widespread formation of thief zones. The thief zone is defined as a laterally continuous stratigraphic unit of relatively high permeability and large pore radius, which has approached residual oil saturation [1]. In reservoirs with thief zones, earlier water breakthrough resulting in uneven sweep of the reservoirs, and the utilization efficiency of the injected water is seriously impaired, which leads to lower oil recovery and more difficulty in undertaking some stimulation measures. Accordingly, the key to enhance oil recovery in the high water cut stage is how to effectively identify thief zones and determine which wells should have their profile modified.

Up to now, there are several approaches to identify thief zones. Chetri et al. presented production logs combined with dynamic data providing inferences on water breakthrough trends thus helping to identify thief zones with high permeability [2]. Al-Dhafeeri et al. identified thief zones using core data and production logging tests (PLTs) [3]. They found these zones contributed more than 50% of the total well production with permeability greater than 20 Darcy. Li et al. described a method of identifying thief zones using integration of production logging test (PLT), nuclear magnetic resonance (NMR) and high-resolution image logs [4]. John et al. detected for the first time the location of thief zones using distributed temperature sensing (DTS) technology combined with production logging tests (PLTs) and water flow logs (WFLs) [5]. Chen et al. studied the thief zones using production logging tests (PLTs) and summarized different types of thief zone distribution [6]. These methods are simple and easily identify thief zones, however, the well logging can only analyze the situation in the near well region and the tests will affect the normal operation of the well.

Feng et al. applied interference well tests to determine the thief zones and calculated the permeability and thickness of thief zones based on a semilog method [7]. Feng et al. characterized thief zones in mature water flooding reservoirs using pressure transient analysis and established a mathematical model for a well intersected by a high-permeability streak. The solution in Laplace space is derived by Ozkan's source function [8]. It is relatively cheap to identify the thief zone by well testing, but these methods must be based on an ideal model, which is quite different from the actual reservoir situation.

Watkins found thief zones through analyzing the time required for tracer-tagged liquids to flow from injection wells to production wells [9]. Ravenne et al. identified the size and distribution of thief zones through building the 3D geocellular model and nested pixel simulations [10]. Shawket et al. investigated the evolution of an injected water front and the effect of different reservoir heterogeneity parameters and gravity on the thief zones [11]. Izgec et al. used modified-Hall analysis (MHA) to discern the characteristics of the thief zone [12]. Ajay et al. took advantage of production logging test (PLT) data, streamline trajectories and tracer data to form an efficient assisted history-matching (AHM) workflow to identify a thief zone [13]. Although these methods are able to quantitatively describe thief zones accurately, they are always time-consuming and expensive.

Wang et al. first used the ISODATA clustering analysis method to determine the thief zone [14]. However, this method only describes the situation near the wellbore, and the selection of the evaluation index is too artificial, lacking a corresponding selection method. Ding et al. presented a methodology of determining the thief zone by using automatic history matching and fuzzy mathematics [15]. This method takes into account the uncertainty of the geology, but the parameters are difficult to obtain.

In order to improve the recognition accuracy of thief zones and to eliminate the interference of human factors on the recognition process, the principal component analysis (PCA) method is proposed to identify thief zones. However, due to the complicated factors which influence the thief zone, when different properties and levels of the indicators are directly evaluated the distinction of recognition results is low. To solve this problem, in this paper we suggested to layer the evaluation indicators and construct a multi-layer weighted principal component analysis (MLWPCA) method based on the PCA method. Combining the reservoir geological data and production monitoring results, the thief zone of the Daqing oilfield is quantitatively evaluated by the above method. The accuracy of the evaluation results is varified by the tracer method.

2. Analysis Method

2.1. Principal-Component-Analysis

The basic idea of the PCA method is to reduce the dimensions of the original data, converting multiple variables into several independent composite variables (principal components), and selecting the number of principal components according to the principle that the contribution rate of variance is more than 85%. These principal components can reflect most of the information of the original

variables and contain information that does not overlap each other. Then the composite principal component score is calculated by using variance contribution rate as weight and the quantitative evaluation is performed. The main steps of principal-component-analysis are:

(a) In order to improve the accuracy of data analysis and eliminate the influence of data dimension, the original data of n samples is standardized by using Equation (1):

$$Z_{ij} = \frac{x_{ij} - \overline{x}_{ij}}{S_j} i = 1, 2, \cdots, n; \quad j = 1, 2, \cdots, p \tag{1}$$

Z_{ij} is the standardized evaluation index; x_{ij} is the evaluation index; $\overline{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$, $S_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \overline{x}_j\right)^2$.

(b) Calculating the correlation coefficient matrix R:

$$R = \left[r_{ij}\right]_p xp = \frac{Z^T Z}{n-1} \tag{2}$$

$$r_{ij} = \frac{\sum z_{kj} \cdot z_{kj}}{n-1}, i, j = 1, 2, \cdots, p$$

R is correlation matrix; Z is standardized matrix of evaluation index.

(c) The eigenvalue λ_j of the correlation coefficient matrix R is calculated, and the number of principal components m is determined according to the principle of variance contribution rate greater than 85% ($\sum_{j=1}^{m} \lambda_j / \sum_{j=1}^{p} \lambda_j \geq 85\%$).

(d) Calculating the principal component loading:

$$l_{ij} = \sqrt{\lambda_j} a_{ij} i = 1, 2, \cdots, m; \quad j = 1, 2, \cdots, p \tag{3}$$

λ_j is the eigenvalue of the correlation-coefficient-matrix; a_{ij} is the orthogonalized unit of the eigenvector of the correlation matrix.

(e) Calculating the principal component score F_i:

$$F_i = a_{1j} X_1 + a_{2j} X_2 + \cdots a_{pj} X_j i = 1, 2, \cdots, m; \quad j = 1, 2, \cdots, p \tag{4}$$

F_j is the score of the ith principal component; X_i is the factor score of the ith principal component.

(f) Calculating the composite score Y_1:

$$Y_1 = \frac{\lambda_1}{\sum_{i=1}^{m} \lambda_i} \times F_1 + \frac{\lambda_2}{\sum_{i=1}^{m} \lambda_i} \times F_2 + \cdots + \frac{\lambda_i}{\sum_{i=1}^{m} \lambda_i} \times F_i i = 1, 2, \cdots, m \tag{5}$$

Y_1 is the comprehensive score of the principal-component-analysis-method.

2.2. Multi-Layer Weighted Principal- Component-Analysis

Because the PCA method has low discrimination in analyzing multi-level and multi-angle evaluation systems, this paper constructs the multi-layer weighted principal component analysis (MLWPCA) method. Its core idea is to divide the index subsystem based on factor analysis, then analyze the main components of each index subsystem, and weight each index subsystem by factor score. Finally, we synthesize the main component analysis result of each index subsystem to obtain a comprehensive score Y_2. Compared to the PCA method, the MLWPCA method divides

the total system into several subsystems in which the number of indicators becomes smaller without changing the number of evaluation sample points. According to the law of large numbers, the larger the number of the evaluation samples is relative to the indicator, the more stable the covariance matrix is, and the higher the evaluation accuracy is. This has greatly improved the stability and credibility of the evaluation results of the thief zone. Moreover, unlike the PCA method which uses a covariance matrix to describe the correlation between the indicators, the MLWPCA method uses the factor rotation load matrix obtained by the maximum orthogonal rotation method to describe the correlation between the indicators. In this way, the variance of the more important indicators in the evaluation system can be elongated and received more attention in the evaluation, which makes the classification of the thief zone more explicit. The main steps of the multi-layer weighted principal-component- analysis are:

(a) Carrying out factor analysis on the standardized matrix Z, selecting m principal factors according to the principle that the cumulative variance contribution rate is more than 75 percent, and dividing the system into m subsystems, where each subsystem comprises p indexes.

(b) Index subsystem weight calculation formula is:

$$\omega_i = \frac{\sum_{j=1}^{p} \beta_{ij} e_i}{\sum_{i=1}^{m} \sum_{j=1}^{p} \beta_{ij} e_i}, i = 1, 2, \cdots m; \quad j = 1, 2, \cdots, p \tag{6}$$

ω_i is the subsystem weight; j is the number of indicators; β_{ij} is each factor score coefficient; e_i is variance contribution rate, usually $\sum_{i=1}^{m} e_i \geq 80\%$.

(c) Principal-component-analysis is carried out on each index subsystem, and the comprehensive score Y_2 is weighted according to the corresponding weight:

$$Y_2 = \omega_1 \times Y_{21} + \omega_2 \times Y_{22} + \cdots + \omega_m \times Y_{2m} \quad i = 1, 2, \cdots m; \tag{7}$$

Y_2 is the comprehensive score of the multi-layer weighted principal-component-analysis- method; Y_{2i} is the composite score of the subsystem i.

3. Example

Taking the Daqing oilfield as the evaluation target, based on the reservoir geological data and production monitoring results, the related thief zone evaluation index is selected, and the MLWPCA method is used to evaluate the thief zone, and the evaluation results are verified by tracer tests.

3.1. Overview of Research Blocks

The Daqing oilfield is heterogeneous sandstone reservoirs with positive rhythmic deposition. The burial depth of reservoir is 780~1300 m. The average effective thickness is 43.5 m. The average porosity is 18%. The original oil saturation is 52~61%. The original formation pressure of the reservoir is 11.07 Mpa. The difference of ground saturation pressure is 8.23 MPa. The temperature of the oil layer is 42.7~51 °C. The density of underground crude oil is 0.89 g/cm^3. There are 11 water injection wells and 36 production wells.

Since 1978, the Daqing oilfield has been developed and experienced three stages. The depletion development mode was used in the early stage. At this stage, the formation pressure dropped rapidly with no stable period. The second stage started with water injection. As the water injected increased, the liquid production in the oil field increased and the decline rate slowed down. The third stage is the full water injection stage. The oil production and water content are all increased greatly at this stage. At present, the Daqing oilfield has entered the high water cut stage. The water cut rose sharply (the comprehensive water cut in 2016 was 71%) and the recovery rate was low (the geological

reserves recovery rate in 2016 was 7.23%). The maximum daily water injection for a single well is about 1000 m^3/d. Average oil pressure of the injection well is 9.8 MPa. The comprehensive water content is 89.8%.

From the geological point of view, the oil reservoir has large thickness and is obviously affected by the gravitational differentiation of oil and water. The reservoir is highly heterogeneous. These characters provide a geological basis for the development of thief zones. According to the production process, the oil field has high water content with a cumulative annual growth rate of water production up to 15%. After the water injection capacity is enhanced, the daily oil production does not increase significantly. The analysis shows a large scale thief zone has appeared in the reservoir. The reservoir heterogeneity is aggravated. The injected water ineffective circulated. The water drive sweep volume is greatly reduced, which will seriously reduce the final recovery. Therefore, it is of great significance to analyze and identify the thief zone to improve the development effect of the oil field. The development status of Daqing oilfield in 2016 is shown in Table 1.

Table 1. Development status of Daqing oilfield in 2016.

Annual Oil Production (10^4 m^3)	Daily Oil Production (m^3)	Cumulative Oil Production (10^4 m^3)	Cumulative Water Production (10^4 m^3)	Recoverable Reserve Recovery Degree (%)
44.3	1327	116.5	97.2	35.27

3.2. Evaluation Index Selection

In sandstone reservoirs, due to their larger porosity, permeability, and effective thickness, reservoir heterogeneity is prominent. Gravity differentiation between oil and water has great effects on the reservoir which will form thief zones relatively easily. After the thief zone is formed, the resistance of fluid flow through the reservoir decreases and the underground conductivity is enhanced, which may cause a decrease of the pressure difference between injection and production wells, rapid increase of water cut, and a significant increase of the liquid productivity index. If the gray correlation degree is used to characterize the interwell connectivity, the connectivity between injection and production wells will be significantly enhanced after the formation of the thief zone. Based on the basic theory of reservoir engineering combining the characteristics of thief zone, nine evaluation indexes are selected according to systematic, scientific and representative principles, as shown in Table 2.

Table 2. Evaluation index.

No.	Index	Unit
x_1	Effective thickness	m
x_2	Porosity	%
x_3	Permeability	μm^2
x_4	Permeability variation coefficient	%
x_5	Interwell connectivity	1
x_6	Water content	%
x_7	Apparent injectivity index	$m^3/d{\cdot}MPa$
x_8	Injection-production pressure difference	MPa
x_9	Liquid productivity index	10^3 $m^3/d{\cdot}MPa$

The connectivity between injection and production well can be calculated by Formulas (8)~(11). The time series of water injection for injection wells is:

$$X_0 = \{x_0(t), t = 1, 2, \cdots, n\} \quad (8)$$

X_0 is the time series of daily water injection of the injection well, m^3/d; x_0 is the volume of daily water injection under different times, m^3/d; t is water injection time, d; n is the total number of water injection days, d.

The oil production time series of a production well connected around is:

$$X_i = \{x_i(t),\ \ t = 1, 2, \cdots, n\} \tag{9}$$

X_i is the time series of daily oil production of production well, m^3/d; x_i is the volume of daily oil production under different water injection times, m^3/d; t is water injection time, d; n is the total number of water injection days, d.

The correlation coefficient of sequence X_i and X_0 is:

$$\zeta(t) = \frac{\min\limits_i \min\limits_i \Delta_i(t) + \rho \max\limits_i \max\limits_i \Delta_i(t)}{\Delta_i(t) + \rho \max\limits_i \max\limits_i \Delta_i(t)} \tag{10}$$

$\min\limits_i \min\limits_i \Delta_i(t)$ represents the minimum of absolute difference, m^3/d; $\max\limits_i \max\limits_i \Delta_i(t)$ represents the maximum of absolute difference, m^3/d; $\Delta_i(t) = |x_0(t) - x_i(t)|$ is absolute difference, m^3/d; $\rho \in (0, 1)$ is resolution coefficient, usually the value is 0.1~0.5.

The correlation degree is defined as:

$$r_i = \frac{1}{n}\sum_{k=1}^{n} \zeta_i(t) \tag{11}$$

r_i is the correlation degree between subsequence i and sequence 0, and n is the sequence length. The data required to calculate connectivity of water injection and oil production are obtained by actual measurements onsite.

The coefficient of permeability variation is calculated by Equation (12):

$$V_K = \frac{\sum\limits_{i=1}^{n} \left(K_i - \overline{K}\right)^2 \div n}{\overline{K}} \tag{12}$$

V_K is the coefficient of permeability variation, %; K_i is the permeability of sample i, μm^2, $\overline{K}$ is the average permeability of all samples, μm^2; n is the number of samples. The permeability data are obtained by well logging.

The apparent injectivity index and the liquid productivity index are calculated by Equations (13) and (14), respectively:

$$A_W = \frac{Q_W}{P_W} \tag{13}$$

$$A_O = \frac{Q_L}{P_O} \tag{14}$$

A_K is apparent injectivity index, $m^3/d{\cdot}MPa$; A_O is liquid productivity index, $m^3/d{\cdot}MPa$; Q_w is the volume of daily water injection from injection well, m^3/d; Q_L is the volume of daily oil production from production well, m^3/d; P_w is wellhead pressure of injection well, MPa; P_O is wellhead pressure of production well, MPa. The volume of daily water injection and daily oil production, and the wellhead pressure data involved in the calculation are obtained from actual on-site measurements.

3.3. Principal-Component-Analysis-Method

The standardized original data of evaluation index for the Daqing oilfield is shown in Figure 1. The principal-component-analysis is performed after that. According to the principle that the cumulative contribution rate of eigenvalues is greater than 85%, three principal components are selected (Eigenvalue > 1). The eigenvalues of the three principal components were 5.499, 1.433 and 1.058 respectively (the Scree plot is shown in Figure 2). The contribution rates of eigenvalue were

61.104%, 15.919% and 11.754% respectively. The cumulative contribution rate of eigenvalue was 88.777%. The principal-component-analysis analyzes the correlation coefficient matrix shown in Table 3.

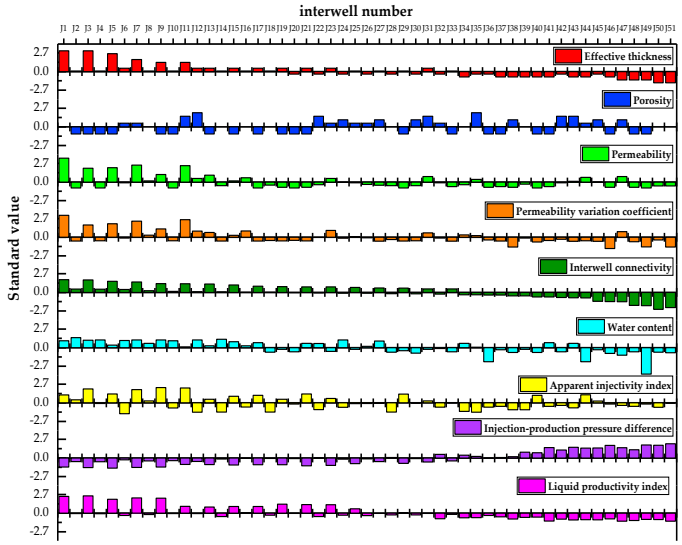

Figure 1. Standardized data.

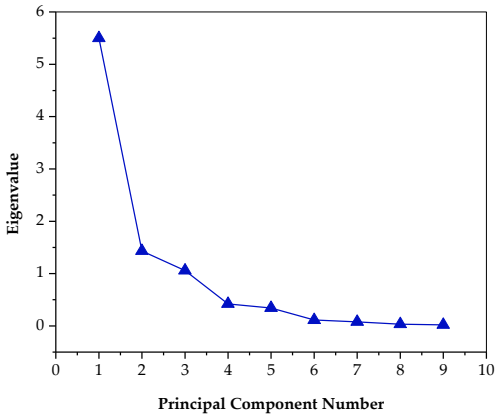

Figure 2. Scree plot.

Table 3. Correlation coefficient matrix.

No.	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1	1.000	−0.103	0.749	0.788	0.887	0.503	0.596	−0.815	0.813
x_2	−0.103	1.000	0.259	0.189	−0.131	0.057	−0.312	−0.153	−0.255
x_3	0.749	0.259	1.000	0.930	0.539	0.272	0.514	−0.417	0.648
x_4	0.788	0.189	0.930	1.000	0.654	0.399	0.476	−0.584	0.721
x_5	0.887	−0.131	0.539	0.654	1.000	0.579	0.534	−0.949	0.869
x_6	0.503	0.057	0.272	0.399	0.579	1.000	0.108	−0.545	0.431
x_7	0.596	−0.312	0.514	0.476	0.534	0.108	1.000	−0.466	0.708
x_8	−0.815	−0.153	−0.417	−0.584	−0.949	−0.545	−0.466	1.000	0.849
x_9	0.813	−0.255	0.648	0.721	0.869	0.431	0.708	−0.849	1.000

The correlation coefficient matrix (Table 4) shows that the injection-production-pressure has negative correlation with the other eight evaluation indexes. This is because with the thief zone, the flow resistance of the fluid through the formation decreases resulting in

decreased injection-production-pressure. At the same time, an evident negative correlation between injection-production-pressure and interwell connectivity indicates that pressure plays a significant role in driving fluid flow to a local area. A clearly positive correlation between permeability and permeability coefficient is observed for the remaining eight evaluation indexes. This is because the permeability coefficient is obtained by the ratio of the permeability standard deviation and the average permeability, but the physical meaning of these two evaluation indexes is different. Permeability is mainly used to evaluate the development of the thief zone near the well, while the permeability coefficient is used to evaluate the heterogeneity of the entire formation.

Calculating the communalities of the three principal components shows they retain at least 85% of the information from the original data except for the index of liquid productivity and water content information. It shows that the principal-component-analysis (PCA) has good effects on dimension reduction and simplifying the original complex multi-dimensional evaluation system. The results of the calculation of the communalities are shown in Table 4.

Table 4. Communalities calculation result.

No.	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
Communalities	0.923	0.865	0.949	0.909	0.925	0.801	0.797	0.881	0.941

The load values of the three principal components are calculated (the load diagram is shown in Figure 3). The functional expression of each principal component is obtained according to the load values as follows.

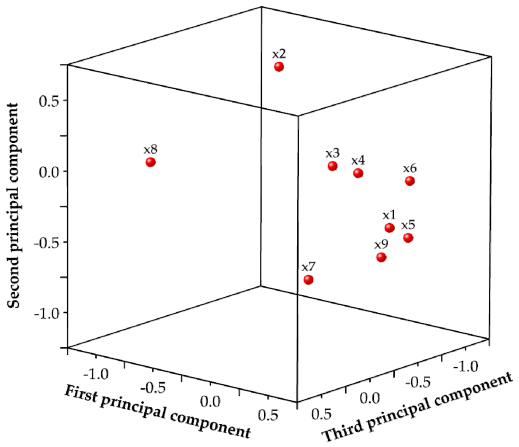

Figure 3. Loading diagram.

$$Y_{11} = 0.410 \times Z(x_1) - 0.039 \times Z(x_2) + 0.327 \times Z(x_3) + 0.359 \times Z(x_4) + 0.394 \times Z(x_5) + 0.238 \times Z(x_6) + 0.287 \times Z(x_7) - 0.370 \times Z(x_8) + 0.404 \times Z(x_9)$$

$$Y_{12} = 0.000 \times Z(x_1) + 0.768 \times Z(x_2) + 0.397 \times Z(x_3) + 0.336 \times Z(x_4) - 0.107 \times Z(x_5) + 0.090 \times Z(x_6) - 0.260 \times Z(x_7) + 0.159 \times Z(x_8) - 0.163 \times Z(x_9)$$

$$Y_{13} = 0.006 \times Z(x_1) - 0.101 \times Z(x_2) + 0.356 \times Z(x_3) + 0.188 \times Z(x_4) - 0.227 \times Z(x_5) - 0.672 \times Z(x_6) + 0.483 \times Z(x_7) + 0.293 \times Z(x_8) + 0.072 \times Z(x_9)$$

According to the scores of principal components, the comprehensive evaluation scores Y_1 (shown in Table 4) of the thief zone between injection and production wells in Daqing oilfield are obtained, ranging from −2.600 to 4.704. The thief zone developed with the increase of comprehensive

score. When the comprehensive score is negative, the seepage channel between injection and production wells is in good condition and the water drive power is sufficient. When the comprehensive score is greater than 0.0 but less than 1.0, the thief zone is moderately developed. At this time, close observation and appropriate measurements are needed to avoid further development into a big pore throat. When the comprehensive score is greater than 1.0, the thief zone is well developed and water channeling has occurred. A large amount of injected water is circulating inefficiently between injection and production wells. According to Figure 4, the Daqing oilfield has 10 well developed thief zones, accounting for 19.61% of the total number of thief zones; and eight moderately developed thief zone, accounting for 15.69% of the total number of thief zones. Improvements, such as profile control and water plugging, are needed for the injection-production wells with thief zones to avoid further development which may influence the final recovery ratio of the oilfield.

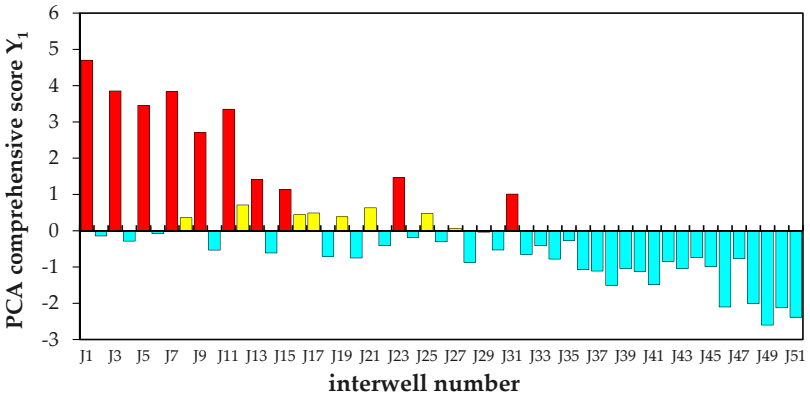

Figure 4. PCA comprehensive score.

3.4. Multi-Layer Weighted Principal-Component-Analysis-Method

The PCA method cannot accurately extract and analyze the development of the thief zone in the same class. In order to improve the accuracy of the evaluation results using PCA, the MLWPCA method was constructed. Firstly, nine indexes are classified according to the result of factor analysis. According to the principle of cumulative contribution rate, greater than 75%, two factors are selected. The contribution rate is 61.104% and 15.919% respectively. The rotated load matrix is shown in Table 5. The score coefficient matrix obtained by the maximum orthogonal rotation method shows that the load values of x_1, x_3, x_4, x_7 and x_9 evaluation indexes are higher in factor 1. These indexes are grouped into index subsystem 1. In factor 2, x_2, x_5, x_6, x_8, the index load is higher. These indicators are divided into index subsystem 2. Using Equation (8) for weight calculation the weights of subsystem 1 and 2 are 0.777 and 0.223, respectively.

Table 5. Rotated load matrix and score coefficient matrix.

No.	Index	Rotated Load Matrix		Score Coefficient Matrix	
		Factor 1	Factor 2	Factor 1	Factor 2
x_1	Effective thickness	0.926	0.610	0.135	0.107
x_2	Porosity	0.119	−0.054	0.104	0.013
x_3	Permeability	0.726	0.173	0.103	−0.202
x_4	Permeability variation coefficient	0.859	0.360	0.201	−0.064
x_5	Interwell connectivity	0.519	0.880	−0.331	0.283
x_6	Water content	0.033	0.777	−0.596	0.558
x_7	Apparent injectivity index	0.719	0.076	0.315	−0.263
x_8	Injection-production pressure difference	−0.420	−0.802	0.089	−0.330
x_9	Liquid productivity index	0.703	0.566	0.134	0.068

PCA was carried out for the evaluation index of the above two thief zone subsystems. The contribution rates of subsystem 1 and subsystem 2 were 90.302% and 86.157%, respectively. The synthesis scores Y_{21} and Y_{22} (Y_{22}' and Y_{22}'') are expressed in the following equations. Based on these scores Y_{21} and Y_{22}, the synthesis score Y_2 of thief zone is calculated shown in Figure 5. The thief zones of the Daqing oilfield are evaluated by the comprehensive score Y_2.

$$Y_{21} = 0.476 \times Z(x_1) + 0.451 \times Z(x_3) + 0.461 \times Z(x_4) + 0.375 \times Z(x_7) + 0.466 \times Z(x_9)$$

$$Y_{22}' = -0.104 \times Z(x_2) + 0.617 \times Z(x_5) + 0.484 \times Z(x_6) - 0.611 \times Z(x_8)$$

$$Y_{22}'' = 0.950 \times Z(x_2) - 0.025 \times Z(x_5) + 0.307 \times Z(x_6) + 0.056 \times Z(x_8)$$

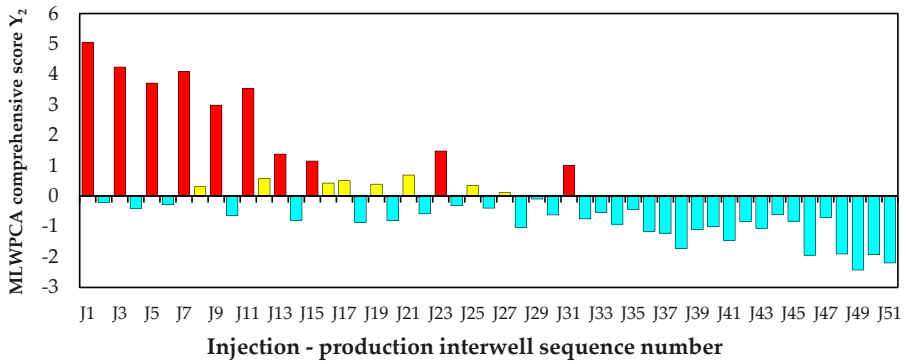

Figure 5. MLWPCA comprehensive score.

Through the comprehensive score Y_2, it can be seen that there are 10 well developed thief zones in Daqing oilfield, eight moderate developed thief zones, and the remaining 34 thief zones are not formed. By comparing and analyzing the original basic data, we found that the effective thickness and the coefficient of permeability variation are the main indexes affecting the development of thief zone in subsystem 1. Such a thief zone has a large effective thickness. The formation of thief zones is clearly affected by the oil-water gravity differentiation. The reservoir is highly heterogeneous, which provides the geological basis for developing thief zones. Subsystem 2 developed a more advanced thief zone than subsystem 1. Interwell connectivity is the main indicator affecting the development of dominant channels in subsystem. 2. Such a thief zone has large throat radius, strong fluid diversion between wells, and very fast rate of further deterioration. Therefore, the corresponding measurements should be taken as soon as possible to control it. It can be concluded from the above analysis that the MLWPCA method is more targeted and differentiated than the traditional PCA method. It is an effective evaluation method for thief zone determination worthy of popularization.

By comparing with the thief zone identification results using the PCA method in Section 3.3, we found that the two methods give consistent identification results for the thief zone. The thief zones, J1, J3, J5, J7, J9, J11, J13, J15, J23 and J31, are seriously developed; while the thief zones, J8, J12, J16, J17, J19, J21, J25, and J27, are moderately developed. Comparing the composite scores of the thief zones between the two methods (as shown in Table 6), we found more obvious differences in the comprehensive score for the thief zones obtained using the MLWPCA method than from PCA method. This can highlight the differences in grade among different thief zones. This can also make the evaluation results more specific and differentiated.

In the process of thief zone evaluation, an interwell tracer test was used to verify the accuracy of the MLWPCA method. As shown from the test results (Table 6), tracer was detected at 10 well developed thief zones by the MLWPCA method, and the average tracer breakthrough time was 6.4 months. Tracer was detected in seven of the total of eight moderately developed thief zones with an average breakthrough time of 10.6 months. Tracer was not detected for J8 interwell which may be due

to the relatively small amount of injected tracer not reaching the oil well. According to the tracer test, the accuracy rate of evaluating the thief zones using the MLWPCA method is 94.44%. This method is proved to be effective in identifying thief zones.

Table 6. Test results of the tracer test.

Interwell Number	Starting Date	Ending Data	Tracer Type	Tracer Amount (kg)	Comment	PCA Comprehensive Score	MLWPCA Comprehensive Score
J1	17-1-2016	23-3-2017	NH_4SCN	5.7	Tracer seen on 23 April 2016	4.704	5.044
J2	17-1-2016	23-3-2017	I^{135}	3.2	No tracer seen	−0.145	−0.212
J3	17-1-2016-	23-3-2017	NH_4SCN	4.9	Tracer seen on 4 June 2016	3.855	4.242
J4	3-3-2015	4-7-2016	NH_4NO_3	4.1	No tracer seen	−0.288	−0.399
J5	17-1-2016	23-3-2017	I^{135}	3.7	Tracer seen on 3 August 2016	3.452	3.714
J6	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	2.7	No tracer seen	−0.082	−0.283
J7	17-1-2016	23-3-2017	NH_4SCN	5.5	Tracer seen on 29 July 2016	3.833	4.101
J8	3-3-2015	4-7-2016	NH_4NO_3	4.4	No tracer seen	0.360	0.306
J9	17-1-2016	23-3-2017	I^{135}	3.1	Tracer seen on 11 September 2016	2.708	2.991
J10	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	2.1	No tracer seen	−0.534	−0.635
J11	17-1-2016	23-3-2017	I^{135}	3.1	Tracer seen on 7 August 2016	3.354	3.528
J12	9-5-2014	15-7-2016	NH_4SCN	7.8	Tracer seen on 3 December 2015	0.711	0.576
J13	17-1-2016	23-3-2017	NH_4SCN	5.3	Tracer seen on 28 November 2016	1.416	1.381
J14	9-5-2014	15-7-2016	NH_4SCN	7.2	No tracer seen	−0.614	−0.790
J15	9-5-2014	15-7-2016	NH_4SCN	6.3	Tracer seen on 30 July 2015	1.143	1.153
J16	3-3-2015	4-7-2016	NH_4NO_3	3.7	Tracer seen on 8 February 2016	0.442	0.431
J17	3-3-2015	4-7-2016	NH_4NO_3	4.1	Tracer seen on 31 January 2016	0.490	0.509
J18	3-3-2015	4-7-2016	NH_4NO_3	4.5	No tracer seen	−0.703	−0.861
J19	3-5-2014	15-7-2016	I^{135}	4.6	Tracer seen on 3 May 2016	0.386	0.391
J20	2016-1-17	23-3-2017	I^{135}	3.3	No tracer seen	−0.752	−0.806
J21	9-5-2014	15-7-2016	NH_4NO_3	5.5	Tracer seen on 30 December 2015	0.632	0.691
J22	17-1-2016	23-3-2017	I^{135}	3.7	No tracer seen	−0.409	−0.583
J23	17-1-2016	23-3-2017	NH_4SCN	5.7	Seeing tracer on 27 September 2016	1.467	1.483
J24	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	3.1	No tracer seen	−0.181	−0.308
J25	16-9-12016	19-12-2017	$C_{28}H_{20}N_3O_5$	2.2	Seeing tracer on 5 September 2017	0.476	0.348
J26	3-3-2015	4-7-2016	NH_4NO_3	4.1	No tracer seen	−0.306	−0.393
J27	9-5-2014	15-7-2016	NH_4NO_3	5.7	Seeing tracer on 4 June 2016	0.061	0.111
J28	9-5-2014	15-7-2016	NH_4SCN	7.1	No tracer seen	−0.874	−1.023
J29	2014-5-9	15-7-2016	NH_4SCN	7.8	No tracer seen	−0.040	−0.096
J30	9-5-2014	15-7-2016	I^{135}	4.3	No tracer seen	−0.530	−0.619
J31	17-1-2016	23-3-2017	NH_4SCN	5.2	Tracer seen on 1 December 2016	1.012	1.002
J32	17-1-2016	23-3-2017	I^{135}	3.2	No tracer seen	−0.650	−0.745
J33	9-5-2014	15-7-2016	NH_4NO_3	4.7	No tracer seen	−0.409	−0.534
J34	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	2.7	No tracer seen	−0.784	−0.920
J35	2014-5-9	15-7-2016	NH_4NO_3	5.1	No tracer seen	−0.263	−0.433
J36	9-5-2014	15-7-2016	NH_4NO_3	4.3	No tracer seen	−1.075	−1.160
J37	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	3.5	Tracer seen on 5 October 2017	−1.112	−1.217

Table 6. *Cont.*

Interwell Number	Starting Date	Ending Data	Tracer Type	Tracer Amount (kg)	Comment	PCA Comprehensive Score	MLWPCA Comprehensive Score
J38	9-5-2014	15-7-2016	NH_4NO_3	4.3	No tracer seen	−1.504	−1.724
J39	9-5-2014	15-7-2016	NH_4NO_3	5.1	No tracer seen	−1.037	−1.100
J40	3-3-2015	4-7-2016	NH_4NO_3	4.3	No tracer seen	−1.119	−1.009
J41	3-3-2015	4-7-2016	NH_4NO_3	3.9	No tracer seen	−1.485	−1.454
J42	3-3-2015	4-7-2016	NH_4NO_3	4.3	No tracer seen	−0.844	−0.827
J43	9-5-2014	15-7-2016	I^{135}	5.2	No tracer seen	−1.046	−1.054
J44	9-5-2014	15-7-2016	I^{135}	4.7	No tracer seen	−0.741	−0.611
J45	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	2.5	Tracer seen on 6 July 2017	−0.989	−0.823
J46	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	2.7	No tracer seen	−2.101	−1.942
J47	16-9-2016	19-12-2017	$C_{28}H_{20}N_3O_5$	2.1	No tracer seen	−0.763	−0.696
J48	2014-5-9	15-7-2016	NH_4NO_3	5.1	No tracer seen	−2.010	−1.902
J49	9-5-2014	15-7-2016	I^{135}	3.7	No tracer seen	−2.600	−2.430
J50	3-3-2015	4-7-2016	NH_4NO_3	3.9	Tracer seen on 2 March 2016	−2.126	−1.925
J51	3-3-2015	4-7-2016	NH_4NO_3	4.5	No tracer seen	−2.385	−2.196

4. Conclusions

In this paper, a MLWPCA method is applied to the evaluation of thief zones. The theoretical basis and calculation steps of this method are introduced. At the same time, the rationality and effectiveness of this method are verified by tracer tests. From the calculation and evaluation results we reached the following conclusions:

(a) The thief zones of the Daqing oilfield are significantly affected by effective thickness, permeability variation coefficient and interwell connectivity. For the severe developed thief zones, J1, J3, J5, J7, J9, J11, J13, J15, J23 and J31, it is urgent to take corresponding profile modification and water shutoff treatment to prevent them from getting worse. For the moderately developed thief zones, J8, J12, J16, J17, J19, J21, J25, J27, closer observation and adjustment of water injection intensity and pressure should be performed to avoid further development of these thief zones.

(b) Through verification by tracer tests, the recognition accuracy of the MLWPCA method for thief zones is 94.44%.

(c) The thief zone evaluation results are consistent between the MLWPCA method and the traditional PCA method. However, more obvious differences in composite scores of the thief zone are obtained from the MLWPCA method. This better highlights the differences in the thief zone at different levels. Moreover, it solves the problem that the traditional PCA method cannot accurately evaluate multi-angle and multi-level evaluation systems. This method can also provide a reference for the evaluation of thief zones in other oilfields.

The data of this study covers a wide range which makes it difficult to obtain these data. This affects the selection of indicators. In the future, when conditions permit, we will further improve the indicator system for the needs of evaluating the thief zones.

Author Contributions: B.H. and R.X. contributed to all parts of the process of this study: developing the methodology, designing the experiments, and writing the paper; L.W. conducted the physical experiment; C.F. and Y.W. analyzed the data and revised the paper.

Acknowledgments: This work is supported in part by the University Nursing Program for Yong Scholars with Creative Talents in Heilongjiang Province (No. UNPYSCT-2017033); the Subsidize of Returned Overseas Students in Heilongjiang Province; the Scientific Research Start-up Funds for Postdoctoral Researchers Settling in Heilongjiang Province; the Northeast Petroleum University Youth Science Foundation (No. NEPUQN2014-27) and the Scientific Research Project of Heilongjiang Provincial Education Department (No. 12541092).

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

References

1. Bane, R.K.; Parker, R.A.; Storbeck, W.G.; Sunde, R.L. Reservoir management of the Fullerton Clearfork unit. In Proceedings of the SPE Permian Basin Oil and Gas Recovery Conference, Midland, TX, USA, 16–18 March 1994.
2. Chetri, H.B.; Al-Anzi, E.; Al-Rabah, A.; Al-Dashti, H.; Al-Mutawa, M.; Chakravarthi, R.; Brown, M.; Isby, J.; Clark, A. Lessons learnt and experiences gained during two years of field monitoring, data integration and reservoir management: A case history of the Mauddud waterflood, North Kuwait. In Proceedings of the SPE Offshore Europe, Aberdeen, UK, 2–5 September 2003.
3. Al-Dhafeeri, A.M.; Nasr-El-Din, H.A. Characteristics of high-permeability zones using core analysis, and production logging data. *J. Pet. Sci. Eng.* **2007**, *55*, 18–36. [CrossRef]
4. Li, B.J.; Hamad, N.; Jim, L.; Mansoor, A.R.; Ihsan, G.; Mohammed, A.K. Detecting thief zones in carbonate reservoirs by integrating borehole images with dynamic measurements. In Proceedings of the SPE Annual Technical Conference and Exhibition, Denver, CO, USA, 21–24 September 2008.
5. John, D.; Hans, V.D.; Maersk, O.; Arve, O.N. Interwell communication as a means to detect a thief zone using DTS in a Danish Offshore well. In Proceedings of the SPE Offshore Technology Conference, Houston, TX, USA, 6–9 May 2013.
6. Chen, Q.; Gerritsen, M.Q.; Kovscek, A.R. Effects of reservoir heterogeneities on the steam assisted gravity drainage process. *SPE Reserv. Eval. Eng.* **2008**, *11*, 921–932. [CrossRef]
7. Feng, Q.; Wang, S.; Gao, G.; Li, C. A new approach to thief zone identification based on interference test. *J. Pet. Sci. Eng.* **2010**, *75*, 13–18. [CrossRef]
8. Feng, Q.; Wang, S.; Zhang, W.; Song, Y.; Song, S. Characterization of high-permeability streak in mature waterflooding reservoirs using pressure transient analysis. *J. Pet. Sci. Eng.* **2013**, *110*, 55–65. [CrossRef]
9. Watkiis, W.R. How to diagnose a thief zone. *SPE Soc. Pet. Eng.* **1973**, *25*, 839–840.
10. Ravenne, C.; Coury, Y.; Cole, J. Characterisation of reservoir heterogeneities and super permeability thief zones in a major oilfield in the Middle East. In Proceedings of the SPE 16th World Petroleum Congress, Calgary, AB, Canada, 11–15 June 2000.
11. Shawket, G.; Younes, B.; Moutaz, S. Thief zones and effectiveness of water-shut-off treatments under variable levels of gravity and reservoir heterogeneity in carbonate reservoirs. In Proceedings of the SPE EUROPEC/EAGE Annual Conference and Exhibition, Barcelona, Spain, 14–17 June 2010.
12. Izgec, B.; Kabir, S. Identification and characterization of high-conductive layers in waterfloods. *SPE Reserv. Eval. Eng.* **2009**, *14*, 113–119. [CrossRef]
13. Ajay, S.; Khan, H.; Majhi, S.; Al-Otaibi, B. Integration of PLT and tracer data using pattern recognition for efficient assisted history matching of heterogeneous North Kuwait carbonate reservoir. In Proceedings of the SPE Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE, 7–10 November 2016.
14. Wang, S.; Jiang, H. Determine level of thief zone using fuzzy ISODATA clustering method. *Transp. Porous Media* **2010**, *86*, 483–490. [CrossRef]
15. Ding, S.W.; Jiang, H.Q. Identification and characterization of high-permeability zones in waterflooding reservoirs with an ensemble of methodologies. In Proceedings of the SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Nusa Dua, Bali, Indonesia, 20–22 October 2015.