Application Of Comprehensive Evaluation Method In Oil Well Classification

Fada Huang², Yifang Tang¹*, Liang Huang², Haitao Zhang²

¹School of Sciences, Southwest Petroleum University, Chengdu 610500, China
²Baikouquan Oil Production Factory, Xinjiang Oilfield Company, Karamay 834011, China

Abstract: China's oil resources are very rich, but the effect of oil exploitation is not good. Precise prediction of oil production plays an extremely important role in formulating a reasonable oil planning plan and determining the direction of investment. In order to improve the accuracy of oil well exploitation, it is necessary to establish a scientific oil production forecasting system. First, pre-treatment research on the data of Baikouquan Oilfield to ensure the quality of oilfield data. Secondly, the main controlling factors affecting the production of oil wells are analyzed, and the single evaluation is carried out by the coefficient of variation method, the entropy method and the mean square error method. Then, the pre-examination test of the consistency test is performed on the single evaluation result, and the combination of the evaluation methods is selected. The combined evaluation method is the average method, and the fuzzy Borda method.

1 Introduction

The comprehensive evaluation refers to a method of simultaneously analyzing and judging multiple levels and multiple factors of the evaluated object using a certain evaluation method [1]. There are many ways to comprehensively evaluate, but the purpose is to recognize things. Usually, several objects are comprehensively analyzed according to certain methods, and then several objects are evaluated to select the best or the worst objects.

As early as 1888, Edgeworth published a "Statistics of Examinations" academic paper at the Journal of the Royal Statistical Society [2], and began to discuss how different parts of the exam should be weighted. In 1913, Sperman published "Correlation of Sums" in the Britkh Journal of Psychology [3]. Or Differences’ paper discusses the role of different weights in the statistical synthesis process; the German economist Etienne Laspeyres (1834-1913) proposed the Laspey index (referred to as the Laplaceyres Index), using the selected base period. Price or quantity to measure the price or quantity of the current period. The comprehensive evaluation theory and method are formed in the above various theories. With the advent of these theories, decision science has risen rapidly, and the decision-making goal has changed from a simple one to a plurality, and then multiple decision-making objectives of multiple decision makers have emerged. In order to solve these problems, domestic and foreign scholars have successively proposed data envelopment analysis methods, ELECTEE methods, network analytic hierarchy process (ANP) and grey system theory.

The main purpose of the research examples in this paper is to use the comprehensive evaluation method to analyze the main controlling factors affecting the oil well production, and then predict the oil well production. Therefore, for the convenience of discussion and analysis, we simply divide the oil wells into two categories according to the output value. That is, effective wells (>2 squares) and invalid wells (<2 squares). Based on the above analysis, this paper will discuss from two perspectives. One is the comprehensive analysis of the main controlling factors of oil well production, based on the comprehensive evaluation index system, the comprehensive evaluation analysis.

*Corresponding Author: Yifang Tang; email: tangyifang_swpu@163.com; phone:15828532038

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
between indicators; the second is based on the analysis of the main control Factors are used to predict production categories.

2 Evaluation model

2.1. Single evaluation model

1. Variation coefficient method steps
   (1) Calculate the coefficient of variation \( V \)
   \[ V_i = \frac{\sigma_i}{\bar{x}_i} (i = 1, 2, ..., n) \]
   (1)

   Where \( V_i \) is the coefficient of variation of the index of item \( i \), \( \sigma_i \) is the mean squared error of the index of item \( i \), \( \bar{x}_i \) is the average of the index of item \( i \), and \( n \) represents the number of indicators.

   (2) Normalization of coefficient of variation
   \[ W_i = \frac{V_i}{\sum_{i=1}^{n} V_i} \]
   (2)

2. Mean variance method steps
   (1) Find the mean square error \( s \)
   \[ s(k) = \left[ \frac{1}{m-1} \sum_{j=1}^{m} (x_{ij} - \bar{x}(k))^2 \right]^{1/2} (k = 1, 2, ..., n) \]
   (3)

   Where \( x_{ij} \) is the \( k \) observation of the \( i \) indicator, \( \bar{x}(k) \) is the evaluation value of the \( k \) indicator, \( m \) is the number of observation samples, and \( n \) is the number of indicators.

   (2) Normalization of mean square error
   \[ W_i = \frac{s(k)}{\sum_{k=1}^{n} s(k)} \]
   (4)

2.2. Combined evaluation model

The steps of the three evaluation models of the average method, the fuzzy Borda method and the dispersion maximization method are respectively introduced below.

1. Average method

   Let \( y_{ij} \) be the evaluation conclusion obtained by the \( i \) evaluation method with the \( j \) evaluation method, then the combined evaluation value of the \( i \) evaluation index is:
   \[ \bar{y}_i = \frac{1}{n} \sum_{j=1}^{n} y_{ij} \]
   (5)

   Where \( n \) represents the number of single evaluation methods.

2. fuzzy Borda method

   The calculation process is as follows:
   (1) Calculate the membership relationship of each evaluation object under each single evaluation method, and let \( x_{ih} \) be the evaluation value or ranking order of the \( i \) evaluation index under the \( k \) single evaluation method, and \( u_{ik} \) is the \( i \) evaluation index. The degree of membership close to "good" under the \( k \) single evaluation method, then
   \[ u_{ik} = \frac{x_{ih} - \min \{x_{ih}\}}{\max \{x_{ih}\} - \min \{x_{ih}\}} \times 0.9 + 0.1, (i = 1, 2, ..., n, k = 1, 2, ..., m) \]
   (6)

   (2) Let the fuzzy frequency of the evaluation index \( x_i \) be:
   \[ f_{ih} = \sum_{k=1}^{m} \delta_{ih} u_{ik} \delta_{ih} = \begin{cases} 1 & \text{If } x_i \text{ ranks } h \text{ in } n \text{ evaluation indicators} \\ 0 & \text{If } x_i \text{ does not rank } h \text{ in the } n \text{ evaluation indicators} \end{cases} \]
   (7)

   Then, there are fuzzy weights \( W_{hi} = \frac{f_{hi}}{R_i} \), where
   \[ R_i = \sum_{k=1}^{m} f_{ki} \cdot \]
   (3)

   (3) Let \( Q_h \) be the score of the evaluation index in the single evaluation method, which is ranked in the \( h \) position, and the ranking of the evaluation index is converted into the score:
   \[ Q_h = \frac{1}{2} (n - h)(n - h + 1) \]
   (8)

   In fact, \( Q_h \) is a fixed value.

   (4) Find the fuzzy Borda number under each evaluation index:
   \[ u_{ik} = \frac{x_{ih} - \min \{x_{ih}\}}{\max \{x_{ih}\} - \min \{x_{ih}\}} \times 0.9 + 0.1, (i = 1, 2, ..., n, k = 1, 2, ..., m) \]
   (9)

3 Application and analysis

3.1. Single evaluation model result

The raw data of 92 wells and 22 parameters that have been fracturing in Baikouquan are preprocessed. The pretreatment mainly includes: data screening, missing value processing, outlier processing, data standardization, gray correlation analysis, attribute selection, and many more.

Coordinated by Kendall:
\[ \sum_{i=1}^{n} R_i^2 = 5134 \quad W = \frac{12 \cdot 5134 - 3 \cdot 3^3 \cdot 12 \cdot (12 + 1)^2}{3^2 \cdot (12^3 - 12)} = 0.444 \]

It can be seen that the \( W \) value is relatively small,
indicating that the three single evaluation models are not compatible and cannot be combined. Then, for the same reason, we calculate the Kendall coordination coefficient of the two-two single evaluation, as shown in the following table:

| Method                | Coefficient of variation | Entropy method | Mean variance method |
|-----------------------|--------------------------|----------------|---------------------|
| Coefficient of variation | 1                        | 0.8461         | 0.5279              |
| Entropy method        | 0.8461                   | 1              | 0.3741              |
| Mean variance method  | 0.5297                   | 0.3741         | 1                   |

It can be seen from the above table that the coefficient of variation method and the entropy method have strong consistency, indicating that the two single evaluation methods can be combined.

3.2. Combined evaluation model results

(1) Results of a combined evaluation

| Method                | Mean method | Fuzzy Borda method |
|-----------------------|-------------|--------------------|
| Mean method           | 1           | 0.979              |
| Fuzzy Borda method    | 0.979       | 1                  |

It can be seen from the above table that the Spearman correlation coefficients of the results of a combined evaluation using the mean method, the fuzzy and Borda method are all greater than. That is to say, the results of a combined evaluation have a significant positive correlation, and the degree of correlation is very high, indicating that the two methods have good compatibility, a high degree of consistency, and meet the requirements of the second combination evaluation.

(2) Preferred results of a combined evaluation

According to the above rules, the sum of squared errors of the mean value method, the fuzzy Borda method and the dispersion maximization method are calculated. The results are shown in Table 5:

| Method                | Mean method | Fuzzy Borda method |
|-----------------------|-------------|--------------------|
| Mean method           | 0.0017      | 0.0071             |
| Fuzzy Borda method    |             |                    |

It can be seen from the above table that the squared sum of the errors the mean method a is almost equal, which is obviously superior to the fuzzy Borda method. In summary, we choose the top eight factors as the main controlling factors affecting oil well production (oil layer thickness, half seam length, formation pressure, total liquid volume, casing pressure, water content, number of fracturing sections, displacement arrangement relationship)

3.3. Classification and Prediction Based on Comprehensive Evaluation and Analysis of Main Control Factors

We select the main control factor as the independent variable in the discriminant prediction model, and the following results are obtained by discriminant analysis:

| Hashtag | 69 | 70 | 71 | 72 | 73 | 74 | 75 |
|---------|----|----|----|----|----|----|----|
| Actual category | Invalid well | Invalid well | Effective well | Invalid well | Effective well | Invalid well | Effective well |
| Forecast category  | Effective well | Invalid well | Effective well | Invalid well | Effective well | Invalid well | Effective well |

It can be seen from the table that only the first well in the seven wells is predicted to be wrong, and the prediction accuracy is close to 86%. The training data here is mainly to help us calculate the discriminant function to facilitate the prediction. So this article does not pay attention to the fitting effect, only focus on the prediction effect.

4 Conclusion

(1) The parameters affecting the production of oil wells are various, and it is very unrealistic for us to analyze all the factors directly. Therefore, people usually choose the parameters and choose the main factors, so that they will not have a big impact on the results but greatly reduce the workload.

(2) Discriminant classification of oil wells by discriminant analysis is an effective method for classification identification of sample wells. The prediction results can be used to identify and judge oil wells, so that it can effectively determine whether the well has value.
Acknowledgments

This thesis was supported financially by the National Science and Technology Major Demonstration Project 70 (2017ZX05070).

References

1. Yusuf, S., Peto, R., Lewis, J., Collins, R., & Sleight, P. (1985). Beta blockade during and after myocardial infarction: an overview of the randomized trials. Progress in Cardiovascular Diseases., 7(5):35-371.
2. Edgeworth, F. Y. (1923). Statistics of examinations. Journal of the Royal Statistical Society., 6(1):9-60.
3. SPEARMAN, & C. (1913). Correlations of sums or differences. British Journal of Psychology, 1904-1920, 5(4):417-426.
4. Levner, E., Alcaide, D., & Sicilia, J. (2007). Multi-attribute Text Classification Using the Fuzzy Borda Method and Semantic Grades. International Workshop on Fuzzy Logic & Applications: Applications of Fuzzy Sets Theory.
5. Sun, L. (2011). The Evaluation of the Construction Degree of Dalian’ Technology Innovation System Based on Fuzzy Borda Method. International Conference on Management & Service Science.
6. Rienzi, L., Ubaldi, F., Iacobelli, M., Ferrero, S., Minasi, M. G., & Martinez, F., et al. (2002). Day 3 embryo transfer with combined evaluation at the pronuclear and cleavage stages compares favourably with day 5 blastocyst transfer. Human Reproduction, 17(7):852-1855.
7. Sullivan, G. J., Ohm, J., Han, W. J., & Wiegand, T. (2013). Overview of the high efficiency video coding (hevc) standard. IEEE Transactions on Circuits & Systems for Video Technology, 22(12):1649-1668.
8. Song, B., Hao, S., Murakami, S., & Sadohara, S. (1996). Comprehensive evaluation method on earthquake damage using fuzzy theory. Journal of Urban Planning & Development, 122(1):1-17.