



Factor Analysis-Based Optimal Selection of Rock-Breaking Bit Applied in Deep Layer of Songliao Basin

Zuguang Li (Corresponding author), Yinghu Zhai & Xu Zhang

MOE Key Laboratory of Petroleum Engineering in China University of Petroleum

Beijing 102249, China

E-mail: lizuguang2009@gmail.com

Abstract

Correctly choosing rock-breaking bit is critical in oil drilling. The optimal model of selecting a bit has been established based on the Factor Analysis Theory. Through selecting primitive variables, and using SPSS (Statistical Package for the Social Sciences) to get factor loading matrix, factor rotation and factor score, we have reasonably evaluated and optimized the selection of rock-breaking bits applied in deep layer drilling of Songliao Basin. The calculated results are consistent with the observation from actual applications. It indicates that the method studied here is reasonably reliable and valuable for broader applications.

Keywords: Bit selection, Factor analysis, Songliao Basin, Volcanic rock

1. Introduction

The deep layers of Songliao Basin are rich in natural gas. It is imperative to extensively explore the natural gas in this region (Zhao, Wenzhi, 2004, p. 45-51). However, the geological structure of this region is complex. Complex volcanic system was broadly developed (Yang, Minghe, 2008, p. 76-87). Therefore, to achieve a safe, efficient and low-cost oil drilling process has been the major focus for scientists and engineers for a long time. One of the critical issues is correctly choosing rock-breaking bits.

Good progress has been made in the research of bit selection. Many selection modes have been proposed (Zhang, Hui, 2007, p. 1-5). Bit selection involves many variables, and is essentially an issue of multivariate statistical analysis. Many multivariate methods have been developed. Factor analysis is a widely applied multivariate statistical analysis method (He, Xiaoqun, 2004, p. 167-180; Bryan F J Manly, 1986, p. 1-15; G A F Seber, 1984, p. 21-27). In view of the actual characteristics of bit selection in oil drilling, the author has developed a factor analysis-based bit selection method.

2. Principle and model of factor analysis

Factor analysis is a multivariate statistical method which treats a system of multiple variables of intricate relationships by a few generic variables (He, Xiaoqun, 2004, p. 167-180; Bryan F. J. Manly, 1986, p. 1-15). The basic idea is to group the variables according to their correlations, i.e. group those highly correlative variables in the same group and lightly correlative variables in different groups. Each group represents a basic structure which is called “common factor”, and is expressed with an unobservable generic variable.

Suppose that there are n samples and p observation indexes per sample, where the observation indexes have significant correlations. Let X represent the standardized variable vector and F represent the standardized common factor. If both X and F satisfies the following conditions:

(1) X (observable random vector) = $(X_1, X_2, \dots, X_p)$, $E(X)$ (mean vector) = 0, $\text{cov}(X)$ (covariance matrix) = Σ , and Σ (covariance matrix) = R (correlation matrix).

(2) F (unobservable variable) = $(F_1, F_2, \dots, F_m)$, ($m < p$), and the components of vector F are mutually independent.

(3) $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$, is independent from F , and the components of ε are mutually independent.

“ $X = AF + \varepsilon$ ” is a basic model of factor analysis (He, Xiaoqun, 2004, p. 167-180).

3. Procedure of factor analysis-based bit selection

3.1 Selection of primitive variables and their standardized processing

Select variables that represent different types of bits at different geological layers, such as bit pressure, drilling time, drilling footage and unit price of bit. The established matrix (X') of primitive variables is as following expression.

$$\begin{cases} X'_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + \varepsilon_1 \\ X'_2 = a_{21}F_1 + a_{22}F_2 + \cdots + a_{2m}F_m + \varepsilon_2 \\ \cdots \cdots \cdots \\ X'_p = a_{p1}F_1 + a_{p2}F_2 + \cdots + a_{pm}F_m + \varepsilon_p \end{cases} \quad (1)$$

Here, F_i represents the i^{th} common factor, and a_{ij} represents factor loading.

In order to avoid the influence of different dimension and magnitude to the final statistical results, the primitive variables need to be standardized as the following (2).

$$X_i = \frac{X'_i - \overline{X'_i}}{\sqrt{\text{var}(X'_i)}} \quad (2)$$

3.2 Determination of factor loading matrix

There are many methods for the determination of factor loading matrix, such as main-component analysis, principal axis factor method, least square method and maximum likelihood method (He, Xiaoqun, 2004, p. 167-180). The principal axis factor method is used in this paper.

From formula (2), we can determine the matrix X -related correlation matrix R as following formula (3).

$$R = AA' + \Sigma\varepsilon \quad (3)$$

Here, A represents factor loading matrix, and $\Sigma\varepsilon$ represents a pair of diagonal matrix with its diagonal elements being the variance of special factors thereof.

Assume that $\lambda_1^* \geq \lambda_2^* \geq \cdots \geq \lambda_m^*$ are the characteristic roots of correlation matrix R , and $\gamma_1^*, \gamma_2^*, \dots, \gamma_m^*$ are corresponding orthonormalized eigenvectors. Therefore, an axis factor of factor loading matrix A is determined by the following formula (4).

$$\tilde{A} = (\sqrt{\lambda_1^*}\gamma_1^*, \sqrt{\lambda_2^*}\gamma_2^*, \dots, \sqrt{\lambda_m^*}\gamma_m^*) \quad (4)$$

3.3 Factor rotation

Factor rotation is determined based on the varimax rotation method in this paper.

First, we take the plane of two common factors into account. Assume that the factor loading matrix is represented by the following expression (5).

$$A = \begin{Bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \cdots \cdots \cdots \\ a_{p1} & a_{p2} \end{Bmatrix} \quad (5)$$

Let T be an orthogonal matrix by the following expression (6).

$$T = \begin{Bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{Bmatrix} \quad (6)$$

Let $B = (b_{ij}) = AT$, then the following expression (7) is obtained.

$$B = \begin{Bmatrix} a_{11} \cos \phi + a_{12} \sin \phi & -a_{11} \sin \phi + a_{12} \cos \phi \\ a_{21} \cos \phi + a_{22} \sin \phi & -a_{21} \sin \phi + a_{22} \cos \phi \\ \cdots \cdots \cdots \\ a_{p1} \cos \phi + a_{p2} \sin \phi & -a_{p1} \sin \phi + a_{p2} \cos \phi \end{Bmatrix} \quad (7)$$

The squared values of two lines of elements in the rotated loading matrix are standardized to 0 and 1. Therefore, The sample variances V_1 and V_2 of the above-mentioned squared values ($b_{11}^2, b_{21}^2, \dots, b_{p1}^2$) and ($b_{12}^2, b_{22}^2, \dots, b_{p2}^2$) should be maximized. For this reason, the orthogonal rotation angle (Φ) is required to satisfy the equation of “ $V_1+V_2=\max$ ”, that is

$$V = \sum_{j=1}^2 \left[\frac{1}{p} \sum_{i=1}^p (b_{ij}^2)^2 - \left(\frac{1}{p} \sum_{i=1}^p b_{ij}^2 \right)^2 \right] = \max \quad (8)$$

As a general rule, take two factors every time and turn all of them into rotating pairs when the number of common factors is more than two.

3.4 Calculation of factor score and comprehensive analysis

The regression equation (9) is established with the common factor as the dependent variable and the primitive variable as the independent variable.

$$F_j = \beta_{j1}X_1 + \beta_{j2}X_2 + \dots + \beta_{jp}X_p \quad (9)$$

Based on the least-square regression approach, the estimated value of F is determined by the equation “ $F=A'R^{-1}X$ ”. Here, A' represents the transpose of the rotated factor loading matrix; R^{-1} represents the inverse matrix of the primitive variable-related matrix; and X represents the primitive variable vector.

The applications of different types of bits at different layers are evaluated on the basis of the weighted sum of all factor scores for each bit. The bits are in turn optimally selected based on the obtained data.

4. Examples of calculation

4.1 Geologic characteristics of deep layers in the north part of Songliao Basin

Songliao Basin, in the northeast part of China, is a large-scale oil-contained meso-cenozoic continental basin. The basin basement consists of the middle-kata metamorphic rock series of Proterozoic and early Paleozoic eras, and low metamorphic rock series of late Paleozoic eras and volcanic rock. From bottom to top, it consists of the upper Jurassic Huoshiling formation, lower Cretaceous Shahezi formation, Yingcheng formation, Dengloulou formation, Quantou formation, upper Cretaceous formation, Yaojia formation, Sifangtai formation, Mingshui formation and Cenozoic formation. The formation from Shahezi to Yingcheng is the deep major hydrocarbon source rocks, and Dengloulou formation is the minor hydrocarbon source rocks.

4.2 Examples for bit optimal selection

Different types of bits applied in the deep Dengloulou formation at Xujiaweizi region of Songliao Basin are statistically analyzed. The applications of bits in the well X3 at this region are shown in table 1.

The bits applied in the Dengloulou formation are optimally selected based on the method studied in this paper. The factor analysis can be achieved by the “Factor Analysis module” in SPSS. The SPSS-analyzed factor scores of various bits are shown in table 2.

The various bits are optimally selected on the basis of the weighted sum of all factor scores for each bit. The results are shown in table 3.

The results in table 3 are basically consistent with the observation from actual applications. It indicates that the method is practicable.

5. Conclusion

(1) Factor analysis is an important multivariate statistical analysis method, and has a bright future for applications in practical engineering tasks.

(2) The above-mentioned SPSS is powerful, convenient and easy to use.

(3) The factor analysis-based optimal selection results of bits basically reflect the actual on-site use. It indicates that the method is relatively reliable and valuable for broader applications.

References

- Bryan F. J. Manly. (1986). *Multivariate Statistical Methods: A Primer*. New Mexico: Chapman & Hall/CRC Press LLC, p. 1-15.
- G. A. F. Seber. (1984). *Multivariate Observations*. New Jersey: John Wiley & Sons, Inc., p. 21-27.
- He, Xiaoqun. (2004). *Multivariable Statistical Analysis*. Beijing: Press of People's University of China, p. 167-180.

Yang, Minghe. (2008). *Study on lithology identification and anti-drilling properties of volcanic in formation of Xujiaweizi*. Beijing: China University of Petroleum, p. 76-87.

Zhang, Hui & Gao Deli. (2007). Review on drill bit selection methods. *Oil Drilling and Production Technology*, 27(4), 1-5.

Zhao, Wenzhi, Li Jianzhong, Zou, Caineng, et al. (2004). Geologic characteristic and exploration direction of Songliao Basin. In Jia, Chengzao (Eds.), *Proc. of Gas Exploration of Songliao Basin Conference*. Beijing: Petroleum Industry Press, p. 45-51.

Table 1. The applications of bits in the well X3 at Denglouku formation

Serial number	Type of bit	Drilling footage (m)	Drilling time (h:min)	Drilling speed (m/h)	Drilling pressure (kN)	Rotation speed (r/min)	Displacement (L/s)	Pumping Pressure (MPa)
1	HJS537GR	129.16	105:55	1.22	260	132	46	22.0
2	HJS537GR	151.60	119:15	1.27	260	132	46	22.0
3	HJS537GR	116.66	129:10	0.90	280	65	45	20.0
4	HJS537GR	149.31	121:05	1.23	280	65	45	20.0
5	HJT547G	40.69	32:10	1.26	200	50	30	19.0
6	HJT617GL	115.12	79:15	1.45	180	50	30	17.0
7	HJT617GL	74.41	52:05	1.43	180	50	30	18.0
8	TD53HR	130.43	56:20	2.32	180	78	32	18.0
9	TD53KPRDH	85.87	64:25	1.33	180	78	32	18.0
10	HJS537GR	187.42	105:25	1.78	180	78	32	17.0
11	HJ517G	163.59	81:15	2.01	180	70	33	17.7
12	HJ517G	158.05	103:25	1.53	180	70	34	17.0

Table 2. The SPSS-analyzed factor scores of various bits applied in Dengloulou formation

Serial number	Type of bit	Main factor 1	Main factor 2	Main factor 3
1	ATJ33	0.572	-0.652	0.368
2	ATJ44	1.685	0.365	0.369
3	ATM33	-0.438	0.655	-0.358
4	DQ506	0.537	0.325	2.355
5	FG45YODPD	0.580	0.365	0.258
6	H617	2.367	0.562	0.354
7	HA617L	0.235	0.255	0.355
8	HA617LC	0.234	-0.352	-0.158
9	HA637LC	0.531	1.365	-0.269
10	HJ517G	0.756	2.365	0.452
11	HJ617	-0.747	-0.258	-0.254
12	HJ637	0.721	0.354	0.324
13	HJS537GR	1.256	0.324	2.325
14	HJT547G	0.835	1.185	0.254
15	HJT617GL	0.243	1.365	-0.365
16	HJT737GDG	-0.543	0.368	-0.487
17	SC-279	-0.124	-0.369	-0.452
18	TD53HR	0.281	0.254	-0.154
19	TD53KPRDH	-0.251	0.324	0.104

Table 3. The optimal selection results of bits

Formation segment	Selection results
Dengloulou 4 th	HJT547G, HJS537GR, HJ517G
Dengloulou 3 rd	HJS537GR, HJ517G, HJT617GL, HJT547G
Dengloulou 2 nd	HJS537GR, HJT617GL, HJ517G