

PAPER • OPEN ACCESS

Research on energy saving of Coalbed Methane Field pumping wells based on Data Mining

To cite this article: Xueqin Li 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **310** 032032

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

Research on energy saving of Coalbed Methane Field pumping wells based on Data Mining

Xueqin Li^{1,2*}

¹ Petrochina Huabei Oilfield CBM Branch Company, Jincheng, Shanxi Province, 048000, China

²CBM pilot test base of China National Petroleum Corporation, Jincheng, Shanxi Province, 048000, China

*Corresponding author's e-mail: mcq_lxq@petrochina.com.cn

Abstract. Generally, the CBM well has the characteristics of large variation of the water rate and unstable pumping load, corresponding equipment type, status and parameters adjusted frequently, the utilization rate of energy is generally low, however, the energy transfer relationship of pumping unit system is complex and the coupling between parameters is strong, the theoretical calculation method requires high energy consumption calculation parameters for each node, and the effective dynamic analysis cannot be realized; the experience analysis method has high workload and high requirement for analysts. In southern Qinshui basin CBM Wells, for example, by analyzing the relationship between production parameters and the energy consumption of each node, a multi variable data mining algorithm is constructed for the energy consumption index of the common pumping unit system, according to the energy consumption data characteristics of pumping unit, the single well is classified, and the adjustment optimization strategy of pump replacement, pumping-unit replacement, intermittent pumping and equipment maintenance is put forward according to the different types of energy consumption bias. Verified by practice, the application of this technology can improve the timeliness and accuracy of energy consumption data analysis of CBM pumping well, effectively reduce the cost of energy consumption analysis on mechanical production Wells, and is helpful to improve the intelligent level of CBM gas field production.

1. Foreword

In the development of CBM (coalbed methane), the pumping system is the main power equipment of the gas field, and its efficiency is directly affecting the cost of development. The production of CBM Wells belong to the step-down drainage, during the early stages of production, the amount of water will be larger. After a period of drainage, the water rate will gradually decrease without coal seam water supply, even the phenomenon of dry pumping. In response to the characteristics of the big changes in water production, it is necessary to adjust the ground production equipment frequently. Once the adjustment is not in time, there will be a "waste one's talent on a petty job" situation[1]. In addition, some of the drainage Wells have a large amount of water, which is difficult to reduce. If the energy consumption analysis is ignored, the adoption of an unsuitable response method, such as continuous keep raising the stroke frequency, will lead to inefficient systems and a lot of wasted energy[2].

According to statistics, the average system efficiency of oil pumping units in the oil field of China is 27.6%, and only 0.6%~7.4% in the Southern of Qinshui Basin CBM gas field, which is far from the



pumping unit energy conservation standard of 25%, and the present situation of the low energy utilization needs to be improved. How to achieve the maximum economic benefit of energy saving, consumption reduction and system optimization? The energy consumption analysis of the pumping system is critical. For the analysis of energy consumption of pumping unit, mounts of scholars have done research, which can be classified into two aspects: based on theoretical calculation and empirical data. The theoretical calculation method is to calculate the loss of each node of the system accurately from the Angle of energy[3], and the empirical analysis method is based on the current, work graph, pump efficiency and other indicators to determine the energy conversion rate of each node[4].

These two kinds of methods used in CBM well pumping unit energy consumption analysis, with the following problems:

- The method of drainage in CBM field by continuously discharging water from coal seam to reduce pressure and achieve methane gas desorption, is different from the way in oil field which is used to maintain the formation pressure. At different stages, the pumping unit has a fluctuating load, and the corresponding node energy consumption is dynamic, so the node parameter is difficult to evaluate.

- The accurate model has higher parameter requirements, and the solution is generally more complex. Affected by various factors, the timeliness and practicability of the theoretical calculation method are poor.

- Empirical analysis relies on the degree of understanding of the analyst on the index data such as the balance rate, pump efficiency, the balance of the motor and so on. There is no fast energy evaluation for a lot of wells.

Based on data mining and big data analysis, this paper innovatively constructs a set of multi-dimensional scale mechanical production system optimization model, applies k-means clustering analysis to obtain independent sample groups with different energy consumption bias, clarifies the direction of energy saving measures, and provides new technical ideas for energy saving and consumption reduction of pumping unit system.

2. Data mining and big data optimization model

In addition to converting the energy consumption of the pumping unit into a payload. In the process of energy transfer from motor, belt, gearbox, connecting rod mechanism, rod pipe to pump, there is a large amount of useless loss. and the energy conversion efficiency of each part determines the overall efficiency of the pumping system[5]. As shown in figure 1, the system efficiency is decomposed into each node to pump efficiency, sucker rod efficiency, belt - gearbox efficiency, motor - belt efficiency, packing box efficiency, etc.

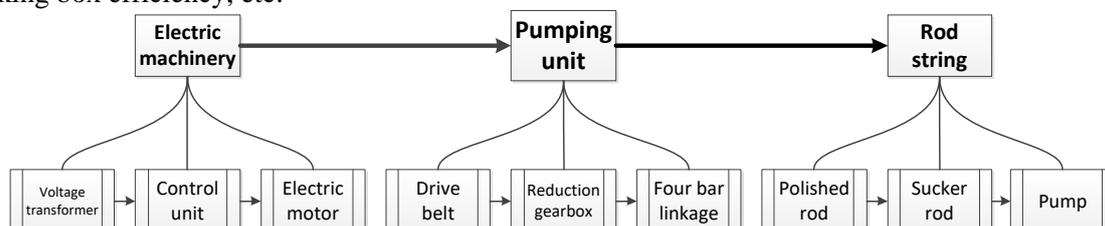


Figure 1. Schematic diagram of the energy flow of a pumping unit

2.1. The selection of energy consumption analysis index

The main consideration factors for energy consumption of ground system are matching degree of motor and pumping unit, matching degree of pumping unit production parameters and load, working condition of pumping unit ("five rate" of pumping unit), tightness of belt and packing. Underground factors mainly include eccentric wear of pipe string, leakage rate of pump, insufficient liquid supply, gas influence and so on. Nine indexes with high correlation of energy consumption and easy to obtain were used for substitution analysis: load utilization rate, submergence, wellbore pressure, liquid yield, pump discharge coefficient, unit consumption per ton of liquid of 100 meters, frequency of stroke,

degree of balance, and motor power utilization rate. The corresponding relationship between these indicators and the node or device factors that can be reflected is shown in table 1.

Table1 The corresponding relationship between the energy consumption indexes and the influencing factors

Ordinal number	Indicators	Node factors
1	Load utilization	polish rod、Sucker rod、 Tubing、 Pump
2	Depth of pump	Sucker rod、 Tubing、 Pump efficiency
3	Wellbore pressure	Pump
4	Liquid production capacity	Coalbed liquid supply capacity
5	Drainage coefficient of pump	Sucker rod、 Tubing、 Pump
6	Power consumption of liquid per ton per 100 meters	Overall system performance
7	Jig frequency	Sucker rod、 Tubing、 Motor
8	Degree of balance	Motor
9	Motor power utilization	Motor、 Control unit、 Transformer

2.2. Big data optimization model construction

In this study, we extracted various production parameters related to the energy consumption of pumping units, established the big data analysis model, and optimized the traditional algorithm for energy consumption analysis of pumping units. The big data analysis model is shown in figure 2:

- Extract the production parameters with higher correlation with unit consumption;
- Conduct principal component analysis to reduce its coupling;
- K-means clustering analysis was used to classify single well types.
- Different types of single well deviation analysis to determine the high energy consumption node;
- According to the high energy consumption node, the corresponding optimization measures are formulated.

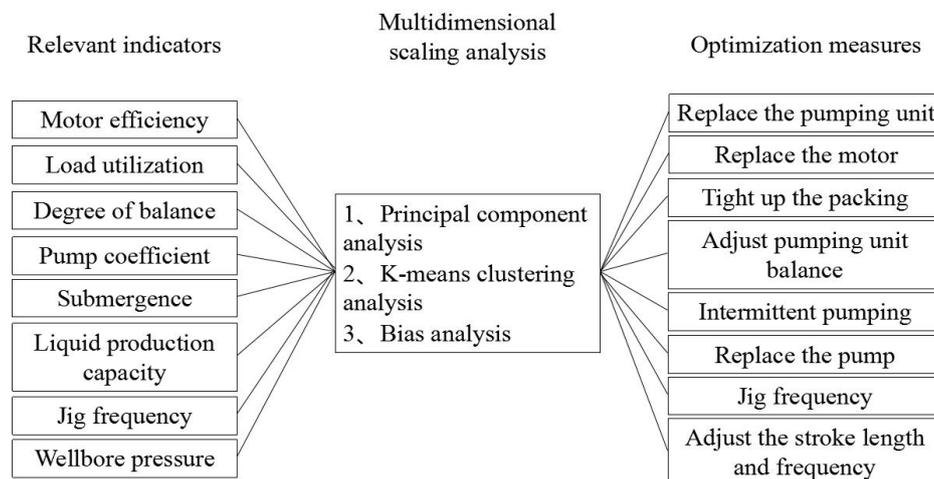


Figure 2. Schematic diagram of big data optimization model unit

2.3. Principal component analysis

These nine indexes reflect the node energy consumption of the pumping unit from different aspects, but the data has multiple dimensions and strong coupling. How to analyze the node energy consumption characteristics from the index data is actually a multi-parameter optimization problem. Taking the production performance data of 504 pumping Wells in the coalbed methane field of qinshui basin as an example, the energy consumption characteristics of a single well are extracted through data mining.

In the actual production, the quantitative calculation of energy consumption does not require much, which is qualitative judgment [9]. Due to strong coupling, the indicators are not independent. Here, principal component analysis is carried out to extract independent variables from data to achieve dimensional reduction. Firstly, principal component analysis was carried out on the above nine index data. Table 2 shows the calculation results modeled by SPSS software. Then three principal components which eigenvalue is greater than 1 were extracted, and nine indexes were converted into three principal components according to the conversion coefficient (table 3).

Table2 Analysis results of main causes of SPSS

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.276	25.294	25.294	2.276	25.294	25.294
2	1.978	21.978	47.271	1.978	21.978	47.271
3	1.440	15.997	63.269	1.440	15.997	63.269
4	0.900	10.004	73.273			
5	0.761	8.460	81.733			
6	0.660	7.329	89.061			
7	0.419	4.652	93.714			
8	0.338	3.755	97.468			
9	0.228	2.532	100.000			

Table3 SPSS main cause analysis of each parameter

	3 extracted component Matrix		
	Component		
	1	2	3
Jig frequency	0.612	-0.306	-0.037
Wellbore pressure	-0.115	0.400	0.616
Liquid production capacity	0.328	0.765	-0.373
Load utilization	0.786	0.381	0.060
Depth of pump	-0.186	0.464	0.724
Degree of balance	-0.533	0.018	-0.250
Discharge coefficient	-0.347	0.624	-0.571
Motor power utilization	0.785	0.234	-0.014
Power consumption of liquid per ton per 100 meters	0.327	-0.578	-0.056

2.4. K-means clustering analysis

Although the principal component analysis simplifies the indicators and reduces the coupling, it is difficult to explore the practical significance [10]. After projecting the principal component index data into the multi-dimensional space, as shown in figure 3, it is impossible to distinguish and discover the characteristics of samples by senses. It is necessary to get the classification of different energy consumption characteristics through cluster analysis [11].

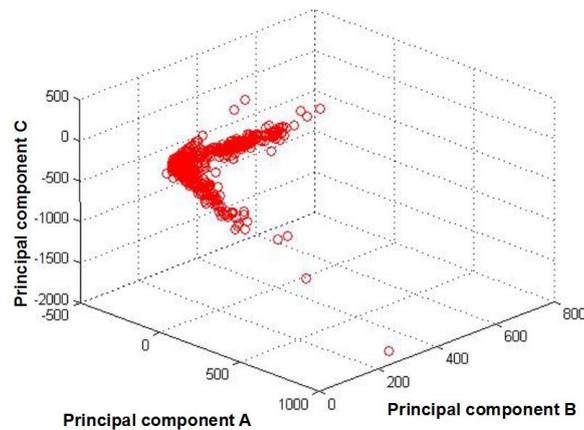


Figure 3. Multidimensional spatial display of energy consumption index data

For multi-index classification, there is still no clear and unique method to solve the problem of how to determine the number of the best classification [12]. The ideal classification scheme needs to maximize the distance between classes and minimize the distance between members in the class. In this paper, k-means clustering algorithm based on Euclidean distance model is adopted for analysis.

Considering from practicality, class spacing and class stability, the classification number we need is within the range of 3 to 5, that is, plan A is divided into 3 categories, plan B is divided into 4 categories, and plan C is divided into 5 categories. The enumeration method is adopted to calculate whether the original sample classification changes under certain data increment, so as to measure the stability of classification. The most stable classification scheme is the scheme with the highest degree of satisfaction. In this paper, we use Cross Validation method to compare three kinds of classification schemes of clustering results. After the classification is obtained by using the clustering calculation of 50% of the 504 sample records, 30%, 40% and 50% (152, 203 and 252) are used as the increment to form 3 sample populations for clustering calculation respectively. The corresponding results are shown in table 4. From the perspective of stability, the category of plan A is the most stable and the classification is reasonable.

Table4 Cluster analysis evaluation

Classification scheme	Classification number	Average distance between classes				Degree of polymerization				Stability of clustering results			
		25%	50%	75%	100%	25%	50%	75%	100%	50%	75%	100%	Mean change
A	3	137	140	357	357	84	84	137	175	0	11	13	12
B	4	449	369	925	932	78	71	85	94	39	209	1	83
C	5	424	478	1036	1038	60	67	81	91	19	51	1	24

Three categories are selected as parameters for k-means clustering calculation. Principal component samples and clustering center results are shown in figure 7. Three different colors are used to mark different classifications in the figure. It can be seen that the sample principal component data of the three categories are basically the extension along a certain direction, which can be interpreted as the performance of the same category with the same eigenvalue.

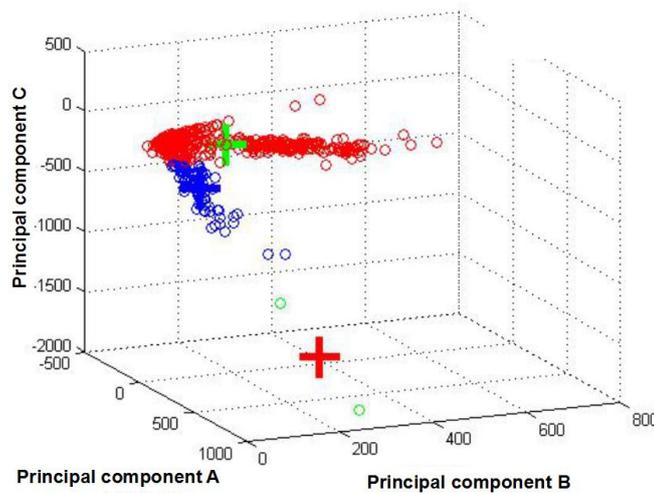


Figure 4. K-means clustering analysis results

2.5. Multidimensional scale bias analysis

According to the three categories obtained by the above clustering analysis, multidimensional scale deviation analysis was carried out to evaluate each single well within a certain category, and the energy consumption characteristic deviation was defined:

- TypeI: the number of such type of wells is small and the characteristic is not obvious, there is no obvious energy consumption bias;
- TypeII: (FIG. 5, left) the Power consumption of liquid per ton per 100 meters is highly correlated with the power utilization rate of the motor and the displacement coefficient of the pump, while it is less affected by other factors;
- TypeIII: (FIG. 5, right) the Power consumption of liquid per ton per 100 meters is highly correlated with the degree of balance, and the power utilization rate of the motor is closely related to the load utilization rate.

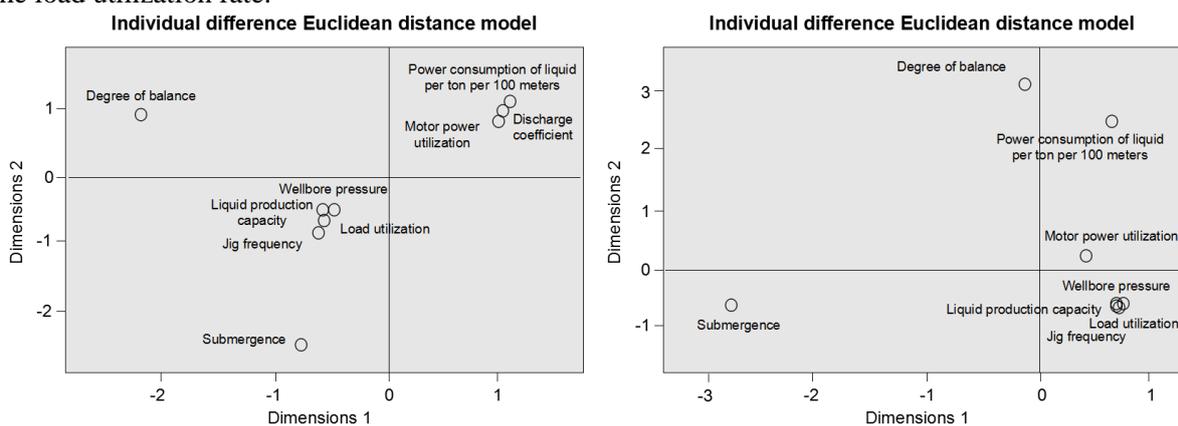


Figure 5. Bias analysis results of type II energy consumption well (left) and type III (right)

3. Optimization measures

According to the results of the bias analysis, the optimization and maintenance work with different emphasis was carried out for different classified Wells. The system is optimized from three aspects: the matching degree of water quantity of single well and equipment, the state of equipment and working system. The specific optimization strategies and measures are shown in table 5.

Table5 Optimization strategy of single well with different types of energy consumption

Type	Optimization strategy	Well number	The implementation
Type I	Abnormal	2	Pipe string leakage operation 2 wells
Type II	Adjustment of working system	422	Intermittent pumping was performed in 115 Wells; Slow down stroke frequency in 213 Wells; Replace 5 Wells with small pumps.
Type III	Equipment state	80	Adjust pumping balance in 45 Wells; Adjust pumping unit in 10 Wells; Replace the inverter motors in 25 Wells.

- TypeI: According to the empirical analysis, there are problems of string leakage in two Wells. After operation and construction, the damaged oil pipe is replaced and the index returns to normal;
- TypeII: There are a lot of single Wells with insufficient liquid supply in coalbed methane field, and the methods of interval pumping, adjustment of pump diameter, length and frequency of stroke are mainly used to reduce the operating time of the pumping unit, so as to reduce unit energy consumption. Relying on the above - well automation system, the single well in the south of qinshui basin can realize large - area intelligent drainage and production control. According to the working system of pumping units in 422 Wells, 115 single Wells were randomly pumped and the stroke times of 213 single Wells were reduced after demonstration;
- TypeIII: Check the working condition of 80 single well pumping units, carry out adjustment and balance, packing and other routine maintenance on the well, and effectively reduce the unit consumption.

4. Conclusions

- The relationship between the Power consumption of liquid per ton per 100 meters and relative production indexes is studied for CBM pumping Wells. Using big data analysis to evaluate the energy consumption characteristics of a large number of single Wells is an effective way to reduce the cost of dynamic energy consumption analysis of coalbed methane Wells, which makes up the deficiency of theoretical calculation method.
- For multi-factor and high-coupling energy consumption related indicators, it is proposed for the first time to comprehensively apply principal component analysis, k-means clustering analysis, multi-dimensional scale deviation analysis and other data mining techniques for analysis, which can reduce the work intensity of empirical analysis method and help improve the intelligent control level of gas fields.
- According to the data mining analysis, the main reasons affecting the high unit consumption of Southern Qinshui Basin CBM pumping well are unstable water quantity and low load. It is difficult for conventional pumping equipment to achieve efficient operation level, so it is urgent to find new energy saving pumping equipment.

Acknowledgements

Program of fund: Key state science and technology project: Key Technologies for High Efficiency Production Increase and Drainage of Coalbed Methane (2016ZX05042); Demonstration Project for High-rank Coalbed Methane Efficient Development in Qinshui Basin(2017ZX05064); Project funded by CNPC (2018D-5006-37).

References

- [1] LIU Shiqi, ZHAO Xianzheng, SANG Shuxun, YANG Yanhui, LI Mengxi, HU Qiuqia, YANG Yanlei. Cooperative control of working fluid level and casing pressure for coalbed methane production:a case study of Fanzhuang block in Qinshui Basin[J]. Acta Petrolei Sinica, 2015, 36(s1): 97-108.

- [2] NING Qian. Optimal intermodulation design of pumping unit type matching [J]. *Energy Conservation in Petroleum & Petrochemical Industry*, 2015, 56(10): 1-4.
- [3] DONG Chunhua, YANG Yingchun. Analysis of factors affecting pumping efficiency of pumping well and Countermeasures [J]. *Engineering Technology*, 2016, 8(12): 240-240.
- [4] LI Jie. Oil well pumping conditions analysis and energy consumption evaluation method based on micro-control charts [D]. Qinhuangdao: Yanshan University, 2016.
- [5] WANG Hai, YU Jia, NI Yongjun, SONG Baiqiang. Analysis of the eccentric wearing prevention of pumping unit for CBM wells in south Qinshui Basin [J]. *China Coalbed Methane*, 2014, 11(6):41-43.
- [6] XING Mingming, DONG Shimin, CUI Yang, CHEN Peiyi. Simulation model of transmission efficiency of the belt-driving systems of beam pumping units [J]. *Engineering Mechanics*, 2013, 30(7):242-247.
- [7] WANG Shiliang. Causes and analysis of common faults in pumping unit reducer [J]. *Petro & Chemical Equipment*, 2014 17(4):58-59.
- [8] CHEN Xiaohe, CHEN Juncong, YUAN Zhou. The relationship between balance ratio and energy consumption of pumping well [J]. *Oil-Gas Field Surface Engineering*, 2014, 33(1):5-6.
- [9] TIAN Lianyu. Determination of energy consumption node standard for daily management of pumping well [J]. *Technology Supervision in Petroleum Industry*, 2014, 30(6): 28-31.
- [10] MA Guojun, WANG Shuibo, PEI Qingqi, ZHAN Yang. Research on parallel coordinate visualization technology based on principal component analysis and K-means clustering [J]. *Chinese Journal of Network and Information Security*, 2017, 3(8): 18-27.
- [11] ZHANG Xuefeng, LIU Peng. An improved K-means algorithm by weighted distance based on the maximum between-cluster variation [J]. *Journal of Shandong University(Natural Science)*, 2010, 45(7):28-33.
- [12] WANG Xiaochu, WANG Shitong, BAO Fang, JIANG Yizhang. Intra-class-distance-sum-minimization based classification algorithm [J]. *Journal of Electronics & Information Technology*, 2016, 38(3):532-540.