# Method of automatic interpretation of information about the geological structure in the process of exploratory wells drilling

**Vladimir Morkun**

*Vice-Rector for research, Doctor of Science,*
*Professor of Computer Science, Automation and Control Systems department,*
*Kryvyi Rih National University*

**Vitaliy Tron**

*PhD, Senior lecturer of Computer Science, Automation and Control Systems department,*
*Kryvyi Rih National University*

**Dmitriy Paraniuk**

*Postgraduate student of Computer Science, Automation and Control Systems department,*
*Kryvyi Rih National University*

Abstract
Method of geologic model forming based on the information on the operational characteristics of the drilling process using the results of the clustering of such process characteristics as torque and speed of drilling is proposed.
Keywords: AUTOMATION OF DRILLING, DATA CLUSTERING, THE GEOLOGICAL MODEL

In problem solving of the energy-efficient automated process control at different stages of extraction and processing of iron ore a major role plays the presence of operational information on disturbing factors [1-6]. Thus, information on the process course can be obtained not only by direct measurement, but also on the basis of the interpretation of indirect signs [1-2]. Studies revealed a high degree of correlation between the parameters obtained as a result of monitoring of the drilling rig, and the data of geophysical well testing [1]. During the study the monitoring of the following parameters: drilling speed, rotation speed, traction, torque was carried out. In [2] besides named parameters the possibility of an axial load using in the identification of the ore rocks geological structure in the drilling process is studied. Thus, the recognition of ore material varieties during drilling is advisable to carry out by several characteristics using static clustering characteristics of the drilling process at the different depths: speed and torque; with pre-normalization of data to improve the efficiency of clustering methods [7, 9].

Among the crisp clustering methods the K-mean and K-medoid methods are most effective, which determine the fitting of each set of ore piece characteristics to one of *c*-clusters to minimize a sum of squares within the cluster [10,11]

$$\sum_{i=1}^{c}\sum_{k\in A_i}\left\|x_k - v_i\right\|_2 \qquad (1)$$

where $A_i$ – is a set of objects (reference points) in *i*-th cluster and $v_i$ - is a mean as indicated by the cluster points *i*. According to the algorithm of K-mean clustering the $v_i$ called as clusters centers

$$v_i = \frac{\sum_{k=1}^{N_i} x_k}{N_i},\ x_k \in A_i \qquad (2)$$

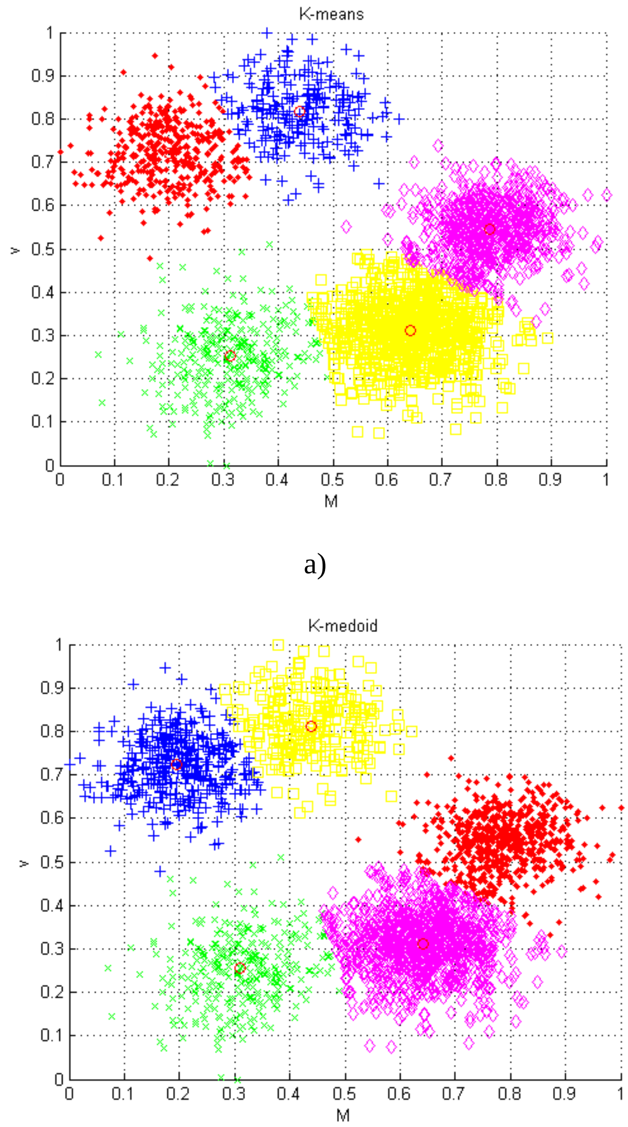
where $N_i$ – is the number of objects in $A_i$. The result of the clustering of the drilling process measured characteristics, performed by K-means algorithm is shown in Fig. 1 *a*. According to the K-medoid clustering method the clusters centers are the average data nearest objects in a single cluster $V=\{V_i \in X | 1 \le i \le c\}$ [10,11]. The clustering result of iron ore raw materials samples performed by the K-medoid algorithm is shown in Fig. 1*b*. Additional quality assessment of clustering was performed using the following scalar measures of integrity. The partition coefficient (PC) measures the "overlap" between the clusters [11]

$$PC(c) = \frac{1}{N}\sum_{i=1}^{c}\sum_{j=1}^{N}(\mu_{ij})^2 \qquad (3)$$

where $\mu_{ij}$ – is the data point membership function *j* in cluster *i*. a) K-means algorithm; b) K-medoid algorithm

The segmentation coefficient (SC) [12]

$$SC(c) = \sum_{i=1}^{c}\frac{\sum_{j=1}^{N}(\mu_{ij})^m\left\|x_j - \upsilon_i\right\|^2}{N_i\sum_{k=1}^{c}(\mu_{ij})^m\left\|\upsilon_k - \upsilon_i\right\|^2} \qquad (4)$$




a)

b)

**Figure. 1**. The result of the drilling process characteristics clustering

The Xie-Beni Index (Xie-Beni, XB) defines the quantitative estimation of the ratio of the total variation in clusters and clusters separation [13]

$$XB(c) = \frac{\sum_{i=1}^{c}\sum_{j=1}^{N}(\mu_{ij})^m\left\|x_j - \upsilon_i\right\|^2}{N\min_{i,j}\left\|x_j - \upsilon_i\right\|^2} \qquad (5)$$

The Dunn's Index (DI): is used to identify the "compact and well-separated clusters" [7]

$$DI(c) = \min_{i\in c}\{\min_{i\in c, i\ne j}\{\frac{\min_{x\in C_i, y\in C_j} d(x,y)}{\max_{k\in c}\{\max_{x,y\in C} d(x,y)\}}\}\}$$

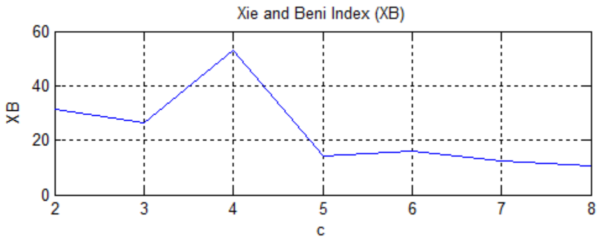
(6)

Alternative Dunn’s Index (ADI)

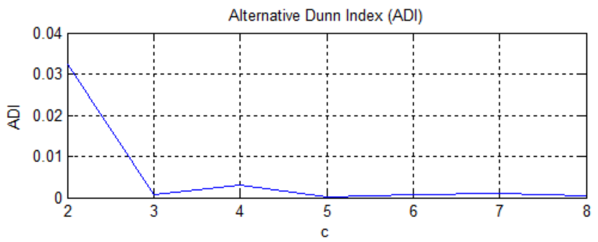
$$ADI(c) = \min_{i \in c}\{\min_{j \in c, i \neq j}\{\frac{\min_{x_i \in C_i, x_j \in C_j} \left| d(y,\upsilon_j) - d(x,\upsilon_j) \right|}{\max_{k \in c}\{\max_{x,y \in C} d(x,,y)\}}\}\}$$

(7)

A graphical representation of the dependences of above indicators of clustering quality on the number of clusters is shown in Fig. 2. The XB Index varies strongly enough (fig. 2*a*), which does not allow to determine the optimal number of clusters. Thus, starting with the value of 5clusters or more this index decreases relatively slower than on the interval of 2-5 clusters. The alternative Dunn's Index value, dependence of which on the number of clusters is shown in Fig. 2b, decreases strongly enough with increasing the number of clusters to 5, for more of clusters this index shows a slight increase.




a)




b)

**Figure. 2.** Clustering quality indexes: *a*) XB; *b*) ADI

The comparison results of clustering methods by indicators considered above are given in Table 1

**Table 1.** The clustering methods comparison results

| | PC | SC | XB | DI | ADI |
|---|---|---|---|---|---|
| K-means | 1 | 0.4785 | 7.463 | 0.0099 | 0.0010 |
| K-medoid | 1 | 0.4789 | 12,354 | 0.0047 | 0.0014 |

The studies have shown that the best clustering results are achieved with the process monitoring results distribution by K-Means method for five clusters, which corresponds to the results of geological sampling of the nearest to the investigated wells drill holes.

**References**

1. Scoble M. J., Peck J., Hendricks C. (1989). Correlation between Rotary Drill Performance Parameters and Borehole Geophysical Logging. *Mining Science and Technology*, No8, p.p. 301-312.
2. Segui, J. B., Higgins M. (2002). Blast Design Using Measurement While Drilling Parameters. *Fragblast*, Vol. 6, No3 – 4, pp. 287 – 299
3. Morkun V., Morkun N., Tron V. (2015). Identification of control systems for ore-processing industry aggregates based on nonparametric kernel estimators, *Metallurgical and Mining Industry*, No1, pp. 14-17.
4. Morkun V., Tron V. (2014). Automation of iron ore raw materials beneficiation with the operational recognition of its varieties in process streams. *Metallurgical and Mining Industry*, No6, p.p. 4-7.
5. Morkun V., Tron V. (2014). Ore preparation energy-efficient automated control multi-criteria formation with considering of ecological and economic factors, *Metallurgical and Mining Industry*, No5, p.p. 8-11.
6. Morkun V. S., Morkun N. V., Pikilnyak A.V. (2014). Iron ore flotation process control and optimization using high-energy ultrasound, *Metallurgical and Mining Industry*, No2, p.p. 36-42.
7. B. Balasko , J. Abonyi and B. Feil. (2010). Fuzzy clustering and data analysis toolbox for use with MATLAB, Available at: http://www.fmt.vein.hu/softcomp
8. Morkun V., Morkun N., Pikilnyak A. (2014). The gas bubble size distribution control formation in the flotation process, *Metallurgical and Mining Industry*, No4 , p.p. 42-45.
9. Pospelov D.A. *Nechetkiye mnozhestva v modelyakh upravleniya i iskusstvennogo intellekta* [Fuzzy sets in management models and artificial intelligence], Moscow, Nauka, 1986
10. Shtovba S.D. Vvedeniye v teoriyu nechetkikh mnozhestv i nechetkuyu logiku [ntroduction to the theory of fuzzy sets and fuzzy logic], Available at: //http:matlab/exponenta.ru/fuzzylogic/book1

11. Bezdek J. C. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, 1981.
12. Bensaid A.M., Hall L.O., Bezdek J.C., Clarke L.P., Silbiger M.L., Arrington J.A., Murtagh R.F. (1996). Validity-guided (Re) Clustering with applications to imige segmentation. *IEEE Transactions on Fuzzy Systems*, No.4, p.p. 112-123.
13. Xie X. L. and Beni G. A. (1991).Validity measure for fuzzy clustering. *IEEE Trans. PAMI*, Vol.3, No8, p.p. 841-846.