

## Case study in North Iraq: Cluster analysis of well logs using IP program

Muhammed Adil Ismail<sup>1</sup>, Ali Hasan Abdulla<sup>2</sup>, Fawzi Al-beyati<sup>3</sup>

<sup>1,2</sup> Petroleum Engineering Department, University of Kirkuk, Iraq

<sup>3</sup>Kirkuk Technical College, North Technical University, Iraq

---

### ABSTRACT

Diminish of error ratio to expecting the lithology of reservoir formation is considered as mean concerning of geology nowadays. The aim of research is to minimize the wrong reading data extract the lithology facies of the reservoir by preparing using Interactive Petrophysics program (IP) to make cluster analysis of multi well logs data. Log data of Kirkuk oil field used in this program included: Gamma ray (GR), sonic, neutron (N), bulk density (BD), Caliper and Resistivity logs. The basic concept of this method is to utilized K-mean clustering theory. Several wells have been conducted to emphasize this method which shows dropping in error ratio of confining the lithology and hydrocarbon existing as 41 wells have precisely predicted of methane prevailing.

---

**Keywords:** Cluster analysis, Well log, Petrophysics, Lithology, Gamma ray

---

### *Corresponding Author:*

Mohammed Adil Ismail  
Department of Petroleum Engineering  
Collage of Engineering, Kirkuk University  
Kirkuk, Iraq, 36013  
E-mail: [muhammedchakmakchi@uokirkuk.edu.iq](mailto:muhammedchakmakchi@uokirkuk.edu.iq)

---

## 1. Introduction

Cluster analysis is a category of statistical approaches whose main purpose is to group objects (e.g., participants, products, or other entities) based on their features. It's a way to combine documents based on characteristics that make them identical. If illustrated symmetrically, the points inside clusters will be close together, whereas the interval among clusters will be further apart. Information clustering is a technique of defining natural clusters or clusters within multi - dimensional data dependent on some measure of similarity [1], [2].

A researcher can be presented with a vast number of findings that may not be relevant until they are grouped into manageable categories. Cluster analysis will critically execute this data reduction technique by minimizing information from the entire sample population to information on groups [3]. Cluster analysis has no theoretical base from which to draw inferences from a survey to a group, and several claim that it is just an exploratory tool. Nothing ensures novel solutions, since participation of clusters with any number of solutions depends on many of the elements of the process, and many multiple approaches could be reached by altering one or more aspects.

### 1.1. Cluster analysis in geophysics

Clustering is considered as one of the methods to establishing petrophysical correlations [4]. Fundamentally, clustering analysis is a technique to categorize data according to a combination of petrophysical, geophysical, petrographic, geological, or other features. It can be also estimating the optimum length of drilled well using cluster analysis approaches through nominated an optimum petrophysical properties and the area of high potential hydrocarbon accumulation [5]. The most intuitive clustering example is rock typing. Since the totality of points belonging to a single cluster exhibits some physical and/or geological generality, there is reason to believe that the petrophysical bonds will be harder expressed within a cluster than in the total dataset. Note that

when identifying lithologies, interpreters primarily consider the relative changes in geophysical parameters, i.e., changes in logging curve signatures [6].

Several published sources describe the tasks of geological classifications that have been developed to predict the physical parameters of rocks using cluster analysis algorithms [7],[8],[9], however, in our work, we will use clustering for geomechanically modeling.

In the general case, this is a research problem, however, with all the advantages of the clustering approach, incorrect identification of a point (i.e., when a correlation for a cluster is applied to a point that is erroneously (or accidentally) assigned to another cluster) may cause problems. To reduce the negative effect of incorrect cluster identification, the so-called fuzzy clustering algorithms should be applied which essentially fit the weight of each cluster membership in each point; in controversial cases (unclear membership, i.e., of several clusters with close weights simultaneously), we can choose the most reasonable and consistent version of the correlation.

## 2. Material and methods

Interactive Petrophysics program (IP) has been used. Collecting log data of wells was first step to present in IP. K-mean clustering theory which is considered as main concept of cluster analysis contain following steps:

- Cluster means are pre-determined by manual input or automatic seeding.
- K-mean clustering assigns each point to mean by calculation of the sum of the variance ranges for the datasets of the average.
- Allocate the item to the cluster with the least variation.
- The loop continuous until the cluster mean no longer move with each run.

Furthermore, the steps of work include preparing the well data including well log data, drilling mud and cemented as well as selecting the top and bottom of logged formation to the cluster analysis by correction logs as shown in Fig.1. In this step, it will find petrophysical properties of rock formation such as (porosity and permeability) and save it in the database.

Figure 1. correction input logs by SLB correction

While Fig. 2 is showing inputting clustering parameters represented by well logs data as following symbols (GR, DT, SONIC, RHOB, CALL...etc.) and show the plots of any well after input data through choose the interval of logged formation depth. Picking discriminators data of all well logs for 1 and 2 model build as shown in Fig. 3.

	Use	Default	Log	Well	Well
	Curve	Name		1	2
Well Name →	←			(2) well1	(3) well3
Input Curve 1 →	✓	RHOGB		RHOGB	RHOGB
Input Curve 2 →	✓	*gammaray		GrC	GrC
Input Curve 3 →	✓	*neutron		NPHI	NPHI
Input Curve 4 →	✓	*sonic		DT	DT
Input Curve 5 →	✓	*deepres	✓	LLD	LLD
Input Curve 6 →					
Input Curve 7 →					
Input Curve 8 →					
Use Well →		for Model Build		✓	✓
Top Depth		for Model Build		2231.4	2239.5
Bottom Depth		for Model Build		3453.2	4051.4
Use Well →		for Model Run		✓	✓
Top Depth		for Model Run		2231.4	2239.5
Bottom Depth		for Model Run		3453.2	4051.4
Show Plot		for Model Run		Show Plot	Show Plot
Discriminator →	CrV 1				
Discriminator →	CrV 2				

Figure 2. Input data of cluster analysis

DISCRIMINATORS

Use Discriminator 1 for model build       Use Discriminator 1 for model run

Use Data when Disc 1 is > 7.5 and < 9.5

and

Use Discriminator 2 for model build       Use Discriminator 2 for model run

Use Data when Disc 2 is > 40 or < [ ]

Figure 3. Input discriminators data

After inputting well data in IP program (cluster analysis), the next step will be driven randomness consolidation and editing output curves to show the result of editing clusters mean and lithology facies in multi well plot as shown in Fig.4, and Fig. 5.

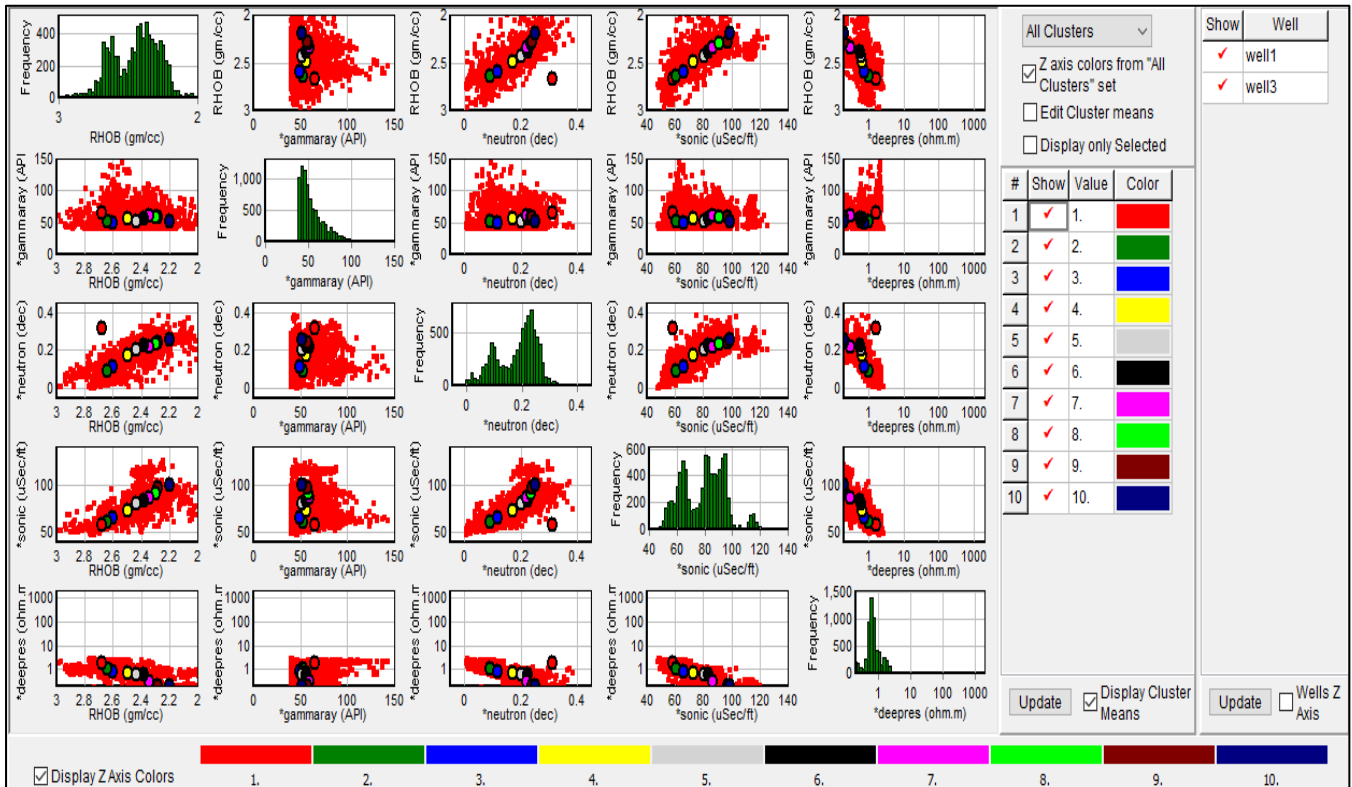
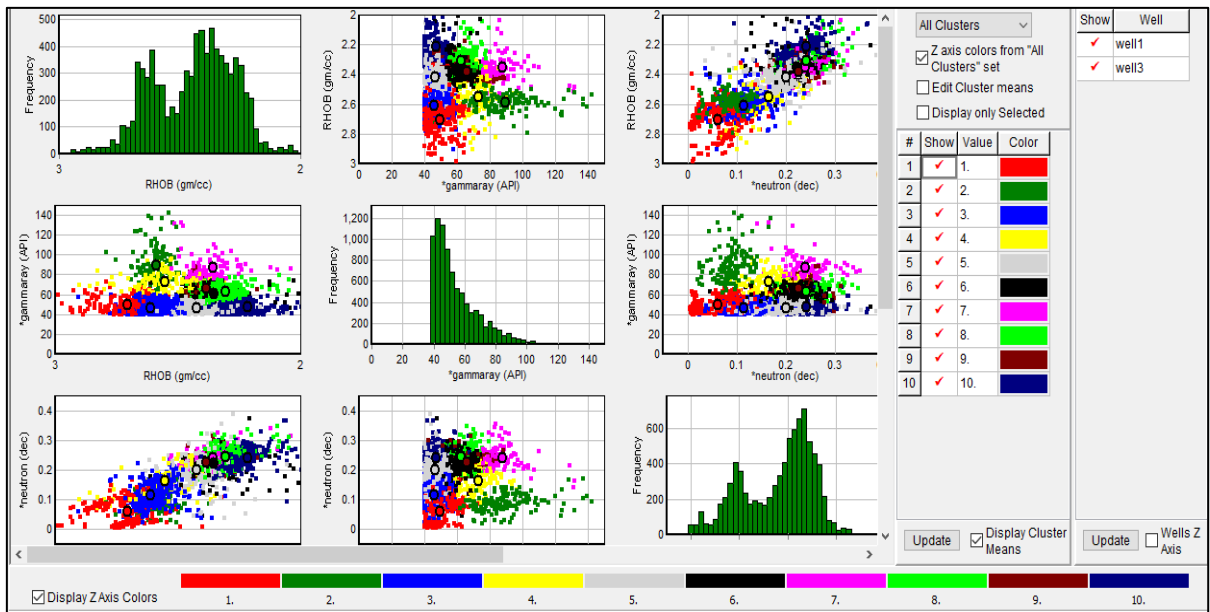


Figure 4. Editing wrong cluster mean data



Figures 5. Plot of cluster editing result

Final step is to increase resolution and improving filtration of output due to select the bulk density cluster as illustrated in Fig. 6. For instance, of previous procedure, Fig. 7 shows a result of one of inputting data of well logs which is neutron log data and the outputting of visualization data.

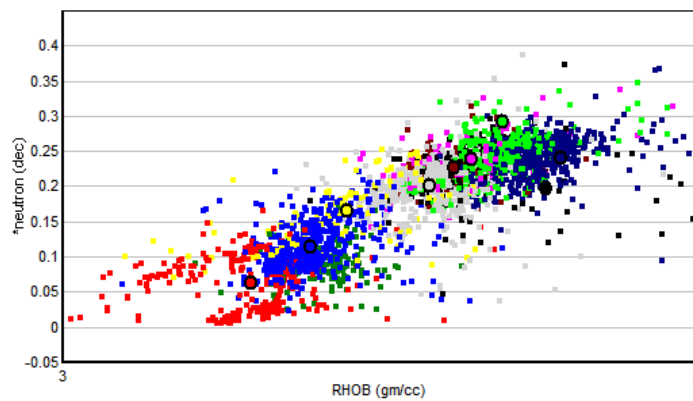


Figure 6. Bulk density clustering

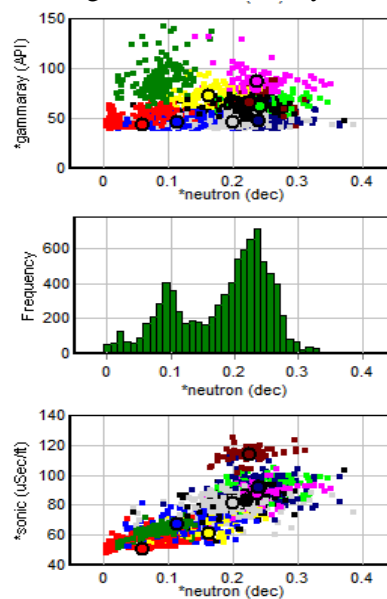


Figure 7. Result of clustering of neutron log

Hint: Editing other logs similar to above logs.

### 3. Result and discussion

It is clear to show that the result of outputting of cluster analysis which presents an obviously simulation among the lithology layers. This could be done through picking correct cluster data by consolidated a group of data for specific well then applied for rest well (see Fig. 8). Whereas Fig. 9 and Fig 10 shows the number of cluster randomness and editing outputting faces for easily recognize of lithology. Where, the comparison among the result of well data present slightly different in interface of lithology which could be neglected.

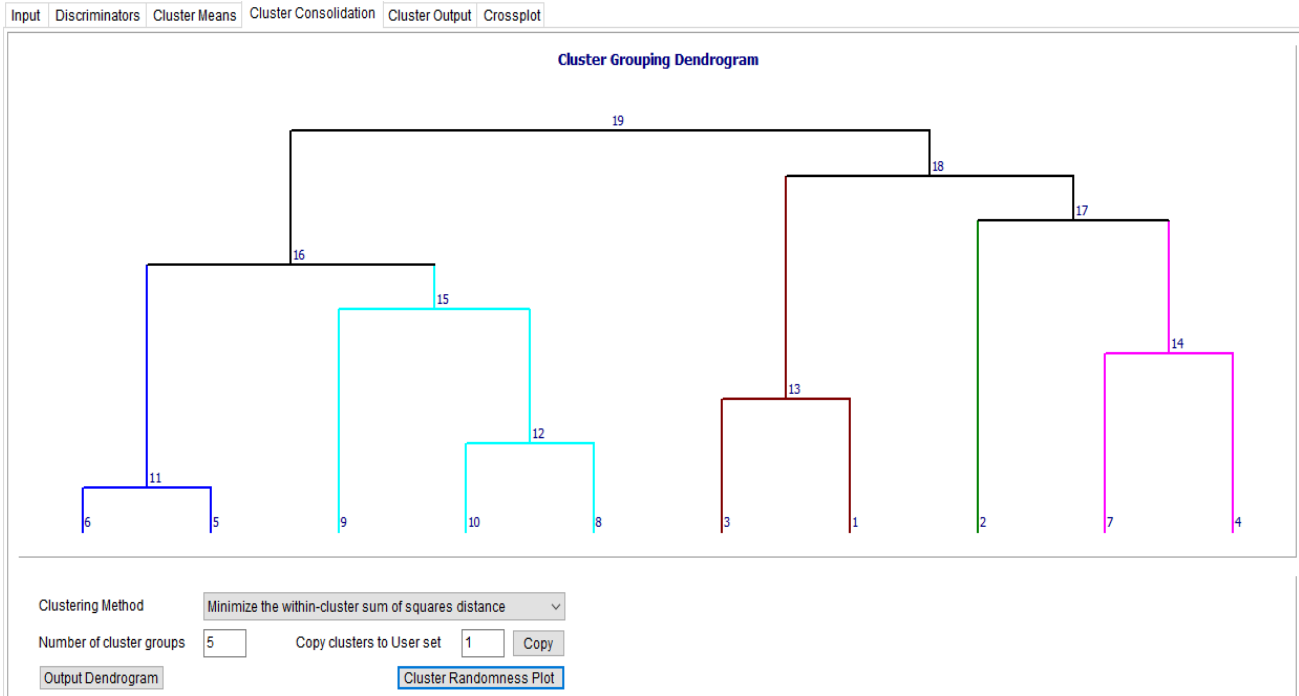


Figure 8. Cluster consolidation of well

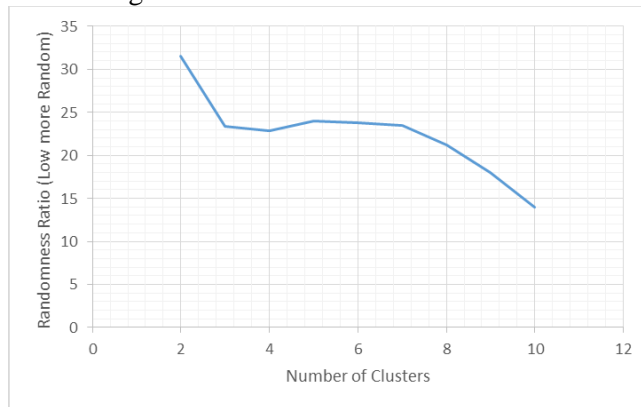
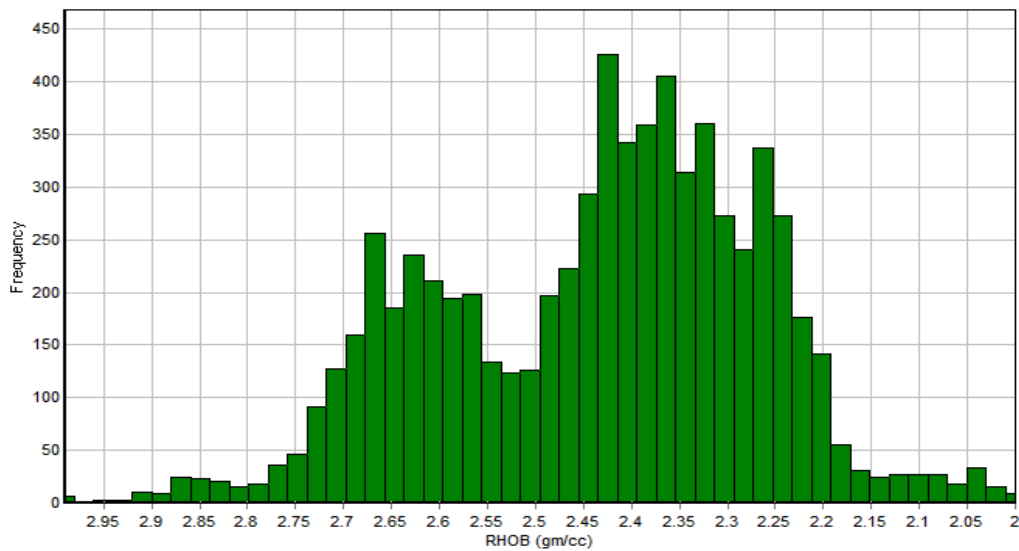
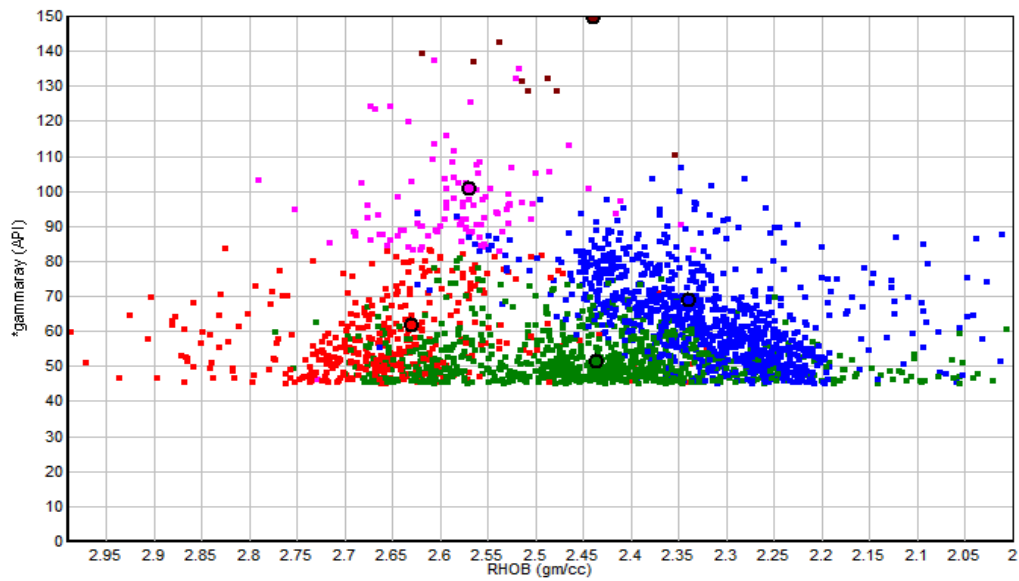


Figure 9. Plot of cluster groups randomness of well

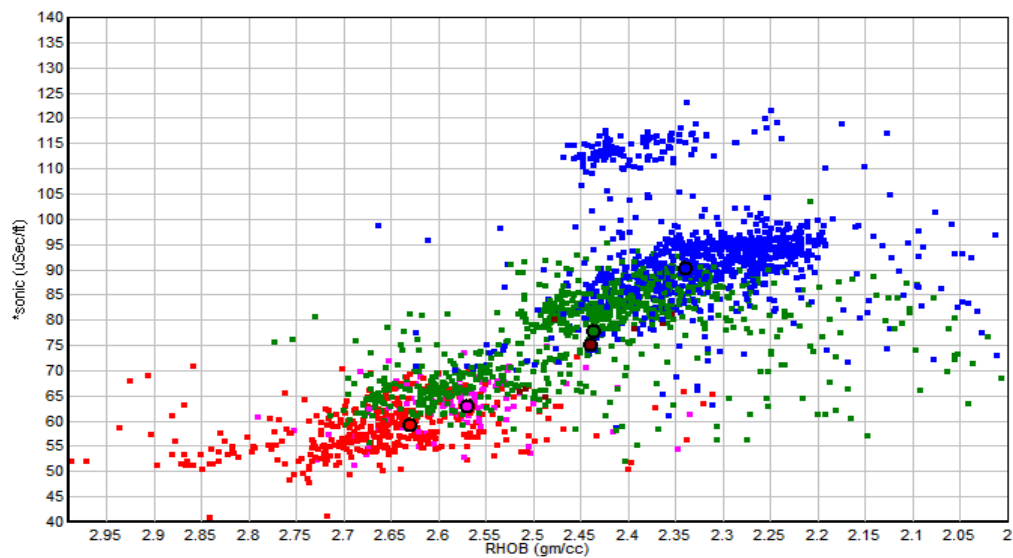
	All Clusters		User set 1		User set 2		User set 3	
Crv Name	L_FaciesAll		L_Facies		L_Facies2		L_Facies3	
Output	✓		✓		✓			
Cluster	Value	Shading	Value	Shading	Value	Shading	Value	Shading
1	1.	Red	1.	Blue	1.	Blue	1.	Blue
2	2.	Green	2.	Green	2.	Green	2.	Green
3	3.	Blue	1.	Blue	1.	Blue	1.	Blue
4	4.	Yellow	3.	Magenta	3.	Yellow	3.	Yellow
5	5.	Grey	4.	Grey	4.	Grey	4.	Grey
6	6.	Black	4.	Grey	4.	Grey	4.	Grey
7	7.	Magenta	3.	Magenta	5.	Magenta	5.	Magenta
8	8.	Cyan	5.	Blue	4.	Grey	4.	Grey
9	9.	Brown	5.	Blue	4.	Grey	4.	Grey
10	10.	Blue	5.	Blue	4.	Grey	4.	Grey

Figure 10. Cluster output facies editing

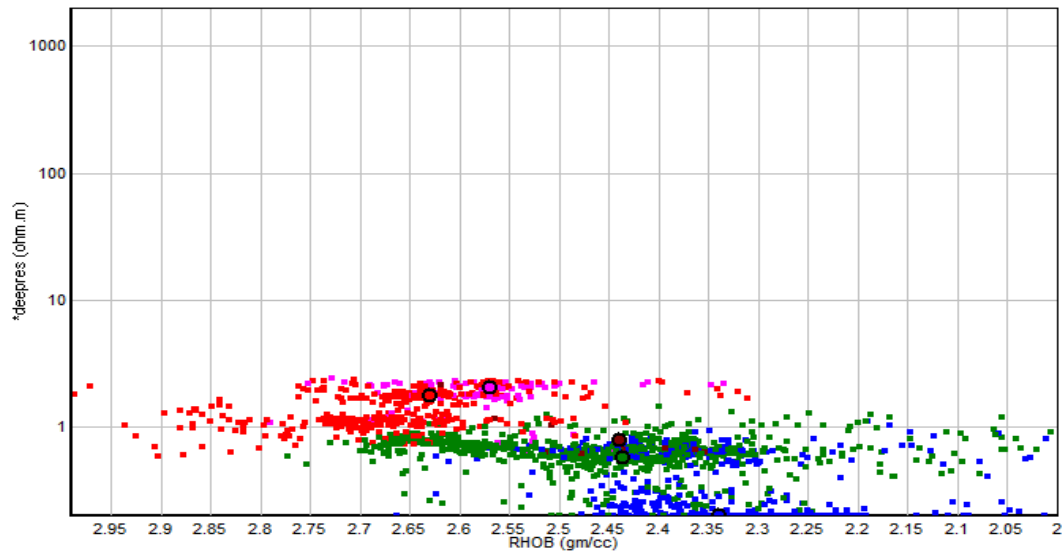
To improve visualization data, we will remove wrong or annoying data to concentration, correlate and mitigate the overlap among of lithology layers (show Fig. 11 - 12).



Figures 11. The distribution of bulk density log data and removal the wrong data



Figures 12a. The distribution of bulk density log data and removal the wrong data



Figures 12b. The distribution of bulk density log data and removal the wrong data

As a result, Fig. 13 illustrated final lithology facies of one well as an example, and it is clear to distinguish among the lithology layers and the exact area of hydrocarbon by choosing the correct depth. Thus, building a structure of reservoir by stratigraphic correlation can be done through this method.

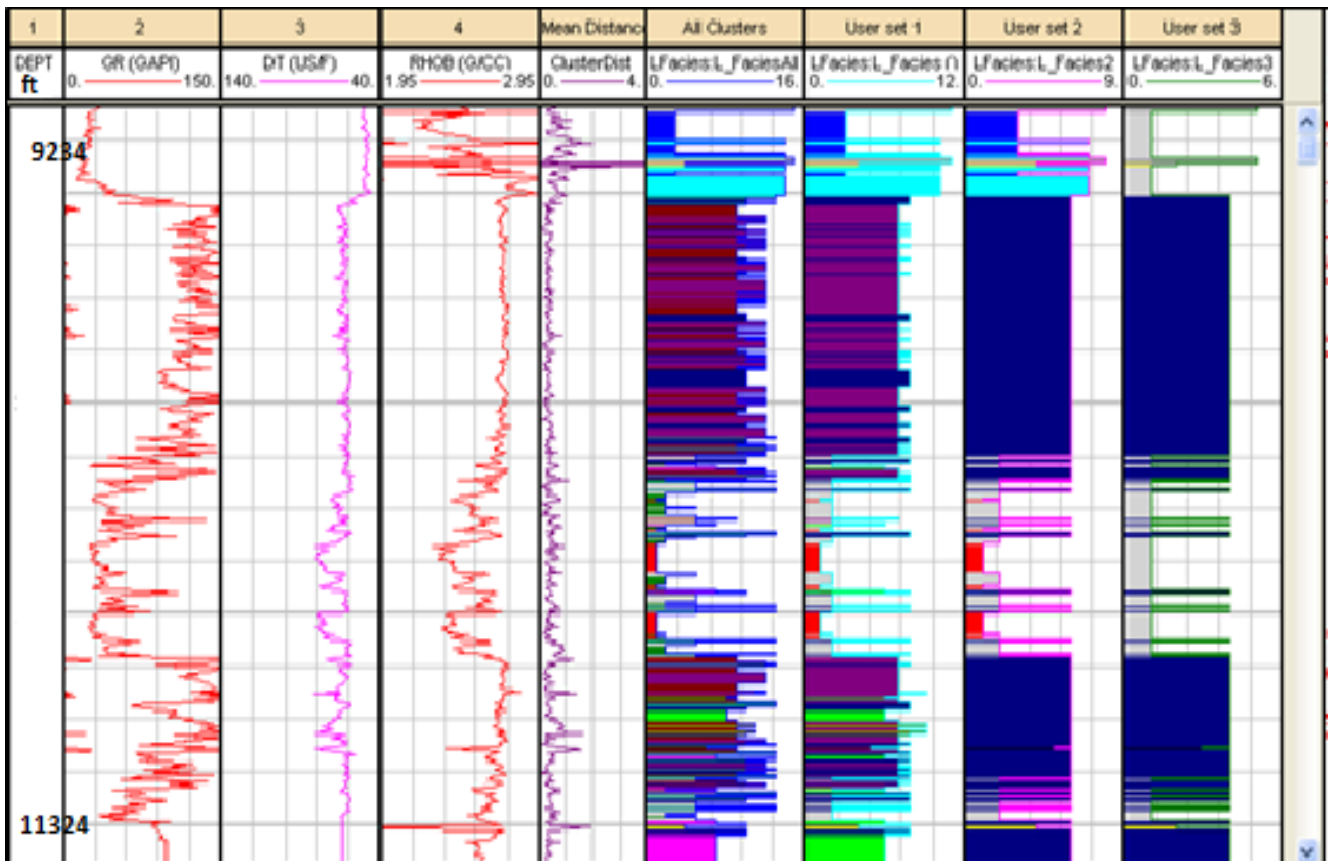


Figure 13. The result of clustering of well logs and show lithology facies

#### 4. Conclusions

A technique used to evaluate well log clusters and identify electro facies for exploration purposes has been introduced. This approach is intended to classify areas with increased effectiveness for deposition of petroleum products. A random test conducted on 41 wells has shown that the designation of electro facies reliably forecasts interval output of methane.

Categorization was carried out on 395 wells in a 28-town district and the findings were used to classify natural gas aggregation and potential areas. This method is especially well suited for exploration in developed reservoirs where high quantities of well data are available and where reservoir structure makes traditional log evaluation less efficient. Various domain concepts that can be introduced into the logistic system are rapidly examined. Heterogeneous data is expressed in several aspects. Hydrocarbon ontology is properly defined for the representation of spatial and temporal operating in various aspects. This methodology enables the recognition and analysis of a variety of data categories in the data mining phase between oil discovery and development data aspects, either geographically based servers or on a local computer. Dimensions assigned graphically (data visualization) to the structure, reservoir and other reservoir resources depicted in various clusters indicate unique data alignments or trends.

## References

- [1] S. Omid, T. Seyed, and A. Mohammad, "Prediction of the Rock Mass Diggability Index by Using Fuzzy Clustering-Based, ANN and Multiple Regression Methods," *Rock Mechanics and Rock Engineering* 47, Vol. 2, 717–32. <https://doi.org/10.1007/s00603-013-0426-3>, March 2014.
- [2] K. Toropetsky, V. Ulyanov, G. Borisov, R. Kurmangaliev, N. Kayurov, and V. Arzhantsev, "A review of petrophysical correlations for building one-dimensional geological and geomechanical models in problems of wellbore stability while drilling," *Automation, Telemetry, and Communication in Oil Industry*, No.12, pp. 31-44, 2016.
- [3] K. Vorontsov, "Lectures on clustering and multidimensional scaling algorithms," *Moscow: Moscow State University* 18, 2007.
- [4] M. Abbas, and B. Morteza, "Applying Evolutionary Optimization Algorithms for Improving Fuzzy C-Mean Clustering Performance to Predict the Deformation Modulus of Rock Mass," *International Journal of Rock Mechanics and Mining Sciences* No. 113, Jun 2019.
- [5] A. Al-bayati, M. Ismail, and F. Terzi, "Optimum horizontal well length considering reservoir properties and drainage area," *Period. Eng. Nat. Sci.*, vol. 8, no. 3, pp. 1836–1846, 2020, doi: 10.21533/pen.v8i3.1621.
- [6] C. Aggarwal, and C. Reddy, "Data Clustering: Algorithms and Applications," *N.-Y.: Chapman and Hall/CRC*, 652 p, 2014.
- [7] Y. Kashnikov, and S. Ashikhmin, "Rock mechanics in the development of hydrocarbon fields," *Monograph. –Moscow, Mining Book PH*, 496 p, 2019.
- [8] K. Toropetsky, N. Kayurov, A. Cheremisin, M. Lushev, M. Samoilov, V. Ulyanov, and G. Borisov, "Building 1D physical-mechanical models and solving the problems of wellbore and near-wellbore stability," *Automation, Telemetry, and Communication in Oil Industry*, No. 11, pp. 29 – 4, 2016.
- [9] K. Toropetsky, N. Kayurov, V. Eremin, E. Lukyanov, V. Ulyanov, and G. Shevtsov, "Well construction support technology," *Automation, Telemetry, and Communication in Oil Industry*, No.12, pp. 4-15, 2017.