No.12 Journal of Petroleum Research & Studies

Using Well Logs Data in Logfacies Determination By Applying the Cluster Analysis technique for Khasib Formation, Amara Oil Field, South Eastern Iraq

Buraq Adnan Hussein Al-Baldawi
Department of Geology, University of Baghdad.

Abstract:

Reservoir facies determination is the most important job in oil industry which dominantly relies on the major properties of rocks. Fundamental properties of rocks are usually understood by their detailed description in the field (lithofacies analysis) and laboratory (petrofacies analysis). The facies (lithofacies and petrofacies) determination in most subsurface studies is impractical, due to lack of cores and cuttings. In such situations, where the wire line logs are the only data available, the logfacies or electrofacies are determined instead. In this study, the available logs (gamma ray, density,Neutron, and sonic logs) for four wells (Am-1, Am-4, Am-5, and Am-6) were used to predicate logfacies of Khasib Formation in Amara oil field. Using Cluster Analysis in this study practiced the logfacies determination in each unit of khasib Formation in Amara oil Field using Interactive Petrophysics software. The types of input data into Interactive Petrophysics software to determine logfacies are available logs which are used to create 15 clusters and determine five groups of logfacies. In this paper, the vertical variations of logfacies for Khasib Formation are carried out based on five groups of logfacies. These groups of logfacies are classified based on responses of well logs after divided Khasib Formation into three units :(upper, middle and lower Khasib). In order to estimate the ratio of existence of logfacies, the histogram of logfacies was built for each unit of Khasib Formation in studied wells.

Keywords: Logfacies, Cluster analysis, Mishrif Formation, Amara Oil Field.

Introduction:

Well logs are principal sources of subsurface geological information. They provide significant information on mineralogical composition, texture, sedimentary structures and petrophysical properties such as porosity and permeability. By compiling data from various well logs, one can discriminate sedimentary units with comparable log characteristics. The sedimentary units which defined on this basis and characterized from wire line logs are known as electrofacies or logfacies in the literatures (Serra 1986). Logfacies analysis is the most important tool in petroleum industry, sedimentological and depositional environment study of the bearing rocks, especially where wire line logs are only reliable data available (Serra 1986). Logfacies analysis can be carried out manually or automatically using mathematical techniques. Multivariate cluster analysis (as the best method of data grouping) is one of the most accurate and affective methods in oil bearing reservoir. The method is applied on both detrital and carbonate rocks (Gill *et al*. 1993). The purpose of present study is classify the gamma ray, neutron log,density, and sonic logs into logfacies types. The cluster analysis method is used to perform the logfacies classification based on attempts to identify clusters of well logs responses with similar characteristics. This classification does not require any artificial subdivision of the data population but follows naturally based on the unique characteristics of well-log measurements, reflecting minerals and lithofacies within the logged interval. This classification of logfacies in my study is done by Interactive Petrophysics software. As well as, the vertical distribution of logfacies for Khasib Formation in Amara oil Field is carried out based on classification of these clusters into five groups of logfacies. In order to obtain presence of logfacies that predicated from cluster analysis ,the histogram of logfacies types of each zone of khasib Formation are carried out using . Interactive Petrophysics software.

Geology of studied Area:

Amara oil field is located at south eastern Iraq in Missan governorate, about 10 Km south western Amara city and about 25 Km east of Al-Rafedain structure (Abu-Amoud structure), and 30 Km south eastern Al-Kumait structure figure (1). Amara structure is assumed to be a low-relief dome though slightly W-E elongated, having dimensions of approximately 16 Kms width (from west to east side) by 5 Kms length (from south to north) as defined from Amara area 2D seismic lines (Missan Oil Company). The Khasib Formation is one of the important reservoirs in the Amara oil field, and it is deposited at the base of megasequence AP9 in the late Touronian- Early Campanian sequence (Jassim and Goff., 2006).The late Touronian- Early Campanian sequence represented by Khasib, Tanuma, Sa`ady, and Komitan Formations, There are two main facies in this sequence: middle shelf represented by Khasib, Tanuma, and Sa`ady formations, However influence of the outershelf environment is obvious with the planktonic abundance (Aqrawi., 1996) and basinal depositional environment of the equivailant Komitan Formation (Jassim and Goff., 2006). Khasib reservoir is mainly composed of Limestone rocks from the late Cretaceous Period. According to Bellen et al,. (1959) the lower boundary of the formation is disconformable with Mishrif Formation. The upper boundary with theTanuma Formation is gradational.

Methodology and Clustering Procedure:

Similar facies may have different log responses due to diverse factors that affect the logs. Since using statistical methods and procedures are mandatory, in clustering procedure data are grouped with minimum distance and maximum homogeneity. It is obvious that distinct geological parameters can be related to a group of data, as logfacies, which be used by geologists for further interpretation. For this calculation, all log readings are considered as "observations" and the used logs as the "values of the observations".In cluster analysis smallest distances are connected together to build a pair. Usually the number of logfacies are less than the number of readings, since then, pairs of vectors (pairs of log readings) are linked together to build a cluster (logfacies). The lower-rank clusters are linked together to build higher-rank types. This procedure continues until a single cluster (representing the whole data) is built. There are various methods to link two clusters. In some of which the minimum distance of the cluster components are utilized to link them.The clustering module carried out in two stages using Interactive petrophysics software (IP): Firstly, the data (gamma ray log, porosity, water saturation) is divided up into manageable data clusters. The number of clusters should be enough to cover all the different data ranges seen on the logs. 15 to 20 clusters would appear to be a reasonable number for most data sets. The second step, which is more manual, is to take these 15 to 20 clusters and group them into a manageable number of geological facies. This may involve reducing the data to 5 to 6 clusters. The first stage of 'Facies Clustering' uses the K-mean statistical technique to cluster the data into a known entered number of clusters. For this to work an initial guess has to be made of the mean value of each cluster for each input log. The initial guess can affect the results and in order to get good results the initial values should cover the total range of the logs. K-mean clustering works by assigning each input data point to a cluster. The routine tries to minimize the within-cluster sums of squares of the difference

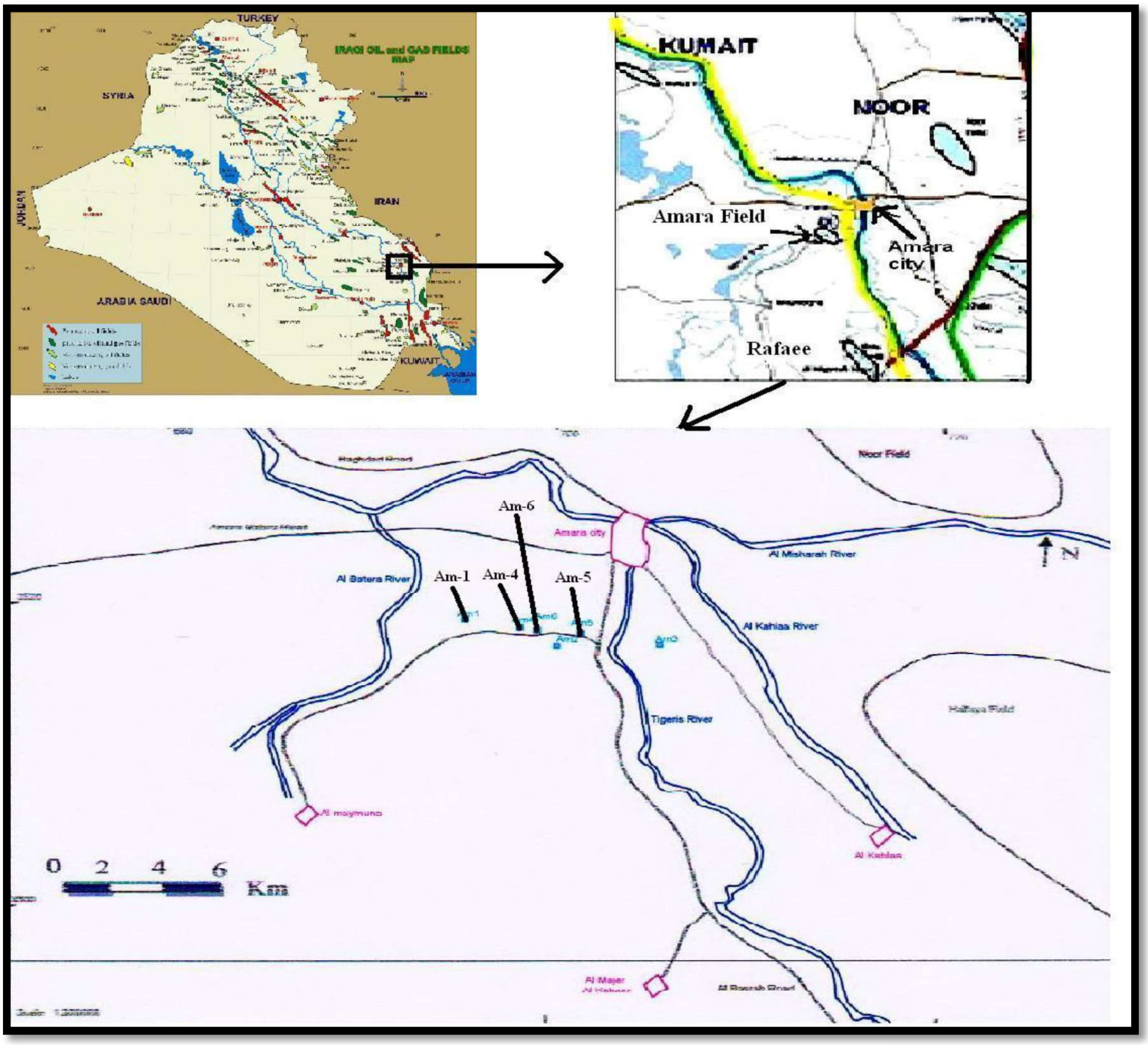


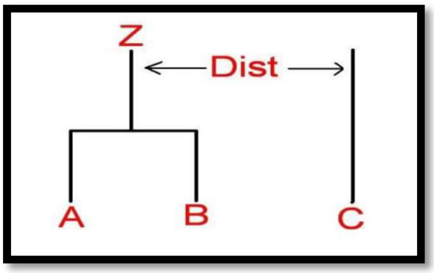
Fig.(1) Location map of studied Area (modified from Missan oil company (2002)).

Between the data point and the cluster mean value. The routine works by calculating the sum of the squares difference for a data point and each cluster mean and assigning the point to the cluster with the minimum difference. Once all the data points have been assigned to the clusters the new mean values in each cluster are calculated. Using the new mean values the routines starts again re-assigning the data to the clusters. This loop continues until the mean values do not change between loops. These then become the results. All input log data is normalized (standardized) before starting so that each input log has the same dynamic range. The normalization is done by calculating the mean and standard deviation of the log and then normalizing the data by subtracting the mean and dividing by the standard deviation.

Stage-2 Cluster Consolidation

Cluster consolidation can be done completely manually by using the crossplot and log plot output to group data, or a hierarchical cluster technique can be used to group the data. Hierarchical clustering works by computing the distances between all clusters and then merging the two clusters closest together. The new cluster distance to all other clusters is then recomputed and the two closest clusters merged again. This process continues until you have only one cluster. The results can be plotted as a dendrogram. The dendrogram shows how the clusters were merged and the order they were merged. The numbers at the top of each branch give the merging order. The original results from the K-mean clustering are shown at the base of the plot.

IP software has five different clustering methods which decide how the clusters are merged. The different methods will show considerably different results. The five methods differ in how the distance calculation is updated after two clusters have been joined. If we assume, in the diagram below, cluster "A" and "B" have just been joined to form cluster "Z", and we need to calculate the distance of "Z" with another cluster called "C".



For the different techniques the calculations are made as follows:

1- Minimum distance between all objects in clusters - the distance from Z to C is the minimum of the distances (A to C, B to C).

2- Maximum distance between all objects in clusters - the distance from Z to C is the maximum of the distances (A to C, B to C).

3- Average distance between merged clusters - the distance from Z to C is the average distance of all objects that would be within the cluster formed by merging clusters and C.

4- Average distance between all objects in clusters - the distance from Z to C is the average distance of objects within cluster Z to objects within cluster C.

5- Minimize the within-cluster sum of squares distance - clusters are formed so as to minimize the increase in the within-cluster sums of squares. The distance between two clusters is the increase in these sums of squares if the two clusters were merged.

In general, 'Minimum distance between all objects in clusters' will yield long thin clusters while 'Maximum distance between all objects in clusters' will yield clusters that are more spherical. 'Average distance between merged clusters' and 'Minimize the within-cluster sum of squares distance' tend to yield clusters that are similar to those obtained with 'Average distance between all objects in clusters'. The default method 'Minimize the within-cluster sum of squares distance' gives good results for separating out the different log lithologies into different clusters as shown in figure (2) of dendrogram**.**

The grouping of the clusters into a known number of groups is easily done by stopping the grouping at a certain cutoff level.

It is possible to analyze the groupings to decide at which level adding another cluster gives more information or is just adding noise. The 'Cluster Randomness Plot' provides this information.

The figure (3) of Cluster Randomness Plot' calculates for each number of cluster groups the perceived randomness of the data. The higher the value the less random the clusters are - i.e. more structured data.

The randomness is calculated by first calculating the average number of depth levels per cluster - i.e. the average thickness of a cluster layer. This is performed on the original log data. Then the theoretical average thickness is calculated assuming the clusters to be assigned randomly at each depth level. The randomness is the ratio of the two. A value of 1 would be totally random, higher values less random.

Av. Thickness = Number of depth levels / Number of cluster layers

Random Thickness = $\Sigma\ p_i / (1 - p_i)$

Where $\mathbf{p}_i$ is the proportion of depth levels assigned to the ith cluster.

Randomness index = Av. Thickness / Random Thickness

The plot is interpreted by picking the number of clusters that are least random (highest peaks).

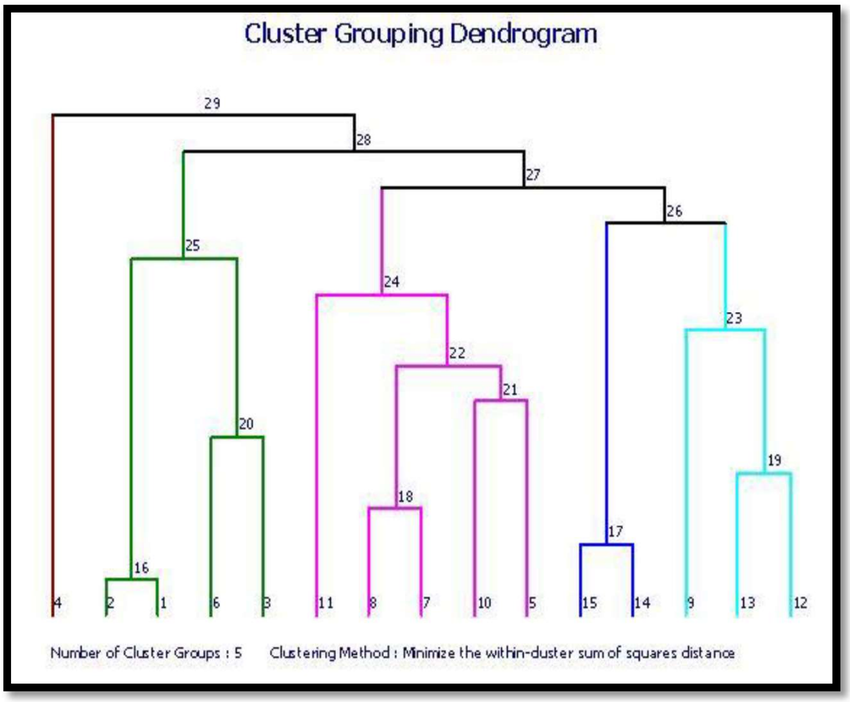


Fig. (2) Cluster Grouping Dendrogram of Khasib Formation.

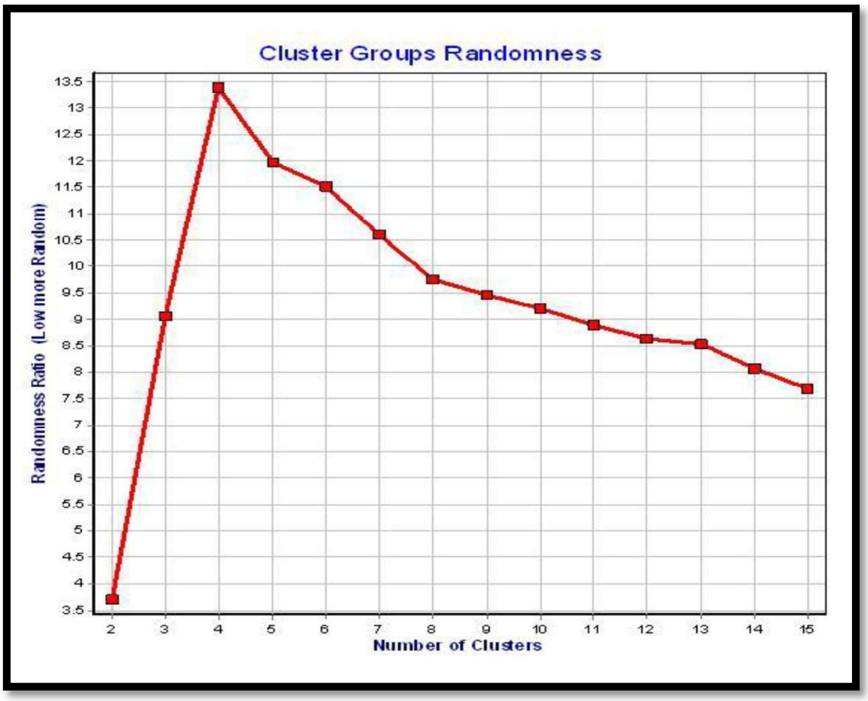


Fig. (3) Cluster Groups Randomness of Khasib Formation.

Results and conclusions:

The figure of dendrogram shows the Khasib Formation is divided into 15 clusters and then created five groups of logfacies based on input data (gamma ray log, density log, sonic log, and neutron log) that used to build clusters. Each cluster is characterized by mean values of gamma ray reading, density reading, sonic reading, and neutron reading. The results of 'Cluster Means' are presented in table (1) that shows the mean values plus other statistics of used data for each cluster (15 cluster). After classification of used data into 15 clusters, the groups of logfacies were created from these clusters. The Khasib Formation was divided into five groups. Each group of logfacies is characterized by well logs response and each group may be containing one cluster or more than one cluster. Figure (4) shows histograms and cross-plots between used well logs as generated by k-means cluster analysis for groups of Khasib Formation. The properties of these groups of logfacies are presented in table (2). This table shows that logfacies-1 is perfect and favorite reservoir facies in Khasib Formation. The vertical distribution of logfacies for Khasib Formation in each well under study is illustrated in figures (5-8). Al-Baldawi (2014) divided Khasib Formation into three units based on interpretation of well logs and petrophysical properties. These units are: upper khasib, middle khasib, and lower khasib. Correct correlation of stratigraphic units using well logs is necessary to make reliable cross sections, and to conduct regional facies analysis Figure (9) shows correlation section of khasib units and distribution of logfacies for each unit.

Table (1):- The results of 'Cluster Means' plus other statistics of used data for each cluster.

Cluster No.	Points	Cluster Spread	GR		RHOB		DT		NPHI	
			Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
1	182	0.6539	8.011	2.9	2.24	0.037	86.00	4.282	0.213	0.024
2	160	0.5396	9.559	2.89	2.299	0.033	80.83	2.866	0.191	0.018
3	177	0.7808	16.48	4.68	2.327	0.055	69.55	3.171	0.186	0.025
4	33	2.556	60.66	16.2	2.282	0.074	103.2	14.18	0.269	0.073
5	189	2.556	31.37	4.22	2.377	0.036	76.97	3.006	0.158	0.027
6	253	0.6443	17.32	4.21	2.385	0.036	75.81	2.716	0.148	0.019
7	290	0.6247	26.91	3.6	2.406	0.039	68.59	2.207	0.156	0.015
8	309	0.5518	27.76	4.36	2.450	0.035	73.15	3.05	0.11	0.013
9	427	0.6602	13.28	3.01	2.455	0.038	67.98	3.568	0.119	0.019
10	245	0.6829	39.28	4.61	2.409	0.045	70.00	3.213	0.156	0.023
11	233	0.7025	40.78	5.82	2.514	0.042	69.01	2.541	0.105	0.018
12	226	0.706	25.4	3.70	2.499	0.047	64.77	3.78	0.103	0.018
13	272	0.6441	18.57	4.14	2.575	0.041	65.20	3.046	0.068	0.016
14	250	0.6575	27.37	3.78	2.646	0.037	60.28	3.364	0.05	0.018
15	290	0.5696	17.91	3.88	2.662	0.023	59.69	3.131	0.034	0.014

Table (2):- The properties for groups of clusters.

Group Name	Group sample	Clusters of group	Logfacies properties
Logfacies-1		1,2,3,6	Low Gamma ray that ranges between (8-17), Density reading ranges between (2.2-2.3), high sonic log, neutron log reads proximately the highest reading in the formation.
Logfacies-2		4	The highest reading of gamma ray, neutron and sonic log.
Logfacies-3		5,7,8,10,11	High reading of gamma ray and sonic log, fair to good reading of neutron porosity, density log ranges between (2.3-2.5).
Logfacies-4		9,12,13	Medium to high reading of gamma ray, low to medium sonic and neutron reading.
Logfacies-5		14,15	Very low reading of neutron and sonic logs, gamma ray reading ranges between 17 to 27 API.

In order to estimation the ratio of existence of each logfacies in units of Khasib Formation, the histogram of logfacies was built for each well under study. Figure (10) of histogram for well Am-1 shows logfacies-3 is predominant in middle Khasib unit in well Am-1 but logfacies -5 is predominant in upper khasib unit where as logfacies -1and 2 is highest ratio in lower Khasib. Figure (11) of histogram for well Am-4 shows logfacies-1 and 4 are predominant in all units of Khasib Formation in well Am-4 with the least ratio of logfacies -3 and 5 where as logfacies -2 is not present in Khasib Formation. Figures (12) and (13) of histogram for well Am-5 amd Am-6 show logfacies -3 is dominant in middle Khasib unit with the different ratios of logfacies -1 and 4 but upper Khasib unit contain logfacies-5 where as logfacies -1 is dominant in lower Khasib unit. In general, the logfacies -3 is the most spread in wells Am-1,Am-5, and Am-6 but logfacies -1, that represent perfect facies in Khasib Formation ,is dominant in well Am-4 because location of well Am-4 is in the center of Amara anticline figure (14) that represent perfect location of Amara wells.

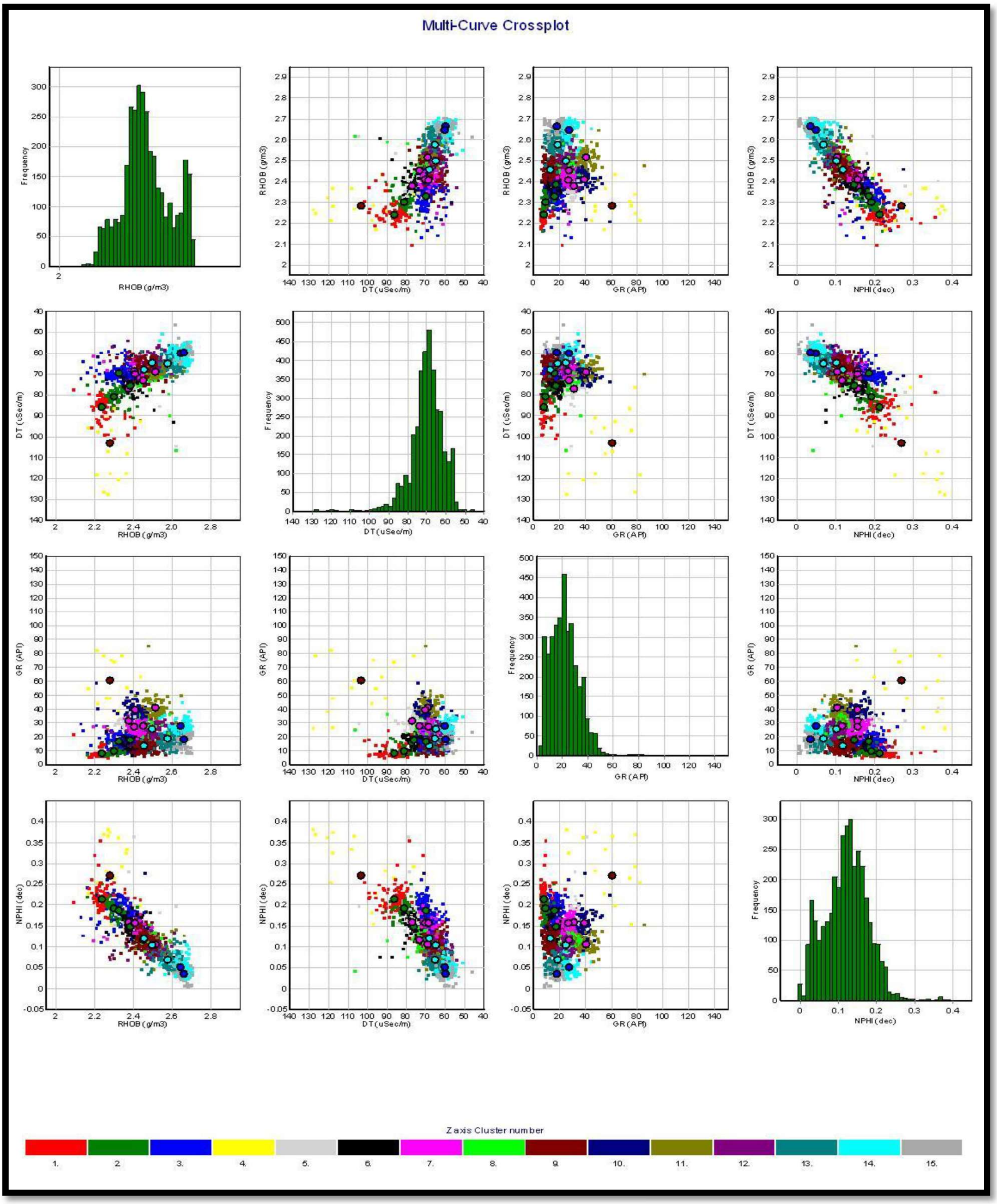


Fig. (4) Cross-plots and histograms between gamma ray, Density, sonic, and Neutron logs as generated by k-means cluster analysis for groups of Khasib Formation.

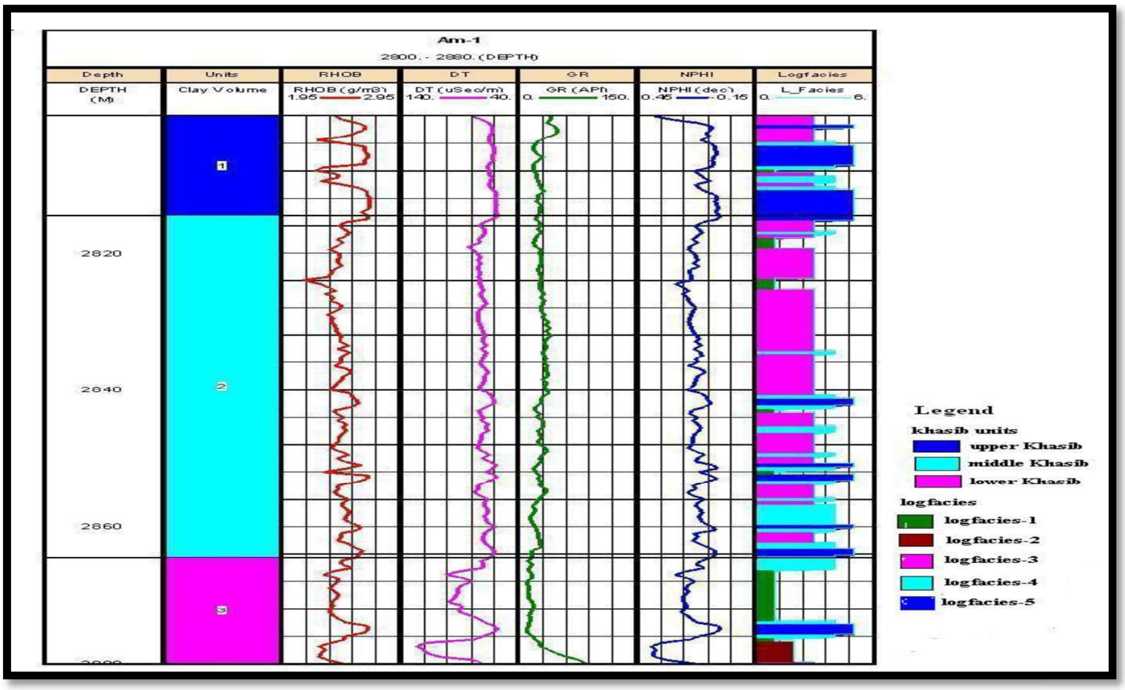


Fig. (5) Stratagraphic column of Khasib Formation in well Am-1.

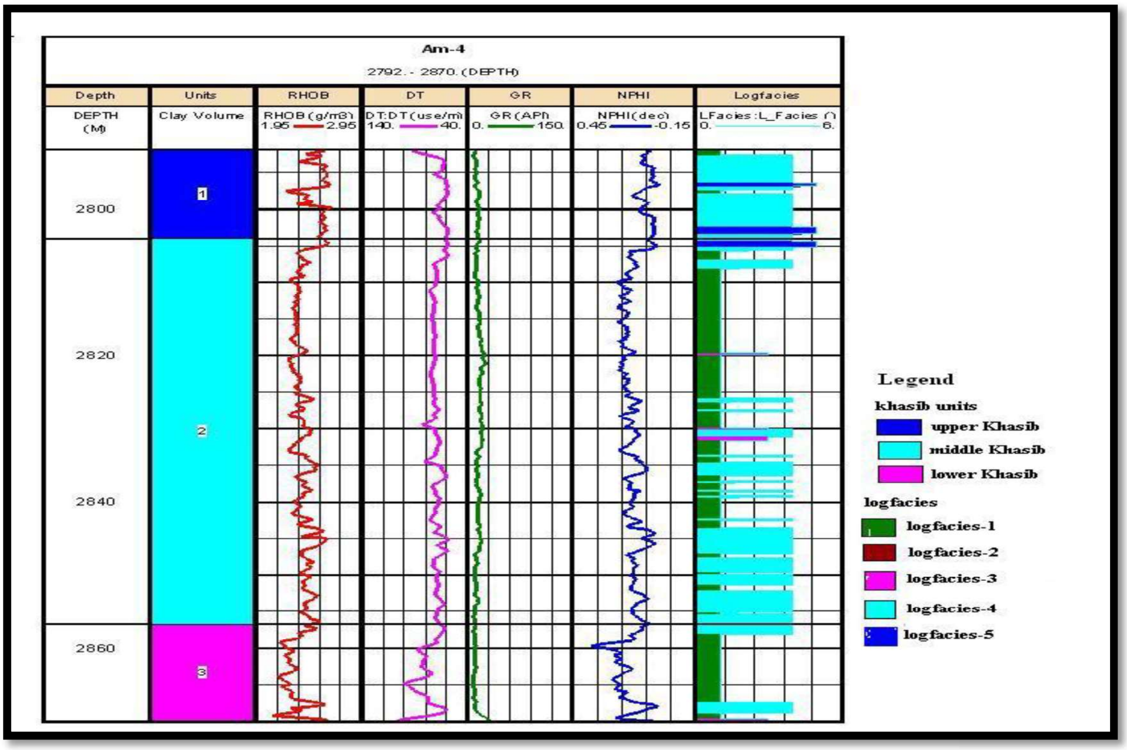


Fig. (6) Stratagraphic column of Khasib Formation in well Am-4.

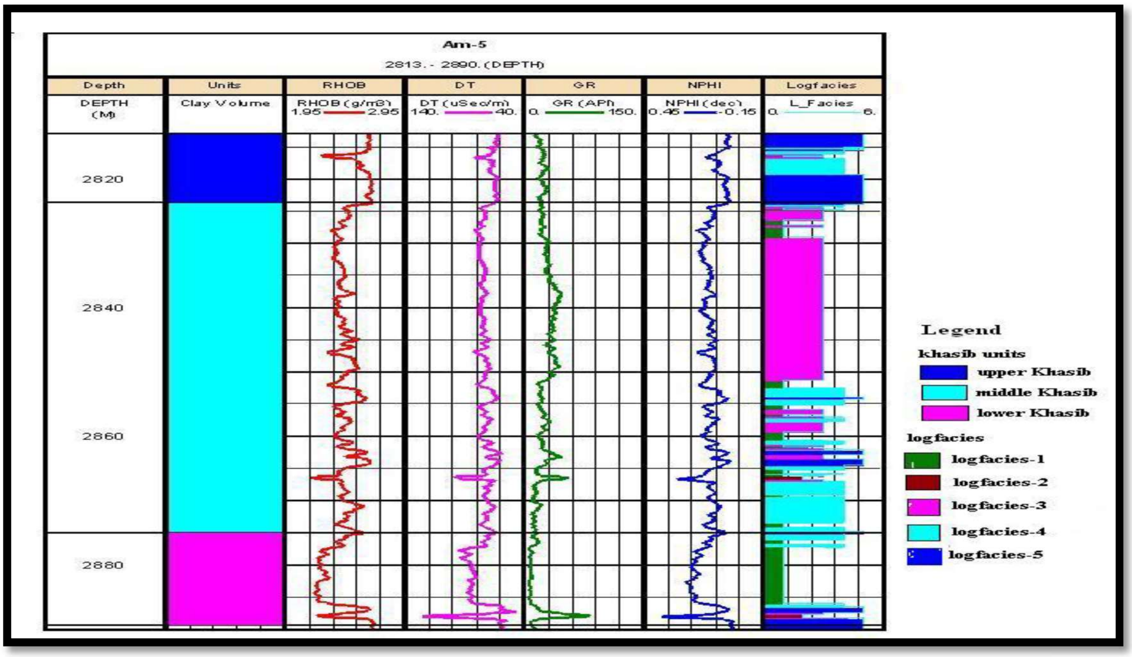


Fig. (7) Stratagraphic column of Khasib Formation in well Am-5.

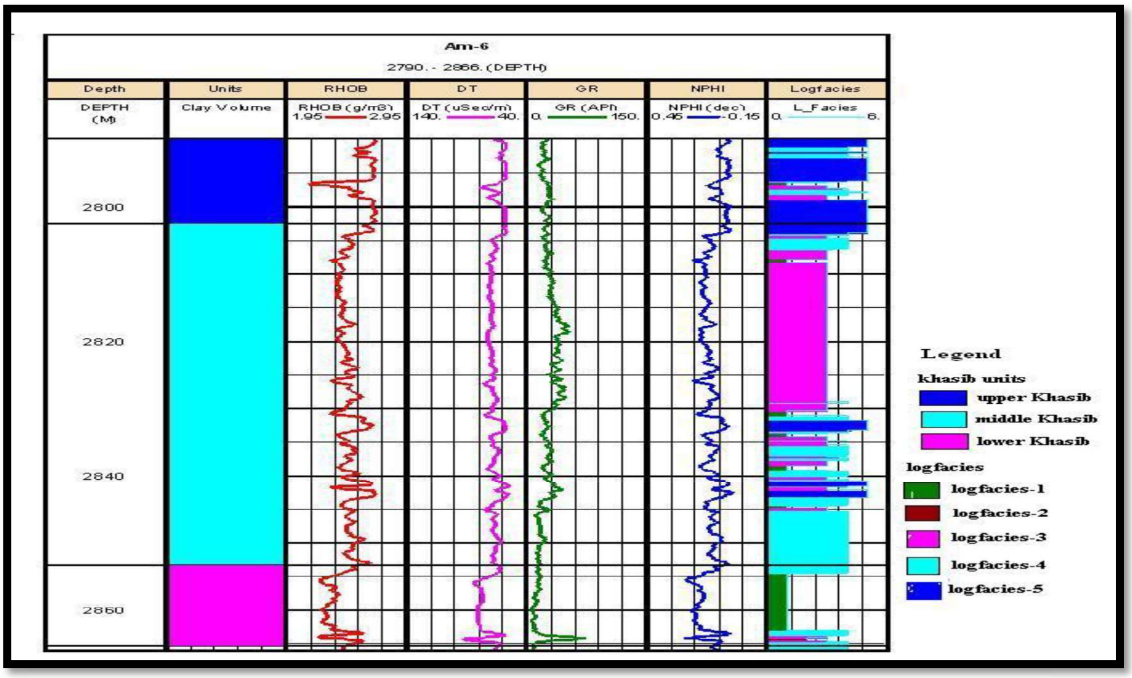


Fig. (8) Stratagraphic column of Khasib Formation in well Am-6.

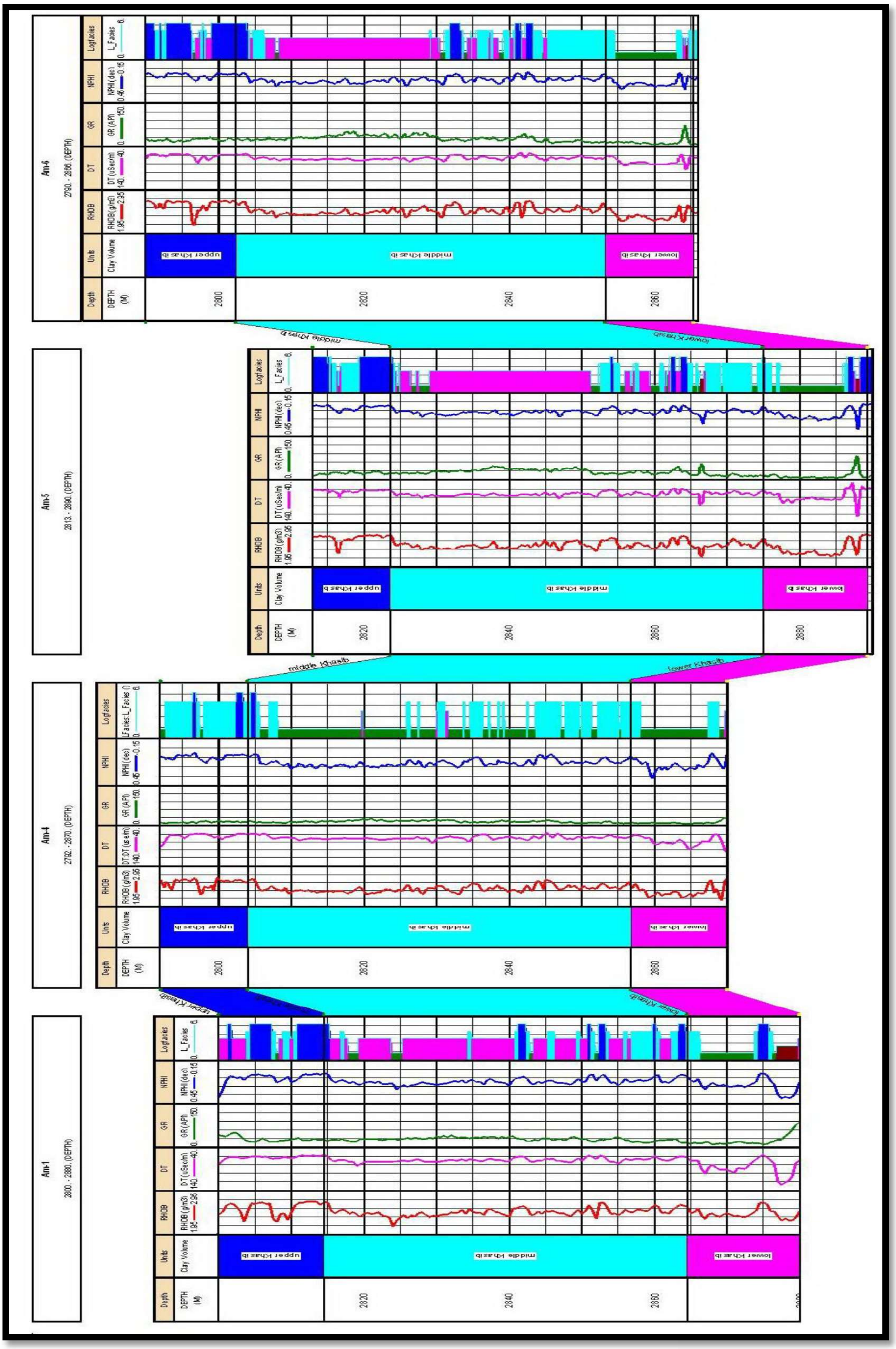
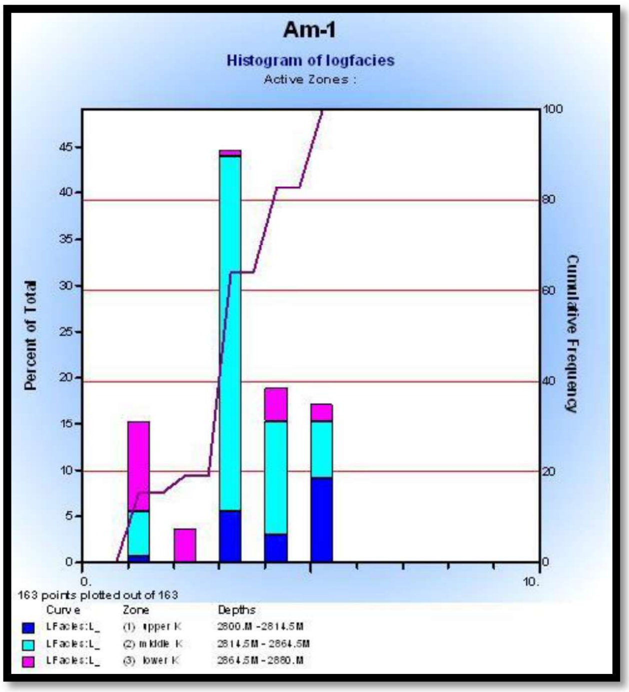


Fig. (9) Correlation section of khasib units and distribution of logfacies for each unit.



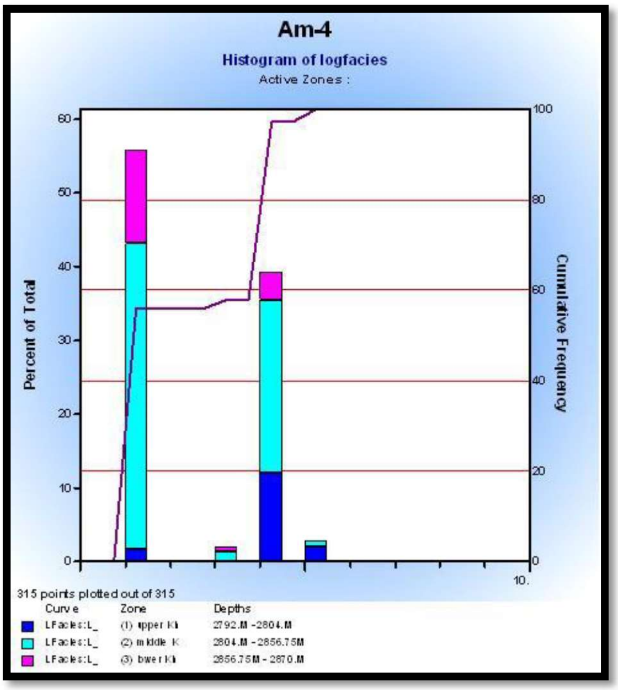
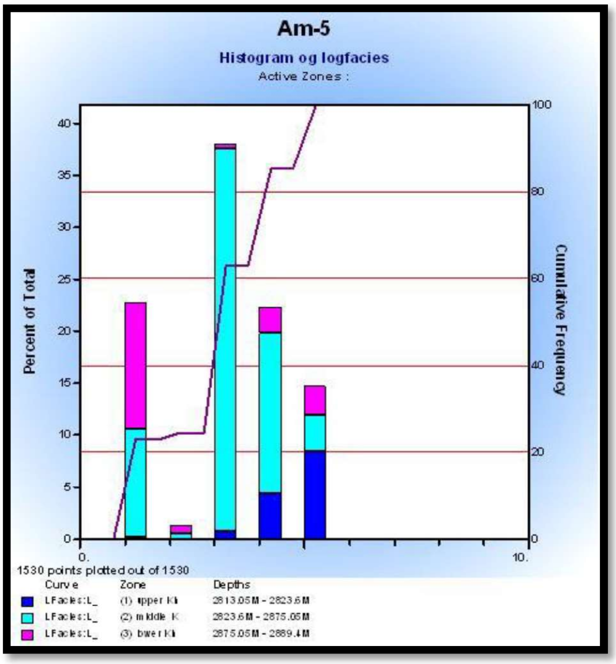


Fig. (10) histohram of logfacies in well Am-1. Fig. (11) histohram of logfacies in well Am-4.



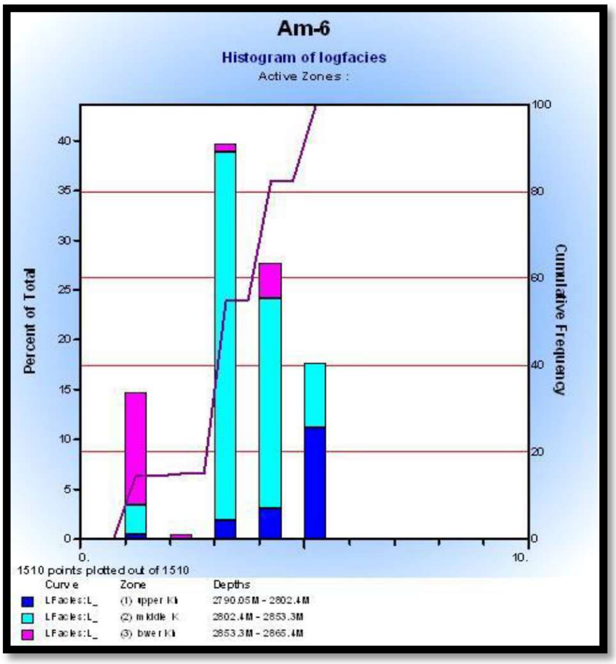


Fig.(12) histohram of logfacies in well Am-5. Fig. (13) histohram of logfacies in well Am-6.

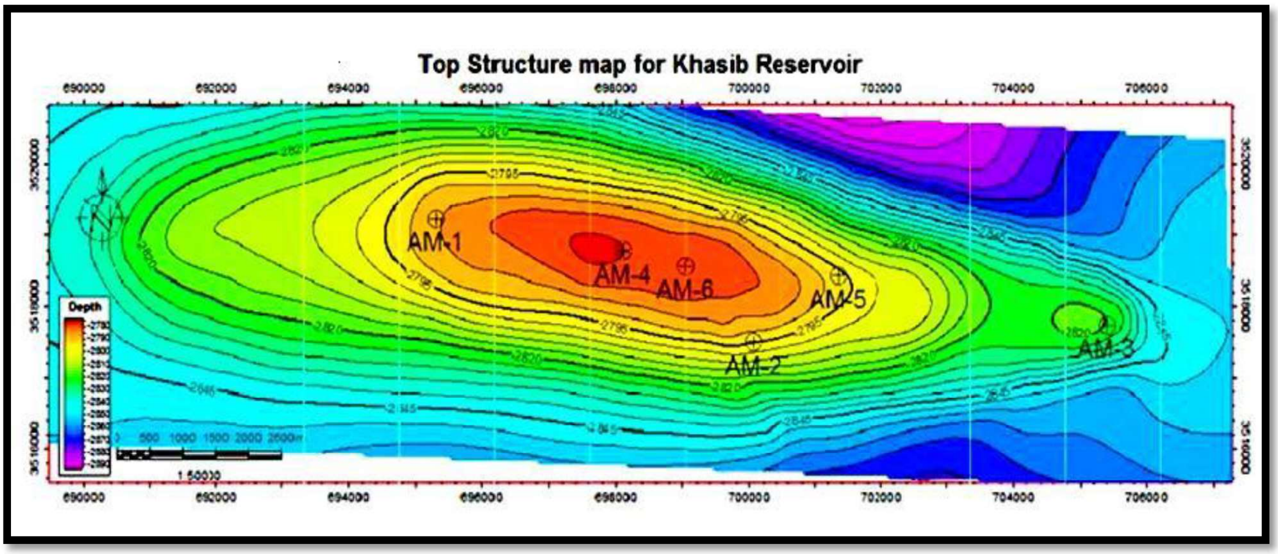


Fig. (14) Structure map of Khasib Formation in Amara oil field with the location of its wells.

References:

1. Al-Baldawi, B.A.,. Petrophysical evaluation study of Khasib Formation, Amara oil field, South Easren Iraq, Arab J Geosci,online(2014).
2. Aqrawi A. A. M.,; Carbonate- siliclastic sediments of the Upper Cretaceous (Khasib, Tannuma, and Sa`di formations) of Mweopotamian basin. Marine and Petroleum Geology, Vol. 13 No.7, pp. 781-790, (1996).
3. Bellen, R. C., Dunnigton, H. V., and Wetzel, R. and Morton, D. M.,; Lexique Stratigraphique international: 03 10Asie (Iraq), 333p, (1959).
4. Gill D., Shomrony A., Fligelman H.: Numerical zonation of log suites and logfacies recognition by multivariate clustering. AAPG Bull. 77(10): (1781-1791).(1993).
5. Jassim S. Z., and Goff J. C., Geology of Iraq. Dolin, Prague and Moravian Museum, Brno. 341p. (2006).
6. Missan Oil Company. Final well reports for (Am-1, Am-4, Am-5, Am-6) Reservoir and Fields Development Department.
7. Serra O: Fundamentals of well log interpretation. The interpretation of logging data, Elsevier, Amsterdam, Vol. 2, (1986).