Petroleum Res. J. Vol. 13 (2001); P. 57–64, 10 Figs.

The Role of Principal Component Analysis (PCA) in the Delineation of Lithofacies Based on Well Logs

Omar E. Suleiman* and Zoltan Kelemen*

دور تحليل المركبات الأساسية في وصف السحنات الصخري باستخدام سرود الآبار

عمر محمد سليمان وزلطان كيليمان

تم استعمال مجموعة من سرود الآبار في حوض غدامس للمقارنة بين فعالية الطرق المختلفة المستخدمة في تعيين حدود السحنات الصخرية بالبئر. وقد وجد أن التحليل الإحصائي الكمي للتركيبة الصخرية باستعمال ستة سرود أكثر فعالية، حيث أنه يستعمل المعلومات الجيولوجية خاصة المتعلقة منها بأنواع المعادن الداخلة في تركيبة الصخور أو التكاوين المختلفة.

وكنتيجة لذلك فهي تعطي الأحجام الجزئية للمعادن الداخلة في التركيبة الصخرية والمسامية، وحجوم الصخر الطيني ومعادن أكسيد الحديد. ويمكن تقسيم أي مرحلة من العمق إلى أجزاء من صخور مسامية وأخرى يتداخل فيها حجر رملي وطفلة على هيئة رقائق، وطفلة غير منفذة.

تم تطبيق التحليل البياني للمركبات الأساسية في نفس البئر لتحليل السحنات الصخرية. كما تم إختبار الأصناف المختلفة لهذه السحنات على أساس الرسم البياني التقاطعي بين المركب الأساسي الأول والمركب الأساسي الثاني ورسمت الحدود بين الطبقات أو التكاوين المختلفة بناء على ذلك.

طبقت طريقة التحليل الاحصائي الكمي في حالتين، أحدهما باستخدام ستة سرود، لتعيين التركيبة الصخرية، والحالة الأخرى باستخدام قياس أشعة جاما، وقياس التوصيل الكهربائي. وفي كلتا الحالتين تم تعيين السحنات الصخرية بناء على تحليل المركبات الأساسية.

وتمت مقارنة النتائج مع طريقة التحليل التفصيلي للتركيبة الصخرية، كما استخدمت هذه الطريقة أيضاً في تقدير الحجم الجزئي للمسامية والطفلة ومعادن أكاسيد الحديد.

Abstract: *A rich set of well logs were measured in a well in Ghadamis Basin, Libya. These data are used to compare the effectiveness of the different methods used to delineate lithofacies boundaries in the borehole.*

Quantitative statistical lithology analysis, based on six logs, is found to be the most effective. It utilizes geological information, particularly the knowledge of mineral types occurring in the formations. It provides volume fractions of minerals including porosity, volumes of shale and ferroan minerals. The depth interval can be divided into segments of porous reservoir rock, sandwich-type development with alternating sand and shale laminae, and impermeable shale.

The statistical method of principal component analysis (PCA) is applied for lithofacies analysis

---

* Petroleum Research Centre, P. O. Box 6431, Tripoli, Libya.

in the same well. Lithofacies categories are selected on the crossplots of the first principal component, (PC1) vs. the second principal component, (PC2). Formation boundaries are delineated based on lithofacies analysis.

Two PCAs were applied: one from the six logs, also used for lithology interpretation, the other from two inputs: gamma ray and conductivity. Lithofacies were delineated based on both of the PCAs and compared with the results of detailed lithology interpretation. PCA is also used for volumetric estimation of porosity, volumes of shale and ferroan minerals.

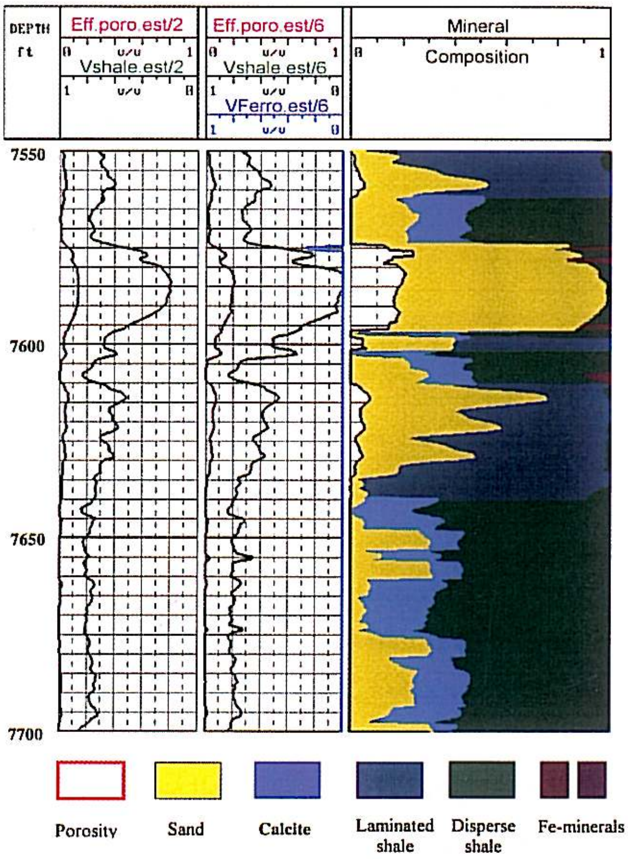

Fig. 1a. Strip log of lithological interpretation (well A1-NC162, 7550′ - 7700′).

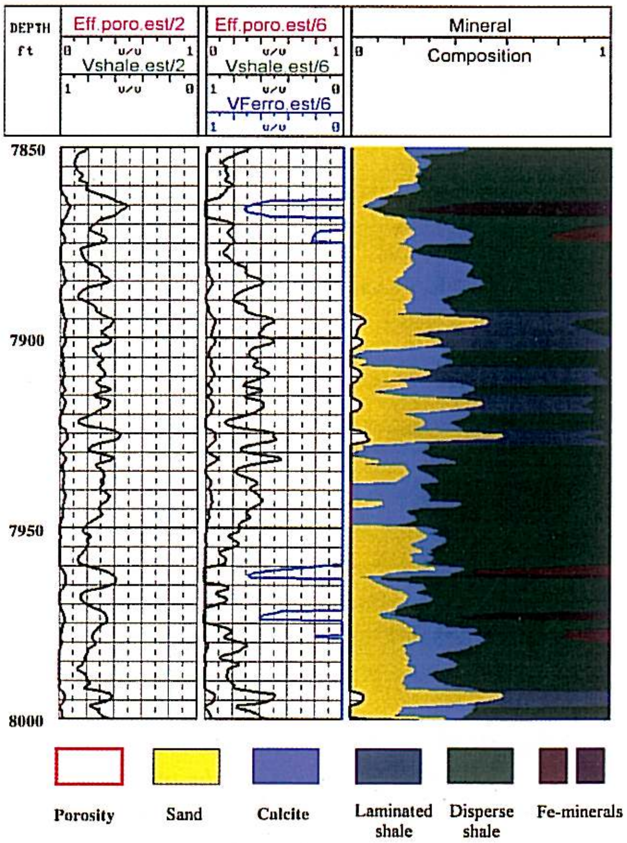

Fig. 1b. Strip log of lithological interpretation (well A1-NC162, 7850′ - 8000′).

THE TASK OF LITHOFACIES ANALYSIS

Lithofacies are typical associations of rock forming minerals, *e.g.* sandstones, shales, carbonates, shaly sandstones *etc.* Recognition and understanding of lithofacies are important in geological research and oil exploration. Well log measurements, supported by cores, provide information on lithology. The delineation of lithofacies categories *vs.* depth creates the base for the determination of formation boundaries which is crucial in basin and reservoir analysis.

The methods applied for this task are exhibited in well A1-NC162 in the Ghadamis Basin of western Libya. A rich set of well logs was measured in this well, and subsequently a detailed analysis of lithology was performed. Statistical qantitative lithology interpretation, based on six input well logs, was carried out. The input logs are: bulk density (ρ_b), neutron porosity (Φ_n), photoelectric effect (Pe), sonic Δt, potassium and thorium content (the last two are components of spectral gamma ray measurement).

In the lithological interpretation 12 different minerals or lithological rock components were determined. Different rock models were applied with five or six components simultaneously; at each depth site the model, minimizing the statistical error (incoherence) was selected. The lithology (in simplified form) is displayed on the upper track of the strip log (Fig. 1a and 1b). Porosity, sand, calcite, laminated shale, dispersed shale (clay), ferroan minerals (oxides

& hydroxides) are displayed in variation diagrams.

Some typical rock developments can be recognized on the strip log. Clean sandstone with 15 – 20 % effective porosity and subordinate shale appears at the drill depth of 7575′ – 7595′. A strip of shale with no porosity, 50 – 60 % dispersed clay with some calcite and silt is positioned above it. Sandwich-type development of alternating layers of sand and impermeable shale are present in the interval 7610′ – 7640′. Porosity increases upwards from 1 % to 7 %, while the amount of shale decreases from 95 % to 25 vol. %, the rest consisting mainly of sand.

Investigating another section of the well, a rock development typical of high concentration of ferroan minerals occurs; the volume fraction of ferroan minerals reaches 60 % at 7962′ and 7974′ and even 80 % at 7866′.

PRINCIPAL COMPONENT ANALYSIS

In basin analysis studies, where several thousand feet thick intervals of a large number of wells (over hundred) are studied, the detailed lithology interpretation is not practical. The set of available input well logs may be very heterogeneous. Other information (geological description, core data) may be inadequate and the amount of necessary work is enormous. For this task, a simpler, but more efficient way of lithofacies analysis, was developed (Marriott, 1980) in the Petroleum Research Centre.

The multivariate statistical method of PCA Marriott, 1980 is used to extract the information contained in the well logs input and is capable of revealing some hidden factors which are reflected in the values of the well logs. These factors depend on the geological environment and on the set of input data. In well A1-NC162 the same six well logs are used as inputs which were applied for the statistical lithology interpretation. Therefore, the dominant factors are porosity, shaliness and presence of ferroan minerals.

Table 1 presents the most important mathematical parameters involved in the principal component analysis in the well of this study. The matrix of correlation coefficients reveals the internal relationships between the input variables. Strong correlations occur between ρb and Pe, Φn and Δt, thorium and potassium *etc.*

Table 1. Principal component analysis from 6 input logs (well A1-NC162)

Variable	Mean	St. dev.
RHOB	2.5721	0.1306
PEF	3.6727	1.3271
NPHI	0.2523	0.0594
DT	69.2614	4.5393
THOR	11.6414	3.6456
POTA	0.0189	0.0070

Correlation matrix

Variable:	RHOB	PEF	NPHI	DT	THOR	POTA
RHOB	1.000	0.617	0.555	-0.049	0.680	0.569
PEF	0.617	1.000	0.509	0.301	0.356	0.118
NPHI	0.555	0.509	1.000	0.628	0.736	0.575
DT	-0.049	0.301	0.628	1.000	0.306	0.131
THOR	0.680	0.356	0.736	0.306	1.000	0.735
POTA	0.569	0.118	0.575	0.131	0.735	1.000

Eigenvalues and eigenvectors

	PC1	PC2	PC3	PC4	PC5	PC6
	3.373	1.189	0.947	0.220	0.164	0.108
RHOB	0.433	-0.407	0.338	-0.097	-0.276	0.669
PEF	0.342	0.158	0.747	0.287	0.338	-0.321
NPHI	0.488	0.266	-0.103	-0.076	-0.713	-0.408
DT	0.248	0.781	-0.185	0.086	0.194	0.500
THOR	0.484	-0.143	-0.234	-0.654	0.482	-0.177
POTA	0.403	-0.330	-0.478	0.684	0.176	-0.048

The other section of this Table contains the eigenvalues (reflecting the importance of the principal components) and the eigenvectors (describing how the PC-s are constructed from the inputs). It is clear that the first three principal components concenrate more than 90 % of the information contained in the inputs.

DETERMINATION OF LITHOFACIES BY PCA

All points of the investigated depth interval are shown on the crossplot of principal components PC2 *vs.* PC1 (Fig. 2). The points are colour coded according to the lithofacies categories shown on the legend of the figure.

The methodology behind creation of the lithofacies categories is the following:

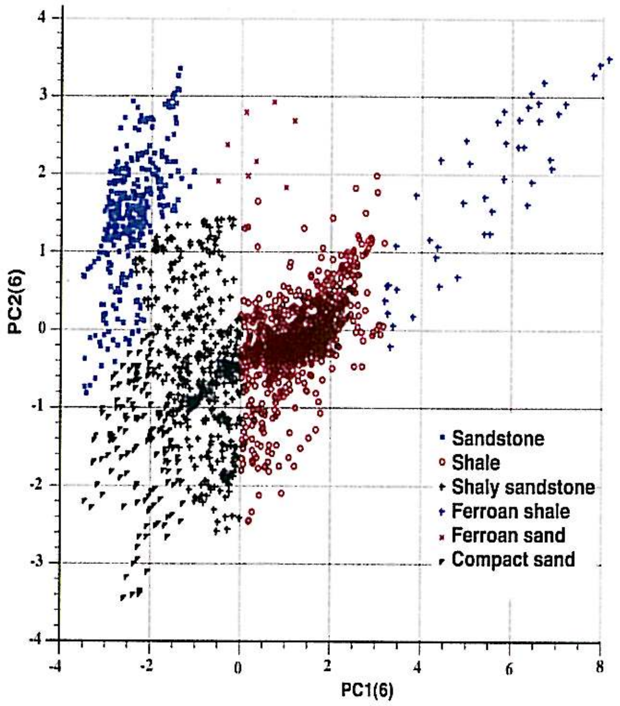

Fig. 2. PC2 vs PC1 (6 logs) (well A1-NC162).

- Recognize the clusters of concentrated points on the crossplot of PC2 *vs.* PC1;
- Identify some typical points of the crossplot on the strip log of detailed lithology evaluation; if necessary, check values of input logs;
- Determine the basic lithology type of the clusters - *e.g.* sandstones, shales, ferroan shales *etc.*
- Carefully draw the discriminating lines between categories.

The crossplot of PC3 *vs.* PC1 is presented (Fig. 3). Here the main categories of shale, shaly sand and sandstone can be distinguished, but less clearly, on a compressed scale. However, the group of ferroan shales is well separated.

Lithofacies analysis is also carried out in older wells where the set of measured well logs is poor. An experiment was made in well A1-NC162 with a two-input PCA involving only gamma ray and deep induction conductivity. Figure 4 shows the crossplot of PC2 *vs.* PC1 in the investigated interval; the points are colour coded by lithofacies categories. Here the categories of shale, sandy shale and compact sand are somewhat mixed, however, the facies of clean sandstones is well defined. This combination is not favourable for the separation of ferroan mineral concentrations.

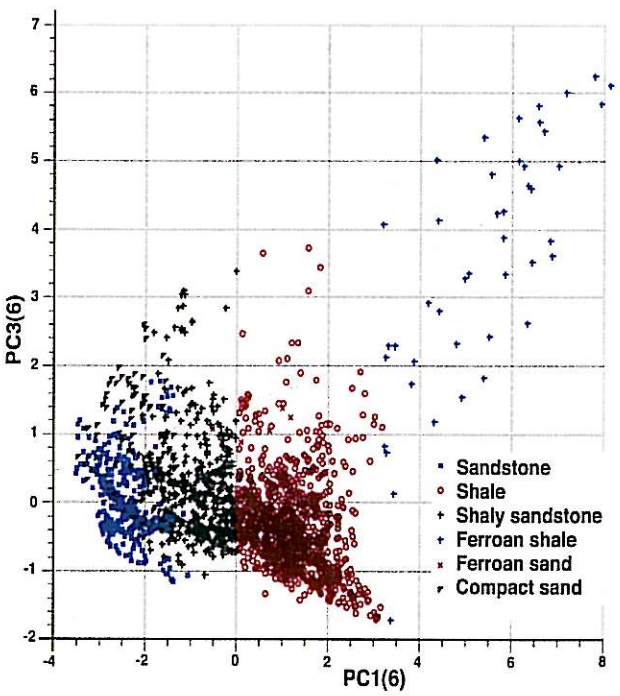

Fig. 3. PC3 vs. PC1 (6 logs) (well A1-NC162).

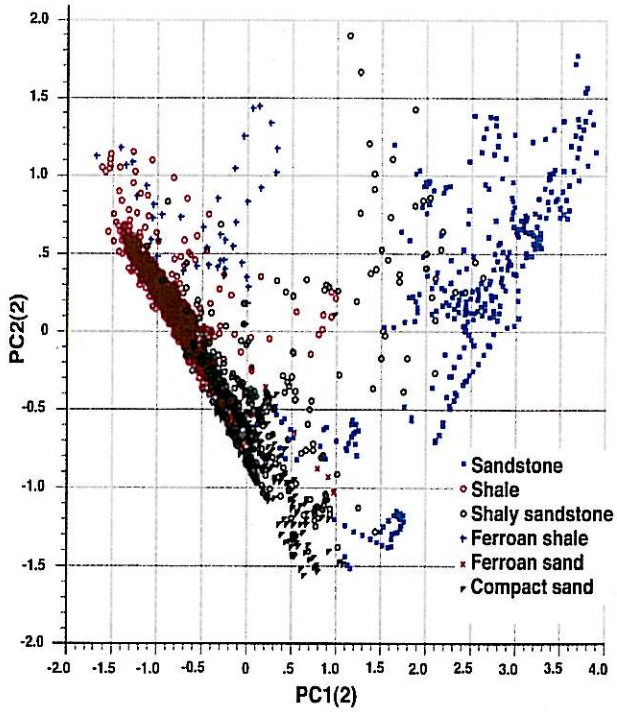

Fig. 4. PC2 vs. PC1 (PCA from conductivity +SGR) (well A1-NC162).

ESTIMATION OF ROCK COMPONENTS FROM PCA

On the previous figures, it has been shown that, there is a strong relationship between the values of principal components and the lithological composition of the rock. In particular, the main trend from impermeable shales through shaly sandstones to porous clean sandstones is clearly visible on the crossplot of PC2 *vs.* PC1. This strong relationship encourages us to estimate the value of porosity and shale from the principal components.

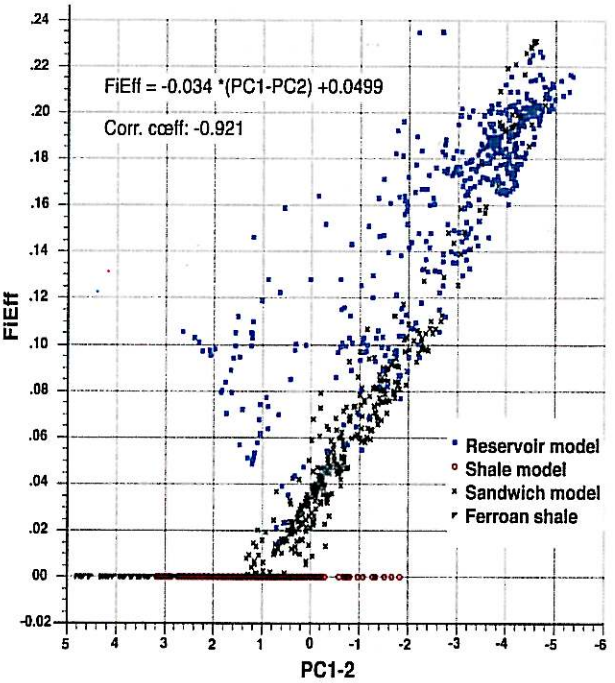

Fig. 5. Effective porosity vs PC1-PC2 (6 logs) (well A1-NC162).

The trend of increasing porosity and decreasing shaliness is going from the lower right to the upper left (Fig. 2). This direction corresponds to the decreasing values of the difference PC1-PC2. Indeed, if the crossplot of effective porosity *vs.* PC1-PC2 (increasing from right to left) is constructed (Fig. 5), then a strong correlation with a coefficient r = – 0.92 is achieved.

Applying the line of regression for the estimation of effective porosity from PC1-PC2, the crossplot of figure 6 is achieved. The shape is similar to that of Fig. 5 but the possibility of negative porosity values, from the estimation, is eliminated. The error of the estimation is 0.0237 which is a relatively small value.

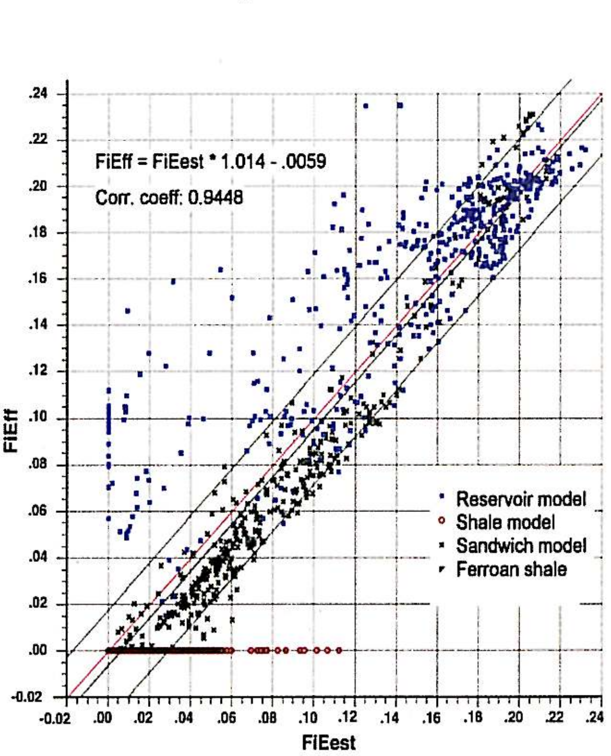

Fig. 6. Effective porosity vs. its estimation from PC1-PC2 (6 logs) (well A1-NC162).

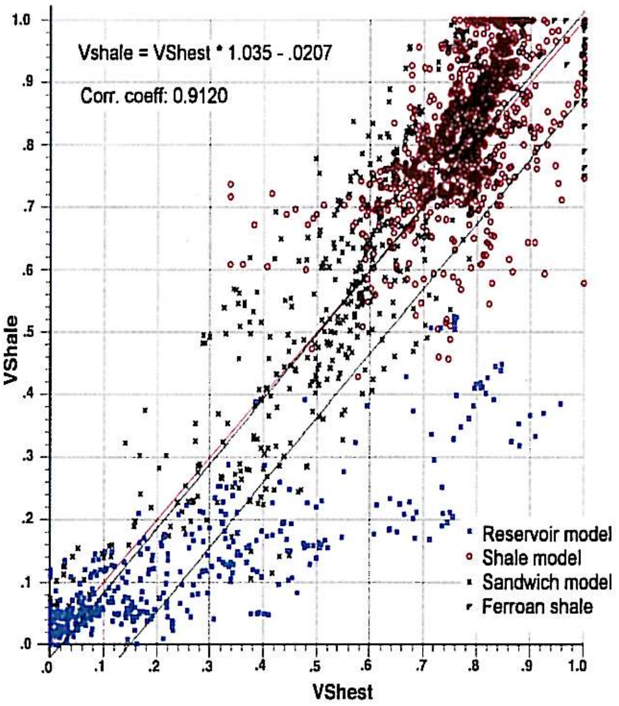

Fig. 7. Volume of shale vs. its estimation from Pc1-PC2 (6 logs) (well A1-NC162).

Estimation of the volume of shale from PCA has also been carried out. Here the difference of principal components PC1-PC2 is applied as the base of the estimation. The line of regression is used for the estimation of shale volume with the necessary restrictions (Vshale >= 0 and Vshale <= 1). Figure 7 shows the crossplot of volume of shale vs. its estimation from PC1-PC2 . The correlation coefficient is positive: r = 0.912 .

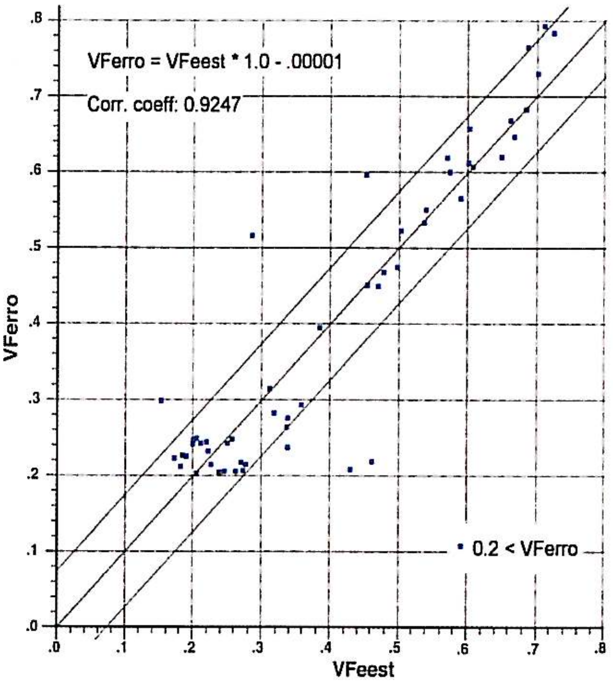

Fig. 8. Volume of ferroan minerals vs. its estimation from PC3 (6 logs) (well A1-NC162).

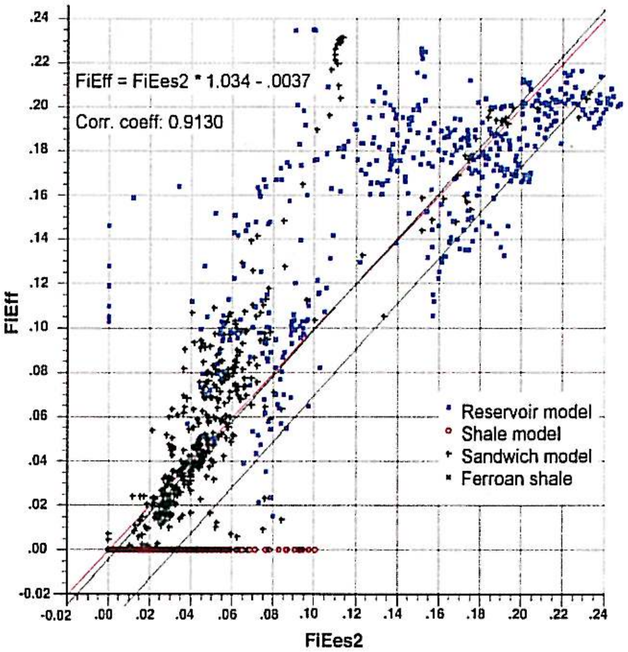

Fig. 9. Effective porosity vs. its estimation from PC1 (2 logs) (well A1-NC162).

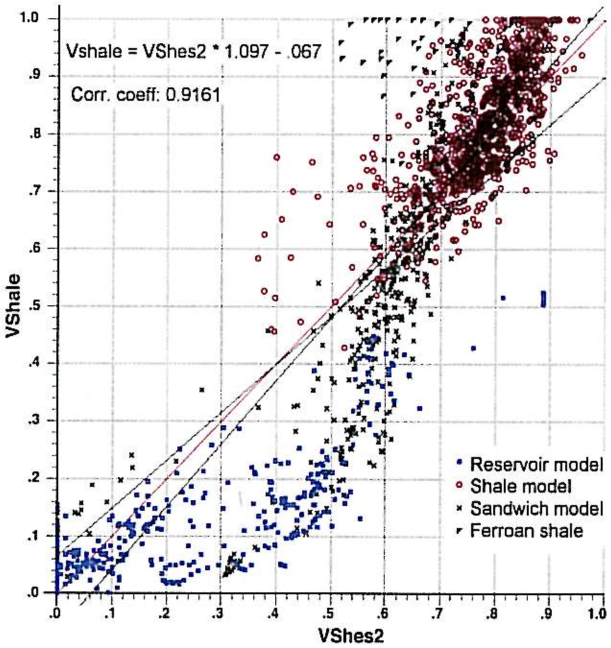

Fig. 10. Volume of shale vs. its estimation from PC1 (well A1-NC162).

As the crossplot of PC3 *vs.* PC1 (Fig. 3) suggests, the third principal component PC3 shows the largest effect of ferroan minerals. This indicates that PC3 is the best estimator of volume of ferroan minerals. In this estimation we restricted the estimation only to the depth sites where the amount of ferroan minerals exceeded 20 %. The results are shown on figure 8.

Figure 4 shows that 2 input PCA from gamma ray and conductivity are suitable for the separation of porous clean sandstones. This suggests that quantitative estimation of porosity and shale volume from PCA is effective. The technique of estimation is similar to that applied from 6-input PCA, however, the correlation coefficient is weaker. The result is shown in figures 9 and 10.

CONCLUSIONS

The method of principal component analysis for lithofacies determination is extensively used in Petroleum Research Centre, Tripoli for basin analysis. Its advantages are:

- High productivity;
- Simplicity: no sophisticated rock models are necessary;
- The method is robust; effects of assumptions of the user are minimal.

In practical use, the method of PCA should be calibrated on detailed quantitative analysis of lithology. In a comprehensive analysis of a hundred wells in a baslin with similar geological conditions, a few wells should be selected where the set of input well logs and other geological information is the richest. In these wells quantitative lithological interpretation as well as principal component analysis should be made. The selection of lithofacies categories is based on the lithological interpretation. These categories are recognized on the crossplots of principal components. In other wells, only PCA is carried out and similar patterns are searched for on the crossplots to determine lithofacies categories, porosity and shalines from PCA alone.

REFERENCE

[1] Marriott, F. H. C., 1980. *The Interpretation of Multiple Observations*. Academic Press, pp. 18 – 25