CLUSTER ANALYSIS OF
THE BALAKHANY VIII RESERVOIR UNIT
WITH SPECTRAL GAMMA RAY LOGS
AZERI-CHIRAG-GUNASHLI FIELD, OFFSHORE AZERBAIJAN

A Thesis

Presented to

the Faculty of the Department of Earth and Atmospheric Sciences

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Melahat Asli Surek

December 2013

CLUSTER ANALYSIS OF THE BALAKHANY VIII RESERVOIR UNIT WITH SPECTRAL GAMMA RAY LOGS AZERI-CHIRAG-GUNASHLI FIELD, OFFSHORE AZERBAIJAN

_____

Melahat Asli Surek

APPROVED:

_____

Dr. Regina Capuano, Chairman

_____

Dr. Donald Van Nieuwenhuise, Co-Chairman

_____

Dr. Shuhab Khan

_____

Dr. Jeffrey Yarus

_____

Dean, College of Natural Sciences and Mathematics

ACKNOWLEDGMENTS

I would like to express my appreciation to my advisor, Dr. Donald Van Nieuwenhuise, for his guidance, support, and dedication to this research. He always encouraged me and made me believe that I can achieve. I also would like to thank to my thesis committee members Dr. Regina Capuano, Dr. Shuhab Khan, and Dr. Jeffrey Yarus for their help and contributions to this project.

Special thanks to Turkish Petroleum Corporation, for giving me the chance to do my Masters in the USA with a scholarship.

Finally, my deepest gratitude goes to my family and my friends for their support through my education.

CLUSTER ANALYSIS OF
THE BALAKHANY VIII RESERVOIR UNIT
WITH SPECTRAL GAMMA RAY LOGS
AZERI-CHIRAG-GUNASHLI FIELD, OFFSHORE AZERBAIJAN

A Thesis

Presented to

the Faculty of the Department of Earth and Atmospheric Sciences

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Melahat Asli Surek

December 2013

ABSTRACT

The objective of this research was to discriminate the source of reservoir sediments coming from different paleo-river systems, both in space and time, to infer the provenance of the sediments. Spectral gamma ray logs were used to identify different compositional suites of sediments vertically and laterally that could affect the multiple sources of Balakhany VIII sands of the South Caspian Basin (SCB) and the impact of those compositional suites on reservoir quality. Three radioactive elements, uranium, thorium, and potassium, all related to different mineralogical suites, were used to test how these affect effective porosity and permeability.

The Lower Pliocene Productive Series containing thick fluviodeltaic rocks are the major hydrocarbon reservoirs in the Azeri-Chirag-Gunashli field. Basin isolation, uplift of the basin margins, and fall of base level resulted in the discharge of several river systems into the SCB; most notably the paleo-Kura, paleo-Volga, and paleo-Amu-Darya. The Balakhany Suite is part of the Productive Series, and the Balakhany VIII, which is the focus of this study, differs from other sub-units by the prevalence of fluvial sediments facies.

Cluster analysis was used to understand the distribution of spectral gamma ray log elements. Vertical and lateral clusters from different wells among the study area were used to infer the different petrofacies and the source of the reservoir sediments. It is

important to understand the mineralogical preferences of Th, U, and K in relation to deposition, transport, weathering, and diagenesis.

In contrast to previous studies, the results of this study show the dominant sediment contributors are the paleo-Amu-Darya and paleo-Uzboy Rivers, rather than the paleo-Volga as previously though. The dominant systems brought sediments from the eastern side of the study area from outcrops of sediments within the Balkhan, Kopet Dag, Pamir and Tian-Shan mountain ranges. Major petrofacies are established by the clustering routines. Results of 13 wells giving consistent results for sands and all lithologies showing dominant sediment contribution from eastern side. Sand-rich petrofacies A having better reservoir qualities are observed mostly in the western side of the study area, and is derived from N-S draining paleo-Volga. The shalier petrofacies B, having less but good porosity and permeability averages and higher K average, is derived from eastern mud-dominated sediment contributors, the paleo-Amu-Darya and paleo-Uzboy rivers.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I: INTRODUCTION

1.1 Introduction, Background, and Location

The Azeri-Chirag-Gunashli (ACG) field lies approximately 120 km southeast of Baku, in the South Caspian Basin (SCB). This field covers an area of 432 square kilometers and has estimated oil reserves in excess of 5.4 billion barrels of oil. The field is located in water depths ranging from 100 m to 400m. The ACG field is approximately 40 km in length and 11.5 km wide and it lies at boundary of Middle and South Caspian (the Apsheron Ridge) (Figure 1.1).

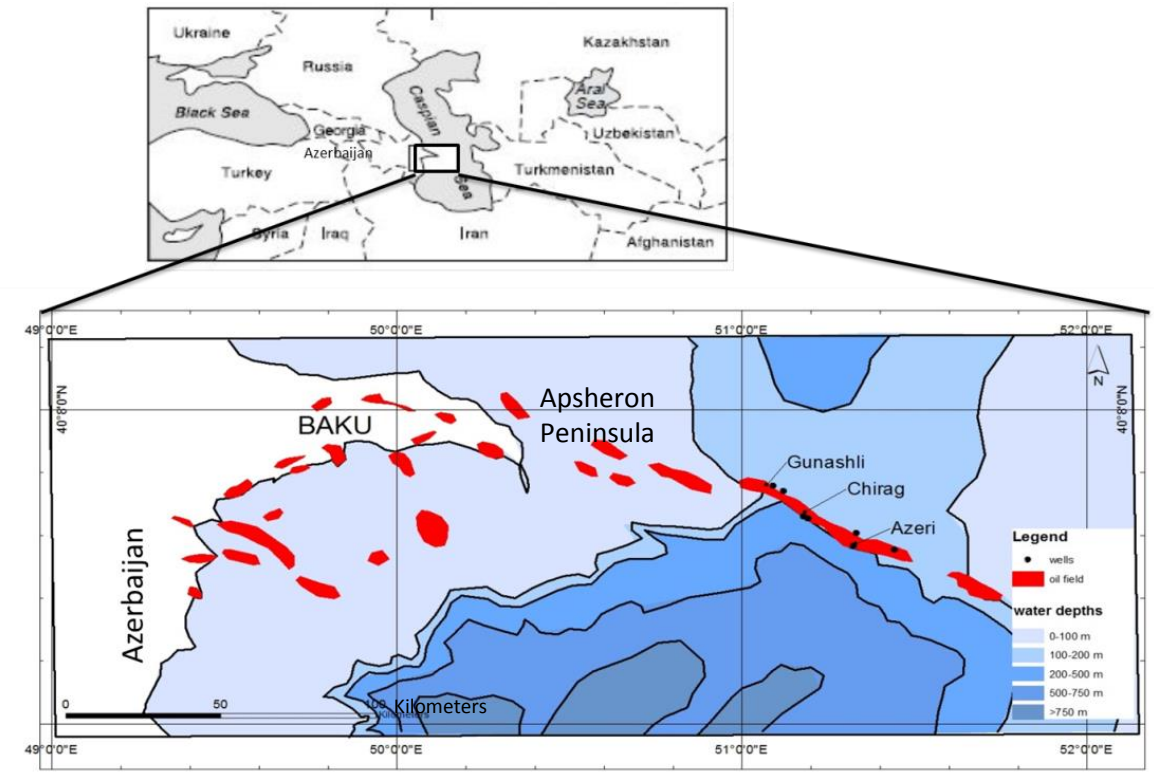


Figure 1.1: The Apsheron Peninsula and part of South Caspian Basin showing the oil fields that lie on the Apsheron Ridge (Modified from Reynolds et al., 1998).

Uneven topography, natural gas seeps, mud volcanoes, gas charged sediments, and subsea mudflows are the characteristics of the area. Below the sea bed, 2500 m and 3000m are the depths where the primary oil bearing zones occur (BP EIA report, 2007). A total of 1.9 billion barrels of oil were produced by ACG from 1997 to 2011 (BP sustainability report, 2011).

Sediment provenance studies are important, both to help locate reservoir rocks, and to know their compositions, which can impact their diagenesis as well as the development or obstruction of porosity and permeability. Measurements of compositional and textural properties of sediments can reveal the characteristics of source areas (Weltje and Eynatten, 2004). Sand provenance in the South Caspian Basin is a debatable issue as reservoir sediments were supplied via different paleo-river systems.

The South Caspian Basin is one of the world`s most prolific oil and gas provinces with onshore and offshore fields. SCB has been associated with oil for centuries. Probably, the first oil exploration well in the world was drilled arguably in SCB in 1848 (Narimanov and Palaz, 1995). According to U.S Geological Survey (2010), the remaining total undiscovered resources for the South Caspian Basin are 12.67 billion bbl. Active oil and gas seeps and "the presence of numerous gas-driven mud volcanoes suggest that hydrocarbons are still migrating and possibly generating within the basin" (Knapp et al., 2007).

The Lower Pliocene Productive Series, with a non-marine clastic succession reaches 5-7 km containing thick fluviodeltaic rocks, are the major hydrocarbon reservoirs in the

South Caspian Basin. The best quality reservoirs lie within Pereriva and Balakhany Suites, which are comprised of highly-amalgamated, low-sinuosity braided-fluvial deposits characterized by channelized, fine-to medium-grained sandstones (Hinds et al., 2004). The Balakhany VIII sub-suite is the main concern of this study (Figure 1.2).

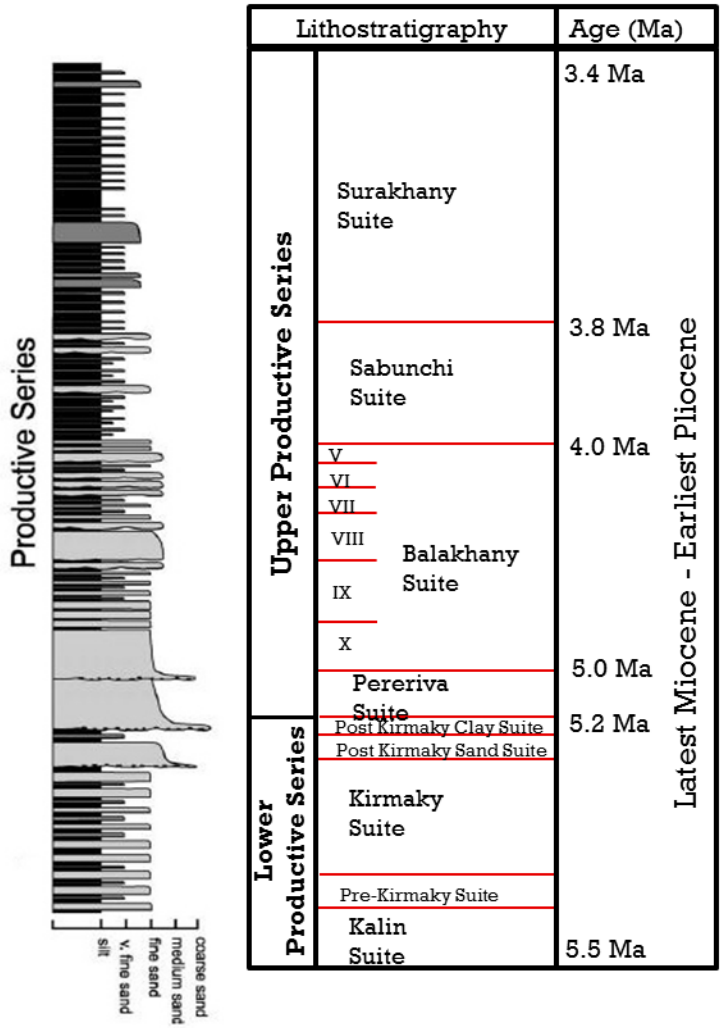


Figure 1.2: Schematic stratigraphic column of the Productive Series (Modified from Hinds et al., 2004 and Vincent et al., 2010).

Productive Series strata in the South Caspian Basin were primarily supplied by three major river systems, the paleo-Volga, the paleo-Amu-Darya, and the paleo-Kura. In addition to these, there are number of minor river systems for sediment supply (Vincent et al., 2010).

Chapter 1 of this thesis focuses on the introduction, background, location, research objectives, data availability, and brief review of previous studies.

Chapter 2 gives information about the geological setting of the SCB. This includes the regional setting and tectonic configuration of the SCB.

Chapter 3 discusses the stratigraphy, depositional facies, and reservoir geology of the Productive Series deposition. The Balakhany Suite is emphasized because of its reservoir potential. The paleogeography of the SCB is also discussed.

Chapter 4 describes the methodology of cluster analysis and how it operates by using NTSYS pc Version 2.2 (Numerical Taxonomy and Multivariate Analysis Systems) software (Rohlf, 2005).

Chapter 5 gives the results of various cluster analyses and their summary statistics.

Chapter 6 discusses the results and what they suggest regarding the sediment provenance in the study area and conclusions that can be postulated.

1.2 Research Objectives

The purpose of this research is to discriminate the source of reservoir sediments that potentially come from different paleo-river systems, both in space and time, using Spectral Gamma Ray (SGR) log data and porosity-permeability data for Balakhany VIII sands in the Azeri-Chirag-Gunashli field. Sand provenance is a major issue in exploration in South Caspian Basin. Most of the previous provenance studies involve obtaining and analyzing geochemistry of core data such as heavy minerals or other compositional components. However, when core data is not available, SGR log data have the potential to be used in such studies. SGR data are more readily available and are cheaper than coring. SGR log data have the potential to fill the gaps and provide additional insight over standard core analyses.

The natural radioactivity of a formation is recorded by gamma ray logs. Three elements and their decay chains are responsible for the radiation emitted by rock which are potassium (K), thorium (Th), and uranium (U). These three radioactive elements from SGR logs and their elemental ratios can be related to different mineralogical suites which may reflect rock compositions that could be conducive of effective porosity and permeability or not.

During initial weathering, erosion, and transportation of sediments, the compositional and textural characteristics of the initial detritus are changed. Further weathering, transportation, and post-depositional processes may have a considerable effect on final

grain assemblages (Weltje and Eynatten, 2004). Nevertheless, radioactive elements can give information about the source, given that the final spectral gamma response reflects the initial composition and the expected changes based on weathering conditions imposed by climate and distance of transportation. Therefore, it is important to understand the mineralogical preferences of Th, U, and K in relation to deposition, transportation, weathering, and diagenesis in order to interpret the patterns related to source terranes and their compositional alteration through succession from source to final deposition (Hesselbo, 1996).

In this study, a cluster analysis approach is used that groups a set of objects in such a way that similar objects or SGR samples will be in the same group or cluster. This helps us visualize the lateral and vertical distribution of similar SGR compositions to infer the source of reservoir sediments. Average porosity and permeability for individual clusters are calculated to see if there is a relationship between compositional end members and reservoir quality.

1.2 Data Availability

A total of 92 wells from BP were available in the study area. However, only 23 have spectral gamma ray logs (SGR). The logs are in log ASCII standard (.las) format. IHS`s PETRA and Hampson-Russell`s Geoview software are used for log analysis. All well logs with spectral gamma ray were screened and only 13 of them (Figure 1.3) and (Table 1.1) have SGR values for Balakhany VIII sub-suite, the main objective of the study. Some wells have very low potassium readings or no potassium readings at all.

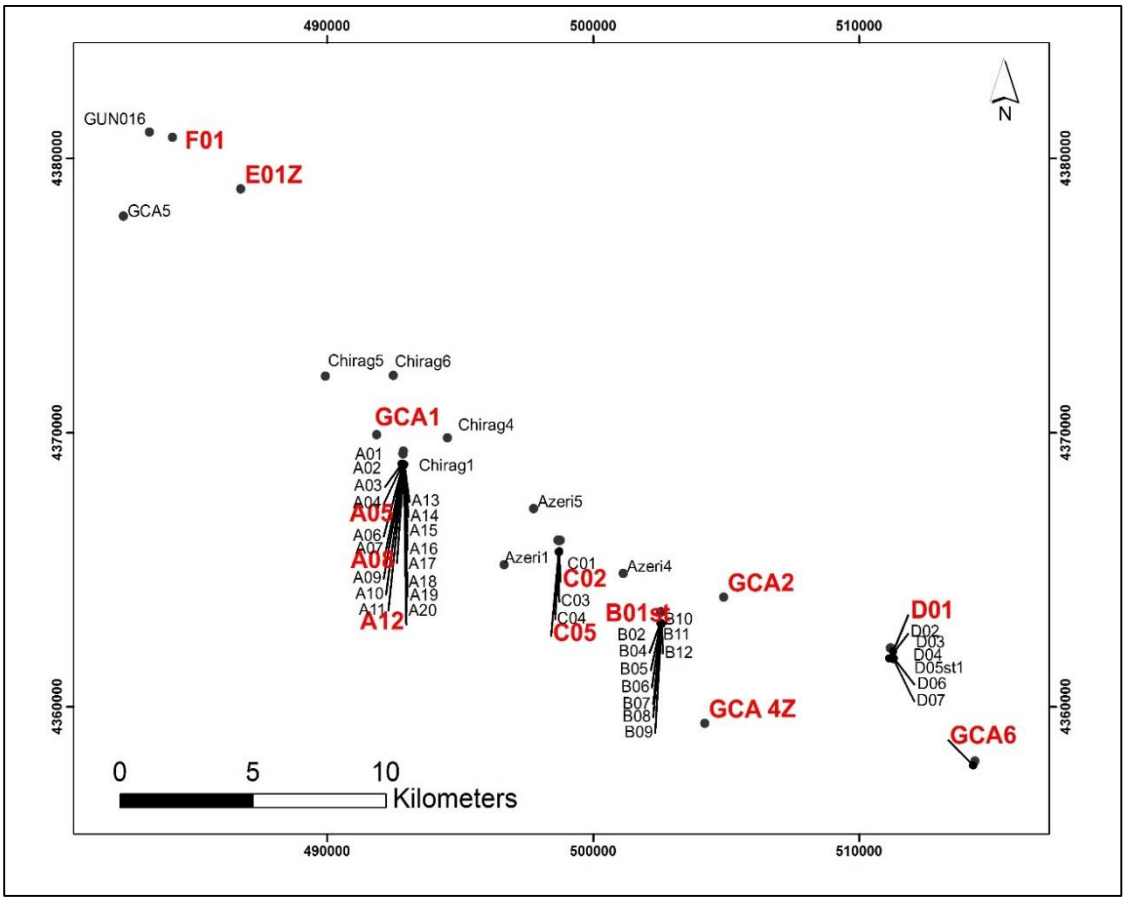


Figure 1.3: Map showing locations of all wells. The 13 key wells used in this study are highlighted in red. See Table 1.1 for names and locations of the key wells.

WELL NAME	EASTING (X)	NORTHING (Y)
E01Z	486758	4378897
F01	484180	4380776
GCA1	491862	4369928
A05	492858	4369240
A08	492860	4369232
A12	492856	4369233
C02	498689	4366086
C05	498685	4366081
B01ST	502549	4363465
GCA2	504913	4364007
GCA4Z	504203	4359404
D01	511186	4362153
GCA6	514363	4358028

Table 1.1: Coordinates of 13 key wells (Coordinate system: WGS 1984 UTM zone 39N).

1.4 Previous Works

Baghirov (2007) focused on defining the geometry of individual channels using measurements of outcrop analogs, well-to-well correlations, empirical equations involving maximum channel depth and channel belt width, and amplitude analysis of 3D seismic data. He established a stratigraphic framework and investigated heterogeneities within the Balakhany VIII. His models show that the Balakhany VIII interval is represented by superimposed channel/channel belt sandstone bodies that are stacked vertically and have laterally restricted geometries, rather than flat, sheet-like sandstone bodies seen on the original correlation.

Esedo (2009) focused on improving stratigraphic correlations and tried to investigate the feasibility of discerning multiple sand sources for the Balakhany VIII sub-suite. He defined chronostratigraphic surfaces (Figure 1.4) and identified clay mineral types within the sand-rich suite by using elemental ratios. Using Th/K crossplot method, mixed layer clays, illite, and smectite have been found to be the dominant clay mineral types. The observed differences in the mineral distribution in sediments are attributed to provenance. In addition to main sediment sources, which are a sandy paleo-Volga and a shalier paleo-Amu Darya from the East, he stated that there exists a third compositional population that potentially could be the product of a NE-SW-flowing paleodrainage system.

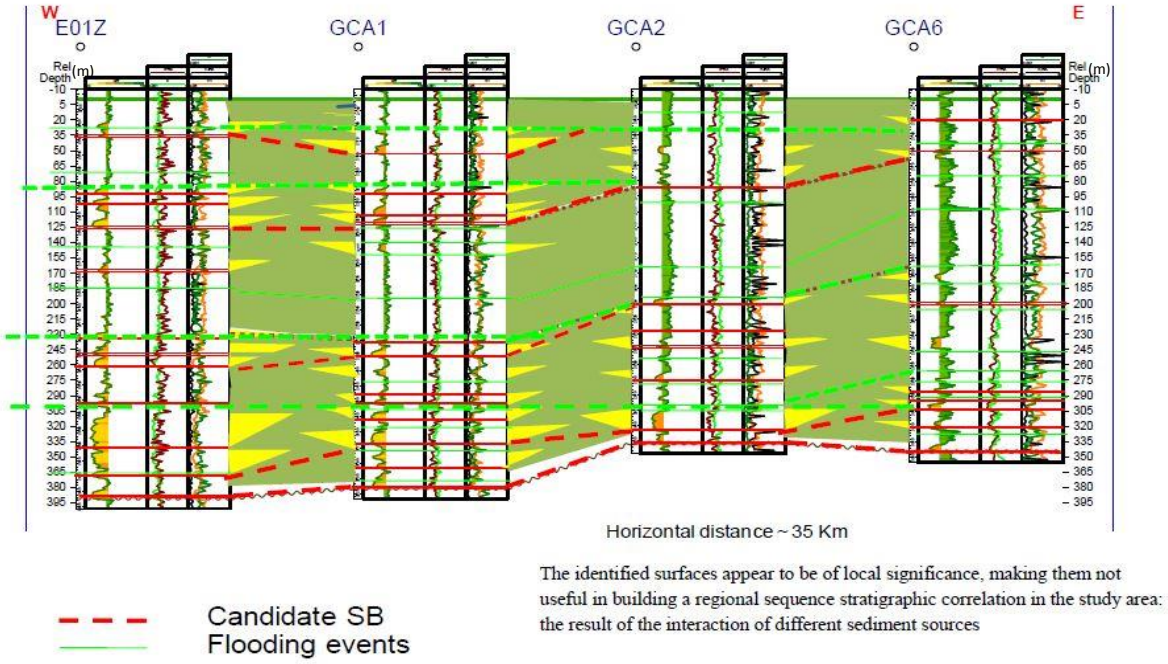


Figure 1.4: Correlation panel for some of the key wells in the study area identified from the SGR logs yellow color shows channels (Esedo, 2009).

Various authors discussed the use of spectral gamma ray log data in sediment characterization and stratigraphic correlation studies. Adams and Weaver (1958) inferred sedimentary processes of deposits by using Th/U ratios. Hurst (1990) worked in the Norwegian continental shelf and used Th/K ratios and cross plots for mineral identification. Doveton and Marriman (2004) have used Th/K and Th/U ratios to determine the oxidation reduction state of depositional environment together with the clay mineral type. Anjos et al. (2008) used Th, U, and K as tracers of the mineralogical properties of beach sands in order to understand provenance and transport processes of sediments along the southeastern Brazilian coast. They used Th/U ratios as proxy for the redox conditions of the depositional environment and proved the usefulness. Positive correlation with the geological evolution of sandy coastal deposits was observed.

Various authors have used cluster analysis in geologic and geochemical studies. Templ et al. (2008) applied cluster analysis to regional geochemical data to show problems and possibilities. They concluded that cluster analysis can be helpful in obtaining an overview of data sets that have many observations and variables as an exploratory method. Cluster analysis can be used both to structure the variables and to group the observations. In our study, cluster analysis is used to group many samples into meaningful sub groups. Pirkle et al. (1984) applied cluster analysis to a large geologic aerial radiometric data set in Copper Mountain, Wyoming and identified four stable clusters which are consistent with prior knowledge of the geology of the area. Cluster analysis results of this study is combined with the geologic knowledge of the South Caspian Basin in this research.

This study aims to understand the distribution of spectral gamma ray log elements uranium, thorium, and potassium and their elemental ratios in order to identify compositional end members and differentiate vertical and lateral clusters to discriminate sediment sources by using cluster analysis method. Using spectral gamma ray logs related to different mineralogical suites, the expected outcome is to understand if that could help or hinder effective porosity and permeabilities for Balakhany VIII reservoir unit.

CHAPTER II: GEOLOGIC SETTING

2.1 Regional Setting

The Caspian Sea is a highly prospective region for oil and gas exploration and production. It covers an area of 371,000 km^2 and it is the world`s largest lake by area and volume. The North Caspian Basin, the Central Caspian Basin, and the South Caspian Basin are the three major basins within the Caspian Basin. The South Caspian Basin is the major focus of this study (Esedo, 2009).

The South Caspian Basin (SCB) is a remnant of the back-arc basin on the margin of Tethys paleo-ocean. It is formed related to the subduction of Neotethyan oceanic crust under southern Eurasia. The SCB has been one of the major hydrocarbon producers over the last 150 years. Principal oil source rocks in the SCB are considered to be early Miocene Maykop Suite and late Miocene Diatom Suite. The SCB is a tectonically active basin and a good place to work on the sediment supply patterns (Morton et al., 2003). The basin is filled by thick (15-28km) sedimentary series and has an oceanic-type crust. Most published models suggest original spreading ages of this oceanic crust to be between Jurassic and Paleogene (Vincent et al., 2010). The Eocene Arabia-Eurasia collision resulted in closure of the Caspian Basin (Hinds et al., 2007) (Figure 2.1). At this time the Caspian Sea had a sea-level regime of its own, independent of global eustasy, as it had already lost its connection with the world`s oceans around 50-60 million years ago

during Paleocene (Kroonenberg et al., 2005). The basin is characterized by complex fold patterns deformed in the basin`s interior; many folds began to form near the end of Productive Series deposition (Devlin et al., 1999).

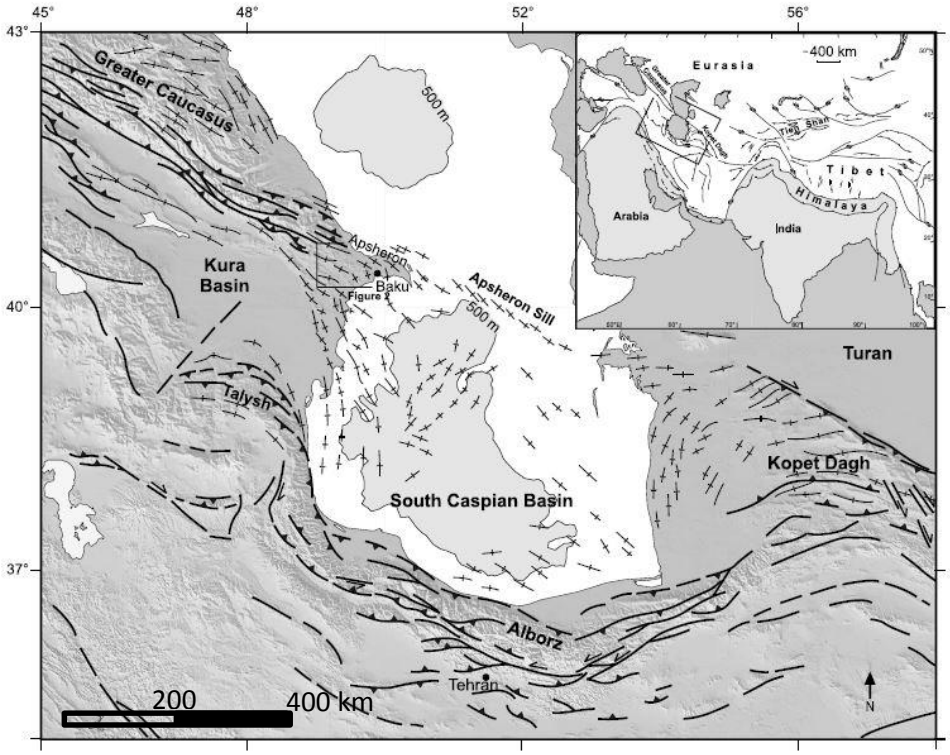


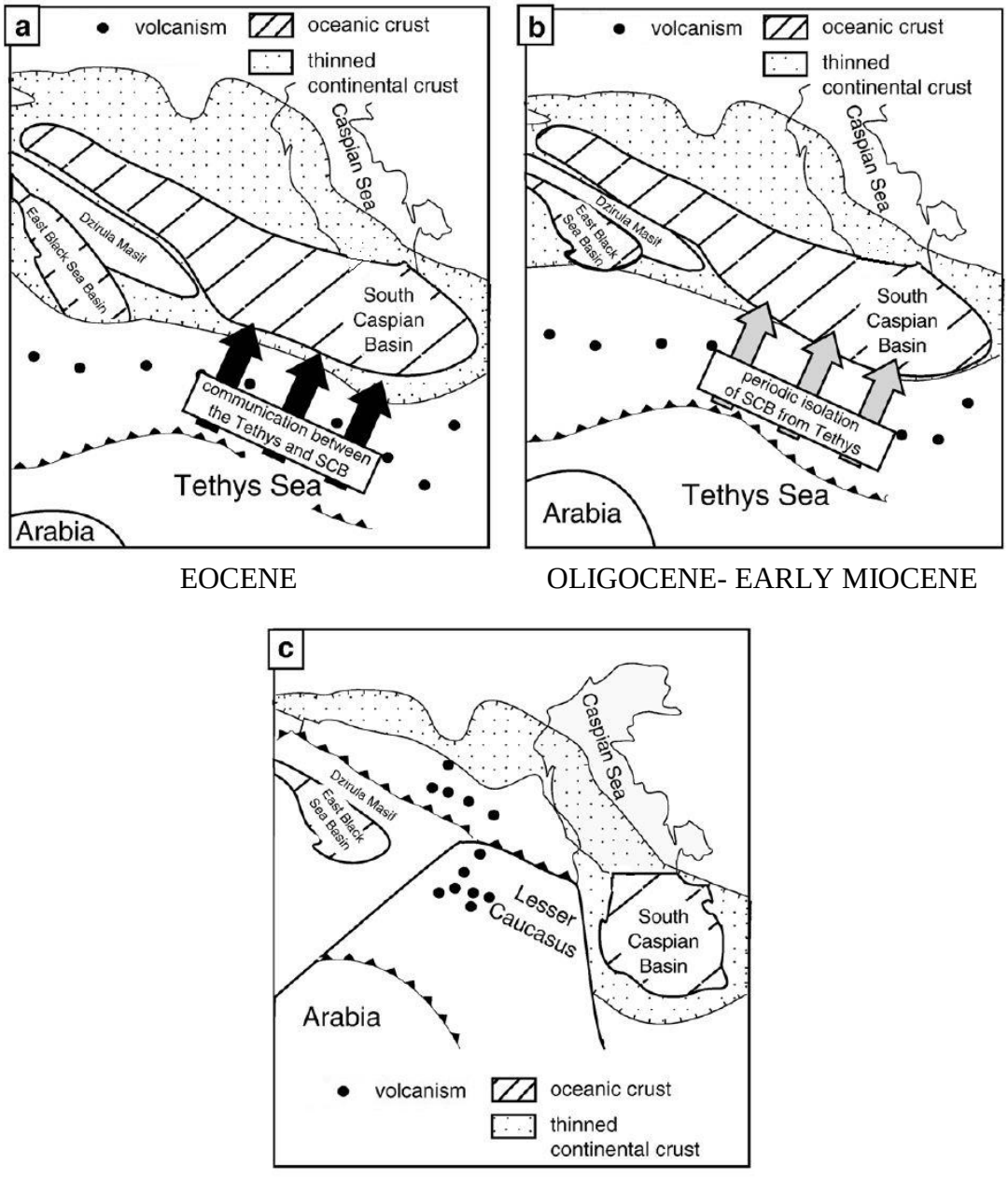
Figure 2.1: Neotectonic structures of the South Caspian Basin and its position in the broad Arabia-Eurasia collision zone (Hinds et al., 2007).

The Basin is also characterized by high subsidence and a very high rate of sediment accumulation (~4.5 km/m.y.), low sediment compaction, and a relatively low geothermal gradients (15º/100m), which has led to the preservation of reservoir quality porosity and permeability properties to depths as great as 12 km (Smith-Rouch, 2006). An important regional Middle Eocene rifting event that may have also affected the southern part of the

basin was recognized by Vincent et al. (2010). Structures are oriented northwest to southeast in the western part and northeast to southwest in the eastern part. In the western and central part of the basin, tectonic movements along Apsheron Ridge have created shale diapirs and mud volcanoes. In the eastern part slumps and growth faulting are the main characteristics. Structures deflect to northeast-southwest as they approach convergent wrench system (Smith-Rouch, 2006).

2.2 Tectonic Configuration of the South Caspian Basin

The South Caspian Basin was nearly at its largest aerial extent during the Eocene. The basin began to be isolated as an effect of subduction of Neotethyan oceanic crust under southern Eurasia occurring to the south (Figure 2.2 a). Arc volcanism and the impact of the Arabian Peninsula also affected the isolation. Continued convergence of the Arabian Peninsula, in combination with eustatic variation, periodically isolated the basin from the Tethys. Source rock of the Maykop Suite was formed during the isolation of SCB from Tethys (Figure 2.2 b). Finally, the basin was isolated from both the Tethys and the Black Sea, and began to take on its current configuration (Figure 2.2 c) (Hudson et al., 2008).



EOCENE OLIGOCENE- EARLY MIOCENE

PLIOCENE

Figure 2.2: Tectonic configuration of the South Caspian Basin (Hudson et al., 2008).

CHAPTER III: STRATIGRAPHIC FRAMEWORK AND PALEOGEOGRAPHY

3.1 Stratigraphy and Depositional Facies

A dramatic drop in base level (around 600m to 1500m) and the isolation of the South Caspian Basin in the latest Miocene initiated the deposition of the Productive Series (Reynolds et al., 1998). The Productive Series is the low-stand wedge of the most dramatic sea-level fall the Caspian ever experienced. Basin isolation, uplift of the basin margins, and fall of base level resulted in the discharge of several major river systems, most notably paleo-Volga, paleo-Amu-Darya and paleo-Kura into the restricted South Caspian Basin. As a result of the dramatic increase of sediment supply, progradation of the fluvio-deltaic Productive Series accumulated to a thickness of up to 7 km over an interval of ~2 Ma occurred (Vincent et al., 2010) (Figure 3.1).

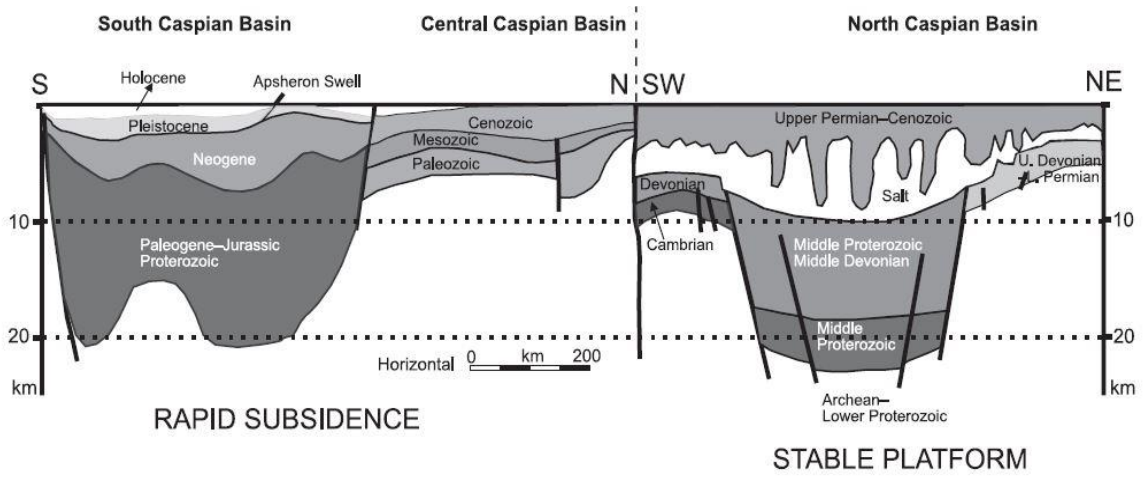


Figure 3.1: N-S cross section through the Caspian basins, showing the extreme thickness of Neogene in the South Caspian Basin and extensive Permian salt diapirs in the stable North Caspian Basin (Kronnenberg et al., 2005).

Subsidence and sedimentation balanced each other during the deposition of the Productive Series. When the rate of sedimentation and water supply roughly equaled the accommodation created over time, a balance-filled lake was formed (Bohacs et al., 2000). Evaporation is also important as it creates accommodation space. The Caspian Sea is a closed-lacustrine system that shows characteristics of a balance-filled lake type. Most of the Productive Series sediments in the northern part of the South Caspian Basin was carried by paleo-Volga River system, and is believed to be derived from Russian platforms-Urals. Rivers draining the Greater Caucasus mountain range may also have been contributing to the paleo-Volga system (Figure 3.1). At the eastern part of the basin, the paleo-Amu-Darya and paleo-Uzboy Rivers that drained the Pamir and Tian Shan mountain ranges are believed to have been the major sediment contributors. At the

western part of the basin sediments were supplied via other systems such as paleo-Kura (Reynolds et al., 1998).

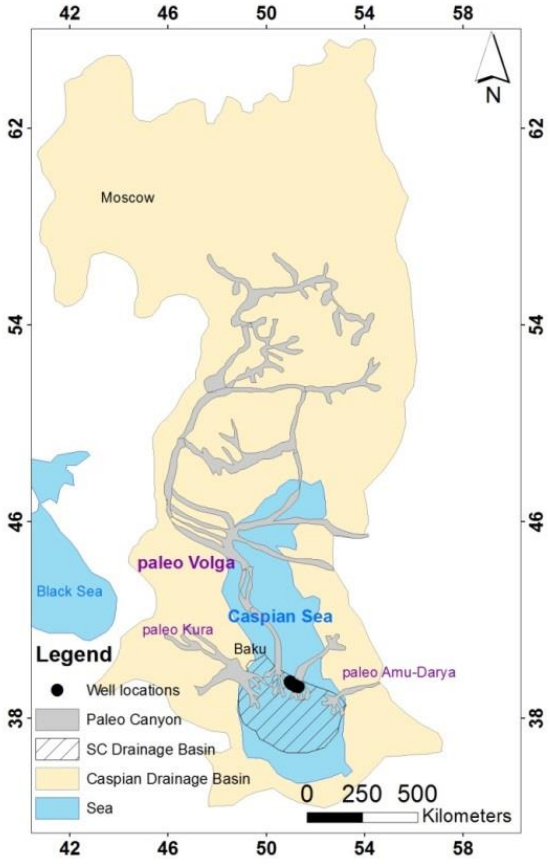


Figure 3.2: Sketch of drainage systems during deposition of Productive Series (Modified from Kroonenberg et al., 2005).

Uplift of the sediments deposited by the paleo-river systems during ongoing compressional deformation has resulted in the formation of anticlinal hydrocarbon traps. Productive Series deposition occurred around 5.5 Ma until 3.4 Ma, over ~2 million years, from latest Pliocene to early Pliocene.

The interaction between the paleo-river systems paleo-Volga, paleo-Kura, and paleo-Amu Darya and the Caspian Sea level produced different sedimentary assemblages and stacking patterns at different parts of the Caspian Basin (Abreu and Nummendal, 2007) (Figure 3.3). Horizontal arrows at the base of the figure show the influence of the deltaic systems in the different regions.

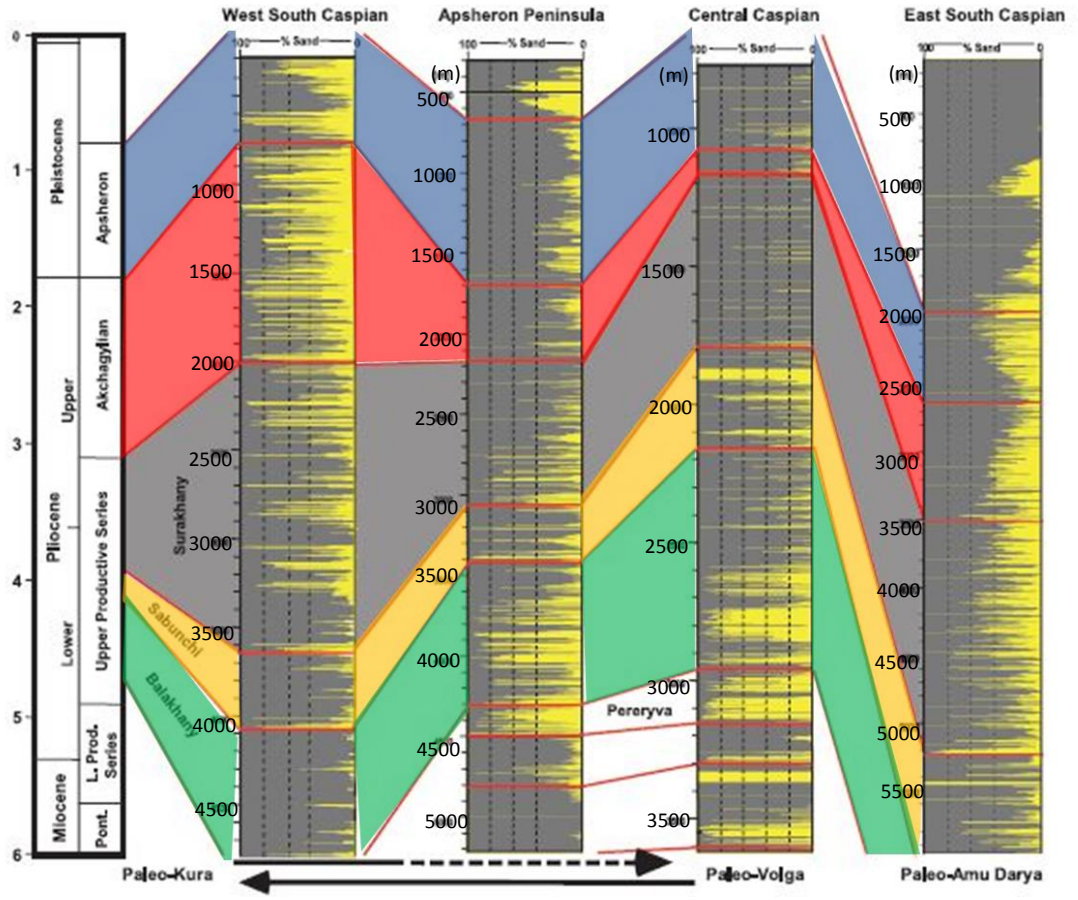


Figure 3.3: The influence of main paleo-river systems in the Caspian Basin (modified from Abreu and Nummendal, 2007).

Based on gross lithological characteristics, the Pliocene Productive Series divided into nine suites. In the South Caspian Basin, stratigraphic sequences are named as suites

rather than formations. According to Caster (1934) and Gray et al. (1972), a suite refers to lithological units that are smaller than formational rank. However the Russian use of suite is equivalent to a formation, as both are defined as mappable stratigraphic units. The Kalin, pre-Kirmaky, Kirmaky, Post-Kirmaky Sand, Post-Kirmaky Clay, Pereriva, Balakhany, Sabunchi, and Surakhany are the nine suites of the Productive Series (Figure 1.2). The Balakhany Suite is subdivided into six lithologically defined subunits, Balakhany X to Balakhany V from base to top, reflecting the nature of their occurrences in the subsurface. Units and subunits can be correlated regionally on wireline logs and have distinct palynological signatures (Figure 1.2). Pre-Kirmaky to post-Kirmaky clay suites are defined as the lower Productive Series, whereas Pereriva and overlying suites are defined as the upper Productive Series (Vincent et al., 2010) (Figure 1.2). The lower Productive Series consist of sandstone and mudstone interpreted as channelized and sheet-flood fluvial deposits, intercalated with mudstones that represent short-lived lacustrine regional transgressions of the Caspian Sea. The Upper Productive Series consist of sandstone prone intervals interpreted as deposited during periods of increased fluvial discharge and sediment supply; mudstone intervals are interpreted as being deposited during periods of decreased discharge and therefore coarse-grained sediment starvation (Hinds et al., 2004)

The Productive Series shows no variations in thickness across anticlinal fold axes on regional offshore seismic lines that indicate deformation did not begin until after deposition in the late Pliocene. As a result of post-Productive Series folding and uplift,

except the lowermost Kalin suite, the entire Productive Series was exposed in and around Apsheron Peninsula.

Hydrocarbon exploration has been focused offshore in the Caspian Basin. Because of the unconsolidated nature of the sedimentary succession, core recovery has been hampered and has prevented a detailed appreciation of the depositional environments and the architecture of the reservoir units. Fortunately, the Productive Series crops out extensively over the Apsheron peninsula and is thought to be a good analog for its offshore equivalents (Hinds et al., 2007).

A limited number of papers have been written on the sedimentology of Productive Series. Reynolds et al. (1998) described the first modern, process-based interpretation of Productive Series. Over the Aspheron peninsula, the paleo-Volga River formed a major sandy braid delta. The delta prograded, retrograded, and backsteped as a result of fluctuations in sediment supply and in base level. Laterally extensive shales corresponding to maximum base-level surfaces and flooding surfaces were deposited during phases of backstepping. The shales formed intraformational seals within petroleum accumulations by subdividing the succession. Between the shales four sandstone facies are recognized, namely: fluvial, delta plain, proximal delta front, and distal delta front. This interpretation was considered to represent the repeated juxtaposition of proximal and distal fluvio-deltaic environments in response to high frequency base-level fluctuations.

Hinds et al. (2004) reinterpreted the Productive Series on the Apsheron peninsula as being deposited in more fluvial dominated settings and proposed climatic control on sediment supply and lake level. New depositional models predict either the presence of extensive sand-rich braid deltas down dip of coeval multi-storey, multilateral channels, or the gradual termination of fluvial systems into silt-rich progradational and shallow lacustrine facies. Hinds et al. (2004) worked on paleocurrent directions. Different suites show variable directions. Paleocurrent vectors within the Balakhany Suite are mostly SSE.

Aliyeva (2003) studied geophysical log diagrams and outcrops of the Lower Pliocene Productive Series at the western flank of the South Caspian Basin in order to interpret paleofacies settings of their accumulations. Within the major productive Kirmaky, post-Kirmaky sand, post-Kirmaky shale, and Balakhany formations twenty high-frequency cycles of sea-level fluctuations in the paleo-Caspian Sea were identified. Most of the cycles were responsible for dramatic sea-level falls. With the study of climate-sensitive trace elements such as Sr and Ba, the importance of the impact of climate in high-frequency fluctuations of the Caspian Sea was understood. Tectonic processes also played an important role in the formation of small-scale sedimentary cycles. Influence of short-term cycles of sea-level fluctuations observed in lithological succession of sediments of the Productive Series. Vertical and lateral heterogeneity is seen in structure of hydrocarbon reservoirs.

Some case studies have focused on reservoir quality within the Productive Series. Morton et al. (2003) have done heavy-mineral analysis and provide a clear indication as to the provenance of particular sandstones. Sandstones in the basin are attributed to deposition by particular paleofluvial systems. The Paleo-Kura is the main sediment supplier for the western part of the South Caspian Basin, whereas paleo-Volga is main sediment supplier for the Productive Series strata in the north of the basin. However, their study area was located at the western onshore part of the SCB. Our study will investigate the sediments in offshore wells which are affected by eastern sediment sources that are most probably the paleo-Amu-Darya and the paleo-Uzboy.

3.2 Stratigraphy of Balakhany Suite

The Balakhany Suite is subdivided into lithologically defined subunits Balakhany X to Balakhany V from base to top. It is locally more than 300 m in thickness with an overall fining upwards trend. Even numbered successions represent more sandstone-rich intervals whereas odd numbered ones are mudstone dominated (Kroonenberg et al., 2005). According to Aliyeva (2003), the Balakhany VIII horizon includes evidence of three drastic sea-level fluctuations that produced braided river systems. Horizon VIII differs from others by this prevalence of sediments dominated by fluvial facies.

The average porosity range for the Balakhany VIII is 19 to 22%. Balakhany VIII contains 13% of the recoverable reserves. The API gravity of the bulk of the recoverable oil ranges from 32 to 36 and is low in sulphur and asphaltenes (Wethington et al., 2002).

Hinds et al., (2007) interpreted the Balakhany Suite as braided fluvial deposits but display features consistent with increased sinuosity relative to Pereriva. The features from which they make interpretation include predominantly fine-grain size and reduced amalgamation of sharp-based fining upwards units for all formations. The appearance of significant mudstone horizons and extensive intraform mud clast horizons are related to muds; whereas paleocurrent orientation displaying the relatively higher dispersal, increased preservation of ripple cross lamination, and evidence of lateral accretion are restricted to sandstone rich intervals.

3.3 Paleogeography

Abreu and Nummendal (2007) studied the stratigraphic evolution of the South and Central Caspian basins from the upper Miocene to the Holocene. By using two-dimensional (2-D) seismic surveys offshore Azerbaijan and Turkmenistan, seismic sequence interpretations were integrated with well log and outcrop data. Paleogeographic reconstruction of the Caspian Sea for the Pliocene and Quaternary was done by using interpretations from seismic and well log data (Figure 3.4).

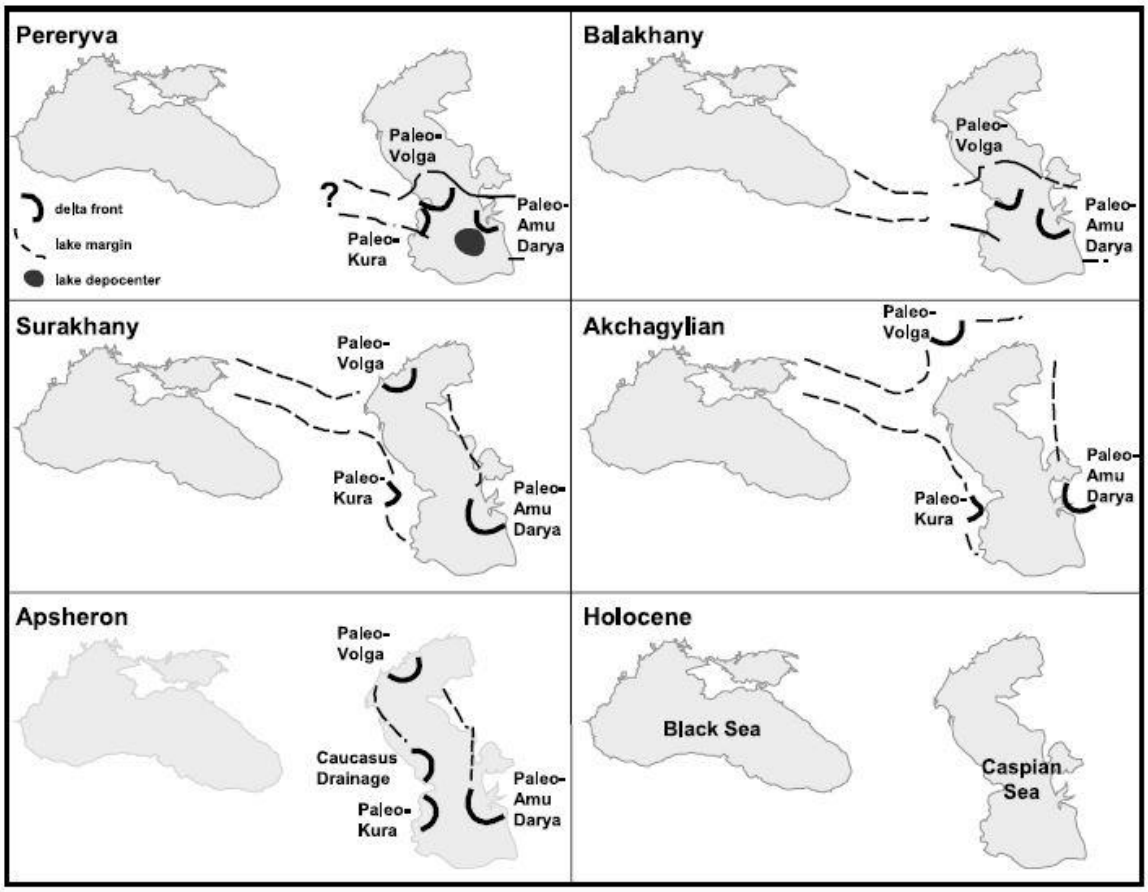


Figure 3.4: Schematic paleogeographic reconstruction of the Caspian Sea for the Pliocene and Quaternary, showing the change in the position of three main deltaic systems and the locations of the possible connections between the Caspian Sea and Black Sea (Abreu and Nummendal, 2007).

During deposition of the Balakhany Suite, when the paleo-Volga was backstepping, the paleo-Amu-Darya was prograding and becoming the most important sediment source for the South Caspian during Surakhany. The Paleo-Kura aggraded on the southern margin of western SCB during deposition of Pereriva and Balakhany Suites (Abreu and Nummendal, 2007). The relative importance of each depositional sequence or suite is shown in Figure 3.4.

CHAPTER IV: METHODOLOGY-CLUSTER ANALYSIS

In statistics, the placing of objects into more or less homogenous groups in a manner so that the relation between groups is revealed is called classification (Davis, 1979). Cluster analysis is an exploratory method of data mining categorized as classification. The principal aim of cluster analysis is the grouping of a collection of a number of objects or entities into subsets, such that the objects within each subset have same statistical relationship but the objects in one subset are different than those in another (Pirkle et al., 1984).

Grouping the similar samples on which many measurements have been made and measuring the degree of similarity between the groups is desirable in some geologic studies. Cluster analysis, a technique developed by psychologists as a method of searching for relationships in a data set, is probably one of the most useful statistical tools available for geologists (Parks et al., 1966).

In this study, cluster analysis is used to understand the distribution of spectral gamma ray log elements. Clusters among different wells located in the study area are used to infer the source of reservoir sediments. In the South Caspian Basin, there are at least three different paleo-river systems from which the reservoir sediments might have been transported.

A good outcome of cluster analysis will result in a number of clusters. Samples in a cluster are very similar to each other, whereas samples in different clusters are different

from those in the other clusters. Results of the cluster analysis can be easily understood and interpreted as they are displayed in the form of two-dimensional hierarchical diagram called a dendogram (Templ et al., 2008). The vertical axis shows the samples whereas horizontal axis showing similarity coefficient between individual samples (Figure 4.1).

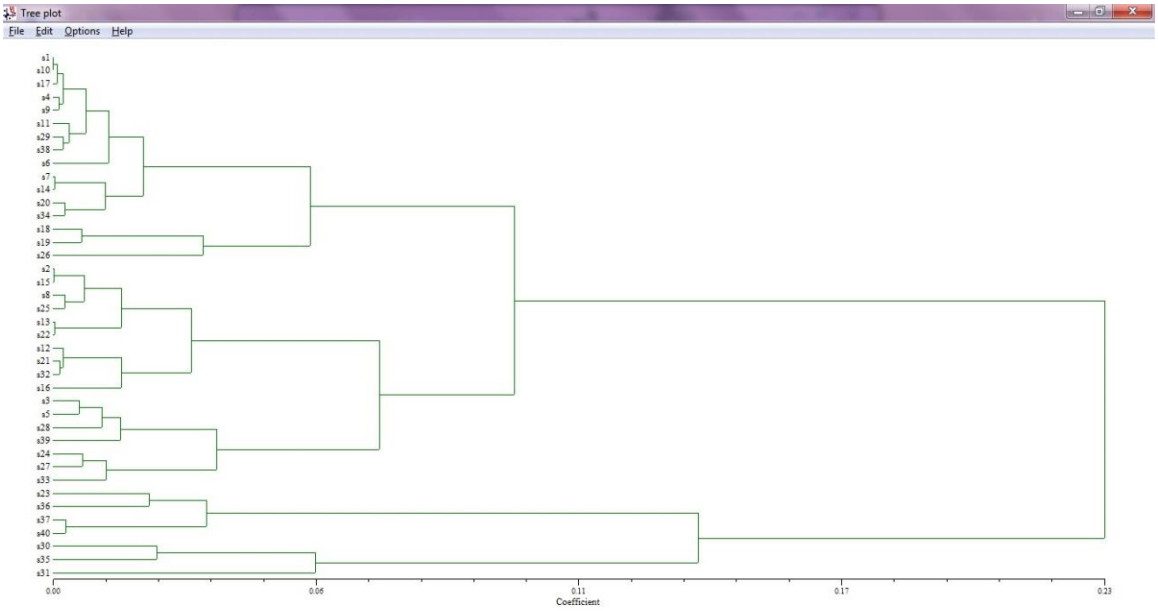


Figure 4.1: Dendogram showing the results of cluster analysis of 40 samples.

Cluster analysis requires a two-step process. First, a similarity analysis should be done between all samples and the results are shown in symmetric matrix called a similarity (or dissimilarity) matrix; then most similar things are clustered first and then the collective results of linking similar samples and sample clusters are shown on dendograms. Cluster analysis is a straightforward and logical analysis that does a pair-by-pair comparison between samples, objects or variables. There are two approaches to cluster analysis to choose from, depending on whether you want to link samples or variables:

- R-mode analysis is used to see which variables are more often co-occur.
- Q-mode analysis is used to see which samples are most similar or dissimilar (Parks,1966).

The spectral gamma ray log elements K, Th, and U are the three main variables in this study. The main purpose is to cluster different samples at various depths of various wells according to elemental compositions. Samples are distributed spatially (well to well) and temporally (by depth). Cluster analysis allows us to compare similar samples that are spatially and temporally distributed. For that reason, Q-mode analysis is chosen as the parameter of interest so that we could see variation from one side of the basin to the other and possible shifts through time.

If data were available with more variables, such as geochemistry data for 40 elements from core analysis, it could be useful to use R-mode analysis in order to see which variables more frequently co-occur and perhaps are dependent upon one another.

4.1 Clustering Methods

A multitude of different clustering methods exist. The purpose is to group observations into clusters. If each observation is allocated into only one cluster, this is called partitioning. Partitioning will result in a pre-defined number of clusters used by an investigator. On the other hand, it is also possible to construct a hierarchy of partitions which is called hierarchical clustering. Hierarchical methods and partitioning methods are the most common clustering methods and will be discussed in detail below.

4.1.1 Hierarchical Methods

Hierarchical methods start by treating n objects to be clustered as n clusters individually; in other words, each object is initially considered as a cluster. The (n-1)/2 distances or similarities among the n objects are compared, and by grouping the two nearest or similar objects together new clusters are constructed. This procedure is repeated until all n objects are in one cluster (Pirkle et al., 1984). Distance in cluster analysis has nothing to do with geographic distance between two observations. It is a measure of similarity or dissimilarity between observations in the multivariate space defined by entered variables (Templ et al., 2008). An agglomerative algorithm starts with each observation having its own class. Then, at each step of the algorithm, most similar classes are combined. At the end of the process there is only one cluster which contains all observations. A distance or similarity matrix is an input to most of the hierarchical clustering algorithms. Computational modules such as distance matrix will be discussed later in this chapter. Cluster results for hierarchical methods displayed are as graphics called dendograms (Templ et al., 2008). As there are three main paleo-river systems for the sediment provenance in the South Caspian Basin, we were looking for three main clusters in each analysis.

4.1.2 Partitioning Methods

In contrast to hierarchical clustering methods, partitioning methods start with an initial division of the data pre-defined by the user. After that, by using various iterative schemes, a grouping is determined that optimizes a measure reflecting the homogeneity of the clusters (Pirkle et al., 1984). K-means is the most popular partitioning algorithm. It

aims to minimize the average squared distance between the observations and their cluster centroids.

Other clustering methods include model-based and fuzzy-clustering methods (Templ et al., 2008).

Model-based clustering is a method based on defining models that describe the shape of clusters and it is not based on distances between the observations. The density of a multivariate normal distribution with a certain mean and covariance is described for each of the individual clusters. The cluster shape is determined by the choice of covariance matrix. Expectation- Maximization (EM) algorithm is used for the estimation of these types of cluster models.

In fuzzy-clustering methods, the observations are distributed in a certain degree among all clusters; they are not clearly or discreetly allocated. Membership coefficients to all clusters is determined for each observation. This will provide information on how strong each individual observation is associated with each cluster. Resulting number of clusters should be pre-defined based on purpose of user just as for partitioning methods (Templ et al., 2008).

In this study, a hierarchical method is used and results are displayed as dendograms.

4.2 Data Preparation and Analysis

Thirteen wells with spectral gamma ray readings for the Balakhany VIII sub-suite are the focus of this study. The Balakhany VIII samples are clustered with Q-mode analysis for 13 wells and their associated depths.

Although cluster analysis of a small data set is relatively simple, it becomes more difficult when the number of samples increases, as the calculation process is more complex. In addition, graphic routines for construction of dendograms becomes complex (Davis, 1973). For that reason, one sample was chosen from each 10 meters in all well logs. Selection was based on a gamma ray log reading which can be used as a lithology indicator. Gamma ray value less than 60 API can be thought of as sandstones in a potential reservoir rock (Rider, 1996). Usually, the lowest gamma ray value is chosen for each 10m interval. If no gamma ray value less than 60 API is present, a random sample is selected and considered to be shales in the reservoirs. Raw data sets with sample numbers, depths, gamma ray values, and the three variables (K, Th and U) are presented in Table 4.1 for all 13 wells used for the Balakhany VIII sub-suite evaluation.

NTSYS 2.2 pc Numerical Taxonomy and Multivariate Analysis System program is used for the cluster analysis (Rohlf, 2005). There are many statistical computational modules included in NTSYS 2.2 pc.

Sample #	Depths between	Actual Depth(m)	Gamma Ray (API)	K(%)	Th(ppm)	U(ppm)
s1	2560m-2570m	2560.47	73.26	1.92	7.13	0.97
s2	2570m-2580m	2572.05	71.84	1.83	7.60	1.19
s3	2580m-2590m	2583.03	30.94	1.00	2.17	1.06
s4	2590m-2600m	2591.26	32.00	1.20	3.25	0.84
s5	2600m-2610m	2608.02	72.85	1.81	6.59	2.01
s6	2610m-2620m	2612.29	69.24	2.15	7.64	1.02
s7	2620m-2630m	2628.29	70.78	2.05	9.77	0.79
s8	2630m-2640m	2633.93	68.70	1.88	7.65	1.29
s9	2640m-2650m	2645.51	30.27	1.12	5.27	0.37
s10	2650m-2660m	2656.18	26.15	0.97	2.40	0.71
s11	2660m-2670m	2663.19	33.76	1.25	4.73	0.56
s12	2670m-2680m	2679.04	25.17	0.75	3.11	0.62
s13	2680m-2690m	2689.56	68.84	1.97	5.88	1.65
s14	2690m-2700m	2695.65	71.12	1.99	7.27	1.12
s15	2700m-2710m	2701.59	44.12	1.34	5.88	0.84
s16	2710m-2720m	2719.27	49.88	1.48	4.65	1.37
s17	2720m-2730m	2728.57	39.19	1.38	6.34	0.47
s18	2730m-2740m	2737.10	55.61	1.90	8.03	0.35
s19	2740m-2750m	2745.49	67.68	2.12	8.56	0.42
s20	2750m-2760m	2756.31	71.46	1.91	7.12	1.17
s21	2760m-2770m	2761.49	74.73	1.60	7.09	1.29
s22	2770m-2780m	2777.03	51.14	1.42	8.15	0.86
s23	2780m-2790m	2788.46	29.91	0.99	3.05	1.18
s24	2790m-2800m	2799.13	32.87	1.06	3.98	1.05
s25	2800m-2810m	2802.64	25.67	0.96	3.48	0.71
s26	2810m-2820m	2816.35	28.72	1.16	2.35	0.78
s27	2820m-2830m	2828.70	51.70	1.33	7.97	1.24
s28	2830m-2840m	2832.20	43.57	1.43	3.84	1.57
s29	2840m-2850m	2847.14	29.72	1.09	3.63	0.61
s30	2850m-2860m	2855.21	24.94	0.56	2.54	1.21
s31	2860m-2870m	2869.23	32.55	1.03	2.43	1.38
s32	2870m-2880m	2878.68	21.99	0.79	2.65	0.68
s33	2880m-2890m	2887.52	21.36	0.83	2.43	0.84
s34	2890m-2900m	2896.36	23.91	1.00	1.84	0.88
s35	2900m-2910m	2903.22	22.10	0.57	2.46	1.14
s36	2910m-2920m	2911.30	28.39	0.75	3.48	1.08
s37	2920m-2930m	2923.18	21.08	0.71	2.01	0.90
s38	2930m-2940m	2931.57	66.44	2.20	7.84	1.11
s39	2940m-2950m	2944.98	23.22	0.76	2.95	0.82
s40	2950m-2960m	2951.84	75.26	2.00	6.75	2.71

Table 4.1: Raw data for well E01Z.

4.2.1 Standardization

Closed-data sets can be explained as if the individual variables are not independent of each other but are related. For example, if the unit is expressed as a percentage, ppm, or mg/kg, they sum up to a constant like 100% or 1. Geochemical data are closed data which are expressed in units like wt% or ppm. Major elements are measured in %, whereas trace elements are measured in ppm. Transferring elements to one common unit is not a solution, as major elements occur in much greater amounts than trace elements (Templ et al., 2008). Therefore, multivariate statistical methods may deliver biased results with these types of data. Because of that reason, appropriate data transformation and standardization have to be considered prior to performing cluster analysis (Aitchison, 1986). Structure of the natural clusters could be distorted in the direction of major elements if the variables are not normalized (Pirkle et al., 1984). The entire raw data matrix should be normalized column by column in order to give equal weight to each of the variables.

The computational module STAND was used in NTSYS to standardize all variables. Standardization (STAND) performs a variety of linear transformations on the variables in a data matrix. It is used to reduce the effects of different scales of measurements in different units. In our case Th and U values are given in ppm, whereas K is in percentage. The most universal method, the z-transformation, is used. In z-transformation the raw data are subtracted with the mean and then divided by the standard deviation of the data. Basically, $X_{std} = \frac{X - mean}{Std\ deviation}$

The first step for all the cluster analysis in this study was standardization of the variables (Table 4.2).

	K	TH	U
1	-0.4284	1.1428	-0.7144
2	-0.4840	1.1499	-0.6659
3	-0.6203	1.1536	-0.5333
4	-0.4319	1.1434	-0.7115
5	-0.6135	1.1539	-0.5405
6	-0.4098	1.1398	-0.7300
7	-0.4425	1.1449	-0.7024
8	-0.4912	1.1506	-0.6595
9	-0.4302	1.1431	-0.7129
10	-0.4285	1.1429	-0.7143
11	-0.4167	1.1410	-0.7243
12	-0.5302	1.1535	-0.6232
13	-0.5079	1.1520	-0.6442
14	-0.4419	1.1448	-0.7029
15	-0.4844	1.1500	-0.6655
16	-0.5469	1.1542	-0.6072
17	-0.4272	1.1426	-0.7155
18	-0.3762	1.1335	-0.7573
19	-0.3679	1.1318	-0.7640
20	-0.4588	1.1471	-0.6883
21	-0.5283	1.1533	-0.6250
22	-0.5074	1.1520	-0.6446
23	-0.6600	1.1505	-0.4906
24	-0.5728	1.1547	-0.5818
25	-0.4946	1.1509	-0.6564
26	-0.3279	1.1228	-0.7949
27	-0.5650	1.1546	-0.5896
28	-0.6296	1.1531	-0.5235
29	-0.4228	1.1420	-0.7191
30	-0.8684	1.0933	-0.2250
31	-0.8016	1.1206	-0.3190
32	-0.5264	1.1532	-0.6268
33	-0.5830	1.1547	-0.5717
34	-0.4553	1.1466	-0.6913
35	-0.8469	1.1032	-0.2563
36	-0.6842	1.1477	-0.4635
37	-0.7084	1.1439	-0.4355
38	-0.4199	1.1415	-0.7216
39	-0.6035	1.1543	-0.5508
40	-0.7115	1.1434	-0.4319

Table 4.2: Standardized data matrix for well E01Z.

4.2.2 Similarity for Interval Data

Most clustering methods use a measure of similarity between the observations in order to determine the group membership. To express the similarity, generally distances between the observations in the data space are used. As mentioned before, distance in cluster analysis has nothing to do with geographic distance between two observations. It is a measure of similarity or dissimilarity between observations in the multivariate space defined by entered variables (Templ et al., 2008). Small distances indicate that the two objects are similar, whereas large distance indicate dissimilarity. There are many different distance measures.

The computational module SIMINT is used in NTSYS for this study. The SIMINT module computes a variety of similarity and dissimilarity coefficients for interval measured data. The input is a standardized rectangular data matrix and the output is a symmetric similarity or dissimilarity matrix, depending on the coefficient employed (Table 4.3). Different coefficients were tried and it was decided to use the default DIST measure for the cluster analysis. The different coefficients that were tried are as follows:

DIST: Average taxonomic distance (Default in NTSYS)

$$E_{ij} = \sqrt{\frac{1}{n}\sum_k (x_{ki} - x_{kj})^2}$$

EUCLID: Euclidian distance $E_{ij} = \sqrt{\sum_k (x_{ki} - x_{kj})^2}$

MANHAT: Average Manhattan distance $M_{ij} = \frac{1}{n}\sum_k \left|x_{ki} - x_{kj}\right|$

CORR: Pearson product-moment correlation coefficient (a commonly used coefficient)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0000														
2	0.0428	0.0000													
3	0.1525	0.1098	0.0000												
4	0.0026	0.0402	0.1498	0.0000											
5	0.1468	0.1041	0.0057	0.1442	0.0000										
6	0.0141	0.0569	0.1665	0.0167	0.1608	0.0000									
7	0.0107	0.0321	0.1418	0.0081	0.1361	0.0248	0.0000								
8	0.0484	0.0055	0.1042	0.0457	0.0985	0.0625	0.0376	0.0000							
9	0.0014	0.0415	0.1511	0.0013	0.1454	0.0155	0.0094	0.0470	0.0000						
10	0.0001	0.0427	0.1524	0.0025	0.1467	0.0142	0.0106	0.0483	0.0013	0.0000					
11	0.0089	0.0518	0.1613	0.0116	0.1557	0.0052	0.0197	0.0573	0.0103	0.0090	0.0000				
12	0.0791	0.0363	0.0735	0.0765	0.0678	0.0932	0.0684	0.0308	0.0778	0.0790	0.0881	0.0000			
13	0.0615	0.0186	0.0911	0.0588	0.0855	0.0756	0.0507	0.0131	0.0601	0.0614	0.0704	0.0177	0.0000		
14	0.0103	0.0326	0.1422	0.0077	0.1365	0.0244	0.0004	0.0381	0.0089	0.0102	0.0192	0.0689	0.0512	0.0000	
15	0.0431	0.0003	0.1095	0.0405	0.1038	0.0572	0.0324	0.0053	0.0418	0.0430	0.0520	0.0360	0.0184	0.0328	0.0000

Table 4.3: Distance matrix showing dissimilarity coefficients for first 15 samples of well E01Z out of 40 samples.

4.2.3 Sequential Agglomerative Hierarchical Nested Cluster Analysis (SAHN)

The SAHN computational module was used in NTSYS in order to complete the cluster analysis. Output matrix from the SIMINT module is used as an input into this module. There are different approaches for the SAHN clustering methods. Both approaches have been tried and default UPGMA method was ultimately chosen. Attempted SAHN strategies were:

UPGMA: Unweighted pair-group method with arithmetic averages

WPGMA: Weighted pair-group method with arithmetic averages

The SAHN module is the standard algorithm for agglomerative clustering and operates as follows:

1. The first steps are choosing a similarity coefficient or a distance function, then Q or R mode, at which point a similarity or distance coefficients is calculated , then search the input matrix for the pair of objects (i, j) that are most similar (or least dissimilar). As previously mentioned, I used Q-mode and compared samples. In the case of this study the objects are samples.

2. Merge these samples into a new cluster.

3. Update the matrix to reflect the deletion of the pair of samples, i and j, that were merged and the addition of a new "sample" corresponding to the new cluster. Similarities or dissimilarities have to be computed between the existing objects and the new cluster (the different SAHN methods differ only in the formulas used at this step).

4. Go back to step 1, above, if the size of the new matrix is greater than 2 x 2 - otherwise stop. Note that 2 objects are deleted and one is added at each step so this algorithm must terminate when there are no additional samples or clusters to combine..

By using the linkage matrix obtained from cluster analysis (Table 4.4), a dendogram showing the clustering results will be obtained (Figure 4.1).

	Sample	Linkage
1	1	0.0001
2	10	0.0010
3	17	0.0023
4	4	0.0013
5	9	0.0072
6	11	0.0036
7	29	0.0022
8	38	0.0120
9	6	0.0196
10	7	0.0004
11	14	0.0114
12	20	0.0026
13	34	0.0560
14	18	0.0062
15	19	0.0328
16	26	0.1004
17	2	0.0003
18	15	0.0067
19	8	0.0027
20	25	0.0149
21	13	0.0004
22	22	0.0301
23	12	0.0023
24	21	0.0015
25	32	0.0149
26	16	0.0711
27	3	0.0057
28	5	0.0107
29	28	0.0147
30	39	0.0356
31	24	0.0064
32	27	0.0115
33	33	0.2288
34	23	0.0210
35	36	0.0334
36	37	0.0027
37	40	0.1404
38	30	0.0227
39	35	0.0571
40	31	0.0000

Table 4.4: Linkage matrix obtained at the end of SAHN module. This matrix shows paired-linkages between the samples and it is used to construct a dendogram.

In summary, there are choices to make in the overall process of cluster analysis. According to Meyer (2011), the steps are as follows:

1. *Parameter of Interest:* R-mode (compare variables) and Q-mode (compare samples as a whole) are the two parameters of interest.

   Q-mode is used in this study as the aim is to compare spectral gamma ray elements values relative to spatial and lateral position of sample locations.
2. *Clustering Method:* Hierarchical and K-means are the two main methods.

   The hierarchical method was used in this study.
3. *Input Data:* Is the raw data adequate or should the data be normalized first?

   Input data is normalized as the variables have very different units.
4. *Similarity/Distance Measure:* Correlation and distance are the two main measures.

   A distance measure is used in order to compute similarity of interval data.
5. *Progression Method:* Agglomerative or divisive are the two progression methods for cluster linkage.

   The agglomerative method is used in this study. (Hierarchical method)
6. *Linkage:* Single, complete, average and centroid are the different types of linkage.

   Linkage controls how the distance measurements are calculated and therefore how clusters are merged in an agglomerative analysis. In UPGMA in NTSYS averages are used.
7. *Attributes:* Monothetic (one variable), polythetic (more than one variable), and omnithetic (all variables available) are the number of variables.

   Attributes are polythetic as there are more than one (three) variables.

CHAPTER V: RESULTS

In order to understand the provenance of reservoir sediments which are different paleo-river systems, a cluster analysis method as discussed above is used. Different analyses have been accomplished. Many strategies were tried in search of temporal and spatial relationships of SGR data between and within wells. The most important ones will be discussed below. Through this process we are looking at different ways to discern geographic or spatial patterns that could help us visualize the spectral gamma ray compositions as affected by three or more sediment provinces.

This process took several trial runs to evaluate the data set in terms of different sample relationships within a well for stratigraphic or temporal relationships. Additional trials were run to evaluate well-to-well samples to determine spatial relationships between samples and establish petrofacies laterally and vertically. This was a key component to establishing sediment sources from different directions starting with east-west end member wells and one centrally located well.

Finally, cluster analysis trials were made to evaluate sandstone-shale relationships to petrofacies. Then, porosity and permeability of different petrofacies were compared. All trials eventually conditioned the relationships between samples in three dimensions and time, while at the same time validating the consistency of the petrofacies established by the clustering routines.

5.1 Analysis #1: Clustering Three Wells Separately

Firstly, two end-member wells (E01Z at the northwestern part and GCA6 at the southeastern part) (figure 5.1) and one well in between these wells, GCA1 are clustered separately. Azeri-Chirag-Gunashli field is located in between northern sediment supplier paleo-Volga and eastern sediment supplier paleo-Amu-Darya river systems. The western sediment supplier, the paleo-Kura, might also be affecting the field (Figure 5.1). End-member wells are picked in order to see compositional differences through time or depth in each well. Dendograms for these three wells are obtained; the vertical axes show the sample numbers and horizontal axes shows the dissimilarity between clusters. These dendograms are colored to see the relation between different samples at various depths. Samples having same color belong to same cluster (Figure 5.2, Figure 5.3 and Figure 5.4). Samples were chosen at every ten meters for Balakhany VIII sub-suite in each well. Potassium (%), thorium (ppm), and uranium (ppm) were the three variables.

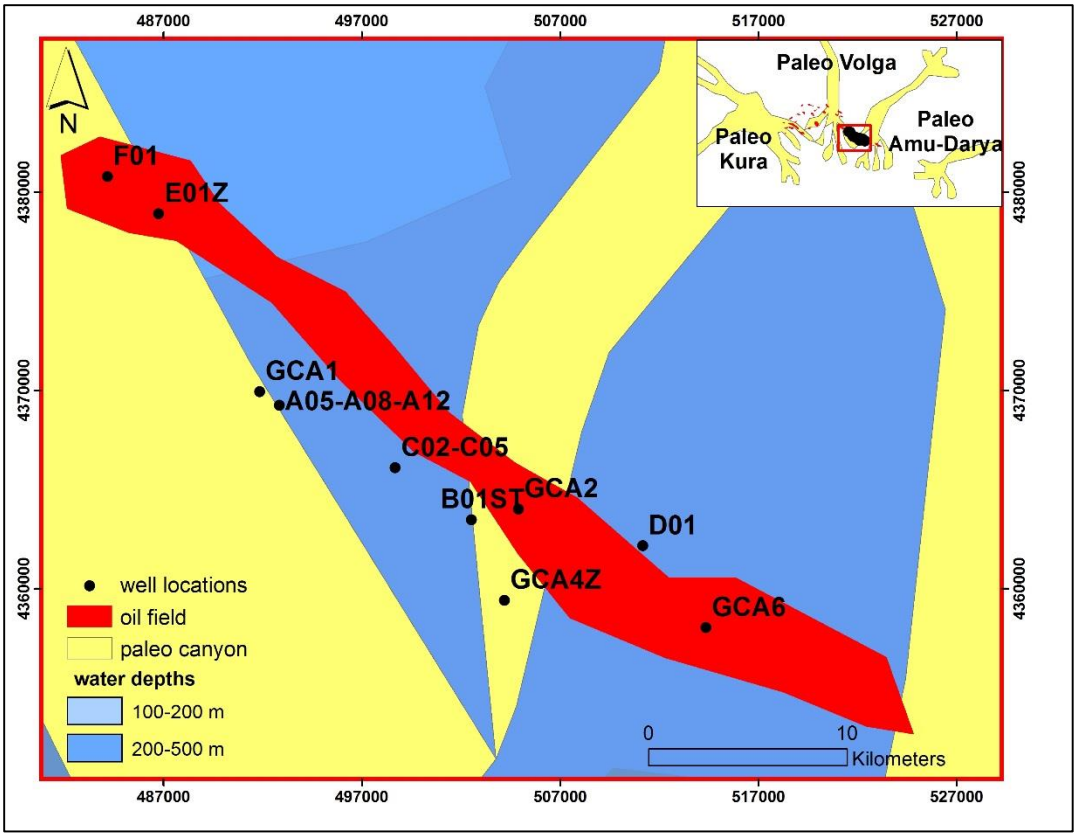


Figure 5.1: ACG field, paleo canyons, and well locations.

Three different clusters are obtained for each of the three wells. There was a characteristic outlier sample in well GCA06. This outlier is eliminated for the future analysis. For each well, a red and blue cluster grouped together to link to yellow clusters. This indicates red and blue clusters are more related to each other and are major clusters, whereas the yellow cluster can be thought of as a less related cluster. The dissimilarity coefficients between the samples are higher for the yellow cluster, which indicates less similarity between these samples. Clusters obtained from this analysis are used to see the temporal relationship of all samples in each well individually. They will then be used in analysis 2 to see the relationship between the three wells.

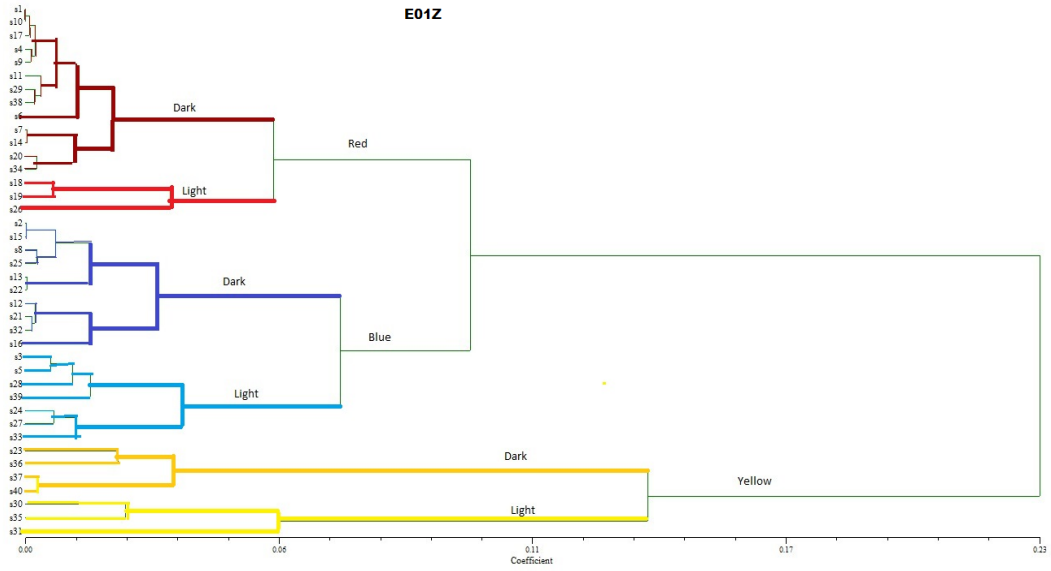


Figure 5.2: Colored dendogram for northwestern end member well E01Z.

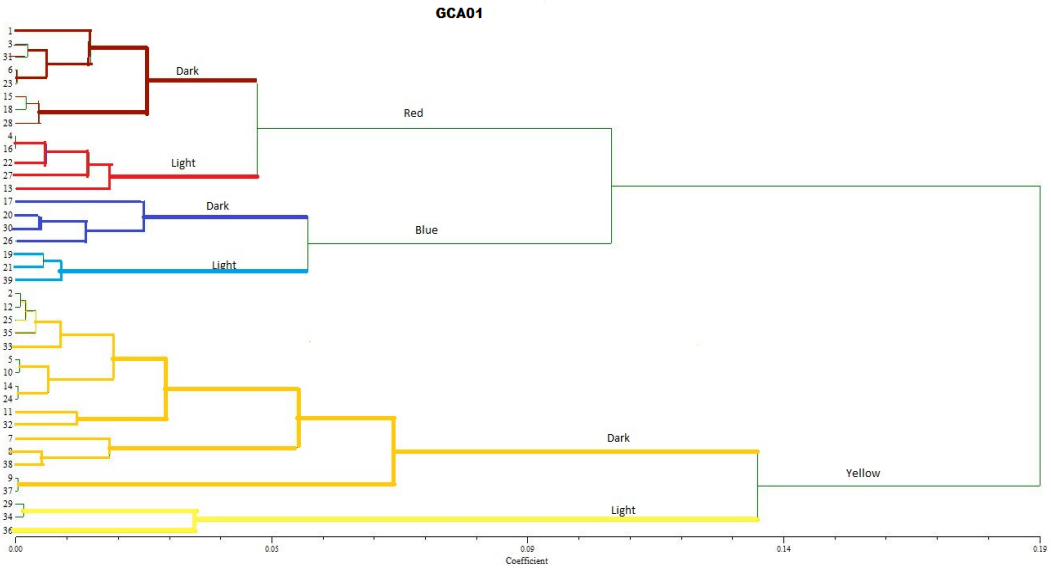


Figure 5.3: Colored dendogram for well GCA01.

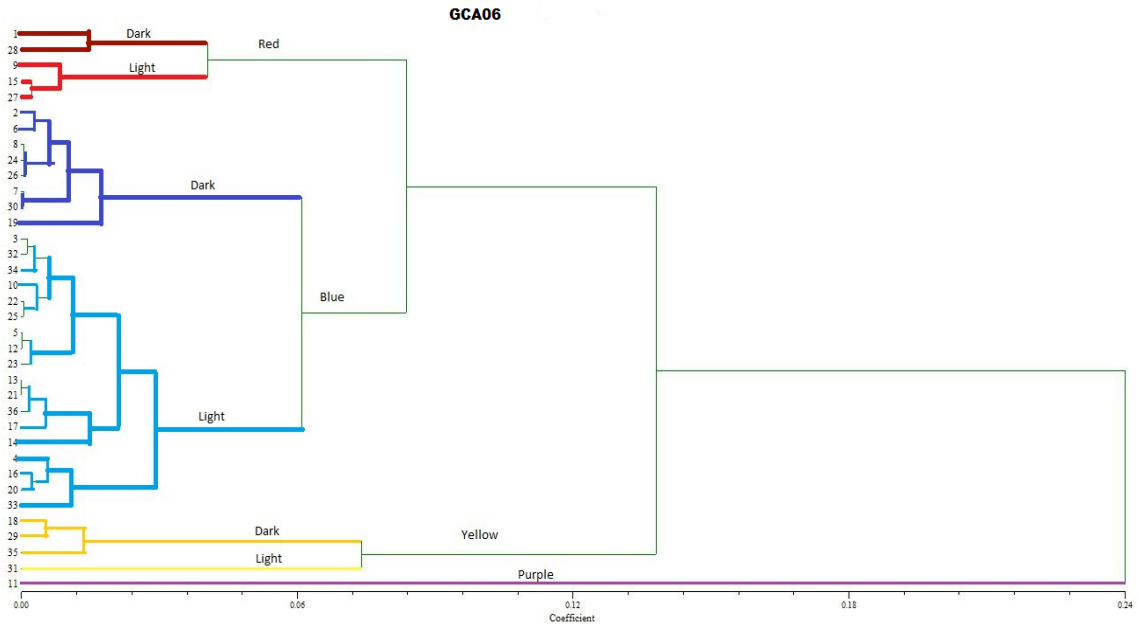


Figure 5.4: Colored dendogram for southeastern end member well GCA06.

5.2 Analysis #2: Clustering Three Wells Together by Using Averages

Six different clusters are obtained for each well from the first analysis (dark red, light red, dark blue, light blue, dark yellow, and light yellow). In total, there were 18 clusters from 3 wells. By taking the averages of variables from each individual cluster for three wells, 18 different samples are obtained. In order to understand which clusters are related to each other and to see their distribution within the wells, cluster analysis has been applied to these 18 samples. By doing this analysis, the relationship between all samples from three wells are interpreted. Results of the analysis are shown on well logs in Figure 5.5. Different colors for clusters designate the different petrofacies with each color showing similar petrophysical properties within themselves. The dark-green and light-green cluster petrofacies 1&2 were the two major clusters, whereas the pink and brown cluster

petrofacies were outliers with a lesser degree of similarity. The E01Z well is characterized by a light-green cluster, petrofacies 2 and a dark-green cluster petrofacies 1. GCA06 is characterized by a dark-green cluster petrofacies 3. The major cluster for the GCA01 is petrofacies 3, and it is followed by a petrofacies 1. Petrofacies 3 is the cluster with the higher porosity and permeability and has the lowest K averages, which indicates more weathering and longer transportation distance for samples in GCA01 relative to the other two wells (Table 5.1). It can be concluded that the contribution of the petrofacies 1 is increasing as we move to the southeastern side of the SCB. This can be interpreted as the effect of sediment sourced from the east by the paleo-Amu-Darya. Each of these wells is characterized by a different cluster. This analysis strongly suggests deposition from different river systems. The results of this cluster analysis were promising, so we were encouraged to try different parameters and to add more wells to the analysis.

	GR (API)	K(%)	Th(ppm)	U(ppm)	Permeability (md)	Porosity (%)	Th/K	Th/U
Petrofacies 1	68.87	1.20	7.66	1.93	68.60	0.10	7.37	5.07
Petrofacies 2	63.63	1.64	6.79	1.72	93.56	0.13	4.10	4.23
Petrofacies 3	38.82	0.46	3.81	1.53	205.88	0.22	6.62	2.45
Petrofacies 4	35.15	0.53	2.78	1.60	110.88	0.22	5.81	1.72
Petrofacies 5	26.53	0.72	2.47	1.24	267.74	0.23	3.75	2.01
Petrofacies 6	53.86	1.64	5.24	1.76	217.37	0.19	3.37	2.86

Table 5.1: Averages of clusters for analysis #2.

5.3 Analysis #3: Clustering All Samples from Three Wells

All the samples from the same three wells (E01Z, GCA01, and GCA06) are clustered together in order to compare the results with the second analysis, which is done by using the averages of individual clusters from each well. Almost the same results were obtained (Figure 5.6). This might indicate the reliability of the cluster analysis method. Averages of the petrofacies are also very similar to these from analysis #2 (Table 5.1 and Table 5.2). Out of 115 samples, only 7 samples belong to different cluster petrofacies from the previous analysis #2 (Figure 5.5 and Figure 5.6).

	GR (API)	K(%)	Th(ppm)	U(ppm)	Permeability (md)	Porosity (%)	Th/K	Th/U
Petrofacies 1	67.74	1.13	7.55	1.93	73.12	0.11	7.71	5.04
Petrofacies 2	64.35	1.63	6.85	1.75	85.02	0.13	4.14	4.19
Petrofacies 3	36.95	0.45	3.55	1.48	226.46	0.23	6.11	2.38
Petrofacies 4	35.15	0.53	2.78	1.60	110.88	0.22	5.81	1.72
Petrofacies 5	26.53	0.72	2.47	1.24	267.74	0.23	3.75	2.01
Petrofacies 6	53.86	1.64	5.24	1.76	217.37	0.19	3.37	2.86

Table 5.2: Averages of clusters for analysis #3.

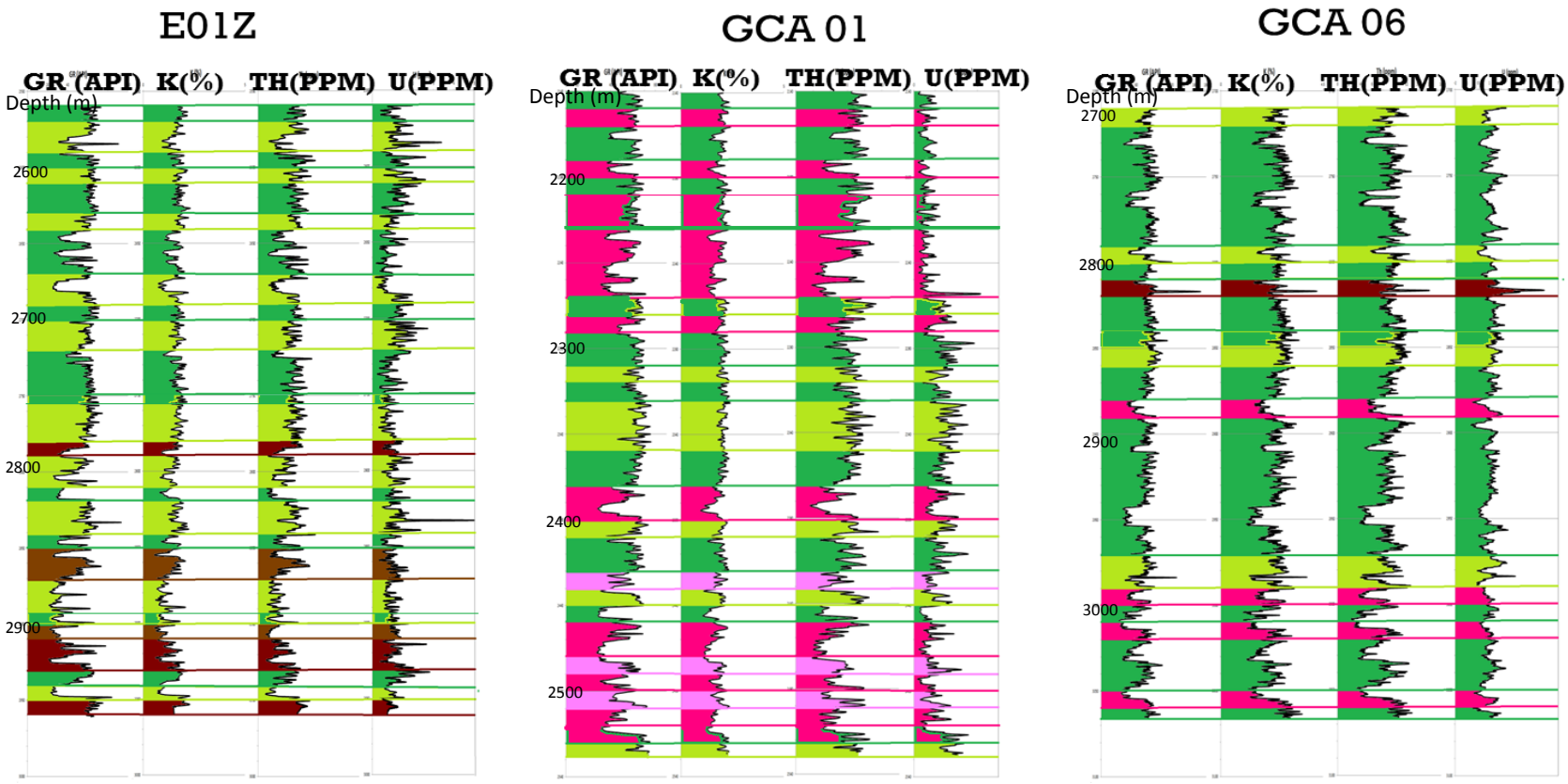


Figure 5.5: Results of cluster analysis obtained by using averages from separate clusters from each well. E01Z is dominated by a light-green cluster petrofacies, GCA01 is by a dark-pink cluster petrofacies and GCA06 is by a dark-green cluster petrofacies. Colors designate different petrofacies as averages shown in Table 5.1.

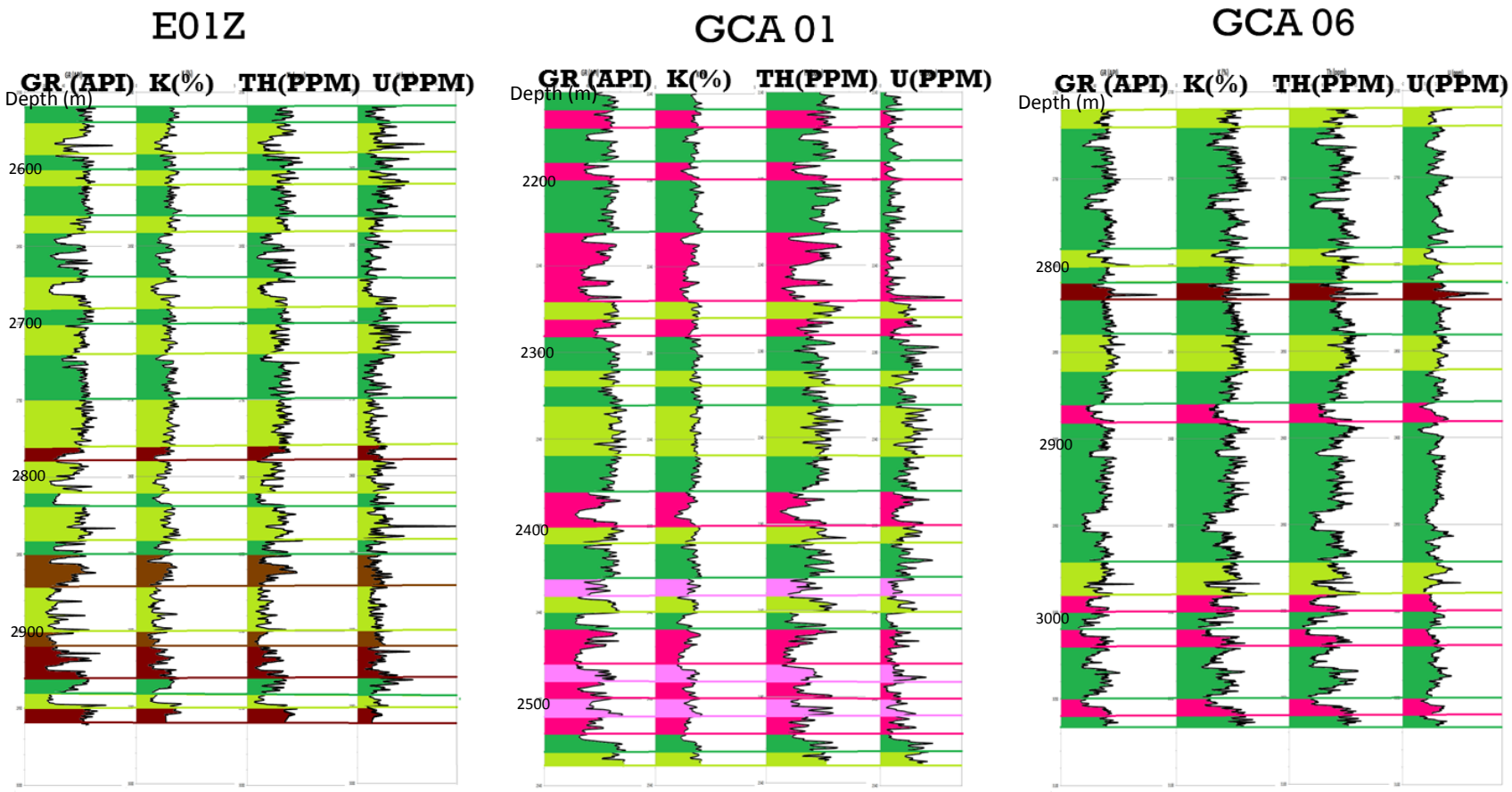


Figure 5.6: Results of cluster analysis obtained by clustering all samples from all three wells together. Almost same results as the previous analysis are obtained. Colors designate different petrofacies as averages shown in Table 5.2.

5.4 Analysis #4: Adding Gamma Ray as a Variable

As a new trial, the same 3 wells were again used; however gamma ray readings were added as a fourth variable into the analysis. Firstly, data from the three wells were clustered separately, but this time with four variables which are K (%), Th (ppm), U (ppm), and GR(API). Three dendograms are obtained and colored for each well like it is done in the analysis #1. After that, using the averages of individual clusters from different wells allowed a new analysis to see the relationship between the three wells, as it was done in analysis #2 (Figure 5.7). Most of the samples were found in the same cluster petrofacies at the end of this analysis, the gamma ray was over affecting the results, even though values are normalized. The reason for that might be the over effect of the total gamma ray reading, as all variables in the analysis essentially add up to the total gamma ray in spite of the use of different units. The formula of GR (API) is as follows:

API= 16*K(%) + 8*U(ppm) + 4*Th(ppm) (Ellis,1987)

Outlier clusters are located around the depths at which the sequence boundaries are identified by previous work by Esedo (2009). These results might be the indication of the power of spectral gamma ray element in identifying the sequence boundaries.

Another attempted analysis was to separate sandstones from other formations by using 60 API gamma ray reading as a cut off. API values lower than 60 are generally thought of as potential reservoir sandstones. However, gamma ray values were again over affecting the analysis even though all the variables were normalized.

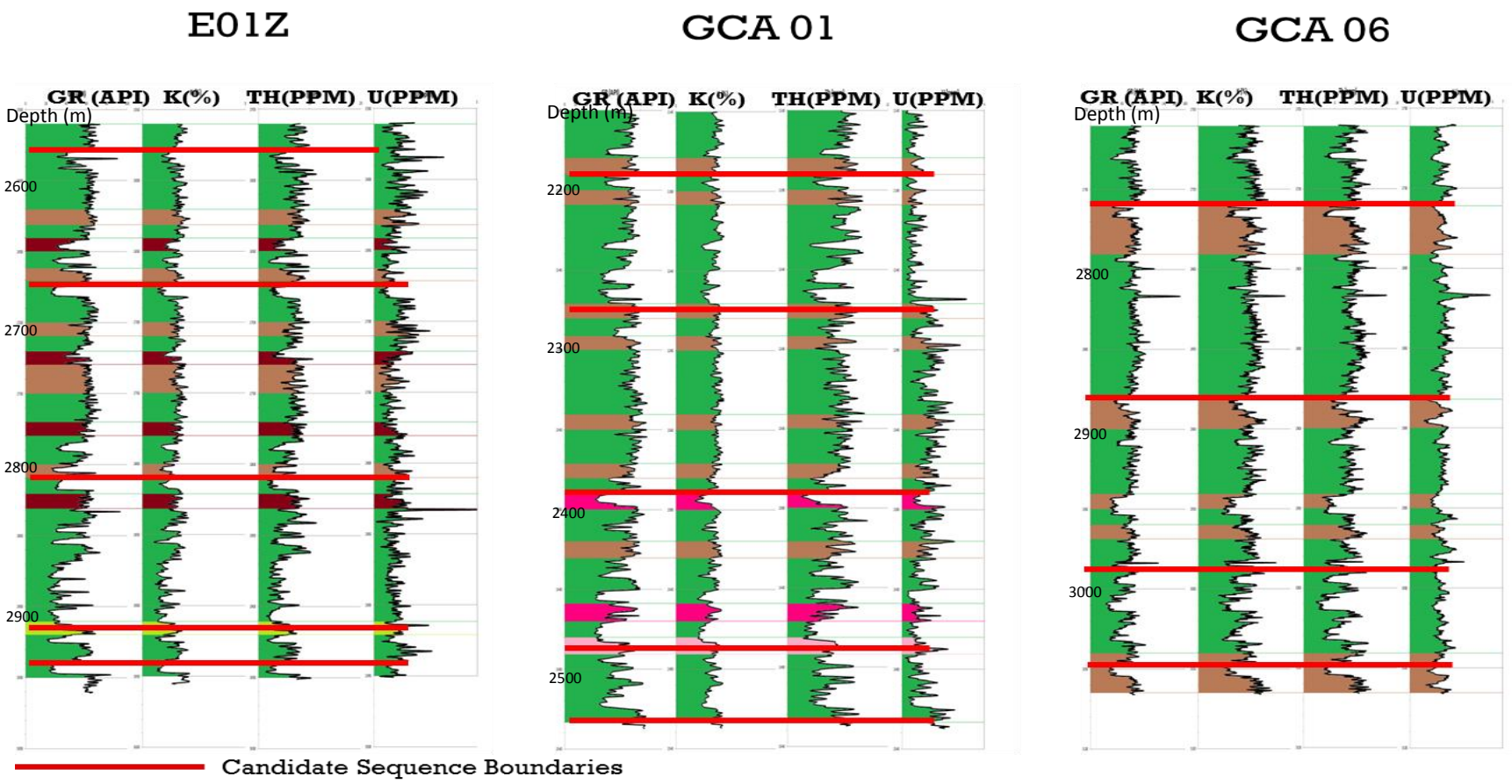


Figure 5.7: Results of cluster analysis by using three variables K(%), Th (ppm), U (ppm) and GR (API). GR was over influencing the results. Colors on well logs designate different petrofacies.

5.5 Analysis #5: Adding Th/K and Th/U Ratios as Variables

Another trial was to use the elemental ratios Th/K and Th/U as variables in the analysis. Elemental ratios might be indicators of transportation distances. However, the results were confusing and many different clusters and layers were present. Using the ratios of the same elements in the analysis was not reasonable and for this reason K(%), Th(ppm) and U(ppm) were decided to be used as the variables for further analysis in this research. This is probably because the ratio were dependent variables.

5.6 Analysis #6: Clustering GR<60 API

The main defining properties of reservoir rocks are obviously porosity and permeability especially of sandstones in conventional resource plays. Another analysis has been done again with three end member wells (E01Z, GCA01, and GCA06) by using three variables but this time for samples having GR<60 API. Northwestern E01Z is sandier than southeastern GCA01. Results of the cluster analysis were giving different cluster petrofacies for each well (Figure 5.8). E01Z was dominated by a dark-red petrofacies, GCA01 by a blue petrofacies, and GCA06 by a dark-blue petrofacies. Red clusters have higher porosity and permeability which indicate better reservoir quality, and are only present in the western part of the study area.

A blue cluster has good reservoir quality and the lowest K averages, which indicates weathering and longer transportation distance for samples in GCA01. This suggests influence could be from the paleo-Volga river system (Table 5.3).

	GR (API)	K(%)	Th(ppm)	U(ppm)	Permeability (md)	Porosity (%)	Th/K	Th/U
Dark-Red Petrofacies	35.14	1.10	4.27	0.97	226.08	0.20	3.82	4.57
Light-Red Petrofacies	26.46	0.82	2.84	1.05	362.27	0.25	3.51	2.67
Dark-Blue Petrofacies	34.33	0.92	4.08	0.93	221.88	0.21	5.82	6.01
Blue Petrofacies	37.35	0.46	3.66	1.49	216.72	0.22	10.31	2.43
Light-Blue Petrofacies	35.15	0.53	2.78	1.60	110.88	0.22	5.81	1.72
Dark-Yellow Petrofacies	23.52	0.56	2.50	1.17	373.56	0.25	4.44	2.13
Light-Yellow Petrofacies	32.55	1.03	2.43	1.38	56.11	0.20	2.36	1.76

Table 5.3: Averages of clusters for analysis #6.

Results of the preliminary cluster analysis with three wells were promising, both for all samples analyses #2 and #3 and also for the sandstones analysis #6. After trials of using different variables and elemental ratios, it was decided to use three elements K(%), Th(ppm), and U(ppm) as the three variables for further analysis.

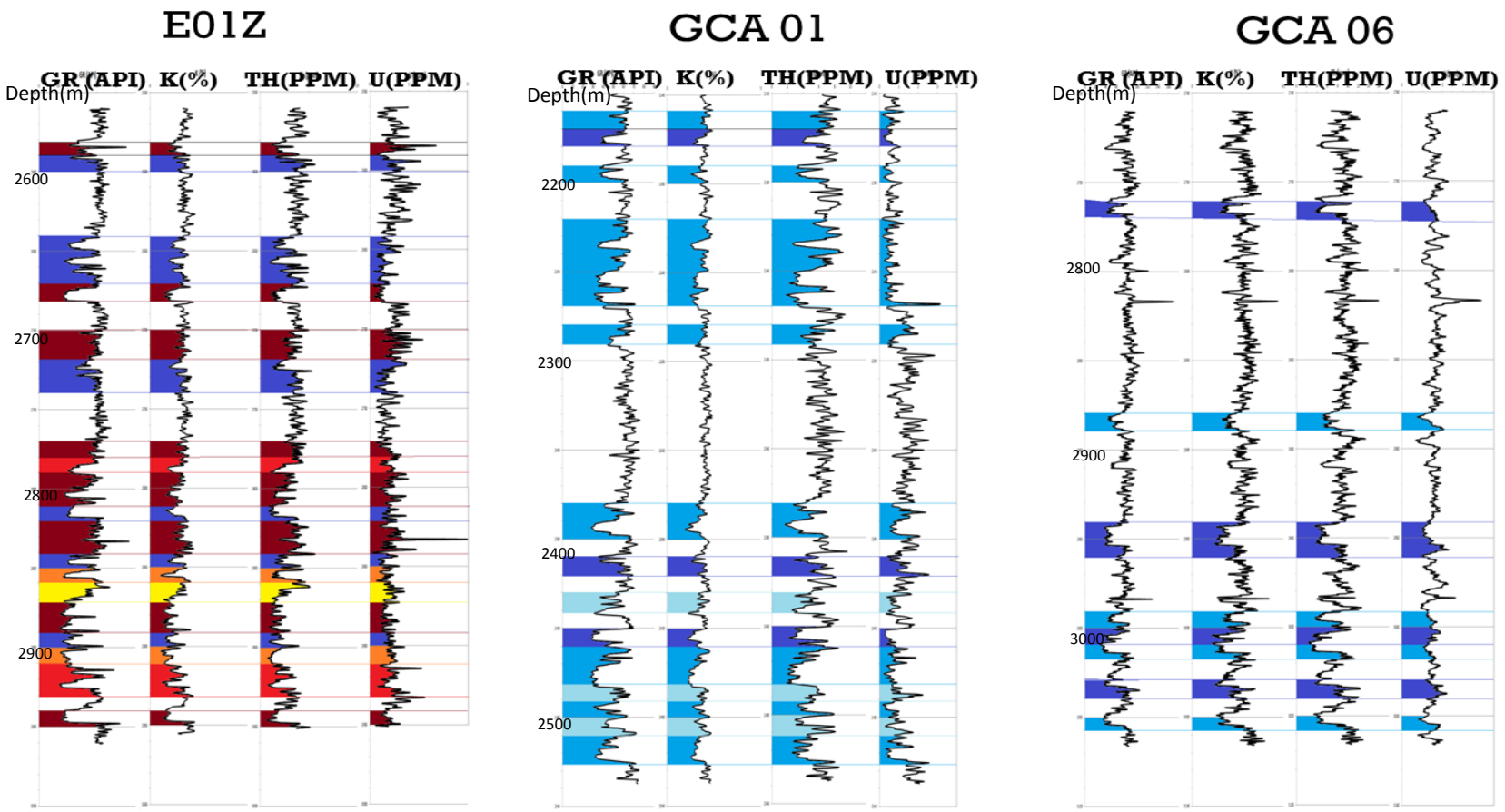


Figure 5.8: Results of cluster analysis for samples having GR<60 API. 3 wells are dominated by different cluster petrofacies. Colors designate different petro facies as averages shown in Table 5.3.

5.7 Analysis #7: Clustering Closely Located Wells

Three of the wells which are A05, A08, and A12 are located very close to each other. Maximum distance between them is 9 meters. In order to see the distribution of the different clusters temporally, samples of these wells are clustered.

All the wells are characterized by red cluster petrofacies, However A05 and A08 are dominantly dark red, whereas A12 is light red (Figure 5.9). Samples in the blue petrofacies have higher GR values that indicate shalier formations. When light and dark-red petrofacies are compared, dark red has higher Th values, indicating deeply weathered sediments from within their provenance (Ellis, 1987).

	GR (API)	K(%)	Th(ppm)	U(ppm)	Permeability (md)	Porosity (%)	Th/K	Th/U
Dark-Red Petrofacies	50.07	1.61	5.94	0.74	19.34	0.14	3.70	9.14
Light-Red Petrofacies	41.52	1.10	3.46	0.99	118.62	0.21	3.19	3.92
Dark-Blue Petrofacies	96.37	1.98	7.99	1.56	0.03	0.03	4.05	5.54
Light-Blue Petrofacies	75.88	2.14	9.14	0.63	0.05	0.04	4.31	17.19
Brown Petrofacies	99.68	1.84	6.59	4.24	0.02		3.58	1.55

Table 5.4: Averages of clusters for analysis #7.

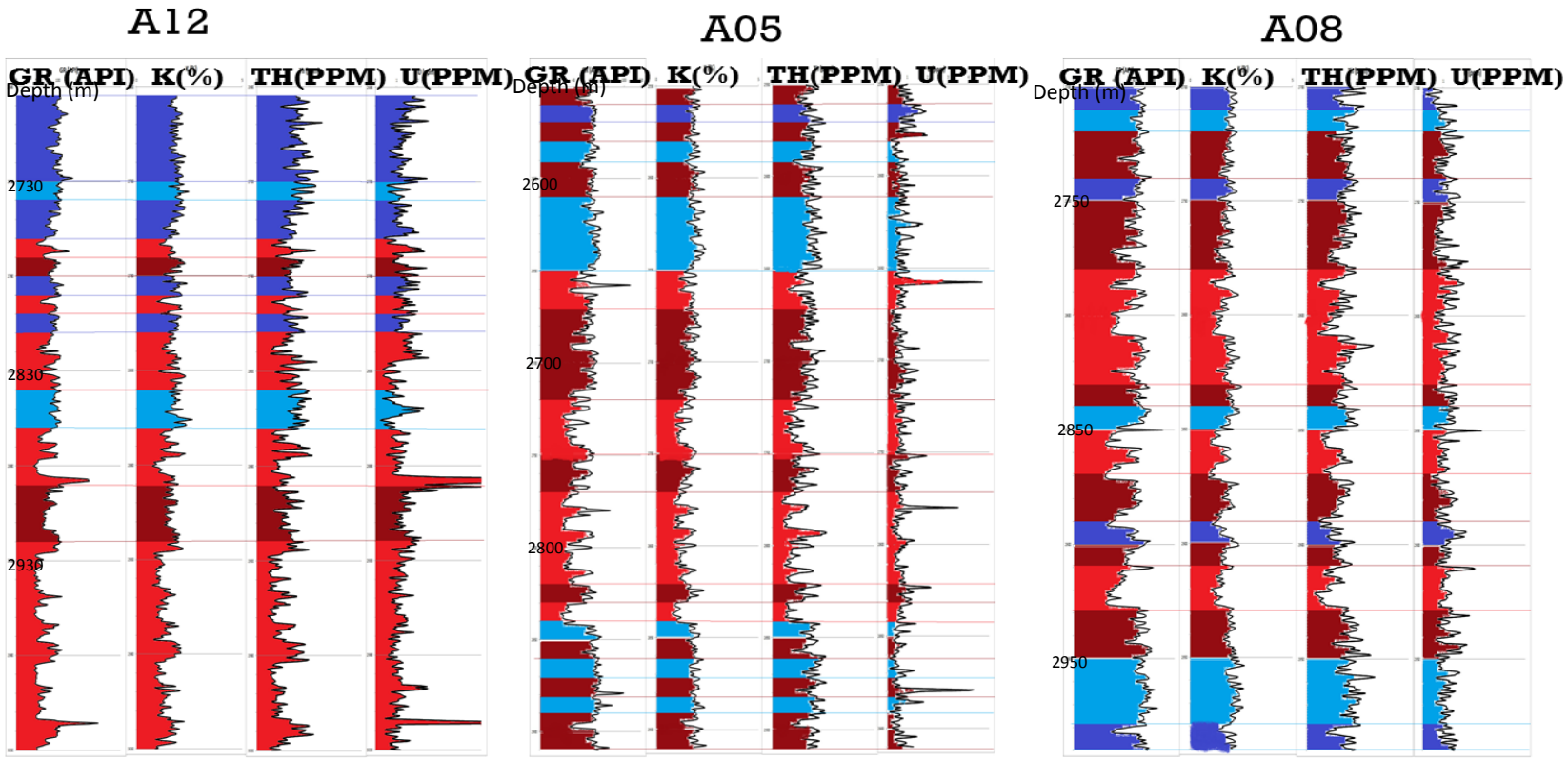


Figure 5.9: Clustering results of closely located wells. Sandier red cluster is dominant one whereas shalier blue one is the minor. Colors designate different petrofacies as averages shown in Table 5.4

There were a total of 13 well logs available to use in the cluster analysis for the Balakhany VIII reservoir unit in this research. To do further analysis, all of the 13 well logs were clustered individually. Individual dendograms were obtained from the NTSYS software. The next step was to do this analysis for all the wells together to see the distribution of the different clusters within the study area. Two different analyses were done for all 13 wells, the first one included all samples and the second one was only for samples having GR less than 60API.

5.8 Analysis #8: Clustering All Samples from All Wells

There were total 421 samples from the 13 wells. Figure 5.10 is the dendogram for that analysis. A green cluster is the first major cluster, whereas a pink cluster is the second one to develop. A yellow cluster is the third cluster with obviously a lesser degree of similarity between its samples.

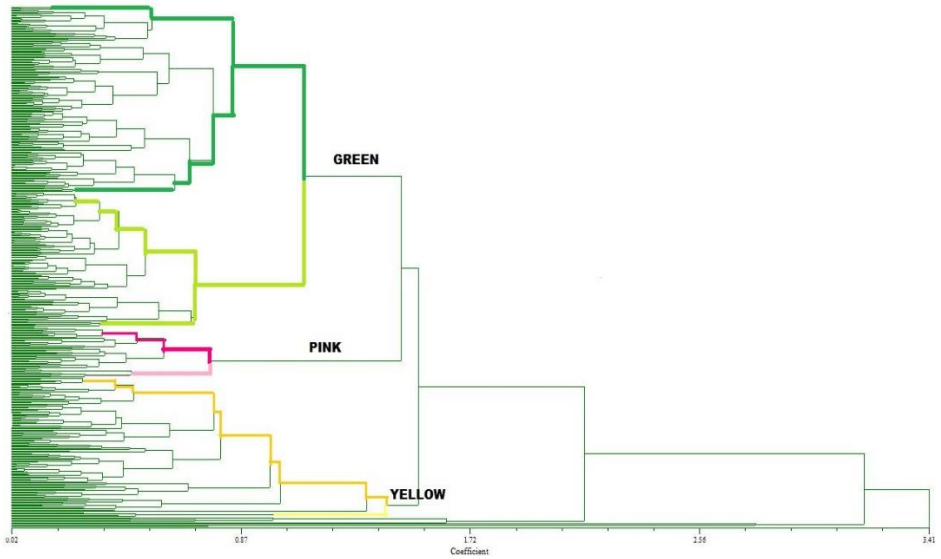


Figure 5.10: Dendogram for 421 samples from 13 wells with three main clusters.

Results are also shown on well logs (Figure 5.11 a,b,c,d and Figure 5.12). The green petrofacies have the better quality reservoirs with higher averages of porosity and permeability. The dark-yellow cluster has better reservoir quality than the pink clusters. Green clusters also have the lowest K averages, which indicate relatively highly weathering and longer transportation distance from the provenance. K average of yellow cluster is lower than the pink cluster. Details of different clusters and their probable provenances will be discussed in the discussion section.

	GR (API)	K (%)	Th(ppm)	U(ppm)	Permeability (md)	Porosity (%)	Th/K	Th/U
Dark-Green Petrofacies B	51.21	1.65	5.46	0.91	158.00	0.17	3.32	5.99
Light-Green Petrofacies A	36.86	0.86	3.23	1.11	293.95	0.21	3.76	2.91
Dark-Pink Petrofacies	71.15	2.19	8.25	0.68	0.25	0.06	3.77	12.11
Light-Pink Petrofacies	83.69	2.58	10.40	0.59	0.02	0.02	4.03	17.76
Dark-Yellow Petrofacies	66.26	1.66	7.83	2.55	18.49	0.12	4.72	3.07
Light-Yellow Petrofacies	110.67	2.11	12.64	3.40	0.02	0.01	6.00	3.72
Outliers	77.81	1.54	10.65	3.40	0.58	0.10	6.90	3.13

Table 5.5: Averages of clusters for analysis #8.

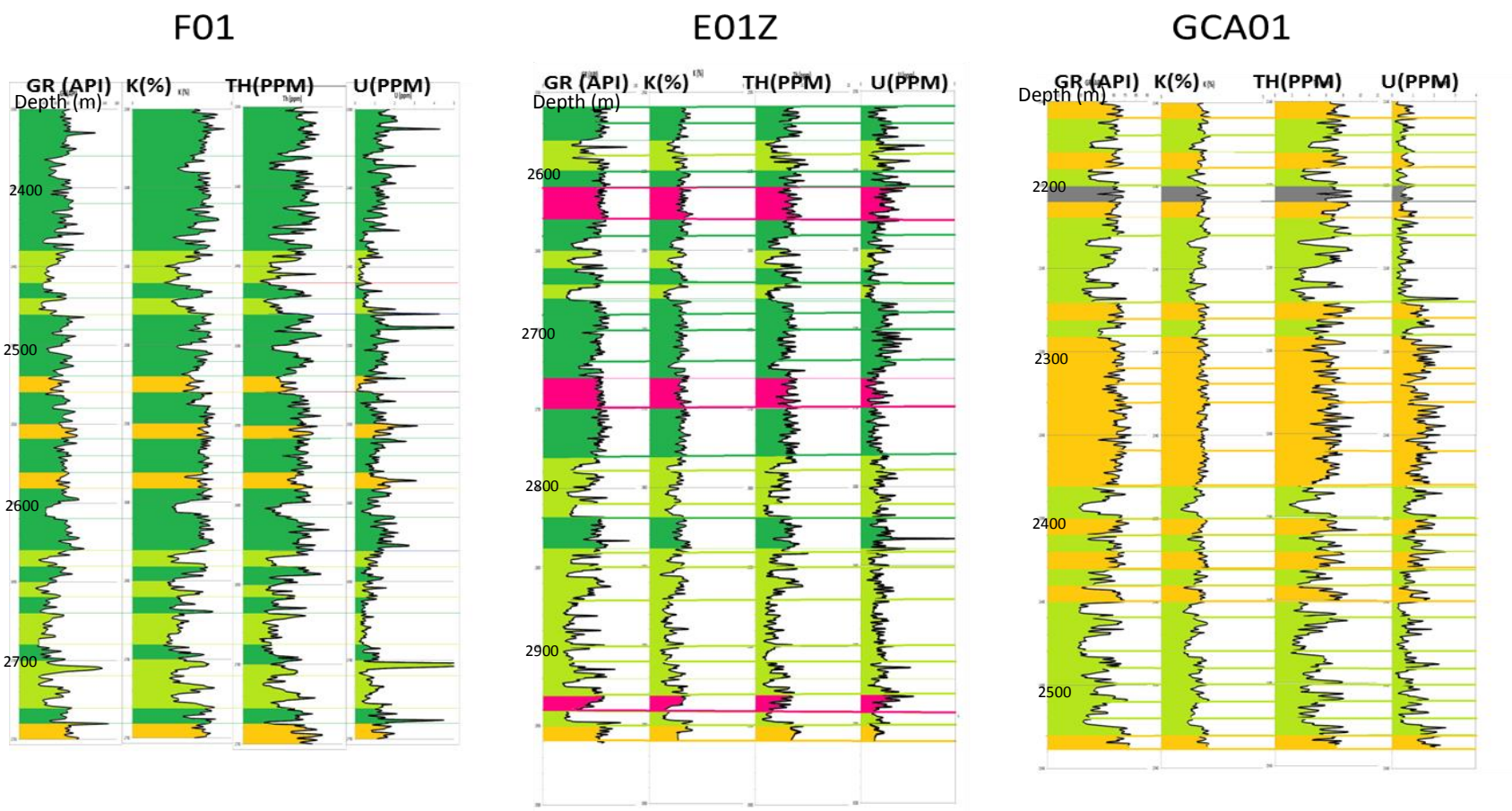


Figure 5.11 a: Cluster analysis results for all samples showing GR and SGR logs. Colors designate different petrofacies as averages shown in Table 5.5.

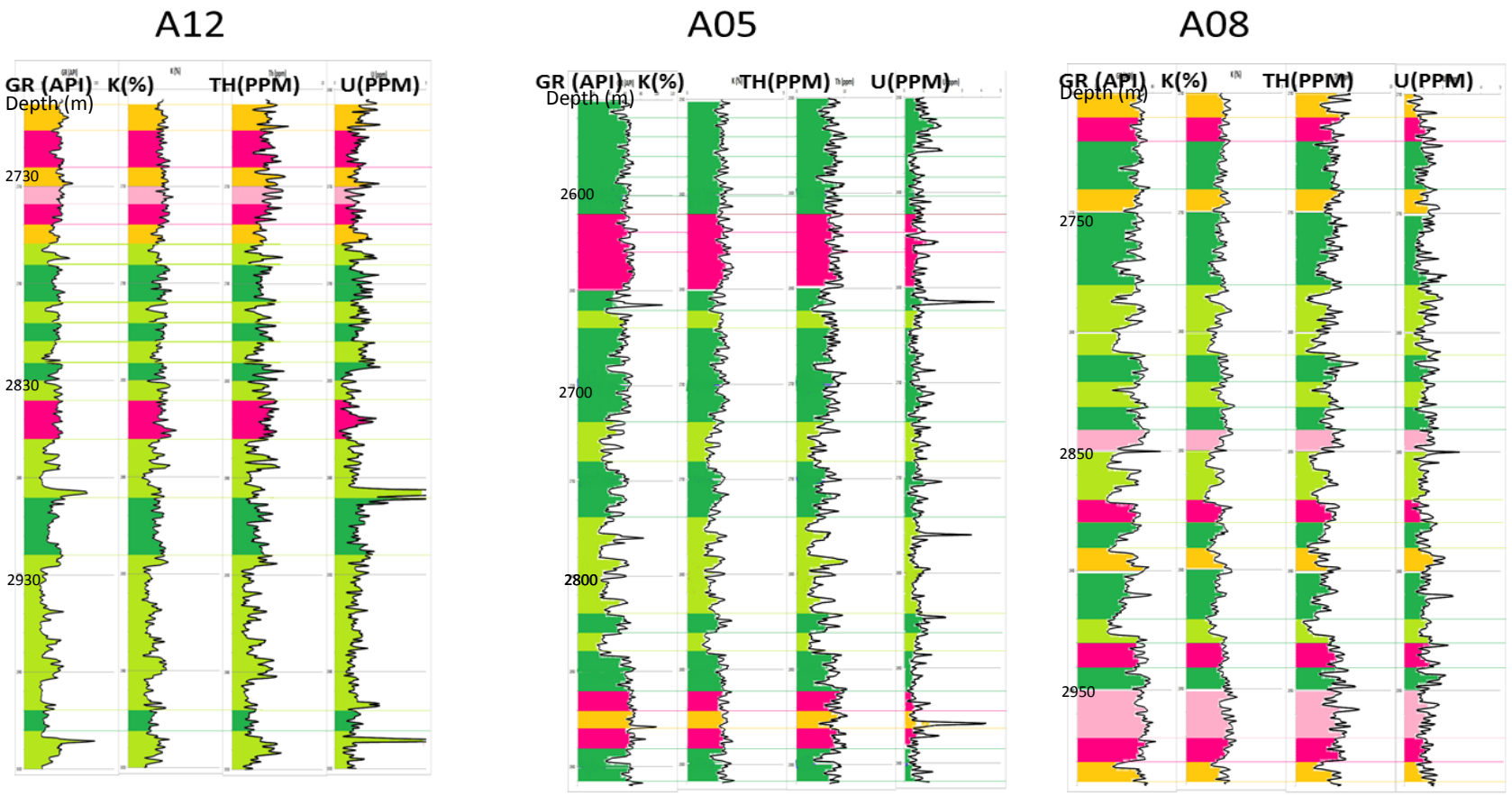


Figure 5.11 b: Cluster analysis results for all samples showing GR and SGR logs continued. Colors designate different petrofacies as averages shown in Table 5.5.

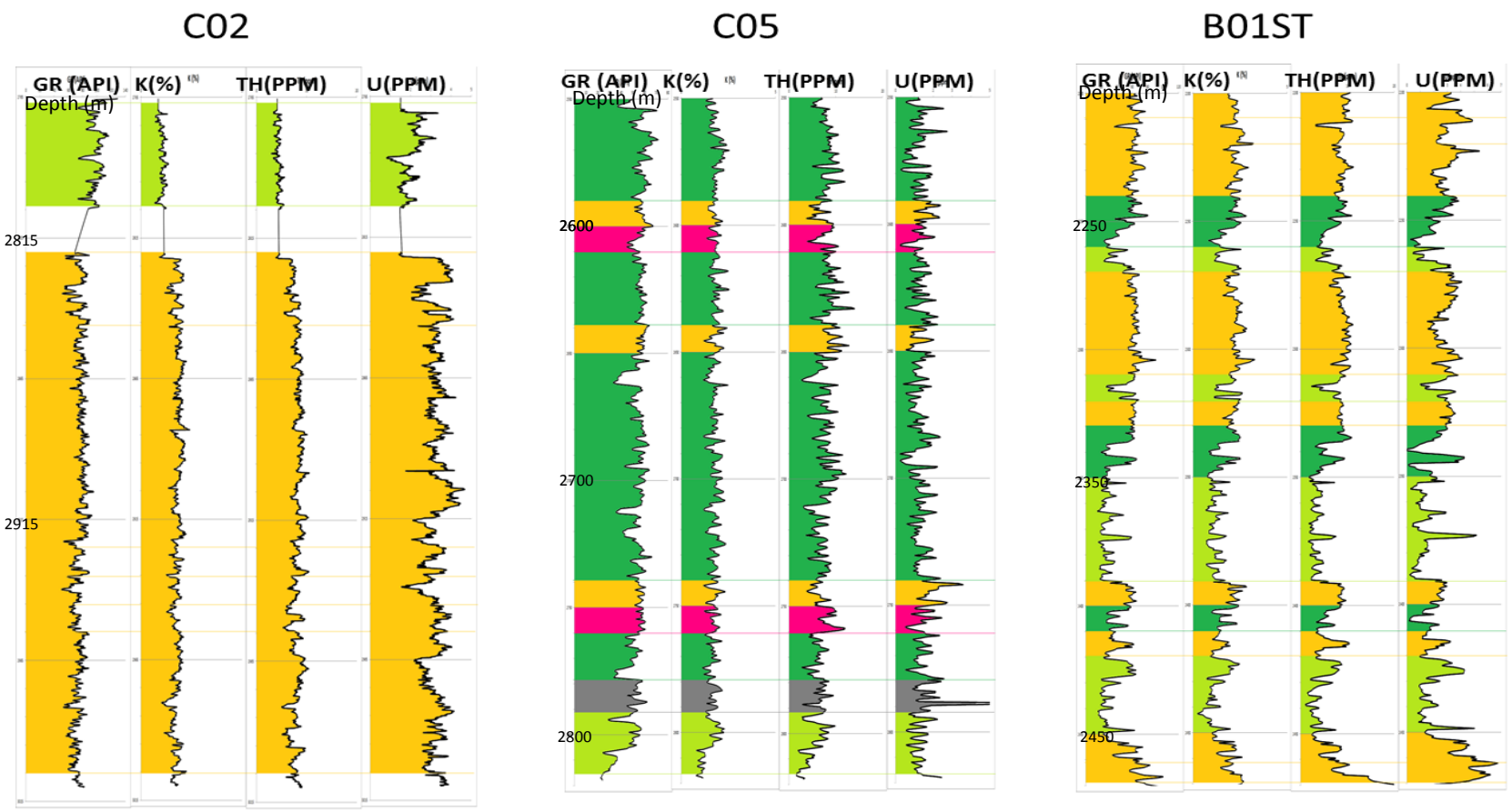


Figure 5.11 c: Cluster analysis results for all samples showing GR and SGR logs continued. Colors designate different petrofacies as averages shown in Table 5.5.

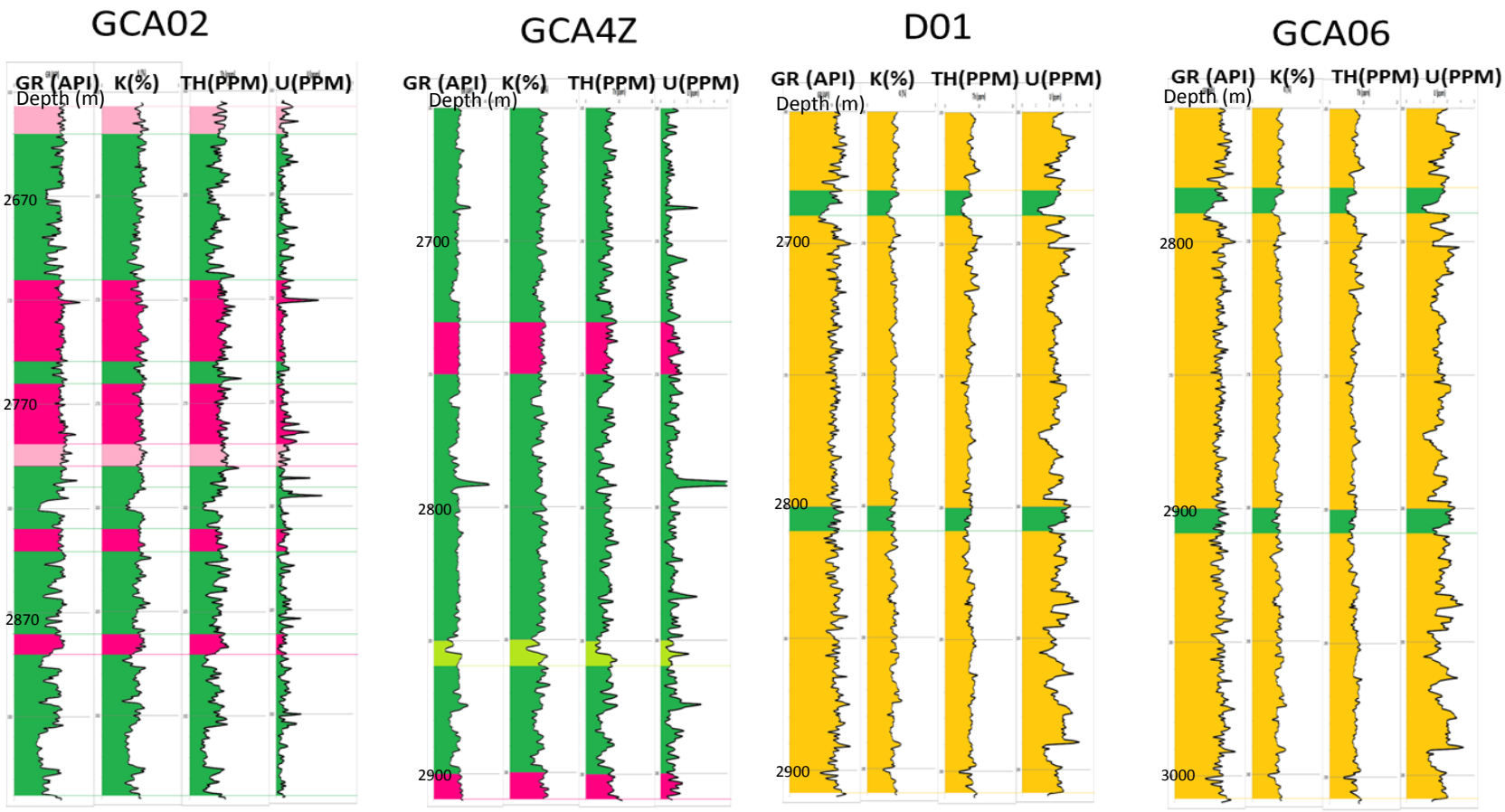


Figure 5.11 d: Cluster analysis results for all samples showing GR and SGR logs continued. Colors designate different petrofacies as averages shown in Table 5.5.

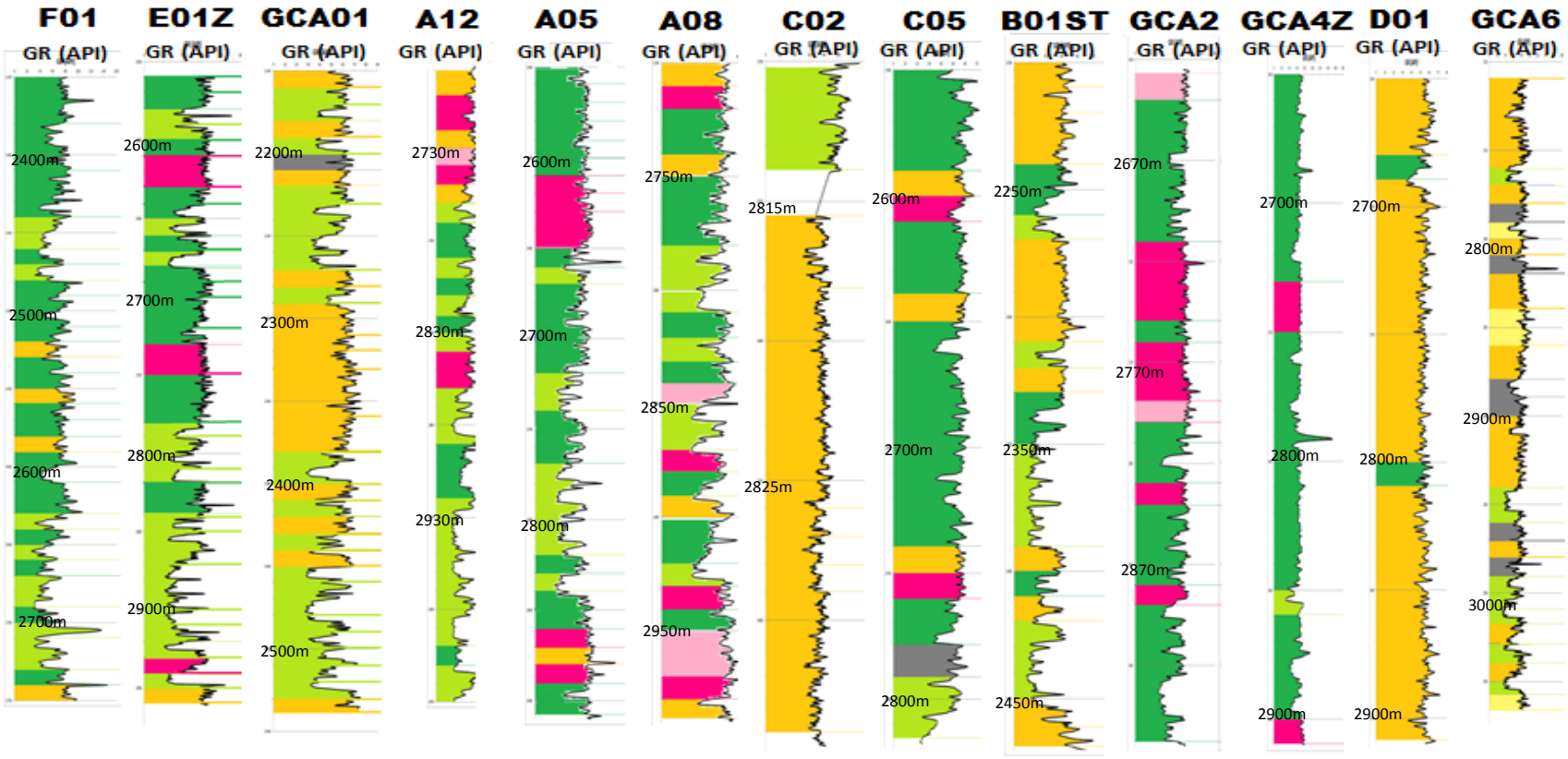


Figure 5.12: GR logs showing cluster analysis results for all samples for all 13 well locations in the study area. Colors designate different petrofacies as averages shown in Table 5.5.

5.9 Analysis #9: Clustering All Samples with GR<60 API from All Wells

There were a total of 282 samples which have GR readings less than 60API from the 13 wells. Figure 5.13 is the dendogram for that analysis. The blue cluster is the first major cluster to develop whereas the red cluster is the second one. Blue and red clusters are linked together to link to third gray cluster or low similarity cluster.

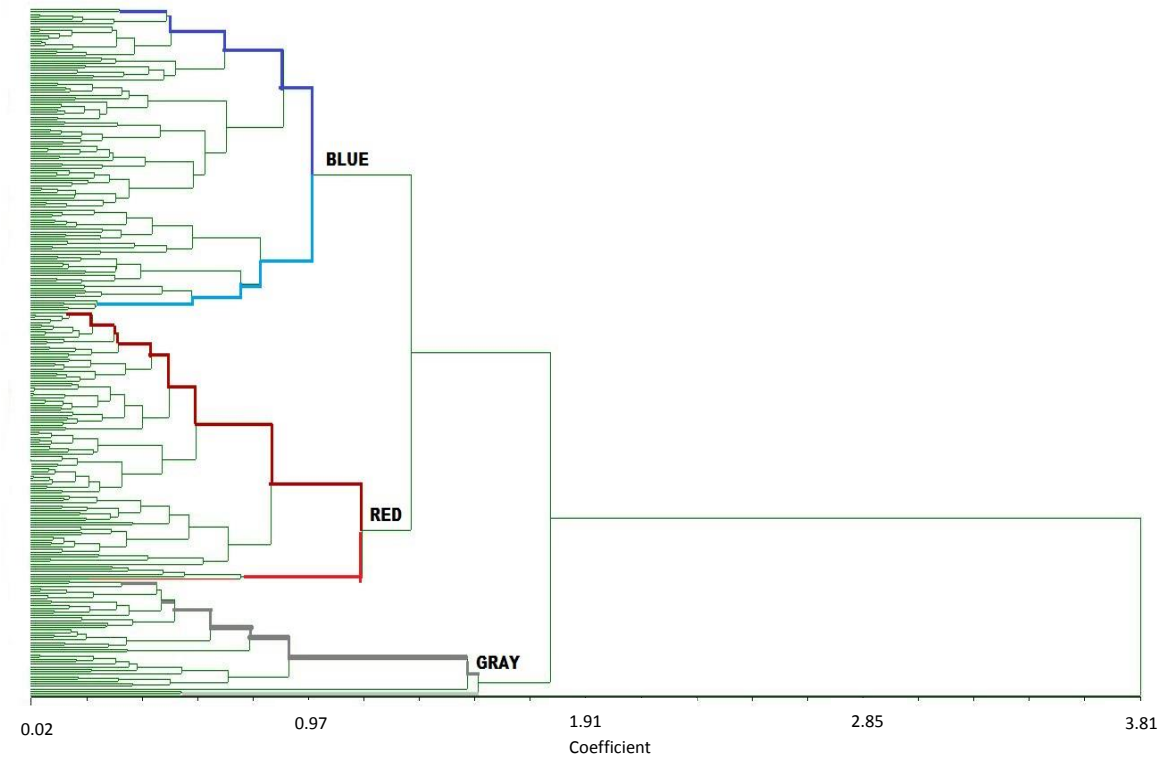


Figure 5.13: Dendogram for all 282 sand samples from 13 wells with three main clusters.

Results are also shown on well logs (Figure 5.14 a,b,c,d and Figure 5.15). According to porosity and permeability averages (Table 5.6), the red petrofacies and the dark-blue petrofacies have the highest quality reservoirs. They are followed by the light-blue petrofacies with ~79 mD permeability and 0.12% porosity.

	GR (API)	K (%)	Th(ppm)	U(ppm)	Permeability (md)	Porosity (%)	Th/K	Th/U
Dark-Blue Petrofacies B	44.40	1.64	4.78	0.63	246.00	0.20	2.91	7.55
Light-Blue Petrofacies	49.83	1.62	6.88	1.17	78.80	0.12	4.24	5.88
Dark-Red Petrofacies A	34.57	0.85	3.21	1.09	287.17	0.21	3.78	2.94
Light-Red Petrofacies	43.81	1.33	3.59	2.35	294.10	0.15	2.70	1.53
Dark-Gray Petrofacies	49.23	1.81	7.56	2.60	3.23	0.18	4.17	2.90
Light-Gray Petrofacies	48.79	1.89	9.85	0.65	2.85	0.13	5.21	15.15
Outliers	45.01	1.38	2.93	6.62	0.28	0.13	2.12	0.44

Table 5.6: Averages of cluster for analysis #9.

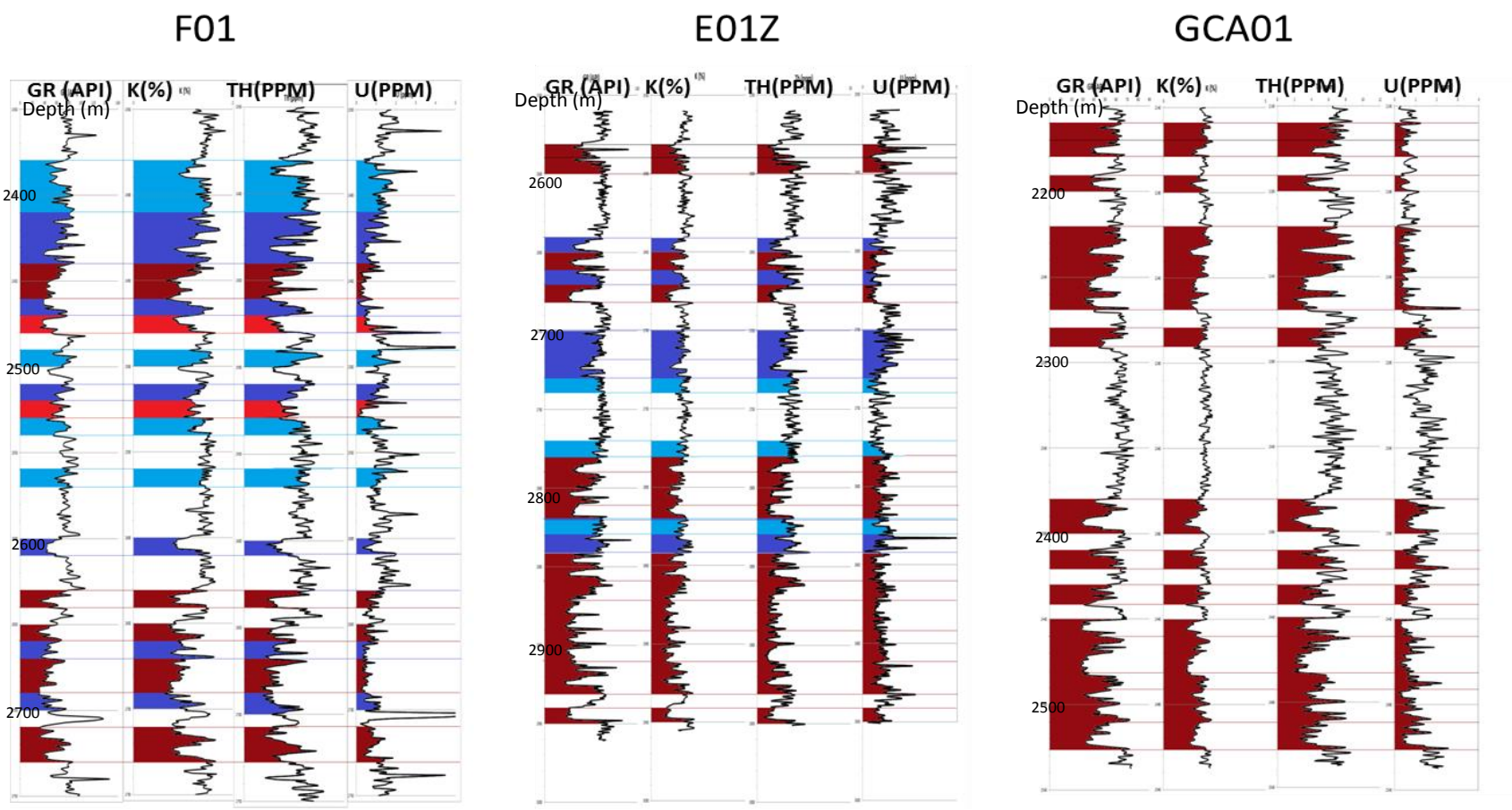


Figure 5.14 a: Cluster analysis results for all samples GR< 60 API showing GR and SGR logs. Colors designate different petrofacies as shown in Table 5.6.

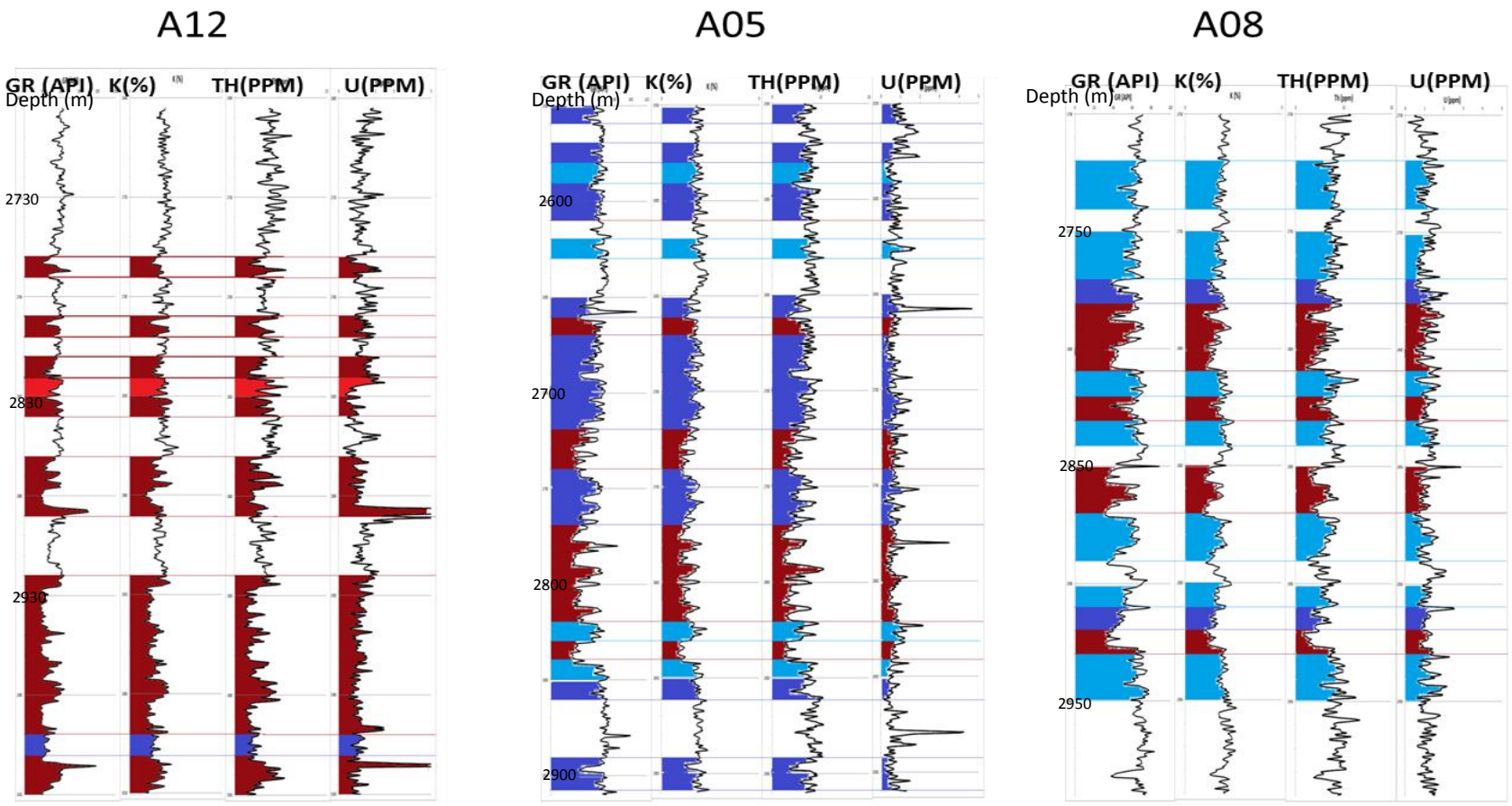


Figure 5.14 b: Cluster analysis results for all samples GR<60 API showing GR and SGR logs continued. Colors designate different petrofacies as shown in Table 5.6.

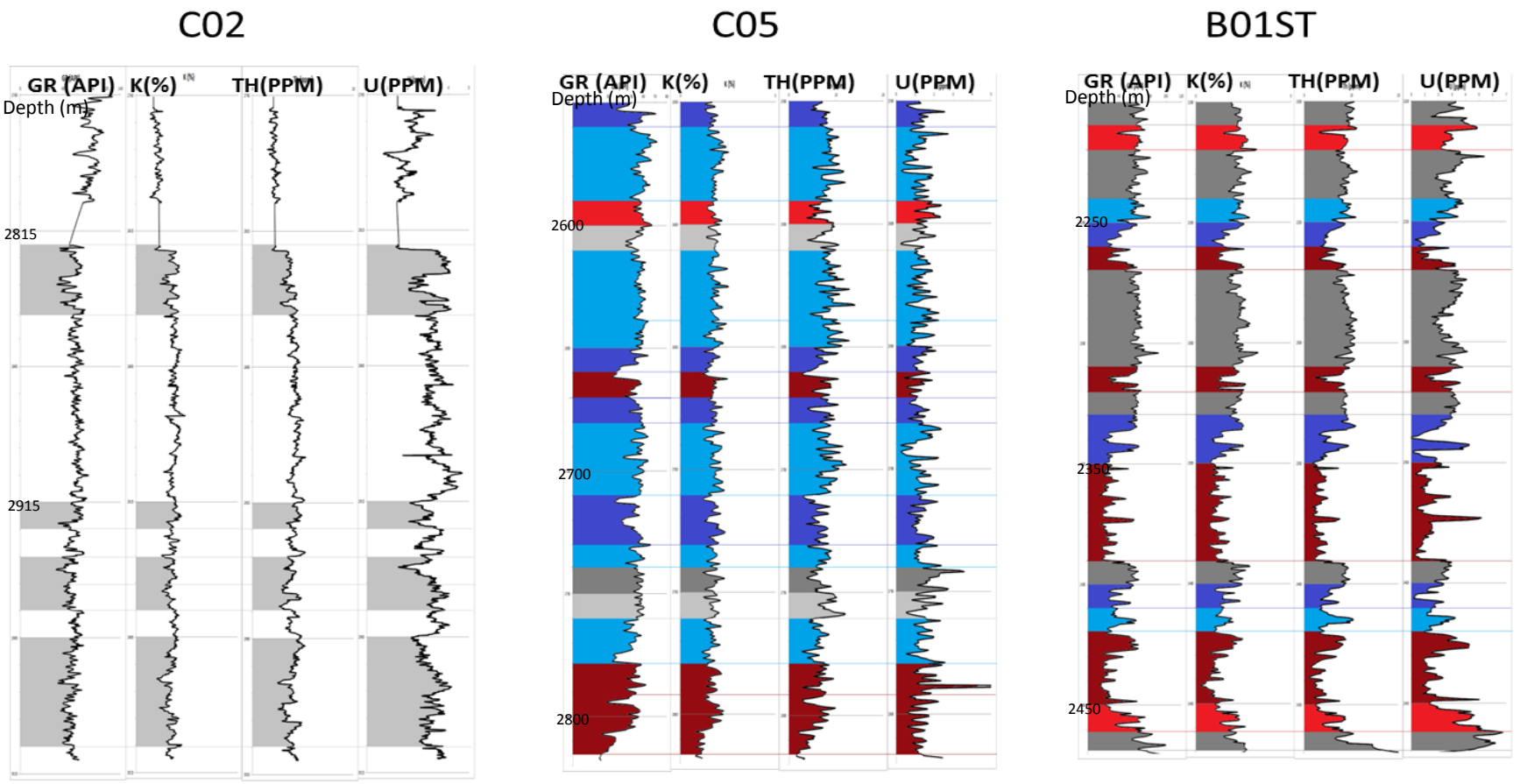


Figure 5.14 c: Cluster analysis results for all samples GR<60 API showing GR and SGR logs continued. Colors designate different petrofacies as shown in Table 5.6.

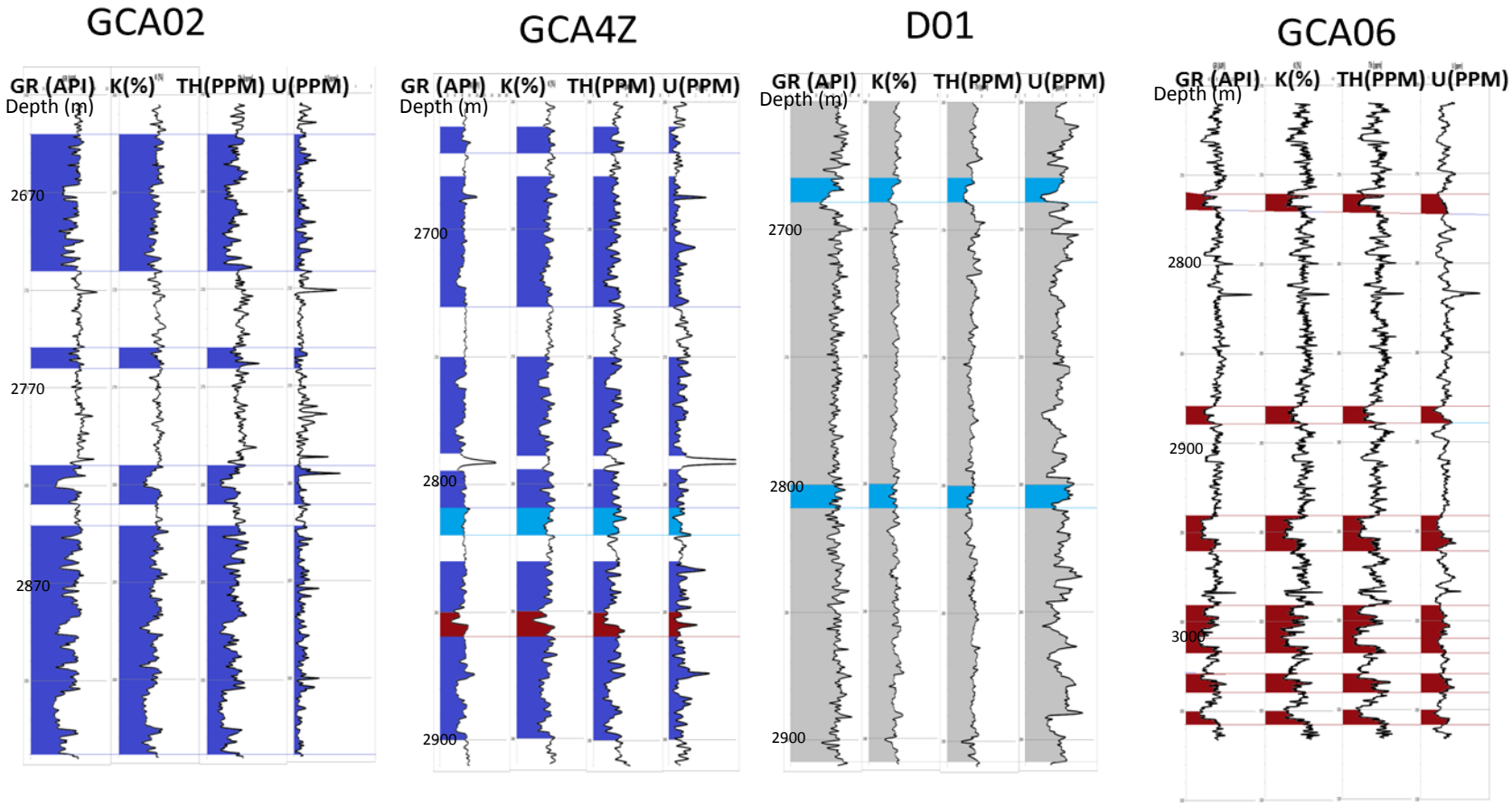


Figure 5.14 d: Cluster analysis results for all samples GR<60 API showing GR and SGR logs continued. Colors designate different petrofacies as shown in Table 5.6.

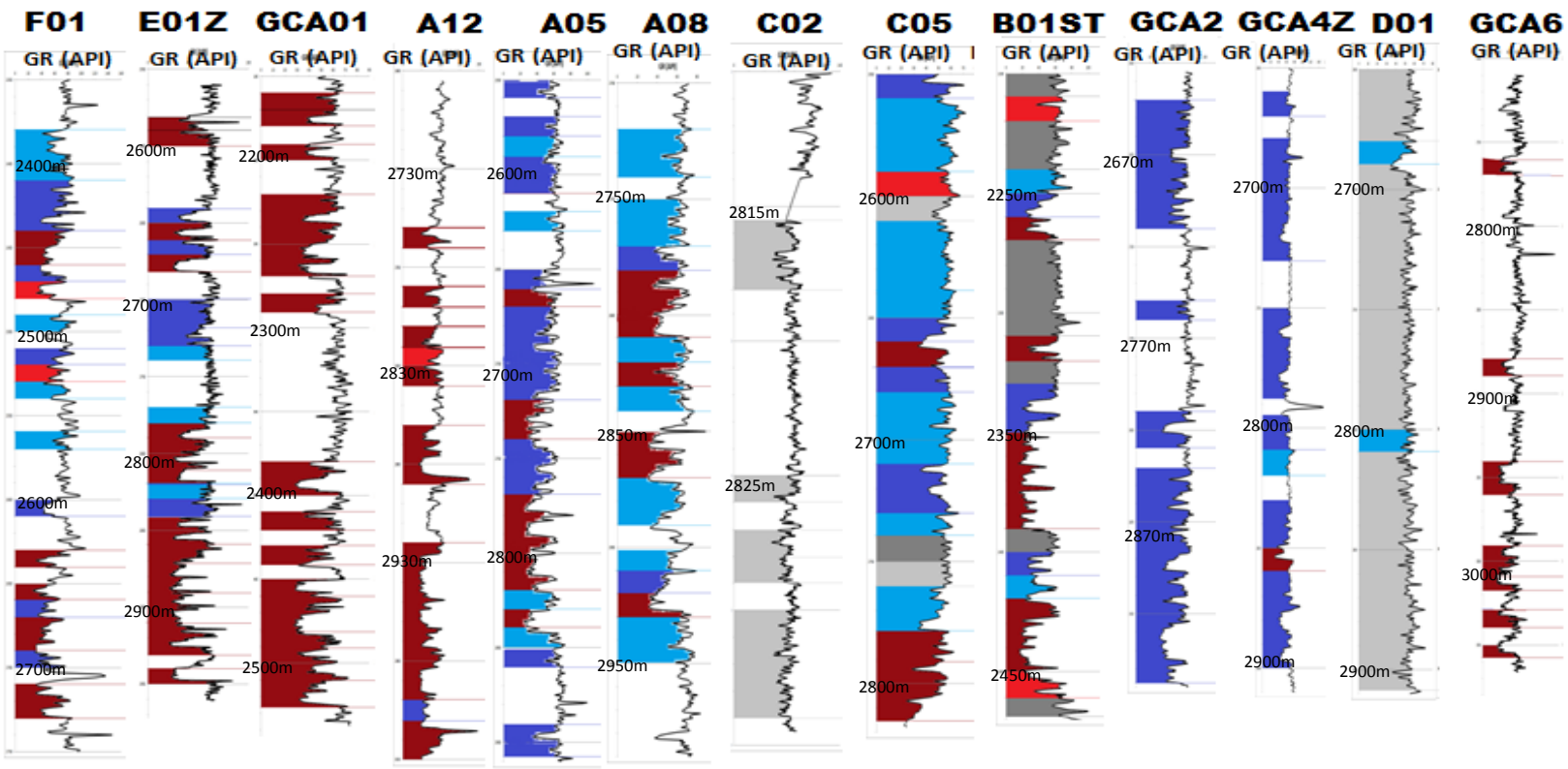


Figure 5.15: GR logs showing cluster analysis results for all samples GR<60API for all 13 well locations in the study area. Colors designate different petrofacies as shown in Table 5.6.

CHAPTER VI: DISCUSSION AND CONCLUSION

6.1 Discussion of the Results

Many authors and publications have been concerned with the sand provenance of the Productive Series in the South Caspian Basin (e.g. Allen et al., 2006; Morton et al., 2003; Hinds et al., 2007). Different methods such as heavy mineral analysis and zircon age constraints are discussed by them. Their main focus has been on the paleo-Volga, which is currently the most obvious sediment contributor to the South Caspian Basin.

However, in geologic time, sediments have been transported into the SCB from the north by the paleo-Volga, from the west by paleo-Kura and from the east by paleo-Amu-Darya and paleo-Uzboy. There are only a few references mentioning the paleo-Uzboy River. Sediments of the paleo-Volga are known to have better reservoir qualities than the sediments of paleo-Kura. Sediments derived from the paleo-Volga are dominated by high proportions of sand-sized quartz and are relatively well sorted and well rounded. On the other hand, sediments derived from the paleo-Kura have much higher proportions of feldspar and lithics and poor sorting, which result in relatively poor permeabilities (Morton et al., 2003). Large volumes of mud-rich sediments were delivered into the eastern part of the basin by paleo-Amu-Darya. In the offshore areas in the South Caspian Basin, the reservoir sequences that deposited by paleo-Amu-Darya are thicker than 4 km (Smith-Rouch, 2006).

Different cluster analyses were tried in order to understand the sediment provenance by using elemental compositions. Two analyses were done with the samples from all available wells (13) for this project, and the results were significant enough to distinguish the compositional end members. Cluster analysis results of 13 wells for all samples (analysis #8) and all sandstone samples (analysis #9) are very similar to each other and show the same probable sources. The light-green petrofacies A (Table 5.5) has the best reservoir qualities with 293.95 mD average permeability and 0.21% average porosity for all samples. This cluster is equal to the dark-red petrofacies in the analysis of sand-rich samples (Table 5.6) with 287.17 mD permeability and 0.21% porosity. Our results suggest these clusters have samples derived mainly from the paleo-Volga river system with better reservoir qualities. The light-green petrofacies A is mostly seen in the western wells of the study area and also the dark-red petrofacies is mostly observed in the western part of the study area (Figure 6.1). The dark-green cluster petrofacies B has the second best reservoir quality of all clusters with 158 mD average permeability and 0.17% average porosity. This cluster is almost equal to the dark-blue cluster in sand-rich sample analysis with 246 mD average permeability and 0.20% average porosity. According to these results, the dark-green and dark-blue clusters (Figure 6.1) are derived mainly from the eastern paleo-river systems which are the paleo-Amu-Darya and paleo-Uzboy Rivers. Petrofacies B are mostly observed at the eastern wells; however they are also present at the western side. This suggests conformance with the previous interpretations that the paleo-Amu-Darya delta was prograding while the paleo-Volga was backstepping or retrograding during the deposition of the Balakhany Suite (Abreu and Nummendal,

2007). Different clusters are intercalating each other in the study area, especially for the wells located in the middle. A probable reason for this is that the relative importance of each deltaic system for sediment sources changed through time through different phases of progradation and sediment-load volumes. According to Abreu and Nummendal (2007), paleo-Kura delta was aggrading or in balance with accommodation space during deposition of the Balakhany Suite. Different paleo-river systems were active during the deposition of the Productive Series.

The pink and yellow clusters of all samples (Figure 5.12 and 6.1) have GR averages more than 60 API with very low permeabilities and porosities. Among them the dark-yellow cluster is the most common one with relatively higher porosity average 0.12%. The dark-yellow cluster in all samples analysis #8 is almost equal to the gray clusters in sandstone analysis #9 (Figure 6.1). When the logs are analyzed, the dark-yellow cluster is seen where there are no abrupt changes in GR and SGR logs, which are interpreted as where the river channels are not present. This cluster facies is observed in some parts of well GCA01 and B01st and most of the parts of C02 and D01.

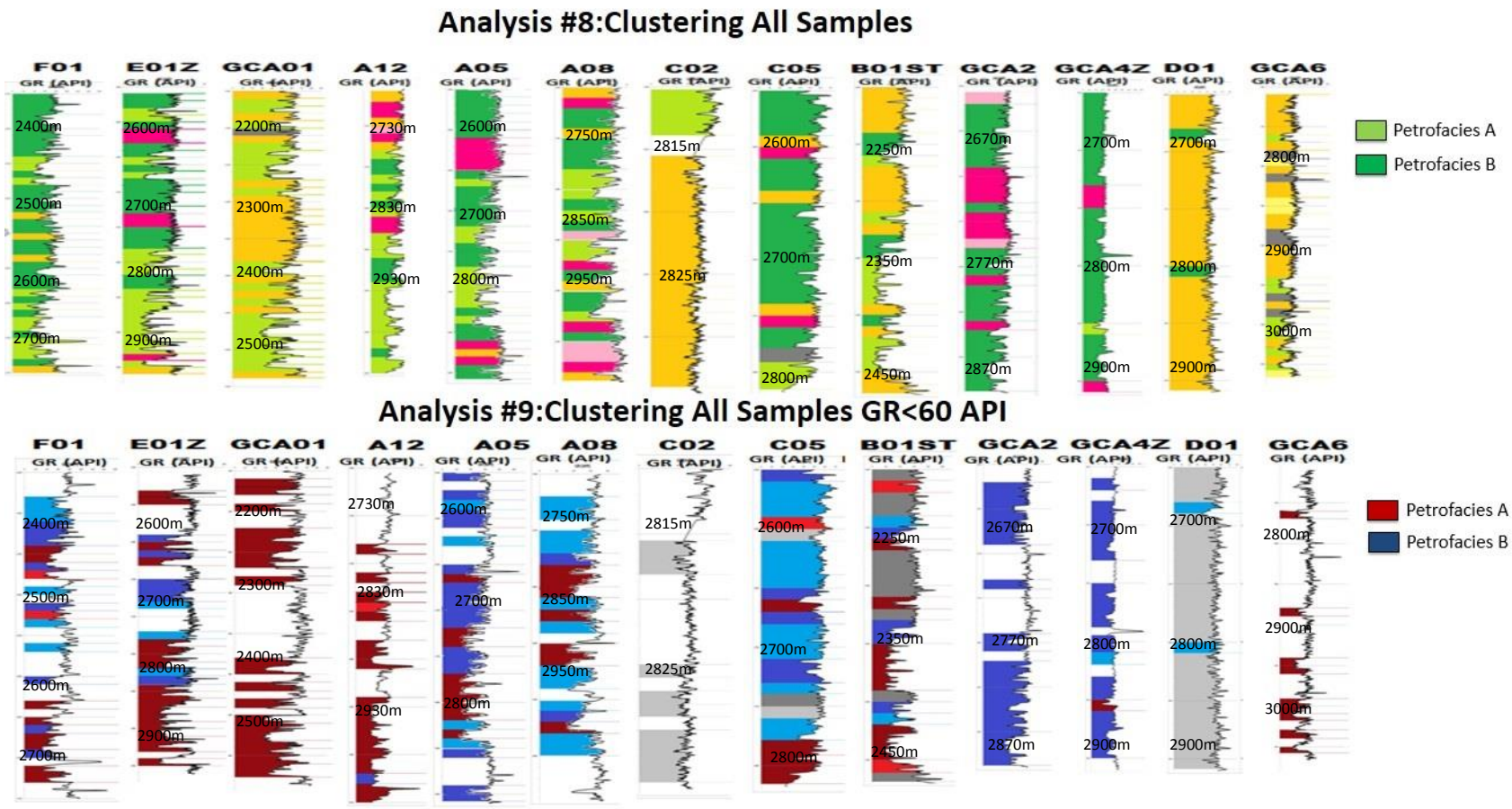


Figure 6.1: GR logs showing cluster analysis results for analysis #8 & #9. Petrofacies A is most probably derived from paleo-Volga, whereas petrofacies B is most probably derived from eastern sediment contributors paleo-Amu-Darya and paleo-Uzboy.

Averages for porosity, permeability, and spectral gamma ray elements are calculated for the Balakhany VIII reservoir unit for all the wells (Figure 6.2).

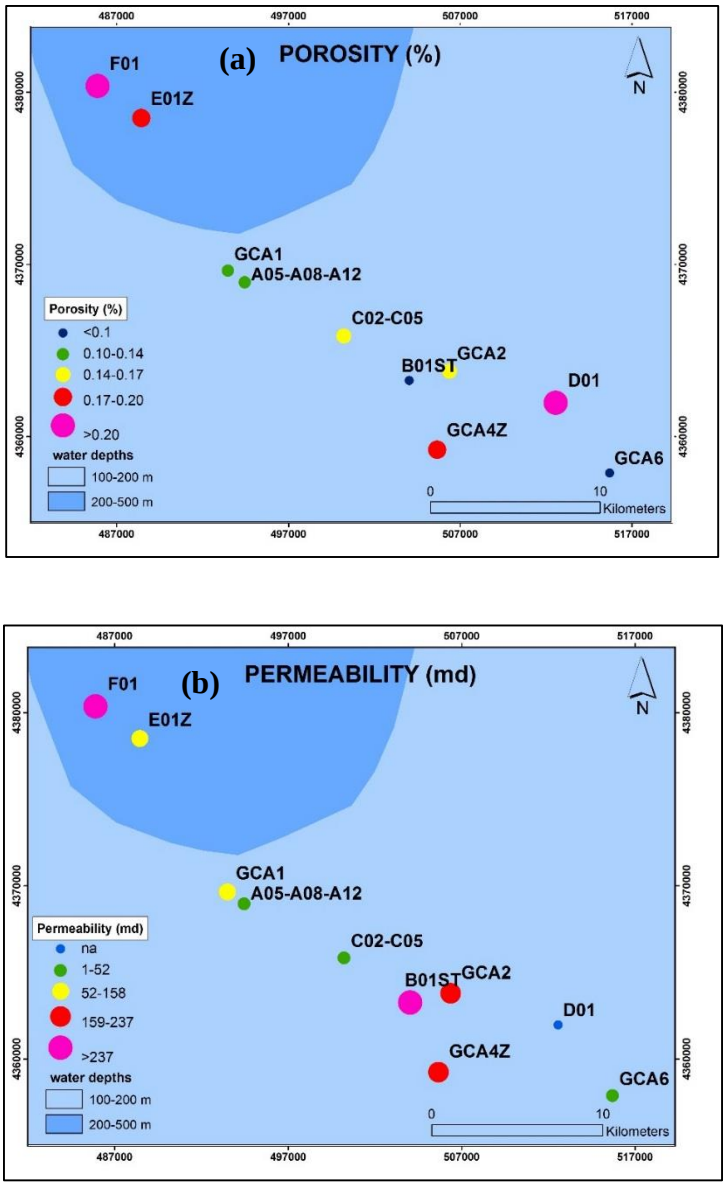


Figure 6.2: (a) Average porosity (%) and **(b)** permeability (md) values for the Balakhany VIII sub suite in the study area. Northwestern well F01 has highest porosity and permeability values.

Wells F01, E01Z, GCA2, and GCA4Z are the wells with relatively good porosity and permeabilities. Cluster results of these wells are also distinct. Provenance for the GCA2 and GCA4Z is most probably eastern paleo-river systems, whereas for the F01 and E01Z is paleo-Volga.

Our study area is located in between the southeastern edge of paleo-Volga and southwestern edge of unnamed paleo-river, which is probably paleo-Uzboy. These two paleo-river systems might be the most important sediment contributors for our study area.

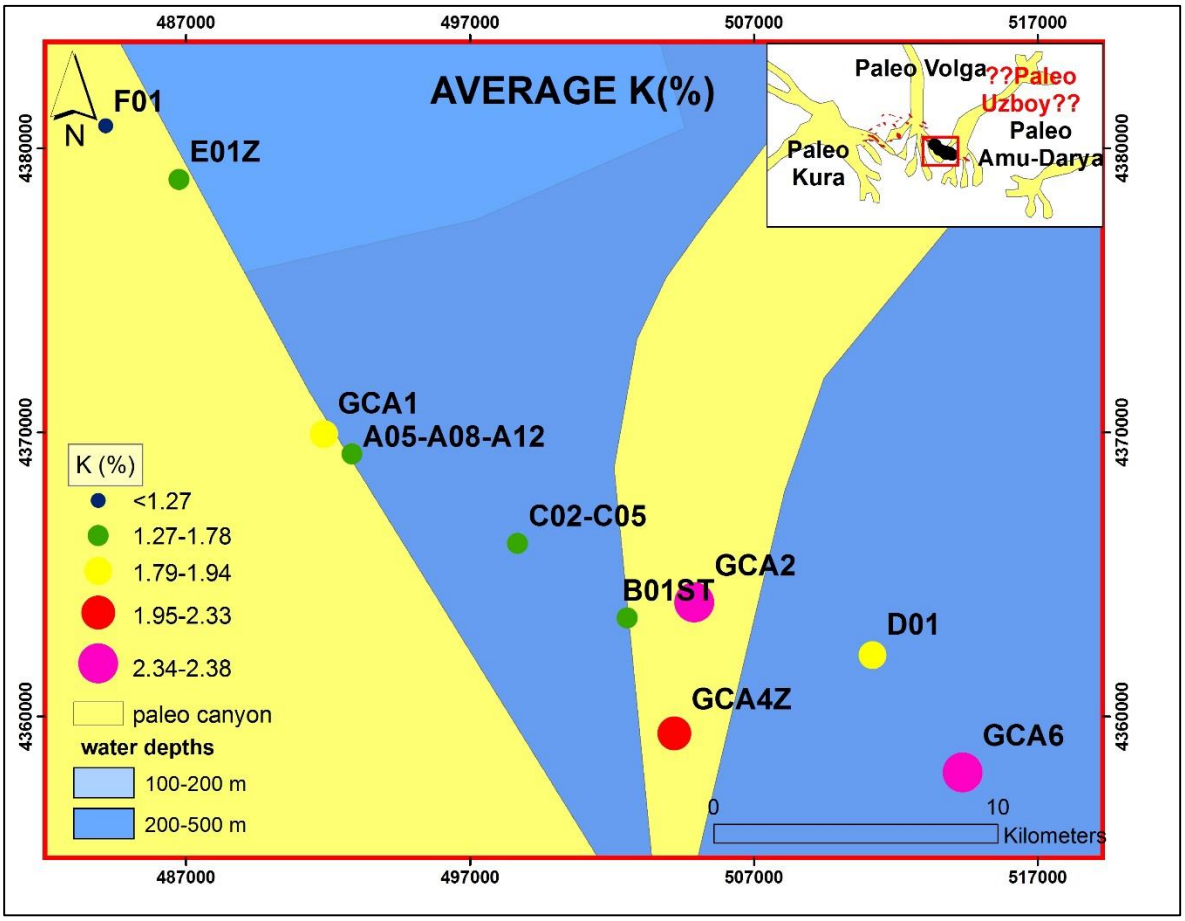


Figure 6.3: Location of the wells within paleo-river systems. Red rectangle in the small map shows the extent of the map. Average K (%) is shown on the map for all wells.

The Uzboy River is a branch of Amu-Darya River which is flowing into the Caspian Sea. The Uzboy River received water from the Amu-Darya and the Sarikamysh River directly from the Tien-Shan mountain ranges, hence from a drainage basin under a climate totally different from that of the Volga River (Ferronsky et al., 1999).

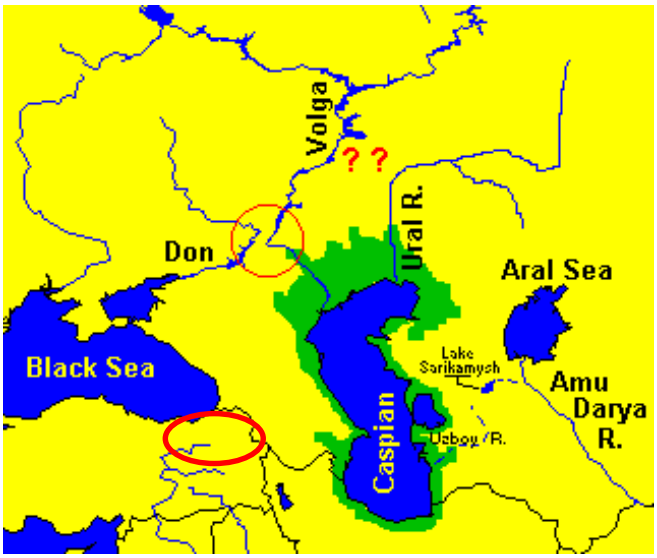


Figure 6.4: River diversions in the Caspian region (Dutch, 2011).

According to K (%) averages, the highest values are observed in the southeastern part of the study area, whereas the lowest ones are limited to the northwest; k-dominance tends to suggest that the sediment is derived from a provenance that is relatively unweathered and has shorter transportation distances (Ellis, 1987). Our results are consistent with that situation. As it is suggested, the paleo-Uzboy and paleo-Amu-Darya Rivers which have shorter transportation distances, are most probable sediment provenances for the southeastern wells. Wells GCA2 and GCA4Z have especially distinct and obvious cluster

results, indicating sediment provenance is the eastern river systems. The Paleo-Volga, with a longer transportation distance, should be the main sediment contributor for the northwestern F01 and E01Z wells.

When we examine the averages of spectral gamma ray elements throughout the study area, higher values are observed in wells in the southeastern part of the study area (Figure 6.5). This is an indication there is a high eastern sediment contributor that is most probably the paleo-Amu-Darya and the paleo-Uzboy rivers.

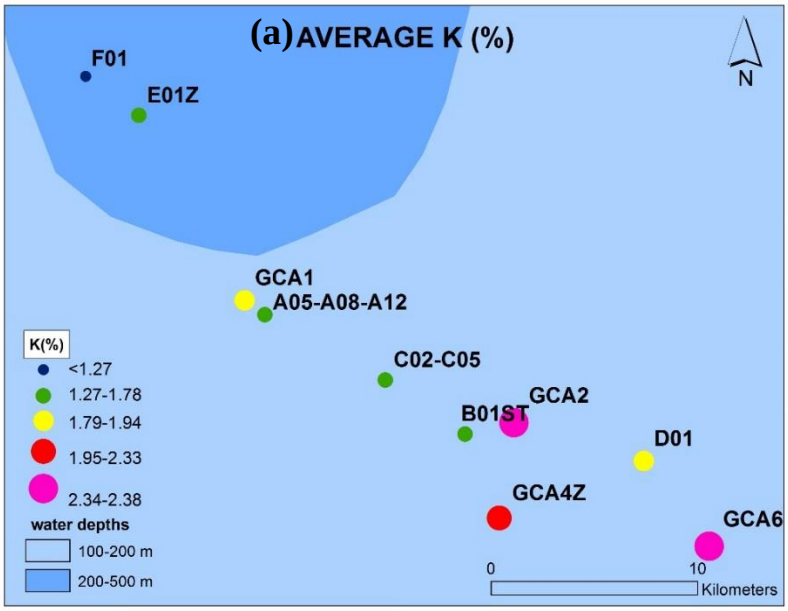
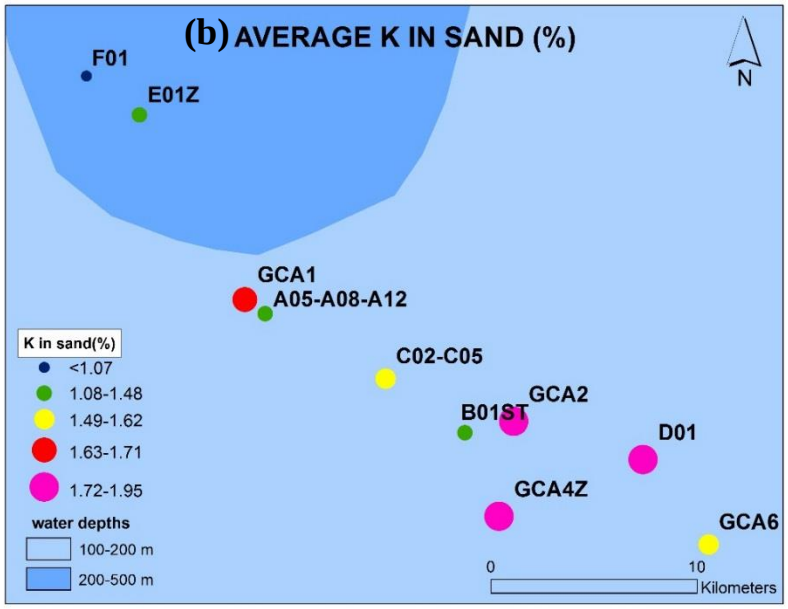


Figure 6.5 a: Averages of K (%) for all wells in the study. See location in Figure 6.2 and Table 1.1



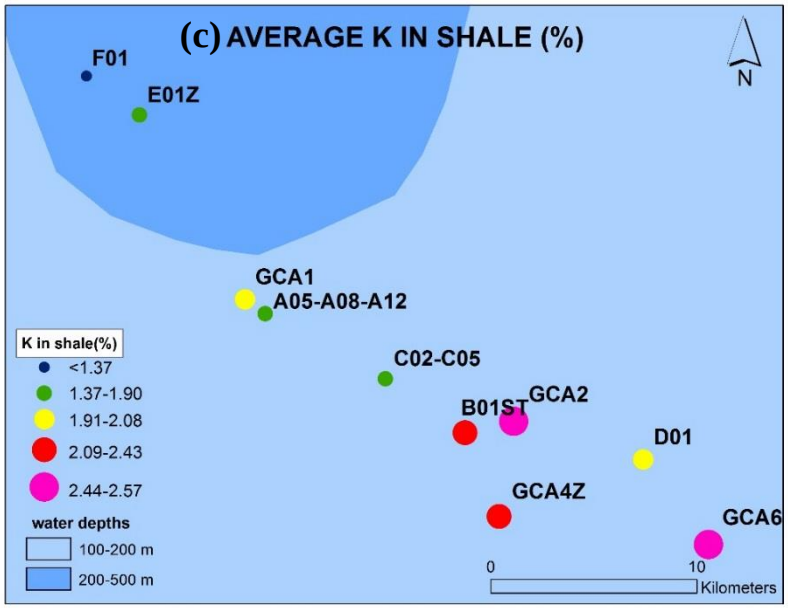
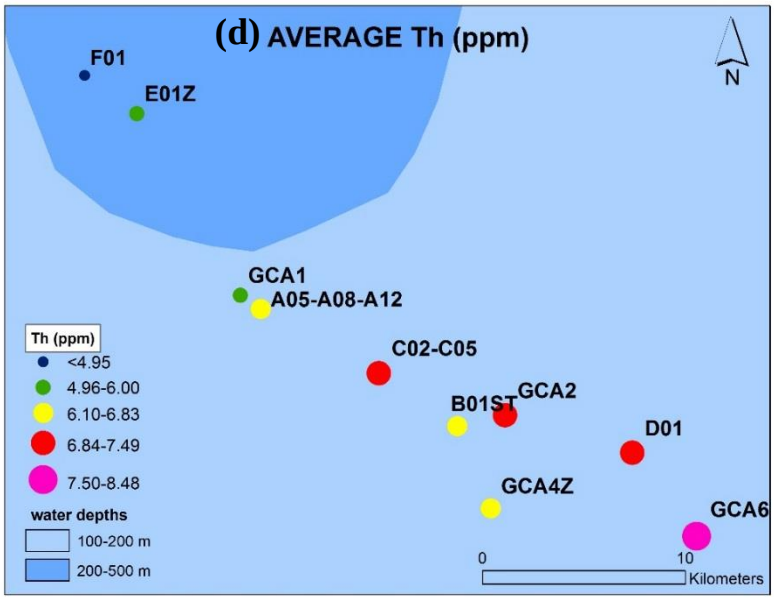


Figure 6.5 b&c: Averages of K (%) in sand and in shale for all wells in the study. See location in Figure 6.2 and Table 1.1



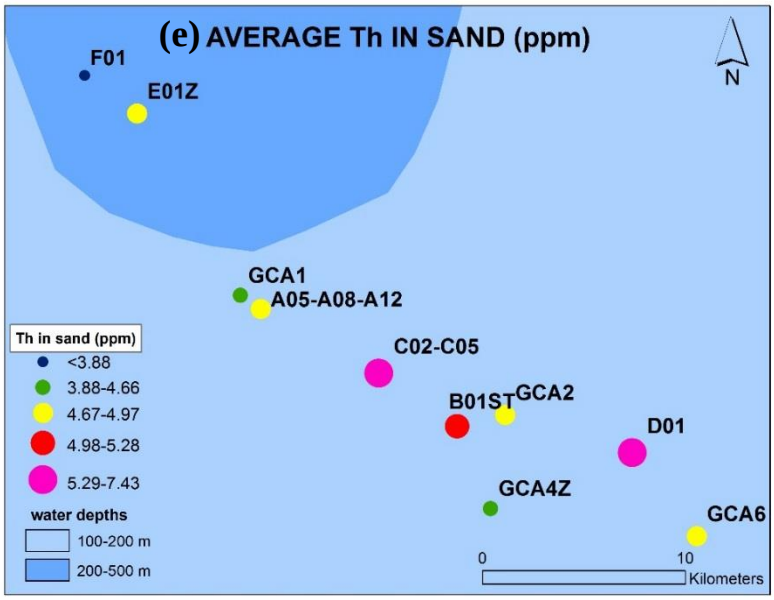
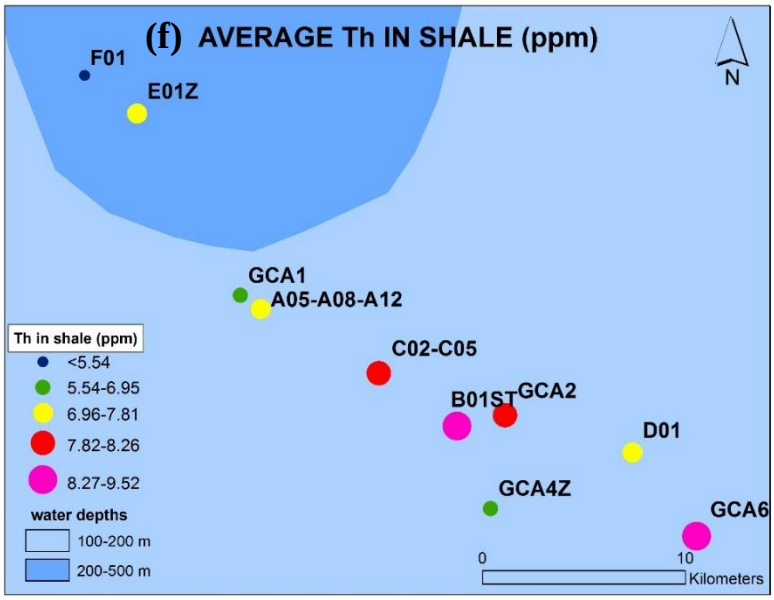


Figure 6.5 d&e: Averages of Th (ppm) and averages in sands for all wells in the study. See location in Figure 6.2 and Table 1.1



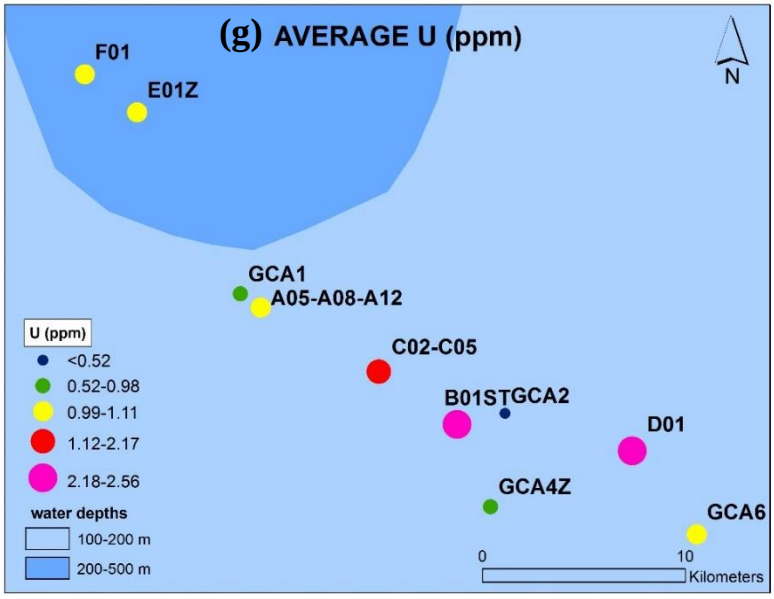
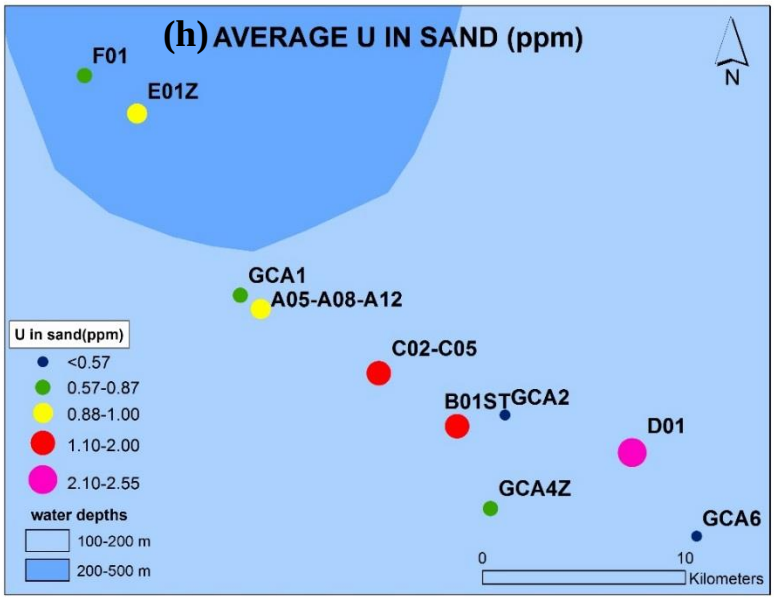


Figure 6.5 f&g: Averages of Th (ppm) in shale and averages of U (ppm) for all wells in the study. See location in Figure 6.2 and Table 1.1



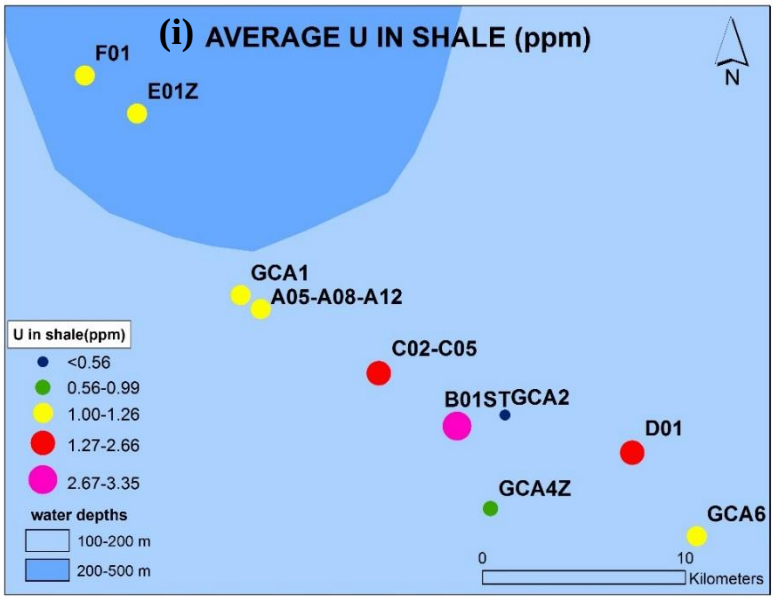


Figure 6.5 h&i: Averages of U (ppm) in sand and in shale for all wells in the study. See location in Figure 6.2 and Table 1.1

When we look at the N-S section through the South Caspian basin (Figure 3.1), sediment thicknesses are extremely high when compared to the Central Caspian Basin. If the paleo-Volga was the main sediment contributor to the basin for the Productive Series, then the Central Caspian Basin should also have more thickness for Neogene.

When all of the attributes are considered, the results of this study are consistent with the paleogeography of the area. During the deposition of Balakhany, the paleo-Volga was backstepping and the paleo-Amu-Darya was prograding (Figure 3.4). According to our results, there is a strong signal for the sediment provenance for the Balakhany VIII sub-suite being from the east, which is paleo-Amu-Darya and paleo-Uzboy Rivers.

Cluster analysis was an effective method to distinguish compositional end members by using spectral gamma ray elements. Multiple trials were done to achieve useful clusters. Analysis was done by using two different data sets for all 13 wells. The first one including all samples no matter what the lithology was, and, the second one including only samples having gr<60 api, which can be generalized as sandstones. Those analyses gave very similar results, showing the reliability of the method. The cluster analysis method also worked well to distinguish sandy formations from others.

6.2 Conclusion

Cluster analysis was applied to spectral gamma ray logs results from the Balakhany VIII sub-suite in the Azeri-Chirag-Gunashli offshore field in the South Caspian Basin. The main purpose was to discriminate the source of reservoir sediments that came from different paleo-river systems, in order to determine the provenance of reservoir sands and to compare their reservoir properties.

Many different analyses have been attempted by changing variables and sample sets to answer different questions. The most effective and consistent results are obtained by limiting the variables to the three spectral gamma ray elements. Cluster results are compared with porosity and permeability averages for each of the clusters. Different clusters are assigned as different petrofacies as they are clustered according to their similar petrophysical properties. Petrofacies established by clustering routines were consistent in different analysis showing the reliability of our method. Sand-rich petrofacies A (analysis #8 and #9), having better reservoir qualities observed mostly in the western side of the study area, is found to be derived from N-S draining paleo-Volga. However, shalier petrofacies B (analysis #8 and #9), having less but good porosity and permeability averages and higher K average, is derived from eastern mud-dominated sediment contributors, the paleo-Amu-Darya and paleo-Uzboy. A strong eastern signal is observed for the sediment provenance from the results of this study. The Paleo-Amu-Darya and paleo-Uzboy Rivers were the main sediment contributors during most of the Balakhany Suite deposition. Our results are consistent with Abreu and Nummendal,

(2007), who interpreted that paleo-Volga was backstepping and the paleo-Amu-Darya was prograding during the deposition of Balakhany Suite. The distribution of SGR elements within different well locations along with porosity and permeability data also support a strong eastern sediment contribution into the ACG field. The dominant systems brought sediments from the eastern side of the study area as sediments from the Greater Balkhan, Kopet Dag, Pamir and Tian Shan mountain ranges were uplifted.

When all the results are considered, cluster analysis was an efficient method to distinguish compositional end members and compare the reservoir qualities. SGR log data were effective to use in cluster analysis. In spite of the limited number of variables, cluster analysis allowed us to obtain meaningful results. Coring operations are much more expensive than the logging. If cluster analysis is done with chemical and heavy mineral results from core analysis, more detailed results might be obtainable. However, in this study, SGR data with cluster analysis proved to be useful in distinguishing compositional end members sediment provenances and their relationship to porosity and permeability trends.

Future work could provide additional insights in the study area. If core data made available in the future, additional multivariate methods can be tried to improve this study.

REFERENCES

Abreu, V., and Nummedal, D., 2007, Miocene to Quaternary sequence stratigraphy of the South and Central Caspian Basins: in, Yilmaz, P.O., Isaksen, G.H., eds, oil and Gas of the Greater Caspian area: AAPG Studies in Geology 55, p. 65 – 86.

Adams, J.S.A., and Weaver, C.E., 1958, Thorium – to uranium ratios as indicators of sedimentary processes: Example of concept of geochemical facies: AAPG Bulletin, vol. 42. No. 2, p. 387 – 430.

Aitchison, J., 1986. The Statistical Analysis of Compositional Data: Wiley,NY, 416p.

Allen, M.B., Morton, A.C., Fanning, C.M., Ismail-Zadeh, A.J., and Kroonenberg, S.B., 2006, Zircon age constraints on sediment provenance in the Caspian region: Journal of the Geological Society, London, vol. 163, p. 647-655

Aliyeva, E.G., 2003, Reservoirs of the Lower Pliocene Productive Series at the Western Flank of the South Caspian Basin, Lithology and Mineral Resources: vol.40 no.3, p.267-278

Anjos, R.M, Carvalho, C., Veiga, R., and Macario, K., 2008, Provenance and transport processes of sediments along the southern Brazilian coast: in A.S. Paschoa, ed., Natural Radiation Environment-8th International Symposium, Buzios, Rio de Janeiro (Brazil), AIP publishing, p. 473-476

Baghirov, E., 2007, Vertical and Lateral Distribution and Continuity of Balakhany VIII Formation of Azeri Field, Offshore Azerbaijan: M.S. Thesis, University Of Houston, 80 p.

Bohacs, K.M., Carroll, A.R., Neal, J.E., and Mankiewicz, P.J., 2000, Lake-basin type, source potential, and hydrocarbon character: an integrated sequence-stratigraphic-geochemical framework: in, E.H. Gierlowski-Kordesch, K.R. Kelts, eds., Lake Basins through Space and Time: AAPG studies in Geology 46, p. 3 – 34.

BP EIA report, 2007, Azeri, Chirag, Gunashli Full Field Development -Produced Water Disposal Project- Environmental and Socio-economic Impact Assessment: Internal, unpublished BP report, 350 p.

BP in Azerbaijan, Sustainability Report, 2011: Internal, unpublished BP report, 60p.

Brunet, M.-F., Korotaev, M.V., Ershov, A.V., and Nikishin, A.M., 2003. The South Caspian Basin: a review of its evolution from subsidence modelling: Sedimentary Geology vol. 156, p. 119-148.

Caster, K.E., 1934, The stratigraphy and paleontology of northwestern Pennsylvania, Part 1: Bulletins of American Paleontology, vol. 21, no. 71, 185 p.

Davis, J.C., 1973, Statistics and Data Analysis in Geology: John Wiley& Sons Inc., Cluster Analysis, Chapter 7, p.456-473

Devlin, W., Cogswell, J., Gaskins, G., Isaksen, G., Pitcher, D., Puls, D., Stanley, K., and Wall, G. (1999). South Caspian Basin: young, cool, and full of promise: GSA Today, vol. 9, p. 1–9.

Doveton, J.H., and Merriman, D.F., 2004, Borehole petrophysical chemostratigraphy of Pennsylvanian black shales in the Kansas subsurface: Chemical Geology, vol. 206, p. 249 – 258.

Dutch, S., 2011, University of Wisconsin-Green Bay, Earth SC-202 Physical Geology Lecture notes, retriewed from: http://www.uwgb.edu/dutchs/earthsc202notes/piracy.htm

Ellis, D.V., 1987, Well Logging for Earth Scientists: New York, Elsevier, 532 p.

Esedo, R., 2009, Spectral Gamma Ray Log Chracterization of Balakhany VIII Reservoir Unit of Azeri Chirag Gunashli Field, Offshore Azerbaijan: M.S. thesis, University of Houston, 92 p.

Ferronsky, V.I., Polyakov, V.A., Kuprin, P.N., and Lobov, A.L., 1999. The nature of variations in the level of the Caspian Sea (based on bottom-sediment data): Water Resources, vol. 26, no. 6, p. 583 –596.

Gary, M., McAfee, R., Jr., and Wolf, C.L., eds., 1972, Glossary of Geology: Washington, D.C., American Geological Institute, 805 p.

Hesselbo, S.P., 1996, Spectral gamma-ray logs in relation to clay mineralogy and sequence stratigraphy, Cenozoic of the Atlantic margin, offshore New Jersey: in, Mountain, G.S., Miller, K.G., Blum, P.,& Twitchell, D. eds Proceedings of the Ocean Drilling Program, Scientific Results, 150, New Jersey continental slope and rise p.411-422.

Hinds, D.J., Aliyeva, E., Allen, M.B., Davies, C.E., Kroonenberg, S.B., Simmons, M.D., and Vincent, S.J., 2004, Sedimentation in a discharge dominated fluvial-lacustrine system: the Neogene Productive Series of the South Caspian Basin, Azerbaijan: Marine and Petroleum Geology, v.21, p.613-638

Hinds, D. J., Simmons M.D., Allen, M.B. and Aliyeva, E., 2007, Architecture variability in the Pereriva and Balakhany Suites of the Neogene Productive Series, Azerbaijan: Implications for reservoir quality: in, Yilmaz, P.O and Isaksen, G.H ed., Oil and gas of the Greater Caspian area: AAPG Studies in Geology 55, p. 87–107.

Hudson, S. M., Johnson, C. L., Efendiyeva, M. A., Rowe, H. D., Feyzullayev, A. A., and Aliyev, C. S., 2008, Stratigraphy and geochemical characterization of the Oligocene-Miocene Maikop Series: Implications for the paleogeography of eastern Azerbaijan: Tectonophysics, v. 451, p. 40–55.

Hurst, A., 1990, Natural gamma-ray spectrometry in hydrocarbon-bearing sandstones from the Norwegian Continental Shelf: in, A. Hurst, M.A. Lovell, A.C. Morton, eds, Geological Applications of Wireline Logs, Geological Society, Special Publications, No. 48, p. 211 – 222.

Knapp, J. H., C. C. D. Knapp, J. A. Connor, J. H. McBride, and M. D. Simmons, 2007, Deep seismic exploration of the South Caspian Basin: Lithosphere-scale imaging of the world's deepest basin: in, P. O. Yilmaz and G. H. Isaksen, eds, Oil and gas of the Greater Caspian area: AAPG Studies in Geology 55, p. 47-49.

Kroonenberg, S. B., Simmons, M. D., Alekseevski, N. I., Aliyeva, E., Allen, M. B., Aybulator, D. N., Baba-Zadeh, A., Badyukova, E. N., Davies, C., Hinds, D. J., Hoogendoorn, R. M., Huseynov, D., Ibrahimov, B., Mamedov, P., Overeem, I., Rusakov, G. V., Suleymanova, S., Svitoch, A. A. and Vincent, S. J., 2005, Two deltas, two basins, one river, one sea. The modern Volga Delta as an analogue of the Neogene Productive Series, South Caspian Basin: River Deltas - Concepts, Models and Examples, 83. Society of Economic Paleontologists and Mineralogists, Tulsa, p. 231-256

Meyer, F.J., 2010, University of Alaska Fairbanks, Statistics and Data Analysis in Geology lecture notes on Cluster Analysis http://avo-ftp.images.alaska.edu/TEMP/geos430_geostats/updated_files_2011/Lecture_pdf_2011/GeoStats_1.pdf

Morton, A., Allen, M., Simmons, M., Spathopoulos, F., Still, J., Hinds, D., Zadeh, A.I. and Kroonenberg, S., 2003, Provenance patterns in a neotectonic basin: Pliocene and Quaternary sediment supply to the South Caspian: Basin Research, vol.15,p.321-337

Narimanov, A. A., and Palaz, I., 1995, Oil history, potential converge in Azerbaijan: Oil & Gas Journal, May 22, p. 32–39.

Parks, J.M., 1966, Cluster analysis applied to multivariate geologic problems: Journal of Geology, vol.74, p. 703-715

Pirkle, F.L., Howell, J.A., Wecksung, G.W., Duran, B.S., and Stablein, N.K., 1984, An example of cluster analysis applied to a large geologic data set: Aerial radiometric data from Copper Mountain Wyoming: Mathematical Geology, vol. 16, no. 5, p.479-498

Reynolds, A. D., Simmons, M. D., Bowman, M. B. J., Henton, J., Brayshaw, A. C., Ali-Zade, A. A., Guliyev, I. S., Suleymanova, S. F., Ateava, E. Z., Mamedova, D. N., and Koshkarly, R. O.,1998, Implications of outcrop geology for reservoirs in the Neogene Productive Series: Apsheron Peninsula, Azerbaijan: AAPG Bulletin, vol. 82, p 25–49.

Rider, M., 1996, Geological Interpretation of Well Logs, 2nd edition, Whittles Publishing, Caitheness, 428p.

Rohlf, F.J., 2005, NTSYSpc Numerical Taxonomy and Multivariate Analysis System Version 2.2 Getting Started Guide: Exeter Software, E. Setauket, New York, 42p. www.exetersoftware.com/cat/ntsyspc/ntsyspc.html

Smith-Rouch, L.S., 2006, Oligocene–Miocene Maykop/Diatom Total Petroleum System of the South Caspian Sea Basin Province, Azerbaijan, Iran, and Turkmenistan: U.S. Geological Survey Bulletin 2201-I, 27 p.

Templ, M., Filzmoser, P., and Clemens, R., 2008, Cluster analysis applied to regional geochemical data: Problems and possibilities: Applied Geochemistry, vol. 23, p 2198-2213

US Geological Survey, 2010, World Petroleum Resources Project Fact Sheet, Assessment of Undiscovered Oil and Gas Resources of the North Caspian Basin, Middle Caspian Basin, North Ustyurt Basin, and South Caspian Basin Provinces, Caspian Sea Area, USGS, 2010, 4p.

Weltje, G.J., and Eynatten, H.V., 2004, Quantitative provenance analysis of sediments: review and outlook: Sedimentary Geology, vol 171, p1-11

Wethington, W.B., Reddick, C., and Husseynov, B., 2002, Development of a super giant, ACG Field, Offshore Azerbaijan: An overview: in, AAPG Annual Convention, Houston, Texas, AAPG Search and Discovery Article #90007, 8p.

Vincent, S.J., Davies, C.E., Richards, K. and Aliyeva, E., 2010, Contrasting Pliocene fluvial depositional systems within the rapidly subsiding South Caspian Basin; A case study of the palaeo-Volga and palaeo-Kura river systems in the Surakhany Suite, Upper Productive Series, onshore Azerbaijan: Marine and Petroleum Geology, vol.27, p.2079-2106