

**RESERVOIR CHARACTERIZATION USING WAVELET
TRANSFORMS**

A Dissertation

by

NESTOR RIVERA VEGA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2003

Major Subject: Petroleum Engineering

**RESERVOIR CHARACTERIZATION USING WAVELET
TRANSFORMS**

A Dissertation

by

NESTOR RIVERA VEGA

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

Jerry L. Jensen
(Chair of Committee)

Robert R. Berg
(Member)

Akhil Datta-Gupta
(Member)

Andrew K. Chan
(Member)

Hans C. Juvkam-Wold
(Head of Department)

December 2003

Major Subject: Petroleum Engineering

ABSTRACT

Reservoir Characterization Using Wavelet Transforms. (December 2003)

Nestor Rivera Vega, B.S., Universidad Industrial de Santander;

M.S., The University of Texas at Austin

Chair of Advisory Committee: Dr. Jerry L. Jensen

Automated detection of geological boundaries and determination of cyclic events controlling deposition can facilitate stratigraphic analysis and reservoir characterization. This study applies the wavelet transformation, a recent advance in signal analysis techniques, to interpret cyclicity, determine its controlling factors, and detect zone boundaries. We tested the cyclostratigraphic assessments using well log and core data from a well in a fluvio-eolian sequence in the Ormskirk Sandstone, Irish Sea. The boundary detection technique was tested using log data from 10 wells in the Apiay field, Colombia.

We processed the wavelet coefficients for each zone of the Ormskirk Formation and determined the wavelengths of the strongest cyclicities. Comparing these periodicities with Milankovitch cycles, we found a strong correspondence of the two. This suggests that climate exercised an important control on depositional cyclicity, as had been concluded in previous studies of the Ormskirk Sandstone.

The wavelet coefficients from the log data in the Apiay field were combined to form features. These vectors were used in conjunction with pattern recognition techniques to perform detection in 7 boundaries. For the upper two units, the boundary was detected within 10 feet of their actual depth, in 90% of the wells. The mean detection performance in the Apiay field is 50%.

We compared our method with other traditional techniques which do not focus on selecting optimal features for boundary identification. Those methods resulted in

detection performances of 40% for the uppermost boundary, which lag behind the 90% performance of our method.

Automated determination of geologic boundaries will expedite studies, and knowledge of the controlling deposition factors will enhance stratigraphic and reservoir characterization models. We expect that automated boundary detection and cyclicity analysis will prove to be valuable and time-saving methods for establishing correlations and their uncertainties in many types of oil and gas reservoirs, thus facilitating reservoir exploration and management.

DEDICATION

To my beloved wife and to the memory of my father and mother

ACKNOWLEDGMENTS

I gratefully acknowledge the time, patience, and guidance of my advisor, Dr. Jerry Jensen, during the development of this research. His desire to implement recent engineering techniques in the oil industry, and to integrate diverse disciplines such as electrical engineering, computer science, geology, and petroleum engineering were important factors in the successful application of the new methods proposed.

I thank Cesar Vasquez of Ecopetrol, Colombia for his agreement to use the data from the Apiay field for my investigation. I greatly appreciate the time and effort devoted by Norman Yepes of Ecopetrol to gather and send the data required for this study.

Thanks to Dr. Andrew Chan for teaching me the fundamentals of wavelets. I also want to express my gratitude to Dr. Walter Ayers for his suggestions and comments on the geologic aspects of my research. Thanks to Dr. Datta Gupta for his suggestions to improve the techniques developed. Also, thanks to Dr. Robert Berg for his comments on the applications of my research.

Thanks to my colleague Jorge Nieto for his illustration of important issues of the reservoir characterization of the Apiay field.

I want to express my deepest gratitude to my wife Stella. Her support and encouragement greatly contributed to my finishing successfully this doctoral program. To my children, Javier and Julian, for their love and patience.

Finally, I wish to thank Dr. Jensen and the A&M petroleum engineering department for the financial support, which was a main factor in achieving my dream of obtaining a doctoral degree in petroleum engineering from one of the finest institutions.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
DEDICATION	v
ACKNOWLEDGMENTS.....	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES.....	ix
LIST OF TABLES	xiii
CHAPTER	
I INTRODUCTION	1
1.1 Statement of the problem.....	1
1.2 Objectives and methods.....	3
1.3 Organization of the dissertation.....	4
II LITERATURE REVIEW	5
2.1 Well-log data analysis.....	5
2.2 Signal processing methods	6
2.3 Other wavelet methods in oilfield applications	7
III WELL-LOG PROCESSING METHODS	8
3.1 Cyclicity detection	8
3.1.1 Fourier analysis	9
3.1.2 Wavelet analysis.....	13
3.1.3 Spatial statistics	16
3.1.4 Comparison of cyclicity detection methods	17
3.2 Multiresolution analysis.....	20
3.3 Principal component analysis	22
IV CYCLOSTRATIGRAPHY	24
4.1 Reasons for geological cycles.....	24
4.2 Irish Sea example.....	25
4.2.1 Wavelet analysis.....	27
4.2.2 Cyclostratigraphic interpretation.....	28
4.3 Discussion.....	32
4.4 Conclusions.....	33
V MULTIPLE-WELL CORRELATIONS.....	35
5.1 Apiay field	37

CHAPTER	Page
5.2 Well-log response to boundaries	39
5.2.1 T1/K1U1 boundary	42
5.2.2 K1L2/K2 boundary	43
5.3 Boundary recognition	43
5.3.1 Pre-processing	45
5.3.2 Feature extraction	46
5.3.2.1 Definition of features.....	47
5.3.2.2 Window length for feature extraction.....	51
5.3.3 Evaluation of features.....	52
5.3.4 Dimensionality reduction	54
5.3.5 Depth selection	61
5.3.5.1 Single-well training set.....	64
5.3.5.2 Multiple-well training set	68
5.4 Effect of number of signals.....	77
5.5 Effect of wavelet decomposition	78
5.5.1 Discrete wavelet transform (dwt).....	79
5.5.2 Continuous wavelet transform (cwt).....	84
5.6 Choice of signals for boundary detection in Apiay area fields.....	88
5.7 Comparison with other techniques	90
5.8 Conclusions.....	92
VI SUMMARY AND CONCLUSIONS	95
VII FUTURE WORK.....	97
NOMENCLATURE.....	98
REFERENCES.....	99
APPENDIX A	104
APPENDIX B	111
VITA	115

LIST OF FIGURES

FIGURE	Page
3.1 Fourier analysis of a sinusoidal signal with random noise.....	11
3.2 Continuous wavelet transform using the Morlet wavelet.....	15
3.3 Wavelet and Fourier analysis of a small discontinuity.....	16
3.4 Signal analysis results for an abrupt change in frequency, e.g. parasequence boundary.....	18
3.5 Signal analysis results for two superimposed sedimentary cycles (e.g. Milankovitch cyclicity).	19
3.6 Multiresolution analysis indicating a) the filtering process and b) the multiresolution diagram.	21
4.1 Well log, core permeability and Morlet wavelet scaleograms for GR and LLD logs, well 110/8a-5.	28
4.2 Wavelet spectra for zone 1 (Sabkha).....	30
4.3 Wavelet spectra for zone 2 (Dune).....	30
4.4 Wavelet spectra for zone 3 (Fluvial).	31
4.5 Wavelet spectra for all zones combined.	31
5.1 Area of Apiay field and wells analyzed in this study.....	38
5.2 Stratigraphic cross section A-A' (NE-SW) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are GR and R _t	40
5.3 Stratigraphic cross section B-B' (NE-SW) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are GR and R _t	41
5.4 GR-features at the T2/K1U1 boundary for well A-04, using a window length $w = 40$ ft.....	45
5.5 GR-trend angle at the T2/K1U1 boundary for well A-04, using a window length $w = 40$ ft.....	50
5.6 GR- wavelet amplitude as a function of wavelength in wells A-03 and A-09.....	52
5.7 Features from GR for T2/K1U1 boundary; window size 40 ft. Dots represent lower window features and crosses upper window features.	53

FIGURE	Page
5.8 Dimensionality reduction process. The example for 5 well logs, window size of 40 ft, 2 samples/ft, and four features per log, demonstrates how the 400-dimensional input matrix is reduced to 20 dimensions with feature extraction; then to 10 dimensions by selection of optimal features; and lastly to 2 dimensions by PCA.	56
5.9 PCA for the GR at T2/K1U1 boundary, window size = 40 ft (red crosses upper, blue circles lower part).	60
5.10 Feature extraction in observation well. Features are extracted at each depth point using the window length, w . The window is shifted along the depth range. The boundary is assigned to the depth of the midpoint of the window with the features closest to those of the training set.	63
5.11 Error distribution for well A-03 for the T2/K1U1 boundary detection, using the raw GR. The training well is A-15 (Example 1). The features shown are: (a) upper mean and (b) lower C_V	65
5.12 Normalized probability distributions for the T2/K1U1 boundary detection, using the raw GR. The training well is A-15. The true boundary depths are indicated with a tick on the right hand side.	67
5.13 Feature vectors for raw signals in Apiay field at the K1L2/K2 boundary. Dots represent lower window features and crosses upper window features.	69
5.14 Principal components for different combinations of features for the raw signals in Apiay field at the K1/K2 boundary. (d) represents the optimal features from the 5-signal 20 features. In this case, the optimal features are GR-mean and GR-trend angle.	71
5.15 Differential-features-coefficient of variation, dC_V , of the raw-signal features in Apiay field at the K1/K2 boundary. For each signal the features from left to right are: mean, CV, Min/Max, trend angle.	73
5.16 Differential-features-coefficient of variation, dC_V , of the raw-signal features for 4-training wells (A-05, A-11, A-13, A-15) at the K1/K2 boundary. For each signal the features from left to right are: mean, C_V , Min/Max, trend angle.	74
5.17 Differential-features-coefficient of variation, dC_V , and Mahalanobis distance variance for the training set A-05, A-11, A-13, A-15 at the K1/K2 boundary. The features are ordered in ascending dC_V values. The optimal number of features is 12, although 10 or 11 also give good performance.	75

FIGURE	Page
5.18 Normalized probability distributions for the K1/K2 boundary detection, using 5 well logs. The training well are A-05, A-11, A-13, A-15 A-15. The true boundary depths are indicated with a tick on the right hand side. (a) shows the distribution using all 20 features and (b) displays the distribution using the optimal number of features (12).....	76
5.19 dwt on the GR of well A-15, using D4 and 4 levels of decomposition. Shown are the signal, the approximation and the 4 details.	80
5.20 Features for the GR, approximation, and details at the K1L2/K2 boundary in Apiay.....	81
5.21 Differential-features-coefficient of variation, dC_V , for the GR, approximation and 4 levels of detail at the K1/K2 boundary, using all wells in Apiay. For each signal the features from left to right are: mean, CV, Min/Max, trend angle.	82
5.22 Differential-features-coefficient of variation, dC_V , and Mahalanobis distance variance for all wells at the K1/K2 boundary, using the GR and 4 levels of decomposition. The features are ordered in ascending dC_V values. The optimal number of features is 6.....	83
5.23 Morlet scaleogram for well A-03. The cwt is applied for the GR and R_t	85
5.24 cwt features for the GR at the K1L2/K2 boundary. The first row shows the 4 features for the raw signal. The other rows show the mean and CV for the wavelets coefficients at different wavelengths.	86
5.25 Differential-features-coefficient of variation, dC_V , for the GR, and wavelength coefficients, for various wavelengths, at the K1/K2 boundary, using all wells in Apiay. For each signal the features from left to right are: mean, CV, Min/Max, trend angle.....	87
A.1 Paleocurrent vectors for cross-stratified fluvial sandstones (from Herries and Cowan, 1997).	109
A.2 Well correlation for the East Irish Sea Basin (from Thompson & Meadows, 1997).	110
B.1 Stratigraphic cross section (N-S) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are GR and R_t	111
B.2 Stratigraphic cross section A-A' (NE-SW) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are ρ_{in1} and ρ_b	112
B.3 Stratigraphic cross section B-B' (NW-SE) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are ρ_{in1} and ρ_b	113

FIGURE	Page
B.4 Stratigraphic cross section (N-S) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are Phinl and ρ_b	114

LIST OF TABLES

TABLE	Page
4.1	Comparison of cyclic frequencies of sequence-stratigraphic units with Milankovitch cycles (after Mitchum and Van Wagoner, 1991).....26
4.2	Wavelengths and ratios in well 110/8a-5.32
5.1	Features for raw GR at the true K1U1 boundary for a 40-ft window.57
5.2	Normalized features for raw GR at the true K1U1 boundary for a 40-ft window.58
5.3	Covariance matrix of standardized feature vectors.58
5.4	Eigenvalues and Eigenvectors of covariance matrix in order of decreasing eigenvalues. The Eigenvectors represent the coefficient of the PCs.....59
5.5	Z-score transform of the feature vectors (PCs).60
5.6	Proportion of boundaries identified within 10 ft of the true position, using A-15 as the training.68
5.7	Detection performance and mean detection error at the K1L2/K2 boundary. Training wells: A-05, A-11, A-13, A-15.....78
5.8	Detection performance and mean detection error at the K1L2/K2 boundary, using dwt on the well-logs. Training wells: A-05, A-11, A-13, A-15.84
5.9	Detection performance and mean detection error at the K1L2/K2 boundary, using cwt on the well-logs. Training wells: A-05, A-11, A-13, A-15.87
5.10	Detection performance for the T2/K1U1 boundary in Apiay.89
5.11	Optimal boundary detection for Apiay.....90
5.12	Comparison of new method with traditional DA techniques.92

CHAPTER I

INTRODUCTION

1.1 Statement of the problem

The characterization of petroleum reservoirs involves data analysis to generate a 3-dimensional (3D) reservoir framework describing rock properties and fluid parameters. The resulting model is used to quantify hydrocarbons and determine the development scheme appropriate for the reservoir. These integrated studies involve various disciplines using different data types. The most comprehensive data sets include well logs, core analyses, 3D seismic data, hydrocarbon composition, and pressure and production tests. Core analysis and 3D seismic play a key role in the identification and correlation of genetic stratigraphic units. However, in practice, only a small percentage of wells are cored. In addition, 3D seismic data were not recorded for the great majority of older fields. On the other hand, well logs are usually run on every well. If a technique could identify and correlate stratigraphic units based only on well log responses, it would greatly contribute to improved reservoir characterization for fields with limited core analysis data and lacking 3D seismic information.

To accomplish this difficult task, well log responses should be optimally processed and transformed. Most well log responses contain numerous nonstationary or transitory characteristics, including cyclicities, trends, and abrupt changes. Recent developments in signal analysis indicate that wavelet transforms are an analysis tool well suited to multiscale, nonstationary processes. Well-log wavelet transformations can identify cyclicities and important events on individual wells. By using pattern recognition techniques, improved multiple-well stratigraphic correlations can be developed by identifying the formation tops and unconformities present on each well.

This dissertation follows the style and format of *The Log Analyst*.

This approach would generate the multiple-well correlations automatically. This would be an improvement compared to the traditional correlations, using visual inspection and unprocessed signals.

The most widely automated well-log correlation methods adopt the algorithms used in dipmeter processing. The measure of similarity between two log traces is the standard correlation coefficient. A segment (window length) of the first log trace is moved by small increments (step length) past the trace of the second log. The shifting is performed for a pre-set search range (search length). At each step the correlation coefficient is computed as a function of depth in a correlogram. A perfect match of the two traces is achieved at a positive correlation of unity. Lateral changes in thickness degrade the correlation coefficient slightly. Major changes generate problems that are tackled by selective stretching of one of the logs (Doveton, 1994, Chapter 6).

In our research, we applied the wavelet transformation and extracted statistical features for a window at the boundary of interest. We found that the optimal window is 40 ft, 20 ft above the boundary and 20 ft below the boundary. The boundary-feature matrix may consist of a single or multiple wells in the training set. Then, we proceeded to select the optimal features and perform dimensionality reduction using principal component analysis (PCA). Finally, using step length of 1ft and search length of 200 ft, we computed probabilities to determine the boundary depth in the observation well.

Other attributes that can be extracted from wavelet transformations, and closely related to stratigraphic units, are the determination of factors controlling deposition (Prokoph and Agterberg, 1999 and 2000), sedimentation rates and preservation potential at higher resolution than other techniques such as biostratigraphy, chronostratigraphy, or sequence stratigraphy (Prokoph and Agterberg, 2000).

To test these methods, we used two data sets, one set from a gas well in the Irish Sea and another from a prolific oil field in Colombia. For the Irish Sea, we found strong indications of Milankovitch cycles during deposition. For Apiay, our approach resulted in boundary detection performance of up to 90% (the boundary was detected within 10 ft of their actual depth, in 90% of the wells).

1.2 Objectives and methods

The main objectives of this research were the following:

1. Cyclostratigraphy: Apply wavelet decomposition of well-logs and assess the wavelengths of the strongest cyclicities to determine if wavelets can aid cyclostratigraphic analysis.
2. Boundary detection: Apply wavelet analysis of well-logs to the detection of boundaries to assess the value of the wavelet analysis.
3. Validate the cyclostratigraphic procedure with one well in the Irish Sea.
4. Evaluate the boundary detection algorithm using 10 wells in the Apiay field.

All computer programs were written in Matlab. The main advantages of Matlab are: **1)** is a matrix oriented programming language, which is very convenient for our multi-dimensional problem. **2)** Uses many built-in functions saving considerable time in the computer program designs. This a very useful for the wavelet and statistical functions. **3)** The graphics options are extensive.

The following approach was used in our research:

1. Literature review: We reviewed current techniques for cyclostratigraphy and multiple-well correlations using well-logs; evaluated the usage of pattern recognition and wavelet transformation for well-logs; studied signal processing techniques for 1-D signals; and examined analogous applications in other fields, such as computer science and electrical engineering.
2. Knowledge acquisition: We studied and developed computer programs for Fourier analysis, wavelet transformation, and pattern recognition for general 1-D and 2-D applications. We also researched the techniques for multidimensionality reduction.
3. Data collection: We gathered data from one well in the Irish Sea and 10 wells from the Apiay field, Colombia.

4. Choice of wavelets: We selected various wells to compare the results of different choices of wavelet families.
5. Pattern recognition: We tested different approaches for pattern recognition, feature selection, and dimensionality reduction. This phase helped us on the selection of the best techniques to tackle our specific problem.
6. Model validation. After selecting the best techniques for boundary detection, we did a comprehensive validation using different combinations of well-logs and type of wavelet transformation. The validation was performed for all 7 boundaries in Apiay.

1.3 Organization of the dissertation

Chapter I is the introduction. Chapter II presents the literature review on signal processing techniques, cyclostratigraphy, multiple-well correlation, and pattern recognition.

Chapter III discusses the signal processing techniques and compares the different methods. Chapter IV covers cyclostratigraphy for one well in the Irish Sea. Chapter V develops the method for multiple-well correlation using the Apiay field. Chapter VI presents the summary and conclusions, and Chapter VII proposes ideas for future work.

CHAPTER II

LITERATURE REVIEW

2.1 Well-log data analysis

The use of well-log data for sedimentology and stratigraphy has been explored extensively in the last two decades. In most wells, usually two or more logs are recorded covering the entire interval of interest at a very high sampling rate (commonly 2 samples per foot). These advantages have led to various geological applications. A widely used method is the study of sedimentary successions.

The work by Serra and Abbot (1982) introduced the concept of electrofacies, which “is primarily a method of describing rock in terms of its log characteristics”. Serra and Abbot treated each log as a vector component of the n -dimensional set of n distinct logs. They proposed manual and automatic methods and the main objectives included the recognition of principal electrofacies traversed by a well, identification of abrupt changes in sedimentation and sedimentary markers, well-to-well correlation, and fine-layer analysis to derive layer-by-layer permeability. Electrofacies analysis is intended to identify zones with similar petrophysical characteristics.

Many studies have continued to use the concept of electrofacies for automatic facies estimation. Supervised and unsupervised techniques using neural networks have been proposed (Saggaf and Nebrija, 2000). For permeability correlation, techniques such as principal component analysis (PCA) and cluster analysis have been utilized (Lee et al., 2002).

Doveton (1994, Chapter 6) presented a historical perspective of lateral correlation and interpolation of logs. Most methods of automated correlation of wireline logs use a single feature to compare the similarity between two well-logs. This feature is the standard correlation coefficient operating on a pre-set window. Lateral correlation is considered a difficult task which can be made more consistent if interpreters use the same criteria.

The task of cyclicity determination has been approached by semivariograms, Fourier analysis and more recently by wavelet transformation. Jennings et al. (2000) and Jensen et al. (2000) have used the semivariogram (SV) of petrophysical data to study periodicities. The SV determines the degree of similarity between sample pairs as a function of separation distance. SV's can also be employed to detect cyclity (Jennings et al., 2000). However, as in the case of the Fourier transform, the localization of the cyclic events in space is not possible.

Prokoph and Agterberg (2000) performed Morlet wavelet analysis to gamma-ray well logs to locate discontinuities and determine high frequency sedimentary cycles. They found a correlation of the ratio of the predominant gamma ray cycles with the ratio of Milankovitch spectra, suggesting that climatic cycles were important factor in the deposition of the Egret member, offshore Canada.

2.2 Signal processing methods

Most well-log responses contain numerous nonstationary or transitory characteristics, including cyclicities, trends, and abrupt changes. These characteristics are often the most important part of the signal. For many applications, the best representation of transient signals is in the frequency domain by using spectral analysis methods. However, the most common methods for well-log data analysis operate on the space domain.

The Fourier transform has been the traditional method for spectral analysis. However, the introduction of wavelets in the 1980s generated new ways to do signal processing. Wavelet analysis is more effective in representing nonstationary signals. Thus, wavelets have gained popularity and new applications are being investigated (Goswami and Chan, 1999).

Wavelet transformation allows multiresolution analysis, in which a signal can be represented by a sum of components at different resolutions. The capability of representing a signal in several levels of resolution is the major strength of wavelet analysis (Goswami and Chan, 1999). Features can be extracted for each one of the

resolutions and be used for pattern recognition. Goswami and Chan (1999) describe one example using acoustic signals (1-D) to identify several faults in the wheel bearings of railroad cars. Goumas et al. (2002) used similar approach to classify washing machines vibration signals. They performed the wavelet decomposition, extracted statistical features from the various levels of resolution, combined the features, performed dimensionality reduction using PCA, and then proceed to the classification.

Wavelet analysis has been applied to detect cyclicity in climate time series (Lau and Weng, 1995).

2.3 Other wavelet methods in oilfield applications

Other wavelet transform applications in the oil industry include denoising and conditioning of well pressure data (Athichanagorn et al. (1999) and Gonzalez et al. (1999), upscaling of rock properties (Panda et al., 2000), seismic data processing (Chakraborty and Okaya, 1995), measurement-while-drilling data compression to be transmitted through mud-pulse telemetry (Bernasconi et al., 1999), and identification of wellbore and reservoir anomalies from pressure transient data (Soliman et al., 2001).

We did not find in the literature wavelet decomposition methods for boundary identification purposes or multiple-well correlation. However, wavelets have been applied to production data to estimate the degree of interference, preferential flow paths and the existence of flow barriers (Jansen and Kelkar, 1997). Ours is a novelty approach in reservoir characterization.

CHAPTER III

WELL-LOG PROCESSING METHODS

In this chapter we review the aspects of signal processing and classification methods related to cyclicity and boundary detection. For cyclicity detection, we cover Fourier analysis, wavelet transform, and the semivariogram. We illustrate some examples to compare the methods.

The wavelet analysis is performed in two ways. For cyclicity determination, we use the coefficients from the continuous wavelet transform, and for boundary detection, in addition to the processing of these coefficients, we perform multiresolution analysis based on the discrete wavelet transform.

Lastly, we cover principal component analysis. We use this dimensionality-reduction technique to perform the pattern recognition required in boundary detection. Principal component analysis has been used by other researchers to define electrofacies from well-logs and predict permeability (Lee et al., 2002); and analysis of lithofacies and shale volume estimation (Doveton, 1994, Chapter 4).

3.1 Cyclicity detection

The oil industry has applied different techniques to quantify cyclicity. The Fourier transform has been a natural choice for cyclicity assessment because of its familiar basis functions and its widespread use for such analysis (e.g., Box and Jenkins, 1976, Chap. 2). Despite advantages to analysis in the spectral domain (e.g., Gelhar, 1993, Chap. 2), some authors (Jennings et al., 2000; Jensen et al., 2000) have preferred the spatial domain, using the semivariogram (SV) to study periodicities. The SV determines the degree of similarity between sample pairs as a function of separation distance. As we will see, the localization of the cyclic events in space is not possible with either method. Wavelet analysis is a more recent tool used to detect cyclicity, e.g., in climate time series (Lau and Weng, 1995) and sedimentary cyclicities (Prokoph and Barthelmes, 1996). In particular, Prokoph and Agterberg applied Morlet wavelet analysis to gamma-ray well

logs to locate discontinuities and determine high-frequency sedimentary cycles. Wavelets do provide localization in space, which is often needed for cyclostratigraphic analysis. We briefly review the properties of all three methods below.

3.1.1 Fourier analysis

Spectral analysis methods can be used to help in the interpretation of well logs. The best known method of spectral analysis is the Fourier transform (e.g., Paupolis, 1962). The Fourier transform of a finite-energy function $f(t)$ of a real variable t is defined by the integral

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \quad (3.1)$$

where ω is the fundamental frequency, and $j = \sqrt{-1}$.

To maintain the convention used in signal processing, we use t to represent time or space. A typical well-log is a function of depth, and the most common sampling rate, F_s , is 2 samples/foot. This customary number is governed by the resolution and limitations of logging tools. By using the sampling theorem (Proakis and Manolakis, 1996, Chapter 1), the maximum frequency F_{max} that can be deciphered, is given by $F_{max} = F_s/2$. The period of the cycles is the reciprocal of the frequency. For space domain the period in feet/cycle is known as the wavelength, wl . Thus for $F_s = 2$ samples/ft, the minimum wavelength becomes 1 ft. However, considering the limitations of the tools, we consider that cyclicities from well-logs will be confident for wavelengths larger than 3 ft. Other data sets, such as probe permeameter, with sampling rates of up to 20 samples/ft, can detect cycles as short as a few inches.

Since Equation 3.1 requires integration, it is appropriate for functions described analytically. Well-logs are sampled signals (discrete in space) and require numerical algorithms to compute the spectrum. The discrete variable $f(n)$, where n is an integer, is discretized in space while the frequency variable (ω) is continuous. The discrete Fourier transform is given by (Goswami and Chan, 1999, Chapter 3)

$$\hat{f}(\omega) = \sum_{n=-\infty}^{\infty} f(n)e^{-jn\omega} \quad (3.2)$$

The inverse Fourier transform recovers $f(n)$ by

$$f(n) = \int_{-\pi}^{\pi} \hat{f}(\omega)e^{jn\omega} d\omega \quad (3.3)$$

The frequency vector, ω , in cycles per foot is given by

$$\omega = \frac{n}{N} F_s \quad (3.4)$$

where N is the length (total number of samples) of $f(n)$. The wavelength variable is calculated as

$$wl = \frac{1}{\omega} \quad (3.5)$$

Two uses of the Fourier transform for well-logs are cyclicity and filtering. To demonstrate the steps and use of Fourier transform for well logs we selected a signal with added noise (Example 3.1a). The sine function has a wavelength $wl = 40$ ft with a sampling rate $F_s = 1$ sample/ft. The total number of samples is $N = 256$ (Figure 3.1). The following are the steps for Fourier analysis on example 3.1:

- 1) Calculate the frequency vector ω using equation 3.4. The Fourier coefficients are symmetric about $\omega(N/2)$. Thus, the length of ω is $N/2$.
- 2) Compute the Fourier transform using equation 3.2. This generates 128 coefficients.
- 3) Calculate the magnitude of the Fourier coefficients, $|\hat{f}(\omega)|$. In general the coefficients are complex, and the magnitude is given by the norm. For example the magnitude of the coefficient $3 + 4j$ is 5.
- 4) Plot the Fourier coefficients as a function of frequency (Figure 3.1b). The largest magnitude is at a frequency of 0.027 cycles/ft, equivalent to $wl = 37$ ft (very close to the original sinusoid).

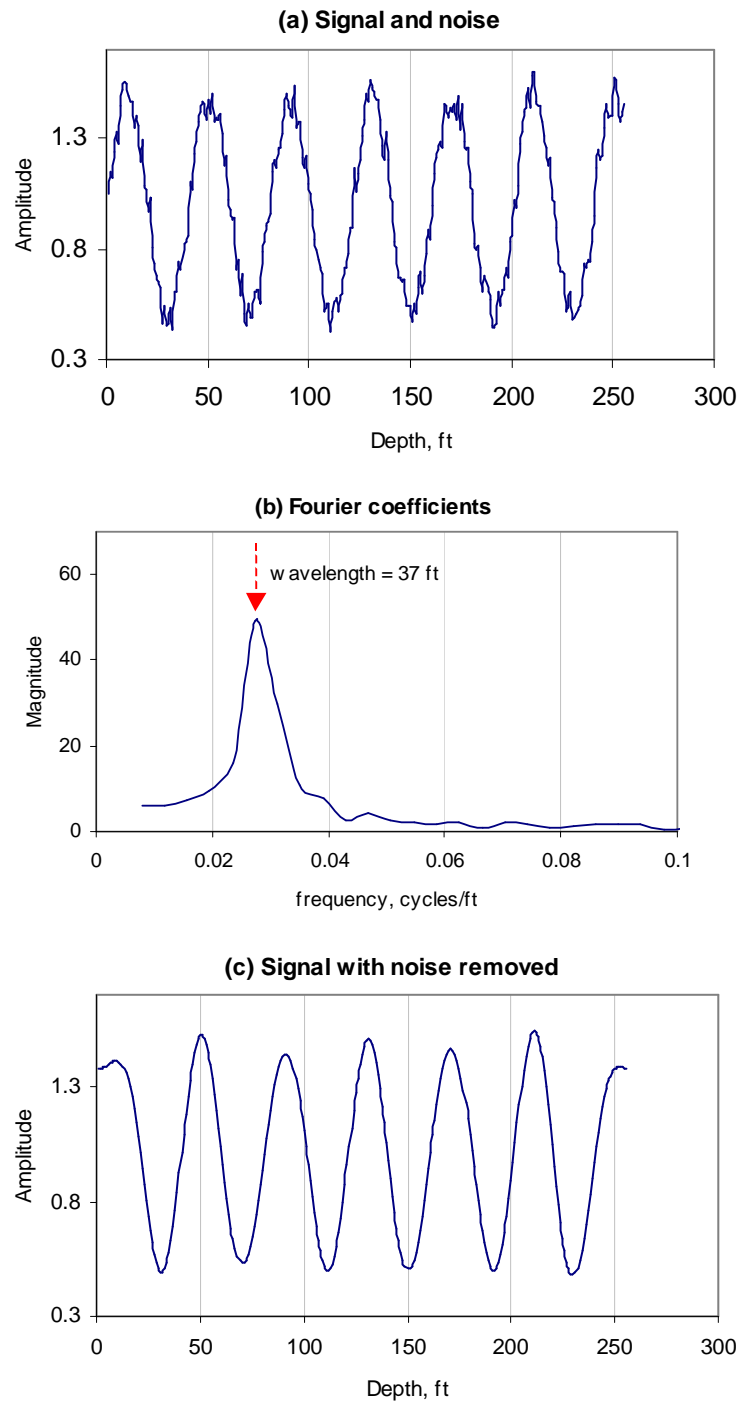


Figure 3.1 Fourier analysis of a sinusoidal signal with random noise.

- 5) The noise has high frequency. To remove the noise keep the low frequency Fourier coefficients (low pass filter). Here, we keep $\hat{f}(\omega)$ for $1 \leq n \leq 10$ and set $\hat{f}(\omega) = 0$ for $11 \leq n \leq 128$. Apply the inverse Fourier transform (Eq. 3.3) to the filtered $\hat{f}(\omega)$. The resulting filtered signal is given in Figure 3.1c.

The discrete Fourier transform (DFT) plays a key role in digital signal processing, including filtering, correlation analysis, and spectrum analysis. The computation of the DFT involves N^2 operations. The fast Fourier transform (FFT) algorithm takes advantage of the symmetry and periodicity of the Fourier matrix, to reduce the complex multiplications to $N \log_2 N$ when N is a power of 2. Any sequence can be augmented with zeros, such that $N = 2^p$, where p is an integer (Proakis and Manolakis, 1996, Chapter 6).

The Fourier transform has the limitation that it can be evaluated at only one frequency at a time, i.e. the Fourier spectrum does not provide any spatial-domain information about the signal. When looking at a Fourier transform it is impossible to tell when a particular event took place. If the signal does not change much over time (stationary signal), this drawback is not important. However, most well-log responses contain numerous non-stationary or transitory characteristics: drift; trends; abrupt changes; and beginnings and ends of events. These characteristics are often the most important part of the signal, and Fourier analysis is unable to detect them. The short-term Fourier transform solves this deficiency by windowing the signal into a two-dimensional representation of space and frequency (Goswami and Chan, 1999, Chapter 4). The information obtained is a function of the size of the window. The drawback of the short-time Fourier transform is that, once a particular window size is chosen, that window is the same for all frequencies. A large window provides good frequency resolution but a poor spatial resolution, whereas a shorter window provides a better spatial resolution with poor frequency resolution. The fixed window size of the short-time Fourier transform is a disadvantage if the signal contains spectral characteristics at different scales.

3.1.2 Wavelet analysis

Wavelet analysis represents an advance from Fourier analysis in the form of a windowing technique with flexible regions. The wavelet transform uses a window function whose radius increases in space (reduces in frequency) while resolving the low-frequency contents of a signal (Goswami and Chan, 1999, Chapter 4).

The integral (continuous) wavelet transform of a function $f(t)$ with respect to some local base function (wavelet) ψ is defined as

$$W_{\psi} f(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad a > 0 \quad (3.6)$$

where ψ^* is the complex conjugate of ψ . The parameters b and a are called translation (shifting) and dilation parameters, respectively. The wavelet behaves like a window function. At any scale a , the wavelet coefficients $W_{\psi a, b}$ are obtained by convolving $f(t)$ and a dilated and translated version of the wavelet. To be a window and to recover $f(t)$ from its inverse wavelet transform (IWT), $\psi(t)$ must satisfy

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (3.7)$$

The continuous wavelet transform (CWT) provides space-scale analysis and not space-frequency analysis. However, proper scale-to-frequency transformation allows analysis that is very close to space-frequency analysis. By reducing the scaling parameter a , the support of the wavelet is reduced in space and hence covers higher frequencies ranges and vice versa. Therefore, $1/a$ is a measure of frequency. The parameter b indicates the location of the wavelet window along the space axis. Thus, by changing (b, a) , the wavelet coefficients $W_{\psi} f(b, a)$ can be computed on the entire space-frequency plane. A required condition is that all wavelets must oscillate, giving them the nature of small waves and hence the name wavelets. The wavelet transform (WT) is an analysis tool well suited to the study of multiscale, nonstationary processes occurring over finite spatial and temporal domains.

The CWT separates out the frequency components of a signal. It is therefore important that the wavelet used gives the best resolution in frequency. The shape of the wavelet coefficients at some scale should resemble a sinusoid at the corresponding pure frequencies. The best wavelet for this purpose is the Morlet wavelet with its Gaussian modulated complex decaying exponential. The graphical representation of the wavelet coefficients for the different scales (wavelengths) as a function of depth is the scaleogram.

One of the most widely used continuous wavelets in geophysics and petrophysics is the complex Morlet wavelet, which consists of a plane wave modified by a Gaussian envelope, given by

$$\psi(t) = e^{i^2/2} e^{-j5.336t} \quad (3.8)$$

Figure 3.2 shows how the wavelet coefficients are processed, using the Morlet wavelet. The combination of the various vectors of coefficients at different scales (wavelengths) form the scaleogram. The depths with the strongest coefficients indicate the position where that particular wavelength is taking place. The scaleogram provides a good space-frequency representation of the signal. The conversion from scale to wavelength for the Morlet wavelet is given by

$$wl = \frac{1.25a}{F_s} \quad (3.9)$$

One major advantage afforded by wavelets is the ability to perform local analysis. Consider a sinusoidal signal with a small discontinuity – one so tiny that is barely visible (Fig. 3.3a). A plot of the Fourier coefficients of this signal shows nothing particularly interesting, a flat spectrum with one peak representing a single frequency. The Fourier analysis represented by the energy of the coefficients shows a single high energy peak at 65 feet/cycle (Fig. 3.3d). However a plot of wavelet coefficients, known as a scaleogram (Fig. 3.3c), clearly shows more detail and identifies the exact location in space of the discontinuity at depth 3085 feet, in addition to the detection of the low frequency cyclicity of the signal. The wavelet analysis is able to reveal aspects at small scales (high frequency) and large scales (low frequency).

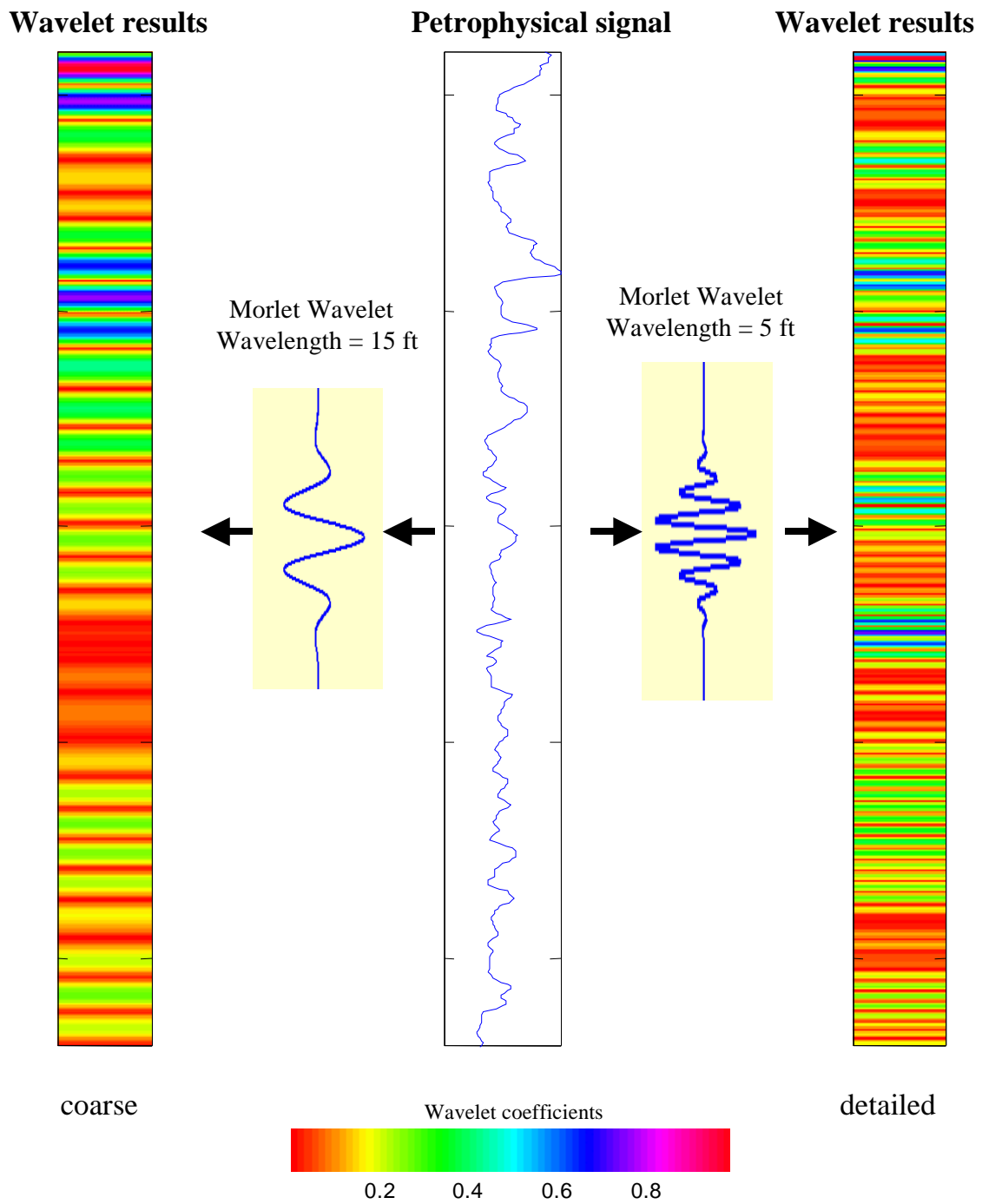


Figure 3.2 Continuous wavelet transform using the Morlet wavelet.

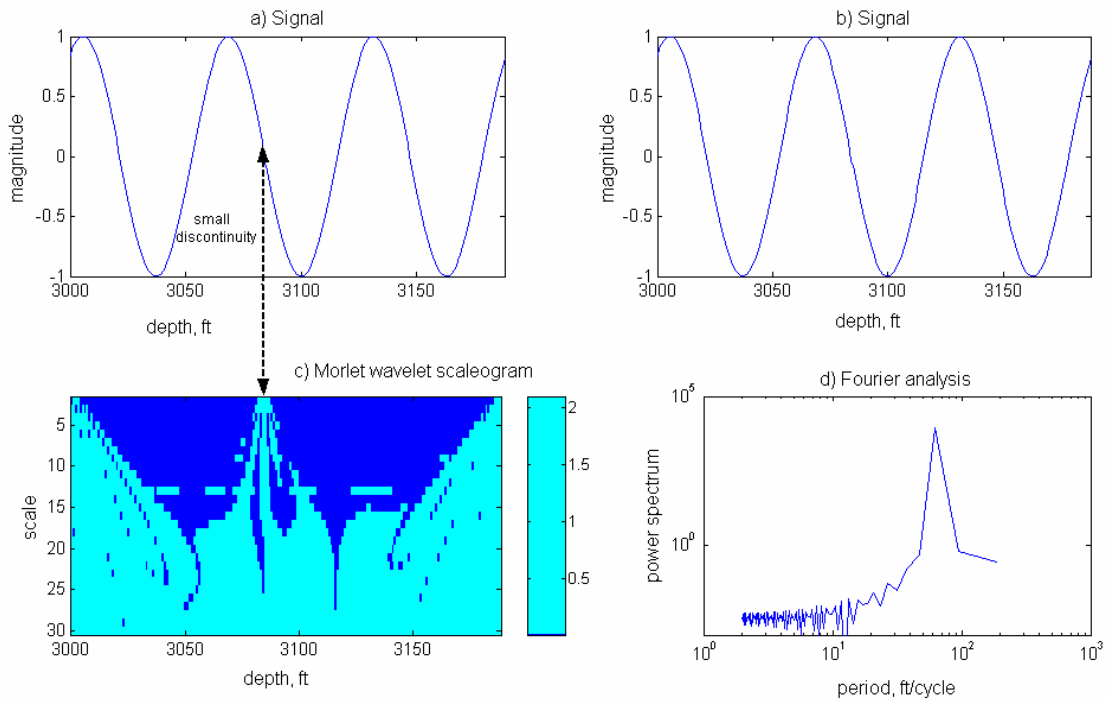


Figure 3.3 Wavelet and Fourier analysis of a small discontinuity.

3.1.3 Spatial statistics

A standard tool used in geostatistical analysis and modelling of rock properties is the SV which estimates the variability of a property as a function of distance. The SV measures the weakness of the relationship between measurements of Z located a certain distance apart. Its standard estimator is

$$\gamma(k) = \frac{\sum_{i=1}^{n-k} (z_i - z_{i+k})^2}{2(n-k)\text{Var}(z)} \quad (3.10)$$

where the n samples Z_1, Z_2, \dots, Z_n , $\text{Var}(Z)$ is the variance of Z , and k is termed the lag and is the number of spacings between the samples.

The semivariance, $\gamma(k)$, is plotted against the separation distance. When $k = 0$ the semivariance approaches a nugget value which may be zero. When k is large, the data are not correlated, the SV approaches the sample variance, i.e. the normalized semivariance given in equation 4 approaches one. SV analysis can be used to determine the spatial correlation of a petrophysical property.

3.1.4 Comparison of cyclicity detection methods

The three different signal analysis methods described above are used for different objectives. To demonstrate the advantage of wavelet analysis over the Fourier transform and the SV for cyclostratigraphic interpretation, we will consider two examples that may occur in strata. The first instance is a signal with an abrupt change in frequency, which can represent a parasequence boundary. Figure 3.4 shows the results generated by the three different signal analysis methods. The synthetic signal has two components, one from 3000 to 3300 ft with wavelength of 100 ft, and the second from 3300 to 3512 ft with wavelength of 40 ft. The SV indicates the two components; the Fourier analysis more clearly identifies the presence of the two frequencies. However, neither of these techniques can identify the locations of the two components and where the frequency changes. On the other hand, the scaleogram identifies the two components and defines the exact location of the transition.

The second example is a signal with two superimposed sedimentary frequencies, which can correspond to Milankovitch cyclicalities (Fig. 3.5). The wavelengths are 40 and 100 ft, respectively. The SV and Fourier analysis identify the presence of both components but cannot define the location in the depth scale. The scaleogram identifies the presence of the two frequencies, quantifies the wavelengths and produces output that clearly identifies the fact that the two cycles are superimposed. Comparing Figs. 3.4 and 3.5, the SV and Fourier analysis generate similar results for the two cases of superimposed cycles and two zones with different wavelength cycles. These examples suggest that wavelet analysis may be more fruitful than either Fourier or SV analysis for cyclostratigraphic studies.

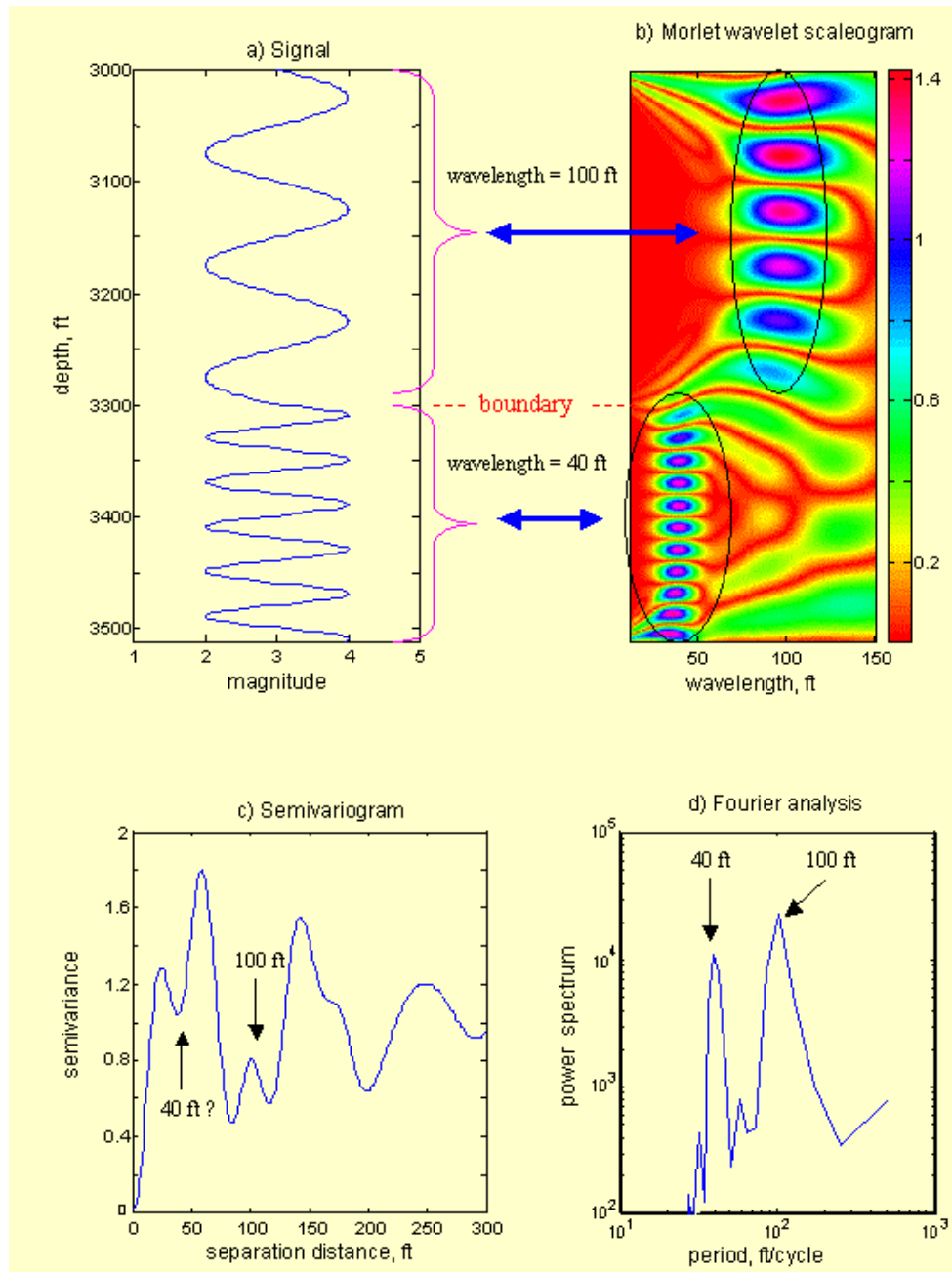


Figure 3.4 Signal analysis results for an abrupt change in frequency, e.g. parasequence boundary.

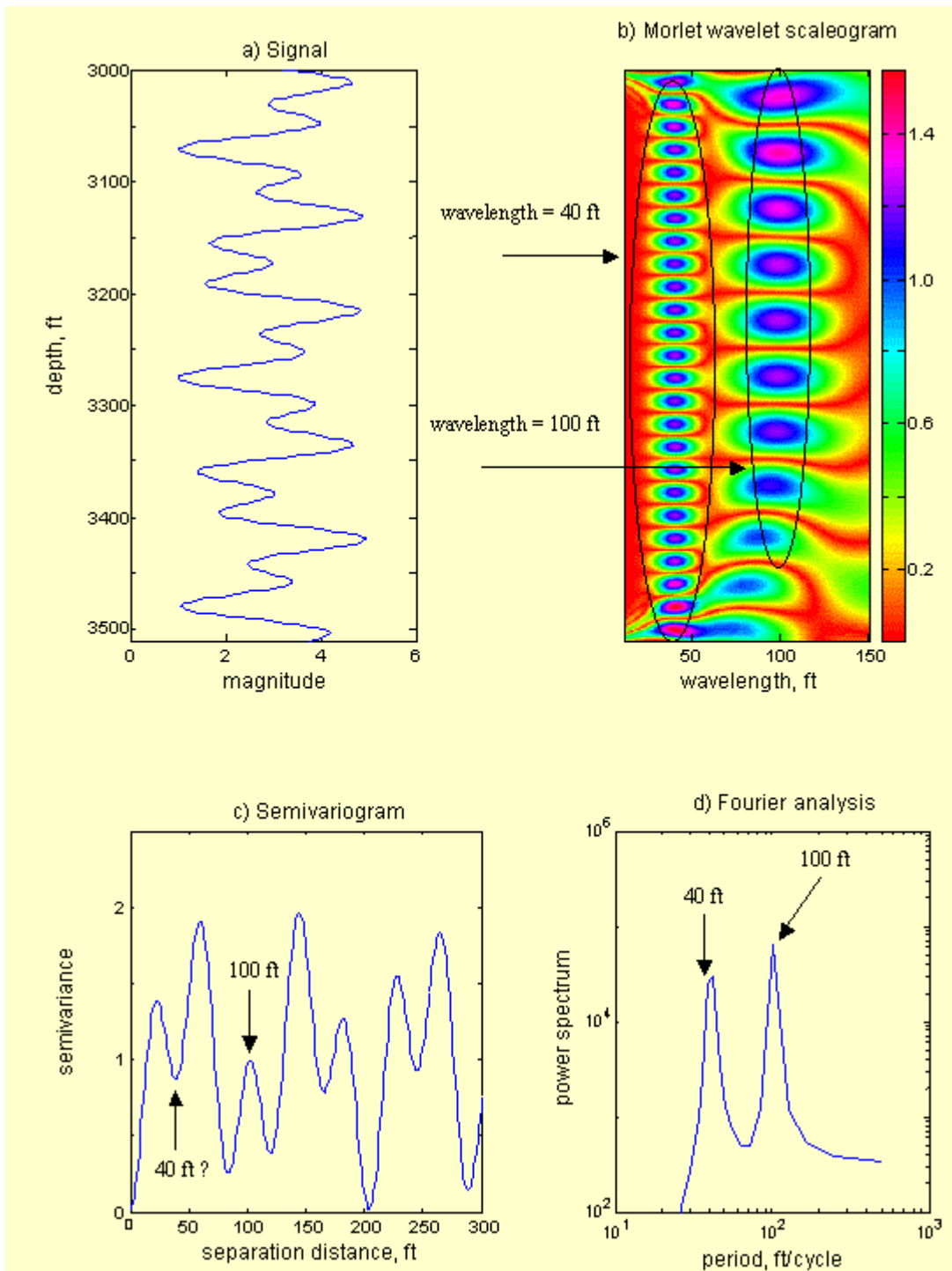


Figure 3.5 Signal analysis results for two superimposed sedimentary cycles (e.g. Milankovitch cyclicity).

3.2 Multiresolution analysis

The CWT is useful for cyclicity determination. However, for other applications, it was found that the CWT was not easy to apply. Meyer discovered that by using certain discrete values of the two parameters a, b , an orthonormal basis can be obtained. The basis is of the form (Jensen and la Cour-Harbo, 2000, Chapter 1):

$$\left\{ 2^{s/2} \psi(2^s t - k) \right\}_{s, k \in \mathbb{Z}}$$

Thus, a will be of the form 2^{-s} and b of the form $k2^{-s}$. With these values of a and b , and using Eq. 3.6, the discrete wavelet transform (DWT) becomes

$$W_{\psi} f(k2^{-s}, 2^{-s}) = 2^{s/2} \int_{-\infty}^{\infty} f(t) \psi(2^s t - k) dt \quad (3.11)$$

By discretizing the function $f(t)$, and assuming the sampling rate to be 1, the integral of (3.10) can be written as (Goswami and Chan, 1999, Chapter 4)

$$W_{\psi} f(k2^{-s}, 2^{-s}) \approx 2^{s/2} \sum_n f(n) \psi(2^s n - k) \quad (3.12)$$

The DWT plays an important role in dividing a complicated signal into several simpler ones and analyze them separately. This concept is known as multiresolution analysis (MRA). Here, the function is decomposed at various levels of approximations and resolutions. As a result, a well-log can be decomposed into a low frequency approximation and several medium-to-high frequency details. Each individual approximation or detail can be analyzed separately, depending on the application. The approximation after 4 levels of resolution can be considered as the general form of a well-log and the details the particular high frequency sedimentation processes for each well. Figure 3.6a displays the filtering process at each level of resolution and Fig. 3.6b shows the decomposition diagram for multiresolution. One example of a DWT decomposition is shown in Fig. 5.19 (Chapter V).

The simplest orthogonal wavelet is the Haar wavelet, which is based on a box function. Other widely used orthogonal wavelets are the Daubechies wavelets with different orders that are compactly supported. There are also semiorthogonal wavelets such as splines wavelets, and biorthogonal wavelets. For our research, we did not

observe significant differences when using different wavelet families. We chose the Daubechies wavelets for multiresolution analysis.

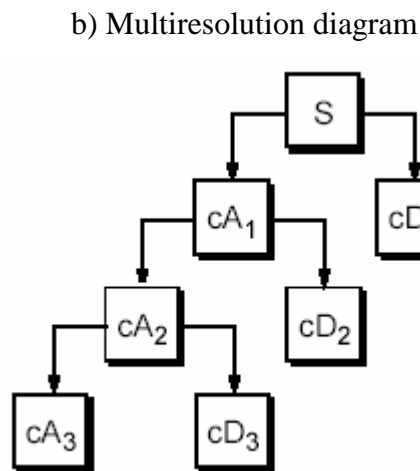
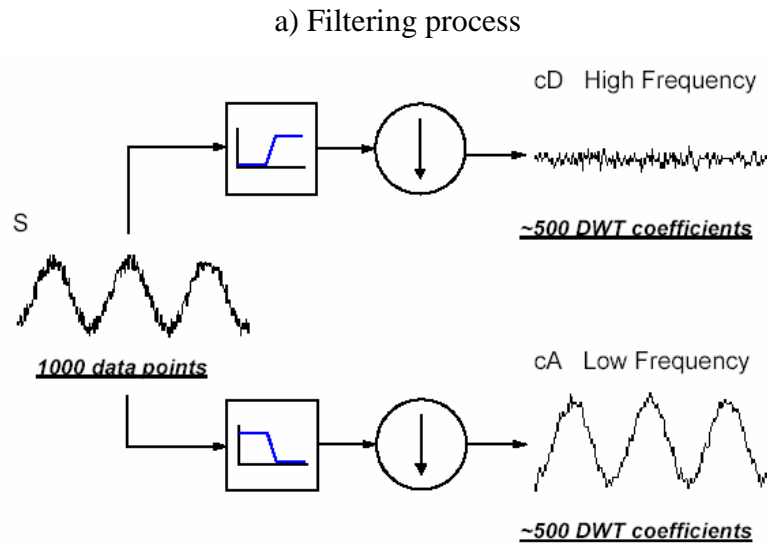


Figure 3.6 Multiresolution analysis indicating a) the filtering process and b) the multiresolution diagram.

3.3 Principal component analysis

Principal component analysis (PCA) is one of the oldest and most widely used of multivariate techniques. The fundamental idea of the method is to describe the variation of a set of multivariate data in terms of a set of uncorrelated variables, each of which is a linear combination of the original variables. The new variables are obtained in order of importance so that the first principal components account for the maximum variation in the original data. The typical task is to determine whether the first few components account for most of the variation in the original data. If they do, then they can be used to summarize the data with minimum loss of information, thus providing a reduction in the dimensionality of the data, which might simplify later analyses (Everitt and Dunn, 1992, page 45). A very useful analysis is the plotting of the first two principal components.

The goal of PCA is to map vector \mathbf{x}^n in a d -dimensional space (x_1, \dots, x_d) onto vector \mathbf{z}^n in a M -dimensional space (z_1, \dots, z_M) , where $M < d$ (Bishop, 1995, Chapter 8). The vector \mathbf{x} can be represented as linear combinations of a set of d orthonormal vector \mathbf{u}_i

$$\mathbf{x} = \sum_{i=1}^d z_i \mathbf{u}_i \quad (3.13)$$

The coefficients z_i are obtained by

$$z_i = \mathbf{u}_i^T \mathbf{x} \quad (3.14)$$

which can be considered as a simple rotation of the coordinate system from the original x 's to a new set of coordinates given by the z 's. If we keep only a subset $M < d$ of the basis vectors \mathbf{u}_i , we will use only M coefficients z_i . The remaining coefficients are replaced by constants b_i so that each vector \mathbf{x} is approximated by an expression of the form

$$\tilde{\mathbf{x}} = \sum_{i=1}^M z_i \mathbf{u}_i + \sum_{i=M+1}^d z_i \mathbf{u}_i \quad (3.15)$$

This accomplishes the dimensionality reduction since the original vector \mathbf{x} with d degrees of freedom is approximated by a new vector \mathbf{z} which has $M < d$ degrees of freedom.

The covariance matrix Σ of the set \mathbf{x}^n is given by

$$\Sigma = \sum_n (\mathbf{x}^n - \bar{\mathbf{x}})(\mathbf{x}^n - \bar{\mathbf{x}})^T \quad (3.16)$$

where the mean vector $\bar{\mathbf{x}}$ is

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}^n \quad (3.17)$$

The minimum error in the approximation of Eq. 3.15 is obtained by discarding the $d-M$ smallest eigenvalues of the covariance matrix, and their corresponding eigenvectors (Bishop, 1995, Chapter 8).

The PCA algorithm starts with the computation of the mean of the vectors \mathbf{x}^n and then subtracting off this mean. Then the covariance matrix is calculated and eigenvalues and eigenvectors are found. The eigenvector corresponding to the M largest eigenvalues are retained and the input vectors \mathbf{x}^n are projected onto the eigenvectors to give the components of the transformed vectors \mathbf{z}^n in the M -dimensional space (Bishop, 1995, Chapter 8). The \mathbf{z} vectors are also known as the z -scores. An example describing the PCA steps is presented in section 5.3.4, Chapter V.

CHAPTER IV

CYCLOSTRATIGRAPHY

Cyclostratigraphy is defined by Schwarzacher (1993) as the branch of geology that uses analysis of cycles to construct and define the stratigraphic framework. Cycles in rock successions are common and represent repetitive stratigraphic and depositional sequences. Eustasy, sediment influx and climate are some of the factors influencing sequence architecture (Nystuen, 1998). By detecting the periodicity of stratigraphic successions, it is possible to subdivide hydrocarbon-bearing intervals into zones for reservoir modeling (Moller and van de Wel, 1998) and to map the zones across a field. This may contribute to properly up-scaled reservoir flow properties, such as the vertical-to-horizontal permeability ratio, and to recognition of important flow baffles and barriers.

4.1 Reasons for geological cycles

Two explanations of the origin of sedimentary cycles are changes in water depth, which are largely caused by tectonic movements, and changes in climate. These causes may overlap and an analysis of the sedimentary record may not permit a clear distinction between the two. For example, it is unclear whether an increase in current energy represents the shallowing of an environment or an increase in storm activity. Similarly, an increased terrestrial influx can be caused either by the uplift of some source area, or by an increased run off due to increased precipitation (Schwarzacher, 1993).

Subdivision of the stratigraphic record by unconformities and bounding discontinuities, and the relationship of these surfaces to relative sea level fluctuations are very important in stratigraphic analysis. Five orders of cyclic sea level change have been defined, ranging from hundreds of millions to tens of thousands of years. Among these five orders, the fourth- and fifth-order cycles, with durations of much less than one million years, reflect a regular cyclic control (Plint et al., 1993).

Sea level changes related to orbitally driven climate cycles are generated by orbital rhythms known as Milankovitch cycles. Three causes of Milankovitch cyclicity are precession, obliquity, and eccentricity. Precession refers to the fact that the earth wobbles like a spinning top; the axis of rotation sweeps out a cone with a period of about 21,000 (21 ky) years. Obliquity refers to changes in the tilt of the earth's axis with respect to the plane in which it orbits (41ky years). Eccentricity refers to changes in shape of the earth's orbit around the sun, from more circular to more elliptical (100 and 400 ky). These variables, acting in combination, control incoming solar radiation and affect global climate, resulting in changes of sea level (Plint et al., 1993).

These Milankovitch cycles overlap with fourth-order sequence stratigraphic cycles (Table 4.1). The parasequences (high-frequency stratigraphic cycles) are a result of changes between high and low stands. In contrast, the influence of the astronomical variations of the Milankovitch cycles on cyclic sediments is not always evident (Schwarzacher, 1998).

4.2 Irish Sea example

For cyclostratigraphic analysis we use the data of well 110/8a-5, located in the Morecambe gas fields in the East Irish Sea. Production is from the Ormskirk sandstone. The geological description of this formation is presented in Appendix A.

The formation in this well is divided into three main zones: Zone 1 (4030 – 4127 ft.), zone 2 (4127 – 4187 ft), and zone 3 (4187 – 4270). Zone 1 is sabkha, subdivided into low-permeability evaporitic and high-permeability non-evaporitic intervals. Zone 2 is subdivided into two high-permeability eolian dune and sandsheet units separated by a low permeability playa unit. Zone 3 is subdivided into high-permeability channel sand units separated by low-permeability silts and clay units. In general, the eolian sands exhibit the best reservoir quality throughout the reservoir, followed by the fluvial channel-fill sands. Playa lake, fluvial channel abandonment, and clay drape deposits have the poorest reservoir quality, are non-reservoir and are baffles and barriers to fluid flow. The sabkha deposits are extremely heterogeneous and exhibit a wide range of

petrophysical properties since these deposits encompass a spectrum of sub-facies ranging from eolian wind-ripple deposits to playa margin deposits (Thomas et al., 1998).

Table 4.1 Comparison of cyclic frequencies of sequence-stratigraphic units with Milankovitch cycles (after Mitchum and Van Wagoner, 1991).

CYCLICITIES OF STRATIGRAPHIC UNITS			
LOG FREQ- UENCY (Millions of years)	CYCLIC		NON-CYCLIC
	SEQUENCE- STRATIGRAPHIC	MILANKOVITCH	
100	FIRST-ORDER TECTONIC CONTROL		
10	SECOND-ORDER SUPERCYCLE (STACKED SEQUENCES)		
1.0	THIRD-ORDER FUNDAMENTAL CYCLE (SEQUENCE, COMPOSITE SEQUENCE)	ECCENTRICITY	
0.1	FOURTH-ORDER PARACYCLE (PARASEQUENCE HIGH-FREQUENCY SEQUENCE)	ECCENTRICITY OBLIQUITY PRECESSION	
0.01	FIFTH-ORDER ? PARACYCLE (PARASEQUENCE HIGH-FREQUENCY SEQUENCE)		(BEDSETS, BEDS, LAMINASETS LAMINAE)

The data comprise wireline logs, as well as core analyses. The conventional well logs available for analysis are gamma ray (GR), dual laterolog (LLS and LLD), microspherical focused log (MSFL), neutron porosity (PHIN), bulk density (RHOB), and photoelectric factor (PEF). Core plug porosity and permeability measurements are available for the interval 4040-4270 ft. Probe permeameter measurements were available for some sections of zones 1 and 3.

4.2.1 Wavelet analysis

Morlet wavelet analysis was performed on the well 110/8a-5 data (Figure 4.1). Given the large variation of the resistivity and permeability data, the analysis of these signals was performed on the logarithm of the signal.

Depending on the log, zones 1 and 3 show some cyclicity at 5-7 ft, corresponding to the thickness of channelized or evaporite cemented deposits. The GR scalogram indicates the presence of weak cyclicities at 5 and 19 ft for zone 1. The GR does not respond to the evaporitic sabka facies in zone 1 very clearly. These facies are primarily characterized by low permeability and high resistivity readings. For the same zone, the LLD shows much stronger cycles at 6-8 ft and at 17-21 ft. For the fluvial channels of zone 3, the GR scaleogram shows a strong response, including a change from 4-5 ft cycles at the base to about 20 ft at the top. Here the DLL shows a similar evolution as the system moves from humid to more arid conditions. For zone 2, the GR response is dominated by the one playa horizon at 4148 - 4153 ft. The resistivity scaleogram shows a strong component at 6 to 8ft wavelength. Clearly, several different measurements are needed here to assess the spectral character.

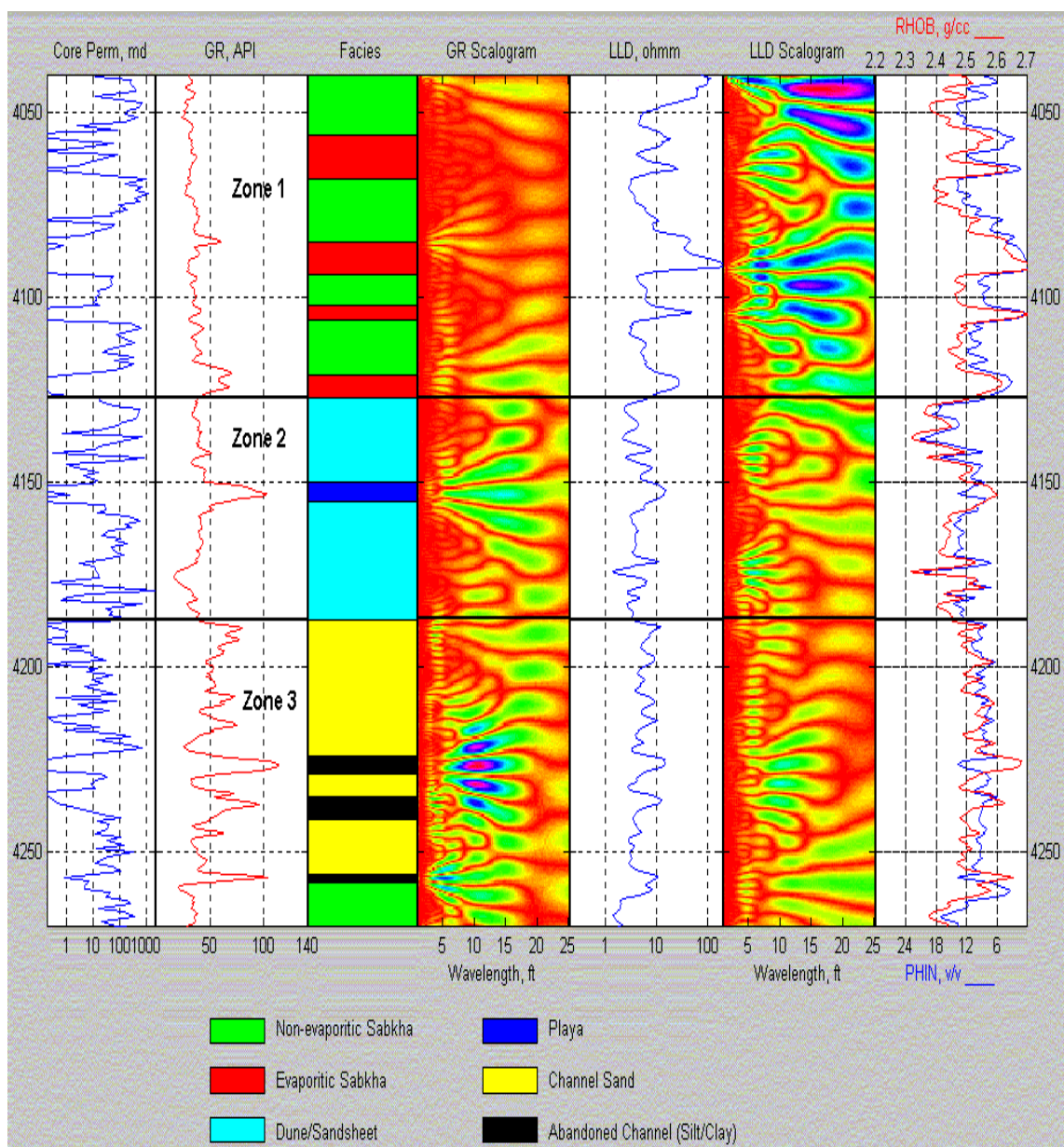


Figure 4.1 Well log, core permeability and Morlet wavelet scaleograms for GR and LLD logs, well 110/8a-5.

4.2.2 Cyclostratigraphic interpretation

Section 4 showed that wavelet transformation can be a useful tool for cyclostratigraphy, especially for the case of superimposed cycles. One way to quantify cyclicity is by

determining the wavelength with the maximum coefficient-energy. This is found by computing the arithmetic average of the wavelet coefficients at each scale (wavelength) for a desired zone. A plot of wavelet spectra energy versus wavelength adequately identifies the dominant wavelengths for each zone (Figs 4.2 – 4.5).

For this analysis, the wavelet transform was applied to plug (approx. 1 ft. sampling spacing) and probe permeability measurements, where they were available, in addition to the well logs. The high sampling rate (20 samples/ft) of the probe permeability measurement allows the detection of high frequency cycles. On the other hand, the conventional well logs, sampled every 0.5 ft and having a measurement vertical resolution of a few feet, would detect cyclicities greater than 3 ft.

Figures 4.2 – 4.5 show the wavelet spectra for zones 1, 2 and 3, respectively. The three components with peaks for the sabkha unit (zone 1) are approximately 5-6 ft, 11 ft, and 22 ft; the probe permeability shows strong responses at 1.5 and 6 ft. For the dunes (zone 2), three peak wavelengths of 4-6, 9, and 23-24 ft are close to those of zone 1. (The GR peak at 15 ft is caused by the one playa unit in this interval.) The fluvial unit (zone 3) shows 4-6, 11, and 21 ft as the dominant wavelengths. For this zone, the probe permeameter shows additional wavelengths at 4 and 10 ft.

Figure 4.5 displays the wavelet spectra for all zones combined. The major component wavelengths are 5-7, 10-12, and 20-23 ft. These peaks are similar in wavelength to those detected for each one of the three zones. This suggests that there is an overall cyclostratigraphic character for the three zones of the Ormskirk Sandstone. Table 4.2 shows the important wavelengths and corresponding ratios.

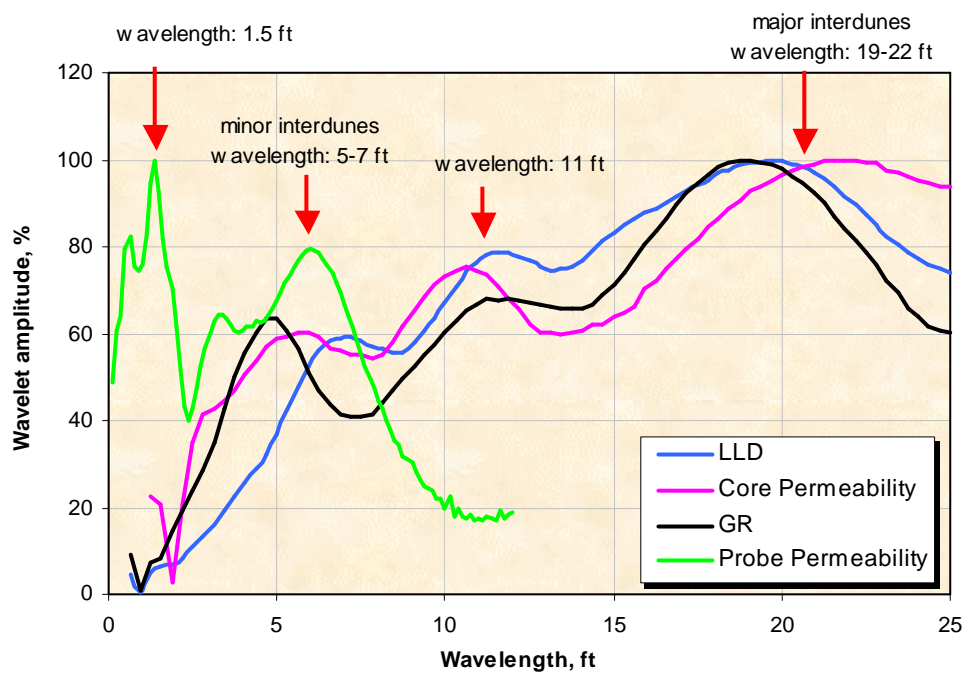


Figure 4.2 Wavelet spectra for zone 1 (Sabkha).

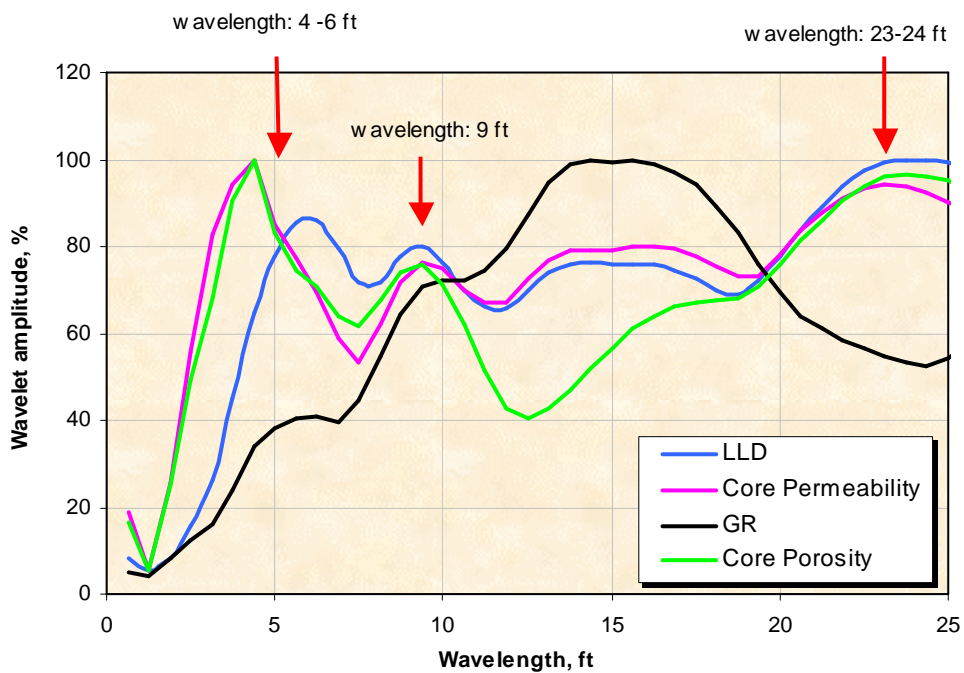


Figure 4.3 Wavelet spectra for zone 2 (Dune).

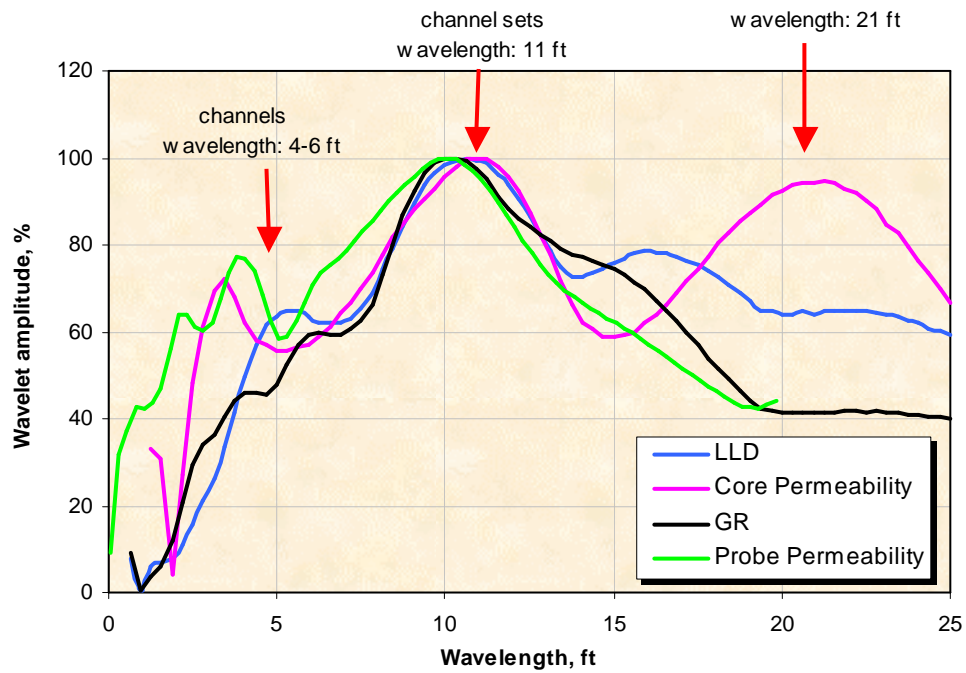


Figure 4.4 Wavelet spectra for zone 3 (Fluvial).

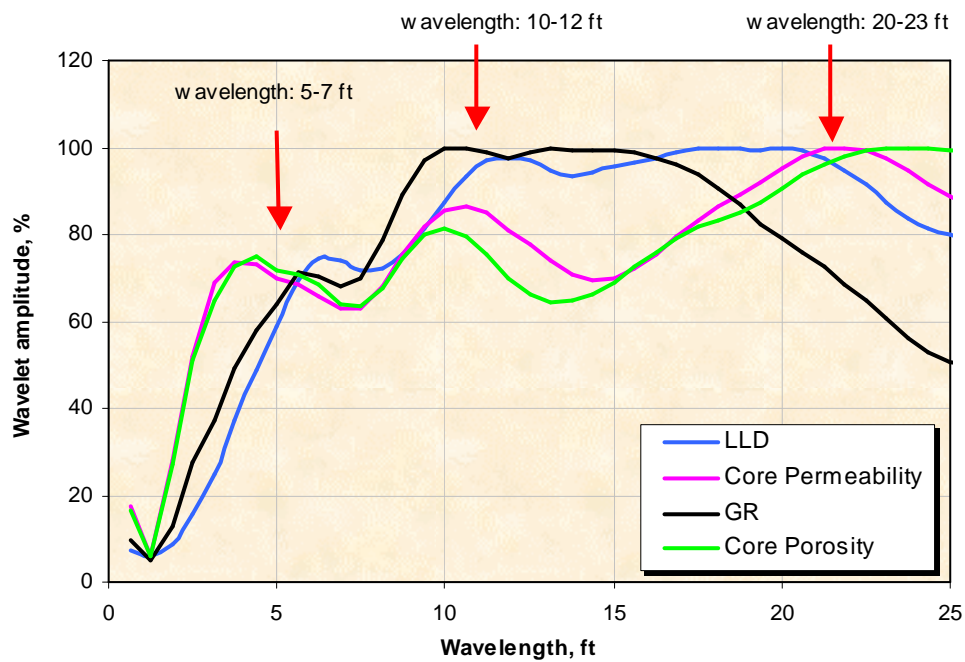


Figure 4.5 Wavelet spectra for all zones combined.

Table 4.2 Wavelengths and ratios in well 110/8a-5.

	Important wavelengths, ft			Wavelength ratios
Zone 1	5-6	11	22	1 : 2 : 4
Zone 2	4-6	9	23-24	1 : 2 : 5
Zone 3	4-6	11	21	1 : 2 : 5

4.3 Discussion

The Milankovitch precession, obliquity and eccentricity cycles have periods of 21, 41, and 100 ky respectively, occurring at a 1: 2: 4.8 period ratio. The dominant wavelengths for the sabkha (zone 1) occur at a 1: 2: 4 wavelength ratio; for the dune (zone 2) at a 1: 2: 5 ratio, and for the fluvial unit (zone 3) at a 1: 2.4: 4.7. For the three units combined the wavelength ratio is 1: 2: 4. The dune unit is the zone that matches most closely the Milankovitch period ratios. The similarity of these ratios with the Milankovitch values strongly suggests that Milankovitch cyclicities are an important factor controlling deposition in the Ormskirk Sandstone. The influence of the precession and obliquity cycles are more evident than the eccentricity cycles. However, for the dune, the wavelet analysis indicates the presence of all three cycles. Hence, climate appears to be an important influence on the deposition.

One way to support this interpretation is appropriate is by calculating the sedimentation rates obtained from the wavelet analysis and comparing with values from previous studies. The Sherwood Sandstone Group deposition spanned the Scythian (Lower Triassic), estimated to have had a duration of 5 Ma (Herries and Cowan, 1997). The thickness of a typical section for this formation is 800 ft. This gives an average sedimentation rate of 5 cm per thousand years (5 cm/ky). This calculation is within the range 1.5 – 6 cm/ka for observed overall sedimentation rates (Anstey and O'Doherty, 2002). For the dune unit (zone 2), the GR, resistivity, core permeability, and core porosity coincide on the dominant wavelength of 9 ft. Taking this peak to correspond to

the obliquity cycle of 41 ky gives a calculated sedimentation rate of 6.7 cm/ky, agreeing well with the published sedimentation rate of 5 cm/ky. Similar results are obtained for the other units and for the different dominant wavelengths. The fluvial unit (zone 3, Fig. 4.4), for example, shows a dominant wavelength of 4 ft for the core and probe permeabilities. Relating this wavelength with the precession cycle (21 ky), a sedimentation rate of 5.8 cm/ky is obtained.

These results support the hypothesis that the Milankovitch cyclicity controlled the deposition in the Ormskirk Sandstone. The presence of these cycles indicate climatic control of deposition with corresponding changes in the sea level, as suggested by earlier researchers (Meadows and Beach, 1993, and Herries and Cowan, 1997).

4.4 Conclusions

The following are general conclusions for cyclostratigraphy:

1. Wavelet analysis generates useful information from well-log responses. It is the appropriate tool for automated cyclostratigraphy. Two other methods commonly used, Fourier transform and the semivariogram, cannot be used to interpret cyclicity in the presence of superimposed cycles.
2. Log and core measurements, responding to different rock and fluid properties, can give different scaleogram results. The combined results from several logs is desirable to define zone boundaries or assess cyclicities.

The following conclusions are specific for well 110/8a-5:

3. Wavelet spectral analysis is consistent with geological definitions of the three main zones in the well used for this study.
4. Amplitudes of spectral peaks appears to correspond with the relative importance of controlling influences on the deposystem.
5. Wavelet coefficients clearly reflect the different orders of cyclicity that occurred during sedimentary deposition. The ratios of cycle wavelengths (~5, ~10, and ~21 ft) correspond to ratios of the Milankovitch cycles 21, 41 and 100 k years.

6. Sedimentation rates calculated from the association of dominant wavelengths and Milankovitch cycles is about 6 cm/year for the Ormskirk Sandstone. This corresponds well with the sedimentation rates of 5 cm/year from published studies. This result supports the hypothesis that Milankovitch cyclicity influenced deposition in the Ormskirk sandstone.

CHAPTER V

MULTIPLE-WELL CORRELATIONS

Geologists and petrophysicists use well logs to correlate formations across fields and basins. Multiple-well correlation has usually been a laborious manual process involving large amounts of visual and qualitative analysis. This task is slow, labor-intensive and inconsistent. Correlations of a set of logs made by several interpreters may differ considerably. Multiple-well correlation is complex. Some units may be laterally continuous over great distances while others may pinch out. Correlations can be disrupted by unconformities, hiatuses, and faults. Cyclic repetition of lithologies can add difficulty to the correlation process. In many cases the units cannot be identified uniquely (Doveton, 1994, Chapter 6, p. 127).

Multiple-well correlation using well logs is usually validated with core and cuttings analysis. Wireline logs measure physical properties that are indirectly related to composition and texture of rocks. Therefore, well logs generate lithostratigraphic correlations that usually do not match coeval units of chronostratigraphy. Correlations track lithofacies in their migrations through space and time (Doveton, 1994, Chapter 6, p. 127).

Well-logs do not determine lithology or grain size directly. Consequently, lithological and textural interpretations are based on calibration of log responses with core or cuttings. However, logs can be used for qualitative lithological information and facies recognition and mapping. The typical logs used for lithofacies interpretation are the SP, resistivity and GR. In most cases the SP and GR have similar patterns (e.g., fig. 5.12 in Galloway and Hobday, 1996). The SP is an indirect measure of permeability and in general the amount of permeability is associated with the shale content (permeability decreases with shale content), and the natural radioactivity increases with shale content. On the other hand, the resistivity log measures resistivity of the rock matrix and pore fluids. Because rock matrix resistivity is high compared to formation water, measured resistivity is primarily a function of pore-fluid interaction rather than lithology.

However, if porosity and permeability are low, as in a tightly cemented or highly compacted bed, the resistivity readings increase considerably. Resistivity may thus be used to measure thickness of sand bodies in freshwater zones or in facies successions characterized by very low intergranular porosity (Galloway and Hobday, 1996).

The choice of log that is best suited for a given stratigraphic correlation has not received much attention. A poor choice would be use the GR to correlate shale-free carbonate sequences. Common well-logs measure for the most part shale content and porosity. Resistivity logs tend to emphasize porosity fluctuations and shale stringers within sandstones. One author has considered the acoustic velocity the best log for subsurface correlation. This log is less affected by borehole conditions or fluids and can discriminate geologic features (Doveton, 1994, Chapter 6, p. 127).

Log patterns may be used at three levels of interpretation: **1)** determination of vertical facies successions and bedding architecture at the scale of sequences and depositional systems. In a progradational succession, in which upward coarsening is the result of the increasing number and thickness of sand beds, both SP and GR display a progressive deflection from the shale baseline to the sand baseline, reflecting the increasing proportion of permeable sand and decreasing proportion of radioactive clay. **2)** recognition and mapping of log facies within correlative stratigraphic intervals. Comparison of many logs through a genetic stratigraphic unit will lead to recognition of recurrent log motifs or electrofacies (see example in figure 2.5, Galloway and Hobday, 1996). Idealized SP or GR logs of vertical sequences for the various depositional models are proposed in the literature. **3)** delineation of lateral and vertical textural relationships of individual sand bodies. However, textural changes within sand sized sediment are least likely to be displayed by the logs. The characteristic upward-coarsening and upward-fining textural patterns of progradational and lateral accretion bedding geometries are recognized in well logs.

The steps involved in facies architecture include the generation of a structural model and a petrophysical model, using biostratigraphy, seismic, core analysis, and well logs. The integration techniques involve in many cases expensive data acquisition

processes such as seismic 3D, coring operations, special core analysis, special techniques, and a considerable amount of time of an integrated reservoir characterization team. The contribution of this research is the development and testing of a technique to automatically identify the units in different wells, starting with the detailed reservoir characterization performed on a single well or a few wells. We do not attempt to perform the core description and integration leading to the identification of different units. We start from a given characterization provided by geoscientists in a few “training” wells, understand the well log characteristics for each unit defined for different signals (GR, ρ_b , Phinl, R_t, R_{xo}), extract the important features and generate the unit identification for all wells in the field where the basic set of well-logs have been acquired.

We want to test whether wavelets will aid in boundary detection. We perform our analysis for both the well logs and the wavelet transform of the logs and compare the results. This process involves feature extraction, dimensionality reduction, and pattern recognition techniques.

5.1 Apiay field

The Apiay field is located in the oil-prolific Llanos basin. The main producing reservoirs are located in Upper Cretaceous (Coniacian-Campanian) sandstones, denominated as K2 (lower) and K1 (upper) units. The original oil in place is estimated in 298 MMBLS with recoverable reserves of 107 MMBLS. The K1 unit is overlain by the T2 unit (upper Eocene) which consists of estuarine deposits, with characteristic fining upwards from the gamma-ray log (Malagon, 1997). The information available for our project comes from 10 vertical wells (Figures 5.1-5.3).

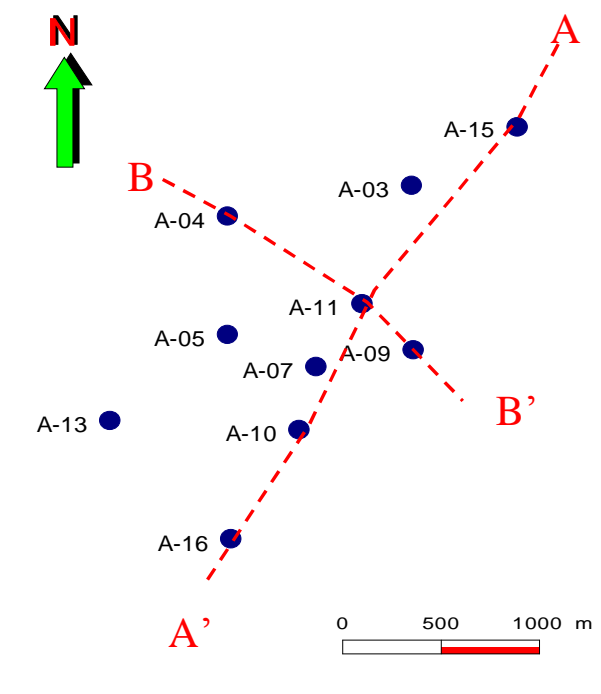


Figure 5.1 Area of Apiay field and wells analyzed in this study.

The operational unit K2 represents the best reservoir and is composed of massive braided fluvial sandstones that produce 25 °API oil from an average net pay thickness of 210 ft. The K1 unit is compartmentalized and represents continuous and progressive transgressive units, formed by the deposition of sandstones and shales having a more marine-like character. These individual sands are more laterally and vertically heterogeneous compared to the underlying K2 unit. The unit complexity is reflected in the varying oil gravities and reservoir pressures. The K1 unit has been divided into six stratigraphic units that are correlated throughout the field. From bottom to top, they are the K1L2, K1L2, K1M3, K1M2, K1M1, and K1U1. The K1L2 and K1L1 units were deposited in a transitional marine environment of a fluvial dominated delta. Major facies include distributary channels and interdistributary bays. The K1M3 unit was deposited in a mesotidal shoreface with barrier inlets-and tidal channel-facies. The

K1M2 units evolved in a fluvial-dominated delta with distributary channels and channel mouth bar facies. The K1M1 unit was deposited in a mesotidal shorezone with abundant tidal inlets and some small tidal barriers. The lithological record of the K1 operational unit ends with the late Eocene deposition of the K1U1 unit over the upper Cretaceous units. The K1U1 consists of braided fluvial stream facies (Nieto, 1997).

The stratigraphic framework described above is the result of the integrated stratigraphic studies performed by Nieto (1997) and Malagon (1997). These studies incorporated data from well logs, 3D seismic, core analysis, production, pressure, biostratigraphy, and palinology. Figures 5.2 and 5.3 show NE-SW and NW-SE stratigraphic cross sections displaying the units from the integrated studies, using the GR and R_t logs. All other wells and Phinl and ρ_b logs are shown in Figures B1 to B4, Appendix B. Five of the 10 wells were cored: A-03, A-09, A-011, A-13, and A-16.

In this Chapter we are assuming that the boundary picks from the integrated study are correct and represent the “true” locations. The boundary with the highest level of confidence in the integrated study corresponds to the T2/K1 boundary.

5.2 Well-log response to boundaries

To investigate the level of importance of each signal to the boundary definitions in Apiay, we take the logs of all wells in the field and make the analysis. The logs selected for analysis in this project are: R_t , R_{xo} , Phinl , GR, and ρ_b . The resistivity logs typically range over several orders of magnitude, so the logarithmic transformation is applied to perform visual inspection of these signals (Bassiouni, 1994). For this research, all analyses, displays, and transformations of the resistivity logs (R_t and R_{xo}) are performed on the natural logarithm of the signal.

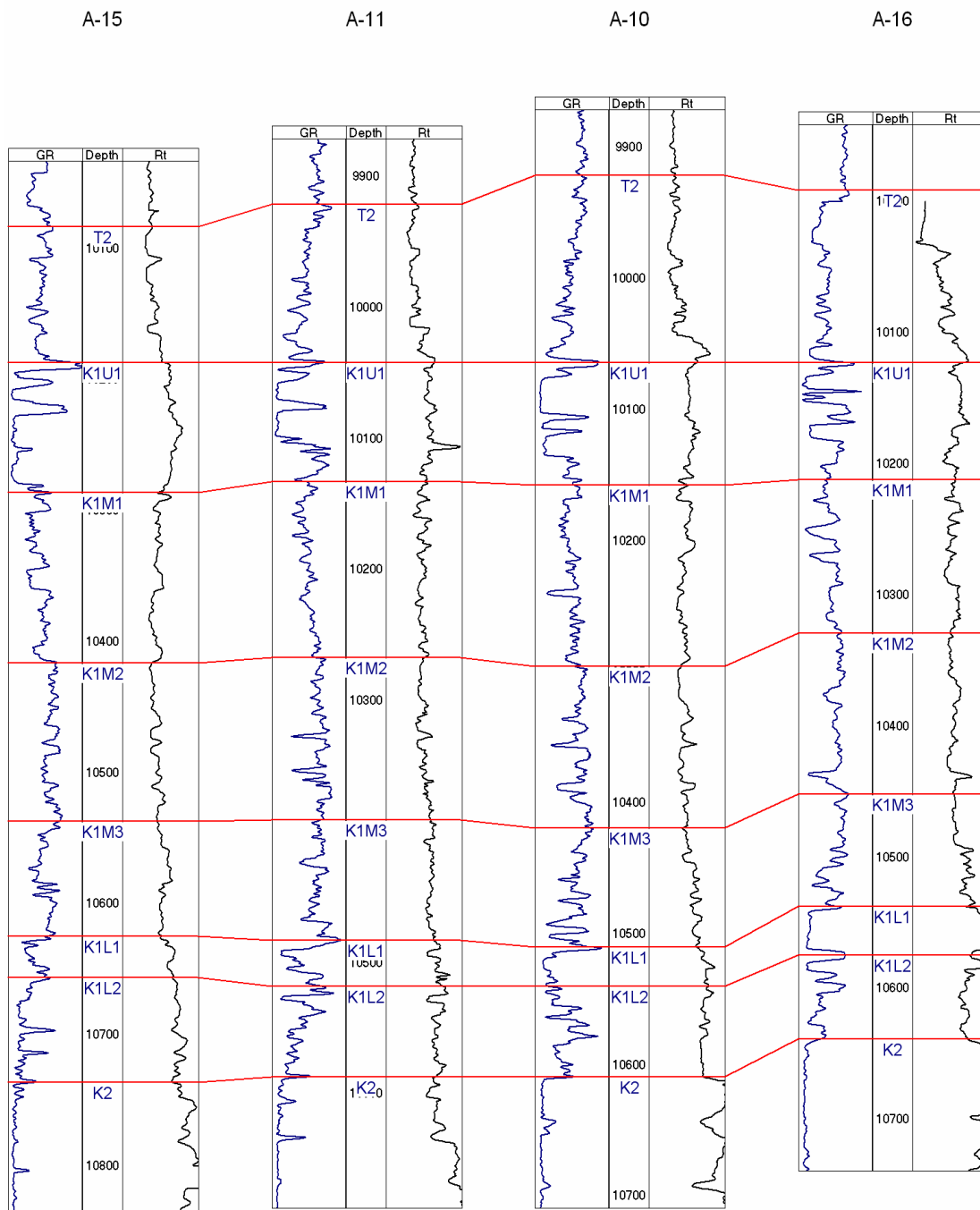


Figure 5.2 Stratigraphic cross section A-A' (NE-SW) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are GR and Rt.

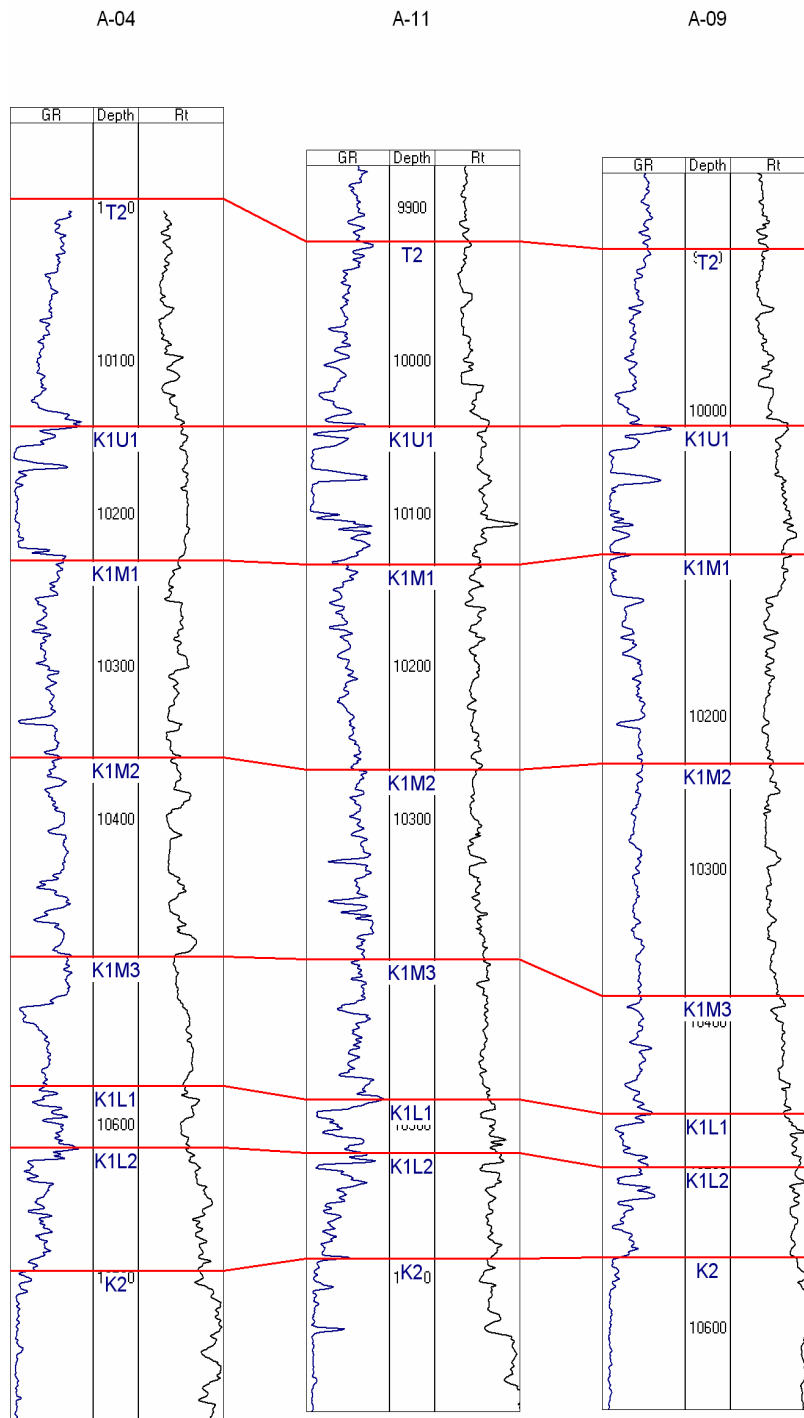


Figure 5.3 Stratigraphic cross section B-B' (NE-SW) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are GR and R_t.

Well logs can be used to construct stratigraphic cross sections or perform electrofacies analysis. It is important to make the distinction between these two techniques. Electrofacies allocate rock regions with similar petrophysical behavior, typically influenced by distinctive porosity and permeability correlations (e.g., Lee et al., 2002). Rock regions do not necessarily correspond to the same geologic time, and the main use of electrofacies is the generation of permeability distributions. Boundary identification is essential in the definition of stratigraphic units and involves time. A series of boundaries forms a stratigraphic framework for the purposes of geologic modeling. The stratigraphic framework is mainly a deterministic model and is coarse, whereas the detailed geologic modeling is highly stochastic and various plausible detailed models are obtained. In many cases the dynamic modeling evaluates several realizations. In this process, the stratigraphic model typically remains unchanged.

To achieve the boundary identification, we need to extract the characteristic features in the neighborhood of each boundary. This pattern recognition process depends on the quality of the reservoir characterization. In the case of Apiay, all wells have logs and their corresponding petrophysical interpretations. Boundary identification will improve for wells with core analysis and having good matches between stratigraphic markers and seismic traces. For Apiay, the well with a high level of confidence in the seismic response for the K1 intra-horizons is A-15. This is a priori information typically known for oil fields.

The general characteristics at the neighborhood of the most important boundaries in Apiay, leading to the selection of relevant features are as follows (Figures 5.2, 5.3 and B1-B4):

5.2.1 T1/K1U1 boundary

The GR is increasing upwards (fining upwards) in the K1U1 section while decreasing upwards (coarsening upwards) in the T2 portion. The variability in the GR readings is higher for the K1U1 section. The span of the GR is higher in the K1U1. The mean GR readings are very similar for both sections. The variability of the Rt is higher in the T2

section. The Phinl is increasing upwards in the K1U1 section, while decreasing upwards in T2.

5.2.2 K1L2/K2 boundary

The GR mean values are smaller in K2 than in K1L2. The GR shows fining upwards in K2 and coarsening upwards in K1L2. The K2 contains intermediate gravity oil, in contrast to light oil in the K1L2 units.

The T1/K1U1 boundary is in the neighborhood of very well organized sequences and the confidence on the allocation of this boundary from the integrated study is very high (more than 90%). This boundary is an unconformable surface recognizable on seismic sections (Gutierrez et al., 1996).

On the other hand the location of the K1L2/K2 boundary is a much more difficult task due to the considerable amount of erosion experienced in the K2 unit. Even though 3D seismic contributed significantly to improvements on the determination of this boundary, some uncertainty remains about its location at the wells (Nieto, 2003).

5.3 Boundary recognition

In this study we use just well logs, which are the typical data available for every well. We use the wells with the completed data analysis (well logs, petrophysics, core, and seismic) as the training set. To perform the boundary recognition using well-logs we will have 1D signals (one for each log) above and below the boundary. If we analyze a single signal in a window of $w/2$ feet below and $w/2$ feet above the boundary, we will have a vector $x = (x_1, \dots, x_e)$ where e is the number of log readings, and $e = w \times 2$ for a sampling rate of 2 samples/ft. In our classification problem, we are to develop an algorithm which will assign the upper portion of the boundary to class C_1 and the lower portion of the boundary to class C_2 . If we increase the number of logging curves to nl , then the input vector will have $e \times nl$ dimensions. Each depth position would generate a variable with nl values each. A typical 40-ft window would generate an 80-variable input (80-dimensional space) for each signal.

The presence of large numbers of input variables can lead to severe problems for pattern recognition methods (Bishop, 1995). One technique used to improve the pattern recognition performance is by combining the input variables to make a smaller number of new variables called features. These can be constructed based on the particular problem being tackled (Bishop, 1995). In our case, for each well log we will have nf features. Each one of the nf features is a scalar summarizing the data for a given depth range. In our boundary detection problem, we perform the analysis in the neighborhood of the boundary, which leads to the definition of a window length (in depth units). Since we are interested in defining features that show contrast between the upper and lower sections of the boundary, we need to define upper-section and lower-section features. Thus, the upper section of the window will result in nf upper features and the lower section will also result in nf lower features. Figure 5.4 illustrates the concept of window length and features, using the T2/K1U1 boundary.

In addition to the reduction of dimensions, the definition of features is necessary to satisfy the requirement that C_1 must be located on top of C_2 ; thus, bed sets in the neighborhood of a boundary can be assigned to unique class. This is not possible when using readings at all depth points. The typical non-stationary characteristics of well-logs would lead to the assignment of different classes within a bed set.

When using a large number of signals (e.g., 5 well-logs), the resulting number of dimensions of the extracted features is still too large (more than 10 dimensions). We can further reduce the dimensionality by using PCA.

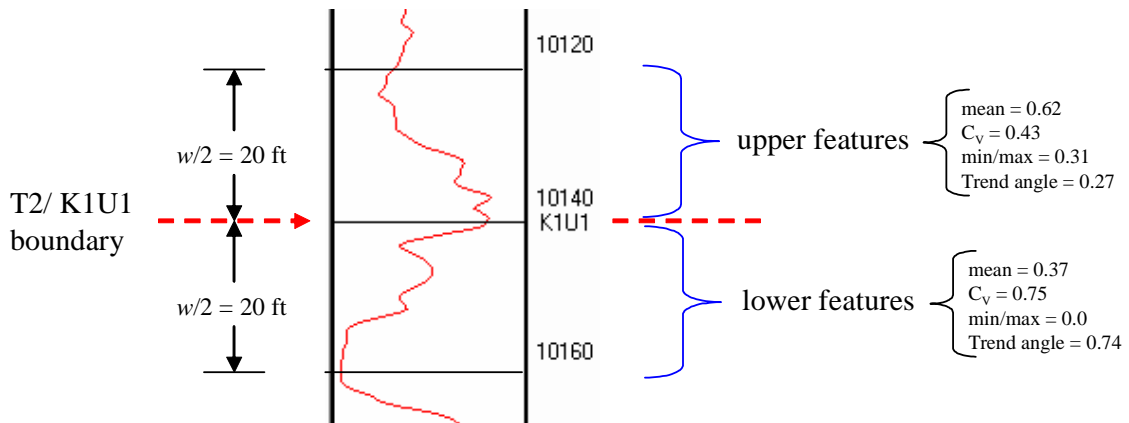


Figure 5.4 GR-features at the T2/K1U1 boundary for well A-04, using a window length $w = 40$ ft.

5.3.1 Pre-processing

Well-logs may contain abnormally high or low readings. Since we are interested in the signal response in a window, it is convenient to smooth those abnormal readings as a pre-processing step before extraction of features. The signal can be smoothed using a moving average filter (e.g., Goumas et al, 2002) prior to analysis for feature evaluation. The signal-averaging filter smoothes rapid fluctuations in the signal. This is also known as a moving average filter and is implemented as follows (Goumas et al, 2002): Let $t(n)$ and $y(n)$ be the input and output signal respectively. For each data point $k \in [\xi+m, N-m]$ of $t(n)$, compute the value

$$y(k) = \frac{\sum_{l=k-m}^{k+m} t(l)}{2m+1} \quad (5.1)$$

of $y(n)$, where ξ is the starting data point of $t(n)$, upon which the filter operates, N the length of input signal, and $2m$ the width of the filter, in depth units, on the input signal.

Considering the well-log tool resolutions, one reasonable choice for these signals is $m = 1$ to 2 ft, which implies a filter operating on a 2- to 4-ft window, or 5 to 9 data points for the typical 2 samples/ft acquisition rate. For this project we performed the smoothing using a filtering-window width of 4 ft.

The pattern recognition process involves the analysis of signals with different units. As a second pre-processing step, we normalize all the log readings of the smoothed signal from zero to one for all depths. These normalized signals, ranging from 0 to 1, are convenient considering that we are more interested in the relative changes observed on both sides of the boundary than the absolute values. The normalized variable for well-log i (where $i = 1, 2, \dots, n$ represent GR, R_t , R_{xo} , $Phinl$, and $Rhob$) is given by

$$x_i^j = \frac{y_i^j - \min(y_i)}{\max(y_i) - \min(y_i)} \quad (5.2)$$

where $j = 1, 2, \dots, J$ are the depth points of signal i .

5.3.2 Feature extraction

To generate the required features, we can develop measures of similarity and abrupt variations, based on the observations presented above for the Apiay boundaries. Here, statistics to be used as features include the sample variability, represented by the coefficient of variation, C_v , and the sample mean, $\hat{\mu}$. The max/min ratio is a third feature. A possible fourth feature is the trend angle, which focuses on differentiating funnel and bell shapes that may occur, for example, in the GR for fining and coarsening upwards respectively. This feature reflects some of the techniques used in stratigraphic analysis. For example, the definition of many stratigraphic units is based on the fining and coarsening characteristics of the bed sets.

5.3.2.1 Definition of features

Now, we proceed to describe the four features we have selected. The first feature, the sample mean for the upper-section and the lower-section of the window are given by

$$\overline{x_i^{upp}} = \frac{\sum_{j=1}^b x_i^j}{b} \quad (5.3)$$

and

$$\overline{x_i^{low}} = \frac{\sum_{j=b+1}^J x_i^j}{J - b + 1} \quad (5.4)$$

where b is the number of depth points from the top of the window up to the boundary and J is the total number of depth points in the window. By construction $\overline{x_i^{upp}}$ and $\overline{x_i^{low}}$ $\in [0,1]$.

The second feature, the coefficient of variation is given by

$$C_{V_i}^{upp} = \frac{\sqrt{\text{Var}(x_i^{upp})}}{\overline{x_i^{upp}}} \quad (5.5)$$

and

$$C_{V_i}^{low} = \frac{\sqrt{\text{Var}(x_i^{low})}}{\overline{x_i^{low}}} \quad (5.6)$$

where $Var(x)$ is the variance of x . C_V is a dimensionless measure of sample variability. This quantity is preferred to the sample variance or standard deviation for data from different populations or sources because of its dimensionless characteristics (Jensen et al. 2000). One alternative to C_V is the LC_V which uses L-moments (Hosking and Wallis, 1997). The first two L-moments are defined as follows,

$$\hat{\lambda}_1 = \frac{1}{I} \sum_{i=1}^I X_i \quad (5.7)$$

$$\hat{\lambda}_2 = \frac{2}{I} \left(\sum_{i=2}^I \frac{i-1}{I-1} X_i \right) - \hat{\lambda}_1 \quad (5.8)$$

$$LC_V = \frac{\hat{\lambda}_2}{\hat{\lambda}_1} \quad (5.9)$$

In our testing the C_V and LC_V applied to the GR gave similar results. Based on this, we selected the C_V as the feature to be used in our analysis for coefficient of variation.

The third feature, the minimum- to-maximum ratio, is a measure of abrupt signal deviation, which may occur in drastic lithology changes from shale to sandstone. The min/max ratio, R , is given by

$$R_i^{upp} = \frac{\max(x_i^{upp})}{\min(x_i^{upp})} \quad (5.10)$$

and

$$R_i^{low} = \frac{\max(x_i^{low})}{\min(x_i^{low})} \quad (5.11)$$

The fourth feature is the trend angle, θ , of the signal in the analysis window. The trend angle is computed with respect to the boundary with depth values decreasing upwards. If either the upper or lower portion of the window is coarsening upwards (GR decreasing upwards) the trend angle will be $< 90^\circ$. For fining upwards (GR increasing upwards) the angle will be $> 90^\circ$. The angle is computed using the regression parameter $\hat{\beta}_1$ obtained from linear regression using the least squares method on the corresponding window section (upper or lower). The linear model is given by

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (5.12)$$

where X is the predictor representing the normalized signal value, Y is the response representing the normalized distance from the boundary increasing upwards, and ε is a random error. The predicted responses are

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X \quad (5.13)$$

The regression parameter is solved as (Jensen et al., 2000)

$$\hat{\beta}_1 = \frac{\sum Y_i X_i - \frac{1}{I} (\sum X_i) (\sum Y_i)}{\sum X_i^2 - \frac{1}{I} (\sum X_i)^2} \quad (5.14)$$

The trend angles for the sections above and below the boundary are given by

$$\theta_i^{upp} = \begin{cases} \frac{\tan^{-1}\left(\hat{\beta}_1^{upp}\right)}{\pi} & , \text{ if } \hat{\beta}_1^{upp} \leq 0 \\ 1 - \frac{\tan^{-1}\left(\hat{\beta}_1^{upp}\right)}{\pi} & , \text{ otherwise} \end{cases} \quad (5.15)$$

and

$$\theta_i^{low} = \begin{cases} \frac{\tan^{-1}(\hat{\beta}_1^{low})}{\pi} & , \text{ if } \hat{\beta}_1^{low} \leq 0 \\ 1 - \frac{\tan^{-1}(\hat{\beta}_1^{low})}{\pi} & , \text{ otherwise} \end{cases} \quad (5.16)$$

Then trend angle is normalized from 0 to 1 representing the range 0° to 180° . Thus, $\theta \in [0, 0.5]$ for decreasing upwards and $\theta \in [0.5, 1]$ for increasing upwards. Figure 5.5 illustrates the steps in calculating the GR-trend angle at the T2/K1U1 boundary, for well A-04. The upper section of the boundary shows a well defined coarsening upwards trend, while the lower section displays a fining upwards behavior.

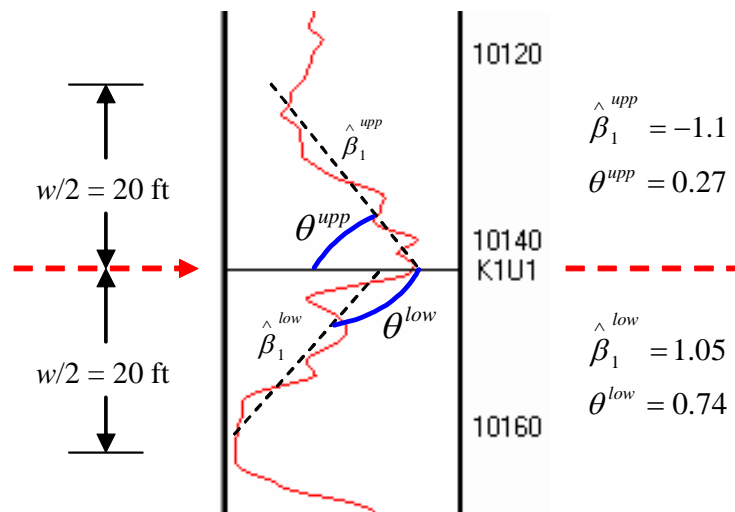


Figure 5.5 GR-trend angle at the T2/K1U1 boundary for well A-04, using a window length $w = 40$ ft.

5.3.2.2 Window length for feature extraction

The values of the extracted features depend on the window length, w , selected. Our algorithm incorporates w as a sensitivity parameter to be optimally determined from the boundary detection performance. However, when incorporating multiple well-logs and wavelet transformation, the process is computationally very intensive and becomes impractical. The best way to perform the boundary detection is by using a unique window length. We chose $w = 40$ ft, based on the following observations:

- 1) We selected $w = 10, 20, 40,$ and 60 ft to find the optimal window length. We discarded values below 10 because we are not interested in definition of beds. Values above 60 ft are not practical because they may define sequences which are beyond the detail of bed sets required in boundary detection.
- 2) We performed boundary detection for the T2/K1U1 boundary for the window lengths mentioned above. The results show that $w = 40$ ft gives the best performance for boundary detection. The performance for $w = 10$ ft is very poor, and the performance of $w = 60$ ft is better than $w = 20$ ft.
- 3) Cyclostratigraphic analysis in Apiay reveals that the wavelengths of the strongest cyclicities are close to 20 ft. Figure 5.6 shows the GR-wavelet amplitudes of wells A-03 and A-09 for four stratigraphic units. Most of the energy is at 20 ft or smaller wavelengths. There are components, however, with wavelengths over 20 ft so that, had we examined the effects of the value of w more closely, we might have gotten better results with w somewhat larger than 20ft.
- 4) The strongest cyclicity of 20 ft corresponds to $w/2 = 20$ ft, which leads to the optimal window length of $w = 40$ ft. This coincides with the detection performance described above.
- 5) The window needs to be long enough to capture trends within the zones, but short enough to ensure that it is capturing within-zone features. Given the characteristics of Apiay the half window, $w/2 = 20$ ft, satisfies those requirements.

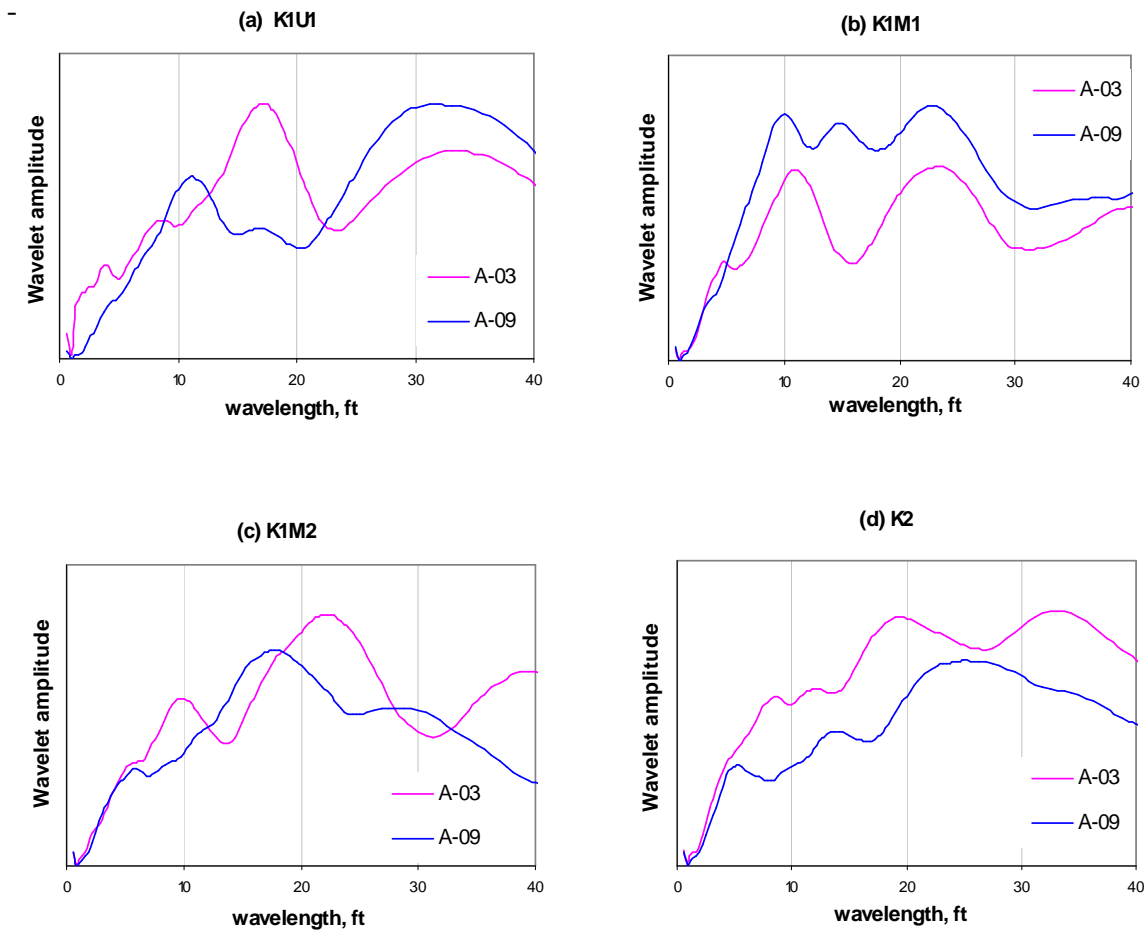


Figure 5.6 GR- wavelet amplitude as a function of wavelength in wells A-03 and A-09.

5.3.3 Evaluation of features

The gamma-ray is the signal most widely used for correlation. We begin our analysis of the Apiay data using this signal. Our research will focus on the boundary detection of the six operating units of the K1 reservoirs, including the top and the base. The K1 unit is overlain and underlain by the T2 and K2 units, respectively. One of the main characteristics of the T2 unit is the fining upwards profile observed in the GR log.

To start the process, we use the GR for the T2/K1U1 boundary, which is one of the most important boundaries in the Apiay field. Figure 5.7 displays the features extracted for all wells at the T2/K1U1 boundary for a 40-ft window. The best features for boundary identification are those which change the most between the upper and lower sections of the window, and that are consistently different for all wells. Finding features with these characteristics is a key factor in recognizing boundaries.

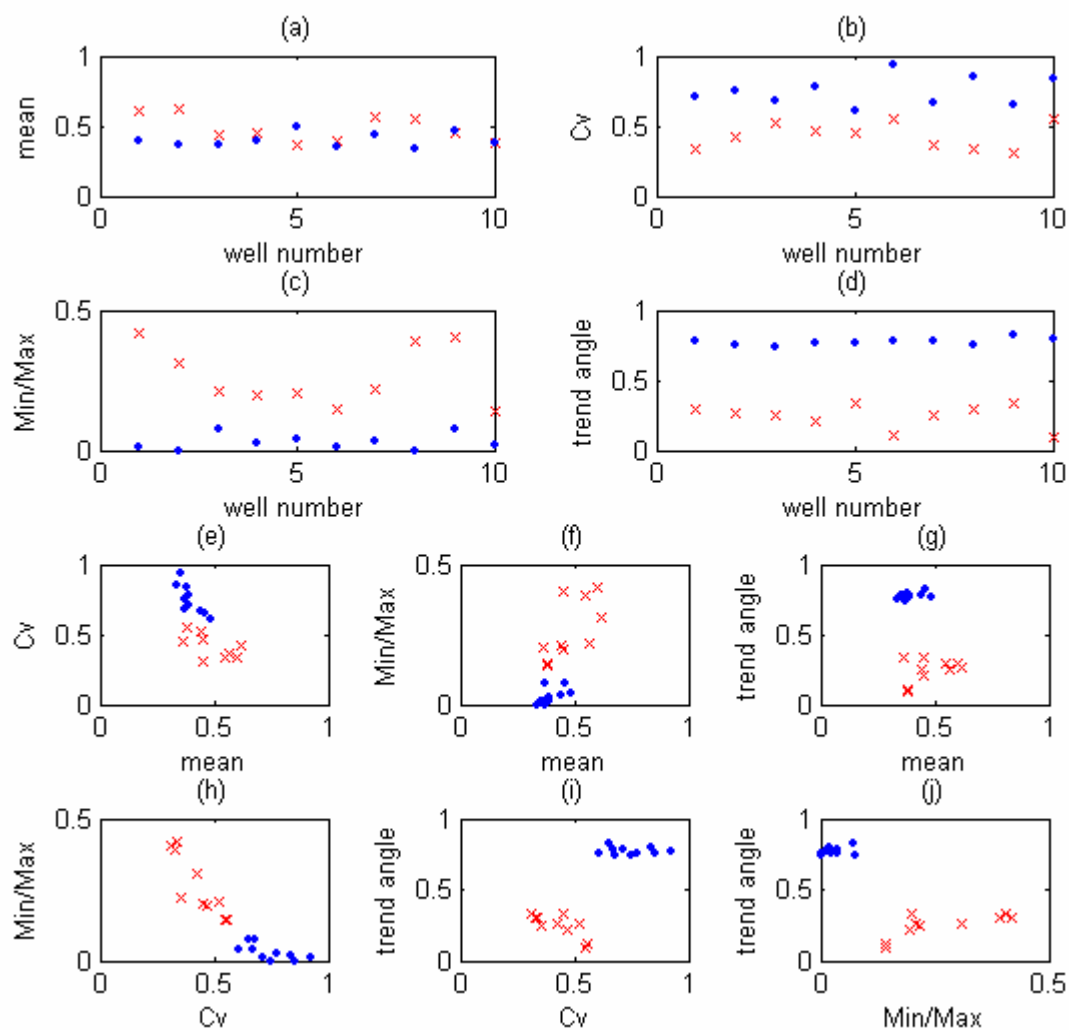


Figure 5.7 Features from GR for T2/K1U1 boundary; window size 40 ft. Dots represent lower window features and crosses upper window features.

The best feature to distinguish the T2/K1U1 boundary in Apiay is the trend angle (Fig 5.7d). By plotting pairs of features, the pairs trend angle-CV (Fig 5.7i) and trend angle-Min/Max (Fig 5.7f) display the best combinations for class separation (boundary identification). Here, the mean is a poor discriminator because the stratigraphic units on both sides of the boundary are mainly controlled by the coarsening upwards profile of the T2 unit and the and fining upwards trend of the K1U1 unit in the proximity (20 ft) of their boundary. The GR-mean values for such contrasting bed sets may be very similar.

The separation of the features above and below other boundaries in the Apiay field are less clear than the T2/K1U1 boundary. The quality of the separation of classes decreases using 20-ft windows, and improves a little bit for 60-ft windows. Based on this, it seems reasonable to choose $w = 40$ ft as the optimal window size. The Ph_{nl} and R_t are logs that show potential for successful pattern recognition at this boundary, although not so robust as the GR since the separation of classes between the upper and lower sections of the boundary is less evident.

5.3.4 Dimensionality reduction

Using a single unprocessed signal generates a 4-feature vector (4 dimensions, $d = 4$). In the previous section we showed graphically the potential performance of the individual and pairs of features (single and two-dimensional plotting) for the T2/K1U1 boundary. Bishop (1995) described the problems which could arise in performing pattern recognition in high-dimensional spaces, and the improvements which can be achieved by first mapping the data into a lower dimensionality space. The dimensionality reduction is achieved using principal component analysis.

The steps for dimensionality reduction (Figure 5.8) are:

- 1) *Feature extraction.* Each feature will have two values: the upper feature (from the upper window) and the lower feature. For two samples per foot the original dimensionality is reduced from $2w \times nl$ to $nl \times nf$. The input matrix will reduce

from $2T \times (2w \times nl)$ to $2T \times (nl \times nf)$. For 5 well-logs and 4 features per well log, the dimensions will reduce from 400 to 20.

- 2) *Optimal-feature selection.* For each feature, the difference between the upper window value and the lower window value will generate a difference, df . The analysis of the variance of df leads to the ranking of the features by relevance. The features having df vectors with the smallest coefficient of variation, dC_V , are the most desirable features for boundary determination. Further PCA on the training set will lead to the optimal number of features nf_{opt} . The optimal features will consist of the nf_{opt} features with the smallest dC_V s. This process will have two goals: select the optimal features and reduce dimensions. The dimension are reduced from $nl \times nf$ to nf_{opt} , where $nf_{opt} \in [2, d]$. The steps for selection of optimal features are described in section 5.3.5.2.
- 3) *PCA.* Applying PCA on the optimal feature matrix, F , and selecting the first two PCs, will reduce the dimensions from nf_{opt} to 2.

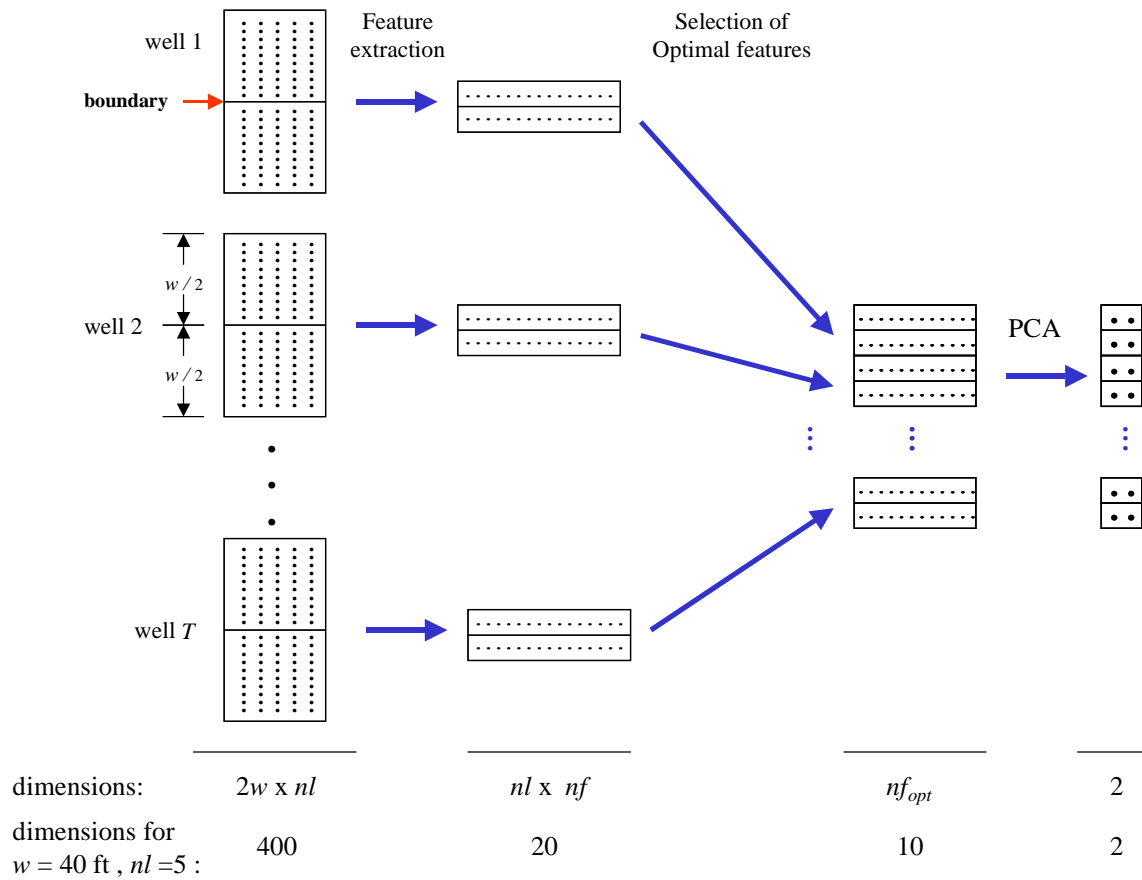


Figure 5.8 Dimensionality reduction process. The example for 5 well logs, window size of 40 ft, 2 samples/ft, and four features per log, demonstrates how the 400-dimensional input matrix is reduced to 20 dimensions with feature extraction; then to 10 dimensions by selection of optimal features; and lastly to 2 dimensions by PCA.

The PCA for the features at the T2/ K1U1 boundary proceeds as follows. For the raw GR ($nl = 1$), and using all 10 wells in Apiay ($T = 10$):

- 1) Extract the four features (mean, CV, min/max, trend angle) at the boundary using a 40-ft window ($nf = 4$, $w = 40 \text{ ft}$). This will create 4 features for the upper section and 4 features for the lower section. For a single signal, the dimension $d = nl \times nf = 1 \times 4 = 4$. For each type of feature, create the feature vector for the

upper and lower sections. The resulting matrix Ft will be of size $2T \times d$ (20×4). (Table 5.1).

- 2) Compute the mean and standard deviation (sd) for each feature vector (Table 5.1).

Table 5.1 Features for raw GR at the true K1U1 boundary for a 40-ft window.

	well	mean	Cv	min/max	trend angle
upper features	A-03	0.602	0.334	0.418	0.298
	A-04	0.616	0.425	0.308	0.265
	A-05	0.441	0.519	0.211	0.254
	A-07	0.449	0.469	0.195	0.214
	A-09	0.366	0.450	0.201	0.336
	A-10	0.386	0.556	0.143	0.113
	A-11	0.564	0.358	0.219	0.251
	A-13	0.552	0.332	0.392	0.296
	A-15	0.450	0.307	0.404	0.336
	A-16	0.386	0.552	0.142	0.086
lower features	A-03	0.388	0.711	0.014	0.785
	A-04	0.370	0.750	0.000	0.743
	A-05	0.371	0.682	0.077	0.740
	A-07	0.387	0.779	0.026	0.759
	A-09	0.486	0.613	0.037	0.760
	A-10	0.353	0.929	0.011	0.772
	A-11	0.440	0.668	0.036	0.783
	A-13	0.339	0.854	0.000	0.754
	A-15	0.461	0.656	0.073	0.825
	A-16	0.379	0.841	0.020	0.793
	mean	0.439	0.589	0.146	0.508
	sd	0.085	0.189	0.142	0.277

- 3) Standardize the feature vectors by subtracting the mean and dividing by the standard deviations calculated in step 2 above. This will result in feature vectors with zero mean and standard deviation of 1. The standardized feature vectors are presented in Table 5.2.

4) Compute the covariance matrix for the normalized feature vectors (Table 5.3).

Table 5.2 Normalized features for raw GR at the true K1U1 boundary for a 40-ft window.

		Standardized features			
well		mean	Cv	min/max	trend angle
upper features	A-03	1.929	-1.352	1.909	-0.760
	A-04	2.088	-0.867	1.137	-0.879
	A-05	0.018	-0.372	0.451	-0.918
	A-07	0.117	-0.635	0.342	-1.063
	A-09	-0.865	-0.737	0.385	-0.621
	A-10	-0.633	-0.178	-0.024	-1.426
	A-11	1.475	-1.227	0.511	-0.929
	A-13	1.328	-1.362	1.724	-0.765
	A-15	0.131	-1.492	1.810	-0.620
A-16	-0.633	-0.198	-0.028	-1.523	
lower features	A-03	-0.611	0.645	-0.933	1.000
	A-04	-0.817	0.849	-1.028	0.849
	A-05	-0.809	0.491	-0.485	0.836
	A-07	-0.621	1.002	-0.846	0.907
	A-09	0.550	0.125	-0.766	0.908
	A-10	-1.020	1.799	-0.951	0.953
	A-11	0.010	0.417	-0.778	0.992
	A-13	-1.182	1.402	-1.028	0.887
	A-15	0.260	0.355	-0.512	1.144
A-16	-0.716	1.334	-0.891	1.026	
mean		0	0	0	0
sd		1	1	1	1

Table 5.3 Covariance matrix of standardized feature vectors.

	mean	Cv	min/max	trend
mean	1.000	-0.728	0.720	-0.421
Cv	-0.728	1.000	-0.916	0.783
min/max	0.720	-0.916	1.000	-0.741
trend	-0.421	0.783	-0.741	1.000

- 5) Compute the eigenvectors and eigenvalues of the covariance matrix. The eigenvectors ordered in decreasing eigenvalues (Table 5.4) constitute the principal components (PCs). The first and second eigenvalues explain 79.4 and 14.4% of the variance respectively. The first two PCs explain 93.8% of the variance of the data, which is good enough to perform further analysis (Jolliffe, 2002, Chapter 6). The PCs will be contained in matrix PC .
- 6) Transform the boundary-feature matrix Ft to its corresponding z -score matrix Zt . The z -scores are the transformed data resulting from PCA. Zt is obtained by matrix multiplication of the standardized feature vectors and the matrix of PC coefficients (PC). Thus, $Zt = Ft \times PC$ (Table 5.5).
- 7) The visual representation of the first two principal components (first two z -score vectors) is shown in Figure 5.9. There is a very clear separation between the upper and lower features.

Table 5.4 Eigenvalues and Eigenvectors of covariance matrix in order of decreasing eigenvalues. The Eigenvectors represent the coefficient of the PCs.

	PC1	PC2	PC3	PC4
	Coefficients			
	-0.449	0.742	-0.489	-0.097
	0.543	0.042	-0.277	-0.791
	-0.536	0.002	0.612	-0.582
	0.465	0.670	0.557	0.160
eigenvalues	3.174	0.581	0.166	0.079
Variation, %	79.4	14.4	4.2	2.0

Table 5.5 Z-score transform of the feature vectors (PCs).

	well	PC1	PC2	PC3	PC4
upper features	A-03	-2.976	0.869	0.177	-0.350
	A-04	-2.427	0.927	-0.574	-0.319
	A-05	-0.879	-0.615	-0.141	-0.117
	A-07	-1.075	-0.651	-0.264	0.122
	A-09	-0.506	-1.087	0.517	0.344
	A-10	-0.463	-1.431	-0.450	-0.012
	A-11	-2.035	0.421	-0.585	0.381
	A-13	-2.615	0.420	0.357	-0.178
	A-15	-2.127	-0.378	1.112	0.016
	A-16	-0.517	-1.498	-0.500	-0.009
lower features	A-03	1.590	0.242	0.106	0.252
	A-04	1.773	-0.004	0.008	0.142
	A-05	1.279	-0.021	0.428	0.106
	A-07	1.699	0.187	0.013	-0.095
	A-09	0.653	1.019	-0.266	0.439
	A-10	2.388	-0.044	-0.052	-0.618
	A-11	1.100	0.688	-0.044	0.280
	A-13	2.256	-0.225	0.054	-0.255
	A-15	0.883	0.972	0.099	0.175
	A-16	2.000	0.211	0.007	-0.303

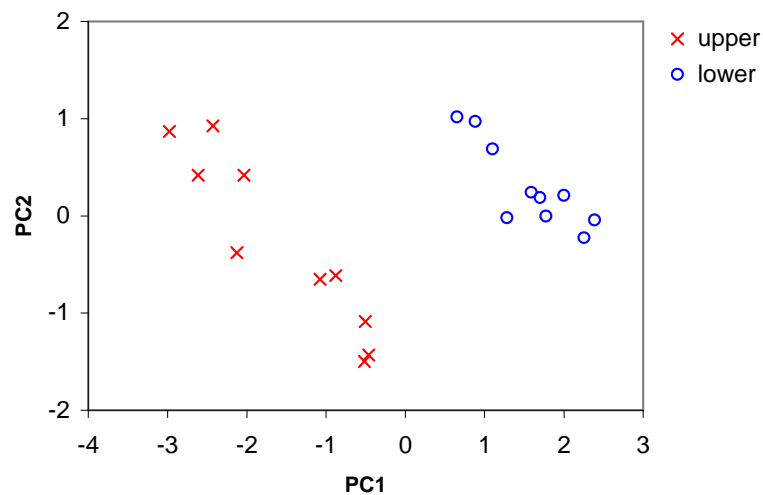


Figure 5.9 PCA for theGR at T2/K1U1 boundary, window size = 40 ft (red crosses upper, blue circles lower part).

5.3.5 Depth selection

In our research we perform the boundary recognition assuming that the various boundaries along the well are independent events. Thus, our process operates on one boundary at a time. The steps involved in boundary detection are as follows:

- 1) Choose well-logs to be used in the pattern recognition and perform smoothing operations on each signal. The signals should include the GR or SP. The number of well-logs, nl , times the number of features per signal, nf , will determine the total number of features $d = nl \times nf$.
- 2) For each one of the observation wells (wells where the boundary is to be identified), select the depth range spanning all boundaries to be computed. The top depth is $depth_t$ and the bottom depth is $depth_b$. Here, we are performing the pattern recognition with precision of 1 ft. Thus, for the depth range $[depth_t, depth_b]$ the total number of depth values is $H = depth_b - depth_t + 1$.
- 3) Choose the window length, w , in depth units. The window will be equally split into upper and lower sections with size $w/2$ each.
- 4) Determine the number of depth points in the observation well as $M = H - w$. The resulting depth vector $r \in [depth_t + w/2, depth_b + w/2]$ with length M .
- 5) For each depth point of r compute the depth-upper and-lower features. This will generate a $2M \times d$ matrix O for the observation well. The upper features are stored in rows 1 to M , and the lower features in rows $M + 1$ to $2M$.
- 6) Select desired boundary.
- 7) For the requested boundary and the observation well, select a search-depth range. The pattern recognition algorithm will assign the boundary depth within this range. In our analysis, we used search-depth ranges with size of 200 ft.
- 8) Define the training wells (T wells). This set in general will consist of the best characterized wells, where the boundaries depths, b_t ($t = 1, 2, \dots, T$), have been defined.
- 9) Extract the boundary-upper and -lower features for each one of the training wells. The boundary-upper features will be computed for the depth range $[b_t - w/2, b_t]$,

and the boundary-lower features will be processed for the depth range $[b_t, b_t + w/2]$. This will result in a $2T \times d$ training-feature matrix F_t . The upper features are stored in rows 1 to T , and the lower features in rows $T+1$ to $2T$.

- 10) Select the optimal features by variance analysis on matrix F_t . This will produce the optimal number of features nf_{opt} . The optimal training-feature matrix F will be of size $2T \times nf_{opt}$.
- 11) Perform dimension reduction on the training wells by applying principal component analysis (PCA) on the matrix F . This will generate the coefficients of the principal components (PCs) contained in the square matrix PC of size $d \times d$. Transform matrix F to Z_t (z-score matrix) by $Z_t = F \times PC$. By choosing the first two PCs (PC1 and PC2), the resulting z-score matrix Z will be of size $2T \times 2$.
- 12) Using the first two PCs, transform the feature matrix O in the observation wells to generate the matrix of z-scores $O_z = O \times PC$.
- 13) Compute the probability that the upper-z-scores at each depth point belongs to the boundary-upper z-scores. Perform similar probability calculation for the lower z-scores. The total probability is obtained by multiplying the probability of the upper z-scores by the probability of the lower z-scores.
- 14) Assign the boundary on the observation wells at the highest-probability depth.
- 15) Repeat steps 6 to 14 for other boundaries.

Figure 5.10 shows the shifting process involved in the boundary detection in an observation well. The window length is the same as the one defined for the training set. Features at each depth point (usually every foot) are calculated, the window is shifted one foot and feature extraction is repeated until covering the whole depth range.

Different approaches can be implemented for testing of the pattern recognition procedure. The first approach is to select a well with high interpretation confidence, for example A-11 well (cored well and good seismic match) or A-15, and then attempt to recognize the boundaries in all other wells. A second approach is to use the information from multiple wells (two or more) and recognize the boundaries for all other wells. For

single-well training set, steps 10 and 11 are not required, and step 12 is performed on the feature vectors directly.

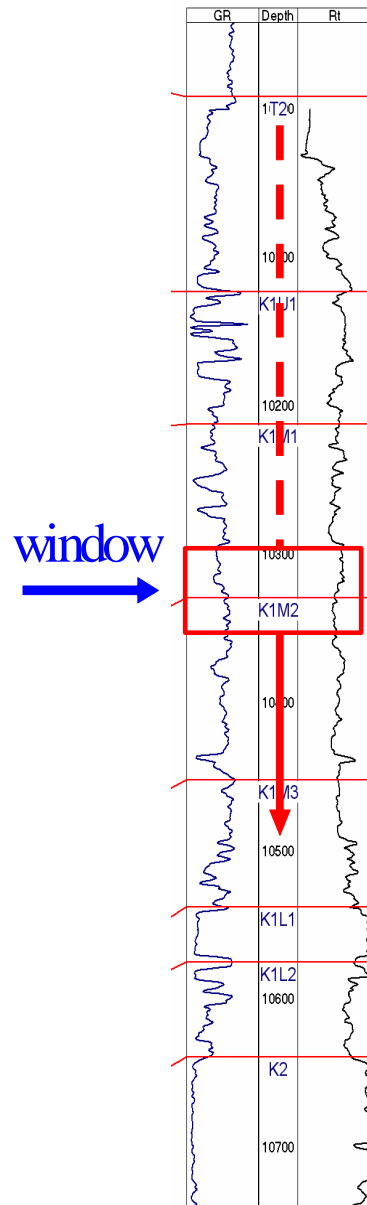


Figure 5.10 Feature extraction in observation well. Features are extracted at each depth point using the window length, w . The window is shifted along the depth range. The boundary is assigned to the depth of the midpoint of the window with the features closest to those of the training set.

In practice, the boundary recognition can be performed starting from the information of very well characterized wells (the training set). This task is accomplished by comparing the boundary features of the training set with the features at a given depth in the test well. A probability value is assigned at each depth. After scanning all depths in the test well a probability curve as a function of depth is obtained. The depth with the highest probability is chosen as the likeliest location for the boundary.

Increasing the number of wells in the training set gives more insight on the characteristics of the features representing the upper and lower sides of a geological boundary. We address the single well training set because this is a common task in boundary identification. The approaches for single well and multiple-well training sets vary slightly.

5.3.5.1 Single-well training set

Consider the case of d extracted features. Treating the upper and lower sections of the boundary as two independent features, the resulting boundary features f_k where $k = 1, 2, \dots, 2d$, define the boundary for the single-well training set. At each m depth point (where $m = 1, 2, \dots, M$) in the observation well, we have test features t_{mk} where $m = 1, 2, \dots, M$; and $k = 1, 2, \dots, 2d$. The error row vector ε_m at each depth point is given by

$$\varepsilon_m = t_m - f \quad (5.17)$$

To illustrate the case of the single-well training set, we selected well A-15, the GR, and the T2/K1U1 boundary (Example 1). There are four extracted features, resulting in eight boundary features for the two half windows. The boundary identification in well A-03 in the interval 10,000 to 10,793 ft results in a matrix error of size 794 x 8. The boundary features of the training well A-15 are shown in table 5.1. Figure 5.11 displays the errors for two features. In the vicinity of well A-03 true boundary (10,135 ft) error values of zero are obtained at 10,127, 10,154, and 10,227 for the upper mean (Fig. 5.11a). For the lower C_v , zero errors are obtained at 10,134, 10,175, and 10,208 ft. Additional zero errors are obtained at depths more than 400 ft below the true boundary.

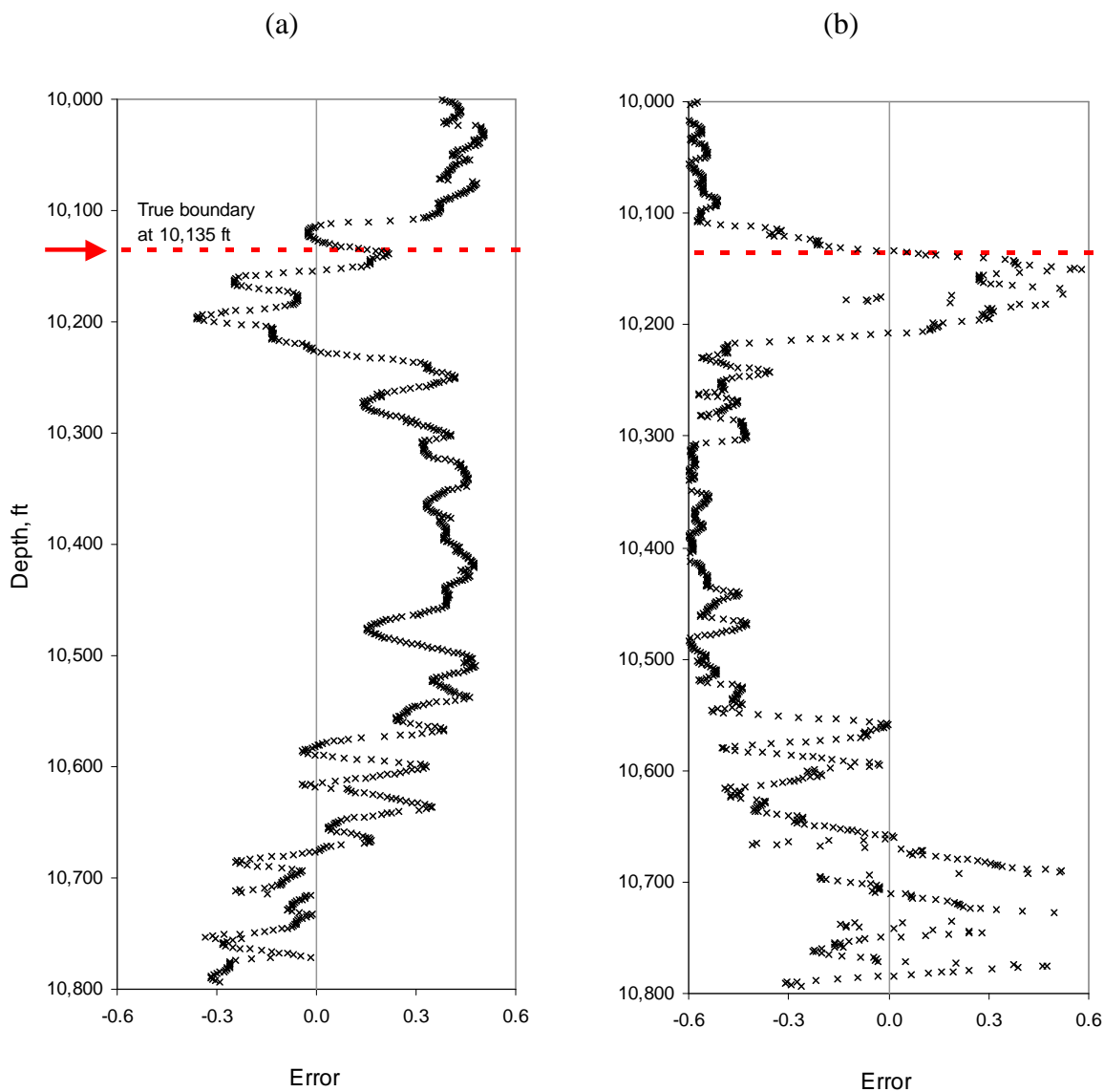


Figure 5.11 Error distribution for well A-03 for the T2/K1U1 boundary detection, using the raw GR. The training well is A-15 (Example 1). The features shown are: (a) upper mean and (b) lower C_V .

To detect the boundary we need to model the conditional probability p for all depths given the feature vector f at the boundary of the training well. Assuming that the distributions of the different features are independent events, the probability that the

feature vector being tested, t_m (at depth point m), corresponds to the feature vector of the training vector f is given by

$$p(t_m | f) = \prod_{k=1}^{2d} p(t_{mk} | f) \quad (5.18)$$

Assuming that the errors ε_{ik} have a normal distribution with zero mean and standard deviation σ , the probability distribution required in Equation 5.18 is given by

$$p(t_{mk} | f) = \frac{1}{(2\pi\sigma_k^2)^{1/2}} \exp\left(\frac{-\varepsilon_{mk}^2}{2\sigma_k^2}\right) \quad (5.19)$$

Figure 5.12 shows the normalized probabilities computed for all wells for Example 1. Two wells (A-07 and A-16) show the highest probability at depth values more than 400 ft below the true boundary. In practice, it is obvious that the T2/K1U1 boundary cannot be located at such depths. To address this we can impose a search range within the vicinity of the true boundary similar to those used in field applications. Performing the analysis on a search range of 200 ft (100 ft above and 100 ft below the true boundary) we find that using A-15 as the training well (Example 1) will result in proper detection (error in boundary detection less than 10 ft) on 7 out of 9 wells (78%).

The other two wells A-07 and A-16 give an error of 21 ft both. Adding the R_t to the analysis reduces the proper detection to 67%. Table 5.6 presents the proper detection results (error less than 10 ft) using different signals and combinations. For the T2/K1 boundary, the GR is by far the best signal for boundary detection. The combination of GR and the other signals (5 signals in total) improved the detection from 78 to 89%. For the K1/K2 boundary the best performance is achieved by combining GR and R_t . Still the boundary identification at this boundary is not as good as the T2/K1. These results are consistent with the integrated reservoir study, in which there is a high degree of

confidence in the determination of the T2/K1 boundary, but less confidence in the determination of the K1/K2 boundary, as described in section 5.2.

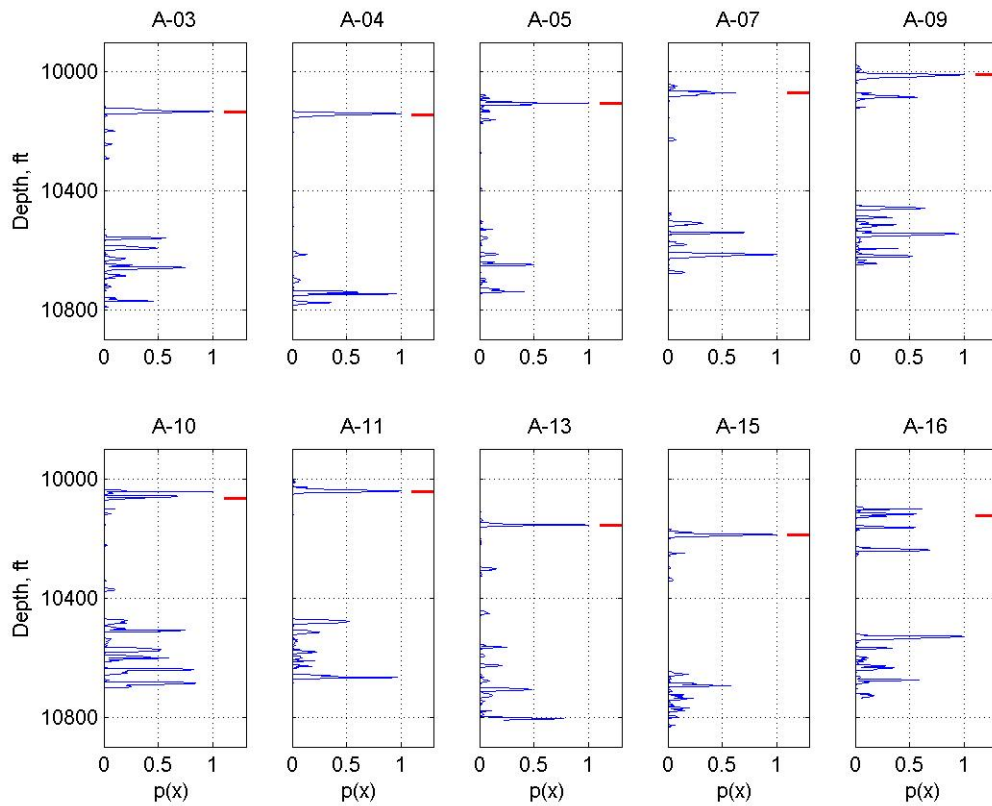


Figure 5.12 Normalized probability distributions for the T2/K1U1 boundary detection, using the raw GR. The training well is A-15. The true boundary depths are indicated with a tick on the right hand side.

Table 5.6 Proportion of boundaries identified within 10 ft of the true position, using A-15 as the training.

Signal	Detection performance(%)	
	T2/K1U1	K1/K2
GR	78	11
R _t	11	11
R _{xo}	0	22
Phinl	33	0
Rhob	0	0
GR and R _t	67	33
GR, R _t , R _{xo} , Phinl, and Rhob	89	11

5.3.5.2 Multiple-well training set

Having a single well in the training set imposes restrictions in the analysis. There is no way to know if similar patterns are consistently present in other wells. Increasing the number of wells in the training set can contribute to more confidence in the boundary determination. However, the training set should include the best characterized wells, to avoid the risk of diminishing the detection ability of the process.

For the case of a single well, and based on current practices and the results presented in Table 5.6, the GR features, in general, should be considered first for boundary identification. Table 5.6 shows that adding more signals could lead to deterioration of the boundary identification (T2/K1 boundary) or improvement (K1/K2 boundary). For the optimal boundary identification we need to know in advance the type of well-logs or combinations that would perform most efficiently. First, we will develop the process for raw signals only and then we will expand to include wavelet transformation. As an example we will use K1/K2, which is more difficult to detect than T2/K1.

Using all the signal available for Apiay field (GR, R_t , R_{x0} , Phinl , and R_{hob}), and extracting four features from each raw signal (mean, CV, min/max, and trend angle) will result in 20 feature vectors per upper or lower window (Figure 5.13). The features that better contribute to the proper boundary identification are the ones showing consistent class separation between the upper and lower portions of the boundary. By visual inspection, we can recognize the GR mean and the GR-trend angle as among the best features.

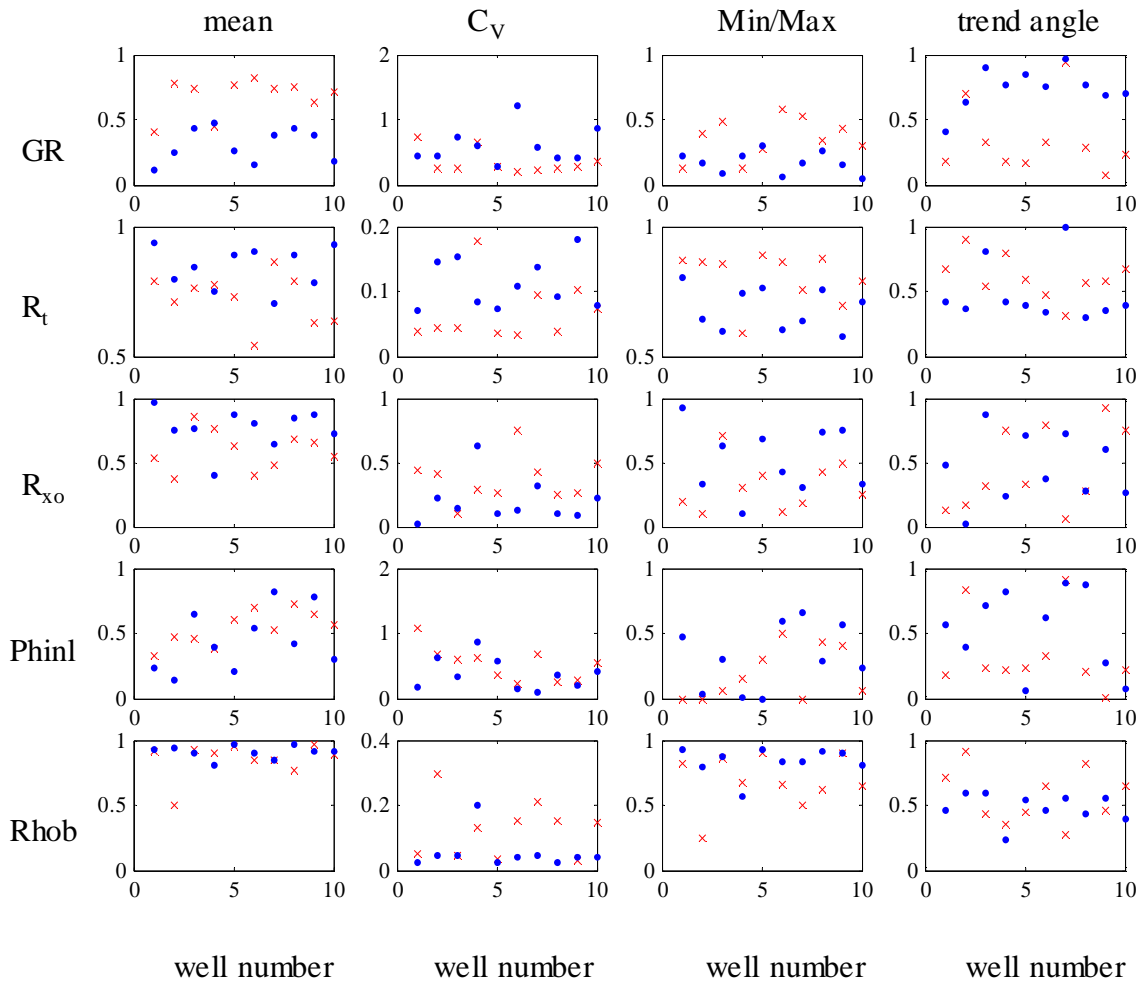


Figure 5.13 Feature vectors for raw signals in Apiay field at the K1L2/K2 boundary. Dots represent lower window features and crosses upper window features.

The dimension reduction is performed using PCA. Figure 5.14 show the first two PCs for the K1/K2 boundary using different combinations of signals. Using the 5 signals (20 features) results in poor class separation of the upper and lower sections of the boundary (Fig. 5.14c) as compared with the 4 GR features (Fig. 5.14a). Adding the 4 Rt-features to the GR-features does not improve the class separation (Fig. 5.14b). We need to develop a method that identifies the maximum class separation between the upper and lower-window features (fig 5.14d).

The pre-selection of features for maximum class separation can be accomplished by computing the coefficient of variation of the difference between the upper and lower window values of each feature, dC_V . If this difference is relatively constant for all wells in the training set, dC_V is small. On the other hand, when the difference is not constant or changes sign frequently, dC_V is high.

The process to determine the optimal features to be used proceeds as follows:

- 1) Create the feature differential vectors by subtracting the lower window value from the upper window value.
- 2) Compute dC_V for each differential vector.
- 3) Order the features in ascending values of dC_V . The best features are the ones with the smallest dC_V s.
- 4) Take the best two features and determine the PCs as described in Section 5.3.4 above.
- 5) Use the upper Z-score transforms (first two PCs) and calculate for each well in the training set the multidimensional distance from the well to the training set (Mahalanobis distance, defined below). This creates a vector of upper-Mahalanobis distances.
- 6) Create the lower-Mahalanobis-distance vector using the lower Z-score transforms.

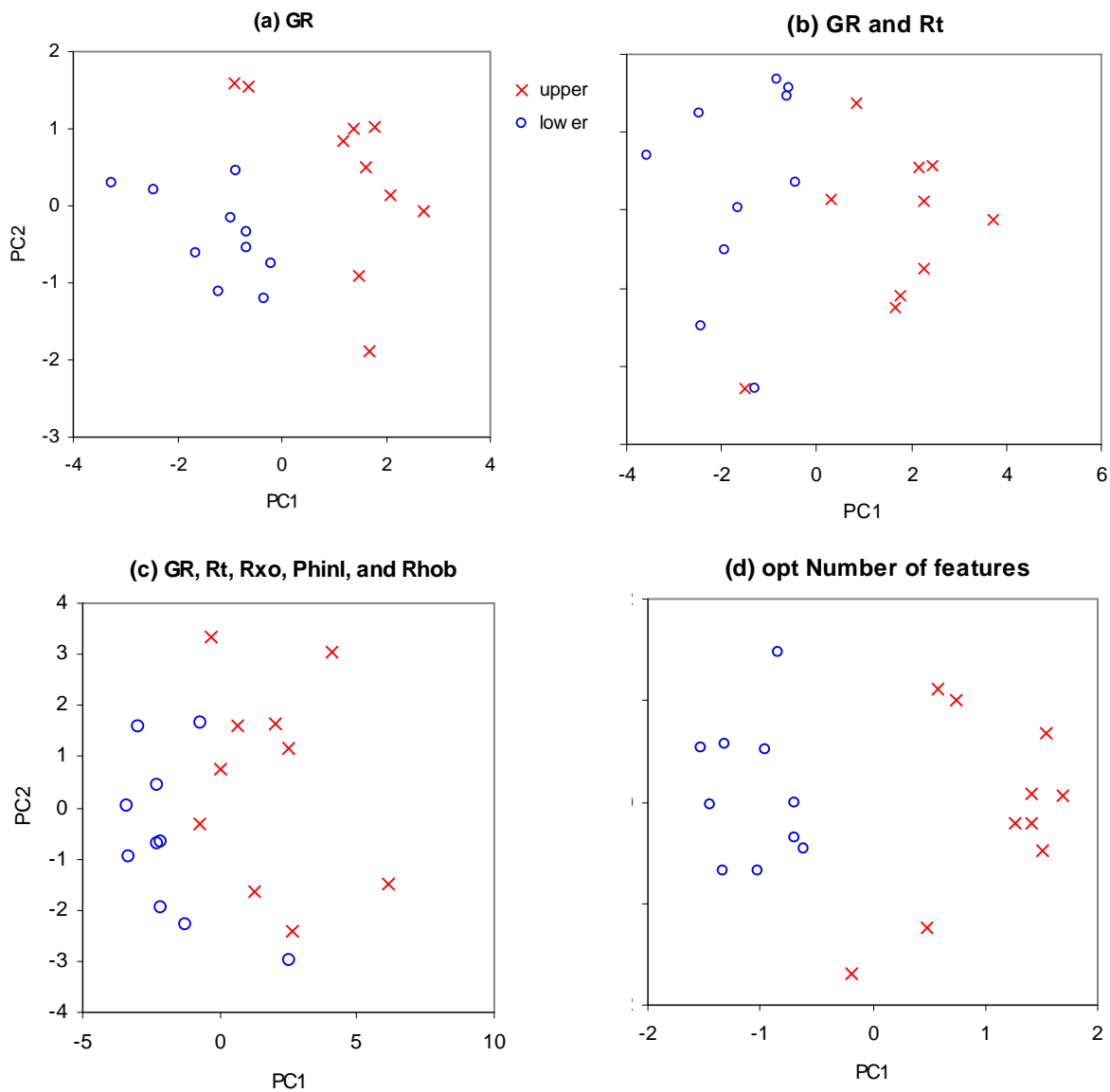


Figure 5.14 Principal components for different combinations of features for the raw signals in Apiay field at the K1/K2 boundary. (d) represents the optimal features from the 5-signal 20 features. In this case, the optimal features are GR-mean and GR-trend angle.

- 7) Compute the total variance of multidimensional distances by adding the upper- and lower- mahalanobis-distance-variances and normalizing by the number of wells in the training set.

- 8) Take the best three features and compute total Mahalanobis distance variance as described in step 7.
- 9) Repeat steps 4 to 7 until using the feature with the largest dC_v .
- 10) Report the optimal number of features at the least total Mahalanobis distance variance.

For d -dimensional data, the Mahalanobis distance, Δ^2 , from vector \mathbf{x} to the d -dimensional mean vector μ is given by (Bishop, 1995)

$$\Delta^2 = (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \quad (5.20)$$

where

$$\mu = \mathcal{E}[\mathbf{x}]$$

$$\Sigma = \mathcal{E}[(\mathbf{x} - \mu)(\mathbf{x} - \mu)^T]$$

Here, $\mathcal{E}[\cdot]$ denotes the expectation and Δ^2 is the distance between an observation and a population mean, and is related to probability density functions. Surfaces of constant probability density are hyperellipsoids on which Δ^2 is constant. The principal axes of the hyperellipsoids are given by the eigenvectors of Σ (Bishop, 1995).

Figure 5.15 shows the dC_v s for the 20 features of the 5 signals for all wells in Apiay field at the K1/K2 boundary. The smallest dC_v is obtained with the GR-mean, followed by the GR-trend angle, R_t -Min/Max, and GR-Min/Max.

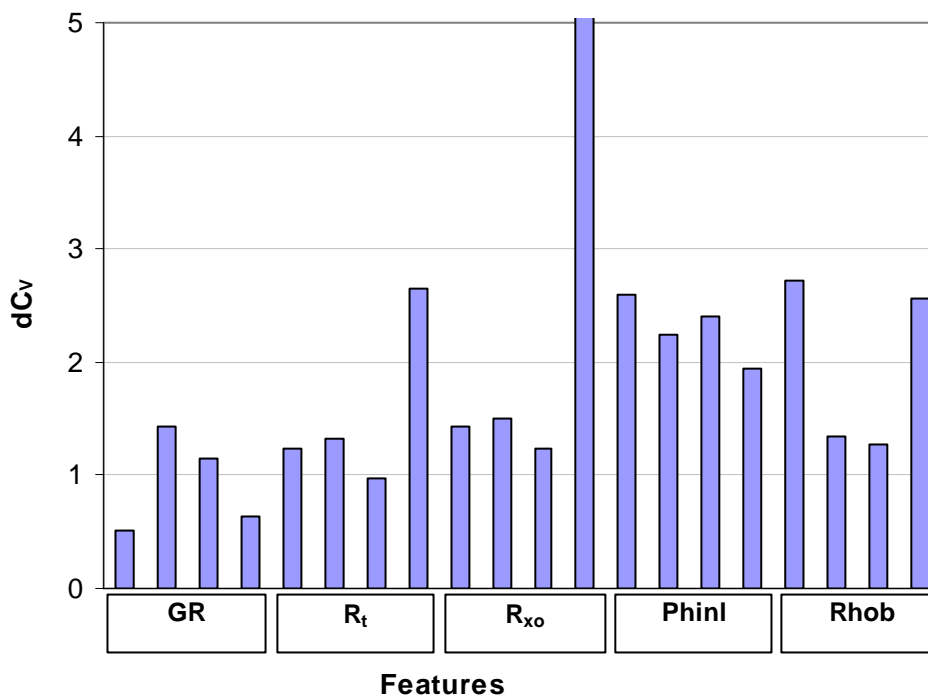


Figure 5.15 Differential-features-coefficient of variation, dC_v , of the raw-signal features in Apiay field at the K1/K2 boundary. For each signal the features from left to right are: mean, CV, Min/Max, trend angle.

To illustrate how the pre-selection of features affects the boundary identification we will use a training set of four wells to identify the K1/K2 boundary. For this case (Example 2) the training wells are A-05, A-11, A-13, and A-15. The dC_v s for the training set are shown in Figure 5.16. The smallest dC_v s are the GR-mean, followed by the R_t -CV, R_t -Min/Max, and GR-Min/Max. These coefficients differ from the ones obtained using all wells in the field (Fig. 5.15). For the purposes of boundary detection we must use the dC_v s from the training set.

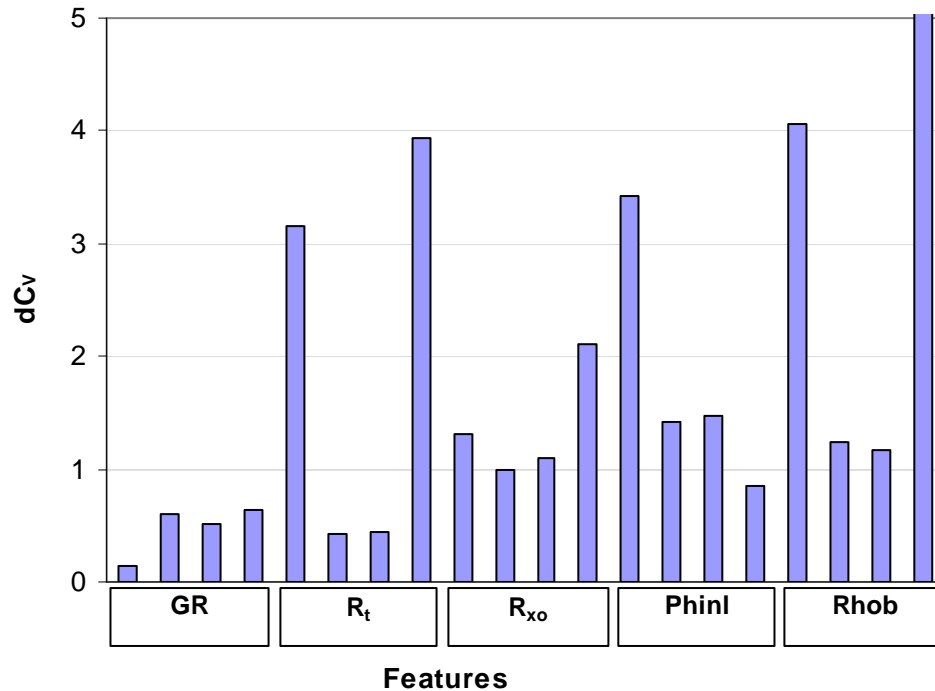


Figure 5.16 Differential-features-coefficient of variation, dC_v , of the raw-signal features for 4-training wells (A-05, A-11, A-13, A-15) at the K1/K2 boundary. For each signal the features from left to right are: mean, C_v , Min/Max, trend angle.

Figure 5.17 shows the dC_v and Mahalanobis distance variance as a function of the features ordered in ascending value of dC_v . The optimal number of features of 12 corresponds to the minimum Mahalanobis distance variance. Also, the dC_v values show a drastic slope increase after 12 features. Figure 5.18 shows the probability distributions using **a)** all the 20 features of the 5 signals, and **b)** using the optimal number of features (12 features with the smallest dC_v s). The pre-selection of optimal type and number of features generates better boundary identification. Without pre-selection of features, 33% of the boundaries in the observation wells are properly identified (within 10 ft of true boundary). The pre-selection of features increases the detection performance to 50%.

c

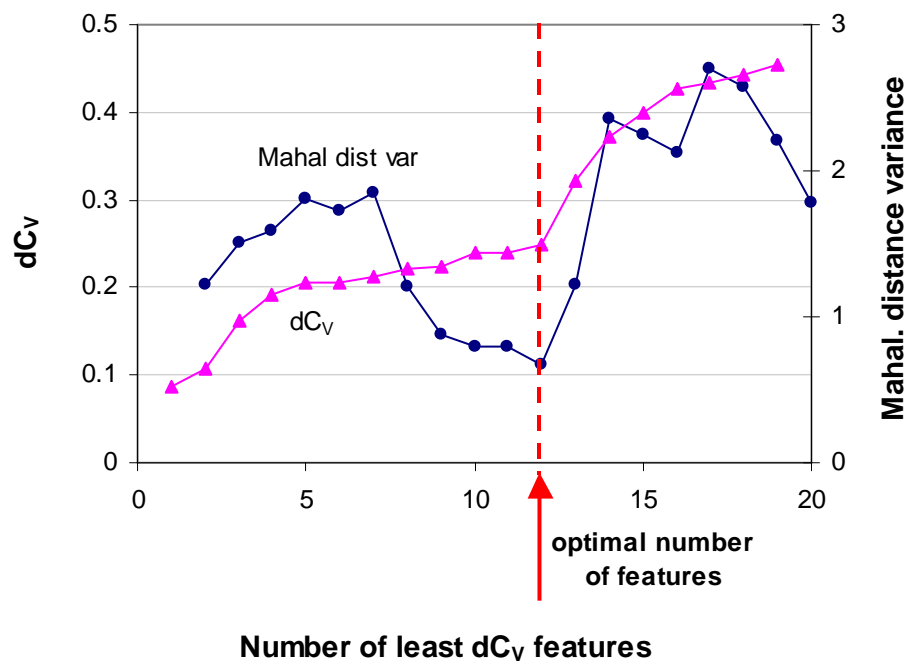


Figure 5.17 Differential-features-coefficient of variation, dC_V , and Mahalanobis distance variance for the training set A-05, A-11, A-13, A-15 at the K1/K2 boundary. The features are ordered in ascending dC_V values. The optimal number of features is 12, although 10 or 11 also give good performance.

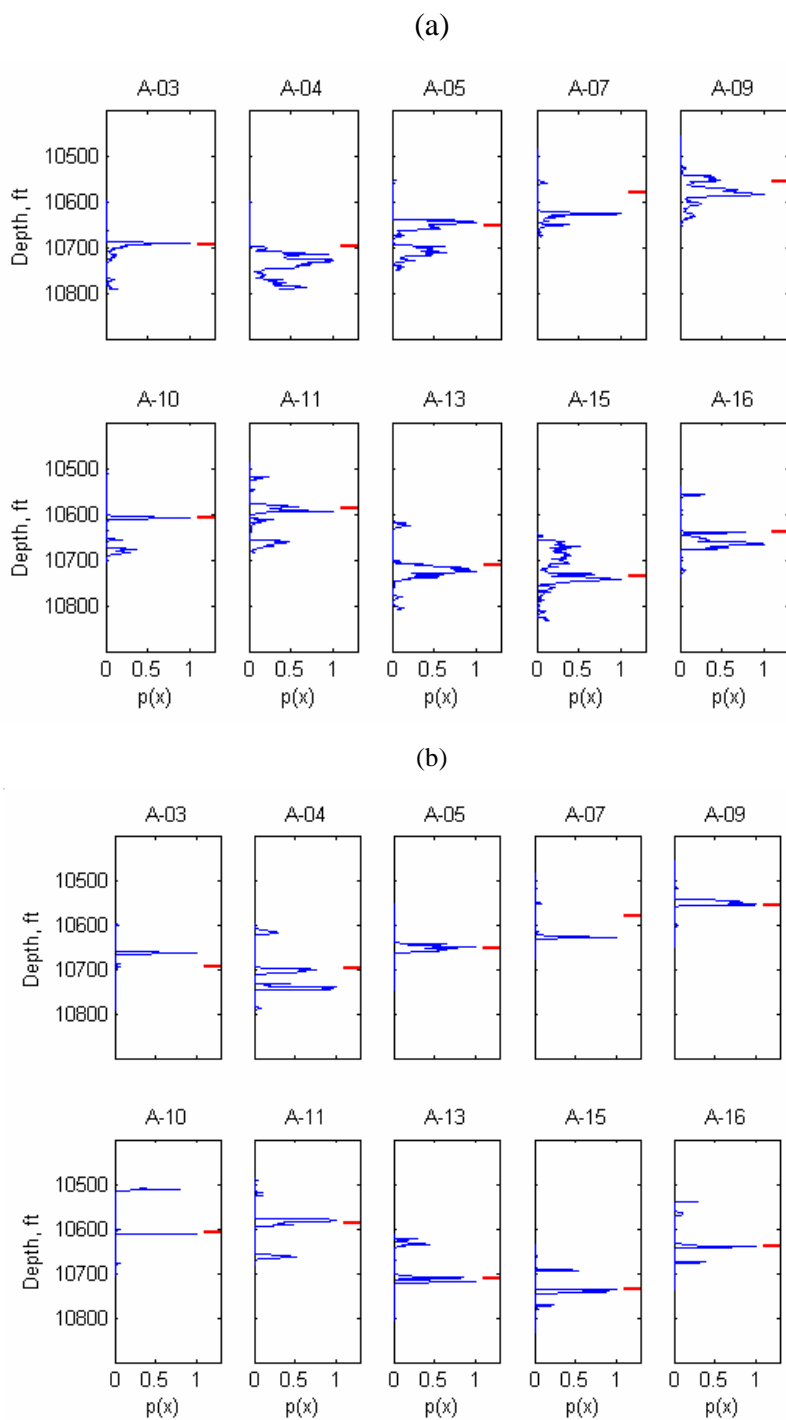


Figure 5.18 Normalized probability distributions for the K1/K2 boundary detection, using 5 well logs. The training wells are A-05, A-11, A-13, A-15, A-15. The true boundary depths are indicated with a tick on the right hand side. (a) shows the distribution using all 20 features and (b) displays the distribution using the optimal number of features (12).

For multiple-well training sets, the probability distribution at each depth m is given by

$$p_m^s(x_m^s) = \frac{1}{(2\pi)^{d/2} |\sum_s|^{1/2}} \exp\left(-\frac{1}{2} \Delta_{m,s}^2\right) \quad (5.21)$$

where the index $\{s = upp, low\}$ represents the upper and lower sections of the window boundary.

Equation 5.21 gives the probability that z -transform of the features of the upper or lower sections at the test depth (vector x) corresponds to the Z -score transforms of feature vector of the true boundaries in the training set. The training-set Z -scores have the mean vector μ with d -dimensions (equal to the number of PCs chosen for analysis). Δ^2 is the Mahalanobis distance from vector \mathbf{x} to the d -dimensional mean vector μ . $|\sum_s|$ is the determinant of \sum_s , a $d \times d$ covariance matrix of the training set Z -scores.

5.4 Effect of number of signals

In section 5.3.5.1 we introduced the expansion on the number of signals to the analysis for single-well training. The addition of signals improved the detection performance for both the T2/K1U1 and the K1L2/K2 boundary. We expand that analysis to the 4-well training set of example 2 at the K1L2/K2 boundary. The detection results are shown in Table 5.7. The addition of the R_t signal improves the detection performance from 17% to 50%. The same occurs when using all 5 signals (GR, R_t , R_{xo} , Phinl, and Rhob). The choice of signals to be used for the detection of the K1L2/K2 boundary is all five signals considering the fact that the mean detection error is smaller than the GR- R_t pair (14.4 vs. 24.1 ft). The pairs GR- R_{xo} , GR-Phinl, and GR-Rhob, do not change the detection performance achieved by the GR only. This is because the optimal selection of features include only the 4 GR features. However, when using all 5 signals the optimal selection of features amounts to 12. The selected features are: 4 GR-features, 2- R_t features, 3 R_{xo} -features, 1 Phinl-feature, and 2 Rhob-features (Fig. 5.16).

Here, we are picking only the highest probability event in the 200 ft interval. The performance would be better if we allowed the consideration of events with smaller probabilities. For instance, if we pick the second highest probability event for the K1/K2 boundary, we would have proper detection for well A-04 (Fig. 5.18).

Table 5.7 Detection performance and mean detection error at the K1L2/K2 boundary.
Training wells: A-05, A-11, A-13, A-15.

Signal	Detection performance(%)	Mean detection error (ft)
GR	17	50.3
GR and R_t	50	24.1
GR and R_{x_0}	17	50.3
GR and $Phinl$	17	50.3
GR and R_{hob}	17	50.3
GR, R_t , R_{x_0} , $Phinl$, and R_{hob}	50	14.4

5.5 Effect of wavelet decomposition

Wavelet transformation of 1-d signals has been used extensively for pattern recognition (Goswami and Chan, 1999). There are two approaches to wavelet transform: the continuous wavelet decomposition (cwt) and the discrete wavelet decomposition (dwt). The most common technique for pattern recognition is the dwt. Here, we perform pattern recognition using both dwt and cwt. The basic concepts of wavelets and the procedures for wavelet decomposition are discussed in Chapter III.

5.5.1 Discrete wavelet transform (dwt)

The discrete wavelet decomposition is a function of the choice of wavelets and the levels of decomposition, nd . The Daubechies family of wavelets has been used extensively because its wavelets coefficients capture the maximum amount of the signal energy (Goumas *et al.*, 2002). The Daubechies family is referred to as D_n , where n is the size of the mother wavelet. In our application we found that D_4 with four levels of decomposition ($nd = 4$) captures the characteristics of the well-logs in Apiay. Consequently all the analyses presented in this chapter are performed using Daubechies 4 (D_4) and 4 levels of decomposition. Figure 5.19 displays the approximation and details for the GR in well A-15. The approximation represents the low frequency reconstruction of the signal. The details represent the high frequency components of the signal. The frequency decreases with the details. From detail 4, the boundaries T2/K1U1 and K1L2/K2 are evident. Detail 3 captures the K1U1/K1M1 and K1M3/K1L1 boundaries. Other boundaries are less evident. The approximation distinguishes the K1U1 unit from the others. One additional use for the approximation is the representation of the general trend of the signal. The general representation can be used as a template for what could be expected in neighboring wells while acquiring the data.

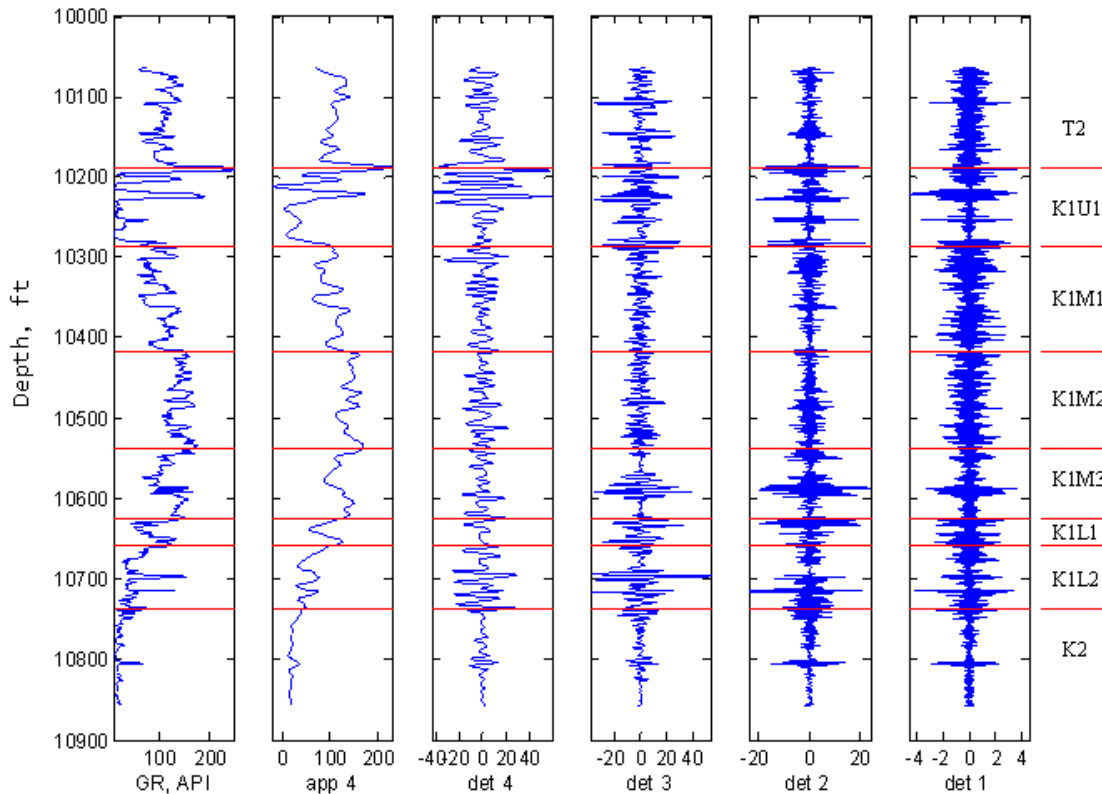


Figure 5.19 dwt on the GR of well A-15, using D4 and 4 levels of decomposition. Shown are the signal, the approximation and the 4 details.

The features extraction from the dwt can be performed similarly to the “raw” signal (un-processed log) described previously. The approximation and details now represent a new set of input data. For nd levels of decomposition we are increasing the number of input signals from 1 (raw signal) to $nd + 2$ signals (4 levels of decomposition plus the approximation and the raw signal). By extracting the same nf features from the decomposition, we will have a total of $nf \times nd + 2$ features for each well-log. If we take the GR and 4 levels of decomposition we will end up with 24 features (24-dimensional input). Figure 5.20 displays the features for the GR at the K1L2/K2 boundary. Figure 5.21 shows the dC_V for all features, and Figure 5.22 presents the dC_V s ordered for

selection of optimal features. The optimal number of features at this boundary is 6. The best feature is the app-mean. The other features selected for pattern recognition in increasing value of dC_V are: GR-mean, GR-trend, app-trend, det1-CV, and GR-Min/Max.

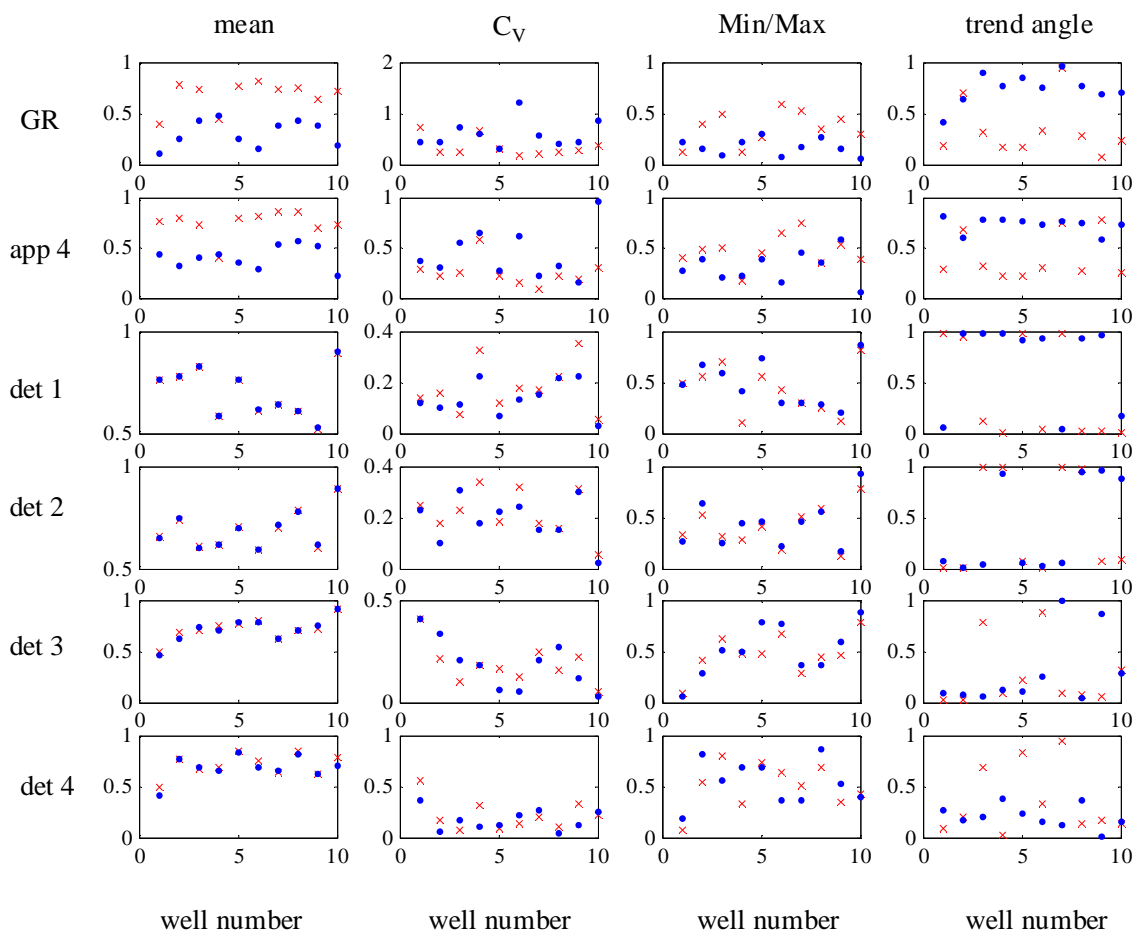


Figure 5.20 Features for the GR, approximation, and details at the K1L2/K2 boundary in Apiay.

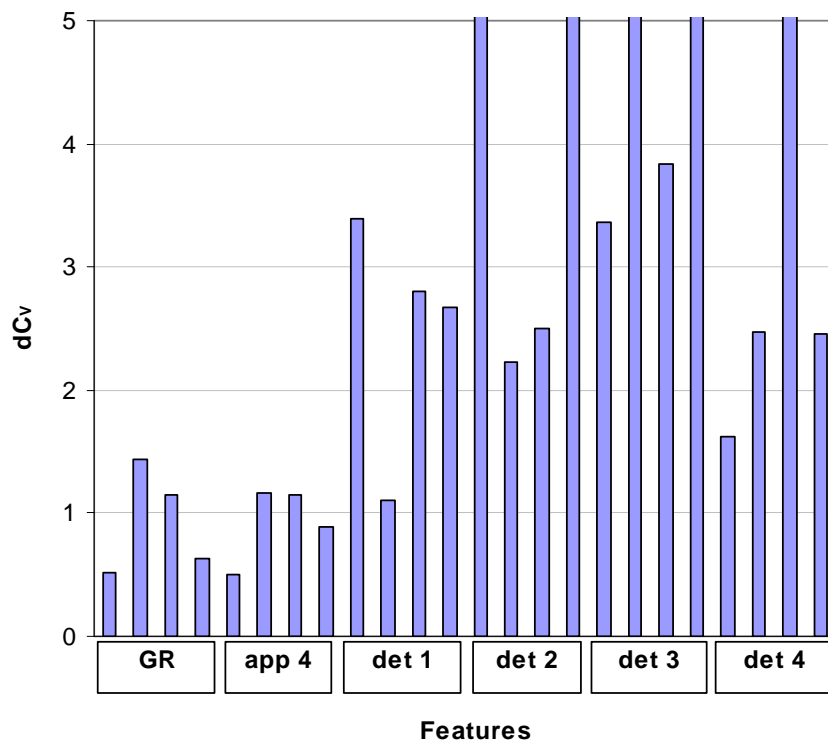


Figure 5.21 Differential-features-coefficient of variation, dC_v , for the GR, approximation and 4 levels of detail at the K1/K2 boundary, using all wells in Apiay. For each signal the features from left to right are: mean, CV, Min/Max, trend angle.

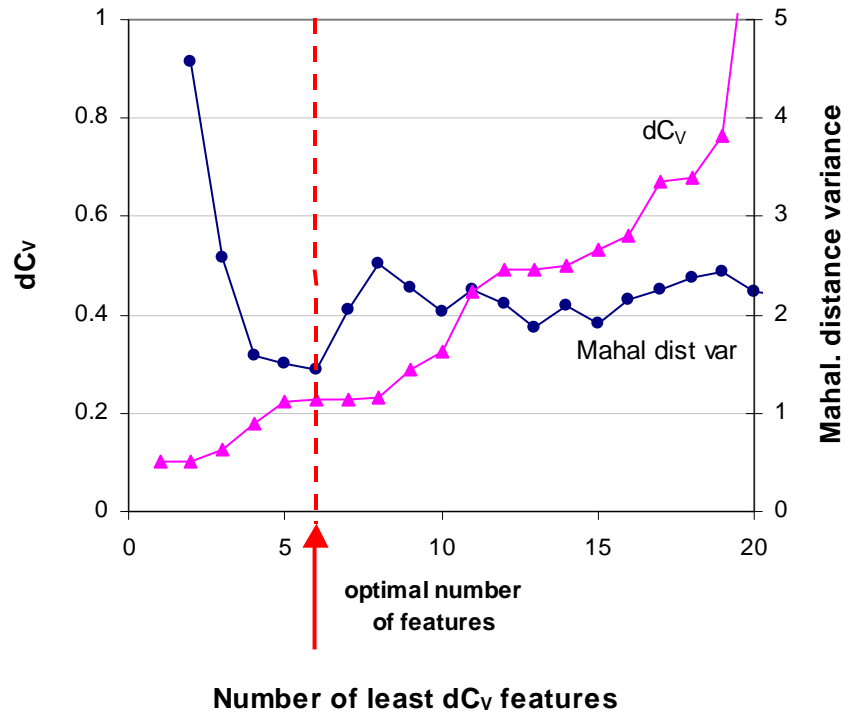


Figure 5.22 Differential-features-coefficient of variation, dC_v , and Mahalanobis distance variance for all wells at the K1/K2 boundary, using the GR and 4 levels of decomposition. The features are ordered in ascending dC_v values. The optimal number of features is 6.

Performing the pattern recognition for example 2 at the K1L2/K2 boundary, and using the discrete wavelet decomposition, generates the detection performance shown in Table 5.8. By comparing these results with the ones obtained using just the raw signal (Table 5.7) we can make the following observations:

- 1) The use of dwt improves the detection for the GR
- 2) For the six combinations of signals, dwt improves the detection in 4 of the cases (67%).
- 3) The best detection for this example is achieved by using the raw signals only, i.e., not using the dwt.
- 4) The optimization of features will select the best combination of signal and determine if wavelet decomposition is required for better detection.

Table 5.8 Detection performance and mean detection error at the K1L2/K2 boundary, using dwt on the well-logs. Training wells: A-05, A-11, A-13, A-15.

Signal and dwt	Detection performance(%)	Mean detection error (ft)
GR	33	37.6
GR and R_t	33	43.0
GR and R_{x0}	33	34.5
GR and Phinl	17	15.3
GR and Rhob	33	37.6
GR, R_t , R_{x0} , Phinl, and Rhob	17	32.1

5.5.2 Continuous wavelet transform (cwt)

The continuous wavelet transform was used in Chapter IV for cyclostratigraphy purposes. In this chapter, we are evaluating the cwt for pattern recognition of geologic boundaries. We use the Morlet wavelet. Fig 5.23 shows the GR- and R_t -scaleograms for well A-03. The GR-scaleogram captures the characteristics of the K1U1 unit and the separation with the overlying and underlying units T2 and K1M1. In addition, the GR-scaleogram distinguishes the K1L2/K2 boundary. The R_t -scaleogram is less revealing.

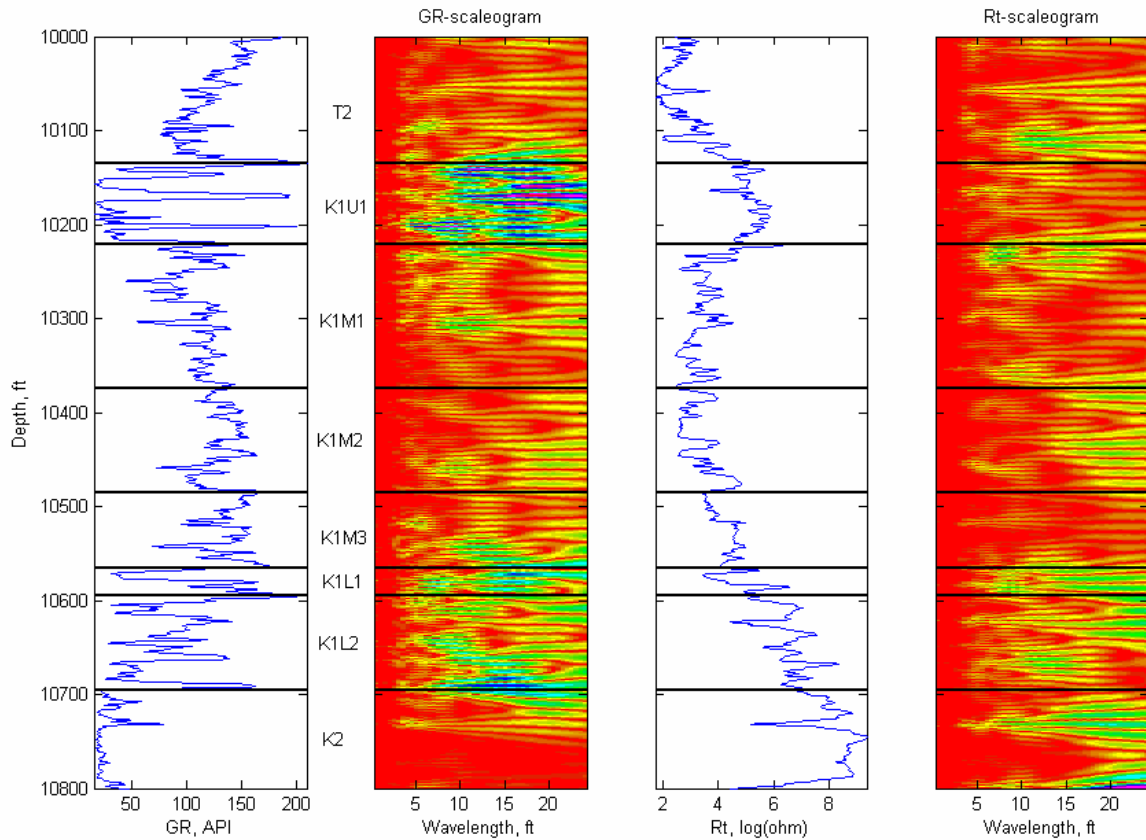


Figure 5.23 Morlet scaleogram for well A-03. The cwt is applied for the GR and R_t .

The cwt features for each one of the well-logs are generated from the scaleogram coefficients as follows:

- 1) The mean, C_v , Min/Max and trend angle of the raw signal will constitute the first 4 features
- 2) Extract six vectors of wavelet coefficients as a function of depth for the wavelengths: 4, 8, 12, 16, 20, and 24 ft (six different scales).
- 3) Define the window, w , as described before.
- 4) Compute the mean and C_v for the upper and lower section of the window. This will generate 12 wavelet-coefficient features, for a total of 16 features (16-dimensional input for each signal).

Figure 5.24 shows the cwt-GR features at the K1L2/K2 boundary and Figure 5.25 the dC_V for the various features. The optimal features (6) are the 4 features of the raw signal, the mean of the 4-ft wavelength-coefficients and the C_V of the 12-ft wavelength coefficients.

The detection performance in Example 2, using cwt, is presented in Table 5.9. The detection performance improves with the cwt compared to the raw signals alone. The combination of all 5 signals and cwt generates the highest detection performance (67%) at the K1L2/K2 boundary.

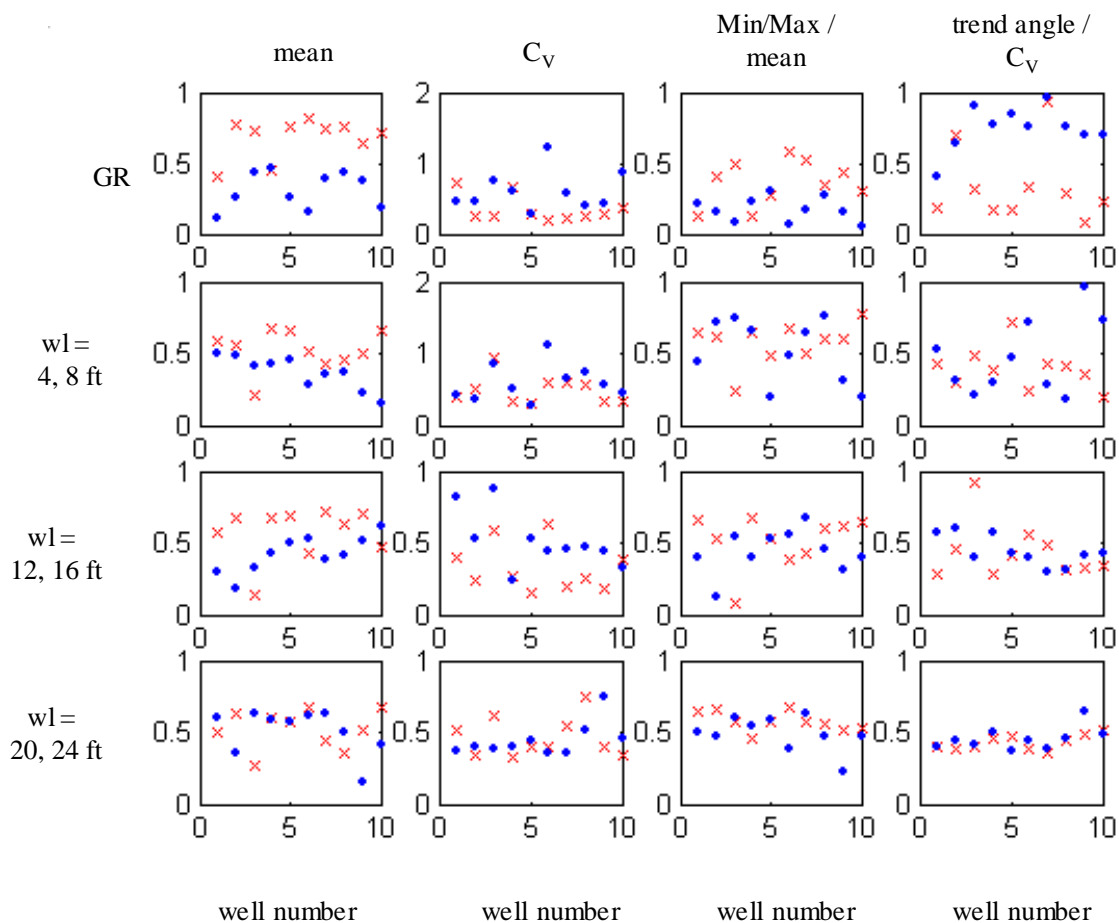


Figure 5.24 cwt features for the GR at the K1L2/K2 boundary. The first row shows the 4 features for the raw signal. The other rows show the mean and CV for the wavelets coefficients at different wavelengths.

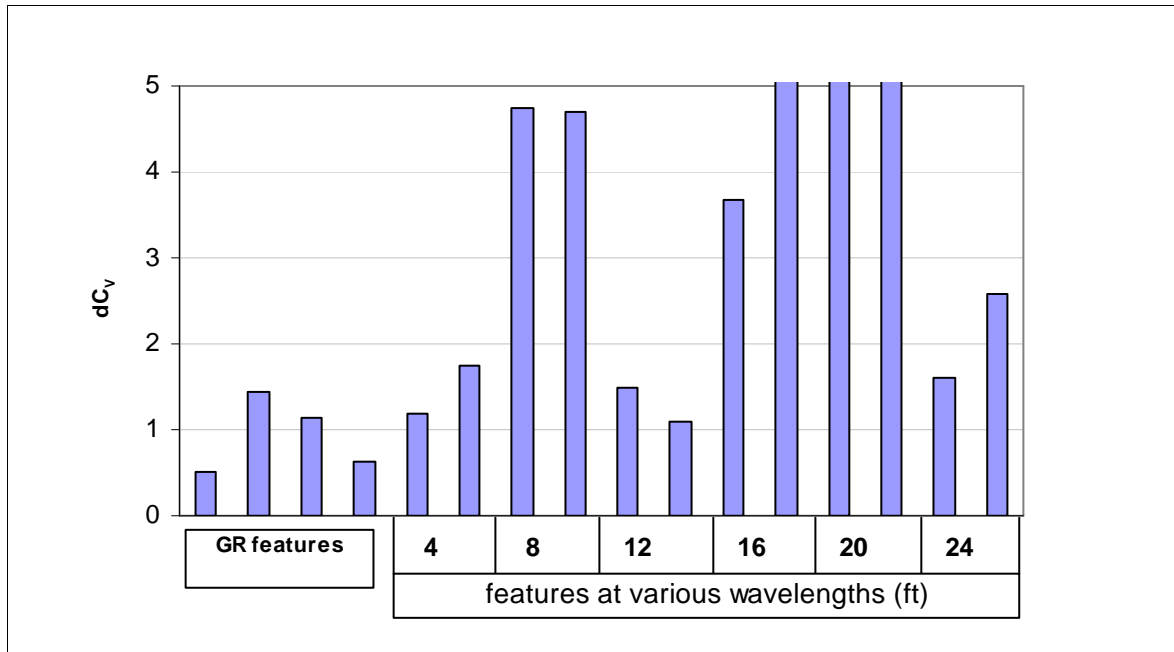


Figure 5.25 Differential-features-coefficient of variation, dC_v , for the GR, and wavelength coefficients, for various wavelengths, at the K1/K2 boundary, using all wells in Apiay. For each signal the features from left to right are: mean, CV, Min/Max, trend angle.

Table 5.9 Detection performance and mean detection error at the K1L2/K2 boundary, using cwt on the well-logs. Training wells: A-05, A-11, A-13, A-15.

Signal and cwt	Detection performance(%)	Mean detection error (ft)
GR	33	36.3
GR and R_t	17	38.6
GR and R_{x0}	50	19.1
GR and $Phinl$	33	16.6
GR and R_{hob}	17	19.1
GR, R_t , R_{x0} , $Phinl$, and R_{hob}	67	16.4

5.6 Choice of signals for boundary detection in Apiay area fields

In the preceding sections we showed that the optimal detection of a given boundary is a function of 3 variables, namely training wells, combination of signals, and pattern recognition technique (raw signal versus wavelet decomposition). This leads us to the need of finding the optimal combination of signals and type of analysis for each boundary. Assuming a high degree of confidence on the integrated study performed for Apiay, we can use these results to predict boundary positions in additional wells in Apiay or in neighboring fields. Thus the training set for this high-confidence data set is extended to include all 10 wells in our study.

Now, we proceed to evaluate our pattern recognition algorithm for all the main 7 boundaries in Apiay, by implementing the leave-one-well-out technique, as follows:

- 1) Choose a boundary.
- 2) Choose 9 out of the 10 wells as the training set, leaving out well # 1.
- 3) Select a signal or combination of signals.
- 4) Select the type of analysis (raw, dwt or cwt)
- 5) Compute the error in detection observed in the well left out.
- 6) Repeat the steps 2 to 5 until all wells are left out
- 7) Repeat steps 1 to 6 for other boundaries

The results for the T2/K1U1 are shown in Table 5.10. The dwt applied to all the 5 signals gives the least error in detection (3.6 ft). The best detection performance in terms of boundaries detected within 10 ft is 90% for the cwt on GR and Rhob.

The detection performance for all boundaries in Apiay are presented in Table 5.11. The boundary at which the best performance is achieved is the T2/K1U1, with 90% detection performance and mean detection error of 3.6 ft. The optimal performance for this boundary is with wavelet analysis. On the other hand, the optimal detection at the K1L2/K2 boundary is achieved with raw signal only (GR and R_t) with 60% detection and mean error of 31.5 ft.

Table 5.10 Detection performance for the T2/K1U1 boundary in Apiay.

		Detection error, ft					
	well	GR	GR, Rt	GR, Rxo	GR, Phinl	GR, Rhob	All 5 logs
raw	A-03	1	1	1	1	1	12
	A-04	-1	0	-1	-3	61	4
	A-05	-2	-1	55	-2	-2	-1
	A-07	0	-1	-1	0	9	-1
	A-09	75	72	7	4	5	5
	A-10	0	0	0	-3	56	0
	A-11	0	1	1	0	0	1
	A-13	0	51	-1	-1	-1	3
	A-15	64	64	64	64	64	64
	A-16	1	-3	0	39	9	4
	mean error, ft	14.4	19.4	13.1	11.7	20.8	9.5
	% wells within 10 ft	80	70	80	80	70	80
dwt	A-03	1	1	6	6	6	6
	A-04	-2	3	-1	-2	65	-1
	A-05	4	4	55	-2	4	-1
	A-07	-3	-1	-2	-1	-2	-1
	A-09	4	8	6	76	76	12
	A-10	0	12	-7	7	-7	0
	A-11	1	11	1	6	1	11
	A-13	0	1	0	1	0	1
	A-15	77	9	2	-4	-1	-2
	A-16	-7	-4	-7	9	6	1
	mean error, ft	9.9	5.4	8.7	11.4	16.8	3.6
	% wells within 10 ft	90	80	90	90	80	80
cwt	A-03	8	7	6	1	6	-2
	A-04	-1	0	-1	-2	7	-2
	A-05	-6	65	-6	-2	-6	-1
	A-07	-1	-2	29	-2	-2	-2
	A-09	-1	13	12	1	-5	13
	A-10	36	-3	-6	33	37	37
	A-11	-3	-4	-3	-3	-3	-3
	A-13	-1	3	-1	-6	-2	1
	A-15	64	5	64	64	7	64
	A-16	-3	12	-3	1	-7	5
	mean error, ft	12.4	11.4	13.1	11.5	8.2	13
	% wells within 10 ft	80	70	70	80	90	70

Table 5.11 Optimal boundary detection for Apiay.

Boundary	optimal type of analysis and statistics			
	Mean error		Detection performance within 10 ft	
	Type of analysis	Error, ft	Type of analysis	Detection Performance, %
T2/K1U1	dwt on 5 logs	3.6	cwt on GR-Rhob	90
K1U1/K1M1	dwt on GR-Rhob	20.9	dwt on GR	40
K1M1/K1M2	GR-Rhob	26.8	GR-Rhob	30
K1M2/K1M3	GR-Phinl	29.7	dwt on GR-R _t	40
K1M3/K1L1	dwt on GR-R _t	21.2	dwt on GR-R _t	50
K1L1/K1L2	dwt on GR-R _t	26.2	dwt on GR-Rhob	30
K1L2/K2	GR-R _t	31.5	GR-R _t	60

5.7 Comparison with other techniques

Our boundary detection method corresponds to a modified version of discriminant analysis (DA). The purpose of DA is the classification of samples into one of the groups established in a training set. Let S be the sample matrix of size $m \times d$, where m is the number of samples and d the number of dimensions. DA classifies each row vector in S into one of the classes of the training set.

In boundary detection we have two classes C_1 and C_2 . C_1 is the section above the boundary and C_2 the section below the boundary. A unique requirement in boundary detection is that C_1 must be located on top of C_2 . This imposes a spatial constraint. That is one of the main reasons for the need of defining features for a given window length on both sides of the boundary. We need to assign the upper section to a unique class C_1 and the lower section to a unique class C_2 .

This is in contrast to the way discriminant analysis is applied for electrofacies purposes, in which the classes do not have spatial constraints. Each electrofacies can be located anywhere in the well one or more times (e.g., Fig. 7 in Lee et al., 2002). One

class can be assigned at each depth point. Therefore, there is not need to window the well-logs, and DA is applied on the well-log readings on a foot by foot basis. The classification for boundary detection is also performed on a foot by foot basis with each depth point having features summarizing the log readings in the window. For example for our window length of 40 ft, the upper section measures 20 ft. At a depth of the window center of 10,050 ft, the upper window features summarize the logs from 10,030 to 10,050 ft. At 10,051 ft, the upper window features are computed for the range 10,031 to 10,051. Here, the input for feature extraction contains 21 log reading vectors, of which 20 vectors are common for the window center depths 10,050 and 10,051 ft.

In boundary detection, the discriminant analysis is performed on the extracted features. Each depth point is composed of features extracted for the window length. At each depth point, we calculate the probability that the upper-section features belong to class C_1 , and that the lower-section features belong to C_2 . Assuming independent events, the total probability is given by $p_i = p_{\text{upp}} \times p_{\text{low}}$.

In this chapter, we showed that our boundary-detection method is robust because of the combination of **a)** feature extraction, **b)** principal component analysis, **c)** automated feature selection, and **d)** wavelet decomposition. A more traditional approach would simply use feature extraction and principal components.

To compare our method with traditional DA, we selected the T2/K1 boundary. This is an unbiased comparison given the fact that this boundary is the easiest to recognize and the confidence on the boundary depths from the integrated study is more than 90%. We used the leave-one-well out technique, in which 9 of the wells are selected as the training set and we calculated the boundary depth in the well left out. We repeated the procedure for all 10 wells.

For the traditional approach, we used two features: the mean and the C_V . The comparisons were performed using the 5 well-logs (GR, R_t , R_{xo} , $Phinl$, and R_{hob}). The results are shown in Table 5.12. The best detection performance is accomplished with the DWT and our previously described method (the new method). The second best is our method on raw signals. If we do not perform the automated pre-selection of

features, the detection error increases considerably for both raw signal and DWT analysis. If we use our method and only two features, the mean and the CV, the detection performance lowers appreciably. These results ratify the robustness of our method due to the use of wavelet transformation and automated optimal selection of features. The fact that the detection is poor with the mean and C_V only reflects the fact that the trend is a powerful feature in boundary detection.

Table 5.12 Comparison of new method with traditional DA techniques.

Detection error for 5 well-logs at T2/K1 boundary (feet)								
well	New method		No feature opt.		2 features, opt.		2 features, no opt.	
	raw	dwt	raw	dwt	raw	dwt	raw	dwt
A-03	12	6	13	12	0	1	26	21
A-04	4	-1	8	7	54	-2	50	17
A-05	-1	-1	58	8	10	62	10	5
A-07	-1	-1	-3	-6	-2	-3	-2	92
A-09	5	12	13	19	19	18	19	22
A-10	0	0	2	-7	-5	-5	59	-7
A-11	1	11	42	41	17	44	45	44
A-13	3	1	58	7	59	1	60	7
A-15	64	-2	1	-2	10	2	11	-1
A-16	4	1	-7	-17	-19	45	-9	-27
Mean error, ft	9.5	3.6	20.5	12.6	19.5	18.3	29.1	24.3
% wells within 10 ft	80	80	50	60	50	60	40	40

5.8 Conclusions

Based on the results presented in this chapter we can draw the following observations and conclusions for boundary detection:

- 1) Boundary determination is a difficult task even when geophysical and core data are available. Making the prediction using just well-logs is a more complex process. However well-logs are the only data acquired for most wells in the oil

industry. Here, we presented a method to predict boundaries starting from the reservoir characterization performed on a few wells.

- 2) The combination of wavelet transformation and pattern recognition techniques generally improves boundary detection.
- 3) Some boundaries are easier to detect than others, using just well logs.
- 4) The combination of the GR and other signals in general, will improve the boundary detection.
- 5) Wavelet transformation in general improves the detection process. However, for two of the seven boundaries in Apiay, the optimal type of analysis was achieved with the raw signals.
- 6) The best detection was achieved for the T2/K1U1 boundary, with a mean error of 3.6 ft and 90% detection within 10 ft of true boundary. This is an excellent case of proof of principle for our boundary detection algorithm, considering that the search range for the boundary was 200 ft. This very good performance was accomplished with wavelet transformation on 5 well-logs.
- 7) The pre-selection of best features before performing pattern recognition is a key step on the boundary detection. This process is automated and is governed by the data in the wells selected for training. Typical pattern recognition applications found in the literature do not perform this pre-processing step.
- 8) We defined features that reflect the geological environment controlling stratigraphic units. In particular, the trend angle is a feature that represents the profiles of well logs. This feature proved to be very useful in the successful identification of the T2/K1 boundary. The upper section (T2) is coarsening upwards and the lower section (K1) is fining upwards in the neighborhood of their boundary. This is the only feature capable of detecting this contrast.
- 9) Using traditional features such as the mean and coefficient of variation results in 50% detection performance at the T2/K1 boundary, which is well below the 90% achieved in the new method. If automated pre-selection of features is not performed the detection diminishes even further (40%).

- 10) The determination of boundaries in Apiay area fields can be performed using the type of analysis and combination of signals presented in Table 5.11.

CHAPTER VI

SUMMARY AND CONCLUSIONS

We have developed a computer program to perform cyclostratigraphic analysis based on the wavelet transformation. We also created a computer program to detect boundaries, starting from the reservoir characterization performed on a few wells. The software incorporates wavelet decomposition (dwt and cwt), feature extraction, dimensionality reduction using PCA, and boundary detection. The following are the main conclusions of our research:

- 1) Wavelet analysis generates useful information from well-log responses. It is a useful tool for automated cyclostratigraphy. Two other methods commonly used, Fourier transform and the semivariogram, cannot be used to interpret cyclicity in the presence of superimposed cycles.
- 2) Amplitudes of spectral peaks appear to correspond with the relative importance of controlling influences on the Ormskirk Formation deposystem. Wavelet coefficients clearly reflect the different orders of cyclicity that occurred during sedimentary deposition.
- 3) Wavelets are useful in boundary detection. As it has been the case for applications in other field of science, performing the analysis in the frequency domain, allows the decomposition of a signal in high and low frequency components at multiple levels. This multiresolution analysis, using wavelet transforms, leads to better pattern recognition of the features in the neighborhood of a stratigraphic boundary.
- 4) The combination of the GR and other signals in general, will improve the boundary detection. Wavelet transformation in general improves the detection process. However, for two of the seven boundaries in the Apiay Field, the optimal type of analysis was achieved with the raw signals.
- 5) The best detection was achieved for the T2/K1U1 boundary, with a mean error of 3.6 ft and 90% detection within 10 ft of the true boundary. This is an excellent

proof of principle for our boundary detection algorithm, considering that the search range for the boundary was 200 ft. This very good performance was accomplished with wavelet transformation on 5 well-logs.

- 6) The pre-selection of best features before performing pattern recognition is a key step on the boundary detection. Typical pattern recognition applications found in the literature do not perform this pre-processing step, resulting in inferior boundary detection performance.

CHAPTER VII

FUTURE WORK

At the conclusion of our research we identify the following tasks for reservoir characterization with wavelets:

- 1) We found that the optimal window length for feature extraction from well-logs in Apiay was 40 ft. We fixed this value in our software. Future programs should incorporate the evaluation of various window lengths. This will increase the computation time.
- 2) We used the Morlet wavelet for cwt and Daubechies (D4) with 4 levels of resolution for dwt. The evaluation using other wavelets can be incorporated. The computation time will increase, and the analysis could be more time consuming.
- 3) Perform 2D wavelet decomposition on image logs to identify important events such as fractures.
- 4) Further work can be performed in coupling the cyclostratigraphic analysis using various well logs. In our research, we performed the analysis using the well-logs independently. The well-logs may be combined using principal component analysis for example to generate a unique Z-score well-log. The cyclostratigraphic analysis can be performed on this transform, resulting in a Z-score scalogram.
- 5) Wavelets in well-test applications have focused on de-noising and data conditioning. Further work could be performed to implement wavelets as a useful tool in well-test interpretation.

NOMENCLATURE

NE = Northeast

NW = Northwest

SE = Southeast

SW = Southwest

GR = Gamma-ray log

R_t = Deep resistivity log

R_{xo} = Shallow resistivity log

Phinl = Neutron-porosity log

Rhob = Bulk density log

Δt = Transit time log

REFERENCES

- Anstey, N. A., and O'Doherty R.F., 2002, Cycles, Layers, and Reflections: Part 1: *The Leading Edge*, January, p. 44-51.
- Athichanagorn, S., Horne, R.N., and Kikani, J., 1999, Processing and Interpretation of Long-term Data from Permanent Downhole Pressure Gauges, paper SPE 56419, *in* SPE Annual Technical Conference and Exhibition, Houston, Oct. 3-6.
- Bassiouni, Z., Theory, 1994, Measurement and Interpretation of Well Logs, SPE Textbook Series Vol. 4, Richardson, TX, 372 p.
- Bernasconi, G, Rampa, V., Abramo F., and Bertelli, L., 1999, Compression of Downhole Data, paper SPE/IADC *in* SPE/IADC Drilling Conference, Amsterdam, 9-11 March.
- Bishop, C.M., 1995, Neural Networks for Pattern Recognition, Oxford Univ. Press, New York, 256 p.
- Box, G.E.P. and Jenkins, G.M., 1976, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco.
- Chakraborty, A. and Okaya, D., 1995, Frequency-Time Decomposition of Seismic Data Using Wavelet-Based Methods: *Geophysics*, v. 60, p. 1906-1916.
- Cowan, G., 1993, Identification and Significance of Aeolian Deposits within the Dominantly Fluvial Sherwood Sandstone Group of the East Irish Sea Basin UK", *in* North, C., and Prosser, J., eds., Characterization of Aeolian and Fluvial Reservoirs: Geol. Soc. Spec. Pub 73, p. 231-245.
- Doveton, J.H., 1994, Geologic Log Analysis Using Computer Methods, AAPG Computer Applications in Geology, No. 2, Tulsa, OK.
- Everitt B.S., and Dunn G., 1992, Applied Multivariate Data Analysis, Oxford University Press, New York, 304 p.
- Galloway, W.E. and Hobday, D.K., 1996, Terrigenous Clastic Depositional Systems – 2nd ed. Springer.
- Gelhar, L.W., 1993, Stochastic Subsurface Hydrology, Prentice-Hall, Englewood Cliffs, N.J.

- Gonzalez, F, Camacho, R, and Escalante, B., 1999, Truncation De-noising in Transient Pressure Tests, paper SPE 56422, *in* SPE Annual Technical Conference and Exhibition, Houston, Oct. 3-6.
- Goumas, S.K., Zervakis, M.E., and Stavrakakis, G.S., 2002, Classification of Washing Machines Vibration Signals Using Discrete Wavelet Analysis for Feature Extraction: *IEEE Trans. Instrumentation and Measurement*, v. 51, no. 3, June 2002, p. 497-508.
- Goswami, J. and Chan, A.K., 1999: Fundamentals of Wavelets, John Wiley, New York.
- Gutierrez, M., Lozano, E., and Sroczynski, 1996, Impact of 3-D Seismic Interpretation on Reservoir Management in the Apiay-Ariari Oil Fields, Llanos Basin, Colombia, *in* P. Weimer and T.L. David, eds., AAPG Studies in Geology No. 42 and SEG Geophysical Development Series No. 5, AAPG/SEG, Tulsa, OK, p. 57-68.
- Herries, R. D., and Cowan, G., 1997, Challenging the Sheetflood Myth: The Role of Water-Table-Controlled Sabkha deposits, *in* Meadows, N. S., et al., eds, Petroleum Geology of the Irish Sea and Adjacent Areas, Geol. Soc. Spec. Pub. 124, p. 253-276.
- Hosking, J.R.M. and Wallis J.R, 1997, Regional Frequency Analysis : An Approach Based on L-Moments, Cambridge U. Press, New York.
- Jansen, F.E., and Kelkar, M.G., 1997, Application of Wavelets to Production Data in Describing Inter-Well Relationships, paper SPE 38876, *in* Annual Technical Conference and Exhibition: Society of Petroleum Engineers, p. 323-330.
- Jennings, J.W., Ruppel, S.C., and Ward, W.B., 2000, Geostatistical Analysis of Permeability Data and Modeling of Fluid-Flow Effects in Carbonate Outcrops: *SPEREE*, v.3, p. 292-303.
- Jensen, A., and la Cour-Harbo A., 2001, Ripples in Mathematics, Springer, New York, 246 p.
- Jensen, J.L., Lake, L.W., Corbett P.W., and Goggin D.J, 2000, Statistics for Petroleum Engineers and Geoscientists, Elsevier, Amsterdam, 338 p.
- Jolliffe, I.T., 2002, Principal Component Analysis, Springer, New York, 487 p.

- Lau K.-M. and Weng, H., 1995, Climate Signal Detection Using Wavelet Transform: How to Make a Time Series Sing: *Bull. of the Am. Meteorological Society*, v. 76, p. 2391-2402.
- Lee, S.H., Khargoria, A., and Datta-Gupta A., 2002, Electrofacies Characterization and Permeability Predictions in Complex Reservoirs: *SPEE*, June, p. 237-248.
- Malagon, C., 1997, Facies Analysis and Stratigraphic Framework of the Lower Oligocene T1 Sandstone, Carbonera Formation, Apiay Field, Colombia, MS thesis, University of Colorado.
- Meadows, N. S., and Beach, A., 1993, Structural and Climate Controls on Facies Distribution in a Mixed Fluvial and Aeolian Reservoir, *in* North, C., and Prosser, J., eds., Characterization of Aeolian and Fluvial Reservoirs, Geol. Soc. Spec. Pub. 73 p. 247-264.
- Mitchum, R.M. and J.C. Van Wagoner, 1991, High-Frequency Sequences and Their Stacking Patterns: Sequence-Stratigraphy Evidence of High-Frequency Eustatic Cycles: *Sedimentary Geology*, v. 70, p. 131-160.
- Moller, N.K. and van de Wel, D., 1998, High Resolution Sequence Stratigraphy as a Basis for 3D Reservoir Modeling, *in* Gradstein, F.M., ed., Sequence Stratigraphy: Concepts and Applications, Elsevier, Amsterdam, p. 315-336.
- Nieto, J.E., 1997, Facies Architecture and Production Optimization in the Upper Cretaceous-Upper Eocene, Operational Unit K1, Apiay Field, Llanos Basin, Colombia, MS thesis, The University of Texas at Austin.
- Nieto, J.E., 2003, Personal communication, graduate student at Texas A&M University, College Station, TX.
- Nystuen, J.P., 1998, History and Development of Sequence Stratigraphy, *in* Gradstein, F.M., ed., Sequence Stratigraphy: Concepts and Applications, Elsevier, Amsterdam, p. 31-116.
- Panda, M.N., Mosher, C.C., and Chopra, A.K., 2000: "Application of Wavelet Transforms to Reservoir-Data Analysis and Scaling, *SPEJ*, p. 92-101.
- Paupolis, A., 1962, The Fourier Integral and Its Applications, McGraw-Hill, New York.

- Plint, A.G., Eyles, N., Eyles, C., and Walker, R.G., 1993, Control of Sea Level Change, *in* Walker, R.G., and James, N.P., eds., *Facies Models – Response to Sea Level Change*, Geological Society of Canada, Toronto, p. 15-25.
- Proakis, J.G., and Manolakis, D.G., 1996, *Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 968 p.
- Prokoph, A. and Agterberg, F.P., 1999, Detection of Sedimentary Cyclicity and Stratigraphic Completeness by Wavelet Analysis: An Application to Late Albian Cyclostratigraphy of the Western Canada Sedimentary Basin: *Journal of Sedimentary Research*, v. 69, July, p. 862-875.
- Prokoph, A. and Agterberg, F.P., 2000, Wavelet Analysis of Well-Logging Data from Oil Source Rock, Egret Member, Offshore Eastern Canada: *AAPG Bull*, v. 84, no. 10, p. 1617-1632.
- Prokoph, A. and Barthelmes, F., 1996, Detection of Nonstationarities in Geological Time Series: Wavelet Transform of Chaotic and Cyclic Sequences: *Computers & Geosciences*, v. 22, p. 1097-1108.
- Saggaf M.M. and Nebrija E.L., 2000, Estimation of Lithologies and Depositional Facies from Wire-Line Logs: *AAPG Bull*, v. 84, no. 10, p. 1633-1646.
- Schwarzacher, W., 1993, Cyclostratigraphy and the Milankovitch Theory, *Developments in Sedimentology* 52, Elsevier, Amsterdam, 225 p.
- Schwarzacher, W., 1998, Stratigraphic Resolution, Cycles and Sequences, *in* Gradstein, F.M., Sandvik, K.O., and Milton, N.J., eds., *Sequence Stratigraphy – Concepts and Applications*, Elsevier, Amsterdam, p. 1-8.
- Serra, O. and Abbot, H.T., 1982, The Contribution of Logging Data to Sedimentology and Stratigraphy: *SPEJ*, February, p. 117-135.
- Soliman, M.Y., Ansah, J., Stephenson, S., and Manda, B., 2001, Application of Wavelet Transform to Analysis of Pressure Transient Data, paper SPE 71571 *in* SPE Annual Technical Conference and Exhibition, New Orleans, 30 Sept-3 Oct.
- Thomas, S., Jensen, J., and Corbett, P., 1998, Application of Small Scale Measurements and Predictions for Accurate Characteristics of Anisotropic Permeability in the

Sherwood Sandstone, Well 110/8a-5 and 110/2a-N1, Heriot-Watt University, Edinburgh.

Thompson, J., and Meadows, N. S., 1997, Clastic Sabkhas and Diachroneity at the Top of the Sherwood Sandstone Group, *in* Meadows, N. S., et al., eds., Petroleum Geology of the Irish Sea and Adjacent Areas, Geol. Soc. Spec. Pub. 124, p. 237-251.

APPENDIX A

GEOLOGIC DESCRIPTION OF THE ORMSKIRK FORMATION

Structure

The East Irish Sea basin is part of a linked array of basins formed as a result of a Permo-Triassic extensional phase related to early Atlantic opening. The sedimentary fill comprises Permian, Triassic and Early Jurassic rocks unconformably overlying Carboniferous sequences. Hydrocarbon prospectivity in the area has centered on the Early Triassic Sherwood Sandstone Group, which forms the reservoir for the Morecambe Field. The northern part of the East Irish Sea basin, where the Morecambe field is located, comprises a series of half-graben bounded by an orthogonal set of NE-SW- and NNW-SSE-trending extensional faults forming an elongate rhomb-shaped (Fig. A.1). This interpretation was proposed by Herries and Cowan (1997) who used dipmeter analysis from the Morecambe fields to infer westerly fluvial paleoflow. This contradicted the Meadows and Beach model (1993), which predicted major northward-flowing fluvial systems along the hanging walls of the basin-bounding faults.

Stratigraphy

The Sherwood Sandstone Group is mainly Scythian (Early Triassic) in age. In the East Irish Sea Basin, the Sherwood Sandstone is overlain by the mudstone- and halite-dominated Mercia Mudstone Group and underlain by the St Bees Shale. The Sherwood Sandstone Group is subdivided into upper and lower formations. The lower is the St Bees Sandstone Formation. The upper formation is the Ormskirk Sandstone which comprises mainly sandstones of mixed fluvial, eolian and sheetflood origins (Herries and Cowan, 1997).

Sedimentology

The upper interval of the St Bees Sandstone and the whole Ormskirk Sandstone comprise a mixed association of facies types, including fluvial channels developed on a

variety of scales, eolian dunes and sandsheets, sandflats subject to episodic unconfined sheetflood events and ephemeral playa lakes. All of these facies are consistent with deposition in a semiarid continental environment.

Fluvial channel systems

The sediments associated with fluvial channel in the Sherwood Sandstone are subdivided into those relating to major, perennial rivers and those associated with ephemeral channels. The sequences interpreted as major fluvial channels systems comprise stacked, erosively-based sets of moderately sorted, medium- and fine-grained sandstone. The sandstones commonly contain clay to silt grade intraclasts. Individual tabular and trough cross-stratified sets are typically stacked to form compound channel-fill sandstone bodies that commonly fine upwards and exhibit an upward decrease in preserved set thickness, occasionally passing up into abandonment fines. This upward fining represents deposition under waning flow conditions, while a variety of ripple forms, typically forming the top to these sequences, represent bar-top modification during low- or falling-stage discharge (Meadows and Beach, 1993).

These major stacked or multi-story sandstone bodies are the depositional products of low- to moderate-sinuosity braided channels containing bars. Some of these sequences also contain intervals of flat-laminated or low-angle, cross-stratified sandstones. The presence of these deposits agrees with the suggestion of perennial rivers subject to significant variation in flow regime (Meadows and Beach, 1993).

There are records of deposition within minor channels with similar sedimentological characteristics to the perennial rivers but with significant differences indicating ephemerality. These features are the preserved thickness of individual and compound bar forms, the absence of intrasets and graded foresets, the common abundance of reworked eolian-type grains, and the interdigitation of these channel sandstones with sheetflood, playa and eolian facies. The interdigitation of facies and scarcity of thick

multi-story sandstone bodies in the sequences comprising this facies type indicate that these channels were short-lived morphological features. The occasional presence of partially flood-reworked eolian dune sandstones and wind-reworked tops of fluvial bars support this hypothesis. These fluvial channels were subject to highly variable discharge regimes, dried out for substantial periods and, when active, may have formed tributaries to the major perennial channel systems (Meadows and Beach, 1993).

Eolian systems

Eolian sandstones are a significant component of the Ormskirk Sandstone Formation. These sediments are either true eolian bedforms or reworked material within other facies, and they are very important due to their consistently high reservoir quality. Eolian deposits occur primarily as either dunes or sandsheets. Moderate-to-high-angle, planar to asymptotically-dipping, cross-stratified sandstones exhibiting fine grain-size differentiated lamination, are interpreted as eolian dunes. Within the Ormskirk Sandstone, dunes are commonly interbedded with other facies, principally eolian sandsheet and sheetflood deposits, and rarely exceed one meter in preserved thickness.

Eolian sandsheets are very finely laminated, flat-to low-angle, cross-stratified sandstones and are interbedded with dune and sheetflood deposits. There is also interdigitation of sandsheet facies with other facies types on a very fine scale with millimeter to centimeter alterations of eolian laminae and silty sheetflood laminae. These very fine facies changes represent subtle variations in the paleo-water table that either allowed or inhibited surficial eolian reworking. The development of more substantial intervals dominated by eolian processes, generally comprising interbedded eolian dune and sandsheet sediments, represents significant episodes of lowered water table conditions, whereas intervals in which eolian sandsheet intercalations form a minor component interbedded with fluvial sheetflood and playa deposits represent episodes of raised water table conditions (Meadows and Beach, 1993).

Sheetflood systems

Sediments attributed to deposition by unconfined fluvial processes form a major element of the Ormskirk Sandstone formation. These sediments rarely form thick, continuous, single facies. Instead they are most commonly interdigitated with eolian, playa and ephemeral channel deposits. These deposits vary from highly porous, flood-reworked, eolian sandstones to low-porosity, finely and irregularly laminated silty and argillaceous fine-grained sandstones. Deposition was from recurrent, rapid, bedload-sedimentation events that were followed by slack-water fall-out of suspension fines. There are also thin, dessicated claystone laminae within these sequences, evidencing the occasional drying out of the sediment surface (Meadows and Beach, 1993).

Playa systems

Claystone-and siltstone-dominated sequences representing deposition within, or associated with, standing bodies of water occur at various horizons within the Ormskirk Sandstone but are most common in the upper parts of the formation. These sediments are volumetrically minor but stratigraphically significant since they constitute the only primary non-reservoir units within the sequence. Intervals interpreted as representing deposition in ephemeral playa lakes are dominated by claystones and siltstones. These sediments were deposited in shallow water. The sediments commonly contain dessication fractures up to 10 cm deep. The majority of playas are assumed to have been non-evaporitic (Meadows and Beach, 1993).

Sequences interpreted as playa margin environments are transitional between the deposits of playa lakes and wet, sheetflood-dominated, sandflats. They exhibit the characteristics of both facies with an interbedding of playa sediments, representing lake expansion, and sheetflood with occasional eolian sandsheet representing lake contraction and drying in the marginal areas (Meadows and Beach, 1993).

Depositional model

Within the eolian-and sabkha-dominated sections of the Ormskirk Sandstone, a repeated arrangement of facies 2-10 m thick is consistently present (Herries and Cowan, 1997). Strongly water-table influenced deposits grade up into eolian sabkha sandstones with a matching increase in both frequency and thickness of interbedded dry eolian deposits. The abundant, thin eolian interbeds suggest that winds transported the bulk of the sediment from dried river beds around the basin margins. These vertically ordered facies patterns form units that can be correlated for at least 20 km in some of the Irish Sea fields. The widespread nature of these drying-upward units is interpreted as a response to basin-wide changes in the balance between sand saturation and the water-table level (Herries and Cowan, 1997).

Herries and Cowan (1997) favor a sand-supply dominated model, which requires distinct seasonality to account for the water-table fluctuations and longer term, cyclic changes in water-table level to generate the drying upward patterns. Climatic modeling suggests that interior lowlands underwent extreme seasonal temperature variations with seasonal precipitation followed by extreme heat and humidity. This would produce the changes in water-table. A mechanism for generating the drying upward units may be the 23000 year precession cycle. Herries and Cowan (1997) conclude that each cycle should be represented by approximately 8 m of preserved sediment, which is within the 2-10 m range of thickness observed for the drying upward depositional units. Support for climatic control comes from evidence for changing wind directions preserved in drying upward units (Herries and Cowan, 1997).

Facies correlations

Thompson and Meadows (1997) evaluated wells representing a northwest to southeast traverse from basin center to margin in the Morecambe field (Fig. A.2). Eolian facies are abundant in wells 110/2-6 and 110/8a-5 and are replaced upwards by progressively wetter sandflat deposits dominated by a range of sheetflood, playa and ephemeral fluvial

channel deposits. A similar wetting upward pattern is evident in 110/13-5 with greater abundance of eolian sandstones. In 110/14-3 the upper sandflat interval is thicker and is dominated by eolian dune and sandsheet deposits. Facies become more eolian towards the basin margin and form part of thicker sandflat sequence, relative to the top fluvial channel datum.

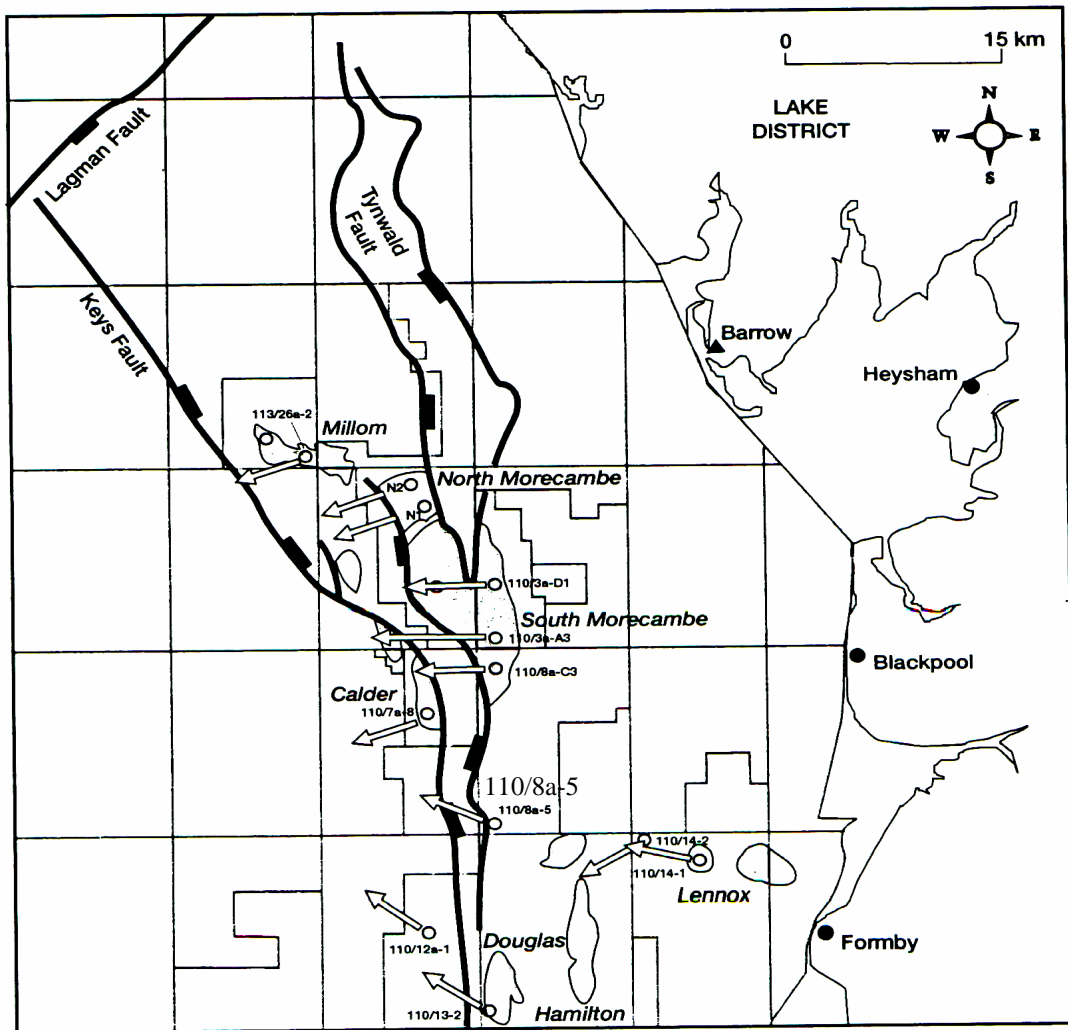


Figure A.1 Paleocurrent vectors for cross-stratified fluvial sandstones (from Herries and Cowan, 1997).

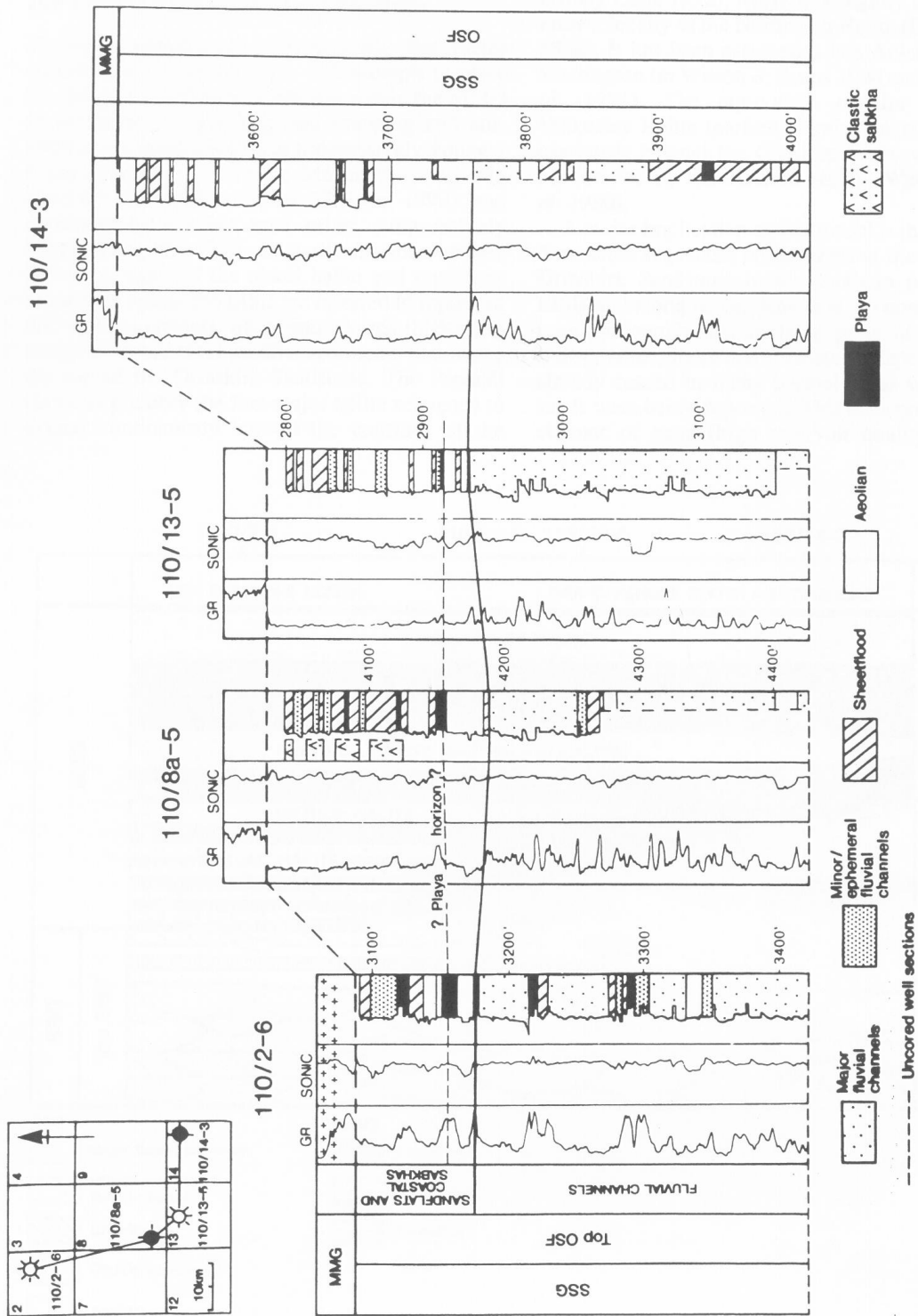


Figure A.2 Well correlation for the East Irish Sea Basin (from Thompson & Meadows, 1997).

APPENDIX B

STRATIGRAPHIC CROSS SECTIONS FOR APIAY FIELD

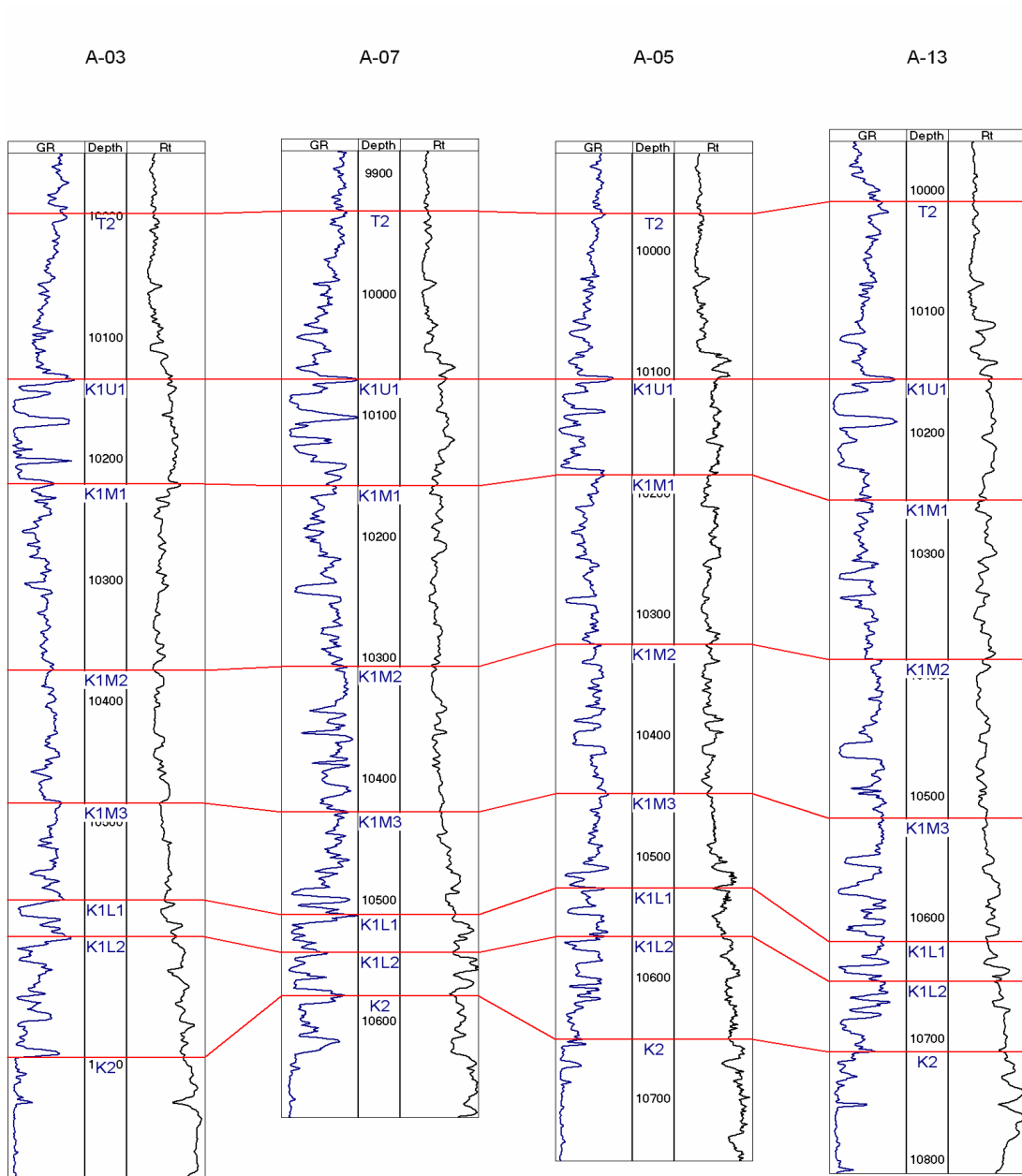


Figure B.1 Stratigraphic cross section (N-S) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are GR and R_t .

A-15

A-11

A-10

A-16

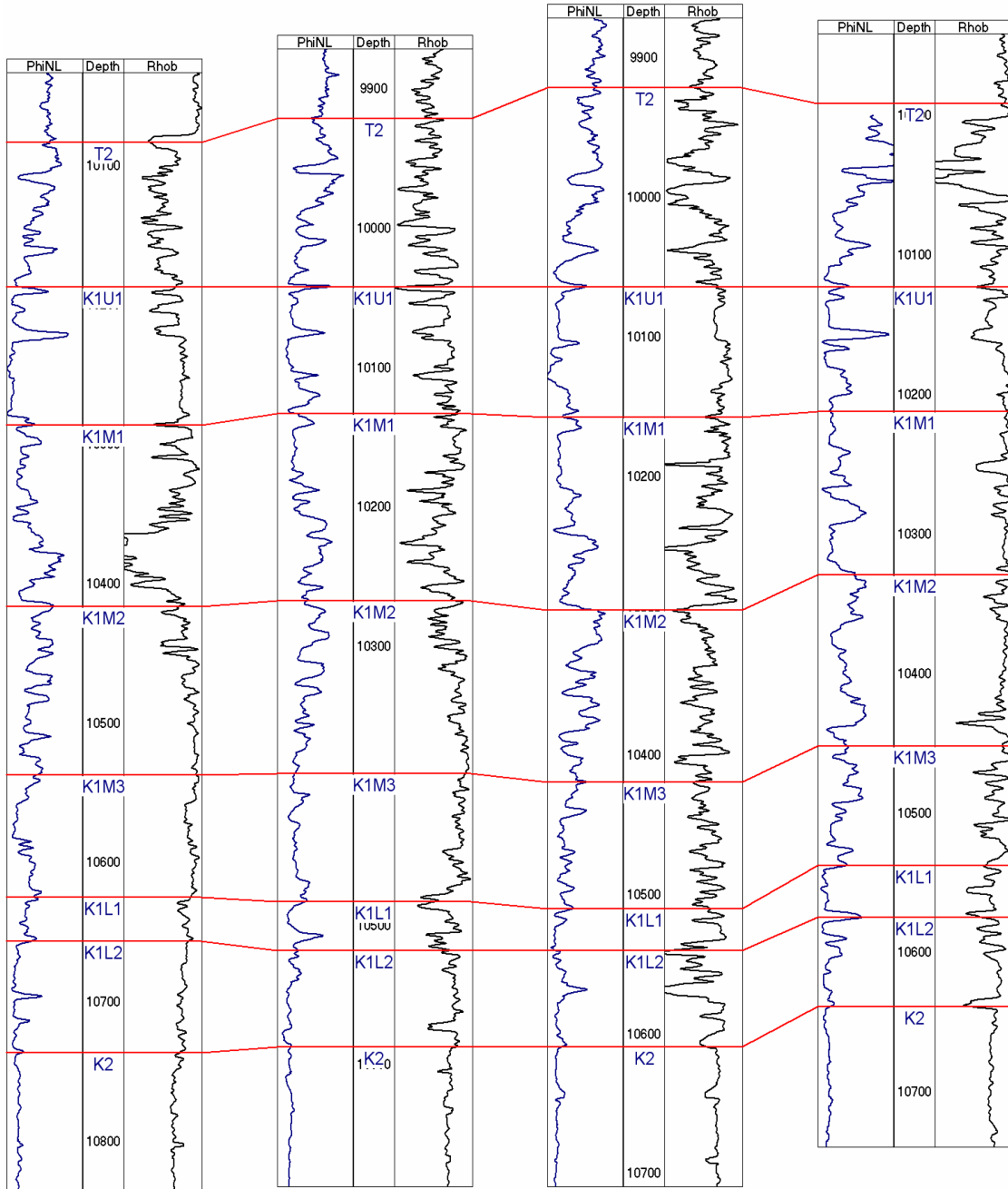


Figure B.2 Stratigraphic cross section A-A' (NE-SW) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are Phi_{nl} and ρ_b.

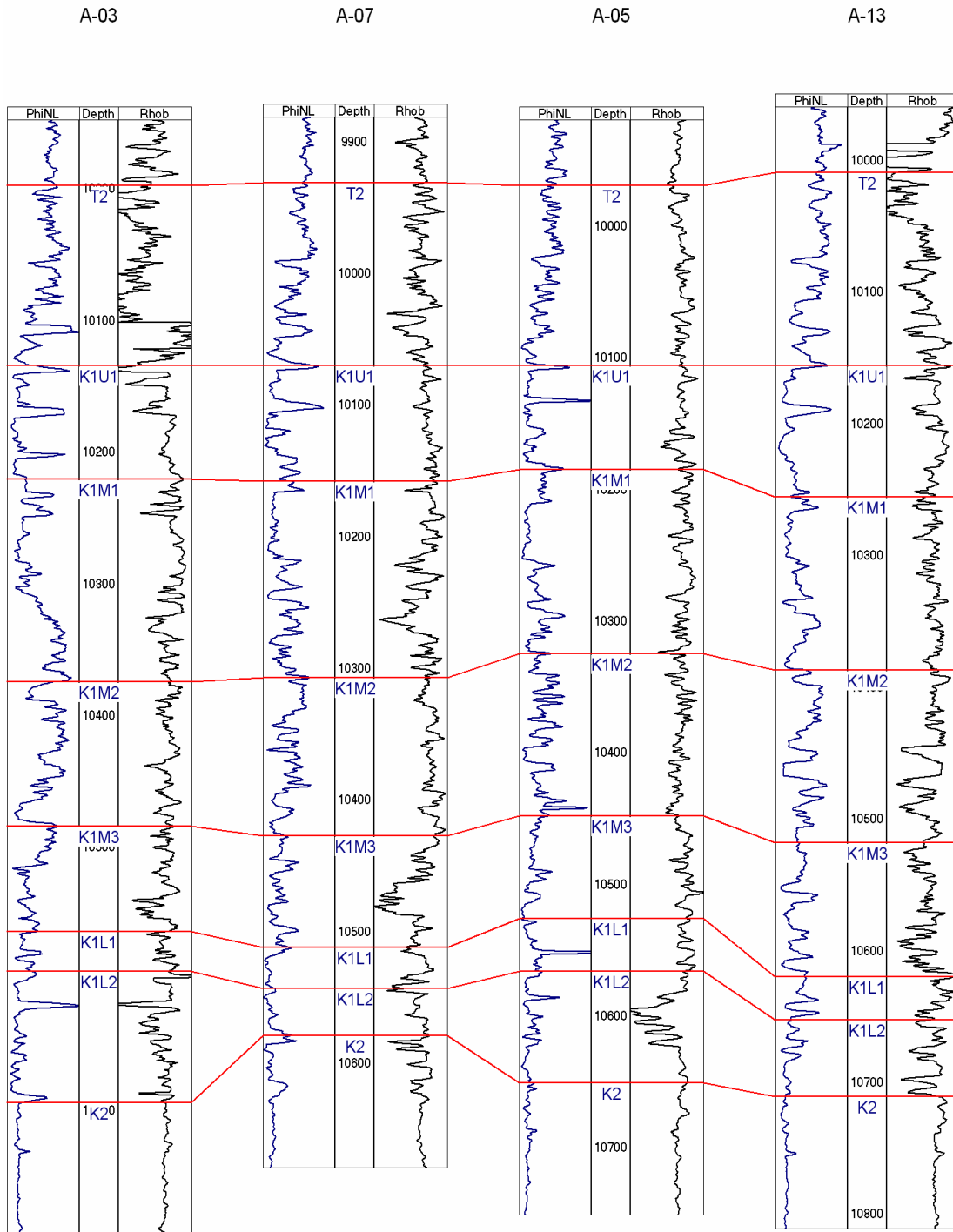


Figure B.4 Stratigraphic cross section (N-S) across the Apiay field displaying the units from reservoir integrated studies. Well logs shown are PhiNL and ρ_b .

VITA

Nestor Rivera Vega was born in Chinavita, Colombia on May 15, 1962, the son of Rafael Rivera and Elvia Vega. After completing his work at Carlos Arturo Torres High School, Santa Rosa de Viterbo, Colombia, in 1978, he entered Universidad Industrial de Santander in Bucaramanga, Colombia, where he graduated with a B.S. degree in petroleum engineering in 1985. In 1986 he started working for the Colombian state oil company, Ecopetrol, in the drilling department until 1993. He worked as a mud engineer, drilling and workover operations engineer, and in well planning.

He obtained his M.S. from The University of Texas at Austin in petroleum engineering in 1995. He worked as a reservoir simulation engineer for Ecopetrol from 1996 to 1999. He started the PhD program in petroleum engineering at A&M in 2000. He worked for Schlumberger USA in summer 2001 and summer 2002.

He joined ChevronTexaco in August 2003 as a reservoir simulation engineer.

His permanent address is:

Carrera 28A # 49A-39 Apt. 403

Bogota, Colombia

e-mail: nrivera@neo.tamu.edu.