



<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study,
without prior permission or charge

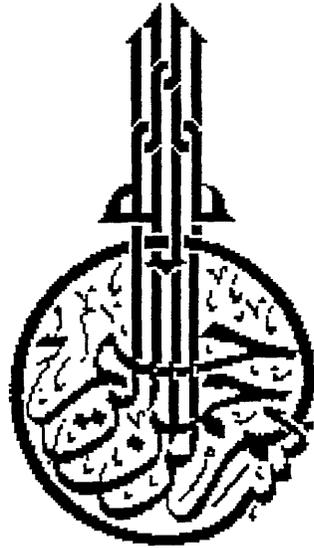
This work cannot be reproduced or quoted extensively from without first
obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any
format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author,
title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk



In the name of Allah, the Beneficent, the Merciful.

ProQuest Number: 11011418

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 11011418

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

DECLARATION

The material in this thesis is the results of independent research by the author undertaken between March 1989 and October 1991 at the Department of Geology and Applied Geology, University of Glasgow. Any published or unpublished papers have been given full acknowledgment in the text.

Mohamed Saleh

Dr Colin Farrow

Department of Geology and Applied Geology,
University of Glasgow, U.K

**Computer Application of Principal Component Analysis to
Boundary Identification and well-to-well Correlation**

by

Mohamed Rahuma Saleh

**Thesis submitted for the degree of Master of Science at the Department of
Geology and Applied Geology, University of Glasgow, U.K.**

October 1991

elias

DEDICATION

To my parents

My dear

My dear

My dear

My dear

ACKNOWLEDGMENTS

First, I wish to express my thanks to the management of Sirte Oil Company and in particular the management of the Exploration Department for giving me this opportunity, for the scholarship, and the release of the data without which this research would not have been possible.

I especially thank Dr Colin Farrow, my supervisor, for his generous support, guidance, constructive advice, and the time during the work on this project.

My thanks also go to Schlumberger Company, London, for the release of the software for reading the well-log magnetic tapes.

I wish also to thank Mr Ahmed El-Gabaili, Sirte Oil Company Manager of London office, for his constant support to complete this study. Many thanks to Mr Fawzi Ahmed, a research student, for his very helpful and constructive remarks.

Most of all, I am very thankful and obliged to my wife for her patience, and continuous encouragement during the last period of this study.

Contents

	Page
Contents	I
List of Figures	V
List of Tables	XII
Abstract	XIV
Chapter 1 Introduction	
1.1 Purpose and scope	2
1.2 Previous work.....	2
Chapter 2 Geological Setting	
2.1 Introduction.....	8
2.2 Petroleum exploration history in the Sirte basin	12
2.3 Area of study.....	12
2.4 Structure.....	13
2.5 Stratigraphy.....	19
2.5.1 Paleozoic.....	19
2.5.2 Mesozoic.....	22
2.5.3 Tertiary.....	24
Chapter 3 Theoretical background	
3.1 Introduction.....	28
3.1.1 Basic Statistical Calculations.....	28
3.2 Principal Component Analysis.....	37
3.3 Smoothing.....	46

	page
3.4 Boundary identification techniques.....	47
3.5 Correlation of well-log sequences.....	54
3.6 Discrete Fourier transform.....	58
3.7 Derivative filtering of data.....	64
3.8 Fast Fourier transform.....	67
3.9 Prediction of stretching & displacement with power spectra	68
3.10 Logarithmic scaling of frequencies.....	70
3.11 Interpolation (stretching) by inverse FFT.....	74
3.12 Determination of displacement.....	74
3.13 Summary.....	76

Chapter 4 Application to Attahaddy Field

4.1 Introduction.....	78
4.2 Analysis of model data.....	81
4.2.1 Using the original variables.....	81
4.2.2 Using the principal components.....	82
4.3 Interpretation of real data.....	88
4.3.1 Introduction.....	88
4.3.2 Correlation between FF7 and FF13.....	90
(a) Correlation of the Sheghega Formation.....	97
(b) Correlation of the Domran Formation.....	100
(c) Correlation of the Ruaga Formation.....	103
(d) Correlation of the Heira Formation.....	105
4.3.3 Correlation between FF13 and FF11.....	109
(a) Correlation of the Etel Formation.....	111

	page
(b) Correlation of the Sheghega Formation.....	115
(c) Correlation of the Domran Formation.....	120
(d) Correlation of the Ruaga Formation.....	120
(e) Correlation of the Heira Formation.....	126
4.3.4 Correlation between FF11 and FF12.....	126
(a) Correlation of the Etel Formation.....	134
(b) Correlation of the Sheghega Formation.....	134
(c) Correlation of the Domran Formation.....	137
(d) Correlation of the Ruaga Formation.....	142
(e) Correlation of the Heira Formation.....	142
4.3.5 Correlation between FF12 and FF10.....	146
(a) Correlation of the Etel Formation.....	148
(b) Correlation of the Sheghega Formation.....	148
(c) Correlation of the Domran Formation.....	154
(d) Correlation of the Ruaga Formation.....	154
(e) Correlation of the Heira Formation.....	158
4.4 Correlation of lithologies within a Formation.....	158
4.5 Characterization of different rock type.....	162
4.6 Discussion of results of real data.....	168
4.7 Summary.....	173
Chapter 5 Software	
5.1 Introduction.....	174
5.2 Program structure.....	177

			page
Chapter 6	Conclusion		190
References			195
Appendix	A		202
Appendix	B		213
Appendix	C		221
Appendix	D		285
Appendix	E		319
Appendix	F		328

List of Figures

	page
 Chapter 2	
Fig. 2.1 Location map of the area studied within the Sirte basin, Libya.	9
Fig. 2.2 Map showing Structural elements of the Sirte basin.	10
Fig. 2.3 Concession map and major tectonic elements in Sirte basin	11
Fig. 2.4 Map showing the major structural elements of the central Sahara in Early Palaeozoic times.	15
Fig. 2.5 Map showing the major structural elements of the central Sahara in Late Palaeozoic and Mesozoic times.	16
Fig. 2.6 Schematic map of major structural elements; showing the present day basin and uplifts of the central Sahara.	18
Fig. 2.7 Formation correlation of the Upper-Cretaceous in the Sirte basin.	20
Fig. 2.8 Stratigraphic section in the Attahaddy field.	21
 Chapter 3	
Fig. 3.1 a A histogram of seven variables in FF13-6.	40
Fig. 3.1 b Pie chart showing the percentage of each eigenvalue in well FF13-6.	40
Fig. 3.2 Plot of the principal components in well FF13-6.	41
Fig. 3.3a A histogram of seven variables in FF13-6.	43
Fig. 3.3b Pie chart showing the percentage of each eigenvalue in well FF13-6.	43
Fig. 3.4 Plot of principal components in well FF13-6	44

	page
Fig. 3.5 Illustration showing the transformation of data by principal component analysis.	45
Fig. 3.6 A diagram showing the affect of window size in filtering. (a) principal component I in well FF13-6. (b) Component I filtered by a moving average filter with window size=50 feet.	48
Fig. 3.7 A diagram showing the affect of window size in filtering. (a) principal component I in well FF13-6. (b) Component I filtered by a moving average filter with window size=150 feet.	48
Fig. 3.8 A diagram showing (a) boundary identification technique (D^2). (a) Small window (100 feet) is used. (b) principal component I in well FF13-6.	51
Fig. 3.9 A diagram showing (a) boundary identification technique (D^2). (a) big window size (250 feet) is used. (b) principal component I in well FF13-6.	52
Fig. 3.10 A diagram showing (a) boundary identification technique (D^2). (a) window size (100 feet) is used. (b) principal component in well FF13-6.	53
Fig. 3.11 Sketch showing the correlation process.	57
Fig. 3.12 Filtering first principal component using high-pass filter. (a) first principal component of two formation in FF13-6. (b) their derivative data.	66
Fig. 3.13 Model data used for cross-correlation of a short with a long series.	69
Fig. 3.14 Interpolation (stretching) technique using FFT.	75

	page
Chapter 4	
Fig. 4.1 Location of the boreholes in the Attahaddy field.	79
Fig. 4.2 Output plot using model data.	83
Fig. 4.3 Output plot using model data of the first principal components of thin (50 feet) and thick formation (300 feet).	84
Fig. 4.4 Output plot using model data of the first principal component of thin formation (120 feet) and thick formation (450 feet).	86
Fig. 4.5 Output plot using model data of the first principal component of thin formation (50 feet) and thick formation (220 feet).	87
Fig. 4.6 Structural cross-section in the Attahaddy field.	89
Fig. 4.7 Location map of the Attahaddy field. The solid lines shows the boreholes used in the analysis.	91
Fig. 4.8 Diagram illustrates the first principal component and its boundaries in FF7. (a) the first principal component. (b) are the boundaries of different formations.	93
Fig. 4.9 Diagram illustrates the first principal component and its boundaries in well FF13. (a) the first principal component. (b) are the boundaries of different formations.	94
Fig. 4.10 Plot of the filtered principal component of two boreholes (FF7-FF13) in the Attahaddy field.	96
Fig. 4.11 plot of correlation function using real data in Attahaddy field.	98
Fig. 4.12 Computer correlation of the Sheghega formation (FF13 and FF7).	99
Fig. 4.13 plot of cross-correlation function of the Sheghega formation using non-filtered data of FF13 and FF7.	101

	page
Fig. 4.14 Cross-correlation of the Domran formation (FF13 and FF7).	102
Fig. 4.15 Cross-correlation of the Ruaga formation (FF13 and FF7).	104
Fig. 4.16 plot of cross-correlation function of the Heira formation using the derivative data of FF13 and FF7.	106
Fig. 4.17 Cross-correlation of Heira formation using the derivative data.	107
Fig. 4.18 Cross-correlation of the Heira formation using the original data (filtered principal components) for stretching .	108
Fig. 4.19 Plot of the first principal components (sampled at 5 feet) of well FF11 and its boundaries.	110
Fig. 4.20 Plot showing the smoothed principal component of the correlated sequences (FF13 and FF11).	112
Fig. 4.21 plot of cross-correlation function of the Etel formation using the derivative data of FF13 and FF11.	113
Fig. 4.22 Cross-correlation of the Etel formation using the derivative data.	114
Fig. 4.23 Cross-correlation of the Etel formation using the non-filtered principal components.	116
Fig. 4.24 plot of cross-correlation using the original data (Gamma Ray) of the Etel formation in FF13 and FF11.	117
Fig. 4.25 Cross-correlation of the Etel formation using the original variables (GR).	118
Fig. 4.26 plot of cross-correlation of the Sheghega formation in FF13 and FF11 using the derivative data.	119
Fig. 4.27 Cross-correlation of the Sheghega formation.	121
Fig. 4.28 plot of cross-correlation of the Domran formation in FF13 and FF11 using the derivative data.	122

	page
Fig. 4.29 Cross-correlation of the Domran formation.	123
Fig. 4.30 plot of cross-correlation of the Ruaga formation in FF13 and FF11 using the derivative data.	124
Fig. 4.31 Cross-correlation of the Ruaga formation.	125
Fig. 4.32 plot of cross-correlation of the Heira formation in FF13 and FF11 using the derivative data.	127
Fig. 4.33 Cross-correlation of the Heira formation using the derivative data.	128
Fig. 4.34 Cross-correlation of the Heira formation using the non-filtered original data (principal components).	129
Fig. 4.35 plot of cross-correlation of the Heira formation in FF13 and FF11 using the derivative data of the non-filtered principal components.	130
Fig. 4.36 Plot of the first principal components (sampled at 5 feet) of well FF12 and its boundaries.	132
Fig. 4.37 Plot showing the smoothed principal component of the correlated sequences (FF12 and FF11).	133
Fig. 4.38 plot of cross-correlation of the Etel formation in FF12 and FF11 using the derivative data.	135
Fig. 4.39 Cross-correlation of the Etel formation.	136
Fig. 4.40 Cross-correlation of the Sheghega formation.	138
Fig. 4.41 plot of cross-correlation of the Domran formation in FF12 and FF11 using the derivative data.	139
Fig. 4.42 Cross-correlation of the Domran formation using the derivative data.	140
Fig. 4.43 Cross-correlation of the Domran formation using the original data (principal components).	141

	page
Fig. 4.44 plot of cross-correlation of the Ruaga formation in FF12 and FF11 using the derivative data.	143
Fig. 4.45 Cross-correlation of Ruaga formation using the derivative data.	144
Fig. 4.46 Cross-correlation of the Heira formation using the original data (principal components).	145
Fig. 4.47 Plot of the first principal components (sampled at 5 feet) of well FF10 and its boundaries.	147
Fig. 4.48 Plot showing the smoothed principal component of the correlated sequences (FF10 and FF12).	149
Fig. 4.49 plot of cross-correlation of the Etel formation in FF10 and FF12 using the derivative data.	150
Fig. 4.50 Cross-correlation of the Etel formation using the derivative data.	151
Fig. 4.51 plot of cross-correlation of the Sheghega formation in FF10 and FF12 using the derivative data.	152
Fig. 4.52 Cross-correlation of the Sheghega formation using the derivative data.	153
Fig. 4.53 plot of cross-correlation of the Domran formation in FF10 and FF12 using the derivative data.	155
Fig. 4.54 Cross-correlation of the Domran formation using the derivative data.	156
Fig. 4.55 plot of cross-correlation of Ruaga formation in FF10 and FF12.	157
Fig. 4.56 Cross-correlation of the Ruaga formation.	159
Fig. 4.57 plot of cross-correlation of the Heira formation in FF10 and FF12.	160
Fig. 4.58 Cross-correlation of the Heira formation.	161
Fig. 4.59 Plot showing the cross-correlation of small bed within the Sheghega formation in well FF13 and the Sheghega formation	

	page
in FF11 using the derivative data.	163
Fig. 4.60 Cross-correlation of small bed within the Sheghega formation of FF13 and the Sheghega formation in FF11 using the derivative data.	164
Fig. 4.61 Cross-correlation of small bed within the Sheghega formation of FF13 and the Sheghega formation in FF11 using the original data (principal components).	165
Fig. 4.62 Filtered principal components of the studies boreholes.	166
Fig. 4.63 Structural cross-sections of the studied wells in the Attahaddy field. (a) cross-section using the geological formation tops. (b) cross-section using the formation tops from PCAXCOR.	170
Fig. 4.64 Correlatable units in the studied wells in the Attahaddy field.	171

Chapter 5

Fig. 5.1 Illustration of the procedures of reading well-log tapes using LIS/A software.	175
Fig. 5.2 Distribution of files in the Sun workstation and programs in in the S system.	178
Fig. 5.3 Flow chart of the main program PCAXCOR.	179

Appendix A

Fig. 1 Location map of the boreholes in the Attahaddy field.	203
--	-----

Appendix B

Fig. 1 Schematic diagrams showing how different well-logging tool work.	214
---	-----

	page
Fig. 2 Symbols used in log interpretation.	215

List of Tables

Chapter 3

Table 3.1 Showing the variance-covariance matrix, eigenvalues, eigenvectors, and the percentage of each eigenvalue to the total variance in well FF13-6.	34
Table 3.2 Showing the diagonal, off-diagonal and the tri-diagonal elements of correlation matrix of a data from well FF13-6.	35
Table 3.3 Showing the correlation matrix, eigenvalues, eigenvectors and the percentage of each eigenvalue to the total variance in well FF13-6.	42

Chapter 4

Table 4.1 Longitude, latitude, and the formation tops in the Attahaddy field.	89
Table 4.2 The geological formation depths and the predicted formation depths of FF7 using the boundary identification technique.	92
Table 4.3 The geological formation depths and the predicted formation depths of FF13 using the boundary identification technique.	95
Table 4.4 The geological formation depths and the predicted formation depths of FF11 using the boundary identification technique.	111
Table 4.5 The geological formation depths and the predicted formation depths of FF12 using the boundary identification technique.	134

Table 4.6 The geological formation depths and the predicted formation depths of FF10 using the boundary identification technique.	146
Table 4.7 Comparison of PCAXCOR correlation of real data in the Attahaddy field to the geologic selection.	172

Appendix B

Table 1 Well-logging tools, their measured parameters, and factors affecting the measurements.	220
--	-----

Appendix E

Table 1 The correlation matrix, the eigenvalues, the eigenvectors, and the percentage of each eigenvalue to the total variance of the original data of well FF7.	323
Table 2 The correlation matrix, the eigenvalues, the eigenvectors, and the percentage of each eigenvalue to the total variance of the original data of well FF13.	324
Table 3 The correlation matrix, the eigenvalues, the eigenvectors, and the percentage of each eigenvalue to the total variance of the original data of well FF11.	325
Table 4 The correlation matrix, the eigenvalues, the eigenvectors, and the percentage of each eigenvalue to the total variance of the original data of well FF12.	326
Table 5 The correlation matrix, the eigenvalues, the eigenvectors, and the percentage of each eigenvalue to the total variance of the original data of well FF10.	327

ABSTRACT

Many studies have employed automated procedures in well-log interpretation to aid the identification of formation boundaries, and perform cross-correlation between formations in different boreholes. These investigations have demonstrated the use of only one well-log variable, usually resistivity or Gamma Ray logs in the process. In conventional well-log interpretation, different well-log variables or a group of variables are used for different tasks. This project makes use of principal components and spectral analysis as the basis for well-log interpretation, including automatic formation boundary identification and cross-correlation using the first principal component of well-log variables.

By transforming a set of well-log data using principal components analysis a single new variable is extracted from the first principal component scores which accounts for a significant amount of the variation within the original data. A further improvement in the results is obtained by passing the data through a moving filter to reduce noise.

Boundary identification is performed by generalized distance (D^2) method.

Cross-correlation between the filtered principal components of two boreholes is then made by matching each formation of one borehole with a part of another. Both the stretch factor which accounts for thickening or thinning between sequences and the relative vertical displacement of the formation are calculated. This requires the calculation of the power spectra,

derived from the fast Fourier transform of the principal component data, with high pass filtering using the derivative filter to obtain the appropriate resolution.

This new technique was applied on model and real well-log data from five boreholes in the Attahaddy field, Libya. Although the Attahaddy field is structurally complex, the method was found to be reliable at predicting both the geological boundaries of the different formations, and the correlation of formations between boreholes.

The distinctive value of this new approach is in its application of the first principal component of the original well-log variables. Such application has many advantages over the previous studies.

PCAXCOR is a new computer program written in F77 to perform all the necessary computation for boundary identification and well-to-well correlation based on principal component analysis. Graphical output of the results uses a number of new functions in the S language.

CHAPTER ONE

Introduction

An understanding of subsurface geology is gained from boreholes and measurements of various parameters by well-logging tools. The information obtained is used to identify boundaries and to establish correlation of strata between different wells. Traditionally boundary identification and well-to-well correlation are performed using only one log, often a resistivity or gamma ray log and requires a thorough understanding of these logs and their properties. In this way, a correlation of these logs is conducted by identification of similar waveforms in the two logs and can be performed either manually or automatically. For manual correlation, where the success of correlation depends on the geologist's ability to recognize similar patterns through visual comparison it is impractical to utilise all the log data. However, using a computer and data reduction techniques like Principal Component Analysis it is possible to make better use of the available data.

1.1 Purpose and scope

This project describes a computer based technique for generating cross-sections from well log data. Principal Component Analysis is used to reduce the complexity of the multivariate well log data to a single new variable combining all the characteristics of the digital data. This new variable is then used to automatically identify important stratigraphic boundaries within each borehole, before correlating pairs of well log data using a combination of statistical cross-correlation which measures the similarity between two signals as a function of time shift and Spectral Analysis.

A computer program called PCAXCOR was developed to perform these calculations and tested on various model data sets, before applying it to some real data from the Attahaddy gas field in Libya.

1.2 Previous work

Geologists have repeatedly succumbed to the temptation to use cross-correlation in subsurface geology, and there are many studies of automatic segmentation and correlation of well-logs by computer.

Zonation is the process of dividing a sequence into relatively uniform segments, each of which is distinctive from adjacent segments. Well logs

may be subdivided into relatively uniform segments that represent zones of constant lithology, corresponding to stratigraphic units.

There are basically two contrasting approaches to zonation. The simplest procedure is *local boundary hunting* which searches for abrupt changes in average values, or equivalently, for the steepest gradients in the sequences (Davis, 1986). A "split-moving window" for defining boundaries between soil zones along transect was developed by Webster (1973). A sequence is examined by iteratively moving a short interval along the sequence. The moving interval is called a window and is split into two parts. A measure called the generalised distance D^2 is calculated for the difference between the segment within the two halves of the window. Webster noted that the performance of the procedure depends upon the variability of the original sequence and the length of the moving window. Webster (1980) has published a FORTRAN program that finds the zone boundaries by this method.

The main objection to local boundary hunting procedures that they are dependent on the size of the window used to identify the boundaries. A long window will average across small zones and may miss short intervals, however, a short window is more sensitive and will identify small zones and may find an inordinate number of boundaries.

Global zonation is a different approach, using procedures that break the sequence into a specific number of segments which are as internally homogeneous as possible and as distinct as possible from adjacent segments. An iterative analysis of variance was first used by Gill (1970). First the sequence is divided into two segments, a short initial segment, and the remainder of the sequence. The partition between the two segments is moved along the sequence to successive positions and at every position the sum of squares within the segments and the sum of the squares between the segments are calculated. The maximum value of the ratio between the two sums divided by the sum of the squares between the segments is considered as the location of the first zonal boundary. Next, the two zones are themselves partitioned by repeating the process to insert an additional boundary which again minimizes the difference ratio. By repeating the process, the entire sequence will be divided into the specific number of zones. Hawkins and Merriam (1973, 1974) used Gill's iterative procedure, but adopted a recursive method and took advantage of Bellman's principle of optimality to ensure that the final set of zone boundaries is the best possible of all set of partitions that might have been chosen. With the nonrecursive procedure, it is always possible that the position selected as the best boundary between two zones is no longer the best when another boundary is inserted into one of the zones. Webster (1973) collected 27 soil properties at 20m intervals. These multiple measurements were compressed by Principal Component Analysis to identify boundaries between segments such that the variance within segments on either side of boundary was minimal. The computational cost of achieving this iterative optimality is very high, and the method is not practical for very long log sequences.

More recently Elek (1988) has used PCA on well-log data for boundary identification and correlation, though the latter is not discussed in any detail. Later Elek (1990) shows how an estimate of porosity can be made from the first principal component of selected variables.

Correlation of subsurface data is the next step required to establish a *framework* into which new data can be fitted and as an aid to understanding the stratigraphy of the area of interest.

A time series is a set of values of a function sampled at equal intervals. Well log data be considered as time series data and can therefore be analysed using time series techniques though problems arise with the variation in thickness and vertical offset of units between different boreholes. Cross-correlation was first used by Weiner (1949) to determine the displacement of two time series on each other in time domain. Jenkins and Watts (1969) have discussed a method for analysing time series in both the time and frequency domain.

Based on an existing computer program, Daskam (1964) described an integrated computer process and emphasized the need for automation in well log analysis. Matuszak (1972) used a normalised cross-correlation function to measure the displacement of similar shaped segments in dipmeter and resistivity curves. The purpose of correlating the curves was to determine the displacement between them. In his investigation he did

not address the problem of thickening and thinning and concluded that more research was needed to refine existing methods or develop new techniques. Haites (1963) described a graphical method based on the fact that thickening (and thinning) of a stratigraphic sequence is common and this creates a stretched (or compacted) log signal. His approach which he called *prospective correlation* considered this effect by giving different degrees of comparison of the depth scale until a correlation was found. No mathematical or automatic processes were involved. Neidell (1969) was the first one who considered this problem in the automated cross-correlation. Neidell proceeded with the correlation after using an interpolated section to compensate for the thinning of beds because cross-correlation can detect only the shift between time series and can not detect the thinning or thickening of the strata. Rudman and Lankeston (1973) attempted to solve this problem by comparing the autocorrelation function and the cross-correlation function of iteratively stretched intervals. An improved method of normalized cross-correlation function and frequency domain was used by Rudman and Henderson (1975). Although both methods were successful, iterative stretching and correlation require considerable computing time. If, in addition, the geologist is unsure which log is to be stretched, the procedure must be performed twice. Without relying on the iterative operations, Kwon (1977) has successfully used a sophisticated algorithm for correlation of well-log data. This procedure indicated the feasibility of correlating well log data and our investigation owes much to his development. He used cross-correlation of power spectra of the logs to identify both the stretching factor and the relative displacement between logs in one simplified operation. Computations were performed in the frequency domain with the frequency intervals transferred to the logarithmic

scale. Interpolation was required to obtain equally spaced power spectra. Given the stretch, the displacement between two curves is computed rapidly by cross-correlation method. Although the technique was successful, however, it depends upon the input logs. Different logs from the same borehole gave different displacements. A difference of 30 feet between the program results and the stratigraphic displacement was considered to be excellent, and a difference of 60 feet was considered to be fair.

CHAPTER TWO

Geological Setting

2.1 Introduction

Libya is situated on the Mediterranean foreland of the African shield. The Sirte basin is located in the central part of Libya and occupies an onshore area of approximately 492,000 Km² (Fig. 2.1).

This area was largely peneplaned by a lengthy and intense erosional phase during the Late Pre-Cambrian. During the Early Cambrian to Middle Devonian, the Caledonian Orogeny created several northwest trending structural elements. One such element was the Calanshio trough, which was later to influence the formation of the Sirte basin (Cain, 1985), (Fig. 2.2, Fig. 2.3). Furthermore, the Sirte basin is a late Mesozoic-Tertiary cratonic rift resulting from crustal extensions of the older basement and Paleozoic rocks. The tectonic evolution of the Sirte basin controlled the sedimentation and provenance of sedimentary material that built the stratigraphic section of the area.

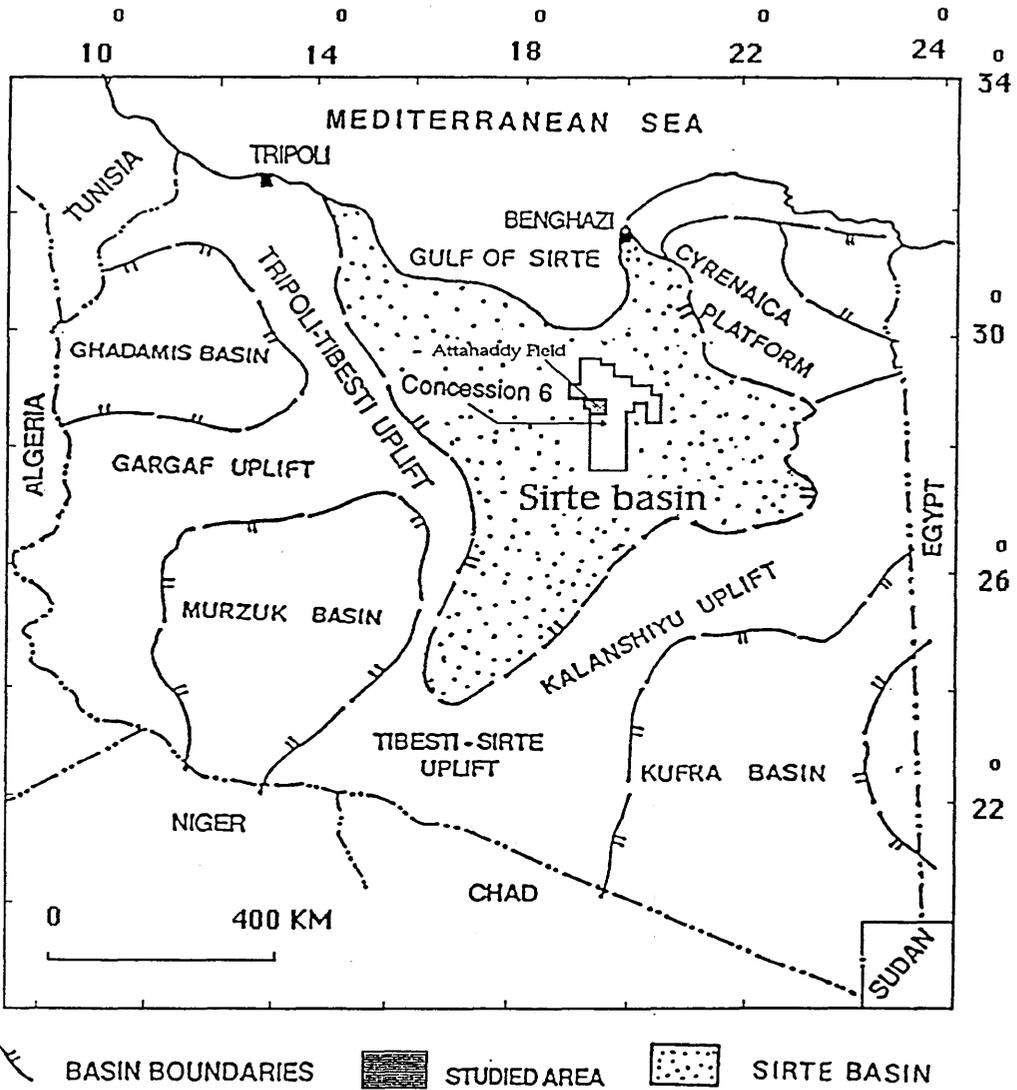


Fig.21 Location map of the Area studied within Sirte Basin, Libya.

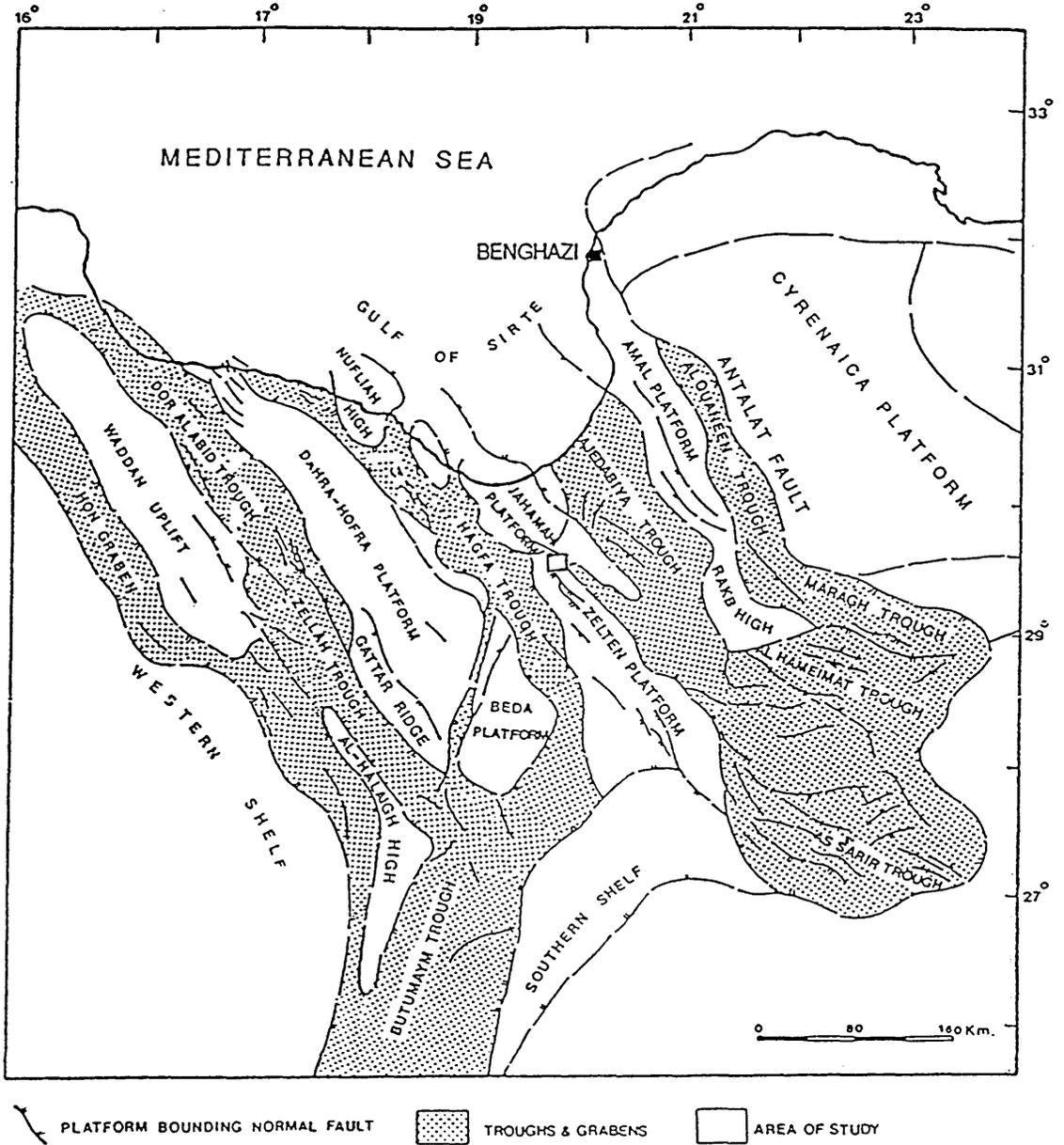


Fig. 2.2 Map showing structural elements of the Sirte Basin, after Mouzoghi & Taleb (1980).

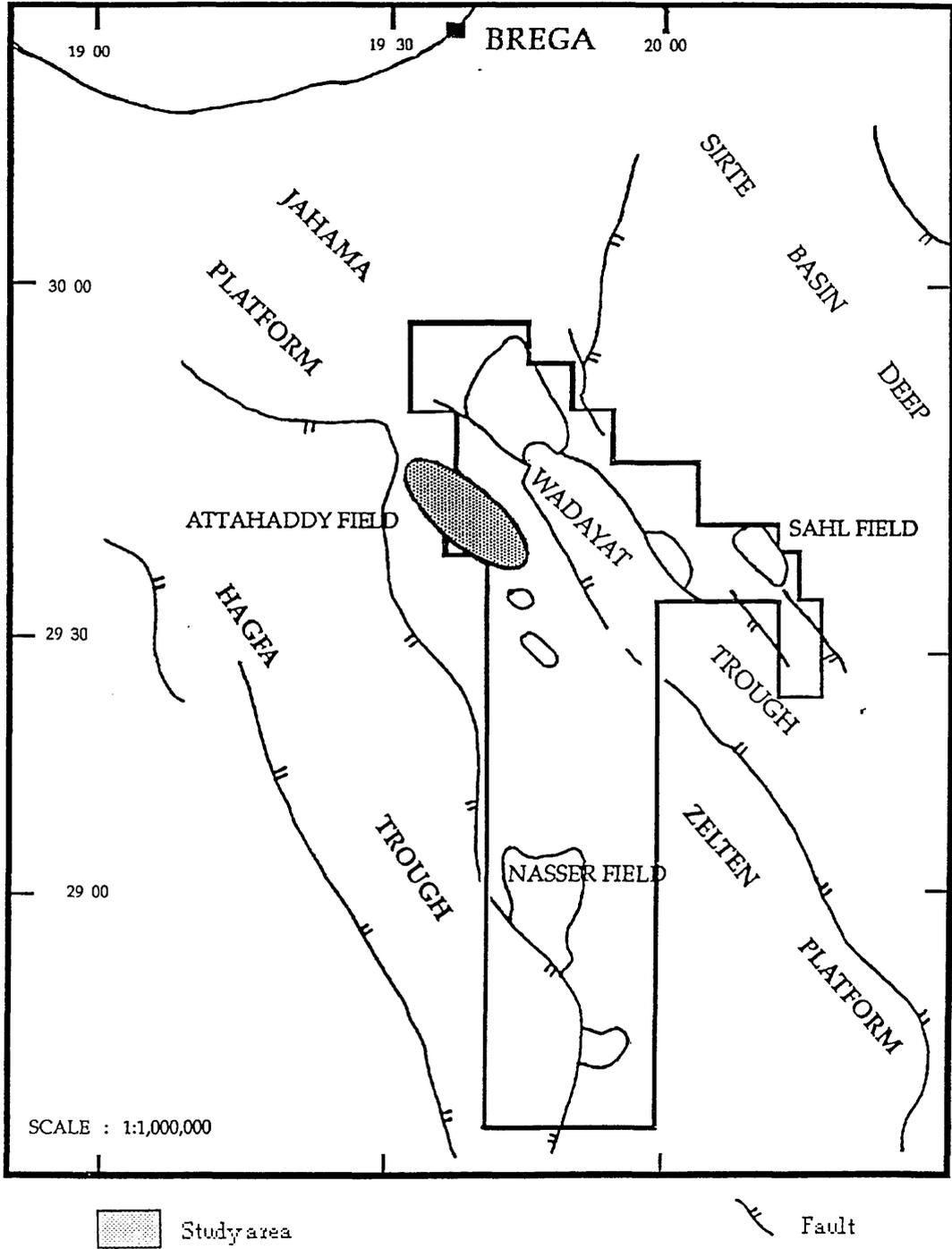


Fig. 2.3 Concession map and major tectonic elements in Sirte basin

2.2 Petroleum exploration history in the Sirte basin

Exploration has taken place in the Sirte basin for over thirty years. From 1956 through to the end of 1977, 3,102 wells had been drilled in the search for hydrocarbons, of which 984 are wildcat, 675 are outpost and 1,442 are development wells. There is estimated to be between 23 to 30 billion barrels of oil and 634 billion cubic metres of gas. This is pooled in 121 reservoirs with 101 fields. Twenty of the major fields are "giant" with reserves in excess of one billion barrels (El-talhi, 1990, Parson et al, 1980).

In 1988, the exploratory drilling in Libya was about 70,136 m and 71.7 rig-months. Most of the activity (93%) took place on the land with 83% concentrated in the Sirte basin.

2.3 Area of study

The Attahaddy field is located in the NW part of concession 6 on the Zelten Platform of the Sirte basin in central Libya (Fig. 2.3). It covers approximately 43,300 acres.

To date sixteen boreholes have been drilled (Appendix A). Two boreholes were drilled by Esso Standard-Libya. FF1-6 was drilled in 1964, and detected gas from a thin Bahi Formation section, and drilled 92 feet into the Gargaf Formation (see Fig. 2.8 for stratigraphy), but this section was not tested. The FF2-6 well was drilled 100 feet into the Gargaf Formation in 1967.

It detected gas, but this test was not conclusive. Exploration enthusiasm has increased following the establishment of the Sirte Oil Company as a national entity in late 1981. A large scale program of modern seismic activity was laid out, then in 1985 two deeper exploration holes FF3 and FF4 located on a structural closure, mapped on the Cambro-Ordovician were drilled resulting in the discovery of gas. These wells were then evaluated and after several outpost and exploration holes had been drilled, a large amount of gas was estimated. The sole reservoir in the Attahaddy field is in the (Cambro-Ordovician) Gargaf Formation, which is composed of highly fractured, dense, massive quartzitic sandstone and quartzite.

2.4 Structure

Since Early Palaeozoic times Libya has been the site of deposition of extensive sheets of continental clastic sediments and several transgressions and regressions by the sea with consequent accumulation of a wide variety of marine sedimentary rocks. Five sedimentary basins were formed in Libya by several tectonic cycles, these basins, (Ghadamis which is known as Hamada, Murzuk, Kufra, Western Desert and Sirte basin) are separated by intervening uplifts (Conant & Goudarzi, 1967), (fig 2.1).

During the Early Cambrian to Middle Devonian, the Caledonian Orogeny formed several north west trending structural elements. Three major axes of uplift are recognised. In the south-west the Tihembika uplift occupies the border region between Libya and Algeria. Towards the north-west, the Tripoli-Tebesti uplift and Haruj uplift define a small and narrow

trough, the Dor-el-Gussa trough. Further to the north-east lies the Calanshio uplift. This structural relief of the early Paleozoic Era is the result of regional stretch in a NE-SW direction, Klitzsch (1971), (fig. 2.4). This epeirogenic movement was accompanied by continental sedimentation during the Cambro-Ordovician, a marine transgression marked the onset of the Silurian and marine sediments were widely deposited through the Middle Devonian.

The next phase was the Hercynian Orogeny, active from the Late Devonian to the Early Triassic. These compressional stresses were resolved as a series of northeast trending basins and uplifts, Klitzsch (1971), (fig 2.5). These uplifts prevented any significant marine transgression by the Permian and Early Mesozoic seas, across the Sahara platform. This became a period of intense erosion and ultimately, only the Cambro-Ordovician clastics were preserved over most of the Tebesti-Sirte uplifts.

During the Triassic, only the northwestern and northeastern corners of Libya (Nefusah uplift, Ghadamis Basin and northern Cyrenaica) were covered by the sea. Terrigenous clastics, carbonates and evaporites were deposited unconformably over the Paleozoic rocks.

During the Jurassic, marine areas did not change very much in the north-west, whilst in northern Cyrenaica the sea extended farther to the south. Shallow water carbonates and clastics were deposited to the south and deeper marine sediments accumulated in the north.

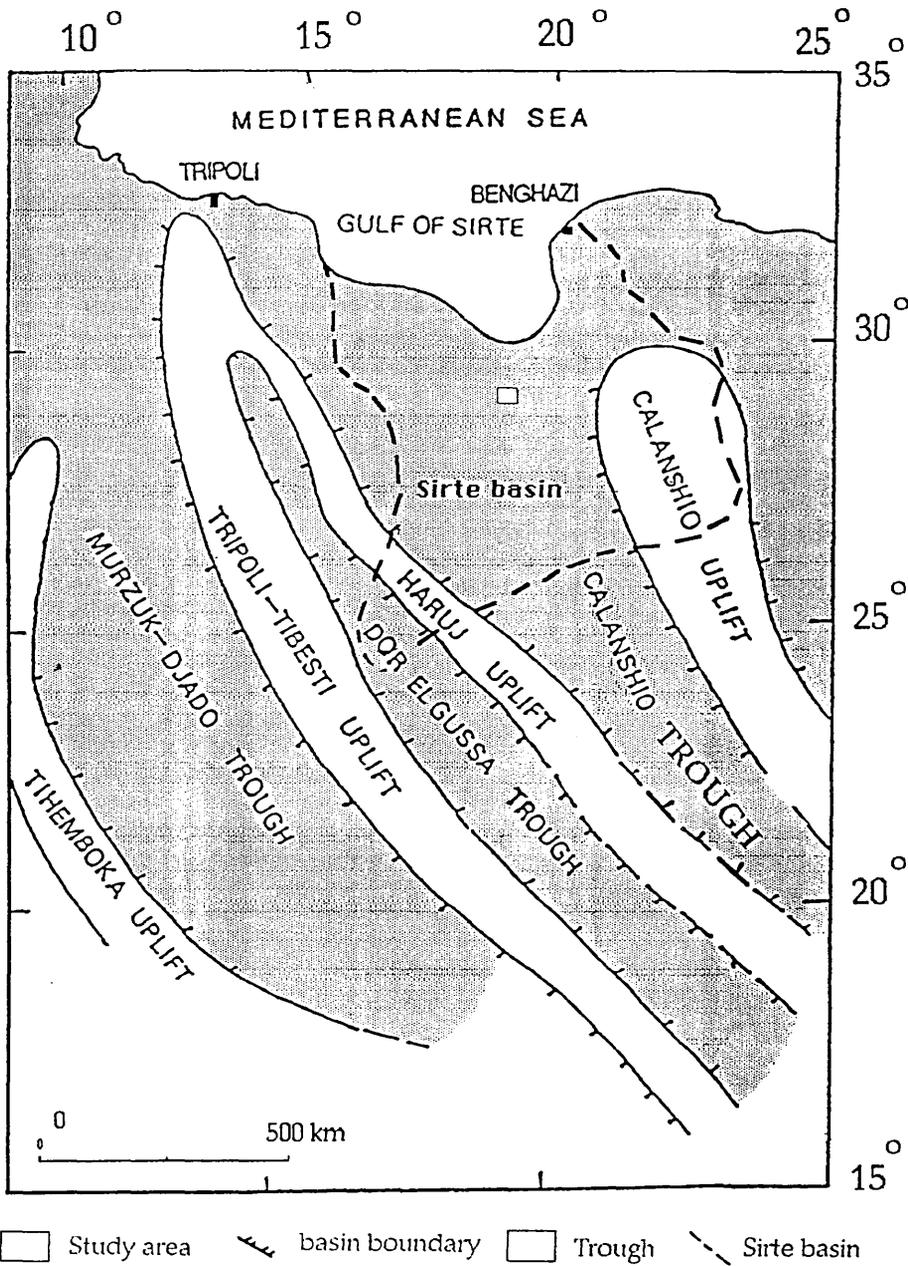


Fig. 2.4 Map showing the major structural elements of the central Sahara in Early Paleozoic times, after Klitsch (1971).

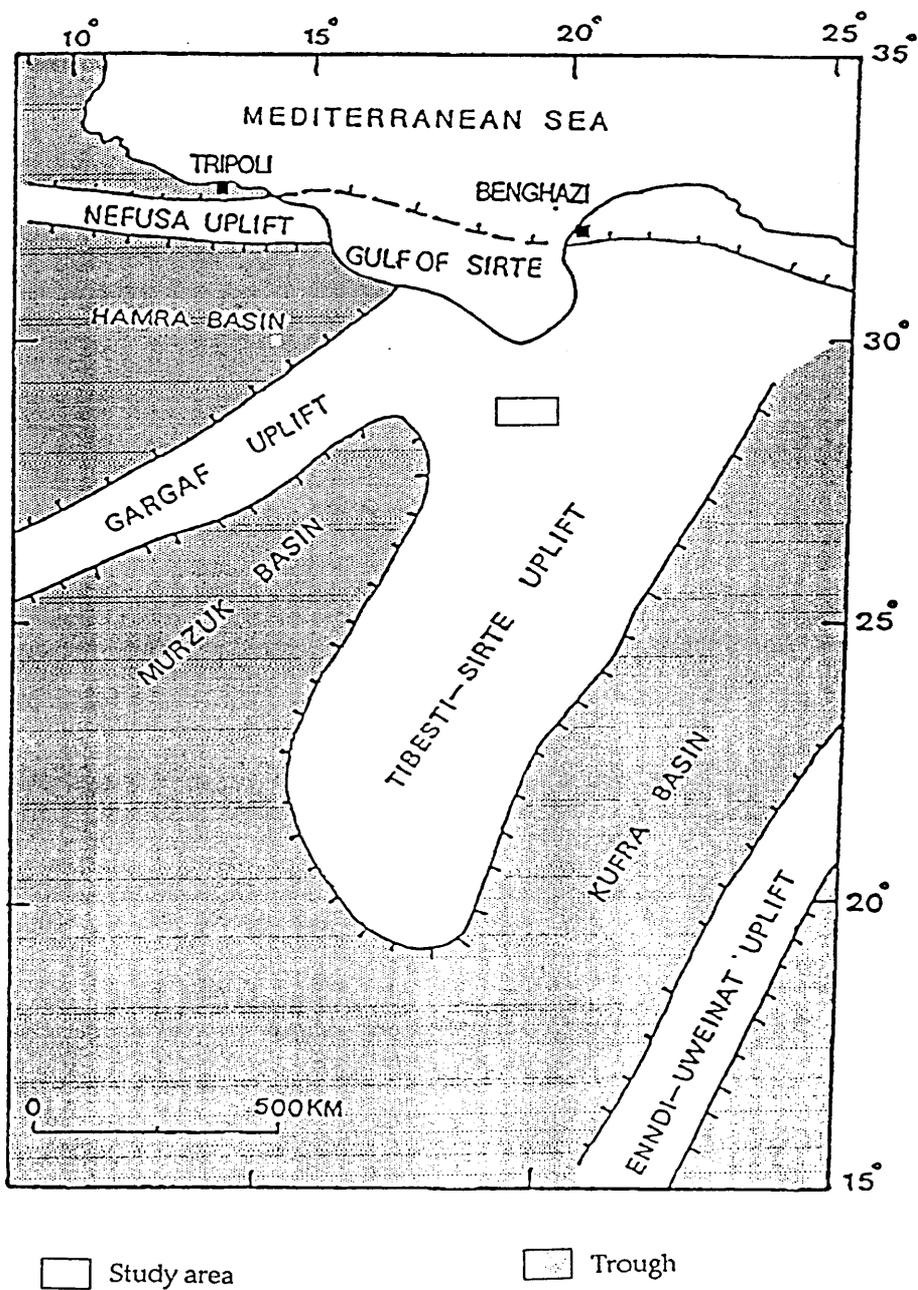


Fig. 2.5 Map showing the major structural elements of the central Sahara in Late Paleozoic and Mesozoic times, after Klitsch (1971).

During the Early Cretaceous (Fig.2.6), an erosional phase was accompanied by right lateral movement, which effects most of north Africa producing a block faulting system. A regression took place in north western Libya, where only continental sediments are recorded whilst northern Cyrenaica was still covered by a shallow sea.

The beginning of the Late Cretaceous (Cenomanian) was characterised by major tensional events, which created the Sirte Basin. Only the major horsts in the Sirte Basin and the Cyrenaica platform remained emergent.

Throughout the Late Cretaceous, the sea continued to advance southwards. By the end of Maastrichtian only a few scattered horst crests remained above the sea as isolated islands (Duronio & Colombi, 1983).

The Tertiary deposits rest conformably on the upper Cretaceous, and are composed entirely of marine sediments. Conditions appear to have remained stable throughout the Tertiary, even with continued rejuvenation of the horst/graben system in the Sirte Basin.

During the late Early Paleocene (Danian), or latest Maastrichtian, the last of the extensive transgressions began, until in late Palaeocene time, the sea stretched between the basins of north Africa and west Africa, Reyment (1966).

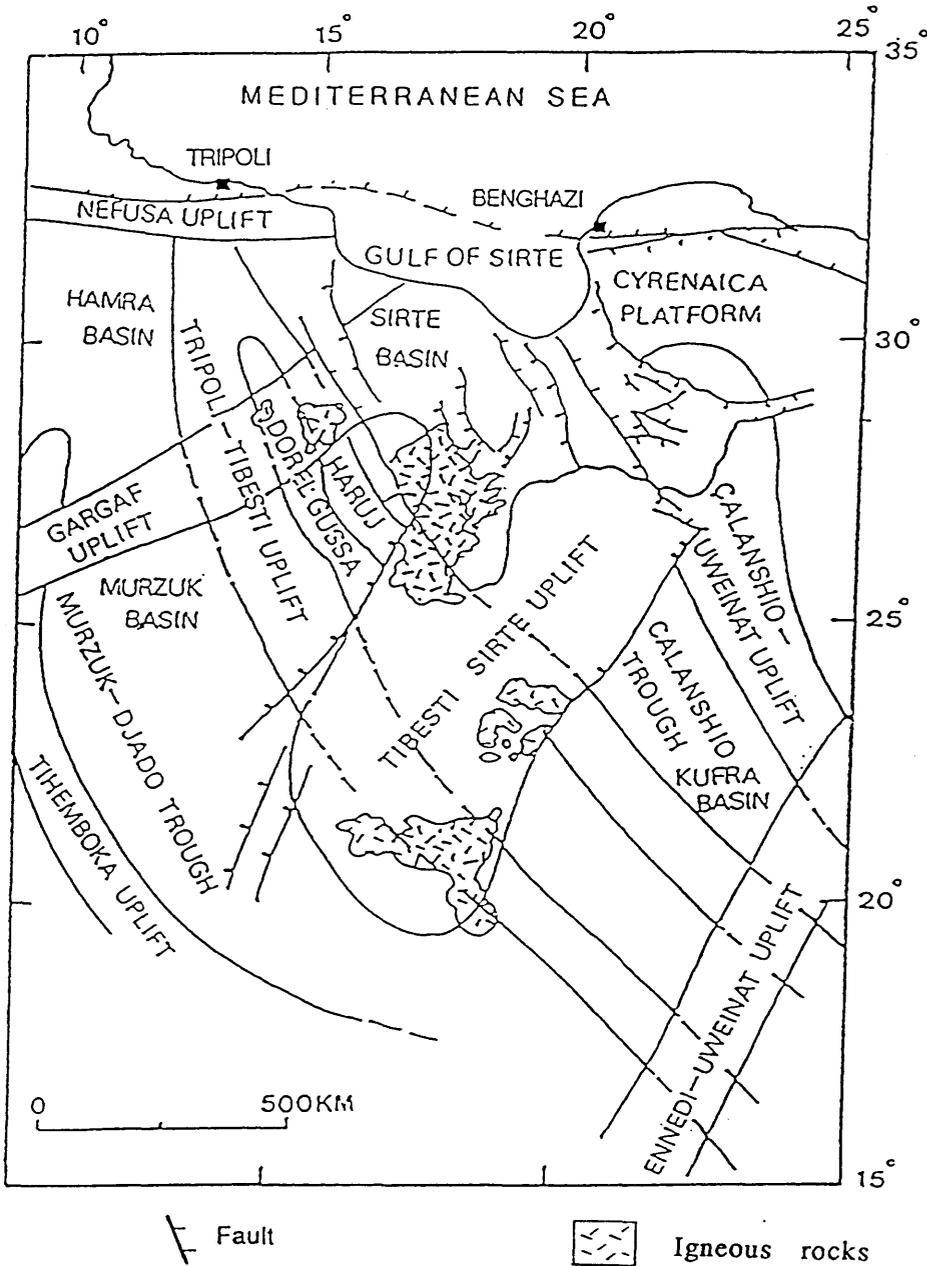


Fig. 2.6 Schematic map of major structural elements, showing the present day basins and uplifts of the central Sahara, resulting from a combination of the effects of two structural regimes, after Klitsch (1971).

2.5 Stratigraphy

The sediment thickness in the Sirte basin is about 2.4 Km, with sediments being mainly Cretaceous and Tertiary (Fig 2.7, Fig. 2.8)

Due to differences in local stratigraphic nomenclature used by different Oil Companies, the nomenclature used here is based on the current Sirte Oil Company's nomenclature.

The stratigraphy of the Attahaddy field (Fig. 2.7, Fig. 2.8) is as following :

2.5.1 Paleozoic

Reconstruction of the Pre-Sirte basin geological history indicates that a moderately thick section of Cambrian to Middle Devonian clastics was deposited, unconformably, on the African shield. However, this Paleozoic section was considerably reduced by subsequent erosion, until only the Cambro-Ordovician Gargaf Formation remained over Concession 6 (Klitsch, 1971).

Gargaf Formation (Cambro-Ordovician)

The Gargaf Formation is composed of very well indurated,

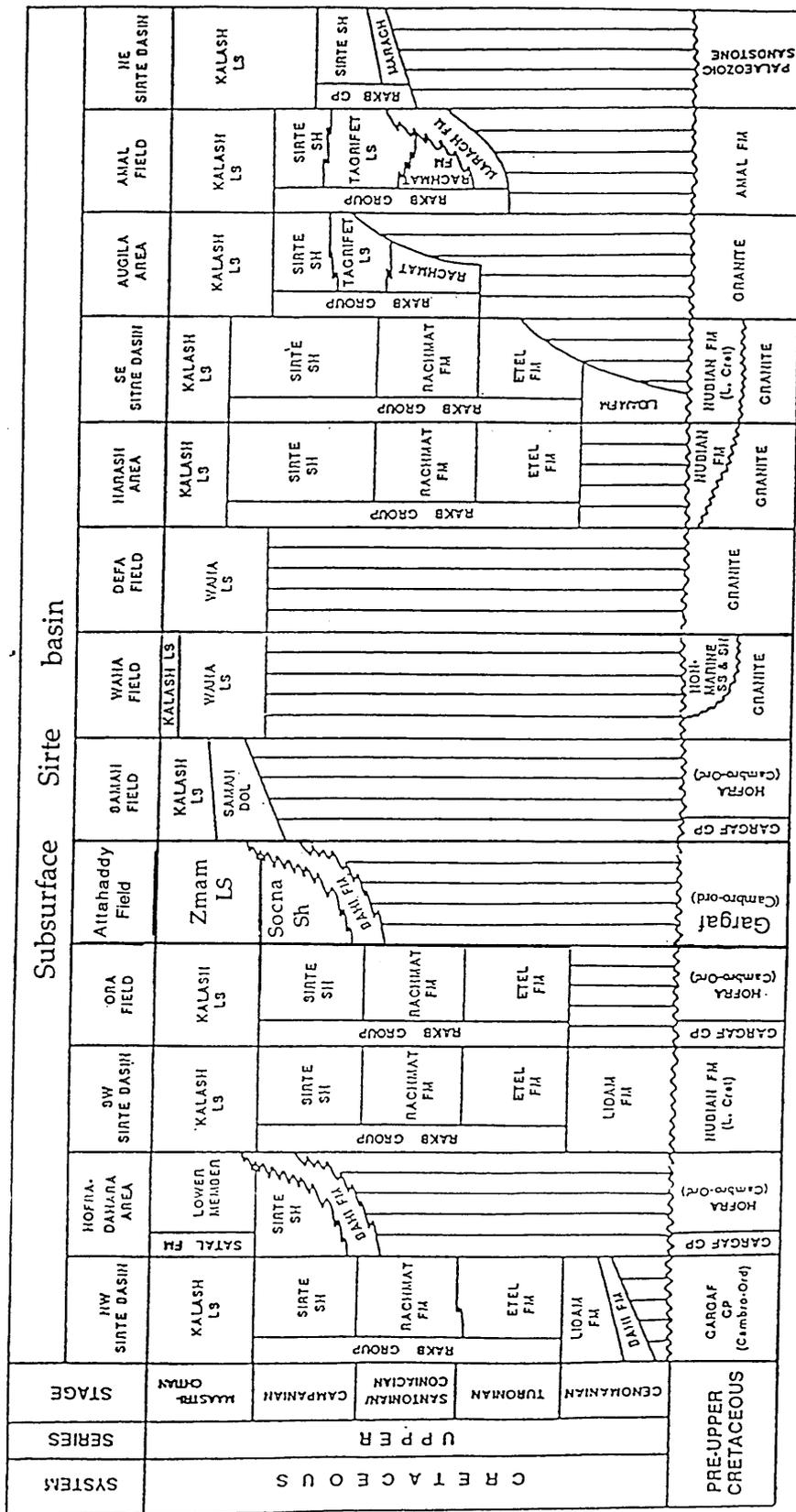


Fig. 2.7 Formation correlation of The Upper Cretaceous in Sirte basin

ERA	System	SERIES	STAGE	GROUP	Formation	MEMBER	LITHOLOGY	
MESOZOIC CENOZOIC	TERTIARY	MIOCENE	Burdigalian					
		OLIGOCENE	Chattian	Gehenna Group	Muailah Formation		Anhy., Ls., ss. & shale.	
			Sannoisian		Etel Formation		shale - l. grey, blocky, m hd.	
		MIDDLE EOCENE	Lutetian	Tamet Group	Sheghega Formation		Limestone - white to l. grey, nummulitic, chackly to soft.	
					Uaddan Group	Domran Formation		Limestone - micritic, l. grey to medium grey brown, tight.
		LOWER EOCENE	Cuisian	FOGHA GROUP	RUAGA FORMATION		Limestone - A/A	
						HEIRA FORMATION		shale - dark grey to greenish-grey, occasionally black calcareous and moderately fissile to blocky
								Micritic L.S.
								shale - grey to brownish grey
		MESOZOIC	CRETACEOUS	UPPER	Danian Maas - trichtian Campanian	Hamada Group equivalent	Zmam FM.	
	Socna FM.						shale - grey to brownish grey	
	Behi FM.						s.s & conglom. white	
PALEOZOIC	CAMBROORDO.		Ashgillian	Gergef Group	Gargaf Formation		Quartzites & Quartz S.S. white, hard, dense, f. grained.	
			Acadian					

Unconformity
 Ls.
 Sh.
 Ss. & Congl.
 Ss.
 Anhy.
 Qtzit.

Fig. 2.8 Stratigraphic section in the Attahaddy field.

translucent, milky-white to very light grey, well sorted, fine grained, angular to subangular quartzitic sandstone and quartzite. Quartz overgrowths are common and have severely reduced the effective porosity of the Gargaf, over much of the area. The Gargaf is considered to be a continental deposit, and is interpreted as a series of fluvial sequences (Cain, 1985). The Gargaf Formation has not been fully penetrated in the Attahaddy field. It is therefore, difficult to comment on it's regional thickness. However, the Gargaf is in excess of 2811 feet at well FF3-6, and 470 feet in FF6-6.

While the original porosity of the Gargaf was destroyed through diagenetic processes, the Caledonian deformation caused the development of high fracture porosity across the crest of these horsts. This phenomenon is well developed in the Attahaddy field, and consequently, the Gargaf forms an important reservoir in this field.

2.5.2 Mesozoic

The Upper Cretaceous sequence rests unconformably on the Paleozoic and is composed of both continental and marine sediments.

Bahi Formation.

The Bahi Formation is composed of moderately well consolidated sandstone and conglomerates. The sandstones are translucent to white, very poorly sorted, with angular to subangular, very fine to coarse grains. The

conglomerates are composed of Gargaf Formation derived pebbles and cobbles. The Bahi Formation is barren of fossils and consequently its age is uncertain (Barr and Weegar, 1972) The upper boundary of the Bahi Formation is rather abrupt with the Socna Formation or other Upper Cretaceous formations. The lower boundary is unconformable with the Paleozoic rocks. In the type area, the Bahi Formation overlies quartzites of Cambrian-Ordovician age. The Bahi Formation is considered as alluvial fan deposits and is associated with the underlying structural highs of the Gargaf.

The Bahi Formation is distributed over some parts of the Attahaddy field, and has an average thickness of 200 feet. In well FF12-6, it reaches a maximum of over 325 feet, whilst to the south in well FF2-6 it is only 17 feet thick.

Socna Formation (Upper Cretaceous)

The Socna Formation is a grey to brownish-grey, calcareous shale, and is typical of a marine environment which took place in the Campanian-Maastrichtian times. The Socna shale is considered to be the prime hydrocarbons source in the Sirte basin. It achieves a regional thickness of approximately 1800 feet in FF14-6, and a thickness of only 55 feet in FF2-6. The average thickness of these shales varies from 500 to 1000 feet.

Zmam Formation (Upper Cretaceous)

The Zmam Formation is composed of micritic limestone, very light grey to light brown, soft to medium hard, argillaceous in part, occasionally fossiliferous. This is a deep water to basinal deposit of the Maastrichtian to the Lower Palaeocene times. It is not normally a reservoir, but since it is tight micritic, it makes a good seismic reflector which maps the top Cretaceous deposits. The Zmam thickness varies from 680 feet in FF14-6 to 97 feet in FF3-6. In other parts of the field the thickness ranges from 300 to 400 feet.

2.5.3 Tertiary

The Tertiary section is predominantly shale and limestones. Sedimentation through the section was controlled by slow tectonic movement and gradual subsidence, thus the depositional environment was deep marine that changed laterally to shallow marine in localized areas..

Heira Formation (Paleocene)

The Heira Formation is composed of a regional thick section of shale. These shales are dark grey to greenish-grey, occasionally black, calcareous and moderately fissile to blocky.

Heira shales are distributed over the study area and reach a thickness

of 2700 feet in well FF14-6 and minimum thickness of 1690 feet in well FF8-8. In general the average thickness of the Heira Formation all over the field is 2000 feet.

Ruaga Formation (Paleocene-lower Eocene)

The Ruaga Formation is composed of limestone and cannot be easily differentiated from the overlying Domran carbonates. The lower boundary is conformable with the Heira shales. Both contacts are conspicuous on the electric logs in the type section. In the southern central Sirte basin, the Ruaga Formation is predominantly a limestone with very subordinate amounts of shale and has been divided according to this distribution by Sirte Oil Co. geologists into a number of members. These members are the Megil shale, Zelten limestone and Cra carbonates. The Zelten member forms the principal reservoir of the Zelten fields south of the Attahaddy Field, however, the Zelten member in Attahaddy Field loses its typical identity as shallow water limestone grades into open marine deposits. The Ruaga Formation is the product of an open marine shelf environment, and this interpretation is supported by the presence of the Zelten member developed, at scattered locations, across some parts of Sirte basin (Cain, 1985). The Ruaga Formation covers all the Attahaddy field and ranges in thickness between 160 to 200 feet.

Domran Formation (Lower Eocene)

The Domran Formation consists of limestones, micritic, light grey to

medium grey-brown and tight and considered to be the product of an open marine shelf to marginal shelf environment that existed during the Tertiary. It reaches a thickness of 1280 feet in FF2-6, and 650 feet in FF7-6. The average thickness of the Domran Formation is approximately 900 feet.

Sheghega Formation (Middle Eocene)

The Sheghega Formation is a thick carbonate unit, distributed across the entire study area. It is composed of limestone, white to light grey, nummulitic in parts to very nummulitic, chalky in part, soft to moderately hard, with porosity varying from poor to good. The depositional environment interpreted for the Sheghega Formation is an open marine shelf to margin setting. These regional carbonates cover all the study area and vary in thickness from 2150 feet to 1800 feet.

Etel Formation (Oligocene-U. Eocene)

The Etel Formation is composed of light grey to greenish-grey, blocky, medium hard to hard, calcareous shale. These deep marine shales are distributed across the entire study area.

Muailah Formation (Oligocene)

The Muailah Formation consists of thin interbedded carbonates, shale, anhydrite and siltstone or very fine grained sandstones. The carbonates are grey to brown with subordinate limestone, the shales contain pyrite in

CHAPTER THREE

Theoretical Background

3.1 Introduction

In this chapter the mathematical background to the treatment of well log data is discussed. Each well-log is first subjected to a principal component analysis to reduce the complexity of the original data to a single new variable or principal component that incorporates most of the variation present in a number of individual well-log variables. The data are then filtered or smoothed to further reduce the noise in the signal, before processing to identify geological boundaries. Finally the processed log data from pairs of wells are compared to establish the connection between geological units, which includes making an allowance for variation in thickness and depth of rock units in the different boreholes.

3.1.1 Basic Statistical Calculations

The following are some definitions of statistical terms used in performing the Principal Component Analysis (PCA). Before calculating the

principal component scores, the eigenvalues and the eigenvectors of the symmetric correlation or variance-covariance matrix must be calculated.

The way to calculate all these entities is described below.

For a sequence of n values of x :

Mean : is defined as the sum of all the observations(x_i) divided by the number of observations(n).

$$\bar{x} = \sum_{i=1}^n x_i / n \dots\dots\dots(3.1)$$

Sum of Squares : is the sum of the squared difference between the variable and its mean.

$$SS = \sum_{i=1}^n (x_i - \bar{x})^2 \dots\dots\dots(3.2)$$

Variance : is the average squared deviation of all observations from the mean.

$$s^2 = \frac{SS}{n-1} \dots\dots\dots(3.3)$$

Standard deviation : is the square root of the variance.

$$s = \sqrt{s^2} \dots\dots\dots(3.4)$$

Covariance : calculating the covariance requires a quantity analogous to the sum of squares called the *corrected sum of products*:

$$SP_{jk} = \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k) \dots\dots\dots(3.5)$$

where x_{ij} is the i -th measurement of variable j , x_{ik} is the i -th measurement of variable k , \bar{x}_j and \bar{x}_k are the mean of variable j and k respectively, and SP_{jk} is the sum of products between variable j and k

then the covariance between variables j and k is

$$COV_{jk} = \frac{SP_{jk}}{n-1} \dots\dots\dots(3.6)$$

Correlation :

The correlation between two variables x and y can be defined as the covariance between two variables divided by their standard deviation :

$$r_{xy} = \frac{\text{COV}_{xy}}{s_x s_y} \dots\dots\dots(3.7)$$

Where COV is the covariance between the two variables and s_x and s_y are the standard deviation of x and y respectively.

The correlation function is used in order to estimate the degree of interrelation between variables in a manner not influenced by measurement units. Because the correlation function is the ratio of the covariance of two variables to the product of their standard deviation, this function is unitless. It ranges from +1 to -1. A correlation of +1 indicates a perfect direct relationship between the variables, whereas a correlation of -1 indicates that one variable changes inversely with relation to the other. Zero correlation, however, indicates the lack of any sort of relationship at all (Davis, 1986).

Eigenvalues and Eigenvectors :

This topic is regarded as the most difficult topic in matrix algebra. The difficulty is not in their calculation, which is cumbersome but no more so than many mathematical procedures. Rather, difficulties arise in developing

a "feel" for the meaning of these quantities (Davis, 1986). The relationship between a data matrix $[A]$ and the vector of eigenvalues λ and matrix of eigenvectors $[U]$ is the solution of the equation :

$$[A][U] = \lambda[U] \dots\dots\dots(3.8)$$

where $[A]$ is the data set matrix,

$[U]$ is the eigenvectors,

λ is the eigenvalue.

In the simple case where $[A]$ is a 2×2 matrix the eigenvectors yield the orientation of the ellipse axes and the eigenvalues represent the magnitude, or lengths, of the axes. The sum of the eigenvalues of the matrix is always equal to the sum of the diagonal elements, or the trace, of the original matrix. There will be as many eigenvectors as there are eigenvalues, or as many as there are rows and columns in the matrix $[A]$. Table 3.1 shows the variance-covariance matrix, eigenvectors and the percentage of each eigenvalue to the total variance.

By solving the simultaneous equations, the eigenvalues and the eigenvectors can be calculated. Although this technique is extendible to any size matrix, finding the roots of a large polynomial can be an arduous task. Usually, eigenvalues are not found by solution of polynomial and quadratic equations as root-searching is usually a very poor computational method

(Press, 1988), but rather by matrix manipulation methods that involve refinement of a successive series of approximations to the eigenvalues (Davis, 1986). These methods are practical because of the great computational speed of digital computers.

The optimum strategy for finding eigenvalues and eigenvectors is, first, to reduce the matrix to a simple form, only then beginning an iterative procedure. For symmetric matrices, the preferred simple form is the tridiagonal matrix (Press, 1986)(Table 3.2).

There are two ways to reduce a symmetric matrix to tridiagonal form. The Givens reduction is a modification of the Jacobi method. The Jacobi method reduces the matrix to a diagonal form, whereas the Givens reduction stops when the matrix is tridiagonal. This allows the procedure to be carried out in a finite number of steps, unlike the Jacobi method which requires iteration to achieve convergence (Press, 1986).

The Givens method is not generally used because the reduction involves taking square roots when the computation is performed. The Householder method is more efficient and more stable. It reduces an $n \times n$ symmetric matrix to tridiagonal form by $n-2$ orthogonal transformation. The Householder reduction method has been used in subroutines TQLI and TRED2 for calculating the eigenvalues and the corresponding eigenvectors of the variance-covariance or correlation matrix (see Appendix C).

THE VARIANCE-COVARIANCE MATRIX IS :

14.990	0.665	2.495	0.802	4.070	-7.147	1.840
0.665	39.397	-1.907	-1.831	-3.705	11.345	0.529
2.495	-1.907	4.615	3.850	6.471	-47.227	-1.594
0.802	-1.831	3.850	3.966	6.516	-46.962	-1.941
4.070	-3.705	6.471	6.516	12.665	-67.801	-3.043
-7.147	11.345	-47.227	-46.962	-67.801	716.010	22.084
1.840	0.529	-1.594	-1.941	-3.043	22.084	2.243

EIGENVECTORS:

VARIABLE	1	2	3	4	5	6	7
SP	-0.01065	0.01880	-0.92359	-0.30802	-0.21578	0.00992	-0.07060
GR	0.01716	0.99536	-0.01549	0.09047	0.02294	0.00226	0.00069
ILS	-0.06585	-0.03211	-0.16217	0.22591	0.35051	0.87118	0.18841
ILM	-0.06543	-0.03080	-0.06123	0.31628	0.11950	0.05458	-0.93474
ILD	-0.09512	-0.07766	-0.32122	0.79240	0.12822	-0.39006	0.29199
DT	0.99044	-0.02889	-0.05211	0.11716	0.01302	0.03318	-0.02172
CALI	0.03077	0.00726	-0.10393	-0.32289	0.89396	-0.29107	-0.00754

EIGENVALUES

729.72363	39.49113	16.71196	6.32060	0.81314	0.71658	0.11020
-----------	----------	----------	---------	---------	---------	---------

PERCENTAGE OF TOTAL VARIANCE CONTRIBUTED BY EACH EIGENVALUE

91.91778	4.97440	2.10508	0.79615	0.10242	0.09026	0.01388
----------	---------	---------	---------	---------	---------	---------

Table 3.1 showing the variance-covariance matrix, eigenvectors, eigenvalues and the percentage of each eigenvalue to the total variance in Well FF13-6. Seven variables are used, Spontaneous Potential (SP), Gamma Ray (GR), Shallow resistivity (ILS), Medium resistivity (ILM), Deep resistivity (ILD), Sonic (DT) and the caliper log (CALI).

THE CORRELATION MATRIX IS :

1.000	0.027	0.300	0.104	0.295	-0.069	0.317
0.027	1.000	-0.141	-0.146	-0.166	0.068	0.056
0.300	-0.141	1.000	0.900	0.846	-0.822	-0.496
0.104	-0.146	0.900	1.000	0.919	-0.881	-0.651
0.295	-0.166	0.846	0.919	1.000	-0.712	-0.571
-0.069	0.068	-0.822	-0.881	-0.712	1.000	0.551
0.317	0.056	-0.496	-0.651	-0.571	0.551	1.000

DIAGONAL ELEMENTS

0.047	0.111	0.935	0.353	1.388	3.165	1.000
-------	-------	-------	-------	-------	-------	-------

OFF-DIAGONAL ELEMENTS

0.000	0.043	0.118	-0.166	-0.227	-1.011	1.184
-------	-------	-------	--------	--------	--------	-------

TRI-DIAGONAL MATRIX IS :

0.047	0.043	0.000	0.000	0.000	0.000	0.000
0.043	0.111	0.118	0.000	0.000	0.000	0.000
0.000	0.118	0.935	-0.166	0.000	0.000	0.000
0.000	0.000	-0.166	0.353	-0.227	0.000	0.000
0.000	0.000	0.000	-0.227	1.388	-1.011	0.000
0.000	0.000	0.000	0.000	-1.011	3.165	1.184
0.000	0.000	0.000	0.000	0.000	1.184	1.000

Table 3.2 Showing the diagonal, off-diagonal and the tri-diagonal elements of correlation matrix of a data from Well FF13-6

3.2 Principal Component Analysis

Principal components are nothing more than the eigenvalues and the eigenvectors of a variance-covariance or correlation matrix.

If 'm' variables are measured on a collection of objects, then the variance-covariance $[S^2]$ or correlation matrix $[R]$ will be a square matrix with m rows and m columns. From either $[S^2]$ or $[R]$, m eigenvalues and m eigenvectors can be extracted. Because the variances are located along the diagonal of the variance-covariance matrix, the total variance is equivalent to finding the trace of the matrix. The sum of the eigenvalues of the matrix is equal to the trace of the matrix.

Since these eigenvalues represent the lengths of the principal semi-axes, the axes also represent the variance of the data set, and each accounts for an amount of the total variance equal to the eigenvalue divided by the trace. Usually the first two semi-axes contain most of the variance, whereas the remaining axes represent only a small amount of the variance.

By making a transformation of the form :

$$y_1 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \dots\dots\dots(3.9)$$

where x_1, x_2 are the original data values for each variable and $\alpha_1, \alpha_2, \dots, \alpha_n$ are the elements of the first eigenvector, a new data set is created which will have a variance exactly equal to the first eigenvalue. A similar transformation:

$$y_2 = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \dots\dots\dots(3.10)$$

where β 's are the elements of the second eigenvector, will create a data set with a variance equal to the second eigenvalue.

Using this transformation, the *principal component scores* are calculated, by projecting the original data set onto their principal axes. This operation in matrix form can be represented as:

$$[X][U] = [P] \dots\dots\dots(3.11)$$

where $[P]$ is the $n \times m$ matrix of principal component scores, $[U]$ is a square matrix of the eigenvectors and $[X]$ is the $n \times m$ matrix of the original observations. If all variables are expressed in the same or commensurate units, the principal components will reflect the relative importance of the different variables. Principal component analysis is sensitive to the magnitude of the measurements.

In well-logging, different units of measurements are expressed for different variables. In this case if , for example, the transit time (DT) is measured in hundred of microseconds per feet and spontaneous potential (SP) is measured in units of tens of millivolts then transit time would have exerted considerably more influence than spontaneous potential.

An obvious way around this difficulty is to standardize all variables so they have a mean of 0.0 and variance of 1.0, then the elements of the variance-covariance matrix will consist of correlation coefficients and the principal components will be in dimensionless form. Subroutine STANDARIZE (Appendix C) standardizes the original variables before calculating the eigenvalues and the eigenvectors. Standardization tends to inflate variables whose variance is small and reduce the influence of variables whose variance is large.

The technique can be illustrated using data from well FF13-6 in the Attahaddy field. Seven variables of electrical well logs are used and the scores for the first principal component compared with one of the original variables. In Figs 3.1, 3.2 and Table 3.1 the variance-covariance matrix is used, measuring the magnitudes of the original variables. The first component whose eigenvalue is 729.72(91.9%) has the most influence on the variance, whereas the last component which has an eigenvalue of 0.11 (0.01%) has negligible influence. In Figs 3.3, 3.4, 3.5 and Table 3.3 the correlation matrix is used. The first component accounts for 57% of the total variance. It is clear from Fig 3.5 how the first principal component can

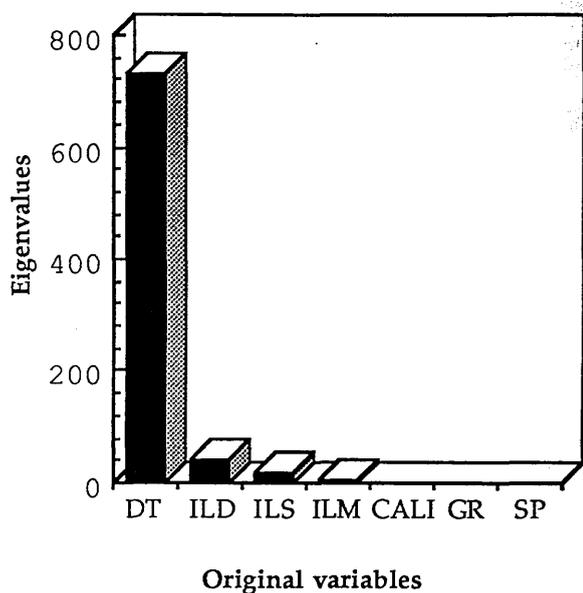


Fig.3.1a Histogram of seven variables in FF13-6

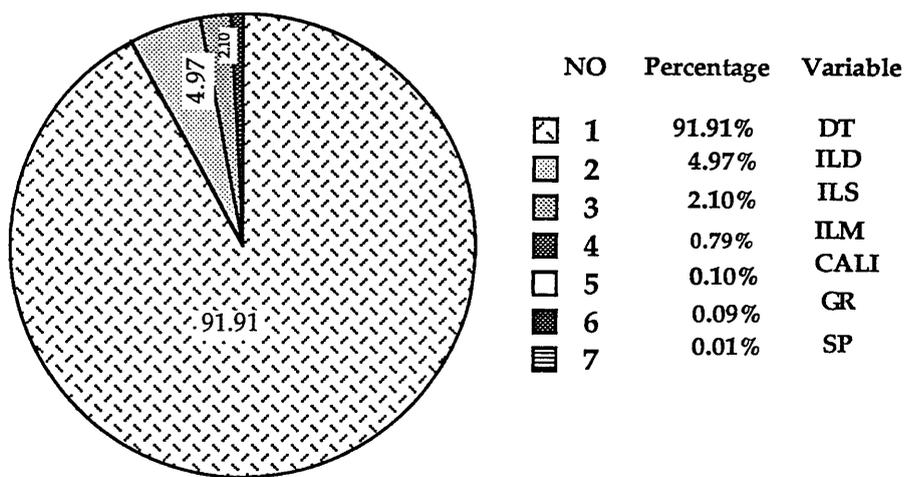


Fig. 3.1b Pie chart showing the percentage of each eigenvalue in Well FF13-6

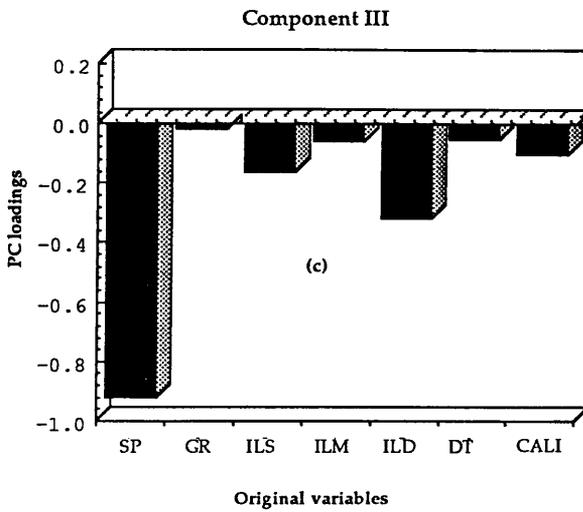
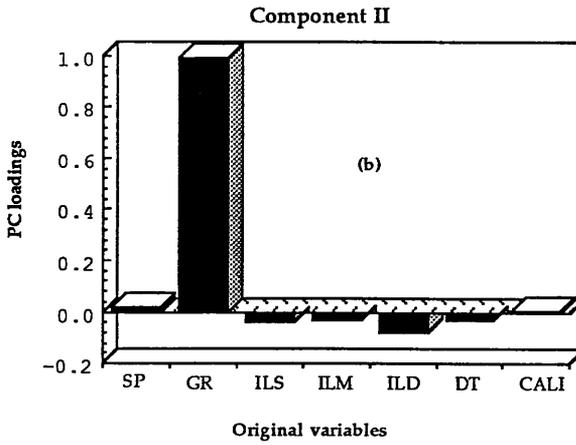
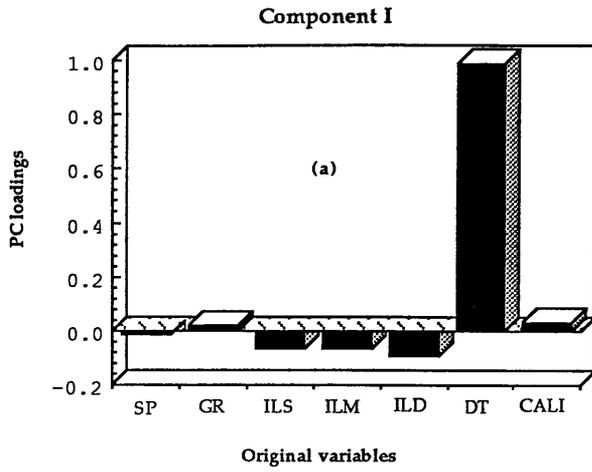


Fig. 3.2 Plot of the Principal Components in Well FF13-6.
 (a) Component I, (b) Component II, (c) Component III.

THE CORRELATION MATRIX IS :

1.000	0.027	0.300	0.104	0.295	-0.069	0.317
0.027	1.000	-0.141	-0.146	-0.166	0.068	0.056
0.300	-0.141	1.000	0.900	0.846	-0.822	-0.496
0.104	-0.146	0.900	1.000	0.919	-0.881	-0.651
0.295	-0.166	0.846	0.919	1.000	-0.712	-0.571
-0.069	0.068	-0.822	-0.881	-0.712	1.000	0.551
0.317	0.056	-0.496	-0.651	-0.571	0.551	1.000

EIGENVECTORS:

VARIABLE	1	2	3	4	5	6	7
SP	-0.08154	-0.83434	0.02123	0.32521	0.40521	0.10562	0.12514
GR	0.08654	-0.04033	0.98870	0.00334	-0.11289	-0.02421	-0.00291
ILS	-0.46505	-0.16861	0.00870	-0.17254	-0.04157	-0.83962	-0.13696
ILM	-0.49005	0.02777	0.02104	-0.09365	-0.26055	0.16626	0.80891
ILD	-0.46207	-0.13473	-0.02434	0.32766	-0.56626	0.32262	-0.48544
DT	0.44493	-0.04225	-0.10345	0.59817	-0.45683	-0.38414	0.27475
CALI	0.34348	-0.50310	-0.10101	-0.62492	-0.47339	0.06320	-0.00983

EIGENVALUES:

4.01994	1.29402	0.98749	0.37209	0.18903	0.11623	0.02116
---------	---------	---------	---------	---------	---------	---------

PERCENTAGE OF TOTAL VARIANCE CONTRIBUTED BY EACH EIGENVALUE:

57.42784	18.48607	14.10700	5.31566	2.70053	1.66048	0.3023
----------	----------	----------	---------	---------	---------	--------

Table 3.3 showing the correlation matrix, eigenvectors, eigenvalues and the percentage of each eigenvalue to the total variance in Well FF13-6. Seven variables are used, Spontaneous Potential (SP), Gamma Ray (GR), Shallow resistivity (ILS), Medium resistivity (ILM), Deep resistivity (ILD), Sonic (DT) and the caliper log (CALI).

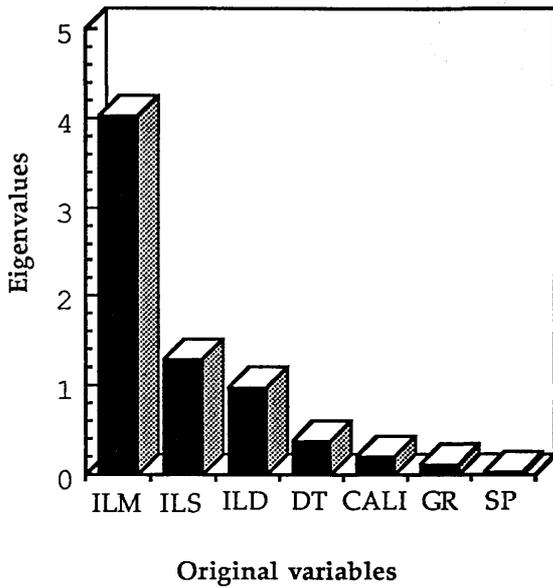


Fig. 3.3a Histogram of seven variables in FF13-6

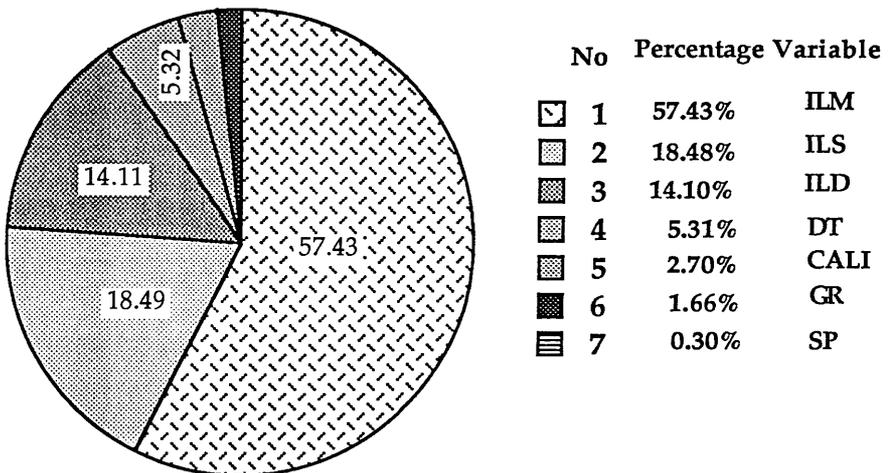


Fig. 3.3b Pie chart showing the percentage of each eigenvalue in Well FF13-6

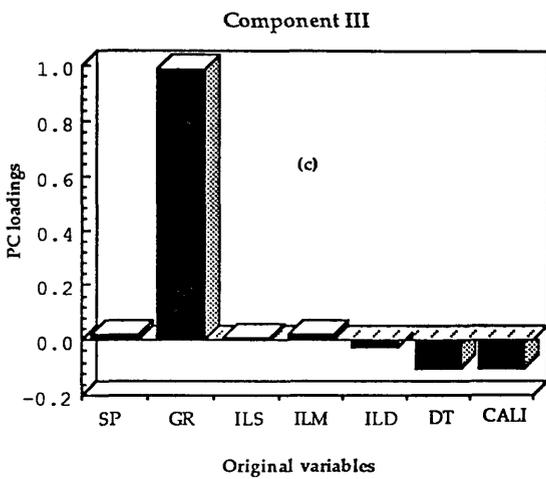
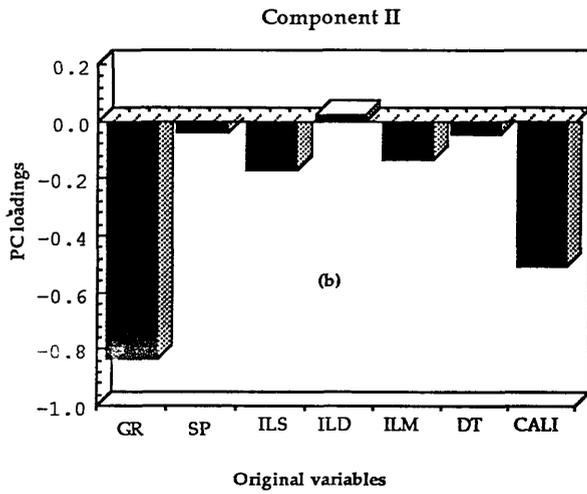
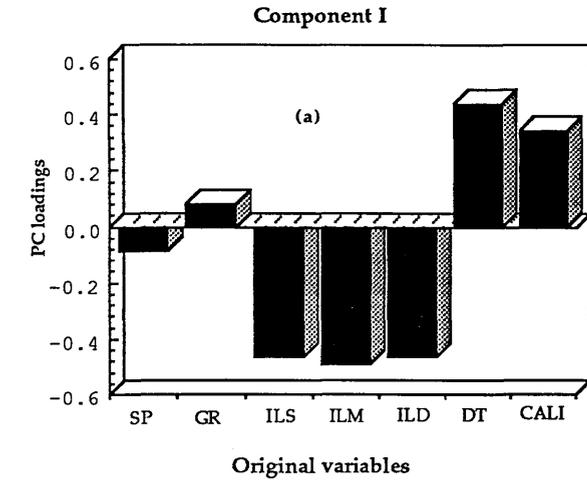


Fig. 3.4 Plot of Principal Components in Well FF13-6. (a) Component I, (b) Component II, (c) Component III.

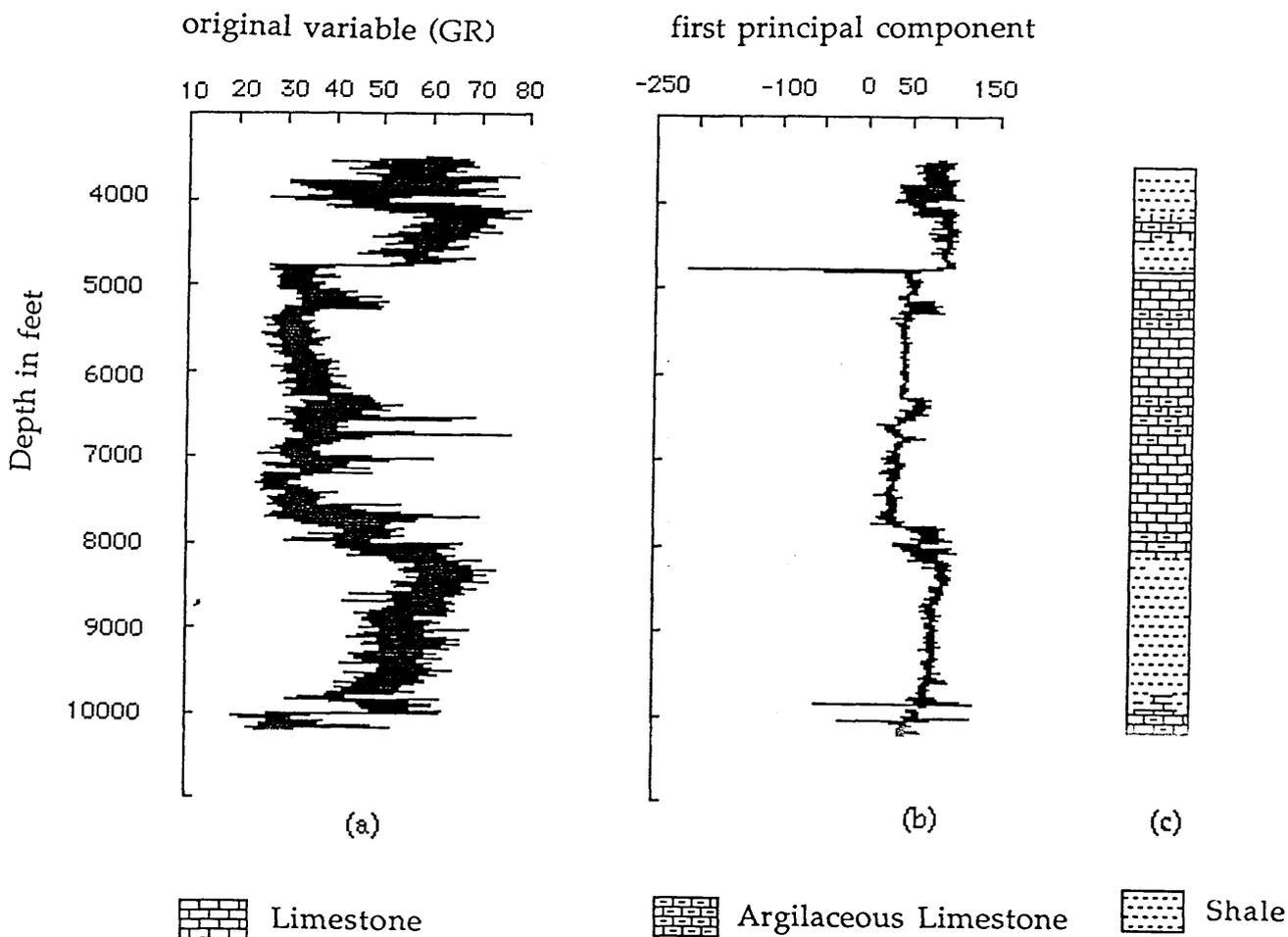


Fig. 3.5 Illustration showing the transformation of data by Principal component Analysis (PCA). (a) original variables of Gamma Ray (GR). (b) is the first principal component scores from data in Table 3.3. (c) is the lithology through well FF13-6 in the Attahaddy field.

represent the original data and reduce the noise signal in the data.

The choice between using the variance-covariance matrix or correlation matrix is dependent on the nature of the problem. For well log data, use of the correlation matrix is preferred as it allows each of the input variables equal importance and influence on the resultant principal components. With the objective of representing a complex sets of data with a single new variable, the first principal component, it is clearly incorrect to allow any one variable to have greater or lesser importance, as each of the well-log variable measures different characteristics of the lithologies in the borehole. It is the aggregate variation of all the variables that is required.

3.3 Smoothing

Field data of most well-logs are characterized by high amplitude low frequency components that are a source of difficulty in the identification of boundaries and in well to well correlation. One way to overcome this problem is to filter the data.

Perhaps the most familiar types of filters used in geology are those designed to reduce the variance in a time series. These are arbitrary filters whose general action is to smooth a data sequence; the output from the filter is a subdued approximation of the input. Most time series data consist of two components, a long-term signal or true part, and superimposed random noise. By its nature, such noise is a short-term component. As the signal tends to be the same from point to point and the noise does not, an average

of several adjacent points will tend to converge on the value of the signal alone.

The simplest smoothing filter is a 'Moving average'. A moving interval or a window is split into two parts, a portion from (i-h) on the sequence to point i and equivalent portion from point i to point (i+h) on the sequence. The window moves to successive positions, at which the average of observations within the window is calculated.

$$\bar{x}_{(i)} = \sum_{i-h}^{i+h} x_{(i)} / n \dots\dots\dots(3.12)$$

It is worthwhile noting that the shape of the resultant curve is severely dependant on the size of the window. If the length of the window is small (Fig. 3.6), the shape of the curve preserves the small edges which is useful if one wants to keep the boundaries between small layers. If, however, a big layers are of interest, big window size will truncate small edges and keep the contact between big layers (Fig. 3.7).

3.4 Boundary identification techniques

The technique used in this study is that of 'local boundary hunting' Webster (1973). An abrupt change in the average values in the sequence is an identification of a change in the properties of the rock type. This simple

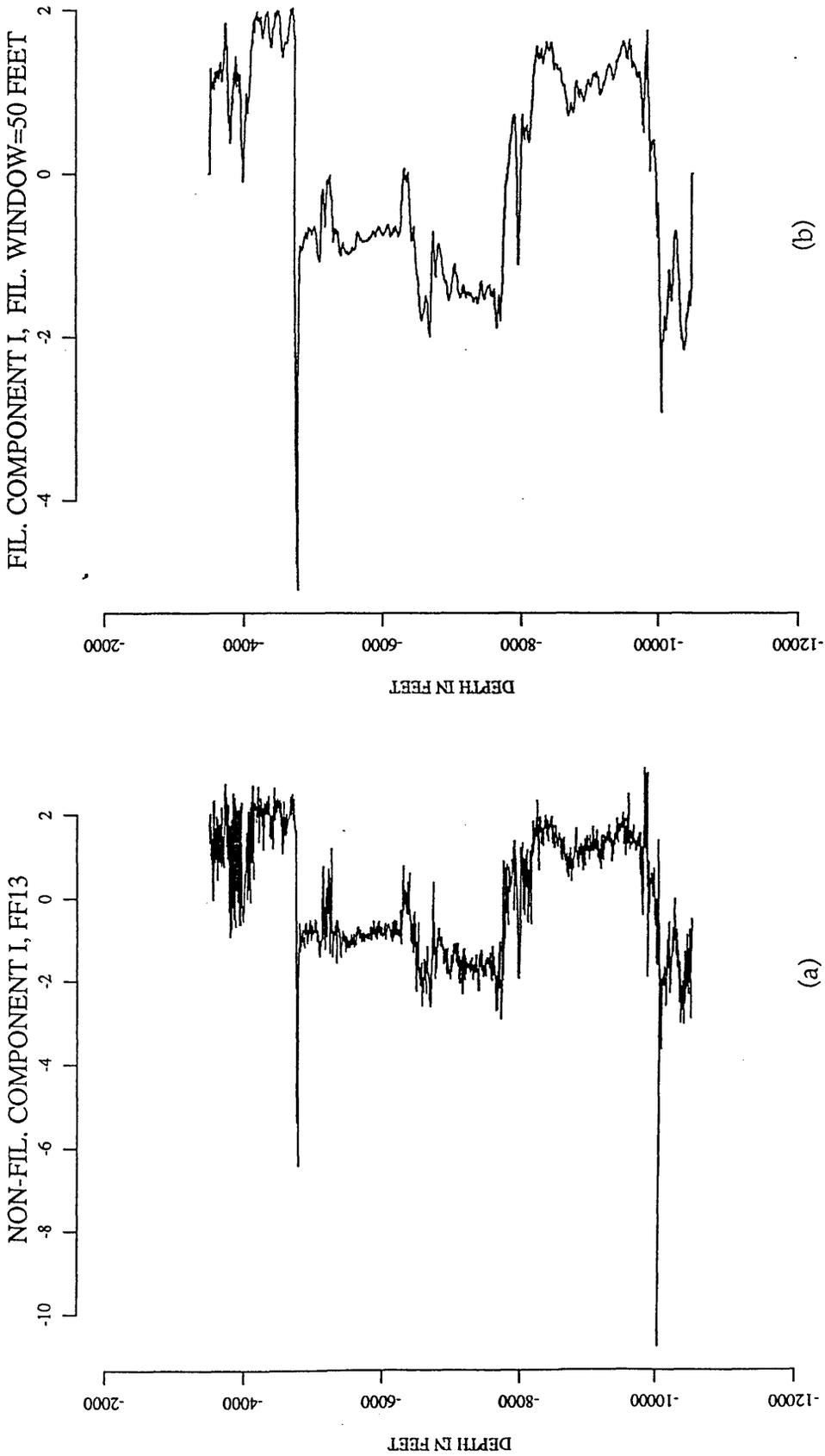


Fig. 3.6 A diagram showing the affect of window size in filtering. (a) principal component I in well FF13-6. (b) Component I filtered by a moving average filter with window size=50 feet.

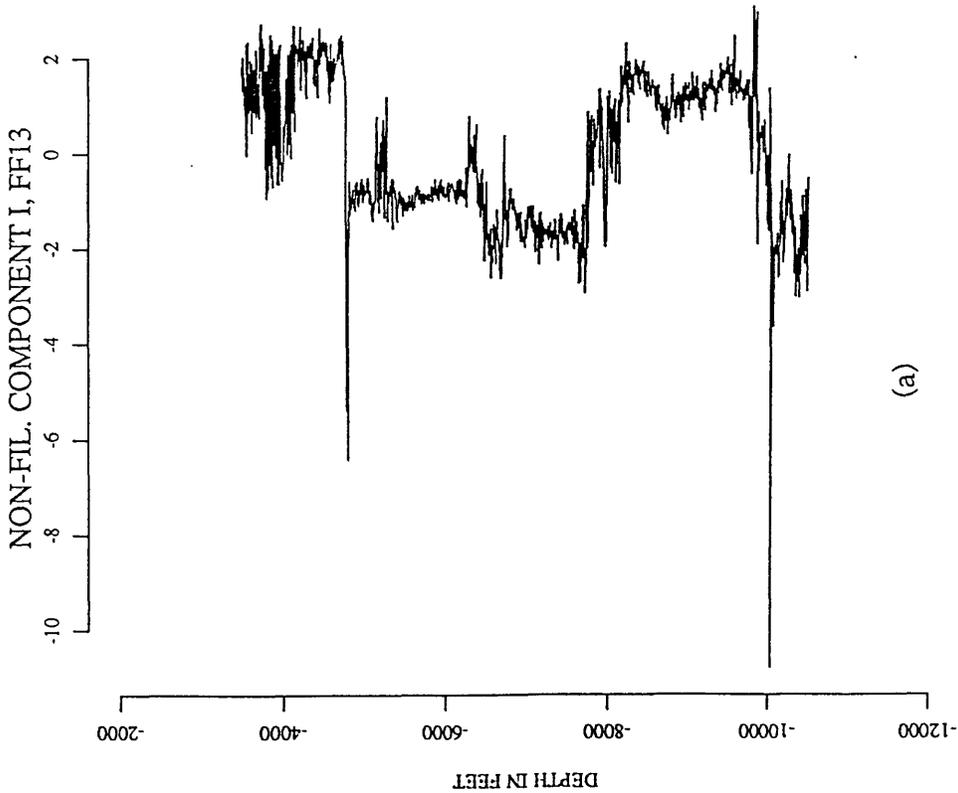
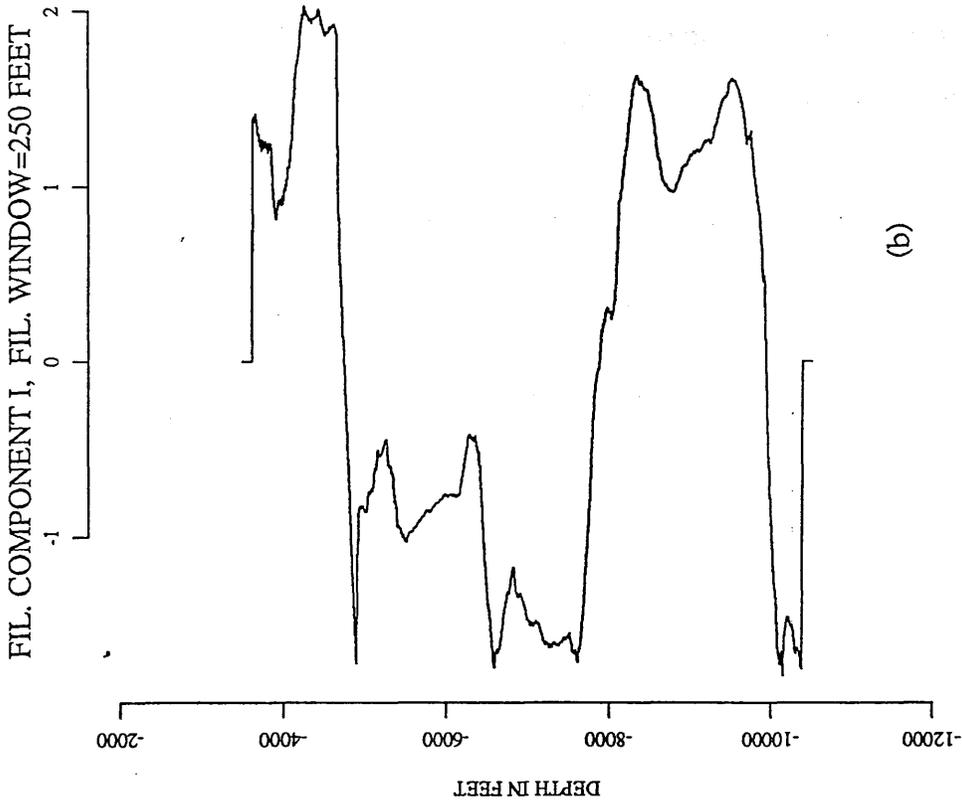


Fig. 3.7 A diagram showing the affect of window size in filtering. (a) principal component I in well FF13-6. (b) Component I filtered by a moving average filter with window size=150 feet.

approach is based on a window with two halves, a segment from point $(i+h)$ on the sequence to point i , and another segment from point i to point $(i-h)$. The generalized distance D^2 , which is the ratio formed by dividing the squared differences between the average values of the two segments by the pooled variance of the sequence in the segments, is calculated between the two halves of the window.

$$D^2 = \frac{(\bar{x}_1 - \bar{x}_2)^2}{s_1^2 + s_2^2} \dots\dots\dots(3.13)$$

where \bar{x}_1, s_1^2 are the mean and the variance of the segment from x_i to x_{i+h} , and \bar{x}_2, s_2^2 are calculated from the segment x_i to x_{i-h} .

The results of this method depend on the length of the window; a long window will miss small intervals whereas a short window creates an irregular, uninterpretable number of boundaries. A sequence of 6500 feet through borehole FF13-6 is examined to show the effect of the window size on the boundary identification technique (Figs. 3.8, 3.9, 3.10). A plot of D^2 with an inordinate number of boundaries is shown when the window size is small (100 feet)(Fig. 3.8). In contrast, using a large window length (250 feet)(Fig. 3.9) misses some important boundaries. For example the Heira Formation at depth 7742 feet has disappeared and the Domran Formation at depth 6788 has moved up to a depth 6700 feet. In Figure 3.10 a window of 150 feet is used. This window length compromises between the long and short lengths and the result is more satisfactory. The geological top of the

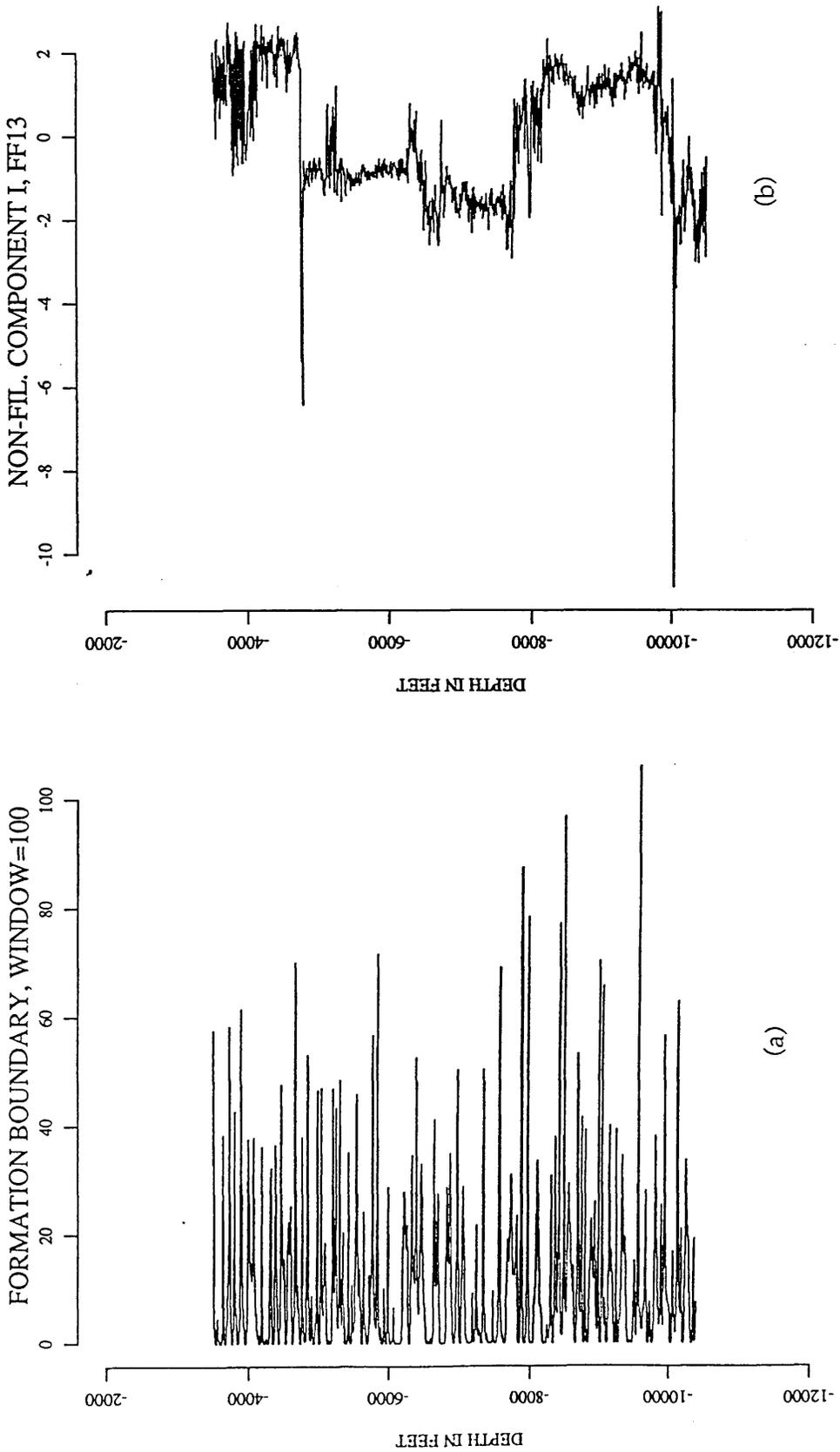


Fig. 3.8 A diagram showing (a) boundary identification technique (D^2). (a) Small window (100 feet) is used. (b) principal component 1 in well FF13-6.

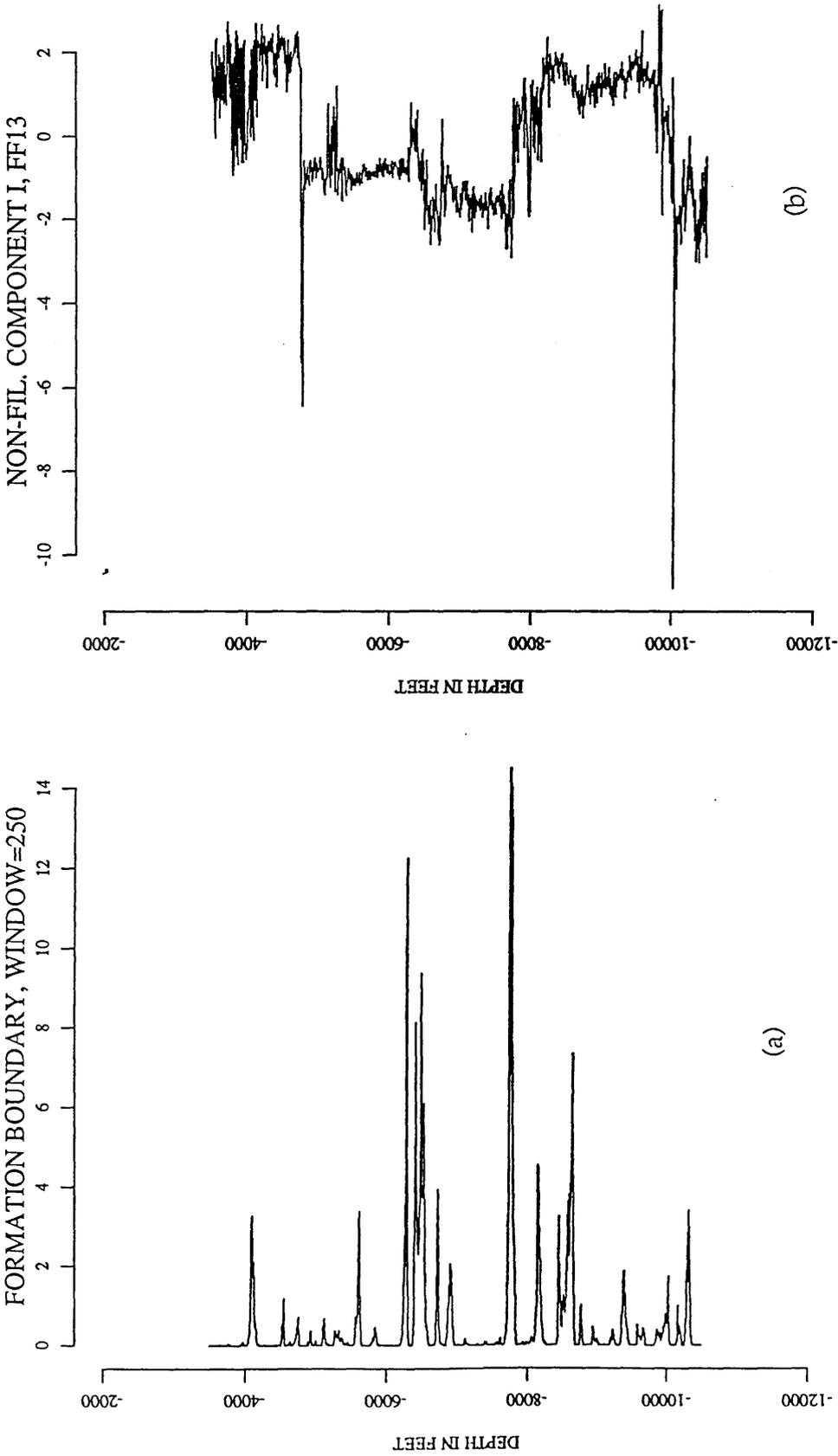


Fig. 3.9 A diagram showing (a) boundary identification technique (D^2). (a) big window size (250 feet) is used. (b) principal component I in well FF13-6.

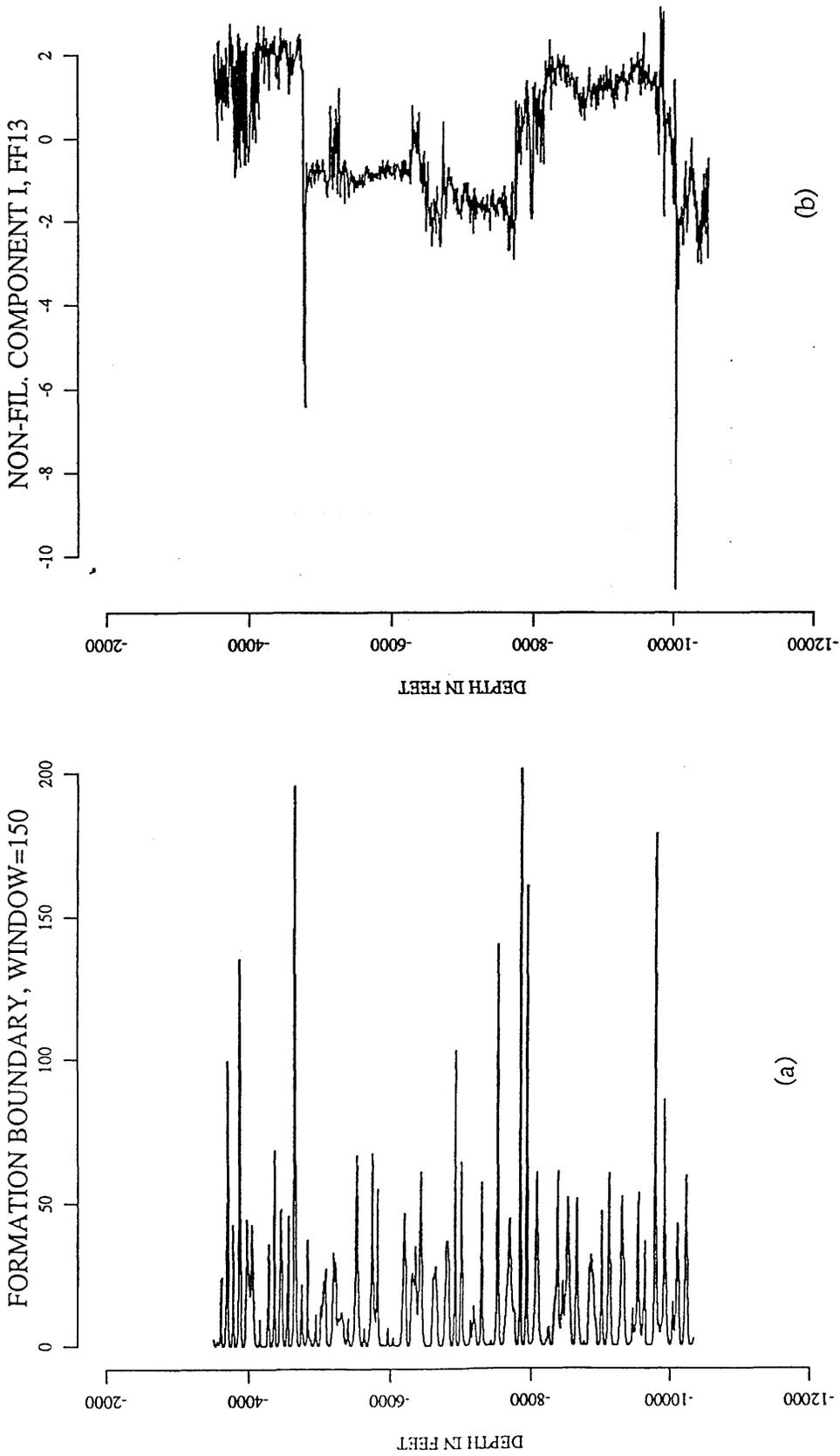


Fig. 3.10 A diagram showing (a) boundary identification technique (D^2). (a) window size (100 feet) is used. (b) principal component I in well FF13-6.

Sheghega Formation is at depth 4742 feet and the top of the formation using this approach is at depth 4720 feet. Also, the geological top of Domran Formation is at a depth of 6788 feet and program result is a depth of 6788 feet, and geological boundary of Heira Formation is 7742 feet and the boundary identification boundary is 7742 feet.

Hunting for boundaries in long sequences using this method requires a preliminary knowledge of the separation between geological boundaries. From experience, a reasonable window size can be set equal to half the average separation between the geological formations.

The alternative global zonation technique which is not applicable for long series as described in chapter one is not used to identify the boundaries in the Attahaddy Field which has long borehole sequences (about 8,000 feet).

3.5 Correlation of well-log sequences

Several techniques of time-series analysis are of importance when considering the correlation of data from pairs of wells. These include autocorrelation, cross-correlation and Fourier analysis.

For well-log data the time component is replaced by depth or distance, without any loss of the effectiveness of the methods.

Autocorrelation: is a measure of the correlation of a time series with itself at a later interval of time (depth).

$$r(v) = \frac{\text{cov}(y_i, y_{i+v})}{s_y^2} \dots\dots\dots(14)$$

Where $r(v)$ is correlation at lag or offset v , y_i is the time series data, and y_{i+v} is the time series data with lag v .

The term $\text{cov}(y_i, y_{i+v})$ is known as the autocovariance.

Cross-correlation: The cross-correlation function of two time series x and y can be defined as:

$$r_{xy}(v) = \frac{\text{cov}(x, y)}{s_x s_y} \dots\dots\dots(3.15)$$

The correlation will be large for some positive value of v if the first series x is a close copy of the second series y but lags it in time or space by v , i.e the first series is shifted to the right of the second. Likewise, the correlation will be large for some negative value of v if the first series leads the second, i.e, is shifted to the left of the second series. The relation that holds when the two series are interchanged is: $r_{xy}(v)=r_{yx}(-v)$.

Two types of cross-correlation functions are considered. If the length of the two series is the same (Fig. 3.11a). The length of the window to be compared is maximum when the time origins of two signals line up with

zero time shift, and then decreases with each time shift v . In this case the cross-correlation function is computed as above by :

$$r_{xy}(v) = \frac{\sum_{n=1}^{N-v} [x_{(n)} - \bar{x}_0] [y_{(n+v)} - \bar{y}_v]}{\sqrt{\sum_{n=1}^{N-v} [x_{(n)} - \bar{x}_0]^2 \sum_{n=1}^{N-v} [y_{(n+v)} - \bar{y}_v]^2}} \dots\dots\dots(3.16)$$

Where $v=0,1,\dots,N-1$, \bar{y}_v is the mean of the series $y_{(n)}$ at time or shift v .

If the two series have different length (Fig. 3.11b) equation (3.16) is modified to consider only a fixed window size equal to the length of the short series (Rudman, Blakely and Henderson, 1975).

$$r_{xy}(v) = \frac{\sum_{n=1}^{L1} x_{(n)} y_{(n+v)} - L1 \bar{x}_0 \bar{y}_v}{\sqrt{(\sum_{n=1}^{L1} x_{(n)}^2 - L1 \bar{x}_0^2)(\sum_{n=1}^{L1} y_{(n+v)}^2 - L1 \bar{y}_v^2)}} \dots(3.17)$$

Where:

$$\bar{x}_0 = \frac{1}{L1} \sum_{n=1}^{L1} x_{(n)} \quad \text{and} \quad \bar{y}_v = \frac{1}{L1} \sum_{n=1}^{L1} y_{(n+v)}$$

and $v=0,1,\dots,L2-L1$ and , $L1$ is the length of the short series.

In this case, the correlation function is obtained by shifting the short

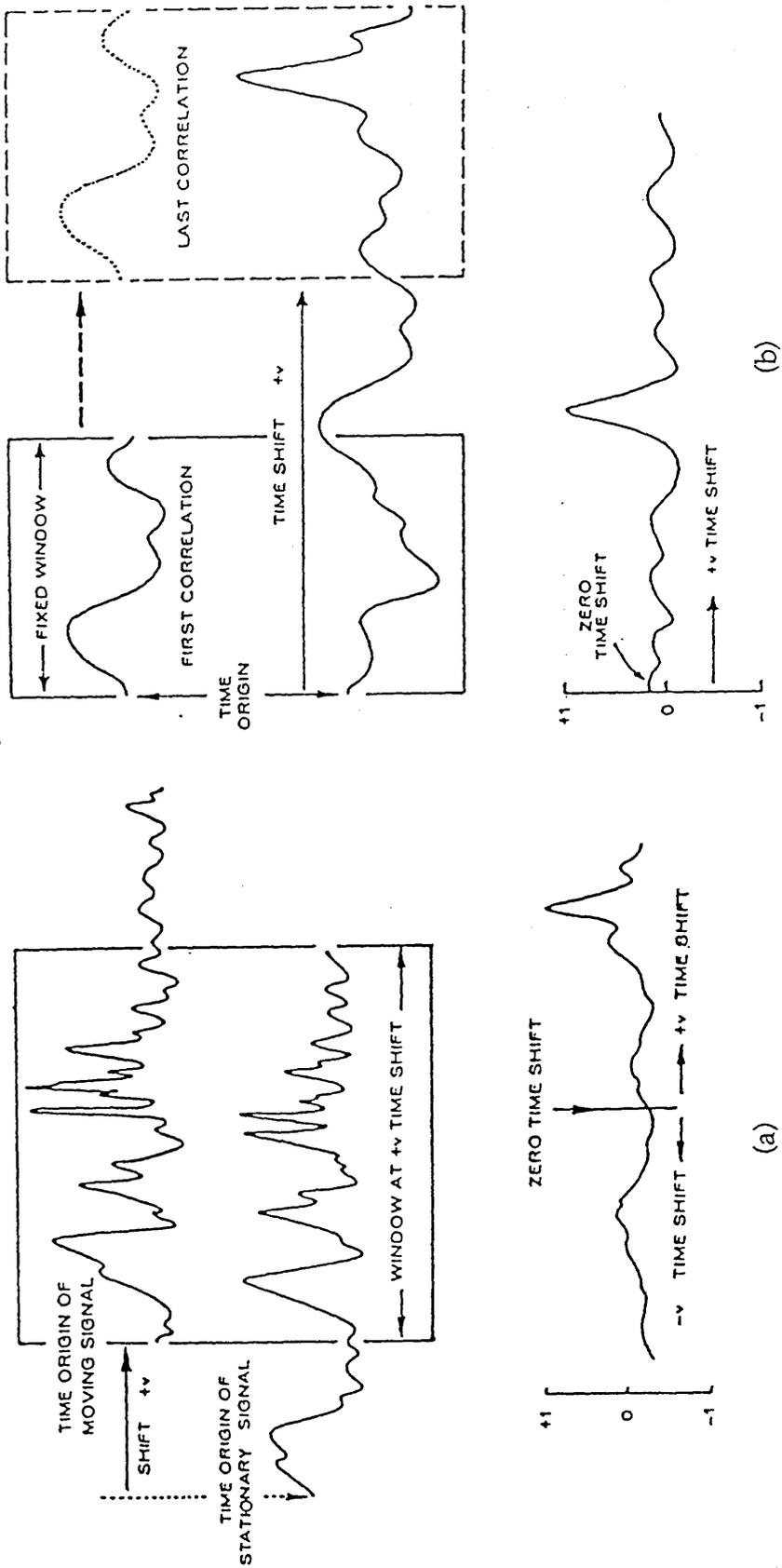


Fig. 3.11 Sketch showing the correlation process. (a) with variable window size and normalized cross-correlation function. (b) with fixed size and normalized cross-correlation function (after Kwon, 1977).

series in one direction. Both processes are used in the cross-correlation of power spectra to predict a stretch factor and cross-correlation of stretched logs to measure a relative displacement between two signals.

3.6 Discrete Fourier Transform (DFT)

The Fourier transform is one of the most powerful tools in signal processing and has long proved its effectiveness. Because this project employs Fourier transforms and its operational properties, it is instructive to review its basic theory.

Analysis process can be either described in the *time domain* by the value of some quantity x as a function of time, e.g $x(t)$, or else in the *frequency domain*, where the process is specified by giving its amplitude X (generally complex) as a function of frequency w , that is $X(w)$. The two functions $x(t)$ and $X(w)$ are two different representation for the same function. One goes back and forth between these two representation by means of Fourier transform equations:

$$X(w) = \int_{-\infty}^{\infty} x(t) e^{-iwt} dt \dots\dots\dots(3.18)$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(w) e^{iwt} dt \dots\dots\dots(3.19)$$

Where $X(w)$ is the discrete Fourier transform, $x(t)$ is the time series, i is the imaginary number $\sqrt{-1}$, w is the frequency and t is time or depth. If x is a function of position (in metres), X will be a function of inverse wavelength (cycle per metre), (Press, 1986).

In the time domain, function $x(t)$ may happen to have one or more special symmetries. It might be real or imaginary or it might be even, $x(t)=x(-t)$, or odd, $x(t)=-x(-t)$. In the frequency domain, these symmetries lead to relationship between $X(w)$ and $X(-w)$. For example, if $x(t)$ is real then $X(-w)=[X(w)]^*$ and if $x(t)$ is imaginary then $X(-w)=-[X(w)]^*$ and so forth (Press, 1988).

In general the function $X(w)$ is a complex quantity and can be expressed in terms of its real (X_R) and imaginary (X_I) parts as:

$$X(w) = X_R(w) + iX_I(w) \dots\dots\dots(3.20)$$

or in terms of amplitude and phase as :

$$X(w) = |X(w)|e^{i\theta(w)} \dots\dots\dots(3.21)$$

Where $|X(w)|$ is the amplitude of $X(w)$ and is given by

$$\sqrt{X_R^2(\omega) + X_I^2(\omega)} \dots\dots\dots(3.22)$$

and $\theta_{(\omega)}$ is the phase spectrum of the Fourier Transform and is given by :

$$\theta(\omega) = \tan^{-1} [X_I(\omega) / X_R(\omega)] \dots\dots\dots(3.23)$$

For automatic (computer) time series analysis and correlation, discrete samples of a continuously recorded signal of finite length is processed, therefore, it is necessary to adapt the analog type of Fourier transform to a discrete sequence.

The Discrete Fourier Transform (DFT) of a sequence of N samples $x(nT)$, $0 \leq n \leq N-1$ is defined as :

$$X(k\omega) = \sum_{n=0}^{N-1} x(nT)e^{-i\omega Tnk} \dots\dots\dots(3.24)$$

$$k=0,1,\dots,N-1$$

Where T is the sampling interval in time or space domain and the frequency increment ω is given as $2\pi / NT$

The time series may be re-covered exactly from the inverse Discrete Fourier Transform

$$x(nT) = \frac{1}{N} \sum_{k=0}^{N-1} X(k\omega) e^{i\omega Tnk} \dots\dots\dots(3.25)$$

$$0 \leq n \leq N-1$$

1/N is included as a scale factor.

In dealing with the Discrete Fourier Transform, there are a few operational properties that are worthwhile mentioning :

(i) Linearity :

If two series $x(nT)$ and $y(nT)$ have periods NT , then the DFT of the sum of the two series is equivalent to the sum of the transform of each series:

$$\begin{aligned} \text{DFT}[x(nT) + y(nT)] &= \text{DFT}\{x(nT)\} + \text{DFT}\{y(nT)\} \\ &= X(\omega k) + Y(\omega k) \dots\dots(3.26) \end{aligned}$$

The transform of a constant times a function is that same constant times the transform of the function :

$$\text{DFT}\{c[x(nT)]\} = c X(kw) \dots\dots\dots(3.27)$$

where c is a constant

(ii) Shift of time series :

If a periodic series $x(nT)$ has Fourier coefficient $X(kw)$, then the DFT of the shifted displaced $x((n+m)T)$ is expressed as a multiplication of $X(kw)$ and an exponential term which contributes to the phase change

$$\begin{aligned} \text{DFT}\{x(n+m)T\} &= \sum_{n=0}^{N-1} x(nT) e^{-iwT(n+m)k} \\ &= \sum_{n=0}^{N-1} [x(nT) e^{-iwTnk}] e^{-iwTmk} \\ &= X(kw) e^{-iwTmk} \dots\dots\dots(3.28) \end{aligned}$$

(iii) Lengthening of a series

Suppose we have samples $x(nT)$, $0 \leq n \leq N-1$, and we create a longer series $y(nT)$, $0 \leq n \leq rN-1$, where r is any integer number and where

$$y(nT) = \begin{cases} x(nT), & 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases} \dots\dots\dots(3.29)$$

The increased length of $y(nT)$ modifies the frequency increment w to w/r and the form of equation 3.24 modifies as follows:

$$Y(k[w/r]) = \sum_{n=0}^{rN-1} y(nT) e^{-i\omega Tnk/r} \dots\dots\dots(3.30)$$

Thus, if k is divisible by r ,

$$Y(k[w/r]) = X([k/r]w) \dots\dots\dots(3.31)$$

(iv) Cross-correlation in the frequency domain

The cross-correlation of two time series in the time domain involves iterative multiplications and summations, however it can be performed by simple multiplication of their Fourier transforms. For a long series, this process is more economic in use of computer time for correlation.

$$\text{DFT} \left[\sum_{n=0}^{N-1} x(nT) y(n+v)T \right] = X^*(kw)Y(kw) \dots(3.32)$$

where the asterisk denotes complex conjugate.

This shows that multiplying the Fourier transform of one function by complex conjugate of the Fourier transform of the other gives the Fourier transform of their correlation (correlation theorem).

(v) Power Spectra

The stretch factor between two signals can be predicted by correlation of the power spectra of the time series rather than correlating the signals themselves. The power spectra of a given series is defined as the square of its amplitude spectrum

$$P_x(kw) = |X(kw)|^2 = X^*(kw) X(kw) \dots\dots\dots(3.33)$$

From equation (3.32), this relationship is obtained by the correlation of a series with itself. Thus the power spectrum of series $x(nT)$ is also defined as the Fourier transform of its autocorrelation function. The unique feature of the power spectrum is the loss of phase information; that is , the displacement or offset between two similar sequences has been eliminated as a pertinent factor (Kwon, 1978).

3.7 Derivative filtering of data (high-pass filter)

The stretch factor between two series as discussed above can be predicted by a cross-correlation of power spectra obtained through the

discrete Fourier transform. However, samples of finite length of well-logs do not adequately represent the long period lithologic variation. Consequently, principal component scores which reflect these changes show poor resolution for the low frequency components. By taking the derivatives of the data, these components are smoothed and filtered to attenuate low frequency components and enhance higher frequency components (Fig. 3.12).

The effect of the derivative filter on the frequency spectrum can be observed by differentiating the inverse Fourier transform formula

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega \dots\dots\dots(3.34)$$

differentiating $x(t)$ gives

$$x'(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} i\omega X(\omega) e^{i\omega t} d\omega \dots\dots\dots(3.35)$$

$$FT[x'(t)] = i\omega FT[x(t)] \dots\dots\dots(3.36)$$

Taking the time derivative corresponds to high-pass filtering in the frequency domain

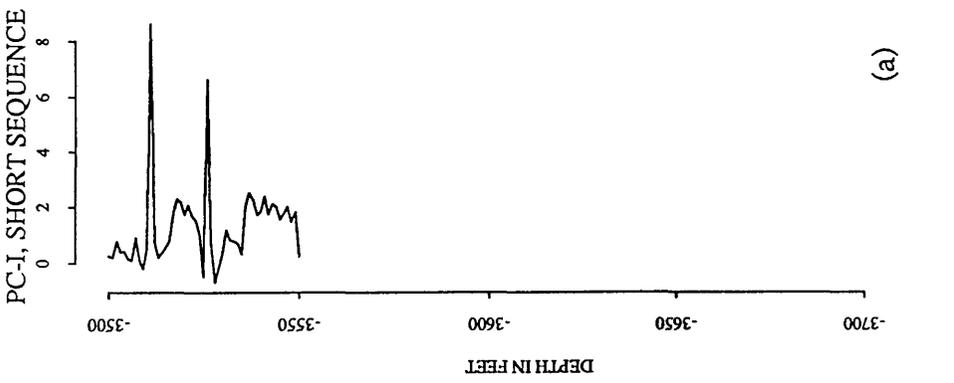
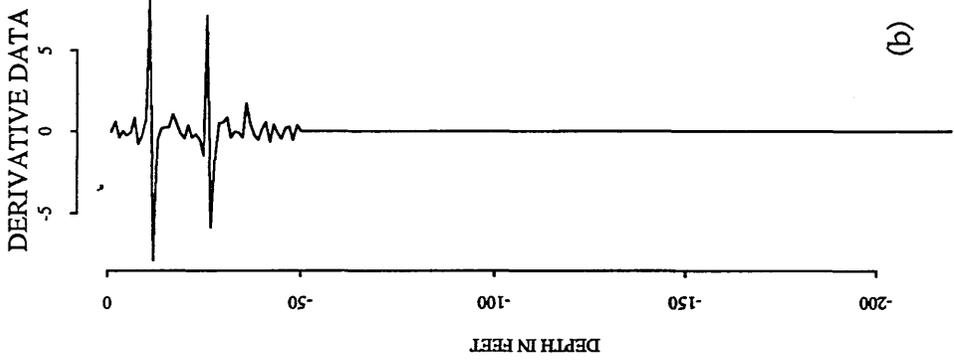
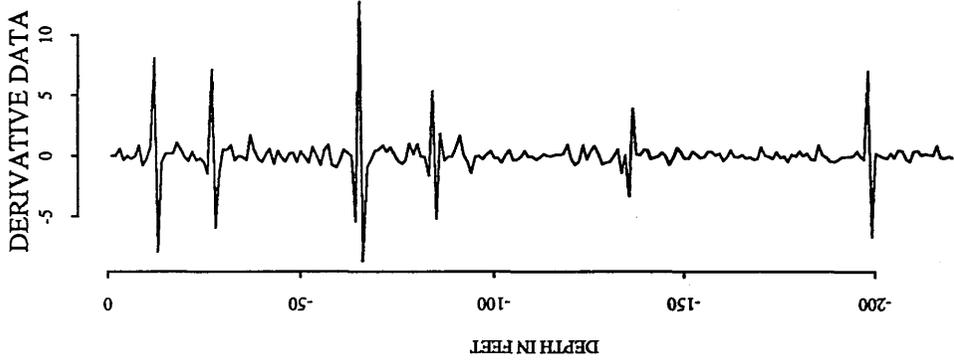


Fig. 3.12 Filtering first principal component using high-pass filter. (a) first principal component of two formations in FF13-6. (b) their derivative data.

3.8 Fast Fourier Transform (FFT)

For a time series that is continuous and sampled at discrete, equally spaced points, the continuous variance (or Power) spectrum may be calculated by either of two methods, the older procedure, calculating the continuous spectrum involves finding the Fourier transform of the autocorrelation of a time series. Developed by Bartlett (1948), this approach achieves the same result as the fast Fourier transform method and is still widely used. However, this method is not applicable when the series is extremely long. A somewhat newer and more widely used approach involves calculating many values of the line spectrum by the fast Fourier transform (FFT). The fast Fourier transform (FFT) is a computer algorithm (FOURT subroutine, appendix C) first introduced by Colley and Turkey (1965) to calculate the discrete Fourier transform faster, as its name implies, than any other available algorithm. The Fourier series in the FFT is the same as in the DFT, and it requires the Fourier relationship to be expressed in complex form

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-2\pi i k n / N} \dots\dots\dots(3.37)$$

where i is the imaginary number $\sqrt{-1}$ And the original time series is

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{+2\pi i k n / N} \dots\dots\dots(3.38)$$

$$n = 0, 1, \dots, N-1$$

Both T (time increment) and w (frequency increment) are omitted because they serve as scale factors which become important only when plotting the results.

The only differences between (3.37) and (3.38) are (i) changing the sign in the exponential, and (ii) dividing (3.38) by N . This means that a routine for calculating Fourier transform can also, with slight modification, calculate the inverse transform.

3.9 Prediction of stretching and displacement with Power Spectra

Consider a time series $x(n)$ of N samples as the short signal and a long series $y(n)$ of L samples (Fig. 3.13). A part of the long series is called $b(n)$ and is equivalent to the short series $x(n)$ with stretch factor ($S=M/N$) and displacement D . The long series $y(n)$ can be represented in the sum of two series $q(n)$, which represent the lengthened series $b(n)$, and noise series $h(n)$. The series $\bar{x}(n)$ with length L is used instead of $x(n)$ for computational reasons. The relationship between the two DFT's $Q(k)$ and $B(k)$ is complicated by the additional zeros in $b(n)$. These effectively change the phase and modify the frequency scaling. However, phase change problem can be avoided by computing the Power Spectra $P_B(k)$ and $P_Q(k)$ from DFT's $Q(k)$ and $B(k)$ respectively

$$P_Q(k) = P_B(k / S') \dots\dots\dots(3.39)$$

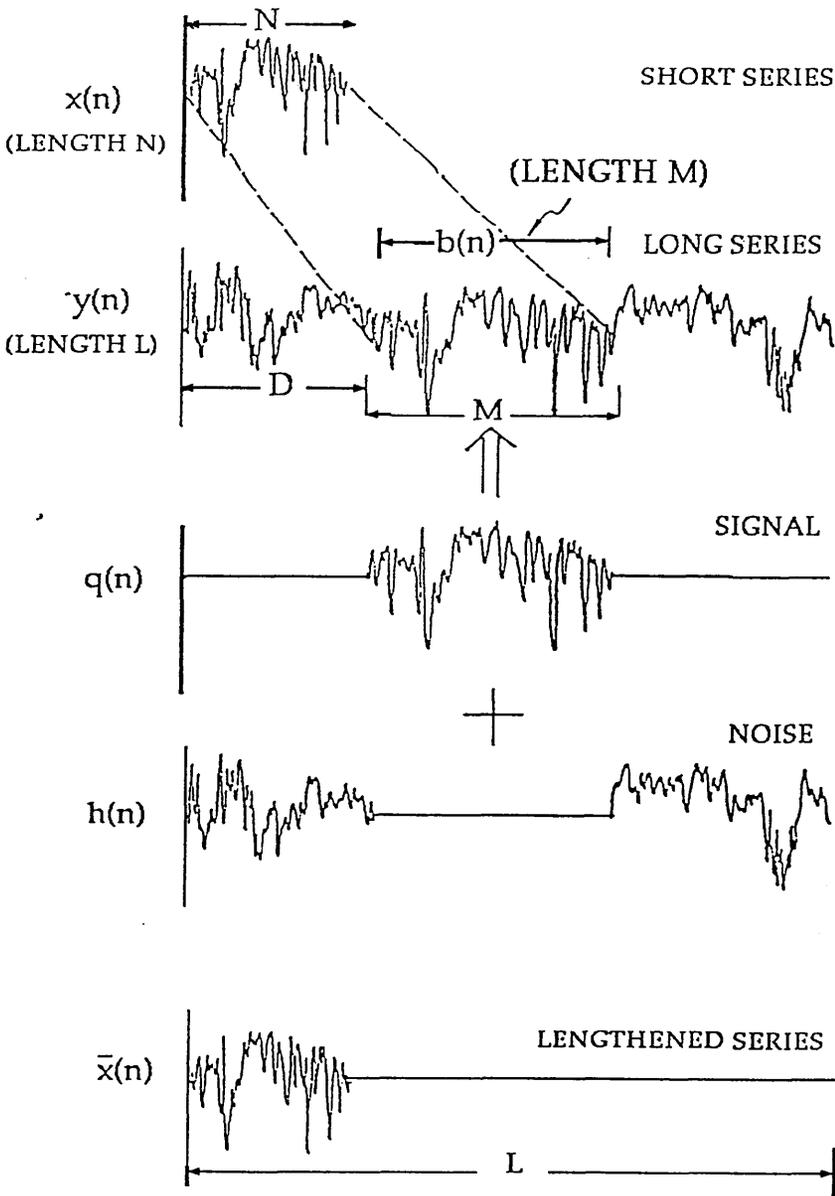


Fig. 3.13 Model data used for cross-correlation of a short series $x(n)$ with a long series $y(n)$. The longer series $y(n)$ is composed of two signals, signal $b(n)$ and noise $h(n)$. Signal $q(n)$ is equivalent to the short series $x(n)$ with a stretch factor $S(=M/N)$ and displacement D . Lengthened series $\bar{x}(n)$ is required for correlation process (modified after Kwon, 1977).

Where S' is a scaling factor = L/M , L is the length of the long series and M is the length of the stretched series.

A similar relationship is derived between $p_{\bar{x}}$ and P_X

$$P_{\bar{x}}(k) = P_x(k / S'') \dots\dots\dots(3.40)$$

Where S'' is a scaling factor= L/N and N is the length of the short series.

From the factor S' , the length of $b^{(n)}$ can be calculated and consequently the stretch factor S between two series.

3.10 Logarithmic Scaling of frequencies

The unknown factor D (Displacement in time domain) is ignored when computing the power spectra. Another problem is the scaling in the frequency domain. The multiplication factor S' and S'' in equation (3.39) and (3.40) will be converted to additive factors if the power spectra are transformed to a logarithmic scale

$$P_Q(\log k) = P_B(\log k - \log S') \dots\dots\dots(3.41)$$

$$P_X(\log k) = P_X(\log k - \log S'') \dots\dots\dots(3.42)$$

Logarithmic scaling of frequencies modifies the Power Spectra by a frequency delay of $\log S'$ or $\log S''$ (Kwon, 1977). By cross-correlation between $P_X(\log k)$ and $P_Q(\log k)$ the factor S' and S'' can be obtained to detect such lag (delay) values.

Unfortunately, the values of power spectra after transforming to logarithmic scale are not at evenly spaced intervals. Before cross-correlation takes place these should be at equal intervals for computer correlation. Kwon (1977) used Lagrange's interpolation method and in this investigation the same method is applied for interpolation to obtain evenly spaced spectra $P_X'(i)$ and $P_Y'(i)$.

The power spectra of the long series $y(n)$ is composed of signal $q(n)$ and noise $h(n)$. The Fourier transform of this series is

$$Y(k) = Q(k) + H(k) \dots\dots\dots(3.43)$$

and the power spectra

$$P_Y(k) = P_Q(k) + P_H(k) \dots\dots\dots(3.44)$$

Assuming zero lag when series are aligned and no noise in the signal, the cross-correlation function between the two interpolated spectra is written in two separate equations as follows :

$$r_{P'_X P'_Y}(-v) \cong \sum_{i=1}^{N-1} P'_X(i+v)P'_Y(i) \dots\dots\dots(3.45)$$

$$r_{P'_X P'_Y}(+v) \cong \sum_{i=1}^{N-1} P'_X(i)P'_Y(i+v) \dots\dots\dots(3.46)$$

where v is a positive integer and i is a dummy variable for the interpolated spectrum. Correlation coefficient $r_{P'_X P'_Y}$ with negative values of v are obtained when $P'_X(i)$ is shifted to the left against a stationary series $P'_Y(i)$ and long the series is assumed to be stretched ($M > N$). The maximum coefficient can be found if :

$$v = \frac{1}{i} \log S'' - \frac{1}{i} \log S' = \frac{1}{i} \log \left(\frac{S''}{S'} \right) = \frac{1}{i} \log \left(\frac{M}{N} \right) \dots\dots(3.47)$$

On the other hand, correlation coefficient $r_{P'_X P'_Y}$ with positive values of v are obtained when $P'_X(i)$ is shifted to the right against a stationary signal $P'_Y(i)$ and the short series is assumed to stretched ($N > M$).

The maximum coefficient can be found if :

$$v = \frac{1}{I} \log \left(\frac{N}{M} \right) \dots\dots\dots(3.48)$$

Using equation 3.39 and 3.40 to transform equation 3.45 to logarithms:

$$r_{p'_x p'_y}(-v) = \sum_{i=1}^{N-v} p'_x \left(i - \frac{1}{I} \log S \right) p'_y \left(i - \frac{1}{I} \log S \right)$$

$$= \sum_{i=1}^{N-v} p'_x \left(i - \frac{1}{I} \log S + v \right) p'_y \left(i - \frac{1}{I} \log S \right) \dots(3.49)$$

where I is the interpolation interval.

Once the shift v is known for the maximum correlation coefficient, the stretch factor S, either M/N or N/M can be deduced from

$$S = 10^{vI} \dots\dots\dots(3.50)$$

3.11 Interpolation (stretching) by inverse FFT

By modifying the DFT of the signal which is obtained using the FFT algorithm, a simple and accurate stretching of time series can be achieved in the frequency domain (Rudman, Blakely and Henderson, 1975). A band limited series with N samples (no frequency component above the Nyquist frequency) can be stretched to M points, $M > N$ by inserting $(M-N)$ zeros in the middle of the DFT values. Because no new frequencies were added above the Nyquist, the inverse transform gives the time series of length M . For example, consider a time series of 8 points (Fig. 3.14). The heavy line indicates the input signal and dashed line is a remainder that the DFT is computed assuming that the signal is cyclically repeated in both directions. The DFT for this signal is also given by 8 frequency components and shows even symmetry about the Nyquist frequency. If we added 8 zeros in the middle, the total number of frequency components increase to 16 points. Twice as many samples of the original signal will be recovered when the signal is transformed back to the time domain. This procedure interpolates 8 new data points into the original time signal which stretches the original series by a factor of 2.

3.12 Determination of Displacement

Given the stretch S between two series, the series can be stretched or resampled using the interpolation method. The displacement D between the short series and the identical part of the long series will be determined from the maximum value of the correlation function computed.

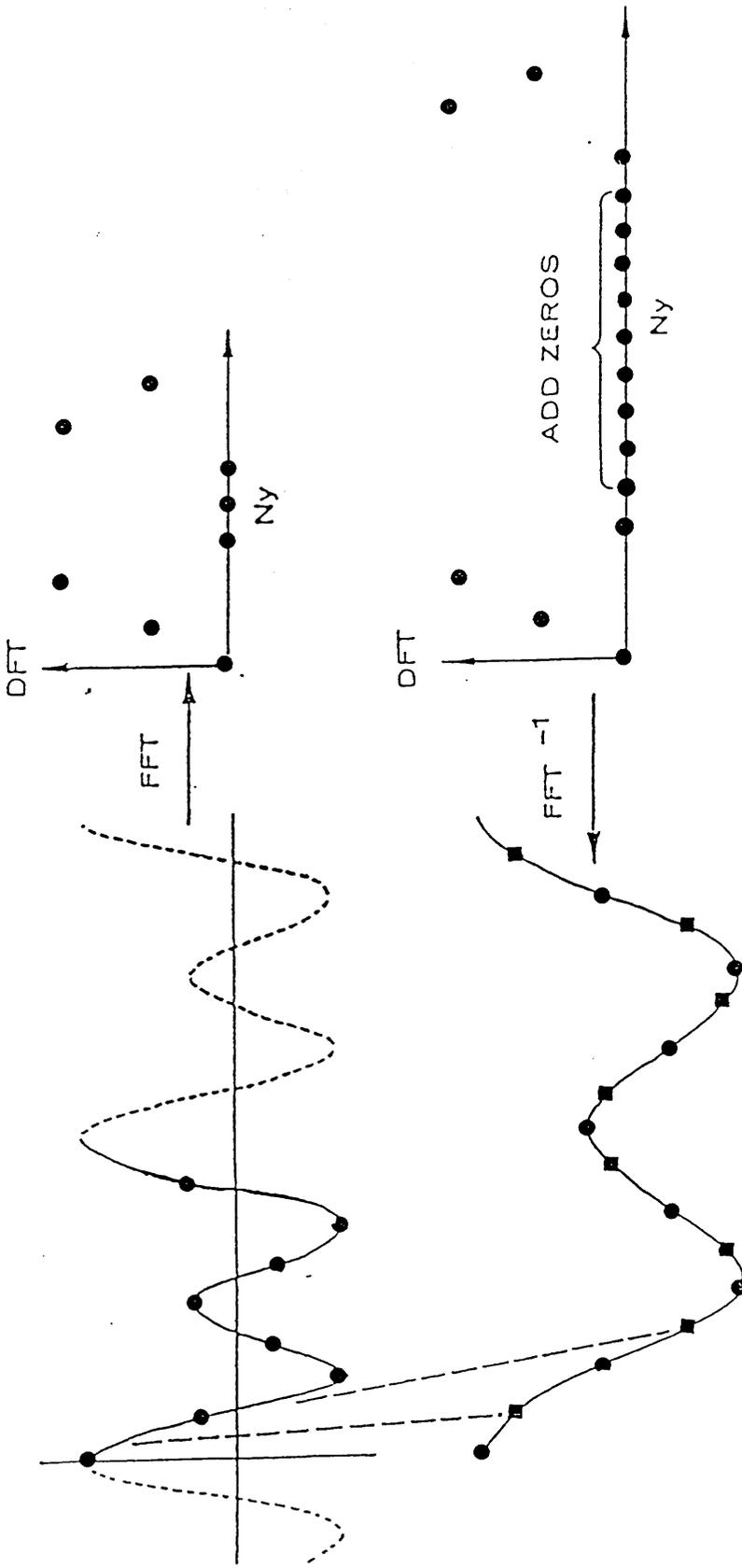


Fig. 3.14 Interpolation (stretching technique using FFT (fast Fourier transform)). A signal of eight samples (heavy lines) is stretched to a signal of 16 samples by inserting eight zeros in the frequency domain. ● is the original point and ■ is the interpolated point. The Nyquist frequency (Ny) is identified on discrete Fourier transform plot (modified after Kwon, 1977).

3.13 Summary

Applying Principal Component Analysis (PCA) and correlation technique to well-log data involves the following steps :

1. Calculate the variance-covariance or correlation matrix of the original variables.
2. From variance-covariance or correlation matrix compute the eigenvalues and the eigenvectors.
3. Calculate principal component scores.
4. Extract and Smooth (optionally) the principal component scores.
5. Identify formation boundaries.
6. Take Fourier transform of the principal component scores and compute power spectra.
7. Transform frequencies to logarithmic scale.
8. Obtain equally spaced (interpolated) power spectra using Lagrange's interpolation method.
9. Cross-correlate between interpolated power spectra.
10. Either, stretch short series and cross-correlate to find the maximum correlation coefficient,
OR stretch long series and cross-correlate to find the maximum correlation coefficient.
11. Find the largest coefficient and determine the optimum stretch.

12. Stretch the shorter series and perform cross-correlation to determine the relative displacement of the two series.

Application to Atmospheric Data

CHAPTER FOUR

Application to Attahaddy Field

4.1 introduction

A Fortran program, PCAXCOR, has been written for applying principal component analysis, boundary identification and correlation of well-log data. The data were obtained from magnetic tapes for different wells in the Attahaddy field (Fig. 4.1). Using model data, constructed from one of the well-logs, the method is first tested using both the original variables of electrical logs and their first principal components. Then complete sequences, sampled at 5 feet intervals, from a number of wells are processed.

The first part of the program is applied to calculate the eigenvalues and the eigenvectors of the correlation matrix of the original variables. These variables are Gamma Ray (GR), Spontaneous potential (SP), Shallow resistivity (ILS), Medium resistivity (ILM), Deep resistivity (ILD), Transit time (DT) and the Calliper (CALI) (Appendix B). These variables are often used in manual Formation boundary identification and well-to-well correlation, and hence are chosen for automated boundary identification.

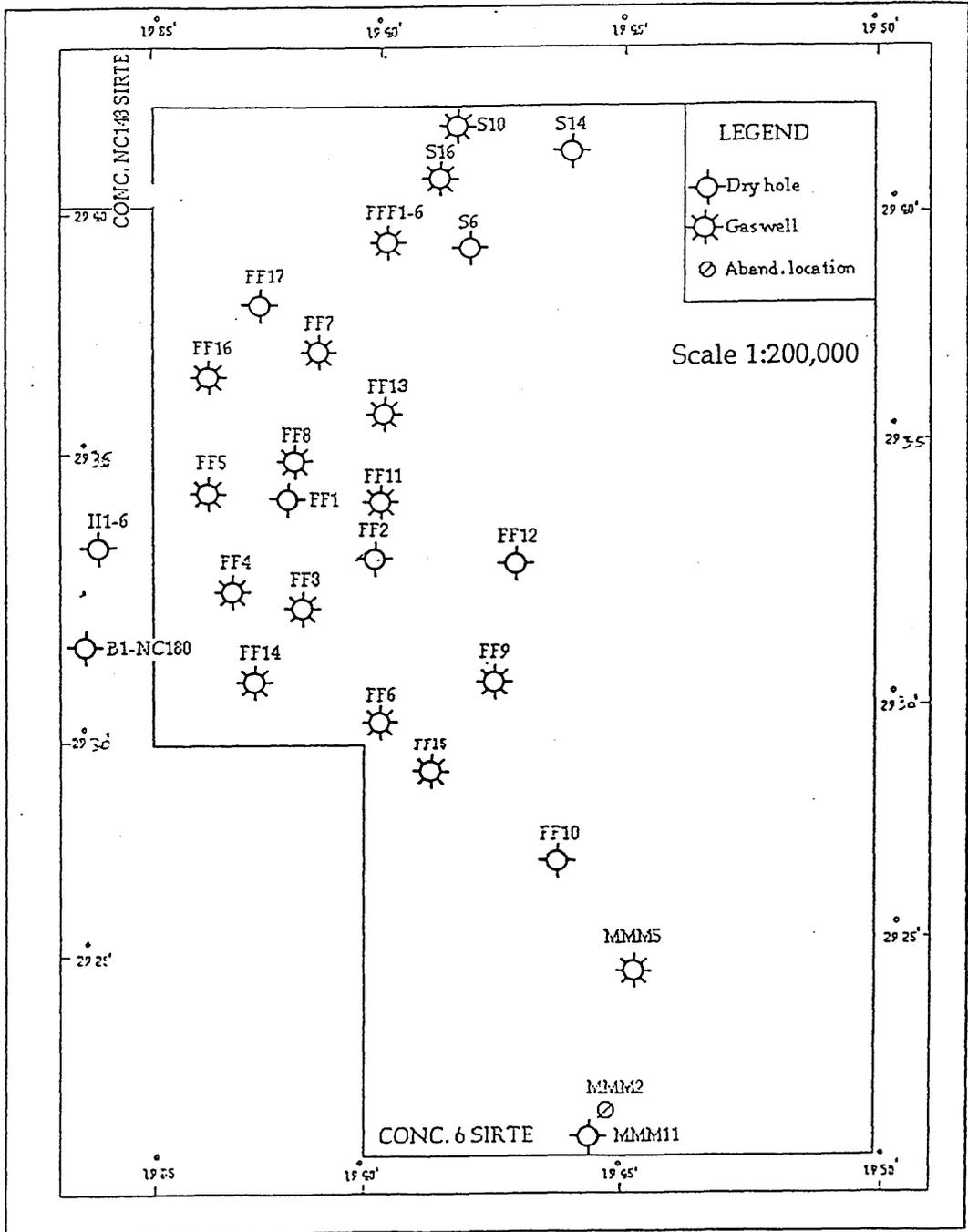


Fig. 4.1 Location of the boreholes in the Attahaddy field.

and cross-correlation. The correlation matrix is used because the variables are expressed in different units of measurements and it is necessary to calculate the different principal components. Using these components, the different boundaries for different Formation in different boreholes are identified. Different well-log data (Neutron, density, ..., etc) and the variance-covariance matrix can be used if all variables are expressed in the same units.

The second part of the program is the determination of the cross-correlation between different formations to examine the effectiveness of using the spectral analysis of the first principal components of real data. In contrast to the spectral analysis method for correlating well-logs developed by Kwon (1977) which applies the spectral analysis to the original data, this new approach makes use of the the first principal component of well-log data as the basis for correlating different boreholes. An empirical measure of successful correlation is determined by comparing the results to the known geological correlations in the field.

Two cross-correlation functions are used. The first uses, cross-correlation of power spectra of the first principal component to determine the stretch factor (S), which in turn represents the variation in the thickness between two formations, and the second uses, cross-correlation of the stretched series to determine the relative displacement (D in Fig. 3.13). This is preceded by the stretching process using the FFT algorithm. Kwon (1977) notes that when applying the stretch method, the derivative filtering (high

pass filtering) is necessary for most real data to obtain appropriate high resolution for the cross-correlation of power spectra. In other words, derivative data is used to obtain the stretch factor, but the original data (principal components) to determine the displacement. He also points out that when dealing with well-logs complicated by the presence of high frequency signals, it is desirable to apply a smoothing filter, such as a moving average, to the data before differentiation in order to obtain more reliable correlation of power spectra.

In this study the filtered components (using moving average filter, SMOOTH subroutine Appendix C) are used before preceding with the cross-correlation. Difficulties encountered in PCAXCOR are also common in other mathematical correlation methods. Therefore several conclusions drawn from the study of the real data are also applicable to the other methods.

4.2 Analysis of model data

4.2.1 Using original variables

A 300 foot sequence of Gamma Ray log (GR), was digitised from borehole FF13-6 in Attahaddy field [depth 4000-4500 feet], and correlated with a short series of GR log (50 feet) from the same borehole and shifted down 50 feet (Fig. 4.2a). The derivatives (high frequency components) of these signals are taken to eliminate the noise affect (Fig. 4.2b). Power spectra of these logs are then calculated using the fast Fourier transform and the

components above the Nyquist are ignored (Fig. 4.2c). The length of the short series in (Fig. 4.2c) has the same number of points as the long series (i.e. \bar{x} in Fig. 3.13). If the spectra are transformed to logarithmic frequencies (Fig. 4.2d), then the transformation converts the scaling effect of Fig. 4.2c into a shift between these spectra. Lagrange's interpolation method then is used to obtain equi-spaced spectra (Fig. 4.2e). In general the ratio of thickening to thinning is not very large and rarely exceeds a value of 2, therefore, the maximum shift is set to terminate at 30 which is derived from equation 3.50 with sampling interval of 0.01 ($I=0.01$). In the case of model data the maximum correlation coefficient (1.0) is observed at 0 (no stretching) (Fig. 4.2f). The cross-correlation of stretched series is then performed to obtain the displacement between the two logs. Figure 4.2g shows the maximum correlation coefficient (1.0) at displacement of 51 feet.

4.2.2 Using Principal Component Scores

A short series (50 feet) of the first principal component obtained from the same borehole (FF13) is compared with a long series (300 feet) with displacement of 50 feet (Fig. 4.3a). The derivatives of the data of the principal component scores are taken (Fig.4.3b) and power spectra are then calculated (Fig. 4.3c). This plot shows that there is a similarity in shape between the two curves, but a prominent scaling effect of frequencies is observed. The frequencies are then transformed to logarithmic scale (Fig. 4.3d) to obtain the shift between the two signals. On a logarithmic scale (base 10), the number of known components in each logarithmic cycle is different, e.g 10 in the first cycle, 90 in the second and 900 in the third. Using the Lagrange interpolation method, 100 samples are interpolated with sampling interval $I=0.01$. The

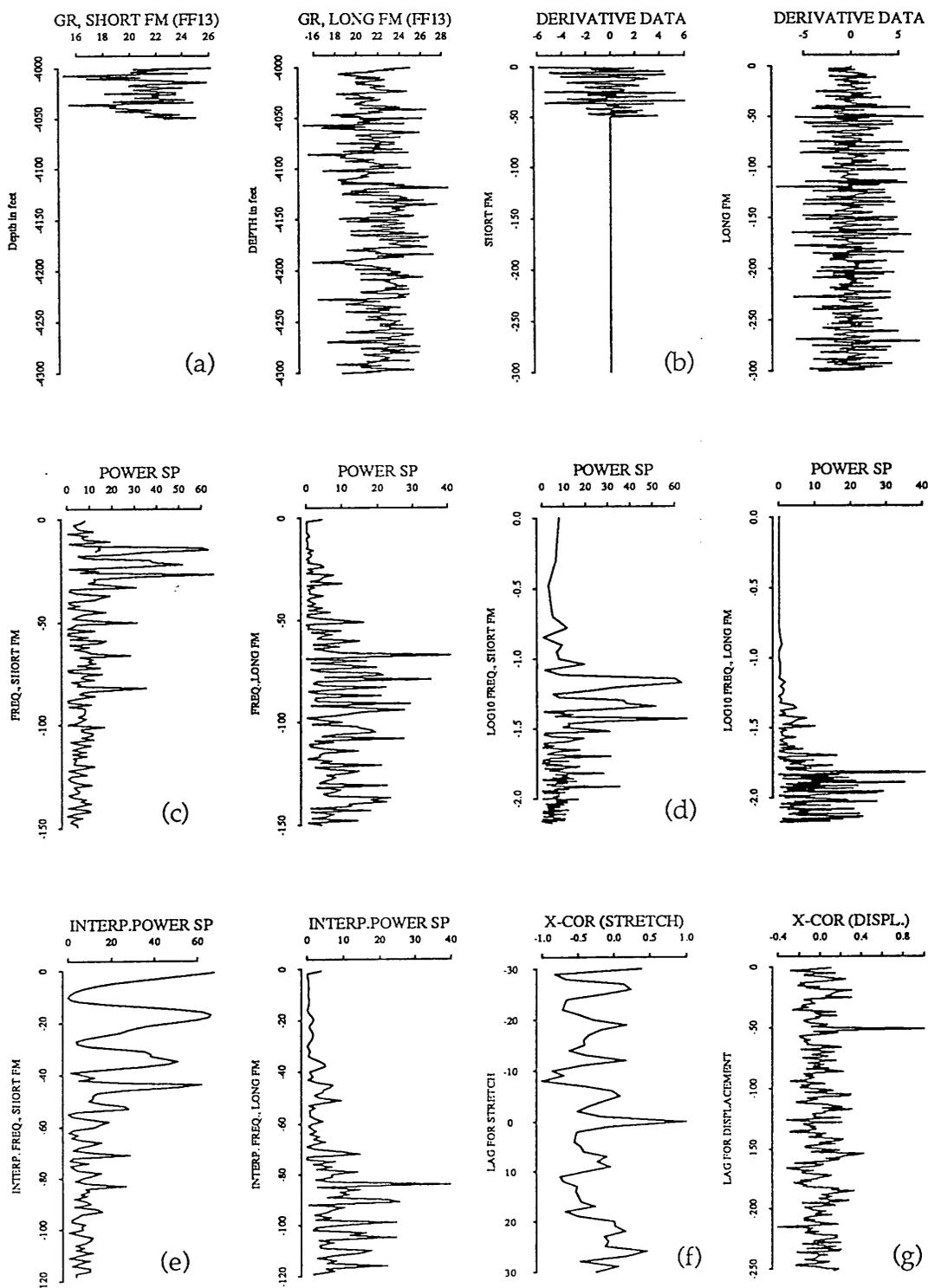


Fig. 4.2 Output plot using model data. (a) Gamma Ray logs of a short log (50 feet) and long log (300 feet). The short log is a part of the long log shifted down 50 feet. (b) Derivative data of both series. (c) is power spectra, and (d) is the power spectra with logarithmically spaced frequencies. (e) is the interpolated power spectra. (f) is the cross-correlation coefficient of power spectra (high peak at 0 which implies no stretching $S=1.0$) and (g) is the cross-correlation coefficient for displacement ($D=51$).

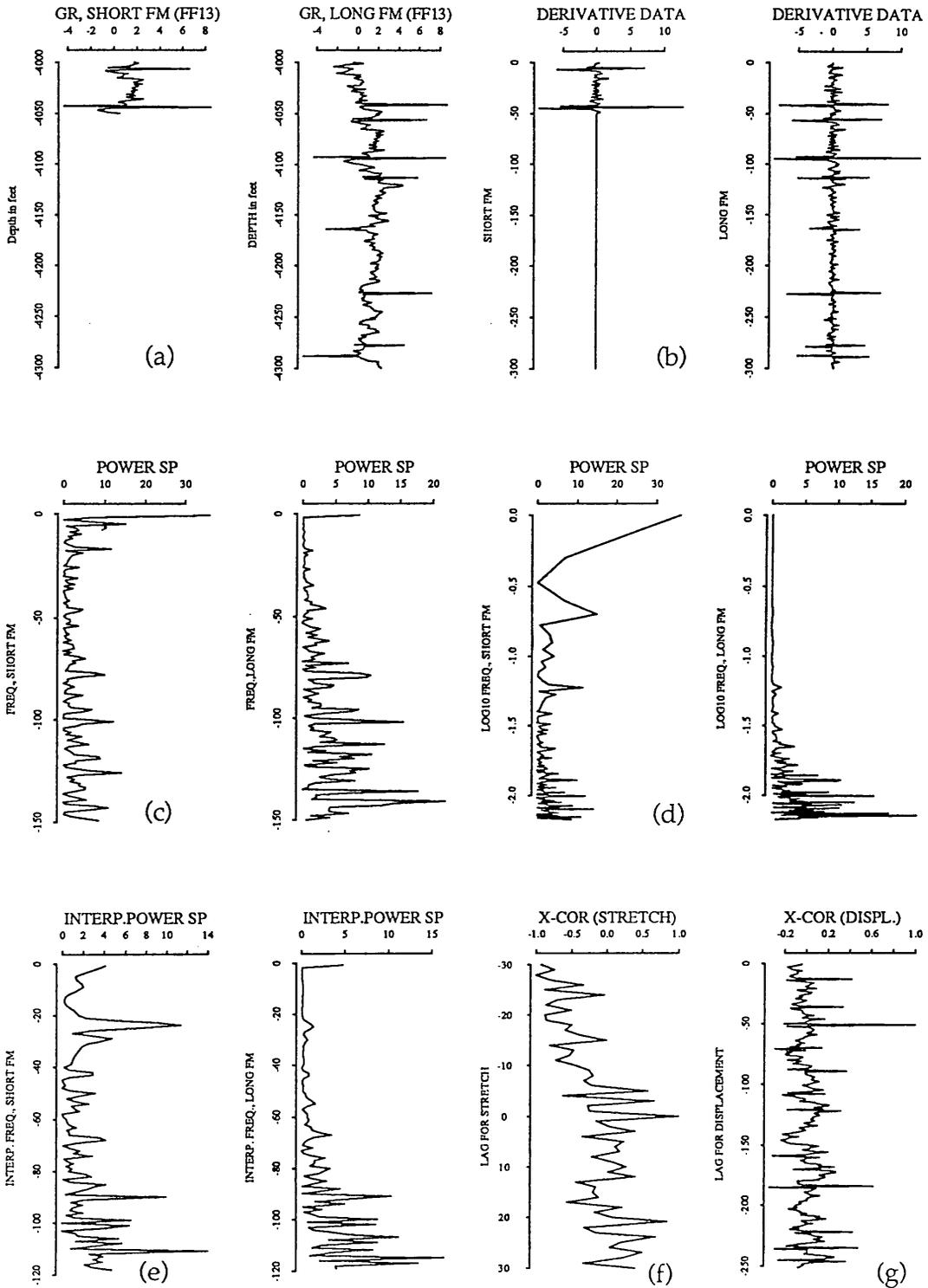


Fig. 4.3 Output plot using model data of the first principal components of thin formation (50 feet) and thick formation (300 feet) and displacement of 50 feet.

first cycle is ignored (Kwon, 1977) because it contains a few data points and it is a waste of computer time to interpolate this component. To obtain equispaced power spectra Lagrang's interpolation is used (Fig. 4.3e). The resultant cross-correlation coefficient between the power spectra is shown in Fig. 4.3f. The cross-correlation coefficient, like the original data (GR), shows a high peak at shift zero which implies that no stretching is involved. A maximum cross-correlation coefficient of 1.0 is observed at a lag of 50 feet (Fig. 4.3g). Appendix E gives the printed output obtained using the principal component scores.

Different thicknesses and different displacements (Fig. 4.4 & Fig.4.5) were tried to test the reliability of this technique and all show promising results. Figure 4.4 shows a test of the program using a thickness of 120 feet for the short series and a thickness of 450 feet for the long series and displacement of 50 feet. In Fig. 4.5, the thickness of the thin Formation is 50 feet and the thickness of the thick Formation is 220 and zero displacement. The results (Figs 4.5f & 4.5g) show a maximum correlation coefficient of 1.0 at zero lag and a maximum correlation coefficient of 1.0 at displacement of 2 feet. (Appendix E)

Although the program results using model data of the original variables of Gamma Ray and model data of the first principal component of all the original variables in Well FF13 in the Attahaddy field are similar, using the first principal component for correlation of well-log data is more reliable. For example in Fig. 4.2 using the original variable of Gamma Ray

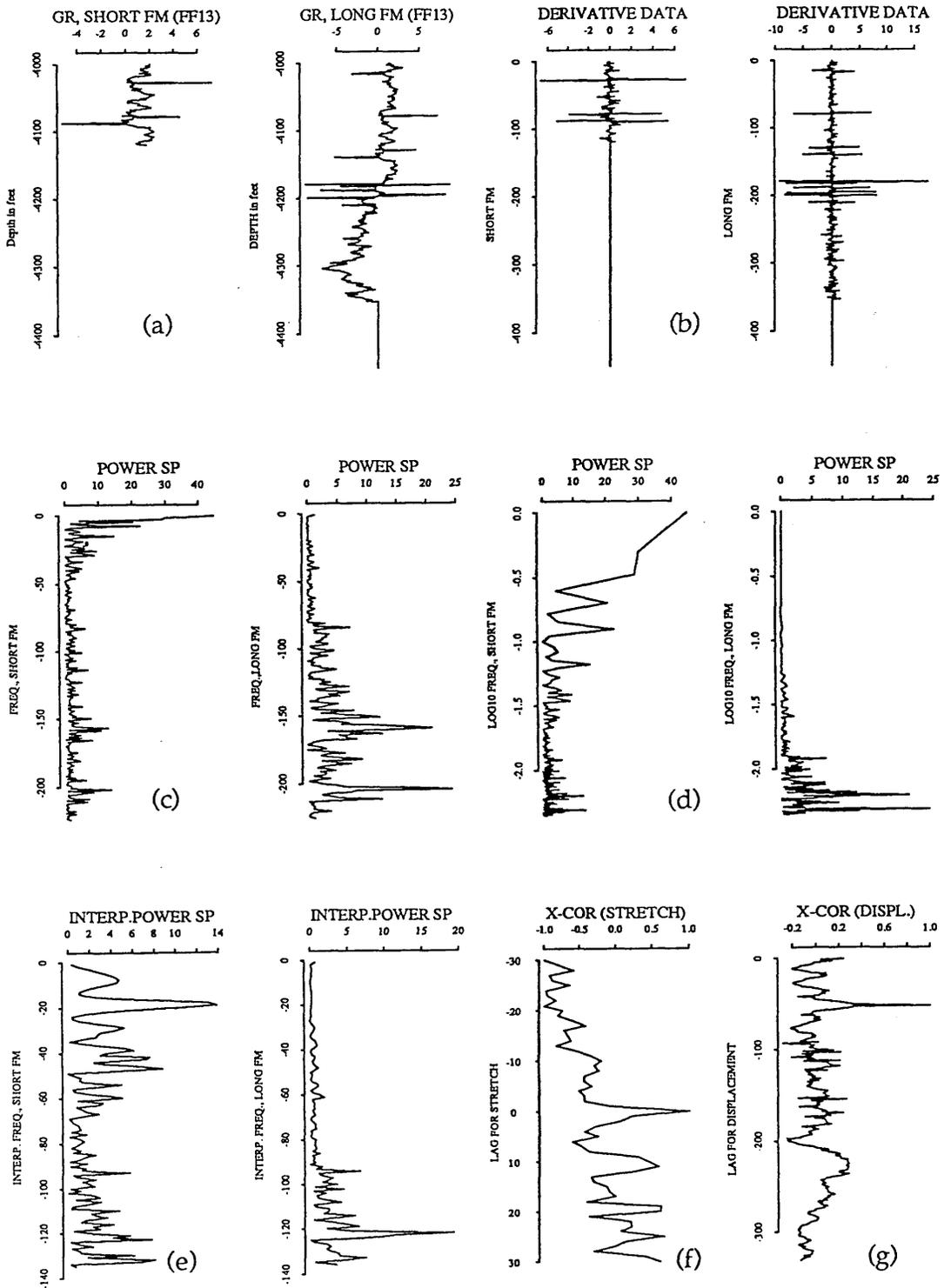


Fig. 4.4 Output plot using model data of the first principal component of thin formation (120 feet) and thick formation (450 feet) and displacement of 50 feet.

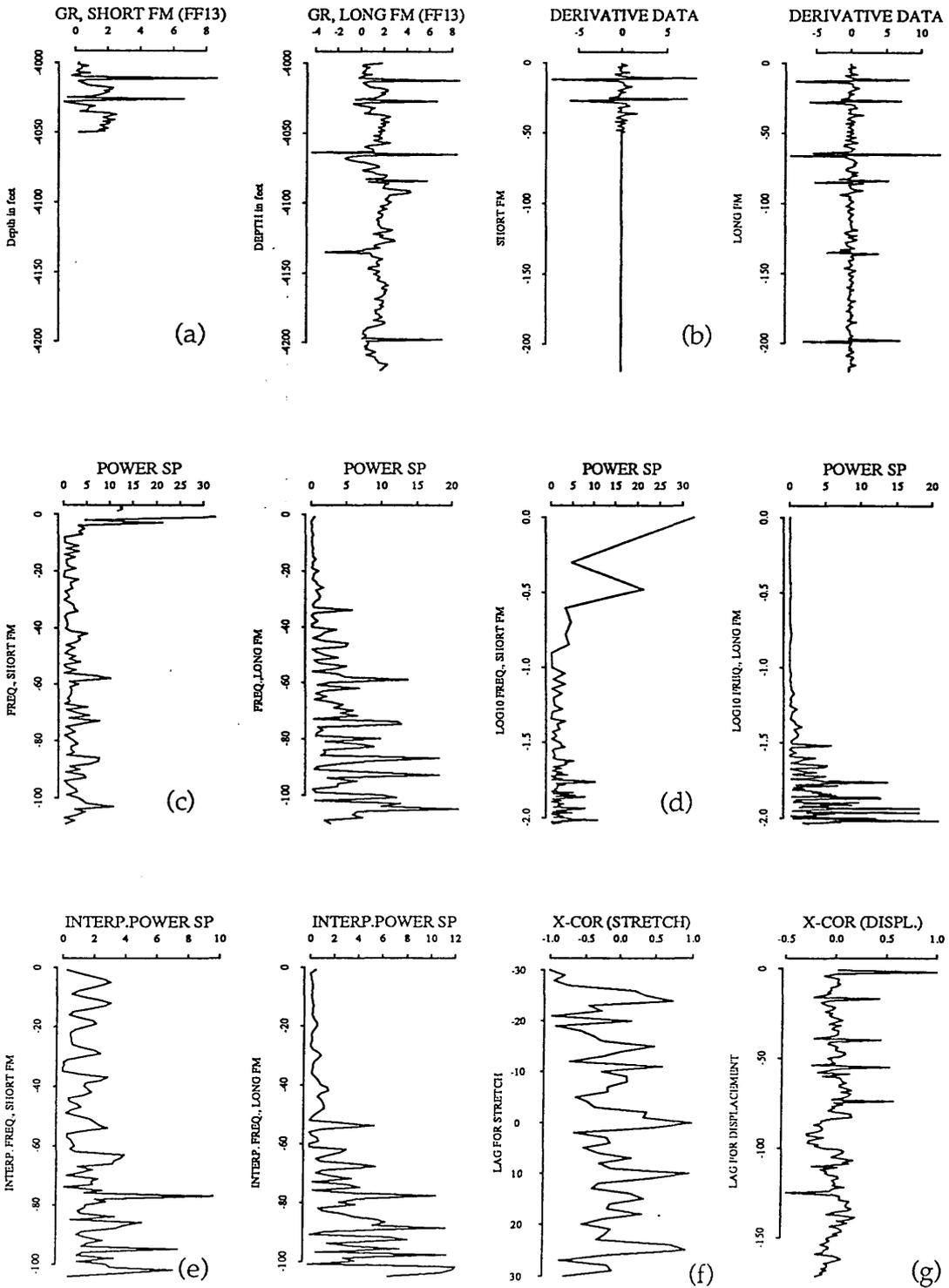


Fig. 4.5 Output plot using model data of the first principal component of thin formation (50 feet) and thick formation (220) without displacement.

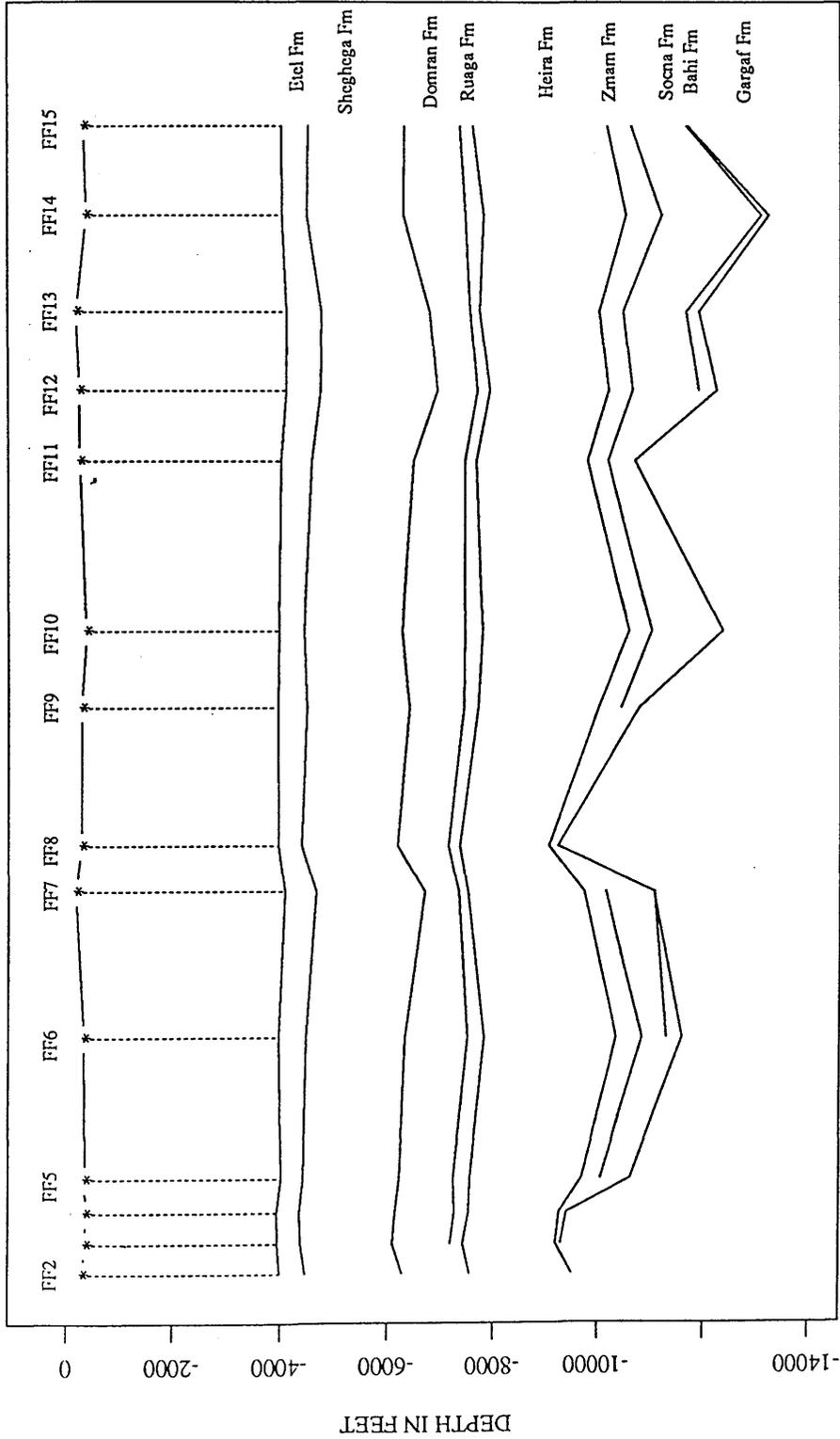
(GR), the derivative data (Fig.4.2b) of both curves are unstable and noisy. Furthermore, the interpolated power spectra (Fig.4.2e), from which the correlation coefficients are calculated, are unstable too. This instability of the original variables is overcome by using the first principal component which gives clear, distinct filtered curves (Fig. 4.3b & 4.3e). The reliability and stability of the first principal component promise fruitful results.

4.3 Interpretation of real data

4.3.1 Introduction

In the previous section the application of PCAXCOR using the first principal components was developed and tested on model data. Empirical results of model data showed that the use of principal component analysis is highly effective in well-to-well correlation. However, real data are complicated by geological variations which do not preserve identical forms nor constant thickness from well to well. This is true especially with geologically complex long sequences and large distances between boreholes such as the Attahaddy field (Fig. 4.6 & Table 4.1). As a result the value of the maximum correlation function will be smaller compared to model data. This is because mathematical correlation gives an entire average of similarity between features of the entire section to be compared. Therefore, the computer selection does not always agree with the geologic selection which considers some other factors when the correlation is made between two logs.

Program PCAXCOR is applied to five boreholes across the Attahaddy field (Fig. 4.7, Appendix A) with variation in distances between successive



HORIZONTAL DISTANCE NOT TO SCALE

Fig. 4.6 Structural cross-section in the Attahaddy field.

Well	Longitude	Latitude	Etel	Sheghega	Domran	Ruaga	Heira	Zmam	Socna	Bahi	Gargaf	KB
Well FF2	193854	293328	3995	4465	6280	-	7560	9496	-	9920	-	297
Well FF3	193741	293252	3951	4382	6099	7194	7427	9192	-	-	9289	362
Well FF4	193612	293301	3940	3463	6142	7266	7538	9264	-	9376	9400	362
Well FF5	193524	293429	4014	4430	6218	7243	7552	9678	10050	-	10635	350
Well FF6	194014	293026	3970	4475	6322	7500	7819	10350	10848	11312	11595	328
Well FF7	193738	293621	4084	4703	6600	7335	7500	9718	10132	11082	11064	173
Well FF8	193657	293433	3960	4392	6165	7140	7348	9038	-	-	9220	287
Well FF9	194150	293054	3948	4485	6396	7430	7696	9998	10444	-	10792	283
Well FF10	194250	292723	3966	4440	6256	7451	7780	10593	11030	-	12345	376
Well FF11	193905	293405	4004	4578	6473	7446	7780	9790	10188	-	10715	248
Well FF12	194206	293301	4015	4738	6925	7691	7924	10199	10669	11908	12243	239
Well FF13	193926	293515	4114	4762	6788	7558	7742	10028	10492	11678	11910	190
Well FF14	193713	293138	4024	4495	6291	7467	7820	10544	11221	13087	13226	375
Well FF15	194059	292942	4013	4508	6298	7357	7609	10175	10642	11674	11712	328

Table 4.1 Longitude & Latitude (in degree, minute, and second) and formation tops (in feet) in the Attahaddy field.

boreholes 2 to 10 kilometres. Five formations are used in studying the boundary identification and cross-correlation. These are: the Etel, Sheghega, Domran, Ruaga and the Hiera formations. The lower part of the sequence comprised of the Zmam, Socna, Bahi and Gargaf formations. These are characterized by very high resistivity values due to the presence of gas and are discarded from the analysis.

4.3.2 Correlation between FF7 and FF13

As the first application of PCAXCOR with real data, two boreholes on the Northern side of the Attahaddy field (Fig.4.7, Appendix A) are used. The distance between the two wells is approximately 3 Km. Well FF7 is a gas well produces from the Gargaf Formation and well FF13 was suspended as a gas well.

The principal components of well FF7 are calculated (Table 1 Appendix E) and the graphical display of its first principal component is shown in Figure 4.8a. The boundaries of the different formations are shown in Figure 4.8b and Table 4.2. The length of the sequence used is from depth of 4460-10946 feet (sampled at 5 feet interval). The Domran Formation at a depth of 6500 feet, the Heira Formation is identified at a depth of 7400 feet, and the Zmam Formation is identified at a depth of 9500 feet.

The boundary identification technique is based on the selection of formation boundaries where the average changes of the sequence values are

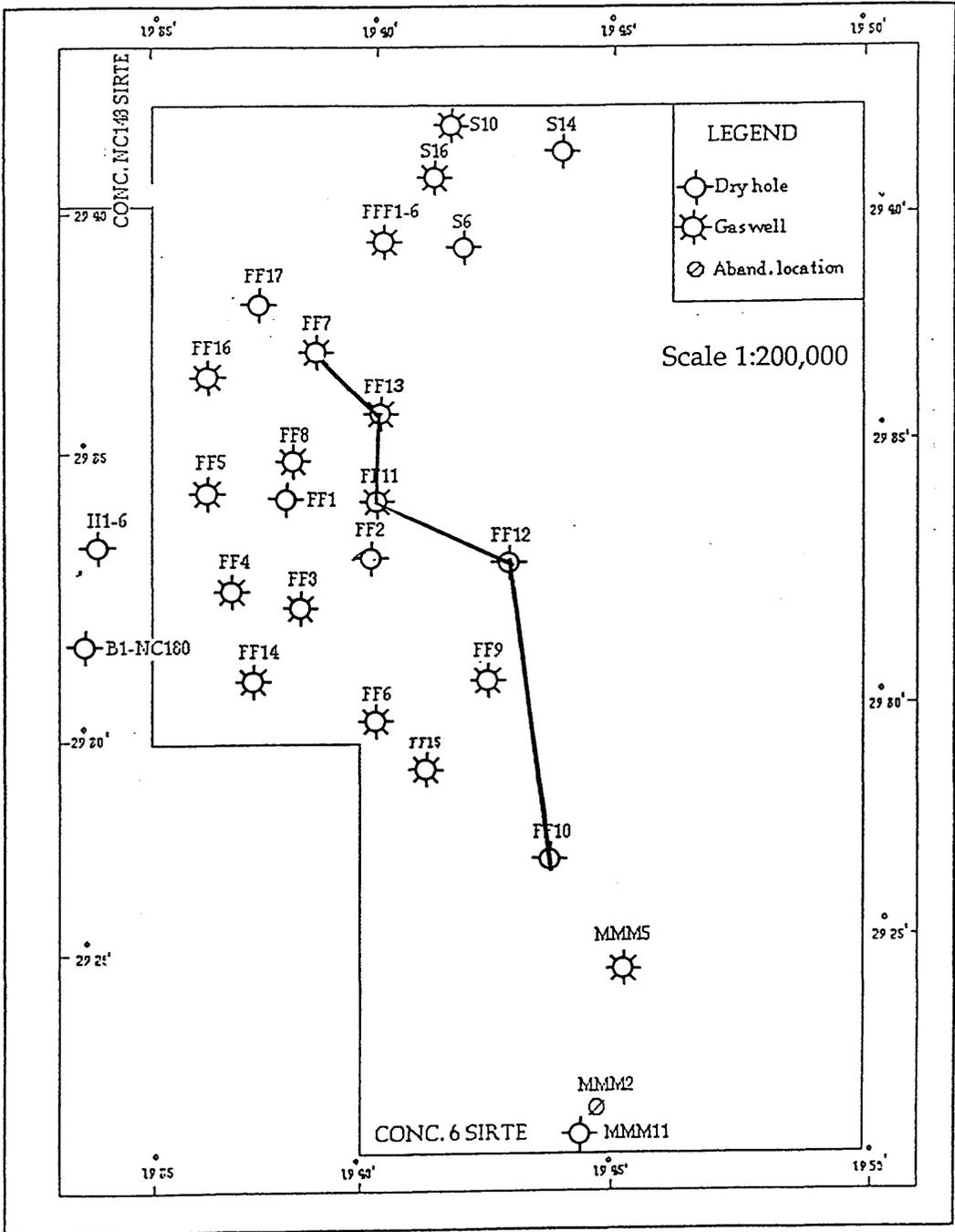


Fig.4.7 Location map of the Attahaddy field. The solid lines shows the boreholes used in the analysis.

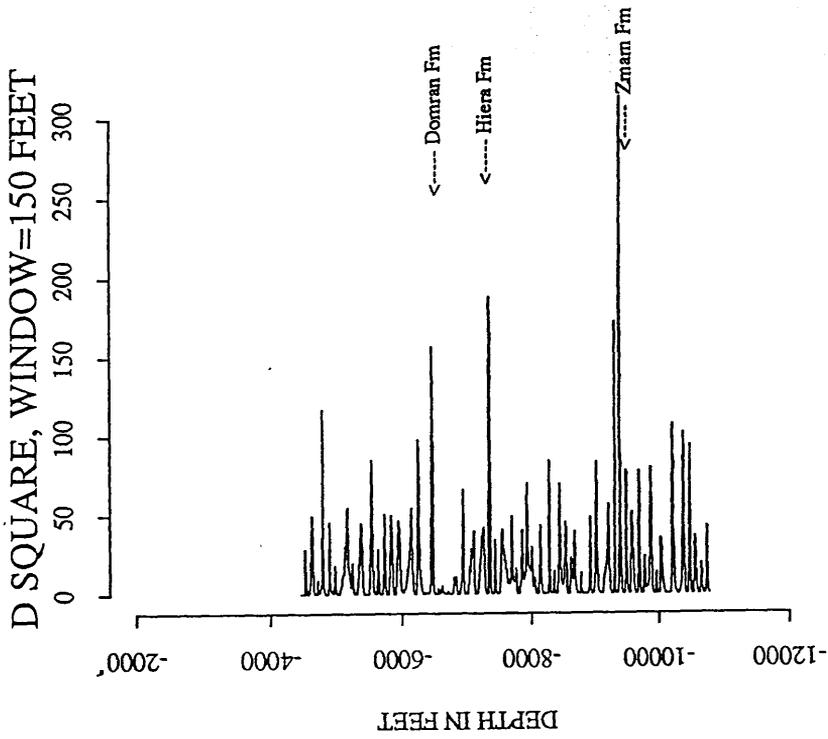
great. However, in the Attahaddy field some Formation boundaries, for example Domran Formation, do not occur at such point, hence some difference between the geological boundaries and the computer selection is observed.

Table 4.2 The Geological Formation depths and the predicted Formation depths of FF7 using the boundary identification technique.

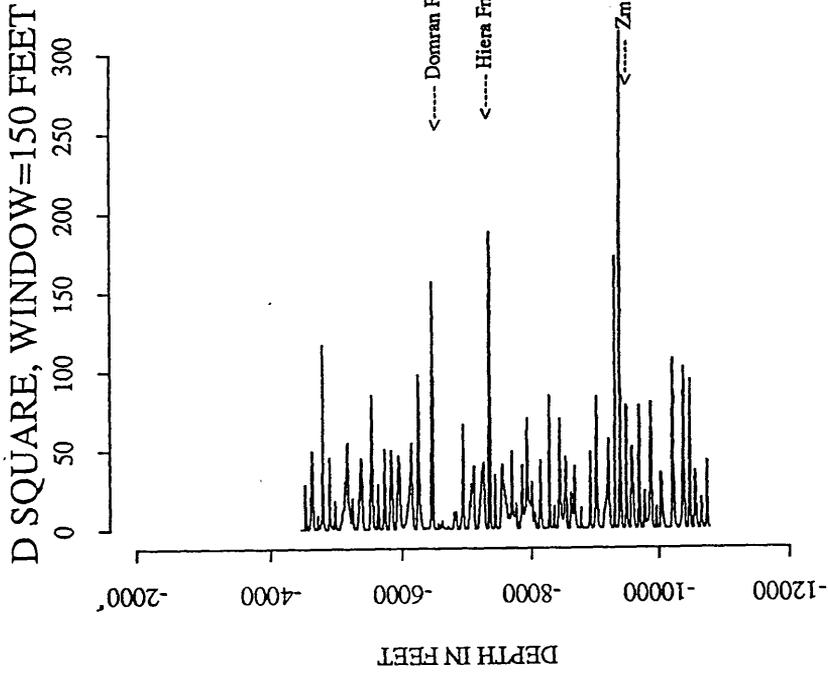
Formation	Geological depth	Predicted depth
Domran	6690	6500
Ruaga	7553	-
Heira	7502	7400
Zmam	9718	9500

Figure 4.9a shows the first principal component for FF13 and its correlation matrix, the eigenvalues, eigenvectors and the percentage of each eigenvalue are shown in Table 4.2 Appendix E.

The formation boundaries of well FF13 are shown in Figure 4.9b and Table 4.3 for the interval between 3500 to 10500 feet (sampled at 5 feet interval). The Etel Formation is identified at depth 4050 feet, the Sheghega Formation at depth 4762 feet, the Domran Formation at depth 6900 feet, the Ruaga Formation at depth 7558 feet, the Hiera Formation at depth 7800, and the Zmam Formation is identified at depth 9800 feet.

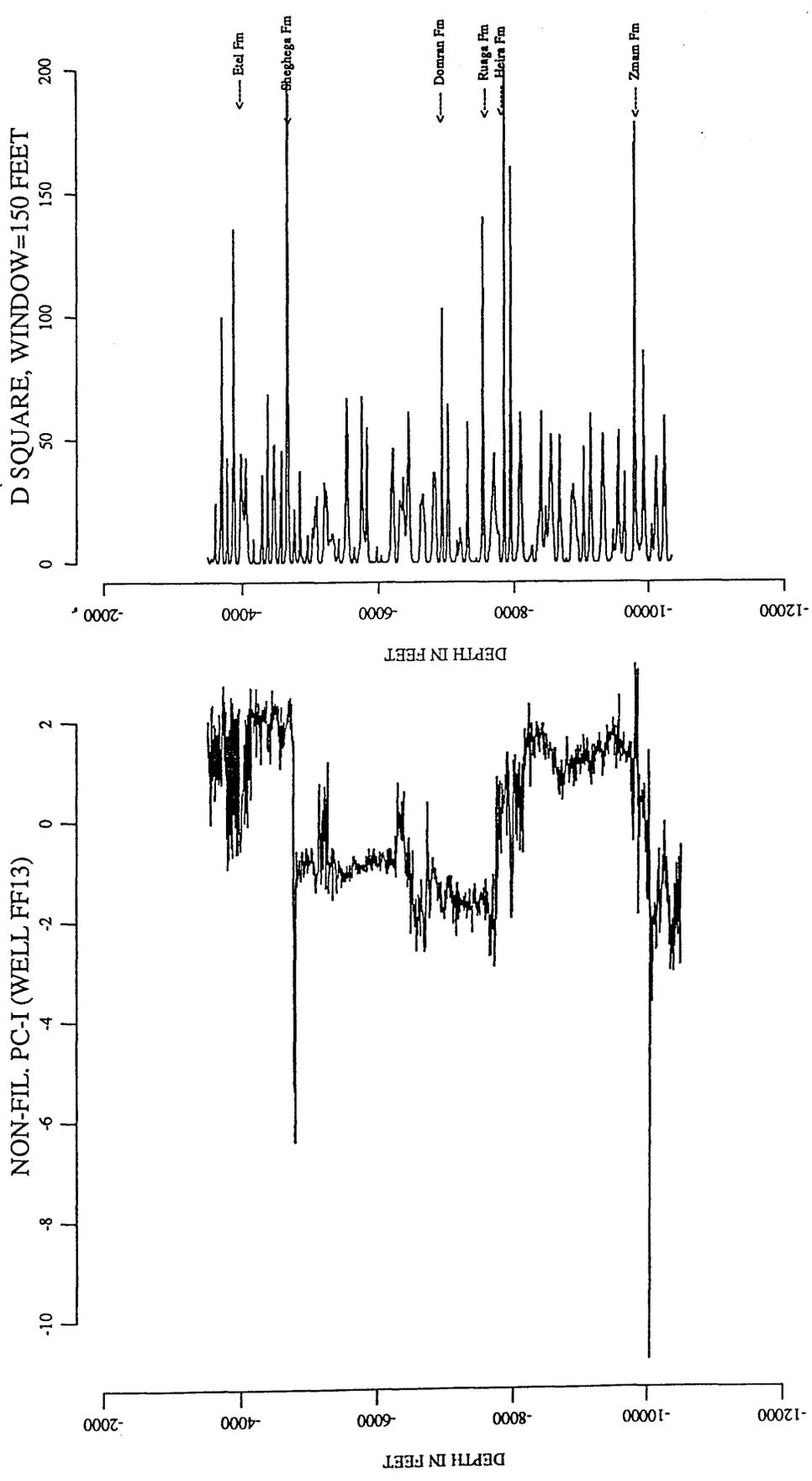


(a)



(b)

Fig. 4.8 Diagram illustrates the first principal component and its boundaries in FF7. (a) the first principal component. (b) are the boundaries of different formations.



(a)

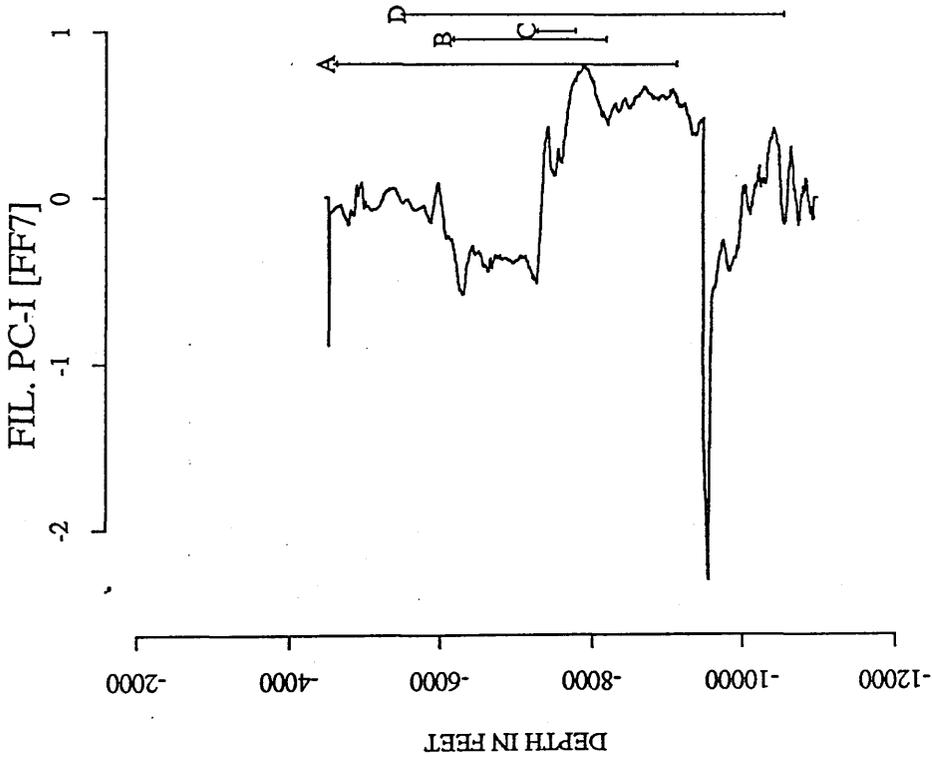
(b)

Fig. 4.9 Diagram illustrates the first principal component and its boundaries in well FF13. (a) the first principal component. (b) are the boundaries of different formations.

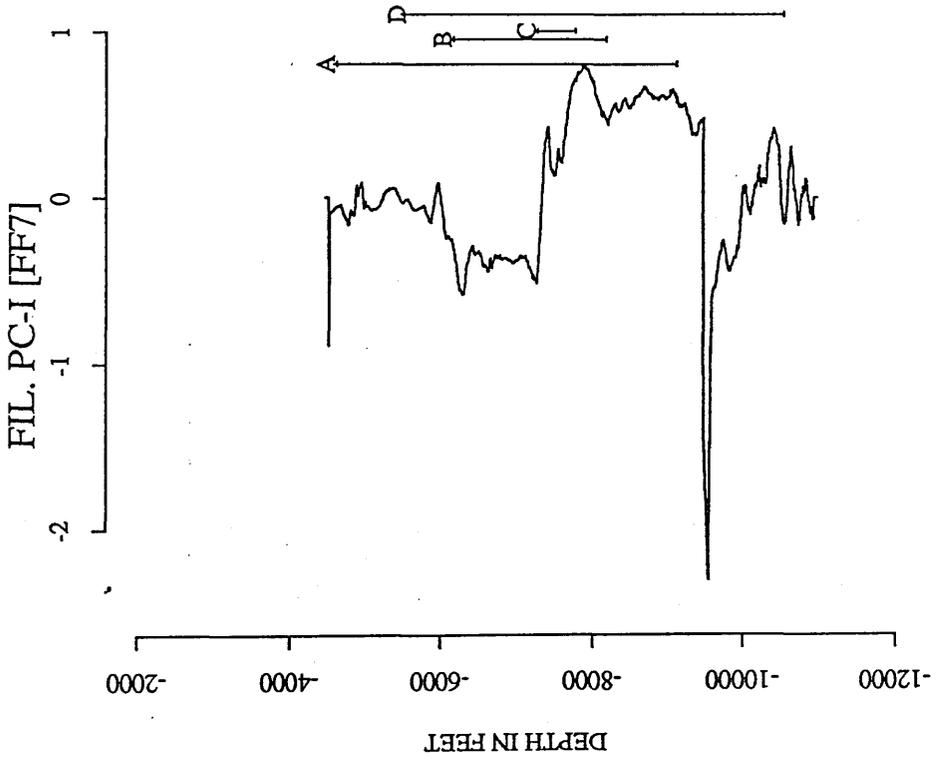
Table 4.3 The Geological formation depths and the predicted formation depths of FF13 using the boundary identification technique.

Formation	Geological depth	Predicted depth
Etel	4114	4050
Sheghega	4762	4762
Domran	6788	6900
Ruaga	7558	7558
Heira	7742	7800
Zmam	10028	9800

After identification of formation boundaries for well FF7 and well FF13, cross-correlation between the smoothed first principal components (Fig. 4.10) of FF7 and that of FF13 is applied. The computer correlation of these components is complicated by sudden change of the apparent magnitude of the sequence values and facies variation which make the visual correlation difficult. For example, some interbeds of limestone in the upper part of the Hiera Formation. Also the variation in lithology from highly calcareous shale and highly argillaceous limestone. Because the thickness in the correlated sequence varies from one formation to another, different window lengths for different formations are used (Fig.4.10).



(a)



(b)

Fig. 4.10 Plot of the filtered principal component of two boreholes (FF7-FF13) in the Attahaddy field. (a) is the first principal component of FF13. (b) is the first principal component of FF7.

4.3.2 (a) Correlation of the Sheghega Formation

The power spectra of the filtered first principal component of the Sheghega Formation in FF13 (indicated by number 1 in Fig. 4.10a) are correlated with power spectra of filtered first component of well FF7 with window length equal approximately double of that of the section in FF13 (Fig. 4.10b, marked by letter A). The derivative data for the short sequence (FF13) and the long sequence (FF7) are calculated (Fig. 4.11b). Figure 4.11c- Figure 4.11e show the power spectra, the transformation to logarithmic scale of power spectra and the interpolated power spectra of both series respectively. The cross-correlation function of these spectra is shown in Figure 4.11f and Figure 4.12b. This function yields a stretch factor of 1.12 (compared with geological stretch of 1.01) for the long sequence. The cross-correlation function for the stretched sequence which has a maximum peak is shown in Fig.4.11g. The relationship between the two series, the cross-correlation function of power spectra and the cross-correlation function of the stretched series is displayed in Figure 4.12. Maximum cross-correlation coefficient for a stretch of 1.12 indicated by an arrow in Fig. 4.12b and a maximum cross-correlation coefficient for displacement yields a maximum peak (0.570) for displacement of 20 feet compared with geological displacement of 59 feet. This result is reasonable for the Sheghega Formation with average thickness of 2000 feet. The small deviation from the geologic correlation is explained by the fact that the computer correlates the average similarity through the selection assuming that the thickening of beds is uniform.

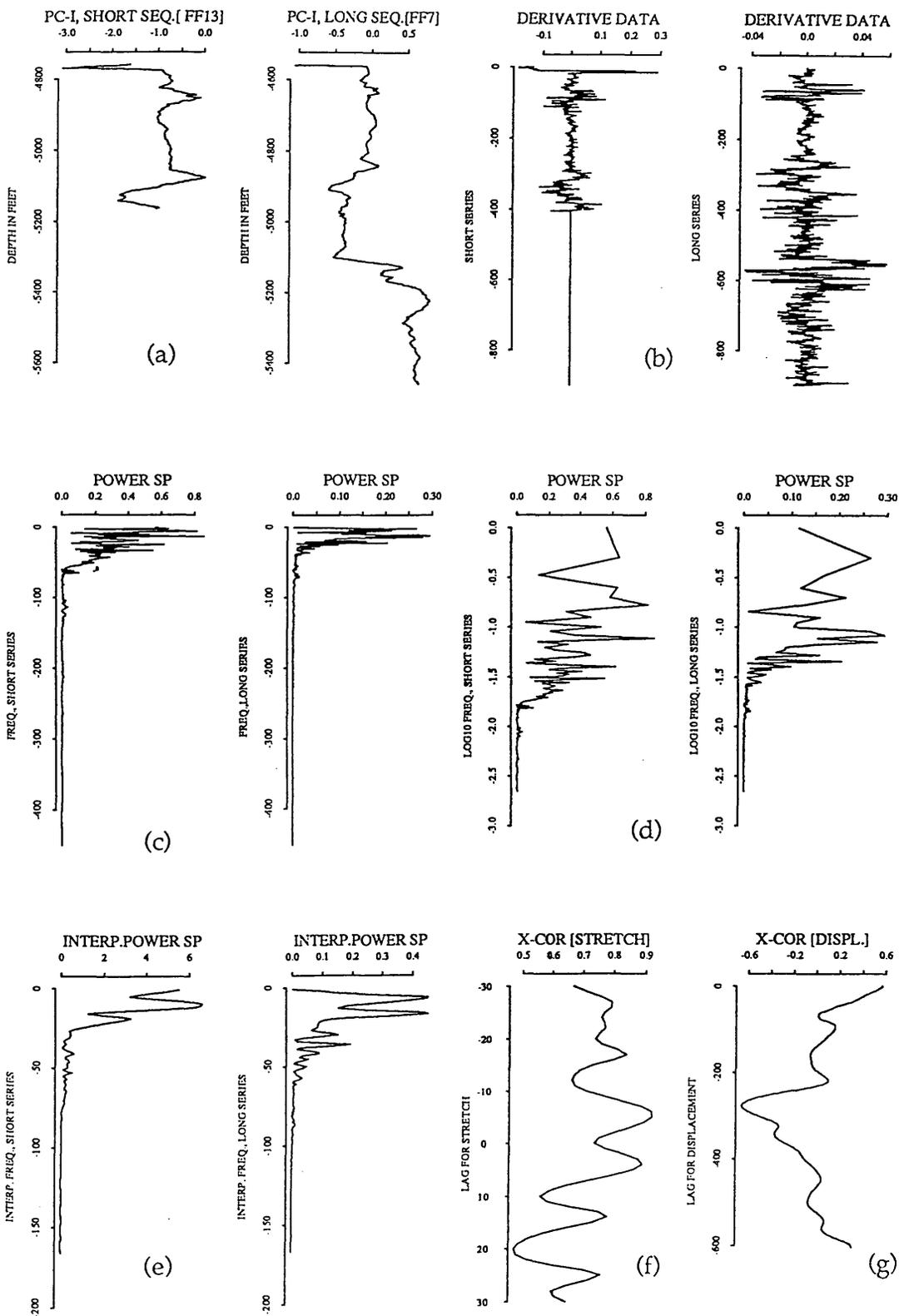


Fig. 4.11 plot of cross-correlation function using real data in Attahaddy field. (a) the smoothed first principal components of the Sheghega Formation in well FF13 is compared with a section (indicated by a letter A in Fig.4.10b). (b) is the derivative of the data. (c) is the power spectra of the derivative data. (d) is the power spectra with logarithmic spaced frequencies. (e) the interpolated power spectra. (f) the cross-correlation function of power spectra. (g) the cross-correlation function of the stretched series.

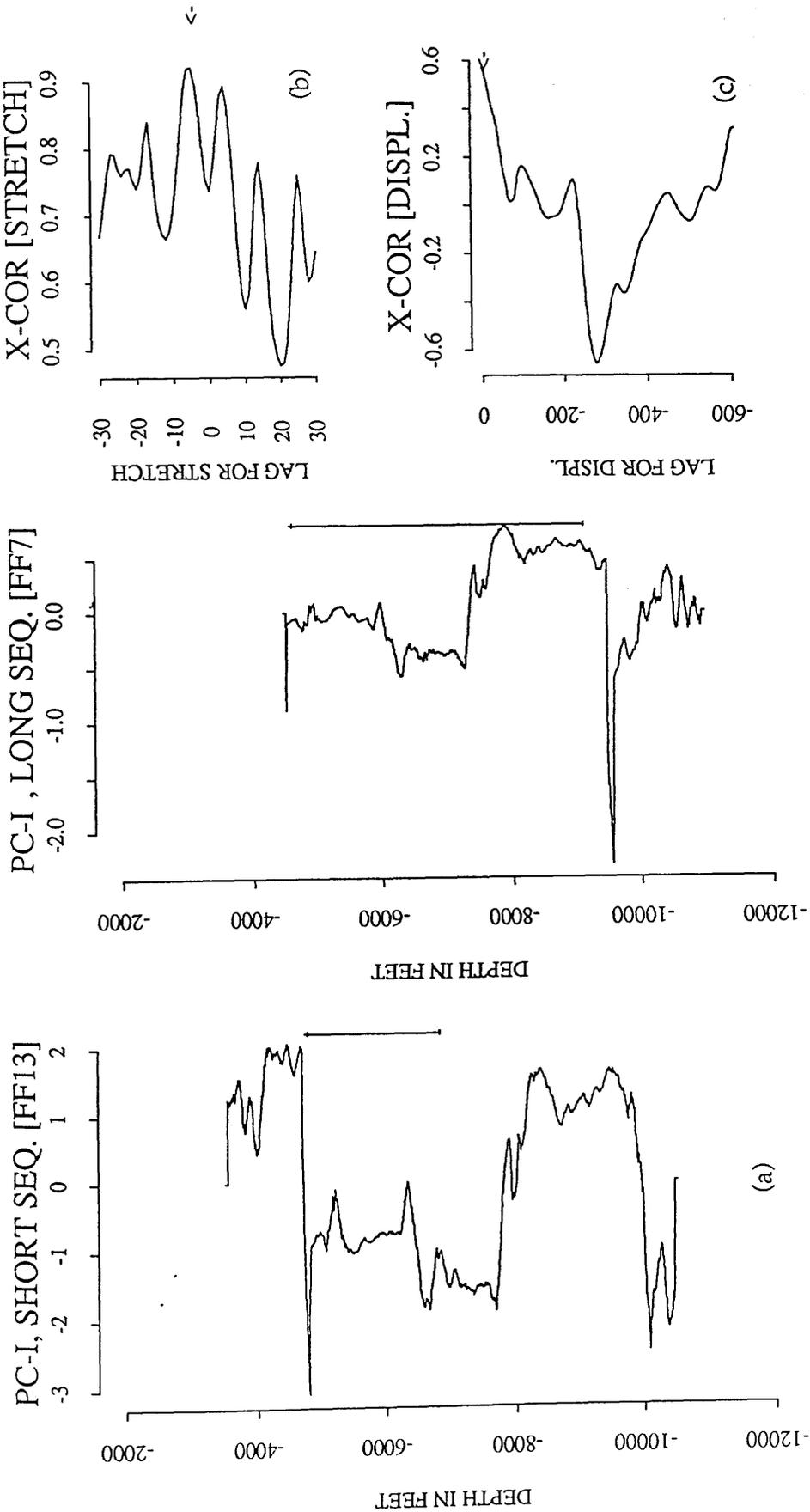


Fig. 4.12 Computer correlation of the Sheghega Formation (FF13 and FF7). (a) the first principal components of FF13 and FF7 (sampled at 5 feet interval). The tie lines show computer correlations, and the dashed lines show geological correlations. (b) the cross-correlation function of power spectra (peak at 5, S=1.12). (c) the cross-correlation function of the stretched series (0.579) at a lag of 20 feet when the long sequence is stretched 1.12 times.

An attempt to correlate the Sheghega Formation using the non-filtered principal component of well FF7 and well FF13 (Fig. 4.13) failed both in predicting the stretch factor and in obtaining the accurate displacement of stretched series. The cross-correlation of power spectra yielded a stretch factor of 1.51 (Fig. 4.13f) compared with geological stretch of 1.01 and cross-correlation function of 0.62 at displacement of 265 feet compared with geological displacement of 59 feet (Fig. 4.13g). This was not unexpected, as the correlation using non-filtered principal component is characterized by noise component both in the time (Fig. 4.13a) and the derivative (Fig. 4.13b) domains.

4.3.2 (b) Correlation of the Domran Formation

Difficulties are encountered in visually correlating the upper part of the Domran Formation and in the computer identification of the upper boundary of the Domran Formation (Fig. 4.9b). This is because the lower part of the Sheghega limestone is similar in the log response to the upper part of the Domraan Formation, consequently, the average change in the log response occurs in the lower part of the Sheghegan Formation and not at the contact between the two formations.

Computer correlation is attempted (Fig. 4.14). The cross-correlation function of power spectra (Fig. 4.14b) yields a peak at $+v=+9$ which corresponds to a stretch of 1.23 compared with 1.19 between FF7 and FF13. The positive sign of v indicates that a portion of the long sequence is

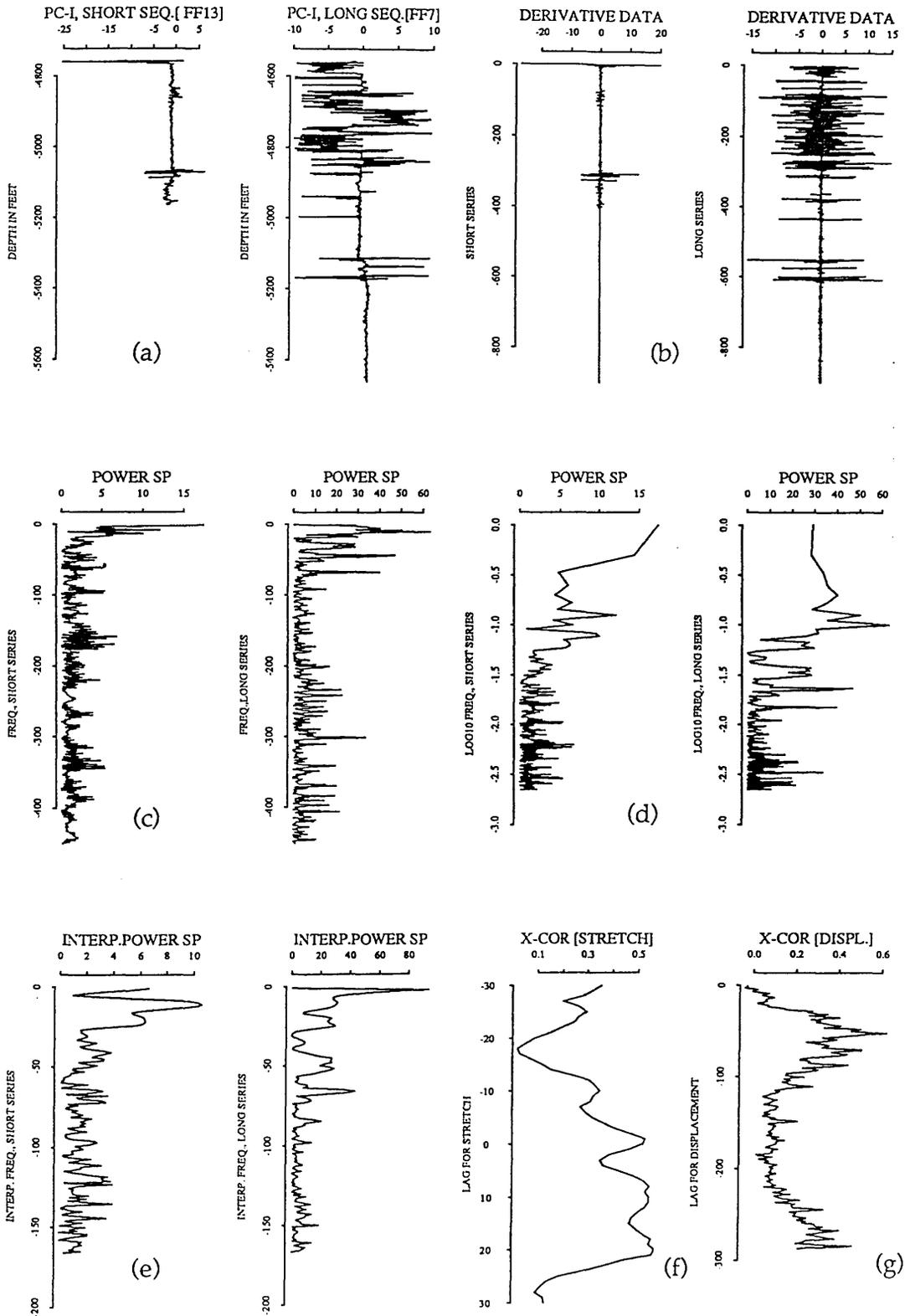


Fig. 4.13 plot of cross-correlation function of the Sheghega Formation using non-filtered data of FF13 and FF7. (a) the non-filtered first principal components of the Sheghega Formation in well FF13. (b) is the derivative of the data. (c) is the power spectra of the derivative data. (d) is the power spectra with logarithmic spaced frequencies. (e) the interpolated power spectra. (f) the cross-correlation function of power spectra. (g) the cross-correlation function of the stretched series.

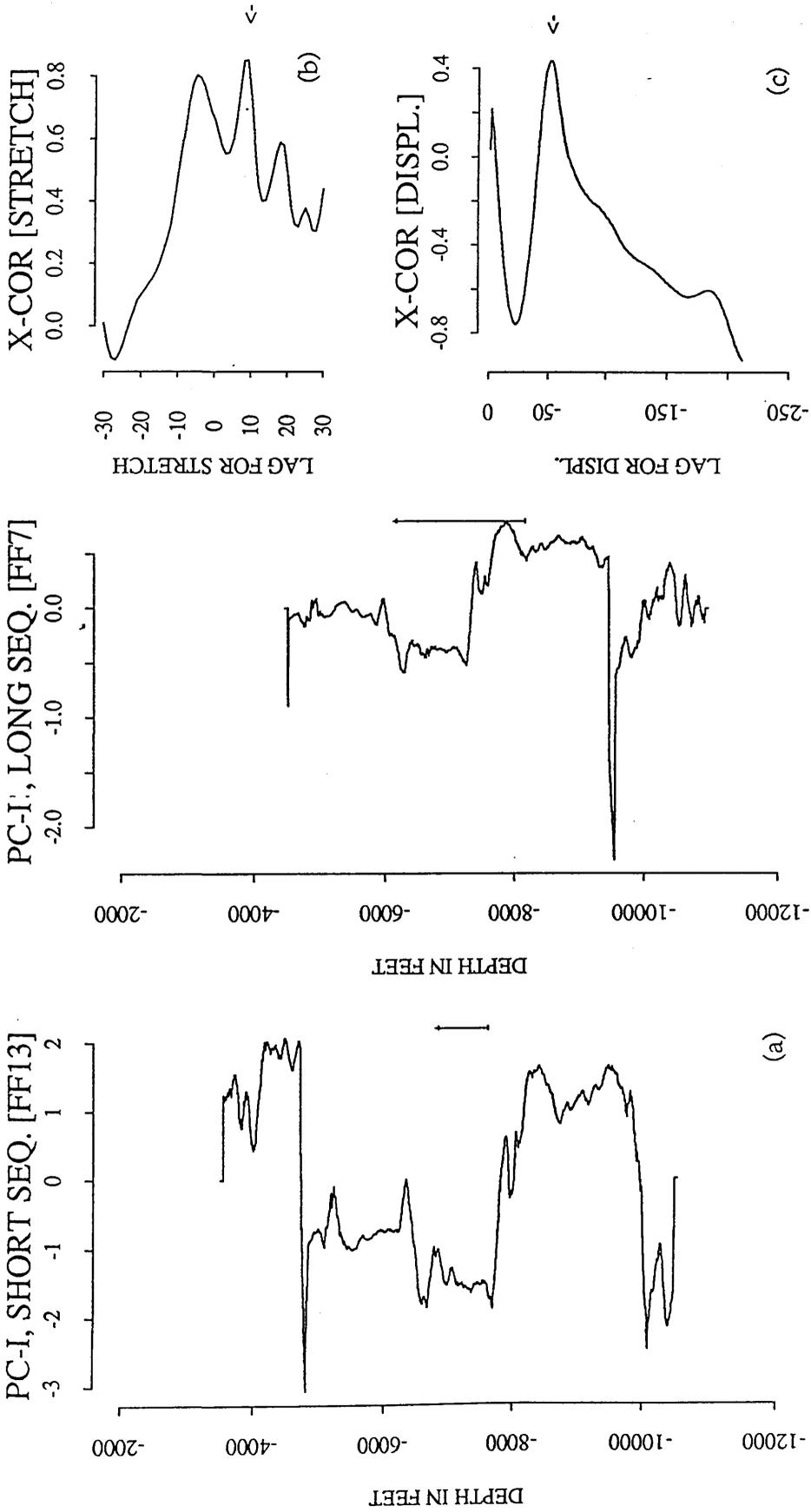


Fig. 4.14 Cross-correlation of the Domran Formation (FF13 and FF7). (a) the first principal components of FF13 and FF7. The tie lines show computer correlations, and the dashed lines show geological correlations. (b) the cross-correlation function of power spectra (peak at $+9$, $S=1.23$). (c) the cross-correlation function (0.433) of the stretched series at displacement of 51 units (265 feet) when the long sequence is stretched 1.23 times.

stretched relative to the short sequence (series). This implies that the direction of the thickening is from well FF13 towards FF7 which agrees with the geological direction. The stretch factor is computed from equation 3.50 :

$$S=10^{9 \times 0.01} = 1.23$$

which is expected for Domraan Formation which has an average thickness of about 700 in the two boreholes. The cross-correlation function of the stretched series has a maximum of 0.433. This reduction in the magnitude of the cross-correlation function results from the fact that the noise signal $h(n)$ (Fig. 3.13) which is not overlapped with the correlative section in the time domain, is transferred to an additive spectrum in the frequency domain. Although the cross-correlation function is the most valid method to detect the desired signal from background noise, the decrease of maximum coefficient value is inevitable for such a case.

4.3.2 (c) Correlation of the Ruaga Formation

Mathematical correlation (Fig. 4.15) agrees with the known stratigraphic correlation for the Ruaga Formation. The maximum peak of the correlation function of the power spectra yields an optimum stretch factor, $S=1.0$ that is no stretch compared with geological thickening of 1.10 for long sequence (FF7) (Fig. 4.15b). The maximum correlation function (0.760) for displacement of 200 feet (40 units) compared with 223 feet is also easily identifiable (Fig. 4.15c).

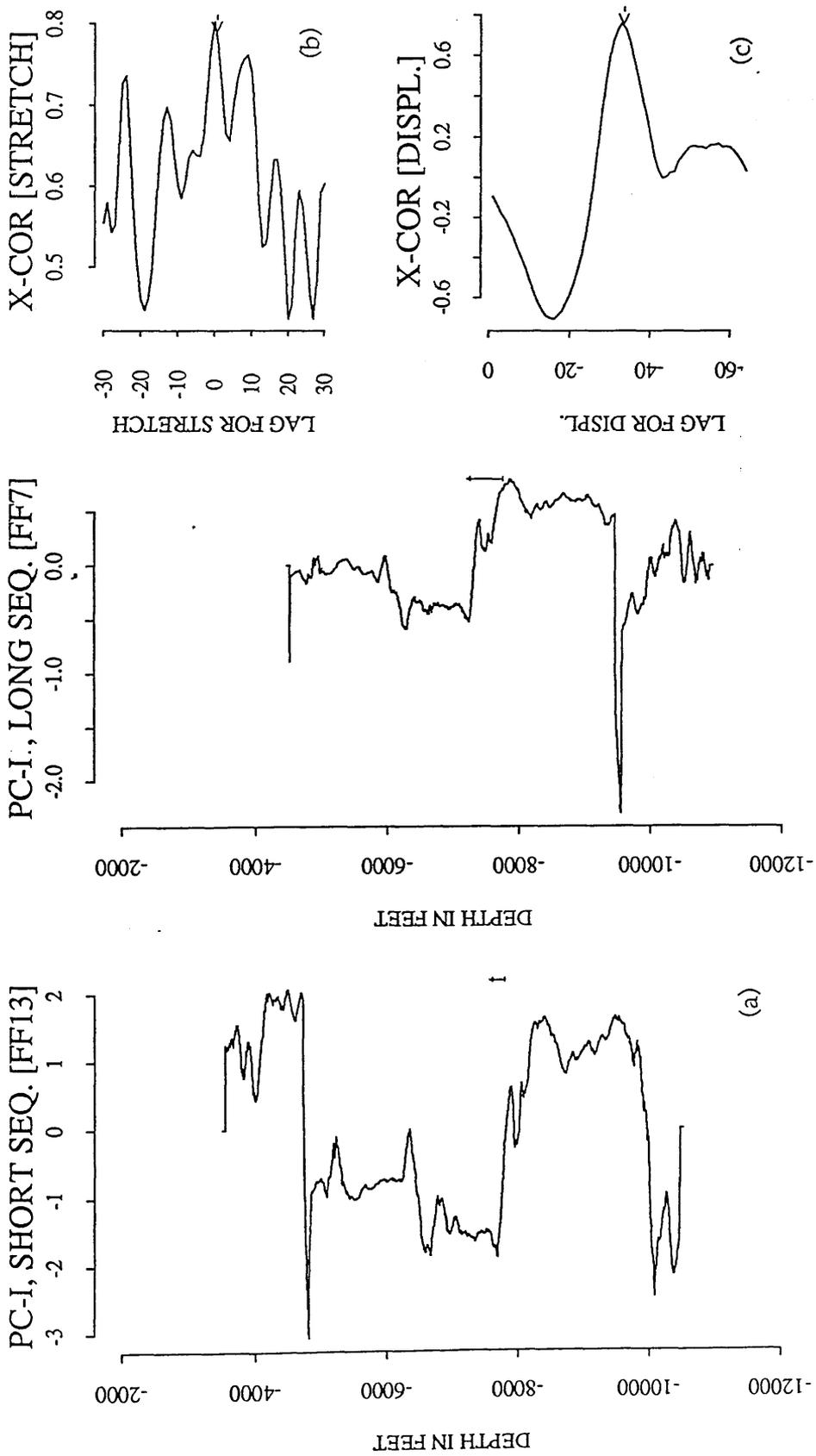


Fig. 4.15 Cross-correlation of the Ruuga Formation (FF13 and FF7). (a) the first principal components of FF13 and FF7. (b) the cross-correlation function of power spectra (peak at 0, S=1, no stretch). (c) the cross-correlation function (0.760) at a lag of 200 feet when the long sequence is stretched 1.0 times.

4.3.2 (d) Correlation of the Heira Formation

Cross-correlation of power spectra resulted in a correlation function which has two peaks, at a lag of $-v=-17$, that is, a stretch factor $S=1.48$, and at a lag of $-v=-28$, that is a stretch factor $S=1.90$ (Fig. 4.16g and Fig. 4.17b). In PCAXCOR, the correlation is based on two choices of maximum correlation functions of power spectra. Although the computer selection of the second peak, $S=1.48$ at a lag of 17 units at displacement of 586 units (2930 feet)(actual displacement times the stretch factor, i.e 1980×1.48) does not agree with the geological thickening (1.04) for the long sequence (i.e FF7, the correlation of the derivative data yields a high value of correlation coefficient (0.748) (Fig.4.17c).

By changing the window size (Fig. 4.18), a higher correlation coefficient (0.867) and a better stretch factor ($S=1.12$) are obtained. Furthermore, the displacement (275 feet) is very close to the geological displacement (240) between FF13 and FF7.

This result demonstrates that the correlation of power spectra is dependant on the selection of the optimum window size used for cross-correlation to give the best stretching of strata and the displacement between two series from the maximum correlation functions. Although the selection of the optimum values from such functions are not always so obvious as when model data is used, the choice by PCAXCOR based on the highest two peaks generally represent a geologically reasonable stretch value and

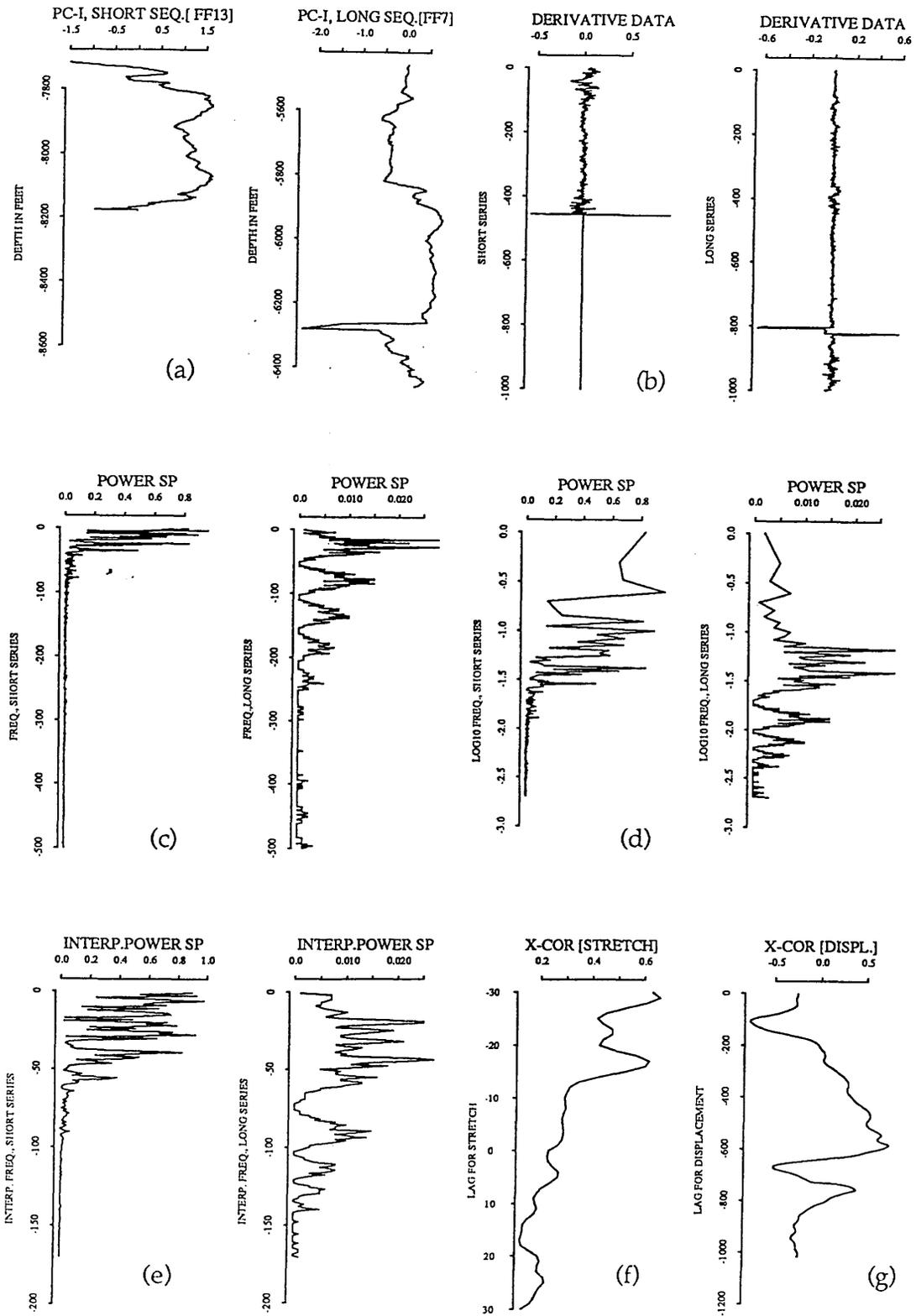


Fig. 4.16 plot of cross-correlation function of the Heira Formation using the derivative data of FF13 and FF7. (a) the first principal components of the Heira Formation in well FF13. (b) is the derivative of the data. (c) is the power spectra of the derivative data. (d) is the power spectra with logarithmic spaced frequencies. (e) the interpolated power spectra. (f) the cross-correlation function of power spectra. (g) the cross-correlation function of the stretched series.

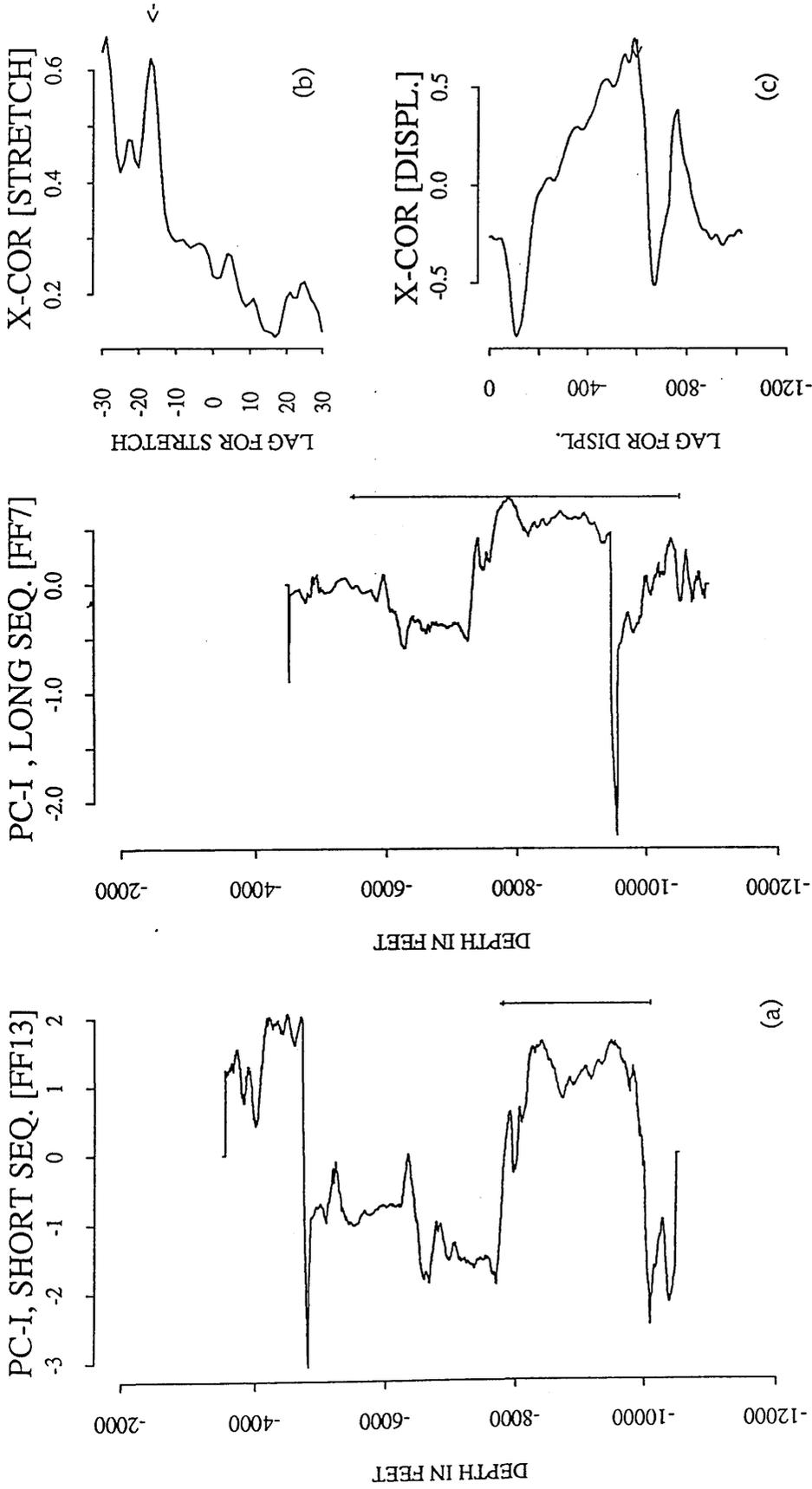


Fig. 4.17 Cross-correlation of the Heira Formation using the derivative data. (a) the first principal components of FF13 and FF7. (b) the cross-correlation function of power spectra at a lag of $-17(S=1.48)$. (c) the cross-correlation function at a displacement of 586 units which is the product of the actual displacement (396) times the stretch factor (1.48). Short log is stretched.

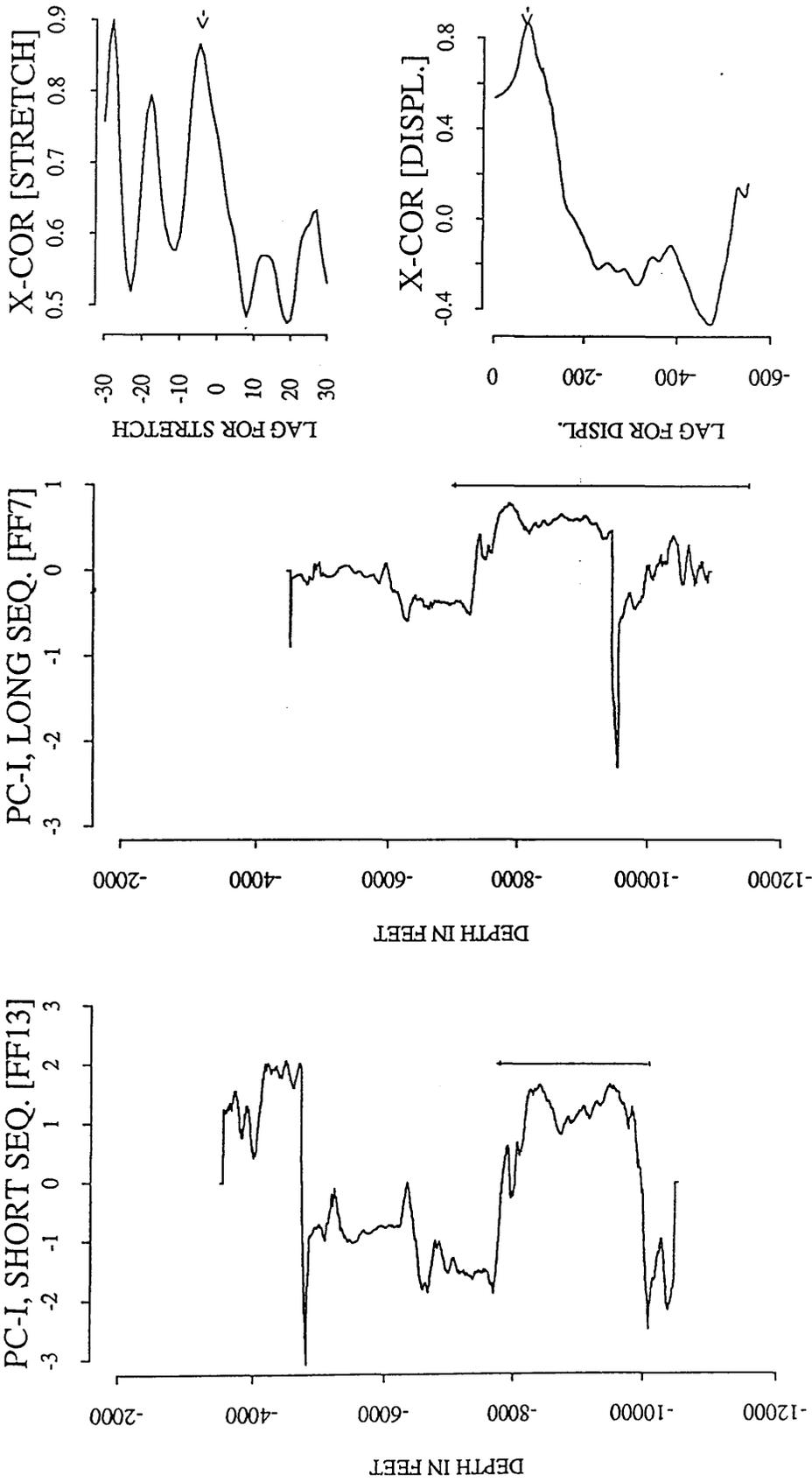


Fig. 4.18 Cross-correlation of the Heira formation using the original data (filtered principal components) for stretching. (a) the first principal components of FF13 and FF7. (b) the cross-correlation function of power spectra ($S=1.12$). (c) the cross-correlation function (0.867) at a lag of 275 feet when the short sequence is stretched 1.12 times.

displacement as will be demonstrated in the following sections.

4.3.3 Correlation between FF13 and FF11

These two wells are about 3 Km distant and the sequence to be analysed is between depth of 3500 and 10500 feet in each borehole. This includes the top of the Etel Shale Formation through to the bottom of the Hiera Shale Formation in which a total of 7000 feet is to be interpreted.

The eigenvalues, the eigenvectors and the percentage of the each eigenvalue of the correlation matrix of the original variables (GR, SP, ILM, ILS, ILM, DT and CALI) of well FF11 is listed in Table 4 Appendix E and its graphical display of the first principal component is shown in Figure 4.19a.

The boundaries of different formations are identified using the boundary identification technique and are shown in Figure 4.19b. The Sheghega Formation is identified at depth of 4578 feet, the Domran Formation at a depth of 6300 feet which was picked at the average change of the curve response at the lower part of the Sheghega Formation and not at the contact between the two formations (depth 6473 feet). The Ruaga Formation is located at depth of 7350 feet. The Hiera Formation is identified at depth of 7600 feet, and the Zmam Formation at depth 10000 feet. The boundaries of well FF13 were identified in section 4.3.2 (Fig.9b)

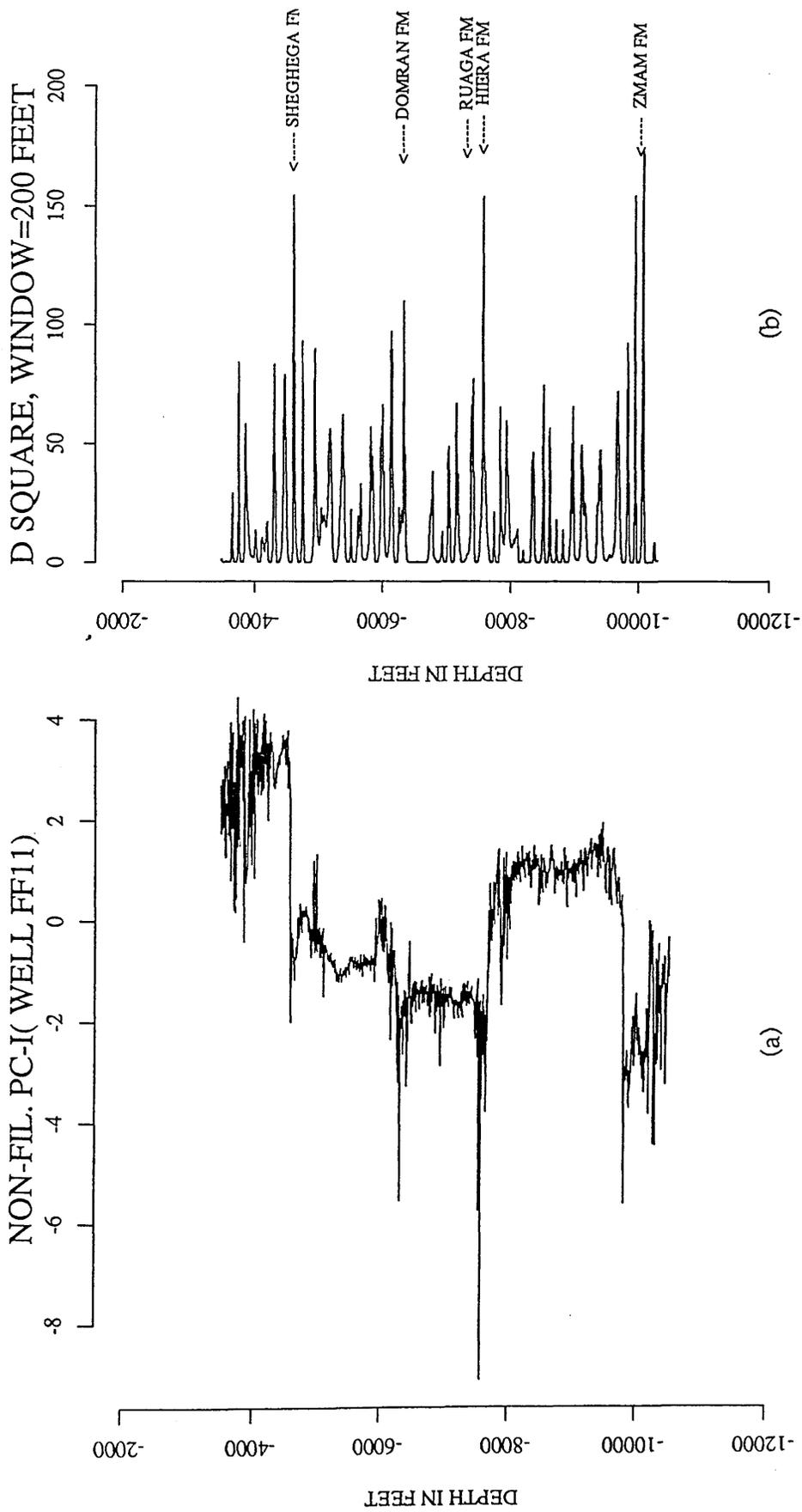


Fig. 4.19 Plot of the first principal components (sampled at 5 feet) of well FF11 and its boundaries. (a) non-filtered principal component of well FF11. (b) boundaries of different formation (window=200 feet)

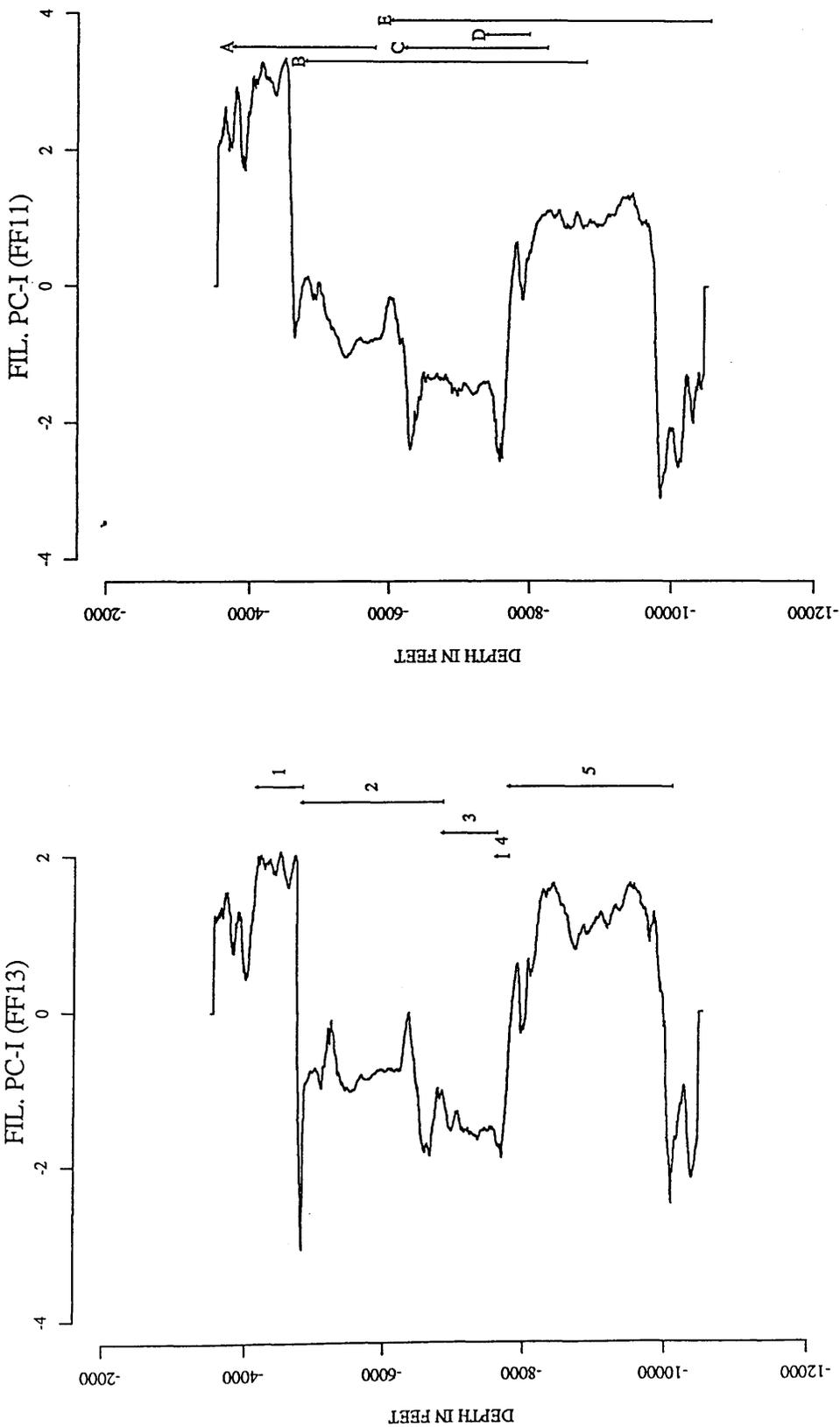
Table 4.4 The Geological formation depths and the predicted formation depths of FF11 using the boundary identification technique.

Formation	Geological depth	Predicted depth
Sheghega	4578	4578
Domran	6473	6300
Ruaga	7446	7350
Heira	7780	7600
Zmam	9790	10000

The first principal component of well FF11 is correlated against that of FF13 with variable window size for different formations (Fig. 4.20).

4.3.3 (a) Correlation of the Etel Formation

The Etel Formation in well FF13 (indicated by number 1 in Fig. 4.20a) is compared with a section from FF11 (indicated by the letter A in Fig. 4.20b). The derivatives of the data are used in calculating power spectra and for final correlation between the stretched sequence (Fig. 4.21). The resultant cross-correlation function of power spectra yields a distinct peak for a lag of +2 which corresponds to a stretch factor of $S=1.05$ (compared with geological stretch of 1.13)(Fig. 4.22b) for long series (FF11) and agrees with known geological thickening from FF13 towards FF11. The cross-correlation function of the stretched series yields a maximum peak (0.808)(Fig. 4.22c) for



(a)

(b)

Fig. 4.20 Plot showing the smoothed principal component of the correlated sequences (FF13 and FF11). (a) smoothed component of well FF13. (b) smoothed component of well FF11. A window (marked 1) in (a) is correlated with a window (marked A) in (b). The window (marked 2) in (a) is compared with the window (marked B) in (b), and so forth.

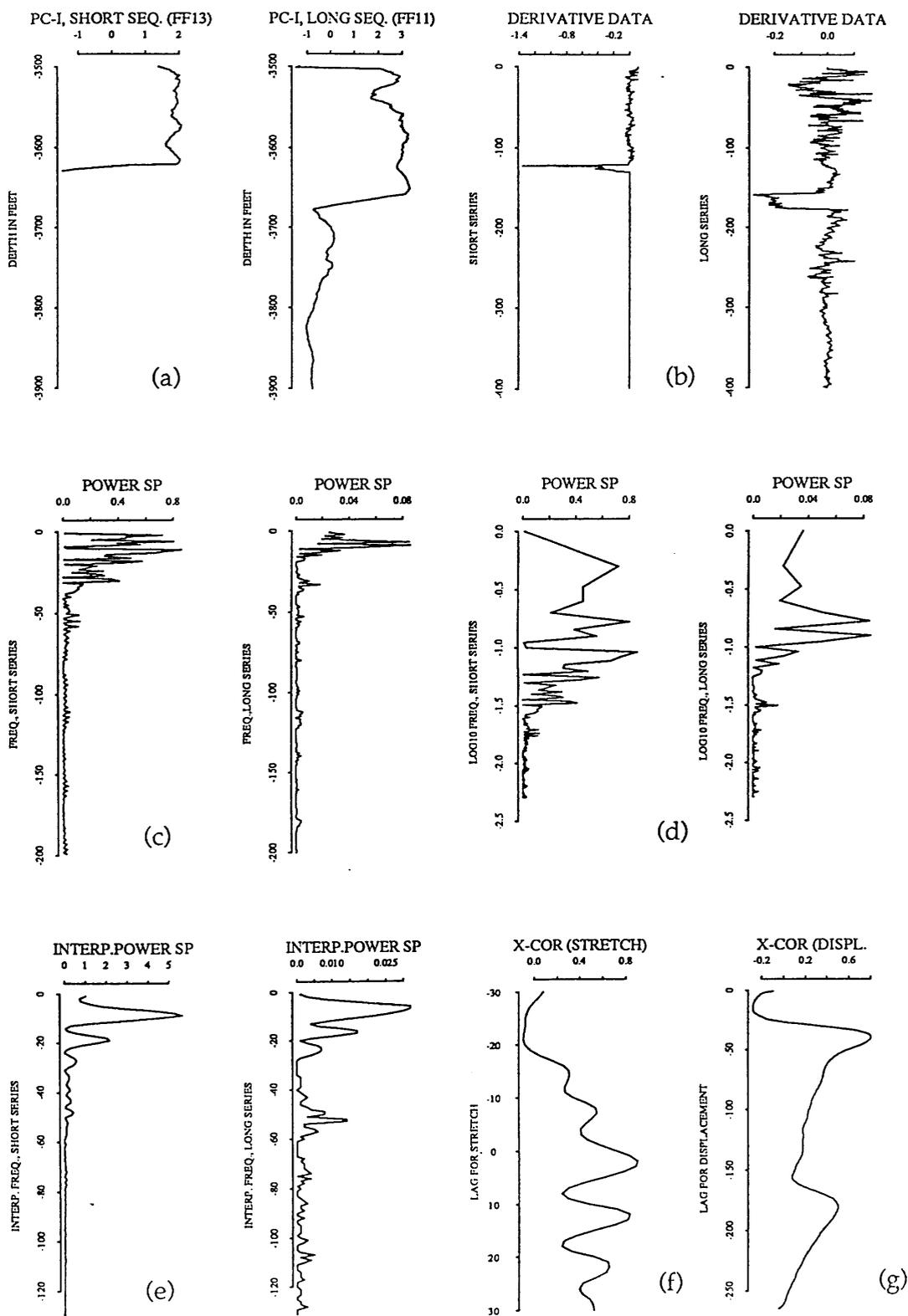


Fig. 4.21 plot of cross-correlation function of the Etel Formation using the derivative data of FF13 and FF11. (a) the first principal components of the Etel Formation in well FF13. (b) is the derivative of the data. (c) is the power spectra of the derivative data. (d) is the power spectra with logarithmic spaced frequencies. (e) the interpolated power spectra. (f) the cross-correlation function of power spectra. (g) the cross-correlation function of the stretched series.

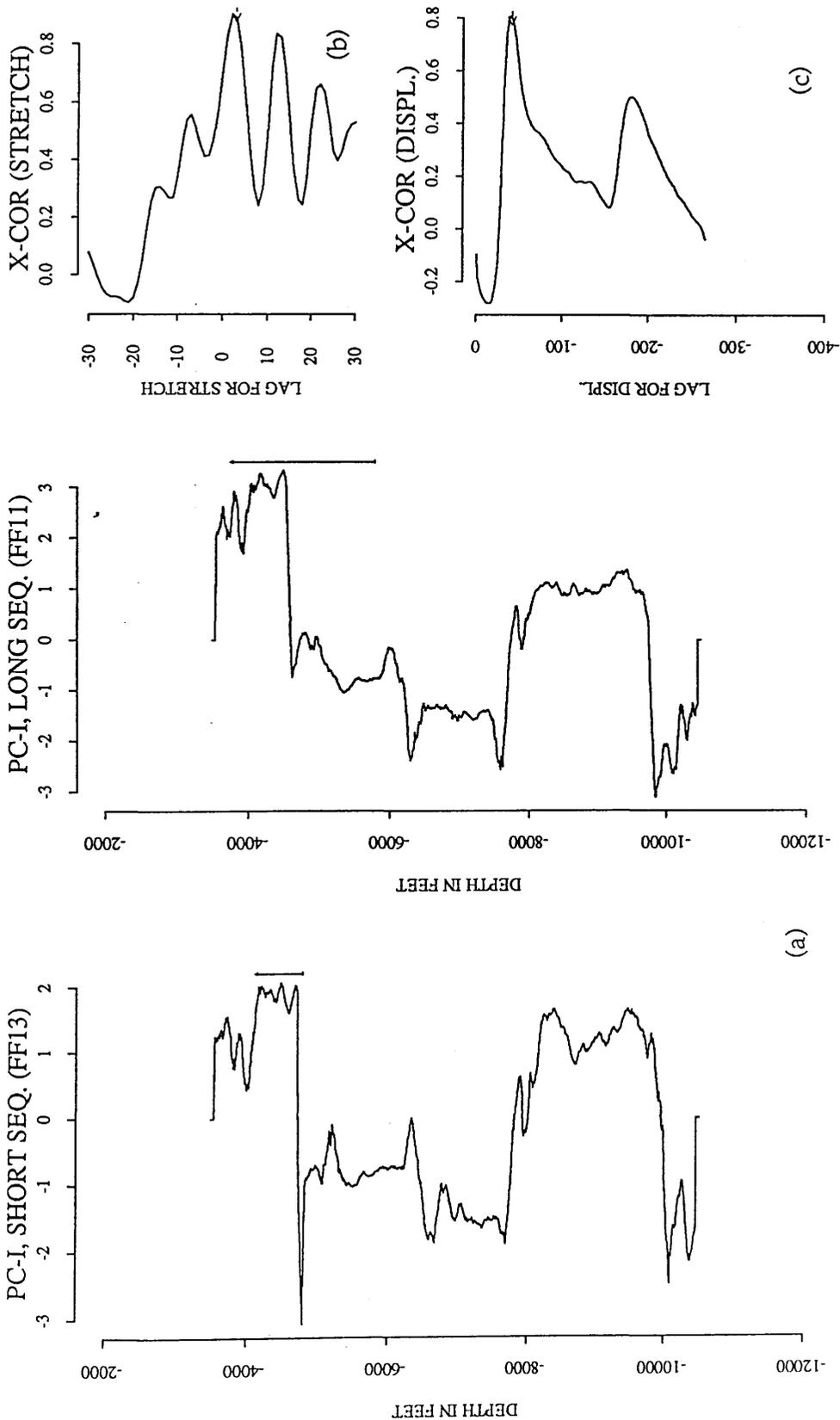


Fig. 4.22 Cross-correlation of the Etel Formation using the derivative data. (a) the first principal components of FF13 and FF11. (b) the cross-correlation function of power spectra ($S=1.05$). (c) the cross-correlation function yields a maximum of 0.808 when the short sequence is stretched 1.05 times at a lag of 30 units (150 feet).

displacement $D=30$ units (150 feet compared with 110 feet). A spurious peak for the cross-correlation function of power spectra is obtained when the non-filtered principal components are used (Fig. 4.23). The cross-correlation function of power spectra yields a sharp peak at a lag $-v=-10$ ($S=1.25$) for short series (sequence) (Fig. 4.23b), and a maximum peak for displacement $D=210$ feet (Fig. 4.23c). Furthermore, an attempt to correlate the original variables (Gamma Ray) of well FF13 and FF11 failed (Figs. 4.24, 4.25) producing a very high stretch factor ($S=1.80$) for long sequence and cross-correlation function for displacement $D=230$ units (1150 feet).

The above results emphasize that the use of the smoothed principal components has the advantage in more accurately predicting the stretch factor over both the non-filtered components and the original variables of well-logs. In addition, the correlation functions of power spectra of the filtered principal component (Fig. 4.22b) against that of the non-filtered (Fig. 4.23b) and that of the original data (Fig. 4.24b), and the correlation functions of the stretched series in Figure 4.22c against that of Figure 4.23c and that of Figure 4.24c are more smooth and symmetric.

4.3.3 (b) Correlation of the Sheghega Formation

The computer successfully correlates the Sheghega Formation of well FF13 (indicated by number 2 in Fig. 4.20a) with the section indicated by the letter B in Figure 4.20b. The high peak of the cross-correlation function of power spectra is shown in Figure 4.26f and Figure 4.27b at a lag of $-v=-2$ ($S=1.05$ compared with geological stretch of 1.05) for the long sequence

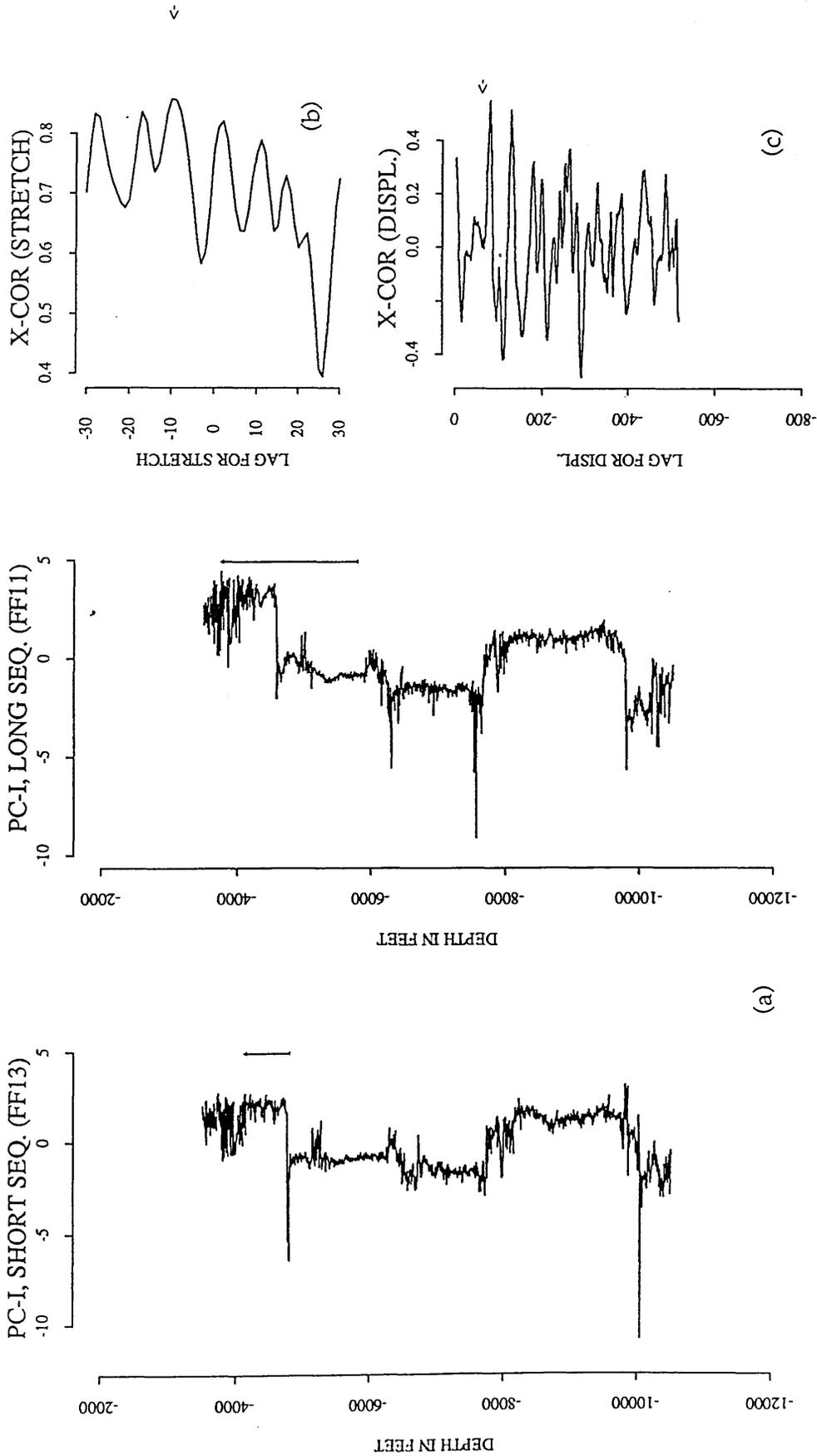


Fig. 4.23 Cross-correlation of the Eitel Formation using the non-filtered principal components. (a) the first principal components of FF13 and FF11. (b) the cross-correlation function of power spectra at a lag of 10 ($S=1.25$). (c) the cross-correlation function (0.565) for displacement of 42 units (210 feet) when the long sequence is stretched.

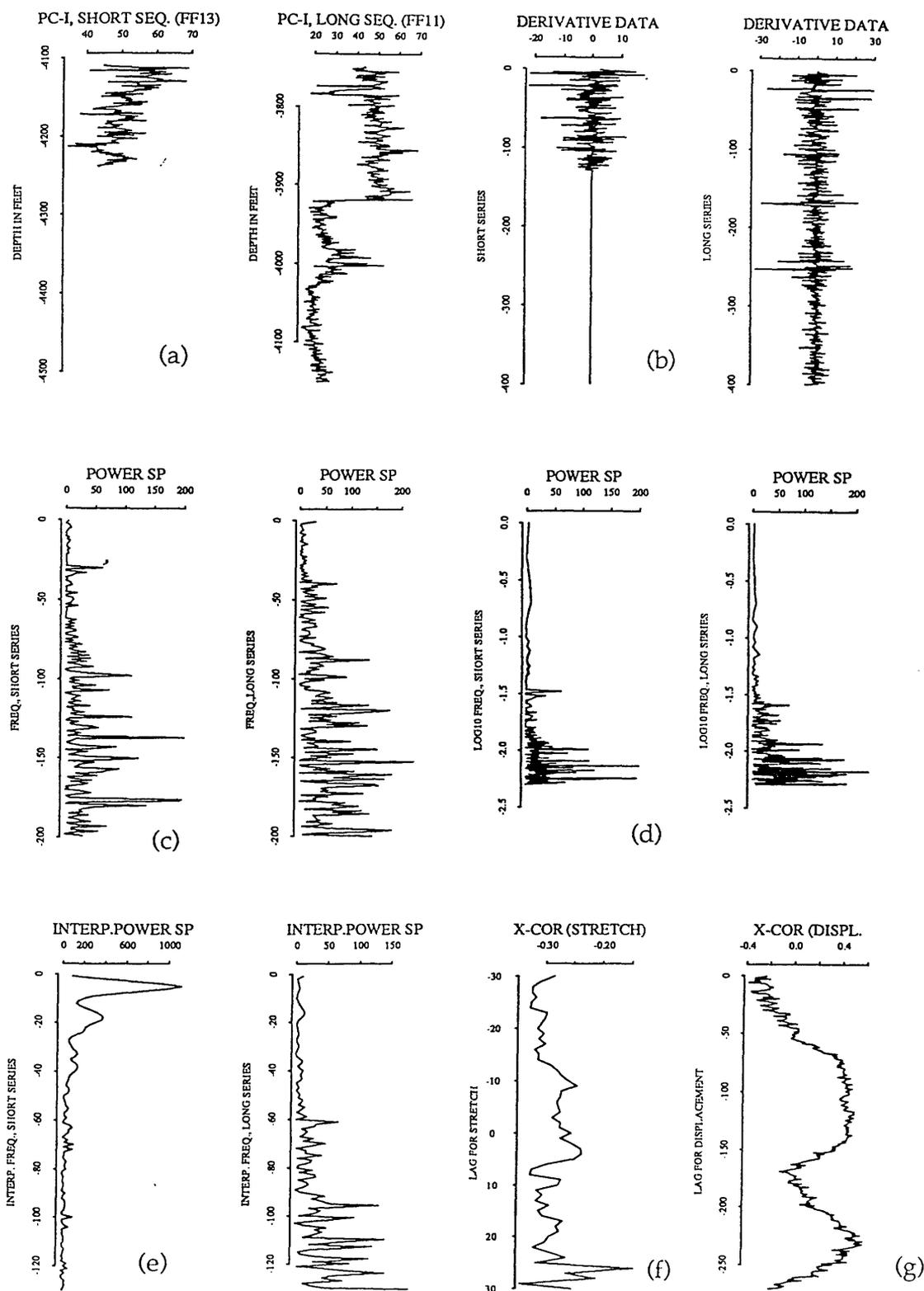


Fig. 4.24 plot of cross-correlation using the original data (Gamma Ray) of the Etel Formation in FF13 and FF11. (a) GR of FF13 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the stretched series. (g) is the cross-correlation function of the displacement series.

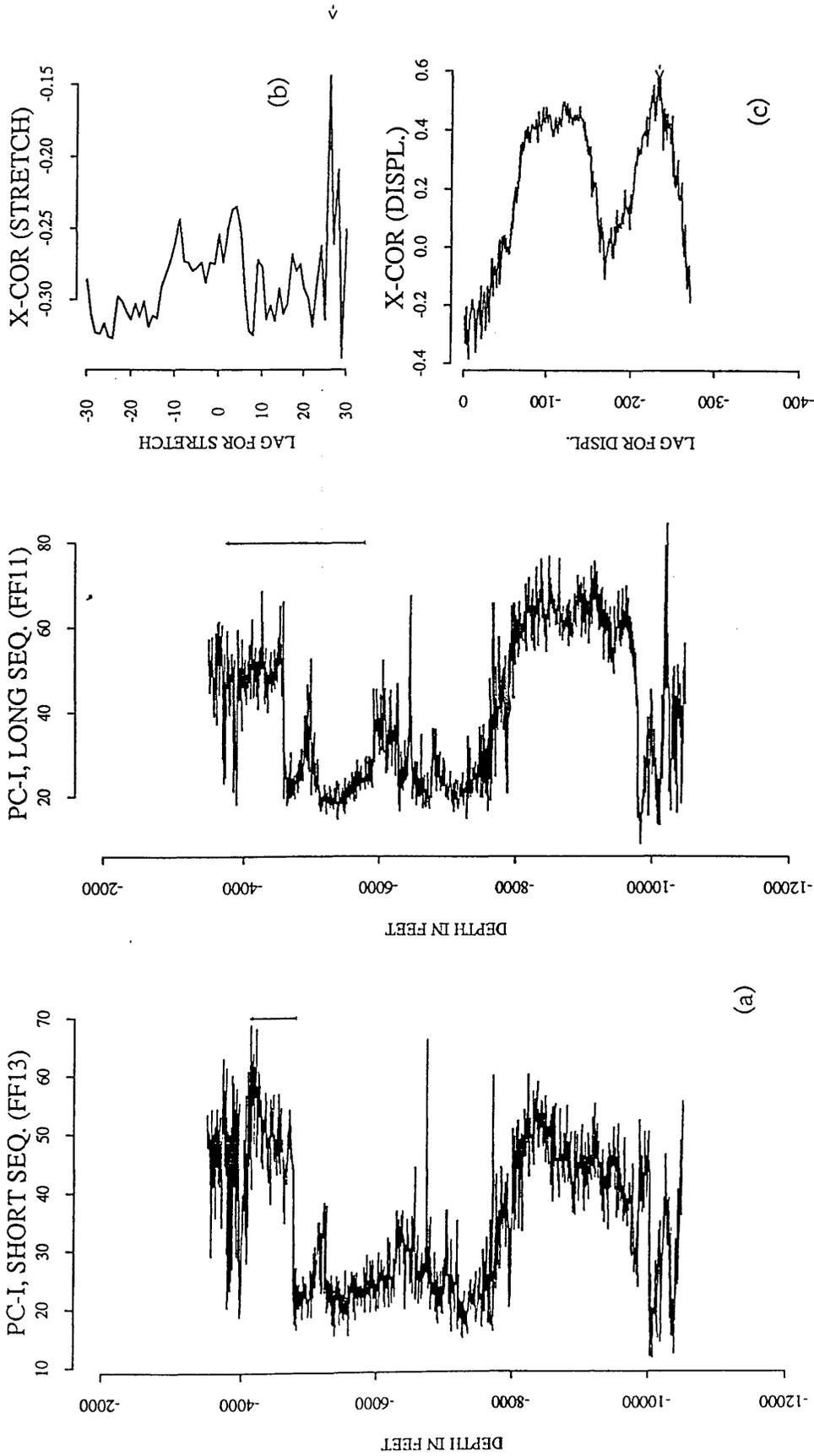


Fig. 4.25 Cross-correlation of the Etel Formation using the original variables (GR). (a) GR of FF13 and FF11 (b) the cross-correlation function of power spectra ($S=1.80$). (c) the cross-correlation function (0.580) for displacement of 230 units (1150 feet). The long sequence is stretched 1.80 times.

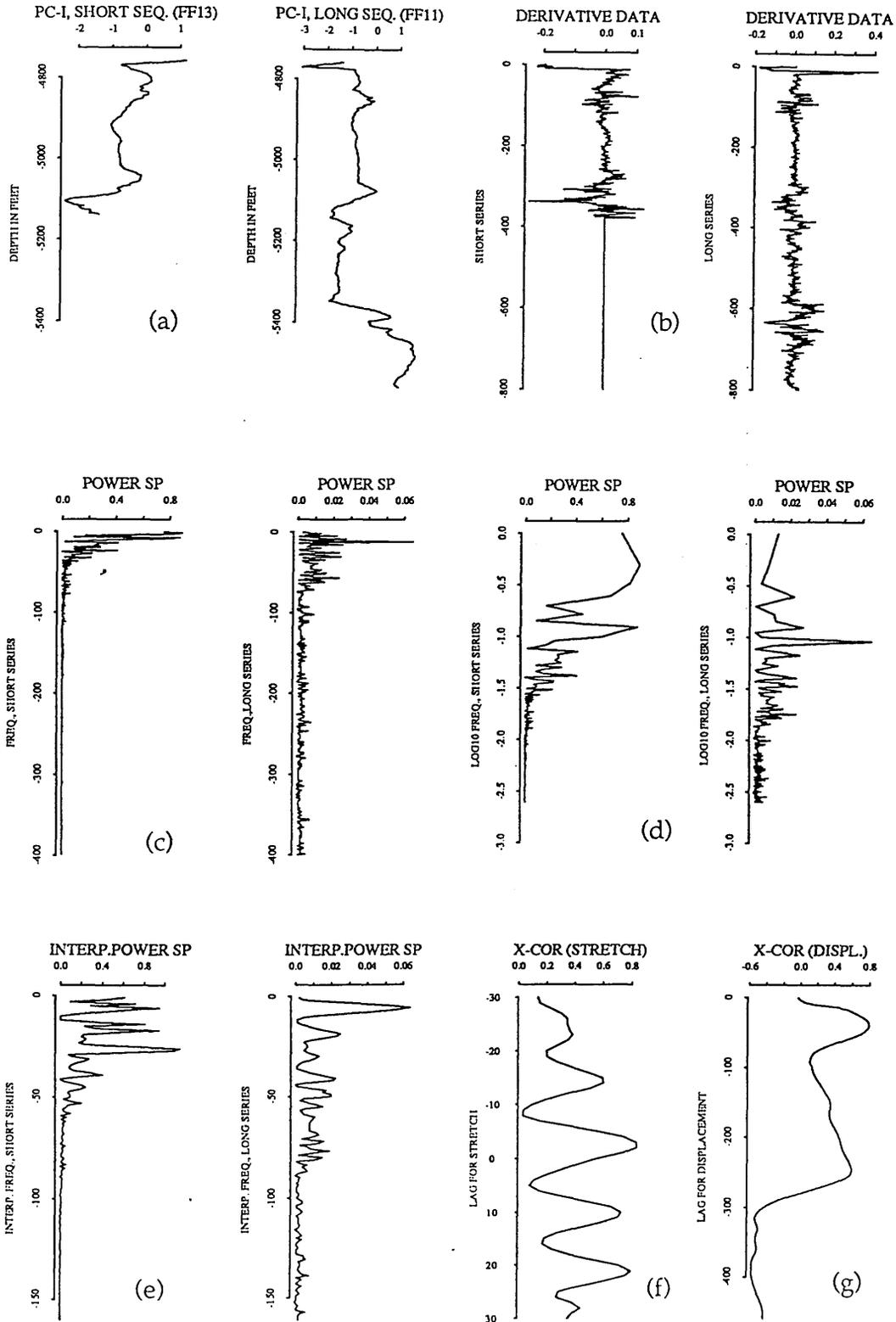


Fig. 4.26 plot of cross-correlation of the Sheghega Formation in FF13 and FF11 using the derivative data. (a) principal components of FF13 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

(series) (Fig. 4.27b), this implies that the direction of the thickening of the strata is towards FF11. The high magnitude (0.799) of the cross-correlation function of the stretched series (Fig. 4.26g and Fig. 4.27c) is corresponding to computer stretch of 39 units (195 feet) compared with the known geological displacement 185 feet.

4.3.3 (c) Correlation of the Domran Formation

Figure 4.28 shows the result of correlating of the Domran Formation (window 3 in Fig. 4.20a) in well FF13 with a portion of long series of well FF11 (window C in Fig. 4.20b). Like the correlation of the Sheghega Formation, the computer selection of the two correlation functions (Fig. 4.28g and Fig. 4.28f) agrees with the known geological correlation. In Figure 4.29b the maximum correlation function has a peak for a stretch factor 1.23 compared with 1.26 at a lag $+v=9$, and a maximum peak (0.645) at displacement of 59 units (295 feet compared with 310 feet) (Fig. 4.29c).

4.3.3 (d) Correlation of the Ruaga Formation

The mathematical correlation of the Ruaga Formation of well FF13 (number 4 in Fig. 4.20a) against a section of well FF11 (letter D in Fig. 4.20b) is displayed in Figure 4.30 and Fig. 4.31. Two comparable peaks of power spectra are observed (Fig. 4.30g and Fig. 4.31b). The top peak corresponds to a stretch of 1.44 at a lag $-v=-16$. The computer selection of the second peak at a lag $-v=-4$ ($S=1.10$) is geologically reasonable compare to geological stretch of 1.20, and displacement of 125 feet compared with 112 feet. The negative sign

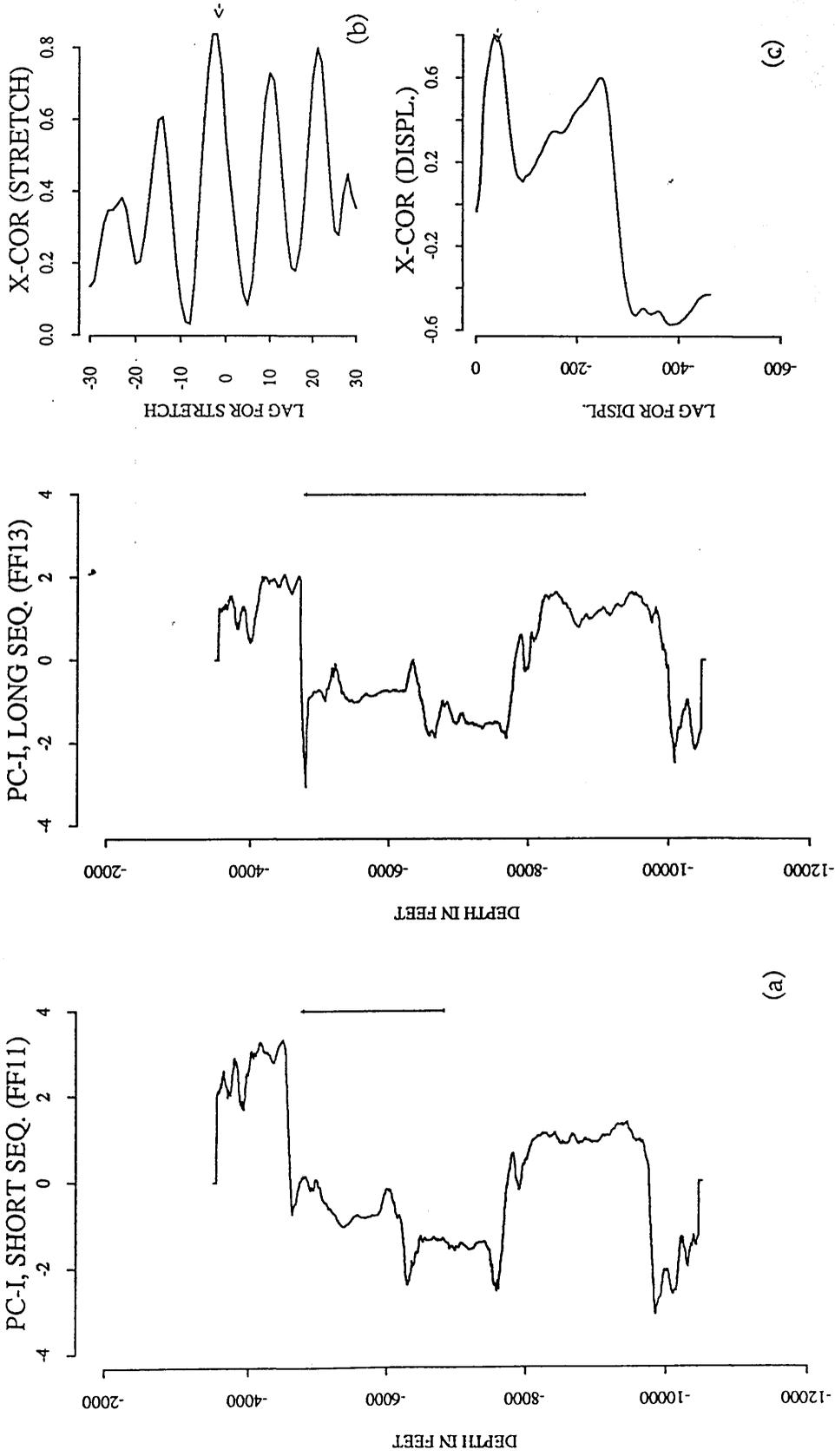


Fig. 4.27 Cross-correlation of the Sheghega Formation. (a) principal components of FF13 and FF11. (b) the cross-correlation function of power spectra at a lag of $-v=-2$ ($S=1.05$). (c) the cross-correlation function (0.799) for displacement of 39 units (195 feet) when short sequence is stretched 1.05 times.

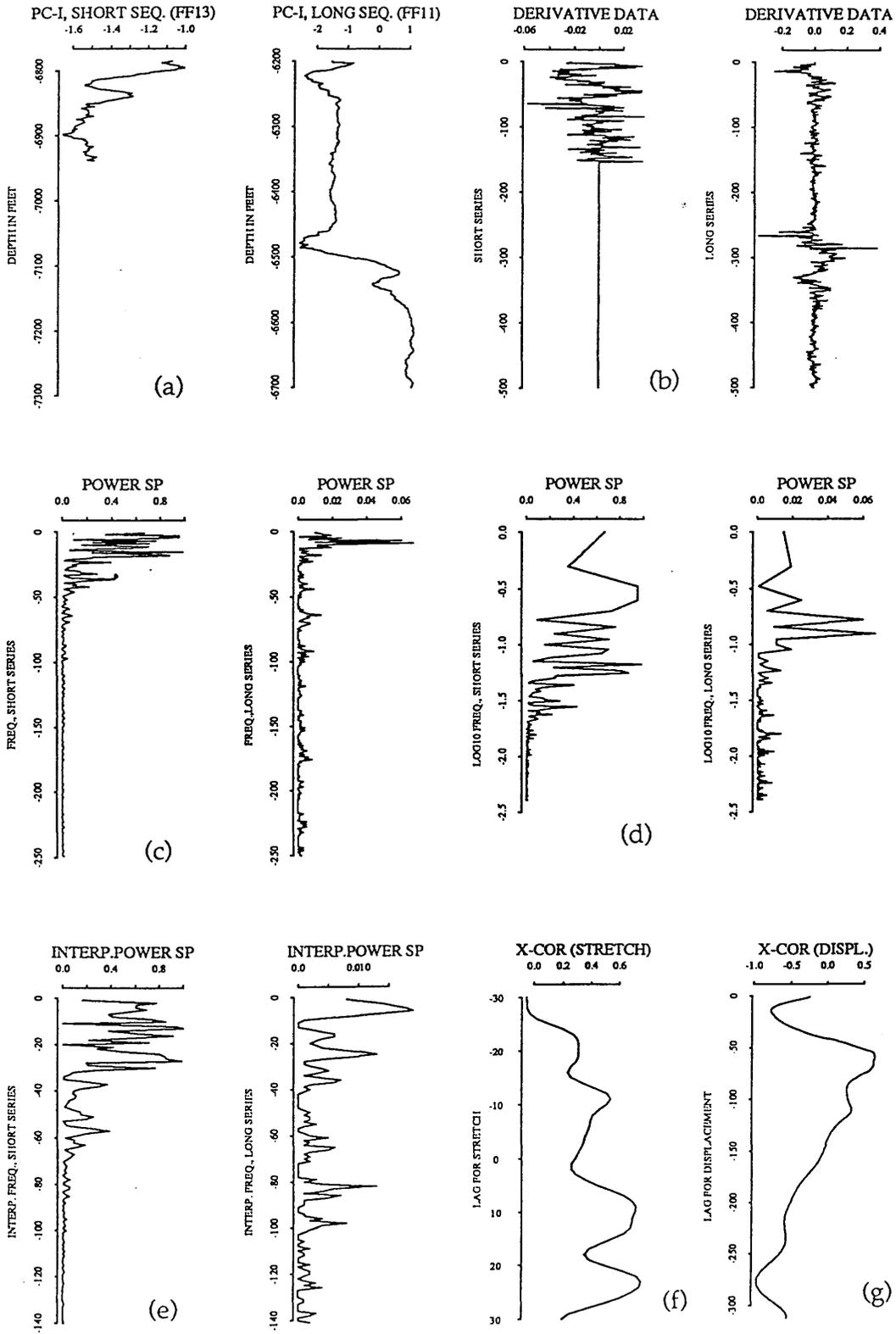


Fig. 4.28 plot of cross-correlation of the Domran Formation in FF13 and FF11 using the derivative data. (a) principal components of FF13 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

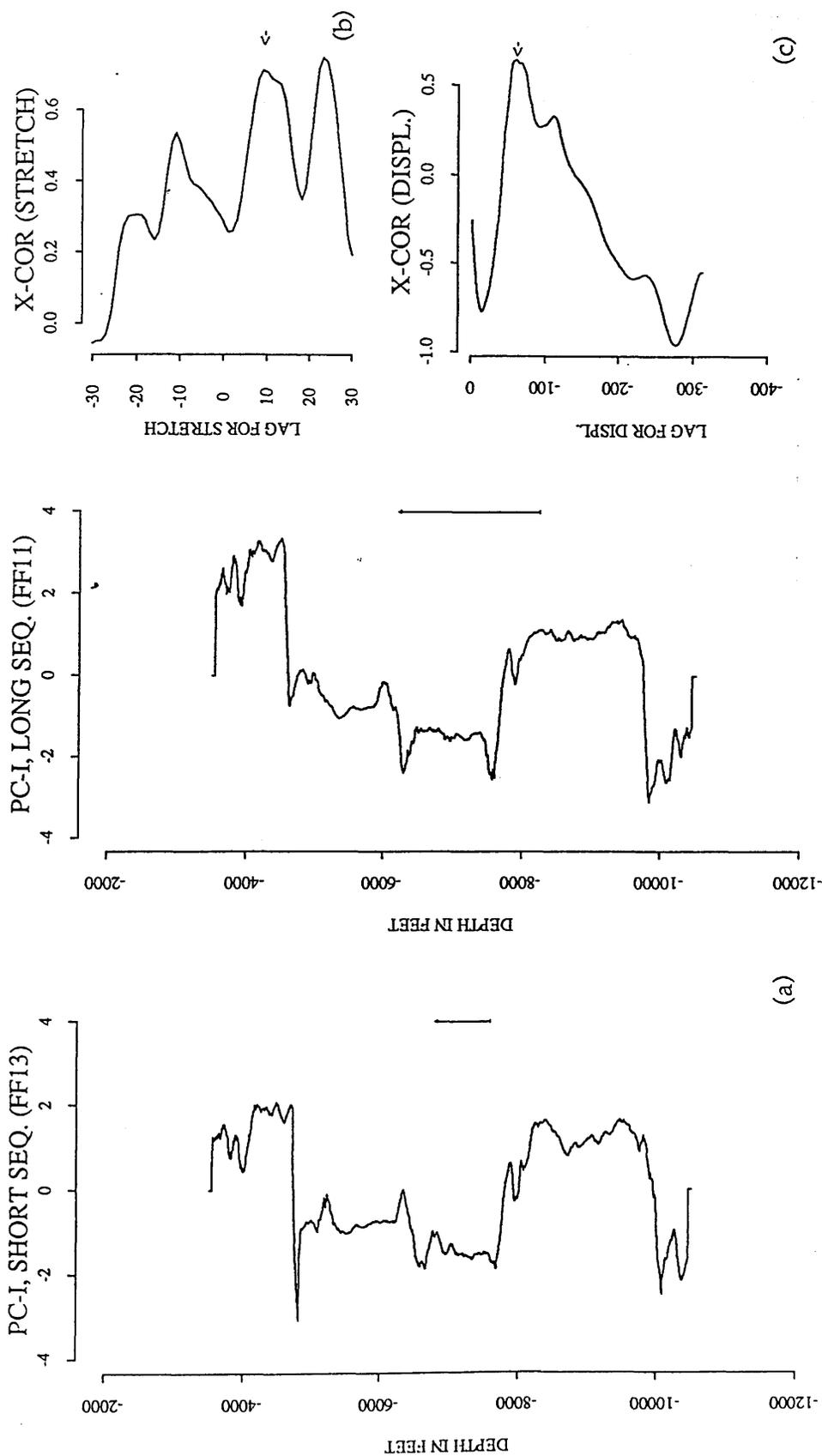


Fig. 4.29 Cross-correlation of the Domran Formation. (a) principal components of FF13 and FF11. (b) the cross-correlation function of power spectra at a lag of $v=9$ ($S=1.23$). (c) the cross-correlation function (0.645) for displacement of 59 units (295 feet) when long sequence is stretched 1.23 times.

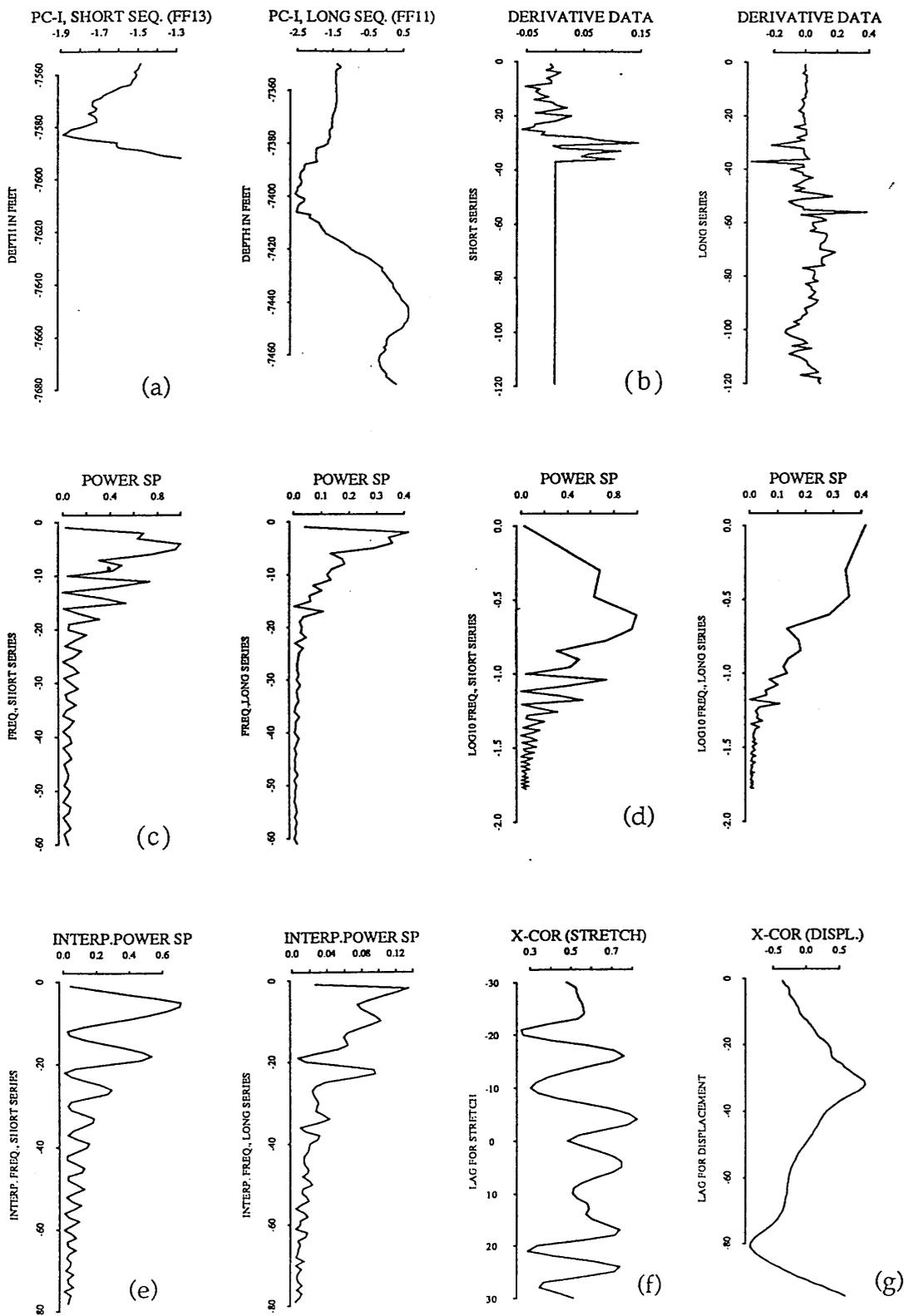


Fig. 4.30 plot of cross-correlation of the Ruaga Formation in FF13 and FF11 using the derivative data. (a) principal components of FF13 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the stretched series. (g) is the cross-correlation function of the power spectra.

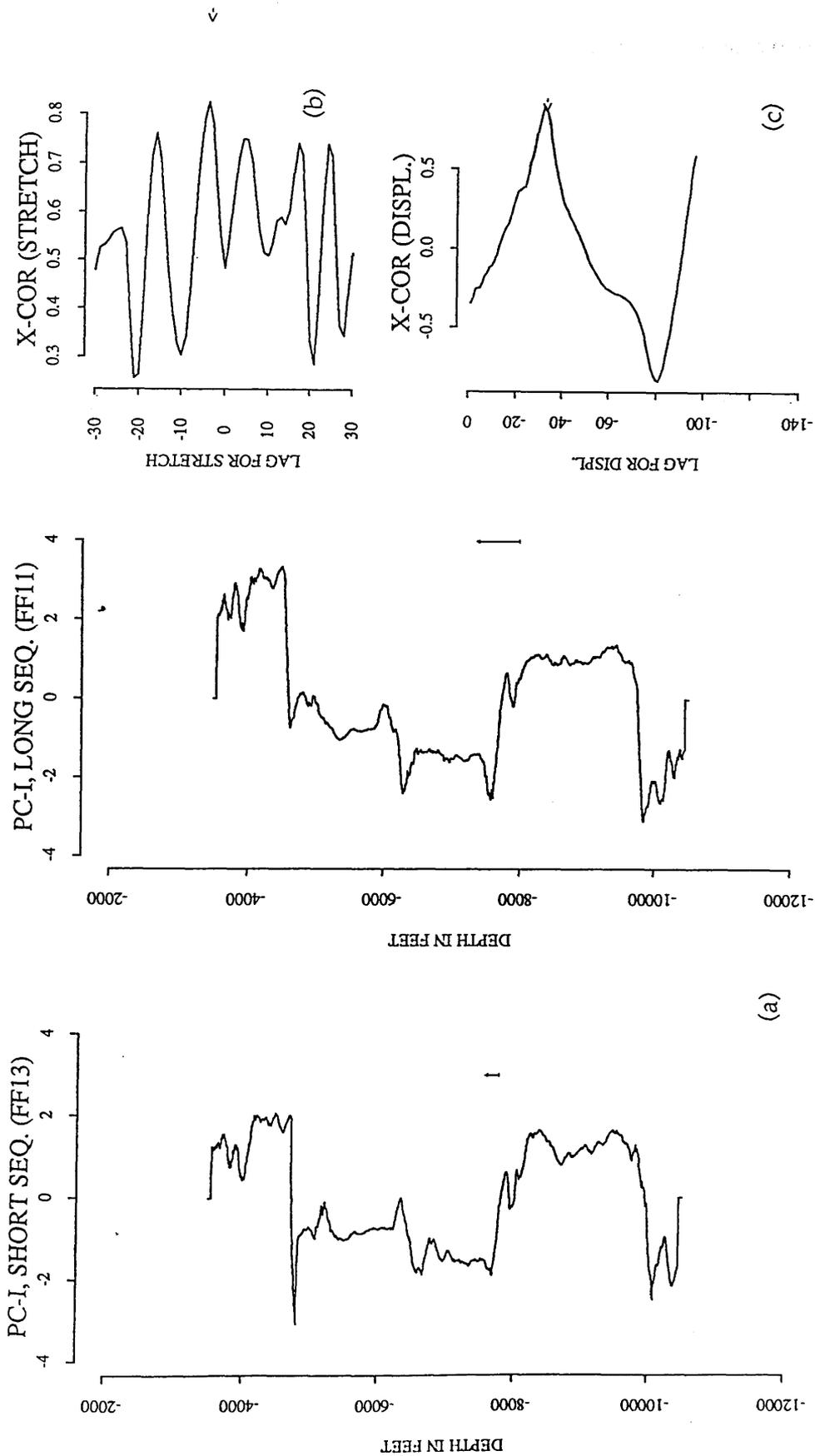


Fig. 4.31 Cross-correlation of the Ruuga Formation. (a) principal components of FF13 and FF11. (b) the cross-correlation function of power spectra at a lag of $-v=-4$ ($S=1.10$). (c) the cross-correlation function (0.893) for a lag of 125 feet when the short sequence is stretched.

of v indicates that the short series (FF13) is a stretched version of the the long series (FF11).

4.3.3 (e) Correlation of the Heira Formation

Figure 4.32 shows the cross-correlation between the Heira Formation in well FF13 (number 5 in Fig. 4.20a) and a window of length E in Figure 4.20b. A high magnitude (0.730) of the cross-correlation function is obtained (Fig. 4.32f and Fig. 4.33c) when correlating the thick sections of the Heira Formation (about 2200 feet). Although the maximum cross-correlation function of the power spectra yields large value of stretch 1.26 compared with 1.08 for the long series (FF11), the computer selection of both the direction of thickening and displacement (Fig. 4.33b and Fig. 4.33c) is satisfactory. An improvement in the stretch factor value (1.17) is obtained when the non-filtered principal components are used in the analysis (Fig.4.34), and the maximum function for correlating stretched series (0.417) and stretch of 1.10 are obtained when the derivative data is used for stretching (Fig. 4.35). However, the direction of the variation in thickness was spurious one and does not agree with the known geological thickening which is from well FF11 towards FF13.

4.3.4 Correlation between FF11 and FF12

Well F11 and well FF12 are located on the eastern flank of the Attahaddy field (Fig. 4.7). The distance between the two boreholes is about 4 Km and a sequence of 7000 feet [3500-10500 feet] is used in the analysis. Well

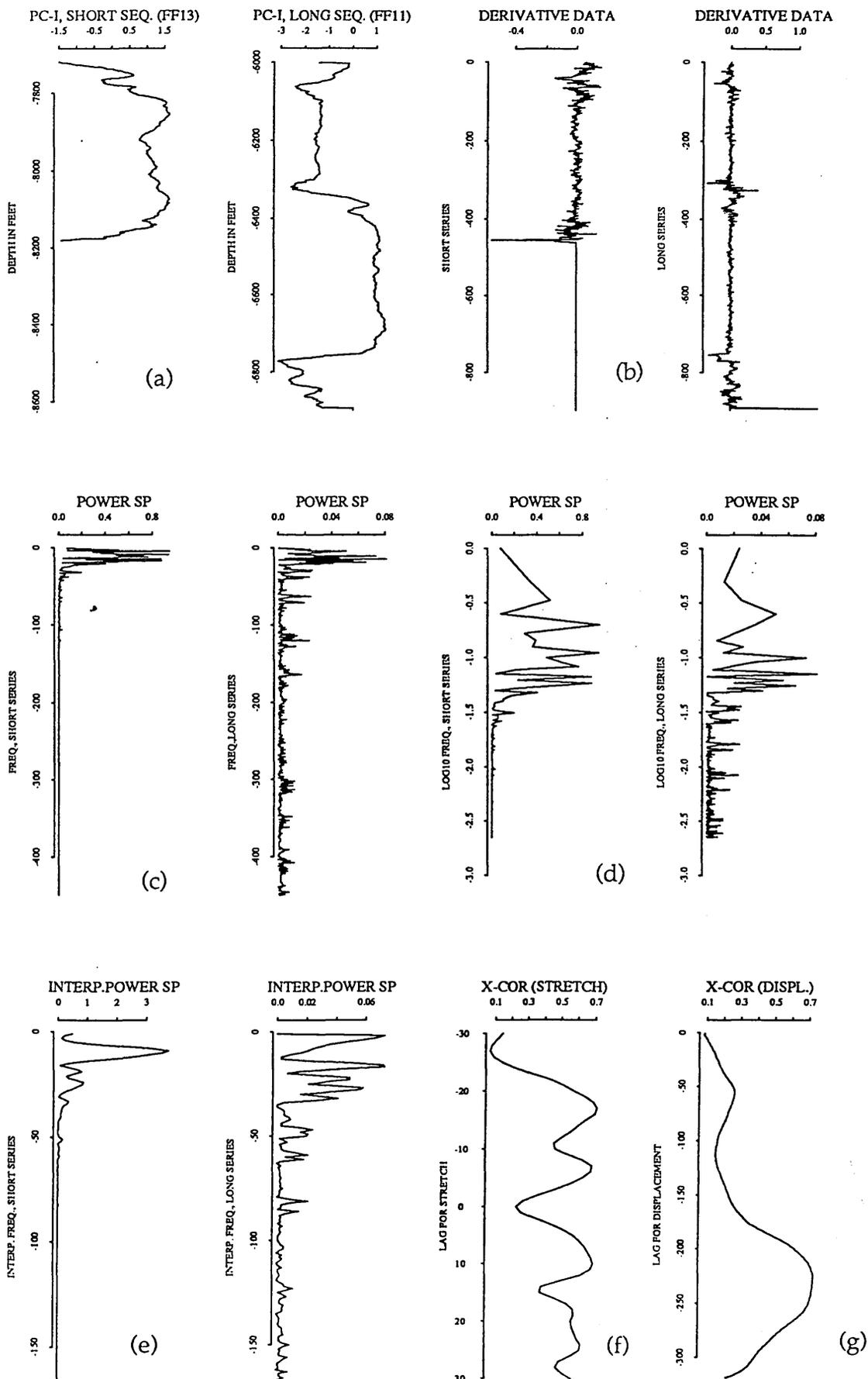


Fig. 4.32 plot of cross-correlation of the Heira Formation in FF13 and FF11 using the derivative data. (a) principal components of FF13 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

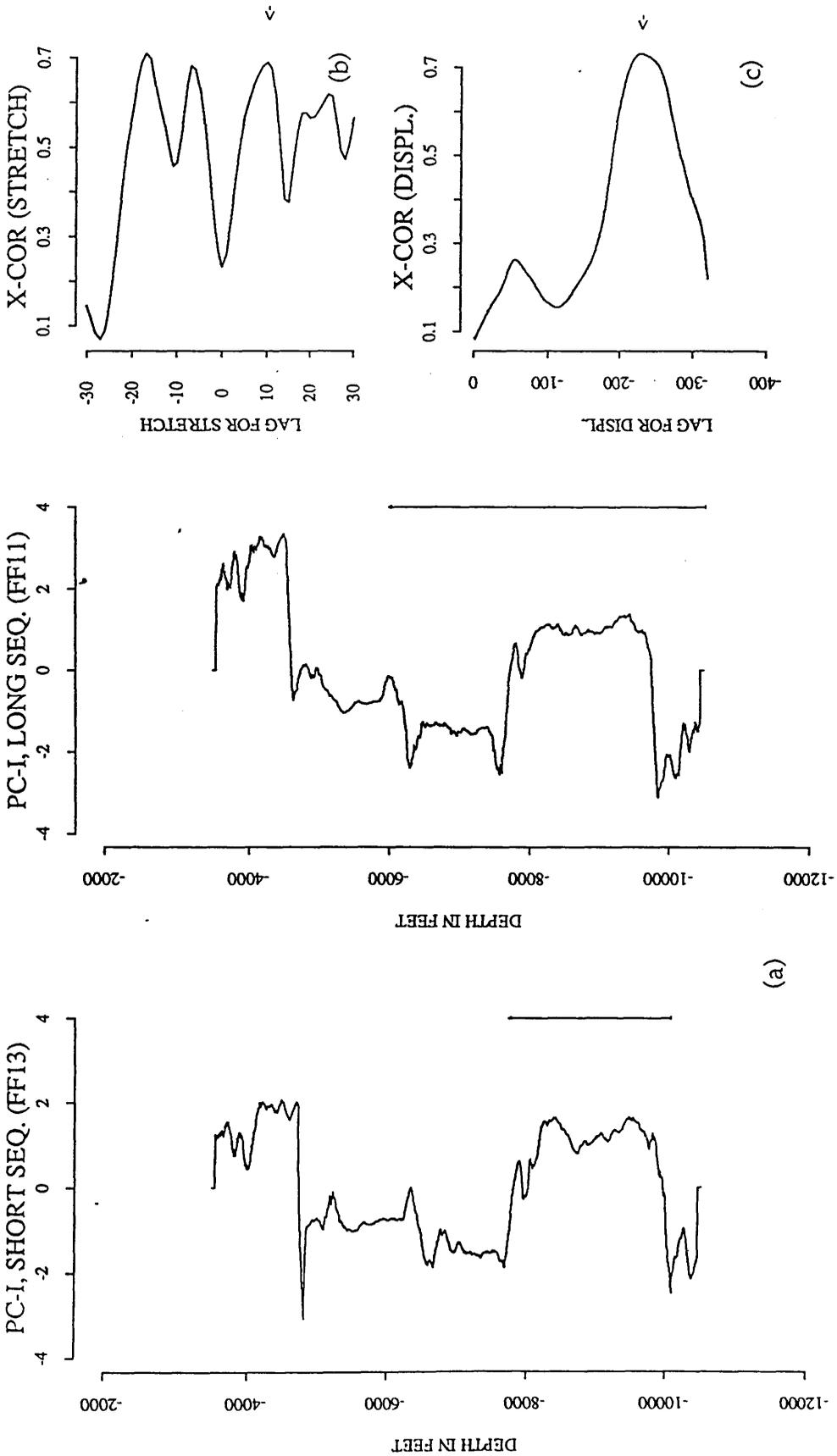


Fig. 4.33 Cross-correlation of the Heira Formation using the derivative data. (a) principal components of FF13 and FF11. (b) the cross-correlation function of power spectra ($S=1.26$). (c) the cross-correlation function (0.730) for a lag of 227 units when the long sequence is stretched 1.26 times.

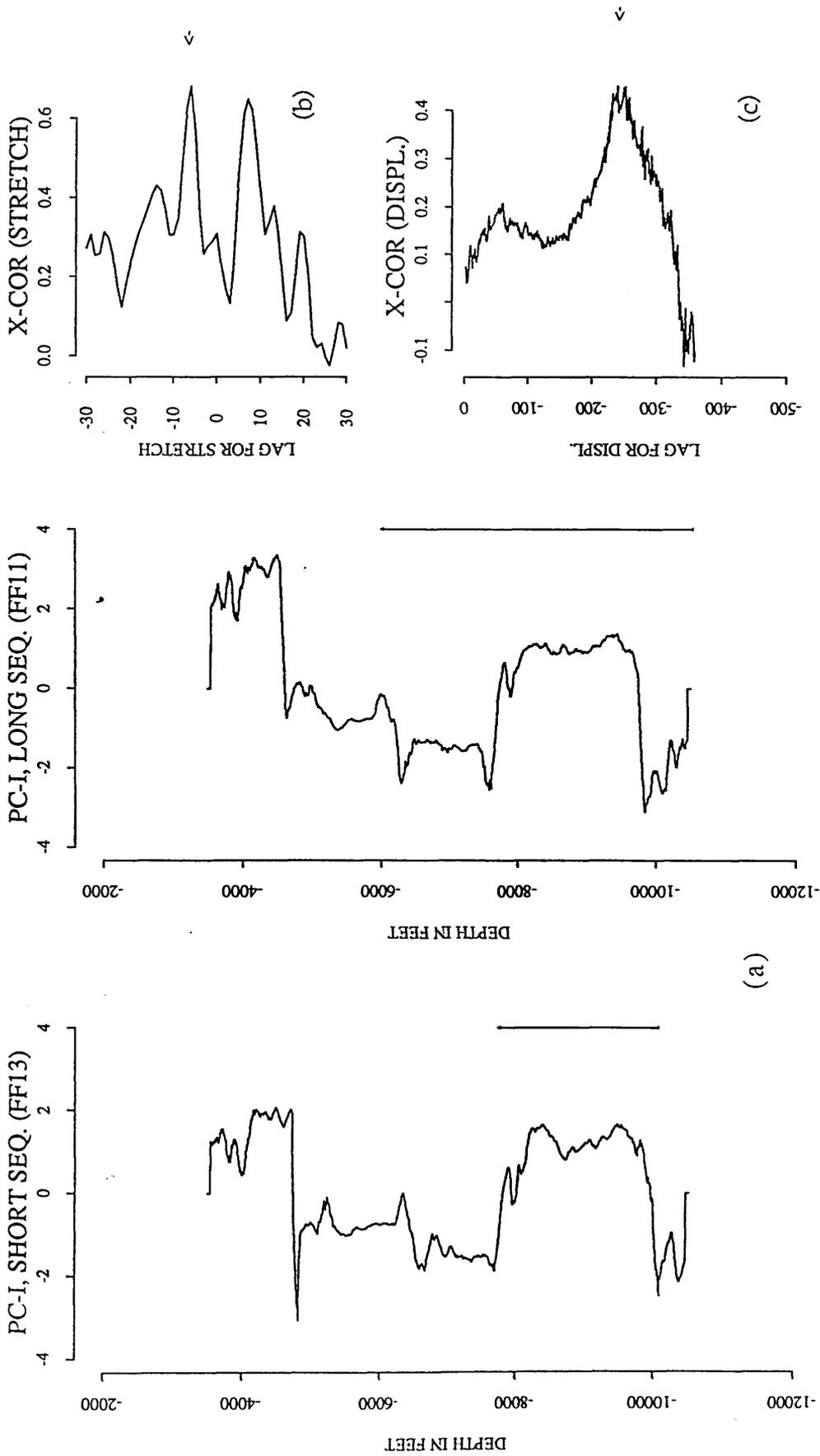


Fig. 4.34 Cross-correlation of the Heira Formation using the original data (principal components). (a) principal components of FF13 and FF11. (b) the cross-correlation function of power spectra ($S=1.17$). (c) the cross-correlation function (0.417) for lag of 238 units (1190 feet) when the short sequence is stretched 1.17 times.

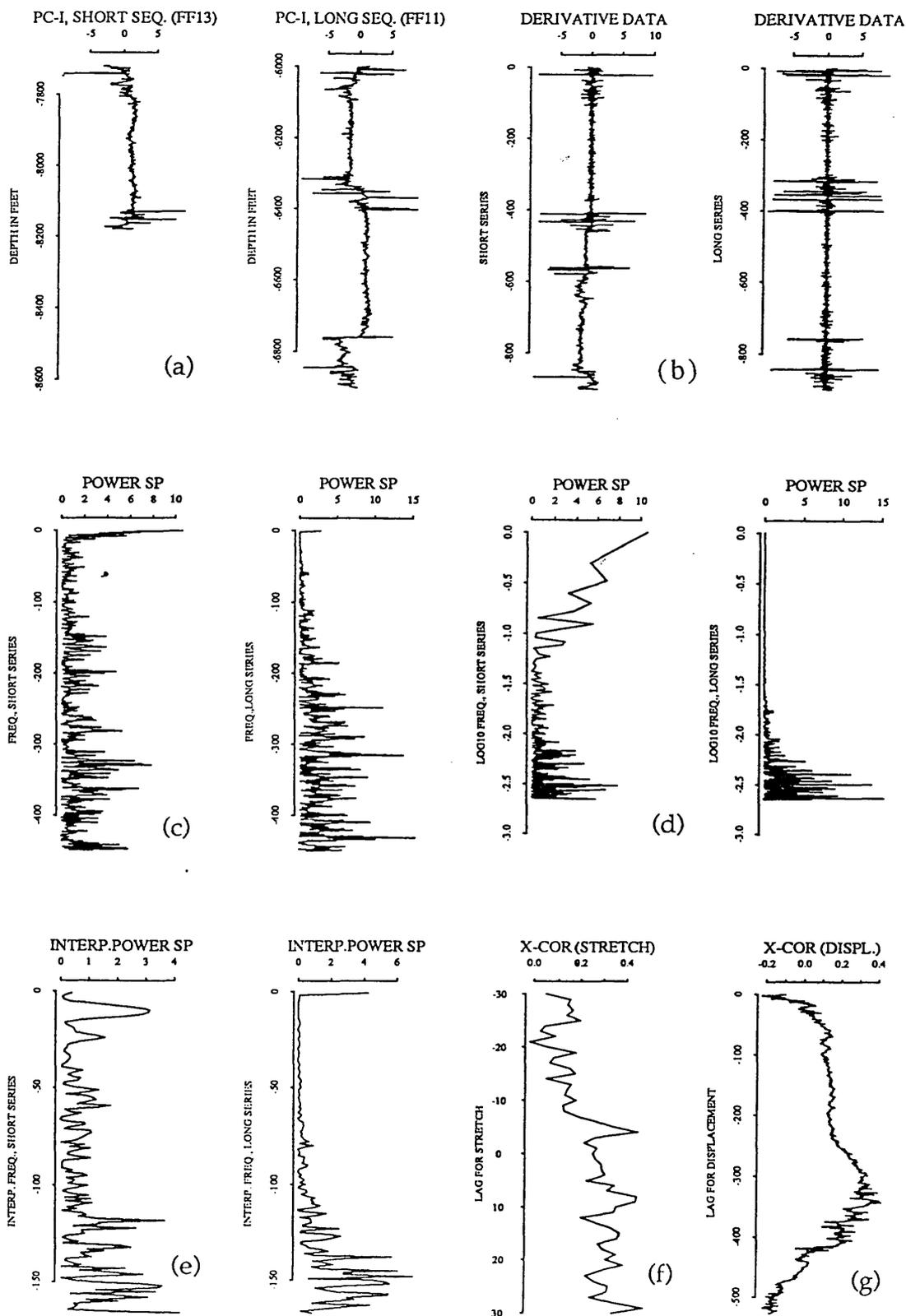


Fig. 4.35 plot of cross-correlation of the Heira Formation in FF13 and FF11 using the derivative data of the non-filtered principal components. (a) non-filtered principal components of FF13 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

FF11 is a gas well producing from the Gargaf Formation and well FF12 lies outside the gas-water contact of the eastern part of the Attahaddy field and is a dry hole.

The eigenvalues, eigenvectors and the percentage of the eigenvalues to the total variance of the correlation matrix of well FF12 are shown in Table 4 Appendix E, and its first principal component is displayed (Fig. 4.36a). The first principal component of well FF11 is shown in Figure 4.19a, and its eigenvalues and eigenvectors are shown in Table 3 Appendix E.

Different Formation boundaries of well FF12 are identified using the boundary identification technique (Fig. 4.36b, Table 4.5). The Sheghega Formation is identified at depth of 4738 feet, the Domran Formation at depth of 6750 feet, the Ruaga Formation is at depth of 7450 feet, the Heira Formation is identified at a depth of 7850 feet, and the Zmam Formation at depth of 10020 feet. The cross-correlation between the two smoothed principal component of well FF11 and FF12 (Fig. 4.37) are then performed.

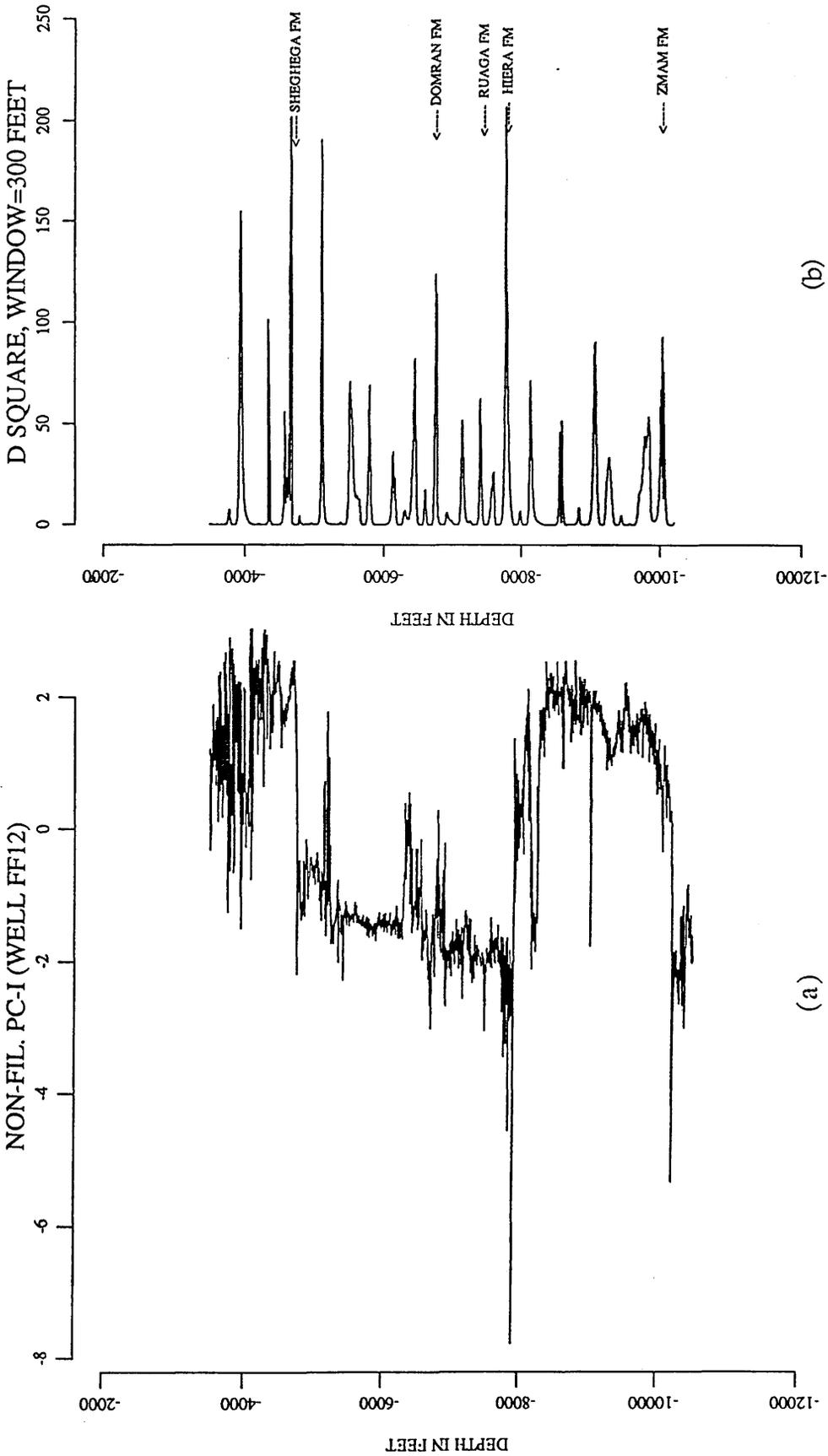
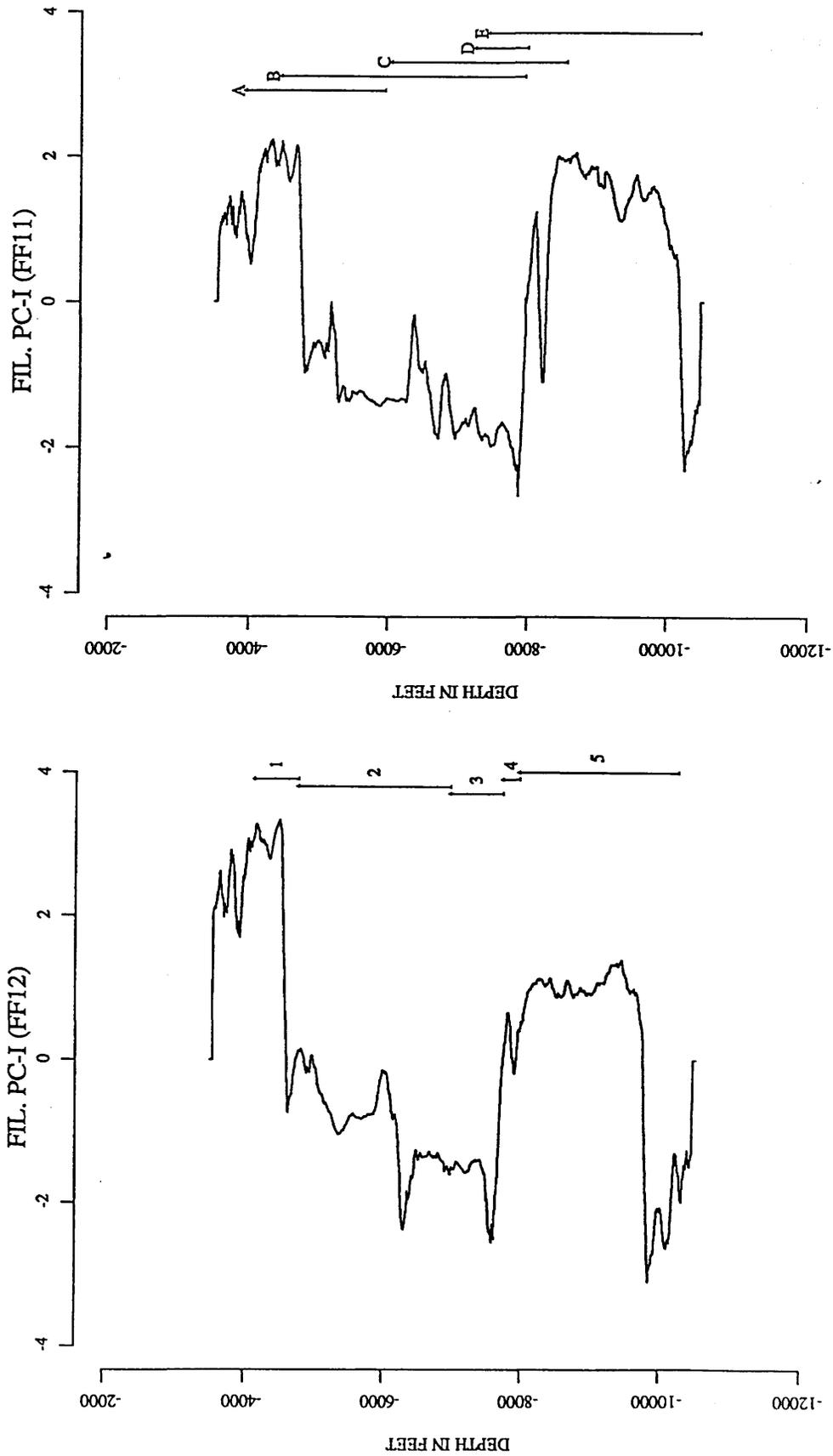


Fig. 4.36 Plot of the first principal components (sampled at 5 feet) of well FF12 and its boundaries. (a) non-smoothed principal components of FF12. (b) boundaries of different formations (window=300 feet).



(b)

(a)

Fig. 4.37 Plot showing the smoothed principal component of the correlated sequences (FF12 and FF11). (a) smoothed component of well FF12. (b) smoothed component of well FF11. A window (marked 1) in (a) is correlated with a window (marked A) in (b). The window (marked 2) in (a) is compared with the window (marked B) in (b), and so forth.

Table 4.5 The Geological Formation depths and the predicted Formation depths of FF 12 using boundary identification technique.

Formation	Geological depth	Predicted depth
Sheghega	4738	4738
Domran	6925	6750
Ruaga	7691	7450
Heira	7924	7850
Zmam	10199	10020

4.3.4 (a) Correlation of the Etel Formation

The Etel Formation of well FF12 (indicated by number 1 in Fig. 4.37a) is compared with a section from well FF11 with a window indicated by the letter A (Fig.4.37b). A maximum of cross-correlation function of power spectra is observed at a lag $v=0$, no stretch compared with 1.10 for the long sequence (FF11) (Fig. 4.38g and 4.39b). The computer correlation of the stretched sequence, with a correlation maximum of 0.71 giving a displacement of 75 feet (Fig. 4.38f and 4.39c) agrees with the known geological correlation of the Etel Formation between well FF12 and FF11.

4.3.4 (b) Correlation of the Sheghega Formation

Cross-correlation of the Sheghega Formation is made between FF12 (number 2 in Fig. 4.37a) and a portion of the principal component of well

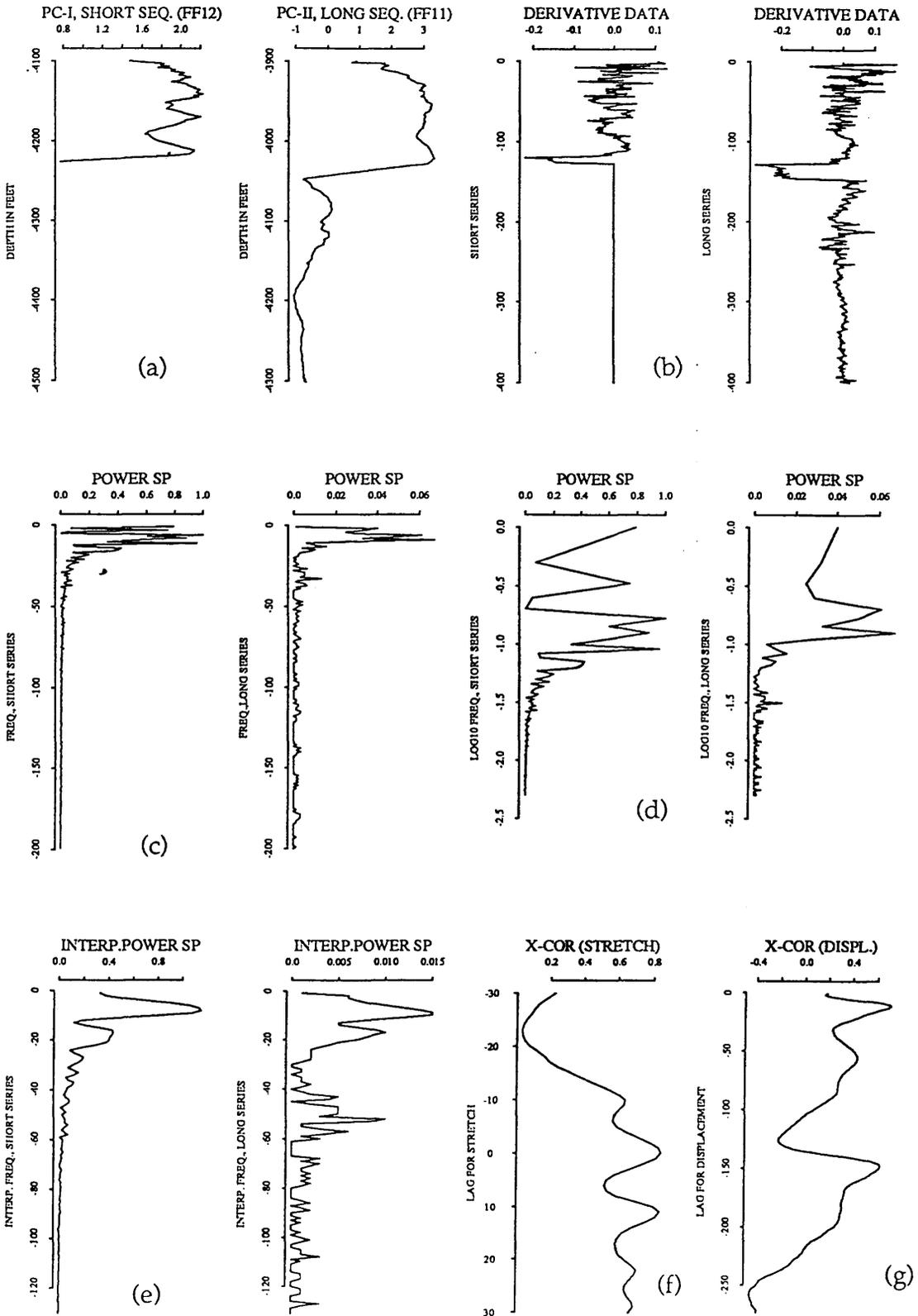


Fig. 4.38 plot of cross-correlation of the Etel Formation in FF12 and FF11 using the derivative data. (a) principal components of FF12 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

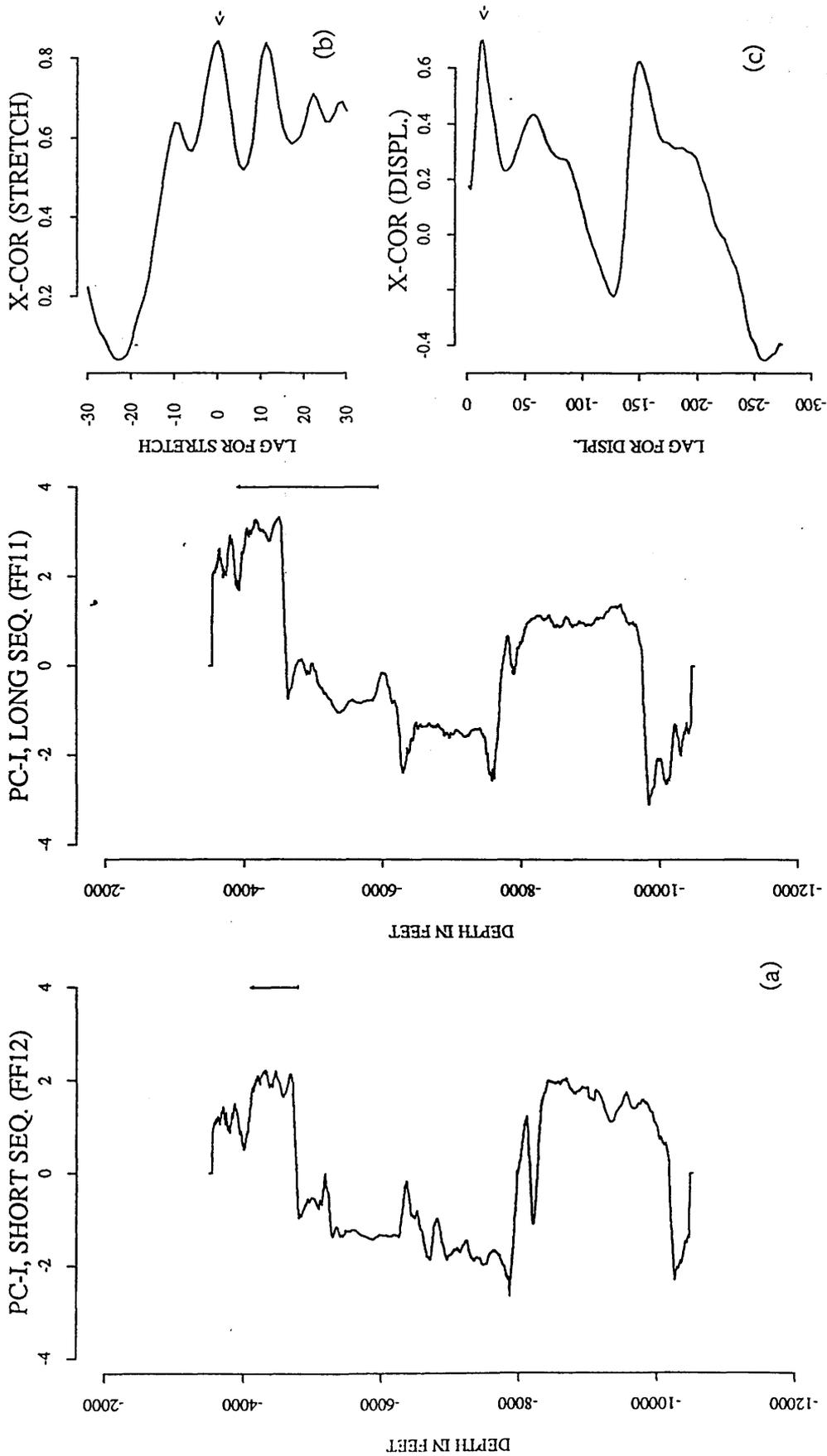


Fig. 4.39 Cross-correlation of the Etel Formation. (a) principal components of FF12 and FF11. (b) the cross-correlation function of power spectra at lag=0 ($S=1$, no stretch). (c) the cross-correlation function for displacement (0.701) at a lag of 15 units (75 feet) without stretching.

FF11 with a window length marked by the letter B using the original data for stretching (Fig. 4.37b). A successful cross-correlation is observed (Fig. 4.40). The maximum value of the cross-correlation function of the power spectra at a lag of 10 (Fig. 4.40b), gives a stretch factor of 1.26 compared with 1.16 which is geologically reasonable. The direction of the thickening of strata is from well FF12 towards FF11 which agrees with the known thickening. Furthermore, the selection of the maximum peak of the cross-correlation function of the stretched series (0.577) (Fig. 4.40c) at displacement of 24 units (140 feet) compared with 32 units (160 feet) confirms the reliability of the correlation.

4.3.4 (c) Correlation of the Domran Formation

The Domran Formation in well FF12 (Fig.4.37a, number 3) is compared with FF11 (Fig.4.37b, symbol C) using both the derivative data and the original data (Fig. 4.41 & Fig. 4.43) respectively. In Figure 4.41g and 4.42b, a stretch factor of 1.32 is predicted compared with 1.27 is observed for long sequence and displacement of 440 feet compared with 452 feet is observed, however, the direction of the thickening of the Domran Formation is towards FF11. When the original principal components data are used for stretching (Fig. 4.43), the direction of thickening of the two formations agrees with the geological thickening but a bigger value of stretch factor (1.38) is obtained (Fig. 4.43b).

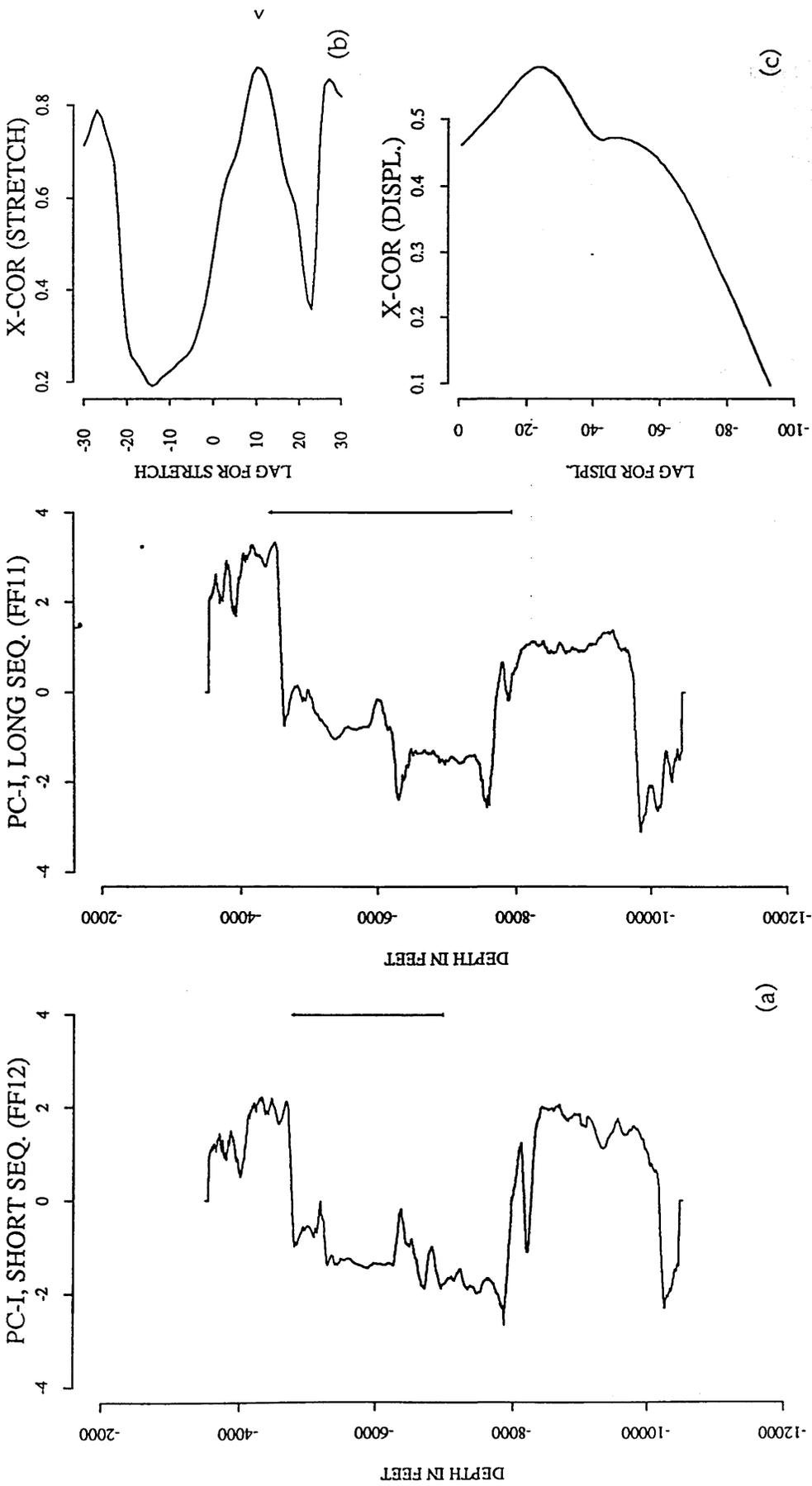


Fig. 4.40 Cross-correlation of the Sheghega Formation. (a) principal components of FF12 and FF11. (b) the cross-correlation function of power spectra at lag $v=10$ ($S=1.26$). (c) the cross-correlation function for displacement (0.577) at a lag of 28 units (140 feet) when the long sequence is stretched 1.26 times.

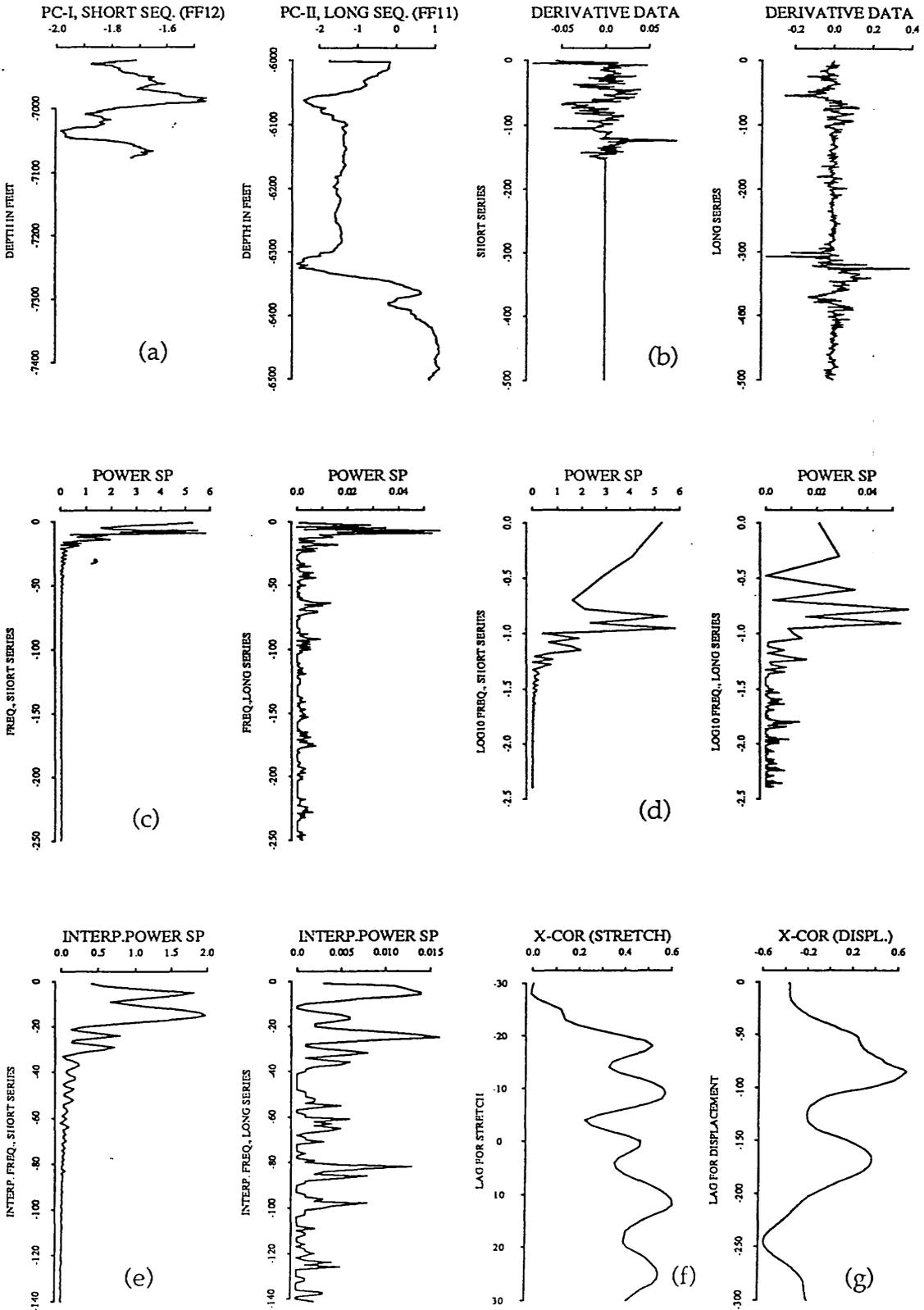


Fig. 4.41 plot of cross-correlation of the Domran Formation in FF12 and FF11 using the derivative data. (a) principal components of FF12 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

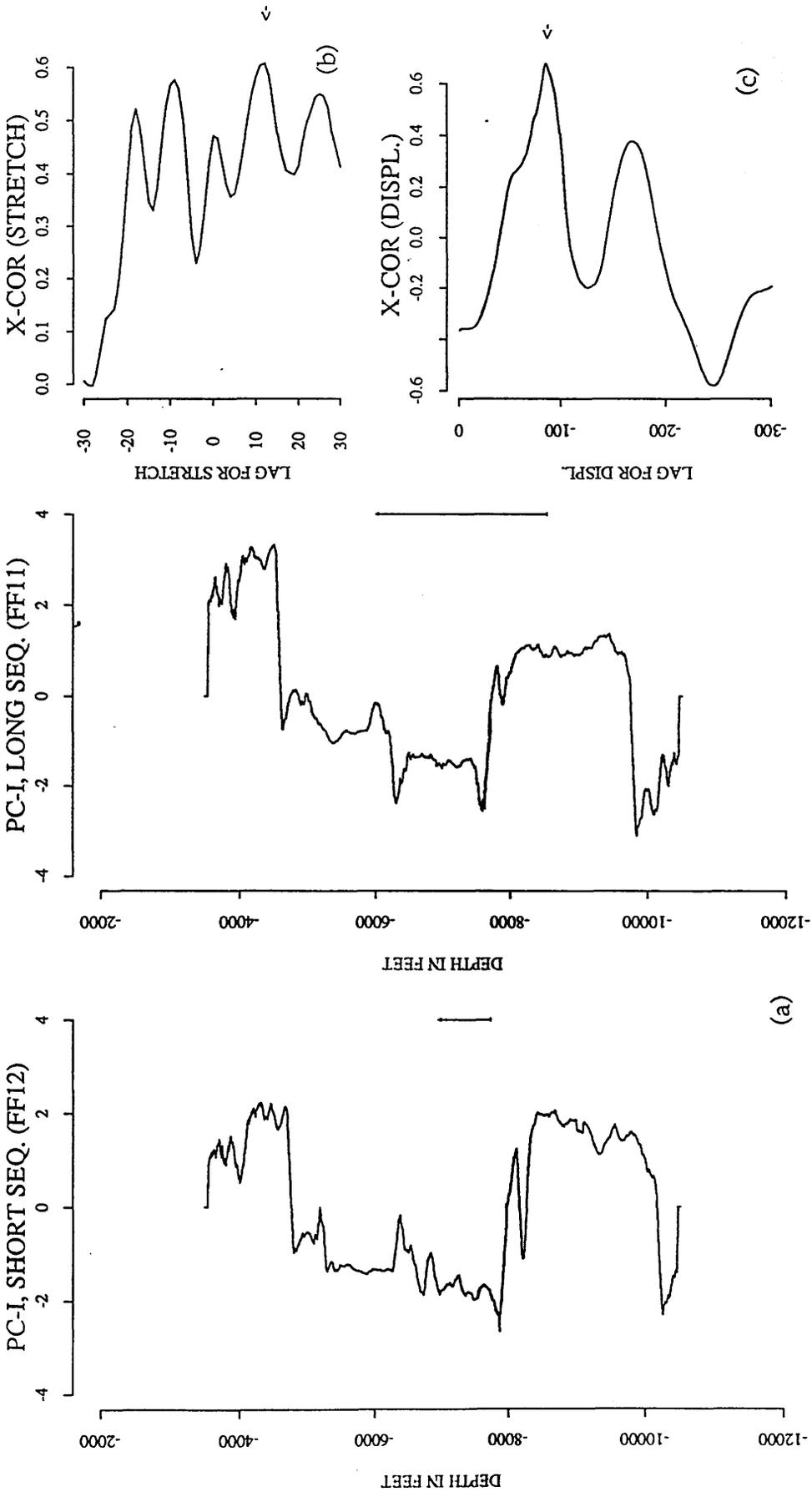


Fig. 4.42 Cross-correlation of the Domran Formation using the derivative data. (a) principal components of FF12 and FF11. (b) the cross-correlation function of power spectra ($S=1.32$). (c) the cross-correlation function (0.678) for displacement of 88 units (440 feet) when the long sequence is stretched 1.32 times.

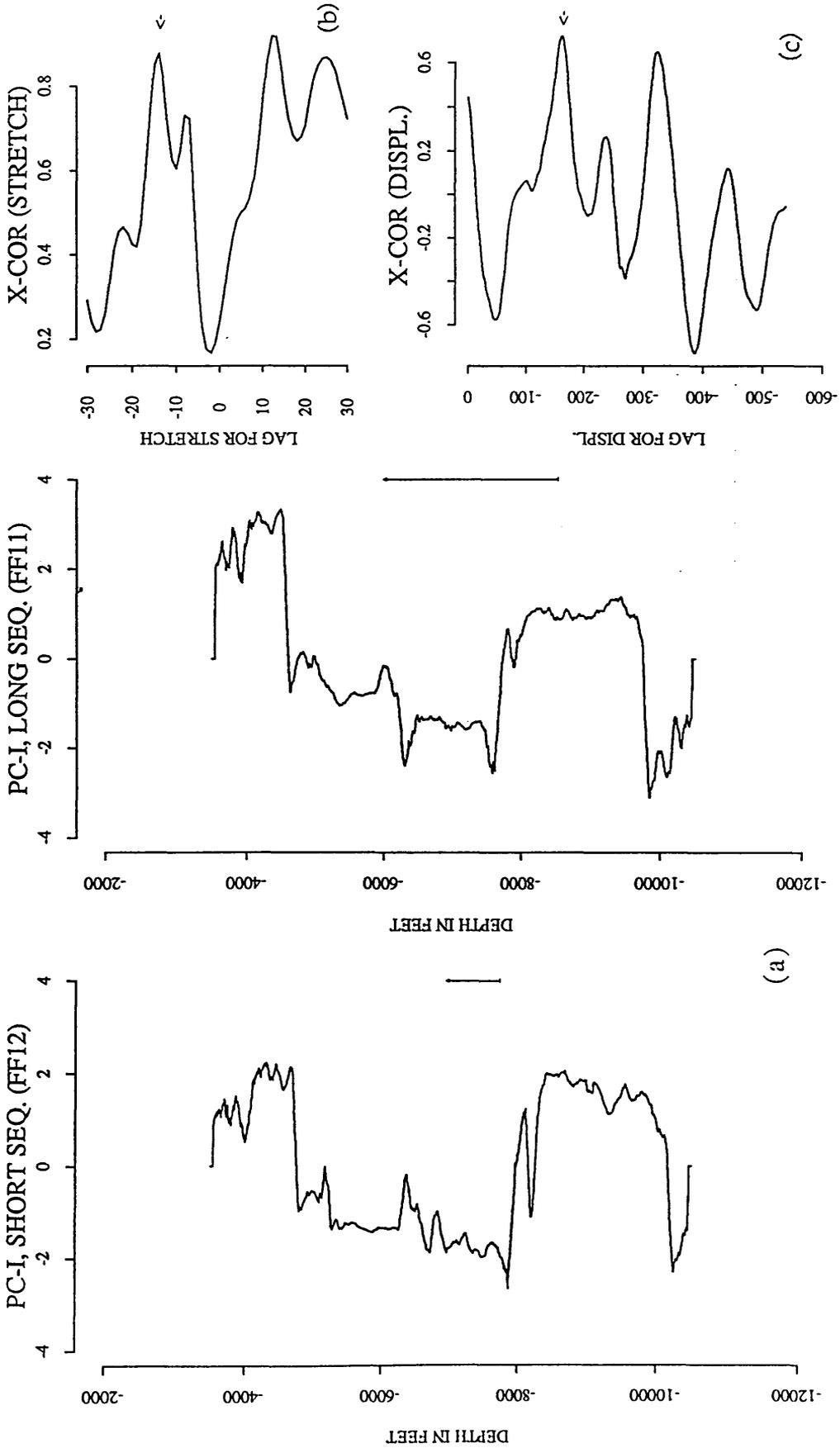


Fig. 4.43 Cross-correlation of the Domran Formation using the original data (principal components. (a) principal components of FF12 and FF11. (b) the cross-correlation function of power spectra ($S=1.38$). (c) the cross-correlation function (0.719) for displacement of 117 units when the short sequence is stretched 1.38 times.

4.3.4. (d) Correlation of the Ruaga Formation

The mathematical cross-correlation of the Ruaga Formation in these boreholes is extremely accurate in predicting the stretch factor as well as the displacement. This correlation is made between the Ruaga Formation in well FF12 (Fig. 4.37a, number 4) and a section from well FF11 (Fig. 4.37b, letter D). A stretch factor of 1.07 compared with 1.02 is obtained from correlating the power spectra (Fig. 4.44f and 4.45b). The cross-correlation function of the stretched series shows a maximum peak of 0.899 at displacement of 250 feet (50 units) compared with geological displacement of 245 feet (Fig. 4.45c).

4.3.4. (e) Correlation of the Heira Formation

Figure 4.46 shows the Heira Formation in FF12 (indicated by number 5 in Fig. 4.37a) when compared with a window length marked by letter E (Fig. 4.37b). The cross-correlation function of power spectra (using the principal components for stretching) yields a peak at a lag of $-v=-3$ which gives a stretch factor of 1.07 which agrees with the expected value, 1.07 (Fig.4.46b). The negative sign indicates that the short series is stretched relative to the long series. The maximum in the cross-correlation function of the stretched series (0.941) is observed in Fig. 4.46c for displacement of 150 feet (30 units) compared with geological displacement of 180 feet.

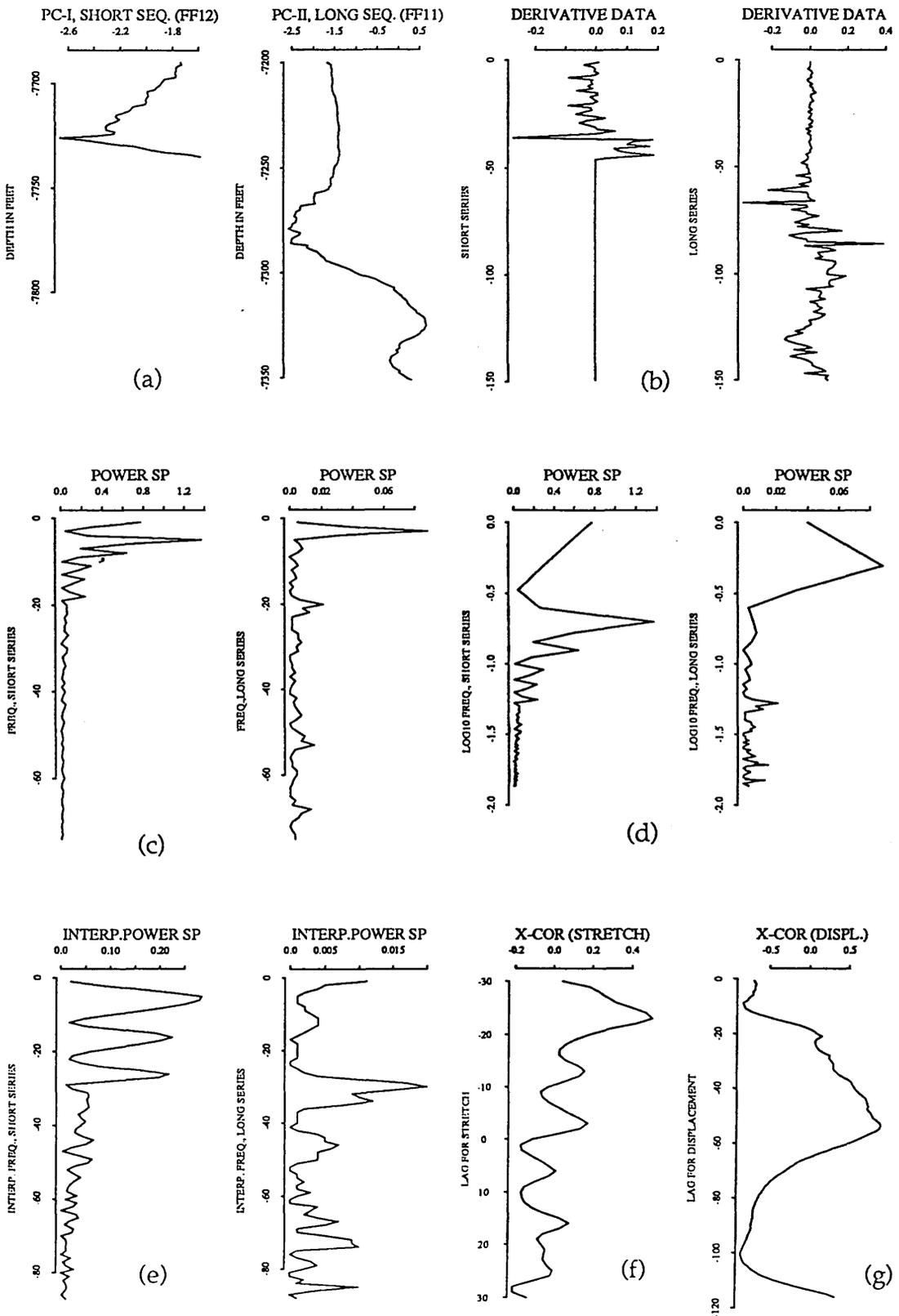


Fig. 4.44 plot of cross-correlation of the Ruaga Formation in FF12 and FF11 using the derivative data. (a) principal components of FF12 and FF11. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

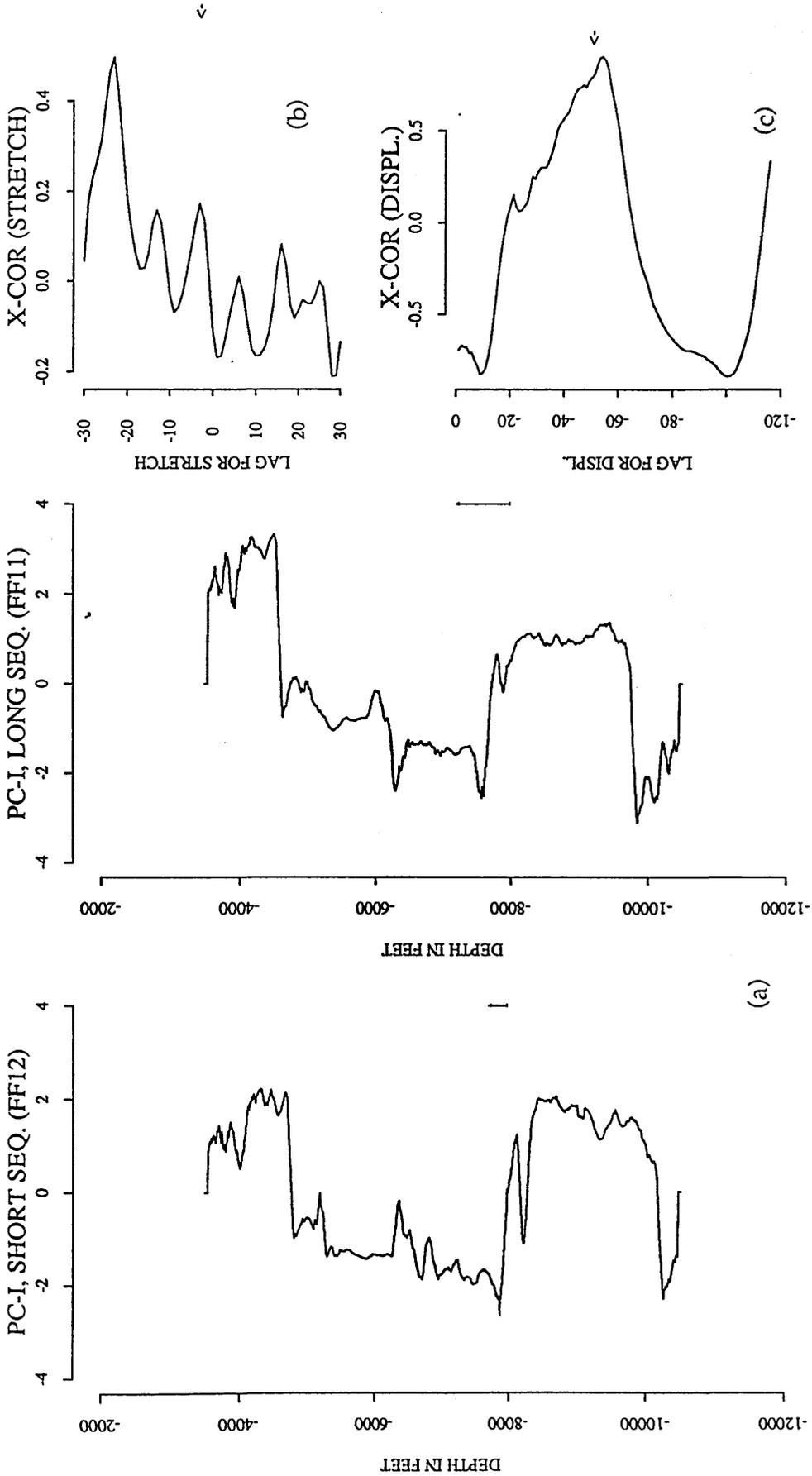


Fig. 4.45 Cross-correlation of the Ruaga Formation using the derivative data (a) principal components of FF12 and FF11. (b) the cross-correlation function of power spectra ($S=1.07$). (c) the cross-correlation function (0.899) for displacement of 50 units (250 feet) when the short sequence is stretched 1.07 times.

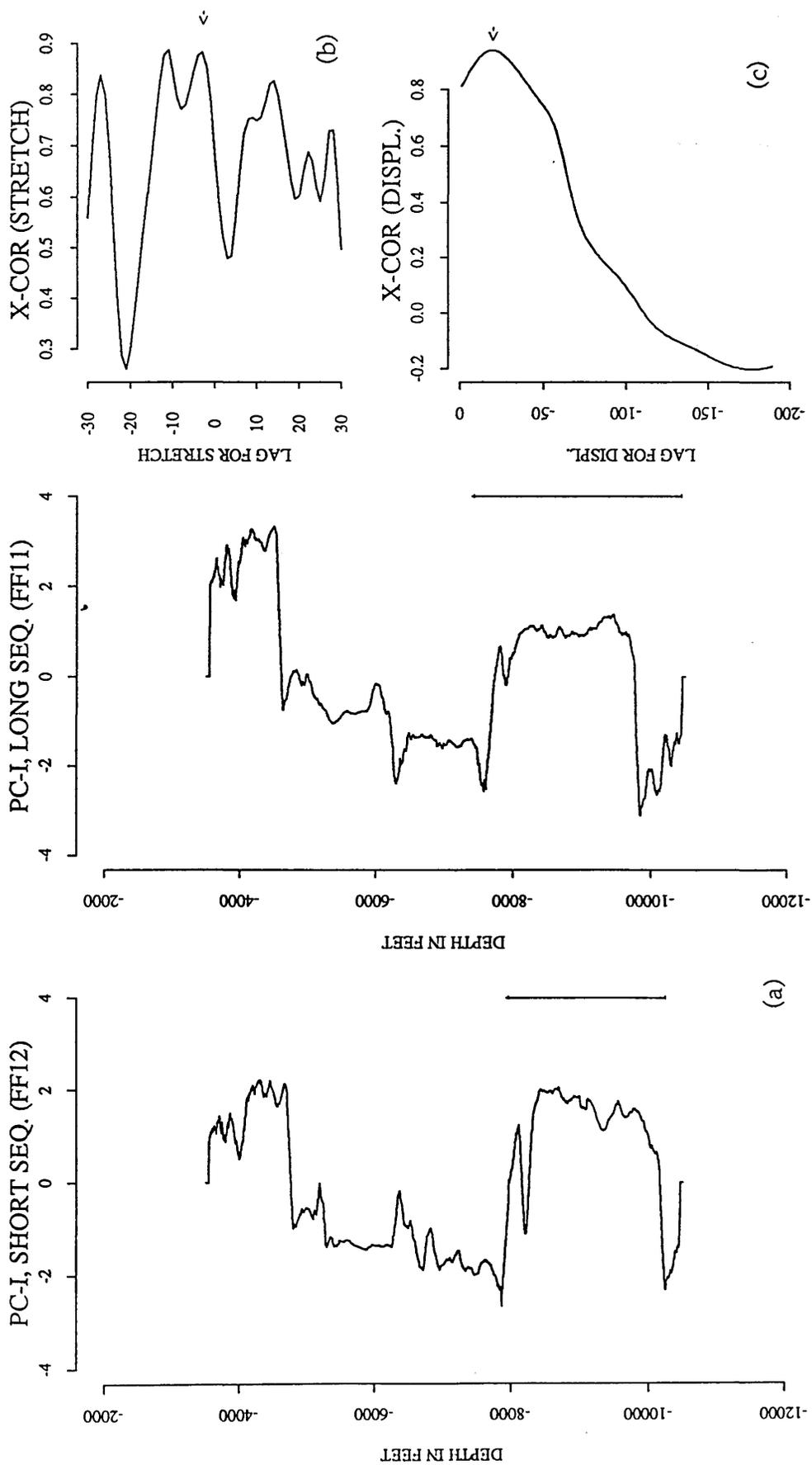


Fig. 4.46 Cross-correlation of the Heira Formation using the original data (principal components) (a) principal components of FF12 and FF11. (b) the cross-correlation function of power spectra ($S=1.07$). (c) the cross-correlation function (0.941) for displacement of 30 units (150 feet) when the short sequence is stretched 1.07 times.

4.3.5 Correlation between FF12 and FF10

As the last application of PCAXCOR on well-log data, boreholes FF12 and FF10 (Fig. 4.7) are subjected to the analysis. Well FF12 delineates the eastern part of the Attahaddy field, and well FF10 was drilled to delineate the southern part of the gas-water contact. The distance between the two wells is about 10 Km, and a total of 7000 feet of each borehole is used.

As with the previous wells, the eigenvalues, eigenvectors and the percentage of each eigenvalue of well FF10 are calculated (Table 5 Appendix E). The first principal component of this borehole is shown in Figure 4.47a. The boundaries are then identified (Fig. 4.47b, Table 4.6). The Sheghega Formation is picked at depth 4440 feet, the Domran Formation is at depth of 5750 feet, the Heira Formation at depth 7670 feet, and the Zmam Formation is identified at depth 10000 feet.

Table 4.6 The Geological Formation depths and the predicted Formation depths of FF10 using the boundary identification technique.

Formation	Geological depth	Predicted depth
Sheghega	4440	4440
Domran	6256	5750
Ruaga	7451	-
Heira	7780	7670
Zmam	10593	10000

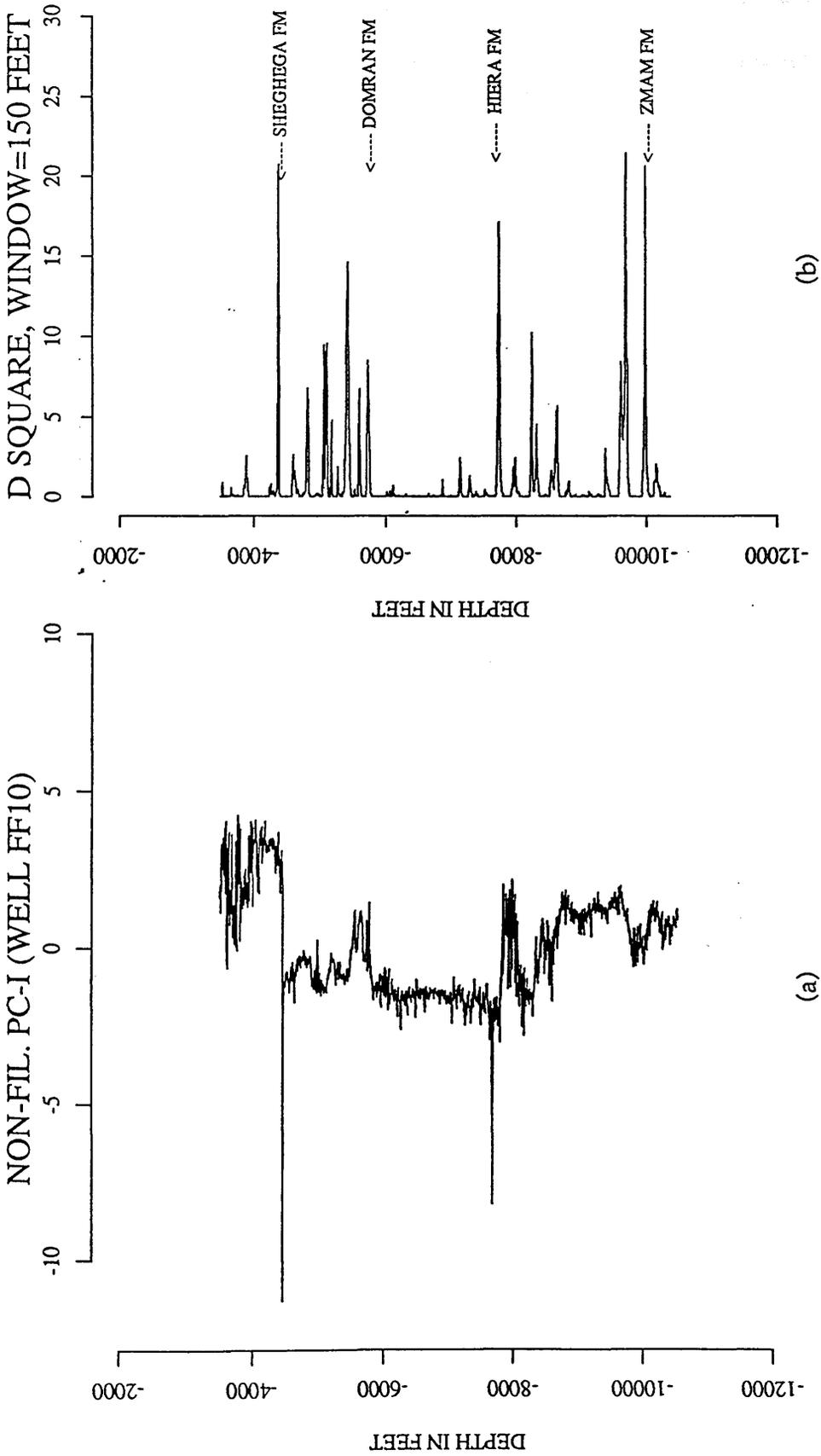


Fig. 4.47 Plot of the first principal components (sampled at 5 feet) of well FF10 and its boundaries. (a) non-smoothed principal components of FF10. (b) boundaries of different formations (window=150 feet).

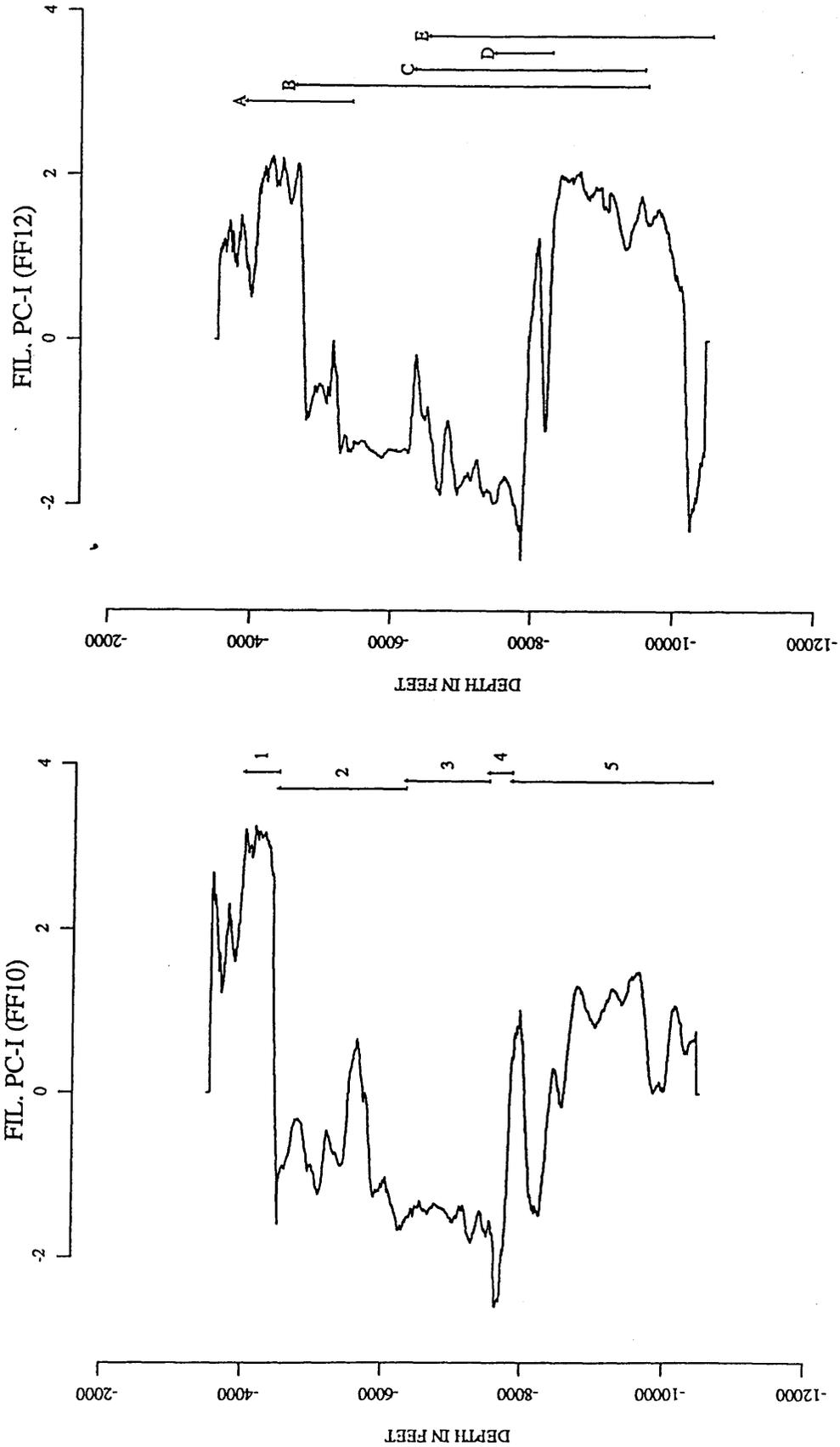
Finally the mathematical cross-correlation between the two filtered principal components is performed with different window lengths for different formations (Fig. 4.48).

4.3.5 (a) Correlation of the Etel Formation

Cross-correlation of the Etel Formation in well FF10 (Fig. 4.48a, number 1) is made against a section of the principal component of well FF12 (Fig. 4.48b, letter A). The resultant cross-correlation function of the power spectra (Fig. 4.49g and Fig. 4.50b) indicates a stretch factor of 1.26 compared with an expected value of 1.3, for the short sequence. The cross-correlation function of the stretched sequence yields an optimum maximum peak of 0.826 (Fig. 4.49g and 4.50c).

4.3.5 (b) Correlation of the Sheghega Formation

The Sheghega Formation in well FF10 (Fig. 4.48a, number 2) is compared with that of well FF12 (Fig. 4.48b, letter B). The cross-correlation function of the power spectra indicates a stretch of 1.17 compared with 1.20, for the long sequence (Fig. 4.51g and 4.52b), the cross-correlation function of the stretched sequence has a maximum peak of 0.707 with a displacement of 39 feet (Fig.4.52c) compared with geological displacement of 59 feet.



(a)

(b)

Fig. 4.48 Plot showing the smoothed principal component of the correlated sequences (FF10 and FF12). (a) smoothed component of well FF10. (b) smoothed component of well FF12. A window (marked 1) in (a) is correlated with a window (marked A) in (b). The window (marked 2) in (a) is compared with the window (marked B) in (b), and so forth.

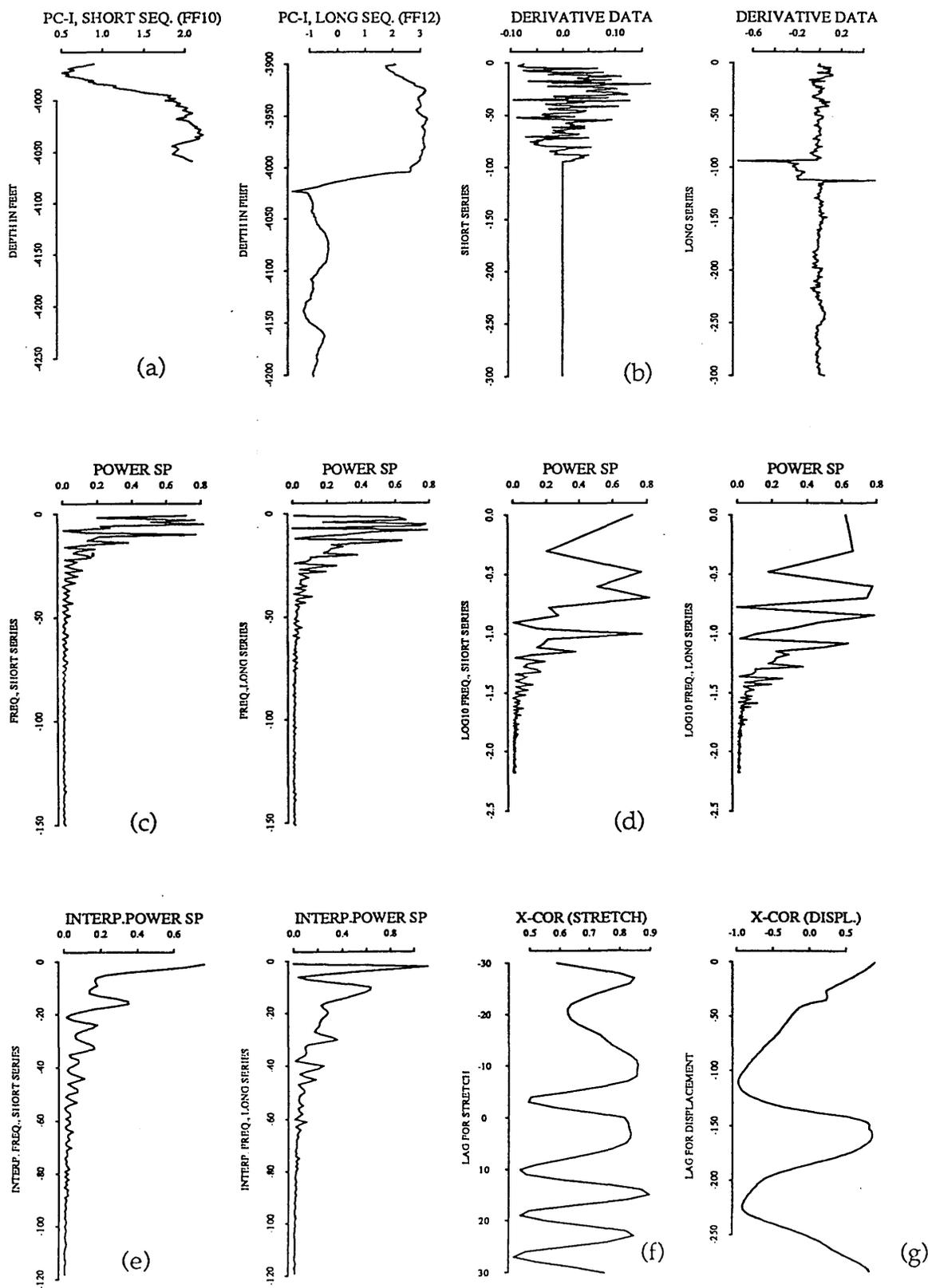


Fig. 4.49 plot of cross-correlation of the Etel Formation in FF10 and FF12 using the derivative data. (a) principal components of FF10 and FF12. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

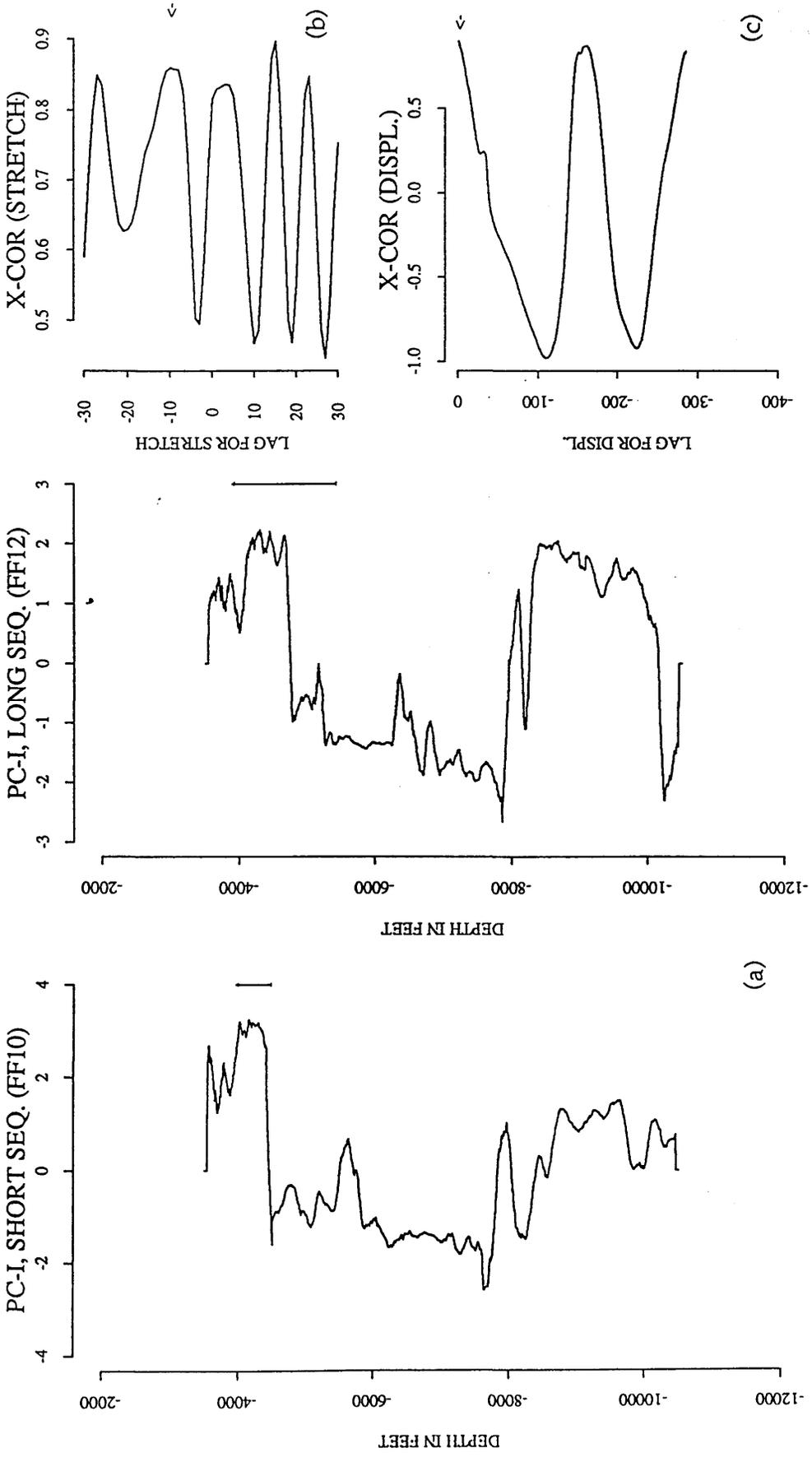


Fig. 4.50 Cross-correlation of the Etel formation using the derivative data. (a) principal components of FF10 and FF12. (b) the cross-correlation function of power spectra at a lag $v=10$ ($S=1.26$). (c) the cross-correlation function (0.826) for displacement of 5 feet when the short sequence is stretched 1.26 times.

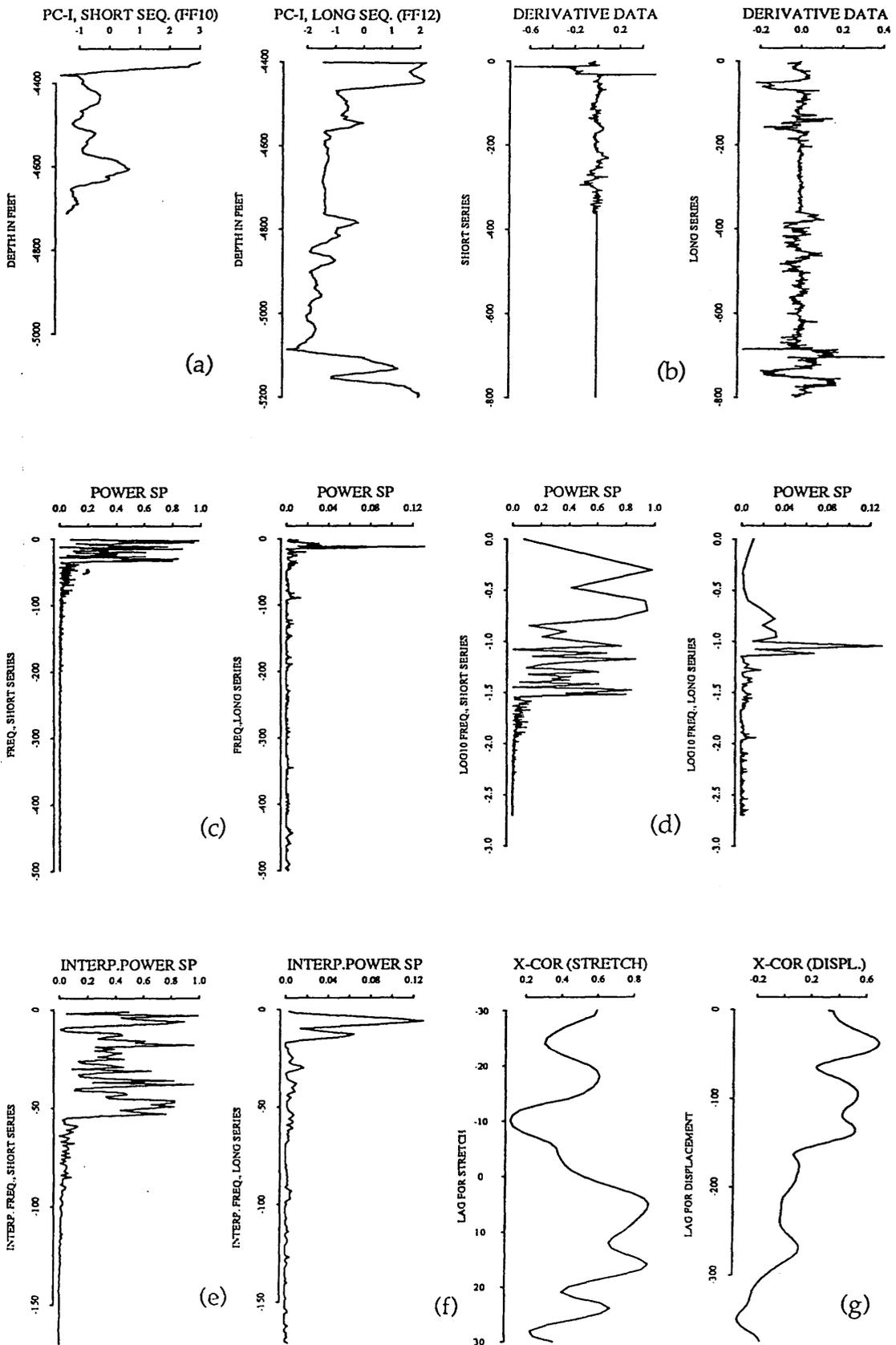


Fig. 4.51 Plot of cross-correlation of the Sheghega Formation in FF10 and FF12 using the derivative data. (a) principal component of FF10 and FF12. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

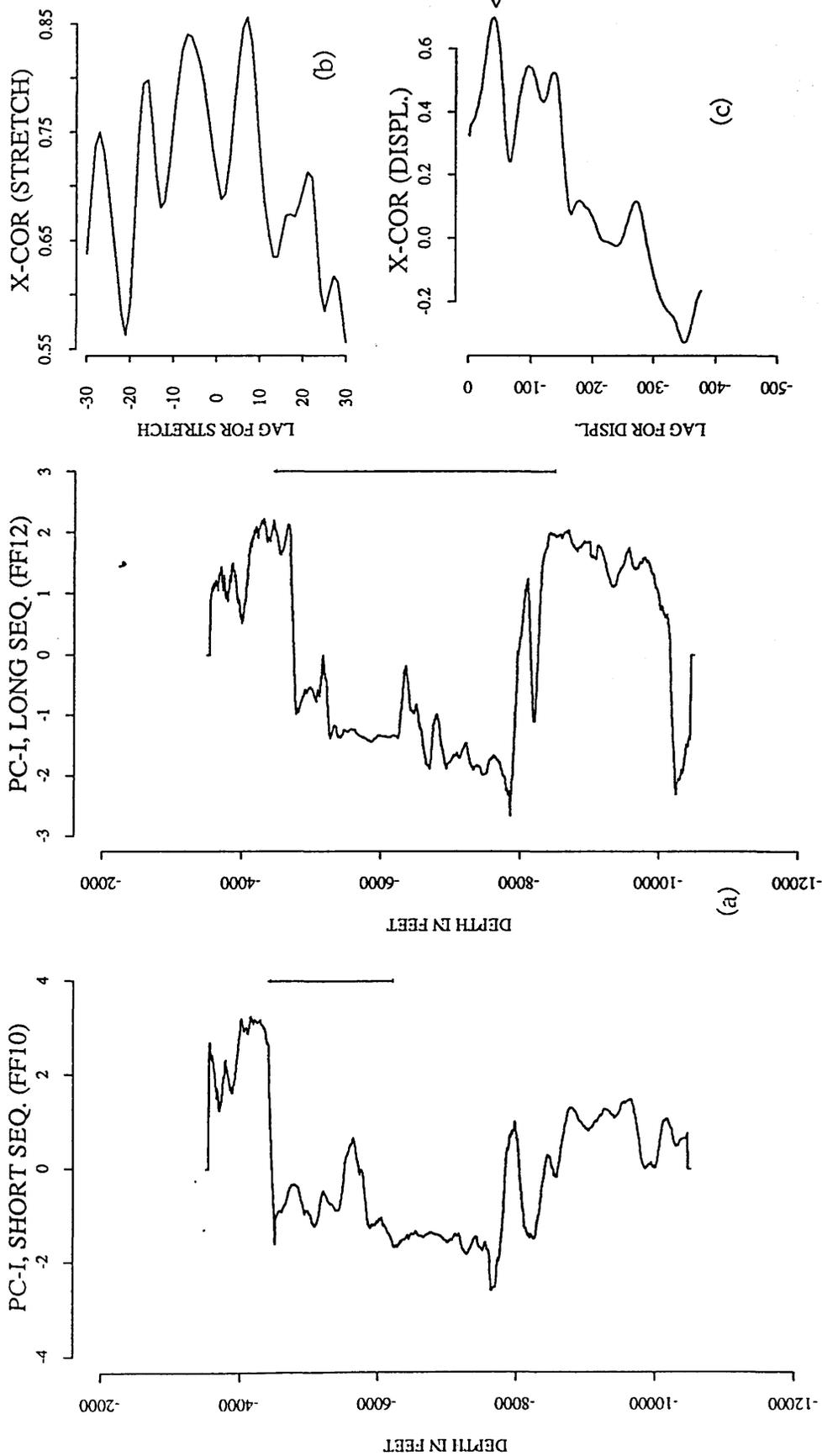


Fig. 4.52 Cross-correlation of the Sheghega formation using the derivative data. (a) principal component of FF10 and FF12. (b) the cross-correlation of power spectra ($S=1.17$). (c) the cross-correlation function (0.707) for displacement of 45 feet when the long sequence is stretched 1.17 times.

4.3.5 (c) Correlation of the Domran Formation

Cross-correlation of the power spectra indicates a thickening of 1.20 when correlating the Domran Formation in FF10 against FF12 (Fig. 4.48a, number 3 against Fig. 4.48b letter C). A maximum of 0.504 (Fig. 4.53g and Fig. 4.54c) is observed when the correlation of the stretched data is made. A stretch factor of 1.20 compared with geological stretch of 1.56 is obtained. The computer selection of the displacement is 100 feet compared with the known displacement of 134 feet. This deviation from the known geological correlation of the stretch factor and small deviation of the displacement for the Domran Formation (1.20 compared with 1.56) is explained by the fact that the computer selection of stretch factor is based on the heighest two peaks in the cross-correlation function of power spectra. In this case both peaks (Fig. 4.52g and Fig.4.53c) were spurious. As mentioned in previous sections, the value of the cross-correlation function is dependent on the average similarity between two signals, however, the Domran Formation in FF10 and FF12 shows no similarity between the two curves (Fig. 4.53a) perhaps due to the long distance (10 Km) between the two boreholes.

4.3.5 (d) Correlation of the Ruaga Formation

In contrast with the Domran Formation and despite the long distance between the correlated wells, the Ruaga Formation in the two boreholes preserve an average similarity in shape of the two curves (Fig. 4.55a). The Ruaga Formation in well FF10 (Fig. 4.48a, number 4) is compared with a window (Fig. 4.48b, D) of the first principal component of well FF12. The computer successfully correlated the two formations with excellent accuracy.

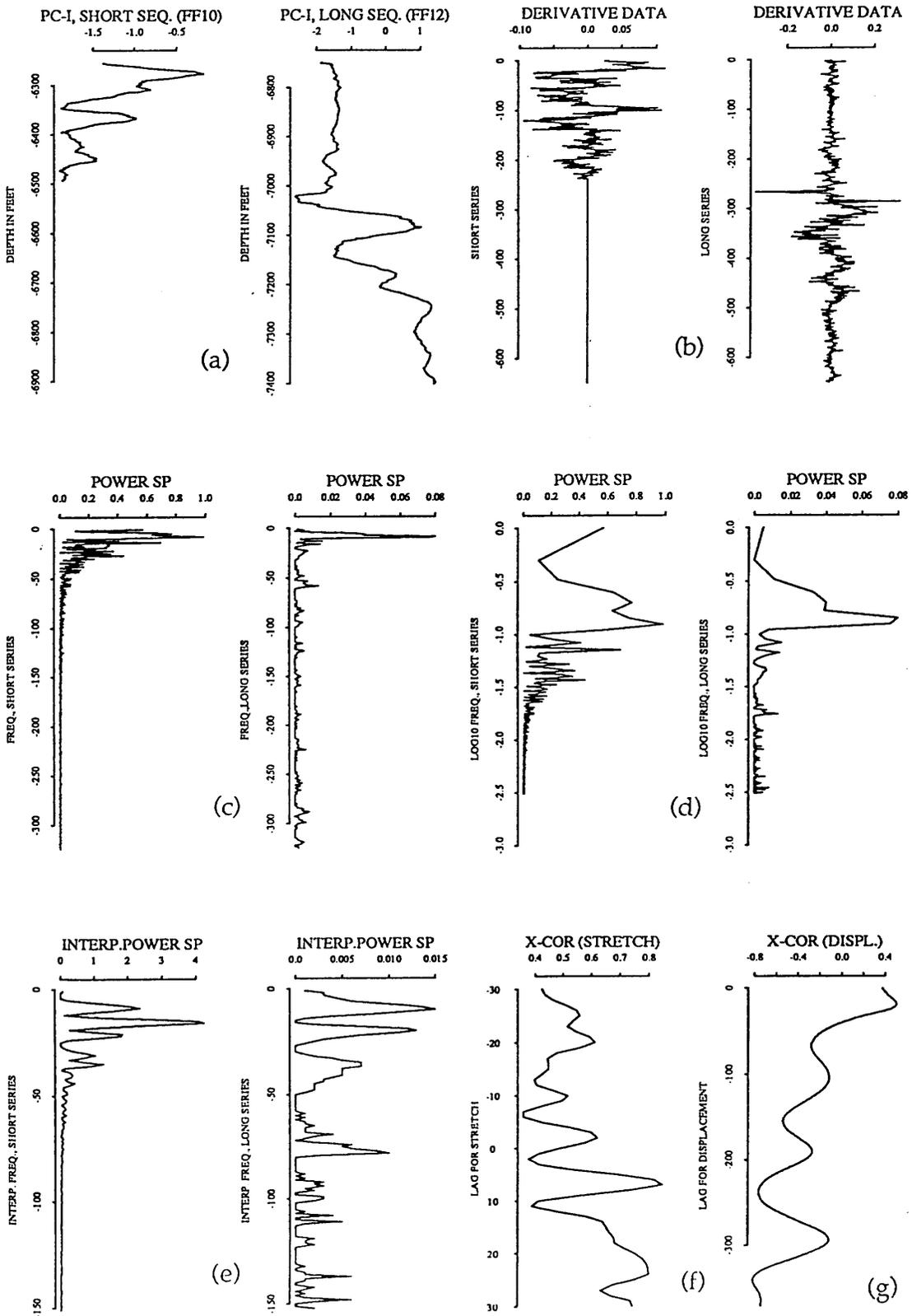


Fig. 4.53 Plot of cross-correlation of the Sheghega Formation in FF10 and FF12 using the derivative data. (a) principal component of FF10 and FF12. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

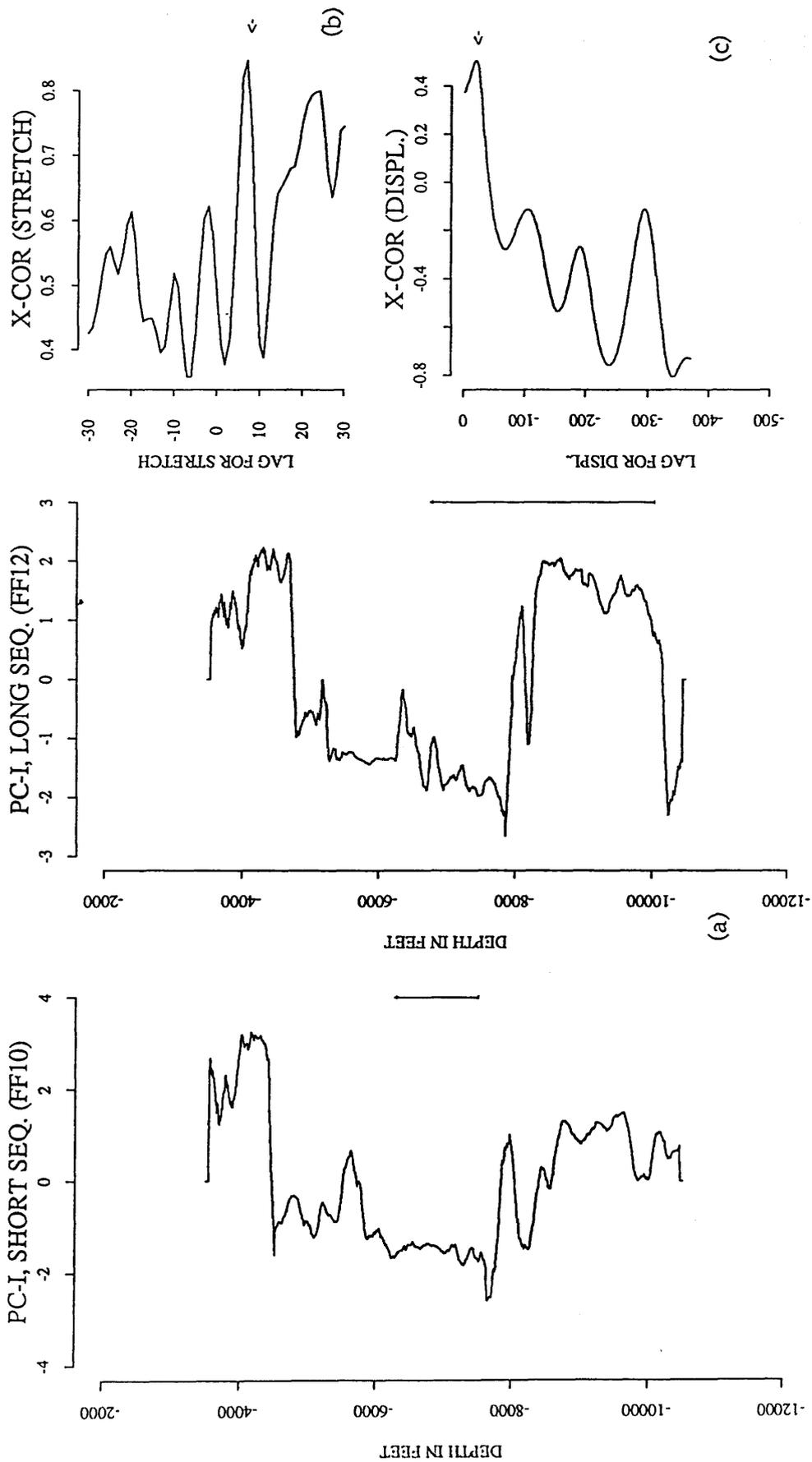


Fig. 4.54 Cross-correlation of the Domran Formation using the derivative data. (a) principal component of FF10 and FF12. (b) the cross-correlation of power spectra ($S=1.20$). (c) the cross-correlation function (0.50) for displacement of 20 units (100 feet) when the long sequence is stretched 1.20 times.

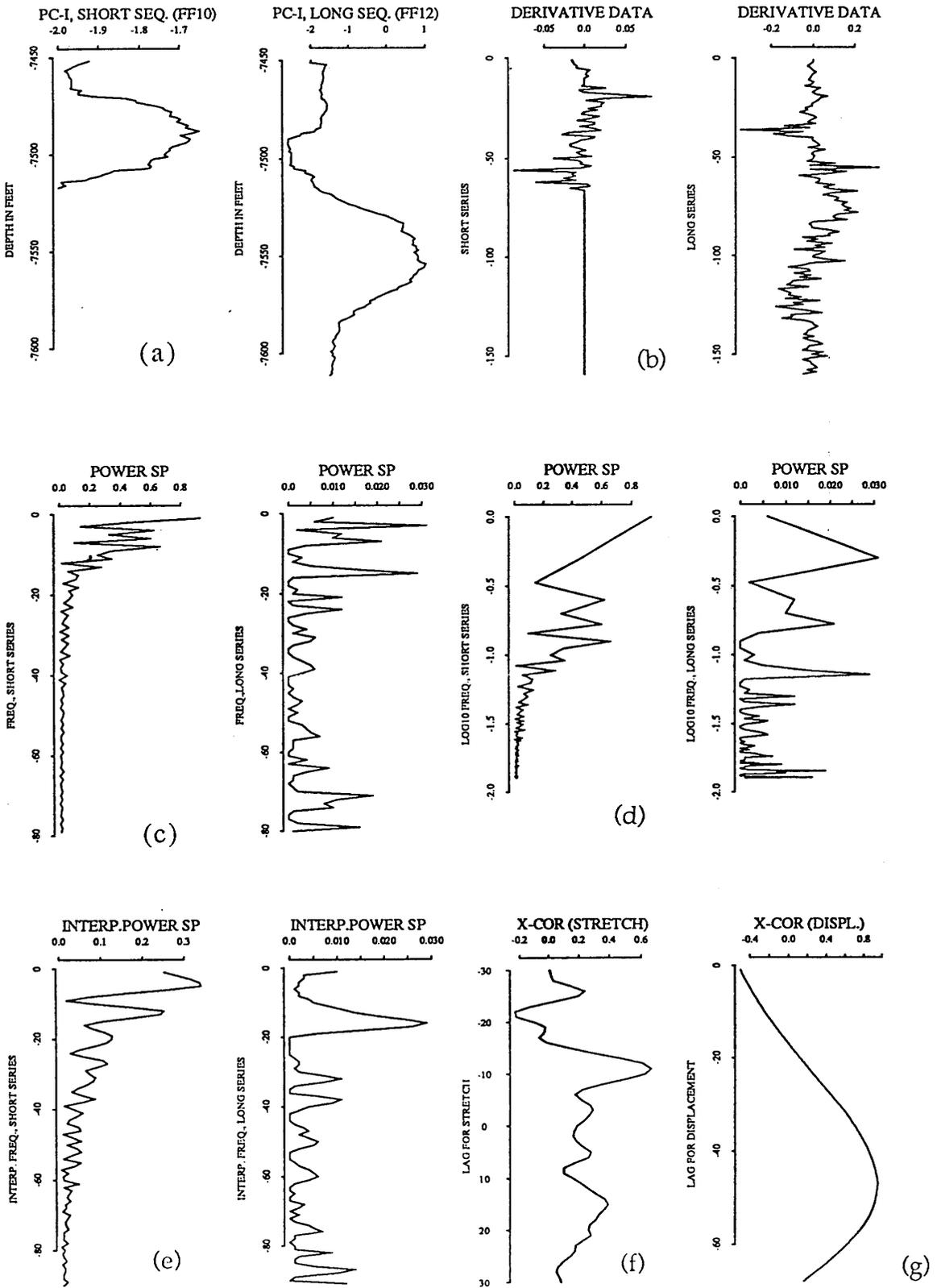


Fig. 4.55 plot of cross-correlation of the Ruaga Formation in FF10 and FF12. (a) principal components of FF10 and FF12. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

Figure 4.55g and Figure 4.56b shows the cross-correlation function of power spectra which indicates a thickening of strata from well FF10 towards well FF12. The stretch factor $-v=-1.41$, compared with expected value of 1.43. Figure 4.55g and Figure 4.56c show a symmetric peak of the cross-correlation function of the stretched series (0.953) and determines the displacement at 235 feet (compared with geological correlation position of 240 feet).

4.3.5 (e) Correlation of Heira Formation

The Heira Formation in well FF12 (Fig. 4.48a, number 5) is correlated with the section in FF10 indicated by letter E (Fig. 4.48b). A stretch factor of 1.17 (compared with 1.19) for long series (FF12) is obtained when correlating the power spectra (Fig. 4.57f and Fig. 4.58b). The cross-correlation function of stretched series yields a maximum (0.739) (Fig. 4.57g and Fig. 4.58c) for a displacement of 144 feet compared with 145 feet which agrees with the known geological correlation between FF10 and FF12.

4.4 Correlation of lithologies within a Formation

A part of well FF13 consists of intercalations of small beds of shale and limestone in the Sheghega Formation at depth of 6200-6500 feet (Fig. 4.9a and Fig. 4.10a). This rock unit is identified by the electrical logs as a zone of high Gamma Ray (GR) and high transit time (DT). The correlation of this small unit is made with the Sheghega Formation in FF11. The thickness of the small unit in this well is about 375 feet. Figures 4.59 and Fig. 4.61 show the computer plot of this rock unit in FF13 and the Sheghega Formation in

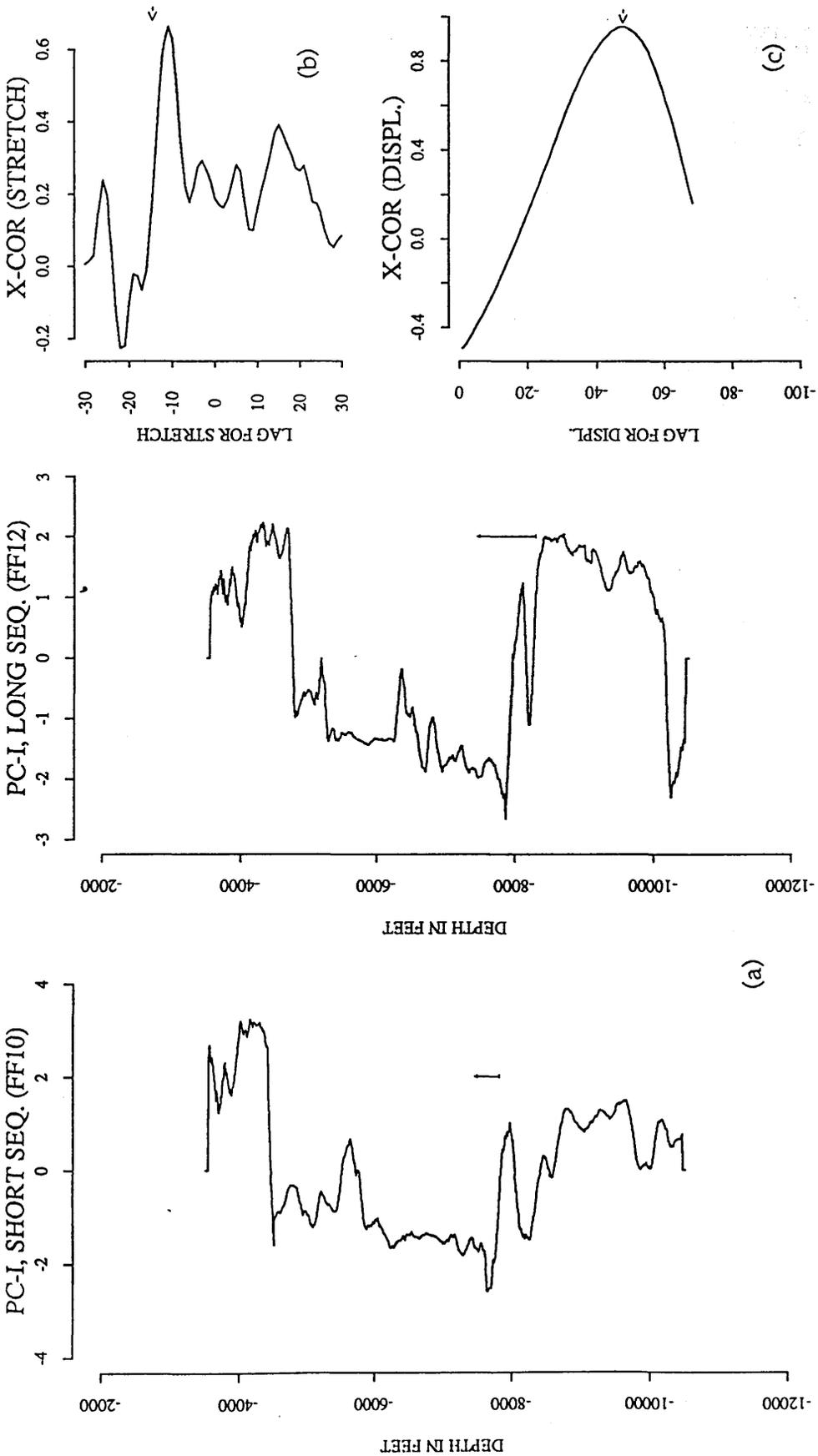


Fig. 4.56 Cross-correlation of the Ruaga Formation. (a) principal components of FF10 and FF12. (b) the cross-correlation function of power spectra ($S=1.41$). (c) the cross-correlation function (0.953) for displacement of 235 feet (47 units) when the short sequence is stretched 1.41 times.

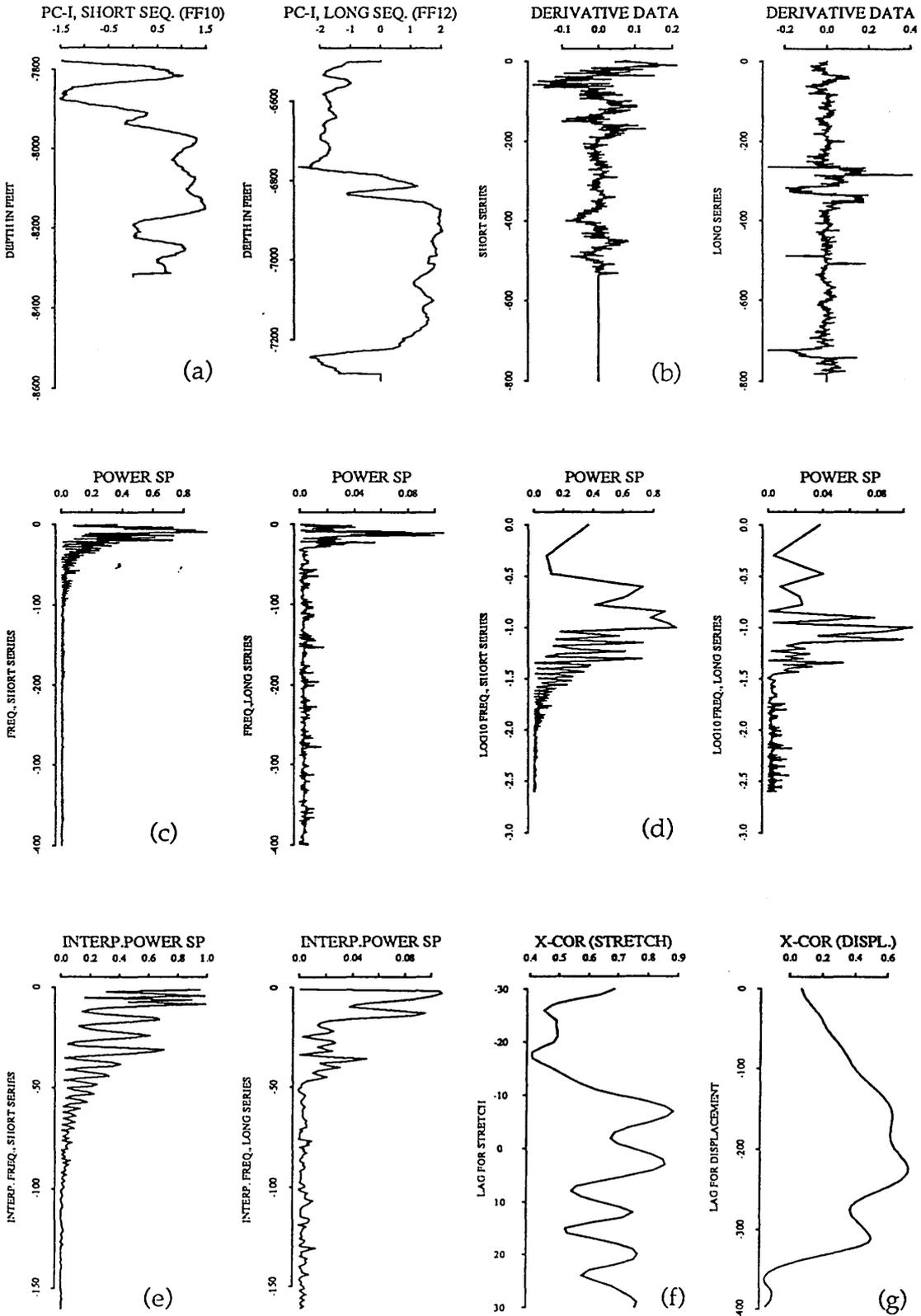


Fig. 4.57 plot of cross-correlation of the Heira Formation in FF10 and FF12. (a) principal components of FF10 and FF12. (b) is the derivative of the data. (c) the power spectra. (d) is logarithmic spaced spectra. (e) is the interpolated spectra. (f) is the cross-correlation function of the power spectra. (g) is the cross-correlation function of the stretched series.

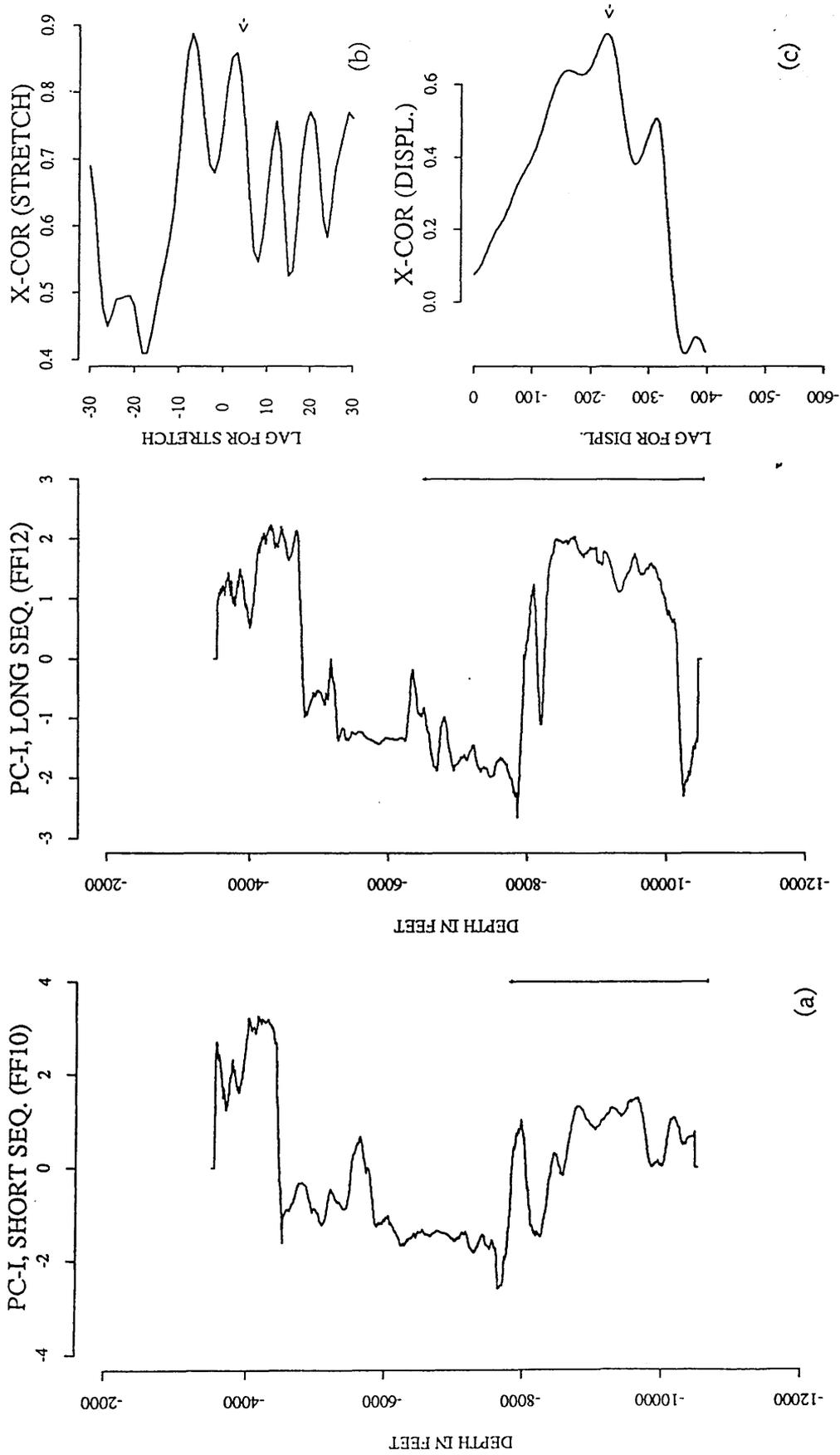


Fig. 4.58 Cross-correlation of the Heira Formation. (a) principal components of FF10 and FF12. (b) the cross-correlation function of power spectra ($S=1.17$). (c) the cross-correlation function (0.739) for displacement of 226 units when the long sequence is stretched 1.17 times.

FF11. The derivative data are first used (Fig. 4.59) for stretching. The cross-correlation function of the power spectra (Fig. 4.59d and Fig. 4.60b) shows a high correlative peak at a lag $-v=-5$ ($S=1.12$ compared with 1.14) for the short series (small bed) and a maximum peak of the cross-correlation function of the stretched series (0.973) for displacement of 1350 feet (compared with geological displacement of 1400 feet) is observed in Figure 4.59e and Figure 4.60c. More accurate results are obtained by using the filtered principal components data (Fig. 4.61). The cross-correlation function of the power spectra (Fig. 4.61b) has a high peak at a lag $-v=-6$ ($S=1.15$) for the short sequence (small bed). The high magnitude of the cross-correlation function (0.958) is identified (Fig. 4.61c) at a displacement of 1400 feet (280 units).

The above results demonstrate the accurate determination of both the stretch factor (thickening and thinning) of beds and the relative displacement between two rock units when the power spectra of the first principal component are used.

4.5 Characterization of different rock type

In principal component analysis, the first principal component of a certain set of variables reflects exactly the behaviour of each individual variable and represents all these variables by one unique measure. The magnitude of the first principal component scores can also be used to characterise different lithologies, (Fig. 4.62b) for example, the Etel Shale Formation (4414-4762 feet) is identified from the Gamma Ray (GR) and transit time curves as shale, and has a mean of 1.70 and standard deviation

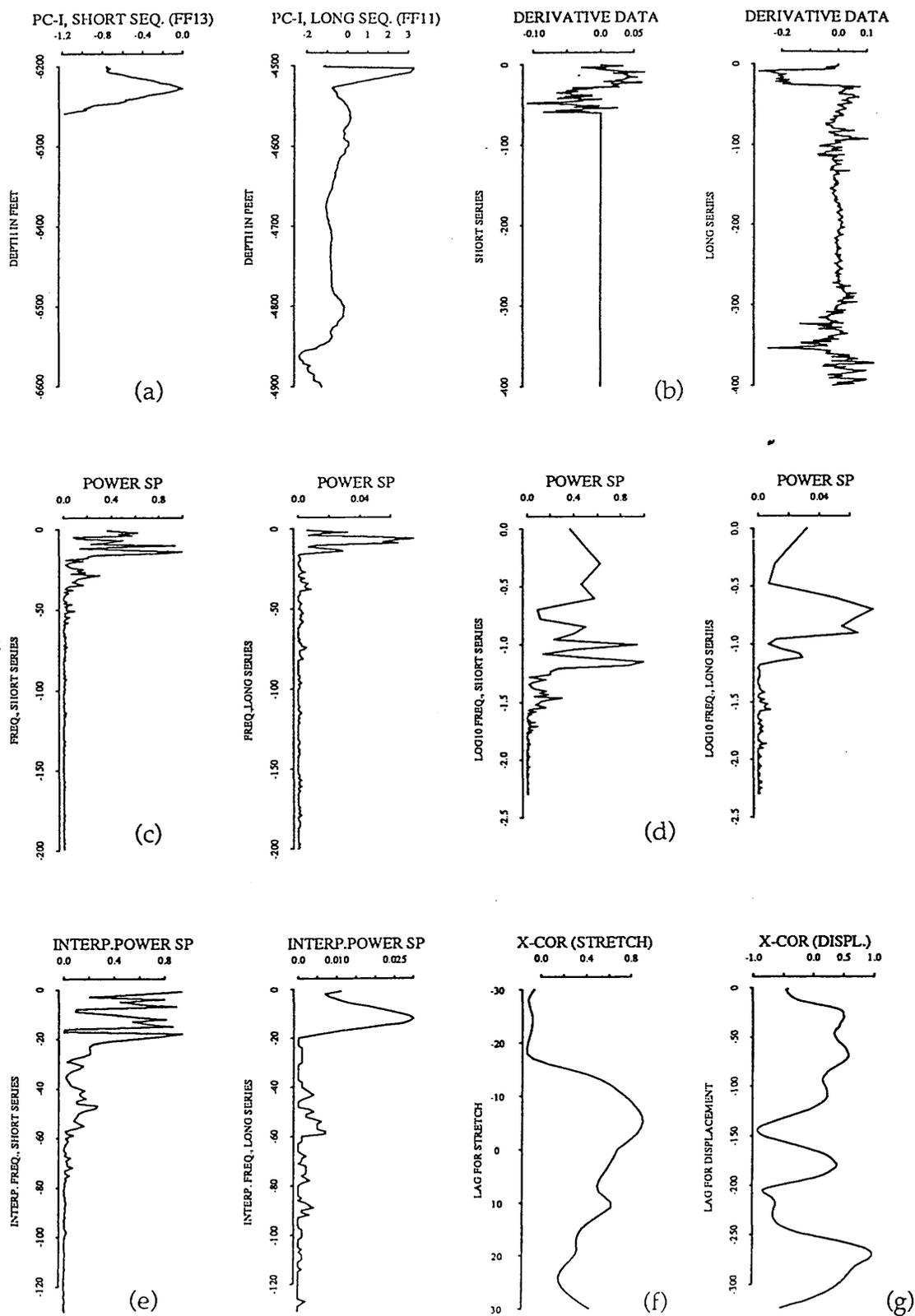


Fig. 4.59 Plot showing the cross-correlation of small bed within the Sheghega Formation in well FF13 and the Sheghega Formation in FF11 using the derivative data. (a) small unit [6200-6500 feet] in well FF13 and the Sheghega Formation in FF11. (b) is the derivative data. (c) the interpolated power spectra. (d) the cross-correlation function of power spectra. (e) the cross-correlation function of the stretched series.

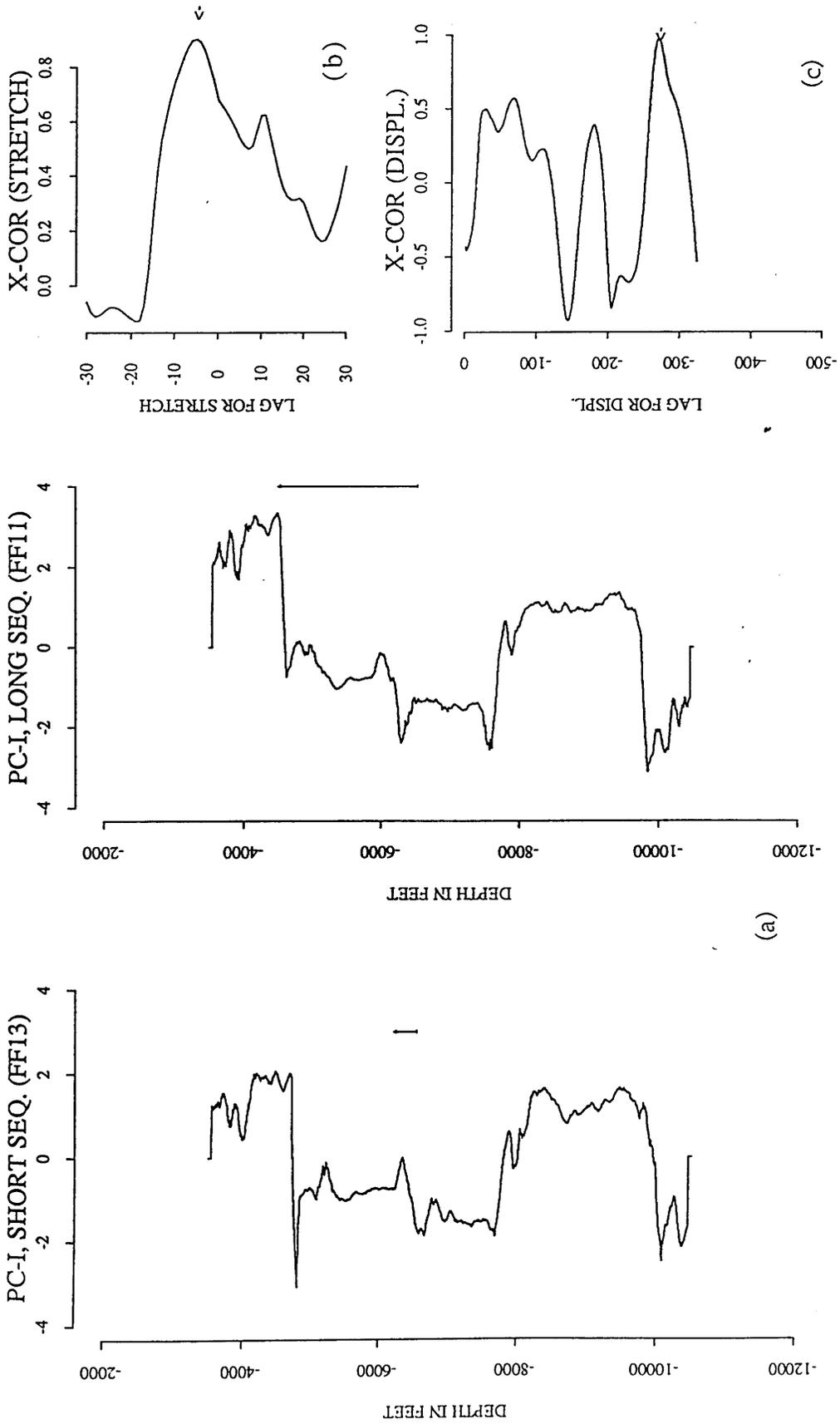


Fig. 4.60 Cross-correlation of small bed within the Sheghega Formation of FF13 and the Sheghega formation in FF11 using the derivative data. (a) principal components of FF13 and FF11. (b) the cross-correlation function of power spectra at a lag $-v=-5$ ($S=1.12$). (c) cross-correlation function of the stretched series (0.973) at displacement $D=270$ units (1350 feet) when the long sequence is stretched 1.12 times.

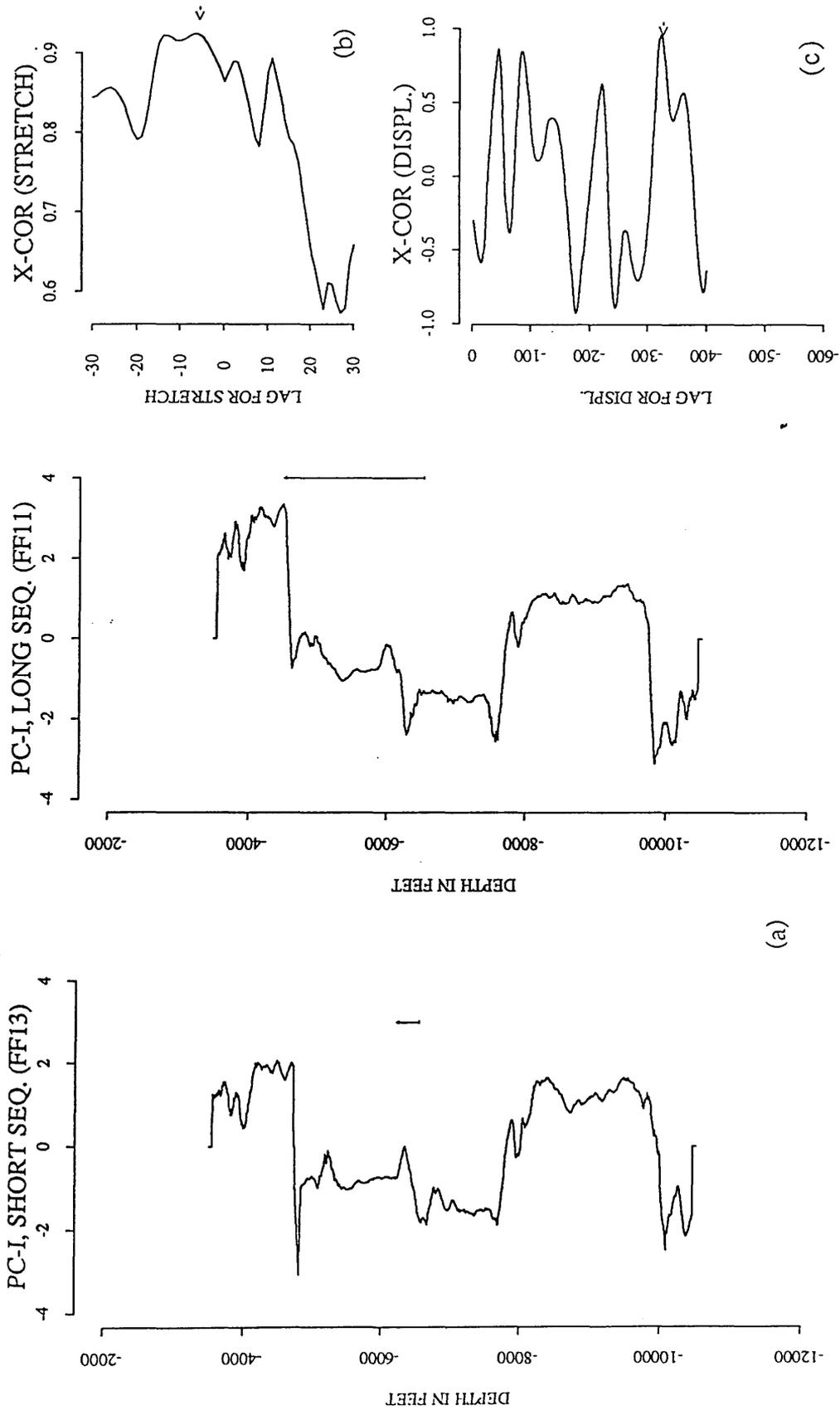


Fig. 4.61 Cross-correlation of small bed within the Sheghega Formation of FF13 and the Sheghega formation in FF11 using the original data (principal components). (a) principal components of FF13 and FF11. (b) the cross-correlation function of power spectra at a lag $-v=-6$ ($S=1.15$). (c) cross-correlation function of the stretched series (0.958) at displacement $D=280$ units (1400 feet) when the short sequence is stretched 1.15.

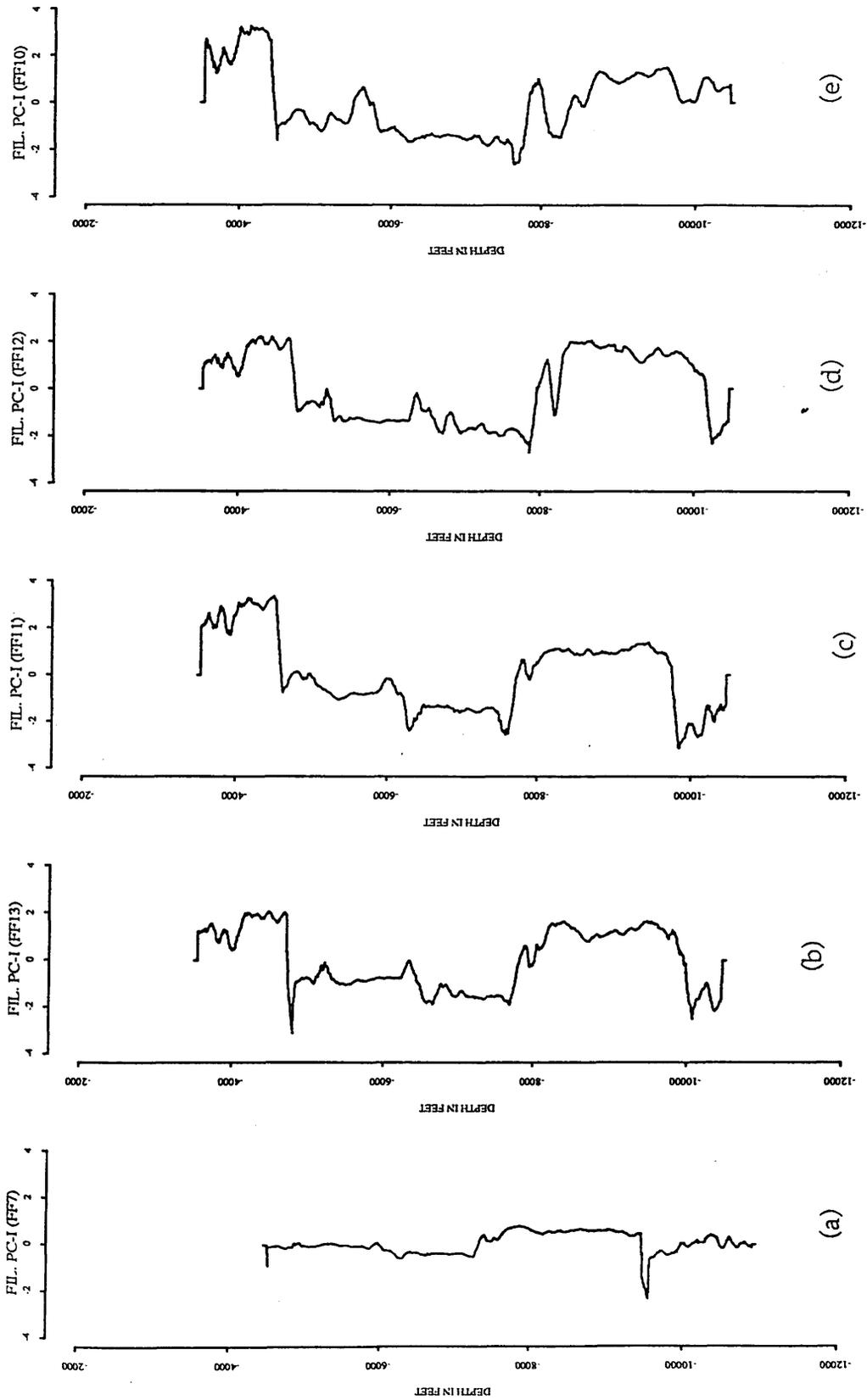


Fig. 4.62 Filtered first principal component of five boreholes from the Attiahaddy field.

value of 0.70 on the principal component, whereas, the Sheghega Formation is identified as limestone unit by low radioactive material (GR) and low transit time (DT), and has a mean of -0.89 and standard deviation of 0.45 on the principal component curve. In addition, the first principal component differentiates between different facies within the same rock unit, for example, within the Hiera Formation. The Formation is highly calcareous at the top part (7750-7950 feet) which has a mean on the principal component equal to 0.04 and a standard deviation of 0.48, and the middle part of the Formation (8550-9250) where the shale becomes less calcareous, the value is shifted above the zero line to have a mean of 1.09 and standard deviation of 0.133. The limestone bed at depth of 7950-8000 feet is identified by a principal component mean and standard deviation corresponding to limestone values (-0.19 and 0.04 respectively). The value of mean of the principal component in the Sheghega Formation is about -0.89 and the value of the standard deviation is about 0.45. In the Domran Formation, however, the values of the mean and the standard deviation of the principal component are shifted below the Sheghega Formation values because the Formation is known to be less argillaceous and more hard than the overlying Sheghega limestone. Both shale intervals occur in the Sheghega Formation between 5100-5250 feet and 6200-6500 feet as indicated by Gamma Ray log as highly argillaceous limestone caused the values of the principal component scores to be shifted towards the shale values (mean of 1.70 and standard deviation of 0.70). In general, any value greater than zero mean is interpreted as shale and any value below zero mean is considered to be limestone. This is common in all the five principal components (Fig.4.62). This accurate representation of different rock type and facies type within the same rock unit may be useful in identifying rock and facies type from the principal

component of the entire sequence. A complete characterization of the principal component scores to identifying different rock type is beyond the scope of this study.

4.6 Discussion of results of real data

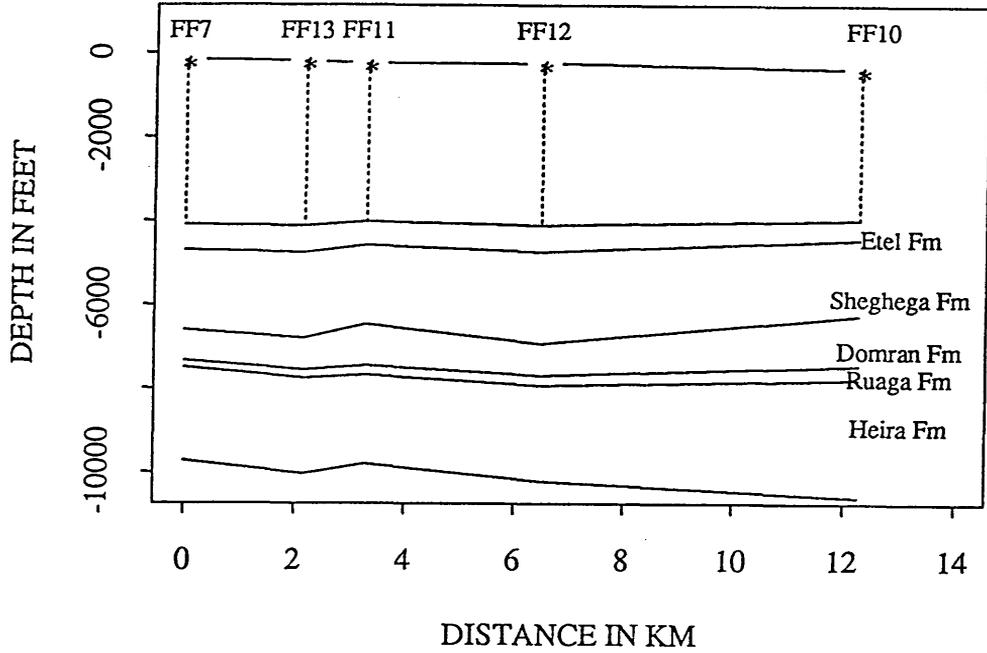
In the previous study by Kwon (1977), the mathematical cross-correlation of was successful in using the spectral analysis of the original data of well-logs. However, this success was limited. The correlation was complicated by the presence of noise signals which are different for different type of logs. Other problem arose because different type of logs record different rock type properties, therefore, each well log is associated with a specific set of frequencies. Kwon (1977) overcame these problems by analysing different frequencies for different log types and modified the program to use multi-log data for the correlation process to improve the reliability of the results over that based on only one type of log. He concluded that the use of the cross-correlation function is of limited use in determining the reliability of computer results using different type of well-logs, as for example, his program failed to correlate the Neutron and Gamma Ray logs from the same borehole.

The noise problem in program PCAXCOR is controlled by filtering the principal component before proceeding with the analysis. In addition, the advanced process of using multi-log data developed by Kwon (1977) is compensated by the impressive improvement of using the single first principal component of all log data. The resultant cross-correlation function

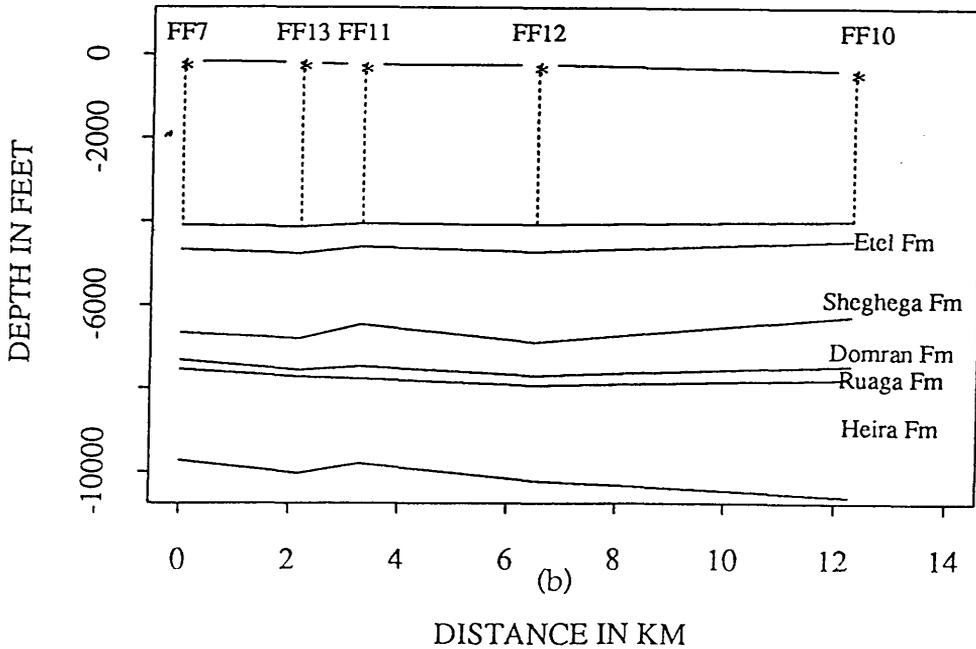
of power spectra of two principal components representing two formations or boreholes is satisfactory in relating the stratigraphic thickening (or thinning) between two wells (Fig. 4.63a against Fig. 4.63b & Fig.64). Table 4.7 shows the evaluation of the program PCAXCOR. A new development in well-log interpretation is the correlation of different rock types within the rock units using the cross-correlation function of the first principal component (Fig. 4.61).

The displacement was determined by stretching (interpolation) the series and cross-correlating such stretched series (sequences). In general, the computer results of the displacement from these series agree with the known geological correlation. There are some occasions where the computer selection of both the stretch factor and the displacement differs slightly, such difference is due to the fact that computer correlation is based on recognising the average similarity between two sequences under processing. However, some formations in different boreholes do not always exhibit distinctive similarities, and hence some errors are observed when correlating such sequences.

The accurate results shown in Figure 4.63 emphasises the value of using the principal components as the basis of well-log analysis.



(a)



(b)

Fig. 4.63 Structural cross-correlation of the studied wells in the Attahaddy field. (a) cross-correlation using geological formation tops. (b) cross-correlation using formation tops from PCAXCOR.

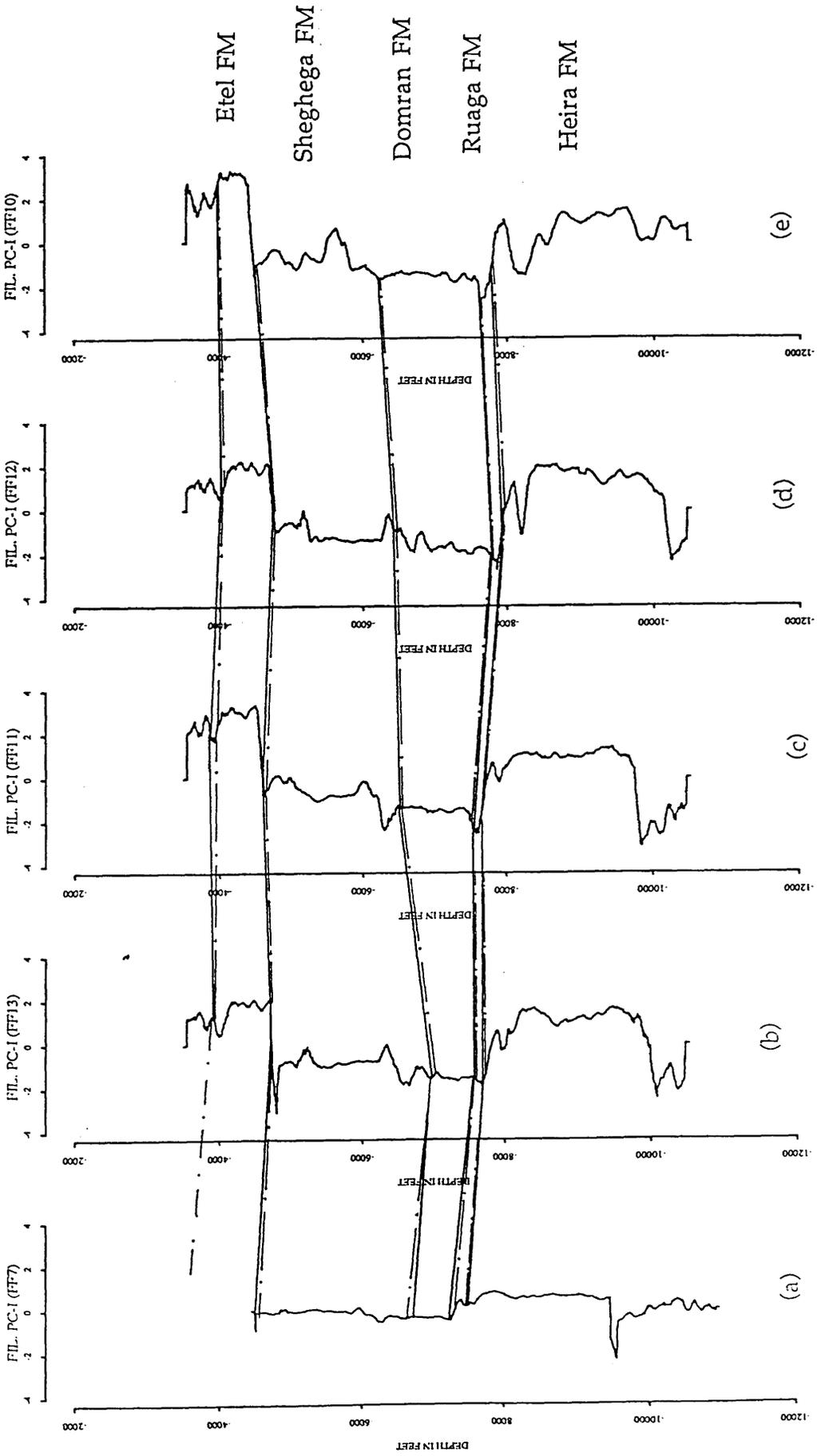


Fig. 4.64 Comparison between the computer correlation and the known geological correlation in the Attahaddy field. The dashed lines are the geological correlation between five boreholes in the study area, and the solid lines represent PCAXCOR correlation of these boreholes.

Correl. between	Formation	Fig.	Geol. Stretch	PCAXCOR Stretch	Geol. Displ.	PCAXCOR Displ.	Diff. in Displ.	Agreement	
FF7-FF13	Sheghega	4.12	1.01	1.12	59	20(0.579)	39	Good	
	Domran	4.14	1.19	1.23	188	265(0.433)	77	Fair	
	Ruaga	4.15	1.10	1.0	223	200(0.760)	23	Good	
	Heira	4.18	1.05	1.12	240	275(0.867)	35	Good	
FF11-FF13	Etel	4.22	1.13	1.05	110	150(0.808)	40	Good	
	Sheghega	4.27	1.05	1.05	185	195(0.799)	10	Excellent	
	Domran	4.29	1.23	1.26	295	310(0.645)	15	Excellent	
	Ruaga	4.31	1.20	1.10	112	125(0.893)	13	Excellent	
	Heira	4.34	1.08	1.17	38	19(0.417)	19	Excellent	
FF11-FF12	Etel	4.39	1.10	1.0	15	75(0.701)	60	Fair	
	Sheghega	4.40	1.16	1.26	160	140(0.577)	20	Excellent	
	Domran	4.42	1.27	1.32	452	440(0.678)	12	Excellent	
	Ruaga	4.45	1.02	1.07	245	250(0.899)	5	Excellent	
	Heira	4.46	1.07	1.07	180	150(0.941)	30	V.Good	
FF10-FF12	Etel	4.50	1.30	1.26	49	5(0.826)	44	Good	
	Sheghega	4.52	1.20	1.17	59	39(0.707)	20	V.Good	
	Domran	4.54	1.56	1.20	134	100(0.50)	34	Good	
	Ruaga	4.56	1.43	1.41	240	235(0.953)	5	Excellent	
	Heira	4.58	1.19	1.17	145	144(0.739)	1	Excellent	
FF11-FF13	Small bed	4.60	1.14	1.12	1400	1350(0.973)	50	Good	
	Small bed	4.61	1.14	1.15	1400	1400(0.958)	0	Excellent	

Table 4.7 Comparison of PCAXCOR correlation of real data in the Attahaddy field to the geologic selection. Both the stretch factor and the displacement values are compared to the known geological stretch (thickening and thinning) and displacement in the study area. Values between brackets are the cross-correlation coefficients for displacements.

4.7 Summary

The generalized aspects drawn from the study of the principal components in boundary identification and cross-correlation can be summarized as follows :

- 1- The eigenvalues, eigenvectors and the principal components of all log data are to be determined.
- 2- The first principal component can be used to identify different rock boundaries and for cross-correlation of different formations.
- 3- Use of smoothed principal component is recommended in order to obtain reliable results.
- 4- Differentiation of principal components is necessary to determine a stretch factor, but a more reliable value of displacement is obtained using the original principal components.
- 5- General direction and degree of thickening between two boreholes can be determined using the cross-correlation of power spectra.
- 6- The displacement (correlation position) is obtained using the cross-correlation of stretched series.
- 7- An individual rock type can be identified within the rock unit from the principal component values by characterizing different lithologies in the borehole.

CHAPTER FIVE

Software

5.1 Introduction

The program PCAXCOR is written in FORTRAN 77 and was developed and tested on a Sun3/260 workstation. Data from the Attahaddy field which is written in LIS (Log International Standard) format was read from magnetic tapes using LIS/A Version 1.19 software (Schlumberger, 1988) which was provided by Schlumberger Company. The magnetic tapes were mounted on the VAX/VMS system run by Glasgow University Computing Services. After reading different magnetic tapes, the data are stored in different files, and then transferred to the Sun workstation run by the Geology and Applied Geology Department. Each file contains a complete set of well-log variables for different boreholes and is used by program PCAXCOR (Fig. 5.1).

The main program PCAXCOR calculates the eigenvalues, the eigenvectors and the variance-covariance or correlation matrix of different well-log variables. This matrix is then used to calculate the principal

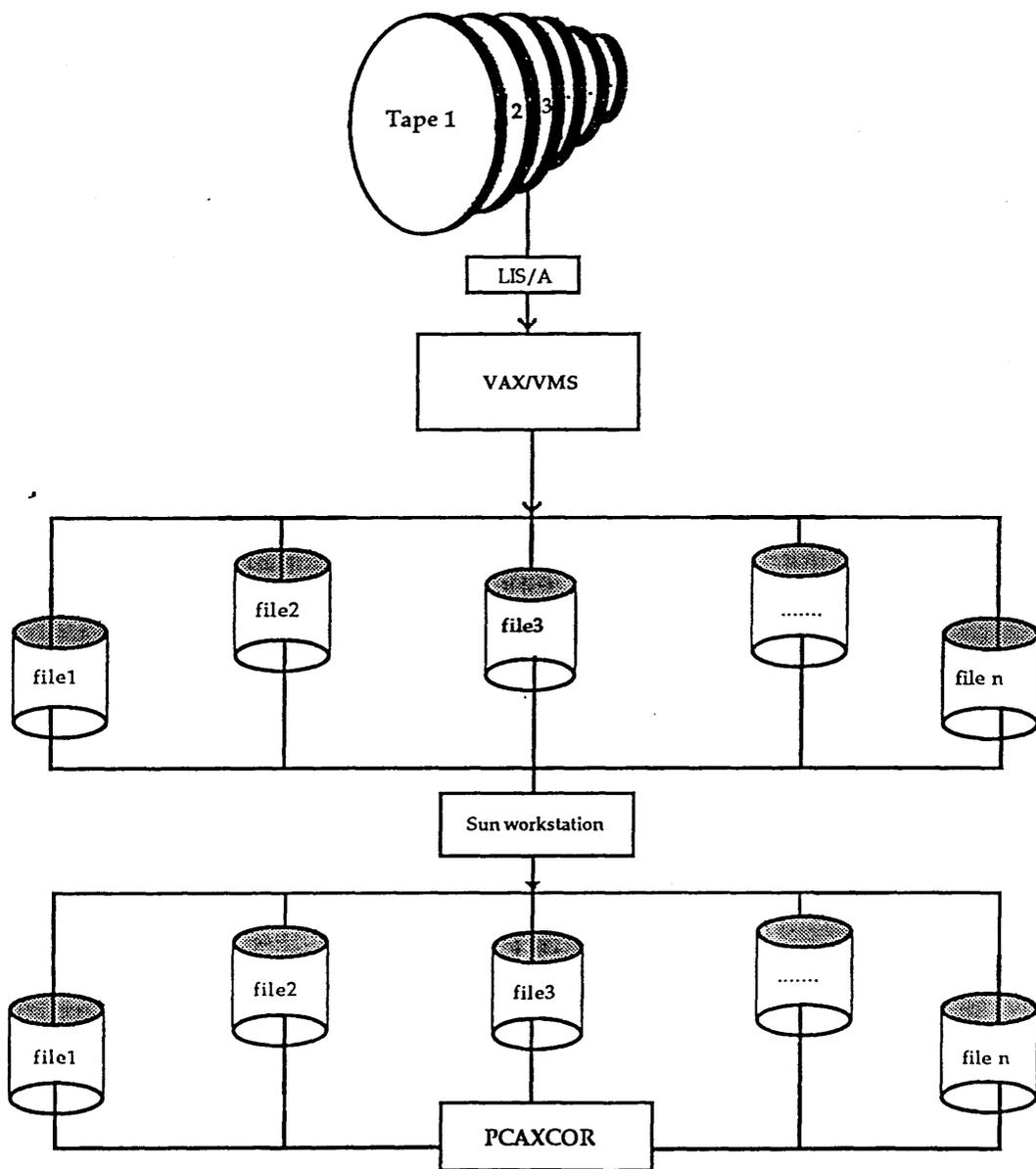


Fig. 5.1 Diagram illustrates procedures of reading well-log tapes using LIS/A software.

component scores from which, after being filtered, boundaries of different formations are identified. Cross-correlation is then applied to moving average filtered principal components to identify the stretch factor (thinning and thickening) of strata, and the displacement between two principal components of two different boreholes.

The subroutines in PCAXCOR program are divided into two groups:

The first group of subroutines (Appendix C) performs the principal component analysis. This includes standardizing the original variables (*STANDARIZE* subroutine), calculating the variance-covariance or correlation matrix (*MULTV7* subroutine), calculating the eigensystem, and calculating the principal component scores (*TRED2* & *TQLI* subroutines). The last two subroutines are those of Press (1988). Subroutine *SMOOTH* filters the first principal component for later use in the boundary identification and cross-correlations. The identification of different formation boundaries is performed using subroutine *BOUNDARY*. Some other utility subroutines are written to output the final results with appropriate format. These are subroutine *EIGENVALUE* which is used to calculate the percentage of the eigenvalue to the total variance, and subroutine *ORGANIZE* which compiles the results output from the first group of subroutines.

The second group of subroutines (Appendix C) performs the cross-correlation between two principal components of two different boreholes. All these subroutines are derived from Kwon (1977) and slightly modified to

suit the requirements of the this project. The main subroutine (*XCOR*) calls the remaining subroutines which perform different tasks, including the determination of the derivatives of the principal components, the interpolation of the data, the calculation of the Fourier transforms, the cross-correlation of power spectra and the stretched series, and scanning for the best stretch factor among the cross-correlation coefficients.

The graphical routines are written using the sophisticated and interactive S language. This is a very powerful and flexible tool for manipulation, analysis and graphical display of data (Farrow, 1991). S also provides a simple interface to the Unix system (Sun workstation) from which different output files from *PCAXCOR* are read interactively to the S system (Fig. 5.2). A number of routines or functions (Becker *et al*, 1988) are written for different graphical outputs (Appendix D).

5.2 Program structure

5.2.1 Calculations programs

The main program *PCAXCOR* (Fig.5.3) utilizes call to all the following subroutines :

STANDARIZE subroutine:

This subroutine is used to standardize the original data matrix (Y), which contains well-log variables, to standard form so it will have a mean

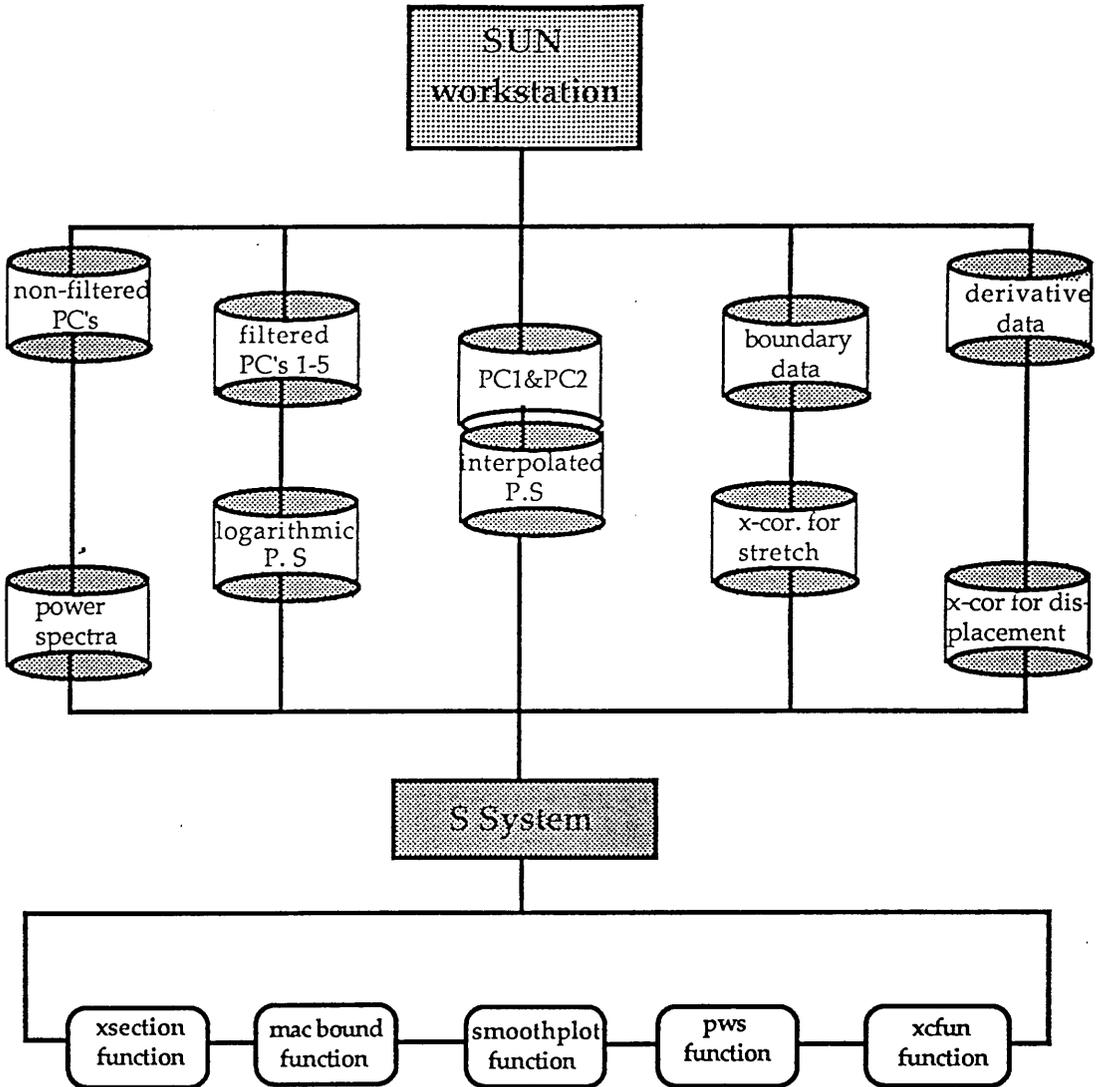


Fig. 5.2 Distribution of files in the Sun workstation and the functions (programs) in the S system.

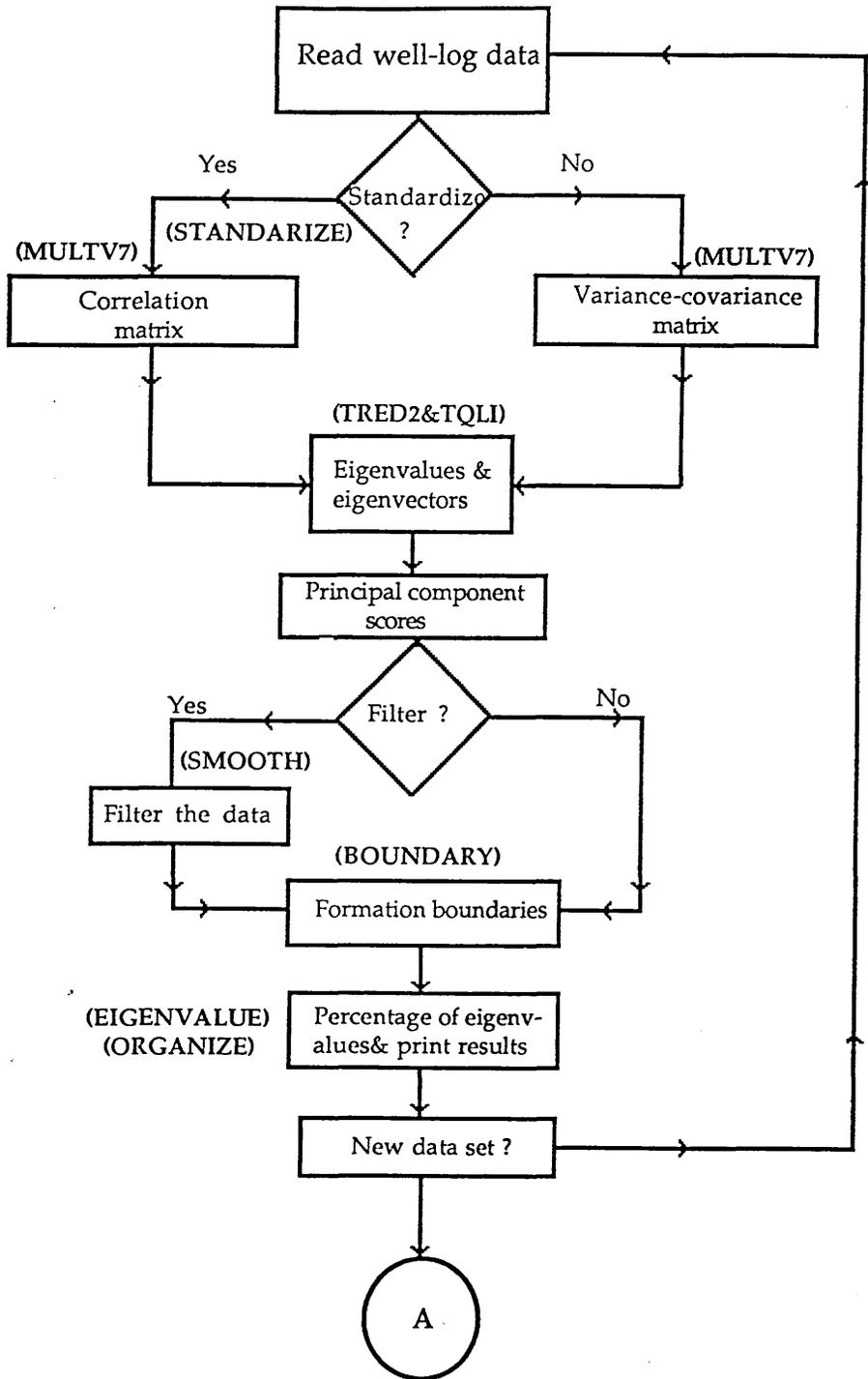


Fig. 5.3a Flow chart of the main program PCAXCOR.

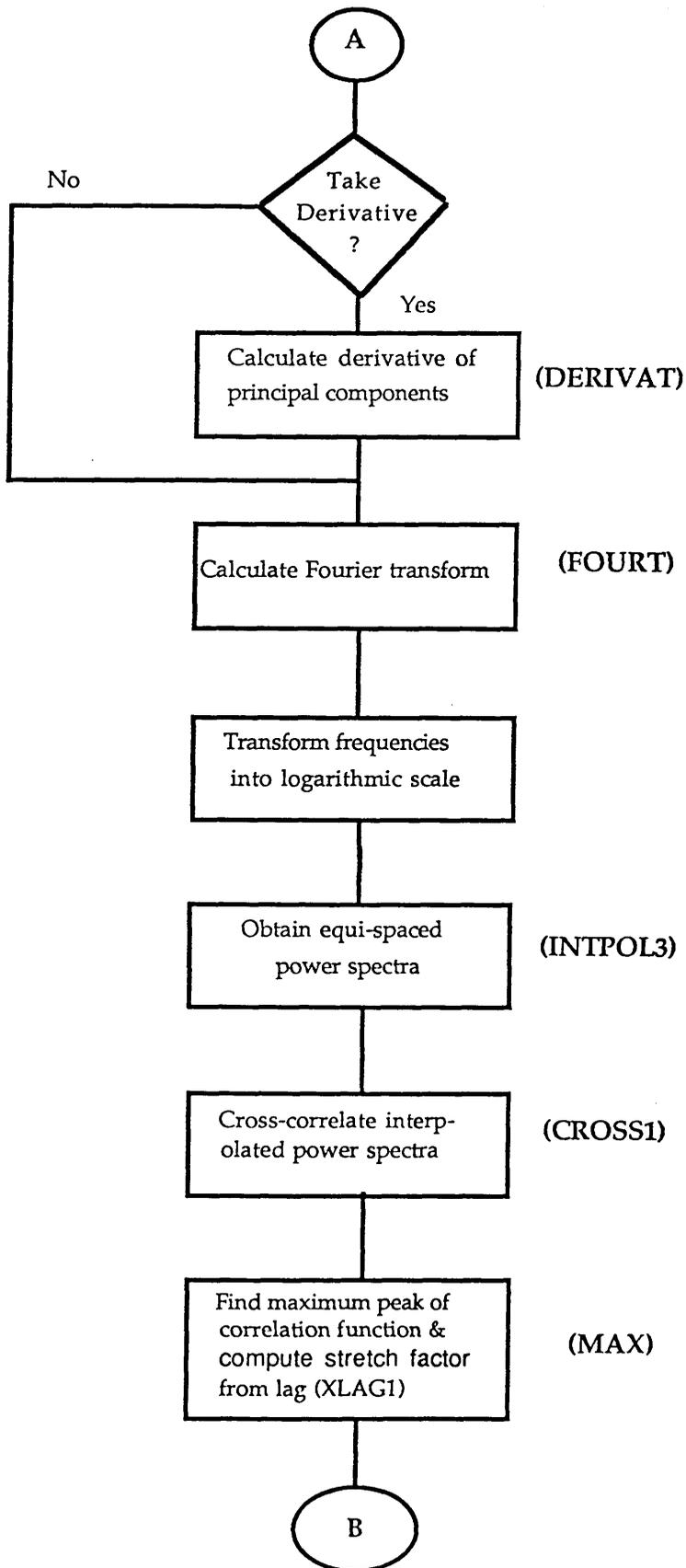


Fig. 5.3b Flow chart of the mainprogram PCAXCOR.

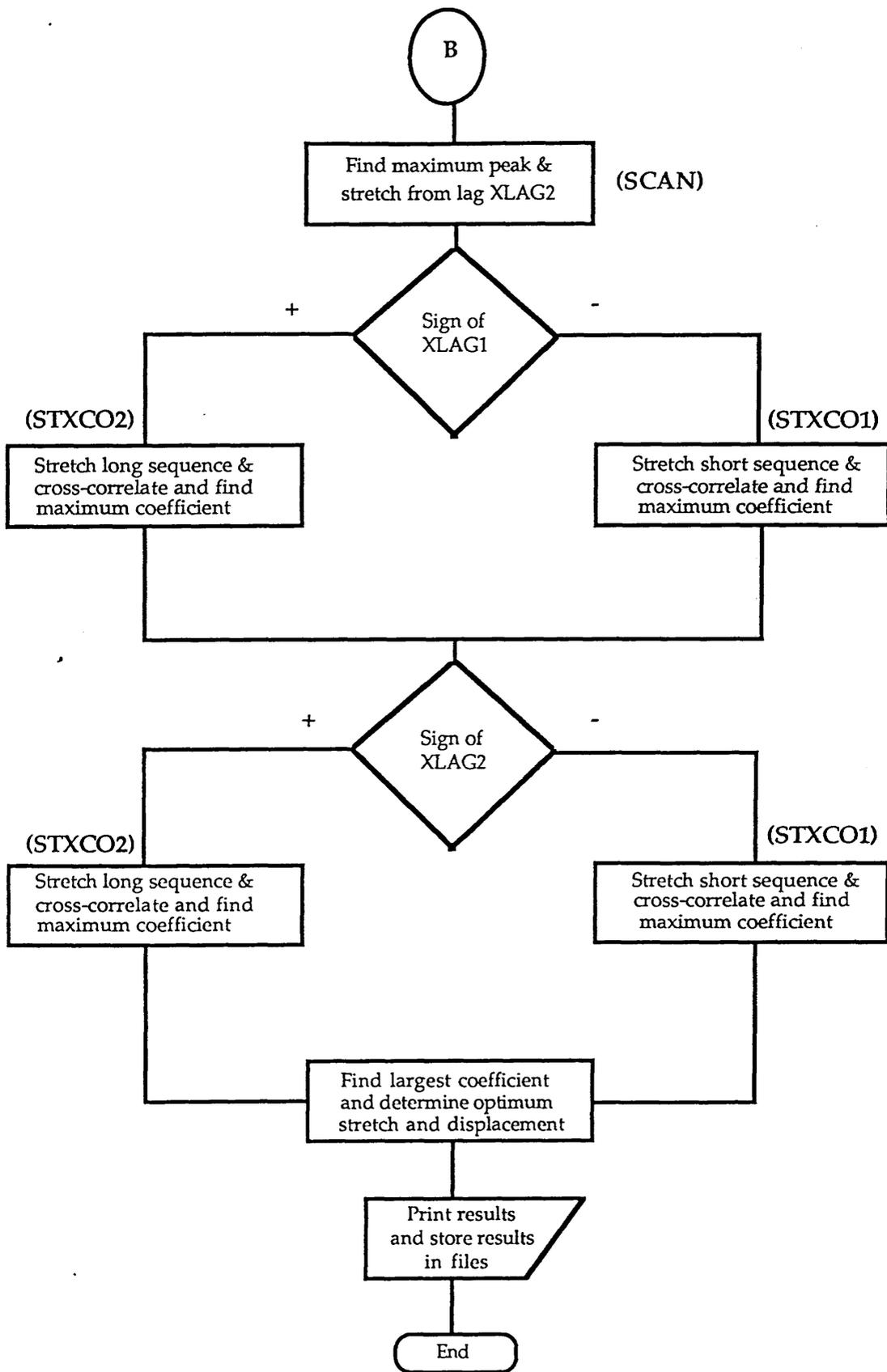


Fig. 5.3c Flow chart of the main program PCAXCOR.

of 0.0 and variance of 1.0. Using the data matrix (Y), the subroutine returns the standardized form (STANDZ) which contains the standardized well-log variables of size a number of columns (NCOL) by a number of rows (NROW). Any size of a matrix can be used.

MULTV7 subroutine

To calculate the variance-covariance or correlation matrix of the original matrix. It uses the standardized data matrix (STANDZ) which has been derived from subroutine STANDARIZE and returns the square symmetric matrix of variance-covariance or correlation matrix (C), which has as many columns and rows as there are columns in the STANDZ.

TRED2 subroutine

Calculates the tridiagonal matrix of the square matrix (C). From *TRED2* come two vectors (D) and (E) which are the diagonal and off-diagonal elements of the input matrix (C). It returns the diagonal (D) and off-diagonal (E) elements of the tri-diagonal matrix (TRI).

TQLI subroutine

D and E from *TRED2* are replaced by the eigenvalues and the corresponding eigenvectors respectively. It returns an $n \times n$ matrix (A) which contains the eigenvectors of the square matrix.

EIGENVALUE subroutine

To calculate the percentage each eigenvalue contributes to the total variance. It returns a one-dimensional vector (PERC) which contains the percentages.

ORGANIZE subroutine

To organize the final output and write it to a file. It tabulates the eigenvalues, eigenvectors and the percentage of each eigenvalue, using the matrix (A) containing the eigenvectors D, one dimensional vector of the eigenvalues and PERC is one-dimensional vector containing the percentage of each eigenvalue.

SMOOTH subroutine

This subroutine is to smooth the first principal component using a moving average filter. It read a vector (PP) which contains the principal component scores and outputs the filtered principal component vector (SMOOTHD). Filtering is optional.

BOUNDARY subroutine

This subroutine is to identify the boundaries of different formations using the Mahalanbis D^2 technique (Davis, 1986). It takes the filtered principal components vector (SMOOTHD) and returns its D squared values.

The maximum peaks of these values represent the positions of formation boundaries.

XCOR subroutine

The subroutine *XCOR* controls the calculation of the cross-correlation between two sequences. It utilizes two cross-correlation processes to determine the stretch factor and the relative displacement. Cross-correlation of power spectra of two sequences identifies the direction and amount of stretch between two series. Cross-correlation of the stretched sequences identifies the relative displacement between these sequences. Outputs consist of a list of the input data, coefficients of the cross-correlation of power spectra and the optimum stretch and displacement values (Appendix E). Other output, for example, derivative data, interpolated power spectra, etc are stored in different files for later graphical manipulation in the S system. Subroutine *XCOR* and all subroutines it is calling are modified from *SPECOR* program (Kwon, 1977). The following subroutines are called by *XCOR*.

DERIVAT subroutine

To replace the principal components by their first derivatives. This is an optional step before calculating the Fourier transforms. *DERIVAT* subroutine takes the first principal component of the short sequence (RLOG1) and the first principal component of the long sequence (RLOG2) and calculates their derivatives, RLOG1 and RLOG2 respectively.

FOURT subroutine

This subroutine calculates the Fourier transforms of the short and long sequence. It takes either the original data (filtered principal components) or the derivative data form of the data and returned their Fourier transform CLOG1 and CLOG2 which are used by subroutine *XCOR* to calculate the power spectra of the two series.

INTPOL3 subroutine

This is to obtain equally spaced power spectra using Lagrange interpolation method. It interpolates power spectra of the short series (RLOG1) and the power spectra of the long series (RLOG2) and returns their interpolated values in Y1P1 and Y1P2 respectively.

CROSS1 subroutine

This is used to cross-correlate between the interpolated power spectra to obtain the stretch factor. The interpolated power spectra of the short and long sequences Y1P1 and Y1P2 from subroutine *INTPOL3* are used by this subroutine. The first call to *CROSS1* by subroutine *XCOR* is to cross-correlate between the interpolated short sequence (Y1P1) and the interpolated long sequence (Y1P2), and cross-correlation coefficients are stored in a vector XCORL. The second call to *CROSS1* is to cross-correlate between the interpolated long sequence (Y1P2) and the interpolated short sequence (Y1P1). The resultant cross-correlation coefficients are stored in XCORS.

Both vectors (XCORL & XCORS) are stored in the vector Y1P1 with a length set equals to the maximum expected stretch factor (S=2).

MAX subroutine

To find the maximum peak in the cross-correlation function of power spectra and compute the corresponding stretch factor. It takes the vector containing the cross-correlation coefficients (Y1P1) and returns its maximum value (PCMAX1).

SCAN subroutine

To find the second peak in the cross-correlation function of power spectra and compute the corresponding stretch factor. The vector Y1P1 which contains the cross-correlation coefficients is input to subroutine *SCAN* which returns the second maximum value of the cross-correlation function of power spectra.

STXCO1 subroutine

This is used to stretch and correlate the first peak of the cross-correlation function assuming that the long series is stretched. It makes calls to *STRETCH*, *CROSS2*, and *MAX* subroutines. It passes the short sequence (RLOG1) to subroutine *STRETCH*. RLOG1 after being stretched (CRLOG1) is passed to subroutine *CROSS2* for cross-correlation with the long sequence (RLOG2), and the maximum value of the cross-correlation function is

determined using subroutine *MAX*.

STXCO2 subroutine

This is used to stretch and correlate the first and the second peak of the cross-correlation function assuming that the short series is stretched. It makes calls to *STRETCH*, *CROSS2*, and *MAX* subroutines. It passes the long sequence (RLOG2) to subroutine *STRETCH*. RLOG1 after being stretched (CRLOG2) is passed to subroutine *CROSS2* for cross-correlation with the short sequence (RLOG1), and the maximum value of the cross-correlation function is determined using subroutine *MAX*.

STRETCH subroutine

This subroutine is to interpolate a time series data with N values to a series with M values in the frequency domain. It makes a call to *FOURT* subroutine to invert the Fourier transforms. The short (RLOG1) and the long (RLOG2) sequences are used by this subroutine, which returns the stretched version of RLOG1 and RLOG2 in vectors CLOG1 and CRLOG2 respectively.

5.2.2 Graphical functions

The graphical functions (Appendix D) which are written in the S language and are used to display the results are:

xsection :

This is used to generate a cross-section in the Attahaddy field using the geological formation tops.

macbound :

This plots the non-filtered first components on one side and the boundaries of different formation on the other side of the diagram.

smoothplot :

To plot the filtered first principal components of the used well-log data along with their window sizes which are used for the cross-correlation process.

pws :

This function is to draw the short and long sequences, the derivative data, the power spectra, the equi-spaced power spectra, the interpolated power spectra, the cross-correlation function of power spectra, and the cross-correlation of the stretched series.

xcfun

To plot the short and the long sequence along with the size of the window used in the cross-correlation process, the cross-correlation function of power spectra, and the cross-correlation function of the stretched series.

prinplot :

To plot different filtered principal components of the studied wells.

CHAPTER SIX

Conclusions

The application of principal component analysis to well-log data was established for better understanding of subsurface geology. By using principal component scores formation boundaries can be identified and well-to-well correlation performed. Because principal component scores contain most of the variance of the original matrix, it is easier to handle these scores than to use all the variables of the raw data matrix.

Conventional well-log variables (Spontaneous potential, Gamma Ray, etc) are often used for boundary identification as well as cross-correlation between different wells. The first principal component of these variables is found to be appropriate for an automatic process to identify formation boundaries, and cross-correlation between the chosen boreholes in the Attahaddy field. Different well-log variables are expressed in different units of measurement. For such a case, the correlation matrix is used for the calculation of the principal component scores. This implies that all variables are expressed in dimensionless form to reduce the affect of variables whose mean is large and inflate variables whose mean is small. This is necessary if

the units of measurement of different well-log variables are not the same. A decrease in the amount of the percentage of the eigenvalues is inevitable for this case.

Filtering the principal component using a moving average filter is necessary before the identification of formation boundaries is obtained to reduce the affect of thin beds. A window size equal to half the expected thickness of the formation is found to to suitable to identify formation boundaries. Although the formation thickness in the Attahaddy field varies from one borehole to another, an average window size of 150 feet gave appropriate results in the study area. If small beds are of interest a smaller window size is appropriate to identify the formation boundaries of these beds.

The filtered principal components are again used for well-to-well correlation. Filtering these components is found to be necessary before proceeding with any correlation. Non filtered scores are tested for correlation and were not as good as the filtered scores. The cross-correlation technique between a number of boreholes in the Attahaddy field is based on spectral analysis of the filtered first principal component scores. Two cross-correlation functions are used. The cross-correlation function of power spectra which gives the degree and direction of stretch, and the cross-correlation function of the stretched principal components which determines the displacement between two formations

Use of principal component scores in well-to-well correlation is a new technique which has many advantages over the previous studies. An important advantage of this approach is the elimination of the noise effect. This is achieved by using moving average filtered scores in the time domain for the cross-correlation function of scores to determine the displacement, and the derivative filter in the frequency domain for the cross-correlation function of power spectra to determine the stretch factor.

Good agreement from this method was obtained for the geological formation boundaries and cross-correlation between boreholes in the study area. Although the geological stretch and displacement vary widely in the Attahaddy field, both the stretch factor (thinning and thickening of strata), and the displacement (correlation position) of the studied wells which were obtained using program PCAXCOR coincide with the known geological stretches and displacements (Fig. 4.63 and Table 4.7). Furthermore, the program can be used to correlate different rock types within the same rock unit and can also be used to identify the general lithological character of formations in the boreholes.

Another advantage of using the first principal component is an increase in the magnitude of the correlation coefficient in well-to-well correlation over the previous studies. In his technique, Kwon (1977) obtained an average value of the cross-correlation coefficients of 0.50 which was considered to be excellent when the original variables of well-log data were used. In this study, the average value of the coefficient is 0.85. This

implies that a very distinctive linear relationship exists between each pair of first principal components used in the correlation processes. The increase in the correlation coefficient is the result of using the first principal components which reduces the noise problems. Kwon (1977) concluded that the reduction in the magnitude of the cross-correlation coefficient was due to the presence of uncorrelatable noise signals in the variables to be correlated. High noise variable gave spurious stretch factors and, hence, wrong displacements. He concluded that care must be taken when dealing with high noise components. In the PCAXCOR program, the noise effects are controlled by using the filtered first principal components before (moving average filter) and after (high-pass filter) the correlation is made between two sequences.

In addition, this study demonstrates the advantage of employing the filtered first principal components over the non-filtered first principal component or the original well-log variables. This approach of using the filtered first principal component in boundary identification and well-to-well correlation is found to be precise, reliable, and gives accurate results related to the geological boundaries and known correlation of the area of study. There are a few occasions when the computer selection does not agree with the manual geological selection. Such deviations occur because PCAXCOR identifies the boundaries at the inflection point between two formations by hunting for a boundary at the abrupt change in the average values of the sequence. However, some of the formation boundaries in the study area do not occur at such a point; for example, the top of the Domran Formation. This illustrates one of the fundamental limitations of the

automated identification of rock boundaries.

The deviation of computed correlation from the known geological correlation in some cases, is explained by the fact that the mathematical correlation by PCAXCOR is made assuming that there is an average similarity between the sequences to be correlated. However, this is not always true in the Attahaddy field, for example the correlation between the Domran formations in well FF10 and well FF12. This is another limitation of using this method.

Nevertheless, this new technique is a useful addition to the current manual methods of boundary identification and well-to-well correlation. The new complete framework of software opens a new era in well-log interpretation and may be made more reliable by further refinement of this method to include automatic identification of rock types from the first principal components.

REFERENCES

Barr, F.T. and Wagar, A.A. 1972. Stratigraphic nomenclature of the Sirte Basin, Libya : *Petroleum Exploration Society of Libya*, Graficle Trevisan Castle Franco, Italy, 179 p.

Bartlett. M. S. 1948. Smoothing Periodogram from time series with contineous spectra: *Nature*, v. 161, p. 686-687.

Becker, A. R., Chambers, M. J. and Wilks, R. A. 1988. *The NEW S Language: A programming Environment for Data Analysis and Graphics*. Wadsworth & Brooks/Cole Advanced Books & Software, California.

Cain, M. 1985. Sirte Oil Company internal report.

Contant, L. C. and Goudarzi, G. H. 1967. Stratigraphic and tectonic framework of Libya: geology and Oil fields of Libya, Algeria and Tunisia, *Am. Assoc. Petroleum Geologists Bull.*, Foreign repe. ser., 1, p. 5-16.

Colley, J. W., and Tukey, J. W. 1965. An algorithm for machine calculation of complex Fourier series: *Math. comput.*, 19, p. 297-301.

Daskman, R. 1964. Automated Well log Analysis and the evaluation of sand-stone bodies in shale sequences: *Col. School. of Mining Quart.*, v. 59, no. 4, p. 517-536.

Davis, J. C. 1986. Statistics and Datta Analysis in Geology: second edition, *Kansas Geological Survey*. John Wiley & Sons In, Inc., New York, 646p.

Duronio, P. & Colombi. L. 1983. Mesozoic rocks of Libya: *Spec. Pap. Petroleum Exploration Society Libya*, S-1-S12.

El-talhi, M. M. 1990. The Geology and Hydrocarbon Potential of the Sirte Basin-Libya: Oxford Polytechnic [High Diploma Thesis], 48p.

Elek, I. 1986. Some Application of the Principal Component Analysis in well-log interpretation: *Hungarian Geophysics*, no. 1,

Elek I. 1988. Some Application of Principal Component Analysis: Well-to-Well correlation, zonation: *GEOBYTE*, p. 46-55..

Elek I. 1990. Fast Porosity Estimation by Principal Component Analysis: Hungarian Hydrocarbon Institute, *GEOBYTE*, v. 5, no. 3, p. 25-34.

Farrow, C. M. 1991. S- Geological Data Analysis and Graphics. *Terra Nova* 3(2) pp 212-214.

Gill, D. 1970. Application of a Statistical Zonation Method to Reservoir Evaluation and Digitized-log Analysis: *The Am. Assoc. Petroleum Geologists Bull.* v. 54, no. 5, p. 719-729.

Haites, T. B. 1963. Perspective correlation: *Am. Assoc. Petroleum Geologists Bull.*, v. 47, no. 4, p. 553-574.

Hawkins, D. M. 1976. Fortran IV program to segment multivariate sequence of data: *Computer & Geosciences*, v. 1, no. 4, p. 339-351.

Hawkins, D.M. and Merriam D. F. 1973. optimal zonation of digitized sequential data: *Jour. Int'l. Assoc., Mathematical Geology*, v. 5, no.4, p. 389-395.

Hawkins, D.M. and Merriam D. F. 1974, zonation of multivariate sequences of digitized geological data: *Jour. Int'l. Assoc., Mathematical Geology*, v. 6, no.3, p. 263-269.

Henderson, G. J. 1973. Correlation and Analysis of Geological Time Series [Ph.D. thesis]: Bloomington, Indiana Univ., 289p.

Jenkins, G. M., and Watts, D. G. 1968. Spectral analysis and its application: San Francisco, Holden-Day, Inc., 525p.

Klitzsch, E. 1971. The structure development of parts of North Africa since Cambrian time: Gray C. (ed) *1st Symposium Geology of Libya*. Univ. Libya, Fac. Sci., p. 253-262.

Kwon, B, 1977. Spectral analysis of geophysical log for correlation [Ph.D. thesis]: Bloomington, Indiana Univ., 255p.

Kwon, B. & Blakely, R. F., and Rudman, A. J. 1978. Fortran Program for Correlation of Stratigraphic Time Series: Part 2. Power Spectral Analysis, Dept. of Natural Resources, *Indiana Geological Survey occasional paper 26*, 50p.

Matuszak, D. R. 1972. Stratigraphic correlation of subsurface geological data by computer: *Mathematical Geology*, v. 4, no.4, p. 331-343.

Merriam, D. F. 1971 Computer Application in Stratigraphic problem solving in decision making in the mineral industry: *Canadian Inst. Min. Metal.*, v. 12, p. 139-147.

Neidell, N. S. 1969. Ambiguity function and the concept of geological correlation: *Kansas Geol. Survey Contr.*, v. 40, p. 19-29.

Parson, M. G., Zagaar, A. M., and Curry, J. J. 1980. Hydrocarbon occurrence in the Sirte basin, Libya: Facts and Principals of World Petroleum Occurance (ed. A. Miall). *Mem. Can. Soc. Petrol. Geol.*, 6, p. 723-732.

Press, H., Flannery, P. B. 1986: Numerical Reciepes, The Art of Scientific computing, Cambridge Univ. Press, 818p.

Reyment, R. A. 1966 Studies on Nigerian Upper Cretaceous and Lower Tertiary Ostracoda, part 3, *Stratigraphical, Paleocological and Biometrical conclusion*. *Stockh. Contr. Geol.*, v. 14, 151p.

Rudman, A. J., and Blakely, R. F. 1976, Fortran program for correlation of stratigraphic time series: *Indiana Geol. Survey Occasional paper 14*, 31p.

Rudman, A. J., Blakely, R. F. and Henderson, G. J. 1975. Frequency domain methods of stratigraphic correlation: *Offshore Technology Conf. paper OTC 2269*, p. 256-269.

Rudman, A. J., and Lankston, R. F. 1973. Stratigraphic correlation of well logs by computer techniques: *Am. Assoc. Petroleum Geologists. Bull.*, v. 57, p. 577-588.

Schlumberger. 1974. Log Interpretation, Applications, volume II-Applications.

Schlumberger. 1981. Customer tape format document.

Schlumberger. 1988. How to use LIS/A: Schlumberger computer product, 2nd edition.

Schlumberger. 1988a. Schlumberger computer products, How to use LIS/A manual, 2nd edition.

Webster, R. 1973 Automatic Soil-Boundary Location from Transect Data: *Mathematical Geology*, v. 5, no. 1.

Webster, R. 1980. A Fortran IV program for segmentation multivariate one-dimensional series: *Computer & Geosciences*, v. 6, no.1, p. 61-68.

Weiner, N. 1949. Exploration, Interpolation and Smoothing of Stationary Time Series, with Engineering application: Cambridge, M. T. Press, p. 163.

APPENDIX A**Borehole information**

The following is some information about the boreholes in the Attahaddy field. This includes the location of the field (Fig. 1), the longitude and latitude, the elevation to the Kelley bushing (KB), the classification of each well ... etc.

WELL FF2-6

Location : 3.7 km SE of FF1-6

Coordinates : 29 33' 28" N : 19 38' 54" E

Elevation KB : 297

Classification: Exploration outpost .

Spudded : 6 Oct. 1967

Completed : 13 Nov. 1967

Completion status : Dry and abandoned .

Total depth : 10,035'

Plugged back T.D : to the surface .

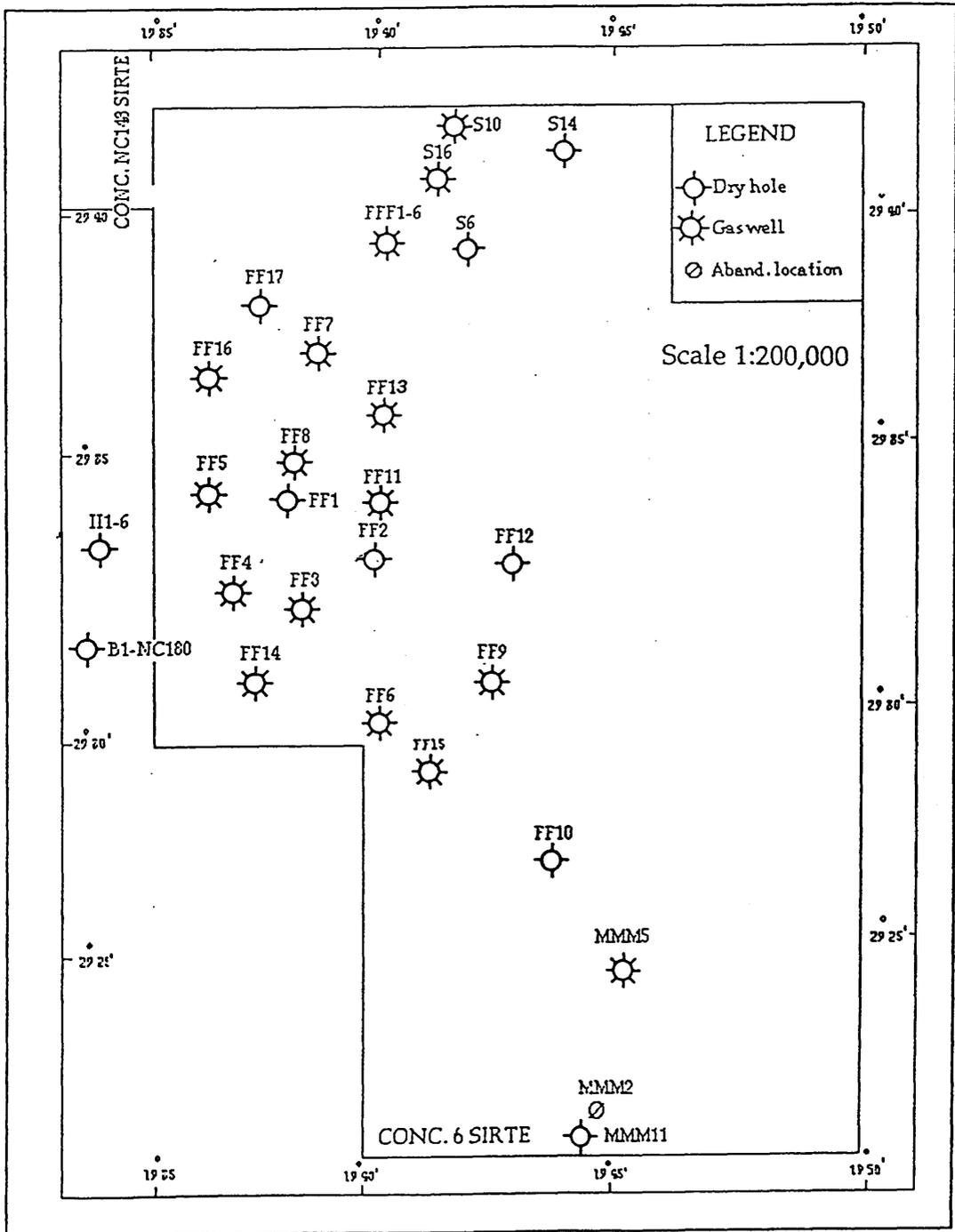


Fig. 1 Location map of the boreholes in the Attahaddy field.

WELL FF3-6

Location : 2.8 km S- SE of FF1-6

Coordinates : 29 32' 52.9" N : 19 37' 41.4"E

Elevation KB : 362'

Classification : outpost .

Spudded : 24 Apr. 1985 Completed :30 Jul. 1985

Completion status : Gas well

Total depth : 12,104'

A total of 10 DST's and production test have been run and have showed a flow rate of 20.2 MMCFG/D with bottom hole pressure of 2995 psi .

WELL FF4-6

Location : SP 430 seismic line V6-27-84

Coordinates : 29 33' 01.059" N : 19 36' 12.946" E

Elevation KB : 362

Classification : outpost .

Spudded : 24 Oct. 1985 Completed : 12 Feb. 1986

Completion status : Gas well

Total depth : 11,170'

This well was drilled with no major problems encountered except lost circulation all the way down from the top of the Gargaf Formation to T.D .

A total of 5 DST's run on the Gargaf Formation showed good reservoir characteristics . A 24 hours flow test has been run , 3320 psi on 1/4" choke pressure with flow rate of 4.58 MMCFG/D have been recorded .

WELL FF5-6

Location : Seismic line V25-84 , SP 230

Coordinates : 29 34' 29.039" N : 19 35' 24.290"E

Elevation KB : 350'

Classification : outpost .

Spudded : 20 Feb. 1986

Completed : 5 May 1986

Completion Status : Gas well .

Total depth : 11,214'

Plugged back T.D : 10,800'

A total of 6 DST's have been run, 3 have failed, fresh water was recovered (cl- 8600 ppm) in the last test .

Maximum surface head pressure 597 psi , BHT : 322 degree F .

WELL FF6-6

Location : Seismic line 192-84 , sp 280

Coordinates : 29 30' 26.818" N : 19 40' 14.216"E

Elevation KB : 328'

Classification : exploration-outpost .

Completion status : Gas well .

Total depth : 12,065'

One DST has been run . It displayed a bottom hole pressure of 3392 psi . Three cores have been cut .

WELL FF7-6

Location : Seismic line 6V-27-84 , sp 300

Coordinates : 29 36' 21.310"N : 19 37' 38.574"E

Elevation KB : 137'

Spudded : 23 Sep. 1986

Completed : 13 Jan. 1987

Classification : outpost .

Completion status : Gas well .

Total depth : 12,594'

Plugged back TD : 12,400'

A total of 5 DST's were run, showing a flow rate of 32 MMCFG/D and surface pressure of 2500 psi has been recorded . Three cores were cut .

WELL FF8-6

Location : Seismic line 6V-27-84 , sp 370

Coordinates : 29 34' 33.508"N : 19 36' 57.339"E

Elevation KB : 287'

Classification : development .

Spudded : 16 Jun. 1986

Completed : 2 Act 1986

Completion status : suspended gas well .

Total depth : 12,018'

Seven DST's accomplished, one core was cut. Average flow rate 9 MMCFG/D .

WELL FF9-6

Location: Seismic line 6V -32- 84 , sp 371

Coordinates : 29 30' 54.190"N : 19 41' 51.632"E

Elevation KB : 283'

Classification : outpost

Spudded : 09 Act. 1986

Completed : 21 Jan. 1987

Completion status : Suspended as non-commercial gas well .

Four DST's have been run, displaying no commercial hydrocarbons . A 20 hours production test was run, indicating 50 psi mean surface pressure on 3/4" choke, flow rate 820 MCFG/D .

WELL FF10-6

Location : Seismic line 6V 218 -E 85

Coordinates : 29 27' 23.954"N : 19 42' 50.183"E

Elevation KB : 376'

Classification : Exploration wildcat

Spudded : 23 Jan. 1987

Completed : 15 Apr. 1987

Completion status : Dry and Abandoned .

Total depth : 12,630' Plugged back T.D : to surface .

Two DST's were run with no hydrocarbon show. Three cores have been cut, indicating that the formation was tight and dense .

WELL FF11-6

Location : Seismic line

Coordinates : 29 34' 05.21"N : 19 39' 05.37"E

Elevation KB : 248'

Classification : outpost

Spudded : 22 Jun. 1987 Completed : 18 Jul. 1987

Completion status : Gas well

Total depth : 12,753'

A total of 4 cores have been cut, 4 DST's were run with average estimated of flow rate of 12-13 MMCFG/D .

WELL FF12-6

Location : Seismic line 6V 206- 85 X , sp 1810

Coordinates : 29 33' 01.495"N : 19 42' 06.320E

Elevation KB : 239'

Classification : outpost exploration .

Spudded : 24 Jul. 1987 Completed : 03 Jan. 1987

Completion status : dry and abandoned .

Total depth : 12,607'

Four cores were cut and two DST's were run, did not show any reservoir characters .

WELL FF13-6

Location : Seismic line 6V 31 -83 , sp 340

Coordinates : 29 35' 15.815" N : 19 39' 26.993"E

Elevation KB : 190'

Spudded : 15 Oct. 1987 Completed : 30 Jan. 1988

Classification: outpost .

Completion status : Suspended as Gas well .

Total depth : 12,524' Plugged back T.D : 12,225'

A total of five DST's were run and production test was run to confirm that there was no potential reservoir in that block .

WELL FF14-6

Location : Seismic line 6V-29-84 , sp 280

Coordinates : 29 31' 28.983"N : 19 37' 13.327"E

Elevation KB : 375'

Classification : Exploratory .

Spudded : 16 Dec. 1987

Completed : 12 Apr. 1988

Completion status : Dry and abandoned .

Total depth : 13,426'

Plugged back TD : 814'

Two DST's were run showed no hydrocarbon accumulation in the block . Three cores have been cut, showing a very tight formation.

WELL FF15-6

Location : Seismic line 6V-216-ext 85 , sp 955

Coordinates : 29 21' 42.44"N : 19 40' 59.98"E

Elevation KB : 328'

Classification : outpost .

Spudded : 18 Apr. 1988

Completed : 22 Aug. 1988

Completion status : Suspended gas well .

Total depth : 12,638.5'

Plugged back T.D : 11,270'

A total of 4 DST's were run , followed by a production test .

The estimated flow rate is 8 MMCFG/D .

APPENDIX B

Well logging principles

The important variables used in downhole well-logging, and used in this project are described. There are different terms of well-log tools for different companies. The names of the tools used through out this study, for example, LDL, CNL, BHCetc are the Schlumberger company trade mark.

Spontaneous Potential (SP)

The Spontaneous Potential (SP) curve is a measure of the difference between the potential of a movable electrode in the borehole (Fig.1a) and the fixed potential of a surface electrode. The unit used is the millivolt. Spontaneous potential is used to:

- 1- detect the permeable beds,
- 2- locate their boundaries and to permit correlation of such beds,
- 3- determine values of formation water resistivity R_W (Fig.2),
- 4- gives qualitative indication of bed shaliness.

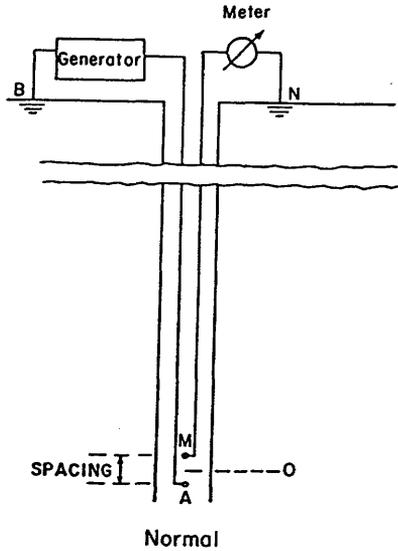
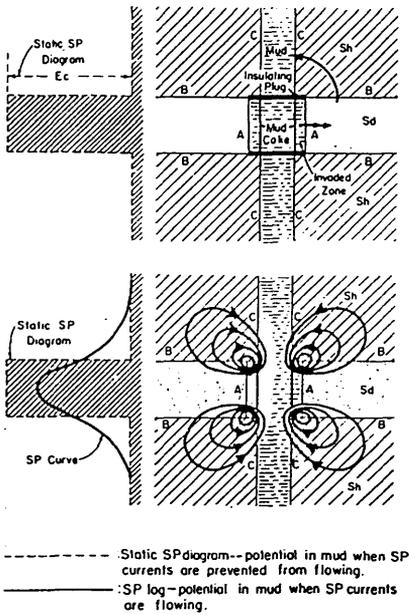


Fig. 1a : Schematic representation of potential current distribution in and around the permeable bed (Schlumberger).

Fig. 1b : Normal resistivity device (Schlumberger)

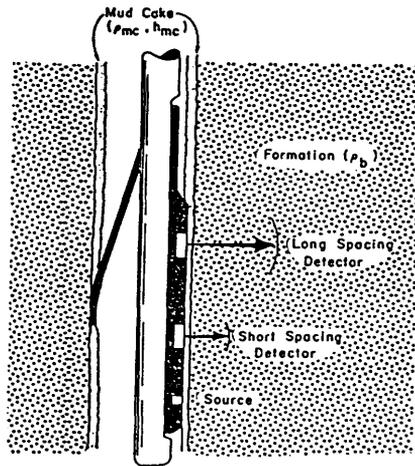
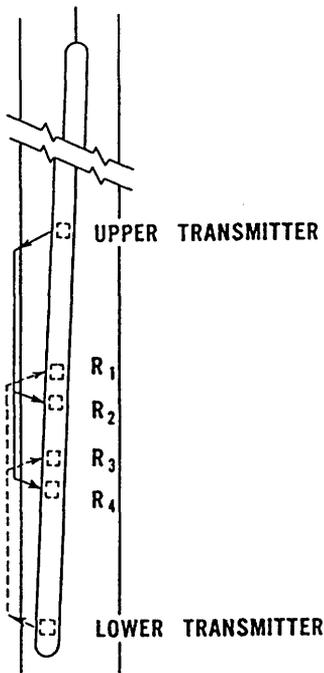


Fig. 1c : Schematic diagram of BHC Sonde (Schlumberger).

Fig. 1d : Schematic drawing of the dual spacing Formation Density Logging device (LDL).

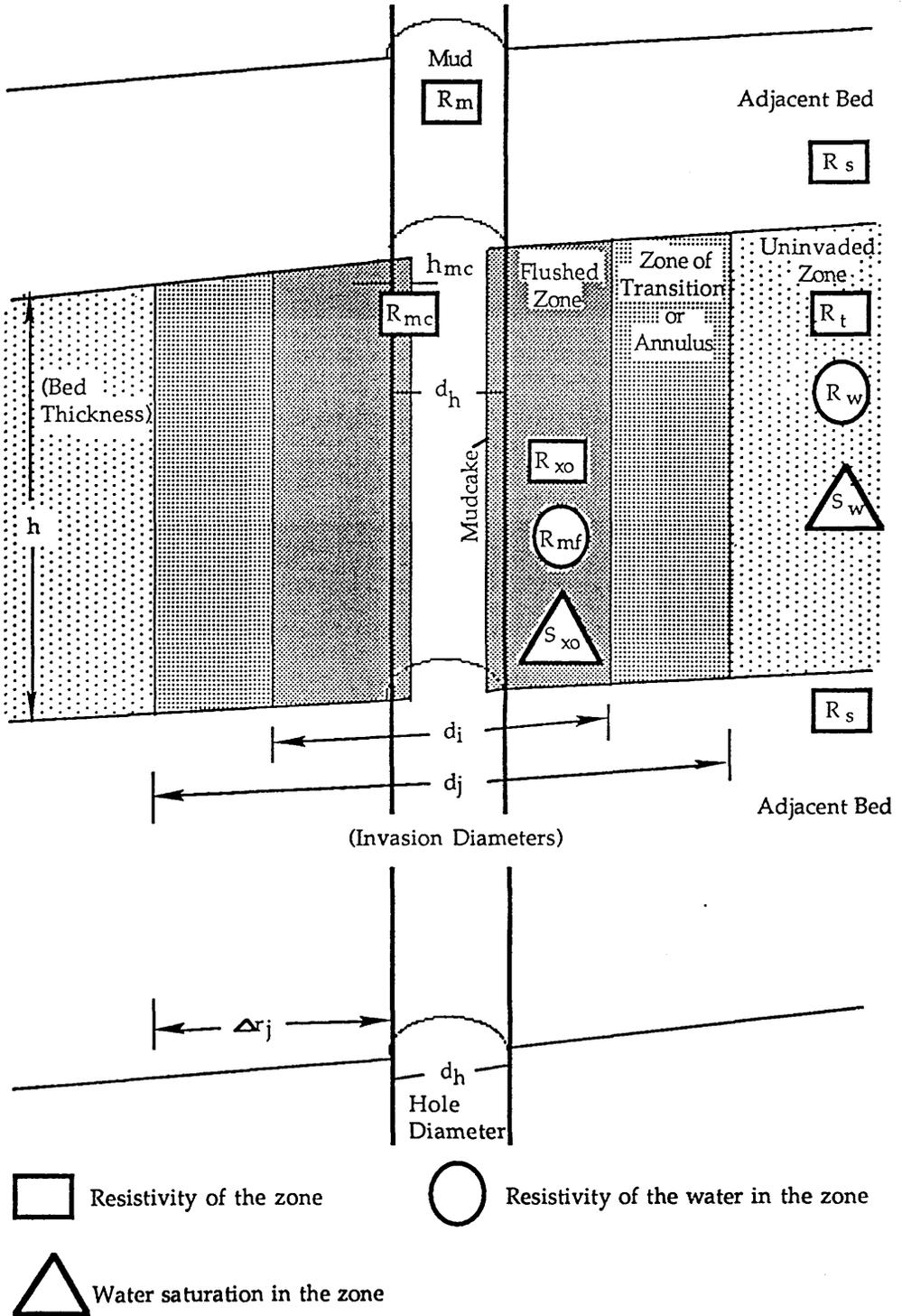


Fig. 2 Symbols used in Log interpretation (Schlumberger).

Gamma Ray (GR)

The basic Gamma Ray log (GR) is a measure of the radioactivity of the formations. Radioactivity arises from the decay of three elements present in the rocks, Uranium U, Thorium Th, and Potassium K, which continuously emit gamma rays, in the form of short bursts of high- energy radiation. These gamma rays are capable of penetrating a few inches of rocks. A fraction of these emitted around the borehole, penetrate the drill mud, and can be detected by a suitable gamma ray sensor. The detector gives a direct pulse for each gamma ray detected. The parameter recorded is the number of pulses per unit of time by the detector (Schlumberger, 1974). The units used are GAPI. The GR log is used to:

- 1- detect permeable beds,
- 2- detection and evaluation of radioactive minerals such as potash or uranium ore,
- 3- delineates non-radioactive minerals,
- 4-aid correlation of cased hole,
- 5- perform well-to-well correlation,
- 6- evaluate shale content, Vsh.

Sonic Logs (BHC)

The BoreHole Compensated (BHC) sonic tool or as widely used DT, is used to detect the travel time in the borehole.

Any solid medium will propagate acoustic waves, the aim is to measure the time of propagation of a sound wave through the formation, over a fixed distance. Basically a transmitter and a receiver are placed some distance away on the sonde (Fig.1b). The log readings are scaled not as a velocity but rather as a transit time (DT). The units used are microsecond per foot. The BHC is used to:

- 1- determine the formation porosity,
- 2-perform well-to-well correlation.

Resistivity Logs

In conventional resistivity (Schlumberger, 1985), currents are passed through the formation via certain electrodes, and voltages are measured (in ohms) between certain others. These measured voltages provide the resistivity determinations. So that there will be a current path between electrodes and formations, the sonde must be run in holes containing electrically conductive mud or water.

There are three types of resistivity curves, shallow, medium and deep depending on the spacing between the electrodes (Fig. 1c). The units used are Ohm and resistivity tools are used to :

- 1- determine different formation resistivities e.g R_t , R_{xo} , R_{mf} . ect (Fig. 2),
- 2- perform well-to-well correlation,
- 3- determine hydrocarbon saturation.

Formation Density Log (LDL)

A radioactive source, applied to the hole wall in a shielded skid, emits medium-energy gamma rays into the formation. These gamma rays may be thought of as high velocity particles which collide with the electrons in the formation. At each collision a gamma ray loses some, but not all, of its energy to the electron, and then continues with diminished energy. This type of interaction is known as Compton Scattering. These scattered gamma rays on reaching a detector, at a fixed distance from the source (Fig.1d), are counted, to give an indication of formation density. The number of Compton-scattering collision is related directly to the number of electrons in the formation. Consequently, the response of the density tool is determined essentially by the electron density (number of electrons per cubic centimetre) of the formation. Electron density is related to the true bulk density, ρ_b , in gms/cc, which in turn depends on the density of the rock matrix material, the formation porosity and the density of the fluids in the pores.

The LDL tool is used as porosity-logging tool. Other uses of density measurement include identification of minerals in evaporite deposits, detection of gas, determination of hydrocarbon density, and evaluation of shaly sand and complex lithologies.

Neutron log

Neutrons are electrically neutral particles, each having a mass almost identical to the mass of an hydrogen atom. High energy (fast) neutrons are continuously emitted from a radioactive source which is mounted in the

sonde. These neutrons collide with nuclei of the formation materials in what may be thought as elastic "billiard-ball" type collision. The amount of energy lost per collision depends on the relative mass of the nucleus which the neutron collides.

The greatest energy loss occur when the neutron strikes a nucleus of particularly equal mass,-i.e, a hydrogen nucleus. Collision with heavy nuclei do not slow the neutron down very much. Thus, the slowing-down of neutrons depends largely on the amount of hydrogen in the formation. When the hydrogen concentration of the material surrounding the neutron source is large, most of the neutrons are slowed down and captured within a short distance of the source. On the the hand, if the Hydrogen concentration is small, the neutrons travel farther from the source before being captured. Accordingly, the counting rate at the detector increases for decreased hydrogen concentration and vice versa (Schlumberger, 1974).

Table 1 summarizes some of the well-logging tools, their measurements, and their uses in well-to-well correlation.

Curve designation	Measured parameter	Factors affecting measurement	Correlating feature
Spontaneous Potential (SP)	DC potential generated at boundary of porous and non porous beds	(1) shale content (2) porosity and/or permeability (3) connate water and mud filtrate salinities (4) fluid saturation (5) hole size (6) formation density	curve shape SP values
Gamma Ray (GR)	naturally occurring gamma rays (normally from radioactive salts concentrated in shales)	(1) hole size (2) operating techniques (statistics) (3) shale content (4) mineralogy	curve shape GR values
Apparent Resistivity	electrical conductivity of rocks	(1) mud salinity (2) borehole size (3) fluid saturation and salinity (4) porosity	curve shape measured value (water saturation)
Acoustic (DT)	formation travel time	(1) lithology (2) porosity (3) fluid saturation (4) mud condition (5) washouts	curved shape measured value (porosity)
Density	electron density (related to bulk density in most common rocks)	as above for acoustic log	as above for acoustic log
Neutron	hydrogen index	(1) hole size (2) mud weight and salinity (3) fluid saturation (4) porosity (5) lithology	curve shape

Table 1 Well-logging tools, their measured parameters, and factors affecting the measurements.

APPENDIX C

PROGRAM PCAXCOR

C
C PROGRAM TO PERFORM PRINCIPAL COMPONENT ANALYSIS
C OF WELL-LOG DATA, BOUNDARY IDENTIFICATION AND WELL
C TOWELL CORRELATION.
C
C 1-MULTV7 SUBROUTINE: TO CALCULATE THE CORRELATION
C , MATRIX OR THE VARIANCE-COVARIANCE MATRIX OF THE
C ORIGINAL MATRIX. THIS SUBROU-
C TINE USES "Y" THE ORIGINAL MATRIX AND OUTPUTS THE
C VAR-COV. OR THE
C CORR. MATRIX "C" WHICH HAS AS MANY ROWS AS
C COLUMNS.

C
C 2- TRED2 SUBROUTINE:
C TRED2 CALCULATES THE TRIDIAGONAL MATRIX OF THE
C SQUARE MATRIX.
C FROM TRED2 COME TWO VECTORS D AND E WHICH ARE THE
C DIA. AND OFF-DIA. ELEMENTS OF "A" . D,E ARE MADE
C ARGUMENTS OF TQLI.
C
C 3- TQLI SUBROUTINE:
C D AND E FROM TRED2 ARE REPLACED BY THE EIGENVALUES
C AND THE CORRESPONDING EIGENVECTORS RESPECTIVELY.
C
C 4- EIGENVALUE SUBROUTINE:
C TO CALCULATE THE PERCENTAGE OF EACH EIGENVALUES
C CONTRIBUTE TO THE TOTAL PERCENTAGE. IT RETURNS ONE-
C DIM. VECTOR "PERC" WHICH CONTAIN THE PERCENTAGES.

C
 C 5- ORGANIZE SUBROUTINE:
 C THIS IS TO ORGANIZE THE FINAL OUTPUT AND WRITE IT TO A
 C IT LISTS THE EIGENVECTORS & EIGENVALUES AND THE
 C PERCENTAGE OF EACH
 C EIGENVALUE TO THE TOTAL VARIANCE."A" TWO-
 C DIMENSIONAL ARRAY
 C CONTAINS THE EIGENVECTORS & "D" IS ONE-DIMENSIONAL
 C VECTOR
 C CONTAINS THE CORRESPONDING EIGENVALUES AND "PERC"
 C IS ONE-
 C DIMENSIONAL VECTOR CONTAINS
 C THE PERCENTAGE OF EACH EIGENVALUE.
 C
 C 6- SMOOTH SUBROUTINE:
 C THIS IS TO FILTER THE DATA USING A MOVING AVERAGE
 C FILTER.
 C
 C BOTH TRED2 & TQLI ARE OBTAINED FROM NUMERICAL
 C RECIPES BY :
 C WILLIAM H. PRESS AND BRIAN P. FLANNERY.
 C SOME MODIFICATIONS ON THESE SUBROUTINES HAVE BEEN
 C MADE.
 C
 C THE MAIN PROGRAM COPIES MATRIX "C" INTO MATRIX "A".
 C "A" IS MADE AS
 C AN ARGUMENT TO "TRED2" AND "TQLI" SUBROUTINES.
 C
 C LIST OF THE INPUTS
 C
 C Y TWO-DIMENSIONAL ARRAY INPUT INTO MULTV7
 C SUBROUTINE.
 C AND CONTAINS THE ORIGINAL MATRIX.
 C C TWO-DIMENSIONAL ARRAY INPUT INTO TRED2 & TQLI
 C SUBROUTINES.
 C IT CONTAINS THE VAR-COV. MATRIX OR CORRELATION
 C MATRIX.....

C A TWO-DIMENSIONAL ARRAY INPUT INTO TRED2
 C
 C ORGANIZE SUBROUTINE...
 C THIS MATRIX IS COPIED FROM "C".ON INPUT IT CONTAINS
 C THE VAR-COV. MATRIX.
 C D ONE-DIMENSIONAL VECTORS INPUT INTO TQLI &
 C EIGENVALUE AND ORGANIZE SUBROUTINES.
 C ON INPUT IT IS DIAGONAL ELEMENTS OF THE TRIDIAG
 C MATRIX.
 C E ONE-DIMENSIONAL VECTOR INPUT INTO TQLI
 C SUBROUTINE.
 C IS SUB-DIAGONAL ELEMENTS OF THE TRIDIAG. MATRIX.
 C P..... ONE-DIMENSIONAL VECTOR INPUT INTO SMOOTH
 C SUBROUTINE.
 C IT CONTAINS THE EIGENVECTORS TO BE SMOOTHED.
 C
 C LIST OF OUTPUTS
 C
 C COR TWO-DIMENSIONAL ARRAY OUTPUTS BY MULTV7
 C SUBROUTINE.
 C IT CONTAINS THE CORRELATION OR THE
 C VAR.COV.MATRIX.....
 C DONE-DIMENSIONAL VECTOR OUTPUTS BY TRED2
 C ORGANIZE
 C SUBROUTINE.....
 C ON OUTPUT IT RETURNS THE EIGENVALUES AND
 C EIGENVALUE PERCENTAGES.....
 C E ONE-DIMENSIONAL VECTOR OUTPUTS BY TRED2
 C SUBROUTINE.
 C TRI TWO-DIMENSIONAL ARRAY OUTPUTS BY TQLI
 C SUBROUTINE.
 C IT IS THE TRI-DIAGONAL MATRIX.
 C F TWO-DIMENSIONAL ARRAY OUTPUTS BY TQLI
 C SUBROUTINE.
 C IT CONTAINS THE PRINCIPAL COMPONENT SCORES
 C A TWO-DIMENSIONAL ARRAY OUTPUTS BY TQLI

```

C   ORGANIZE
C   SUBROUTINE.ON OUTPUT IT CONTAINS THE EIGENVECTORS. .
C   PERC ..... ONE-DIMENSIONAL VECTOR OUTPUTS BY ORGANIZE
C   EIGENVALUE.
C   ..... IT CONTAINS THE PERCENTAGES OF EACH EIGENVALUE.
C   SMOOTHD .. TWO-DIMENSIONAL ARRAY OUTPUTS BY
C   SMOOTH
C   SUBROUTINES. IT CONTAINS THE SMOOTHED PRINCIPAL
C   COMPONENT SCORES.....
C   BOUNDARY...ONE-VECTOR OUTPUTS BY BOUNDARY
C   SUBROUTINE.
C   .....IT CONTAINS THE BOUNDARY OF DIFFERENT
C   FORMATION.
C   .....
C
C
C   PARAMETER(NMAX=8000,NCOL=10,TINY=1.0E-4,NI=4,N=7)
C   PARAMETER(INROW=8000)
C   DIMENSION A(NMAX,NCOL),D(NCOL),E(NCOL)
C   DIMENSION TRI(NMAX,NCOL),Y(INROW,NCOL)
C
C   -----
C   DIMENSION STANDZ(INROW,NCOL)
C   CHARACTER*10 STANDZFILE,ANSYN
C
C   -----
C
C   DIMENSION PERC(NCOL)
C
C   -----
C
C   DIMENSION Y(NMAX,NCOL),C(NMAX,NCOL)
C   CHARACTER*55 PRINC,FILE2,FINALOUT
C   CHARACTER*35 MULFILEIN,MULFILOUT,FIRSTSC
C   CHARACTER*10 CHAR(10),YN,YN2
C
C   BOUNDLYIN FILE CONTAINS THE FIRST PRINCIPAL
C   COMPONENT COPIED. FROM"SMOOTHD".
C   SMOOTHD IS THE FILE CONTAINING THE SMOOTHED DATA

```

```

C   OUTPUTS OF "SMOOTH".
C
REAL SMOOTHD(8000),P(8000,10),PP(8000)
C
DIMENSION DSQUARE(8000)
C
0010 WRITE(6,1000)
1000 FORMAT(/)
    WRITE(6,1001)'INPUT NAME OF THE FILE CONTAINING THE
* ORIGINAL VARIABLES'
1001 FORMAT(/,A)
    READ(5,1002)MULFILEIN
1002 FORMAT(A10)
    WRITE(6,1003)'INPUT THE NO. OF COLUMNS OF THE MATRIX'
1003 FORMAT(/,A)
    READ(5,1004)NCL
1004 FORMAT(I12)
    WRITE(6,1005)'INPUT THE OUTPUT FILE CONTAINS THE VAR-
* COV. OR CORRELATION MATRIX'
1005 FORMAT(/,A)
    READ(5,1006)MULFILOUT
1006 FORMAT(A10)
C
    WRITE(6,1007)'ENTER THE FILE THAT WILL CONTAIN THE
* PRINC. COMP. SCORES'
1007 FORMAT(/,A)
    READ(5,1008)PRINC
1008 FORMAT(A10)
    WRITE(6,1009)'ENTER THE FINAL OUTPUT FILE CONTAINING
* THE FINAL OUTPUT'
1009 FORMAT(/,A)
    READ(5,1010)FINALOUT
1010 FORMAT(A10)
C
    WRITE(6,1011)'INPUT THE FILE THAT WILL CONTAIN THE
* PERCENTAGE'
1011 FORMAT(/,A)

```

```

      READ(5,1012)FILE2
1012  FORMAT(A10)
C
      WRITE(6,1126)'ENTER THE FILE WILL CONTAIN STANDARDISE
*   DATA'
1126  FORMAT(/,A)
      READ(5,1127)STANDZFILE
1127  FORMAT(10A)
C
      WRITE(6,1016)
1016  FORMAT(/,A)
      WRITE(6,1017)'ENTER THE FILE CONTAIN THE PRINCIPAL
*   COMPONENT SCORES'
1017  FORMAT(/,A)
      READ(5,1018)FIRSTSC
1018  FORMAT(A10)
      WRITE(6,1020)'ENTER THE OUTPUT FILE CONTAINS THE
*   BOUNDARIES'
1020  FORMAT(/,A)
      READ(5,1021)BOUNDFILEOUT
1021  FORMAT(A10)
C
      WRITE(6,1019)'ENTER THE BOUNDARY WINDOW'
1019  FORMAT(A)
      READ(5,*)IBWIN
C
C FIRST OPEN THE FILE OF THE ORIGINAL MATRIX THAT HAS BEEN
C USED BY SUBROUTINE "MULTV7.F".NAME THIS MATRIX AS "Y"
C AND USE IT TO CALCULATE THE PRINCIPAL COMPONENT SCORES
C OF THE ORIGINAL MATRIX BY MULTIPLYING. THIS MATRIX BY
C THE CORRESPONDING EIGENVECTORS.
C
C MULFILEIN: A FILE CONTAINS THE ORIGINAL MATRIX (ROW
C DATA)
C MULFILEOUT " " " " VARIANCE-COVARIANCE
C MATRIX(MULTV7).
C FILE2 : " " " " EIGENVALUES PERCENTAGE(EIGENVALUE

```

```

C OUTPUT)
C PRINC: " " " " PRINCIPAL COMPONENT SCORES (PCAXCOR
C OUTPUT)
C FINALOUT: " " " " FINAL OUTPUT FILE CONTAINS FINAL
C RESULT.
C SMOOTHOUT " " " " SMOOTHED PRINCIPAL COMPONENT
C SCORES.
C BOUBDFILEOUT " " " " BOUNDARY OF DIFFERENT
C FORMATIONS.
C -----
  OPEN(4,FILE=MULFILEIN,STATUS="UNKNOWN")
  OPEN(2,FILE=MULFILEOUT,STATUS="UNKNOWN")
  OPEN(10,FILE=PRINC,STATUS="UNKNOWN")
  OPEN(12,FILE=FILE2,STATUS="UNKNOWN")
  OPEN(13,FILE=FINALOUT,STATUS="UNKNOWN")
  OPEN(17,FILE=BOUNDFILEOUT,STATUS="UNKNOWN")
  OPEN(18,FILE=FIRSTSC,STATUS="UNKNOWN")
  OPEN(20,FILE=STANDZFILE,STATUS="UNKNOWN")
C -----
C
C Y IS THE ORIGINAL MATRIX CONTAINS THE ORIGINAL
C VARIABLES. C IS THE MATRIX CONTAINS THE VARIANCE-
C COVARIANCE OR CORRELATION MATRIX.
C
C NOW READ THE ORIGINAL MATRIX Y
C
      II=0
      DO 8 KI=1,INROW
      READ(4,*,END=555)(Y(KI,JI),JI=1,NCL)
      II=II+1
8      CONTINUE
555    CONTINUE
      NROW=II
C
C IF READING MORE THAN ONCE FROM THE SAME FILE,
C REWIND THE FILE

```

```

C   AND CLOSE IT TO BE OPENED AGAIN.
      REWIND(4)
      CLOSE(4)

C -----
C WRITE THE VARIANCE-COVARIANCE OR THE CORRELATION
C   MATRIX TO A FILE
C
C READ THE NAME OF THE VARIABLE FROM THE SCREEN.
C CHAR IS THE NAME OF THE CHARACTERS
C THIS SECTION IS FOR THE USER TO CHOOSE WHETHER TO
C STANDARDISE THE ROW DATA OR NOT AND WHICH, THE VAR-
C COV.
C MATRIX OR CORRELATION MATRIX IS TO BE USED.
C
2111 WRITE(6,1025)'DO YOU WANT TO STANDARDIZE DATA...
      * [Y/N]'
      READ(5,*)YN
      IF(YN.EQ."Y".OR.YN.EQ."y")THEN
C
      CALL STANDARIZE(Y,NROW,NCL,STANDZ)
      DO 1111 J=1,NROW
      WRITE(20,'(7F11.3)')(STANDZ(J,I),I=1,NCL)
1111 CONTINUE
C
      WRITE(6,1025)'STANDARDISE DATA AND CORRELATION
      * MATRIX USED '
C
C
      WRITE(2,*)'THE CORRELATION MATRIX IS :'
      WRITE(13,*)'THE CORRELATION MATRIX IS :'
C
      CALL MULTV7(STANDZ,C,NROW,NCL)
C
      WRITE(6,1515)'STARTING CALCULATE CORRELATION MATRIX'
1515 FORMAT(/,A)
      DO 3333 J=1,NCL
      WRITE(2,'(1X,7F10.3)')(C(I,J),I=1,NCL)

```

```

WRITE(13,'(6X,7F10.3)')(C(I,J),I=1,NCL)
WRITE(6,1025)' ENTER NAME OF CHARACTER =====> 'J
READ(5,FMT=666) CHAR(J)
1025  FORMAT(/ A,I2)
C
      KB=J
3333  CONTINUE
C
C COPY MATRIX "C" INTO "A"
C "C" IS THE MATRIX CONTAINS THE CORRELATION MATRIX
C
      DO 12 J=1,NCL
      DO 11 I=1,NCL
      A(J,I)=C(J,I)
11    CONTINUE
12    CONTINUE
C
      ELSE
5555  WRITE(6,5151)'NON-STANDARISED DATA AND VARIANCE-
      * COVARIANCE MATRIX '
5151  FORMAT(/,A)
C
      WRITE(2,*)' THE VARIANCE-COVARIANCE MATRIX IS :'
      WRITE(13,*)' THE VARIANCE-COVARIANCE MATRIX IS :'
      WRITE(6,5252)'STARTING CALCULATE VARIANCE-
      * COVARIANCE'
5252  FORMAT(/,A)
C
      CALL MULTV7(Y,C,NROW,NCL)
C
      DO 6666 I=1,NCL
      WRITE(2,'(1X,7F10.3)')(C(I,J),J=1,NCL)
      WRITE(13,'(6X,7F10.3)')(C(I,J),J=1,NCL)
      WRITE(6,1027)' ENTER NAME OF CHARACTER =====> 'I
      READ(5,FMT=666) CHAR(I)
1027  FORMAT(/ A,I2)
C

```

```

      KB=I
C
6666  CONTINUE
C
C COPY MATRIX "C" INTO "A"
C "C" IS THE MATRIX CONTAINS THE VAR-COV. MATRIX
C
      DO 24 J=1,NCL
      DO 23 I=1,NCL
      A(J,I)=C(J,I)
23    CONTINUE
24    CONTINUE
C    COPY MATRIX "Y" INTO "STANDZ" FOR LATER CALCULATION
C    OF THE SCORES.
      DO 2020 J=1,NROW
      DO 2021 I=1,NCL
      STANDZ(J,I)=Y(J,I)
2021  CONTINUE
2020  CONTINUE
      ENDIF
2211  FORMAT(A12)
2221  FORMAT(I4)
666   FORMAT(5A)
1029  FORMAT(A)
C
      CALL TRED2(A,NMAX,NCL,D,E)
C
C TEST FORTHE EIGENVALUE.....
C
      WRITE(2,(/1x,A))'DIAGONAL ELEMENTS'
      WRITE(2,(1x,7f11.3))'(D(I),I=1,ncl)
      WRITE(2,(/1x,A))'OFF-DIAGONAL ELEMENTS'
      WRITE(2,(1x,7f11.3))'(E(I),I=1,ncl)
C    CHECK TRANSFORMATION MATRIX .
      DO 6 J=1,NCL
      DO 5 K=1,NCL
      TRI(J,K)=0.0

```

```

DO 4 L=1,NCL
DO 3 M=1,NCL
TRI(J,K)=TRI(J,K)+A(L,J)*C(L,M)*A(M,K)
3 CONTINUE
4 CONTINUE
5 CONTINUE
6 CONTINUE
C HOW DOES IT LOOK ...
WRITE(2,'(/1x,A)')'TRI-DIAGONAL MATRIX LOOKS'
DO 7 I=1,NCL
WRITE(2,'(1X,7F11.3)')(TRI(I,J),J=1,NCL)
7 CONTINUE
CALL TQLI(D,E,NMAX,NCL,A,N)
C WRITE THE EIGENVALUES FOR THE REAL SYMMETRIC
C , MATRIX.
C
C MATRIX "STANDZ" BELOW IS THE STANDARDIZED ROW DATA
C WHEN THE ORIGINAL DATA WAS STANDARDIZED AND ALSO
C "STANDZ" IS THE ORIGINAL ROW DATA WHEN
C STANDARDIZATION HAS NOT TAKE PLACE.(i.e "STANDZ"HAS
C BEEN COPIED FROM ORIGINAL DATA MATRIX "Y").
C
WRITE(2,'(/1x,A)')'EIGENVALUES FOR REAL SYMMETRIC
* MATRIX'
DO 16 I=1,NCL
DO 14 J=1,NROW
P(J,I)=0.0
DO 13 K=1,NCL
P(J,I) = P(J,I)+STANDZ(J,K)*A(K,I)
13 CONTINUE
14 CONTINUE
C
WRITE(2,'(/1X,A,I3,A,F12.3)')'EIGENVALUE',I,'=',D(I)
WRITE(2,'(/1X,A)')'EIGENVECTORS'
C
C WRITE THE EIGENVALUES TO A FILE
C

```

C TEST FOR THE EIGENVALUE.....

C

DO 15 J=1,NCL

IF(ABS(A(J,I)).LT.TINY)THEN

WRITE(2,'(15X,F12.6,A12)')A(J,I),'DIV. BY ZERO'

c WRITE(2,'(1X,2F12.6,A12)')A(J,I)

ELSE

c WRITE(2,'(1X,2F12.6,E14.6)')A(J,I),F(J,K),F(J,K)/A(J,I)

WRITE(2,'(15X,F12.6,)')A(J,I)

ENDIF

15 CONTINUE

16 CONTINUE

C

C WRITE ALL THE PRINCIPAL COMPONENT SCORES TO A FILE AND

C CAST THEM IN A MATRIX(THE LAST COLUMN IS THE FIRST

C PRIC. COMP.)

C

DO 2010 J=1,NROW

WRITE(10,'(7F10.4)')(P(J,I),I=1,NCL)

2010 CONTINUE

C

C NOW WRITE EIGENVALUES PERCENTAGE OF THE TOTAL

C VARIANCE TO A FILE

C

CALL EIGENVALUE(D,PERC,NCL)

C

WRITE(13,'(/1X,A,3X,10(I2,6X),//)')'VARIABLE',(I,I=1,NCL)

C

C CHAR: IS THE NAME OF THE VARIABLES TO BE READ FROM THE

C SCREEN.

DO 32 I=1,NCL

WRITE(12,'(1X,F6.2)')PERC(I)

WRITE(13,'(3X,A3,1X,7F10.5)')CHAR(I),(A(I,J),J=NCL,1,-1)

32 CONTINUE

CALL ORGANIZE(D,PERC,NCL)

C

CLOSE(10)

```

      OPEN(10, FILE=PRINC)
C
C CALL SUBROUTINE "SMOOTH" TO SMOOTH THE PRINCIPAL
C COMPONENT SCORES.
C
C KB REFERENT TO WHICH COMPONENT TO BE USED(IN THIS
C CASE THE LAST IS USED).
C
C ASK FOR WHICH COMPONENT TO BE USED :
C
      PRINT 2001
2001  FORMAT(//, '*** WHICH COMPONENTS TO BE USED 1,2 OR
      *   LAST COMP[3].***', //)
      READ(5,*)COMP
C
C TO USE THE FIRST COMPONENT
C
      IF(COMP.EQ.1)THEN
      DO 2110 I=1,NROW
      READ(10,*)(P(I,J),J=1,NCL)
      PP(I)=P(I,KB)
      WRITE(18,*)PP(I)
2110  CONTINUE
      ENDIF
C
C TO USE THE SECOND COMPONENT
C
      IF(COMP.EQ.2)THEN
      DO 2112 I=1,NROW
      READ(10,*)(P(I,J),J=1,NCL)
      PP(I)=P(I,KB-1)
      WRITE(18,*)PP(I)
2112  CONTINUE
C
      ENDIF
C CHECK THE LAST COMPONENT(THE SMALLEST)
      IF(COMP.EQ.3)THEN

```

```

DO 1112 I=1,NROW
  READ(10,*)(P(I,J),J=1,NCL)
  PP(I)=P(I,KB-KB+1)
  WRITE(18,*)PP(I)
1112 CONTINUE
      ENDIF
C
C THIS SECTION IS TO DECIDE TO USE A FILTER TO SMOOTH THE
C ONE OF THE PRINCIPAL COMPONENT OR TO USE IT DIRECTLY
C WITHOUT SMOOTHING.
C
1114 PRINT 1113
1113 FORMAT(//,'***** DO YOU WANT TO FILTER THE DATA
*   *****, [Y/N]',//)
      READ(5,*)YN2
C
C THIS SECTION PROMPTS FOR THE DATA FOR "SMOOTH"
C SUBROUTINE.
C
      IF(YN2.EQ."Y".OR.YN2.EQ."y")THEN
C
      WRITE(6,1013)'USING SUBROUTINE TO FILTER THE OUTPUT
* DATA ===>'
1013 FORMAT(/,A)
      WRITE(6,1014)'ENTER OUTPUT FILE FOR THE SMOOTHED
* DATA.'
1014 FORMAT(/,A)
      READ(5,1015)SMOOTHOUT
1015 FORMAT(A10)
      WRITE(6,1121)' ENTER THE LENGTH OF THE WINDOW OF THE
* FILTER.'
1121 FORMAT(/,A)
      READ(5,*)LEN
C
      OPEN(15,FILE=SMOOTHOUT,STATUS="UNKNOWN")
C
      CALL SMOOTH(PP,SMOOTHOUT,NROW,LEN)

```

```

C
  DO 220 M=1,NROW
  WRITE(15,'(F12.6)')SMOOTHD(M)
220  CONTINUE
C
C THIS "ENDIF" IS RELATED TO YN2 TO FILTER OR NOT THE DATA.
C
  CLOSE(15)
  OPEN(15,FILE=SMOOTHOUT)
  DO 2122 I=1,NROW
  READ(15,*)SMOOTHD(I)
2122 CONTINUE
C
C IF YOU DO NOT WANT TO FILTER THE DATA THEN:
C THE VECTOR "SMOOTHD" IS THE NON-FILTERED PRINCIPAL
C COMP. SCORES OF THE CHOSEN COMPONENT.
C
  ELSEIF(YN2.EQ."N".OR.YN2.EQ."n")THEN
  PRINT 5222
5222  FORMAT(//,'NON-FILTERED PRINCIPAL COMPONENT WILL BE
*  USED',//)
C
C COPY VECTOR "PP" INTO "SMOOTHD".
C
  REWIND(UNIT=18)
C
  DO 2223 I=1,NROW
  SMOOTHD(I)=PP(I)
2223 CONTINUE
  ENDIF
C CALL THE BOUNDARY SUBROUTINE TO IDENTIFY THE
C BOUNDARIES.
C
  CALL BOUNDARY(SMOOTHD,IBWIN,NROW,DSQUARE)
C
  DO 1045 I=1,NROW
  WRITE(17,*)DSQUARE(I)

```

1045 CONTINUE

C

C IF TO BE CONTINUED TO WORK ON OTHER SET OF DATA

C

WRITE(6,*)' DO YOU WANT TO CONTINUE [Y/N]'

READ*,ANSYN

IF(ANSYN.EQ.'Y'.OR.ANSYN.EQ.'y')GO TO 0010

IF(ANSYN.NE.'Y'.OR.ANSYN.NE.'y')THEN

CONTINUE

ENDIF

C

C

C CALL THE CORRELATION SUBROUTINE TO CORRELATE BETWEEN

C THE TWO LOGS.

C

CALL XCOR(NROW)

C

STOP

END

C

C

C

C

C

C

C

C

C

C

C THIS SUBROUTINE IS TO STANDARDIZE THE DATA(i.e THE DATA

C WILL HAVE A MEAN OF ZERO [0] AND VARIANCE OF ONE [1]

C USING THE EQUATION:

C

C

C

$$X_{IJ}=(X_{IJ}-\text{MEAN}(J))/\text{VAR}(J)$$

C
 C STANDARDIZE SUBROUTINE CALCULATES THE STANDARDIZED
 C DATA FROM THE ORIGINAL VARIABLES(MATRIX "Y") AND
 C RETURN THE OUTPUT DATA(STANDARZIED) IN MATRIX
 C "STANDZ" WHICH WILL BE FED TO THE MAIN PROGRAM TO
 C CALCULATE THE CORRELATION MATRIX, EIGENVALUES,
 C EIGENVECTORS AND THE PRINCIPAL COMPONENT SCORES. NOTE
 C THAT THESE SCORES ARE CALCULATED FROM THE
 C STANDARDIZED DATA RATHER THAN THE ROW DATA.

C
 SUBROUTINE STANDARIZE(Y,NROW,NCL,STANDZ)
 PARAMETER(ICL=10,MAXN=8000)
 C Y=THE ORIGINAL DATA MATRIX WHICH CONTAINS THE
 C ORIGINAL WELL-LOG
 C VARIABLES OF DIMENSION NROW BY NCL.
 C STANDZ=STANDARDIZED DATA MATRIX OF DIMENSION
 C NROW BY NCL.
 C MAXN=DIMENSION PARAMATER FOR THE NUMBER OF ROWS
 C OF THE MATRIX
 C ICL=DIMENSION VECTOR SET BIGGER THAN THE EXPECTED
 C NUMBER OF
 C COLUMNS. BOTH MAXN AND ICL CAN BE ADJUSTED TO ANY
 C NUMBER.
 DIMENSIONSTANDZ(MAXN,ICL),DMEAN(MAXN),VAR(MAXN),
 * STD(MAXN)
 DIMENSION OS(MAXN,ICL),SUM(ICL),Y(MAXN,ICL)

C
 C Y..... IS THE ROW DATA MATRIX.
 C STANDZ..... IS THE STANDARDIZED DATA MATRIX.

C
 DO 25 M=1,NCL
 SUM(M)=0.0
 DO 15 J=1,NROW
 SUM(M)=SUM(M)+Y(J,M)
 15 CONTINUE
 DMEAN(M)=SUM(M)/FLOAT(NROW)
 25 CONTINUE

```

DO 199 M=1,NCL
SUMM=0.0
DO 99 I=1,NROW
OS(I,M)=Y(I,M)-DMEAN(M)
SOS=OS(I,M)**2
SUMM=SUMM+SOS
99 CONTINUE
VAR(M)=SUMM/FLOAT(NROW-1.0)
STD(M)=SQRT(VAR(M))
199 CONTINUE
299 FORMAT(F9.5)
C
DO 18 M=1,NCL
DO 19 I=1,NROW
STANDZ(I,M)=OS(I,M)/STD(M)
19 CONTINUE
18 CONTINUE
C
RETURN
END
C
C SUBROUTINE TRED2
C
C SUBROUTINE TRED2 TO CALCULATE THE TRI-DIAGONAL MATRIX
C OF VAR.-COV. MATRIX OR THE CORRELATION MATRIX.
C
SUBROUTINE TRED2(A,NMAX,NCL,D,E)
C
C A=SQUARE MATRIX OF THE VARIANCE-COVERAINCE OR
C CORRELATION MATRIX.
C NMAX=PARAMETER TO SET THE DIMENSION OF THE NUMBER
C OF ROWS OF DATA.
C NCL=THE NUMBER OF COLOUMNS OF THE DATA MATRIX.
PARAMETER(NCOL=10)
DIMENSION A(NMAX,NCOL),D(NCOL),E(NCOL)
N=NCL
C NCOL= A DIMENSIONAL VECTOR SET BIGGER THAN THE

```

```

C   EXPECTED NUMBER
C   OF COLUMNS. NCL IS THE NUMBER OF COLUMNS OF THE
C   MATRIX.
      IF (N .GT. 1) THEN
        DO 18 I=N,2,-1
          L=I-1
          H=0.0
          SCALE=0.0
              IF(L .GT. 1) THEN
                DO 11 K=1,L
                  SCALE=SCALE+ABS(A(I,K))
11      CONTINUE
              IF(SCALE .EQ. 0.0) THEN
                E(I)=A(I,L)
                    ELSE
                DO 12 K=1,L
                  A(I,K)=A(I,K)/SCALE
                H=H+A(I,K)**2
12      CONTINUE
                F=A(I,L)
                G=+SIGN(SQRT(H),F)
                E(I)=SCALE*G
                H=H-F*G
                A(I,L)=F-G
                F=0.0
                DO 15 J=1,L
C       OMIT THE FOLLOWING IF FINDING ONLY EIGENVALUES
                A(J,I)=A(I,J)/H
                G=0.0
                DO 13 K=1,J
                  G=G+A(J,K)*A(I,K)
13      CONTINUE
                IF(L.GT.J) THEN
                  DO 14 K=J+1,L
                    G=G+A(K,J)*A(I,K)
14      CONTINUE
                ENDIF

```

```

E(J)=G/H
F=F+E(J)*A(I,J)
15 CONTINUE
HH=F/(H+H)
      DO 17 J=1,L
        F=A(I,J)
        G=E(J)-HH*F
        E(J)=G
          DO 16 K=1,J
            A(J,K)=A(J,K)-F*E(K)-G*A(I,K)
16 CONTINUE
17 CONTINUE
        ENDIF
      ELSE
        E(I)=A(I,L)
      ENDIF
      D(I)=H
18 CONTINUE
      ENDIF
C   OMIT THE FOLLOWING IF FINDING ONLY EIGENVALUES.
C
D(1)=0.0
C   BEGIN ACCUMULATION OF TRANSFORMATION MATRIX.
      E(1)=0.0
      DO 23 I=1,N
C   DELETE LINE FROM HERE
      L=I-1
      IF (D(I) .NE. 0.0) THEN
        DO 21 J=1,L
          G=0.0
          DO 19 K=1,L
            G=G+A(I,K)*A(K,J)
19 CONTINUE
          DO 20 K=1,L
            A(K,J)=A(K,J)-G*A(K,I)
20 CONTINUE
21 CONTINUE

```

```

                ENDIF
C TO HERE WHEN FINDING ONLY EIGENVALUES.....
C
      D(I)=A(I,I)
C ALSO DELETE LINES FROM HERE.
C
C RESET ROW AND COLUMNS OF 'A' TO IDENTIFY FOR NEXT
C ITERATION.
      A(I,I)=1.0
      IF(L .GE. 1)THEN
        DO 22 J=1,L
          A(I,J)=0.0
          A(J,I)=0.0
22      CONTINUE
      ENDIF
C
C TO HERE IF FINDING ONLY EIGENVALUES.
C
23      CONTINUE
      RETURN
      END
C
C
C          SUBROUTINE TQLI
C
C SUBROUTINE TQLI TO CALCULATE THE EIGENVALUE AND THE
C EIGENVECTORS.
C
C THIS SUBROUTINE IS TO CALCULATE EIGENVALUES AND THE
C EIGENVECTORS "D" AND "A" RESPECTIVELY FROM THE TRI-DIAG.
C MATRIX "TRI".
C
      SUBROUTINE TQLI(D,E,NMAX,NCL,A,N)
C
C
C      N IS THE DIAGONAL ELEMENTS OF C.
C

```

```

DIMENSION D(N),E(N),A(NMAX,N)
IF(N.GT.1)THEN
DO 11 I=2,N
E(I-1)=E(I)
11 CONTINUE
ENDIF
E(N)=0.0
DO 15 L=1,N
ITER=0.0
1 DO 12 M= L,N-1
C
C LOOK FOR A SINGLE SMALL SUBDIAGONAL ELEMENTS TO SPLIT
C THE MATRIX.
DD=ABS(D(M))+ABS(D(M+1))
IF(ABS(E(M))+DD.EQ.DD)GOTO 2
12 CONTINUE
M=N
2 IF(M.NE.L)THEN
IF(ITER.EQ.70)PAUSE' TOO MANY ITERATION'
ITER=ITER+1
G=(D(L+1)-D(L))/(2.0*E(L))
R=SQRT(G**2+1.0)
G=D(M)+D(L)+E(L)/(G+SIGN(R,G))
C THIS WAS dm-ks
S=1.0
C=1.0
P=0.0
DO 14 I=M-1,L,-1
F=S*E(I)
B=C*E(I)
IF(ABS(F).GE.ABS(G))THEN
C=G/F
R=SQRT(C**2+1.0)
E(I+1)=F*R
S=1.0/R
C=C*S
ELSE

```

```

S=F/G
R=SQRT(S**2+1.0)
E(I+1)=G*R
C=1.0/R
S=S*C
ENDIF
G=D(I+1)-P
R=(D(I)-G)*S+2.0*C*B
P=S*R
D(I+1)=G+P
G=C*R-B

```

C

C OMIT LINES FROM HERE

C

```

DO 13 K=1,N
F=A(K,I+1)
A(K,I+1)=S*A(K,I)+C*F
A(K,I)=C*A(K,I)-S*F

```

13 CONTINUE

C

C TO HERE IN FINDING ONLY EIGENVALUES

C

14 CONTINUE

D(L)=D(L)-P

E(L)=G

E(M)=0.0

GOTO 1

ENDIF

15 CONTINUE

RETURN

END

C

C

SUBROUTINE EIGENVALUE

C

C TO CALCULATE THE PERCENTAGE OF THE EIGENVALUE.

C

SUBROUTINE EIGENVALUE(D,PERC,NCL)

```

C
C   D=EIGENVALUE.
C   PERC=PERCENTAGE OF EACH EIGENVALUE.
C   NCL=NUMBER OF COLOUMNS OF THE INPUT MATRIX.
C
C   PARAMETER(NCOL=10)
C   DIMENSION D(NCOL),SUM(NCOL),PERC(NCOL)
C   NCOL=A DIMENSIONA VECTOR SET BIGGER THAN THE
C   EXPECTED NUMBER OF
C   COLUMNS OF THE MATRIX. THIS CAN BE ADJUSTED TO ANY
C   NUMBER.
C   DO 20 I=1,NCL
C   SUM(I)=0.
C   DO 10 J=1,NCL
C   SUM(I)=SUM(I)+D(J)
10  CONTINUE
20  CONTINUE
C   DO 30 J=1,NCL
C   PERC(J)=(D(J)/SUM(J))*100
30  CONTINUE
C   RETURN
C   END

```

```

C
C
C           SUBROUTINE ORGANIZE
C
C TO ORGANIZE THE FINAL OUTPUT OF THE PRINCIPAL
C COMPONENT SECTION.

```

```

C
C
C   SUBROUTINE ORGANIZE(D,PERC,NCL)
C   DIMENSION D(NCOL),PERC(NCOL)
C
C   D=EIGENVALUE.
C   PERC=PERCENTAGE OF EACH EIGENVALUE.
C   NCL=NUMBER OF COLOUMNS OF THE INPUT MATRIX.
C   NCOL=A DIMENSIONAL VECTOR BIGGER THAN THE EXPECTED

```

C NUMBER OF COLUMNS OF THE MATRIX. THIS IS ADJUSTABLE
 C VECTOR.
 C PERC=OUTPUT FILE WILL CONTAIN THE EIGENVAULES AND
 C THE PERCENTAGE
 C EACH EIGENVALUE TO THE TOTAL VARIANCE.

WRITE(13,'(/1X,A)')'EIGENVALUES'

WRITE(13,'(/,7(F10.3,1X))')(D(I),I=NCL,1,-1)

WRITE(13,'(/1X,A)')'PERCENTAGE OF TOTAL VARIANCE

* CONTRIBUTED BY EACH EIGENVALUE'

WRITE(13,'(/,7(F10.3,1X))')(PERC(I),I=NCL,1,-1)

RETURN

END

C

C

C . MULTV7.F SUBROUTINE

C

C CALCULATES THE VARIANCE-COVARIANCE OR CORRELATION
 C MATRIX.

C

C

SUBROUTINE MULTV7(STANDZ,C,NROW,NCL)

PARAMETER (ICL=10,MAXN=8000)

C

C STANDZ=DATA MATRIX WHICH CONTAINS THE

C STADARDISED DATA.

C C=DATA MATRIX WHICH WILL CONTAIN THE SQUARE

C MATRIX.

C MAXN=A DIMENSIONAL VECTOR BIGGER THAN THE

C NUMBER OF ROWS. THIS

C CAN BE ADJUSTED TO A NUMBER TO SUIT ANY

C REQUIREMENT.

C ICI= A DIMENSIONAL VECTOR BIGGER THAN THE NUMBER OF

C COLUMNS AND

C CAN BE ADJUSTED.

DIMENSIONDMEAN(MAXN),VAR(MAXN),STD(MAXN),SOS(ICL,

* OS(ICL),SUMSP(MAXN,ICL)

DIMENSIONSP(ICL,ICL,MAXN),SUMM(ICL),SUM(ICL),

```

* STANDZ(MAXN,ICL),C(MAXN,ICL)
C
C STANDZ.....IS THE STANDARDIZED DATA MATRIX
C C.....IS THE MATRIX CONTAINS THE VARIANCE-COVARIANCE
C OR CORRELATION MATRIX.
C
      DO 25 M=1,NCL
      SUM(M)=0.0
      do 15 j=1,NROW
      SUM(M)=SUM(M)+STANDZ(j,M)
15  CONTINUE
      DMEAN(M)=SUM(M)/NROW
25  CONTINUE
      DO 199 M=1,NCL
      SUMM(M)=0.0
      DO 99 I=1,NROW
      OS(M)=STANDZ(I,M)-DMEAN(M)
      SOS(M)=OS(M)**2
      SUMM(M)=SUMM(M)+SOS(M)
      VAR(M)=SUMM(M)/(NROW-1.0)
      STD(M)=SQRT(VAR(M))
99  CONTINUE
199 CONTINUE
      DO 200 I=1,NCL
      DO 100 K=1,NCL
      SUMSP(i,k)=0.0
      DO 150 L=1,NROW
      SP(I,K,L)=(STANDZ(L,I)-DMEAN(I))*(STANDZ(L,K)-DMEAN(K))
      SUMSP(I,K)=SUMSP(I,K)+SP(I,K,L)
150 CONTINUE
100 CONTINUE
200 CONTINUE
C
      DO 18 I=1,NCL
      DO 17 J=1,NCL
C "C" IS A TWO DIMENSIONAL ARRAY CONTAINS THE VARIANCE-
C COVARIANCE MATRIX.OR CORRELATION MATRIX.

```

```

C
C(I,J)=SUMSP(I,J)/(NROW-1.0)
C
17 CONTINUE
18 CONTINUE
RETURN
END
C
C          SUBROUTINE SMOOTH
C
C THIS IS SMOOTHING SUBROUTINE TO SMOOTH THE PRINCIPAL
C COMPONENT SCORES.
C
C ALL DIMENSIONAL ARRAYS CAN BE ADJUSTED TO ANY SIZE.
SUBROUTINE SMOOTH(PP,SMOOTHD,NROW,LEN)
REAL SMOOTHD(8000),PP(8000)
INTEGER T
C PP=DATA VECTOR TO BE SMOOTHED
C SMOOTHD=A VECTOR WHICH WILL CONTAIN THE OUTPUT
C SMOOTHED DATA.
C LEN IS SIZE OF THE WINDOW (NO OF SAMPLES)
C PP IS A VECTOR CONTAINING THE DATA TO BE SMOOTHED.
C SMOOTHD IS A VECTOR CONTAINING THE OUTPUT SMOOTHED
C DATA.
C NROW IS THE NUMBER OF DATA SAMPLES TO BE SMOOTHED.
C
M=(LEN-1)/2
DO 434 J=1+(LEN-1)/2,NROW(LEN-1)/2
SUM = 0.0
DO 433 T=-M,M
SUM = SUM + PP(J+T)
433 CONTINUE
SMOOTHD(J) = SUM
SMOOTHD(J)=SMOOTHD(J)/LEN
434 CONTINUE
C
RETURN

```

END

```

C
C
C           SUBROUTINE BOUNDARY
C
C SUBROUTINE BOUNDARY TO IDENTIFY THE FORMATION
C BOUNDARIES.
C THIS TECHNIQUE IS BASED ON:
C
C           DSQUARE=(MEAN1-MEAN2)/(VAR1+VAR2)
C
C
C           SUBROUTINE BOUNDARY(SMOOTHD,IBWIN,NROW,DSQUARE)
C
C ALL DIMENSIONAL ARRAYS CAB BE ADJUSTED TO ANY SIZE.
C
C           DIMENSION SMOOTHD(8000),RMEAN(8000),DSQUARE(8000)
C           DIMENSION SMEAN1(8000),OS1(8000),RVAR(8000),SMEAN2(8000),
C * VAR1(8000),VAR2(8000),OS2(8000)
C
C SMOOTHD=DATA VECTOR FROM WHCIH BOUNDARIES WILL BE
C IDENTIFIED.
C DSQURE=A VECTOR WILL CONTAIN THE OUTPUT BOUNDARIES.
C A 'MOVING WINDOW' IS USED HERE
C SMOOTHED IS A VECTOR OF DATA SEQUNCE TO BE ANALYSED.
C IBWIN IS THE SIZE OF THE WINDOW TO BE USED.
C DSQUARE IS A VECTOR WHICH WILL CONTAIN THE OUTPUT D
C SQUARED VALUES (BOUNDARIES OF FORMATIONS)
C
C           M=(IBWIN/2)
C           DO 30 J=M+1,NROW-M
C           SUM1=0.0
C           SUM2=0.0
C           DO 40 K=J-M,J-1
C           SUM1=SUM1+SMOOTHD(K)
C           SUM2=SUM2+SMOOTHD(K+M-1)

```

```

40  CONTINUE
    SMEAN1(J)=SUM1/M
    SMEAN2(J)=SUM2/M
C
    RMEAN(J)=(SMEAN1(J)-SMEAN2(J))**2
C
    SUMOS1=0.0
    SUMOS2=0.0
    DO 50 K=J-M,J-1
    OS1(K)=(SMOOTH(K)-SMEAN1(J))**2
    OS2(K)=(SMOOTH(K+M-1)-SMEAN2(J))**2
    SUMOS1=SUMOS1+OS1(K)
    SUMOS2=SUMOS2+OS2(K)
50  CONTINUE
    VAR1(J)=SUMOS1/(M-1)
    VAR2(J)=SUMOS2/(M-1)
C
    RVAR(J)=VAR1(J)+VAR2(J)
C
    DSQUARE(J)=(RMEAN(J)/RVAR(J))**2
30  CONTINUE
    RETURN
    END
C
C
C          SUBROUTINE XCOR
C
C THIS SUBROUTINE IS TO PERFORM THE CROSS-CORRELATION
C BETWEEN TWO SEQUENCES (SERIES) OR LOGS.
C
C PROCESSES TO DETERMINE THE STRETCH FACTOR AND
C RELATIVE DISPLACEMENT BETWEEN TWO SEQUENCES (SERIES) OF
C WELL-LOGS. CROSS-CORRELATION (WITH VARIABLE WINDOW
C SIZE) OF THE POWER SPECTRA OF TWO SERIES
C IDENTIFIES THE DIRECTION AND AMOUNT OF STRETCH
C BETWEEN TWO SERIES. THE PROCESS INVOLVES THE
C COMPUTATIONS OF POWER SPECTRA IN THE

```

C FREQUENCY DOMAIN WITH THE FREQUENCY INTERVALS
 C TRASFORMED TO A LOGARITHMIC SCALE. LAGRANGE'S
 C METHOD OF INTERPOLATION OBTAINS
 C EQUALLY SPACED POWER SPECTRA FOR CORRELATION. USING
 C TOP TWO PEAK VALUES OF THE CROSS-CORRELATION FUNCTION
 C OF POWER SPECTRA, SERIES ARE THEN STRETCHED BY THE FFT
 C (FAST FOURIER TRANSFORM) INTERPOLATION METHOD. THE
 C LARGEST COEFFICIENT OBTAINED FROM CROSS-CORRELATION
 C (WITH FIXED WINDOW SIZE) OF EACH SET OF SUCH STRETCHED
 C SERIES DETERMINES THE OPTIMUM DISPLACEMENT AND
 C STRETCH.
 C OUTPUTS CONSIST OF A SCREEN LIST OF THE INPUT DATA,
 C COEFFICIENTS OF THE CROSS-CORRELATION FUNCTION OF
 C POWER SPECTRA AND THE OPTIMUM STRETCH AND
 C DISPLACEMENT VALUES.
 C
 C NROW=NUMBER OF DATA SETS TO BE CORRELATED.
 C LS = NUMBER OF DATA POINTS OF THE SHORT SERIES.
 C LL = NUMBER OF DATA POINTS OF THE LONG SERIES.
 C IDER = 1 DERIVATIVE IS WANTED TO COMPUTE POWER SPECTRA
 C = 0 DERIVATIVE IS NOT WANTED.
 C IORG = 1 PRINCIPAL COMPONENTS ARE WANTED FOR
 C STRETCHING AND FOLLOWING CORRELATION.
 C = 0 DERIVATIVE DATA IS WANTED FOR STRETCHING AND
 C FOLLOWING CORRELATION.
 C SMAX = MAXIMUM ANTICIPATED STRETCH VALUE. TYPICAL
 C VALUE = 2.0
 C FMTOP1 = DEPTH OF THE SHORT SERIES.
 C FMTOP2 = DEPTH OF THE LONG SERIES.
 C
 C THIS SUBROUTINE AND THE SUBROUTINES THAT IS CALLING
 C ARE WRITTEN BY BYUNG-DOO KWON, GEOLOGY DEPARTMENT,
 C INDIANA UNIVERSITY,
 C BLOOMINGTON, INDIANA (USA).
 C
 C ALL DIMENSIONAL ARRAYS CAN BE ADJUSTED TO ANY SIZE.
 C

SUBROUTINE XCOR(NROW)

```

C
DIMENSION RLOG1(12800),RLOG2(12800),YIP1(12800),YIP2(12800)
DIMENSION CLOG1(12800),CLOG2(12800),WORK(12600)
DIMENSION XCORL(12100),XCORS(12100),TITLE(10)
COMPLEX CLOG1,CLOG2
DATA LONG /5H LONG/
DATA SHORT /5HSHORT/

C
CHARACTER*10FILEIN1,FILEIN2,DERIVT,PWSPEC,XCORR,
* TRANFRQ,PARAFILE, ORIGFILE,INTSPEC,CRSTRETCH

C
C READ THE NUMBER OF DATA SETS TO BE CORRELATED
C
C
WRITE(6,*)'ENTER THE INPUT FILE of THE SHORT LOG'
READ*,FILEIN1
WRITE(6,*)'ENTER THE INPUT FILE of THE LONG LOG'
READ*,FILEIN2
WRITE(6,*)'ENTER THE INPUT FILE OF THE PARAMETERS'
READ*,PARAFILE
WRITE(6,*)'ENTER THE OUTPUT FILE NO: 1 [ ORIGINAL DATA ]'
READ*,ORIGFILE
WRITE(6,*)'ENTER OUTPUT FILE NO:2 [ DERIVATIVE DATA ]'
READ*,DERIVT
WRITE(6,*)'ENTER OUTPUT FILE NO:3 [ POWER SPECTRA DATA
* ]'
READ*,PWSPEC
WRITE(6,*)'ENTER OUTPUT FILE NO:4 INTERPOLATED POWER
* SPECTRA DATA]'
READ*,INTSPEC
WRITE(6,*)'ENTER OUTPUT FILE NO:5 [ NORMALIZED CROSS-
* CORRELATION DATA]'
READ*,XCORR
WRITE(6,*)'ENTER OUTPUT FILE NO:6 [ FOR TRANSFORMED
* FREQUENCY ]'
READ*,TRANFRQ

```

C
C

```

OPEN(1,FILE=FILEIN1)
OPEN(2,FILE=ORIGFILE)
OPEN(3,FILE=DERIVT)
OPEN(4,FILE=PWSPEC)
OPEN(7,FILE=XCORR)
OPEN(8,FILE=TRANFRQ)
OPEN(9,FILE=FILEIN2)
OPEN(10,FILE=PARAFILE)
OPEN(11,FILE=INTSPEC)

```

C

C READ PARAMETERS AND INPUT DATA (TWO SHORT AND TWO
C LONG LOGS)

C

C

```

READ(10,*)NSET
DO 290 IJ=1,NSET

```

C INITIALIZE ALL ARRAYS TO ZERO

C SET THE MAXIMUM DATA LENGTH TO 12800 (MAXIMUM DEPTH
C OF THE BOREHOLE).

C

```

DO 10 I=1,12800
RLOG1(I)=0.0
RLOG2(I)=0.0
YIP1(I)=0.0
YIP2(I)=0.0
WORK(I)=0.0
WORK(I+12800)=0.0
CLOG1(I)=CMPLX(0.0,0.0)
CLOG2(I)=CMPLX(0.0,0.0)

```

10 CONTINUE

```

DO 20 I=1,12800

```

```

XCORL(I)=0.0

```

20 XCORS(I)=0.0

C

C READ AND WRITE PARAMETERS AND INPUT DATA

```

C
  READ(10,298) (TITLE(I),I=1,8)
  READ(10,298) ITITLE
  READ(10,*)IDER,IORG,SMAX,FMTOP1,FMTOP2
C
C
C READ THE INFORMATION ABOUT THE DEPTH AND THICKNESS
C FROM DATAFILE.
C [ MFTOP1 ] IS THE TOP OF FORMATION TO BE CORRELATED, AND
C [ THICK1 ] IS IT'S THICKNESS. [ FMTOP2 ] IS THE TOP OF THE
C FORMATION TO BE CORRELATED WITH [ LONG SERIES], AND [
C THICK2 ] IS IT'S THICKNESS.
C
  READ(10,*)THICK1,THICK2
C
C THE ACTUAL DEPTH OF THE FORMATION [ FROM PCA ] IS
C EQUAL TO THE REAL DEPTH ON LOG MINUS THE DEPTH THAT
C THE WELL DATAFILE STARTES FROM. THIS DEPTH CORRECTION
C IS NECESSARY IN ORDER TO GET AN ACCURATE DEPTH FROM
C THE PRINCIPAL COMPONENT SCORES.
C
  WRITE(6,*)'ENTER LENGTH OF FORMATION 1'
  READ*,LS
  WRITE(6,*)'ENTER LENGTH OF FORMATION 2'
  READ*,LL
C
C
C TO READ THE DESIRED FORMATION TO BE CORRELATED THEN
C ADD THE THICKNESS OF THE FORMATION TO THE ACTUAL
C DEPTH
C
  READ(1,302) (RLOG1(I),I=1,LS)
C
  REWIND 1
C
  I2=FMTOP1
  DO 111 I=1,THICK1

```

```

RLOG1(I)=RLOG1(I2)
I2=I2+1
111 CONTINUE
LS=THICK1
C
READ(9,302) (RLOG2(I),I=1,LL)
C
REWIND 9
C
I3=FMTOP2
DO 112 I=1,THICK2
RLOG2(I)=RLOG2(I3)
I3=I3+1
112 CONTINUE
LL=THICK2
C
C KEEP THE ORIGINAL DATA IN A FILE FOR PLOT
C
WRITE(2,'(F10.3)') (RLOG1(I),I=1,LS)
WRITE(2,'(F10.3)') (RLOG2(I),I=1,LL)
WRITE(6,299) (TITLE(I),I=1,8)
WRITE(6,300) ITITLE
WRITE(6,303) LS,LL,IDER,IORG,SMAX,FMTOP1,FMTOP2
WRITE(6,304)
DO 30 I=1,LS
30 WRITE(6,305) I,RLOG1(I),RLOG2(I)
LS1=LS+1
DO 40 I=LS1,LL
40 WRITE(6,306) I,RLOG2(I)
C
C CHECK WHETHER DERIVATIVE IS WANTED
C
IF(IDER.EQ.0) GO TO 80
CALL DERIVAT (RLOG1,LS)
RLOG1(LS+1)=0.0
CALL DERIVAT (RLOG2,LL)
C

```

```

WRITE(6,307)
DO 50 I=1,LS
50  WRITE(6,305) I,RLOG1(I),RLOG2(I)
    LS1=LS+1
    DO 60 I=LS1,LL
60  WRITE(6,306) I,RLOG2(I)
C
C
C KEEP THE DERIVATIVE DATA FOR PLOT
C
    WRITE(3,'(F10.3)') (RLOG1(I),I=1,LL)
    WRITE(3,'(F10.3)') (RLOG2(I),I=1,LL)
80  CONTINUE
C
C CONSTRUCT COMPLEX SERIES AND DO FOURIER TRANSFORM
C
    DO 90 I=1,LL
    CLOG1(I)=CMPLX(RLOG1(I),0.0)
    CLOG2(I)=CMPLX(RLOG2(I),0.0)
90  CONTINUE
    CALL FOURT (CLOG1,LL,1,-1,1,WORK)
    CALL FOURT (CLOG2,LL,1,-1,1,WORK)
C
C COMPUTE POWER SPECTRA (THE SECOND HALF ABOVE NYQUIST
C FREQUENCY IS IGNORED)
C
    NYQ=LL/2+1
    DO 100 I=2,NYQ
    RLOG1(I-1)=(REAL(CLOG1(I))**2+
*   AIMAG(CLOG1(I))**2)/FLOAT(LL)
    RLOG2(I-1)=(REAL(CLOG2(I))**2+
*   AIMAG(CLOG2(I))**2)/FLOAT(LL)
100 CONTINUE
    NN=NYQ-1
    WRITE(6,308)
    DO 110 I=1,NN
110  WRITE(6,309) I,CLOG1(I+1),RLOG1(I),CLOG2(I+1),RLOG2(I)

```

```

        IF (PRALL.EQ.0.0) GO TO 120
120   CONTINUE
C
C KEEP THE POWER SPECTRA IN A FILE FOR PLOT
C
        WRITE(4,'(F10.3)') (RLOG1(I),I=1,NN)
        WRITE(4,'(F10.3)') (RLOG2(I),I=1,NN)
C
C TRANSFORM THE FREQUENCIES INTO A LOGARITHMIC SCALE
C
        DO 130 I=1,NN
130   WORK(I)=ALOG10(FLOAT(I))
        WRITE(8,'(F10.3)') (WORK(I),I=1,NN)
C
C OBTAIN EQUALLY SPACED POWER SPECTRA USING LAGRANGE'S
C INTERPOLATION METHOD
C
        JLAST=NN-2
        DELT=0.01
        CALL INTPOL3 (WORK,RLOG1,RLOG2,YIP1,YIP2,10,JLAST,
*   NLAST,DELT)
C
C KEEP INTERPOLATED SPECTRA IN A FILE FOR PLOT
C
        WRITE(11,'(F10.3)') (YIP1(I),I=1,NLAST)
        WRITE(11,'(F10.3)') (YIP2(I),I=1,NLAST)
C
        WRITE(6,310)
        DO 140 I=1,NLAST
140   WRITE(6,305) I,YIP1(I),YIP2(I)
C
C CROSS-CORRELATE INTERPOLATED POWER SPECTRA TO OBTAIN
C STRETCH VALUES.
C
        LAGMAX=ALOG10(SMAX)/DELT+1.5
        CALL CROSS1 (YIP1,YIP2,XCORL,NLAST,LAGMAX)
        CALL CROSS1 (YIP2,YIP1,XCORS,NLAST,LAGMAX)

```

```

WRITE(6,313)
DO 160 I=1,LAGMAX
  K1=-I+1
  K2=I-1
160  WRITE(6,312) K1,XCORL(I),K2,XCORS(I)
  WRITE(6,311)
  LAGTOT=2*LAGMAX-1
  DO 170 I=1,LAGMAX
  WORK(I)=FLOAT(-LAGMAX+I)
170  YIP1(I)=XCORL(LAGMAX-I+1)
  DO 180 I=2,LAGMAX
  WORK(LAGMAX+I-1)=FLOAT(I-1)
180  YIP1(LAGMAX+I-1)=XCORS(I)
C
C KEEP THE CROSS-CORRELATION FUNCTION OF POWER SPECTRA
C IN A FILE FOR PLOT.
C
  WRITE(7,'(F10.3)') (WORK(I),I=1,LAGTOT)
  WRITE(7,'(F10.3)') (YIP1(I),I=1,LAGTOT)
C
C FIND THE MAXIMUM PEAK IN THE CORRELATION FUNCTION OF
C POWER SPECTRA AND COMPUTE CORRESPONDING STRETCH
C FACTOR.
C
  CALL MAX (YIP1,1,LAGTOT,I1,PCMAX1)
  XLAG1=WORK(I1)
  DEL1=ABS(XLAG1)*DELT
  ST1=10.**DEL1
C
C FIND SECOND PEAK IN THE CORRELATION FUNCTION OF POWER
C SPECTRA AND COMPUTE CORRESPONDING STRETCH FACTOR
C
  CALL SCAN (YIP1,I1,LAGTOT)
  CALL MAX (YIP1,1,LAGTOT,I2,PCMAX2)
  XLAG2=WORK(I2)
  DEL2=ABS(XLAG2)*DELT
  ST2=10.**DEL2

```

```

C
C FROM TWO PEAK VALUES, FIND THE OPTIMUM DISPLACEMENT
C AND STRETCH
C
      IF(XLAG1.GT.0.0) GO TO 190
C
C STRETCHING AND CORRELATING THE FIRST PEAK ASSUMES THE
C LONG SERIES (SEQUENCE) IS STRETCHED
C
      WRITE(6,315) ST1
      CALL STXCO1
      * (RLOG1,RLOG2,CLOG1,WORK,YIP1,LS,LL,ST1,ML1,ID1,
      * CMAX1,IDER,IORG)
      .IF (XLAG2.GT.0.0) GO TO 210
      GO TO 200
C
C STRETCHING AND CORRELATING THE FIRST PEAK ASSUMES THE
C SHORT SERIES (SEQUENCE) IS STRETCHED.
C
190  WRITE(6,314) ST1
      CALL STXCO2 (RLOG1,RLOG2,CLOG1,WORK,YIP1,LS,LL,ST1,ML1,
      * ID1,CMAX1,IDER,IORG)
      IF(XLAG2.GT.0.0) GO TO 210
C
C STRETCHING AND CORRELATING THE SECOND PEAK ASSUMES THE
C LONG SERIES (SEQUENCE) IS STRETCHED.
C
200  WRITE(6,317) ST2
      CALL STXCO1 (RLOG1,RLOG2,CLOG2,WORK,YIP2,LS,LL,ST2,
      * ML2,ID2,CMAX2,IDER,IORG)
      GO TO 220
C
C STRETCHING AND CORRELATING THE SECOND PEAK ASSUMES
C THE SHORT SERIES (SEQUENCE) IS STRETCHED.
C
210  WRITE(6,316) ST2
      CALLSTXCO2 (RLOG1,RLOG2,CLOG2,WORK,YIP2,LS,LL,ST2,ML2,

```

* ID2,CMAX2,IDER,IORG)

C

C COMPARE THE COEFICIENTS OBTAINED FROM CORRELATIONS
C TWO SETS OF STRETCHED SERIES.

C

220 IF(CMAX1.LT.CMAX2) GO TO 230

CMAX=CMAX1

ST=ST1

ML=ML1

ID=ID1

C

WRITE(6,*)

WRITE(6,*)'CROSS-CORRELATION FUNCTION OF STRETCHED
* SERIES'

READ*,CRSTRETCH

OPEN(12,FILE=CRSTRETCH)

C

WRITE(12,'(F10.3)') (YIP1(I),I=1,ML)

IF(XLAG1.GT.0.0) GO TO 240

GO TO 260

230 CMAX=CMAX2

ST=ST2

ML=ML2

ID=ID2

WRITE(12,'(F10.3)') (YIP2(I),I=1,ML)

240 IF(XLAG2.GT.0.0) GO TO 250

GO TO 260

C

C THE FINAL RESULT SUGGESTS THAT THE SHORT SERIES IS
C STRETCHED.

C PLOT THE CORRELATION RESULT.

C

250 ID=FLOAT(ID)/ST+0.5

WRITE(6,318) ST,CMAX,ID

IDEND=FLOAT(ID)+(FLOAT(LS)/ST)

C

GO TO 280

```

C
C THE FINAL RESULT SUGGESTS THAT THE LONG SERIES IS
C STRETCHED.
C
260  WRITE(6,319) ST,CMAX,ID
      IDEND=FLOAT(ID)+(FLOAT(LS)*ST)
C
280  CONTINUE
      REWIND 2
      REWIND 3
290  CONTINUE
C
C FORMATS
C
298  FORMAT(8A10)
299  FORMAT(1H1,8A10,/)
300  FORMAT(3X,A10)
302  FORMAT(F10.3)
303  FORMAT(3X,'LS=',I5,3X,'LL=',I5,3X,'IDER=',I2,3X,'IORG=',I2,
* 3X,'SMAX=',F5.1,/,3X,'DEPTH OF SHORT SERIES =',
* F6.1,' FEET',/3X,'DEPTH OF LONG SERIES =',F6.1,' FEET',/)
304  FORMAT(1H0,10X,'INPUT DATA',/,/,10X,'SHORT SERIES LONG
* SERIES',/)
305  FORMAT(I5,2F10.3)
306  FORMAT(I5,10X,F10.3)
307  FORMAT(//,8X,'DERIVATED DATA',/,/,10X,'SHORT SERIES
* LONG SERIES',/)
308  FORMAT(//,30X,'FOURIER TRANSFORM',/,/,15X,'SERIES
* 1',35X,'SERIES 2',/,/,10X,'REAL',3X,'IMAGINARY',2X,'POWER
* SPECTRUM',7X,'REAL',3X,'IMAGINARY',2X,'POWER
* SPECTRUM',/)
309  FORMAT(I5,3F10.3,10X,3F10.3)
310  FORMAT(//,10X,'INTERPOLATED POWER SPECTRUM ( START
* FROM 10TH OF ORIGINAL )',/,/,10X,'SHORT SERIES LONG
* SERIES')
311  FORMAT(///' STRETCH FACTOR FOUND FROM CORRELATION
* OF POWER SPECTRA')

```

```

312  FORMAT(10X,I5,F15.3,22X,I5,F15.3)
313  FORMAT(//,20X,'NORMALIZEDCORRELATION COEFFICIENTS',/,
*    10X,'( ASSUME LONG SERIES IS STRETCHED )',10X,
*    '( ASSUME SHORT SERIES IS STRETCHED )',//,8X,'LAG
*    NUMBER', 5X,'VALUE OF COEFFICIENT',7X,'LAG NUMBER',5X,
*    'VALUE OF COEFFICIENT',/)
314  FORMAT(//,' FIRST CHOICE - SHORT SERIES I STRETCHED',F6.2,
*    'TIMES')
315  FORMAT(//,' FIRST CHOICE - LONG SERIES IS STRETCHED',F6.2,
*    ' TIMES')
316  FORMAT(/,' SECOND CHOICE - SHORT SERIES IS
*    STRETCHED',F6.2,' TIMES')
317  FORMAT(/,' SECOND CHOICE - LONG SERIES IS
*    STRETCHED',F6.2,' TIMES')
318  FORMAT(///,' FINAL RESULT SUGGESTS THAT SHORT SERIES
*    IS STRETCHED', F5.2,' TIMES',//,' MAXIMUM CORRELATION
*    IS',F5.3,' AT A LAG OF',I5)
319  FORMAT(///,' FINAL RESULT SUGGEST THAT LONG SERIES IS
*    STRETCHED',F5.2,' TIMES',//,' MAXIMUM CORRELATION
*    IS',F5.3,' AT A LAG OF',I3)
      RETURN
      END

C
C          SUBROUTINE MEAN
C
C TO CALCULATE THE MEAN
C
C          SUBROUTINE MEAN (A,N)
C
      DIMENSION A(1)
      TOT=0.0
      DO 10 I=1,N
10     TOT=TOT+A(I)
      AMEAN=TOT/FLOAT(N)
      DO 20 I=1,N
20     A(I)=A(I)-AMEAN
      RETURN

```

END

C
C
C SUBROUTINE MAX
C
C TO FIND THE MAXIMUM CORRELATION COEFFICIENT OF THE
C CROSS-CORRELATION.

C
C
C
C SUBROUTINE MAX (A,M,N,ID,AMAX)

C
C FIND THE MAXIMUM (AMAX) AND ITS POSITION (ID)

C
C , DIMENSION A(1)
C AMAX=0.0
C DO 1 I=M,N
C IF(A(I).GT.AMAX) GO TO 2
C GO TO 1
C 2 AMAX=A(I)
C ID=I
C 1 CONTINUE
C RETURN
C END

C
C SUBROUTINE DERIVAT
C
C TO REPLACE THE DATA BY THEIR FIRST DERIVATIVES.

C
C
C SUBROUTINE DERIVAT (A,N)

C
C DIMENSION A(1)
C N=N-1
C DO 10 I=1,N
C 10 A(I)=A(I+1)-A(I)
C RETURN


```

B4=TXIP-X(J+2)
P1=B2*B3*B4
P2=B1*B3*B4
P3=B1*B2*B4
P4=B1*B2*B3
YIP1(NSEQ)=(C1*P1*RLOG1(J-1))+(C2*P2*RLOG1(J))+
* (C3*P3*RLOG1(J+1))+(C4*P4*RLOG1(J+2))
YIP2(NSEQ)=(C1*P1*RLOG2(J-1))+(C2*P2*RLOG2(J))+
* (C3*P3*RLOG2(J+1))+(C4*P4*RLOG2(J+2))
IF (YIP1(NSEQ).LT.0.) YIP1(NSEQ)=0.0
IF (YIP2(NSEQ).LT.0.) YIP2(NSEQ)=0.0
NSEQ=NSEQ+1
GO TO 2
1 CONTINUE
  NLAST=NSEQ-1
  RETURN
  END

C
C          SUBROUTINE CROSS1
C
C NORMALIZED CROSS-CORRELATION WITH A VARIABLE
C WINDOW SIZE.
C
C
C          SUBROUTINE CROSS1 (A,B,C,L,ML)
C
C          DIMENSION A(1),B(1),C(1)
          ATOT=0.0
          BTOT=0.0
          ASQ=0.0
          BSQ=0.0
          DO 1 I=1,L
            ATOT=ATOT+A(I)
            BTOT=BTOT+B(I)
            ASQ=ASQ+A(I)**2
1          BSQ=BSQ+B(I)**2
          DO 2 J=1,ML

```

```

AB=0.0
N=L-J+1
DO 3 K=1,N
3 AB=AB+(A(K+J-1)*B(K))
  CNUM=AB-(ATOT*BTOT/FLOAT(N))
  CDEN=SQRT((ASQ-(ATOT**2/FLOAT(N)))*
* (BSQ-(BTOT**2/FLOAT(N))))
  IF(CDEN.EQ.0.0) CDEN=100000000.
  C(J)=CNUM/CDEN
  ATOT=ATOT-A(J)
  BTOT=BTOT-B(L-J+1)
  ASQ=ASQ-A(J)**2
  BSQ=BSQ-B(L-J+1)**2
2 CONTINUE
  RETURN
  END

```

C

C

C

SUBROUTINE CROSS2

C

C NORMALIZED CROSS-CORRELATION WITH A FIXED WINDOW
C SIZE.

C

C

```

SUBROUTINE CROSS2 (A,B,C,L1,L2,ML)

```

C

C

```

DIMENSION A(1),B(1),C(1)

```

```

ATOT=0.0

```

```

BTOT=0.0

```

```

ASQ=0.0

```

```

BSQ=0.0

```

```

DO 1 I=1,L1

```

```

ATOT=ATOT+A(I)

```

```

BTOT=BTOT+B(I)

```

```

ASQ=ASQ+A(I)**2

```

1

```

BSQ=BSQ+B(I)**2

```

```

ML=L2-L1+1
DO 2 J=1,ML
AB=0.0
DO 3 K=1,L1
3 AB=AB+(A(K)*B(K+J-1))
CNUM=AB-(ATOT*BTOT/FLOAT(L1))
CDEN=SQRT((ASQ-(ATOT**2/FLOAT(L1)))*
* (BSQ-(BTOT**2/FLOAT(L1))))
IF(CDEN.EQ.0.0) CDEN=100000000.
C(J)=CNUM/CDEN
BTOT=BTOT-B(J)+B(L1+J)
BSQ=BSQ-B(J)**2+B(L1+J)**2
2 CONTINUE
RETURN
END

```

```

C
C
C          SUBROUTINE SCAN
C
C SCAN CORRELATION COEFFICIENTS TO DETERMINE SECOND BEST
C STRETCH FACTOR.

```

```

C
C
C          SUBROUTINE SCAN (A,ID,LAGMAX)

```

```

C
C
DIMENSION A(1)
ID1=ID+1
LMAX=LAGMAX-1
IF (ID1.GE.LAGMAX) GO TO 3
DO 1 I=ID1,LMAX
IF ((A(I+1)-A(I)).LT.0.0) GO TO 2
GO TO 4
2A(I)=-1.0
IF(I.EQ.LMAX) A(LAGMAX)=-1.0
1 CONTINUE
3 A(ID1)=-1.0

```

```

4   LAST=ID-2
   IF (LAST.LT.1) GO TO 7
   DO 5 J=1, LAST
   K=ID-J
   IF((A(K-1)-A(K)).LT.0.0) GO TO 6
   GO TO 8
6   A(K)=-1.0
   IF (K.EQ.2) A(1)=-1.0
5   CONTINUE
7   A(ID-1)=-1.0
8   A(ID)=0.0
   RETURN
   END

```

```

C
C
C
C

```

SUBROUTINE STXCO1

```

C STRETCH THE SHORT SERIES BY FFT INTERPOLATION METHOD
C AND CROSS-CORRELATE WITH THE LONG SERIES
C FIND THE MAXIMUM CORRELATION COEFFICIENT
C
C

```

```

C
C

```

SUBROUTINE STXCO1

```

* (RLOG1,RLOG2,CLOG1,WORK,XCOR,LS,LL,ST,ML1,
* ID1,CMAX1,IDER,IORG)

```

```

C

```

```

DIMENSION RLOG1(1),RLOG2(1),CLOG1(1),WORK(1),XCOR(1)
COMPLEX CLOG1
REWIND 2
REWIND 3

```

```

C
C
C
C

```

```

UNITS 2 CONTIANS THE ORIGINAL DATA
UNIT 3 CONTAINS THE DERIVATIVE DATA

READ(2,302) (RLOG1(I),I=1,LS)
READ(2,302) (RLOG2(I),I=1,LL)
IF (IDER.EQ.0.OR.IORG.NE.0.) GO TO 1

```

```

      READ(3,302) (RLOG1(I),I=1,LS)
      READ(3,302) (RLOG2(I),I=1,LL)
1     M=FLOAT(LS)*ST+0.5
      CALL STRETCH (RLOG1,CLOG1,WORK,LS,M)
      CALL CROSS2 (RLOG1,RLOG2,XCOR,M,LL,ML1)
      CALL MAX (XCOR,1,ML1,ID1,CMAX1)
302  FORMAT(F10.3)
      RETURN
      END

```

```

C
C
C
C

```

SUBROUTINE STXCO2

```

C STRETCH THE LONG SERIES BY FFT INTERPOLATION METHOD
C AND CROSS-CORRELATE WITH THE SHORT SERIES
C FIND THE MAXIMUM CORRELATION COEFFICIENT.
C
C

```

```

C
C

```

```

      SUBROUTINE STXCO2

```

```

* (RLOG1,RLOG2,CLOG2,WORK,XCOR,LS,LL,ST,ML2,
* ID2,CMAX2,IDER,IORG)

```

```

C
C

```

```

      DIMENSION RLOG1(1),RLOG2(1),CLOG2(1),WORK(1),XCOR(1)

```

```

      COMPLEX CLOG2

```

```

      REWIND 2

```

```

      REWIND 3

```

```

C

```

```

C UNIT 2 CONTAINS THE ORIGINAL DATA

```

```

C UNIT 3 CONTAINS THE DERIVATIVE DATA.

```

```

C

```

```

      READ(2,302) (RLOG1(I),I=1,LS)

```

```

      READ(2,302) (RLOG2(I),I=1,LL)

```

```

      IF (IDER.EQ.0.OR.IORG.NE.0.) GO TO 1

```

```

C

```

```

      READ(3,302) (RLOG1(I),I=1,LS)

```

```

      READ(3,302) (RLOG2(I),I=1,LL)

```

```

1   M=FLOAT(LL)*ST+0.5
    CALL STRETCH (RLOG2,CLOG2,WORK,LL,M)
    CALL CROSS2 (RLOG1,RLOG2,XCOR,LS,M,ML2)
    CALL MAX (XCOR,1,ML2,ID2,CMAX2)
302 FORMAT(F10.3)
    RETURN
    END

C
C
C           SUBROUTINE STRETCH
C
C INTERPOLATE TIME SERIES DATA WITH N VALUES TO A SERIES
C WITH M VALUES IN THE FREQUENCY DOMAIN.
C
C
C   SUBROUTINE STRETCH (RA,A,WORK,N,M)
C
C
C   DIMENSION WORK(1),RA(1),A(1)
C   COMPLEX A
C   DO 5 I=1,N
5   A(I)=CMPLX(RA(I),0.0)
    CALL FOURT (A,N,1,-1,1,WORK)
    IF(N.EQ.M) GO TO 50
C
C SEARCH FOR THE NYQUIST
C
C   K=FLOAT(N)/2.+1.5
C   MN=M-N
C   KZ=K+MN-1
C
C TRANSFER THE CONJUGATE PARTS
C
C   DO 10 I=K,N
10  A(M-I+K)=A(N-I+K)
C
C CHECK IF INPUT DATA TOTAL IS EVEN OR ODD

```

```

C
  IF((N/2*2).EQ.N) GO TO 20
  GO TO 30
C
C DIVIDE THE AMPLITUDE OF NYQUIST FREQUENCY BY 2
C FOR THE CASE OF EVEN N.
C
20  A(K)=a(k)/2
    A(K+MN)=a(k)/2
    K=K+1
    IF(M.EQ.(N+1)) GO TO 50
30  CONTINUE
C
C ADD (M-N)ZEROS FOR ODD CASE ,(M-N-1) FOR EVEN CASE
C
    DO 40 I=K,KZ
40  A(I)=0.0
C
C INVERSE F.T.
C
50  CALL FOURT (A,M,1,1,1,WORK)
C
C NORMALIZATION - DIVIDE BY INPUT SIGNAL LENGTH (N)
C
    DO 60 I=1,M
    A(I)=A(I)/FLOAT(N)
    RA(I)=REAL(A(I))
60  CONTINUE
    RETURN
    END
C
C
C          SUBROUTINE FOURT
C
C
C THE COOLEY-TUKEY FAST FOURIER TRANSFORM IN FORTRAN IV
C TRANSFORM(K1,K2,...)=SUM(DATA(J1,J2,...)*EXP(ISIGN*2*PI*SQR(-

```

C 1)*((J1-1)*(K1-1)/NN(1)+(J2-1)*(K2-1)/NN(2)+...)), SUMMED FOR ALL
 C J1, K1 BETWEEN 1 AND NN(1), J2, K2 BETWEEN 1 AND NN(2), ETC.
 C THERE IS NO LIMIT TO THE NUMBER OF SUBSCRIPTS. DATA IS A
 C MULTIDIMENSIONAL COMPLEX ARRAY WHOSE REAL AND
 C IMAGINARY PARTS ARE ADJACENT IN STORAGE, SUCH AS
 C FORTRAN IV PLACES THEM IF ALL IMAGINARY PARTS ARE ZERO
 C (DATA ARE DISGUISED REAL), SET IFORM TO ZERO TO CUT THE
 C RUNNING TIME BY UP TO FORTY PERCENT. OTHERWISE, IFORM =
 C +1. THE LENGTHS OF ALL DIMENSIONS ARE STORED IN ARRAY
 C NN, OF LENGTH NDIM. THEY MAY BE ANY POSITIVE INTEGERS,
 C THO THE PROGRAM RUNS FASTER ON COMPOSITE INTEGERS,
 C AND ESPECIALLY FAST ON NUMBERS RICH IN FACTORS OF TWO
 C ISIGN IS +1 OR -1. IF A -1 TRANSFORM IS FOLLOWED BY A +1 ONE
 C (OR A +1 BY A -1) THE ORIGINAL DATA REAPPEAR, MULTIPLIED BY
 C NTOT (=NN(1)* NN(2)*...).
 C TRANSFORM VALUES ARE ALWAYS COMPLEX, AND ARE
 C RETURNED IN ARRAY DATA, REPLACING THE INPUT. IN
 C ADDITION, IF ALL DIMENSIONS ARE NOT POWERS OF TWO,
 C ARRAY WORK MUST BE SUPPLIED, COMPLEX OF LENGTH EQUAL
 C TO THE LARGEST NON 2**K DIMENSION. OTHERWISE, REPLACE
 C WORK BY ZERO IN THE CALLING SEQUENCE.
 C NORMAL FORTRAN DATA ORDERING IS EXPECTED, FIRST
 C SUBSCRIPT VARYING FASTEST. ALL SUBSCRIPTS BEGIN AT ONE.
 C RUNNING TIME IS MUCH SHORTER THAN THE NAIVE NTOT**2,
 C BEING GIVEN BY THE FOLLOWING FORMULA. DECOMPOSE NTOT
 C INTO 2**K2 * 3**K3 * 5**K5 * LET SUM2 = 2*K2, SUMF = 3*K3 +
 C 5*K5 + .. AND NF = K3 + K5 + THE TIME TAKEN BY A MULTI-
 C NTOT*(T1+T2*SUM2+T3*SUMF+T4*NF). ON THE CDC (FLOATING
 C POINT ADD TIME OF SIX MICROSECONDS), T = 3000
 C +NTOT*(500+43*SUM2+68*SUMF+320*NF) MICROSECONDS ON
 C COMPLEX DATA. IN ADDITION, THE ACCURACY IS GREATLY
 C IMPROVED, AS THE RMS RELATIVE ERROR IS BOUNDED BY 3*2**(-
 C B)*SUM(FACTOR(J)**1.5), WHERE B IS THE NUMBER OF BITS IN
 C THE FLOATING POINT FRACTION AND FACTOR(J) ARE THE PRIME
 C FACTORS OF NTOTC PROGRAM BY NORMAN BRENNER FROM
 C THE BASIC PROGRAM BY CHARLES RADER. RALPH ALTER
 C SUGGESTED THE IDEA FOR THE DIGIT REVERSAL. MIT LINCOLN

C LABORATORY, AUGUST 1967. THIS IS THE FASTEST AND MOST
 C VERSATILE VERSION OF THE FFT KNOWN TO THE AUTHOR
 C SHORTER PROGRAMS FOUR1 AND FOUR2 RESTRICT DIMENSION
 C LENGTHS TO POWERS OF TWO. SEE-- IEEE AUDIO TRANSACTIONS
 C (JUNE 1967), SPECIAL ISSUE ON FFT.

C THE DISCRETE FOURIER TRANSFORM PLACES THREE
 C RESTRICTIONS UPON THE DATA

C 1. THE NUMBER OF INPUT DATA AND THE NUMBER OF
 C TRANSFORM VALUES MUST BE THE SAME.

C 2. BOTH THE INPUT DATA AND THE TRANSFORM VALUES MUST
 C REPRESENT EQUISPACED POINTS IN THEIR RESPECTIVE DOMAINS
 C FREQUENCY. CALLING THESE SPACINGS DELTAT AND DELTAF, IT
 C MUST BE TRUE THAT $\text{DELTAF} = 2 * \text{PI} / (\text{NN}(\text{I}) * \text{DELTAT})$. OF COURSE
 C, DELTAT NEED NOT BE THE SAME FOR EVERY DIMENSION.

C 3, CONCEPTUALLY AT LEAST, THE INPUT DATA AND THE
 C TRANSFORM OUTPUT REPRESENT SINGLE CYCLES OF PERIODIC
 C FUNCTIONS.

C

 SUBROUTINE FOURT(DATA,NN,NDIM,ISIGN,IFORM,WORK)

C

 DIMENSION DATA(1),NN(1),IFACT(32),WORK(1)

 TWOPI=6.283185307

 IF(NDIM-1)920,1,1

1 NTOT=2

 DO 2 IDIM=1,NDIM

 IF(NN(IDIM))920,920,2

2 NTOT=NTOT*NN(IDIM)

C

C MAIN LOOP FOR EACH DIMENSION

C

 NP1=2

 DO 910 IDIM=1,NDIM

 N=NN(IDIM)

 NP2=NP1*N

 IF(N-1)920,900,5

C

C FACTOR N

```

C
5  M=N
   NTWO=NP1
   IF=1
   IDIV=2
10  IQUOT=M/IDIV
   IREM=M-IDIV*IQUOT
   IF(IQUOT-IDIV)50,11,11
11  IF(IREM)20,12,20
12  NTWO=NTWO+NTWO
   M=IQUOT
   GO TO 10
20  IDIV=3
30  IQUOT=M/IDIV
   IREM=M-IDIV*IQUOT
   IF(IQUOT-IDIV)60,31,31
31  IF(IREM)40,32,40
32  IFACT(IF)=IDIV
   IF=IF+1
   M=IQUOT
   GO TO 30
40  IDIV=IDIV+2
   GO TO 30
50  IF(IREM)60,51,60
51  NTWO=NTWO+NTWO
   GO TO 70
60  IFACT(IF)=M
C
C  SEPARATE FOUR CASES :
C  1. COMPLEX TRANSFORM OR REAL TRANSFORM FOR THE 4TH,
C     5TH,ETC.
C     DIMENSIONS:
C  2. REAL TRANSFORM FOR THE 2ND OR 3RD DIMENSION
C     METHOD TRANSFORM HALF THE DATA, SUPPLYING THE BY
C     CONJUGATE SYMMETRY.
C  3. REAL TRANSFORM FOR THE 1ST DIMENSION, N ODD.
C     METHOD TRANSFORM HALF THE DATA AT EACH STAGE,

```

C CSUPPLYING THE OTHER HALF BY CONJUGATE SYMMETRY.
 C 4. REAL TRANSFORM FOR THE 1ST DIMENSION, N EVEN.
 C METHOD TRANSFORM A COMPLEX ARRAY OF LENGTH N/2
 C WHOSE REAL PARTS ARE THE EVEN NUMBERED REAL
 C VALUES AND WHOSE IMAGINARY PARTS ARE THE ODD
 C NUMBERED REAL VALUES. SEPARATE AND SUPPLY THE
 C SECOND HALF BY CONJUGATE SYMMETRY.

C
 70 NON2=NP1*(NP2/NTWO)

ICASE=1

IF(IDIM-4)71,90,90

71 IF(IFORM)72,72,90

72 ICASE=2

IF(IDIM-1)73,73,90

73 ICASE=3

IF(NTWO-NP1)90,90,74

74 ICASE=4

NTWO=NTWO/2

N=N/2

NP2=NP2/2

NTOT=NTOT/2

I=3

DO 80 J=2,NTOT

DATA(J)=DATA(I)

80 I=I+2

90 I1RNG=NP1

IF(ICASE-2)100,95,100

95 I1RNG=NP0*(1+NPREV/2)

C

C SHUFFLE ON THE FACTORS OF TWO IN N. AS THE SHUFFLING
 C CAN BE DONE BY SIMPLE INTERCHANGE, NO WORKING ARRAY
 C IS NEEDED.

C

100 IF(NTWO-NP1)600,600,110

110 NP2HF=NP2/2

J=1

DO 150 I2=1,NP2,NON2

```

IF(J-I2)120,130,130
120 I1MAX=I2+NON2-2
DO 125 I1=I2,I1MAX,2
DO 125 I3=I1,NTOT,NP2
J3=J+I3-I2
TEMPR=DATA(I3)
TEMPI=DATA(I3+1)
DATA(I3)=DATA(J3)
DATA(I3+1)=DATA(J3+1)
DATA(J3)=TEMPR
125 DATA(J3+1)=TEMPI
130 M=NP2HF
140 IF(J-M)150,150,145
145 J=J-M
M=M/2
IF(M-NON2)150,140,140
150 J=J+M
C
C MAIN LOOP FOR FACTORS OF TWO. PERFORM FOURIER
C TRANSFORMS OF LENGTH FOUR, WITH ONE OF LENGTH TWO
C IF NEEDED. THE TWIDDLE FACTOR
C  $W = \text{EXP}(\text{ISIGN} * 2 * \text{PI} * \text{SQRT}(-1) * M / (4 * \text{MMAX}))$ . CHECK FOR
C  $W = \text{ISIGN} * \text{SQRT}(-1)$ 
C AND REPEAT FOR  $W = \text{ISIGN} * \text{SQRT}(-1) * \text{CONJUGATE}(W)$ .
C
NON2T=NON2+NON2
IPAR=NTWO/NP1
310 IF(IPAR-2)350,330,320
320 IPAR=IPAR/4
GO TO 310
330 DO 340 I1=1,I1RNG,2
DO 340 J3=I1,NON2,NP1
DO 340 K1=J3,NTOT,NON2T
K2=K1+NON2
TEMPR=DATA(K2)
TEMPI=DATA(K2+1)
DATA(K2)=DATA(K1)-TEMPR

```

```

DATA(K2+1)=DATA(K1+1)-TEMPI
DATA(K1)=DATA(K1)+TEMPR
340 DATA(K1+1)=DATA(K1+1)+TEMPI
350 MMAX=NON2
360 IF(MMAX-NP2HF)370,600,600
370 LMAX=MAX0(NON2T,MMAX/2)
IF(MMAX-NON2)405,380,380
380 THETA=-TWOPI*FLOAT(NON2)/FLOAT(4*MMAX)
IF(ISIGN)400,390,390
390 THETA=-THETA
400 WR=COS(THETA)
WI=SIN(THETA)
WSTPR=-2.*WI*WI
WSTPI=2.*WR*WI
405 DO 570 L=NON2,LMAX,NON2T
M=L
IF(MMAX-NON2)420,420,410
410 W2R=WR*WR-WI*WI
W2I=2.*WR*WI
W3R=W2R*WR-W2I*WI
W3I=W2R*WI+W2I*WR
420 DO 530 I1=1,I1RNG,2
DO 530 J3=I1,NON2,NP1
KMIN=J3+IPAR*M
IF(MMAX-NON2)430,430,440
430 KMIN=J3
440 KDIF=IPAR*MMAX
450 KSTEP=4*KDIF
DO 520 K1=KMIN,NTOT,KSTEP
K2=K1+KDIF
K3=K2+KDIF
K4=K3+KDIF
IF(MMAX-NON2)460,460,480
460 U1R=DATA(K1)+DATA(K2)
U1I=DATA(K1+1)+DATA(K2+1)
U2R=DATA(K3)+DATA(K4)
U2I=DATA(K3+1)+DATA(K4+1)

```

```

U3R=DATA(K1)-DATA(K2)
U3I=DATA(K1+1)-DATA(K2+1)
IF(ISIGN)470,475,475
470 U4R=DATA(K3+1)-DATA(K4+1)
U4I=DATA(K4)-DATA(K3)
GO TO 510
475 U4R=DATA(K4+1)-DATA(K3+1)
U4I=DATA(K3)-DATA(K4)
GO TO 510
480 T2R=W2R*DATA(K2)-W2I*DATA(K2+1)
T2I=W2R*DATA(K2+1)+W2I*DATA(K2)
T3R=WR*DATA(K3)-WI*DATA(K3+1)
T3I=WR*DATA(K3+1)+WI*DATA(K3)
T4R=W3R*DATA(K4)-W3I*DATA(K4+1)
T4I=W3R*DATA(K4+1)+W3I*DATA(K4)
U1R=DATA(K1)+T2R
U1I=DATA(K1+1)+T2I
U2R=T3R+T4R
U2I=T3I+T4I
U3R=DATA(K1)-T2R
U3I=DATA(K1+1)-T2I
IF(ISIGN)490,500,500
490 U4R=T3I-T4I
U4I=T4R-T3R
GO TO 510
500 U4R=T4I-T3I
U4I=T3R-T4R
510 DATA(K1)=U1R+U2R
DATA(K1+1)=U1I+U2I
DATA(K2)=U3R+U4R
DATA(K2+1)=U3I+U4I
DATA(K3)=U1R-U2R
DATA(K3+1)=U1I-U2I
DATA(K4)=U3R-U4R
520 DATA(K4+1)=U3I-U4I
KMIN=4*(KMIN-J3)+J3
KDIF=KSTEP

```

```

IF(KDIF-NP2)450,530,530
530 CONTINUE
M=MMAX-M
IF(ISIGN)540,550,550
540 TEMPR=WR
WR=-WI
WI=-TEMPR
GO TO 560
550 TEMPR=WR
WR=WI
WI=TEMPR
560 IF(M-LMAX)565,565,410
565 TEMPR=WR
WR=WR*WSTPR-WI*WSTPI+WR
570 WI=WI*WSTPR+TEMPR*WSTPI+WI
IPAR=3-IPAR
MMAX=MMAX+MMAX
GO TO 360

C
C MAIN LOOP FOR FACTORS NOT EQUAL TO TWO. APPLY THE
C TWIDDLE FACTOR
C  $W = \text{EXP}(\text{ISIGN} * 2 * \text{PI} * \text{SQRT}(-1) * (\text{J2}-1) * (\text{J1}-\text{J2}) / (\text{NP2} * \text{IFP1}))$ , THEN
C PERFORM A FOURIER TRANSFORM OF LENGTH IFACT(IF),
C MAKING USE OF CONJUGATE SYMMETRIES.
C
600 IF(NTWO-NP2)605,700,700
605 IFP1=NON2
IF=1
NP1HF=NP1/2
610 IFP2=IFP1/IFACT(IF)
J1RNG=NP2
IF(ICASE-3)612,611,612
611 J1RNG=(NP2+IFP1)/2
J2STP=NP2/IFACT(IF)
J1RG2=(J2STP+IFP2)/2
612 J2MIN=1+IFP2
IF(IFP1-NP2)615,640,640

```

```

615 DO 635 J2=J2MIN,IFP1,IFP2
      THETA=-TWOPI*FLOAT(J2-1)/FLOAT(NP2)
      IF(ISIGN)625,620,620
620 THETA=-THETA
625 SINTH=SIN(THETA/2.)
      WSTPR=-2.*SINTH*SINTH
      WSTPI=SIN(THETA)
      WR=WSTPR+1.
      WI=WSTPI
      J1MIN=J2+IFP1
      DO 635 J1=J1MIN,J1RNG,IFP1
      I1MAX=J1+I1RNG-2
      DO 630 I1=J1,I1MAX,2
      DO 630 I3=I1,NTOT,NP2
      J3MAX=I3+IFP2-NP1
      DO 630 J3=I3,J3MAX,NP1
      TEMPR=DATA(J3)
      DATA(J3)=DATA(J3)*WR-DATA(J3+1)*WI
630 DATA(J3+1)=TEMPR*WI+DATA(J3+1)*WR
      TEMPR=WR
      WR=WR*WSTPR-WI*WSTPI+WR
635 WI=TEMPR*WSTPI+WI*WSTPR+WI
640 THETA=-TWOPI/FLOAT(IFACT(IF))
      IF(ISIGN)650,645,645
645 THETA=-THETA
650 SINTH=SIN(THETA/2.)
      WSTPR=-2.*SINTH*SINTH
      WSTPI=SIN(THETA)
      KSTEP=2*N/IFACT(IF)
      KRANG=KSTEP*(IFACT(IF)/2)+1
      DO 698 I1=1,I1RNG,2
      DO 698 I3=I1,NTOT,NP2
      DO 690 KMIN=1,KRANG,KSTEP
      J1MAX=I3+J1RNG-IFP1
      DO 680 J1=I3,J1MAX,IFP1
      J3MAX=J1+IFP2-NP1
      DO 680 J3=J1,J3MAX,NP1

```

```

J2MAX=J3+IFP1-IFP2
K=KMIN+(J3-J1+(J1-I3)/IFACT(IF))/NP1HF
IF(KMIN-1)655,655,665
655  SUMR=0.
    SUMI=0.
    DO 660 J2=J3,J2MAX,IFP2
    SUMR=SUMR+DATA(J2)
660  SUMI=SUMI+DATA(J2+1)
    WORK(K)=SUMR
    WORK(K+1)=SUMI
    GO TO 680
665  KCONJ=K+2*(N-KMIN+1)
    J2=J2MAX
    SUMR=DATA(J2)
    SUMI=DATA(J2+1)
    OLDSR=0.
    OLDSI=0.
    J2=J2-IFP2
670  TEMPR=SUMR
    TEMPI=SUMI
    SUMR=TWOWR*SUMR-OLDSR+DATA(J2)
    SUMI=TWOWR*SUMI-OLDSI+DATA(J2+1)
    OLDSR=TEMPR
    OLDSI=TEMPI
    J2=J2-IFP2
    IF(J2-J3)675,675,670
675  TEMPR=WR*SUMR-OLDSR+DATA(J2)
    TEMPI=WI*SUMI
    WORK(K)=TEMPR-TEMPI
    WORK(KCONJ)=TEMPR+TEMPI
    TEMPR=WR*SUMI-OLDSI+DATA(J2+1)
    TEMPI=WI*SUMR
    WORK(K+1)=TEMPR+TEMPI
    WORK(KCONJ+1)=TEMPR-TEMPI
680  CONTINUE
    IF(KMIN-1)685,685,686
685  WR=WSTPR+1.

```

```

WI=WSTPI
GO TO 690
686 TEMPR=WR
WR=WR*WSTPR-WI*WSTPI+WR
WI=TEMPR*WSTPI+WI*WSTPR+WI
690 TWOWR=WR+WR
IF(ICASE-3)692,691,692
691 IF(IFP1-NP2)695,692,692
692 K=1
I2MAX=I3+NP2-NP1
DO 693 I2=I3,I2MAX,NP1
DATA(I2)=WORK(K)
DATA(I2+1)=WORK(K+1)
693 K=K+2
GO TO 698
C
C COMPLETE A REAL TRANSFORM IN THE 1ST DIMENSION, N
C ODD, BY CONJUGATE SYMMETRIES AT EACH STAGE.
C
695 J3MAX=I3+IFP2-NP1
DO 697 J3=I3,J3MAX,NP1
J2MAX=J3+NP2-J2STP
DO 697 J2=J3,J2MAX,J2STP
J1MAX=J2+J1RG2-IFP2
J1CNJ=J3+J2MAX+J2STP-J2
DO 697 J1=J2,J1MAX,IFP2
K=1+J1-I3
DATA(J1)=WORK(K)
DATA(J1+1)=WORK(K+1)
IF(J1-J2)697,697,696
696 DATA(J1CNJ)=WORK(K)
DATA(J1CNJ+1)=-WORK(K+1)
697 J1CNJ=J1CNJ-IFP2
698 CONTINUE
IF=IF+1
IFP1=IFP2
IF(IFP1-NP1)700,700,610

```

```

C
C   COMPLETE A REAL TRANSFORM IN THE 1ST DIMENSION, N
C   EVEN, BY CONJUGATE SYMMETRIES.
C
700  GO TO (900,800,900,701),ICASE
701  NHALF=N
      N=N+N
      THETA=-TWOPI/FLOAT(N)
      IF(ISIGN)703,702,702
702  THETA=-THETA
703  SINTH=SIN(THETA/2.)
      WSTPR=-2.*SINTH*SINTH
      WSTPI=SIN(THETA)
      WR=WSTPR+1.
      WI=WSTPI
      IMIN=3
      JMIN=2*NHALF-1
      GO TO 725
710  J=JMIN
      DO 720 I=IMIN,NTOT,NP2
      SUMR=(DATA(I)+DATA(J))/2.
      SUMI=(DATA(I+1)+DATA(J+1))/2.
      DIFR=(DATA(I)-DATA(J))/2.
      DIFI=(DATA(I+1)-DATA(J+1))/2.
      TEMPR=WR*SUMI+WI*DIFR
      TEMPI=WI*SUMI-WR*DIFR
      DATA(I)=SUMR+TEMPR
      DATA(I+1)=DIFI+TEMPI
      DATA(J)=SUMR-TEMPR
      DATA(J+1)=-DIFI+TEMPI
720  J=J+NP2
      IMIN=IMIN+2
      JMIN=JMIN-2
      TEMPR=WR
      WR=WR*WSTPR-WI*WSTPI+WR
      WI=TEMPR*WSTPI+WI*WSTPR+WI
725  IF(IMIN-JMIN)710,730,740

```

```

730 IF(ISIGN)731,740,740
731 DO 735 I=IMIN,NTOT,NP2
735 DATA(I+1)=-DATA(I+1)
740 NP2=NP2+NP2
      NTOT=NTOT+NTOT
      J=NTOT+1
      IMAX=NTOT/2+1
745 IMIN=IMAX-2*NHALF
      I=IMIN
      GO TO 755
750 DATA(J)=DATA(I)
      DATA(J+1)=-DATA(I+1)
755 I=I+2
      J=J-2
      IF(I-IMAX)750,760,760
760 DATA(J)=DATA(IMIN)-DATA(IMIN+1)
      DATA(J+1)=0.
      IF(I-J)770,780,780
765 DATA(J)=DATA(I)
      DATA(J+1)=DATA(I+1)
770 I=I-2
      J=J-2
      IF(I-IMIN)775,775,765
775 DATA(J)=DATA(IMIN)+DATA(IMIN+1)
      DATA(J+1)=0.
      IMAX=IMIN
      GO TO 745
780 DATA(1)=DATA(1)+DATA(2)
      DATA(2)=0.
      GO TO 900

C
C COMPLETE A REAL TRANSFORM FOR THE 2ND OR 3RD
C DIMENSION BY CONJUGATE SYMMETRIES.
C
800 IF(I1RNG-NP1)805,900,900
805 DO 860 I3=1,NTOT,NP2
      I2MAX=I3+NP2-NP1

```

```
DO 860 I2=I3,I2MAX,NP1
IMIN=I2+I1RNG
IMAX=I2+NP1-2
JMAX=2*I3+NP1-IMIN
IF(I2-I3)820,820,810
810 JMAX=JMAX+NP2
820 IF(IDIM-2)850,850,830
830 J=JMAX+NP0
DO 840 I=IMIN,IMAX,2
DATA(I)=DATA(J)
DATA(I+1)=-DATA(J+1)
840 J=J-2
850 J=JMAX
DO 860 I=IMIN,IMAX,NP0
DATA(I)=DATA(J)
DATA(I+1)=-DATA(J+1)
860 J=J-NP0
C
C END OF LOOP ON EACH DIMENSION
C
900 NP0=NP1
NP1=NP2
910 NPREV=N
920 RETURN
END
```

APPENDIX D

Plotting Functions

Introduction

The capabilities of the S system can be expanded by the user, by writing functions in the S language as follows.

`<-` is the assignment operation in S.

```
line.plot<- function(file,label="Example of line.plot")
{
  data<-scan(file)
  plot(data,type='l')
  title(main=label)
}
```

This example reads a vector of data from a Unix file and draws the plot with lines connecting each of the data points. The function may be used by issuing the command:

```
line.plot("test.data")
```

or

```
line.plot("well.FF10","Gamma Ray well FF10")
```

The following are some functions which are written in the S language to plot the results of program PCAXCOR.

xsection

#Plot a cross-section given absolute formation depths.

```
xsection<-function(a = "xsec.in", d = "numoffms", e = 8)
```

#a=datafile containing the information about the boreholes used in the
#analysis.

#d=number of the formations of the sequence.

#e=number of columns of the datafile matrix.

#read the data from the datafile, into a matrix called data:

```
data <- matrix(scan(a, skip = 1), ncol = e, byrow = T)
```

#structure of matrix is as follows: first two columns are Longitude &
#Latitude, respectively. 3rd column is data for the first formation, 4th
#column, is second formation....etc. Last column is the height of the well-
#head above sea level (KB).

So the number of formation is equal to length of one row minus three:

```
nof <- len(data[, 1]) - 3
```

#And number of wells is equal to the length of a column:

```
now <- len(data[, 1])
```

```
pal <- data[, 1]      # This just sets up a vector of the correct lengths.
```

```
pal[1] <- 0          # Set the first value to zero
```

#using Pythagorus theory to calculate positions a long the line of each of the
#wells:

```
for(i in 2:now) {
```

```
  xdist <- data[i, 1] - data[i - 1, 1]
```

```
    ydist <- data[i, 2] - data[i - 1, 2]
```

```
    x <- (xdist) * (xdist)
```

```
    y <- (ydist) * (ydist)
```

#re-assign each value in the pal vector, (position is cumulative)

```
  pal[i] <- pal[i - 1] + sqrt(x + y)
```

```
}
```

```
par(mar = c(9, 4.1, 4.1, 4.1))
```

```
par(adj = .5)
```

```
plot(pal, -data[, d + 2], type = "b", xlim = c(0, 5500), ylim = c(
```

```
-14000, 500), xlab = " HORIZONTAL DISTANCE NOT TO SCALE",
ylab=
```

```
"DEPTH IN FEET", lty = 1, axes = F)
```

```
box()
```

```
axis(2)
```

```
for(i in 3:(d + 1))
```

```
lines(pal, -data[, i], lty = 1, col = 1)
```

```
· for(i in 1:now)
```

```
segments(pal[i], -data[i, 3], pal[i], -data[i, d + 2], lty = 2)
```

```
text(5450, -4300, "Etel Fm", cex = .8)
```

```
text(5450, -5200, "Sheghega Fm", cex = .8)
```

```
text(5450, -6800, "Domran Fm", cex = .8)
```

```
text(5450, -7500, "Ruaga Fm", cex = .8)
```

```
text(5450, -9000, "Heira Fm", cex = .8)
```

```
text(5450, -10200, "Zmam Fm", cex = .8)
```

```
text(5450, -11300, "Socna Fm", cex = .8)
```

```
text(5450, -11750, "Bahi Fm", cex = .8)
```

```
text(5450, -12750, "Gargaf Fm", cex = .8)
```

#Plot the names of the wells :

```

text(13, 300, "FF2 ", cex = .8)

text(414, 300, "FF5 ", cex = .8)

text(1045, 300, "FF6 ", cex = .8)

text(1705, 300, "FF7", cex = .8)

text(1906, 300, "FF8", cex = .8)

text(2551, 300, "FF9", cex = .8)

text(2881, 300, "FF10", cex = .8)

text(3656, 300, "FF11", cex = .8)

text(3957, 300, "FF12", cex = .8)

text(4330, 300, "FF13", cex = .8)

text(4760, 300, "FF14", cex = .8)

text(5148, 300, "FF15", cex = .8)

}

```

xsectioncor function

#Plot the result of the cross-correlation tops of the program PCAXCOR and
#draw a cross-section between the studied boreholes.

```
xsectioncor<-function(a = xsec.cor, d = numoffms, e = 8)
```

#a=datafile containing the information about the boreholes used in the
#analysis.

#d=number of the formations of the sequence.

#e=number of columns of the datafile matrix.

{

#read the data from the datafile, into a matrix called data:

```
data <- matrix(scan(a, skip = 1), ncol = e, byrow = T)
```

#structure of matrix is as follows: first two columns are Longitude&
#Latitude respectively. 3rd column is data for the first formation, 4th
#column, is second formation....etc. Last column is the height of the well-
#head above sea level (KB).

So the number of formation is equal to length of one row minus three:

```
nof <- len(data[1, ]) - 3
```

#And number of wells is equal to the length of a column:

```
now <- len(data[, 1])
```

```
pal <- data[, 1]      # This just sets up a vector of the correct lengths.
```

```
pal[1] <- 0          # Set the first value to zero
```

#using Pythagorus theory to calculate positions a long the line of each of the
#wells:

```
for(i in 2:now) {
```

```
xdist <- data[i, 1] - data[i - 1, 1]
```

```
ydist <- data[i, 2] - data[i - 1, 2]
```

```
x <- (xdist) * (xdist)
```

```
y <- (ydist) * (ydist)
```

#re-assign each value in the T.pal vector, (position is cumulative)

```
pal[i] <- pal[i - 1] + sqrt(x + y)
```

```
par(mar = c(9, 4.1, 4.1, 4.1))
```

```
par(adj = .5)
```

```
plot(pal, -data[, d + 2], type = "b", xlim = c(0, 1400), ylim = c(
-10300, 700), xlab = " DISTANCE IN KM", ylab = "DEPTH IN
FEET", lty = 1, axes = F)
```

```
post <- c(seq(0, 1400, 200))
```

```
axis(1, at = post, labels = F)
```

```
txt <- c("0", "2", "4", "6", "8", "10", "12", "14")
```

```
mtext(txt, at = post, side = 1, line = 1)
```

```
axis(2)
```

```
box()
```

```
for(i in 3:(d + 1))
```

```

lines(pal, -data[, i], lty = 1, col = 1)

for(i in 1:now)

segments(pal[i], -data[i, 3], pal[i], -data[i, d + 2], lty = 2)

text(1300, -4435, "Etel Fm", cex = .8)

text(1300, -5825, "Sheghega Fm", cex = .8)

text(1300, -7076, "Domran Fm", cex = .8)

text(1300, -7762, "Ruaga Fm", cex = .8)

text(1300, -8900, "Heira Fm", cex = .8)

text(3, 550, "FF7", cex = .8)

text(215, 550, "FF13", cex = .8)

text(325, 550, "FF11", cex = .8)

text(644, 550, "FF12", cex = .8)

text(1244, 550, "FF10", cex = .8)

}

```

pws

```

#Plot the original first principal component, their derivatives, power
#spectra, the logarithmic scaled power spectra, the interpolated power
#spectra, the cross-correlation function for stretch and the cross-correlation
#function for displacement.

```

```
pws<-function(a = "c1", b = "c2", d = "c3", e = "c4", f = "c5", g = "c6", h
= "c7")
```

#a=datafile containing the first principal component of the formation to be
#correlated and the first principal component of the sequence to be
#correlated with.

#b=derivative data of the first formation and the long sequence.

#d=the power spectra of the data.

#e=logarithmic scaled power spectra of the data.

#f=the interpolated power spectra.

#g=the cross-correlation function of power spectra for stretch.

#h=the cross-correlation function for displacement.

```
{
a <- readf(a, 1)
print(length(a))
ln1 <- len(a)
b <- readf(b, 1)
d <- readf(d, 1)
e <- readf(e, 1)
```

```
f <- readf(f, 1)
```

```
g <- readf(g, 1)
```

```
h <- readf(h, 1)
```

```
par(mar = c(5, 4.1, 4.1, 4.1, 4.1))
```

```
par(mfrow = c(3, 4))
```

```
#Plot the original data (non-filtered or filtered principal components or  
#original well-log data)
```

```
print("ENTER THICKNESS OF FORMATION 1", quote = F)
```

```
d1 <- read(length = 1, print = F)
```

```
print("ENTER THICKNESS OF FORMATION 2", quote = F)
```

```
d2 <- read(length = 1, print = F)
```

```
dd2 <- (d2 + d2)/2 + 1
```

```
print("ENTER DEPTH OF FORMATION 1")
```

```
df1 <- read(length = 1, print = F)
```

```
print("ENTER DEPTH OF FORMATION 2")
```

```
df2 <- read(length = 1, print = F)
```

```
dd1 <- d1 * 2
```

```
dd1 <- dd1/2 + 1
```

```
plot(a[1:dd1], -df1:(-df1 - d1), type = "l", xlab = "", ylab =
```

```
"DEPTH IN FEET", axes = F, ylim = c(-df1 - dd2, -df1))
```

```
title(main = "PC-I, SHORT SEQ. (FF13)")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

```
plot(a[d1:ln1], -df2:(-df2 - d2), type = "l", xlab = "", ylab = "DEPTH IN
```

```
FEET", axes = F)
```

```
title(main = "PC-I, LONG SEQ. (FF11)")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

```
#Plot the derivative data.
```

```
d3 <- len(b)
```

```
d3 <- d3/2
```

```
plot(b[1:d3], -1:-d3, type = "l", xlab = "", ylab = "SHORT SERIES", axes =  
F)
```

```
title(main = "DERIVATIVE DATA")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

```
dd3 <- d3 * 2
```

```
ddd3 <- dd3/2 + 1

plot(b[d3:dd3], -1:-ddd3, type = "l", xlab = "", ylab = "LONG
SERIES", axes = F)

title(main = "DERIVATIVE DATA")

axis(2, col = 1)

axis(3, col = 1)

#Plot power spectra.

d4 <- len(d)

d4 <- d4/2

plot(d[1:d4], -1:-d4, type = "l", xlab = "", ylab = "FREQ., SHORT
SERIES",

axes = F)

title(main = "POWER SP")

axis(2, col = 1)

axis(3, col = 1)

dd4 <- d4 * 2

ddd4 <- dd4/2 + 1

plot(d[d4:dd4], -1:-ddd4, type = "l", xlab = "", ylab = "FREQ., LONG
```

```

SERIES", axes = F)

title(main = "POWER SP")

axis(2, col = 1)

axis(3, col = 1)

#Plot logarithmic frequencies.

d7 <- len(d)

d7 <- d7/2

xx <- min(g[1:d7])

yy <- max(g[1:d7])

yy1 <- yy + (yy/5)

plot(d[1:d7], -g[1:d7], type = "l", xlab = "", ylab = "LOG10 FREQ., SHORT
SERIES", axes = F, ylim = c(-yy1, -xx))

title(main = "POWER SP")

axis(2, col = 1)

axis(3, col = 1)

dd7 <- d7 * 2

ddd7 <- dd7/2 + 1

xx <- min(g[1:d7])

```

```
yy <- max(g[1:d7])
```

```
yy1 <- yy + (yy/5)
```

```
plot(d[ddd7:dd7], -g[1:d7], type = "l", xlab = "", ylab = "LOG10 FREQ.,  
LONG SERIES", axes = F, ylim = c(-yy1, -xx))
```

```
title(main = "POWER SP")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

#plot the interpolated frequencies.

```
d5 <- len(e)
```

```
d5 <- d5/2
```

```
plot(e[1:d5], -1:-d5, type = "l", xlab = "", ylab =
```

```
"INTERP. FREQ., SHORT SERIES", axes = F)
```

```
title(main = "INTERP.POWER SP")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

```
dd5 <- d5 * 2
```

```
ddd5 <- dd5/2 + 1
```

```
plot(e[d5:dd5], -1:-ddd5, type = "l", xlab = "", ylab = "INTERP. FREQ.,
```

```
LONG SERIES", axes = F)
```

```
title(main = "INTERP.POWER SP")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

```
#Plot cross-correlation function for stretch.
```

```
d6 <- len(f)
```

```
d6 <- d6/2
```

```
dd6 <- d6 * 2
```

```
ddd6 <- dd6/2 + 1
```

```
plot(f[ddd6:dd6], f[1:d6], type = "l", xlab = "", ylab =
```

```
"LAG FOR STRETCH", axes = F)
```

```
post <- c(seq(-30, 30, 10))
```

```
axis(2, at = post, labels = F)
```

```
txt <- c("-30", "-20", "-10", "0", "10", "20", "30")
```

```
mtext(txt, at = post * -1, side = 2, line = 2)
```

```
title(main = "X-COR (STRETCH)")
```

```
axis(3, col = 1)
```

```
#Plot cross-correlation function for displacement.
```

```

d8 <- len(h)

plot(h[1:d8], -1:-d8, type = "l", xlab = "", ylab = " LAG FOR
DISPLACEMENT", axes = F)

title(main = "X-COR (DISPL.)")

axis(2, col = 1)

axis(3, col = 1)

}

```

pws1

#Plot the original principal component, the derivative data, the
#interpolated power spectra, the cross-correlation function for stretch, and
#the cross-correlation function of the displacement.

#

```

pws1<-function(a = "c1", b = "c2", d = "c3", e = "c4", f = "c5", g = "c6", h
= "c7")

```

#

#a=datafile containing the first principal component of the formation to be
#correlated and the first principal component of the sequence to be
#correlated with.

#b=derivative data of the first formation and the long sequence.

#d=the power spectra of the data.

#e=logarithmic scaled power spectra of the data.

#f=the interpolated power spectra.

#g=the cross-correlation function of power spectra for stretch.

#h=the cross-correlation function for displacement.

#

{

a <- readf(a, 1)

ln1 <- len(a)

print(length(a))

b <- readf(b, 1)

d <- readf(d, 1)

e <- readf(e, 1)

f <- readf(f, 1)

g <- readf(g, 1)

h <- readf(h, 1)

par(mfrow = c(2, 4))

#Plot the original data (non-filtered or filtered principal components or the
#original well-log variables).

```
print("ENTER THICKNESS OF FORMATION 1", quote = F)
```

```
d1 <- read(length = 1, print = F)
```

```
print("ENTER THICKNESS OF FORMATION 2", quote = F)
```

```
d2 <- read(length = 1, print = F)
```

```
dd2 <- (d1 + d2)/2 + 1
```

```
print("ENTER DEPTH OF FORMATION 1")
```

```
df1 <- read(length = 1, print = F)
```

```
print("ENTER DEPTH OF FORMATION 2")
```

```
df2 <- read(length = 1, print = F)
```

```
dd1 <- d1 * 2
```

```
dd1 <- dd1/2 + 1
```

```
plot(a[1:dd1], -df1:(-df1 - d1), type = "l", xlab = "", ylab =
```

```
"DEPTH IN FEET", axes = F, ylim = c(-df1 - dd2, -df1))
```

```
title(main = "PC-I, SHORT SEQ. (FF13)")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

```
plot(a[d1:ln1], -df2:(-df2 - d2), type = "l", xlab = "", ylab = "DEPTH IN
```

```
FEET", axes = F)

title(main = "PC-I, LONG SEQ. (FF13)")

axis(2, col = 1)

axis(3, col = 1)

# Plot the derivative data.

d3 <- len(b)

d3 <- d3/2

plot(b[1:d3], -1:-d3, type = "l", xlab = "", ylab = "SHORT
SERIES", axes = F)

title(main = "DERIVATIVE DATA")

axis(2, col = 1)

axis(3, col = 1)

dd3 <- d3 * 2

ddd3 <- dd3/2 + 1

plot(b[d3:dd3], -1:-ddd3, type = "l", xlab = "", ylab = "LONG
SERIES", axes = F)

title(main = "DERIVATIVE DATA")

axis(2, col = 1)

axis(3, col = 1)
```

#Plot the interpolated frequencies.

```
d5 <- len(e)
```

```
d5 <- d5/2
```

```
plot(e[1:d5], -1:-d5, type = "l", xlab = "", ylab = "INTERP. FREQ.,  
SHORT SERIES", axes = F)
```

```
title(main = "INTERP.POWER SP")
```

```
axis(2, col = 1)
```

```
, axis(3, col = 1)
```

```
dd5 <- d5 * 2
```

```
ddd5 <- dd5/2 + 1
```

```
plot(e[d5:dd5], -1:-ddd5, type = "l", xlab = "", ylab = "INTERP. FREQ.,  
LONG SERIES", axes = F)
```

```
title(main = "INTERP.POWER SP")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

#Plot the cross-correlation function for stretch.

```
d6 <- len(f)
```

```
d6 <- d6/2
```

```
dd6 <- d6 * 2
```

```
ddd6 <- dd6/2 + 1
```

```
plot(f[ddd6:dd6], f[1:d6], type = "l", xlab = "", ylab = "LAG FOR  
STRETCH", axes = F)
```

```
post <- c(seq(-30, 30, 10))
```

```
axis(2, at = post, labels = F)
```

```
txt <- c("-30", "-20", "-10", "0", "10", "20", "30")
```

```
  mtext(txt, at = post * -1, side = 2, line = 2)
```

```
  title(main = "X-COR (STRETCH)")
```

```
  axis(3, col = 1)
```

```
# Plot the cross-correlation function for displacement.
```

```
d8 <- len(h)
```

```
plot(h[1:d8], -1:-d8, type = "l", xlab = "", ylab = "LAG FOR  
DISPLACEMENT", axes = F)
```

```
title(main = "X-COR (DISPL.)")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

```
}
```

xcfun

#A function to plot the original variables or principal components (filtered
#or non- filtered), and the two cross-correlation function (for stretch & for
#displacement).

```
xcfun<-function(a = "l18", b = "pp8", d = "c5", e = "c7")
```

```
#
```

```
#a=datafile containing the first principal component of borehole sequence.
```

```
#b=datafile containing the first principal component of another borehole  
#sequence.
```

```
#d=datafile containing the cross-correlation for stretch.
```

```
#e=datafile containing the cross-correlation for displacement.
```

```
#
```

```
{
```

```
  a <- readf(a, 1)
```

```
  b <- readf(b, 1)
```

```
  d <- readf(d, 1)
```

```
  e <- readf(e, 1)
```

```
  par(mar = c(5, 4.1, 4.1, 4.1))
```

```
  par(fig = c(0, .36, 0, 1))
```

```

plot(a, -dep, type = "l", xlab = "", ylab = "DEPTH IN FEET", axes =F,
ylim=

c(-12000, -2000), xlim = c(-4, 4))

title(main = "PC-I, SHORT SEQ. (FF13)")

axis(2)

axis(3)

print("ENTER X-AXIS FOR WINDOW")

xax <- read(length = 1, print = F)

print("ENTER Y-AXIS [TOP] FOR WINDOW")

ytp <- read(length = 1, print = F)

print("ENTER Y-AXIS [BOTTOM] FOR WINDOW")

ybt <- read(length = 1, print = F)

text(xax, ytp, "-")

segments(xax, ytp, xax, ybt)

text(xax, ybt, "-")

par(fig = c(.35, .7, 0, 1))

plot(b, -dep, type = "l", xlab = "", ylab = "DEPTH IN FEET", axes
=F,ylim = c(-12000, -2000), xlim = c(-4, 4))

title(main = "PC-I, LONG SEQ. (FF11)")

```

```
axis(2)
```

```
axis(3)
```

```
print("ENTER X-AXIS FOR WINDOW")
```

```
xax <- read(length = 1, print = F)
```

```
print("ENTER Y-AXIS [TOP] FOR WINDOW")
```

```
ytp <- read(length = 1, print = F)
```

```
print("ENTER Y-AXIS [BOTTOM] FOR WINDOW")
```

```
ybt <- read(length = 1, print = F)
```

```
text(xax, ytp, "-")
```

```
segments(xax, ytp, xax, ybt)
```

```
text(xax, ybt, "-")
```

```
# Plot cross-correlation function for displacement.
```

```
par(fig = c(.68, 1, 0, .6))
```

```
d8 <- len(e)
```

```
xx <- min(1:d8)
```

```
yy <- max(1:d8)
```

```
yy1 <- yy + (yy/2)
```

```
plot(e[1:d8], -1:-d8, type = "l", xlab = "", ylab = "LAG FOR DISPL.", axes
= F, ylim = c(-yy1, -xx))
```

```
title(main = "X-COR (DISPL.)")
```

```
axis(2)
```

```
axis(3)
```

```
print("enter X-axis of X-COR function for displacement", quote = F)
```

```
xax <- read(length = 1, print = F)
```

```
print("enter Y-axis of X-COR function for displacement", quote = F)
```

```
yax <- read(length = 1, print = F)
```

```
text(xax, yax, "<-")
```

```
# Plot cross-correlation function for displacement.
```

```
par(fig = c(.68, 1, .5, 1))
```

```
plot(d[ddd6:dd6], d[1:d6], type = "l", xlab = "", ylab = "LAG FOR
```

```
STRETCH", axes = F)
```

```
title(main = "X-COR (STRETCH)")
```

```
axis(3)
```

```
par(xpd = F)
```

```
post <- c(seq(-30, 30, 10))
```

```

axis(2, at = post, labels = F)

txt <- c("-30", "-20", "-10", "0", "10", "20", "30")

mtext(txt, at = post * -1, side = 2, line = 2)

print("enter X-axis of X-COR function for stretch", quote = F)

xax <- read(length = 1, print = F)

print("enter Y-axis of X-COR function for stretch", quote = F)

yax <- read(length = 1, print = F)

text(xax, yax, "<-")

}

```

macbound

#Plot on one side the first principal components, and on the other side the
#boundaries of different formations.

```

macbound<-function(a = "ll6", b = "ff13.30", d = "dep.30", e =
"formfile")

```

#

#a=datafile containing the non-filtered first principal component of a
certain #borehole.

#b=datafile containing D^2 values (boundaries of the related borehole.

#d=datafile containing the depth of the sequence.

#e=datafile containing formation tops.

#

{

a <- readf(a, 1)

b <- readf(b, 1)

d <- readf(d, 1)

par(fig = c(0, .6, 0, 1))

plot(a, -dep, type = "l", xlab = "(a)", ylab = "DEPTH IN FEET", axes = F,
ylim = c(-12000, -2000))

axis(2, col = 1)

axis(3, col = 1)

title(main = "NON-FIL. PC-I (WELL FF13)")

par(fig = c(.55, 1, 0, 1))

plot(b, -dep.30, type = "l", ylab = "DEPTH IN FEET", xlab = "(b)", axes =
F, ylim = c(-12000, -2000))

axis(2, col = 1)

axis(3, col = 1)

title(main = "D SQUARE")

```

fdep <- readf(e, 2)

text(fdep[1, 1], fdep[1, 2], "<----- Etel Fm", cex = .8)

text(fdep[2, 1], fdep[2, 2], "<----- Sheghega Fm", cex = .8)

text(fdep[3, 1], fdep[3, 2], "<----- Domran Fm", cex = .8)

text(fdep[4, 1], fdep[4, 2], "<----- Ruaga Fm", cex = .8)

text(fdep[5, 1], fdep[5, 2], "<----- Heira Fm", cex = .8)

text(fdep[6, 1], fdep[6, 2], "<----- Zmam Fm", cex = .8)

}

```

smoothplot

#Plot the filtered first principal components of two different wells and
#different windows used in the cross-correlation process.

```
smoothplot<-function(a = "l18", b = "pp8")
```

```

#

#a=datafile of the first principal component of a borehole.

#b=datafile of the first principal component of another borehole.

#

{

```

```
a <- readf(a, 1)
```

```
b <- readf(b, 1)
```

```
par(mar = c(5, 4.1, 4.1, 4.1))
```

```
par(mfrow = c(1, 2))
```

```
plot(a, -dep, type = "l", xlab = "(A)", ylab = "DEPTH IN FEET", axes = F,
ylim = c(-12000, -2000), xlim = c(-4, 3))
```

```
title(main = "FIL. PC-I (FF13)")
```

```
axis(2, col = 1)
```

```
axis(3, col = 1)
```

#This routine read the text input files which contain:

#column 1 the upper range of y-axis of the window.

#column 2 the lower limit of y-axis of the window.

#column 3 the x-axis of the window.

#file "textfile" contains the data to plot windows on the first plot.

```
dat <- readf("textfile1", 3)
```

```
text(dat[1, 3], dat[1, 1], "-")
```

```
segments(dat[1, 3], (dat[1, 1] - 1), dat[1, 3], (dat[1, 2] - 1), lty = 1)
```

```
text(dat[1, 3], dat[1, 2], "-")
```

```
dev1 <- (dat[1, 2] + dat[1, 1])/2  
  
text(dat[1, 3] + .2, dev1, "1")  
  
text(dat[2, 3], dat[2, 1], "-")  
  
segments(dat[2, 3], dat[2, 1], dat[2, 3], dat[2, 2], lty = 1)  
  
text(dat[2, 3], dat[2, 2], "-")  
  
dev2 <- (dat[2, 2] + dat[2, 1])/2  
  
text(dat[2, 3] + .2, dev2, "2")  
  
text(dat[3, 3], dat[3, 1], "-")  
  
segments(dat[3, 3], dat[3, 1], dat[3, 3], dat[3, 2], lty = 1)  
  
text(dat[3, 3], dat[3, 2], "-")  
  
dev3 <- (dat[3, 2] + dat[3, 1])/2  
  
text(dat[3, 3] + .2, dev3, "3")  
  
text(dat[4, 3], dat[4, 1], "-")  
  
segments(dat[4, 3], dat[4, 1], dat[4, 3], dat[4, 2])  
  
text(dat[4, 3], dat[4, 2], "-")  
  
dev4 <- (dat[4, 2] + dat[4, 1])/2  
  
text(dat[4, 3] + .2, dev4, "4")  
  
text(dat[5, 3], dat[5, 1], "-")
```

```

segments(dat[5, 3], dat[5, 1], dat[5, 3], dat[5, 2])

text(dat[5, 3], dat[5, 2], "-")

dev5 <- (dat[5, 2] + dat[5, 1])/2

text(dat[5, 3] + .2, dev5, "5")

plot(b, -dep, type = "l", xlab = "(B)", ylab = "DEPTH IN FEET", axes= F,
ylim = c(-12000, -2000), xlim = c(-4, 4))

axis(2, col = 1)

axis(3, col = 1)

title(main = "FIL. PC-I (FF11)")

```

#file "textfile2" contains data to plot windows on the second plot.

```

dat <- readf("textfile2", 3)

text(dat[1, 3], dat[1, 1], "-")

segments(dat[1, 3], dat[1, 1], dat[1, 3], dat[1, 2], lty = 1)

text(dat[1, 3], dat[1, 2], "-")

dev1 <- (dat[1, 2] + dat[1, 1])/2

text(dat[1, 3] + .2, dev1, "A")

text(dat[2, 3], dat[2, 1], "-")

segments(dat[2, 3], dat[2, 1], dat[2, 3], dat[2, 2], lty = 1)

```

```
text(dat[2, 3], dat[2, 2], "-")

dev2 <- (dat[2, 2] + dat[2, 1])/2

text(dat[2, 3] + .2, dev2, "B")

text(dat[3, 3], dat[3, 1], "-")

segments(dat[3, 3], dat[3, 1], dat[3, 3], dat[3, 2], lty = 1)

text(dat[3, 3], dat[3, 2], "-")

dev3 <- (dat[3, 2] + dat[3, 1])/2

text(dat[3, 3] + .2, dev3, "C")

text(dat[4, 3], dat[4, 1], "-")

segments(dat[4, 3], dat[4, 1], dat[4, 3], dat[4, 2], lty = 1)

text(dat[4, 3], dat[4, 2], "-")

dev4 <- (dat[4, 2] + dat[4, 1])/2

text(dat[4, 3] + .2, dev4, "D")

text(dat[5, 3], dat[5, 1], "-")

segments(dat[5, 3], dat[5, 1], dat[5, 3], dat[5, 2])

text(dat[5, 3], dat[5, 2], "-")

dev5 <- (dat[5, 2] + dat[5, 1])/2

text(dat[5, 3] + .2, dev5, "E")
```

}

readf

#To read different input files to the S system from the Sun workstation.

```
readf<-function(a = m1, b = 3)
```

#

#a=datafile containing any matrix to be read to the S system from a Unix
#file.

#b=number of the columns of that matrix.

#

{

```
file <- matrix(scan(a), ncol = b, byrow = T)
```

#

}

post

#Set up the graphic mode for the Laser printer.

```
post<-function(b = 9.5, d = 6, e = 8, f = 3)
```

#

#b=the width of the plot.

#d=the length of the plot.

#e=point size to be used.

#f=the font type.

{

postscript(hor = T, width = b, height = d, pointsize = e, font = f)

}

Appendix E

1

CROSS-CORRELATION USING PRINCIPAL COMPONENTS OF MODEL DATA IN FF13-6

IS
 LS= 50 LL= 300 IDER= 1 IORG= 1 SMAX= 2.0
 DEPTH OF SHORT SERIES = 50.0 FEET
 DEPTH OF LONG SERIES = 1.0 FEET

INPUT DATA

SHORT SERIES LONG SERIES

1	-0.158	0.000
2	-0.240	-1.125
3	-0.444	-0.194
4	-0.404	0.062
5	-0.187	0.131
50	-0.434	0.172
.	.	.
.	.	.
.	.	.
294		0.246
295		0.188
296		0.030
297		0.058
298		0.059
299		0.010

FOURIER TRANSFORM

SERIES 1

SERIES 2

	REAL	IMAGINARY	POWER SPECTRUM		REAL	IMAGINARY	POWER SPECTRUM
1	-109.803	23.329	42.143		1.963	-2.245	0.030
2	-5.417	31.510	3.419		1.428	-0.687	0.008
3	-4.885	14.159	0.750		1.522	-1.624	0.017
4	2.960	41.209	5.709		-2.197	-1.540	0.024
5	37.902	-63.728	18.388		6.629	3.411	0.186
6	-3.412	-8.095	0.258		1.095	0.070	0.004
7	4.255	49.166	8.145		-6.171	1.300	0.133
8	24.855	-6.670	2.215		2.970	4.879	0.109
9	-17.986	-0.354	1.082		2.518	-4.160	0.079
10	24.068	9.221	2.222		-1.833	3.251	0.047
11	6.772	29.046	2.975		-7.986	0.086	0.213
12	6.513	19.100	1.362		-6.593	2.382	0.164
13	-14.045	4.861	0.739		-1.136	-1.011	0.008
14	7.006	5.499	0.265		-0.416	4.614	0.072
15	-12.232	-0.220	0.501		2.757	-2.479	0.046
16	-4.095	10.651	0.436		-2.442	-1.782	0.031
17	27.974	37.122	7.226		-15.781	8.360	1.067
18	15.760	0.332	0.831		-2.760	9.121	0.304
19	-15.878	25.803	3.070		-8.486	-2.755	0.266
20	-12.941	3.660	0.605		2.903	-0.213	0.028
21	-6.876	3.117	0.191		4.430	-0.202	0.066
22	-7.603	16.135	1.064		-2.006	-4.821	0.091
23	10.893	3.278	0.433		-0.860	3.981	0.055
24	-1.807	2.188	0.027		-0.925	-0.433	0.003
25	6.352	3.906	0.186		-2.747	5.641	0.132
26	-15.958	9.293	1.141		-0.486	-5.741	0.111
27	-1.211	4.139	0.062		1.694	1.355	0.016
28	-2.470	8.722	0.275		-0.962	-2.320	0.021
29	-1.966	7.410	0.197		-2.472	-2.829	0.047
30	5.645	3.831	0.156		-3.833	3.540	0.091
31	-6.250	-8.375	0.365		5.946	1.261	0.124
32	-4.204	11.282	0.485		-4.440	-0.456	0.067
33	1.297	0.221	0.006		3.515	4.381	0.106
34	-6.915	8.737	0.415		0.436	-5.530	0.103
35	6.451	9.679	0.453		-5.304	1.113	0.098

36	4.428	-1.721	0.075	0.576	2.927	0.030
37	-3.494	-0.777	0.043	1.107	-1.192	0.009
38	-5.146	6.418	0.226	-2.804	-3.227	0.061
39	2.543	1.395	0.028	-0.356	3.672	0.046
40	-2.029	3.205	0.048	-0.046	-1.589	0.008
41	-0.710	0.683	0.003	0.136	-0.658	0.002
42	2.753	-2.558	0.047	-0.237	3.777	0.048
43	-1.983	-0.199	0.013	-0.822	0.864	0.005
44	-2.092	6.739	0.167	-5.416	-0.805	0.100
45	1.616	0.120	0.009	-0.503	4.626	0.072
46	-1.171	3.003	0.035	-1.685	0.088	0.010
47	2.230	-0.541	0.018	-1.539	4.111	0.064
48	-4.169	0.302	0.058	-0.566	-0.416	0.002
49	-2.571	2.220	0.039	-2.674	2.110	0.039
50	-0.223	2.317	0.018	-1.976	5.335	0.108
51	-3.111	4.044	0.087	0.221	1.198	0.005
52	0.444	5.795	0.113	-2.460	1.291	0.026
53	-0.067	0.737	0.002	1.188	1.979	0.018
54	-1.297	1.347	0.012	-0.380	0.677	0.002
55	-2.114	0.052	0.015	0.734	2.193	0.018
56	-2.637	2.035	0.037	0.388	1.659	0.010
57	-1.280	3.666	0.050	-0.320	1.371	0.007
58	0.178	2.683	0.024	0.078	1.571	0.008
59	0.434	1.085	0.005	0.074	1.530	0.008
60	-1.011	0.112	0.003	0.124	1.269	0.005
61	-2.434	0.748	0.022	-0.046	1.310	0.006
62	-2.548	1.896	0.034	0.278	1.833	0.011
63	-2.084	3.269	0.050	0.631	0.851	0.004
64	0.040	3.043	0.031	-0.025	1.094	0.004
65	0.574	1.031	0.005	0.202	1.476	0.007
66	-0.683	0.506	0.002	-0.471	0.893	0.003
67	-2.046	-1.001	0.017	1.146	2.881	0.032
68	-2.678	2.715	0.049	-0.727	1.192	0.007
69	-1.747	2.259	0.027	1.473	2.700	0.032
70	-0.429	4.125	0.058	0.188	0.408	0.001
71	0.310	3.046	0.031	0.189	-0.372	0.001
72	2.180	-1.256	0.021	0.753	4.412	0.067
73	-2.147	-0.278	0.016	2.179	-0.027	0.016
74	-1.490	1.370	0.014	-0.262	0.498	0.001
75	-1.773	2.024	0.024	0.367	0.125	0.001
76	-0.635	1.177	0.006	0.743	1.419	0.009
77	-0.981	2.880	0.031	-1.171	-1.551	0.013
78	0.390	1.433	0.007	-2.388	1.691	0.029
79	-1.503	0.756	0.009	0.064	2.155	0.016
80	-2.128	2.354	0.034	0.735	1.179	0.006
81	-1.477	1.751	0.018	3.107	2.063	0.047
82	1.617	3.167	0.042	-0.390	1.097	0.005
83	1.569	1.255	0.013	0.692	0.526	0.003
84	0.316	-0.730	0.002	1.694	0.576	0.011
85	-1.649	-0.014	0.009	1.404	-1.117	0.011
86	-1.728	1.174	0.015	0.361	-0.845	0.003
87	-1.949	2.445	0.033	0.977	-2.125	0.018
88	1.321	1.871	0.018	-1.395	0.303	0.007
89	2.168	-1.925	0.028	0.429	3.861	0.050
90	1.216	-2.337	0.023	-0.313	2.640	0.024
91	-1.483	-2.172	0.023	0.889	1.077	0.007
92	-1.422	0.589	0.008	-2.269	-0.395	0.018
93	-1.208	1.382	0.011	-2.154	0.050	0.016
94	-0.508	-0.551	0.002	-0.127	3.081	0.032
95	-2.940	-0.811	0.031	3.445	0.269	0.040
96	-1.792	0.630	0.012	-0.643	-0.780	0.003
97	-2.694	1.217	0.029	-0.545	-1.051	0.005
98	-2.682	2.006	0.038	-0.229	-0.582	0.001
99	-2.682	1.866	0.036	2.099	-0.321	0.015
100	0.245	1.654	0.009	-0.717	0.561	0.003
101	0.092	-1.158	0.005	0.922	2.262	0.020
102	-1.513	-2.119	0.023	1.361	1.822	0.017
103	-2.771	-0.746	0.028	-0.139	0.813	0.002
104	-2.874	0.545	0.029	-0.683	2.068	0.016
105	-2.528	1.142	0.026	1.236	3.273	0.041
106	-0.924	2.054	0.017	1.197	1.964	0.018
107	-0.603	1.396	0.008	2.235	0.193	0.017
108	-0.584	-0.345	0.002	1.829	0.665	0.013
109	-1.871	-0.253	0.012	1.266	-0.299	0.006
110	-2.463	0.118	0.020	0.953	0.813	0.005

111	-2.105	1.745	0.025	0.559	0.188	0.001
112	-0.818	1.564	0.010	0.818	1.088	0.006
113	0.017	0.632	0.001	0.953	1.098	0.007
114	-0.413	-0.524	0.001	1.130	0.906	0.007
115	-1.685	-0.706	0.011	1.203	0.566	0.006
116	-2.550	0.210	0.022	1.194	0.430	0.005
117	-2.013	1.385	0.020	0.836	0.528	0.003
118	-1.024	1.684	0.013	1.223	0.390	0.006
119	0.000	0.838	0.002	0.991	0.395	0.004
120	-0.259	-0.394	0.001	1.000	0.407	0.004
121	-1.437	-0.778	0.009	1.075	0.347	0.004
122	-2.417	0.023	0.020	1.264	0.206	0.005
123	-2.042	1.189	0.019	1.038	0.154	0.004
124	-0.731	1.222	0.007	0.859	0.731	0.004
125	0.165	0.293	0.000	0.855	0.834	0.005
126	-0.445	-0.855	0.003	1.435	0.456	0.008
127	-1.809	-0.869	0.013	1.525	-0.497	0.009
128	-2.622	-0.209	0.023	1.146	-0.294	0.005
129	-2.244	0.897	0.020	0.570	-0.017	0.001
130	-1.467	1.507	0.015	0.930	-0.313	0.003
131	-0.463	0.679	0.002	0.742	0.157	0.002
132	-1.038	-0.086	0.004	1.446	-0.513	0.008
133	-1.388	-0.655	0.008	0.214	0.165	0.000
134	-2.742	-0.280	0.025	1.325	0.792	0.008
135	-2.407	0.705	0.021	1.203	1.265	0.010
136	-0.602	2.133	0.016	-0.266	-0.405	0.001
137	-0.057	0.933	0.003	0.931	0.378	0.003
138	0.755	0.704	0.004	-0.904	-1.493	0.010
139	-0.791	-1.540	0.010	0.497	1.490	0.008
140	-3.056	0.068	0.031	2.688	-0.932	0.027
141	-2.166	0.026	0.016	0.991	1.564	0.011
142	-0.908	0.946	0.006	0.252	1.031	0.004
143	-0.786	1.430	0.009	1.847	-0.951	0.014
144	0.535	0.269	0.001	-0.504	-0.549	0.002
145	-1.138	-1.059	0.008	1.361	0.841	0.009
146	-1.989	-0.320	0.014	1.398	0.183	0.007
147	-2.054	0.053	0.014	1.605	1.270	0.014
148	-0.519	0.542	0.002	0.287	1.162	0.005
149	0.476	0.691	0.002	0.026	-0.298	0.000

NORMALIZED CORRELATION COEFFICIENTS

(ASSUME LONG SERIES IS STRETCHED)

(ASSUME SHORT SERIES IS STRETCHED)

LAG NUMBER	VALUE OF COEFFICIENT	LAG NUMBER	VALUE OF COEFFICIENT
0	0.900	0	0.900
-1	0.760	1	0.817
-2	0.507	2	0.586
-3	0.327	3	0.378
-4	0.290	4	0.272
-5	0.301	5	0.214
-6	0.253	6	0.144
-7	0.148	7	0.061
-8	0.057	8	0.003
-9	0.019	9	-0.010
-10	0.021	10	0.018
-11	0.035	11	0.063
-12	0.047	12	0.096
-13	0.055	13	0.105
-14	0.060	14	0.109
-15	0.076	15	0.143
-16	0.116	16	0.228
-17	0.178	17	0.333
-18	0.228	18	0.409
-19	0.226	19	0.443
-20	0.175	20	0.443
-21	0.115	21	0.428
-22	0.076	22	0.399
-23	0.064	23	0.381
-24	0.093	24	0.428
-25	0.154	25	0.530
-26	0.179	26	0.569
-27	0.160	27	0.510
-28	0.134	28	0.482

-29	0.171	29	0.520
-30	0.262	30	0.544

STRETCH FACTOR FOUND FROM CORRELATION OF POWER SPECTRA

FIRST CHOICE - LONG SERIES IS STRETCHED 1.00 TIMES

SECOND CHOICE - SHORT SERIES IS STRETCHED 1.82 TIMES

FINAL RESULT SUGGEST THAT LONG SERIES IS STRETCHED 1.00

TIMES MAXIMUM CORRELATION IS 1.00 AT A LAG OF 51

THE CORRELATION MATRIX IS :

1.000	0.608	0.011	0.033	0.017	0.290	-0.409
0.608	1.000	-0.023	-0.041	-0.053	0.697	-0.210
0.011	-0.023	1.000	0.715	0.639	-0.140	-0.102
0.033	-0.041	0.715	1.000	0.816	-0.104	-0.038
0.017	-0.053	0.639	0.816	1.000	-0.113	-0.073
0.290	0.697	-0.140	-0.104	-0.113	1.000	0.251
-0.409	-0.210	-0.102	-0.038	-0.073	0.251	1.000

EIGENVECTORS :

VARIABLE	1	2	3	4	5	6	7
SP	0.06832	0.55669	0.28581	0.64582	-0.39462	0.09737	-0.14650
GR	0.16567	0.61768	-0.11185	-0.27119	0.08338	-0.14309	0.69105
ILS	-0.52919	0.12941	-0.06416	-0.47191	-0.65090	0.22166	-0.05995
ILM	-0.56791	0.14283	-0.14455	0.16975	0.13434	-0.75801	-0.12145
ILD	-0.55401	0.13198	-0.11367	0.24428	0.49401	0.57941	0.15268
DT	0.23032	0.43718	-0.54055	-0.20085	0.18917	0.09989	-0.61444
CALI	0.07316	-0.25074	-0.75874	0.39730	-0.34027	0.02547	0.28603

EIGENVALUES .:

2.517	2.101	1.275	0.410	0.365	0.171	0.162
-------	-------	-------	-------	-------	-------	-------

PERCENTAGE OF TOTAL VARIANCE CONTRIBUTED BY EACH EIGENVALUE

35.953%	30.014%	18.217%	5.862%	5.207%	2.437%	2.309%
---------	---------	---------	--------	--------	--------	--------

Table 1 The correlation matrix, the eigenvectors, the eigenvalues and the percentage of each eigenvalue to the total variance of the original data of well FF7.

THE CORRELATION MATRIX IS :

1.000	0.482	-0.030	-0.025	0.162	0.181	0.236
0.482	1.000	-0.007	-0.004	-0.231	0.639	0.582
-0.030	-0.007	1.000	0.991	0.298	-0.088	-0.055
-0.025	-0.004	0.991	1.000	0.316	-0.084	-0.051
0.162	-0.231	0.298	0.316	1.000	-0.403	-0.292
0.181	0.639	-0.088	-0.084	-0.403	1.000	0.558
0.236	0.582	-0.055	-0.051	-0.292	0.558	1.000

EIGENVECTORS :

VARIABLE	1	2	3	4	5	6	7
SP	0.23331	-0.23696	0.72340	-0.38760	-0.27827	0.37207	0.00206
GR	0.47125	-0.32089	0.13568	-0.13796	0.17025	-0.78008	0.00028
ILS	-0.28677	-0.59621	-0.21268	-0.11548	-0.07936	0.03787	-0.70427
ILM	-0.28689	-0.59984	-0.20073	-0.09312	-0.06041	0.04206	0.70958
ILD	-0.35202	-0.18014	0.55909	0.58154	0.43596	-0.04902	-0.02012
DT	0.48250	-0.19787	-0.22103	-0.01667	0.66746	0.48305	-0.00871
CALI	0.44969	-0.23238	-0.10639	0.68576	-0.49807	0.11877	-0.00418

EIGENVALUES :

2.628	2.020	1.141	0.504	0.426	0.272	0.009
-------	-------	-------	-------	-------	-------	-------

PERCENTAGE OF TOTAL VARIANCE CONTRIBUTED BY EACH EIGENVALUE

37.543%	28.857%	16.302%	7.201%	6.086%	3.883%	0.128%
---------	---------	---------	--------	--------	--------	--------

Table 2 The correlation matrix, the eigenvectors, the eigenvalues and the percentage of each eigenvalue to the total variance of the original data in well FF13.

THE CORRELATION MATRIX IS :

1.000	0.613	-0.217	-0.027	-0.029	0.754	0.789
0.613	1.000	-0.217	0.003	-0.041	0.593	0.270
-0.217	-0.217	1.000	0.125	0.491	-0.411	-0.200
-0.027	0.003	0.125	1.000	-0.009	-0.112	-0.046
-0.029	-0.041	0.491	-0.009	1.000	-0.182	-0.076
0.754	0.593	-0.411	-0.112	-0.182	1.000	0.737
0.789	0.270	-0.200	-0.046	-0.076	0.737	1.000

EIGENVECTORS :

VARIABLE	1	2	3	4	5	6	7
SP	0.50901	-0.24541	0.03065	0.07228	-0.08097	-0.59884	-0.55620
GR	0.38905	-0.14309	0.11509	-0.78337	-0.19484	-0.04588	0.40148
ILS	-0.27747	-0.61005	0.00786	0.10941	-0.70899	0.18012	-0.06082
ILM	-0.06040	-0.13908	0.96629	0.07833	0.18264	0.05835	-0.01990
ILD	-0.14195	-0.70302	-0.22359	-0.12892	0.64602	0.03702	0.01730
DT	0.52599	-0.02533	-0.03546	0.02660	0.04595	0.77393	-0.34591
CALI	0.46054	-0.18144	-0.02804	0.58795	-0.00248	-0.05570	0.63672

EIGENVALUES :

3.092	1.380	1.008	0.744	0.462	0.215	0.099
-------	-------	-------	-------	-------	-------	-------

PERCENTAGE OF TOTAL VARIANCE CONTRIBUTED BY EACH EIGENVALUE

44.170%	19.707%	14.402%	10.626%	6.600%	3.078%	1.417%
---------	---------	---------	---------	--------	--------	--------

Table 3 The correlation matrix, the eigenvectors, the eigenvalues and the percentage of each eigenvalue to the total variance of the original data in FF11.

THE CORRELATION MATRIX IS :

1.000	0.760	-0.044	0.035	0.000	0.424	0.181
0.760	1.000	-0.250	-0.045	-0.026	0.655	0.291
-0.044	-0.250	1.000	0.102	0.017	-0.525	-0.281
0.035	-0.045	0.102	1.000	0.045	-0.073	-0.068
0.000	-0.026	0.017	0.045	1.000	-0.044	-0.025
0.424	0.655	-0.525	-0.073	-0.044	1.000	0.696
0.181	0.291	-0.281	-0.068	-0.025	0.696	1.000

EIGENVECTORS :

VARIABLE	1	2	3	4	5	6	7
SP	0.40677	0.56943	-0.14329	0.15366	-0.04190	-0.63605	-0.24461
GR	0.50483	0.35806	-0.11692	0.13617	-0.14373	0.49483	0.56498
ILS	-0.32103	0.47830	-0.17317	0.09317	0.72187	0.29392	-0.14835
ILM	-0.06365	0.40162	0.38468	-0.82126	-0.10268	0.02996	0.02753
ILD	-0.03239	0.13881	0.86814	0.47388	0.03030	0.02268	-0.00575
DT	0.55431	-0.17159	0.08696	-0.10473	0.08600	0.41305	-0.68320
CALI	0.40530	-0.32979	0.16312	-0.19791	0.66154	-0.30358	0.36250

EIGENVALUES

2.732	1.232	1.011	0.938	0.719	0.238	0.129
-------	-------	-------	-------	-------	-------	-------

PERCENTAGE OF TOTAL VARIANCE CONTRIBUTED BY EACH EIGENVALUE

39.033%	17.607%	14.443%	13.400%	10.268%	3.404%	1.845%
---------	---------	---------	---------	---------	--------	--------

Table 4 The correlation matrix, the eigenvectors, the eigenvalues and the percentage of each eigenvalue to the total variance of the original data in well FF12.

THE CORRELATION MATRIX IS :

1.000	0.032	0.001	0.013	0.014	-0.291	-0.292
0.032	1.000	-0.158	-0.012	-0.010	0.775	0.584
0.001	-0.158	1.000	0.032	0.031	-0.335	-0.199
0.013	-0.012	0.032	1.000	1.000	-0.062	-0.037
0.014	-0.010	0.031	1.000	1.000	-0.060	-0.035
-0.291	0.775	-0.335	-0.062	-0.060	1.000	0.688
-0.292	0.584	-0.199	-0.037	-0.035	0.688	1.000

EIGENVECTORS :

VARIABLE	1	2	3	4	5	6	7
SP	-0.19841	0.03321	0.87026	0.27560	-0.25042	-0.25201	0.00081
GR	0.50718	-0.11669	0.28411	0.37636	0.37573	0.60466	0.00055
ILS	-0.24783	0.02497	-0.37876	0.87653	0.01057	-0.16148	-0.00049
ILM	-0.12679	-0.69518	-0.00169	-0.01912	-0.00130	-0.01495	0.70714
ILD	-0.12542	-0.69551	-0.00035	-0.01897	-0.00197	-0.01519	-0.70707
DT	0.58192	-0.09124	0.00877	0.03135	0.33237	-0.73588	0.00056
CALI	0.52110	-0.09656	-0.13560	0.11128	-0.82797	0.05321	0.00079

EIGENVALUES :

2.560	1.983	1.052	0.894	0.373	0.138	0.000
-------	-------	-------	-------	-------	-------	-------

PERCENTAGE OF TOTAL VARIANCE CONTRIBUTED BY EACH EIGENVALUE

36.567%	28.333%	15.034%	12.773%	5.327%	1.967%	0.000%
---------	---------	---------	---------	--------	--------	--------

Table 5 The correlation matrix, the eigenvectors, the eigenvalues and the percentage of each eigenvalue to the total variance of the original data in well FF10.

APPENDIX F**F1-Abbreviations**

BHC*	BoreHole Compensated log
CALI*	calliper log
CNL*	Compensated Neutron Log
DFT	Discrete Fourier Transform
d_h	hole diameter
DST	drill stem test
DT*	transit time
FFT	Fast Fourier Transform
GAPI	Gamma American Petroleum Institute
GR	Gamma Ray log
ILD*	deep resistivity log

ILM*	medium resistivity log
ILS*	shallow resistivity log
KB	Kelley Pushing
LDL*	Lateral Density Log
LIS*	Log International Standard
LIS/A*	Log International Standard Access
MMCFG/D	million cubic feet of gas per day
PCA	Principal Component Analysis
R_{mc}	mud cake resistivity
R_{mf}	mud filtrate resistivity
R_s	formation resistivity
R_t	true resistivity
R_w	formation water resistivity
SP	Spontaneous Potential log
S_w	water saturation
S_{xo}	flushed zone resistivity
T.D	total depth

F2-Symbols**Description**

A	original well-log variables matrix
c	constant
COV	covariance
D	displacement
i	imaginary number
λ	eigenvalues
P	principal component scores
R	correlation matrix
r_{xy}	correlation of between two variables
S	variance-covariance matrix
s	standard deviation
s^2	variance
SP	corrected sum of products
SS	sum of squares
t	time
U	eigenvectors

w

frequency

 \bar{x}

mean

* Schlumberger mark

