

GEOTILL Inc.

Geotechnical Engineering • Subsurface Exploration • Environmental Services • Construction Testing and Material Engineering

GEOTECHNICAL ENGINEERING LIBRARY

[GEOTILL](#)

USA



GEOTILL

ENGINEERING, INC.

Phone 317-449-0033 Fax 317- 285-0609

info@geotill.com

Toll Free: 844-GEOTILL

Geotechnical, Environmental and Construction Materials Testing Professionals

www.geotill.com

Offices Covering all USA

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

<http://geotill.com/>

NOTE TO USERS

The original manuscript received by UMI contains pages with slanted print. Pages were microfilmed as received.

This reproduction is the best copy available

UMI

**Delineating Geostratigraphy by Cluster
Analysis of Piezocone Data**

**A Thesis
Presented to
The Academic Faculty**

by

Yasser Ali Hegazy

**In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Civil Engineering**

**Georgia Institute of Technology
June, 1998**

UMI Number: 9837762

UMI Microform 9837762
Copyright 1998, by UMI Company. All rights reserved.

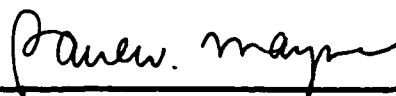
**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

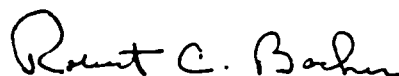
Delineating Geostratigraphy by Cluster

Analysis of Piezocone Data

Approved:



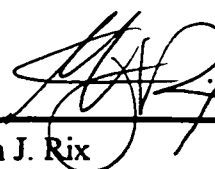
Paul W. Mayne, Chair



Bob Bachus



J. David Frost



Glenn J. Rix



Alexander Shapiro
Date Approved: 29 April 98

For Dalal, Mom, and My Mother in Law

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Paul Mayne for his continuous support during this study. His guidance and deep knowledge in the geotechnical engineering were necessary for the completion of this dissertation. Thanks also to Dr. David Frost for his effective thoughts and advice during different stages of my research. Dr. Glenn Rix provided friendly advice when I needed it. Dr. Bob Bachus offered precious review comments. I also enjoyed several discussions with Dr. Alexander Shapiro whose statistical directions were needed during this research.

I am very grateful for the support, assistance and friendship of Susan Burns and Hani Al Ansari. Appreciation is extended to Gina Martin and Ken Thomas for their help to collect some of the data necessary for this study. I would like to thank Dr. Alan Lutenneger for his sincere aid during my visit to Amherst, Massachusetts.

I owe deep appreciation and thanks to my wife, Dalal, for her love, and patient. She forgot about herself many times and thought just about me. She encouraged me throughout the course of my Ph.D. My thanks and sincere gratitude are expressed to my family for their continuous support.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	xi
ABSTRACT	xxix
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. SOIL STRATIGRAPHY	
2.1 Synopsis	5
2.2 Soil Stratification Based on Borehole Sampling and Laboratory Testing	6
2.3 Stratification by In-Situ Tests	11
2.4 Cone Penetration Testing	12
2.5 Subsurface Stratification Using Piezocone Data	16
2.5.1 Visual Method	17
2.5.2 Profiling Stratigraphy Using Cone Classification Charts	20
2.6 Limitations of Classification Methods Using Cone Data	28
2.6.1 Limitations Due to Size of Database	28
2.6.2 Limitations Due to Equipment and Test Procedure	29
2.6.3 Soil Transitions and Lenses	35
2.7 Summary and Conclusions	41
CHAPTER 3. STATISTICAL IDENTIFICATION OF SOIL LAYER BOUNDARIES	
3.1 Introduction	44
3.2 Autocorrelation Function	48
3.3 Intraclass Correlation Coefficient and Generalized Distance	57
3.4 Application of ICC and D ² Methods to Detect Subtle Changes in Piezocone Data	64
3.4.1 ICC and D ² at Drammen, Norway	64
3.4.1.1 Effect of the Window Width on the Chosen Boundaries	65
3.4.2 ICC and D ² at Lilla Mellösa, Sweden	70
3.4.2.2 Uncertainty of Soil Demarcation	72
3.5 Summary and Conclusions	75

CHAPTER 4. CLUSTER ANALYSIS

4.1	Introduction	77
4.2	Cluster Analysis Definition and Criterion	77
4.3	Current Applications of Cluster Analysis	78
4.4	Cluster Analysis Components	79
4.5	Selection of Variables for Cluster Analysis	80
4.6	Standardization of the Data	82
4.7	Resemblance (Distance) Matrix	90
4.8	Clustering Techniques	97
4.9	Hierarchical Techniques	97
4.9.1	Limitations of Hierarchical Techniques	103
4.10	Partitioning Techniques (k-means)	104
4.10.1	Limitations of Partitioning Techniques	104
4.11	Other Clustering Techniques	106
4.12	Simple Examples of Hierarchical Techniques	107
4.13	Determining the Number of Clusters	114
4.14	Interpretation of Clustering Results	116
4.15	Validity of the Cluster Analysis	122
4.15.1	Visual Interpretation	122
4.15.2	Hypothesis Testing	123
4.15.3	Validation Using Replications	124
4.5.	Summary and Conclusions	124

CHAPTER 5. CLUSTER APPLICATIONS OF PIEZOCONE DATA

5.1	Synopsis	126
5.2	Database	128
5.3	Applications of Cluster Analysis	129
5.3.1	Cluster Analysis at Amherst , Massachusetts	131
5.3.2	Unprocessed and Processed Data	140
5.4	Factors Affecting Cluster Results	142
5.4.1	Data Errors	142
5.4.1.1	Effect of Electrical Noise	143
5.4.1.2	Effect of Procedural Errors	146
5.4.2	Actual Versus Assumed Unit Weight	149
5.4.3	Data Frequency	153
5.4.4	Porewater Pressure Position	154
5.4.4.1	Bothkennar Test Site, Scotland	156
5.4.4.2	Brent Cross, England	161
5.4.5	Scale Effect	167
5.5	Spatial Cluster Analysis	173
5.6	Conclusions	180

CHAPTER 6. DEMARCATING SUBTLETIES IN PIEZOCONE PROFILES BY CLUSTERING

6.1	Synopsis	182
6.2	Clustering Applications	183
6.2.1	Drammen Test Site	184
6.2.2	Gloucester Test Site	190
6.2.3	Hachirogata Test Site	198
6.2.4	Lilla Mellösa Test Site	203
6.2.5	Recife Test Site	208
6.2.6	St. Alban Test Site	213
6.2.7	Tiller Test Site	219
6.2.8	Troll Site	224
6.3	Discussion	226
6.3	Conclusions	232

CHAPTER 7. SUMMARY AND RECOMMENDATIONS

7.1	SUMMARY AND CONCLUSIONS	233
7.2	RECOMMENDATIONS FOR FUTURE RESEARCH	238

APPENDIX A. FIELD EXPERIENCES WITH CONE PENETRATION TESTING

A.1	Preface	243
A.2	Georgia Tech Cone Penetration Systems	243
A.3	Piedmont Residual Soils, Atlanta, Georgia	248
A.4	Bagdad Mine Tailings, Arizona	250
A.5	Amherst Varved Clay, Massachusetts	250
A.6	Dunklin County Alluvial Deposit, Missouri	253
A.7	Opelika Piedmont Soils, Alabama	256
A.8	Summary	256

APPENDIX B. CLUSTERING OF PIEZOCONE DATA

B.1	Synopsis	258
B.2	Amherst Test Site	259
B.3	McDonald's Farm Test Site	264
B.4	Fort Road Test Site, Singapore	271
B.5	Conclusions	285

APPENDIX C. GROWTH OF CLUSTERING

C.1	Synopsis	286
C.2	Effect of Different Data Combinations	286
C.3	Procedural Errors	294
C.4	Actual Versus Assumed Unit Weight	302

C.5	Data Resolution	302
C.6	Porewater Position	320
C.6.1	Bothkennar Test Site, Scotland	321
C.6.2	Brent Cross Test Site, UK	328
C.7	Scale Effect	334
C.8	Spatial Cluster Analysis	342
C.9	Drammen Test Site	350
C.10	Gloucester Test Site, Ontario	355
C.11	Hachirogata Test Site	359
C.12	Lilla Mellösa Test Site	359
C.13	Recife Test Site	363
C.14	St. Alban Test Site, Quebec	369
C.15	Tiller Test Site	369
C.16	Troll Test Site	374
C.17	Summary	380

APPENDIX D CLUSTERING ASSESSMENT OF PIEZOCONES DATA AT DIFFERENT SITES

D.1	Synopsis	382
D.2	Aiken Test Site, South Carolina	383
D.3	South Boston Test Site, Massachusetts	386
D.4	Kagoshima Test Site	396
D.5	Newport News Test Site, Virginia	403
D.6	Onsoy Test Site, Norway	410
D.7	Opelika Test Site, Alabama	414
D.8	Po River Sand, Italy	423
D.9	Penuelas Test Site, Puerto Rico	430
D.10	Maskinongé Test Site, Québec	433
D.11	Taranto Test Site, Italy	440
D.12	Conclusions	449

REFERENCES	450
------------	-----

LIST OF TABLES

Table 2.1	Summary of Available Soil Classification Methods Using Cone Data	21
Table 2.2	Cone Classification Systems Used in Different Countries Summarized from CPT'95 Proceedings	39
Table 3.1	Definitions of Statistical Terms, Modified after Lacasse and Lamballerie (1995)	44
Table 3.2	Applied Statistical Methods for Delineating Soil Stratigraphy Based on Cone Data	45
Table 3.3	Common Fitting Variogram Forms for Autocorrelation Function (Based on Journel and Huijbregts, 1993)	52
Table 4.1	A Comparison between Means and Variances of Piezocone Parameters at Bothkennar, Scotland (Piezocone Data from Nash et al., 1992)	85
Table 4.2	Different Methods of Data Standardization [adapted from Romesburg (1984) and Milligan (1996)]	85
Table 4.3	Summary of Methods for Determining the Resemblance Matrix [Modified from Noursis (1993)]	93
Table 4.4	Four Selected Pairs of Q and B_q of PCPT1 sounding at Amherst, Massachusetts (Piezocone Data from This Study)	94
Table 4.5	Summary of the Similarity Measurements between Four Selected Points of Q and B_q of PCPT1 sounding at Amherst, Massachusetts (Piezocone Data from This Study)	94
Table 4.6	Methods of Hierarchical Type of Clustering [from Everitt (1974), and Gordon (1996)]	100
Table 4.7	Effect of Artificial Outliers on a Rank of a Cluster Analysis Method Using Simulated Data (after Milligan, 1980)	101
Table 4.8	Selected Points of Q and B_q of PCPT1 at Amherst, Massachusetts (Data from this Study)	108
Table 4.9	A Summary of Cosine Measurement (Similarity) between Selected Points of Q and B_q of PCPT1 at Amherst, Massachusetts (Data from this Study)	108

Table 4.10	Summary of Clustering Consequence of Selected Piezocone Data at Amherst Using the Cosine Method for the Similarity and Single Linkage for Grouping the Data	109
Table 4.11	Selected Piezocone Data and Their Standardization Using Zscore Method at Troll, North Sea (Piezocone data from Amundsen et al., 1985)	111
Table 4.12	A summary of the Cosine Measurement (Similarity) Between Selected Points of Q and B_q of PCPT1 at Troll, North Sea (Data from Amundsen et al., 1985)	112
Table 4.13	Summary of Clustering Consequence of Selected Piezocone Data at Troll, North Sea, Using the Cosine Method for the Similarity and Single Link Method for Grouping the Data (Piezocone Data from Amundsen et al., 1985)	114
Table 5.1	Summary of 24 Sites Analyzed Using Cluster Methods	130
Table 5.2	Primary Soil Boundaries Using Three Piezocone Soundings at Amherst	178
Table 6.1	Summary of 8 Clay Sites Analyzed Using SCZ-Type Clustering of Piezocone Data with Subtle Changes	183
Table A.1	Summary of Cone Penetration Tests Performed by the Author During this Research Study	244
Table B.1	Summary of 56 Different Cluster Analyses Performed on Piezocone Data (PCPT1) from Amherst, Massachusetts (Data from this Study)	263
Table B.2	Summary of 56 Different Cluster Analyses Performed on Piezocone Data from McDonald's Farm, British Columbia (Data from this Study)	270
Table B.3	Summary of 56 Different Cluster Analyses Performed on Piezocone Data from Fort Road, Singapore (Data from Chang, 1991)	280
Table C.1	Listing of 12 Sites Analyzed Using SCZ-Type Clustering Method	287
Table D.1	Summary of 10 Sites Analyzed Using Cluster Methods	382

LIST OF FIGURES

Figure 2.1	An Example of A Routine Site Investigation Program Based on Borehole Sounding and Laboratory Tests	7
Figure 2.2	Borderlines between Different Soil Types Using the Unified Soil Classification System (USCS); Modified after Howard (1977)	9
Figure 2.3	Locations of Different Elements of a Piezocone	13
Figure 2.4	Effect of Pore Water Pressure on the Tip Resistance and Sleeve Friction (from Jamiolkowski et al., 1985)	14
Figure 2.5	Piezocone Data and Soil Stratigraphy at Amherst, MA (Data By Author During This Study)	19
Figure 2.6	Piezocone Data and Soil Stratigraphy at McDonald's Farm, Vancouver (Data from Robertson, 1982)	19
Figure 2.7	Piezocone Data and Soil Stratigraphy at Drammen, Norway (Data from Masood et al., 1990)	20
Figure 2.8	Comparison of Piezocone Soundings in Two Different and Distinct sandy clay deposits, yet Identical in Terms of USCS	25
Figure 2.9	Factors Affecting Cone Measurements	27
Figure 2.10	One Example of Cone Classification Charts (Robertson et al., 1986)	28
Figure 2.11	Reliability of Piezocone Measurements in Soft Clay Deposit in Japan Using 8 Different Cones (Data from Tanaka, 1995)	31
Figure 2.12	A Comparison between Electronic 10-cm² and Electrical 15-cm² Piezocones in Piedmont Residual Silts and Sands in Atlanta, Georgia (Data from This Study)	34
Figure 2.13	Transition Zone of the Cone Tip Resistance Between: (I) Soft Layer above Stiff Layer; (II) Stiff Layer above Soft Layer	36
Figure 2.14	Correction of the Tip Resistance in a Transition Zone Between Two Layers (from Lunne et al., 1997)	37

Figure 2.15	Soil Classification Using Piezocone Data at Drammen, Norway (Data from Masood et al., 1990)	42
Figure 3.1	Effect of the Window Width on the Delineation of Soil Boundaries	49
Figure 3.2	Combination Sequence to Calculate Autocorrelation Function (r_{ah}) at Different Distances	51
Figure 3.3	Common Models Fitted to the Autocorrelation Function (Modified after Journal and Huijbregts, 1993)	52
Figure 3.4	Piezocone Data at Surry, Virginia and Subsurface Profile	54
Figure 3.5	Autocorrelation Function of Piezocone Data at Surry, Virginia (Modified from Hegazy et al., 1996)	54
Figure 3.6	Piezocone Data at Opelika, Alabama and Subsurface Profile	56
Figure 3.7	Autocorrelation Function of Piezocone Data at Opelika, Alabama (Modified from Hegazy et al., 1997)	56
Figure 3.8	Movement of the Correlation Window along a Cone Measurement Profile, and the Recorded Interclass Correlation Coefficient ICC	58
Figure 3.9	Concept of the Generalized Distance (D^2) Statistics for Two Samples of q_t and f_t (after Harbaugh and Merriam, 1968)	62
Figure 3.10	Two Classes for which ICC and D^2 Methods Are Unable to Detect the Soil Layer Boundaries (after Wickremesinghe, 1989)	63
Figure 3.11	Piezocone Data at Drammen, Norway (after Masood et al., 1990)	66
Figure 3.12	Autocorrelation Function at Drammen, Norway (Piezocone Data from Masood et al., 1990)	66
Figure 3.13	Effect of Window Width on Determination of ICC and D^2 Using Piezocone Data at Drammen, Norway	67
Figure 3.14	Statistical Soil Boundaries of ICC and D^2 Using Piezocone Data at Drammen, Norway	69
Figure 3.15	Piezocone Data and Soil Profile at Lilla Mellosa, Sweden	71
Figure 3.16	Autocorrelation Function of Piezocone Data at Lilla Mellosa, Sweden	71

Figure 3.17	Effect of Window Width on ICC and D₂ Using Piezocone Data from Lilla Mellosa, Sweden	73
Figure 3.18	Statistical Soil Boundaries of ICC and D² Using Piezocone Data at Lilla Mellosa, Sweden	74
Figure 4.1	Steps of Clustering Analysis	80
Figure 4.2	Piezocone Results in a Uniform Soft Clay at Bothkennar, Scotland (Piezocone Data from Nash et al., 1992)	83
Figure 4.3	Piezocone Data and Soil Profile at Amherst, Massachusetts (Data Collected by the Author During This Study)	87
Figure 4.4	Normalized Piezocone Parameters Q and B_q and Soil Profile at Amherst, Massachusetts (Data from This Study)	87
Figure 4.5	Comparison of Different Standardization Methods at Amherst, Massachusetts (Piezocone Data from This Study)	88
Figure 4.6	Definitions of the Euclidean and the Cosine Measurements in the “Octahedral” Space of Piezocone Data	91
Figure 4.7a	Piezocone Data and Soil Stratigraphy at McDonald’s Farm, British Columbia (Data from Robertson, 1982)	96
Figure 4.7b	Piezocone Data and Soil Stratigraphy at Fort Road, Singapore (Data from Chang, 1991)	96
Figure 4.8	Concept of Grouping Piezocone Data into Clusters	98
Figure 4.9	A Dendogram (Tree) Diagram Summarizes the Consequence of Clustering of Piezocone data at Amherst, Massachusetts	110
Figure 4.10	Piezocone Data and Soil Profile at Troll, North Sea (Data from Amundsen et al., 1985)	113
Figure 4.11	A Dendogram (Tree) Diagram Summarizes the Consequence of Clustering of Piezocone Data at Troll, North Sea (Piezocone Data from Amundsen et al., 1985)	113
Figure 4.12	Proposed Criteria for Evaluating Soil Stratigraphy from Cluster Analysis of Piezocone Data	119
Figure 4.13	Proposed Clustering Analysis of Cone Data	121

Figure 5.1	Soil Classification Using Piezocone Data (PCPT1) at Amherst, Massachusetts (This Study)	133
Figure 5.2	Normalized Piezocone Parameters Q and B_q and Soil Profile at Amherst, Massachusetts (Data from This Study)	134
Figure 5.3	Zscore of Normalized Piezocone Parameters Q and B_q and Soil Profile at Amherst, Massachusetts (Data from This Study)	134
Figure 5.4	Correlation Coefficient between Consecutive Cluster Results at Amherst, Massachusetts, (a) Normal, and (b) Expanded Scales	135
Figure 5.5	Cluster Analysis of Piezocone Data at Amherst, Massachusetts ($N_c = 2$ to 31)	136
Figure 5.6	Cluster Analysis of Piezocone Data at Amherst, Massachusetts ($N_c = 45$ to 100)	138
Figure 5.7	Comparison between Cluster Analysis, Visual Classification and Water Content % at Amherst, Massachusetts	139
Figure 5.8	Effect of Using Raw and Processed Piezocone Data on the Cluster Analysis at Amherst, Massachusetts	141
Figure 5.9	Measured and Filtered q_c Data Collected in Chamber Deposit of Kaolinitic Clay (Data from Mayne et al., 1992)	144
Figure 5.10	Data Noise and Effect on Cluster Analysis of q_c Data in Chamber Deposit of Kaolinitic Clay (Data from Mayne et al., 1992)	145
Figure 5.11	Measured and Filtered Piezocone Data (PCPT1) and Soil Profile at Amherst, Massachusetts (Data from This Study)	147
Figure 5.12	Cluster Analysis of Filtered Piezocone Data (PCPT1) at Amherst, Massachusetts (Data from This Study)	148
Figure 5.13	Effect of Using Filtered Data and Assumed Versus Measured Unit Weight on the Cluster Analysis at Amherst, Massachusetts	150
Figure 5.14	Effect of Data Resolution of the Filtered Data on the Cluster Analysis at Amherst, Massachusetts Using 10-cm² Cone	151
Figure 5.15	Effect of Using Actual γ_t on the Cluster Analysis of Piezocone Data (PCPT1) at Amherst, Massachusetts (This Study)	152

Figure 5.16	Effect of Data Frequency of the Filtered Data on the Cluster Analysis at Amherst, Massachusetts Using 10-cm² Cone	155
Figure 5.17	Piezocone Data and Soil Stratigraphy at Bothkennar, UK (Data from Nash et al., 1992)	157
Figure 5.18	Soil Classification Using Piezocone Data at Bothkennar, UK (Data from Nash et al., 1992)	158
Figure 5.19	Clustering of Types 1 and 2 Piezocone Data at Bothkennar, UK	158
Figure 5.20	Comparison Between Clustering Using Midface Penetration Porewater Pressure (u_b) Versus Shoulder (u_t) with Profiles of Liquidity Index and Overconsolidation Ratio at Bothkennar, UK	160
Figure 5.21	Piezocone Data and Soil Stratigraphy at Brent Cross, UK (Data from Powell et al., 1988)	162
Figure 5.22	Soil Classification Using Piezocone Data at Brent Cross, UK (Data from Powell et al., 1988)	163
Figure 5.23	Clustering of Type 1 and 2 Piezocone Data at Brent Cross, UK	163
Figure 5.24	Comparison Between Stratification from Cluster Results Using u_t and u_b and Soil Properties at Brent Cross, UK	165
Figure 5.25	Piezocone Data from 10- and 15-cm² Cones at Surry, Virginia (Data from Gordon and Mayne, 1987)	168
Figure 5.26	Soil Classification Using Piezocone Data at Surry, Virginia (Data from Grodon and Mayne, 1987)	170
Figure 5.27	Cluster Analysis of Piezocone Data with Different Size Penetrometers at Surry, Virginia, (Data from Gordon and Mayne, 1987)	170
Figure 5.28	Comparison between Cluster Analysis Using 10- and 15-cm² Cones and Borehole Classification at Surry, Virginia	172
Figure 5.29	Test Location and Site Plan of National Geotechnical Experimental Site (NGES) at Amherst, Massachusetts	174
Figure 5.30	Piezocone Data and Soil Stratigraphy at Amherst, Massachusetts (Data from This Study)	175
Figure 5.31	Normalized Piezocone Data and Soil Stratigraphy at Amherst, Massachusetts (Data from This Study)	176

Figure 5.32	Cluster Analysis of 3 Piezocone Soundings at Amherst, Massachusetts (Data from This Study)	177
Figure 5.33	Normalized Data from Three Piezocone Tests at Amherst, Massachusetts (This Study)	179
Figure 6.1	A Representative Piezocone Sounding from Drammen, Norway (Data from Masood et al., 1990)	185
Figure 6.2	Cluster Analysis of Piezocone Data at Drammen, Norway	186
Figure 6.3	Clustering Results of Piezocone Data at Drammen, Norway on a $Q-B_q$ Space (Data from Masood et al., 1990)	186
Figure 6.4	Comparison between Cluster Analysis, Visual Classification and Water Content at Drammen, Norway	188
Figure 6.5	Comparison between Cluster Analysis, Soil Profile, and Undrained Shear Strength at Drammen, Norway	189
Figure 6.6	A Representative Piezocone Data at Gloucester, Ontario (Data from Konrad and Law, 1987)	191
Figure 6.7	Cluster Analysis of Piezocone Data at Gloucester, Ontario	193
Figure 6.8	Clustering Results of Piezocone Data at Gloucester, Ontario on a $Q-B_q$ Space (Data from Konrad and Law, 1987)	193
Figure 6.9	Comparison between Cluster Analysis, Visual Classification, and Water Content at Gloucester, Ontario	194
Figure 6.10	Comparison between Cluster Analysis, Visual Classification, and Vertical Effective Stresses at Gloucester, Ontario	196
Figure 6.11	Comparison Between Cluster Analysis, Visual Classification and Vane Shear Strength at Gloucester, Ontario	197
Figure 6.12	A Representative Piezocone Sounding at Hachirogata, Japan (Data from Tanaka et al., 1992)	199
Figure 6.13	Cluster Analysis of Piezocone Data at Hachirogata, Japan	201
Figure 6.14	Clustering Results of Piezocone Data at Hachirogata, Japan on a $Q-B_q$ Space (Data from Tanaka et al., 1992)	201

Figure 6.15	Comparison Between Cluster Analysis, Visual Classification and Clay Fraction at Hachirogata, Japan	202
Figure 6.16	A Representative Piezocone Sounding at Lilla Mellösa, Sweden (Data from Larsson and MulabdiC, 1991)	204
Figure 6.17	Cluster Analysis of Piezocone Data at Lilla Mellosa, Sweden	206
Figure 6.18	Clustering Results of Piezocone Data at Lilla Mellosa, Sweden on a Q-Bq Space (Data from Larsson and Mulabdic, 1991)	206
Figure 6.19	Comparison Between Cluster Analysis, Visual Classification and Organic Contents at Lilla Mellösa, Sweden	207
Figure 6.20	A Representative Piezocone Sounding at Recife, Brazil (Data from Coutinho and Oliveira, 1997)	209
Figure 6.21	Cluster Analysis of Piezocone Data at Recife, Brazil	211
Figure 6.22	Clustering Results of Piezocone Data at Recife, Brazil on a Q-Bq Space (Data from Coutinho and Oliveira, 1997)	211
Figure 6.23	Comparison Between Cluster Analysis, Visual Classification and Water Content at Recife, Brazil	212
Figure 6.24	Comparison Between Cluster Analysis, Visual Classification and Overconsolidation Ratio at Recife, Brazil	214
Figure 6.25	A Representative Piezocone Sounding at St. Alban, Quebec (Data from Roy et al., 1982)	216
Figure 6.26	Cluster Analysis of Piezocone Data at St. Alban, Quebec	217
Figure 6.27	Clustering Results of Piezocone Data at St. Alban, Quebec on a Q-Bq Space (Data from Roy et al., 1982)	217
Figure 6.28	Comparison Between Cluster Analysis, Visual Classification and Water Content at St. Alban, Quebec	218
Figure 6.29	A Representative Piezocone Data at Tiller, Norway, (Data from Sandven, 1990)	220
Figure 6.30	Cluster Analysis of Piezocone Data at Tiller, Norway	222
Figure 6.31	Clustering Results of Piezocone Data at Tiller, Norway on a Q-Bq Space (Data from Sandven, 1990)	222

Figure 6.32	Comparison Between Cluster Analysis, Visual Classification and Sensitivity at Tiller, Norway	223
Figure 6.33	Piezocone Data at Troll, North Sea (Data from Amundsen et al., 1985)	225
Figure 6.34	Cluster Analysis of Piezocone Data at Troll, North Sea	227
Figure 6.35	Clustering Results of Piezocone Data at Troll, North Sea on a Q-Bq Space (Data from Amundsen et al., 1985)	227
Figure 6.36	Comparison between Cluster Analysis, Visual Classification, and Water Content at Troll, North Sea	228
Figure 6.37	Comparison between Cluster Analysis, Visual Classification, and Sensitivity at Troll, North Sea	229
Figure 6.38	Comparison between Plastic and Lean Clays at Recife and Troll Test Sites	231
Figure 6.39	Evaluation of the Boundaries between Different Fine Soils Defined in Robertson Classification Chart (1991)	231
Figure 7.1	Schematic Diagram of a Geoenvironmental Cone	240
Figure A.1	Tested Sites by the Author During this Research Study	246
Figure A.2	Piezocones penetrometers at the Georgia Institute of Technology including from left to right: (1) a miniature 5-cm² electrical cone, (2) a Davey 10-cm² type-2 electrical piezocone, (3) a Davey 10-cm² type-1 electrical piezocone, (4) a Hogentogler 10-cm² type-2 electronic seismic piezocone, (5) a Hogentogler 10-cm² dual element electronic seismic piezocone, and (6) a Fugro 15-cm² triple element electrical piezocone	246
Figure A.3	Fugro Cone Truck at the Proposed Area of a New Civil Engineering Building on Georgia Tech Campus, Atlanta, Georgia	249
Figure A.4	A Representative Cone Sounding in Piedmont Residual Sandy Silts at Georgia Tech Campus, Atlanta, Georgia	249
Figure A.5	Pushing a Piezocone in a Mine Tailings Dam, Bagdad, Arizona Using a CME-55 Drill Rig.	251
Figure A.6	A Representative Piezocone Sounding in Copper Mine Tailings at Bagdad, Arizona	251
Figure A.7	Seismic Piezocone Test Locations at Amherst, Massachusetts	252

Figure A.8	Pushing a Piezocone in a Varved Clay Deposit, Amherst, Massachusetts, Using a Portable Hydraulic Rig	254
Figure A.9	Summary Results of 15 Seismic Piezocone Soundings Performed at Amherst, Massachusetts	254
Figure A.10	A CME-85 Track-Mounted Drill Rig Used to Push Piezocones in a Sandy Deposit at Route 164, Dunklin County, Missouri	255
Figure A.11	A Representative Piezocone Sounding in a Sandy Deposit at Route 164, Dunklin County, Missouri	255
Figure A.12	Hogentogler Cone Truck During Seismic Piezocone Operation in Piedmont Residual Soils at Opelika, Alabama	257
Figure A.13	Two Seismic Piezocone Soundings in Piedmont Residual Soils at Opelika, Alabama	257
Figure B.1a	Cluster Results of Piezocone Data at Amherst Using Different Cluster Methods	260
Figure B.1b	Cluster Results of Piezocone Data at Amherst Using Different Cluster Methods	261
Figure B.1c	Cluster Results of Piezocone Data at Amherst Using Different Cluster Methods	262
Figure B.2	Piezocone Sounding at McDonald's Farm, British Columbia (Data from Robertson, 1982)	265
Figure B.3	Soil Classification Using Piezocone Data at McDonald's Farm, British Columbia (Data from Robertson, 1982)	266
Figure B.4a	Cluster Results of Piezocone Data at McDonald's Farm Using Different Cluster Methods	267
Figure B.4b	Cluster Results of Piezocone Data at McDonald's Farm Using Different Cluster Methods	268
Figure B.4c	Cluster Results of Piezocone Data at McDonald's Farm Using Different Cluster Methods	269
Figure B.5	Correlation Coefficient Between Consecutive Cluster Results at McDonald's Farm, British Columbia	271

Figure B.6a	Cluster Analysis of Piezocone Data at McDonald's Farm, British Columbia by the Single-Cosine-Zscore Method	272
Figure B.6b	Cluster Analysis of Piezocone Data at McDonald's Farm, British Columbia by the Single-Cosine-Zscore Method	273
Figure B.7	Comparison Between Cluster Analysis and Visual Classification. from Borehole at McDonald's Farm, British Columbia	274
Figure B.8	Piezocone Sounding at Fort Road, Singapore (Data from Chang, 1991)	275
Figure B.9	Soil Classification Using Piezocone Data at Fort Road, Singapore (Data from Chang, 1991)	276
Figure B.10a	Cluster Results of Piezocone Data at Fort Road, Singapore Using Different Cluster Methods	277
Figure B.10b	Cluster Results of Piezocone Data at Fort Road, Singapore Using Different Cluster Methods	278
Figure B.10c	Cluster Results of Piezocone Data at Fort Road, Singapore Using Different Cluster Methods	279
Figure B.11	Correlation Coefficient Between Consecutive Cluster Results at Fort Road, Singapore	281
Figure B.12a	Cluster Analysis of Piezocone Data at Fort Road, Singapore	282
Figure B.12b	Cluster Analysis of Piezocone Data at Fort Road, Singapore	283
Figure B.13	Comparison Between Cluster Analysis, Visual Classification and Water Content % at Fort Road, Singapore	284
Figure C.1	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using q_t and u_b	288
Figure C.2a	SCZ-Type Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using q_t and u_b	289
Figure C.2b	SCZ-Type Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using q_t and u_b	290
Figure C.3	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using q_t and the Ratio u_b/q_t	292

Figure C.4a	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using q_t and the Ratio u_b/q_t	293
Figure C.4b	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using q_t and the Ratio u_b/q_t	295
Figure C.5	Correlation Coefficient Between Consecutive Cluster Results at Amherst Using Normalized Parameters Q, B_q and F	296
Figure C.6a	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Normalized Parameters Q, B_q and F	297
Figure C.6b	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Normalized Parameters Q, B_q and F	298
Figure C.7	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Filtered Piezocone Data	299
Figure C.8a	Cluster Analysis of Filtered Piezocone Data at Amherst	300
Figure C.8b	Cluster Analysis of Filtered Piezocone Data at Amherst	301
Figure C.9	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts	303
Figure C.10a	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Actual Unit weight	304
Figure C.10b	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Actual Unit weight	305
Figure C.11	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Resolution = 10 cm	307
Figure C.12a	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 10 cm	308
Figure C.12b	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 10 cm	309
Figure C.13	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Resolution = 20 cm	311
Figure C.14a	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 20 cm	312

Figure C.14b	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 20 cm	313
Figure C.15	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Resolution = 30 cm	313
Figure C.16	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 30 cm	315
Figure C.17	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Resolution = 40 cm	316
Figure C.18	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 40 cm	317
Figure C.19	Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Resolution = 50 cm	318
Figure D 20a	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 30 cm	319
Figure C.20	Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Resolution = 50 cm	322
Figure C.22a	Cluster Analysis of Type 2 Piezocone Data at Bothkennar, Scotland	323
Figure C.22b	Cluster Analysis of Type 2 Piezocone Data at Bothkennar, Scotland	324
Figure C.23	Correlation Coefficient Between Consecutive Cluster Results at Bothkennar, Scotland Using Type 1 Piezocone Data	325
Figure C.24a	Cluster Analysis of Type 1 Piezocone Data at Bothkennar, Scotland	326
Figure C.24b	Cluster Analysis of Type 1 Piezocone Data at Bothkennar, Scotland	327
Figure C.25	Correlation Coefficient Between Consecutive Cluster Results at Brent Cross, UK Using Type-2 Piezocone Data	329
Figure C.26a	Cluster Analysis of Type-2 Piezocone Data at Brent Cross, UK	330
Figure C.26b	Cluster Analysis of Type-2 Piezocone Data at Brent Cross, UK	331
Figure C.27	Correlation Coefficient Between Consecutive Cluster Results at Brent Cross, UK Using Type-1 Piezocone Data	332
Figure C.28a	Cluster Analysis of Type-1 Piezocone Data at Brent Cross, UK	333

Figure C.28b	Cluster Analysis of Type-1 Piezocone Data at Brent Cross, UK	335
Figure C.29	Correlation Coefficient Between Consecutive Cluster Results at Surry Using 10-cm² cone (Data from Gordon and Mayne, 1987)	337
Figure C.30a	Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 10-cm² Cone	338
Figure C.30b	Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 10-cm² Cone	339
Figure C.31	Correlation Coefficient Between Consecutive Cluster Results at Surry Using 15-cm² cone (Data from Gordon and Mayne, 1987)	340
Figure C.32a	Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 15-cm² Cone	341
Figure C.32b	Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 15-cm² Cone	343
Figure C.33	Correlation Coefficient Between Consecutive Cluster Results of Piezocone Data (PCPT2) at Amherst, Massachusetts	345
Figure C.34a	Clustering of Piezocone Sounding (PCPT2) at Amherst	346
Figure C.34b	Clustering of Piezocone Sounding (PCPT2) at Amherst	347
Figure C.35	Correlation Coefficient Between Consecutive Cluster Results of Piezocone Data (PCPT15) at Amherst, Massachusetts	348
Figure C.36a	Clustering of Piezocone Sounding (PCPT15) at Amherst	349
Figure C.36b	Clustering of Piezocone Sounding (PCPT15) at Amherst	351
Figure C.37	Correlation Coefficient Between Consecutive Cluster Results at Drammen, Norway	352
Figure C.38a	Cluster Analysis of Piezocone Data at Drammen, Norway	353
Figure C.38b	Cluster Analysis of Piezocone Data at Drammen, Norway	354
Figure C.39	Correlation Coefficient Between Consecutive Cluster Results at Gloucester, Ontario.	356
Figure C.40a	Cluster Analysis of Piezocone Data at Gloucester, Ontario	357

Figure C.40b	Cluster Analysis of Piezocone Data at Gloucester, Ontario	358
Figure C.41	Correlation Coefficient Between Consecutive Cluster Results at Hachirogata, Japan	361
Figure C.42a	Cluster Analysis of Piezocone Data at Hachirogata, Japan.	361
Figure C.42b	Cluster Analysis of Piezocone Data at Hachirogata, Japan	362
Figure C.43	Correlation Coefficient Between Consecutive Cluster Results at Lilla Mellosa, Sweden	364
Figure C.44a	Cluster Analysis of Piezocone Data at Lilla Mellosa, Sweden.	365
Figure C.44b	Cluster Analysis of Piezocone Data at Lilla Mellosa, Sweden	366
Figure C.45	Correlation Coefficient Between Consecutive Cluster Results at Recife, Brazil	367
Figure C.46a	Cluster Analysis of Piezocone Data at Recife, Brazil	368
Figure C.46b	Cluster Analysis of Piezocone Data at Recife, Brazil	370
Figure C.47	Correlation Coefficient Between Consecutive Cluster Results at St. Alban, Quebec	371
Figure C.48a	Cluster Analysis of Piezocone Data at St. Alban, Quebec	372
Figure C.48b	Cluster Analysis of Piezocone Data at St. Alban, Quebec	373
Figure C.49	Correlation Coefficient Between Consecutive Cluster Results at Tiller, Norway	375
Figure C.50a	Cluster Analysis of Piezocone Data at Tiller, Norway	376
Figure C.50b	Cluster Analysis of Piezocone Data at Tiller, Norway	377
Figure C.51	Correlation Coefficient Between Consecutive Cluster Results at Troll, North Sea	378
Figure C.52a	Cluster Analysis of Piezocone Data at Troll, North Sea	379
Figure C.52b	Cluster Analysis of Piezocone Data at Troll, North Sea	381
Figure D.1	Piezocone Data at Aiken, South Carolina (Data from Bratton et al., 1993)	384

Figure D.2	Soil Classification Using Piezocone Data at Aiken, South Carolina (Data from Bratton et al., 1993)	385
Figure D.3	Correlation Coefficient Between Consecutive Cluster Results at Aiken, South Carolina	385
Figure D.4a	Cluster Analysis of Piezocone Data at Aiken, South Carolina	387
Figure D.4b	Cluster Analysis of Piezocone Data at Aiken, South Carolina	388
Figure D.5	Comparison Between Cluster Analysis and Visual Classification at Aiken, South Carolina	389
Figure D.6	Piezocone Sounding at Boston, Massachusetts (Data from Sweeney, B. P. and Kraemer, S. R., 1993)	391
Figure D.7	Soil Classification Using Piezocone Data at Boston, Massachusetts (Data from Sweeney, B. P. and Kraemer, S. R., 1993)	392
Figure D.8	Correlation Coefficient Between Consecutive Cluster Results at Boston, Massachusetts	392
Figure D.9a	Cluster Analysis of Piezocone Data at Boston, Massachusetts	393
Figure D.9b	Cluster Analysis of Piezocone Data at Boston, Massachusetts	394
Figure D.10	Comparison Between Cluster Analysis, Visual Classification and Overconsolidation ratio at Boston, Massachusetts	395
Figure D.11	Piezocone Data at Kagoshima, Japan from Takesue et al. (1995)	397
Figure D.12	Soil Classification Using Piezocone Data at Kagoshima, Japan(Takesue et al., 1995)	399
Figure D.13	Correlation Coefficient Between Consecutive Cluster Results at Kagoshima, Japan	399
Figure D.14a	Cluster Analysis of Piezocone Data at Kagoshima, Japan. Cluster Analysis of Piezocone Data at Kagoshima, Japan	400
Figure D.14b	Cluster Analysis of Piezocone Data at Kagoshima, Japan. Cluster Analysis of Piezocone Data at Kagoshima, Japan	401
Figure D.15	Comparison Between Cluster Analysis, Visual Classification Soil Contents and D50 at Kagoshima, Japan	402

Figure D.16	Piezoecone Sounding at Newport News, Virginia(Data from Mayne, 1989)	405
Figure D.17	Soil Classification Using Piezocone Data at Newport News, Virginia (Data from Mayne, 1989)	406
Figure D.18	Correlation Coefficient Between Consecutive Cluster Results at Newport News, Virginia	406
Figure D.19a	Cluster Analysis of Piezocone Data at Newport News, Virginia	407
Figure D.19b	Cluster Analysis of Piezocone Data at Newport News, Virginia	408
Figure D.20	Comparison Between Cluster Analysis, Visual Classification and Fines Content at Newport News, Virginia	409
Figure D.21	Piezoecone Sounding at Onsoy, Norway (Data from Gillespie et al., 1985).	411
Figure D.22	Soil Classification Using Piezocone Data at Onsoy, Norway (Data from Gillespie et al., 1985)	412
Figure D.23	Correlation Coefficient Between Consecutive Cluster Results at Onsoy, Norway	412
Figure D.24a	Cluster Analysis of Piezocone Data at Onsoy, Norway	413
Figure D.24b	Cluster Analysis of Piezocone Data at Onsoy, Norway	415
Figure D.25	Comparison Between Cluster Analysis, Visual Classification, and Undrained Shear Strength at Onsoy, Norway	416
Figure D.26	Piezoecone Sounding at Opelika, Alabama(Data from this Study)	417
Figure D.27	Soil Classification Using Piezocone Data at Opelika, Alabama(Data from this Study)	419
Figure D.28	Correlation Coefficient Between Consecutive Cluster Results at Opelika, Alabama	419
Figure D.29a	Cluster Analysis of Piezocone Data at Opelika, Alabama	420
Figure D.29b	Cluster Analysis of Piezocone Data at Opelika, Alabama	421
Figure D.30	Comparison Between Cluster Analysis, Visual Classification and Soils Content at Opelika, Alabama	422
Figure D.31	Piezoecone Sounding at Po River, Italy (Data from Bruzzi et al., 1986)	424

Figure D.32	Soil Classification Using Piezocone Data at Po River, Italy(Bruzzi et al., 1986)	426
Figure D.33	Correlation Coefficient Between Consecutive Cluster Results at Po River, Italy	426
Figure D.34a	Cluster Analysis of Piezocone Data at Po River, Italy	427
Figure D.34b	Cluster Analysis of Piezocone Data at Po River, Italy	428
Figure D.35	Comparison Between Cluster Analysis, Visual Classification and Fines Content at Po River, Italy	429
Figure D.36	Piezocone Sounding at Penuelas, Puerto Rico (Data from Hegazy and Mayne, 1996)	431
Figure D.37	Soil Classification Using Piezocone Data at Penuelas, Puerto Rico (Data from Hegazy and Mayne, 1996)	432
Figure D.38	Correlation Coefficient Between Consecutive Cluster Results at Penuelas, Puerto Rico	432
Figure D.39a	Cluster Analysis of Piezocone Data at Penuelas, Puerto Rico	434
Figure D.39b	Cluster Analysis of Piezocone Data at Penuelas, Puerto Rico	435
Figure D.40	Comparison Between Cluster Analysis, Visual Classification at Penuelas, Puerto Rico	436
Figure D.41	Piezocone Sounding at Maskinonge, Quebec(Data from Demers et al., 1993)	437
Figure D.42	Soil Classification Using Piezocone Data at Maskinonge, Quebec (Data from Demers et al., 1993)	438
Figure D.43	Correlation Coefficient Between Consecutive Cluster Results at Maskinonge, Quebec	438
Figure D.44a	Cluster Analysis of Piezocone Data at Maskinonge, Quebec	439
Figure D.44b	Cluster Analysis of Piezocone Data at Maskinonge, Quebec	441
Figure D.45	Comparison Between Cluster Analysis, Visual Classification and Water Content % at Maskinonge, Quebec	442

Figure D.46	Piezocone Data at Taranto, Italy (Data from Battaglio et al., 1986)	444
Figure D.47	Soil Classification at Taranto, Italy Using Piezocone Data from Battaglio et al. (1986)	445
Figure D.48	Correlation Coefficient Between Consecutive Cluster Results at Taranto, Italy	445
Figure D.49a	Cluster Analysis of Piezocone Data at Taranto, Italy	446
Figure D.49b	Cluster Analysis of Piezocone Data at Taranto, Italy	447
Figure D.50	Comparison Between Cluster Analysis, Visual Classification at Taranto, Italy	448

ABSTRACT

Soil stratigraphy includes the demarcation of different soil strata, boundaries, lenses, and transitions and is a fundamental first step in geotechnical site investigation. Soil stratification is accomplished using laboratory and in-situ testing methods, such as the piezocone penetration test. Piezocone testing provides three separate readings including tip resistance (q_t), sleeve friction (f_s), and pore pressure (u_b), and has advantages in site investigation because it is fast, economical, and data are collected in the vertical direction every 1 to 5 cm. The data are functions of both soil type and behavior. One of the primary applications of piezocone data is to provide preliminary soil identifications based on visual examination of different trends in cone data profiles and empirical cone-classification charts. However, the simple visual method is subjective and non repeatable, and in some geological conditions, both methods fail to properly indicate major changes in soil type and/or behavior.

A statistical method termed *cluster analysis*, is introduced in this study to: (1) objectively define similar groups in the soil profile, (2) delineate different layer boundaries, (3) identify the lenses and outliers within a sublayer. An interpretation criterion is proposed to define subsurface stratification based on piezocone data in terms of normalized cone parameters: $Q = (q_t - \sigma_{vo}) / \sigma_{vo}'$ and $B_q = (u_b - u_o) / (q_t - \sigma_{vo})$. Piezocone readings are recommended to be standardized using a zscore procedure. For example,

zscore of $q_t = (q_t \text{ at a certain depth} - \text{average value of } q_t \text{ readings}) / \text{standard deviation of } q_t \text{ readings}$. The similarity between different attributes is determined using a cosine coefficient, and the data are grouped using a hierarchical single link (nearest neighbor) technique. A single-cosine-zscore clustering is applied to piezocone parameters from 25 case studies representing different soil types and geological settings and the results are independently verified by available reference in-situ and laboratory data. Cluster analysis is able to detect drastic changes within the stratigraphy not evidenced by visual examination of the unprocessed or processed data, or other cone data interpretation techniques.

CHAPTER 1

INTRODUCTION

Site characterization is a fundamental step in geotechnical engineering that is complex due to the natural makeup, and inherent variability of the soil. The purpose of a site investigation is usually to define the following: subsurface stratigraphy, soil types, and properties (behavior). Laboratory and in-situ testing are used to determine the components of a site investigation. Soil stratigraphy is the focus of this research and indicates the depths of soil boundaries, the number of soil types, and the presence of lenses, or transitions between layers and outliers. An outliers is defined as a soil measurement that do not belong to a specific soil layer and can be different from the whole data. Traditionally, geotechnical site investigations have been accomplished using rotary drilling, augering, and soil sampling practices. While these are important in a site investigation program, they are slow and expensive relative to in-situ methods.

A relatively newer direct-push technology than boreholes, such as the cone penetrometer, offers faster, near continuous, and economical information about the in-situ subsurface stratification. Piezocone testing is now routinely performed as part of site investigation programs because it provides as many as three separate readings, and screens the soil in the vertical direction at short intervals usually between 1 cm and 5 cm. Piezocone data are functions of both soil type and soil behavior.

In current practice, piezocone results are visually examined to interpret the layering in the subsurface profile. The accuracy of this interpretation depends on user experience or by referencing to available empirical classification charts. The simple naked-eye inspection is not always satisfactory because it may be subjective and non-repeatable. Alternatively, empirical classification schemes are used to delineate different soil strata, however they sometimes give erroneous indications of the soil stratigraphies. In some soil profiles, there are drastic changes such as plastic clay over lean clay, and insensitive clay over quick clay. These variations are neither detected by visual examination of unprocessed piezocone data nor by available classification charts. Univariate and multivariate statistical methods (e. g., intraclass correlation coefficient and generalized distance) have also been tried for demarcation of soil boundaries (Hegazy et al., 1996). These too, however, are unable to delineate dramatic soil variations from subtleties in certain piezocone data, particularly clay deposits.

In this research program, a new and powerful statistical method termed *cluster analysis* is applied for analyzing piezocone test data for stratigraphic profiling of geomaterials. Cluster analysis has been previously utilized in the areas of medicine and biology, for example, for the purpose of classifying different plants and animals in similar groups, and adapted in this study for applications to geotechnical data. The developed clustering criterion is applied to piezocone data from 25 sites having different soil types and geological conditions worldwide. These sites are predominantly clay deposits because collecting undisturbed samples in fine grained soils for the purpose of laboratory testing is

more feasible than in the case of sandy materials. Therefore, the behavior of clay soils is better understood than that of the sand soils.

The clustering method is able to detect the inherent correlation between the independent piezocone test measurements and give an objective indication of a soil stratigraphy. The cluster results are verified using backup laboratory and in-situ data at the studied sites. In some surprising examples, the cluster analysis is shown to detect vast distinctions between adjacent soil layers where no apparent changes are evident by visual examination of the cone soundings or by routine processing of cone data. The method can serve to scan for potentially dangerous situations or unusual soils, such as cemented zones, sensitive clays, quick clay deposits, and the like.

An overview of the structure of the thesis is discussed herein. A review of available methods for site characterization and their limitations is given in Chapter 2 with emphasize on piezocone data and cone classification charts. Statistical techniques such as the intraclass correlation coefficient and the generalized distance have been used previously to delineate different soil layer boundaries. A review of the applications of these methods using piezocone data is discussed in Chapter 3. As an alternative and improvement to conventional statistical approaches, *cluster analysis* is introduced in Chapter 4 for applications involving geotechnical site characterization. Different cluster methods are evaluated. Components of cluster analysis are explained and an interpretation criterion of clustering piezocone data is developed.

In this study, clustering is evaluated at 25 sites for which a representative piezocone data at each site are divided into numbers of groups between 2 and 100. At

each site, the statistical subsurface stratigraphy obtained using the cluster technique is validated by comparing the obtained primary and secondary layers, and soil lenses and outliers with reference laboratory and field tests. Different cluster applications including simple examples and a discussion of several factors affecting cluster results are presented in Chapter 5. The verification of clustering as a powerful statistical tool to detect subtle changes in piezocone data where dramatic changes in soil behavior is given in Chapter 6.

In Appendix A, a summary is given of the piezocone tests performed by the author during the term of this study and related field experience. Piezocone data are analyzed at three sites having different soil types and geological settings, and the results of different clustering techniques are discussed in detail in Appendix B. The growth of cluster analyses at 12 sites up to cluster number 100 are summarized in Appendix C. Finally, clustering assessment of piezocone data at 10 different case studies is evaluated in Appendix D.

CHAPTER 2

SOIL STRATIGRAPHY

2.1. Synopsis

Soil stratigraphy is a primary step in geotechnical engineering to identify different soil strata, their boundaries, and soil lenses and transitions. Soil stratigraphy is usually obtained using boreholes, soil sampling and laboratory testing, and/or in-situ testing. In this chapter, an overview of borehole and laboratory methods for soil identification is discussed. However these methods are slow, expensive and in some soils such as clean sands, high quality sampling can be very difficult. Piezocone testing for the purpose of geostratigraphy is the focus of this research because it is fast and economical, provides near continuous data every 1 to 5 cm, and can be performed in different geological conditions including clean sands and quick clays where soil sampling is usually difficult. Soil stratigraphy can be obtained based on piezocone data using two methods including: (1) a simple visual method when different trends of piezocone data with depth indicate different soil layers, and (2) empirical soil classification charts. The simple visual method is subjective and not repeatable and the cone classification charts can not properly indicate the subsurface stratigraphy in some geological settings. These methods are reviewed in this chapter with a discussion of their limitations.

2.2. Soil Stratification Based on Borehole Sampling and Laboratory Testing

Site investigation usually includes the following steps: visual inspection of the site and its topography, reviewing the surficial geology at the site, and finally, direct investigation of the subsurface conditions. In most instances, the latter step is achieved by drilling boreholes and collecting both push and drive samples to discern the various subsurface strata. Soil types and properties are defined both by visual examination of the samples and performing laboratory testing.

Boreholes are usually made by wash, rotary, or auger drilling. Operation details and differences between different drilling methods are discussed, for example, by Terzaghi et al. (1996). Disturbed samples can be obtained using a split spoon sampler every 1.5 m during the advance of a borehole. These soil samples can be used in laboratory index tests, for instance, to determine the grain size distribution, water contents, and other soil index properties. Undisturbed soil samples may be obtained using thin-walled tube samplers (Shelby tube). These samples are commonly used to perform laboratory consolidation and strength tests. More details of soil sampling are given, for example, in American Society of Testing and Materials (ASTM) guide D-4700, and by McGuffey et al. (1996).

The steps of a typical site investigation program based on borehole sampling and laboratory testing are shown in Fig. 2.1. The approximate time needed to complete each component in the program is also given. After field drilling operations, soil samples are transported to the laboratory for testing to determine their physical and mechanical properties. Soil type is the first goal in a geotechnical investigation and can be defined by

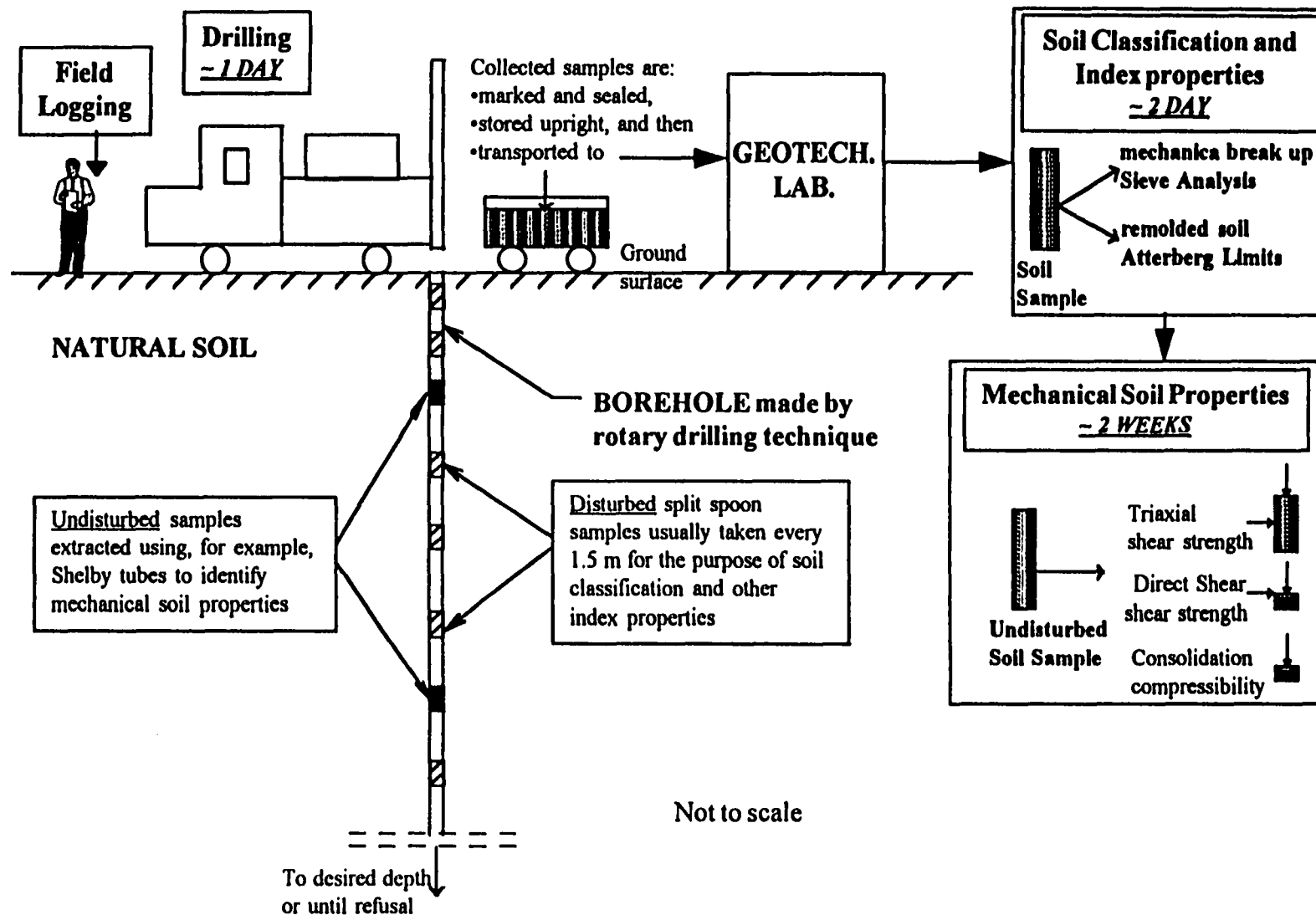


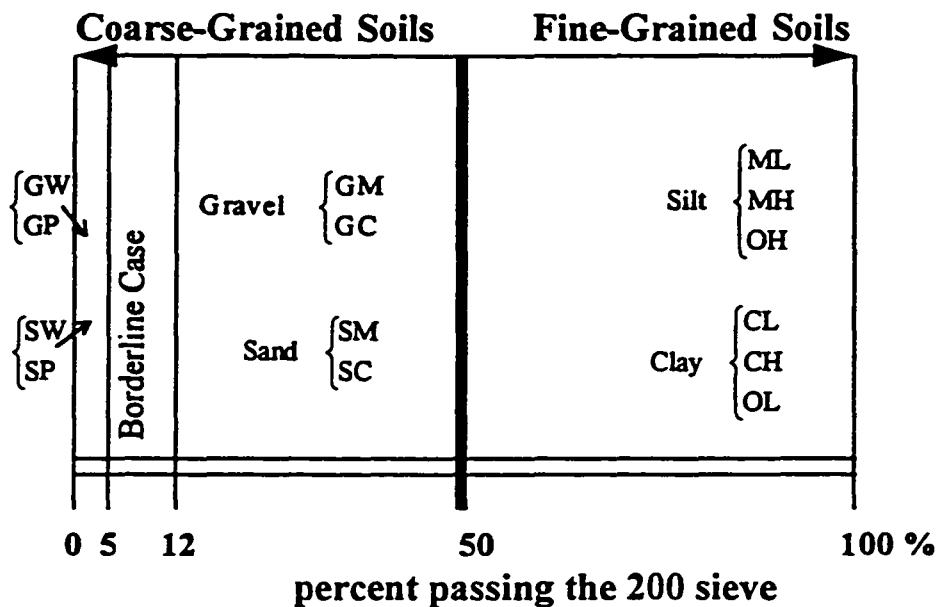
Figure 2.1. An Example of A Routine Site Investigation Program Based on Borehole Sounding and Laboratory Tests.

examining the soils by eye, feeling their texture by hand, smelling their odor (if any), and performing simple tests such as consistency. Guidelines for visual soil description can be found in ASTM guide D-2488. Then, sieve analyses are performed to obtain quantitative information about the percentages of different coarse and fine materials in a soil sample.

The Unified Soil Classification System (USCS), as per ASTM guide D-2487, is the most commonly used approach in current geotechnical practice to define soil types in the laboratory. The method is based on two primary indices which include: (1) grain size distribution, and (2) plasticity tests (known as Atterberg limits). Casagrande (1948) originally developed the USCS and indicated that the method uses borderlines to give a qualitative definition of the soil type and should be accompanied with other means of tests to quantify the engineering soil properties at the expected field conditions.

According to USCS criterion, the major delineation between fine-grained versus coarse-grained soils is defined at a U.S. sieve size number 200, corresponding to, a particle size of 0.075 mm. Figure 2.2 shows a simplified diagram of the classification scheme. Soil particles greater than 0.075 mm are defined as coarse materials (sands and gravels), otherwise, they are known as fines (silt and clay sizes). Coarse materials are divided into three subgroups based on their fines content. For instance, a soil specimen is classified as silty sand (SM) if its fines content is equal to 12 percent to 49 percent, although, this is a major disadvantage of the USCS scheme as explained herein.

Suppose two adjacent soil specimens where the first sample has 51 percent sands and 49 silts, and the second sample has 49 percent sands and 51 percent silts. The former is classified as silty sand (SM) and the latter is classified as sandy silt (ML), however, their



Soil Type: G = Gravel S = Sand M = Silt
C = Clay O = Organics

Soil Gradation: determined on dis-aggregated specimen forced through nested sieves
W = Well Graded P = Poorly Graded

Plasticity: determined on remolded specimens
H = High Plasticity L = Low Plasticity

Figure 2.2. Borderlines between Different Soil Types Using the Unified Soil Classification System (USCS); Modified after Howard (1977).

behavior is essentially the same which indicates a disadvantage of the USCS procedure. Piedmont residual soils is a typical example in which the USCS gives unreliable indices of the soil behavior (Sowers and Richardson, 1983). Moreover, soil samples are destructured and remolded during the process of laboratory soil classification. As a consequence, several important factors are often neglected (Douglas and Olsen, 1981). For example, the effects of soil fabric, geologic origin, structure, mineralogy, sensitivity, and in-situ state-of-stress are ignored.

A description of soil color, odor, texture, consistency, structure, and geology is combined with the USCS results to develop a preliminary model of the site stratigraphy with approximate soil boundaries based on borehole information and experience of the engineer. On large scale projects, it is possible to perform extensive laboratory tests to evaluate detailed soil properties such as permeability, compressibility, and shear strength. A more reliable soil characterization is obtained based on both soil type and soil behavior. Soil samples always have some degree of disturbance and the field conditions are difficult to be completely restored or compensated in the laboratory (Jamiolkowski et al., 1985). Certain soils are very hard or impossible to sample, such as in the cases of quick clays and clean sands. Recovered soil samples represent only a very small volume of the total geologic mass. Therefore, it is necessary that a large number of high quality samples be tested in the laboratory to obtain an accurate delineation of the soil boundaries. This is not usually feasible in most cases due to high cost and excessive testing times.

2.3. Stratification by In-Situ Tests

In-situ tests, such as the cone penetrometer and flat dilatometer, are relatively quick, and economical. In-situ tests can be performed in cohesionless sands in which undisturbed sampling is not feasible. Site investigation based on in-situ penetration testing is more expedient than using laboratory testing. However, there are a few disadvantages to in-situ testing (Jamiolkowski et al., 1985), including: unknown drainage behavior especially in mixed soils, uncontrolled boundary conditions, full or partial displacement, strain rate effects, and disturbance. As a consequence, most available interpretation methods of in-situ testing are empirical due to these limitations. Herein, the piezocone penetration test is the focus of this study with particular emphasis on its use to detect subsurface stratigraphy, as discussed in the following sections.

2.4. Cone Penetration Testing

Piezocone testing provides the most comprehensive logging of soil in the vertical direction compared with other available in-situ tests (Wroth, 1984). Soil stratigraphy is a primary use of piezocone data, which are affected by both soil type and soil behavior (Lunne et al., 1997). In this section, an introduction is given to different piezocone terminology. Then, available methods to delineate soil profiles from cone testing are evaluated, including: (1) simple visual criteria, and (2) empirical cone classification techniques.

In the standard cone penetration test (CPT), two stress measurements are obtained including: (1) cone tip resistance (q_c), and (2) sleeve friction (f_s). Testing procedure are

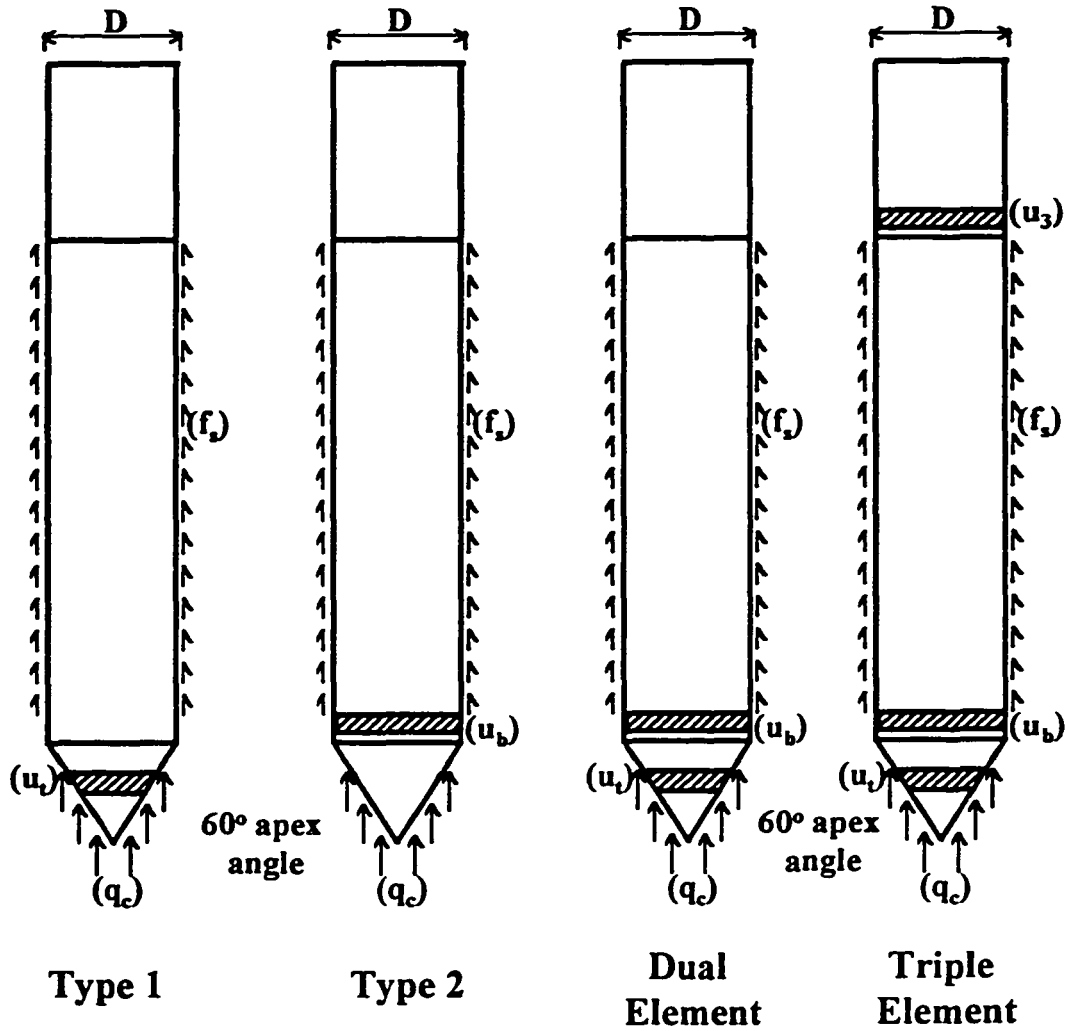
given by ASTM guide D-5778 and Lunne et al. (1997). The cone is designated as piezocone penetrometer test (PCPT) when a third channel, penetration porewater pressure (u), is recorded. The piezocone measurements improve the exploration of minor geological variations within a soil layer such as sand lenses within a clay deposit which may have a major effect on the drainage behavior of the stratum.

A schematic diagram of a piezocone indicating alternative porous element location is shown in Fig. 2.3. The penetration pore pressure has been measured at one or more of three locations: (1) the cone face ($u_1 = u_t$), (2) behind the tip at the shoulder position ($u_2 = u_b$), and/or (3) on the shaft behind the sleeve (u_3) as proposed by Campanella and Robertson (1988).

The water pressure has an effect on both the unprocessed q_c and f_s readings due to the effect of unequal areas of the cross section of the cone at different locations as shown in Fig. 2.4. For all types of cones in practical use, the recorded tip resistance q_c is smaller than the actual intended value and the sleeve friction is larger than the true value (Senneset et al., 1989). Robertson and Campanella (1983) gave the correction of the tip resistance q_c as follows:

$$q_t = q_c + u_b (1 - a_n) \quad \text{(Equation 2.1)}$$

where q_t is the corrected tip resistance, $a_n = A_n/A_c$ is the net area ratio, $A_n = \pi d^2/4$ is the cross-sectional area of the load cell or shaft, and $A_c = \pi D^2/4$ is the projected area of the



Note: Cone diameter, $D = 35.7 \text{ mm}$ (10-cm^2 tip)
 $D = 43.5 \text{ mm}$ (15-cm^2 tip)

- Tip Resistance, q_c
- Sleeve Friction, f_s
- Base Area, $A_b = \pi D^2 / 4$
- Pore Pressure at the Face, $u_1 = u_t$
- Pore Pressure at the Shoulder, $u_2 = u_b$
- Pore Pressure Behind the Sleeve, u_3

Figure 2.3. Locations of Different Elements of a Piezocone.

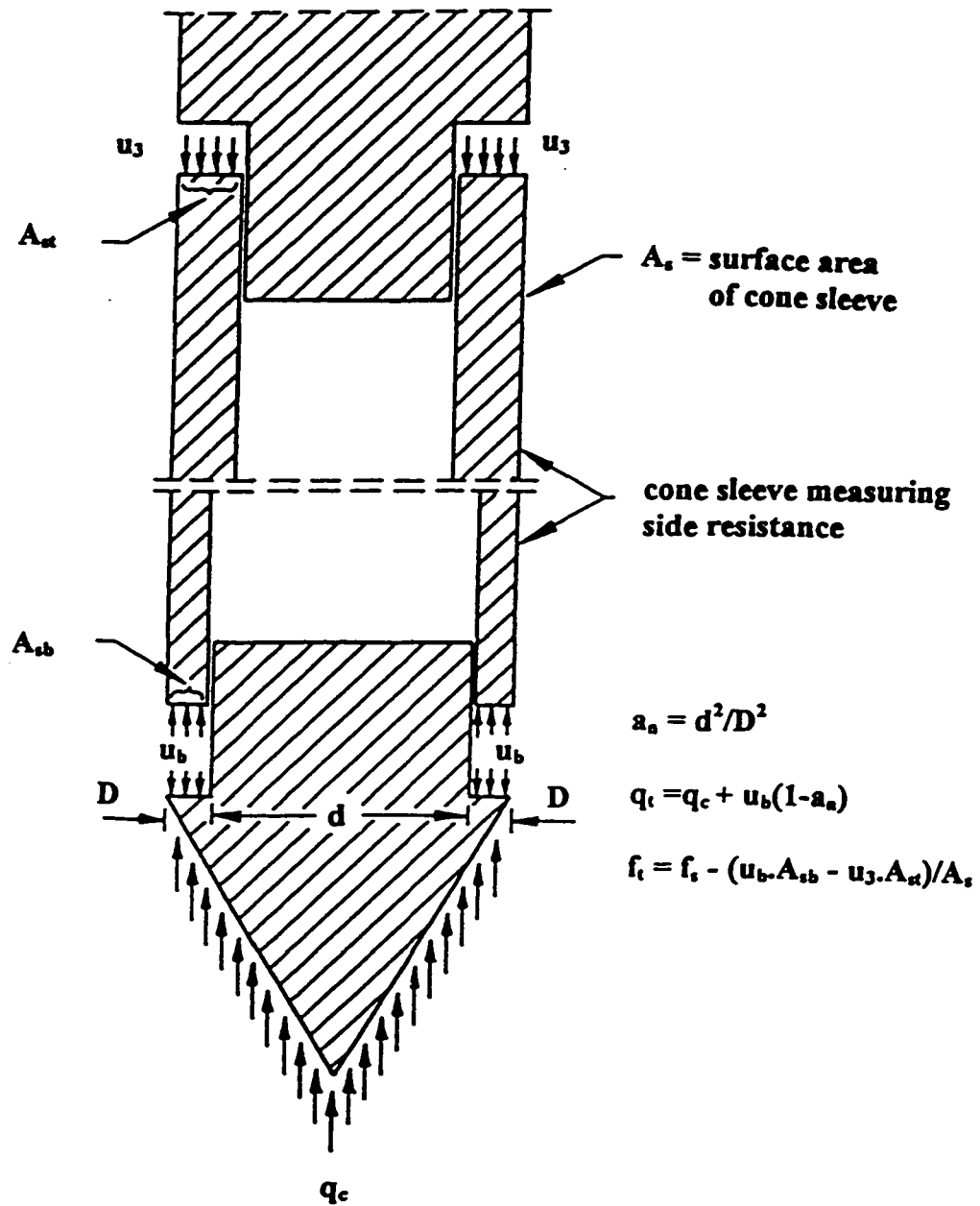


Figure 2.4. Effect of Pore Water Pressure on the Tip Resistance and Sleeve Friction (from Jamiolkowski et al., 1985).

cone as shown in Fig. 2.4. Lunne et al. (1997) gave the correction of the sleeve friction as follows:

$$f_t = f_s - \frac{(u_b \cdot A_{sb} - u_3 \cdot A_{st})}{A_s} \quad (\text{Equation 2.2})$$

where A_{sb} and A_{st} are the cross sectional areas of the friction sleeve at the bottom and the top, respectively, and A_s is the friction sleeve surface area (see Fig. 2.4).

Robertson et al. (1986), experimentally showed that the filter location and soil type affects the magnitude of porewater pressure measurements. Theoretical studies by Levadoux and Baligh (1986) also showed these differences. The maximum pore pressure measurements are on the cone face where there are maximum compression stresses. That is, u_t measurements are always positive. At the shoulder position above the cone and behind the sleeve, there is normal stress relief and large shear stresses. The increase of the normal stresses in saturated soils produces positive pore pressure, however the contribution of the shear stresses in u_b measurements might be positive or negative based on the dilatancy of the soils. Therefore, measured porewater pressures can be positive or negative at the shoulder position (u_b) and behind the sleeve (u_3).

In the case of normally-consolidated to moderately overconsolidated fine grained soils and loose silts and sands, porewater pressures measured on the cone face and shoulder are positive due to contraction of the soil. In the case of heavily overconsolidated clays and dense silts and sands, porewater pressures on the cone face are

also positive due to soil compression. However, excess pore water pressures (penetration pore pressure minus hydrostatic pressure) at the shoulder and shaft positions are negative due to soil dilatancy and/or fissuring (Mayne et al., 1990). In all cases, u_t is greater than u_b which is greater than u_3 . According to Campanella and Robertson (1988) and Lunne et al. (1997), the recommended porous filter position is at the u_b location for these reasons: (1) the filter is less exposed to damage or wear and its compressibility has less effect on the measurements, (2) the shoulder readings are necessary to convert measured tip resistance q_c to corrected resistance q_t , and (3) during a dissipation test, measured pore pressure is less affected by the hydraulic pressure on the rods. However, in the case of highly dilatant soils (where $u_b < 0$), a filter at the face location (i.e. u_t) is better for indicating the soil variation in certain soil stratigraphy and because the $q_c \rightarrow q_t$ correction is not significant for these soils (Mayne et al., 1990). To accommodate all possible cases, Juran and Tumay (1989) recommended the use of a dual-element cone with simultaneous measurements both on the cone face and shoulder.

In this study, u_b data are considered the reference piezocone measurements and therefore used in the statistical analysis. The effect of alternatively using face pore pressure data (u_t) on the clustering of data is explained later for comparative purposes.

2.5. Subsurface Stratification Using Piezocone Data

Piezocone tests are used to demarcate layering in the soil profile in the vertical direction. A subjective soil stratigraphy is thus obtained based on visual inspection of the

piezocone data or through use of empirical CPT classification charts. Both criteria are evaluated herein.

2.5.1. Visual Method

A simple stratigraphic technique is to interpret the variation of a vertical piezocone profile by eye. In order to illustrate this technique, piezocone data at Amherst, Massachusetts (this study) and soil stratigraphy from boreholes (Lally, 1993) are shown on Fig. 2.5. The soil consists of a 2-m silty clay fill, 2-m sandy silt crust, and 1-m varved silty clay underlain by soft normally consolidated varved clay to the termination depth of 14.5 meters. By examining the trends of the vertical profiles of the corrected tip resistance q_t and the sleeve friction f_s , four different groups of data are detected. The visual boundaries are approximately defined at depths of 2.0 m, 4.0 m, and 5.5 m which are in good agreement with the borehole boundaries. The u_b measurements are negative down to 2.6 m and not able to detect the difference between the fill and the crust during the penetration above the groundwater table (GWT). The porous element may have, in fact, become desaturated in this zone which caused a delay in the pore pressure transducer response beneath the groundwater table. A variation of the u_b trends is observed at a depth of 4 m which indicates the boundary between the crust and the varved clay. Therefore, q_t and f_s profiles are able to allocate the variation within a clay profile, but a primary soil boundary is missed using the u_b profile due to desaturation effect. Note also the decays at approximate 1 m intervals during stops in penetration in order to place additional rods.

Another example of visual-based stratigraphic profiling is via the piezocone data at McDonald's Farm, south of Vancouver, British Columbia (Robertson, 1982), as shown on Fig 2.6. The soil profile consists of a 2-m of soft clay and silt layer, 11-m of loose to medium coarse sand, and 2-m of silty sand underlain by 15 m of soft clayey silt extending to the termination depth of the sounding at 30 meters. Four layers are evident looking at the vertical profiles of the corrected tip resistance q_t , sleeve friction f_s , and pore pressure u_b . The soil boundaries are almost at depths of 2 m (between the upper soft clay and silt layer and the sand layer), 13 m which indicates the top of the silty sand layer, and 15 m where the soft lower clay layer starts. The defined soil boundaries are supported by the reference soil stratification as obtained from adjacent soil test borings and shown on Fig. 2.6. Therefore, the piezocone data are able to detect successfully the differences between sands, clays and mixed soils, as well as the layering sequence and depths of their interfaces.

In some cases, the piezocone data have only subtle changes, whereas actually dramatic changes occur within the soil profile. For example, Fig. 2.7. shows representative piezocone data and borehole stratigraphy from Masood et al. (1990) at Drammen, Norway (Lacasse and Lunne, 1982). The soil deposit consists of silty clay between the depths of 4 m and 5 m and plastic clay between the depths of 5 m and 10 m underlain by a lean clay down to a depth of 16 m. By visually inspecting the vertical profiles of the tip resistance q_t and the sleeve friction f_s , one possible interpretation is made by the author as follows: a soil boundary located at 12 m where there is a slight decrease of both readings. Looking at the pore water pressure u_b , a possible soil boundary is

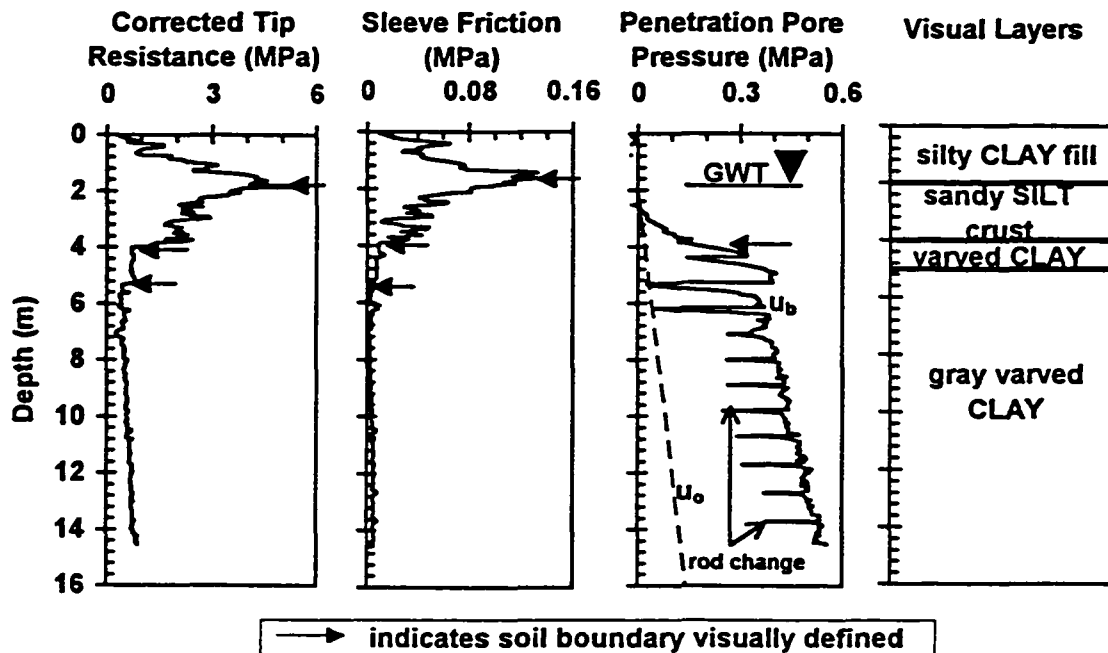


Figure 2.5. Piezocone Data and Soil Stratigraphy at Amherst, MA (Data by Author During This Study).

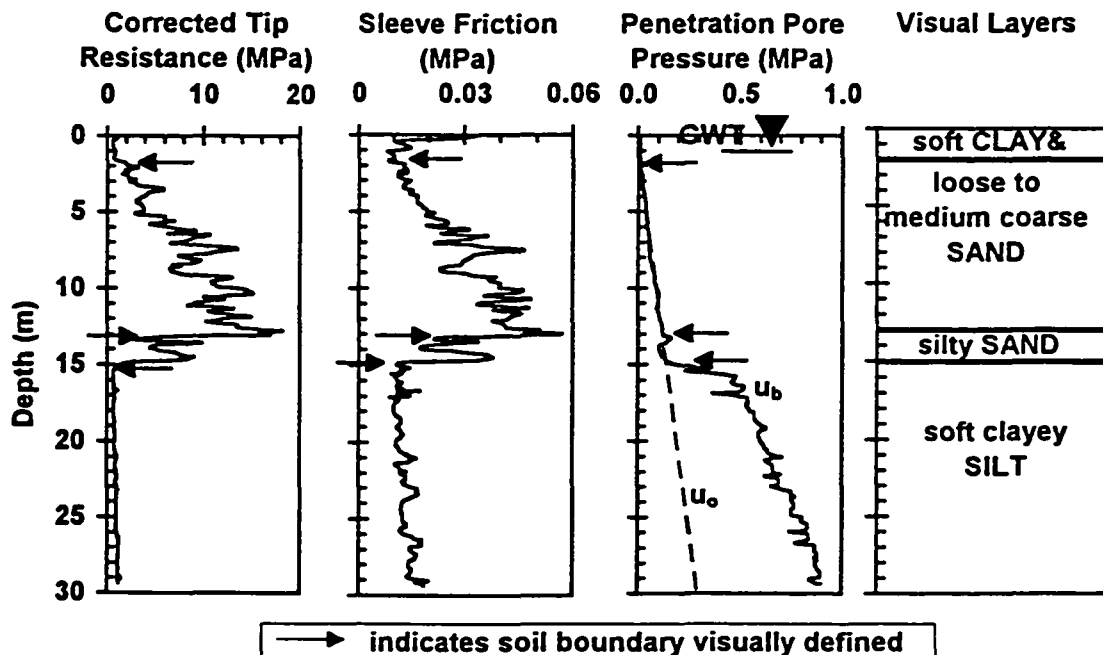
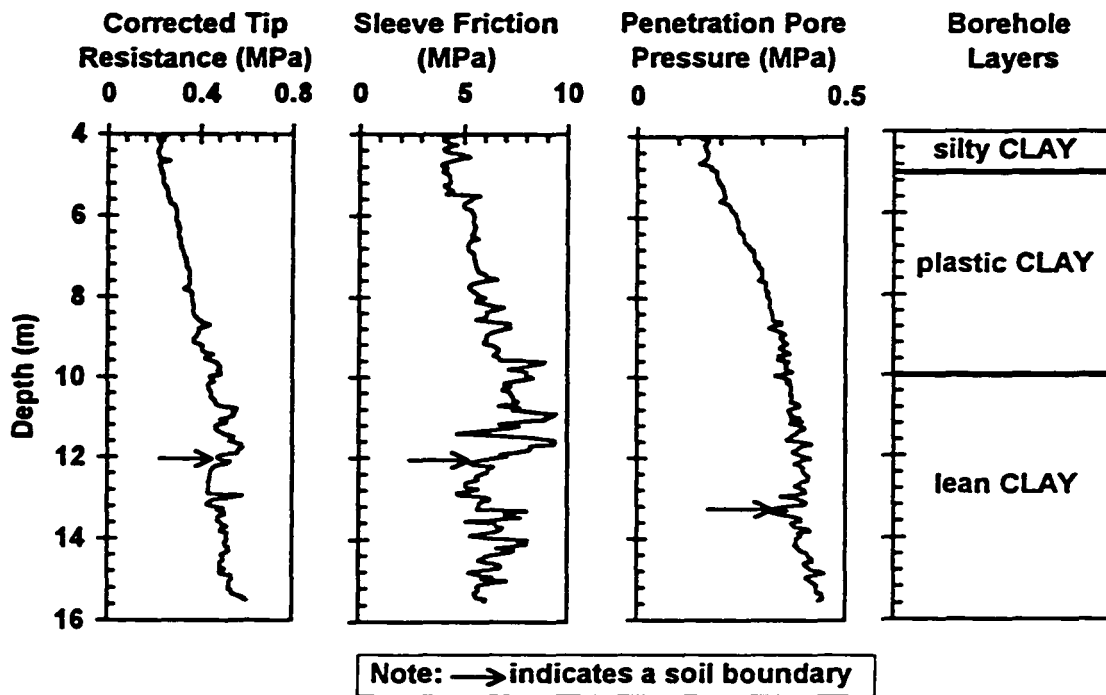


Figure 2.6. Piezocone Data and Soil Stratigraphy at McDonald's Farm, Vancouver (Data from Robertson, 1982).



**Figure 2.7. Piezocone Data and Soil Stratigraphy at Drammen, Norway
(Data from Masood et al., 1990).**

indicated at a depth of 13 m. Therefore, the major change in the soil plasticity and strength (discussed later in this report) at a 10-m depth is not seen by naked-eye.

2.5.2. Profiling Stratigraphy Using Cone Classification Charts

Cone measurements have been used to develop empirical classification charts, such as those presented by Begemann (1965) and Searle (1979). Table 2.1 summarizes these methods based on electrical cone data. Unprocessed, partially processed cone data, and normalized parameters are correlated with different soil types. A discussion of their different forms is included herein. The cone tip resistance increases by increasing the size of soil particles from fine grained soils to coarse grained soils. The tip resistance increases

Table 2.1. Summary of Available Soil Classification Methods Using Cone Data.

Used data	Reference	Remarks
q_c and f_s	Fugro symposium (1972) ⁽¹⁾	<ul style="list-style-type: none"> Correlates a soil type with percentage of particles less than 16 microns.
q_c and f_s / q_c	Douglas and Olsen (1981)	<ul style="list-style-type: none"> Directions of increase of fine contents, particle size, void ratio (e_o), liquidity index (LI), and in-situ stress parameter (K_o) are given.
$(q_c - \sigma_{vo}) / \sigma_{vo}$ and $(u_b - u_o) / u_o$	Jones et al. (1981)	<ul style="list-style-type: none"> Based on data in natural soil deposits and mine tailings. σ_{vo} = total vertical stress u_o = hydrostatic pore pressure
$(q_t - \sigma_{vo})$ and $(u_b - u_o)$	Jones and Rust (1982)	<ul style="list-style-type: none"> A description of fine grained soils consistency and coarse soils density is given.
q_c and f_s / q_c	Robertson and Campanella (1983)	<ul style="list-style-type: none"> The method only identifies the soil type but not the soil behavior.
q_t and $B_q = \frac{u_b - u_o}{q_t - \sigma_{vo}}$	Senneset and Janbu (1985)	<ul style="list-style-type: none"> A description of fine grained soils consistency and coarse soils density is given.
q_o , f_s / q_o , and B_q	Robertson et al. (1986)	<ul style="list-style-type: none"> Directions of increase of sensitivity, overconsolidation ratio (OCR), relative density (D_r) and e_o are given.
q_t and B_q	Parez and Fauriel (1988)	<ul style="list-style-type: none"> A description of fine grained soils consistency and coarse soils density is given.
q_c and f_s / q_c	Erwig (1988)	<ul style="list-style-type: none"> The method gives a description of fine grained soils consistency and coarse soils density.
$q_{cl} = \frac{q_c}{(\sigma_{vo}')^n}$ and $FR_1 = \frac{f_s / \sigma_{vo}'}{q_c / (\sigma_{vo}')^n}$	Olsen and Malone (1988)	<ul style="list-style-type: none"> This iterative method gives a guideline of soil plasticity, overconsolidation ratio, metastable condition and sensitivity. where $n = 0.6$ for sands, 0.8 for silts and 1.0 for clays. σ_{vo}' = Effective vertical stress.
q_t and B_q	Senneset et al. (1989)	<ul style="list-style-type: none"> The method gives a description of fine grained soils consistency and coarse soils density.
$Q = \frac{q_t - \sigma_{vo}'}{\sigma_{vo}'}, B_q$ and $F = \frac{f_s}{q_t - \sigma_{vo}}$	Robertson (1990, 1991) ⁽²⁾	<ul style="list-style-type: none"> The method gives an indication of the soil behavior including age, cementation, OCR, ϕ' and sensitivity.
$q_t / (\gamma_w h)$ and $B_p = \frac{(u_b - u_o)}{q_t - \gamma_w h}$	Chang-hou et al. (1990)	<ul style="list-style-type: none"> An indication of OCR and soil sensitivity is given. γ_w = water unit weight, and z = the depth of a q_t reading from the ground surface.
$Q(1 - B_q)$ and F	Jefferies and Davies (1991)	<ul style="list-style-type: none"> This method gives an indication of the soil plasticity, dilation, sensitivity and OCR.
<u>Method A:</u> $q_m = (q_t - \sigma_{vo})$ and B_q <u>Method B:</u> q_m and $(1/B_q - f_s / \sigma_{vo}')$	Larsson and Mulabdic (1991)	<ul style="list-style-type: none"> Method A is for fine grained soils and method B is for all soils. Method A gives an indication of OCR, soil consistency and sensitivity of fine grained soils. Method B gives an indication of the consistency of fine grained soils and density of coarse soils.

⁽¹⁾ As referenced by Delft soil mechanics laboratory (1985); ⁽²⁾ Cited in Wroth (1988) first.

Table 2.1. (Cont'd) Summary of Some Soil Classification Methods Using Cone or Piezocone Data.

Used data	Reference	Remarks
q_t , B_q , and t_{50}	Jian et al. (1992)	<ul style="list-style-type: none"> This method uses information of pore pressure dissipation tests which are usually available at few depths. t_{50} = the time required for 50 % dissipation of the excess pore pressure.
q_{c1} and f_t/q_c	Olsen and Mitchell (1995)	<ul style="list-style-type: none"> This iterative method gives a guideline of soil plasticity, overconsolidation ratio, metastable condition and sensitivity.
Soil index $U = f(x,y)$, and in-situ state index $V = f(x, y)$	Zhang and Tumay (1996a,b)	<ul style="list-style-type: none"> Indicates a significant overlap between adjacent soil types. $x = 0.1539 (f_t/q_c) + 0.8870 \log (q_c) - 3.35$ $y = -0.2957 (f_t/q_c) + 0.4617 \log (q_c) - 0.37$

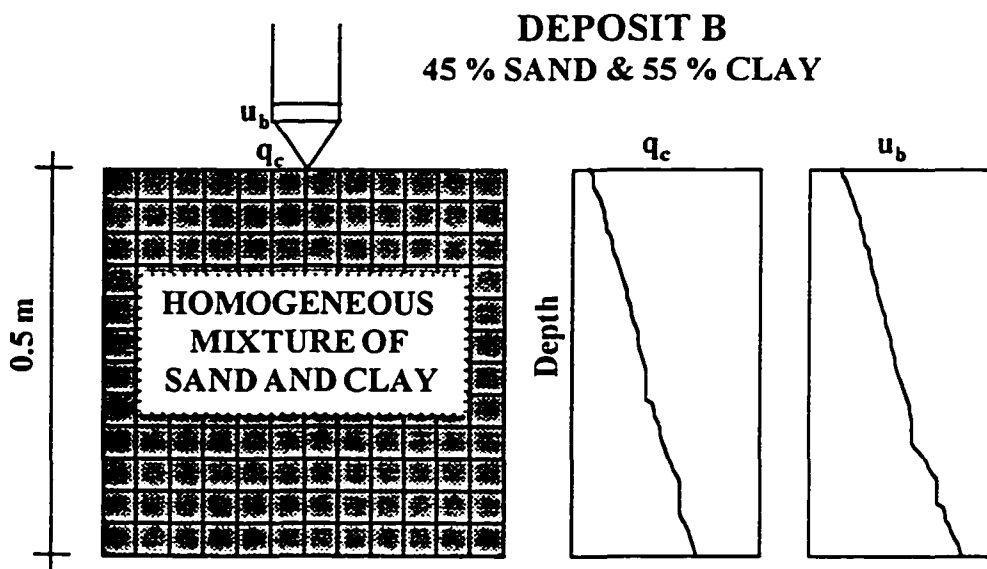
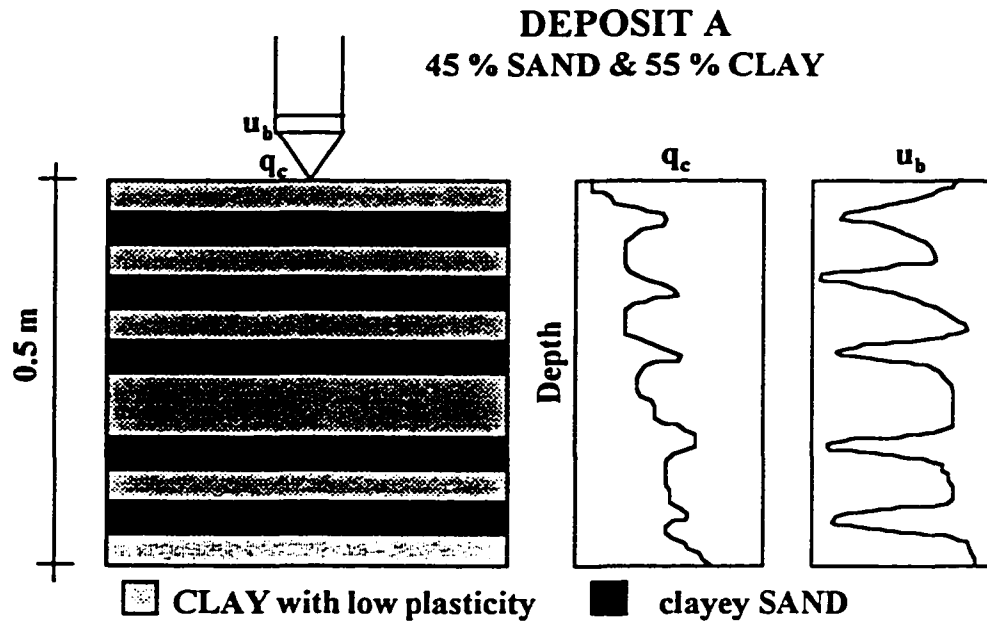
with depth in the same soil layer due to the effect of the total and vertical effective stresses and this explains the scatter plot of the cone data representing the same soil type (Robertson et al. 1986). To reduce this effect, the total vertical stress is subtracted from the tip resistance at the same depth which is termed as net tip resistance $= q_t - \sigma_{vo}$. Wroth (1988) recommended a normalized tip resistance parameter $Q = (q_t - \sigma_{vo}) / \sigma'_{vo}$ as a function of soil shear strength and overconsolidation ratio for the estimation of soil behavior. The derived normalized parameter (Q) reduces the scatter effect (Robertson, 1990). This parameter is also justified using limit equilibrium and cavity expansion theories (Olsen, 1994). Moreover, Olsen (1994) recommended to normalize the tip resistance with respect to the vertical effective stress [$q_{c1} = q_c / (\sigma'_{vo})^n$] based on the stress focus concept which is defined as the intercept of different q_c profiles of different soil types with depth at a very large vertical effective stress. The stress exponent (n) is defined as the slope of the trend of $\log q_c$ profile with respect to \log effective vertical stress. The

stress exponent is approximately equal to 0.6 for sands and 1.0 for clays. However, the stress exponent should be assumed at the beginning and a representative value of a soil type can be obtained by performing 3 to 9 iterations (Olsen and Mitchell, 1995). Also, the tip resistance is not corrected which is significant in the case of identifying normally to moderately overconsolidated clays.

The sleeve friction readings usually follow the trend of q_c readings with depth as shown in Figure 2.6. Therefore, using f_s may not indicate additional information about a soil type identified based on just q_c readings. However, the friction ratio ($FR = f_s/q_c$) is a better parameter to use with q_c measurements because a higher FR indicates a fine material and a lower FR indicates a coarse material. For example at McDonald's Farm test site, the average FR in the upper sand layer is equal to 0.4 percent and the average FR in the lower clay layer is equal to 1.6 percent. Olsen and Mallone (1988) indicated that both the tip resistance and the sleeve friction were functions of the vertical effective stress and recommended to use the ratio of the normalized q_c and f_s [$FR_1 = (f_s/\sigma'_{vo}) / (q_c/(\sigma'_{vo})^n)$] based on the stress focus concept. However, n should be assumed and the tip resistance is not corrected. Wroth (1988) proposed a friction ratio parameter [$F = f_s/(q_t - \sigma'_{vo})$] for soil interpretation based on piezocone data. The excess pore pressure ($u_b - u_o$) is normalized with respect to u_o to reduce the scatter of the data representing the same soil type, due to the increase of u_b with the vertical effective stress (Jones et al., 1981). Wroth (1984, 1988) recommended the normalized pore pressure parameter [$B_q = (u_b - u_o)/(q_t - \sigma'_{vo})$] as a direct function of the overconsolidation ratio to estimate soil behavior based on piezocone

data. Houlsby (1988) proposed to combine Q and B_q in one parameter $[Q(1-B_q)-1 = (q_t - u_b)/\sigma'_{vo}]$ as a direct function of the overconsolidation ratio. Jefferies and Davies (1991) used $[Q(1-B_q)]$ as a cone parameter for identification of both soil type and behavior.

A major advantage of using the cone data to identify the soil as compared with other methods such as USCS is that they are functions of both soil index and behavioral properties under field conditions (Mayne et al., 1995; Lunne et al., 1997). Moreover, the cone data can identify a quick variation of the soil type. For example, two soil layers consisting of identical 45 percent sands and 55 percent clays with two different compositions as follows: (1) the sands are deposited interchangeably with the clays, and (2) the clays and the sands form a homogeneous deposit as depicted in Fig. 2.8. The visual inspection of the two soil samples can help to identify the difference in their soil compositions. Using the USCS indicates identical two deposits of sandy clay. However, the cone measurements can clearly indicate the major differences between the two deposits as follows: in the case of deposit (1), the troughs of the penetration pore pressure identify the locations of the sand seams interbedded within the clay, and in the case of deposit (2), the smoothness of the profiles of both cone records implies the homogeneity of the soil mixture. De Ruiter (1982) noted that the porewater pressure measurements can indicate soil lenses in the order of the thickness of the porous element (approximately 5 mm or 0.1 times a cone diameter). Note also for deposit A that q_c readings have more variability of that of q_c of deposit B which is an indication of the soil variability of the former deposit. However, there are no sharp peaks in the q_c profile of deposit A at the positions of the sand seams because q_c readings are affected by both soils above and below the cone tip. A



For both deposits
 Passing sieve no. 200 = 55% & Retained sieve no. 200 = 45 %
 The two deposits are identical sandy CLAYS using USCS.

Figure 2.8. Comparison of Piezocone Soundings in Two Different and Distinct sandy clay deposits, yet Identical in Terms of USCS.

transition zone of 10 to 20 times the diameter of the cone is required in order to get a full response of q_c at a boundary between two soil strata (Robertson and Campanella, 1989). This facet is discussed in more detail later in this chapter.

Figure 2.9 summarizes different factors affecting the cone measurements, therefore, classification methods based on cone data give an indication of the soil behavior. For example, Villet and Mitchell (1981) conducted 80 cone penetration soundings in sand deposits prepared in chamber tests. The sand used was composed mainly of quartz and feldspar grains. The tip resistance was found to be a function of the void ratio or relative density, the particles size, and the state of stress including the vertical and horizontal effective stresses.

In all classification schemes, empirical boundaries are identified between data groupings of different soil types. For instance, a soil classification chart (Robertson et al., 1986) is shown in Fig. 2.10 in which the tip resistance q_t and the friction ratio FR (f/q_t) and the normalized pore pressure parameter $B_q = (u_b - u_o)/(q_t - \sigma_{vo})$ are used. A subsurface stratification can be obtained by plotting measured cone data on a classification chart. The elevations of the boundaries between different soil layers are approximately equal to the depths of the data at the empirical borders between different soil-type groups. Therefore, a preliminary vertical subsurface profile can be identified including information of the soil-type and behavior. In the next section, limitations of available CPT classification techniques are discussed.

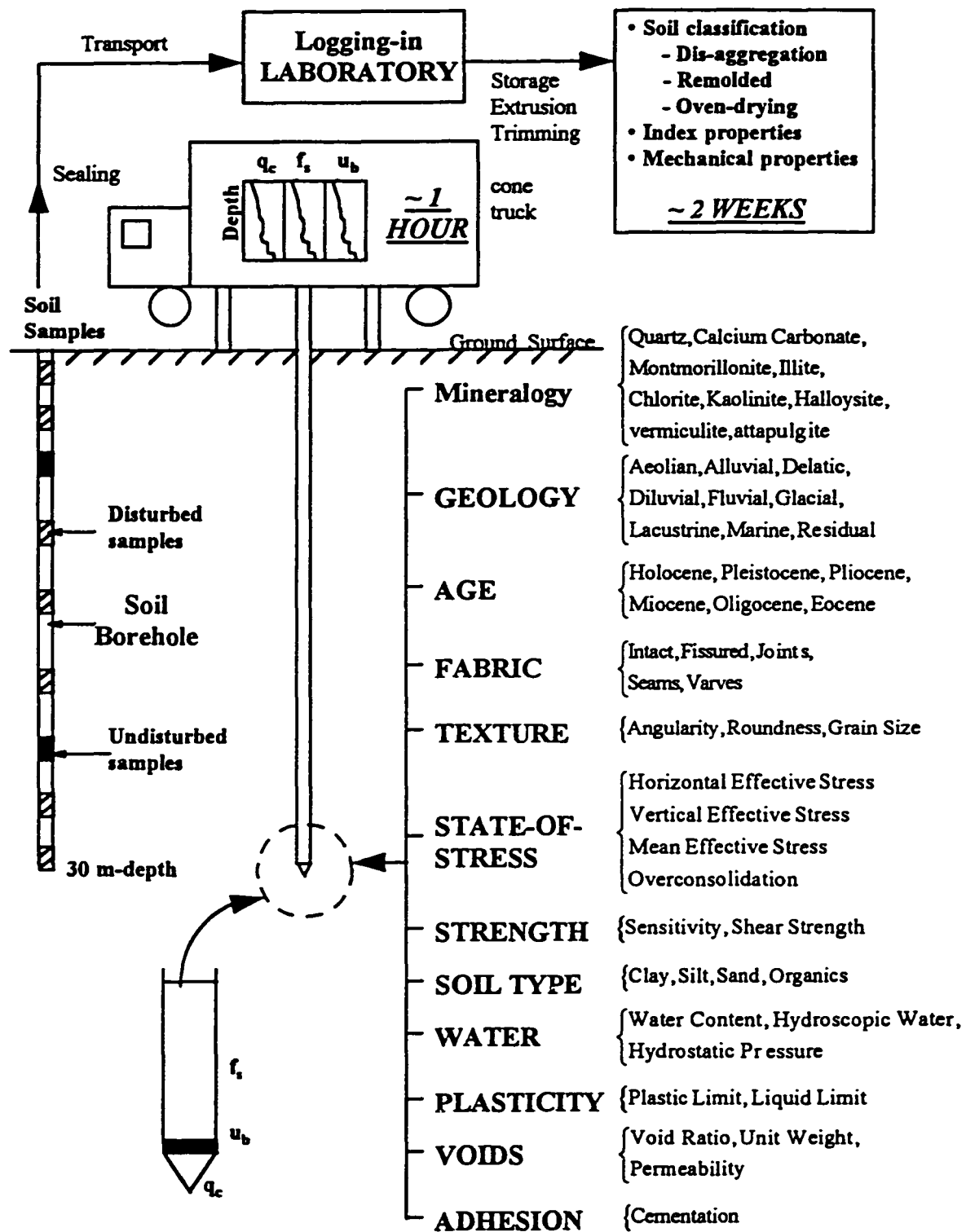


Figure 2.9. Factors Affecting Cone Measurements.

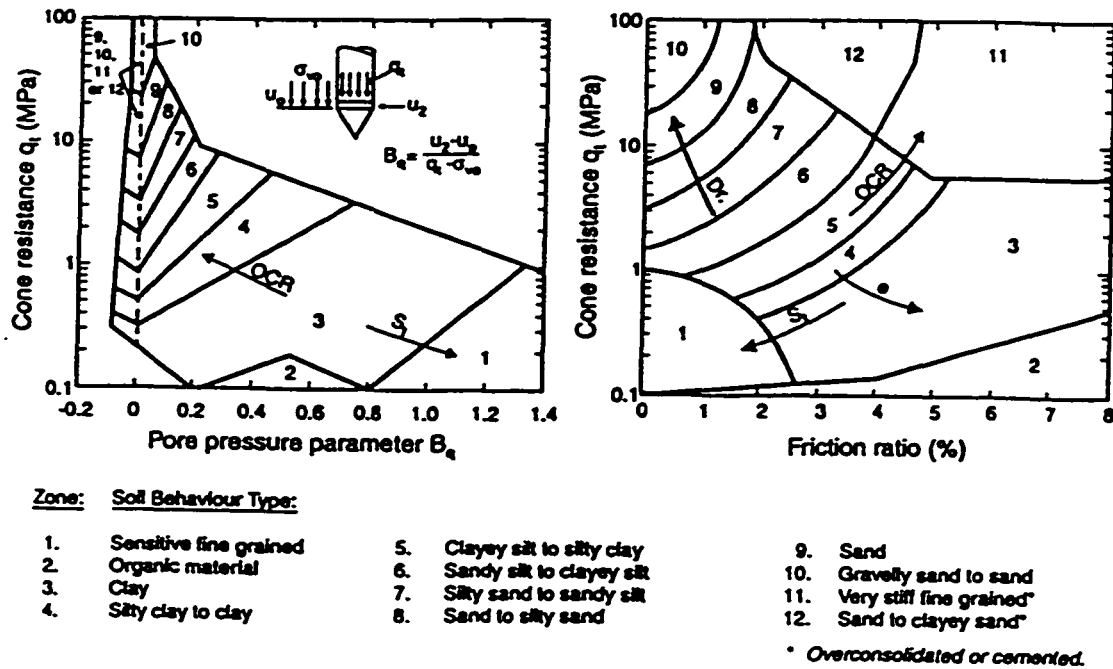


Figure 2.10. One Example of Cone Classification Charts (Robertson et al., 1986).

2.6. Limitations of Classification Methods Using Cone Data

The proposed classification methods in Table 2.1 have limitations due to the following items: limited size of database, different equipment and testing procedures, and soil transitions and lenses neglected in interpretation.

2.6.1. Limitations Due to Size of Database

Each of the soil classification methods are limited to the geological origins of the specific cone database used for their development. For instance, the method by Douglas and Olsen (1981) represents observed soil behavior in the Western United States. The method by Larsson and Mulabdic (1991) represents a summary of soil behavior in Swedish

plastic and organic clay soils. Therefore, there is no one universal method that can be used in all geological conditions. Robertson et al. (1986) recommended that the method be appropriately modified to local experience. Certain soil properties are neglected during the development of the proposed classification methods. For instance, the soil density is not implicitly included in the Robertson (1990) chart therefore a loose sand layer might be indicated as silty sand or sandy silt.

In some early methods, the measured data (specifically, q_c) were used without correction for porewater pressure effects (e.g., Lunne et al., 1986). This error can be significant, especially in normally to moderately overconsolidated clay soils (Robertson and Campanella, 1989). The proposed indirect CPT methods are also limited to a certain depth of a sounding, for instance, the Olsen and Mitchell (1995) method is limited to a depth equal to 40 m, which is a common terminal depth of a cone sounding in the United States. Therefore, these methods should be used only as a preliminary guideline of soil type in a site investigation program (Lunne et al., 1997).

2.6.2. Limitations Due to Equipment and Test Procedure

In cone penetration testing, there are limitations attributed to the design of the cone and its manufacture which affect the repeatability of the data and cause systematic errors. For example, Lunne et al. (1986) quantified the accuracy and repeatability of four piezocone measurements (q , f , u , and u_b) at a relatively homogeneous soft clay deposit at Onsoy test site, Norway. The data were collected using 14 cone types belonging to different organizations. The cone tip resistance and the sleeve friction data were corrected

to account for the thermal and pore pressure effects. The data were corrected using equations 2.1 and 2.2 to compensate the effect of the pore pressure and also the thermal zero shift which changed the output signals at zero load were subtracted from q_t and f_t readings. The variability of the q_t readings were significantly reduced, however, the f_t data were still dispersed and their reliability was questionable. The vertical profiles of penetration porewater pressure indicate better repeatability than that of q_t and f_t at each of the filter positions u_t and u_b . Note that the inherent soil variability at different test locations could also contribute to cone measurements and affect their repeatability.

The coefficient of variation (COV) was calculated for each cone measurement collected using a certain cone at a defined depth interval (Lunne et al., 1986). Then the COVs using the 14 cones were compared at the same depth interval. The coefficient of variation of q_t readings was between 1 and 10 percent. The pore pressure readings using different filter locations indicated more repeatability and their COV was less than 5 percent. However, the corrected sleeve friction readings showed the least repeatability and their COV was as high as 32 percent. A similar evaluation was conducted by Lunne et al. (1986) at Haga clay test site, and Holmen sand test site, Norway. At both sites, the sleeve friction was less reliable and repeatable compared with the tip resistance and pore pressure.

Tanaka (1995) performed a similar study in Japanese soft clay using 8 different cone penetrometers, all available commercially. The readings of q_t and u_b indicated much more repeatability than the f_t readings as depicted in Fig. 2.11. The differences between q_t measurements were attributed to the thermal effect which can vary from a cone to another

and the ambient temperature conditions for different tests. The differences between the u_b measurements were due to the locations of the filter which varied between 7 mm and 31.1 mm above the cone and also due to different thickness and material of the filters. However, very large scatter was observed in f_s readings which were measured very close to the lower level of the capacity of a sleeve friction transducer. Therefore, the fluctuation of a transducer response within its tolerance might have significantly contributed to the variability of the sleeve friction measurements. Using sleeve friction f_s readings that depend on the specific commercial cone type adds uncertainty to the soil classification interpretation. Therefore, f_s measurements have been excluded, in general, from the reminder of this study so as to not to affect the objectivity of a defined soil stratigraphy.

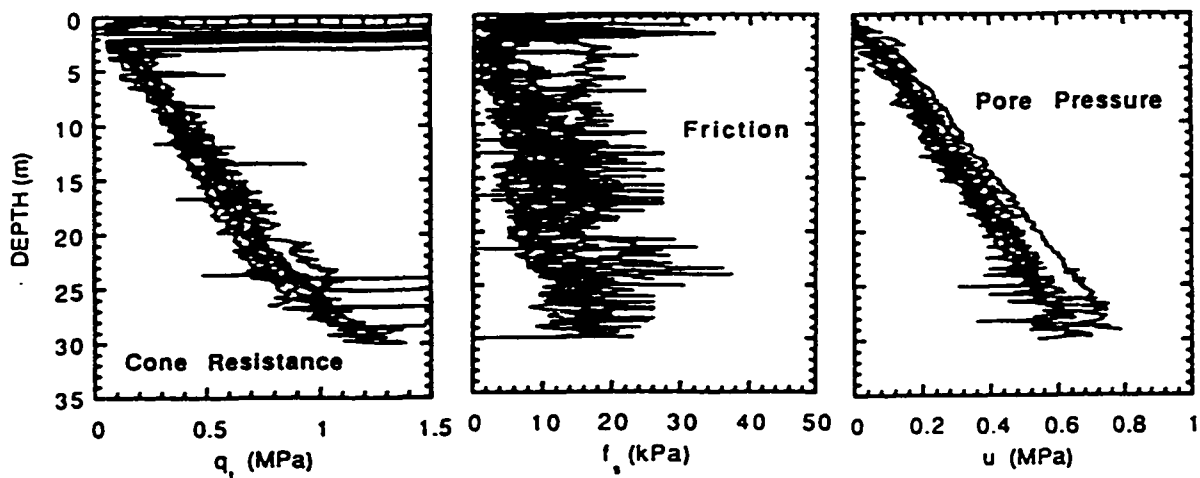


Figure 2.11. Reliability of Piezocone Measurements in Soft Clay Deposit in Japan Using 8 Different Cones (Data from Tanaka, 1995).

Systematic errors affecting the cone measurements are due to three primary factors summarized herein: (1) Small scale variation which means that a geological evidence might be missed if the cone data are measured at relatively large intervals. Therefore the frequency of collecting the data should be taken as high as possible. In most practical applications, the piezocone data are obtained every 1 to 5 cm; (2) Procedural errors due to frequent stops to add cone-rods could result in spikes in the pore pressure profile due to dissipation towards the hydrostatic pressure, although, the readings of q_c and f_t could slightly decrease due to the stress relief on the cone-rods. These errors should be carefully removed from the data because a true geological evidence might be missed; (3) Electrical noise depends on the accuracy of the load cells used in the cone to measure different soil stresses and the location of magnifying the data whether inside the cone or at the ground surface. The effect of the electrical noise can be reduced by: (a) Using high resolution load cells with minimum fluctuation ranges, and (b) amplifying the cone records inside the cone especially in the cases of deep soundings, and penetration in soft clays; and (4) Measurement error sources were summarized by Post and Nebbeling (1995) as follows: penetration depth and verticality, calibration uncertainty, geometrical variation, errors in load transfer, eccentric loading, and temperature. The effects of the measurement errors can be minimized by taking extra care of the cone operation and maintenance as summarized by Lunne et al. (1997).

In piezocone tests, the filter material can affect the pore water pressure measurements due to squeezing, clogging or abrasion (Robertson and Campanella, 1989). The chosen filter should have an adequate permeability. Its permeability should be high

enough to allow a fast response of the pore pressure transducer. It also should be low enough to resist clogging due to the entry of soil grains and air, and maintain saturation. The filter can be made of sintered bronze, carborundum, ceramic, plastic, teflon, or stainless steel (Burns, 1997). For example, hard plastic elements are usually suitable for use in coarse soils and show minor wear. However they can be squeezed in case of dense soils especially at the apex location and their compression can generate a high positive pore pressure (Campanella and Robertson, 1988). A ceramic filter is more brittle than the plastic and stainless steel filters and can be damaged during penetration in dense sands. The stainless steel elements can be clogged due to abrasion during penetration of a dense sand layer. Plastic elements were used by the author during the term of this study and they gave satisfactory results in different soil types including clays, silts, sands, and soil mixtures. The filters should be fully saturated to obtain reliable pore water pressure readings. Water, silicon, or glycerin can be used to saturate the filters (Campanella and Robertson, 1988).

The cone size has some effect on the collected piezocone data. Two sizes including 10-cm² and 15-cm² cones are now common in geotechnical practice (Lunne et al., 1997) and permitted by ASTM guide D-5778. The scale effect of the cone dimensions on q_c , f_s , and/or u_b measurements has been studied previously. De Ruiter (1982) and Lima and Tumay (1991) suggested no significance difference in q_c readings collected using 10- and 15-cm² cones. Juran and Tumay (1989) found no significant difference in q_c and u_b but indicated a possible 20 % increase in f_s by using a 10-cm² base cone instead of 15-cm² base cone. Moreover, Hegazy et al. (1996) confirmed that the cone size has a negligible

influence on the spatial analysis of piezocone data using geostatistical analysis. For example, two cone soundings performed by the author in Piedmont residual soils in Atlanta, Georgia are shown in Fig. 2.12. The two soundings were separated by a distance of 2 meters. The data were collected using a Hogentogler electronic 10-cm² cone and Fugro electrical 15-cm² cone. In the case of the electronic cone, the signals were magnified by an amplifier in the cone before they were transferred to the data acquisition system at the ground surface through a cable. In the case of electrical cone, the volt readings were amplified at the ground surface after traveling through a cable between the cone and a data acquisition system. The electronic noise due to carrying the recorded

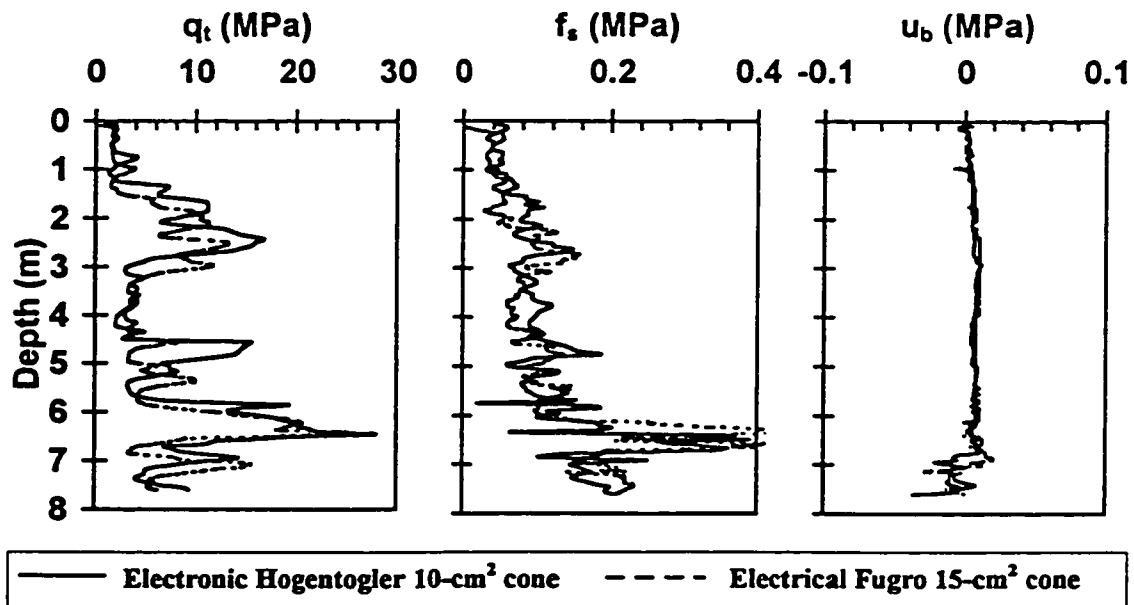


Figure 2.12. A Comparison between Electronic 10-cm² and Electrical 15-cm² Piezocones in Piedmont Residual Silts and Sands in Atlanta Georgia (Data from This Study).

electrical signals through the cable has usually less effect on the measured data in the electronic cones than in the electrical cones especially in soft soils.

The coefficient of variations (COV) are less than 8 percent, 10 percent and 3 percent for q_t , f_t and u_b readings at different depths which indicates a similar variability of both q_t and f_t readings and a relatively better repeatability of pore pressure readings. The ranges of coefficient of variations of q_t and u_b readings are in agreement with those indicated by Lunne et al. (1986) for data collected using 14 different cones at Onsoy, Norway. However, the range of the COV of f_t readings using Hogentogler and Fugro cones is much less than that determined by Lunne et al. (1986) at three Norwegian sites. Sowers and Richardson (1983) indicated a drastic heterogeneity in the Piedmont residual soils which could contribute to the difference between the two cones readings presented herein in addition to cone-type and size.

2.6.3. Soil Transitions and Lenses

Available classification methods for the CPT are unable to categorize soil transitions and lenses. The soil interface above and below the tip affects q_c readings (Mitchell and Brandon, 1998). Treadwell (1976) and Schmertmann (1978a) found that the full response of q_c in a chamber test pushing the cone from softer to stiffer soil can be obtained in a distance equal to 10 to 20 times the diameter of the cone as depicted in Fig. 2.13. However, in the case of moving the cone from stiffer to softer soils, the full response of q_c can be obtained in a distance smaller than 10 times the diameter of the cone as shown in Fig. 2.13.

Campanella and Robertson (1989) noted the scale effects using 10-cm² and 15-cm² cones on q_c readings at a transition zone. Using a 15-cm² cone, the peaks of q_c at stiff lenses were not completely retrieved, but the troughs of q_c at soft lenses were reproduced. Vreugdenhil et al. (1994) used an elastic analysis to quantify the transition zone between two layers. A correction was proposed for the tip resistance measured in a transition zone between two soil layers as shown in Fig. 2.14. The full response of q_c was found to be a function of the ratio of the modulus of the two soils. The greater the stiffness ratio of two layers, the thicker the transition zone between them. Transition zones are generally hard

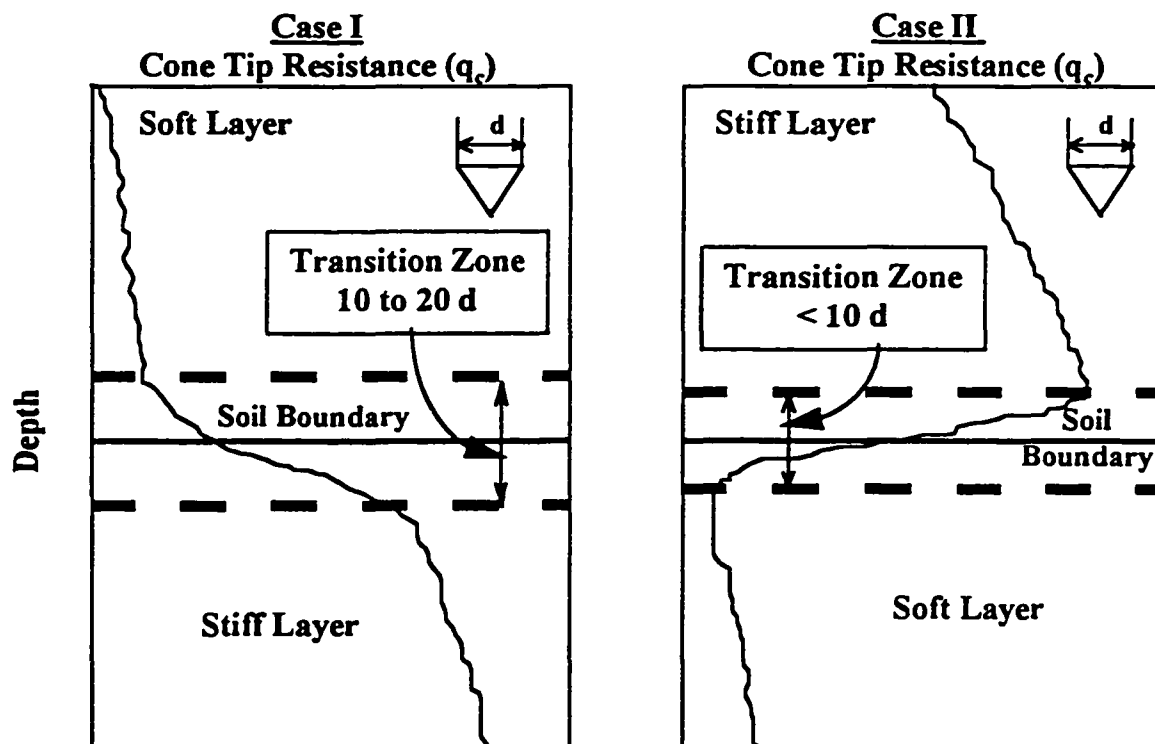


Figure 2.13. Transition Zone of the Cone Tip Resistance Between: (I) Soft Layer above Stiff Layer; (II) Stiff Layer above Soft Layer.

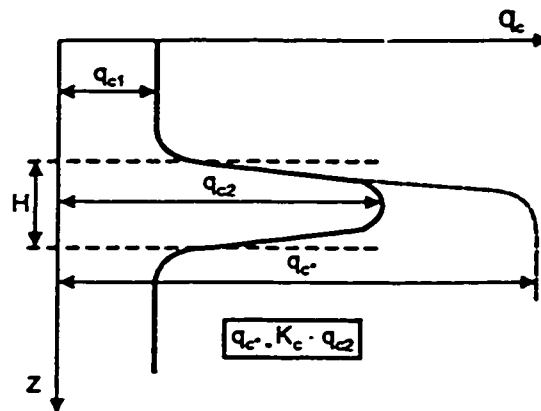
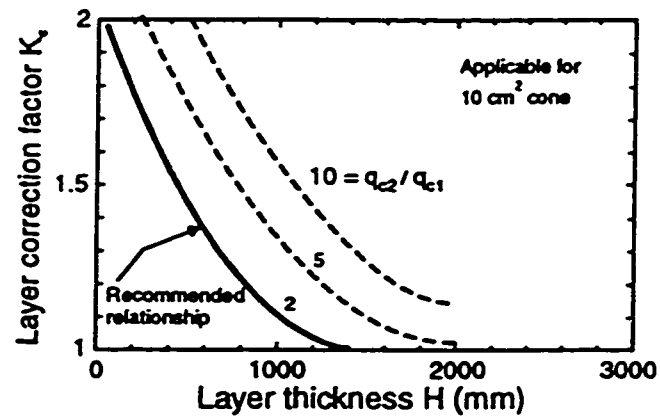


Figure 2.14. Correction of the Tip Resistance in a Transition Zone Between Two Layers (from Lunne et al., 1997).

to quantify because they are based on soil types and stiffness, and cone size. Therefore, they are not considered in the available classification methods.

Juran and Tumay (1989) noted that existing classification methods are based on piezocone data collected in thick layers. Consequently, these methods should be used cautiously to identify thin layers of sands or clays within thick strata of clays or sands, respectively. The difference between the pore pressure measured at two or more positions was suggested to be included to improve the existing classification techniques to avoid the effect of boundary drainage conditions. For example, Juran and Tumay (1989) noted a major difference between pore pressure readings in a thick sand deposit and thin sand inclusions within a clay layer. Moreover, information from dissipation tests at different depths would help also to verify the identity of different soil types. This was also recommended by Senneset et al. (1989), Robertson (1990). Jian et al. (1992) used the time required for 50 percent dissipation of the penetration pore pressure at different soils in their classification scheme.

The most common cone classification methods used in 32 countries worldwide are summarized in Table 2.2 based on their national reports which were published in *Proceedings of the International Symposium on Cone Penetration Tests (CPT '95)*. The classification methods proposed by Robertson et al. (1986) and Robertson (1990,1991) are the most popular in use in current geotechnical practice. The two methods are predominantly utilized in 9 countries including Brazil, Canada, China, Iceland, Japan, Norway, Poland, Singapore, Malaysia, Spain, Switzerland, Turkey, and United States.

Table 2.2. Cone Classification Systems Used in Different Countries Summarized from CPT'95 Proceedings.

Country	Cone classification method(s)	Cone data
Australia	Cone charts based on q_c and FR such as Douglas and Olsen (1981), and Robertson et al. (1983).	q_c and FR*
Brazil	Robertson et al. (1986)	q_t , FR and B_q
Canada	Robertson et al. (1986) Robertson (1990,1991)	q_t , FR and B_q Q , B_q and F
China	Robertson et al. (1986)	q_t , FR and B_q
Denmark	Jorgensen and Denver (1987) Luke (1994)	q_c and FR* q_c and FR*
Egypt	Sabri and Dakhli (1995), mechanical cone	q_c and FR*
Finland	Halkola and Tornqvist (1995)	q_c
France	Parez and Fauriel (1988)	q_t and B_q
Germany	Begemann (1965), mechanical cone	q_c and f_s
Iceland	Robertson (1990,1991)	Q , B_q and F
India	Mahesh and Vikash (1995)	q_c
Lithuania	Furmonavicius and Dagys (1995)	q_c
Malaysia	Robertson (1990,1991)	Q , B_q and F
Republic of Ireland	Douglas and Olsen (1981) Jones and Rust (1982) Robertson et al. (1983) Senneset and Janbu (1985)	q_c and FR* ($q_t - \sigma_{vo}$) and ($u_b - u_o$) q_c and FR* q_t and B_q
Italy	Direct inspection of the pore pressure measurements, in addition to cone charts based on q_c and FR such as Douglas and Olsen (1981), and Robertson et al. (1983).	q_c , FR* and u_b
Japan	Robertson et al. (1986) Senneset and Janbu (1984)	q_t and FR q_t , u_b
Netherlands	Computer code: UNICLASS (Peuchen et al., 1995)	q_c and f_s
New Zealand	Douglas and Olsen (1981) Jones and Rust (1982) Robertson et al. (1983) Senneset and Janbu (1985)	q_c and FR* ($q_t - \sigma_{vo}$) and ($u_b - u_o$) q_c and FR* q_t and B_q
Nigeria	George and Ajayi (1995)	FR*
Norway	Senneset et al. (1989) Robertson et al. (1986) Robertson (1990,1991)	q_t and B_q q_t , FR and B_q Q , B_q and F
Poland	Olsen and Malone (1988) Senneset et al. (1989) Robertson (1990,1991)	q_{cl} and FR _l q_t and B_q Q and B_q and F
Romania	Marcu and Culita (1995)	q_c
Russia	Trofimenkof et al. (1995)	q_c
Singapore	Robertson (1990,1991)	Q , B_q and F
Slovenia	Robertson and Campanella (1983)	q_c and FR*
Spain	Robertson and Campanella (1986)	q_t and FR
Sweden	Larsson and Mulabdic (1991)	q_t , f_s , u_b and B_q

Table 2.2. (Cont'd) Cone Classification Systems Used in Different Countries Summarized from CPT'95 Proceedings.

Country	Cone classification method(s)	Cone data
Switzerland	Robertson et al. (1986) Robertson (1990,1991)	q_t , FR and B_q Q , B_q and F
Turkey	Robertson and Campanella (1986)	q_t , FR and B_q
United Kingdom	Douglas and Olsen (1981) Jones and Rust (1982) Robertson et al. (1983) Senneset and Janbu (1985)	q_c and FR* ($q_t - \sigma_{vo}$) and ($u_b - u_o$) q_c and FR* q_t and B_q
United States	Robertson et al. (1986) Robertson (1990,1991)	q_t , FR and B_q Q , B_q and F
Vietnam	Tuong et al. (1992)	q_c

Notes:

q_c = cone tip resistance

q_t = corrected cone tip resistance

f_s = sleeve friction

f_t = corrected sleeve friction

u_b = penetration porewater pressure behind the tip

FR* = friction ratio = f_s/q_c

FR = friction ratio = f_s/q_t

σ_{vo} = total vertical stress

σ_{vo}' = effective vertical stress

u_o = hydrostatic porewater pressure

Q = normalized tip resistance = $(q_t - \sigma_{vo})/\sigma_{vo}'$

F = friction ratio = $f_s/(q_t - \sigma_{vo})$

B_q = normalized pore pressure = $(u_b - u_o)/(q_t - \sigma_{vo})$

q_{ci} = normalized tip resistance = $q_c/(\sigma_{vo}')^n$

FR₁ = friction ratio = $(f_s/\sigma_{vo}')/[q_c/(\sigma_{vo}')^n]$

n = 0.6 for sands, 0.8 for silts and 1.0 for clays

Note that some countries are using both methods such as in Canada, and the United States. The Robertson et al. method (1986) was based on the corrected cone tip resistance q_t , the friction ratio $FR = f_s/q_t$, and the normalized pore pressure parameter B_q . The Robertson charts (1991) were developed using the normalized cone parameters that were recommended by (Wroth, 1984 and 1988; Jamiolkowski et al., 1985; Houlsby, 1988; Robertson, 1990; Mayne, 1991) for soil behavior interpretation based on piezocone data, and defined as follows:

$$\text{Normalized tip resistance parameter } Q = \frac{q_t - \sigma_{vo}}{\sigma_{vo}'} \quad (\text{Equation 2.4})$$

$$\text{Normalized pore pressure parameter } B_q = \frac{u_b - u_o}{q_t - \sigma_{vo}} \quad (\text{Equation 2.5})$$

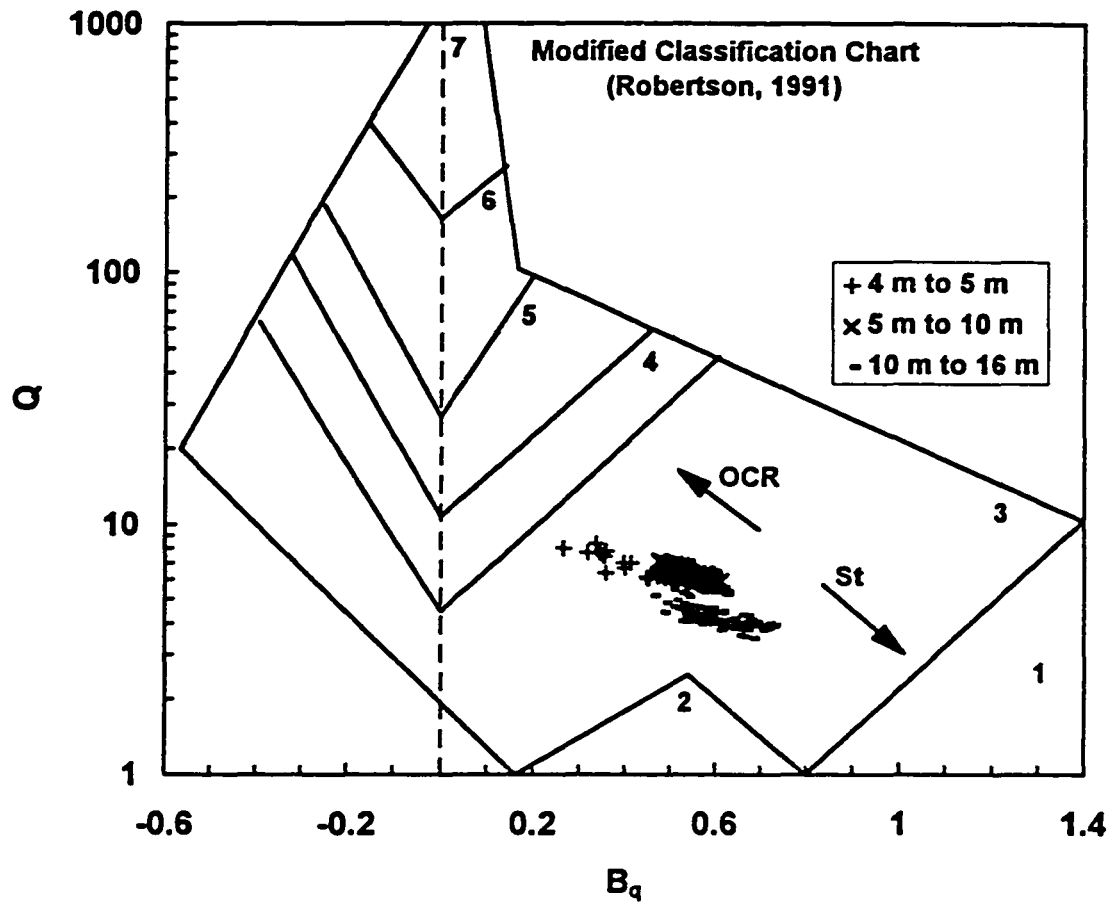
$$\text{Friction ratio } F = \frac{f_s}{q_t - \sigma_{vo}} \quad (\text{Equation 2.6})$$

where σ_{vo} is the total vertical stress, σ_{vo}' is the vertical effective stress, and u_o is the hydrostatic pore water pressure.

Therefore, the Robertson technique (1991) is chosen to detect the subsurface stratification of the studied sites in this report (Chapters 4 and 5). For example, piezocone data at Drammen, Norway (see Fig. 2.7) are examined. The derived normalized parameters Q and B_q are shown in Fig. 2.15. The soil is classified as one clay unit and the major changes in the soil plasticity at a depth of 10 m are not recognized. Therefore, a more objective method is needed to quantify a soil stratigraphy at a certain site. The technique of *cluster analysis* is introduced in this study.

2.7. Summary and Conclusions

In this chapter, different laboratory and in-situ methods for delineating the vertical layering and subsurface stratigraphy in soil exploration were evaluated. A common



- | | |
|---|--|
| 1. Sensitivity Fine Grained | 5. Sand Mixtures: Silty Sand to Sandy Silt |
| 2. Organic Soils-Peats | 6. Sands: Clean Sand to Silty Sand |
| 3. Clays: Clay to Silty Clay | 7. Gravelly Sand to Clayey Sand |
| 4. Silt Mixtures: Clayey Silt to Silty Clay | |

Figure 2.15. Soil Classification Using Piezocone Data at Drammen, Norway (Data from Masood et al., 1990).

method is to conduct borehole soundings and extract soil samples to be tested in the laboratory. However, soil sampling is usually discrete and intermediate information about the soil is missed. Performing extensive laboratory studies are often time-consuming and expensive.

In-situ tests including the cone penetration test are quick, continuous, and economical. Using cone testing, a larger volume of soil is investigated in its field conditions compared with laboratory testing. Therefore, cone testing has become a routine tool in current geotechnical practice and its primary use is for soil stratification using visual methods and/or cone classification schemes. Available cone soil classification methods and their limitations have been reviewed in this chapter. The readings of the sleeve friction channel were excluded from the scope of this study because they were shown by previous studies to be more dependable on the penetrometer-type and manufacture than those the tip resistance and pore pressure. In some geological settings, the simple visual method and/or the empirical cone classification approaches can not detect subtle changes in the cone data and major changes in the soil profile can be missed.

CHAPTER 3

STATISTICAL IDENTIFICATION OF SOIL LAYER BOUNDARIES

3.1. Introduction

A primary goal of a site investigation program is the delineation of subsurface stratification and soil type. Statistical approaches provide an objective and repeatable systematic technique for defining layer boundaries and the detection of small lenses and inclusions. The definitions of different statistical terms are given in Table 3.1. Statistical

Table 3.1. Definitions of Statistical terms, Modified after Lacasse and Lamballerie (1995).

Statistical Term	Definition
Autocorrelation function	Description of the correlation of the residuals about a trend in a certain distance.
Filtering	Removing elements from the raw data that indicate small inhomogeneties or errors.
Generalized distance	A multivariate similarity measurement between two groups of data.
Histogram	Graphical presentation of measurements frequency in a certain range.
Hierarchical cluster	A method for dividing the data into correlated groups.
Intraclass correlation coefficient	A univariate similarity measurement between two groups of data.
Mean (average)	Measure of the most likely value of a random variable.
Regression analysis	Fitting a correlation to data points by minimizing sum squares of the residuals.
Residuals	Algebraic measure of distance between a data point and a trend.
Scale of fluctuation	Distance within which there is a strong correlation of data representing a soil property.
Smoothing	Elimination of small inherent variability within raw data in a chosen distance.
Standard deviation	Measure of dispersion of a random variable with respect to the mean value.
Trigonometric polynomial	Numerical technique to fit a polynomial function of a certain degree to a large number of data points based on least squares method.

methods applied to cone data for the purpose of soil stratigraphy identification are summarized in Table 3.2 including goal of the analysis, and processed cone data.

Table 3.2. Applied Statistical Methods for Delineating Soil Stratigraphy Based on Cone Data.

Statistical methods	Analysis performed to delineate:	Remarks	Used cone data	Reference
Mean, standard deviation, autocorrelation function and scale of fluctuation	homogeneous sublayers within a soil stratum.	<ul style="list-style-type: none"> Number of soil layers and locations of their boundaries are not defined. 	q_c	Vanmarke (1977)
Hierarchical cluster analysis	homogeneity of physical and strength parameters of a clay deposit	<ul style="list-style-type: none"> Grouped cone data represented the variation of soil properties with depth. 	q_c	Mlynarek and Lunne (1987)
Filtering, smoothing, mean and standard deviation	number of primary soil layers based on visual examination of a cumulative profile of a cone parameter.	<ul style="list-style-type: none"> Data are filtered and some thin geological evidence are missed. Data trends are subjectively indicated. Univariate analysis. 	q_c and FR	Harder and Bloh (1988)
Mean, standard deviation and histogram	homogeneous sublayers within a soil stratum.	<ul style="list-style-type: none"> Soil stratigraphy should be estimated in advance. Univariate analysis. 	q_c and FR	Ghinelli and Vannucchi (1988)
Regression analysis	number of soil layers by optimization of different trends using a least square errors method.	<ul style="list-style-type: none"> A prior estimate of soil boundaries is needed. Univariate analysis. 	q_t , f_s and u_b	Wickremesinghe (1989)
Intraclass correlation coefficient (ICC) and generalized distance (D^2)	soil boundaries, lenses and transitions.	<ul style="list-style-type: none"> Both methods are unable to properly detect the soil stratigraphy in some soil conditions. The ICC is a univariate analysis and the D^2 is a multivariate analysis. 	q_t , f_s and u_b	Wickremesinghe (1989), Zhang (1996), Hegazy et al. (1996, 1997)
Trigonometric polynomials	soil boundaries at the inflection point of the polynomial fitting.	<ul style="list-style-type: none"> This technique is supplemented by a soil identification scheme to optimize the polynomial degree. Univariate analysis. 	q_t and $(q_t - \sigma_{vo})$	Vermeulen and Rust (1995)

Statistical methods applied to cone data can be separated in two groups as follows:

(1) univariate statistical methods in which one variable such as q_t can be analyzed, including mean, standard deviation, histogram, scale of fluctuation, regression analysis, trigonometric polynomials, and intraclass correlation coefficient, and (2) multivariate statistical methods in which more than one variable such as q_t and u_b can be analyzed, including autocorrelation function, clustering technique, and generalized distance.

The mean, standard deviation, and histogram are conventional statistical methods for indication of the homogeneity of a soil stratum. The number of soil layers should be predefined or estimated before performing the analysis. In the case of analyzing the large cone data sets, usually normal or lognormal statistical distribution is fitted to the histogram of the data (Lacasse and Lamballerie, 1995). Then, representative mean and standard deviation of the cone data are estimated. A large value of the coefficient of variation (standard deviation/mean) relatively indicates more variability of a soil layer compared with other layers. Regression analysis was also applied to cone data for the purpose of identifying soil boundaries (Wickremesinghe, 1989). First, a number of soil layers in a certain profile was estimated and then linear trends were fitted to the cone data of each layer using the least square method. Then, based on the variances of the residuals of the fitted trends, the positions of soil boundaries were estimated. These steps were repeated until the variances of the residuals were minimum or until a predefined number of iterations. Wickremesinghe (1989) indicated improper identification of geostratigraphy in some studied cases. Vermeulen and Rust (1995) fitted q_c data with a trigonometric

polynomial function of a degree n . In this analysis, the number of soil layers was defined using a soil classification chart and the proper n was determined by trial and error to fit the predefined soil stratigraphy. The autocorrelation function can only define a distance within a soil layer or a soil profile in which measured data such as cone readings are strongly correlated.

In the case of applying the mean, standard deviation, histogram, regression analysis and polynomials functions to the cone data, initial estimate of the number of soil layers and the locations of their boundaries is required either visually or by using classification charts. However, the intraclass correlation coefficient (ICC), the generalized distance (D^2) and cluster analysis have been applied to cone data without prior knowledge of a geostratigraphy. Therefore, these methods are discussed in more detail. A preliminary application of clustering analysis was performed by Mlynarek and Lunne (1987), and homogeneous sublayers within a clay deposit were successfully identified based on cone data. A thorough evaluation of clustering techniques has been carried out for the purpose of delineating soil stratifications in this study, and detailed discussion is given in the next chapters.

Both ICC and D^2 have been applied to demarcate geostratigraphies based on cone data (Wickremesinghe, 1989; Wickremesinghe and Campanella, 1991; Zhang, 1994; Hegazy et al., 1996 and 1997). The two methods were able to properly define soil boundaries in several case studies, however, there are limitations associated with their applications that might cause misinterpretation of specific soil profiles as discussed herein. The analysis of ICC and D^2 depends on the correlation distance which can be selected

using the autocorrelation function. The main or secondary soil boundaries depend on the chosen cut-off or statistical hypothesis used. Some examples will be presented in this chapter to discuss the validity and limitations of using ICC and D² methods to delineate different soil types.

3.2. Autocorrelation Function

The autocorrelation function is a description of the correlation of piezocone data in a certain distance. A soil stratigraphy is divided into statistically homogeneous zones or layers (windows) using the piezocone data. The width of this window should be equal to or less than the smallest thickness of a layer. Within a half of each window, the piezocone data and soil properties are highly correlated. In other words, the minimum layer which can be detected is equal to half of a window width, as depicted in Fig. 3.1. Therefore, all layer boundaries could be detected. Webster (1973) suggested the window width size be obtained using an autocorrelation function (r_{ah}). For instance, r_{ah} of q_t is in the following form (Christian et al., 1994):

$$r_{ah} = \frac{n}{n-m} \frac{\sum_{i=1}^{n-m} (q_{t(i)} - q_{t(av)}) (q_{t(i+m)} - q_{t(av)})}{\sum_{i=1}^n (q_{t(i)} - q_{t(av)})^2} = \frac{\text{cov}_h(q_t)}{\text{var}(q_t)} \quad (\text{Equation 3.1})$$

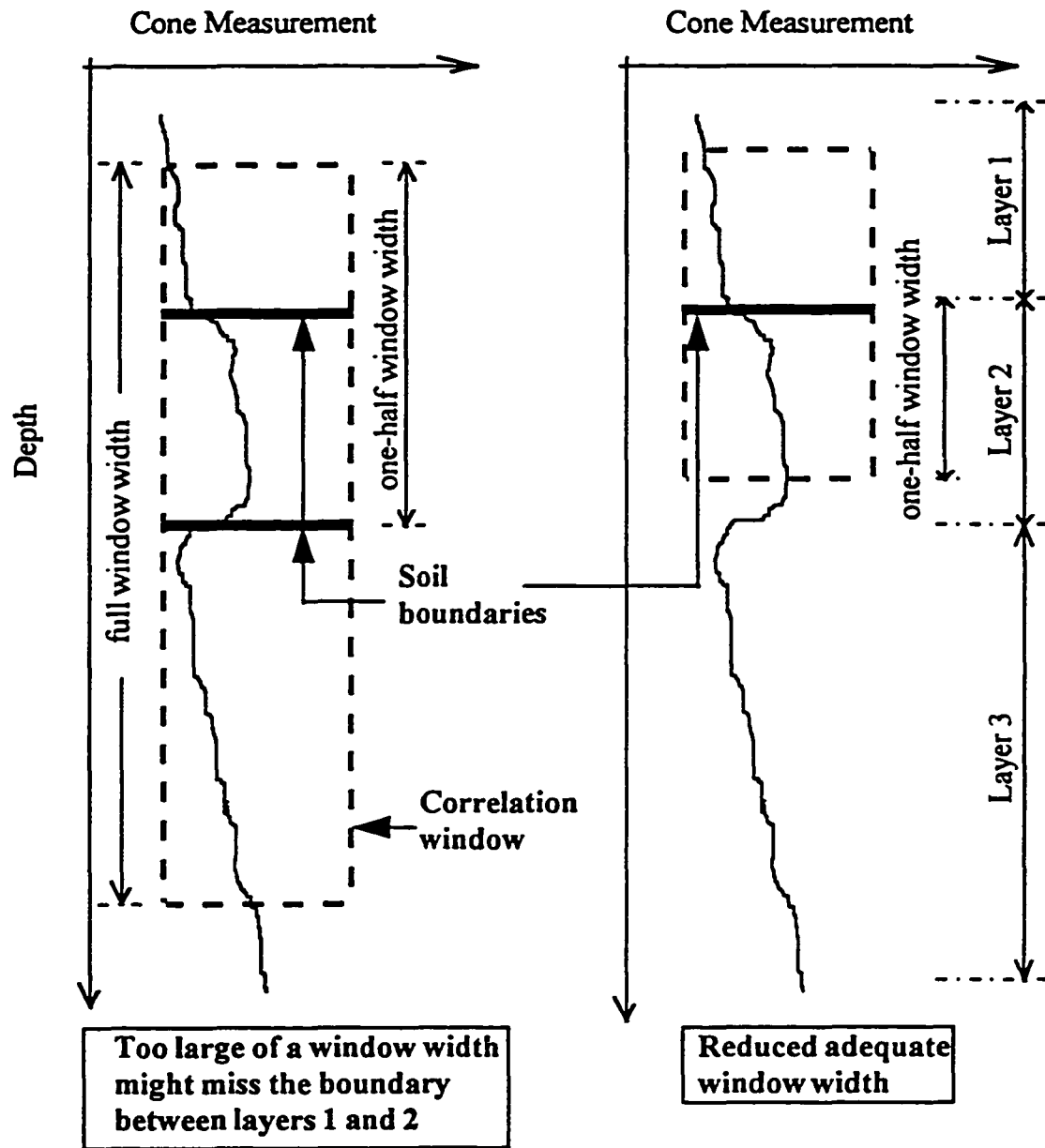


Figure 3.1. Effect of the Window Width on the Delineation of Soil Boundaries.

in which n is the number of q_t points in a sounding, m is the number of measurement intervals between two points, h is equal to m multiplied by data interval, $q_{(i)}$ is the corrected tip resistance at a depth i , $q_{(av)}$ is the average of all q_t measurements. Christian et al. (1994) also noted that r_{ah} is the ratio of the covariance (cov_h) between the cone data separated by a distance h divided by the variance (var) of all data. For instance, Figure 3.2 shows an example of correlating consecutive q_t data with depth using r_{ah} .

Then, a model is fitted to the experimental r_{ah} value in the following form:

$$r_{ah}^* = 1 - \frac{\gamma(h)}{\text{var}(x)} \quad (\text{Equation 3.2})$$

in which $\gamma(h)$ is a simple expression “termed” variogram, $\text{var}(x)$ is the variance of piezocone data such as q_t , and h is the distance between the data. Some fitting models are summarized by Journel and Huijbregts (1993) and include: nugget, experimental, Gaussian, exponential, and linear models, as shown on Fig. 3.3. The forms of these variograms are summarized in Table 3.3. Note that the theoretical variograms are conditional positive definite functions to ensure that all calculated variances are positive (Journel and Huijbregts, 1993). The nugget model indicates there is no correlation between the data and needs only one component [the sill (S)] to be defined. The sill is the actual dispersion or variability of the measurements. The linear model also has only one parameter [the slope (b)] and indicates a drift (trend) in the data. The spherical,

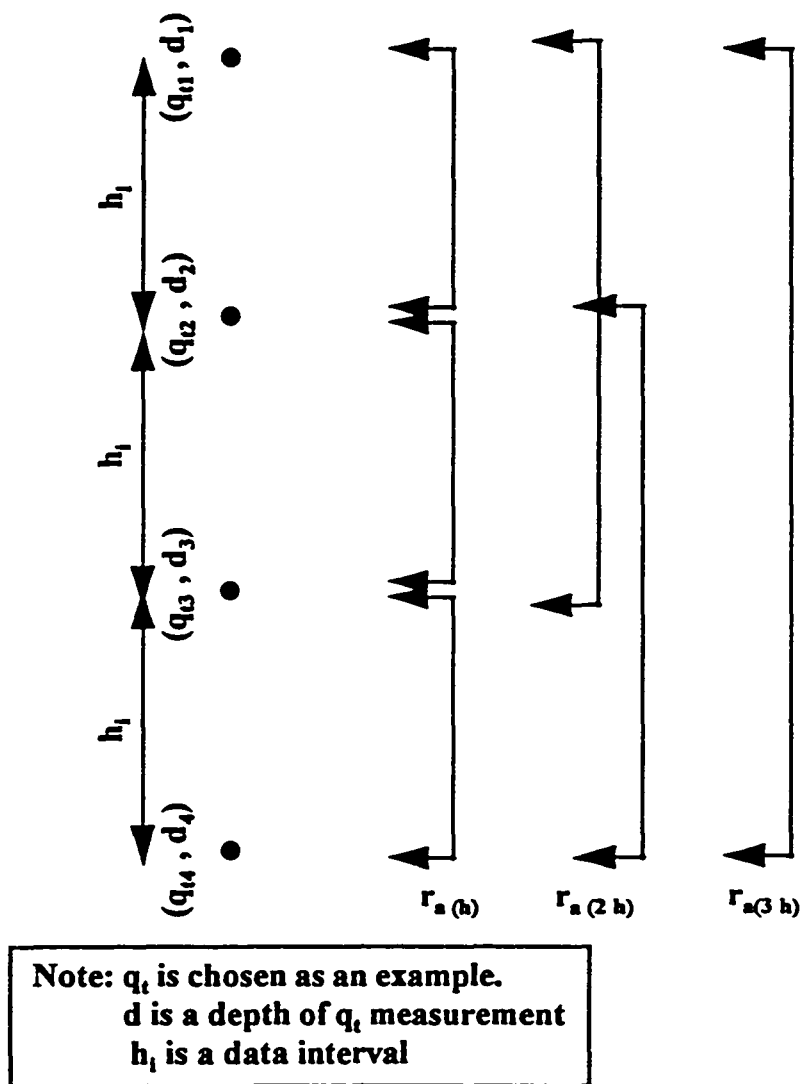


Figure 3.2. Combination Sequence to Calculate Autocorrelation Function (r_s) at Different Distances.

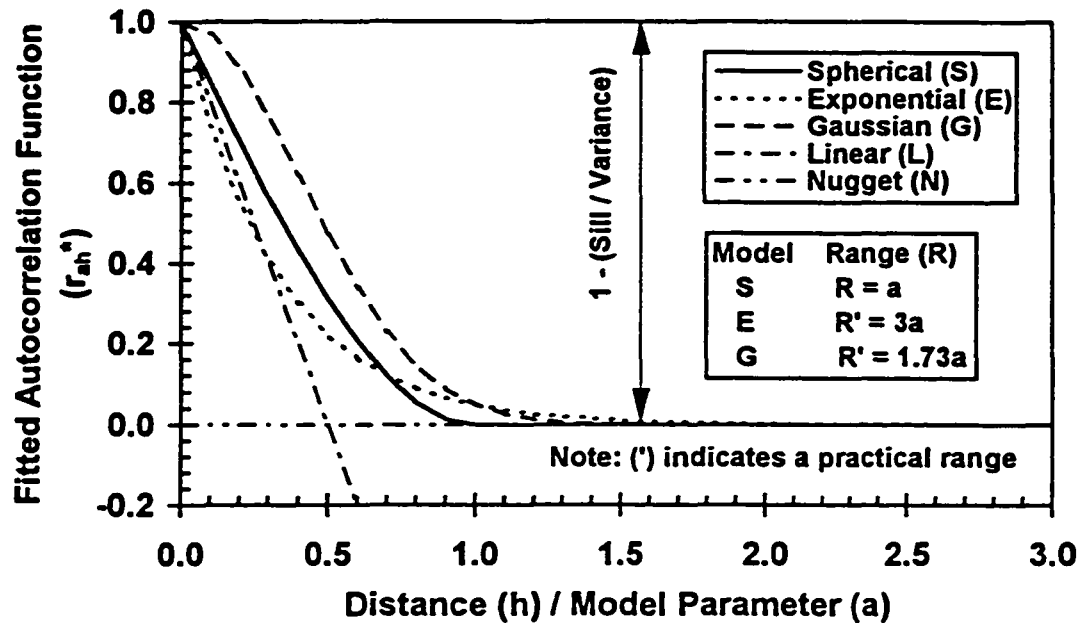


Figure 3.3. Common Models Fitted to the Autocorrelation Function
(Modified after Journel and Huijbregts, 1993).

Table 3.2. Common Fitting Variogram Forms for Autocorrelation Function
(Based on Journel and Huijbregts, 1993).

Name	Variogram form	Remarks
Nugget	$\gamma(h) = S$	h = distance between data S = sill
Linear	$\gamma(h) = bh$	b = slope
Spherical	$\gamma(h) = S \left(1.5 \left(\frac{h}{a} \right) - 0.5 \left(\frac{h}{a} \right)^3 \right)$	$R = a$ = range Note: at $h = R$, $\gamma(h) = S$
Exponential	$\gamma(h) = S \left(1 - \exp \left(\frac{-h}{a} \right) \right)$	$R' = 3a$ = practical range Note: at $\gamma(h) = 0.95 S$
Gaussian	$\gamma(h) = S \left(1 - \exp \left(\frac{-h^2}{a^2} \right) \right)$	$R' = 1.73a$ = practical range Note: at $\gamma(h) = 0.95 S$

Notes: R = range from theoretical point
 R' = practical range to limit asymptote

exponential, and Gaussian models each have two parameters: the sill and a distance parameter (a). The maximum correlation distance between the data is “termed” the range (R) which thus represent the minimum thickness of a layer within a soil profile. In case of a spherical model, R is equal to a. However in case of exponential and Gaussian models, the sill is reached asymptotically and a practical range can be taken equal to $R' = (3 a)$ and $R' = (1.73 a)$, respectively, at $\gamma(h) = (0.95 S)$.

For example, Hegazy et al. (1996) applied two-dimensional geostatistical analysis to delineate different soil layers at Surry, Virginia. Piezocone data were collected using both 10-cm² and 15-cm² tip penetrometers (Gordon and Mayne, 1987). The geological formation at the site consists of recent alluvial deposits underlain by interbedded Atlantic coastal plain sediments of clays, silts, sands, and gravels of Pleistocene Age. These layers are underlain by preconsolidation clays of Miocene/Pliocene age, known locally as the Yorktown Formation. The subsurface stratigraphy consisted of two clay layers with an intermediate sand layer. A representative piezocone sounding (CP5) using a 10-cm² cone and the soil profile are shown in Fig. 3.4. The autocorrelation function (r_{ah}) was determined for both q_t and u_b as shown in Fig. 3.5. Different models summarized in Table 3.3 were fitted to the experimental r_{ah} and exponential models were chosen because they gave the least weighted sum square errors. Each point of r_{ah} is given a weight (w) defined as follows:

$$w = \frac{n_{p(i)}}{n_{p(1)}} \quad \text{for } i = 1 \text{ to } (n-1) \quad (\text{Equation 3.3})$$

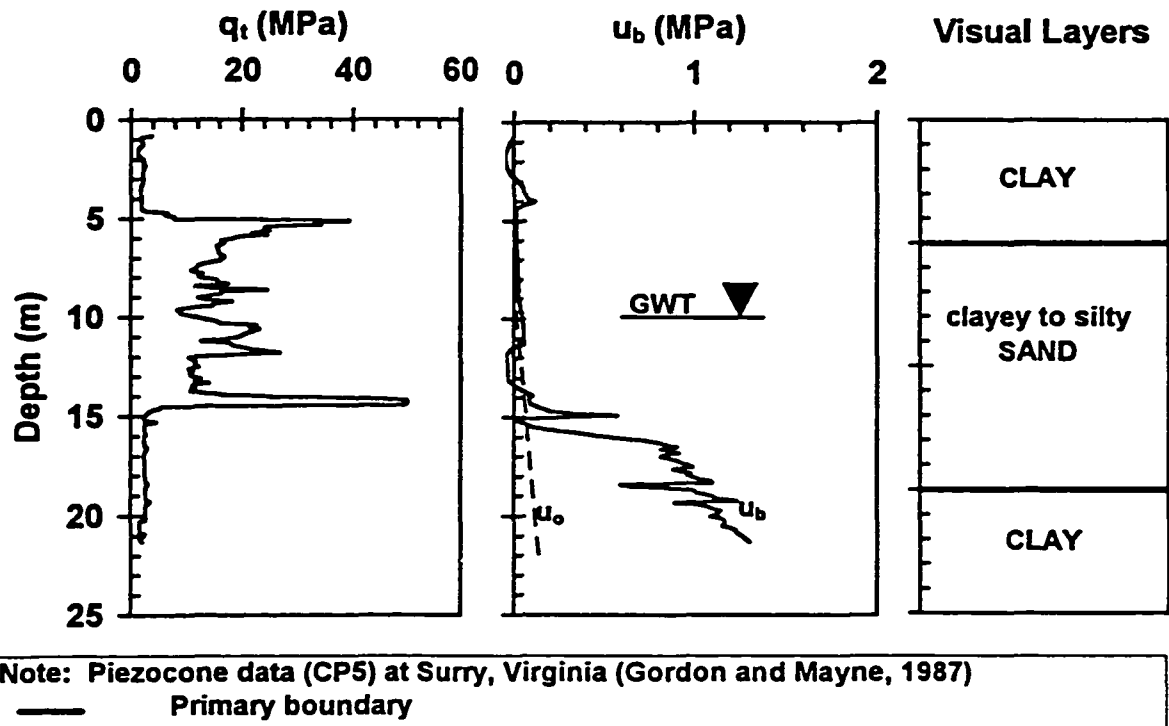


Figure 3.4. Piezocone Data at Surry, Virginia and Subsurface Profile.

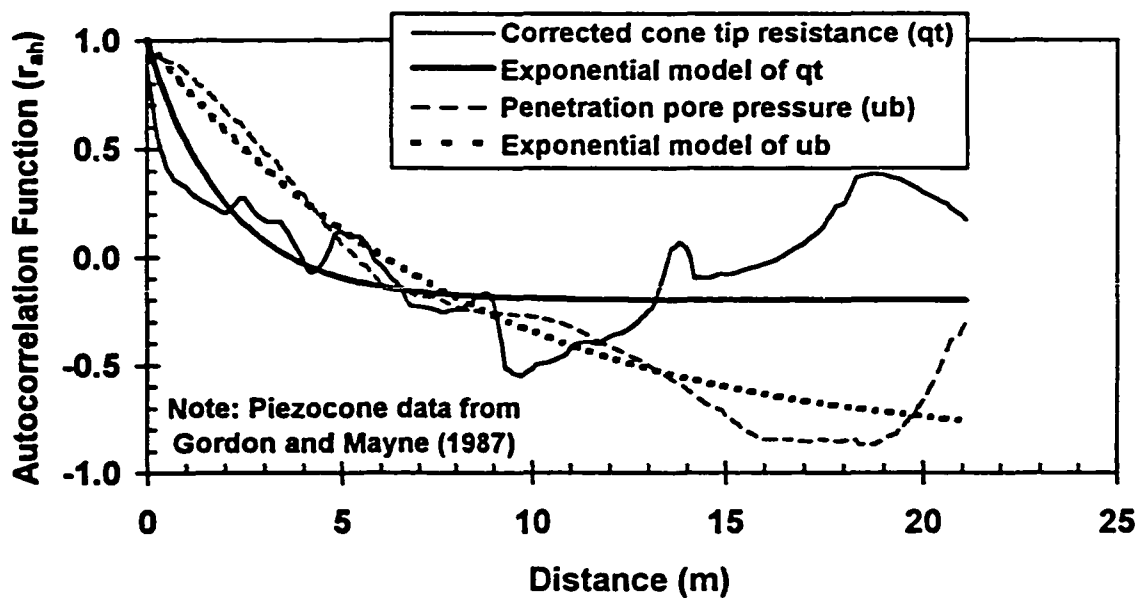


Figure 3.5. Autocorrelation Function of Piezocone Data at Surry, Virginia (Modified from Hegazy et al., 1996).

in which n_p is the number of data pairs at a certain distance (h), n is the number of data set and i is a distance number where h_i equal to the shortest distance between pairs of data and h_{n-1} is the largest distance. For instance, $n_{p(1)} = n(n-1)/2$. More weights are given to r_{ah} values at shorter distances because the closer data are more likely to be correlated. The fitted models are shown in Fig. 3.5. The exponential model of q_t has a zero nugget, sill/variance = 1.2, and range = 6 m. The components of the porewater pressure ($u_2 = u_b$) in terms of an exponential model are as follows: nugget = zero, sill/variance = 1.9, and range = 18 m. It is evident that the penetration porewater pressure readings are relatively more correlated than the tip resistance records because they have larger ranges. The ratio of the range of the former to the range of the latter is equal to 3. The correlation window width was chosen equal to 3 m, half of the smaller range.

Also, Hegazy et al. (1997) performed a three-dimensional geostatistical study at Opelika, Alabama. The site is located in the Piedmont geologic province and the soil profile consists of an upper silty sandy clay layer underlain by a soil mixture of sand and silt, all of which are residual soil byproduct, caused by the in-plane weathering of schist and gneiss. A representative piezocone sounding collected by the author is shown in Fig. 3.6. The experimental autocorrelation functions (r_{ah}) of q_t and u_b are shown in Fig. 3.7. Different models were fitted to r_{ah} and exponential variograms gave the least weighted sum square errors. The fitted model of q_t has the following parameters: nugget = zero, sill/variance = 1.5, and range = 9 m. The exponential model of u_b has a zero nugget,

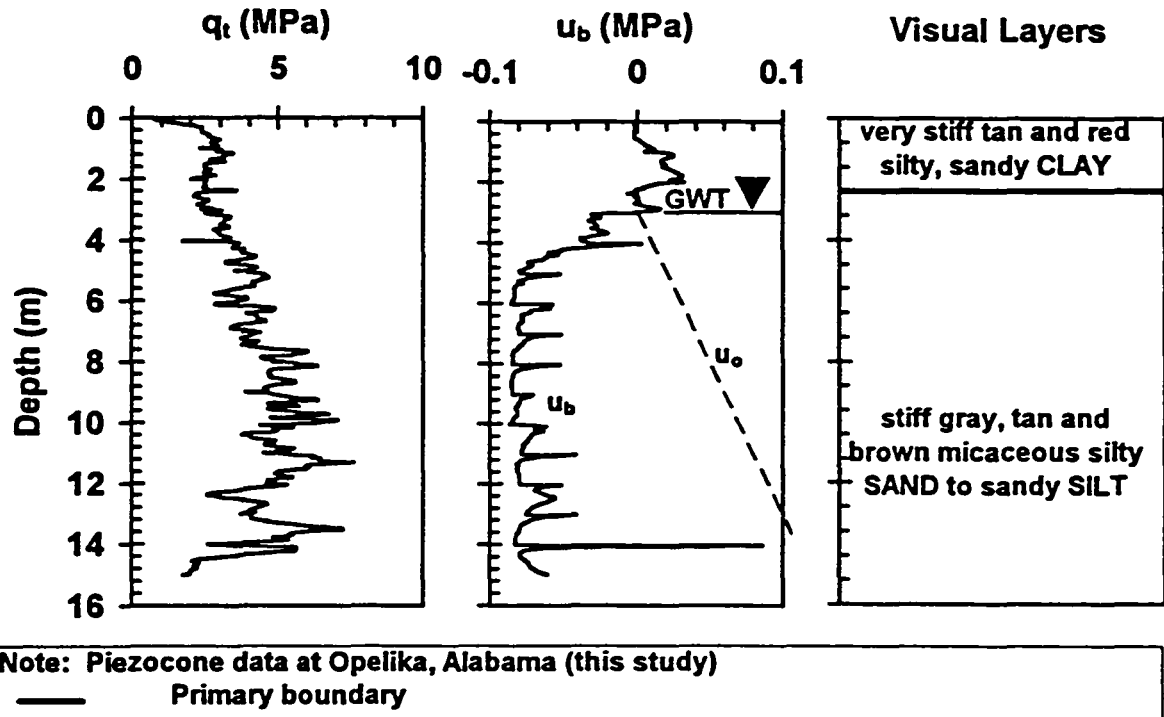


Figure 3.6. Piezocone Data at Opelika, Alabama and Subsurface Profile.

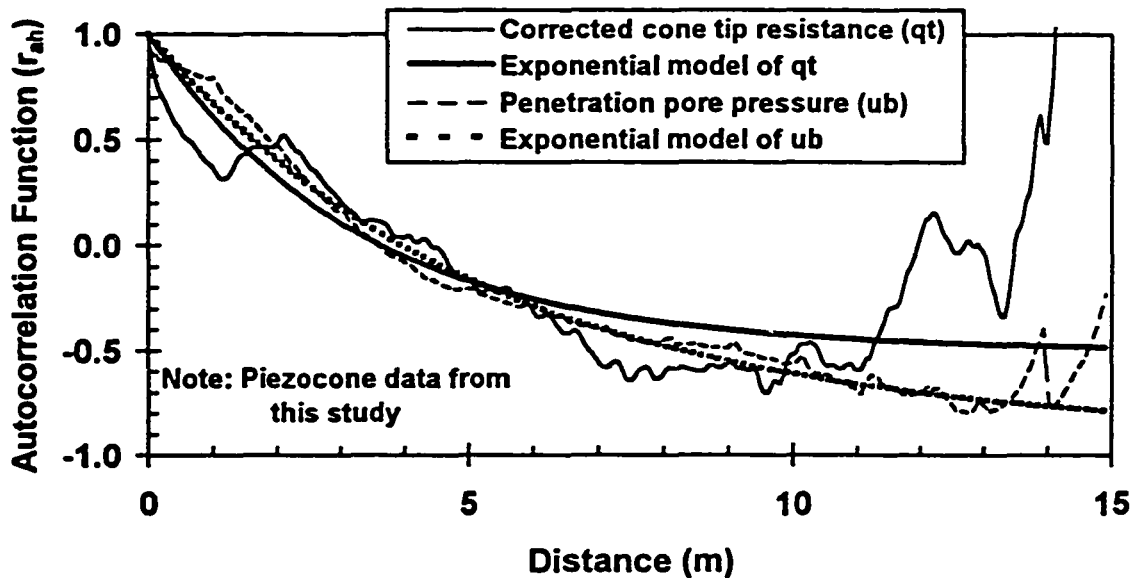


Figure 3.7. Autocorrelation Function of Piezocone Data at Opelika, Alabama (Modified from Hegazy et al., 1997).

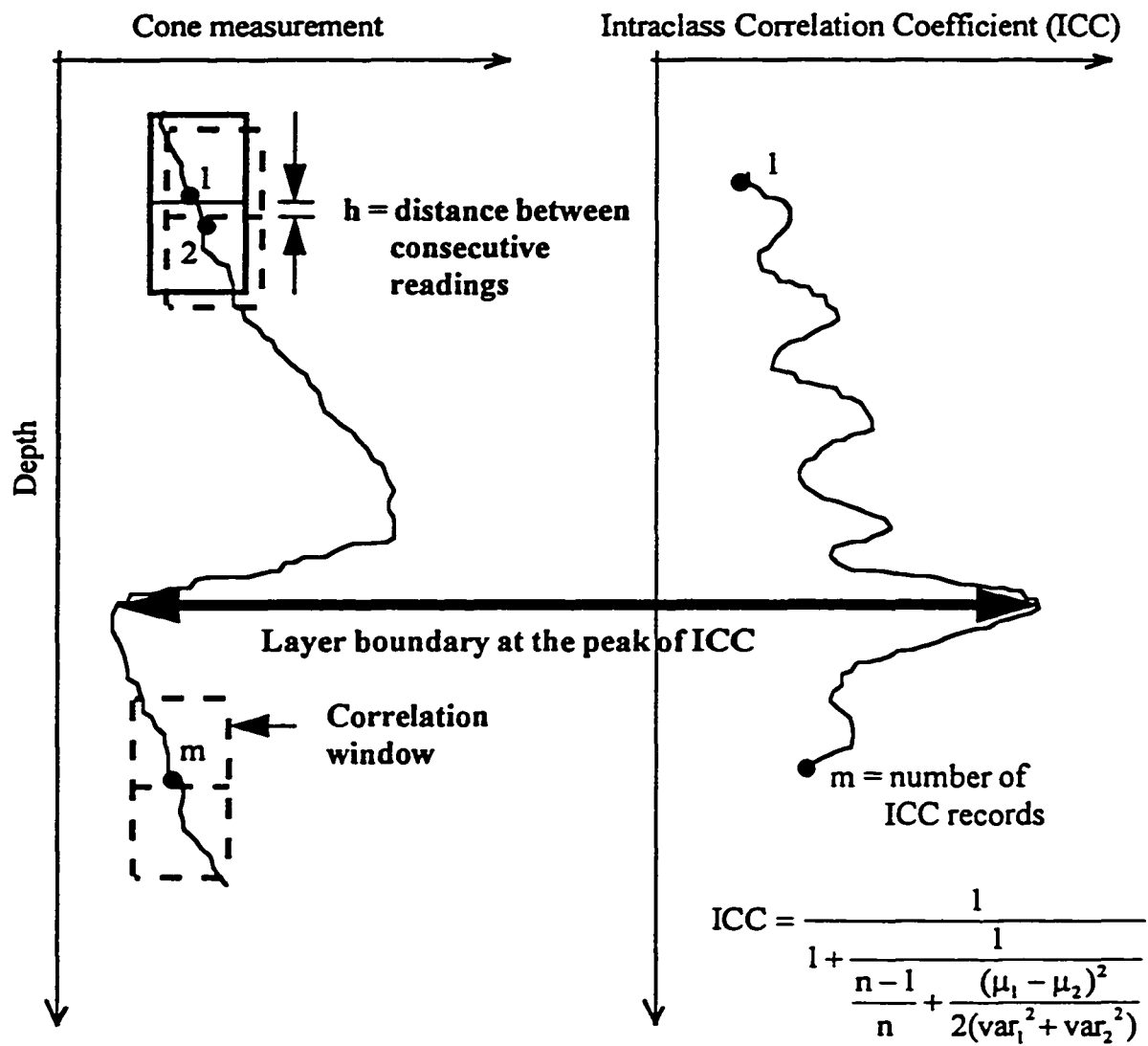
sill/variance = 1.9, and range = 15 m. It is also noticed that the correlation distance of the porewater pressure is larger than that of the tip resistance. The correlation window width was chosen equal to 4.5 m, half of the smaller range.

3.3. Intraclass Correlation Coefficient and Generalized Distance

The size of the window is related to the range (R). An empirical window width was suggested by Webster (1973) to be about 0.5 R to 0.67 R. Then, this window is moved along each variable of the piezocone sounding as shown in Fig. 3.8. The average and variance of the data points of the upper and lower halves of the window are compared using a univariate analysis such as ICC or multivariate analysis such as D². In the case of univariate analysis, one variable such as Q or B_q is used to define the boundaries. However, in the case of multivariate analysis, one or more variables can be included. Zhang (1996) introduced the form of ICC as follows:

$$ICC = \frac{1}{1 + \frac{1}{\frac{n-1}{n} + \frac{(\mu_1 - \mu_2)^2}{2(\text{var}_1^2 + \text{var}_2^2)}}} \quad (\text{Equation 3.4})$$

In which var₁ and var₂ are the variances of the lower and upper halves, respectively, μ₁ and μ₂ are the averages of the upper and lower halves, respectively, and n is the number of data sets in a window. The larger the square difference of the two means and the smaller the sum of the two variances, the larger the ICC value which means that the upper half of



Notes: var_1 = variance of the upper half of a window
 var_2 = variance of the lower half of a window
 μ_1 = mean of the upper half of a window
 μ_2 = mean of the lower half of a window
 n = number of points in a window

Figure 3.8. Movement of the Correlation Window along a Cone Measurement Profile, and the Recorded Intraclass Correlation Coefficient ICC.

a window belongs to a soil layer and the lower half of this window belongs to another soil layer. The boundary layers are defined at the peaks of ICC as shown on Fig. 3.8. The maximum value of ICC is 1 where $[(\mu_1 - \mu_2)^2 / (\text{var}_1^2 + \text{var}_2^2)] \gg 1$ and the minimum value of ICC is $(n - 1) / (2n - 1)$ where $\mu_1 = \mu_2$. However, the maximum value of ICC has not reached a value of 1 based on this study and previous applications (Hegazy et al., 1996 and 1997) because the window width is usually chosen small enough to detect soil boundaries. The minimum value of ICC is approximately equal to 0.5 if the number of points (n) in a window width are ≥ 6 . Zhang (1994) recommended n to be larger than 15 based on a parametric study, therefore the minimum value of ICC is 0.5 in the case of cone data applications.

Harbaugh and Merriam (1968) introduced the generalized distance (D^2) in the following form:

$$D^2 = d^T W^{-1} d \quad (\text{Equation 3.5})$$

in which d is the vector of differences between the means of the two window halves, d^T is the transpose of the vector d, and W is the variance-covariance matrix. The vector d of piezocone data is expressed as follows:

$$d = \begin{bmatrix} \mu_{qtU} - \mu_{qtL} \\ \mu_{ubU} - \mu_{ubL} \end{bmatrix} \quad (\text{Equation 3.6})$$

In which, μ_U and μ_L are the averages of the upper and lower halves of a window of a cone measurement. The matrix W of piezocone data is defined as follows:

$$W = \begin{bmatrix} SD_{pq_t}^2 & cov_{pq_t u_b} \\ cov_{pq_t u_b} & SD_{pu_b}^2 \end{bmatrix} \quad (\text{Equation 3.7})$$

in which SD_p^2 is the pooled variance and cov_p is the pooled covariance of the upper and lower halves of a window in the following form, for instance, for q_t and u_b :

$$SD_{pq_t}^2 = \frac{n}{2n-1} (SD_{q_t U}^2 + SD_{q_t L}^2) \quad (\text{Equation 3.8})$$

$$cov_{pq_t u_b} = \frac{n}{2n-1} (cov_{q_t u_b U} + cov_{q_t u_b L}) \quad (\text{Equation 3.9})$$

in which n is the number of q_t or u_b measurements in a half window, and cov is the covariance of q_t and u_b in the same half of a window and determined as follows:

$$cov_{q_t u_b} = \frac{\sum_{i=1}^n q_{ti} u_{bi}}{n} - \frac{\sum_{i=1}^n q_{ti}}{n} \frac{\sum_{i=1}^n u_{bi}}{n} \quad (\text{Equation 3.10})$$

The concept of D^2 for two variables, for instance, q_t and u_b , is illustrated in Fig.

3.9. The soil boundaries are defined at the peaks of D^2 . Wickremesinghe and Campanella (1991) found that the results obtained using D^2 to be more reliable than ICC because the soil layers are given based on the analysis of more than one measurement.

Wickremesinghe (1989) suggested that the primary boundary be defined where $ICC \geq 0.8$ and a secondary boundary be given when $0.80 > ICC > 0.65$. However, it was found that the boundaries were given at different levels of D^2 based on the soil type because D^2 was calculated using the cross correlation between the piezocone measurements. A parametric study was also performed to evaluate the effect of the window width on the given soil boundaries. The location of the primary boundaries were relatively insensitive to the window width effect, however, the secondary boundary which was defined at a lower level of peaks were more sensitive. It was found that too large of a window width might hide secondary boundaries whereas too small of a window width might cause noise and difficulty in interpreting some of the boundaries. Zhang (1994) recommended the window width to be equal to 1.5 m based on a parametric study using cone data at 6 sites.

Wickremesinghe (1989) performed a statistical study of piezocone data collected in Canada using both ICC and D^2 and found that neither can detect soil boundaries in two special cases depicted in Fig. 3.10. In one case, an upper layer has a gradient that is underlain by a layer with gradient in opposite direction. In the other case, there is a minor change in the averages of two consecutive layers as depicted in Fig. 3.10. The two

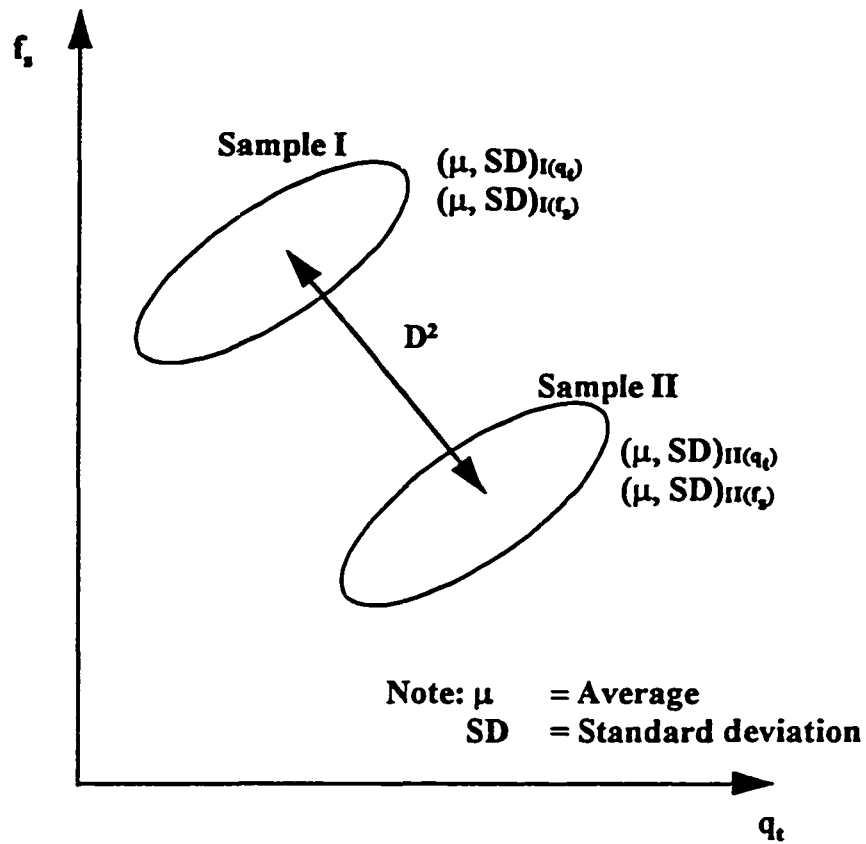


Figure 3.9. Concept of the Generalized Distance (D^2) Statistics for Two Samples of q_t and f_t (after Harbaugh and Merriam, 1968).

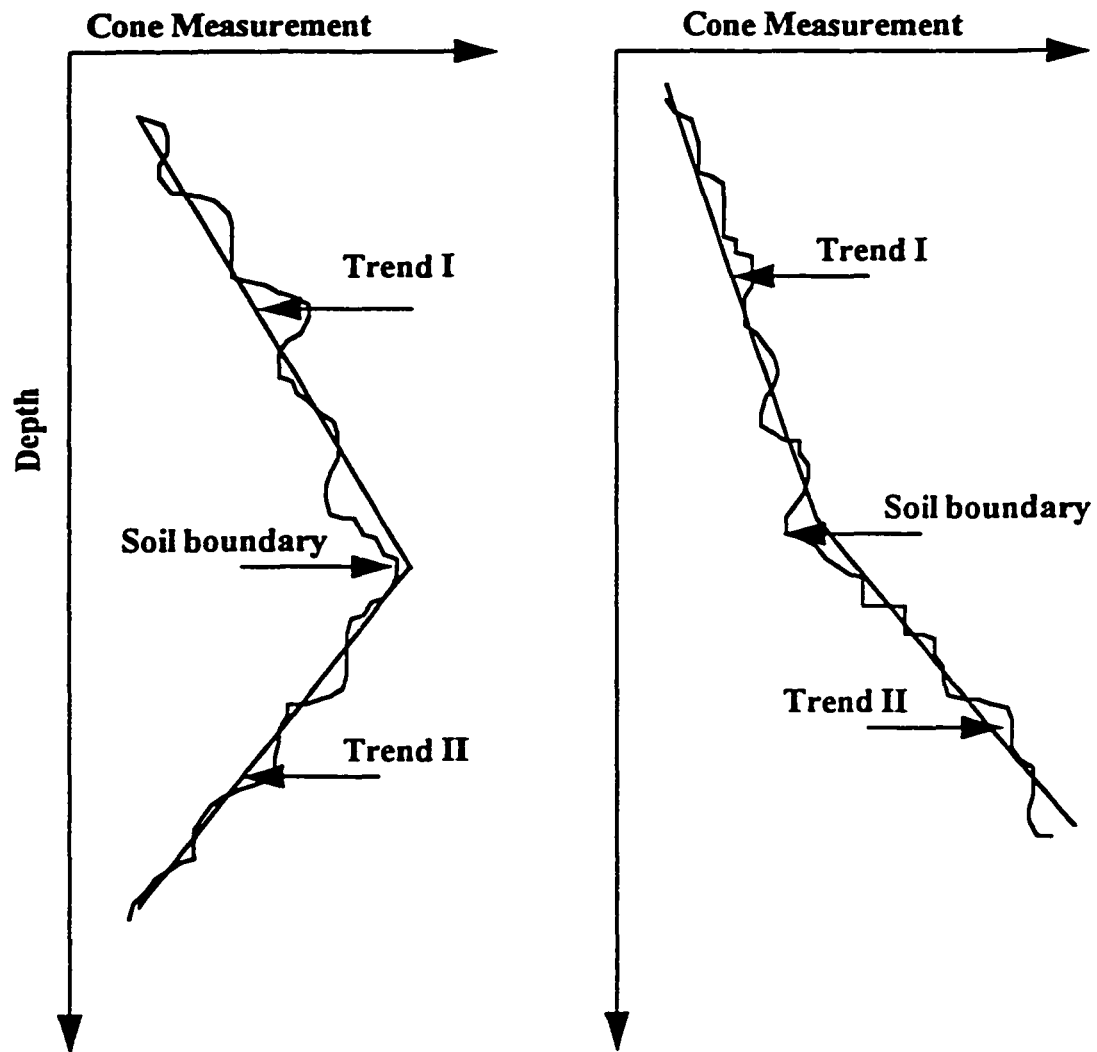


Figure 3.10. Two Cases for which ICC and D^2 Methods Are Unable to Detect the Soil Layer Boundaries (after Wickremesinghe, 1989).

methods do not give information about the correlation between the defined soil layers, therefore the number of soil types at a certain stratigraphy can not be known. Furthermore, neither method is able to detect some major changes in a soil stratigraphy as discussed in the following section.

3.4. Application of ICC and D² Methods to Detect Subtle Changes in Piezocone Data

In some soil stratigraphies, clay layers have dramatic changes in their mechanical and physical properties. For instance, major changes are indicated in vertical soil profiles at the following cases: the clay plasticity and sensitivity at Drammen, Norway (Lacasse and Lunne, 1982), the clay water content and overconsolidation ratio at Gloucester, Canada (Konrad and Law, 1987), and the organic content at Lilla Mellösa, Sweden (Larsson and Mulabdić, 1991). However, at these sites, the layer boundaries are difficult or in some cases impossible to detect by a simple visual inspection of piezocone profiles with depth. Also, available classification methods based on piezocone data are not able to detect the drastic differences of clay properties with depth. An evaluation of the detected soil boundaries at Drammen and Lilla Mellösa test sites is given in the following sections.

3.4.1. ICC and D² at Drammen, Norway

An example of ICC and D² statistical analyses is presented using piezocone data at Drammen, Norway. A representative piezocone sounding and the soil stratigraphy at Drammen (Lacasse and Lunne, 1982) are depicted in Fig. 3.11. The soil at the site consists of a 2-m sand layer underlain by clay. Piezocone data are reported only for the

interval from 4 to 16 meters deep (Masood and Mitchell, 1990). The clay plasticity changes with depth, particularly at a 10- m depth. The application of ICC and D^2 to the normalized parameters (Q and B_q) is discussed herein.

First, the experimental autocorrelation functions of Q and B_q are determined and the results are shown in Fig. 3.12. Each point of r_{ah} was given a weight (w) and the sum of the weights was equal to 1. More weight was given to calculated r_{ah} using larger number of data pairs. Then the various models (summarized in Table 3.3) were fitted to the experimental r_{ah} and a model was chosen based on the weighted least square errors criterion. A linear model as shown on Fig. 3.12 was fitted to r_{ah} of Q with a slope equal to (-0.252 1/m) . An exponential model was chosen for r_{ah} of B_q that had a zero nugget, sill/variance = 0.95, and range $R = 2 \text{ m}$.

3.4.1.1. Effect of the Window Width on the Chosen Boundaries

The maximum correlation distance (window width) is taken equal to the minimum R of both models and in this example is equal to 2 m. Statistical soil boundaries are compared using three different window widths. The smallest window width is taken equal to 0.5 m. The intermediate window width is selected equal to half of the correlation distance (1 m) as suggested by Webster (1973). This also agrees with a recommendation by Wickremesinghe (1989). The largest window width is chosen equal to the correlation distance (2 m). A window is moved from point to another on the vertical profile of each normalized parameter and ICC and D^2 are calculated at a window center. The determined profiles of ICC and D^2 are shown on Fig. 3.13 indicating that no soil boundaries can be

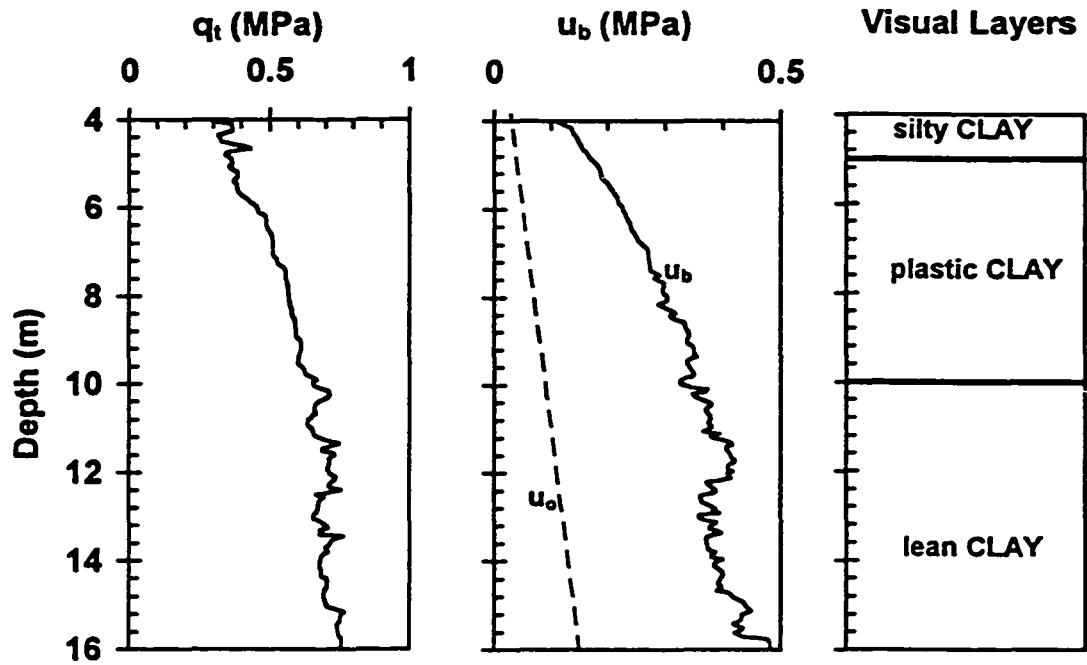


Figure 3.11. Piezocone Data at Drammen, Norway (after Masood et al., 1990)

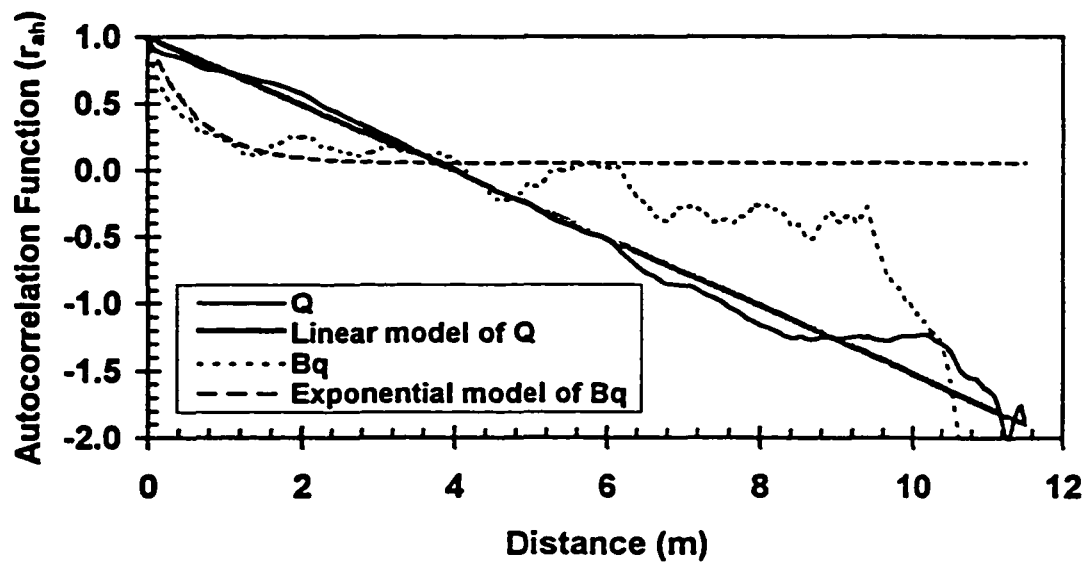


Figure 3.12. Autocorrelation Function at Drammen, Norway (Piezocone Data from Masood et al., 1990).

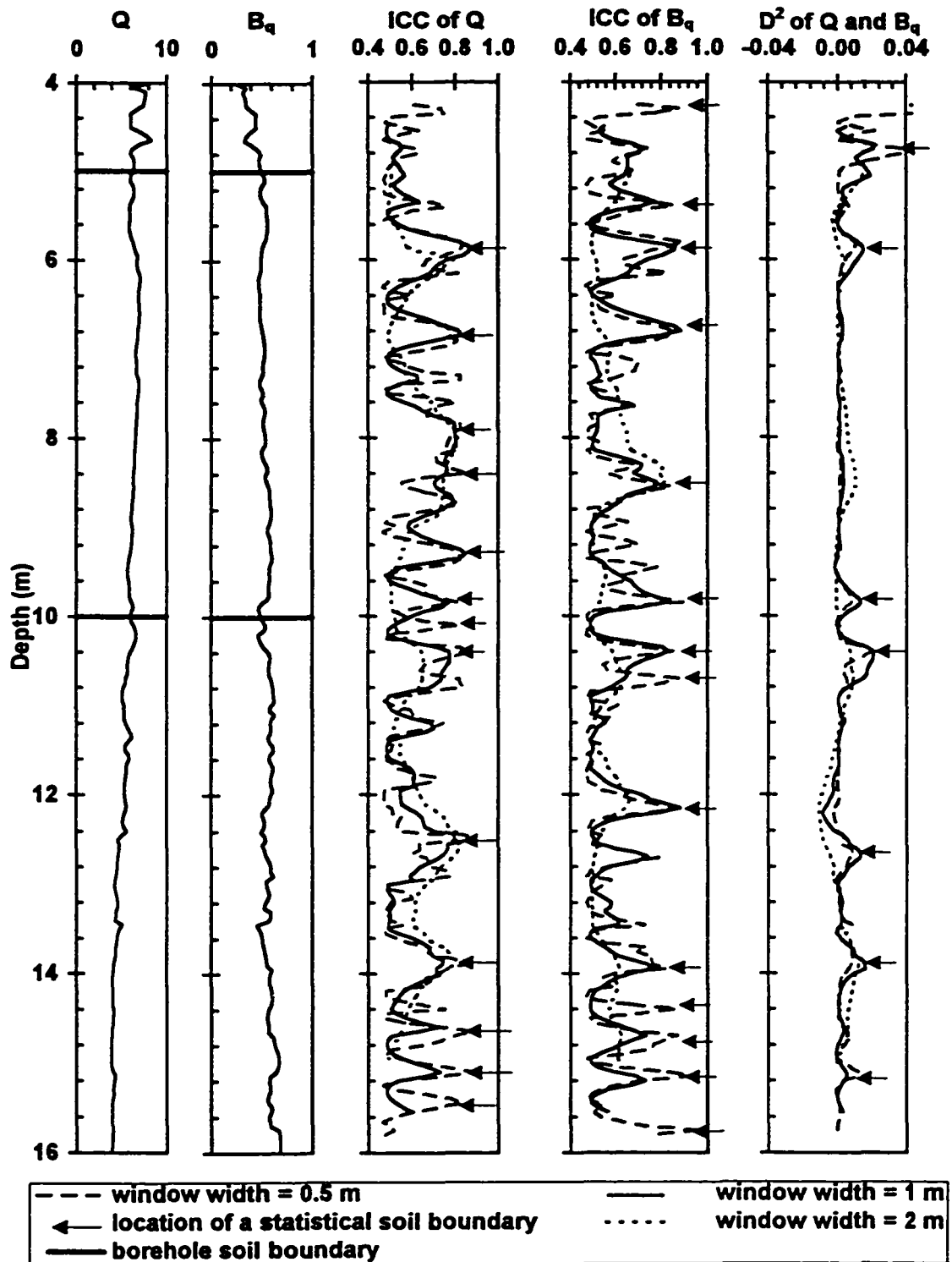


Figure 3.13. Effect of Window Width on Determination of ICC and D^2 Using Piezocone Data at Drammen, Norway.

detected visually from the Q and B_q profiles. By doubling the window width, the statistical profiles become smoother with less peaks. Soil boundaries are determined at the peaks of $ICC > 0.8$ and at $D^2 > (\text{average} + 1.65 \text{ standard deviation})$ of D^2 's within the same profile with a level of significance equal to 0.05 (Hegazy et al., 1997).

A comparison between the borehole and the statistical soil boundaries using the three window widths equal to 0.5 m, 1 m and 2 m is shown in Fig. 3.14. None of the performed analyses satisfies the soil stratigraphy at the site. Applying the ICC analysis to Q profile using a window width equal to 0.5 m, seven statistical boundaries are detected. The one at a depth of 10.4 m is supported by the change from plastic to lean clays, however, there are no variation of the soil properties to indicate the other six boundaries. There are 5 boundaries defined using ICC of Q with a window width equal to 1 m. None of these matches the reference borehole boundaries. Using a window width equal to 2 m indicates two boundaries at depths of 8.0 m and 12.5 m although the plastic and lean clays are separated at 10 m.

The analysis of ICC (B_q) using a window width equal to 0.5 m discovered 7 soil layers within the stratigraphy. Moreover, using a window width equal to 1 m results in 5 primary boundaries. For both analyses, the change from plastic to lean clay might be demarcated between the depths of 9.8 m and 10.7 m. However the other boundaries do not match with any reference boundary and causes uncertainty of the interpretation. By doubling the window width, only one boundary is detected at 8.5 m.

Using D^2 of Q and B_q at window widths equal to 0.5 gives seven soil boundaries where six of them are matching with those defined using a window width equal to 1 m.

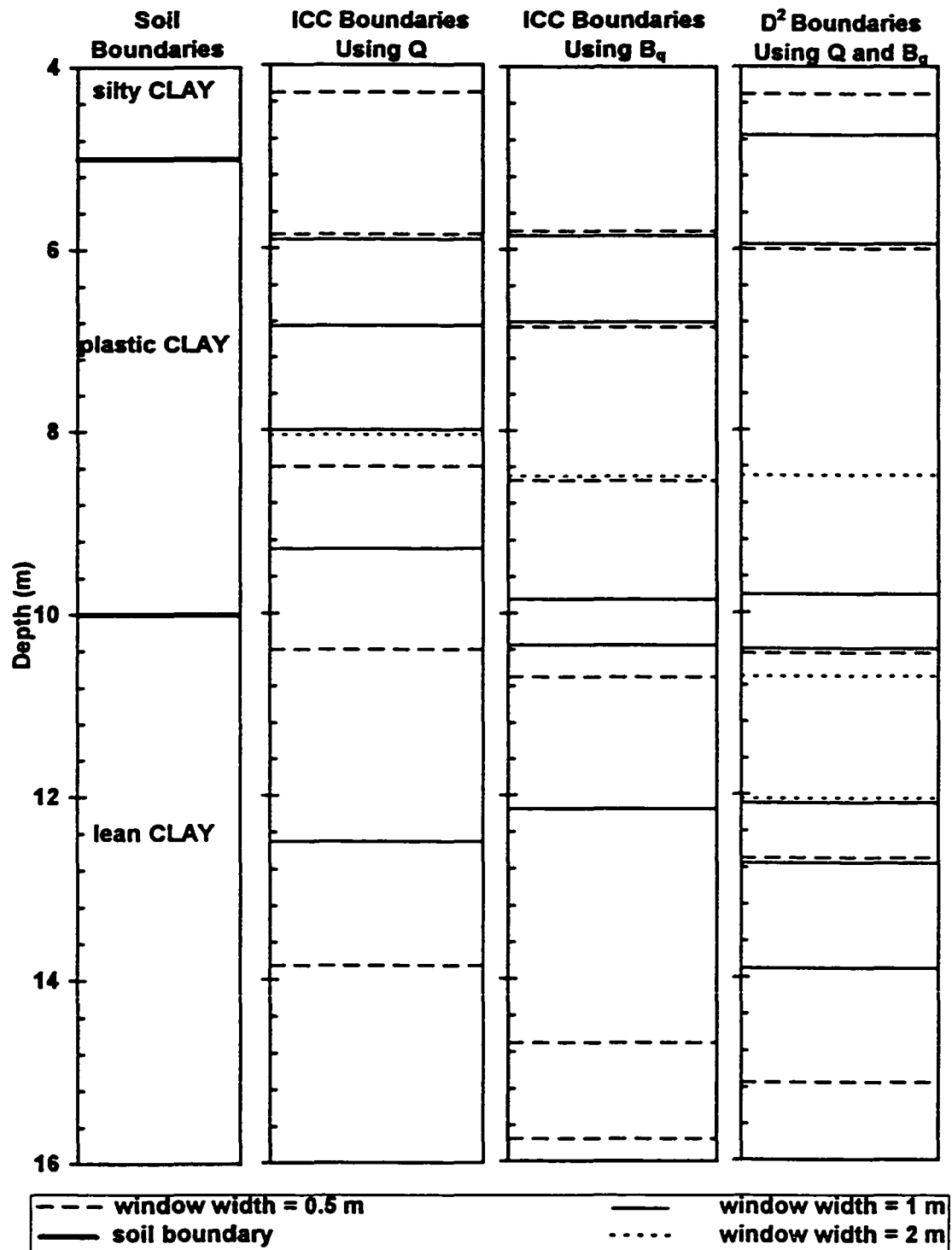


Figure 3.14. Statistical Soil Boundaries of ICC and D^2 Using Piezocone Data at Drammen, Norway.

The change from silty clay to plastic clay and then to lean clay might support the statistical boundaries at 4.7 m and 9.8 m, respectively. However, all other boundaries are unverified boundaries which indicate poor reliability in the analysis. At a window width equal to 2 m, three boundaries are found which do not match the borehole reference boundaries. Also, the depths of the boundaries using ICC of Q , ICC of B_q , and D^2 were different. At this site, both ICC and D^2 indicated improper site stratigraphy and the window width has a very significant effect on the number and elevation of the statistical boundaries.

3.4.2. ICC and D^2 at Lilla Mellösa, Sweden

Piezocone data at Lilla Mellösa test site (Larsson and Mulabdić, 1991) were analyzed using the single and multivariate statistical measurements ICC and D^2 , respectively, to demarcate different soil boundaries. The soil profile consists of organic clays overlying clays and separated at 5.5 m as depicted in Fig. 3.15 which also includes a typical piezocone sounding at the site. To define the correlation distance between the data, experimental autocorrelation functions of Q and B_q are determined as shown in Fig. 3.16. The points of $r_{\Delta h}$ were given higher weights where they were calculated at a relatively short distance between pairs of piezocone data.

Exponential and Gaussian models which give the least square errors compared with other models are chosen to fit the experimental $r_{\Delta h}$ of Q and B_q , respectively, as shown in Fig. 3.16. The parameters of the two models are summarized as follows: (1) the exponential model has nugget = zero, sill/variance = 1.9, and range = 12 m, and (2) the Gaussian model has nugget = zero, sill/variance = 2, and range = 12 m. Based on the

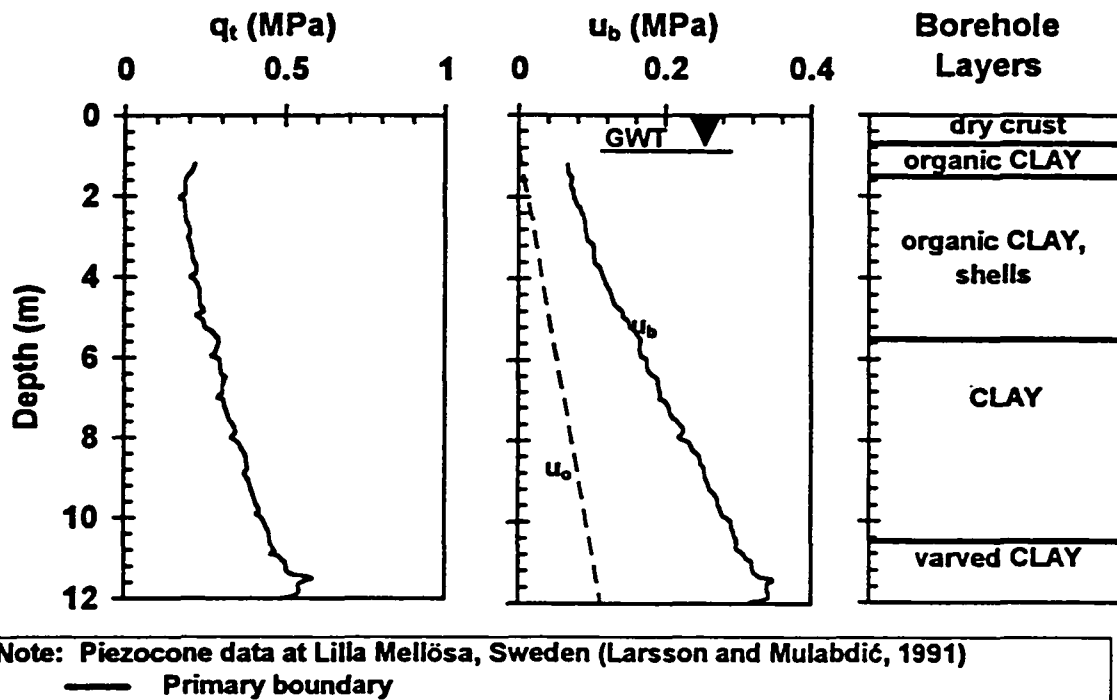


Figure 3.15. Piezocone Data and Soil Profile at Lilla Mellösa, Sweden.

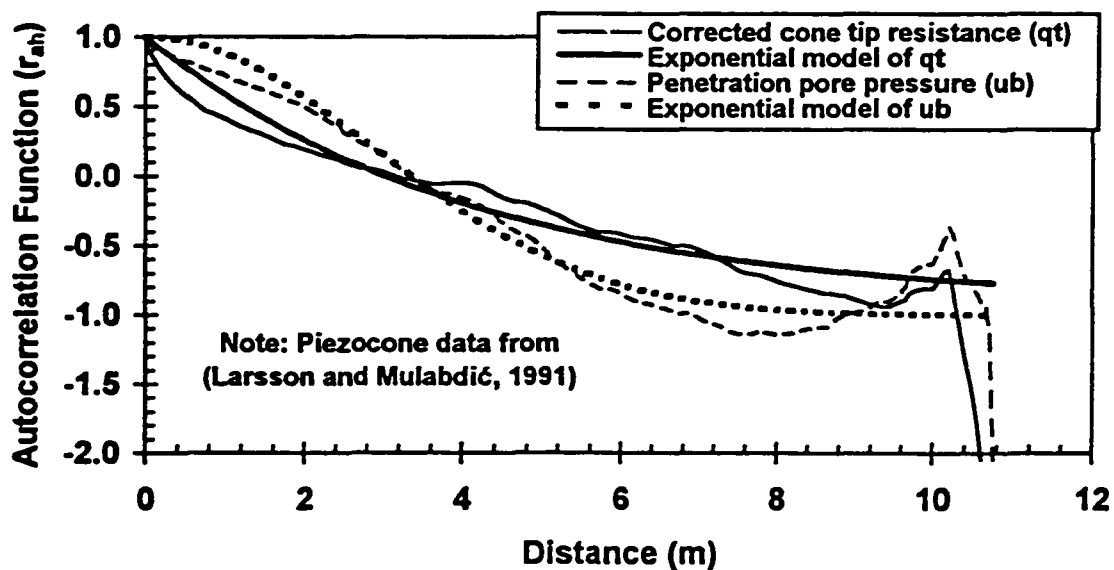


Figure 3.16. Autocorrelation Function of Piezocone Data at Lilla Mellösa, Sweden.

obtained ranges, the data are correlated between the depths of 1 m and 12 m and there is statistical similarity of the soil along the vertical profile, however, this disagrees with the soil stratigraphy depicted in Fig. 3.15.

3.4.2.2. Uncertainty of Soil Demarcation

The correlation distance (12 m) defined using the fitted models is too large to detect changes within the soil stratigraphy. Therefore, a window width of 1 meter is chosen based on a parametric study performed by Wickremesinghe (1989). Also, two other window widths equal to 0.5 m and 2 m are used to study the size effect of a window on the statistical boundaries. A window is moved from point to another on the vertical profile of each normalized parameter and ICC and D^2 are calculated at a window center. The determined profiles of ICC and D^2 are shown on Fig. 3.17 which also indicates one soil layer looking at the profiles of the derived parameters Q and B_q . By increasing the window width, the peaks of the ICC and D^2 profiles decrease. Soil boundaries are determined at the peaks of $ICC > 0.8$ and at $D^2 > [\text{average} + 1.65 \text{ standard deviation}]$ of a generalized distance profile with a level of significance equal to 0.05 (Hegazy et al., 1997).

Figure 3.18 shows a comparison between the borehole and the statistical soil boundaries of the three window widths. None of the performed analyses satisfies the soil stratigraphy at the site. Using a window width equal to 0.5 m, four statistical boundaries are detected and the one at a depth of 5.3 m supported by the change from organic clays to clays. None of the boundaries detected using a window width equal 1 m or 2 m matches the reference boundaries.

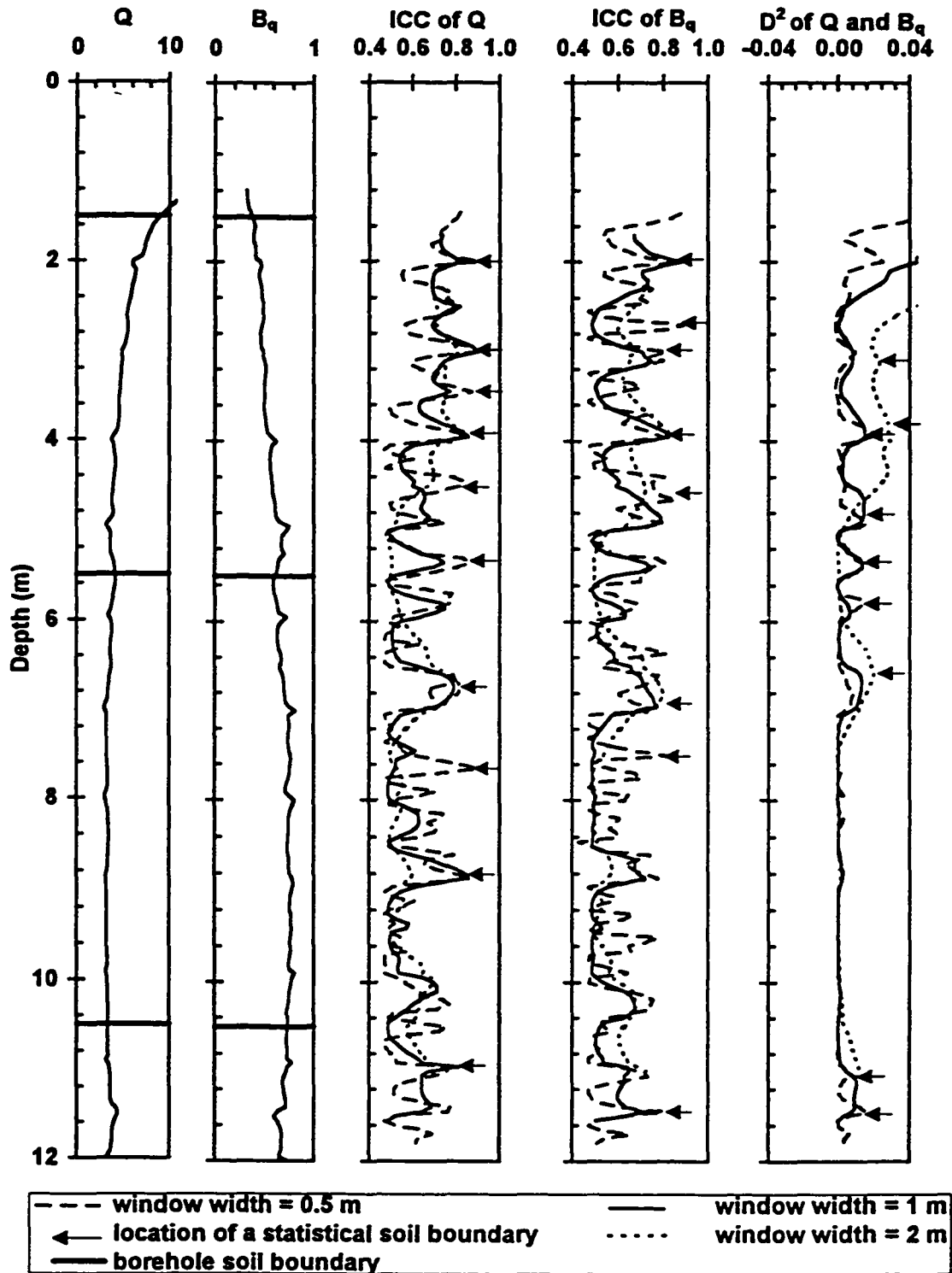


Figure 3.17. Effect of Window Width on ICC and D^2 Using Piezocone Data from Lilla Mellösa, Sweden.

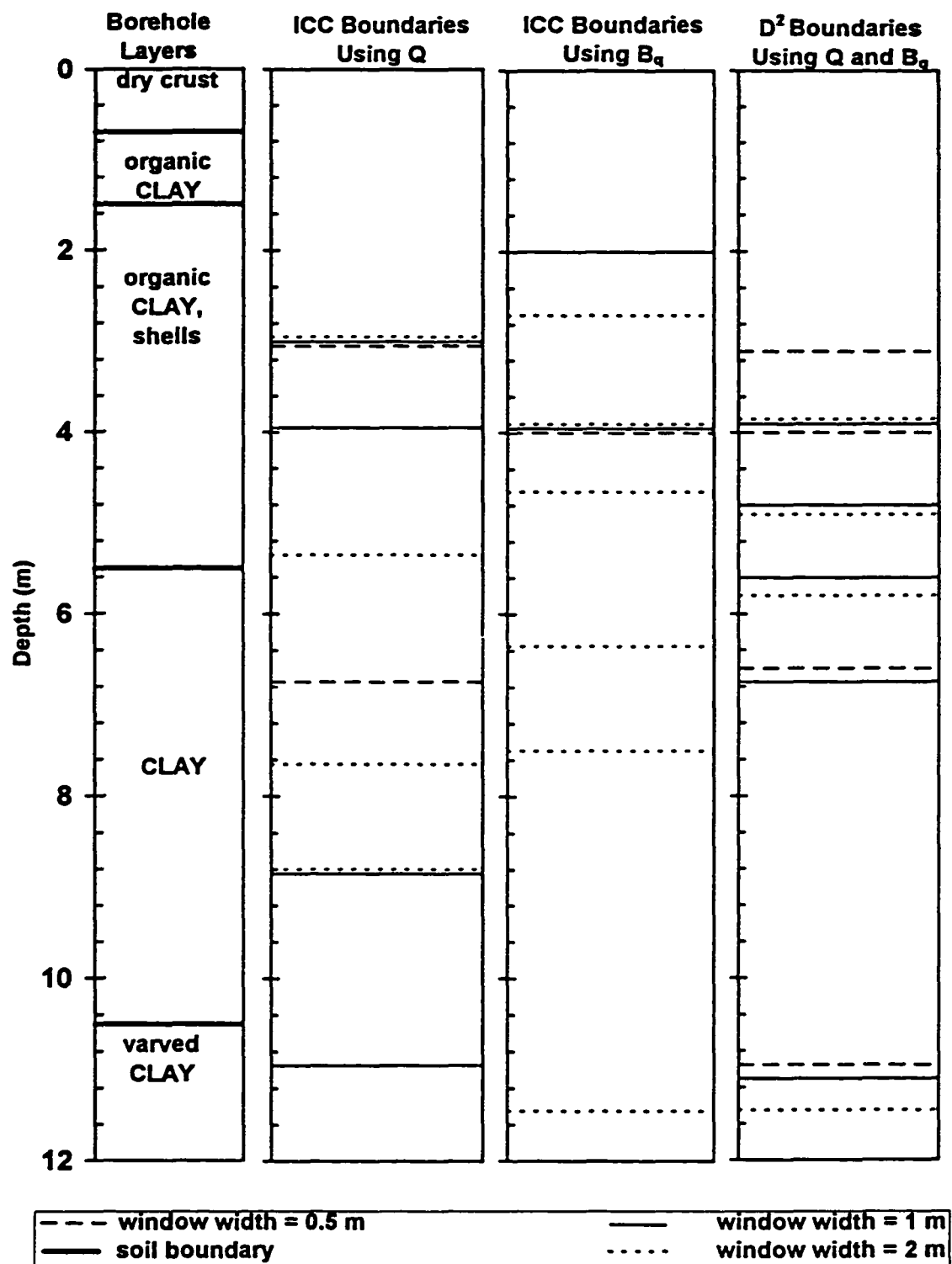


Figure 3.18. Statistical Soil Boundaries of ICC and D² Using Piezocone Data from Lilla Mellösa Sweden.

The profile of ICC (B_q) indicates 7 layers within the stratigraphy, however, the stratigraphy consists only of four layers between the depths of 1 m and 12 m. Using a window width equal to 1 m results in two boundaries at depths of 2 m and 3.95 m which have no support. The analysis using a window width equal to 2 m indicates two layers separated at 4 m, however, the boundary between the organic clays and the underlain clays is at 5.5 m. Moreover, the profile of the D^2 measurements detected 4 boundaries for each window width although none of them is supported by the primary reference boundaries. The statistical boundary at a depth of 5.8 m using a window width equal to 0.5 m is the closest to the borehole boundary at a depth of 5.5 m. Different boundary elevations were obtained using ICC of Q , ICC of B_q , and D^2 .

The chosen window width has a strong effect on the detected statistical boundaries. For too small of window width, more layers were detected than those in a subsurface stratigraphy. For too large of window width, some primary soil boundaries were missed. Choosing a window width equal to 1 m, as recommended by Wickremesinghe (1989), did not properly indicate the soil profiles at either the Drammen or Lilla Mellösa test sites.

3.5. Summary and Conclusions

The intraclass correlation coefficient (ICC) and the generalized distance (D^2) have been used in the statistical analysis of piezocone data to delineate soil stratigraphy. Both the ICC and D^2 methods were applied, for example, at Drammen, Norway, and Lilla Mellösa, Sweden. Neither method resulted in reliable soil boundaries as compared with

adjacent soil boring records and index data. Also, no information is given about the correlation of the soils between these boundaries. The number and locations of statistical soil boundaries depended on the chosen or determined window width using an autocorrelation function. A small window width gave more boundaries than expected. However, some boundaries were missed using too large of a window width.

For the cases examined, the defined soil boundaries using ICC of the normalized parameters Q and B_q did not match with each other nor with those indicated by D^2 . No information was given about the soil in the upper half of the first window and the lower half of the last window. The given boundaries depend on an arbitrary cut off value of ICC or D^2 . Therefore, a more objective and reliable method is needed in order to screen a certain soil stratigraphy.

A more consistent and reliable method termed *cluster analysis* is proposed and a review of clustering methods is given in Chapter 4.

CHAPTER 4

CLUSTER ANALYSIS

4.1. Introduction

The concept of cluster analysis is introduced to objectively define vertical data into similar groups in the soil profile, delineate different layer boundaries, and allocate the lenses and outliers within a sublayer. Different clustering methods are reviewed and the hierarchical techniques are found to be more feasible for the purpose of this study. The cluster analysis requires a selection of variables, standardization methods, similarity measurements, type of cluster method, and number and validity of clusters. These problems are evaluated and recommendations are given herein specific for clustering piezocone data. Moreover, the method can be used for any types of in-situ vertical data for the same purpose (e. g., flat dilatometer and vane shear), as well as multi-channeled soundings with greater numbers of independent readings (e.g., resistivity piezocone).

4.2. Cluster Analysis Definition and Criterion

Cluster analysis is a statistical method for grouping data which have similar mathematical descriptions (Kendall, 1966; Everitt, 1974; Romesburg, 1984). Clustering analysis can be applied to data belonging to the same group in order to divide them into sections based on a predetermined properties (Kendall, 1966). Hartigan (1996) suggested

the term “classification” interchangeably with the term “cluster” because it is familiar in different science fields such as medicine, biology and engineering. Everitt (1974) summarized the uses of cluster analysis, including:

- determining a true typology (systematic classification of types that have characteristics in common),
- model fitting,
- estimation based on groups,
- hypothesis generation and testing, and
- data reduction and exploration.

In this study, soil exploration by piezocone data reduction will be the primary use of the cluster analysis. The method can be easily applied to other types of in-situ geotechnical test data too, such as standard penetration test, flat dilatometer test, and vane shear test. It is especially valuable for the evaluation of piezocone data because these soundings provide continuous records of tip resistance (q_t), and penetration porewater pressure (u_b), and in some instances, additional measurements including soil conductivity and intensity fluorescence signal (Robertson et al., 1998).

4.3. Current Applications of Cluster Analysis

Cluster analysis has been used extensively in taxonomy and biology (e. g., Sokal, 1974), chemistry and medicine (e. g., Solberg et al., 1976), and sociometry and social psychology (e. g., Burt, 1988). More recently, a preliminary application of clustering in

image analysis processing was discussed by Russ (1995). For instance, biologists used it to classify different types of animals and plants having a very wide variety of kinds and properties that can be due to inherent random variability. The analysis includes a large number of data.

Mlynarek and Lunne (1987) performed a statistical study for identification of homogeneity of soil strength and physical properties of a clay deposit based on hierarchical clustering of cone data. Clustering results of their preliminary study were found promising for more comprehensive evaluation in geotechnical applications. In this study, piezocone data are evaluated by cluster analysis to obtain quantitative delineation of the boundaries between different soil types and to define the locations of the thin layers (lenses) and outliers within thick layers. Piezocone data include the inherent soil variability due to differences in soil type, geological formation, thicknesses of sublayers, age and stress history, existence of lenses and seams, soil fabric, and other factors.

Piezocone data usually have high frequency in the vertical direction. For instance, data are collected at 0.01 to 0.05 m intervals. Therefore, there is a similarity in the analogy between piezocone data and other fields of applications of the cluster analysis because large data sets are analyzed to indicate the inherent random soil variability.

4.4. Cluster Analysis Components

Milligan (1996) summarized the steps of applying cluster analysis in research and application as shown in Fig. 4.1. Each step is discussed in detail in the following sections.

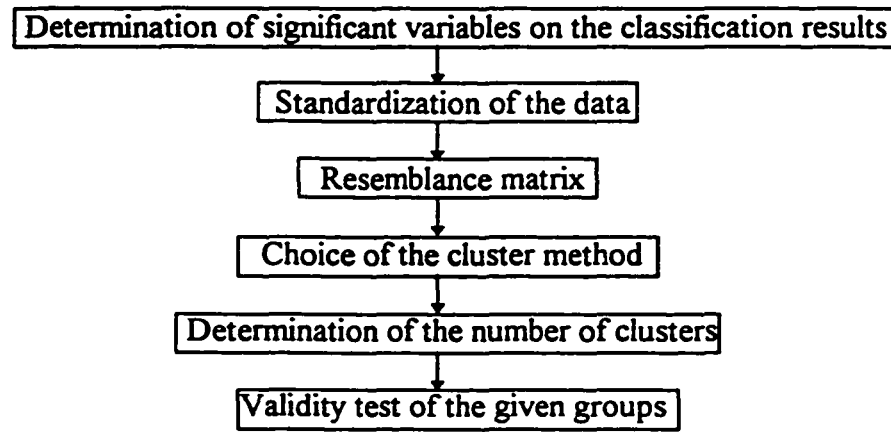


Figure 4.1. Steps of Clustering Analysis.

4.5. Selection of Variables for Cluster Analysis

All significant variables should be included in the cluster analysis to better indicate grouping of analyzed data (Milligan, 1996). Applying statistical methods such as principal components, multivariate analysis, or factoring techniques were not successful in helping to decide the variables added or excluded from the cluster analysis. In this study, two piezocone readings namely, q_t and u_b , are considered for the purpose of soil stratigraphy. The tip resistance readings can distinguish between different soil types and the variation in the strength within the same layer such as the differences between loose and dense sands. The pore pressure readings can also indicate the change between major soil types such as sands and clays. Moreover, pore pressure readings help in the identification of soil lenses within a soil stratum. For example, the peaks within a sand layer indicate clay lenses and the troughs within a clay layer indicate sand lenses. Therefore the two readings are important for the purpose of soil stratigraphy. The two readings are used in this study

in normalized formats $Q = (q_t - \sigma_{vo})/\sigma_{vo}'$ and $B_q = (u_b - u_o)/(q_t - \sigma_{vo})$, as proposed by Wroth (1984; 1988) for the interpretation of soil behavior based on piezocone data. The tip resistance increases by increasing depth and soil vertical stresses even within the same soil layer therefore its normalization in the form of Q helps to reduce this effect. The variation of Q with depth is relatively more related to the variation of soil type and/or behavior (Wroth, 1988; Robertson, 1991). The hydrostatic porewater pressure is a component of the penetration pore pressure but it is usually independent of the changes in the soil profile, and the excess pore pressure represents the variation of soil type and/or behavior. The division of the excess pore pressure ($u_b - u_o$) over the net tip resistance ($q_t - \sigma_{vo}$) helps in better identification of the differences between different soils in a geostratigraphy because q_t and u_b are inversely proportional in the case of primary change in the soil type during penetration. For example, in the case of a clay deposit over a sand deposit, the u_b readings decrease and the q_t readings increase in the sand layer compared with those in the upper clay layer. The differences in the profiles of B_q in the upper and lower layers are relatively more obvious than those between u_b measurements of the same layers. The two derived normalized parameters Q and B_q are functions of soil type and behavior (Wroth, 1984, 1988; Robertson, 1991).

Three separate stresses are measured in piezocone testing (q_t , u_b , and f_t). Lunne et al. (1986) and Tanaka (1995) studied the repeatability of the three piezocone measurements using different commercial penetrometers. They found that q_t and u measurements at the same depth were relatively repeatable for most cone penetrometers.

However, a large scatter of the f_t data at the same depth was observed and indicated the dependency of the measurements on the specific commercial type of cone.

The effect of f_t readings on the cluster analysis is studied at the Amherst test site in Massachusetts and shown to cause incorrect interpretation of the soil stratigraphy as discussed in chapter 5. Therefore, f_t measurements have not been included in the cluster analysis of sites considered herein. However, if the reliability of f_t measurements is improved, clustering can easily accommodate the third reading, as well additional measurements such as resistivity, dielectric readings, or other parameters.

Robertson (1991) proposed a classification method including two charts using: (1) Q and F , and (2) Q and B_q . The chart based on the derived normalized parameters, Q and B_q , is used to indicate different soil layers at the analyzed sites. In order to use these two parameters, the groundwater table (GWT) and the total unit weight (γ_t) must be known, or else assumed.

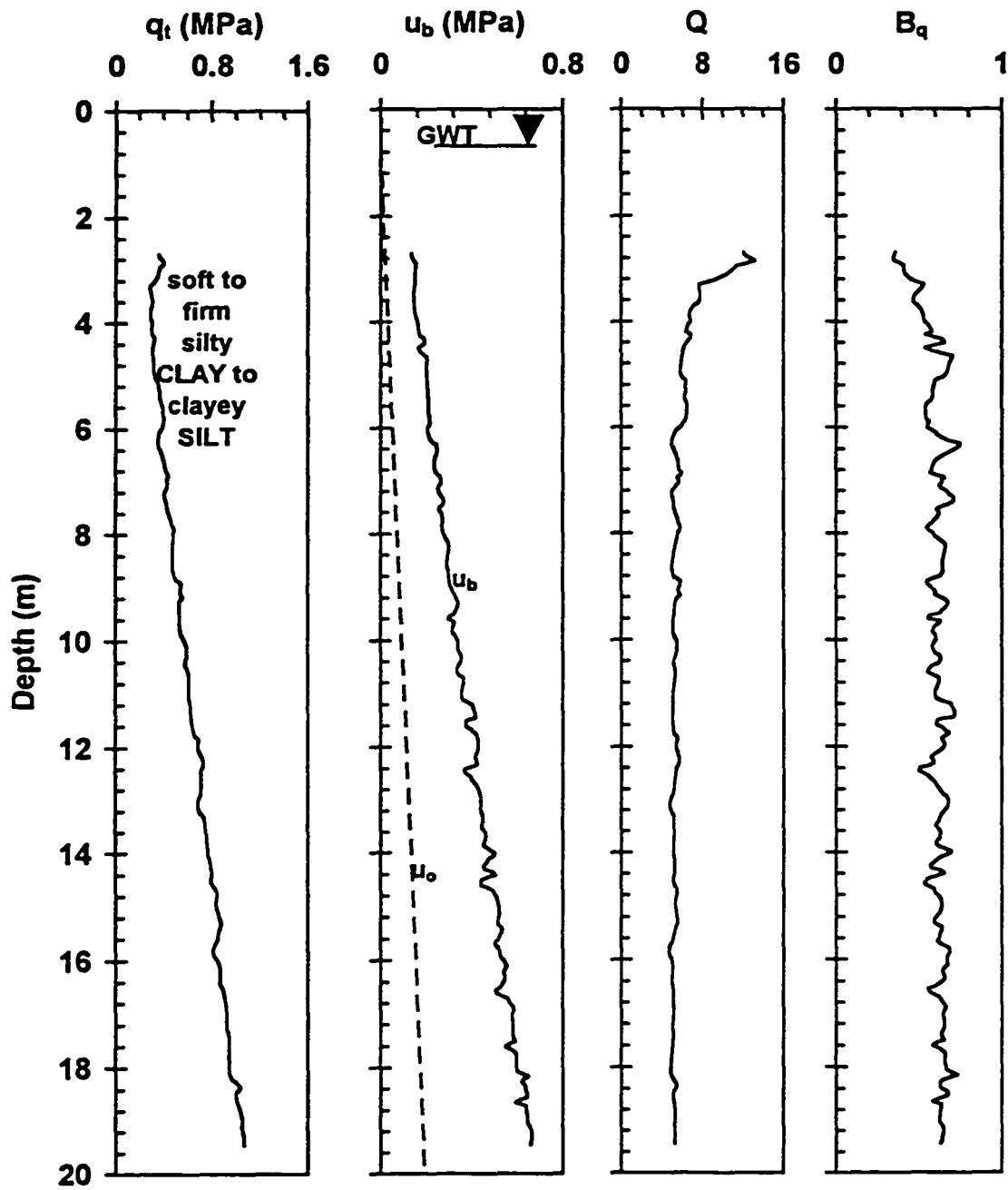
4.6. Standardization of the Data

Standardization is a means to represent the data in dimensionless format to eliminate the effect of using different units for different variables (Romesburg, 1984). Standardization also decreases, for instance, the influence of q_t or Q measurements, which always have larger values than u_b or B_q readings at the same depth, respectively, on the classification results. Standardization is not essential, but it is a logical or an optional step in cluster analysis (Milligan, 1996). The relative influence of different parameters used in clustering can be correlated with the average and standard deviation of each one of them.

The more the value of the mean and standard deviation of a parameter, the more its influence on the obtained groups of data. Therefore, Standardization can be essential in the case of clustering piezocone data to stratify the soil profile because the means and variances of the unprocessed measurements (or the two normalized parameters) are quite different.

For instance, results from a piezocone sounding at a uniform soft clay site at Bothkennar, UK (Nash et al., 1992) are shown in Fig. 4.2. The means and variances of the unprocessed piezocone readings (q_t and u_b) and the normalized parameters (Q and B_q) of the layer between 2.5 m and 19.5 m are summarized in Table 4.1. The ratio [mean (q_t)/mean (u_b)] is equal to 1.7 and the ratio [variance (q_t)/variance (u_b)] is equal to 2.3. While the ratio [mean (Q)/mean (B_q)] is equal to 9.3 and the ratio [variance (Q)/variance (B_q)] is equal to 39.3. Seven common forms of standardization (Romesburg, 1984 and Milligan, 1996) are summarized in Table 4.2. In Table 4.2, z_{ij} is a standardized value of a certain cone variable X_j (for example, q_t) at a certain depth i , x_{ij} is a cone reading, $\text{Max}(X_j)$ is the maximum of a cone variable such as q_t , $\text{Min}(X_j)$ is the minimum of a cone variable, $E(X_j)$ is the average of a cone variable, and $\text{SD}(X_j)$ is the standard deviation of a cone variable. The average and standard deviation are in the following forms:

$$E(X_j) = \text{Mean}(X_j) = \frac{\sum_{i=1}^n x_{ij}}{n} \quad (\text{Equation 4.1})$$



Note (1): $Q = (q_t - \sigma_{vo}) / \sigma_{vo}'$

Note (2): $B_q = (u_b - u_o) / (q_t - \sigma_{vo})$

Figure 4.2. Piezocone Results in a Uniform Soft Clay at Bothkennar, Scotland (Piezocone Data from Nash et al., 1992).

Table 4.1. A comparison between Means and Variances of Piezocone Parameters at Bothkennar, Scotland (Piezocone Data from Nash et al., 1992).

Cone parameter	Mean	Variance
q/p_a^*	6.4	5.48
u_0/p_a^*	3.8	2.36
Q	5.6	1.57
B_q	0.6	0.04

(*) p_a = unit atmospheric pressure $\cong 0.1$ MPa

Table 4.2. Different Methods of Data Standardization [adapted from Romesburg (1984) and Milligan (1996)].

Standardization Method	Formula
None (data in the original format)	$z_{ij} = x_{ij}$
Zscore	$z_{ij} = \frac{x_{ij} - E(X_j)}{SD(X_j)}$
Range	$z_{ij} = \frac{x_{ij}}{\text{Max}(X_j) - \text{Min}(X_j)}$
Rescale	$z_{ij} = \frac{x_{ij} - \text{Min}(X_j)}{\text{Max}(X_j) - \text{Min}(X_j)}$, where $1 \geq z_{ij} \geq 0$
Max	$z_{ij} = \frac{x_{ij}}{\text{Max}(X_j)}$, where $z_{ij} \leq 1$
Mean	$z_{ij} = \frac{x_{ij}}{E(X_j)}$
Stdev	$z_{ij} = \frac{x_{ij}}{SD(X_j)}$

Notes: (1) x_{ij} is a cone reading (e. g., q_i at a certain depth)
(2) X_j is a cone variable (e. g., q_i at any depth)
(3) z_{ij} is a standardized value

$$\text{Stdev}(X_j) = \sqrt{\frac{\sum_{i=1}^n [x_{ij} - E(X_j)]^2}{n - 1}} \quad (\text{Equation 4.2})$$

in which n is the number of data points collected with depth. Note that the standard deviation is the square root of the variance.

The seven available standardization methods are applied to piezocone data at Amherst, Massachusetts (this study). A representative piezocone sounding and the soil profile at the site are shown in Fig. 4.3. The soil at Amherst consists of a clay fill over a silty clay crust, which is underlain by a thick soft varved clay layer (Lally, 1993). Figure 4.4 shows the derived normalized parameters Q and B_q . Figure 4.5 compares the standardized Q and B_q using the other six methods. Looking at the standardized Q , soil boundaries can be defined at approximate depths of 1 m, 2 m, 2.5 m, 3 m, 4 m, and 5.5 m. Note that the standardization using the zscore method and the mean methods can better indicate the difference between the fill and the crust above 4 m and the lower soft clay layer down to a depth of 14.5 m than the other methods. In the case of the zscore method, the soil above 4-m depth is denoted by positive zscores, however, the soil below 4-m depth is denoted by negative zscores. In the case of the mean method, there is a relatively wide range of the means of Q s between 2 and 13 in the upper 4 meters, however, the mean of Q s in the soft clay layer below a depth of m is almost constant and equal to 1. For the stdev method, the standardized data are in the range between 0 and 4 and for the range, rescale, and max methods, the range of the standardized data is between

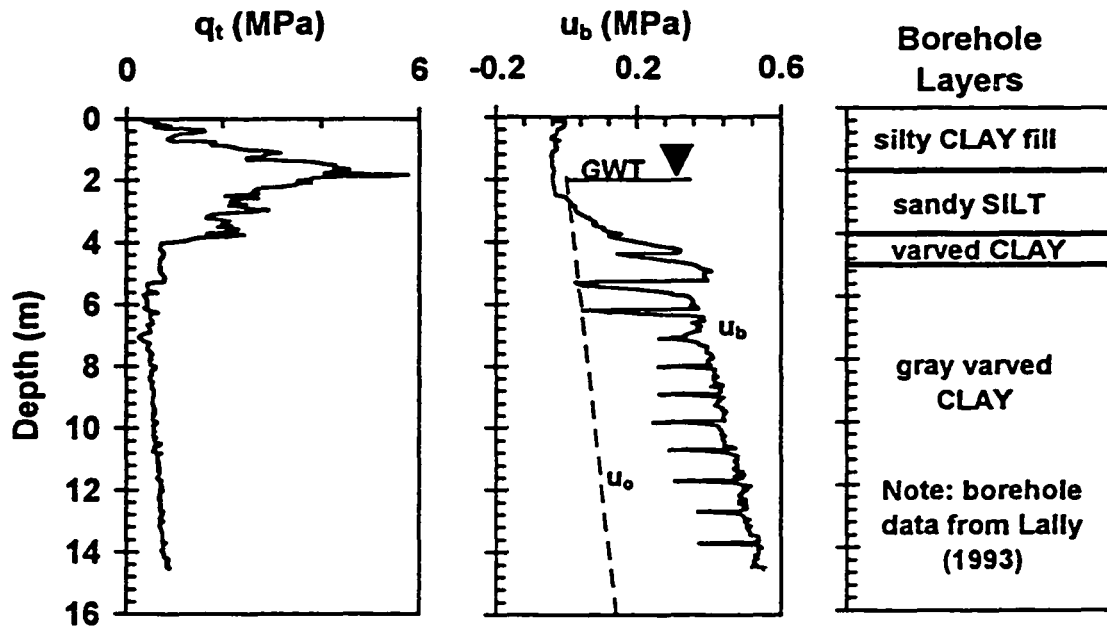


Figure 4.3. Piezocone Data and Soil Profile at Amherst, Massachusetts
(Data Collected by the Author During This Study)

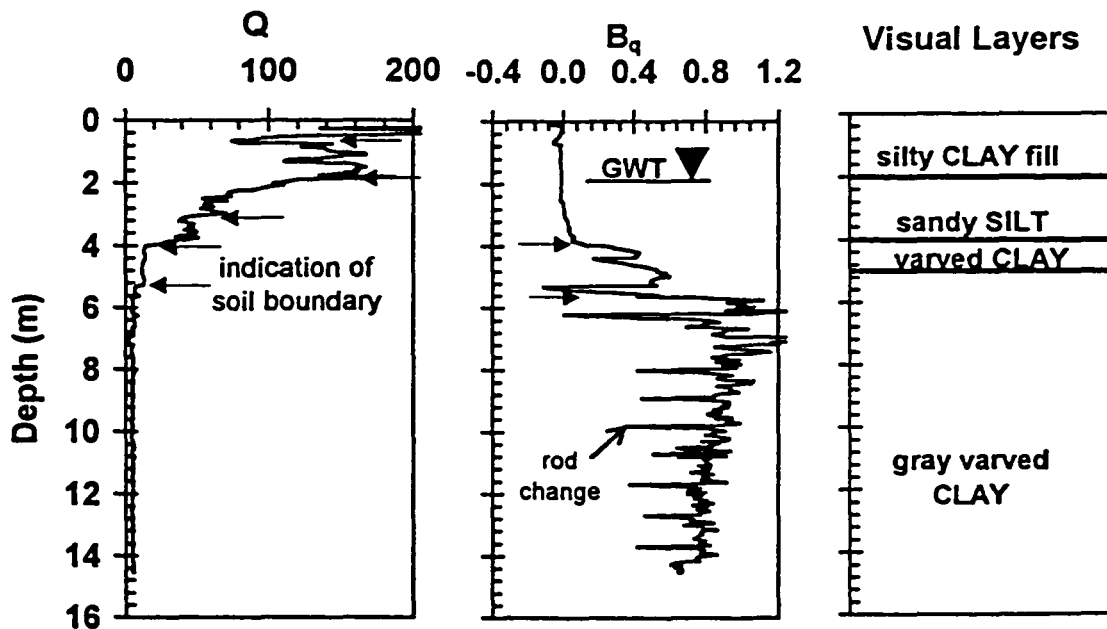


Figure 4.4. Normalized Piezocone Parameters Q and B_q and Soil Profile
at Amherst, Massachusetts (Data from This Study).

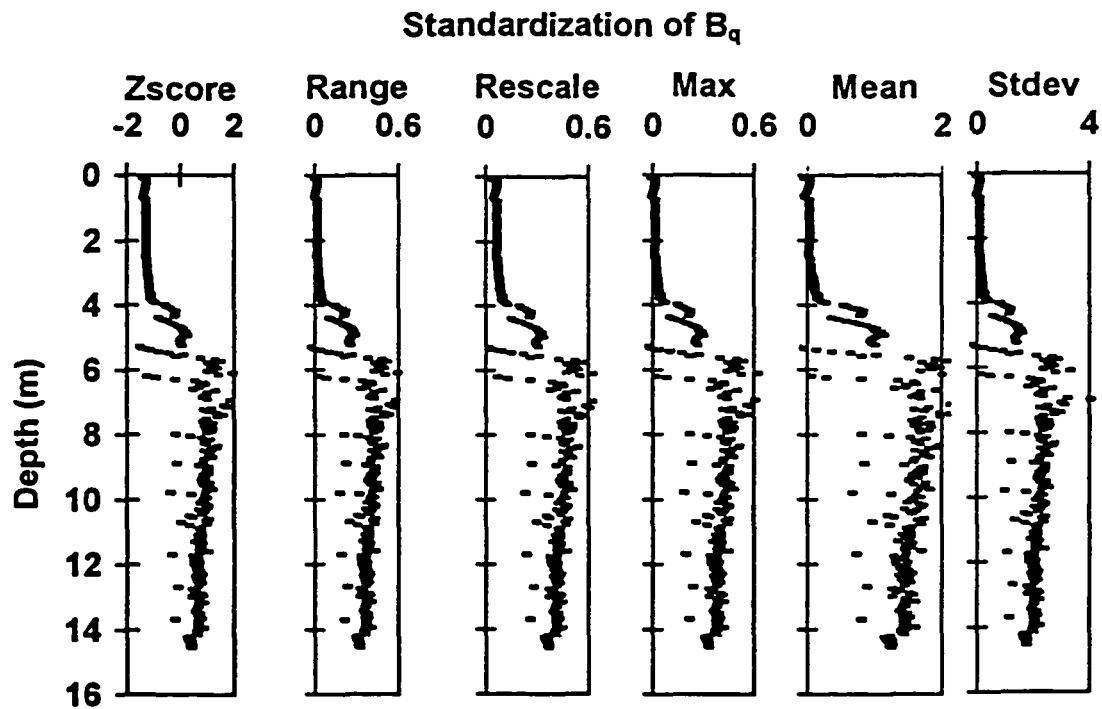
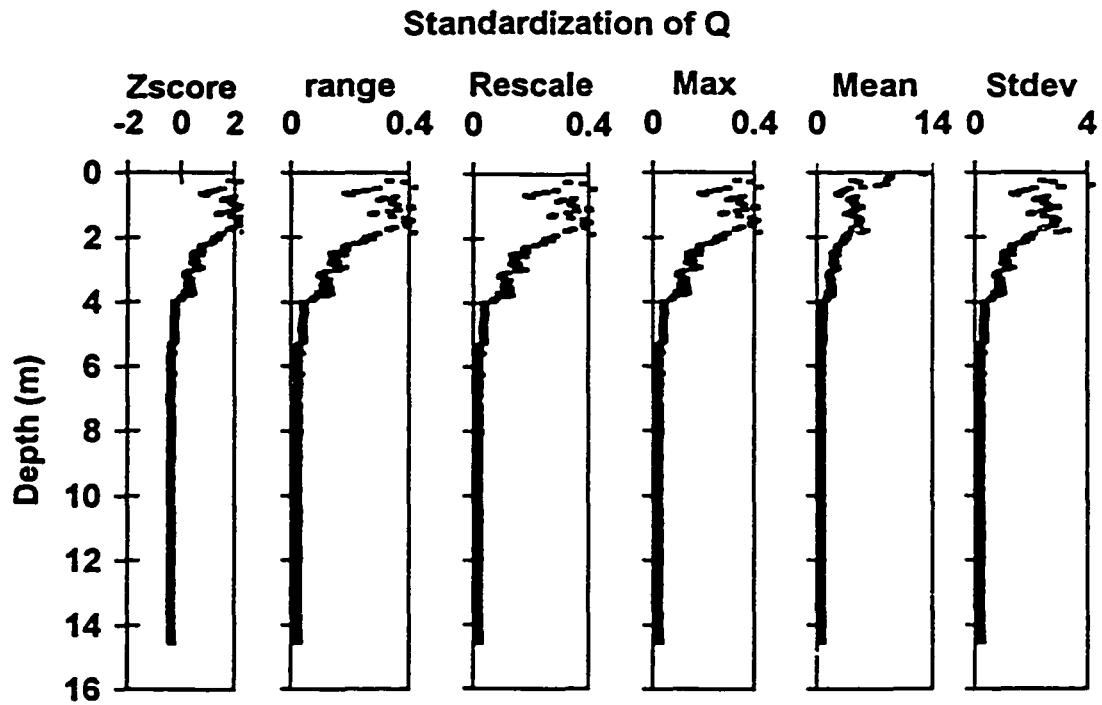


Figure 4.5. Comparison of Different Standardization Methods at Amherst, Massachusetts (Piezocone Data from This Study).

0 and 0.4. Therefore, using zscore or mean methods for standardization of the Q are preferred in this example.

Looking at the standardized B_q , two primary layers are identified with a boundary at 4-m depth. Using the zscore method clearly indicates the difference between the soil above 4 m which has negative zscores and the lower soft clay layer which has positive zscores. In the other 5 methods, the standardized data are positive. The stdev method has a larger range between 0 and 4 of the standardized records than that of the range, rescale, max and mean methods, therefore, it is relatively better to identify the soil stratigraphy. Using zscore and stdev methods are preferred for standardization of B_q at Amherst test site. The zscore technique is suitable for the standardization of both Q and B_q at this site.

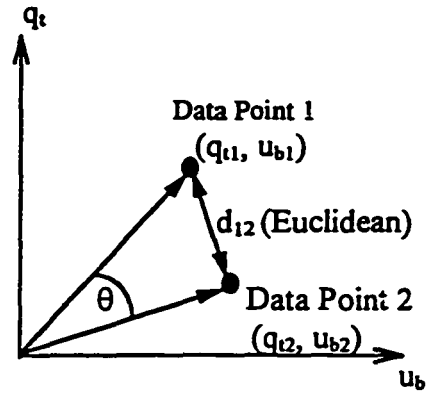
Milligan and Cooper (1988) performed a simulation study on 864 artificial data sets generated from multivariate normal distributions with uncorrelated variables to evaluate different methods of standardization using different cluster methods. The range and rescale standardization formulas are found to have a superior ability to recover predefined groups in the simulated data. However, Milligan (1996) noted that yet a generalized rule has not been established and standardization is optional based on the characteristics of the analyzed data. A parametric study is performed to choose a standardization method of the cone data used in a cluster analysis as discussed in this chapter, and the zscore method is superior to properly indicate a soil stratigraphy at a minimum cluster number. The zscore, for example, for Q is defined in the following format:

$$Zscore = \frac{Q_i - E(Q)}{SD(Q)} \quad \text{(Equation 4.3)}$$

in which Q_i is a normalized tip resistance at a certain depth i , $E(Q)$ is the average of all Q measurements, and $SD(Q)$ is the standard deviation of all Q measurements.

4.7. Resemblance (Distance) Matrix

The resemblance (distance) is a statistical measurement of the similarity between data points. Clustering of a data set depends on the similarity between different records. The most similar data points are grouped in one cluster. Piezocone data can be visualized, for instance, in q_t and u_b space as shown in Fig. 4.6. The two cone measurements at a certain depth (i) forms a vector (q_{ti}, u_{bi}) . Each element in the resemblance matrix can be viewed as a distance or a scale to measure the similarity between two vectors of piezocone data at two different depths. For instance, Euclidean and cosine resemblance measurements are depicted in Fig. 4.6. The former is the magnitude of a vector between two pairs of cone data, and the latter is defined as a cosine of an angle between two vectors in q_t and u_b space. Another way to look at the piezocone data is to put them in a matrix (D) format as follows:



Euclidean distance: $d_{12} = [(q_{t1} - q_{t2})^2 + (u_{b1} - u_{b2})^2]^{0.5}$

$$\text{Cosine: } \cos(\theta) = \frac{q_{t1}q_{t2} + u_{b1}u_{b2}}{(q_{t1}^2 + u_{b1}^2)^{0.5}(q_{t2}^2 + u_{b2}^2)^{0.5}}$$

Figure 4.6. Definitions of the Euclidean and the Cosine Measurements in the “Octahedral” Space of Piezocone Data.

$$D = \begin{bmatrix} (q_t)_{11} & (u_b)_{12} \\ (q_t)_{21} & (u_b)_{22} \\ \vdots & \vdots \\ (q_t)_{n1} & (u_b)_{n2} \end{bmatrix} \quad (\text{Equation 4.4})$$

in which the readings in a row are at the same depth. The distance of resemblance is calculated between pairs of data [e.g., $\{(q_t)_{11}, (u_b)_{12}\}$ and $\{(q_t)_{21}, (u_b)_{22}\}$] at different depths. Table 4.3 includes a summary of eight available methods for calculating the distance matrix. In Table 4.3, var_k is the variance of a variable X_j (e.g., q_t) in the following form:

$$\text{var}_k = \text{Variance}(X_j) = \frac{\sum_{i=1}^n [x_{ij} - E(X_j)]^2}{n - 1} \quad (\text{Equation 4.5})$$

For example, four points of Q and B_q are chosen from a piezocone sounding at Amherst, Massachusetts (see Fig. 4.5). Two points are from the crust clay layer between 2 m and 4 m and two points are from the soft clay layer below 4 m. The points are selected to represent two different layers and listed with their depths in Table 4.4. The similarity between different points is measured using the eight methods listed in Table 4.3

Table 4.3. Summary of Methods for Determining the Resemblance Matrix
[Modified from Noursis (1993)].

Type of Distance	Formula	Reference
Pearson	$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2 / \text{var}_k}$	Sokal & Rohlf (1962)
Block (Manhattan Distance)	$d_{ij} = \sum_{k=1}^n x_{ik} - x_{jk} $	Carmichael & Sneath (1969)
Minkowski	$d_{ij} = \left(\sum_{k=1}^n x_{ik} - x_{jk} ^p \right)^{\frac{1}{p}}$	Jardin & Sibson (1971)
Power (p,r)	$d_{ij} = \left(\sum_{k=1}^n x_{ik} - x_{jk} ^p \right)^{\frac{1}{r}}$	Jardin & Sibson (1971)
Euclidean	$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$	Everitt (1974)
Cosine	$d_{ij} = \frac{\sum_{k=1}^n x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^n (x_{ik})^2 \sum_{k=1}^n (x_{jk})^2}}$	Romesburg (1984)
Squared Euclidean	$d_{ij} = \sum_{k=1}^n (x_{ik} - x_{jk})^2$	Everitt (1984)
Chebychev	$d_{ij} = \text{Max} x_{ik} - x_{jk} $	Noursis (1993)

Notes: (1) d_{ij} is the similarity between two vectors, x_{ik} and x_{jk} , of two cone readings (e. g., q_0) at different depths i and j
(2) p and r are integers.

and the results are summarized in Table 4.5. In all methods except the cosine method, the smallest statistical distance between pairs of data at different depths indicates the most similarity amongst the analyzed data. For instance, using Pearson method, Q and B_q of points 3 and 4 at depths of 6 m and 10 m, respectively, are the most similar measurements.

Table 4.4. Four Selected pairs of Q and B_q of PCPT1 sounding at Amherst, Massachusetts (Piezocone Data from this study).

Point number	Depth (m)	Q	B _q
1	2.0	103.57	-0.011
2	2.5	61.18	-0.016
3	6.0	5.21	0.996
4	10.0	4.62	0.834

Table 4.5. Summary of the similarity measurements between four selected points of Q and B_q of PCPT1 sounding at Amherst, Massachusetts (Data from this study).

Data points	Pearson	Block	Minkowski p = 2	Power p = 1 r = 2	Euclidean	Cosine	Squared Euclidean	Chebychev
1 and 2	0.88	42.39	42.39	6.51	42.39	1.000	1796.77	42.39
1 and 3	2.77	99.37	98.36	9.97	98.37	0.982	9675.81	98.36
1 and 4	2.59	99.79	98.94	9.99	98.95	0.984	9790.73	98.94
2 and 3	2.21	56.99	55.97	7.55	55.98	0.982	3133.91	55.97
2 and 4	1.97	57.41	56.56	7.58	56.56	0.984	3199.32	56.56
3 and 4	0.30	0.75	0.59	0.86	0.61	1.000	0.37	0.58

Also, Q and B_q of points 1 and 3 at depths of 2 m and 6 m are the most dissimilar measurements. In the case of cosine measurement, a relatively higher cosine measurement indicates more similarity between pairs of cone data at different depths. A cosine value of 1 indicates maximum similarity and a cosine value of -1 indicates absolute dissimilarity. For example, points 1 and 2 have maximum similarity as well as points 3 and 4, however, points 1 and 3 have minimum similarity as well as points 2 and 3.

Milligan (1996) claimed that there is no specific rule or theory to base the choice of a similarity distance on. The selection of a distance measurement depends on the type

of data. Seven standardization methods and eight similarity measurements are discussed. A parametric study is performed to examine the 56 available combinations and compare the given soil layers at a certain test site with the determined layers using cluster analysis. Three sites are chosen for this study, namely, (1) Amherst, Massachusetts (this study), (2) McDonald's Farm, British Columbia, and (3) Fort Road, Singapore. Each of these sites have different soil conditions and a different number of soil layers.

The soil at Amherst, Massachusetts (Lally, 1993) consists of a clay fill over a silty clay crust, which is underlain by a thick soft varved clay layer as shown in Fig. 4.3. The soil stratigraphy at McDonald's Farm (Robertson, 1982) consists of a top soft clay layer, sand and silty sand in turn, underlain by a deep deposit of soft clay. The soil deposit at the Fort Road site (Chang, 1991) consists of a soft marine clay and intermediate silt layer. Figures 4.7a and b show the piezocone data and soil stratigraphies from borings at the three sites. These sites were chosen for this parametric study because a visual interpretation of the piezocone data at each site clearly matches with a defined soil boundary based on nearby borehole information. The piezocone data at each site are analyzed using the available 56 cluster methods combinations and the results are discussed in greater detail in Appendix B.

The zscore standardization and the cosine similarity measurements are concluded to give the best indication of the soil stratigraphy at a minimum cluster number at each of the three sites.

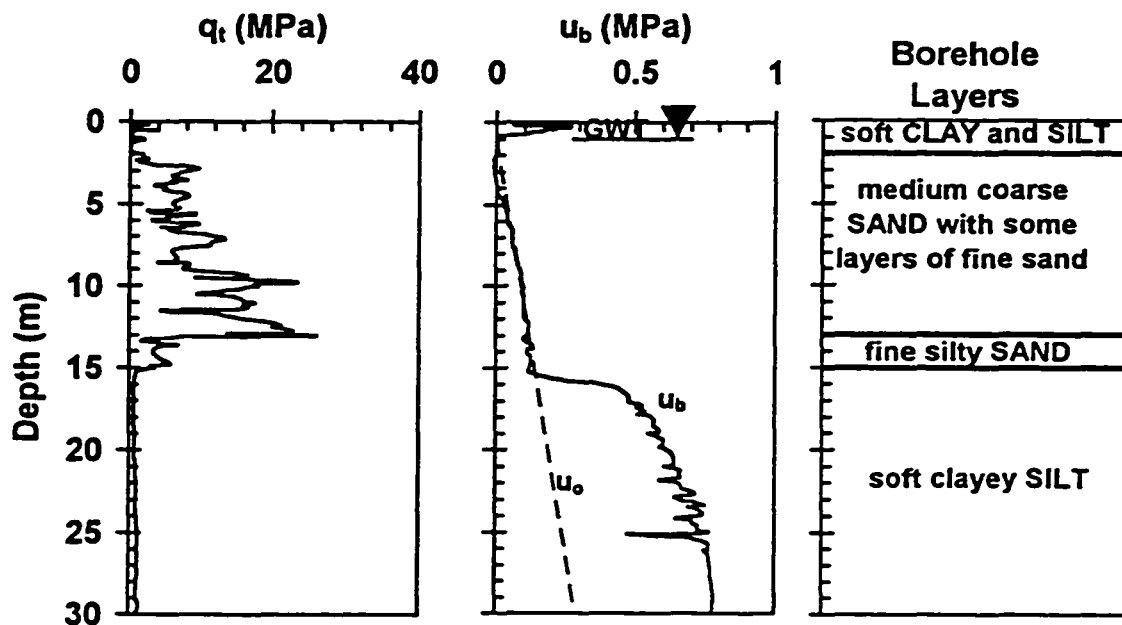


Figure 4.7a. Piezocone Data and Soil Stratigraphy at McDonald's Farm, British Columbia from (Data from Robertson, 1982).

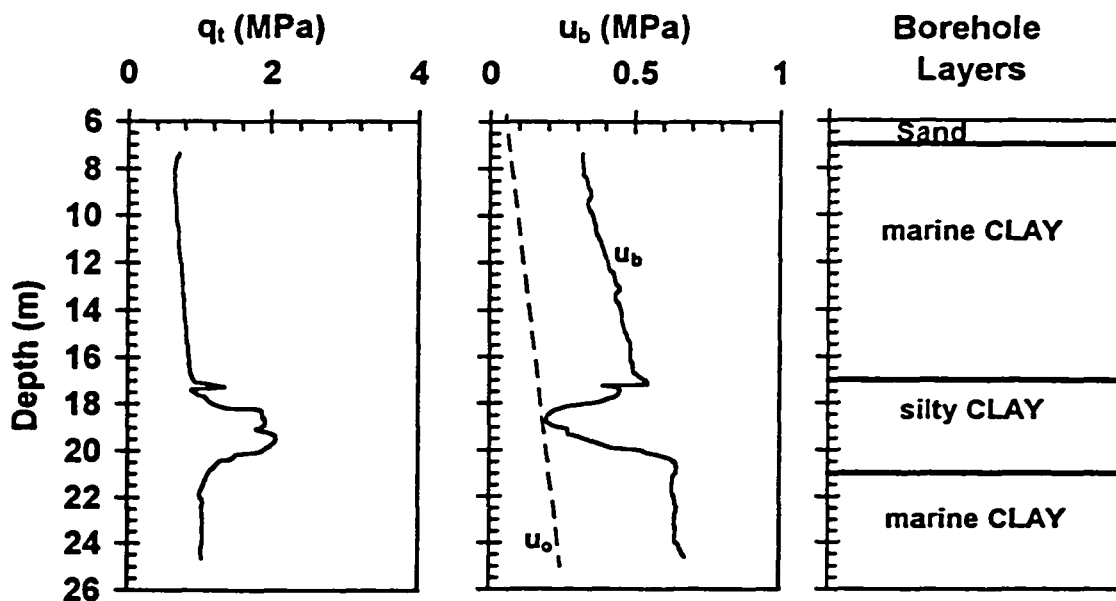


Figure 4.7b. Piezocone Data and Soil Stratigraphy at Fort Road, Singapore (Data from Chang, 1991).

4.8. Clustering Techniques

After data standardization and the calculation of the similarity between different pairs of measurements, the data are divided into correlated groups. Two piezocone vectors would be combined in one cluster if the distance between them is smaller than the distance between either vector and the rest of the vectors. A simple example to visualize this concept is shown in Fig. 4.8. Milligan (1996) reported that there is no single method that can be applied for all types of data and there is no unifying theory of clustering that has been widely-accepted. However, he summarized general rules to aid in choosing the best available clustering method. The chosen method should recover the inherent classification groups, be insensitive to the errors or outliers in a data set, and be automated.

There are several clustering techniques, such as hierarchical, optimization-partitioning (k-means), density, and clumping (overlapping) techniques (e. g., see Everitt, 1974; Arabie and Hubert, 1996; Milligan, 1996). The development of these methods and their limitations is discussed herein.

4.9. Hierarchical Techniques

Hierarchical clustering techniques can be divided into two main groups, namely, agglomerative and divisive methods (Everitt, 1974). In the agglomerative criteria, each datum of the analyzed observations begins in a separate cluster. For example, in the case of a data set of n observations, there are n individual clusters. In the first step, the two observations closest together are joined. In the next step, either a third observation joins

NOTE TO USERS

Page(s) not included in the original manuscript and are unavailable from the author or university. The manuscript was microfilmed as received.

98 - 104

This reproduction is the best copy available.

UMI

partitions (P) of a number of data set (N) into (g) groups be calculated using the following formula:

$$P(N, g) = (g^N - \sum_{i=1}^{g-1} g_{(g-i)} P(N, i)) / g \quad (\text{Equation 4.7})$$

in which, $g \geq 2$, $N \geq g$, $P(N, g)$ is the number of distinct partitions containing exactly g clusters and $g_{(g-i)} = g(g-1)(g-2)\dots(g-i+1)$. Everitt (1974) gave an example as follows: if $N = 19$, and $g = 8$, then the number of possible partitions $P = 1709 \times 10^{12}$. Therefore, the partitioning techniques are not considered feasible, especially with large data sets, if a complete optimization is required, even after the attempt of reducing the number of studied partitions of a data set (Gower, 1967; Scott and Symon, 1971). Moreover, Fisher and Van Ness (1971) evaluated the partitioning cluster analysis and concluded unsatisfactory performance in detecting the underlying clusters. Milligan (1996) suggested that the number of clusters should be specified in advance before performing the k-means clustering which brings subjectivity in the analysis. Based on this discussion, using partitioning techniques is not recommended for clustering piezocone data for the purpose of this study.

4.11. Other Clustering Methods

Clumping techniques: In this method, overlapping would be allowed between different clusters. First the data are divided into two groups by minimizing, for example, a cohesion function proposed by (Parker-Rhodes and Jackson, 1969) as follows:

$$G(A) = \frac{S_{AB}}{S_{AA}} \left[\frac{n_A (n_A - 1)}{S_{AA}} - \frac{S_{AA}}{p[n_A (n_A - 1)]} \right] \quad (\text{Equation 4.8})$$

where A and B are the two groups into which the data are partitioned, A is a smaller cluster than B, S_{xy} is the sum of the similarities between members of any two clusters x and y, n_A is the number of entities in A, and p is an arbitrary parameter to allow the user some control over the size of the overlap. Arbitrary points are chosen as centroids for initial number of clusters (Jones and Jackson, 1967). Then, many iterations are performed and individual data are assigned to the chosen initial clusters by dividing each available cluster into two clusters satisfying the condition of minimizing the cohesion function.

Some disadvantages are associated with this method, including that a diagnostic test be performed for the smaller group of each division. The overlap between the clusters is defined arbitrarily by the user which might lead to significant differences in the results and also the number of iterations are user-dependent. Due to the uncertainty produced during the different steps of clumping analysis, the method is not used for piezocone data in this study.

Density-seeking methods: Gengerelli (1963) sought methods to separate data with lower density than those with higher density (more repeatable data). In other words, the purpose of this method is to find the points that can represent a specific data set. The concept of this method is not applicable to the purpose of defining a soil stratification based on cone data, especially in the case of multilayered profiles. For example, suppose a soil log consists of a layer of sand and a layer of clay, it is meaningful to find a subset of data that can represent both soil-types.

By comparing the limitations of the various cluster methods, the hierarchical techniques are concluded to be the most feasible for clustering piezocone data into correlated groups. Therefore, the hierarchical types of clustering is applied through this study.

4.12. Simple Examples of Hierarchical Clusters

The cluster results using the agglomerative scheme can be visualized via a dendrogram (tree) diagram (Everitt, 1974) which is a summary of the consequence link between two individual points, one point and a group of data, and two groups of data. For example, Table 4.8 includes 10 readings of Q and B_q at chosen depths of PCPT1 sounding at Amherst Massachusetts. The data are standardized using zscore method and the results are listed in Table 4.8. A cosine measurement is used to indicate the similarity between the data and the statistical distances (cosine) between pairs of data at different depths are summarized in Table 4.9. A single-link (nearest neighbor) technique is used to

Table 4.8. Selected Points of Q and B_q of PCPT1 at Amherst, Massachusetts (Data from this Study).

Point No.	Depth (m)	Q	B _q	zscore (Q)	zscore (B _q)
1	1	147.0	-0.011	1.89	-1.40
2	2	103.6	-0.011	1.16	-1.40
3	3	62.4	0.006	0.47	-1.36
4	4	13.7	0.272	-0.34	-0.70
5	5	12.3	0.533	-0.37	-0.05
6	6	5.2	0.997	-0.49	1.10
7	7	2.6	1.584	-0.53	2.56
8	8	4.9	0.944	-0.49	0.97
9	9	4.8	0.928	-0.49	0.93
10	10	4.6	0.834	-0.50	0.88

Table 4.9. A summary of the Cosine Measurement (Similarity) Between Selected Points of Q and B_q of PCPT1 at Amherst, Massachusetts (Data from this Study).

Point number	1	2	3	4	5	6	7	8	9	10
1	1.00	0.97	0.79	-0.12	-0.69	-0.93	-0.74	-0.96	-0.96	-0.98
2	0.97	1.00	0.91	0.11	-0.51	-0.99	-0.88	-1.00	-1.00	-1.00
3	0.79	0.91	1.00	0.51	-0.12	-0.96	-1.00	-0.94	-0.93	-0.91
4	-0.12	0.11	0.51	1.00	0.80	-0.25	-0.58	-0.18	-0.15	-0.11
5	-0.69	-0.51	-0.12	0.80	1.00	0.38	0.04	0.46	0.48	0.52
6	-0.93	-0.99	-0.96	-0.25	0.38	1.00	0.94	1.00	0.99	0.99
7	-0.74	-0.88	-1.00	-0.58	0.04	0.94	1.00	0.91	0.89	0.87
8	-0.96	-1.00	-0.94	-0.18	0.46	1.00	0.91	1.00	1.00	1.00
9	-0.96	-1.00	-0.93	-0.15	0.48	0.99	0.89	1.00	1.00	1.00
10	-0.98	-1.00	-0.91	-0.11	0.52	0.99	0.87	1.00	1.00	1.00

group the data of two classes based on the closest statistical distance between their elements and the consequence of grouping the data is summarized in Table 4.10 and Fig.

4.9. At a step number zero, every point is in an individual cluster, and, obviously, the similarity level between a point and itself is 1 on a scale between zero and 1.

At a step number 1, the closest two points (8 and 9) are grouped in one cluster, and the total number of clusters is dropped by 1. At a step number 2, the next two most similar points (8 and 10) are grouped in a cluster at a similarity level of 0.996 and the points 8, 9 and 10 form one group. The procedure continues until all points are grouped in one cluster at a step number 9. The points numbers 6 to 10 are highly correlated and grouped together at a similarity level ≥ 0.937 which indicates that they have the same soil

Table 4.10. Summary of Clustering Consequence of Selected Piezocone Data at Amherst Using the Cosine Method for the Similarity and Single Linkage for Grouping the Data.

Step	Number of clusters	Similarity level	Joined points		Number of data points in new cluster
0	10	1.000	Each point is one cluster		
1	9	0.999	8	9	2
2	8	0.999	8	10	3
3	7	0.996	6	8	4
4	6	0.973	1	2	2
5	5	0.937	6	7	5
6	4	0.912	1	3	3
7	3	0.795	4	5	2
8	2	0.517	4	6	6
9	1	0.510	1	4	10

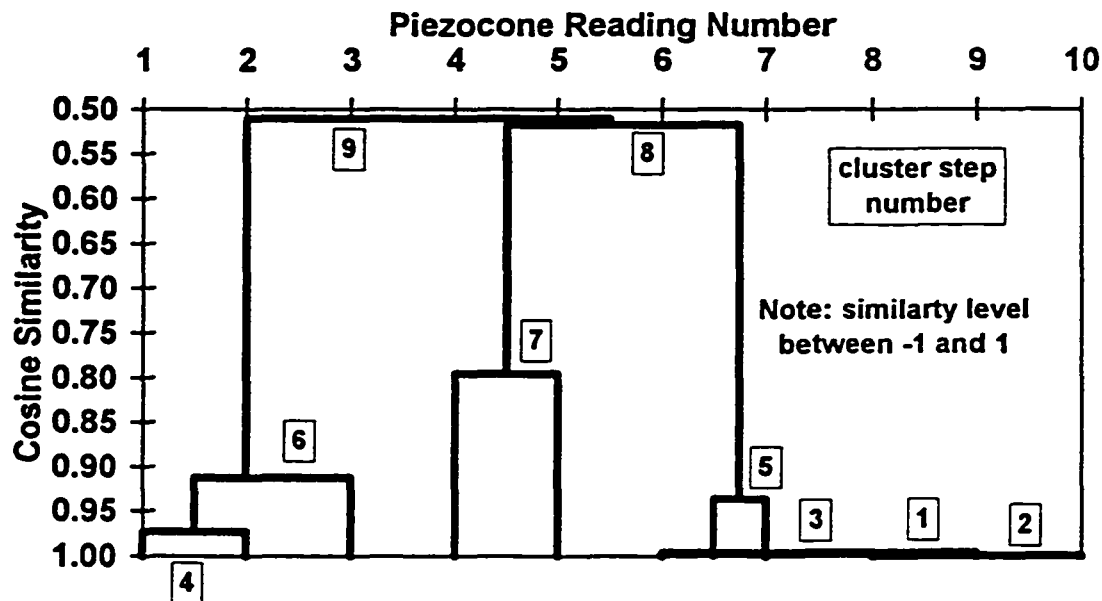


Figure 4.9. A Dendrogram (Tree) Diagram Summarizes the Consequence of Clustering of Piezocone data at Amherst, Massachusetts.

type and/or behavior. Looking at the dendrogram, three groups of data are identified as follows: (1) group 1 includes points from 6 to 10 and its data are highly correlated because their similarity levels ≥ 0.937 , (2) group 2 includes points from 1 to 3 and the data are also highly correlated because they have similarity levels ≥ 0.912 , and (3) group 3 includes points 4 and 5 which are relatively moderately correlated at a similarity level of 0.795. Then the correlation of the three groups is relatively weak at similarity levels equal to 0.510 and 0.517. This suggests that each group belongs to a separate soil stratum. Group 1 represents the soft clay layer, group 2 represents the upper fill layer, and group 3 represents the clay crust layer.

In some cases, there are subtle changes in piezocone data which can not be identified by a simple visual method. For example, piezocone data and the soil stratigraphy at Troll test site, North Sea are shown in Fig. 4.10. A detailed study of the site was performed by Amundsen et al. (1985) and the soil is classified as soft clay in the upper 17 m underlain by a stiff clay down to the termination of the sounding at 45 m. In the upper layer, the average water content and sensitivity are equal to 59 percent and 5.8, respectively, and their averages in the lower layer are equal to 24 percent and 2.2, respectively. Selected points of Q and B_q are listed in Table 4.11 to represent both layers. Cluster analysis is performed to divide the data into correlated groups. First, the data are standardized using a zscore method and the results are summarized in Table 4.11. Then the similarity between pairs of data at different depths is determined using a cosine measurement and the results are listed in Table 4.12.

Table 4.11. Selected Piezocone Data and Their Standardization Using Zscore Method at Troll, North Sea (Piezocone data from Amundsen et al., 1985).

Data No.	Depth (m)	Q	B _q	ZQ	ZB _q
1	3	2.35	1.09	-0.42	3.10
2	6	3.73	0.78	-0.24	0.91
3	9	3.92	0.72	-0.21	0.42
4	12	4.21	0.70	-0.17	0.28
5	15	4.22	0.70	-0.17	0.33
6	20	4.66	0.77	-0.11	0.82
7	25	5.16	0.61	-0.04	-0.33
8	30	5.20	0.58	-0.04	-0.55
9	35	5.04	0.60	-0.06	-0.39
10	40	4.98	0.64	-0.07	-0.15

Table 4.12. A summary of the Cosine Measurement (Similarity) Between Selected Points of Q and B_q of PCPT1 at Troll, North Sea (Data from Amundsen et al., 1985).

Data point	1	2	3	4	5	6	7	8	9	10
1	1.000	0.956	0.633	-0.088	0.091	0.040	-0.983	-0.998	-0.999	-0.987
2	0.956	1.000	0.833	0.208	0.380	-0.256	-0.994	-0.972	-0.970	-0.991
3	0.633	0.833	1.000	0.714	0.828	-0.748	-0.765	-0.680	-0.673	-0.751
4	-0.088	0.208	0.714	1.000	0.984	-0.999	-0.096	0.028	0.037	-0.075
5	0.091	0.380	0.828	0.984	1.000	-0.992	-0.272	-0.152	-0.142	-0.252
6	0.040	-0.256	-0.748	-0.999	-0.992	1.000	0.145	0.022	0.012	0.124
7	-0.983	-0.994	-0.765	-0.096	-0.272	0.145	1.000	0.992	0.991	1.000
8	-0.998	-0.972	-0.680	0.028	-0.152	0.022	0.992	1.000	1.000	0.995
9	-0.999	-0.970	-0.673	0.037	-0.142	0.012	0.991	1.000	1.000	0.994
10	-0.987	-0.991	-0.751	-0.075	-0.252	0.124	1.000	0.995	0.994	1.000

Finally, the data are clustered based on their similarity using a single link method. A summary of the consequence of clustering is given in Fig. 4.11 and Table 4.13. First, each point is in a separate cluster, then the most similar points are grouped together until all points forms one cluster. Points 8 and 9 are grouped first with a similarity level equal to 0.9999 (~1.000) then points 7 and 10 are grouped at a similarity level equal to 0.9997 (~1.000). Points 7 and 8 are the most similar points and a new cluster is formed including points 7, 8, 9 and 10 at a similarity level equal to 0.995. Then points 4 and 5 of the upper soft clay layer are grouped at a similarity level equal to 0.984 and clustering continues until all points forms one cluster. Looking at the dendrogram in Fig. 4.11, two primary groups can be defined as follow: (1) group 1 includes points from 7 to 10 which are highly correlated as validated by a similarity level approximately equal to 1, and (2) group 2 includes points from 1 to 5 which are relatively less correlated compared with group 1 and

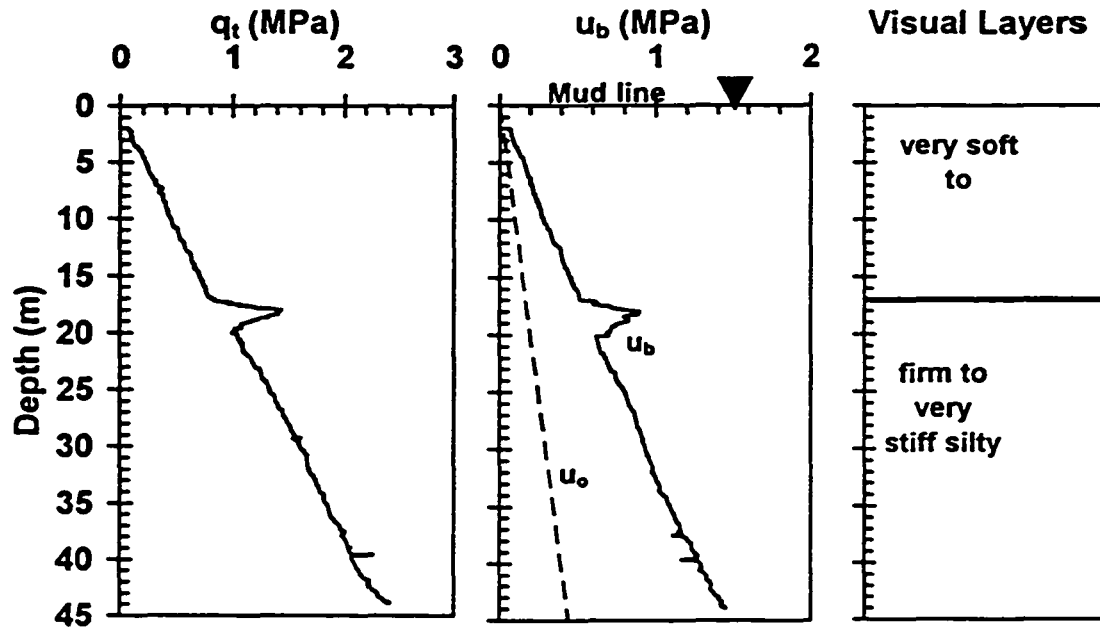


Figure 4.10. Piezocone Data and Soil Profile at Troll, North Sea
(Data from Amundsen et al., 1985).

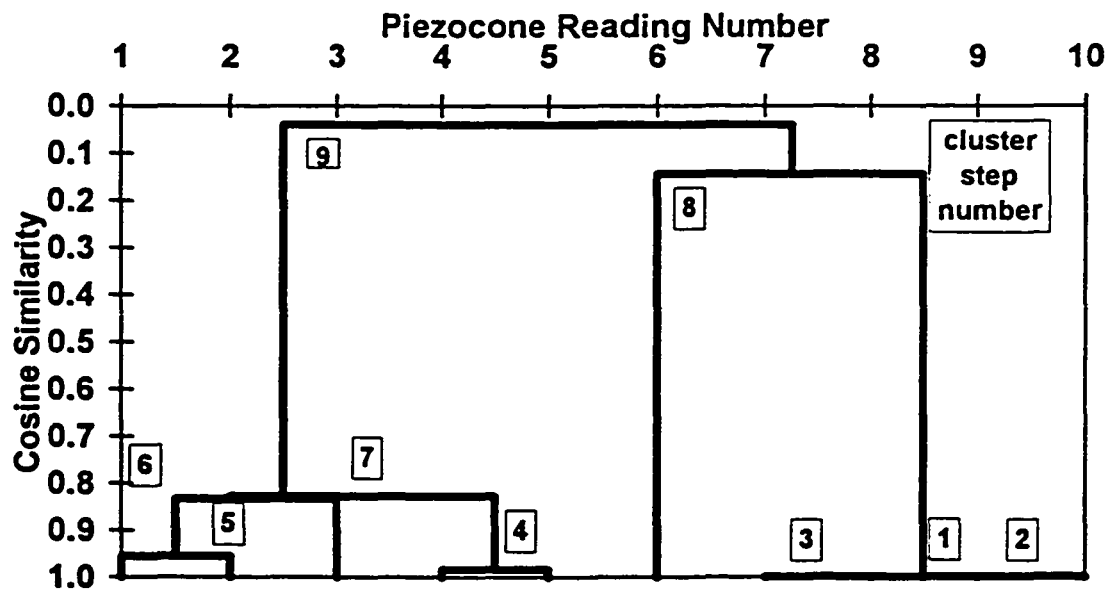


Figure 4.11. A Dendrogram (Tree) Diagram Summarizes the Consequence
of Clustering of Piezocone Data at Troll, North Sea
(Piezocone Data from Amundsen et al., 1985).

Table 4.13. Summary of Clustering Consequence of Selected Piezocone Data at Troll, North Sea, Using the Cosine Method for the Similarity and Single Link Method for Grouping the Data (Piezocone Data from Amundsen et al., 1985).

Step	Number of clusters	Similarity level	Joined points		Number of data points in new cluster
0	10	1.000	Each point is one cluster		
1	9	1.000	8	9	2
2	8	1.000	7	10	2
3	7	0.995	7	8	4
4	6	0.984	4	5	2
5	5	0.956	1	2	2
6	4	0.833	1	3	3
7	3	0.828	1	4	5
8	2	0.145	6	7	5
9	1	0.040	1	6	10

their similarity level is greater than 0.828. A point number 6 does not belong to either of the two groups which are uncorrelated as indicated by a similarity level equal to 0.04. The first group represents the lower stiff clay layer between the depths of 17 m and 45 m and the second group represents the upper soft clay layer between the mudline and a depth of 17 m. Therefore, cluster analysis is able to identify the invisible changes in piezocone data which indicates major changes in soil type and/or behavior.

4.13. Determining the Number of Clusters

Milligan and Cooper (1985) evaluated 30 independent methods to define the number of clusters. The methods were also independent of the cluster technique used. A total of 432 simulated artificial data sets produced from multivariate Gaussian distributions

with independent variables were analyzed in their study to determine predistinguished number of clusters in a range from 2 to 5 groups. The methods were ranked from 1 to 30 according to their performance to recover the predefined clusters. Cooper and Milligan (1988) tested the application of these methods for data sets including outliers and found that these methods were not able to indicate the right number of clusters that could represent the inherent groups within each data set. Milligan (1996) indicated that, yet, there has been not enough development and evaluation of the subject of choosing the method(s) and number of clusters. Also, the characteristics of the analyzed data should be considered in a criterion to define the number of groups representing these data.

If desired, a cluster grouping can provide as many clusters as there are data sets, therefore a criterion must be set to evaluate the fewest number of clusters for consideration. Extensive cluster analysis has been conducted at many sites for a guide and the evaluation ranged from a cluster number $N_c = 2$ to $N_c = 100$. However, the proper soil stratigraphies of the 25 studied cases are detected at $N_c < 15$. A simple criterion is developed to decide which cluster results should be examined. For this, the correlation coefficient (ρ_c) is calculated between the data ranks of each consecutive pair of clusters as follows (Neter et al., 1990):

$$\rho_c = \frac{\sum_{i=1}^n (x_{i(j)} - E(X)_{(j)})(x_{i(j+1)} - E(X)_{(j+1)})}{\sqrt{\sum_{i=1}^n (x_{i(j)} - E(X)_{(j)})^2} \sqrt{\sum_{i=1}^n (x_{i(j+1)} - E(X)_{(j+1)})^2}} \quad (\text{Equation 4.9})$$

where $-1 \leq \rho_c \leq 1$, x_i is a cluster rank of a data vector (i), and j is the cluster number. For instance, if ρ_c between the two clusters j and $j+1$ is equal to 1, that means the cluster $j+1$ does not add more divisions of the data set than cluster j and the analysis can be stopped at cluster j . However, if ρ_c is equal to 0.5, that might indicate significant difference between the assigned cluster ranks for the data. This will suggest that the cluster analysis needs to be performed again until ρ_c approaches a value of 1. Although ρ_c becomes close to 1, it will never equal 1 because there will at least be a difference of one point rank between two successive clusters. Cluster results will be examined at the peaks of ρ_c , or in other words, after a significant change of data groups might occur.

Based on this discussion and because the cone data usually contain outliers either due to soil lenses or data errors, these methods are not used in this study. A simple criterion is developed in this study for choosing a cluster number corresponding to the correct stratigraphy and based on the characteristics of the cone data as functions of the surrounding soils as discussed in the following section.

4.14. Interpretation of the Cluster Results

The cluster analysis provides the following results within a vertical profile: related and unrelated soil layers, seams, lenses, anomalies, and transition zones between different layers. The anomalies can include natural soil inclusions, such as cemented layers, stones,

or voids; or systematic errors related to measurement difficulties, including electrical noise, rod changes, and random events.

Definitions are needed for a minimum layer thickness, a transition zone, and outliers associated with soil lenses or anomalies. First, for the minimum layer thickness, Treadwell (1976) and Schmertmann (1978) found that the cone penetration resistance, q_c , is influenced both by the interfaces ahead and behind the tip. They reported that the full response of q_c in a chamber test moving from softer to stiffer soil can be obtained after interface zone equal to 10 to 20 times the cone diameter. However, in the case of a moving cone from stiffer to softer soils, the full response of q_c can be obtained in a distance smaller than 10 times the cone diameter. Vivitrat (1978) suggested that a soil layer should contain at least 20 points (almost equivalent to a layer thickness between 0.5 m and 1 m for data frequency equal to 2.5 cm and 5 cm, respectively) of the piezocone data to be statistically significant.

Campanella and Robertson (1989) noted the scale effect on the transition zone using 10-cm² and 15-cm² cones. The peaks of q_c were not completely retrieved, but the troughs were reproduced. If the cone has a base area equal to 10 cm² or 15 cm², the interface zone thickness might be equal to 0.35 to 0.70 m or 0.44 to 0.88 m, respectively. The penetration pore pressure might be influenced by a soil zone equal to 0.1 times a cone diameter (DeRuiter, 1982). The soil volume affecting q_c is prevailing and considered in the development of the interpretation criterion in this study. Wickremesinghe (1989) proposed a statistical method to define the soil boundaries based on piezocone data using the intraclass correlation coefficient (ICC) and the generalized distance (D^2). He

suggested a minimum layer thickness equal to 0.5 m. Zhang (1996) further used the ICC and D^2 statistical methods to analyze piezocone data to get the soil boundaries and suggested a minimum layer thickness equal to 0.75 m.

Based on the above discussion, a minimum layer thickness (t) is chosen, $t = 0.5$ m. Two layer definitions are given as follows: a primary layer (designated “A”) has $t \geq 1$ m, while a secondary layer (designated “a”) has $0.5 \text{ m} \leq t < 1 \text{ m}$. Soil mixtures and transitions are denoted by a^* and A^* , which indicates that there is no continuous group of data with $t \geq 0.5$ m or $t \geq 1$ within layers a and A , respectively. Note that a^* and A^* have $0.5 \text{ m} \leq t < 1$ and $t \geq 1$ m, respectively. A summary of the steps to the cluster analysis interpretation for soil layers is shown on Fig. 4.12.

The cluster analysis achieves the ultimate use of the piezocone data (collected, for instance, every 2 cm) by detecting a single point or number of points ($t < 0.5$ m) that are not associated with a soil layer. Those points could represent a soil lense or seam within a layer, or could be an outlier due to an electrical noise, operational errors or natural anomalies in the soil formation, such as gravels, boulders or voids. A transition zone is defined between soil layers or where at least three consecutive cone measurements have ascending or descending order of cluster numbers within a soil mixture. Soil lenses, seams, outliers, and data errors are recognized within a primary stratum by having different cluster number(s) than that of a soil-type surrounding them. A prior knowledge

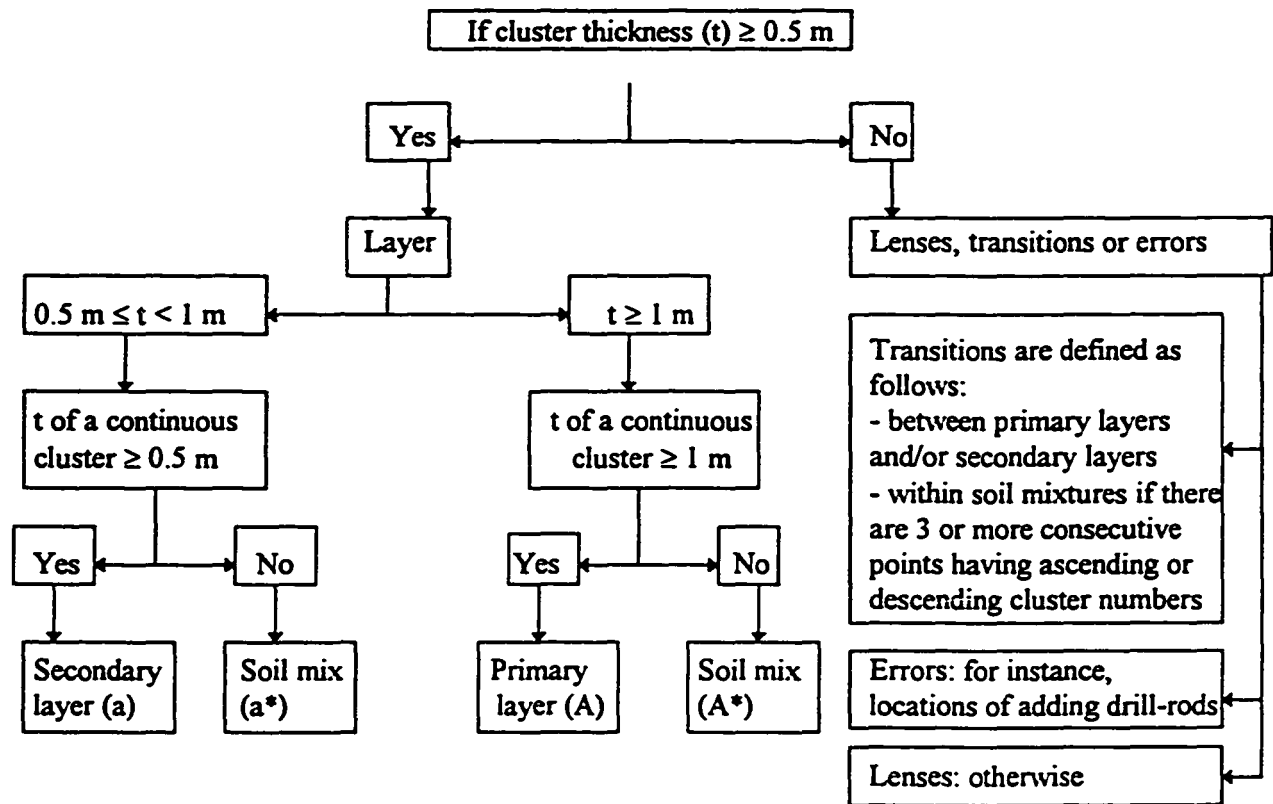


Fig. 4.12. Proposed Criteria for Evaluating Soil Stratigraphy from Cluster Analysis of Piezocone Data.

of the effect of data errors on cone measurements is needed in order to differentiate them from minor geological evidences. For instance, the depths at which cone-rods are added during a cone penetration test should be recorded to be able to identify anomalies detected by clustering results as procedural errors. In the case of soil lenses or seams, clustering analysis acts as a warning signal to point out inherent local variability within a homogeneous soil stratum. The existence of a soil lense might also be confirmed by

comparing cluster results with available index or mechanical soil properties at the same depth.

Electrical noise affects the piezocone readings within a sounding. It is found that this error-type has no significance on the stratigraphy obtained by cluster results because the data representing this error are defined in separate clusters than those indicating secondary and primary layers in a soil profile as discussed later in this chapter.

Operational errors such as those caused by the addition of rods every 1-meter are usually known. Clustering piezocone data is unaffected by this error-type. These prior known errors should be filtered from the original data, if possible, but clustering does not require this for purposes of stratification.

The discussed elements of the proposed statistical clustering analysis of piezocone data for the purpose of geostratigraphies delineation are integrated in Fig. 4.13 and summarized herein. First, the normalized piezocone parameters (Q and B_q) are derived from the measured raw data (q_t and u_b), then standardized using the zscore method. The similarity between different vectors of Q and B_q at different depths are determined using the cosine measurement. Standardized data are grouped into clusters using the single link (nearest neighbor) technique, then the developed interpretation method is applied to demarcate different soil layers and their association. Finally, the interpreted soil stratigraphy is verified by independent information from adjacent boreholes and available back up laboratory and in-situ data.

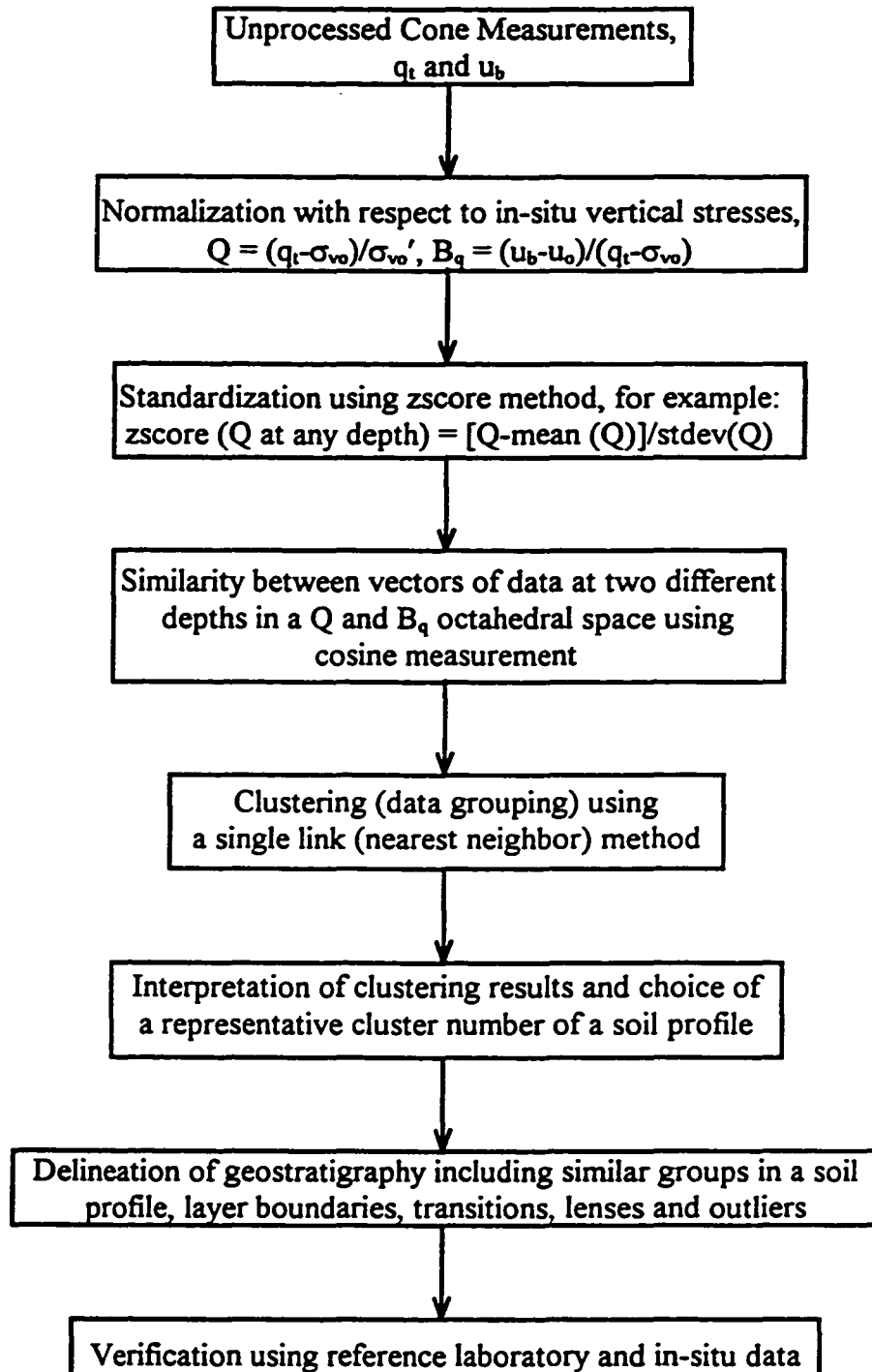


Figure 4.13. Proposed Clustering Analysis of Cone Data.

4.15. Validity of the Cluster Analysis

In this study, clustering analysis of piezocone data is verified by comparing the obtained statistical stratigraphy with independent information obtained from adjacent boreholes, as well as reference laboratory, and in-situ testing data. Statistical validity methods including visual interpretation, hypothesis testing, and replications are available, however, these techniques are not applied in this study because of their limitations discussed in the following sections.

4.15.1. Visual Interpretation

The validity of data classification by cluster analysis has been studied by Anderberg (1973), and Duffy and Quiroz (1991) who suggested the similarity (distance) matrix be arranged into blocks which represent the obtained clusters. The distance measurements within the same block are rearranged into an ascending order. The data are presented graphically using one of some available methods such as tree-diagrams (Kleiner and Hartigan, 1981). For the purpose of defining a subsurface profile using piezocone data, the dendrogram is a confusing interpretation method because of the large number of data involved in the analysis (e.g. 500 points in the case of a 10-m sounding). It is proposed that cluster results be summarized as a vertical profile with depth while the abscissa is assigned the cluster number for each datum (examples shown in Chapters 5 and 6). The internal correlation between the data of the same layer can always be obtained by referring to the dendrogram. A comprehensive discussion on the use of graphical techniques can be found in Jain and Dubes (1988). Then, the researcher can visually judge the validity of the

results by looking, for instance, at different colors or shadows represent different ranges of similarity. If the similarity within the same group is less than the similarity between the other groups, then a valid cluster analysis is applied. This criterion is a user-dependent and brings subjectivity into the final decision. Also, it depends on the same data for both the analysis and the validation, therefore, it is not used to validate clustering results in this research.

4.15.2. Hypothesis Tests

Hypothesis tests can be divided into two main categories including: (1) external validation using independent data set, (Hubert and Baker, 1977), or using variables not included in the analysis to check the given clusters (Milligan, 1996), and (2) internal validation tests based on the goodness of fit between the input data and the resulting classifications (Baker and Hubert, 1975; Milligan, 1981). However, the validity of hypothesis testing is questionable (Milligan and Mahajan, 1980; Jain and Dubes, 1988; Milligan, 1996) because they always give significant results due to the underlying assumption that clusters exist in the data. Therefore, the hypothesis testing methods have not been used in this study.

4.15.3. Validation Using Replications

Morey, Blashifield, and Skinner (1983) developed a replication analysis to check the validity of cluster analysis results. Two replicated data sets for the same variables should exist to perform the analysis. First, for each data set, the same cluster analysis is performed and then the agreement of their cluster results is compared (Hubert and Arabie, 1985). This method might be useful in the case of a simulation study where data of two artificial samples can be produced with predefined cluster structures. However, in the case of studying soil stratigraphy using piezocone data, it is not feasible to repeat the data set because the cone penetration test itself is destructive to the soil mass. The recorded piezocone data are functions of the soil micro- and macro- structures which are functions of the spatial location. Therefore, this method is not suitable for the purpose of this study.

4.16. Summary and Conclusions

In this chapter, review of standardization schemes, similarity measurements, and hierarchical techniques is given. Piezocone data should be standardized because the mean and variance of the direct stress measurements q_t , and u_b vary significantly. Also, the mean and variance of normalized parameters $Q = (q_t - \sigma_{vo})/\sigma_{vo}'$ and $B_q = (u_b - u_o)/(q_t - \sigma_{vo})$ vary greatly. Standardization is recommended to reduce the effect of the larger cone parameter (q_t in the case of using raw measurements or Q in the case of using normalized data) on clustering results. There are 7 different standardization methods and 8 different similarity approaches for grouping the data, thus gives possible 56 combinations. There is no single method that can be applied for all types of data. A parametric study is needed in

order to compare the clustering of piezocone data using the 56 methods as discussed later. The optimal cluster method is selected where the actual soil stratigraphy is minimum cluster number.

The single link (nearest neighbor) clustering method has been found superior over other techniques to indicate a predetermined inherent structure of artificial data sets (Milligan and Cooper, 1985; 1988). Also, it satisfies mathematical conditions such as minimum distortion (Sibson, 1972). Therefore, this method has been adopted for the analysis of piezocone data in detecting a subsurface stratification. The cluster number has been chosen according to a developed criterion based on piezocone data performed in chamber tests and previous statistical studies. To verify the cluster results, the given statistical stratification is compared with independent information of boreholes boundaries, physical, and mechanical properties of the soil at a specific site.

CHAPTER 5

CLUSTER APPLICATIONS OF PIEZOCONC DATA

5.1. Synopsis

In this chapter, a proposed clustering technique for analyzing piezocone data to demarcate a soil profile is evaluated. The single-link (nearest neighbor) method is proposed for dividing the piezocone data into correlated groups. A “zscore” standardization procedure and a cosine similarity measurement were found to be superior in indicating the soil stratigraphy at the three sites. Therefore, a single-cosine-zscore cluster technique is proposed for analyzing piezocone data to delineate soil boundaries and stratigraphies.

According to the interpretation criterion of clustering piezocone data, each soil stratigraphy is separated into three categories according to the thickness (t) of each class. These are defined as follows: (1) a primary stratum has $t \geq 1$ meter, (2) a secondary layer has $1 \text{ meter} > t \geq 0.5 \text{ meter}$, and (3) grouping of soil lenses or transitions if $t < 0.5 \text{ meter}$. A certain number of clusters (N_c) is chosen to represent a subsurface stratification if no new primary layers are detected at higher cluster numbers.

The validity of the single-cosine-zscore method using the normalized parameters Q and B_q is discussed, for example, for varved clay strata at Amherst, Massachusetts (piezocone data collected by the author). Cluster results are compared using selections of

paired and grouped parameters in the analysis, including: (1) unprocessed data [q_t and u_b], (2) partly processed data [q_t and the ratio u_b/q_t], and (3) normalized parameters [Q , F , and B_q]. Clustering of the two normalized parameters Q and B_q , properly indicated the subsurface stratification at Amherst using the lowest cluster number, as compared with the other three data combinations.

The effects of test procedure, data collection, and processing of the data are studied in their representative influence on the clustering. These include:

- measurement errors due to test procedure and electrical noise,
- data frequency,
- porous filter location,
- scale effect (cone size), and
- actual versus assumed unit weight.

Measurement errors due to procedural operations of field piezocone tests at the Amherst site are shown to be insignificant. Electrical noise influence of miniature cone data in kaolinite clay in chamber tests (Mayne et al., 1992) are also shown to be insignificant on the clustering results. A comparison is performed of clustering at Amherst using two sets of Q and B_q which are derived as functions of measured unit weights, and an assumed unit weight. Using the actual unit weights helps to identify the soil profile at a lower cluster number. From the parametric effect, it is found that, data frequency of 5 cm up to 50 cm has no effect on the detected primary layers at Amherst. This suggests the possibility of

using clustering for other in-situ methods such as the flat dilatometer and vane shear test in which the readings are taken at relatively larger spacing than in the case of a cone.

A comparison is performed to investigate the effect of porous filter location for penetration porewater pressures ($u_1 = u_t$) versus ($u_2 = u_b$) at two clay sites: (a) soft clay at the Bothkennar test site, Scotland (Nash et al., 1992) and (b) stiff fissured London clay at the Brent Cross test site, UK (Powell et al., 1988). At the Bothkennar test site, both u_t and u_b are positive, and increase with depth. However, at Brent Cross test site, the u_t data are positive whilst the u_b readings are negative due to dilatancy and fissuring of the overconsolidated clays. Using u_t measurements indicated a soil profile which is better validated by the back-up data than the stratigraphy obtained using u_b readings. The scale effect using 10-cm² and 15-cm² cones in layered clay-sand-clay profiles at Surry, Virginia (Gordon and Mayne, 1987) is shown to have negligible effect on the cluster results.

A spatial cluster analysis was performed on piezocone data from Amherst test site using three soundings oriented in a triangle in a plan view. The derived normalized parameters Q and B_q are compared at the same depths to evaluate the association of the detected statistical layers. Also, the possibility of interpolating the soil boundaries between the locations of the soundings in the space is checked.

5.2. Database

Piezocone data are collected from 24 sites worldwide to represent a variety of different soil types including clays, silts, sands and soil mixtures. Piezocone data collected in a chamber test are also analyzed. Some of the piezocone data are available in their

original digital format, however, the rest of the data were redigitized using AutoCad software (Bertoline, 1994). Table 5.1 includes a summary of the studied sites which are divided into two groups. Group I includes sites where a simple visual examination of the piezocone data can be used to define the soil stratigraphy including layers, lenses, and probable transitions. Group II sites have significant changes in the soil stratigraphy that are not readily apparent from the piezocone data by looking at the variation of different trends in vertical profiles of a cone or by using available interpretation techniques of cone data. The two normalized parameters Q and B_q are used as input for cluster analysis. Verification is provided by backup laboratory index properties such as liquidity index, water content, sensitivity, and/or other measurements on soil samples taken from borings at the site or nearby field tests, such as the vane shear test.

5.3. Applications of Cluster Analysis

In the following sections, cluster analyses of piezocone data are discussed at four sites to illustrate the influence of several factors on the methodology. First, the effect of different data combinations, errors, and frequencies are studied using piezocone data collected by the author at the Amherst test site, Massachusetts. At the same site, a comparison is performed between cluster results based on normalized cone parameters derived using actual and assumed unit weights. Moreover, the validity of spatial cluster analysis is evaluated at Amherst. Electrical noise effects are evaluated on clustering of q_c readings collected in a calibration chamber test of prestressed kaolinite and fine silica. The influence of the filter position at the cone face (u_f) and behind the tip (u_b) on clustering is

Table 5.1. Summary of 24 Sites Analyzed Using Cluster Methods.

Site Name, and Group Number	Location	Soil Type	Reference
(I) Aiken*	South Carolina	Alternative clay, silt and sand	Bratton et al. (1993)
(I) Amherst*	Massachusetts	Fill over clay crust underlain by soft clay	(This study)**
(I) Bothkennar	Scotland	Clay crust over soft silty clay, clayey silt	Nash et al. (1992)
(I) Brent Cross	United Kingdom	Weathered finely fissured clay over unweathered highly fissured clay	Powell et al. (1988)
(II) Drammen	Norway	Plastic clay over lean clay	Masood et al. (1990)
(I) Fort Road	Singapore	Soft clay and intermediate silty clay	Chang (1991)
(II) Gloucester	Ontario	Soft silty clay over clay to silty clay	Konrad and Law (1987)
(II) Hachirogata	Japan	Soft marine clay	Tanaka et al. (1992)
(I) Kagoshima	Japan	Sand to silty sand over silt underlain by sand	Takesue et al. (1995)
(II) Lilla Mellosa	Sweden	Organic clay over clay underlain by a varved clay	Larsson and Mulabdic (1991)
(I) Maskinonge	Quebec	Firm to soft silty clay	Demers et al. (1993)
(I) McDonald's Farm	British Columbia	Sand and silty sand over soft clayey silt	Robertson (1982)
(I) Newport News	Virginia	Soil mixture of sand, silt and clay underlain by sandy clay to clayey sand	Mayne (1989)
(I) Onsoy	Norway	Soft clay	Gillespie et al. (1985)
(I) Opelika*	Alabama	Silty sand to sandy silt	(This study)***
(I) Penuelas*	Puerto Rico	Clayey silt and some sand	Hegazy and Mayne (1996)
(I) Po River	Italy	Medium to coarse sand	Bruzzi et al. (1986)
(II) Recife	Brazil	Organic soft clay (1) and (2)	Coutinho and Oliveira (1997)
(I) South Boston	Massachusetts	Silty clay differs from stiff to soft with depth	Sweeney and Kraemer (1993)
(II) St. Alban	Quebec	Soft very sensitive silty clay changes to clayey silt	Roy et al. (1982)
(I) Surry*	Virginia	Clay over sand underlain by clay	Gordon and Mayne (1987)
(I) Taranto	Italy	Stiff to hard silty clay	Battaglio et al. (1986)
(II) Tiller	Norway	Silty clay over quick clay	Sandven (1990)
(II) Troll	North Sea	Very soft clay over very stiff silty clay	Amundsen et al. (1985)

Notes: (*) means that piezocone data are available in digital format. Otherwise, piezocone data are redigitized using AutoCad software.
 (**) Additional supplementing data obtained from Lally (1993).
 (***) Additional supplementing data obtained from Vinson and Brown (1997).

studied at Bothkennar, Scotland, and Brent cross, England, test sites. Finally, the scale effect on clustering using 10-cm² and 15-cm² cones is evaluated at Surry test site.

5.3.1. Cluster Analysis at Amherst , Massachusetts

One of 15 representative piezocone soundings (PCPT1) conducted by the author in June, 1996 at the National Geotechnical Experimental Site at Amherst, Massachusetts is used for this clustering example. The sounding was presented earlier as Fig. 4.3. Detailed laboratory and in-situ data at the site are given elsewhere, (Lally, 1993, Lutenegeger, 1995). The soil stratigraphy consists of a 2-m thick silty clay fill over a 2-m thick desiccated sandy silt crust underlain by a soft brown-gray varved clay down to 5 m and a deep deposit of normally-consolidated gray varved clay to the termination depth of the sounding at a depth of 14.5 meter.

A visual examination of the q_t measurements indicates four separate layers with interface boundaries at approximately 1.8 m, 4.0, m and 5.2 m. The visual boundaries of the q_t measurements and adjacent borehole soundings are in agreement. For the penetration porewater readings, there is only one apparent boundary visually defined at a 4-m depth using the u_b measurements. The lack of sensitivity for the porewater channel reading is attributed to the penetration in the partially-saturated zone above the groundwater table. Likely, a partial desaturation of the porous element occurred due to suction effects which resulted in a delay of the full response of the penetration pore pressure at later depths.

The piezocone data are first evaluated by an empirical classification scheme using the modified Robertson chart (1991) to define the soil types at the site, as shown on Fig. 5.1, resulting in a sand to sand mixture between the ground surface and 4-m depth. The soil between the depths of 4.0 m and 14.5 m is defined as clay to silty clay with indication that the overconsolidation ratio (OCR) decreases and sensitivity (S_v) increases with depth. A primary soil boundary is identified at approximately a depth of 4 meters between the upper fill and sandy silt crust layers and the lower varved clay layer.

A single-cosine-zscore method is applied to the normalized parameters Q and B_q which are derived from the Amherst data and shown on Fig. 5.2. The standardization of Q and B_q using zscore method is shown in Fig. 5.3. The zscore profile of Q can be divided visually into 6 groups with boundaries at approximately 0.5 m, 2 m, 3 m, 4 m, and 5.5 m. The zscore profile of B_q can be divided by eye into three groups with boundaries at approximately 4 m and 5.5 m. The autocorrelation between consecutive clusters up to $N_c = 100$ is shown on Fig. 5.4. The scale of ρ_c is expanded in the vertical direction for cluster numbers between $N_c = 2$ and $N_c = 50$ showing 17 peaks of ρ_c defined between them. Figure 5.5 shows the cluster results at these peaks up to $N_c = 29$. At $N_c = 2$, two major clusters appear with a boundary at 5.8 m. A third primary cluster (A) with $t \geq 1$ m splits from the upper cluster at $N_c = 3$. Then, up to $N_c = 6$ some points separate indicating non-homogeneity within the three primary clusters.

At $N_c = 8$, another primary cluster is detached between 2.2 m and 3.95 m. For $N_c > 8$, no more primary layers of type A appear, however, more points indicating transitions

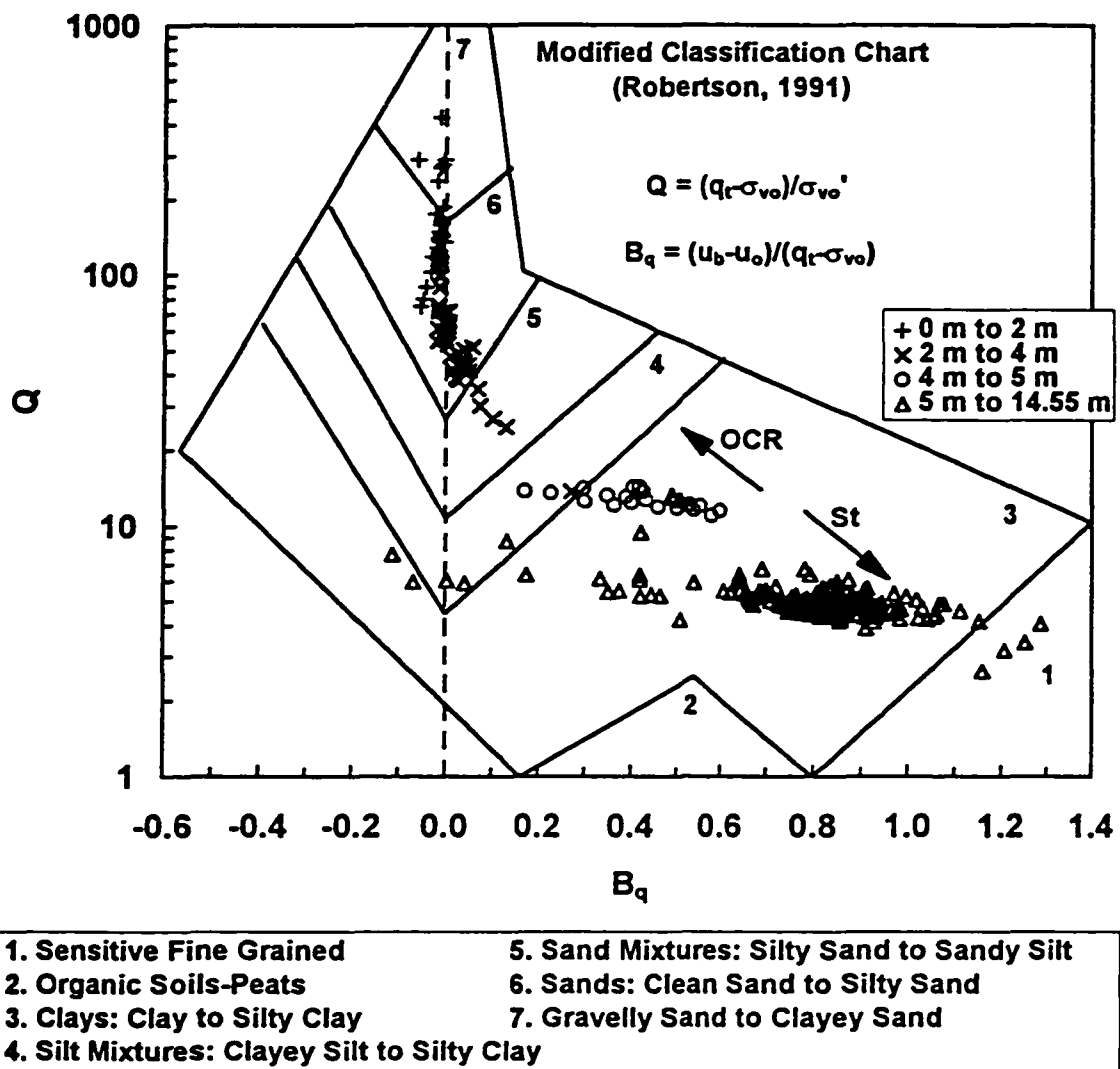


Figure 5.1. Soil Classification Using Piezocone Data (PCPT1) at Amherst Massachusetts (This Study).

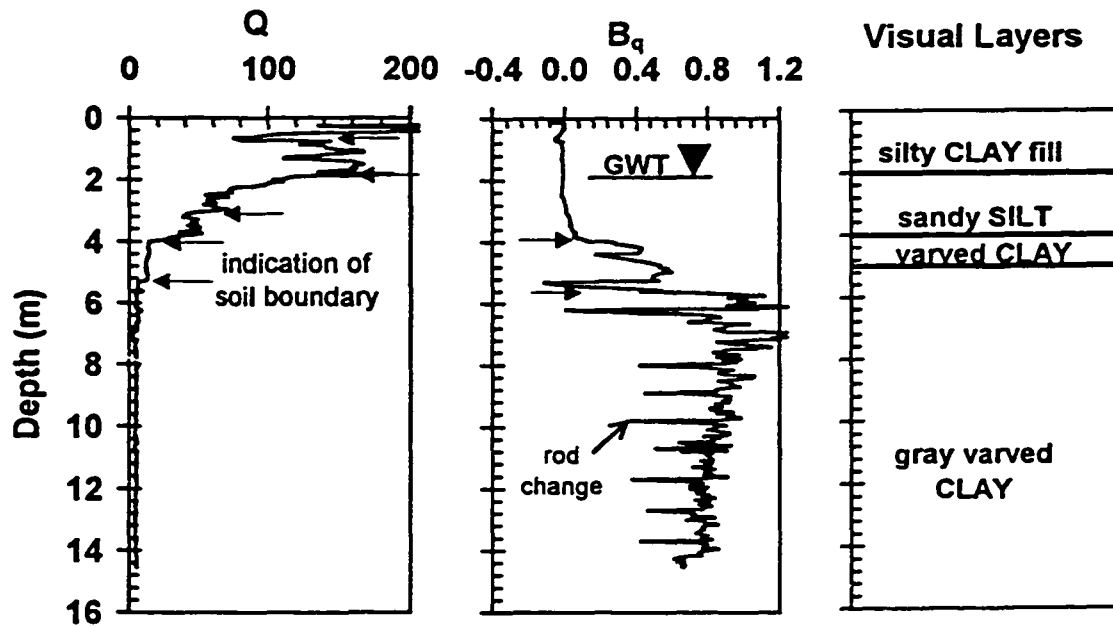


Figure 5.2. Normalized Piezocone Parameters Q and B_q and Soil Profile at Amherst, Massachusetts (Data from This Study).

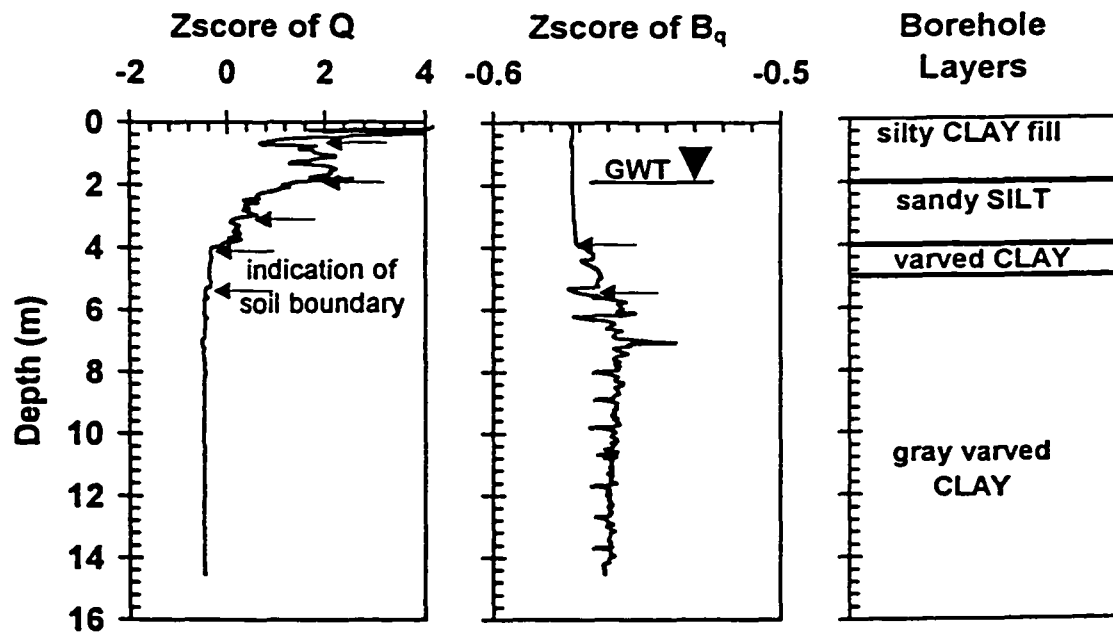
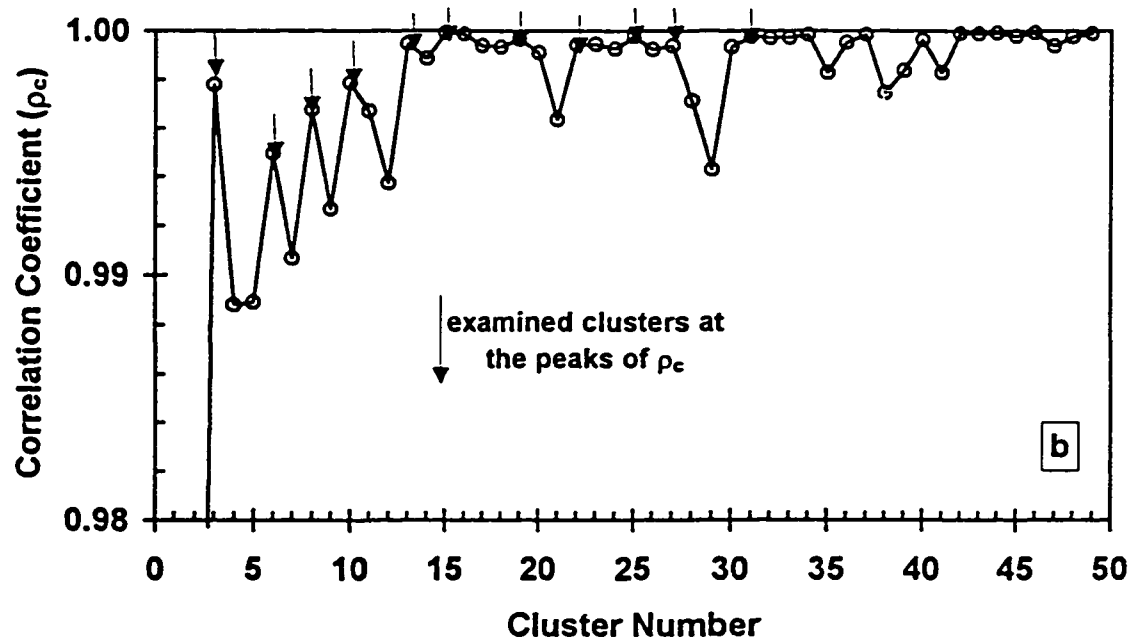
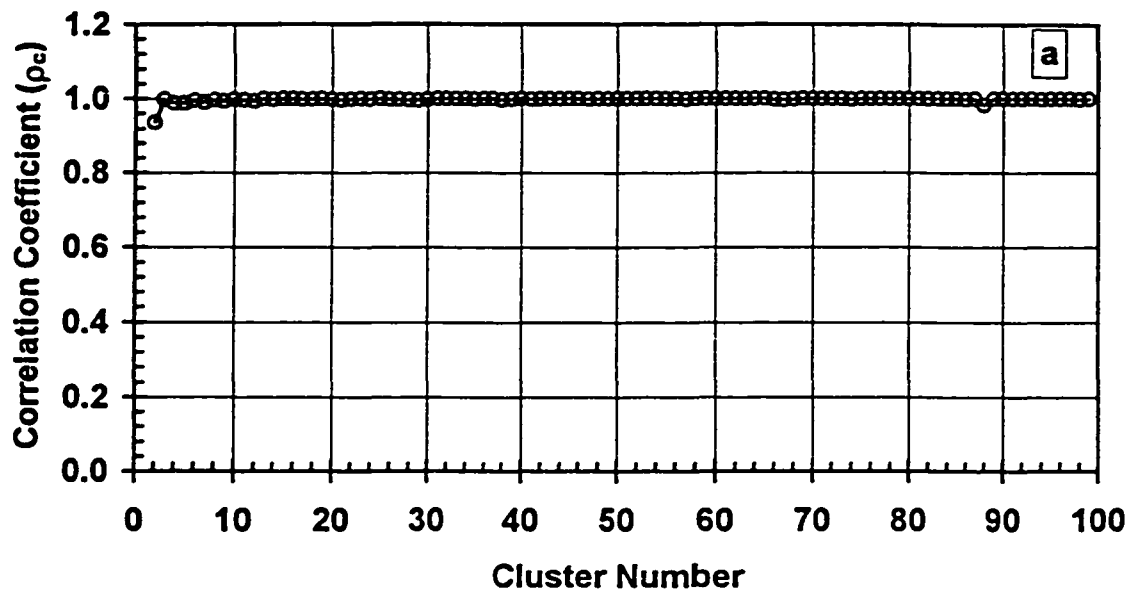
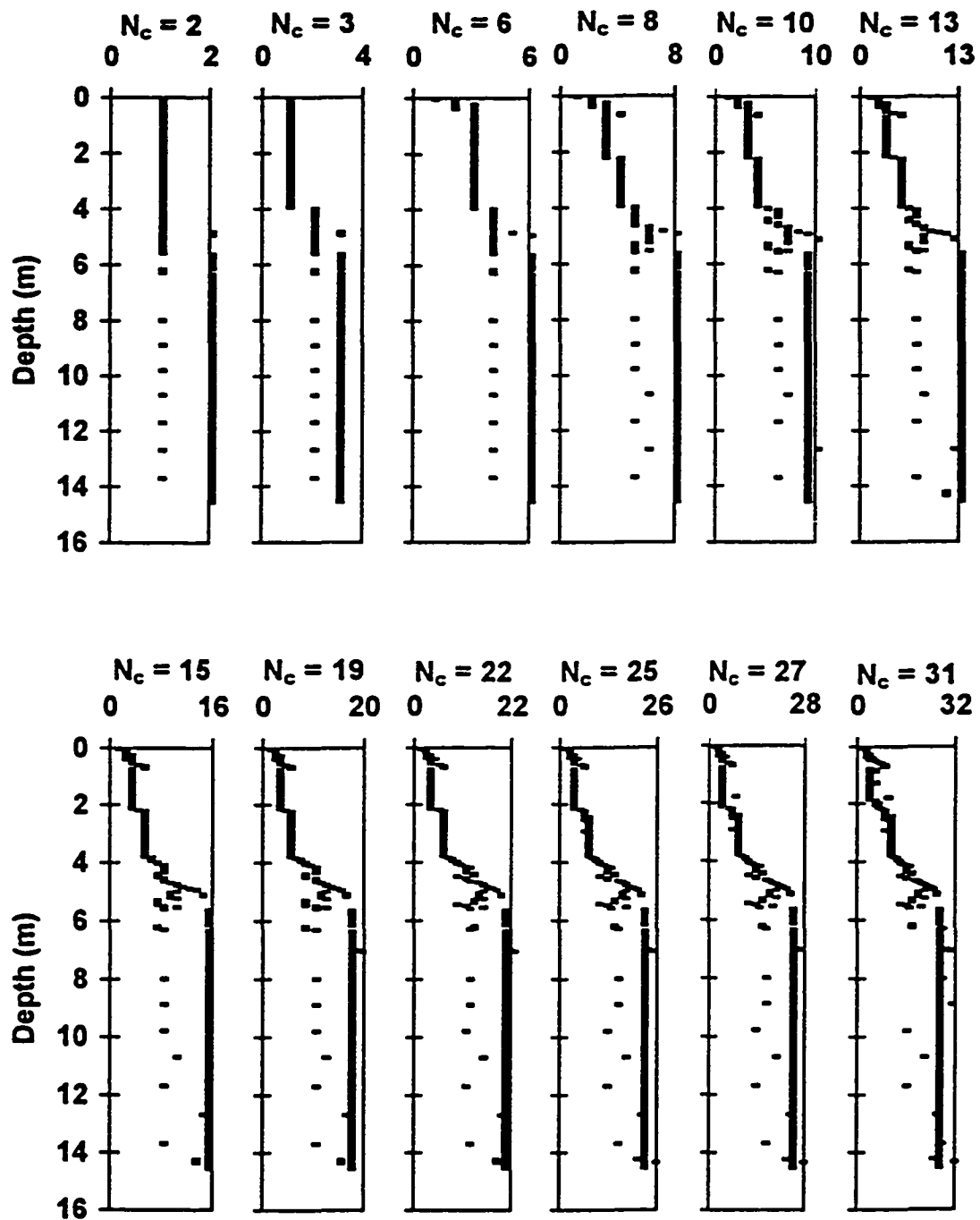


Figure 5.3. Zscore of Normalized Piezocone Parameters Q and B_q and Soil Profile at Amherst, Massachusetts (Data from This Study).



Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).

Figure 5.4. Correlation Coefficient Between Consecutive Cluster Results at Amherst, (a) Normal, and (b) Expanded Scales.

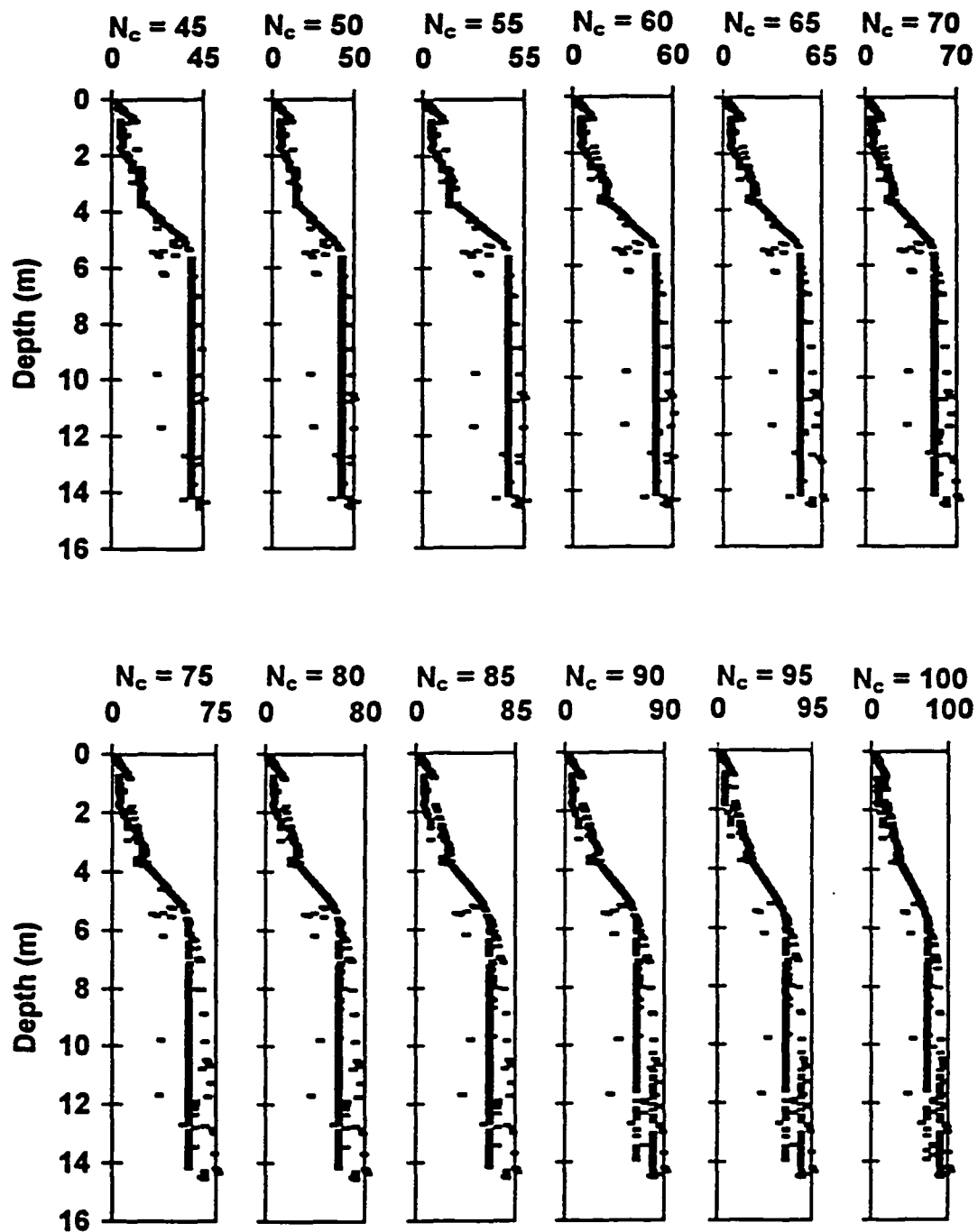


Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).

Figure 5.5. Cluster Analysis of Piezocone Data at Amherst, Massachusetts ($N_c = 2$ to 31).

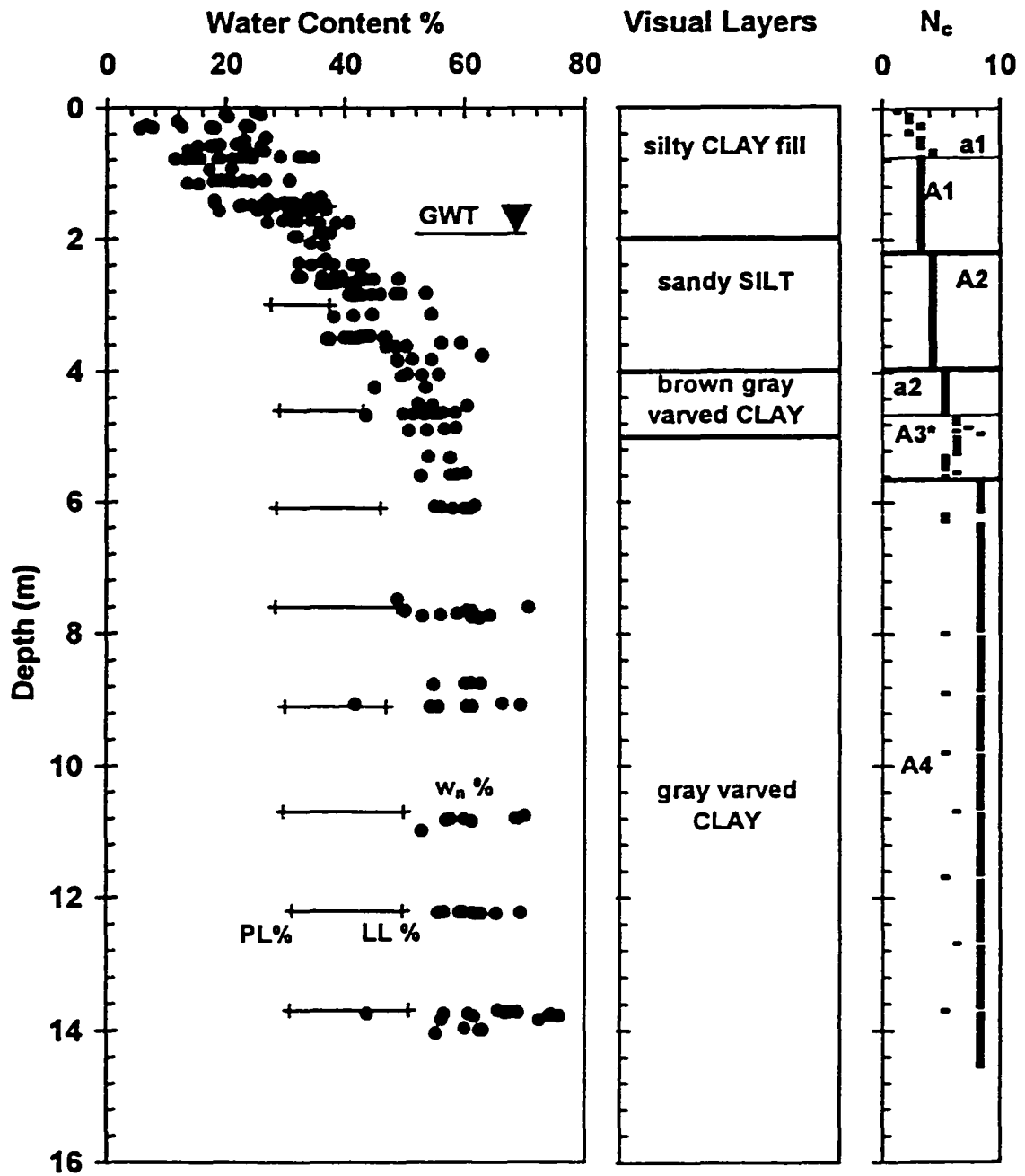
or lenses continue to separate the primary 3 layers. This statement is also valid to the cluster results between $N_c = 45$ and up to $N_c = 100$ which are shown on Fig. 5.6 with increment equal to 5 clusters. Note that at cluster number 88, there is a decrease in ρ_c (see Fig. 5.4) which is indicated by separation of more points from the lower primary layer, especially between the depths of 11 m and 14 m. However as shown in Fig. 5.6. (e.g., $N_c = 90$), the thickness of any consecutive group of points within these depths is less than 0.5 m. Note also that the number of clusters can grow up until each point in the soil is assigned a different cluster number. Therefore, a cluster analysis number 8 is chosen to represent the soil stratigraphy as shown in Fig. 5.7. Furthermore, note that the rod breaks (procedural error caused by creep during rod additions) show up as their own cluster.

The statistical soil stratigraphy at Amherst includes three primary layers (A1, A2 and A4), two secondary layers (a1 and a2) and a mixture soil layer (a3*). Soil boundaries defined by the cluster analysis are in agreement with the visual boundaries defined using borehole soundings as shown on Fig. 5.7. The water content data have a positive trend increasing with depth down to almost 6 m which supports the gradual change of the cluster groups in the same zone. The water content data have a mean value equal to 60 % below 6 m and down to almost 14 m and validates the stability of the continuity of cluster number 8 at the same zone. The undrained shear strengths measured using a field vane shear test have a mean value of $s_w = 30$ kPa below 6 m also in validation of cluster A4.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).

Figure 5.6. Cluster Analysis of Piezocone Data at Amherst, Massachusetts ($N_c = 45$ to 100).



Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).
 SCZ method using Q and Bq, cluster no. 8

— Primary boundary
 — Secondary boundary

Figure 5.7. Comparison Between Cluster Analysis, Visual Classification and Water Content % at Amherst, Massachusetts.

5.3.2. Unprocessed and Processed Data

The cluster results of single-cosine-zscore-type is compared at Amherst using three forms of piezocone data including the following:

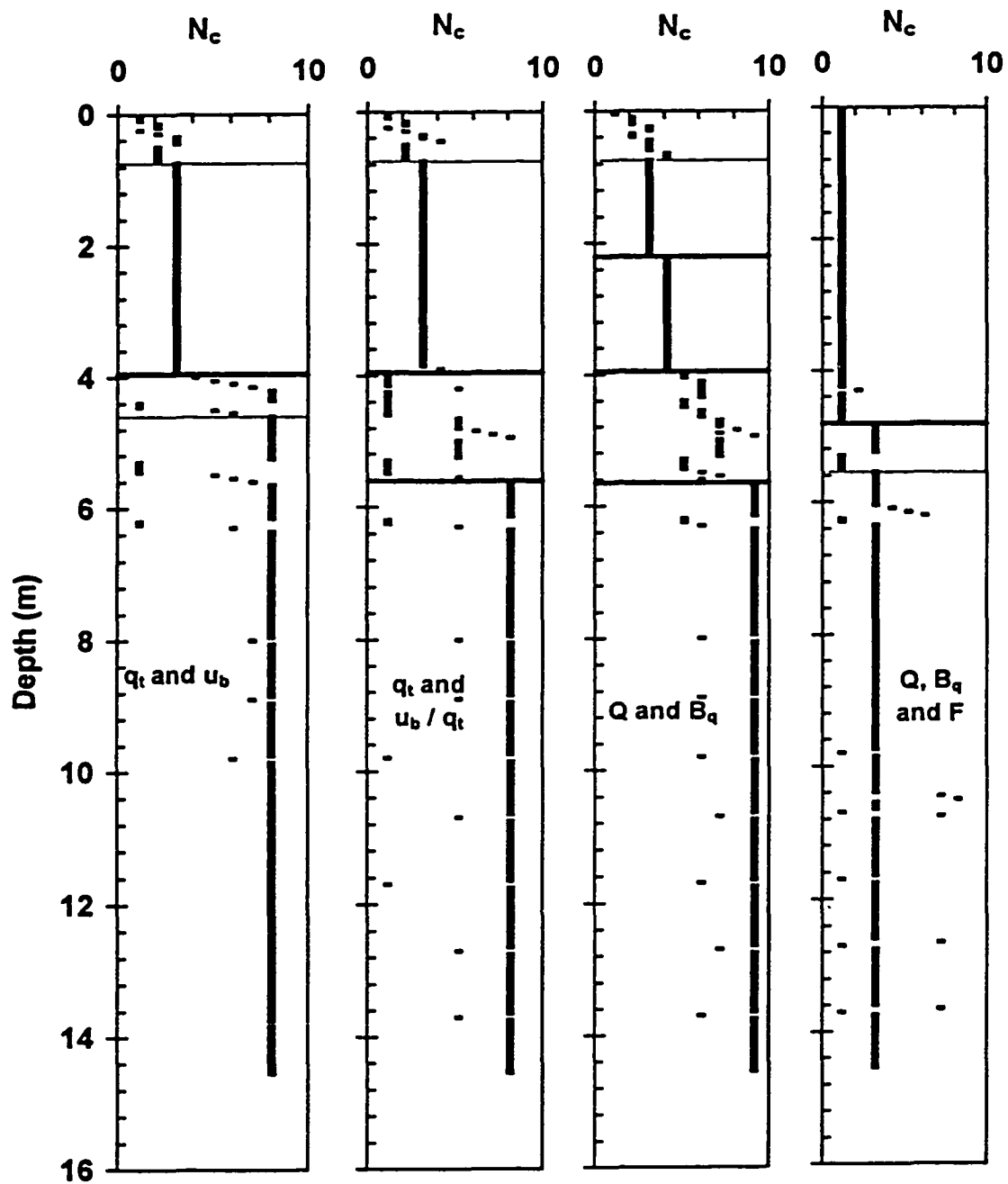
1. unprocessed data: q_t and u_b ,
2. lightly-processed data: q_t and the ratio u_b/q_t , and
3. normalized parameters: Q , B_q and $F = [f_s/(q_t - \sigma_{vo})]$.

The second group of data is closer to the form of Q and B_q neglecting the effect of the in-situ vertical stresses. The presentation of the piezocone data in terms of the normalized parameters, Q , B_q and F was proposed by others (e. g. Wroth, 1988; Robertson, 1991).

Also, this group includes the influence of f_s on the analysis. The cluster results of the three data combinations are compared with the results using Q and B_q at $N_c = 8$ as shown on Fig. 5.8. The growth of their cluster up to $N_c = 100$ is discussed in detail in Appendix C.

Using q_t and u_b indicates a secondary boundary and a primary boundary at 0.7 m and 3.9 m, respectively. However the primary boundaries indicated using Q and B_q at 2.2 m and 5.6 m are not detected. Using q_t and the ratio u_b/q_t gives similar boundaries to those obtained using Q and B_q except for the boundary at 2.2 m which is not retrieved. Note that using Q , B_q and F improperly indicates the soil stratigraphy at Amherst because the primary boundaries at approximately 2 m, 4 m and 5.6 m are not detected. However primary and secondary boundaries are found at 4.8 m and 5.5 m, respectively.

The proper soil stratigraphy is demarcated using the three piezocone combinations at $N_c > 8$ which is discussed in detail in Appendix C and summarized as follows: (1) unprocessed data, q_t and u_b at $N_c = 26$, (2) partly-processed data, [q_t and the ratio u_b/q_t] at



Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).
 SCZ method at $N_c = 8$

— Primary boundary
 — Secondary boundary

Figure 5.8. Effect of Using Raw and Processed Piezocone Data on the Cluster Analysis at Amherst, Massachusetts.

$N_c = 33$ and (3) normalized parameters, Q , B_q and F at $N_c = 30$. Therefore, using Q and B_q is superior to the other investigated combinations of cone data because cluster results represent the soil stratigraphy at Amherst at a relatively small cluster number $N_c = 8$.

5.4. Factors Affecting Cluster Results

Several factors affect piezocone data and their interpretation for the purpose of geostratigraphy including measurement errors, assumed unit weight, data frequency, porous filter position, and cone size. Detailed discussion of the influence of these factors on clustering of piezocone data for the purpose of soil stratigraphy is given in the following sections.

5.4.1. Data Errors

In routine assessment of piezocone data, errors in the data set affect the interpretation by empirical, analytical, and numerical methods. For autocorrelation and variogram analysis, errors can also affect the fitted models and interpolated data. A superiority of the clustering technique is that the method is unaffected by either random or systematic errors as discussed herein in two case studies.

5.4.1.1. Effect of Electrical Noise

An example of the electrical noise error-type is presented with q_c measurements shown in Fig. 5.9 and taken using a miniature electric cone in chamber tests of 50/50 Peerless kaolinite and fine silica (Mayne et al., 1992). The water content of the clay was originally equal to 66 % which was twice the liquid limit. The tested clay consisted of two layers. Both layers were prestressed using a perforated rigid piston and an applied pneumatic pressure to preconsolidate the clay to 50 kPa.

In the first (lower) clay layer, the clay was covered at the top and the bottom by geotextile filter layers. The drainage was allowed at the top through the piston and the bottom through a sand layer (i.e. double drainage) during the first prestressing. To obtain a thicker soil section, a second clay layer was laid over the top of the first one and prestressed also to 50 kPa. Drainage was impeded beneath of the lower clay layer (i.e. single top drainage). Thus, a two layer system was formed from the same 50/50 kaolinite clay. The boundary between the two deposits occurs at 220 mm from the top soil surface and shown in Fig. 5.9.

A single-cosine-zscore cluster-type was applied to the measured and filtered q_c from two penetration tests. The data were filtered using a moving window average technique where a window of 5 data points was chosen and every 5 measurements were replaced by their average at a depth representing the center of each window. The cluster analysis of the measured and filtered data are compared at N_c equal to 2 and appear to be in good agreement as shown in Fig. 5.10. Homogeneous clusters are indicated by cluster number 1 above 22.5 cm, and cluster number 2 below 36 cm using both data. The cluster

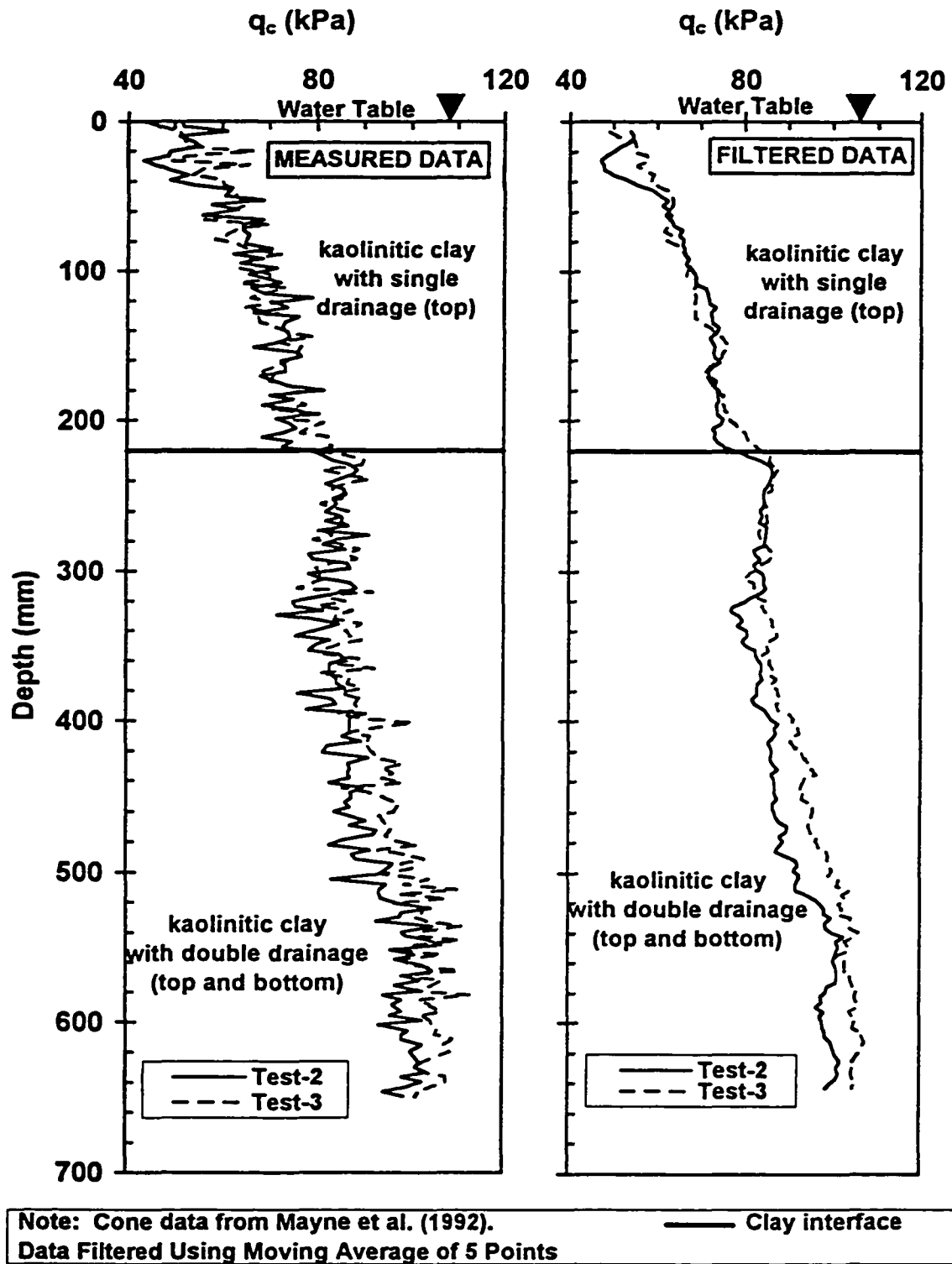


Figure 5.9. Measured and Filtered q_c Data Collected in Chamber Deposit of Kaolinitic Clay (Data from Mayne et al., 1992).

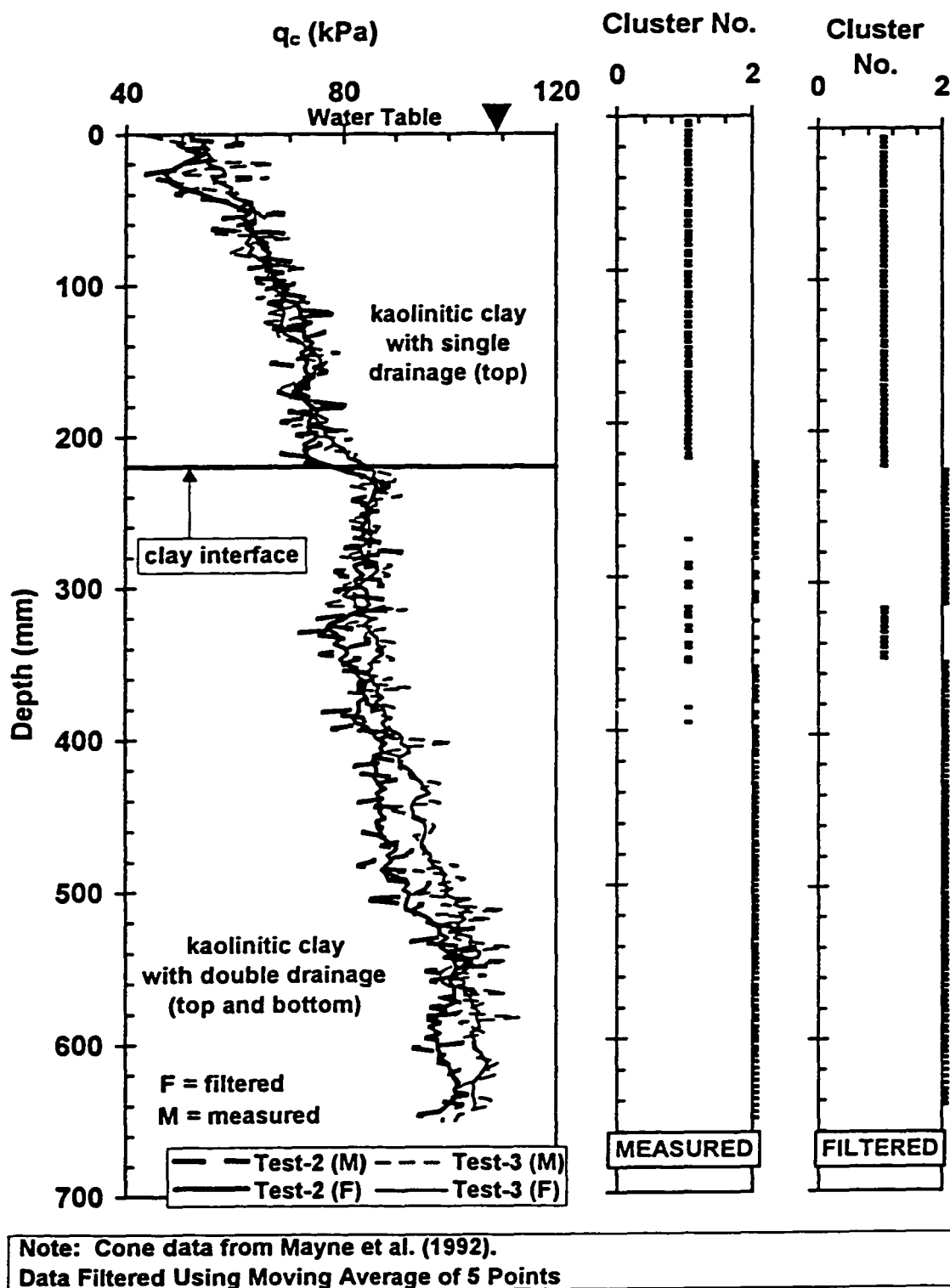
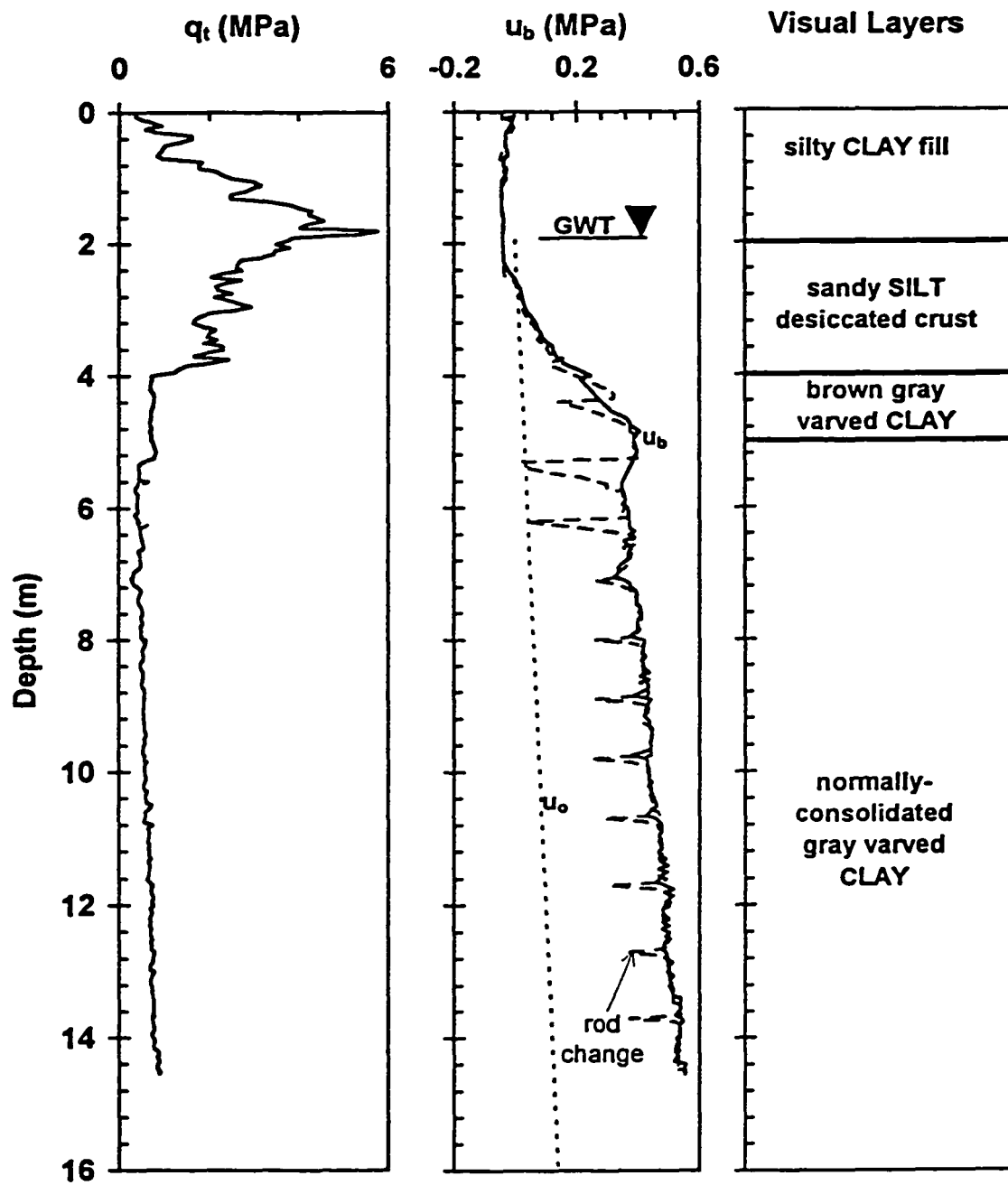


Figure 5.10. Data Noise Effect on Cluster Analysis of q_c Data in Chamber Deposit of Kaolinitic Clay (Data from Mayne et al., 1992).

results of the measured data indicate that the soil between 22.5 cm and 27.5 cm belongs to the clay layer below 36 cm. A soil mixture is indicated between 27.5 cm and 36 cm and this is due to non-homogeneous mixture of the kaolinite and the silica and/or different soil properties. Using the filtered data in the cluster analysis indicated a secondary layer between the depths of 32 cm and 36 cm which represents a change in the soil type, plasticity and/or consolidation behavior.

5.4.1.2. Effect of Procedural Errors

A study was performed to quantify the effect of procedural errors included in piezocone data collected at Amherst due to frequent stops to add a cone-rod. The latter error-type sometimes results in a decrease of the q_c measurements and/or a dissipation of the porewater measurements, especially in clay soils where creep effects can occur. Data measurement errors due to frequent stops during each 1-m of penetration appear in the piezocone results collected at Amherst as depicted in Fig. 5.11. A data filtering criterion proposed by Vivitrat (1978) was used to define the data errors at the rod breaks approximately every one-meter depth. In order to perform the analysis, a window width was chosen equal to 0.5 m and a cone reading was considered an outlier if it is greater than (average + 2 standard deviation) of the data within a window. The data errors were deleted and replaced by linear interpolation between the two data points above and below the removed measurements. Filtered data are shown in Fig. 5.11 in comparison with the measured unprocessed data.

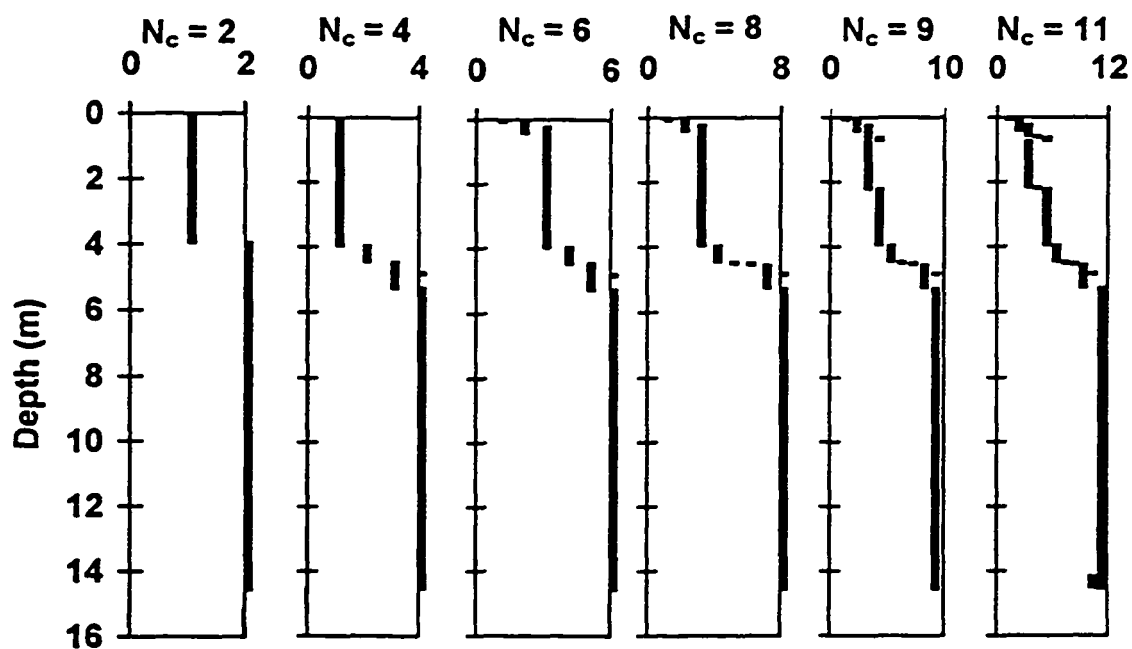


Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).

--- Measured — Filtered

Figure 5.11. Measured and Filtered Piezocone Data (PCPT1) and Soil Profile at Amherst, Massachusetts (Data from This Study).

A cluster analysis is applied to the filtered data using the single-cosine-zscore (SCZ) method. The correlation coefficient between consecutive cluster results is discussed in Appendix C up to $N_c = 100$. The peaks of ρ_c are defined between $N_c = 2$ and 50 and Fig. 5.12 shows the cluster results at the peaks between $N_c = 2$ and 11. At $N_c = 2$, the cluster results divide the stratigraphy into 2 primary layers with a boundary at 3.9 m. Thereafter, soil lenses separate up to $N_c = 6$ and a primary layer splits at $N_c = 9$ between 2.2 m and 3.9 m. Subsequently, no new primary clusters appear up to $N_c = 100$, however, more growth of transition zones and lenses is indicated.



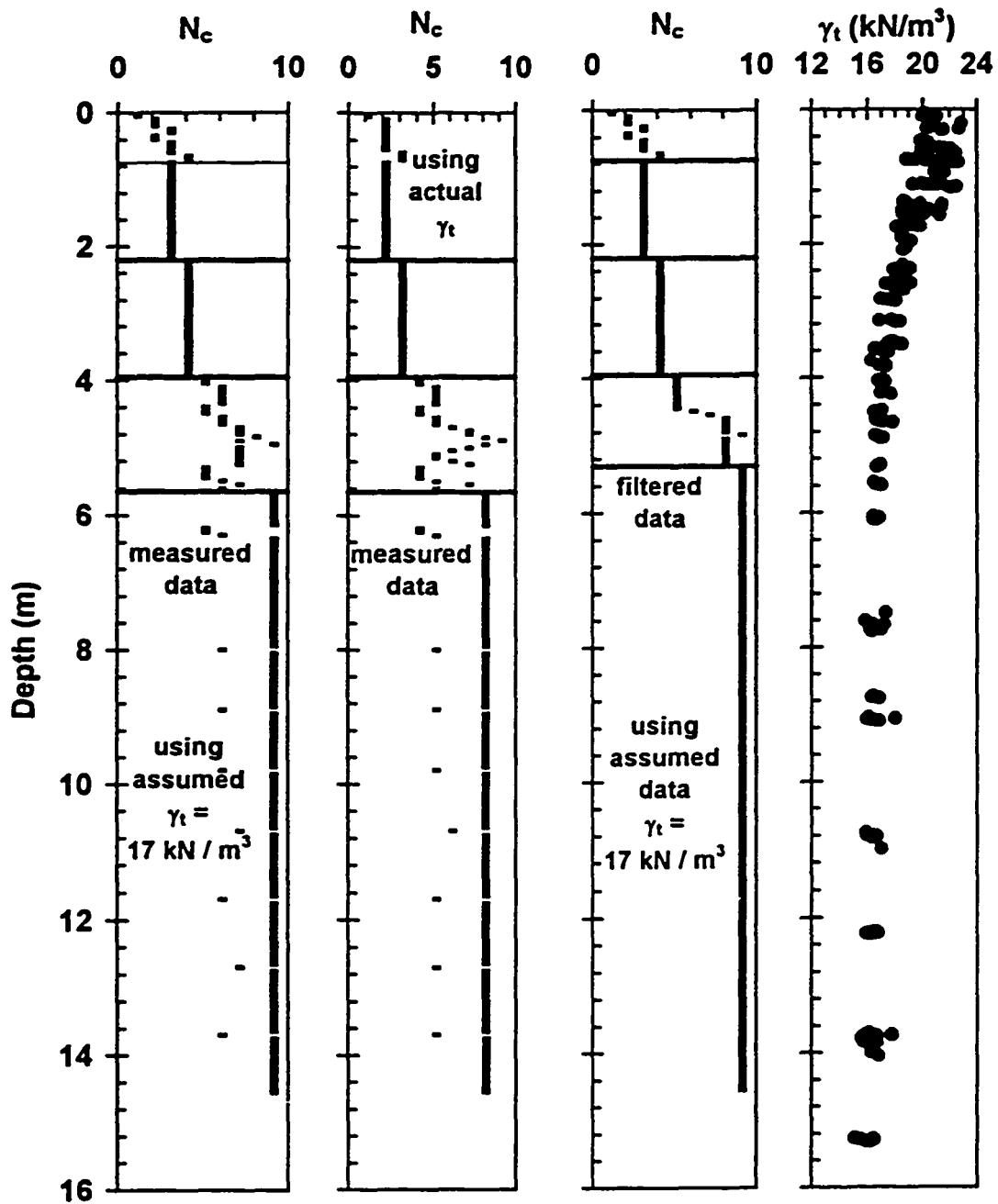
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters

Figure 5.12. Cluster Analysis of Filtered Piezocone Data (PCPT1) at Amherst, Massachusetts (Data from This Study).

A cluster number $N_c = 9$ is chosen to represent the soil stratigraphy. The primary layer boundaries are defined at 2.2 m, 3.9 m and 5.3 m. Figure 5.13 shows a comparison between the cluster results of the measured and filtered data at $N_c = 9$. The cluster results of the filtered data are very similar to those of the measured data. Note that the subsurface stratigraphy was detected at Amherst at cluster number $N_c = 9$ using the filtered data compared with cluster number $N_c = 8$ using the measured data. Therefore, there is no need to filter the piezocone data for clustering for the purpose of soil stratigraphy. Moreover, some soil lenses which can be important in a geotechnical design, can be missed by filtering the data.

5.4.2. Actual Versus Assumed Unit Weight

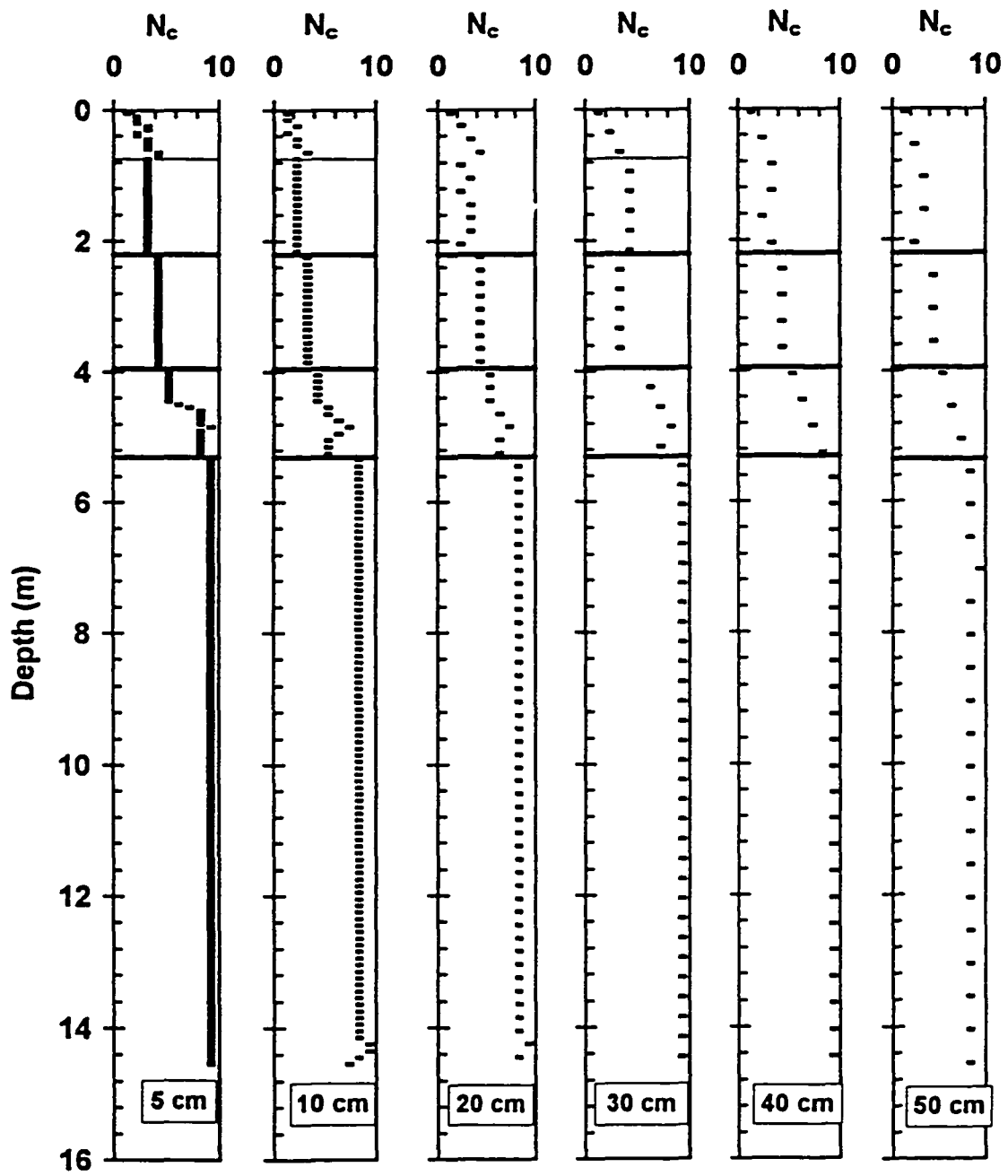
In order to use piezocone normalized parameters $Q = [(q_t - \sigma_{vo})/(\sigma_{vo}')]$ and $B_q = [(u_b - u_o)/(q_t - \sigma_{vo})]$, the total $[\sigma_{vo} = \gamma_t \cdot \text{depth}]$ and effective $[\sigma_{vo}' = \sigma_{vo} - u_o]$ overburden stresses are needed. Their values are calculated at Amherst using the actual unit weight, γ_t which is shown in Fig. 5.13, as opposed to an assumed unit weight. Then a single-cosine-zscore cluster method is applied to the normalized parameters Q and B_q from one piezocone (PCPT1). The correlation coefficient between consecutive cluster results is calculated up to $N_c = 100$ and examined clusters are chosen at the peaks of ρ_c . Figure 5.14 shows the cluster results at the peaks between $N_c = 2$ and 12. At $N_c = 2$, the cluster results are the same as using an assumed constant value of $\gamma_t = 17 \text{ kN} / \text{m}^3$. However, at $N_c = 4$, the upper cluster is divided into 3 primary groups. For $N_c > 4$, points weakly correlated with



Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).
 Note: γ_t = Total unit weight
 SCZ method using Q and B_q at $N_c = 9$

— Primary boundary
 - - - Secondary boundary

Figure 5.13. Effect of Using Filtered Data and Assumed Versus Measured Unit Weight on the Cluster Analysis at Amherst, Massachusetts.



Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).
 SCZ method using Q and B_q at $N_c = 9$
x cm = resolution
 — Primary boundary
 - - - Secondary boundary

Figure 5.14. Effect of Data Resolution of the Filtered Data on the Cluster Analysis at Amherst, Massachusetts Using 10-cm² Cone.

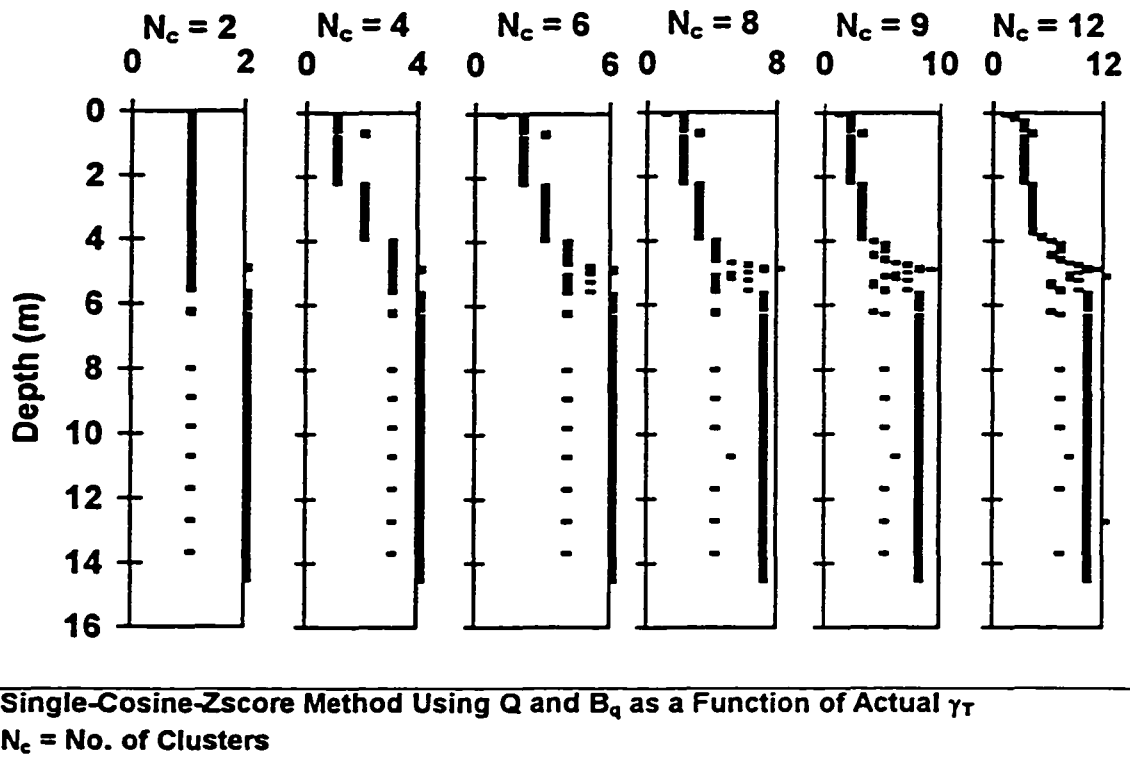


Figure 5.15. Effect of Using Actual γ_t on the Cluster Analysis of Piezocone Data (PCPT1) at Amherst, Massachusetts (This Study).

the 3 groups continue to separate indicating lenses, transitions, or data errors up to $N_c = 12$, however, no new primary layers are detected. The cluster results up to N_c equal to 100 are discussed in Appendix C and indicates more growth of the transition zones without primary group separation. Therefore, a cluster number 4 is chosen to represent the soil stratigraphy at the site. The primary layer boundaries are defined at depths of 2.2 m, 3.9 m and 5.6 m.

Using the actual measured γ_t helped to get very similar cluster results at $N_c = 4$ as those obtained using an assumed γ_t ($17 \text{ kN} / \text{m}^3$) at double the cluster number. Looking at the actual unit weights in Fig. 5.13, the values of γ_t gradually decrease from 23 kN/m^3 to approximately 17 kN/m^3 between the ground surface and a depth of 2.5 m, therefore the values of Q using actual γ_t are less than those using $\gamma_t = 17 \text{ kN/m}^3$. This helps to indicate the difference between the upper fill layer and the clay crust layer between the depths of 2 m and 4 m at a lower cluster number as shown in Figures 5.12 and 5.14. For comparison, Fig. 5.13 indicates a similarity of the primary boundaries at $N_c = 9$ using assumed and actual γ_t . Therefore, there was no need to measure the unit weight at Amherst to apply the cluster analysis, however, the subsurface stratigraphy at the site is detected at larger cluster number using a constant assumed unit weight.

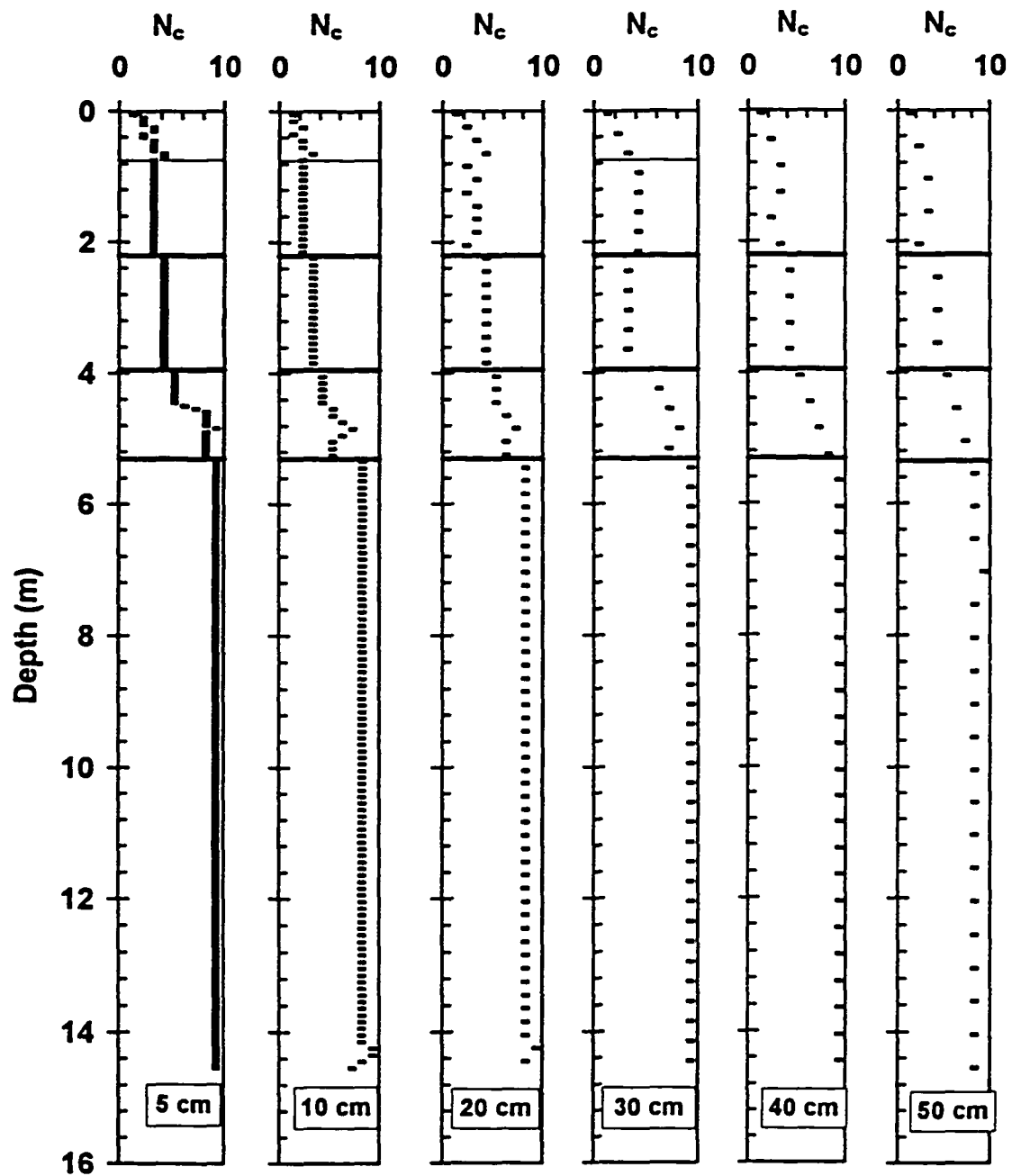
5.4.3. Data Frequency

Piezococone data are commonly measured at small intervals from 1 cm to 5 cm. More details of the inherent variability of a soil stratigraphy can be known by increasing the frequency of collecting the cone data or other means of testing. However, in other laboratory and in-situ testing, data are usually collected at lower frequencies. For instance, flat dilatometer readings are recommended to be taken every 20 cm (Marchetti, 1980). Therefore, SCZ cluster-type is performed at Amherst are using six chosen frequencies where single readings taken every 5 cm, 10 cm, 20 cm, 30 cm, 40 cm and 50 cm. The frequency (r) is not reduced further to be able to distinguish between soil layers ($t \geq 0.5 \text{ m}$), and soil lenses, transitions and outliers ($t < 0.5 \text{ m}$), according to the

interpretation criterion defined in this chapter. Note that the filtered data are used to eliminate the errors due to dissipation effect which might not be available in other in-situ or laboratory methods. The development of clustering of different frequencies is discussed in Appendix C. Cluster results are compared at $N_c = 9$ and Figure 5.15 indicates similarity of the primary boundaries detected using $r = 5$ cm to $r = 50$ cm. The main statistical boundaries are at depths of 2.2 m, 3.9 m, and 5.3 m. However, the secondary boundary at 0.7 m defined using $r = 5$ cm is not identified using $r = 20, 40$ and 50 cm. Therefore, the delineated subsurface stratigraphy is not affected by increasing the distance between the data to 50 cm. This suggests that clustering can be applied to other laboratory or in-situ testing in which the frequency of the data is usually smaller than that of the piezocone readings.

5.4.4. Porewater Pressure Position

It is well recognized that the position of porous element can result in significantly different readings (e. g. Mayne et al., 1990; Lunne et al., 1997). Two case studies are discussed in this section. First, a comparison is performed between the cluster results using ($u_1 = u_t$) and ($u_2 = u_b$) in soft clay at the Bothkennar test site, Scotland, where both pore pressure measurements are positive and increase with depth. Then, a similar study is performed using piezocone data collected in stiff fissured clay at Brent Cross test site, UK, where u_b readings are mostly negative and u_t readings are positive.



Note: Piezocone data (PCPT1) at Amherst, Massachusetts (this study).
 SCZ method using Q and B_q at $N_c = 9$
 x cm = frequency
 — Primary boundary
 — Secondary boundary

Figure 5.16. Effect of Data Frequency of the Filtered Data on the Cluster Analysis at Amherst, Massachusetts Using 10-cm² Cone.

5.4.4.1. Bothkennar Test Site, Scotland

Usually, porewater pressure measurements are taken at the midface u_t or at the shoulder u_b . Porewater pressure measurements on the face of the cone are always positive. However, at the shoulder position, u_b measurements can be positive, zero, or negative (Robertson et al., 1986; Senneset et al., 1988). Piezocone soundings with both u_t and u_b measurements were taken at the Bothkennar test site, Scotland as shown on Fig.

5.16. A very detailed geotechnical study was performed at this site, (for instance, Hawkins et al., 1989; Nash et al., 1992). The soil stratigraphy from extensive boring and sampling information at the site consists of a silty clay crust in the upper 2.5 m underlain by a soft to firm homogeneous silty clay to clayey silt down to 19.5 m. A soil classification by Q and B_{q2} was performed using the Robertson (1991) chart as shown on Fig. 5.17. A soil layer of silty clay to clayey silt is defined between 1.35 m and 2.7 m underlain by a clay to silty clay layer down to 19.5 m. Therefore, Robertson's method properly indicates the subsurface stratification at the site.

A cluster analysis was performed using Q and B_q with the latter term defined as a function of u_b measurements. The clusters at the peaks of ρ_c up to $N_c = 50$ are examined and Fig. 5.18 shows the results between $N_c = 2$ and 6. Two primary clusters appear at $N_c = 3$ with a boundary at 3.7 m. The growth of the cluster analysis is also studied up to $N_c = 100$ and discussed in Appendix C. No new primary clusters with $t \geq 1$ m appear at $N_c > 3$, however, lenses and transitions continue to separate the two main statistical layers. Therefore, cluster number 3 is chosen to delineate the soil stratigraphy. An extensive data

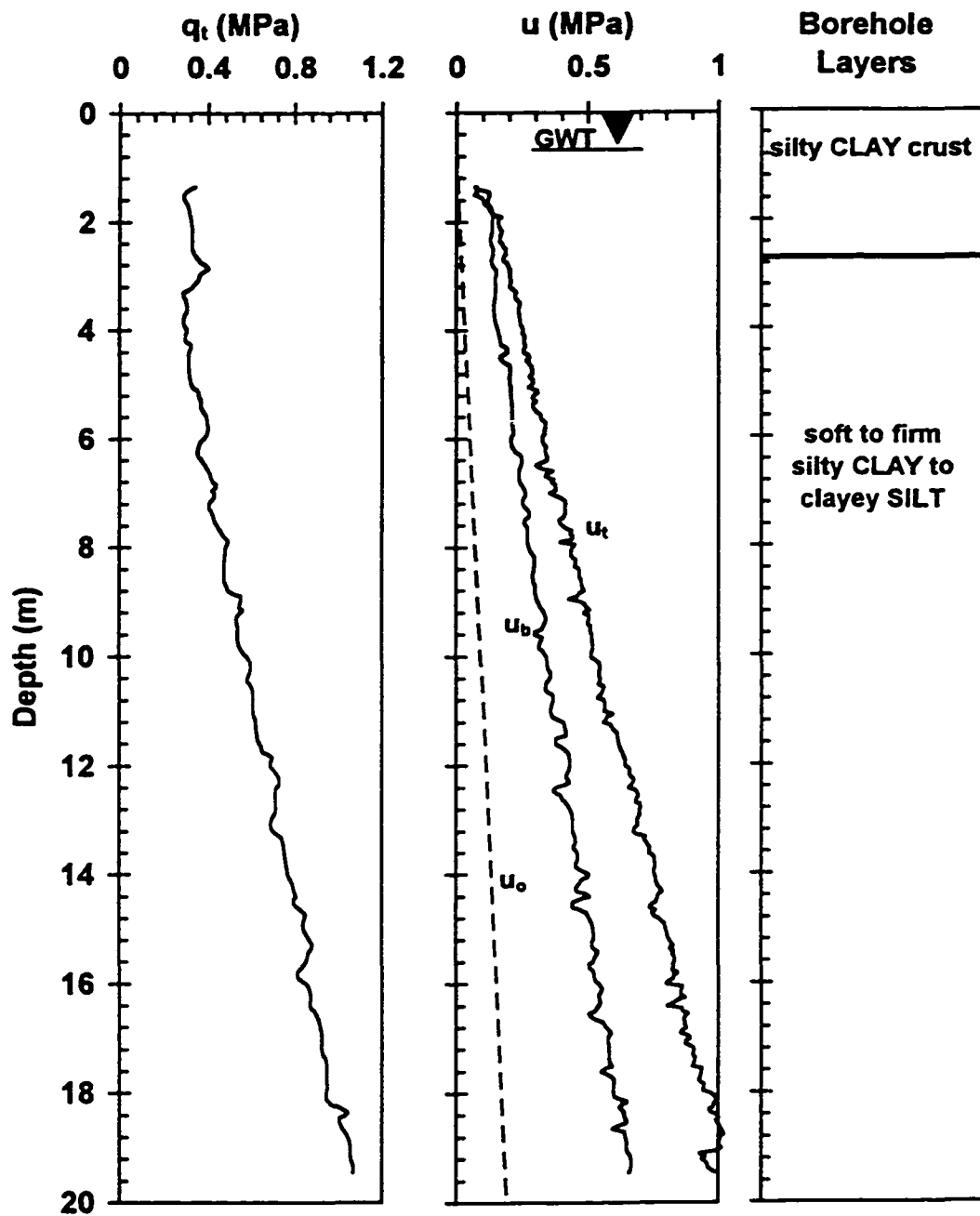


Figure 5.17. Piezocone Data and Soil Stratigraphy at Bothkennar, UK
(Data from Nash et al., 1992).

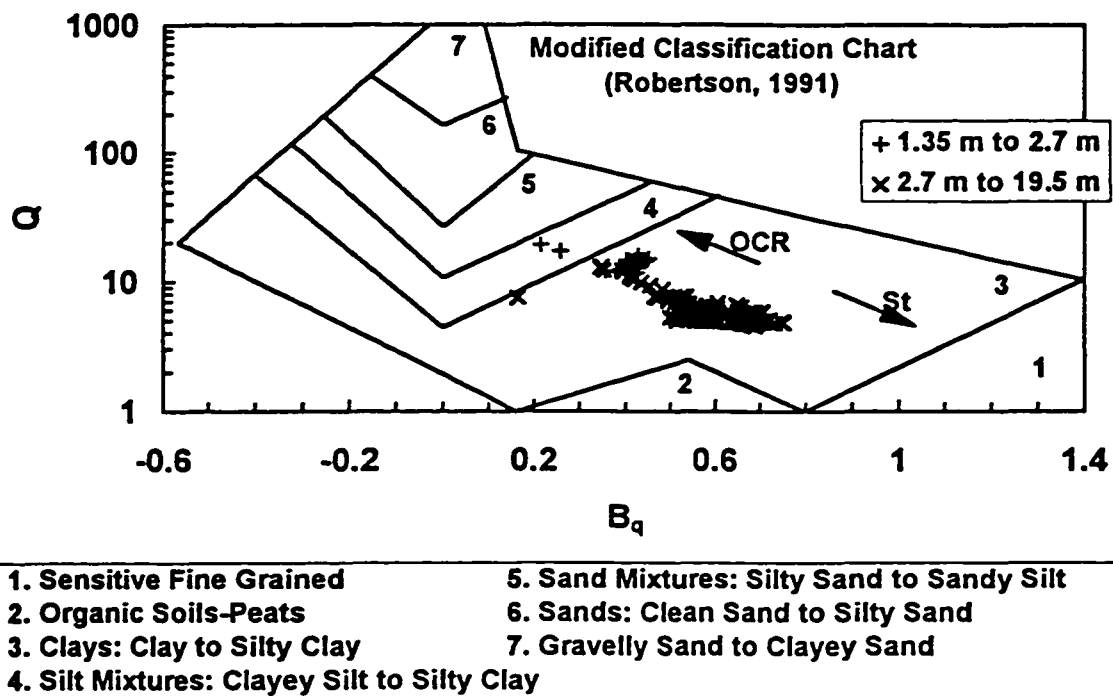


Figure 5.18. Soil Classification Using Piezocone Data at Bothkennar, UK (Data from Nash et al., 1992).

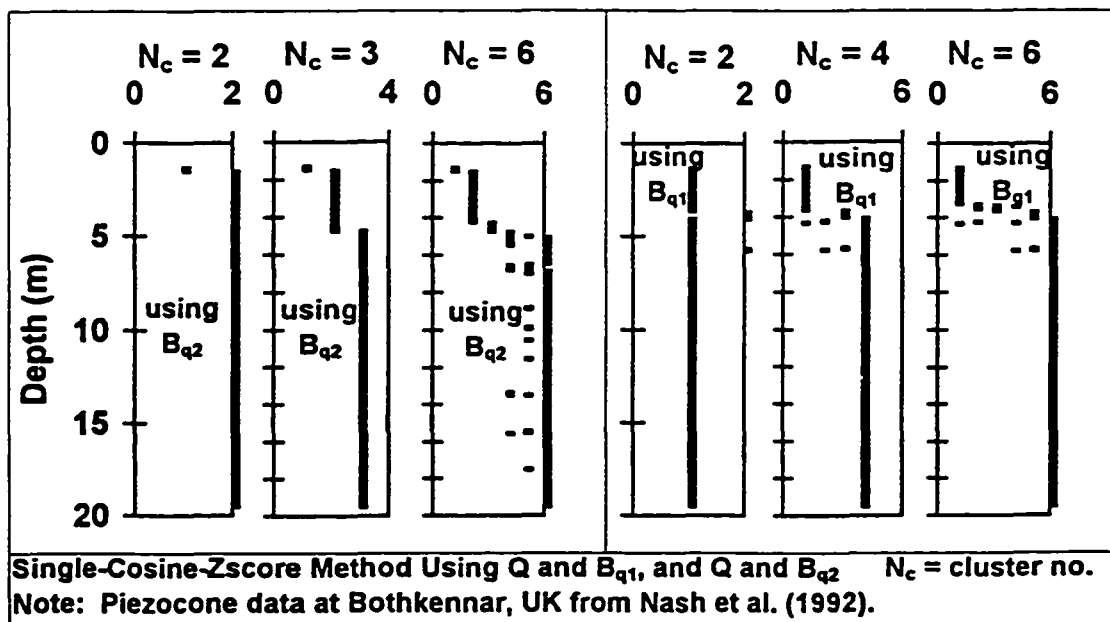
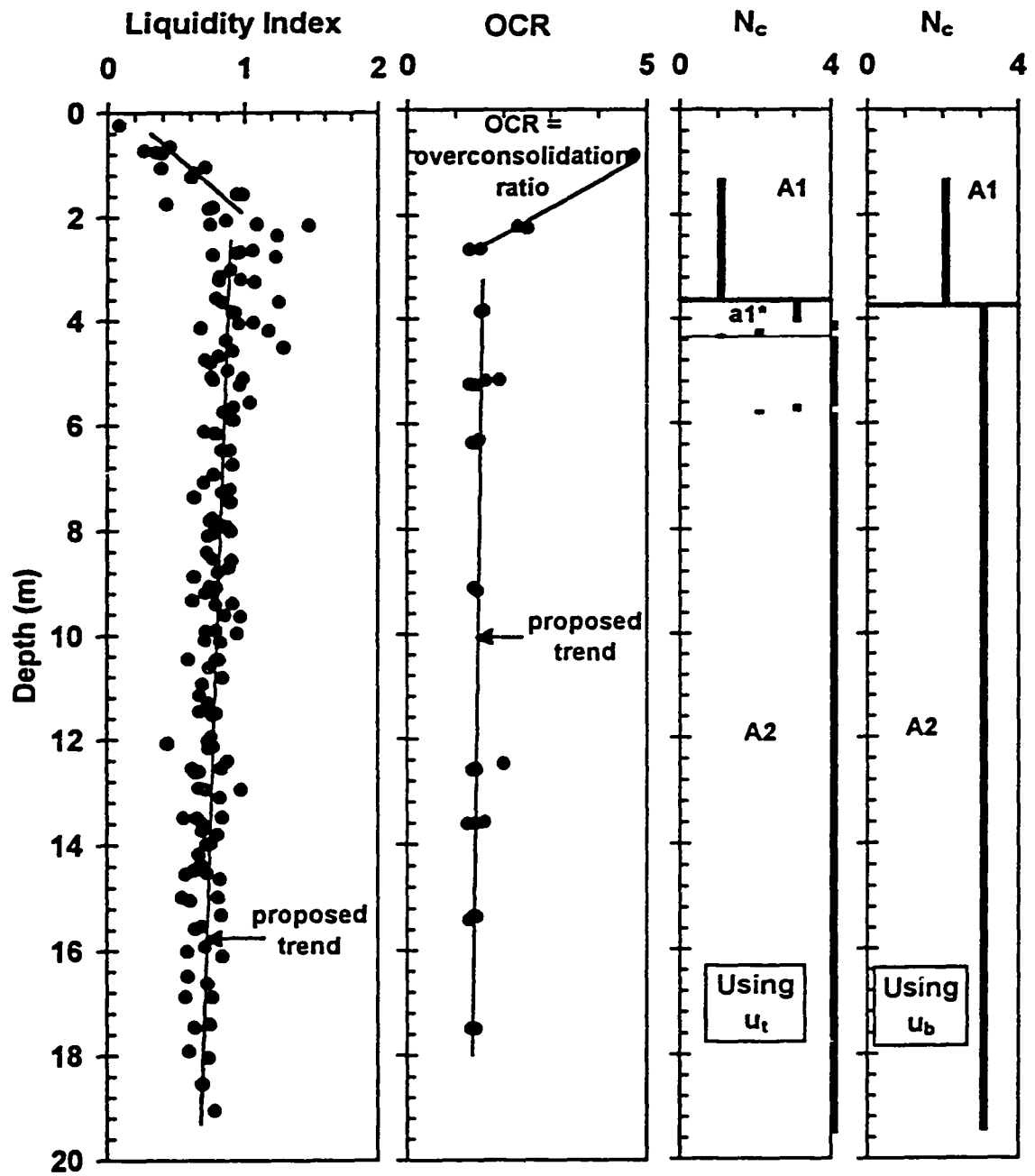


Figure 5.19. Clustering of Types 1 and 2 Piezocone Data at Bothkennar, UK.

sets of liquidity index and overconsolidation ratio confirm the statistical stratification as depicted in Fig. 5.19. Also, fitted trends are included in the figure.

A similar cluster analysis is performed using Q and B_q with the latter defined as a function of u_t readings. The cluster results at the peaks of ρ_c between $N_c = 2$ and 6 are shown on Fig. 5.18. A cluster number 4 is chosen to denote the soil stratigraphy because no new primary clusters split at larger N_c . The cluster results as a function of u_b and u_t indicated the existence of two primary clay layers in the deposit with a boundary at approximately 3.7 m. Figure 5.19 indicates a good agreement for primary layer detection derived from u_t and u_b piezocones used in clustering, except that, a secondary layer $a1^*$ is defined between 3.6 m and 4.4 m using the u_t measurements. The primary layers indicated by clustering are verified by both the trends of liquidity index and overconsolidation ratio. Experimental, statistical and analytical correlations between piezocone data and overconsolidation ratio were established (Wroth, 1988; Chen and Mayne, 1994). The shear strength of the soil increases by decreasing the liquidity index. The cone data are functions of the shear strength of the soil (Campanella and Robertson, 1988), therefore, the cone data are also functions of the liquidity index of a clayey soil. Looking at the lower trend of liquidity index, the data are relatively more scattered between the depths of 2 m and 4 m which can support the existence of a secondary or a transition layer between the two primary layers. Therefore, using u_t in the analysis gives a better indication of the soil stratigraphy at Bothkennar test site.



Note: Piezocone data at Bothkennar, UK from Nash et al. (1992).
 SCZ method using Q and $B_q(u_b)$, cluster no. 3, and $B_q(u_t)$, cluster no. 4
 — Primary boundary — Secondary boundary

Figure 5.20. Comparison Between Clustering Using Midface Penetration Porewater Pressure (u_b) Versus Shoulder (u_t) with Profiles of Liquidity Index and Overconsolidation Ratio at Bothkennar, UK.

5.4.4.2. Brent Cross, England

The soil stratigraphy at Brent Cross test site, UK consists of heavily overconsolidated and fissured London clay (Lunne et al., 1986; Powell et al., 1988). The weathered clay down to a depth of 9 m has fissure spacings between 6 mm and 50 mm, however, the unweathered clay below 9 m and down to 25 m has fissure spacings between 75 mm to 325 mm. An example of piezocone data collected at the site is shown in Fig. 5.20. The porewater pressure measured behind the tip is negative due to soil dilation (Senneset and Janbu, 1989), however the porewater pressure readings at the cone face are positive due to soil compression.

By simply looking at the unprocessed q_t and u measurements, no clear soil boundaries are evident. However, it is noticed that the variability of the q_t measurements increases with depth and the u_b readings are positive down to 2.5 m and then become negative. The soil is classified using the Robertson PCPT chart (1991) as presented in Fig. 5.21. The soil between the depths of 0.5 m and 10.7 m is defined as sandy silt to silty sand to clean sand underlain by clayey silt to silty clay down to a depth of 18.5 m. Therefore, this method is not able to properly identify the primary boundary between the upper and lower soil layers.

A single-cosine-zscore cluster analysis is applied to the Q and B_{q2} from Brent Cross. The clusters at the peaks of ρ_c up to $N_c = 50$ are examined and Fig. 5.22 shows the results between $N_c = 2$ and 7. At $N_c = 5$, three primary clusters appear with boundaries at 2.4 m and 7.9 m and a transition starts to build up between depths of 7.6 m and 7.9 m. The growth of the cluster analysis up to $N_c = 100$ is discussed in Appendix C. No new

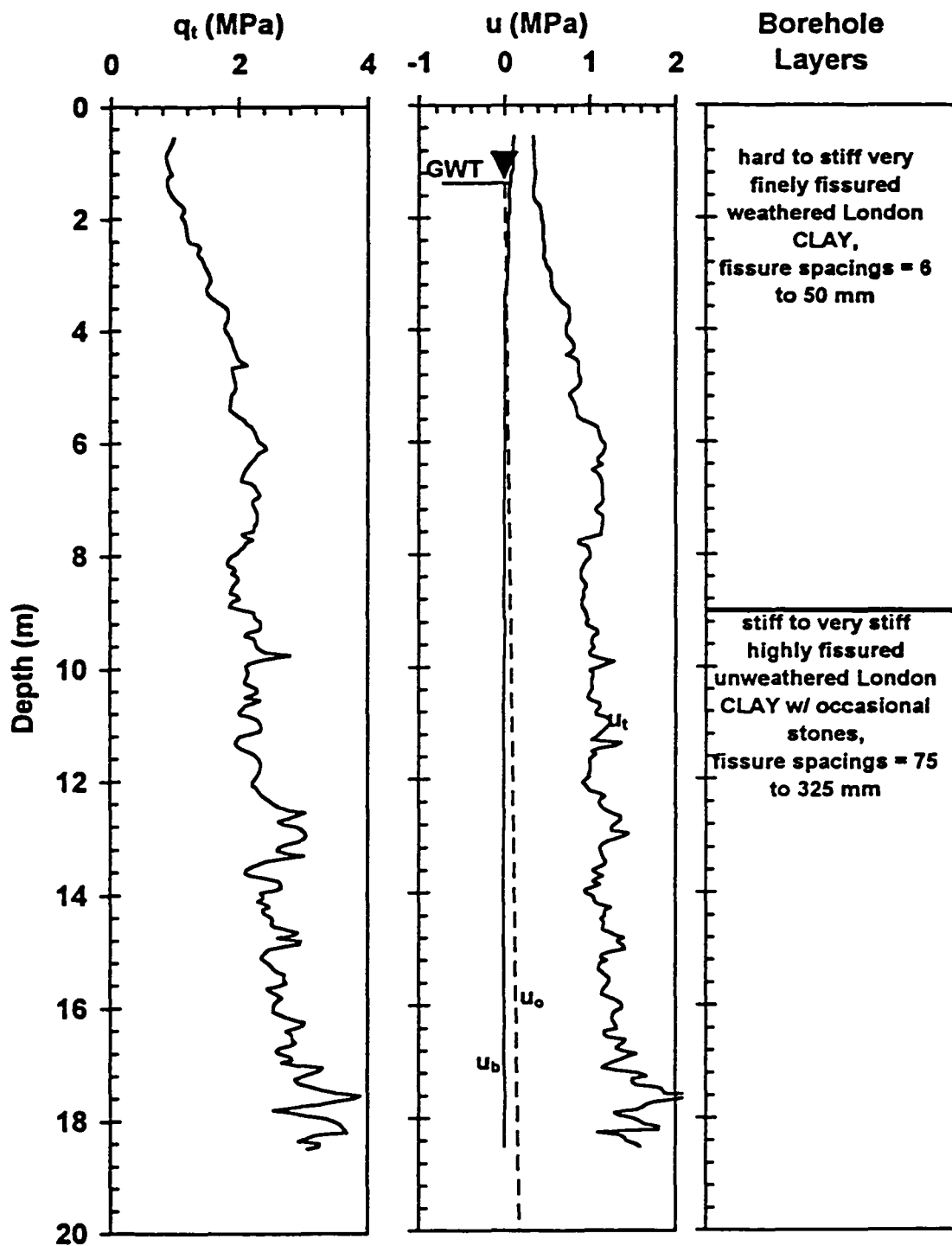


Figure 5.21. Piezocone Data and Soil Stratigraphy at Brent Cross, UK
 (Data from Powell et al., 1988).

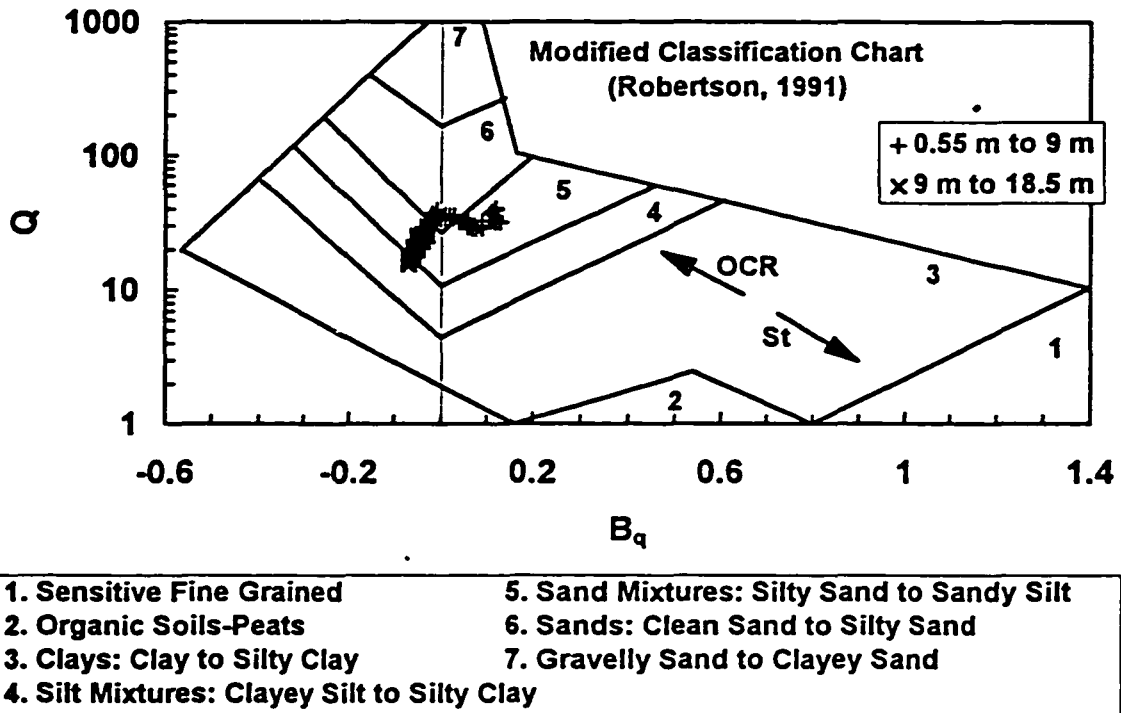


Figure 5.22. Soil Classification Using Piezocone Data at Brent Cross, UK
(Data from Powell et al., 1988).

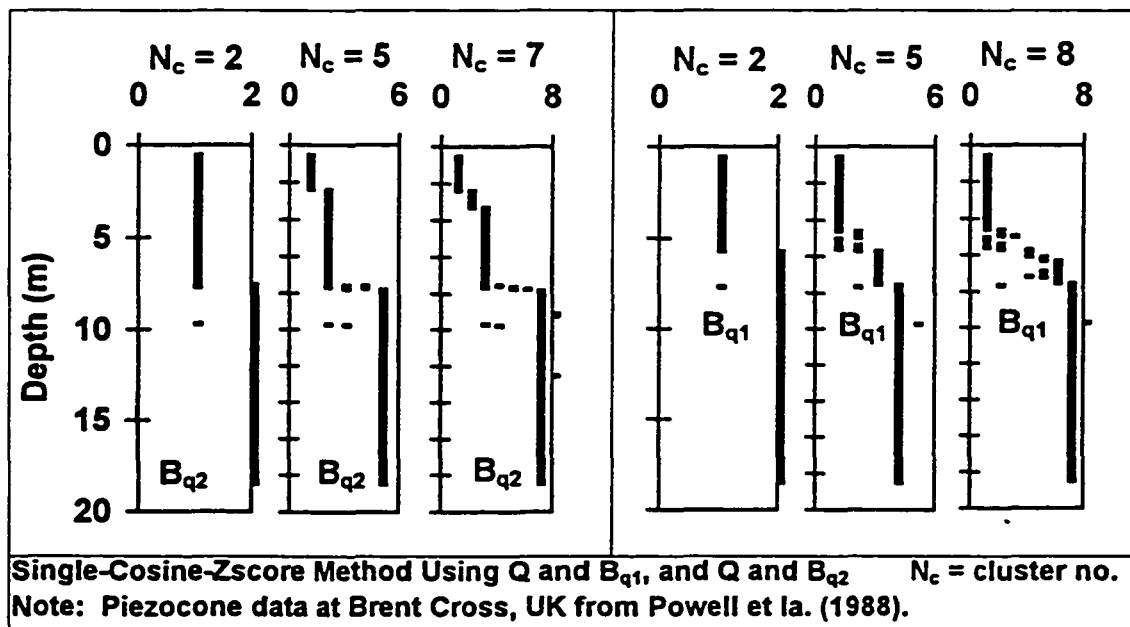
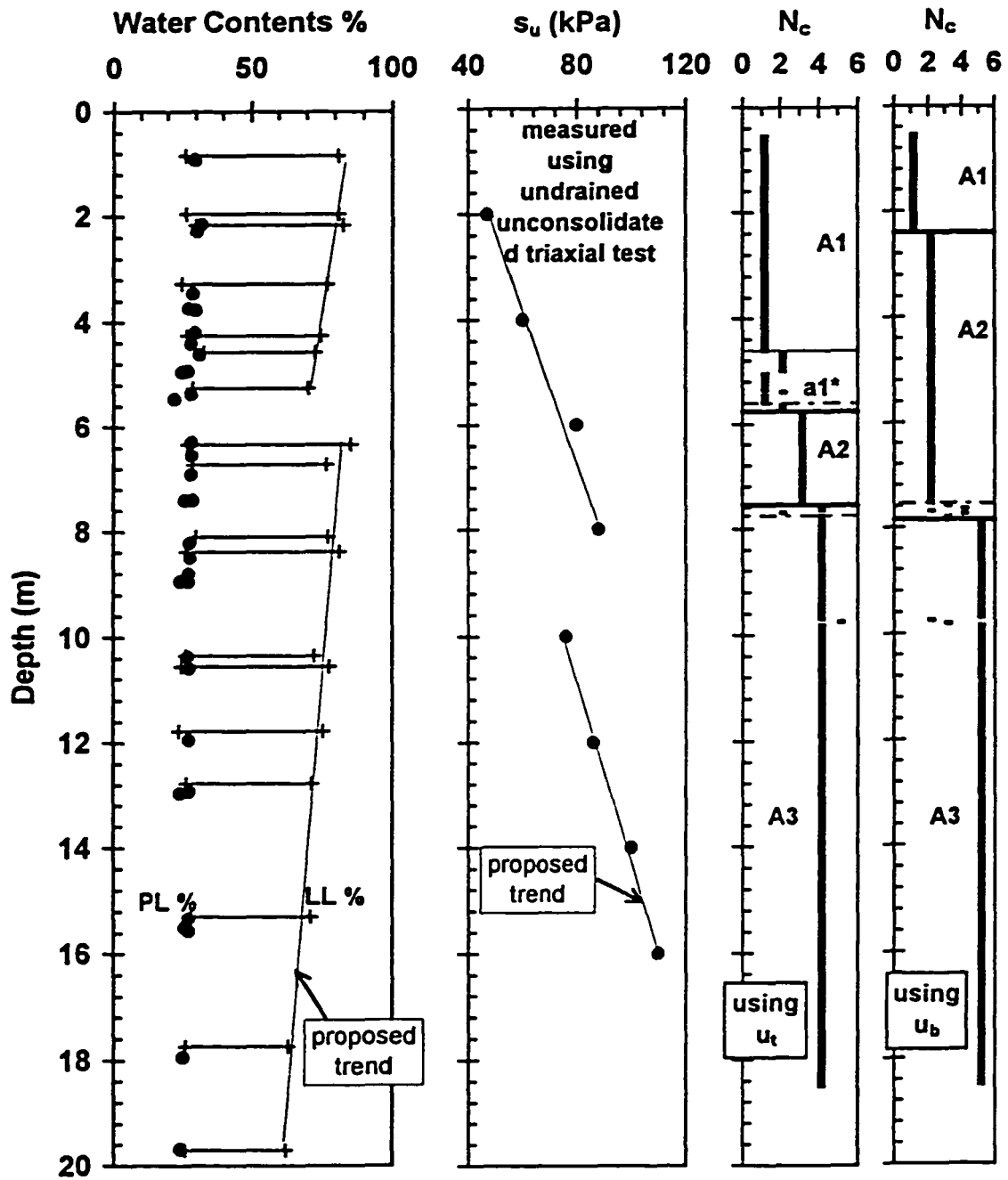


Figure 5.23. Clustering of Type 1 and 2 Piezocone Data at Brent Cross, UK.

primary clusters ($t \geq 1$ m) are separated at N_c greater than 5, however, lenses and transitions continue to separate. Therefore, cluster number $N_c = 5$ is chosen to represent the soil stratigraphy at the site. The upper two primary layers are associated in terms of soil type and/or soil properties because they are denoted by cluster numbers 1 and 2. However, the lower main layer could have different characteristics than the upper two layers because it is assigned cluster number 5 and the transition zone between this stratum and the intermediate deposit increases by increasing N_c .

The vertical profiles of water contents and undrained shear strength (s_u) data measured using undrained-unconsolidated triaxial tests are shown in Fig. 5.23. The natural water contents and the plastic limits slightly change with depth and their averages are the same and equal to 27.5 percent. There are two proposed trends of the liquid limits with depth as shown in Fig. 5.24. For the first trend, the liquid limits change from 81 percent at 0.8 m to 70 percent at 5.2 m. For the second trend, the liquid limits decrease from 85 percent to at 6.3 m to 62 percent at 19.7 m. The water content profiles do not verify the obtained clustering results using u_b measurements to derive B_q .

The undrained shear strength (s_u) data increase with depth down to almost 8 m where their trend is shifted to the left as shown on Fig. 5.23. The variation of s_u with depth confirms the defined soil boundary at 7.9 m which also reflects the change of the soil structure and fissure size between the upper and lower layers. Powell and Quarterman (1988) noted the effect of fissure size on both soil properties and cone readings in clay soils, however, studies should be performed to quantify their effect on stresses measured



Note: Piezocone data at Brent Cross, UK (Powell et al., 1988).
 SCZ method using Q and B_q , cluster no. 5
 — — — Transition boundary
 — Primary boundary
 — Secondary boundary

Figure 5.24. Comparison Between Stratification from Cluster Results Using u_t and u_b and Soil Properties at Brent Cross, UK.

by the cone penetrometer. The upper cluster layer A1 between depths of 0.4 m and 2.4 m is influenced by the transition effect between positive and negative u_b readings.

Another cluster analysis is performed using Q and B_{q1} and the cluster results at the peaks of ρ_c between $N_c = 2$ and 7 are shown on Fig. 5.22. At $N_c = 5$, three primary clusters are separated with boundaries at 5.75 m and 7.70 m. The growth of the cluster analysis up to $N_c = 100$ is discussed in Appendix C. No new primary clusters with $t \geq 1$ m are separated at $N_c > 5$, however, lenses and transitions continue to separate the profile. Therefore, a cluster number 5 is chosen to represent the soil stratigraphy at the site. The lower two primary layers (A2 and A3) are associated in terms of soil type and/or soil properties because they are denoted by cluster numbers 3 and 4. However, the upper main layer (A1) could have different characteristics than these two layers because it is assigned cluster number 1, and a transition zone starts to grow up between A1, and A2. The cluster results using u_t and u_b to calculate B_q are compared as shown in Fig. 5.23. The two primary clusters A1, and A3 defined using u_t data could be verified by the two trends of the liquid limits. The given boundary at almost 7.9 m is indicated by both analyses and verified by the changes of s_u backup data.

At the Brent Cross site, the cluster analysis using both u_t and u_b is able to detect the primary soil boundary between the upper finely fissured and the lower highly fissured clays. The cluster results using u_t is better confirmed by the variation of liquid limits with depth. Therefore, in case of the highly overconsolidated fissured clays at this site, using u_t instead of u_b readings in the cluster analysis is favorable in order to obtain a better estimation of the soil stratigraphy.

5.4.5. Scale Effect

The scale effect using 10-cm² and 15-cm² cones on clustering of piezocone data is evaluated in this section. Both sizes are now common in geotechnical practice (Lunne et al., 1997) and permitted by ASTM guide D-5778. The scale effect of the cone dimensions on q_c , f_s , and/or u_b measurements has been studied previously. Specifically, for the comparison between 10- and 15-cm² cones advanced into the same soil conditions, De Ruiter (1982) and Lima and Tumay (1991) showed no significance difference in q_c . Juran and Tumay (1989) found no significant difference in q_c and u_b but indicated a possible 20 % increase in f_s using 10-cm² base cone compared with 15-cm² base cone. Of course, there are significant differences in the sleeve frictions (f_s) measured by different penetrometers as discussed earlier in section 1.7.2. That variability is the prime reason why f_s (nor normalized form F) has been included herein. Moreover, Hegazy et al. (1996) confirmed that the cone size has a negligible influence on the spatial analysis of piezocone data using geostatistical analysis.

Two representative piezocone soundings conducted at Surry, Virginia, are used in this assessment. A piezocone sounding CP5 was collected using a 10-cm² cone and compared with sounding SP8 that was measured using a 15-cm² cone. Figure 5.24 shows a comparison between both soundings in terms of the measured unprocessed data and the normalized parameters Q and B_q and indicates some apparent difference between the two soundings which could be attributed to scale effect, relative locations of the two soundings in a plan view and inherent soil variability. Looking at the piezocone data alone, three

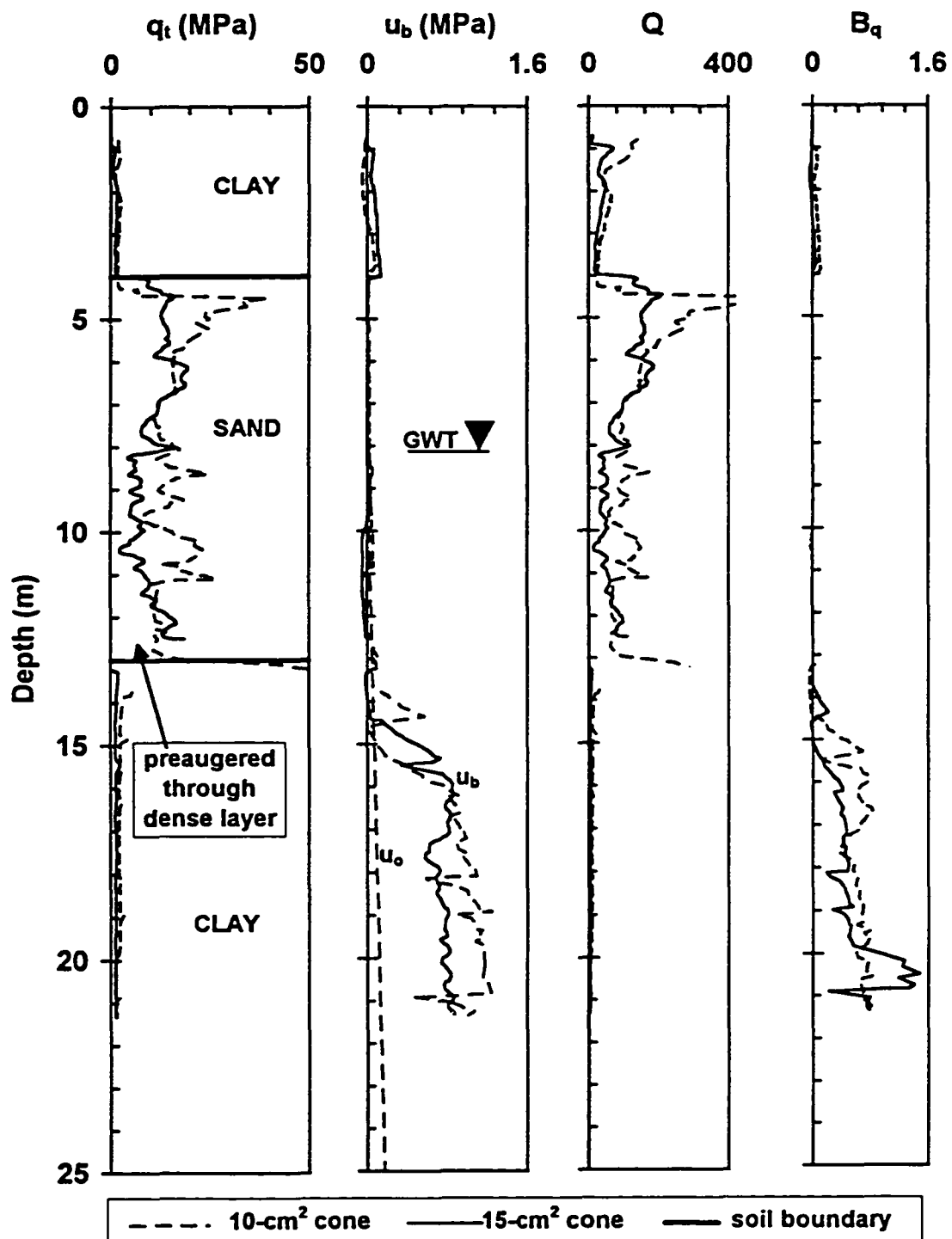
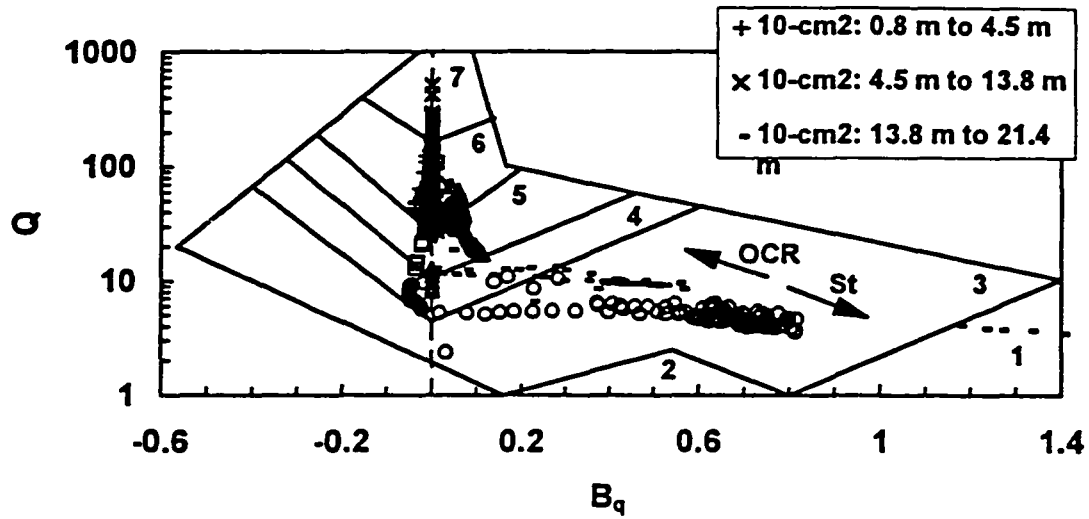


Figure 5.25. Piezocone Data from 10- and 15-cm² Cones at Surry, Virginia (Data from Gordon and Mayne, 1987).

major layers are detected with boundaries at approximate depths of 4.5 m and 13.5 m. Gordon and Mayne (1987) reported the geological formation at the site to consist of recent alluvial deposits underlain by interbedded Atlantic coastal plain sediments of clays, silts, sands, and gravels of Pleistocene age. These layers are underlain by preconsolidated clays of Miocene/Pleistocene age, known locally as the Yorktown Formation. The groundwater table was reported at 8.2-m depth. The soil stratigraphy at the site is determined using the Robertson technique (1991) as shown in Fig. 5.25. Both Q and B_q derived from the two soundings are plotted on the chart. The upper soil between 0.8 and 4.5 m is defined as a soil mixture of clays, silts, and sands. The intermediate layer down to a depth of 13.8 m lies in the sands and silty sands zone and the lower layer between 13.8 m and 21.4 m is classified as clay to silty clay. Therefore, the method is able to give an indication of the different layers in the stratigraphy.

A statistical analysis is performed using SCZ cluster-type analysis with normalized parameters Q and B_q up to cluster number $N_c = 100$. Then, clustering is examined at the peaks of the correlation coefficient of the consecutive clusters. Figure 5.26 shows the cluster results of the CP5 sounding between cluster numbers $N_c = 2$ to 8. At cluster number 2, the data are divided into two groups at a boundary of 15.7-m depth. At cluster number 8, the upper group of data is separated into two primary layers with a boundary at a depth of 4.8 m. A transition zone appears between the depths of 13.8 m and 16 m. For $N_c > 8$, data points continue to separate indicating dissimilarity within the three major groups. Up to $N_c = 100$, as explained in Appendix C in more detail, no new primary



1. Sensitive Fine Grained	5. Sand Mixtures: Silty Sand to Sandy Silt
2. Organic Soils-Peats	6. Sands: Clean Sand to Silty Sand
3. Clays: Clay to Silty Clay	7. Gravelly Sand to Clayey Sand
4. Silt Mixtures: Clayey Silt to Silty Clay	Robertson Modified Chart (1991)

Figure 5.26. Soil Classification Using Piezocone Data at Surry, Virginia
(Data from Gordon and Mayne, 1987).

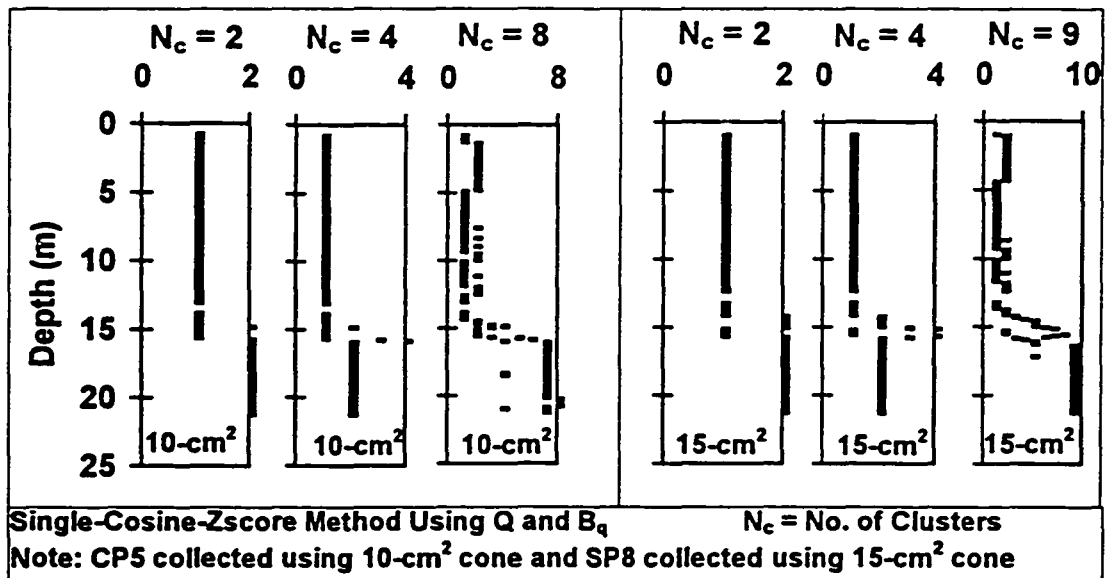
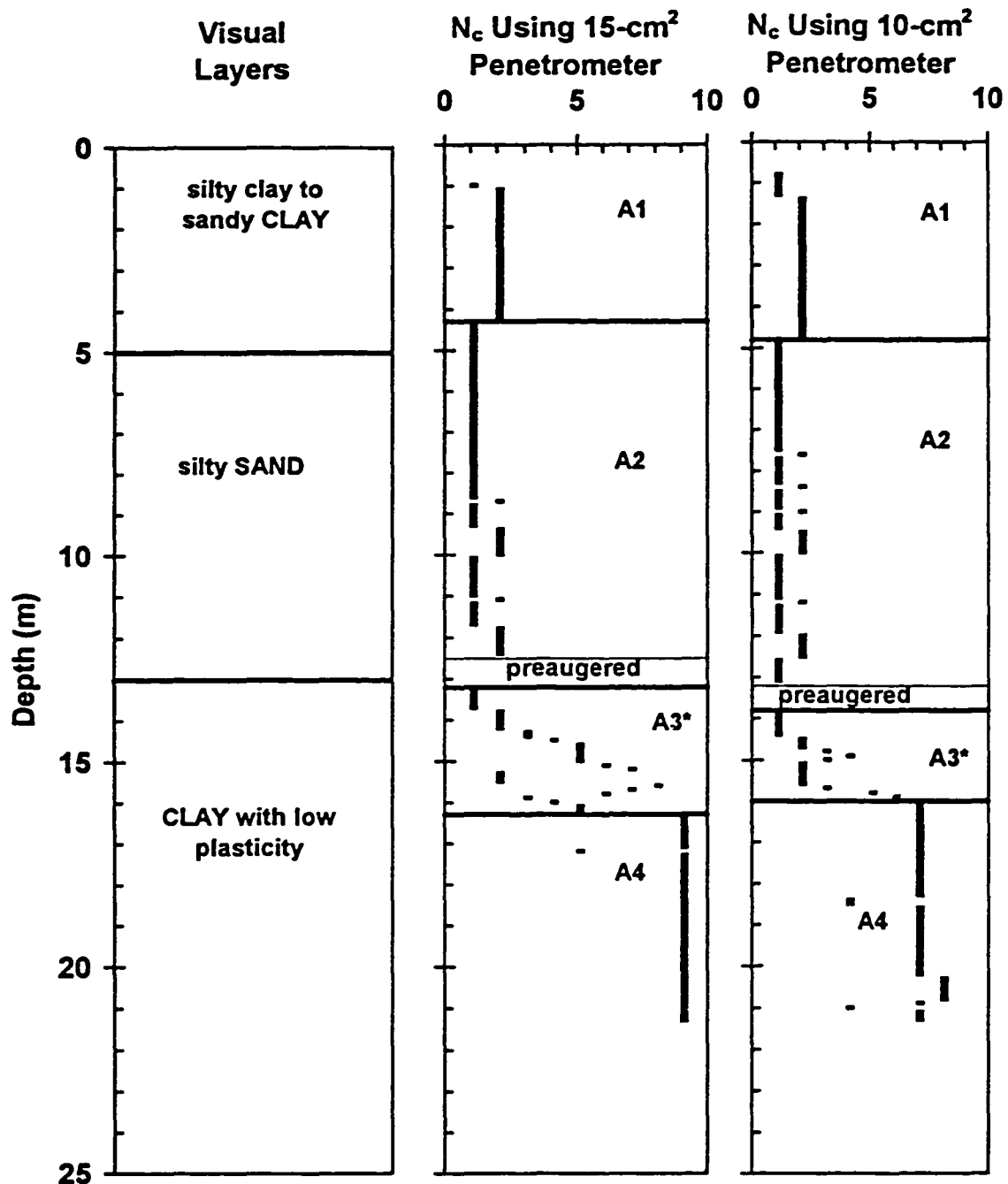


Figure 5.27. Cluster Analysis of Piezocone Data with Different Size
Penetrometers at Surry, Virginia,
(Data from Gordon and Mayne, 1987).

layers ($t \geq 1$ m) are detected. Therefore, a cluster number 8 is selected to delineate the soil stratigraphy.

A similar cluster analysis is performed using Q and B_q of SP8 sounding up to cluster number 100. Clustering is checked at the peaks of the correlation coefficient between successive clusters and Fig. 5.26 shows the results between cluster numbers $N_c = 2$ and 9. A more detailed discussion of the analysis up to $N_c = 100$ is given in Appendix C. At cluster number 2, two main groups of data appear with a boundary at a depth of 14.2 m. In addition, a transition zone is shown between the depths of 14.2 m and 15.7 m. By increasing the cluster number, more points are detached indicating non-homogeneity within the primary layers. At $N_c = 9$, the upper group of data is divided into two statistical layers at a depth of 4.3 m. The transition zone grows up between the depths of 13.2 m and 16.3 m. For higher clusters, no new layers ($t \geq 1$ m) are discovered, therefore, cluster number 9 is chosen to represent the subsurface stratigraphy.

A comparison was performed between cluster results using both 10-cm² and 15-cm² cones as shown in Fig. 5.27 which also includes a representative borehole profile at the site. There is a 0.5-m difference between the elevations of the statistical and borehole boundaries which could be due to inherent soil variability at the relative location of each sounding. Clustering of both cone soundings indicated a transition layer A3* because of gradual change of the pore pressure from negative or zero in the sand layer to positive in the clay layer. Therefore, the scale effect on clustering is concluded to be essentially small, or negligible.



Piezocene data at Surry, Virginia (Gordon and Mayne, 1987)
 Single-Cosine-Zscore Method Using Q and B_q — Primary boundary
 Note: for SP8, cluster no. 9 and for CP5, cluster no. 8

Figure 5.28. Comparison Between Cluster Analysis Using 10-cm² and 15-cm² Cone and Borehole Classification at Surry, Virginia.

5.5. Spatial Cluster Analysis

Clustering can be applied to two- and three-dimensional sets of data. The author collected 15 piezocone data sets at Amherst, Massachusetts, all using a 10-cm² penetrometer. The use of clustering in spatial analysis is studied using three piezocone test results from Amherst. Their relative locations in a plan view are shown in Fig. 5.28. A summary of the unprocessed piezocone data (q_t and u_b) and their derived normalized parameters, (Q and B_q) are shown in Figures 5.29 and 5.30, respectively. It is noticed that the unprocessed q_t measurements are scattered in the upper 4 m due to variability in the clay fill and crust, however, they are more consistent below this level. Due to penetration above the ground water table, the pore pressure readings are either negative or zero in the upper zone. The depth of this zone varies from one sounding to another with an average of about 4 m. In fact, the porous element might have become desaturated in this zone and a delay of the full response of the pore pressure transducer appears to have occurred in piezocone test number 15.

A single-cosine-zscore cluster method is applied to the set of Q and B_q of each sounding. The detailed growth of the data grouping between $N_c = 2$ and 100 is discussed in Appendix C. The interpretation criterion discussed in this chapter is applied to cluster results and cluster numbers $N_c = 8, 10$, and 9 are chosen to represent the subsurface stratigraphy at the locations of piezocone soundings 1, 2, and 15, respectively. Figure 5.31 includes the chosen clusters, the detected primary boundaries and the transitions between the main layers. The cluster number and the depths of the defined soil boundaries

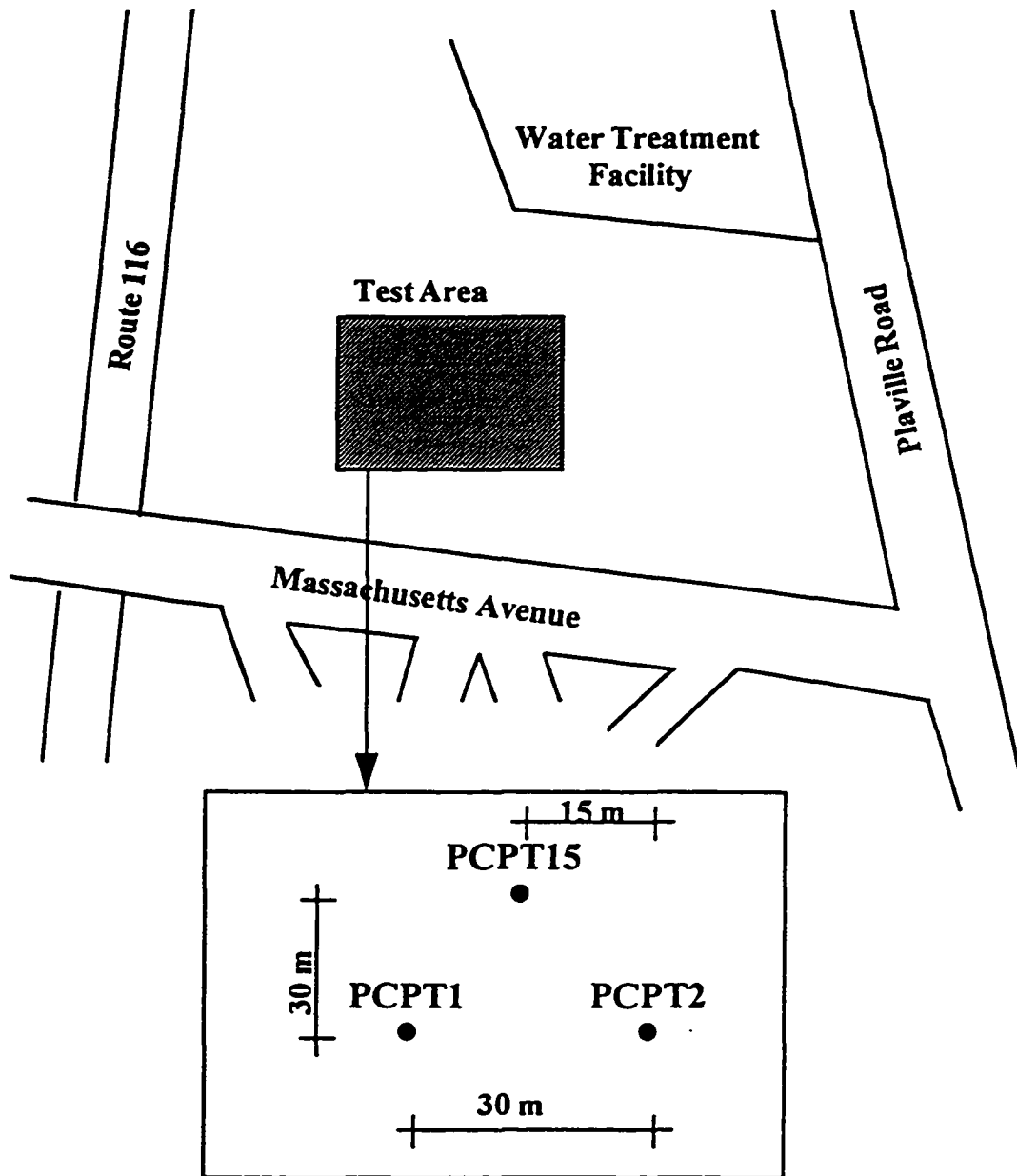


Figure 5.29. Test Location and Site Plan of National Geotechnical Experimental Site (NGES) at Amherst, Massachusetts.

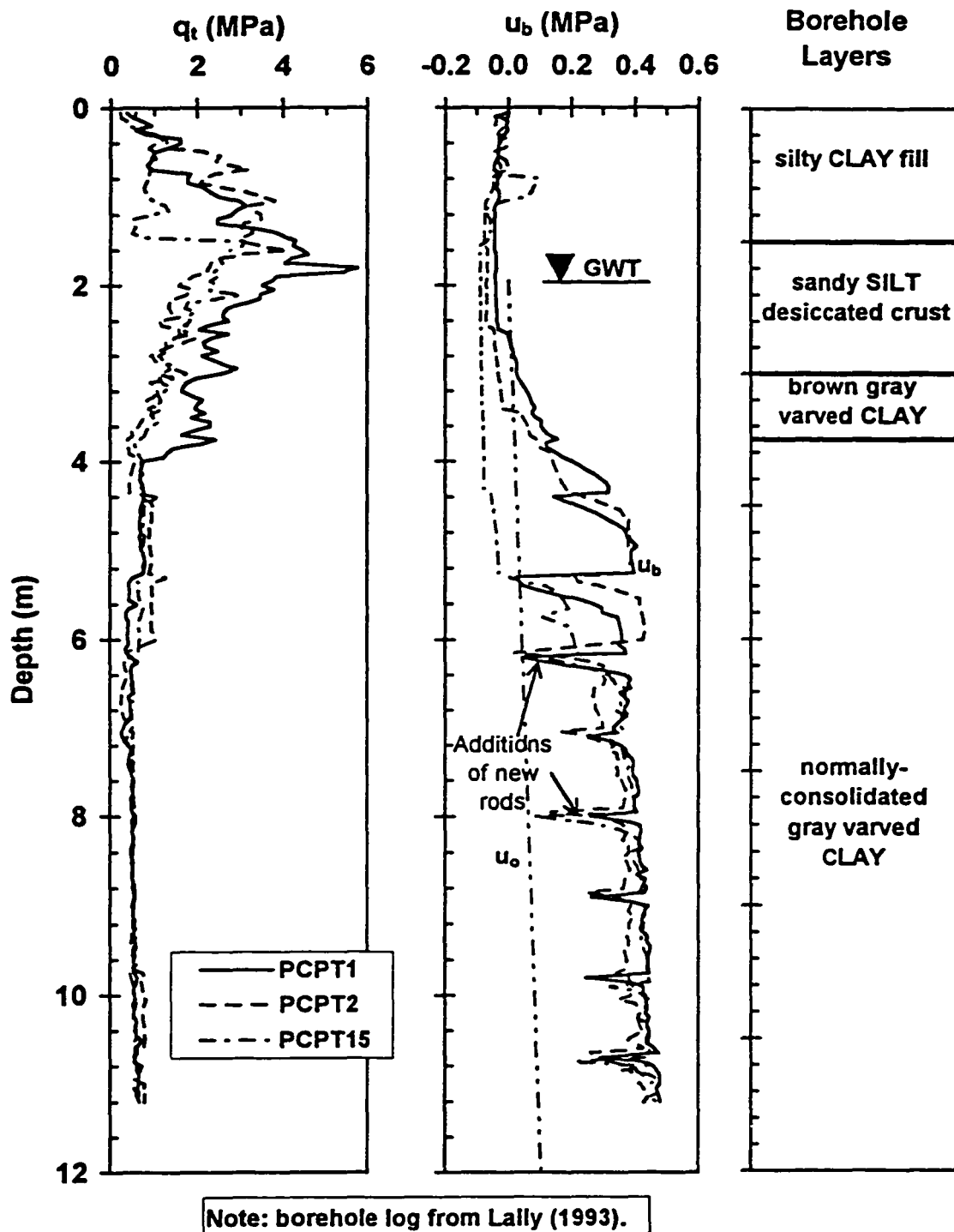
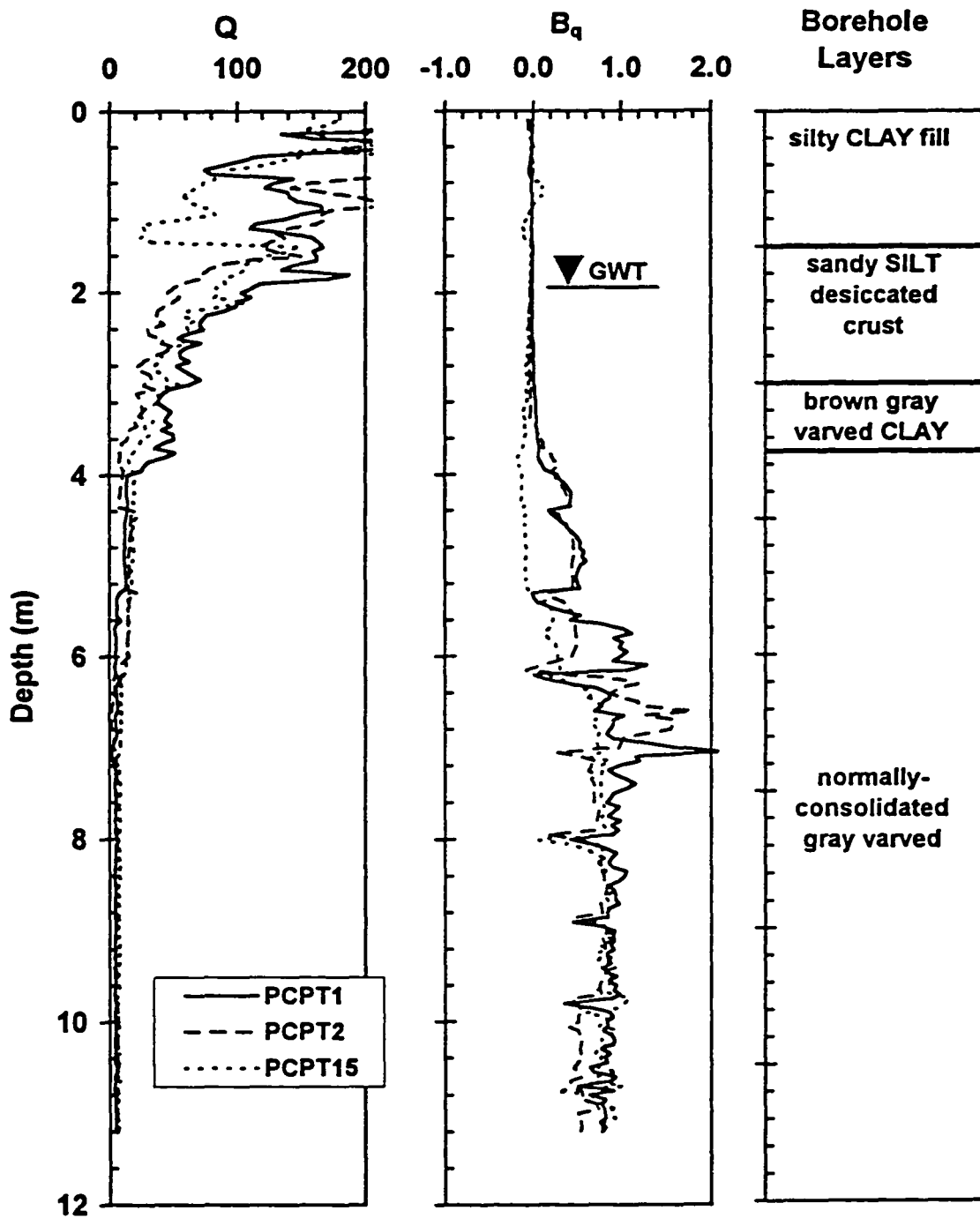
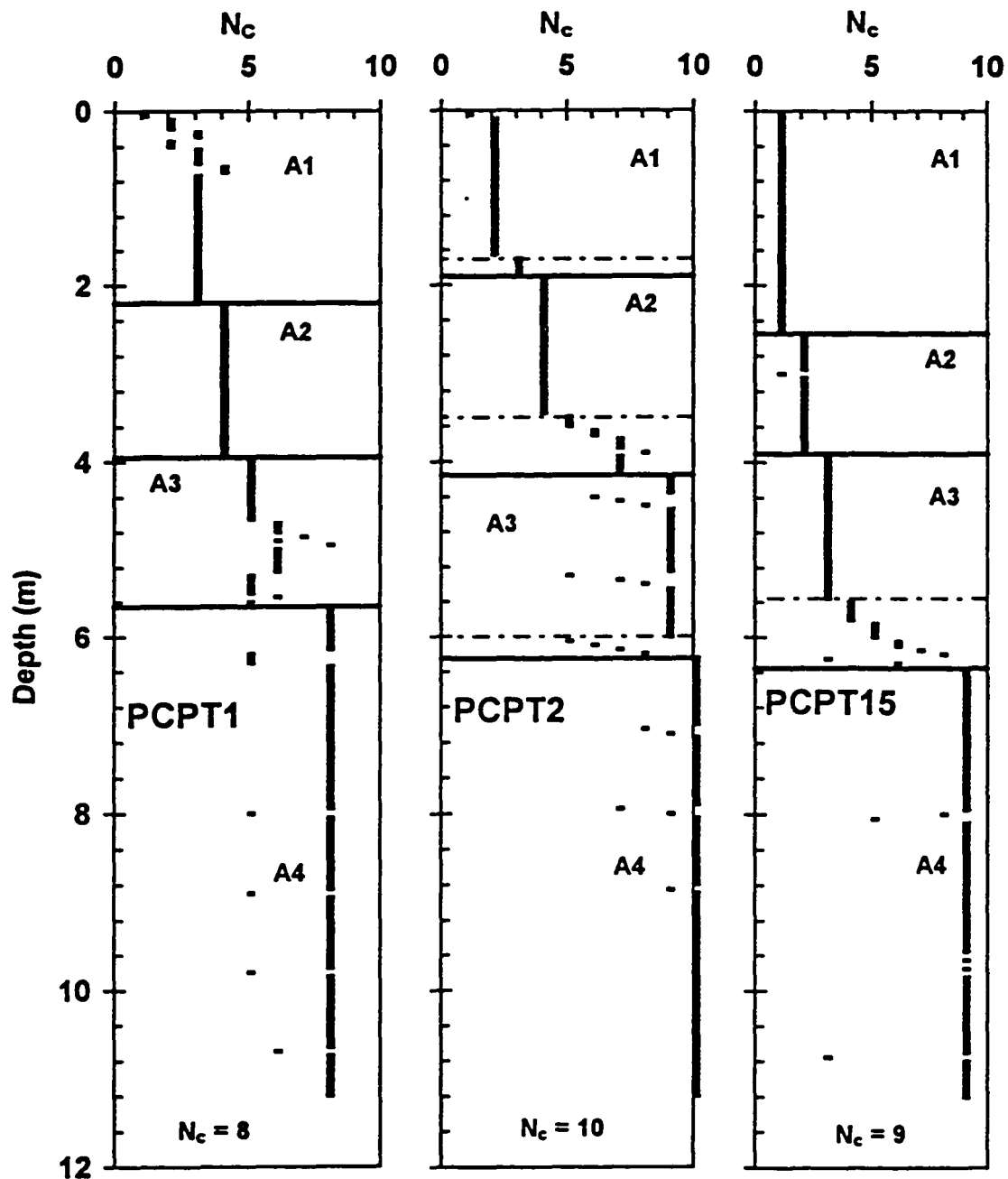


Figure 5.30. Piezocone Data and Soil Stratigraphy at Amherst Massachusetts (Data from This Study).



Note: borehole log from Lally (1993).

Figure 5.31, Normalized Piezocone Data and Soil Stratigraphy at Amherst Massachusetts (Data from This Study).



Note: Piezocone data at Amherst, Massachusetts (this study).
 SCZ method using Q and B_q
 — Primary boundary
 - - - Transition boundary

Figure 5.32. Cluster Results of 3 Piezocone Soundings at Amherst Massachusetts (Data from This Study).

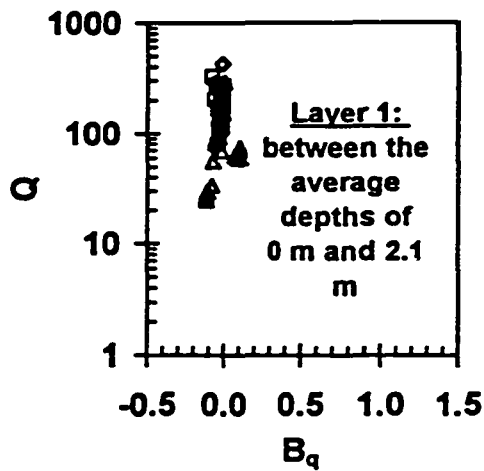
changes at the locations of the three soundings and indicates inherent soil variability and heterogeneity in the horizontal direction. Clustering detects four primary layers at each soundings which are summarized in Table 5.2.

In order to study the similarity of the soil types between the statistical soil boundaries, the piezocone data of the three tests are plotted in a (B_q , Q) space for each statistical layer as shown in Fig. 5.32. For better clarification of the correlation between different defined layers, the normalized data at the transition zones and the data errors

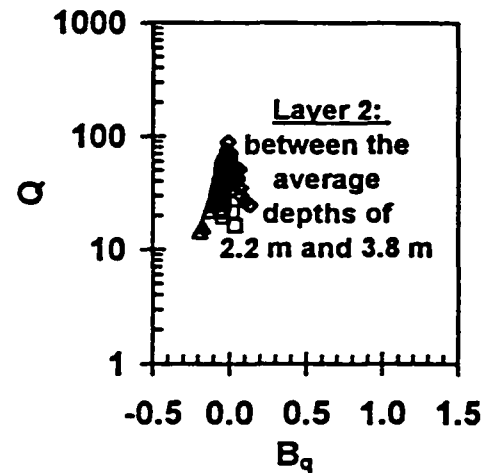
Table 5.2. Primary Soil Boundaries Using Three Piezocone Soundings at Amherst.

Piezocone sounding	Depth in meter		
	Boundary 1	Boundary 2	Boundary 3
PCPT1	2.2	3.9	5.6
PCPT2	1.7	4.1	6.2
PCPT15	2.5	3.9	6.3
Average depth	2.1	4.0	6.1

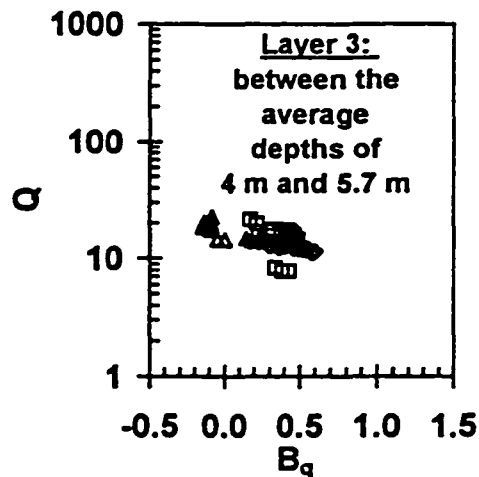
indicated by single cluster points are not included in Fig. 5.32. For the upper fill layer defined between the depths of zero m and 2.5 m, the data appear to have one cluster. In this zone, the normalized pore pressure parameter B_q has a small range and an average equal to -0.02, however the normalized tip resistance Q has a wide range between 27 and 327 and indicates a significant variability within the fill layer. The limits of layer 2 are between the depths of 1.9 m and 3.9 m. The normalized parameters of piezocone soundings 1, 2, and 15 indicate a similar soil type. The average of B_q is equal to -0.04 and the range of Q between 15 and 90 represents the scatter within the clay crust. In the



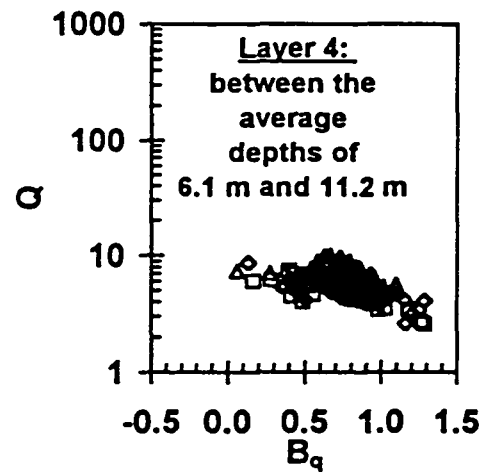
◊ PCPT1 from 0 m to 2.2 m
 ◻ PCPT2 from 0 m to 1.7 m
 Δ PCPT15 from 0 m to 2.55 m



◊ PCPT1 from 2.2 m to 3.95 m
 ◻ PCPT2 from 1.9 m to 3.5 m
 Δ PCPT15 from 2.5 m to 3.9 m



◊ PCPT1 from 3.9 m to 5.6 m
 ◻ PCPT2 from 4.1 m to 6 m
 Δ PCPT15 from 3.9 m to 5.5 m



◊ PCPT1 from 5.6 m to 11.2 m
 ◻ PCPT2 from 6.3 m to 11.2 m
 Δ PCPT15 from 6.4 m to 11.2 m

Figure 5.33. Normalized Data from Three Piezocone Tests at Amherst Massachusetts (This Study).

overlap zone between layers 1 and 2, there could be association between them in terms of soil type and/or behavior. In case of layer number 3 between the depths of 3.9 m and 6.0 m, the normalized parameters of the three soundings are shown to form one group. The normalized tip resistance Q has an average equal to 15 and the normalized pore pressure B_q varies between -0.2 and 0.6. The negative B_q measurements of sounding 15 are due to the saturation delay of the porous element after penetrating the upper partially saturated fill and the clay crust. A soil layer 4 is located between 5.6-m depth and 11.2-m depth and the range of Q is between 2 and 10 and B_q varies between 0.05 and 1.30. The piezocone data indicate a spatial uniformity in this layer and a similarity of soil type and/or behavior at the locations of the three soundings. Therefore, cluster analysis can be applied in more than one dimension to detect associated soil groups within a subsurface stratification.

5.6. Conclusions

The statistical technique of clustering is applied to in-situ piezocone test data for purposes of subsurface stratigraphy and the determination of layer interfaces. Data processing has been adopted using a single-cosine-zscore criterion based on a parametric study and analysis of normalized piezocone data, $Q = [(q_t - \sigma_{vo})/\sigma_{vo}']$ and $B_q = [(u_b - u_o)/(q_t - \sigma_{vo})]$. The validity of clustering to detect soil layer boundaries, lenses, transitions, and number of soil types is studied using piezocone data at several sites. Cluster results appear not to be affected by electrical noise in the piezocone data or by procedural errors due to frequent stops to add successive rods. The normalized parameters Q and B_q gave a better indication of the soil stratigraphy than other data combinations such as: (1)

unprocessed readings (q_t and u_b), (2) slightly processed readings (q_t and the ratio u_b/q_t), and (3) normalized parameters (Q , B_q and F). Note that the sleeve friction parameter F is defined as follows: $[f_s/(q_t - \sigma_{vo})]$.

Using the actual unit weight in the derivation of the normalized parameters Q and B_q helped to reduce the cluster number delineating a soil stratigraphy. Clustering can be applied to other laboratory and in-situ measurements collected at a lower frequency than that of the piezocone data. Similar primary layers were defined using data frequency varied between 5 cm and 50 cm. Alike clustering results were obtained in a soft clay deposit using positive pore pressure readings measured at the face of the cone (u_t) or behind the tip (u_b). However, in the case of overconsolidated clays, using positive u_t measurements resulted in a more proper soil profile than using negative u_b readings affected by soil dilation, and fissures. The scale effect using 10-cm² and 15-cm² cones has a negligible influence on cluster results. Clustering can be applied in more than one dimension to indicate the spatial correlation of different soil strata and continuity of similar soil layers.

CHAPTER 6

DEMARCATING SUBTLETIES IN PIEZOCONES PROFILES BY CLUSTERING

6.1. Synopsis

In this chapter, cluster analysis is used to analyze piezocone data in soil formations where the in-situ tests show no readily apparent stratigraphic differences but in fact there are drastic changes in the vertical profiles of index and mechanical properties of the soil. Also, using the simple visual method, it would be difficult or impossible to decide the demarcations in these types of geological formations. In some soil stratigraphies, empirical piezocone classification methods fail to distinguish the subtle changes in piezocone data and consequently the boundaries between vastly different soil types.

Piezocone data are analyzed using single-cosine-zscore (SCZ) clustering type at eight sites in which fairly substantial changes in the soil stratigraphies would be missed and undetected by routine available methods. In these cases, clustering strongly implied that a significant layering pattern occurred in the profile, although no visible evidence was apparent. The sites had undergone extensive geotechnical studies and include eight clay deposits at which clustering was verified by available back-up data of physical and mechanical soil properties. A summary of the investigated sites is given in Table 6.1.

Table 6.1. Summary of 8 Clay Sites Analyzed Using SCZ-Type Clustering of Piezocone Data with Subtle Changes.

Site Name	Location	Soil Type	Reference
Drammen	Norway	Plastic clay over lean clay	Lacasse and Lunne (1982); Masood et al. (1990)
Gloucester	Ontario	Soft silty clay over clay to silty clay	Konrad and Law (1987)
Hachirogata	Japan	Soft marine clay	Tanaka et al. (1992)
Lilla Mellösa	Sweden	Organic clay over clay underlain by a varved clay	Larsson and Mulabdić (1991)
Recife	Brazil	Organic soft clay over another soft clay	Coutinho and Oliveira (1997)
St. Alban	Quebec	Soft very sensitive silty clay changes to clayey silt	Roy et al. (1982)
Tiller	Norway	Silty clay over quick clay	Sandven (1990)
Troll	North Sea	Very soft clay over very stiff silty clay	Amundsen et al. (1985)

6.2. Clustering Applications

Representative piezocone soundings at the eight sites were visually examined, however, it was difficult to define the breaks between unlike soil layers. Moreover, the traditional Robertson classification chart (1991) was not satisfactory for the delineation of the differences between different clay units using the derived normalized parameters $Q = (q_t - \sigma_{vo})/\sigma_{vo}'$ and $B_q = (u_2 - u_o)/(q_t - \sigma_{vo})$. Note that this is one of two charts proposed by Robertson (1991) for piezocone data interpretation. The second chart was developed based on the normalized parameters Q and $F = f_s/(q_t - \sigma_{vo})$, however, it has not been used in this study because of the unreliability associated with f_s readings as explained earlier in section 2.7.2. Clustering using a single-cosine-zscore (SCZ) method was applied to

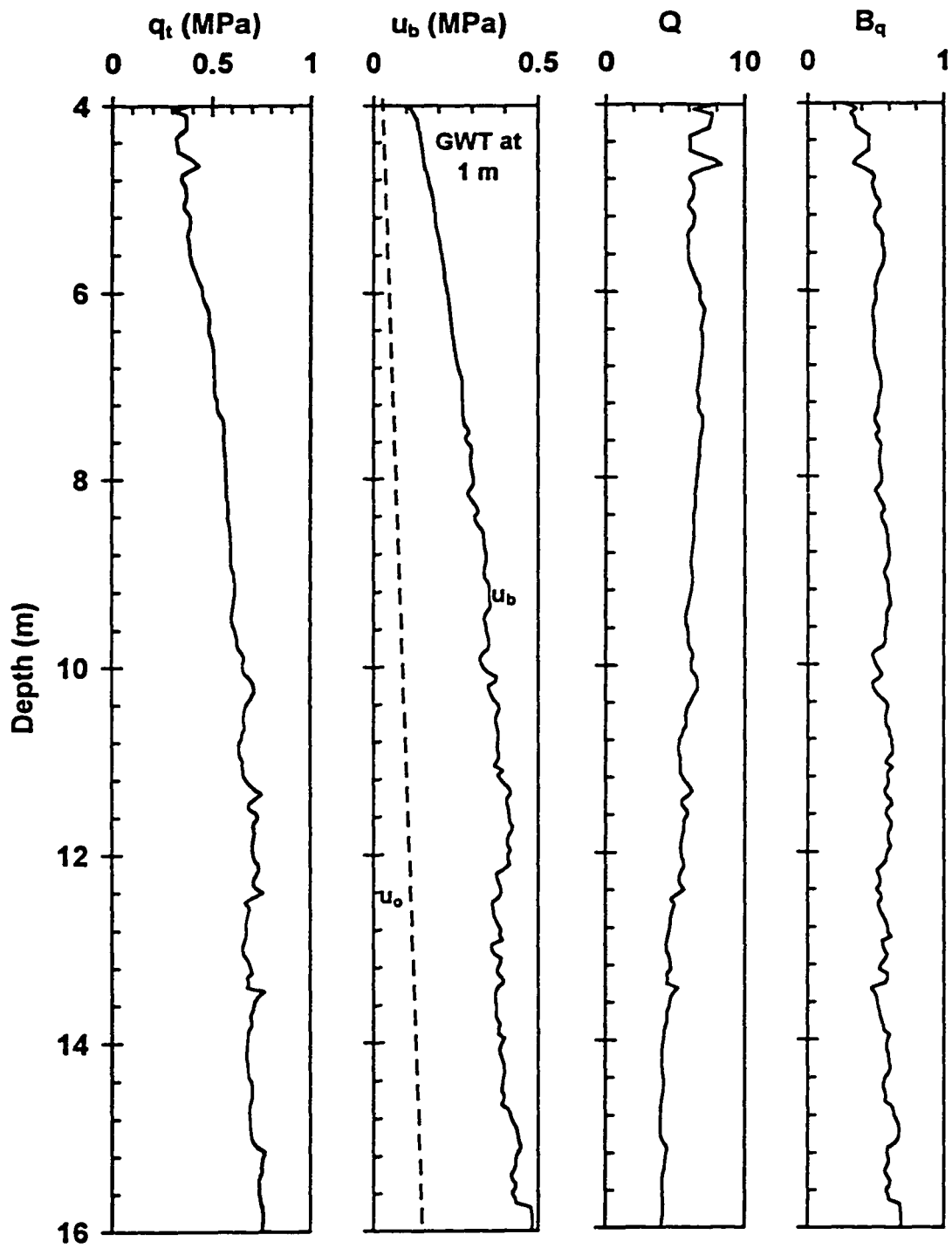
piezocone data having these subtle changes and the results were able to demarcate the dramatic changes in soil types and/or properties as discussed in the following sections.

6.2.1. Drammen Test Site

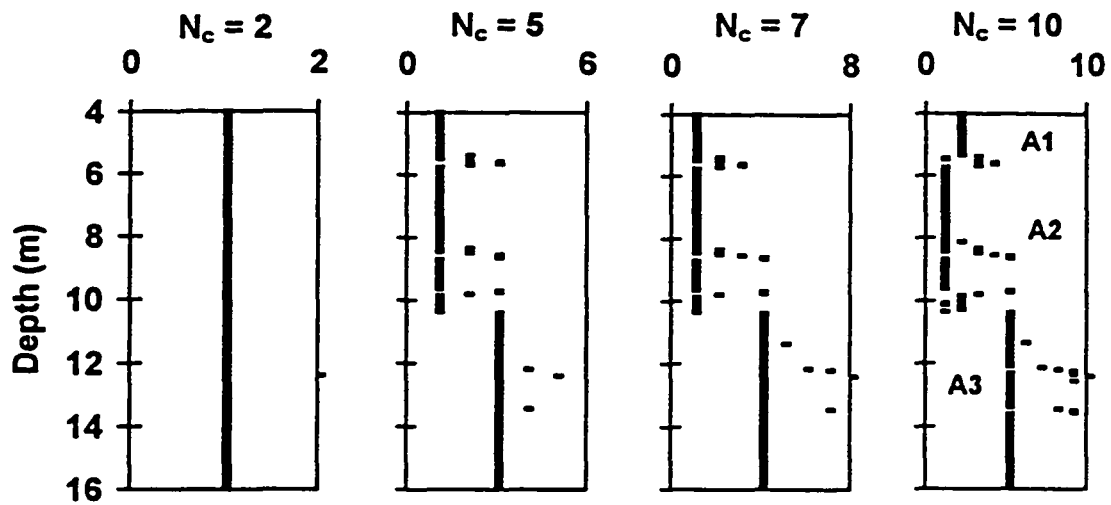
A representative piezocone sounding from the Drammen test site in Norway (Masood et al., 1990) is shown in Fig. 6.1. Visual examination of the data indicates no obvious layering in the vertical profile. The site has been tested extensively by laboratory and in-situ means (Lacasse and Lunne, 1982). The soil stratigraphy in the zone of interest lies between the depths of 4 m and 16 m and consists of the following: a silty clay layer from depths of 4 m to 5 m, a plastic clay layer from 5 m to 10 m, underlain by a lean clay layer from 10 m to 16 m. The groundwater table is determined to be at a depth of 1 m. The derived normalized parameters Q and B_q are also shown in Fig. 6.1.

A single-cosine-zscore cluster method is applied to the piezocone data using Q and B_q and the cluster results are examined from $N_c = 2$ to 100, as discussed in detail in Appendix C. The growth of data clustering at the peaks of ρ_c is shown on Fig. 6.2 between $N_c = 2$ to 10. At $N_c = 2$, the soil is grouped in one cluster except for a lense detected at 12.4 m. Subsequently two groups are separated with a boundary at 10.4 m until $N_c = 10$ where a primary layer separates from the upper group. For larger clusters up to $N_c = 100$, only transitions and lenses breaks off the three primary clusters. Therefore, a cluster number 10 is chosen to represent the soil stratigraphy at the site.

The data representing the primary groups identified by clustering are plotted in a Q - B_q space as shown in Fig. 6.3 which also includes boundaries of Robertson Chart. The

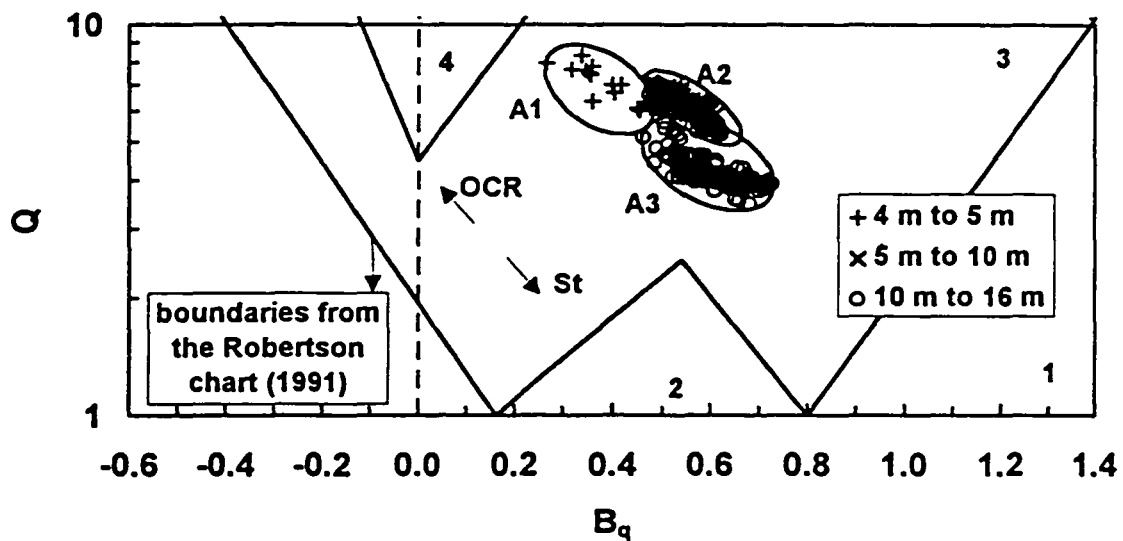


**Figure 6.1. A Representative Piezocone Sounding from Drammen, Norway
(Data from Masood et al., 1990).**



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Drammen, Norway (Masood et al., 1990).

Figure 6.2. Cluster Analysis of Piezocone Data at Drammen, Norway.



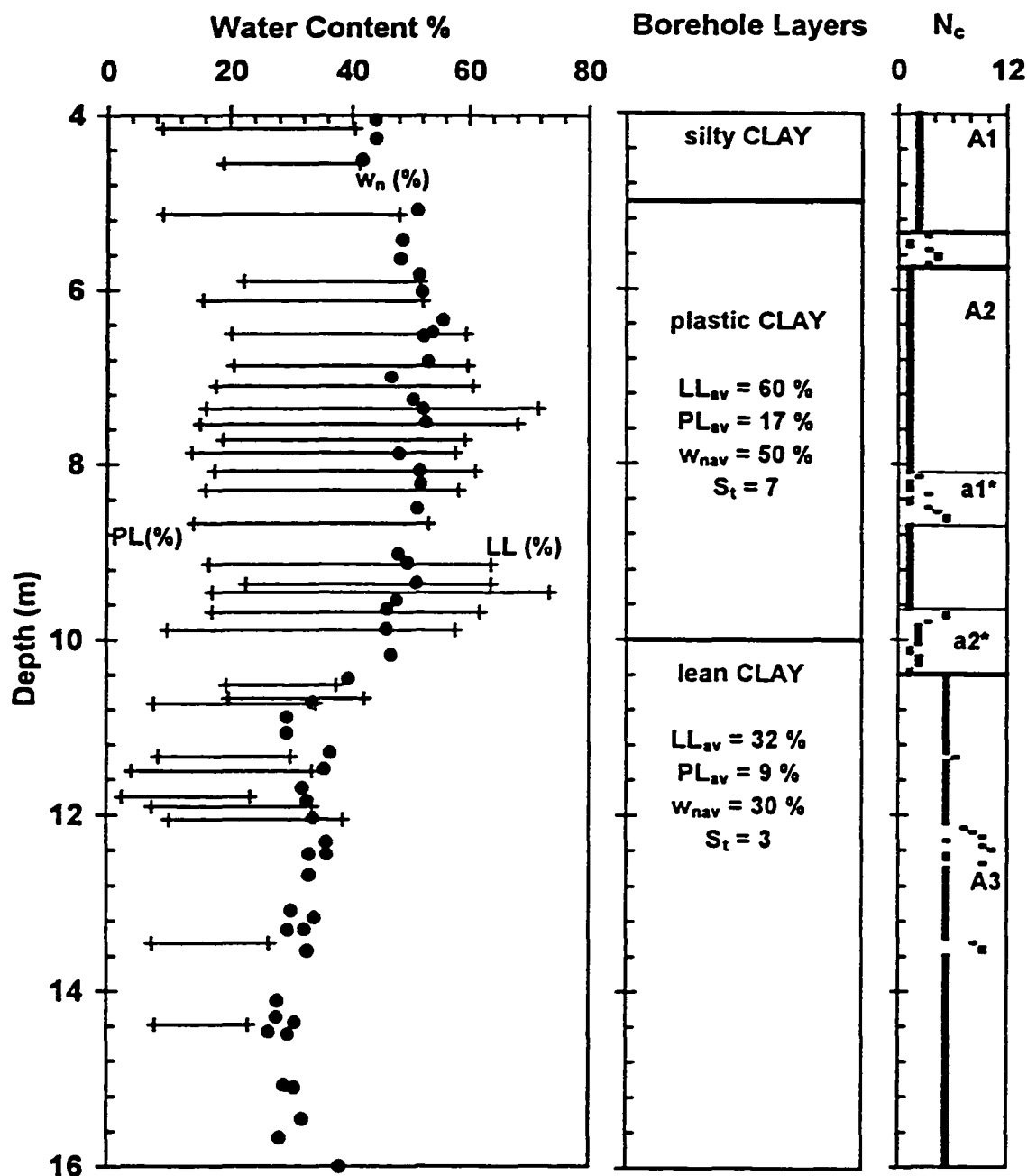
- | | |
|---------------------------|---|
| 1. Sensitive Fine Grained | 3. Clays: Clay to Silty Clay |
| 2. Organic Soils-Peats | 4. Silt Mixtures: Clayey Silt to Silty Clay |

Figure 6.3. Clustering Results of Piezocone Data at Drammen, Norway on a Q - B_q Space (Data from Masood et al., 1990).

soil is preliminary classified as a one clay layer and the cluster results indicated that there are three primary sublayers which have different properties, within the soil profile. This information is useful for the engineer because it helps him to identify the number of soil samples and their depths in a preliminary site investigation program. For example, in this case study, three samples can be extracted at approximately the middle of each primary layer to examine their physical and mechanical properties.

The stratigraphy obtained by clustering is verified by water content and plasticity measurements as shown in Fig. 6.4. Note that the undrained shear strength is correlated with the plasticity index and liquidity index which is a function of the water content (Bjerrum, 1973). The cone parameters are functions of the undrained shear strength (Wroth, 1984; 1988). The index data are divided visually into three groups with approximate boundaries at 6 m and 10.5 m. The water contents increase with depth from 44 percent at 4 m to 55 percent at 6 m and then decrease to 33 percent at 10.7 m. They have an average equal to 32 percent below 10.7 m and down to 16 m. In the latter zone, the water contents have a range between 26 percent and 38 percent.

The undrained shear strength (s_u) measured by vane shear tests also validate a primary soil boundary at almost 10.5 m between the plastic and the lean clays. Figure 6.5 shows two trends of s_u measurements above and below this boundary, although the data do not confirm cluster A1 between 4.0 m and 5.3 m. Also, the statistical soil profile is verified by the variation of soil sensitivity which, on the average, decreases from 7 to 3 in association with clusters A2 and A3, respectively. Therefore at this site, the cluster analysis is able to discover the differences between silty, plastic, and lean clays at



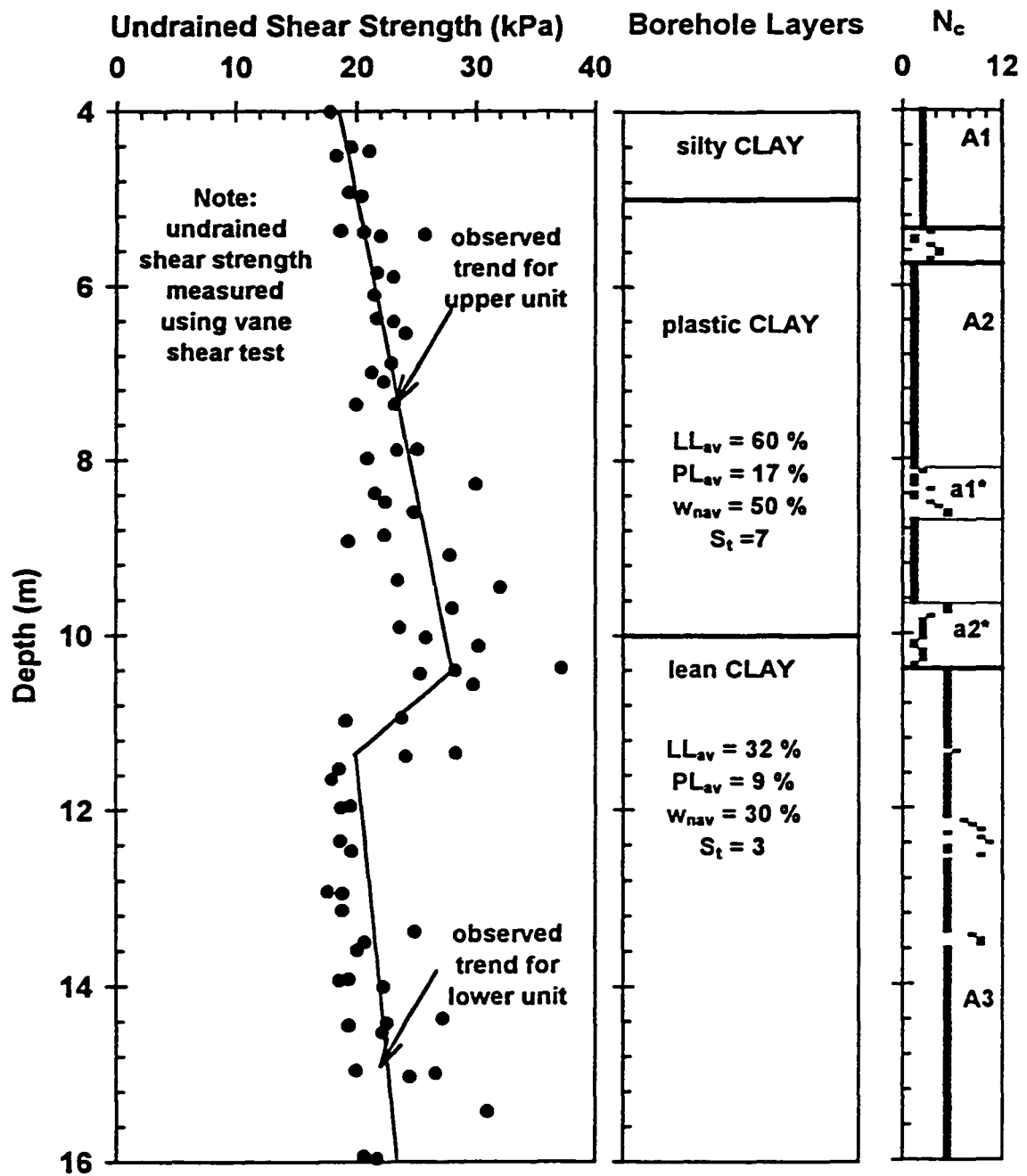
Note: Piezocone data at Drammen, Norway (Masood et al., 1990).

SCZ method using Q and B_q, cluster no. 10

— Primary boundary

— Secondary boundary

Figure 6.4. Comparison Between Cluster Analysis, Visual Classification and Water Content at Drammen, Norway.



Note: In-situ data at Drammen, Norway (Lacasse and Lunne, 1982).
 SCZ method using Q and B_q , cluster no. 10

— Primary boundary
 — Secondary boundary

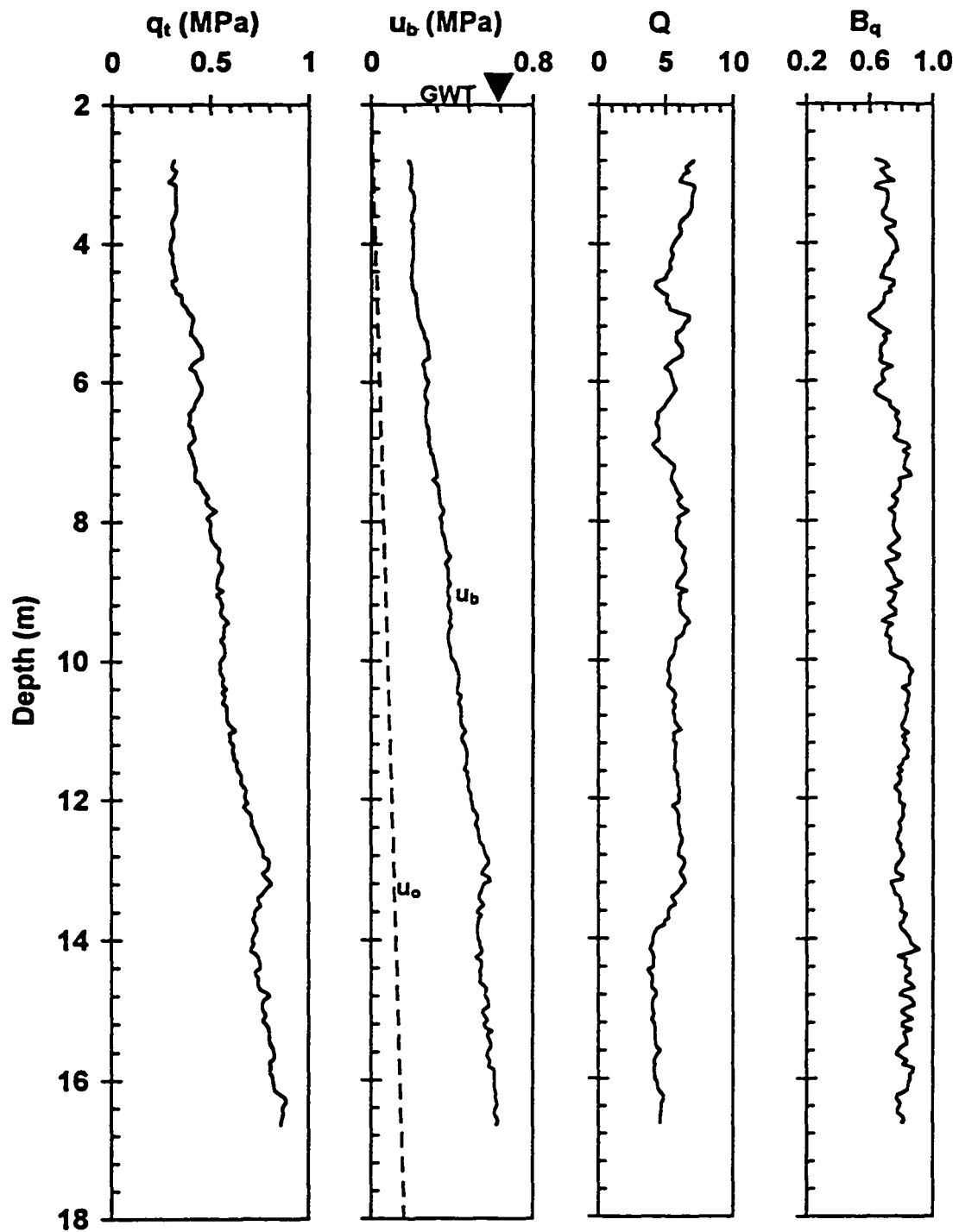
Figure 6.5. Comparison Between Cluster Analysis, Soil Profile and Undrained Shear Strength at Drammen, Norway.

Drammen test site. It also can be supplement the classification chart which can identify the soil type, to objectively indicate the major changes in the soil properties of a clay deposit.

6.2.2. Gloucester Test Site

Results of a representative piezocone sounding at the Gloucester site in Ontario, Canada is shown in Fig. 6.6 (Konrad and Law, 1987). The site has been extensively studied by Bozozuk and Leonards (1972). The soil at the site consists of soft marine clay deposits of the Champlain Sea. The stratigraphy at the site is the result of different stages of deposition and erosion. Below 2.8 m where the piezocone measurements start, the soil stratification consists of three layers as follows: a soft silty clay layer down to 7.0 m overlying a clay layer down to 13.8 m, which is underlain by a silty clay layer until the termination depths of exploration at 16.6 meters. Occasional small stones appear in the middle and lower layers. The groundwater table is at 2-m depth.

Results of the derived normalized parameters Q and B_q are also shown in Fig. 6.6. Looking at the unprocessed q_t and u_b readings, one interpretation by the author is the existence of a soil boundary visually defined at a depth of 13 m. By inspecting the Q profile, layer boundaries are detected approximately at 4.6 m, 7 m, 10.5 m and 14 m. Visual examination of the B_q profile indicates the probable occurrence of soil boundaries at 6 m and 10 m. Therefore, different answers are obtained by the visual examination of piezocone vertical profiles.



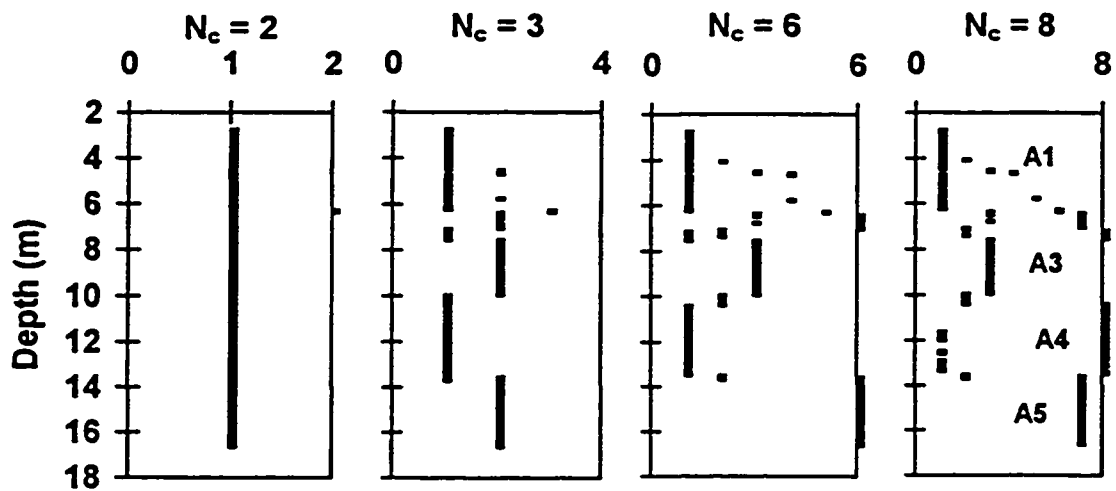
**Figure 6.6. A Representative Piezocone Data at Gloucester, Ontario
(Data from Konrad and Law, 1987).**

A single-cosine-zscore cluster analysis is applied to the normalized parameters Q and B_q . The growth of data groups is studied between cluster number $N_c = 2$ and 100 as discussed in Appendix C. The cluster results at the peaks of ρ_c are shown in Fig. 6.7 between cluster number $N_c = 2$ and 8. At $N_c = 2$, the stratigraphy is shown as one primary layer except a soil lense is detected between 6.3 m and 6.35 m. At $N_c = 3$, the data are separated into 4 groups with alternative cluster numbers 1 and 2. Then some points detached from cluster number 3 and the chain between the second and the fourth groups (sorted from the top to the bottom) breaks. The second group is assigned a cluster number 3 and the fourth group is assigned a cluster number 6 (see Fig. 6.7, $N_c = 6$).

At higher clusters, a similar break happens in the chain between groups 1 and 3. The third group is assigned a cluster number 8 (see Fig. 6.7, $N_c = 8$). Subsequently, no primary layers ($t > 1$ m) split at $N_c > 8$ which is chosen to represent the soil stratigraphy at the site. The first and second data groups, which are assigned cluster numbers 1 and 3, are relatively correlated in terms of soil type and/or properties. The same applies to the third and fourth data groups which are assigned cluster numbers 7 and 8.

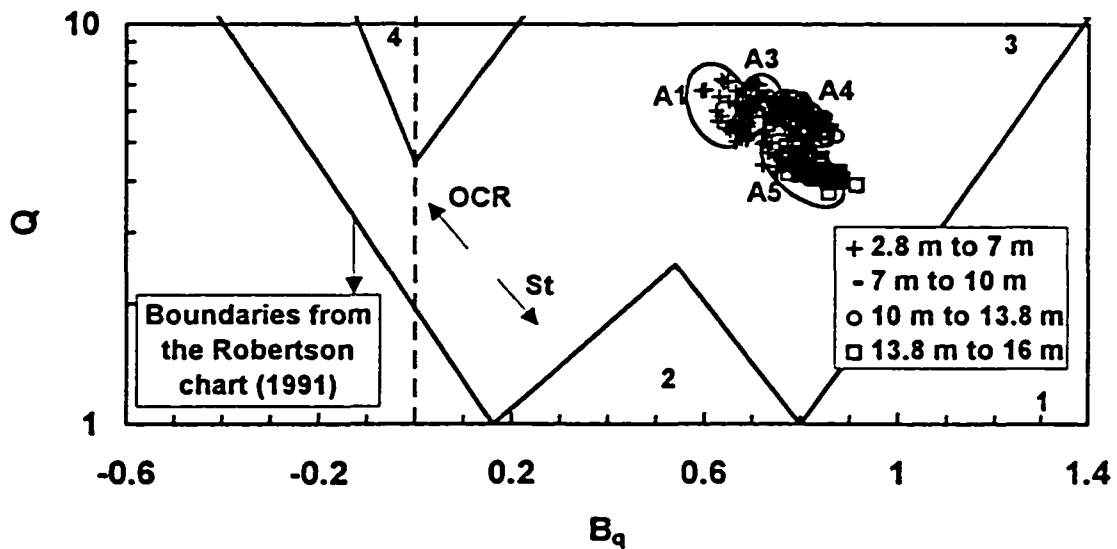
The soil is classified as one clay layer using the Robertson (1991) chart based on Q and B_q as shown in Fig. 6.8 which also includes the four primary clusters. Clustering indicates that there is a variation in the properties of the four sublayers within the clay deposit which helps to properly plan a site investigation.

The cluster results are independently confirmed by water content (w_n) readings as shown on Fig. 6.9 and the data are divided into four groups with boundaries at depths of 7.0 m, 10.4 m, and 14.0 m. The upper two groups from 2.8 m to 7.0 m, and from 7.0 m



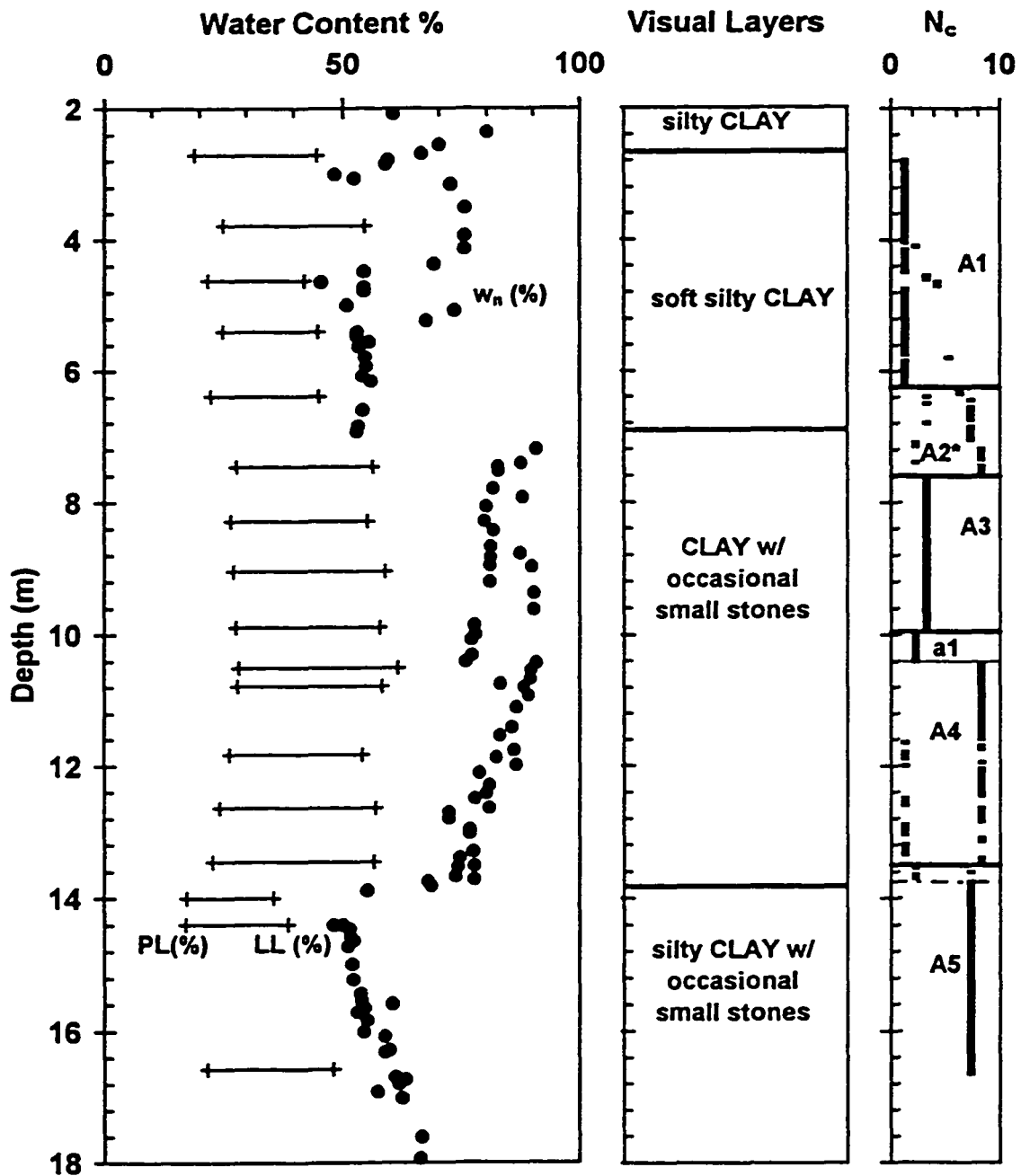
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Gloucester, Canada (Konrad and Law, 1987).

Figure 6.7. Cluster Analysis of Piezocone Data at Gloucester, Ontario.



1. Sensitive Fine Grained	3. Clays: Clay to Silty Clay
2. Organic Soils-Peats	4. Silt Mixtures: Clayey Silt to Silty Clay

Figure 6.8. Clustering Results of Piezocone Data at Gloucester, Ontario on a Q - B_q Space (Data from Konrad and Law, 1987).



Note: Piezocone data at Gloucester, Ontario from Konrad and Law (1987).
 SCZ method using Q and Bq, cluster no. 8

— Primary boundary
 - - - Transition boundary
 — Secondary boundary

Figure 6.9. Comparison Between Cluster Analysis, Visual Classification and Water Content at Gloucester, Ontario.

to 10.4 m have average water contents equal to 59 percent and 83 percent, respectively. In the third group, the w_a decreases from 91 percent at 10.4 m to 59 percent at 13.8 m. However in the fourth group, the w_a increases from 48 percent at 14.4 m to 61 percent at 16.7 m.

The preconsolidation stresses (σ_p') from laboratory oedometer tests also support the cluster results as shown on Fig. 6.10. The normalized cone parameters are functions of the overconsolidation ratio of the soil (Wroth, 1984; 1988; Chen and Mayne, 1994). These yield stresses indicate different stages of deposition, erosion and redeposition cycles as noted by Bozozuk and Leonards (1972). The data are divided into five groups each has a trend which better fits the data than the three trends suggested by Bozozuk and Leonards (1972). Each data group supports a corresponding cluster as follows: groups 1, 2, 3, 4 and 5 verify clusters A1, A2*, A3, A4 and A5, respectively.

The undrained shear strength (s_u) measured using vane shear tests also verifies the primary layers delineated by the clustering. Bozozuk and Leonards (1972) proposed three positive trends of s_u measurements as shown on Fig. 6.11. The upper trend between 2 m and 6.8 m supports cluster A1. The lower trend between 14 m and 18 m supports cluster A5. However, the proposed middle trend does not represent the s_u measurements between depths of 6.8 m and 14 m. The data are grouped above the trend between 8 and 10 m and below the trend below 8 m and down to 13 m. Therefore, these two data groups are better fitted with linear trends shown in Fig. 6.11 and support clusters A3 and A4. The cluster results can indicate major variations of physical and mechanical properties

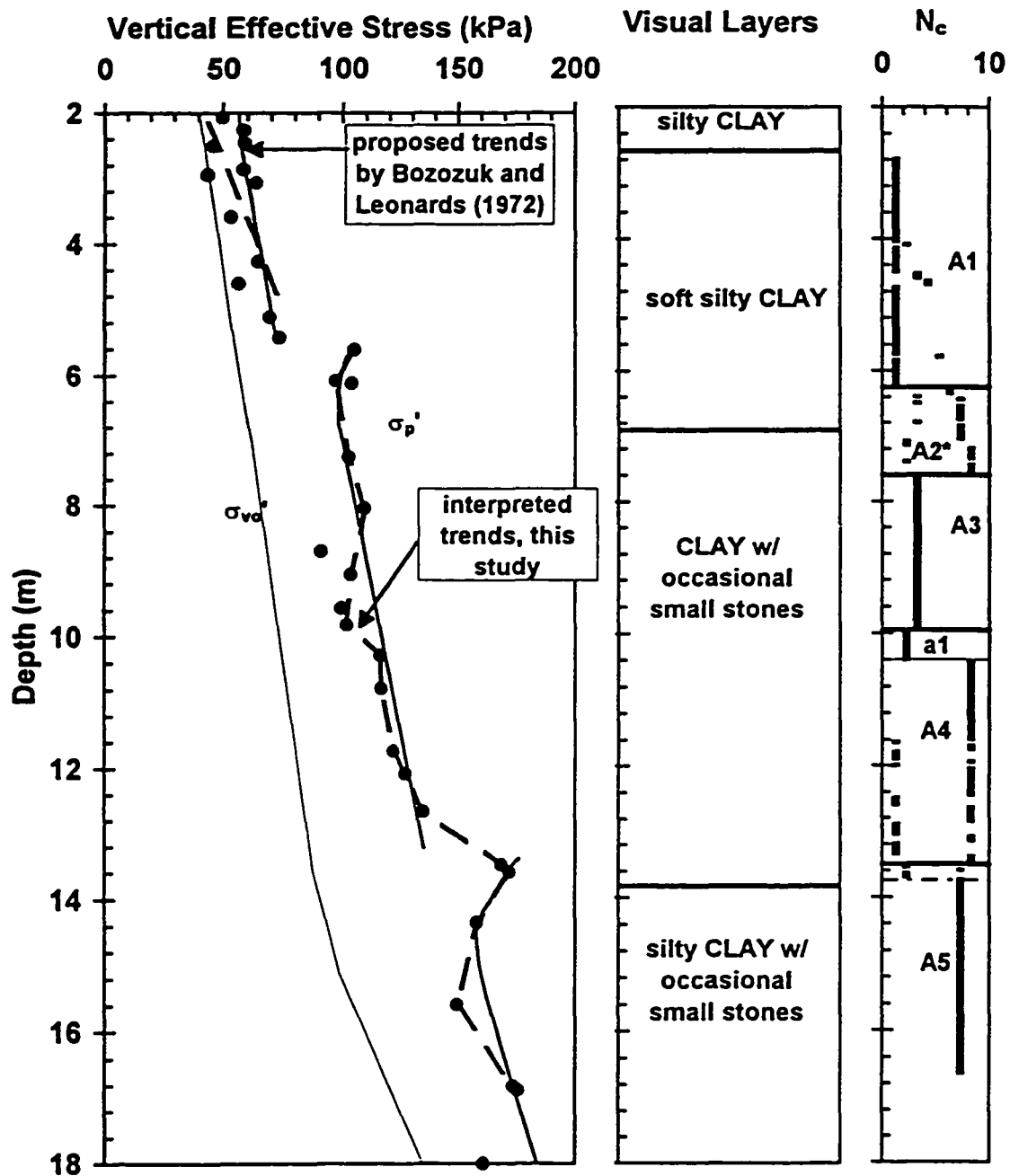


Figure 6.10. Comparison Between Cluster Analysis, Visual Classification and Vertical Effective Stresses at Gloucester, Ontario.

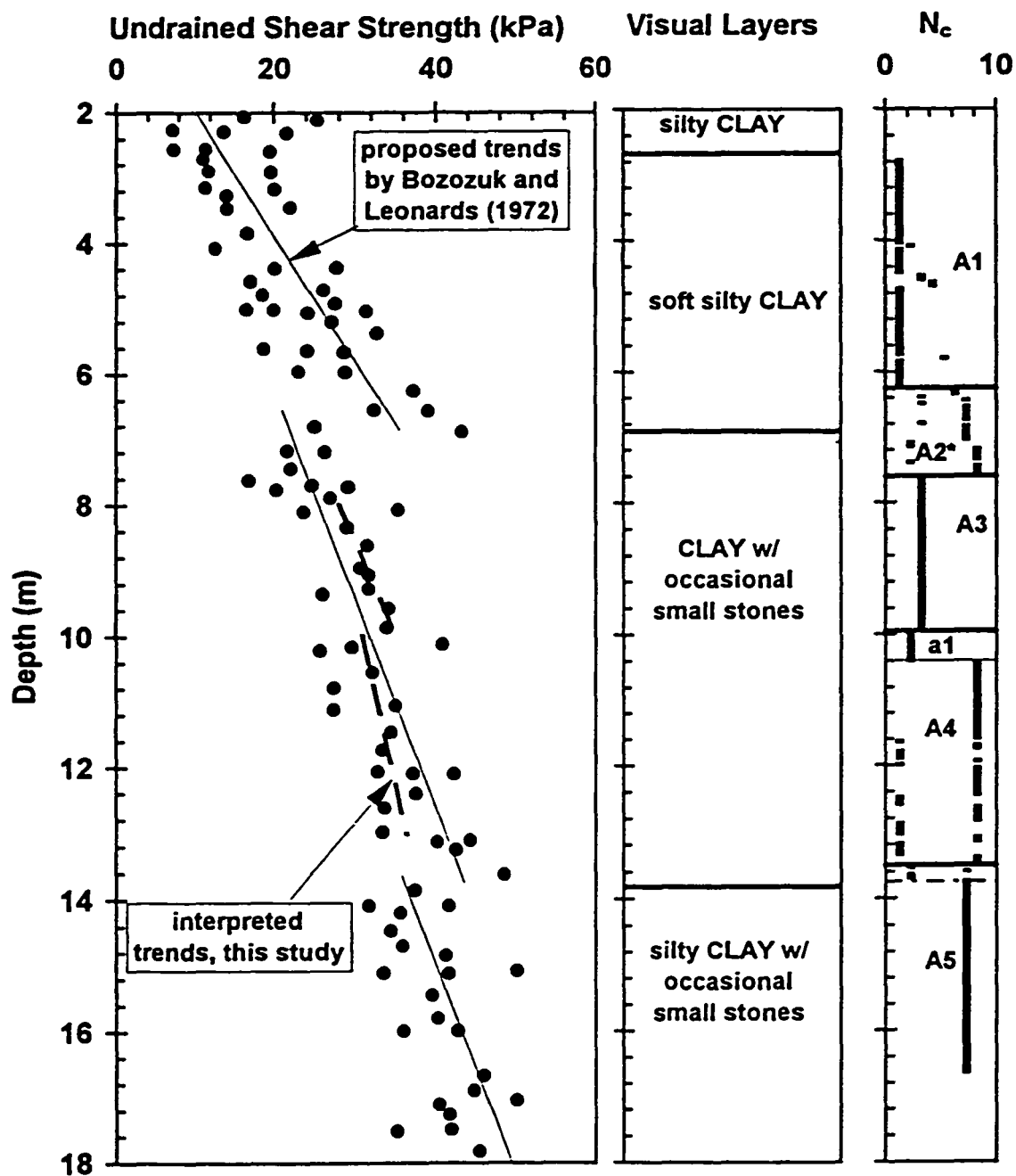


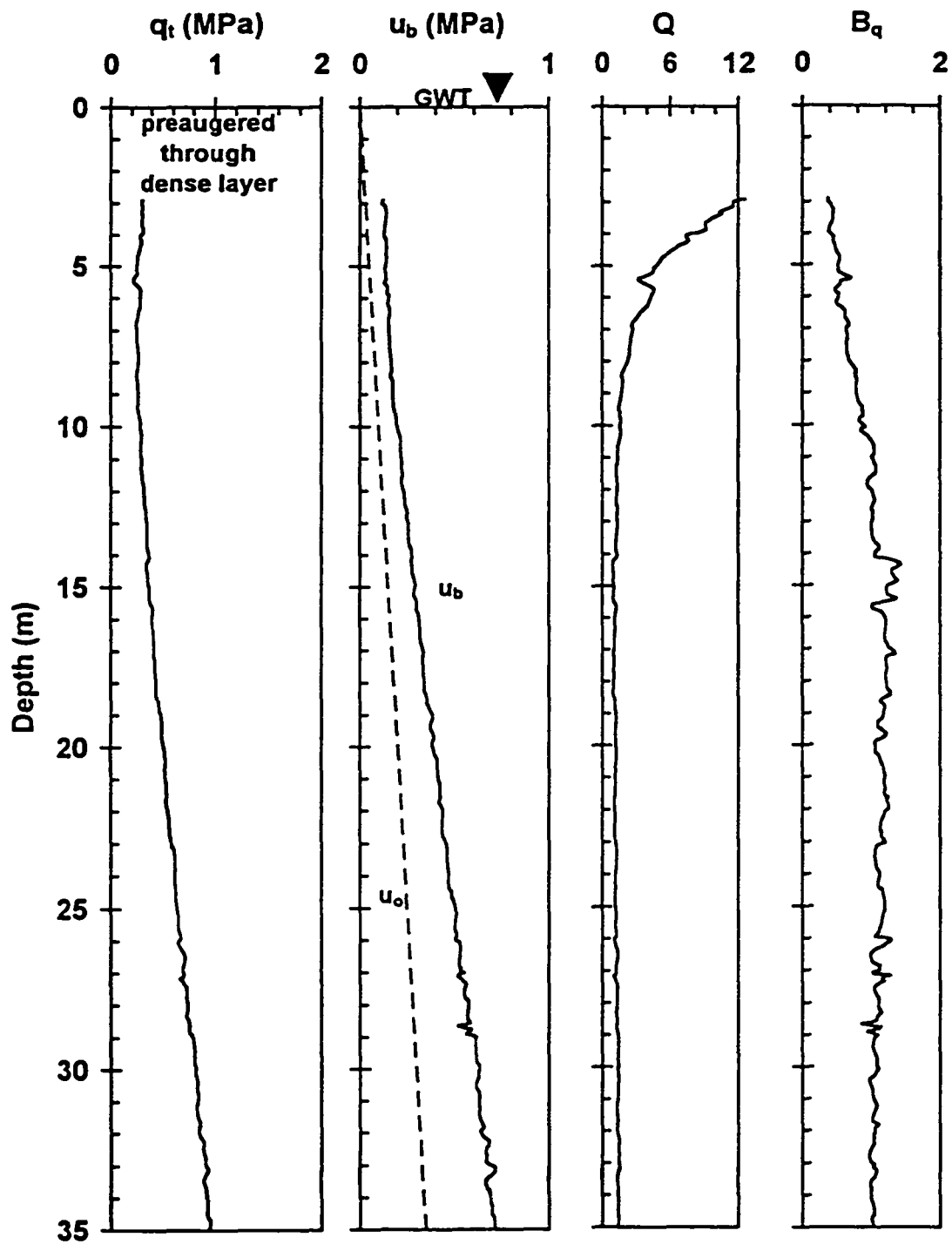
Figure 6.11. Comparison Between Cluster Analysis, Visual Classification and Vane Shear Strength at Gloucester, Ontario.

within the soil profile which can supplement the cone classification chart for a reliable identification of a geostatigraphy including the soil type.

6.2.3. Hachirogata Test Site

Raw piezocone readings from Hachirogata, Japan have been reported by Tanaka et al. (1992) and are presented along with the normalized parameters Q and B_q in Fig. 6.12. The soil profile consists of a deep deposit of a soft marine clay and the groundwater table is found at the ground surface. The visual inspection of q_t and u_b indicates one soil layer consisting of clay. Looking at the Q profile, there are two noticed trends of data, however, it is very difficult to define the boundary between them by eye. By examining B_q profile, a soil boundary is defined at approximately 15 m where there are positive and negative trends of B_q above and below this depth, respectively.

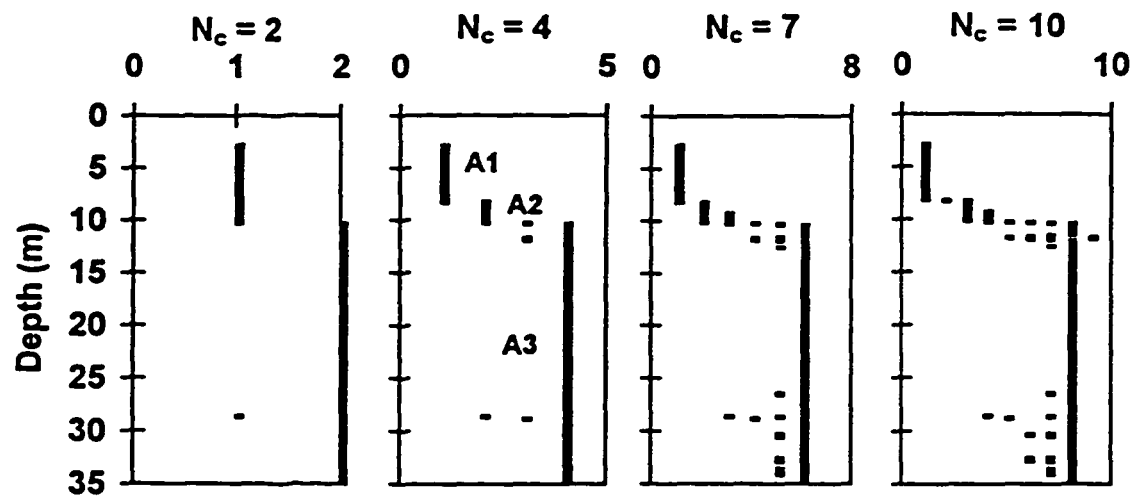
A single-cosine-zscore cluster analysis is applied to the piezocone data using Q and B_q and the cluster results are examined between $N_c = 2$ and 100 as discussed in Appendix C. The growth of data grouping at the peaks of ρ_c is shown in Fig. 6.13 between cluster number $N_c = 2$ and 10. At $N_c = 2$, two primary clusters appear with a boundary at 10.3 m. Then at larger clusters, soil lenses disjoin the lower layer indicating soil lenses or outliers within the deposit. A soil transition also starts to build up around the depth of 10.3 m. At $N_c = 4$, the upper cluster splits into two primary groups with a boundary at a depth of 8.3 m. At higher clusters, no new primary layers ($t \geq 1$ m) separate, therefore, a cluster number 4 is chosen to represent the subsurface stratigraphy at the site.



**Figure 6.12. A Representative Piezocone Sounding at Hachirogata, Japan
(Data from Tanaka et al., 1992).**

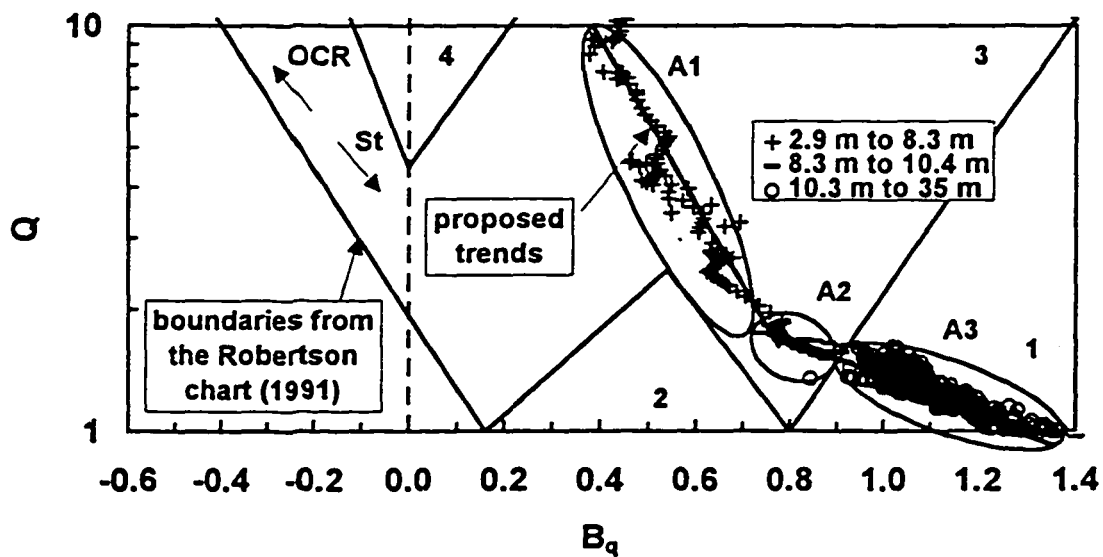
Using Robertson's chart (1991), the soil is classified as clay to a silty clay between depths of 2.9 m to 10.4 m underlain by a sensitive clay down to a depth of 35 m as shown on Fig. 6.14. The classification method is able to detect the effect of the clay fraction on the measured piezocone data. Note also that the Q and B_q data are visually divided into two trends separated at a depth of 6.4 m at which the clay fraction dramatically starts to increase from approximately 30 percent to 70 percent at a depth of 10 m. Therefore, identification of different trends in a Q - B_q space helps to indicate major changes in clay properties.

Clustering is supported by the variation of the clay fraction (< 0.005 mm) with depth as shown on Fig. 6.15. Note that in geotechnical practice in the United States, the clay fraction is defined to be less than 0.002 mm (ASTM guide D-2487). Also, note that the cone data are functions of the mean particle size and the fines content (Jamiolkowski et al., 1985). The clay fraction (CF^*) is divided into three groups. The first group validates a cluster A1 where the CF^* increases from 20 percent at 3 m to 25 percent at 7 m. In the second group, the CF^* has a relatively steeper transition trend whilst it increases to 65 percent at 10.2-m depth. This verifies the statistical primary layer A2. The last group of CF^* between the depths of 10.2 m and 35 m has a mean equal to 66 percent and a range varies between 49 percent and 78 percent. The latter group supports the soil deposit denoted by A3. The cluster results properly demarcates the soil stratigraphy at the site. The Robertson's method is also able to detect two primary layers in this soil deposit verified by the available clay fraction data.



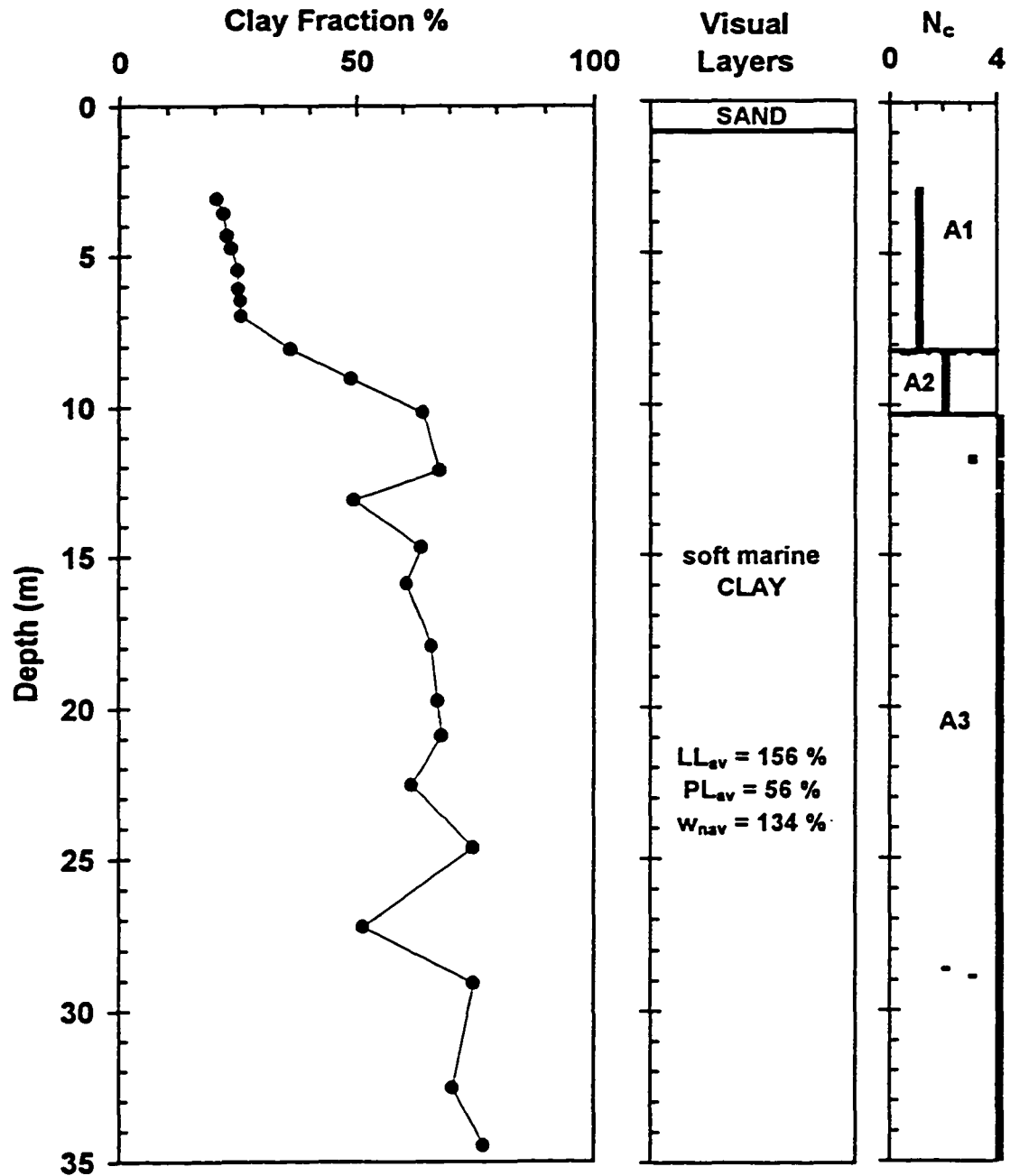
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Hachirogata, Japan from Tanaka et al. (1992).

Figure 6.13. Cluster Analysis of Piezocone Data at Hachirogata, Japan.



1. Sensitive Fine Grained	3. Clays: Clay to Silty Clay
2. Organic Soils-Peats	4. Silt Mixtures: Clayey Silt to Silty Clay

Figure 6.14. Clustering Results of Piezocone Data at Hachirogata, Japan on a Q - B_q Space (Data from Tanaka et al., 1992).



Note: Piezocone data at Hachirogata, Japan from Tanaka et al. (1992).
 SCZ method using Q and B_q , cluster no. 4

— Primary boundary
 - - - Transition boundary

Figure 6.15. Comparison Between Cluster Analysis, Visual Classification and Clay Fraction at Hachirogata, Japan.

6.2.4. Lilla Mellösa Test Site

The Lilla Mellösa site is located in Sweden and a representative piezocone sounding from Larsson and Mulabdić (1991) is shown in Fig. 6.16. The soil profile consists of 14 m of post-glacial organic to slightly organic clays overlaying a thin sand layer sitting on a bed rock. The groundwater table is near 0.8 m depth. The soil stratigraphy between 1.2 m and 12 m (where piezocone data are available) consists of organic clay with shells down to 5.5 m overlying a clay to 10.5 m depth, underlain by varved clay which becomes more pronounced with depth down to 12 m.

The derived normalized piezocone parameters Q and B_q are shown on Fig. 6.16. Looking at the q_t , u_b or B_q profiles, a single primary soil layer is interpreted by the author. A soil boundary perhaps could be observed at 2-m depth by examining the Q profile.

A single-cosine-zscore cluster type is applied to the normalized parameters Q and B_q and cluster analysis is performed from $N_c = 2$ to 100 as discussed in Appendix C. The cluster results at the peaks of ρ_c are shown in Fig. 6.17 from cluster number $N_c = 2$ to 12. Two primary soil layers are given at cluster number 2 and a transition layer is also defined between them from 5.9 m to 6.35 m. Then some points which are less correlated with the two primary clusters separate indicating soil lenses and transitions but no a new primary statistical group ($t \geq 1$ m) is detected up to $N_c = 100$. Therefore a cluster number $N_c = 2$ is chosen to represent the soil stratigraphy at the site.

The soil is classified using Robertson (1991) chart yielding a uniform layer of clay or silty clay. The soil sensitivity appears to increase, and/or the overconsolidation ratio

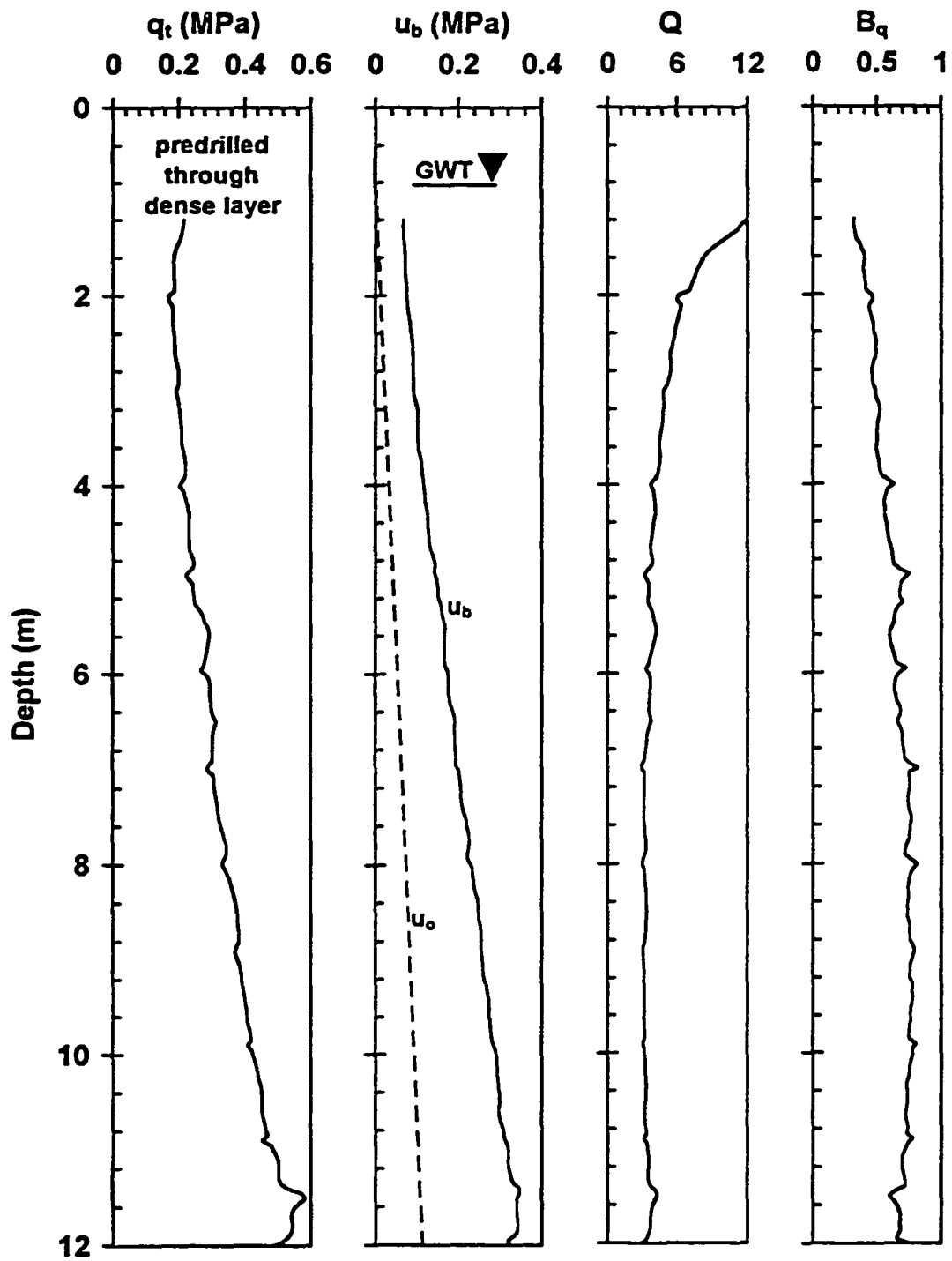
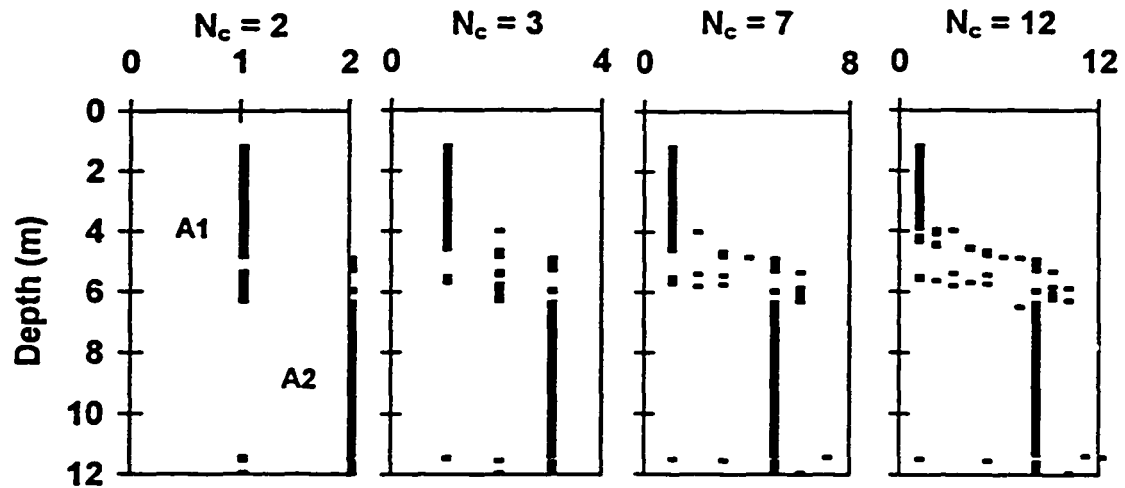


Figure 6.16. A Representative Piezocone Sounding at Lilla Mellösa, Sweden (Data from Larsson and Mulabdić, 1991).

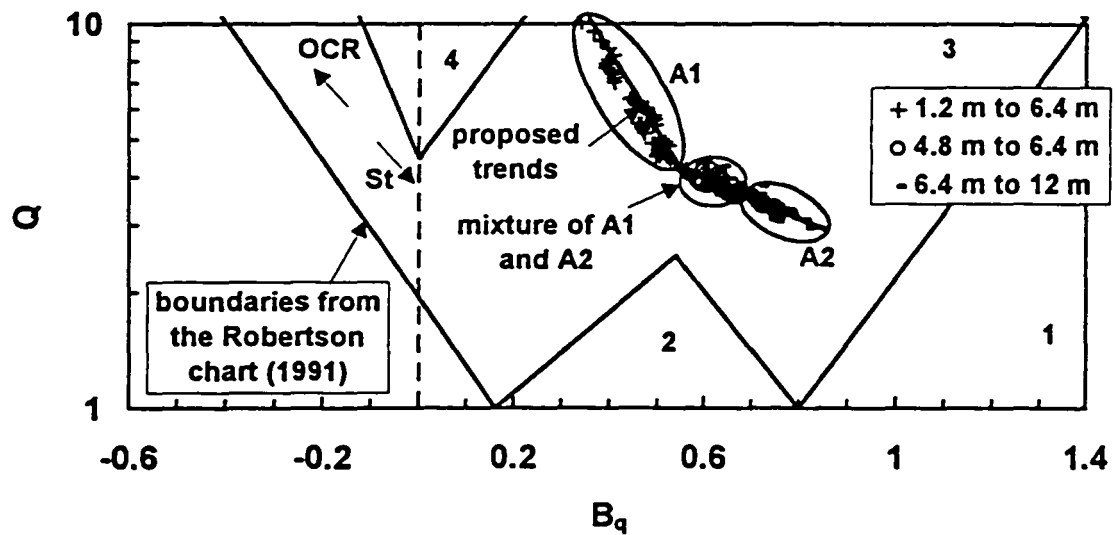
appears to decrease with depth as depicted in Fig. 6.18 which also includes the identified primary groups. Note that the data in a $Q-B_q$ space are visually divided into two trends approximately separated at a depth of 4.8 m which is above the boundary between the organic clay and the clay by a distance equal to 0.7 m. Therefore, these notified trends can give an indication of the change of the organic contents within the soil profile.

The cluster results are validated by the independent measurements of organic contents with depth as shown on Fig. 6.19. In case of organic soils, the void ratio and compressibility are greater than that in the case of clay soil. Also, the shear strength in an organic clay is less than that in a clay soil. Cone data are functions of the soil compressibility and shear strength (Robertson and Campanella, 1989), therefore they are correlated with the organic contents in a soil. The organic contents are divided into three groups. In the first group, the organic contents decrease from 5.5 percent at 1.8 m to 3 percent at 4.6 m which supports the primary cluster A1. The second group is a transition zone defined between 4.6 m and 5.8 m where organic contents remain at the level of 3 percent. This group supports the transition zone in which cluster numbers alternates between 1 and 2 between the two primary layers. In the third group, the organic contents decrease from 3 percent at 5.8 m to 0.9 percent at 11.5 m which supports the primary cluster A2. Note that the borehole boundary at 10.5 m between the clay and varved clay is not detected. This may be attributed to the relative location of the borehole in the vicinity of the piezocone sounding and the gradual increase of the varves from 10.5 m to 12 m.



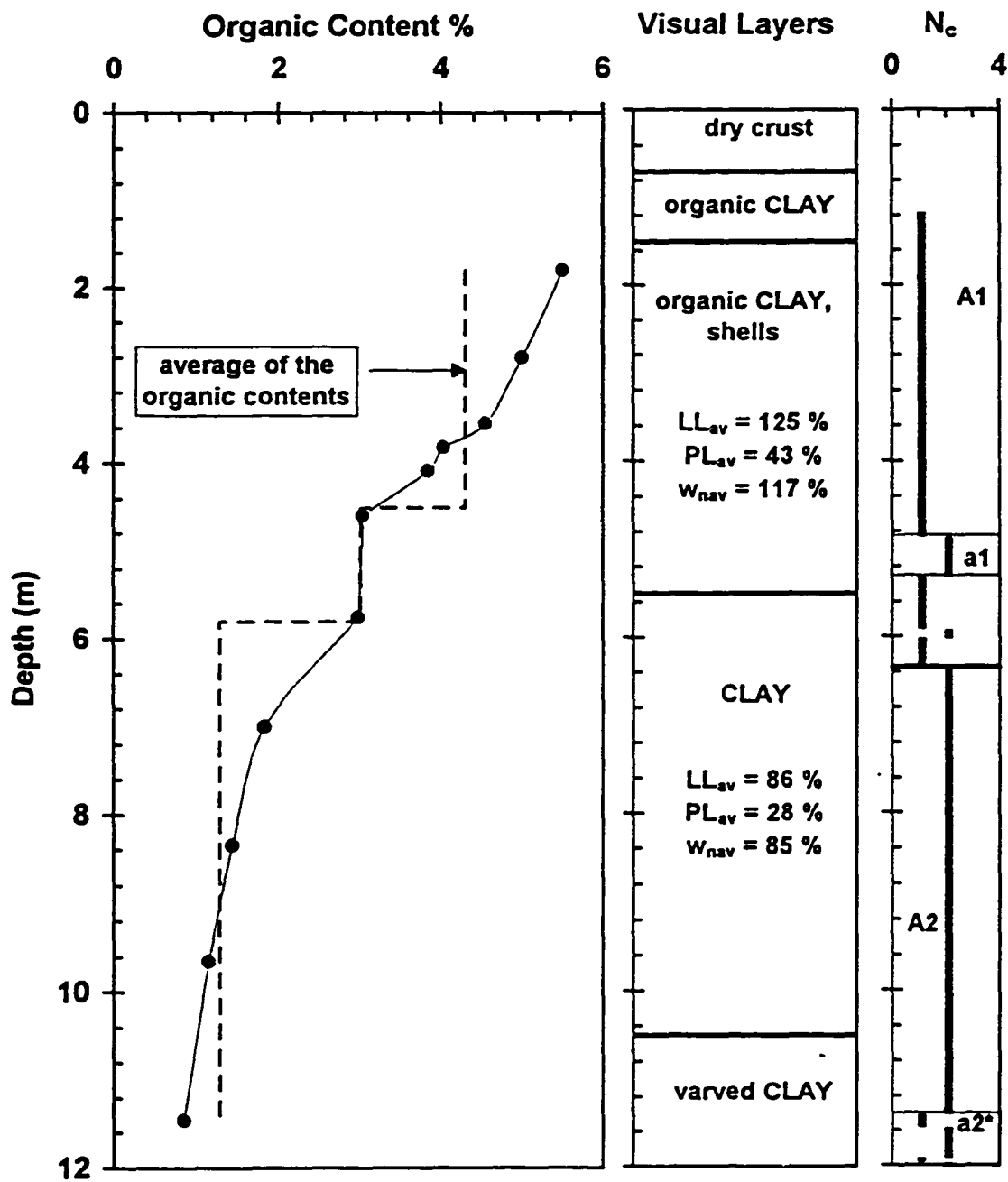
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Lilla Mellösa, Sweden (Larsson and Mulabdić, 1991).

Figure 6.17. Cluster Analysis of Piezocone Data at Lilla Mellösa, Sweden.



1. Sensitive Fine Grained	3. Clays: Clay to Silty Clay
2. Organic Soils-Peats	4. Silt Mixtures: Clayey Silt to Silty Clay

Figure 6.18. Clustering Results of Piezocone Data at Lilla Mellösa, Sweden on a Q - B_q Space (Data from Larsson and Mulabdić, 1991).



Note: Piezocone data at Lilla Mellösa, Sweden (Larsson and Mulabdić, 1991).
 SCZ method using Q and B_q , cluster no. 2

— Primary boundary
 - - - Secondary boundary

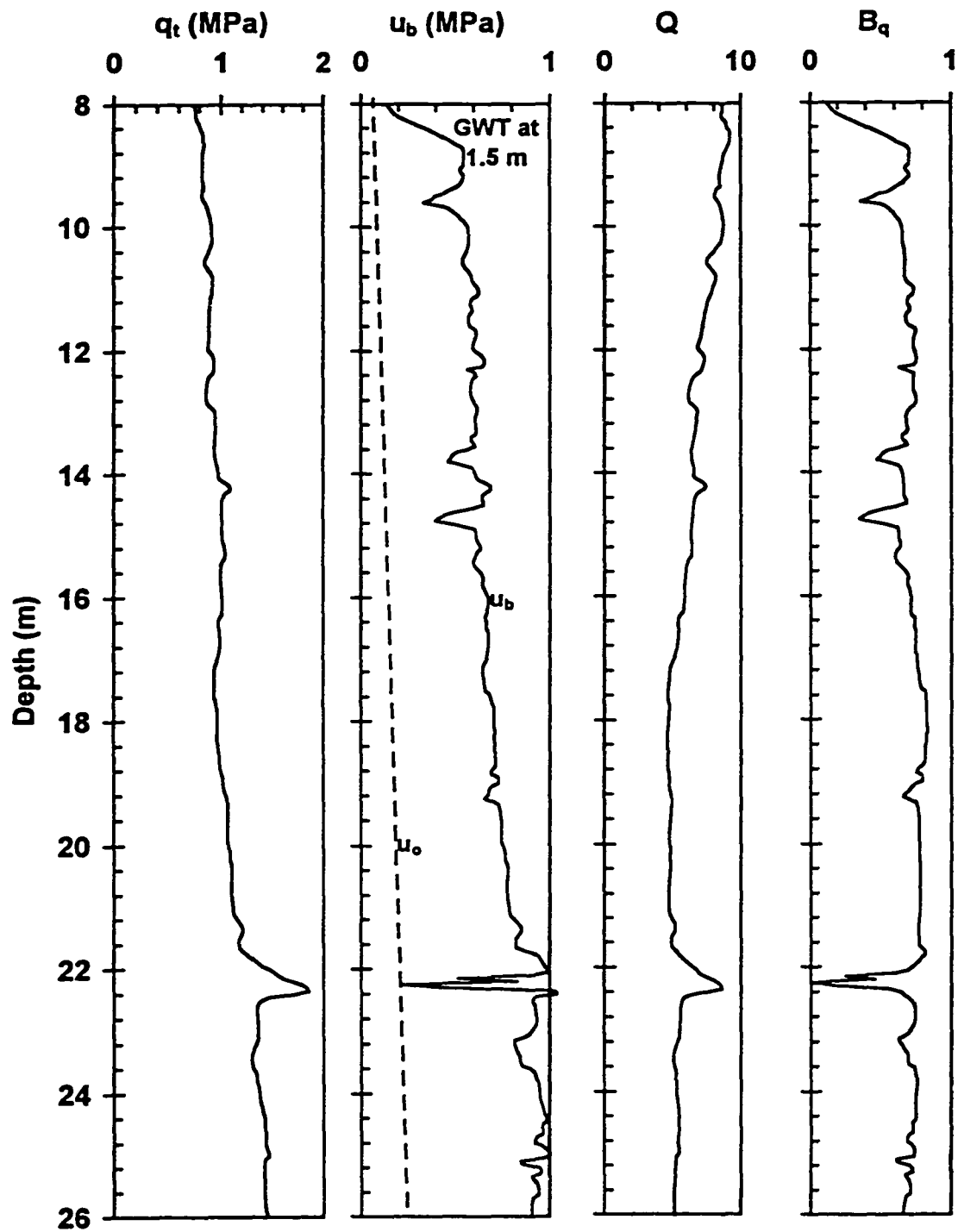
Figure 6.19. Comparison Between Cluster Analysis, Visual Classification and Organic Contents at Lilla Mellösa, Sweden.

6.2.5. Recife Test Site

The Recife site is located in the northeastern coast of Brazil and the soil was formed in the Quaternary period with an age of almost 10,000 years (Coutinho and Oliveira, 1997). The soil deposit is affected by both salt and freshwater, and organic soft clays are common in this area and located below the groundwater table. The soil stratigraphy at Recife test site consists of an upper 6 to 7 m of sand to sandy clay underlain by an organic soft clay down to a depth of almost 26 m. The groundwater table is at a depth of $1.5 \text{ m} \pm 0.5 \text{ m}$ fluctuation.

A representative piezocone sounding at the site and the derived normalized parameters Q and B_q are shown in Fig. 6.20. The data are presented over the depths of interest between 8 m and 26 m where the soils are comprised of organic soft clays. Looking at the unprocessed u_b and normalized B_q readings, a secondary layer is defined between 8 m and 8.8 m, and soil lenses are detected at 9.6 m, 13.8 m, 14.8 m, and 22.2 m. The soil lense at a depth of 22.2 m is also discovered by visual examination of q_t and Q readings.

A single-cosine-zscore cluster method is applied to Q and B_q data and the cluster results are examined between $N_c = 2$ and 100 as discussed in Appendix C. Figure 6.21 shows the growth of the cluster analysis at the peaks of ρ_c between $N_c = 2$ and 10. At N_c equal to 2, two primary data groups are detected with a boundary at 15.6 m and also a soil lense is found between 21.9 m and 22.4 m. For $N_c > 2$, no new primary clusters ($t > 1 \text{ m}$) appear up to $N_c = 100$, however, transitions and lenses are detected indicating non-

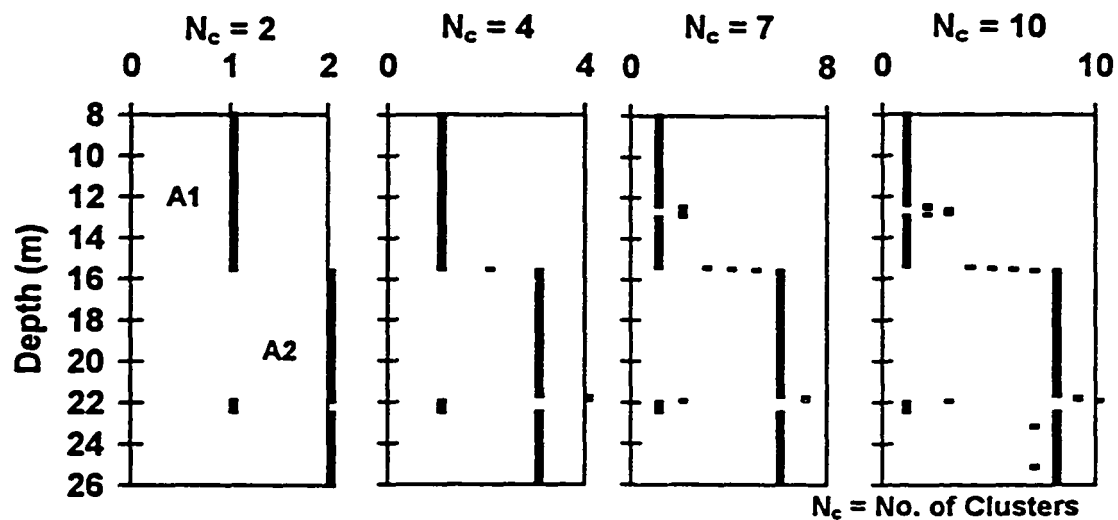


**Figure 6.20. A Representative Piezocone Sounding at Recife, Brazil
(Data from Coutinho and Oliveira, 1997).**

homogeneity within the two primary clusters. Therefore cluster number 2 is chosen to designate the soil stratigraphy at the site.

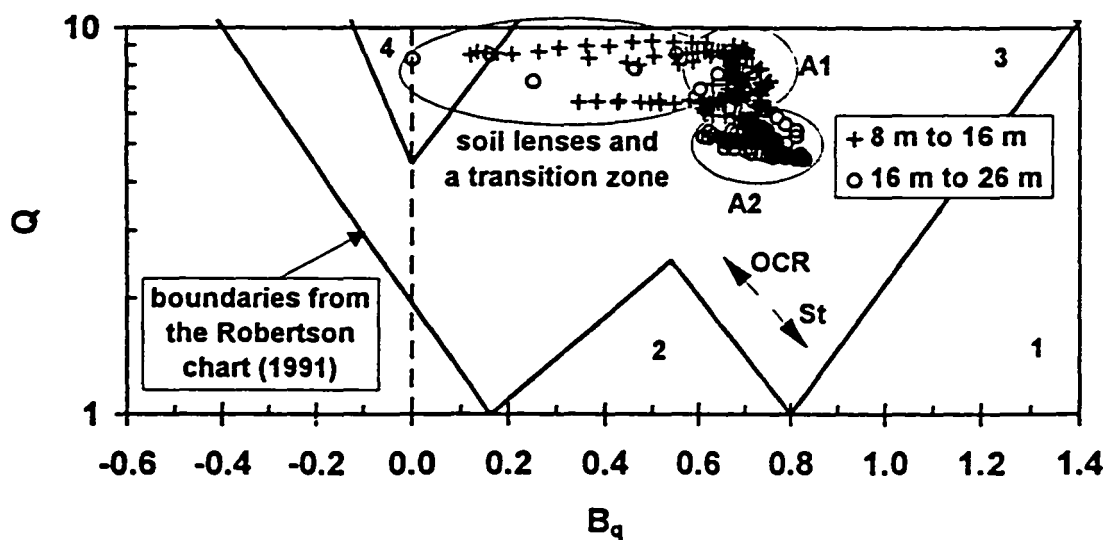
The normalized parameters Q and B_q are processed using Robertson (1991) chart and a single clay soil type is indicated. The normalized pore pressure significantly decreases between the depths of 22.25 m and 22.40 m which identifies a soil lense. The two primary groups A1 and A2 as well as soil lenses and a transition zone is shown in Fig 6.22. The clustering results indicate that there is a change of the properties of the two layers A1 and A2. The classification method suggests that the overconsolidation ratio decreases and/or sensitivity increases moving the cone from A1 to A2. This information is very useful in a preliminary site investigation and two representative soil samples of A1 and A2 can be tested to distinguish the difference between their properties such as the soil plasticity and overconsolidation ratio.

The cluster results do not agree with either the visual examination of the piezocone data or the CPT classification scheme by which one clay soil type is defined. However, the statistical soil stratigraphy defined at N_c equal to 2 is validated by defined soil layers based on independent laboratory tests, including soil water content (w_n) and plasticity measurements as shown in Fig. 6.23. The organic clay soil between the depths of 8 m and 15 m has an average water content, liquid limit, and plastic limit in the order of 88 percent, 119 percent, and 42 percent , respectively. Their averages decrease to 58 percent, 67 percent, and 33 percent, respectively, at the lower organic clay layer below 15 m and down to 24.5 m.



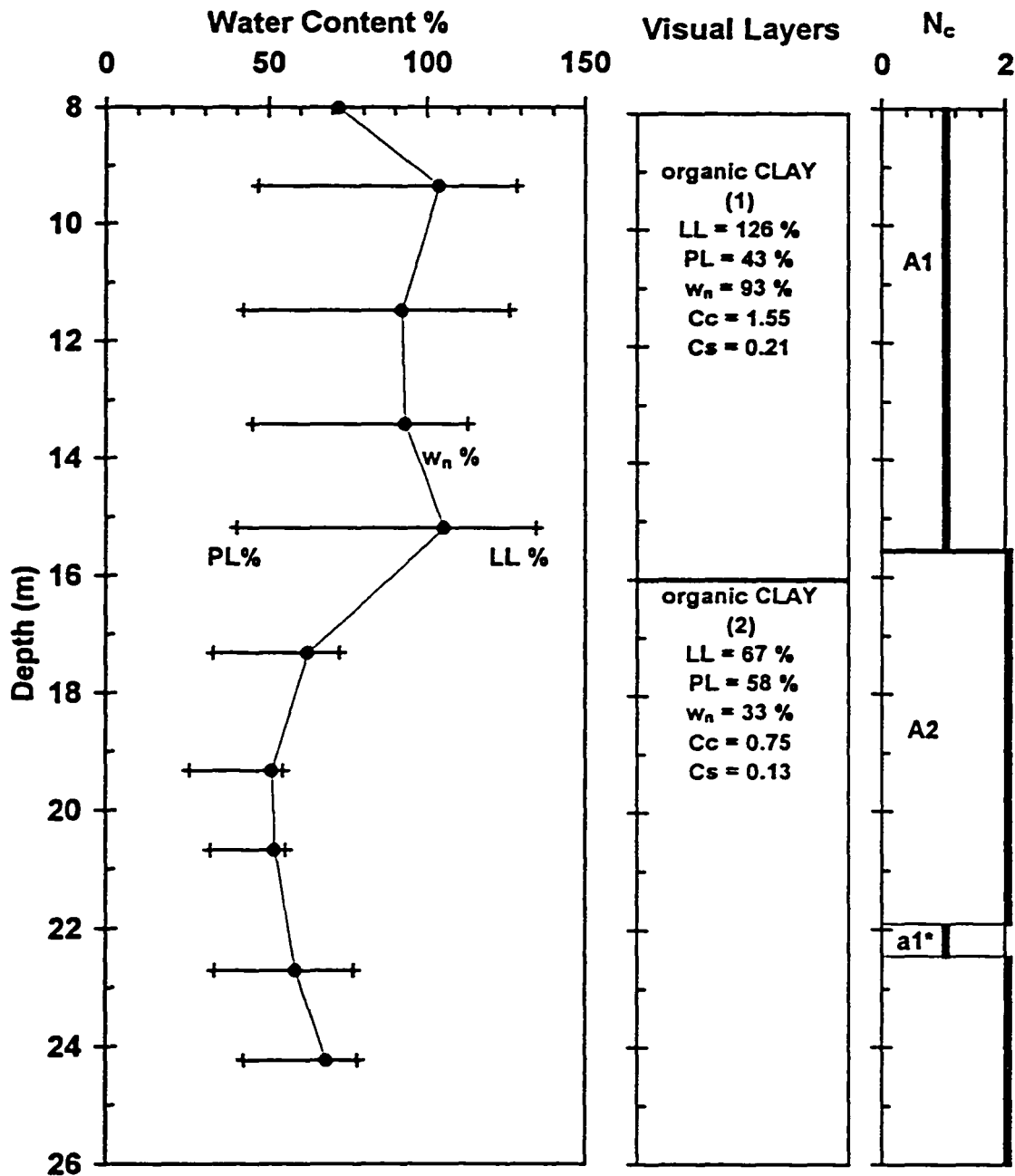
Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data at Recife, Brazil from Coutinho and Oliveria (1997).

Figure 6.21. Cluster Analysis of Piezocone Data at Recife, Brazil.



1. Sensitive Fine Grained
 2. Organic Soils-Peats
 3. Clays: Clay to Silty Clay
 4. Silt Mixtures: Clayey Silt to Silty Clay

Figure 6.22. Clustering Results of Piezocone Data at Recife, Brazil on a Q - B_q Space (Data from Coutinho and Oliveria, 1997).



Note: Piezocone data at Recife, Brazil from Coutinho and Oliveira (1997).
 SCZ method using Q and B_q, cluster no. 2

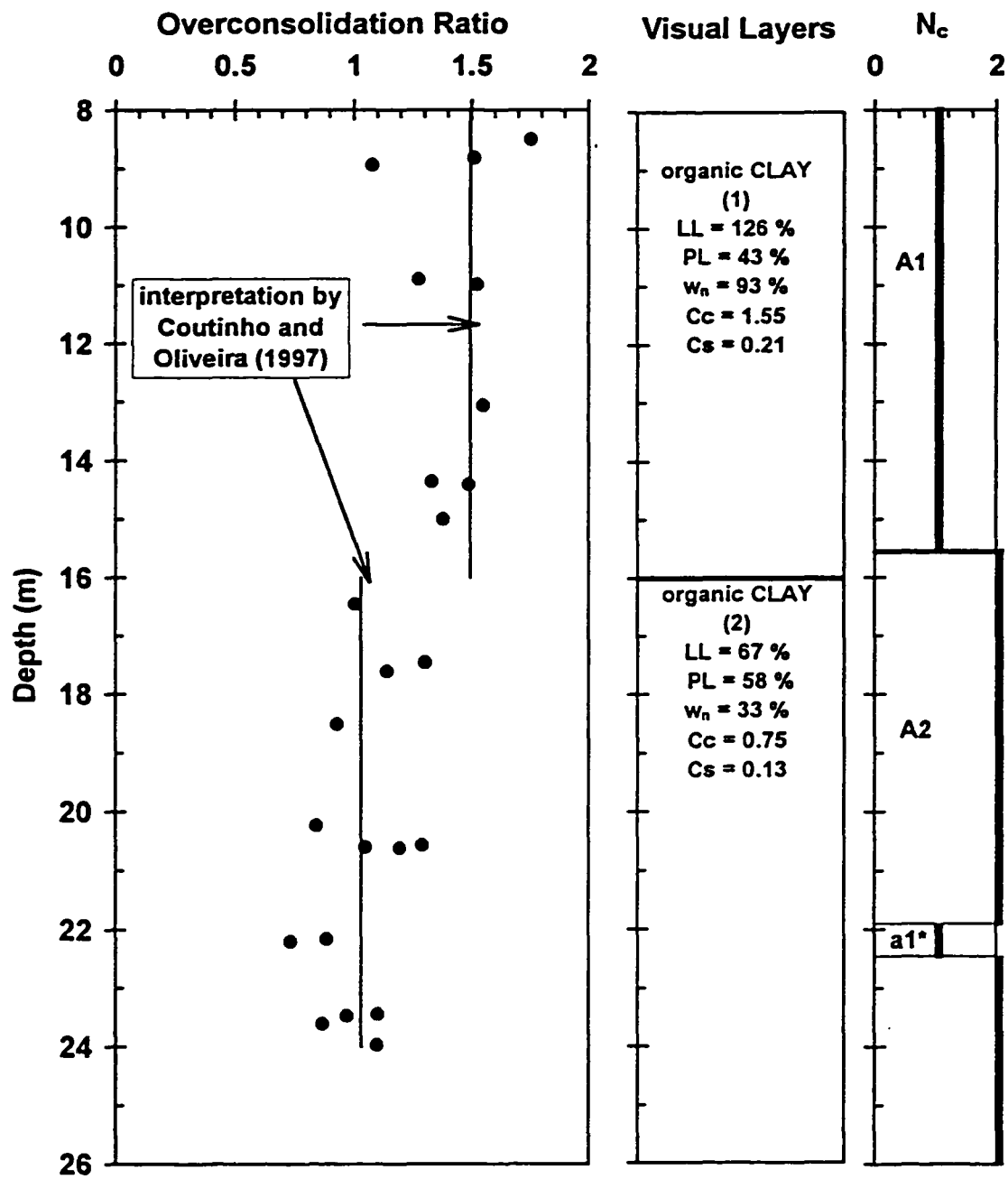
— Primary boundary
 — Secondary boundary

Figure 6.23. Comparison Between Cluster Analysis, Visual Classification and Water Content at Recife, Brazil.

Mechanical properties from laboratory consolidation tests on undisturbed samples also support the cluster results. For instance, Figure 6.24. shows two groups of derived overconsolidation ratios (OCR). The division at 16 m is the interpretation by Coutinho and Oliveira (1997). The average of the upper group between depths of 8 m and 16 m is equal to 1.5 and that of the lower group down to 24 m is equal to 1. Coutinho and Oliveira (1997), indicated that the average compression index (C_c) in the upper layer was twice as much that of the lower layer, and they are equal to 1.55 and 0.75, respectively. Also, the average swelling indices (C_s) of the two layers were different and equal to 0.21 for the upper layer and 0.13 for the lower layer. Therefore, cluster analysis of piezocone data at Recife test site, Brazil, was able to detect the changes of the soil properties within the organic clay deposit.

6.2.6. St. Alban Test Site

The Saint Alban test site is located in Quebec, Canada (Roy et al., 1982). The soils were formed during the late Pleistocene period in the Champlain Sea and have become slightly overconsolidated with a constant overconsolidation ratio 2.2. The soils have been subjected to a slight geological preconsolidation and a significant quasi-preconsolidation which could be due to a secondary consolidation and/or cementation of the clay. The soils consist of the following clay profile: top soil and a weathered clay crust to depths of 1.6 meters underlain by sensitive clay. The groundwater table is located approximately at 0.7 m.



Note: Piezocone data at Recife, Brazil from Coutinho and Oliveira (1997).
 SCZ method using Q and Bq, cluster no. 2

—— Primary boundary
 —— Secondary boundary

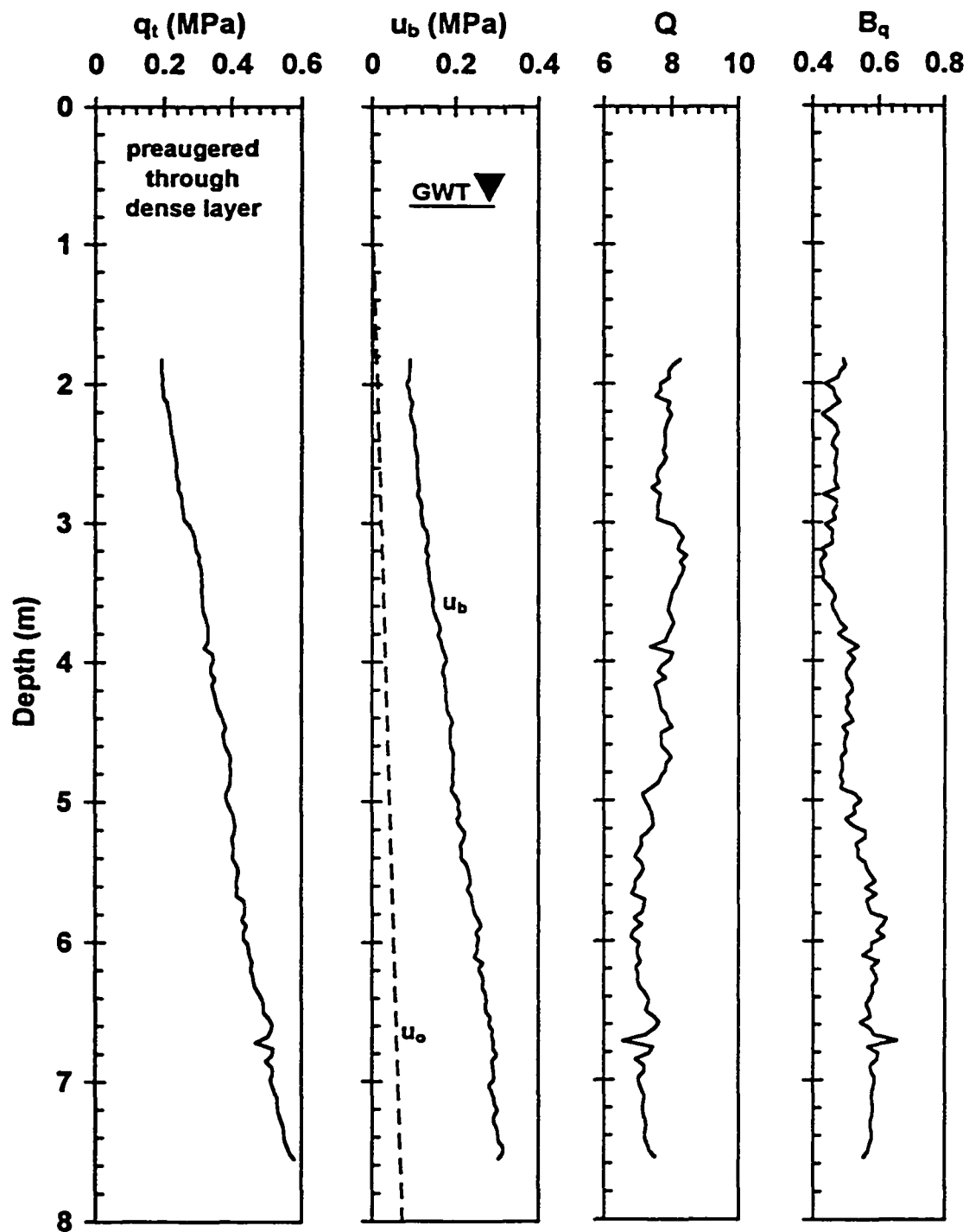
Figure 6.24. Comparison Between Cluster Analysis, Visual Classification and Overconsolidation Ratio at Recife, Brazil.

A representative piezocone sounding and its derived normalized parameters are shown in Fig. 6.25. The results indicate by eye a single soil layer and a probable small soil lense at almost 6.7 meters depth. A single-cosine-zscore cluster analysis was applied to the Q and B_q normalized parameters and the growth of clustering is inspected between $N_c = 2$ and 100 as discussed in more detail in Appendix C.

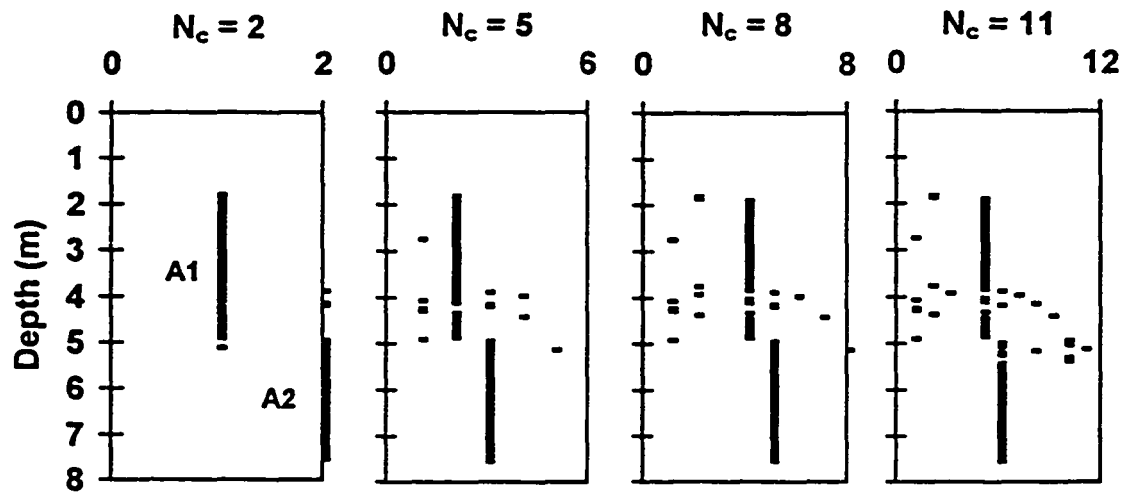
The cluster results at the peaks of ρ_c are shown in Fig. 6.26 between $N_c = 2$ and 11. At N_c equal to 2, two primary clusters are separated with a boundary at a depth of 5.2 m. Two soil lenses appear in the upper layer at 3.90 m, and 4.20 m to 4.25 m. A part of a soil transition between the two primary layers starts to build up at a depth of 5.15 m. At larger cluster numbers, more soil lenses and transitions continue to disjoin the two original clusters but no a new primary group ($t \geq 1$ m) is detected. Therefore, a cluster number 2 is selected to stand for the subsurface stratification at the site.

The soil is classified using the CPT chart to indicate a single clay to silty clay layer as shown in Fig. 6.27. The two primary identified clusters are shown in Fig. 6.27 and indicate a change in the clay properties at a depth of 4.8 m which helps the engineer to decide the number and depths of soil samples. Therefore clustering is a useful complementary tool to the classification chart for the purpose of reliable and objective soil stratigraphy.

The cluster results are independently confirmed by the natural water contents (w_n) and clay plasticity data as shown in Fig. 6.28. The first trend which supports cluster A1 is between 1.8 m and 4.7 m where the w_n decreases from 91 percent to 43 percent, respectively. Then a transition is indicated between 4.7 m and 5.2 m where the w_n

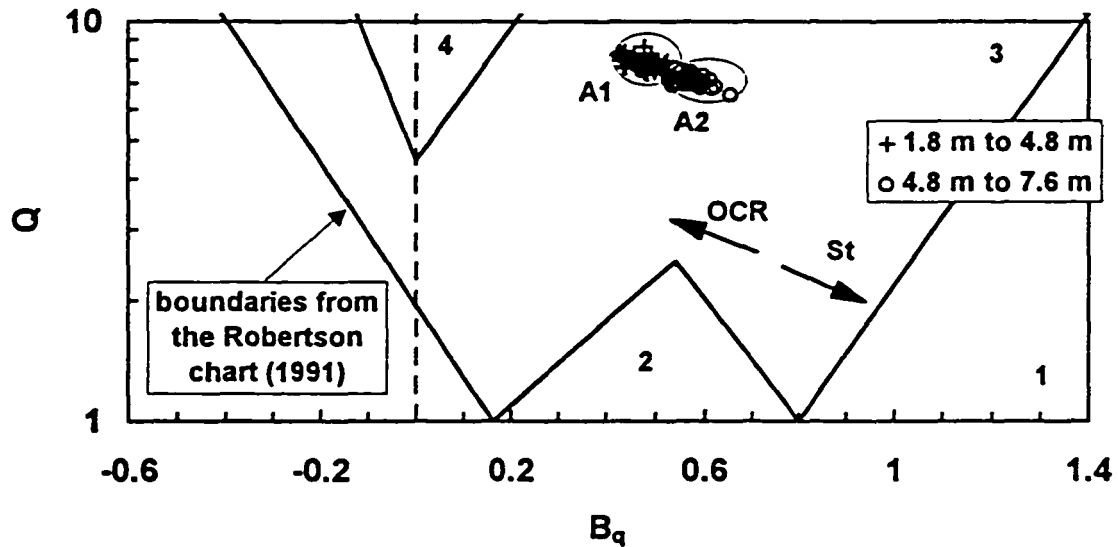


**Figure 6.25. A Representative Piezocone Sounding at St. Alban, Quebec
(Data from Roy et al., 1982).**



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at St. Alban, Quebec from Roy et al. (1982).

Figure 6.26. Cluster Analysis of Piezocone Data at St. Alban, Quebec.



1. Sensitive Fine Grained	3. Clays: Clay to Silty Clay
2. Organic Soils-Peats	4. Silt Mixtures: Clayey Silt to Silty Clay

Figure 6.27. Clustering Results of Piezocone Data at St. Alban, Quebec on a Q - B_q Space (Data from Roy et al., 1982)

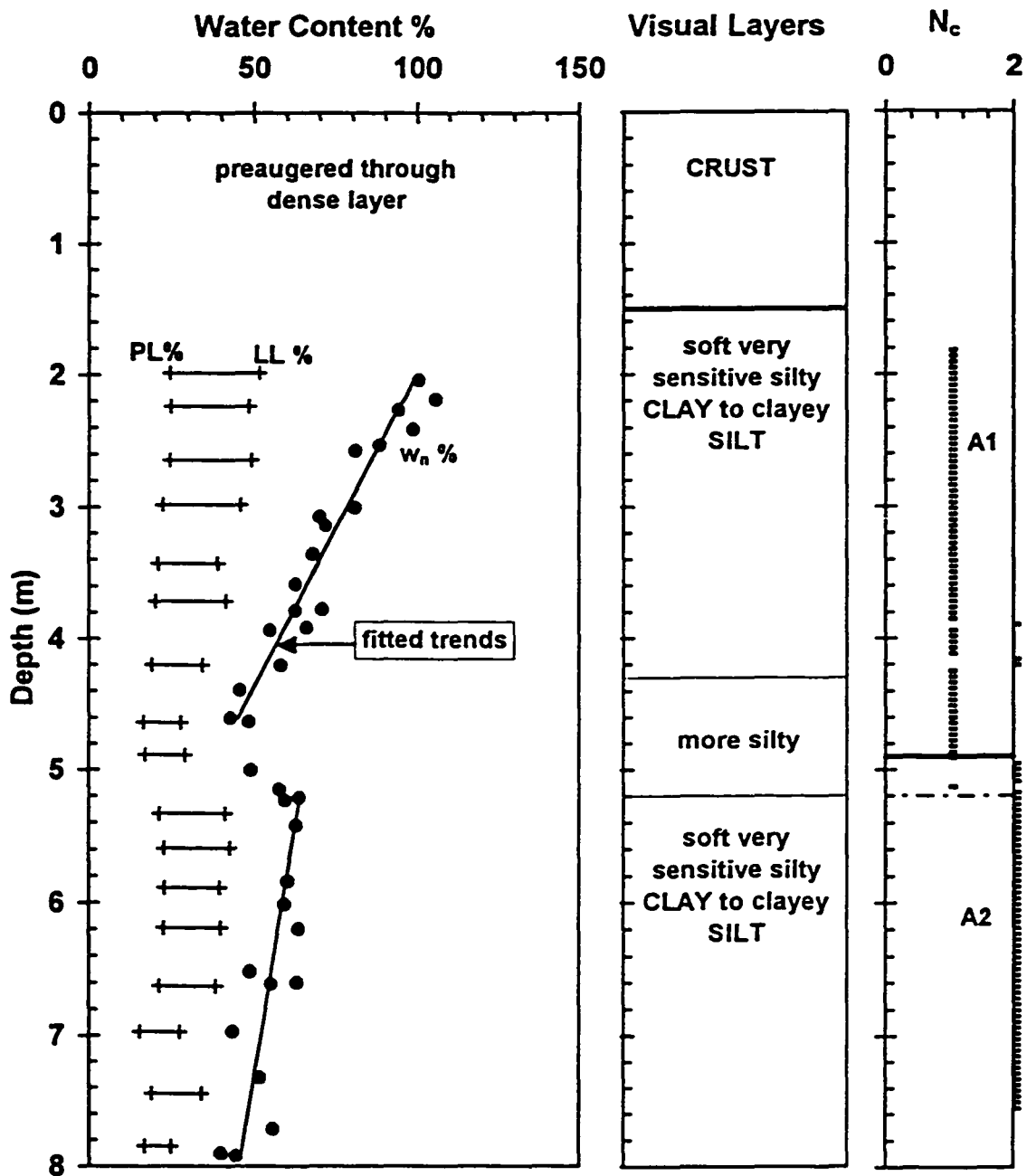


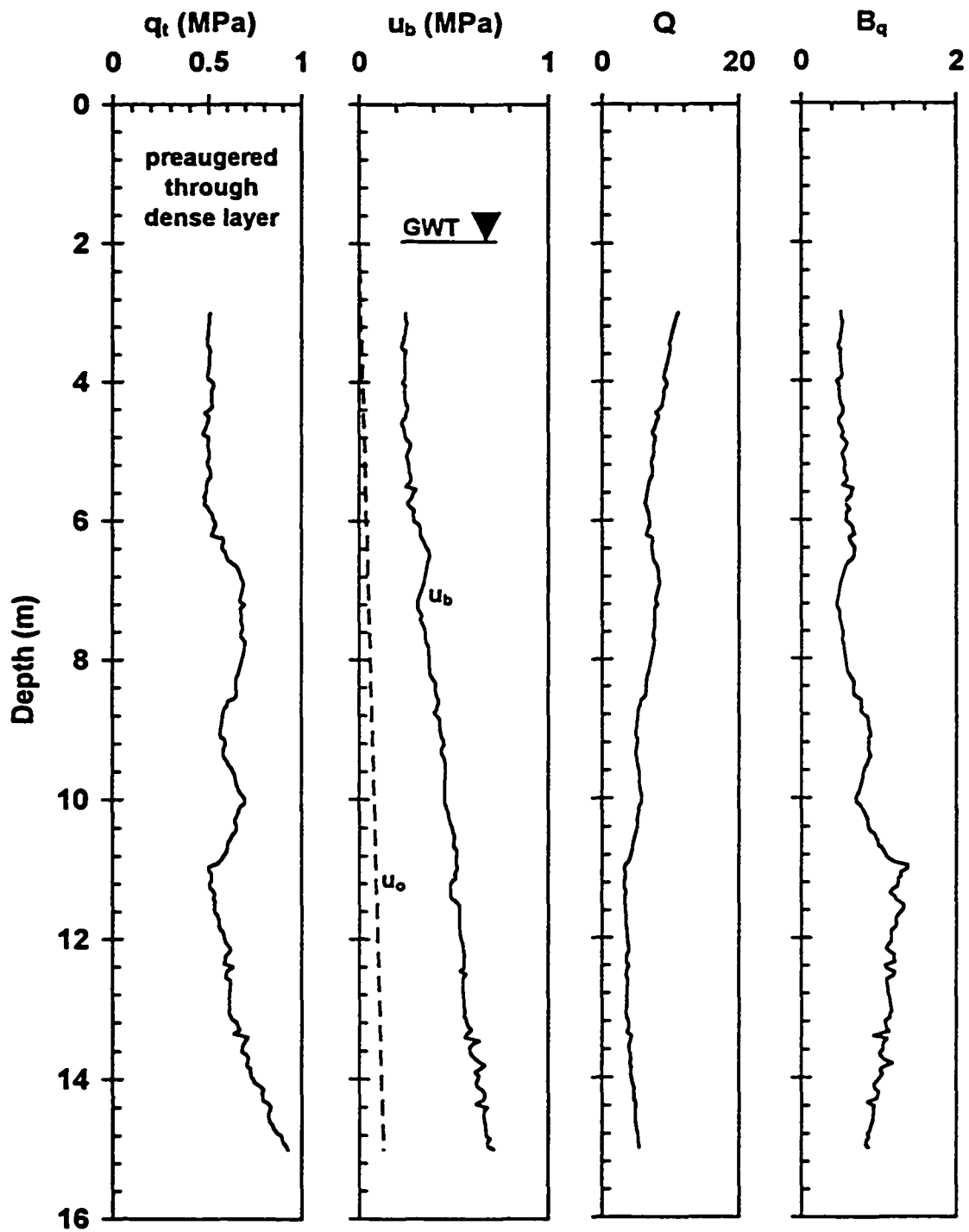
Figure 6.28. Comparison Between Piezocone Cluster Analysis, Visual Classification, and Water Content at St. Alban, Quebec.

increases from 43 percent to 64 percent. In the second trend which supports cluster A2, the w_a decrease from 64 percent at 5.2 m to 41 percent at 7.9 m.

6.2.7. Tiller Test Site

The Tiller clay site is located in Trondheim, Norway and was studied by Sandven (1990). The soil stratification below 3 m where the piezocone data and reference index properties are available consists of: 6 m of silty clay over 6 m of quick clay. A representative piezocone data and their derived normalized parameters Q and B_q are shown in Fig. 6.29. Soil boundaries might be visually defined where the trends of the unprocessed data q_t and the normalized parameters Q and B_q vary at approximate depths of 7 m, 9 m, 10 m and 11 m. However, by inspecting the unprocessed data u_b , a soil boundary might be delineated at a 6.5-m depth. There is no obvious or single interpretation can be given to the soil profile by merely looking at the piezocone data thus different interpretations can be obtained.

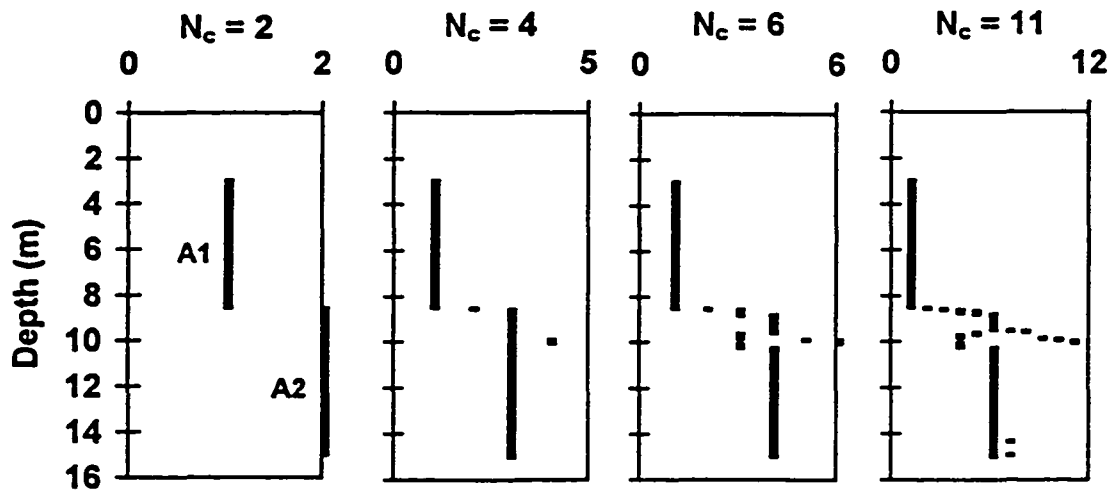
A single-cosine-zscore cluster analysis is applied to the Q and B_q normalized parameters and the growth of the cluster analysis between $N_c = 2$ and 100 is discussed in Appendix C. The cluster results at the peaks of ρ_c are shown in Fig. 6.30 between $N_c = 2$ and 11. At N_c equal 2, the data are grouped into two clusters with a boundary at 8.5 m where the soil type drastically changes from silty clay to quick clay. For higher clusters, no a new primary layer ($t \geq 1$ m) appears, however, some points which are less associated with the two primary clusters separate indicating soil transitions and lenses. Therefore cluster number 2 is selected to identify the soil profile at the site.



**Figure 6.29. A Representative Piezocone Data at Tiller, Norway,
(Data from Sandven, 1990).**

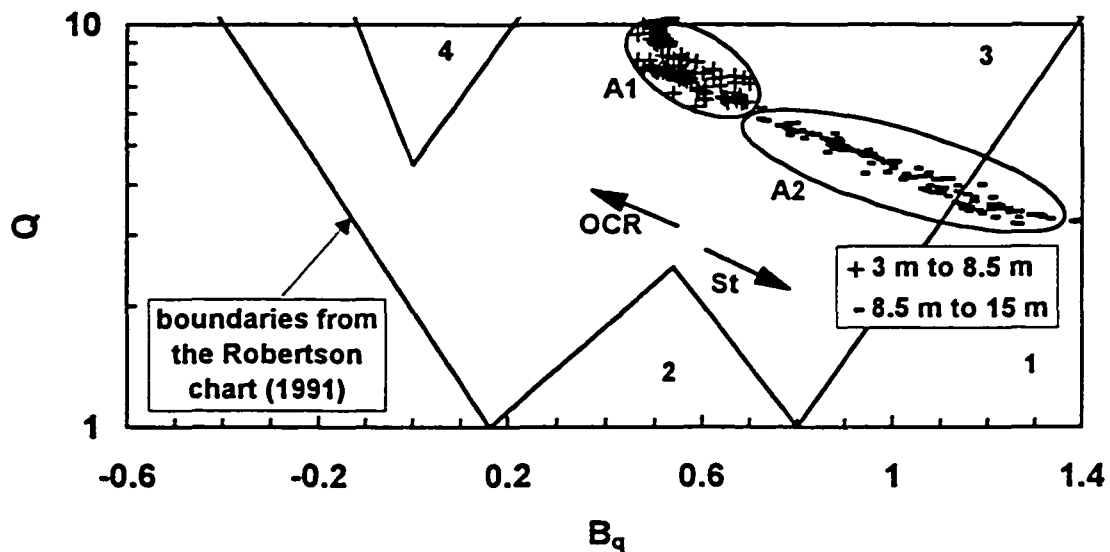
The piezocone chart is used to evaluate soil types as shown in Fig. 6.31. The soil between 3 m and 10.8 m is noted as clay to silty clay and a sensitive clay layer is detected between 10.8 m and 13.8 m. The soil below 13.8 m and down to 15 m is classified as clay to silty clay. However, the latter layer was classified as quick clay based on reference laboratory and field tests (Sandven, 1990). Moreover, the proposed boundary at 10.8 m is shifted 1.8 m below the actual boundary between the silty clay and quick clay layers. The two primary groups A1 and A2 are shown on Fig. 6.31 and indicate a significant difference of the properties of the two identified layers. The classification chart shows that moving the cone from layer A1 to A2, the sensitivity increases and/or the overconsolidation ratio decreases. Therefore combining the results of clustering analysis and the classification chart objectively indicates a preliminary information of soil stratigraphy and soil behavior. Note that the coefficient of variations of Q and B_q between the depths of 3.0 m and 8.5 m are equal to 0.35 and 0.30, respectively, and those of Q and B_q between the depths of 8.5 m and 15.0 m are equal to 0.17 and 0.16, respectively. The dispersion of the data in the lower quick clay layer is approximately equal to a half of that in the upper clay layer. Therefore, the data of the quick clay layer are grouped at a larger similarity levels than that of the clay layer which caused their separation at a low cluster number equal to 2.

The cluster results are confirmed by sensitivity data shown in Fig. 6.32, and measured using a fall cone test. Note that the cone parameters are functions of the soil sensitivity (Robertson et al., 1986). The sensitivity data are divided into 2 groups with a boundary at 8.5 m. The average sensitivity of the upper silty clay group is equal to 9



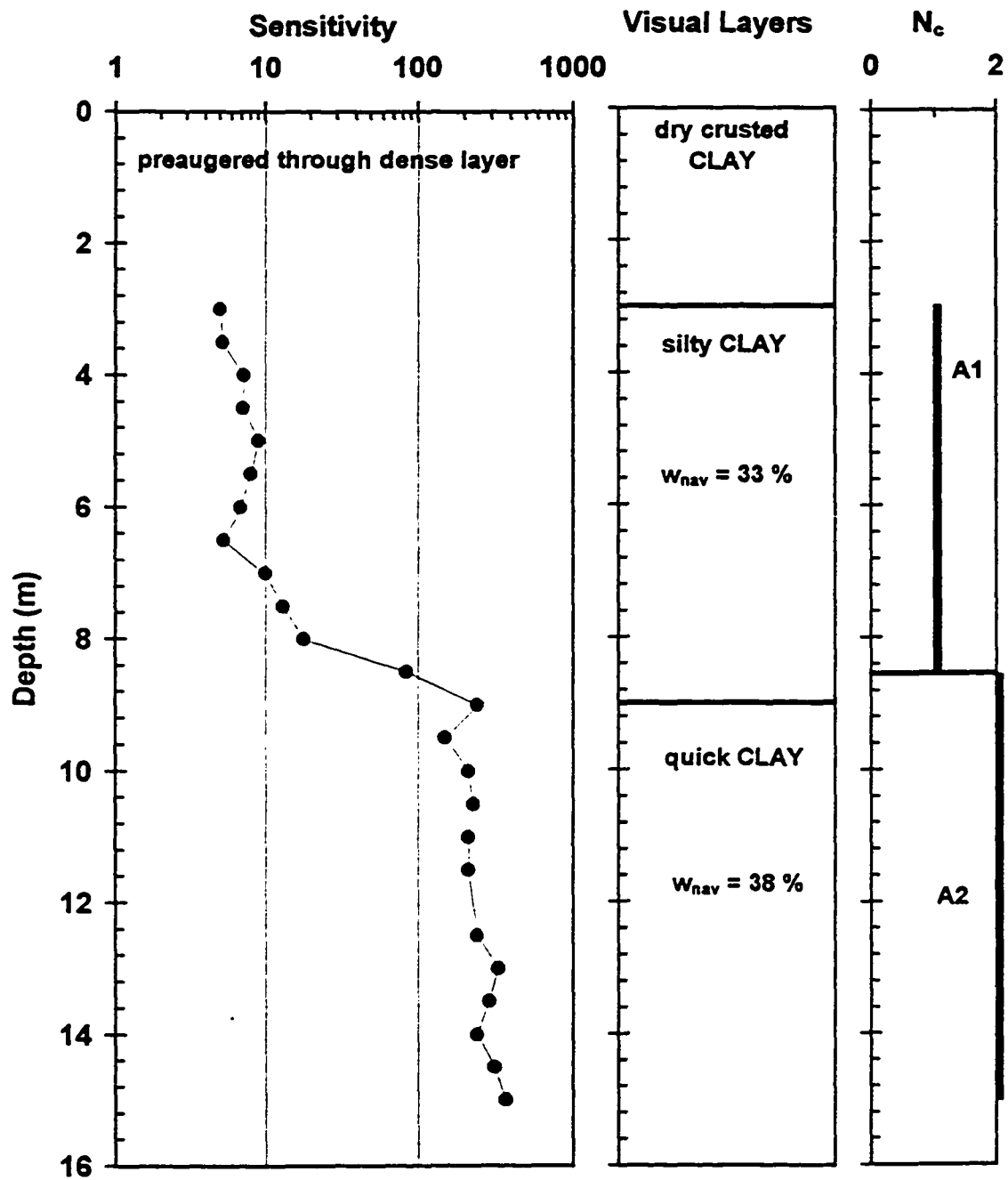
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Tiller, Norway from Sandven (1990).

Figure 6.30. Cluster Analysis of Piezocone Data at Tiller, Norway.



1. Sensitive Fine Grained	3. Clays: Clay to Silty Clay
2. Organic Soils-Peats	4. Silt Mixtures: Clayey Silt to Silty Clay

Figure 6.31. Clustering Results of Piezocone Data at Tiller, Norway on a Q - B_q Space (Data from Sandven, 1990).



Note: Piezocone data at Tiller, Norway from Sandven (1990).

SCZ method using Q and B_q , cluster no. 2

Primary boundary

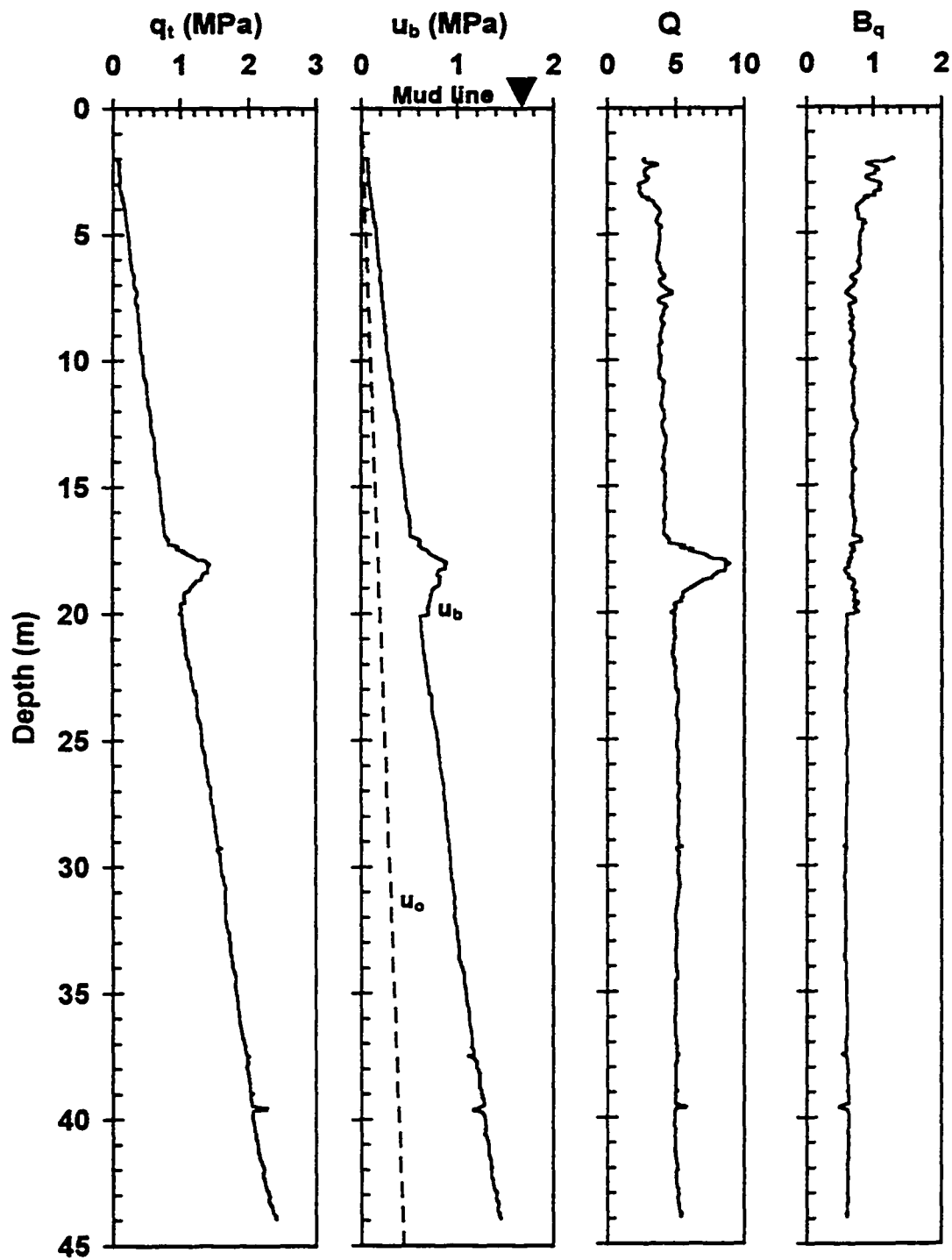
Figure 6.32. Comparison Between Cluster Analysis, Visual Classification and Sensitivity at Tiller, Norway.

between the depths of 3 m and 8 m. In the lower quick clay group, the average sensitivity dramatically increases to 240 between the depths of 8.5 m and 15 m. The former and latter sensitivity groups validate the statistical layers A1 and A2, respectively.

6.2.8. Troll Site

A laboratory and in-situ test program was performed for the offshore Troll site in the North Sea (Amundsen et al., 1985). The soil stratification in the zone of interest below the mud line and down to 45 m consists of 17 m of very soft to firm clay over 28 m of firm to very stiff silty clay. Typical piezocone data at the site and their derived normalized parameters Q and B_q are shown in Fig. 6.33. Looking at the piezocone profiles, one possible interpretation is that a single layer is indicated having an intermediate silty or sandy zone between depths of 17 and 20 m.

A single-cosine-zscore cluster analysis is applied to the piezocone data using the normalized parameters Q and B_q . The cluster analysis is evaluated between $N_c = 2$ and 100 as discussed in Appendix C. The cluster results at the peaks of ρ_c are shown in Fig. 6.34 between $N_c = 2$ and 12. At N_c equal to 2, two primary clusters are delineated indicating a major difference in the soil types and/or properties above 17.3 m and below 20 m. In the latter intermediate zone, the cluster numbers alternates between 1 and 2 pointing to a transition zone between the two major layers. For higher clusters, no new primary groups ($t \geq 1$ m) are detached, therefore, cluster number 2 is selected to indicate the soil profile at the site.



**Figure 6.33. Piezocone Data at Troll, North Sea
(Data from Amundsen et al., 1985).**

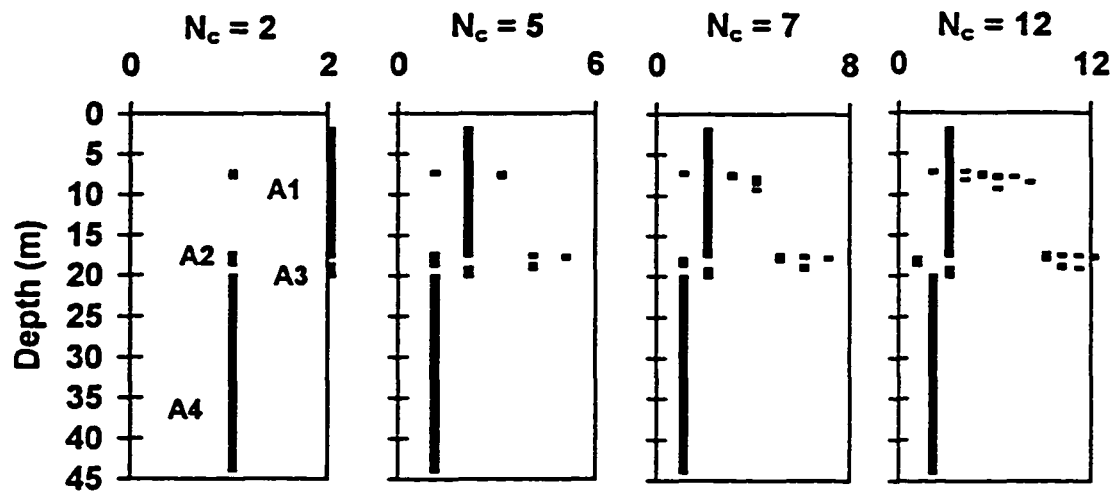
The piezocone classification chart suggests a thick clay or silty clay layer below 3.3 m and down to 45.0 m as shown in Fig. 6.35. The four primary layers identified by clustering are shown in Fig. 6.35. Clustering helps to objectively demarcate different zones which have different physical and/or mechanical properties within a clay deposit. Therefore, it is a valuable complementary tool to the classification chart for the proper delineation of a soil stratigraphy.

The cluster results are verified by water contents and plasticity measurements as shown in Fig. 6.36. The data are divided into two groups with a boundary at 18 m. The water contents and the liquid limits have an equal average equal to 59 percent from 2 m to 18 m, however, the plastic limits have a an average equal to 24 percent. The three measurements have constant averages with depth equal to 24 percent, 37 percent, and 16 percent, respectively, between the depths of 18 m and 45 m.

Also, the sensitivity readings taken using a fall cone test validate the clustering as shown in Fig. 6.37 which includes two groups of sensitivity data with a boundary at 16-m depth. The average sensitivity decreases from 5.8 in the upper clay to 2.2 in the lower clay. Therefore, clustering properly demarcated the subsurface profile at the site. However, neither the naked eye nor the empirical classification method is able to detect the drastic changes in the behavior of the two clay deposits.

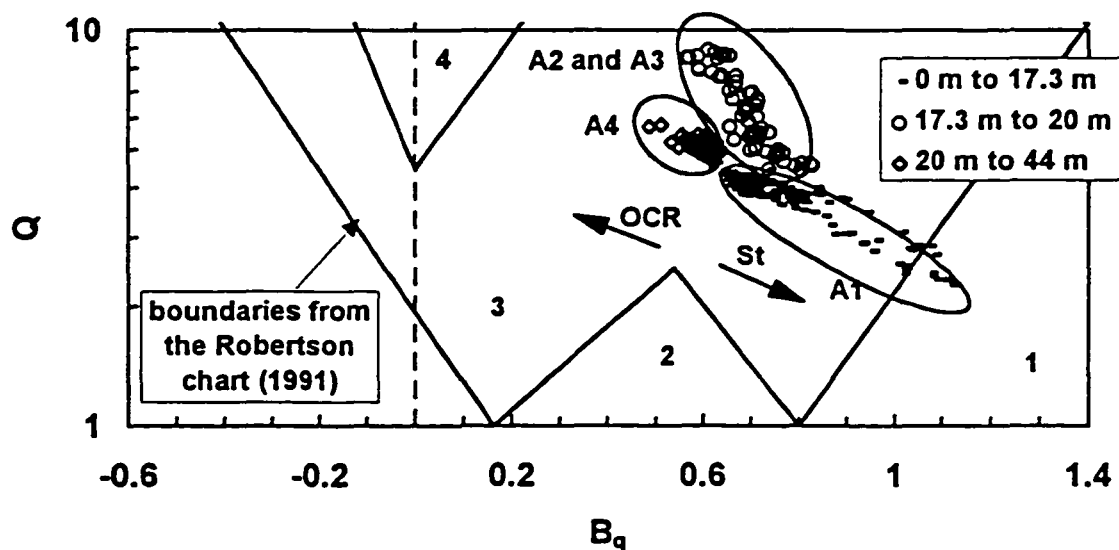
6.3. Discussion

Clustering was used to successfully demarcate eight clay sites where the piezocone data have subtle changes and the differences between different soil types and/or properties



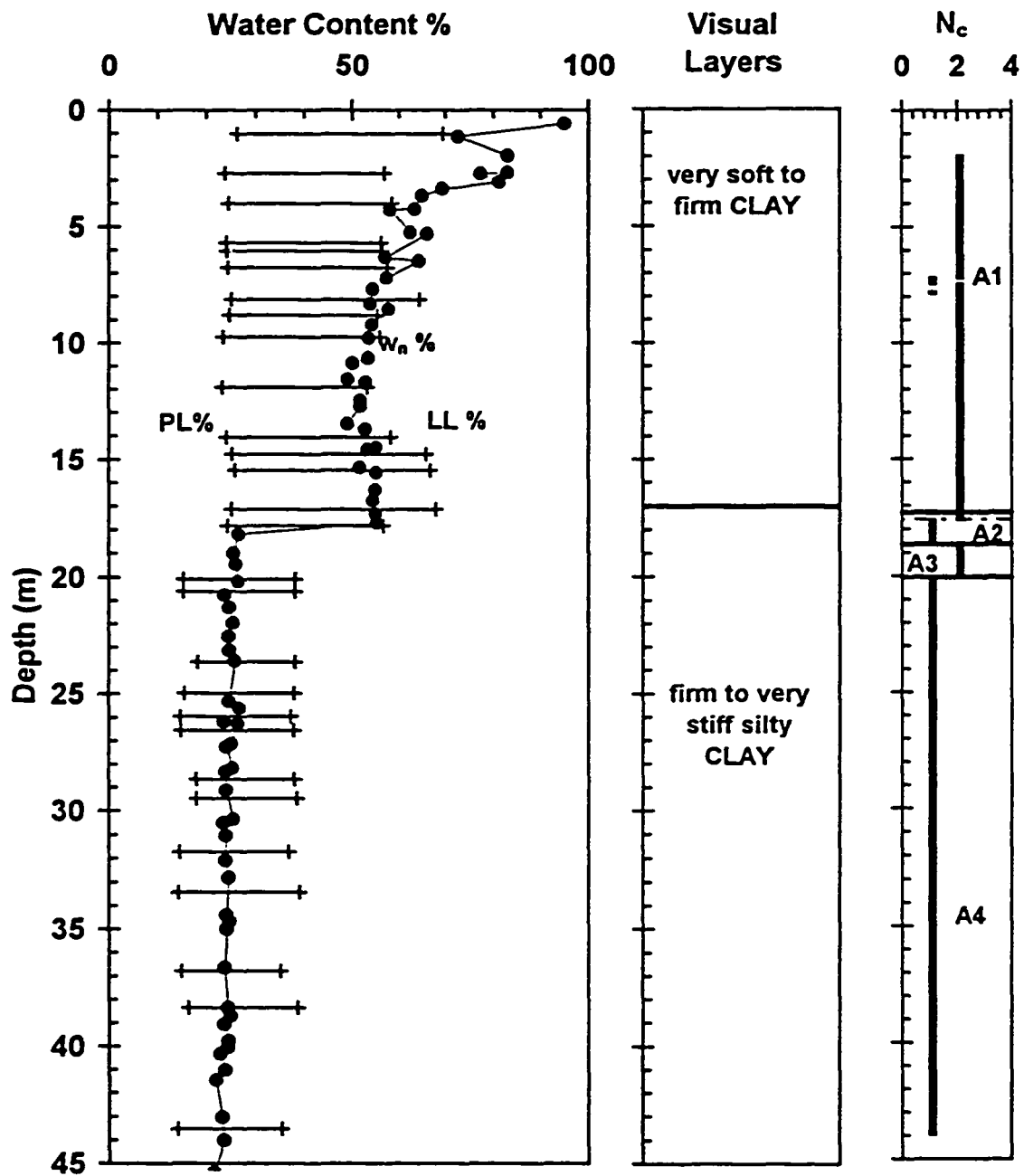
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Troll, Norway from Amundsen et al. (1985).

Figure 6.34. Cluster Analysis of Piezocone Data at Troll, North Sea.



1. Sensitive Fine Grained
 2. Organic Soils-Peats
 3. Clays: Clay to Silty Clay
 4. Silt Mixtures: Clayey Silt to Silty Clay

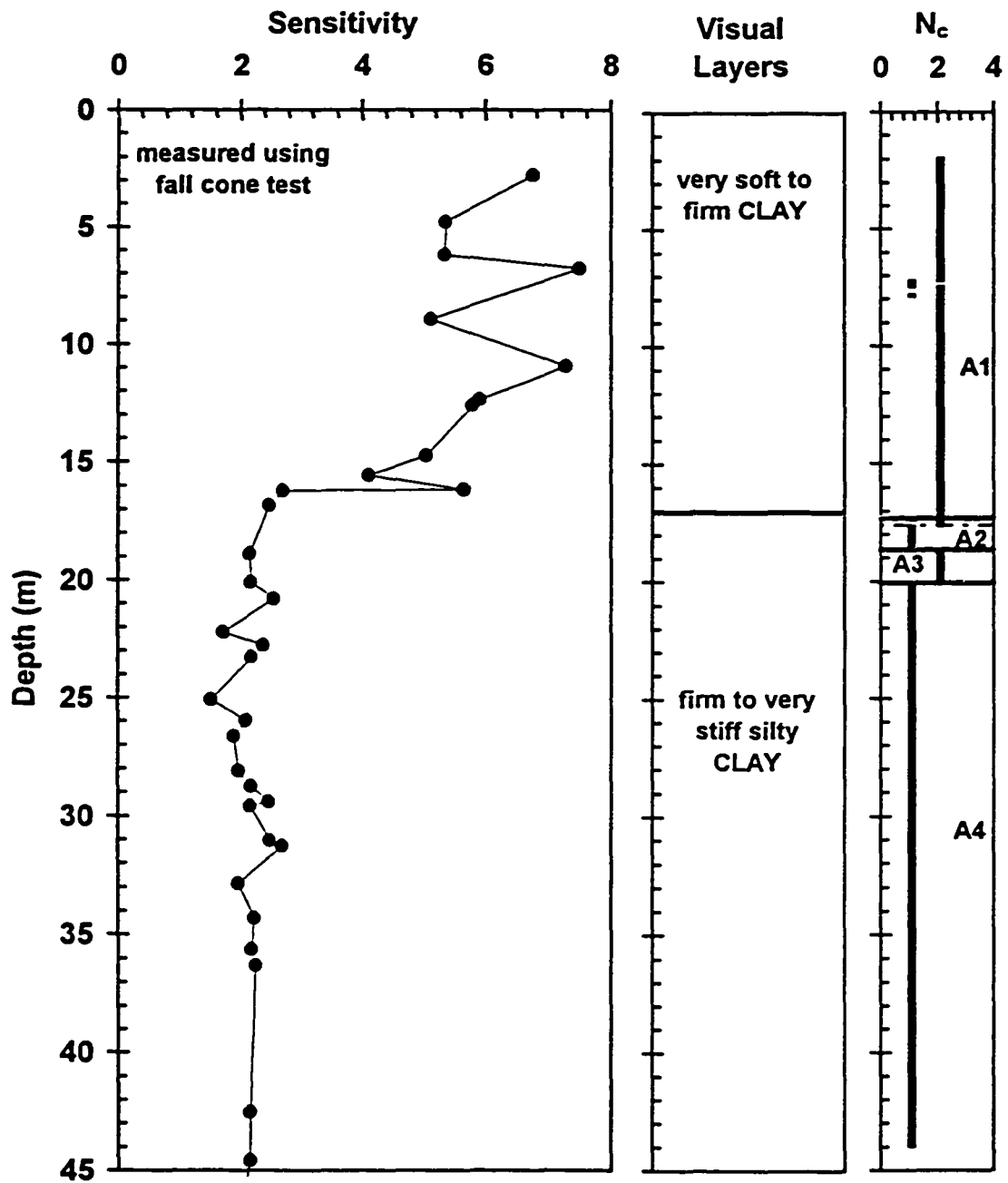
Figure 6.35. Clustering Results of Piezocone Data at Troll, North Sea on a Q - B_q Space (Data from Amundsen et al., 1985).



Note: Piezocone data at Troll, Norway from Amundsen et al. (1985).
 SCZ method using Q and B_q, cluster no. 2

—— Primary boundary
 - - - Transition boundary

Figure 6.36. Comparison Between Cluster Analysis, Visual Classification and Water Content at Troll, North Sea.



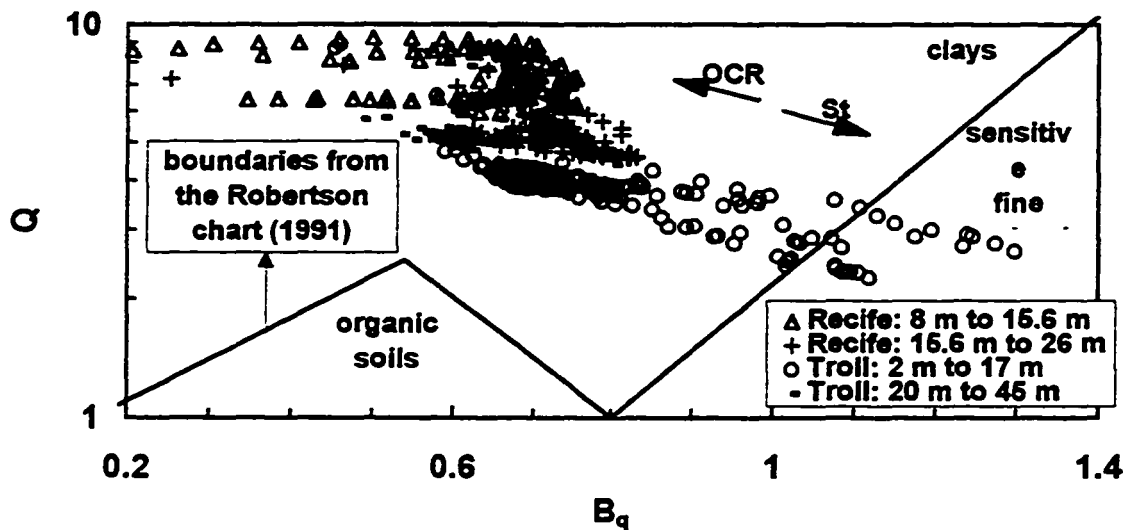
Note: Piezocone data at Troll, Norway from Amundsen et al. (1985).
 SCZ method using Q and Bq, cluster no. 2

Primary boundary
 Transition boundary

Figure 6.37. Comparison Between Cluster Analysis, Visual Classification and Sensitivity at Troll, North Sea.

can not be visually detected. An empirical CPT classification chart was chosen to indicate the soil stratigraphy at the eight sites. Clustering is a valuable supplementary tool to the classification chart to obtain an objective and reliable soil stratigraphy at each of the studied cases. However, it is not feasible to divide the clay zone into smaller clusters where each cluster represents a specific soil property such as plastic or lean clay in all geological conditions. For example, Fig. 6.38 shows the normalized parameters Q and B_q of piezocone data at Recife, Brazil, and Troll, North sea, tests sites. At both sites, the clay deposits consist of plastic clay over lean clay with boundaries at depths of 15.6 m and 17.0 m for Brazil (Coutinho and Oliveira, 1997) and Troll (Amundsen et al., 1985), respectively. There is an overlap between the lean clay at Brazil and the plastic clay at Troll in the Q and B_q diagram as depicted in Fig. 6.38. Cone data are functions of soil properties such as plasticity, compressibility and strength, and it is not easy to separate the effect of one soil property on the cone readings in a Q - B_q space.

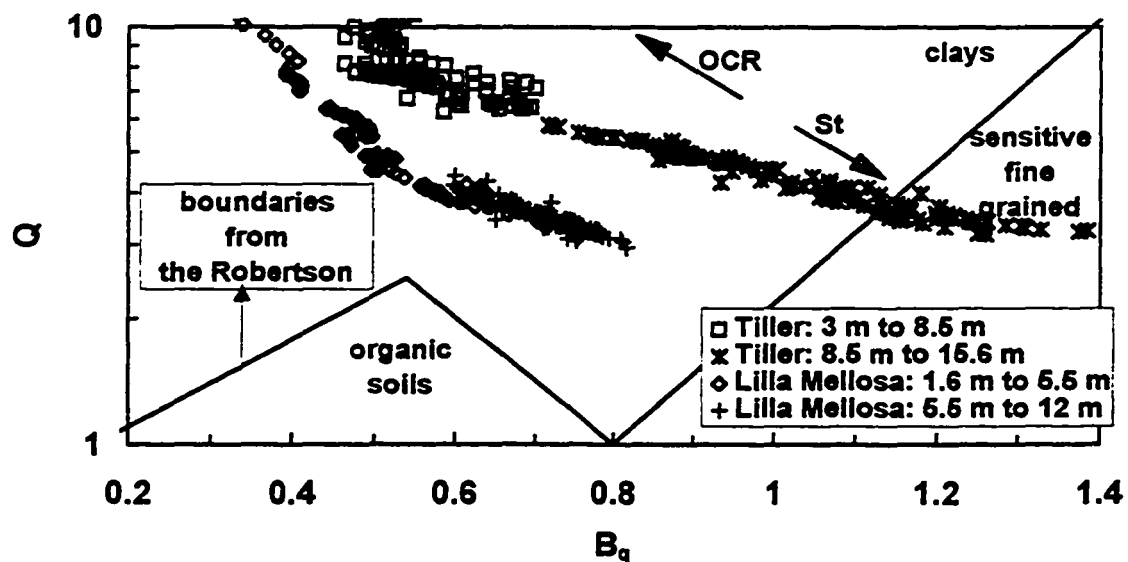
The proposed boundaries between sensitive fine grained, organic soils, and clays are subjective and can not represent some soil profiles. For example, the normalized parameters Q and B_q of piezocone data from Sandven (1990) at Tiller, Norway, are shown in Fig. 6.39 in which the boundary between clays and sensitive clays is suggested at 10.8 m. However, the upper clay and the lower quick clay are separated at a depth of 8.5 m based on reference data. Also, Q and B_q derived of piezocone data from Larsson and Mulabdić (1991) at Lilla Mellösa are shown in Fig. 6.39. The soil profile consists of organic clay over clay, however, the soil is determined as one clay cluster using the piezocone chart.



Note: Piezocone data at Recife, Brazil from Coutinho and Oliveira (1997).

Note: Piezocone data at Troll, North Sea from Amundsen et al. (1985).

Figure 6.38. Comparison between Plastic and Lean Clays at Recife and Troll Test Sites.



Note: Piezocone data at Tiller, Norway from Sandven (1990).

Note: Piezocone data at Lilla Mellosa, Sweden from Larsson and Mulabdic (1991).

Figure 6.39. Evaluation of the Boundaries between Different Fine Soils Defined in Robertson Classification Chart (1991).

6.4. Conclusions

Clustering using the single-cosine-zscore method is able to detect the subtle changes within the piezocone record and thus implicate major stratigraphic units within vertical soil profiles in some sites worldwide. The association between different strata is noted and also, soil lenses and transitions are indicated within a subsurface stratigraphy. Moreover, the proposed method is objective, repeatable and for all case studies, a representative soil profile was discovered at cluster number < 15 . The obtained statistical primary layers are verified by reference data of index and/or mechanical soil properties. However, simple naked-eye examination of the piezocone data is unable to detect these facets. Cluster results can serve as a supplement to available cone classification methods such as the Robertson chart (1991). First, the number of soil layers in a certain stratigraphy and the association between them are defined by cluster analysis, then a classification chart is used to identify the soil types. This results in a more reliable geostratigraphy.

CHAPTER 7

SUMMARY AND RECOMMENDATIONS

7.1. SUMMARY AND CONCLUSIONS

Site stratigraphy is a fundamental first step in a geotechnical investigation to define layering as well as the presence of any soil lenses, outliers, and inclusions. In a traditional site exploration program, borehole soundings are performed and soil samples are extracted to be tested in the laboratory. Extensive testing of high quality undisturbed samples is essential to quantify the physical and mechanical soil properties, although, laboratory methods are rather slow and expensive.

In-situ methods such as the piezocone tests are relatively fast, economical, and provide immediate results. Nowadays, piezocone tests are popular because they measure three different stresses (q_t = cone tip resistance, f_s = sleeve friction, and u_b = penetration pore pressure) in a single sounding and provide continuous information about a soil profile, usually, every 1 cm to 5 cm. Soil logging is a primary use of the piezocone data which are functions of both soil type and behavior. The soil profile is identified by visually examining the unprocessed piezocone data coupled with the use of empirical based cone classification charts. The former technique is subjective, not repeatable and dependent on the engineer experience, and the latter methods are empirical and rely on the characteristics of the database used in their development.

Statistical methods such as the intraclass correlation coefficient (ICC) and the generalized distance (D^2) have been used to objectively identify a subsurface profile using piezocone data. However, in some case studies, these methods improperly indicate the locations of soil boundaries and cannot identify possible subtleties existing in the piezocone data. Also, the association between different soil strata is not given.

An alternative statistical method termed *cluster analysis* is evaluated in this study for the purpose of soil stratigraphy based on piezocone data. Clustering is a powerful statistical method to objectively define similar groups of data in the soil profile, delineate different layer boundaries and transitions, and identify the lenses and outliers within a sublayer. Clustering techniques including hierarchical, optimizing-partitioning (k-means), density or mode-seeking, and clumping are evaluated in this study. Hierarchical clustering methods are selected as the best suitable means for analyzing the piezocone data because no preliminary estimation of the inherent groups within the analyzed data is needed, and no overlapping is permitted between identified clusters. For data grouping, six different hierarchical routines are discussed and the single link (nearest neighbor) method is selected to cluster piezocone data because it is mathematically stable which means that data outliers or errors do not affect the primary clusters of the analyzed data. Standardization is essential in the case of clustering piezocone data to reduce the predominant effect of q_t which has a larger value than u_b at the same depth, on the analysis. Seven standardization methods and eight similarity techniques are evaluated. A parametric study is performed at three sites to compare the 56 different clustering routines, and a single-cosine-zscore method is shown to be the best technique for properly delineating a geostratigraphy at a

minimum cluster number compared with other clustering methods. While the piezocone data are standardized using zscore technique, the similarity between pairs of data at different depths is evaluated using cosine measurement, and the data are clustered into correlated groups using the single link method.

A cluster interpretation criterion is developed based on cone data representing key sites, previous analytical and statistical studies of extensive piezocone data and cone results collected in chamber tests. A soil layer is defined if a cluster thickness (t) is ≥ 0.5 m and a soil lense, transition or outlier is detected if a cluster thickness is < 0.5 m. A primary layer is delineated if $t \geq 1$ m and a secondary layer is demarcated if $1 \text{ m} > t \geq 0.5$ m. A cluster number (N_c) is chosen to represent a soil stratigraphy if no primary layers are detected at higher cluster numbers. Cluster results are verified by comparing obtained statistical stratification with independent information of boreholes boundaries, and reference physical and mechanical properties of the soil at a specific site.

Two normalized cone parameters $Q = (q_t - \sigma_{vo}) / \sigma_{vo}'$ and $B_q = (u_b - u_o) / (q_t - \sigma_{vo})$ are chosen as input for clustering analysis. The Q and B_q were recommended for soil type and soil behavior interpretation based on piezocone data (Wroth, 1984; 1988; Robertson, 1991). The readings of the sleeve friction channel were excluded from the scope of this study because they are dependent on the cone-type and manufacture, and indicated a much higher variability compared with the records of the tip resistance and pore pressure. Cluster results obtained at Amherst, Massachusetts (data from this study) using Q and B_q are compared with those using different forms of piezocone data as follows: (1)

unprocessed data (q_t and u_b), (2) partially processed readings (q_t and the ratio u_b/q_t), and (3) the normalized cone parameters (Q , B_q and F). The statistical soil stratigraphy obtained using Q and B_q matches the borehole soil profile at Amherst, however, at the same cluster number, one or more primary soil boundary is not identified using other forms of the data.

Piezococone database of 25 case studies representing different soil types and geological conditions are compiled for this study. These sites are mainly clay deposits in which extracting undisturbed samples for laboratory studies is more feasible than in the case of sandy soils. Part of the data were collected by the author at three sites within the United States including Amherst, Massachusetts, Opelika, Alabama, and Penuelas, Puerto Rico, and other data were available in the literature and redigitized using AutoCad software. Cluster analysis using the single-cosine-zscore technique is able to properly detect the geostratigraphy at each of the 25 case studies. Cluster analysis is successfully applied to indicate changes in soil properties within a continuous clay or sand profile. It also indicates the differences between different soil types including sand, silt and clay. Soil lenses, seams, transitions are objectively delineated and the similarity of different clusters is demarcated. Clustering is performed for the studied cases up to a cluster number equal to 100 which means dividing a data set into 100 groups. The interpretation criterion is applied at each cluster number for each site and a statistical soil profile is chosen where there is no primary layer separates at higher cluster numbers. At all case studies, a representative soil stratigraphy is identified at cluster number $N_c < 15$, therefore, for a

preliminary investigation, clustering is recommended to be performed up to $N_c = 15$ in future studies.

Different factors affecting the cone measurements and available cone interpretation techniques including empirical and analytical methods are evaluated. The studied factors include data errors and outliers, cone size, data frequency, penetration porewater position, and spatial variability. Clustering is advantageous over other cone interpretation techniques for the purpose of soil stratigraphy because cluster results are independent of systematic errors or data outliers. In other statistical methods such as the autocorrelation function, data should be filtered first for a proper interpretation of piezocone data, however, some geological evidences which can be important for a geotechnical design are missed. Clustering is not affected by the scale effect using 10-cm² and 15-cm² cones and similar cluster results are obtained. Similar cluster results are obtained for a range of data frequency of 5 cm to 50 cm which suggests that cluster analysis can be applied for other laboratory and in-situ tests for which data are collected at larger intervals. Identical primary boundaries are obtained using both the penetration porewater pressure on the face (u_t) or behind the tip (u_b) at a soft clay deposit. However, in the case of heavily overconsolidated fissured clays, it is recommended to use the pore pressure at the face to properly identify a soil stratigraphy. A preliminary spatial cluster analysis is performed and cluster results are used to predict the association and continuity of different strata in two- and three-dimensions which is promising for future cluster applications in more than one dimension.

A major contribution of this research is that the cluster analysis is able to delineate subtleties in the piezocone profiles and thus discover drastic changes in soil types and/or properties in a soil stratigraphy which are not readily evident by visual examination of either the unprocessed data nor processed parameters, or by using available interpretation techniques. Several case studies involving sensitive clays, preconsolidated layered deposits, and cohesive materials of varying plasticity are used to illustrate the approach. Cluster analysis supplements available cone classification charts to obtain an objective and reliable preliminary soil stratification.

7.2. RECOMMENDATIONS FOR FUTURE RESEARCH

Cluster analysis has a great potential and use beyond that discussed in this research. In future studies, the author recommends the following aspects be considered.

- Clustering delineates different groups of piezocone data which are functions of soil type and behavior. Therefore cluster results are proposed to be directly correlated with soil properties. Laboratory calibration chamber tests using two or more soil layers are recommended. A parametric study can be performed for each soil property of interest such as overconsolidation ratio by changing one parameter at a time. For example, two clay layers can be prepared at two different levels of preloading. In this case, the soil type is the same and the initial conditions of the two layers before preloading are identical. Cone data are collected in both layers and a single-cosine-zscore cluster analysis is performed. Then, these steps are repeated several times and a comparison is performed between zscore records and different similarity

measurements of all tests. Several ranges of a certain soil property can be correlated with several ranges of zscore and/or similarity levels at which data are collected in a group. In a similar study the correlation between cluster results and relative density in sands can be evaluated.

- The effect of increasing the frequency and resolution on the analysis and the obtained statistical stratigraphy is proposed for a future study. Local details and variations within a soil layer can be detected and verification is required by examination of continuous samples in the vicinity of a cone penetration sounding.
- Cluster analysis is performed using Q and B_q to analyze the subtleties in piezocone data which were not obvious by visually examining the data or by plotting them in a Q and B_q space. Therefore alternative cone parameters should be explored for a better estimation of the soil behavior. For example, Houlsby (1988) recommended to use $(q_t - u_b)/\sigma'_{vo}$ for interpretation of soil type and soil behavior using piezocone data. Evaluation of this parameter and other data collected using a multi-sensor cone for identifying a soil profile is recommended for a future study. Figure 7.1 shows different readings that can be collected using a cone penetrometer with different sensors including tip resistance, pore pressure, resistivity, heat flow, fluorescence, and soil moisture content. Moreover, the sleeve friction readings can be accommodated in the analysis if there is enhancement in their repeatability and reliability. Different combinations of these readings are proposed to be studied for the purpose of soil stratigraphy and correlation with soil properties using cluster analysis.

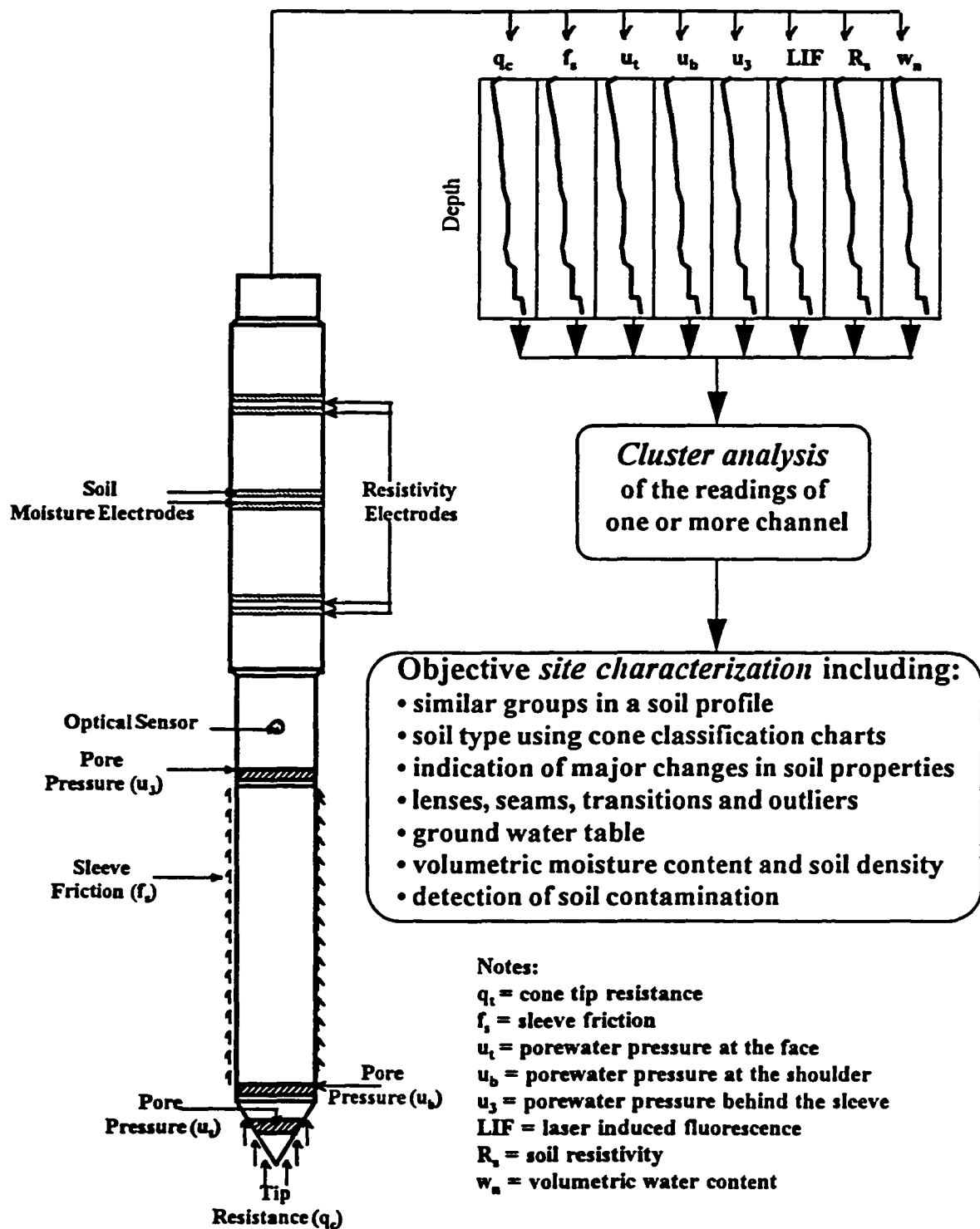


Figure 7.1. Schematic Diagram of a Geoenvironmental Cone.

- A study of the clustering validity is needed at more well-documented sand sites.

Clustering results could indicate, for instance, different levels of soil density which could be helpful for a preliminary estimation of liquefaction susceptibility of loose sands, or evaluation of a site remediation quality such as, in the case of vibratory compaction. Validity of two- and three-dimensional clustering analyses needs to be verified in more case studies where spatial cone penetration data are available for the purpose of evaluating the continuity of different soil strata and their association. The application of clustering analysis can be extended to other means of laboratory tests such as water contents, overconsolidation ratio, and shear strength, and in-situ testing such as dilatometer, vane shear, and standard penetration tests for the purposes of subsurface stratification and exploring the association between different soil types. A future study is needed to examine the validity of the single-cosine-zscore method for the analysis of these data.

- Several statistical and probability techniques have been applied to analyze laboratory and in-situ geotechnical parameters, including geostatistical (Hegazy et al., 1997), reliability (Christian et al., 1994), neural network (Chandler, 1996), and Monte Carlo simulation (Najjar and Basheer, 1996) methods for the purposes of identifying uncertainties in soil characterization. The application of the statistical methods in geosystems requires the division of the data into correlated homogeneous groups (soil layers) and then the inherent variability is determined within each layer. Therefore, clustering analysis can serve as a preliminary step to objectively delineate associated groups of the evaluated data. Clustering could supplement image analyses for

identification of local and volumetric void ratio, and soil fabric, including particles size, shape and angularity.

APPENDIX A

FIELD EXPERIENCES WITH CONE PENETRATION TESTING

A.1. Preface

Cone penetration tests (CPT) were performed by the author during the term of this study for the following purposes: (1) gain field experience in the conduct and collection of piezocone data; (2) obtain results needed for geostatistical analyses; and (3) assist others in the procurement of CPT data for their purposes. Table A.1 summarizes the locations, types, and numbers of cone tests, and objective mission associated with the work performed. Figure A.1 shows a mapping of the tested sites by the author during the term of this research study between the years of 1993 and 1997. This appendix outlines several of these field trips and testing programs to document the author's in-situ testing experiences in understanding the difficulties and uncertainties associated with cone penetration tests.

A.2. Georgia Tech Cone Penetration Systems

Cone tests were performed according to ASTM guide D-5778 using commercial Hogentogler electronic penetrometers (Robertson and Campanella, 1989), however, in two case studies, Fugro electrical cones (DeRuiter, 1982) and Davey electrical cones (Chen and Mayne, 1994) were used. There are currently five penetrometers located at

Table A.1. Summary of Cone Penetration Tests Performed by the Author During this Research Study.

Site location	No. of tests	Test-type	Average sounding depth (m)	Purpose of testing ⁽¹⁾	Tests performed for the benefit of:	Cone Data Presented in:
Atlanta, Georgia	3 10 3	CPT ⁽²⁾ PCPT SPCPT	11	<ul style="list-style-type: none"> Studying the scale effect of 10- and 15-cm² cones. Foundation design of pedestrian bridge at 10th Street. New civil engineering building at Georgia Tech. Field demonstration for graduate students. 	This study; Georgia Tech	This study and Internal reports from 1993 to 1997
Bagdad, Arizona	1 3 3	PCPT SPCPT PPDT	22	<ul style="list-style-type: none"> Checking the static and dynamic stability of mine tailings dam. 	AGRA Earth and Environmental, Phoenix	Mayne et al. (1994)
Penuelas, Puerto Rico	3 8 6	PCPT SPCPT PPDT	26	<ul style="list-style-type: none"> Evaluating soil improvement after surcharge loads. Checking the soil liquefaction potential at the site. 	Law Engineering, Houston; GeoCim, Inc., San Juan*	Mayne et al. (1995, 1997); Hegazy and Mayne (1996)
Richmond, Virginia	3 3	PCPT PCPT ⁽³⁾	12	<ul style="list-style-type: none"> Evaluation of soil properties for foundations design of a bridge. 	Virginia Geotechnical Services, P.C., Richmond	Hegazy and Mayne (1995)
Dunklin County, Missouri	3	SPCPT	12	<ul style="list-style-type: none"> Dynamic stability of existing and proposed highway bridges in New Madrid earthquake region. 	MDHT, Jefferson City**	Mayne et al. (1996)
Opelika, Alabama	2	SPCPT	12	<ul style="list-style-type: none"> Three-dimensional geostatistical analysis of CPT data in residual soils. Other research purposes. 	This study; Auburn University	Hegazy et al. (1997) and this study
Amherst, Massachusetts	15	SPCPT	14	<ul style="list-style-type: none"> Spatial analysis of SPCPT data. 	This study	This study

⁽¹⁾ In all case studies, cone data were also used for soil stratigraphy.

⁽²⁾ Data collected using 15-cm² Fugro electrical cone.

⁽³⁾ Piezocone data were collected using 10-cm² Davey electrical cone.

Note: Other tests were performed using 10-cm² electronic Hogentogler cones.

(*) Other companies were involved in this project including Dames & Moore, Black & Veatch, and ENRON.

(**) MDHT = Missouri Department of Highways and Transportation

CPT = cone penetration test; PCPT = piezocone penetration test; SPCPT = seismic piezocone penetration test; PPDT = pore pressure dissipation test

Georgia Tech which are operated using two different data acquisition systems, including: (1) a computer-based commercial Hogentogler system (Robertson and Campanella, 1984), and (2) a notebook-based system consisting of separate component elements (Chen and Mayne, 1994). Electrical power input necessary for both computer systems was obtained from an electrical power generator in the case of using a drill-rig for penetration, or from a battery in the case of using a cone-truck for penetration.

The two Hogentogler electronic penetrometers which were used to conduct most of the tests summarized in Table A.1 are 10-cm² cones and have two different capacities, namely, 50 kN and 100 kN. These two cones are pictured in Fig. A.2 along with other penetrometers, including a miniature 5-cm² cone, type-1 electrical 10-cm² Davey piezocone, type-2 electrical 10-cm² Davey piezocone, and triple-element electrical 15-cm² Fugro piezocone. Note that in case of electronic cones, the voltage responses of different cone channels are amplified inside the cone, however, in the case of electrical cones, the analog measurements of different cone channels are amplified at the ground surface after being transmitted through a cable to the data acquisition system.

Careful laboratory calibration of the load cells and transducers was performed by the author to obtain factors used to convert output voltages of each cone channel to measured soil stresses in the field, including: (1) cone tip resistance (q_c), (2) sleeve friction (f_s), and (3) pore pressure (u_b). Moreover, calibration was also performed to determine the net area ratio of penetrometers used to conduct the tests. Calibration setup of the cone tip resistance and sleeve friction load cells, and pore pressure transducer at Georgia Tech is discussed in detail by Chen and Mayne (1994).

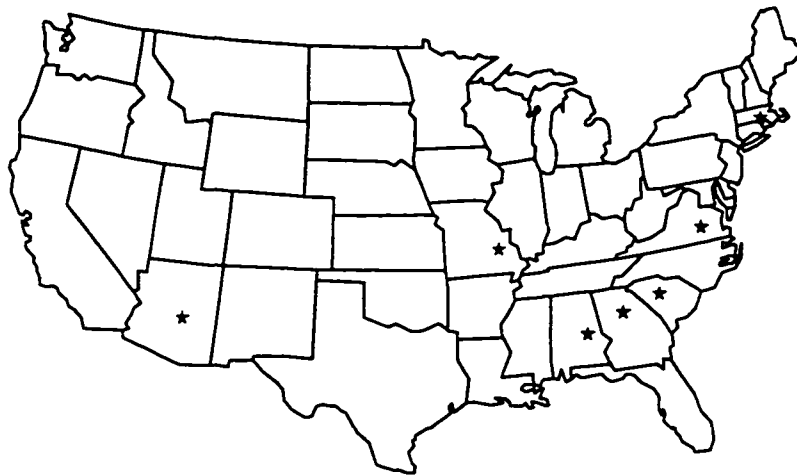


Figure A.1. Tested Sites by the Author During this Research Study.

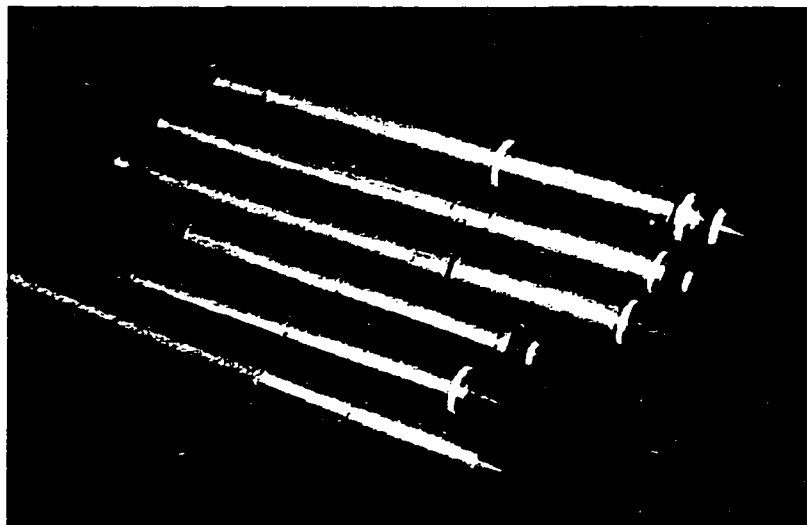


Figure A.2. Piezocones penetrometers at the Georgia Institute of Technology including from left to right: (1) a miniature 5-cm² electrical cone, (2) a Davey 10-cm² type-2 electrical piezocone , (3) a Davey 10-cm² type-1 electrical piezocone, (4) a Hogentogler 10-cm² type-2 electronic seismic piezocone, (5) a Hogentogler 10-cm² dual element electronic seismic piezocone, and (6) a Fugro 15-cm² triple element electrical piezocone.

Penetration depth was continuously measured by ultrasonic depth sensor using the notebook-based system (Chen and Mayne, 1994). In the case of using a Hogentogler system, a depth was measured when a proximity switch sent a signal to the data acquisition system (Hogentogler manual, 1995). The proximity switch faces metal pins mounted at equidistant increments on a cylindrical depth-wheeler which is in contact with the cone-rods during penetration.

Each of the Hogentogler penetrometers is supplemented by a uniaxial geophone oriented horizontally with a natural frequency of 28 Hz to allow seismic measurements of soil shear wave velocity, V_s (Robertson et al., 1986). A seismic shear wave source consisting of a rigid beam or a platform was used. It was weighted by rear wheels of a vehicle, or rear pads of a drill-rig or a cone-truck to ensure a good contact with the ground surface and minimize the loss of energy. A sledge hammer was used to strike the source to generate a horizontally polarized shear wave at the surface. Two different systems were used to record generated shear waves, including: (1) a four-channel Hewlett-Packard 54601A oscilloscope (Kates, 1996), and (2) a Hogentogler computer simulating an oscilloscope (Hogentogler manual, 1995). In the first system, a trigger geophone with a natural frequency of 14 Hz was used to determine the start time of the shear wave. In the Hogentogler system, an electrical step trigger of the type suggested by Hoar and Stokoe (1978) was used. When a hammer contacts a metal pad on the shear source, a generated voltage pulse travels to the simulated oscilloscope which starts to measure the travel time of the shear wave to the receiver. Using both systems, the travel time of the wave from the source to the receiver was determined. Pseudo-interval method

was used to calculate the shear wave velocity that is the difference in travel distance of two successive test depths divided by the difference in their travel times.

In the following sections, selected examples of cone testing experiences by the author are discussed.

A.3. Piedmont Residual Soils, Atlanta, Georgia

In Atlanta, cone testing was performed on different occasions from 1993 to 1997 by the author for a variety of purposes, including: (1) field demonstrations for undergraduate and graduate students, (2) studying scale effect of 10-cm² and 15-cm² penetrometers, (3) foundation design of a proposed pedestrian bridge crossing 10th Street, and (4) providing supplemented information for the new civil engineering Sustainable Education Building on Georgia Tech campus. The Atlanta area is underlain by residual silts and sands of the Piedmont Geology. Figure A.3 shows the Fugro cone truck in operation on the campus. Here, three CPTs were advanced for a new civil engineering office. A representative cone sounding is shown in Fig. A.4 including q_c and f_s profiles. The groundwater table is at 7-m depth.



Figure A.3. Fugro Cone Truck at the Proposed Area of a New Civil Engineering Building on Georgia Tech Campus, Atlanta, Georgia.

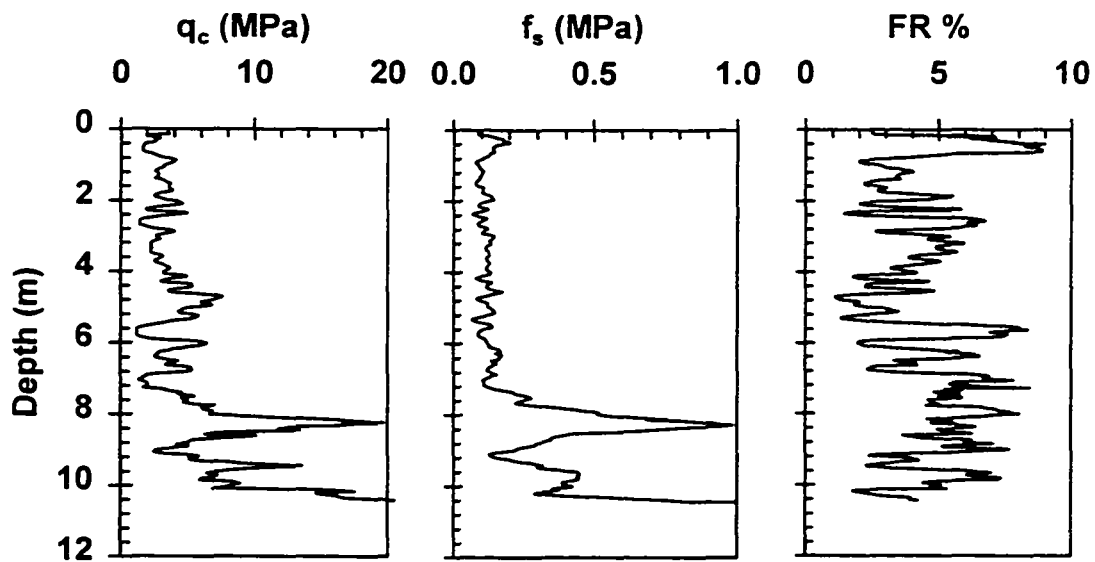


Figure A.4. A Representative Cone Sounding in Piedmont Residual Sandy Silts at Georgia Tech Campus, Atlanta, Georgia.

A.4. Bagdad Mine Tailings, Arizona

Seismic piezocone tests were performed at an existing copper mine tailings dam in western Arizona. The test locations were made accessible by constructing a 3.6-m thick and 180-m long sand fill ramp over a geotextile layer laid on the top of the tailings. The tests were conducted for soil characterization, estimation of soil properties, and the evaluation of dynamic stability of the tailings dam during possible earthquake events. The tailings consist of about 45 percent cycloned sands and 55 percent silt-sized particles, and their specific gravity is between 2.66 and 2.72 (Vidic et al., 1995). The cone was pushed using a CME-55 drill rig as pictured in Fig. A.5. A representative piezocone sounding is shown in Fig. A.6. The three channels (q_c , f_s , and u_b) indicate the variable nature of the tailings deposit with a quick variation of the soil type as noticed from the alternating peaks and troughs in the three piezocone channels.

A.5. Amherst Varved Clay, Massachusetts

The Amherst test site is one of the five national geotechnical experimental sites (NGES) in the United States and was tested by the author during the summer of 1996. The soil profile consists of 2 m of clay fill, 2 m of clay crust, 1 m of brown-gray varved silty clay over gray varved clay (Lally 1993). The ground water table is almost at 1.9 m. The purpose of testing was to perform spatial geostatistical analysis on fifteen seismic piezocone soundings arranged as shown in Fig. A.7. The cone was pushed in the ground using the small portable hydraulic penetration rig loaned by Prof. A. J. Lutenegeger and pictured in Fig. A.8. The results of the seismic piezocone soundings are summarized in



Figure A.5. Pushing a Piezocone in a Mine Tailings Dam, Bagdad, Arizona Using a CME-55 Drill Rig.

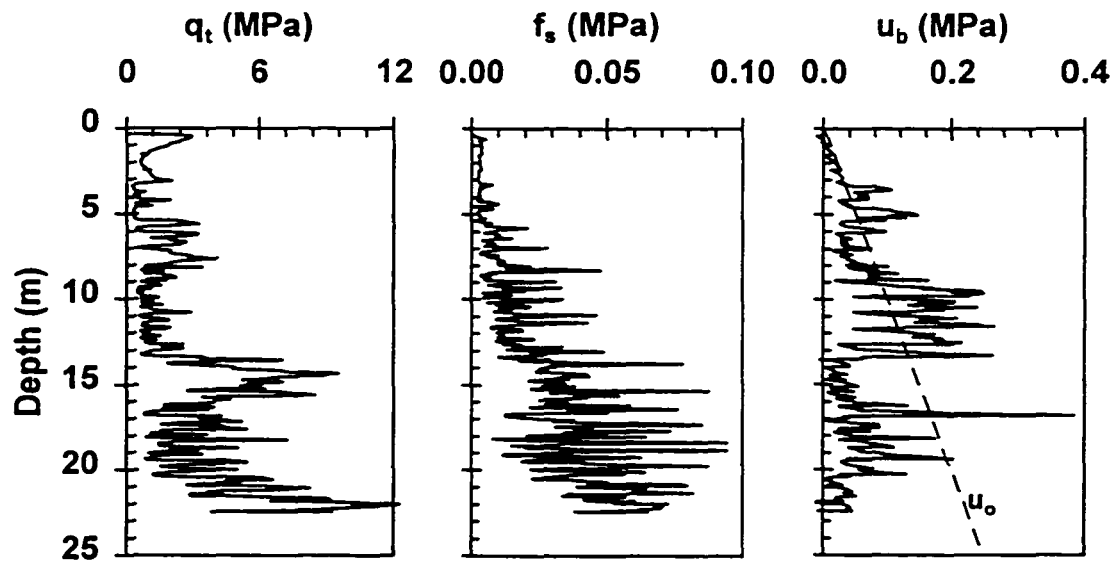


Figure A.6. A Representative Piezocone Sounding in Copper Mine Tailings at Bagdad, Arizona.

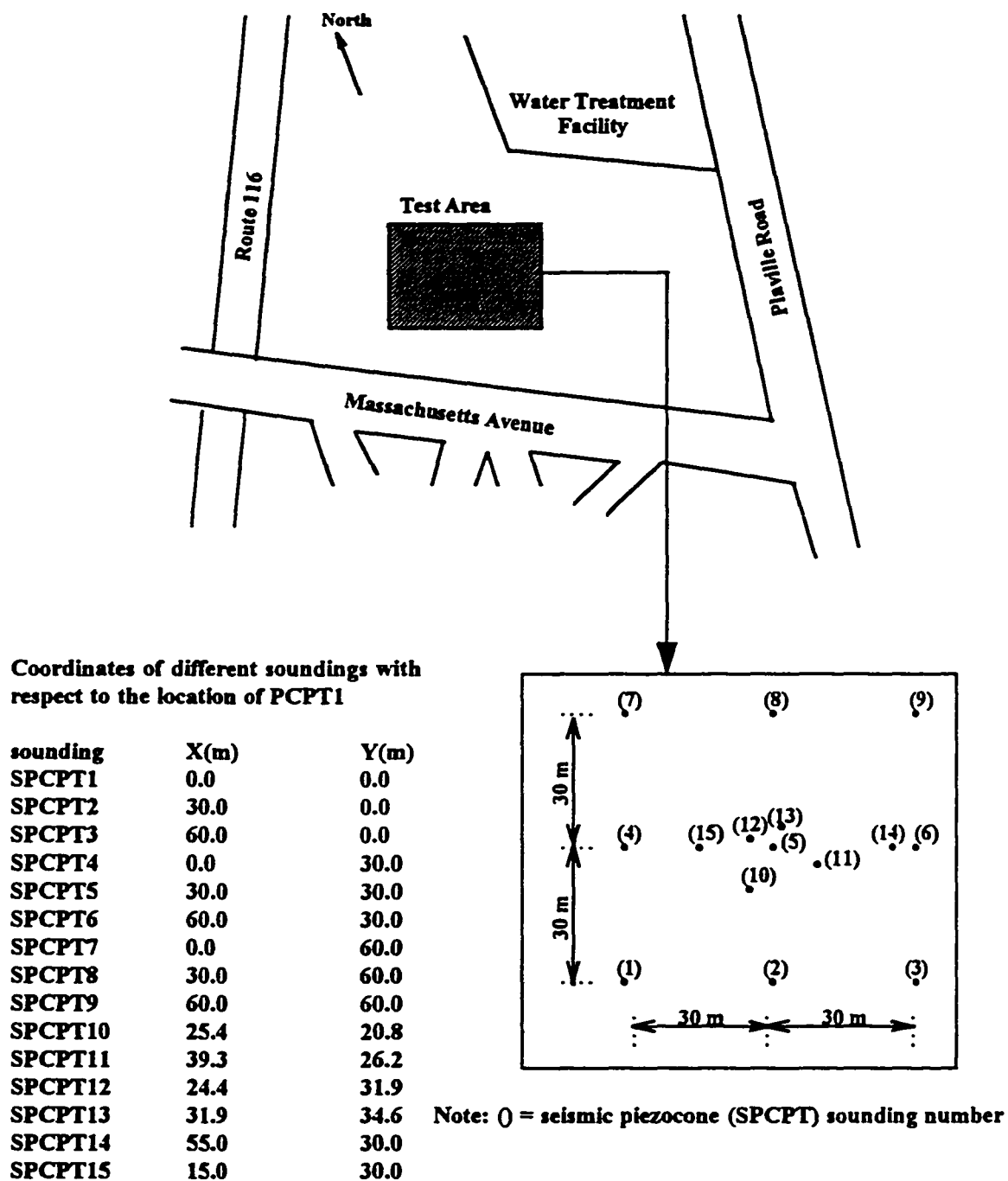


Figure A.7. Seismic Piezocone Test Locations at Amherst, Massachusetts.

Fig. A.9. These indicate a high variability in the upper 4 m of clay fill and desiccated clay crust. The similarity of the seismic piezocone data in the lower depths imply a spatial homogeneity of the normally-consolidated varved clay deposit. In many instances, there was a delay of the response of the pore pressure transducer below the groundwater table due to the possibility of desaturating the filter as the cone penetrated the upper vadose zone. Note also the significant dissipation of the pore pressure at the depths of rod-additions.

A.6. Dunklin County Alluvial Deposit, Missouri

Seismic piezocone tests were performed at three sites in Dunklin County, southeastern Missouri near the 1811-1812 New Madrid earthquake region. The tests were used to evaluate the risk susceptibility of roadway bridges by the Missouri Department of Highway & Transportation (MDHT) during future possible earthquakes. The piezocone tests were conducted with a CME-850 track-mounted drill rig shown in Fig. A.10. A representative type-1 seismic piezocone sounding at the Route 164 site is shown in Fig. A.11, which includes the vertical profiles of q_c , f_t , u_t , and V_s . The soil profile at this site predominantly consists of loose to dense sands and the groundwater table is at a depth of 1.5 meter.



Figure A.8. Pushing a Piezocone in a Varved Clay Deposit, Amherst Massachusetts, Using a Portable Hydraulic Rig.

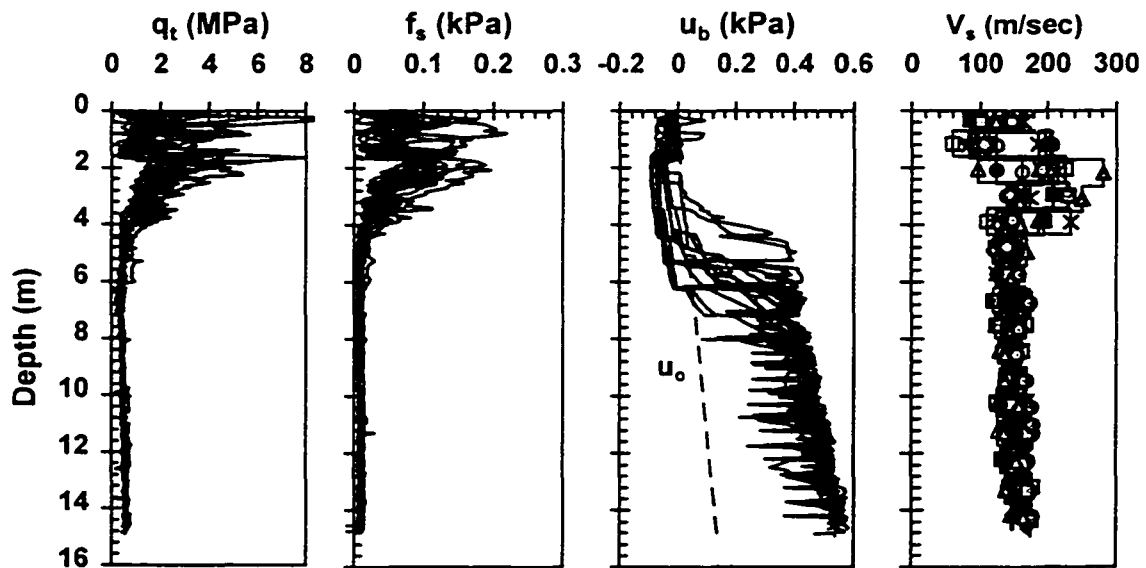


Figure A.9. Summary Results of 15 Seismic Piezocone Soundings Performed at Amherst, Massachusetts.

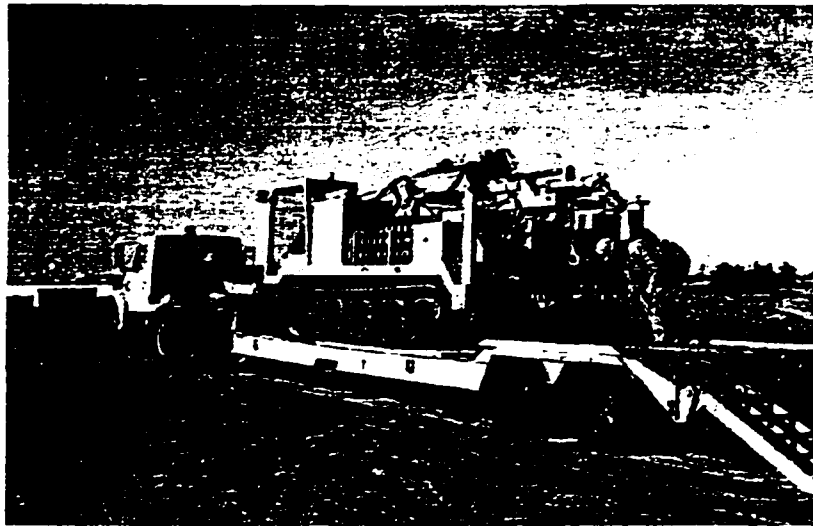


Figure A.10. A CME-85 Track-Mounted Drill Rig Used to Push Piezocones in a Sandy Deposit at Route 164, Dunklin County, Missouri.

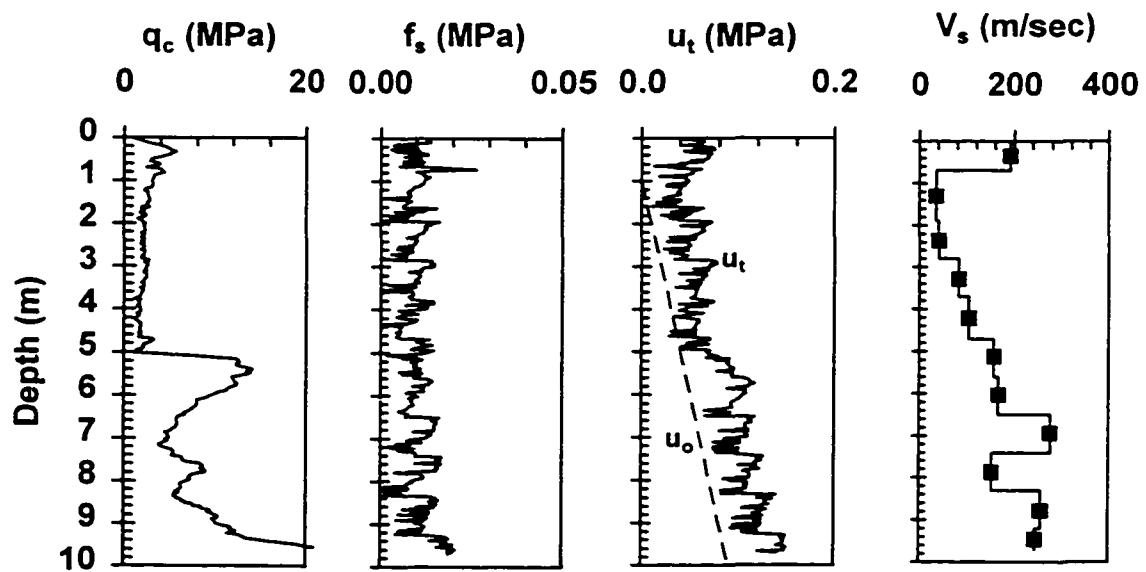


Figure A.11. A Representative Piezocone Sounding in a Sandy Deposit at Route 164, Dunklin County, Missouri.

A.7. Opelika Piedmont Soils, Alabama

Seismic piezocone data were conducted at Opelika, Alabama for the purpose of performing spatial geostatistical analyses (Hegazy et al. 1997), and to supplement other research programs. The site is located in the southern Piedmont province and the residual soils were derived from the in-place weathering of schist and gneiss bedrock. On the average, the soil consists of 50 percent silty fines and 50 percent sands. The groundwater table is located at a depth of 3 m. The penetrometer was pushed in the ground using a Hogentogler-type cone truck operated by Williams and Associates of Tampa, Florida as shown in Fig. A.12. A summary of two seismic piezocone soundings is shown in Fig. A.13 and indicates good agreement between the records of q_t , f_s , u_b and V_s of both tests C41 and C42. Note the negative penetration porewater pressures throughout the sounding depths, well below the water table.

A.8. Summary

Numerous cone soundings were performed by the author at different test sites in a variety of soil types and conditions for research purposes and assisting others on geotechnical projects. Testing was performed using friction cones, piezocones, and seismic cones. The gained experiences during this extensive testing program helped the author to evaluate some of the uncertainties involved in the interpretation of the cone data, specifically, the subjectivity of estimated soil stratigraphies based on cone data.



Figure A.12. Hogentogler Cone Truck During Seismic Piezocone Operation in Piedmont Residual Soils at Opelika, Alabama.

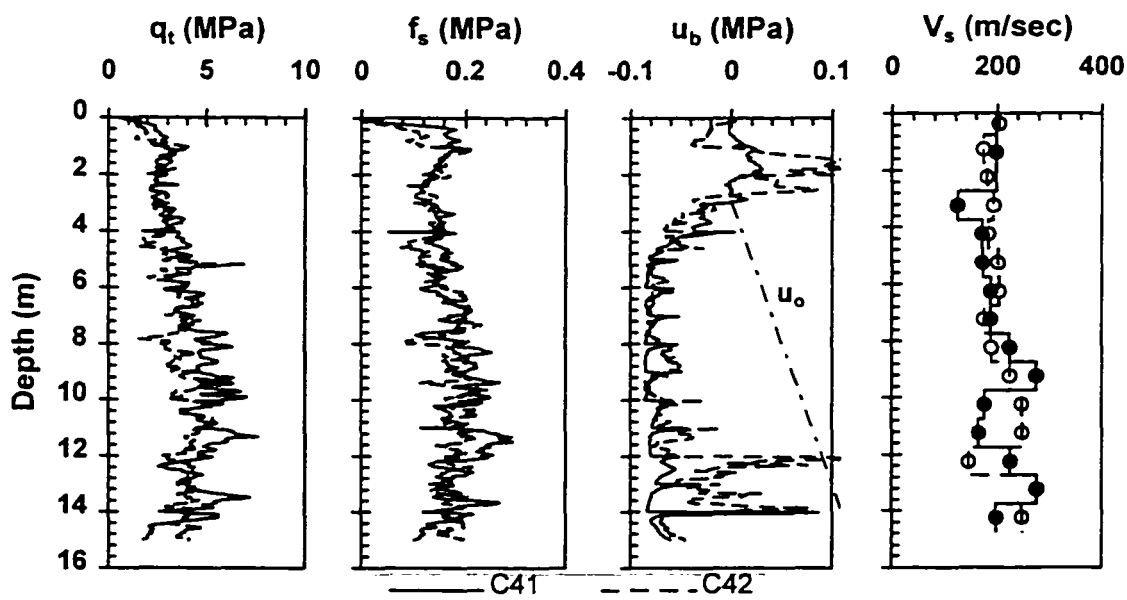


Figure A.13. Two Seismic Piezocone Soundings in Piedmont Residual Soils at Opelika, Alabama.

APPENDIX B

CLUSTERING OF PIEZOCONE DATA

B.1. Synopsis

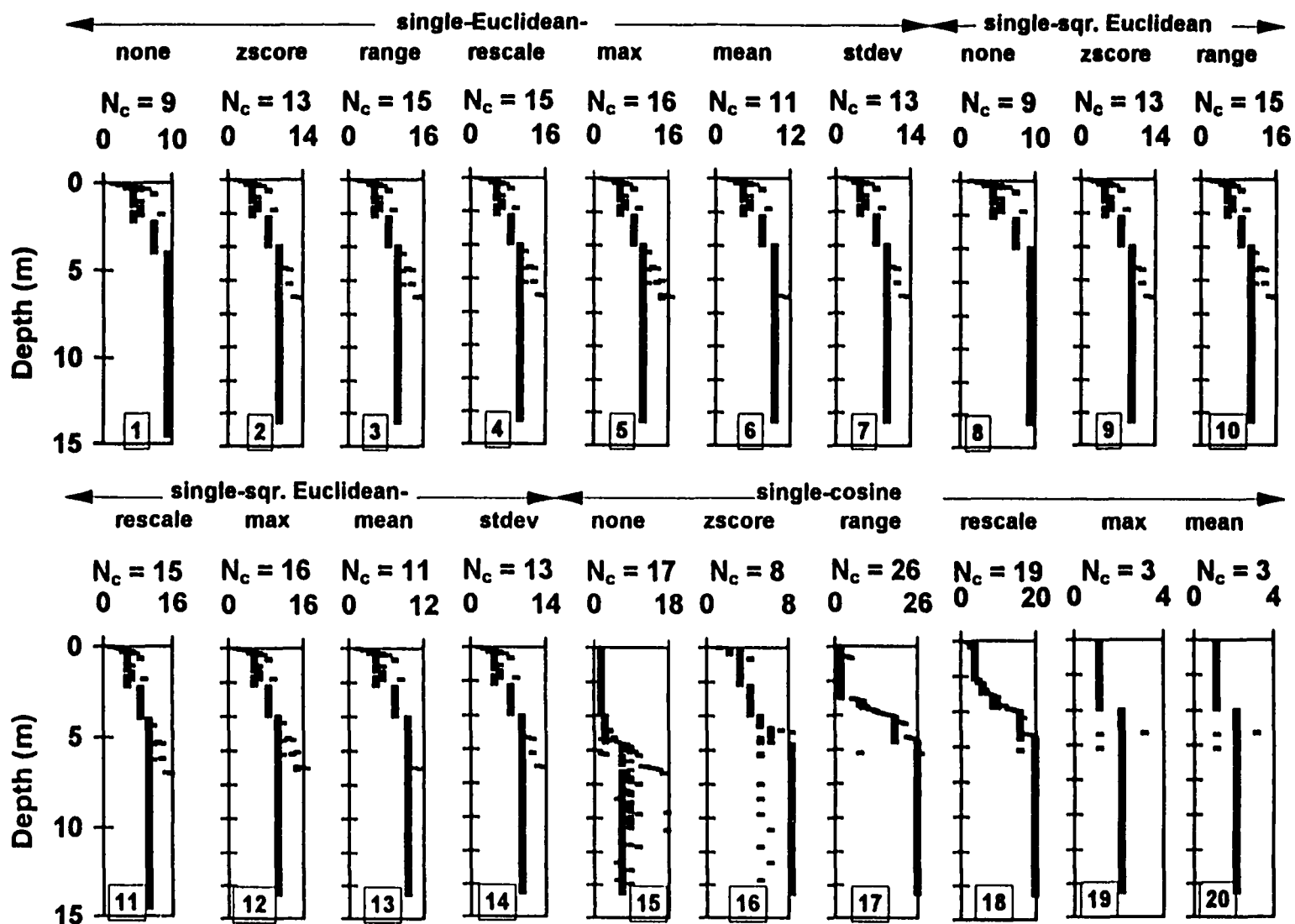
In this section, a parametric study is performed to select an appropriate method for clustering piezocone data for the purpose of delineating a soil stratigraphy. There are 7 methods to standardize piezocone data (see Table 4.2), and 8 methods to measure the similarity between two vectors of them measured at two different depths (see Table 4.3). The total available methods of standardization, and measuring the resemblance of piezocone data are 56. While, for the purpose of dividing the data into correlated groups, the single (nearest neighbor) method has been recommended in this study.

Cluster analysis is applied 56 times for piezocone data of one sounding at each of the following three sites: (1) Amherst, Massachusetts (piezocone data from this study), (2) McDonald's Farm, British Columbia (data from Robertson, 1982), (3) Fort Road, Singapore, (data from Chang, 1991). Clustering is performed 56 times for data of one piezocone sounding collected at each site. Data are divided into 100 groups for each analysis and a subsurface stratigraphy is interpreted using the proposed criterion discussed in Chapter 4. The results of different clustering methods are compared at each site. A better clustering is chosen where a soil stratigraphy including location of soil boundaries and number of soil types is properly retrieved at a minimum cluster number (N_c).

The subsurface stratigraphy at the three sites can easily be detected by inspecting the piezocone profiles by eye. The soil stratigraphy at Amherst includes 2 m of fill, 2 m of clay crust, 1 m of silty clay overlying a soft varved clay. The profile at McDonald's Farm contains 2 m of soft clay, 11 m of sand, 2 m of silty sand and 15 m of soft clayey silt. The soil at Fort Road consists of a deposit of a soft marine clay with an intermediate silty clay layer of a thickness equal to 4 m.

B.2. Amherst Test Site

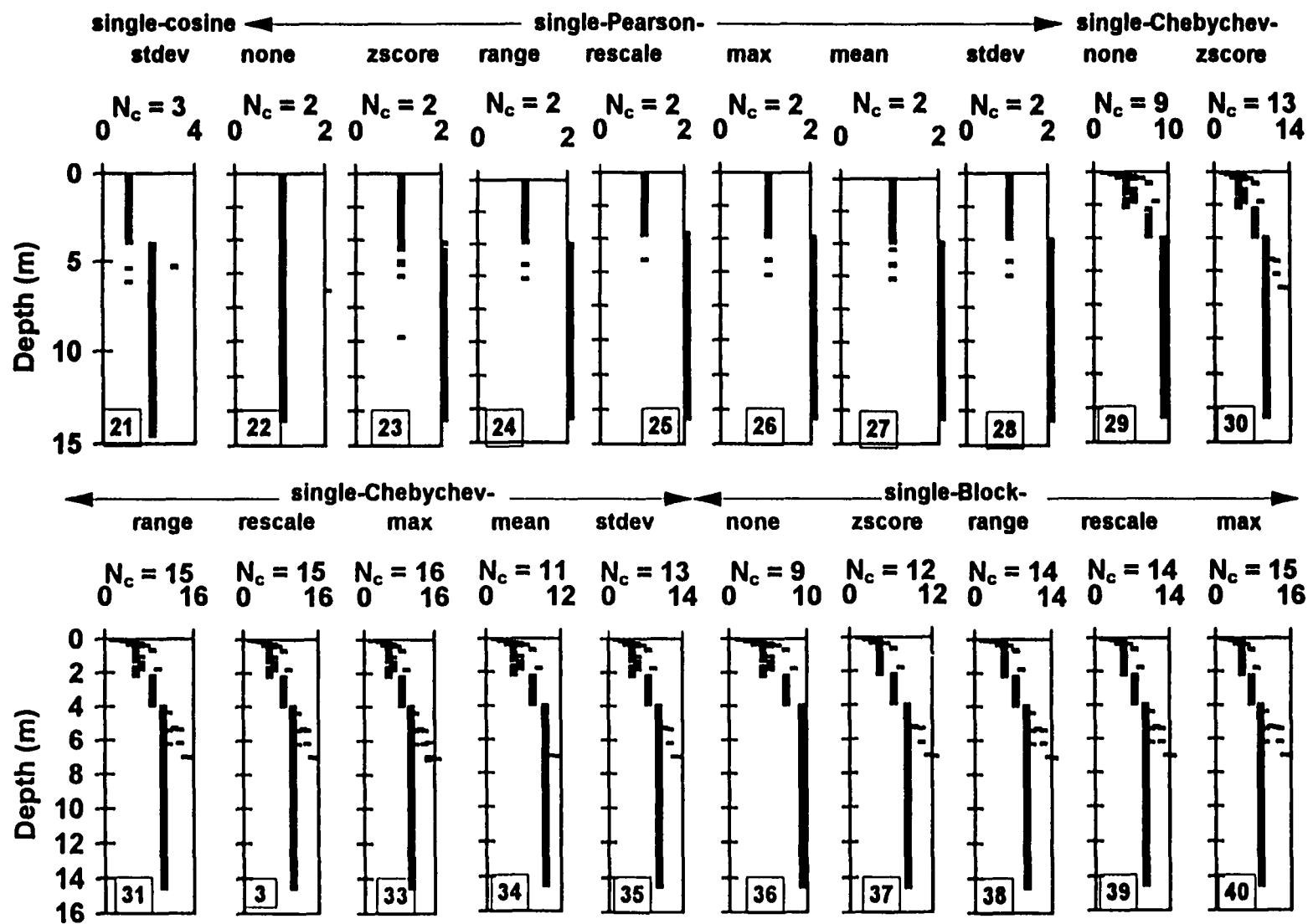
Clustering is performed for piezocone data of a PCPT1 sounding at Amherst. The vertical soil profile consists of a 2-m thick silty clay fill over a 2-m thick desiccated sandy clay crust underlain by a soft brown-gray varved clay down to 5 m and a deep deposit of normally-consolidated gray varved clay to the termination depth of the piezocone sounding at a depth of 14.5 meter (Lally, 1993). Clustering results of 56 different analyses are shown in Fig. B.1 and the statistical layers are compared with the given soil profile by Lally, 1993. A summary of clustering procedures are given in Table B.1 in which a proper detection of the four layers in the profile is denoted by "Y" and a poor estimation of the stratigraphy is denoted by "N". The single-cosine-zscore (labeled number 16) was the only method able to properly detect the soil profile at a minimum cluster number $N_c = 8$. Detailed single-cosine-zscore analysis of PCPT1 was previously given in Chapter 5.



Note: Piezocone data at Amherst, Massachussets (this study)

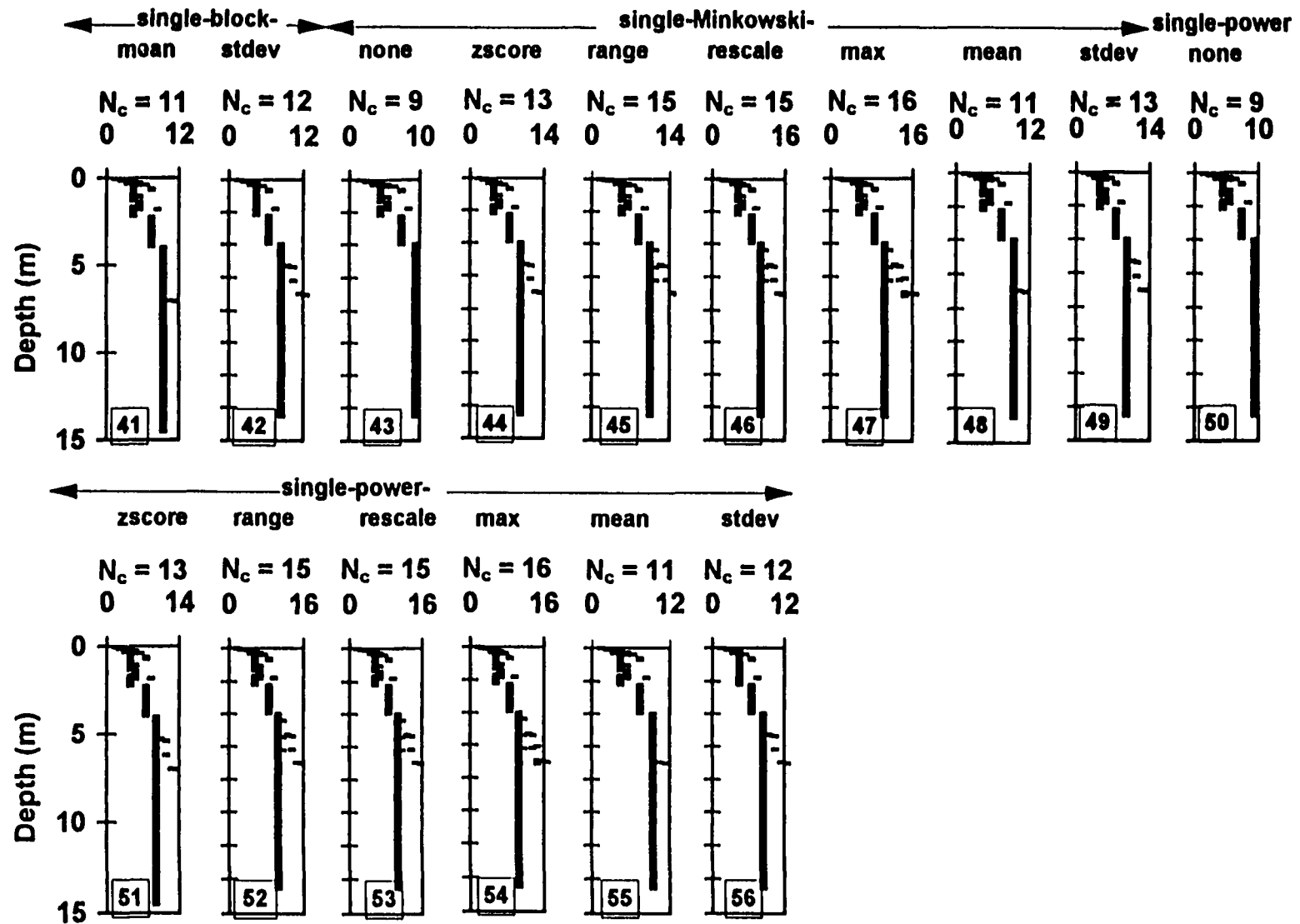
[no.] = cluster method no. (see Table C. 1)

Figure B.1a. Cluster Results of Piezocone Data at Amherst Using Different Cluster Methods.



Note: Piezocone data at Amherst, Massachusetts (this study) [no.] = cluster method no. (see Table C. 1)

Figure B.1b. Cluster Results of Piezocone Data at Amherst Using Different Cluster Methods.



Note: Piezocone data at Amherst, Massachusetts (this study)

[no.] = cluster method no. (see Table C. 1)

Figure B. 1c. Cluster Results of Piezocone Data at Amherst Using Different Cluster Methods.

Table B.1. Summary of 56 Different Cluster Analyses Performed on Piezocone Data (PCPT1) from Amherst, Massachusetts (Data from this Study).

Method	Chosen Cluster	Results	Method	Chosen Cluster	Results
(1) None-Euclidean	9	N	(29) None-Chebychev	9	N
(2) Zscore-Euclidean	13	N	(30) Zscore-Chebychev	13	N
(3) Range-Euclidean	15	N	(31) Range-Chebychev	15	N
(4) Rescale-Euclidean	15	N	(32) Rescale-Chebychev	15	N
(5) Max-Euclidean	16	N	(33) Max-Chebychev	16	N
(6) Mean-Euclidean	11	N	(34) Mean-Chebychev	11	N
(7) Stdev-Euclidean	13	N	(35) Stdev-Chebychev	13	N
(8) None-sqr. Euclidean	9	N	(36) None-Block	9	N
(9) Zscore-sqr. Euclidean	13	N	(37) Zscore-Block	12	N
(10) Range-sqr. Euclidean	15	N	(38) Range-Block	14	N
(11) Rescale-sqr. Euclidean	15	N	(39) Rescale-Block	14	N
(12) Max-sqr. Euclidean	16	N	(40) Max-Block	15	N
(13) Mean-sqr. Euclidean	11	N	(41) Mean-Block	11	N
(14) Stdev-sqr. Euclidean	13	N	(42) Stdev-Block	12	N
(15) None-Cosine	17	N	(43) None-Minkowski*	9	N
(16) Zscore-Cosine	8	Y	(44) Zscore-Minkowski*	13	N
(17) Range-Cosine	26	N	(45) Range-Minkowski*	15	N
(18) Rescale-Cosine	19	N	(46) Rescale-Minkowski*	15	N
(19) Max-Cosine	3	N	(47) Max-Minkowski*	16	N
(20) Mean-Cosine	3	N	(48) Mean-Minkowski*	11	N
(21) Stdev-Cosine	3	N	(49) Stdev-Minkowski*	13	N
(22) None-Pearson	2	N	(50) None-Power**	9	N
(23) Zscore-Pearson	2	N	(51) Zscore-Power**	13	N
(24) Range-Pearson	2	N	(52) Range-Power**	15	N
(25) Rescale-Pearson	2	N	(53) Rescale-Power**	15	N
(26) Max-Pearson	2	N	(54) Max-Power**	16	N
(27) Mean-Pearson	2	N	(55) Mean-Power**	11	N
(28) Stdev-Pearson	2	N	(56) Stdev-Power**	12	N

Legend:

() = number of cluster analysis

N = one or more soil boundary is missed

(*) power $p = 2$ and root $r = 1/2$

Notes:

Stdev = standard deviation

Y = soil profile is properly detected

(**) power $p = 1$ and root $r = 1/2$

Standardization method⁽¹⁾

None (raw data)	Max
Zscore	Mean
Range	Stdev
Rescale	

Similarity measurement⁽²⁾

Euclidean	Chebychev
Squared Euclidean	Block (Manhattan Distance)
Cosine	Minkowski
Pearson	Power (p,r)

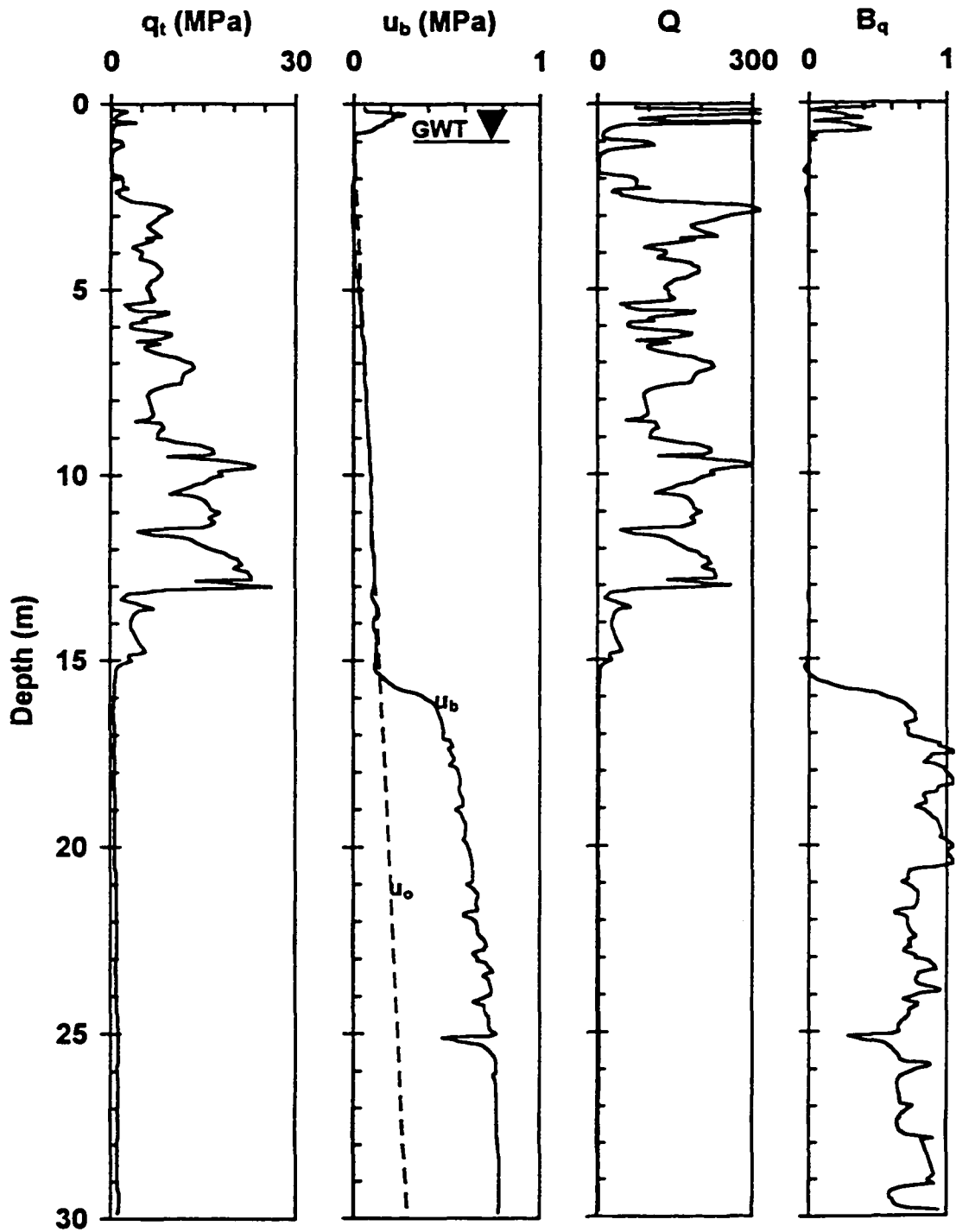
⁽¹⁾ Definitions of standardization methods were previously given in Table 3.2.

⁽²⁾ Definitions of similarity measurements were previously given in Table 3.3.

B.3. McDonald's Farm Test Site

A representative piezocone sounding (Robertson, 1982) at McDonald's Farm is shown in Fig. B.2. In addition to the two raw readings (q_t and u_b), Fig. B.2 also includes the derived normalized parameters Q and B_q . Looking at the piezocone profiles, four soil layers were delineated with boundaries at depths of 2 m, 13 m and 15 m. Similar results were obtained using the Robertson's classification chart (1991) as shown in Fig. B.3. The results of 56 clustering techniques are grouped in Fig. B.4. A summary of the trial clustering methods and their evaluations is given in Table B.2. The single-cosine-zscore (SCZ) method properly indicated the subsurface stratification including four soil layers at the smallest cluster number $N_c = 14$. A detailed study of clustering using SCZ method is discussed herein.

Clustering was performed up to $N_c = 100$ and the correlation coefficient (ρ_c) between consecutive clusters is shown in Fig. B.5 up to $N_c = 50$. Cluster results are examined at the peaks of ρ_c and are shown on Fig. B.6a between $N_c = 2$ and $N_c = 35$. At $N_c = 4$, the data are divided into two primary groups at a depth of 15.7 m. Then at larger cluster numbers, a transition layer appears between the two layers and some data points separate from the two cluster indicating non homogeneity. At $N_c = 14$, two statistical layers with thickness (t) > 1 m separate from the upper primary group. At larger clusters, no new primary clusters ($t > 1$ m) appear as shown in Fig. B.6b, therefore, $N_c = 14$ is



**Figure B.2. Piezocone Sounding at McDonald's Farm, British Columbia
(Data from Robertson, 1982).**

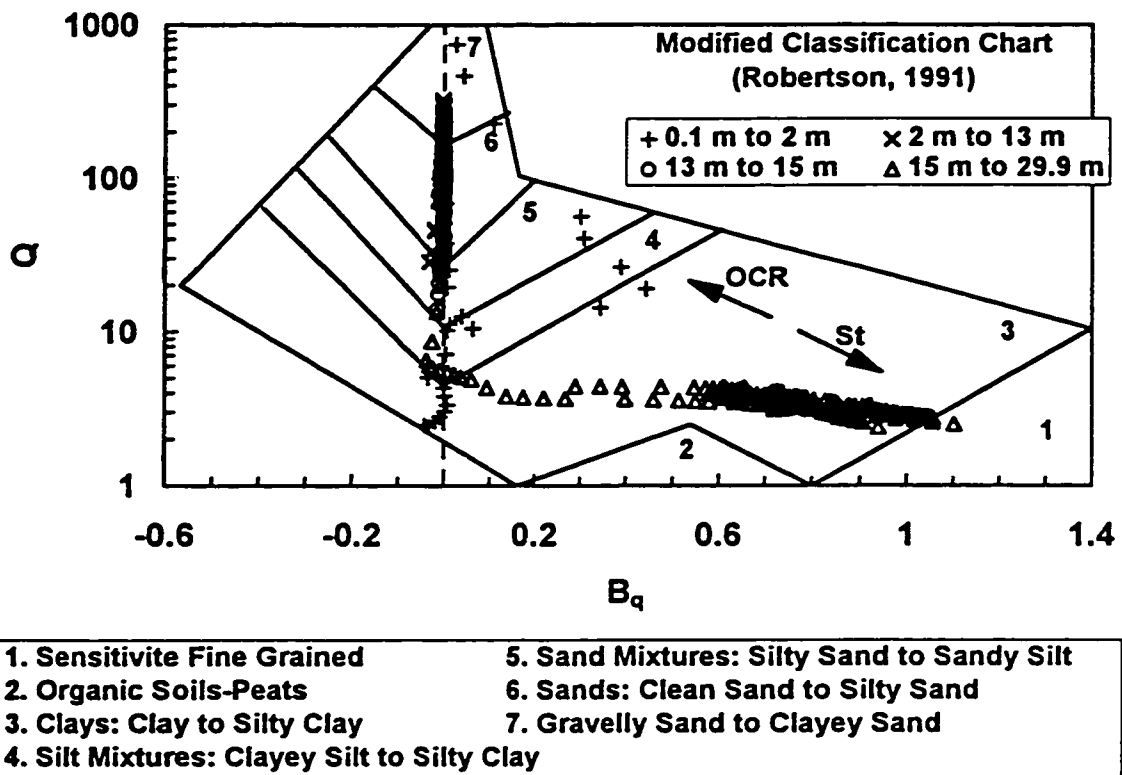
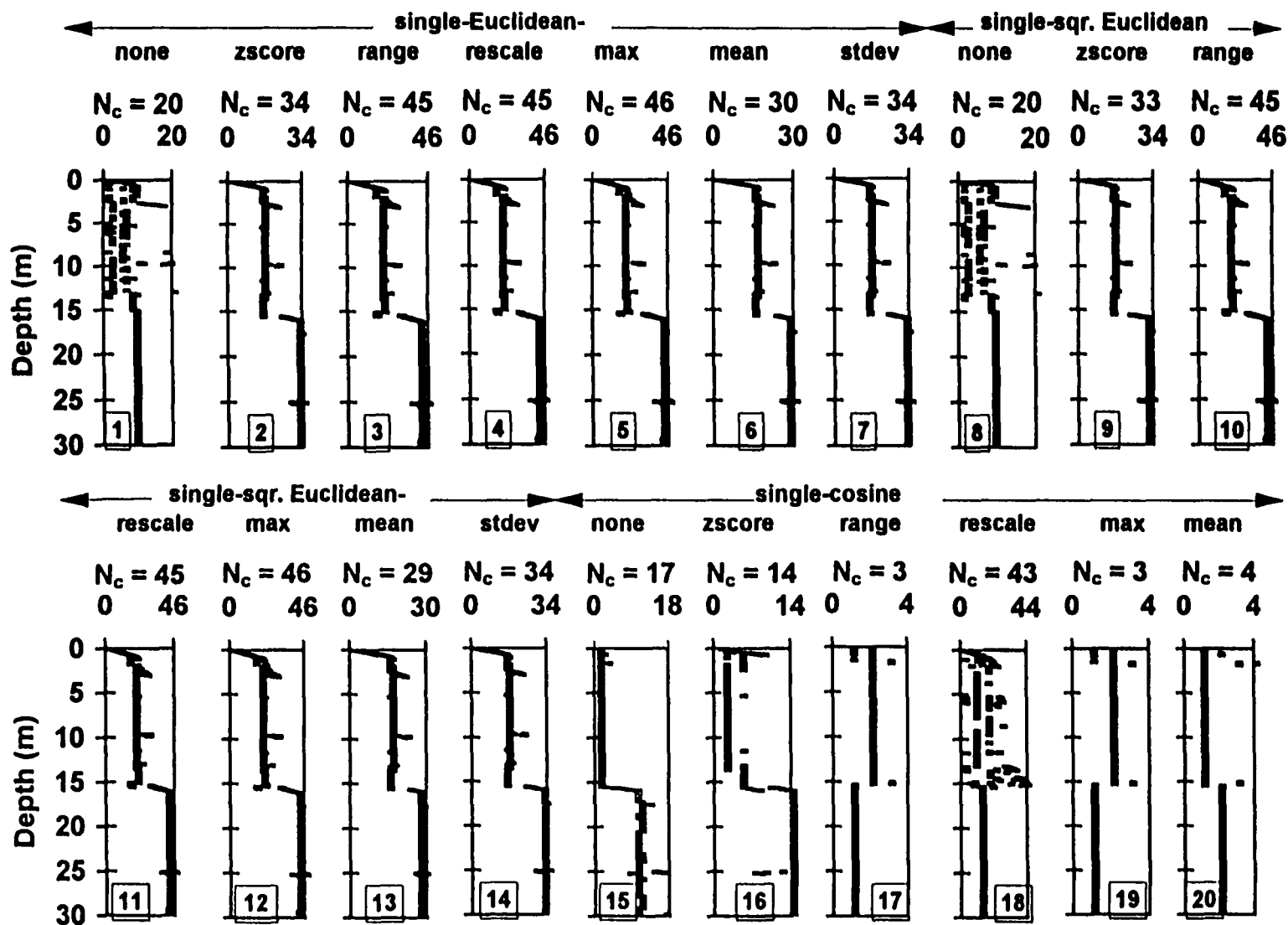


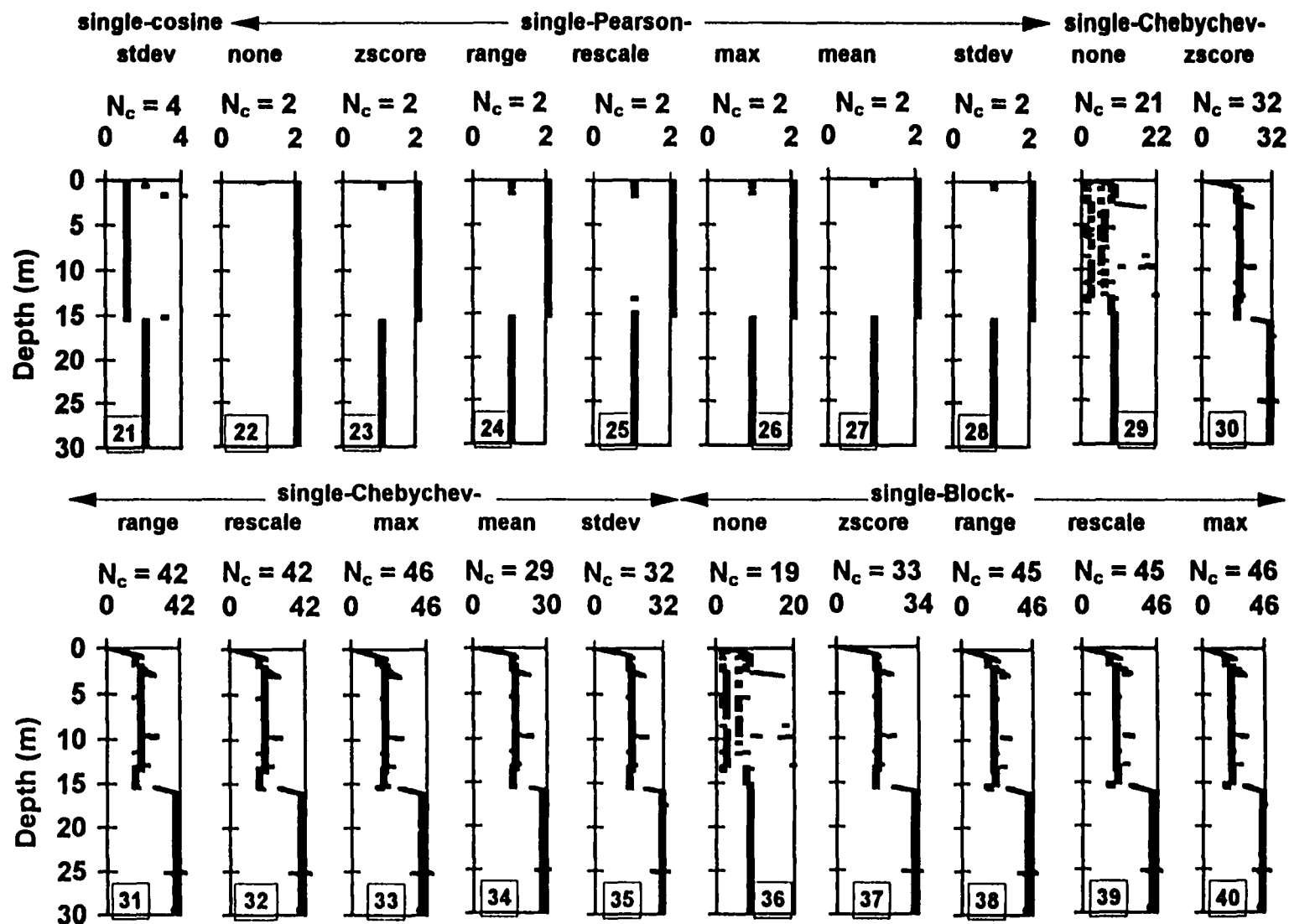
Figure B.3. Soil Classification Using Piezocone Data at Mcdonald's Farm, British Columbia (Data from Robertson, 1982).

chosen to represent the profile. Figure B.7 shows a good agreement between the borehole soil boundaries and the statistical boundaries at $N_c = 14$. The upper layer A1* represents a soil mixture and is more associated with layer A3. Layers A2 (sands) and A4 (clayey silts) are denoted by cluster numbers 2 and 14, respectively, which indicates non-association between the two groups.



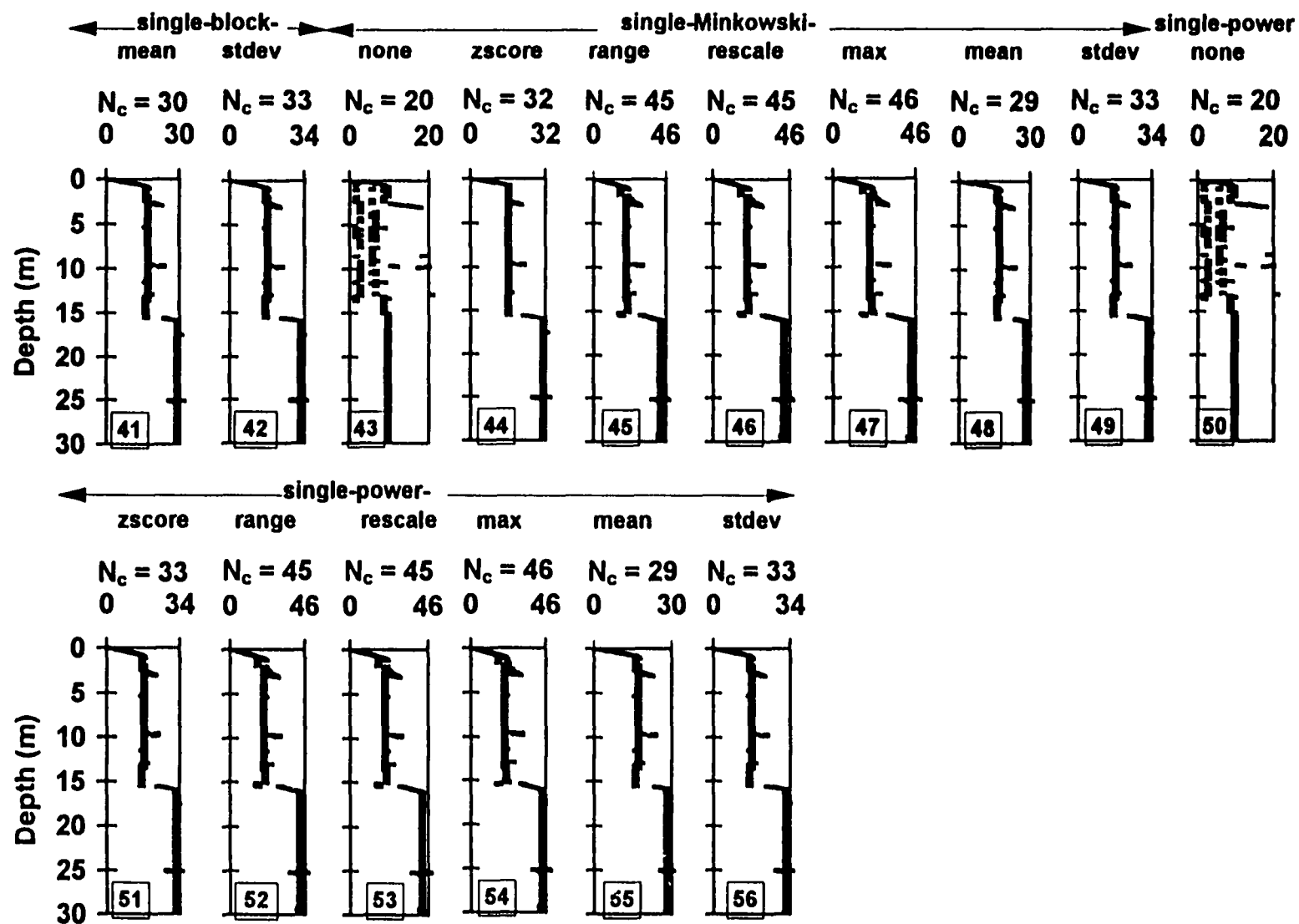
Note: Piezocone data at McDonald's Farm, British Columbia (Robertson, 1982). [no.] = cluster method no. (see Table C. 2)

Figure B.4a. Cluster Results of Piezocone Data at McDonald's Farm Using Different Cluster Methods.



Note: Piezocone data at McDonald's Farm, British Columbia (Robertson, 1982). [no.] = cluster method no. (see Table C. 2)

Figure B. 4b. Cluster Results of Piezocone Data at McDonald's Farm Using Different Cluster Methods.



Note: Piezocone data at McDonald's Farm, British Columbia (Robertson, 1982). [no.] = cluster method no. (see Table C. 2)

Figure B. 4c. Cluster Results of Piezocone Data at McDonald's Farm Using Different Cluster Methods.

Table B.2. Summary of 56 Different Cluster Analyses Performed on Piezocone Data from McDonald's Farm, British Columbia (Data from this Study).

Method	Chosen Cluster	Results	Method	Chosen Cluster	Results
(1) None-Euclidean	20	N	(29) None-Chebychev	21	N
(2) Zscore-Euclidean	34	Y	(30) Zscore-Chebychev	32	Y
(3) Range-Euclidean	45	Y	(31) Range-Chebychev	42	Y
(4) Rescale-Euclidean	45	Y	(32) Rescale-Chebychev	42	Y
(5) Max-Euclidean	46	Y	(33) Max-Chebychev	46	Y
(6) Mean-Euclidean	30	Y	(34) Mean-Chebychev	29	Y
(7) Stdev-Euclidean	34	Y	(35) Stdev-Chebychev	32	Y
(8) None-sqr. Euclidean	20	N	(36) None-Block	19	N
(9) Zscore-sqr. Euclidean	33	Y	(37) Zscore-Block	33	Y
(10) Range-sqr. Euclidean	45	Y	(38) Range-Block	45	Y
(11) Rescale-sqr. Euclidean	45	Y	(39) Rescale-Block	45	Y
(12) Max-sqr. Euclidean	46	Y	(40) Max-Block	46	Y
(13) Mean-sqr. Euclidean	29	Y	(41) Mean-Block	30	Y
(14) Stdev-sqr. Euclidean	34	Y	(42) Stdev-Block	33	Y
(15) None-Cosine	17	N	(43) None-Minkowski*	20	N
(16) Zscore-Cosine	14	Y	(44) Zscore-Minkowski*	32	N
(17) Range-Cosine	3	N	(45) Range-Minkowski*	45	Y
(18) Rescale-Cosine	6	N	(46) Rescale-Minkowski*	45	Y
(19) Max-Cosine	3	N	(47) Max-Minkowski*	46	Y
(20) Mean-Cosine	4	N	(48) Mean-Minkowski*	29	Y
(21) Stdev-Cosine	4	N	(49) Stdev-Minkowski*	33	Y
(22) None-Pearson	2	N	(50) None-Power**	20	N
(23) Zscore-Pearson	2	N	(51) Zscore-Power**	33	Y
(24) Range-Pearson	2	N	(52) Range-Power**	45	Y
(25) Rescale-Pearson	2	N	(53) Rescale-Power**	45	Y
(26) Max-Pearson	2	N	(54) Max-Power**	46	Y
(27) Mean-Pearson	2	N	(55) Mean-Power**	29	Y
(28) Stdev-Pearson	2	N	(56) Stdev-Power**	33	Y

Legend:

() = number of cluster analysis

N = one or more soil boundary is missed

(*) power $p = 2$ and root $r = 1/2$

Stdev = standard deviation

Y = soil profile is properly detected

(**) power $p = 1$ and root $r = 1/2$

Notes:

Standardization method⁽¹⁾

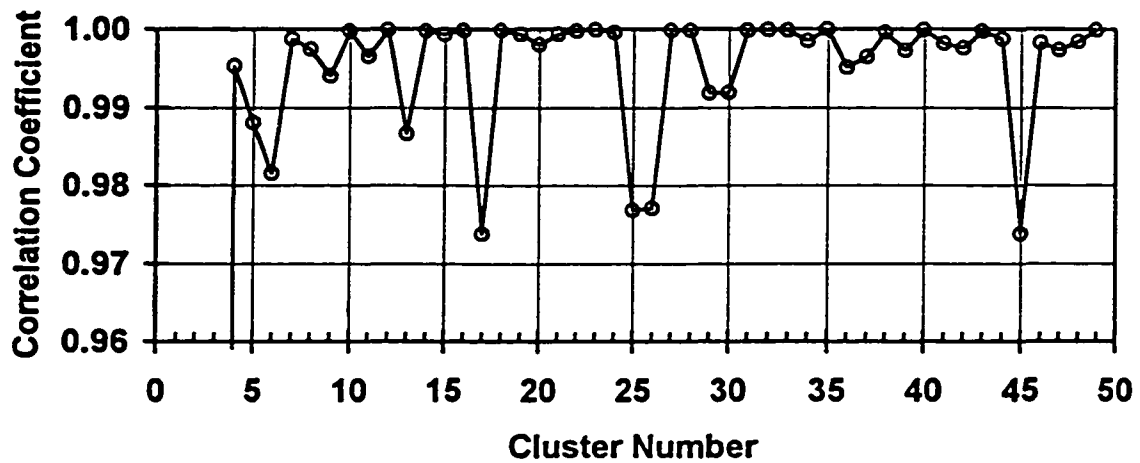
None (raw data)	Max
Zscore	Mean
Range	Stdev
Rescale	

Similarity measurement⁽²⁾

Euclidean	Chebychev
Squared Euclidean	Block (Manhattan Distance)
Cosine	Minkowski
Pearson	Power (p,r)

⁽¹⁾ Definitions of standardization methods were previously given in Table 3.2.

⁽²⁾ Definitions of similarity measurements were previously given in Table 3.3.



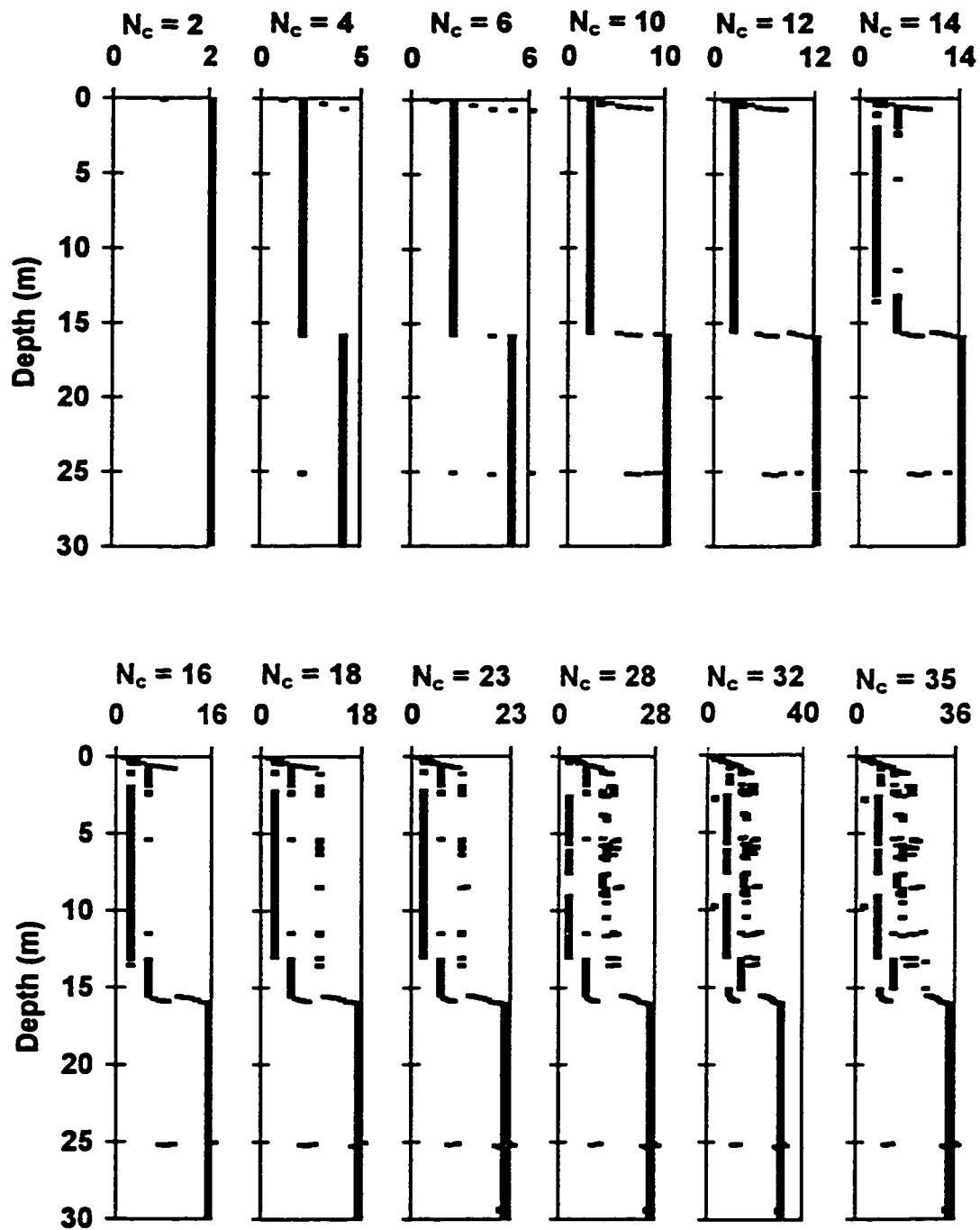
Single-Cosine-Zscore Method Using Q and B_q
Note: Piezocone data at McDonald's Farm, British Columbia (Robertson, 1982)

Figure B.5. Correlation Coefficient Between Consecutive Cluster Results at McDonald's Farm, British Columbia.

B.4. Fort Road Test Site, Singapore

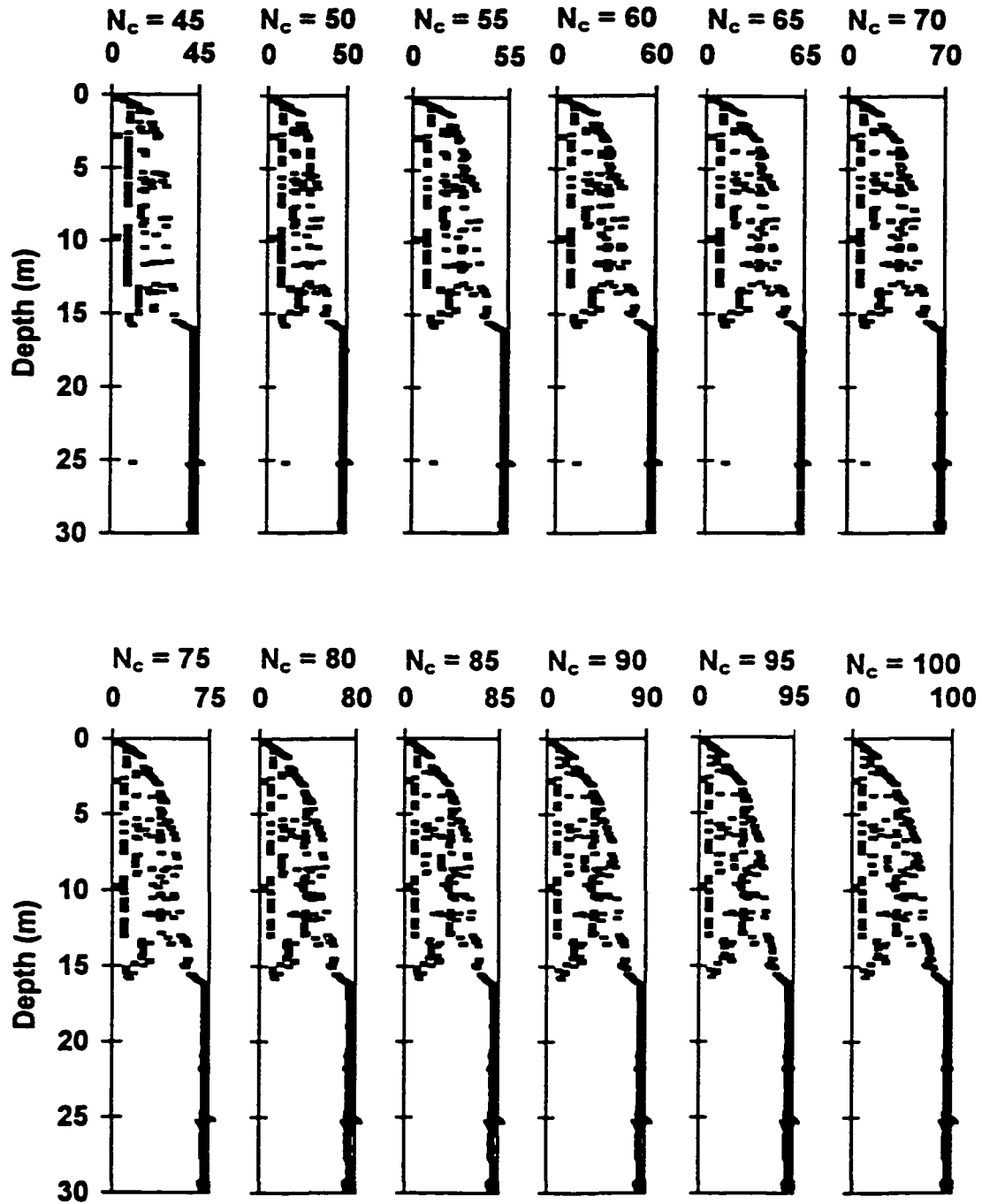
A representative piezocone sounding (Chang, 1991) at Fort Road, Singapore is shown in Fig. B.8. Three layers were delineated visually and separated at approximate depths of 17 m and 20.5 m. This observation is confirmed by soil classification results using the Robertson's chart (1991) as shown in Fig. B.9. The normalized parameters Q and B_q between the depths of 17 m and 21 m are grouped in zones 3 and 4 and can be separated visually from the group of the overlaying and underlying soils which are clustered in zone 3. Cluster analysis was performed using 56 different methods and the results are summarized in Fig. B.10.

Table B.3 indicates good agreement between the defined soil profile by Chang (1991) and the statistical layers using 52 techniques including the single-cosine-zscore



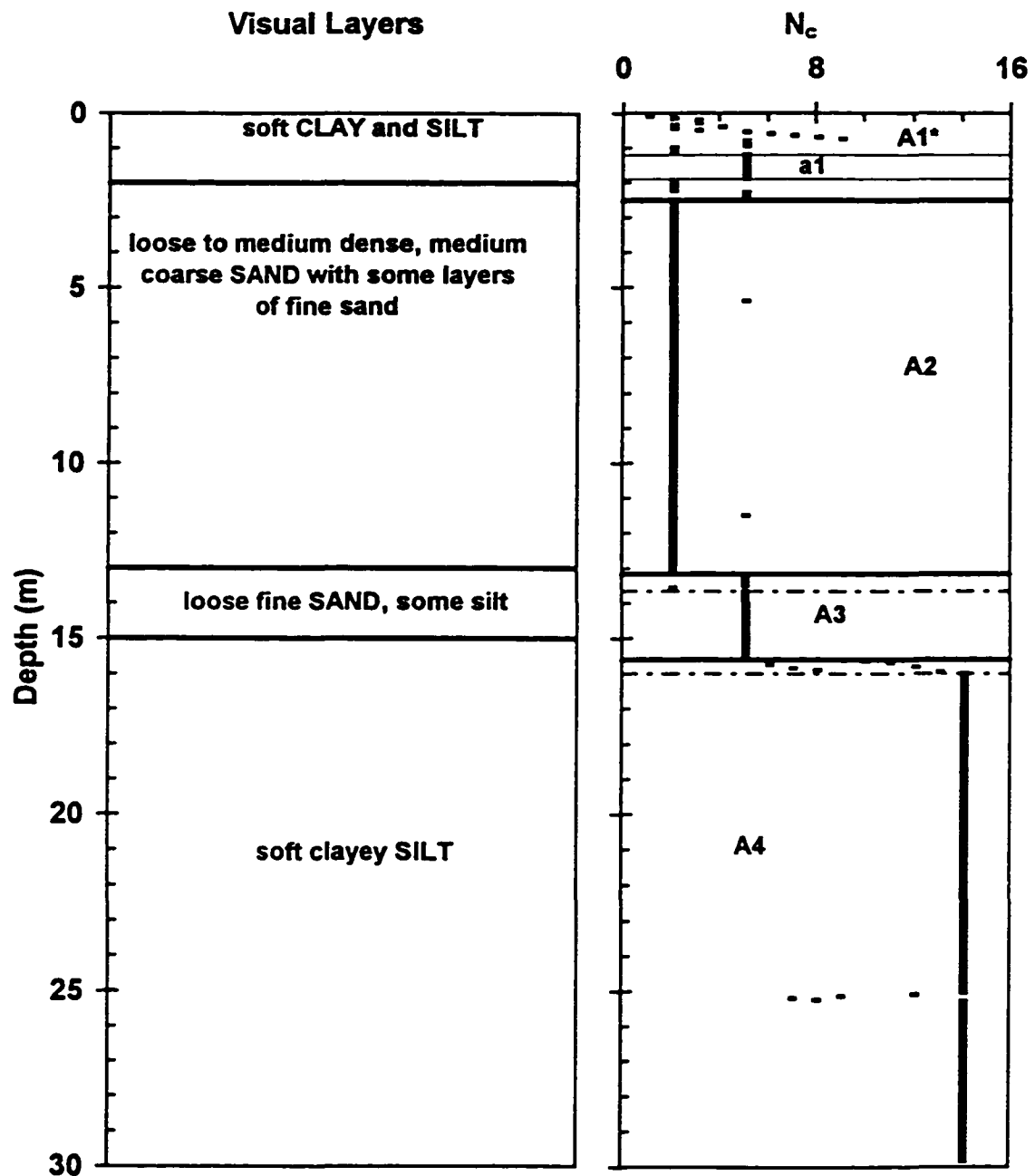
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at McDonald's Farm, British Columbia (Robertson, 1982)

Figure B. 6a. Cluster Analysis of Piezocone Data at McDonald's Farm, British Columbia by the Single-Cosine-Zscore Method.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at McDonald's Farm, Canada (Masood and Mitchell, 1993).

Figure B. 6b. Cluster Analysis of Piezocone Data at McDonald's Farm, British Columbia by the Single-Cosine-Zscore Method.



**Figure B. 7. Comparison Between Cluster Analysis and Visual Classification.
 from Borehole at McDonald's Farm, British Columbia..**

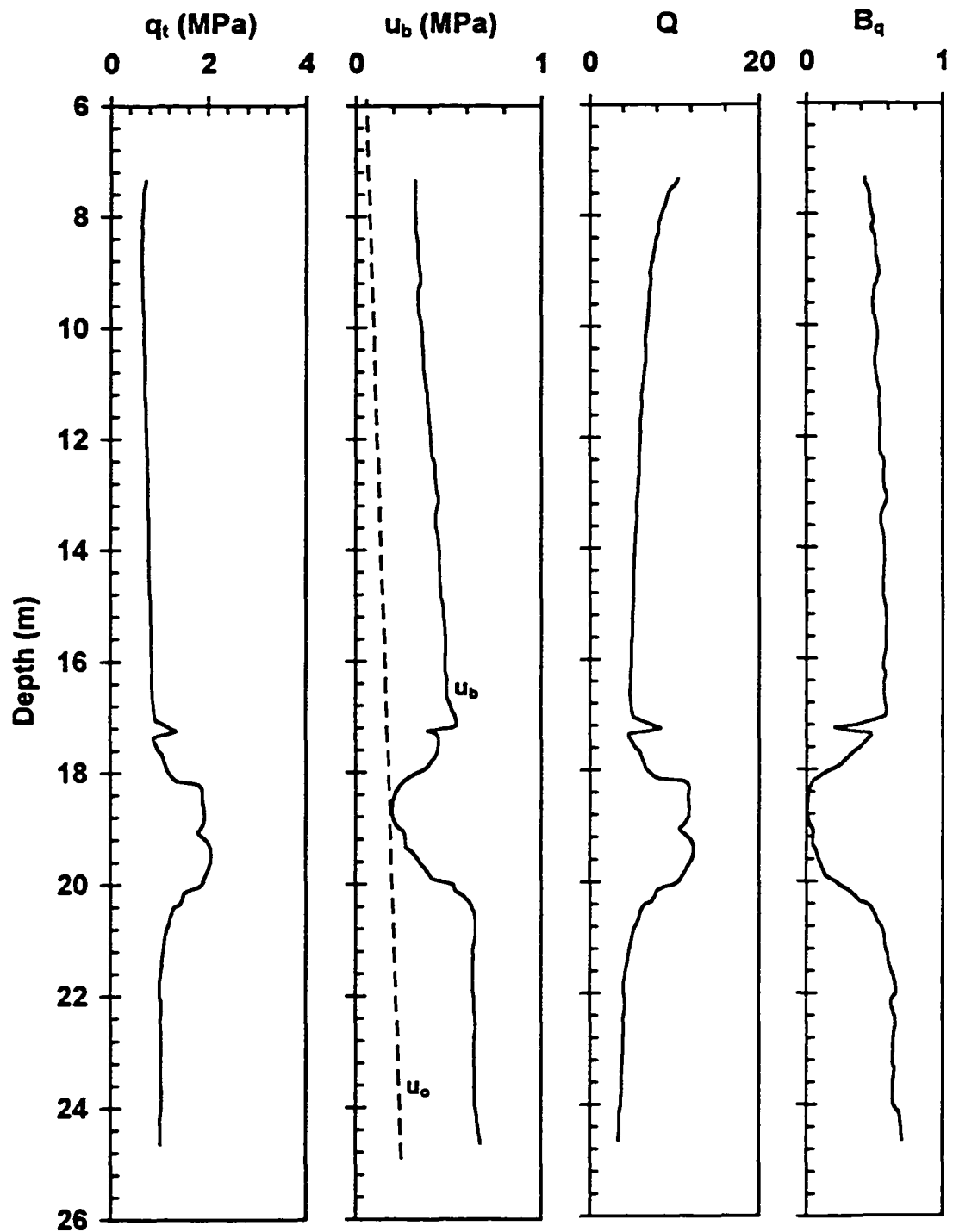


Figure B.8. Piezocone Sounding at Fort Road, Singapore (Data from Chang, 1991).

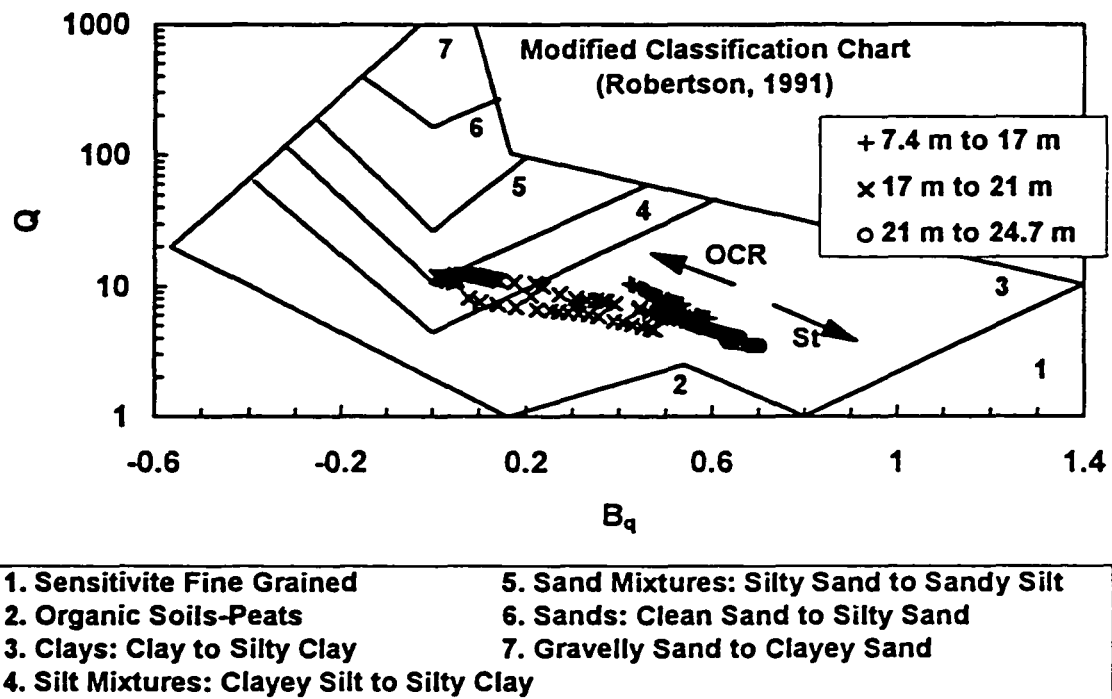
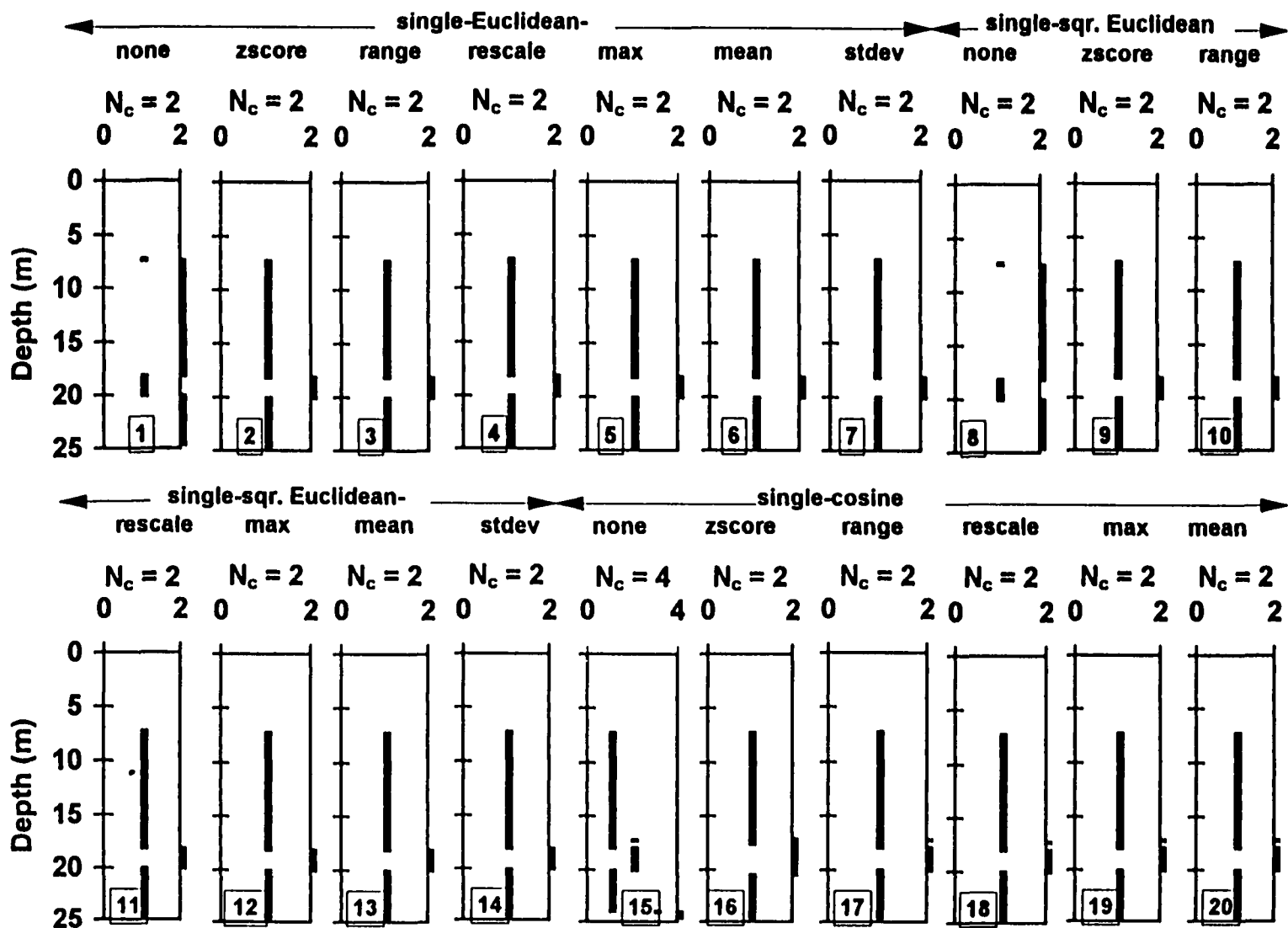


Figure B.9. Soil Classification Using Piezocone Data at Fort Road, Singapore (Data from Chang, 1991).

(SCZ) method. However, in the other four methods denoted by numbers 23, 24, 27 and 28, an unverified layer was indicated between the depths of 7.4 m and 10.5 m. Clustering using SCZ method is discussed in details herein.

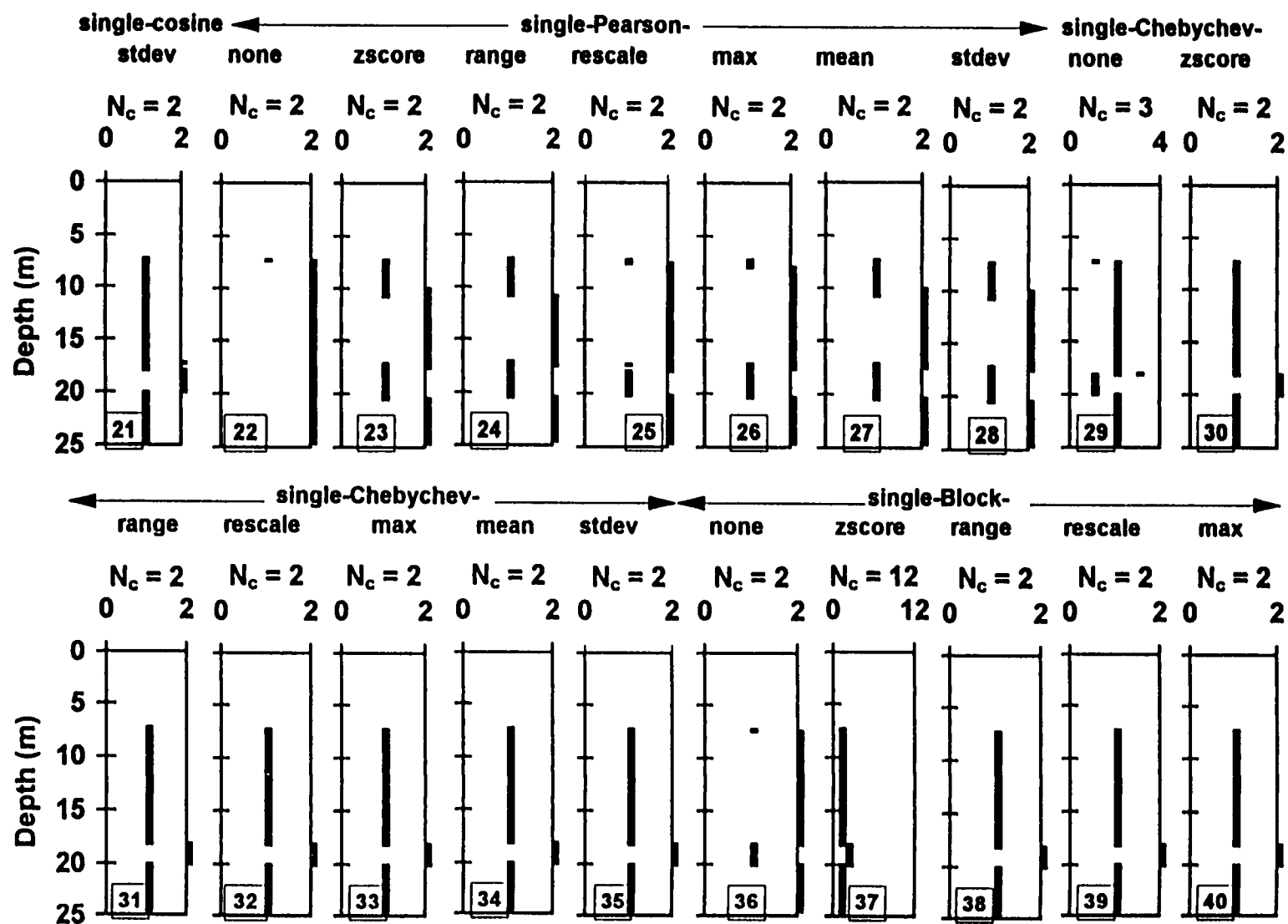
Clustering was performed using the single-cosine-zscore (SCZ) method up to cluster number $N_c = 100$. The correlation coefficient (ρ_c) between the consecutive clusters are shown in Fig. B.11 between cluster numbers $N_c = 2$ and 50. Cluster results are examined at the peaks of ρ_c and shown in Fig. B.12a for N_c between 2 and 36. At $N_c = 2$, one soil type is predominant and denoted by cluster number $N_c = 1$ with intermediate soil



Note: Piezocone data at Fort Road, Singapore (Chang, 1991).

[no.] = cluster method no. (see Table C. 3)

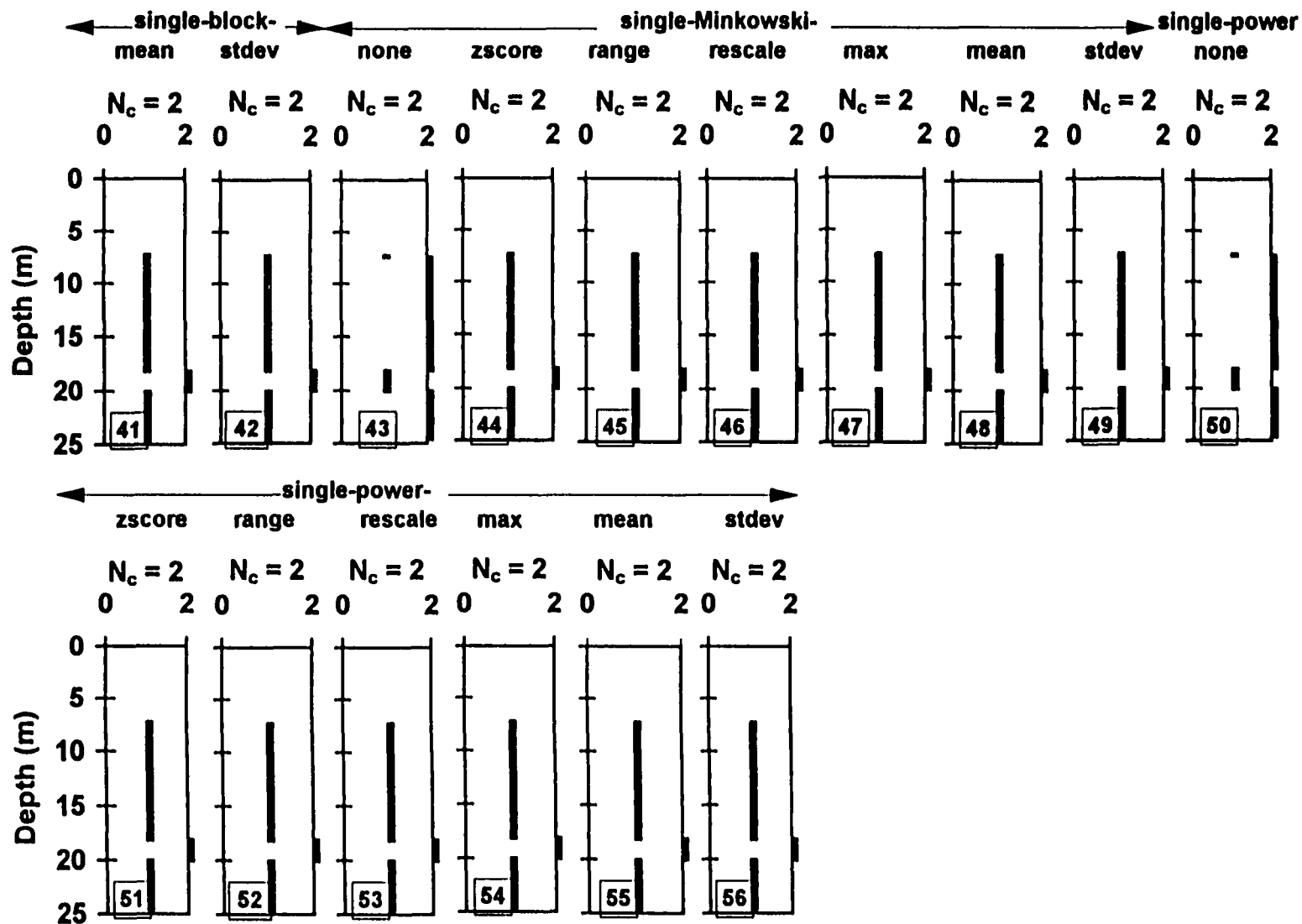
Figure B. 10a. Cluster Results of Piezocone Data at Fort Road, Singapore Using Different Cluster Methods.



Note: Piezocone data at Fort Road, Singapore (Chang, 1991).

[no.] = cluster method no. (see Table C. 3)

Figure B.10b. Cluster Results of Piezocone Data at Fort Road, Singapore Using Different Cluster Methods.



Note: Piezocone data at Fort Road, Singapore (Chang, 1991).

[no.] = cluster method no. (see Table C. 3)

Figure B. 10c. Cluster Results of Piezocone Data at Fort Road, Singapore Using Different Cluster Methods.

Table B.3. Summary of 56 Different Cluster Analyses Performed on Piezocone Data from Fort Road, Singapore (Data from Chang, 1991).

Method	Chosen Cluster	Results	Method	Chosen Cluster	Results
(1) None-Euclidean	2	Y	(29) None-Chebychev	3	Y
(2) Zscore-Euclidean	2	Y	(30) Zscore-Chebychev	2	Y
(3) Range-Euclidean	2	Y	(31) Range-Chebychev	2	Y
(4) Rescale-Euclidean	2	Y	(32) Rescale-Chebychev	2	Y
(5) Max-Euclidean	2	Y	(33) Max-Chebychev	2	Y
(6) Mean-Euclidean	2	Y	(34) Mean-Chebychev	2	Y
(7) Stdev-Euclidean	2	Y	(35) Stdev-Chebychev	2	Y
(8) None-sqr. Euclidean	2	Y	(36) None-Block	2	Y
(9) Zscore-sqr. Euclidean	2	Y	(37) Zscore-Block	2	Y
(10) Range-sqr. Euclidean	2	Y	(38) Range-Block	2	Y
(11) Rescale-sqr. Euclidean	2	Y	(39) Rescale-Block	2	Y
(12) Max-sqr. Euclidean	2	Y	(40) Max-Block	2	Y
(13) Mean-sqr. Euclidean	2	Y	(41) Mean-Block	2	Y
(14) Stdev-sqr. Euclidean	2	Y	(42) Stdev-Block	2	Y
(15) None-Cosine	2	Y	(43) None-Minkowski*	2	Y
(16) Zscore-Cosine	2	Y	(44) Zscore-Minkowski*	2	Y
(17) Range-Cosine	2	Y	(45) Range-Minkowski*	2	Y
(18) Rescale-Cosine	2	Y	(46) Rescale-Minkowski*	2	Y
(19) Max-Cosine	2	Y	(47) Max-Minkowski*	2	Y
(20) Mean-Cosine	2	Y	(48) Mean-Minkowski*	2	Y
(21) Stdev-Cosine	2	Y	(49) Stdev-Minkowski*	2	Y
(22) None-Pearson	2	Y	(50) None-Power**	2	Y
(23) Zscore-Pearson	2	Y	(51) Zscore-Power**	2	Y
(24) Range-Pearson	2	Y	(52) Range-Power**	2	Y
(25) Rescale-Pearson	2	Y	(53) Rescale-Power**	2	Y
(26) Max-Pearson	2	Y	(54) Max-Power**	2	Y
(27) Mean-Pearson	2	Y	(55) Mean-Power**	2	Y
(28) Stdev-Pearson	2	Y	(56) Stdev-Power**	2	Y

Legend:

() = number of cluster analysis

N = one or more soil boundary is missed

(*) power $p = 2$ and root $r = 1/2$

Notes:

Stdev = standard deviation

Y = soil profile is properly detected

(**) power $p = 1$ and root $r = 1/2$

Standardization method⁽¹⁾

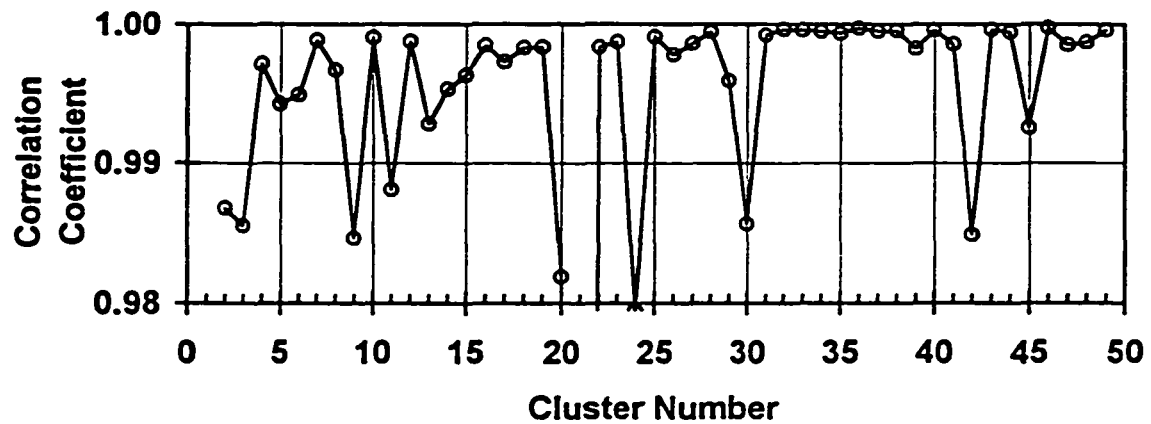
None (raw data)	Max
Zscore	Mean
Range	Stdev
Rescale	

Similarity measurement⁽²⁾

Euclidean	Chebychev
Squared Euclidean	Block (Manhattan Distance)
Cosine	Minkowski
Pearson	Power (p,r)

⁽¹⁾ Definitions of standardization methods were previously given in Table 3.2.

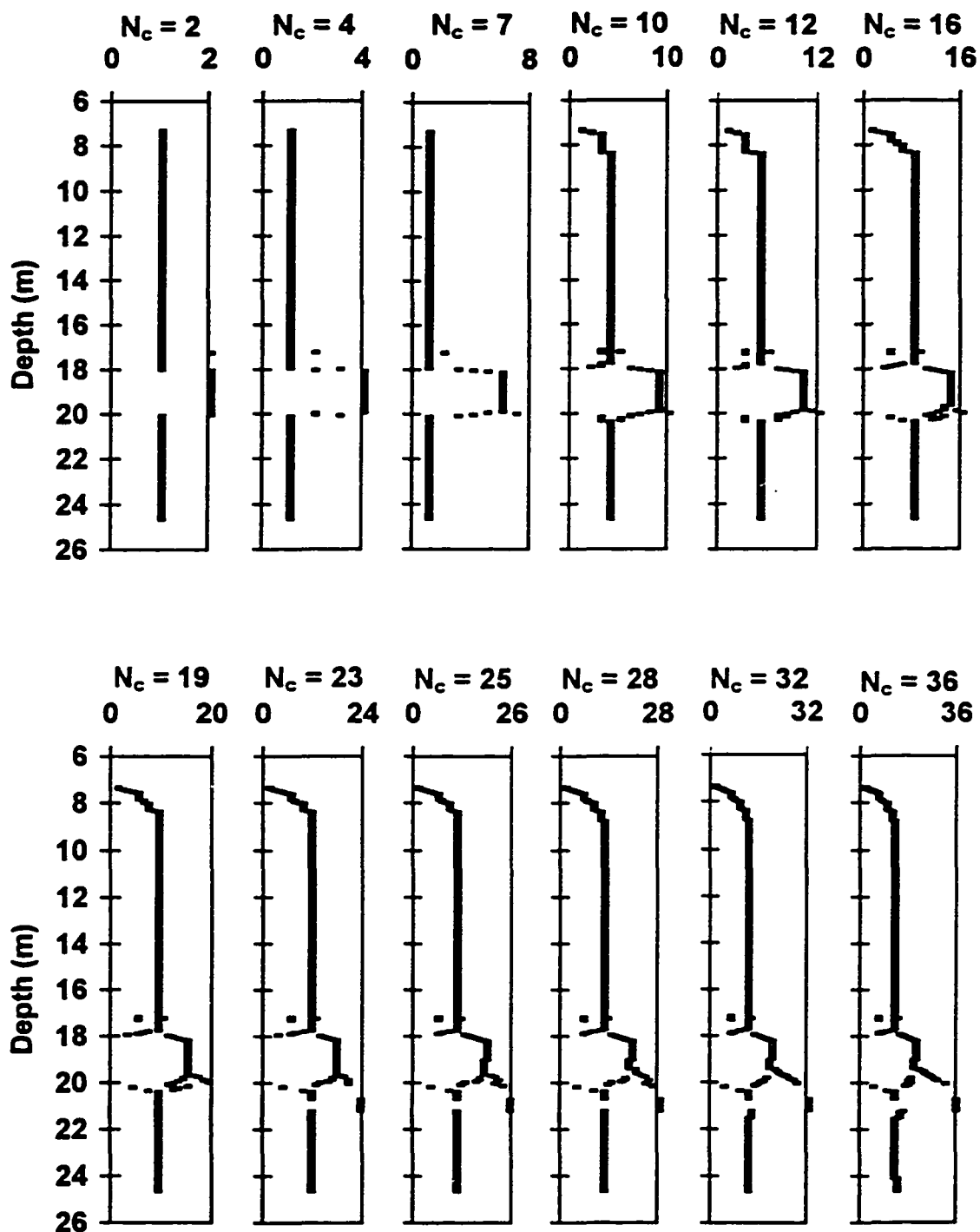
⁽²⁾ Definitions of similarity measurements were previously given in Table 3.3.



Single-Cosine-Zscore Method Using Q and B_q
Note: Piezocone data at Fort Road, Singapore (Chang, 1991).

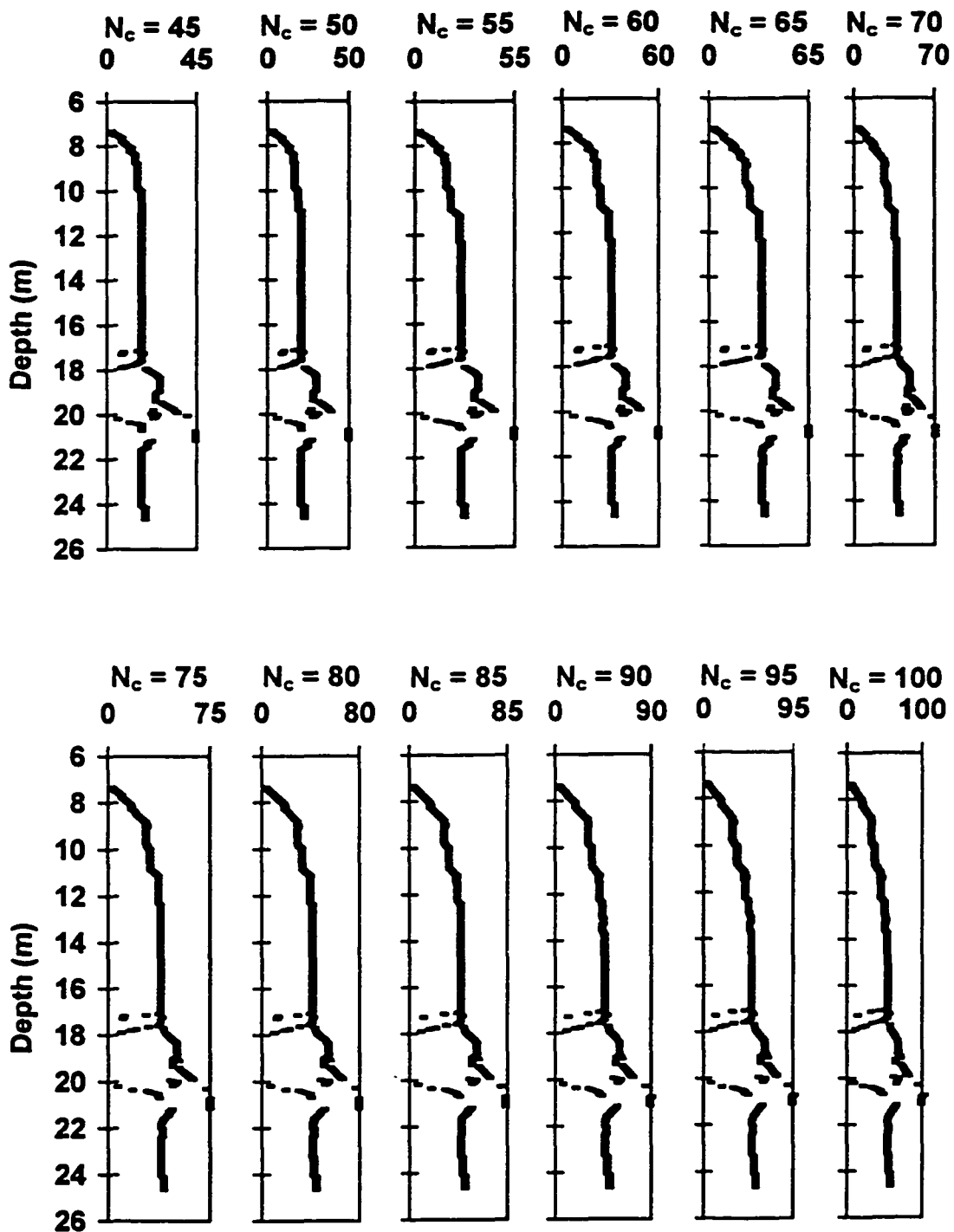
Figure B.11. Correlation Coefficient Between Consecutive Cluster Results at Fort Road, Singapore.

layer denoted by cluster number $N_c = 2$ between the depth of 17.7 m and 21 m. For larger cluster numbers, points separate from these two primary clusters indicating transition between them and inherent non-homogeneity. The separated points have the largest dissimilarity with a primary statistical cluster. This scenario continues from $N_c = 45$ to $N_c = 100$ as noticed in Fig. B.12b. The clustering is confirmed by the water contents profile as shown in Fig. B.13. The average water content is equal to 62 % from 7.3 m to 17 m and decreases to 33 % down to 21 m. Then, the average water contents increases to 51 % between the depths of 17 m and 24.5 m.



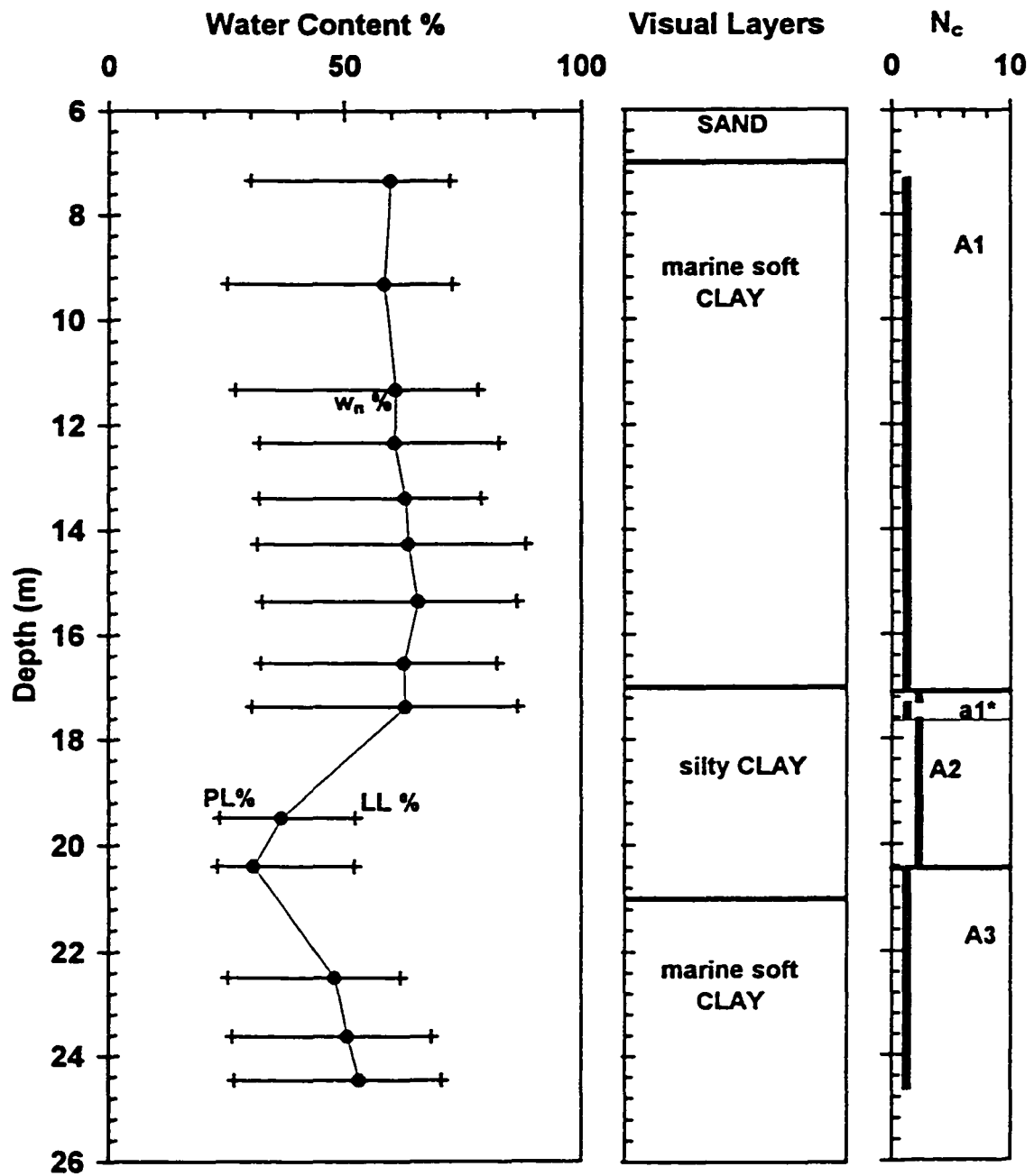
Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data at Fort Road, Singapore (Chang, 1991).

Figure B.12a. Cluster Analysis of Piezocone Data at Fort Road, Singapore.



Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data at Fort Road, Singapore (Chang, 1991).

Figure B.12b. Cluster Analysis of Piezocone Data at Fort Road, Singapore.



Note: Piezocone data at Fort Road, Singapore (Chang, 1991).

SCZ method using Q and B_q , cluster no. 2

— Primary boundary
— Secondary boundary

Figure B.13. Comparison Between Cluster Analysis, Visual Classification and Water Content % at Fort Road, Singapore.

B.5. Conclusions

Piezocone data were analyzed using 56 different clustering combinations of standardization methods and similarity measurements. A total number of 168 clustering analyses was performed at three sites to detect changes in stratigraphic profiles. These results were compared with physical and visual data from adjacent borehole determinations of strata layers at the sites. The zscore standardization method and the cosine similarity measurements were the only two techniques that properly defined the subsurface stratification for the minimum cluster number at each of the three sites. Also, the single link (nearest neighbor) cluster method was used successfully to divide the data into correlated groups. Therefore the statistical single-cosine-zscore (SCZ) method is recommended for clustering of piezocone data in delineating subsurface strata profiles.

APPENDIX C

GROWTH OF CLUSTERING

C.1. Synopsis

In this section, the growth of the cluster results is discussed up to cluster number $N_c = 100$ at twelve sites summarized in Table C.1. Cluster analysis was performed at each site using the normalized parameters Q and B_q . The correlation coefficient (ρ_c) was determined between consecutive clusters and clustering was examined at the peaks of ρ_c . According to the interpretation criterion discussed in Chapter 4, a cluster number was chosen to represent a subsurface stratigraphy where no new primary groups ($t > 1$ m) separated.

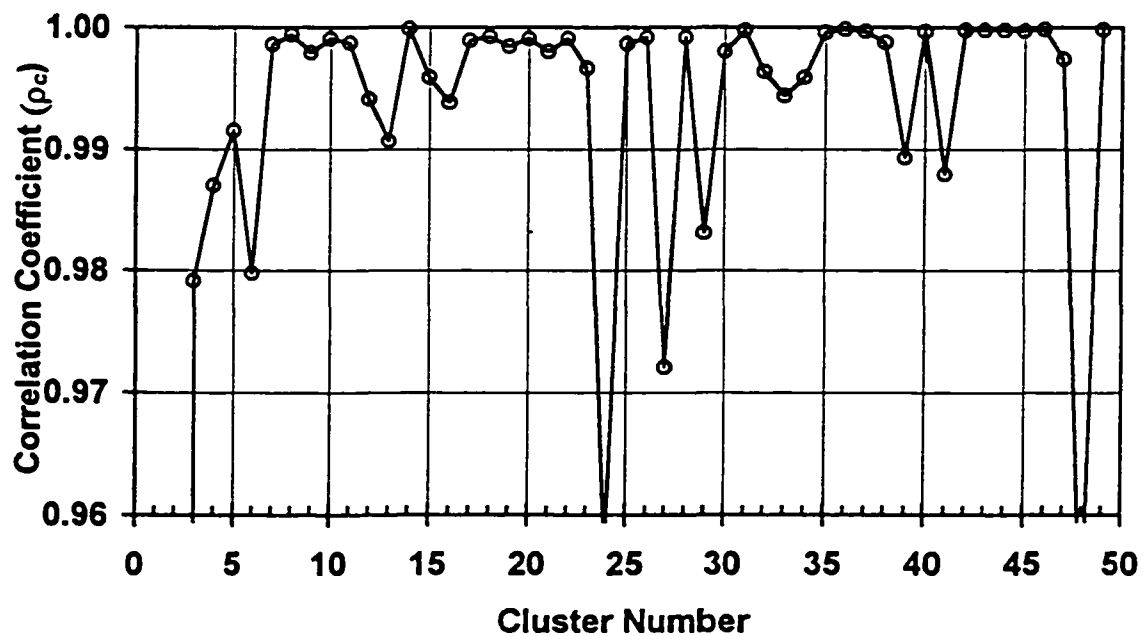
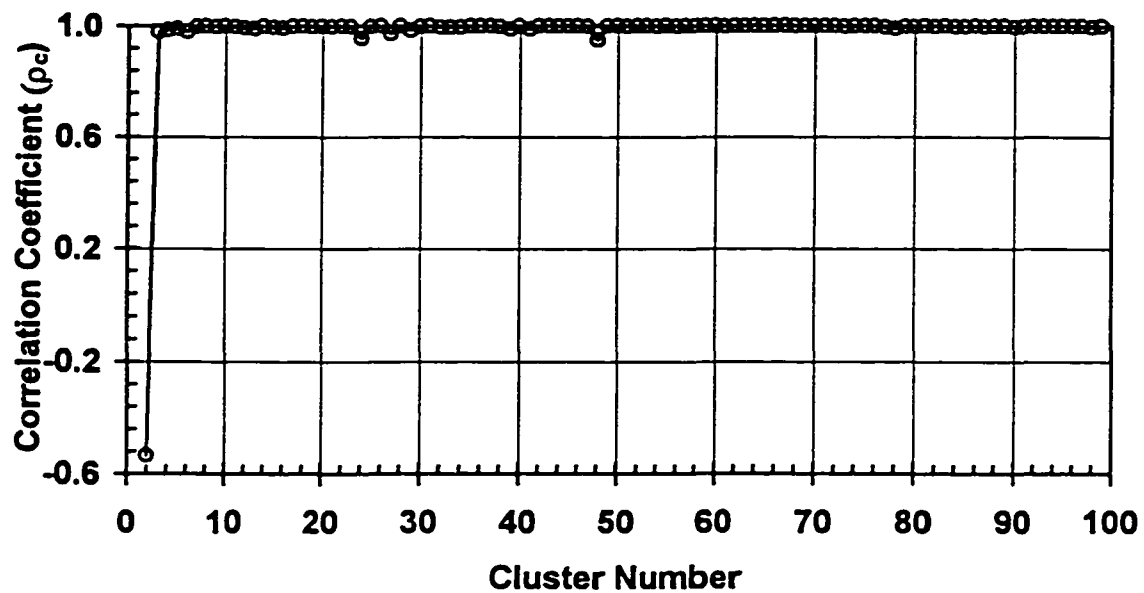
C.2. Effect of Different Data Combinations

A single-cosine-zscore clustering method is used to analyze three sets of piezocone data combinations defined as follows: (1) unprocessed data, q_t and u_b ; (2) partially processed data, q_t and the ratio u_b/q_t ; and (3) the derived normalized parameters, Q and B_q , where $Q = (q_t - \sigma_{vo})/\sigma_{vo}'$ and $B_q = (u_b - u_o)/(q_t - \sigma_{vo})$. The study was performed using piezocone data sounding number (PCPT1) collected at Amherst, Massachusetts by the author. The piezocone sounding was previously presented in Fig. 5.2. First, clustering

Table C.1. Listing of 12 Sites Analyzed Using SCZ-Type Clustering Method.

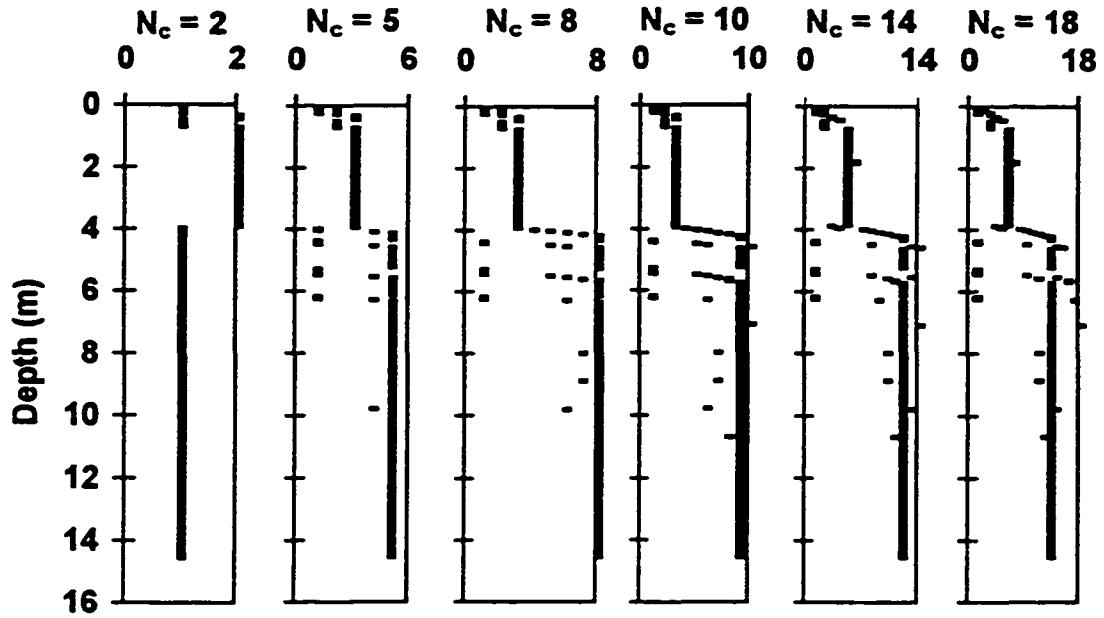
Site Name	Location	Soil Type	Reference
Amherst	Massachusetts	Fill over clay crust underlain by soft clay	(This study)
Bothkennar	Scotland	Clay crust over soft silty clay, clayey silt	Nash et al. (1992)
Brent Cross	United Kingdom	Weathered finely fissured clay over unweathered highly fissured clay	Powell et al. (1988)
Surry	Virginia	Clay over sand underlain by clay	Gordon and Mayne (1987)
Drammen	Norway	Plastic clay over lean clay	Masood et al. (1990)
Gloucester	Ontario	Soft silty clay over clay to silty clay	Konrad and Law (1987)
Hachirogata	Japan	Soft marine clay	Tanaka et al. (1992)
Lilla Mellösa	Sweden	Organic clay over clay underlain by a varved clay	Larsson and Mulabdić (1991)
Recife	Brazil	Organic soft clay (1) and (2)	Coutinho and Oliveira (1997)
St. Alban	Quebec	Soft very sensitive silty clay changes to clayey silt	Roy et al. (1982)
Tiller	Norway	Silty clay over quick clay	Sandven (1990)
Troll	North Sea	Very soft clay over very stiff silty clay	Amundsen et al. (1985)

using q_t and u_b is discussed. The correlation coefficient between consecutive clusters is shown in Fig. C.1. Clustering is checked at the peaks of ρ_c and the results are shown in Fig. C.2a between cluster numbers $N_c = 2$ and 36. At cluster number 2, two primary clusters are delineated at 3.9 m and a soil mixture appears at a depth of 0.75 m. For $N_c > 2$, no new major statistical layers separate from the grouping, however, some points are denoted different cluster numbers than those of the two dominant clusters indicating transitions, outliers or soil lenses. The difference of the cluster numbers given to the two major layers increases by increasing N_c , that suggests no association between them in terms of soil types and/or properties. Figure C.2b includes cluster results between $N_c = 45$



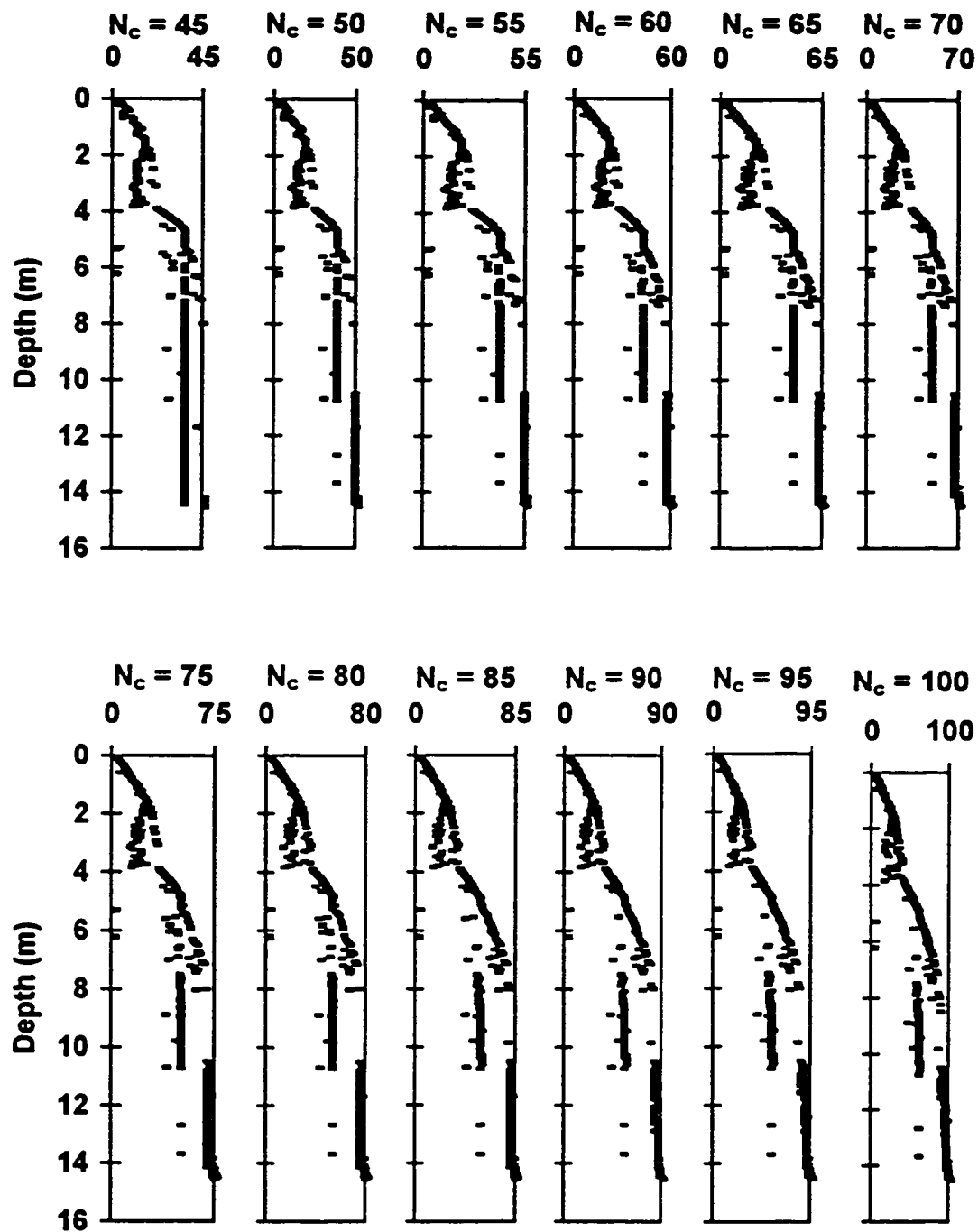
Single-Cosine-Zscore Method Using q_t and u_2
 Note: Piezocone data at Amherst, Massachusetts (this study).

Figure C.1. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using q_t and u_2 .



Single-Cosine-Zscore Method Using q_t and u_b $N_c =$ No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (this study).

Figure C.2a. SCZ-Type Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using q_t and u_b .



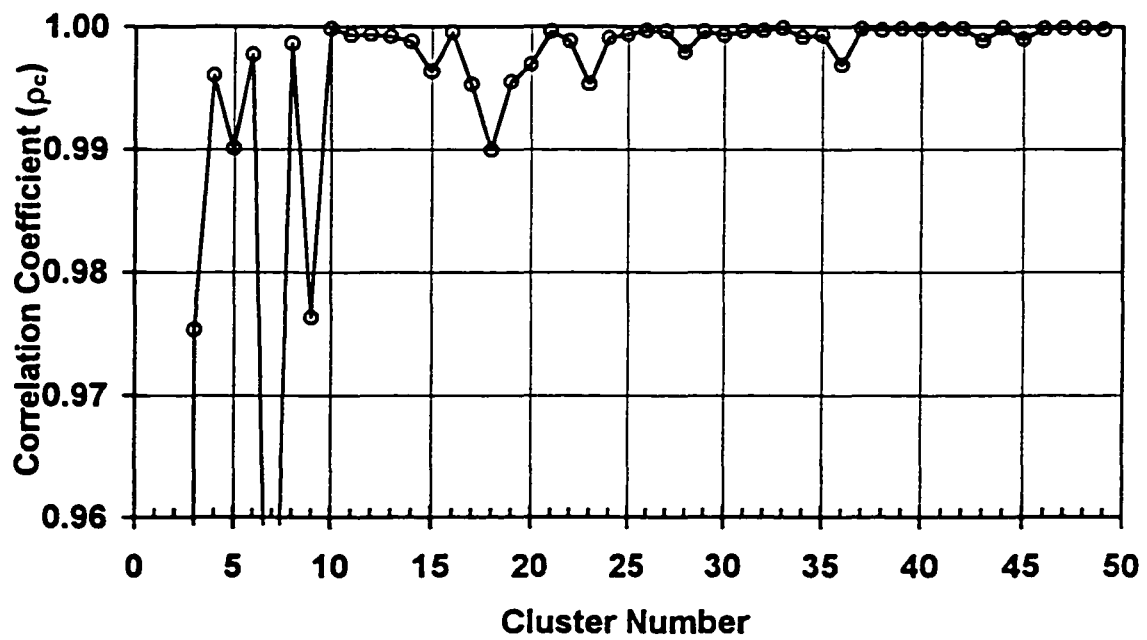
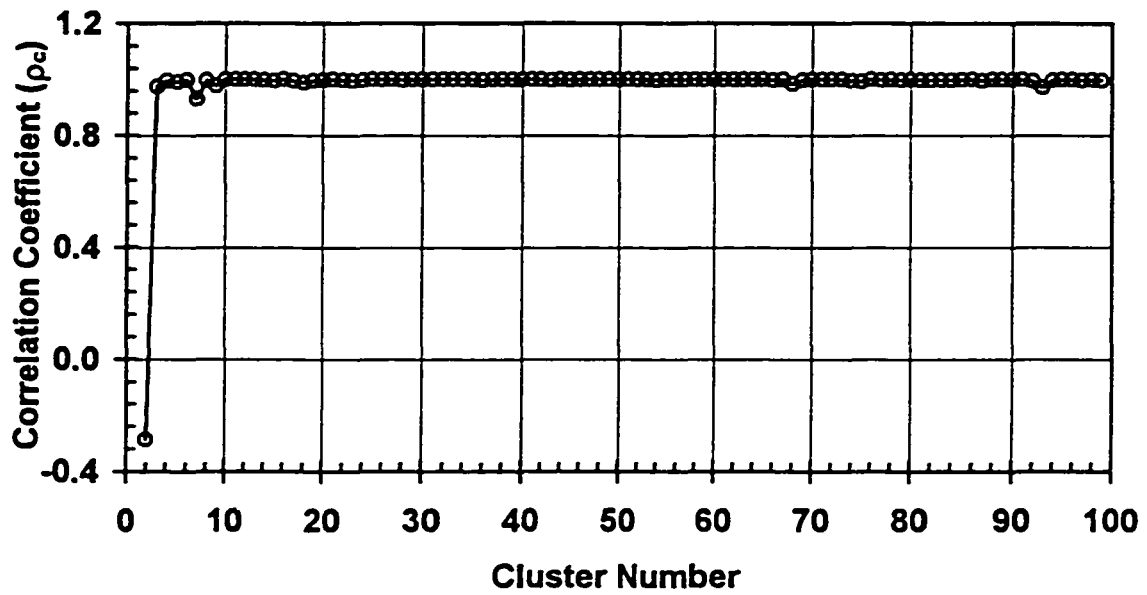
Single-Cosine-Zscore Method Using q_t and u_b N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (this study).

Figure C.2b. SCZ-Type Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using q_t and u_b .

and 100 every 5 increment interval and no new primary layers were discovered. Therefore cluster number 2 is chosen to represent the subsurface stratigraphy.

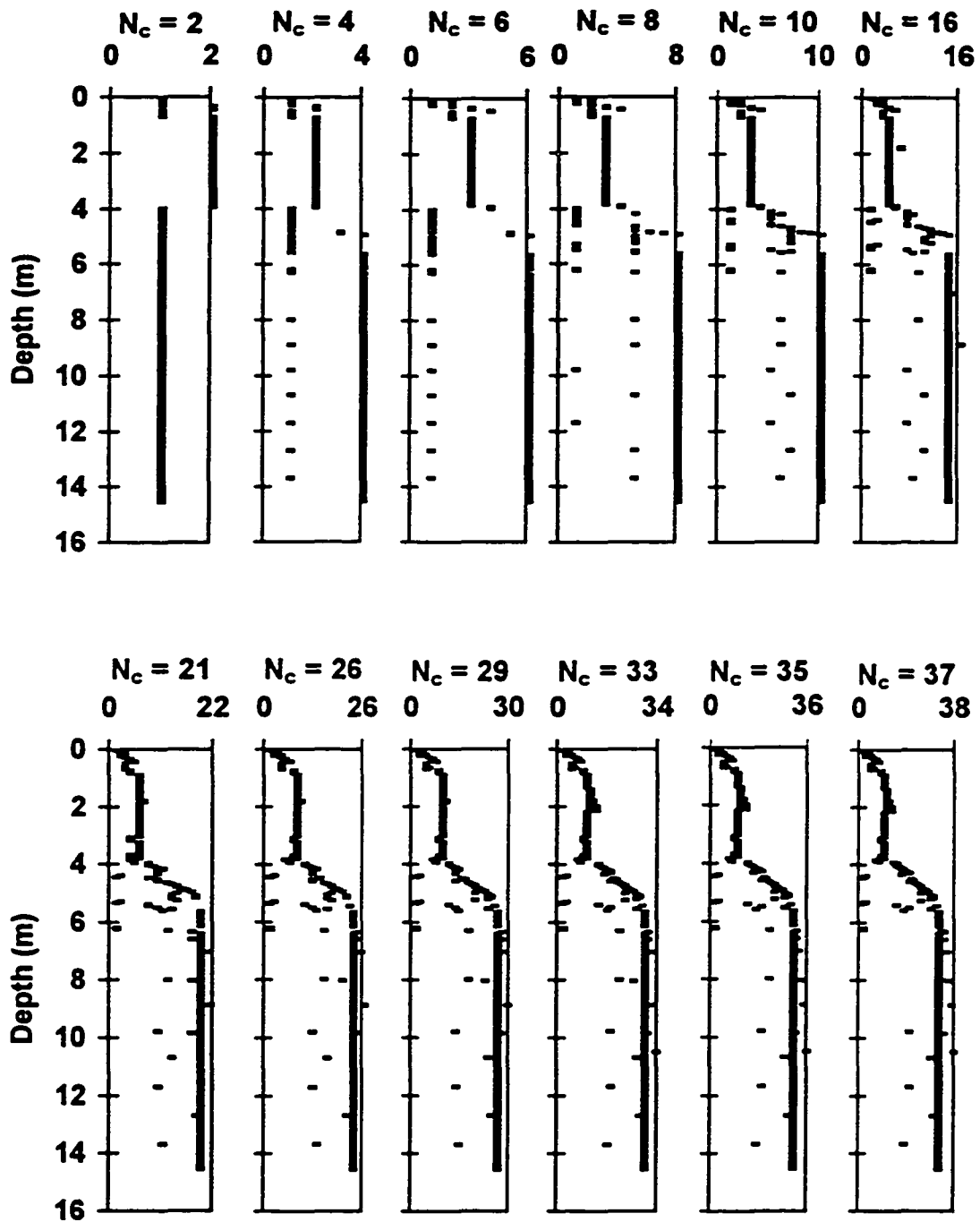
A similar clustering study is performed using partially processed data q_t and u_t/q_t up to cluster number $N_c = 100$. The correlation coefficient between consecutive clusters is shown in Fig. C.3 and the cluster results are examined at the peaks of ρ_c up to $N_c = 50$. Data grouping is shown in Fig. C.4a between cluster numbers 2 and 37. At cluster number 2, the data are divided into two primary groups with a boundary at a depth of 3.9 m. At cluster number 4, the lower statistical layer is separated into two primary layers at a depth of 5.6 m. Also, some points are separated at frequent depths at the locations where penetration stopped while successive cone rods were added. For larger cluster numbers, no new primary layers ($t > 1$ m) are detected, however more individual or small groups of points ($t < 0.5$ m) continue to depart from the grouping indicating less association with major clusters. This also applies up to cluster number $N_c = 100$ as shown in Fig. C.4b. Therefore, cluster number 4 is chosen to represent the subsurface stratigraphy.

Soil stratigraphy is derived based on clustering of three normalized parameters Q , B_q and F from the Amherst piezocone results and the correlation coefficient between consecutive clusters is shown in Fig. C.5. Cluster results are examined at the peaks of ρ_c up to cluster number $N_c = 50$ and Figure C.6a shows clustering between $N_c = 2$ and 32. Up to $N_c = 5$, the data are predicted to be in one group except few points implying non-homogeneity. At cluster number 8, the data are divided into two primary groups delineated at a depth of 4.8 m. For $N_c > 8$, no new major statistical layers ($t > 1$ m)



Single-Cosine-Zscore Method Using q_t and u_v/q_t
 Note: Piezocone data at Amherst, Massachusetts (Data from this study).

Figure C.3. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using q_t and the Ratio u_v/q_t .



Single-Cosine-Zscore Method Using q_t and u_v/q_t $N_c =$ No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

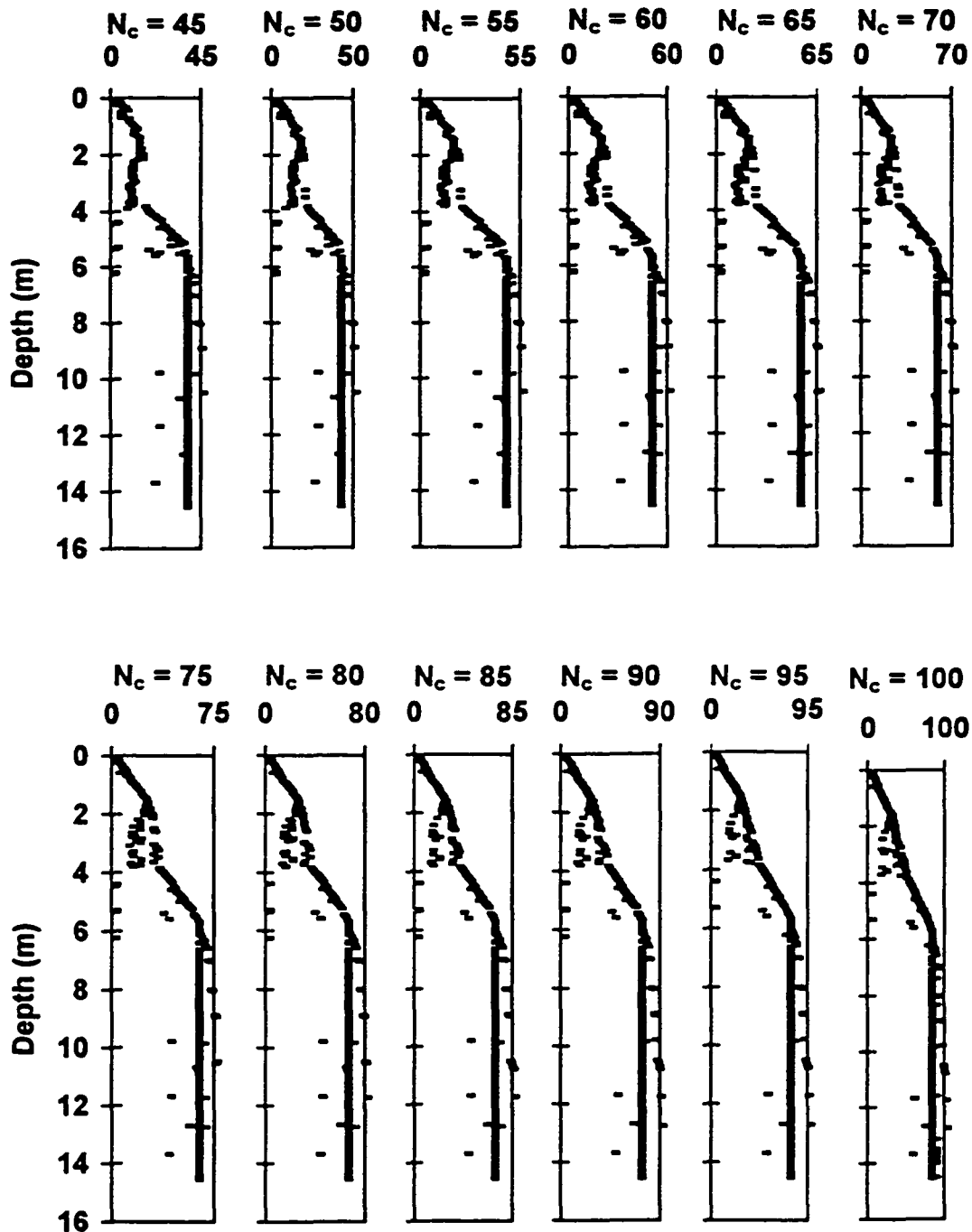
Figure C.4a. Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using q_t and the Ratio u_v/q_t .

appear, however, more points separate implying dissimilarity within the two primary layers. The latter statement is also valid up to $N_c = 100$ as shown in Fig. C.6b. Therefore, cluster number 8 is chosen to represent the subsurface stratification.

C.3. Procedural Errors

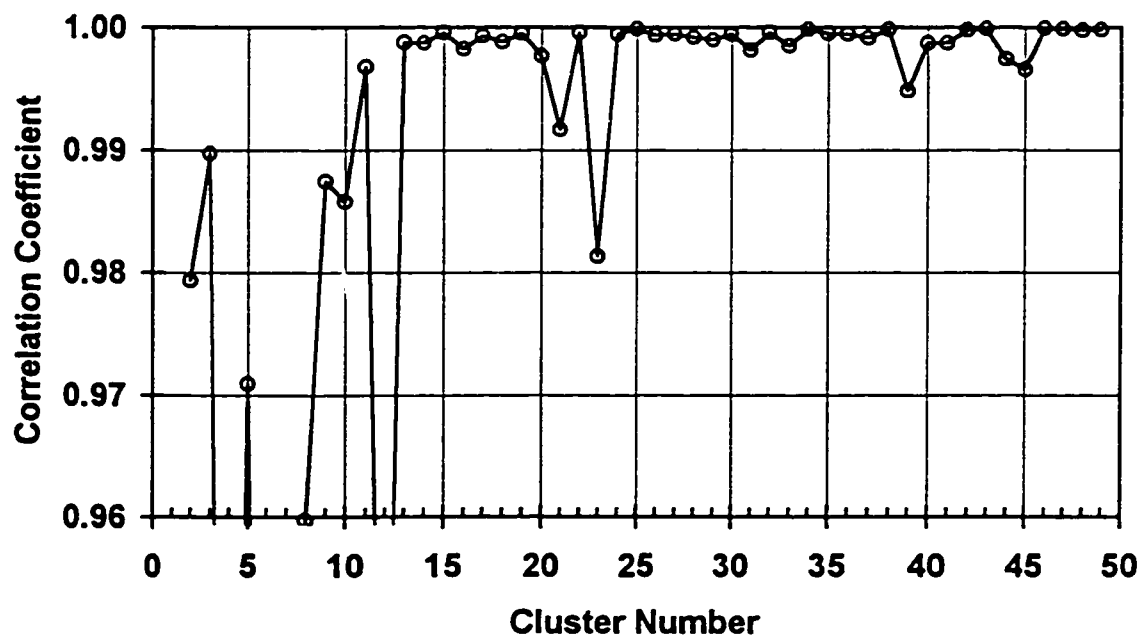
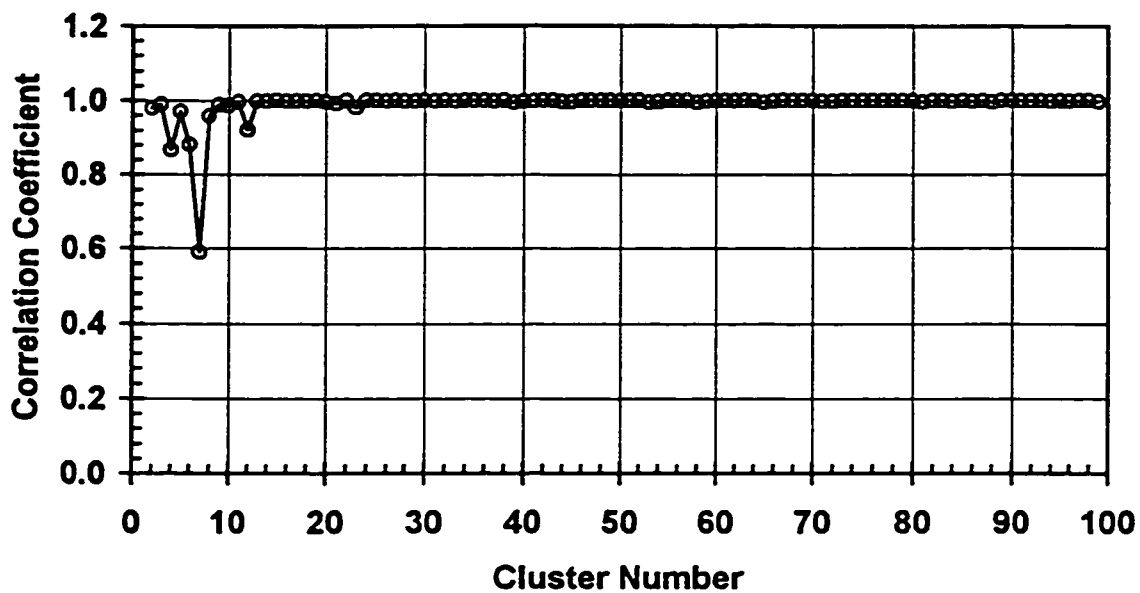
Piezocone data (sounding number PCPT1) collected at Amherst, Massachusetts by the author in June of 1996 include measurement errors due to frequent stops during each 1-m of penetration. A data filtering criterion proposed by Vivitrat (1978) was used to delineate the depths of the data errors at the rod breaks using a window width equal to 0.5 m. A cone reading was considered an outlier if it is greater than (average + 2 standard deviation) of the data within a window. The data errors were deleted and replaced by linear interpolation between the two data points above and below the removed measurements.

Filtered data are divided into correlated groups using single-cosine-zscore method. The correlation coefficient between consecutive cluster results is shown in Fig. C.7 up to $N_c = 100$. The peaks of ρ_c are defined between cluster number $N_c = 2$ and 50 and Fig. C.8a shows the cluster results at the peaks between $N_c = 2$ and 26. At $N_c = 2$, the data are divided into two major groups with a boundary at a depth of 3.95 m. Thereafter, soil lenses separate up to $N_c = 7$ and a primary layer separates at $N_c = 9$ between 2.20 m and 3.95 m. Subsequently, no new primary clusters appear up to $N_c = 100$ as shown in Fig. C.8b which includes clustering between $N_c = 45$ and 100 every 5 increments. However,



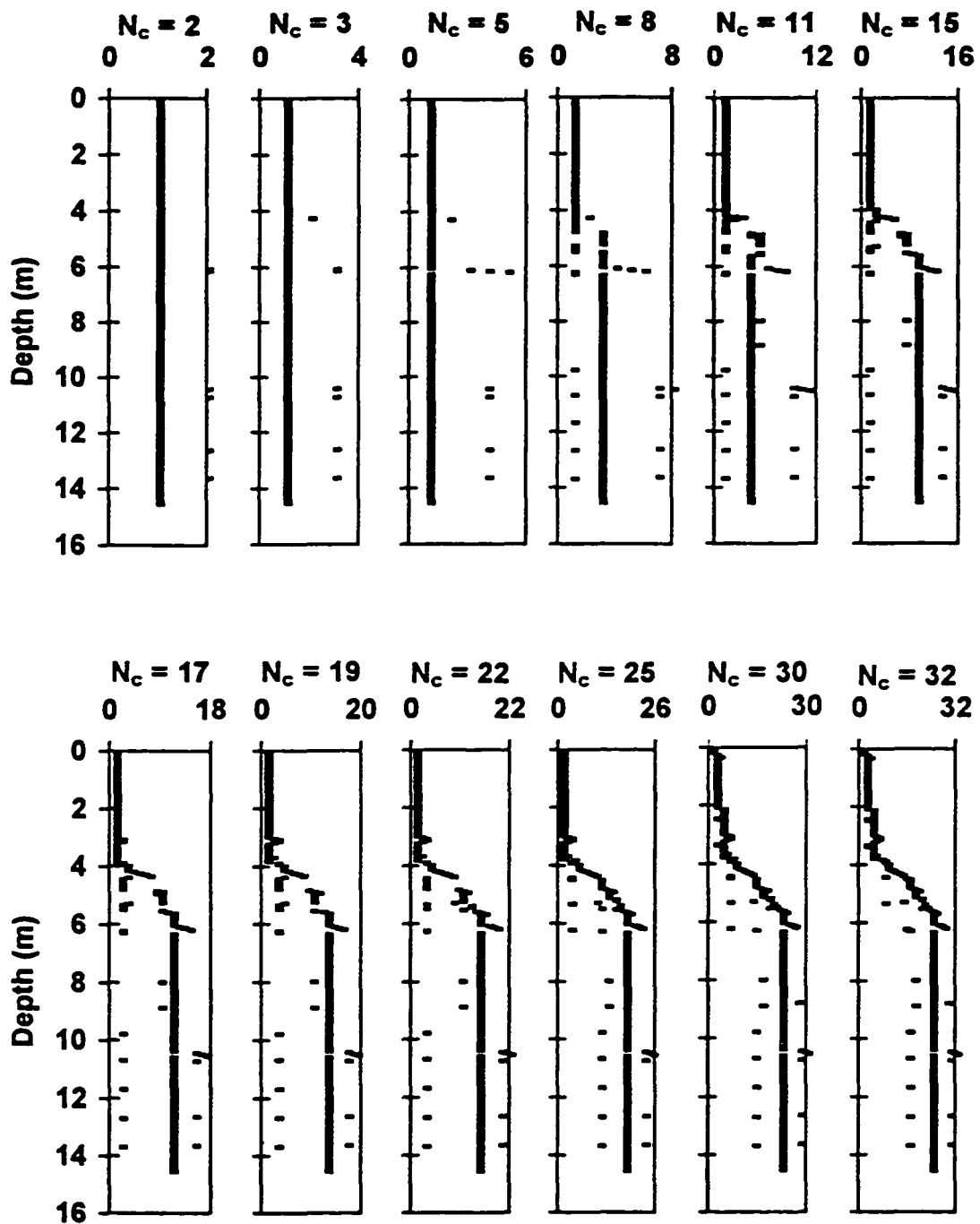
Single-Cosine-Zscore Method Using q_t and u_b . N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

**Figure C.4b. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
 Using q_t and the Ratio u_b/q_t .**



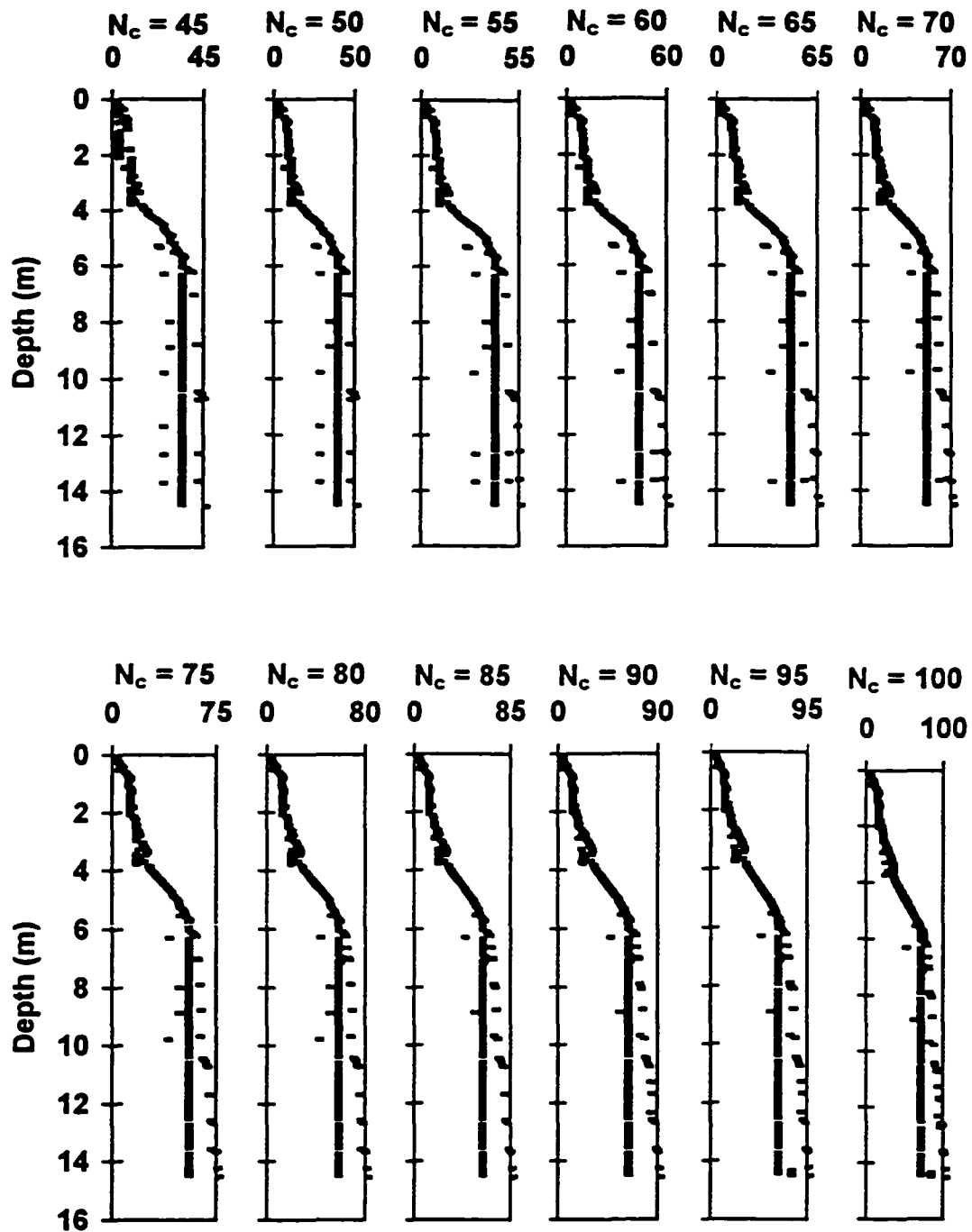
Single-Cosine-Zscore Method Using Q , B_q and F
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.5. Correlation Coefficient Between Consecutive Cluster Results at Amherst Using Normalized Parameters Q , B_q and F .



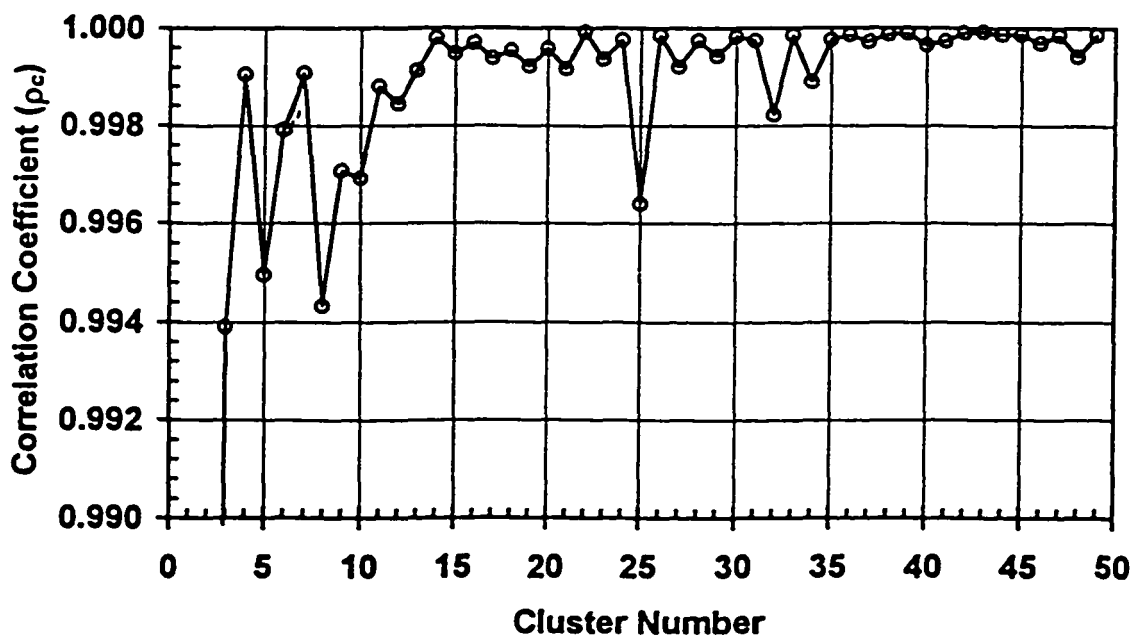
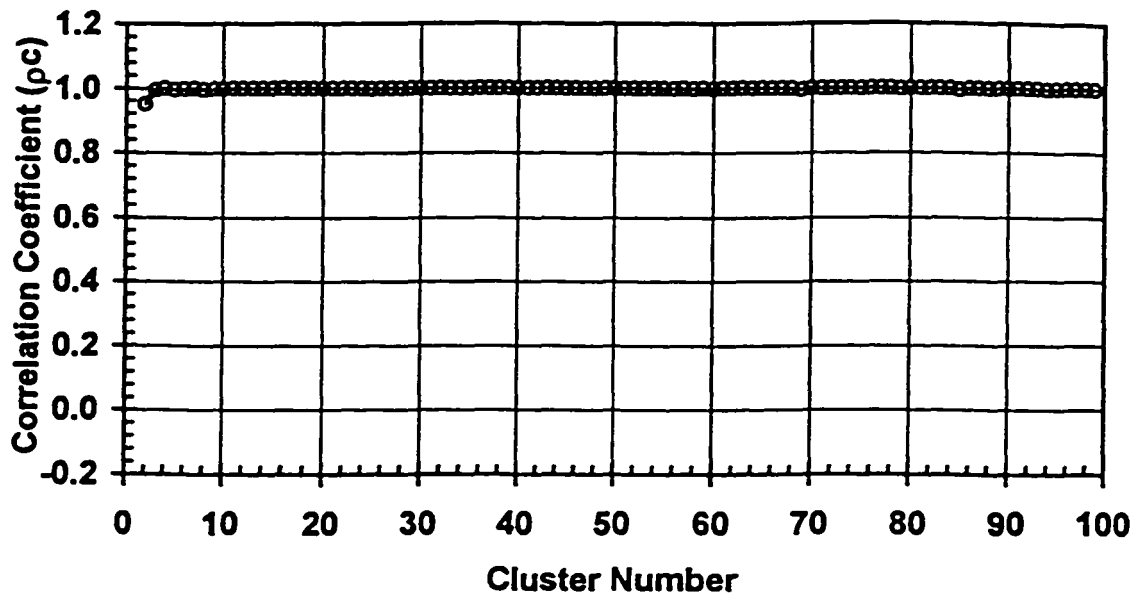
Single-Cosine-Zscore Method Using Q , B_q and F N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.6a. Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Normalized Parameters Q , B_q and F .



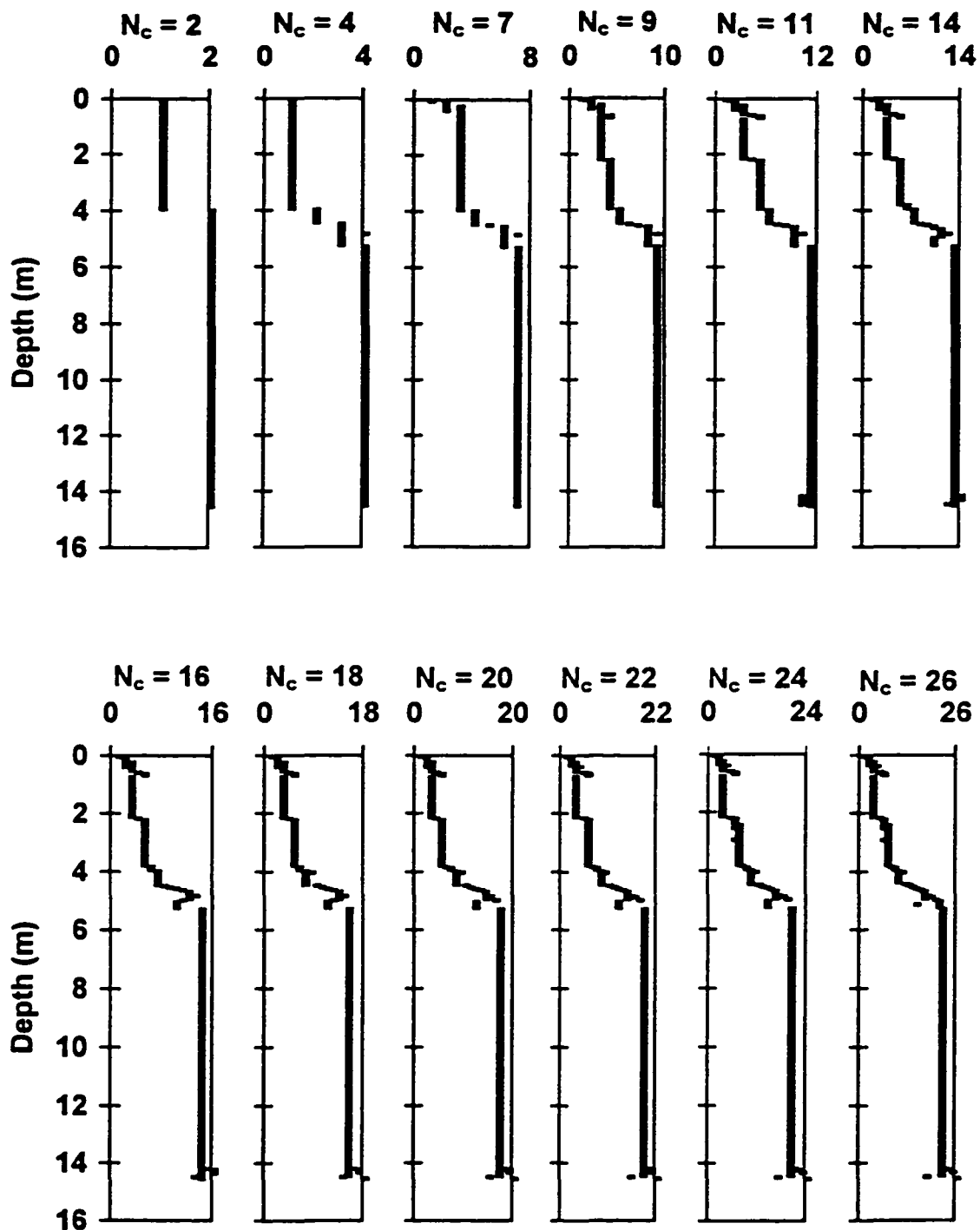
Single-Cosine-Zscore Method Using Q , B_q and F N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.6b. Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Normalized Parameters Q , B_q and F .



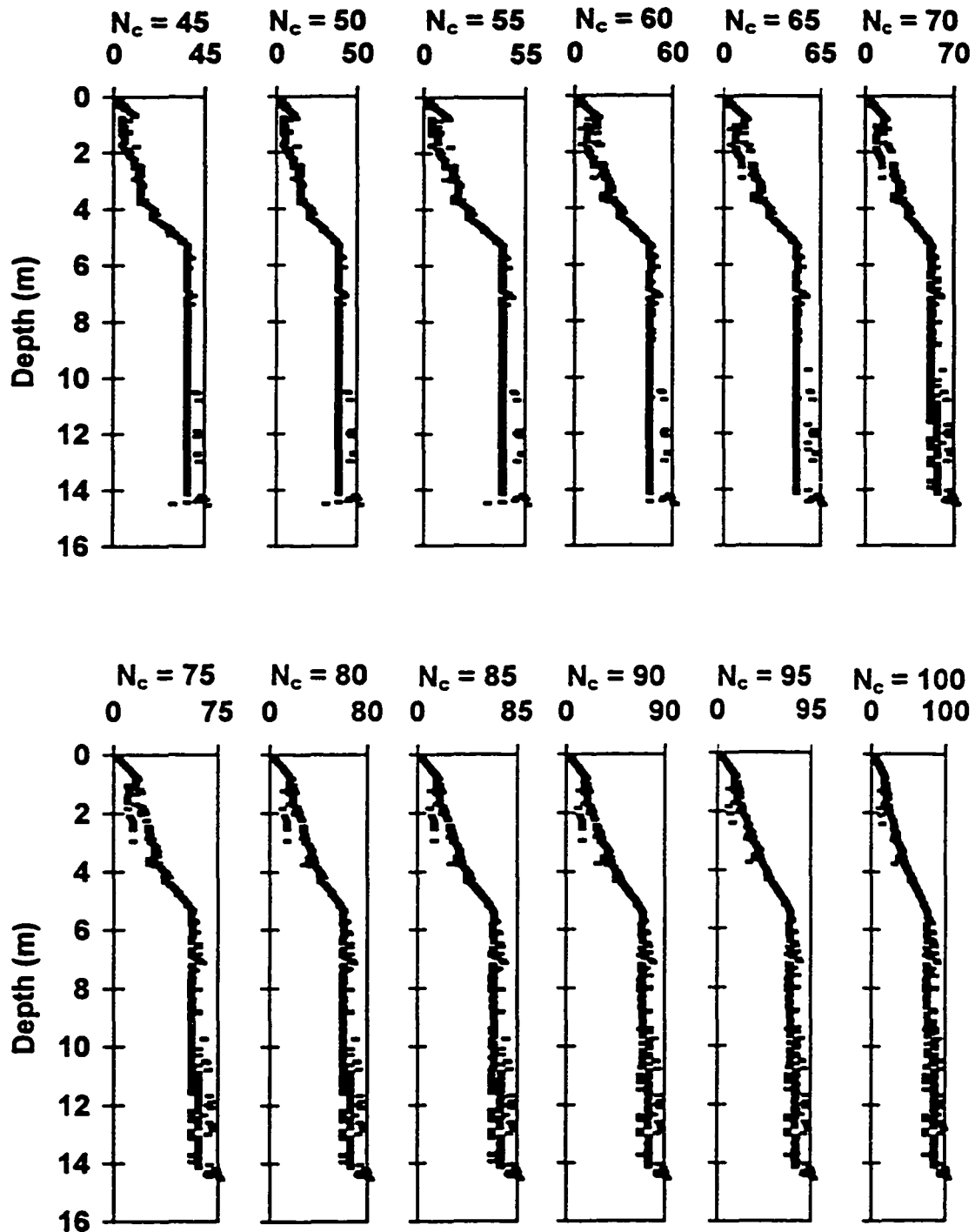
Single-Cosine-Zscore Method Using Q and B_q
 Note: Filtered PCPT data at Amherst, Massachusetts (data from this study).

Figure C.7. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Filtered Piezocone Data.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Filtered PCPT data at Amherst, Massachusetts (data from this study).

Figure C.8a. Cluster Analysis of Filtered Piezocone Data at Amherst.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Filtered piezocone data at Amherst, Massachusetts (this study).

Figure C.8b. Cluster Analysis of Filtered Piezocone Data at Amherst.

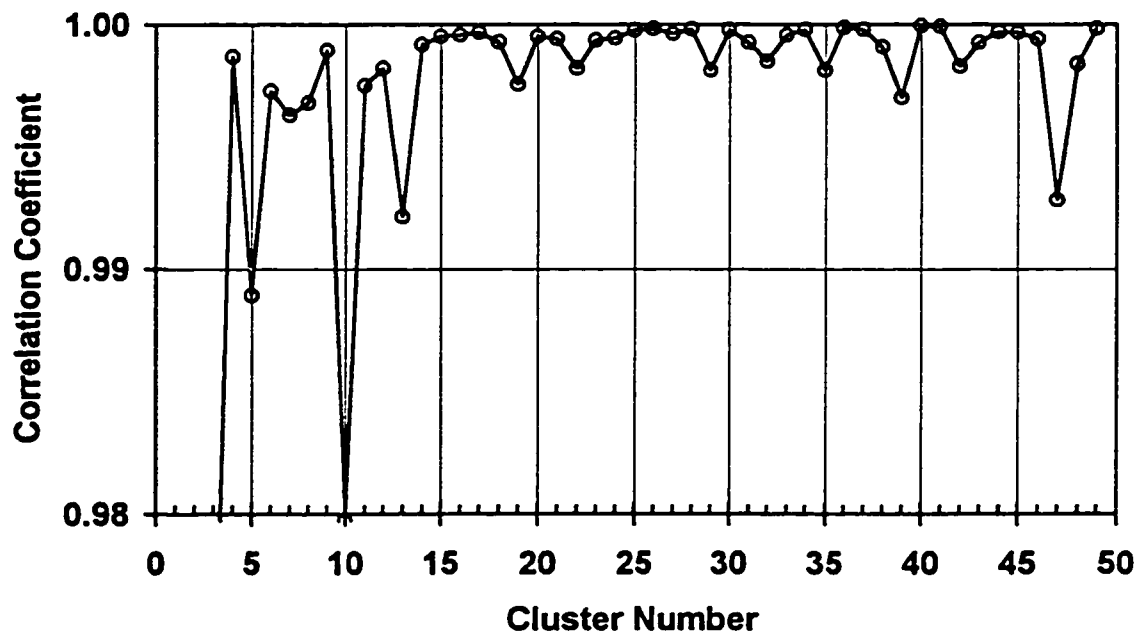
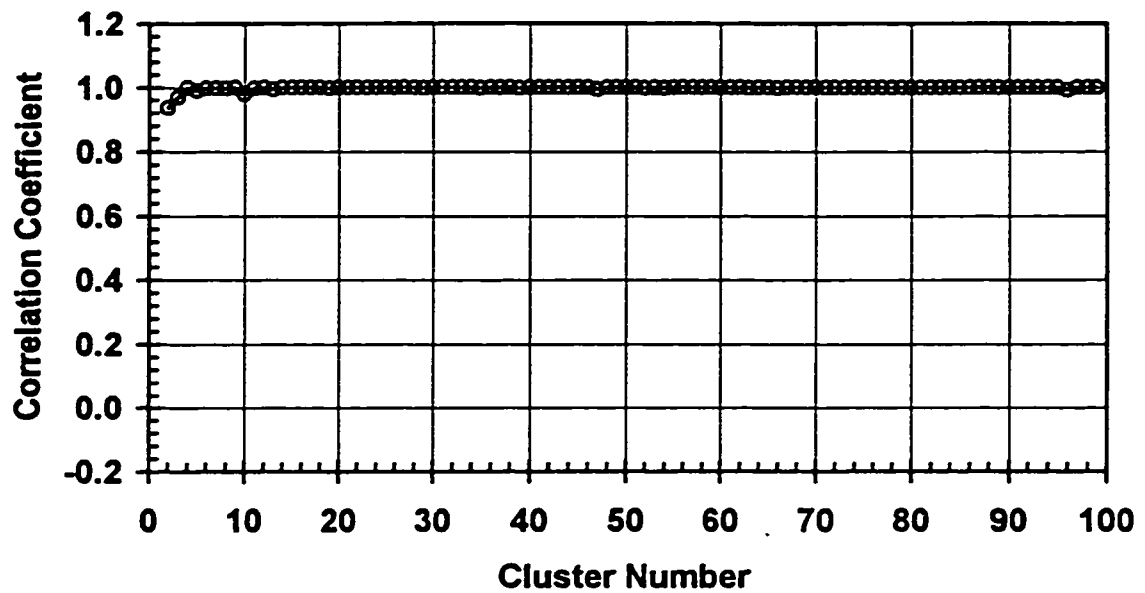
more growth of transition zones and lenses is indicated. Therefore a cluster number $N_c = 9$ is chosen to represent the subsurface stratigraphy.

C.4. Actual Versus Assumed Unit Weight

Total and effective stresses are calculated using measured unit weights at Amherst and the normalized parameters Q and B_q of sounding PCPT1 are derived as a function of actual γ_t . Then a single-cosine-zscore cluster method is used to divide the piezocone parameters Q and B_q into correlated groups. The correlation coefficient between consecutive cluster results is calculated up to cluster number $N_c = 100$ as shown in Fig. C.9. Examined clusters are chosen at the peaks of ρ_c and Fig. C.10a shows the cluster results at the peaks between $N_c = 2$ and $N_c = 36$. At $N_c = 2$, two major groups appear and delineate at a depth of 5.6 m. At $N_c = 4$, the upper cluster separates into 3 primary groups with boundaries at 2.2 m and 3.9 m. For $N_c > 4$, points indicating non-homogeneity continue to separate from the grouping and represent lenses, transitions, and/or data errors. Then the growth of clustering is shown in Fig. C.10b between $N_c = 45$ and 100 every 5 increments and no new primary layers ($t > 1$ m) are detected. Therefore, a cluster number 4 is chosen to represent the soil stratigraphy at the site.

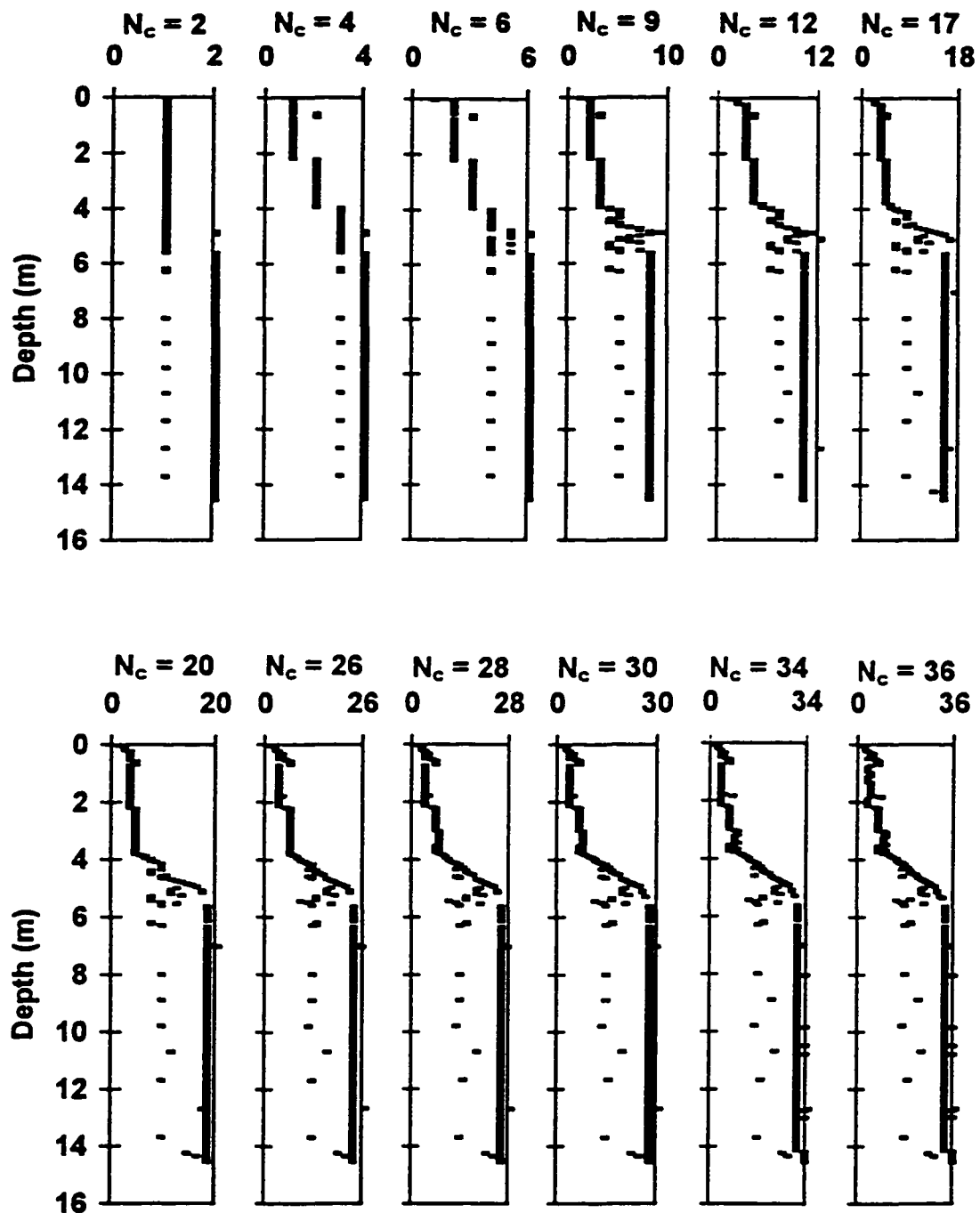
C.5. Data Frequency

The effect of data frequency is studied at Amherst using piezocone data of sounding PCPT1. Usually, the cone readings are collected every 1 cm to 5 cm, however, the purpose of this study is to indicate the possibility of using cluster methods to analyze



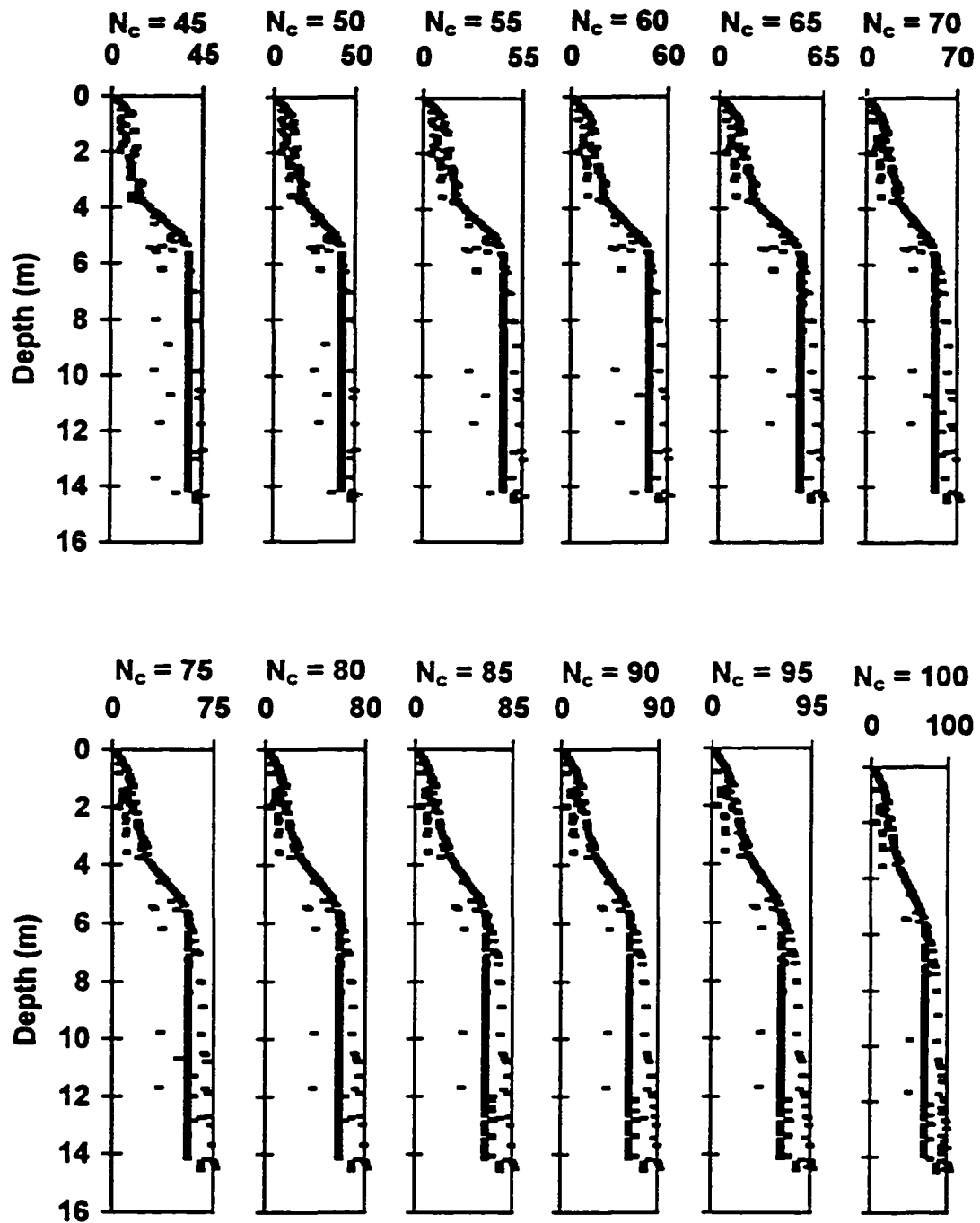
Single-Cosine-Zscore Method Using Q and B_q as a Function of Actual γ_t
Note: Piezocone data at Amherst (data from this study).

Figure C.9. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts.



Single-Cosine-Zscore Method Using Q and B_q as a Function of Actual γ_t
 Note: Piezocone data at Amherst (data from this study). N_c = No. of Clusters

Figure C.10a. Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Actual Unit weight.

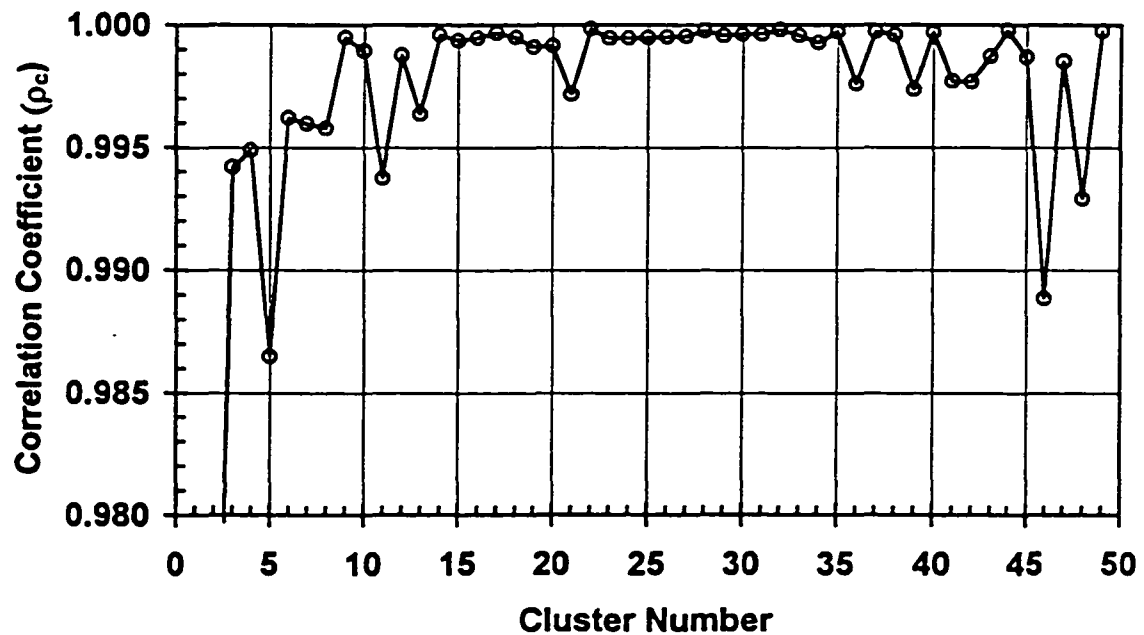
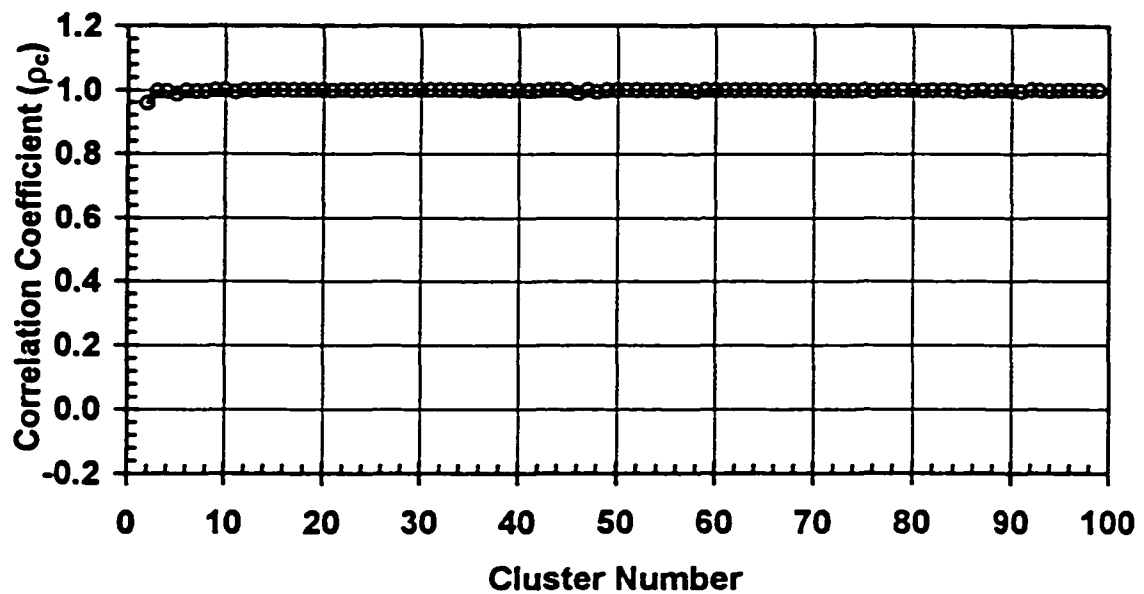


Single-Cosine-Zscore Method Using Q and B_q as a Function of Actual γ_t
 Note: Piezocone data at Amherst (this study). N_c = No. of Clusters

**Figure C.10b. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
Using Actual Unit weight.**

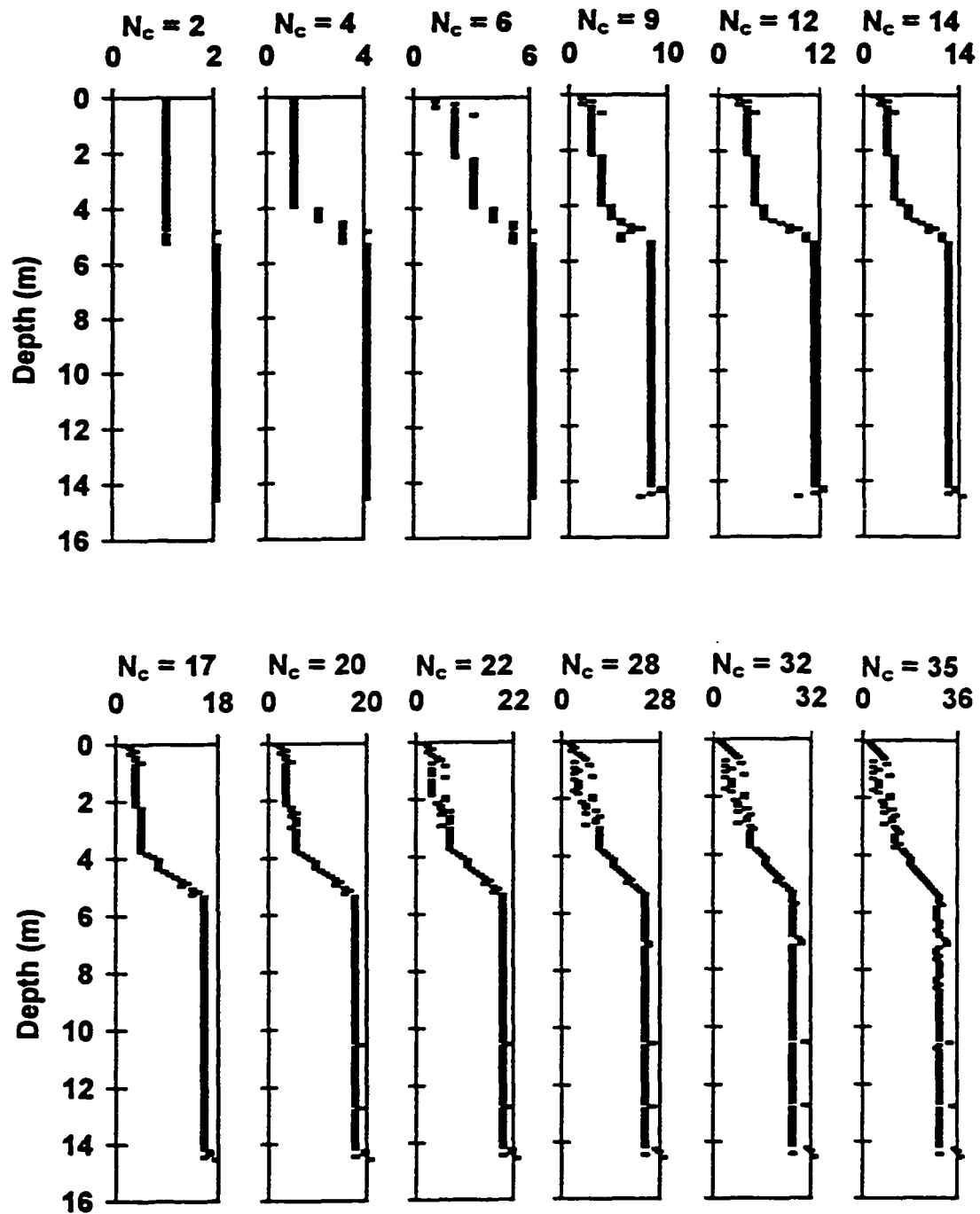
other laboratory and in-situ testing collected at larger distances. Filtered piezocone data based on the Vivitrat (1978) method are used to avoid the dissipation effect during adding cone-rods. The reason for filtering the data is because that the data errors due to dissipation effect do not exist in the case of laboratory and field testing other than piezocone testing. Five different frequencies are used including 5 cm, 10 cm, 20 cm, 30 cm, 40 cm and 50 cm. The growth of the clustering using the first frequency (5 cm) is discussed in detail in a previous section in this appendix.

A single-cosine-zscore method is used to delineate the data into correlated groups. For frequency = 10 cm, the correlation coefficient between consecutive clusters are shown in Fig. C.11 up to cluster number $N_c = 100$. Cluster results are examined at the peaks of ρ_c up to $N_c = 50$ and Figure C.12a shows data groups between $N_c = 2$ and 35. At cluster number 2, the data are divided into two groups with a boundary at 5.6 m. At cluster number 4, a transition layer appear between 3.9 m and 5.6 m and then at cluster number 6, the upper main cluster is separated into two primary clusters at a boundary depth of 2.2 m. Subsequently, no new primary layers appear for higher clusters, however, some points continue to separate indicating non homogeneity within the main statistical groups of data. Also, this statement is valid up to $N_c = 100$ as shown in Fig. C.12b which includes clustering between cluster numbers 45 and 100 every 5 increments. Therefore cluster number 6 is chosen to represent the subsurface stratigraphy. It suggests more association in soil type and/or properties between the upper two clusters down to 3.9 m than that below 5.6 m.



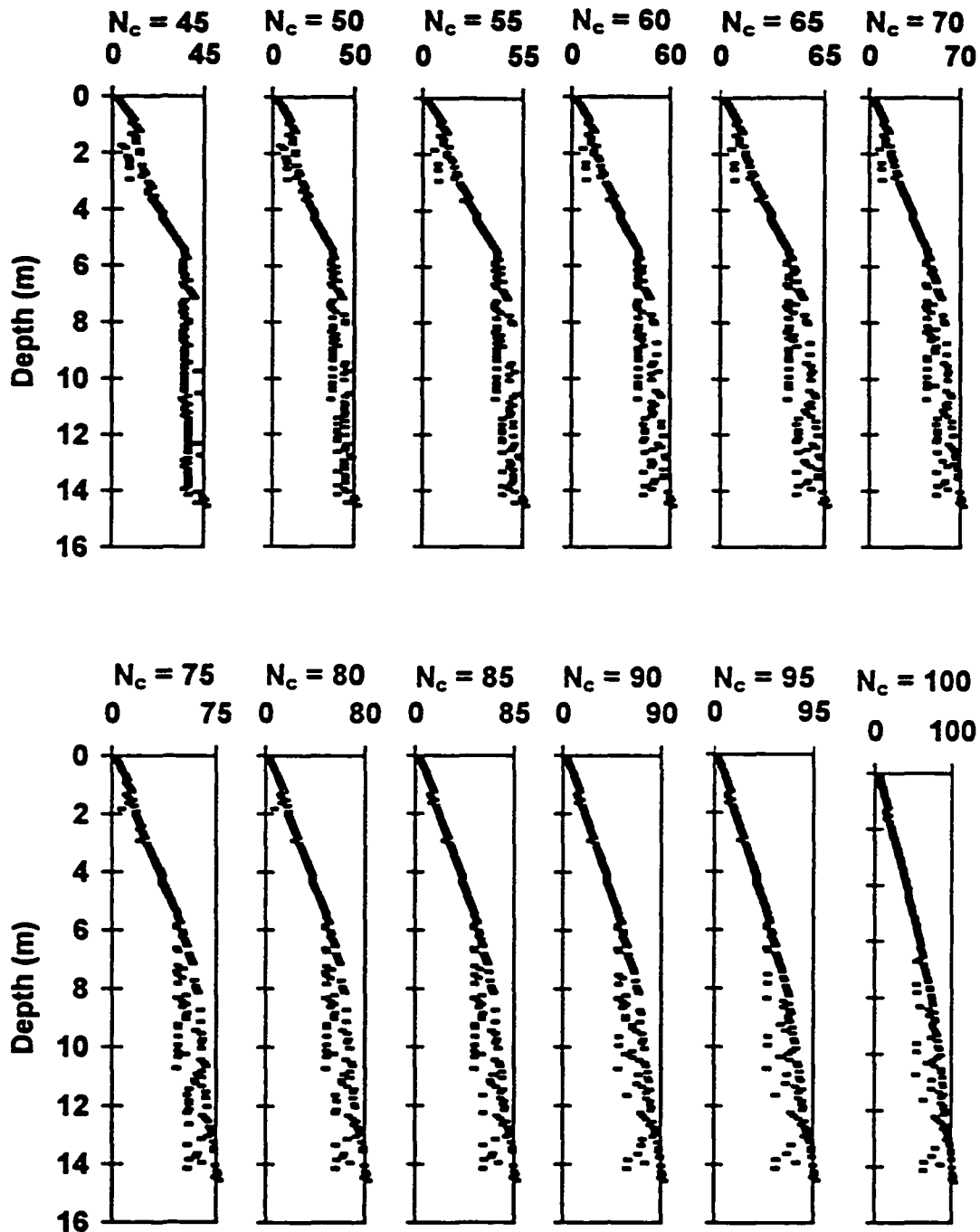
Single-Cosine-Zscore Method Using Q and B_q
Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.11. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Frequency = 10 cm.



Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.12a. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
 Using Data Frequency = 10 cm.



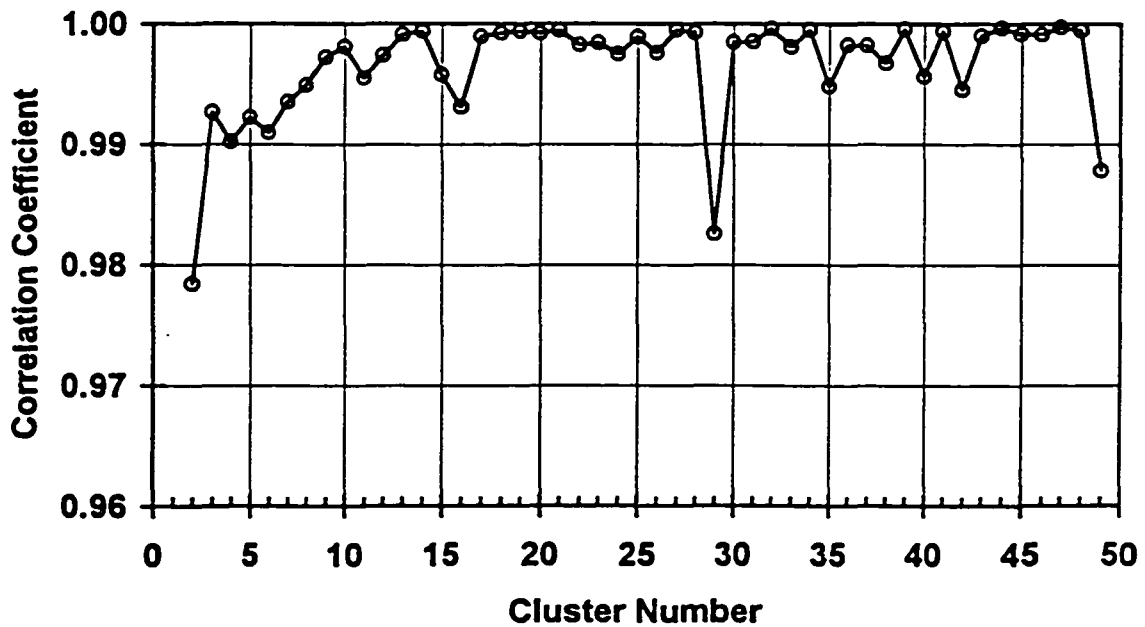
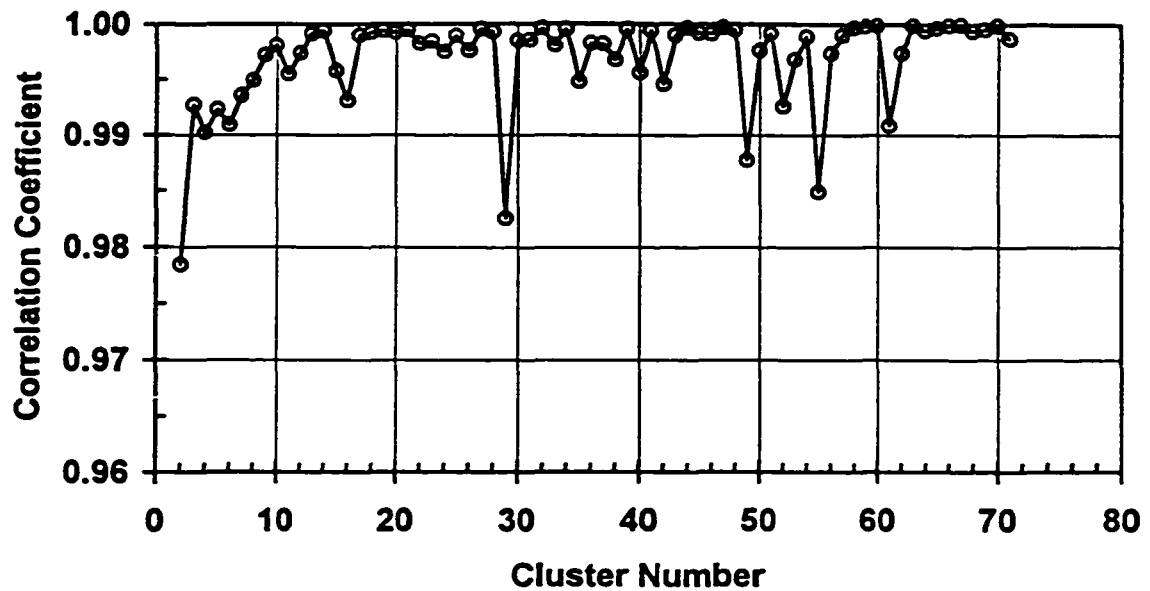
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (this study).

Figure C.12b. Cluster Analysis of Piezocone Data at Amherst, Massachusetts Using Data Frequency = 10 cm.

For frequency = 20 cm, the correlation coefficient between consecutive clusters is shown in Fig. C.13 up to cluster number $N_c = 73$ (number of data points). The cluster results are examined at the peaks of ρ_c up to $N_c = 50$ and Figure C.14a shows clustering between $N_c = 2$ and 36. At cluster number 2, the data are divided into two primary groups with a boundary at a depth of 5.6 m. For $N_c \geq 3$, a transition zone started to build up between the two main clusters. At cluster number 7, the upper statistical layer is separated into two primary layers at a depth of 2.1 m. At cluster number 16, the chain of the data below 5.6 m breaks into two associated layers at a depth of 11.2 m. The latter boundary is not supported by other back-up laboratory or field data. The reason for that might be due to decreasing the number of data from 291 points at 5 cm frequency to 73 points at 20 cm frequency.

For clarification, consider a data set of two points which belong to the same soil layer, each point would be assigned a different cluster number at $N_c = 2$. This would suggest to keep the cluster number as a minimum not to get unverified layers in a subsurface stratigraphy. For data frequency = 20 cm, a maximum considered number of data groups could be equal to 10. Subsequently, no new primary layers appear for higher clusters up to $N_c = 73$ as shown in Fig. C.14b which includes clustering between $N_c = 45$ and 70 every increment of 5 clusters. Cluster number 7 is chosen to represent the subsurface stratification.

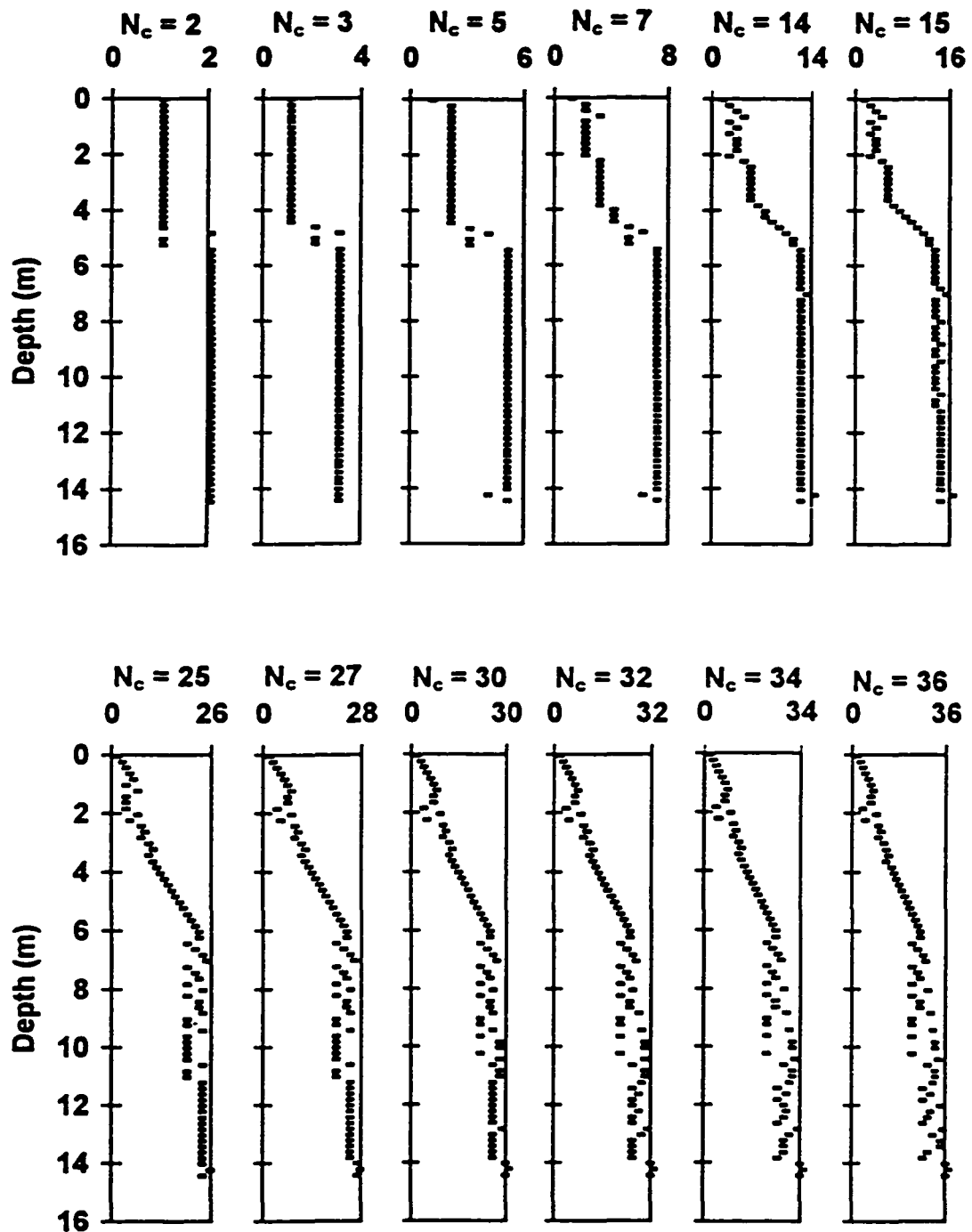
For data frequency = 30 cm, the correlation coefficient between consecutive clusters is shown in Fig. C.15 up to $N_c = 49$ (number of data points). Cluster results are



Single-Cosine-Zscore Method Using Q and B_q

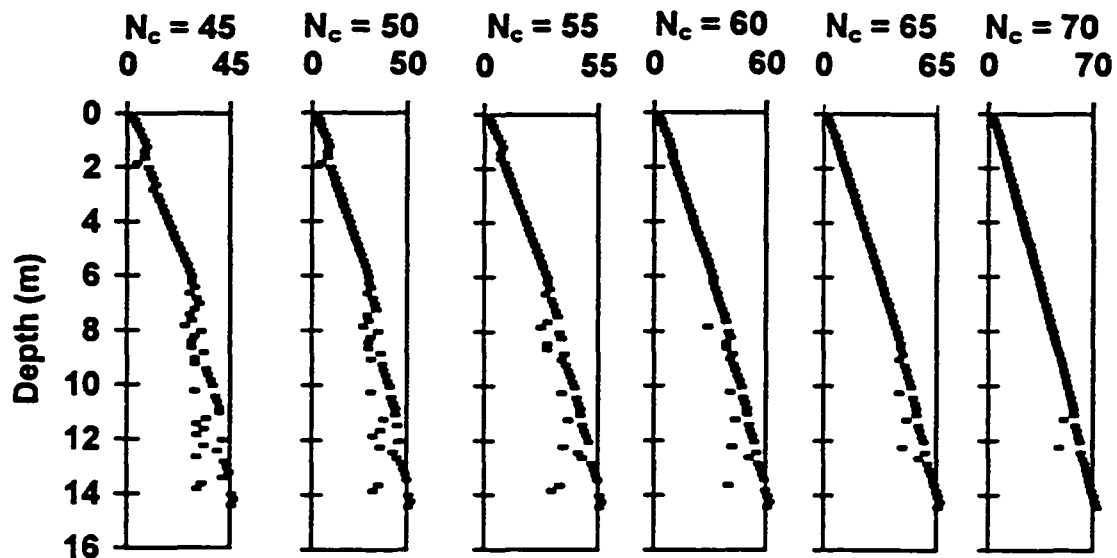
Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.13. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Frequency = 20 cm.



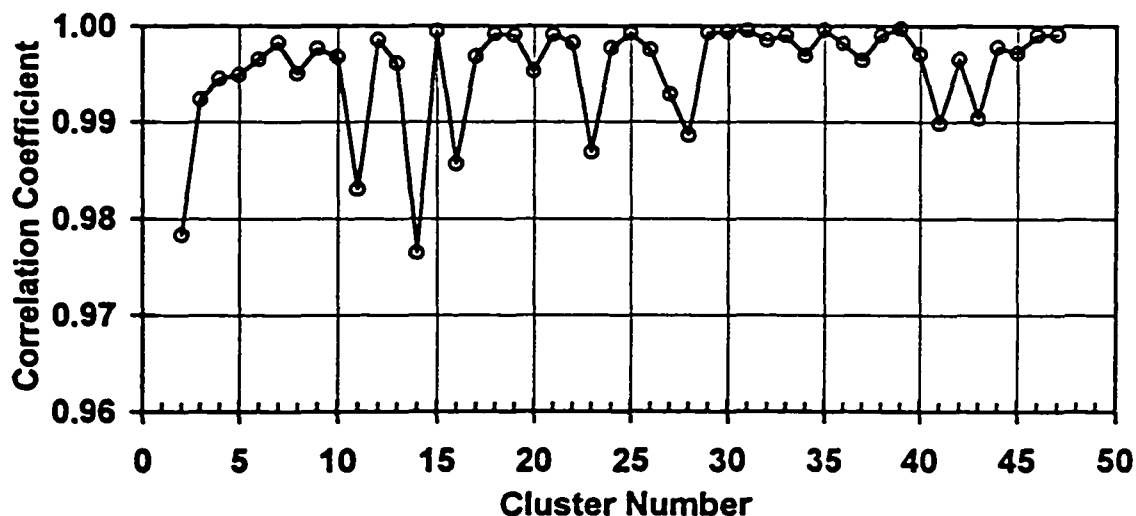
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

**Figure C.14a. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
 Using Data Frequency = 20 cm.**



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.14b. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
 Using Data Frequency = 20 cm.



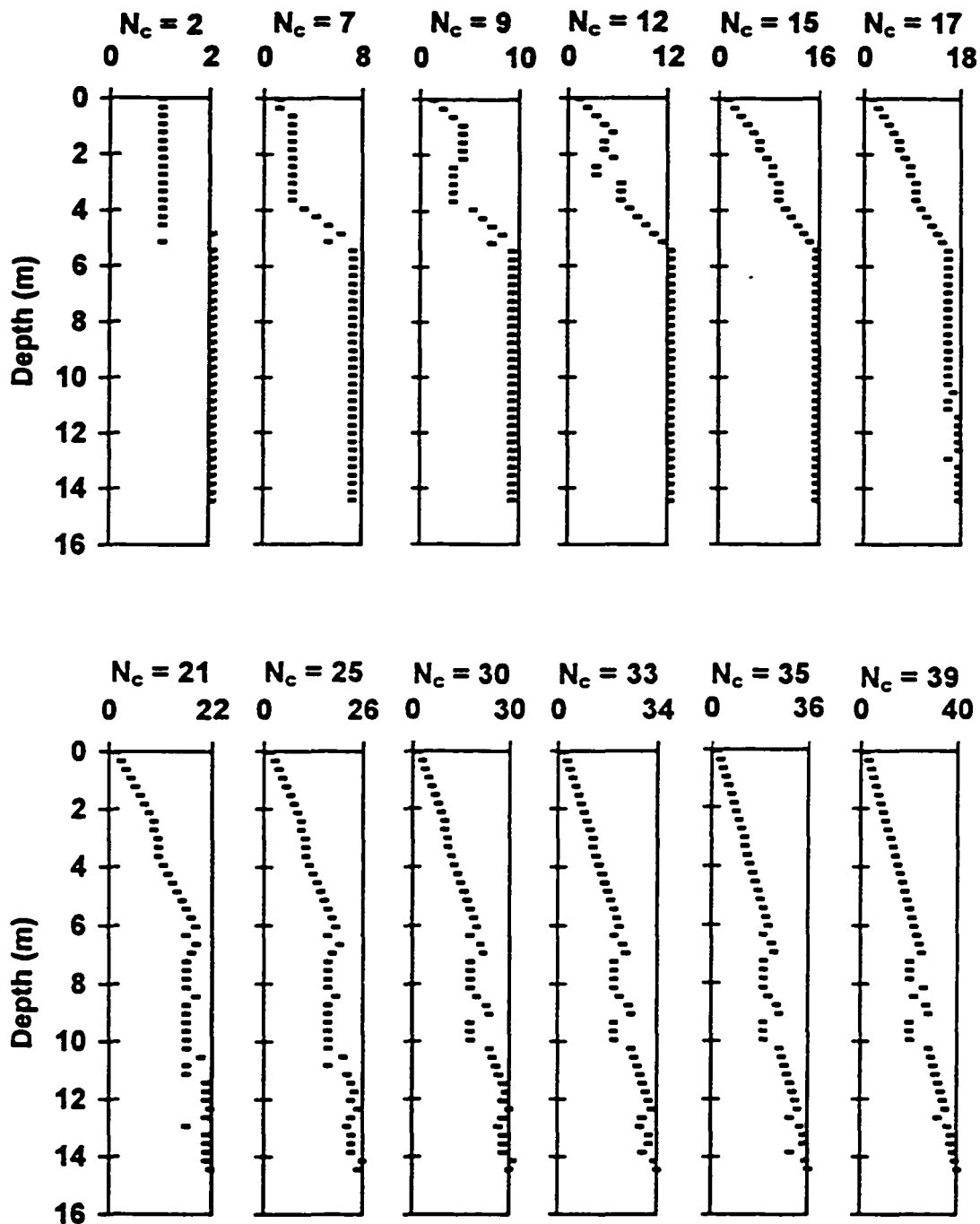
Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data at Amherst, Massachusetts (this study).

Figure C.15. Correlation Coefficient Between Consecutive Cluster Results at
 Amherst, Massachusetts Using Data Frequency = 30 cm.

examined at the peaks of ρ_c and shown in Fig. C.16 between cluster number $N_c = 2$ and 39. At cluster number 2, data are divided into two major groups at a depth of 5.6 m. For higher clusters, a transition zone appear , for instance, between the depths of 3.9 m and 5.6 m at cluster number 7. At cluster number 9, the upper cluster is separated into two main clusters at a depth of 2.2 m.

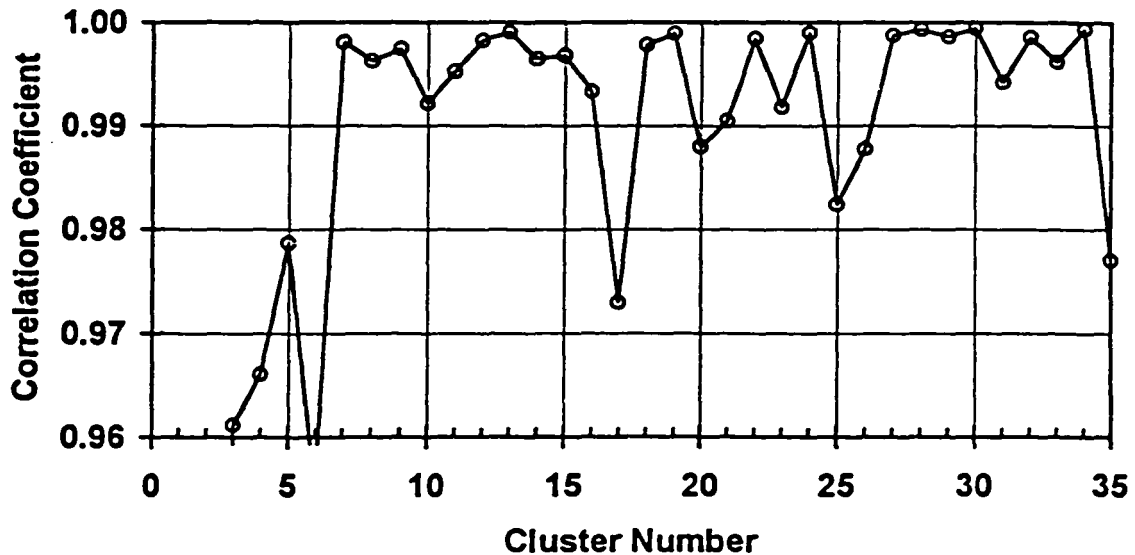
At cluster number 17, the cluster below a depth of 5.6 m is divided into primary groups at a depth of 11.2 m, however, this boundary is not verified by other laboratory or field data. The last two layers are strongly associated in terms of soil type and/or properties because they are denoted cluster numbers 16 and 17 and the separation occur at a relatively large cluster number (17) compared with the number of data points (49). For higher clusters, no new primary layers are detected, however, more points continue to separate from the grouping until each point is assigned a single cluster number at $N_c = 49$. In this case, cluster number 9 is chosen to represent the subsurface stratigraphy considering cluster number 10 is the cut-off of the analysis.

For frequency = 40 cm, the correlation coefficient between consecutive clusters is shown in Fig. C.17 up to cluster number $N_c = 37$ (number of data points). Cluster results are examined at the peaks of ρ_c and Figure C.18 shows them between $N_c = 2$ and 32. At cluster number 2, the data are divided into two groups delineated at a depth of 5.6 m. A transition zone grows up, for instance, between the depths of 3.9 m and 5.6 m at cluster



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (this study).

**Figure C.16. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
 Using Data Frequency = 30 cm.**



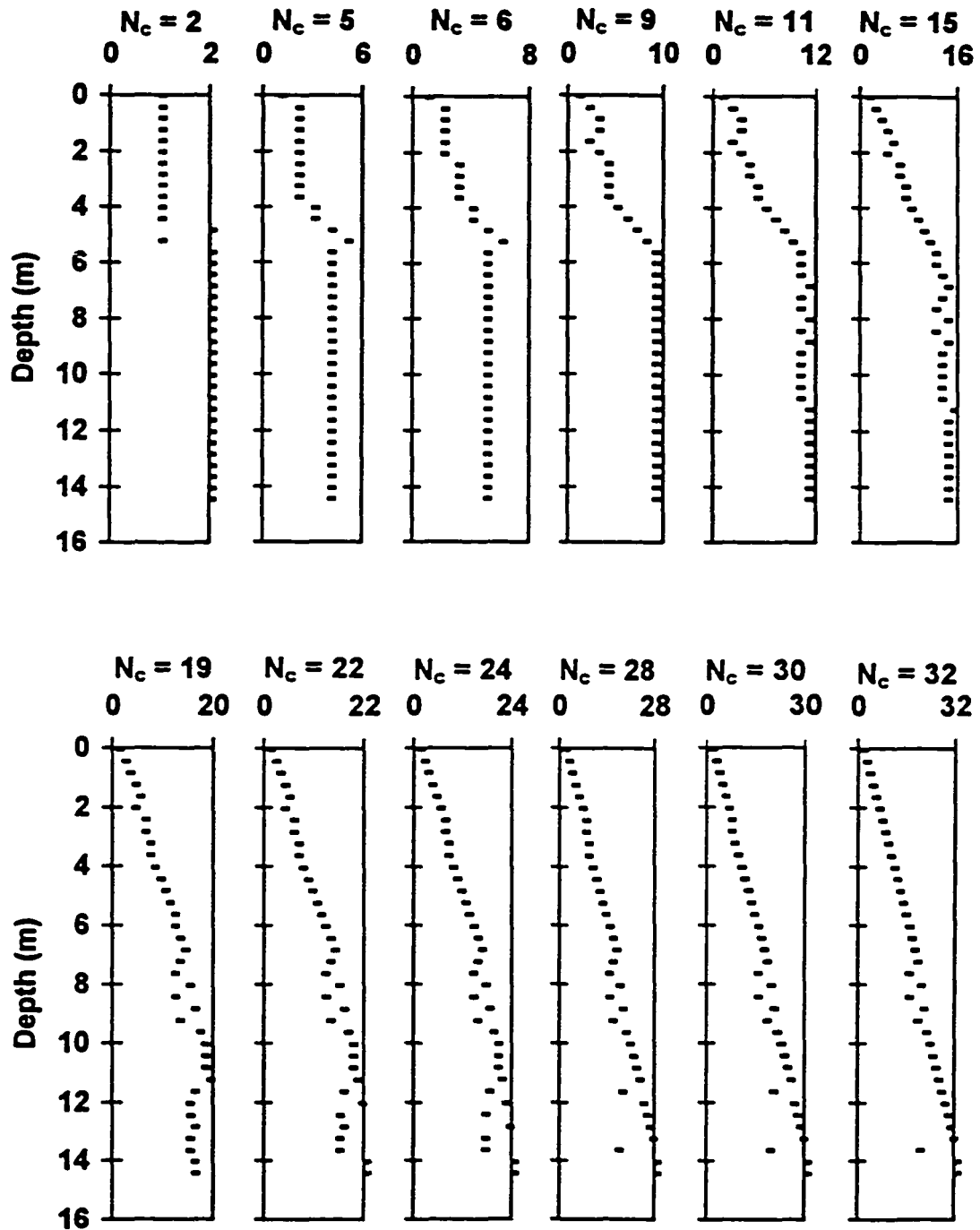
Single-Cosine-Zscore Method Using Q and B_q

Note: Piezocone data at Amherst, Massachusetts (data from this study).

Figure C.17. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Frequency = 40 cm.

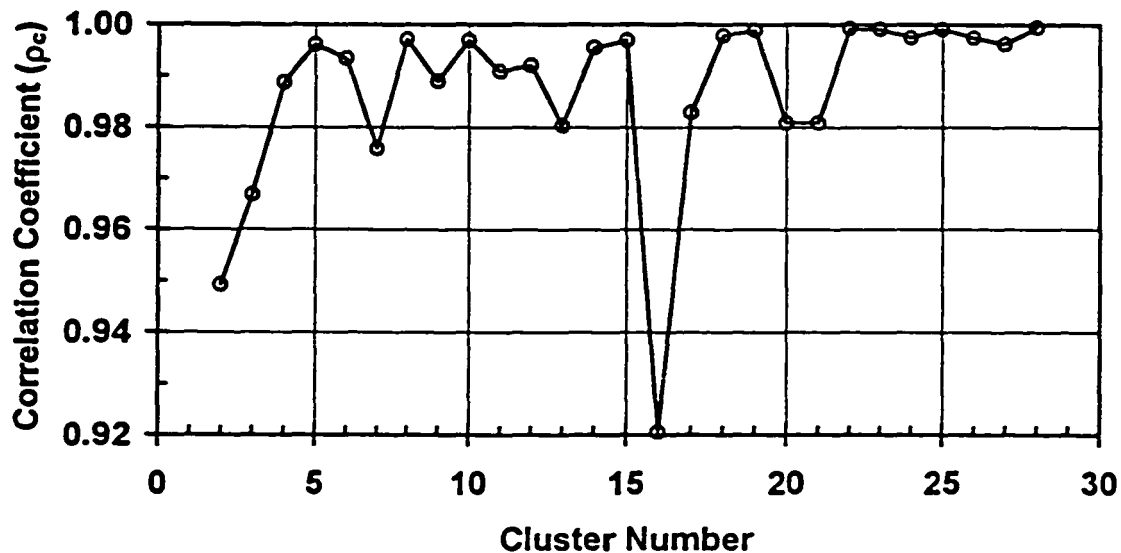
number 5. At cluster number 6, the upper layer separates into primary layers at a depth of 2.2 m. At cluster number 11, the layer below a depth of 5.6 m is divided into groups at a depth of 11.2 m. The latter two layers are assigned cluster numbers equal to 10 and 11 and the separation relatively occur at a large cluster number, therefore, they are related in terms of soil type and/or properties. Subsequently, data points continue to separate until each point forms a cluster at $N_c = 37$. By ignoring the statistical boundary at 11.2 m due to lack of data, cluster number 6 is chosen to represent the subsurface stratification.

For data frequency = 50 cm, the correlation coefficient between consecutive clusters is shown in Fig. C.19 up to $N_c = 30$ (number of data points). Clustering is



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

**Figure C.18. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
 Using Data Frequency = 40 cm.**



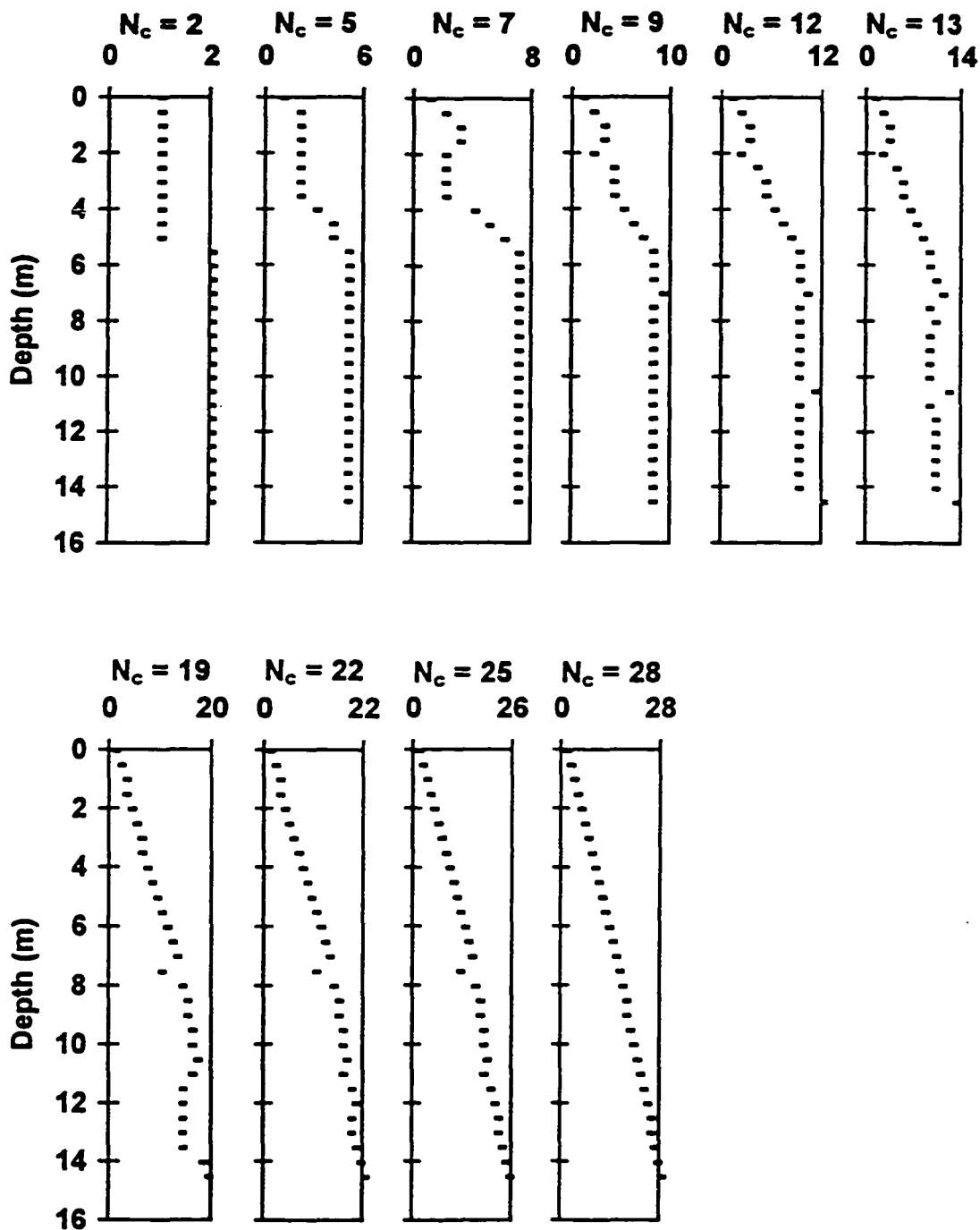
Single-Cosine-Zscore Method Using Q and B_q

Note: Piezocone data at Amherst, Massachusetts (this study).

Figure C.19. Correlation Coefficient Between Consecutive Cluster Results at Amherst, Massachusetts Using Data Frequency = 50 cm.

examined at the peaks of ρ_c and Figure C.20 shows the results between $N_c = 2$ and 30. At cluster number 2, the data are divided into two major groups at a depth of 5.6 m. A transition zone appear between the depths of 3.9 m and 5.6 m at cluster number 5. At cluster number 7, the upper layer separates into two primary layers at a depth of 2.2 m. For higher cluster numbers, points continue to depart the main clusters until each point is assigned a cluster number at $N_c = 30$. At a relatively large cluster ($N_c = 13$), a statistical boundary is delineated at 11.2 m, however, it is ignored due to lack of evidence. Cluster number 7 is chosen to represent the subsurface stratification.

The frequency effect is studied up to 50 cm to distinguish between soil lenses and transitions ($t < 0.5$ m) and soil layers ($t \geq 0.5$ m) according to the criterion defined in



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Amherst, Massachusetts (data from this study).

**Figure C.20. Cluster Analysis of Piezocone Data at Amherst, Massachusetts
 Using Data Frequency = 50 cm.**

Chapter 4. The previous analyses suggests to put a restriction on the upper limit of examined cluster results based on available number of data to avoid unsupported soil layers and boundaries. In the discussed examples, a cluster number $N_c = 10$ is suitable to be a cut-off in case of data frequency between 20 cm (data points = 73) and 50 cm (data points = 30).

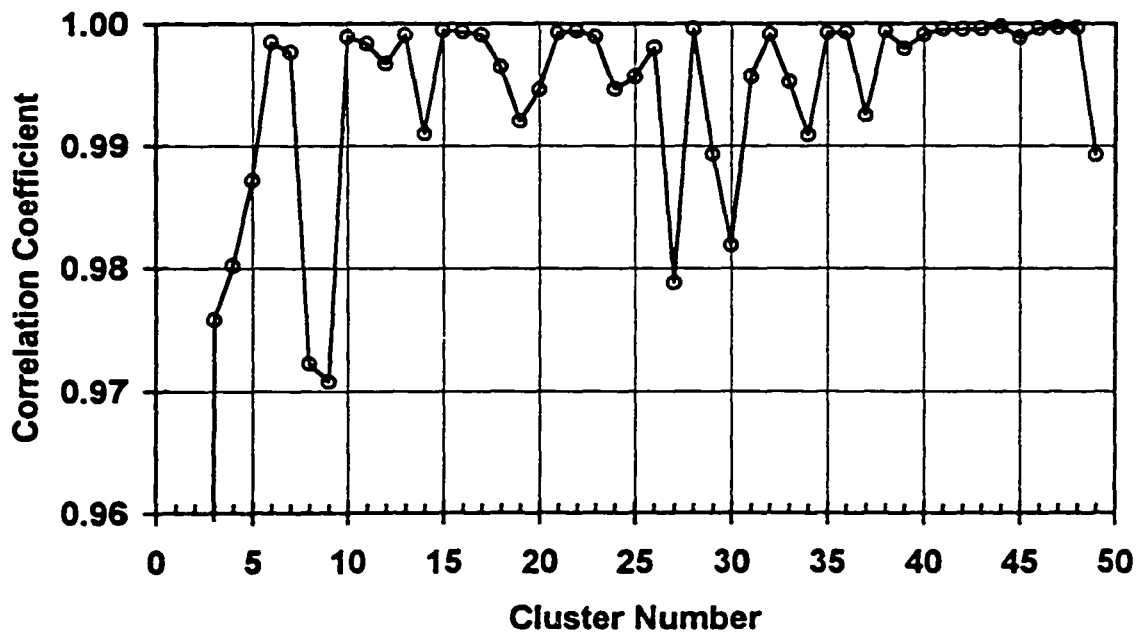
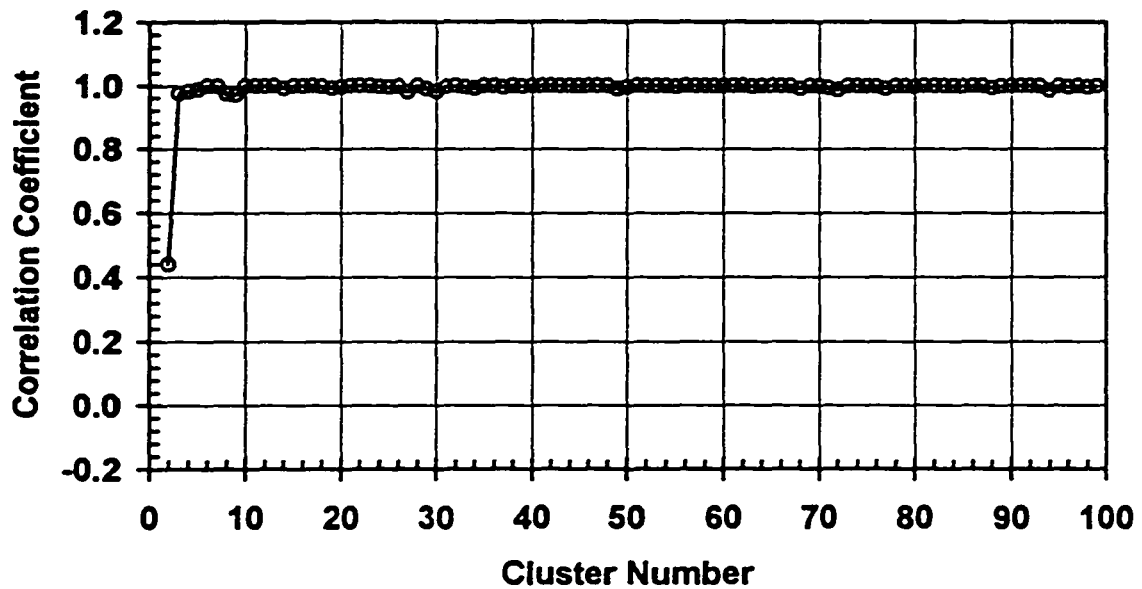
C.6. Porewater Position

The position of the porous element at the face or behind the cone tip has a significant effect on the measured porewater pressure. Two case studies are discussed in this section to evaluate the effect of u_t and u_b measurements on the growth of the cluster analysis. First, piezocone data from Nash et al., (1992) were collected in soft clays at Bothkennar, Scotland where both u_t and u_b readings were positive. Second, piezocone data from Powell et al. (1988) were collected in stiff clays at Brent Cross, UK where u_t readings were negative, however, u_b readings were positive. These facets of u_t and u_b penetration porewater pressures are characteristic in clays (Mayne et al., 1990). A single-cosine-zscore method is applied to normalized parameters Q and B_q at both sites.

C.6.1. Bothkennar Test Site, Scotland

Piezocene data at the Bothkennar test site were analyzed using SCZ clustering technique, and presented previously in Fig. 5.16 (Nash et al., 1992). The correlation coefficient between consecutive clusters of Q and B_{q2} is shown in Fig. C.21 up to cluster number $N_c = 100$. Clustering is examined at the peaks of ρ_c and Figure C.22a shows the cluster results between $N_c = 2$ and 38. At cluster number 3, the data are divided into two primary clusters with a boundary at a depth of 3.7 m. At higher clusters, a transition zone appear between the two layers between the depths of 3.7 m and 4.5 m. Soil lenses and outliers continue to separate from the two main layers indicating non homogeneity. No more primary layers are detected ($t > 1$ m) up to cluster number 100 as shown in Fig. C.22b which includes cluster results between $N_c = 45$ and 100 every 5 increments.

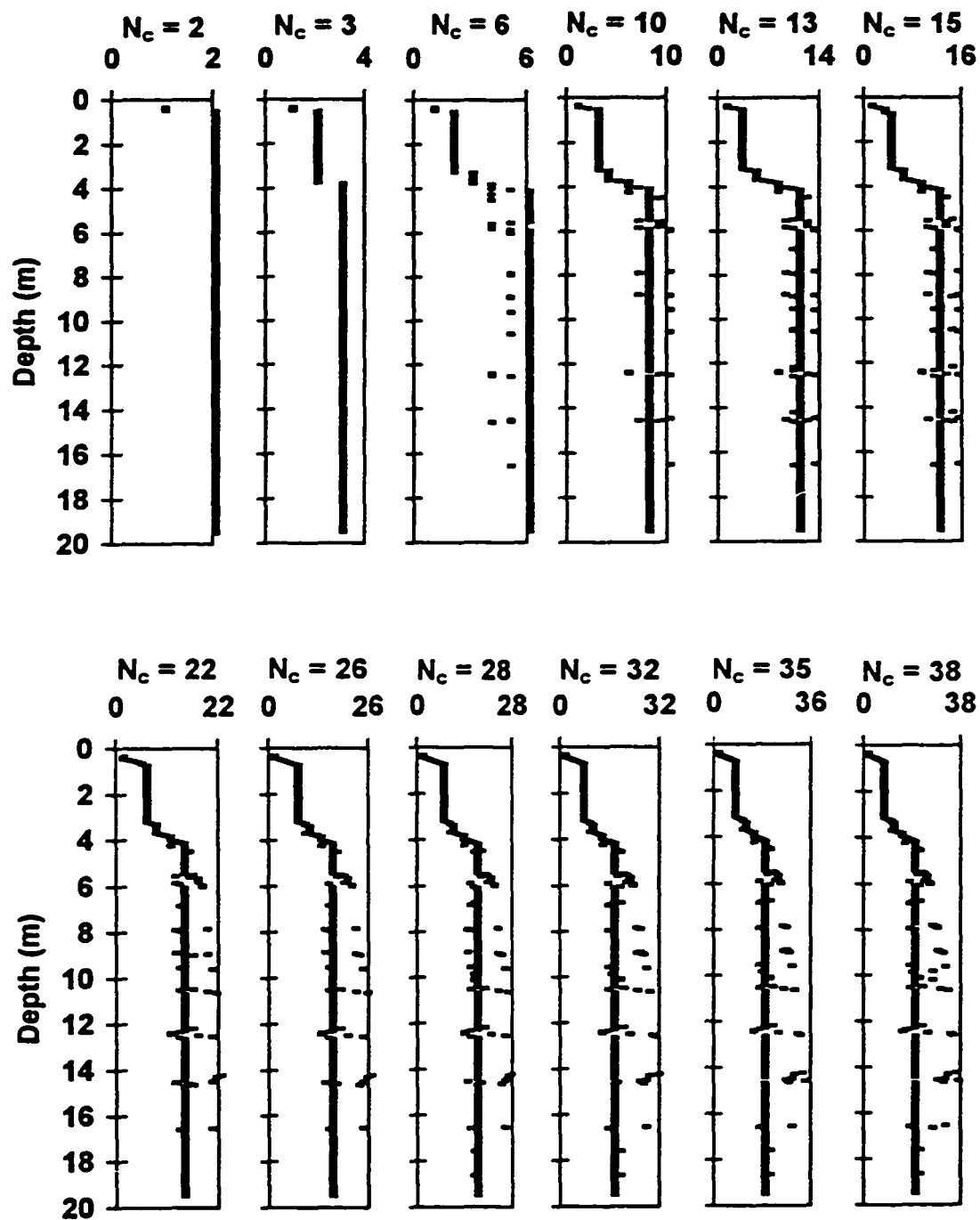
A similar analysis is performed using normalized parameters Q and B_{q1} and the correlation coefficient between consecutive clusters is shown in Fig. C.23. Cluster results are studied at the peaks of ρ_c and Fig. C.24a shows them between cluster number $N_c = 2$ and 31. At cluster number 4, the data are divided into two groups with a boundary at a depth of 3.7 m. Also, a transition zone appears between the depths of 3.7 m and 4.4 m. For higher clusters, no new primary layers are discovered, however, points continue to separate indicating lenses or outliers within the major statistical layers. Also, this statement applies to clustering up to 100 as shown in Fig. C.24b which includes the cluster results between $N_c = 45$ and 100 every 5 increments. Therefore, a cluster number 4 is chosen to represent the subsurface stratification. Note that, the same soil stratigraphy was obtained using u_t or u_b in clustering.



Single-Cosine-Zscore Method Using Q and B_{q2}

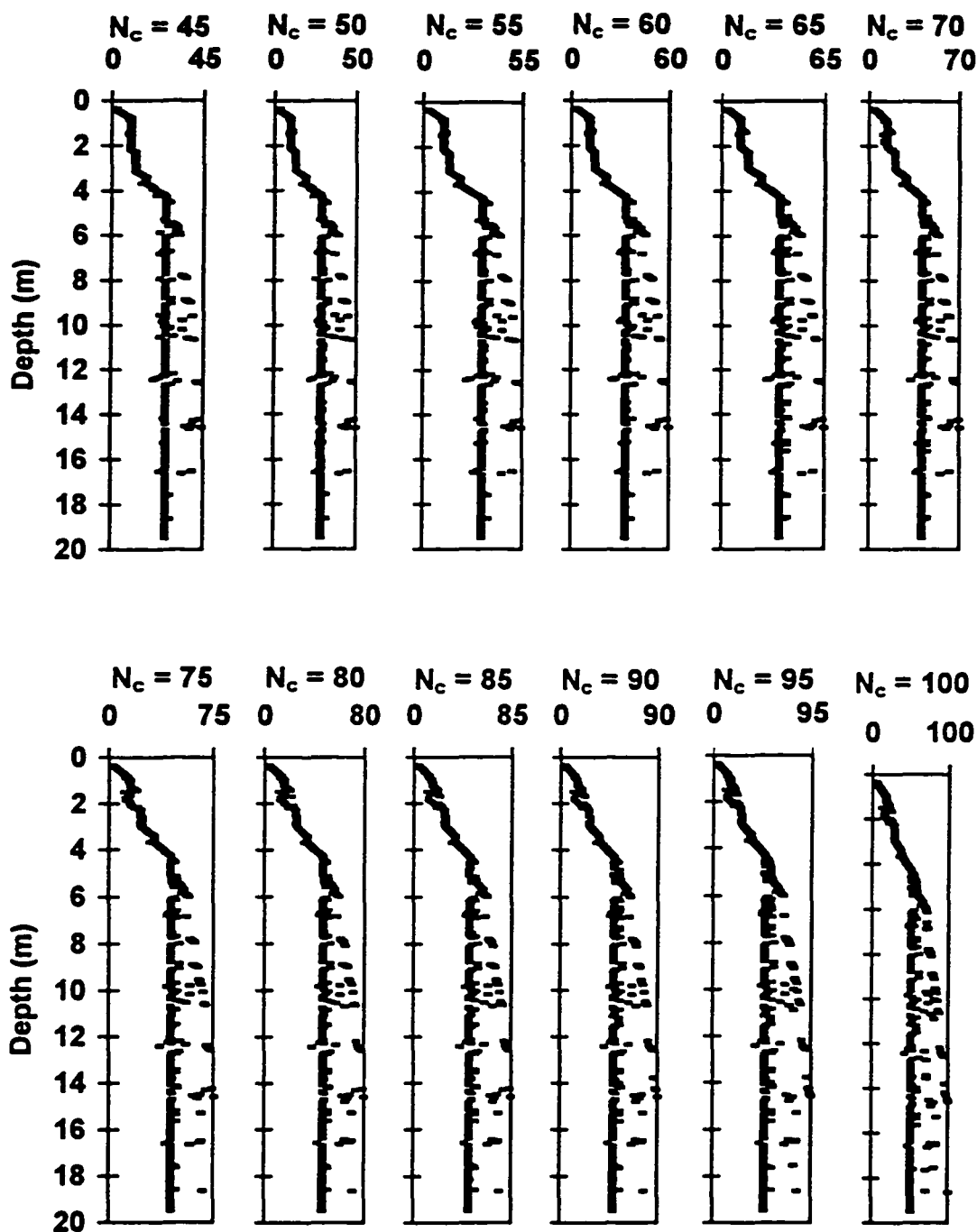
Note: Piezocone data at Bothkennar, Scotland (Nash et al., 1992).

Figure C.21. Correlation Coefficient Between Consecutive Cluster Results at Bothkennar, Scotland Using Type 2 Piezocone Data.



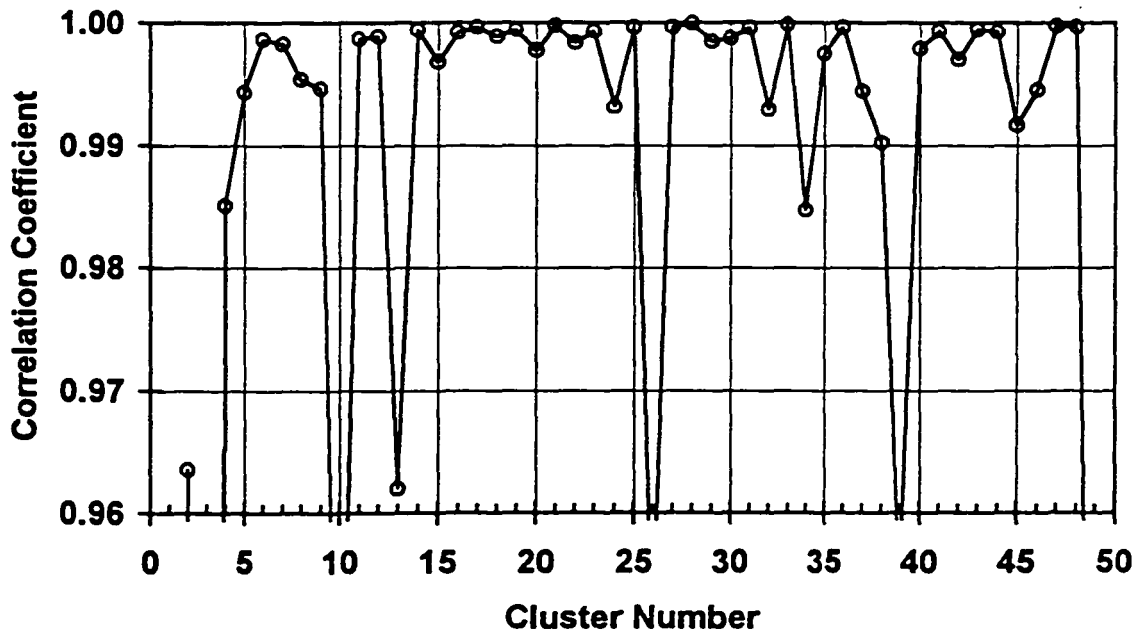
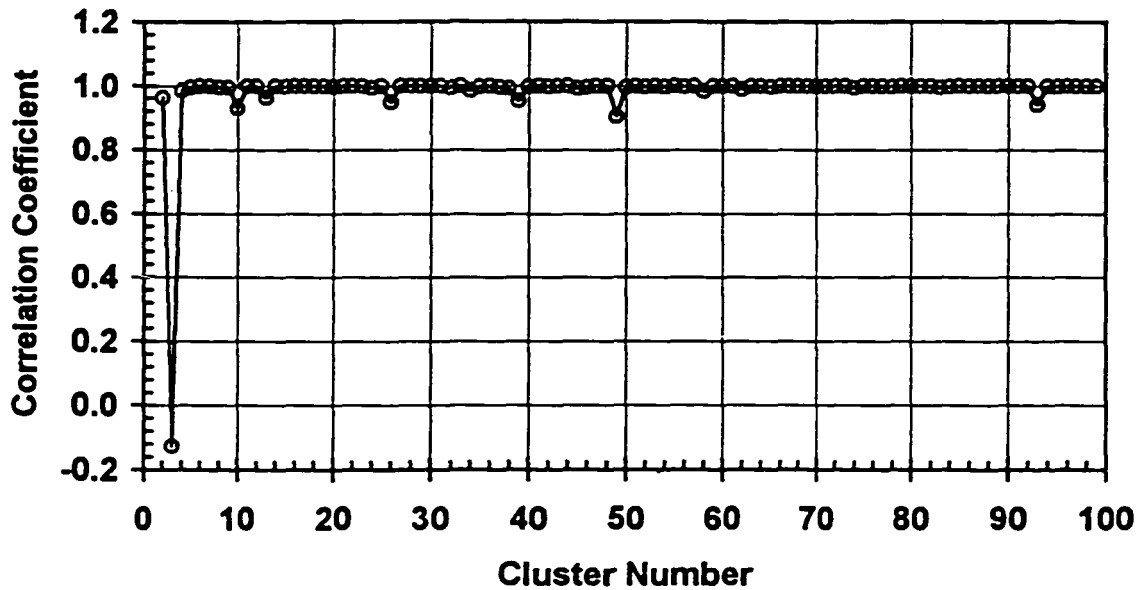
Single-Cosine-Zscore Method Using Q and B_{q2} N_c = No. of Clusters
 Note: Piezocone data at Bothkennar, Scotland (Nash et al., 1992).

Figure C.22a. Cluster Analysis of Type 2 Piezocone Data at Bothkennar, Scotland.



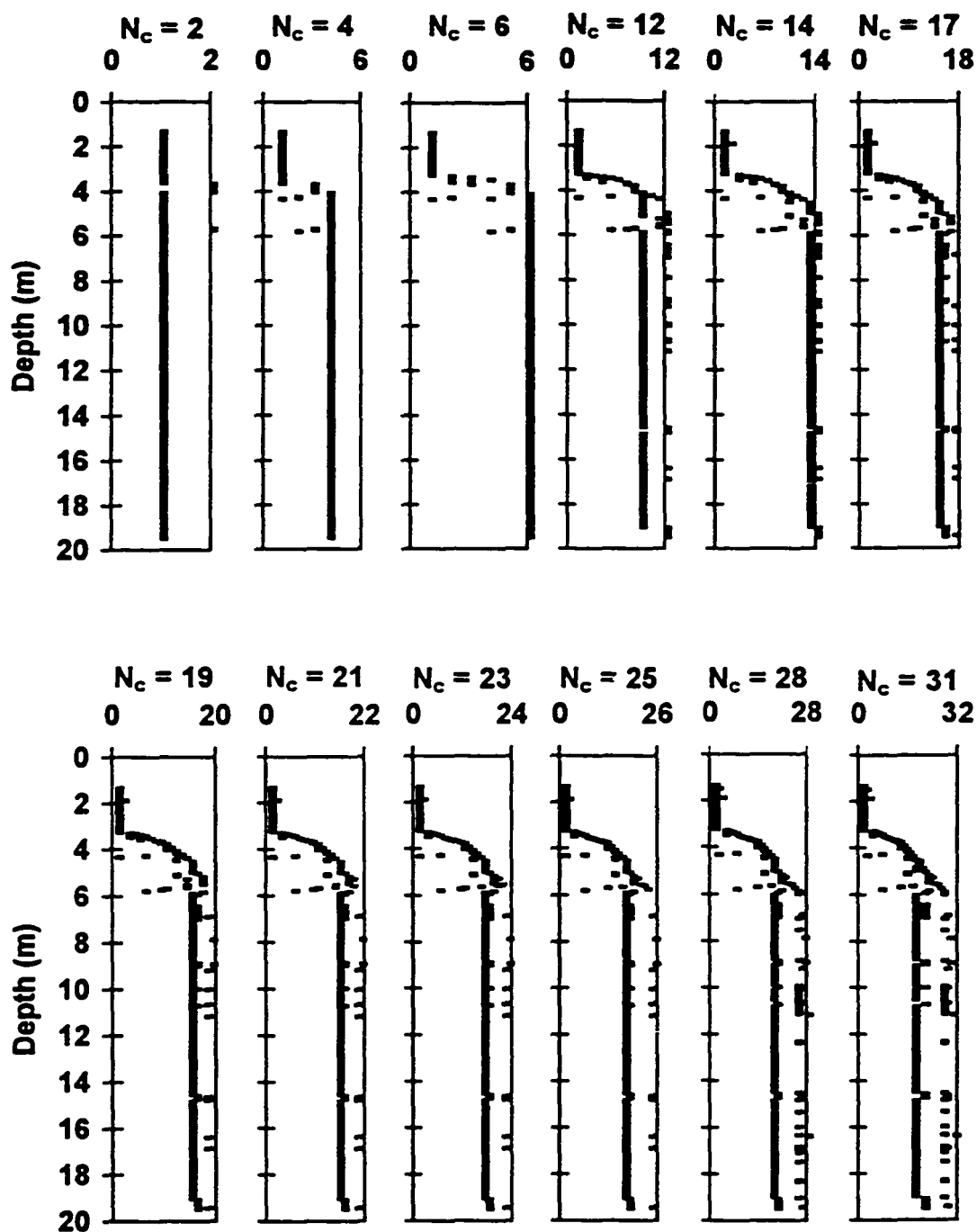
Single-Cosine-Zscore Method Using Q and B_{q2} N_c = No. of Clusters
 Note: Piezocone data at Bothkennar, Scotland (Nash et al., 1992).

Figure C.22b. Cluster Analysis of Type 2 Piezocone Data at Bothkennar, Scotland.



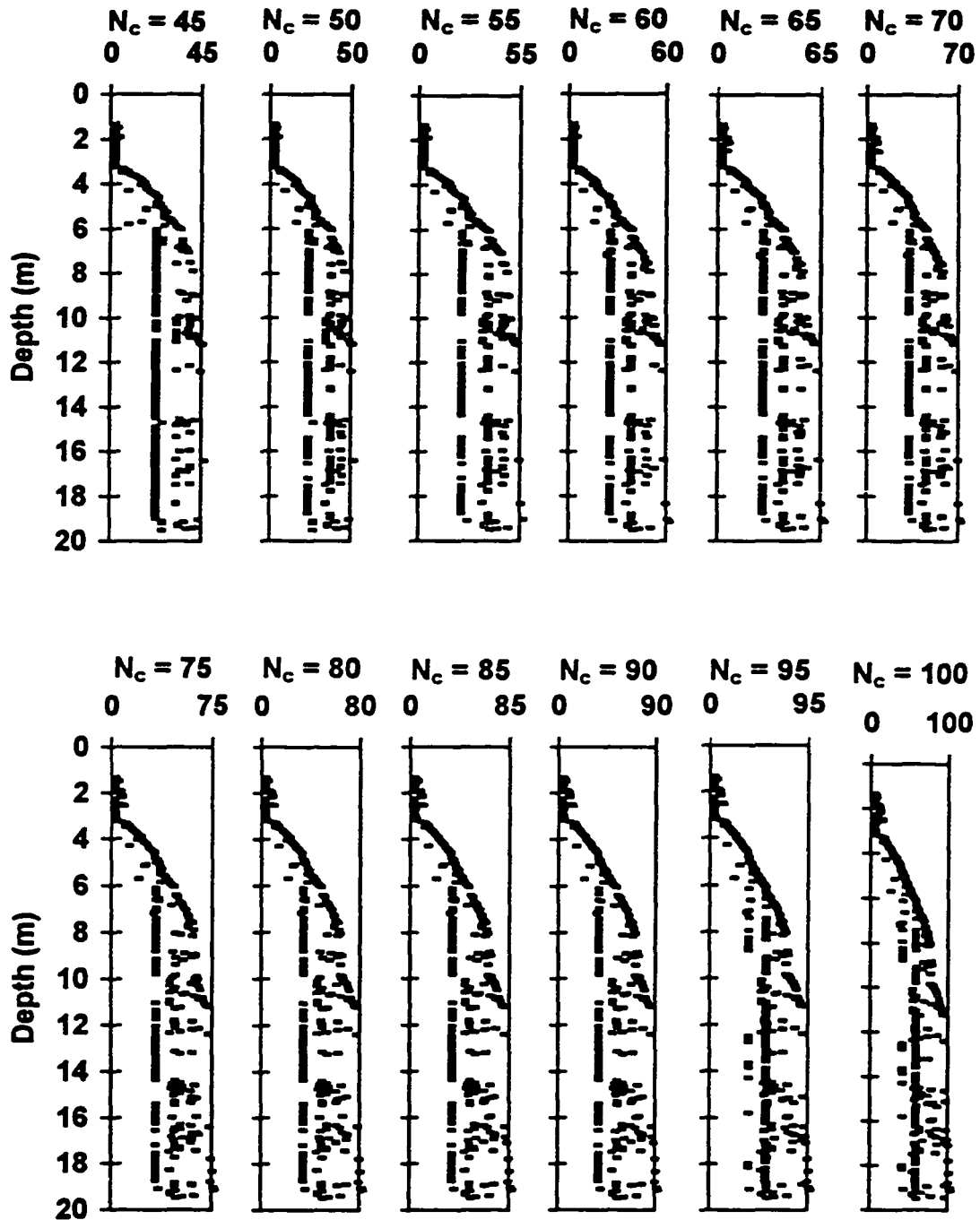
Single-Cosine-Zscore Method Using Q and B_{q1}
 Note: Piezocone data at Bothkennar, Scotland (Nash et al., 1992).

Figure C.23. Correlation Coefficient Between Consecutive Cluster Results at Bothkennar, Scotland Using Type 1 Piezocone Data.



Single-Cosine-Zscore Method Using Q and B_{q1} N_c = No. of Clusters
 Note: Piezocone data at Bothkennar, Scotland (Nash et al., 1992).

Figure C.24a. Cluster Analysis of Type 1 Piezocone Data at Bothkennar, Scotland.



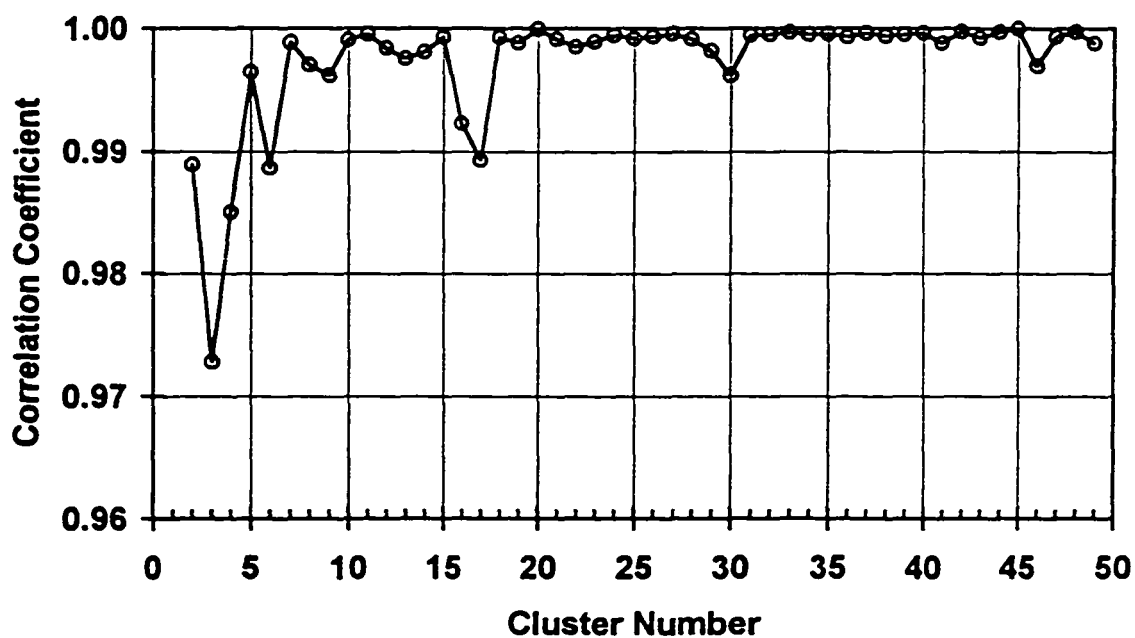
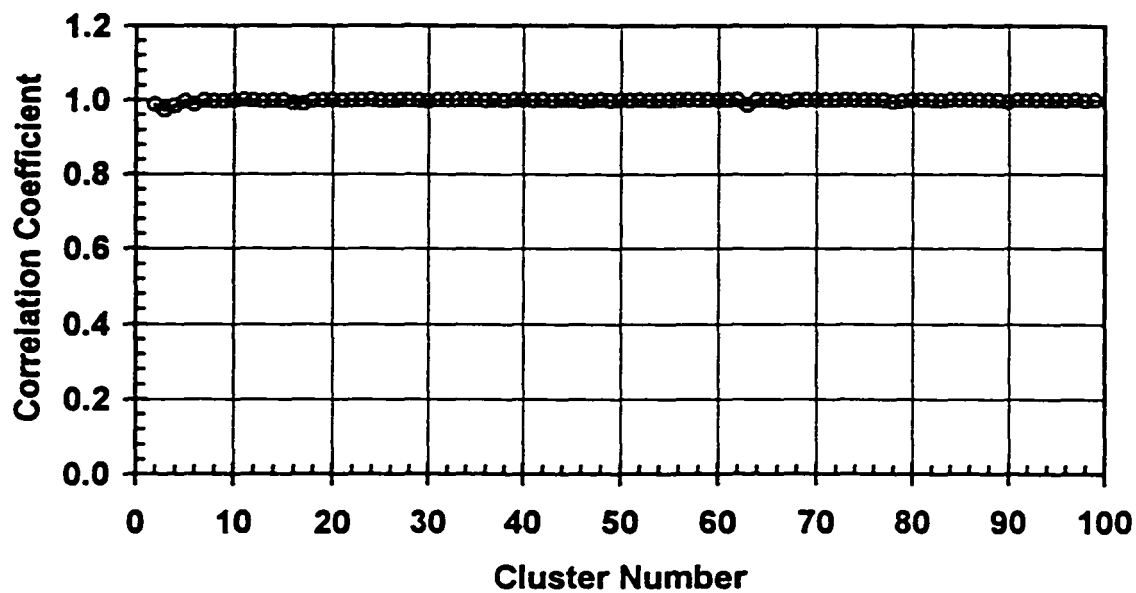
Single-Cosine-Zscore Method Using Q and B_{q1} N_c = No. of Clusters
 Note: Piezocone data at Bothkennar, Scotland (Nash et al., 1992).

Figure C.24b. Cluster Analysis of Type 1 Piezocone Data at Bothkennar, Scotland.

C.6.2. Brent Cross Test Site, UK

Piezocone data at the Brent Cross test site were analyzed using a single-cosine-zscore clustering technique, and presented previously in Fig. 5.20 (Powell et al., 1988). Statistical analysis is applied to Q and B_{q2} and the correlation coefficient between consecutive clusters is shown in Fig. C.25. Clustering is examined at the peaks of ρ_c up to cluster number 50 and the results are shown in Fig. C.26a between cluster number $N_c = 2$ and 37. At cluster number 2, the data are divided into two groups and their boundary is at a depth of 7.8 m. At cluster number 5, the upper cluster is separated into two primary layers at a depth of 2.4 m. A transition zone appear between depths of 7.6 m and 7.8 m. At cluster number 7, a secondary layer is detected between depths of 2.4 m and 3.3 m however no new primary layers are detected. More points continue to separate from the main groups up to cluster number 100 indicating dissimilarity within the same statistical layer. Figure C.26b shows the cluster results between $N_c = 45$ and 100 every 5 increments and new primary layers ($t > 1$ m) are discovered. Therefore, cluster number 5 is chosen to represent the subsurface stratigraphy. Note that the upper two main layers are denoted by two consecutive cluster numbers 1 and 2 which implies that they have related soil types and/or properties.

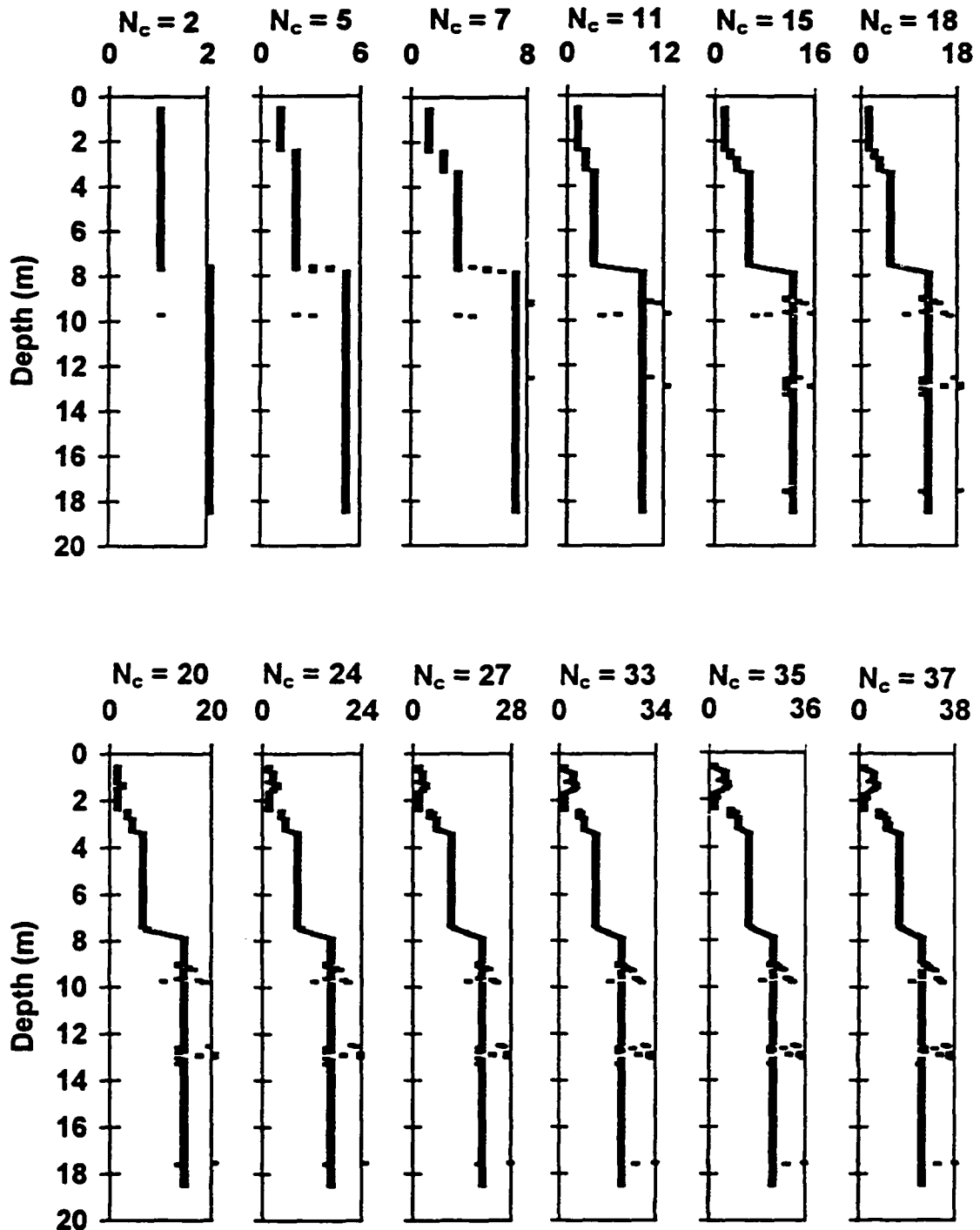
A similar clustering is performed using Q and B_{q1} and the correlation coefficient between consecutive clusters is shown in Fig. C.27 up to cluster number $N_c = 100$. Cluster results are examined at the peaks of ρ_c and shown in Fig. C.28a between $N_c = 2$ and 37. At cluster number 2, the data are separated into two main clusters at a depth of



Single-Cosine-Zscore Method Using Q and B_{q2}

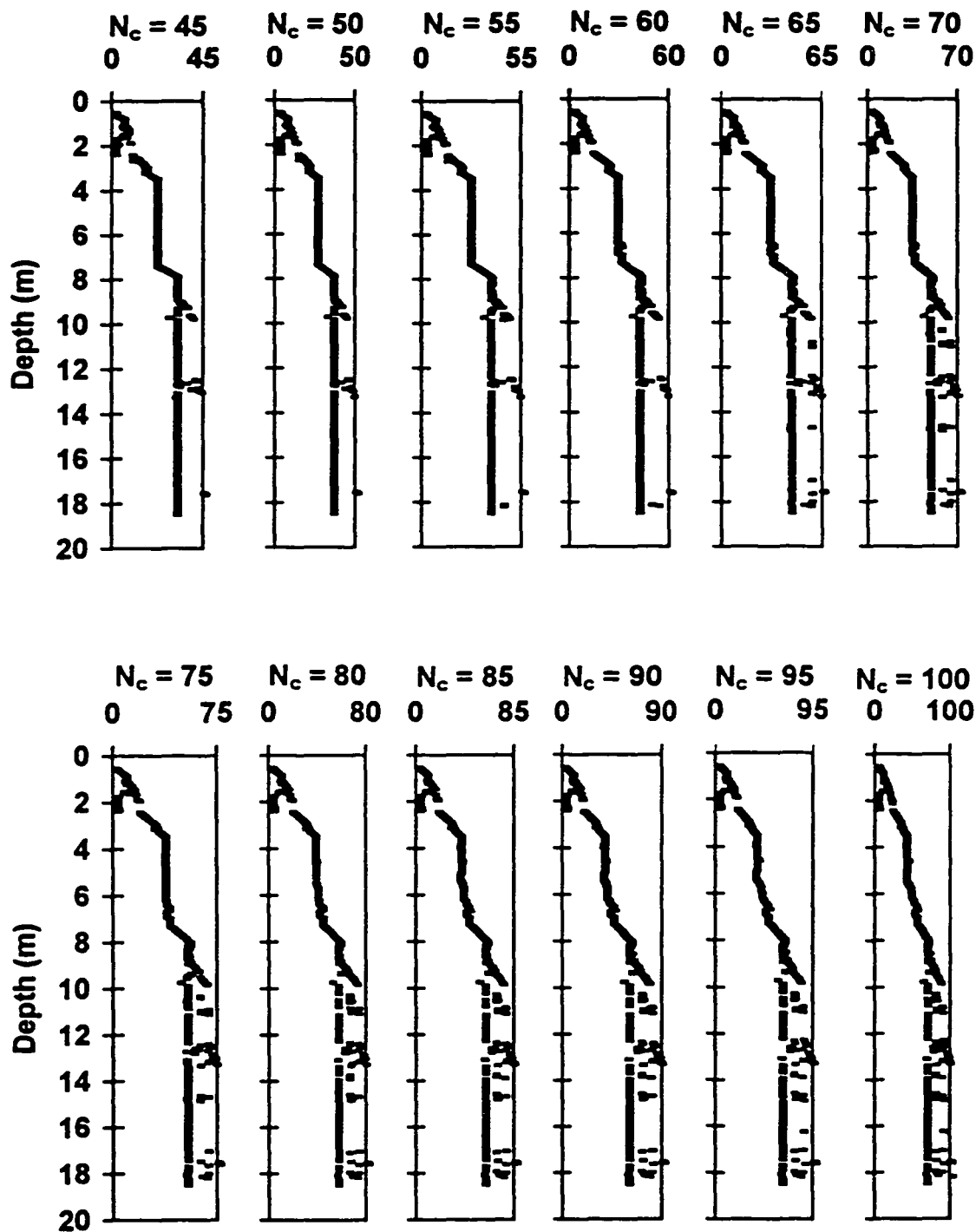
Note: Piezocone data at Brent Cross, UK (data from Powell et al., 1988).

Figure C.25. Correlation Coefficient Between Consecutive Cluster Results at Brent Cross, UK Using Type-2 Piezocone Data.



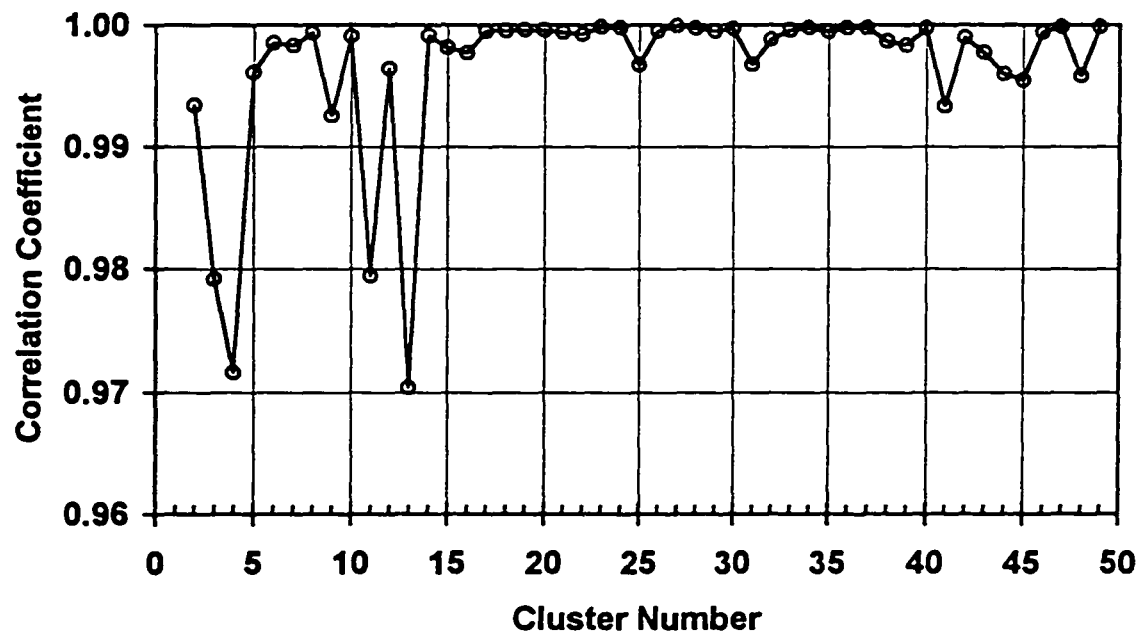
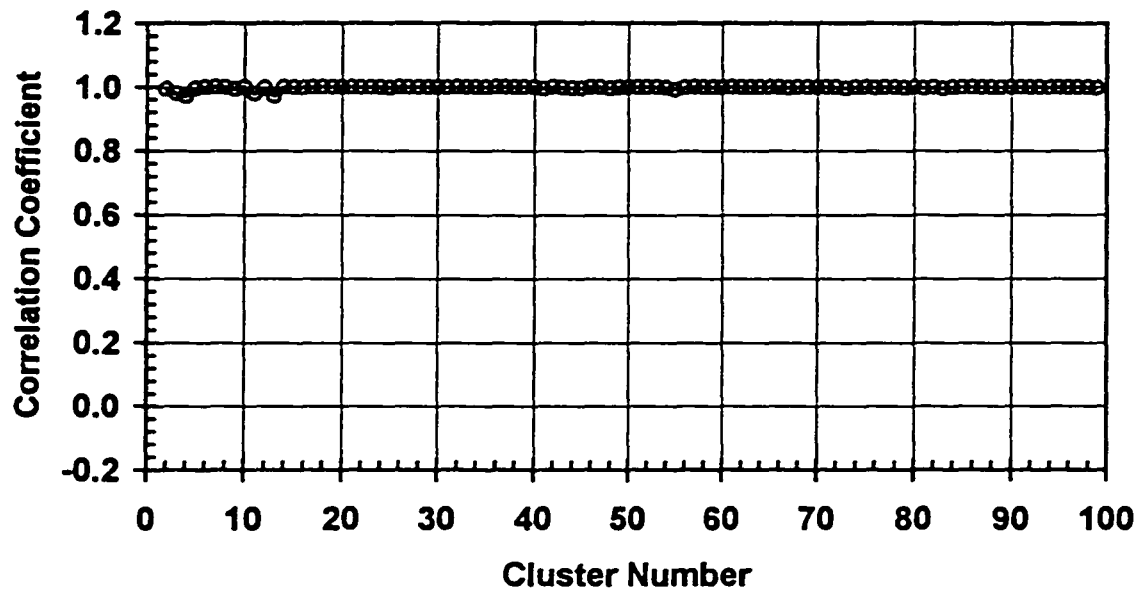
Single-Cosine-Zscore Method Using Q and B_{q2} N_c = No. of Clusters
 Note: Piezocone data at Brent Cross, UK (data from Powell et al., 1988).

Figure C.26a. Cluster Analysis of Type-2 Piezocone Data at Brent Cross, UK.



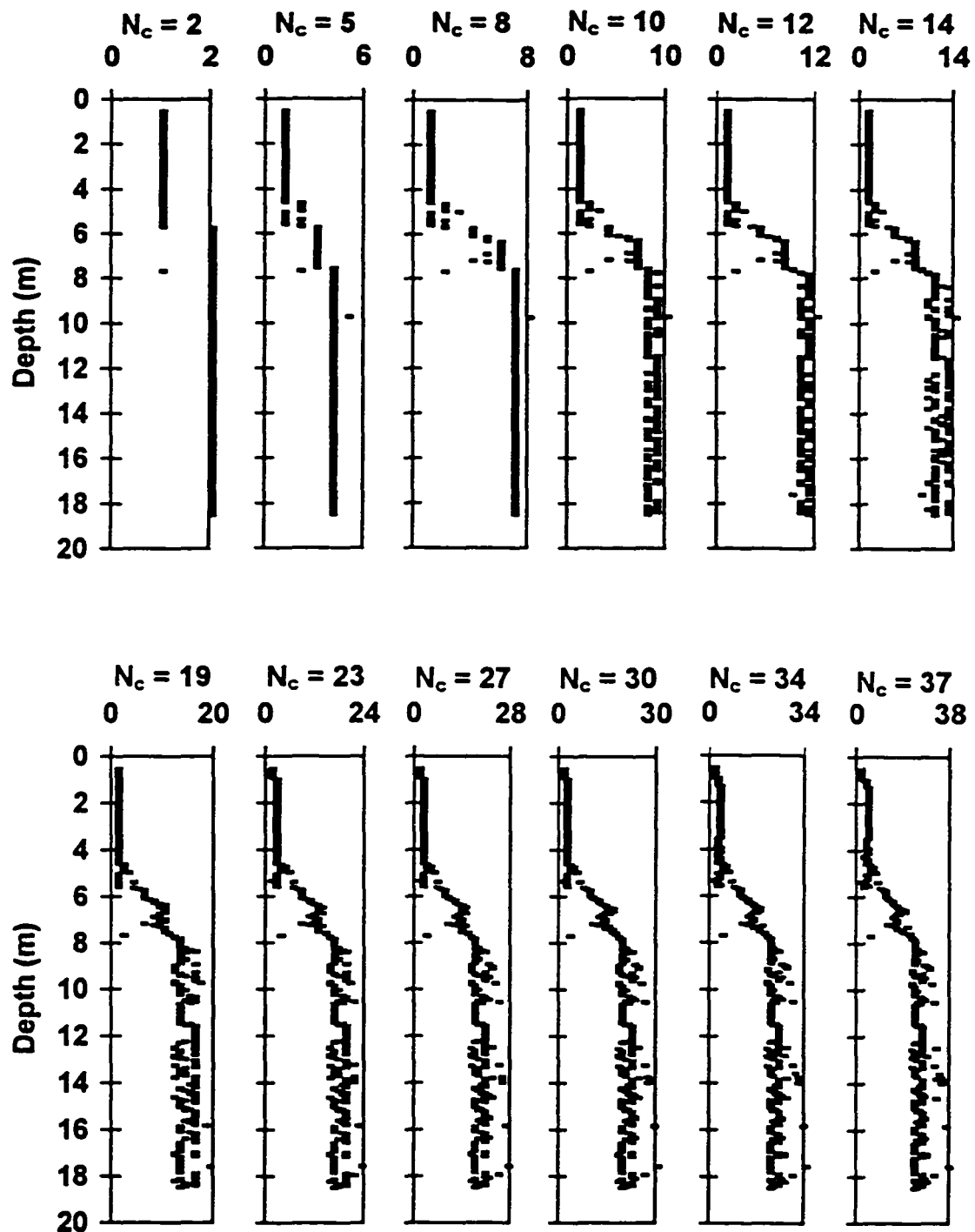
Single-Cosine-Zscore Method Using Q and B_{q2} N_c = No. of Clusters
 Note: Piezocone data at Brent Cross, UK (data from Powell et al., 1988).

Figure C.26b. Cluster Analysis of Type-2 Piezocone Data at Brent Cross, UK.



Single-Cosine-Zscore Method Using Q and B_{q1}
 Note: Piezocone data at Brent Cross, UK (Powell et al., 1988).

Figure C. 27. Correlation Coefficient Between Consecutive Cluster Results at Brent Cross, UK Using Type-1 Piezocone Data.



Single-Cosine-Zscore Method Using Q and B_{q1} N_c = No. of Clusters
 Note: Piezocone data at Brent Cross, UK (Powell et al., 1988).

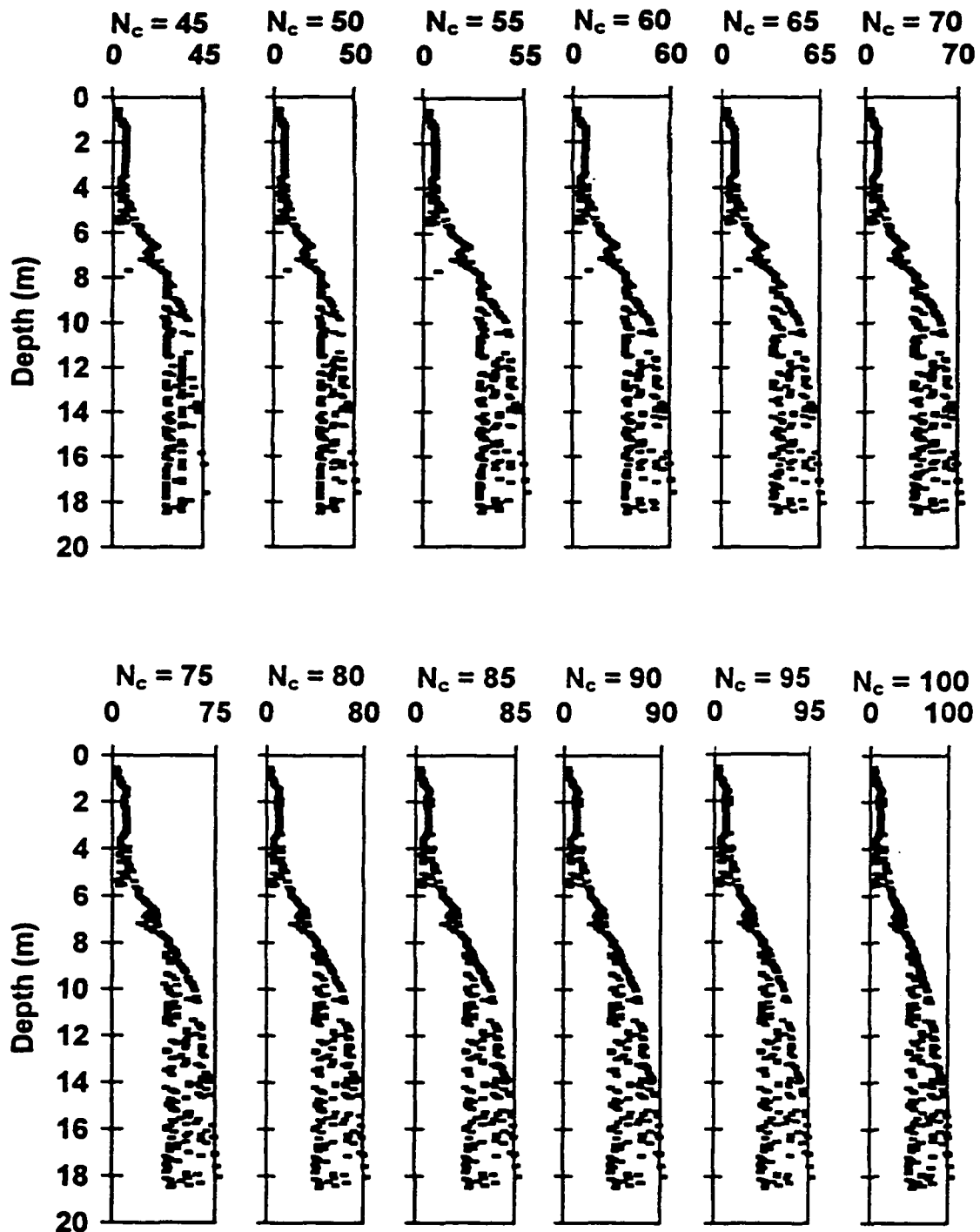
Figure C.28a. Cluster Analysis of Type-1 Piezocone Data at Brent Cross, UK.

5.7 m. At cluster number 5, the lower layer is divided into two primary layers with a boundary at a depth of 7.7 m. A transition layer appears between the depths of 4.6 m and 5.5 m where the cluster numbers alternates between 1 and 2. For $N_c > 5$, points continue to separate from the major statistical groups indicating dissimilarity within a layer. Up to $N_c = 100$ as shown in Fig. C.28b which includes cluster results between $N_c = 45$ and 100 every 5 increments, no new primary layers ($t > 1$ m) are detected. Therefore, cluster number 5 is chosen to represent the subsurface stratigraphy indicating more association between the lower two groups below a depth of 5.7 m.

Both cluster analyses using u_t and u_b measurements detected a soil boundary at almost 7.7 m, however not matching boundaries were defined above this level as follows: (1) at a depth of 2.4 m using Q and B_{q2} and (2) at a depth of 5.7 m using Q and B_{q1} . Back-up laboratory data such as liquid limits were more supportive to the analysis using u_t readings as discussed in Chapter 5.

C.7. Scale Effect

Piezocone data were collected at Surry, Virginia (Gordon and Mayne, 1987) using both 10-cm² and 15-cm² cones, and penetration porewater pressures were measured behind the tip (u_b). Both cone sizes are now common in geotechnical practice (Lunne et al., 1997) and permitted by ASTM guide D-5778. Piezocone records were previously presented in Fig. 5.24. The penetration porewater pressure alternated between negative and positive values in the upper 15 m during the penetration in the upper clay layer and the intermediate sand layer. This could be explained that the porous element could have been



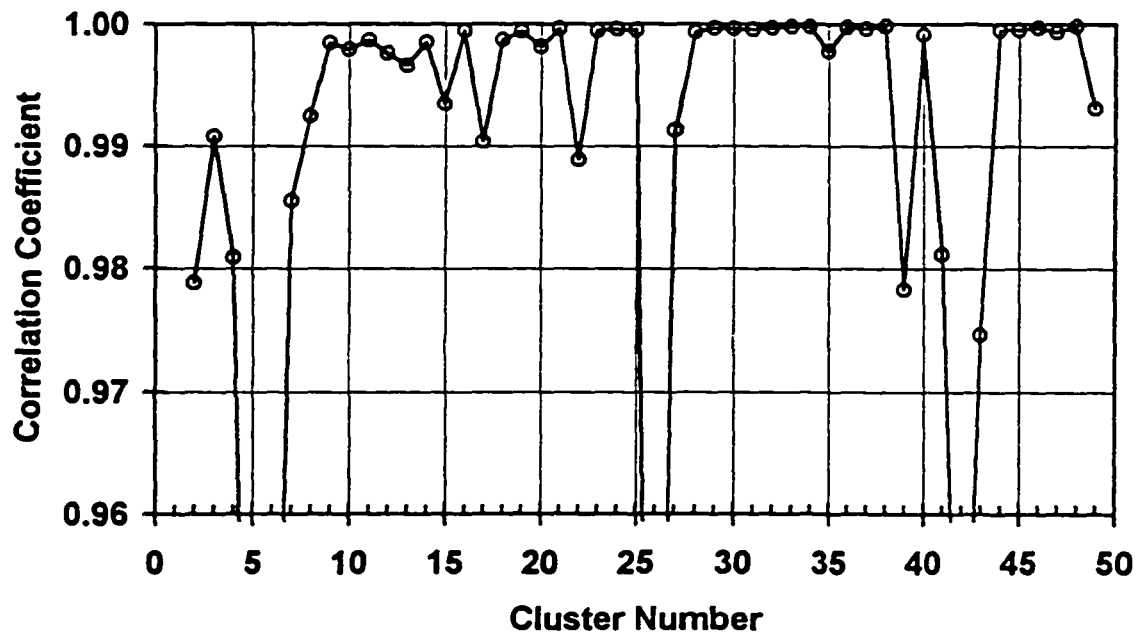
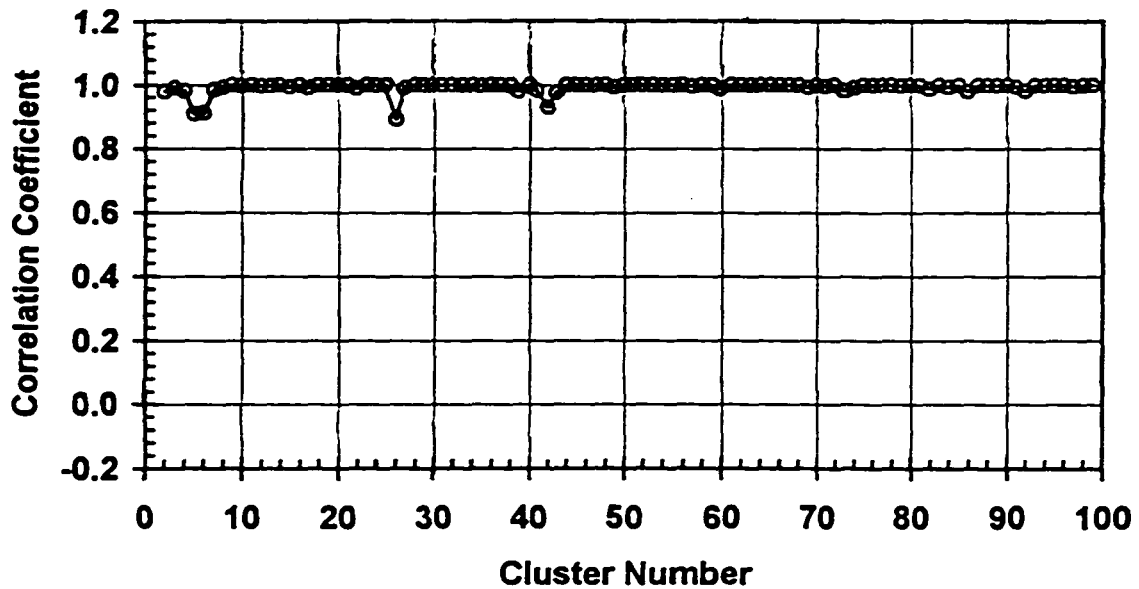
Single-Cosine-Zscore Method Using Q and B_{qt} N_c = No. of Clusters
 Note: Piezocone data at Brent Cross, UK (Powell et al., 1988).

Figure C.28b. Cluster Analysis of Type-1 Piezocone Data at Brent Cross, UK.

desaturated in the vadose zone above the groundwater table which was defined at a depth of 8.2 meter.

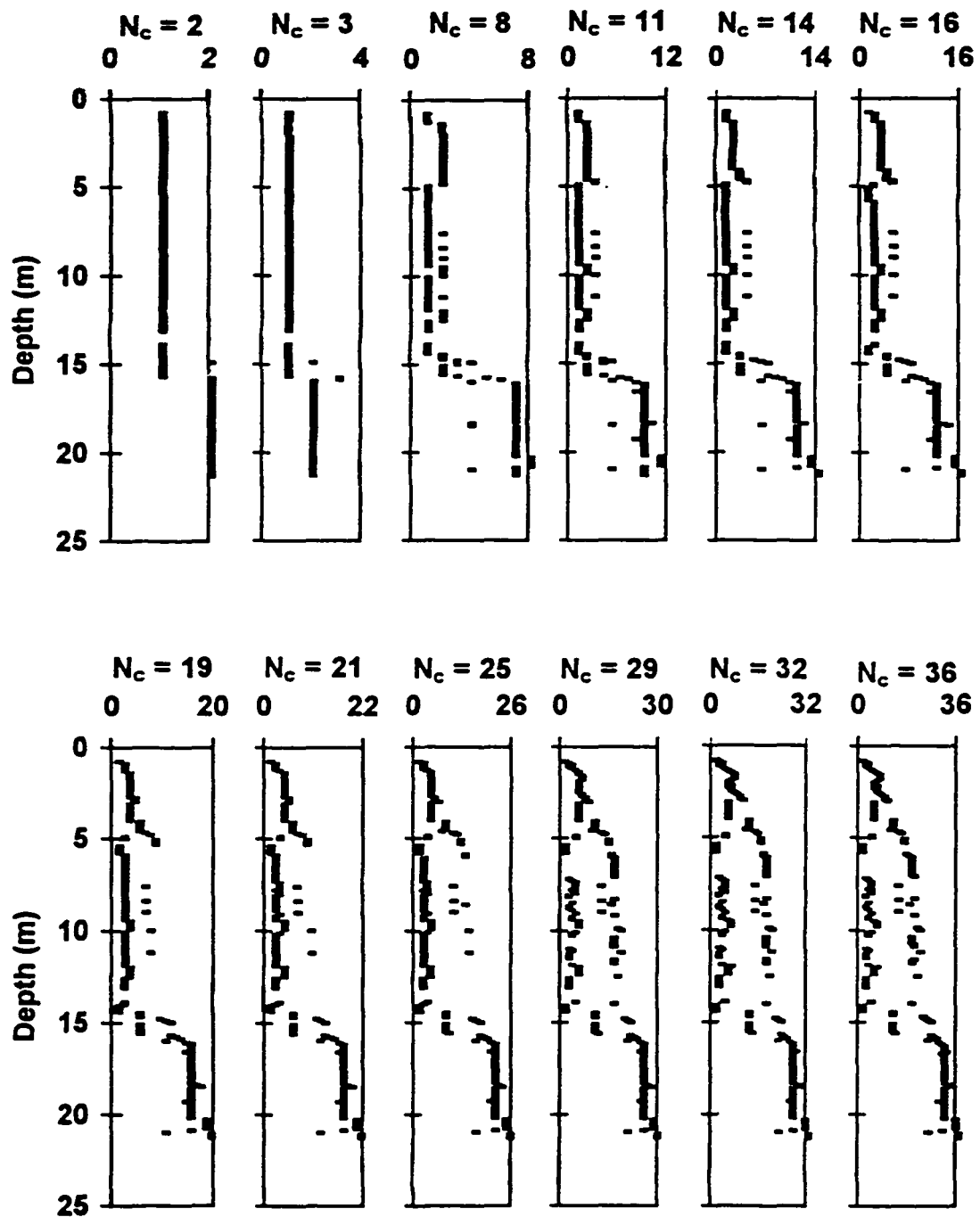
In this section, the effect of the cone size on the growth of cluster analysis is discussed. A single-cosine-zscore method is applied to normalized parameters Q and B_q of piezocone sounding CP5 collected using 10-cm² cone. The correlation coefficient between consecutive clusters is shown in Fig. C.29 up to cluster number $N_c = 100$. The cluster results are examined at the peaks of ρ_c up to cluster number 50 and shown in Fig. C.30a between cluster number $N_c = 2$ and 36. At cluster number 2, the data are divided into two groups with a boundary at a depth of 15.7 m. At cluster number 8, the upper group of data is separated into two primary layers with a boundary at a depth of 4.8 m. A transition layer appears between the depths of 13.8 m and 16 m. For higher clusters, data points continue to separate indicating dissimilarity with the three main groups. Up to $N_c = 100$, no new primary layers ($t \geq 1$ m) are discovered as shown in Fig. C.30b which includes cluster results between $N_c = 45$ and 100 every 5 increments. Therefore, cluster number 8 is chosen to represent the subsurface stratigraphy.

A similar cluster analysis is performed using piezocone data (SP8) collected using a 15-cm² cone. The correlation coefficient between consecutive clusters is shown in Fig. C.31 up to cluster number $N_c = 100$. The cluster results are examined at the peaks of ρ_c and shown in Fig. C.32a between cluster numbers 2 and 39. At cluster number 2, the data are divided into two primary groups with a boundary at a depth of 14.2 m. A transition layer appears between 14.2 m and 15.7 m. The upper group is separated into two primary



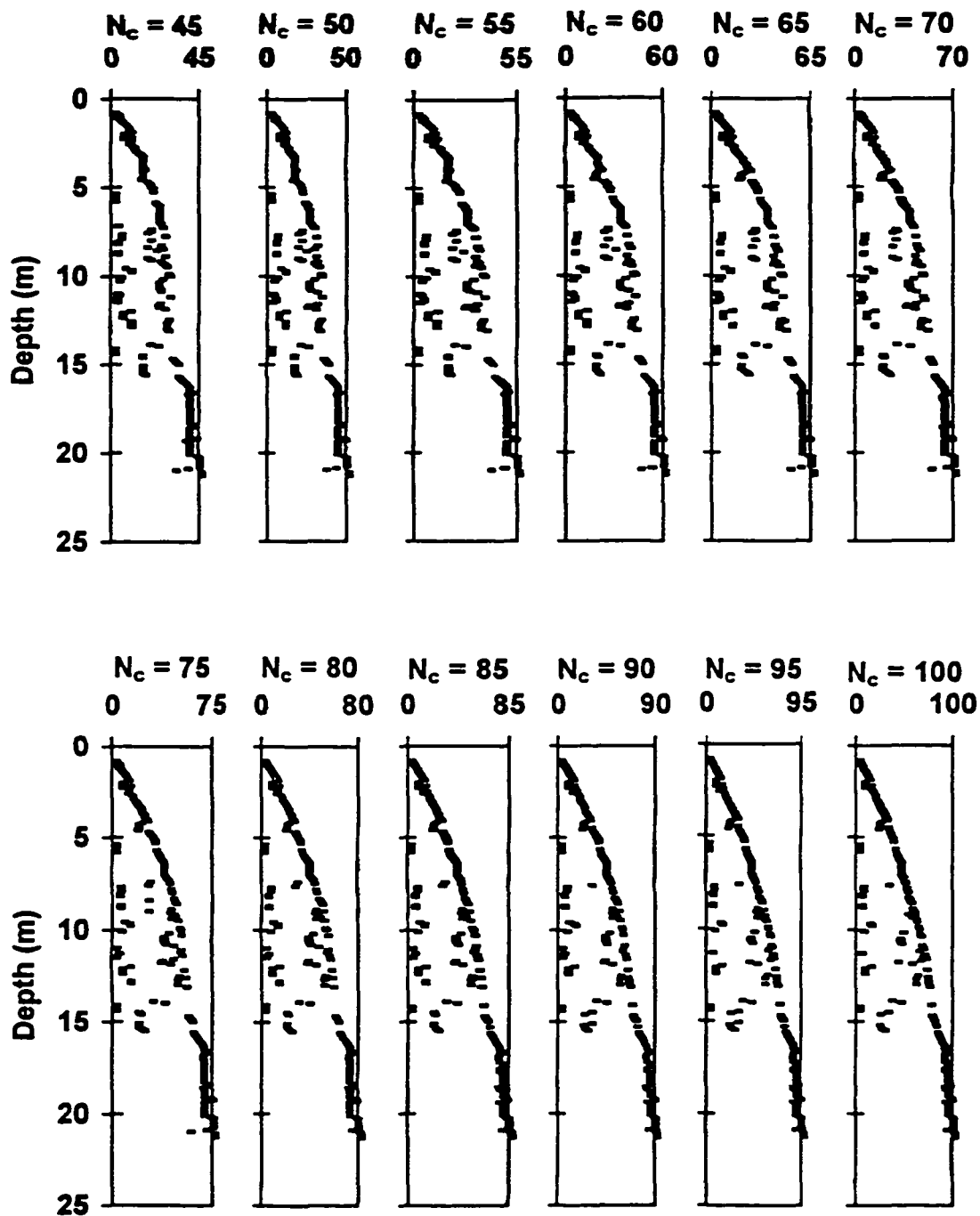
Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data (CP5) Using 10-cm² cone at Surry, Virginia.

Figure C.29. Correlation Coefficient Between Consecutive Cluster Results at Surry Using 10-cm² cone (Data from Gordon and Mayne, 1987).



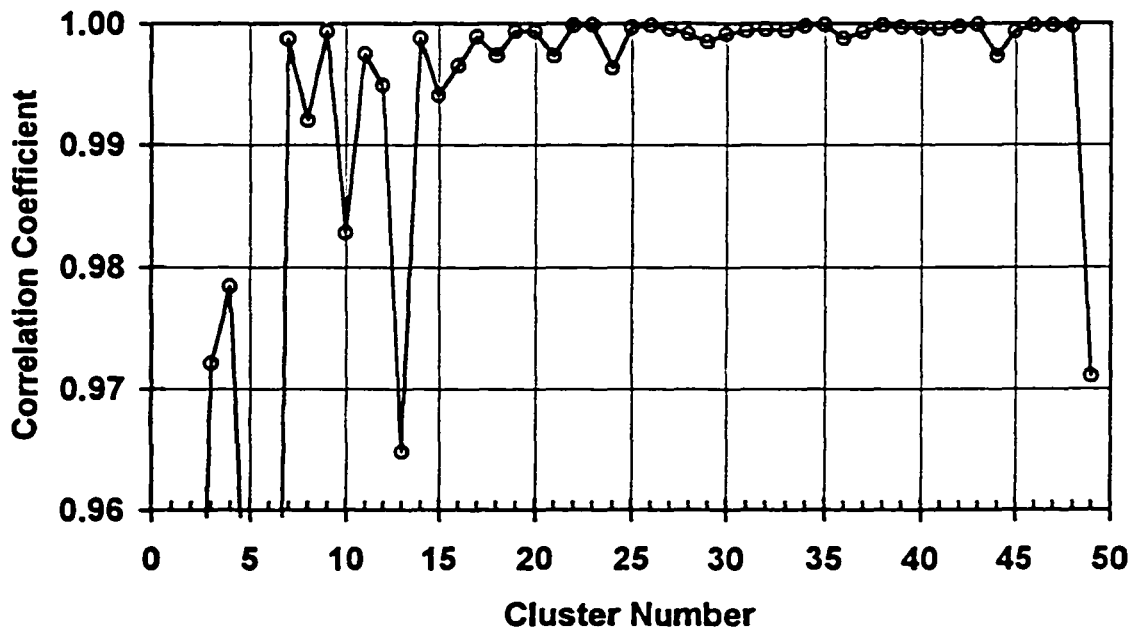
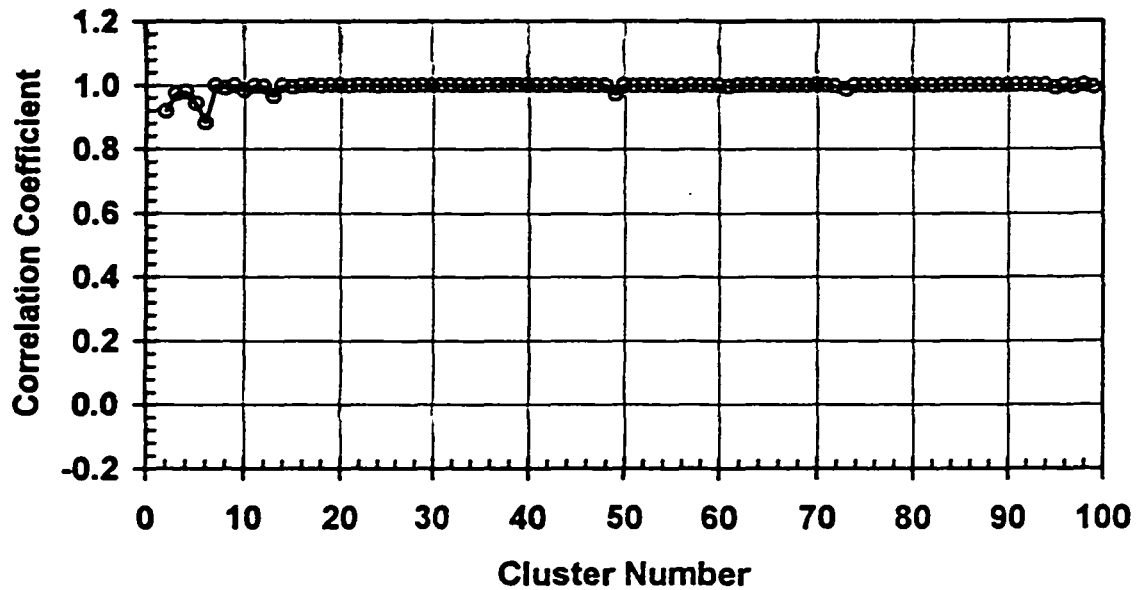
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (CP5) Using 10-cm² cone at Surry, Virginia.

Figure C.30a. Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 10-cm² Cone.



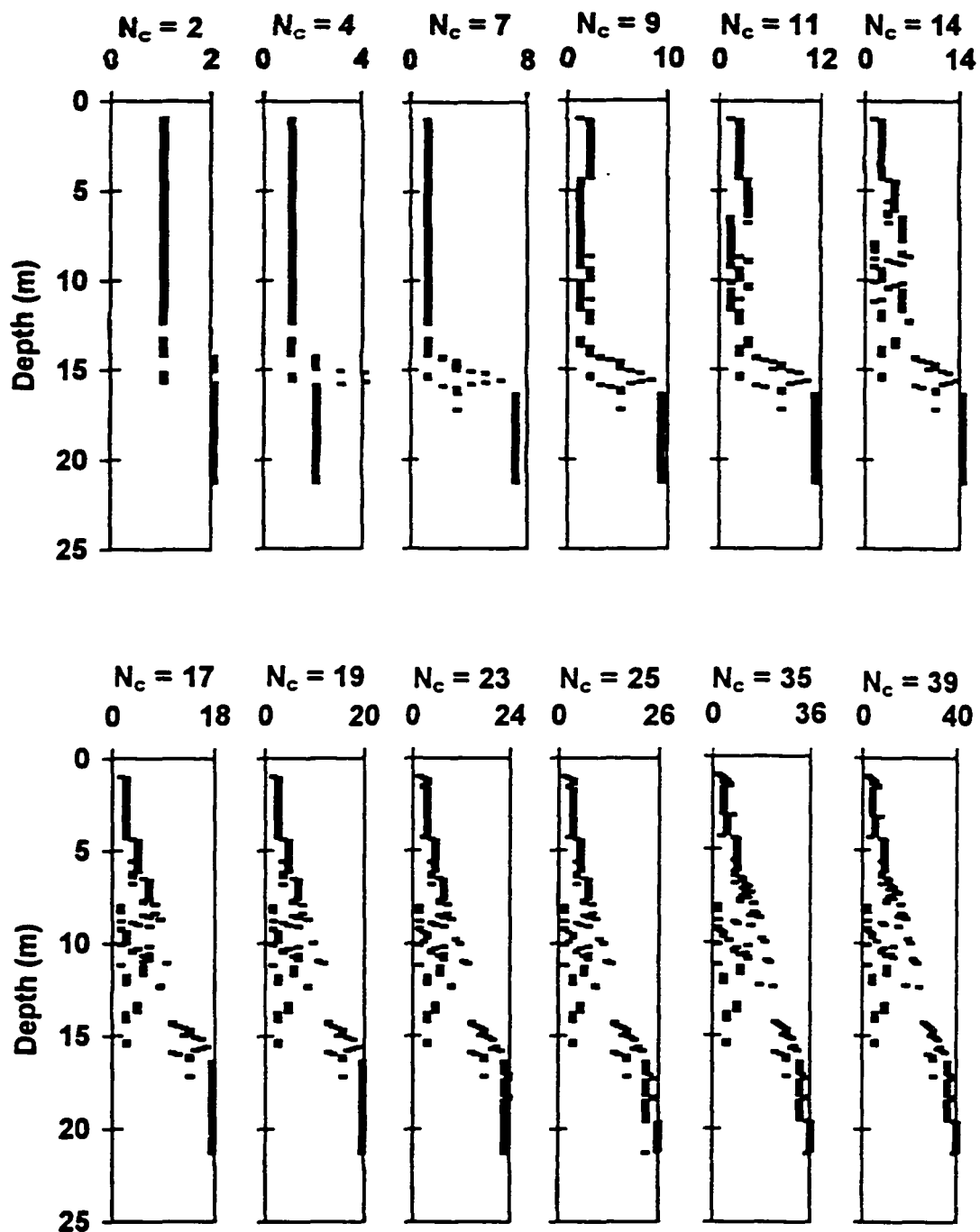
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (CP5) Using 10-cm² cone at Surry, Virginia.

Figure C.30b. Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 10-cm² Cone.



Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data (SP8) Using 15-cm² cone at Surry, Virginia.

Figure C.31. Correlation Coefficient Between Consecutive Cluster Results at Surry Using 15-cm² cone (Data from Gordon and Mayne, 1987).



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (SP8) Using 15-cm² cone at Surry, Virginia.

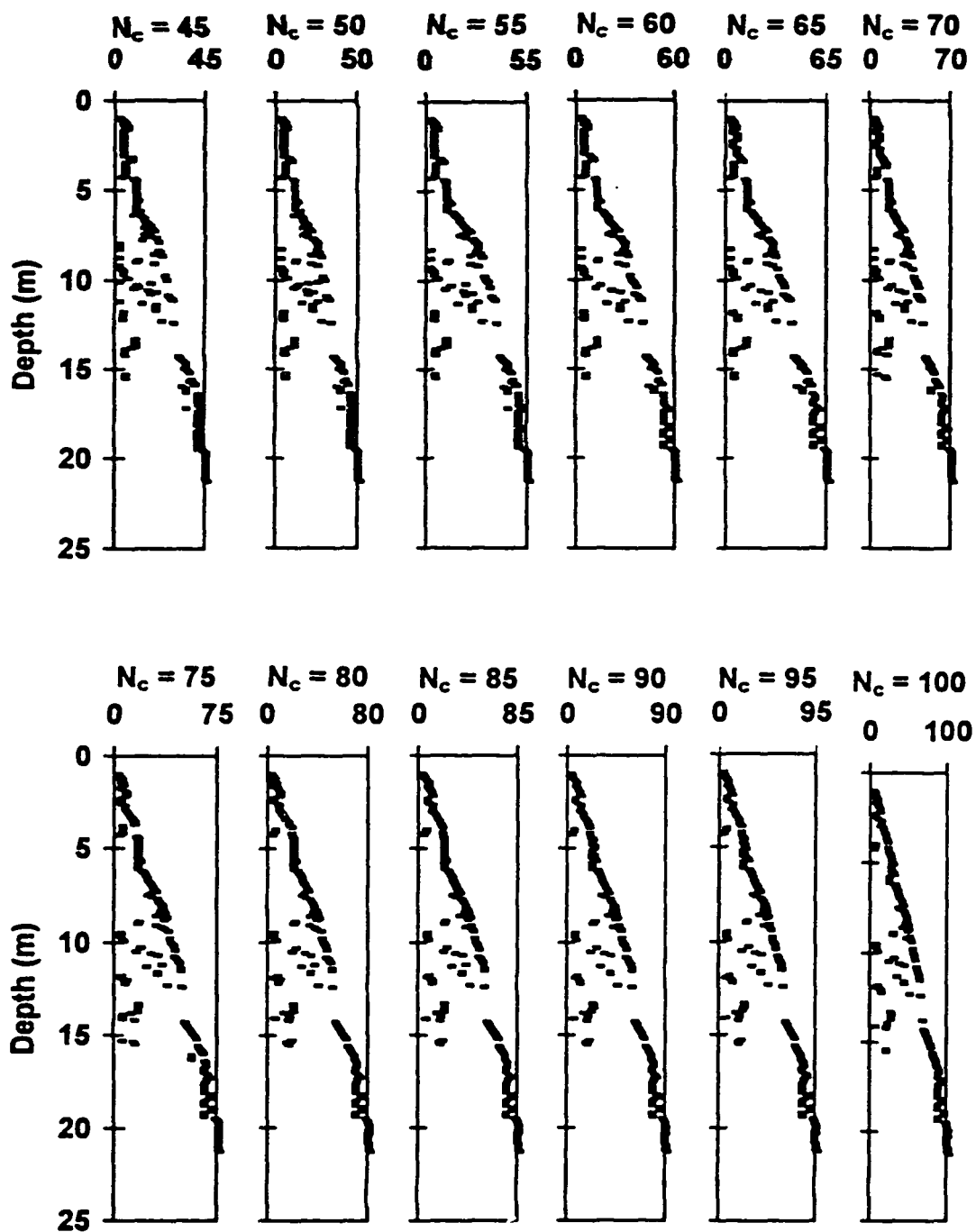
Figure C.32a. Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 15-cm² Cone.

layers with a boundary at a depth of 4.3 m. The transition layer is extended between the depths of 13.2 m and 16.3 m. For $N_c > 9$, data points continue to separate indicating dissimilarity with the three main groups. Up to $N_c = 100$, no new primary layers ($t > 1$ m) are discovered as shown in Fig. C.32b which includes cluster results between $N_c = 45$ and 100 every 5 increments. Therefore, cluster number 9 is chosen to represent the subsurface stratification.

The soil profile obtained using both cone soundings was almost similar which implies an apparent negligible effect of the cone size on the cluster analysis. In both analyses, clustering suggests that the first and third layers are more homogeneous than the middle layer which contains more points separated from the main cluster and indicate soil lenses or outliers. The upper two layers are related to each other because they are denoted two consecutive cluster numbers 1 and 2. However, this is not supported by the fact that the upper layer is clay and the middle layer is sand. This misleading indication is due to the effect of the negative or zero porewater pressure down to a depth of 15 meter.

C.8. Spatial Cluster Analysis

Clustering is performed on three individual piezocone data of soundings numbers PCPT1, PCPT2 and PCPT15 at Amherst, Massachusetts using single-cosine-zscore method. In this section, data grouping of PCPT2 and PCPT15 soundings is discussed, while, that of PCPT1 sounding was given in detail in Chapter 5. Vertical profiles of piezocone data were previously shown in Fig. 5.29.

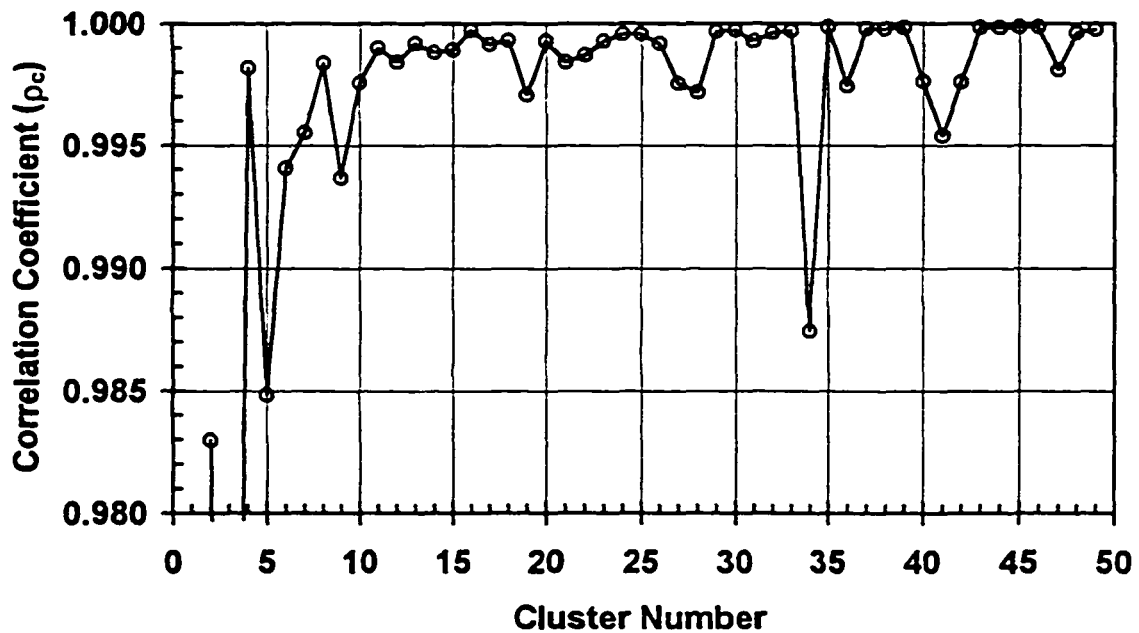
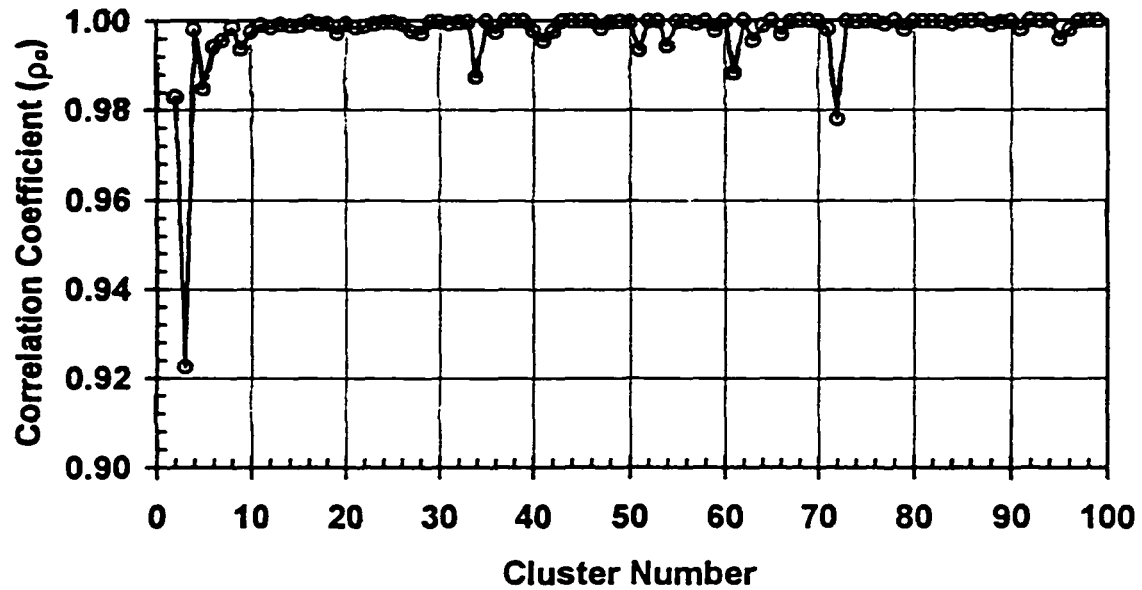


Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (SP8) Using 15-cm² cone at Surry, Virginia.

Figure C.32b. Cluster Analysis of Piezocone Data from Gordon and Mayne (1987) at Surry, Virginia Using 15-cm² Cone.

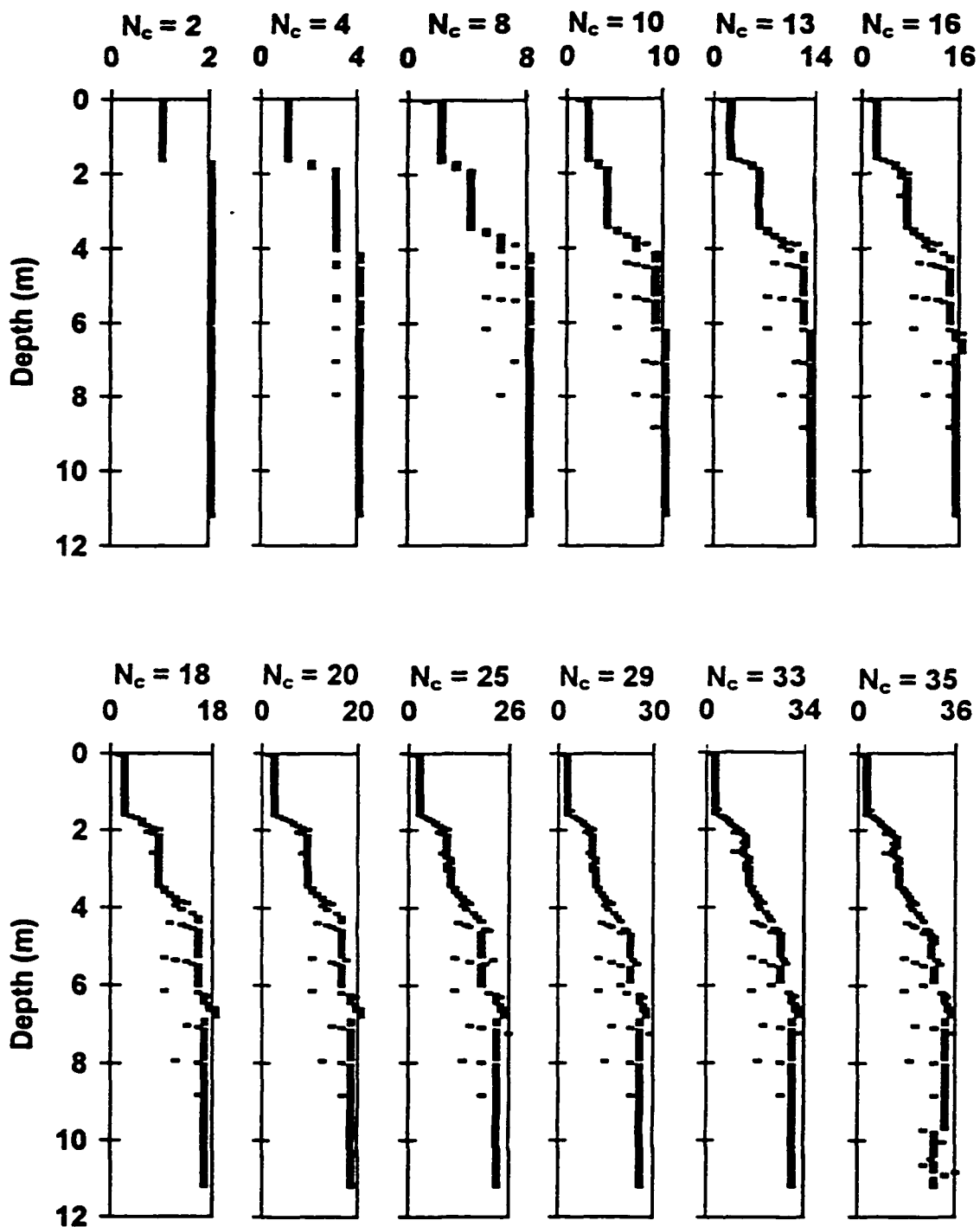
The correlation coefficient between consecutive clusters using normalized parameters Q and B_q of sounding PCPT2 is shown in Fig. C.33 up to $N_c = 100$. The cluster results are reviewed at the peaks of ρ_c and shown in Fig. C.34a between cluster number $N_c = 2$ and 36. At cluster number 2, data are divided into two groups with a boundary at 1.6 m. At cluster number 4, the lower group is separated into two primary layers at a depth of 4 m. A transition zone appears between the depths of 1.6 m and 1.8 m. Also, some points denoted by cluster number 3 below a depth of 4 m indicate the locations pauses in penetration for adding cone-rods. At cluster number 10, another primary layer is detected between the depths of 4 m and 6 m. More points continue to separate from the main clusters representing dissimilarity within a statistical layer. However, up to cluster number 100, no new layers are detected as shown in Fig. C.34b which includes the cluster results between $N_c = 45$ and 100 every 5 increments. Therefore, cluster number 10 is chosen to represent the subsurface stratigraphy.

A similar analysis is performed using normalized parameters Q and B_q of PCPT15 and the correlation coefficient between consecutive clusters is shown in Fig. C.35. Clustering is examined at the peaks of ρ_c and cluster results are shown in Fig. C.36a between cluster number $N_c = 2$ and 39. At cluster number 2, the data are divided into two groups with a boundary at a depth of 5.4 m. A transition zone appears between the depths of 5.4 m and 6.2 m. A soil lense or an outlier appears at a depth of 8 m. At cluster number 4, a similar profile is obtained except few points are separated from the two main layers. At cluster number 9, the upper layer is divided into three primary layers and their boundaries are defined at depths of 2.5 m and 3.9 m. Up to $N_c = 100$, no new main layers



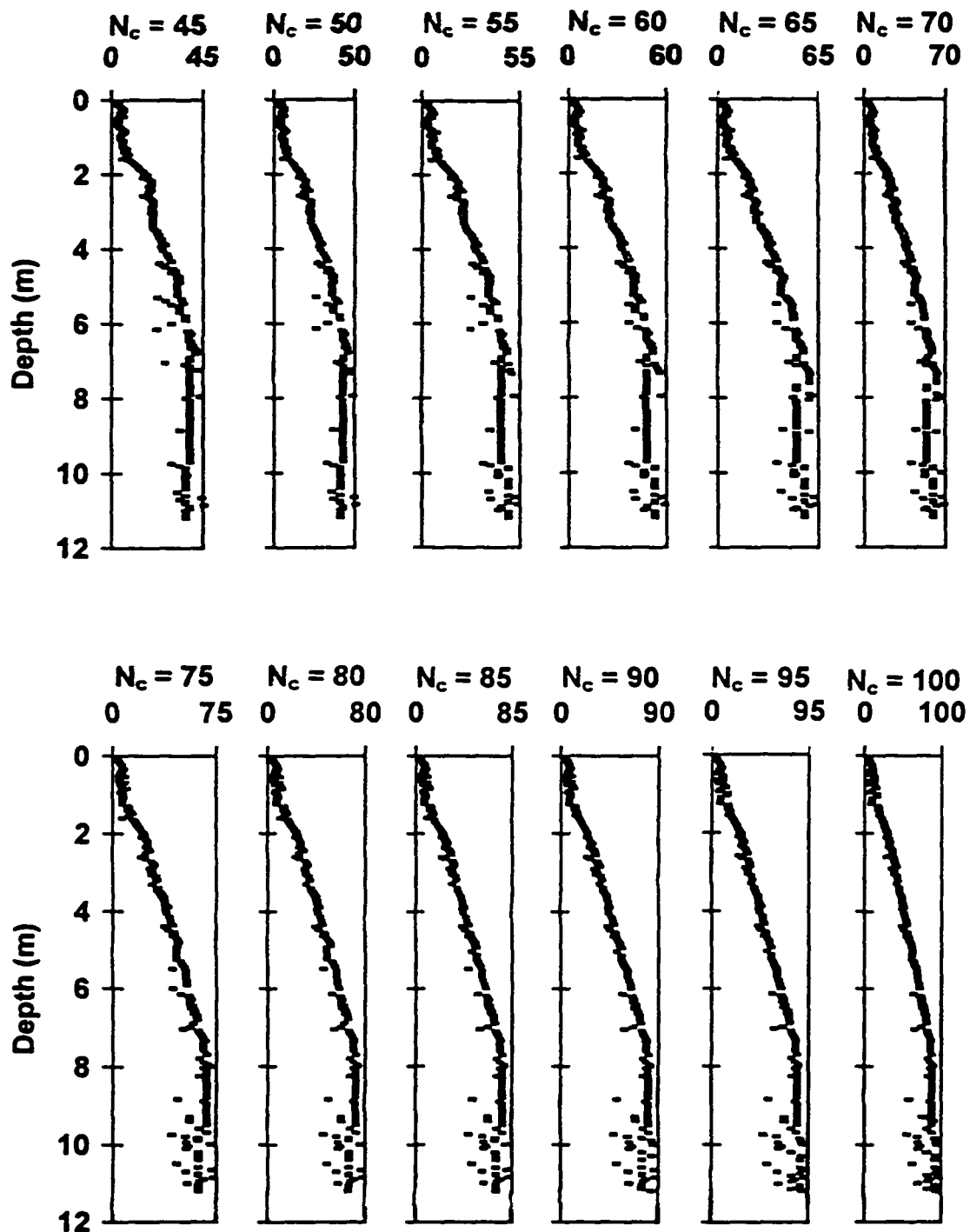
Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data (PCPT2) at Amherst, Massachusetts (data from this study).

Figure C.33. Correlation Coefficient Between Consecutive Cluster Results of Piezocone Data (PCPT2) at Amherst, Massachusetts.



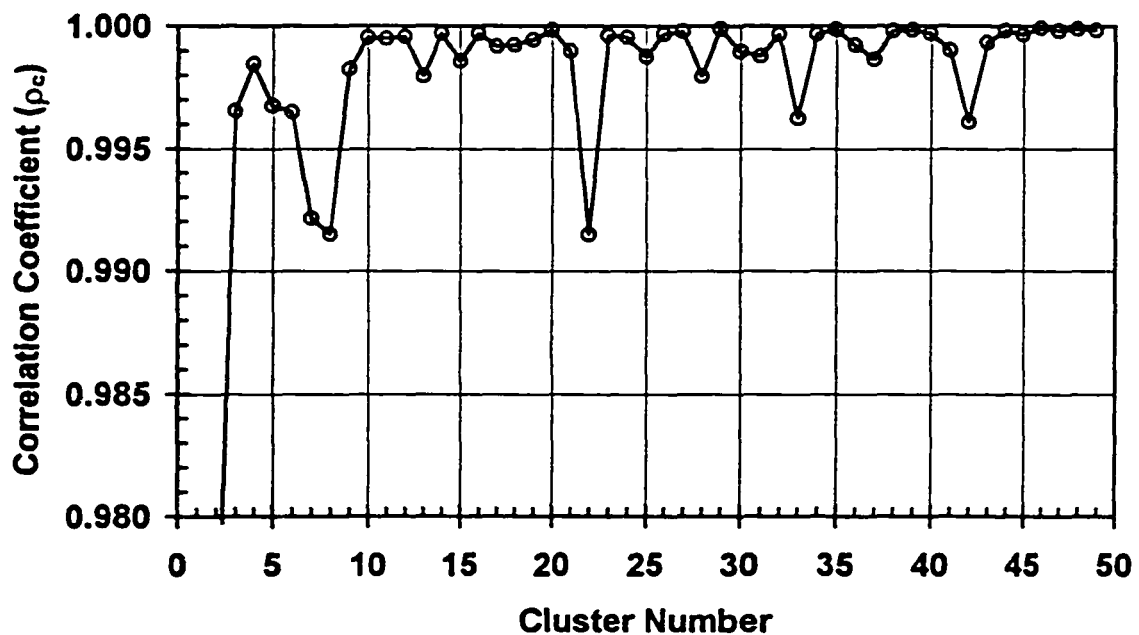
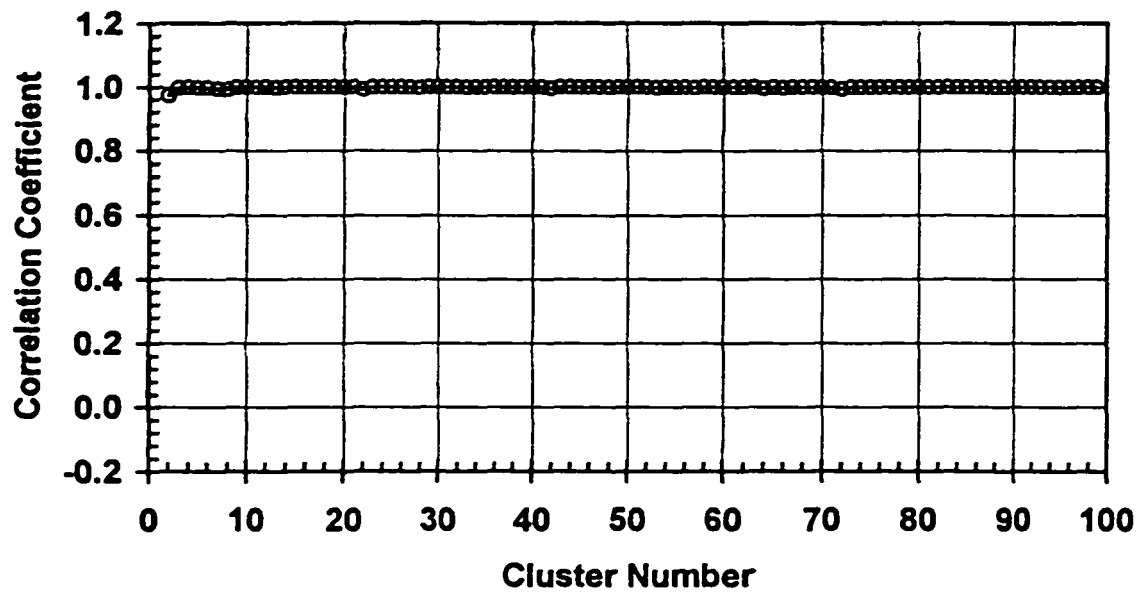
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (PCPT2) at Amherst, Massachusetts (data from this study).

Figure C.34a. Clustering of Piezocone Sounding (PCPT2) at Amherst.



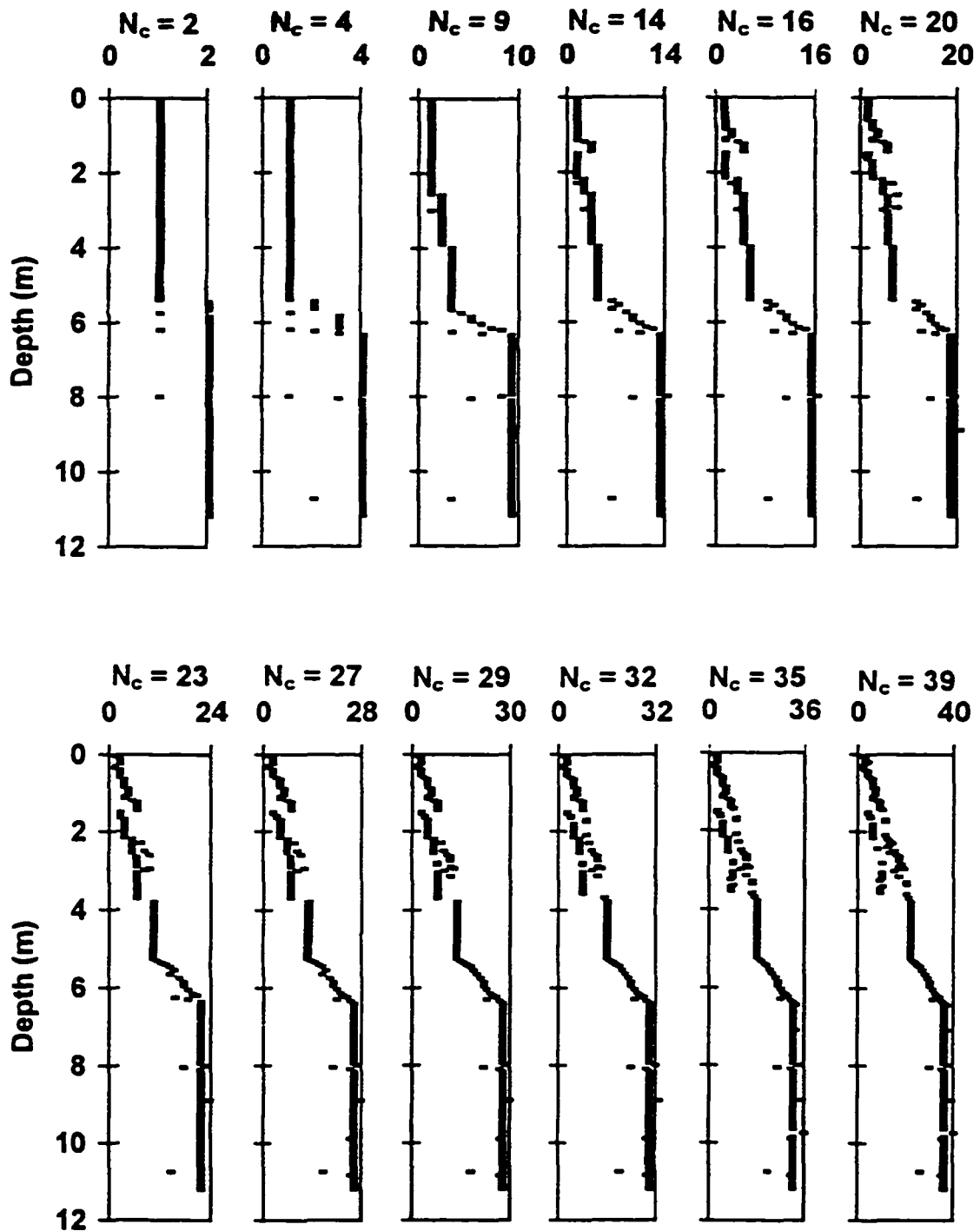
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (PCPT2) at Amherst, Massachusetts (data from this study).

Figure C.34b. Clustering of Piezocone Sounding (PCPT2) at Amherst.



Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data (PCPT15) at Amherst, Massachusetts (this study).

Figure C. 35. Correlation Coefficient Between Consecutive Cluster Results of Piezocone Data (PCPT15) at Amherst, Massachusetts.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (PCPT15) at Amherst, Massachusetts (data from this study).

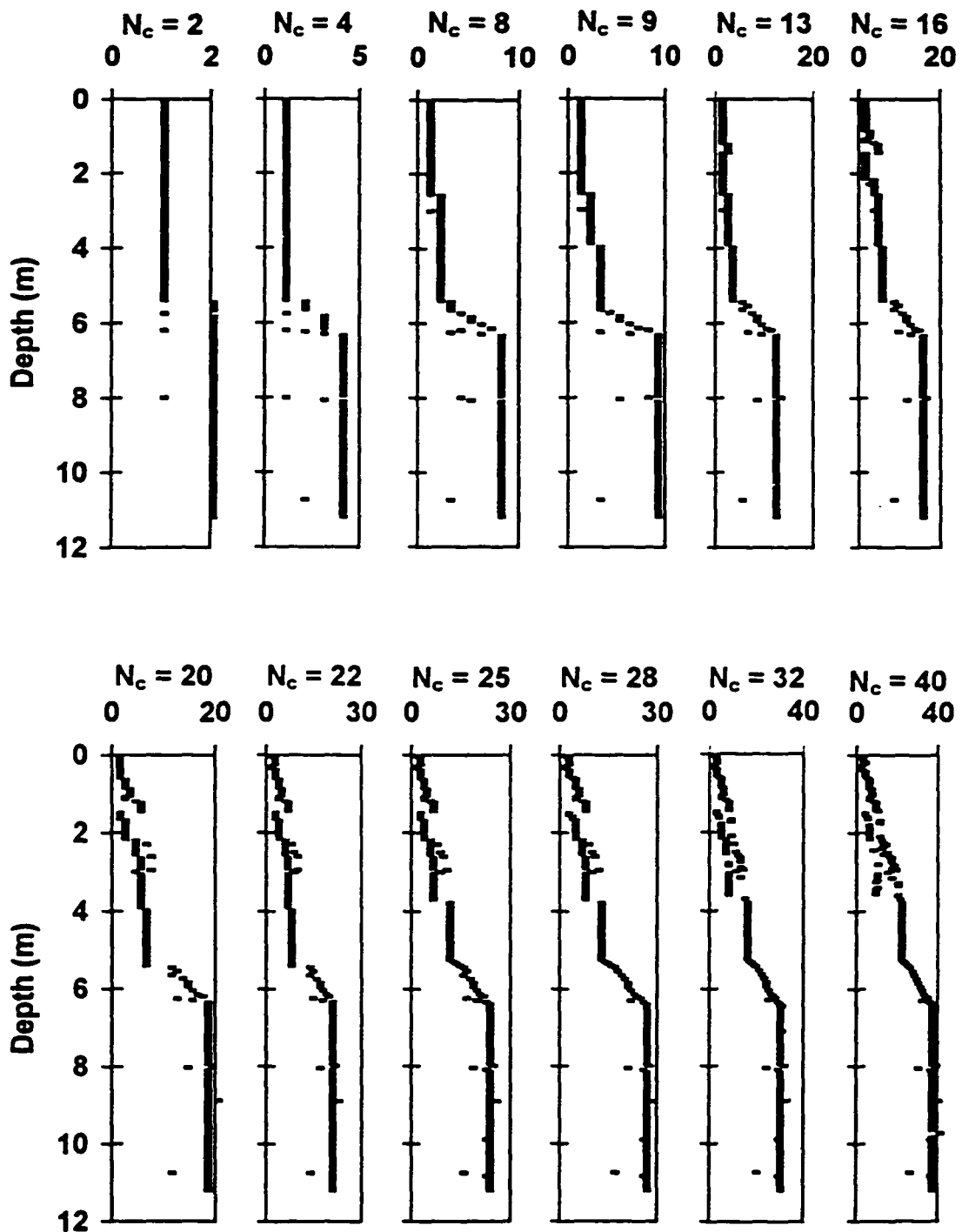
Figure C.36a. Clustering of Piezocone Sounding (PCPT15) at Amherst.

($t > 1$ m) are detected as seen in Fig. C.36b which shows the cluster results between $N_c = 45$ and 100 every 5 increments. However, data points continue to separate from the four data groups indicating dissimilarity within the same statistical layer. Therefore, cluster number 9 is chosen to represent the subsurface stratification.

Using piezocone data of both soundings, there is a similarity between the number of layers and the demarcation of different boundaries. However, the soil layer defined between the depths of 4 m and 6 m is suggested to be related to the underlying soil using PCPT2 and associated with the overlying soil in case of PCPT 15. This might reflect a some change in the soil type and/or properties from location to another due to inherent soil variability.

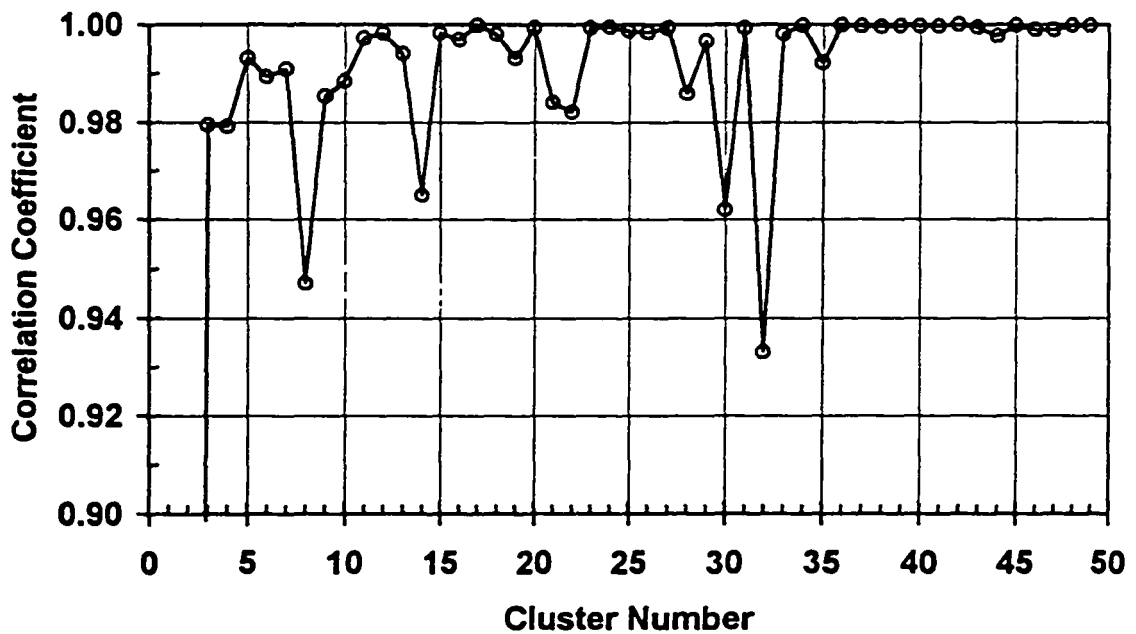
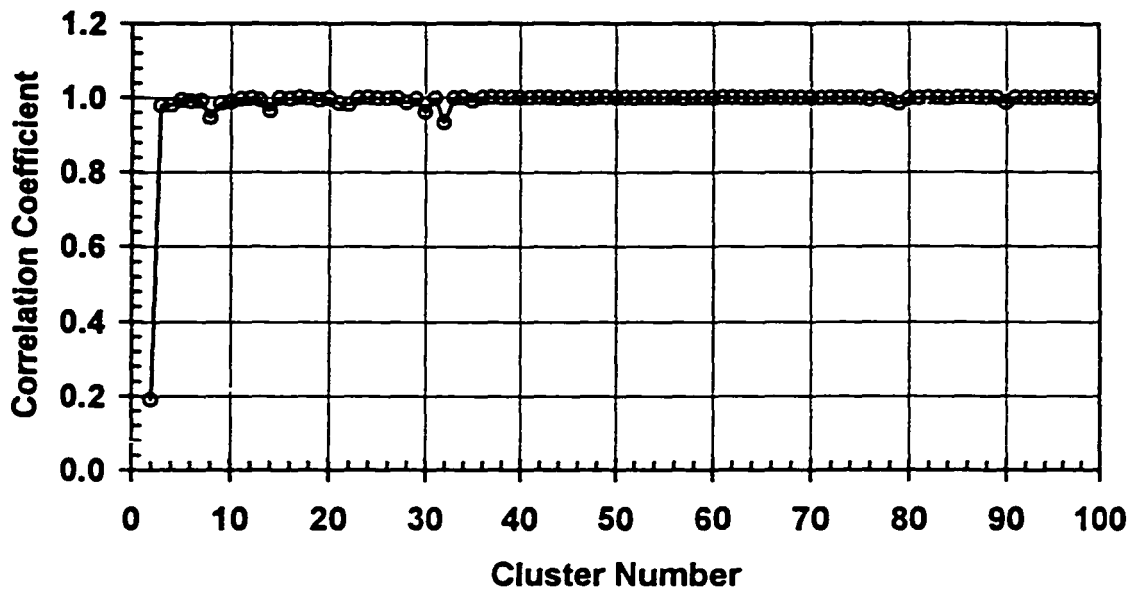
C.9. Drammen Test Site

A single-cosine-zscore cluster method is applied to piezocone data from the Drammen site (Masood et al., 1990) using Q and B_q . Piezocone data were previously shown in Fig. 6.1. The correlation coefficient between consecutive cluster results is shown in Fig. C.37 up to cluster number $N_c = 100$. The growth of data clustering at the peaks of ρ_c is shown in Fig. C.38a between $N_c = 2$ to 31. At $N_c = 2$, the soil is grouped in one cluster except of a lense detected at 12.4 m. Subsequently two groups are separated with a boundary at 10.4 m. At $N_c = 10$, a primary layer separates from the upper group with a boundary at a depth of 5.3 m. For larger clusters up to $N_c = 100$, only transitions and lenses breaks off the three primary clusters as seen in Fig. C.38b which shows cluster



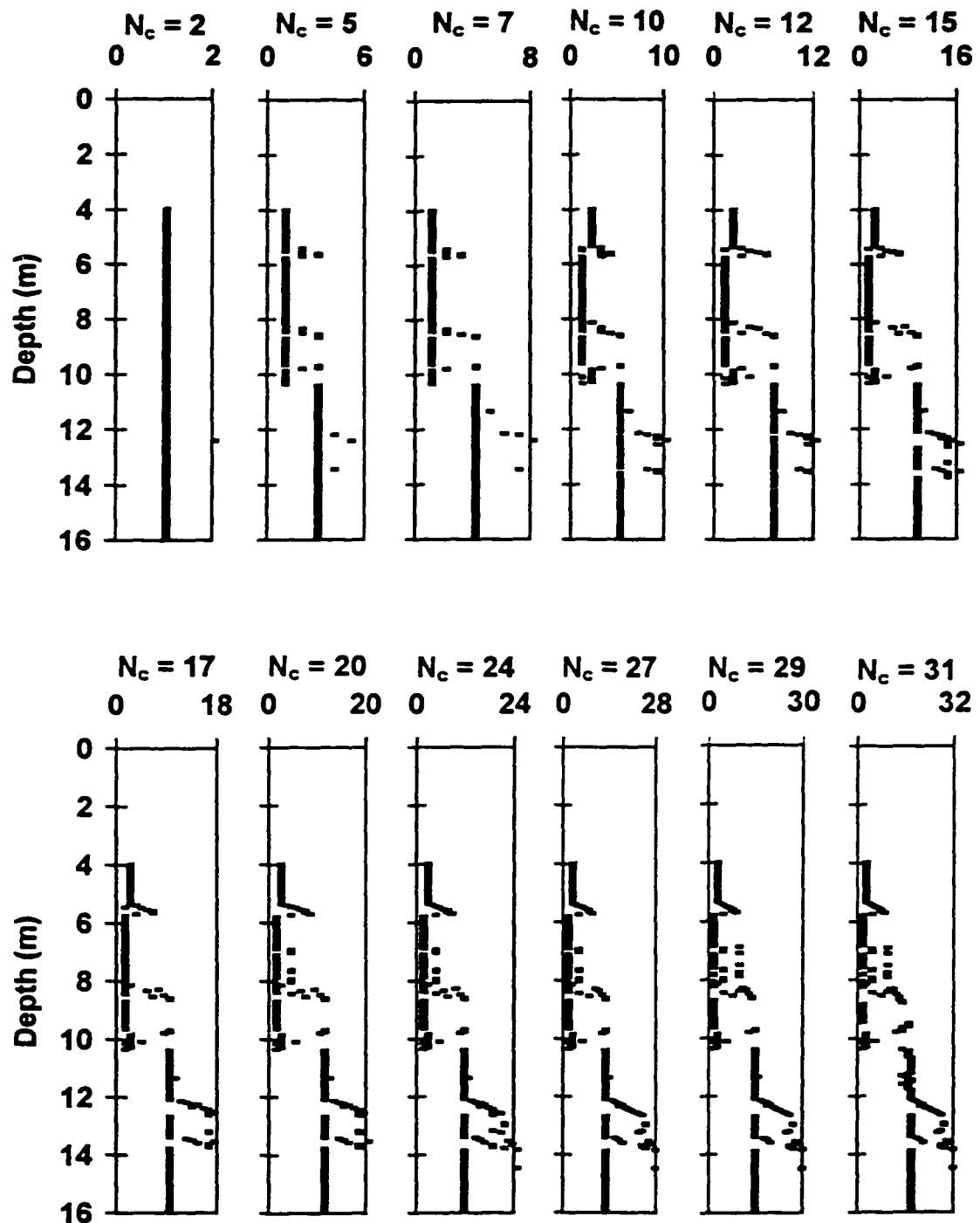
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data (PCPT15) at Amherst, Massachusetts (data from this study)

Figure C.36b. Clustering of Piezocone Sounding (PCPT15) at Amherst.



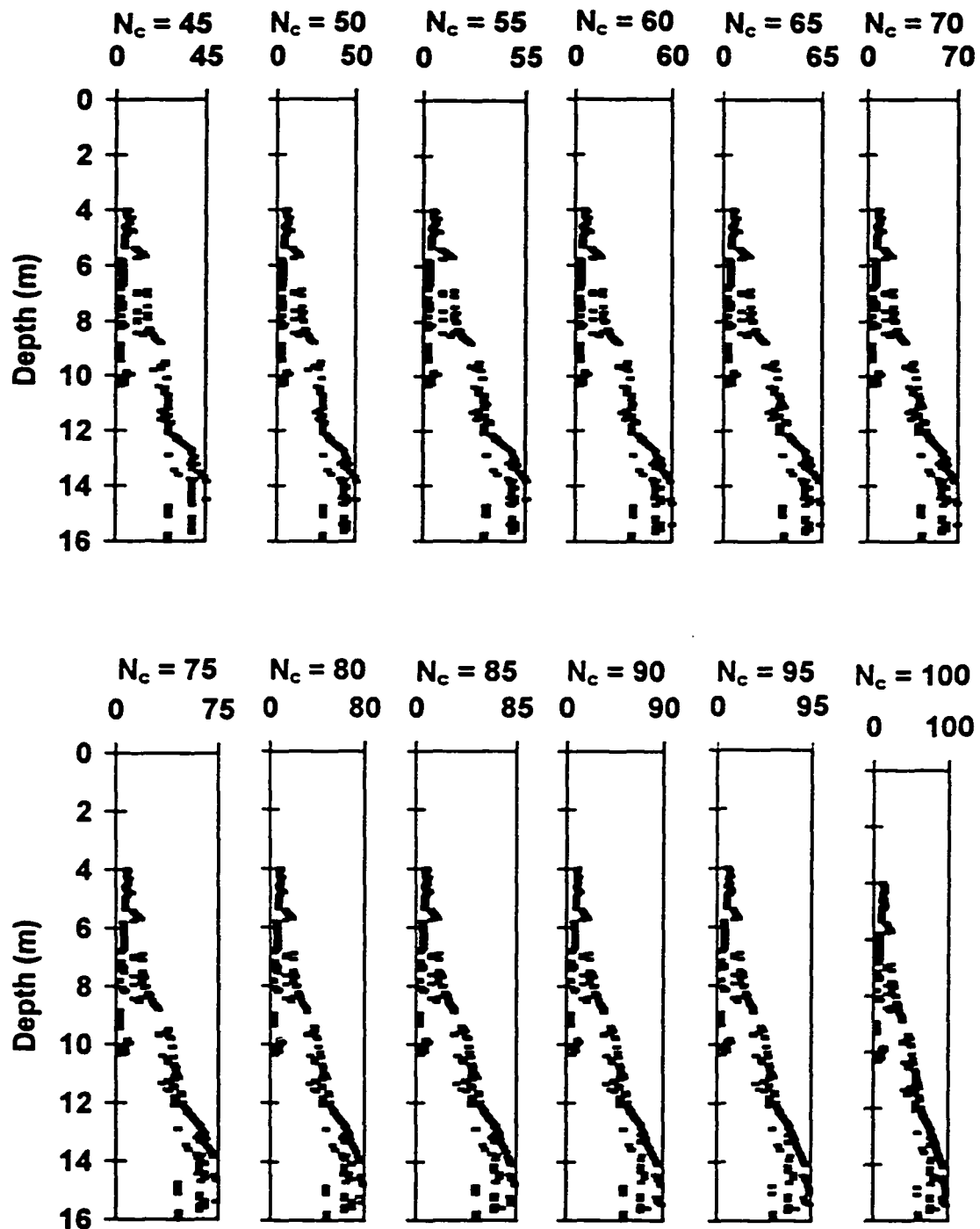
Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data at Drammen, Norway (Masood et al., 1990).

Figure C.37. Correlation Coefficient Between Consecutive Cluster Results at Drammen, Norway.



Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data at Drammen, Norway (Masood et al., 1990).

Figure C.38a. Cluster Analysis of Piezocone Data at Drammen, Norway.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Drammen, Norway (Masood et al., 1990).

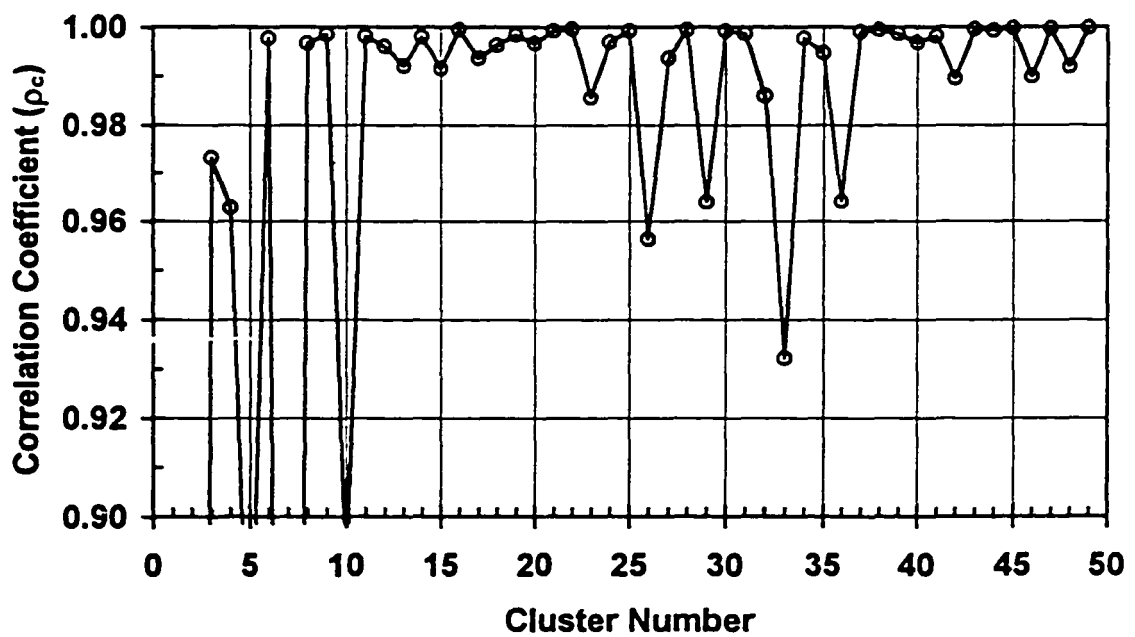
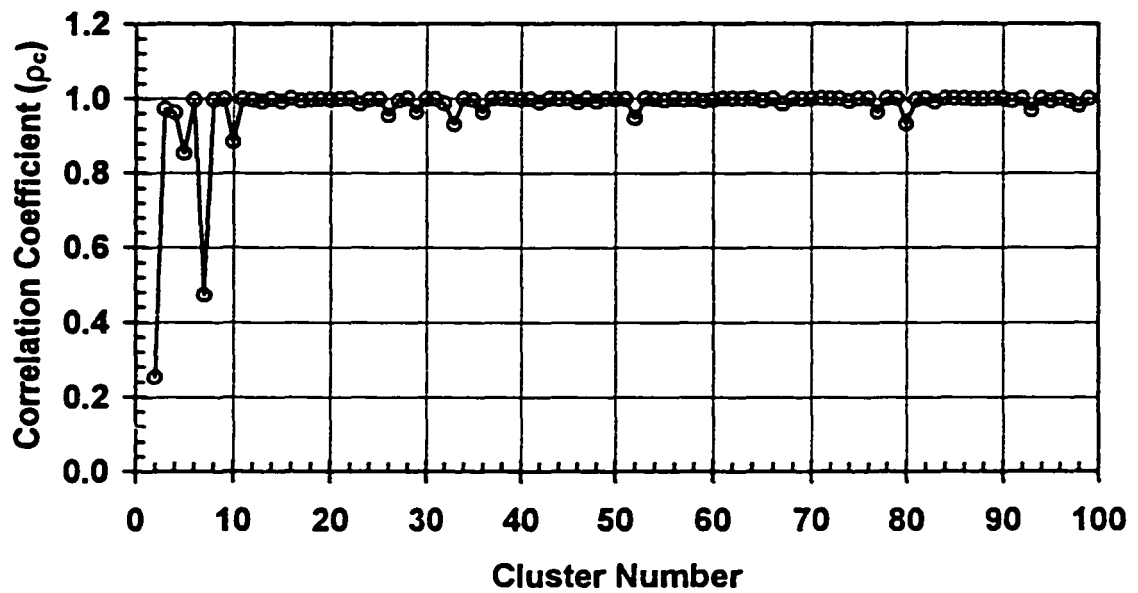
Figure C.38b. Cluster Analysis of Piezocone Data at Drammen, Norway.

results between $N_c = 45$ and 100 every 5 increments. Therefore, a cluster number 10 is chosen to represent the soil stratigraphy at the site.

C.10. Gloucester Test Site, Ontario

A single-cosine-zscore cluster method is applied to piezocone data from the Gloucester test site in Ontario (Konrad and Law, 1987) using Q and B_q . Piezocone data were previously presented in Fig. 6.6. The correlation coefficient between consecutive cluster results are shown in Fig. C.39 up to $N_c = 100$. The growth of data groups is studied at the peaks of ρ_c . The cluster results are shown in Fig. C.40a between cluster number $N_c = 2$ and 30. At $N_c = 2$, the stratigraphy is shown as one primary layer except a soil lense is detected between 6.3 m and 6.35 m. At $N_c = 3$, the data are separated into 4 groups with alternative cluster numbers 1 and 2. Then some points separate from cluster number 3 and the chain between the second and the fourth groups (sorted from the top to the bottom) breaks. The second group is assigned a cluster number 3 and the fourth group is assigned a cluster number 6 (see Fig. C.40a, $N_c = 6$).

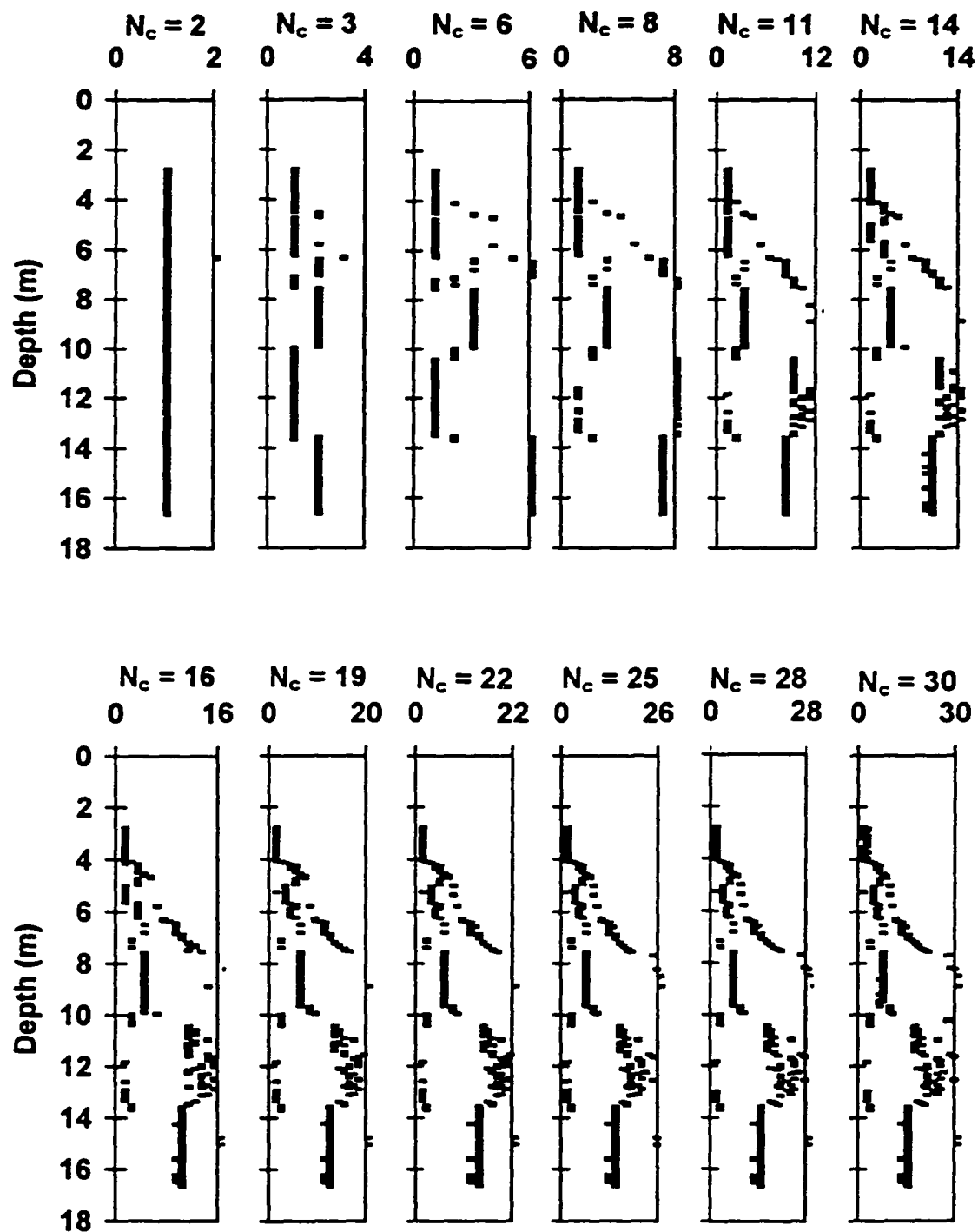
At higher clusters, a similar break happens in the chain between groups 1 and 3. The third group is assigned a cluster number 8 (see Fig. C.40a, $N_c = 8$). Subsequently, no primary layers ($t > 1$ m) separate at $N_c > 8$ as also shown in Fig. C.40b which includes cluster results between $N_c = 45$ and 100 every 5 increments. Therefore cluster number 8 is chosen to represent the soil stratigraphy at the site. The first and second data groups, which are assigned cluster numbers 1 and 3, are relatively correlated to each other. The



Single-Cosine-Zscore Method Using Q and B_q

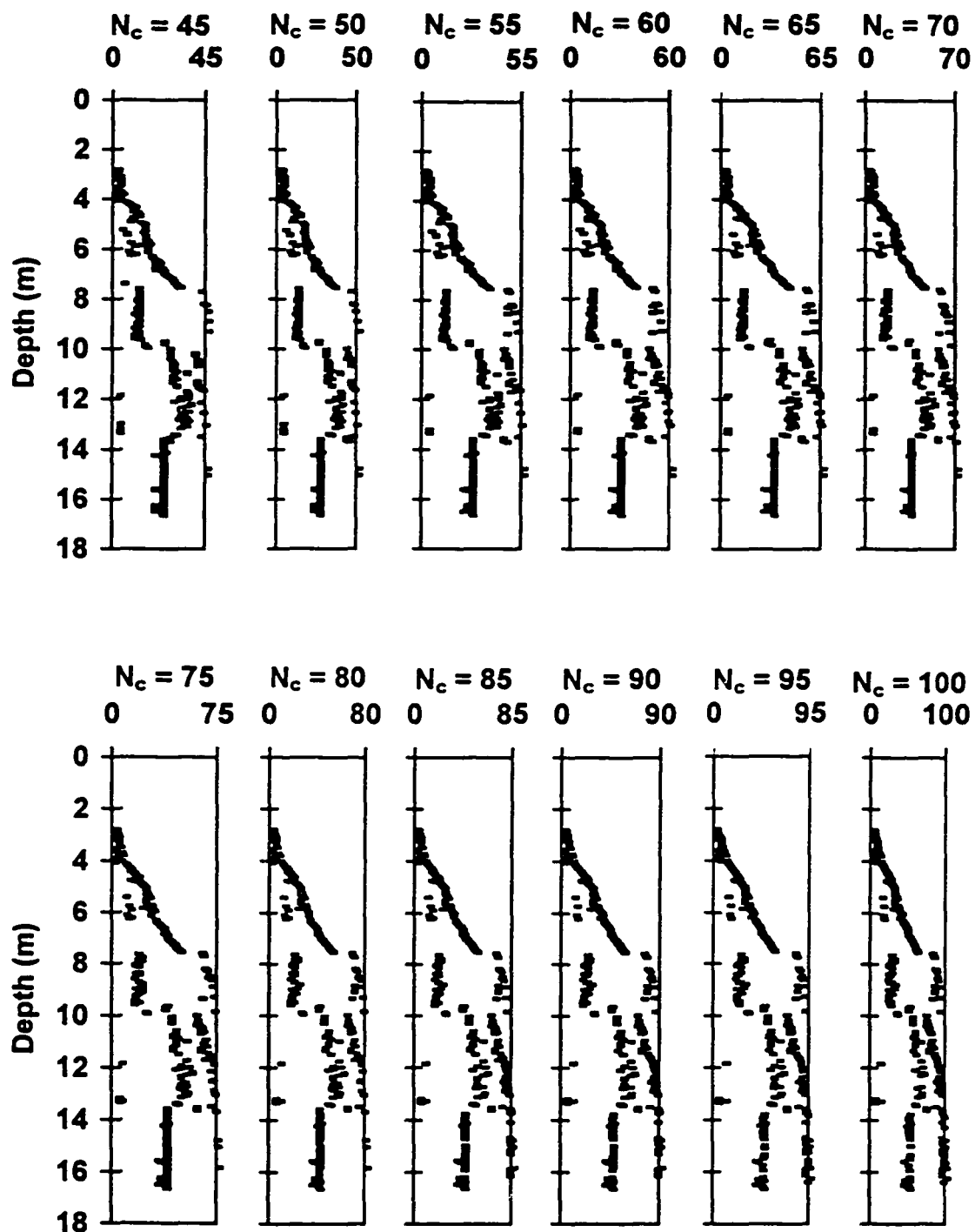
Note: Piezocone data at Gloucester, Ontario (Konrad and Law, 1987).

Figure C.39. Correlation Coefficient Between Consecutive Cluster Results at Gloucester, Ontario.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Gloucester, Ontario (Konrad and Law, 1987).

Figure C.40a. Cluster Analysis of Piezocone Data at Gloucester, Ontario.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Gloucester, Ontario (Konrad and Law, 1987).

Figure C.40b. Cluster Analysis of Piezocone Data at Gloucester, Ontario.

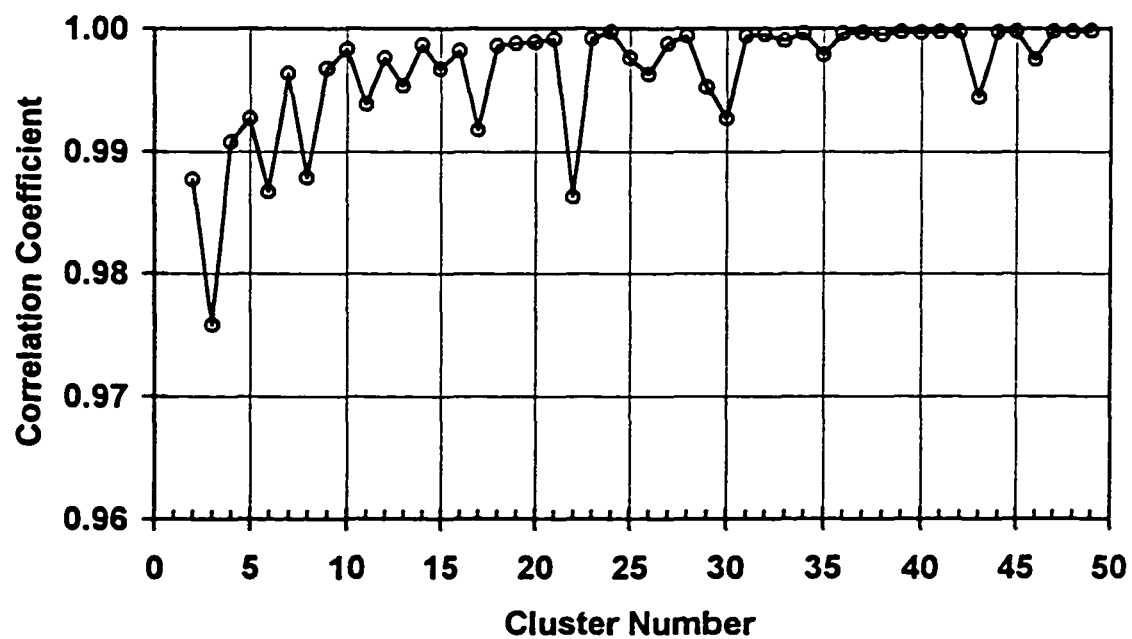
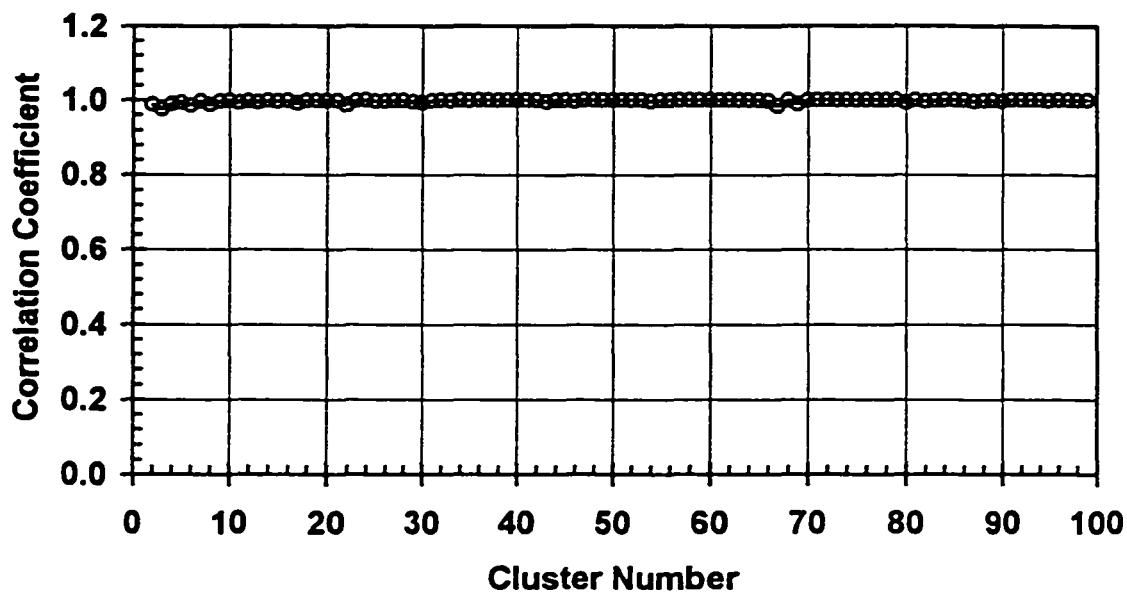
same applies to the third and fourth data groups which are assigned cluster numbers $N_c = 7$ and $N_c = 8$.

C.11. Hachirogata Test Site

A single-cosine-zscore (SCZ) cluster analysis is applied to piezocone data from Hachirogata test site (Tanaka et al., 1994) using Q and B_q . Piezocone data were previously shown in Fig. 6.12. The correlation coefficient between consecutive clusters is shown in Fig. C.41 up to cluster number $N_c = 100$. The growth of cluster results is examined at the peaks of ρ_c and shown in Fig. C.42a between $N_c = 2$ and 34. At $N_c = 2$, two primary clusters appear with a boundary at 10.5 m. Then at larger clusters, soil lenses separate from the lower layer indicating dissimilarity with the primary cluster. A soil transition also starts to build up above and below 10.5 m. The upper cluster breaks into two primary clusters at $N_c = 4$ with a boundary at a depth of 8.3 m. The main cluster between the depths of 3 m and 8.3 m breaks into four smaller secondary layers up to $N_c = 17$ indicating a transition of soil type and/or properties. However, up to $N_c = 100$, no new primary clusters ($t \geq 1$ m) are discovered as seen in Fig. C.42b which includes cluster results between $N_c = 45$ and 100 every 5 increments. Therefore a cluster number 17 is chosen to represent the soil stratigraphy at the site.

C.12. Lilla Mellösa Test Site

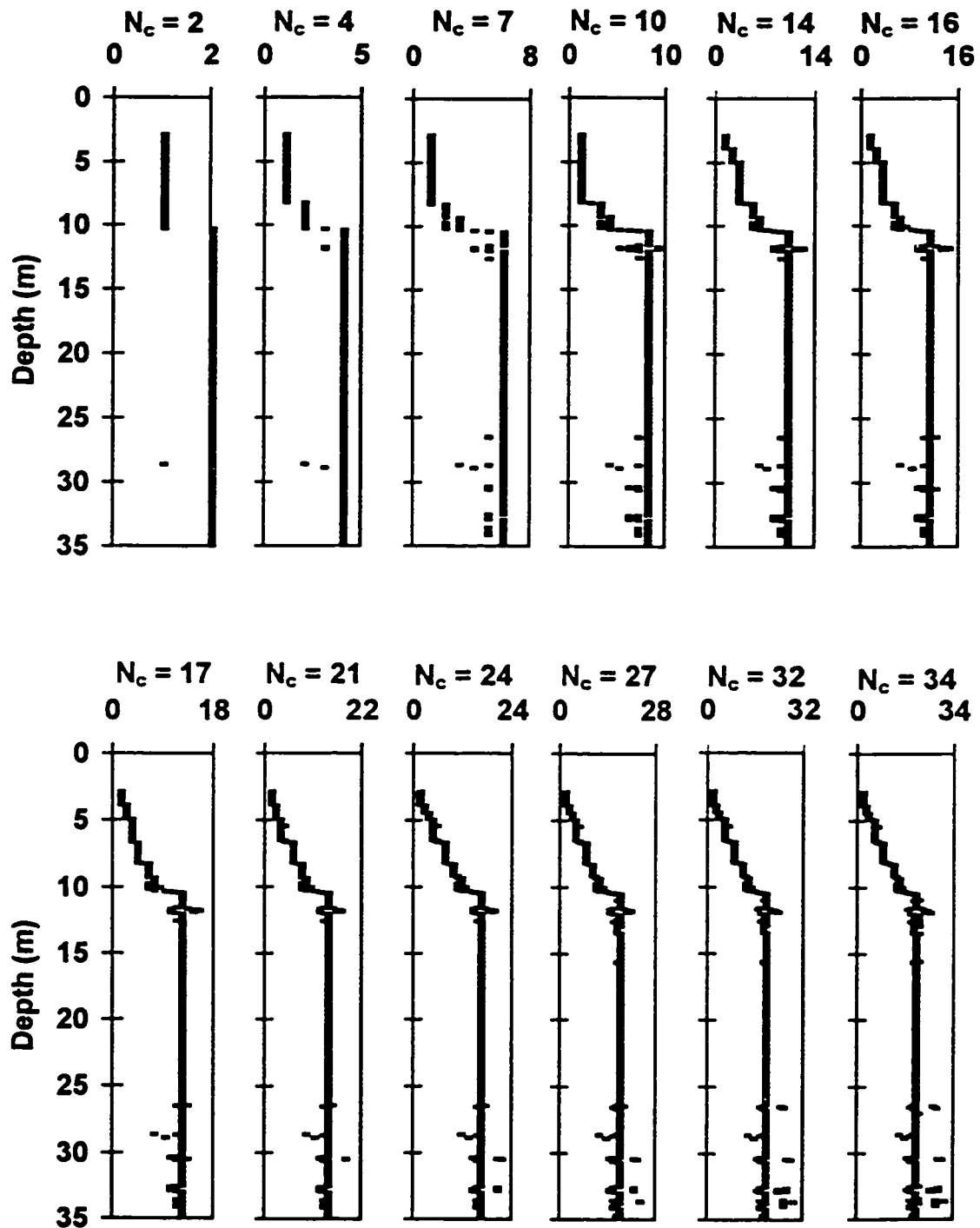
A single-cosine-zscore cluster analysis is applied to representative piezocone data from Lilla Mellösa test site (Larsson and Mulabdić, 1991) using the normalized parameters



Single-Cosine-Zscore Method Using Q and B_q

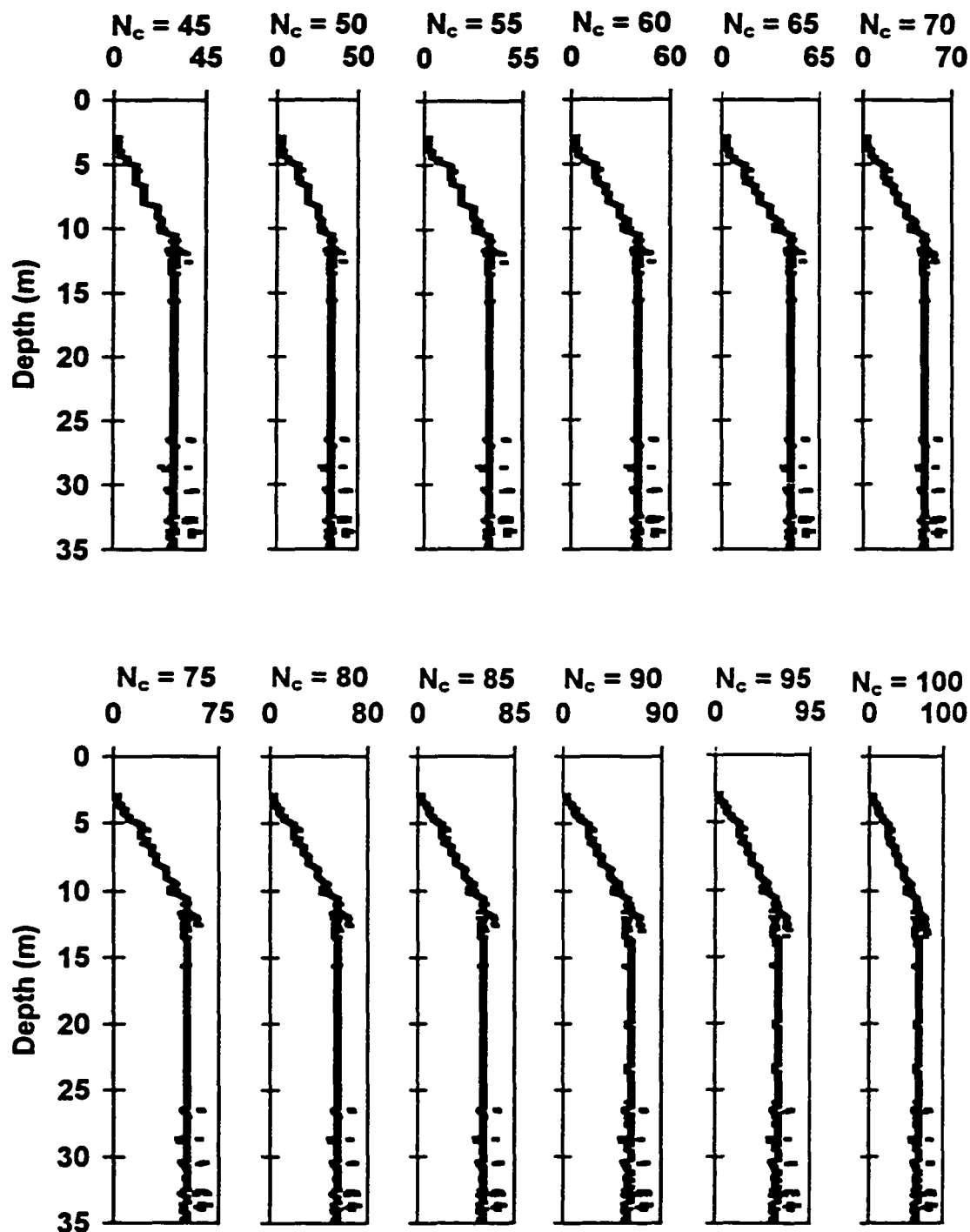
Note: Piezocone data at Hachirogata, Japan from Tanaka et al. (1992).

Figure C.41. Correlation Coefficient Between Consecutive Cluster Results at Hachirogata, Japan.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Hachirogata, Japan from Tanaka et al. (1992).

Figure C.42a. Cluster Analysis of Piezocone Data at Hachirogata, Japan.



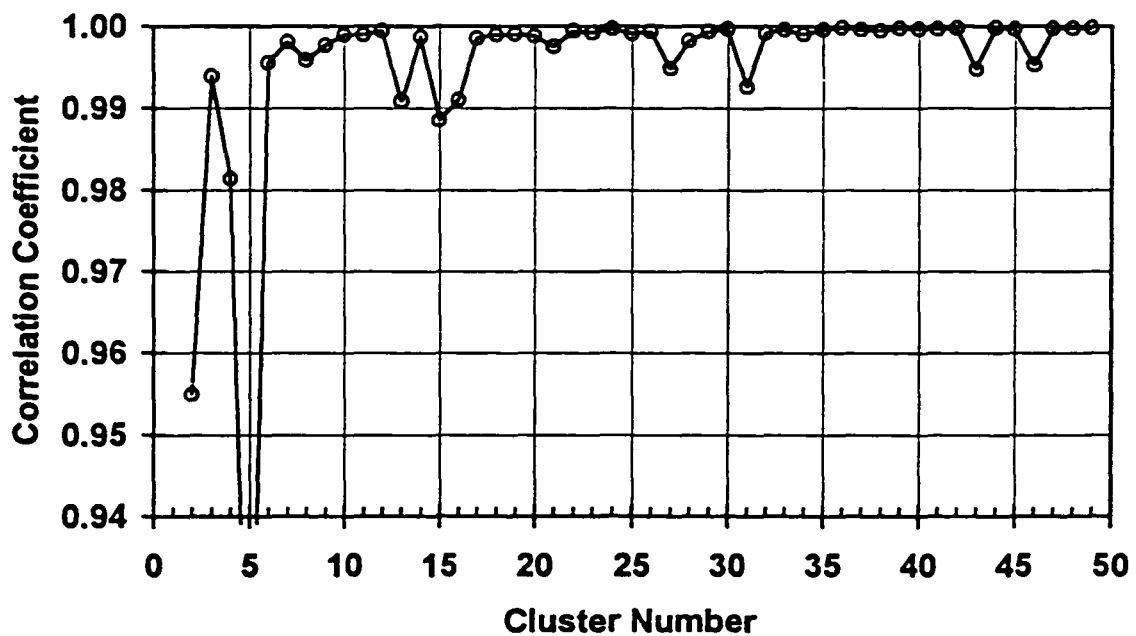
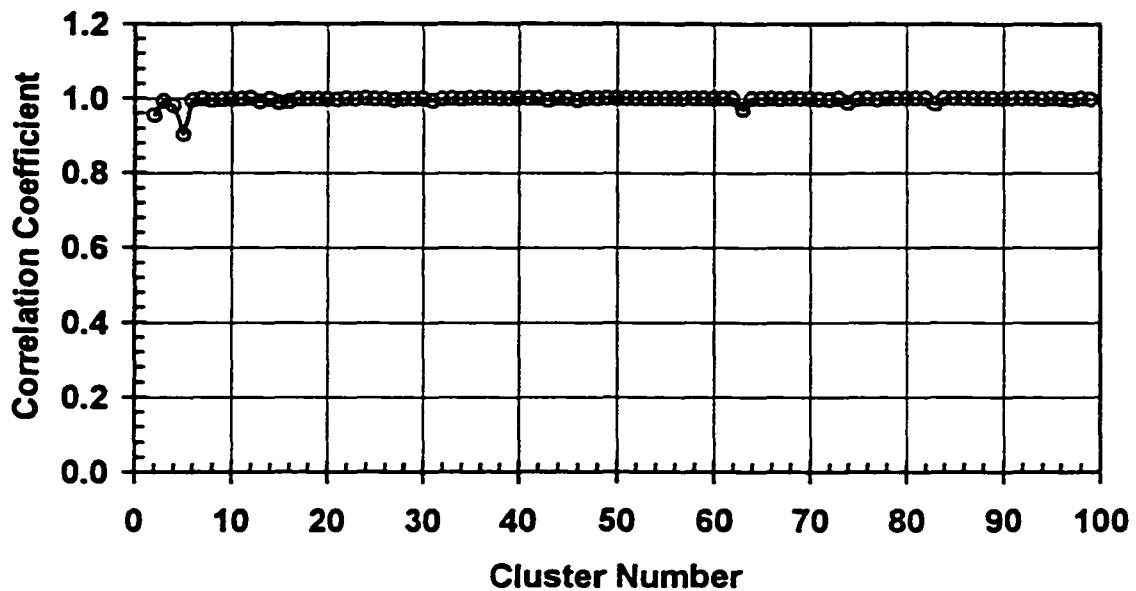
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Hachirogata, Japan from Tanaka et al. (1992).

Figure C.42b. Cluster Analysis of Piezocone Data at Hachirogata, Japan.

Q and B_q . Piezocone data were previously shown in Fig. 6.16. The correlation coefficient between consecutive clusters is shown in Fig. C.43 up to cluster number $N_c = 100$. Clustering is examined at the peaks of ρ_c and cluster results are shown in Fig. C.44a between $N_c = 2$ and 36. Two primary soil layers are given at cluster number 2 and a transition layer is also defined between them from 5.9 m to 6.35 m. Then some points having a relatively large dissimilarity with the two primary clusters separate indicating soil lenses, transitions or outliers. Up to $N_c = 100$ as shown in Fig. C.44b which includes cluster results between 45 and 100 every 5 increments, no new primary clusters ($t \geq 1$ m) are detected. Therefore a cluster number 2 is chosen to represent the soil stratigraphy at the site.

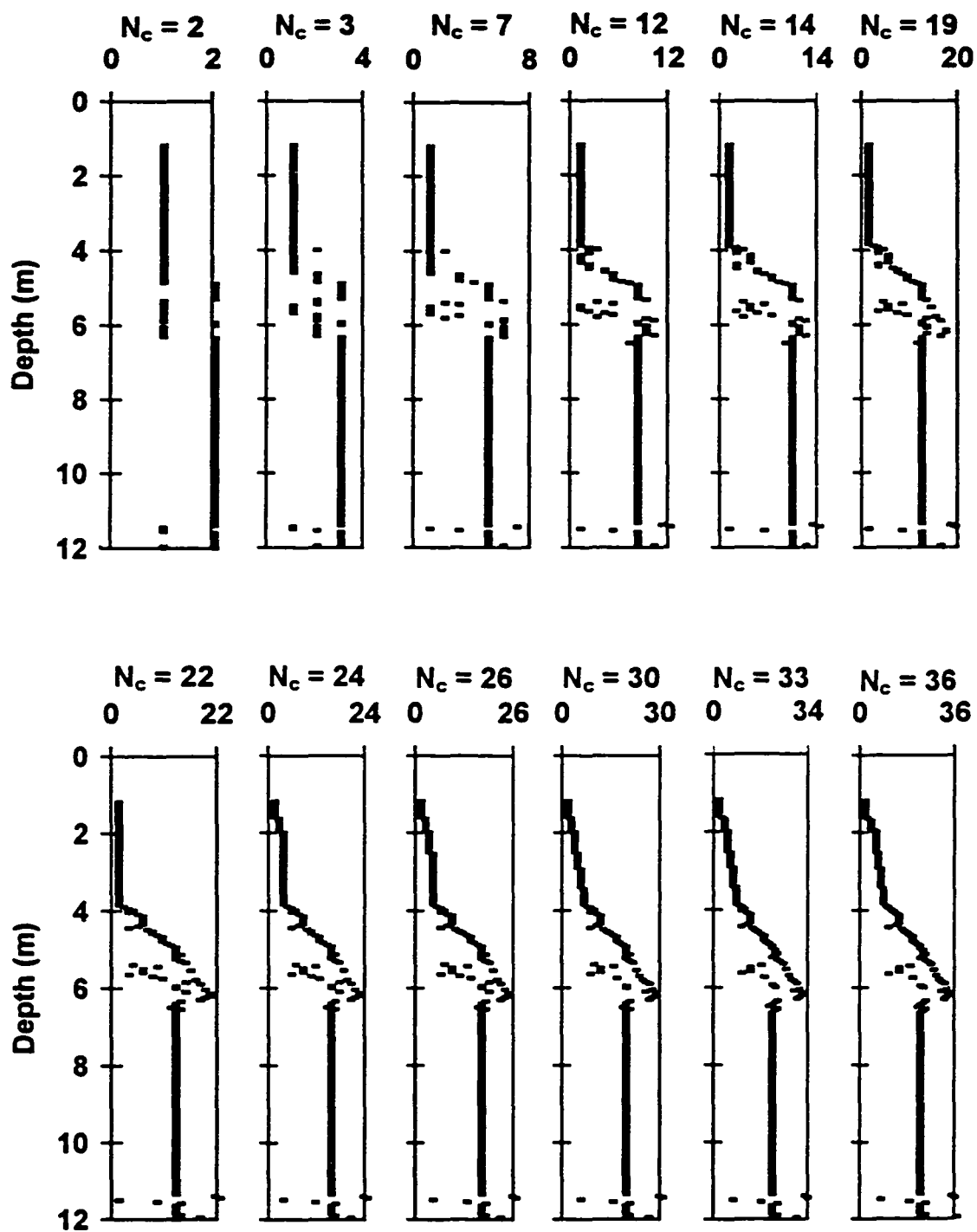
C.13. Recife Test Site

A single-cosine-zscore cluster method is applied to a representative piezocone data from Recife test site (Coutinho and Oliveira, 1997) using the derived normalized parameters Q and B_q . Piezocone data were previously presented in Fig. 6.20. The correlation coefficient between consecutive clusters is shown in Fig. C.45 up to cluster number $N_c = 100$. The cluster results are examined at peaks of ρ_c and shown in Fig. C.46a between $N_c = 2$ and 35. At N_c equal to 2, two primary data groups are detected with a boundary at 15.6 m and also a soil lense is found between 21.9 m and 22.4 m. For higher clusters, some points having relatively large cosine measurements continue to separate from the two main statistical layers. However, up to $N_c = 100$, no new primary clusters (t



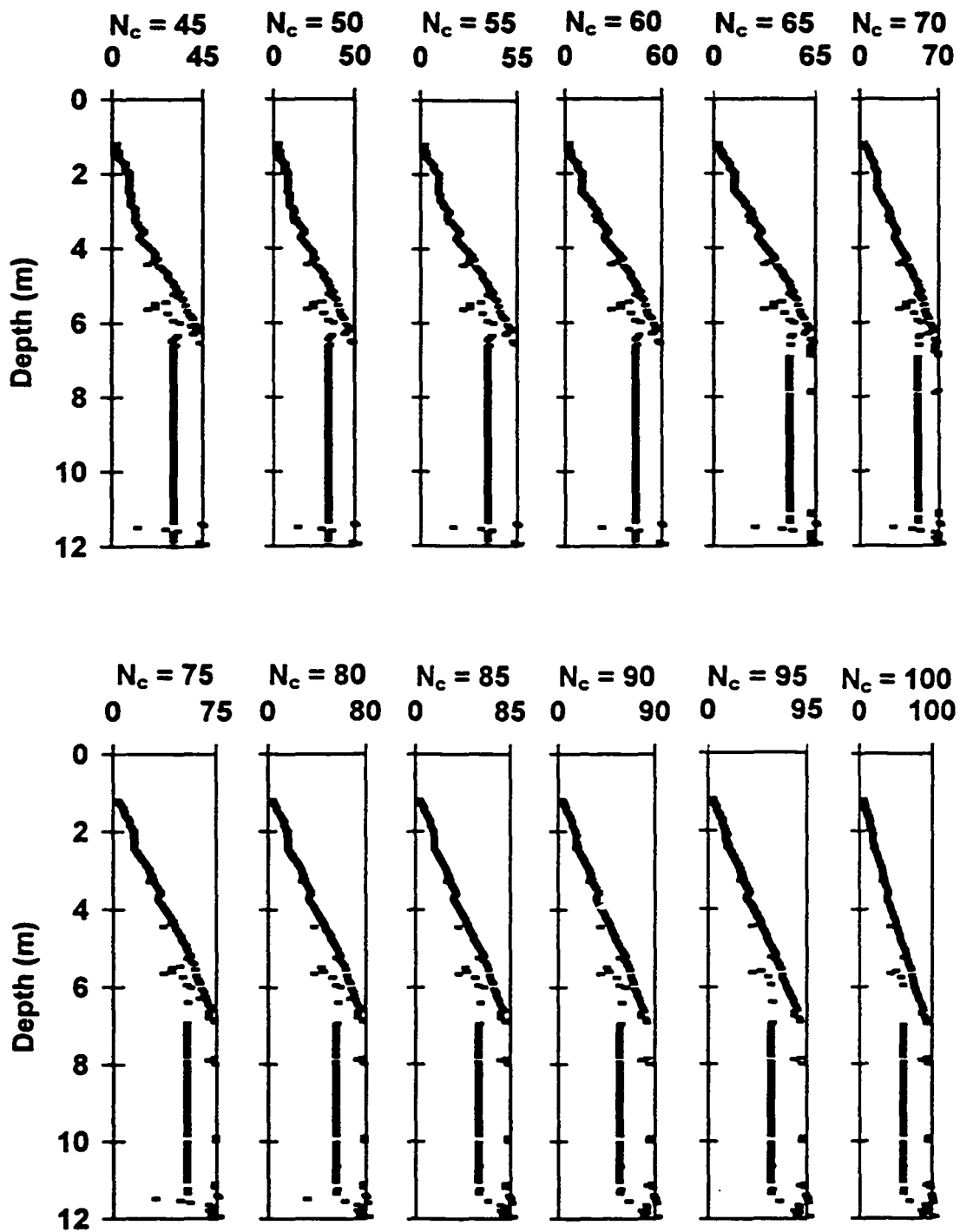
Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data at Lilla Mellosa, Sweden (Larsson and Mulabdic, 1991).

Figure C.43. Correlation Coefficient Between Consecutive Cluster Results at Lilla Mellosa, Sweden.



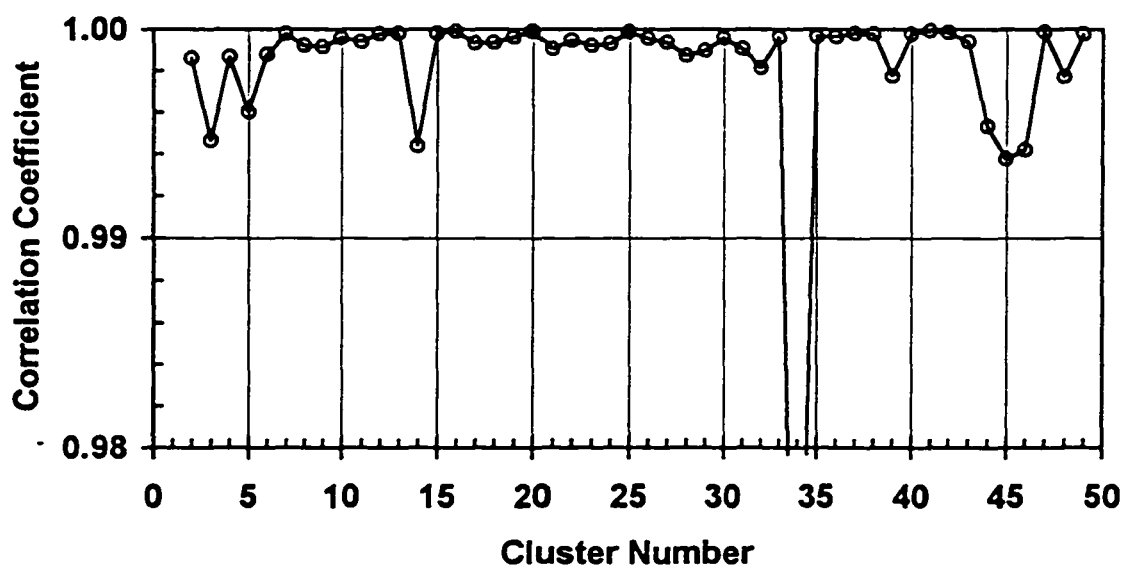
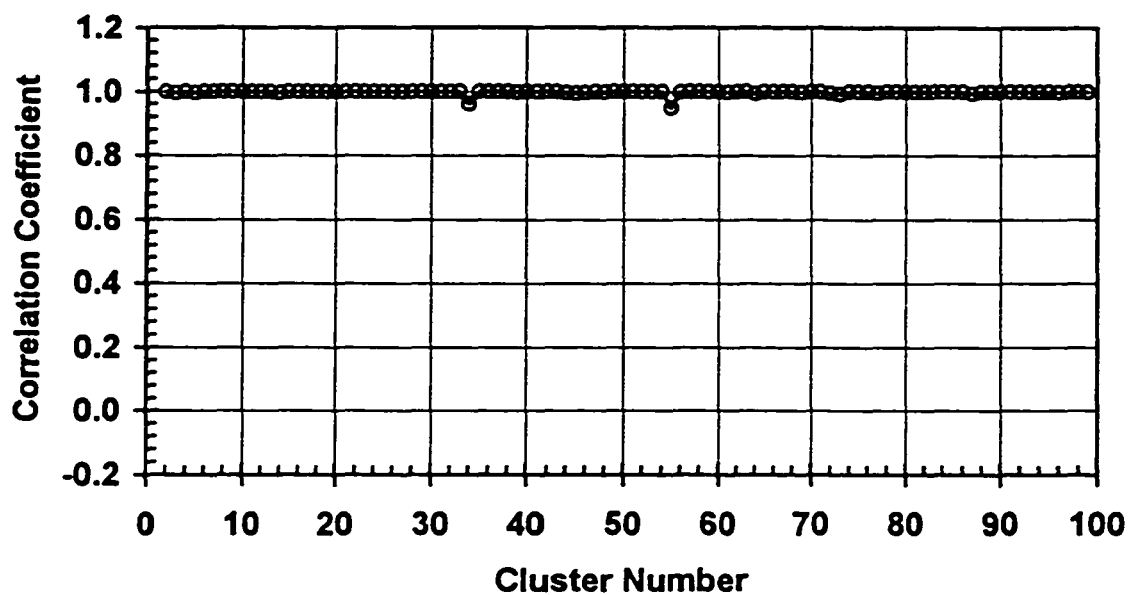
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Lilla Mellosa, Sweden (Larsson and Mulabdic, 1991).

Figure C.44a. Cluster Analysis of Piezocone Data at Lilla Mellosa, Sweden.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Lilla Mellosa, Sweden (Larsson and Mulabdic, 1991).

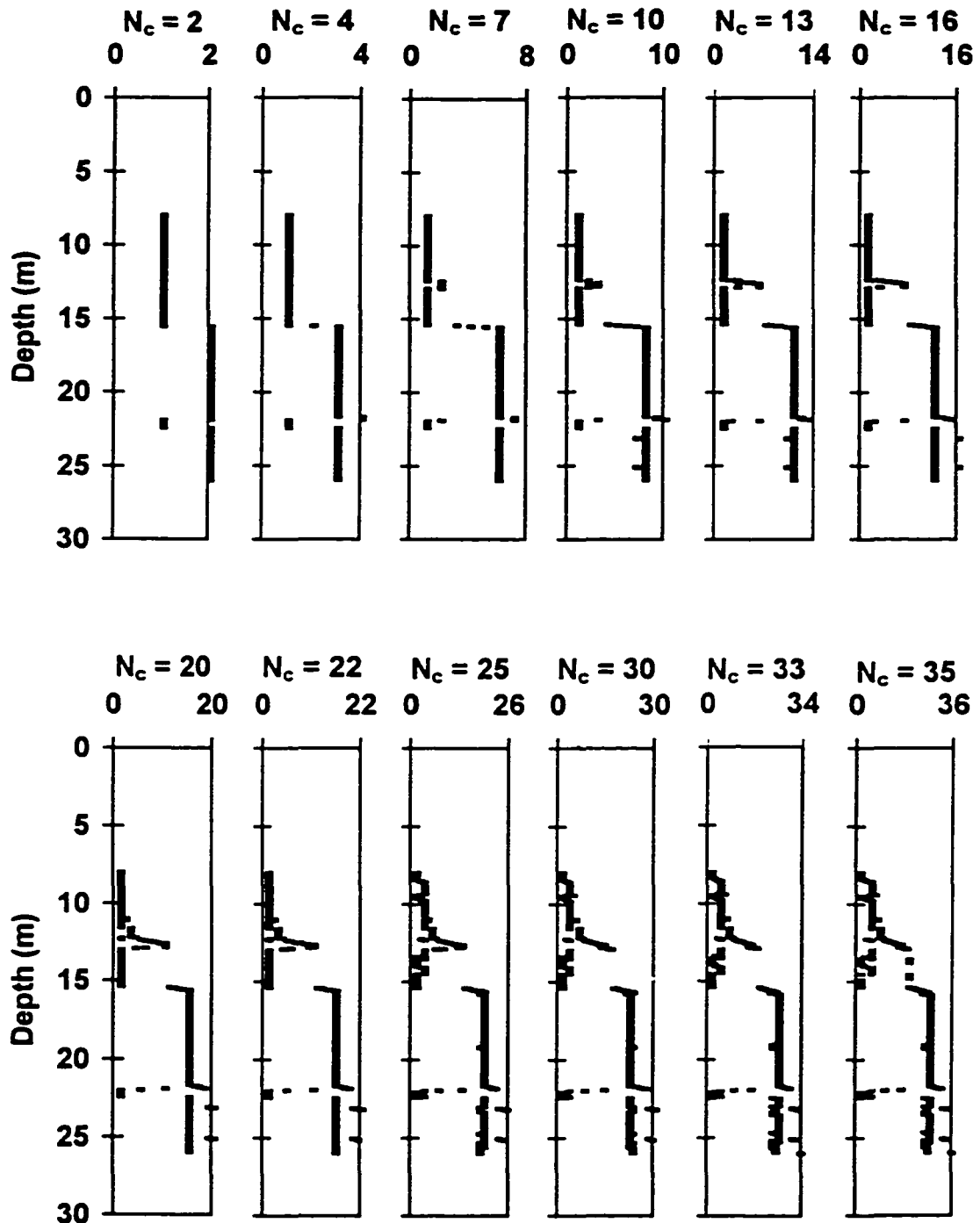
Figure C.44b. Cluster Analysis of Piezocone Data at Lilla Mellosa, Sweden.



Single-Cosine-Zscore Method Using Q and B_q

Note: Piezocone data at Recife, Brazil from Coutinho and Oliveria (1997).

Figure C.45. Correlation Coefficient Between Consecutive Cluster Results at Recife, Brazil.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Recife, Brazil from Coutinho and Oliveria (1997).

Figure C.46a. Cluster Analysis of Piezocone Data at Recife, Brazil.

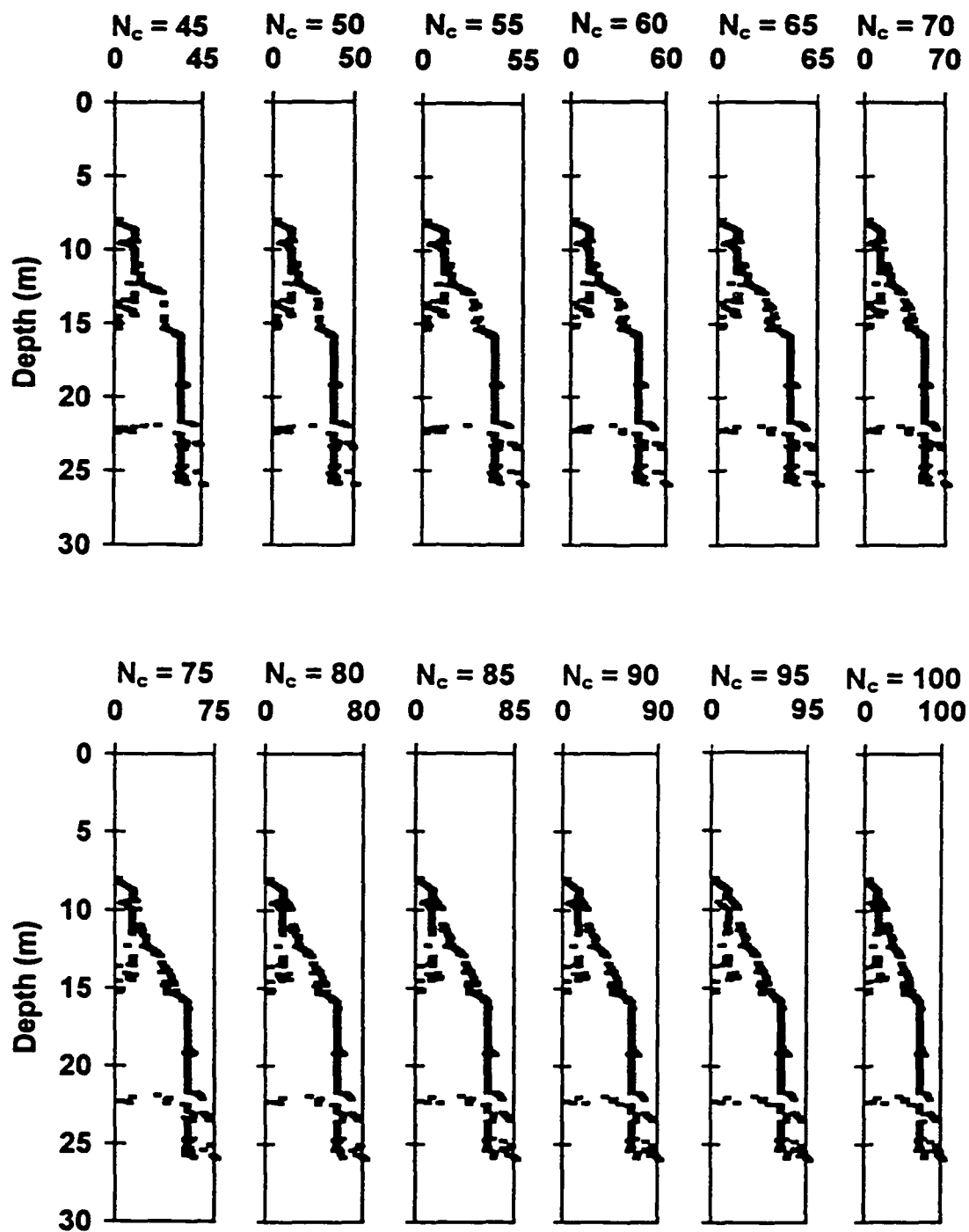
≥ 1 m) are discovered as shown in Fig. C.46b which includes cluster results between $N_c = 45$ and 100 every 5 increments. Therefore cluster number 2 is chosen to represent the subsurface stratification at the site.

C.14. St. Alban Test Site, Quebec

A single-cosine-zscore cluster analysis is applied to representative piezocone data from St. Alban test site (Roy et al., 1982) using the normalized parameters Q and B_q . Piezocone data were previously shown in Fig. 6.25. The correlation coefficient between consecutive clusters is shown in Fig. C.47 up to cluster number $N_c = 100$. The cluster results are examined at the peaks of ρ_c and shown in Fig. C.48a between $N_c = 2$ and 34. At N_c equal to 2, two primary clusters are separated with a boundary at a depth of 5.2 m. Two soil lenses appear in the upper layer at 3.90 m, and 4.20 m to 4.25 m. A part of a soil transition between the two primary layers starts to build up at a depth of 5.15 m. At larger cluster numbers, more soil lenses and transitions continue to separate of the two original clusters but no a new primary group ($t \geq 1$ m) appears. Also, the latter is applied for clusters between 45 and 100 as shown in Fig. C.48b. Therefore, a cluster number 2 is chosen to represent the soil stratigraphy at the site.

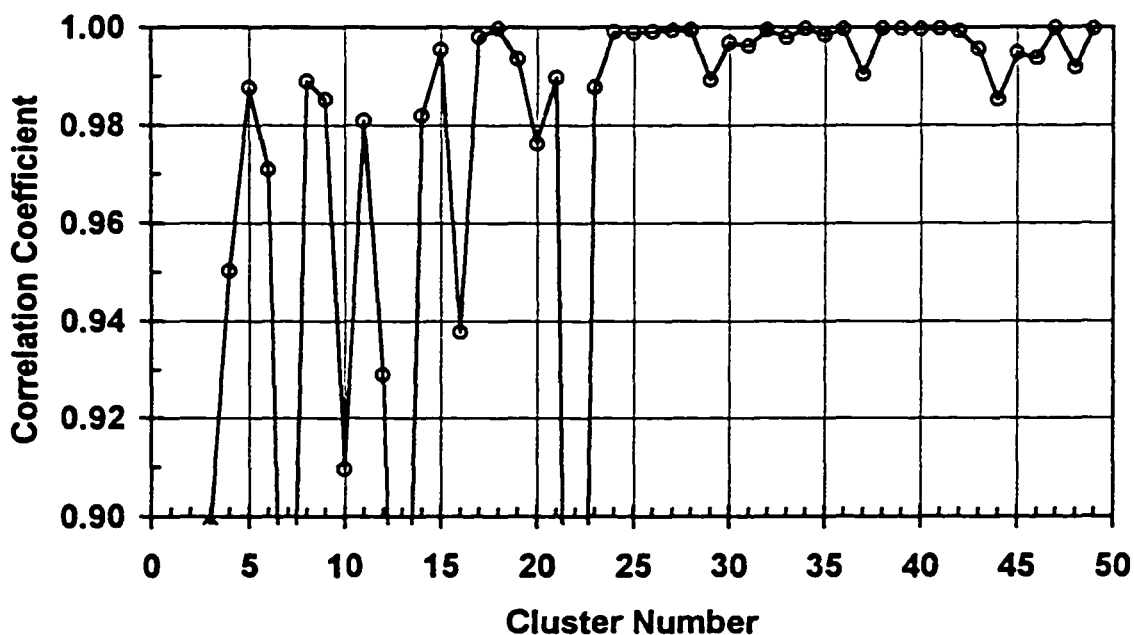
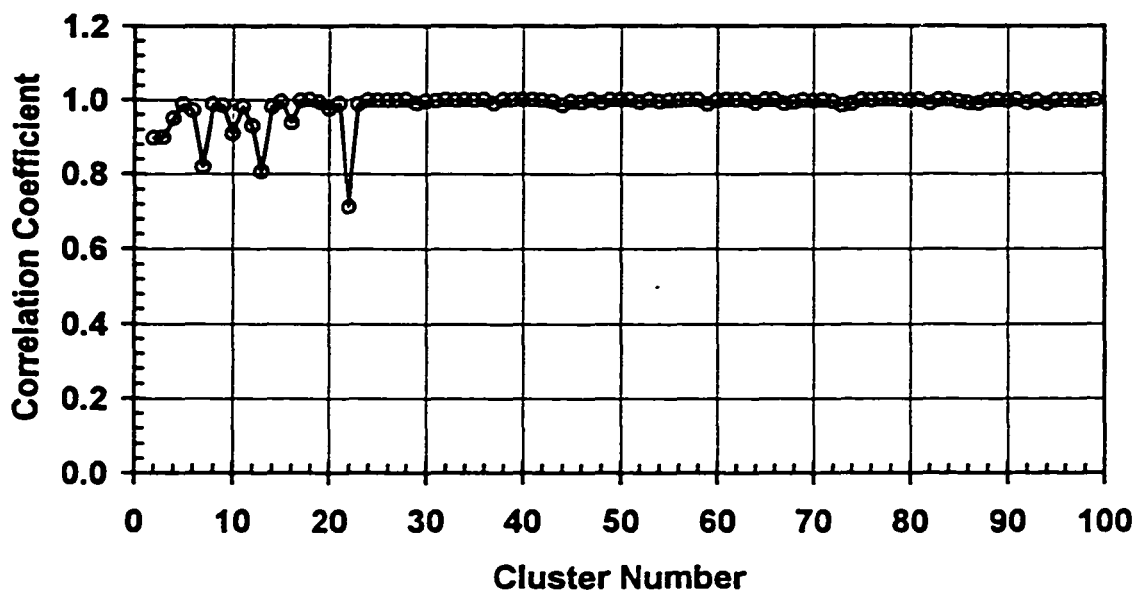
C.15. Tiller Test Site

A single-cosine-zscore cluster analysis is applied to representative piezocone data from Tiller test site (Sandven, 1990) using the normalized parameters Q and B_q .



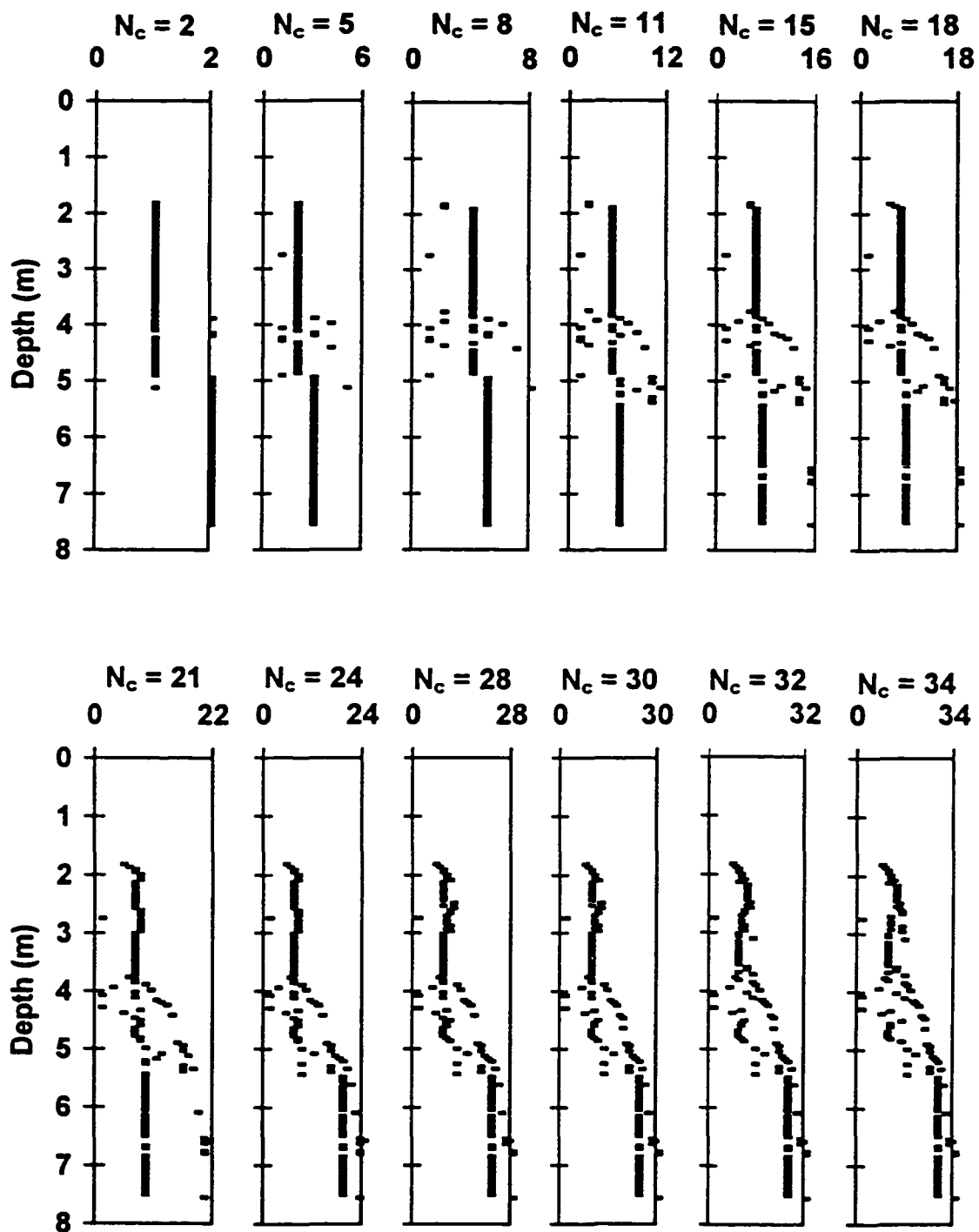
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Recife, Brazil from Coutinho and Oliveria (1997).

Figure C.46b. Cluster Analysis of Piezocone Data at Recife, Brazil.



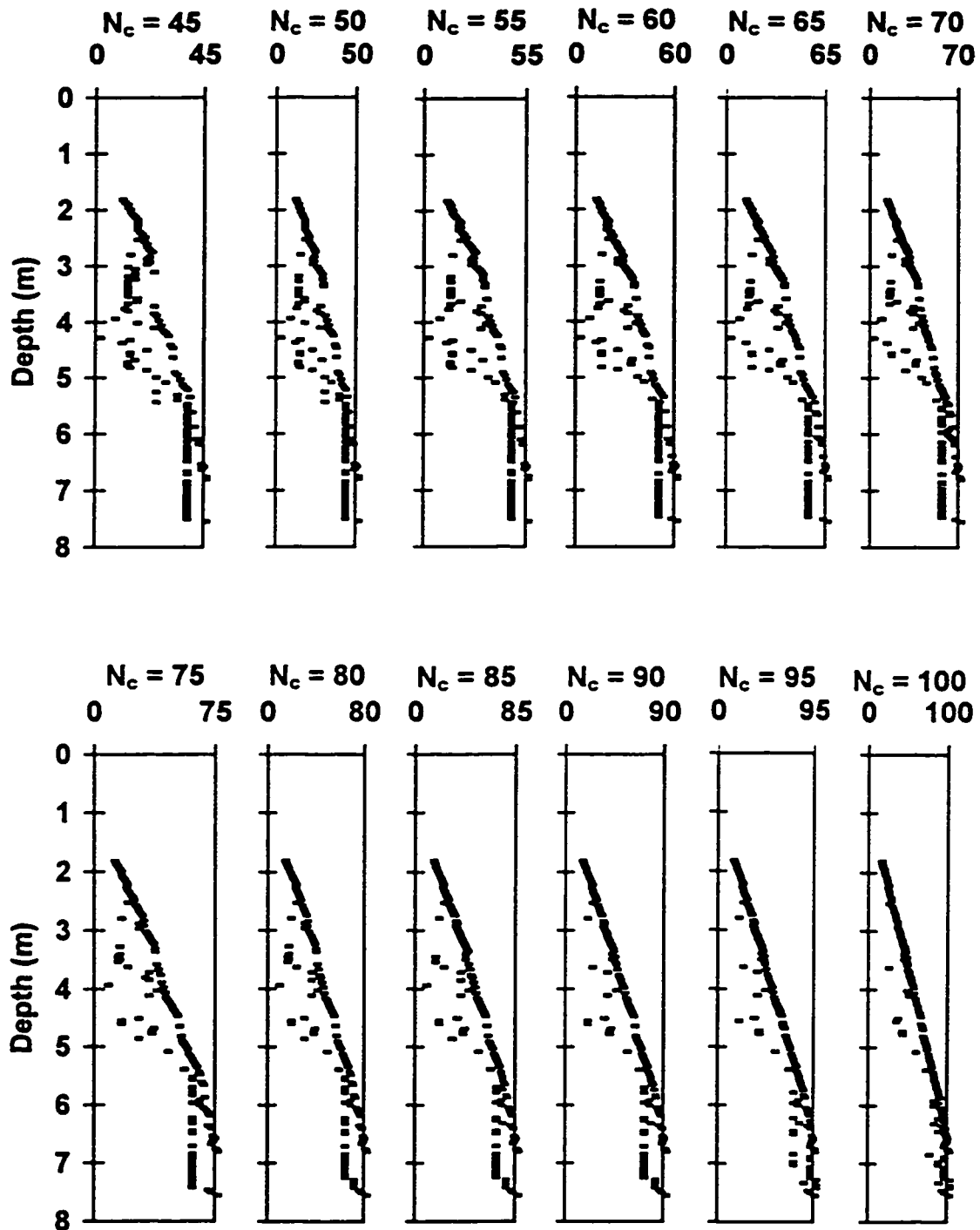
Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data at St. Alban, Quebec from Roy et al. (1982).

Figure C.47. Correlation Coefficient Between Consecutive Cluster Results at St. Alban, Quebec.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at St. Alban, Quebec from Roy et al. (1982).

Figure C.48a. Cluster Analysis of Piezocone Data at St. Alban, Quebec.



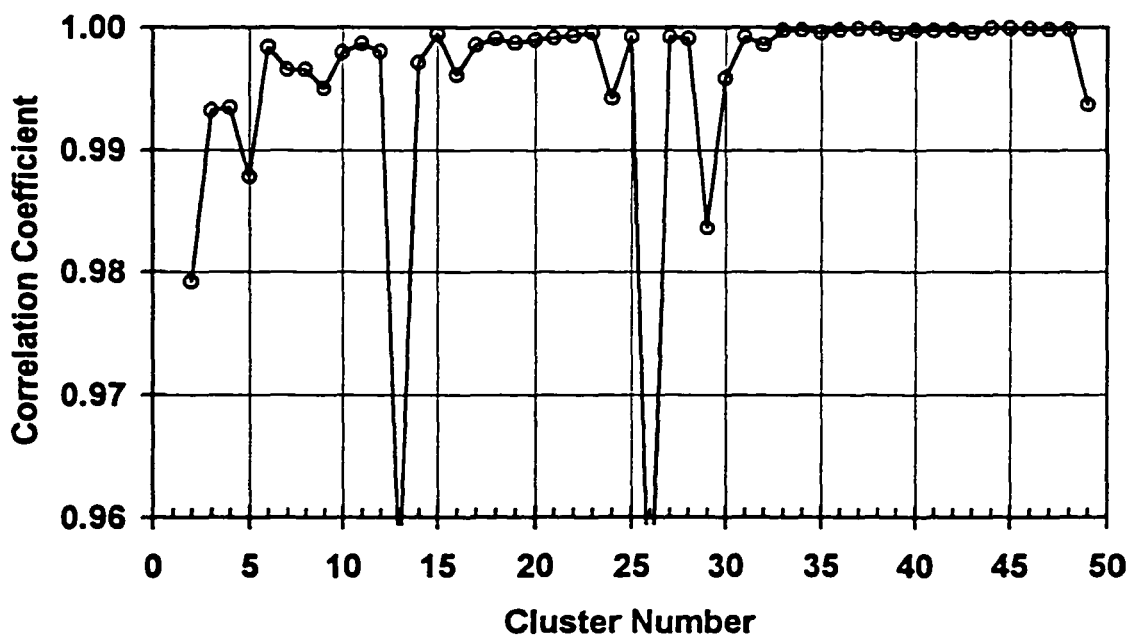
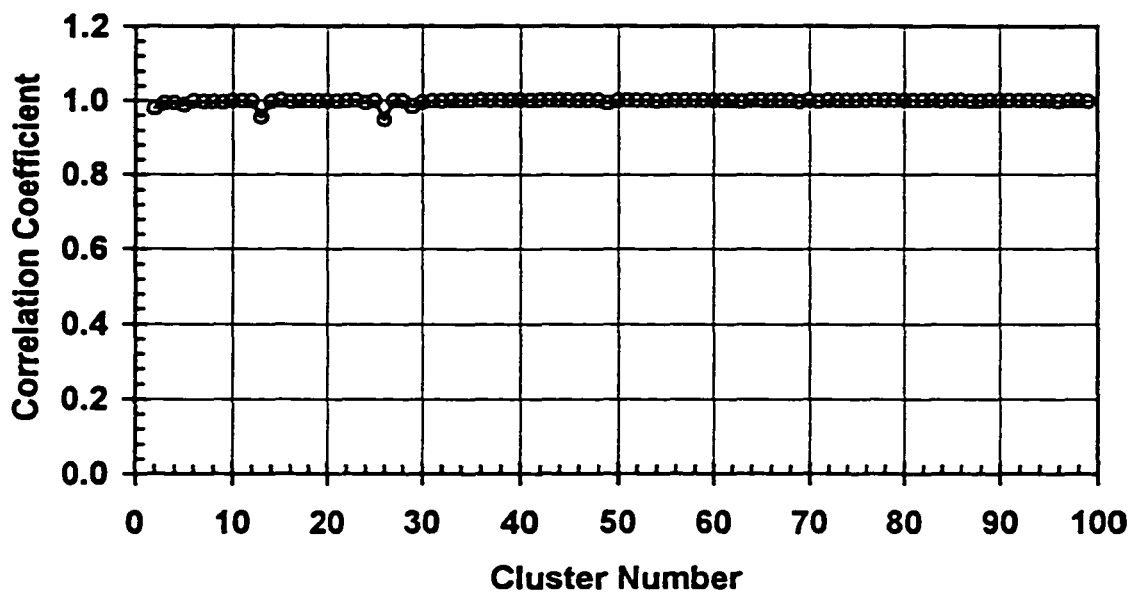
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at St. Alban, Quebec from Roy et al. (1982).

Figure C.48b. Cluster Analysis of Piezocone Data at St. Alban, Quebec.

Piezocone data were previously presented in Fig. 6.29. The correlation coefficient between consecutive clusters is shown in Fig. C.49 up to cluster number $N_c = 100$. The cluster results are studied at the peaks of ρ_c and shown in Fig. C.50a between $N_c = 2$ and 38. At N_c equal 2, the data are grouped into two clusters with a boundary at 8.5 m. At higher clusters, some points which have relatively large cosine measurements (dissimilarity) with the two primary clusters separate indicating soil transitions, lenses or outliers. However, up to $N_c = 100$, no new primary clusters ($t \geq 1$ m) are detected as shown in Fig. C.50b which includes cluster results between $N_c = 45$ and 100 every 5 increments. Therefore a cluster number 2 is chosen to represent the soil stratigraphy at the site.

C.16. Troll Test Site

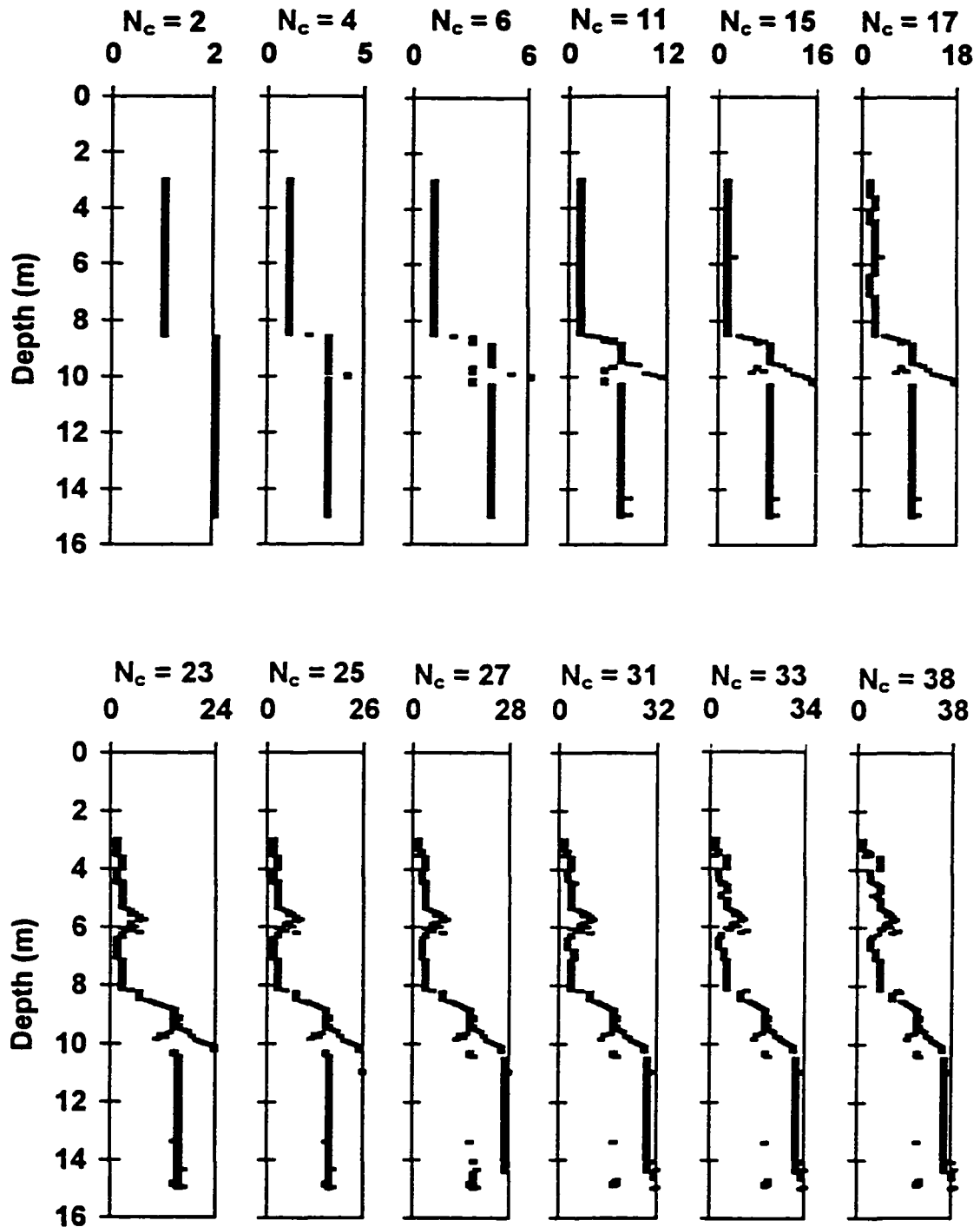
A single-cosine-zscore cluster analysis is applied to a representative piezocone data from Troll test site (Amundsen et al., 1985) using the derived normalized parameters Q and B_q . Piezocone data were previously shown in Fig. 6.33. The correlation coefficient between consecutive clusters is shown in Fig. C.51 up to cluster number $N_c = 100$. Cluster results are studied at the peaks of ρ_c and shown in Fig. C.52a between $N_c = 2$ and 41. At N_c equal to 2, two primary clusters are separated indicating a major difference in the soil types and/or properties above 17.3 m and below 20 m. In the latter intermediate zone, the cluster numbers alternates between 1 and 2 indicating a transition zone between the two major layers. For higher cluster numbers, some points having dissimilarity with



Single-Cosine-Zscore Method Using Q and B_q

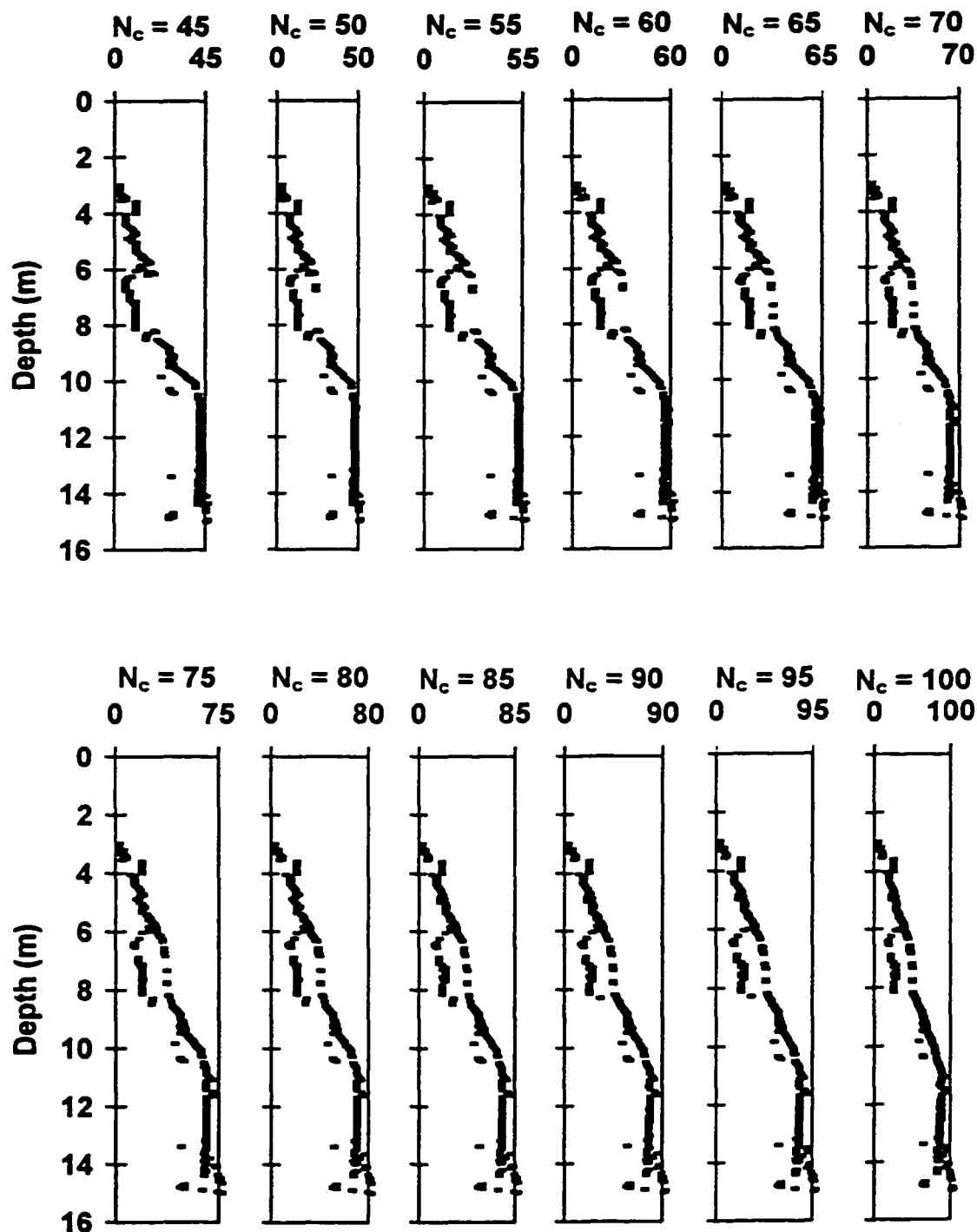
Note: Piezocone data at Tiller, Norway from Sandven (1990).

Figure C.49. Correlation Coefficient Between Consecutive Cluster Results at Tiller, Norway.



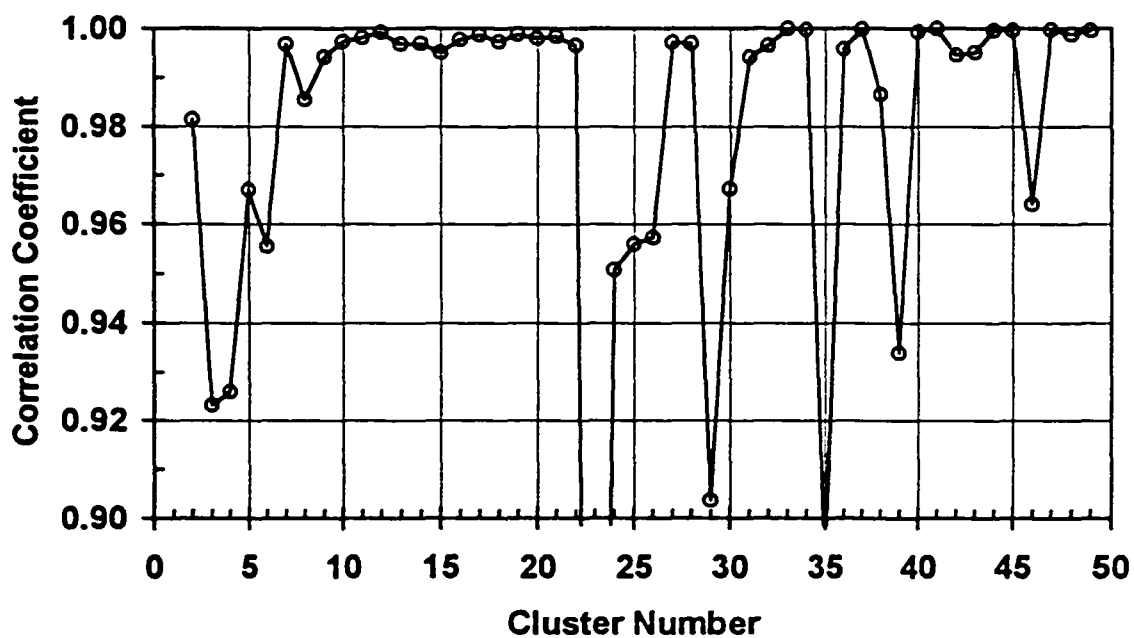
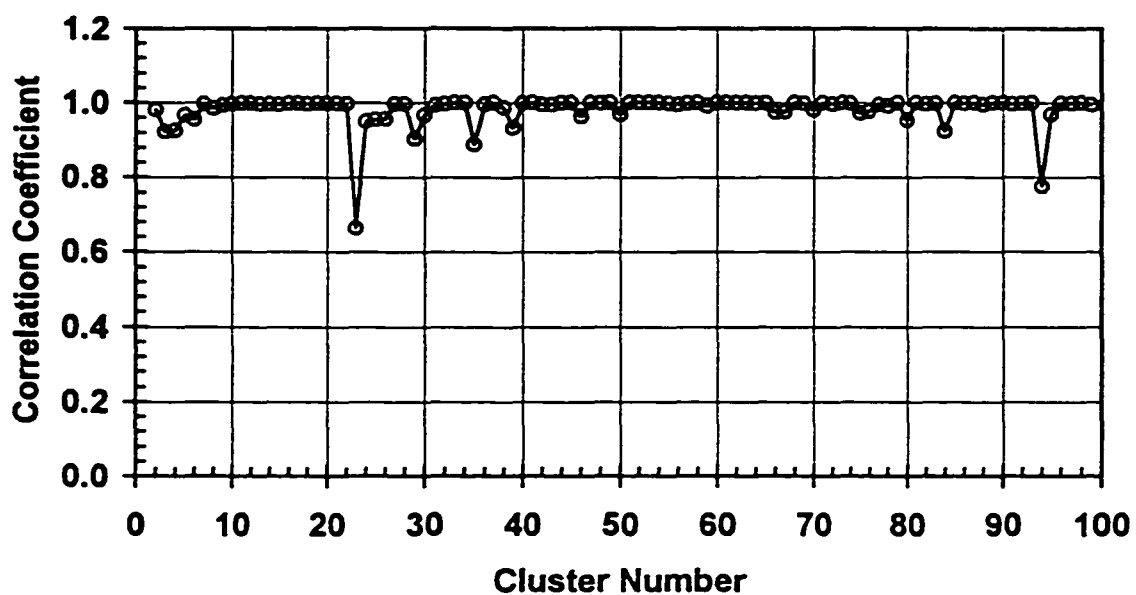
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Tiller, Norway from Sandven (1990).

Figure C.50a. Cluster Analysis of Piezocone Data at Tiller, Norway.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Tiller, Norway from Sandven (1990).

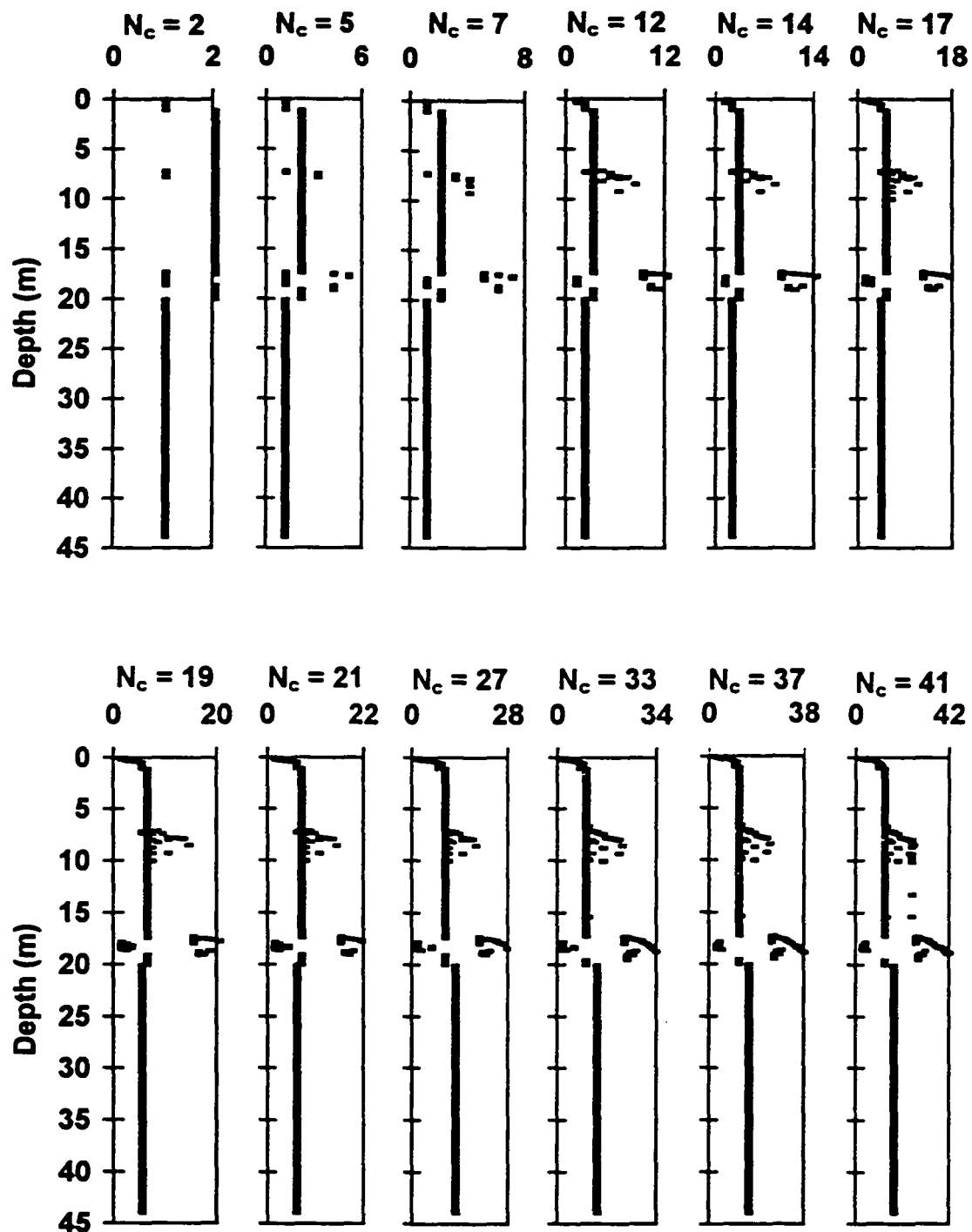
Figure C.50b. Cluster Analysis of Piezocone Data at Tiller, Norway.



Single-Cosine-Zscore Method Using Q and B_q

Note: Piezocone data at Troll, North Sea from Amundsen et al. (1985).

Figure C.51. Correlation Coefficient Between Consecutive Cluster Results at Troll, North Sea.



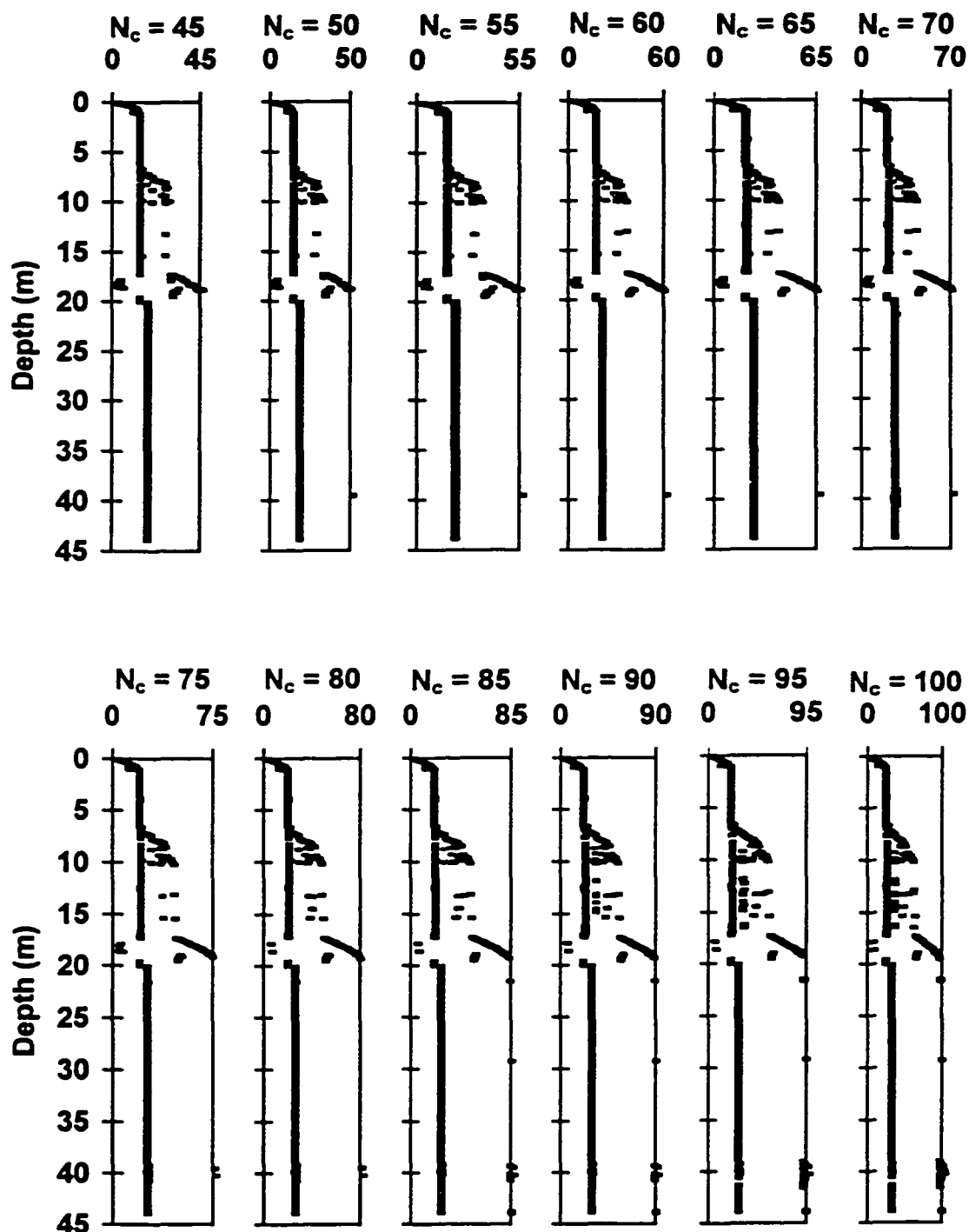
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Troll, North Sea from Amundsen et al. (1985).

Figure C.52a. Cluster Analysis of Piezocone Data at Troll, North Sea.

the two main statistical layers continue to separate indicating non homogeneity. However, no new primary layers ($t \geq 1$ m) are discovered up to cluster number 100 as shown in Fig. C.52b which presents clustering between $N_c = 45$ and 100.

C.17. Summary

Piezocone data of 12 sites were divided into correlated groups using single-cosine-zscore (SCZ) method. Clustering of the normalized parameters Q and B_q was studied between cluster number $N_c = 2$ and $N_c = 100$. Cluster results were checked at the peaks of the correlation coefficient between consecutive clusters. Then a certain cluster number was chosen to represent subsurface stratification at a site according to the interpretation criterion discussed in Chapter 4. At all sites, the soil profile was detected properly at $N_c < 15$. Therefore, it is recommended that the normalized piezocone parameters Q and B_q be analyzed up to 15 divisions instead of 100.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Troll, North Sea from Amundsen et al. (1985).

Figure C.52b. Cluster Analysis of Piezocone Data at Troll, North Sea.

APPENDIX D

CLUSTERING ASSESSMENT OF PIEZOCONE DATA AT DIFFERENT SITES

D.1. Synopsis

Cluster analysis using single-cosine-zscore method is applied to piezocone data at 10 sites summarized in Table D.1 to detect a subsurface stratigraphy. Normalized piezocone data Q and B_q are divided into correlated groups up to cluster number $N_c = 100$. A cluster is chosen to indicate a soil profile based on the criterion described in Chapter 4. Clustering is validated by comparing the obtained primary and secondary layers, and soil lenses and outliers with available laboratory and field tests.

Table D.1. Summary of 10 Sites Analyzed Using Cluster Methods.

Site Name	Location	Soil Type	Reference
Aiken	South Carolina	Alternative clay, silt and sand	Bratton et al. (1993)
South Boston	Massachusetts	Silty clay differs from stiff to soft with depth	Sweeney and Kraemer (1993)
Kagoshima	Japan	Sand to silty sand over silt underlain by sand	Takesue et al. (1996)
Newport News	Virginia	Soil mixture of sand, silt and clay underlain by sandy clay to clayey sand	Mayne (1989)
Onsoy	Norway	soft clay	Gillespie et al. (1985)
Opelika	Alabama	silty sand to sandy silt	(This study)
Po River	Italy	medium to coarse sand	Bruzzi et al. (1986)
Penuelas	Puerto Rico	clayey silt and some sand	Hegazy and Mayne (1996)
Maskinonge	Quebec	firm to soft silty clay	Demers et al. (1993)
Taranto	Italy	stiff to hard silty clay	Battaglio et al. (1986)

D.2. Aiken Test Site, South Carolina

The Savannah River Site in Aiken, South Carolina, is located in the Atlantic Coastal Plain and the soil deposits consist of marine and fluvial sediments of sandy, silty, and soils. The soils contain, from the top to the bottom, Hawthorne, Branwell, and McBean formations of tertiary age (Castro and Reeves, 1991). The Tuscaloosa formation exists between the McBean and the rock surface. The subsurface stratigraphy consists of alternate layers of sands, clayey sands and sandy clays with boundaries at approximate depths of 9 m, 12 m, 21 m, 27.5 m, 38.5 m and 45.5 m (Bratton et al., 1993).

A representative piezocone sounding at the site (unprocessed q_t and u_b) and derived normalized parameters Q and B_q are shown in Fig. D.1. Alternate layers of fine and coarse soils are detected looking at piezocone profiles. The boundaries are approximately at 9 m, 12 m, 22 m, 27 m, 30 m, 33 m, 38m, 44.5 m, 47 m, 49 m. The soils were classified using the Robertson's chart (1991) as shown in Fig. D.2. There is a good agreement between the subsurface stratigraphy obtained from the chart and that proposed by Bratton et al. (1993) based on sampling, laboratory results and in-situ testing.

A single-cosine-zscore clustering method was applied to Q and B_q parameters and the correlation coefficient between consecutive clusters (ρ_c) is shown in Fig. D.3. Cluster results are examined at the peaks of ρ_c and shown in Fig. D.4a between cluster number $N_c = 2$ and 36. At cluster number 2, the data were grouped in one layer except one point at 5.9 m. At $N_c = 5$, the data are separated into two major soil types indicated by cluster numbers 1 and 3. Then points declustered from the primary groups representing transition zones between different layers and non-homogeneity within the same layer. At $N_c = 9$, the

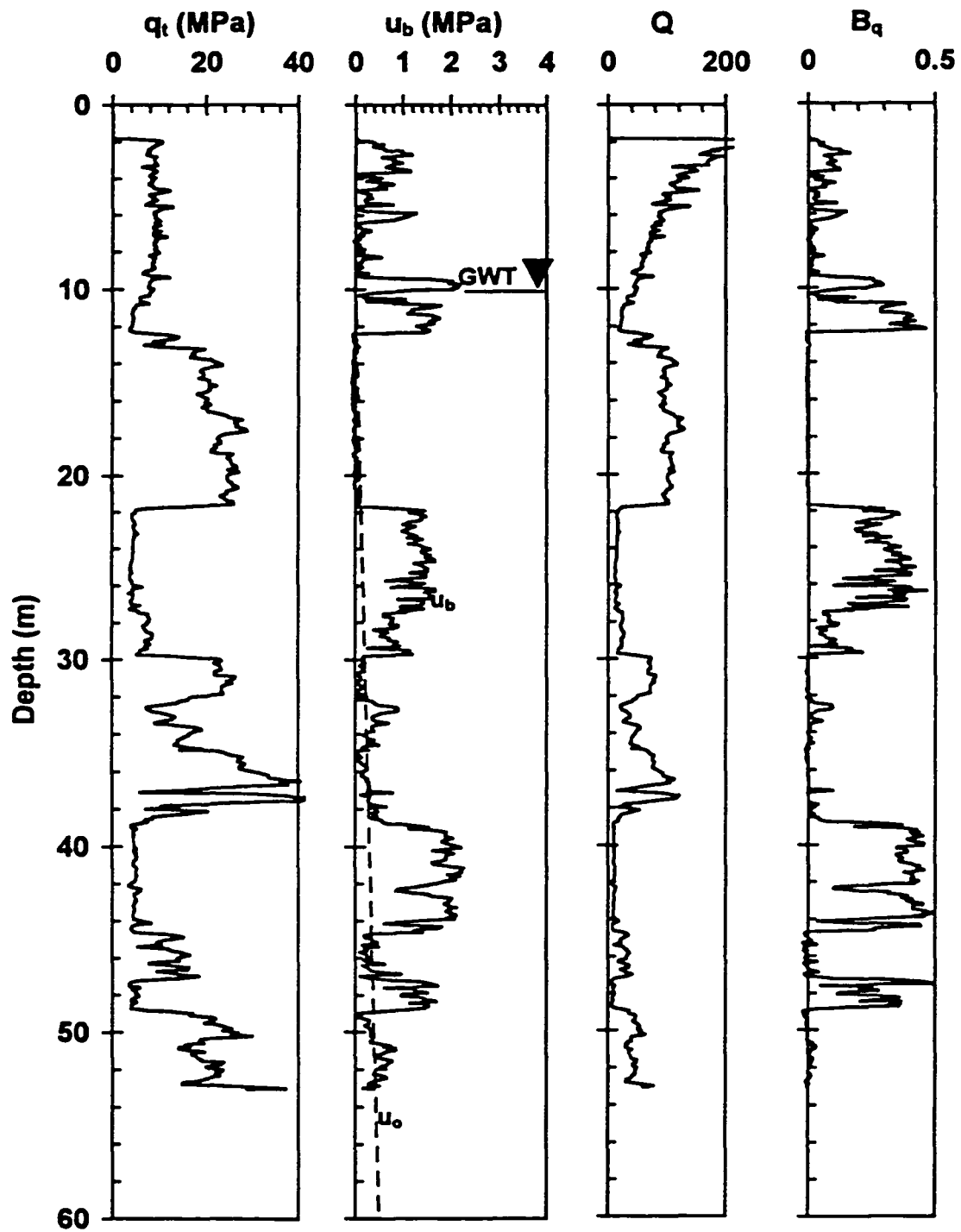
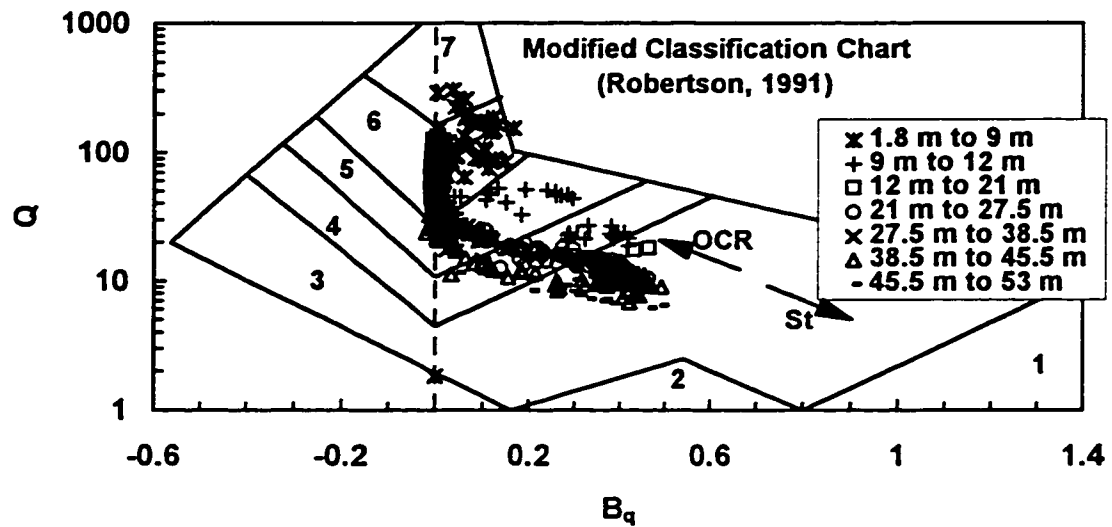
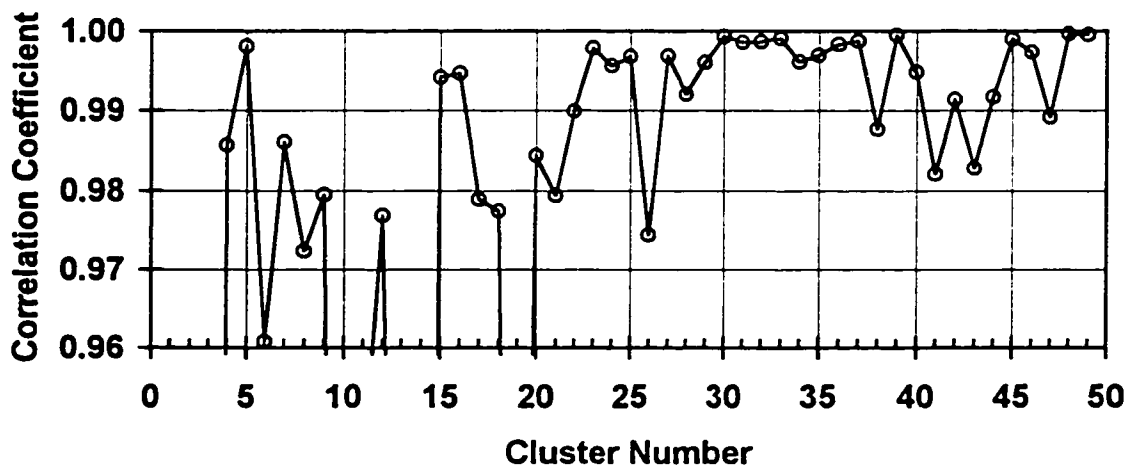


Figure D.1. Piezocone Data at Aiken, South Carolina (Data from Bratton et al., 1993).



- | | |
|---|--|
| 1. Sensitive Fine Grained | 5. Sand Mixtures: Silty Sand to Sandy Silt |
| 2. Organic Soils-Peats | 6. Sands: Clean Sand to Silty Sand |
| 3. Clays: Clay to Silty Clay | 7. Gravelly Sand to Clayey Sand |
| 4. Silt Mixtures: Clayey Silt to Silty Clay | |

Figure D.2. Soil Classification Using Piezocone Data at Aiken, South Carolina (Data from Bratton et al., 1993).



Single-Cosine-Zscore Method Using Q and B_q

Note: Piezocone data at Aiken, South Carolina (Bratton et al., 1993).

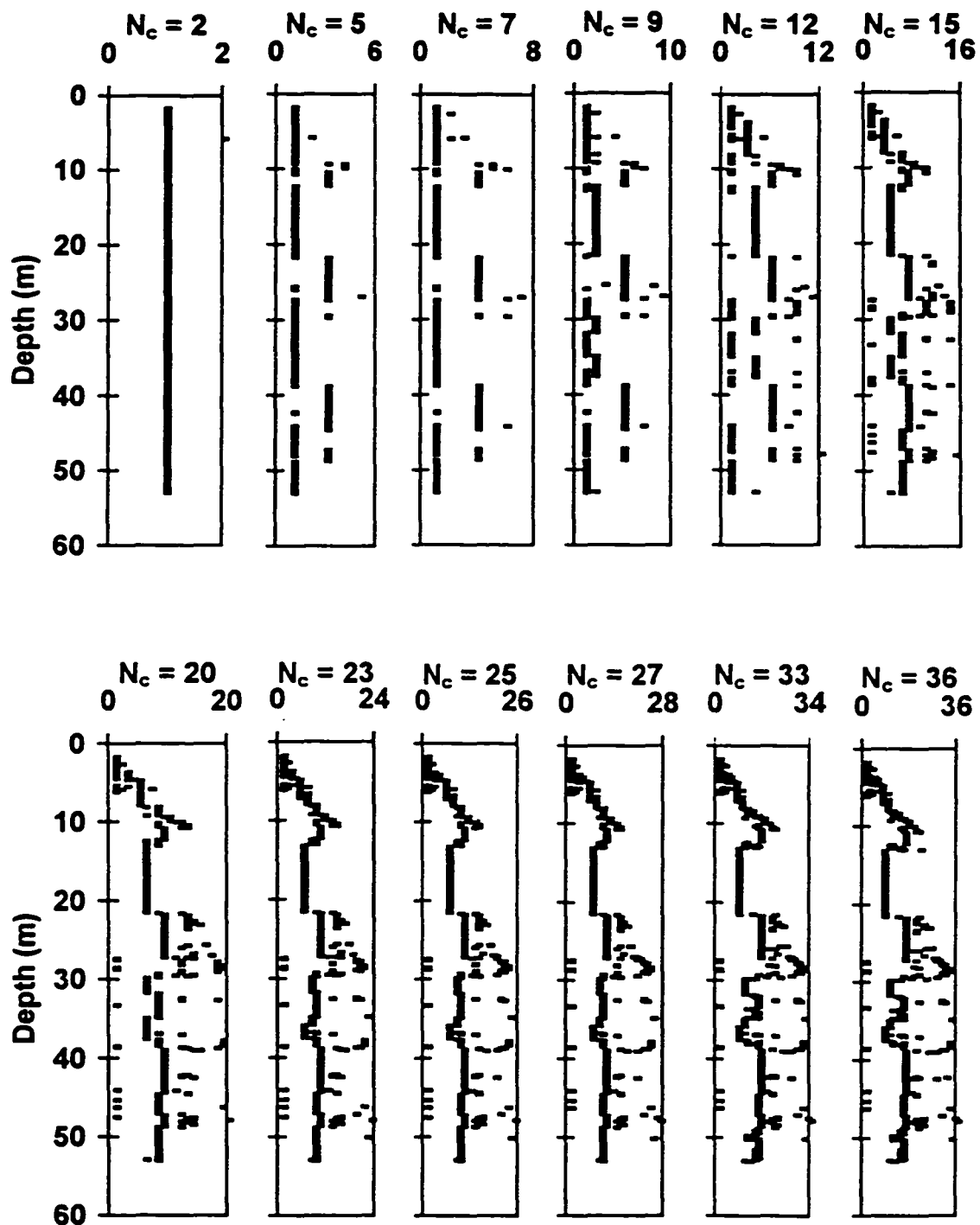
Figure D.3. Correlation Coefficient Between Consecutive Cluster Results at Aiken, South Carolina.

upper cluster between 1.8 m and 21 m is divided into two groups and the cluster between 27.7 m and 39 m is divided into 5 groups. For larger cluster numbers, no new primary layers (thickness > 1m) appears up to cluster number $N_c = 100$ as shown in Figures D.4a and D.4b. Therefore cluster number 9 is chosen to represent the soil profile at the site. The major statistical layers ($t > 1$ m) denoted by clusters numbers 1 and 2 are associated together in terms of soil type and/or properties. For instance, the clusters between 1.8 m and 9.2 m ($N_c = 1$), and between 13.2 m and 21.6 m ($N_c = 2$) are both classified as silty sand (Bratton et al., 1993). However, the statistical layer between 21.9 m and 27 m ($N_c = 5$) is less associated with the upper two layers and classified as silty clay.

Clustering was successful in detecting the subsurface stratigraphy as shown in Fig. D.5. Soil boundaries and the relation between different soil layers are in agreement with those defined based on laboratory and field tests. The soil layer, for instance, between 27.5 m and 38.5 m was divided by clustering into 5 associated groups which were denoted by alternating cluster numbers, 1 and 2. This is an indication of heterogeneity or inconsistency of soils type and/or properties within this layer which is supported by the variation of the soil description from silty sand to clean sand.

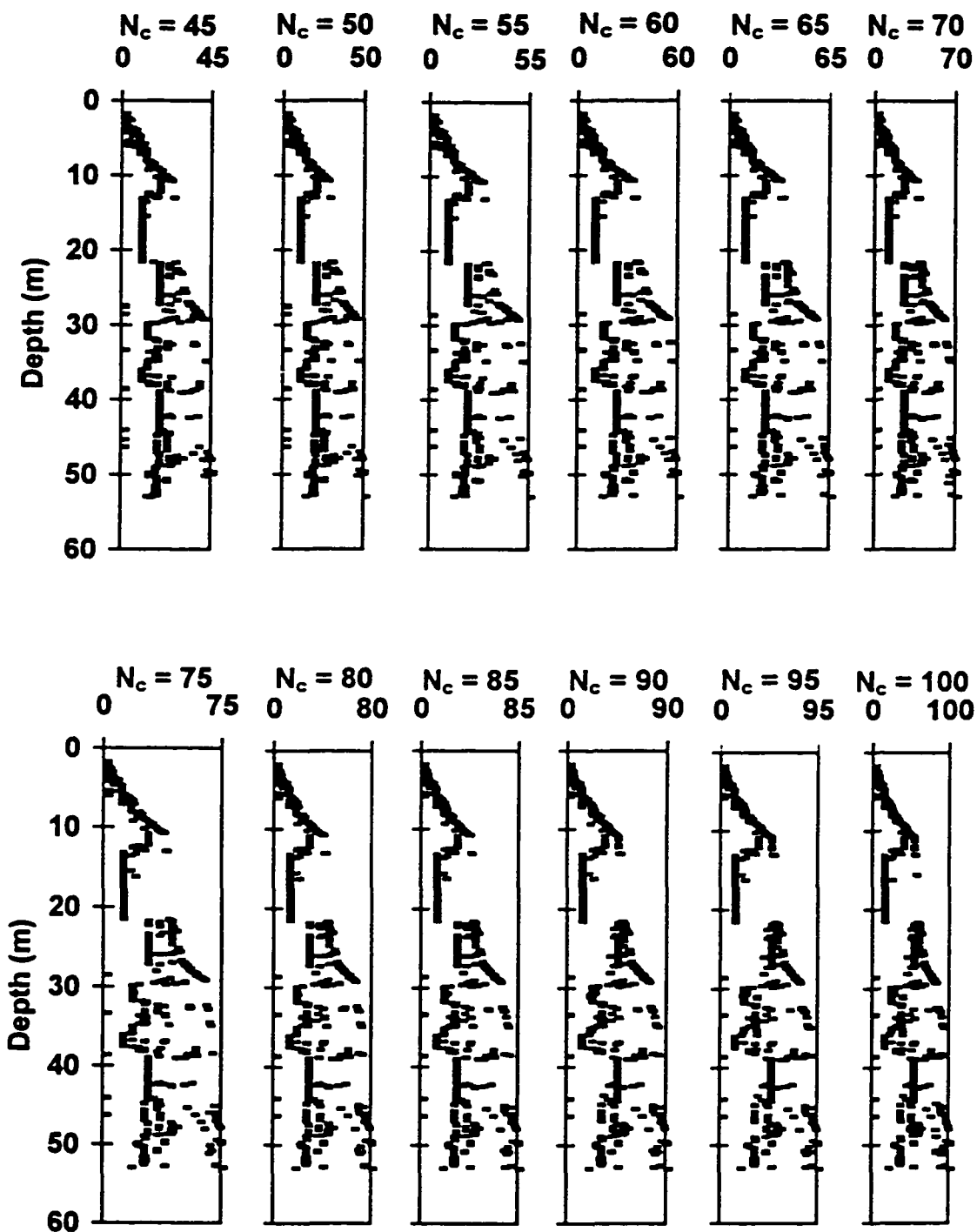
D.3. South Boston Test Site, Massachusetts

Extensive laboratory and in-situ geotechnical investigations were performed for a design of a tunnel at South Boston, Massachusetts (Sweeney and Kraemer, 1993). The geological features at the site consists of the following, from the ground surface downward: cohesive fill, marine deposit (sand over clay), glaciomarine and glacial till



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Aiken, South Carolina from Bratton et al. (1993).

Figure D.4a. Cluster Analysis of Piezocone Data at Aiken, South Carolina.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Aiken, South Carolina from Bratton et al. (1993).

Figure D.4b. Cluster Analysis of Piezocone Data at Aiken, South Carolina.

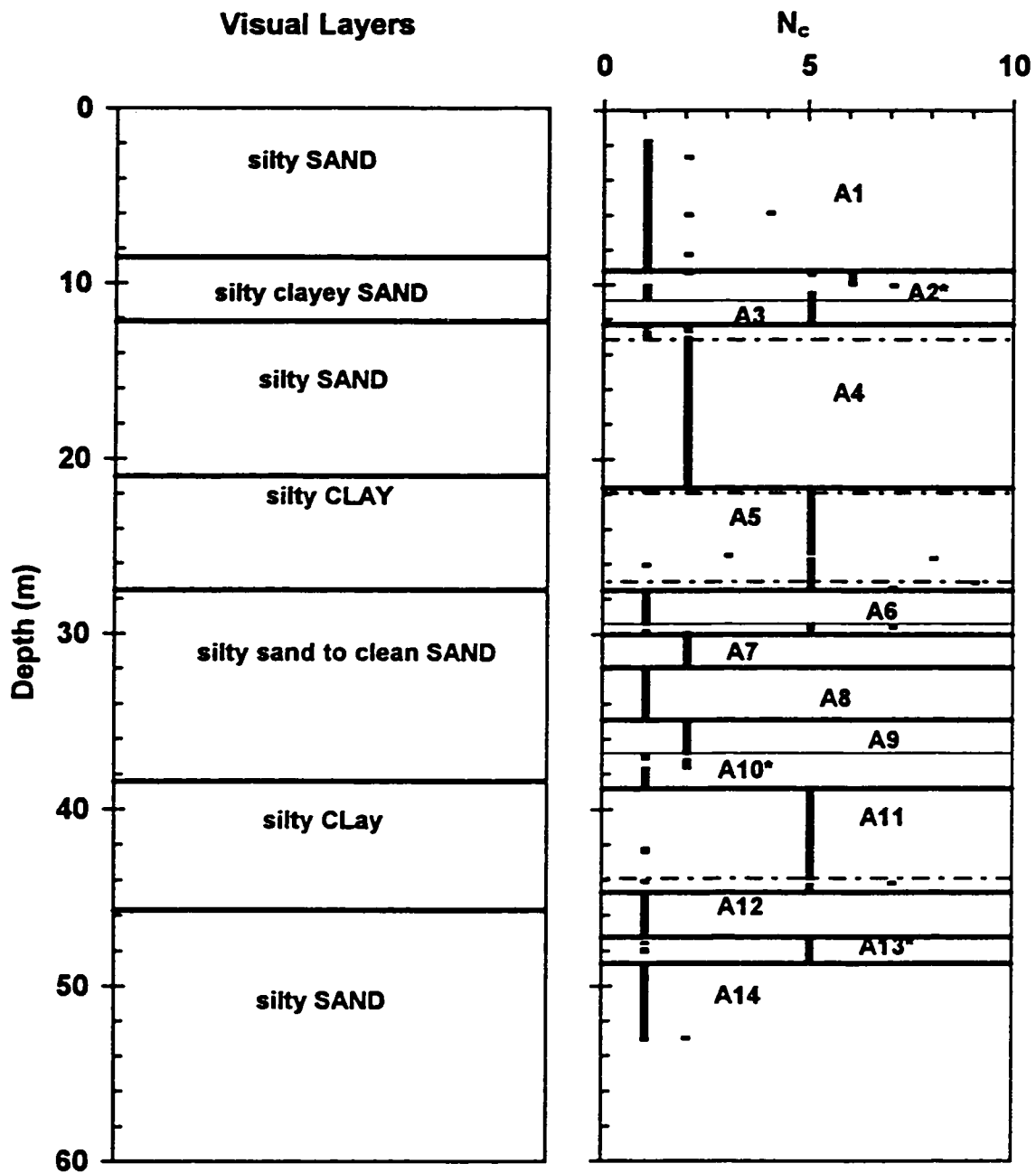


Figure D.5. Comparison Between Cluster Analysis and Visual Classification at Aiken, South Carolina.

deposits. The piezocone data were recorded in the clay marine deposit in which the consistency changes from stiff to soft with depth. The ground water table is at a depth of 3 m. A representative piezocone sounding and derived normalized parameters Q and B_q are shown in Fig. D.6. Two primary layers separated at almost 28 m are detected by looking at the piezocone profiles. The Robertson (1991) classification chart also indicated two soil types delineated at a depth of 24 m as shown in Fig. D.7. The upper layer was denoted number 3-type (clay to silty clay) and the lower layer was described as number 1-type (sensitive fine grained).

Clustering was applied to the normalized parameters Q and B_q using single-cosine-zscore method to discover correlated groups of the data. The correlation coefficient is calculated between consecutive clusters up to cluster number $N_c = 50$ as shown in Fig. D.8. Clustering was examined at the peaks of ρ_c and Fig. D.9a shows the cluster results between $N_c = 2$ and 43. The piezocone data were divided into two main groups starting at $N_c = 4$ after which no new primary layers ($t > 1$ m) separated. The two primary layers are denoted by cluster numbers 1 and 2. Transition layers with mixed cluster numbers from 1 to 4 is detected between depths of 26.15 m to 27.35 m. However some data points continue to separate indicating no association with the two main layers or transition zone. The same note is applied for clusters from $N_c = 45$ up to $N_c = 100$ as shown in Fig. D.9b. Therefore, cluster number 4 is chosen to represent the soil profile.

Clustering is supported by the primary layers defined based on laboratory and field testing and the overconsolidation ratio (Sweeney and Kraemer, 1993) as shown in Fig. D.10. The overconsolidation ratio decreases from 8 at a depth of 7.2 m to 1 at a depth of

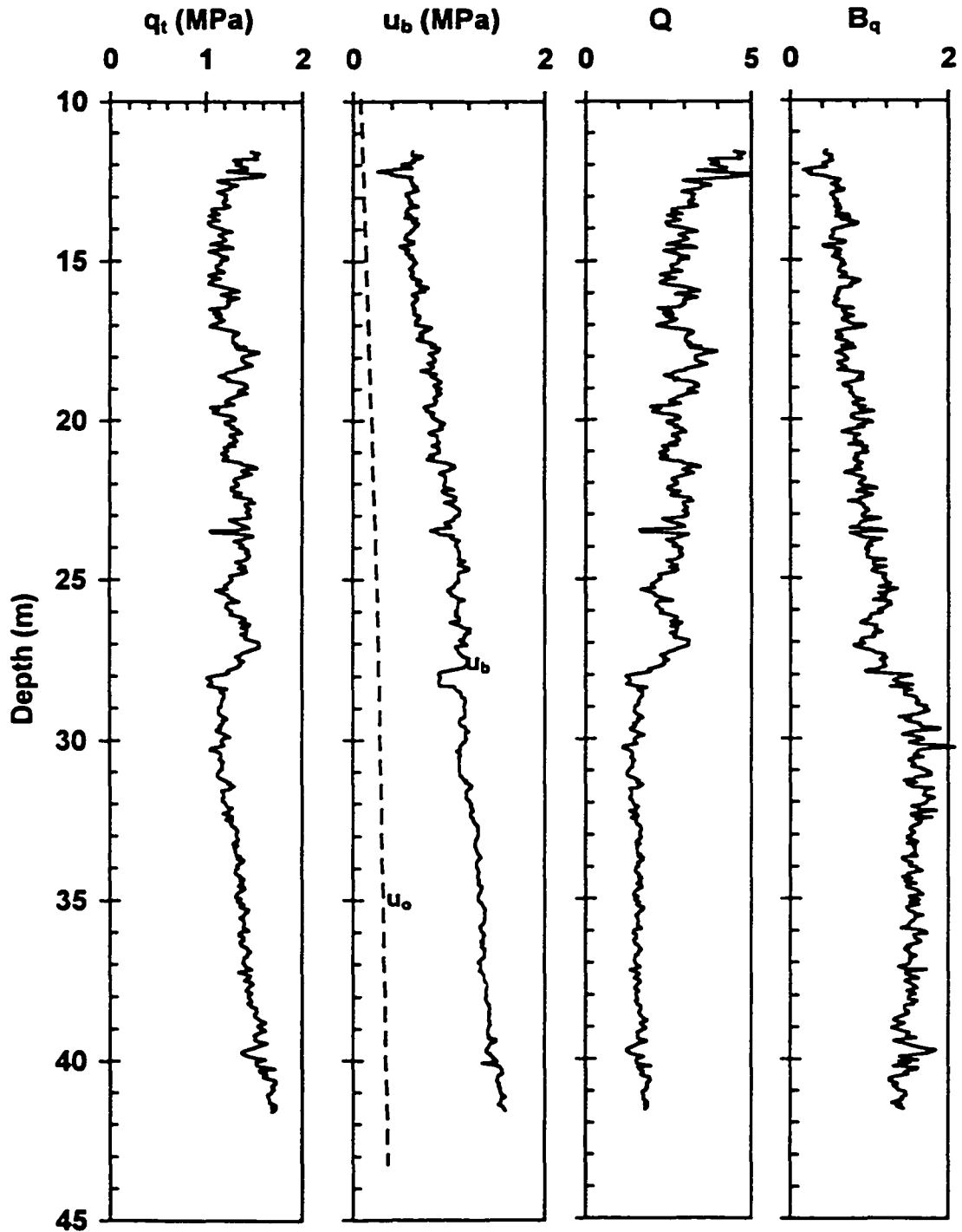


Figure D.6. Piezocone Sounding at Boston, Massachusetts (Data from Sweeney, B. P. and Kraemer, S. R., 1993).

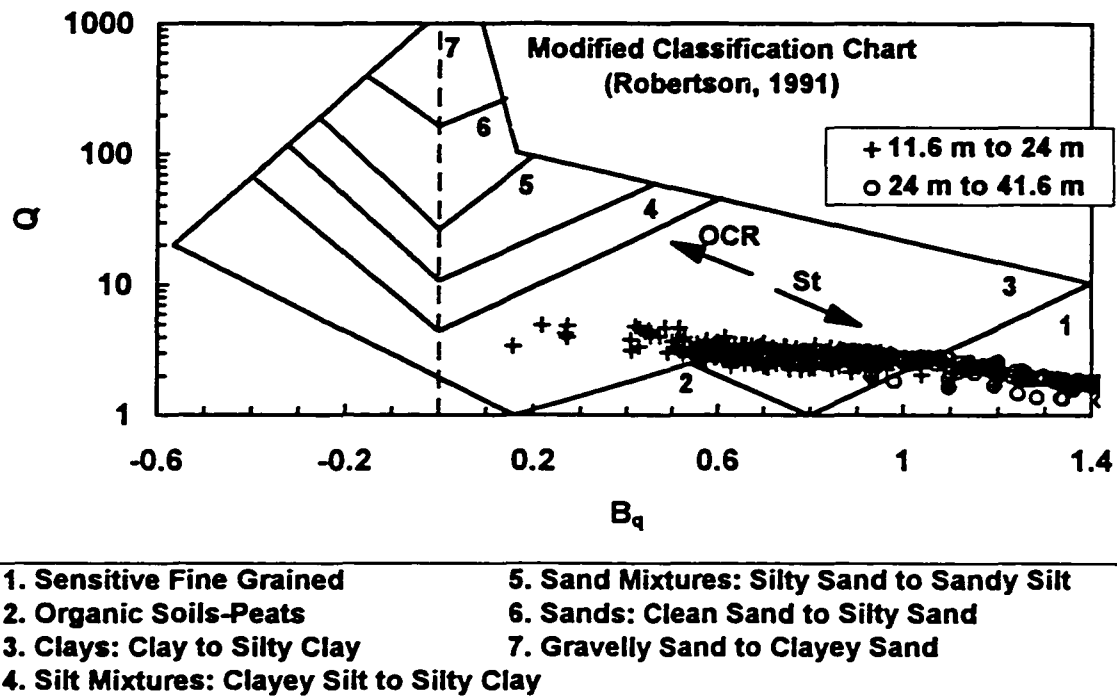


Figure D.7. Soil Classification Using Piezocone Data at Boston, Massachusetts (Data from Sweeney, B. P. and Kraemer, S. R., 1993).

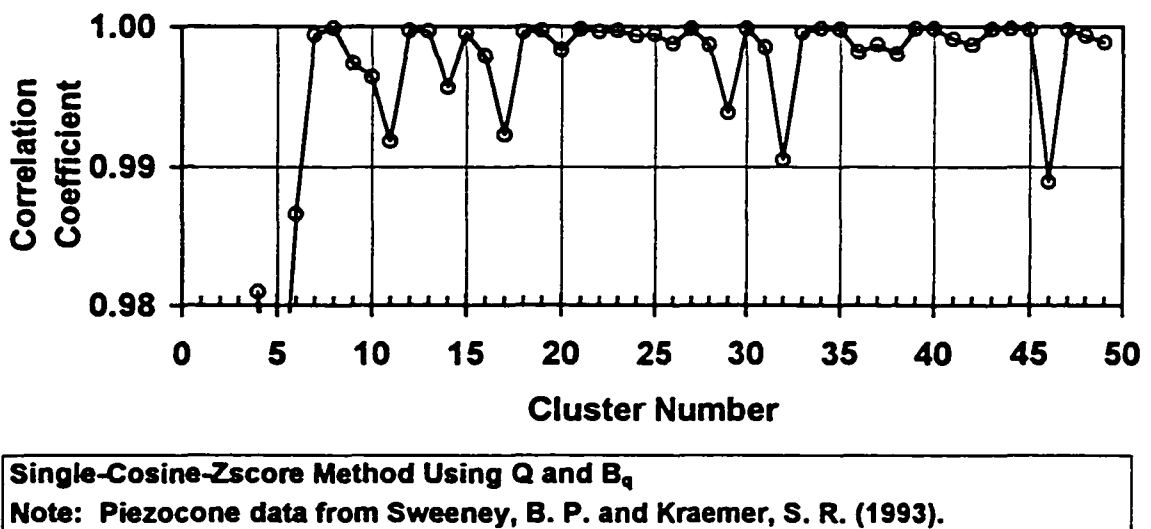
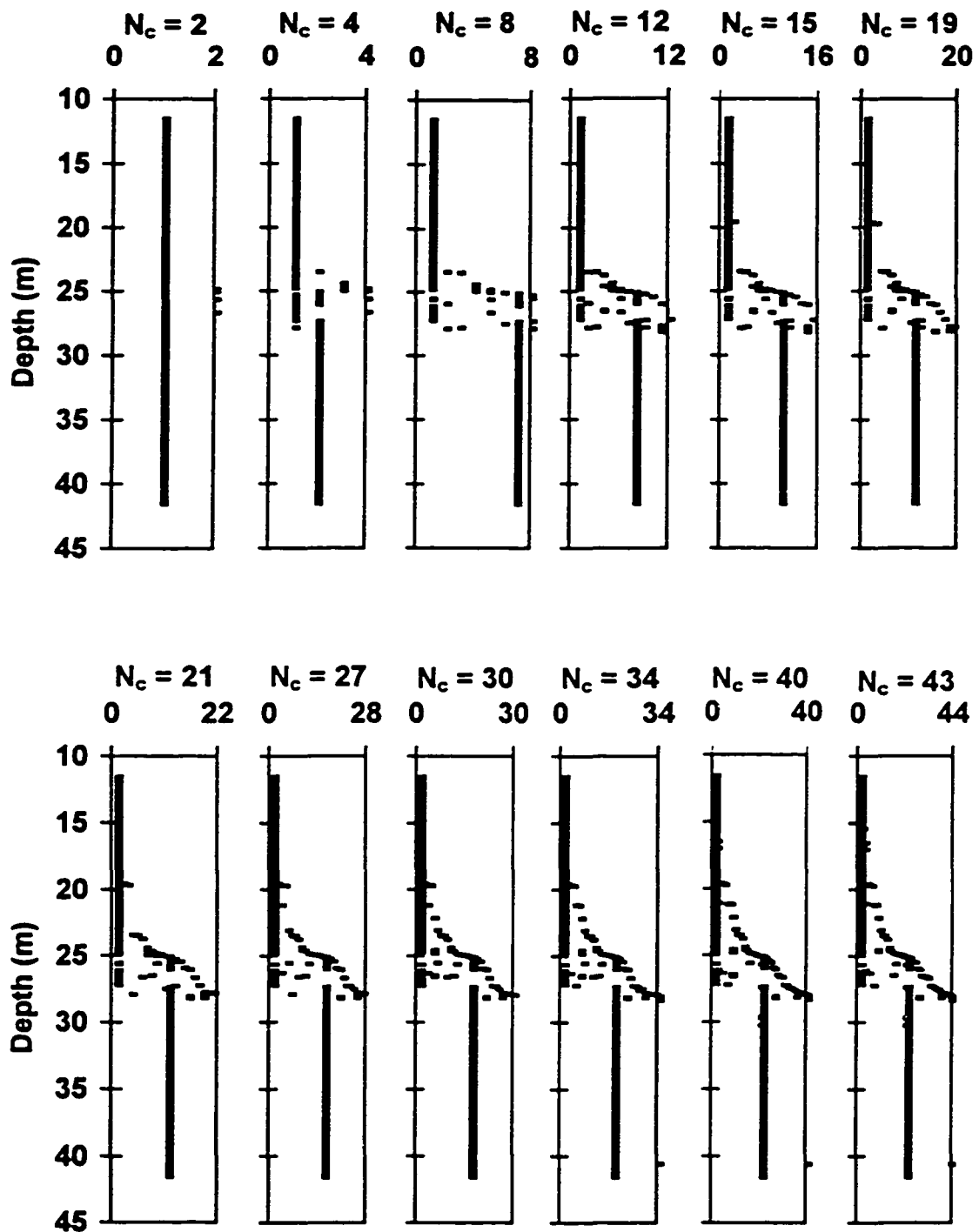
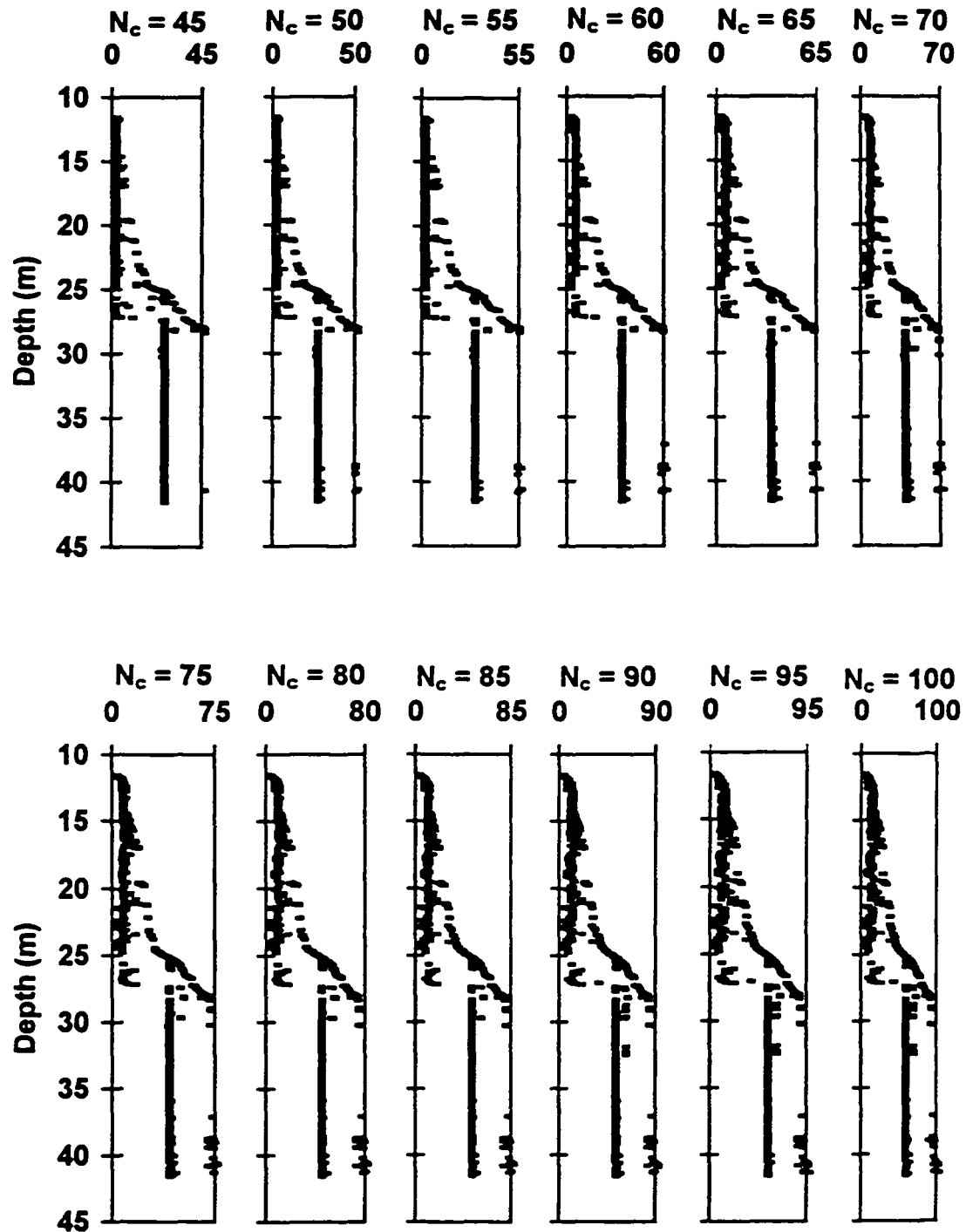


Figure D.8. Correlation Coefficient Between Consecutive Cluster Results at Boston, Massachusetts.



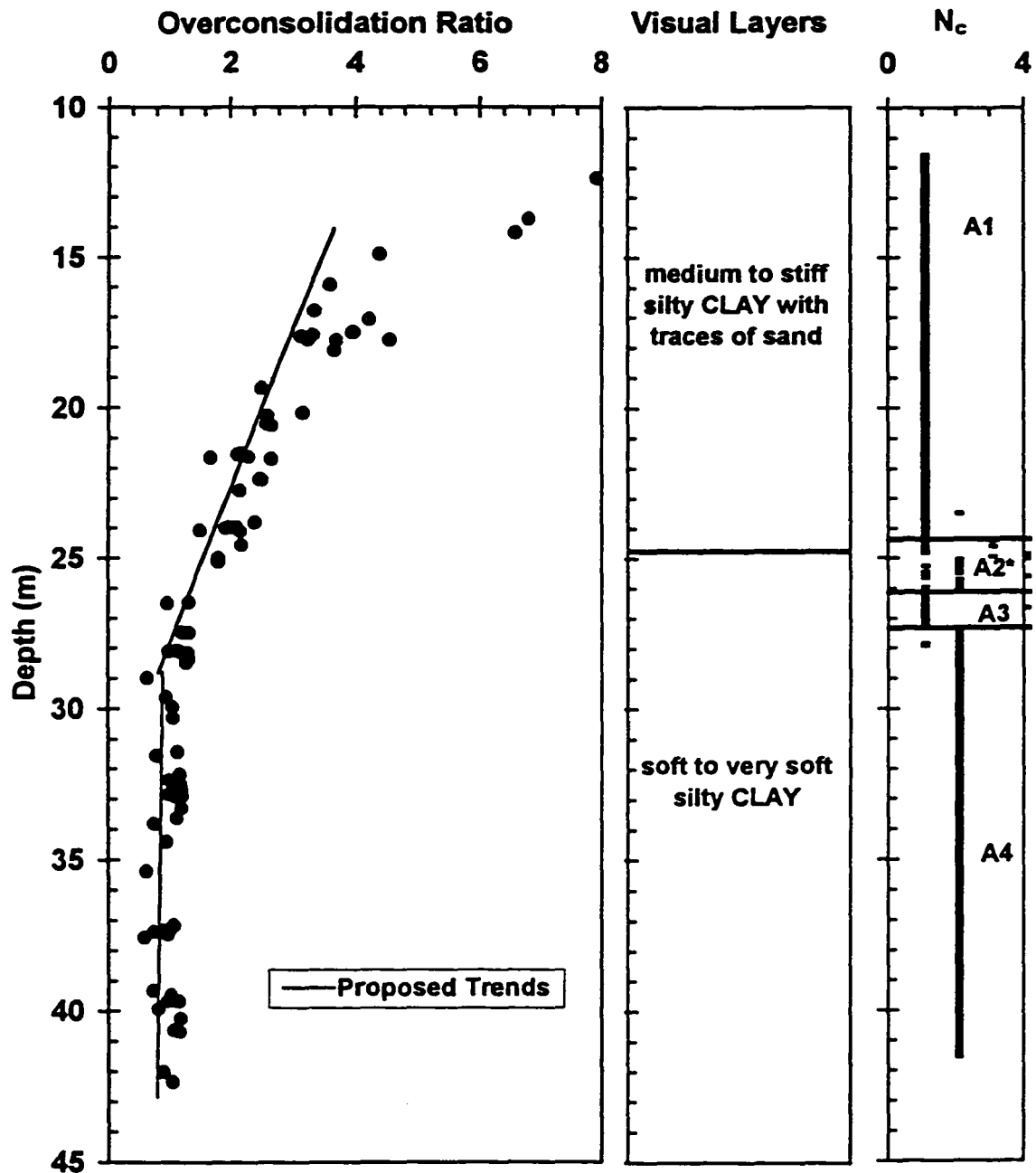
Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data from Sweeney, B. P. and Kraemer, S. R. (1993).

Figure D.9a. Cluster Analysis of Piezocone Data at Boston, Massachusetts.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data from Sweeney, B. P. and Kraemer, S. R. (1993).

Figure D.9b. Cluster Analysis of Piezocone Data at Boston, Massachusetts.



Note: Piezocone data from Sweeney, B. P. and Kraemer, S. R. (1993).
 SCZ method using Q and B_q , cluster no. 4

Figure D.10. Comparison Between Cluster Analysis, Visual Classification and Overconsolidation ratio at Boston, Massachusetts.

26.5 m. Then, it has an average equal to 1 down to a depth of 42.5 m, indicating normally consolidated clay.

D.4. Kagoshima Test Site

The Kagoshima site is located in southern Kyushu, Japan. A geotechnical study was performed by Takesue et al. (1996) at the site to evaluate the validity of traditional soil testing including the piezocone test in volcanic soils which are denoted “Shirasu” in Japan. The Shirasu consists primarily of vesiculate and volcanic glass grains which are easily crushable and the range of their specific gravity is lower than that of quartz sands. The mechanical properties of the Shirasu are affected by its unique physical properties.

The soil stratigraphy at the site consists of the following: 7 m of fill, 41 m of alluvium which is divided into three layers denoted as clean quartz sands, silty sand (Shirasu) and silt (Shirasu) with boundaries at 25 m and 36 m and 17 m of diluvium [pumice and sand (Shirasu)]. Piezocone data and their normalization parameters Q and B_q are shown in Fig. D.11. Looking at the data, three major layers are defined with boundaries at approximately 36 m, 48 m. Two soil lenses are also detected near depths of 35 m and 60 m. Therefore, visual inspection of the data does not give a complete proper soil stratification at the site.

Based on the CPT chart (see Fig. D.12), the upper 32 m of soil is defined as clean sand to silty sand and a clay to silty clay layer is detected between 37 m and 47 m. Transitions layers are given including a soil mixture of sand and silt between 32 and 37 m, and 47 m and 49 m. A soil layer of 16 m of sand with some silt is defined below 49 m. A

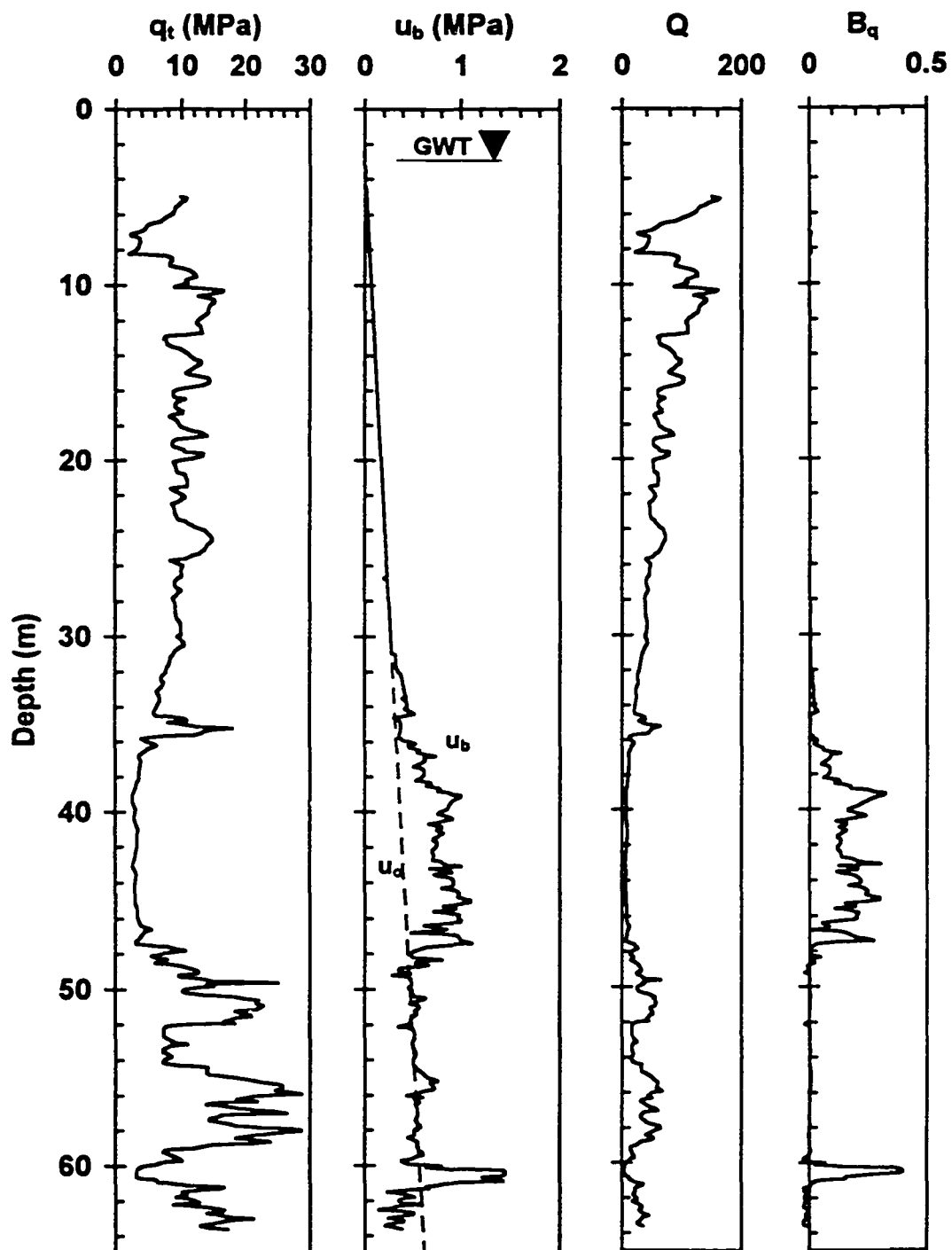


Figure D.11. Piezocone Data at Kagoshima, Japan from Takesue et al. (1995).

silt soil lense and a clay to silty clay soil lense are discovered near 7 m and 60 m, respectively. Therefore the method is able to detect the primary layers at the site except the boundary at 25 m and also a class number 6 is given for both of the upper clean sand and the Shirasu (volcanic) pumice and sand.

A single-cosine-zscore cluster analysis is applied to the normalized parameters Q and B_q . The correlation coefficient between consecutive clusters is shown in Fig. D.13 up to cluster number $N_c = 100$. Clustering is examined at the peaks of ρ_c and shown in Fig. D.14a between $N_c = 2$ and 28. At N_c equal to 2, Two soil types and three data groups are defined with boundaries at 36.4 m and 47.5 m. At cluster number 5, two transition zones appear as follows: (1) between the depths of 36 m and 38.2 m and (2) between the depths of 46.4 m and 48.3 m. For larger cluster numbers, primary boundaries, soil lenses and more transitions break off the three major clusters. At cluster number 13, the upper cluster is divided into 3 major groups denoted by cluster numbers 1, 2 and 3 with boundaries at 15.8 m, 25.6 m and 35.9 m. The upper group between the depths of 5 m and 15.8 m has an intermediate primary layer denoted by cluster number 2 between the depths of 7 m and 8.2 m. For higher clusters, no new primary layers are discovered up to $N_c = 100$ as shown in Fig. D.14b which includes cluster numbers between $N_c = 45$ and 100 every 5 increments. Therefore cluster number 13 is chosen to represent the subsurface stratigraphy.

The cluster results are verified by both the soil contents and the mean particle size (D_{50}) as shown in Fig. D.15. The boundary between layers A1 and A2 is matching with the boundary between the fill and the clean sand layers and indicates also the end of a

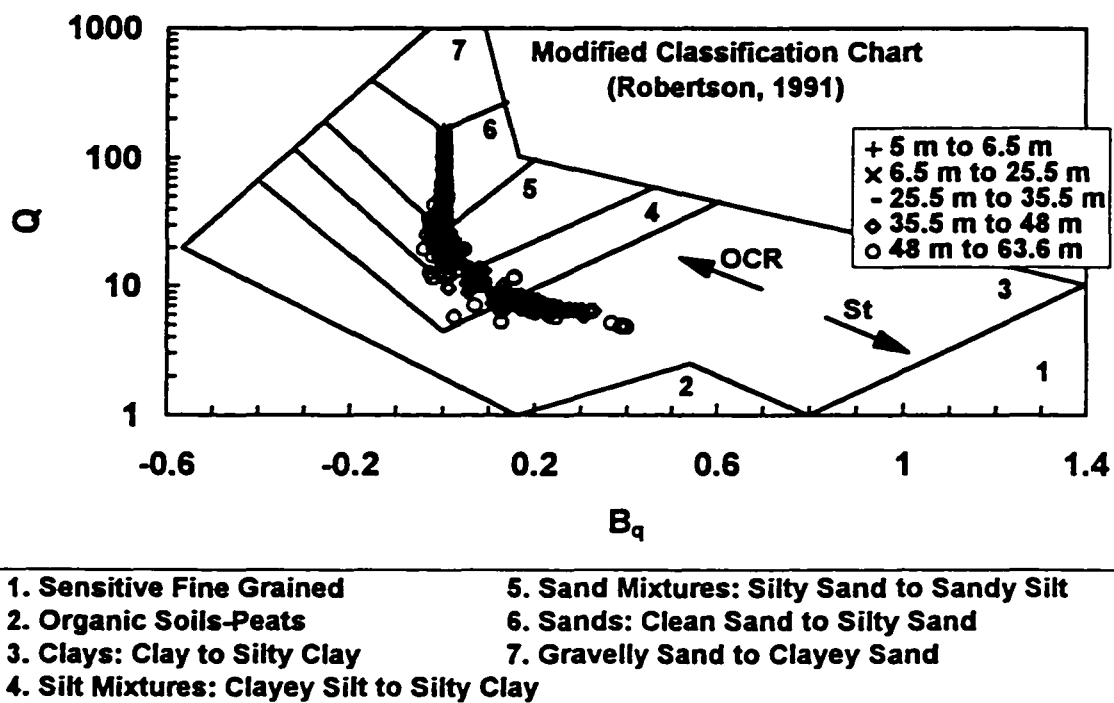
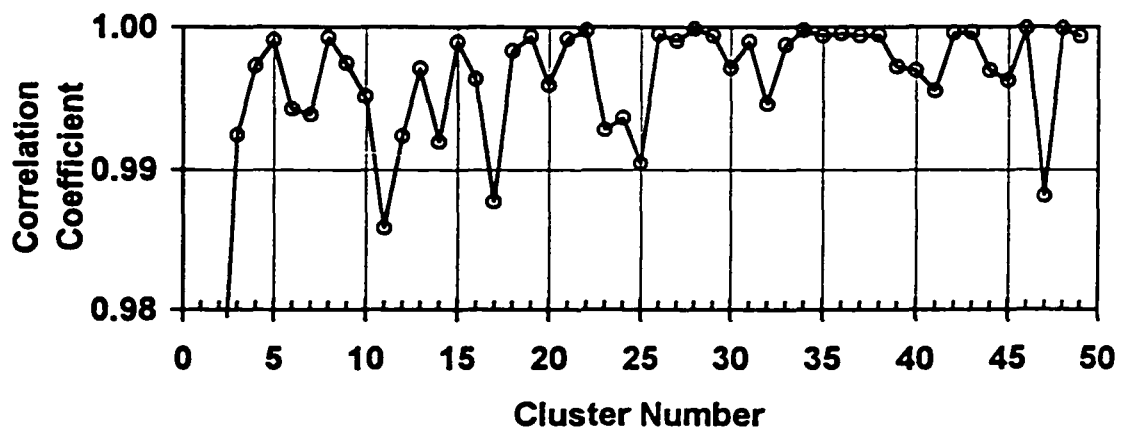


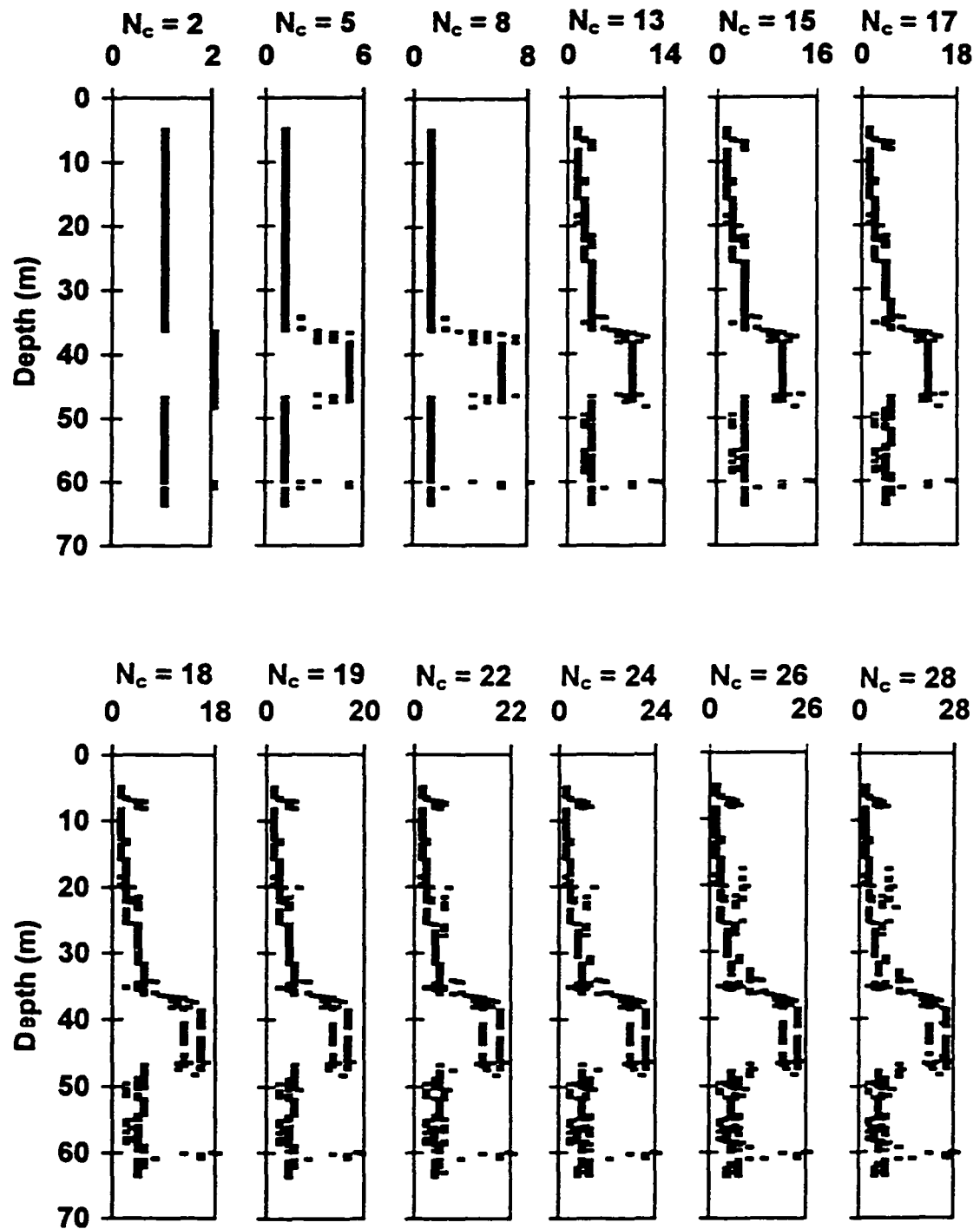
Figure D.12. Soil Classification Using Piezocone Data at Kagoshima, Japan (Takesue et al., 1995).



Single-Cosine-Zscore Method Using Q and B_q

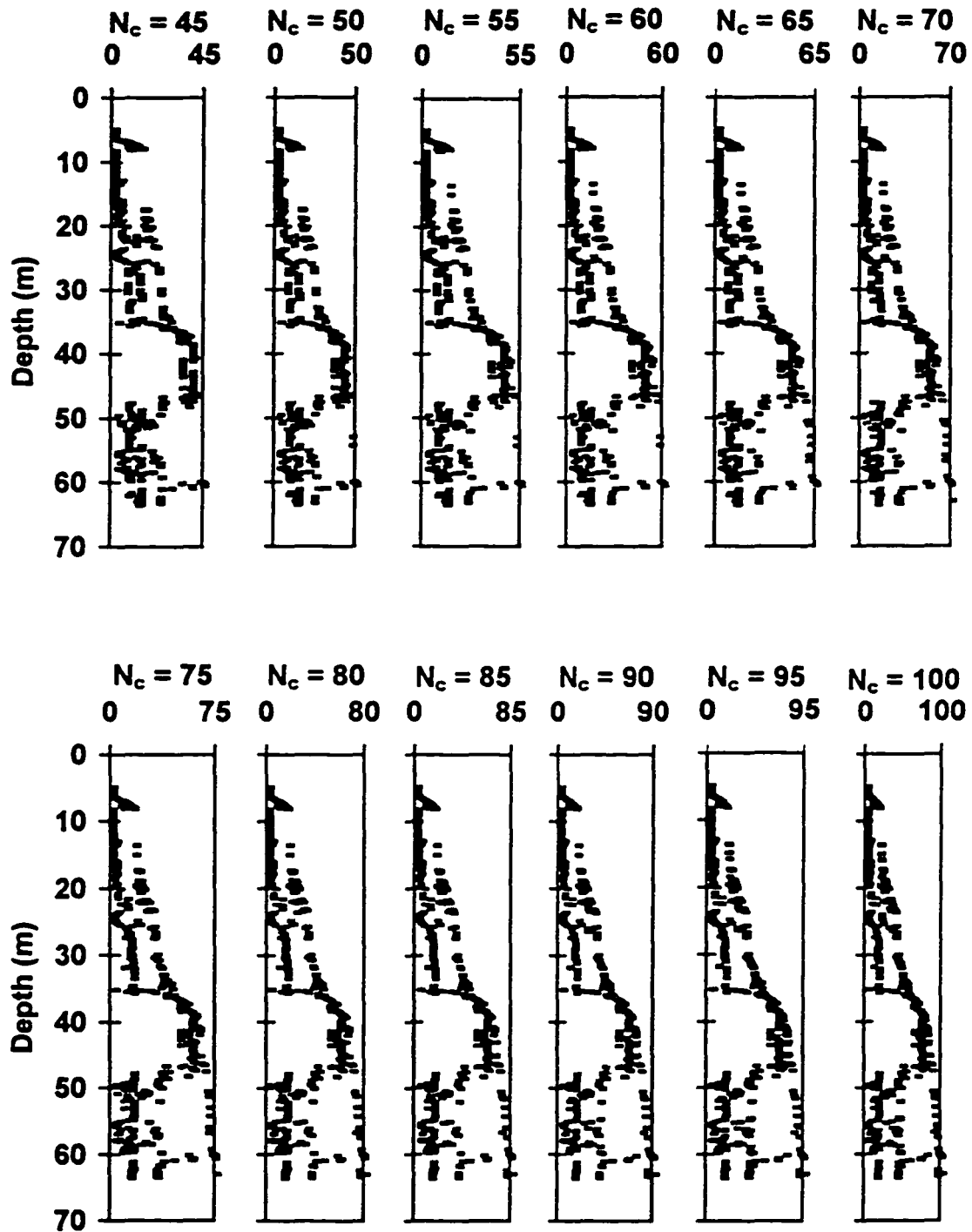
Note: Piezocone data at Kagoshima, Japan (Takesue et al., 1995).

Figure D.13. Correlation Coefficient Between Consecutive Cluster Results at Kagoshima, Japan.



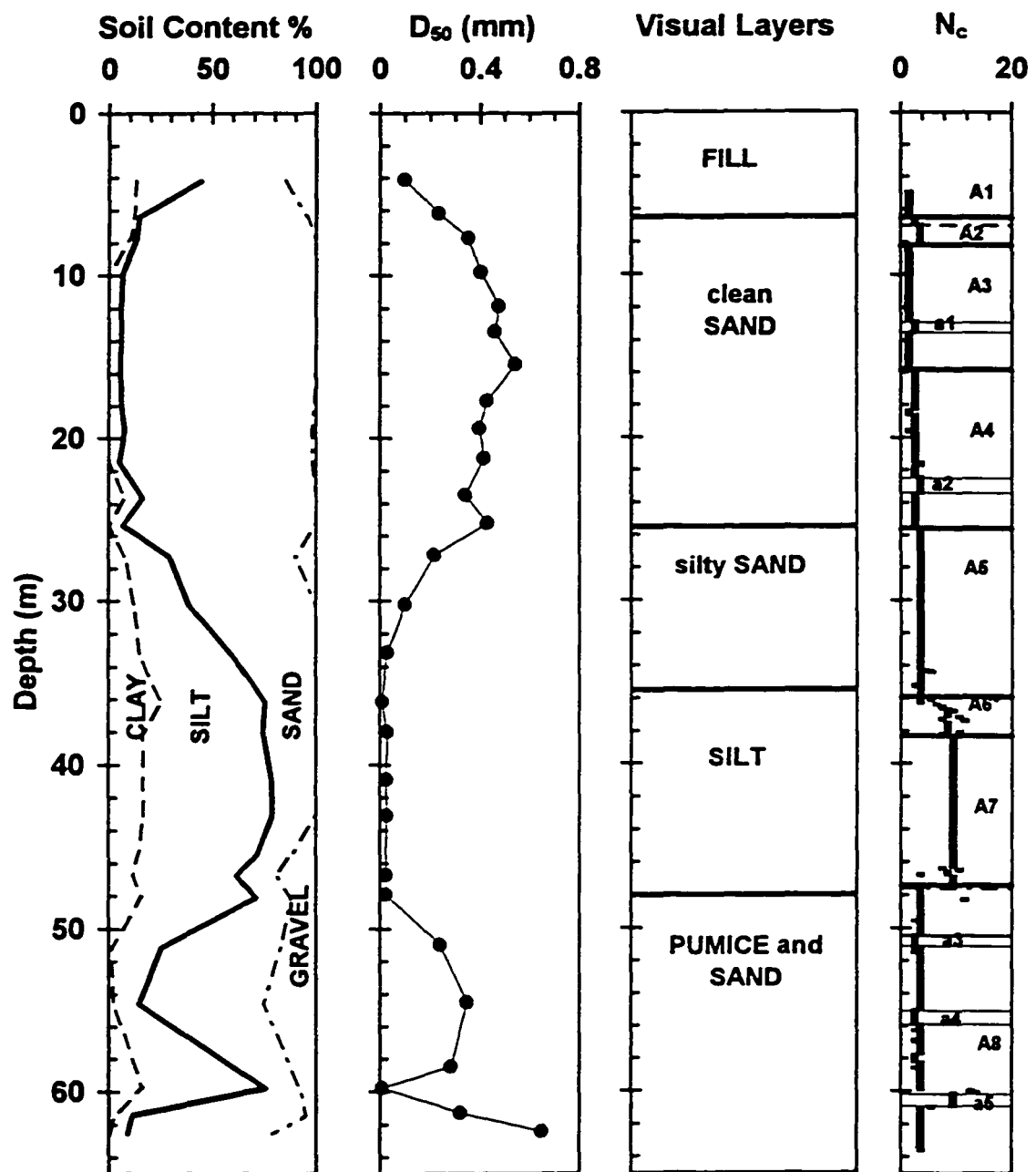
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Kagoshima, Japan from Takesue et al. (1995).

Figure D.14a. Cluster Analysis of Piezocone Data at Kagoshima, Japan.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Kagoshima, Japan (Takesue et al., 1995).

Figure D.14b. Cluster Analysis of Piezocone Data at Kagoshima, Japan.



Note: Piezocone data at Kagoshima, Japan from Takesue et al. (1995).
 SCZ method using Q and B_q , cluster no. 13

— Primary boundary
 - - - Transition boundary

Figure D.15. Comparison Between Cluster Analysis, Visual Classification
 Soil Contents and D_{50} at Kagoshima, Japan.

steep negative trend of the fine contents (FC) from 47 percent to 15 percent. Then along layer A2, the FC remains almost constant with average equal to 6 percent. However, D_{50} increases from 0.4 mm at 9.8 m to 0.54 mm at 15.4 m.

Layer A4 is given cluster number 2 which indicates different soil type or property than layer A3. This is supported by a negative trend of D_{50} which decreases from 0.54 mm at 15.4 m to 0.43 mm at 25.2 m. Layer a2 is also supported by the sudden decrease of D_{50} from 0.41 mm at 21.2 m to 0.34 mm at 23.5 m. Layers A5 is supported by the positive and negative trends of FC and D_{50} , respectively. Layer A6* suggests a soil mixture.

Layer A7 indicates fine grained soils because the fines content is greater than 50 percent and D_{50} decreases to its minimum average equal to 0.02 mm. Layer A8 has a relatively more inherent variability than other layers. Three secondary layers and soil lenses are detected within this layer and supported by peaks and troughs of fines content and/or D_{50} vertical profiles. For example, the secondary layer a6 is verified by the decrease of D_{50} from 0.28 mm at 58.5 m to 0.01 mm at 59.5 m and then back to 0.32 mm at 61.5 m.

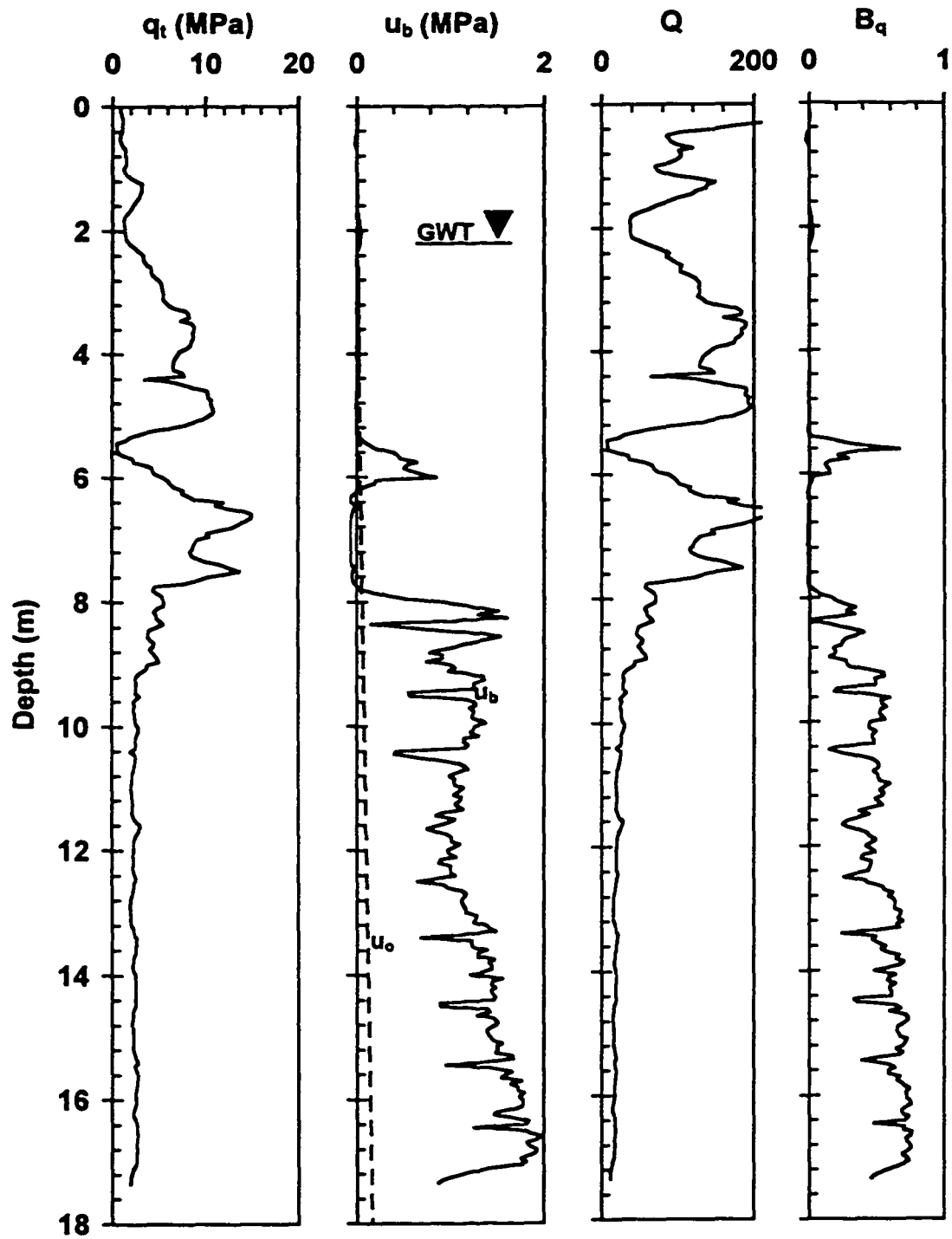
D.5. Newport News Test Site, Virginia

The geological setting at the Newport News site consists of Norfolk formation deposited over Yorktown formation (Mayne, 1989). The Norfolk formation was deposited in the late Pleistocene age due to a rise in the sea level. It consists of a heterogeneous mixture of sands, silts and clays. Lenses of organic-rich clay also are found in this formation. The Yorktown formation was deposited in a warm shallow continental

shelf environment during the Miocene and/or the Pliocene ages. The deposit includes very fine sandy clay to very clayey sand. The groundwater table is located at a depth of 2.1 m.

A representative piezocone sounding and normalized derived parameters are shown in Fig. D.16. Looking at the piezocone profiles, two primary layers are detected and their boundary is at 8 m. A soil lense is also seen in the upper layer at a depth of 5.5 m. The frequent dissipation of the penetration pore pressure indicates the locations where cone-rods were added at approximately every 1-m interval. The subsurface profile is also estimated successfully using the Robertson chart (1991) as shown in Fig. D.17. Three layers were detected as follows: (1) a clean sand to silty sand layer between the depths of 0 m to 8 m with a clay to silty clay lenses located between 5.4 m to 5.7 m, (2) a transition layer described as silty sand to sandy silt mixture between the depths of 8 m and 9.5 m and (3) a clayey silt to silty clay layer down to a depth of 17.4 m.

Clustering is applied to the normalized parameters Q and B_q using single-cosine-zscore method and the correlation coefficient between successive clusters up to cluster number $N_c = 50$ is shown in Fig. D.18. The clusters are examined at the peaks of ρ_c and the results are shown in Fig. D.19a between cluster numbers $N_c = 2$ and 39. At cluster number 2, two primary layers, a soil lense between the depths of 5.4 m and 5.7 m and a transition layer between the depths of 8 m and 9 m are discovered. At larger cluster numbers up to $N_c = 100$ (see Fig. D.19a and D.19b), no new primary layers ($t > 1$ m) are detected, therefore, cluster number 2 is chosen to represent the soil profile. The interpreted clustering is supported by both the subsurface stratigraphy based on laboratory and in-situ testing, and fines content with depth as shown in Fig. D.20. In the upper 8 m,



**Figure D.16. Piezocone Sounding at Newport News, Virginia
(Data from Mayne, 1989).**

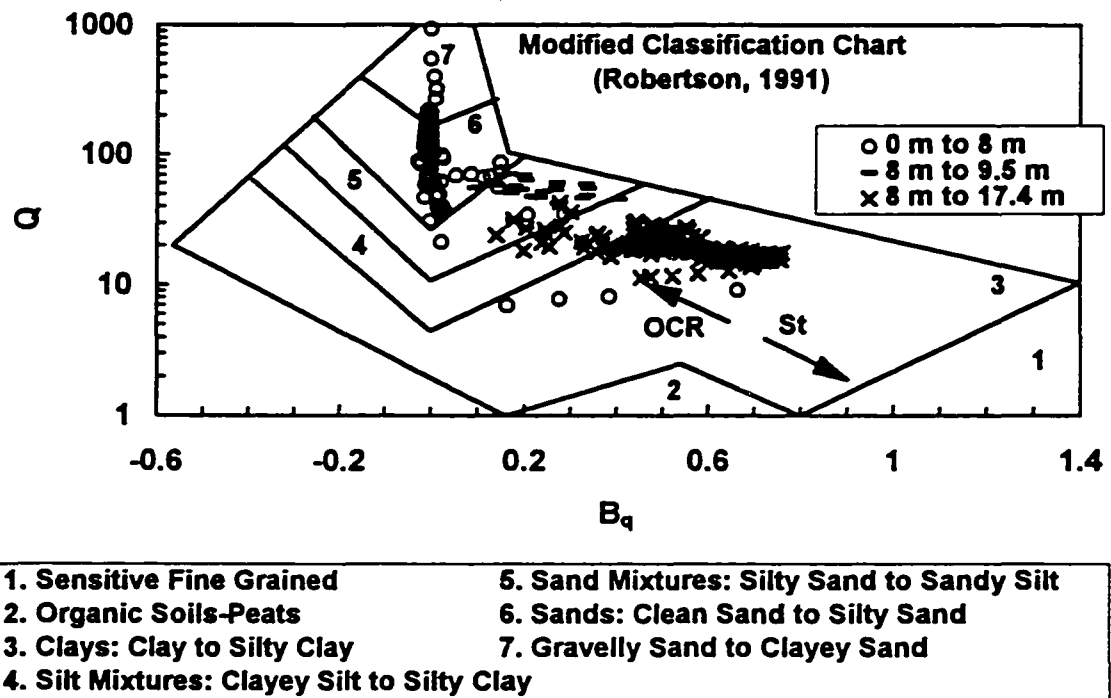


Figure D.17. Soil Classification Using Piezocone Data at Newport News, Virginia (Data from Mayne, 1989).

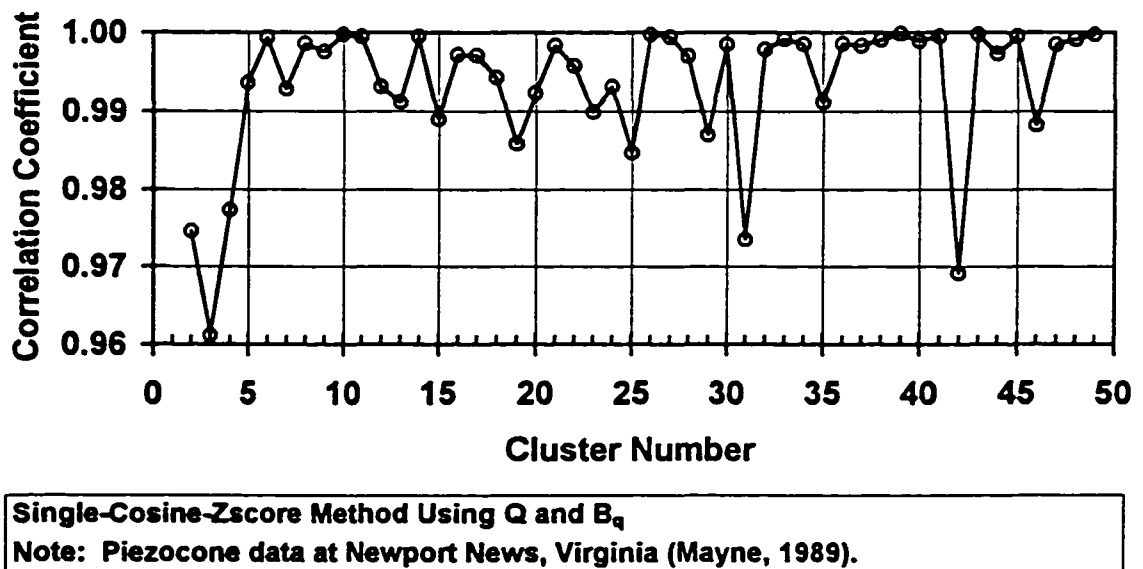
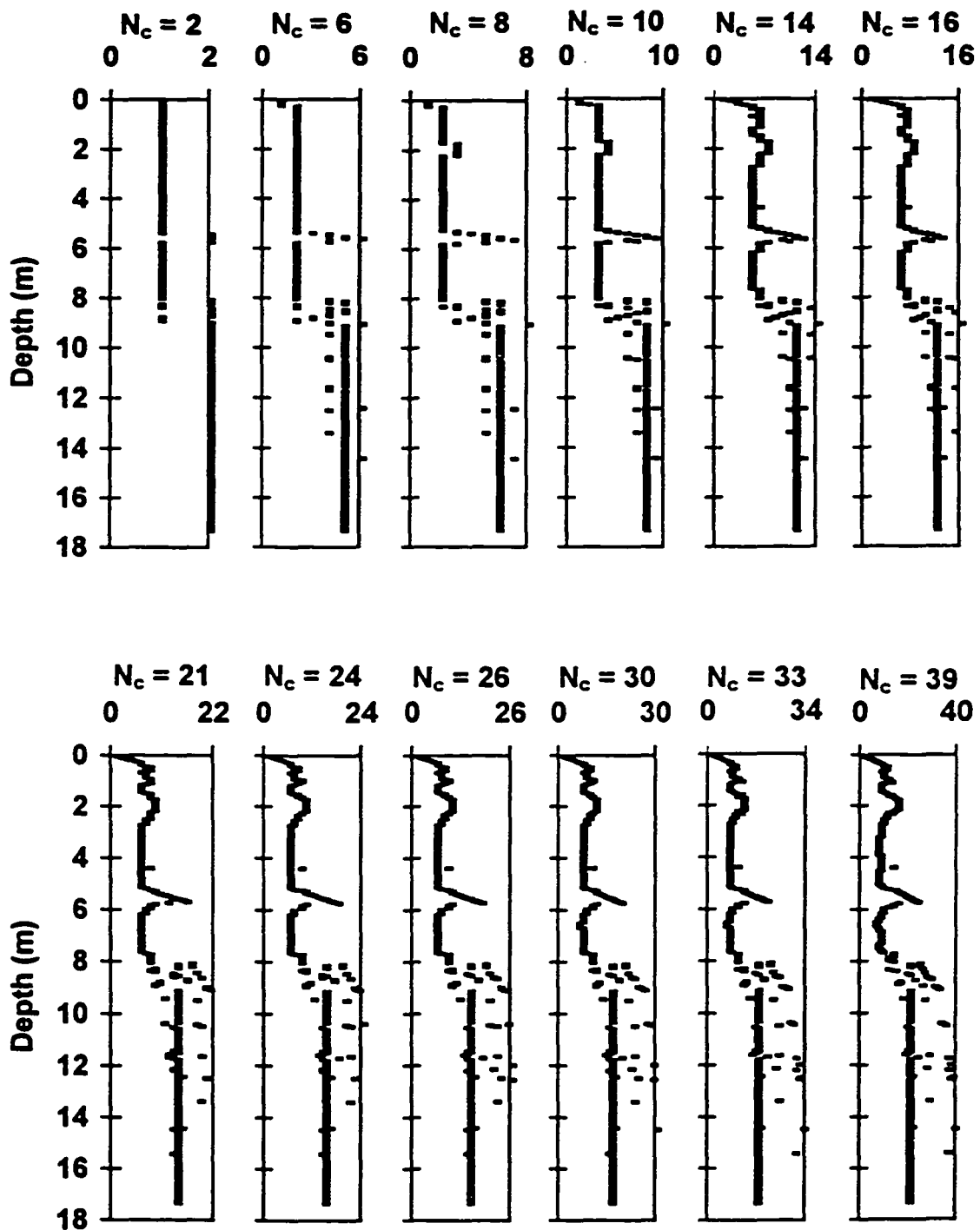
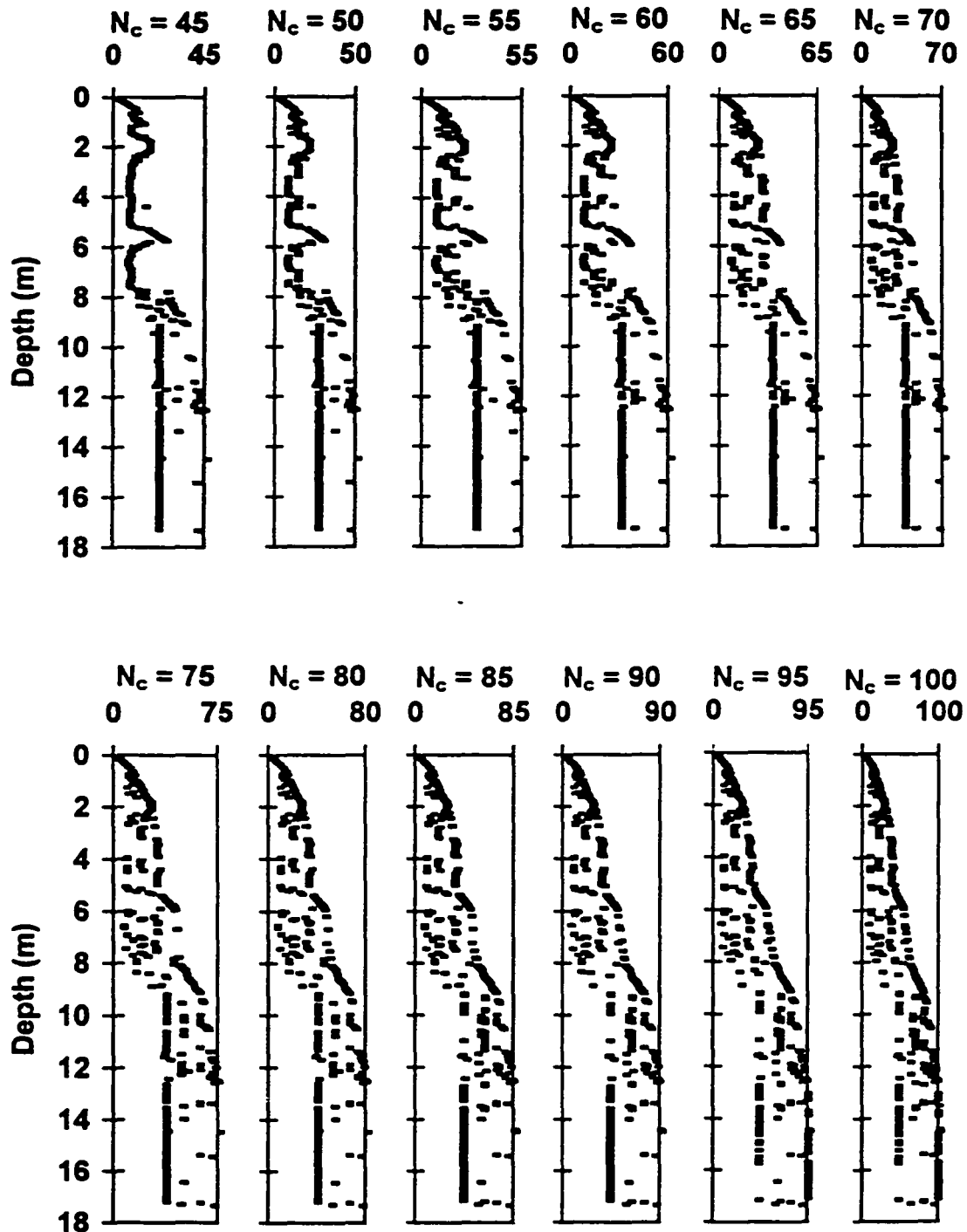


Figure D.18. Correlation Coefficient Between Consecutive Cluster Results at Newport News, Virginia.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Newport News, Virginia (Mayne, 1989).

Figure D.19a. Cluster Analysis of Piezocone Data at Newport News, Virginia.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Newport News, Virginia (Mayne, 1989).

Figure D.19b. Cluster Analysis of Piezocone Data at Newport News, Virginia.

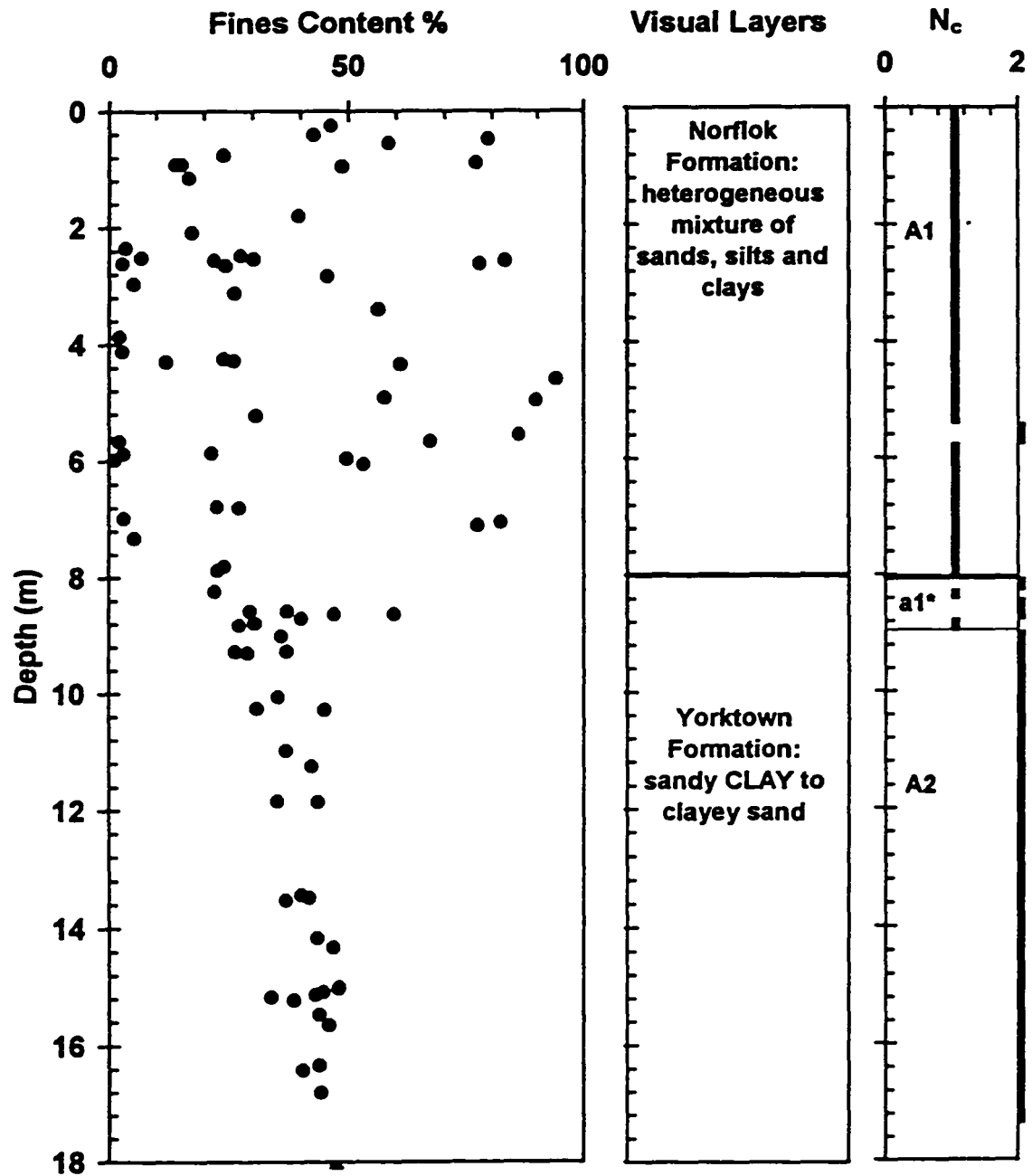


Figure D.20. Comparison Between Cluster Analysis, Visual Classification and Fines Content at Newport News, Virginia.

the fines content (FC) has an average equal to 36 percent and a coefficient of variation equal to 0.6 indicating a large scatter compared with the lower layer which has an average fines content equal to 40 percent with a coefficient of variation equal to 0.14.

D.6. Onsoy Test Site, Norway

Clustering of piezocone data from the Onsoy clay in Norway was conducted. The soil profile downward from the ground surface consists of 1 m of weathered crust and 8 m of soft clays with iron spots, organic matter and fragments of shell underlain by 22 m of soft medium-plastic clays (Gillespie et al., 1985). The groundwater table is at the ground surface. A representative piezocone sounding and the derived normalized parameters Q and B_q are shown in Fig. D.21. Looking at the piezocone profiles, the trend of q_t increases with depth, and that of Q and B_q decrease with depth although there is no sharp change in any of the profiles. Visually, it is not obvious where to define the demarcation(s) between different soil layers. The classification chart is used to explore the subsurface stratigraphy as shown in Fig. D.22, however, boundaries were detected at 3.2 m and 3.9 m. The soil below 3.9 m is classified a continuous layer of clays to silty clays. Therefore, both discussed schemes could not properly detect the depths of soil boundaries and the association between different layers.

Clustering is applied to Q and B_q parameters using single-cosine-zscore technique and the correlation coefficient between successive clusters are shown in Fig. D.23 up to cluster number $N_c = 50$. The cluster results are examined at the peaks of ρ_c and shown in Fig. D.24a between cluster numbers $N_c = 2$ and 33. The data are separated at a depth of

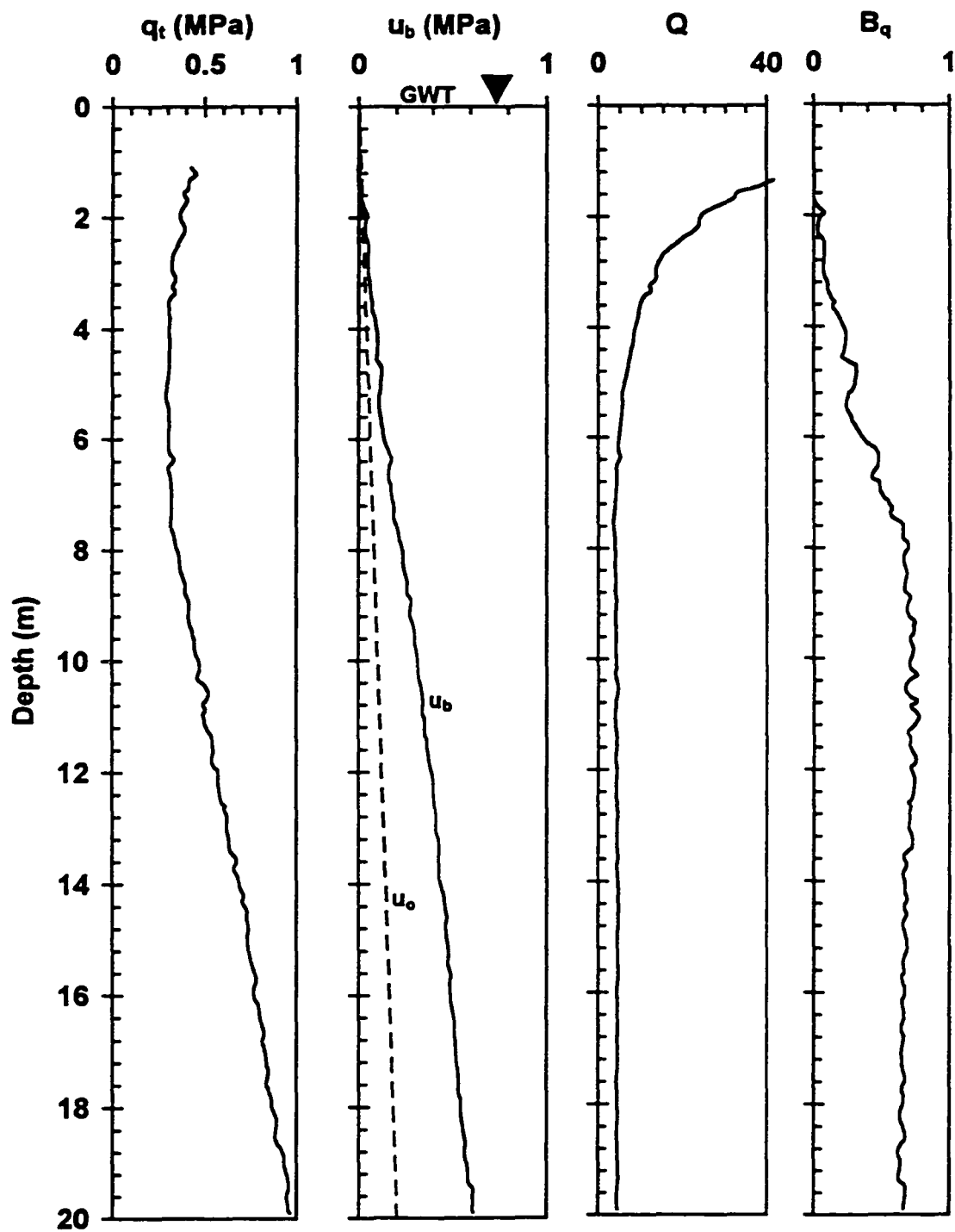


Figure D.21. Piezocone Sounding at Onsoy, Norway
(Data from Gillespie et al., 1985).

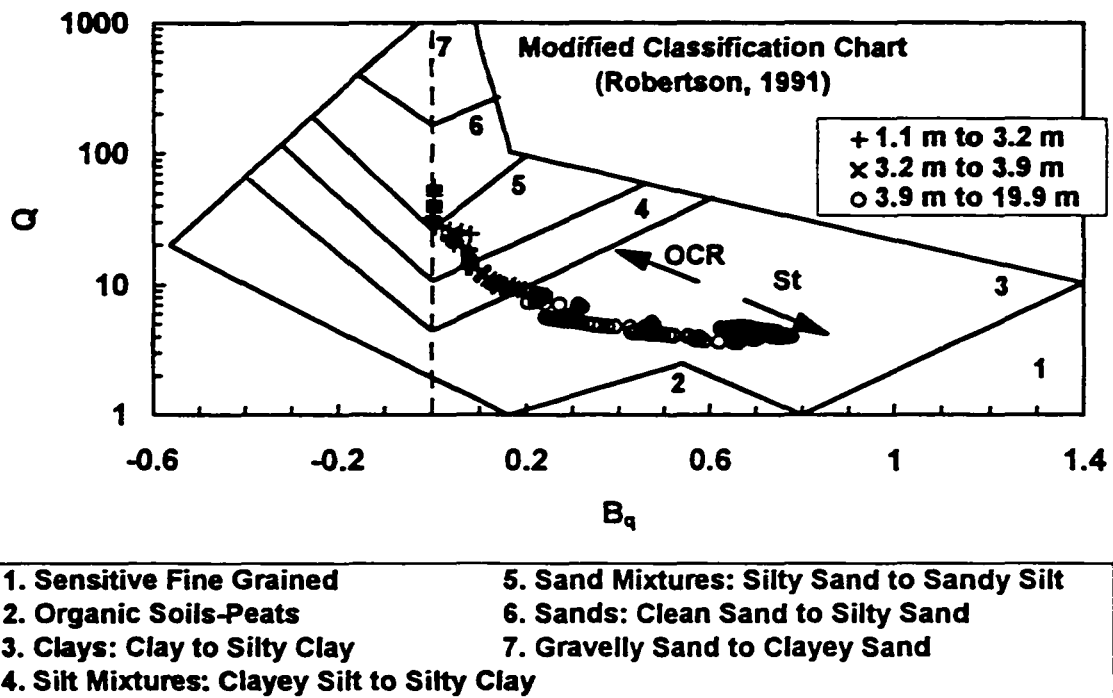
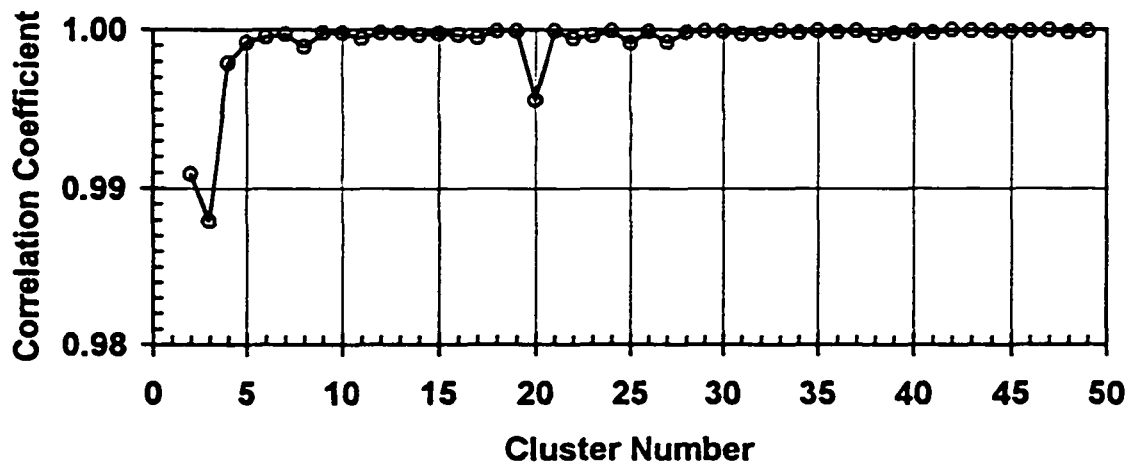
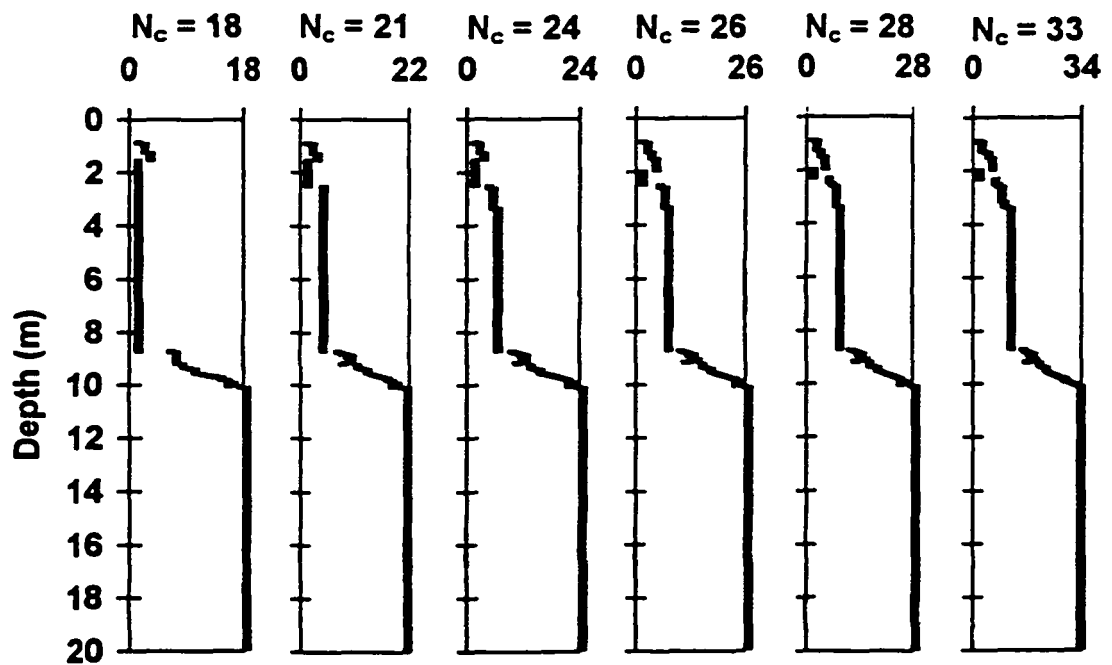
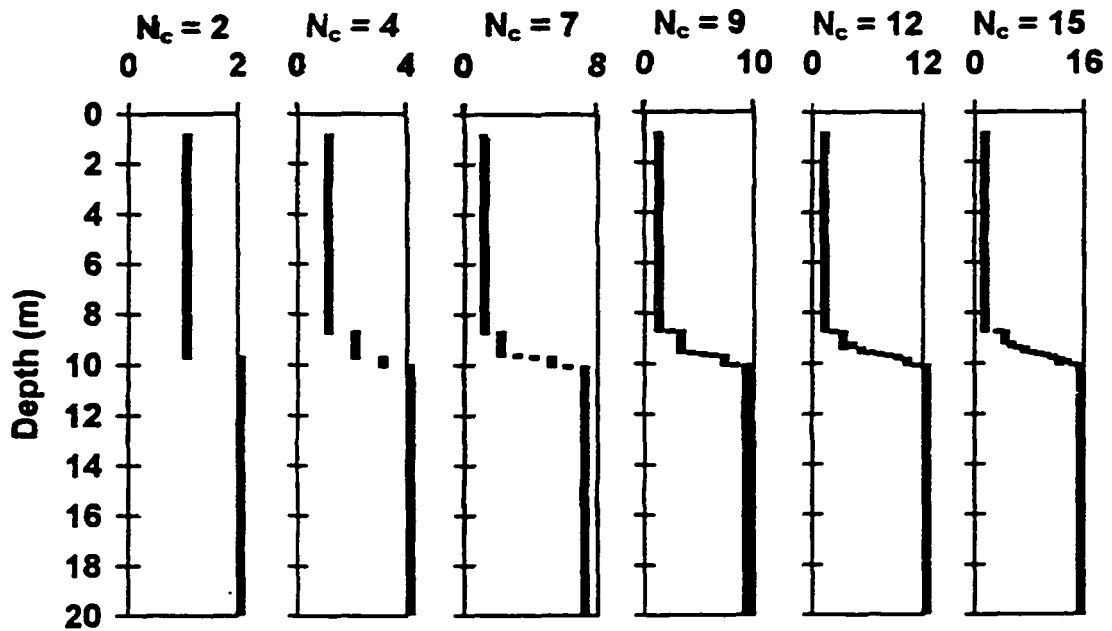


Figure D.22. Soil Classification Using Piezocone Data at Onsoy, Norway (Data from Gillespie et al., 1985).



Single-Cosine-Zscore Method Using Q and B_q
 Note: Piezocone data at Onsoy, Norway (Gillespie et al., 1985).

Figure D.23. Correlation Coefficient Between Consecutive Cluster Results at Onsoy, Norway.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Onsoy, Norway (Gillespie et al., 1985).

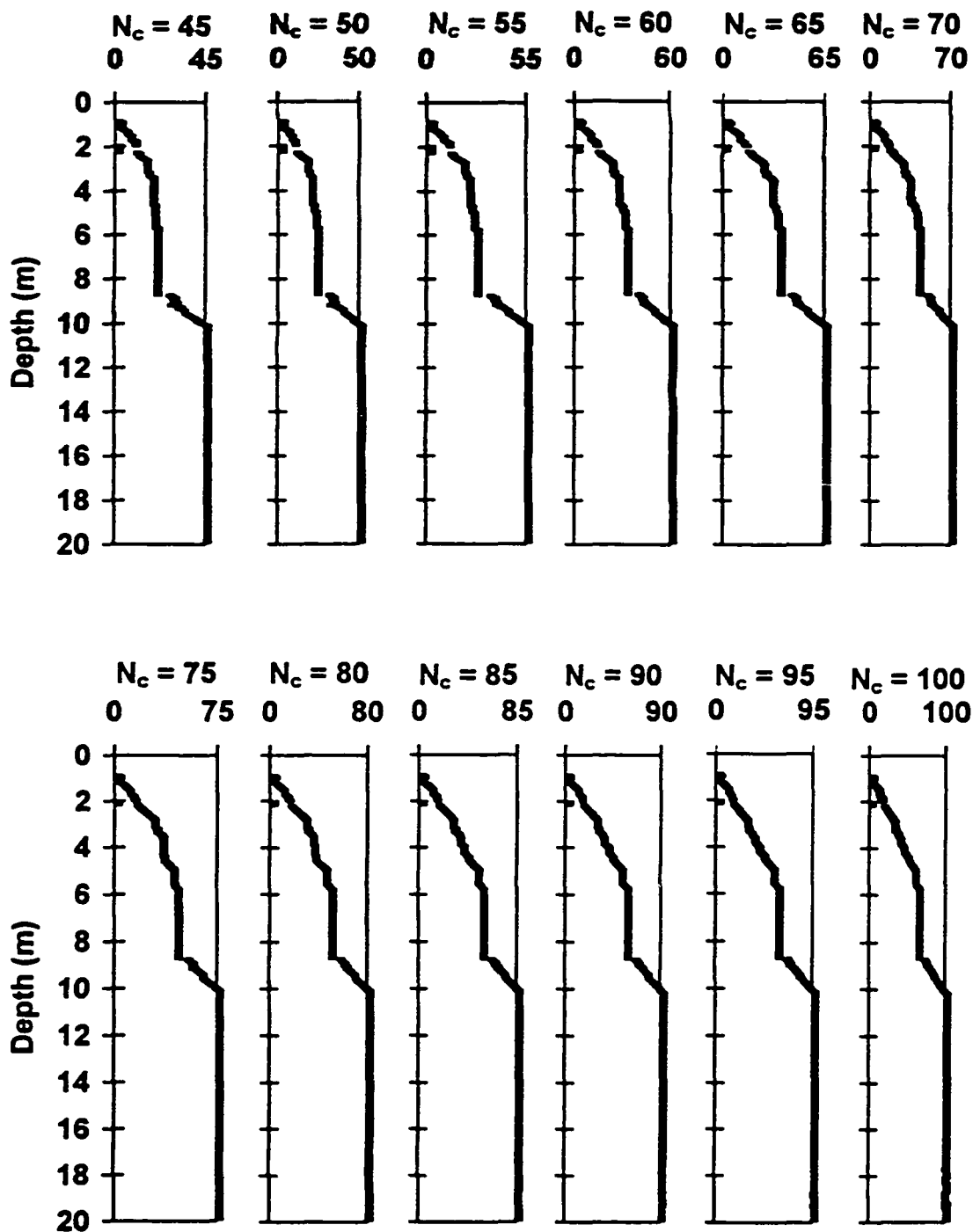
Figure D.24a. Cluster Analysis of Piezocone Data at Onsoy, Norway.

9.7 m into two primary clusters. At larger cluster numbers, some points are denoted as different clusters other than the two main groups which indicates a transition zone or less association between the data at the upper 2 m of the top soil. Also for N_c up to 100 as seen in Figures D.24a and D.24b, no more primary layers ($t > 1$ m) separates, therefore cluster number 2 is chosen to indicate the soil profile.

Clustering is supported by the soil layering and boundaries defined based on laboratory and field testing including undrained shear strength measurements obtained by field vane tests as shown in Fig. D.25. The statistical boundary detected at 9.65 m is matching with the boundary between the upper and lower marine clays. The undrained shear strength (s_u) readings have two trends with depth as follows: first s_u changes slightly from 11 kPa at 1.8 m to 14 kPa at 9.6 m and then increases dramatically up to 35 kPa at 20 m.

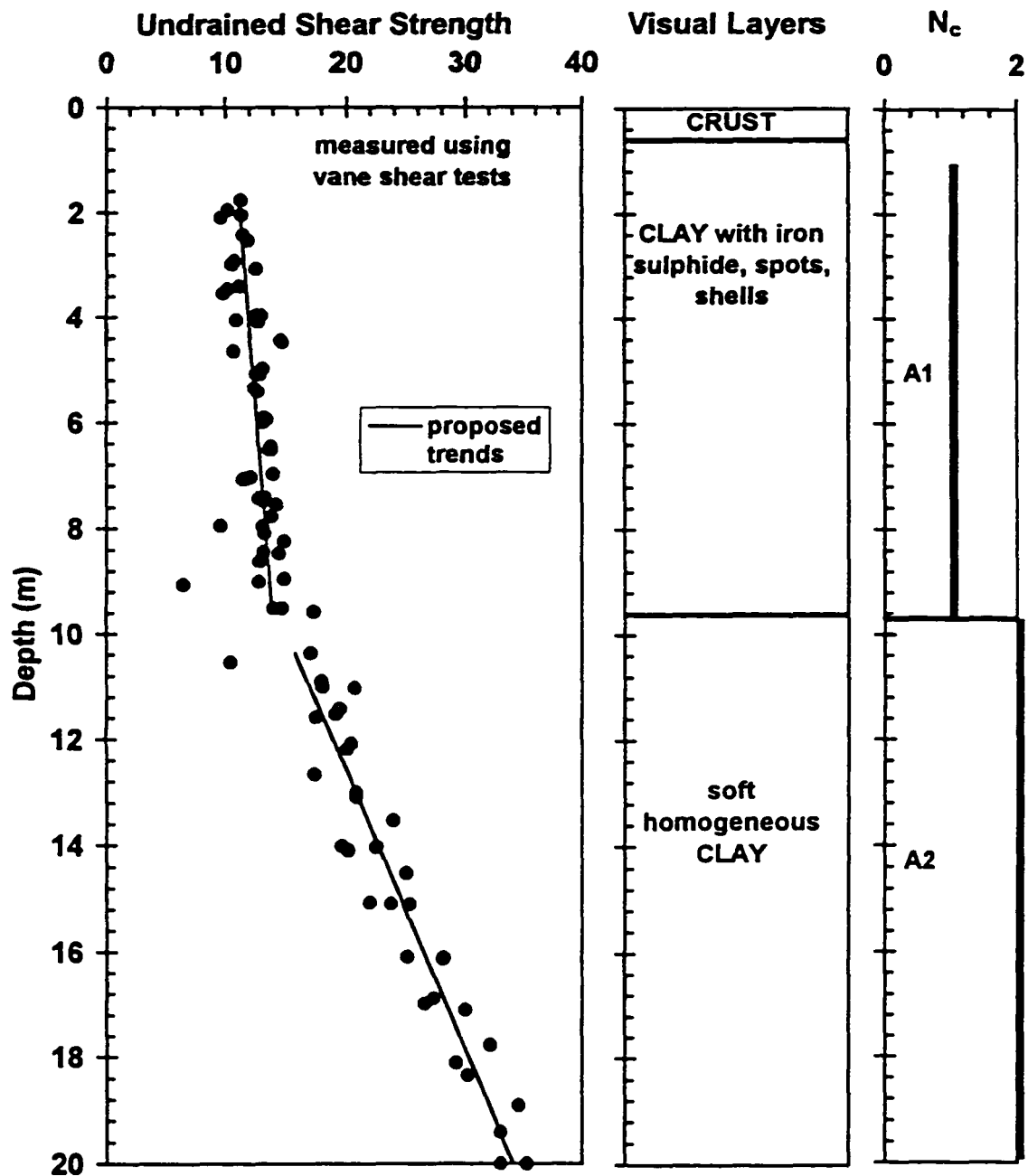
D.7. Opelika Test Site, Alabama

The Opelika site in Alabama resides in the southern Piedmont province which was formed from Precambrian to Paleozoic era high-grade metamorphic and igneous rocks (Vinson and Brown, 1996). The residual soils are primarily of the Wacoochee Complex. The site geology is classified as either Halawaka Schist or Phelps Creek Gneiss. The groundwater table at the time of testing was at a depth of 3 m. Representative piezocone data and the normalized parameters Q and B_q are shown in Fig. D.26. The vertical profiles suggest that the subsurface stratigraphy consists of one primary layer. The decrease of q_t and B_q readings below 14.5 m might indicate a secondary layer. The



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Onsoy, Norway (Gillespie et al., 1985).

Figure D.24b. Cluster Analysis of Piezocone Data at Onsoy, Norway.



Note: Piezocone data at Onsoy, Norway (Gillespie et al., 1985).
 SCZ method using Q and B_q , cluster no. 2 — Primary boundary

Figure D.25. Comparison Between Cluster Analysis, Visual Classification, and Undrained Shear Strength at Onsoy, Norway.

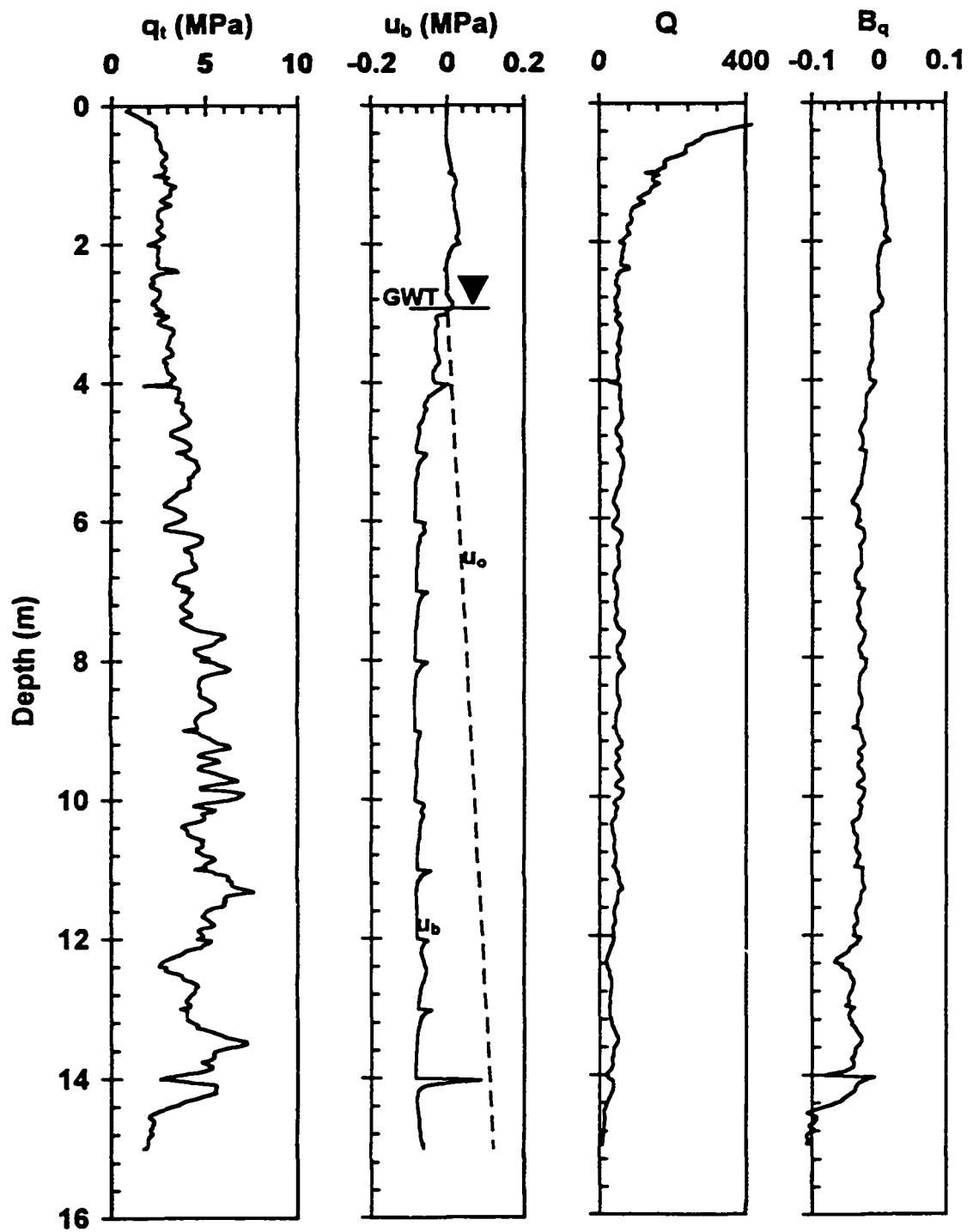
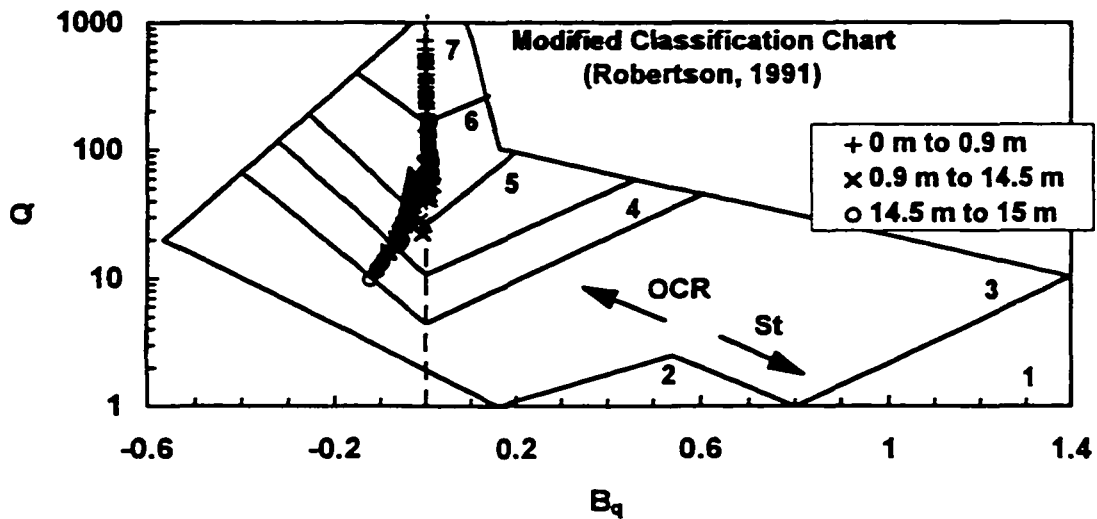


Figure D.26. Piezocone Sounding at Opelika, Alabama
(Data from this Study).

Robertson chart (1991) was used to delineate soil types as shown in Fig. D.27. Three subsurface layers were detected as follows: (1) gravelly sand to clayey sand between the ground surface and a depth of 0.9 m, (2) sand and silty sand to sandy silt between the depths of 0.9 m and 14.5 m and (3) clayey silt to silty clay below 14.5 m down to a depth of 15 m.

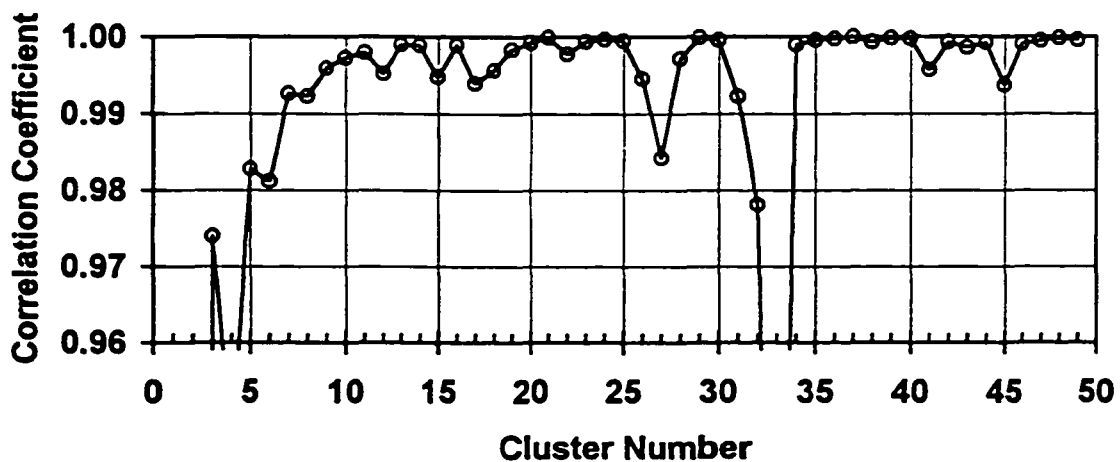
A single-cosine-zscore cluster method is applied to the normalized parameters Q and B_q up to cluster number $N_c = 100$. The correlation coefficient is determined between consecutive clusters as shown in Fig. D.28 up to $N_c = 50$. Clustering is examined at the peaks of ρ_c and Figure D.29a shows the cluster results between cluster numbers $N_c = 2$ and 39. At $N_c = 2$, one primary layer denoted by cluster number 2 was detected and one secondary layer denoted by cluster number 1 was discovered between the ground surface and a depth of 0.75 m. At $N_c = 3$, another primary layer was defined between the depths of 0.75 m and 3 m. Also, some soil lenses and/or outliers ($t < 0.5$ m) denoted by cluster number 2 are seen in Fig. D.29a. For $N_c > 2$, no new primary layers ($t > 1$ m) separated which is also true in case of clustering between $N_c = 45$ and 100 as shown in Fig. D.29b. Therefore, cluster number 3 was chosen to indicate the subsurface stratification.

Clustering is verified by the visual description from 8 boreholes and the average soil contents obtained using sieve analysis as shown in Fig. D.30. In the upper 2 m, the average fines and sands contents are equal to 60 percent and 40 percent, respectively. The clay fraction (CF) decreases from 23 percent at a depth of 1 m to 8 percent at a depth of 2 m. Below a depth of 3 m, the soil almost contains an equal mixture of fines and sands



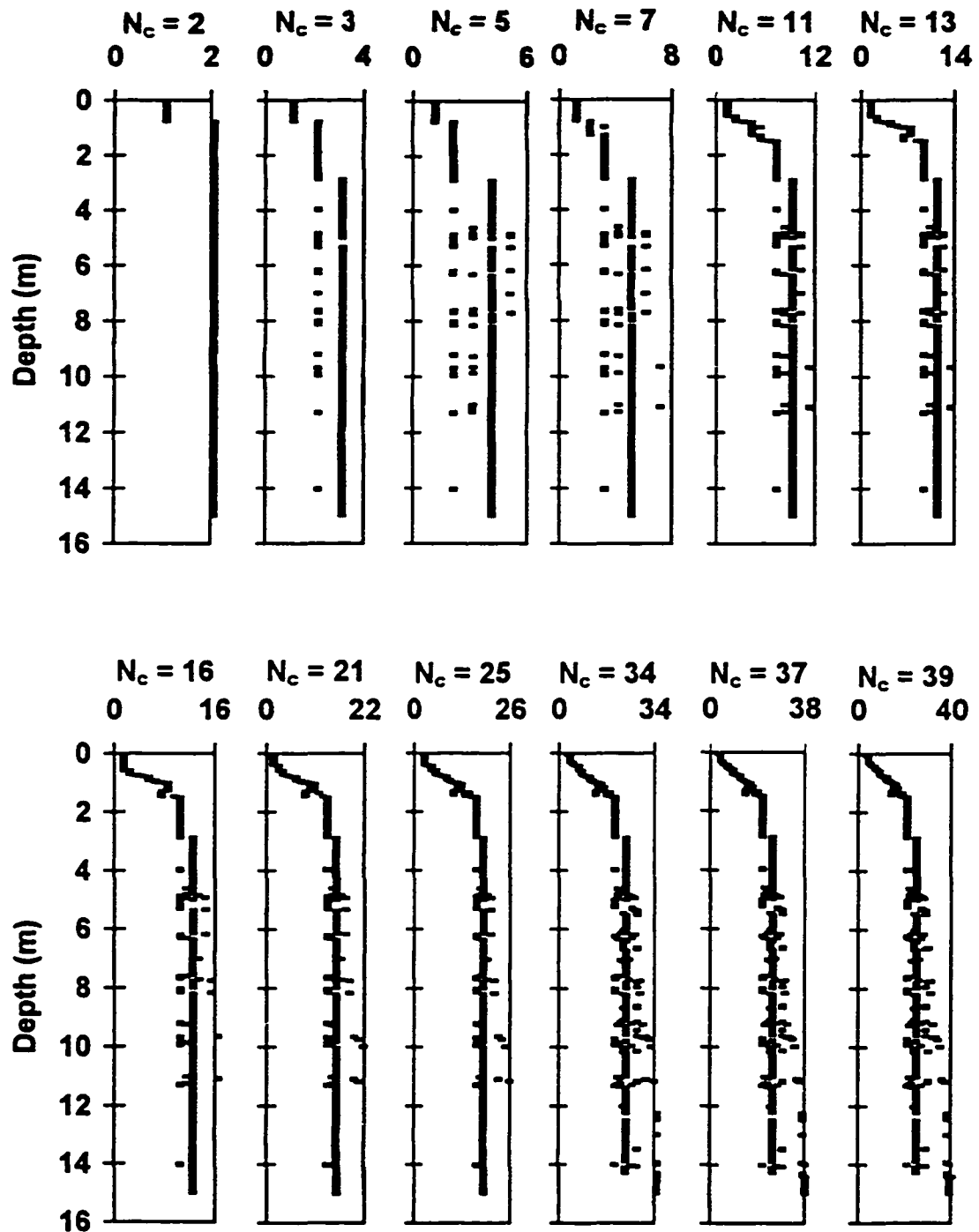
- | | |
|---|--|
| 1. Sensitive Fine Grained | 5. Sand Mixtures: Silty Sand to Sandy Silt |
| 2. Organic Soils-Peats | 6. Sands: Clean Sand to Silty Sand |
| 3. Clays: Clay to Silty Clay | 7. Gravelly Sand to Clayey Sand |
| 4. Silt Mixtures: Clayey Silt to Silty Clay | |

Figure D.27. Soil Classification Using Piezocone Data at Opelika, Alabama (Data from this Study).



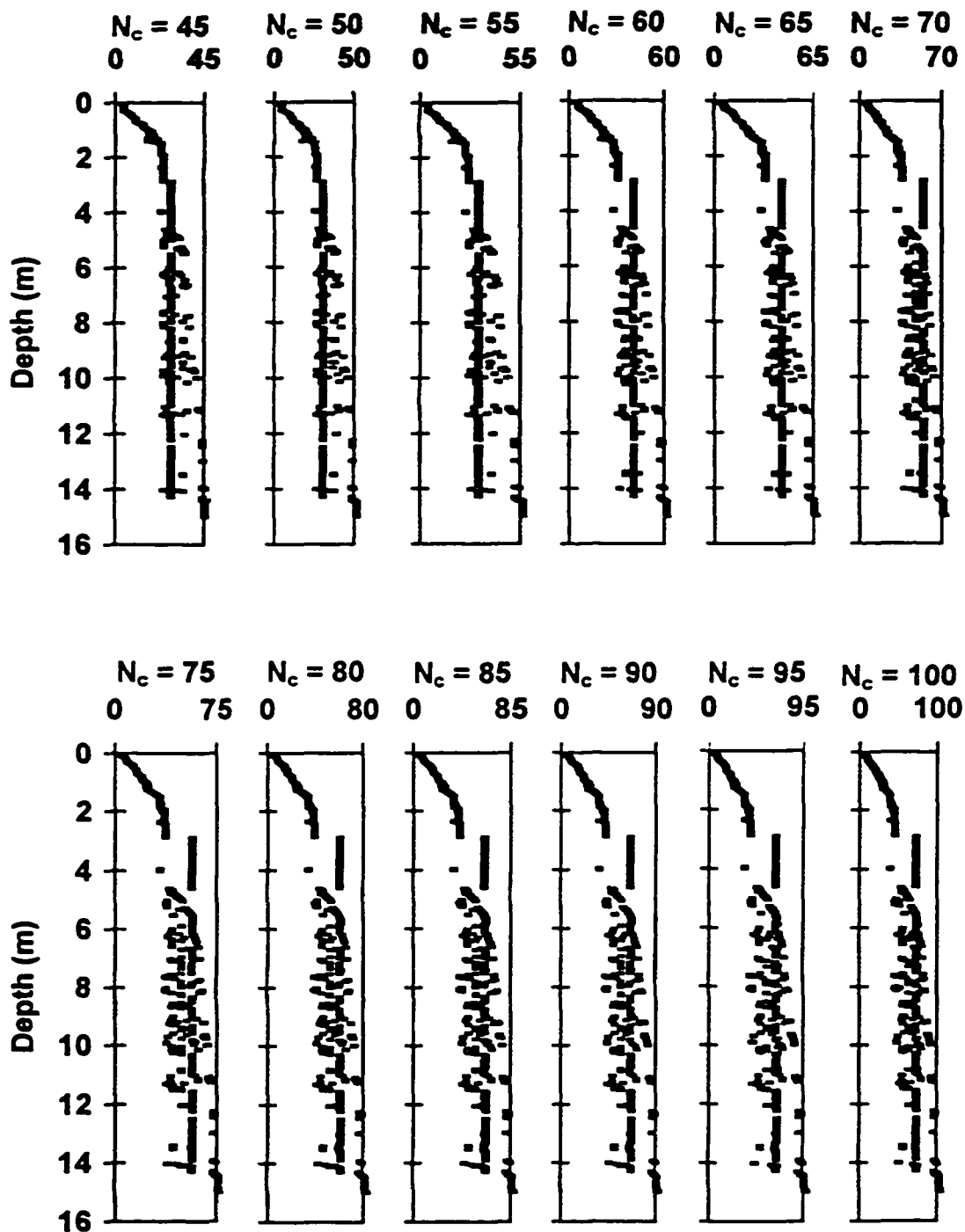
Single-Cosine-Zscore Method Using Q and B_q
Note: Piezocone data at Opelika, Alabama (this study).

Figure D.28. Correlation Coefficient Between Consecutive Cluster Results at Opelika, Alabama.



Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data at Opelika, Alabama (this study)

Figure D.29a. Cluster Analysis of Piezocone Data at Opelika, Alabama.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Opelika, Alabama (this study)

Figure D.29b. Cluster Analysis of Piezocone Data at Opelika, Alabama.

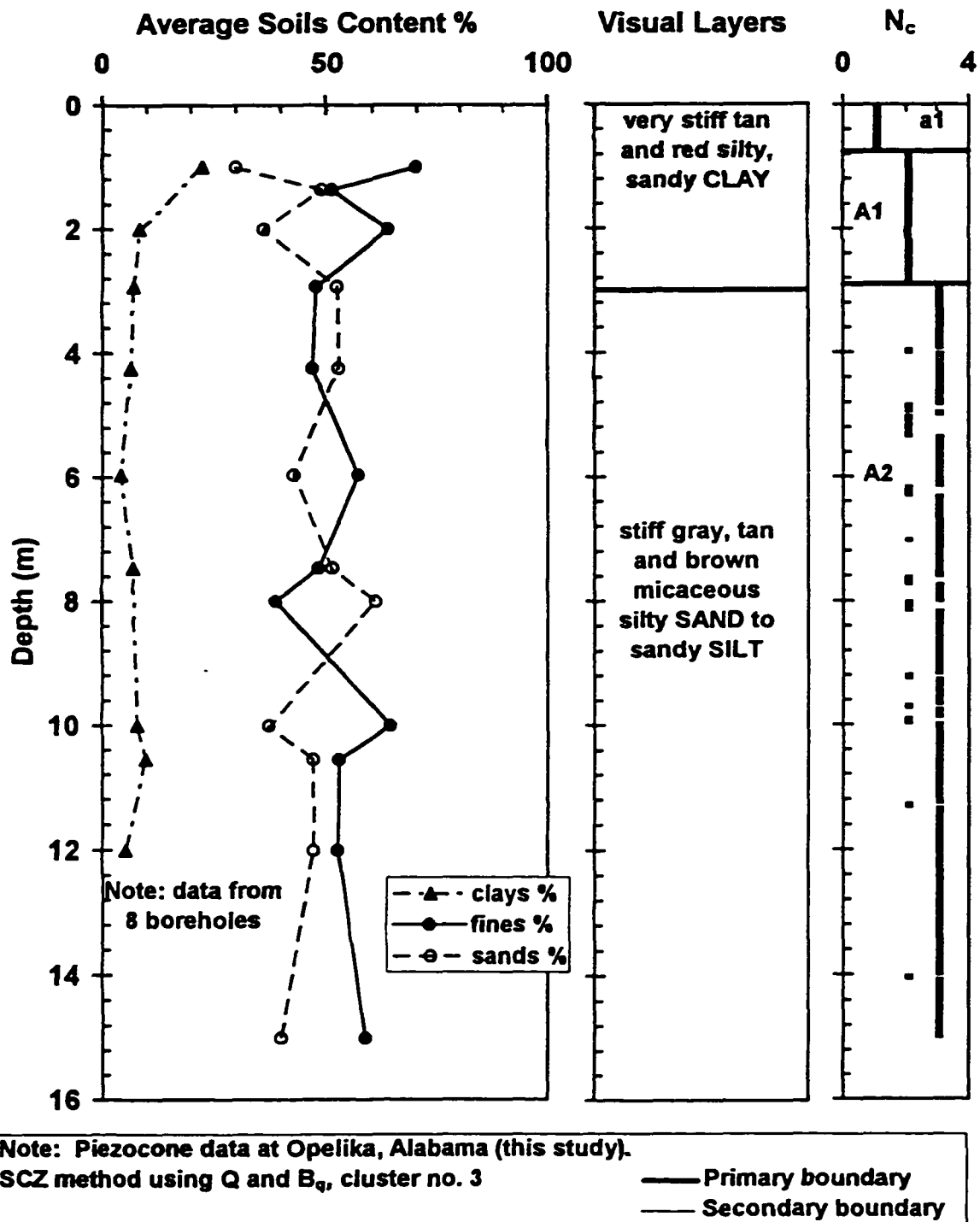


Figure D.30. Comparison Between Cluster Analysis, Visual Classification and Soils Content at Opelika, Alabama.

with averages equal to 52 percent and 48 percent, respectively. The average clay fraction is equal to 7 percent.

D.8. Po River Sand, Italy

Po River site in Italy has been extensively tested and reported in the literature (e.g., Jamiolkowski et al., 1985). The upper 15 m of the geological setting is a very recent Holocene deposit and consists of clean to slightly silty sand overlain by a 5 m to 7 m layer of clayey silt (Bruzzi et al., 1986). The underlying layer is a glacial deposit and consists of 30 m to 40 m of clean to slightly silty sand with embedded thin lenses of cohesive soils. Both deposits experienced minor mechanical overconsolidation due to erosion of the Holocene deposit. The groundwater table is at a depth of 2.2 m.

A representative piezocone sounding and the derived normalized parameters Q and B_q are shown in Fig. D.31. Looking at the piezocone profiles, soil lenses are seen at the spikes of q_t , u_b , Q and/or B_q at depths of 5.5 m, 8 m, 11.2 m, 14 m, 20.5 m, 21.5 m, 23 m and 24.5 m. A layer of fine soils is detected at depths of 23.5 m to 25 m. The soil is classified using the Robertson's chart (1991) as shown in Fig. D.32. The soil between the depths of 5 m to 30 m is denoted as clean sands to silty sands except an intermediate layer between the depths of 23 m to 25 m is defined as silty clays to clayey silts with some sands.

A single-cosine-zscore clustering is applied to the normalized data Q and B_q to divide the data into correlated groups. The correlation coefficient between consecutive

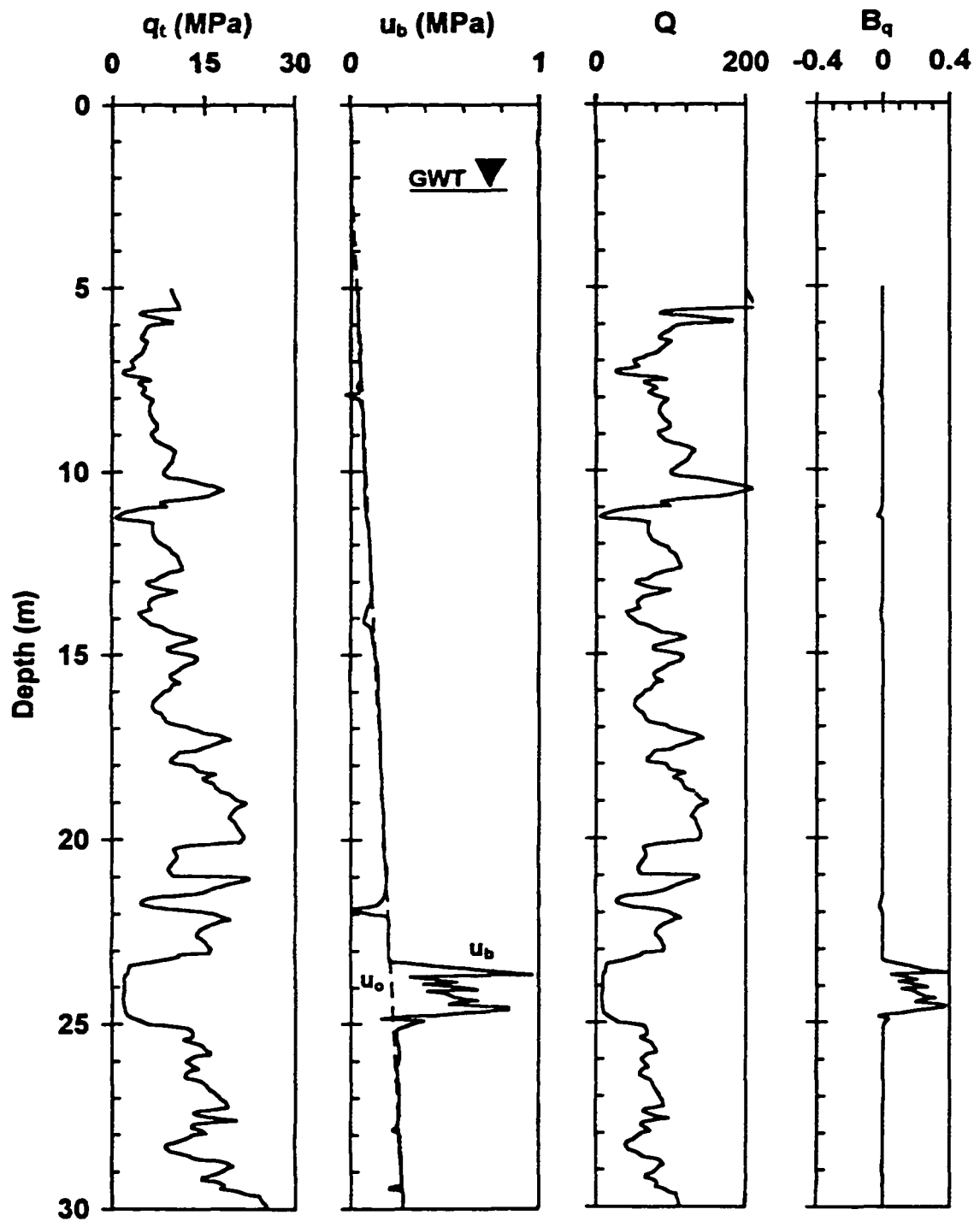
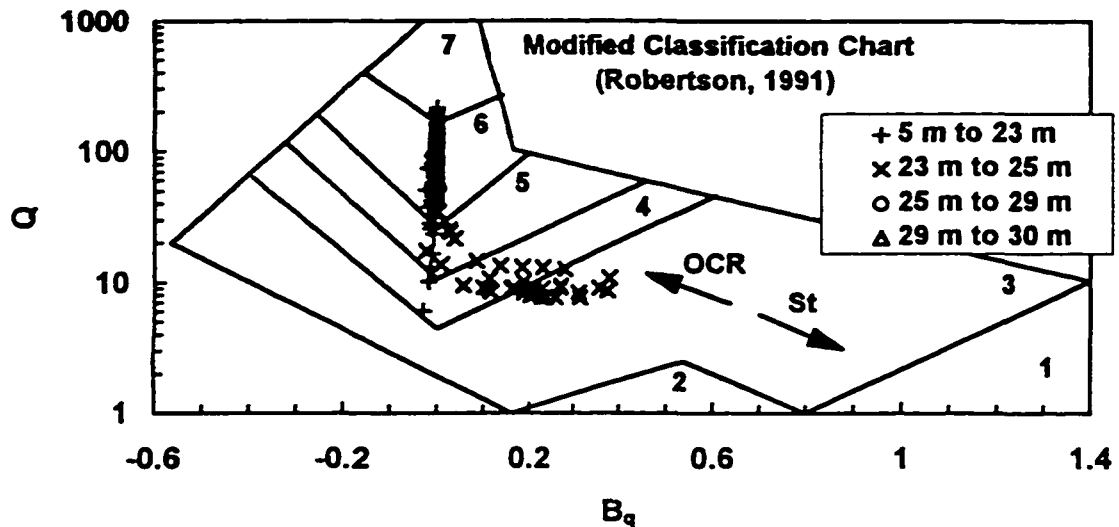


Figure D.31. Piezocone Sounding at Po River, Italy
 (Data from Bruzzi et al., 1986).

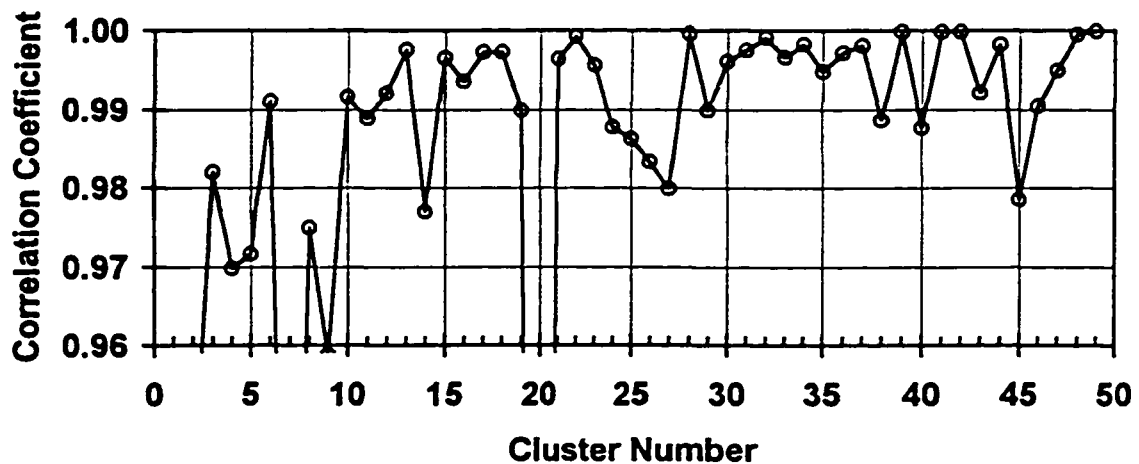
clusters up to cluster number $N_c = 50$ is shown in Fig. D.33 and the clusters are examined at the peaks of ρ_c . Cluster results are shown in Fig. D.34a between cluster numbers $N_c = 2$ to 34. At cluster number 2, the profile is divided into one major layer and an intermediate primary layer between 23.3 m and 24.8 m. A soil lense is also detected at depths of 11 m to 11.3 m. Up to $N_c = 8$, some data points separated from the two layers indicating soil lenses and transition zone between the two layers. At $N_c = 8$, an upper primary layer is detected between the depths of 5 m and 6.2 m and a secondary layer is also discovered between the depths of 29.5 m to 30 m. The former layer is a transition from the silty clays to the sands and the latter layer is a transition from the sands to the lower clayey silts. These two layers and soil lenses are grouped in cluster number 1 that indicates a similarity in their soil type and/or behavior. There is association between clusters, for instance, A1 and A2 which are grouped in clusters 1 and 2, respectively. However there is less association between the data grouped in clusters 1 and 7 such as a1 and a2, respectively. At clustering greater than 8 groups, no new primary layers ($t > 1\text{m}$) is detected, therefore cluster number 8 is chosen to represent the subsurface profile. This note is also valid for clustering between $N_c = 45$ and 100 as seen in Fig. D.34b.

Cluster results are confirmed by both soil stratigraphy based on field and laboratory testing including fines content measurements as shown in Fig. D.35. The statistical layers A1 and a3 are an indication of the transition zones between fine and coarse soils. The primary and secondary layers, A3* and a2, are supported by the increase of the fines content from almost 5 percent to 30 percent.



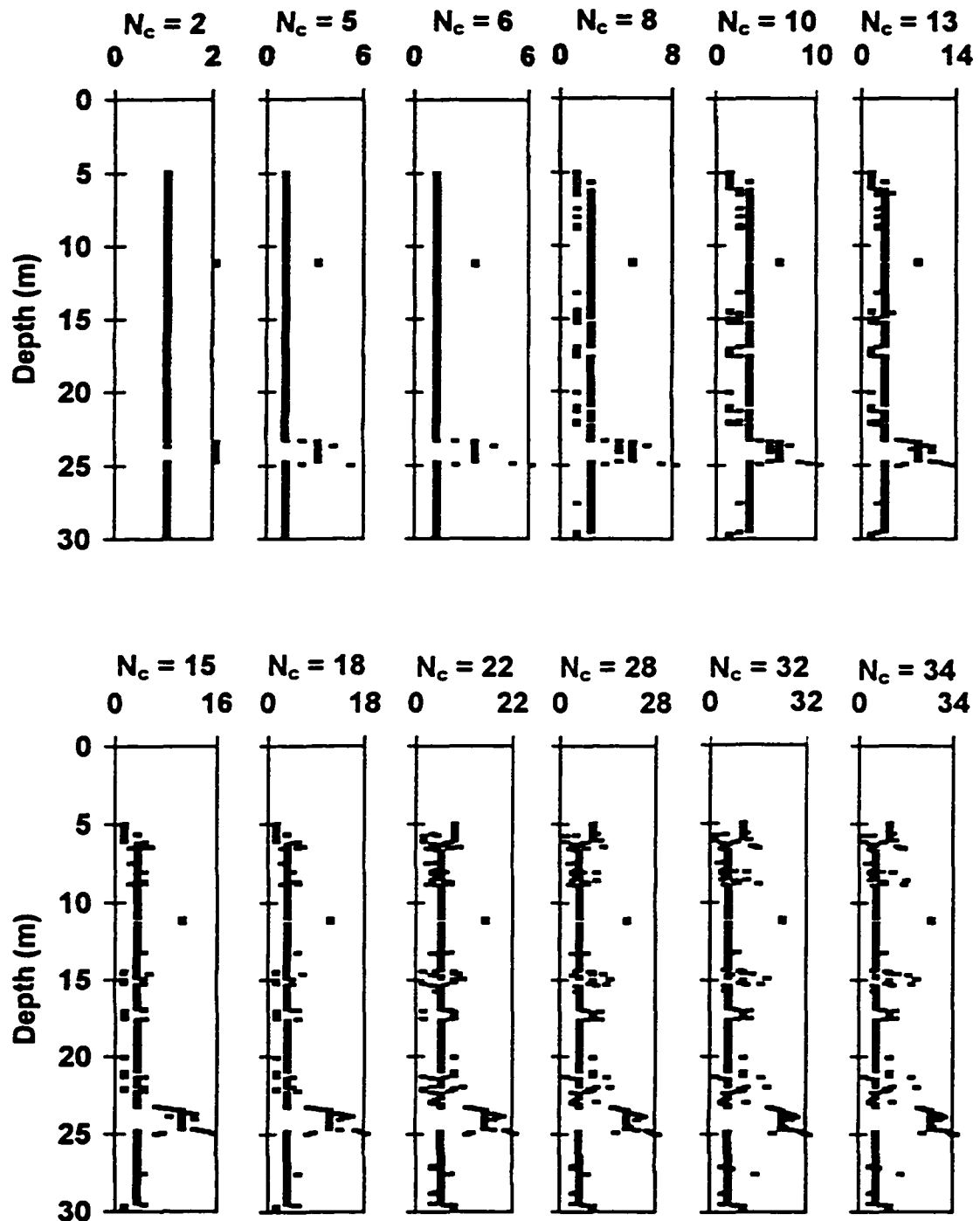
1. Sensitive Fine Grained	5. Sand Mixtures: Silty Sand to Sandy Silt
2. Organic Soils-Peats	6. Sands: Clean Sand to Silty Sand
3. Clays: Clay to Silty Clay	7. Gravelly Sand to Clayey Sand
4. Silt Mixtures: Clayey Silt to Silty Clay	

Figure D.32. Soil Classification Using Piezocone Data at Po River, Italy (Bruzzi et al., 1986).



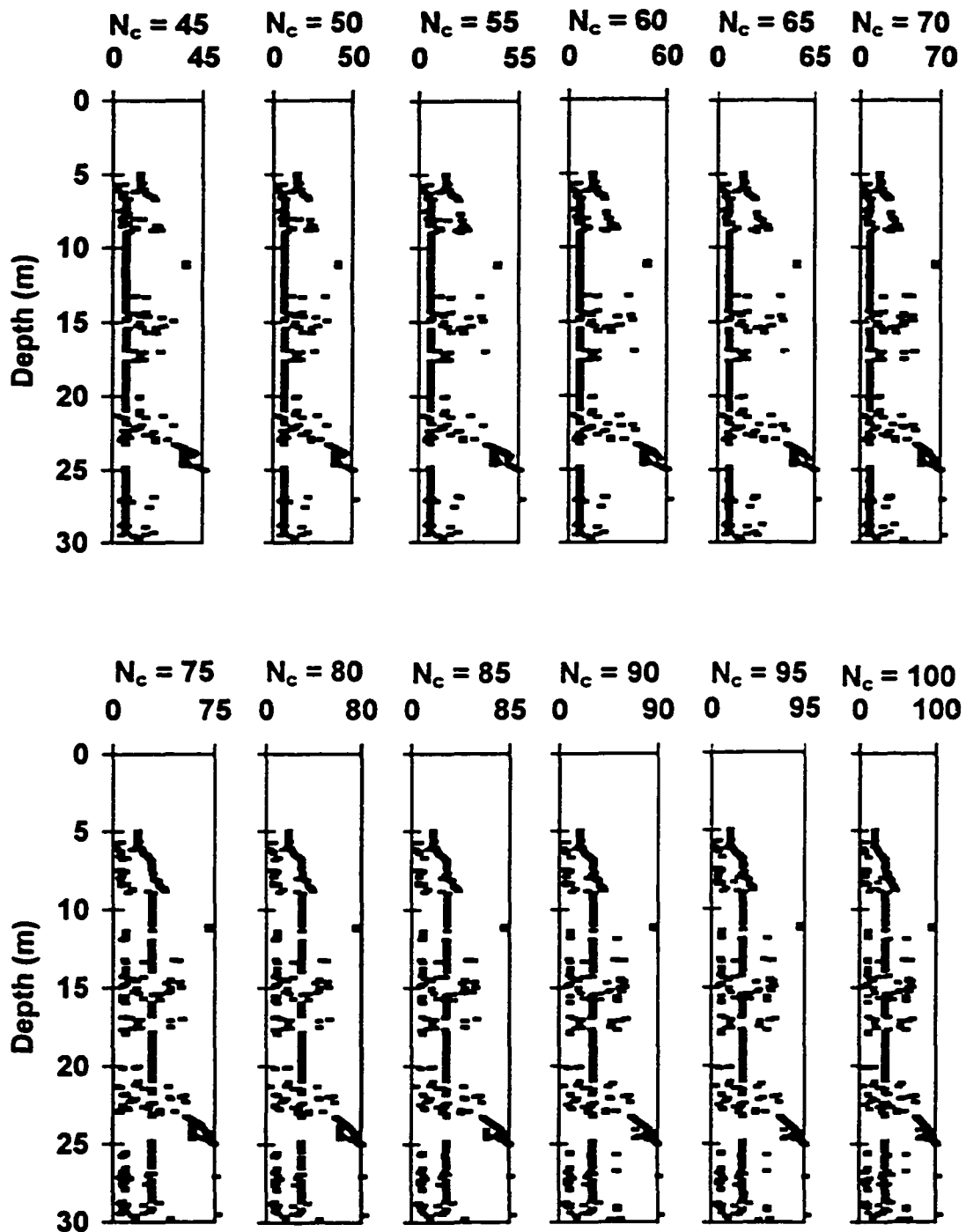
Single-Cosine-Zscore Method Using Q and B_q
Note: Piezocone data at Po River, Italy (Bruzzi et al., 1986).

Figure D.33. Correlation Coefficient Between Consecutive Cluster Results at Po River, Italy.



Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data at Po River, Italy (Bruzzi et al., 1986).

Figure D.34a. Cluster Analysis of Piezocone Data at Po River, Italy.



Single-Cosine-Zscore Method Using Q and Bq N_c = No. of Clusters
 Note: Piezocone data at Po River, Italy (Bruzzi et al., 1986).

Figure D.34b. Cluster Analysis of Piezocone Data at Po River, Italy.

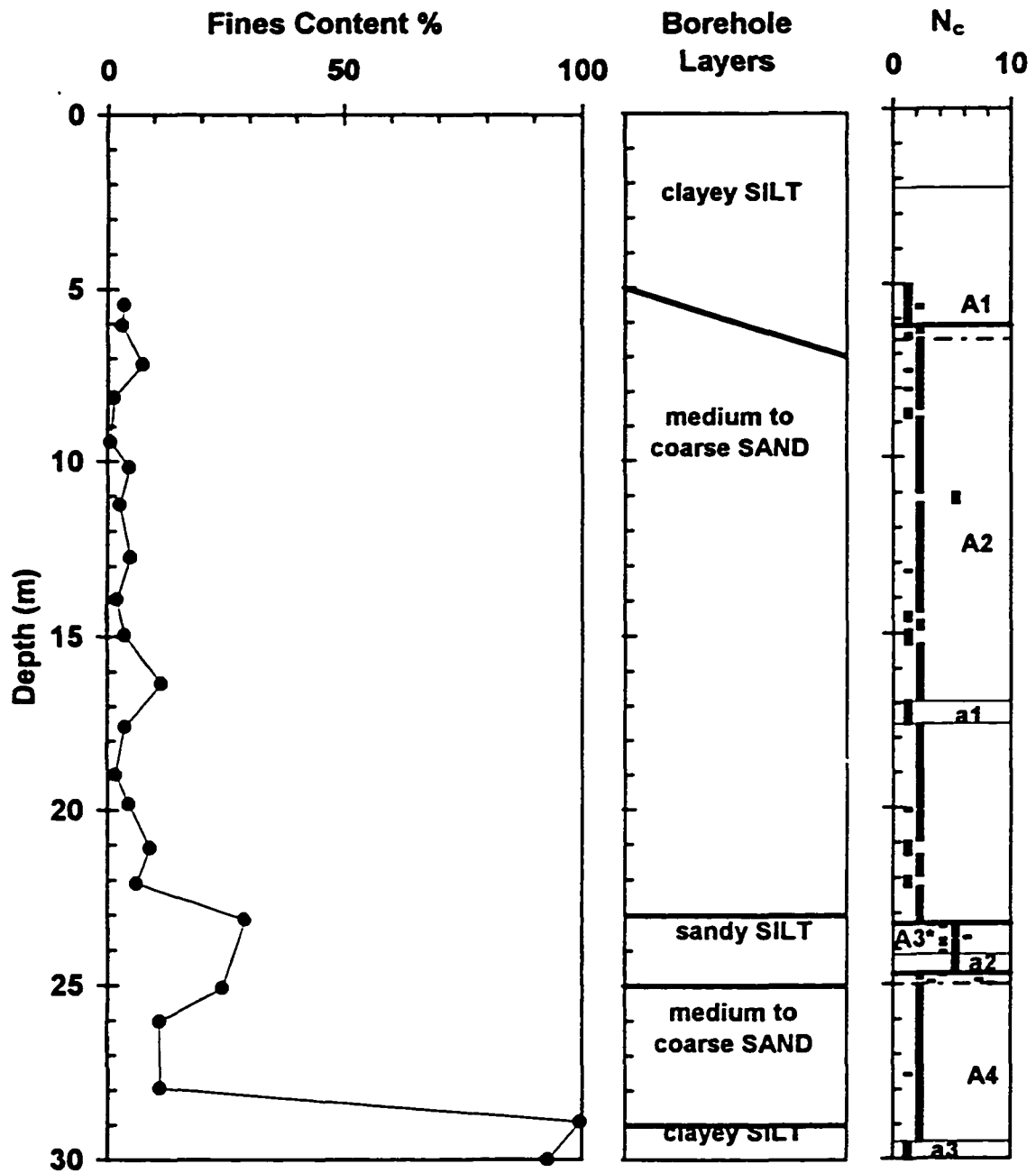
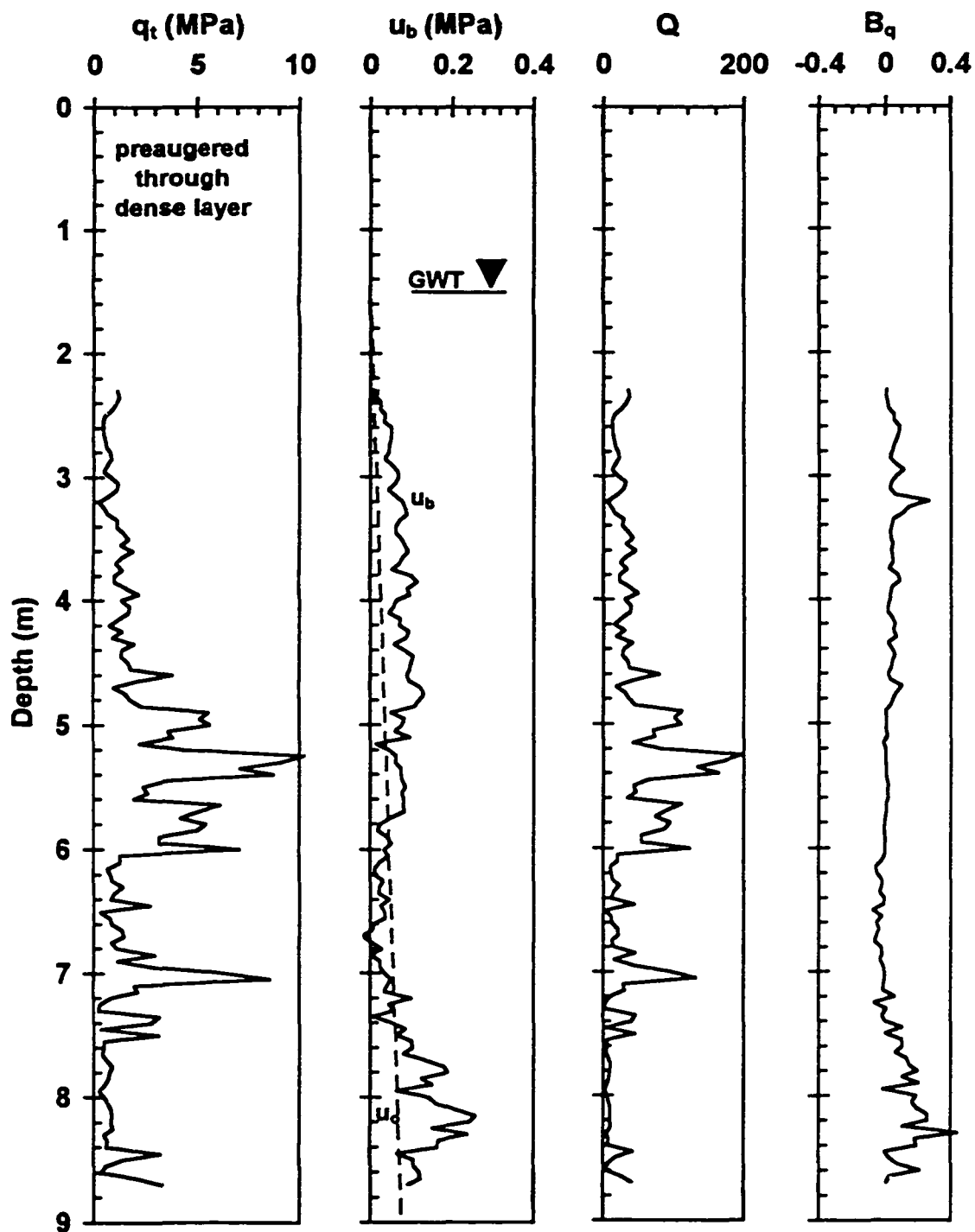


Figure D.35. Comparison Between Cluster Analysis, Visual Classification and Fines Content at Po River, Italy.

D.9. Penuelas Test Site, Puerto Rico

Piezocone tests were performed by the author at Penuelas test site, Puerto Rico for the benefit of a local consulting firm as part of a large conducted geotechnical investigation. The subsurface stratigraphy in the upper 10 meters consists of 2-m of silty sand fill overlain 8-m of a Holocene marine deposit (Chen, 1995). The latter deposit includes interchangeable soil layers of soft clayey silts, and loose silty clayey sands to sandy clayey silts. The groundwater table is at approximately 1.5 m. A representative piezocone sounding (Hegazy and Mayne, 1996) and the derived normalized parameters Q and B_q are shown in Fig. D.36. Looking at q_t and Q profiles, three layers are detected and separated at depths of 4.5 m and 6 m. The upper and lower layers appear to be associated. Soil lenses are also discovered at the sharp peaks or troughs of q_t and Q within a stratum, for instance, at 5.1 m, 5.6 m, and 7.0 m. However, soil stratification is not visually apparent using either the u_b or B_q profile.

The Robertson's chart (1991) is used to classify the soil as shown in Fig. D.37. Four primary layers were detected as follows: (1) silty sand to sandy silt between the depths of 2.3 m to 4.6 m, (2) clean sand to silty sand between the depths of 4.6 m and 6 m, (3) silty sand to sandy silt between the depths of 6 m to 7.2 m and (4) clayey silt to silty clay between the depths of 7.2 m and 8.7 m. Clustering was applied to the normalized parameters Q and B_q using single-cosine-zscore method up to cluster number $N_c = 100$. The correlation coefficient between consecutive clusters is shown in Fig D.38 for cluster numbers between 2 and 50. Cluster results are examined at the peaks of ρ_c and shown in



**FigureD.36. Piezocone Sounding at Penuelas, Puerto Rico
(Data from Hegazy and Mayne, 1996).**

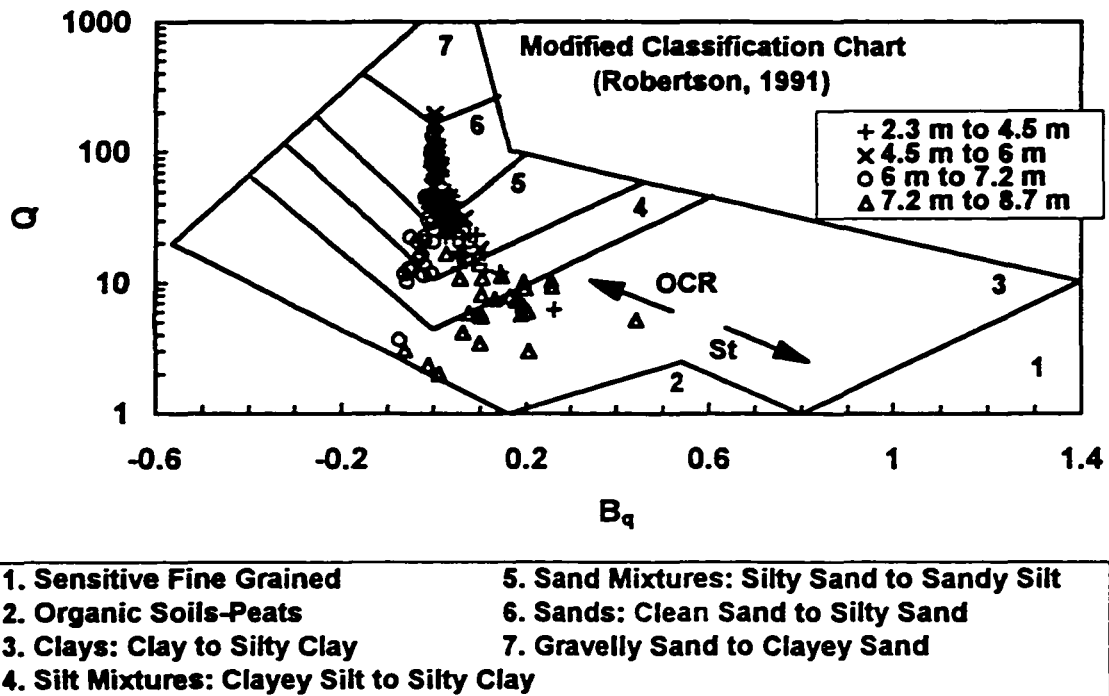
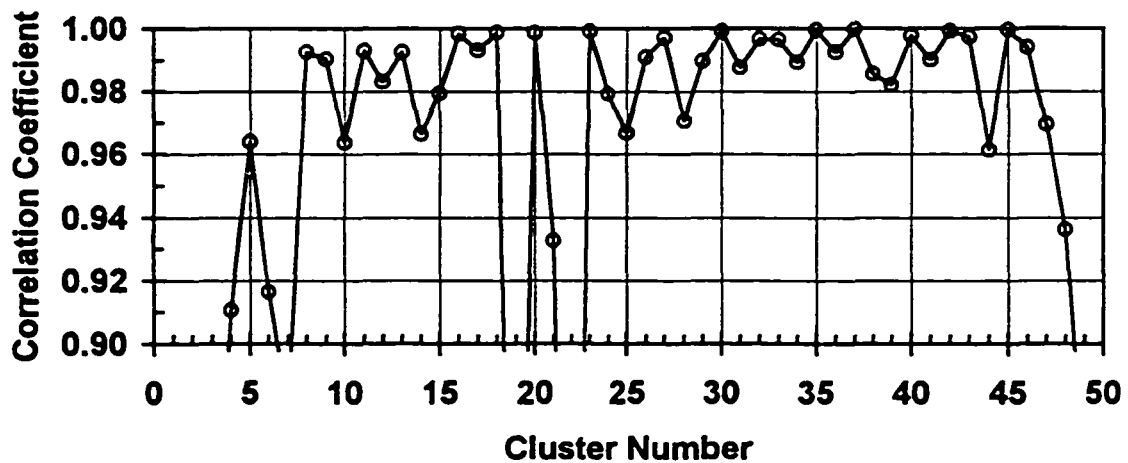


Figure D.37. Soil Classification Using Piezocone Data at Penuelas, Puerto Rico (Data from Hegazy and Mayne, 1996).



Single-Cosine-Zscore Method Using Q and B_q

Note: Piezocone data at Penuelas, Puerto Rico (Hegazy and Mayne, 1996).

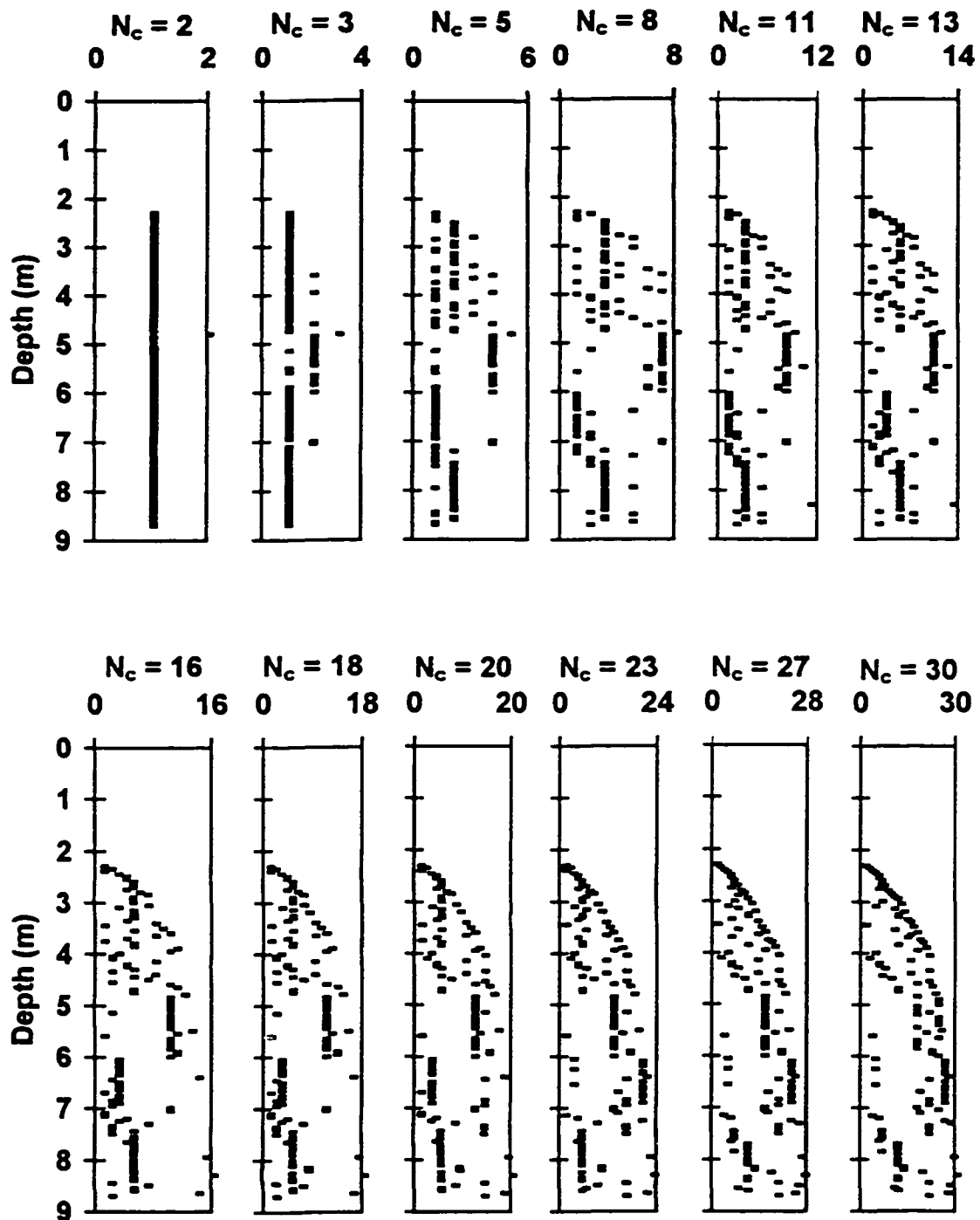
Figure D.38. Correlation Coefficient Between Consecutive Cluster Results at Penuelas, Puerto Rico.

Fig. D.39a between $N_c = 2$ and $N_c = 30$. At $N_c = 3$, three primary groups are detected and denoted by cluster numbers 1 and 2 alternatively. Soil lenses and/or data outliers are also discovered, for instance, at 3.6 m, 3.95 m and 4.8 m. For $N_c > 3$, no new primary layers ($t > 1$ m) separated from the major groups of data. This statement also is valid for clustering up to $N_c = 100$ as shown in Fig. D.39b. Therefore cluster number 3 is chosen to represent the subsurface stratigraphy. Cluster results are supported by visual definition of soil strata as shown in Fig. D.40.

D.10. Maskinongé Test Site, Québec

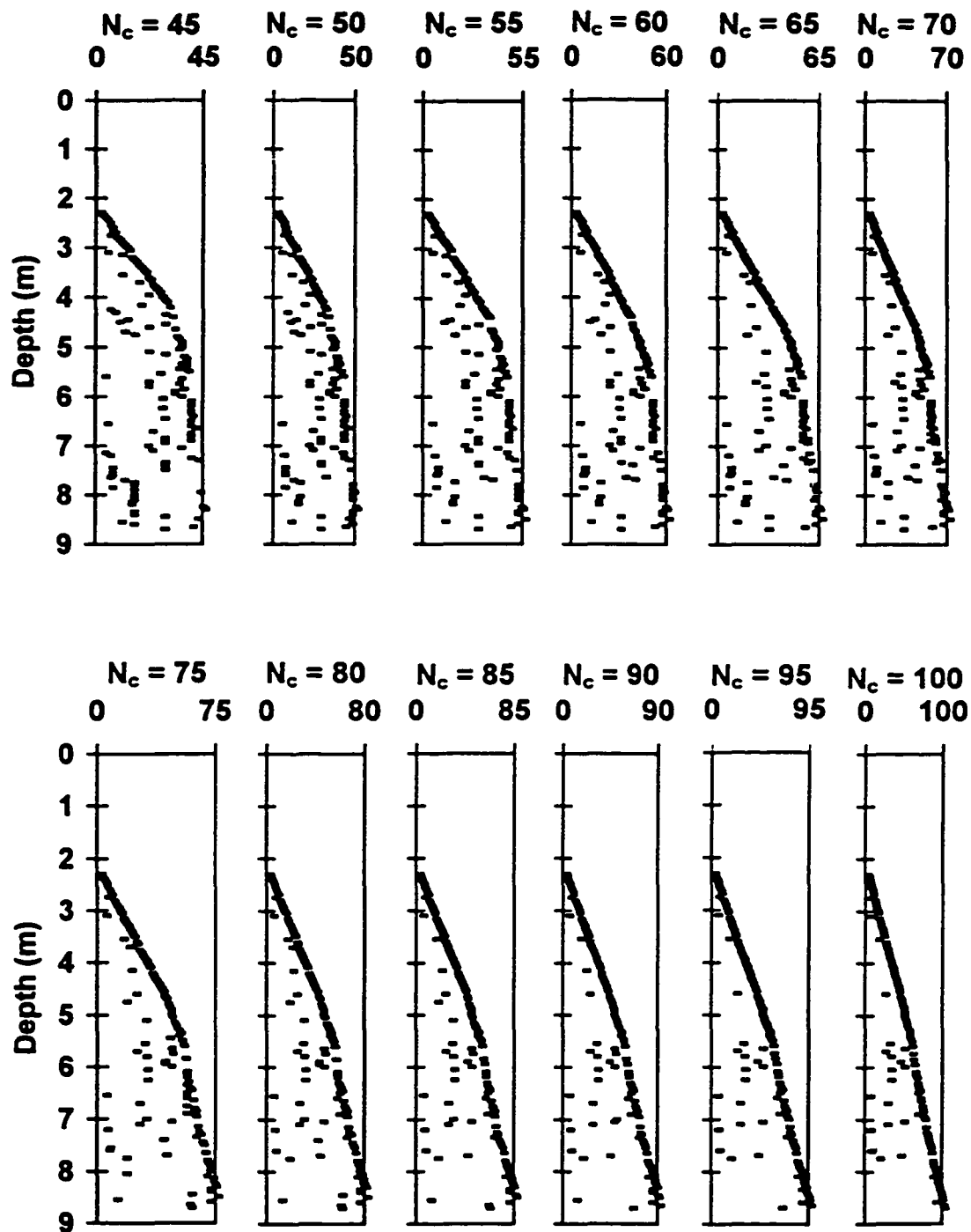
Extensive geotechnical investigation was performed to evaluate a landslide area at Maskinongé, Québec (Demers et al., 1993). The subsurface stratigraphy consists of a clay deposit of post-glacial Champion sea. The upper 2.7 m consists of loose stratified silty sand with silty and clayey layers underlain by a thick homogeneous deposit of firm gray silty clay with horizontal black bends of 1 to 2 cm thick and some fossil marine shells. The groundwater table is at a depth of 3 m. Representative piezocone sounding and the derived normalized parameters Q and B_q are shown in Fig. D.41. Two trends are observed of Q profile and delineated at 2.8 m, however, q_t , u_b and B_q profiles indicate one continuous stratum. The classification chart resulted in one clay to silty clay layer as shown in Fig. D.42.

Cluster analysis of Q and B_q is performed using single-cosine-zscore method up to cluster number $N_c = 100$. The correlation coefficient of consecutive clusters up to $N_c = 50$ is seen in Fig. D.43 and the cluster results are examined at the peaks of ρ_c . Figure D.44a



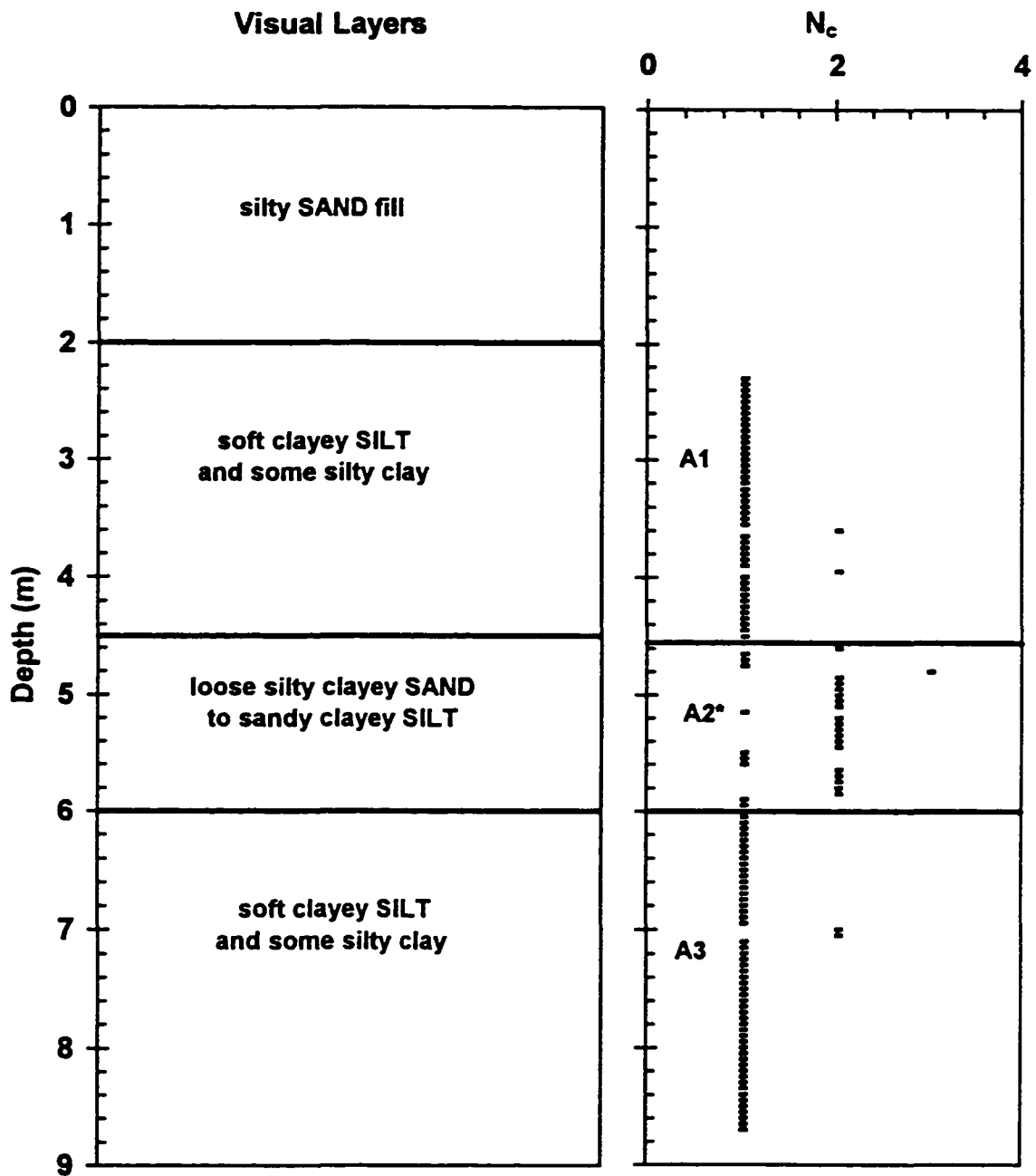
Single-Cosine-Zscore Method Using Q and B_q $N_c =$ No. of Clusters
 Note: Piezocone data at Penuelas, Puerto Rico (Hegazy and Mayne, 1996).

Figure D.39a. Cluster Analysis of Piezocone Data at Penuelas, Puerto Rico.



Single-Cosine-Zscore Method Using Q and Bq N_c = No. of Clusters
 Note: Piezocone data at Penuelas, Puerto Rico (Hegazy and Mayne, 1996).

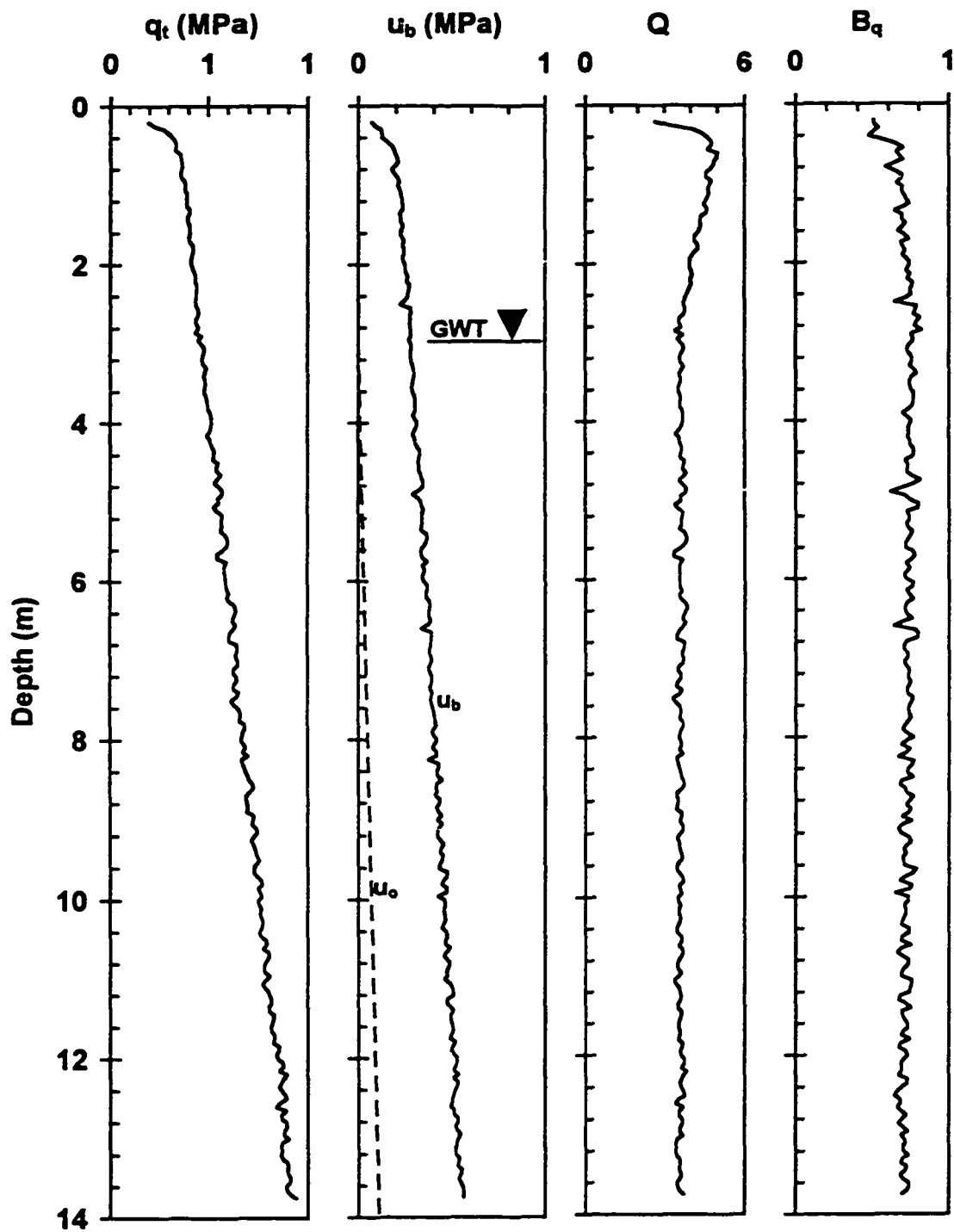
Figure D.39b. Cluster Analysis of Piezocone Data at Penuelas, Puerto Rico.



Note: Piezocone data at Penuelas, Puerto Rico (Hegazy and Mayne, 1996).
 SCZ method using Q and Bq, cluster no. 3

— Primary boundary
 — Secondary boundary

Figure D.40. Comparison Between Cluster Analysis, Visual Classification at Penuelas, Puerto Rico.



**Figure D.41. Piezocone Sounding at Maskinonge, Quebec
(Data from Demers et al., 1993).**

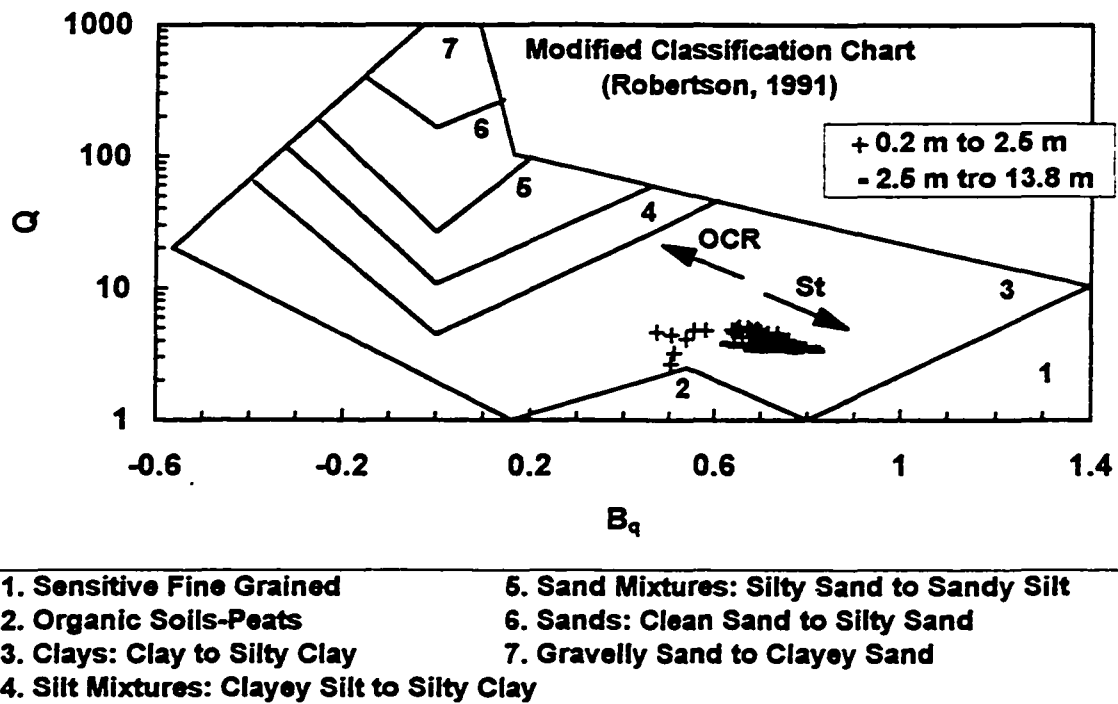


Figure D.42. Soil Classification Using Piezocone Data at Maskinonge, Quebec (Data from Demers et al., 1993).

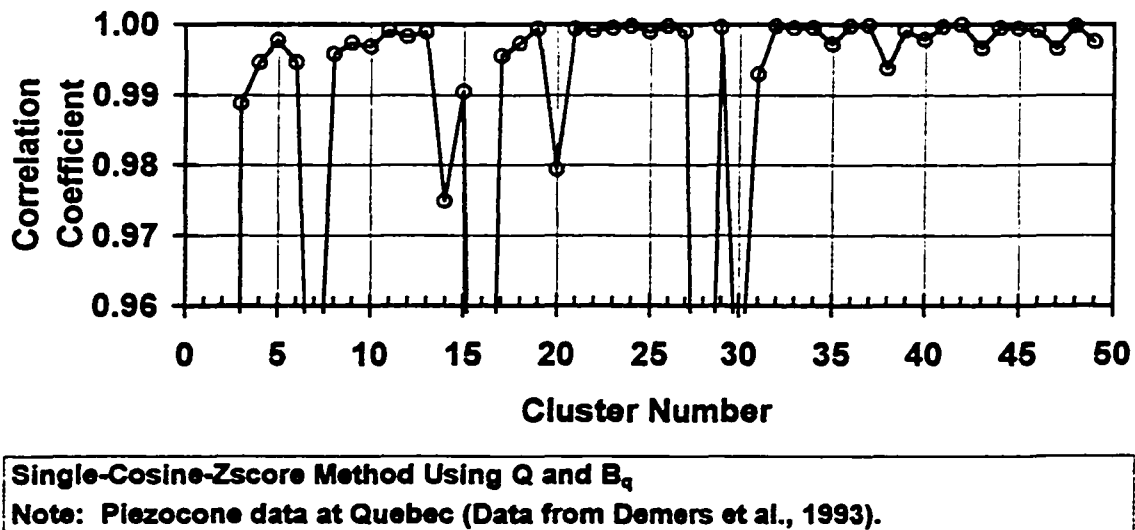
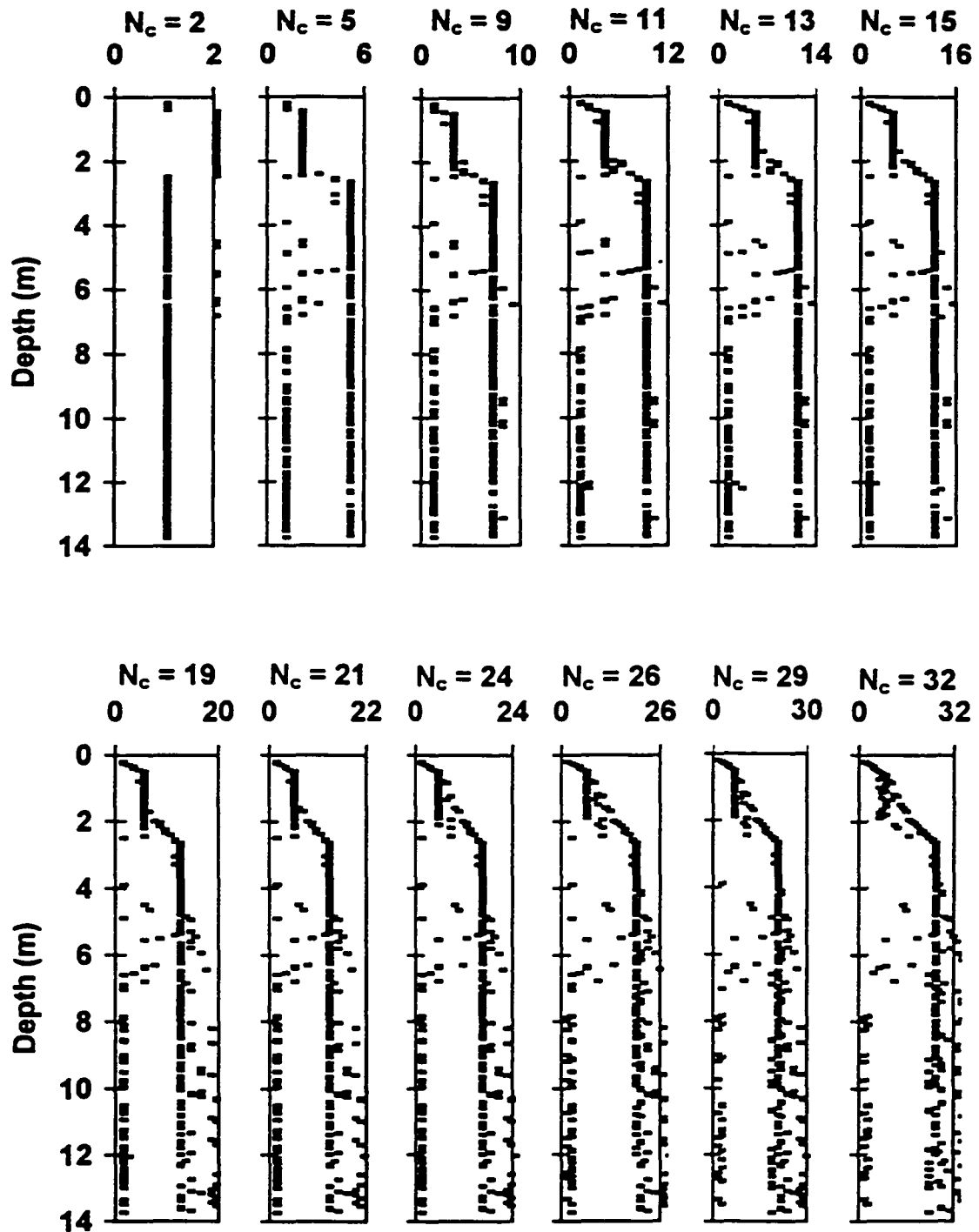


Figure D.43. Correlation Coefficient Between Consecutive Cluster Results at Maskinonge, Quebec.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Maskinonge, Quebec (Data from Demers et al., 1993).

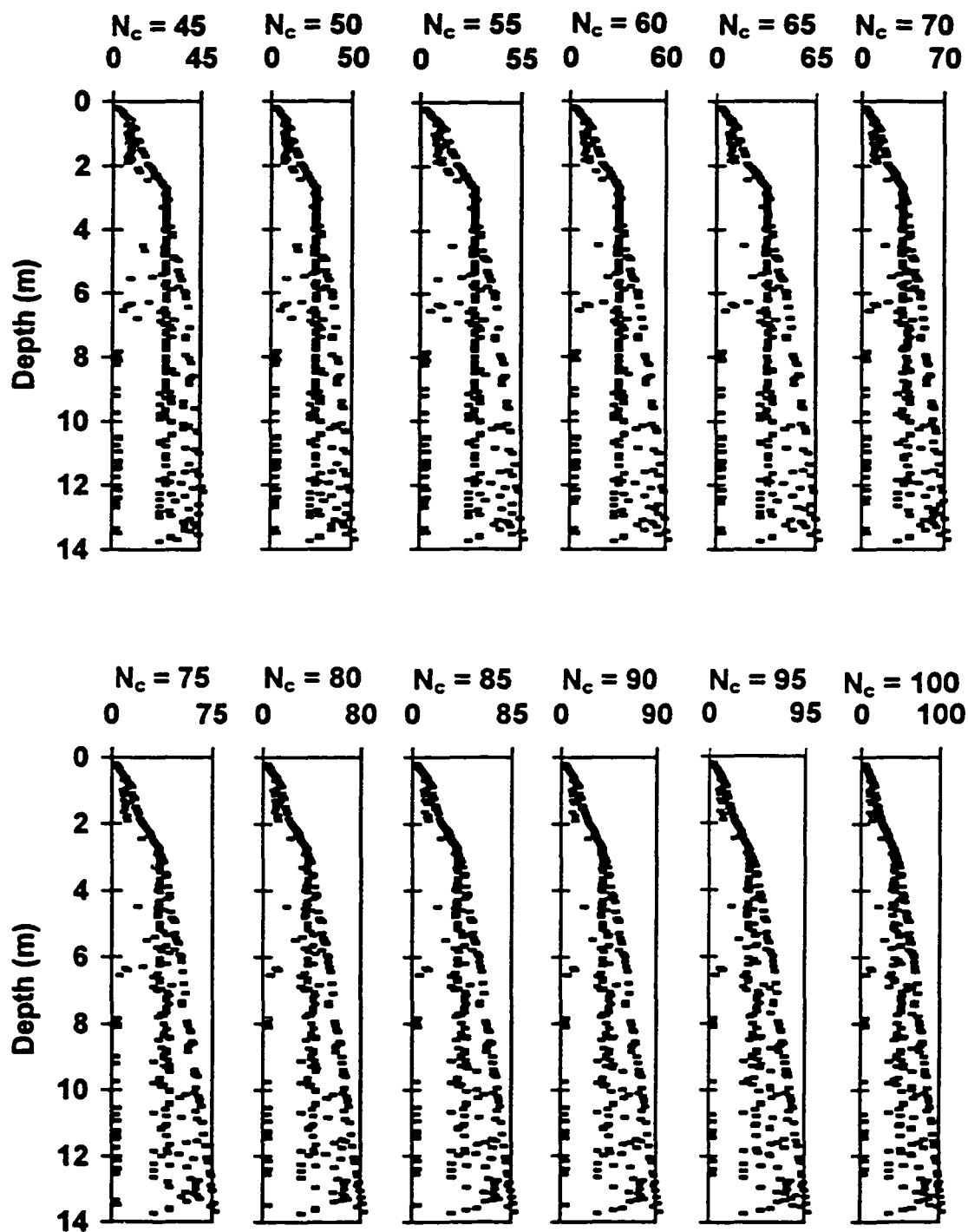
Figure D.44a. Cluster Analysis of Piezocone Data at Maskinonge, Quebec.

shows the clustering between $N_c = 2$ and 32. At $N_c = 2$, two primary clusters separated at a depth of 2.5. The upper and lower primary layers are denoted by cluster numbers 1 and 2, respectively. Soil lenses and/or data outliers also are seen in the lower deposit, for instance, at depths of 4.6 m, 5.5 m and 6.4 m. At $N_c = 5$, the lower major layer is divided into two interchangeable clusters, 1 and 5, which indicates non homogeneity within the deposit. However, no new primary layer ($t > 1$ m) separated for $N_c > 2$. This statement is valid for clustering between $N_c = 45$ and 100 as shown in Fig. D.44b. Therefore cluster number 2 is chosen to represent the subsurface stratigraphy.

The clustering results are validated by the laboratory determination of fines content and water contents measurements as seen in Fig. D.45. In the upper 2.5 m, the averages of fines content and water contents are equal to 55 percent and 27 percent, respectively. Their averages below 2.5 m increase to 99 percent and 74 percent, respectively.

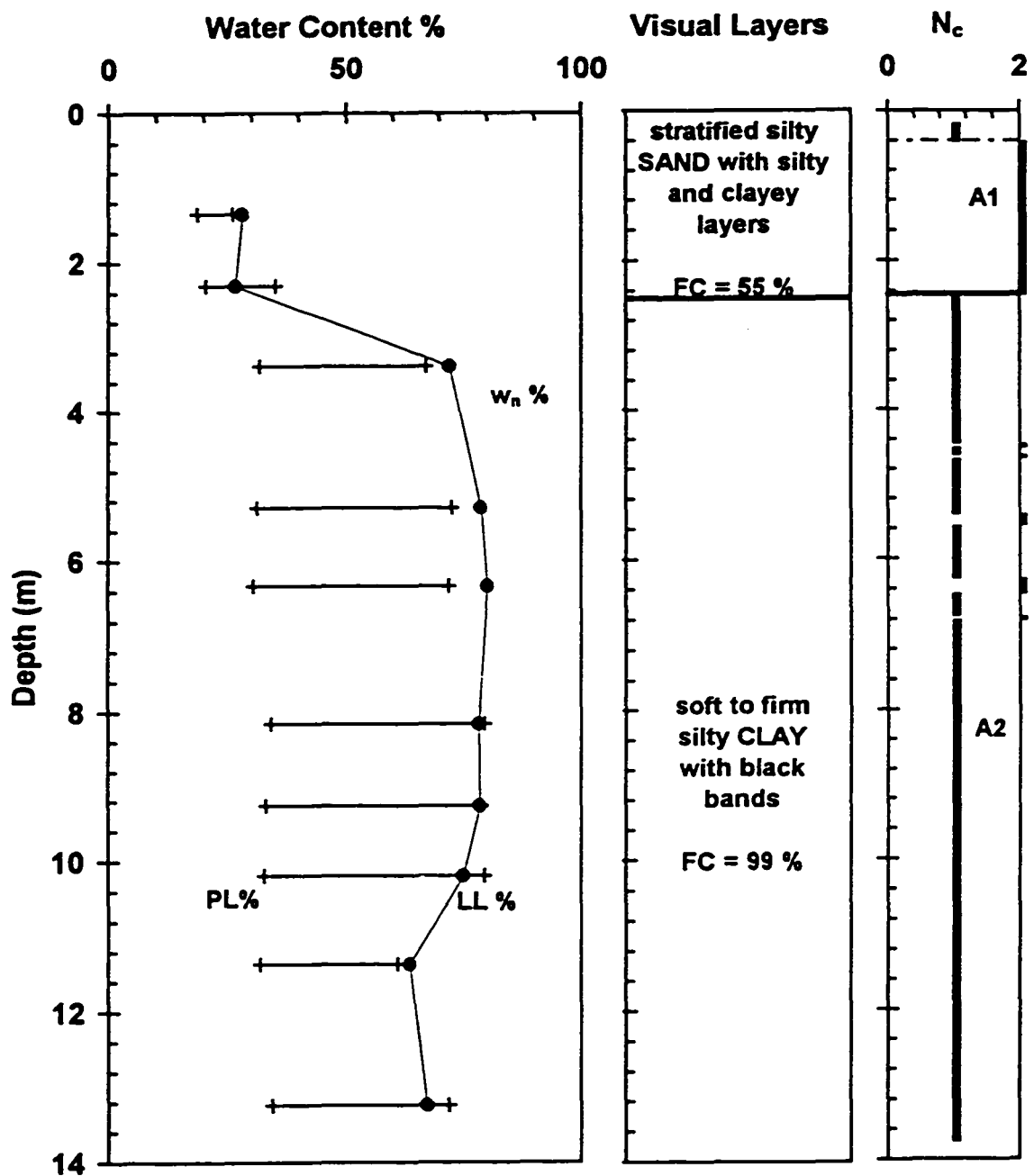
D.11. Taranto Test Site, Italy

Taranto test site is located in Italy and the stratigraphy at the site consists of highly overconsolidated cemented clay deposit with microfissures (Battaglio et al., 1986). The deposit is divided into two layers between the depths of 8 m and 20 m as follows: (1) stiff to hard weathered silty clay and (2) hard to very hard silty clay. The boundary depth between them changes between almost 10 m and 12 m. The overconsolidation ratio (OCR) decreases with depth from nearly 50 at a depth of 4 meter to 20 at a depth of 20 meter. The groundwater table is at a depth of 1 meter.



Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Maskinonge, Quebec (Data from Demers et al., 1993).

Figure D.44b. Cluster Analysis of Piezocone Data at Maskinonge, Quebec.



Note: Piezocone data at Maskinonge, Quebec (Demers et al., 1993).
 SCZ method using Q and Bq, cluster no. 2

— Primary boundary
 - - - Transition boundary

Figure D.45. Comparison Between Cluster Analysis, Visual Classification and Water Content % at

A representative piezocone sounding and the normalized parameters Q and B_q are shown in Fig. D.46. Looking at u_b and B_q profiles, a soil boundary is detected at a depth of 10 m and Q profile indicates another boundary at a depth of almost 12 m. The Robertson chart (1991) was used to delineate different soil layers as shown in Fig. D.47. Two layers were detected as follows: (1) silty sand to sandy silt between the depths of 8.2 m and 18.0 m and (2) clayey silt to silty clay between the depths of 18.0 m and 18.7.

A single-cosine-zscore method was applied to the normalized parameters Q and B_q to delineate the subsurface stratigraphy. Clustering is applied up to cluster number $N_c = 100$ and the correlation coefficient between consecutive clusters is shown in Fig. D.48 up to $N_c = 50$. Clusters results were examined at the peaks of ρ_c and shown in Fig. D.49a up to $N_c = 37$. At cluster number 2, the data are divided into two primary groups separated at 12 m. Up to $N_c = 12$, some points continue to separate from the two major statistical layers indicating non homogeneity within a layer. At cluster number 12, another primary layer separates from upper group of data. For larger clustering, no more clusters having a thickness $t > 1$ m appear up to $N_c = 100$. Figure D.49b shows the cluster results between $N_c = 45$ and 100 every 5 clusters interval. Therefore, cluster number 12 is chosen as an indication of the subsurface stratification. Also, note that the upper two layers are more associated because they are denoted by cluster numbers 4 and 5, however the lower layer is denoted by cluster number 7.

The primary statistical boundaries are in agreement with the soil boundaries defined based on laboratory and field testing as shown in Fig. D.50. Layer A1* indicates the weathered silty clay layer and layer A2 could represent a transition between A1* and

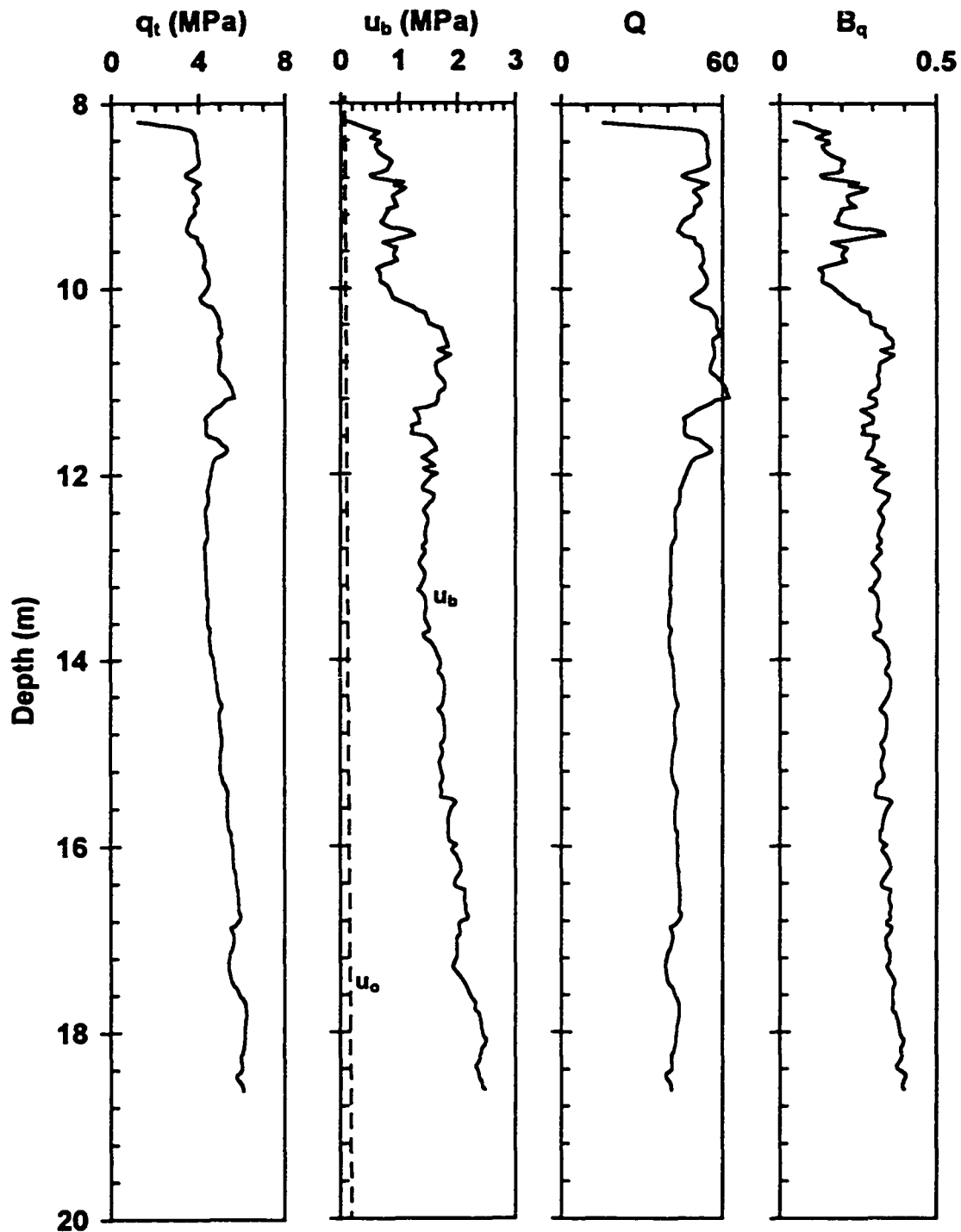
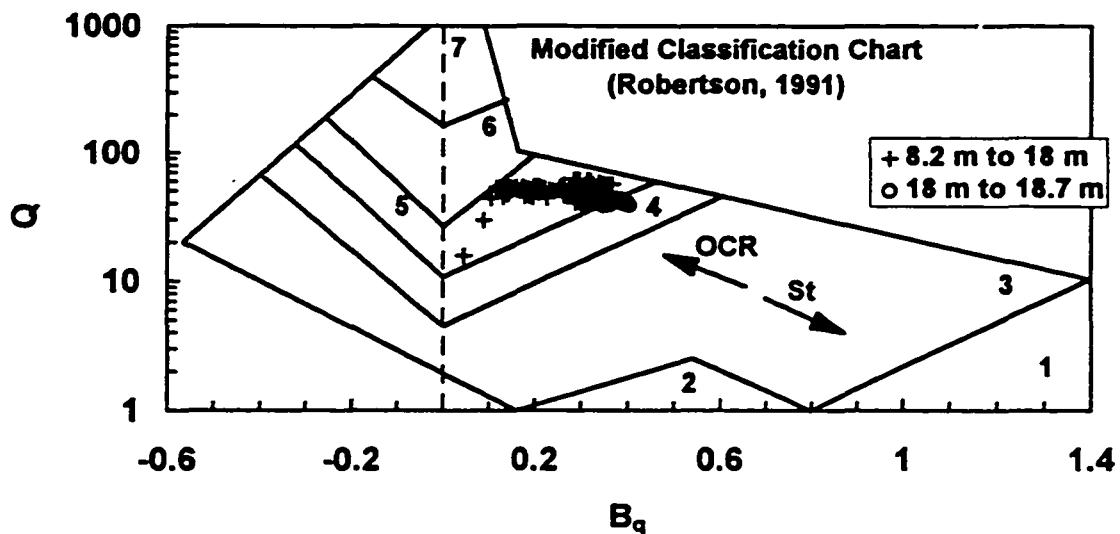
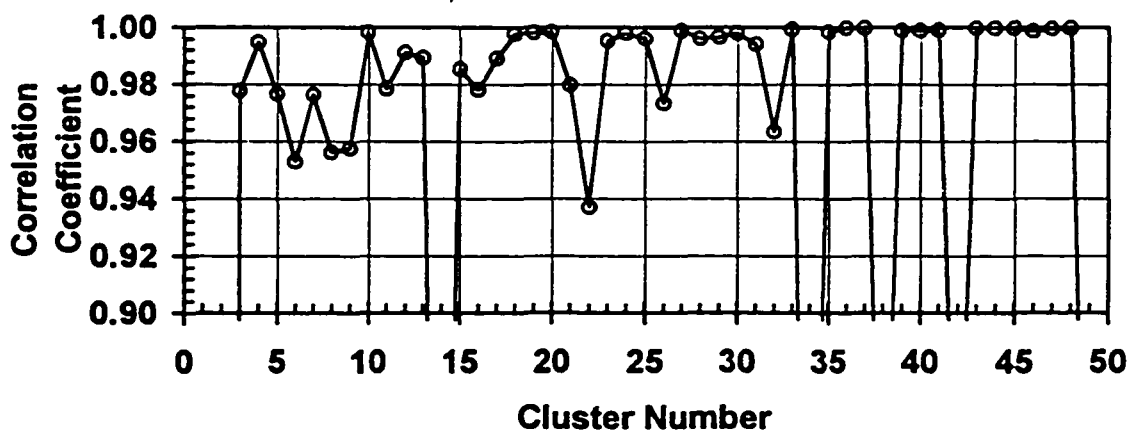


Figure D.46. Piezocone Data at Taranto, Italy (Data from Battaglio et al., 1986).



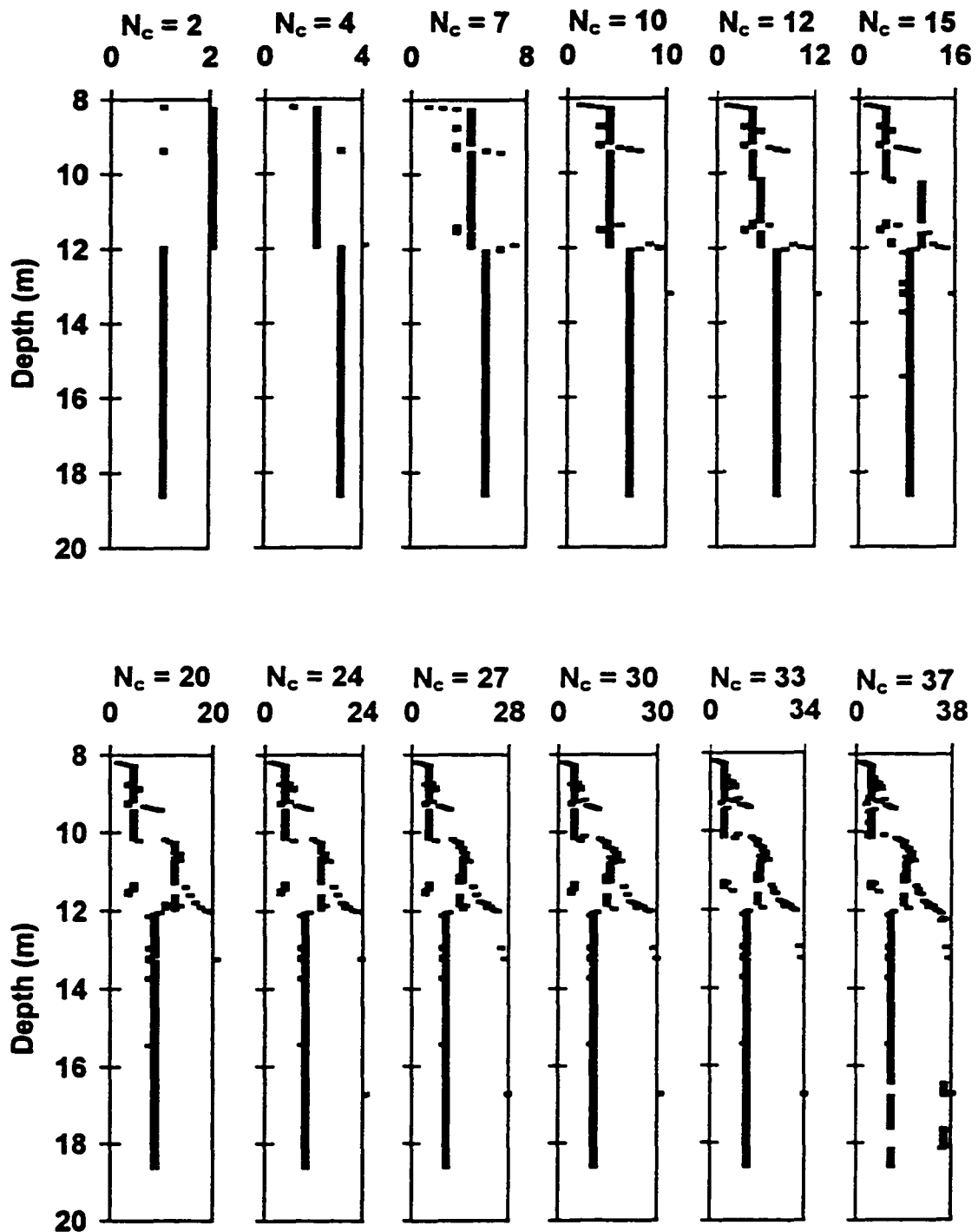
- | | |
|---|--|
| 1. Sensitive Fine Grained | 5. Sand Mixtures: Silty Sand to Sandy Silt |
| 2. Organic Soils-Peats | 6. Sands: Clean Sand to Silty Sand |
| 3. Clays: Clay to Silty Clay | 7. Gravelly Sand to Clayey Sand |
| 4. Silt Mixtures: Clayey Silt to Silty Clay | |

Figure D.47. Soil Classification at Taranto, Italy Using Piezocone Data from Battaglio et al. (1986).



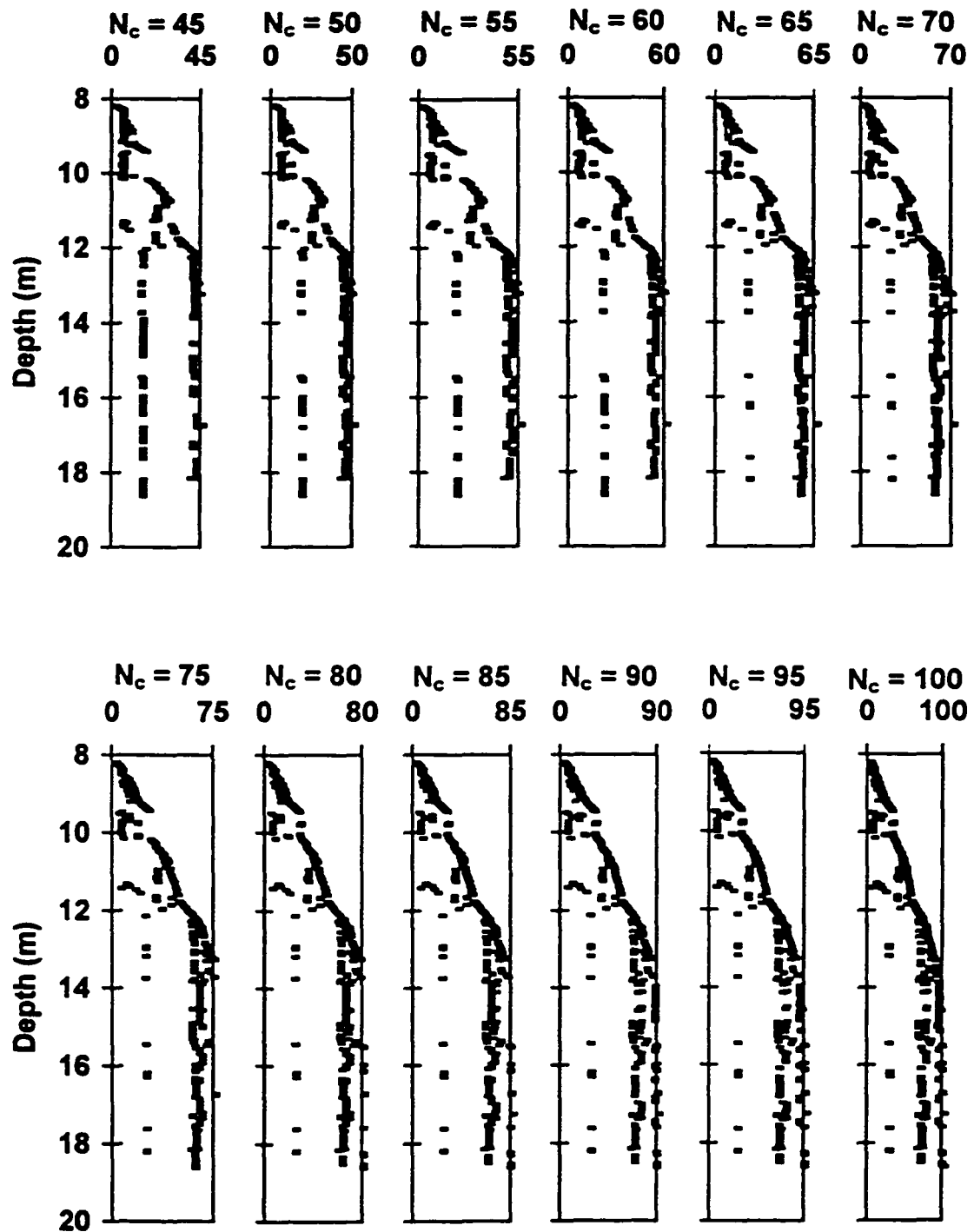
Single-Cosine-Zscore Method Using Q and B_q
Note: Piezocone data at Taranto, Italy (Battaglio et al., 1986).

Figure D.48. Correlation Coefficient Between Consecutive Cluster Results at Taranto, Italy.



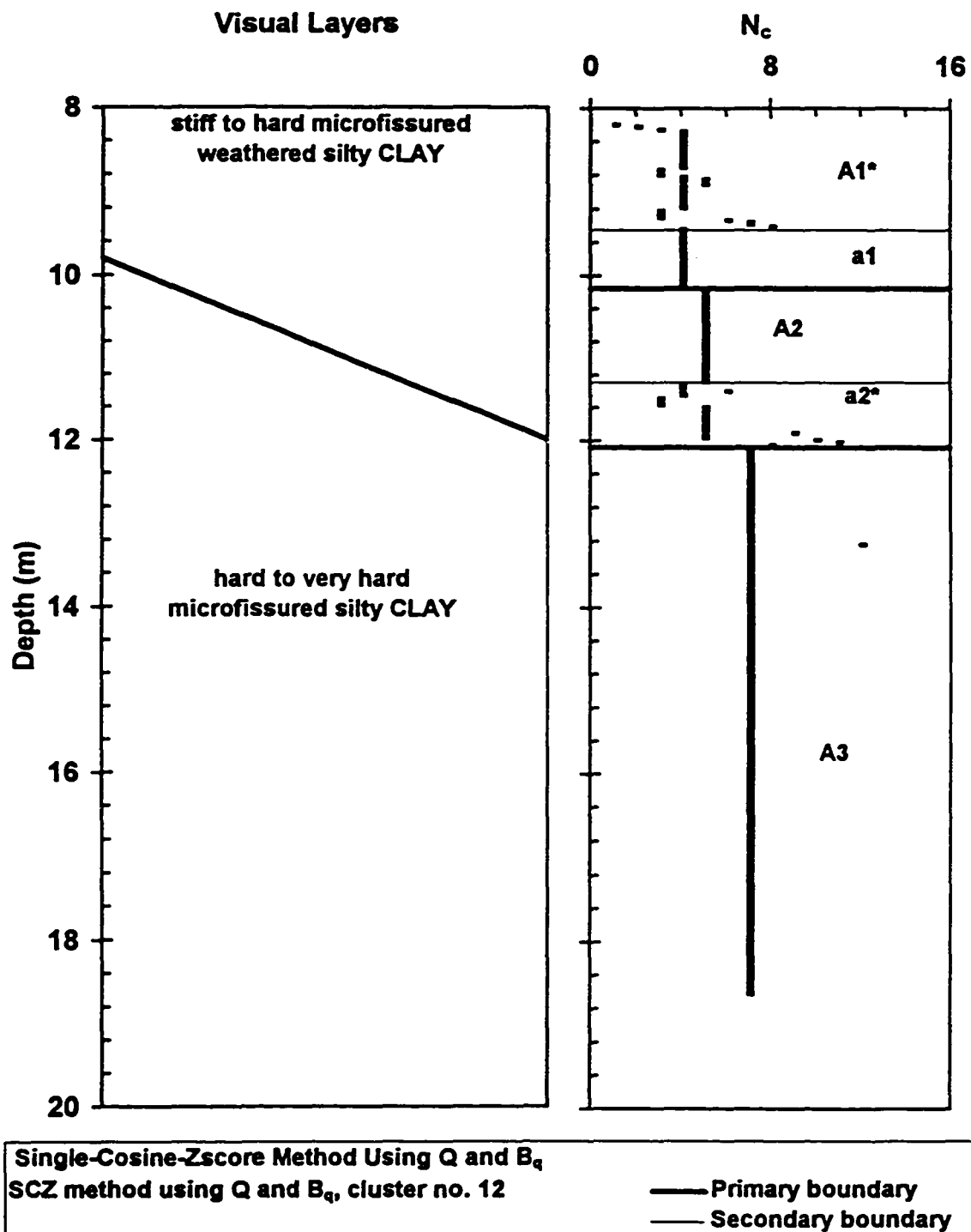
Single-Cosine-Zscore Method Using Q and B_q N_c = No. of Clusters
 Note: Piezocone data at Taranto, Italy (Battaglio et al., 1986).

Figure D.49a. Cluster Analysis of Piezocone Data at Taranto, Italy.



Single-Cosine-Zscore Method Using Q and Bq N_c = No. of Clusters
 Note: Piezocone data at Taranto, Italy (Battaglio et al., 1986).

Figure D.49b. Cluster Analysis of Piezocone Data at Taranto, Italy.



FigureD.50. Comparison Between Cluster Analysis, Visual Classification at Taranto, Italy.

A3, however it is more correlated with A1*. Layer A3 represents the silty clay layer below 12 meter.

D.12. Conclusions

Cluster analysis was applied successfully using single-cosine-zscore (SCZ) method at 10 sites including different soil types and geological conditions. Cluster results were validated by visual delineation from adjacent soil borings and available back-up laboratory and field testing data. Performing the analysis up to cluster number $N_c = 100$ was deemed to be unnecessary because subsoil stratigraphies at all studied sites were properly detected at $N_c < 15$.

REFERENCES

- American Society for Testing and Materials (1995). Standard test method for performing electronic friction cone and piezocone penetration testing of soils {D5778-95}. Annual Book of Standards, 4.08, ASTM, Philadelphia.
- Amundsen, T., Lunne T., Christophersen, H. P., Bayne, J. M. and Barnwell, C. L. (1985). Advanced deep-water soil investigation at the Troll East Field. *Proceedings of an International Conference, Offshore Site Investigation, Society for Underwater Technology*, London, UK, Vol. 3: 165-186.
- Anderberg, M. R. (1973). *Cluster analysis for applications*. Academic Press, New York: 359 p.
- Arabie, P. and Hubert, L. J. (1996). An overview of combinatorial data analysis. *Clusters and Classification*, P. Arabie, L. J. Hubert and G. De Soete (Eds.), World Scientific, New Jersey: 5-64.
- Baker, F. B. and Hubert, L. J. (1975). Measuring the power of hierarchical cluster analysis. *Journal of the American Statistical Association*, 70: 31-38.
- Battaglio, M., Bruzzi, D., Jamiolkowski, M. and Lancellotta, R. (1986). Interpretation of CPTs and CPTUs. *Proceedings of the 4th International Geotechnical Seminar*, Singapore, Nanyang Technical Institute, Singapore: 129:143.
- Begemann, H.K.S. Ph. (1965). The friction jacket cone as an aid in determining the soil profile. *Proceedings, 6th International Conference on Soil Mechanics and Foundation Engineering*, Montreal, 1: 17-20.
- Bertoline G. K. (1994). *Autocad for Engineering Graphics*. Prentice Hall, New Jersey: 750 p.
- Bjerrum, L. (1973). Problem of soil mechanics and construction of soft clays. *Proceedings of 8th International Conference on Soil Mechanics and Foundations Engineering*, Moscow, Vol. 3: 111-159.
- Bozozuk, M. and Leonards, G. A. (1972). The Gloucester test fill. *Proceedings of the ASCE Specialty Conference: Performance of Earth and Earth-Supported Structures*, West Lafayette, 1(1): 299-317.

Bratton, W. L., Timian, S. M. and Fisk, B. E. (1993). Piezo/seismic-cone penetration tests at the in-tank precipitation facility. *ARA Report* No. 5769-ITP, Applied Research Associates, Savannah river site, Aiken.

Bruzzi, D., Ghionna, V., Jamiolkowski, M., Lancellotta, R. and Manfredini, G. (1986). Self-boring pressuremeter in Po River sand. *Proceedings, The Pressuremeter and its Marine Applications: Second International Symposium, ASTM Special Technical Publication* 95, Texas A&M University: 57-74.

Burns, S. E. (1997). Development, adaptation, and interpretation of cone penetrometer sensors for geoenvironmental subsurface characterization. Doctor of Philosophy *Thesis*, Georgia Institute of Technology.

Burt, R. S. (1988). Some properties of structural equivalence measures derived from sociometric choice data. *Social Networks*, Vol. 10, 1-28.

Campanella R. G. and Robertson, P. K. (1988). Current status of the piezocone test. *Proceedings, International Symposium on Penetration Testing, ISOPT-1, Penetration Testing 1988, Orlando, Balkema, Rotterdam*, 1: 93-116.

Campanella, R. G. and Weemes, I. A. (1989). Development and use of an electrical resistivity cone for groundwater contamination studies. *Proceedings, 42 Canadian Geotechnical Conference, Winnipeg*: 1-11.

Carmichael, J. W. and Sneath, P. H. A. (1969). Taxometric Maps. *System Zoology*, 18, 402-415.

Casagrande, A. (1948). Classification and identification of soils. *Transactions, American Society of Civil Engineering, New York*, 113: 901-991.

Chandler, D. S. (1996). Monte Carlo simulation to evaluate slope stability. *Proceedings: Uncertainty in the Geologic Environment: From Theory to Practice, Geotechnical Special Publications* No.58, Madison, ASCE, New York, 1: 474-493.

Chen, B. S. and Mayne, P. W. (1994). Profiling the overconsolidation ratio of clays by piezocone tests. *Report* No. GIT-CEEGeo-94-1, Georgia Institute of Technology, Atlanta: 280 p.

Cheng-hou, Z., Greeuw, G., Jekel, J. and Rosenbrand, W. (1990). A new classification chart for soft soils using the piezocone test. *Engineering Geology*, 29: 31-47.

Christian, J. T., Ladd, C. C. and Baecher, G. B. (1994). Reliability applied to slope stability analysis. *Journal of Geotechnical Engineering*, American Society of Civil engineering, New York, 120(12): 2180-2207.

Cooper, M. C. and Milligan, G. W. (1988). The effect of measurement error on determining the number of clusters in cluster analysis. W. Gaul and M. Schader (Editors), *Data, Expert Knowledge and Decisions*, Berlin: Springer-Verlag: 319-328.

Coutinho, R. Q. and Oliveira, J. T. R. (1997). Geotechnical characterization of a Recife soft clay-Laboratory and in-situ tests. *Proceeding: XIVth International Conference on Soil Mechanics and Foundation Engineering*, Hamburg, Vol. 1, Balkema, Rotterdam: 69-72.

Demers, D. and Leroueil, S. (1993). In situ testing in a landslide area at Maskinongé, Québec. *Proceedings*, 4th Annual Canadian Geotechnical Conference, Saskatoon, Sask: 465-474.

DeRuiter, J. (1982). The static cone penetration test. State-of-the-Art Report, *Proceedings of the second European Symposium on Penetration Testing (ESOPT-II)*, Amsterdam, Balkema, Rotterdam, 2: 389-405.

Douglas, B. J. and Olsen, R. S. (1981). Soil classification using electric cone penetrometer. Cone Penetration Testing and Experience. *Proceedings of the ASCE National Convention*, St. Louis, MI: 209-227.

Duffy, D. E. and Quiroz, A. J. (1991). A permutation-based algorithm for block clustering. *Journal of Classification*, 6, 65-91.

Everitt, B. (1974). *Cluster analysis*. A Halsted Press Book, John Wiley & Sons, New York: 122 p.

Finke, K., Hegazy, Y. A. and Mayne, P. W. (1997). *Report of 2nd series of seismic piezocone tests at Ponce, Puerto Rico*, to GeoCim, Inc., San Juan, by Georgia Tech.

Fisher, L. and Van Ness, J. W. (1971). Admissible clustering procedures. *Biometrika*, 58: 91-104.

Forgey, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classification. *Biometrika*, 21: 768-769.

Fortier, J. J. and Solomon, H. (1966). Clustering procedures, *Proceedings Symposium of Multivariate Analysis*. Dayton Ohio (P. R. Krishnaian, ed.), New York: Academic Press: 493-506.

- Furmonavicius, L. and Dagys, A. (1995). Cone penetration testing in Lithuania. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 1: 125-132.
- Gengerelli, J. A. (1963). A method of detecting subgroups in a population and specifying their membership. *Journal of Psychology*, 5: 456-468.
- George, E. A. and Ajayi, L. A. (1995). National report on cone penetration testing-Nigeria. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 1: 149-162.
- Ghinelli, A. and Vannucchi, G. (1988). Statistical analysis of cone penetration testing: an application. *Proceedings, First International Symposium on Penetration Testing, Penetration Testing 1988 (ISOPT-1)*, De Ruiter (editor), Balkema, Rotterdam, 1: 757-769.
- Gordon, A. D. (1996). Hierarchical classification. *Clusters and Classification*, P. Arabie, L. J. Hubert and G. De Soete (Eds.), World Scientific, New Jersey: 65-122.
- Gordon, D. M. and Mayne, P. W. (1987). Report of Geotechnical exploration, Surry radwaste facility. Law Engineering *Report* No. W7-5768-A to Virginia Power, Richmond, Virginia.
- Gower, J. C. (1967). A comparison of some methods of cluster analysis. *Biometrika*, 23, 623-628.
- Halkola, H. and Hornqvist, J. (1995). National report, Finland. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 1: 63-66.
- Harbaugh, J. W. and Merriam, D. F. (1968). *Computer applications in stratigraphic analysis*. John Wiley and Sons, New York: 282 p.
- Harder, H. and Bloh, G. (1988). Determination of representative CPT-parameters. *Proceedings of the Penetration Testing in the UK, Geotechnology Conference*, Birmingham: 237-240.
- Hartigan, J. A. (1996). Introduction, *Clusters and Classification*, P. Arabie, L. J. Hubert and G. De Soete (Eds.), World Scientific, New Jersey: 1-3.

Hawkins, A. B., Larnach, W. J., Liloyd, M. I and Nash, D. F. T. (1989). Selecting the location, and the initial investigation of the SERC soft clay test bed. *Journal of Engineering Geology*, 22, 281-316.

Hegazy, Y. A. and Mayne, P. W. (1995). *Report of piezocone tests at Richmond, Virginia, to Virginia Geotechnical Services, P. C., Richmond, by Georgia Tech.*

Hegazy, Y. A. and Mayne, P. W. (1996). *Report of initial seismic piezocone tests at Ponce, Puerto Rico, to GeoCim, Inc., San Juan, by Georgia Tech.*

Hegazy, Y. A., Mayne, P. W and Rouhani, S. (1996). Geostatistical assessment of spatial variability in piezocone tests. *Proceedings: Uncertainty in the Geologic Environment: From Theory to Practice, Geotechnical Special Publications No.58, Madison, ASCE, New York, 1: 254-268.*

Hegazy, Y. A., Mayne, P. W and Rouhani, S. (1997). Three dimensional geostatistical evaluation of seismic piezocone data. *Proceeding: XIVth International Conference on Soil Mechanics and Foundation Engineering, Hamburg, 1, Balkema, Rotterdam: 683-686.*

Hoar, R. J. and Stohoe, K. H. (1978). Generation and measurement of shear waves in-situ. *Dynamic Geotechnical Testing. ASTM STP 654: 3-29.*

Houlsby, G. T. (1988). Piezocone penetration test. *Proceedings of the Geotechnology Conference: Penetration Tesying in the UK, Birmingham, Thomas Telford, London: 141-146.*

Hubert, L. J. and Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2: 193-218.

Hubert, L. J. and Baker, F. B. (1996). The comparison and fitting of given classification schemes. *Journal of Mathematical Psychology*, 16: 233-253.

Jain, A. K. and Dubes, R. C. (1988). *Algorithms for clustering data.* Englewood Cliffs, Prentice-Hall, New Jersey.

Jamiolkowski, M., Ladd, C. C., Germaine, J. T. and Lancellotta, R. (1985). New developments in field and laboratory testing of soils. State-of-the art report. *Proceedings of the 11th International Conference on Soil Mechanics and Foundation Engineering, San Francisco, Vol. 1, Balkema, Rotterdam: 57-153.*

Jardine, N. and Sibson, R. (1971). *Mathematical taxonomy.* Wiley, London: 286 p.

Jefferies, M. G. and Davies, M. P. (1991). Soil classification using the cone penetration test: Discussion. *Canadian Geotechnical Journal*, 28(1): 173-176.

- Jian, D., Xiaoling, Z., Longgen, Z and Lianyang, Z. (1992). Some in-situ and laboratory geotechnical test techniques used in China. *Proceedings, U.S.-China Workshop on Cooperative Research in Geotechnical Engineering*, (Du Jian and Tien H. Wu, editors), Tonji University, Shanghai and The Ohio State University, Columbus, Ohio.
- Jones, G. A. and Rust, E. (1982). Piezometer penetration testing CUPT. *Proceedings of the Second European Symposium on Penetration testing (ESOPT-II)*, Amsterdam, Balkema, Rotterdam, Vol. 2: 607-613.
- Jones, K. S. and Jackson, D. M. (1967). Current approaches to classification and clump finding at the Cambridge Language Research Unit. *Computational Journal*, 10, 29-37.
- Jorgensen, M. Denver, H. (1992). CPT-interpretation. *Proceedings, 11th Nordiske Geotechnikermode*, Aalborg, Denmark, DGF Bulletin 9.
- Journel, A. G. and Huijbregts, C. J. (1993). *Mining geostatistics*. Academic Press Inc., San Diego, California: 600 p.
- Juran, I. and Tumay, M. T. (1989). Soil stratification using the dual pore-pressure piezocone test. *Transportation Research Record*, No. 1235, National Academy Press, Washington D.C.: 68-78.
- Kates, J. (1996). Development and implementation of a seismic flat dilatometer test for small- and high-strain soil properties. Master of Science *Thesis*, Georgia Institute of Technology: 173 p.
- Kendall, M. G. (1966). *Discrimination and classification in Multivariate Analysis*, P. R. Krishnaiah, ed., Academic Press, New York.
- Kennard, R. W. and Stone, L. A. (1969). Computer aided design of experiments. *Technometrics*, 11, 137-148.
- King, B. F. (1966). Market and industry factors in stock price behavior. *Journal of Business*, Vol. 39, 139-190.
- King, B. F. (1967). Step wise clustering procedures. *Journal of American Statistical Association*, Vol. 62, 86-101.
- Kleiner, B. and Hartigan, J. A. (1981). Representing points in many dimensions by trees and castles (with comments and rejoinder). *Journal of the American Statistical Association*, 76: 260-276.

Konrad, J. M. and Law, K. T. (1987). Undrained shear strength from piezocone tests. *Canadian Geotechnical Journal*, 24: 392-405.

Lacasse, S. and Lamballerie, J. Y. N., de (1995). Statistical treatment of CPT data. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 2: 369-377.

Lacasse, S. and Lunne, T. (1982). Penetration tests in two Norwegian clays. *Proceedings of the European Symposium on Penetration Testing (ESOPT) II*, Amsterdam, Vol. 2: 661-669.

Lally, M. J. (1993). A field and laboratory investigation of geotechnical properties for design of a seasonal heat storage facility. Master of Science *Thesis*, Department of Civil and Environmental Engineering, University of Massachusetts: 223 p.

Lambson, M. and Jacobs, P. (1995). The use of the laser induced fluorescence cone for environmental investigations. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 2: 29-34.

Larsson, R. and Mulabdic, M. (1991a). Piezocone tests in clay. *Report No. 42*, Swedish Geotechnical Institute, Linköping: 240 p.

Levadoux, J. N. and Baligh, M. M. (1986). Consolidation after undrained piezocone penetration I: Prediction, *Journal of Geotechnical Engineering*, ASCE, 112(7): 707-726.

Lima, D. C. and Tumay, M. T. (1991). Scale effects in Cone Penetration tests. *Geotechnical Engineering Congress 1991*, 1, American Society of Civil Engineering (Geotechnical Special Publication 27), New York: 8-51.

Luke, K. (1994). The use of CPT in Danish soils - with special emphasis on measuring the undrained shear strength. Ph.D. *Thesis*, Soil Mechanics Laboratory Aalborg University.

Lunne, T., Eidsmoen, T. J., Gillespie, D. and Howland, J. D. (1986). Laboratory and field evaluation of cone penetrometers. Use of In Situ Tests in Geotechnical Engineering, Blacksburg, (*Proceedings, In Situ' 86*) Geotechnical Special Publication No. 6, American Society of Civil Engineering, New York: 714-729.

Lunne, T., Robertson, P. K. and Powell, J. J. M. (1997). *Cone Penetration Testing in Geotechnical Practice*, Blackie Academic and Professional, Chapman & Hall, London: 312 p.

- Lutenegger, A. J. (1995). Geotechnical behavior of overconsolidated surficial clay crusts. *Transportation Research Record*, No. 1479, National Academy Press, Washington, D. C.: 61-74.
- MacNaughton-Smith, P., William, W. T., Dale, N. B. and Mockett, L. G. (1964). Dissimilarity analysis. *Nature*, London, 202: 1034-1035.
- Mahesh, D. D. and Vikash, J. (1995). State of the art of CPT in India. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 1: 87-95.
- Marchetti, S. (1980). In-situ tests by flat dilatometer. *Journal of Geotechnical Engineering*, American Society of Civil Engineering, 107(3): 832-837.
- Marcu, A. and Culita, C. (1995). State-of-practice on CPT in Romania (national report). *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 1: 175-182.
- Masood, T. and Mitchell, J. K., Vaslestad, J., Lunne, T. and Makkelbost, K. H. (1990). Testing with lateral stress cone, special dilatometer and stepped blade at three sites in Drammen. *Report No. 521600-1*, Norwegian Geotechnical Institute, Oslo.
- Masood, T., Mitchell, J. K., Vaslestad, J., Lunne, T. and Mokkelbost, K. H. (1990). *Report No. 5216001-1*, Testing with Lateral Stress Cone, Special Dilatometer and Stepped Blade at Three Sites in Drammen, Norwegian Road Research Laboratory, University of California, Berkeley, and Norwegian Geotechnical Institute: 1-70.
- Mayne, P. W. (1989). Site Characterization of Yorktown Formation for New Accelerator. *Foundation Engineering: Current Principles* (Geotechnical Special Publication 22), American Society of Civil Engineering, New York, 1: 1-15.
- Mayne, P. W. (1991). Determination of OCR in Clays by Piezocone Tests Using Cavity Expansion and Critical State Concepts. *Soils and Foundations*, 31(2): 65-76.
- Mayne, P. W. (1991). Tentative method for estimating σ_{ho}' from q_c data in sands. *Proceedings of the International Symposium on Calibration Chamber Testing*, Potsdam, New York, 1991, Elsevier, New York: 249-256.
- Mayne, P. W., Kulhawy, F. H. and Kay, J. N. (1990). Observations on the development of pore water stresses during piezocone penetration on clays. *Canadian Geotechnical Journal*, 27(4): 418-428.

Mayne, P. W., Kulhawy, F. H. and Trautmann, C. H. (1995). Laboratory modeling of laterally-loaded drilled shafts in clays. *Journal of Geotechnical Engineering*, ASCE, 121(12): 827-835.

Mayne, P. W., Webb, D. L., Hegazy, Y. and Burns, S. (1995). *Report of piezocone tests at 10th street for the proposed pedestrian bridge, Atlanta, GA, by Georgia Tech.*

Mayne, P. W., Burns, S. E. and Hegazy, Y. A. (1994). *Report of seismic piezocone tests at Bagdad, Arizona, to AGRA Earth and Environmental, Inc., Phoenix, by Georgia Tech.*

Mayne, P. W., Burns, S. E. and Hegazy, Y. A. (1995). *Report of piezocone tests at Ponce, Puerto Rico, to Law Engineering, Houston, by Georgia Tech.*

Mayne, P. W., Burns, S. E., Hegazy, Y. A. and Kates, G. (1995). *Report of seismic piezocone, and flat dilatometer tests at Pemiscot and Dunklin Counties, Missouri, to Missouri DOT, Jefferson City, by Georgia Tech.*

McQuitty, L. L. (1960). Hierarchical syndrome analysis. *Educational and Psychological Measurement*, 20: 293-304.

Milligan G. W. and Mahajan, V. J. (1980). A note on procedures for testing the quality of a clustering of a set of objects. *Decision Sciences*, 11, 669-677.

Milligan, G. W. (1980). An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika*, 45, 325-342.

Milligan, G. W. (1981). A review of Monte Carlo tests of cluster analysis. *Multivariate Behavioral Research*, 16: 379-407.

Milligan, G. W. (1996). Clustering validation: results and implications for applied analyses. *Clusters and Classification*, P. Arabie, L. J. Hubert and G. De Soete (Eds.), World Scientific, New Jersey: 341-375.

Milligan, G. W. and Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50, 159-179.

Milligan, G. W. and Cooper, M. C. (1988). A study of variable standardization. *Journal of Classification*, 5, 181-204.

Mitchell, J. K. and Brandon, T. L. (1998). Analysis and Use of CPT in Earthquake Environmental Engineering. *Proceedings, First International Conference on Site Characterization (ISC '98)*, Atlanta, Balkema, Rotterdam, in print.

Mlynarek, Z. and Lunne, T. (1987). Statistical estimation of homogeneity of a North Sea overconsolidated clay. *Proceedings, 5th International Conference on Application of Statistics and Probability in soil and Structural Engineering (ICASP5)*, Vancouver, 2: 961-968.

Morey, L. C., Blashfield, R. K. and Skinner, H. A. (1983). A comparison of cluster analysis technique within a sequential validation framework. *Multivariate Behavioral research*, 18: 309-329.

Najjar, Y. M. and Basheer, I. A. (1996). A neural network approach for site characterization and uncertainty prediction. *Proceedings: Uncertainty in the Geologic Environment: From Theory to Practice, Geotechnical Special Publications No.58*, Madison, ASCE, New York, 1: 134-148.

Nash, D. F. T., Powell, J. J. M. and Liroyd, I. M. (1992). Initial investigation of the soft clay test site at Bothkennar. *Geotechnique*, 42(2): 163-181.

Neter, J., Wasserman, W. and Kutner, M. H. (1990). *Applied Linear Statistical Models: Regression Analysis of Variance, and Experimental Designs*. Richard D. Irwin, Inc., Boston, Massachusetts: 1012 p.

Norusis, M. J. (1993). *SPSS professional statistics 6.1*. SPSS Inc., p 385.

Olsen, R. S. (1994). Normalization and prediction of geotechnical properties using the cone penetrometer test (CPT). *Report No. GL-94-29*, US Army Corps of Engineers and Waterways Experiment Station, Washington, DC: 292 p.

Olsen, R. S. and Malone, P. (1988). Soil classification and site characterization using the cone penetrometer test. *Proceedings, First International Symposium on Penetration Testing (ISOPT 1)*, Orlando.

Olsen, R. S. and Mitchell, J. K. (1995). CPT stress normalization and prediction of soil classification. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, Vol. 2: 257-262.

Parker-Rhodes, A. F. and Jackson, D. M. (1969). Automatic classification the ecology of higher fungi. In *Numerical Taxonomy*. (A. J. Cole, ed.), 181-215. New York: Academic Press.

Peuchen, J., Harteveld, J., Golightly, C. R. and Haland, G. (1995). Tunnel route investigation for Euopipe landfall. *Proceedings, Eleventh European Conference on Soil Mechanics and Foundation Engineering*. Copenhagen, 8: 131-136.

Post, M. L. and Nebbeling, H. (1995). Uncertainties in cone penetration testing. *Proceedings of the International Symposium on Cone Penetration Testing, CPT'95*, Linköping, Sweden, Swedish Geotechnical Society, Vol. 2: 73-78.

Powell, J. J. M. and Quarterman, R. S. T. (1988). The interpretation of cone penetration tests in clays, with particular reference to rate effects. *Proceedings of the International Symposium on Penetration Testing, ISPT-1*, Orlando, Vol. 2, Balkema, Rotterdam: 903-910.

Powell, J. J. M., Quarterman, R. S. T. and Lunne, T. (1988). Interpretation and use of the piezocone test in UK clays. *Proceedings of the Geotechnology Conference: Penetration Testing in the UK*, Birmingham, Thomas, Telford, London: 151-156.

Raschke, S. A. and Hryciw, R. D. (1997). Viscon cone penetration (Vis CPT) for direct subsurface soil observation. *Journal of Geotechnical and Geoenvironmental Engineering*, American Society of Civil engineering, New York, 23(11).

Robertson, P. K. (1982). In-situ testing of soil with emphasis on its application to liquefaction assessment. Doctor of Philosophy *Thesis*, Department of Civil Engineering, University of British Columbia: 395 p.

Robertson, P. K. (1990). Soil classification using the cone penetration test. *Canadian Geotechnical Journal*, 27(1): 151-158.

Robertson, P. K. (1991). Soil classification using the cone penetration test: Reply. *Canadian Geotechnical Journal*, 28(1): 176-178.

Robertson, P. K. and Campanella, R. G. (1983b). Interpretation of cone penetration test: Part II: Clay. *Canadian Geotechnical Journal*, 20(4): 734-745.

Robertson, P. K. and Campanella, R. G. (1984). *Guidelines for Use and Interpretation of other Electronic Cone Penetration Test*, Hogentogler & Company, Inc.

Robertson, P. K. and Campanella, R. G. (1989). *Guidelines for geotechnical design using CPT and CPTU*. Soil Mechanics Series No. 120, Department of Civil Engineering, University of British Columbia: 193 p.

Robertson, P. K., Campanella, R. G., Gillespie, D. and Greig, J. (1986). Use of piezometer cone data. Use of In Situ Tests in Geotechnical Engineering, Blacksburg, (Proceedings, In Situ' 86) Geotechnical Special Publication No. 6, American Society of Civil Engineering, New York 1263-1280.

Robertson, P. K., Lunne, T. and Powell, J. J. M. (1998). Geo-environmental applications of penetration testing. *Proceedings, First International Conference on Site Characterization (ISC '98)*, Atlanta, Balkema, Rotterdam, in print.

Romesburg, H. C. (1984). *Cluster analysis for researchers*. Wadsworth, Inc., Lifetime Learning Publications, Belmont, California, 334 p.

Roy, M., Tremblay, M., Tavenas, F. and La Rochelle, P. (1982). Development of a quasi-static piezocone apparatus. *Canadian Geotechnical Journal*, 19(2): 180-188.

Sabri, M. M. and Dakhli, A. H. (1995). National report for Egypt. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 3: 141-146.

Schmertmann, J. H. (1978). Guidelines for cone penetration test, performance and design. *Report TS-78-209*, Federal Highway Administration, Washington, D. C.: 145 p.

Scott, A., J. and Symons, M. J. (1971a). On the Edwards and Cavalli-Sofza method of cluster analysis. *Biometrics*, 27, 217-219.

Senneset, K. and Janbu, N. (1984). Shear strength parameters obtained from static cone penetration tests. *American Society of Testing and Material, STP 883*, Symposium, San Diego.

Senneset, K. and Janbu, N. (1985). Shear strength parameters obtained from static cone penetration tests. *Strength Testing of Marine Sediments, Laboratory and In Situ Measurements*. Symposium, San Diego, ASTM Special Technical Publication, STP 883: 41-45.

Senneset, K., Sandven, R. and Janbu, N. (1989). The evaluation of soil parameters from piezocone tests. *Transportation Research Record*, No. 1235, National Academy Press: 24-37.

Senneset, K., Sandven, R., Lunne, T., By, T. and Amundsen, T. (1988). Piezocone tests in silty soils. *Proceedings of the International Symposium on Penetration Testing (ISOPT-1)* Orlando, Balkema, Rotterdam, Vol. 2: 955-966.

Shinn, J. D., Hull, R. L., Timian, D. A. and Morey, R. M. (1998). Development of a CPT probe to determine volumetric soil moisture content. *Proceedings of International Conference on Site Characterization (ISC '98)*, Atlanta, Balkema, in print.

Sibson, R. (1972). Order invariant methods for data analysis. *Journal of statistical Society, Series B*, 34, 311-349.

- Sneath, P. H. A. (1957). The application of computers to taxonomy. *Journal of General Microbiology*, 17: 201-226.
- Sokal, R. R. (1974). Classification: purposes, principles, progress, prospects. *Science*, 185, 1115-1123.
- Sokal, R. R. and Mickener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, 38: 1409-1438.
- Sokal, R. R. and Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *Taxon*, 11, 33-40.
- Solberg, H. E., Strede, S., Elgjo, K., Bolmhoff, J. P., and Gjone, E. (1976). Classification of liver diseases by clinical chemical laboratory results and cluster analysis. *The Scandinavian Journal of Clinical and Laboratory Investigation*, 36: 81-85
- Sowers, G. F. and Richardson, T. L. (1983). Residual soils of Piedmont and Blue Ridges. *Transportation Research Record*, No. 919, Washington, D. C.: 10-16.
- Stiensra, P. and van Deer, J. K. (1994). Field data collection techniques-unconventional sounding and sampling techniques. *Proceedings*, Engineering Geology of Quaternary Sediments, 20-year Jubilee Symposium of the Ingeokring, N. Rengers (editor), Balkema, Rotterdam: 41-56.
- Sweeney, B. P. and Kraemer, S. R. (1993). *Report* on special laboratory and in situ testing program at Central Artery (I-93)/Tunnel (I-90), Haley and Aldrich Inc., Boston, Massachusetts.
- Takesue, K., Hikari, S. and Makiyara, Y. (1996). Cone penetration testing in volcanic soil deposits. *Proceedings* of an International Conference, Advances in Site Investigation Practice, Thomas Telford, London: 452-463.
- Tanaka, H. (1995). National Report - the current state of CPT in Japan. *Proceedings* of the International Symposium on Cone Penetration Testing, CPT' 95, Linköping, Sweden, Swedish Geotechnical Society, Vol. 1: 115-124.
- Tanaka, H., Sakakibaea, Goto, K., Susuki, K. and Fukazawa, T. (1992). *Report*, Properties of Japanese normally consolidated marine clays obtained from static piezocone penetration test, The Port and Harbor Research Institute, Ministry of Transport, Nagase, Yokosuka, Japan, 31(4): 64-92.

Terzaghi, K., Peck, R. B. and Mesri, G. (1996). *Soil Mechanics in Engineering Practice*. A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York: 549 p.

Tien, N. T. (1995). Cone penetration testing in Vietnam. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 1: 277-281.

Treadwell, D. D. (1976). The influence of gravity, prestress, compressibility and layering on soil resistance to static penetration. Doctor of Philosophy *Thesis*, University of California, Berkeley.

Trofimenkov, Y. G., Yulachkin, B. I., Mariupolsky, L. G. and Ryzhkhov, I. B. (1995). Cone penetration testing in Russia. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 1: 183-192.

Vanmarcke, E. H. (1977). Probabilistic modeling of soil profiles. *Journal of Geotechnical Engineering*, American Society of Civil Engineering, 103(11): 1227-1246.

Vermeulen, N. and Rust, E. (1995). CPTU profiling: a numerical method. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 2: 343-350.

Vidic, S. D., Beckwith, G. H. and Mayne, P. W. (1995). Profiling mine tailings with CPT. *Proceedings of the International Symposium on Cone Penetration Testing, CPT' 95*, Linköping, Sweden, Swedish Geotechnical Society, 2: 607-612.

Vinson, J. L. and Brown, D. A. (1997). Site characterization of the Spring Villa geotechnical test site and a comparison of strength and stiffness parameters for a Piedmont residual soil. *Report No. IR-97-04*, Highway Research Center, Herbert Engineering Center, Auburn University, Alabama: 385 p.

Villet, W. C. B. and Mitchell, J. K. (1981). Cone resistance, relative density and friction angle. *Cone Penetration Testing and Experience*; Session at the ASCE National Convention, St. Louis: 178-207.

Vivitrat, V. (1978). Cone penetration in clays. Doctor of Philosophy *Thesis*, Department of Civil Engineering, Massachusetts Institute of Technology: 427 p.

Vreugdenhil, R., Davis, R. and Berrill, J. (1994). Interpretation of cone penetration results in multilayered soils. *International Journal for Numerical and Analytical Methods in Geomechanics*, Vol. 18: 585-599.

Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58: 236-244.

Webster, R. (1973). "Automatic soil boundary location from transect data", *Journal of Mathematical Geology*, Vol. 5, No. 1: 27-37.

Wickremesinghe, D. S. (1989). Statistical characterization of soil profiles using in-situ tests. Doctor of Philosophy *Thesis*, Department of Civil Engineering, University of British Columbia: 288 p.

Wickremesinghe, D. S. and Campanella, R. G. (1991). "Statistical methods for soil layer boundary location using the cone penetration test", *Proceedings: 6th International Conference on Applications of Statistics and Probability in Civil Engineering*, Vol. 2, Mexico City: 636-643.

Wroth, C. P. (1984). The interpretation of in situ soil test. 24th Rankine Lecture, *Geotechnique*, 34(4): 449-489.

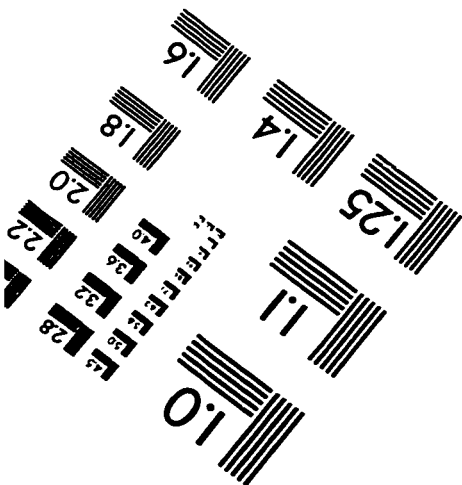
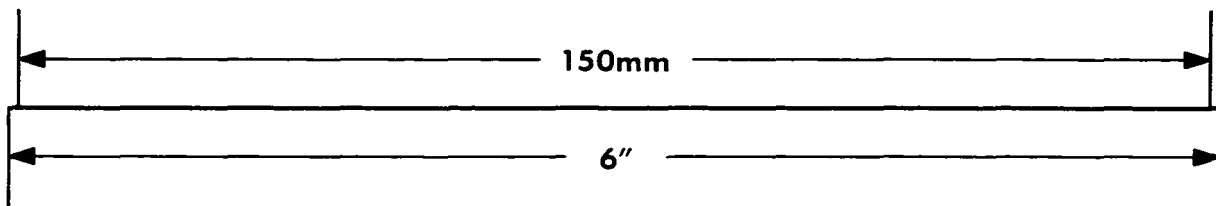
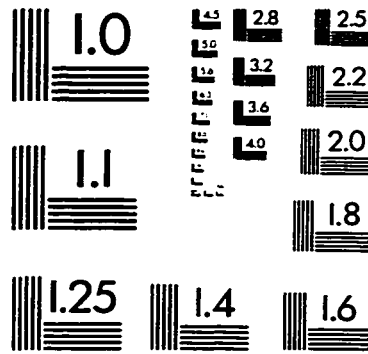
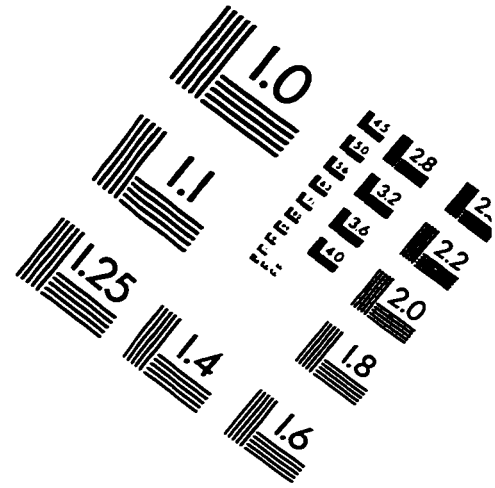
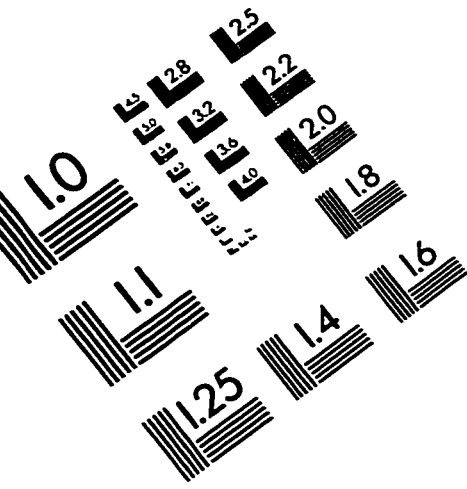
Wroth, C. P. (1988). Penetration testing-a more rigorous approach to interpretation. *Proceedings of the International Symposium on Penetration testing (ISOPT-1)*, Orlando, Balkema, Rotterdam, Vol. 1: 303-311.

Zhang, Z. (1994). Use of uncertainty methodology in identification and classification of soils based on CPT. Doctor of Philosophy *Thesis*, Department of Civil Engineering, Louisiana State University: 199 p.

Zhang, Z. and Tumay, M. T. (1996a). "The reliability of soil classification derived from cone penetration test", *Proceedings: Uncertainty in the Geologic Environment: From Theory to Practice*, Geotechnical Special Publications No.58, Madison, ASCE, New York, Vol. 1: 383-408.

Zhang, Z. and Tumay, M. T. (1996b). Simplification of soil classification charts derived from the cone penetration test. *Geotechnical Testing Journal*, ASTM, 19(2): 203-216.

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc.
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc.. All Rights Reserved

