NTNU – Trondheim
Norwegian University of
Science and Technology

Interpretation methods of CPTU and RCPTU with special focus on soft soils

Assessment between classical approach and
data mining techniques

Stanislaw Puakowski

Geotechnics and Geohazards
Submission date: June 2015
Supervisor: Steinar Nordal, BAT
Co-supervisor: Rolf Sandven, Multiconsult

Norwegian University of Science and Technology
Department of Civil and Transport Engineering

NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF CIVIL AND TRANSPORT ENGINEERING

Report Title: Interpretation methods of CPTU and RCPTU with special focus on soft soils in Norway	Date: 17th June 2015 Number of pages (incl. appendices): 192			
	Master Thesis	X	Project Work	
Name: Stanislaw Puakowski				
Professor in charge/supervisor: Professor Steinar Nordal, NTNU. Department of Transport and Civil Engineering				
Other external professional contacts/supervisors: Dr Rafal Ossowski, Gdansk University of Technology, Department of Civil Engineering Professor Rolf Sandven, Multikonsult				

Abstract:
The purpose of the following work is to explain numerous methods of interpretation of CPT (Cone Penetration Test) and to find the most practical one for the soft soils. Due to the complex nature of the problem and strict requirement of credible results, following parameters have been taken into account: pore pressure measurement (CPTU) and resistivity (RCPTU). Clay deposits are frequently found on numerous construction sites in Norway. Presence of the weak clays or quick clays tremendously increases possibility of a failure due to extremely weak and unstable structure of particles. The results obtained from the soundings have been compared to the laboratory investigations on proper samples, which have given leading parameters of the soils. Data mining models were constructed on the basis of four databases from three different investigation sites. With soundings readings and laboratory results combined some hidden correlations are seek between them. In addition, a background of CTP technology including history, development and possible applications in civil engineering is presented in the thesis.

Keywords:

1. Cone Penetration Test
2. Soft soil
3. Interpretation
4. Data mining

Preface

Geotechnical investigation is essential in every advanced project in civil engineering. The primary objective is to ensure stability of a object during the construction stage and in a long-term perspective. To achieve that, an *in situ* geotechnical survey must be conducted. A necessary condition for a successful design is to obtain high-quality data, which will allow to assess soil conditions and setting of layers in a precise manner.
The core issue is the lack of ability to classify quick clay by direct methods - in this case by Cone Pressuremeter Tests. To distinguish quick clay deposits a sampling and laboratory research is required or futher *in situ* tests. Such solution generates additional costs, requires more time and labour.
The purpose of the following work is to explain numerous methods of interpretation of CPT (Cone Penetration Test) and to find the most practical one for soft soils. Due to the complex nature of the problem and the needs of acquiring credible results following parameters have been taken into account: Pore Pressure Measurement (CPTU) and Resistivity Module (RCPTU). Clay deposits are frequently found on construction sites in Norway, which should be dealt with caution and care. Presence of the weak clays or quick clays tremendously increases possibility of a failure due to extremely weak and unstable structure of particles. The results obtained from the soundings have been compared to the laboratory investigations on proper samples, which have given leading parameters of the soils. Data mining models were constructed on the basis of four databases from three different investigation sites. With soundings readings and laboratory results combined some hidden correlations are seek between them. In addition, a background of CTP technology including history, development and possible applications in civil engineering is presented in the thesis.

Acknowledgements

During my Norwegian exchange, which lasted for two semesters, I have encountered many supportive people, without whom I could not finish this thesis. By sharing knowledge, experience and kindness they kept me motivated and focused on the task at hand.
I would like to express my gratitude to Professor Steinar Nordal for guidance and forbearance - with his invaluable support I managed to face many scientific challenges. I am also very grateful for Professor Rolf Sandven, who was so kind and provided necessary data for this project. Without his input I wouldn't be able to make any progress at all.
The practical part of this Master's Thesis wouldn't be possible without NTNU personnel, who participated in *in situ* tests in Tiller investigation site. Thus, I thank Jønland Jan and Winther Gunnar - engineers from the Department of Civil and Transport Engineering. Obtained results were necessary for completing the database, which was crucial in my further research.
Moreover, I want to give my thanks to staff from Gdansk University of Technology, Department of Geotechnics, Geology and Maritime Engineering: Professor Lech Bałachowski for supplying me with geotechnical papers and articles, PhD Rafał Ossowski for supervising my progress over data mining.
Finally, I wish to memorialize supervisor of my bachelor's thesis, the Head of Department Professor Zbigniew Sikora, who has drawn His last breath on 12th of May 2015. He was a great tutor and mentor for many generations of engineers. It is a great honor to be one of them.

Summary

The purpose of the following work is to explain numerous methods of interpretation of CPT (Cone Penetration Test) and to find the most practical one for soft soils. Due to the complex nature of the problem and the needs of acquiring credible results following parameters have been taken into account: pore pressure measurement (CPTU) and resistivity (RCPTU).
The present master's thesis focuses on comparison of post-processed results from different interpretation methods with laboratory data. In the beginning most popular approaches for sensitive soils are presented: soil classification charts, undrained shear strength, sensitivity and resistivity measurements from RCPTu.
Further, an experimental method of machine learning is explained. Three solutions has been choosen with different classifying algorithms. This ensures separate origin of the calculated results, which should simplify overall analysis of the models. The data mining software called WEKA is used for the calculations. Possible combinations of testing, verification process and modeled soil profiles are presented in the process.
All the research is concluded in final comparison of obtained results to laboratory data and most accurate predictions are selected.

Table of contents

List of figures

List of tables

List of attachments

Nomenclature

a_n	$-$ value of n attribute in neural network algorithm
A	$-$ area of cross $-$ section of effective resistive unit
A_c	$-$ area of a cone
A_s	$-$ area of a friction sleeve
B	$-$ diameter of the cone
B_q	$-$ pore pressure parameter
c	$-$ cohesion
c_v	$-$ classifying value, usually nominal
d	$-$ depth as an attribute in data mining techniques
f_s	$-$ unit sleeve friction resistance
F_r	$-$ normalized friction ratio
G	$-$ shear modulus
i	$-$ inclination of rods
L	$-$ path lenght
N_c	$-$ Terzaghi's bearing capacity factors
N_q	
N_γ	$-$ cone factors
N_{De}	
N_{ke}	
N_{kt}	$-$ constant related to friction ratio
$N_{\Delta u}$	
q_c	$-$ cone resistance
q_e	$-$ effective cone resistance
q_t	$-$ corrected cone resistance
Q_t	$-$ normalized cone resistance
R	$-$ measured resistance
R_f	$-$ friction ratio
S_i	$-$ value of sample in a decistion tree algorithm
S_t	$-$ sensitivity
s_u	$-$ undrained shear strength
s_{ur}	$-$ remoulded undrained shear strength
t_v	$-$ threshold value in neural network algorithm
u_0	$-$ in situ pore water pressure
u_2	$-$ pore water pressure, directy taken from CPT data
Δu_{2c}	$-$ excess pore water pressure behind the cone
w_n	$-$ weight for n attribute in neural network algorithm
$x_{j,i}$	$-$ j attribute of the sample in a decistion tree algorithm
z	$-$ depth
z_{GWL}	$-$ depth of ground water level

α $-$ *cone apex angle*
α $-$ *calibration coefficient, individual per probe*
β $-$ *angle used in the solution by Janbu and Senneset* (1974)
γ $-$ *total unit weight*
λ $-$ *cone roughness indicator* ($1 = rough\ cone, 0 = smooth\ cone$)
ρ $-$ *resistivity*
σ_v' $-$ *vertical effective stress on the surface of the cone*
σ_{v0} $-$ *vertical initial total stress*
ϕ' $-$ *effective friction angle*

1. Introduction

1.1. Objective of the thesis

The goal of the following thesis is to present and review various methods for interpretation of data obtained by CPT - Cone Penetration Test. The main idea is to differ layers of soil, define its properties and detect sensitive clays with quick clays in particular. With usage of different ways of interpretation it is possible to examine and choose, which one of them is most accurate for this kind of purpose. One of the methods presented in this thesis is special owing to interdisciplinary and experimental approach: data mining. One of the chapters is fully dedicated to techniques and tools of machine learning. As an author of this thesis I hope, that the Reader will find following paper useful and interesting.

1.2. Motivation

The CPT sounding is a worldwide popular *in situ* method of sounding for soil investigation. It allows collecting data in continuous manner - it is easy to be processed and presented in form of graph or in any linear manner. It is an easy, fast and efficient tool, which every geotechnical engineer should be familiar with. This work was written to explain and discuss basic rules, technical issues (like equipment or testing procedures) and to present different methods of interpretation of CPT results. Geotechnical investigation is always an important issue for engineering projects, no matter the class or scale of those. Thereby every civil engineer should acquaint him or herself with following work.
Moreover, also aspects of familiarizing Reader with unique soils was taken into account. Norway is one of a few northern areas, where quick clay occurs - a type of soft clay, sensitive and unstable. In appropriate conditions, deposits of quick clay may liquefy and cause massive displacement of higher placed soil layers. One of the most famous and well-documented cases of quick clay liquefaction was landslide in 1978 at Rissa. Rising awareness of such peculiar soil and showing possible methods of detecting is also intended by the author.

1.3. Use of materials

Following thesis is continuation of work from the same author; course TBA 4510 - geotechnical engineering specialization project, which was finished in autumn semester 2015 at NTNU.
Moreover, chapter 7 would be impossible to complete without using book "Data mining - practical machine learning tools and techniques" by Ian. H.

Witten, Eibe Frank and Mark. A. Hall. Large amount of quotes was taken from that volume, because of plain and clear explanations of machine learning techniques. The mentioned book is basically a manual for data mining software called WEKA, which was used in the following research.

1.4. History of CPT

The very first cone penetrometer was invented in 1932 by Peter Barentsen - a civil servant at Department of Public Works in Holland. A steel rod of 19 mm diameter with 10 cm^2 cone with 60° angle was initially in use and was operated manually by single man. When body-weight was applied, rod was pushed into the soil to depths of maximum 10-12 meters and readings of penetration resistance q_c were red with use of manometer. Due to fact, that achieveable cone resistance values were limited by the weight of operator, this method was generally used for determing thickness of weak sendimental soils like peats and clays.

Figure 1.1 *The first manual CPT apparatus*

In consideration of extending CPT range in more consistent layers of soils the pushing force was increased. In 1935, the director of the "Laboratorium voor Grondmechanica" in Delft - T.K. Huizinga has developed deep CPT apparatus with reaction force of 10 tons. It took about 3 days to complete the test, because of method of applying the ballast. It required to dig a hole filled with 6m^3 sand with wooden floor at the bottom, which was connected to the device on the ground surface. It was a time consuming method, though for the first time in history it allowed to compute bearing capacity of wooden piles, instead of driving test piles.
Subsequent upgrades improved shape and type of cone: Vermeiden in 1948 added a conical mantle to the jacket cone; Begemann in 1953 connected so called "adhesion jacket", which lately was transformed into friction sleeve. Significant progress happened in 1965, when H.K.S. Begemann classified soils and their relation with friction ratio (f_s/q_c). Next step for upgrading the equipment was introducing in order: pushing rigs with greater force, hydraulic jacks instead of gravity load and replacement mechanical cones with electric ones.

1.5. CPTu

Piezocone is essentially a standard electrical CPT cone with built-in module for measuring in situ pore water pressure u_0 during the penetration. Armed with stainless steel porous tip, the conventional probe measured pore pressures in the vicinity of CPT sounding. It was developed by Norwegian Geotechnical Institute in 1974 and utilized by Nilmar Janbu and Kåre Senneset in their research. Year later Swedish scientist Tortensson performed dissipation test with piezocone developed by himself. Development of CPTU gathered pace in 1980ies all around the globe: Canada, Netherlands, United States and of course Scandinavia. Furthermore, adding next parameter in Cone Penetration Test allowed to interpret data with tremendously greater accuracy than before and determine soil types with more characteristics.

1.6. RCPTu

For electrical resistivity (or conductivity) measurement of soil a piezocone with resistivity module was combined in the mid 1970ies in Holland. Initially this method was developed to detected contamination of groundwater. However, evaluation of density, porosity and corrosive properties of soil become another reason to use this technique.
There are two types of resistivity probes: operating in soil or in water. Soil type probe consist of standard 10 cm^2 cone, friction sleeve and additional module mounted in the back. It is equipped with the set of two or four isolated electrodes with insulating collar. First outer electrode conducts current and the second outer is grounded. With the voltage applied, the two inner measure resistivity of a soil. The resistivity module with a set of 4 electrodes with plastic insulation is in total about 350 mm long.
Data used in my thesis was obtained from RCPTu soundings with the use of resistivity module with 4 electrodes.

Figure 1.2 C*onnected resistivity module to piezocone, photographed by S.P.*

2. Analytical models for CPT

The key factor in determining soil properties is their relation with CPT measurements. For validation of those correlations, many theories has been developed since 1960ies. Relationship between cone resistance and soil properties may be treated in two ways. At the beginning assumptive soil properties are used to calculate cone resistance - this solution is used for determining end-bearing capacity of piles or estimating process of liquefaction. On the other hand, process of calculation can be back tracked. Given measured cone resistance as an input the soil properties are back-calculated.
There can be differed an abundant number of correlations and, what is worse, not a single one is precise nor universal due to complex nature of a problem. As a result of high stresses and strains caused by cone penetration, initial soil conditions are unidentified. Each of available methods should be used accordingly with caution. Moreover, every experienced engineer should bear this fact in mind and choose a method of interpretation to the appropriate circumstances - whether probed soil is cohesive or cohesionless; conditions are drained or undrained; boundary conditions are fitting or mismatched. Following project describes briefly few of the most popular methods.

2.1. Bearing capacity theory

This is one of the first and most recognizable theory used for CPT data analysis. The cone tip is treated as circular, deep foundation, whereas cone tip resistance q_c defines ultimate load in subsoil during the failure. Analytical approach is based on classic equation of bearing capacity (Terzaghi 1943, Meyerhof 1951, Brinch Hansen 1970):

$$cN_c + \sigma_v' N_q + \frac{\gamma B N_\gamma}{2} = q_t - u_2 = q_e \quad \textbf{(1)}$$

$$c - cohesion$$

$$N_c, N_q, N_\gamma - Terzaghi's\ bearing\ capacity\ factors$$

$$\sigma_v' - vertical\ effective\ stress\ on\ the\ surface\ of\ the\ cone$$

$$\gamma - total\ unit\ weight$$

$$B - diameter\ of\ the\ cone$$

$$q_t - corrected\ cone\ resistance$$

$$u_2 - \ pore\ water\ pressure, directy\ taken\ from\ CPT\ data$$

$$q_e - \ effective\ cone\ resistance$$

With given a relatively small cone diameter to the other parameters the third term of the equation can be assumed as minimal and thus neglected. Next assumption describes failure, which develops under undrained conditions. Assumptions for the cohesive soils:

- angle of internal friction is zero $ø = 0°$
- undrained shear strength is equal to cohesion $s_u = c$ (from Mohr-Coulomb criterion)
- parameter N_q is neglected - influence of depth is already taken into account by N_c coefficient, making undrained shear strength a primary factor of effective cone resistance

The simplified equation is formed as following:

$$q_e = \ s_u N_c + \sigma_{v0}$$

$$\sigma_{v0} - vertical\ initial\ total\ stress$$

Next step in calculations is choosing or calculating cone factor accordingly to selected methods:

Table 1 - *Bearing capacity solutions*

Authors	**Cone factor or main conclusions**
Cohesive soils	
Meyerhof (1961)	$N_C = 1{,}15 \cdot (6{,}28 + a + cot\frac{a}{2})$
Durgunoglu and Mitchell (1975)	$N_C = 1{,}2 \cdot (2{,}443 + 3{,}303\lambda + \sin[(1-\lambda)\frac{\pi}{2}]$
Houlsby and Wroth (1982)	q_C increases indefinitely with depth
Koumoto and Kaku (1982)	q_C is constant if depth is greater than the cone diameter
Cohesionless soils	
Janbu and Senneset (1974)	$N_q = \frac{1+sinø'}{1-sinø'} \exp[(\pi - 2\beta)tanø']$
Durgunoglu and Mitchell (1975) Chen and Juang (1996)	$N_q = 0{,}194 \cdot \exp(7{,}629tanø')$
Di Simone and Golia (1988) Koumoto (1988)	Cone factors for plane strain cases are much less than for axisymmetric cases. Cone roughness has great influence on the value of N_C

There are two methods for analyzing bearing capacity approach:

- **limit equilibrium method**: basic assumption is that cone resistance is treated as deep, circular foundation, which is about to fail. Similarity to Terzaghi assumptions is essential. Initially, the failure mechanism is assumed; afterwards the global equilibrium is analyzed and in the end the critical load is calculated.
- **slip - line method**: this solution combines equations of equilibrium with yield criterion. As an effect a set of differential equations of plastic equilibrium is obtained, which allows to plot a slip line. Multiple slip lines create a network and critical load is possible to be calculated.

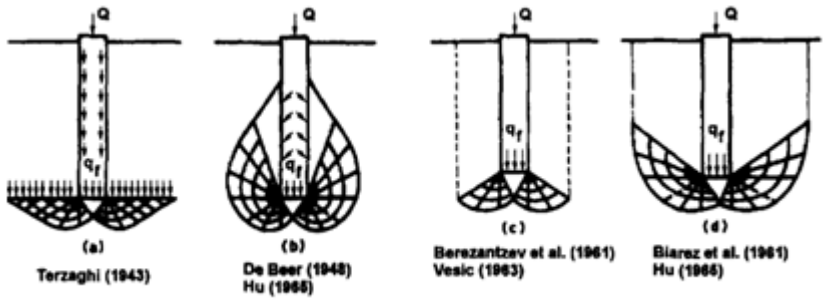


Figure 2.1 *Assumed failure mechanism for deep foundation*

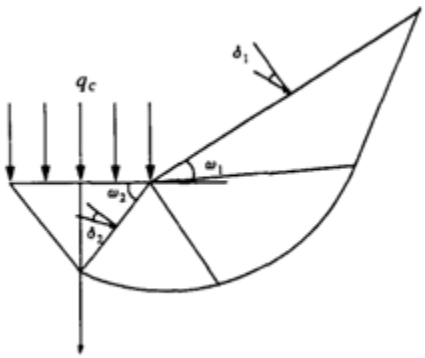


Figure 2.2 *Slip line network for Wedge and cone penetration analysis*

Limitations for bearing capacity methods:

1. Influence of soil deformation is ignored on the cone resistance - its value relies on soil stiffness and compressibility
2. "The bearing capacity approach ignores the influence of the cone penetration process on the initial stress states around the shaft. In particular, the horizontal stress tends to increase around the cone shaft after cone penetration and the influence of this change [...] is not considered in bearing capacity analysis." (Yu and Mitchell, 1998)

2.2. Cavity expansion

In 1945 Alan W. Bishop has connected cavity expansion theory with CPT: he noticed, that the pressure required to create a deep hole in elasto-plastic soil is proportional to pressure, which expands cavity of the same value under equivalent conditions. Two conditions must be fulfilled for calculating cone resistance:

- theoretical limit pressure solutions of cavity expansions must be established in given soil conditions
- relationship between cavity expansions limit pressures to cone resistance must be found

This theory was extended by Vesic (1972) by following Mohr-Coulomb criterion. Accuracy of results depends on yield criterias and stress-strain models of a soil (cohesive or non-cohesive). Large number of researchers linked limit pressure solutions to more practical values like cone resistance or pile end bearing. Figure below represents expansion of a cavity between soil particles.

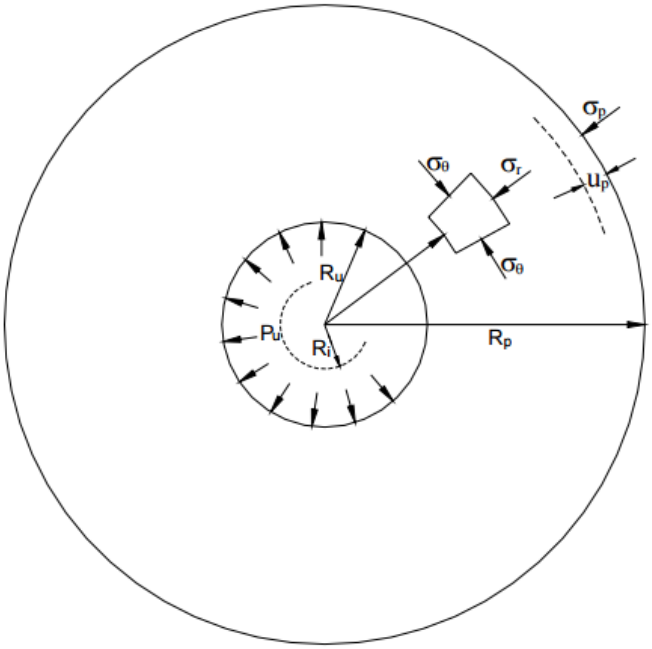


Figure 2.3 *Expansion of cavity radius from initial value to limited by ultimate internal cavity pressure*

Limitations for cavity expansion theory:

1. Influence of dilatancy is neglected, causing a tendency towards undervaluation of cone resistance.
2. "All cone factors N_k derived from cavity expansion solutions depends on the rigidity index I_r of soil" (Salgado, Prezzi, Kim, 2006)

Table 2 - *Cone factor N_k derived using different cavity expansions methods (after Yu et al. 1998)*

G/S_u	Ladanyi and Johnston (1974)	Vesic (1977)	Baligh (1975)	Yu (1993)	Yu (1993)
	rough cone	rough cone	rough cone	Smooth cone	Partly rough cone
50	8,3	9,1	15,9	8,5	10,4
100	9,2	10,0	16,6	9,3	11,2
200	10,1	10,9	17,3	10,1	12,0
300	10,6	11,5	17,7	10,6	12,5
400	11,0	11,9	18,0	10,9	12,8

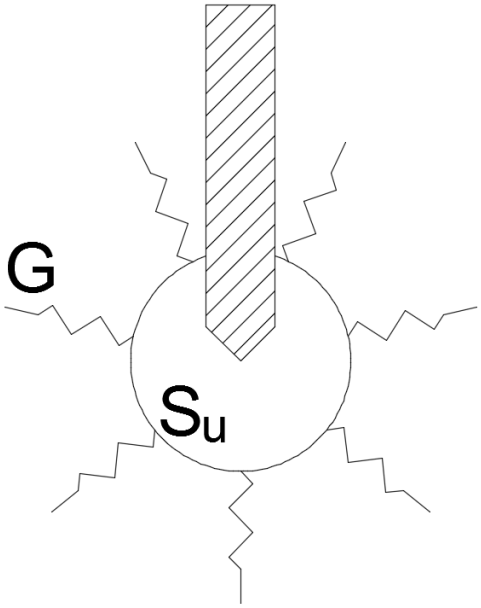


Figure 2.4 *Model of plasticized soil in the vicinity of cone at the state of failure*

2.3. Steady state deformation

In this method penetration of cone is "treated as steady state flow of soil past fixed cone penetrometer". Most of the models assume soil as a ideally plastic matter, yet other include strain hardening critical state. Steady state deformation method is still under development and at the moment its usage is restricted to undrained clays. Obtaining initial conditions of flow field for frictional soils proved to be quite problematic.

2.4. Incremental finite elements

Incremental displacement FEM is used for analyzing cone penetration test in two approaches:

- **Small Strain Models**. Cone is modeled in pre-bored hole and the soil has initial undisturbed properties. Then an incremental plastic collapse is calculated with following assumption: cone resistance is equal to failure load. *De facto* situation is quite different: the shaft of a cone tends to buildup high lateral stresses during the penetration. Thus, value of cone resistance will be greater than calculated by this model.
- **Large Strain Models**. In this case stress increase around the cone shaft in included, because it is possible to model vertical displacements generated from the penetration of a cone and simulate changes of initial stress conditions. Models for cohesive soils presented by Budhu and Wu in 1991-92 include elements without given thickness and frictional interfaces on surface of a cone.

Despite of rapid development of Finite Elements Methods there still doubts about the accuracy of the cone factors, especially in clays. For example, results acquired from van den Berg (1994) model of a circular footing on undrained clays were about 23% higher than the exact solution.

2.5. Calibration chamber testing

For years enormous calibration chambers were used to establish empirically values of cone penetration factors. There are three types of factors, divided by their correlation with leading parameter of soil:

- relative density
- friction angle
- state parameter

Next parameters, which have major influence, are: size of a chamber and assumed type of boundary conditions. "For example, if a flexible boundary (i.e. constant pressure applied) is used in the chamber testing, then the cone resistance measured in the chamber will be lower than what would be measured in the field for the same soil at same initial conditions. On the other hand, if a rigid boundary (i.e. zero displacement) is used in the chamber

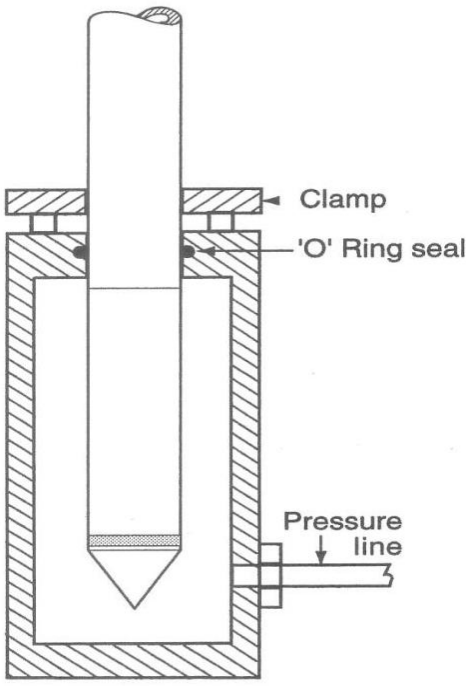


Figure 2.5 *Scheme of calibration chamber*

testing, the cone resistance measured in the chamber would be higher than that measured in the field." (*Analysis of cone resistance: review of methods, by H. S. Yu and J. K. Mitchell).*

3. Soil classification and measurement of properties

3.1. Normalized parameters

A standard CPTu cone measures cone tip resistance, sleeve friction and pore water pressure. However, study on solely raw data from sounding can give false results. For instance, vertical stresses increases with depth, as well as cone tip resistance tends to rise at the same time. Interpretation of such readings could cause errors - increase of mentioned parameters can change of their supposed classification. This problem was very visible in deep soundings especially: like in a thick, normally consolidated soil deposits or offshore test. In order to specify soil stratigraphy, evaluate parameters of sub-soil material in geotechnical design and evaluate soil-behavior in a rational manner, obtained data from soundings must be post-processed.
Following equations represent step by step calculations:

- normalized cone resistance Q_t :

$$q_t = q_c + (1 - \alpha) \cdot u_2 \quad \textbf{(3)}$$

$$\sigma_{vo} = \gamma_{soil} \cdot z \quad \textbf{(4)}$$

$$\sigma'_{vo} = \sigma'_{vo} - u_0 \quad \textbf{(5)}$$

$$u_o = \gamma_w \cdot (z - z_{GWL}) \quad \textbf{(6)}$$

$$Q_t = \frac{q_t - \sigma_{vo}}{\sigma'_{vo}} \quad \textbf{(7)}$$

$$q_c - directly\ taken\ from\ CPT\ data$$

$$\alpha - calibration\ coefficient, individual\ per\ probe$$

$$u_2 - directy\ taken\ from\ CPT\ data$$

$$u_0 - pore\ water\ pressure\ in\ situ$$

$$z - depth$$

$$z_{GWL} - depth\ of\ ground\ water\ level$$

- normalized friction ratio F_r :

$$F_r = \frac{f_s}{q_t - \sigma_{v0}} * 100\% \quad \textbf{(8)}$$

$$f_s - \ directy\ taken\ from\ CPT\ data$$

- pore pressure parameter B_q:

$$u_o = \gamma_w \cdot (z - z_{GWL}) \quad \textbf{(9)}$$

$$q_t = q_c + (1 - \alpha) \cdot u_2 \quad \textbf{(10)}$$

$$\sigma_{vo} = \gamma_{soil} \cdot z \quad \textbf{(11)}$$

$$B_q = \frac{u_2 - u_0}{q_t - \sigma_{v0}} = \frac{\Delta u_{2c}}{q_n} \quad \textbf{(12)}$$

$$u_2 - \ directy\ taken\ from\ CPTU\ data$$

$$\Delta u_{2c} - \ excess\ pore\ water\ pressure\ behind\ the\ cone$$

3.2. Undrained shear resistance

Due to numerous empirical and theoretical researches there is a wide variety of possible calculations of undrained shear resistance. Multiple parameters must be taken into account, such as:

- stress history,
- soil anisotropy,
- type of failure,
- strain rate.

Strength anisotropy is very influential factor for sensitive clays. However, for all types of soils the most crucial of all is assumption of a design problem, which affects final value of s_u. Analytical approaches were described in chapter 2 of this thesis. To calculate undrained shear resistance a theoretical cone factor can be taken - usually its value is already defined. In this case a Karlsrud et al. (1996) approach was choosen. Most common values are $N_{kt} \in (6{,}18)$ and mentioned cone factor is related to pore-pressure ratio:

$$s_u = \frac{q_t - \sigma_{vO}}{N_{kt}} \quad \textbf{(13)}$$

$$N_{kt} = 18{,}7 - 12{,}5 \cdot B_Q \quad \textbf{(14)}$$

Most common values are $N_{ke} \in (1{,}10)$ and mentioned cone factor is influenced by pore-pressure ratio:

$$s_u = \frac{q_e}{N_{ke}} = \frac{q_t - u_2}{N_{ke}} \quad \textbf{(15)}$$

$$N_{ke} = 13{,}8 - 12{,}5 \cdot B_Q \quad \textbf{(16)}$$

Most common values are $N_{\Delta u} \in (6{,}8)$ and mentioned cone factor is not linked with pore-pressure ratio:

$$s_u = \frac{\Delta u}{N_{\Delta u}} = \frac{u_2 - u_0}{N_{\Delta u}} \quad \textbf{(17)}$$

$$N_{\Delta u} = 1{,}8 + 7{,}25 \cdot B_Q \quad \textbf{(18)}$$

3.3. Sensitivity

This parameter defines the ratio of undrained shear strength in undisturbed conditions and totally remolded undrained shear strength. Weak clays and quick clays in particular are marked by high value of this ratio, due to low sleeve friction. Vane test is an efficient method for measuring sensitivity *in situ*:

$$S_t = \frac{N_s}{R_f} \quad \textbf{(19)}$$

$$R_f = \frac{f_t}{q_t} - \ friction\ ratio$$

$$N_s - \ assumed\ constant$$

3.4. Resistivity

In this method resistivity of soils is measured indirectly - by the use of two or four electrodes at the constant supplied current. If we assume that soil is homogenous and isotropic medium and is "plugged" in a perfect electrical source, then we obtain following equation of soil resistivity:

$$\rho = \frac{A}{L} \cdot R \quad \textbf{(20)}$$

$$L - path\ lenght$$

$$A - area\ of\ cross - section\ of\ effective\ resistive\ unit$$

$$R - measured\ resistance$$

With the probes dimensions given, the parameters L and A are calculated. The only variable left is soil resistance R. It is influenced by pore water pressure and presence of ions - which indicates concentration of chemical compounds like salts, acid etc. Nowadays, RCPTU is used for detecting marine clays (which characterize high salinity) and for determining soil and ground water contamination.

Figure 3.1 *Disconnected resistivity module*

3.5. Robertson classification chart

One of most popular and recognizable solutions include usage of all three parameters from CPTu data. System, which was initially designed by Peter K. Roberston in 1986, consist of two graphs represented in Figure 3.1. One common feature is similarity of vertical axis, which represents value of normalized friction ratio Q_t in logarithmic scale. Different marked areas allows to classify approximately type of soil:

- sandy soils - relatively high Q_t, low F_r and very low B_q
- soft clays - low Q_t , high F_r and moderate B_q
- organic peat - minor Q_t , very high F_r and high B_q
- sensitive soil - low Q_t , low F_r and very high B_q
- soils with high OCR (horizontal stresses) - tendency towards high Q_t, high F_r and low B_q

In sands and non-cohesive soils tip resistance lowers with diameter of particles. Fine-grained material has relatively higher sleeve resistance at the expense of pore pressure. Impermeable or saturated soils do indicate high values of pore water pressure parameter.
"Generally, soils that fall in zones 6 and 7 represent approximately drained penetration, whereas, soils in zones 1,2,3 and 4 represent approximately undrained penetration. Soils in zones 5,8 and 9 may represent partially drained penetration. An advantage of pore pressure measurements during cone penetration is the ability to evaluate drainage conditions more directly" (T. Lunne, P.K. Robertson, John J.M. Powell, 1997).

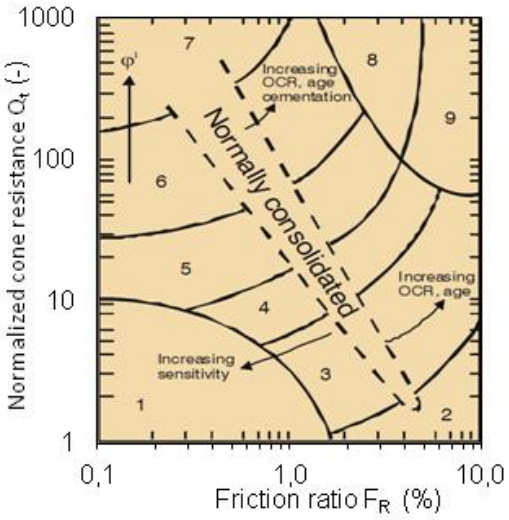


Figure 3.2 *Soil behaviour type classification chart based on normalized cone resistance and friction ratio (after Robertson, 1990)*

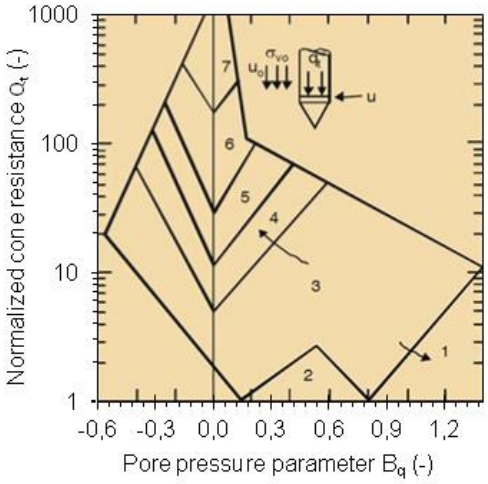


Figure 3.3 *Soil behaviour type classification chart based on normalized cone resistance and pore pressure parameter (after Robertson, 1990)*

Table *3* - *Soil classification for Robertson charts*

Zone	Soil behaviour type	Zone	Soil behaviour type	Zone	Soil behaviour type
1	Sensitive, fine grained	4	Silt mixtures clayey silt to silty clay	7	Gravelly sand to sand
2	Organic soils-peats	5	Sand mixtures; silty sand to sand silty	8	Very stiff sand to clayey sand
3	Clays-clay to silty clay	6	Clean sands to silty sands	9	Very stiff fine grained

3.6. Senneset classification chart

This chart was developed in response to older classification solutions, which based on sleeve friction and cone resistance only. It was a new answer to undependable and inaccurate measurements, caused by the effect of water pressure on unequal end areas of differently designed cones. Moreover, many CPTu test and studies have shown, that analysis of sleeve friction is sometimes discrepant with pore water pressure and cone resistance.
Initial version of chart proposed by Senneset and Janbu (1985) used measured cone resistance q_c instead of total cone resistance q_t in the new one.
It is the only chart plotted in non-logarithmic scale, thus vertical axis starts from initial value equal to zero.

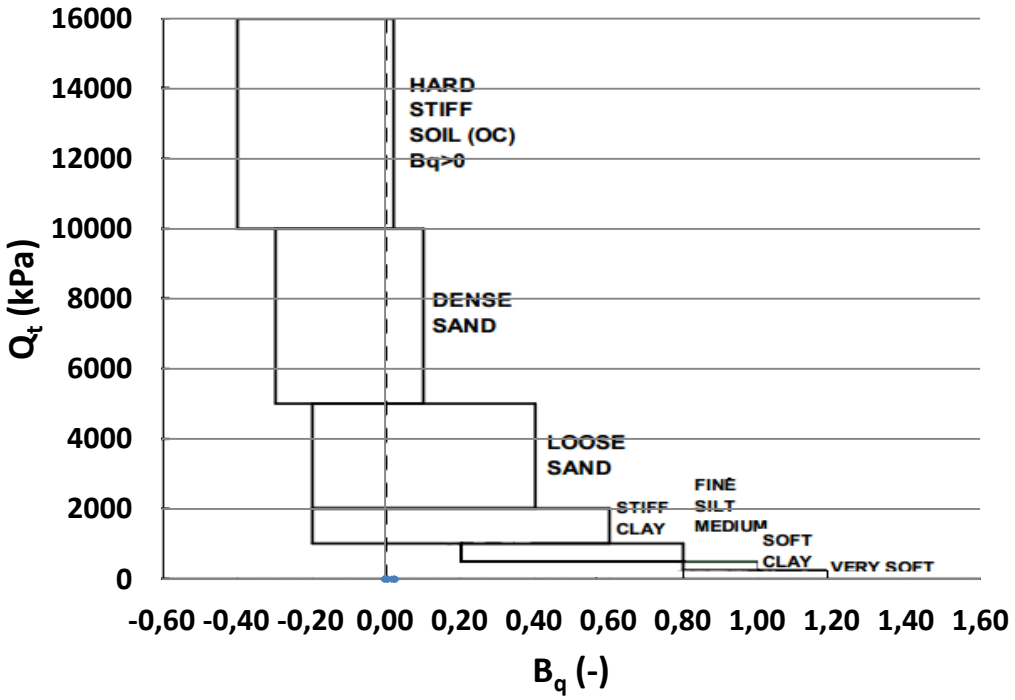


Figure 3.4 CPTu c*lassification chart (Senneset, 1989)*

Table *4* - *Soil classification for Senneset chart*

Soil state	Clay		Silt		Sand	
	a	tanø′	a	tanø′	a	tanø′
	[kN/m^2]	[-]	[kN/m^2]	[-]	[kN/m^2]	[-]
Soft-loose	5 - 10	0,35 - 0,45	0 - 5	0,50 - 0,60	0	0,55 - 0,65
Medium	10 - 20	0,40 - 0,55	5 - 15	0,55 - 0,65	10 - 20	0,60 - 0,75
Stiff, dense	20 - 50	0,50 - 0,60	15 - 30	0,60 - 0,70	20 - 50	0,70 - 0,90

3.7. Eslami and Fellenius classification chart

This particular chart was invented for investigating data from CPT and CPTu in pile design. Original data base, which was used for defining following graph, included cone penetrometer data correlated with sampling, laboratory data and borings from 20 investigation sites in 5 countries. For soil classification more basic parameters are used: sleeve friction f_s and effective cone resistance q_e, which proved to be more reliable for plotting boundaries of soil types than cone resistance q_c. It is a simple chart for quick profiling analysis of CPT data.

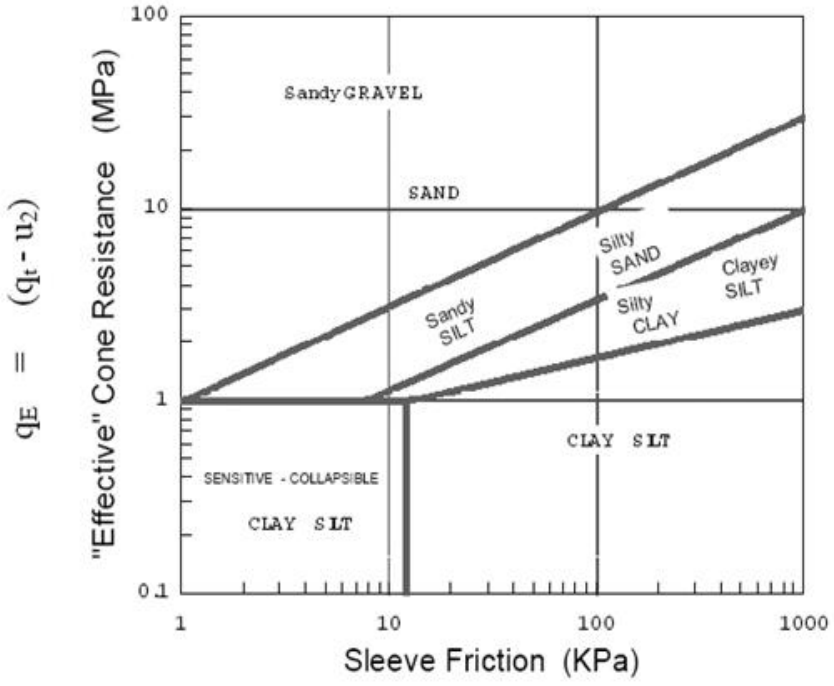


Figure 3.5 *CPTu classification chart (Eslami and Fellenius, 1989)*

Table 5 - *Soil classification for Eslami and Fellenius chart*

Zone	Soil type
1	Sensitive and collapsible clay and/or silt
2	Clay and/or silt
3	Silty clay and/or clayey silt
4	Sandy silt and/or silty sand
5	Sand and/or sandy gravel

3.8. Schneider classification chart

This method was developed in 2008 and is based on Robertson's charts. During his theoretical studies James K. Schneider has taken into account few analytical problems: influence of initial yield stress ratio and degree of consolidation on soil behaviour during loading. "Increases in YSR and degree of consolidation during loading tend to result in an increase in normalized cone tip resistance and decrease in pore pressure parameter, which are typically used for soil classification by piezocone." His research proved, that " for many cases the influence of YSR and partial consolidation have opposite effects when plotting data as Q against $\Delta u_2/\sigma'_{v0}$ (=Bq·Q). Therefore, charts of Q plotted against Δu2/σ'v0 are more useful for evaluation of soil type than conventional plots of Q against Bq." (Schneider, 2008)
Proposed system eliminates use of net cone resistance $q_{net} = q_t - \sigma_{v0}$ in both axes of charts, thus data are not plotted by function of its own self parameter, which reduces distortion in soil classification.

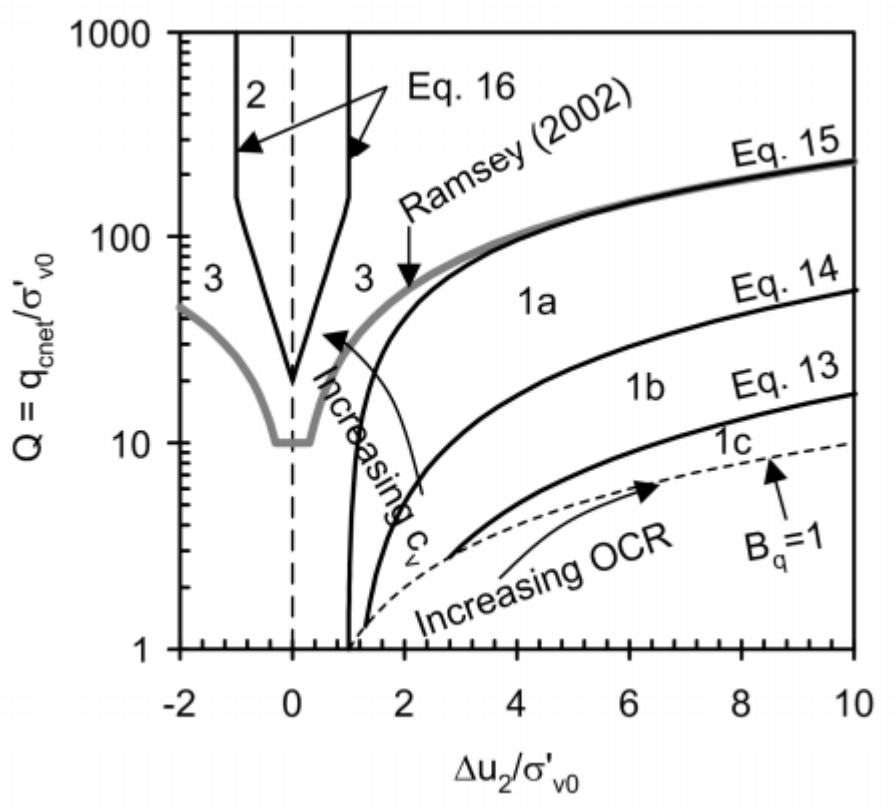


Figure 3.6 *CPTu classification chart (Schneider, 2008)*

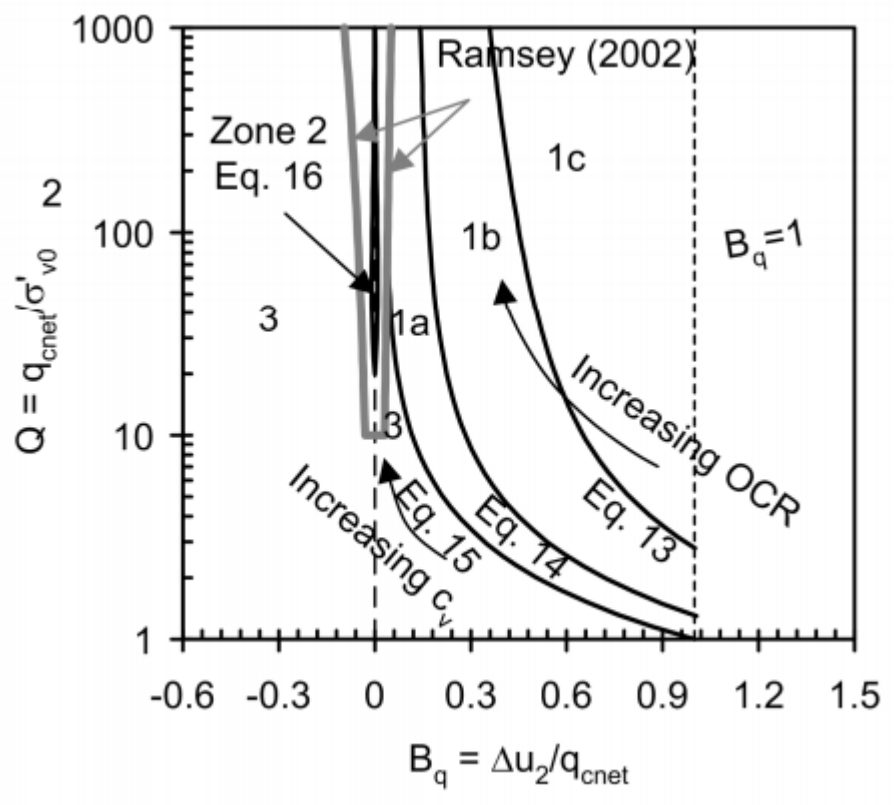


Figure 3.7 *CPTu classification chart (Schneider, 2008) with classic normalized parameters used by e.g. Robertson*

Table 6 - *Soil classification for Schneider charts*

Zone	Soil type
1a	Silts and low I_r clays
1b	Clays
1c	Sensitive clays
2	Esentially drained sands
3	Transitional soils

3.9. Summary on classification charts

The data from CPTu and RCPTu soundings, which were conducted in investigation sites in Søn-Trøndelag, Norway, produced comparable outcome of soil profiles. It is possible to recognize a linear layout of results on graphs and familiar formation of layers in the profile. For best visualization and formulation of conclusions a representative sounding was choosen from all available data.

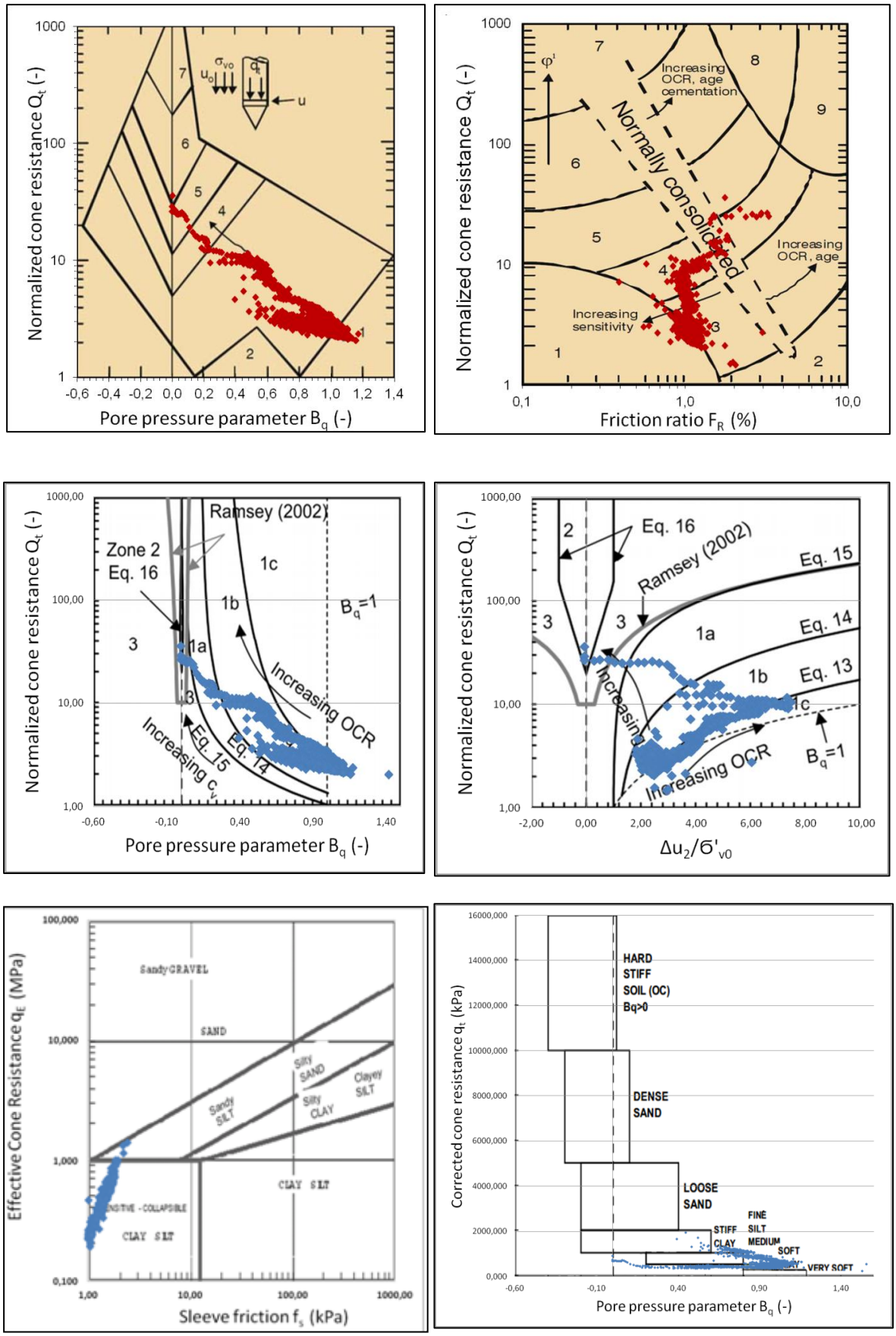


Figure 3.8 *Exemplary results for S1 RCPTu sounding from Klett*

Based on results from analyzed and presented classification charts following conclusions are formed:

1. Only two methods do not plot cone resistance versus its own value in different forms: Eslami & Fellenius (1997) and Schneider (2008). These profiling charts gave better results in cohesive soils: clays, silty clays and sensitive clays, which deposits are common in researched area.
2. Except for Eslami & Fellenius (1997) method, all of specified classification charts requires adjustments for estimating effective stress σ'_{vo} and total stress σ_{vo} due to usage of basic non-normalized parameters.
3. Utility of Senneset (1989) classification chart is more complicated in this case. The reason for disputable classification of soils is visible after visual comparison to the other solutions: in this single chart a great amount of data is plotted on non-classified area. Furthermore, boundaries of classified soils outline less than 40% of whole graph. Even more, the results are plotted in a natural scale instead of logarithmic - this solution causes lesser dispersion of data in the chart.
4. Highest density of data points was obtained by Elsami & Fellenius (1997) method - however, it was caused by MPa instead of kPa for effective cone resistance q_E. Moreover, logarithmic scales used for both axes. On the other hand, alternative solutions represent transition of layers in a more visible manner - Schneider (2008) in particular.
5. The best performance system for sensitive soils include usage of Qt versus $\Delta u2/\sigma'v0$ parameters. A newly proposed chart by Schneider detected and distinguished sensitive layers in more effective way. This graph with Robertsons classic solutions were choosen for data representation of soundings.

4. Equipment and testing site

4.1. Description of an equipment

A whole set of CPT consist of a series of rods with cone penetrometer installed at the end, pushing equipment and finally system for data collection and storage. Standard cone has tip with inclination of 60°, 10cm^2 area of base, diameter of 35,7 mm and 150 cm^2 area of friction sleeve. Both intermittent and continuous data can be collected of following parameters:

1) q_c - **cone resistance**; calculated by total pushing force on a cone Q divided by projected area of a cone A_c
2) f_s - **sleeve friction**; total force acting on the friction sleeve divided by surface area of the friction sleeve A_s
3) u_2 - **pore water pressure**; available only in CPTU piezocones; can be measured by numerous sensors, leading one is located behind the cone
4) *i* - **inclination**; best quality data is retrieved from vertical CPT, non-vertical requires correction of data; 1 degree of deflection is normally acceptable, rapid deflection with magnitude greater than 5 may result in damage of the equipment

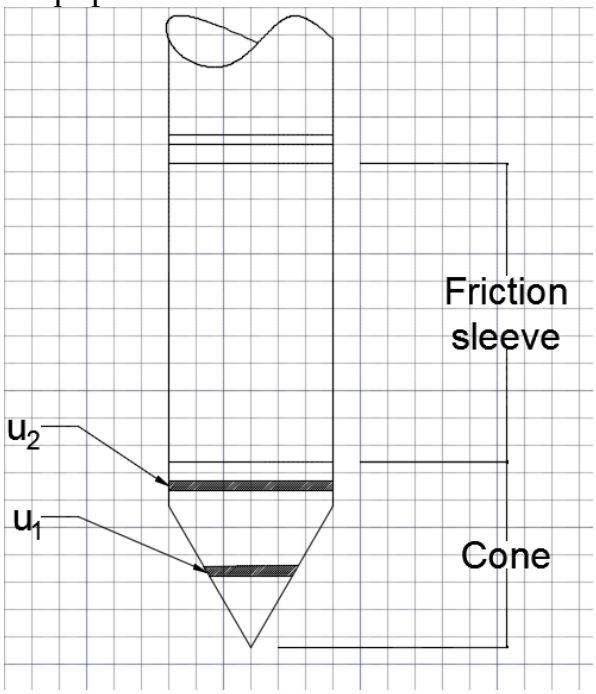


Figure 4.1 Scheme of piezocone

For CPTU tests was used type TE2 cone with u_2 sensor), section area 10cm^2 and diameter equal to 35,7mm. Geotech AB set, probe type Nova.
Measuring accuracy was regulated according to:
Revised NGF Message 5 (juni 2010); EN-ISO standard 22476-1

4.2. Fallan investigation site

Fallan is located in Sør-Trøndelag, Melhus community, about 35 kilometers south from Trondheim. It is located about 110 meters above the sea level. Local terrain is most diversified among other investigation sites. Steepness in the vicinity of soundings In this project data from soundings 2,5,7 and 4 were used. Area was investigated for prospective project of a new expressway, which would decrease the amount of traffic on European route E6. Vertical axis of road is perpendicular to the indicated row. According to map from Berggrunnsgeologidata base (Figure 4.2) there are sedimentary soils in the area - origin of them is linked to last glacial period.
At the end of Vistulian glaciation, about 10'000 years ago ice sheets begun to shrink. Massive amount of soil materials, which had been trapped in the glacier, has been released. Clay and silt particles has retained in the maritime waters, which had deeply penetrated inland. With a salty water as a bonding agent, they are formed in layers under the topsoil. As a result, quick clays are not usually placed at the surface and they are likely located in thin deposits. Area of occurrence is strictly limited by the highest sea level.
More detailed localization of borings and soundings along with topographic map can be found in appendixes A and B.

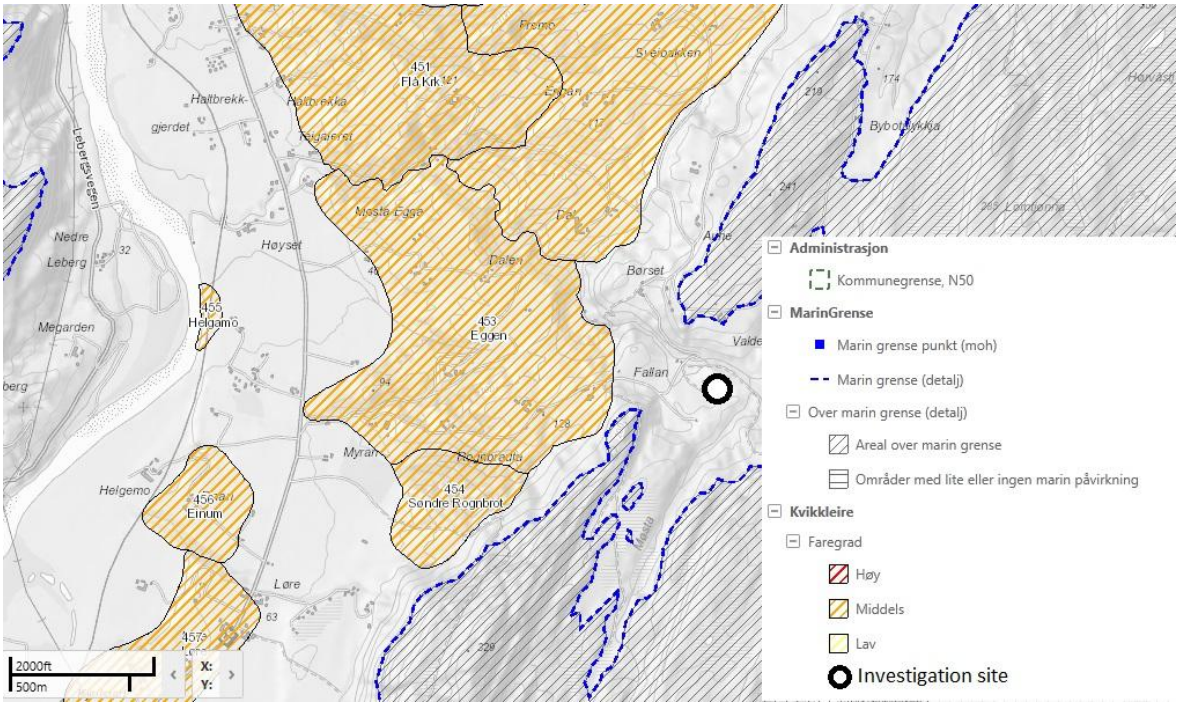


Figure 4.2 *Risk map of Fallan for quick clay slides developed by NVE (www.skrednett.no)*

4.3. Klett testing site

Klett is located in Sør-Trøndelag, Heimdal community, approximately 13 km south from Trondheim. It is located about 30 meters above the sea level. In south from Klett there is intersection of European routes E6 and E39. In the southeastern region a construction site is planned for motorway expansion. According to NGU data moraine, marine sedimentary and organic soils are present in the vicinity of area.
More detailed localization of borings and soundings along with topographic map can be found in appendixes C and D.

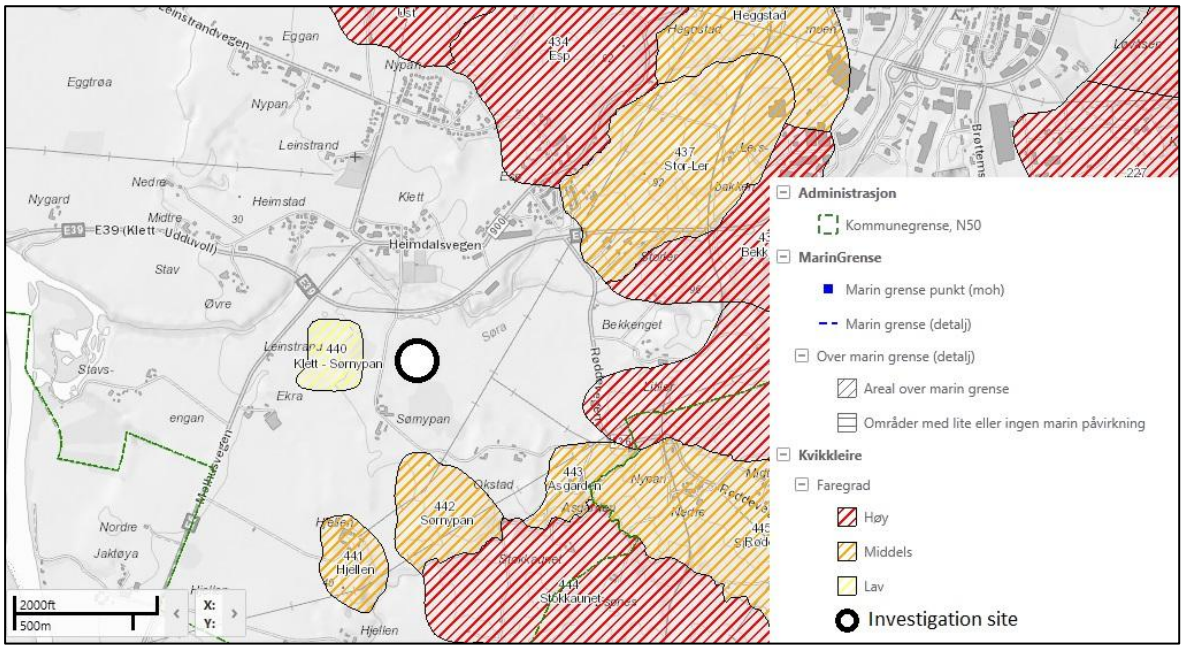


Figure 4.3 *Risk map of Klett for quick clay slides developed by NVE (www.skrednett.no)*

4.4. Tiller investigation site

The Tiller site (sometimes also referred as Kvenild due to the name of local village) is about 10 km south from city center of Trondheim. It is located at about 125 m above the sea level. Local clay deposits are very well investigated - from early 1980's this place was tested by numerous variants of ground investigation techniques. Most of the research was undertaken by NTNU with cooperation of NGI.

This site is also worth considering of historical events: in 1816 a major landslide occurred causing massive displacement of soil and leaving 15 casualties.
In relatively small proximity there is Klett test site; it is expected to receive similar results from both sites. Thanks to vast amount of collected data and experience on Tiller investigation site it is a great opportunity to check credibility of obtained results from methods referred in the following thesis.

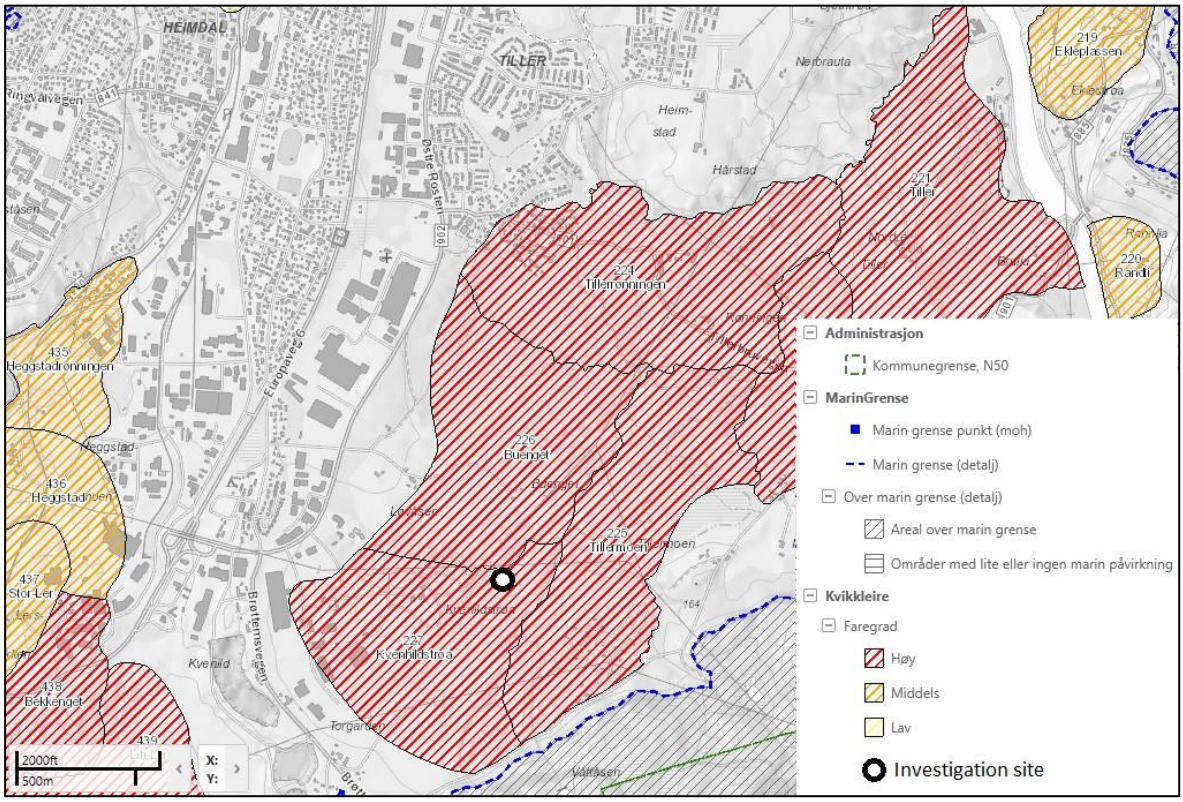


Figure 4.4 *Risk map of Tiller for quick clay slides developed by NVE (www.skrednett.no)*

Moreover, author of this thesis supervised R-CPTu soundings on 18th March 2015. For mentioned tests a heavy type geo-rig was used, equipped with two different resistivity modules. In the test participated: Jønland Jan, Senior Engineer and Winther Gunnar, Staff Engineer - technical personnel from NTNU. Pre-drilling on site was performed to 1,5 m.

Figure 4.5 *Heavy geo-rig, photo by S.P.*

Objectives of the investigations were:

- acquire data for soil profile and to complete database for Tiller investigation site
- verify newly calibrated resistivity module, which belongs to Multikonsult company, with one from the university

Two soundings were performed in a close vicinity to ensure comparable data. Results showed a layer of sensitive soil below 6-7 meter depth. Collected resistivity values are in 94 % similar - which proves a proper calibration of module. Unfortunately, readings of pore water pressure are imprecise, thus using them in the following thesis would be unsuitable. For data analysis results from first soundings were taken into account.

Figure 4.6 *Dismantling a rod from the resistivity module after sounding, photo by S.P.*

5. Interpretation of data

5.1. Fallan BP 2 - RCPTU

5.1.1. Plan of the investigation site

This section presents data from Fallan investigation site. Laboratory data are available for soundings 2 and 4. Results are discussed in chapters 8 and 9.

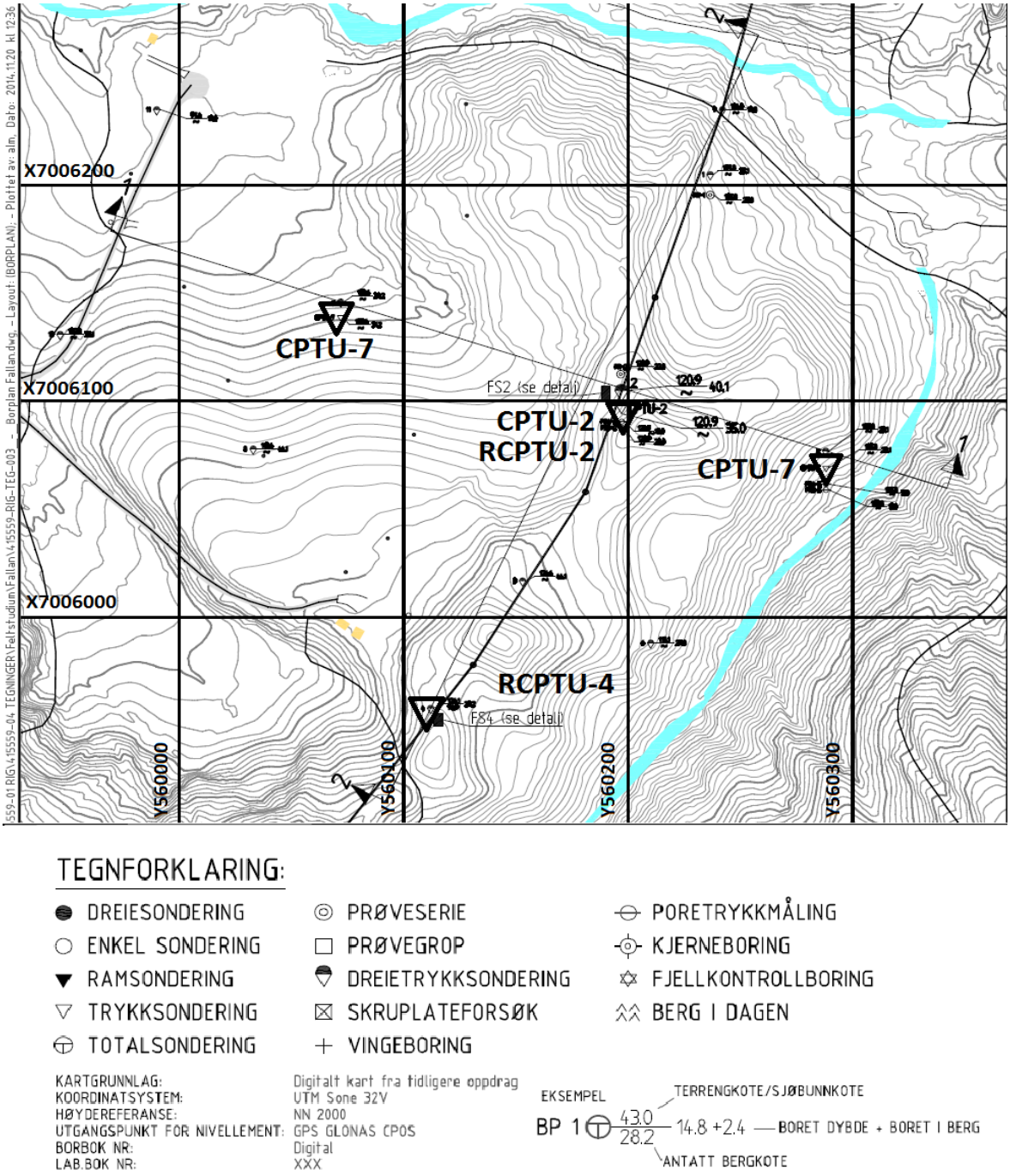


Figure 5.1.1 *Borings and soundings at Fallan investigation site*

5.1.2. Classification charts

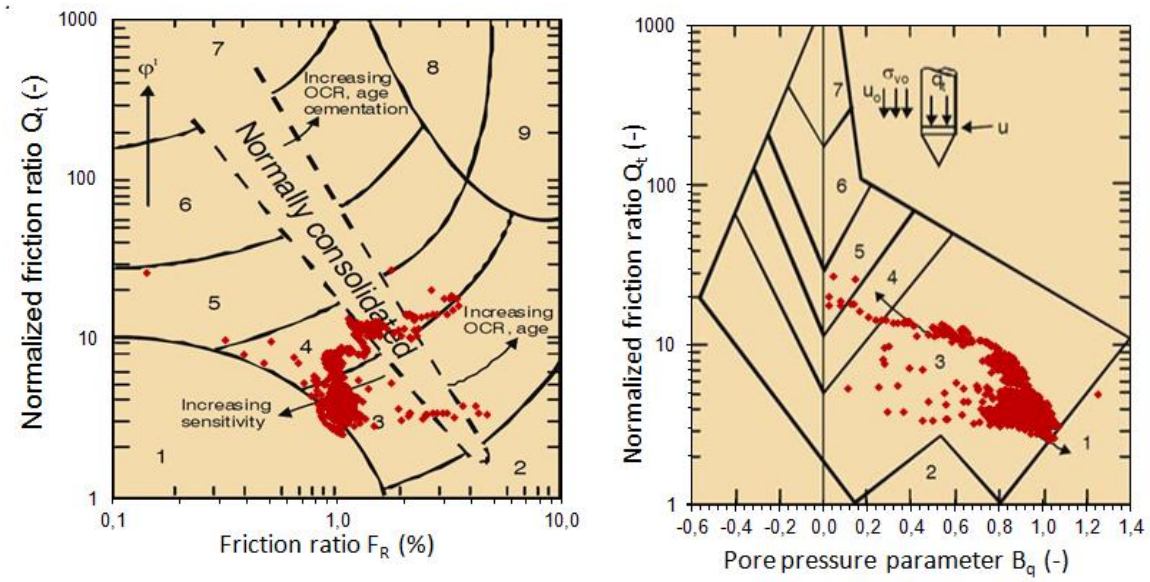


Figure 5.1.2 *Robertson classification charts (1990) for Fallan BP2*

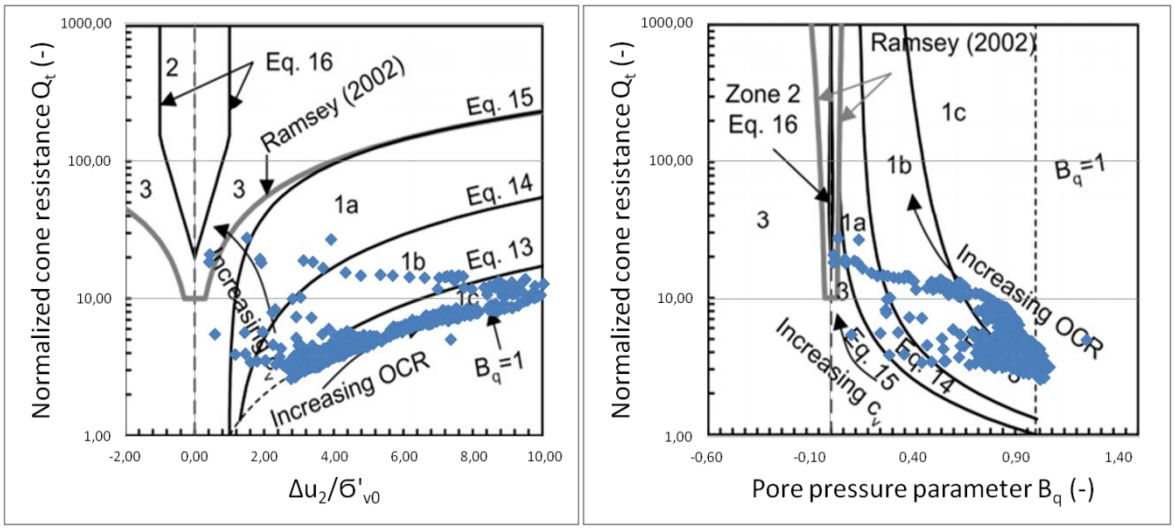


Figure 5.1.3 *Schneider classification charts (2008) for Fallan BP2*

Comment:

The soil profile for this investigation site is verified with laboratory data - samples were taken from BP2 and BP4 points. Initial 4-6 meters layer is non-sensitive clay (St ≈ 10), then it transits into sensitive and finally to quick clay.

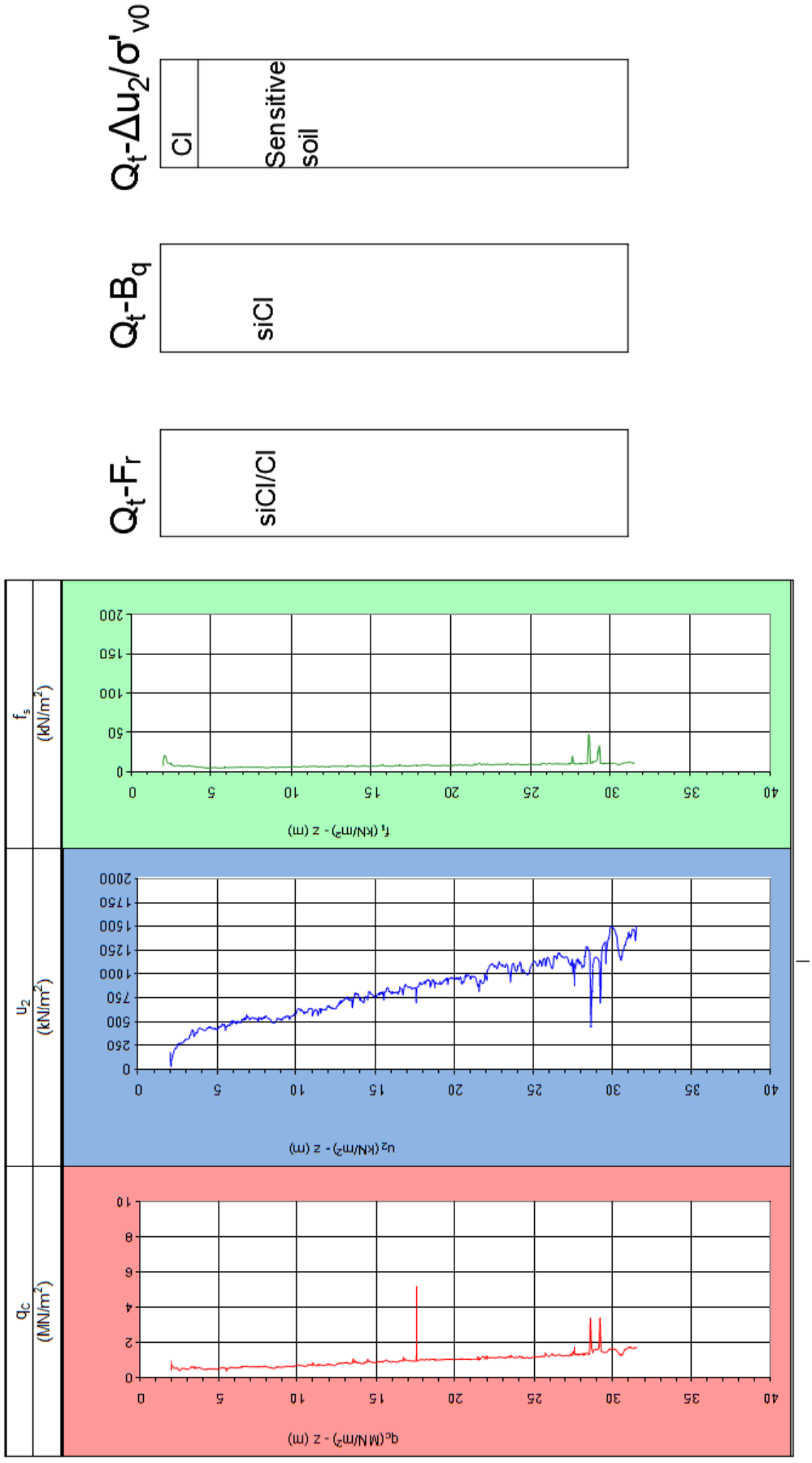


Figure 5.1.4 *Sounding parameters with interpreted profiles for Fallan BP2*

5.1.3. Undrained shear resistance

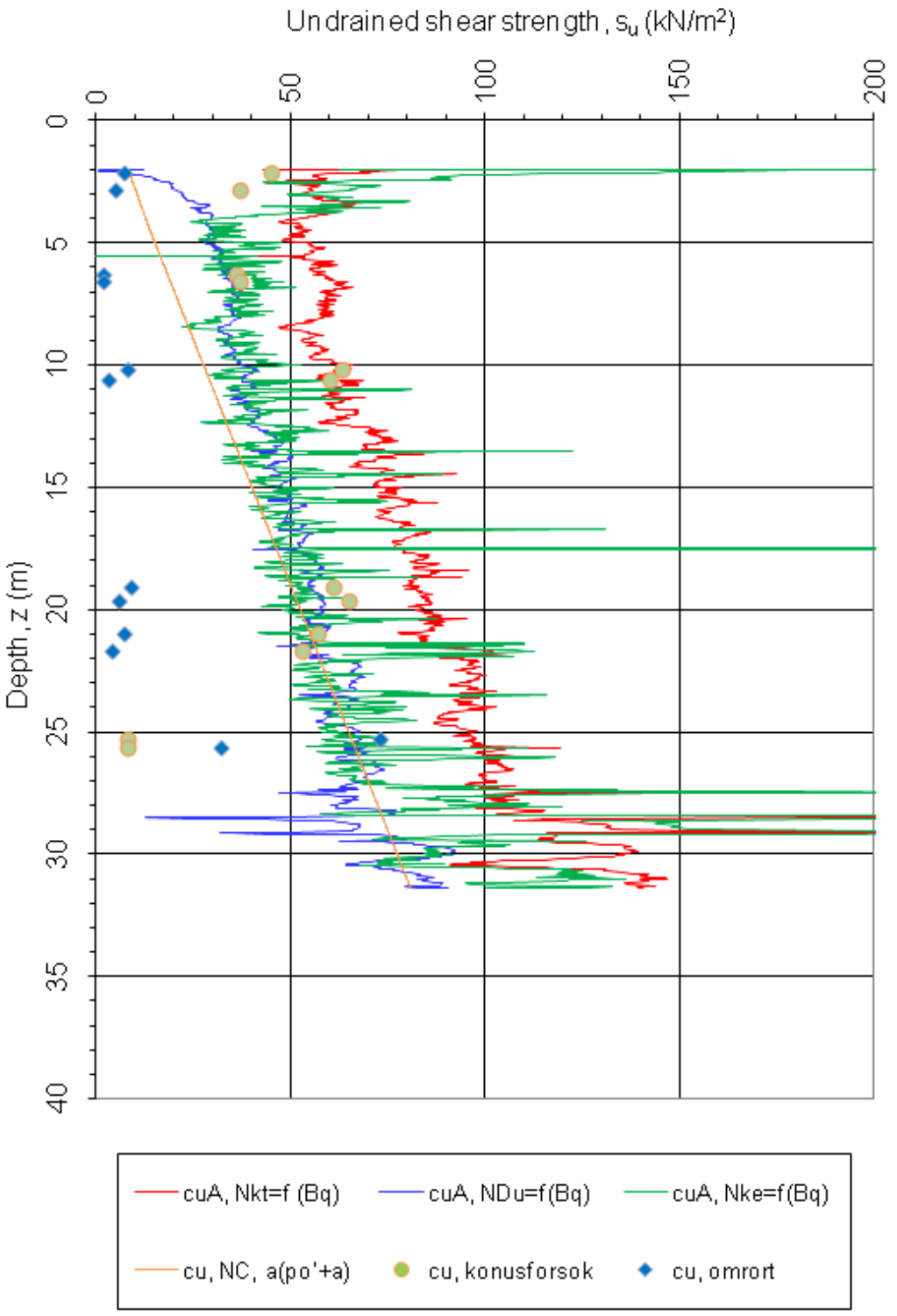


Figure 5.1.5 *Undrained shear resistance - interpreted and laboratory results for Fallan BP2*

5.1.4. Sensitivity

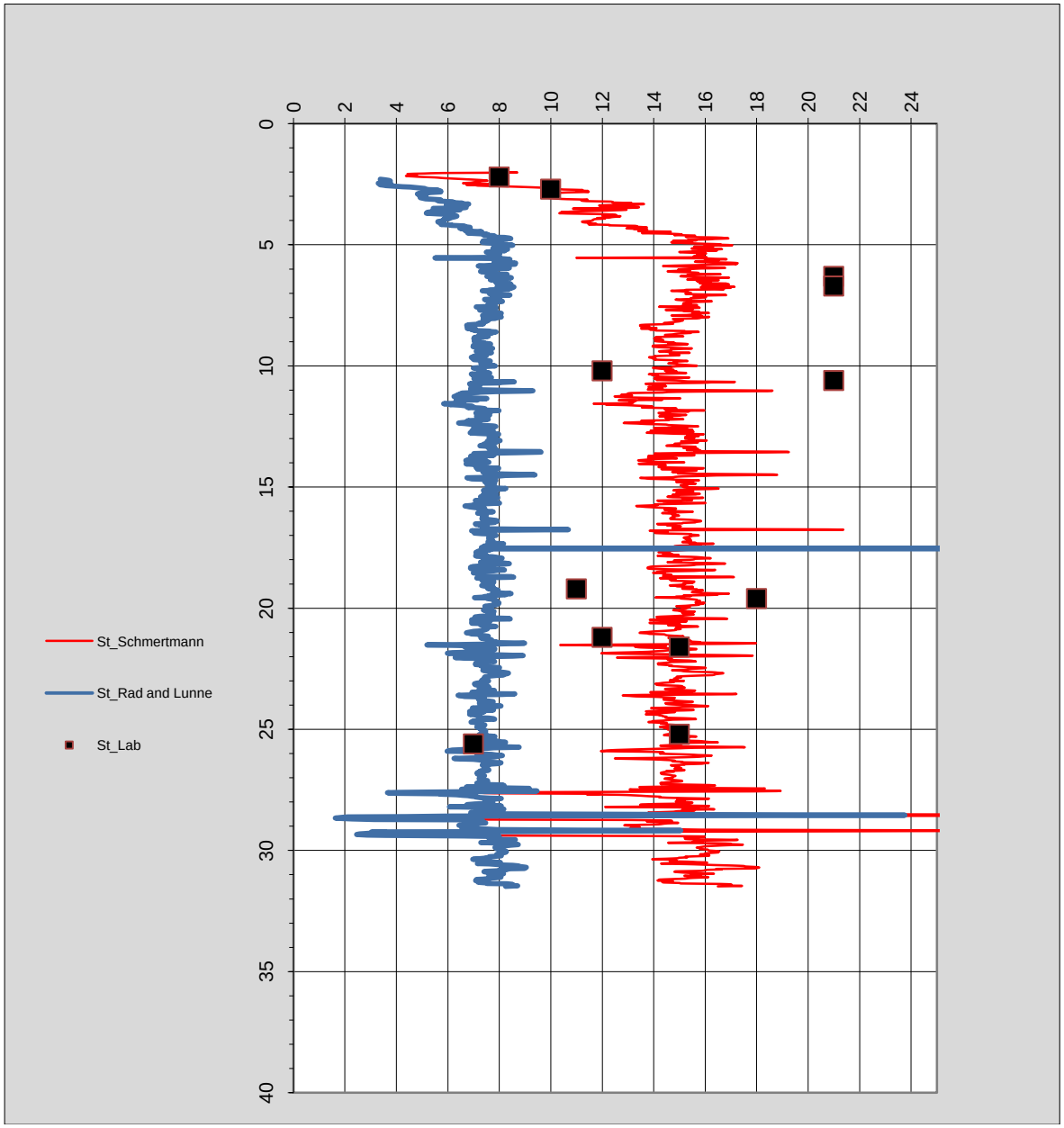


Figure 5.1.6 *Sensitivity with different parameters and laboratory results for Fallan BP2*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.1.5. Resistivity

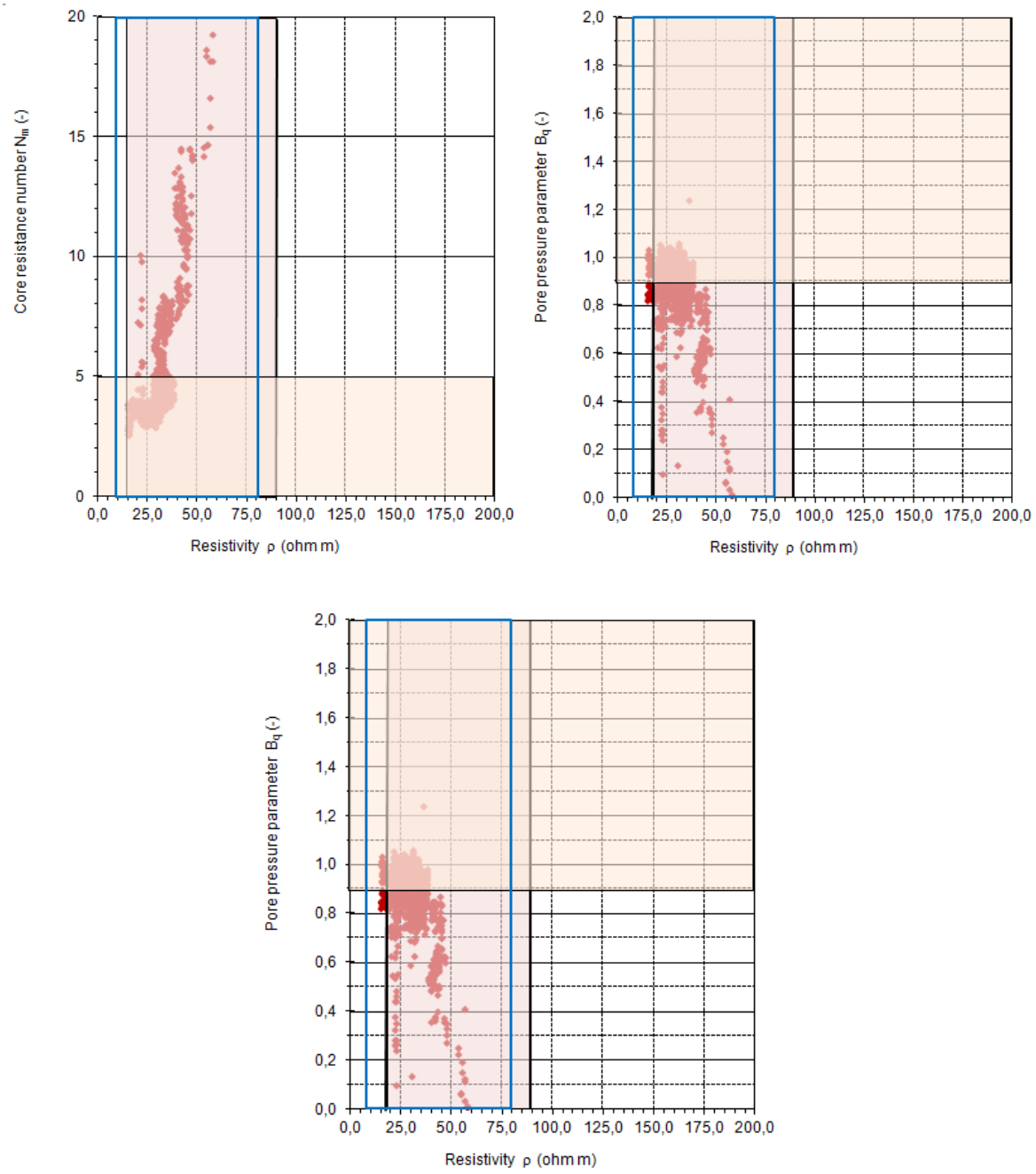


Figure 5.1.7 *Resistivity results in relation to N_m, R_f and B_q for Fallan BP2*

Comment:
Marked areas are related to soil classification tables, which are discussed in Chapter 10 of this project. Highlights represents values for quick clay.

5.2. Fallan - CPTU 2

5.2.1. Classification charts

After calculations following types of soils and profiles were acquired:

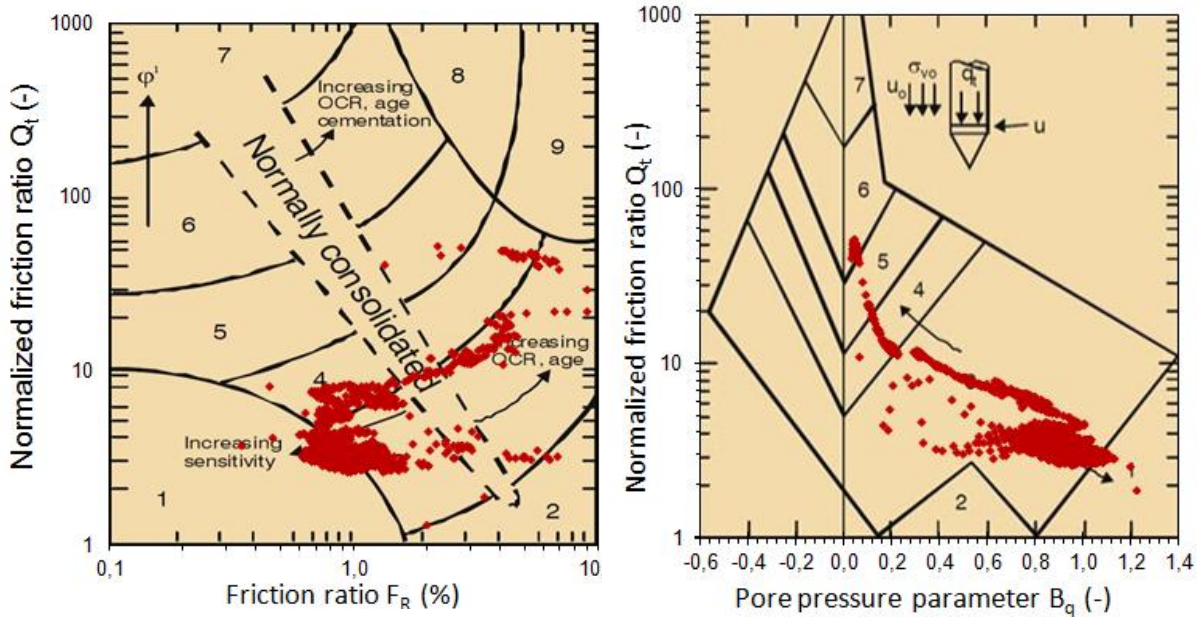


Figure 5.2.1 *Robertson classification charts (1990) for Fallan CPTu2*

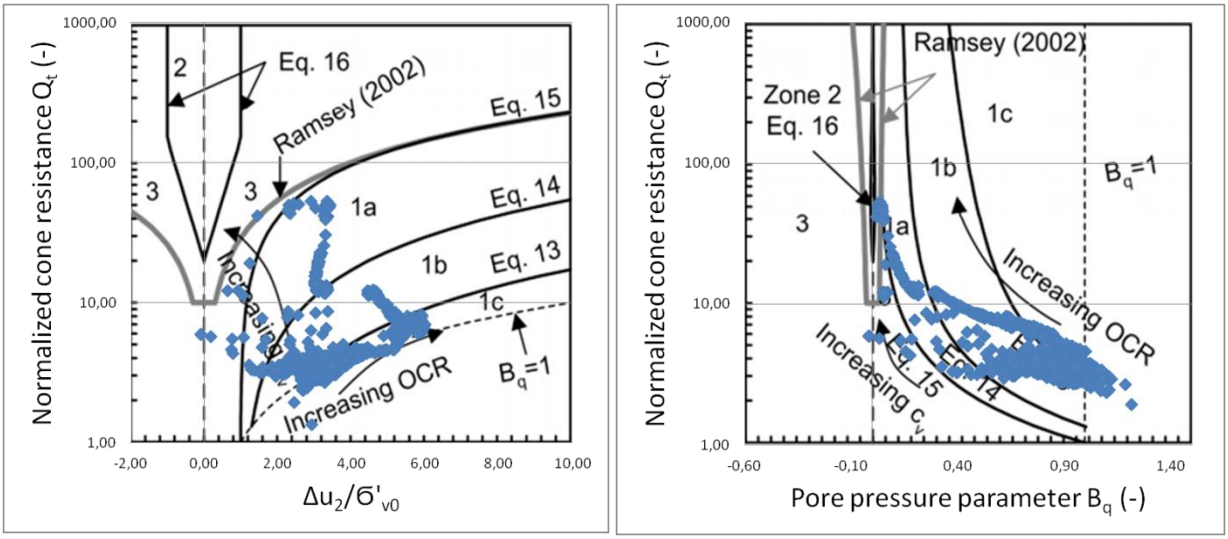


Figure 5.2.2 *Schneider classification charts (2008) for Fallan CPTu2*

Comment:
Sounding was ceased after encountering an obstacle at approx. depth of 35m.

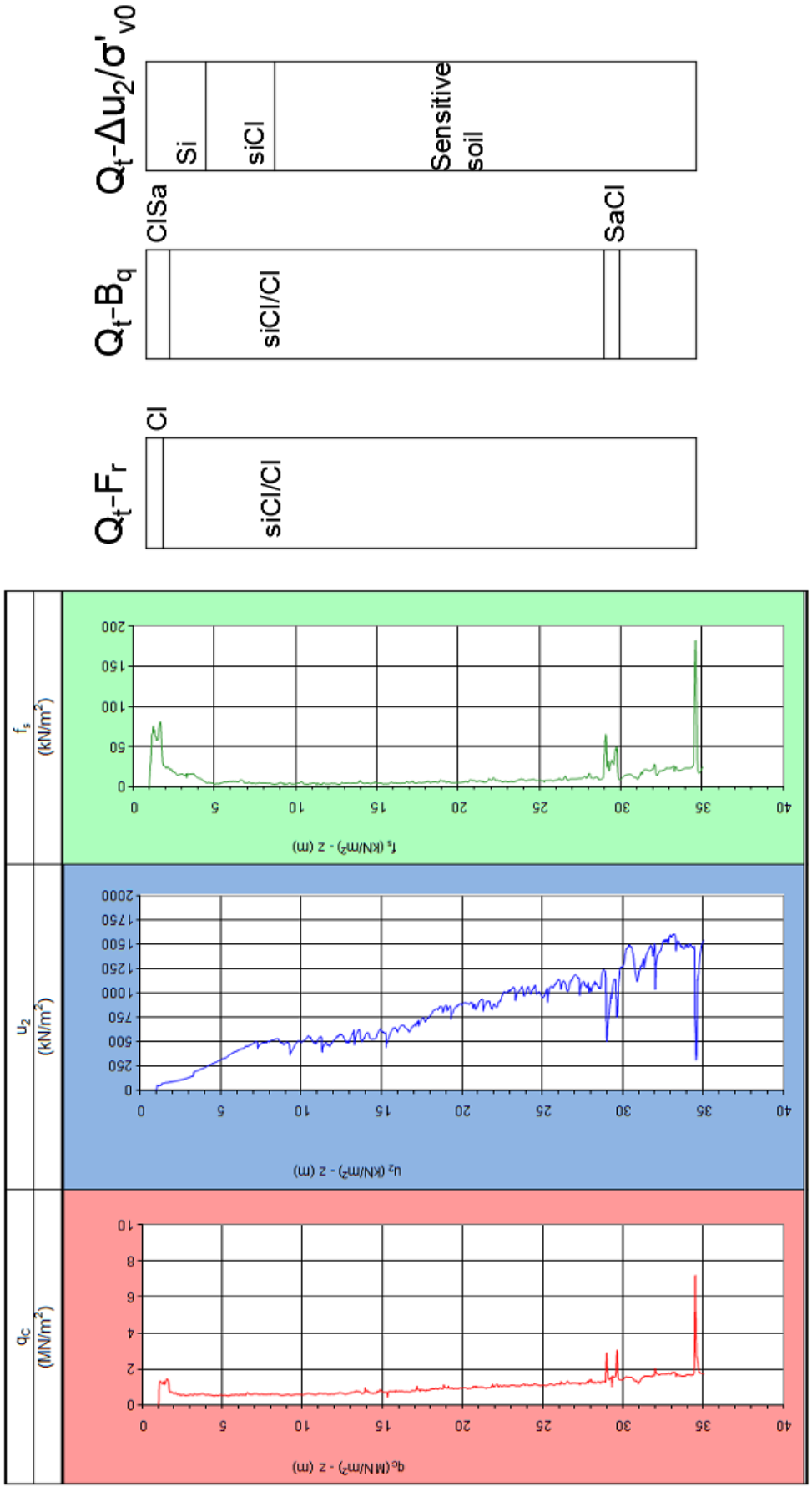


Figure 5.2.3 *Sounding parameters with interpreted profiles for Fallan CPTu2*

5.2.3. Undrained shear resistance

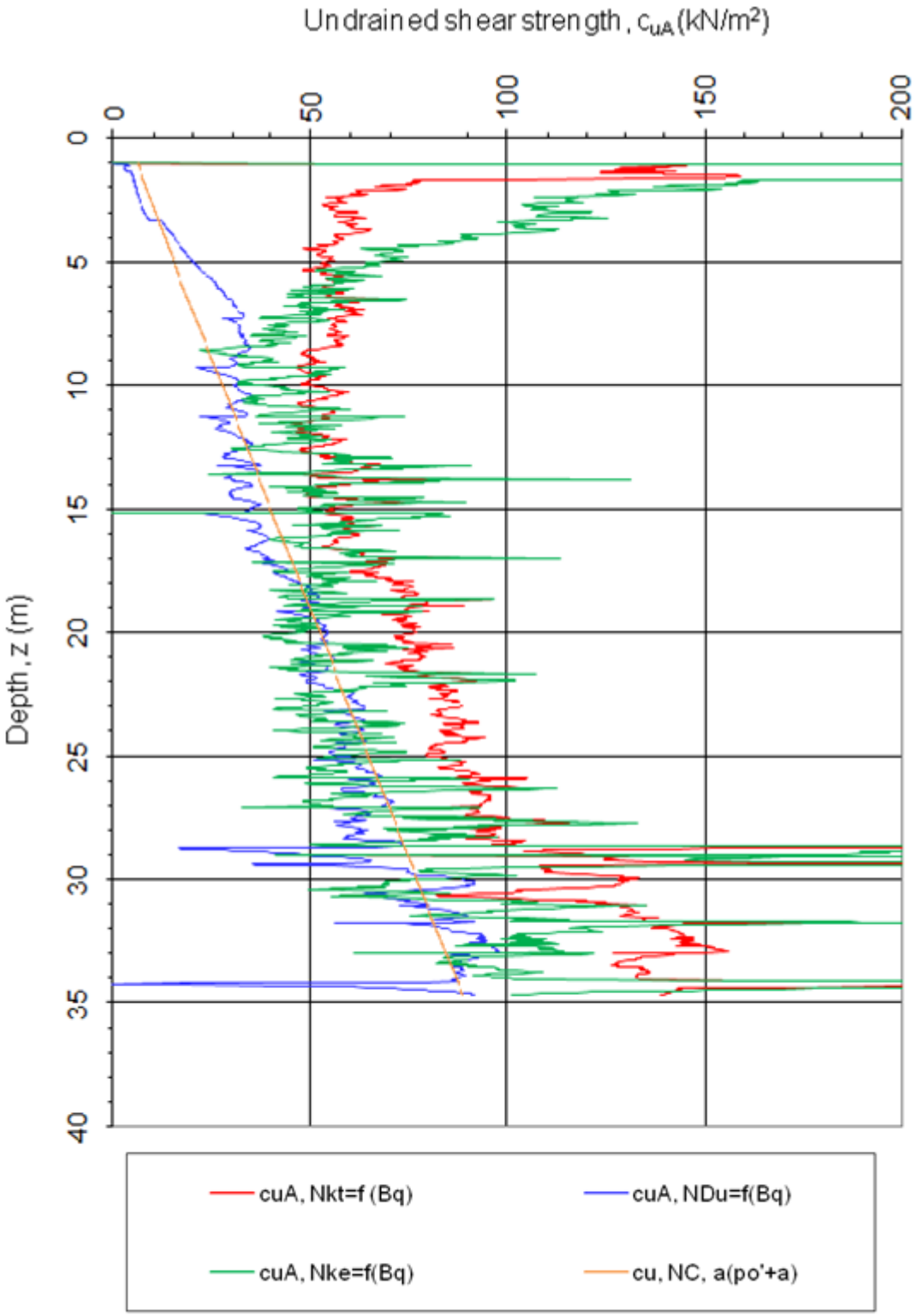


Figure 5.2.4 *Undrained shear resistance for interpreted results for Fallan CPTu2 results*

5.2.4. Sensitivity

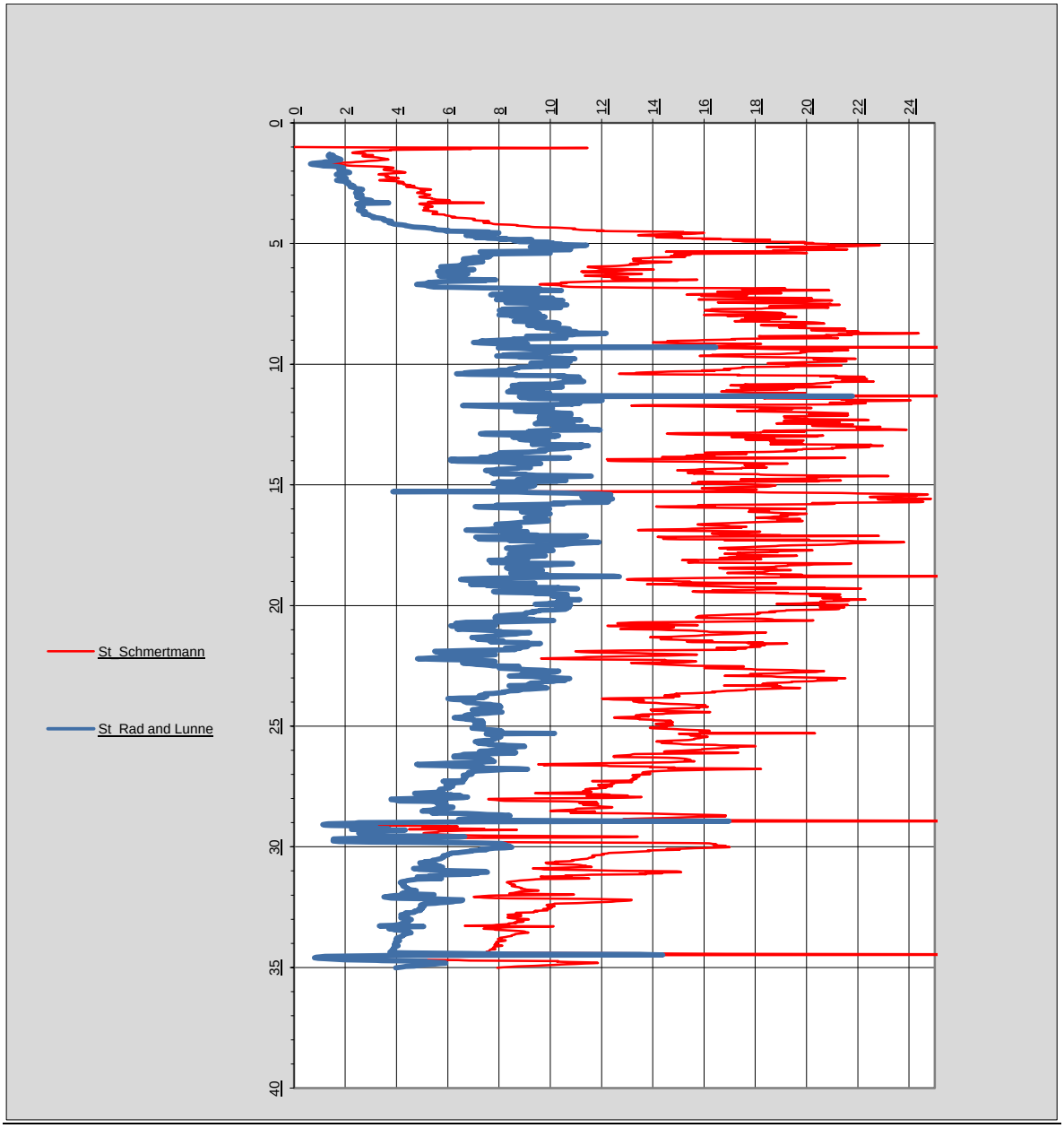


Figure 5.2.5 *Sensitivity with different parameters, results for Fallan CPTu2*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.3. Fallan - RCPTU 4

5.3.1. Classification charts

After calculations types of soils and profiles were acquired:

Figure 5.3.1 *Robertson classification chart (1990) for Fallan RCPTu4*

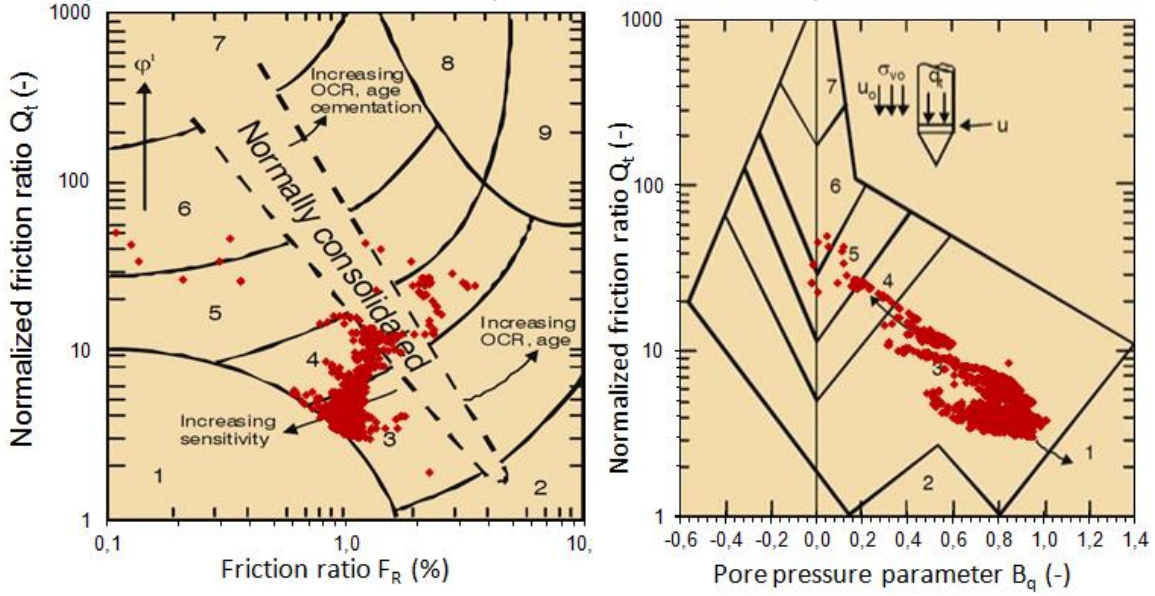


Figure 5.3.1 *Robertson classification chart (1990) for Fallan RCPTu4*

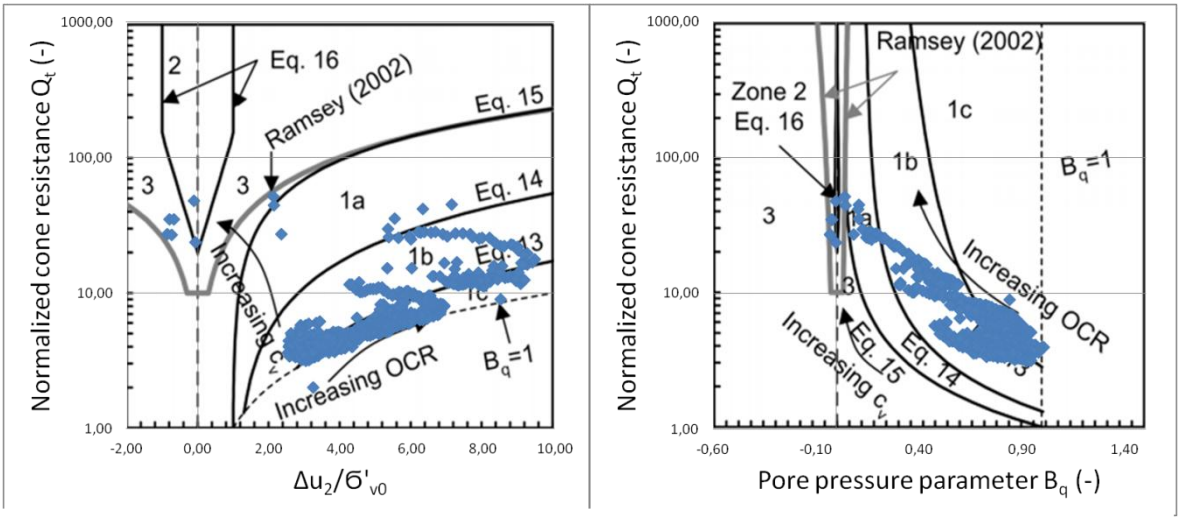


Figure 5.2.2 *Schneider classification charts (2008) for Fallan RCPTu4*

Comment:
Sounding was ceased after encountering an obstacle at approx. depth of 29m.

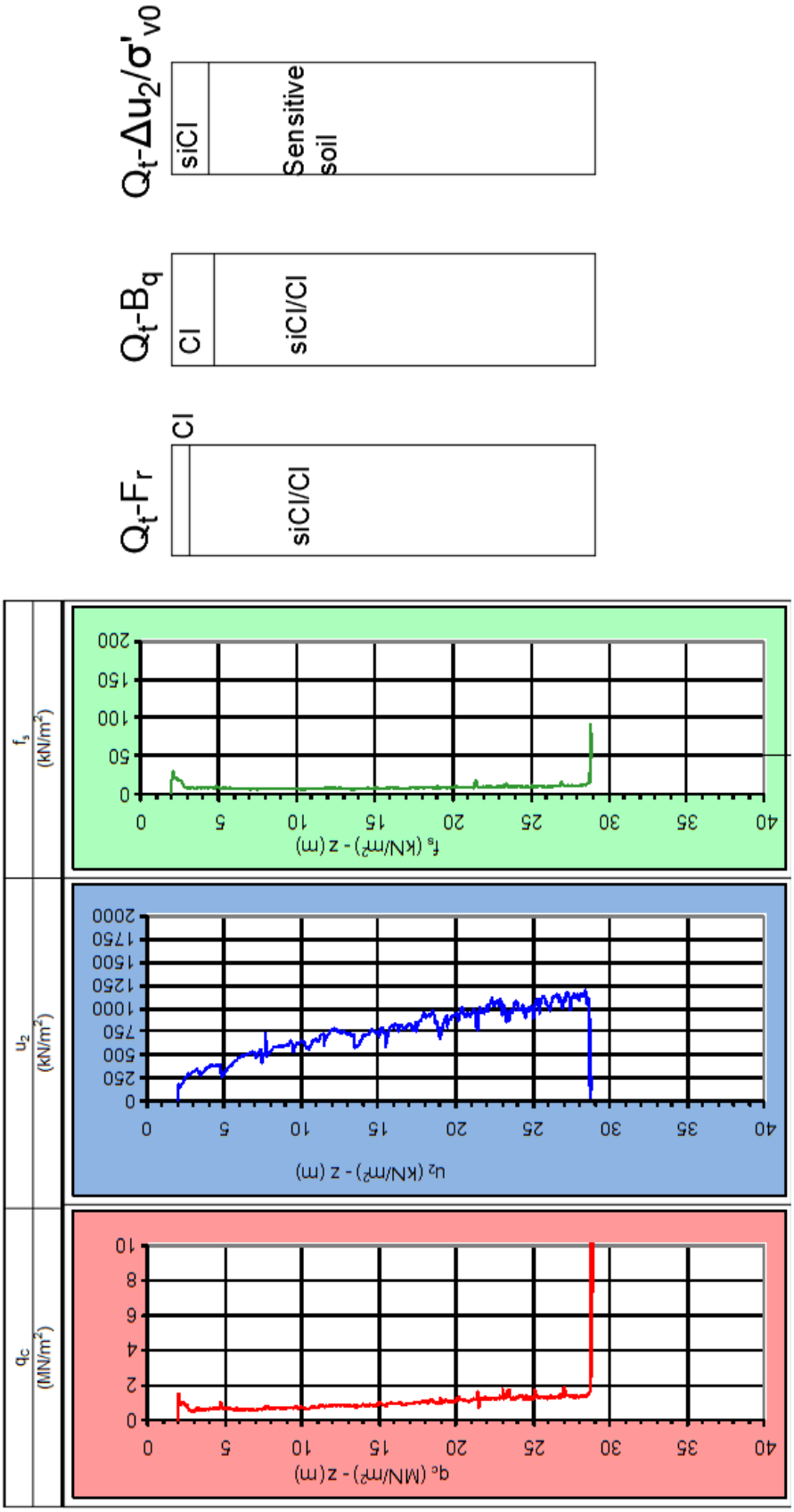


Figure 5.3.3 *Sounding parameters with interpreted profiles for Fallan RCPTu4*

5.3.2. Undrained shear resistance

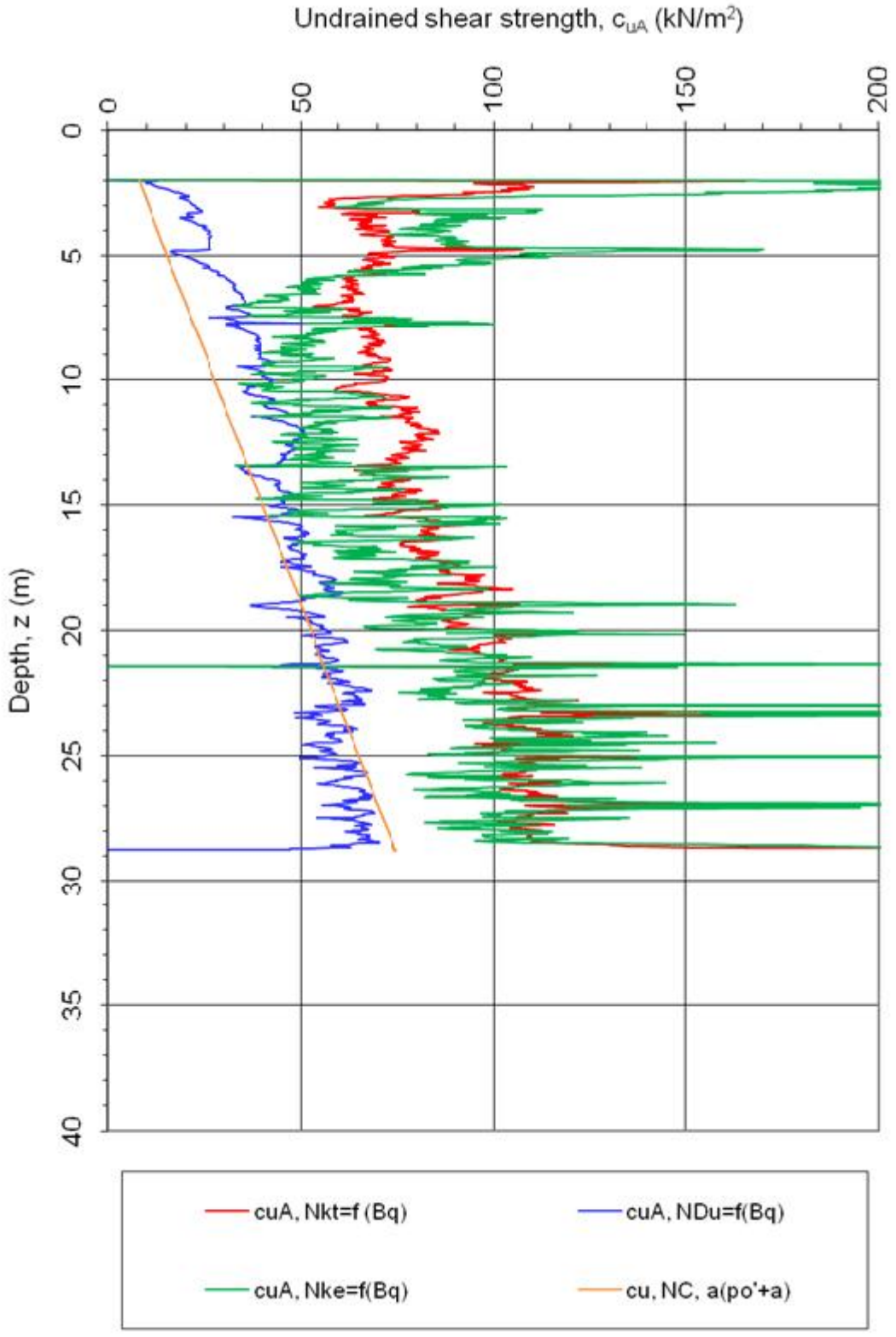


Figure 5.3.4 *Undrained shear resistance for interpreted results for Fallan RCPTu4 results*

5.3.3. Sensitivity

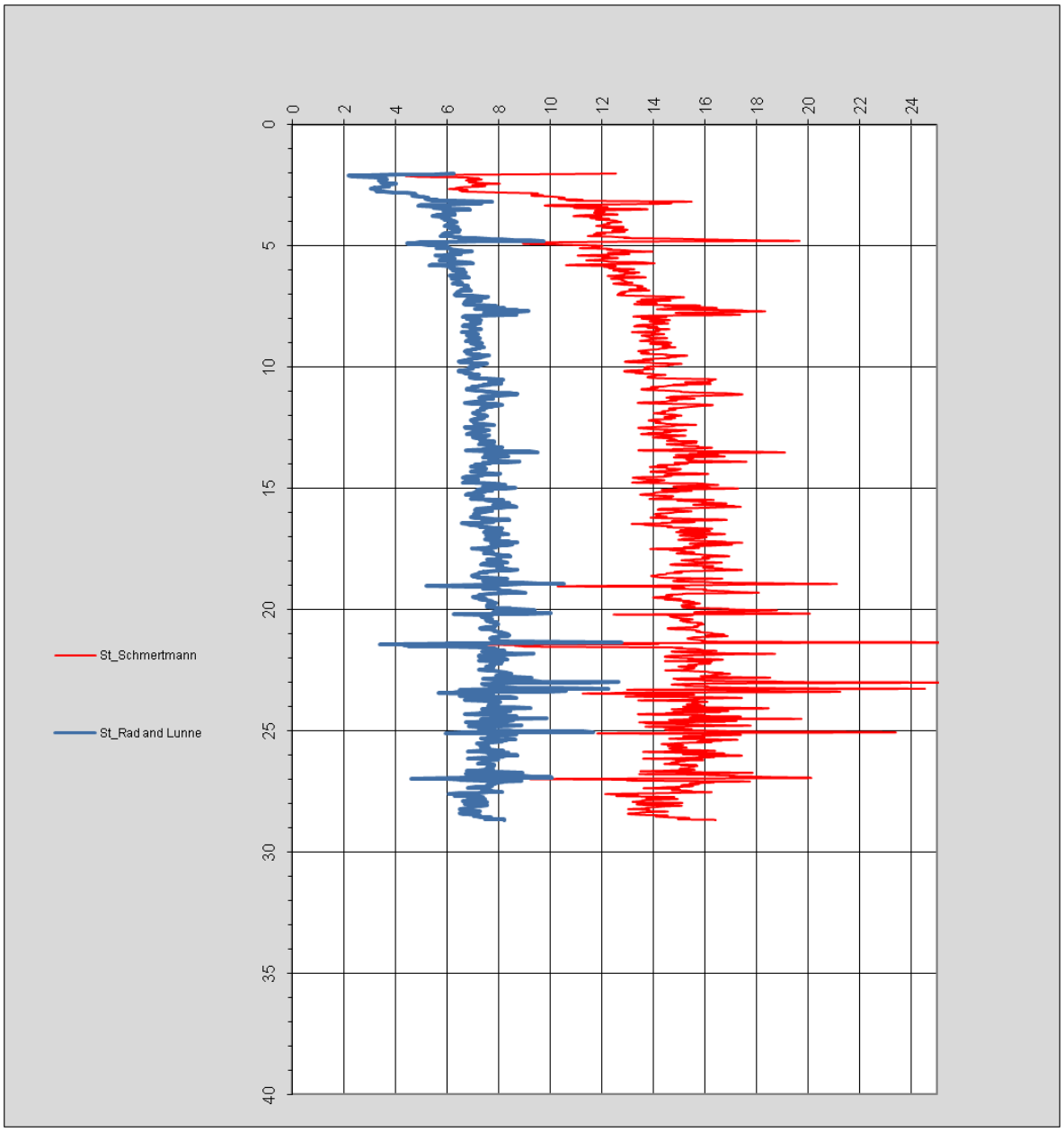


Figure 5.3.5 *Sensitivity with different parameters, results for Fallan RCPTu4*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.3.4. Resistivity

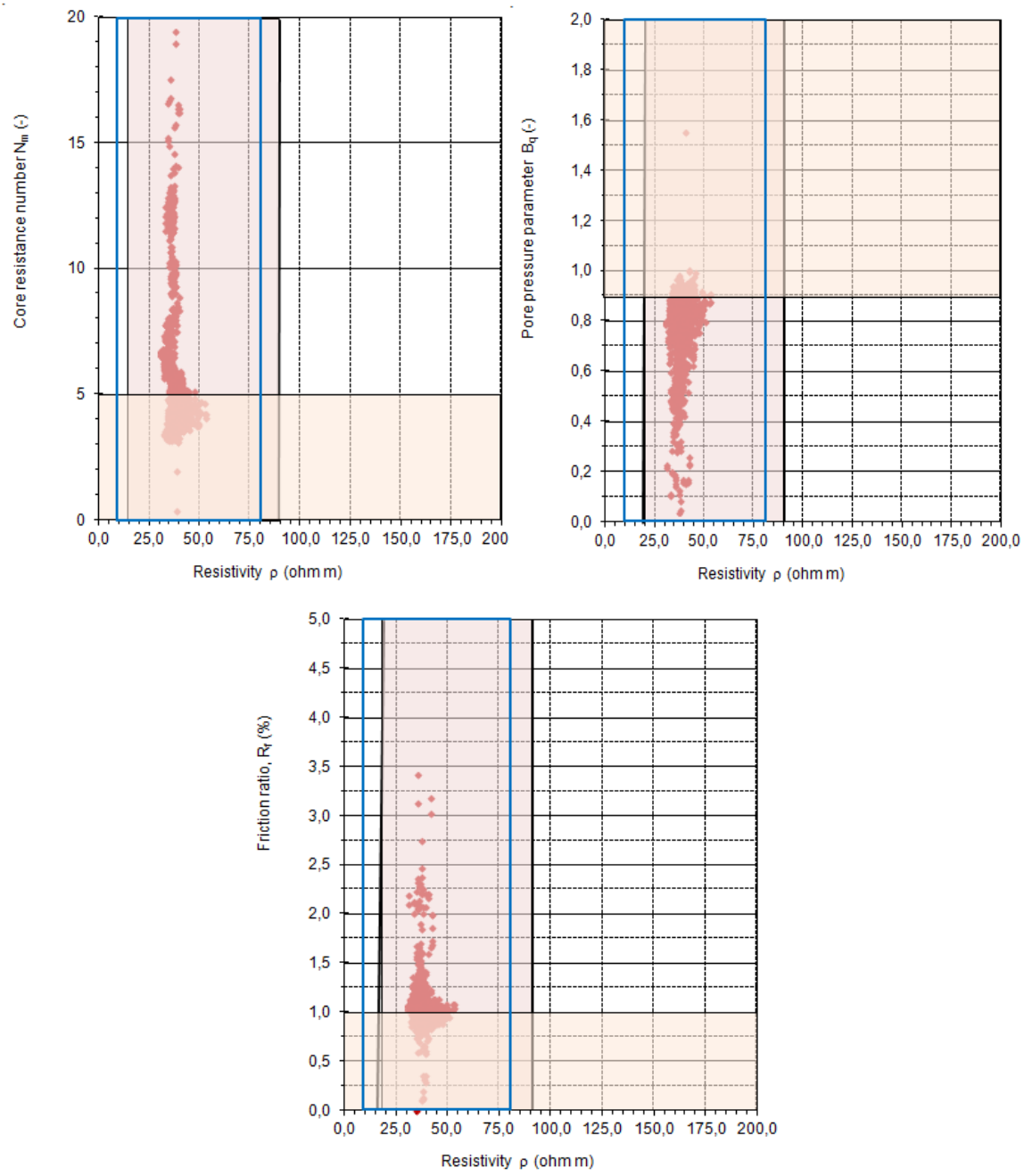


Figure 5.3.6 *Resistivity results in relation to N_m, R_f and B_q for Fallan RCPTu4*

Comment:
Marked areas are related to soil classification tables, which are in Chapter 9 of this project. Highlights represents values for quick clay.

5.4. Fallan - CPTU 5

5.4.1. Classification charts after Robertson

After calculations types of soils and profiles were acquired:

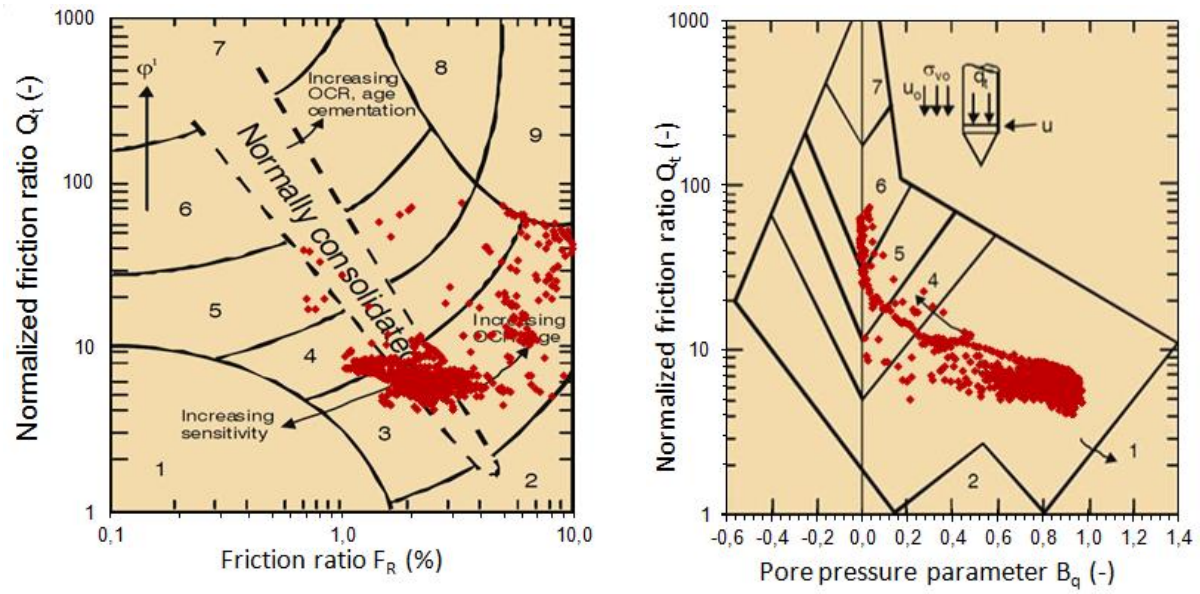


Figure 5.4.1 *Robertson classification charts (1990) for Fallan CPTu5*

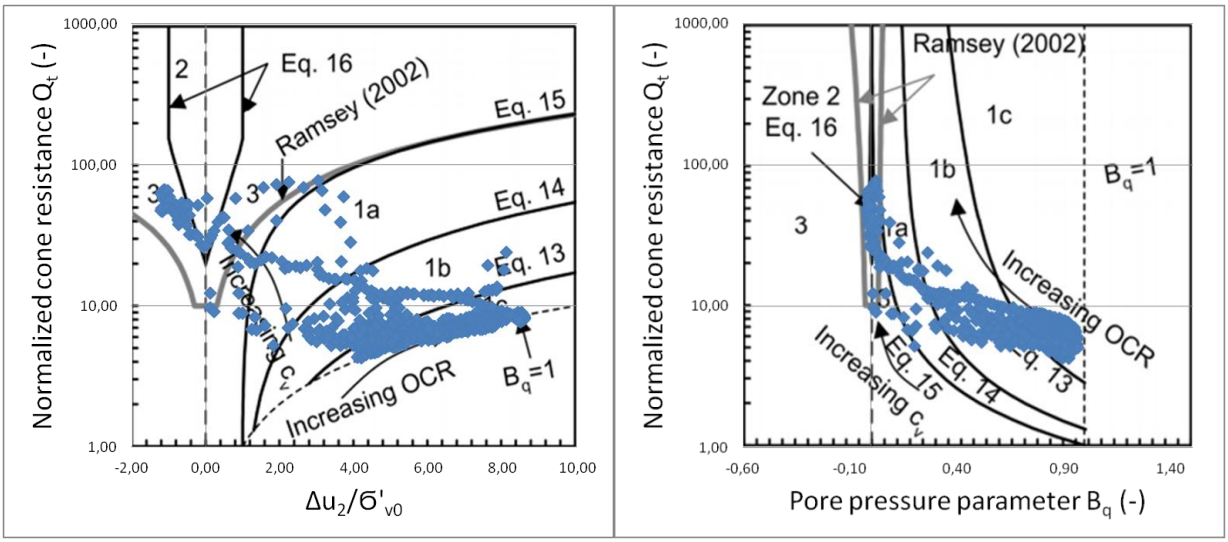


Figure 5.4.2 *Schneider classification charts (2008) for Fallan CPTu5*

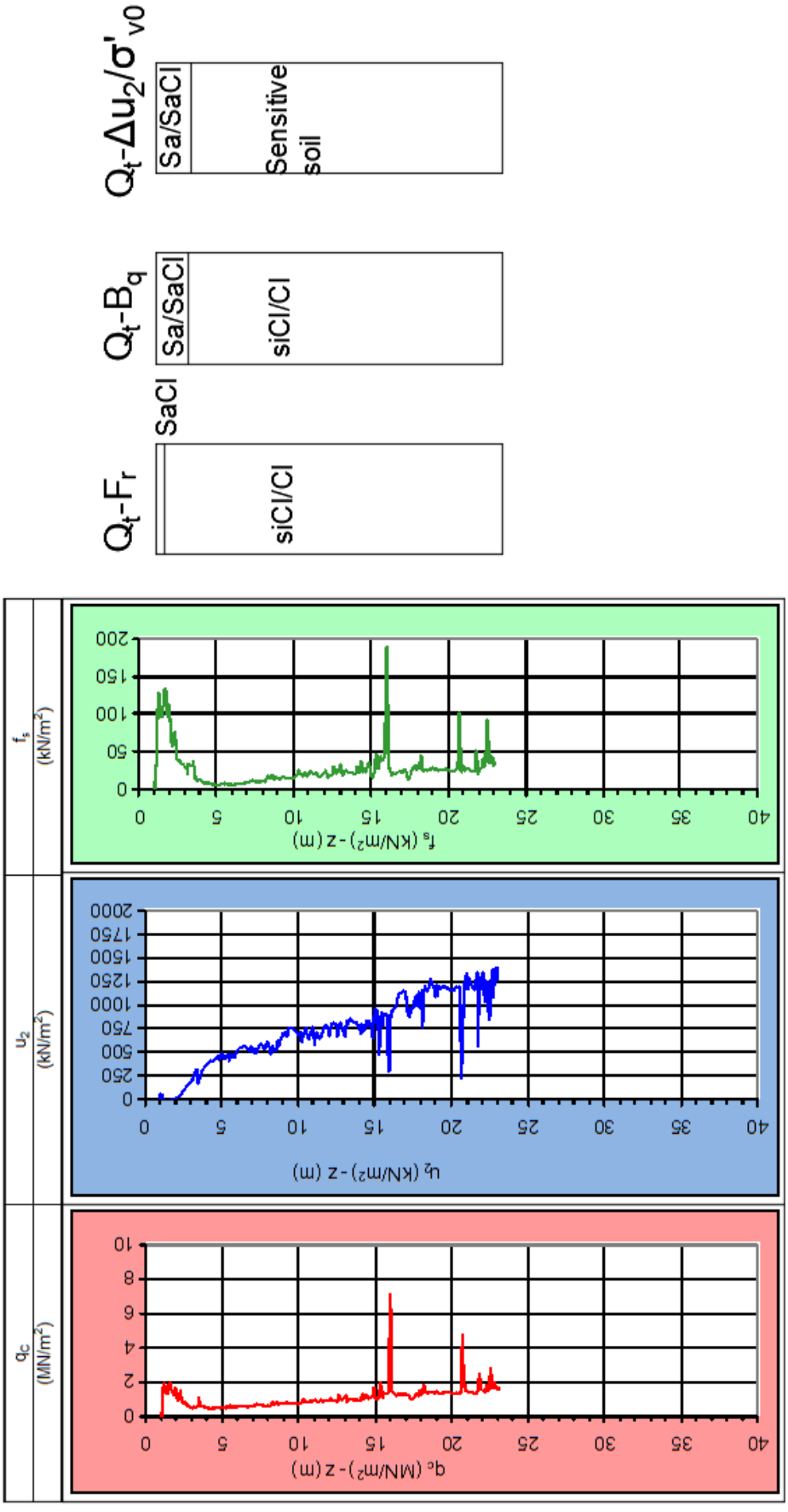


Figure 5.4.3 *Sounding parameters with interpreted profiles for Fallan CPTu5*

5.4.2. Undrained shear resistance

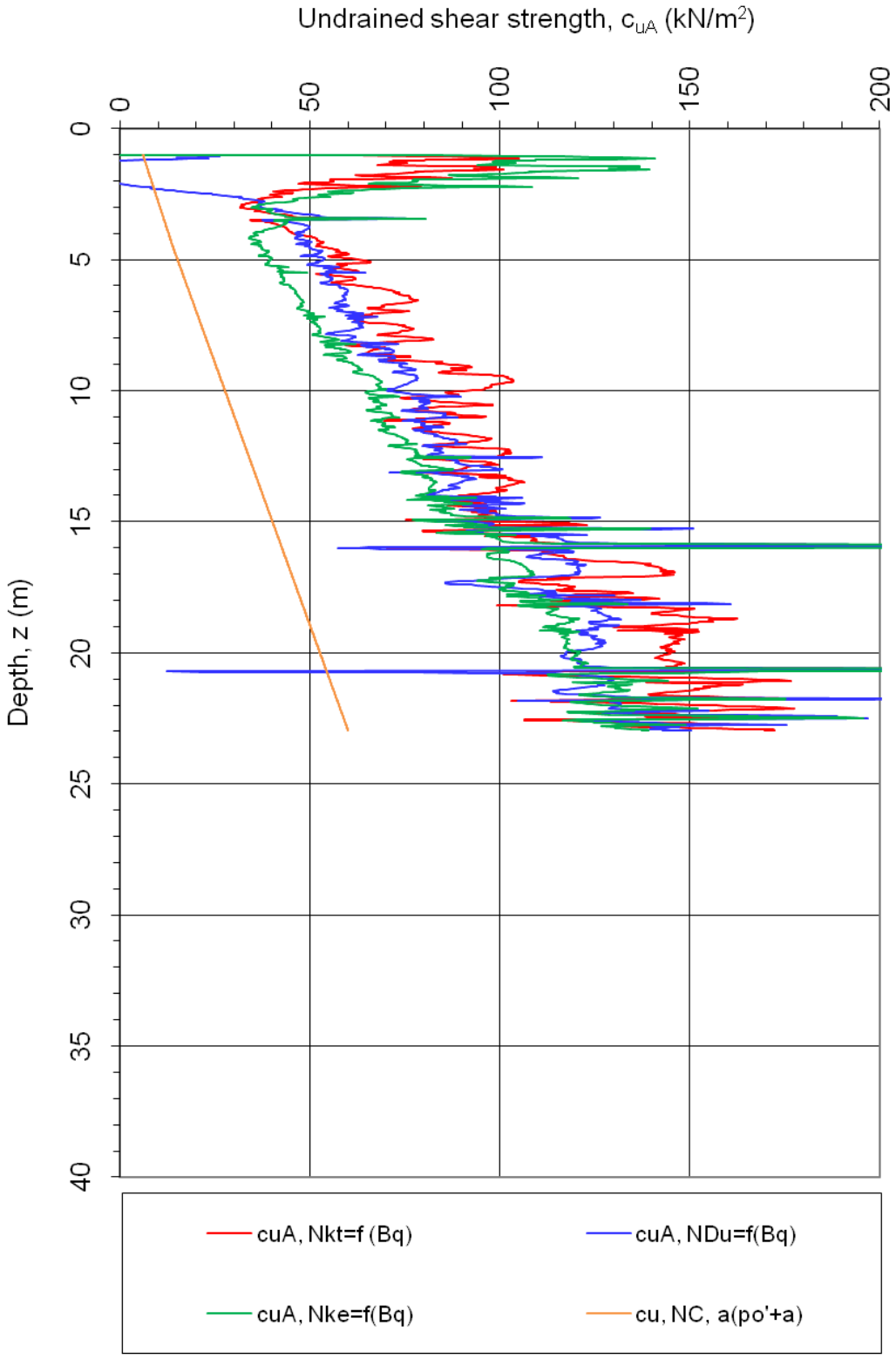


Figure 5.4.4 *Undrained shear resistance for interpreted results for Fallan CPTu5 results*

5.4.3. Sensitivity

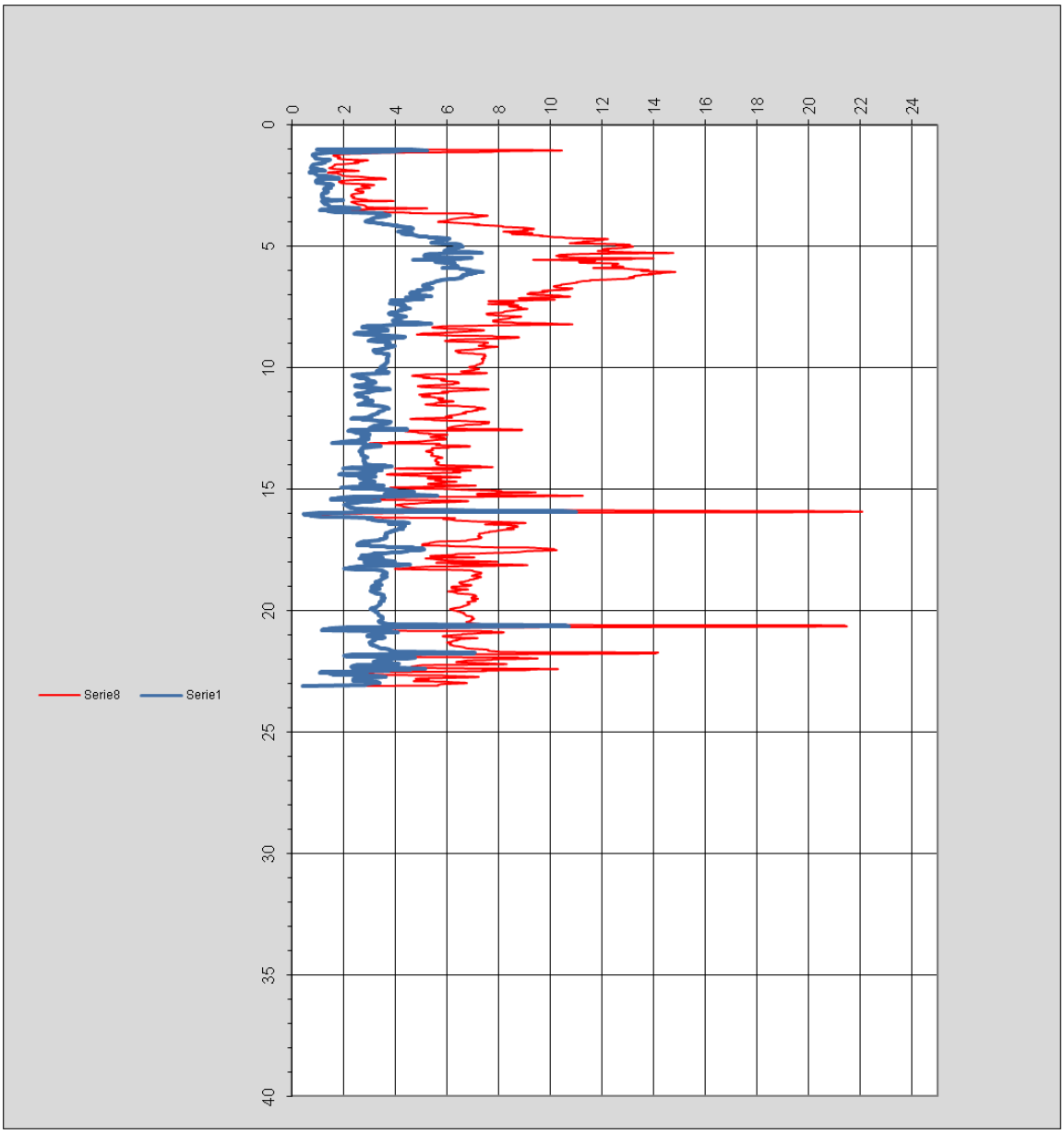


Figure 5.4.5 *Sensitivity with different parameters, results for Fallan CPTu5*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.5. Fallan - CPTU 7

5.5.1. Classification charts

After calculations types of soils and profiles were acquired:

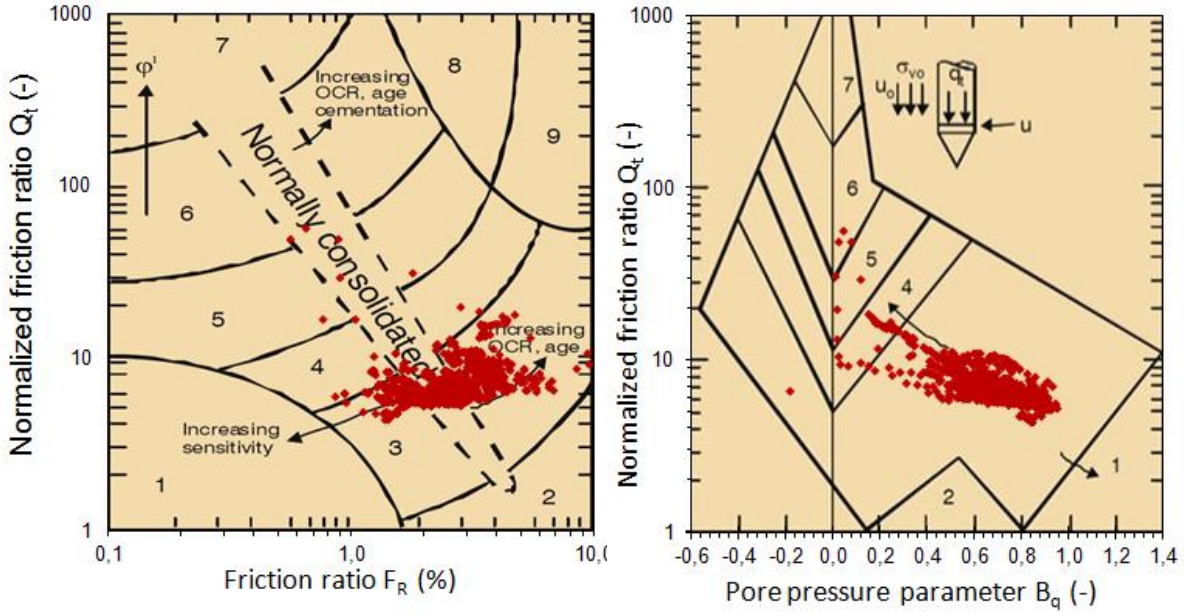


Figure 5.5.1 *Robertson classification charts (1990) for Fallan CPTu7*

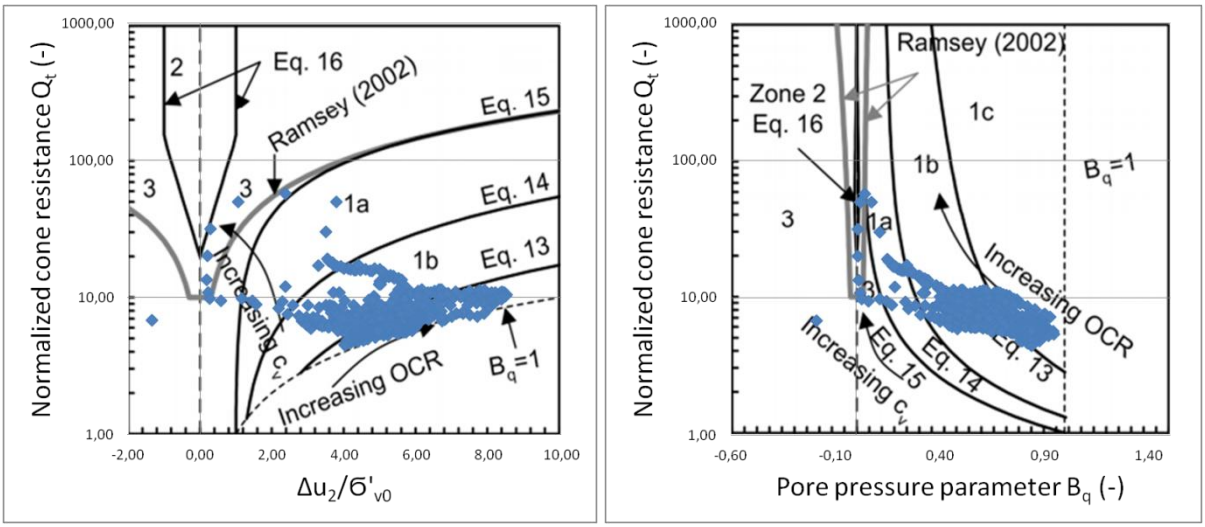


Figure 5.5.2 *Schneider classification charts (2008) for Fallan CPTu7*

Comment:
Sounding was ceased after encountering an obstacle at approx. depth of 35m.

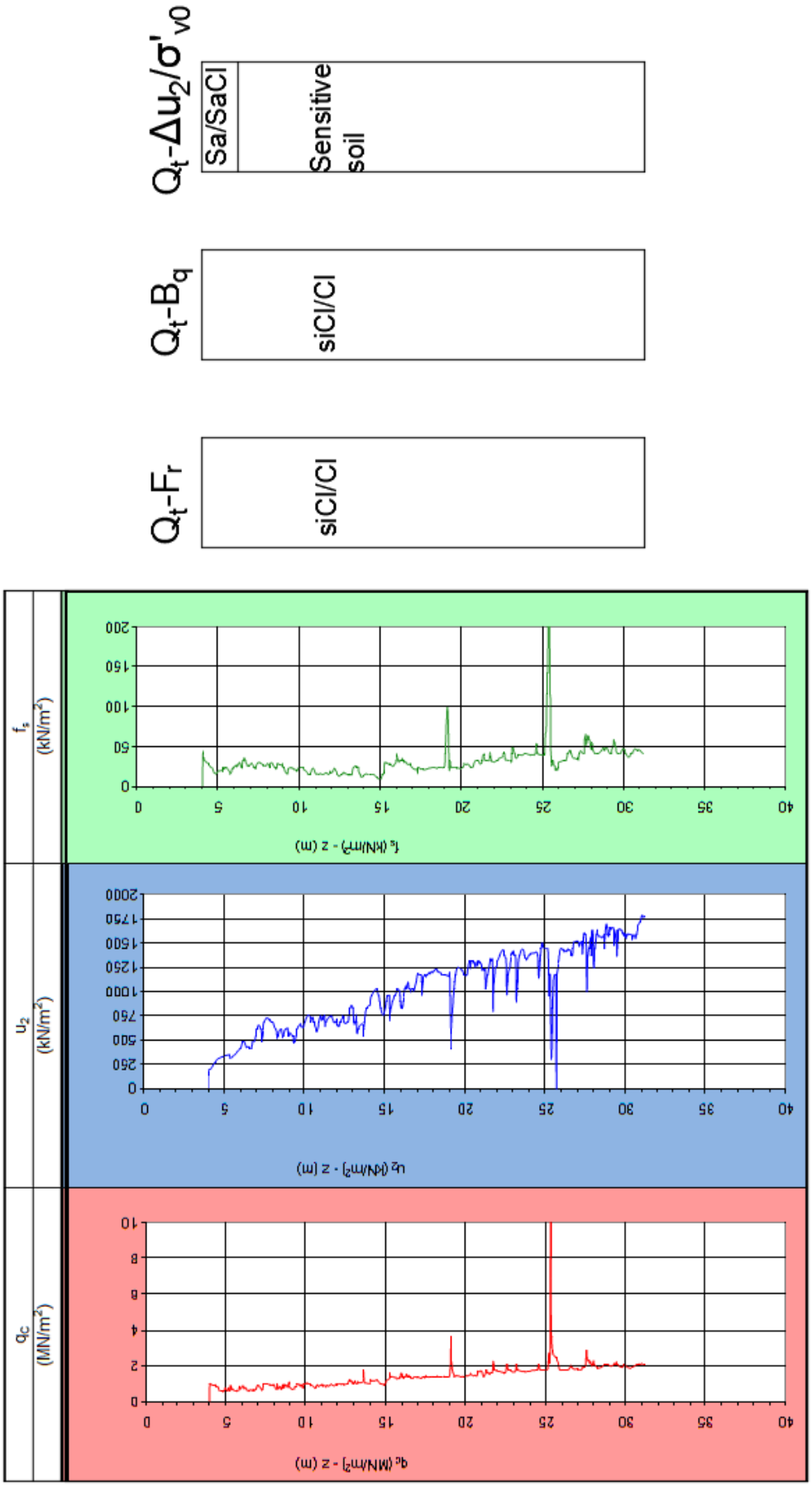


Figure 5.5.3 *Sounding parameters with interpreted profiles for Fallan CPTu7*

5.5.2. Undrained shear resistance

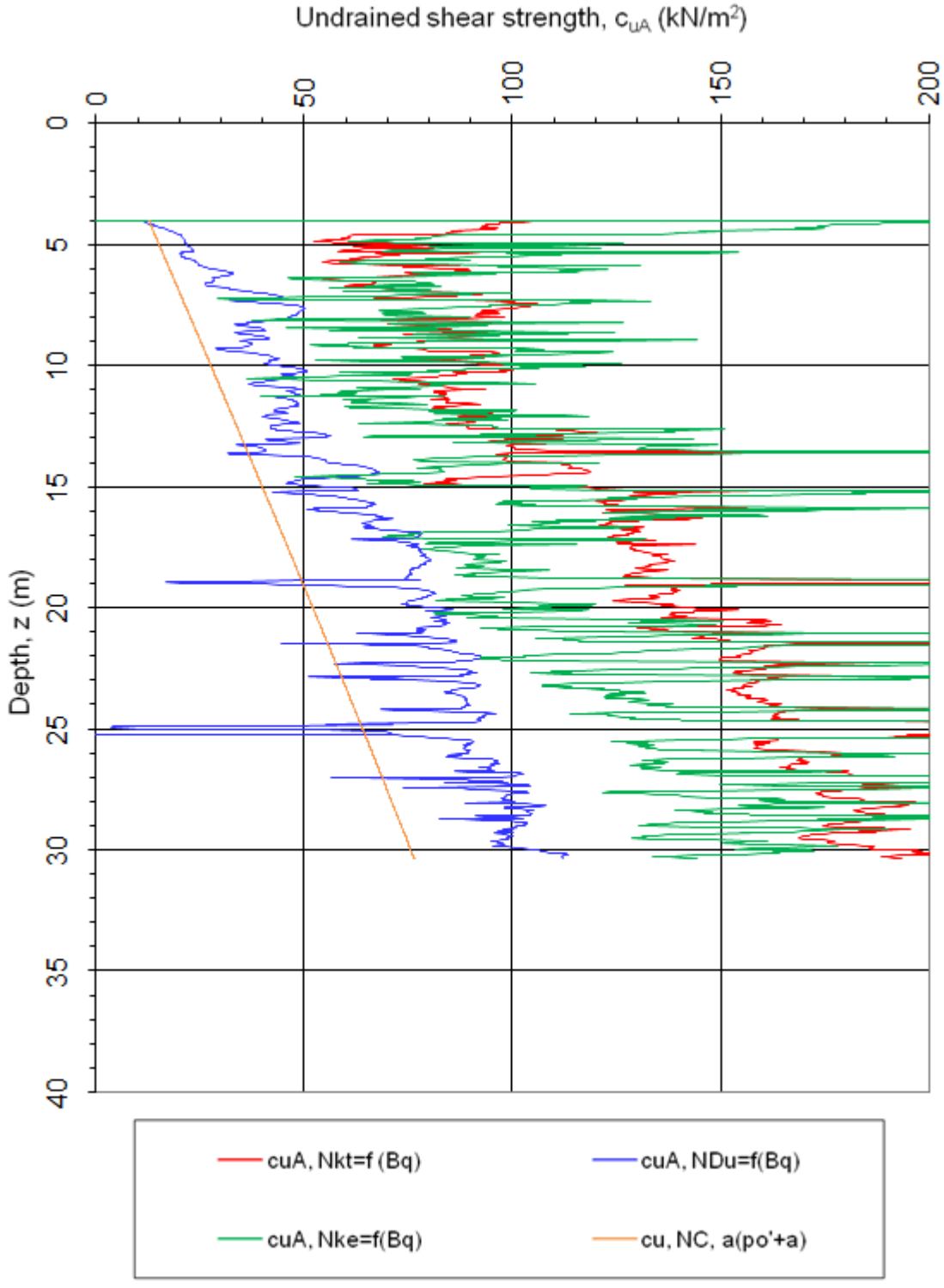


Figure 5.5.4 *Undrained shear resistance for interpreted results for Fallan CPTu7 results*

5.5.3. Sensitivity

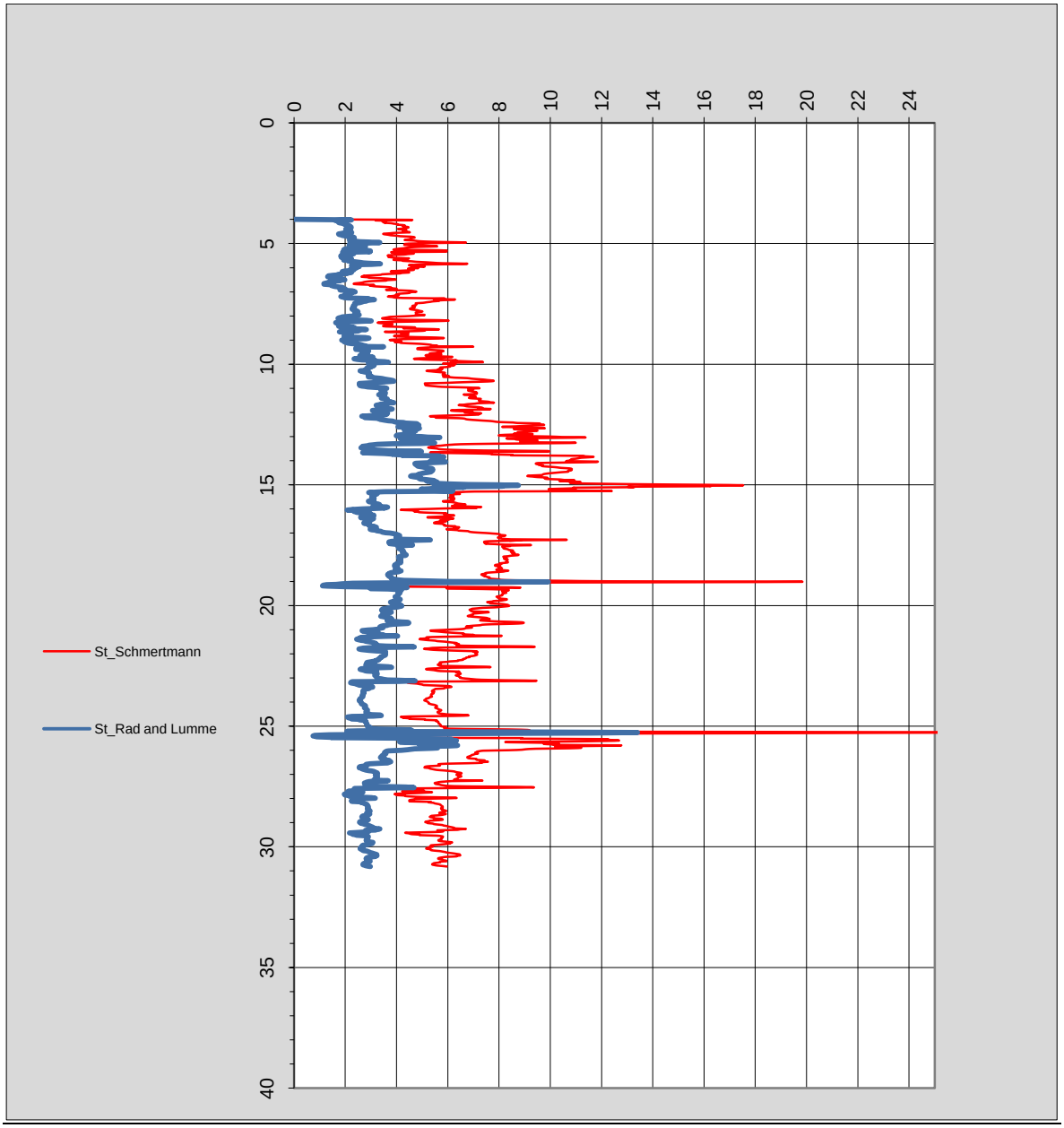


Figure 5.5.5 *Sensitivity with different parameters, results for Fallan CPTu7*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.6. Klett - RCPTU S1

5.6.1. Boring location plan

This section presents data from Klett investigation site. Laboratory data are available only for soundings S1 and S2. All acquired results will be discussed in chapters 8 and 9.

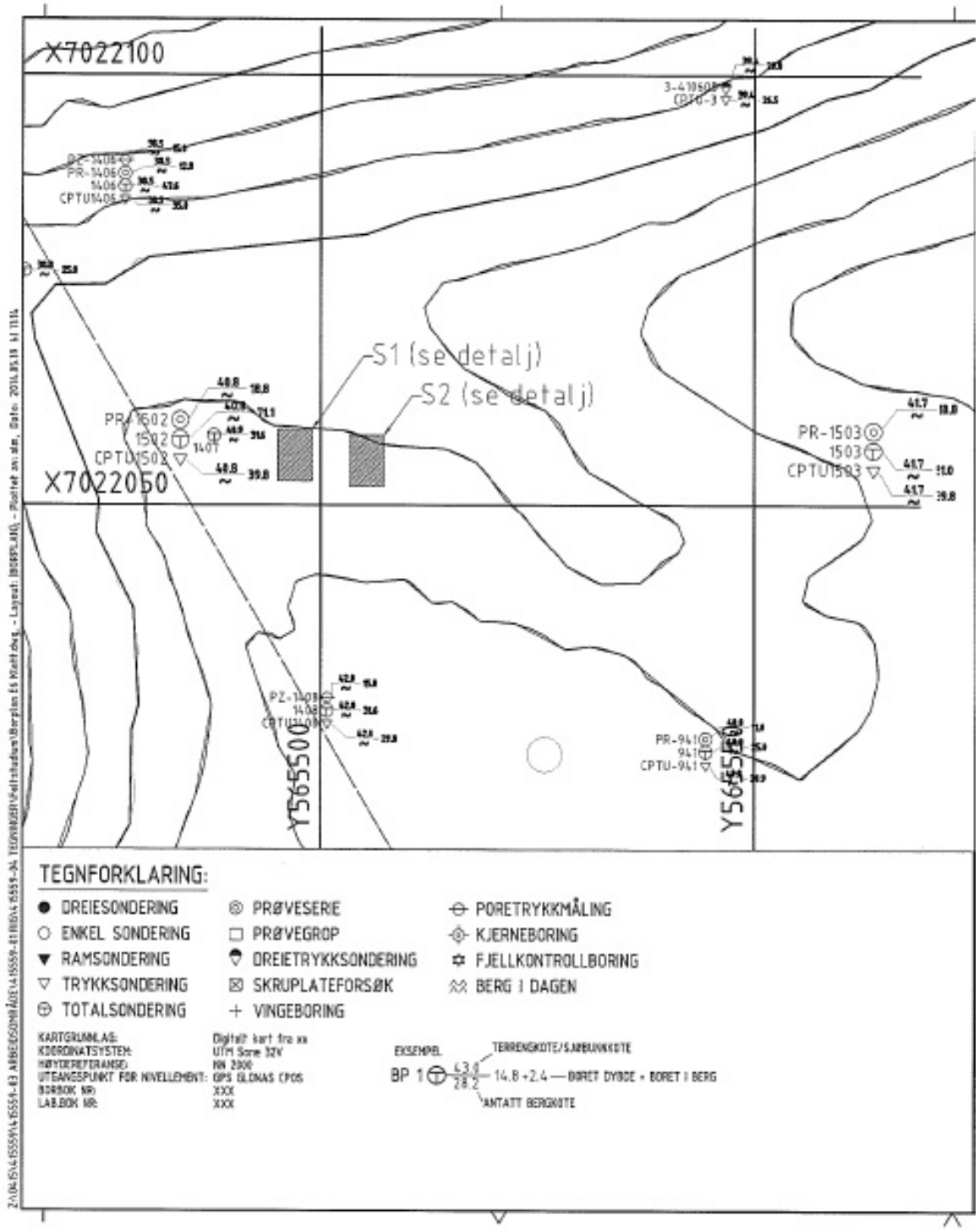


Figure 5.6.1 *Borings and soundings at Klett investigation site*

5.6.2. Classification charts

After calculations following types of soils and profiles were acquired:

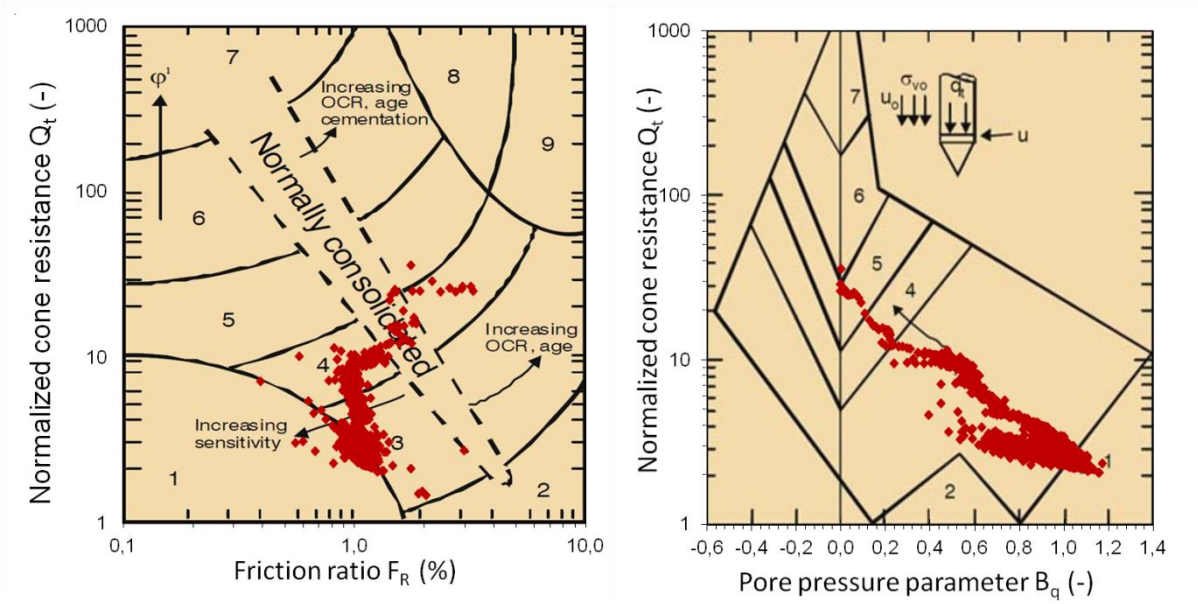


Figure 5.6.2 *Robertson classification chart (1990) for Klett RCPTu S1*

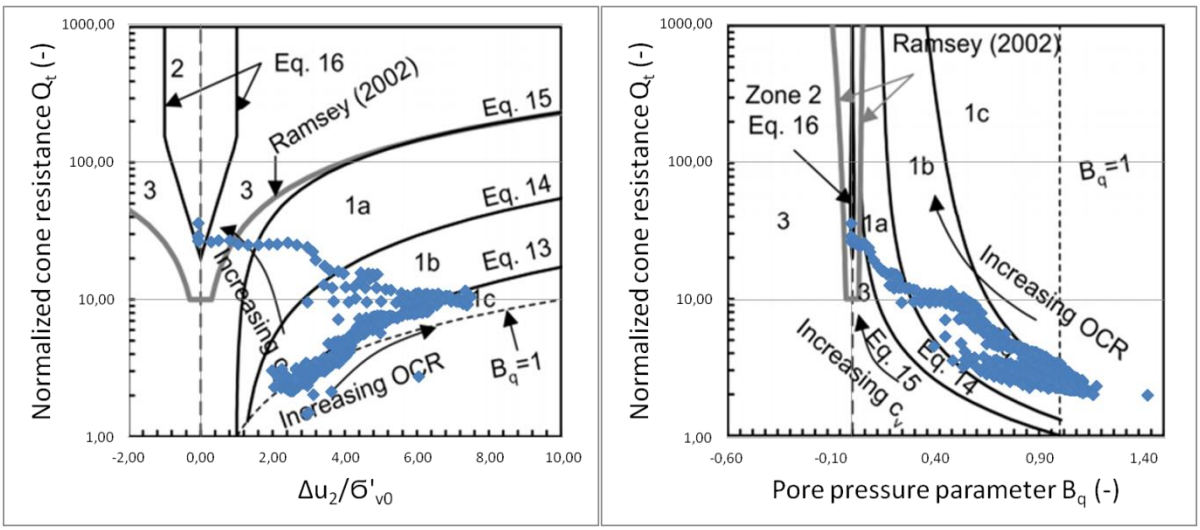


Figure 5.6.3 *Schneider classification charts (2008) for Klett RCPTu S1*

Comment:

Results plotted on classification charts presents a gradual change of soil with the depth in the profile - few initial readings indicates layer of sandy clay near the surface. Soil profile for this investigation site is verified with laboratory data - samples were taken from S1 and S2 points. Initial 4-8 meters layer is silty clay ($St \approx 5$), then it transits into sensitive and finally to quick clay.

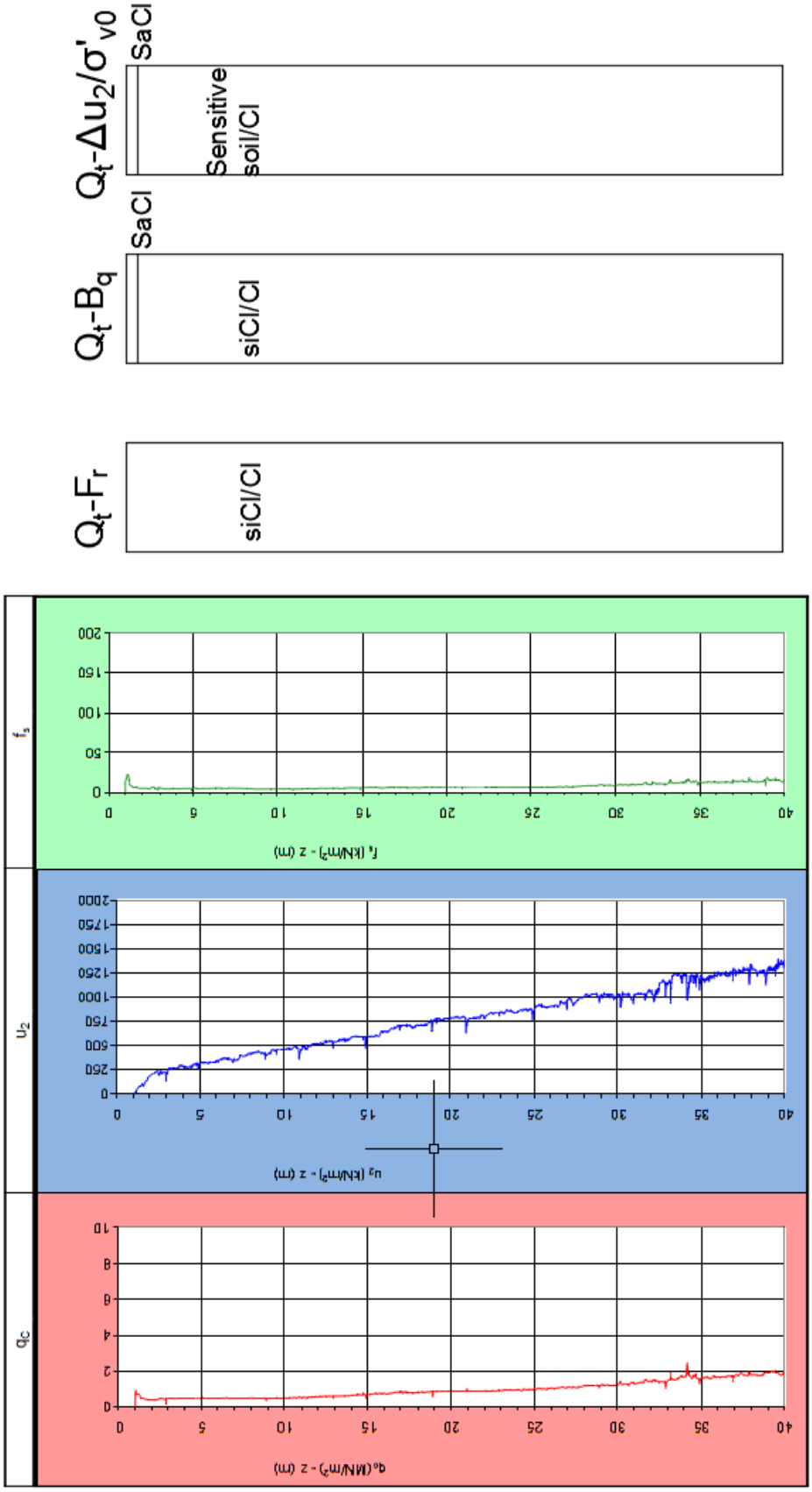


Figure 5.6.4 *Sounding parameters with interpreted profiles for Klett RCPTu S1*

5.6.3. Undrained shear resistance

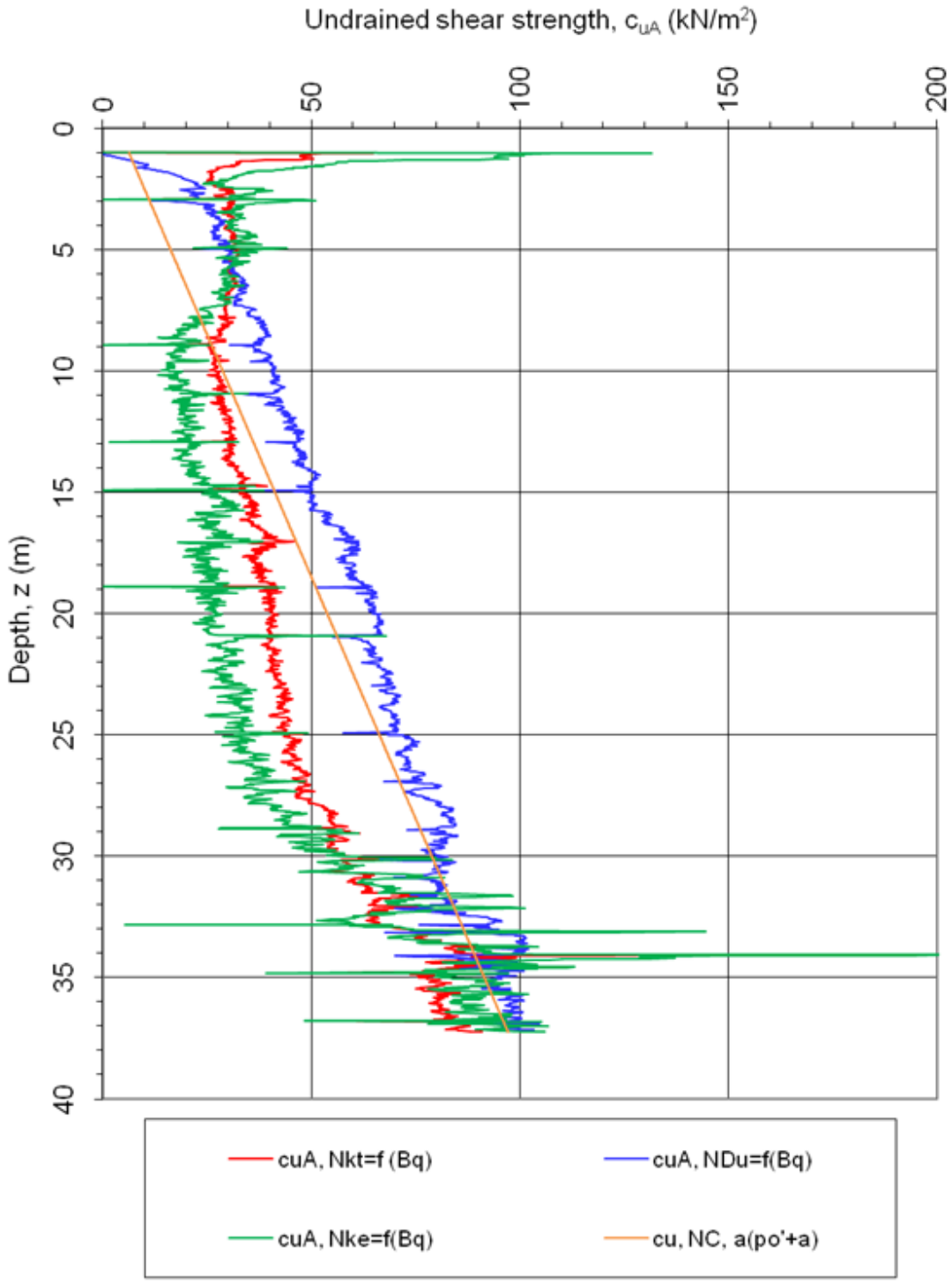


Figure 5.6.5 *Undrained shear resistance for interpreted results results for Klett RCPTu S1*

5.6.4. Sensitivity

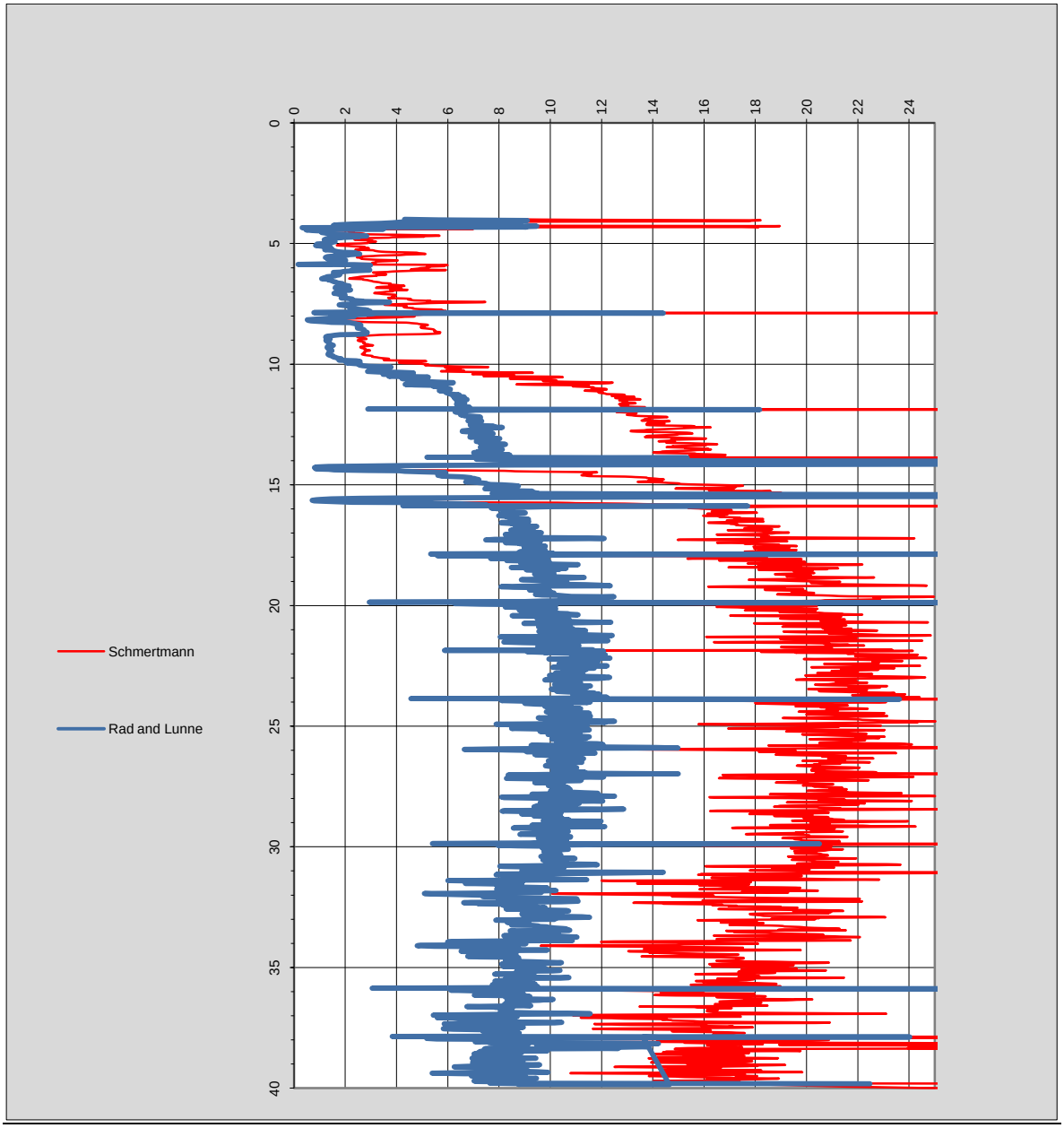


Figure 5.6.6 *Sensitivity with different parameters, results for Klett RCPTu S1*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.6.5. Resistivity

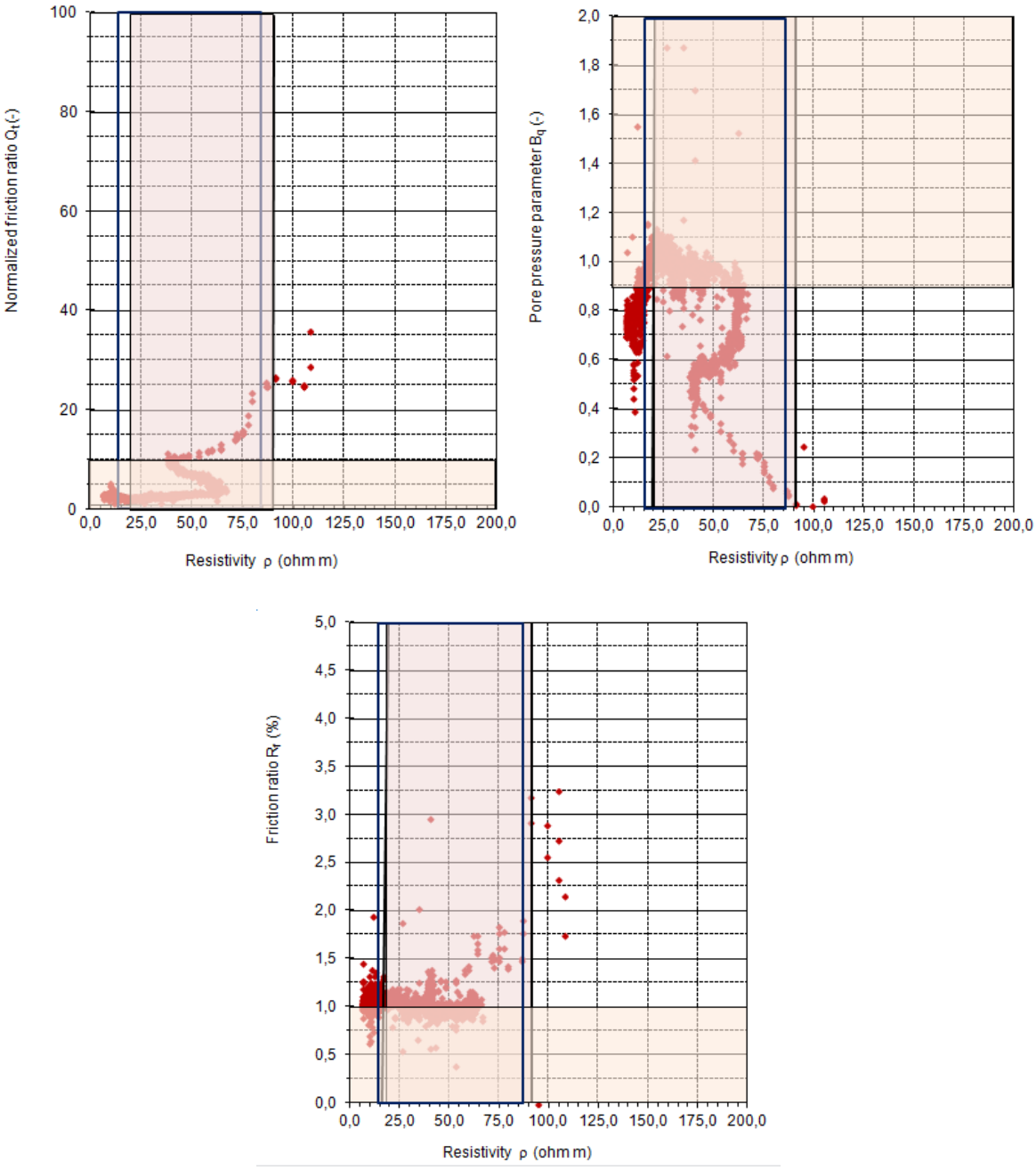


Figure 5.6.7 *Resistivity results in relation to N_m, R_f and B_q for Klett RCPTu S1*

Comment:
Marked areas are related to soil classification tables, which are in Chapter 9 of this project. Highlights represents values for quick clay.

5.7. Klett - RCPTU S2

5.7.1. Classification charts

After calculations types of soils and profiles were acquired:

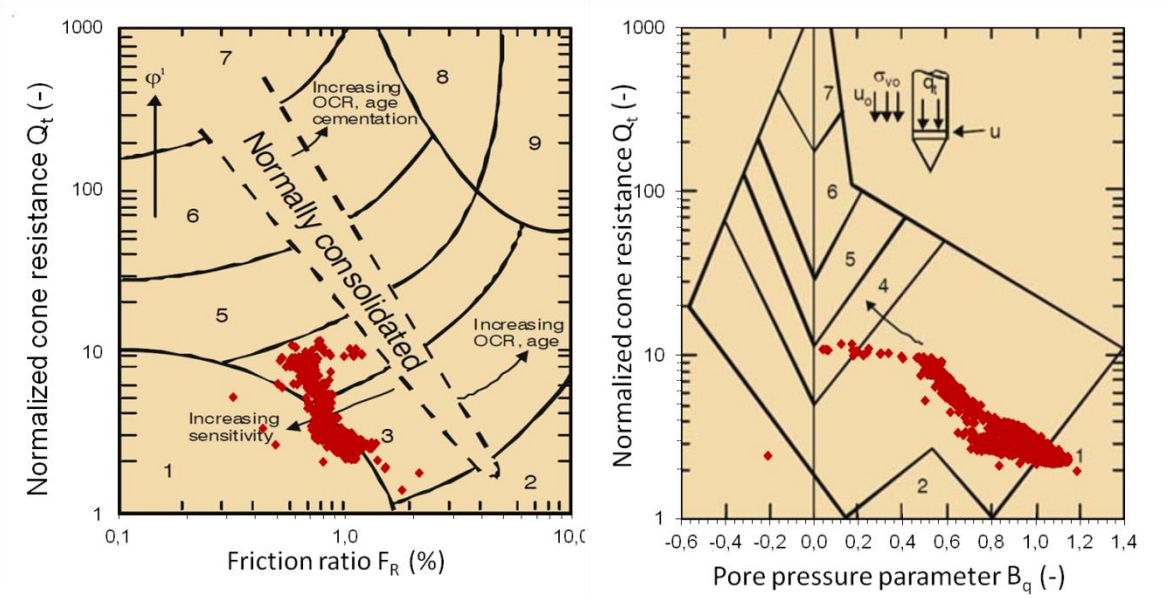


Figure 5.7.1 *Robertson classification chart (1990) for Klett RCPTu S2*

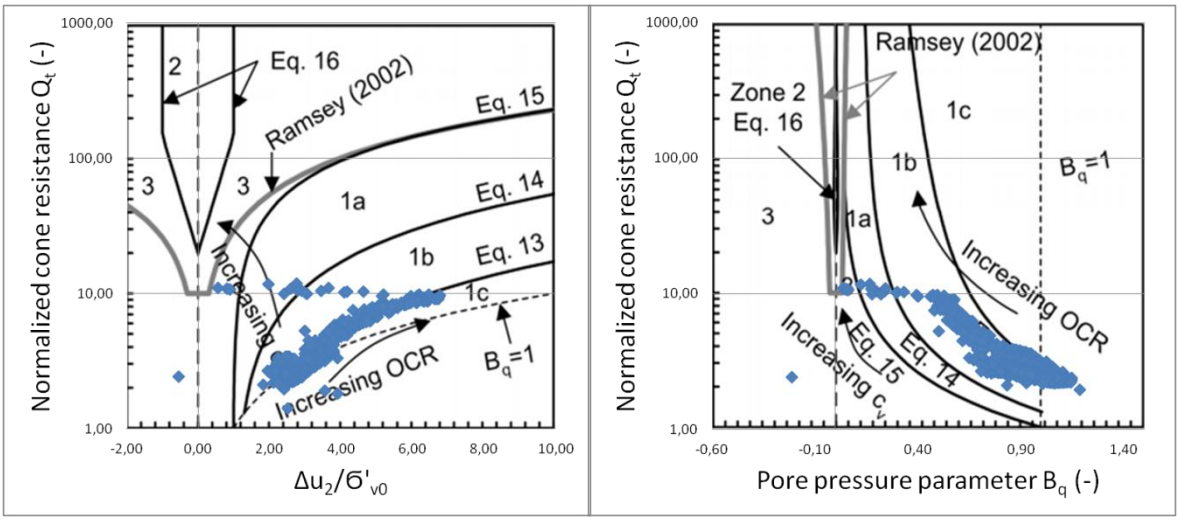


Figure 5.7.2 *Schneider classification charts (2008) for Klett RCPTu S2*

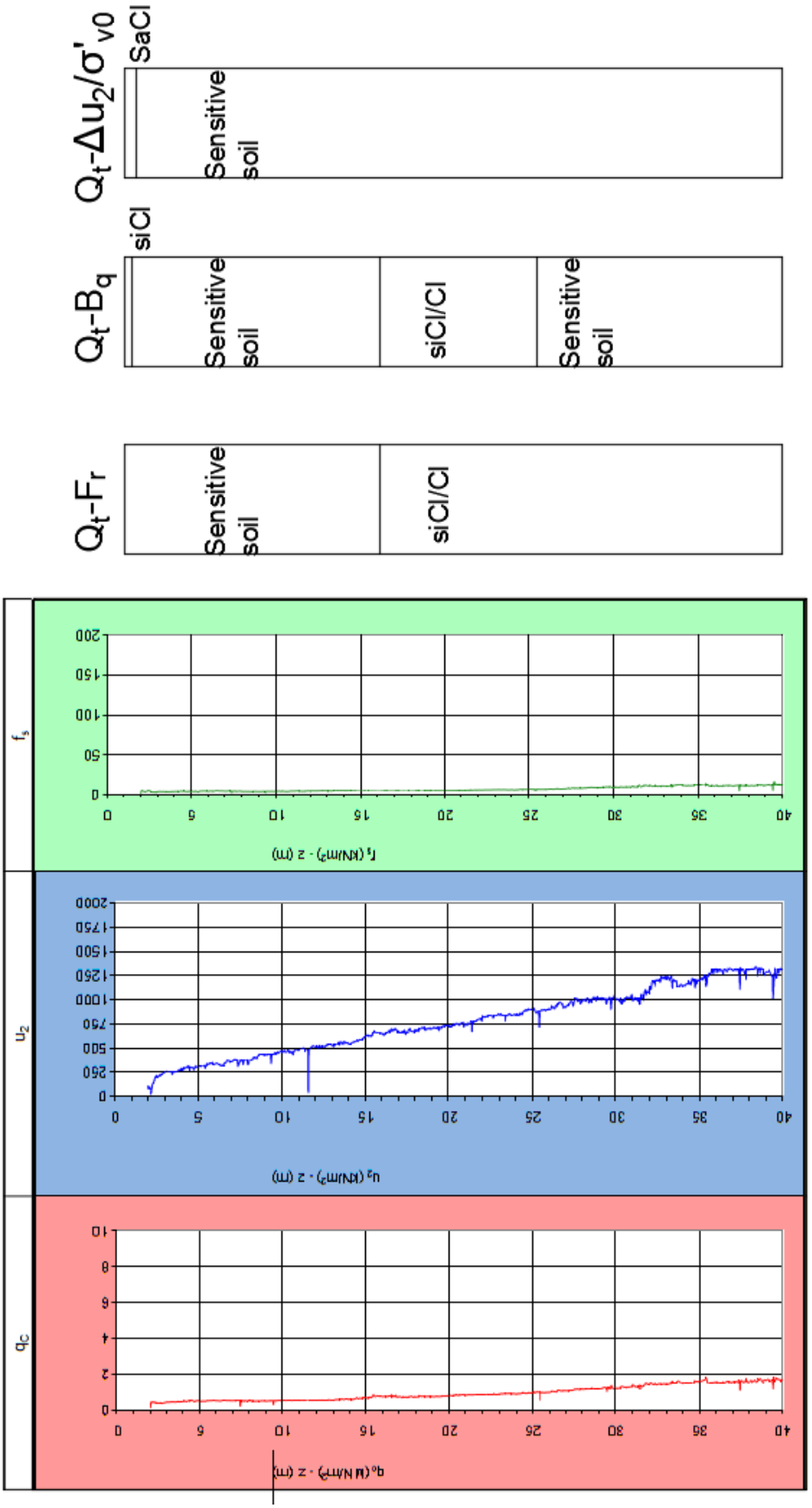


Figure 5.7.3 *Sounding parameters with interpreted profiles for Klett RCPTu S2*

5.7.2. Undrained shear resistance

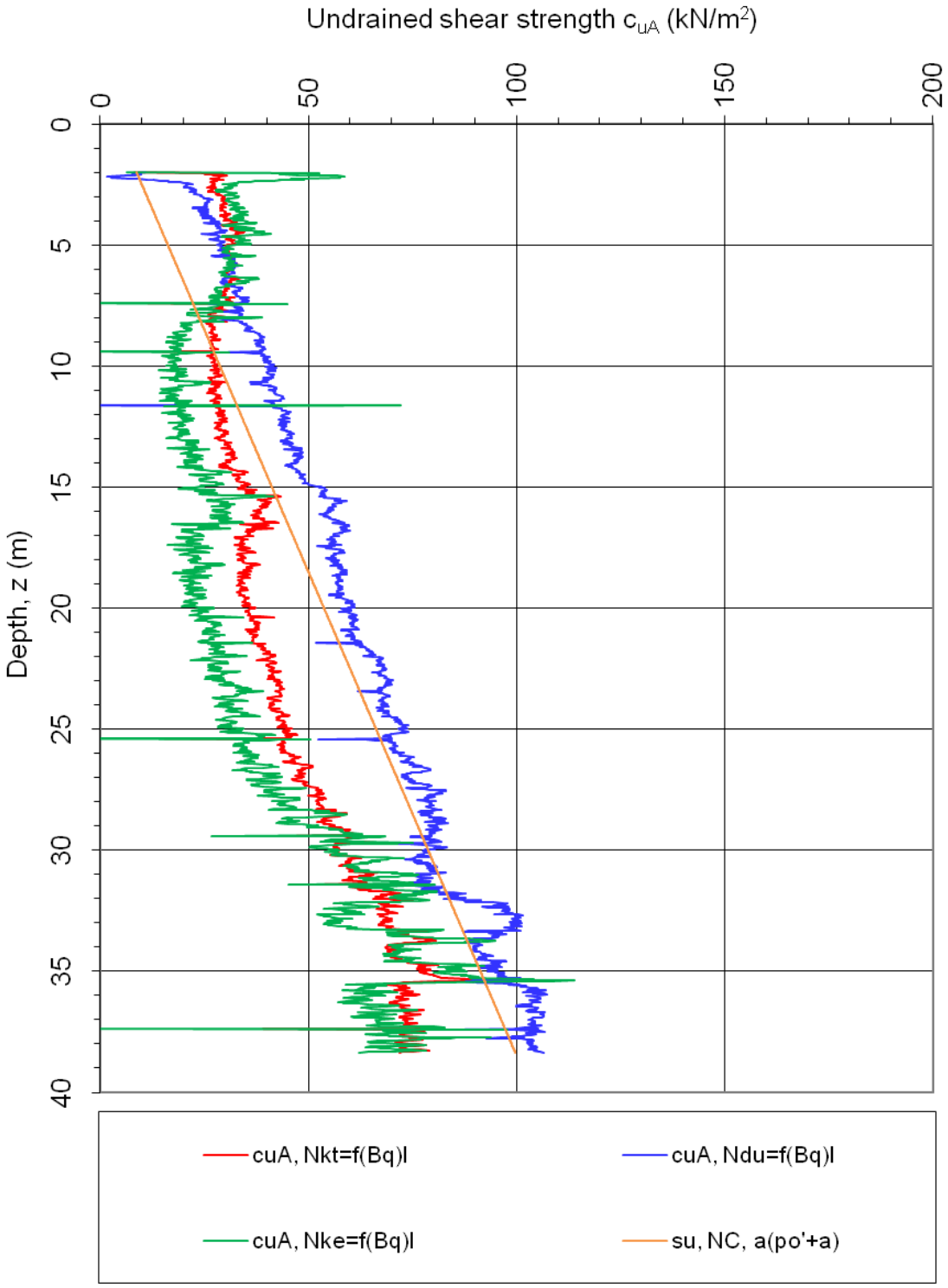


Figure 5.7.4 *Undrained shear resistance for interpreted results results for Klett RCPTu S2*

5.7.3. Sensitivity

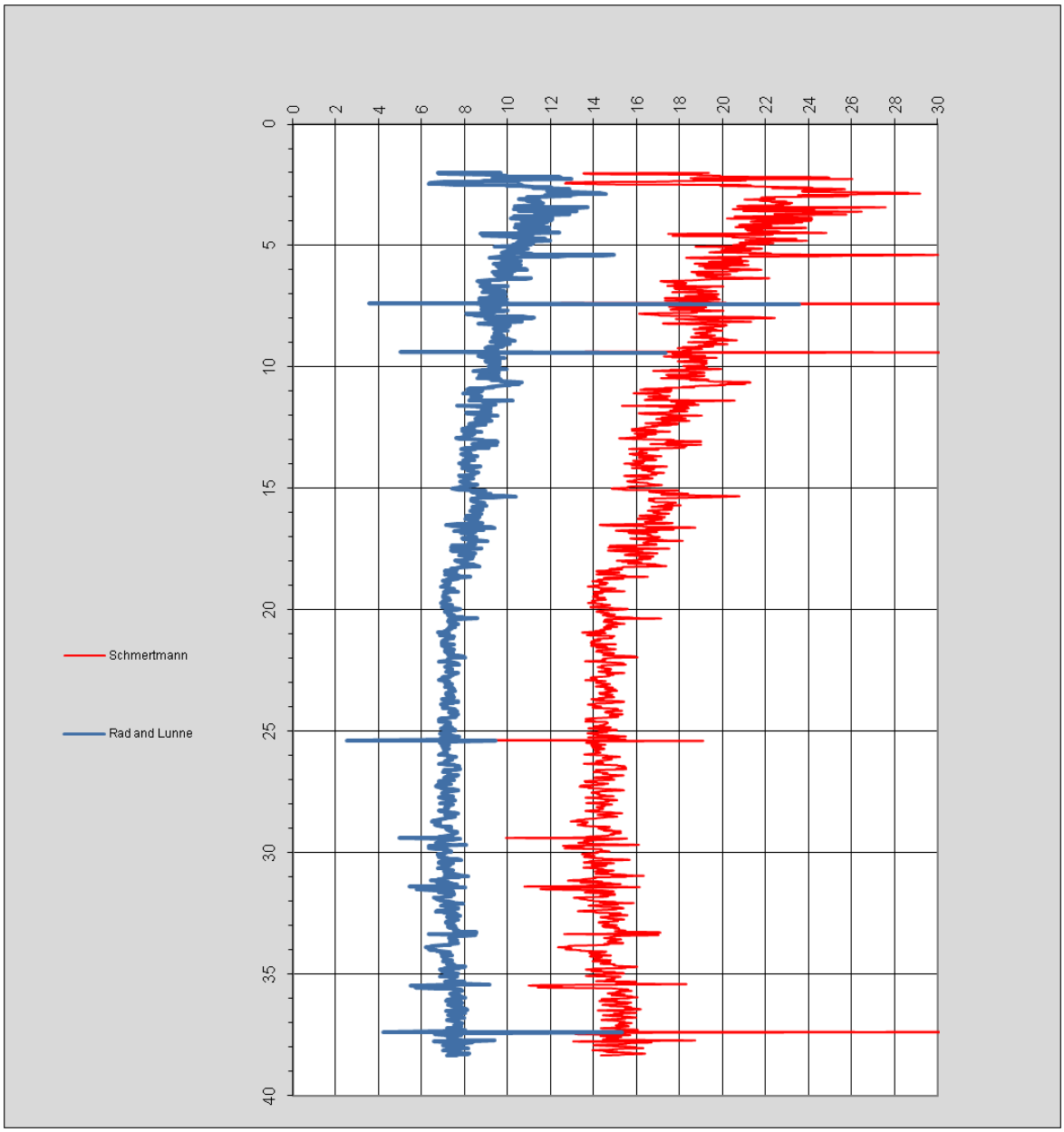


Figure 5.7.5 *Sensitivity with different parameters, results for Klett RCPTu S2*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.7.4. Resistivity

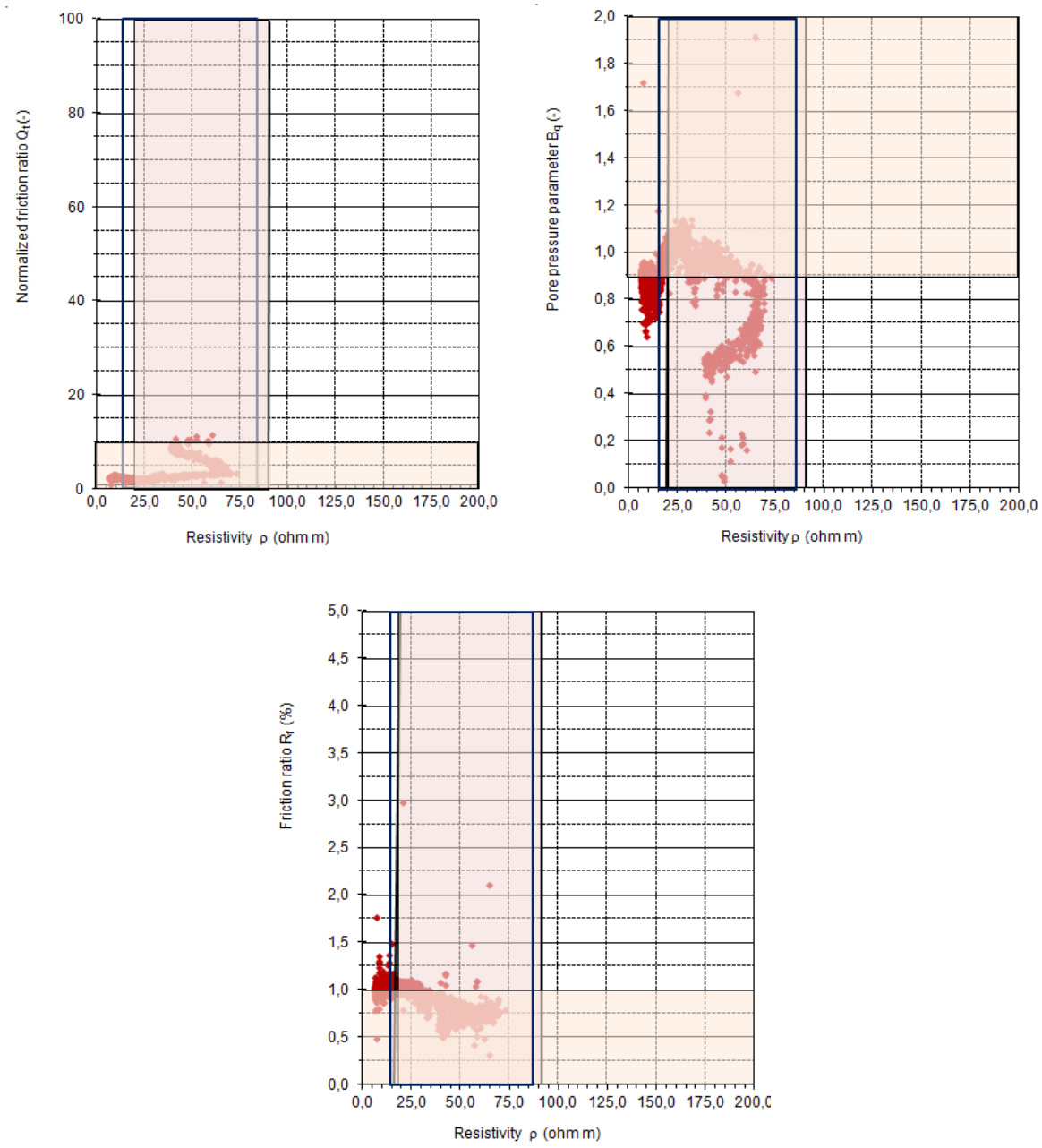


Figure 5.7.6 *Resistivity results in relation to N_m, R_f and B_q for Klett RCPTu S2*

Comment:
Marked areas are related to soil classification tables, which are in Chapter 6 of this project. Highlights represents values for quick clay.

5.8. Klett - CPTU 1502

5.8.1. Classification charts

After calculations types of soils and profiles were acquired:

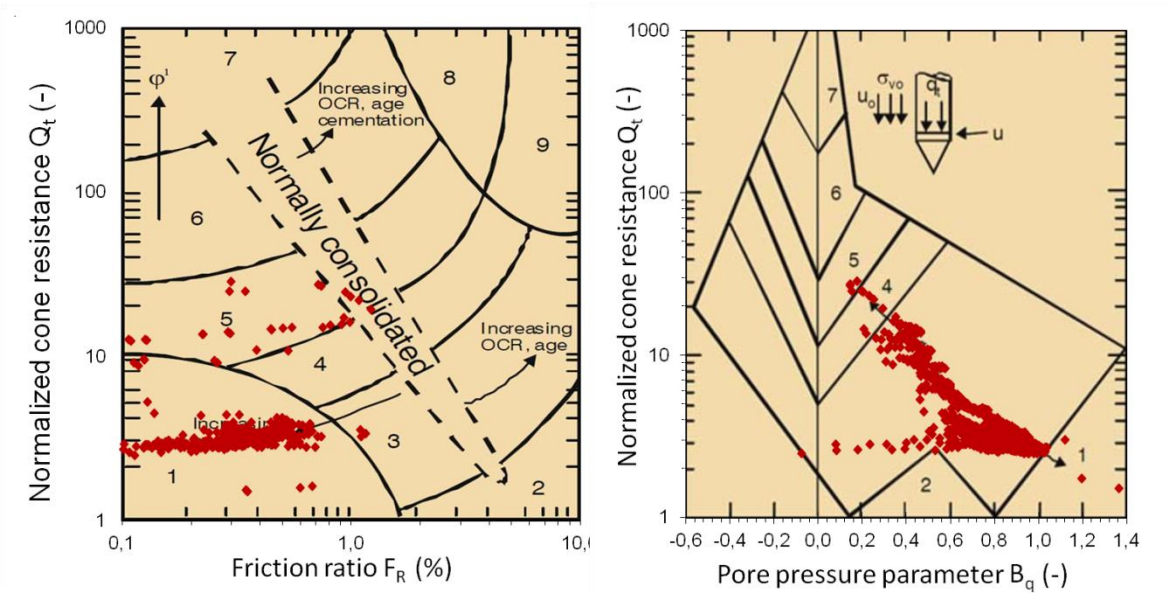


Figure 5.8.1 *Robertson classification chart (1990) for Klett CPTu 1502*

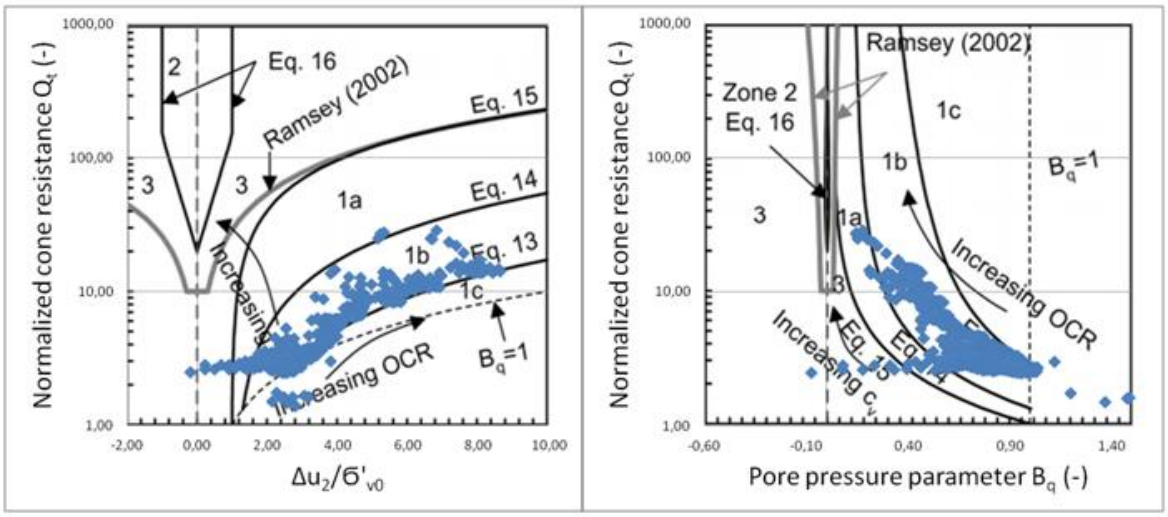


Figure 5.8.2 *Schneider classification charts (2008) for Klett CPTu 1502*

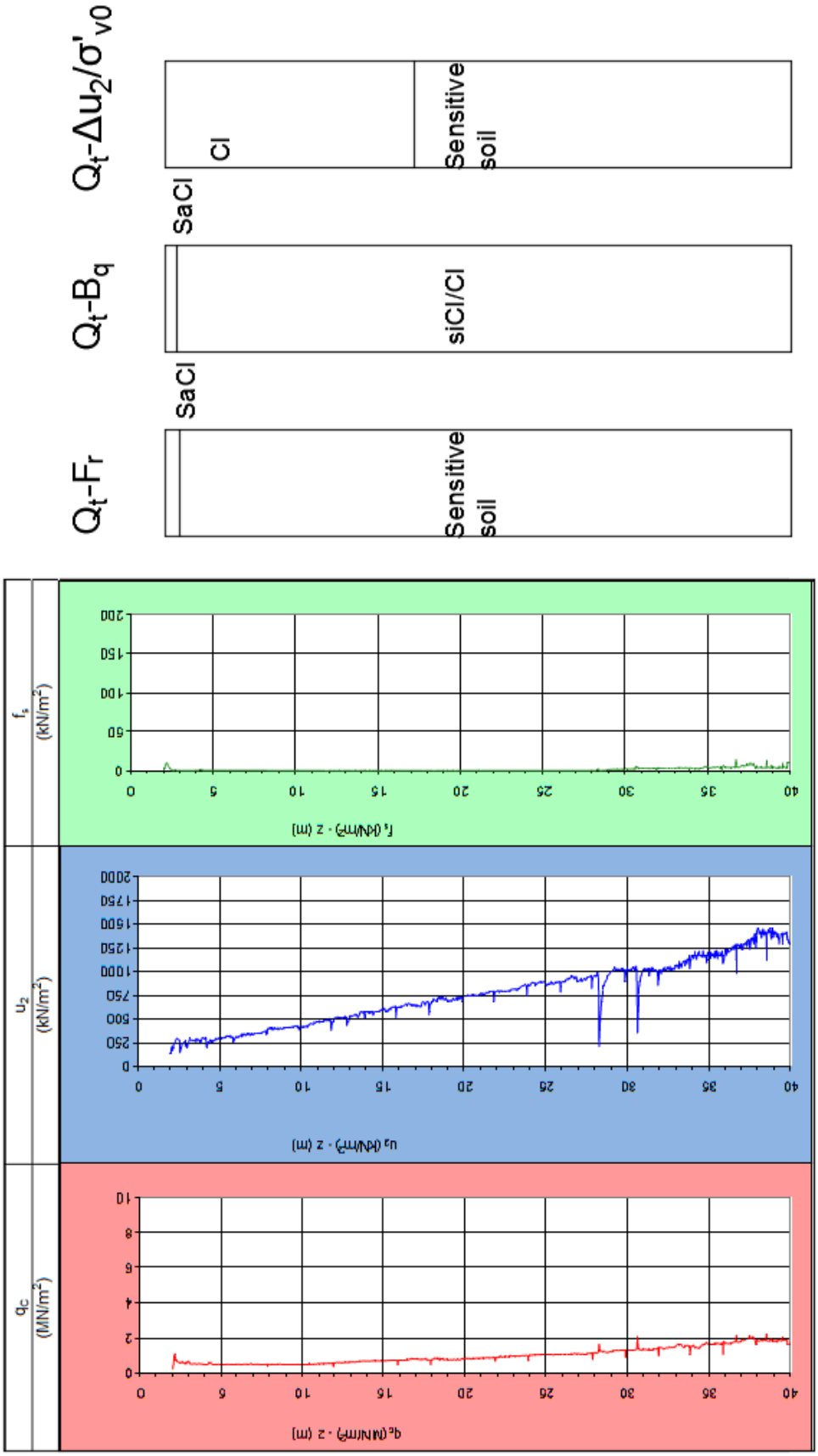


Figure 5.8.3 *Sounding parameters with interpreted profiles for Klett CPTu 1502*

5.8.2. Undrained shear resistance

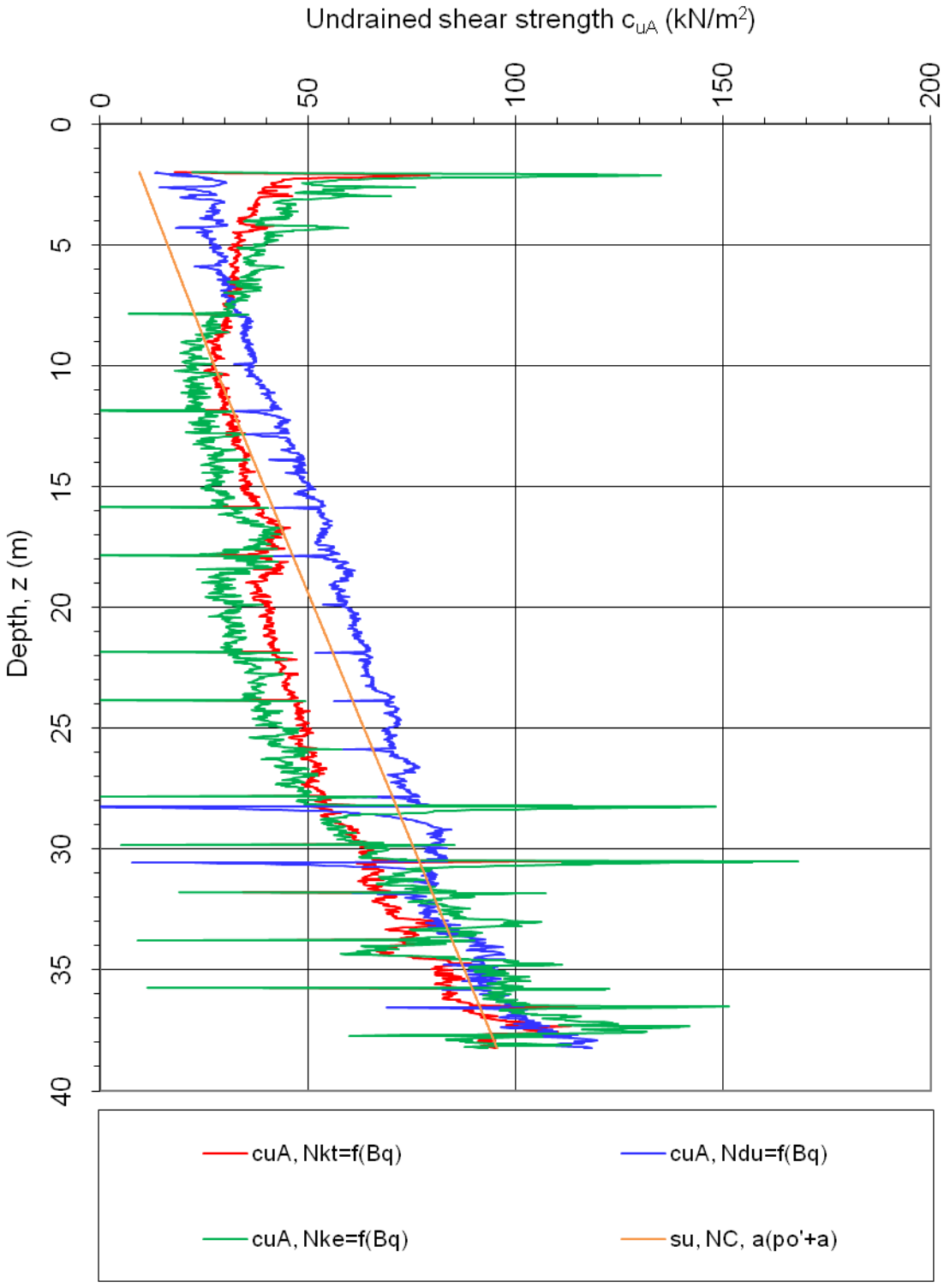


Figure 5.8.4 *Undrained shear resistance for interpreted, results for Klett CPTu 1502*

5.8.3. Sensitivity

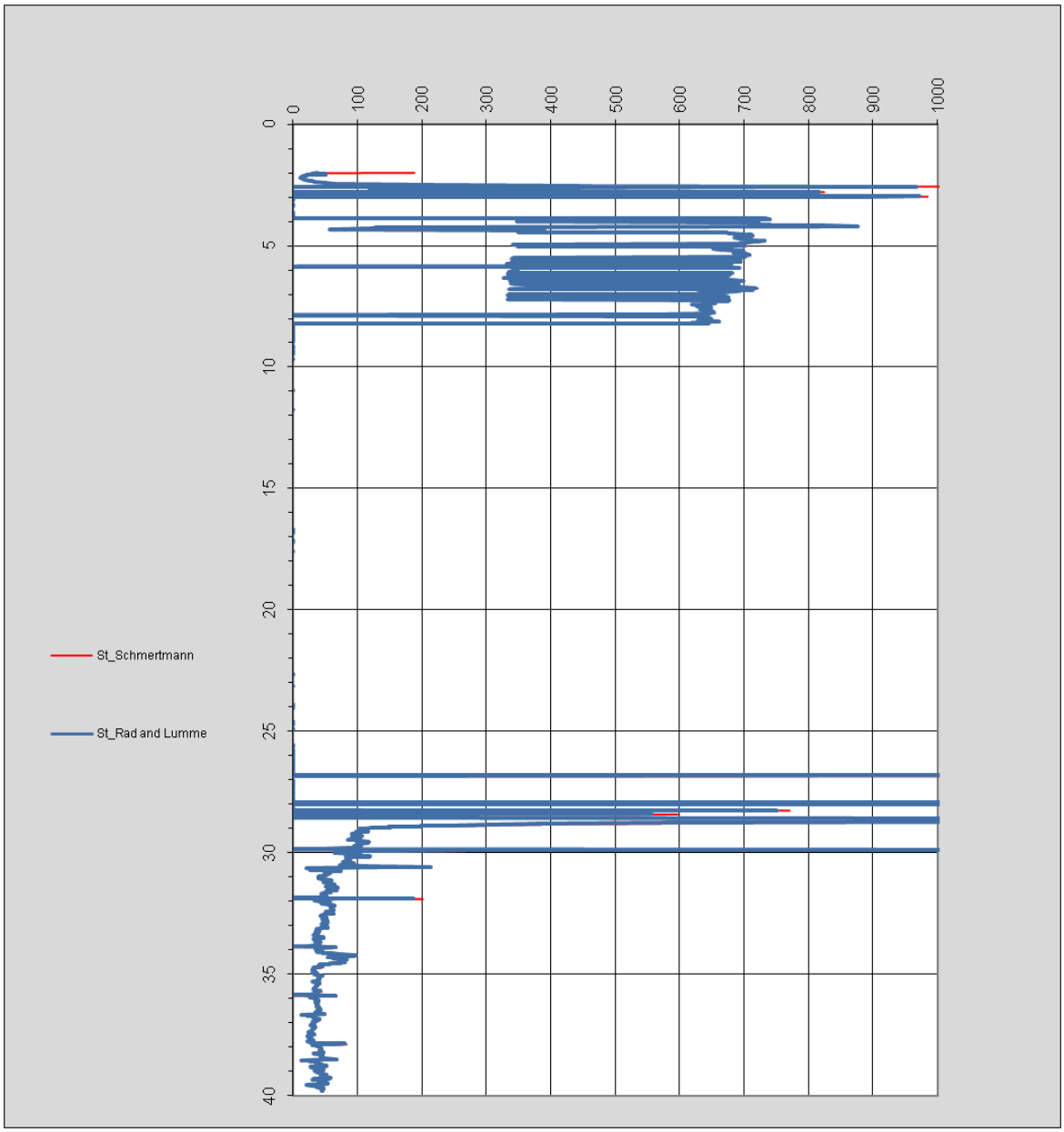


Figure 5.8.5 *Sensitivity with different parameters, results for Klett CPTu 1502*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account $N_s = 7{,}5$.

In this case, sleeve friction was so minimal, that sensitivity values could not be plotted (see Eq. 19 in chapter 3). Scale for horizontal axis was changed for more comprehensible view.

5.9. Klett - CPTU 1503

5.9.1. Classification charts

After calculations types of soils and profiles were acquired:

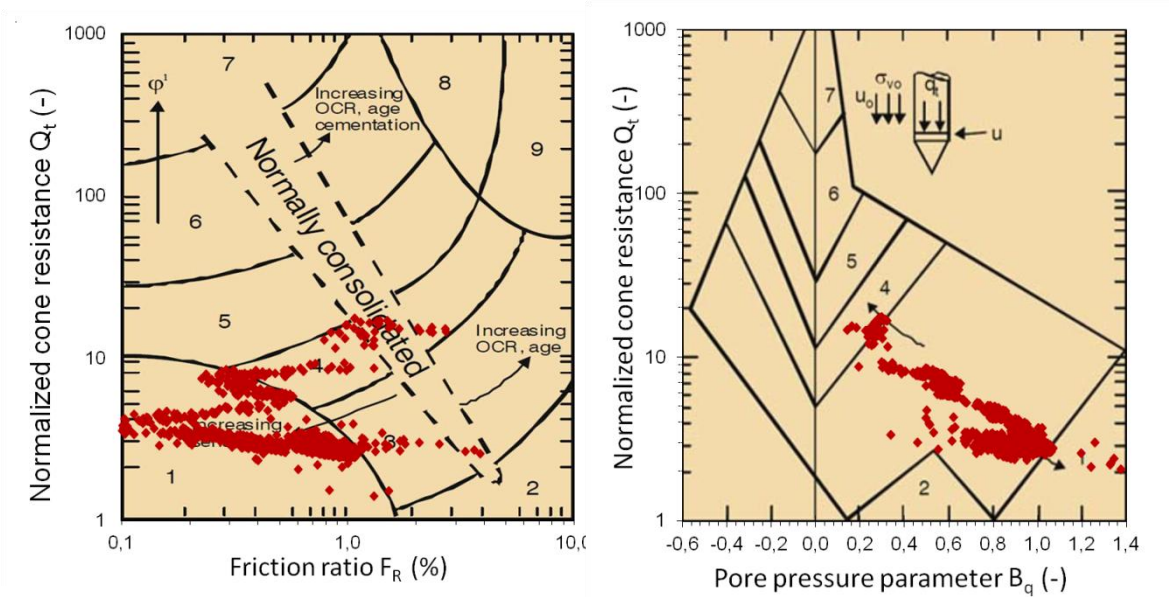


Figure 5.9.1 *Robertson classification chart (1990) for Klett CPTu 1503*

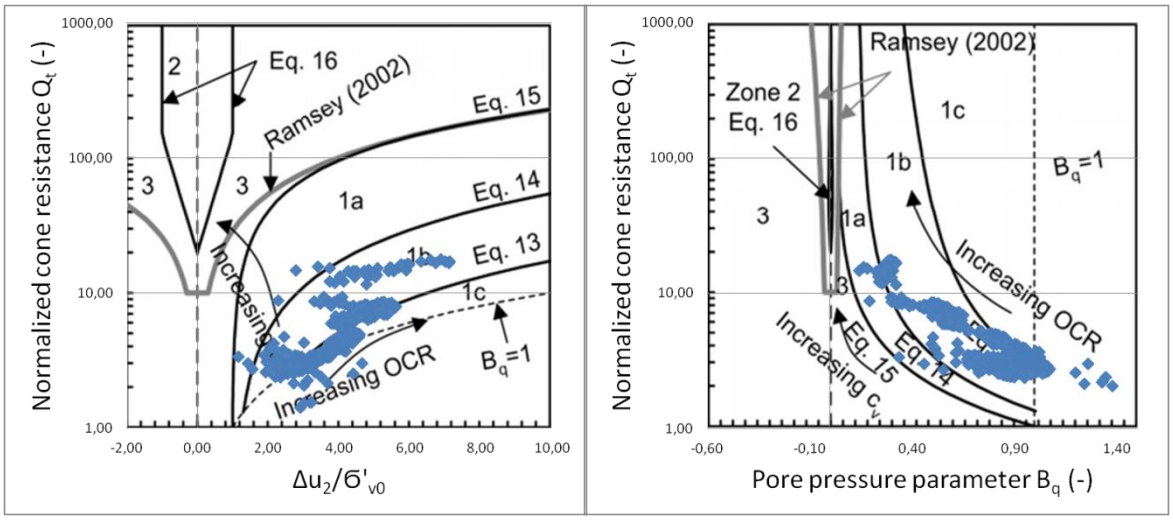


Figure 5.9.2 *Schneider classification charts (2008) for Klett CPTu 1503*

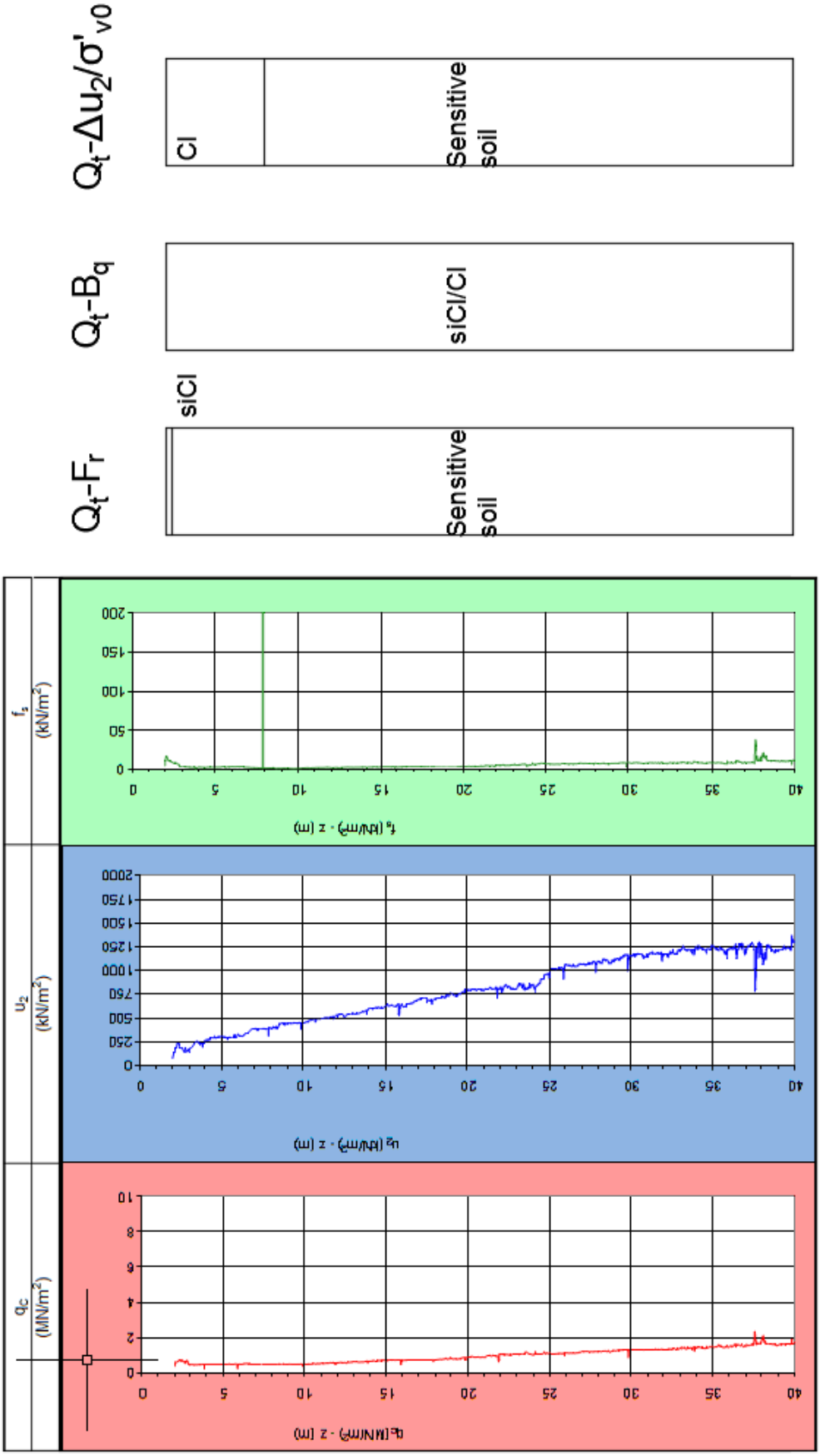


Figure 5.9.3 *Sounding parameters with interpreted profiles for Klett CPTu 1503*

5.9.2. Undrained shear resistance

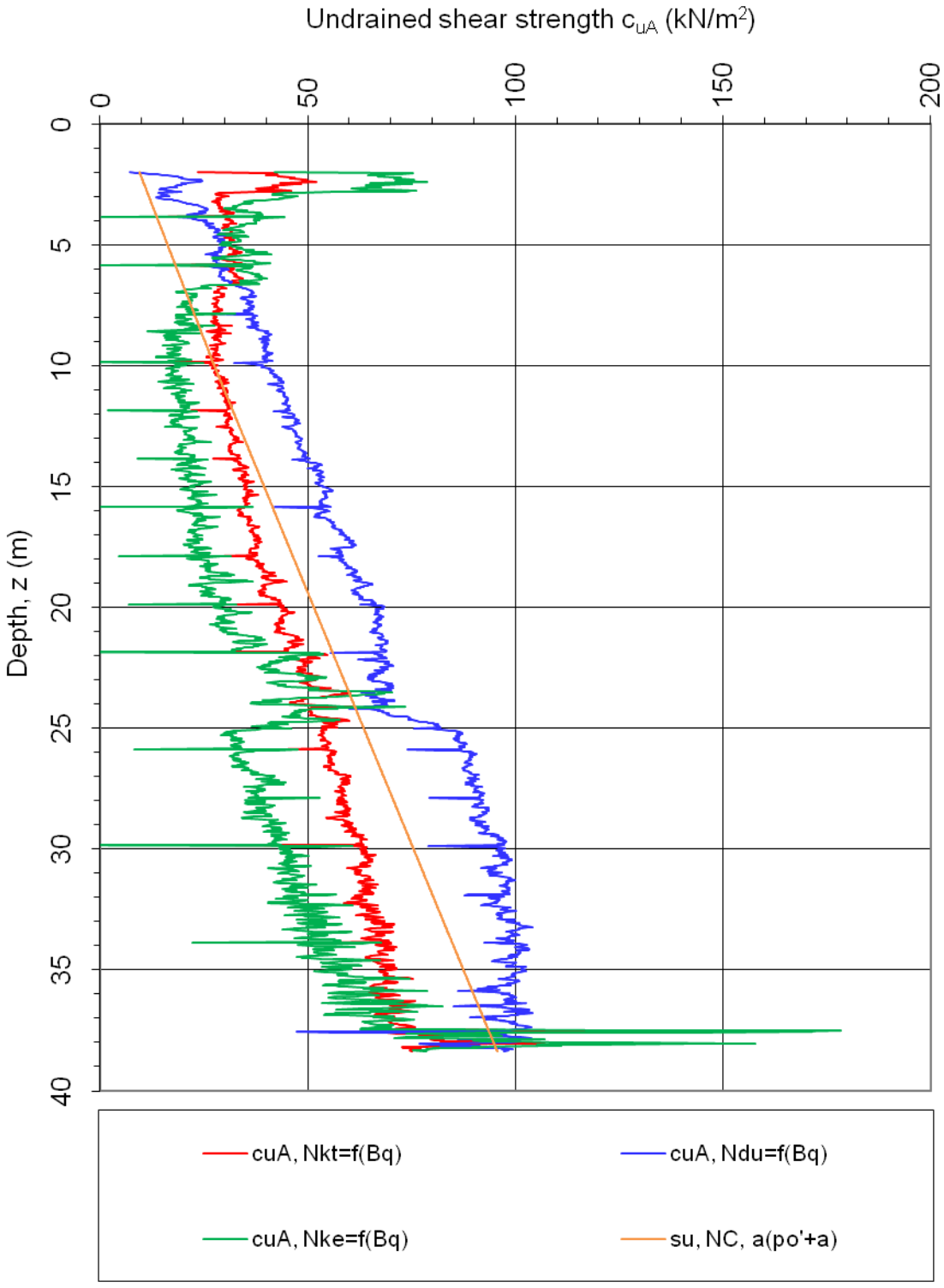


Figure 5.9.4 *Undrained shear resistance for interpreted results for Klett CPTu 1503*

5.3.3. Sensitivity

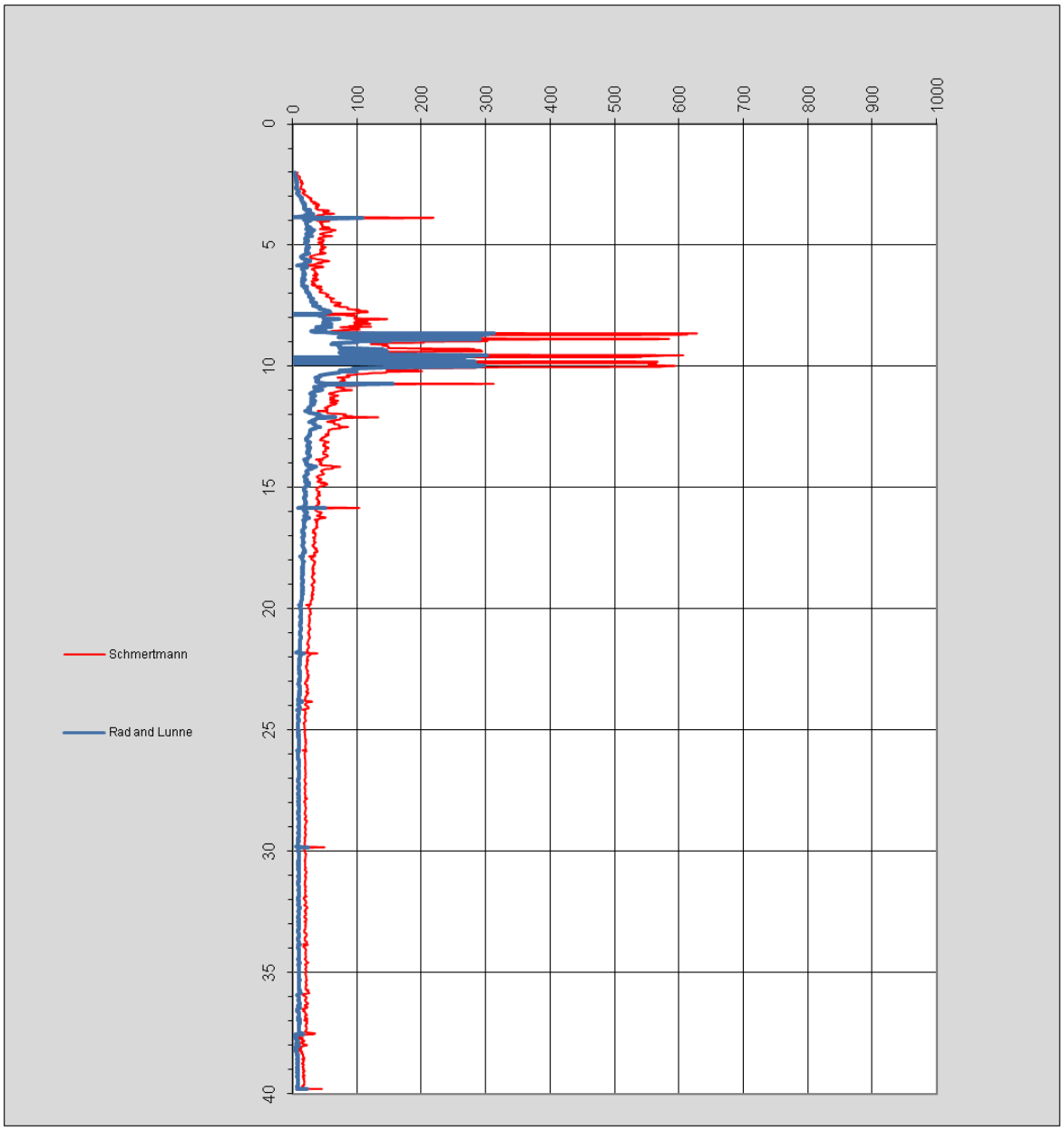


Figure 5.9.5 *Sensitivity with different parameters, results for Klett CPTu 1503*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

Scale for horizontal axis was changed for more comprehensible view due to disturbances of sleeve friction on a depth 10 m. Average value for sensitivity below 20 m is about 30 for Schmertmann and 15 for Rad and Lunne.

5.10. Klett - CPTU 1504

5.10.1. Classification charts

After calculations types of soils and profiles were acquired:

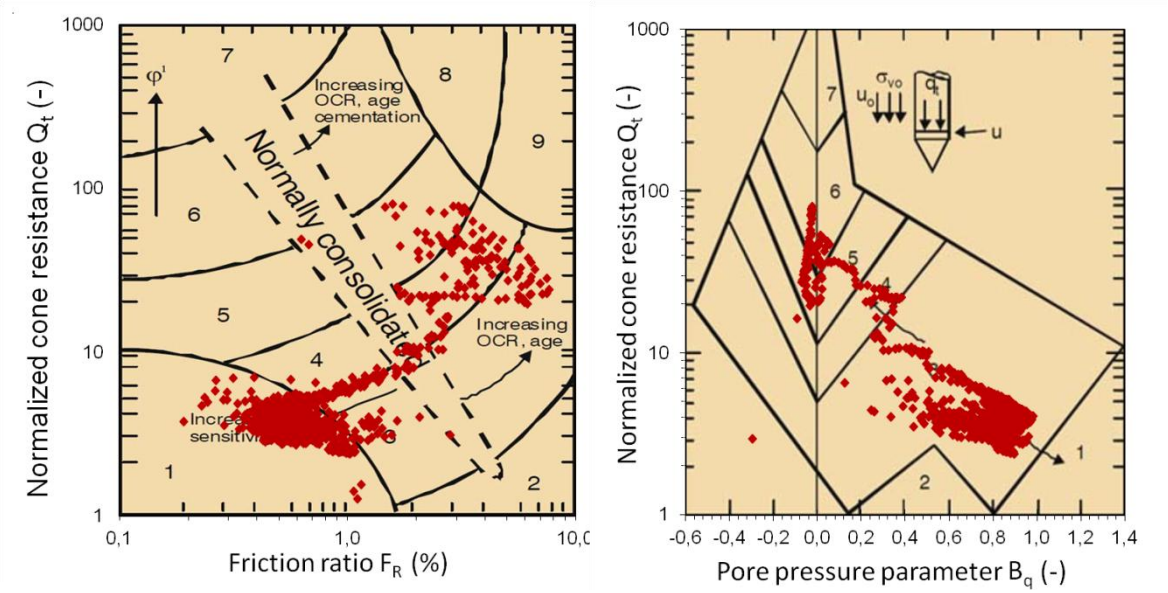


Figure 5.10.1 *Robertson classification chart (1990) for Klett CPTu 1504*

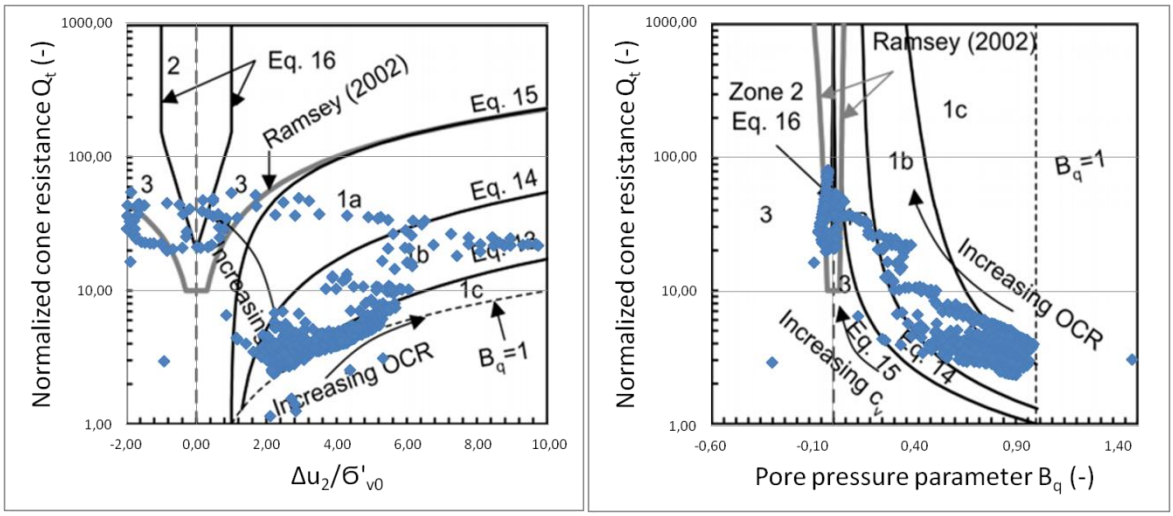


Figure 5.10.2 *Schneider classification charts (2008) for Klett CPTu 1504*

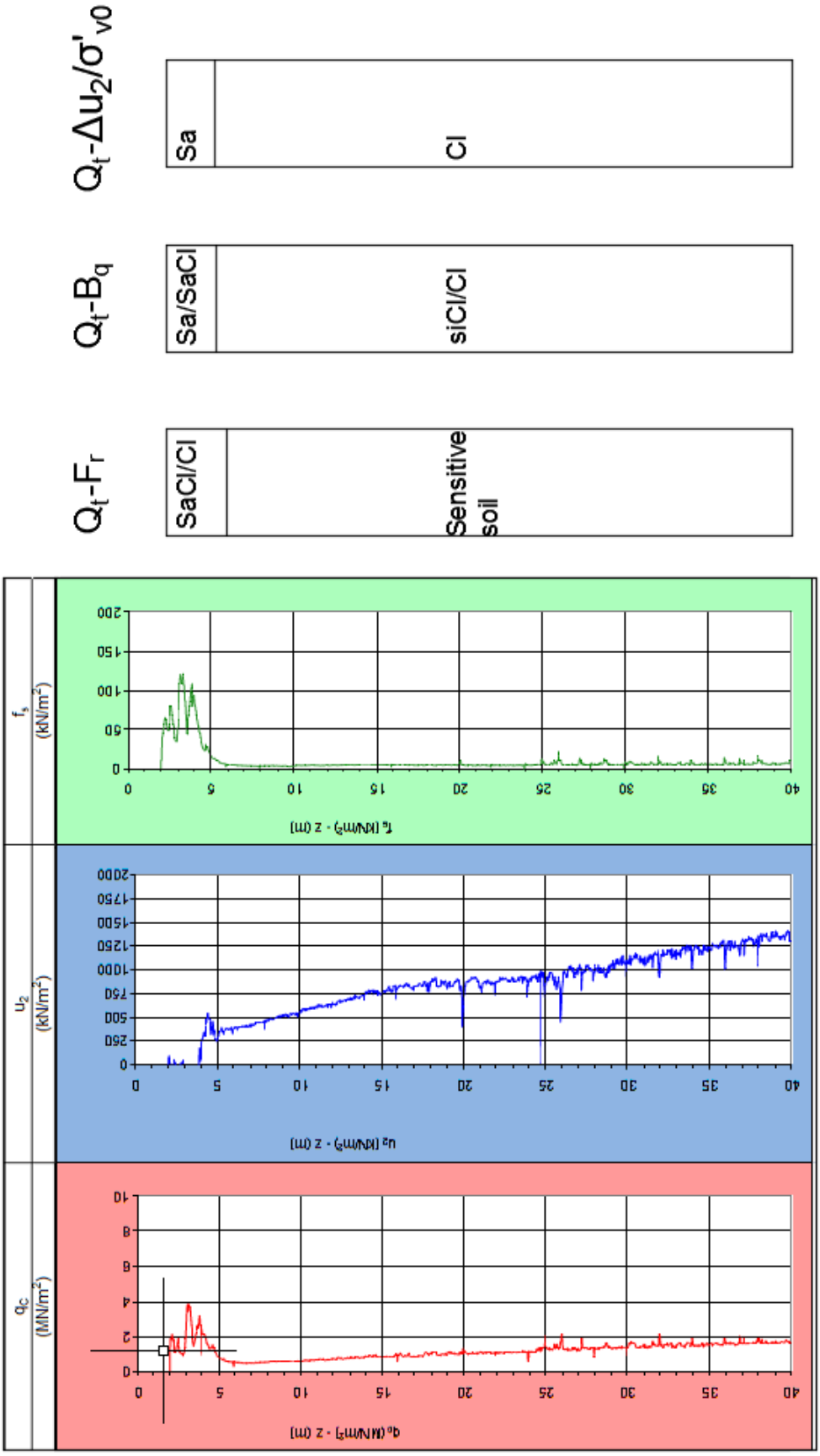


Figure 5.10.3 *Sounding parameters with interpreted profiles for Klett CPTu 1504*

5.10.2. Undrained shear resistance

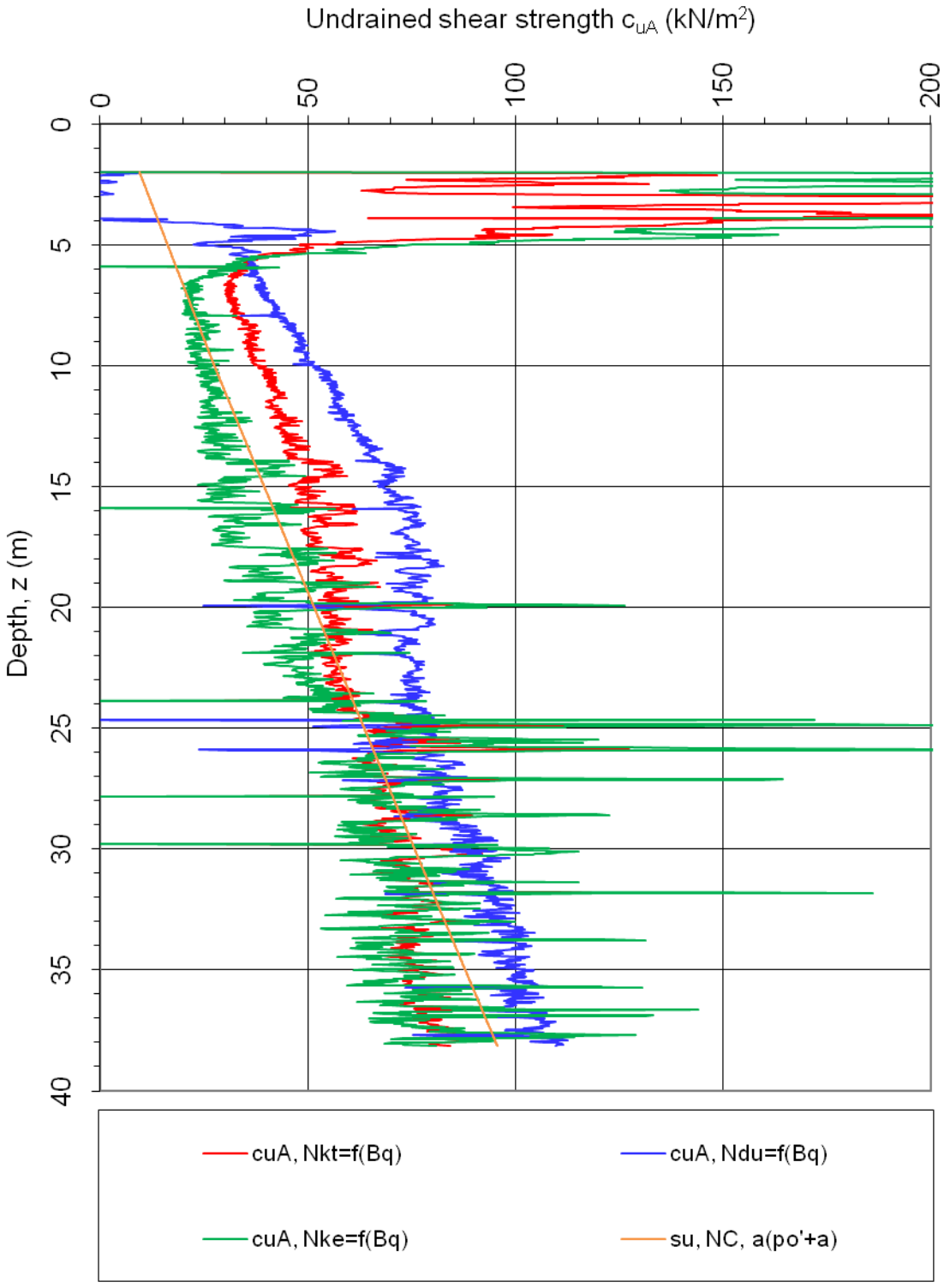


Figure 5.10.4 *Undrained shear resistance for interpreted results for Klett CPTu 1504*

5.10.3. Sensitivity

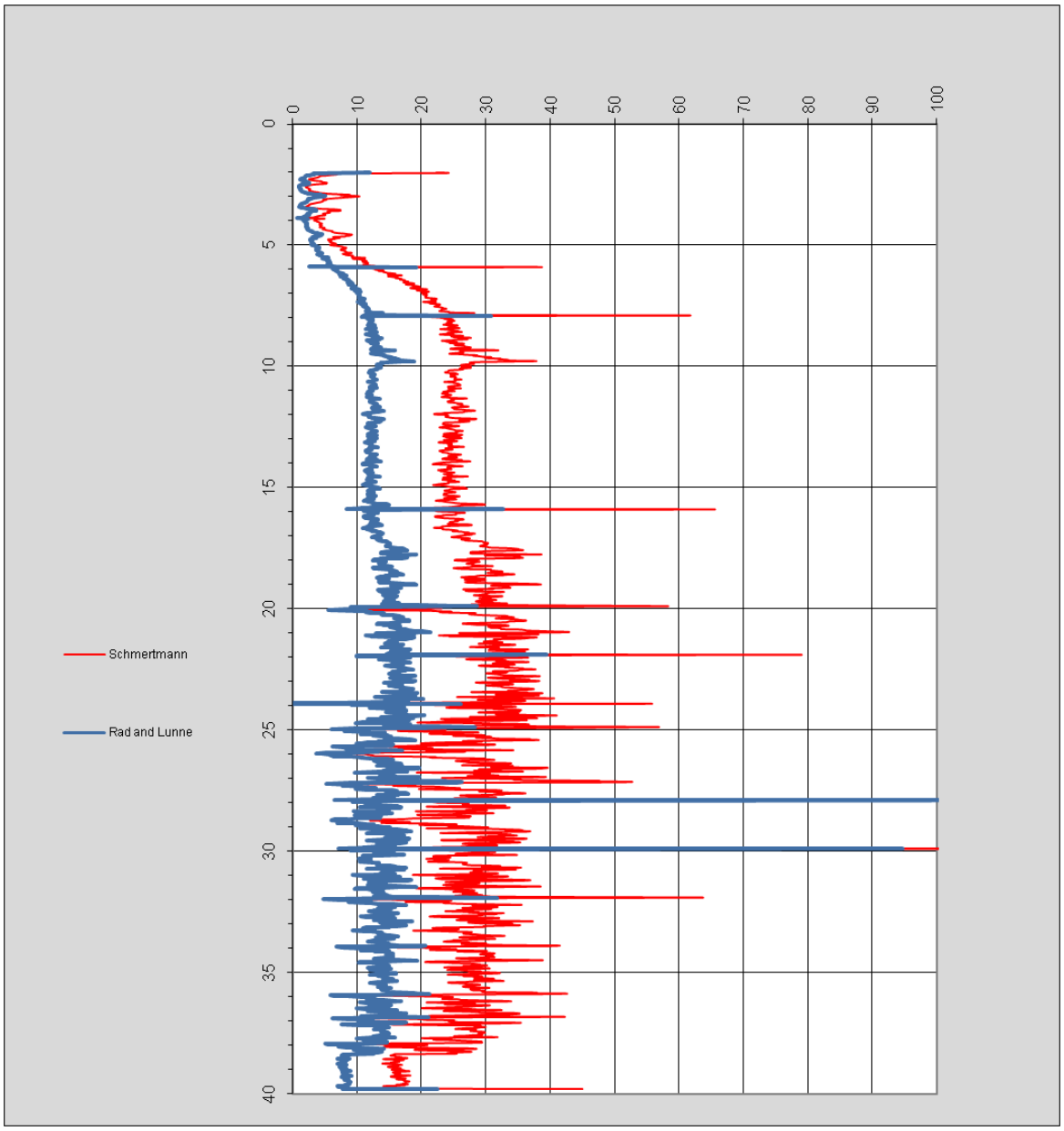


Figure 5.10.5 *Sensitivity with different parameters, results for Klett CPTu 1504*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

Scale for horizontal axis was changed for more comprehensible view - average values are comparable to soundings S1 and S2.

5.11. Klett - CPTU 1505

5.11.1. Classification charts

After calculations types of soils and profiles were acquired:

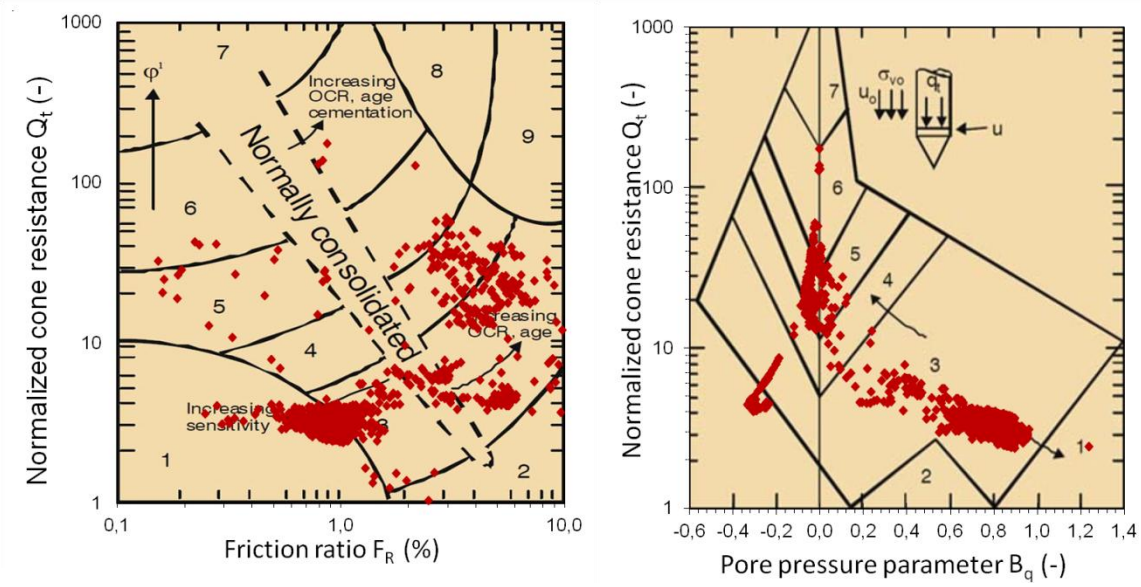


Figure 5.11.1 *Robertson classification chart (1990) for Klett CPTu 1505*

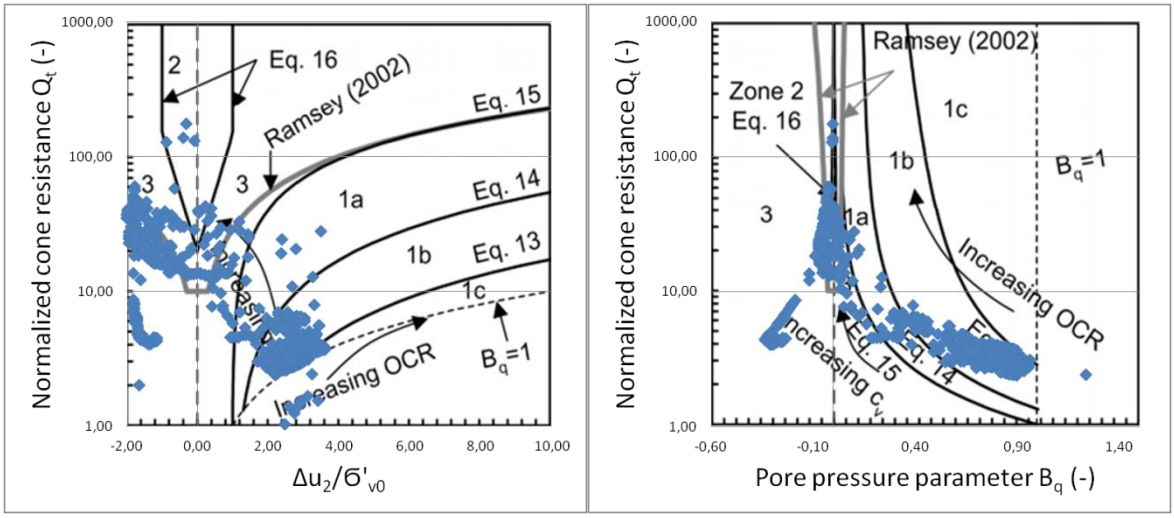


Figure 5.11.2 *Schneider classification charts (2008) for Klett CPTu 1505*

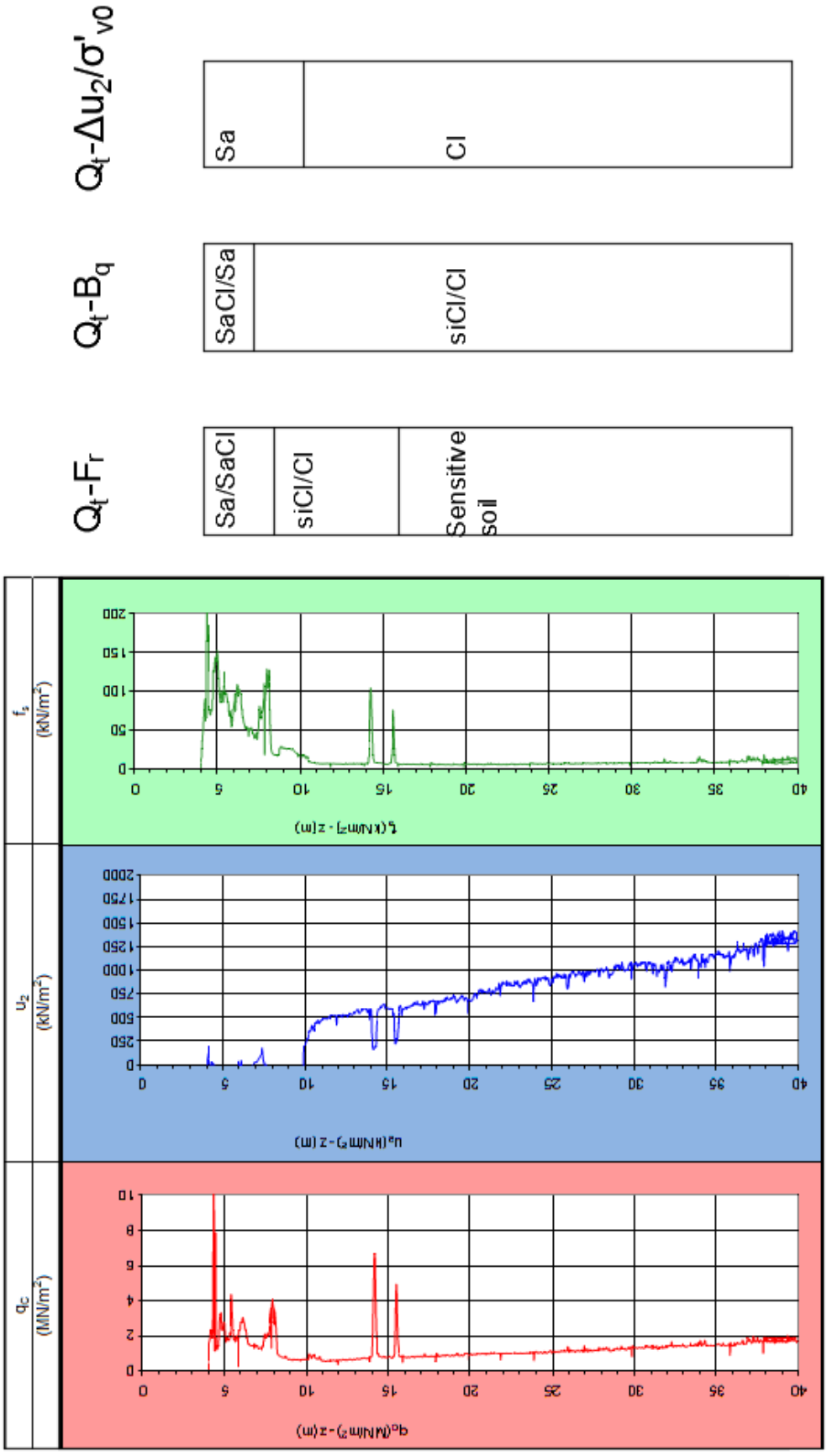


Figure 5.11.3 *Sounding parameters with interpreted profiles for Klett CPTu 1505*

5.10.2. Undrained shear resistance

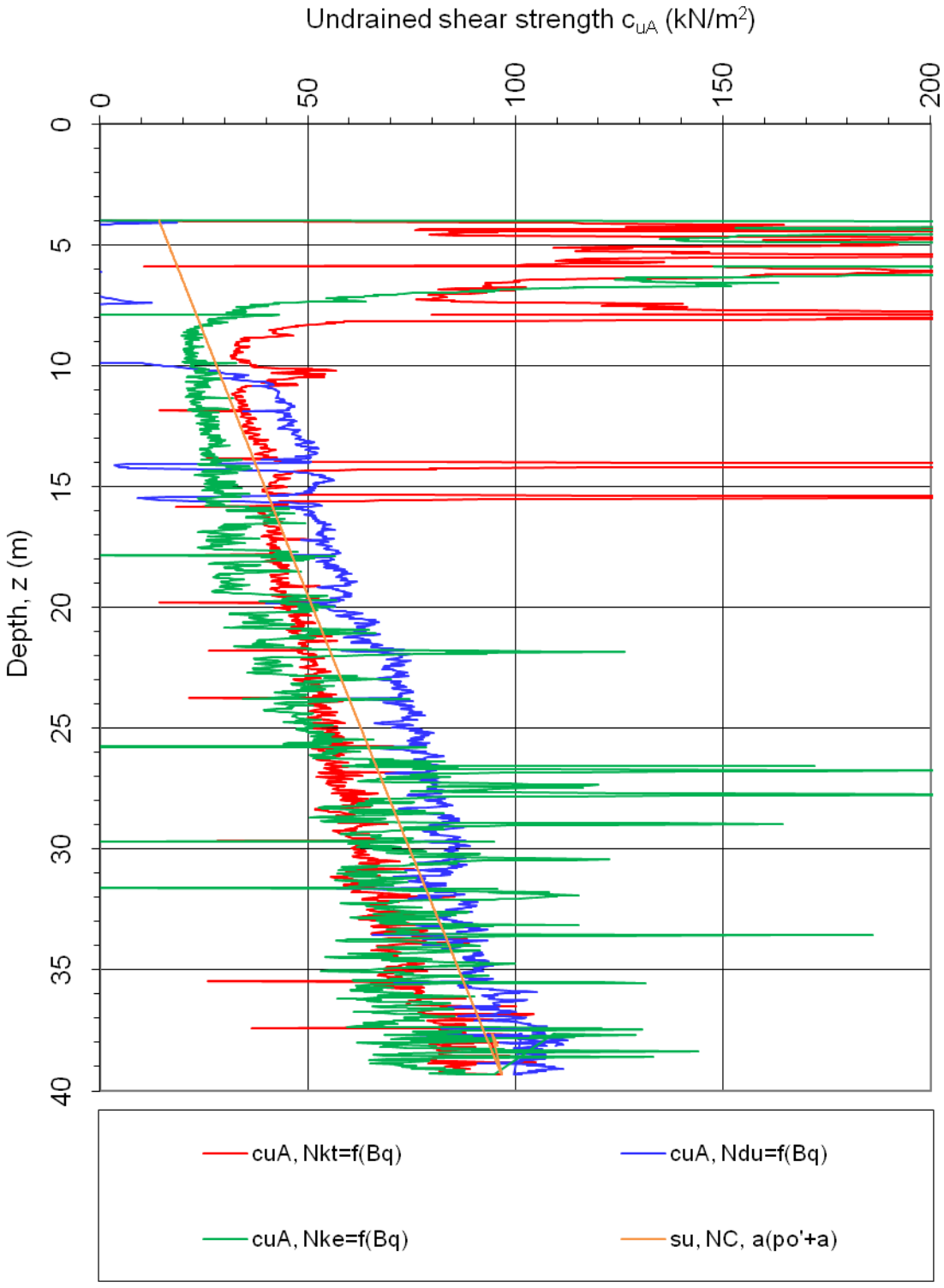


Figure 5.11.4 *Undrained shear resistance for interpreted results for Klett CPTu 1505*

5.10.3. Sensitivity

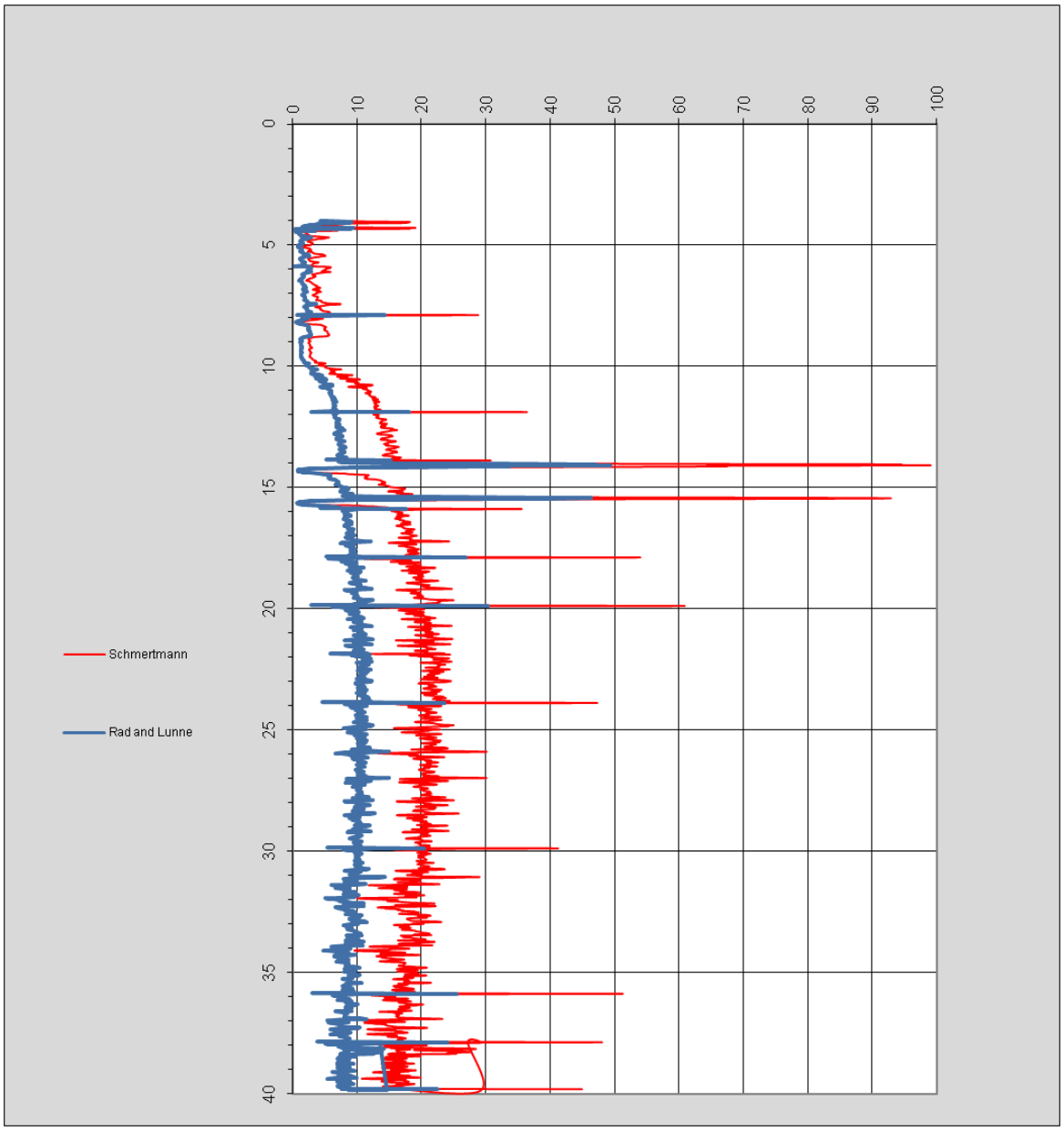


Figure 5.11.5 *Sensitivity with different parameters, results for Klett CPTu 1505*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

Scale for horizontal axis was changed for more comprehensible view - average values are comparable to soundings S1, S2 and CPTu 1504.

5.12. Klett - CPTU S1

5.12.1. Classification charts

After calculations types of soils and profiles were acquired:

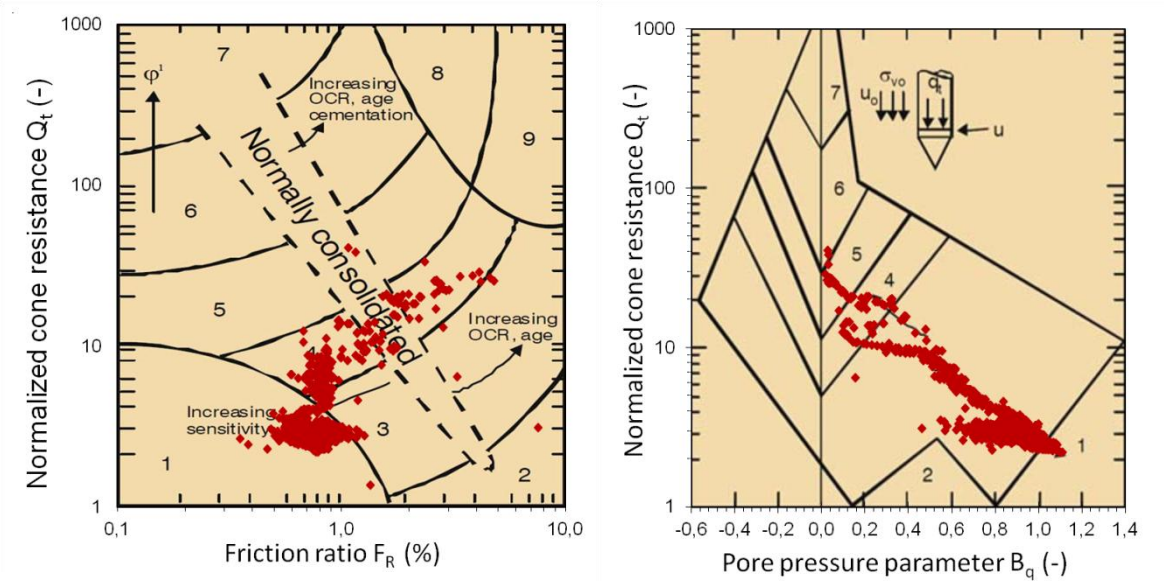


Figure 5.12.1 *Robertson classification chart (1990) for Klett CPTu S1*

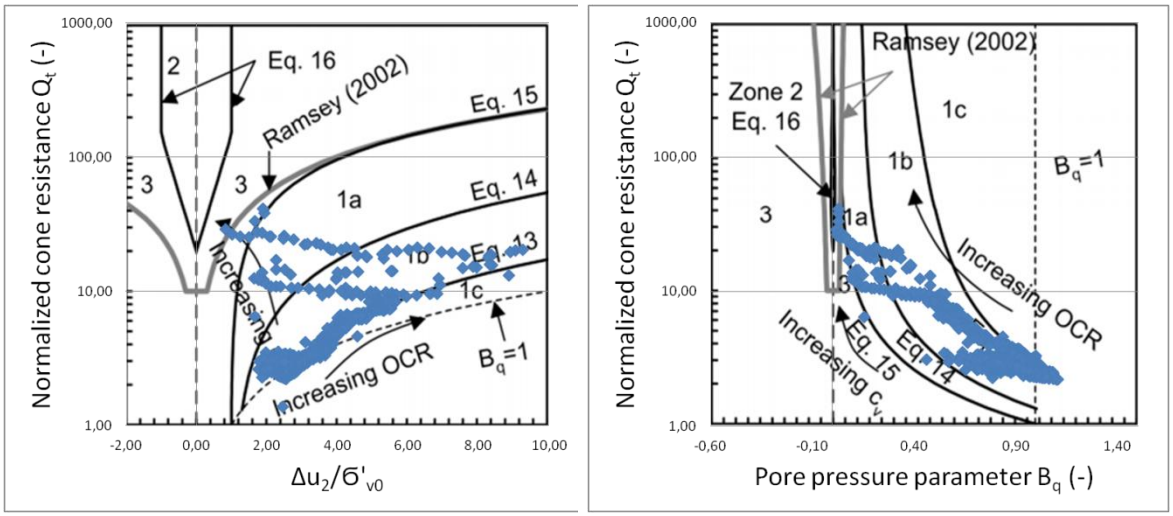


Figure 5.12.2 *Schneider classification charts (2008) for Klett CPTu S1*

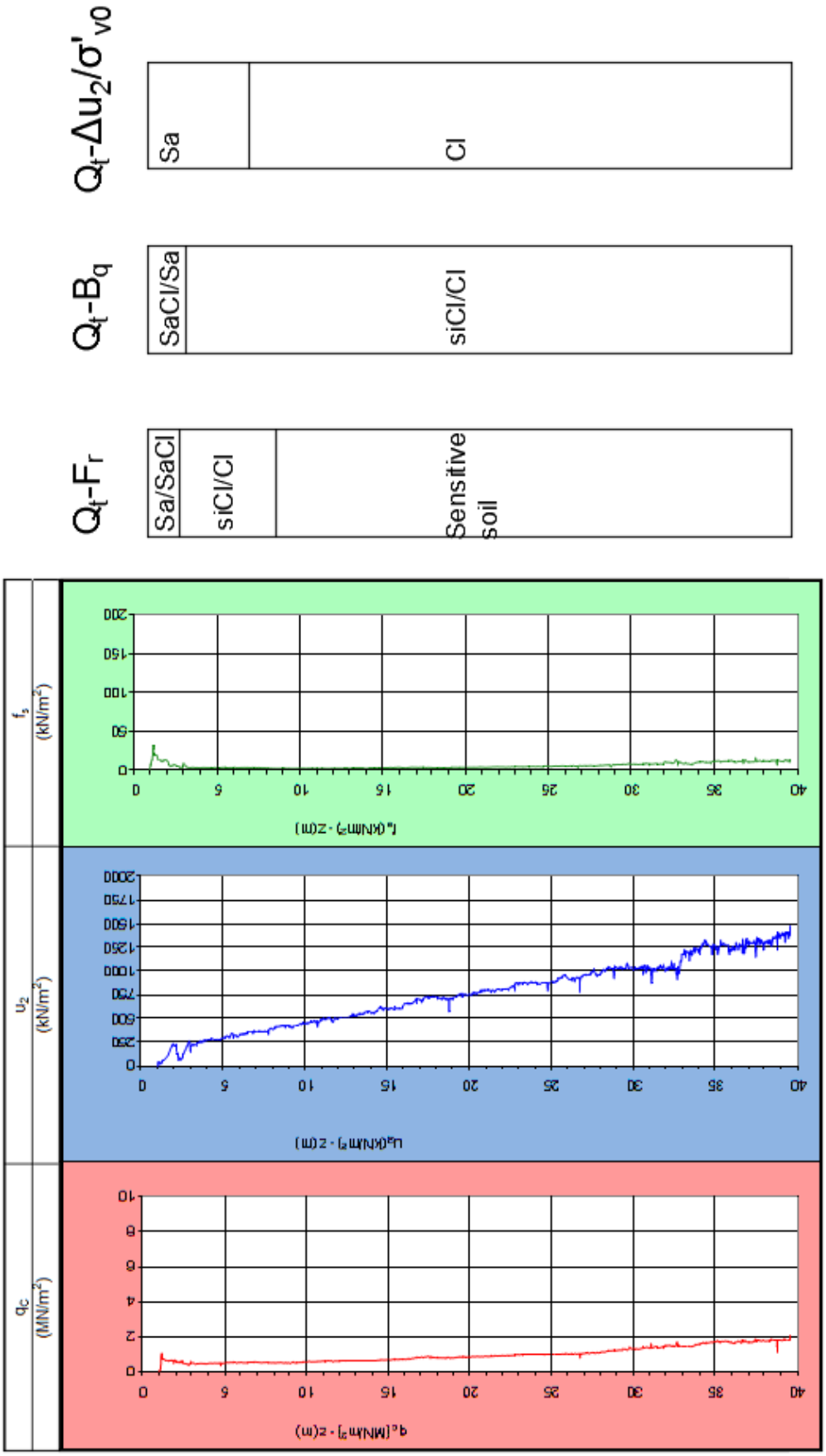


Figure 5.12.3 *Sounding parameters with interpreted profiles for Klett CPTu S1*

5.12.2. Undrained shear resistance

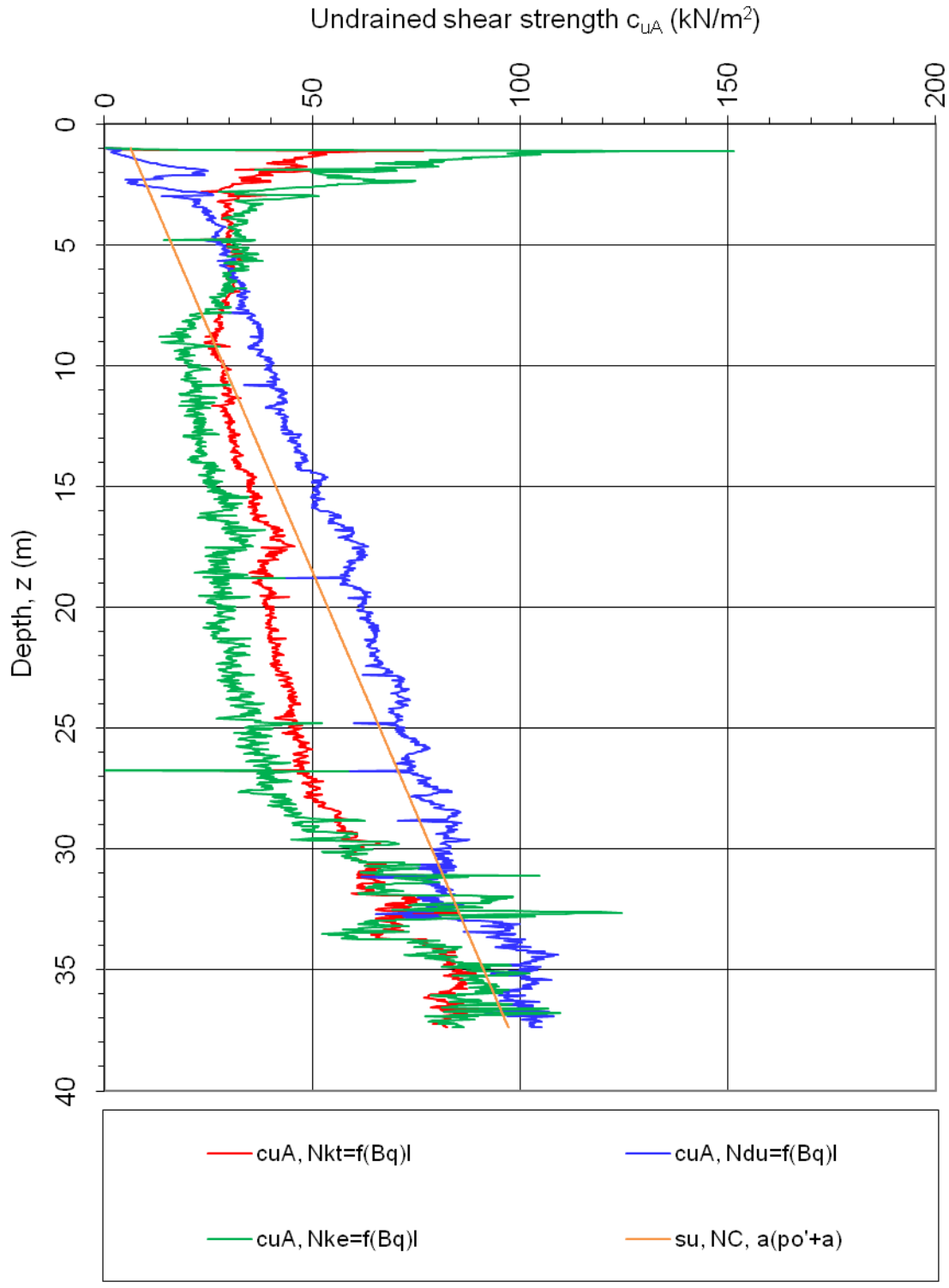


Figure 5.12.4 *Undrained shear resistance for interpreted results for Klett CPTu S1*

5.12.3. Sensitivity

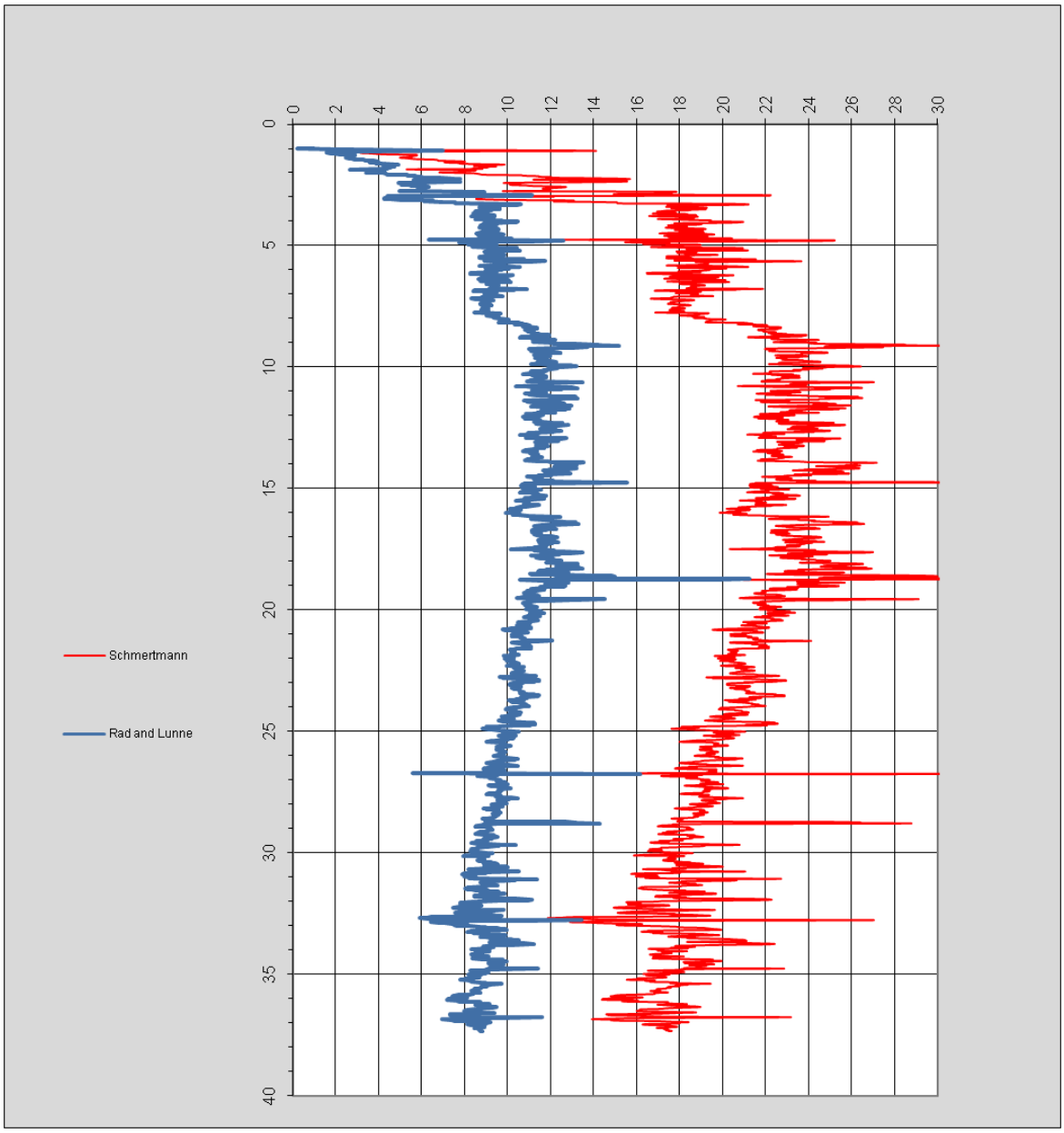


Figure 5.12.5 *Sensitivity with different parameters, results for Klett CPTu S1*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.

5.13. Tiller - RCPTu B1

5.13.1. Classification charts

After calculations types of soils and profiles were acquired:

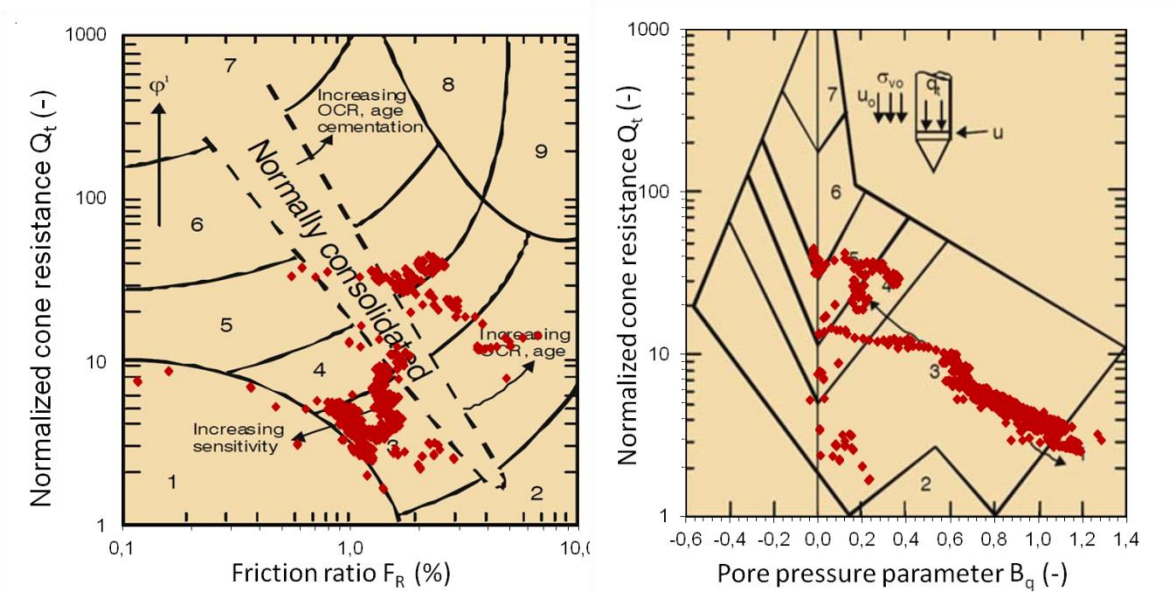


Figure 5.13.1 *Robertson classification chart (1990) for Tiller RCPTu B1*

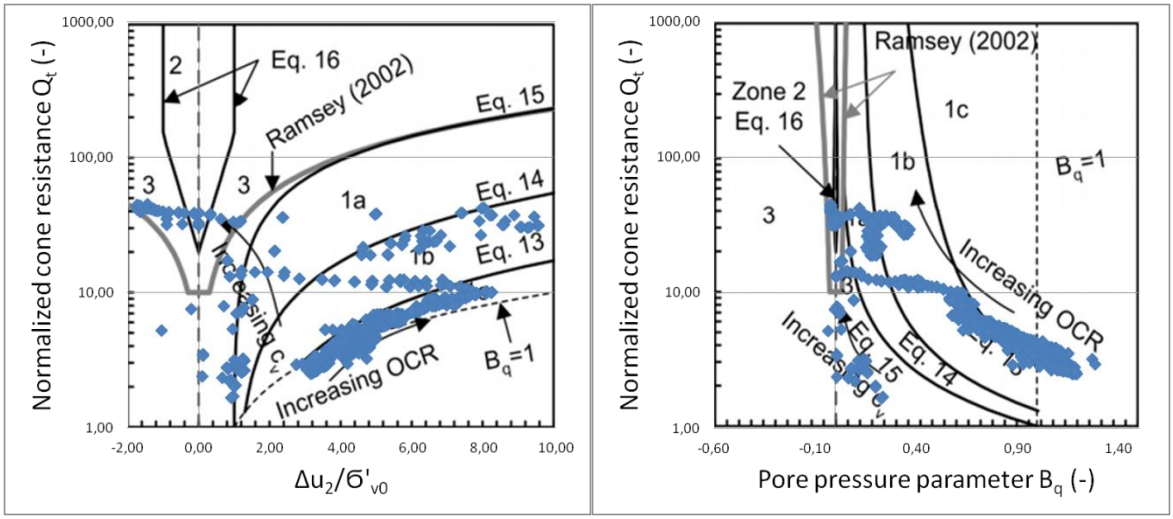


Figure 5.13.2 *Schneider classification charts (2008) for Tiller RCPTu B1*

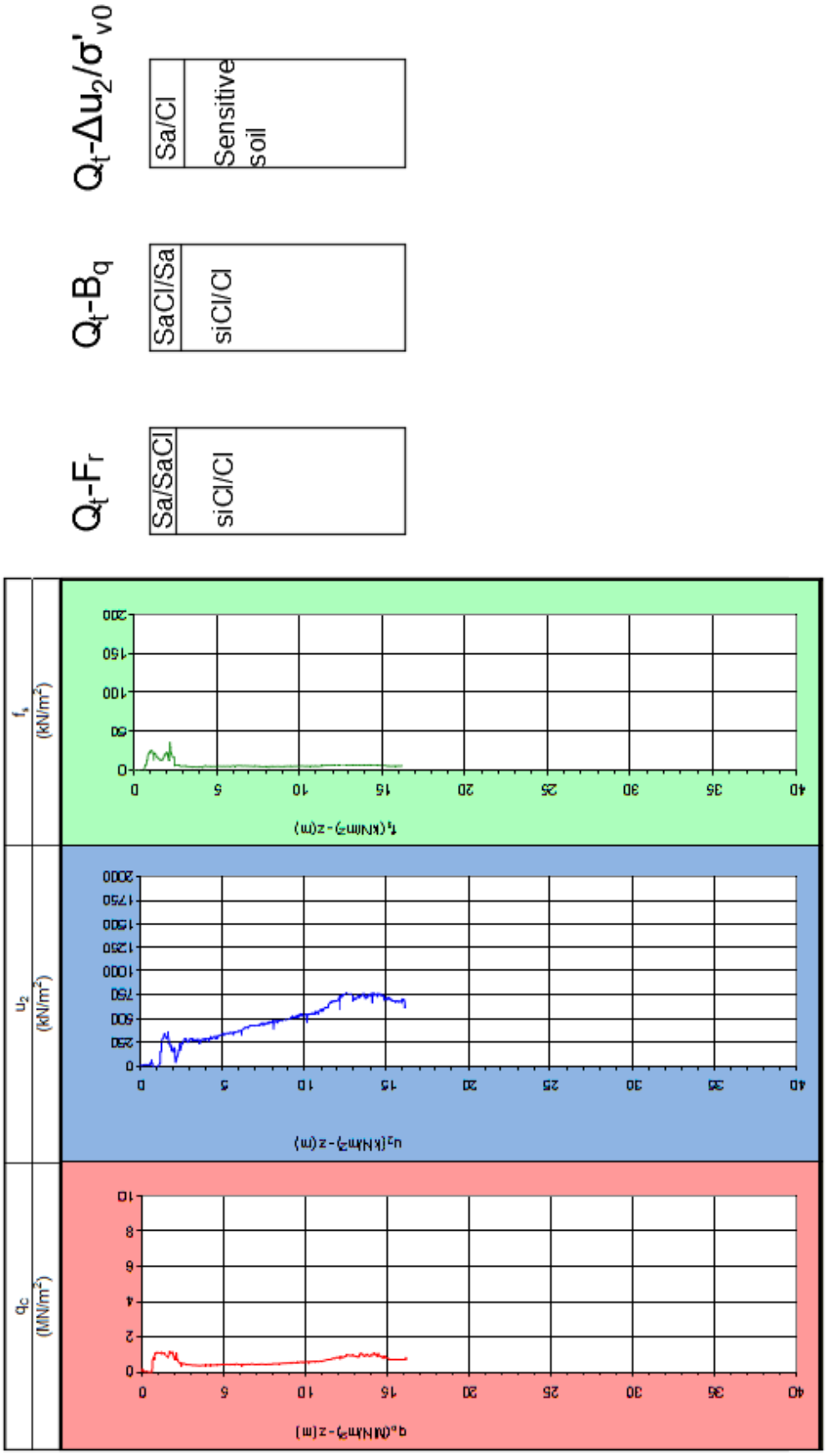


Figure 5.13.3 *Sounding parameters with interpreted profiles for Tiller RCPTu B1*

5.13.2. Undrained shear resistance

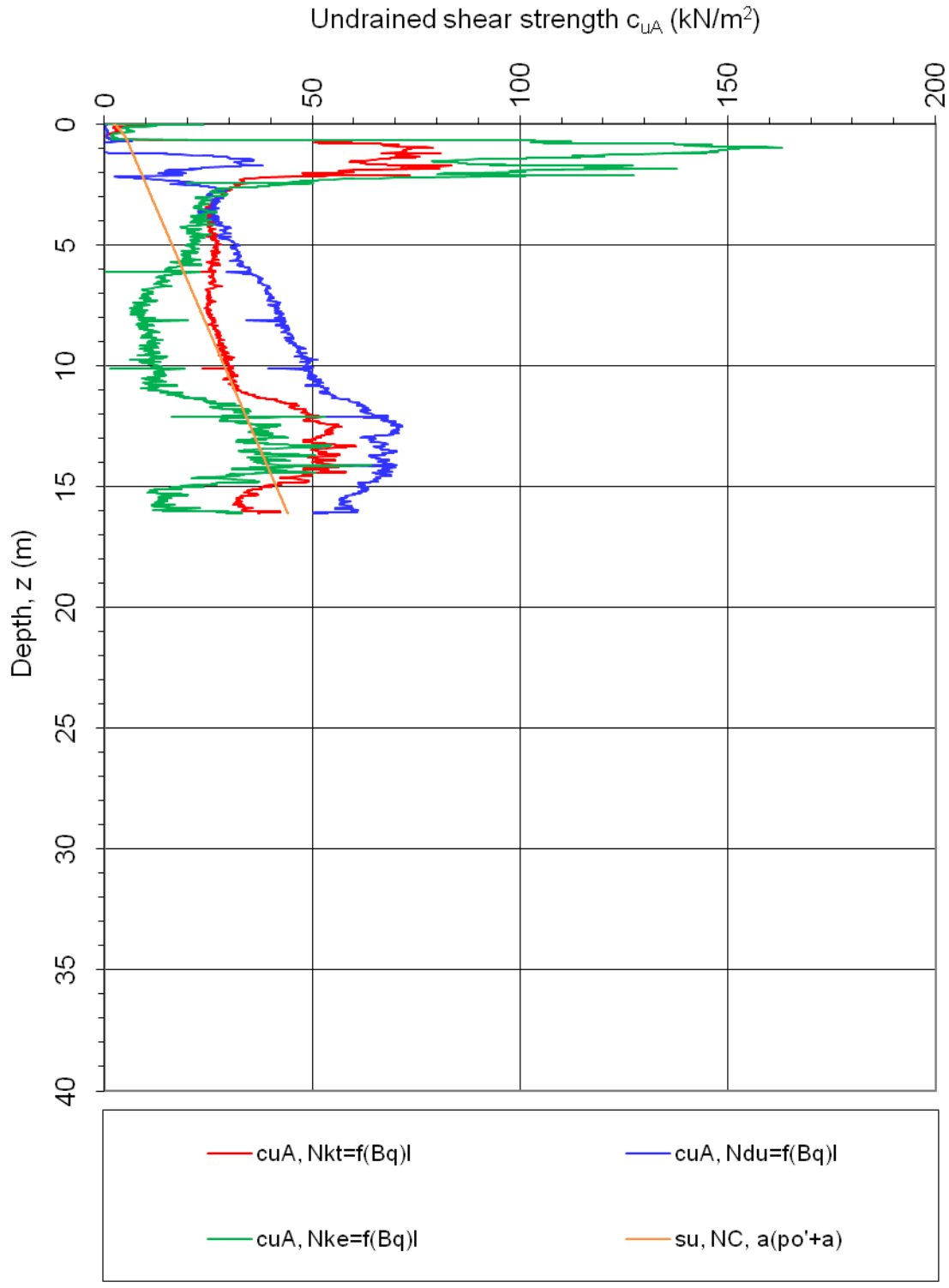


Figure 5.13.4 *Undrained shear resistance for interpreted results for Tiller RCPTu B1*

5.13.3. Sensitivity

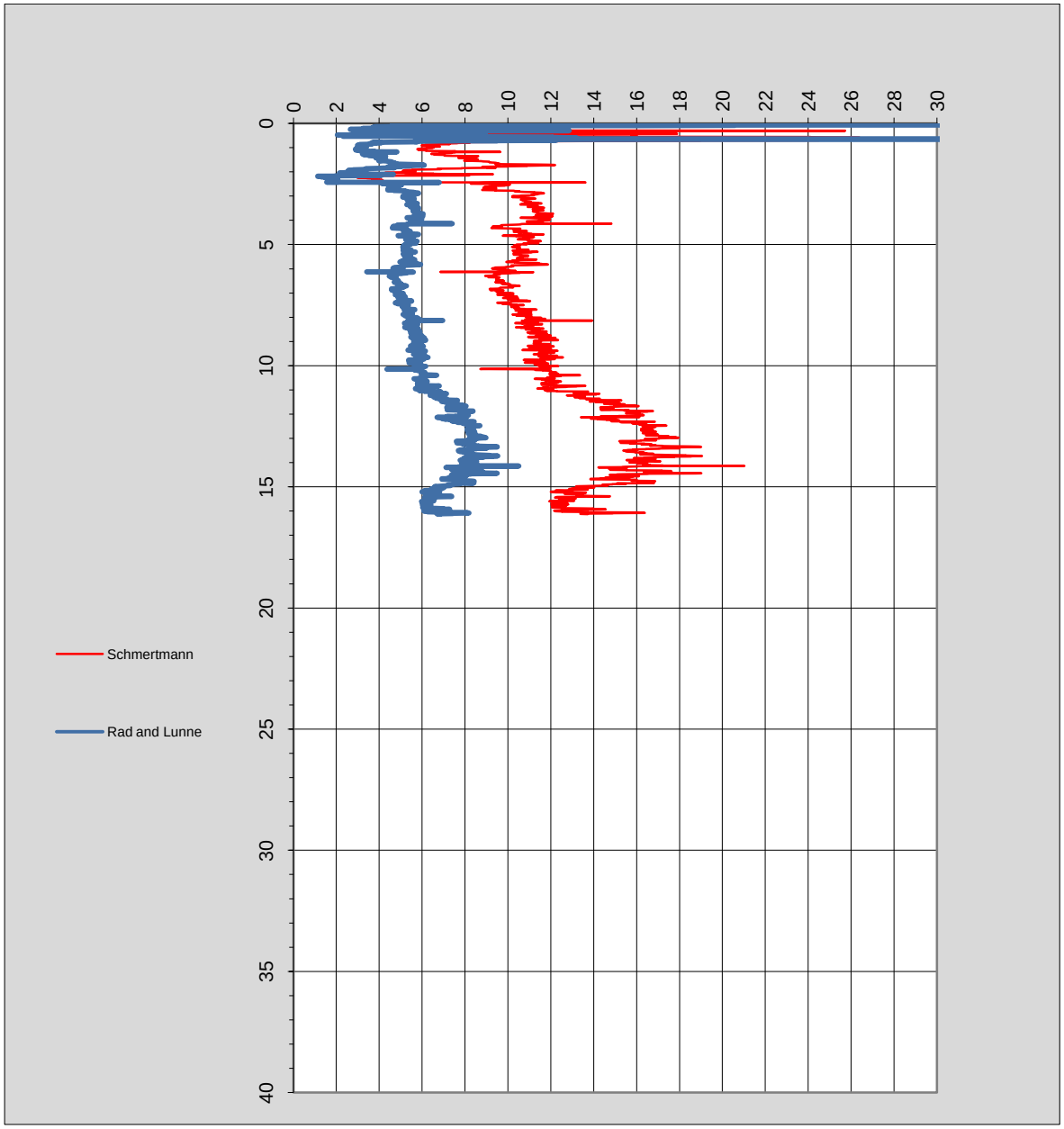


Figure 5.13.5 *Sensitivity with different parameters , results for Tiller RCPTu B1*

Comment:

From Schmertman research N_s is equal to 15, however this value is for mechanical type CPT. Rad and Lunne suggested assuming range of parameter from 5 to 10. In graph the average was taken into account N_s = 7,5.
Sensitivity values are lower than results from Fallan and Klett.

5.13.4. Resistivity

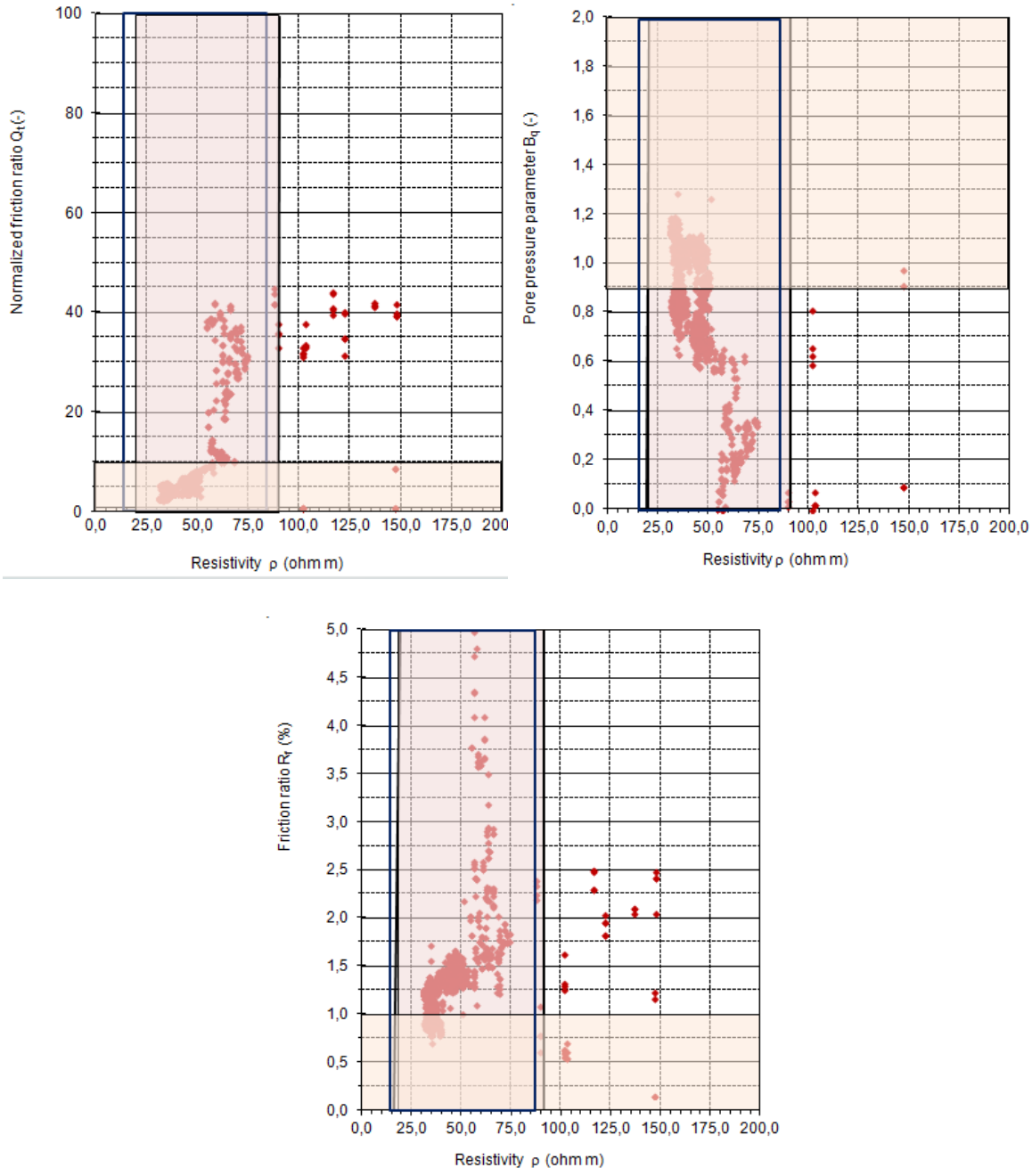


Figure 5.13.6 *Resistivity results in relation to N_m, R_f and B_q for Tiller RCPTu B1*

Comment:
Marked areas are related to soil classification tables, which are in Chapter 6 of this project. Highlights represents values for quick clay.

6. Data mining approach

6.1. Introduction

The modern world we live in is govern by information. Phenomena's like human behaviour, shopping preferences, global economy, probability of failure in engineering design - all of these and much more can be compared to a certain mechanism, which is ruled by known or undiscovered patterns. The ultimate goal for a researcher is to comprehend the data, determine the patterns and use them in a form of theorem or equations to predict values in the future.

A versatile potential of data mining can be more clear after presenting a classic example: "[...] the problem is fickle customer loyalty in a highly competitive marketplace. A database of customer choices, along with customer profiles, holds the key to this problem. Patterns of behavior of former customers can be analyzed to identify distinguishing characteristics of those likely to switch products and those likely to remain loyal. Once such characteristics are found, they can be put to work to identify present customers who are likely to jump ship. This group can be targeted for special treatment, treatment too costly to apply to the customer base as a whole. More positively, the same techniques can be used to identify customers who might be attracted to another service the enterprise provides, one they are not presently enjoying, to target them for special offers that promote this service." (Witten, Frank, Hall, 2011)

All available data is in electronic format and the process of discovering patterns must be automatic or more frequently semi-automatic. Amount of opportunities for finding relations rises in a staggering rate, due to overwhelming growth of quantity and size of databases and a tremendous number of available solutions: algorithms, data transformations, optimizations, numerous evaluation methods and etc. Decisions and choices of a researcher must be supplemented with vast knowledge of concepts, principles and mechanisms of practiced algorithms, mathematical understanding of models and finally recognition the nature of a problem.

Secondly, the idea of machine learning should be described to emphasize the difference between this concept and data mining. The process of machine learning is similar in the initial stage. However, instead of presenting calculated data to the user, it is used for further improvement of a model in order to obtain enhanced accuracy and thus higher quality of data. The process of learning in this case has only practical, nontheoretical sense. One of the most famous examples is a pattern recognition system, which is so called "artificial neural networks". A topology of such web is usually compared to a set of organic neurons, resembling operation of a human brain. Such a network

consist of a specific amount of processors, which operate parallelly - within boundaries of knowledge and accessible database.
However, no philosophical aspects of a phenomena of artificial intelligence are concerned in this thesis. Nevertheless, it is recommended for the reader to acquaint himself/herself with this controversial, yet peculiar issue.
A computer software choosen for data analysis is called WEKA - developed at the University of Waikato, New Zealand. The name is acronym for "Waikato Environment for Knowledge Analysis" and the software is free under the GNU General Public License.

6.2. Definitions of basic terms

Before proceeding any further it is extremely crucial to clarify few basic terms and phrases. It is necessary due to interdisciplinary nature of a choosen method. With the knowledge given in this subchapter it is easier to comprehend any further concepts of data mining process and challenging environment of WEKA software.
" The input takes the form of ***concepts***, ***instances***, and ***attributes***. [...] A *concept description* [...] is the result of the learning process. The information that the learner is given takes the form of a set of *instances*. Each instance is characterized by the values of attributes that measure different aspects of the instance. [...] There are many different types of attributes, although typical data mining schemes deal only with numeric and *nominal*, or categorical, ones." (Witten, Frank, Hall, 2011)
Instances are divided by comma signs; attributes assign values in a given vertical order and in all analysis only two types of attributes are used: numeric for sounding parameters (depth, pore pressure parameter or resistivity) and nominal, which determines presence of quick clay at the certain depth. This attribute represents data from other sources than soundings.

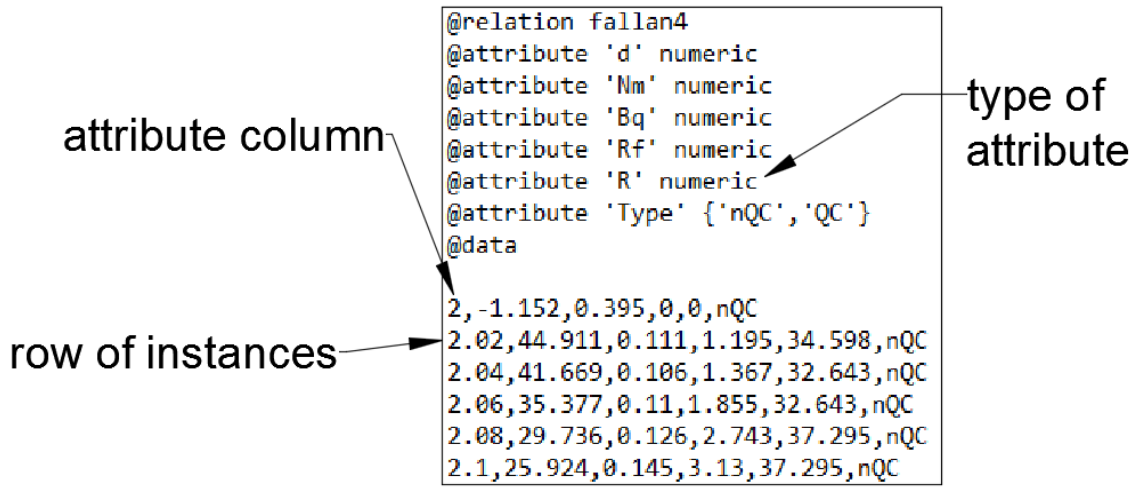


Figure 6.1 *View on typical WEKA file*

Database is a term for assembled data:

1. basic group of attributes consist of post-processed CPTu coefficients
2. measurements from resistivity module if available
3. results from other sources if available

If under certain circumstances a database lacks any informations, which were mentioned before in p.2 and p.3, then a ***missing value*** is added to the set. It is indicated as a question mark symbol "?". It is crucial to fill up blank positions, because software will treat such database as defective. Moreover, one of the machine learning techniques requires two databases: one for the training data and another for predicting values for any missing attributes. In such case, the second database will not contain any data about quick clay, so all instances in 'Type' attribute will be marked as "?". Use of training sets is described in next subchapter.

Number of instances depends on the depth of soundings: each row of values refers to soil properties every 0,02 cm - measurement rate of CPT rig.

In case of limited database a certain technique is used for verification of created algorithm: some amount of data is used for training and the rest is for testing. It is very crucial to ensure credibility of training set - quality of data will affect accuracy of a model after all. In order to meet this requirement: "Each class in the full dataset should be represented in about the right proportion in the training and testing sets. If, by bad luck, all examples with a certain class were omitted from the training set, you could hardly expect a classifier learned from that data to perform well on examples of that class — and the situation would be exacerbated by the fact that the class would necessarily be overrepresented in the test set because none of its instances made it into the training set!" (Witten, Frank, Hall, 2011) To avoid such situation a random sampling of data should be commenced - it is a process of "stratification".

However, previously mentioned solution is too primitive for uneven distribution in test and training test. Much safer and efficient results gives "**cross-validation**" technique: data base is split into approximately equal number of parts, which are specified by the user - "folds". Each single one is used for training and the rest for testing. After that a whole process is repeated with another part swapped until all are used as a training set. Standard approach include stratification combined with cross-validation in 10 folds. With successful implemented stratification whole process is referred as "**stratified tenfold cross-validation**".

On the other hand, every user have to bear in mind, that "Stratification reduces the variation, but it certainly does not eliminate it entirely. When seeking an accurate error estimate, it is standard procedure to repeat the cross-validation process 10 times that is, 10 times tenfold cross-validation— and average the results. This involves invoking the learning algorithm 100 times on datasets that are all nine-tenths the size of the original. Getting a

good measure of performance is a computation-intensive undertaking." (Witten, Frank, Hall, 2011)

6.3. Field of work

6.3.1. Use of data

In this thesis selected data mining techniques are used for detecting layers of quick clay in soil profile. Created data files consist of post-processed parameters from CPTu and RCPTu soundings:

- normalized cone resistance - "N_m" (Eq. 7)
- pore water pressure parameter - "B_q" **(**Eq. 9)
- normalized friction ratio - "R_f" (Eq. 8)
- resistivity - "R" - in case of CPTu this attribute is ignored (Eq. 21)

Sounding database was implemented with results from additional sources: laboratory data, *in situ* tests and boreholes, which allowed to divide soil into two categories. This key parameter is used for classification of data - in all models it sets boundaries for calculations, is linked to the other attributes and is presented as an output along with results (amount of classified layers):

- detection parameter - { "nQC", "QC"}
- depth - "d"

During the studies it become clear, that for available sounding results for depth parameter "d" had insignificant influence or even caused distortion in produced outcome. It was directly caused by the character of quick clay deposits: layers encountered on comparable depth with high thickness, without detection of transition layers in between. This observation confirms similarity of profiles in selected investigation sites. For that reason parameter "d" was ignored in following models: tree classifier and neural network. Research proved, that depth attribute has no impact on clustering method.

6.3.2. Decision trees

"A “divide-and-conquer” approach to the problem of learning from a set of independent instances leads naturally to a style of representation called a *decision tree*. [...] Nodes in a decision tree involve testing a particular attribute. Usually, the test compares an attribute value with a constant. Leaf nodes give a classification that applies to all instances that reach the leaf, or a set of classifications, or a probability distribution over all possible classifications. To classify an unknown instance, it is routed down the tree according to the values of the attributes tested in successive nodes, and when a leaf is reached the instance is classified according to the class assigned to the leaf."

In the case of numeric attributes, "the test at a node usually determines whether its value is greater or less than a predetermined constant, giving a two-way split. Alternatively, a three-way split may be used, in which case there are several different possibilities. If *missing value* is treated as an attribute value in its own right, that will create a third branch. An alternative for an integer-valued attribute would be a three-way split into *less than*, *equal to*, and *greater than*. An alternative for a real-valued attribute, for which *equal to* is not such a meaningful option, would be to test against an interval rather than a single constant, again giving a three-way split: *below*, *within*, and *above*. A numeric attribute is often tested several times in any given path down the tree from root to leaf, each test involving a different constant." (Witten, Frank, Hall, 2011)

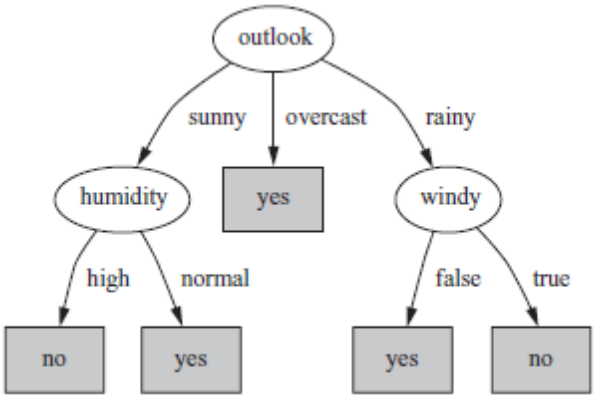


Figure 6.2 *Example of decision tree, which determines weather conditions for sport activity*

WEKA software offers a wide choice of tree classifiers, which initially were narrowed to the most useful for classifying enormous amount of instances. After series of testing a J48 tree was choosen along with its altered version: J48graft for user-friendly interface, straightforward results presented in a convenient form, short processing time and finally, for sufficient accuracy. "J48 Class for generating a pruned or unpruned C4.5 decision tree. For more information, see Ross Quinlan (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA." (WEKA description)
"C4.5 algorithm creates decision tree from a training data, which consist of already classified samples:

$$S = S_1, S_2, \dots, where \tag{21}$$

$$[S_i]_p = \begin{bmatrix} x_{1,i} \\ \cdots \\ x_{p,i} \end{bmatrix} \tag{22}$$

$$S_i - \ value\ of\ sample$$

$$x_{j,i} - \ j\ attribute\ of\ the\ sample$$

Each sample consist of a p-dimensional vector, while x_j represents attributes of features of the sample [...] as well as the class in which S_i falls."

6.3.3. Clustering

An accurate visualization of clusters output is a diagram with instances. Each cluster can be described as a mathematical set, which consist of certain amount of numbers grouped with a given pattern. There is a possibility, that one instance can belong to more than one cluster, which causes overlapping subsets (Figure 6.3b). Even more, an instance may belong to a group with a certain probability or in hierarchical manner (Figure 6.3b).

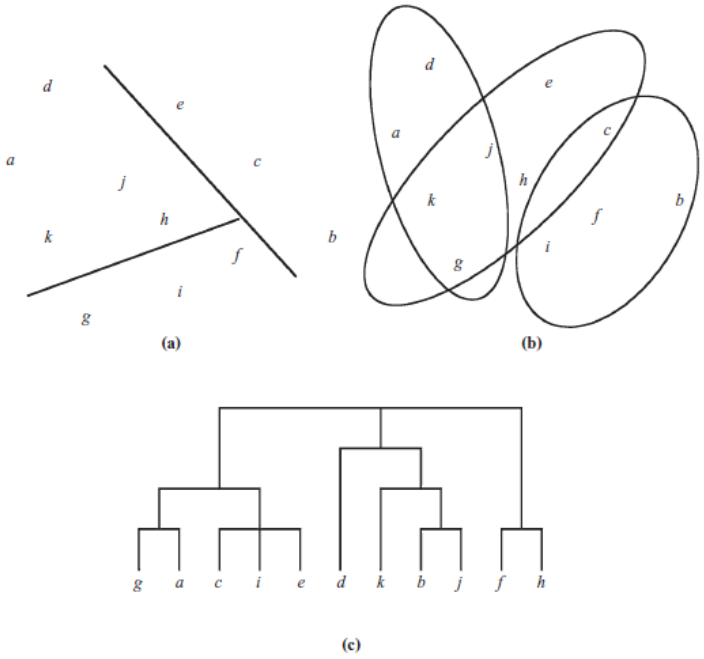


Figure 6.3 *Possible visualization of cluster classification*

In this method instances are group regardless the class they belong to, usually when there is no need to classify given data. Cluster group instances of closest resemblance in subsets. Those domains are govern by a concealed mechanism, which in specific way distinguishes instances in their natural group from the disqualified ones.
"*DBScan* uses the Euclidean distance metric to determine which instances belong together in a cluster (Ester et al., 1996), [...] it can determine the number of clusters automatically, find arbitrarily shaped clusters, and incorporate a

notion of outlier. A cluster is defined as containing at least a minimum number of points, every pair of points of which either lies within a user-specified distance (ε) of each other or is connected by a series of points in the cluster that each lie within a distance ε of the next point in the chain. Smaller values of ε yield denser clusters because instances must be closer to one another to belong to the same cluster. Depending on the value of ε and the minimum cluster size, it is possible that some instances will not belong to any cluster. These are considered outliers." (Witten, Frank, Hall, 2011)
DBScan was chosen due to convenient output format and promising convergence of clustered data with laboratory data. Despite the fact, that computation is performed in a longer time than other clusterers, very rarely layers from various depths are grouped together.

6.3.4. Neural networks

In order to clarify idea behind neural network and mechanism of calculations a sequence of terms and definitions must be presented beforehand.
First of all, a typical neural network is build from "neurons", which are connect with each other. Such a basic unit is called a "perceptron" - node, which links all input parameters with the classified data. Classification of data is dictated with following algorithm:

$$f(t_v + w_0 * a_0 + w_1 * a_1 + \ldots + w_n * a_n) = c_v \quad \textbf{(23)}$$

$$a_n - value\ of\ n\ attribute$$

$$w_n - weight\ for\ n\ attribute$$

$$t_v - threshold\ value$$

$$c_v - classifying\ value, usually\ f_v \in < 0{,}1 >$$

Result of a perceptron defines type of a class in output - if the value is between 0 and 1, then response is formulated as probability. Threshold value is computed at the very moment of calculation of the model. It is the only data generated by the algorithm - others are given by the user beforehand.

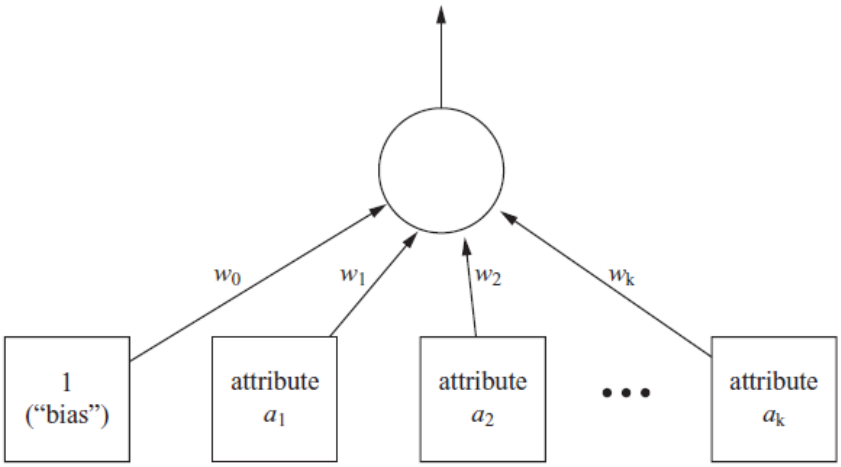


Figure 6.4 *Perceptron as a fragment of a neural network*

Threshold value prevents from dividing by zero case by the "numeric membership function". Two most popular cases include hyperbolic tangent or sigmoid functions:

$$f(x) = tanh \quad \textbf{(25)}$$

$$f(x) = \frac{1}{1+e^{-x}} \quad \textbf{(26)}$$

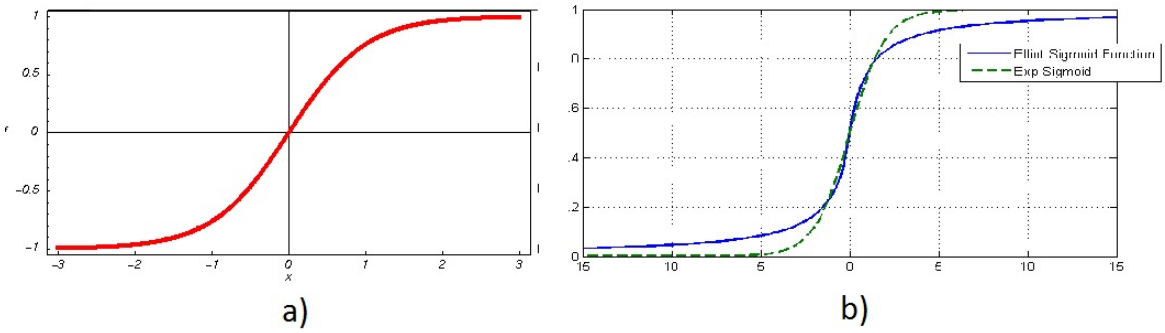


Figure 6.5 *Visualization of membership functions*

For a neural network classifier more traditional solution was choosen: a multi-layer perceptrons, which is constructed from perceptron modules linked in a hierarchical order. In this case only a sigmoid function is in use. To build a classifier with logical connectives multiple perceptrons are required. Such a model has similar expressive power as a decision tree; "[...] each node in the hidden layer has the same role as a leaf in a decision tree or a single rule in a set of decision rules." (Witten, Frank, Hall, 2011) Figure 6.6 represents a simple a multi-layer perceptron network with single layer of "hidden layer". Hidden neurons contains units, which have no direct connection with the environment.

To understand multilayer perceptron mechanism two additional aspects must be disclosed:

- origin of attributes weights - they are calculated by a "backpropagation" algorithm, which includes differentiating the error function. Well-trained network is distinguished by minimized squared error of output, which directly translates into accuracy of the classifier.
- what structure of model is most efficient - alas, this can be determined only be experimentation. Sometimes a character of instances in data-base requires just a single hidden layer with few nodes for optimal results.

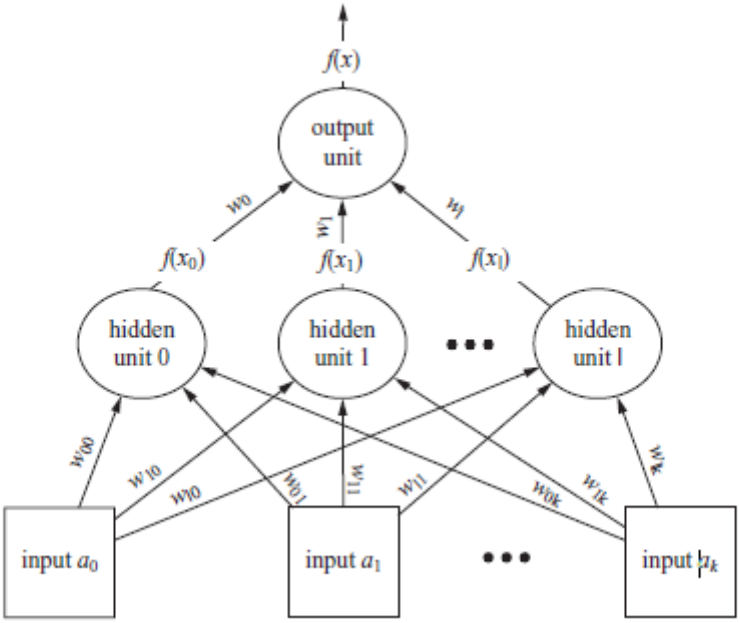


Figure 6.6 *Example of neural network*

7. Data mining results

7.1. Tree classification

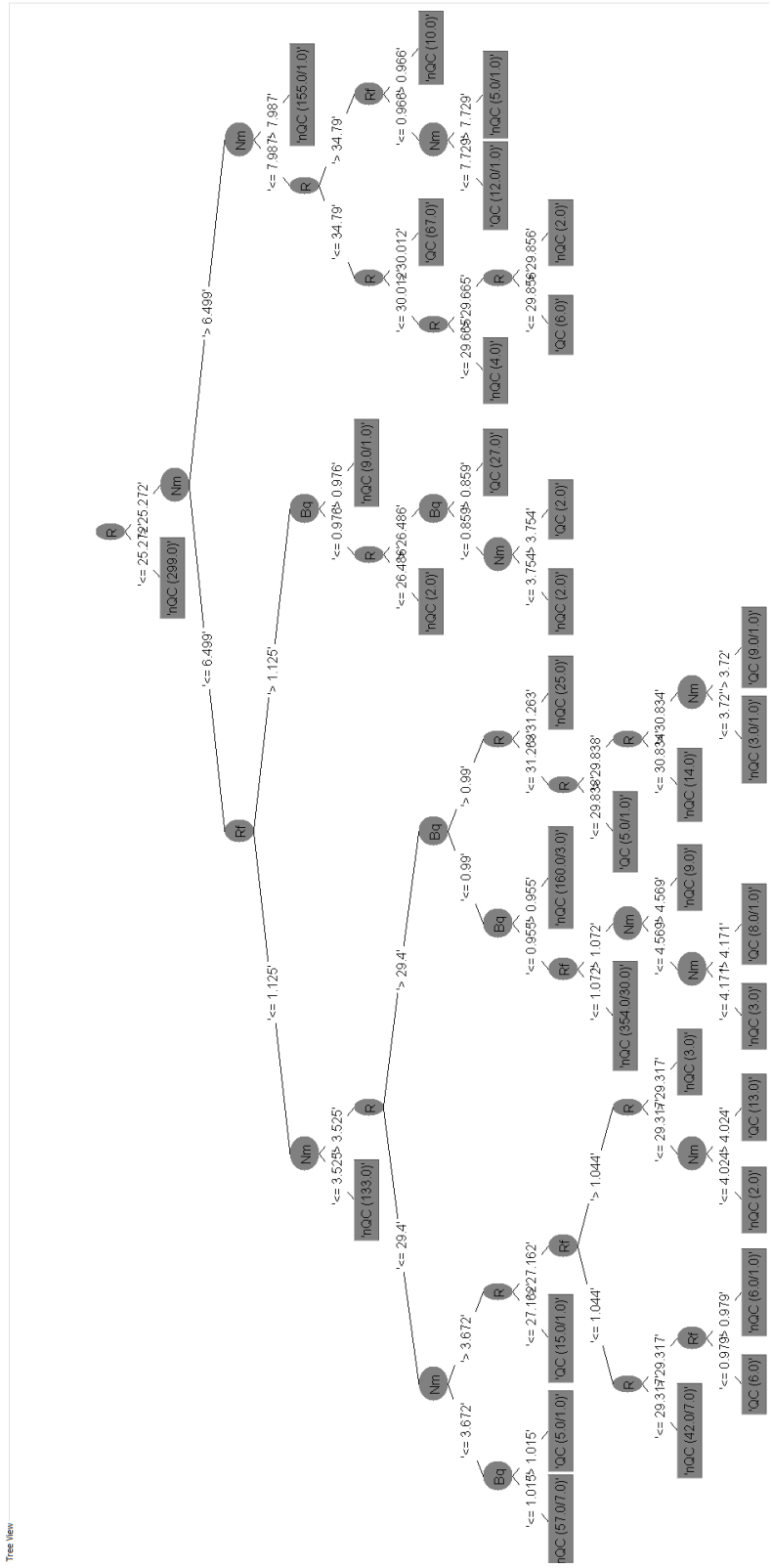


Figure 7.1 *Tree classifier J48 for Fallan BP 2*

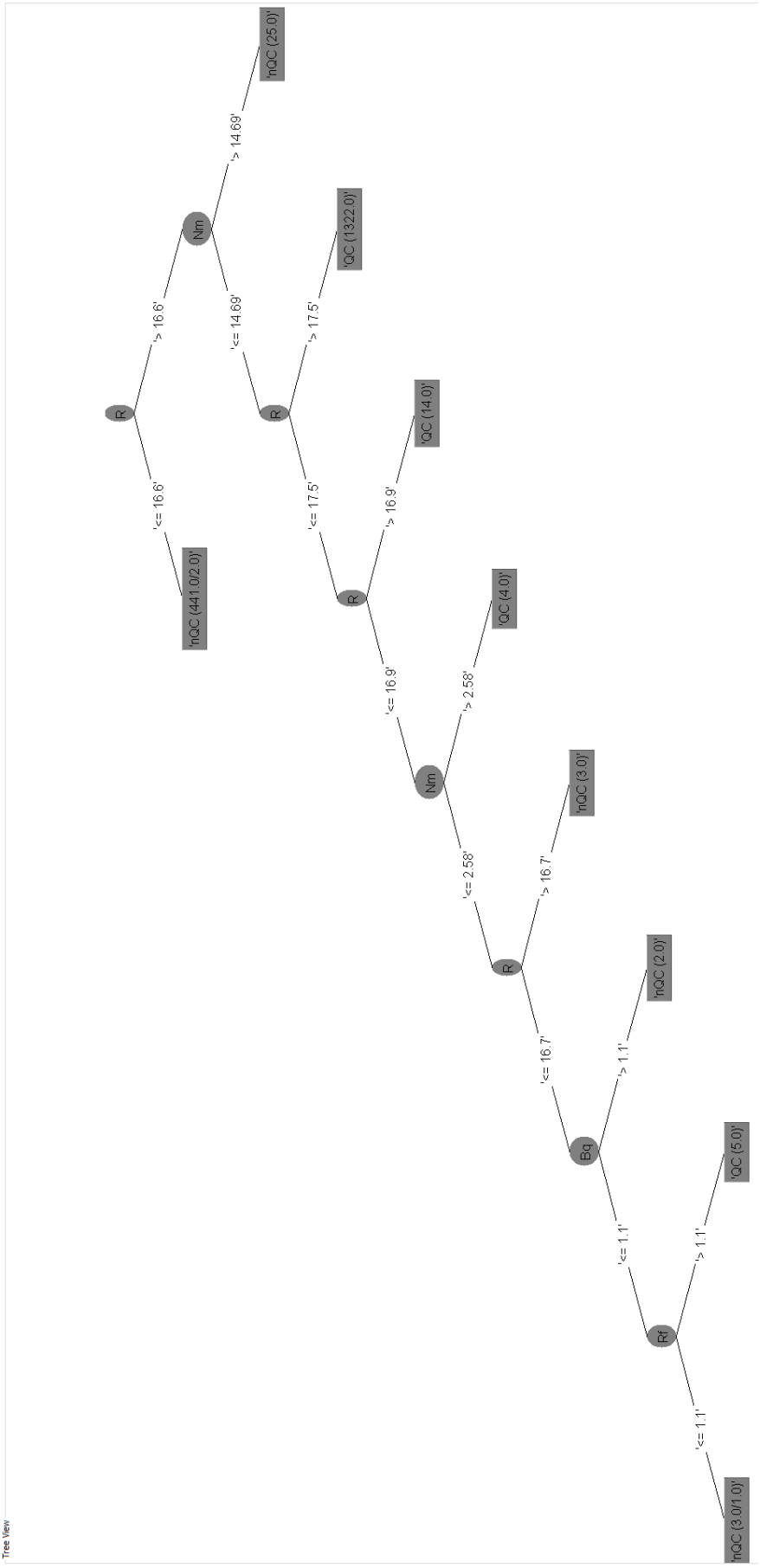


Figure 7.2 *Tree classifier J48 for Klett S1*

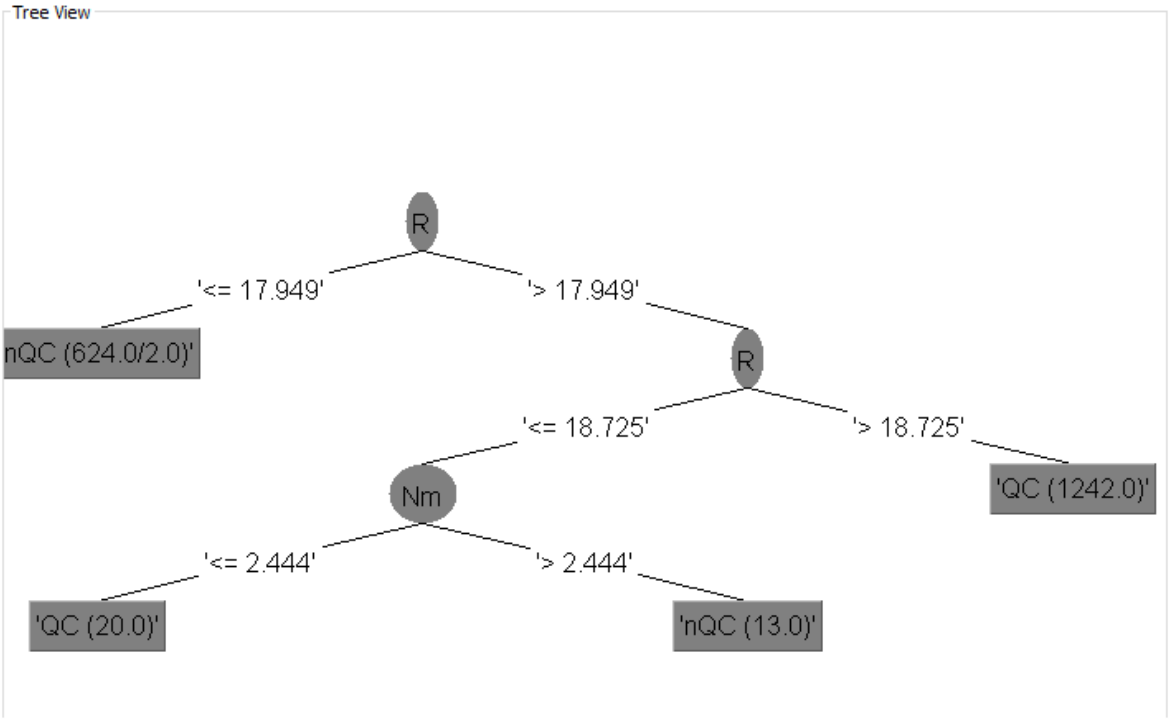


Figure 7.3 *Tree classifier J48 for Klett S2*

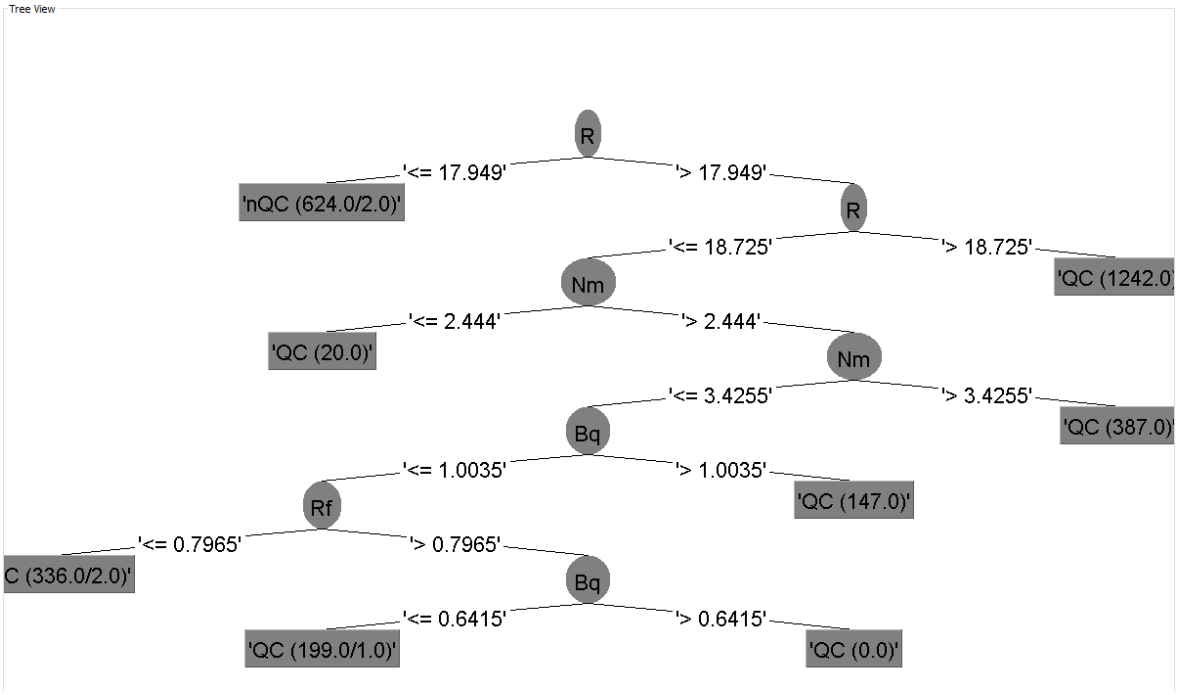


Figure 7.4 *Tree classifier J 48 graph for Klett S2*

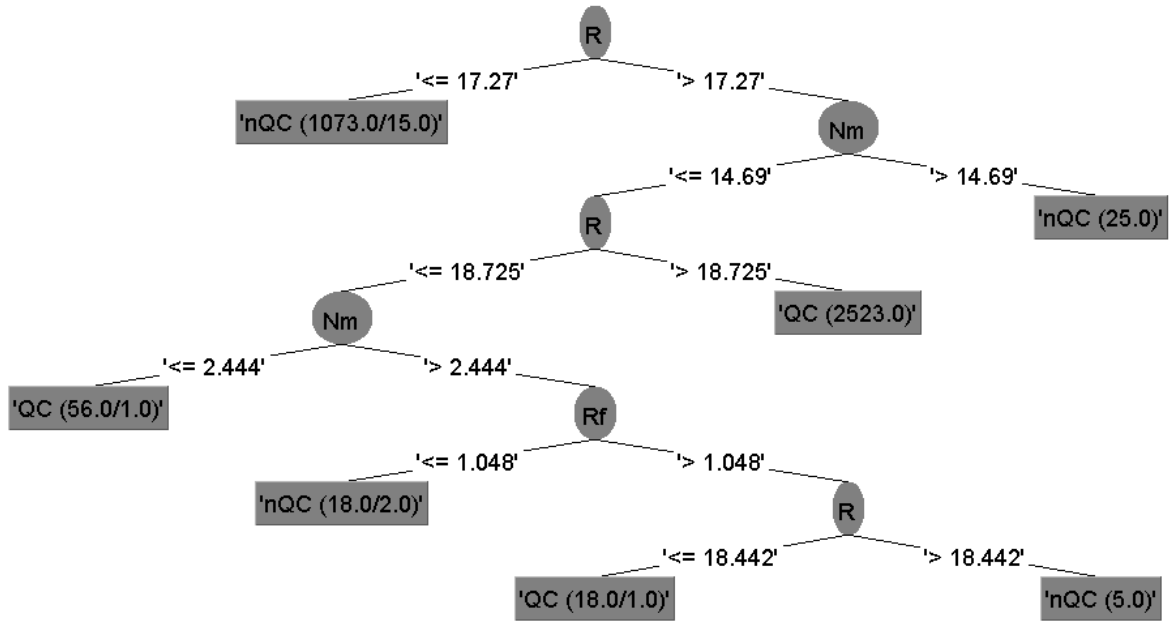


Figure 7.5 *Tree classifier J48 for combined database Klett S1 and S2*

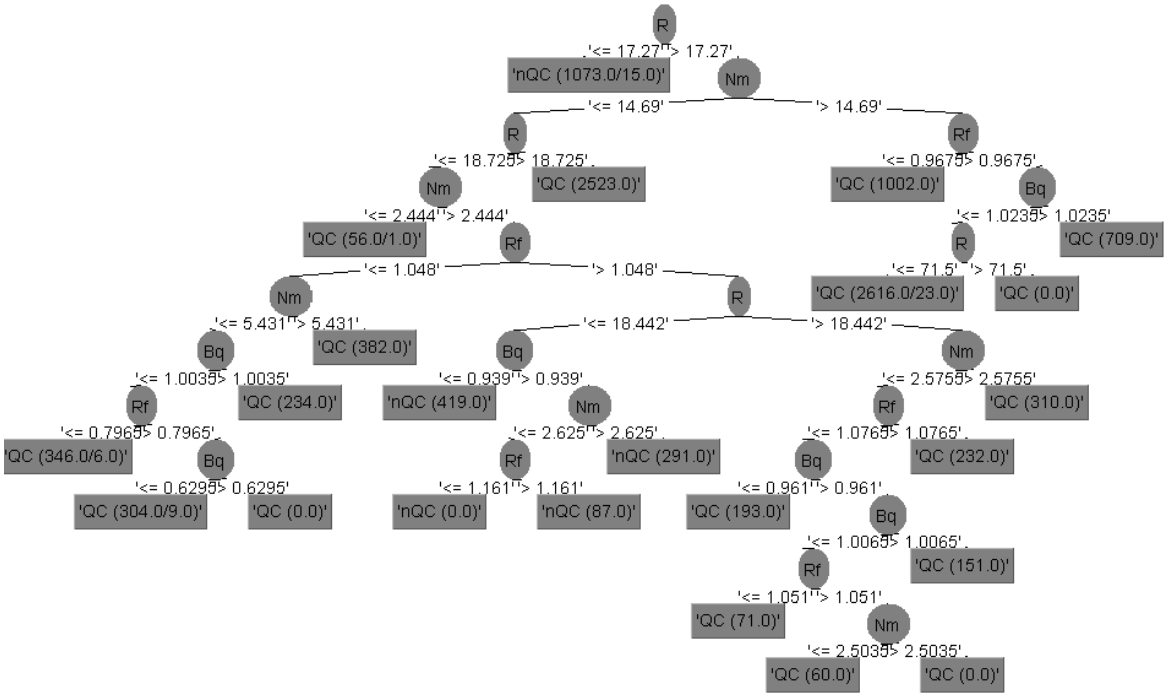


Figure 7.6 *Tree classifier J 48 graph for combined database Klett S1 and S2*

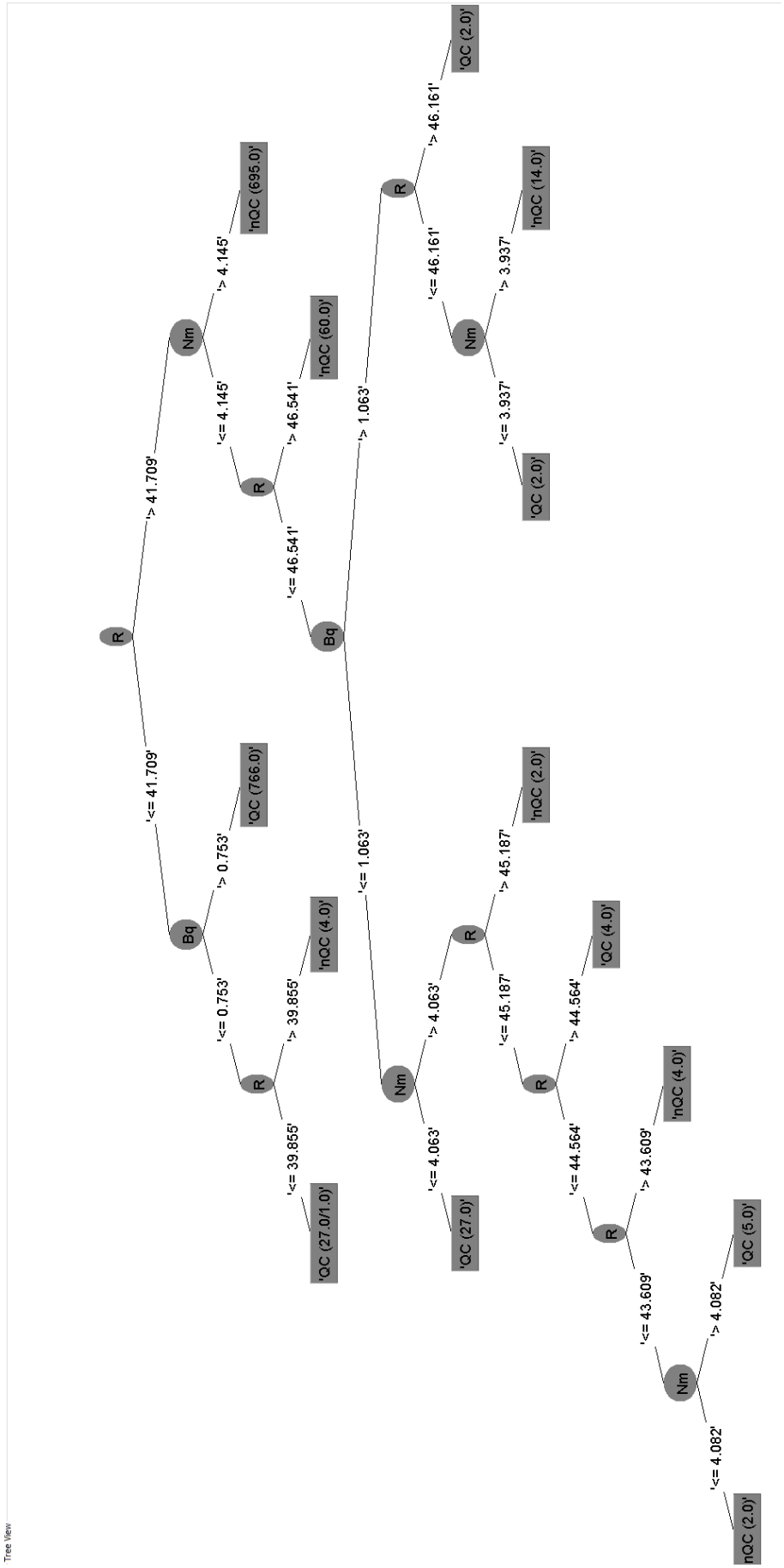


Figure 7.7 *Tree classifier J48 for Tiller*

Table 7 - *Accuracy of J48 and J48 graph classifier for datasets implemented with laboratory data*

Training set			Fallan	Klett					Tiller
			RCPTu	RCPTu					RCPTu
Test set			2	S1	S2	S2*	S1&S2	S1&S2*	B1
Fallan	RCPTu	2	**96,07**	87,25	88,94	88,40	89,62	88,87	91,18
Klett	RCPTu	S1	28,75	**99,84**	96,26	96,26	99,06	99,06	52,34
		S2	35,60	97,42	**99,89**	**99,89**	99,89	99,89	44,13
Tiller	RCPTu	B1	35,07	61,40	51,61	51,61	61,40	54,09	**99,94**

Note: Initial accuracy for S1&S2 case is 99,49; S2* and S1&S2* columns represents J48graphs

For evaluation of model's quality, a quantity of leaves, their content and repeatability of attributes are taken into account. Number of instances is given in the brackets in every single leaf. Small sized leafs do not have a significant impact on the performance of the model, especially if placed at the end of a classifier. More complex classifiers tends to consist of high number of leafs with small number of instances.

Basing on the presented models the following observations can be made:

1. Model trained on Tiller dataset gives unexpectedly high reliability for Fallan sounding - 91,18 %. However, for the reversed case the accuracy of 35,07% is obtained.
2. Most complicated tree is trained on the Fallan results. When used for classifying datasets from different investigation sites, a tremendously lower accuracy is acquired.
3. Friction ratio R_f does not appear in some models at all. Even if this parameter is in use, then is placed only in low-stage leafs.
4. Classifier trained on the combined database from Klett gives better results than trained on single sounding. Moreover, better performance is also noted for data from the same investigation site.

For the purpose of simulation most succeeding models with highest accuracy are choosen:

- for Fallan CPTu and RCPTu - models J48 trained on Fallan and Tiller datasets
- for Klett CPTu - model J48 trained on combined Klett database

Predicted profiles are presented on the next pages.

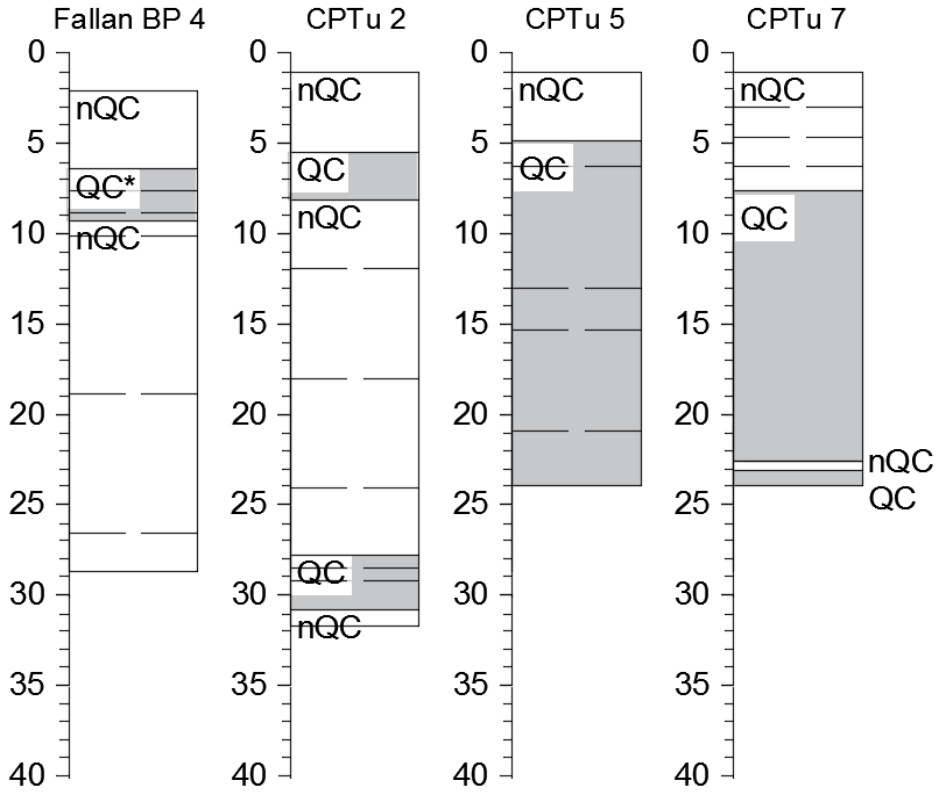


Figure 7.8 *Predictions based on Fallan BP 2 model - dashed lines indicate data distortion*

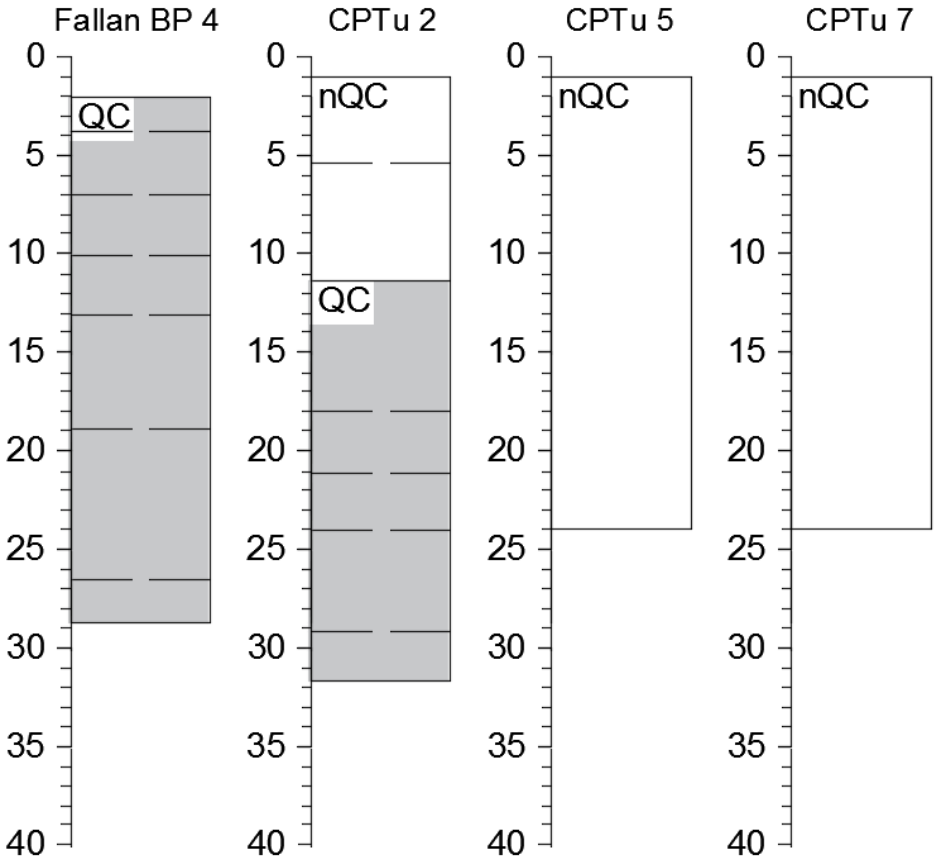


Figure 7.9 *Predictions based on Tiller 1 model - dashed lines indicate data distortion*

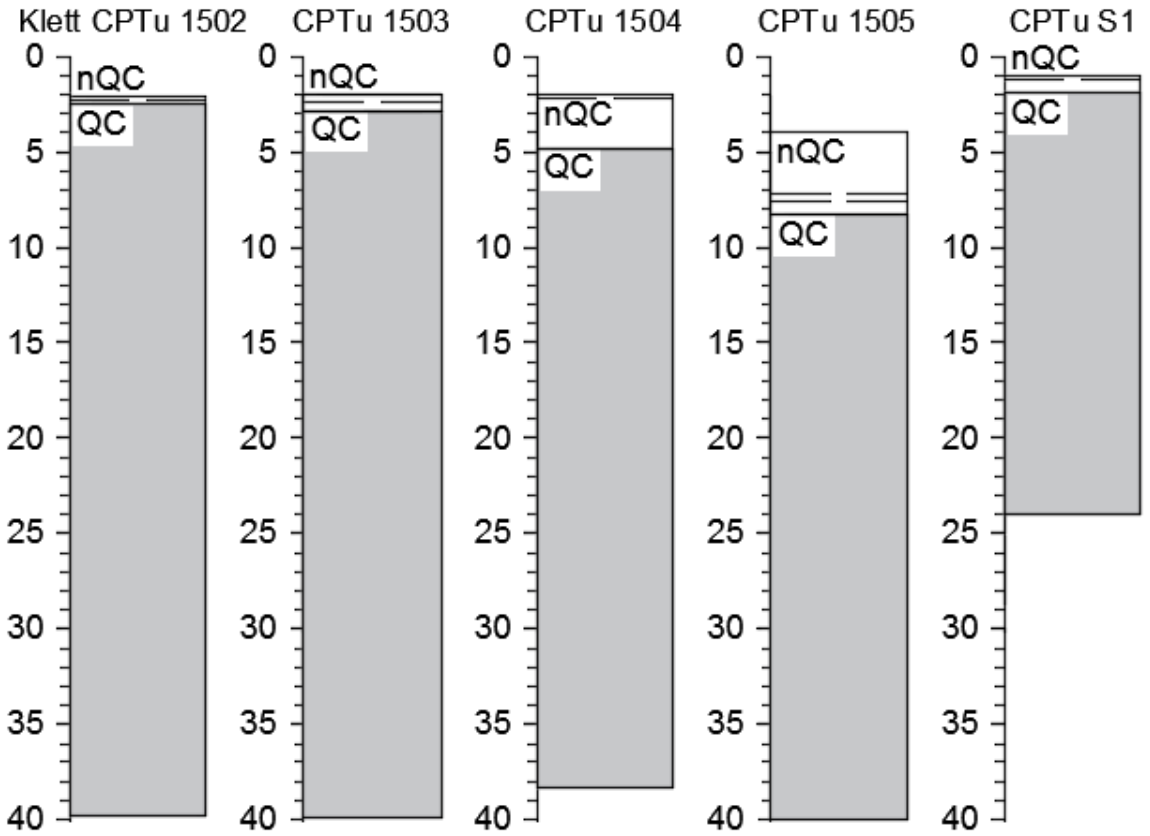


Figure 7.10 *Predictions based on Klett 1.2 combined model - dashed lines indicate data distortion, which in this case are minimal*

Profiles from Fallan and Tiller models do not present reliable predictions due to:

- lack of similarity in the set of layers between profiles from the same model - it is expected to acquire comparable results
- intensity of data distortion - packs of instances, which forms a thin layer (2-10 cm) classified to an opposite soil group
- diversity of layers from Fallan model; this indicates a problem in data classification, which might be caused by a complex structure of the trained model
- homogenous profiles from Tiller model proved tendency towards generalization of data; origin of this trend can be pointed as a trait of Tiller investigation site - simple set of layers with undifferentiated parameters

On the other hand, Klett model gave dependable results, which are comparable to the laboratory data. Few distortions can be treated as an indication of a transition layer. Obtained profiles have similar characteristics, which is feasible for a set of soundings from a single investigation site.

7.2. Clustering

7.2.1. Fallan 2 models

First of all, the following subchapter contains graphs generated by WEKA software. Represented charts plot calculated soil profiles. However, the vertical axis is inverted and scaled to the largest depth value - maintaining any proportions between various results is very laborious.
At this stage, it is important to refer to p. 6.3.1 of this thesis - the matter of the depth parameter in different data mining techniques. Analysis of two models on a single Fallan RCPTu database presents identical results (with included and excluded depth attribute). Model b) on Figure 7.5 recognizes layout of soil layers without given depth. In this case clusters are plotted versus instance number.

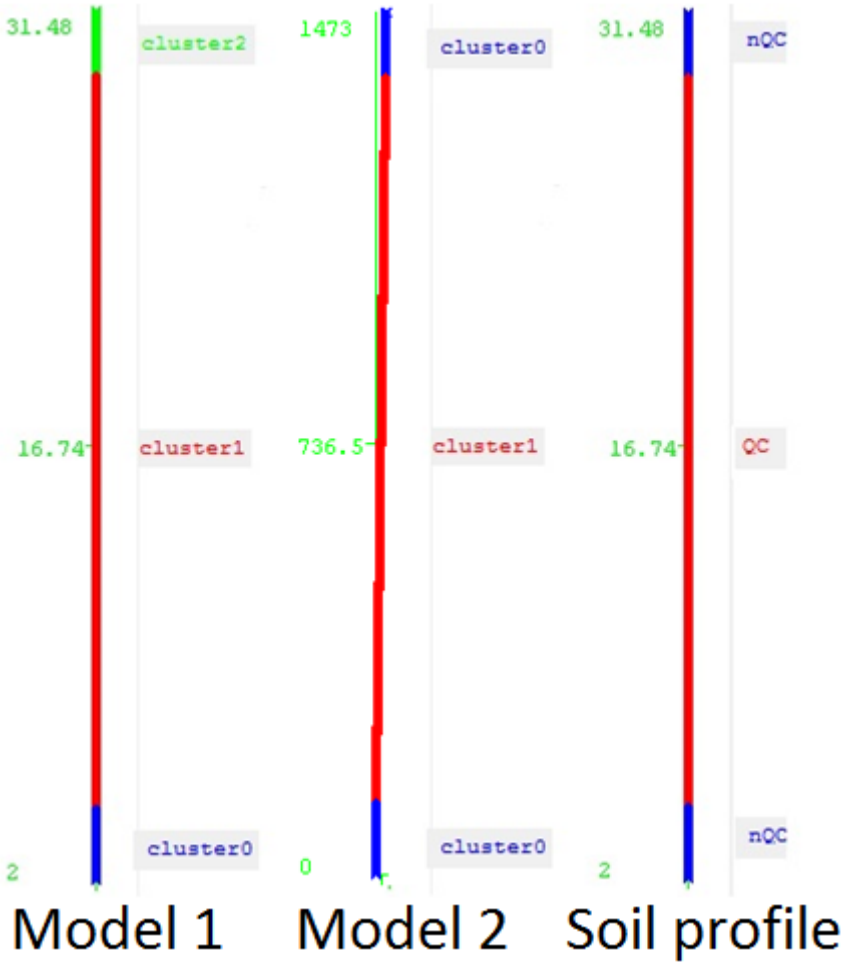


Figure 7.11 *Cluster assignments in trained models 1) & 2) compared with soil class attribute 3) in Fallan BP 2*

Table 8 - *Division of instances in trained models 1 and 2*

Clustered instances				
Number of cluster	Model 1		Model 2	
	Number	[%]	Number	[%]
0	128	9	232	16
1	1242	84	1242	84
2	104	7	-	-

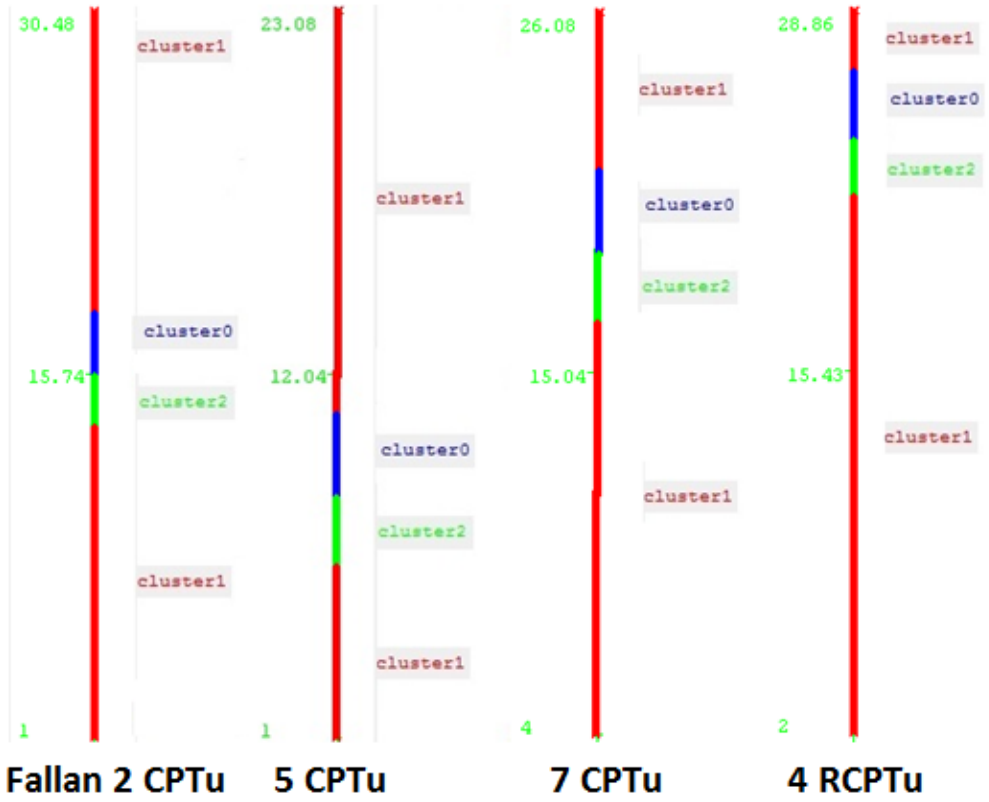


Figure 7.12 *Calculated cluster assignments in CPTu data by Fallan model*

Table 9 - *Division of instances for modeled soundings*

Sets tested by model trained on Fallan BP 2								
Number of cluster	Fallan 2 CPTu		Fallan 5 CPTu		Fallan 7 CPTu		Fallan 4 RCPTu	
	Number	[%]	Number	[%]	Number	[%]	Number	[%]
0	128	9	128	12	128	12	128	10
1	1242	84	873	79	873	79	1111	83
2	104	7	104	9	104	9	104	8

Comment:
Clusters 0 and 2 contain a constant number of instances despite of an evident difference in values between the attributes. In addition, none of the results resembles the original soil profile (Fig. 7.11). It is a severe flaw for a trained model, which disqualifies it from any practical applications.

7.2.2. Klett models

Two databases with investigated soil profile give broader range of possible models and forms of a verification. Three related models were formed for soundings from Klett investigation site. It is expected, that analysis of output will give more or less details of data dependency and trends for combined databases.

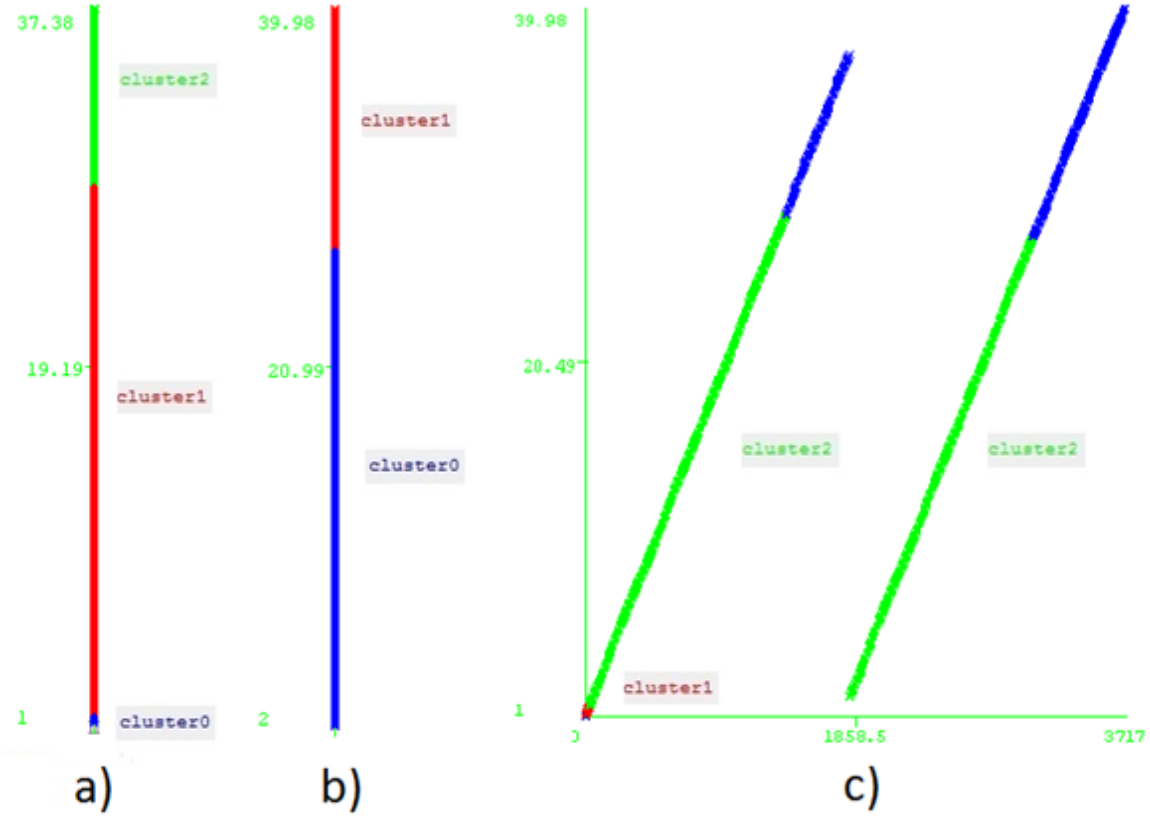


Figure 7.13 *Cluster assignment for a) Klett S1, b) Klett S2 and c) combined*

Table 10 - *Division of instances for separated and combined databases*

Clustered instances						
Number of cluster	Model Klett 1		Model Klett 2		Combined model	
	Number	[%]	Number	[%]	Number	[%]
0	25	1	1263	67	1081	29
1	1348	74	635	33	25	1
2	445	24	-	-	2612	70

Comment:
Combined data is plotted: depth (vertical axis) versus number of instance (horizontal). WEKA software treats databases in a simple way - thus numeration was continued after last instance from Klett S1. Consequently, considered data was plotted in form of two perpendicular segments.

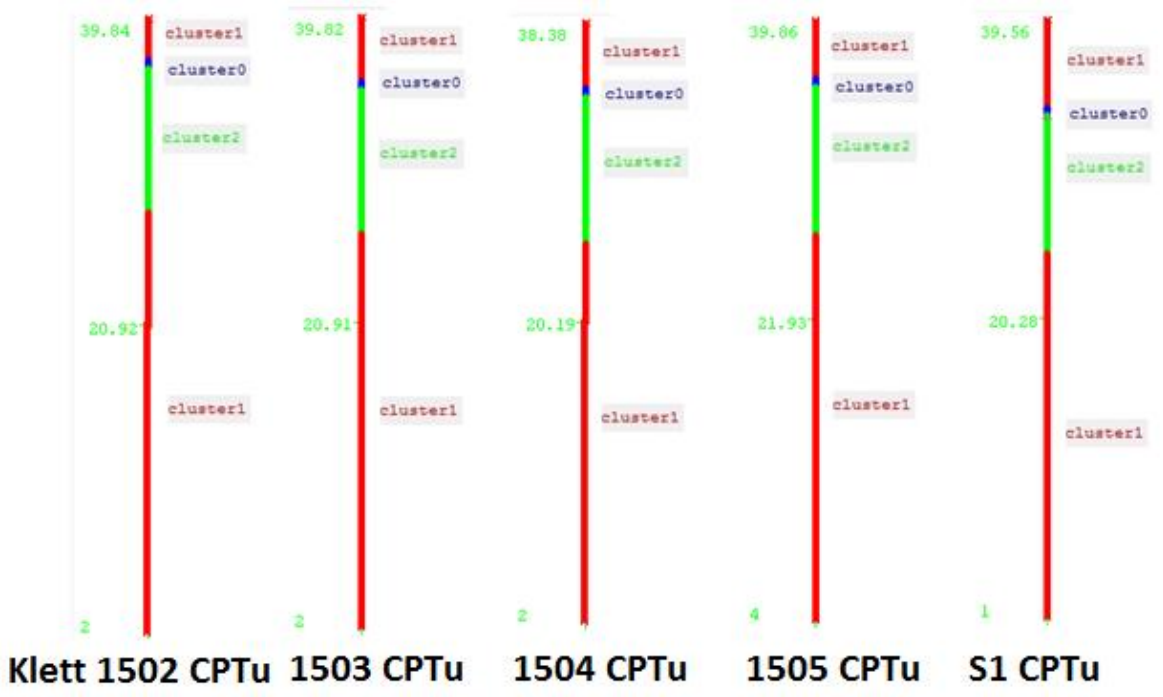


Figure 7.14 *Cluster assignments in CPTu data for Klett 1 model*

Table 11.1 - *Division of instances for modeled soundings*

Sets tested by model trained on Klett 1						
Number of cluster	Klett 1502 CPTu		Klett 1503 CPTu		Klett 1504 CPTu	
	Number	[%]	Number	[%]	Number	[%]
0	50	3	25	1	25	2
1	1395	74	1420	75	1348	74
2	445	24	445	24	445	24

Table 11.2 - *Division of instances for modeled soundings*

Sets tested by model trained on Klett 1			
Klett 1505 CPTu		Klett S1 CPTu	
Number	[%]	Number	[%]
25	1	25	1
1322	74	1457	76
445	24	445	23

Comment:
Data from borings indicates a lower surface of soft clay layer (cluster 1) at the approximate depth 27-28 m below the surface. Calculated results predicts this surface at 27-29 m depth, which is considered as accurate prognosis. Nonetheless, this model detects another layer of quick clay below 37-39 m, which requires further investigating.

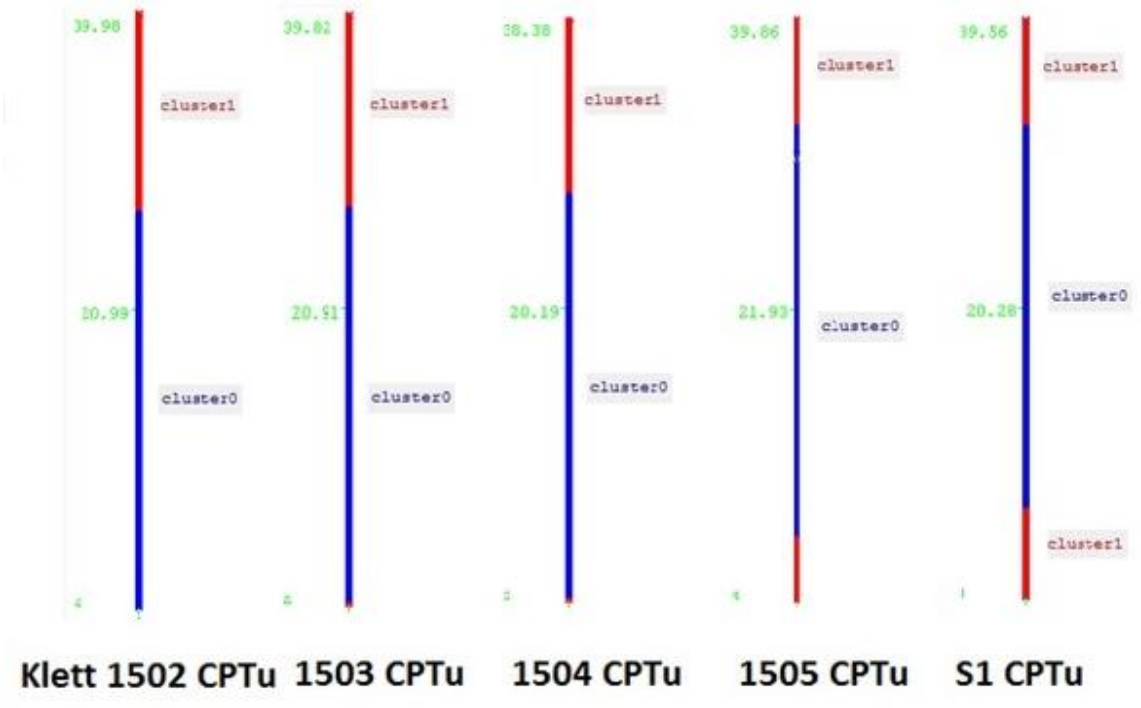


Figure 7.15 *Cluster assignments in CPTu data for Klett 2 model*

Table 12.1 - *Division of instances for modeled soundings*

Sets tested by model trained on Klett 2						
Number of cluster	Klett 1502 CPTu		Klett 1503 CPTu		Klett 1504 CPTu	
	Number	[%]	Number	[%]	Number	[%]
0	1263	67	1263	67	1263	69
1	628	33	627	33	555	31

Table 12.2 - *Division of instances for modeled soundings*

Sets tested by model trained on Klett 2			
Klett 1505 CPTu		Klett S1 CPTu	
Number	[%]	Number	[%]
1263	70	1263	66
529	30	664	34

Comment:
Presented results plot quick clay soil as cluster 0. Depth and thickness of layers are comparable to boring profiles. This is the only model, which reveals non-quick clay layer near the surface.

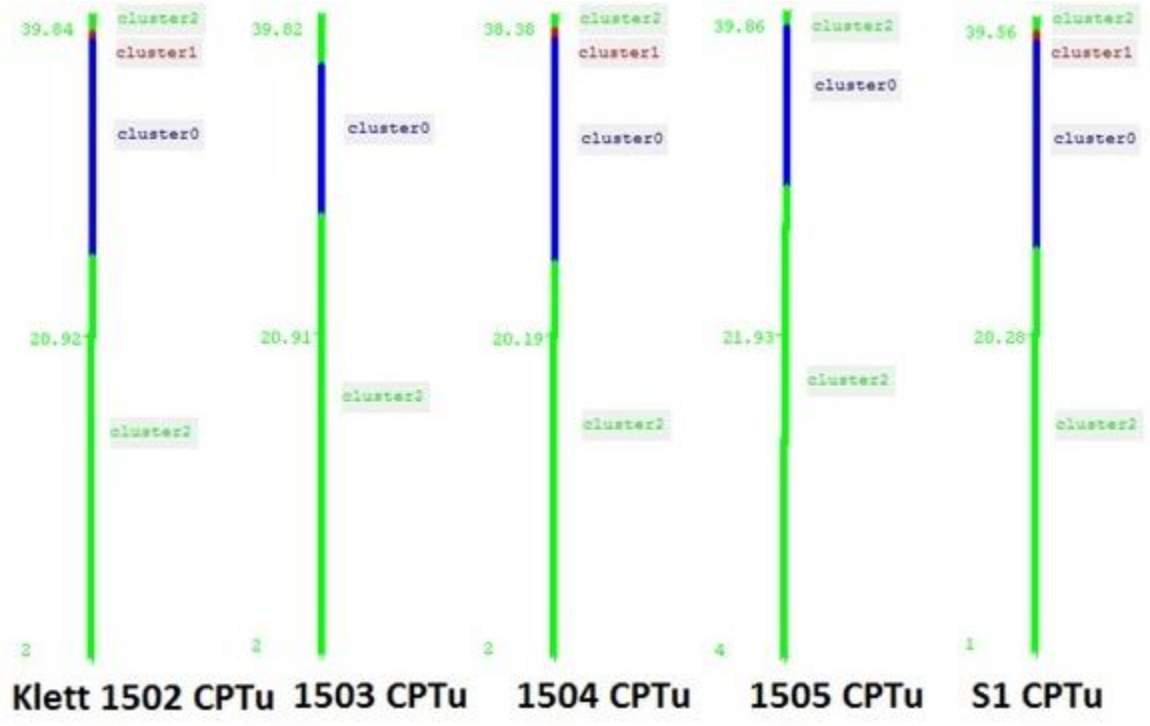


Figure 7.16 *Cluster assignments in CPTu data for combined model*

Table 13.1 - *Division of instances for modeled soundings*

Sets tested by combined model						
Number of cluster	Klett 1502 CPTu		Klett 1503 CPTu		Klett 1504 CPTu	
	Number	[%]	Number	[%]	Number	[%]
0	446	24	445	26	636	35
1	25	1	0	-	25	1
2	1421	75	1446	74	1158	64

Table 13.2 - *Division of instances for modeled soundings*

Sets tested by combined model			
Klett 1505 CPTu		Klett S1 CPTu	
Number	[%]	Number	[%]
445	25	636	33
0	-	25	1
1348	75	1267	66

Comment:
Combined model reveals layout, which is similar to results from Klett 1. Both clusters indicates an additional layer of quick clay below 37-39 meters. Secondly, changes in transitional layer at depth 26-28 meters are analogous in both models.

7.3. Neural networks

7.3.1. Presentation and accuracy of models

Following research includes four schematics for neural networks, which are trained on different datasets. All versions are tested by the remaining datasets and a final model with a best performance is choosen.

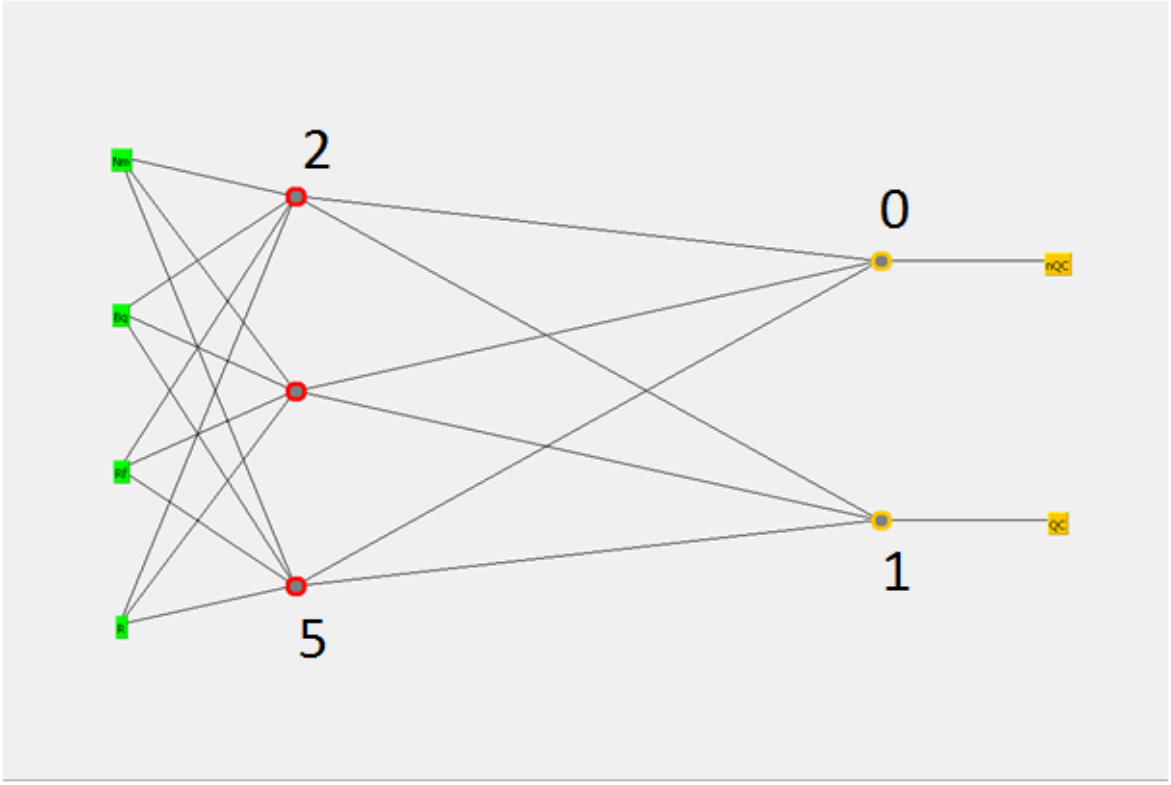


Figure 7.17 *Version 1 - ordinal number of outermost neurons is indicated*

Comment:
It is a basic version of neural network, which was generated initially by the multilayered perceptron classifier. This is one of the few classifiers, which do not require a prepared outline of network from the user. Amount of attributes in the input layer (green marked) affects generated number of nodes in a first hidden layer.

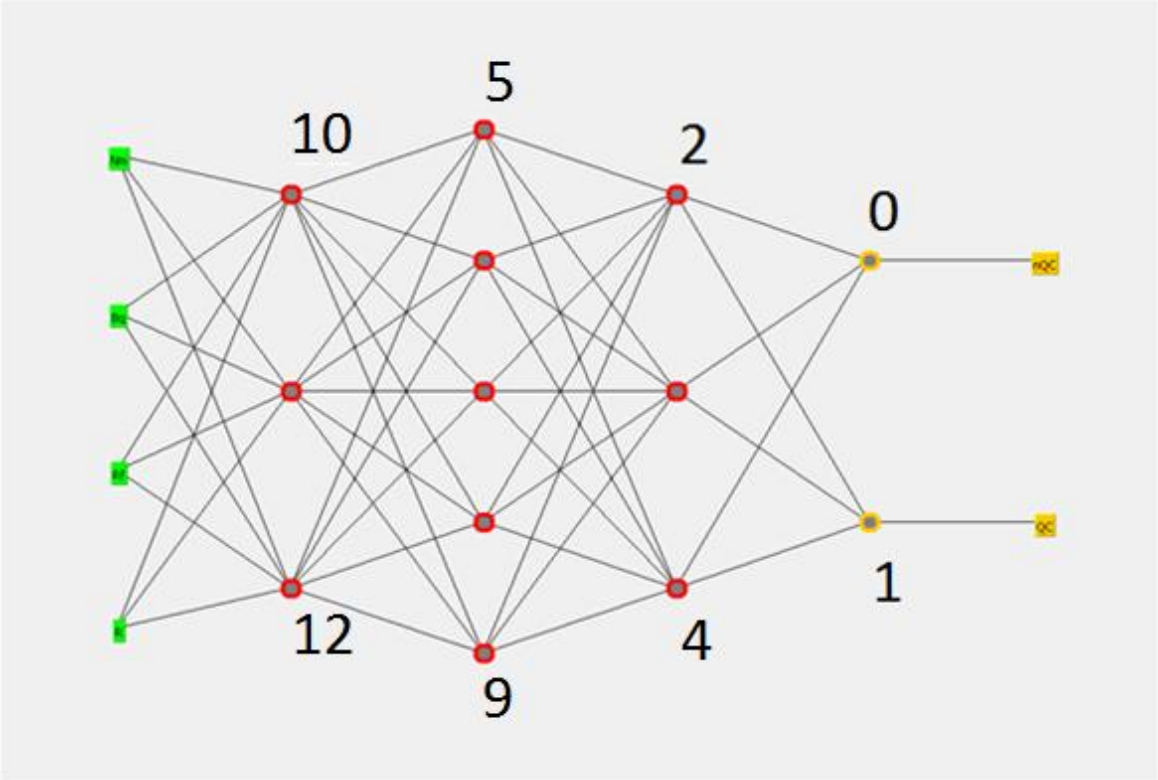


Figure 7.18 *Version 2 - ordinal number of outermost neurons is indicated*

First step: Two additional hidden layers are added in a symmetrical order.

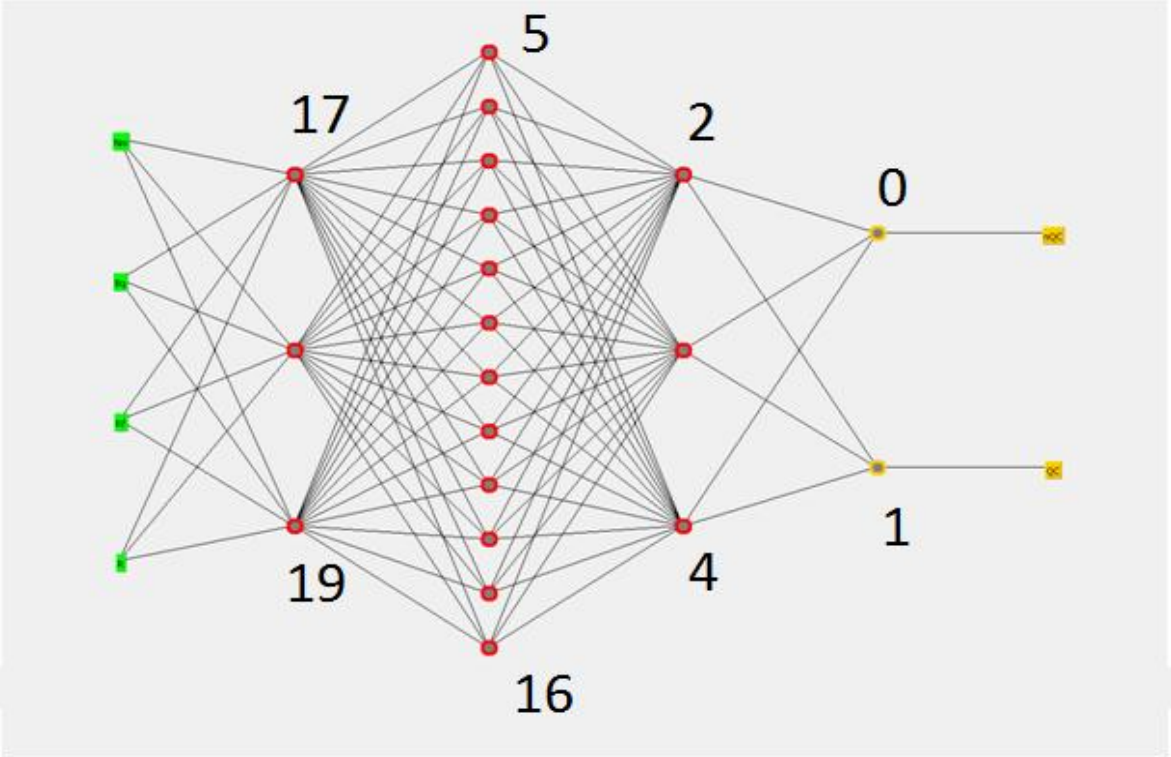


Figure 7.19 *Version 3 - ordinal number of outermost neurons is indicated*

Second step: Number of nodes in the middle hidden layer is increased.

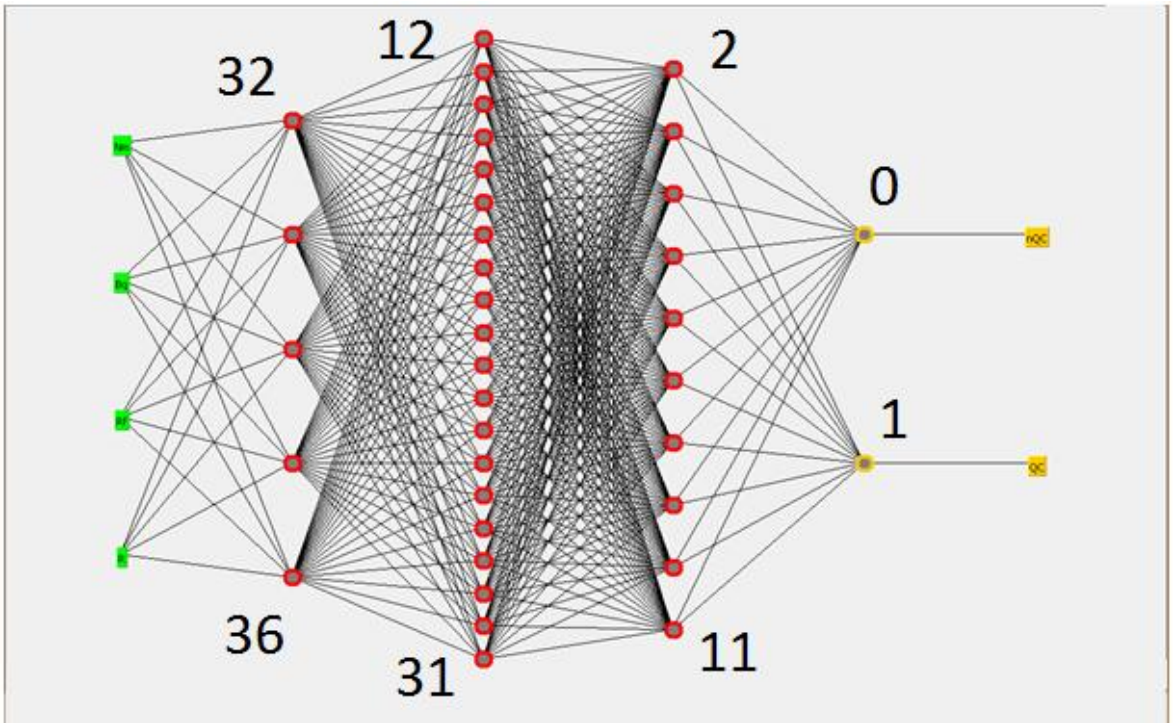


Figure 7.20 *Version 4 - ordinal number of outermost neurons is indicated*

Final step: Number of neurons in middle and last hidden layers increased.

Idea behind the creation of a precise version of network is to apply different modifications. With available training data a pattern between changes is investigated on the basis of accuracy fluctuation.

Table 14 - *Verified accuracy for all versions (values in %)*

Training set			Fallan	Klett			Tiller
			RCPTu	RCPTu			RCPTu
Test set			2	S1	S2	S1&S2	B1
Fallan	RCPTu	2		83,85V1 83,92V2 83,92V3 83,85V4	90,23V1 90,57V2 90,57V3 90,23V4	86,50V1 87,52V2 83,92V3 87,25V4	89,55V1 89,76V2 89,62V3 87,11V4
Klett	RCPTu	S1	70,45V1 70,81V2 57,17V3 86,26V4		96,15V1 95,99V2 95,93V3 96,21V4	97,20V1 98,57V2 97,42V3 98,46V4	67,40V1 68,17V2 69,21V3 71,41V4
		S2	77,88V1 65,14V2 51,55V3 91,36V4	95,00V1 95,00V2 95,00V3 95,00V4		98,95V1 98,95V2 95,00V3 99,15V4	65,88V1 65,77V2 65,67V3 71,14V4
Tiller	RCPTu	B1	81,54V1 73,92V2 79,86V3 74,41V4	56,75V1 57,62V2 57,50V3 66,34V4	60,72V1 61,77V2 61,77V3 60,90V4	54,46V1 66,17V2 57,50V3 65,80V4	

Table 15 - *Correctly classified instances of training sets (values in %)*

Training set			Fallan	Klett			Tiller
			RCPTu	RCPTu			RCPTu
Verification	Nodes		2	S1	S2	S1&S2	B1
Tenfold cross-validation	2,0,0	V1	94,71	97,31	99,37	97,90	97,03
	3,5,3	V2	92,70	98,08	99,16	98,47	97,34
	3,12,3	V3	93,96	97,69	99,16	97,61	96,59
	5,20,10	V4	95,66	95,00	99,21	98,47	97,27

Initial accuracy of a model on training set is more crucial then accuracy obtained on test set. A network, which was assembled imprecisely, will generate more distortion during the testing stage.
It is important to indicate a single combined training database - the highest accuracy was obtained by the test sets from the same investigation site. Furthermore, reliability is lower in case of using a single database: S1 versus S2 and vice-versa. This fact indicates a comparable quality of data from those particular soundings and borings.
Last and most complicated version gave most dependable results - a tremendously high discrepancy was detected in Fallan training set tested on data from Klett. Moreover, the highest accuracy in this research was also obtained by version 4 - with the use of combined Klett database. It is even comparable to initial accuracy of models trained by Klett S2. This might be caused by the implemented changes of a model - increased number of nodes near the output (two last hidden layers). It is most efficient enhancement applied to the network.
Unfortunately, no other relations between individual versions of network can be formulated.

For the purpose of simulation most succeeding models with highest accuracy are choosen:

- for Fallan CPTu and RCPTu - models V3 trained on Klett S2 and Tiller datasets
- for Klett CPTu - model V4 trained on combined Klett database

Predicted profiles are presented on the next pages.

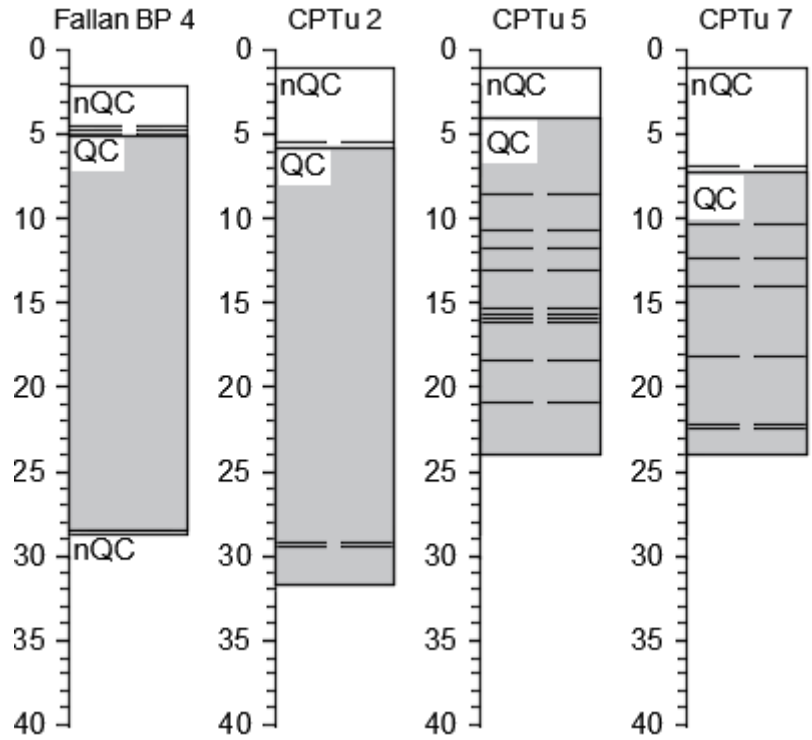


Figure 7.21 *Predictions based on Klett S1 model - dashed lines indicate data distortion*

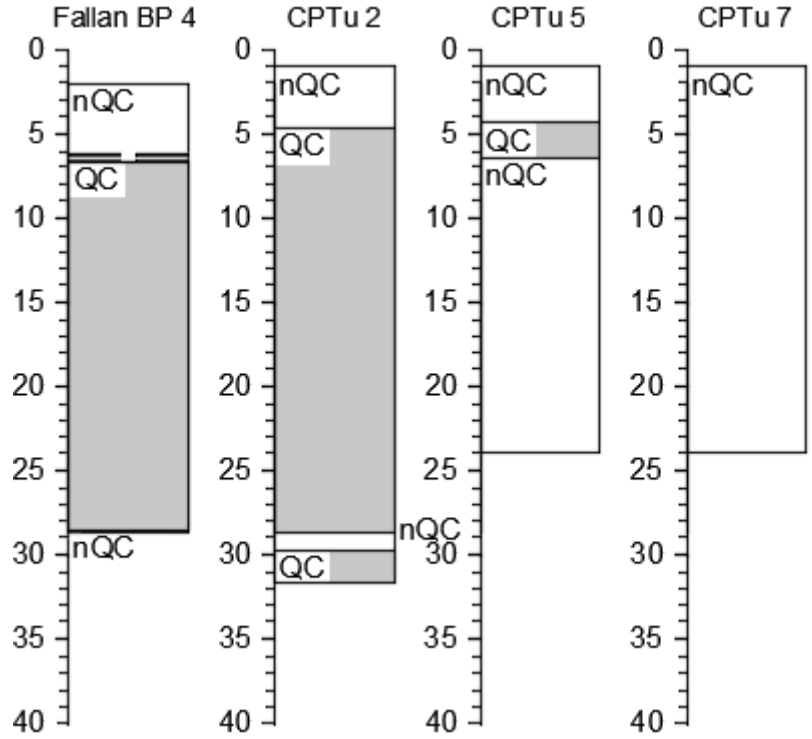


Figure 7.22 *Predictions based on Tiller B1 model - dashed lines indicate data distortion*

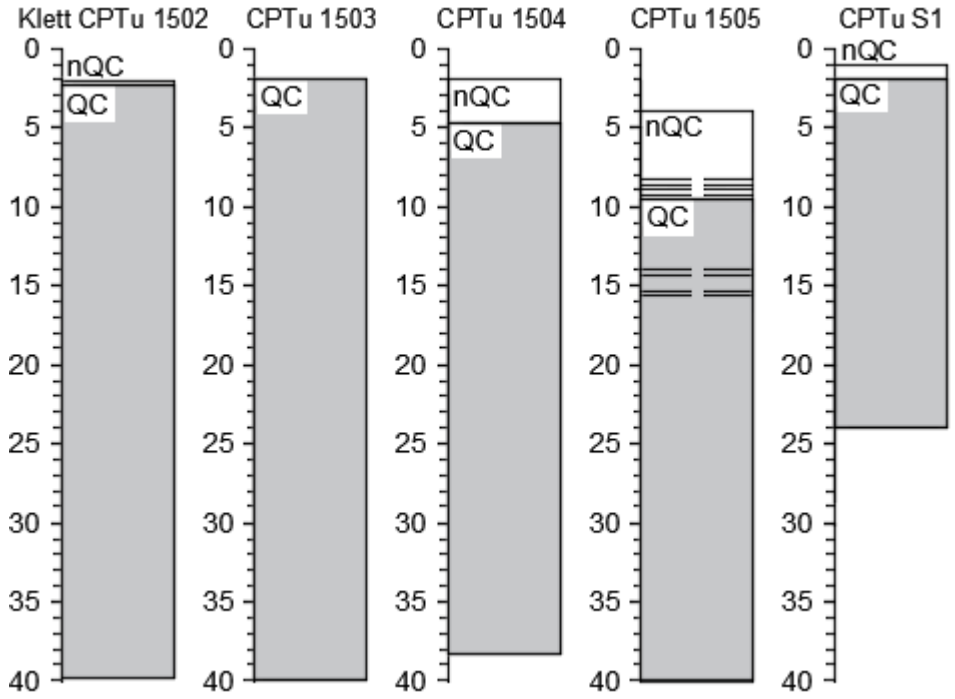


Figure 7.23 *Predictions based on Klett combined model - dashed lines indicate data distortion*

Predicted profiles from Fallan represent following dependency:

- results for BP2 and CPTu 2 are very similar for two models, minor differences are in the transition layers
- data distortion is large in profiles calculated by Klett-trained model; despite of that fact, these results are much more probable than the one modeled after Tillers training set

Klett results are considered as reliable; they are comparable to the laboratory data. Data distortions strongly suggest presence of a transition layer. Obtained profiles have similar characteristics, which is expected for a set of soundings from a single investigation site.
Profiles from Fallan are more diverse - the Tiller model gives inconsequential results. For that reason the Klett S1 model is treated as acceptable despite of multiple data distortions in the profile.

8. Discussion of results

8.1. CPTu and R-CPTu

After comparing post-processed results from soundings with laboratory data the following conclusions are made:

1. Classification chart - Schneider (2008) has the best performance and highest accuracy. In addition, identification of transition layers is a valuable feature, which proved useful in the profile assessment (see profile Fallan 2 CPTu on Fig. 5.2.3). At the same time, Robertson (1990) diagram Q_t-B_q is more responsive to interlayers. Both mentioned solutions utilize excess pore water pressure Δu_{2c} parameter, which apparently benefits the overall performance of the charts. Sensitive clays are more responsive to the changes of pore water pressure than to sleeve friction.
   Figures 8.2 - 8.9 represent CPTu results, which are classified by the laboratory data. It is clearly visible, that used classification charts do not distribute results according to their class. Because of that, this method gives only estimated soil profile.

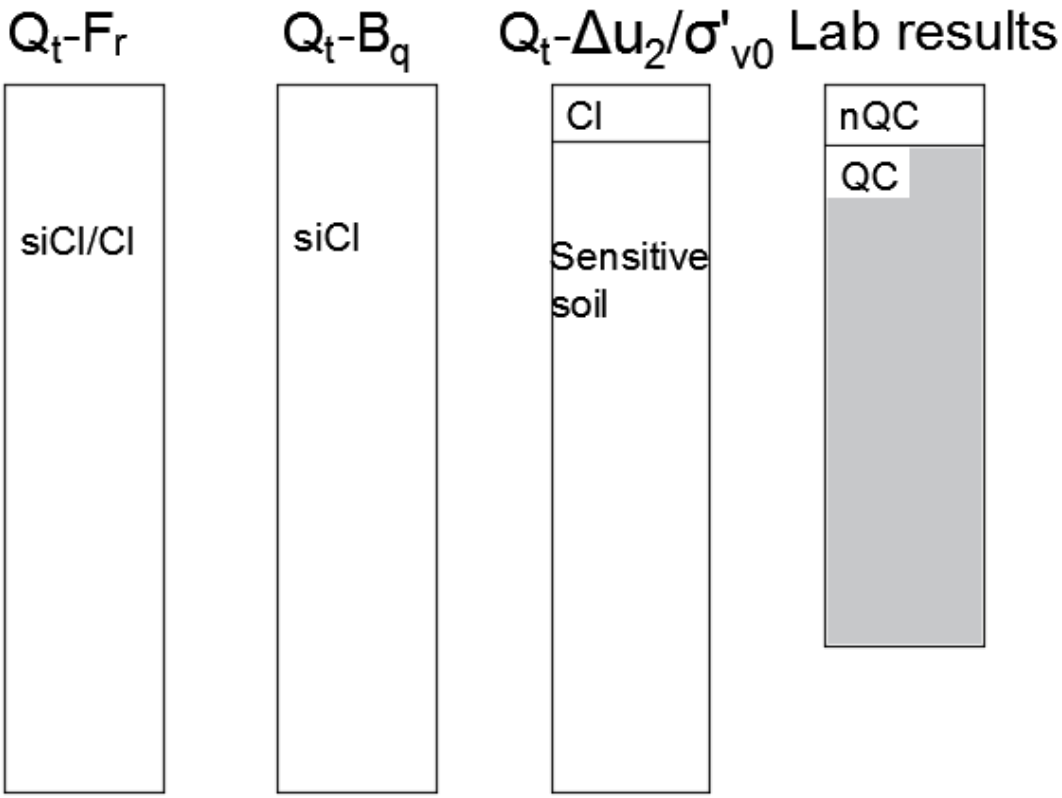


Figure 8.1 *Results from Fallan BP2 RCPTu and sampling*

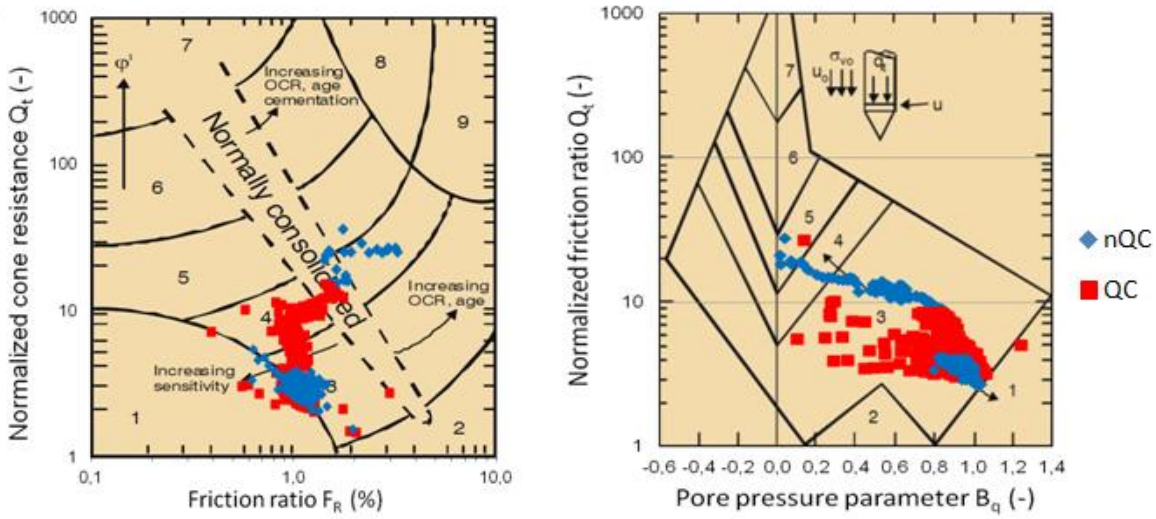


Figure 8.2 *Fallan BP 2 soil classes from laboratory data plotted on Robertson charts (1990)*

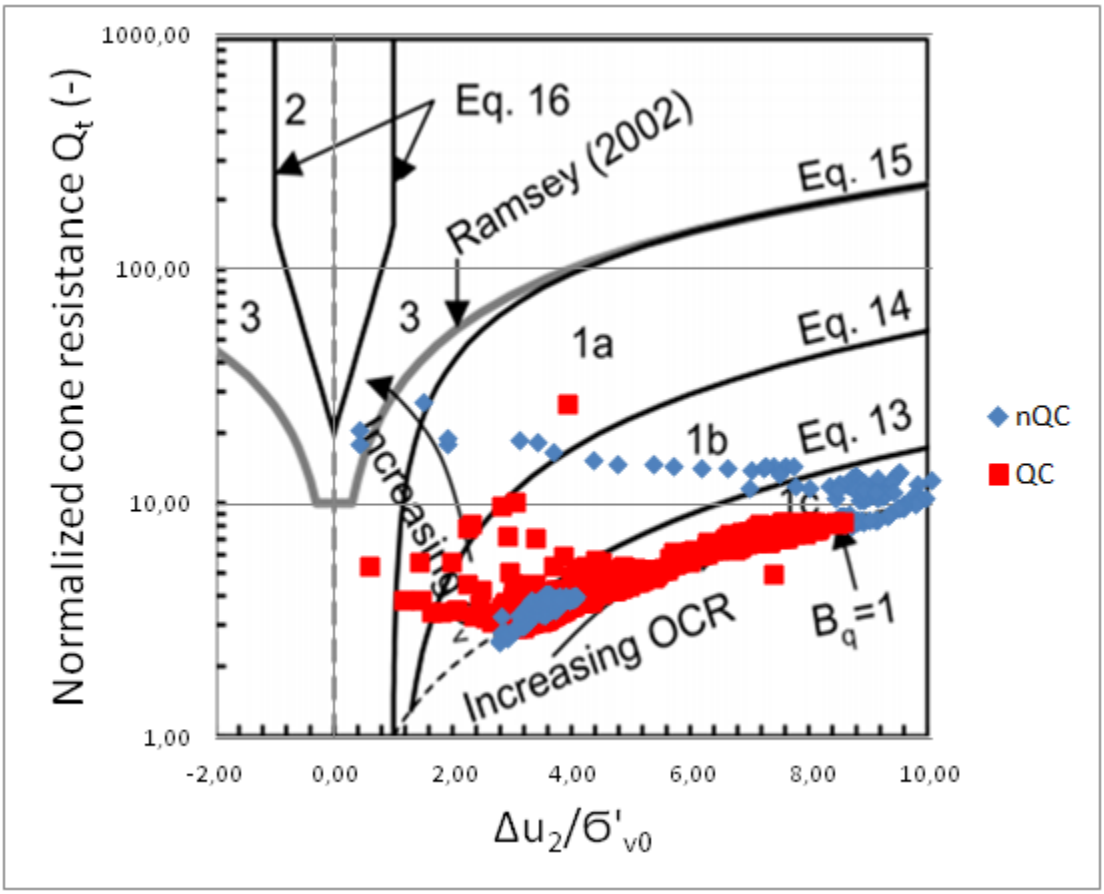


Figure 8.3 *Fallan BP 2 soil classes from laboratory data plotted on Schneider charts (2008)*

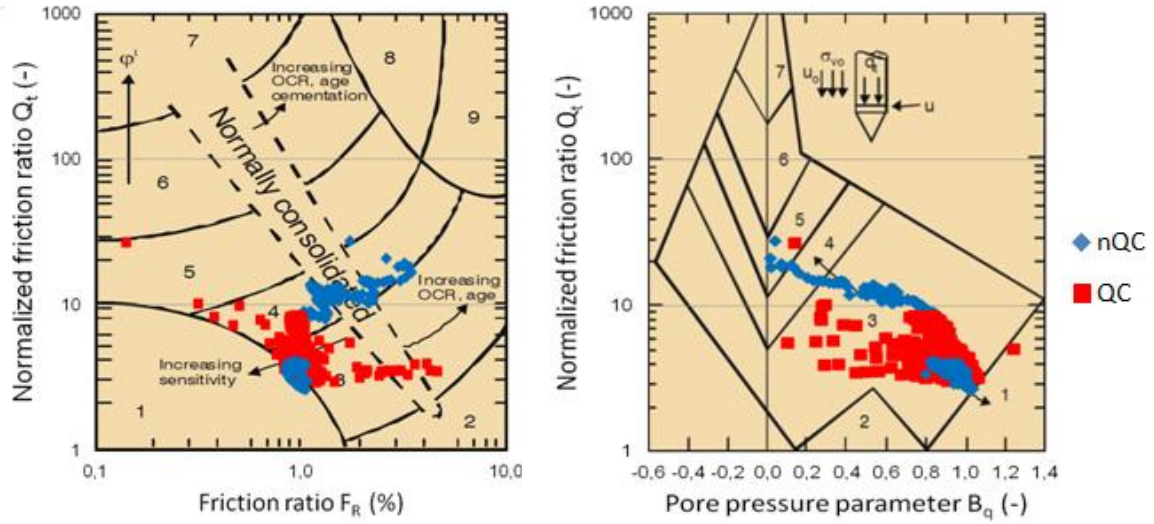


Figure 8.4 *Klett S1 soil classes from laboratory data plotted on Robertson charts (1990)*

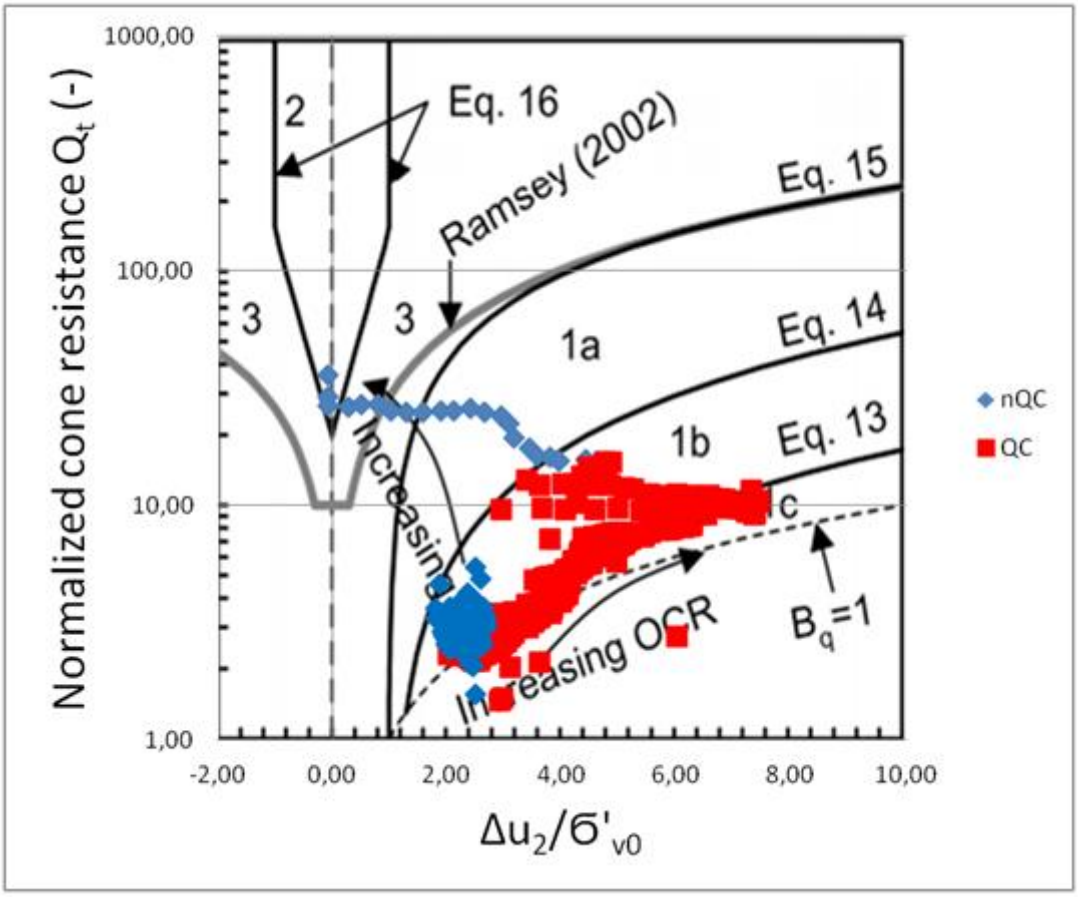


Figure 8.5 *Klett S1 soil classes from laboratory data plotted on Schneider charts (2008)*

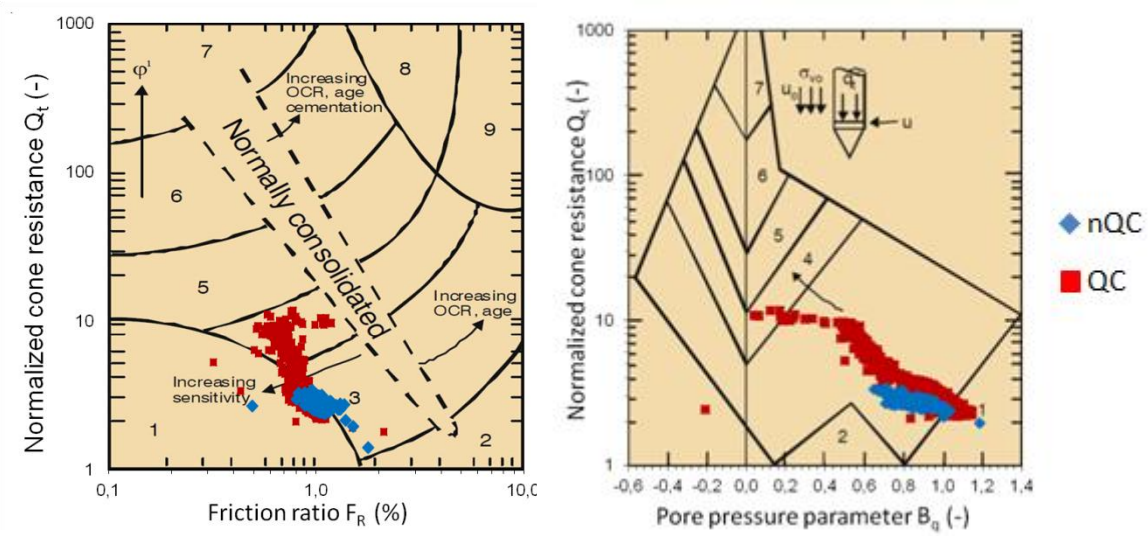


Figure 8.6 *Klett S2 soil classes from laboratory data plotted on Robertson charts (1990)*

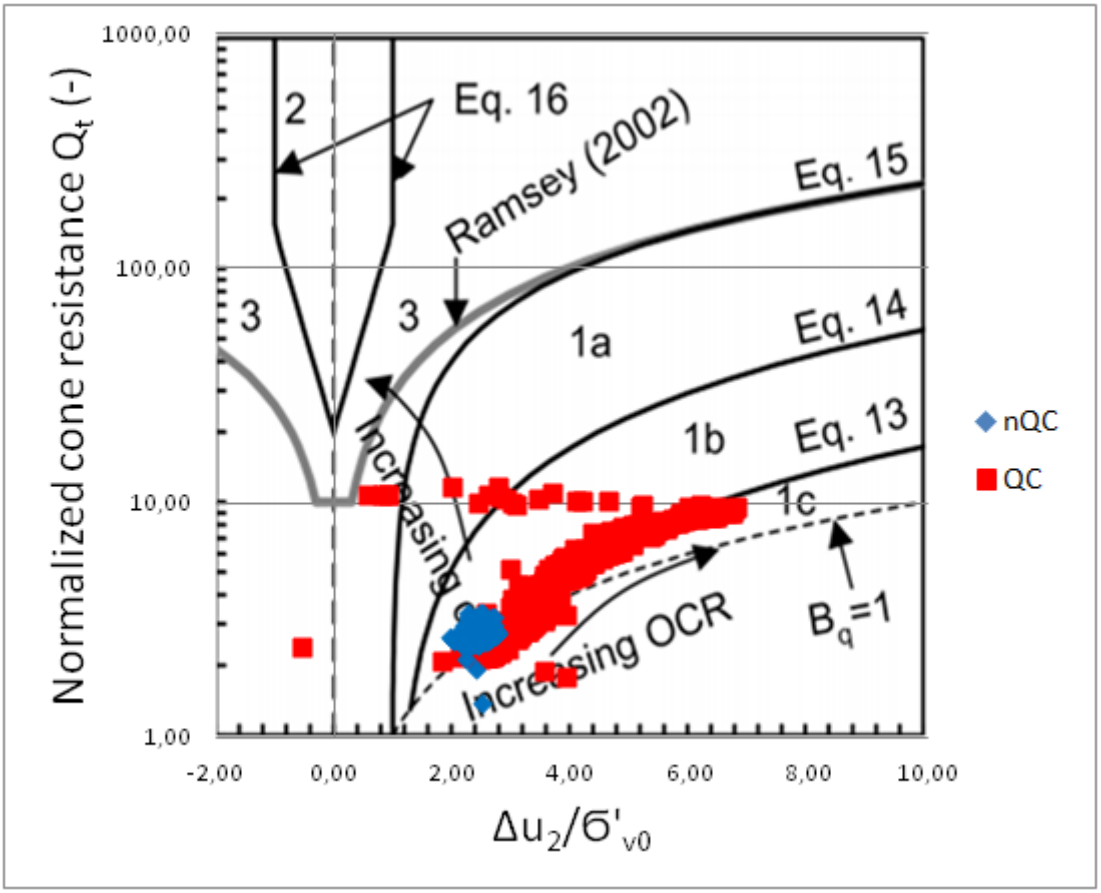


Figure 8.7 *Klett S2 soil classes from laboratory data plotted on Schneider charts (2008)*

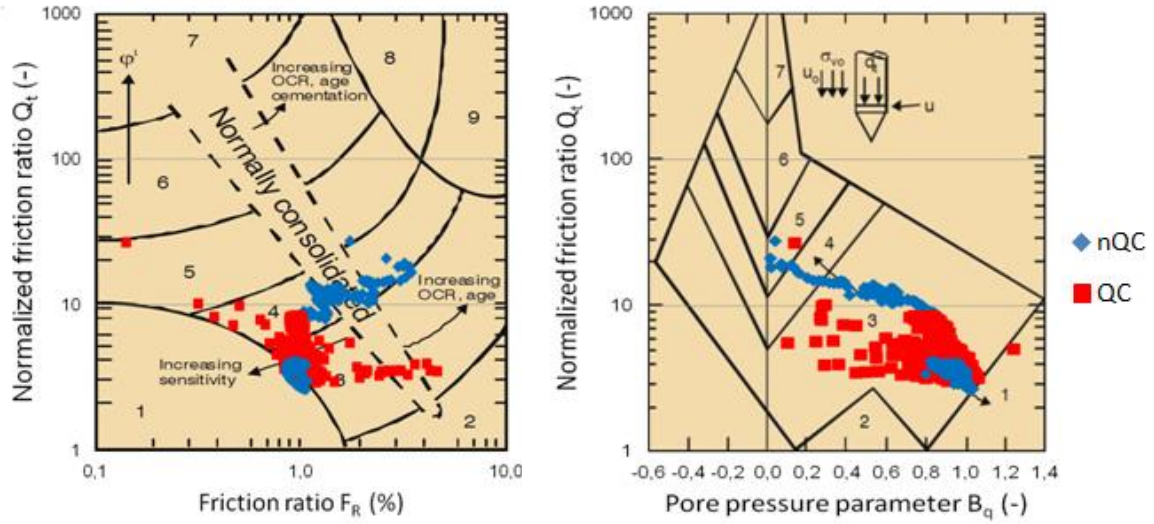


Figure 8.8 *Tiller B1 soil classes from laboratory data plotted on Robertson charts (1990)*

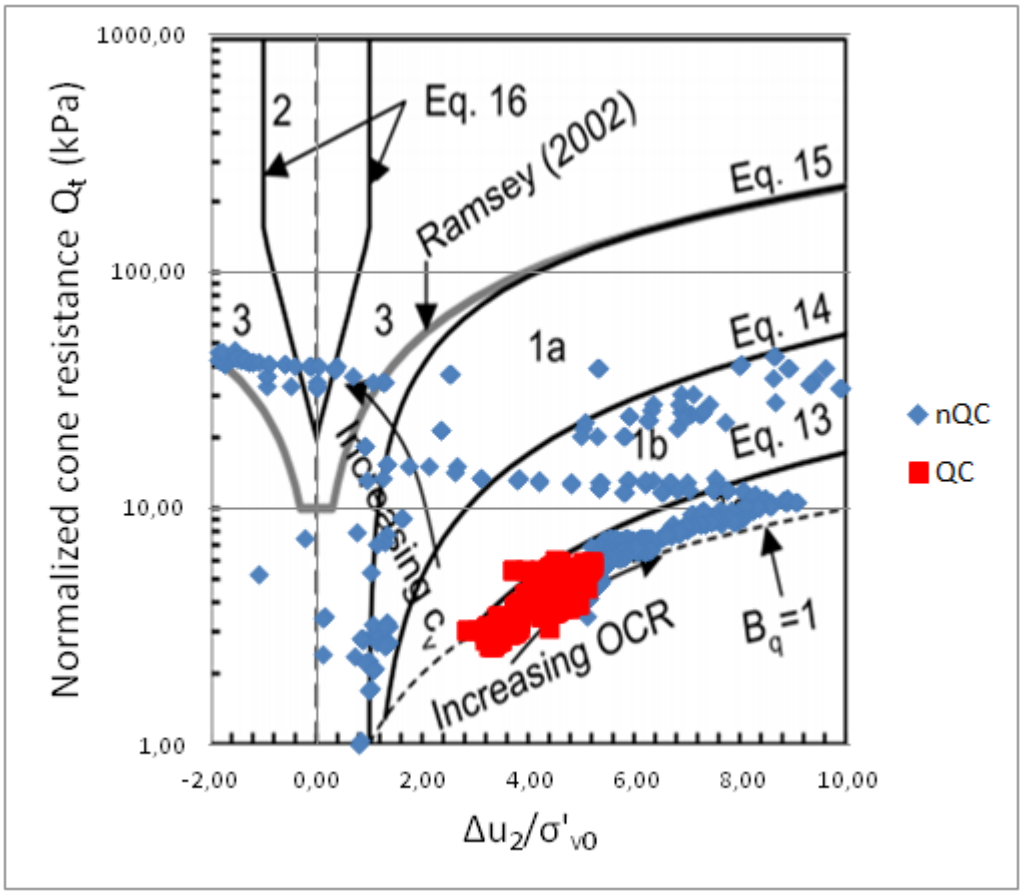


Figure 8.9 *Tiller B1 soil classes from laboratory data plotted on Schneider charts (2008)*

2. Undrained shear strength - most effective approaches employ total cone resistance N_{kt} and excess pore pressure N_{du} factors (Eq. 13,14,17,18). Method with N_{du} tends to overestimate s_{uA} as well as Nkt, which usually gives lesser results. Solution, which implements effective cone resistance N_{ke} is vulnerable to changes of total cone resistance and pore water pressure. Unfortunately, most of the results tend to give non-linear results, which are higher than the laboratory data, so this parameter should be analyzed with caution.
3. Sensitivity graphs are not reliable - despite of detected sensitive soils by soundings and borings charts do not indicate no diametrical changes. Only in few cases, like Klett CPTu 1503 (Fig. 5.3.5), there is detected diversity in the profile. Even though, those results cannot be compared to the remaining data. Parameter of sensitivity S_t marks sensitive clay if friction ratio R_f is very minimal. This fact excludes any practical application. According to Norwegian standards:

$$S_t > 30 \; for \; sensitive \; soil \qquad \textbf{(27)}$$

4. Available laboratory data and R-CPTu measurements reveal following typical values of resistivity of quick clay:
   - Fallan (15-40 Ωm)
   - Klett (20-50 Ωm)
   - Tiller (25-45 Ωm)

   Solberg classification is more accurate for the surveyed investigation sites. However, the analyzed graphs cannot be used to classify quick clay itself - they should be treated as a tracer of a potential deposit.

Table 16 - *Soil classification by Solberg (2008)*

Soil	Resistivity	Note
Salt/intact marine clay	1-10 Ωm	
Leached, possible quick clay	10-80 Ωm	
Dry crust clay, slide deposits	> 80 Ωm	
Silt, mettet	50 - 200	
Sand, mettet	200 - 1000	

Table 17 - *Soil classification by Berger (1980)*

Soil	Resistivity	Note
Clay, salt	1 - 20	
Clay, leached	20 - 90	
Clay, dry crust	70 - 300	
Silt, wet	50 - 200	
Sand, saturated	200 - 1000	

8.2. Data mining

After selecting most accurate models for each data mining technique the produced profiles are compared with each other. It is essential for the analysis to remember an individual origin of each model. To visualize expected outcome of the simulated profiles the borehole data is included and compared with the predicted results.

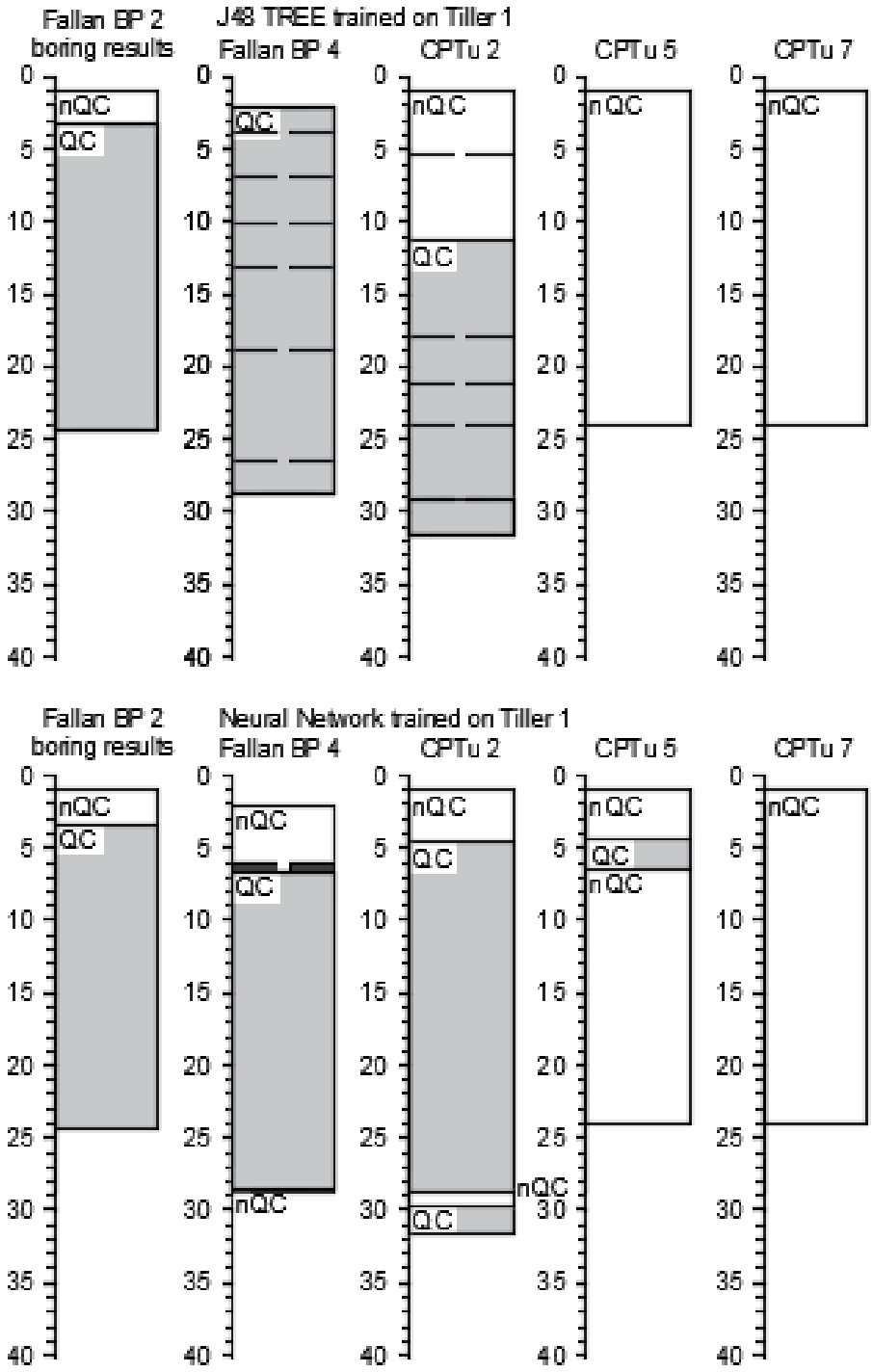


Figure 8.10 *Results for Fallan investigation site from tree classifier and neural network approach*

Comment:

Despite the fact, that presented profiles of CPTu 5 and 7 soundings are unacceptable, the resemblance between the results is striking. Similarity is definitely caused by the choice of the dataset for training - which in this case is Tiller S1.

Lack of results from clustering method for Fallan investigation site is justified by full incongruity of the produced profiles.

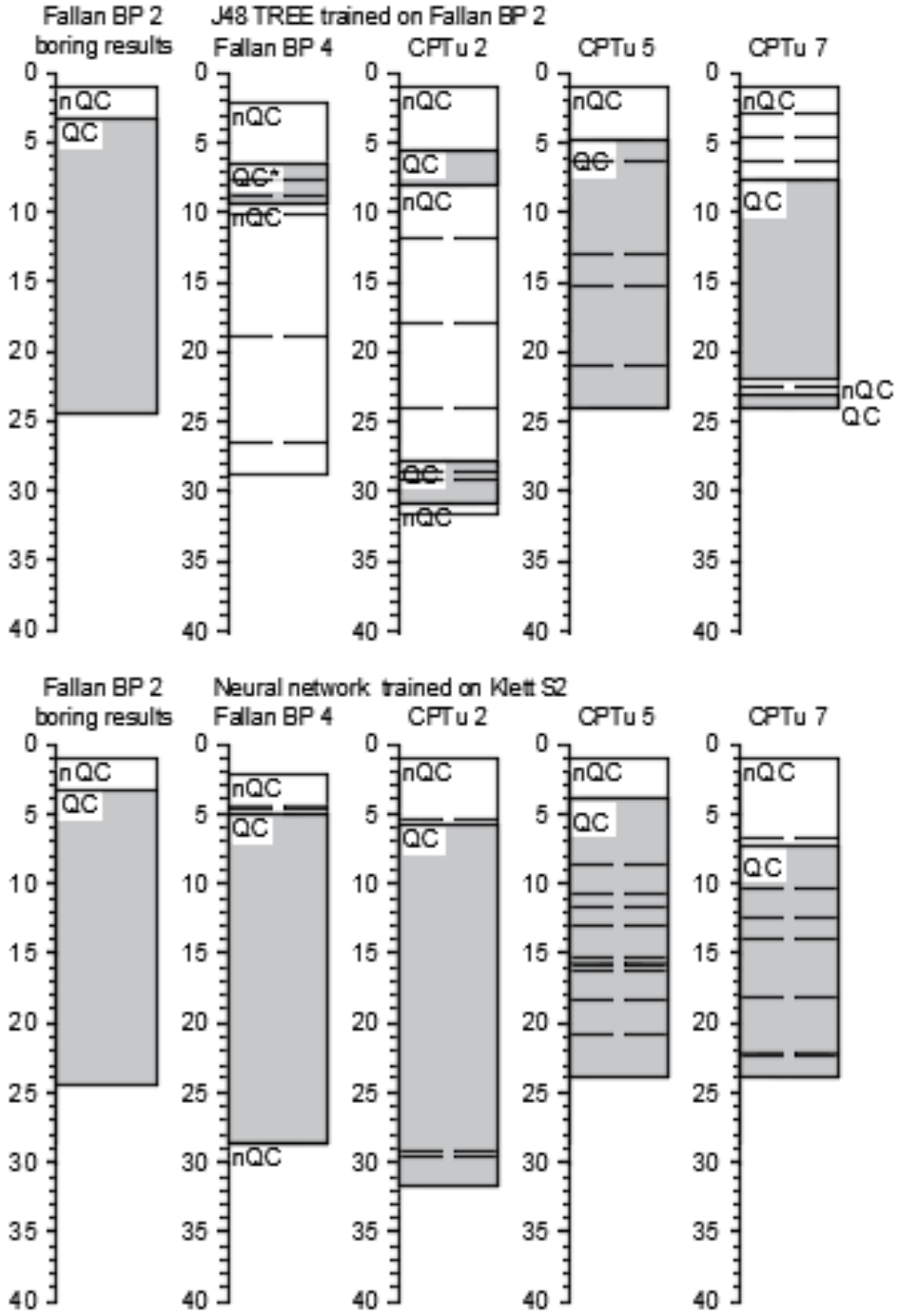


Figure 8.11 *Results for Fallan investigation site from tree classifier and neural network approach*

Comment:

With different training dataset the neural networks results are comparable to the laboratory data. For Fallan investigation site the neural network approach gave more accurate profiles than a tree classifier. It was expected to achieve lower reliability by the more primitive data mining algorithm. A model trained on Klett S2 dataset detects data distortion in CPTu soundings 2,5,7 below 15,0 m. Origin of the interference can be caused by obstacles, which resulted in the extreme readings from piezocone. Another considered reason is an oversensitivity of a model - a defect of mismatched training set with the tested database.

The stated dilemma can be solved with comparison of acquired results from the models with the CPTu readings.

8.3. Soundings and data mining techniques

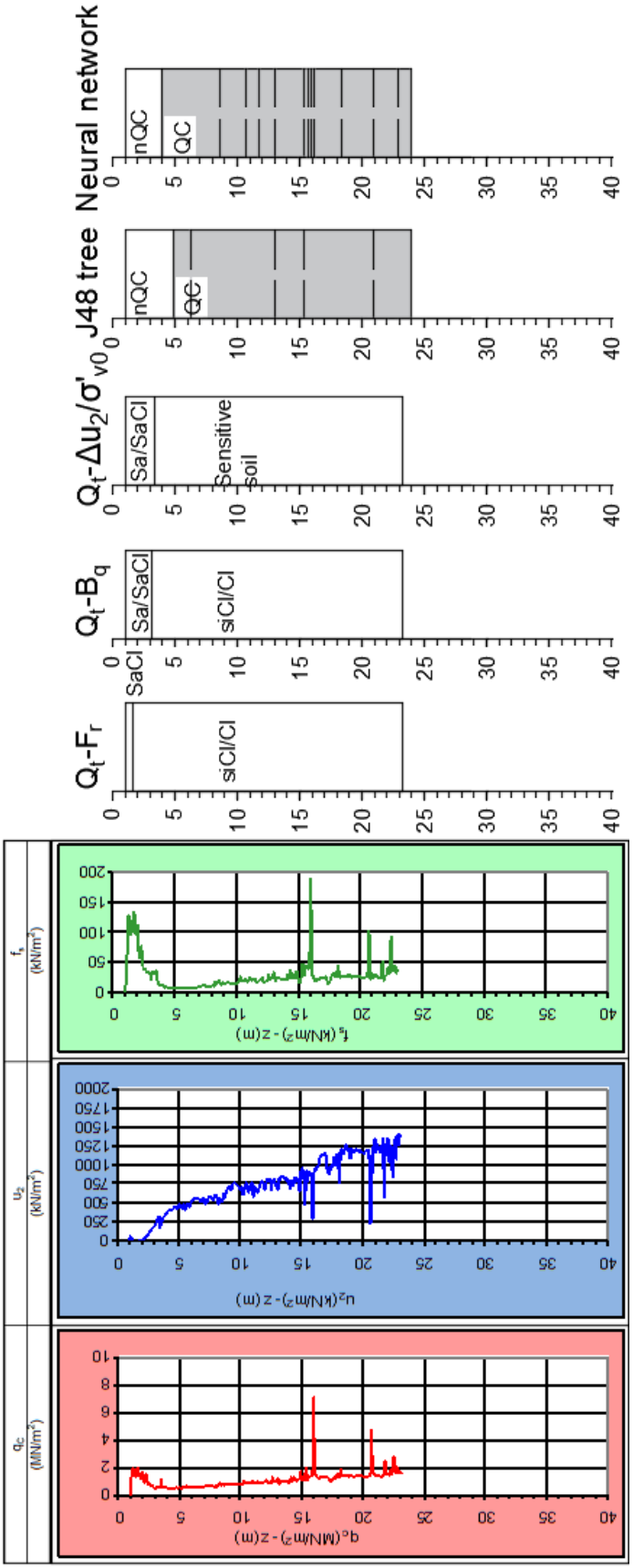


Figure 8.12 *Results for exemplary sounding: Fallan CPTu 5 with profiles from different methods*

On the basis of analyzed soundings with corresponding profiles, it appears that the data distortion is caused by interferences generated by the piezocone - locations of peaks in the graphs fit almost ideally. Presented readings on Figure 8.12 represent homogenous soil layers with randomly distributed gravel particles. It is safe to conclude, that a neural network approach is more responsive to the changes of CPTu readings than tree classifier. Therefore, a hypothesis of an oversensitive model is disproved.
Remaining soundings with corresponding profiles are listed in the appendix.

9. Conclusion

9.1. Sounding post-processing

A study of the most commonly used, classical interpretation charts reveals that the most reliable chart for identifying sensitive soils is Schneider (2008). A profile formulated by the classification chart can be compared and verified by undrained strength and resistivity readings. For the deposits identified by laboratory tests as quick clay. Calculated values for Klett and Fallan can be treated as legitimate. However, given methods provide disputable results, which cannot differ quick clay deposits with a high accuracy.
The one, most universal and dependable parameter, which classifies quick clay in all possible cases, is remolded undrained shear strength s_{ur}. Unfortunately, it can be obtained only via sampling and laboratory testing.

9.2. Machine learning

As excepted, the best performance is achieved by Neural Network models. Version 4, which has most advanced and complex structure, proved to be most reliable by acquiring highest average accuracy in all investigation sites. However, it is suggested to use a model, which is trained on the dataset from the same investigation site as the input soundings. Figure 8.2 from the previous chapter represents mismatched dataset. Determination of a sample training set, collecting high quality samples with laboratory results and choosing appropriate algorithm are key factors, which influence precision of a model. Figure 9 presents compatibility of trained database with studied data.

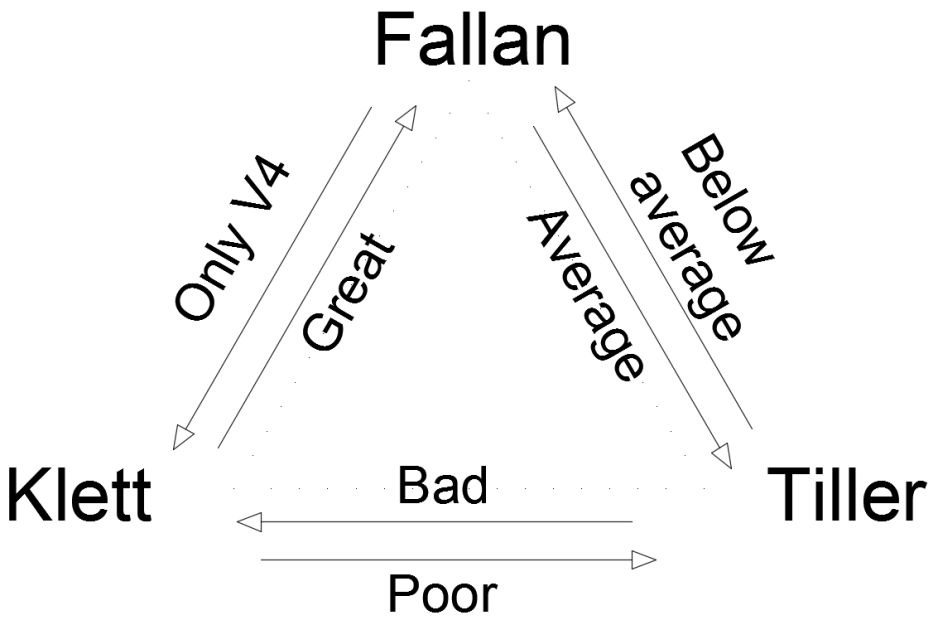


Figure 9 *Neural network application chart*

An ultimate verification process include additional borings in the points, where soundings were carried out. It is most effective way for checking simulated profiles and consequently validate the reliability of models.

Furthermore, author of this thesis proposes for future work designing and testing alternative classifiers, algorithms or other data mining methods. Conducted studies are experimental and enlarging available databases along with number of performed tests, implementing new solutions and data verification processes would unquestionably increase the quality of the output.

9. References

Literature:

1. A. Gylland, M. Long, A. Emdal, and R. Sandven. 2013. *Characterization and engineering properties of Tiller clay. Engineering Geology* 164:86-100

2. A. T. Iliesi, A. L. Tofan, D. L. Presti. 2012. *Use of cone penetration tests and cone penetration tests with porewater pressure measurement for difficult soils profiling,* Technical University of Iasi, Italy

3. B. H. Fellenius, A. Eslami. 2000. *Soil profile interpreted from CPTu data,* Ottawa

4. C. R. Daniel, H. L. Giacheti, J. A. Howie, R. G. (D.) Campanella. 1999. *Resistivity piezocone (RCPTU) data interpretation and potential applications,* The University of British Columbia

5. C. R. Daniel, R.G. (D.) Campanella, J. A. Howie. 2003. *Specific depth cone resistivity measurements to determine soil engineering properties,* The University of British Columbia

6. GEOTECH. 2014. *Users manual CPT GEOTECH NOVA*

7. Geotechnical research group. *Practical applications of the cone penetration test. A manual on interpretation of seismic piezocone test data for geotechnical design,* The University of British Columbia

8. H. S. Yu, J. K. Mitchell. 1998 *Analysis of cone resistance: review of methods,* ASCE

9. H. S. Yu. 2000. *Cavity Expansion Methods in Geomechanics*

10. H. S. Yu. 2006. *The first James K. Mitchell Lecture, In situ soil testing: from mechanics to interpretation,* The University of Nottingham, UK

11. J. A. Schneider, M. F. Randolph, P. W. Mayne, N. R. Ramsey. 2008. *Analysis of factors influencing soil classification using normalized piezocone tip resistance and pore pressure parameters,* ASCE

12. J. A. Schneider. 2009. *Seperating influences of yield stress ratio (YSR) and partial drainage on piezocone response,* The University of Western Australia

13.J. W. Chen, C. H. Juang. 1996. *Determination of drained friction angle of sands from CPT,* ASCE

14.K. Karlsrud, T. Lunne, D.A. Kort, S. Strandvik. 2005. *CPTU Correlations for Clays,* NGI

15.K. K. Kim, M. Prezzi, R. Saldago. 2006. *Interpretation of cone penetration test in cohesive soils*, Purdue University West Lafayette, Indiana

16. K. Kumar, G. S. M. Thakur. 2012. *Extracting explanation from artificial neural networks,* Jalandhar (Punjab), India

17.K. Rankka, Y. A. Sköld, C. Hultén, R. Larsson, V. Leroux, T. Dahlin, 2004. *Quick clay in Sweden, Report 65,* SGI, Linköping

18.M. Long, S. Donohue, J. S. L'Heureux, I. L. Solberg, J. S. Rønning, R. Limacher, P. O'Connor, G. Sauvin, M. Rømoen, I. Lecomte. 2012. *Relationship between electrical resistivity and basic geotechnical parameters for marine clay,* NRC Research Press

19.M. Rømoen, A. A. Pfaffhuber, K. Karlsrud. *Resistivity on marine sediments retrieved from RCPTU soundings: a Norwegian case study,* NGI, Oslo

20.P. W. Mayne, B. R. Christopher, J. DeJong. 2001. *Manual on subsurface investigations,* NHI, Washington DC

21.R. Sandven. *Revidert NGF-veiledning for CPTU*

22.S. J. Hong, M. J. Lee, J. J. Kim & W. J. Lee. *Evaluation of undrained shear strength of Busan clay using CPT*

23.T. Lunne, P.K. Robertson, J.J.M. Powell. 1997.z *Cone penetration testing in geotechnical practise*

24.Y. Kumar, G. Sahoo. 2012. *Analysis of Bayes, neural network and tree classifier or classification technique in data mining using WEKA*

25.http://www.cpt10.com/ - 2nd International Symposium on Cone Penetration Testing official website

26.http://www.cs.waikato.ac.nz/ - official website of The University of Waikato, New Zealand

27. http://www.geotech.eu/ - CPT equipment and manual

28. http://www.gouda-geo.com/ - CPT equipment and guidelines

29. http://www.ibm.com/ - data mining manual

30. http://netbeans.dzone.com/ - articles about neural networks

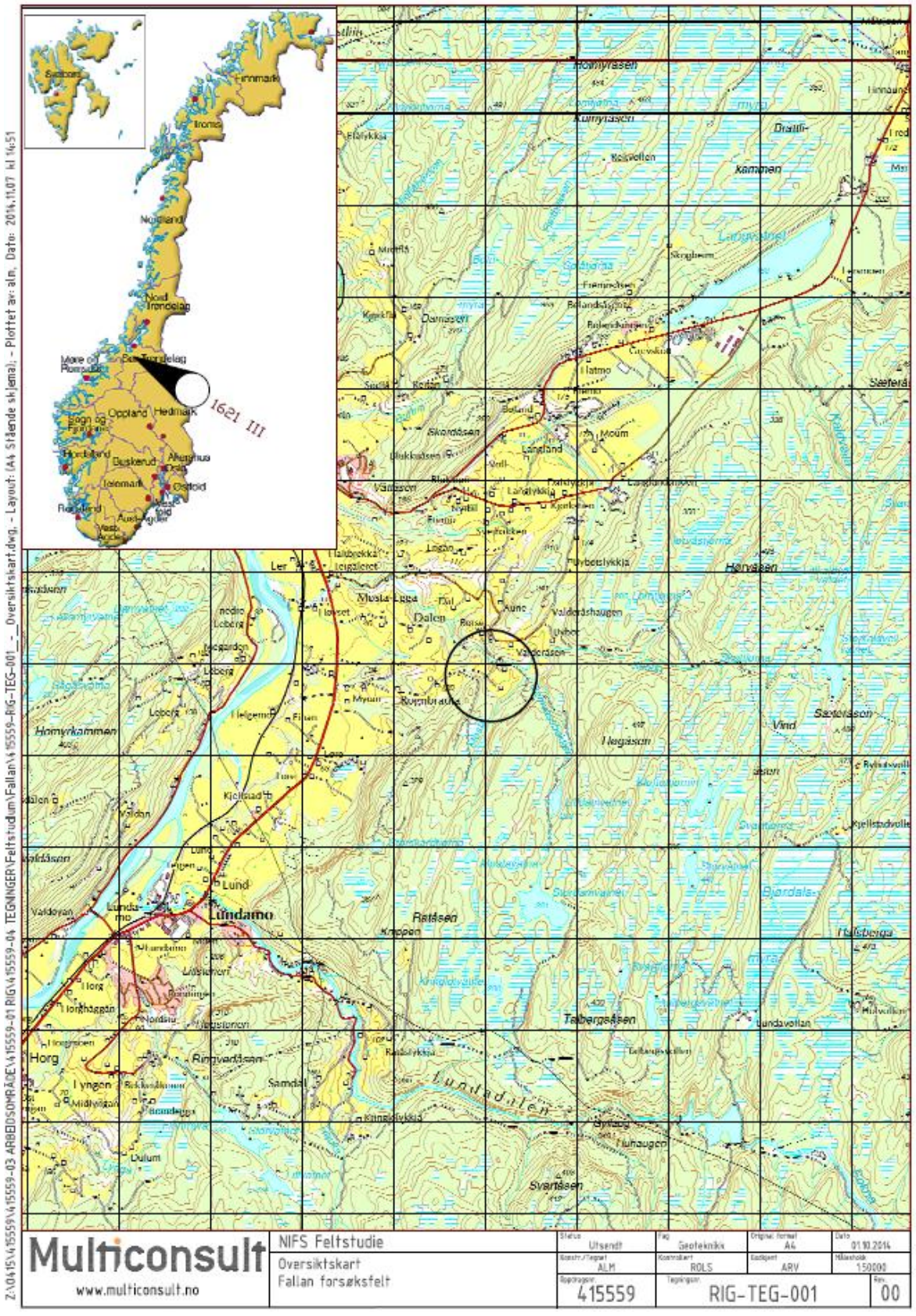
31. http://www.opensounding.sourceforge.net/ - additional data mining software

32. http://www.wikipedia.org/ - definitions and theories

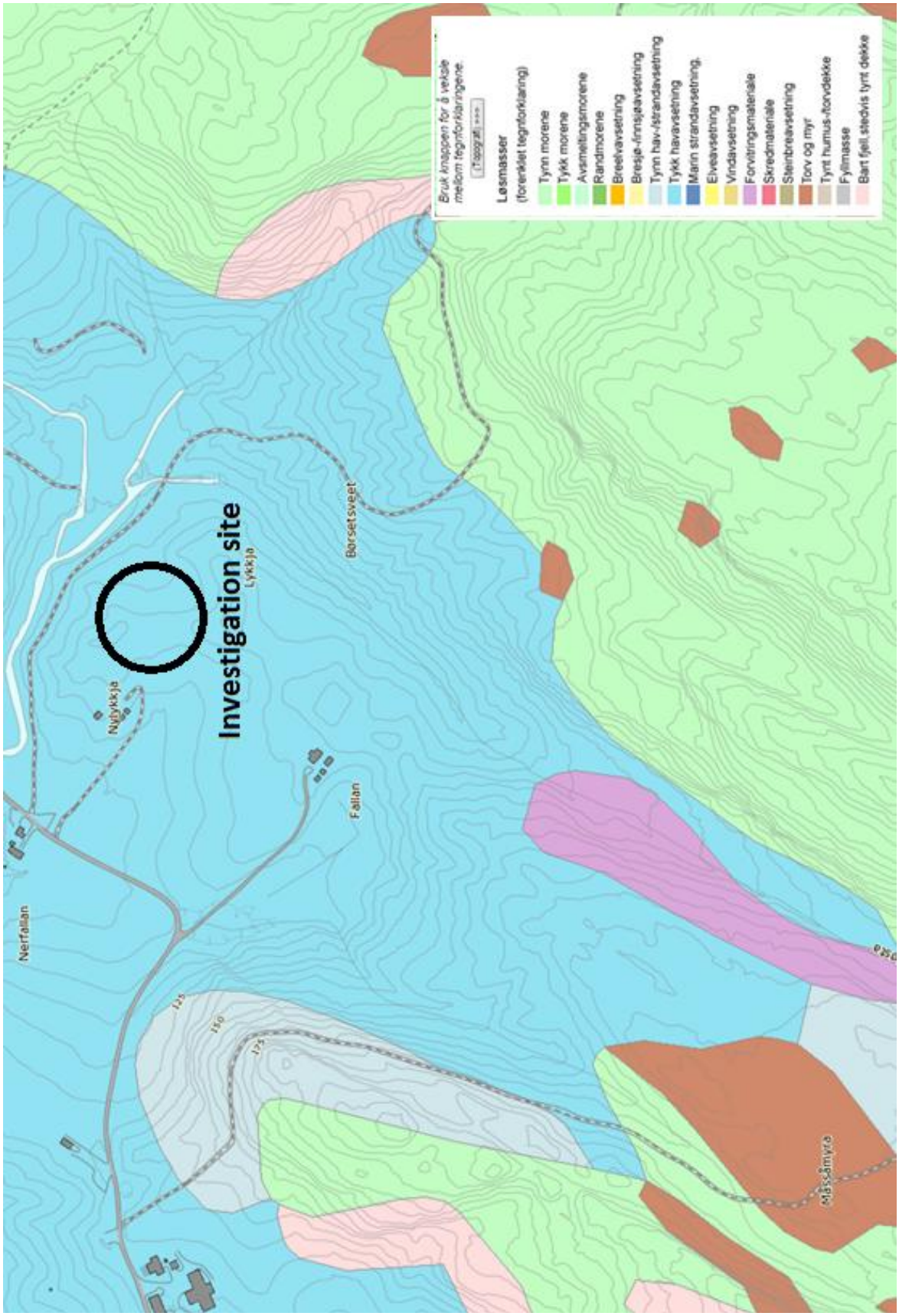
33. https://www.youtube.com/ - Prof. Ian Witten lectures on the "WekaMOOC" account

Appendices

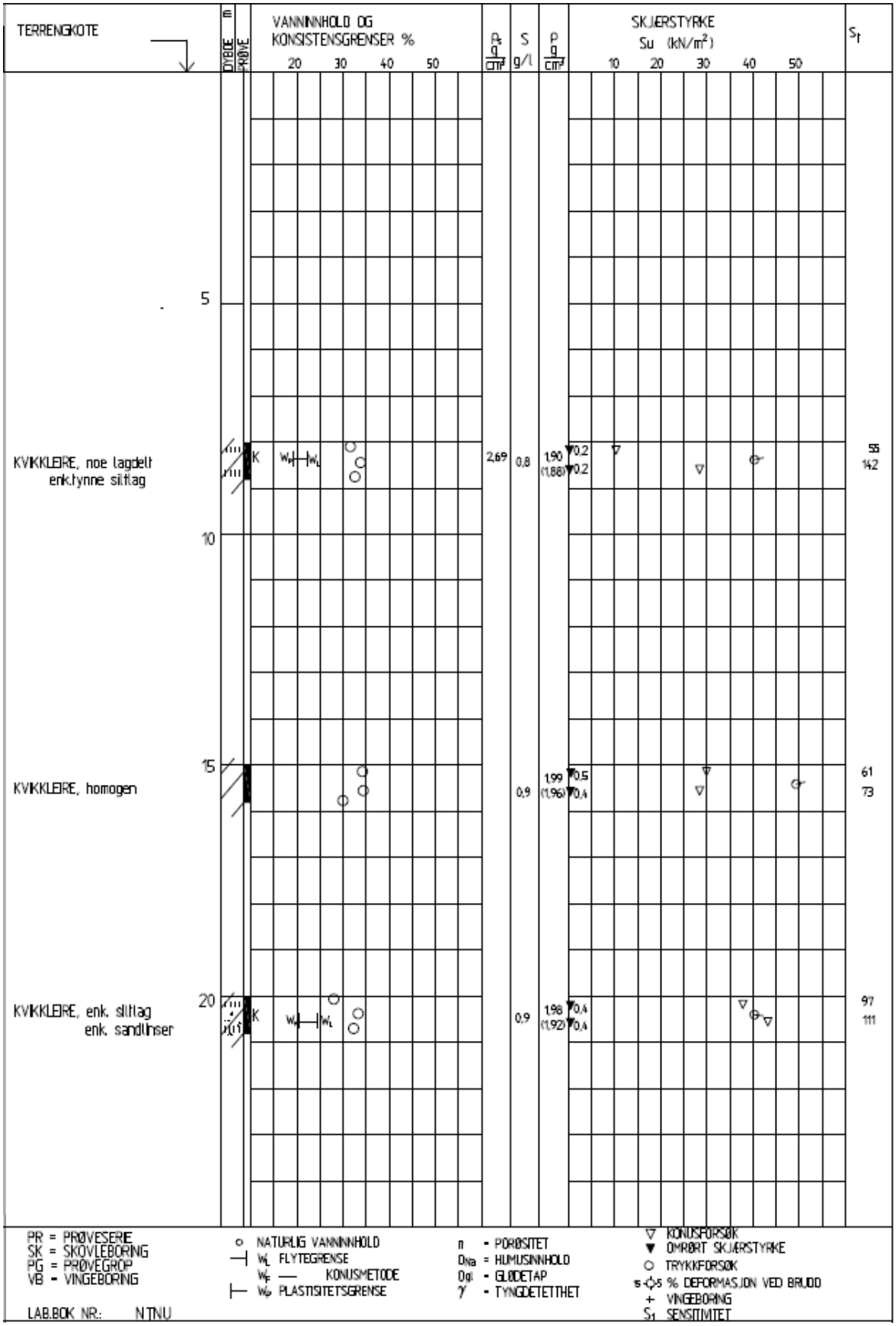
A. Overview map of Fallan by Multiconsult



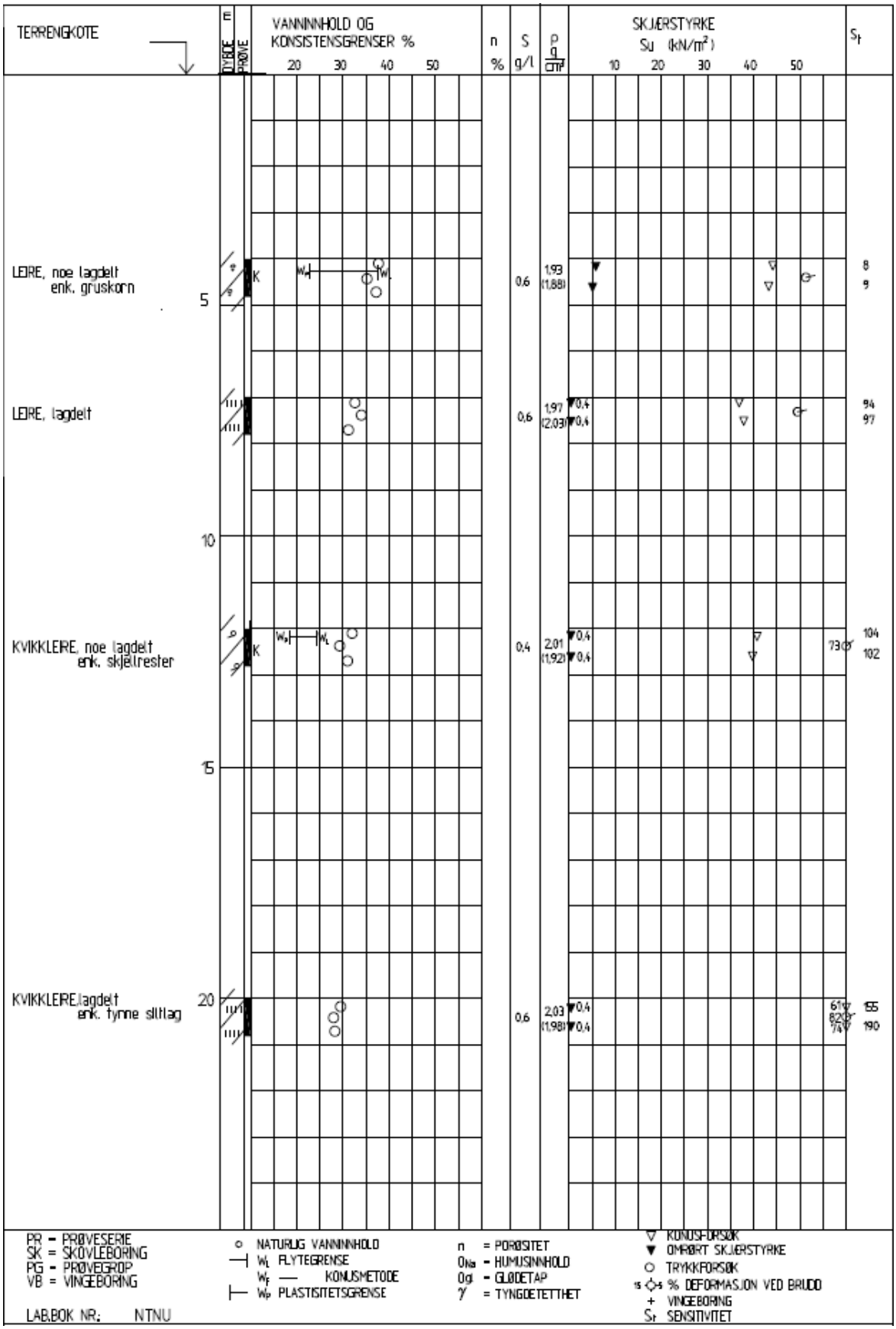
B. NGU Geological map of Fallan



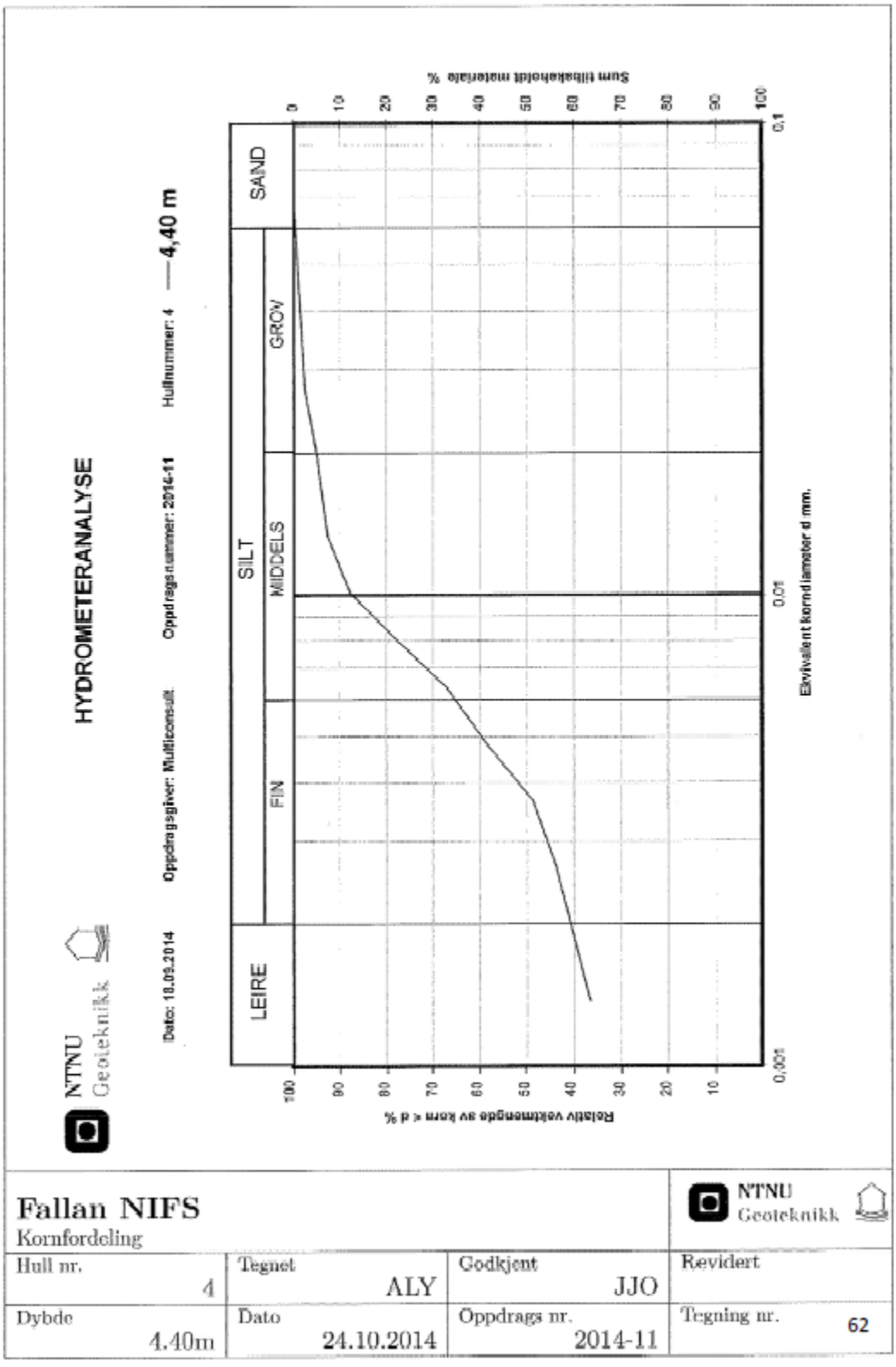
C. Laboratory results from Fallan BP 2 by Multiconsult



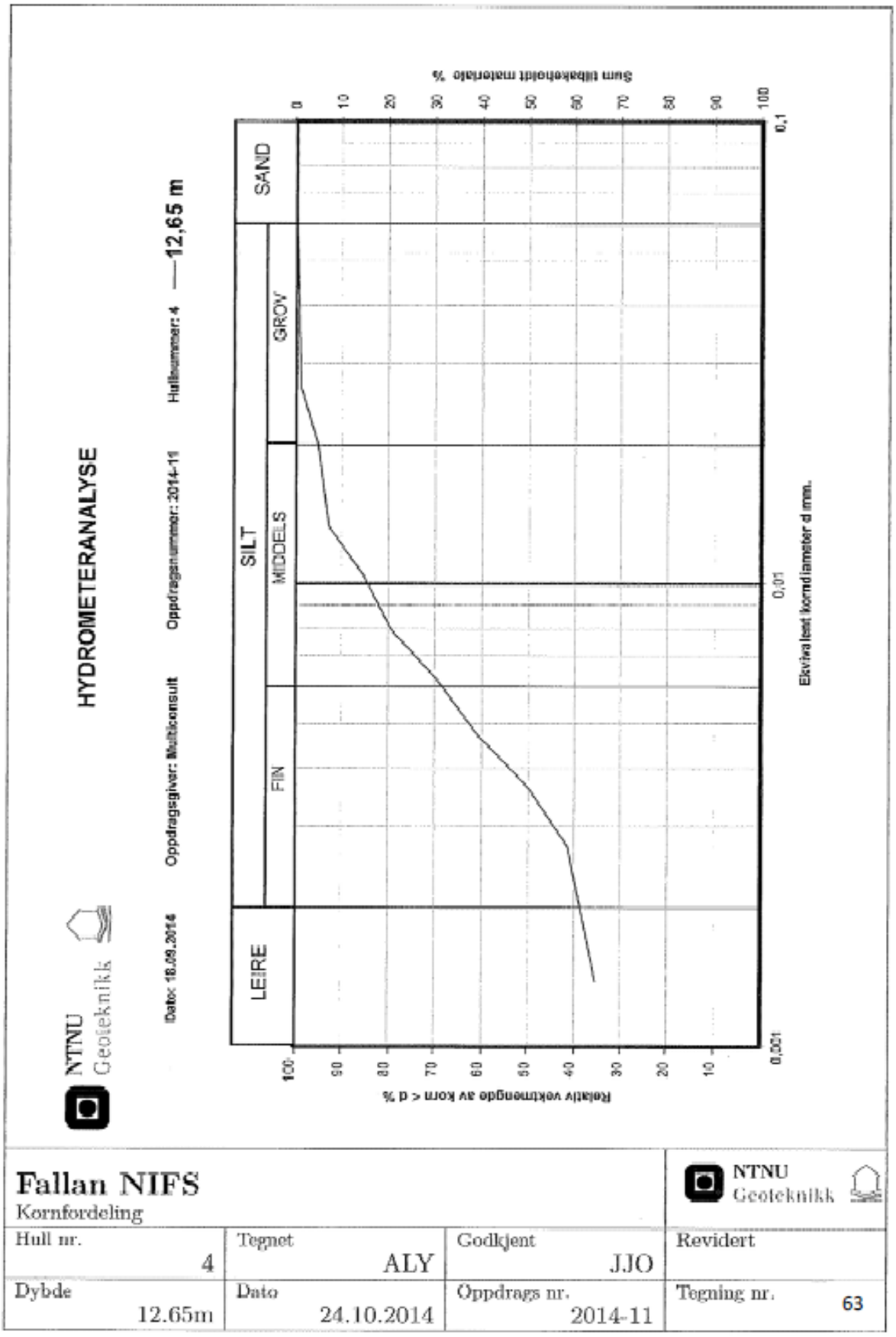
D. Laboratory results from Fallan BP 4 by Multiconsult



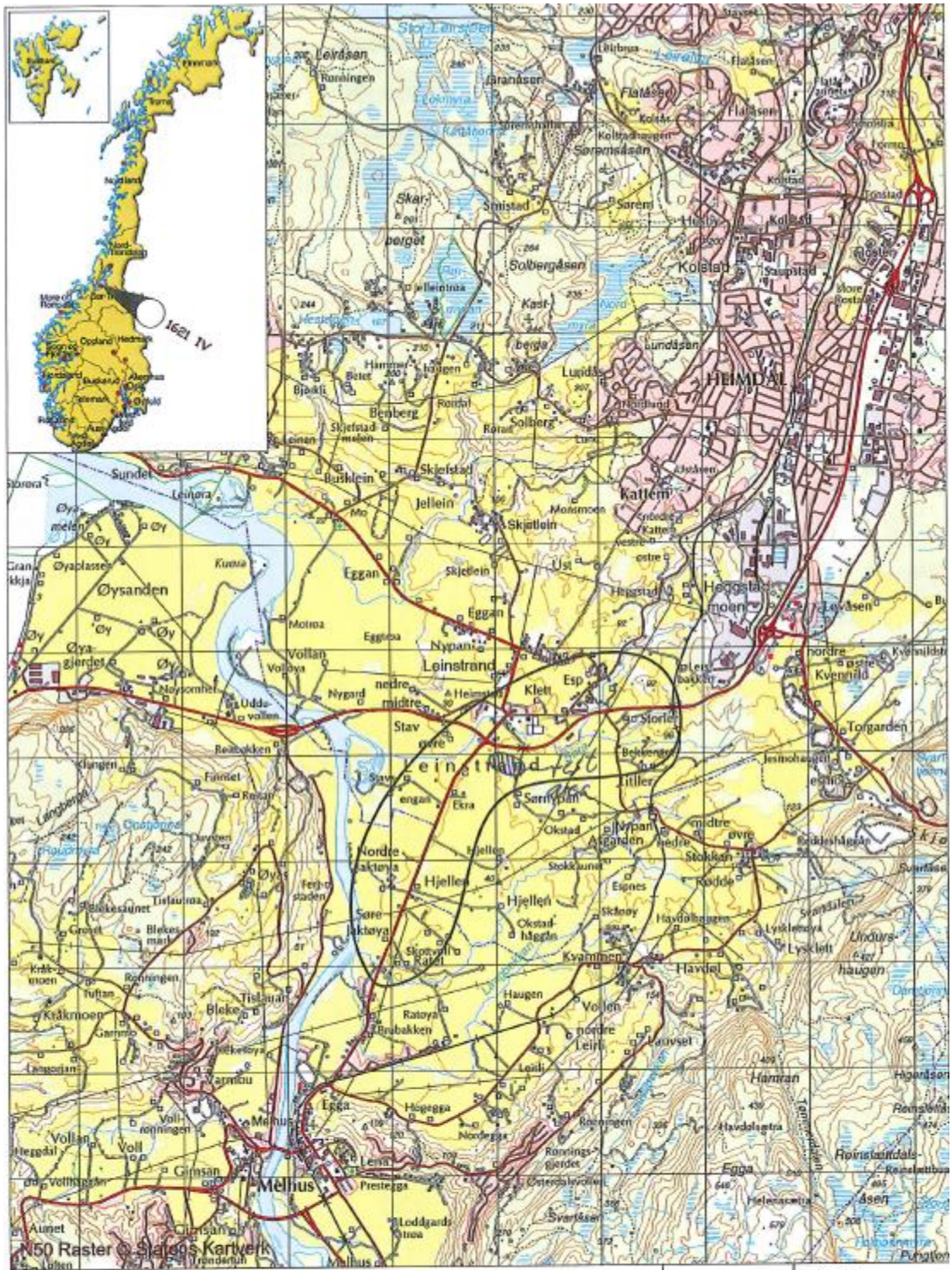
E. Sieve analyzis of sampled non-sensitive clay from Fallan BP 4



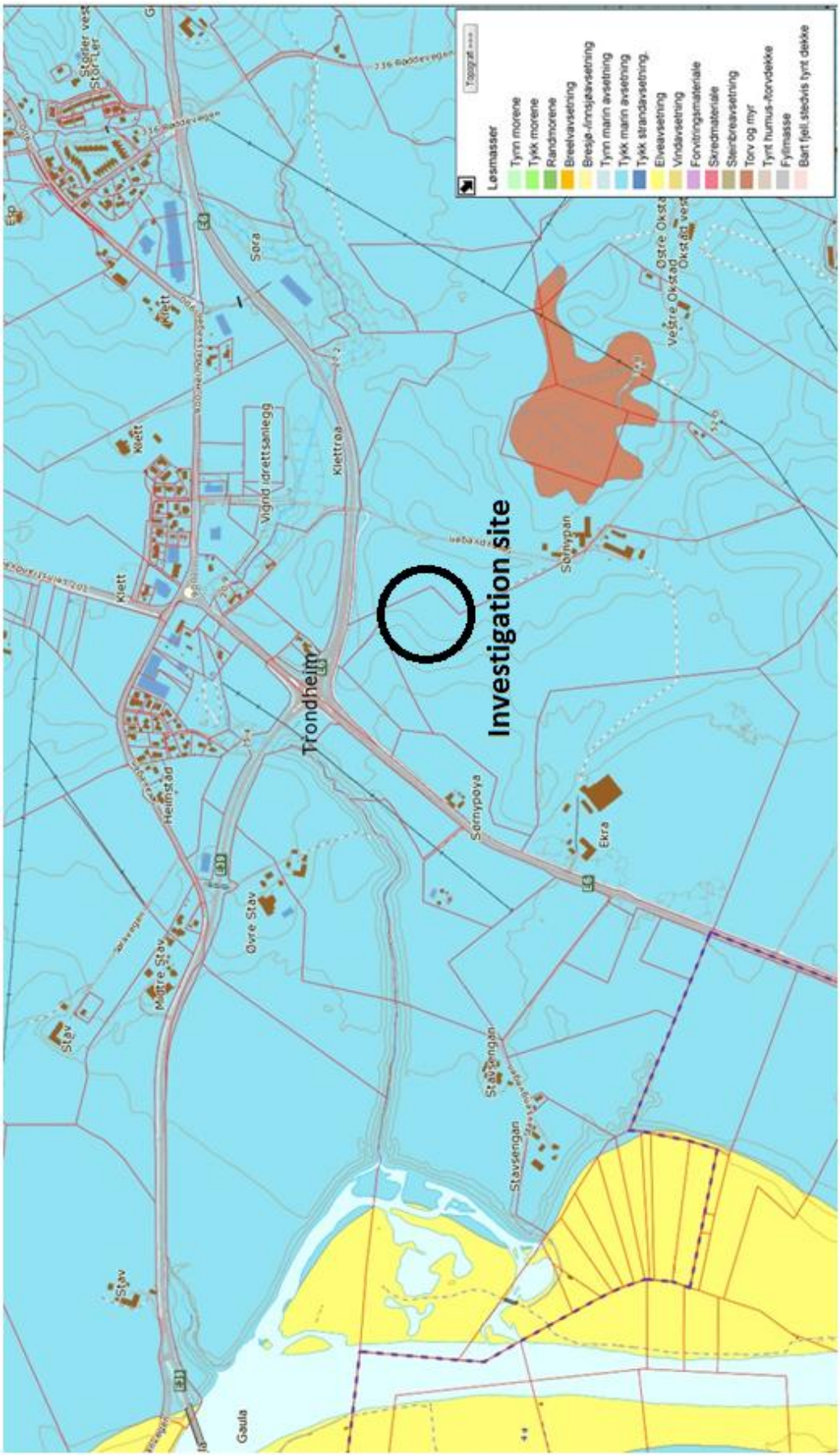
F. Sieve analyzis of sampled quick clay from Fallan BP4



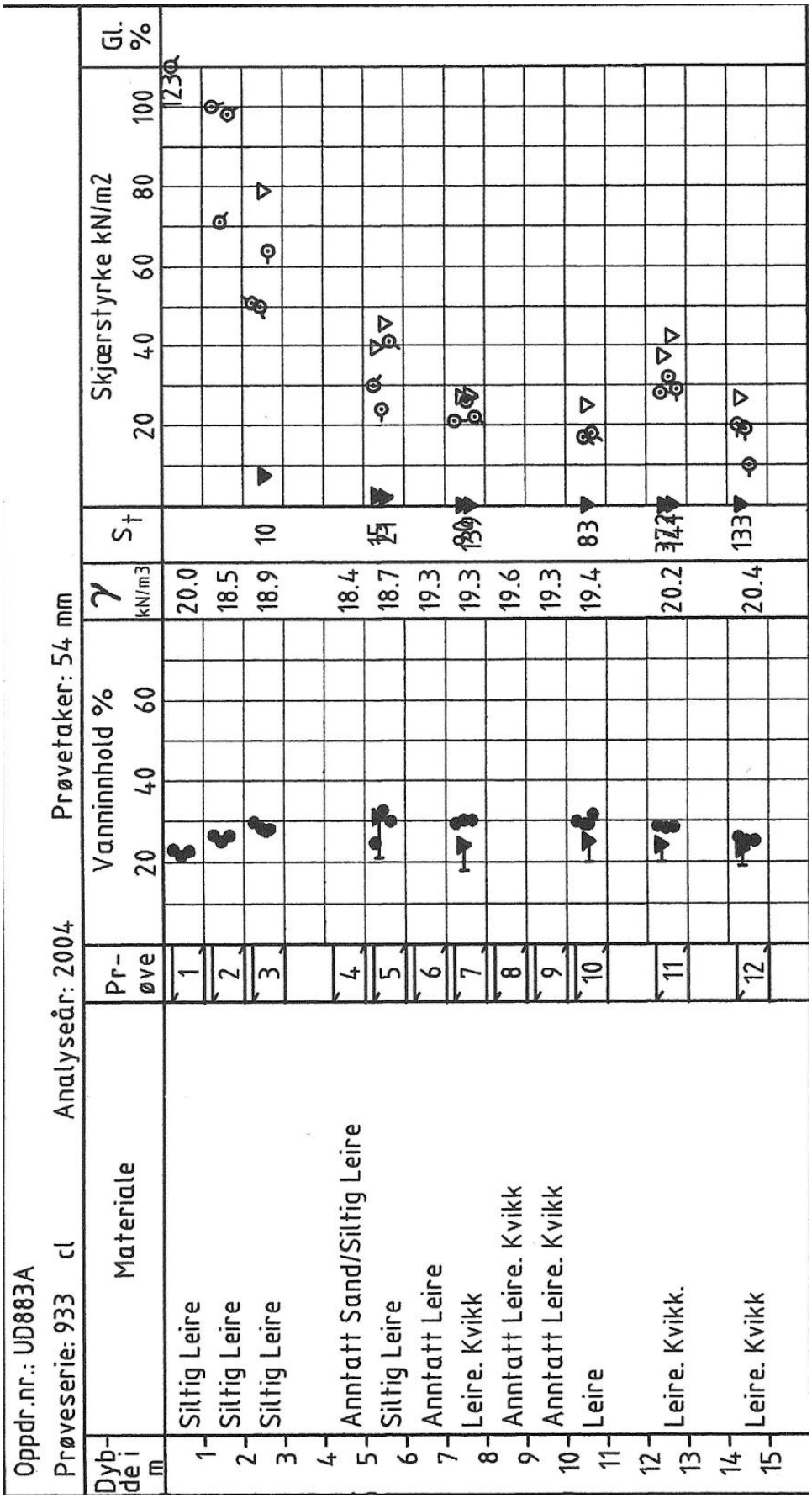
G. Overview map of Klett



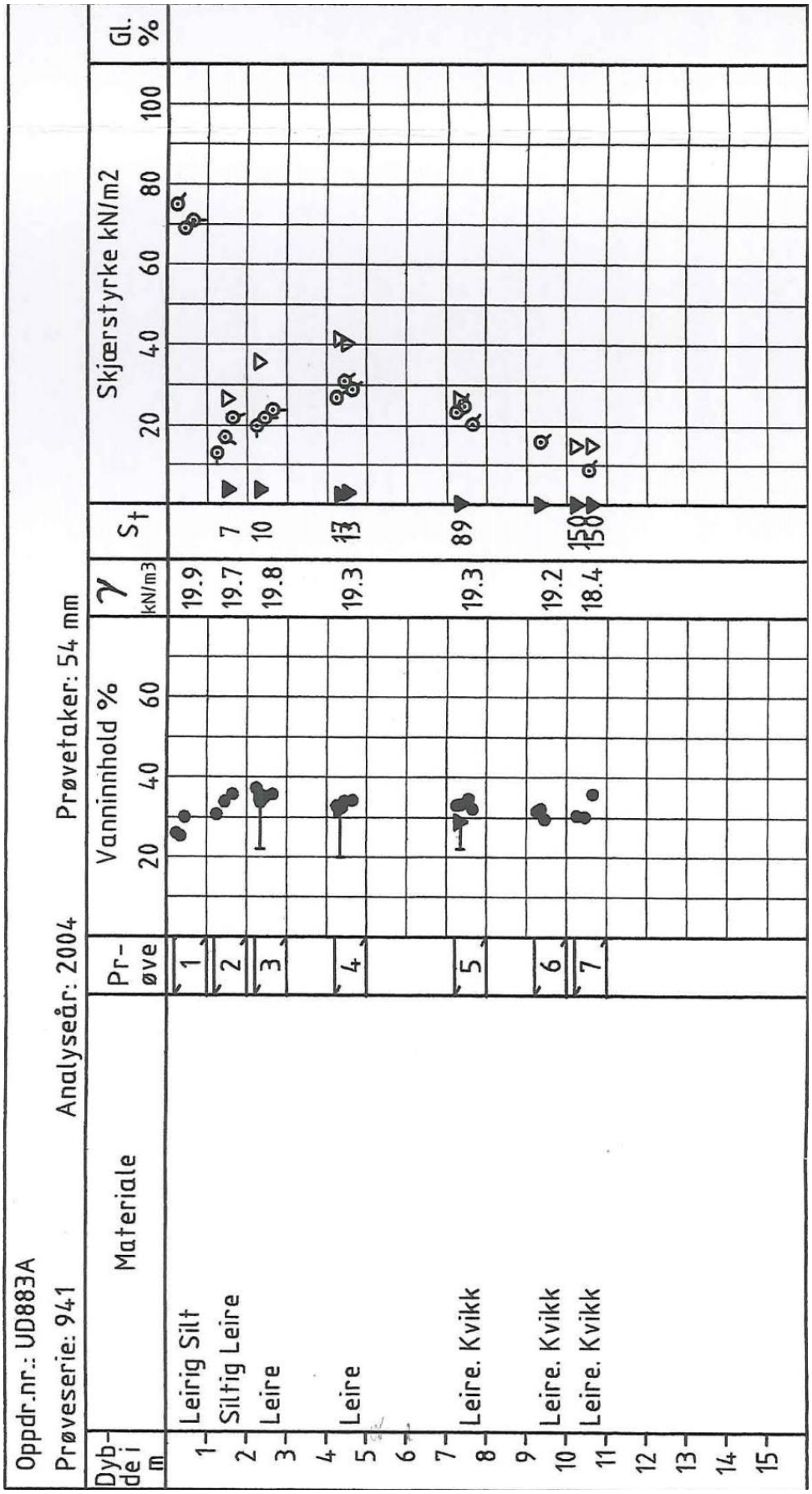
H. NGU Geological map of Klett



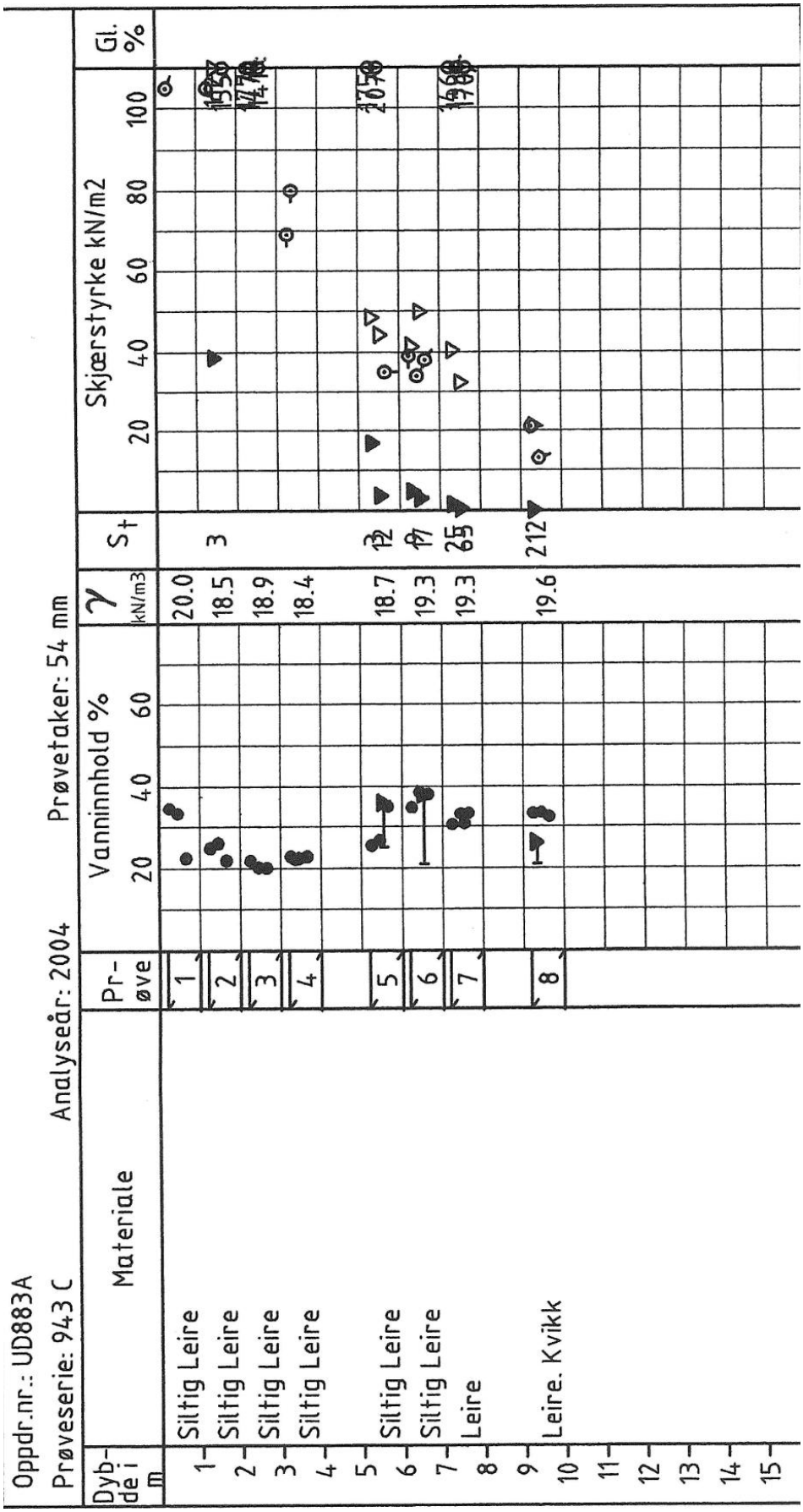
I. Laboratory results from Klett BP 933 by Multiconsult



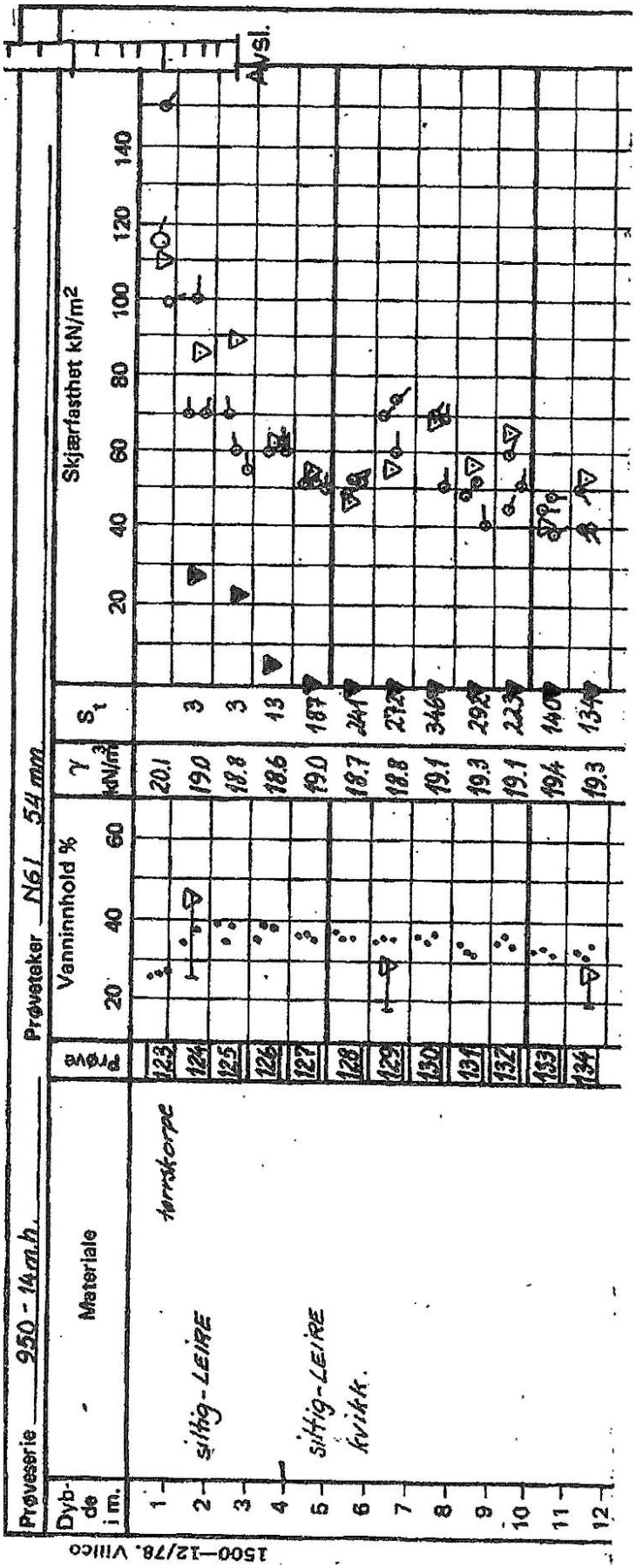
J. Laboratory results from Klett BP 941 by Multiconsult



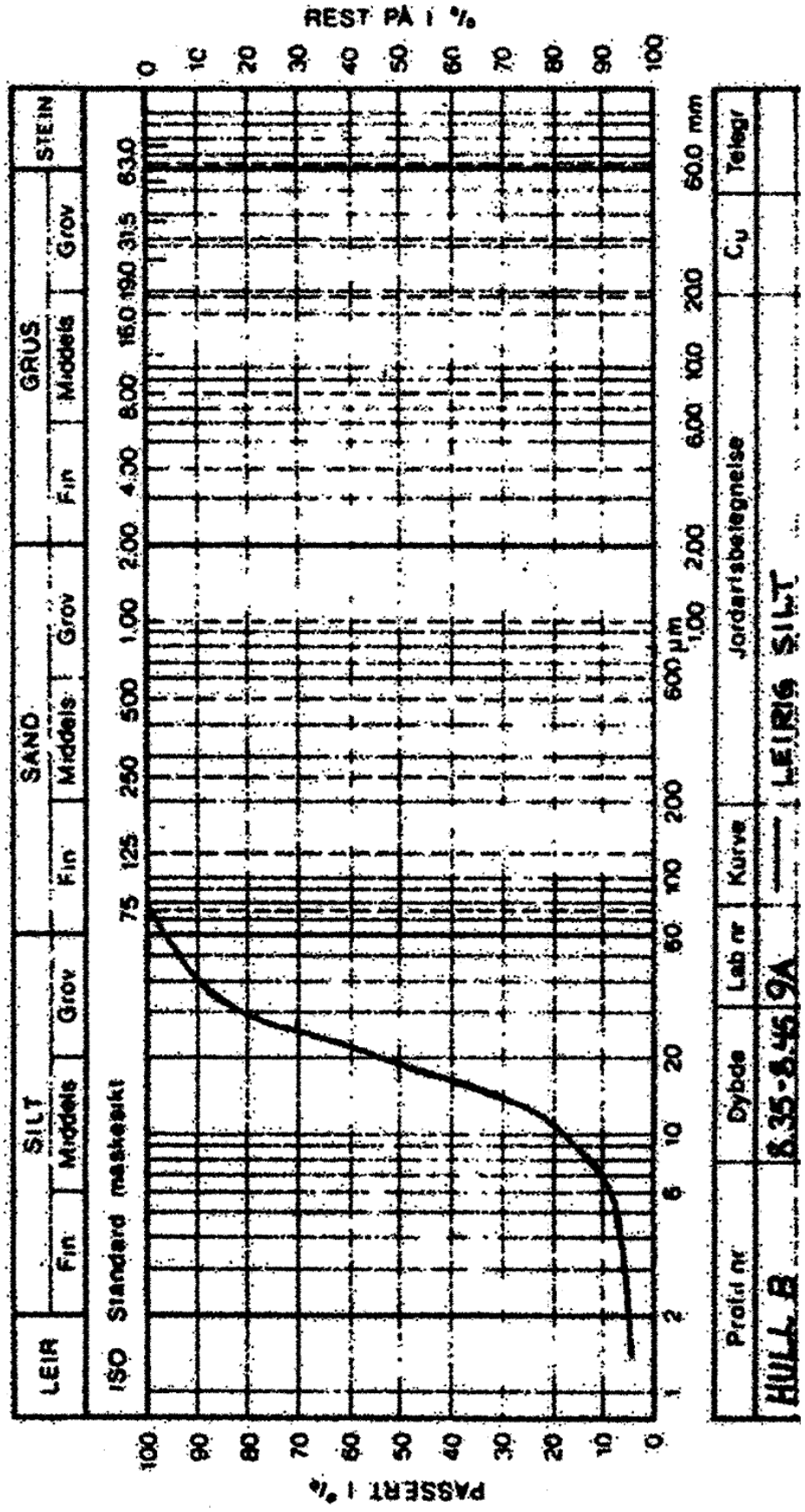
K. Laboratory results from Klett BP 943 C by Multiconsult



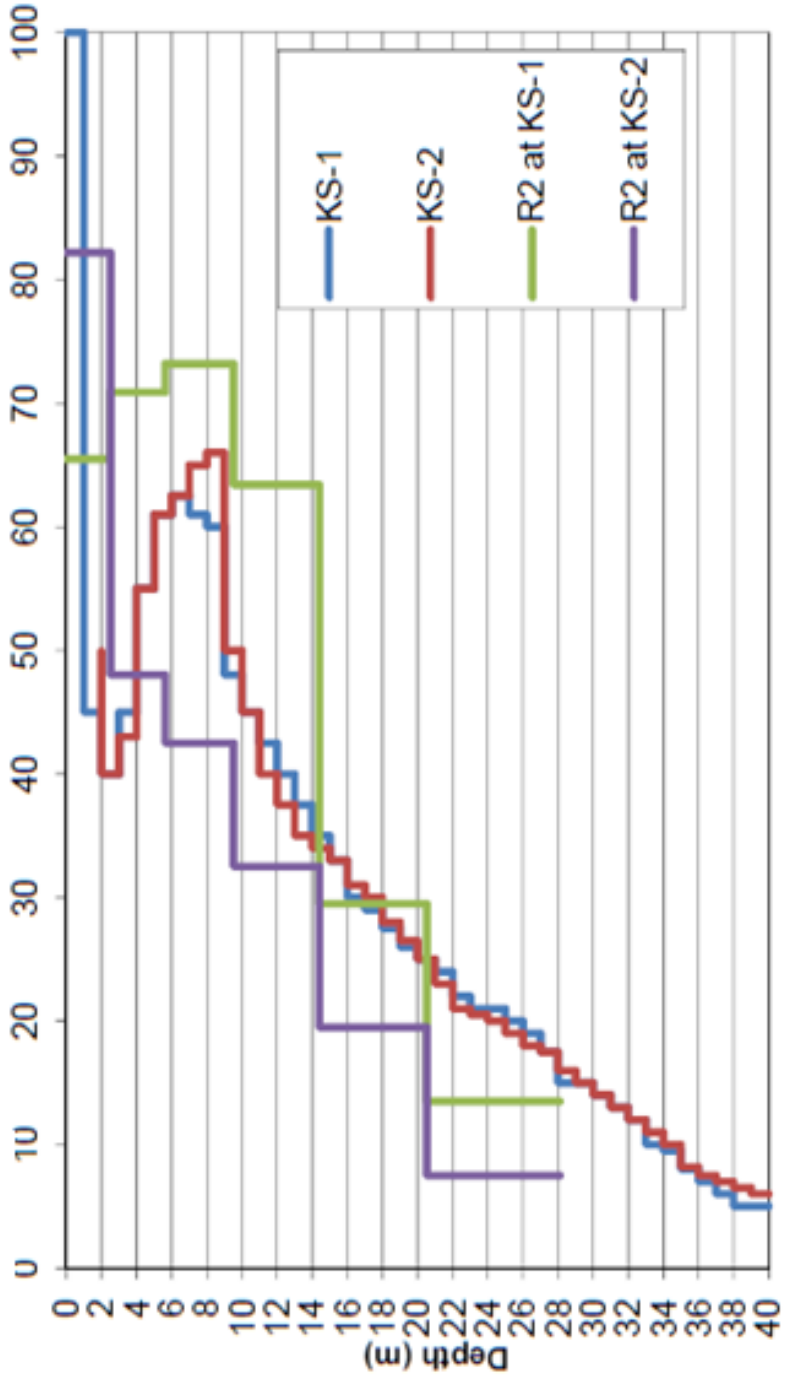
L. Laboratory results from Klett BP 950 by Multiconsult



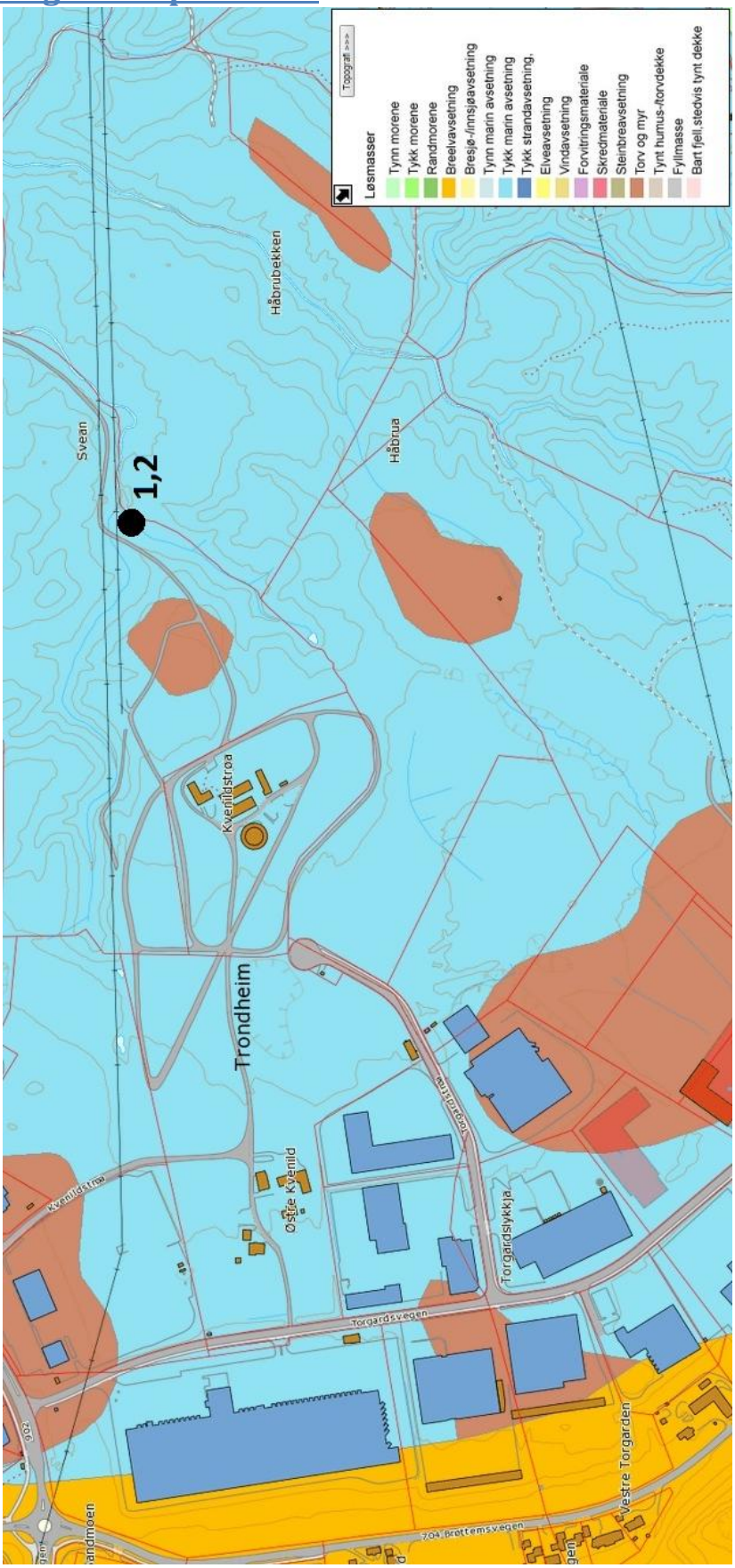
M. Sieve analyzis of sampled quick clay from Klett by Multiconsult



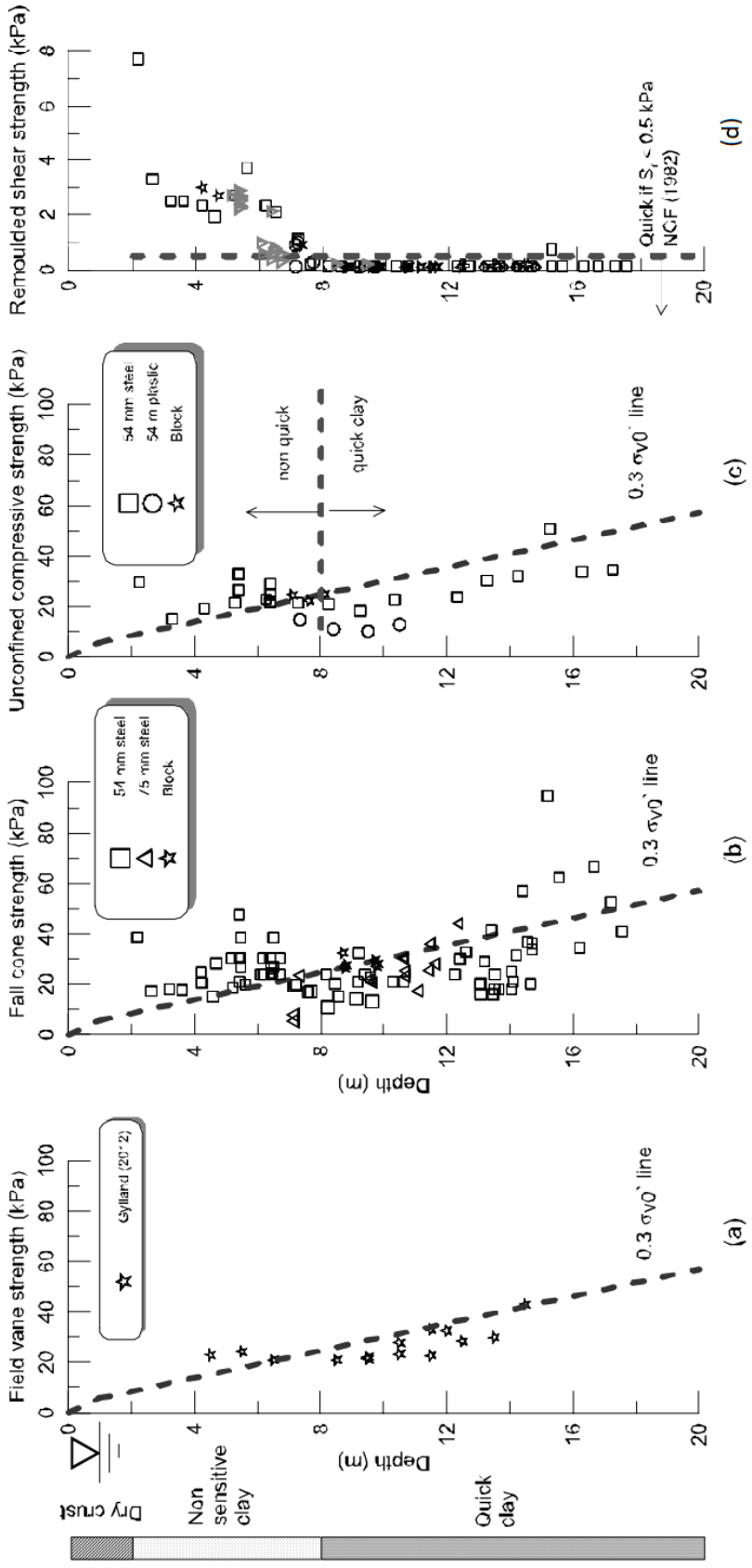
N. Resistivity profiling with MASV and RCPTu from Klett S1 and S2 by Anex Geoservice Limited for Multiconsult



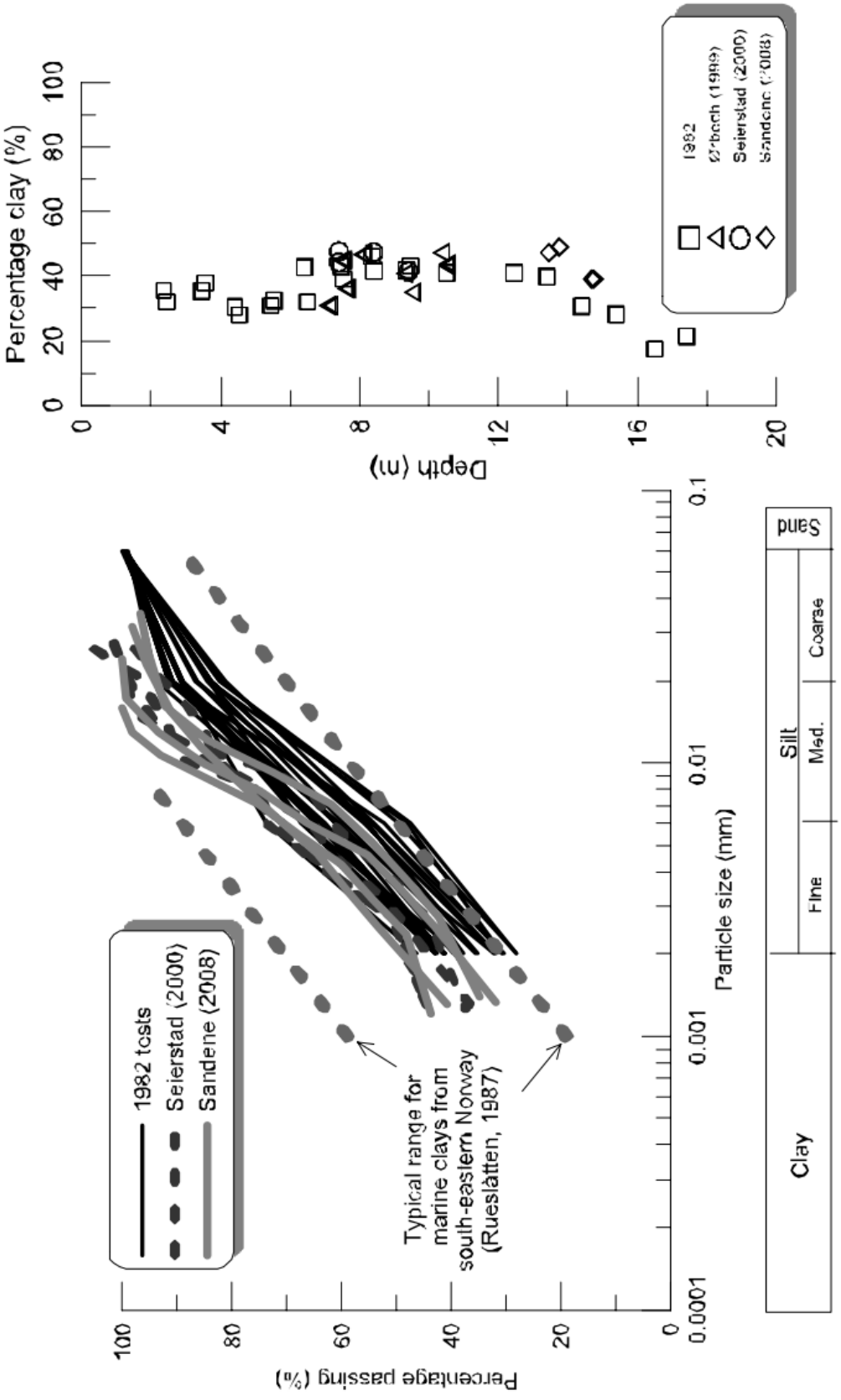
O. NGU Geological map of Tiller



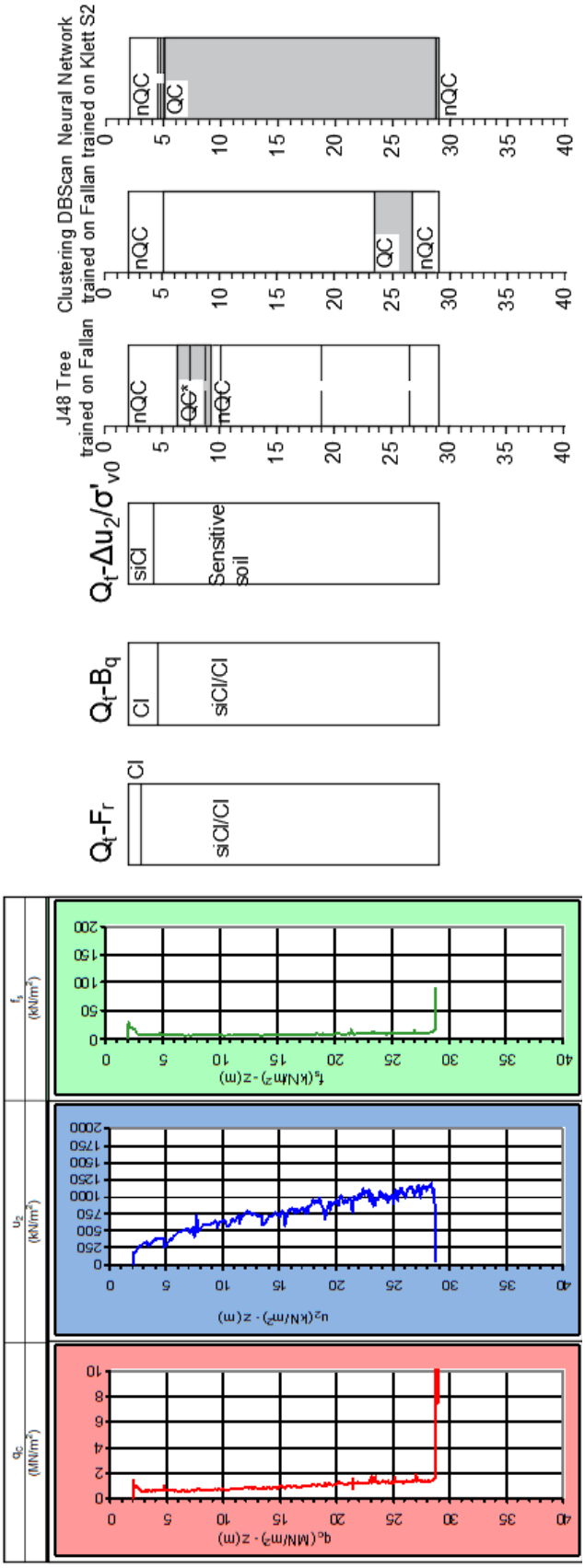
P. Undrained shear strength from index tests and remoulded shear strength with depth by Gylland et al. on Tiller clay



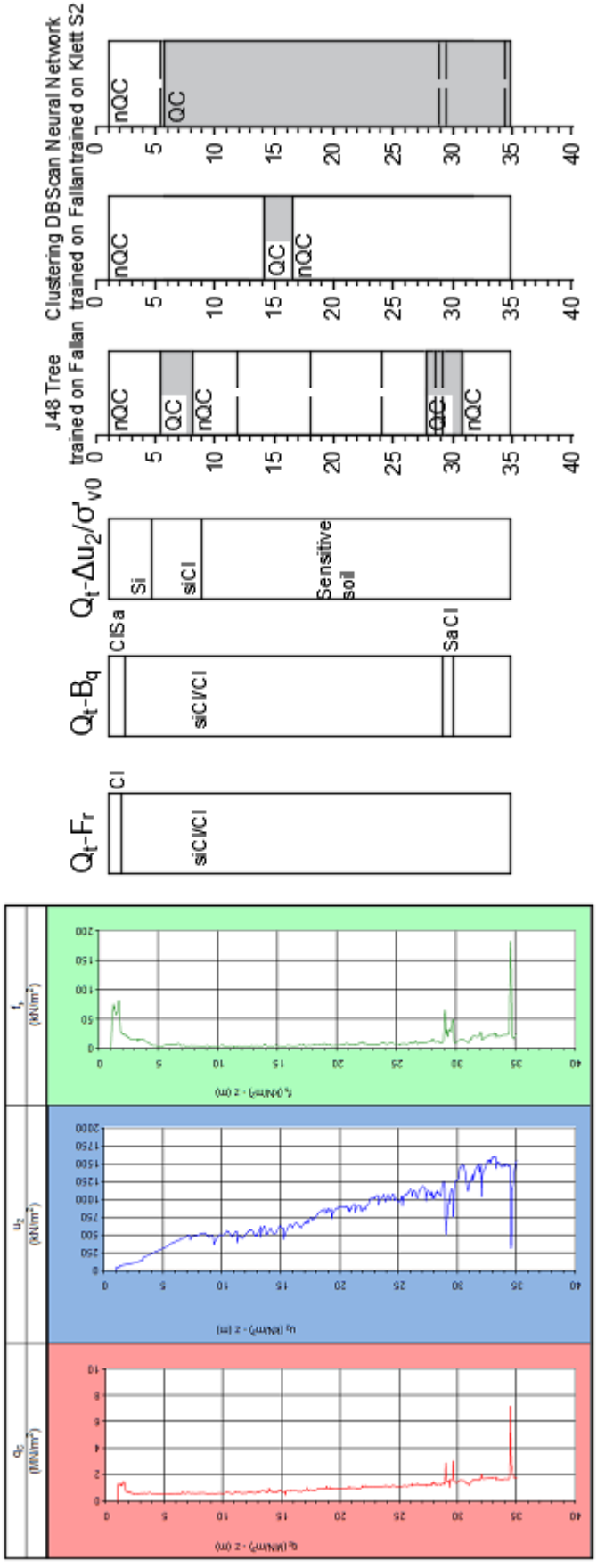
Q. Particle size distribution curve and clay content with depth by Gylland et al. on Tiller clay



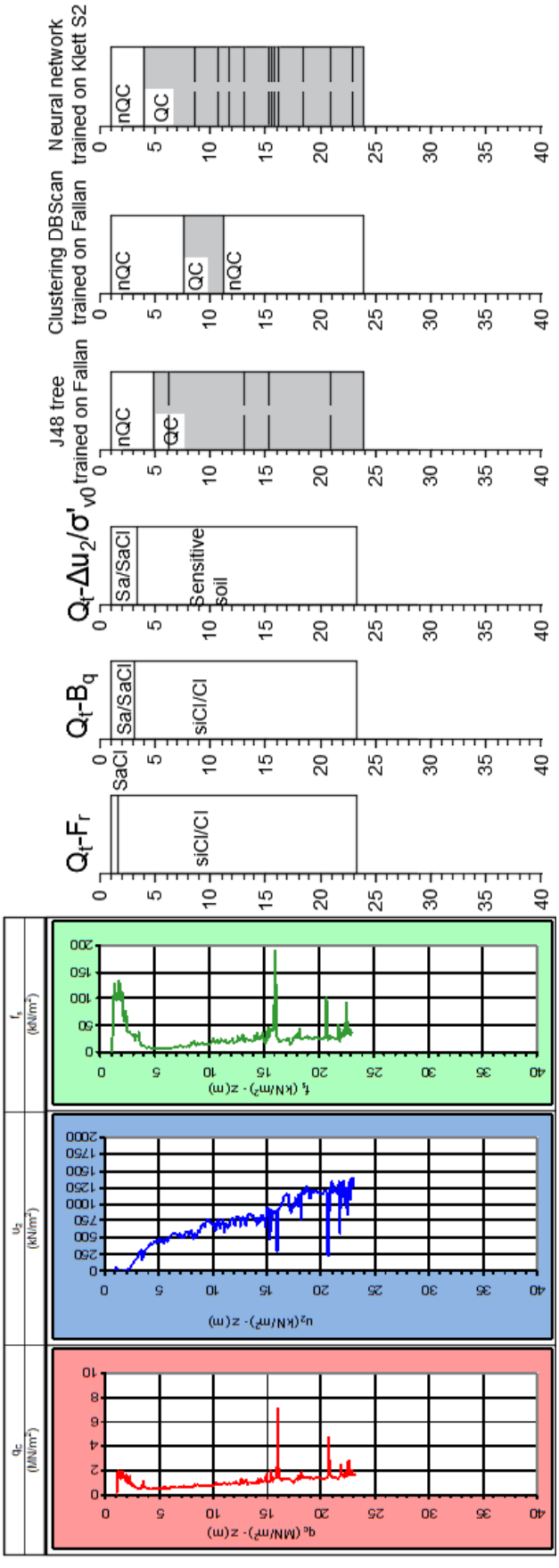
R. Results for Fallan RCPTu BP 4 from different methods



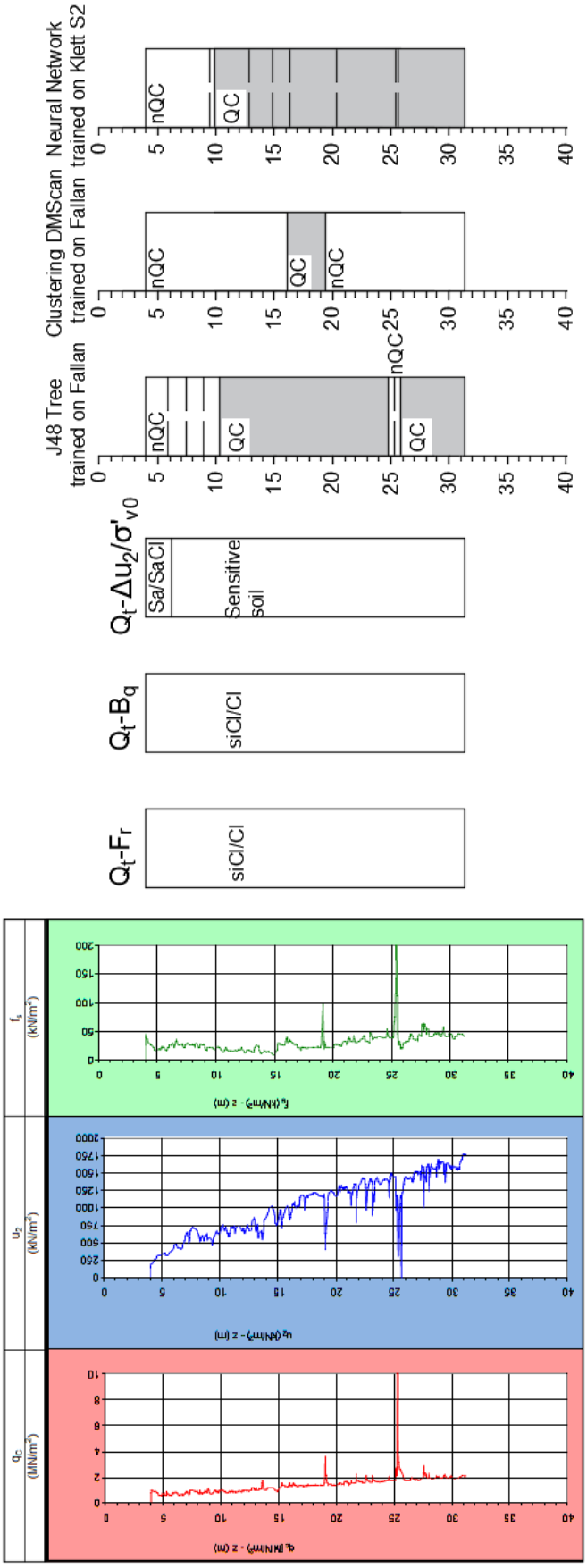
S. Results for Fallan CPTu 2 from different methods



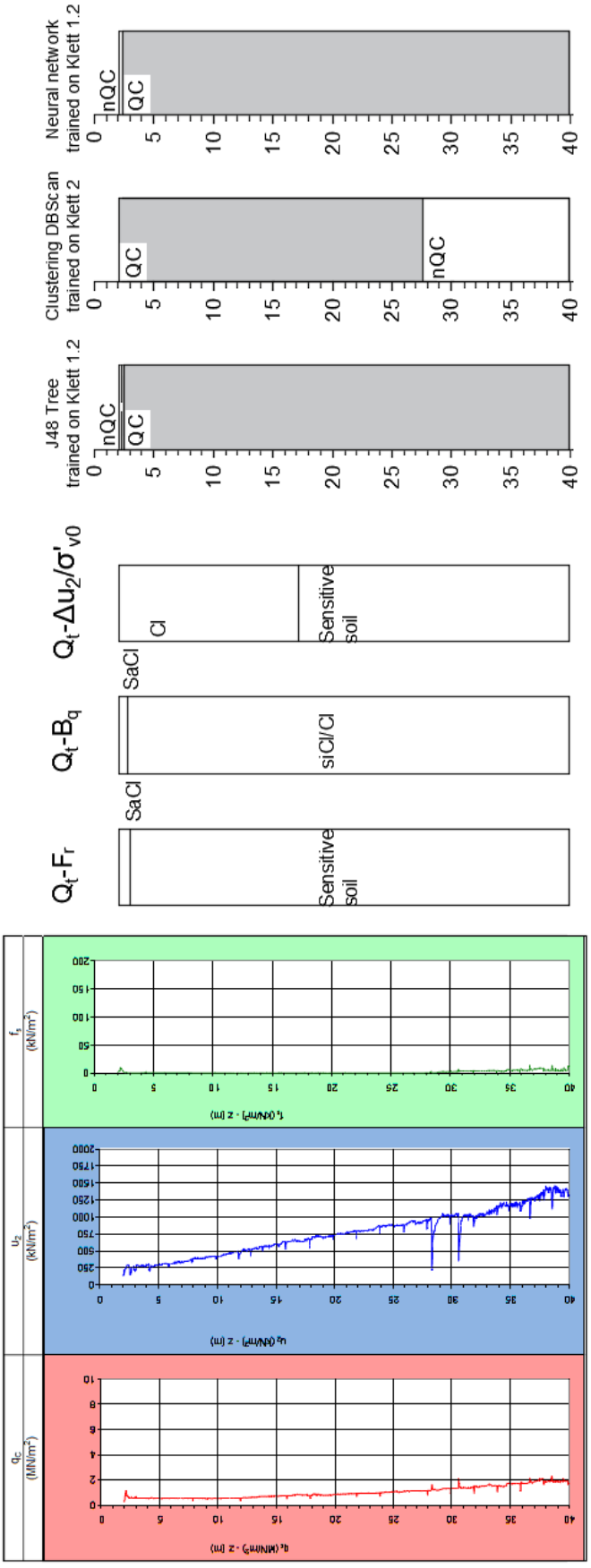
T. Results for Fallan CPTu 5 from different methods



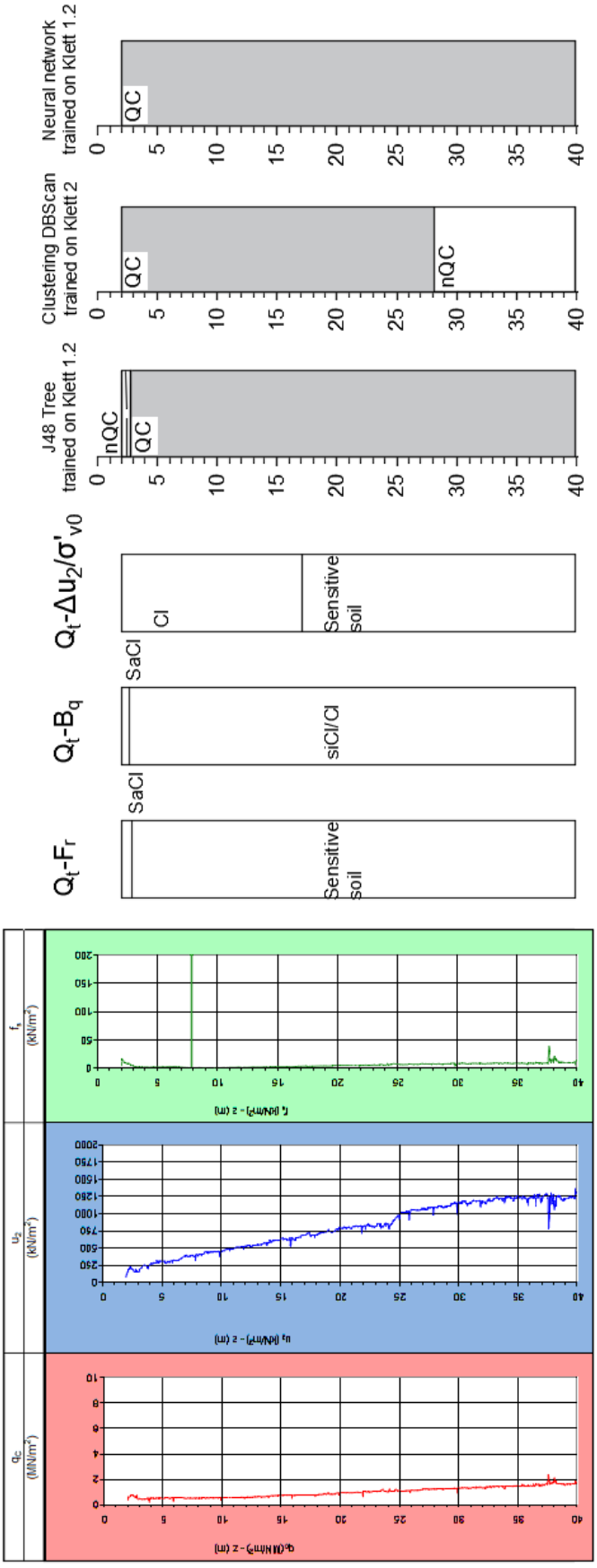
U. Results for Fallan CPTu 7 from different methods



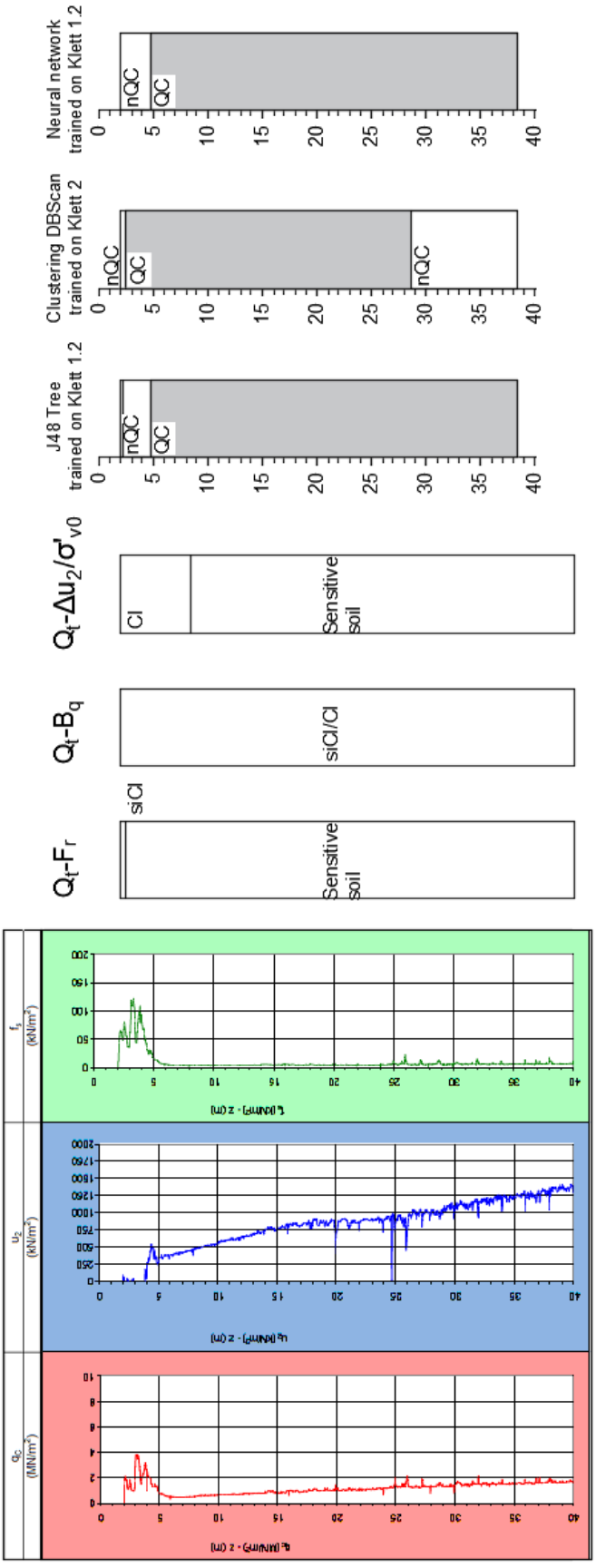
V. Results for Klett CPTu 1502 from different methods



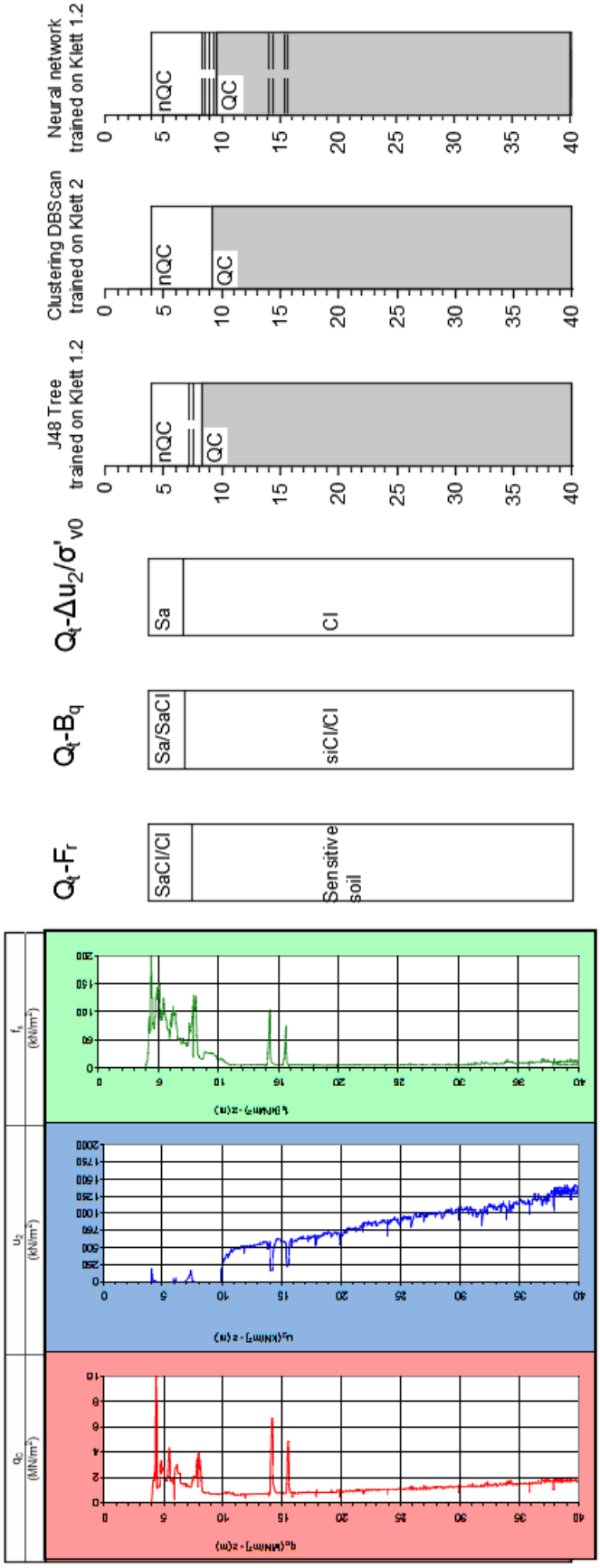
W. Results for Klett CPTu 1503 from different methods



X. Results for Klett CPTu 1504 from different methods



Y. Results for Klett CPTu 1505 from different methods



Z. Results for Klett S1 from different methods