Rapid exploration of (large) sets of cone penetration test data

DOMINIQUE NGAN-TILLARD[1], GERARD KRUSE[2], JOS MACCABIANI[3] & JORDY MOLLÉ[4]

[1] *Delft University of Technology. (e-mail: d.j.m.ngan-tillard@citg.tudelft.nl)*
[2] *Geodelft. (e-mail: g.a.m.kruse@geodelft.nl)*
[3] *Geodelft. (e-mail: j.maccabiani@geodelft.nl)*
[4] *Delft University of Technology. (e-mail: j.mollé@student.tudelft.nl)*

Abstract: In the Netherlands, large databases of cone penetration tests (CPTs) have been developed by municipalities, knowledge institutes and geotechnical engineering consultants. They contain hundreds of thousands of digital CPT traces and are worth exploring during the early planning of any civil and environmental engineering projects. However, time and geotechnical expertise to process the mass of information available, manually, in a sufficiently objective way is lacking. Grouping in the depth-cone parameters space, the CPT traces which look alike by fuzzy-c means clustering, permits rapid identification of the main features of the subsurface. CPT traces can also be trimmed before clustering to reveal more subtle features of the subsurface.

The clustering tool has been developed using the geotechnical data gathered for the construction of a new neighbourhood in Rotterdam, The Netherlands. The use of the tool has helped to reveal a buried Holocene channel and its associated crevasse splay channels. The performance of the tool was further verified in two other sites: the Betuwe line, a freight railway line between Rotterdam container port and Germany and Bergambacht where a large scale dike failure test took place.

The tool was found to be robust and efficient at the regional and local scales. Its simplicity should facilitate its use within the civil engineering community. It allows a better integration of the ground factor into the sustainable (re-)development of urban quarters and the selection of industrial sites having priority for decontamination.

Résumé: Aux Pays-Bas, de larges banques d'essais pénétrométriques ont été développées par les municipalités et les bureaux de consultants en géotechnique. Elles contiennent des centaines de milliers de traces pénétrométriques et sont intéressantes à explorer pendant la phase préliminaire de tout projet de génie civil ou de protection de l'environnement. Toutefois, le temps et l'expertise géotechnique manquent pour traiter manuellement la masse d'information disponible de façon objective. Grouper dans l'espace profondeur-paramètres pénétrométriques, les traces pénétrométriques qui se ressemblent à l'aide de la méthode du regroupement flou permet une identification rapide des principaux éléments architecturaux du sous-sol. Les traces pénétrométriques peuvent aussi être coupées avant d'être regroupées de façon à révéler des éléments plus subtils du sous-sol.

L'outil de regroupement flou a été développé en utilisant les données géotechniques collectées pour la construction d'un nouveau quartier de Rotterdam aux Pays-Bas. L'utilisation de l'outil a permis de révéler la présence d'un chenal Holocène enfoui et de crevasses dans ses rives. La performance de l'outil a été vérifiée sur 2 autres sites: la ligne de Betuwe, une ligne de fret de chemin de fer entre le port container de Rotterdam et l'Allemagne et Bergambacht, le site d'un test de rupture de digue grandeur nature.

L'outil est robuste et efficace aux échelles locales et régionales. Sa simplicité devrait faciliter son utilisation par la communauté de génie civil. Il permet une meilleure intégration du facteur sol dans le développement durable des quartiers urbains et la sélection de sites industriels à décontaminer en priorité.

Keywords: Cone penetration test, site investigation, data analysis.

INTRODUCTION

By 2005, the municipality of Rotterdam, The Netherlands, had to draw a picture of its industrial heritage in order to quantify the risk of contaminant migration around 50000 industrial sites that have been inventoried on its territory. A good knowledge of the subsurface was required to discriminate areas that present adverse ground water flow conditions and necessitate immediate remedial work from areas showing low risk of contamination spreading. Rather detailed data on the coastal and riverine lowland deposits that constitute Rotterdam shallow subsurface, is available: the geotechnical database counts more than 25000 CPTs and every year about a thousand new CPTs enter the database. However, time and geotechnical expertise to process this mass of information manually in a sufficiently objective way is lacking. Within the framework of the Delft Cluster research programme sponsored by the Dutch Government, techniques have been developed to optimize and automate the interpretation of CPT logs. The bottlenecks in interpretation of CPT logs have been identified and the desirable features for the new tools have been established. In this paper one of these new tools is presented and its effectiveness is demonstrated using data from the Nesselande area, a residential zone under development near Rotterdam. The tool consists of grouping the CPT logs which look alike in a population of CPTs. It can easily be implemented to process a large database of CPTs and is found to be most useful in analyzing regional patterns in field data.

PILOT SITE : NESSELANDE VINEX

Nesselande Vinex location is a new housing area developed east of the city of Rotterdam, the Netherlands and is underlain by extensive peat and soft clay deposits (Figure 1). In order to develop a geotechnical schematization of the subsurface for settlement predictions and pile foundations design, between 1996 and 2000, the 2.5 by 3.5 km^2 of Nesselande have been probed by CPT trucks at more than 565 locations and drilled at about 60 locations. The high density of site investigation reflects the sensitivity of the project to the expected variability of Nesselande subsurface. Based on the lithological composition of the individual boreholes performed by the Dutch Geological Survey Department, TNO-NITG for regional mapping purposes, the Late Pleistocene and Holocene deposits that make up the upper 10 m of soil in the Nesselande area can be divided into four main sequences (Weerts, 2001):

1. Fluvial braiding environment; deeper than about 10 m below the surface, coarse grained sandy deposits,
2. Fluvial meandering environment; overbank fines, sandy channel deposits and some peat from 5 to 10 m below the surface,
3. Fluvial anastomosing environment; overbank fines and peat from 5 m below the surface to the present surface,
4. Marine tidal basin environment and peat marsh; clayey tidal deposits and peat at the surface to a maximum depth of 5 m below the surface.

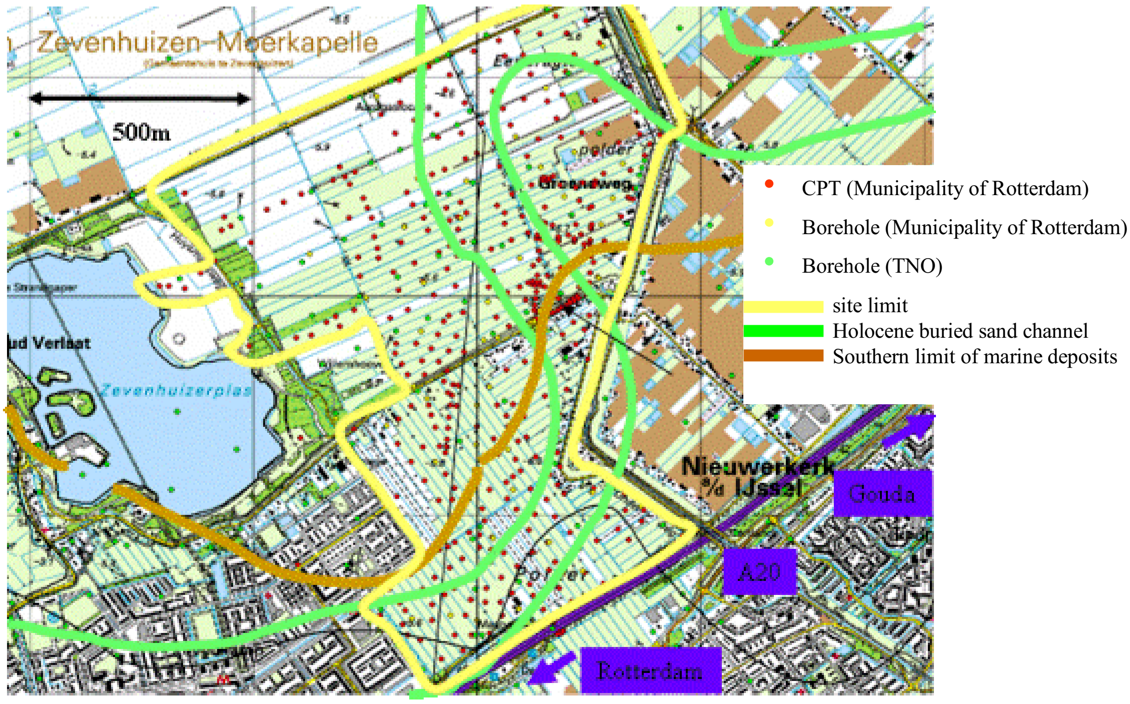

Figure 1. Nesselande Vinex location: its situation, site investigation and main geological features.

The predominantly soft Holocene layers show relevant heterogeneity in the vertical and horizontal directions. Only some units have a strong lateral continuity. This continuity exists for peaty horizons which constitute, in some places, the base and top of the fluvial meandering second sequence and the top of the anastomosing third sequence. In the overbank deposits of the second and third sequences, sandy crevasses deposits including sandy clayey channel fill sequences can be expected. Figure 1 displays the geological features shown on the 1: 50,000 geological map of Nesselande (RGD, 1995, 38W), (RGD, 1998, 37O), i.e., the buried channel of a Holocene river as well as the southern extent of the marine deposits at ground surface level. Figure 2 presents the geology of Nesselande along a South-North section. The deposits of the 4 sequences mentioned above appear on this section.

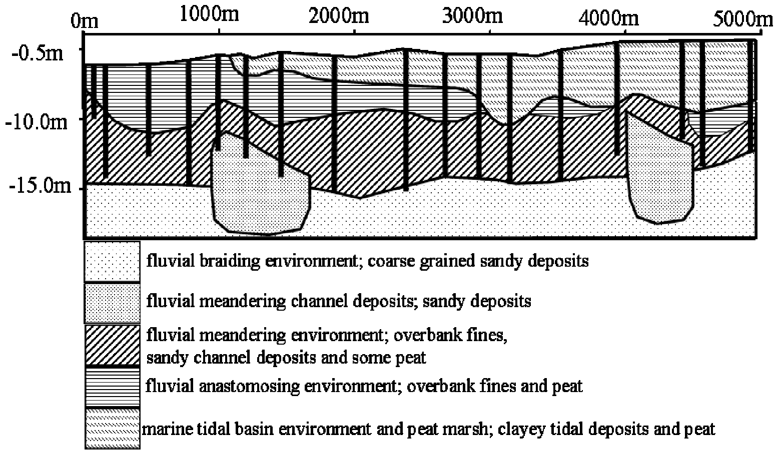

Figure 2. South-North geological section across Nesselande Vinex location

CLUSTERING DEPTH SERIES

A technique, called fuzzy C-means clustering, introduced by Bezdek (1973), has been successfully applied to the CPT data. This technique groups data by minimizing the variance within clusters, while maximizing the variance between clusters. Each data point belongs to a cluster, with a certain degree of membership. The necessity to know the optimal number of groups present in the data and the theoretical need to normalize and standardize the data before processing are discussed in (Maccabiani et al, 2003).

Fuzzy C-means clustering has been applied to the CPT logs of Nesselande in two spaces separately: the cone resistance-depth and friction ratio-depth spaces in order to group logs with similar cone resistance, qc and friction ratio, Rf traces. After preliminary processing, more than 500 CPT logs truncated between 0 m NAP and –16 m NAP were divided by cluster analysis into two groups, Square and Circle. Figure 3 presents logs of representatives of both groups. As shown in Figure 4, most of the members of the Square group are concentrated within the lines indicating on the 1:50000 geological map the position of the large buried Holocene channel. The agreement between the geological model and the spatial distribution of the members of the Square cluster is remarkable. The few points of divergence between the zones are explained by the difference in density of information. Observations made during earthworks revealed the cluster pattern to be more accurate than the geological model, which was based on a less dense observation grid.

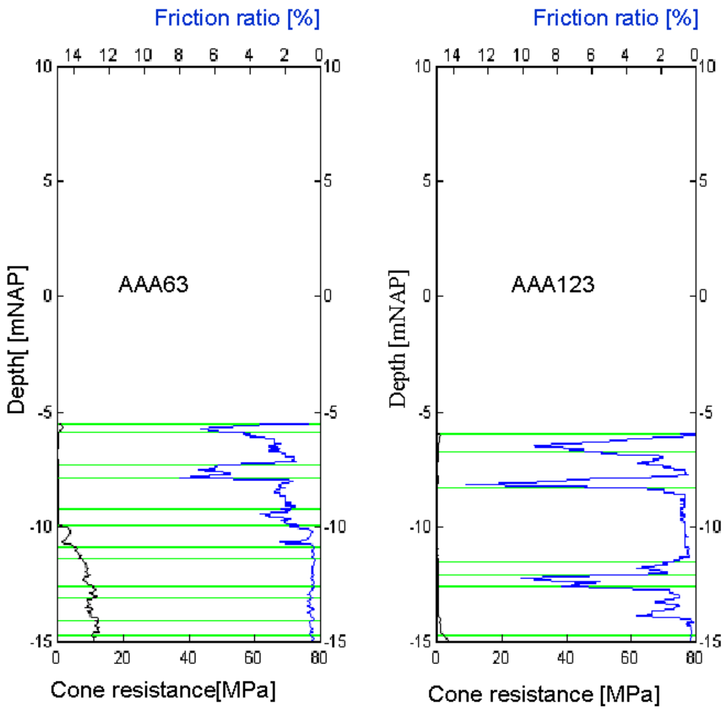

Figure 3. Best representatives of the Green Square (left) and Dot (right) clusters resulting from the cluster analysis. The Cone resistance (qc) and friction ratio (Rf) are plotted versus depth in black and blue, respectively.

Fuzzy C-means clustering was subsequently used to detail certain sections in specific areas, revealing more subtle features. The detailed analyses comprised truncating the cone penetration profiles to sections between –10.5 mNAP and –12.5 m NAP in an apparently more complicated area before grouping them. The clustering of the sections revealed an apparent crevasse splay deposit in the overbanks of the larger Holocene channel. This crevasse is schematized in Figure 5. A further detailed analysis did concern processing of the first meters of the friction resistance profiles. This analysis revealed a diffuse nature of the transition between marine and river deposits at the spatial resolution of CPT logs.

In brief, applying Fuzzy C-means clustering to the CPT logs in the space of the cone parameters-depth is simple and efficient. A large number of CPT logs have been analyzed jointly, detailing the Nesselande geological model for geotechnical engineering applications. For each group identified, a specific analysis taking into account the statistical properties of the group can be tailored. For example, the expert can decide to segment the CPTs of a same population with the same number of segments.

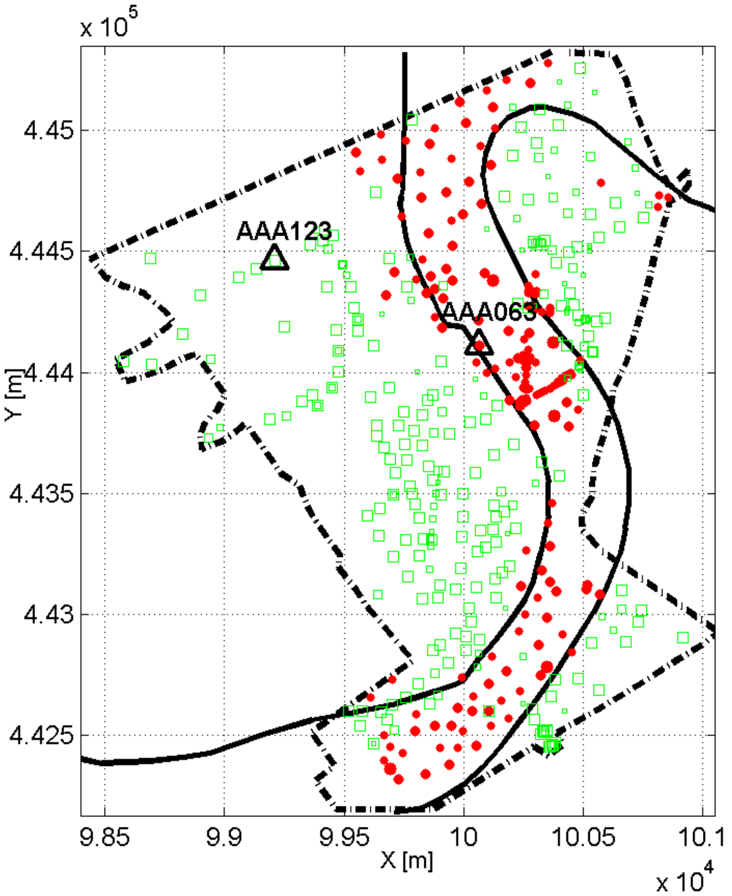

Figure 4. Results of the fuzzy c-means clustering analysis performed in the cone resistance-depth space and location of best representatives. The size of the symbols is proportional to the degree of membership. Green square: inside the channel, red dot: outside

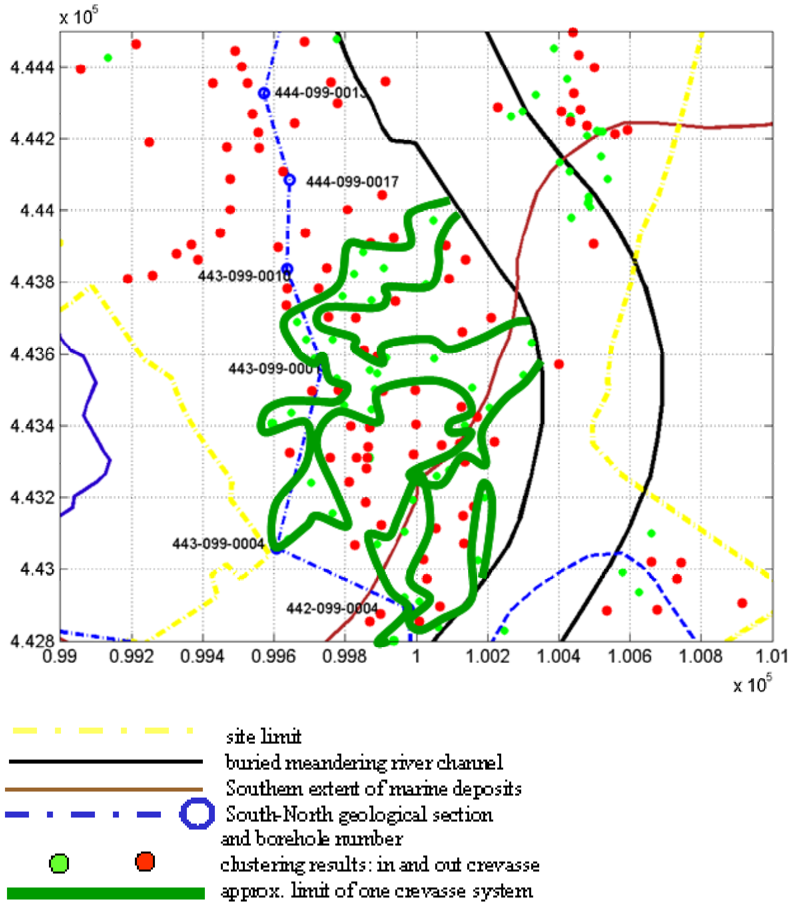

Figure 5. Crevasse revealed in the banks of a Holocene channel by applying Fuzzy c-means clustering to CPT logs truncated between –10.5 and –12.5 mNAP. Green dot: inside crevasse, red dot: outside.

PERFORMANCE VERIFICATION

The performance of the tool was further verified in two other sites: km 7 to 17 of the Betuwe line, a freight railway line between Rotterdam container port and Germany and Bergambacht where a large scale dike failure test took place.

In the former case, the fuzzy-c means of depth series highlighted the eastwards decrease of the depth to Pleistocene and allowed a rapid identification of the Holocene stream belts anastomosing or meandering under the railroad track. In the latter case, the tool revealed the presence of an unexpected river gulley filled in with silty clay often mixed with sand layers under the inlands toe of the dike.

CONCLUSIONS

With a view to characterizing the subsurface and preparing a detailed lithological model based on CPTs the applicability of several quantitative procedures was explored. In this paper, one of the tools is presented: the fuzzy C-means clustering of depth series.

By applying fuzzy C-means clustering of depth series, CPT logs with similar patterns among Nesselande large set of CPTs were identified and patterns in the regional distribution of the clusters were easily analyzed. The tool was found to be robust and efficient at the regional and local scales using the data of 2 other sites.

In a country where huge databases of digital CPT traces exist, the applicability of the tool is high. Models of the near-surface can be updated at the exploratory stage of the processing of field data with limited computational effort and time. Then, the ground factor can be integrated in the sustainable (re-)development of urban quarters even before new site investigation data becomes available. In the case of Nesselande VINEX, aligning infrastructure lines along the main buried sand channel revealed by CPT clustering could have been explored in order to reduce maintenance costs.

The fuzzy C-means clustering tool was also found to be efficient to track more subtle architectural features of the near-surface such as sand lenses. It can be of valuable assistance in the selection industrial sites having priority for decontamination.

The simplicity of the tool should facilitate its use within the civil engineering community.

Acknowledgements: The research was partly funded by the Delft Cluster research programme of the Dutch Government and by the Municipality of Rotterdam. The authors want to thank all DC partners and, especially the municipality of Rotterdam (Jos Streng and Ard den Outer).

Corresponding author: Dr Dominique Ngan-Tillard, Delft University of Technology, Stevinweg 1- Post box 5048, Delft, 2600 GA, Netherlands. Tel: +31 15 278 33 25. Email: d.j.m.ngan-tillard@citg.tudelft.nl.

REFERENCES

BEZDEK J.C. 1973. Fuzzy Mathematics in pattern classifications. Ph.D. thesis, Applied Math. Center, Cornell University, Ithaca.

MACCABIANI, J., NGAN-TILLARD, D.J.M. & VERHOEF P.N.W. 2003. Assisting data interpretation: the cone penetration test, Characterization of the Shallow Subsurface: Implications for Urban Infrastructure and Environmental Assessment, Spa, Belgium, Rosenbaum & Turner (eds), Springer, pp. 292-300.

RGD. 1995. Geologische kaart Ondiepe Ondergrond, Gorinchem West, GK38/W.

RGD. 1998. Geologische kaart Ondiepe Ondergrond, Rotterdam Oost, GK37/O.

Weerts H.J.T. 2000. Geology of the Nesselande area: Facies distribution and lithology of the Late Weichselian and Holocene deposit. TNO-report NITG 01-205-A, Utrecht.