

Annals of Warsaw University of Life Sciences – SGGW
 Land Reclamation No 41 (2), 2009: 119–130
 (Ann. Warsaw Univ. of Life Sciences – SGGW, Land Reclam. 41 (2), 2009)

Automated soil profile generation system for the dilatometer test (DMT)

PIOTR BILSKI¹, SIMON RABARIJOELY²

¹Department of Applied Informatics, Warsaw University of Life Sciences – SGGW

²Department of Geotechnical Engineering, Warsaw University of Life Sciences – SGGW

Abstract: *Automated soil profile generation system for the dilatometer test (DMT).* The paper presents an automated method of the soil profile generation using the soft computing method. This method can be used as the main source of knowledge for the future soil classification module, which will be able to identify geotechnical layers without the interaction with the user. The method uses a graph clustering algorithm, which groups together measurements from the Dilatometer Test. The histogram-based similarity measure is used to automatically adjust number of generated categories (soil geotechnical layers). Test site, algorithm and its application to the generation of soil profiles in the Warsaw University of Life Sciences campus are described. Results are discussed to provide a more general insight into application of intelligent algorithms into the geotechnical data analysis.

Key words: soil profiles, soft computing methods, clustering, Dilatometer Test.

INTRODUCTION

The geotechnical examination of soils is an important aspect of the construction of buildings and structures such as tunnels, bridges, etc. With the introduction of geotechnical probes (including CPT or DMT), the soil analysis became faster and cheaper than the traditional method, requiring drilling boreholes. Unfortunately, to correctly identify geotechnical layers based on the measurement data, a wide human

expert's knowledge is required. The latter is usually represented using charts describing soil structure (Marchetti 1980). Their accuracy is limited as they are usually prepared locally and are applicable only in the nearby area. For instance, as most DMT charts were prepared for Western Europe and North America, their usefulness in Eastern Europe is questionable. To create an accurate chart for the preconsolidated soils (such as present in Central Poland), multiple probings are required. This in turn would take a lot of time and be expensive. Novel methods of classification, based on the expert's knowledge (Zhang and Tumay 1996) are proposed. Unfortunately, when the latter is unavailable, additional methods of extracting information from the measurement data sets are required.

The data can be analyzed by the computer algorithm, creating the soil profile, i.e. the set of different categories assigned to measurements at particular depths. They are then analyzed by the geotechnical engineer, who can verify accuracy of the generated profile. Such knowledge can be further used to construct the soil identification module, which will classify the geotechnical layer based on the measured quantities at particular depths. Similar works were

done before (Hashash et al. 2004, Shahin et al. 2005), but new approaches must be proposed.

The following paper presents a generic method for automated soil profile generation. It is verified at the location of Warsaw University of Life Sciences (WULS), Warsaw, Poland. The data delivered for the computer algorithm are obtained using DMT probe, although the method is universal. It was implemented using measurements taken at four different locations. The research presented here is a continuation of the experiments published before (Rabarajoely et al. 2007, Bilski and Rabarajoely 2009). The expected result of the research is the method for extracting knowledge about the soil types and layers from the probes' measurement data.

The organization of the paper is as follows. In the next section the measurement approach and test site are presented. Further, the computer algorithm applied for the soil profile generation is introduced. The next section contains the description of tests and verification of the proposed approach. Finally, conclusions and future prospects are discussed.

DESCRIPTION OF THE DMT METHOD AND TEST SITE

Among multiple non-invasive methods of the soil analysis, the dilatometer test (DMT) is a relatively new, yet powerful probe-based method. It exploits the measurement of the working gas pressure on the membrane (installed in the head of the blade) at the preselected depths, as the probe is inserted into the ground

with constant speed of 0.02 m/s. Two measured parameters are used further by the clustering method: pressure A (the gas pressure in the first phase of the movement of the membrane) and pressure B (the gas pressure obtained at the additional inclination of the centre of the membrane toward the ground). Further details of the DMT probe operation can be found in Lechowicz and Rabarajoely (2000) (1) and Lechowicz and Rabarajoely (2000) (2).

The test site was the Warsaw University of Life Sciences Campus. During its expansion, the Department of Geotechnical engineering did a research aimed at assessing the conditions of the foundation of newly designed buildings. One of the analyzed objects was the laboratorial-didactic building No 34, for which both geotechnical documentation and supervision were conducted. The building, dimension 57×120 m consists of three dilated parts (Fig. 1). Vertically, they contain one-storey ground level (the -1st floor) located at 1.2 m below ground level and a five-storey building above ground level (the 1st to the 5th floor). An auditorium of double-storey height was designed at the central part of the building. The plan of the building includes foundation with the location of the boreholes and CPT and DMT tests.

The profiles of p_0 and p_1 (calculated based on pressures A and B) pressures (measured) and index parameters I_D , K_D and E_D calculated based on the pressures for the selected DMT experiment are in Figure 2. The traditional chart for the soils investigated in the test site is in Figure 3 (where letters A, B, C and D indicate particular lines separate soil states – Marchetti 1980). The borders between

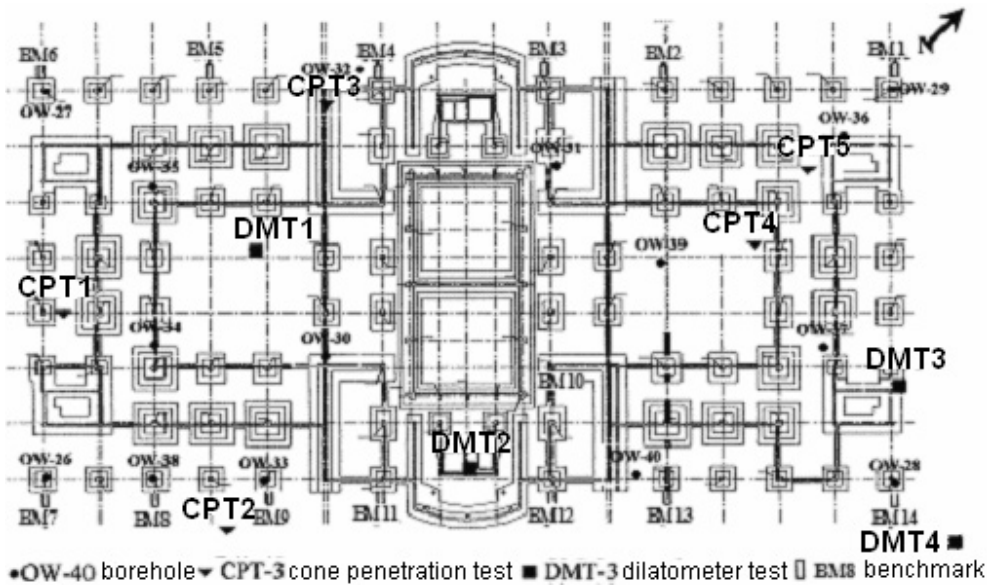


FIGURE 1. Location of boreholes and in situ tests in the foundation of WULS building

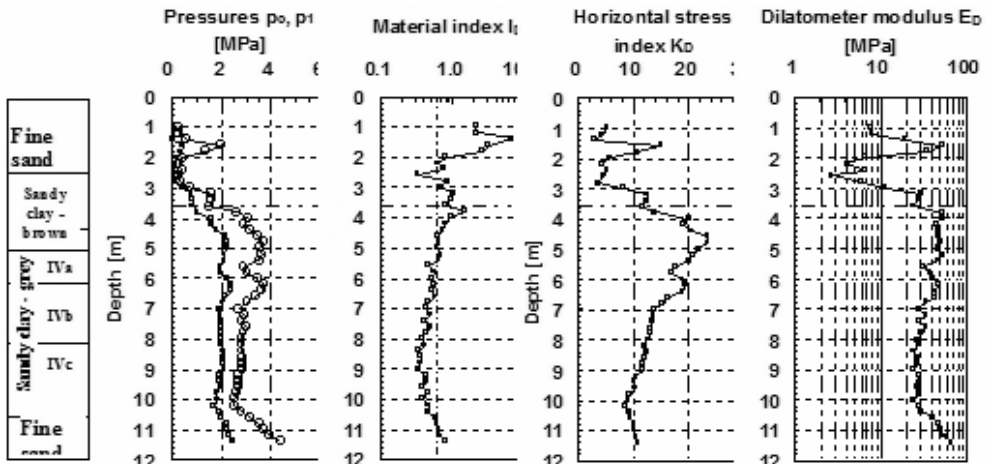


FIGURE 2. Index parameters p_0 , p_1 , I_D , K_D and E_D profiles from DMT in WUoLS, building No 34

the soils are expressed by the vertical lines, while the small black triangles and squares are the measured parameters. The geological layers are not accurately

separated using the chart (triangles are identified as silts and clays). Therefore the artificial intelligence method could improve the analysis.

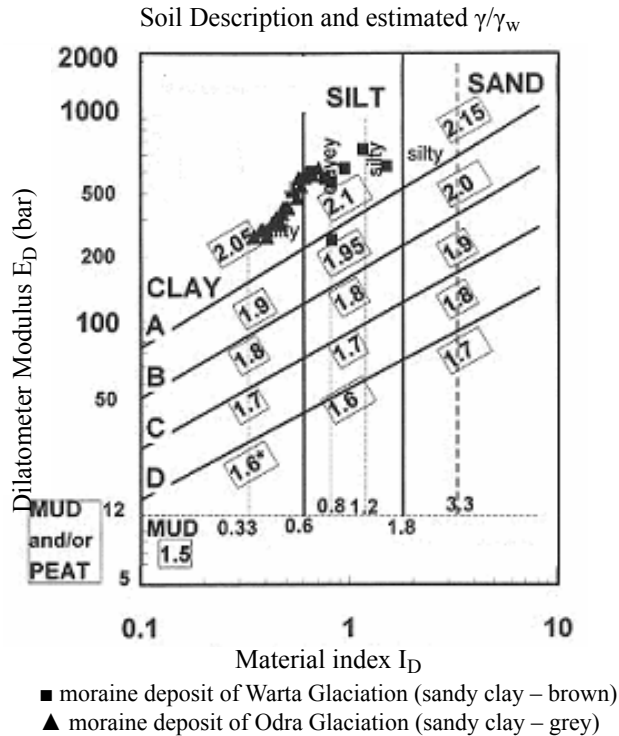


FIGURE 3. Marchetti and Crapps (1981) chart

DESCRIPTION OF THE ALGORITHM FOR THE SOIL PROFILE GENERATION

In the experiment, an automated clustering method is implemented to extract knowledge about geotechnical layers in the examined soil. This knowledge covers also information about the characteristic traits of every soil category, which can be further used to distinguish between the categories. At this stage, the output of the algorithm must be verified by the geotechnical engineer. If the process is done appropriately, an expert system may be designed to analyze the measurements done by the DMT test and in the real-time mode determine the

soil type at the particular depth. This is analogous to the manual usage of the soil profiles. To obtain such characteristics of the expert system, a generalization ability of the machine learning method must be present. It is the ability to correctly classify the soil categories based on the measurements which were not present in the learning data set. The latter was exploited before, for instance in the neural networks approach (Hegazy et al. 2002, Deka and Chandramouli 2006, Kang et al. 2006) or Bayesian networks (Mello and Brow 1999, Garbulewski et al. 2008).

The graph clustering algorithm is a simple and computationally efficient algorithm of grouping the measurement sets. It requires measurements taken at the

particular site, at different depths. Every set of measurements is further called an “example” or “experiment”. As the DMT test takes pressure measurements at every 0.02 m, the average training data set (covering about ten meters of depth) contains no more than five hundred examples (experiments). A set of measurements at every depth (in the DMT case various pressures) represents one of the soil categories. The algorithm reflects the example as the node of the graph. The edges are established between these nodes, which are similar regarding the measurements – belong to the same soil category. Because the number of the measured parameters can be different for various methods, the similarity must be assessed considering all of them. The idea of using the algorithm is presented in Figure 4, where measurements from

the DMT test (a) are nodes grouped into clusters (b).

Note that the algorithm only groups the experiments into clusters, labeling the similar ones with the same name. However, the geotechnical engineer must analyze the clusters and identify the particular soil types.

The learning data sets are matrices with four columns: ordinal number of the measurement, depth at which the measurement was taken and two measured parameters. In the presented DMT investigation, the latter are: pressure A and pressure B (in bars). The example of the learning data set is presented in Table 1.

The clustering algorithm analyses every experiment in the data set and assigns a label to it according to the similarity between the experiments.

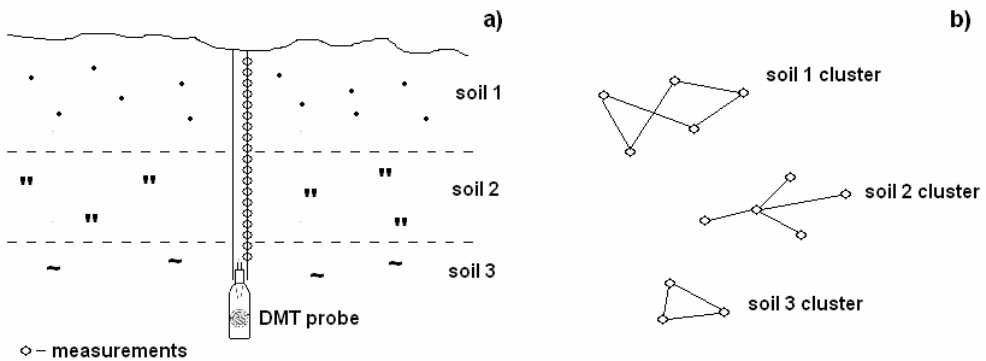


FIGURE 4. Illustration of the graph clustering application to the DMT measurements: data gathering (a) and result of the algorithm (b)

TABLE 1. Fragment of the learning data set for the clustering algorithm

Ident.	Depth [m]	Pressure A [bar]	Pressure B [bar]
1	0.3	2.0	11.0
2	0.9	2.0	11.0
3	1.7	2.4	10.1
4	2.3	3.5	11.3

Two different measurements can be considered similar if the difference between their values is smaller than the predefined threshold. Selection of the latter strongly influences the structure of the soil. Therefore an adaptive method of threshold selection was proposed. The algorithm is presented in Figure 5.

```

procedure graphclust
input  $DT$ 
begin
 $C \leftarrow \emptyset$ 
  for  $i=2$  to  $m-1$ 
     $s_i \leftarrow \text{findsimilar}(p_i)$ 
     $\text{add}(C, s_i)$ 
  end
   $c \leftarrow \text{gather}(C)$ 
end

```

FIGURE 5. Graph clustering algorithm

The algorithm works on the data set DT (here it is a table of geotechnical measurements). For each (i -th) column (containing one parameter, for example pressure A) a *findsimilar* procedure is executed. It performs a histogram analysis of all values in the attribute, assigning a category to the experiments that are close to each other, (to groups in the histogram). The result of the procedure is the set of partial categories s_i assigned to the measurements at the particular depths regarding i -th parameter. The partial categories are stored in the matrix C . They are transformed for every depth into one final category using the procedure *gather*. The same global category will be assigned to the depths, at which for every parameter the partial category was identical. For example, if for two depths (1.6 m and 2.8 m) the partial categories are, respectively, 3 and 6, these depths will be considered as containing identical

soil type. The exemplary histogram for the stamp is presented in Figure 6. Here, stamp's values form six clusters, which are numbered subsequently (from 1 to 6). The arrows indicate the calculated thresholds. The latter are values in the middle of the empty interval between the borders of two closest not empty intervals

(t_j^{\max} being the maximum possible value of the j -th not empty interval and t_{j+1}^{\min} – the minimum possible value of the $j+1$ -th non-empty interval) and all measurements inside the interval form one group, being a candidate for the separate soil type.

$$\theta_i = \frac{t_j^{\max} + t_{j+1}^{\min}}{2} \quad (1)$$

The key issue here is the number of the intervals used to divide the values of the measured parameters. It may affect the result of the profile generation – the larger number of intervals, the larger number of soil classes will be obtained. The value proposed for tests was $p/2$, where p was the number of the depths at which the measurements were taken. This way the method individually determines the similarity for each test site.

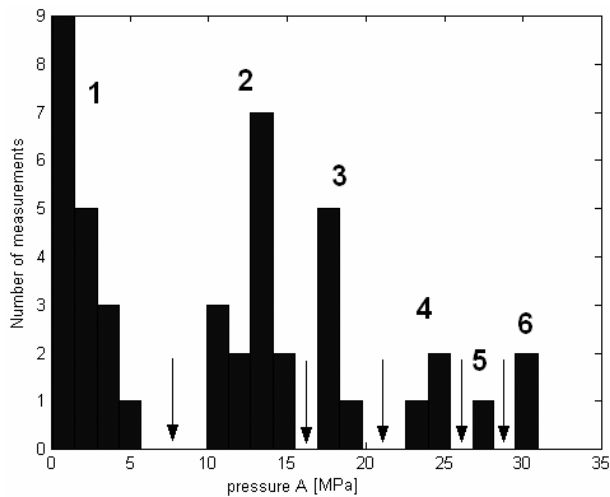


FIGURE 6. Example of the histogram for the measurement data with six clusters (categories) identified

The algorithm was implemented in MATLAB computing environment, which is sufficient for the research purposes. In the next stage of the research the computer program will be created using a high level programming language.

TEST RESULTS

The proposed algorithm was verified on the data obtained from the DMT test in four different points in the vicinity of the building No 34 at the campus (Fig. 2). Although test points are close to each other, the profile of the ground is not constant, but varies according to the processes shaping the form of the soil during the history (i.e. glaciation). Therefore there is a non-uniform ground structure, so the measurements even in the neighbor locations are different for the same soil type. These different data sets were used as the input data for the algorithm, which produces pairs: $\{depth,$

$category\}$ for every measurement. For the visualization purposes these are further drawn on the diagram to produce a soil profile. The aim of the experiment was to find the optimal threshold values, i.e. values that ensure the soil categorization consistent with profile obtained using DMT charts (used here as the reference values). In the following research the test sites were regarded as DMT1, DMT2, DMT3 and DMT4 (also Fig. 2).

The detailed results are presented for two sites, DMT2 and DMT3, which are the most representative and present two different situations in the soil structure. The measured parameters and their statistical distributions are presented in Figures 7 and 8. In most cases pressures can easily be separated into two groups, there are also some depths, at which there are more categories. The former indicate two main geological layers, i.e. sands and clays. However, the task of the engineer is to assess the characteristics of the geotechnical layers, which are more important for the building construction.

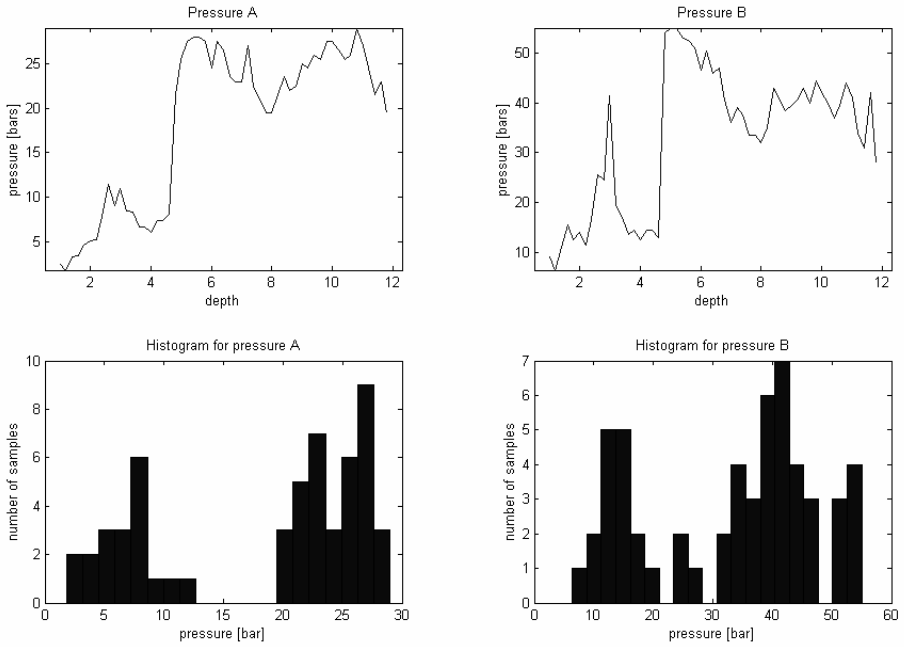


FIGURE 7. Measured parameters (pressure A and B) in the DMT2 test site and their histograms

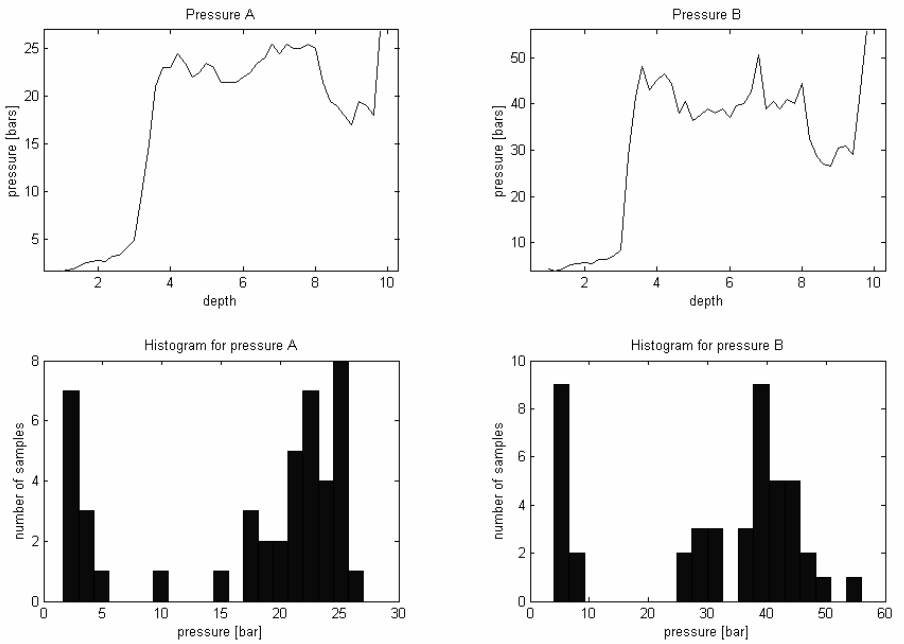


FIGURE 8. Measured parameters (pressure A and B) in the DMT3 test site and their histograms

In the DMT2 site, the diversity of the soil is greater, leading to multiple categories, with small numbers of measurements (Fig. 9a). The main soil formations are “1” and “2”, however there are also multiple additional categories in the transient area between them (“3”). The latter indicates that the soil structure is not uniform and in most cases is the mixture of two or more basic soils. Also, there is a small formation of single-element categories close to the surface (“4”), which is probably related to the existence of the object such as rock. These areas will manifest at the junction of main categories. The DMT3 site (Fig. 9b) is more uniform, as there are only three main categories, which relate to sands (“1”) and clays (“2” and “3”). The differences between the borders of geological layers obtained using DMT and borehole methods are below one

m., which is acceptable as the DMT measurements are taken every 20 cm and sites of DMT and borehole investigations are located at the distance of about ten meters. Results from other sites are similar. Figure 9 also contains geological layers, which have borders close to the obtained geotechnical layers. They are however, only for the visualization purposes, as the engineer will be mostly interested in the geotechnical layers.

The main conclusions drawn from the presented investigations are as follows. Geotechnical layers can be represented as the uniform categories only to some extent. They usually consist of some subcategories, which have separate physical characteristics. Also, additional categories emerge when two different geotechnical layers meet. Therefore another alternative to the discrete clustering (when there is the finite number

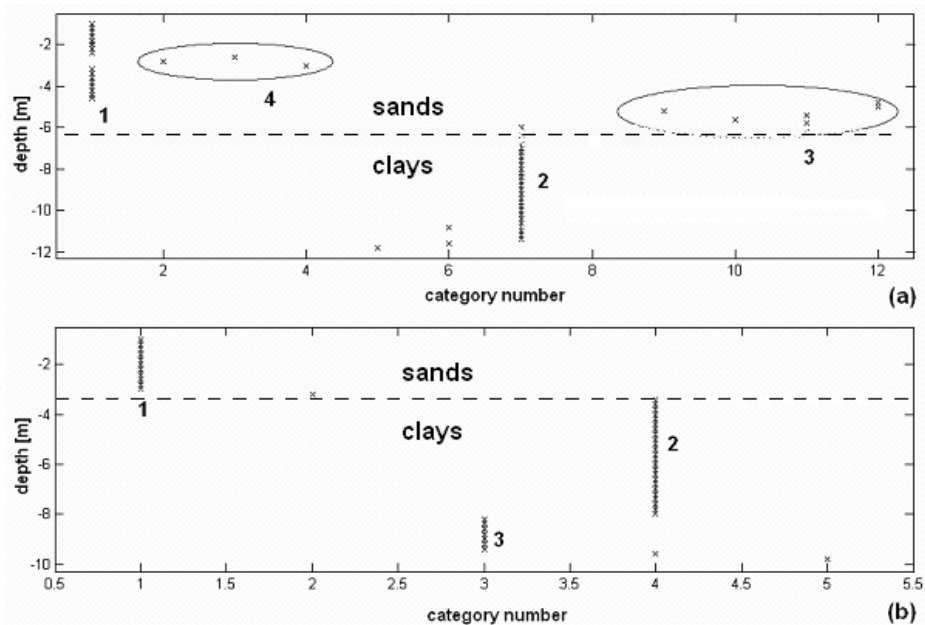


FIGURE 9. Soil profiles obtained using the algorithm for DMT2 (a) and DMT3 (b)

of easily distinguishable categories) is fuzzy clustering and classification – (Zhang and Tumay 1999), which shows more complex structure of soils.

Another problem is the locality of the measurements. Characteristic features of the particular soil layers strongly depend on their geographical location and history. For instance, the same soil behaves differently depending on the events it was part of in the past (such as existence of the glacier). Therefore the resented method of soil identification is probably applicable only in narrow parts of geographical locations. To verify this phenomenon more data are required from other locations than WULS.

Measurements taken for the same soil type in various (though close) locations are coherent, but the differences in ranges exist (especially for the fine sands). It is clear that the soil profile is different in every location, so the pressure ranges also must vary. Although the soil types are easily distinguishable from each other based on the two pressure measurements, to successfully identify other soil types, additional information may be required (for example, more parameters). The same is for distinguishing between the subcategories of the soil, for example brown and grey clays. This problem will be investigated further.

CONCLUSIONS AND FINAL REMARKS

The DMT soil analysis was performed at four different locations of the Warsaw University of Life Sciences. From every location, a different data set was constructed. Subsequently, an automated

graph clustering algorithm was used to divide experiments into soil categories. The critical operation, i.e. the threshold selection was implemented based on the histogram analysis. It allowed using the intuitive sense of similarity between two points (measurements at the particular depths). Then, the geotechnical engineer verified the algorithm output selecting the threshold assuring the best soil classification and identifying the categorized soils. The outcome of the algorithm is satisfactory as it is able to generate the categorization of geotechnical layers close to the one obtained using drilling boreholes. The differences are caused by the fact that measurements were taken every 0.02 m, which limits the accuracy of the measurement method. Also, the test sites of DMT investigations and boreholes were located at about ten meters from each other, which justifies differences in borders of soil categories up to one meter. The applied algorithm allowed gaining additional knowledge about the soil structure in the tested locations, revealing that the DMT charts have the limited accuracy.

REFERENCES

- BILSKI P., RABARIJOELY S. 2009: Automated soil categorization using the CPT and DMT investigations, Proc. 2nd Int. Conf. on New Developments in Soil Mechanics and Geotechnical Engineering ZM2009, Nicosia, Northern Cyprus, May 28–30 2009, pp. 368–375.
- DEKAP, CHANDRAMOULI V. 2006: Fuzzy neural network model for hydrologic flow routing. *Journal of Hydrologic Engineering*, 10(4), pp. 302–314.

- GARBULEWSKI K., RABARIJOELY S., JABLONOWSKI S. 2008: Bayesian approach in interpretation of CPT/DMT tests. Proc. of the 11th Baltic Sea Geotechnical Conference, Geotechnics in Maritime Engineering, Gdańsk, Poland, pp. 391–398.
- HASHASH Y.M.A., JUNGS., GHABOUSSI J. 2004: Numerical implementation of a neural network based material model in finite element analysis, International Journal for Numerical Methods in Engineering, 59, pp. 989–1005.
- HEGAZY Y.A., and MAYNE P.W. 2002: Objective Site Characterization Using Clustering of Piezocone Data, Journal of Geotechnical and Geoenvironmental Engineering, Vol. 12; pp. 986–996.
- KANG J.Y., CHOI B.I., LEE H.J. 2006: Application of artificial neural network for predicting plain strain fracture toughness using tensile test results, Fatigue & Fracture of Engineering Materials and Structures 29(4), pp. 321–329.
- LECHOWICZ Z., RABARIJOELY S. 2000: Numerical analysis of organic soil behaviour during dilatometer test. Annals of Warsaw Agricultural Univ. – SGGW, Land Reclam., No 29, pp. 41–49.
- LECHOWICZ Z., RABARIJOELY S. 2000: Dilatometer C-reading to help determine stratigraphy of organic subsoil. Annals of Warsaw Agricultural Univ. – SGGW, Land Reclam., No 29, pp. 71–78.
- LUTENEGGER A.J., CABIR M.G. 1988: Dilatometer C-reading to help determine stratigraphy, Proc. Int. Sym. on Penetration Testing ISOPT-1, Orlando, 1, pp. 549–553.
- MARCHETTI S. 1980: In Situ Tests by Flat Dilatometer. J. Geotech. Eng. Div., ASCE, 106, GT3, pp. 299–321.
- MARCHETTI S., CRAPPS D.K. 1981: Flat dilatometer manual. Internal report of GPE.
- MELLO K.L., BROWN S.D. 1999: Novel hybrid classification method employing Bayesian networks, Journal of Chemometrics, John Wiley & Sons, 13(6), pp. 579–590.
- RABARIJOELY S., BILSKI P., FALKOWSKI T. 2007: The usage of the graph clustering algorithm to the recognition of geotechnical layers, Annals of Warsaw University of Life Sciences – SGGW. Ann. Warsaw Univ. of Life Sciences – SGGW, Land Reclam., No 38, 2007, pp. 57–68.
- SHAHIN M.A., JAKSA M.B., MAIER H.R. 2005: Neural network based stochastic design charts for settlement prediction, Can. Geotech. Jour. (42): pp. 110–120.
- ZHANG Z., TUMAY M. 1996: Simplification of soil classification charts derived from the cone penetration test, Geotechnical testing Journal, Vol. 19, No 2, pp. 203–216.
- ZHANG Z., TUMAY M. 1999: Statistical to fuzzy approach toward CPT soil classification, Journal of Geotechnical and Geoinvernonmental Engineering, Vol. 125, No 3, pp. 179–186.

Streszczenie: *System automatycznej generacji profilu gruntowego na podstawie badań dylatometrycznych (DMT).* W artykule przedstawiono metodę automatycznej generacji profili gruntów z wykorzystaniem metody heurystycznej. Może ona zostać wykorzystana jako źródło wiedzy dla projektowanego modułu klasyfikacji gruntów, który będzie w stanie identyfikować warstwy gruntów na podstawie pomiarów geotechnicznych bez udziału użytkownika. Metoda wykorzystuje algorytm clusteringu grafowego, który grupuje pomiary uzyskane techniką DMT na poszczególnych głębokościach. Grupowanie odbywa się na podstawie pomiaru podobieństwa pomiędzy pomiarami. Zaliczane są one do tej samej kategorii (warstwy geotechnicznej), jeśli różnica pomiędzy ich wartościami jest mniejsza od wartości progowej. Ta ostatnia jest uzyskiwana automatycznie, na podstawie analizy histogramu. W artykule przedstawiono metodę pomiaru geotechnicznego, scharakteryzowano lokalizację (kampus SGGW), w której przeprowadzono eksperymenty oraz opisano algorytm generacji profilu. Przedstawiono wyniki eksperymentów w postaci uzyskanych profili gruntów oraz wyciągnięto wnioski na temat możliwych zastosowań metody w przyszłości.

Acknowledgments: This research was supported by Grant no 3951/B/T02/2008/35 from the Ministry of Science and Higher Education, Warsaw, Poland.

MS. received – November 2009

Authors' addresses:

Piotr Bilski
Katedra Zastosowań Informatyki
Wydział Zastosowań Informatyki i Matematyki
SGGW
02-776 Warszawa, ul. Nowoursynowska 159,
Poland
piotr_bilski@sggw.pl

Simon Rabarijoely
Katedra Geoinżynierii
Wydział Inżynierii i Kształtowania Środowiska
SGGW
02-776 Warszawa, ul. Nowoursynowska 159,
Poland
Simon_rabarijoely@sggw.pl