Proceedings of the 7th International Conference on Geotechnical and Geophysical Site Characterization
Barcelona, 18 - 21 June 2024

Enhanced, statistically-controlled integrated model of geo-data based on CPTU

Jędrzej Wierzbicki [1,2#], *Katarzyna* Stefaniak[1,2]

[1]*Adam Mickiewicz University in Poznań, Institute of Geology, ul. B. Krygowskiego 12, 61-680 Poznań, Poland*
[2]*SoilSpot sp. z o.o., ul. Uniwersytetu Poznańskiego 10, 61-614 Poznań*
[#]*Corresponding author: jwi@amu.edu.pl*

ABSTRACT

The paper presents the possibilities of geostatistical analysis of geotechnical data using cluster analysis. As a result of the analysis, geological data from boreholes were replaced by digital data corresponding to parameters measured during CPTU static sounding. A unified database with a depth resolution of 2 cm was obtained, allowing it to be used to build geostatistical geotechnical models, e.g. using kriging. Techniques providing statistical control of the homogeneity of geotechnical separations, based only on measured data and not on interpolated or correlated data, were used to create the model. The technique used makes it possible to create an integrated geo-engineering model of the subsoil even for small sites with limited data.

Keywords: data analysis, cluster, CPTU, statistical model.

1. Introduction

Geostatistical modelling is becoming an increasingly common practice in geotechnical studies. Modern modelling techniques allow advanced interpolation of data and the substitution of one piece of information for another, more useful in geotechnical analysis (Vanneste et al. 2022). However, the picture obtained from these methods cannot always be considered to be statistically controlled information to such an extent as to allow, for example, risk analysis (Wierzbicki et al. 2016). An additional issue also remains how the reliability of the model is affected by, for example, the density of the distribution of survey points and their abundance. This issue is of particular relevance with regard to subsoil investigations of linear structures, such as roads, or small structures where we have, for example, only four test points.

This paper presents the possibilities of geostatistical analysis of the subsoil data prepared for a small road structure - a bridge over a 10 m wide watercourse. Recognition of the subsoil properties was carried out by means of geological surveys and CPTU static soundings, with two static soundings and three test holes only. The analysis took the actual CPTU results and transformed them so that it was possible to replace the borehole profiles with CPTU profiles. This resulted in a unified database with a resolution of 2 cm depth increments, allowing it to be used for the construction of geostatistical geotechnical models, e.g. using kriging, or the IDW method (Młynarek et al 2007a). Techniques providing statistical control of homogeneity of geotechnical separations, based only on measured data and not interpolated or correlated, e.g. using artificial neural networks, were used to build the model.

2. Test site

The study area is located in the vicinity of the town of Suwałki in north-western Poland (Fig. 1). The object for which the reconnaissance was carried out is a bridge in the course of the DW662 road, crossing the small Szczerbenka river at this location, with a channel width of approximately 10 m. The river flows in a larger valley, left over from glacial processes during the last glaciation (Vistulian), which receded from the area about 12,000 years ago. The youngest layers are organic soils (peats), which lie in layers of varying thickness on fluvioglacial sands and overconsolidated glacial till. Due to the low bearing capacity of the organic layers and the high ground water level, as well as the foundation technology of the bridge abutments (piles), a compacted sand working platform with a thickness of up to 2 m was made prior to the survey. Both natural processes and engineering activities therefore meant that the soils in the subsoil were characterised by a pronounced degree of pre-consolidation at the time of the survey.

3. Tests carried out

The soil reconnaissance included two static soundings and 3 boreholes with sampling for laboratory tests (Fig. 1). One of the boreholes was located in the immediate vicinity of the CPTU sounding, allowing this point to be taken as a reference for further analyses.

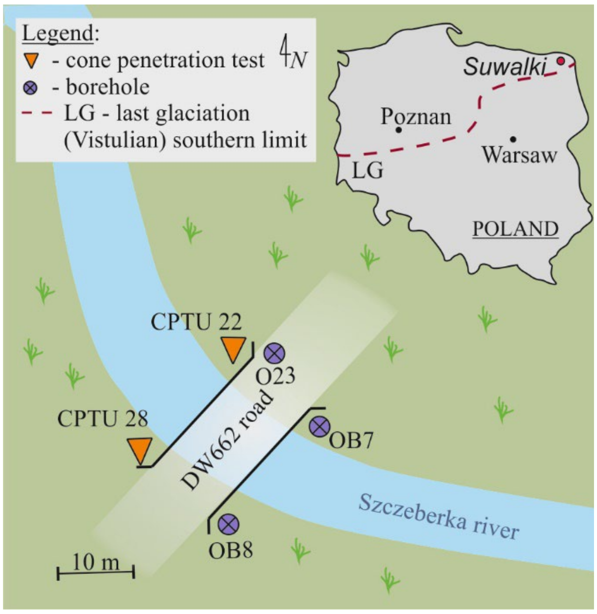

Figure 1. Location of the study site and sketch of the study area.

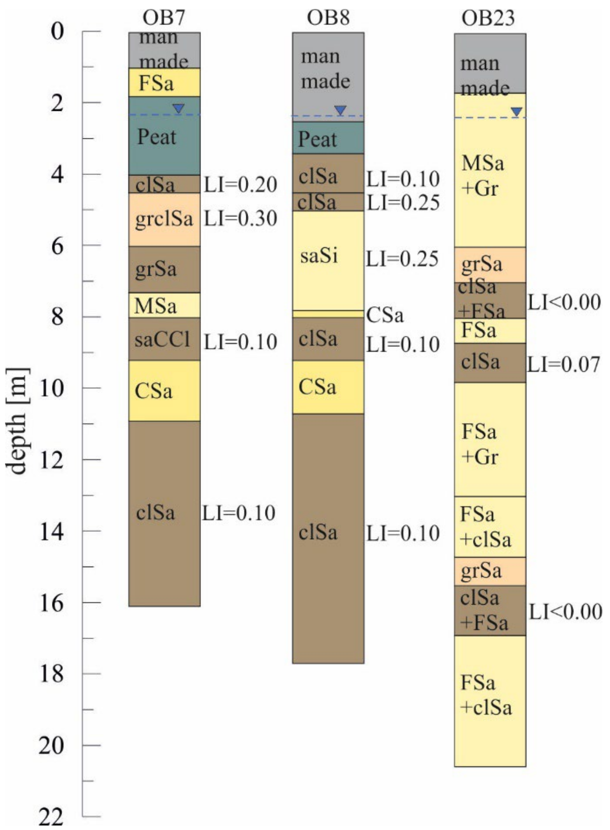

Figure 2. Borehole profiles at the study site (soil type symbols according to ISO14688-2, LI - liquidity index).

The borehole results confirm the general pattern of the geological structure, with the characteristic features of the individual series being a high content of gravel fraction in fluvioglacial sediments and a high content of sand and gravel fraction in glacial till (Fig. 2). The laboratory tests carried out indicate fairly low values for peat strength and deformation parameters and high values for these parameters in the case of glacial till (table 1). On the other hand, the interpretation of the CPTUs indicates that the glacifluvial sands are predominately dense (according to the interpretation proposed by Jamiolkowski et al. 2003) (Fig. 3).

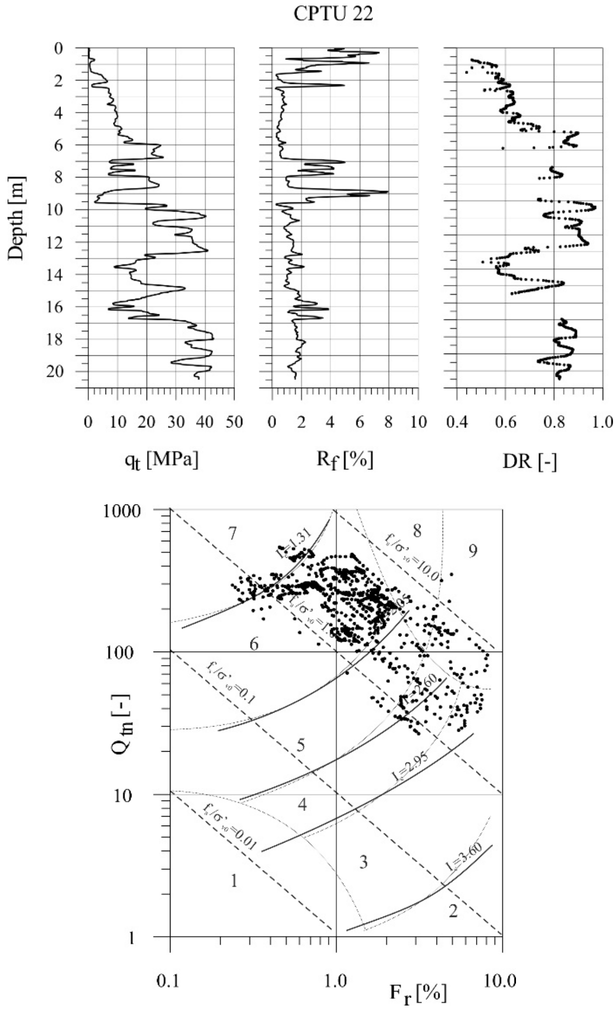

Figure 3. Results of the CPTU 22 with calculated density ratio (DR) (Jamiolkowski et al. 2003) and plotted in the Soil Behaviour Chart (Robertson and Cabal 2012).

4. Method

The relatively small amount of data at our disposal (1604 CPTU measurements - q_t, f_s, u_2, taken at a frequency of every 2 cm of depth in two profiles) and the drilling data, as basic information on the layout and type of layers, make it difficult to carry out an objective and statistically controlled analysis of the subsoil structure. A solution in such a case may be the creation of an integrated geoengineering model according to the concept of Wierzbicki et al. (2016). Such a model does not require full numerical information about the subsoil as an input - drilling data are replaced by CPTU data and can then be used to build a model of the distribution of any geotechnical parameter in the subsoil, using known interpolation techniques (Młynarek et al. 2013). The method involves the following procedural steps:

Step 1 - collect all CPTU data into a single set and cluster them using the cluster analysis method. In this step, it is also necessary to select the CPTU parameters that are included in the analysis. This results in a dozen or more solutions.

2nd step - selecting the most favourable solution, whereby the term "most favourable" may have different meanings. For this purpose, one can use the well known CH procedure (Caliński & Harabasz, 1974) (a purely statistical criterion) or the somewhat more intuitive method of Wierzbicki (2007), where one can control various aspects of the obtained solutions in terms of the statistical parameters of the distribution.

3rd step - substitutions of the data in the individual clusters into a depth-dependent functional form of the given CPTU parameter.

4th step - fitting the individual clusters and thus functions with specific forms to the drill profiles. In this step it is extremely helpful to have at least one reference point (CPTU + drilling).

5th step - replacing the drilling profile with the profile of the selected CPTU feature, resulting from the calculation of the value of the feature at the corresponding depth.

It should be noted that each cluster has its basic statistical parameters controlled during the analysis, such as the mean value, standard deviation, coefficient of variation, sample distribution or assumed confidence interval.

This has the advantage of resulting in a dataset that is statistically defined by the assumed (acceptable) variability of the selected parameter.

5. Results

The procedure outlined above was applied to the analysis of the collected data. In step 1, four leading parameters were selected against which clustering was performed:

$$q_t = q_c + (1-a)u_2 \tag{1}$$

where: a - the cone's surface area ratio, as specified by the manufacturer.

$$q_n = q_t - \sigma_{v0} \tag{2}$$

where: σ_{v0} – vertical geostatic stress state.

$$F_r = R_{fn} = \frac{f_s}{q_t - \sigma_{v0}} 100\% \tag{3}$$

$$Q_{tn} = \frac{q_n}{p_a}\left(\frac{p_a}{\sigma'_{v0}}\right)^n \tag{4}$$

where: p_a – atmospheric pressure, n – preconsolidation dependent exponent (for normally consolidated soil equal to 1).

The choice of parameters was guided by the use of Q_{tn} i F_r parameters in the general analysis of static sounding results (Robertson and Cabal, 2012) and the use of q_t i q_n parameters in the interpretation of soil geotechnical properties such as angle of internal friction ϕ' or compressibility modulus M (Lunne et al. 1997). Thus, it was assumed that the determined homogeneous groups of measurements, would characterise soil layers similar in terms of geotechnical properties. The values of these parameters were standardised in order to avoid excessive influence of any of them on the clustering results, although, of course, it is possible to select any of the parameters as the leading one at this point of the analysis. Clustering was carried out using the k-means method (Młynarek et al. 2007b), applying a 50-fold solution iteration. This resulted in 15 solutions, ranging from a split into five clusters to a split into 20 clusters.

The k-means method on its own does not indicate the optimal solution and requires an independent evaluation, which is carried out in step 2 of the analysis. In the present example, the weighted average coefficient of variation method (Wierzbicki 2007) was chosen, introducing a certain modification to it. In the original method, the average coefficient of variation of the selected parameter is determined in relation to the total number of data:

$$CV_{av} = \frac{\sum_{1}^{n}(l_n CV_n)}{l} \tag{5}$$

where: CV_{av} – weighted average coefficient of variation of the parameter in the profile in the next step of the analysis, n – number of clusters separated, l_n – number of data grouped in the n-th cluster, l – total number of data, CV_n - coefficient of variation of the parameter in the n-th cluster.

In the modified method, it is the mean coefficient of variation of the product of the number of clusters and the total number of data

$$CV1_{av} = \frac{\sum_{1}^{n}(l_n CV_n)}{nl} \tag{6}$$

where: $CV1_{av}$ – weighted average coefficient of variation of a parameter in a profile relative to the number of clusters.

While the CV_{av} value indicates whether the next step of the analysis (generating 1 more cluster) caused a change in the average coefficient of variation in the whole set of clusters (the lower the variability of a given solution, the better), the $CV1_{av}$ value allows us to observe to what extent a decrease in the variability of the solution is only due to an increase in the number of clusters. This is important because intuitively we are looking for the solution that, with the smallest possible number of clusters (in effect geotechnical layers), will give the most statistically homogeneous clusters. The degree of change of $CV1_{av}$ can be conveniently defined as the gradient of the graph line between two consecutive points, defined by successive cluster analysis solutions. Such a parameter is defined in this analysis as ${}_{grad.}CV1_{av}$.

The results presented for solutions in the range of 8 to 20 clusters (Fig. 4) show that the largest decrease in CV_{av}

is observed until 12 homogeneous groups (clusters) are separated. Thereafter, the value of CV_{av} no longer changes significantly from step to step of the analysis. However, when assessing the $CV1_{av}$ value (Fig. 5), it can be seen that the effect of increasing the number of clusters can still have a significant impact on the decrease in variability up to around step 14 of the analysis (considering the values of Q_{tn} and F_r).

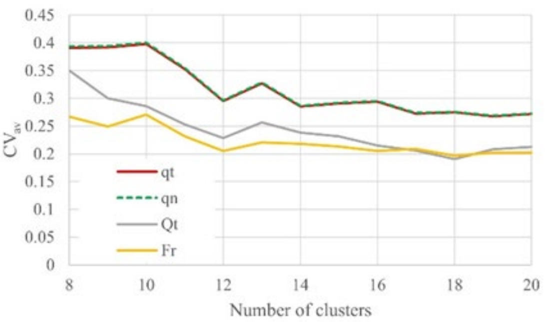

Figure 4. Changes in CV_{av} values of individual CPTU parameters in relation to the increase in the number of separate clusters.

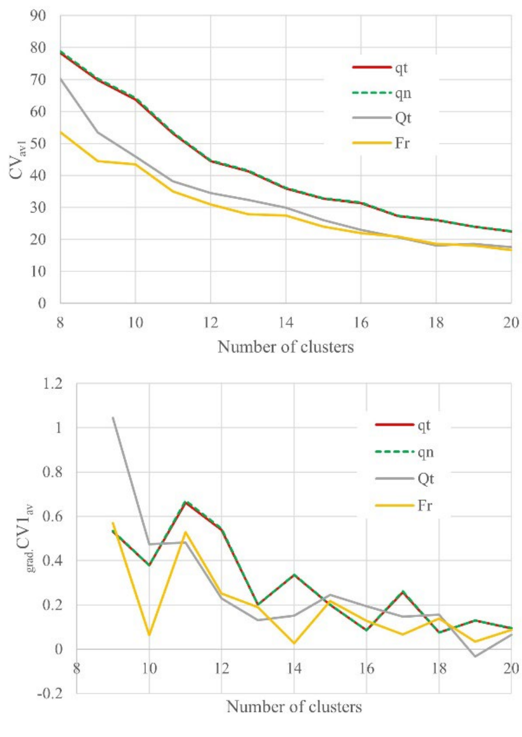

Figure 5. Changes in $CV1_{av}$ values and $_{grad.}CV1_{av}$ of individual CPTU parameters in relation to the increase in the number of separate clusters.

Following these indications, solution 14 was adopted as the most effective. The separated clusters are presented in the soil behaviour chart (Robertson and Cabal 2012) (Fig. 6) and their statistical parameters are included in the diagrams (Fig. 7).

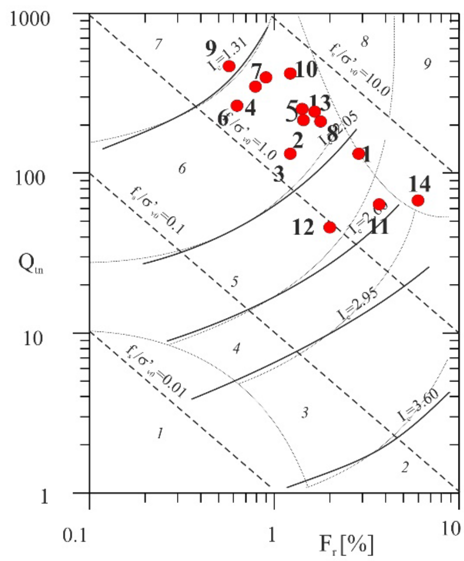

Figure 6. Location of individual clusters in the Soil Behaviour Chart (Robertson and Cabal 2012).

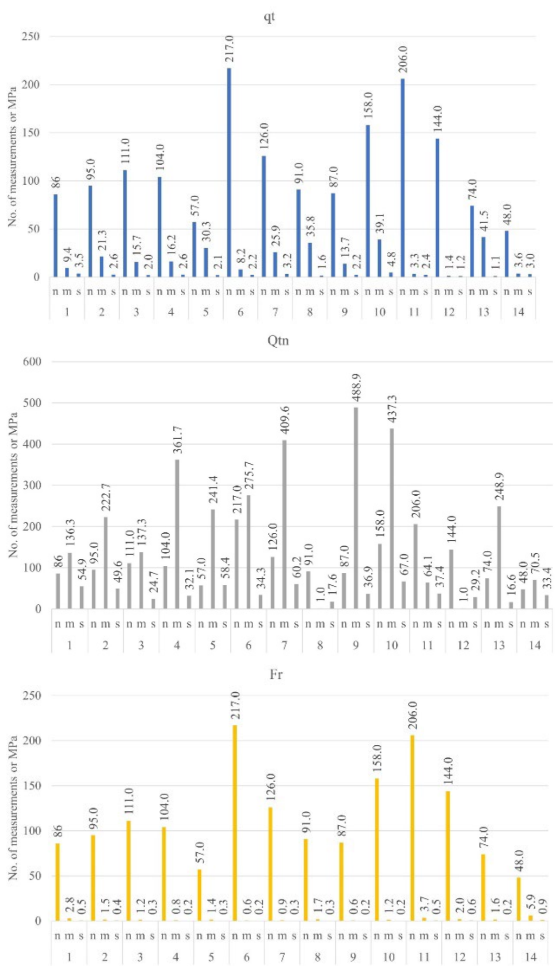

Figure 7. Basic statistical parameters of the separated clusters (from 1 to 14): n – number of data in cluster, m - mean, s - standard deviation, for the parameters: q_t, Q_{tn}, F_r.

Subsequently, for each CPTU parameter in each cluster, the function of the dependence of its value on the depth of measurement z was determined. The form of a power polynomial function with a degree from 4 to 10 was adopted (Fig. 8).

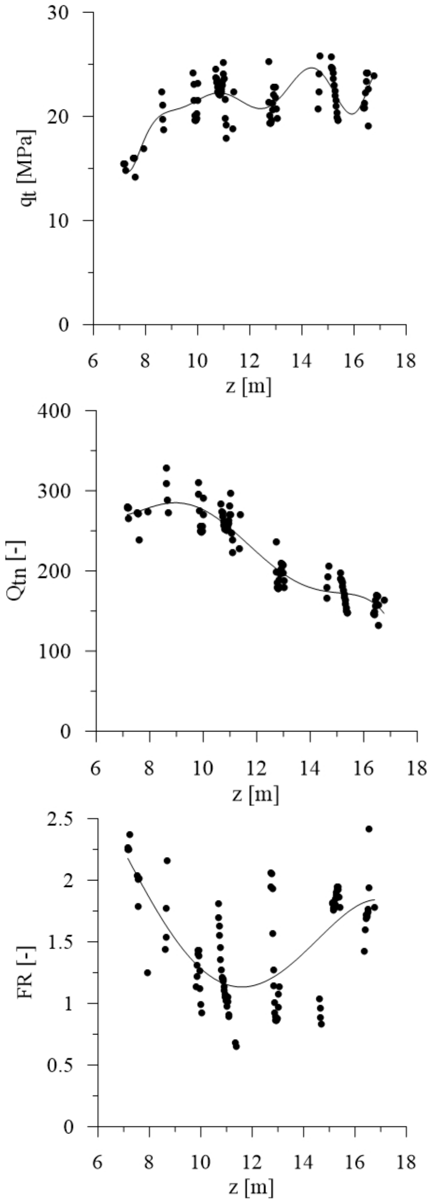

Figure 10. Polynomial functions approximating the values of q_t, Q_{tn}, F_r of the data of cluster No. 2 relative to depth z.

Next, the assignment of individual CPTU clusters, along with the functions of each parameter, was made to the layers separated during geological drilling. For this purpose, lithological and *LI* and depth information were used, but mainly data from the reference node (borehole OB23 and CPTU OB22). Since the parameter functions depend only on the depth z, this allowed us to obtain a continuous picture of the changes in the CPTU parameters in the drilling profiles (Fig. 11).

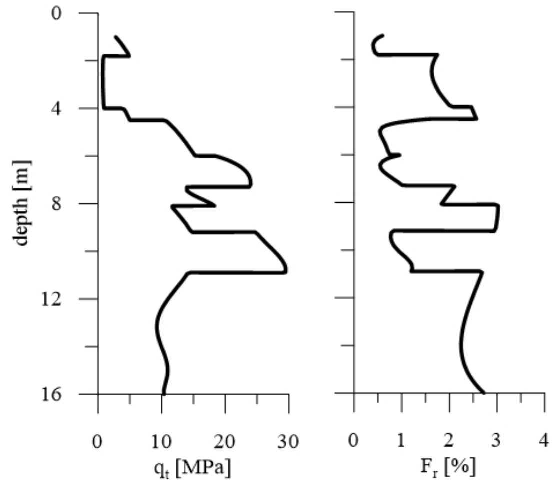

Figure 11. An example of the OB7 borehole profile converted to CPTU data: q_t and F_r.

The effectiveness of the adopted methodology can be indirectly seen by analyzing the artificially generated static sounding based on the OB 23 borehole profile against the nearby CPTU 22 (Fig. 12). The artificially generated CPT results deviate in precision from the real ones (especially in the case of F_r values) but can be considered to reflect well the values of CPTU parameters and the general trends of changes in soil properties with depth.

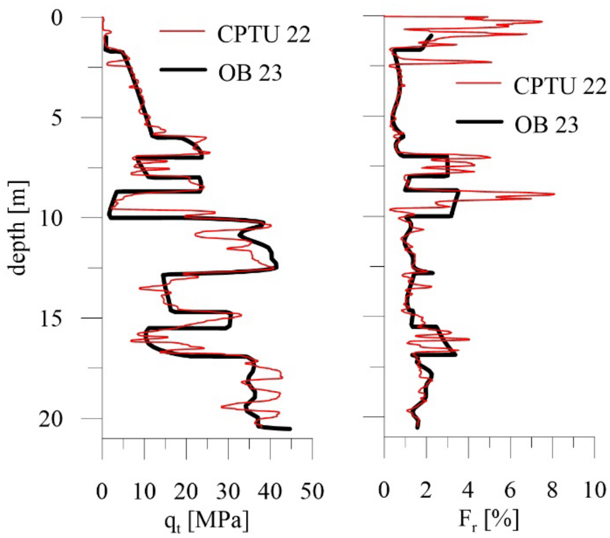

Figure 12. The OB23 borehole profile converted to CPTU data: q_t and F_r, on the background of the nearest CPTU 22 results.

The results of the analysis can be used to create a geostatic model using any interpolation method, for example IDW (inverse distance to the power) as described by Młynarek et al. (2007) (Fig. 13).

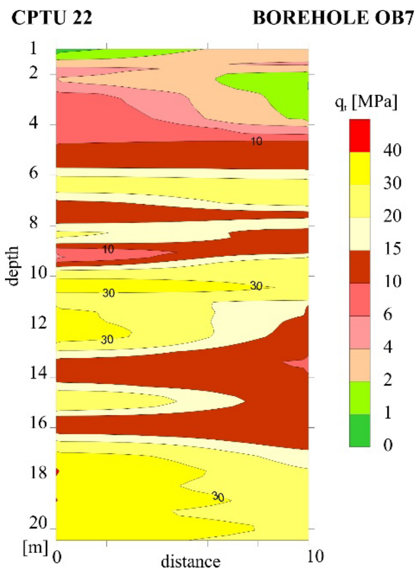

Figure 13. Geostatistical model of q_t for analysed example, calculated by the method inverse distance to the power.

6. Discussion

The resulting picture of changes in CPTU parameters in the borehole profiles is obviously different from the real sounding results. Particularly noteworthy is the absence of a so-called transition zone between geotechnical layers (i.e. a relatively smooth transition from one layer to another) (Robertson and Fear 1995). On the other hand, such a solution does not pretend to be anything, in particular a true CPTU sounding, it is merely a prospecting of the statistical parameters of the distribution of the selected sounding parameter on the drilling results. Thus, as long as we have correlated the clusters and geological layers from the drilling well, we can obtain data in a statistically controlled manner corresponding to the average CPTU results in the area.

Another problem that arises when converting borehole data to CPTU results is the effect of the transition zone, at the boundary between lithological layers. This issue has been written about by Robertson and Fear (1995) or Boulanger and DeJong (2018), among others, who attempted to determine the "true" values of cone resistance in a given layer, unabated by the influence of neighboring layers. In this context, it should be noted that, although the proposed method eliminates the problem of the transition zone (lithological boundaries unambiguously affect the abrupt change in the results of artificial CPT), but at the same time provides parameter values that are in a sense averaged over a given group of layers.

7. Conclusions

The presented analysis methodology is a valuable alternative to the classical approach to geotechnical and geological-engineering analyses, carried out with limited resources and research possibilities, thus for small projects.

The method offers the possibility to use the generated data together with the real data, e.g. to delineate geotechnical layers in a statistically controlled manner, e.g. assuming a specific confidence level of a geotechnical parameter (as e.g. in the Instruction of the General Directorate for National Roads and Motorways).

In cases of small construction sites, and a limited number of test profiles, the method presented makes it possible to create a subsoil model using statistical interpolation techniques. It is worth noting that in the example given, without the use of the data transformation method, it would not have been possible to build a geostatistical model based on just one CPTU sounding.

Regardless, it represents a slightly different approach to generating artificial digital data than the increasingly common machine learning methods. The undoubted advantage of this approach is the possibility of continuously controlling the statistical distribution parameters of the parameters under study. In this context, it can be a valuable alternative to big data methods.

References

Caliński, T. and J. Harabasz, 1974. „A dendrite method for cluster analysis.“ Communication in Statistics, 3: 1-27.

Boulanger R.W. and J.T. DeJong, 2018. „Inverse filtering procedure to correct cone penetration data for thin-layer and transition effects.“ In Cone Penetration Testing 2018 edited by Hicks, Pisano and Peuchen. CRC Press Taylor & Francis Group, 25-44.

Jamiolkowski M., D. C. F. Lo Presti, M. Manassero. 2003. “Evaluation of Relative Density and Shear Strength of Sands.” In Symposium on Soil Behavior and Soft Ground Construction Honoring Charles C. "Chuck" Ladd, 2003, 1-37.

Lunne T., P.K. Robertson, J.J.M. Powell. 1997. Cone Penetration Testing in Geotechnical Practice. Blackie Academic EF Spon/Routledge Publishers, New York.

Młynarek, Z., J. Wierzbicki and W. Wołyński. 2007a. „An approach to 3D subsoil model based on CPTU results.“ In Geotechnical Engineering in Urban Enviornment edited by V. Cuellar et. al. Vol. 3, 1721-1726. Millpress, Rotterdam.

Młynarek, Z., J. Wierzbicki and W. Wołyński. 2007b. „Efficiency of selected statistical criteria in determination of geotechnical parameters from CPTU.“ Studia geotechnica et Mechania. vol. 29, No. 1-2, 2007, 137-149.

Robertson, P. K. and Fear, C. E. 1995. „Liquefaction of sands and its evaluation.“ In Proc., 1st Int. Conf. on Earthquake Geotechnical Engineering, edited by K. Ishihara, A. A. Balkema.

Robertson P.K. and K.L. Cabal. 2012. Guide to Cone Penetration Testing. Greg Drilling & Testing, Inc., Signal Hill, California.

Vanneste, M., G. Sauvin, J.-R.Dujardin, C.F. Forsberg, R.T. Klinkvort and Ragnhild R.C. 2022. “Data-Driven Ground Models: The Road to Fully-Integrated Site Characterization and Design.” In *VSOE 2021* edited by C. HansenD. V. K. Huynh et al., LNCE 208, 3–2.

Wierzbicki J., A. Smaga, K. Stefaniak and W. Wołyński. 2016. “3D mapping of organic layers by means of CPTU and statistical data analysis.” In *Geotechnical and Geophysical Site Characterisation 5* edited by Lehane, Acosta-Martínez and Kelly, Australian Geomechanics Society, Sydney, Australia, 1481-1486.

Wierzbicki, J. 2007. „Determination of homogenous geotechnical layers in strongly laminated soil by means of CPTU and cluster analysis.“ In Geotechnical Engineering in Urban Enviornment edited by V. Cuellar et. al., Vol. 5, 575-579. Millpress, Rotterdam.