Budapest University of Technology and Economics
Faculty of Civil Engineering
Pál Vásárhelyi Doctoral School of Civil Engineering and Earth Sciences

Statistical analysis of soil parameters in geotechnical design

István Kádár M.Sc.

Theses of Ph.D. dissertation

Supervisor:
Nagy László Ph.D.

Budapest, 2018.

1. INTRODUCTION

The reliability of the results provided by geotechnical design primarily depends on the input soil parameters used in the calculation. Estimation of these parameters is one of the critical issues in geotechnical engineering. Natural soils have high heterogeneity, especially compared to artificial building materials. The soil behavior may differ even in a small area, and there is significant variety within a single layer which is considered homogeneous.

Applications based on probabilistic calculations have a greater emphasis in geotechnics with the introduction of Eurocode 7 standard. The accurate knowledge of the soil properties is a principle by reliability-based design. In order to provide characteristic values – as input parameters of design – series of experiments are carried out. The amount of data available for statistical analysis varies by tasks with different magnitude. For smaller design projects, for large industrial investments or for controlled laboratory testing, the set of data is different in the analysis. Therefore, it is advisable to choose the degree of statistical processing and to determine the acceptable uncertainty based on the available set of data.

In most geotechnical problems probably the most important soil properties are shear strength and shear strength parameters. Shear strength is defined by the complex laws of nature and its knowledge is required for most geotechnical tasks. From the perspective of design and control the number of experiments, the mean and standard deviation of the data have to be known. The precision of the determination of shear strength and shear strength parameters has a direct effect on the characteristic values, used in the calculation, which has an influence on the principle of economy. The low number of measurements can negatively affect the safety of the structure. Too many measurements lead to uneconomical design. To resolve this issue, the deeper knowledge of different parameters' coefficient of variation provides a basis. The coefficient of variation (relative standard deviation) has a major role during the statistical evaluation process of soil properties. Uncertainty of a soil parameter can be described with the coefficient of variation. Based on its value different measurements from different sites can be compared. Neither a low number, nor unnecessarily many experiments are satisfying due to excessive uncertainty or diseconomy. To resolve this problem, the knowledge of coefficient of variation is a powerful tool.

Detailed calculations can be performed with the knowledge of the variability of input design parameters. During my research I examined the bearing capacity of shallow foundations and stability of flood protection dikes on probability basis. During the development of geotechnics, various safety considerations have evolved and spread. The development in theoretical approaches resulted in more and more sophisticated safety ideologies and calculation methods. The safety level of these methods compared to each other is nontrivial. The comparison of different methods' and different standards' global safety is possible based on calculating the probability of failure for all cases.

In my research, emphasis is put on examining the goodness-of-fit of different distributions of soil parameters, which I believe will soon become a routine task in civil engineering practice. The growing IT capacity and the user-friendly software products with statistical tools ensure the background.

2. OBJECTIVES AND CONTENTS OF THE DISSERTATION

2.1. THE REASON FOR THE CHOICE OF THE TOPIC AND OBJECTIVES

In Hungary, the use of probability-based solutions in the field of civil engineering, and particularly in the field of geotechnics, is undervalued. The importance and application of probabilistic theories have more emphasis since the introduction of MSZ EN 1997-1: 2006 (Eurocode 7) standard by 1 December 2006. In reliability-based design it is principal to determine the input soil parameters as precisely as possible. Therefore, we carry out experiments and series of experiments and identify characteristic values with the use of statistical tools. The statistical processing and evaluation of available data have become an engineering task. The amount of data varies considerably in different tasks. For routine design applications, for large industrial investments or for a controlled series of laboratory experiments, the data set for analysis is different.

The probability of failure is an input parameter for risk calculation. With calculating the probability of a failure, we can characterize a whole system, while the deterministic factor of safety values refer only to one part of the system. The safety of different structures and different calculation methods can be compared based on computing the probability of failure.

2.2. Contents

In the dissertation, after the introduction of the topic and the description of the methodology, I show my results in four different chapters:

- the coefficient of variation of the soil parameters and recommendation of a new classification method with the possible limits of the given intervals;
- calculating the probability of failure in case of bearing capacity of shallow foundations and comparison of the different methods (Eurocode 7, MSZ 15004-89/2.3.1, Brinch Hansen and Meyerhof);
- comparison of different standards' (Eurocode 7 and MSZ 15292: 1997) safety level by calculating the probability of failure in case of flood protection dikes;
- analysis of the distribution of certain soil parameters and the effect of the applied distribution function on the characteristic value.

3. DATA SETS AND APPLIED METHODS

The results presented in the dissertation are primarily based on three large data sets. These data sets are my own series of laboratory experiments, the Ajka data set, and the Paks data set.

During my laboratory research, I performed direct shear tests on five different types of soil (sand, clayey sand, sandy silty clay, high plasticity clay and fly ash) with two different vertical loads (100 kPa, 200 kPa), which means a total of 300 tests.

The Ajka data set includes the results of shear strength experiments carried out during the examination of the red mud catastrophe. A total of 60 shear strength experiments were collected for four different soil layers (gravel, clayey silty sand, high plasticity clay, fly ash).

The Paks data set includes laboratory test results of 39 boreholes and 39 CPTs. Laboratory tests mean altogether 5192 measurements, of which I focused on the experiments related to shear strength. Sufficient number of experiments of shear strength were made in two different types of sand layer (mainly direct shear tests, but also triaxial shear tests). The examination of the CPT results is mainly related to these two sand layers, complemented with examination of the gravel layer measurements. The total length of the CPT soundings was 1338 m, the minimum depth was 23.0 m and the maximum depth was 40.2 m, while the average was 34.3 m.

During the analysis of the data sets, basically Microsoft Excel with @RISK statistical and simulation extension helped my work. Modelling of flood protection dikes was carried out with the use of GeoStudio 7.10. – GeoSlope module and Plaxis finite element software. The probability of failure for shallow foundations was calculated in @RISK environment.

The value of coefficient of variation has a major role in the dissertation. The coefficient of variation (C_v)[1] is the ratio of standard deviation (s) to the mean value ($\bar{x}$):

$$C_v = \frac{s}{\bar{x}}$$

If samples taken from the data set are continuously examined, and for a certain *k* number of samples the mean value and the standard deviation for *k + 1, k + 2, ...* are calculated, we find that the value of C_v is less and less fluctuating. This is called the stabilization tendency. The coefficient of variation has a major role in the statistical evaluation of soil parameters and geotechnical problems. Besides the stabilization tendency, with the help of its value the different parameters of a soil can be characterized. Values for coefficient of variation from research studies all over the world showed that their value for the same parameter is within a relatively narrow range. It has the advantage of having reliable data on the examined parameter even before before the exploration or laboratory testing.

For the analysis of different soil parameters' distribution functions and to determine the goodness-of-fit, I used the Kolmogorov-Smirnov test besides the Pearson classification.

[1] Other notations: COV or v_x.

4. THESES

THESIS 1

The coefficient of variation is a good indicator by the determination of soil properties. The different properties of a soil have different coefficient of variation values. However, the same soil property of different soils cannot be characterized with one certain coefficient of variation.

Based on my own laboratory experiments and data based on Hungarian and international literature references I defined 5 quality classes for the accuracy of determination of shear strength and shear strength parameters for different soils.

These quality classes are presented in Table 1 and Figure 1. For example, in case of sand soil, if the coefficient of variation is below 0.08, it refers to Class 1, [0.08-0.15[interval means Class 2, [0.13-0.25[interval means Class 3, [0.25 -0.40[intervals means Class 4, and Class 5 refers to a coefficient of variation 0.40 or higher.

Class	C_v (φ)	C_v (c)
Class 1	< 0.08	< 0.3
Class 2	[0.08-0.15[	[0.30-0.40[
Class 3	[0.15-0.25[	[0.40-0.50[
Class 4	[0.25-0.40[	[0.50-0.60[
Class 5	0.40 ≤	0.60 ≤

Table 1. Quality classes offered for internal friction angle and cohesion

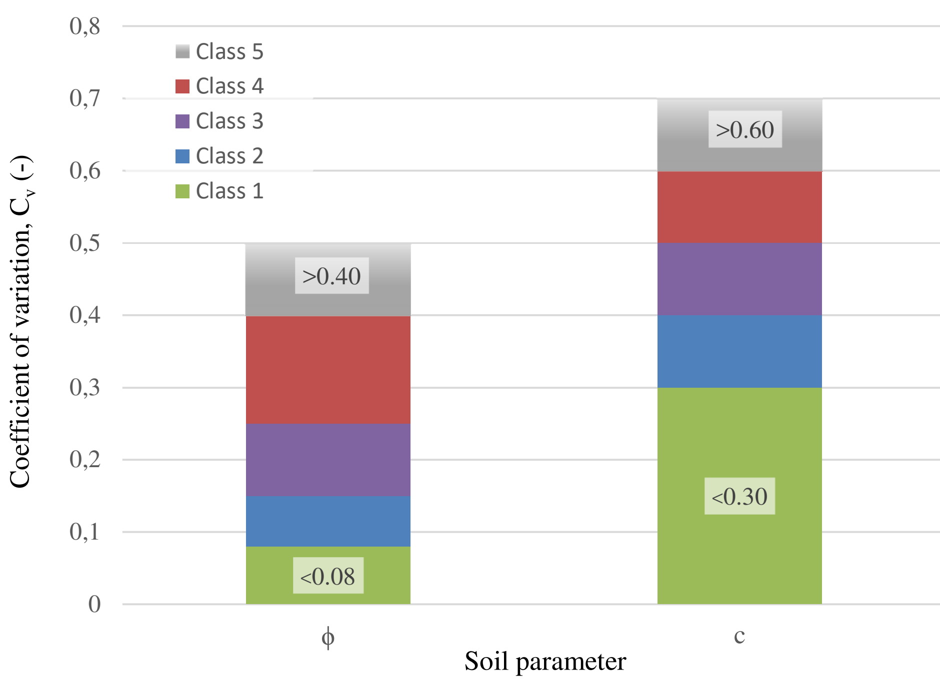

Figure. 1. Quality classes offered for internal friction angle and cohesion

Publications related to the thesis: I. Kádár, L. Nagy, A. Takács (2010a), I. Kádár (2013), I. Kádár (2014).

THESIS 2

I examined the change in the coefficient of variation value with increasing experimental number.

Based on CPT results, I found that the increase in the number of measurements for a soil parameter even beyond any reasonable limit does not result in convergence of the coefficient of variation to zero, but convergence is observed. The coefficient of variation value to which the soil parameter converges by increasing the number of measurements is called the fundamental uncertainty of the examined soil parameter for the given number of measurements.

The thesis can be illustrated by CPT tip resistance measurements, due to the available large data set. Figure 2 shows the coefficient of variation values of CPT tip resistance values measured in Pleistocene medium sand layer determined based on different sample numbers (n = 10 pcs, n = 100 pcs, n = 1000 pcs, n = 10467 pcs), (for different soils see Table 2). Each point related to a certain sample number refers to the coefficient of variation of a randomly selected n-element sub-set of the whole set of data.

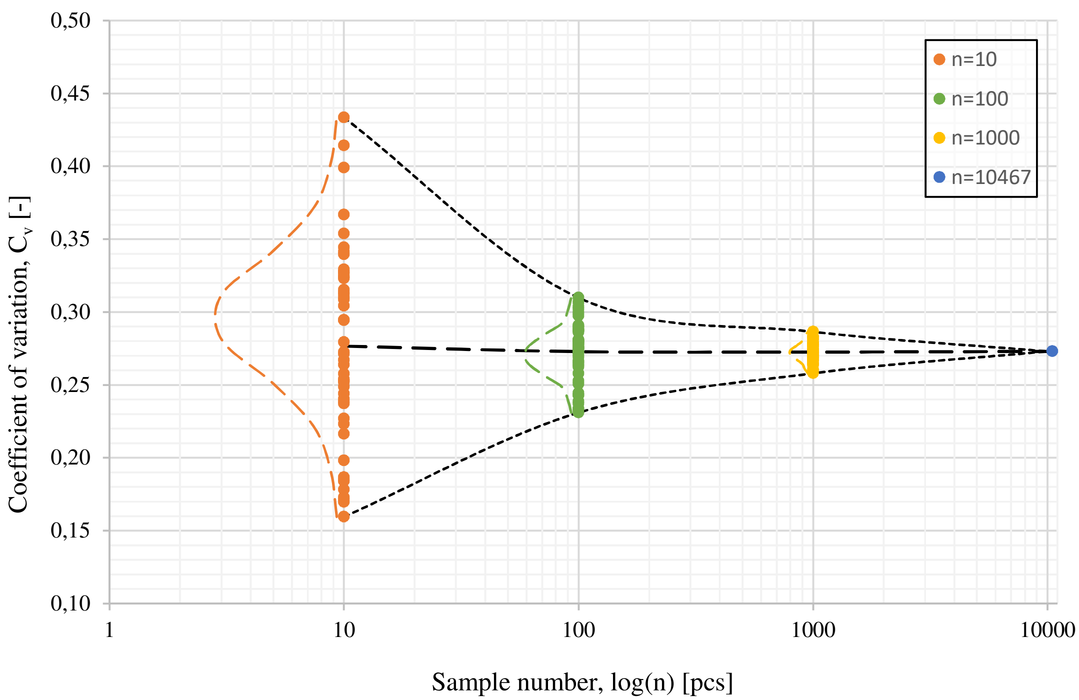

Figure. 2. Coefficient of variation values determined based on different sample numbers in case of Pleistocene medium sand layer

Soil	Sample number (pcs)	Fundamental uncertainty (C_v)
Eolian fine sand	9195	0.318
Pleistocene medium sand	10467	0.273
Gravel	18709	0.441

Table 2. Fundamental uncertainty of different soils' CPT tip resistance

Publications related to the thesis: I. Kádár, L. Nagy (2016).

THESIS 3

I determined the shear strength and shear strength parameters of sand, clayey sand, sandy silty clay, high plasticity clay soils and fly ash (Fig. 3) with the use of direct shear tests.

A large number of laboratory experiments were carried out with direct shear tests to determine the shear strength and the shear strength parameters. I found that the coefficient of variation of shear strength (τ) is smaller than the coefficient of variation of internal friction angle (φ) and cohesion (c) derived from the same tests.

It means that with the processing method additional uncertainty is created. I suggest the use of shear strength values instead of using the shear strength parameters which are affected by excessive error.

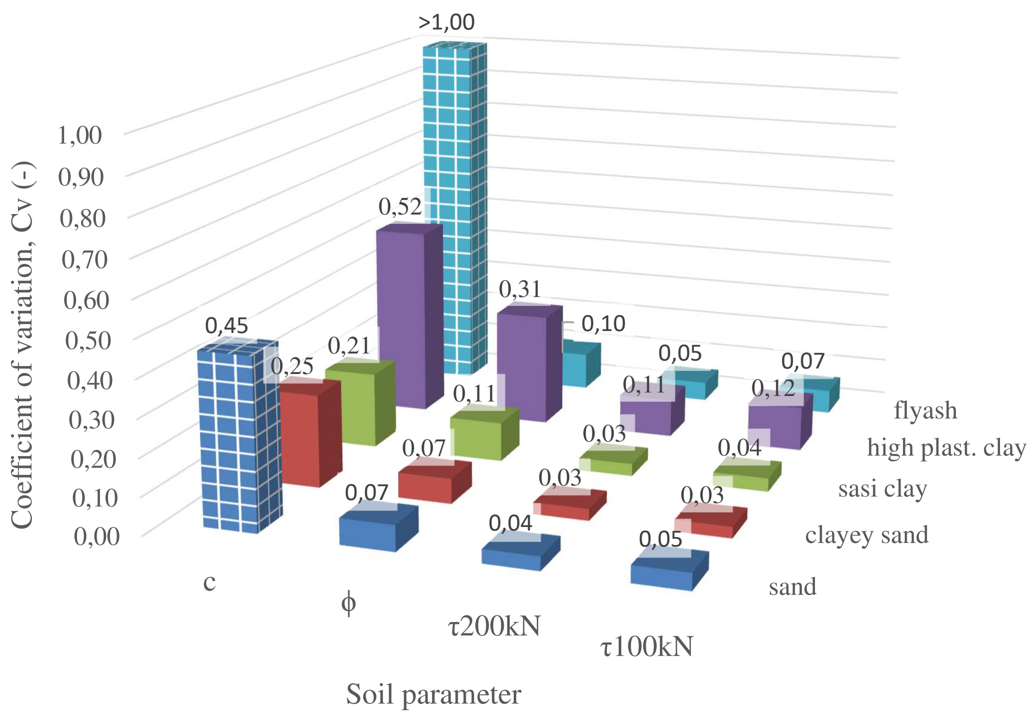

Figure 3. Coefficient of variation values of shear strength and shear strength parameters in case of different soils

Publications related to the thesis: I. Kádár, L. Nagy (2010a), I. Kádár, L. Nagy, A. Takács (2010b), I. Kádár, L. Nagy (2012a), I. Kádár, L. Nagy (2012b), I. Kádár, L. Nagy (2014).

THESIS 4

The conventional slope stability expressed by the factor of safety does not give information about the reliability. MSZ 15292: 1997 and MSZ EN 1997-1: 2006 standards also have a difference in the required safety (partial) factors (Table 3). Consequently, the two standards' deterministic safety factors do not indicate which standard provides a higher level of security. I determined the probability of failure based on the two standards' different input parameters.

Based on the examination of slope stability of flood protection dikes I determined that the different standards' levels of safety can be compared based on computing the probability of failure.

The calculation example showed that the MSZ 15292: 1997 standard gives the higher security level (P_f = 0.0056 for design flood level). Calculations based on MSZ EN 1997-1: 2006 have a higher probability of failure (P_f = 0.0252 for design flood level). The ratio does not change if the water level equals the crest level. The partial factor (1.25) in the international code is different than the partial factor in the Hungarian National Annex (1.35). With partial factor of 1.25, P_f = 0,0825 is the calculated probability of failure.

Standard	Safety factor		Input data
	design flood level	water level at crest	
MSZ 15292:1997	1.50	1.30	mean value
MSZ EN 1997-1:2006 (EU)	1.25	1.25	characteristic value
MSZ EN 1997-1:2006 (HU)	1.35	1.35	characteristic value

Table 3. Safety (partial) factors and input data for different standards

Publications related to the thesis: I. Kádár, L. Nagy (2015a), I. Kádár, L. Nagy (2015b), I. Kádár, L. Nagy (2017a).

THESIS 5

Four calculation methods were investigated to determine the bearing capacity of shallow foundations (Meyerhof, Brinch Hansen, MSZ 15004-89 / 2.3.1, MSZ EN 1997-1: 2006. In Meyerhof's method the bearing capacity factors are not available in closed formula.

Based on the examination of different calculation methods for bearing capacity of shallow foundations I determined that the different methods' levels of safety can be compared based on computing the probability of failure.

Calculation results (Tables 4-5) show significant differences, not only in the calculation methods, but also in terms of which shear strength parameters are considered as probability variables. The highest probability of failure is always calculated according to the MSZ EN 1997-1: 2006 standard.

Calculation method	Probability variable		
	φ	c	φ és c
MSZ EN 1997-1:2006	1.824%	0.018%	3.358%
MSZ 15004-89/2.3.1	0.074%	≈ 0	0.247%
Brinch Hansen	0.524%	≈ 0	1.131%
Meyerhof	2.020%	≈ 0	3.060%

Table 4. Probability of failure values (silty sand; 500 000 simulations)

Calculation method	Probability variable		
	φ	c	φ és c
MSZ EN 1997-1:2006	0.226%	0.512%	2.459%
MSZ 15004-89/2.3.1	0.204%	0.009%	0.484%
Brinch Hansen	0.135%	0.111%	1.238%
Meyerhof	≈ 0	0.033%	1.180%

Table 5. Probability of failure values (medium plasticity clay; 500 000 simulations)

Publications related to the thesis: I. Kádár, L. Nagy (2010b), I. Kádár, L. Nagy, Nagy R. (2016).

THESIS 6

The examination of distribution fitting of CPT tip resistance showed that the assumption of normal distribution is not correct in most cases. Eurocode 7 derivates the characteristic value from the mean value. In approximately three-quarters of the examined cases, the lognormal distribution function gives a better approximation than normal distribution. The mean value and the mode of lognormal distribution may differ significantly (Figure T4). The mode of lognormal distribution gives a more accurate approximation of the expected value than the normal distribution's mean value (Fig. 4).

In case of CPT test results when the measured data follows lognormal distribution, the characteristic value can be defined according to the following equation:

$$X_k = e^{\mu-\sigma^2} \cdot (1 - k \cdot \sigma)$$

where X_k is the characteristic value, μ is the mean and σ is the standard deviation. Value of k can be assumed based on Table 6.

Condition	**k**
$1.36 \cdot \sqrt{n} > \mathrm{D}_\mathrm{n}^{0.05}$	0
$1.36 \cdot \sqrt{n} < \mathrm{D}_\mathrm{n}^{0.05}$	0.1

Table 6. k value for determination of characteristic value ($\mathrm{D}_\mathrm{n}^{0,05}$ is the result of Kolmogorov–Smirnov-test on 95% confidence level; n is the number of samples)

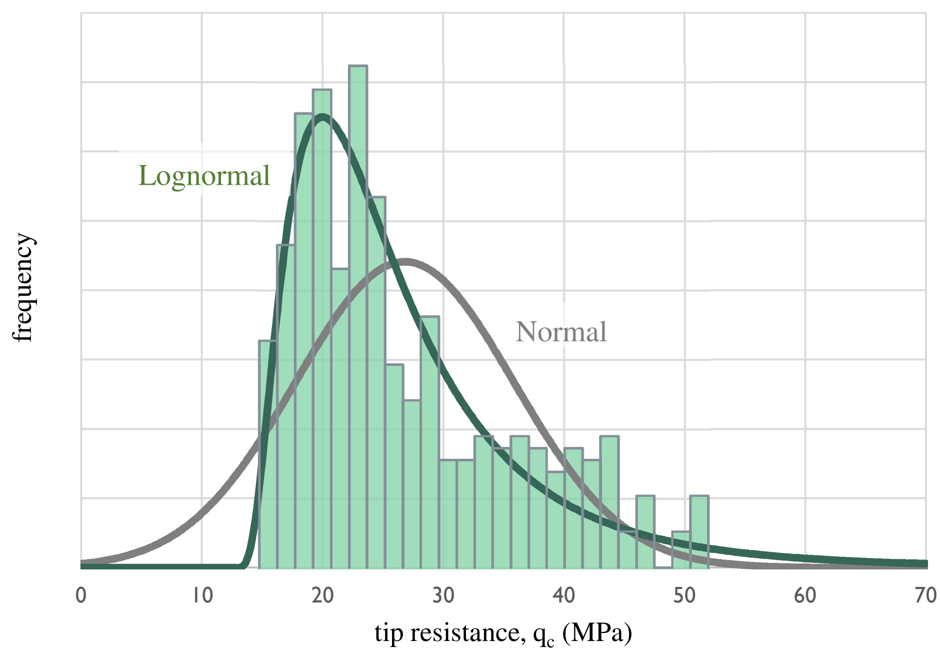

Figure 4. Example for fitting Normal and Lognormal distribution function on CPT tip resistance measurements

Publications related to the thesis: I. Kádár, L. Nagy (2017b), I. Kádár, L. Nagy (2018a), I. Kádár, L. Nagy (2018b).

5. THE AIM AND APPLICABILITY OF THE DISSERTATION

In the dissertation I summarized the results of a large number of laboratory tests and in-situ investigations, which serve as a useful set of data for researchers committed to this topic. I consider the coefficient of variation values particularly important for practice.

The defined quality classes can be determined to further soil parameters and connection can be established with geotechnical categories included in the Eurocode standard, with importance categories of buildings and with the degree of damage caused by failure. The coefficient of variation linked to a geotechnical category can be used as a criterion to control and check whether the field explorations and laboratory experiments have been properly detailed.

Further development can be the separation of the uncertainty of soil parameters and the numerical determination of these components. The errors of the measurement method, the measuring instrument, the personnel or the soil sample itself, etc. are superimposed on the basis of the propagation of uncertainty and the resulting error gives the fundamental uncertainty.

Progress can be made by considering the correlation of internal friction angle and cohesion in our calculation methods. Using the two shear strength parameters in combination can reduce their higher coefficient of variation compared to the shear strength itself. It means that our existing calculation methods (e.g. slope stability, bearing capacity of shallow foundations) might be revised. In the reliability calculations I applied the available deterministic formulas, in which the chosen geotechnical parameters were treated as probability variables. This semi-probabilistic approach can be extended by taking the uncertainties associated with the subsoil, with the foundation and with the superstructure into account with respect to their interaction as well.

Examining the goodness-of-fit of different distributions of soil parameters will soon become a routine task in civil engineering practice. The growing IT capacity and the user-friendly software products with statistical tools ensure the background.

PUBLICATIONS

OWN PUBLICATIONS ON THE TOPIC OF DISSERTATION

I. Kádár (2013): Some characteristic values of the stability analysis of MAL dams. Proceedings of the Second Conference of Junior Researchers in Civil Engineering. Location and date of conference: Budapest, Hungary, 2013.06.17-2013.06.18., pp. 100-104.

I. Kádár (2014): Talajmechanikai paraméterek meghatározása a matematikai statisztika eszközeivel. XVIII. Nemzetközi Építéstudományi Konferencia. Location and date of conference: Csíksomlyó, Romania, 2014.06.12-2014.06.15., pp. 137-140.

I. Kádár, L. Nagy (2010a): Determination of the Statistical Analysis of Shear Strength and Shear Strength Parameters. Proceedings of the XIVth Danube-European Conference on Geotechnical Engineering. Location and date of conference: Bratislava, Slovakia, 2010.06.02-2010.06.04. Bratislava: Slovak University of Technology. ISBN:978 80 227 3279 6, pp. 1-7.

I. Kádár, L. Nagy (2010b): Alapadatok az Eurocode 7 alkalmazásához: A nyírószilárdság statisztikai paraméterei. Mélyépítő Tükörkép Magazin, 2010: (October) pp. 46-47.

I. Kádár, L. Nagy (2012a): Talajfizikai jellemzők változékonyságának jellemzése a műszaki ellenőrzés szempontjából. Műszaki Ellenőr Magazin, 2012: (May) pp. 38-40.

I. Kádár, L. Nagy (2012b): Hogyan vegyük figyelembe a különböző talajok nyírószilárdsági paramétereit? Magyar Hidrológiai Társaság, XXX. Országos Vándorgyűlés. Location and date of conference: Kaposvár, Hungary, 2012.07.04-2012.07.06. ISBN:978-963-8172-29-7, pp. 1-10.

I. Kádár, L. Nagy (2014): Tanulságok a nyírószilárdsági paraméterek statisztikai értékeléséből. Geotechnika 2014 Konferencia. Location and date of conference: Ráckeve, Hungary, 2014.10.13-2014.10.15. ISBN:978-615-80006-2-8. 8 p.

I. Kádár, L. Nagy (2015a): Evaluation and Application of Characteristic Values Based on Eurocode 7 Design Methodology. Geotechnical Safety and Risk V. Location and date of conference: Rotterdam, Netherlands, 2015.10.13-2015.10.16. ISBN:978-1-61499-579-1, pp. 492-496. *DOI: 10.3233/978-1-61499-580-7-496*.

I. Kádár L. Nagy (2015b): The Statistical Analysis of Hydrological Disasters. Geotechnical Safety and Risk V. Location and date of conference: Rotterdam, Netherlands, 2015.10.13-2015.10.16. ISBN:978-1-61499-579-1, pp. 499-504. *DOI: 10.3233/978-1-61499-580-7-503*.

I. Kádár, L. Nagy (2016): Geotechnikai vizsgálatok minimálisan szükséges számának meghatározása. Műszaki Ellenőr Magazin V.:(4.) p. 42. 4 p.

I. Kádár, L. Nagy (2017a): Comparison of different standards based on computing the probability of failure of flood protection dikes. PERIODICA POLYTECHNICA-CIVIL ENGINEERING 61:(1) pp. 146-153. *DOI: 10.3311/PPci.9501*.

I. Kádár, L. Nagy (2017b): The examination of different soil parameters' coefficient of variation values and types of distributions. Proceedings of the 6th INTERNATIONAL YOUNG GEOTECHNICAL ENGINEERS' CONFERENCE. Location and date of conference: Seoul, South-Korea, 2017.09.16-2017.09.17. pp. 98.

I. Kádár, L. Nagy (2018a): Comparison of determination of oedometric modulus based on CPT and laboratory testing in case of pleistocene sand layers. Proceedings of XVI. Danube-European Conference on Geotechnical Engineering: Geotechnical Hazards and Risks. *DOI: 10.1002/cepa.749*.

I. Kádár, L. Nagy (2018b): CPT szondázási eredmények értékelése különböző eloszlások figyelembevételével. Műszaki Ellenőr Magazin (accepted).

I. Kádár, L. Nagy, Nagy R. (2016): Síkalapok teherbírásának megbízhatósági vizsgálata. Geotechnika 2016 Konferencia. Location and date of conference: Ráckeve, Hungary, 2016.10.10-2016.10.12., 13 p.

I. Kádár, L. Nagy, A. Takács (2010a): Talajok nyírószilárdságának statisztikai értékelése. Mérnökgeológia - Kőzetmechanika 2010 Konferencia. Location and date of conference: Budapest, Hungary, 2010.03.25 Budapest: Műegyetemi Kiadó, 2010. pp. 107-112. ISBN:978-963-313-001-8.

I. Kádár, L. Nagy, A. Takács (2010b): A nyírószilárdság karakterisztikus értékének statisztikai meghatározása. Geotechnika 2010 Konferencia. Location and date of conference: Ráckeve, Hungary, 2010.10.26-2010.10.27. ISBN:978-963-89016-0-6, 7 p.

OTHER PUBLICATIONS

Bán Z., I. Kádár, Nagy G. (2015a): Malajzia multifunkcionális alagútja, a SMART. Mérnök Újság XXII.:(12) p. 28. 4 p.

Bán Z., I. Kádár, Nagy G. (2015b): Malajzia "okos" alagútja: a SMART, Műszaki Ellenőr Magazin IV:(11) pp. 43-46.

Berinkei O., I. Kádár (2013): Mobil árvízvédekezés. Magyar Építőipar 63:(5) pp. 213-219.

Ivicsics F., Horkai A., I. Kádár, L. Nagy, Nagy M., Nagy Z., Vona M. (2016): Mobil árvízvédelmi falak minősítése. Mérnöki Kamara Nonprofit Kft., 169 p.

I. Kádár (2009): A nyírószilárdság és a nyírószilárdsági paraméterek statisztikai jellemzése. BME Építőmérnöki Kar TDK, Építőanyagok és geotechnika szekció, pp. 1-41.

I. Kádár (2011): A nyírószilárdság és a nyírószilárdsági paraméterek statisztikai jellemzése (absztrakt). XXX. OTDK Műszaki Tudományi Szekció Kiadványa, ISBN: 978-963-720-85-5. Konferencia helye és ideje: Baja, Hungary, 2011.04.27-2011.04.29.

I. Kádár (2015): Mobile Flood Protection Walls. POLLACK PERIODICA: An International Journal For Engineering and Information Sciences 10:(1) pp. 133-142. *DOI: 10.1556/Pollack.10.2015.1.13.*

I. Kádár, Mahler A., Móczár B., A. Takács (2015a): Kötött talajok szilárdsági és alakváltozási jellemzőinek statisztikai értékelése. XIX. Nemzetközi Építéstudományi Konferencia: ÉPKO 2015. Location and date of conference: Csíksomlyó, Románia, 2015.06.04-2015.06.07., pp. 89-92.

I. Kádár, Mahler A., Móczár B., A. Takács (2015b): Gyöngyös környéki agyagok szilárdsági és alakváltozási paraméterei karakterisztikus értékeinek meghatározása. 4. Kézdi Árpád Emlékkonferencia. Location and date of conference: Budapest, Hungary, 2015.05.21 Budapest: BME Geotechnika és Mérnökgeológia Tanszék. ISBN:978-963-313-180-0, pp. 115-122.

I. Kádár, L. Nagy (2015c): Az árvízvédekezés "fantom" szerkezetei: mobil árvízvédelmi fal. Műszaki Ellenőr Magazin IV:(12) 4 p.

I. Kádár, Szatmári T., J. H. van den Berg, D. Woods (2016): Comparison of deep foundation solutions for embankments with sensitivity analysis using finite element method. Proceedings of 6th European Geosynthetics Congress, Ljubljana. Location and date of conference: Ljubljana, Slovenia, 2016.09.25-2016.09.28., pp. 1226-1237.

I. Kádár, Szatmári T., J. H. van den Berg, D. Woods (2017): Gevoeligheidsanalyse toepassing van geokunststoffen voor fundering van aardebanen, GEOTECHNIEK 21:(2) pp. 40-45.