Magazine of Civil Engineering

journal homepage: http://engstroy.spbstu.ru/

ISSN 2712-8172

Research article

UDC 624.131

DOI: 10.34910/MCE.122.5

Validation metrics for non-linear soil models using laboratory and in-situ tests

R.F. Sharafutdinov ✉

Gersevanov Research Institute of Bases and Underground Structures (NIIOSP), JSC Research Center "Stroitelstvo", Moscow, Russia

✉ *linegeo@mail.ru*

Keywords: numerical modeling, verification, validation, laboratory tests, in-situ testing, statistical analysis, back analysis

Abstract. The article discusses the application of statistical metrics for the validation of comprehensive non-linear soil models. The assessment was carried out on the basis of triaxial, oedometer, consolidation and plate load tests of sandy and clay soils. Validation of non-linear soil models was divided depending on the problem type: strength-type problem and strain-type problem. For a strength-type problem the indicators of failure points should be compared. In the course of strain-type problem the stress-strain curves should be compared. Average ratio of experimental data to calculated and coefficient of variation showed the highest efficiency for standard triaxial and oedometer tests, as they allow taking into account the specificity of the deviation and its variability. Other statistical metrics are less effective in geotechnical engineering. Validation according to consolidation tests is recommended to be performed based on the analysis of the time of 100% primary consolidation and the slope of the consolidation curve during the creep phase. In the course of stress paths analysis (for example, in the course of unloading and further reloading), the advantage should be given to visual assessment. Acceptable values of validation metrics for geotechnical engineering are proposed. The specific values of deviations should be determined by the analyst depending on the required accuracy of calculations, the responsibility of the construction object and the assessment of the risk of an accident.

Acknowledgements: The author extended appreciation to the staff of the Gersevanov Research Institute of Bases and Underground Structures (NIIOSP) for aid and the materials provided.

Citation: Sharafutdinov, R.F. Validation metrics for non-linear soil models using laboratory and in-situ tests. Magazine of Civil Engineering. 2023. 122(6). Article no. 12205. DOI: 10.34910/MCE.122.5

1. Introduction

Comprehensive geomechanical models of soils (hereinafter "non-linear models") that take into account physical and geometric non-linearity, plastic behavior, hardening, filtration and rheological behavior are widely used for geotechnical analysis. Such models are: Duncan-Chang, Hardening soil, Cam-Clay, Soft soil, etc. [1–6]. This became possible due to the widespread introduction of Finite Element Method (FEM) software, such as PLAXIS, MIDAS GTS, Z-Soil, etc.

Correct description confirmation of the soil's real behavior model is carried out on the basis of validation [7, 8], by comparing the model with experimental or reference data.

The model should be selected depending on the type of the problem, the soil type and the factors that should be taken into account in the calculation [9]. When using non-linear models, it is recommended to evaluate the adequacy of the simulated environment's behavior for the required loading trajectory. For the first approximation, it is advisable to compare the results of back analysis with real soil tests. It shows the possibilities and boundaries of models for specific tasks [10]. However, the actual behavior may differ from the laboratory one due to the influence of the stress state and the history of its formation, the real density of the soil in the array, anisotropy and so on.

However, the use of non-linear models may not be sufficient for exact analysis. The fact is that different FEM software has various model implementations [9], which is confirmed by research [11, 12]. It is customary to compare the reliability and applicability of models with known solutions [13–15]. However, the listed works perform a comparison in analysis of different groups of specialists, different approaches, prerequisites, and models. Brinkgreve&Engin admit a discrepancy in the results of various models of no more than 10 %. They also indicate that non-linear soil models should be reliable and repeatable which counts as the partial responsibility of software developers [9].

Russian geotechnical regulatory documents enact the verification and validation of soil models. Thus, in accordance with Russian Rules of Constructions SP 22.13330.2016, to obtain reliable and accurate results using non-linear models, it is necessary to perform their validation (i.e. prove the adequacy of their application) by comparing the calculation results with monitoring data, laboratory and in-situ soil tests. However, geotechnical codes and publications do not contain clear validation metrics. The regulations of SP 22.13330.2016 are only limited to the requirements for validation of bearing capacity based on its upper and lower estimates. So the model can be considered acceptable if the obtained calculation results are in the range between the upper and lower estimates of the bearing capacity. In this case, the difference between the values of the upper and lower estimates should be no more than 10% of the calculated value. This approach is considered valid, but Russian Rules of Constructions SP 22.13330.2016 does not explain whether this refers to average or partial values, or how to take into account variability, etc.

V.N. Shirokov and M.P. Golub [16] suggested using the maximum deviation from experimental values by no more than 15 % as a metric for the possibility of using the model. The approach is similar to Interstate Standard GOST 20522-2012, however, the authors here perform a deviation that is relative to particular values at different stages of plate load test.

Model validation has found wide application in modeling practice. Analysis of most publications [9, 10, 13–15] shows that the assessment of the adequacy of models in geotechnical engineering is carried out mainly qualitatively, not quantitatively.

Requirements for quantifying the adequacy of models using various validation metrics can be seen in related disciplines: hydrodynamics, thermodynamics [17] and solid mechanics [18–19]. In Russia, matters of validation are described in Interstate Standards GOST R 57188-2016 and GOST R 57700.2-2017 and are used for software certification. The validation metric depends on the modeled process and the corresponding experiment. Preference should be given to metrics that contain an assessment of variability based on confidence level.

The validation metrics should determine the maximum acceptable difference between the analysis and the experiment and it also should take into account: the accuracy of the model, the limitations associated with obtaining experimental data (accuracy of sensors, equipment, cost of testing, etc.), the stage of engineering evaluation (conceptual solutions, final project) and the consequences of model non-compliance with the requirements of reliability and safety. At the same time, the validation requirements should include recommendations for action in case if results are not acceptable. In such cases, it may be: model improvements, refinement of experimental data, mitigation of validation metrics, etc.

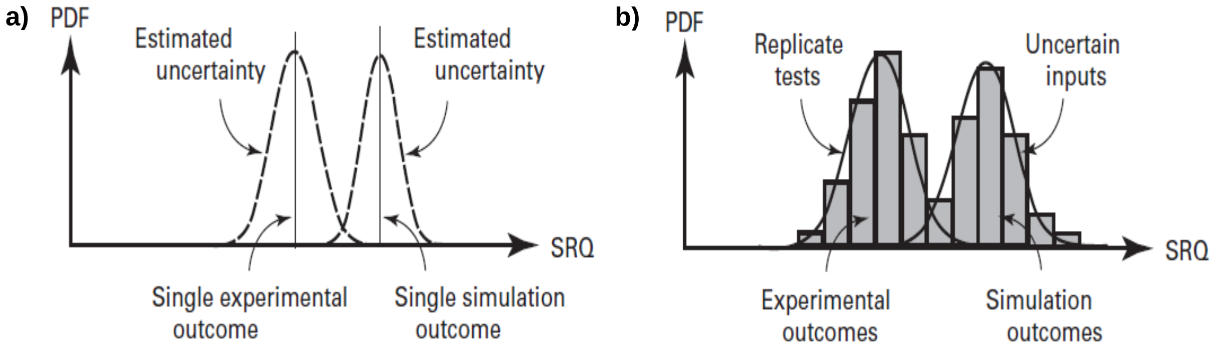
In the manual ASME V&V 10.1-2012 two validation approaches are considered: (a) when the system response quantities (SRQ) are limited and the information about the error is nonexistent or received from experts in the subject area[1] (Fig. 1a); (b) when there is a series of system response quantities and error data available (Fig. 1b). The errors in the simulation results are detected using probabilistic analysis with undefined model inputs that are obtained from various types of repeated tests.

Validation is based on a comparison of cumulative distribution functions (CDF). The validation metrics is the area enclosed between the experimental and model CDF, normalized by the absolute mean of the experimental outcomes (Fig. 2) [20]. The relative area is calculated by the formula:

$$M^{SRQ} = \frac{1}{|\bar{y}_o|}\int_{-\infty}^{\infty}\left|F_{y_c}(y) - F_{y_o}(y)\right| dy, \tag{1}$$

[1] For example, in the Russian GOST 20522-2012 engineering-geological elements are separated, taking into account the limitation of the coefficient of variation ν ≤ 0.30 in mechanical characteristics. The actual ν can be taken as the error of the experimental data.

where M^{SRQ} is relative difference in area of cumulative distribution functions (CDF), y_o are experimental values; y_c are calculated values; $F_{y_c}(y)$ is the CDF of calculated values; $F_{y_o}(y)$ is the CDF of observed (experimental) values.



(a) when the SRQ are limited and the information about the error is nonexistent or received from experts in the subject area

(b) when there is a series of system response quantities and error data available

Figure 1. Illustration of two approaches to validation.

This metric, in fact, is a relative absolute error. The M^{SRQ} parameter has positive value and approaches zero, in the case when both CDF functions coincide. When two CDF functions do not intersect, the integral of equation (1) represents the absolute difference between the mean values. In the case of deterministic values, the CDF are stepwise, and the area between them is simply the absolute value of the difference. In solid mechanics, it is customary to limit M^{SRQ} ≤ 0.1. The application of this metric and the limit value in geotechnics requires verification and clarification, because soils have heterogeneity that is not comparable with other materials (metal, reinforced concrete, etc.).

This approach allows taking into account the error of experimental data as well as the model error. The approach is suitable for comparing the final experimental and calculated data. However, to compare calculation results of non-linear models, the approach is not the most convenient and requires special mathematical processing to calculate the CDF. Also, in order to apply the approach, it is necessary to know the distribution of the results of the model curve, which requires numerous series of calculations and is tied to certain geotechnical model [21–25].

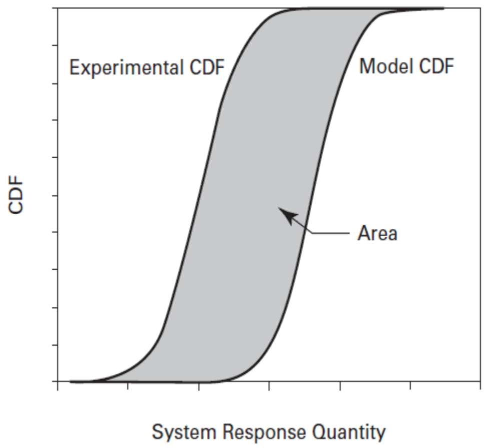


Figure 2. Validation metric for the area of the space enclosed between the experimental and model CDF curves.

For quantitative comparison of calculated and observed designations, S.S. Vyalov [26] proposed using the Pearson's correlation coefficient ρ, the ratio of experimental data to theoretical is $y_i = y_{oi}/y_{ci}$ and their coefficient of variation. Thus, the relationship is considered strong at ρ > 0.8; with a valid description of the experimental model values, the average ratio of experimental data to calculated $y_a = 1$, and the coefficient of variation ν should have the smallest value. However, the approaches used by S.S.

Vyalov were developed mainly for linear and linearized functions. The results of calculations based on comprehensive models are usually non-linear and these approaches cannot be used directly. Notwithstanding this fact, they were taken as the basis for these studies.

Generally, the lack of clear validation metrics in geotechnical engineering causes a subjective analysis using non-linear models. Practice shows that nowadays such analysis is performed mainly qualitatively, based on visual assessment. Approaches used in solid mechanics and hydro-thermodynamics are not widely used in geotechnical engineering. The lack of strictly selected approaches and metrics affects the reliability and repeatability of analysis results.

It is worth noting that the validation of non-linear soil models is part of the overall validation process of the computational model, which is advisable to regulate by special documents containing information about the accuracy of analysis. Performing validation of a non-linear model is a necessary, but not the only one condition for positive verification of the model [9]. When performing validation of the soil model, the type of analysis for which it is performed should be taken into account.

This paper considers the application of validation metrics of comprehensive non-linear soil models. The results of evaluating the applicability of the metrics for triaxial, oedometer test, consolidation and in-situ soil tests are presented. The most effective validation metrics depending on the type of analysis are revealed. Acceptable validation metrics are proposed.

2. Methods and Materials

2.1. The validation metrics of non-linear soil models

The following metrics were considered for the validation of non-linear soil models in the present research.

Visual assessment allows us to qualitatively assess the description of experimental data performed by the model.

The relative difference in area CDF M^{SRQ}, is described in the introduction and is determined by the formula (1).

The mean average error MAE characterizes the average between the absolute experimental and calculated values and is determined by the formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{ci} - y_{oi}\right|. \tag{2}$$

The mean absolute percentage error $MAPE$ is the mean average error, divided by the actual values, determined by the formula:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_{ci} - y_{oi}}{y_{oi}}\right| 100\%. \tag{3}$$

The lower the value of the parameters M^{SRQ}, MAE and $MAPE$, the better the model describes experimental values.

The coefficient of determination R^2 is an explanatory variable and characterizes how much the model explains the experimental data. The parameter is widely used in regression analysis and is determined by the relationship:

$$R^2 = 1 - \frac{D\left[y \middle| x\right]}{D\left[y\right]}, \tag{4}$$

where $D\left[y \middle| x\right]$ is the dispersity of the model error (the difference between the experimental data and the model); $D\left[y\right]$ is the dispersity of the experimental data.

$$D\left[y|x\right]=\frac{\sum\limits_{i=1}^{n}\left[\left(y_{ci}-y_{oi}\right)-MAE\right]^2}{n-1};\tag{4a}$$

$$D\left[y\right]=\frac{\sum\left[y_{oi}-\overline{y_o}\right]^2}{n-1}.\tag{4b}$$

The parameter ranges from 0 to 1; the closer to 1, the higher the strength of the relationship.

The theoretical correlation ratio η along with the coefficient of determination is an explanatory variable.

$$\eta=\sqrt{1-\frac{D\left[y|x\right]}{D\left[y\right]}}.\tag{5}$$

The parameter is in the range from 0 to 1; proximity to 1 indicates a strong relationship. For the linear function $|\rho|=\eta$. The parameter η is convent for the Cheddock scale [27] (Table 1).

The Cheddock scale characterizes the link strength of factors in a fairly wide range and is used in various disciplines as well. For the validation of non-linear models, it is advisable to limit the parameter η to "strong" and "very strong" strength on the Cheddock scale.

Table 1. Strength of relationship between factors and responses on the Cheddock scale [27].

Value η	0 ≤ η < 0.3	0.3 ≤ η < 0.5	0.5 ≤ η < 0.7	0.7 ≤ η < 0.9	0.9 ≤ η < 1	1
Value R^2	0 ≤ R^2 < 0.09	0.09 ≤ R^2 < 0.25	0.25 ≤ R^2 < 0.49	0.49 ≤ R^2 < 0.81	0.81 ≤ R^2 < 1	1
Interpretation	Weak (absent)	Moderate	Notable	Strong	Very strong	Functional

The average value of the experimental data to the calculated y_a *ratio* characterizes the relative deviation of the model from the experimental data and were calculated by the formula:

$$y_a=\frac{1}{n}\sum\limits_{i=1}^{n}y_i=\frac{1}{n}\sum\limits_{i=1}^{n}\frac{y_{oi}}{y_{ci}},\tag{6}$$

where y_i is the value of the ratio of the partial experimental value y_{oi} to the partial calculated value y_{ci}.

A model with a value close to 1 will be considered adequate.

The coefficient of variation ν characterizes the scatter of y_i and was calculated by the formula:

$$\nu=\frac{\sigma_y}{y_a},\tag{7}$$

σ_y is the standard deviation *of* y_i.

The lower the ν value, the lower the dispersion.

2.2. Characteristics of the investigated soils

The research was performed for four sandy soils and three clay soils of the Moscow region.

Medium and fine sands, of alluvial and fluvioglacial genesis of Quaternary-age were considered. (Fig. 3). The depth of sampling varied from 8–20 m. Disturbed soil samples were collected in the Moscow construction site at Kosino and Lermontovsky Prospekt metro stations. Sands S(1)...S(3) were characterized as: homogeneous (C_u = 2.0...2.8); sands S(4) heterogeneous (C_u = 6.3).

The S(1)-S(4) sands were subjected to laboratory tests by consolidated isotropic drained triaxial (d = 50 mm, h = 100 mm) and oedometer tests (d = 87 mm, h = 25 mm). Reconstituted sand specimens were prepared from the disturbed soil samples by compacting sand in the air followed by saturation. Additionally, the sands S(1), S(2) and S(4) were subjected to a plate load test with a foot area of 0.6 m^2 (d = 276 mm) followed by unloading and loading according to GOST 20276.1-2020.

Jurassic-age clay soils were represented as very stiff and stiff clay and firm-stiff loam. Undisturbed samples of clay soils were collected from a depth of 20-45 m at construction sites in Moscow at Kosino and Lermontovsky Prospekt S(5) and at Minskaya S(6), S(7) metro stations.

Specimens S(5) were subjected to oedometer and consolidation tests (d = 87 mm, h = 25 mm). The S(6) and S(7) were subjected to consolidation tests.

The studies were performed for the non-linear Hardening soil model (Table 2) and Soft Soil Creep model (Table 3). For the Hardening soil model, the power factor m was obtained from triaxial m_s and oedometer m_{od} tests. The parameters were determined from the performed tests.

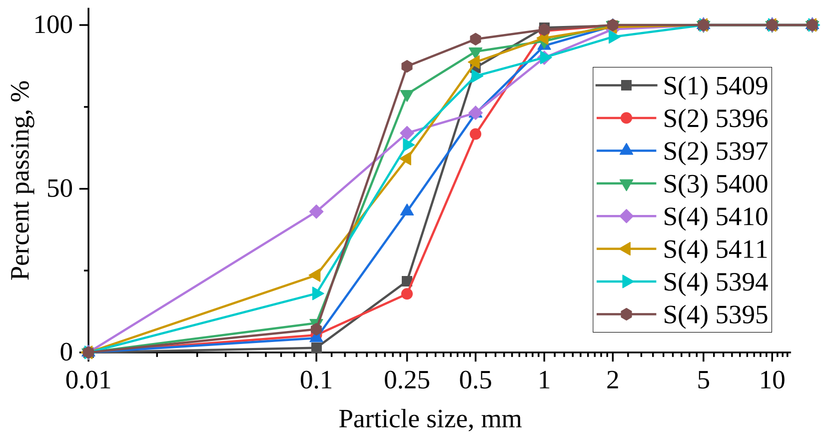


Figure 3. Cumulative curve of the particle size distribution, plotted for the studied sands.

Table 2. Parameters of the Hardening Soil model.

Soil type	γ, kN/m^3	e	p^{ref}, kPa	E_{oed}^{ref}, MPa	E_{50}^{ref}, MPa	E_{ur}^{ref}, MPa	ν_{ur}	m_s	m_{od}	c, kPa	φ,°	ψ,°
S(1) Medium sand	19.0	0.62	75	14.04	14.5	91.0	0.39	0.751	1.12	11.5	35.8	6.6
S(2) Medium sand	20.9	0.60	75	12.64	13.21	85.3	0.42	0.741	1.21	11.1	35.7	6.44
S(3) Fine sand	20.8	0.51	75	4.97	8.99	76.8	0.41	0.634	1.56	8.8	36.8	6.8
S(4) Fine sand	20.0	0.60	75	7.4	10.3	83.6	0.39	0.544	1.18	21.8	33.5	5.3

Note: γ is unit weight; e is void ratio; p^{ref} is reference pressure; E_{oed}^{ref} is reference primary oedometer stiffness; E_{50}^{ref} is reference primary stiffness from CID triaxial tests at the 50 % stress level; E_{ur}^{ref} is reference unloading/reloading stiffness from CID triaxial tests; ν_{ur} is unloading/reloading Poisson's ratio; m_s is power for stress dependent stiffness from triaxial tests; m_{od} is power for stress dependent stiffness from oedometer tests; c is cohesion; φ is friction angle; ψ is dilatancy angle.

Table 3. Parameters of the Soft Soil Creep model.

Soil type	γ kN/m^3	e	φ,°	c, kPa	OCR	POP, MPa	λ^*	κ^*	μ^*	K_f, m/day	B
S(5) Clay I_L<0.25	17.8	1.15	12.8	179	3.85	1.275	0.0995	0.0128	0.00087	2.14E-06	0.39
S(6) Clay I_L<0	19.7	0.69	23	73	2.6	0.810	0.052	0.0069	0.00131	8.14E-06	0.43
S(7) Loam 0.25≤ I_L <0.5	19.5	0.66	27	49	3.7	1.100	0.04	0.0046	0.00097	8.93E-06	0.47

Note: OCR is overconsolidation ratio; POP is pre-overburden pressure; λ^* is modified compression index; κ^* is modified swelling index; μ^* is modified creep index; K_f is coefficient of filtration; B is Skemton's coefficient.

2.3. Characteristic of the back-analysis

The applicability of the validation metrics was evaluated for the Hardening soil (HS) for sandy soils and Soft Soil Creep (SSC) for clay soils. Both models are implemented in the PLAXIS software.

The HS model was evaluated using the results of: consolidated isotropic drained triaxial, oedometer and in-situ plate load tests. The results of the triaxial tests were evaluated for strength-type and strain-type problems; for the other tests, the results were evaluated by the strain-type problem.

The SSC model was evaluated using the results of oedometer and consolidation tests.

Calculation diagrams for triaxial and odometer tests were constructed in the SoilTest module.

Calculation curves for plate load tests were obtained on the basis of back-analysis. For this purpose, the axisymmetric problem was solved in accordance with the computational scheme shown in Fig. 4a. The depth of the plate load test and the loading steps were taken in accordance with the tests performed beforehand. The plate load tests were performed under the stress-controlled loading mode.

The model consolidation curves were obtained on the basis of numerical modeling of oedometer test. For this purpose, a specimen with a height of 25 mm and a diameter of 87 mm was simulated in the axisymmetric formulation, with two-way filtration, which corresponds to the tests performed (Fig. 4b).

The strength-type problem was validated by comparing the calculated and experimental maximum deviator stress $q_{\max}$ at different confining pressures. This allows a comparison of stresses during soil. For strength-type problems, the MAE, $MAPE$, y_a and ν metrics were used. The R^2 and η metrics were not used because their analysis required functions (a group of points) rather than partial SRQ values.

Validation by strain-type problem and by consolidation tests was performed visually and using the M^{SRQ}, MAE, $MAPE$, R^2, η, y_a and ν metrics described in Section 2.1. For the consolidation tests, additionally the comparison of the 100th primary consolidation time t_{100} was performed. For the parameters specified $y_{a(t_{100})}$, $\nu_{(t_{100})}$, $MAE_{(t_{100})}$ and $MAPE_{(t_{100})}$ were calculated.

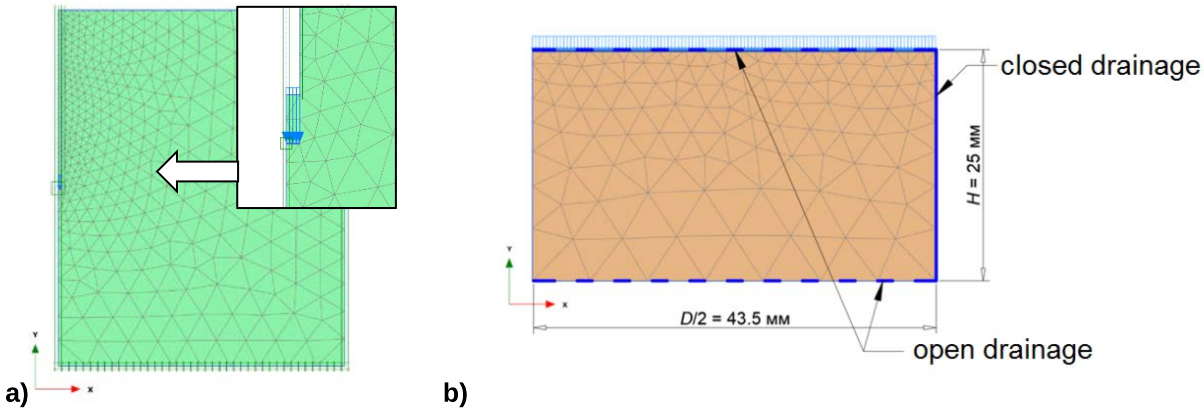


Figure 4. Calculation scheme to validation using back-analysis: (a) – by plate load test; (b) – by consolidation tests.

3. Results and Discussion

3.1. Validation by triaxial test data

Fig. 5 shows the results of triaxial tests of sands at different confining pressure. It can be noted that the HS model implemented in the PLAXIS software describes the behavior of the soil under triaxial compression with high reliability. A similar character of loading curves and failure load was noted.

The results of the *validation by strength-type problem* were shown in Table 4. The estimate of the y_a parameter indicates that in the sands S(1)-S(3) the calculated $q_{\max}$ differs from the experimental ones by less than 5 %. The exception was the heterogeneous sand S(4), where the y_a deviation did not exceed 10 %, and the coefficient of variation did not exceed 20 %. This deviation is due to the influence of the silt particles content (see Fig. 3).

The MAE and $MAPE$ analyses give a similar picture. In homogeneous S(1)-S(3) sands, MAE = 10.0–52.6 kPa; in heterogeneous S(4), MAE = 42.0–254.3 kPa. Relative to the confining pressure, the values obtained were 5 to 15.5 % for S(1)-S(3) and 22 to 56 % for S(4), which are significant for strength-type problems. However, the MAE metric shows only the absolute error, with no reference to the absolute values of the system response. The more advanced $MAPE$ metric, normalized to observed values, does not exceed 5 % in S(1)-S(3) and lies between 7.4 % and 18.0 % in S(4). In this regard, the $MAPE$ metric is more illustrative. The MAE metric is presentable for analysis, but does not characterize the degree of influence of the error. MAE and $MAPE$ metrics always have positive characteristics and consider only the degree of deviation itself, without taking into account its nature (upward or downward).

The parameters y_a, ν and $MAPE$ are the most indicative for strength-type problem estimation. They show not only the nature of the deviation, but also its variability. The MAE itself can act only as an auxiliary characteristic and cannot be the main characteristic when performing validation.

It should be noted that the high values of the validation metrics obtained confirm the thesis about the necessity of determining the strength characteristics of the HS model on the basis of triaxial tests [28].

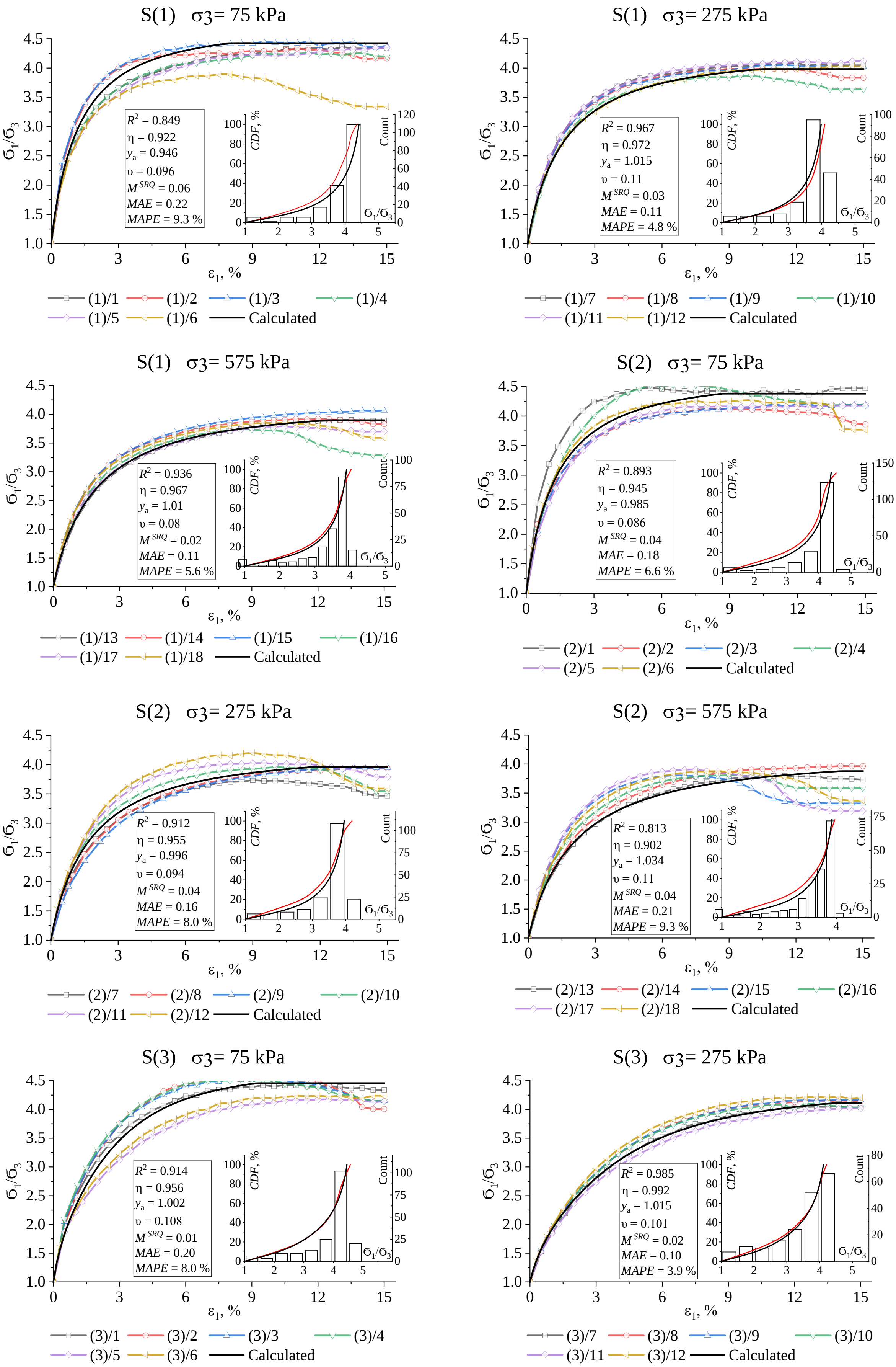
During the *validation by strain-type problem* based on visual assessment, it is obtained that the model describes most of the triaxial tests on the suitable level: the character and absolute values are similar. Differences are observed only in S(4), where the model curve describes observations "in the middle" (in the range ε_1 < 2–3 %), but visually very wide deviations from the experimental data associated with heterogeneity were noted.

Analysis using statistical metrics also indicates high reliability of the experimental data description. For homogeneous sands S(1)-S(3) R^2 = 0.813...0.987, η = 0.902...0.994; the lowest values were obtained for S(2) at σ_3 = 575 kPa, which was also noted visually – the model plot runs along the lower boundary of experimental diagrams. Nevertheless, the link between the calculated and observed values on the Cheddock scale is characterized as "very strong". For heterogeneous sand S(4) R^2 = 0.429...0.775 and η = 0.655...0.881, the relationship is characterized as "notable" and "strong", which is confirmed visually.

The ratio y_a for homogeneous and heterogeneous sands lies in the range of 0.904...1.090 and for most tests deviates by no more than 10 %. The largest deviations of the ratio are y_a = 0.904 for S(1) at σ_3 = 75 kPa, as confirmed by other metrics. The coefficient of variation ν for S(1)-S(3) sands lies in the range of 0.05...0.11 at $MAPE$ = 3.9–9.3 %, for heterogeneous sands it is shown in the range of 0.214...0.261 at $MAPE$ = 15.1–22.9 %. It is worth noting that the deviations obtained lie within the range of ≤ 0.30, which are considered permissible by GOST 20522-2012. This acceptable limit has been used in geotechnical engineering of the USSR and Russia for many years and its effectiveness was confirmed by many years of safe construction experience.

The parameter M^{SRQ} lies in the range of 0.01-0.06 for S(1)-S(3) and 0.05-0.15 for S(4). In S(1)-S(3) homogeneous sands, with minor deviations between calculated and experimental data ($MAPE$ < < 5–6 % and ν < 0.10), M^{SRQ} does not exceed the limits allowed in solid mechanics [19]. At the same time, for heterogeneous soils S(4) with moderate deviation y_a <10 % and values ν < 0.26, the parameter M^{SRQ} can lie both within the acceptable limits and exceed them. In this case the deviations y_a and ν can be considered acceptable still. In this regard, for soils, as materials with variability of mechanical properties (in comparison with other materials), the limiting values of M^{SRQ} from related disciplines are not suitable.

It is worth noting that, as in the case of the $MAPE$ metric, the parameter M^{SRQ} characterizes the absolute value of the deviation. However, the experimental deviations can be in different directions from the model, and the calculated curve can pass between the set of experimental curves and generally adequately describe the soil behavior. This is the advantage of the pair of parameters y_a and ν, which signal whether the model gives a safe or potentially dangerous result.


S(1) σ3= 75 kPa
R² = 0.849
η = 0.922
y_a = 0.946
υ = 0.096
M^SRQ = 0.06
MAE = 0.22
MAPE = 9.3 %
(1)/1 (1)/2 (1)/3 (1)/4 (1)/5 (1)/6 Calculated
S(1) σ3= 275 kPa
R² = 0.967
η = 0.972
y_a = 1.015
υ = 0.11
M^SRQ = 0.03
MAE = 0.11
MAPE = 4.8 %
(1)/7 (1)/8 (1)/9 (1)/10 (1)/11 (1)/12 Calculated
S(1) σ3= 575 kPa
R² = 0.936
η = 0.967
y_a = 1.01
υ = 0.08
M^SRQ = 0.02
MAE = 0.11
MAPE = 5.6 %
(1)/13 (1)/14 (1)/15 (1)/16 (1)/17 (1)/18 Calculated
S(2) σ3= 75 kPa
R² = 0.893
η = 0.945
y_a = 0.985
υ = 0.086
M^SRQ = 0.04
MAE = 0.18
MAPE = 6.6 %
(2)/1 (2)/2 (2)/3 (2)/4 (2)/5 (2)/6 Calculated
S(2) σ3= 275 kPa
R² = 0.912
η = 0.955
y_a = 0.996
υ = 0.094
M^SRQ = 0.04
MAE = 0.16
MAPE = 8.0 %
(2)/7 (2)/8 (2)/9 (2)/10 (2)/11 (2)/12 Calculated
S(2) σ3= 575 kPa
R² = 0.813
η = 0.902
y_a = 1.034
υ = 0.11
M^SRQ = 0.04
MAE = 0.21
MAPE = 9.3 %
(2)/13 (2)/14 (2)/15 (2)/16 (2)/17 (2)/18 Calculated
S(3) σ3= 75 kPa
R² = 0.914
η = 0.956
y_a = 1.002
υ = 0.108
M^SRQ = 0.01
MAE = 0.20
MAPE = 8.0 %
(3)/1 (3)/2 (3)/3 (3)/4 (3)/5 (3)/6 Calculated
S(3) σ3= 275 kPa
R² = 0.985
η = 0.992
y_a = 1.015
υ = 0.101
M^SRQ = 0.02
MAE = 0.10
MAPE = 3.9 %
(3)/7 (3)/8 (3)/9 (3)/10 (3)/11 (3)/12 Calculated
σ1/σ3
ε1, %
CDF, %
Count

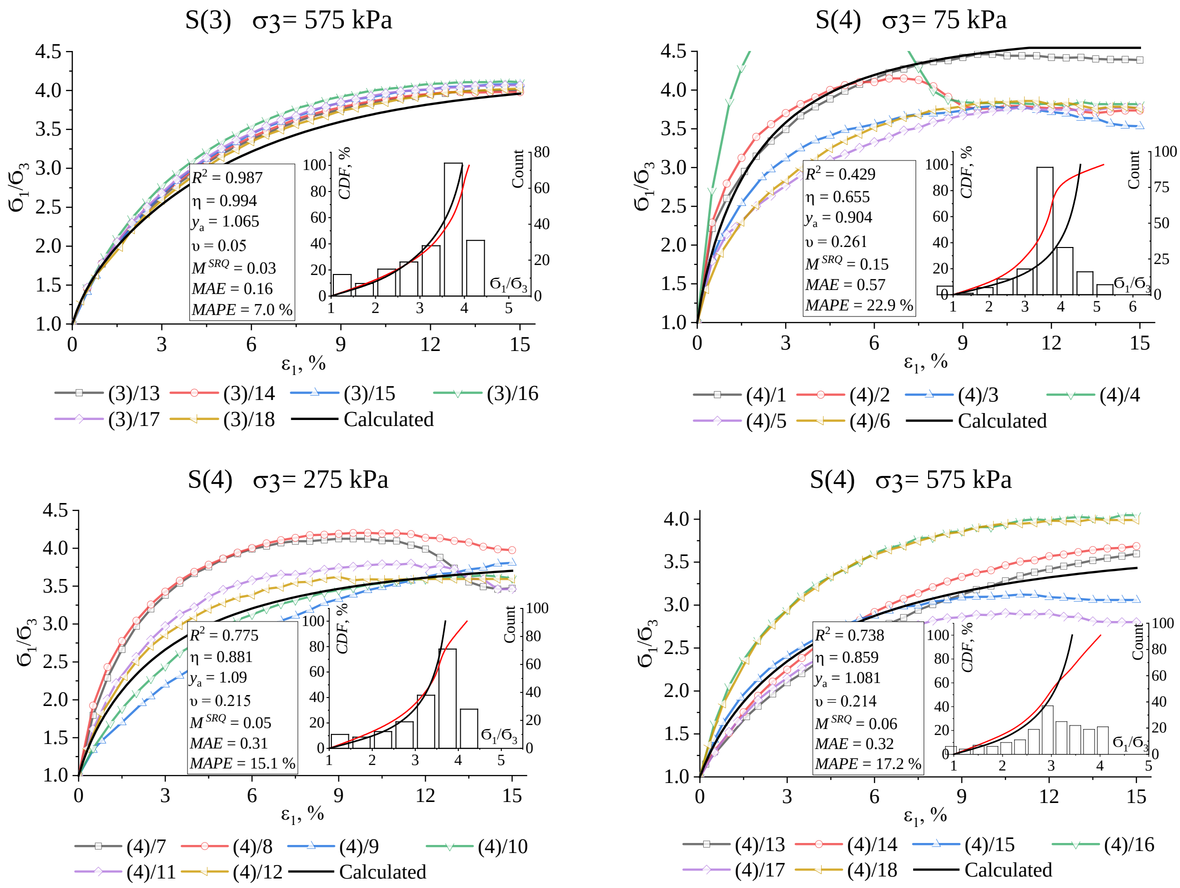


Figure 5. Results of validation by triaxial tests of sands S(1)-S(4) at different confining pressure.

Table 4. Results of calculation of the validation metrics for strength-type problem defined by triaxial tests.

Soil type	Confining pressure σ_3, kPa	$q_{max} = (\sigma_1 - \sigma_3)_{max}$			
		MAE, kPa	$MAPE$, %	y_a	ν
S(1) Medium sand	75	10.0	4.4	0.97	0.06
	275	23.6	2.8	1.02	0.03
	575	47.3	2.9	1.0	0.04
S(2) Medium sand	75	11.6	4.7	0.98	0.05
	275	25.4	3.1	1.01	0.05
	575	32.0	1.9	1.00	0.024
S(3) Fine sand	75	10.3	4.1	0.99	0.05
	275	16.2	1.9	1.00	0.023
	575	52.6	3	1.03	0.017
S(4) Fine sand	75	42.0	18.0	0.91	0.182
	275	62.0	7.4	1.07	0.085
	575	254.3	17.3	1.07	0.181

In general, the combination of all metrics allows one to perform a sufficiently complex analysis of the experimental data deviation from the calculated ones. It is reasonable to use the parameters y_a and ν as the main validation metrics according to the results of triaxial tests for strength-type problems and strain-type problems. R^2, η, $MAPE$ and M^{SRQ} parameters can also be used as additional parameters for strain-type problems validation. Using only the explanatory variables R^2 and η is insufficient, because although they can give low values in some cases, y_a can be close to 1 and ν can be within the acceptable limits.

3.2. Validation on oedometer test data

Oedometer tests are widely used to solve compression-type problems [29]. For the HS model, the m coefficient can be derived from triaxial m_s and oedometer m_{od} tests. Therefore, computational diagrams were constructed for both cases. Fig. 6 shows the experimental and calculated compression curves for soils S(1)...S(5).

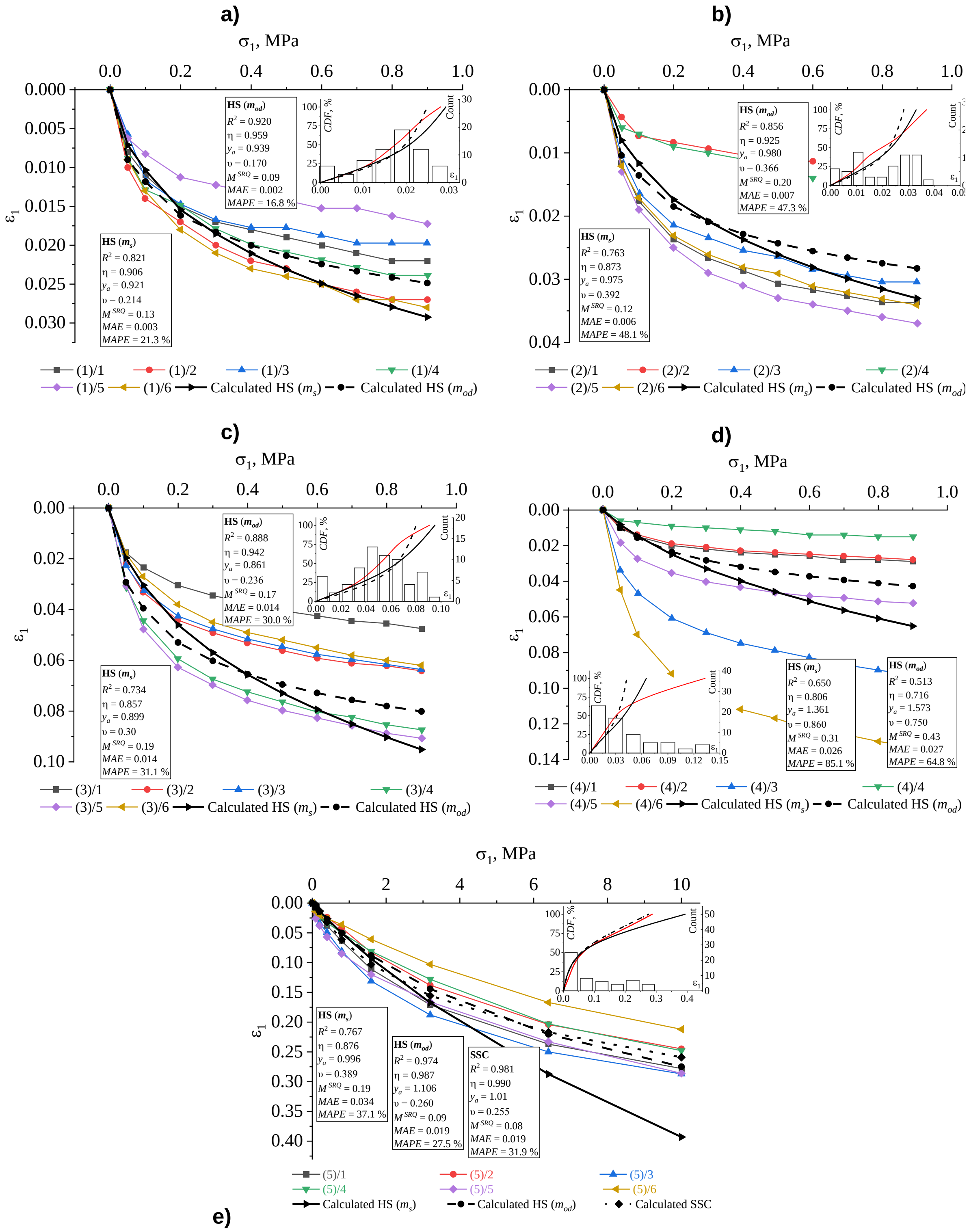


Figure 6. Results of validation by oedometer compression tests:
a – S(1); b – S(2); c – S(3); d – S(4); e – S(5).

A visual comparison of the curves indicates their similar nature. However, the test results can be divided into two groups.

The first group, which includes soils S(1), S(3) and S(5), is characterized by a "heap" arrangement of experimental curves. Here the metrics indicate a well-done description of the experiments by the model: R^2 = 0.734...0.981, η = 0.887...0.990 ("very strong" strength of relationship); y_a = 0.861...1.106, ν = 0.17...0.389, M^{SRQ} = 0.09-0.19. At the same time the parameter $MAPE$ = 21.3–37.1 % indicates a significant scatter.

The second group consists of soils S(2) and S(4), characterized by a wide scatter of experimental curves. Visual variation was confirmed by the values of the explanatory variables, which are 15–30 % lower than for the first group (R^2 = 0.513...0.856, η = 0.716...0.925). Although R^2 and η correspond to "strong" strength of relationship, other metrics suggest very significant deviations (y_a = 0.975...1.573, ν = 0.366...0.860, $MAPE$ = *47*.3–85.1 % and M^{SRQ} = 0.12–0.43). The deviations obtained for S(2) and S(4), as a rule, are confined to the first stage of loading application and at the subsequent stages the deformation practically does not differ from the calculated one (experimental and model curves are parallel). The indicated effect may be due to heterogeneity (in the case of S(4)), as well as to a loose fit of the oedometer stamps to the samples and does not characterize compressibility.

It can be seen that R^2 and η can indicate a high strength of relationship, while the model will deviate significantly from the experimental diagrams. As in the case of triaxial tests, the use of the R^2 and η metrics is not sufficient for a reliable statistical comparison and must be considered in correlation with y_a, ν and $MAPE$, M^{SRQ}.

It is worth noting that for the compression-type problem using the HS model, the calculations with m_{od} describe the tests better. The value of R^2 in the case of m_{od} is 10–30 % higher than in the case of m_s, with similar y_a. At the same time, the scatter of the data in the case of m_{od} is lower by 25–35 %. This confirms the thesis about the necessity of setting the parameter m depending on the problem to be solved. At the same time, for clayey soil S(5), the SSC model gives a more reliable description than HS.

3.3. Validation by consolidation test data

Fig. 7 shows the results of the consolidation tests of soils S(5), S(6) and S(7).

Based on the visual evaluation of the diagrams, it can be noted that the model curves satisfactorily describe the experimental data; significant differences were observed in the absolute deformations. The transition zones from primary consolidation to secondary consolidation (creep) in the model and experimental diagrams coincide; a similar slope of the curves in the creep stage was observed. The explanatory variables lie in the range: R^2 = 0.143...0.958 and η = 0.379...0.979 and indicate the strength of relationship from "moderate" to "very strong". At the same time, the deviation parameters lie in a satisfactory range: y_a = 0.815...1.390, ν = 0.02...0.327, M^{SRQ} = 0.04...0.28. Thus, as for other types of tests, the explanatory variables do not allow us to objectively characterize the model accuracy with the required reliability.

The values of the $MAPE$ parameter are close to the coefficient of variation. However, when comparing one experimental diagram with one model curve (e.g., Fig. 7b) $MAPE$ = 25.1 %, although the graphs run parallel (y_a = 0.815, ν = 0.02), i.e. no variability is observed. For consolidation tests, $MAPE$ does not show the objectivity of the estimate, because it characterizes the error in measured strains. On the contrary, the coefficient of variation does not take into account the mean relative values of the error, but considers their variability.

Wide scatter of strains of samples is caused by differences on conditionally instantaneous sections of compressibility at t < 1 min, which is characterized by closing of pore spaces with gas content. This is evidenced by the low values of the Skempton's coefficient B = 0.39–0.47, which should be taken into account when modeling performing [8]. Another reason for non-uniform precipitation is the uneven surface of the samples in contact with the plate load test (similar to the oedometer tests). Meschyan S.R. showed that such effects often occur in stiff clays, where it is difficult to ensure an absolutely smooth surface of the specimen contact with the stamps during specimen preparation [30].

The task of any consolidation analysis is to predict strains over time. Therefore, the absolute values of settlements in the validation should play a secondary role. For the tests in question, the comparison by t_{100} showed a satisfactory match: mean deviation $y_{a(t_{100})}$ = 0.82...1.352, $\nu_{(t_{100})}$ = 0.05...0.196, $MAPE_{(t_{100})}$ = 9.2...25.3 %.

In view of the above, validation based on consolidation test data should be performed on the basis of 100 % primary consolidation time as well as the slope of the consolidation curve during the creep stage. Comparison of calculated and experimental strain values at different points in time is an auxiliary element of validation.

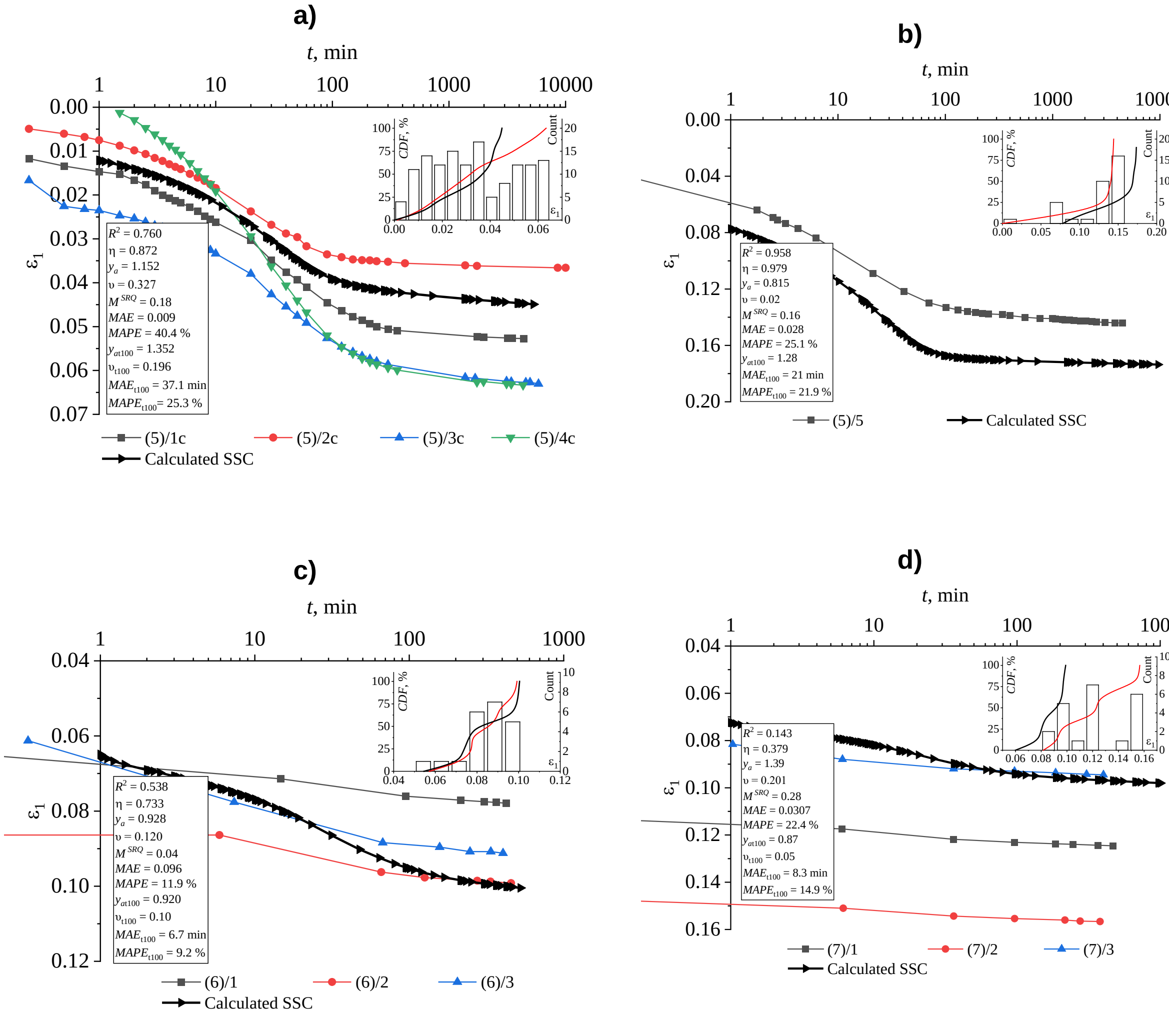


Figure 7. Results of validation by oedometer consolidation test: a – S(5), σ_1 = 0.55 MPa; b – S(5), σ_1 = 4 MPa; c – S(6), σ_1 =1.24 MPa; d - S(7), σ_1 =1.13 MPa.

3.4. Validation by plate load test data

Back analysis of the plate load test tests were performed for sands S(1), S(2), and S(4) (Fig. 8).

A visual comparison shows a good strain match on the primary loading up to 250 kPa. On the unloading path it can be seen that the stiffness E_{ur} is higher than that of the plate load tests. Coincidences on the unloading path are observed only for soil S(2) at a depth of 8.8–9.0 m. The simulated reloading path also differs from the experimental data.

In the analysis based on the validation metrics it can be noted that the differences between the model and calculated curves are insignificant: R^2 and η above 0.9, y_a = 0.923–1.20 at ν = 0.121...0.239, M^{SRQ} < 0.19. A separate deviation of variation coefficient ν = 0.39 is revealed for heterogeneous sand S(4) with "ideal" y_a = 1.017 and moderate $MAPE$ = 25 %.

When analyzing comprehensive stress paths (e.g., with unloading and reloading), the considered metrics do not always allow an objective evaluation of the analysis results. For simple oedometer and triaxial tests, the validation metrics characterized negligible deviations which were confirmed visually. For plate load tests with two stress paths, the metrics lie in satisfactory ranges, while a good visual match was observed in only 30 % of the curve sections.

Such significant discrepancies are explained by peculiarities of the HS model. First, the model parameters are determined on the basis of laboratory tests. Damage of samples during sampling, transportation and storage, lack of reliable data on density (in course of sand) and real stress state in the massif (for accurate laboratory tests) reduce the reliability of modeling. This causes the necessity to confirm the applied model parameters on the basis of plate load and pressuremeter tests. Secondly, the HS model takes into account isotropic hardening and the yield surface that can expand due to plastic straining but changes its position after the change of stress paths in principal stress space. In the process of unloading/reloading or non-linear stress paths, the position of yield surface changes, which affects the deformation behavior of soil. Therefore, for comprehensive stress paths, the model should be evaluated on the basis of visual assessment, and its permissibility should be evaluated by an analyst.

Considering the above features, the results of the back analysis can be called satisfactory. In the case of simple loading, validation approaches for plate load tests are similar, with triaxial and oedometer tests: first of all, they rely on a joint analysis together with y_a, ν, $MAPE$ and M^{SRQ}, and use the metrics R^2 and η as auxiliary. In case of a comprehensive loading trajectory, the analysis should be based on a visual evaluation and the metrics should have an auxiliary function.

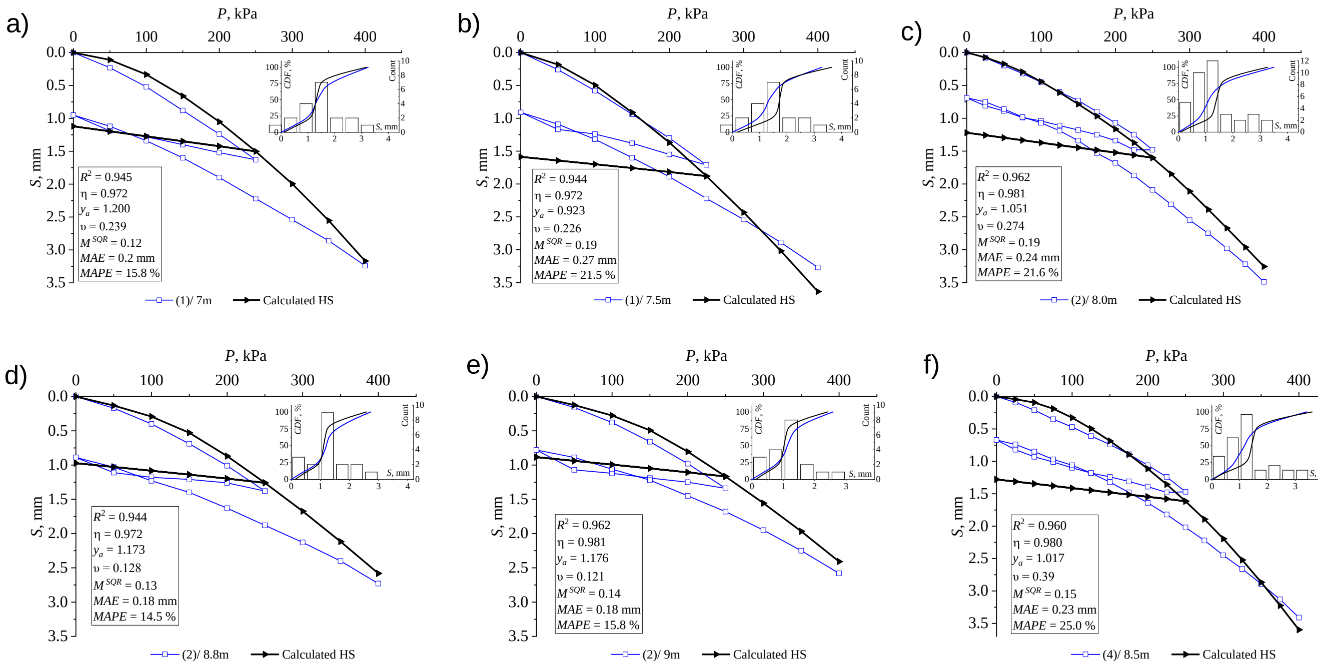


Figure 8. Results of validation by plate load tests for: a, b – S(1) at depths of 7.0 and 7.5 m; c, d, e – S(2) at depths of 8.0, 8.8 and 9.0 m; f – S(4) at depth of 8.8 m.

3.5. Acceptable values of validation metrics

In related disciplines, validation issues have advanced considerably over the past few years. In particular, there have been published works by Roache [31], Coleman and Steele [32], Oberkampf and Roy [33], Hazelrigg and Saari [34] which consider various aspects of validation. An important factor influencing the result of any validation is the experience and competence of the analyst performing the calculation and back analysis [34]. As a rule, the more experienced analyst performs validation for a particular task, the higher the probability of a successful result of verification. The biggest challenge is assessing the competence of the analyst. Validation issues in geotechnical engineering are complicated by the fact that geotechnical engineering is a branch of knowledge at the intersection of several specialties: civil engineer, engineer-geologist and mining engineer [35]. In order to obtain the required competence it is necessary to have at least 10–15 years of experience. In this regard, the issue of applying independent validation metrics and assigning their admissible limits in geotechnical engineering is quite acute.

In these conditions, visual and qualitative metrics used for geotechnical analysis, it is reasonable to supplement with independent metrics, which, for example, correspond to the provisions in other disciplines [17–19]. The analysis has shown that the validation metrics accepted in related disciplines are rather restricted and not suitable for geotechnical engineering. The influence of heterogeneity in the soil massif forces to use less stringent boundary metrics. However, to ensure high accuracy of predictions, it is necessary to provide a good description of the processes as a whole and to introduce restrictions on the scatter of data.

The analysis shows that the validation in geotechnical engineering can be conditionally divided depending on the problem to be solved: the strength-type or the strain-type analysis. The metrics used for estimation also depend on the problem to be solved (Table 5).

Table 5. Validation metrics depending on the test type and problem type to be solved.

Type of test	Validation metric	
	Main	Auxiliary
Triaxial tests - strength-type problem	y_a, ν and $MAPE$	MAE
- strain-type problem	Visual, y_a and ν	R^2, η, M^{SRQ} and $MAPE$
Oedometer tests	Visual, y_a and ν	R^2, η, M^{SRQ} and $MAPE$
Consolidation tests	Visual, $y_{a(t_{100})}$ and $\nu_{(t_{100})}$	y_a, ν, R^2, η, M^{SRQ} and $MAPE$
In-situ soil tests (plate load)	Visual, y_a and ν	R^2, η, M^{SRQ}, and $MAPE$
Stress path	Visual	y_a, ν, R^2, η, M^{SRQ} and $MAPE$

For the strength-type problem, the correctness of the model's description of soil failure is evaluated. To compare the experimental and simulated soil failure, the metrics y_a, ν and $MAPE$ proved to show better results. Considering that the strength-type problem determines the object safety, the peculiarities of the experimental data acquisition and models accuracy, the deviation value y_a, in the author's opinion, should not exceed 5–10 %.

The strain-type problem validation (triaxial, oedometer or plate load tests) assesses the correctness of the model's repeatability of soil deformation under load. Here it is reasonable to perform a visual assessment of stress-strain curves supported by the proposed metrics. As in the case of strength-type problem validation, the y_a and ν metrics showed the greatest efficiency. Metrics M^{SRQ}, $MAPE$, R^2 and η can be used as auxiliary. For the case of a comprehensive stress path (with unloading and reloading, along an atypical trajectory, etc.), visual assessment should be preferred. Model and experimental stress-strain curves in different load ranges may have different deviations (up or down). Because of this, it is more difficult to provide strict values of validation metrics. On the other hand, inaccuracies in the strain analysis affect the operation of the object (settlements, inclinations, crack opening, etc.) and do not affect safety. In view of the above, in the author's opinion, for the strain validation the deviation value y_a should not exceed 10–15 %, and M^{SRQ} should not exceed 0.3. For the explanatory variables R^2 and η were proposed to focus on the "very strong" and "strong" strength of relationship according to the Cheddock scale (Table 1). The use of a lower strength of relationship is inappropriate, as it indicates a bigger error value of the model. For comprehensive stress paths, the acceptability of the model should be evaluated by the analyst.

In case of the consolidation model validation, the advantage should be given to the behavior on a timeline. The estimation of absolute value strains should play a secondary role. Therefore, it will be correct to compare the time of the 100th primary consolidation t_{100} using the metrics $y_{a(t_{100})}$ and $\nu_{(t_{100})}$. Limit values of the validation metrics during consolidation are quite difficult to assign, since tests are usually performed under laboratory conditions on specimens sampled from the soil massif. Consolidation tests are the most sensitive to the quality of sampling, especially from large depths [36]. In this regard, the validation metrics should be assigned by the analyst, taking into account the statistical variability of specific experimental data.

For strength-type and strain-type problems the value of the variation coefficient ν must not exceed 0.3. The value of $\nu \leq 0.30$ is justified by many years of experience in the application of this limit in the USSR and Russia in the allocation of engineering-geological elements (Interstate Standard GOST 20522-2012). The use of a stricter coefficient of variation may be justified for materials with high reproducibility of properties (e.g., systematically erected embankments of homogeneous materials, concrete, metals, etc.). In the case of natural soils, this may lead to limitations on the use of non-linear models in heterogeneous soils. However, it may be justified in the case of high requirements for analysis accuracy.

The values of the $MAPE$ parameter when comparing a group of curves have close values to ν. When comparing with single experimental values, the analysis by MAE and $MAPE$ can give a more objective assessment. However, it is common in geotechnical engineering to consider statistical variability [37], in connection with which the analysis by ν is more preferable. Therefore, allowable values of MAE and $MAPE$ should be determined by the calculator for specific tasks.

For particularly critical objects (unique high-rise buildings, hydraulic structures, etc.) more stringent limits of validation metrics can be considered. Specific values of deviations should be determined by the analyst depending on the required accuracy of analysis, responsibility of the object and accident risk assessment. However, in the author's opinion, the application of more stringent metrics in soils, taking into account the above factors, is not justified.

If the model does not meet the specified validation metrics, additional correction factors can be introduced. As a first approximation, we can consider applying an additional correction factor equal to the value of the maximum deviation y_a.

Based on the estimates made for the proposed validation metrics for various tests, Table 6 was developed. The proposed metrics and their allowable values will make formalizing the performance of validation for geotechnical tasks possible.

Table 6. Acceptable values of validation metrics depending on analysis accuracy requirements and the problem type to be solved.

Requirements for analysis accuracy	Strain-type problem and consolidation					Strength-type problem	
	y_a	ν	M^{SRQ}	R^2	η	y_a	ν
Normal*	$0.85 \geq y_a \geq 1.15$	$\nu \leq 0.30$	$M^{SRQ} \leq 0.3$	$0.81 > R^2 \geq 0.49$	$0.90 > \eta \geq 0.70$	$0.90 \geq y_a \geq 1.10$	$\nu \leq 0.30$
High	$0.90 \geq y_a \geq 1.10$		$M^{SRQ} \leq 0.2$	$R^2 \geq 0.81$	$\eta \geq 0.90$	$0.95 \geq y_a \geq 1.05$	

Note: "*" is allowed for preliminary analysis with high requirements for analysis accuracy

Validation of non-linear models and their parameters using the specified acceptable values is of course a necessary but not sufficient metric for solving the problem of limiting equilibrium in the soil mass. Further, it is necessary to evaluate the influence of application of the validation metrics of Table 6. On the results of analysis of full FEM models using non-linear soil models for specific problems the deterministic [21–23] or stochastic (e.g., using the Monte Carlo method) [24, 25, 38] approaches can be used. It is also necessary to evaluate the applicability of validation metrics on more empirical data and monitoring results.

4. Conclusions

Geotechnical analyses are often performed using non-linear soil models. In order to obtain reliable results using non-linear models it is required to perform their validation. However, model validation in geotechnical engineering is performed expertly or visually, without applying special quantitative metrics. Model quantification requirements are widely used in related disciplines (hydro- and thermodynamics, solid mechanics). At the same time, permissible validation metrics accepted in related disciplines are rather restricted and are not applicable in geotechnical engineering, because soils, in comparison with other materials, have significant heterogeneity. Based on the analysis, the following conclusions were obtained.

- The following validation metrics can be used in geotechnical engineering: visual, the relative area between the calculated and model cumulative distribution function M^{SRQ}, mean average error MAE, mean absolute percentage error $MAPE$, coefficient of determination R^2, theoretical correlation ratio η, average ratio of experimental data to calculated y_a, coefficient of variation ν.

- Validation of the soil models can be divided depending on the problem to be solved: strength-type or strain-type problem. In the first case, the indicators at failure are compared (for example, for triaxial compression it is suggested to use the maximum deviator stress $q_{\max}$). In the second case, the stress-strain curves are compared.
- The metrics y_a and ν showed the greatest efficiency in the strength-type and strain-type problems. They take into account not only the nature of the deviation, but also its variability. Parameters M^{SRQ}, MAE and $MAPE$ take into account the absolute error, but ignore the direction of deviation. The deviations of the experimental data can be in different directions from the model, and yet, the model can adequately describe the behavior of the soil. The advantage of the pair of parameters y_a and ν is that they signal whether the model gives us a conservative or potentially dangerous result. The use of the R^2 and η metrics is not sufficient for reliable validation and should be considered together with y_a, ν, $MAPE$ and M^{SRQ}.
- Validation from oedometer consolidation test data should be based on an analysis of the 100 % primary consolidation time and the slope of the consolidation curve during the creep phase. Ancillary elements of validation should be presented as comparison of calculated and measured strain values at different points in time.
- When analyzing comprehensive stress paths (e.g., with unloading and reloading), the considered metrics do not always allow an objective assessment of the simulation results. Therefore, the advantage should be given to visual assessment.
- Acceptable values of validation metrics for geotechnical engineering are proposed (Table 6). The proposed metrics and their admissible values will allow formalizing the execution of validation for geotechnical tasks. At that, specific values of deviations should be determined by an analyst depending on the required accuracy of analysis, object responsibility and accident risk assessment. The application of the proposed metrics will allow systematization of the validations performed by different specialists and thereby increase the reliability of geotechnical analysis. In case the model does not fit into the required validation metrics, additional correction factors can be introduced.

In the future, it is advisable to: consider the applicability of these metrics to other soil test methods, evaluate the effect of metrics application on the results of full FEM models using non-linear soil models, and test the proposed validation metrics on a larger volume of data.

References

1. Duncan, J.M., Chang, C.-Y. Nonlinear Analysis of Stress and Strain in Soils. Journal of the Soil Mechanics and Foundations Division. 1970. Vol. 96 (5). Pp. 1629–1653. DOI: 10.1061/JSFEAQ.0001458
2. Schanz, T., Vermeer, P., Bonnier P. The hardening soil model: Formulation and verification. Proceedings of the Plaxis Symposium. Beyond 2000 in Computational Geotechnics. Rotterdam. Balkema, 1999. Pp. 281–290.
3. Roscoe, K.H., Burland, J.B. On the generalized stress-strain behaviour of wet clay. Papers of a Conference held in Cambridge. Engineering Plasticity. 1968. Pp. 535–609.
4. Muir Wood, D. Soil behaviour and critical state soil mechanics. UK. Cambridge University Press, 1991. 480 p. DOI: 10.1017/CBO9781139878272
5. Potts, D.M., Zdravkovic, L. Finite element analysis in geotechnical engineering: theory. London. Thomas Telford, 1999. 458 p.
6. Britto, A.M., Gunn, M.J. Critical state soil mechanics via finite elements. Chichester. Ellis Horwood Ltd., 1987. 488 p.
7. Ter-Martirosyan, A.Z., Sidorov, V.V., Ermoshina, L.Y. Opredelenie i verifikacziya parametrov modeli slabogo grunta v uchetom polzuchesti [Determination and verification of parameters of the soft soil with account for creep]. Vestnik MGSU. 2018. 6. Pp. 697–708. DOI: 10.22227/1997–0935.2018.6.697-708
8. Shulyatiev, O.A., Isaev, O.N., Nayatov, D.V., Sharafutdinov, R.F. Prognoz razvitiya deformaczij osnovaniya mnogofunkczional`nogo zhilogo kompleksa [Forecast of base strains development for a multifunctional residental complex]. Geotechnika. 2017. No. 2. Pp. 38–49.
9. Brinkgreve, R.B.J., Engin, E. Validation of geotechnical finite element analysis. Proceeding of the 18th International Conference on Soil Mechanics and Geotechnical Engineering. Paris, France. 2013. Pp. 677–682.
10. Szilvágyi, Z., Ray, R.P. Verification of the Ramberg-Osgood material model in midas GTS NX with the modeling of torsional simple shear tests. Periodica Polytechnica Civil Engineering. Budapest University of Technology and Economics, 2018. 62 (3). DOI: 10.3311/PPci.11191

11. Razvodovsky, D., Skorikov, A. Problemy` i vozmozhny`e puti razvitiya normativnoj literatury` v oblasti proektirovaniya svajny`kh fundamentov [Problems and possible ways of development of pile design building codes]. Bulletin of Science and Research Center "Stroitelstvo". 2020. 26 (3). Pp. 74–85.
12. Pruška, J. Comparison of geotechnic softwares – Geo FEM, Plaxis, Z-Soil. Proceeding of the XIII ECSMGE. Prague, Czech Republic. 2003. Pp. 819–824.
13. Anderson, K.H. et al. Suction anchors for deepwater applications. Frontiers in Offshore Geotechnics. Gourvenec & Cassidy. London, Taylor & Francis. 2005. Pp. 3–30. DOI: 10.1201/NOE0415390637.ch1
14. Jeffries, R.M. Interclay II project. A coordinated benchmark exercise on the rheology of clays. Final report. 1995. 362 p.
15. Schweiger, H.F. Results from numerical benchmark exercises in geotechnics. Proceeding of the 5th European Conf. Numerical Methods in Geotechnical Engineering. Paris. 2002. Pp. 305–314.
16. Shirokov, V.N., Golub, M.P. Shtampovy`e ispy`taniya kak komponent modelirovaniya v srede «PLAXIS» deformirovaniya svyazny`kh dispersny`kh gruntov v osnovanii sooruzhenij [Stamping tests as a component of modeling of cohesive disperse soils deformations in the structure bases using "PLAXIS" program]. Engineering survey. 2018. No. 12. Pp. 16–24. DOI: 10.25296/1997-8650-2017-12-16-24
17. ASME V&V 20-2009. Stanard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer. The American Society of Mechanical Engineers.
18. ASME V&V 10-2006. Guide for Verification and Validation in Computational Solid Mechanics. The American Society of Mechanical Engineers.
19. ASME V&V 10.1-2012. An Illustration of the Concepts of Verification and Validation in Computational Solid Mechanics. The American Society of Mechanical Engineers.
20. Ferson, S., Oberkampf, W.L., Ginzburg, L. Model validation and predictive capability for the thermal challenge problem. Computer Methods Applied in Mechanics and Engineering. 2008. 197 (29-32). Pp. 2408–2430. DOI: 10.1016/j.cma.2007.07.030
21. Kolybin, I.V., Astryab, V.V. Determinirovanny`j podkhod k oczenke chuvstvitel`nosti MKE` modelej pri sovmestnom raschete sooruzheniya s osnovaniem [A deterministic approach to assessing the sensitivity of FEM models in the joint design calculation with the base]. Proceedings of the Gersevanov Research Institute of Bases and Underground Structures (NIIOSP). 2011. Pp. 156–169.
22. Barvashov, V.A., Boldyrev, G.G., Utkin, M.M. Raschet osadok i krenov sooruzhenij s uchetom neopredelennosti svojstv gruntovy`kh osnovanij [Calculation of settlements and tilts of engineering structures taking into account uncertainly of foundation soil properties]. Geotechnika. 2016. No. 1. Pp. 12–29.
23. Barvashov, V.A. Sensitivity of the "bed-structure" system. Soil Mechanics and Foundation Engineering. 2007. 44 (3). Pp. 87–93. DOI: 10.1007/s11204-007-0016-z
24. Sheinin, V.I., Lesovoi, Y.V., Mikheev, V.V., Popov N.B. An approach to reliability assessment in engineering calculations of foundation beds. Soil Mechanics and Foundation Engineering. 1990. 27 (1). Pp. 32–36. DOI: 10.1007/BF02306679
25. Skorikov, A., Pavlovskiy, N. Stokhasticheskij podkhod k oczenke nadezhnosti rezul`tatov rascheta osnovanij i fundamentov v zavisimosti ot chuvstvitel`nosti MKE`-modelej [Stochastic approach to assessing the reliability of the calculation results of bases and foundations depending on the sensitivity of the FEM models]. Bulletin of Science and Research Center "Stroitelstvo." 2021. 29 (2). Pp. 101–111. DOI: 10.37538/2224-9494-2021-2(29)-101-111
26. Vyalov, S.S. Rheological Fundamentals of Soil Mechanics. Elsevier, 1986. 564 p.
27. Chaddock, R.E. Principles and methods of statistics. Boston, Now-York, 1925. 471 p.
28. Sharafutdinov, R.F., Isaev, O.N., Morozov, V.S. Experimental Studies of Cohesionless Subsoil Dilatancy Under Conditions of Triaxial Compression. Soil Mechanics and Foundation Engineering. 2021. 57(6). Pp. 465–472. DOI: 10.1007/s11204-021-09694-3
29. Schanz, T., Vermeer, P.A. On the Stiffness of Sands. Pre-failure deformation behaviour of geomaterials In Jardine, R., Davies, M., Hight, D., Smith, A., Stallebras, S., editors. London. Telford, 1998. Pp. 383–387.
30. Meschyan, S.R. Experimental Rheology of Clayey Soils. Taylor & Francis, 1995. 460 p.
31. Roache, P.J. Fundamentals of Verification and Validation. Albuquerque, NM. Hermosa Publishers, 2009. 476 p.
32. Coleman, H.W., Steele, W.G. Experimentation, Validation, and Uncertainty Analysis for Engineers. 3rd ed. Hoboken, NK. John Wiley, 2009. 336 p.
33. Oberkampf, W.L., Roy, C.J. Verification and Validation in Scientific Computing. Cambridge University Press, 2010. 790 p.
34. Hazelrigg, G.A., Saari, D.G. Toward a Theory of Systems Engineering. Journal of Mechanical Design. 2022. 144 (1). 8 p. DOI: 10.1115/1.4051873
35. Lees, A.S. A competency based approach to managing skills in geotechnical numerical analysis. Proceeding of the XVI ECSMGE. 2015. Pp. 4283–4287.
36. Tanaka, H., Ritoh, F., Omuka,i N. Quality of samples retrieved from great depth and its influence on consolidation properties. Canadian Geotechnical Journal. 2002. 39 (6). Pp. 1288–1301. DOI: 10.1139/t02-064
37. Ermolaev, N.N., Mikheev, V.V. Nadezhnost` osnovanij i sooruzhenij [Reliability of the bases of structures]. Leningrad. Stroyizdat, 1976. 152 p.
38. Sheinin, V.I., Mikheev, V.V., Popov, N.B., Lesovoi Yu.V. Probabilistic calculation for the bed under an individual foundation in accordance with the second group of limiting states. Soil Mechanics and Foundation Engineering. 1991. 28 (2). Pp. 79–83. DOI: 10.1007/BF02316019

Information about authors:

Rafael Sharafutdinov, *PhD in Technical Sciences*
ORCID: https://orcid.org/0000-0002-5806-7190
E-mail: linegeo@mail.ru

Received 18.12.2022. Approved after reviewing 12.07.2023. Accepted 07.09.2023.