DOI 10.52209/1609-1825_2022_4_162 UDC 552.52

Dependence of the Clay Rocks Swelling on Their Geological and Physical Characteristics

[1]***VDOVKINA Darya,** *Master, doctoral student, daryavdovkina@gmail.com,*
[1]**PONOMAREVA Marina,** *Cand. of Tech. Sci., Associate Professor, mv_ponomareva18@mail.ru,*
[1]**PONOMAREVA Yekaterina,** *PhD, Senior Lecturer, evmussina1992@gmail.com,*
[1]*NPJSC «Abylkas Saginov Karaganda Technical University», Kazakhstan, Karaganda, N. Nazarbayev Avenue, 56,*
**corresponding author.*

Abstract. *The active development of the construction industry leads to the fact that territories composed of specific soils are covered. The number of storeys and the complexity of the constructed structures are increasing. Limited financial and time resources require more efficient methods of pre-survey of the development site. In this regard, the requirements for engineering and geological research are increasing. Areas composed of swelling soils require special attention. Swelling clay rocks have the ability to change their volume when wet and thus pose a serious threat to the stability of structures and foundations. The measurement of swelling characteristics takes a long period of time and requires a significant investment of money. The purpose of this study is to establish the relationship between the physical and geological characteristics of clay soils, which require minimal resources, and their relative swelling using machine learning methods. As input data for building a relative swelling model of clay rocks, the physical characteristics of soils were used, such as water content, liquid limit, plastic limit, plasticity index, liquidity index, solid particles density, soil density, dry soil density, buoyant density, void ratio, maximum water capacity, soil's moisture level. According to these characteristics, three relative swelling models were built using the following machine learning methods: Decision tree, Gradient Boosting, Random Forest. The models were evaluated using the Mean Squared Error parameter, which measures the standard deviation between predicted and true values. As a result of evaluating this criterion, the Random Forest model was chosen to build the relative swelling model. The presence of a relative swelling predictive model will save money and time resources at the preliminary stage of engineering and geological studies.*

Keywords: *clay soils, montmorillonite clays, swelling, Random Forest model, Quaternary deposits, Neogene deposits.*

Introduction

Currently, there is an active development of the construction industry. Intensive urban development leads to the development of territories with complex geological conditions [1]. One of the important stages of pre-design work in the construction of buildings and structures is geotechnical investigation. The purpose of geotechnical investigation is a comprehensive survey of the future construction site, specifically the research of the territory geological structure, groundwater monitoring, determination of the physical and stress-related characteristics, forecasting hazardous natural and man-made processes. Particular attention should be paid to subsidence and swelling soils, since the subfoundation must be designed taking into account these features.

Subfoundation built with swelling soils should be designed taking into account the ability of such soils to increase in volume with increasing water content – to swell. With a subsequent decrease in water content in swelling soils, the reverse process occurs – shrinkage [2]. The main swelling soils feature is drastic in their bearing capacity reduction during soaking. Underestimation of swelling soils causes damage to many industrial and civil structures. Such soils are distributed all over the world and numerous damages that have been received as a result of their swelling properties have been documented in different countries, such as Saudi Arabia, Germany, France, Poland, Italy, Spain, USA [3], [4], [5]. Including such damage was recorded in the city of Karaganda, the Republic of Kazakhstan. In 2012, four 5-storey sections of the eight-section Besoba apartment complex collapsed [6], so the issues discussed in this article are very important and relevant. An increase in soil's water content can occur as a result of the following processes [7]:

- due to groundwater level rise or infiltration – soil moistening by industrial or surface waters;
- accumulation of moisture under structures in a zone limited in depth due to violation of the natural conditions of evaporation during the territory development (shielding of the surface).

Determining the swelling properties of clay rocks in the laboratory is costly and time consuming. In this regard, it is important to derive the correlation dependences of soil swelling parameters on their physical characteristics.The availability of predictive models to correlate swelling properties with physical characteristics would significantly reduce the cost and time spent, at least at the preliminary design stage [8]. To this date, numerous studies and analysis of factors that affect soil swelling have been carried out. A generalization of the study's results on the correlation of the selected parameters is presented in Table 1 [1,9-12].

In this regard, this study is devoted to determining the relationship between the relative swelling of clay soils in Karaganda city and their physical and geological characteristics using machine learning methods.

Research methods

As a result of geotechnical investigation in Karaganda city, 399 samples of clayey rocks (clays and loams) were selected and examined in the laboratory and the values of the following characteristics were obtained: water content (w), liquid limit (W_L), plastic limit (W_P), plasticity index (I_P), liquidity index (I_L), solid particles density (p_s), soil density (p), dry soil density (p_{dry}), buoyant density (p_b), void ratio (e), maximum water capacity (W_p), soil's moisture level (S_r), relative swelling (ε_{sw}). Tests in the laboratory to determine the physical characteristics of the soil and relative swelling were carried out according to the standards applied in the territory of the Republic of Kazakhstan. Swelling characteristics are determined by the results of testing samples in consolidometer when the soil is saturated with water. For testing, samples of undisturbed structure with natural water content or samples of disturbed structure with specified values of density and water content are used.

The areas of work where geotechnical investigation were carried out is orographically part of watershed between the Nura and Sherubainur rivers. In general, the relief of the site is a wavy plain, complicated by a small hill. Clays exposed in the studied areas belong to the Neogene period, loams to the Quaternary. Lower Quaternary deposits are brown and greenish-gray loams with interlayers and lenses of gravels, sands and sandy loams. The thickness of the Neogene sediments occurs throughout the territory of the Central Kazakhstan shallow hills. Here, green, greenish-gray montmorillonite clays with manganese «buckshot», places with gypsum and marl lenses, which belong to the Aral Formation of the Lower-Middle Miocene age, are widely developed. The predictive model producing of relative swelling was carried out using machine learning methods such as Decision tree (DT), Gradient Boosting (GB), Random Forest (RF). The high-level programming language «Python» with «Anaconda» distribution kit was used as tools for creating the model. The following parameters were used as input data: physical characteristics of rocks, groundwater level, sampling depth, soil type. At the same time, 80% of the data was used to train the model, and 20% were used to validate the constructed model, according to the Pareto theorem. To select the parameters that have the most pronounced relationship with relative swelling, the method of recursive feature exclusion was used. The adequacy of the regression models was assessed by the Mean Squared Error (MSE) indicator, which measures the standard deviation between predicted and true values. When this indicator tends to zero, the result is considered the best. Also, to select the best model, in addition to the average value of the MSE indicator, a box plot of the MSE indicator was used, which visualizes the difference in the distribution of this indicator between the three models.

Results

The database of laboratory studies consists of 350 samples of clay rocks and was compiled as a result of geotechnical investigation in Karaganda, Kazakhstan. Clay rocks are represented by loams and clays of the Quaternary and Neogene periods, respectively, which were territorially selected in areas located within Karaganda city. To select the soil's physical characteristics that have the most pronounced relationship with relative swelling, the recursive feature selection method was used, which selected the following soil properties: W, W_L, W_P, I_P, I_L, p_s, S_r (Figure 1).

In addition to the selected physical characteristics, the input data were:

- type of soil – loam and clay. The lithology differentiation of rocks makes it possible to take into account the genesis and age of rocks;

Table 1 – Empirical dependencies for assessing soil swelling

№	Correlation
1	$S = 1.00 + 0.006 \cdot (C + PI - w)$
2	$S = -432.06 + 7.73 \cdot C + 0.12 \cdot CEC + 0.46 \cdot PI + 4.30\,\gamma_{dry} - 1.18\,w$
3	$S = 24.5 \cdot (P)^{-0.26} \cdot (PI \cdot C)^{1.26} \cdot [F_i - 7.1 \cdot (P)^{0.22} \cdot (PI \cdot C)^{0.78}]$
4	$S = -31.321 + 0.592\,C + 0.717\,PI - 0.807\,CEC - 0.891\,w + 2.668\,\gamma_{dry}$ $S = -9.567 + 0.606\,C + 0.636\,PI - 0.792\,w - 0.487\,\gamma_{dry} + 6.289\,I_L$ $S = -19.856 + 0.595\,C + 0.686\,PI - 0.769\,CEC$
5	$S = 25.202 + 0.643 \cdot W_L - 2.089 \cdot w_n$ $S = 29.692 + 0.914 \cdot PI - 2.089 \cdot w_n$

- groundwater level and sampling depth. They take into account the position of the clayey stratum relative to the aquifer in natural environment.

Based on the selected input data, a predictive swelling model was developed using three machine learning methods. Models were evaluated using the MSE parameter (Table 2) and the MSE box plot (Figure 2).

Analyzing the box plot, which displays the median, quartiles, maximum and minimum values, outliers for all three methods, it can be visually noted that the Random Forest method has the best distribution of values. Table 2 shows that the model was developed by the Decision tree method has the highest MSE and is 101.20, which indicates the absence of any relationship between the studied parameters.

Table 2 – Results of data processing by machine learning methods

Model	MSE value
Decision tree	101.20
Random Forest	50.10
Gradient Boosting	52.40

The Random Forest method has the lowest and, therefore, the best MSE indicator and is 50.10. The MSE of the model was developed using the Gradient Boosting method takes an intermediate value between the smallest and the largest MSE value and is 52.40.

Figure 3 shows the results of swelling prediction using the Random Forest method. The blue color

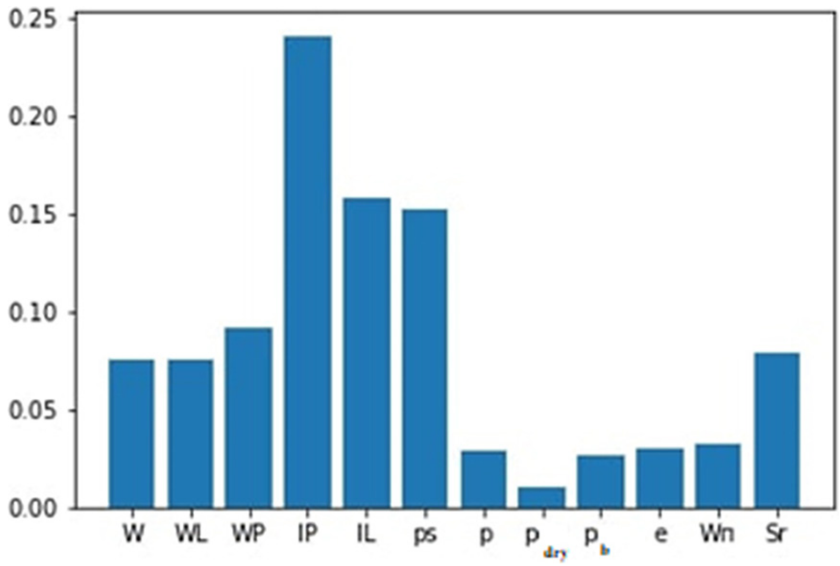
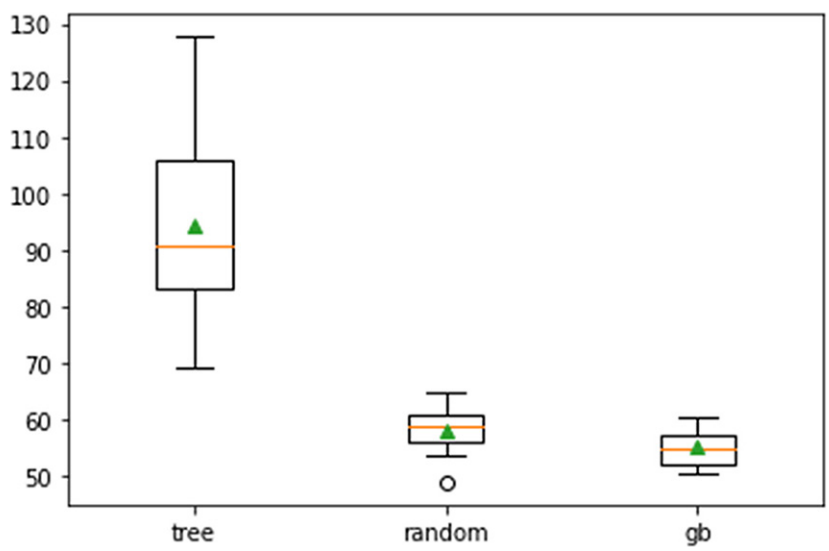


Figure 1 – Feature Selection for mathematical model development



tree – method «Decision tree», random – method «Random Forest», gb – method «Gradient Boosting»

Figure 2 – MSE box plot

indicates the measured values that were obtained through laboratory tests, the red color indicates the predicted values obtained as a result of developing a predictive model.

For plotting, 70 clay relative swelling values (20% of the data used for validation) were used, which were obtained as a result of developing a Random Forest model.

The scatterplot (Figure 4) shows the correlation between predicted and measured relative swelling values. The coefficient of determination (R^2) was 0.60. Correlation coefficient (R) – 0.77. This suggests that the relationship between the variables (relative swelling and physical, geological characteristics) is high.

Conclusions

Based on the results of this study, the following conclusions can be drawn:

- in this study, machine learning methods were used for modeling aimed at offering a new approach to predicting the relative swelling of clay rocks without laboratory studies: the relative swelling (ε_{sw}) was estimated based on geological characteristics (rock type, bedding rocks relative to the aquifer, rock age) and physical characteristics (w, W_L, W_P, I_P, I_L, p_s, S_r) for loams and clays of Karaganda city based on 350 samples obtained as a geotechnical investigation results;

- comparison of Decision tree, Random Forest, Gradient Boosting models based on MSE and

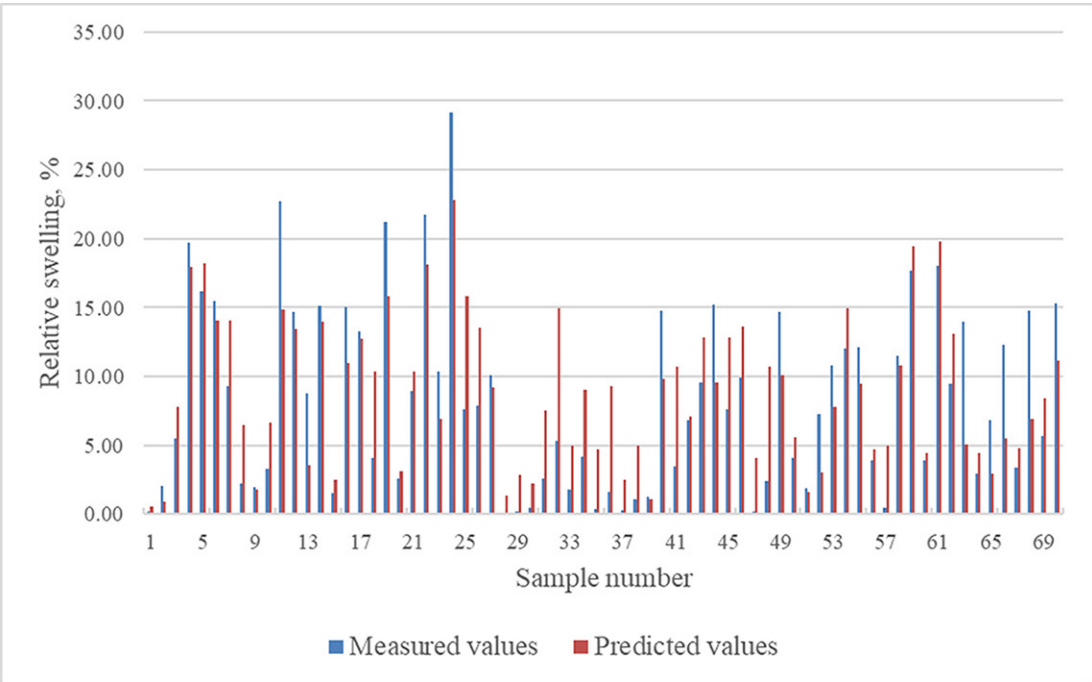


Figure 3 – Histogram comparing measured and predicted relative swelling values

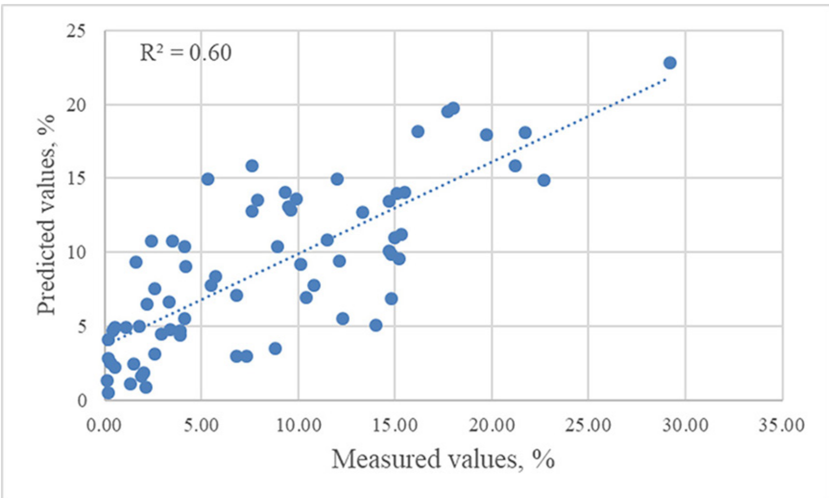


Figure 4 – Scatterplot of measured and predicted relative swelling values

scatterplot. To predict the relative swelling, the Random Forest method was chosen, as it has the lowest MSE and the best distribution in the scatterplot;

- this study showed that the model which was developed using the Random Forest method is the most effective for modeling the relative swelling of clay rocks compared to models developed using the Decision tree and Gradient Boosting methods;

- coefficient of determination, which amounted to 0.60, indicates that it is possible to apply the developed model in geotechnical investigation for the preparing of pre-project documentation.

REFERENCES

1. Gawriuczenkow I., Wojcik E. Prediction of swell pressure in Neogene clays from Warsaw, based on the swell index // Geol. Geophys. Environ, 2018. – Vol. 44, no. 2. – 219 p. https://doi.org/fvv7.
2. Mohamed B., Mekhti D., Abdelhamid C. A method for predicting the deformation of swelling clay soils and designing shallow foundations that are subjected to uplifting // Acta Geotech. Slov., 2016. – Vol. 13, no. 1. – 65-75 pp.
3. Al-Mhaidib A.I. Swelling behaviour of expansive shales from the middle region of Saudi Arabia // Geotech. Geol. Eng., 1998. – Vol. 16, no. 4. – 291-307 pp. https://doi.org/cbkt9x.
4. Driscoll R C.M. Subsidence damage to domestic buildings: lessons learned and questions remaining. BRE Press, 2000. – 32 p.
5. Butscher C. Swelling potential of clay-sulfate rocks in tunneling in complex geological settings and impact of hydraulic measures assessed by 3D groundwater modeling // Eng. Geol. Elsevier B.V., 2017. – Vol. 221. – 143-153 pp. https://doi.org/f95982.
6. Akt rassledovaniya prichin obrusheniya doma no. 7 v mikrorajone Besoba [Electronic resource] // news. 2012. URL: http://ekaraganda.kz/?mod=news_read&id=5678 (дата обращения: 12.05.2022).
7. Stroitel`ny`e normy` i pravila: SN RK 5.01-02-2013 – Osnovaniya zdanij i sooruzhenij. Astana: AO «KazNIISA», 2015. 20 p.
8. Ermias B., Vishal V. Application of Artificial Intelligence for Prediction of Swelling Potential of Clay-Rich Soils // Geotech. Geol. Eng. Springer International Publishing, 2020. – Vol. 38, no. 6. – 6189-6205 pp. https://doi.org/fvv9.
9. Sabtan A.A. Geotechnical properties of expansive clay shale in Tabuk, Saudi Arabia // J. Asian Earth Sci., 2005. – Vol. 25, no. 5. – 747-757 pp. https://doi.org/bxd7jn.
10. Erzin Y. Artificial neural networks approach for swell pressure versus soil suction behaviour // Can. Geotech. J., 2007. – Vol. 44, no. 10. 1215-1223 pp. https://doi.org/dnwvzn.
11. Zumrawi M.M.E. Prediction of swelling characteristics of expansive soils, 2012. – Vol. 1, no. 2. – 55-62 pp.
12. Erzin Y., Gunes N. The unique relationship between swell percent and swell pressure of compacted clays // Bull. Eng. Geol. Environ, 2013. Vol. 72, no. 1. – 71-80 pp. https://doi.org/fvwb.

Сазды тау жыныстарының ісінуінің олардың геологиялық және физикалық сипаттамаларына тәуелділігі

[1]***ВДОВКИНА Дарья Игоревна,** магистр, докторант, daryavdovkina@gmail.com,*
[1]***ПОНОМАРЕВА Марина Викторовна,** т.ғ.к., доцент, mv_ponomareva18@mail.ru,*
[1]***ПОНОМАРЕВА Екатерина Вадимовна,** PhD, аға оқытушы, evmussina1992@gmail.com,*
[1]*«Әбілқас Сағынов атындағы Қарағанды техникалық университеті» КеАҚ, Қазақстан, Қарағанды, Н. Назарбаев даңғылы, 56,*
**автор-корреспондент.*

***Аңдатпа.** Құрылыс саласының белсенді дамуы белгілі бір топырақтардан тұратын аумақтардың қамтылуына әкеледі. Салынған құрылымдардың қабаттылығы мен күрделілігі артып келеді. Шектеулі ақша және уақытша ресурстар құрылыс алаңын алдын-ала барлаудың тиімді әдістерін қажет етеді. Осыған байланысты инженерлік-геологиялық зерттеулер жүргізуге қойылатын талаптар артуда. Қожыма топырақтардан тұратын аумақтар ерекше назар аударуды қажет етеді. Ісінген сазды тау жыныстар ылғалданған кезде олардың көлемін өзгерту қасиетіне ие және осылайша құрылымдар мен іргетастардың тұрақтылығына айтарлықтай қауіп төндіреді. Ісіну сипаттамаларын өлшеу ұзақ уақытты қажет етеді және қаражаттың айтарлықтай үлесін алады. Бұл зерттеудің мақсаты машиналық оқыту әдістерін қолдана отырып, олардың салыстырмалы ісінуінен ең аз ресурстар жұмсалатын саз балшықты топырақтардың геологиялық және физикалық сипаттамаларының байланысын анықтау болып табылады. Сазды тау жыныстарының салыстырмалы ісіну моделін құру үшін кіріс деректер ретінде топырақтың физикалық сипаттамалары пайдаланылды, мысалы, табиғи ылғал, аққыштық шекарасындағы ылғал, жаймалап жазу шекарасындағы ылғал, иілімділік саны, аққыштық көрсеткіші, топырақ бөлшектерінің тығыздығы, топырақтың тығыздығы, құрғақ топырақтың тығыздығы, қалыпты жағдайдағы тығыздық, кеуектілік коэффициенті, жалпы ылғал сыйымдылығы, ылғалдылық дәрежесі. Осы сипаттамаларға сәйкес машиналық оқытудың келесі әдістерімен салыстырмалы ісінудің үш моделі жасалды: Decision tree, Gradient Boosting, Random Forest. Модельдерді бағалау болжамды және шынайы мәндер арасындағы орташа квадраттық ауытқуды өлшейтін Mean Squared Error параметрі бойынша жүргізілді. Осы критерийді бағалау нәтижесінде Random Forest моделі салыстырмалы ісіну моделін құру үшін таңдалды.*

Кілт сөздер: *сазды тау жыныстар, монтмориллонит саздар, ісіну, Random Forest моделі, төрттік шөгінділер, неогендік шөгінділер.*

Зависимость набухания глинистых пород от их геологических и физических характеристик

[1]****ВДОВКИНА Дарья Игоревна,*** *магистр, докторант, daryavdovkina@gmail.com,*
[1]***ПОНОМАРЕВА Марина Викторовна,*** *к.т.н., доцент, mv_ponomareva18@mail.ru,*
[1]***ПОНОМАРЕВА Екатерина Вадимовна,*** *PhD, старший преподаватель, evmussina1992@gmail.com,*
[1]*НАО «Карагандинский технический университет имени Абылкаса Сагинова», Казахстан, Караганда, пр. Н. Назарбаева, 56,*
**автор-корреспондент.*

Аннотация. *Активное развитие строительной отрасли ведет к тому, что охватываются территории, сложенные специфическими грунтами. Увеличивается этажность и сложность возводимых конструкций. Ограниченные денежные и временные ресурсы требуют более эффективных методов предварительной разведки застраиваемой площадки. В связи с этим повышаются требования к проведению инженерно-геологических исследований. Особого внимания требуют территории, сложенные набухающими грунтами. Набухающие глинистые породы обладают свойством изменять свой объем при увлажнении и тем самым представляют серьезную угрозу для устойчивости конструкций и фундаментов. Измерение характеристик набухания требует длительного периода времени и значительного вклада денежных средств. Целью данного исследования является установление взаимосвязи геологических и физических характеристик глинистых грунтов, на получение которых затрачиваются минимальные ресурсы, от их относительного набухания с применением методов машинного обучения. В качестве входных данных для построения модели относительного набухания глинистых пород использовались физические характеристики грунтов, такие как естественная влажность, влажность на границе текучести, влажность на границе раскатывания, число пластичности, показатель текучести, плотность частиц грунта, плотность грунта, плотность сухого грунта, плотность во взвешенном состоянии, коэффициент пористости, полная влагоемкость, степень влажности. По данным характеристикам было построено три модели относительного набухания следующими методами машинного обучения: Decision tree, Gradient Boosting, Random Forest. Оценка моделей производилось по параметру Mean Squared Error, который измеряет среднеквадратичное отклонение между прогнозируемыми и истинными значениями. В результате оценки данного критерия, для построения модели относительного набухания была выбрана модель Random Forest. Наличие прогностической модели относительного набухания позволит сэкономить денежные и временные ресурсы на предварительной стадии инженерно-геологических исследований.*

Ключевые слова: *глинистые породы, монтмориллонитовые глины, набухание, модель Random Forest, четвертичные отложения, неогеновые отложения.*

REFERENCES

1. Gawriuczenkow I., Wojcik E. Prediction of swell pressure in Neogene clays from Warsaw, based on the swell index // Geol. Geophys. Environ, 2018. – Vol. 44, no. 2. – 219 p. https://doi.org/fvv7.
2. Mohamed B., Mekhti D., Abdelhamid C. A method for predicting the deformation of swelling clay soils and designing shallow foundations that are subjected to uplifting // Acta Geotech. Slov., 2016. – Vol. 13, no. 1. – 65-75 pp.
3. Al-Mhaidib A.I. Swelling behaviour of expansive shales from the middle region of Saudi Arabia // Geotech. Geol. Eng., 1998. – Vol. 16, no. 4. – 291-307 pp. https://doi.org/cbkt9x.
4. Driscoll R C.M. Subsidence damage to domestic buildings: lessons learned and questions remaining. BRE Press, 2000. – 32 p.
5. Butscher C. Swelling potential of clay-sulfate rocks in tunneling in complex geological settings and impact of hydraulic measures assessed by 3D groundwater modeling // Eng. Geol. Elsevier B.V., 2017. – Vol. 221. – 143-153 pp. https://doi.org/f95982.
6. Akt rassledovaniya prichin obrusheniya doma no. 7 v mikrorajone Besoba [Electronic resource] // news. 2012. URL: http://ekaraganda.kz/?mod=news_read&id=5678 (дата обращения: 12.05.2022).
7. Stroitel`ny`e normy` i pravila: SN RK 5.01-02-2013 – Osnovaniya zdanij i sooruzhenij. Astana: AO «KazNIISA», 2015. 20 p.
8. Ermias B., Vishal V. Application of Artificial Intelligence for Prediction of Swelling Potential of Clay-Rich Soils // Geotech. Geol. Eng. Springer International Publishing, 2020. – Vol. 38, no. 6. – 6189-6205 pp. https://doi.org/fvv9.
9. Sabtan A.A. Geotechnical properties of expansive clay shale in Tabuk, Saudi Arabia // J. Asian Earth Sci., 2005. – Vol. 25, no. 5. – 747-757 pp. https://doi.org/bxd7jn.
10. Erzin Y. Artificial neural networks approach for swell pressure versus soil suction behaviour // Can. Geotech. J., 2007. – Vol. 44, no. 10. 1215-1223 pp. https://doi.org/dnwvzn.
11. Zumrawi M.M.E. Prediction of swelling characteristics of expansive soils, 2012. – Vol. 1, no. 2. – 55-62 pp.
12. Erzin Y., Gunes N. The unique relationship between swell percent and swell pressure of compacted clays // Bull. Eng. Geol. Environ, 2013. Vol. 72, no. 1. – 71-80 pp. https://doi.org/fvwb.