Proceedings 10th NUMGE 2023
10th European Conference on Numerical Methods in Geotechnical Engineering
Zdravkovic L, Kontoe S, Taborda DMG, Tsiampousi A (eds)

https://doi.org/10.53243/NUMGE2023-92

The prediction of soil cracking caused by desiccation using artificial intelligence and statistical analysis

A. Baghbani[1], M.J. Yaghoubi[2], F. Daghistani[3] S. Costa[1], T. Baumgartl[4]

[1] *Deakin University, Geelong, VIC, Australia*

[2] *Australian Road Research Board, Melbourne, VIC, Australia*

[3] *La Trobe University, Melbourne, VIC, Australia*

[4] *Federation University, Churchill, VIC, Australia*

ABSTRACT: Desiccation cracking affects soil hydromechanical behaviour and can cause serious hazards. Several factors affect desiccation cracking. Multiple and non-linear factors make desiccation cracking difficult to predict. Artificial intelligence (AI) has proven effective for predicting parameters based on different numbers of variables, but rarely for predicting soil cracking. Based on drying tests and generated databases, this study investigated if AI methods could predict two-dimensional soil desiccation cracking. Multiple linear regression (MLR) and support vector machines (SVM) were used to predict two outputs, including vertical shrinkage and cracks and shrinkage intensity factor (CSIF). Based on four input parameters, soil thickness, liquid limit, plasticity index, and shrinkage limit, mathematical models were built. In MLR, coefficient of determination (R^2) and mean absolute error (MAE) were 0.803 and 2.396 for predicting CSIF, respectively, and 0.768 and 2.608 for predicting vertical shrinkage. CSIF was predicted with R^2 of 0.963 and MAE of 1.225, and vertical shrinkage with R^2 of 0.970 and MAE of 0.960 by SVM. The soil thickness was the most significant input parameter in the sensitivity analysis.

Keywords: Desiccation cracking; Support vector machines; Artificial intelligence; Multiple linear regression; Soil shrinkage

1 INTRODUCTION

Desiccation cracking is a natural phenomenon caused by soil moisture loss. Generally, this phenomenon occurs near the earth's surface in arid and semi-arid regions. Cracks can be characterized by a variety of dimensions, sizes, and depths. Cracks can affect the soil's physical and hydrological properties. These changes may pose dangers, such as landslides. Soil cracks have drawn the attention of engineers, geotechnical researchers, and other experts (Nguyen et al., 2023; Baghbani et al., 2023a).

Dry cracking can be influenced by a number of factors (Tang et al., 2008; Prat et al., 2008; Sahebzadeh et al., 2017; Baghbani et al., 2022a). One of the influencing factors is the thickness of the soil layer. Tang et al. (2008) conducted a series of drying tests on eight groups of clay soil samples. With an increase in soil thickness, the mean length, width, aggregate area and crack intensity factor CIF increase. As sample thickness increases, Prat et al. (2008) found that the surface crack ratio increases, whereas the total crack length decreases. In addition, soil type and plasticity index play a significant role in soil desiccation cracking. During drying, clay-rich soils are more susceptible to volumetric shrinkage strains (Baghbani et al., 2022b). According to Yesiller et al. (2000), the soils with the highest fines content and highest plasticity index have the highest crack intensity factor (CIF) and crack width.

Multiple effective parameters in drying cracks and their nonlinear effects have made it challenging to predict drying cracks and develop a mathematical model for soil cracking. Artificial intelligence is widely used for predicting parameters without prior knowledge based on various parameters. Geotechnical engineering has extensively used artificial intelligence for slope stability (Baghbani et al., 2022b), soil dynamic (Baghbani et al., 2023b, Daghistani et al., 2023) and foundation (Samek et al., 2021; Samui and Sitharam, 2011). To predict the number of soil cracks, Choudhury and Costa (2019) used artificial neural networks (ANNs). Choudhury and Costa (2019) used a 16-set database with initial moisture content, specimen layer thickness, and specimen size as inputs. ANN predicted the training database with 0.88 accuracy (correlation coefficient or R) and 0.023 error (Mean absolute error or MAE), and the test database with R 0.93 and MAE 0.092. To predict crack number, Baghbani et al. (2022a) performed a series of drying tests, combined the results with existing studies in the literature, and developed a database of 31 sets. The specimen layer thickness, moisture content, and size of the specimen were input into this database. Baghbani et al. (2022d) found that the CART model could predict the number of drying cracks with an R^2 of 0.989 and RMSE of 1.285 for the testing database.

Despite existing research, artificial intelligence has not been used to predict two-dimensional drying cracks in soil. In addition, in the last two studies, the output parameter was the number of cracks, whereas the cracks and shrinkage intensity factor (CSIF) parameter is a better indicator of soil cracking. In this paper, a statistical method, namely multiple linear regression (MLR), and an artificial intelligence method, namely support vector machines (SVM), are used for the first time to predict the CSIF for clays. A database containing 15 sets from drying tests was used for this purpose. Database inputs were soil thickness, liquid limit, plasticity index and shrinkage limit, and database outputs were CSIF and vertical shrinkage.

2 METHODOLOGY

2.1 Materials

Three types of clay have been used in this study. Clay 1 and 2 were taken from Yallourn mine overburden in Victoria, Australia. Clay 3 was collected from North Warragul, located in eastern Victoria. The liquid limit (LL), plasticity index (PI) and linear shrinkage (LS) were measured according to Australian standard (AS, 2008). Table 1 shows the Atterberg limits results of these three clays.

Table 1. Results of Atterberg limit tests for studied clays.

	Clay 1	**Clay 2**	**Clay 3**
LL	57.5	64.9	65.7
PI	33.4	46.0	11.9
LS	12.4	19.2	14.0

2.2 Desiccation cracking test

An optical approach was used to study crack formation. Figure 1 shows how two fans were used to speed up the drying process, and a scale was used to measure the water loss. For the purposes of image analysis, an automated camera was installed one meter above the sample, taking photographs in various stages (every 30 minutes) of crack initiation and propagation.

Five different test groups were conducted using five different moulds with thicknesses of 5, 10, 15, 20, and 25 mm and the same diameter of 100 mm. In this test, the soil samples had an initial water content equal to the liquid limit of clays. Petroleum jelly was applied to the sides of the moulds, which allowed clay to freely separate from the mould walls. After filling the mould, the samples were tapped to prevent air bubbles from forming. Every 30 minutes, a calliper was used to measure the vertical shrinkage of the dry sample.

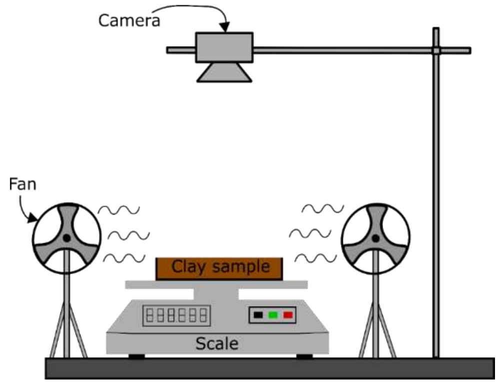


Figure 1. Schematic drawing of the drying test setup

2.3 Shrinkage and cracks quantification

Images were analysed to determine the initial clay area, dry clay area, crack and shrinkage area. The cracks and shrinkage intensity factor (CSIF) was calculated using Equation 1. These values were measured when the clay sample was fully dry at the end of each test.

$$CSIF = \frac{Cracks\ and\ shrinkage\ area\ (CSA)}{Initial\ clay\ area} \tag{1}$$

2.4 Mathematical modelling

2.4.1 Multiple linear regression (MLR)

Multiple linear regression (MLR) is one of the most widely used regression methods. It is one of the statistical methods to be studied before applying other statistical or machine-learning methods. Using two or more independent input variables, this method predicts an output.

2.4.2 Support vector machines (SVM)

A support vector machine (SVM) is one of the most widely used methods in AI. Based on mathematical equations, Boser et al. (1992) introduced a hyperplane method to separate data input nodes. The performance of this method depends on the hyperplane location. When hyper-planes achieve the most positive vectors and separate the most data points, they are most effective.

3 RESULTS AND DISCUSSION

3.1 Experimental

At the end of each test, when the clay samples were fully dried, a photograph was taken to calculate the cracks and shrinkage of the clay samples. Figure 2 illustrates photos of dry clays under different conditions. According to the results, clay 1 and 3 did not crack at initial thicknesses of 15 to 25 mm, and shrinkage occurred only in clays. In addition, the number of cracks decreased as the initial thickness increased. Furthermore, a thin layer of clay was found to contain a significant number of thin cracks, while a thicker layer contained a greater number of large cracks.

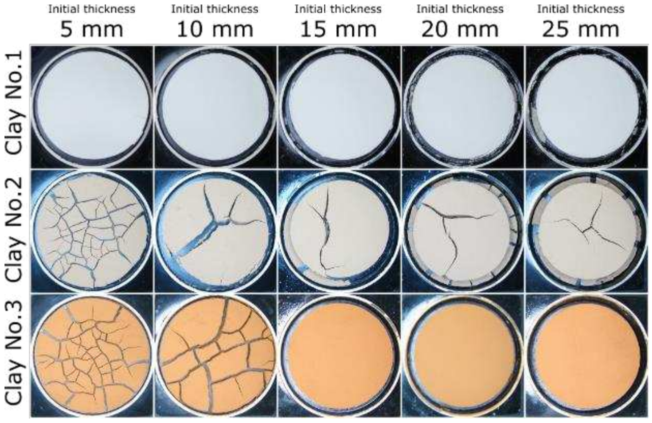
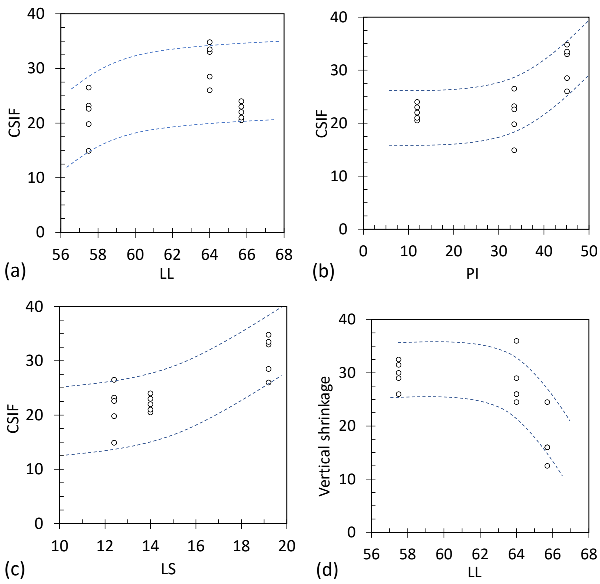


Figure 2. Desiccation cracking test results

Figure 3 shows CSIF and vertical shrinkage parameters as a function of soil type (LL, PI, LS). Figure 3a shows CSIF increases with increasing LL. A slope analysis of this graph reveals that the slope is higher at the beginning and gradually decreases. CSIF increases with soil PI, as shown in Figure 3b. In this graph, the slope increases with PI increase. The effect of LS on CSIF is illustrated in Figure 3c. The results indicate that CSIF increases with LS. In this graph, the slope almost rises as LS increases. Figure 3d shows the effect of LL on vertical shrinkage. This figure shows that vertical shrinkage decreases with increasing LL of the soil. According to Figures 3e and 3f, vertical shrinkage increases with PI and LS.



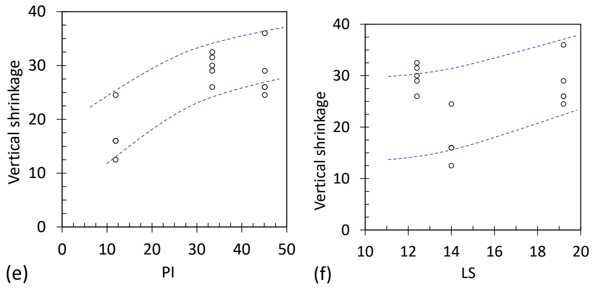


Figure 3. Effect of type of clays on CSIF and vertical shrinkage.

Figure 4 illustrates the effect of soil thickness on CSIF and vertical shrinkage. According to Figure 4a, CSIF has increased with the increase in soil sample thickness, while based on Figure 4b, vertical shrinkage has decreased.

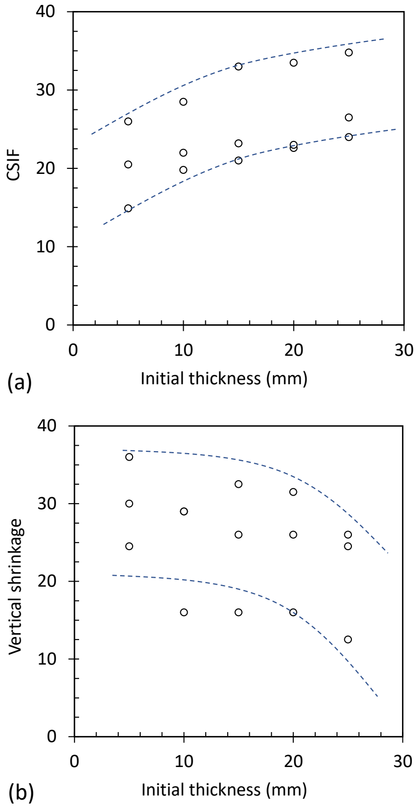


Figure 4. Effect of sample initial thickness on CSIF and vertical shrinkage.

3.2 Data processing

Database inputs and outputs in this study had different measurement units. This issue may affect the accuracy of artificial intelligence models. This issue has been addressed by using linear normalization. Based on Equation 3, all input and output values will be transferred between zero and one.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3)$$

This equation has four terms, X_{max}, X_{min}, X and X_{norm}, which are the maximum, minimum, actual and normalized values, respectively.

The next step was to randomly classify the existing databases into two categories: training and test databases. Table 2 shows the values of important statistical parameters, such as maximum, minimum, mean, and standard division, for these two databases. The table indicate that the statistical characteristics of two databases are similar, which increases the model's accuracy.

Table 2. Results of Atterberg limit tests for studied clays.

Variable	Observations		Minimum		Maximum		Mean		Std. deviation	
Database	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**	**Train**	**Test**
Vertical shrinkage	12	3	16.0	12.5	36.0	29.0	24.7	21.5	20.2	14.8
CSIF	12	3	14.9	24.0	33.5	34.8	23.5	30.6	4.7	5.8
LL	12	3	57.5	64.0	65.7	65.7	61.9	64.6	3.9	1.0
PI	12	3	11.9	11.9	45.1	45.1	29.2	34.0	13.6	19.2
LS	12	3	12.4	14.0	19.2	19.2	14.6	17.5	2.8	3.0
Initial thickness	12	3	5.0	15.0	25.0	25.0	13.3	21.7	6.8	5.8

Table 3. Results of Pearson parameter for all parameters.

Parameters	LL	PI	LS	Initial thickness	CSIF	Vertical shrinkage
LL	**1**	-0.436	0.558	-0.203	0.268	-0.677
PI	-0.436	**1**	0.503	0.013	0.513	0.847
LS	0.558	0.503	**1**	-0.183	0.729	0.118
Initial thickness	-0.203	0.013	-0.183	**1**	0.460	-0.417
CSIF	0.268	0.513	0.729	0.460	**1**	0.231
Vertical shrinkage	-0.677	0.847	0.118	-0.417	0.231	1

3.3 Mathematical modelling results

In order to assess the accuracy of the proposed models, the coefficient of determination (R^2) and the mean absolute error (MAE) between the predicted and measured values have been calculated. A definition of the coefficient of determination (R^2) and mean absolute error (MAE) can be found in Equations 4 and 5.

$$R^2 = \left[\frac{\sum_N (X_m - \bar{X}_m)(X_p - \bar{X}_p)}{\sum_N (X_m - \bar{X}_m)^2 \sum_N (X_p - \bar{X}_p)^2}\right]^2 \quad (4)$$

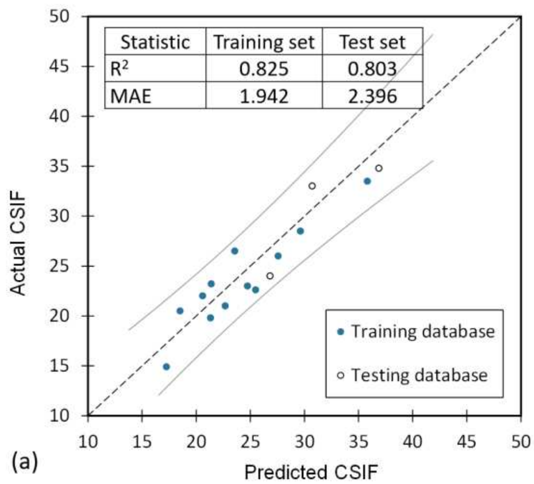
$$MAE = \frac{\sum_N |(X_m - X_p)|}{N} \quad (5)$$

where X_m, X_p, $\overline{X_p}$, $\overline{X_m}$ are actual values, predicted values, the average of actual values and the average of predicted values, respectively and N is the number of datasets. A model with a coefficient of determination (R^2) of 1 and a mean of absolute error (MAE) of 0 is considered to be the best model.

3.3.1 Multiple linear regression (MLR)

A summary of the results for Pearson parameter (r) for all input and output parameters can be found in Table 3. When this parameter approaches 1/-1, it indicates that the two parameters are linearly related. Despite the fact that if parameter A becomes 0, it means that the two discussed parameters do not have a linear relationship. According to Table 3, there is no strong linear relationship between any two parameters of the input and output parameters. Thus, it is expected that the linear regression method will be unable to predict with high accuracy the two output parameters, CSIF and vertical shrinkage, based on the input parameters of the database.

For both parameters CSIF and vertical shrinkage as well as for both test and training databases, Figure 5 displays the values predicted by the MLR against the actual values of the laboratory experiments output. Additionally, Figure 5 shows the accuracy of the MLR model in predicting both parameters CSIF and vertical shrinkage for both training and test databases. According to these results, MLR model was not capable of predicting both outputs, i.e., CSIF and vertical shrinkage, and more complex and powerful models, such as artificial intelligence (AI) models, are required.



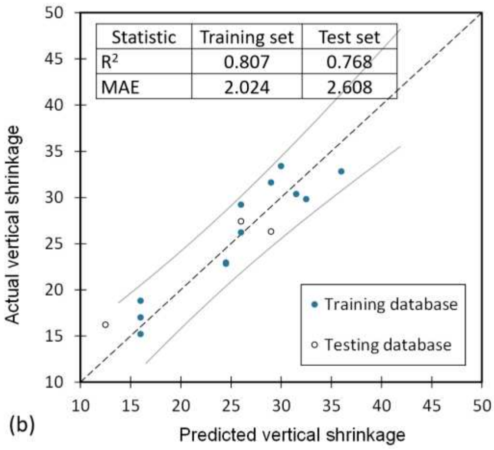


Figure 5. MLR results for predicting (a) CSIF and (b) vertical shrinkage

3.3.2 Support vector machines (SVM)

To find the most accurate model, influence parameters were changed extensively. Figure 6 shows the difference between the AI model's prediction and the actual values for both CSIF and vertical shrinkage. SVM model accuracy (R^2) and error (MAE) for training and test databases are also shown. According to the training database, the R^2 and MAE of the SVM model for predicting CSIF were 0.951 and 0.874, respectively, and for predicting vertical shrinkage, they were 0.974 and 0.599. According to the testing database, the best SVM model's accuracy (R^2) and error (MAE) for predicting CSIF were 0.963 and 1.225, respectively, and for predicting vertical shrinkage, R^2 and MAE were 0.970 and 0.960. Based on these results, it appears that SVM model predicted both parameters CSIF and vertical shrinkage accurately.

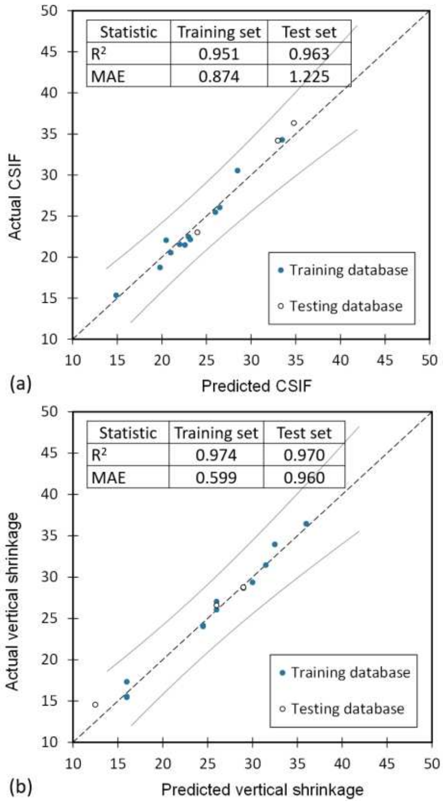


Figure 6. SVM results for predicting (a) CSIF and (b) vertical shrinkage

3.4 The variable importance of input parameters

As a result of developing the best artificial intelligence model, the importance of parameters has been investigated. For this purpose, each parameter has been increased or decreased by +-100% while other parameters have been left unchanged. The maximum error created is then calculated. In general, the greater the error for a parameter, the more sensitive the model will be to that parameter, and the greater the importance of that parameter. Figure 7 and Table 4 summarize the results of the parameter's importance. Based on the obtained results, the linear shrinkage parameter is the most important for predicting both output parameters, namely shrinkage and cracking. In addition, the results indicate that the liquid limit parameter was the least important of the four model inputs in predicting both outcomes.

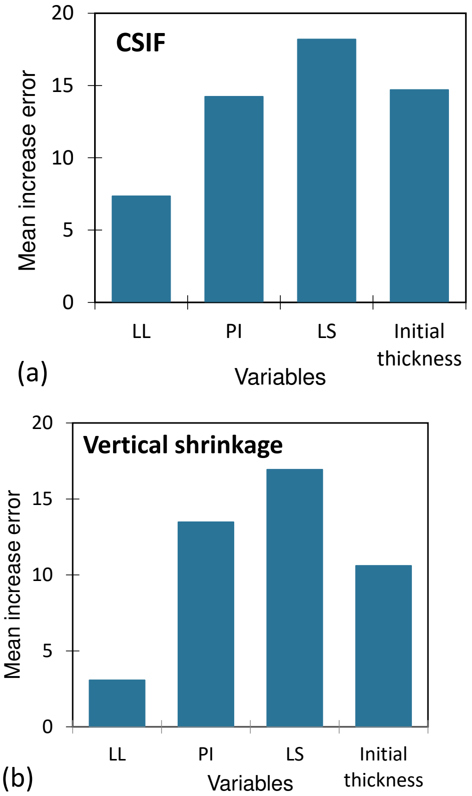


Figure 7. Effect of type of clays on CSIF and vertical shrinkage.

4 CONCLUSIONS

This study investigated crack patterns in three types of clayey soils and their impact on hydraulic properties, flow preference, and solute transport. Fifteen desiccation cracking tests were conducted on clay samples ranging from 5 mm to 25 mm in initial thickness. Photos were taken every 30 minutes during drying, and the crack and shrinkage intensity factor (CSIF) parameter was calculated using ImageJ software based on the final photo of each fully dried sample. Vertical shrinkage was also measured for each test. The experimental results were analyzed, and two models, SVM (AI) and MLR (statistical), were developed. The key findings of the study are as follows:

- Crack patterns in clayey soils have significant implications for hydraulic properties, flow preference, and solute transport.
- Thin layers of clay tend to have numerous thin cracks, while thicker layers exhibit fewer but larger cracks. Thinner clay samples show higher crack density, whereas thicker layers experience more shrinkage.
- The CSIF and vertical shrinkage increase with higher plasticity index (PI) and linear shrinkage (LS) values of the clays. However, vertical shrinkage decreases with increasing liquid limit (LL), while CSIF increases.
- The CSIF increases with an increase in the initial thickness of the clay sample, whereas vertical shrinkage decreases.
- The MLR model demonstrates reasonable accuracy in predicting both CSIF and vertical shrinkage, with an R^2 value of 0.803 for CSIF and 0.768 for vertical shrinkage in the testing database. The corresponding mean absolute errors (MAEs) are 2,396 and 2,608, respectively.
- The SVM model shows high accuracy in predicting CSIF and vertical shrinkage, outperforming the MLR model. It achieves an R^2 of 0.963 and an MAE of 1.225 for CSIF, and an R^2 of 0.970 and an MAE of 0.960 for vertical shrinkage in the testing database.
- Based on importance analysis, the parameter of linear shrinkage (LS) is identified as the most significant factor for both CSIF and vertical shrinkage prediction. Conversely, the parameter of liquid limit (LL) is found to have the least importance.

5 REFERENCES

AS 1289.3.4.1 2008. Soil classification test- Determination of the linear shrinkage of a soil Standard method 2008, *Standards Australia*.

Baghbani, A., Costa, S., Choundhury, T., Faradonbeh, R.S. 2022a. Prediction of Parallel Desiccation Cracks of Clays Using a Classification and Regression Tree (CART) Technique. In Proceedings of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR), Newcastle, Australia.

Baghbani, A., Choudhury, T., Samui, P., Costa, S. 2023a. Prediction of secant shear modulus and damping ratio for an extremely dilative silica sand based on machine learning techniques. *Soil Dynamics and Earthquake Engineering*, **165**, 107708.

Baghbani, A., Choudhury, T., Costa, S., Reiner, J. 2022b. Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. *Earth-Science Reviews*, **228**, 103991.

Baghbani, A., Abuel-Naga, H., Shirani Faradonbeh, R., Costa, S., Almasoudi, R. 2023b. Ultrasonic Characterization of Compacted Salty Kaolin–Sand Mixtures Under Nearly Zero Vertical Stress Using Experimental Study and Machine Learning. *Geotechnical and Geological Engineering*, **41**, 2987–3012.

Baghbani, A., Costa, S., O'Kelly, B.C., Soltani, A., Barzegar, M., 2023. Experimental study on cyclic simple shear behaviour of predominantly dilative silica sand. International Journal of Geotechnical Engineering, **17**(1), 91-105.

Boser, B.E., Guyon, I.M., Vapnik, V.N. 1992. A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory,* 144-152.

Choudhury, T., Costa, S. 2019. Prediction of parallel clay cracks using neural networks–a feasibility study. In Contemporary Issues in Soil Mechanics: *Proceedings of the 2nd GeoMEast International Congress and Exhibition on Sustainable Civil Infrastructures, Egypt 2018–The Official International Congress of the Soil-Structure Interaction Group in Egypt (SSIGE),* 214-224. Springer International Publishing.

Nguyen, M.D., Baghbani, A., Alnedawi, A., Ullah, S., Kafle, B., Thomas, M., Moon, E.M., Milne, N.A. 2023. Investigation on the suitability of aluminium-based water treatment sludge as a sustainable soil replacement for road construction. *Transportation Engineering*, **12**, 100175.

Prat, P., Ledesma, A., Lakshmikhanta, M.R., Levatti, H.U., Tapia, J. 2008, October. Fracture mechanics for crack propagation in drying soils. *12th International Conference on Computer Methods and Advances in Geomechanics* 2008. Leibniz Information Centre for Science and Technology University Library (TIB).

Sahebzadeh, S., Heidari, A., Kamelnia, H., Baghbani, A. 2017. Sustainability features of Iran's vernacular architecture: A comparative study between the architecture of hot–arid and hot–arid–windy regions. *Sustainability*, **9**(5), 749.

Samek, W., Montavon, G., Lapuschkin, S., Anders, C.J., Müller, K.R. 2021. Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, **109**(3), 247-278.

Samui, P., Sitharam, T.G. 2011. Machine learning modelling for predicting soil liquefaction susceptibility. *Natural Hazards and Earth System Sciences*, **11**(1), 1-9.

Tang, C., Shi, B., Liu, C., Zhao, L., Wang, B. 2008. Influencing factors of geometrical structure of surface shrinkage cracks in clayey soils. *Engineering geology*, **101**(3-4), 204-217.

Yesiller, N., Miller, C.J., Inci, G., Yaldo, K. 2000. Desiccation and cracking behavior of three compacted landfill liner soils. *Engineering geology*, **57**(1-2), 105-121