



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Shrink–swell index prediction through deep learning

This is the Published version of the following publication

Teodosio, Bertrand, Wasantha, PLP, Yaghoubi, Ehsan, Guerrieri, Maurice, an Staden, RC and Fragomeni, Salvatore (2022) Shrink–swell index prediction through deep learning. *Neural Computing and Applications*, 35. pp. 4569-4586. ISSN 0941-0643

The publisher's official version can be found at
<https://link.springer.com/article/10.1007/s00521-022-07764-7>
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/46823/>



Shrink–swell index prediction through deep learning

B. Teodosio¹ · P. L. P. Wasantha^{1,2} · E. Yaghoubi^{1,2} · M. Guerrieri^{1,2} · R. C. van Staden^{1,2} · S. Fragomeni^{1,2}

Received: 20 April 2022 / Accepted: 30 August 2022 / Published online: 26 October 2022
© The Author(s) 2022

Abstract

Growing application of artificial intelligence in geotechnical engineering has been observed; however, its ability to predict the properties and nonlinear behaviour of reactive soil is currently not well considered. Although previous studies provided linear correlations between shrink–swell index and Atterberg limits, obtained model accuracy values were found unsatisfactory results. Artificial intelligence, specifically deep learning, has the potential to give improved accuracy. This research employed deep learning to predict more accurate values of shrink–swell indices, which explored two scenarios; Scenario 1 used the features liquid limit, plastic limit, plasticity index, and linear shrinkage, whilst Scenario 2 added the input feature, fines percentage passing through a 0.075-mm sieve (%fines). Findings indicated that the implementation of deep learning neural networks resulted in increased model measurement accuracy in Scenarios 1 and 2. The values of accuracy measured in this study were suggestively higher and have wider variance than most previous studies. Global sensitivity analyses were also conducted to investigate the influence of each input feature. These sensitivity analyses resulted in a range of predicted values within the variance of data in Scenario 2, with the %fines having the highest contribution to the variance of the shrink–swell index and a relevant interaction between linear shrinkage and %fines. The proposed model Scenario 2 was around 10–65% more accurate than the preceding models considered in this study, which can then be used to expeditiously estimate more accurate values of shrink–swell indices.

Keywords Shrink–swell index · Reactive soils · Artificial intelligence · Deep learning · Sensitivity analysis

1 Introduction

The reactivity of soils is a characteristic that affects the mechanical properties of most clayey grounds [1]. Reactive soils undergo substantial volume changes in response to the variations in soil moisture content through swelling and shrinking. The shrink–swell ground movement leads to distresses concerning infrastructures built on and in the vicinity of reactive soils [2, 3]. The damages to lightweight structures such as pavements, underground pipes, and residential structures due to these shrink–swell ground movements are well known [4, 5]. The severe damage

brought by reactive soils has been recorded in Australia, China, Egypt, India, Israel, South Africa, the UK, and the USA, totalling a significant annual financial loss [6, 7]. In the UK, the impact of reactive soils summed up to £3 billion from 1991 to 2001 due to the effect of droughts, making it the most damaging geohazard in the region [8]. In the USA, the cumulative rehabilitation costs were more than twice the financial loss incurred from natural disasters due to floods, hurricanes, tornadoes, and earthquakes, amounting to around US\$ 15 billion per year [8]. Li et al. [9] found that the damage caused by reactive soils in Australia was mostly to lightweight structures even though no combined estimates are reported in the literature. Approximately 20% of the Australian land can be categorised as moderately to very highly reactive soils, with six out of eight major cities being affected, causing geohazards to structures and infrastructures [6].

The shrink–swell soil index (I_{ss}) is a soil parameter commonly used to determine the potential characteristic surface movement (y_s) of sites having reactive soils in

✉ B. Teodosio
bertrand.teodosio@vu.edu.au

¹ Institute of Sustainable Industries and Liveable Cities, Victoria University, Melbourne, VIC 3011, Australia

² College of Engineering and Science, Victoria University, Melbourne, VIC 3011, Australia

Australia [10]. This index is determined through laboratory testing using undisturbed soil samples collected from the field. The I_{ss} index has been used in Australia for more than 15 years and the Australian Standards, AS 1289 7.1.1, provides standardised testing procedures. Estimates of y_s have generally been successful in determining the dimensions of residential slabs and footings.

Determining the value of I_{ss} requires the collection of undisturbed soil samples for the swell test and the simplified core shrinkage test. Undisturbed soil sampling costs relatively higher and is more difficult to implement than disturbed sampling. Determining I_{ss} takes a longer time to obtain results, which can take more than four days involving around two hours of hands-on experimental work depending on the skill level of the individual performing the tests [1]. This can be a significant waiting period for most practitioners and researchers, and the results are sensitive to instrumental conditions, skills and experience of the tester, and changing ambient conditions.

Several studies had attempted to correlate I_{ss} with other soil properties such as the Atterberg limits (i.e. liquid limit (LL), plasticity limit (PL), plasticity index (PI), and linear shrinkage (LS)) to estimate I_{ss} indirectly. However, most studies have found sub-par correlation coefficients ($R^2 < 0.80$) between I_{ss} and the Atterberg limits, which ranged from 0.20 to 0.53 [9]. For instance, Young and Parmar [11] performed more than 300 laboratory tests to correlate I_{ss} with more common soil indices such as gravimetric soil moisture content (ω), LL, PL, PI, and LS that resulted in low correlation factors. Earl [12] suggested that the Atterberg limits alone may not be reliable to estimating values of I_{ss} based on his investigations using clay samples from the Shepparton geologic formation. Reynolds [13] performed similar correlation analyses for a dataset of clay samples collected from Central, Southeast, and Northwest Queensland for a pavement design application and also reported weak relationships. Similar weak correlations were found in the investigations conducted by Zou [14] and Li et al. [9], where soil samples were collected from 47 study sites from 37 suburbs in Melbourne. However, these investigations employed simple univariate regressions that limit the capturing of nonlinear relationships between extracted features or variables. Contrarily, Jayasekera and Mohajerani [15] found a relatively stronger correlation with R^2 values that ranged from 0.85 to 0.91. However, their investigation had a low variance dataset, with I_{ss} values that varied from 3.8 to 5.5, which limited the predictive capacity of their model [15].

Recent advances in data engineering and data science have expanded the application of artificial intelligence (AI) techniques to many disciplines [16]. AI refers to the ability of a machine or robot to display intelligence comparable to humans by learning through experiences in performing a

specific task with improving measured performance [17]. Machine learning (ML) is a subset of AI referring to algorithms capable of learning and improving performance without explicit programming or hard coding (Fig. 1a). ML tasks include recognising objects, understanding speech, responding to a conversation, solving problems, optimising solutions, greeting people, and driving a vehicle [18–20]. Rumelhart et al. [21] initially proposed shallow learning that initiated ML applications (Fig. 1b). These shallow neural networks restrict algorithmic support and are unable to train multiple hidden layers due to limitations in the computing power and available data [22].

Recent applications of AI in geotechnical engineering include geotextile [23, 24], tunnelling [25], geothermal energy [26], unsaturated flow [27], geo-structural health monitoring [28, 29], liquefaction [30], nanotechnology [31], carbon sequestration [32], and soil properties and behaviour prediction [33–35]. The ML techniques applied in these past investigations include artificial neural network (ANN), support vector machine (SVM), genetic algorithms (GA), fuzzy logic, image analysis, and adaptive neuro-fuzzy inference systems (ANFIS). One of the emerging ML techniques in geotechnical engineering is deep learning (DL), an implementation of ANN with multiple hidden layers as presented in Fig. 1c. It allows computation of more complex features of the input layer [36, 37]. Each hidden layer computes a nonlinear transformation of the preceding layer [38]. This deep network can have a substantially greater representation of the extracted features that can learn significantly more complex functions than a shallow network [22]. However, understanding the implementation of DL can be a challenge due to its complex network. It is beneficial for users to understand the implementation of algorithms to enable the accurate and confident application of complex models such as DL [22, 36, 38].

The application of AI to investigate the properties and nonlinear behaviour of reactive soil is currently not well explored, although there has been an increased application of AI in geotechnical engineering in general [39]. As linear correlations between I_{ss} and Atterberg limits had potential, most model accuracy and data ranges of the previous studies found sub-par results. DL has the potential to give better accuracy in the prediction of I_{ss} using the Atterberg limits given the ability of neural networks to handle complex nonlinear scenarios. The application of Atterberg limits can be used by promoting adaptive nonlinear models and providing insightful findings [16]. Therefore, the objective of this study is to employ DL to predict more accurate values of I_{ss} using different combinations of soil properties including LL, PL, PI, LS, and %fines. This can contribute to an efficient process of calculating the maximum potential characteristic surface movement, y_s . This

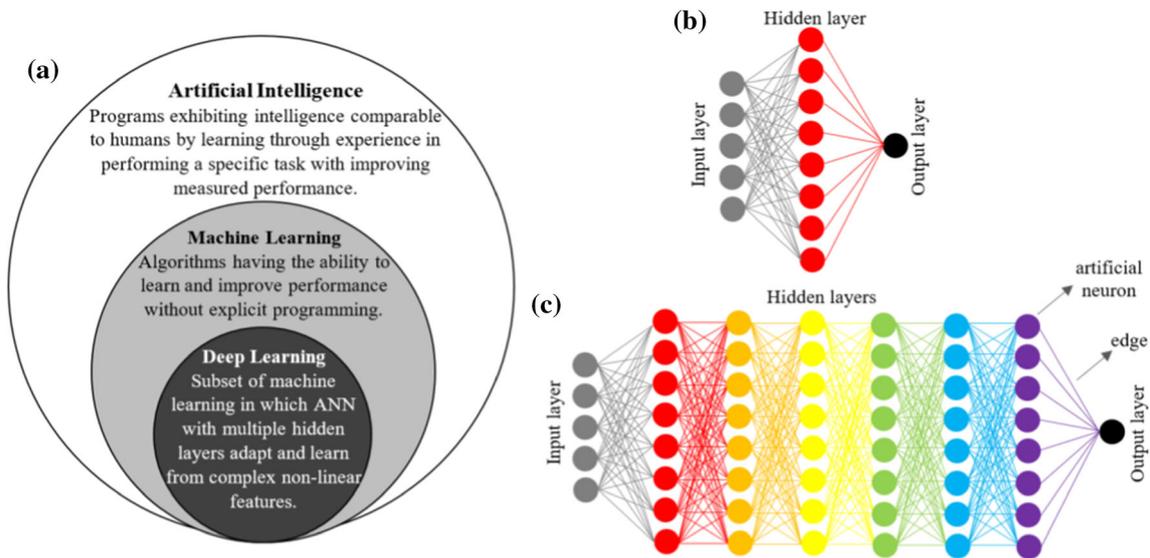


Fig. 1 Defining **a** the difference between artificial intelligence (AI), machine learning (ML), and deep learning (DL), **b** a sample shallow artificial neural network (ANN), and **c** a sample DL network

study also carries out a sensitivity analysis to identify the relative influence of each input variable, LL, PL, PI, LS, and %fines, on the targeted output, I_{ss} in %/pf. The sensitivity analyses elaborate on how the specific DL prediction mechanism functions with respect to the input features, increasing the applications.

2 Methodology

The concept of ANN is analogous to the neural network of a human brain in the way it processes information and establishes logical relationships [40]. A collection of connected nodes called artificial neurons comprise a network comparative to those of a human brain. The artificial neurons are connected by links called edges to transmit signals from a single neuron to other neurons. These signals are represented by real numbers. Each node and edge have weights that serve as a correlation factor that adjusts signals as learning occurs. The main function of this network is to obtain the lowest value of a loss or cost function, $L(y, \hat{y})$, that will give the optimum weights. The DL process is influenced by two main considerations: the architecture of the neural network and the learning process of the implemented algorithm.

2.1 Deep learning architecture

In this current study, a DL network was used to obtain the acceptable weights for I_{ss} prediction and comprised of an input layer, ten hidden dense layers, and an output layer. The input and the output values are the expected number of

input and output neurons, which indicate the size of the matrices for the calculation. The input layer contained the input vector extracted from a dataset and the number of artificial neurons in the input layer was determined by the input features extracted. In this study, two scenarios were considered. The first scenario, or Scenario 1, used the input features LL, PL, PI, and LS, whereas the second scenario, or Scenario 2, also used %fines.

Earl [12] and Reynolds [13] suggested that LL, PL, PI, and LS were not sufficient to be employed for estimating the values of I_{ss} . Thus, Earl [12] added clay and silt fraction, and Reynolds [13] included California Bearing Ratio (CBR), per-cent swell, and other polynomial features as inputs. The resulting values of R^2 were higher compared to those with Atterberg limits alone and ranged from 0.51 to 0.78 [13] and from 0.54 to 0.82 [12]. However, the data samples they used were limited ($n < 30$). In addition, the dataset had low variance ($0.1 < I_{ss} < 4.0$), and the models were not tested or validated for their prediction capacity. These improved outcomes led to the development of Scenario 2 of this study, which included %fines as an additional input to predict I_{ss} . The addition of the CBR by Reynolds [13] as an input did not result in greater R^2 compared to the addition of; thus, in the current study, the CBR was not considered.

In this current study, the number and size of hidden dense layers were determined by trial and error. The hidden dense layers connect each neuron, receiving input from its preceding layer. The dimension of the first five hidden dense layers was restrained to eight since greater dimensions resulted in the model calculation divergence (i.e. erratic and high values of calculated loss). The dimension

of the last five hidden dense layers was increased to 128 to generate more nonlinearity in the relationship between neurons. The increase in the number of hidden layers and the value of the dimension, depending on a specific scenario, often leads to improved accuracy [41]. The consequence is the time inefficiency in performing the DL algorithm, specifically for the training period to obtain the optimum weights. The output layer concludes the DL learning process using one neuron.

2.2 Deep learning process

The learning process of a DL neural network comprises five main stages, (1) pre-processing, (2) random initialisation, (3) forward propagation, (4) backward propagation, and (5) evaluation. The DL process implemented in this study is summarised in Fig. 2.

Pre-processing a dataset is an essential step that can increase the accuracy of a DL network training and

validation. The common pre-processing techniques applied to previous DL networks include the removal of data entries with outliers and missing values, creation of polynomial features, implementing feature scaling, and employing normalisation to a dataset [42].

The dataset used in this study was extracted from the five studies conducted by [12–14, 43, 44], as presented in Appendix 1. A total of 169 and 116 data entries were collected for Scenarios 1 and 2, respectively, with a summary description presented in Fig. 3 and Table 1. Outliers were determined using the interquartile range (IQR), calculated as

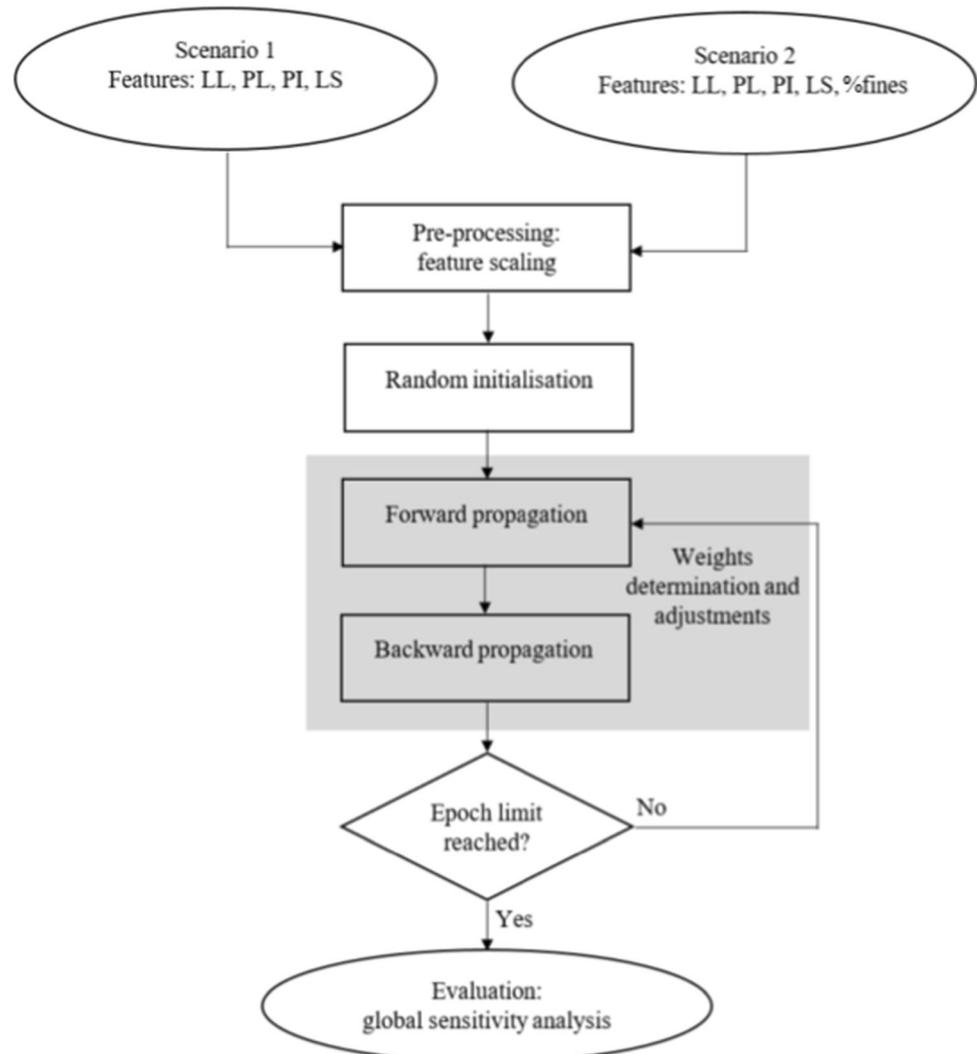
$$\text{Outlier}_{lb} = Q1 - 1.5\text{IQR} \quad (1)$$

$$\text{Outlier}_{ub} = Q3 + 1.5\text{IQR}, \quad (2)$$

where IQR is described as

$$\text{IQR} = Q3 - Q1, \quad (3)$$

Fig. 2 Deep learning (DL) neural network process for two scenarios



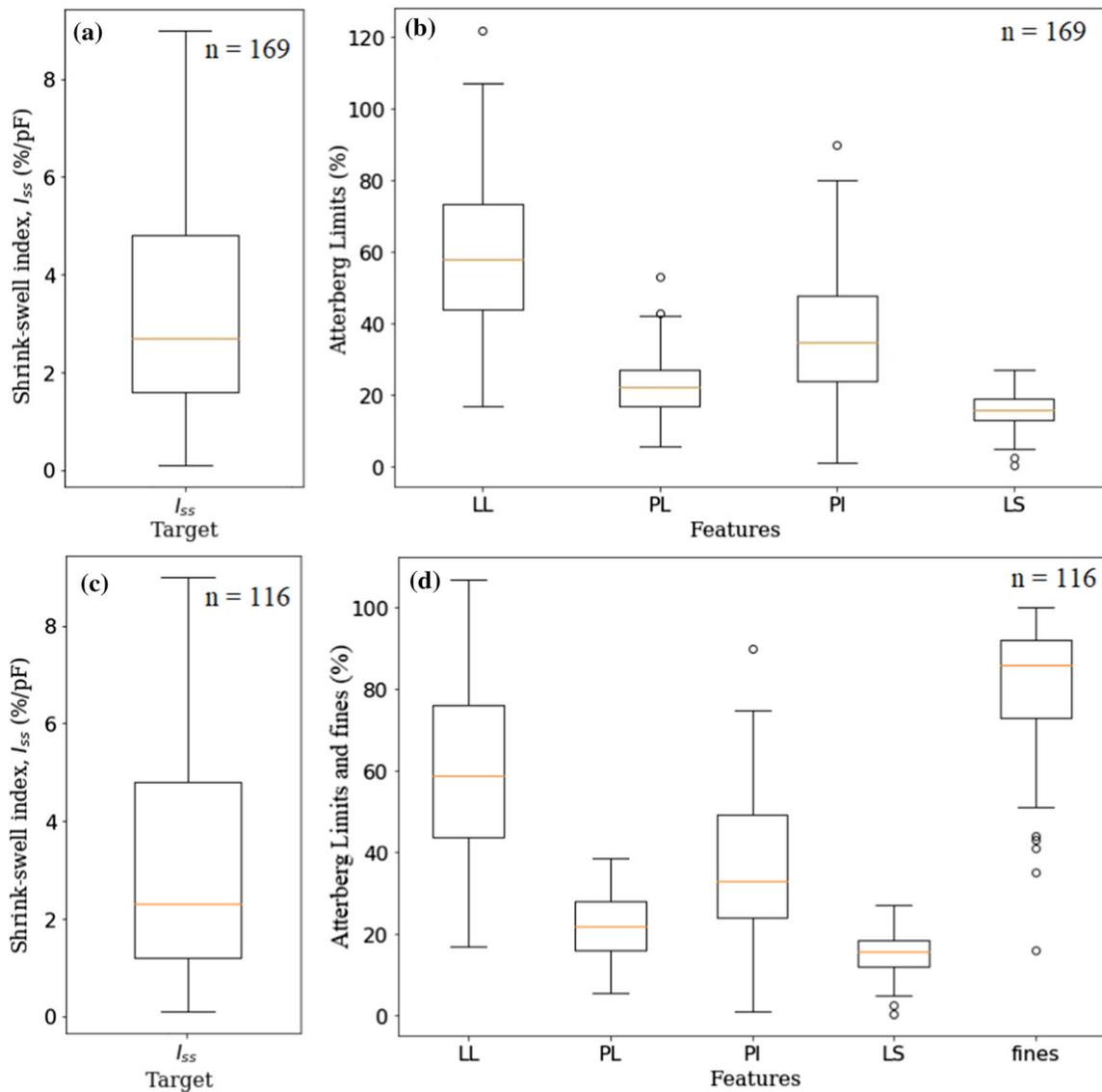


Fig. 3 Boxplot of the dataset for Scenario 1 showing **a** the target I_{ss} and **b** the features LL, PL, PI, and LS, and Scenario 2 showing: **c** the target I_{ss} and **d** the features LL, PL, PI, LS, and fines, where

I_{ss} = shrink–swell index, LL—Liquidity limit, PL = Plasticity limit, PI = Plasticity index, LS = Linear shrinkage, and n is the number of data entries or samples

where $Q1$ is the first quartile, $Q3$ is the third quartile, and subscripts lb and ub indicate the lower bound and the upper bound outliers.

The circles in Fig. 3 represent the outliers less than or more than the calculated values of $Outlier_{lb}$ and $Outlier_{ub}$. A comparison between DL with and without the outliers showed comparable results indicating a negligible effect of the outliers. Therefore, in this study, the complete dataset without removing the outliers was used. Data entries with missing values of I_{ss} , LL, PL, PI, LS, and %fines were omitted for Scenario 2. Polynomial features are commonly created to incorporate a nonlinear relationship between the target and the input features. Polynomial features are added to improve the accuracy of linear models with limited

features or when one feature is dependent on another. The use of polynomial features was initially implemented in the dataset but did not result in a noteworthy effect on the accuracy of the algorithm. Therefore, these features were not implemented in the DL of this study. The entire dataset was randomly split into two; one for training and the other for testing, with a ratio of 70–30%, as listed in Table 1. The 70–30% ratio was considered the most suitable division for training and validating neural network models with small dimensional datasets [45].

Two feature scalings, standard scaling and min–max scaling, were tested to check if the DL process would improve. The feature scaling using the standard scaling was

Table 1 Descriptive statistics of the target (I_{ss}) and features for Scenario 1 and Scenario 2

	Scenario 1 with four features: LL, PL, PI, and LS										
	Training data (70%)					Testing data (30%)					
	I_{ss} (%/pf)	LL (%)	PL (%)	PI (%)	LS (%)	I_{ss} (%/pf)	LL (%)	PL (%)	PI (%)	LS (%)	
Count	118	118	118	118	118	51	51	51	51	51	
\bar{x}	3.0	58.1	22.4	35.9	15.4	3.6	62.2	23.8	38.4	16.0	
s	1.9	19.1	6.8	14.8	4.3	2.4	21.9	8.5	16.6	5.2	
Min	0.1	25.0	11.0	11.0	5.0	0.1	17.0	5.6	1.0	0.5	
$Q1$	1.6	43.5	16.9	23.8	12.8	1.5	46.0	17.6	24.9	13.3	
Median	2.6	55.7	21.9	32.9	15.5	3.6	64.6	22.6	37.6	16.5	
$Q3$	4.2	71.1	26.9	45.6	18.3	5.1	76.5	29.8	49.3	19.6	
Max	7.4	122.0	42.0	90.0	27.0	9.0	107.0	53.0	75.0	25.2	

	Scenario 2 with four features: LL, PL, PI, LS, and %fines											
	Training data (70%)						Testing data (30%)					
	I_{ss} (%/pf)	LL (%)	PL (%)	PI (%)	LS (%)	%fines (%)	I_{ss} (%/pf)	LL (%)	PL (%)	PI (%)	LS (%)	%fines (%)
Count	81	81	81	81	81	81	35	35	35	35	35	35
\bar{x}	3.0	57.2	21.4	35.7	14.8	80.3	2.9	61.5	23.2	38.3	15.9	80.6
s	2.1	19.8	6.5	15.4	4.9	16.2	2.2	19.8	6.5	16.2	4.5	14.1
Min	0.1	17.0	5.6	1.0	0.5	16.0	0.5	31.0	14.0	13.8	7.0	41.0
$Q1$	1.2	42.0	16.0	23.0	11.0	73.0	1.2	46.2	17.5	28.0	13.2	73.0
Median	2.5	59.0	21.9	34.6	16.2	85.0	2.0	58.8	24.0	32.4	15.0	85.0
$Q3$	4.8	72.2	25.0	49.2	18.2	92.0	4.7	75.9	28.7	48.6	18.1	90.0
Max	9.0	105.0	36.0	70.0	27.0	100.0	8.9	107.0	38.6	75.0	27.0	99.0

Table 2 Upper and lower bounds to generate the input features for Scenarios 1 and 2 using the scheme by Saltelli [51]

	Input features for Scenario 1 and Scenario 2				
	LL (%)	PL (%)	PI (%)	LS (%)	%fines (%)
Lower bound (min)	15	5	0	0	1
Upper bound (max)	130	60	100	40	100

observed to be more applicable to the study and was employed using the following equation:

$$\text{Standardscaling} = \frac{x - \bar{x}}{s}, \quad (4)$$

where x is the data entry, \bar{x} is the mean value of a feature, and s is the standard deviation of a feature.

The values of \bar{x} and s for the input parameters LL, PL, PI, and LS are presented in Table 1, for both training and testing data. Applying normalisation was considered and implemented in preliminary model runs. However, normalising resulted in a lesser accuracy of the results, compared to standard scaling.

The random initialisation method of He et al. [46] was used to initialise the DL process. Following this method, the weights were randomly initialised with values close to

zero and then multiplied by $\sqrt{\frac{2}{\text{size}^{L-1}}}$, where size^{L-1} was the number of neuron in layer $L - 1$. Multiplying this term helps consider the nonlinearity of the activation functions. This initialisation proposed by He et al. [46] was specifically used together with the Rectified Linear Units (ReLU) activation, solving learning inefficiency and vanishing gradient issues. The loss function employed in the DL process is the mean squared error (MSE), which is a commonly used function for regression. The calculated loss is the mean overseen data of the squared differences between a true value in the dataset and a predicted value calculated by the DL algorithm described as

$$L(y, \hat{y}) = \frac{1}{N} \sum_{i=0}^N (y_i - \hat{y}_i)^2 \quad (5)$$

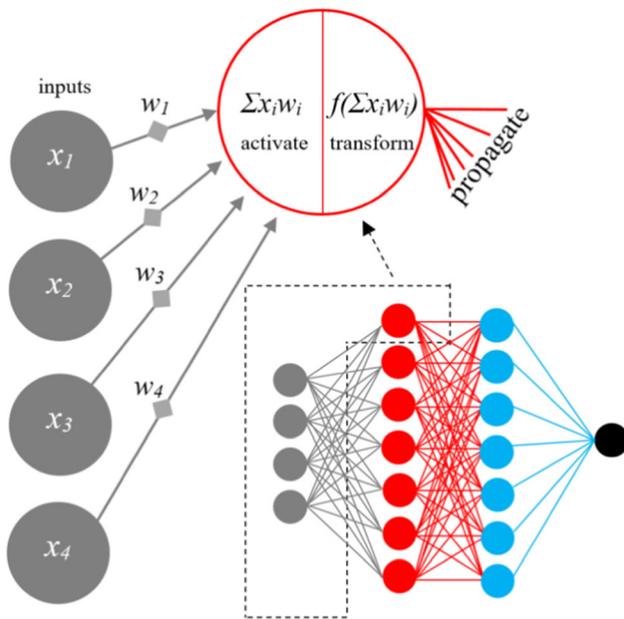


Fig. 4 Typical biologically inspired artificial neuron showing activation and transformation

where y_i is the actual value, \hat{y}_i is the predicted value, and N is the total number of data entries.

ReLU by Nair and Hinton [47] was implemented as the activation function for the forward and backward propagation, which reads

$$f(x) = \max(0, x) = \begin{cases} x_i, & \text{if } x_i \geq 0 \\ 0, & \text{if } x_i < 0 \end{cases} \quad (6)$$

where x_i is the input value of feature i .

A simplified representation of the usage of the activation function is presented in Fig. 4. Ridge Regression or L2 Regularisation was also used since the preliminary DL runs experienced overfitting. L2 regularisation adds a squared magnitude of coefficient as penalty term to the loss function defined as

$$L(y, \hat{y}) = \frac{1}{N} \sum_{i=0}^N (y_i - \hat{y}_i)^2 + \lambda \sum_{i=0}^N w_i^2 \quad (7)$$

where λ is a hyperparameter for regularisation, and w_i is a weight of a feature.

The value of λ is taken to be greater than zero. Taking the value of λ too high may lead to larger weights and underfitting. After fine-tuning the hyperparameter λ , the value was specified as 1.00 for Scenario 1 and 2.35 for Scenario 2.

The Adaptive Moment (Adam) estimation by Kingma and Ba [48] was used in the DL neural network. The Adam stochastic optimisation is widely used due to its benefits of straightforward implementation, computational efficiency, and lower required memory. The Adam method combines

the momentum gradient descent method and the Root Mean Squared Propagation (RMSprop), which is modelled as

$$Vdw = \beta_1 Vdw + (1 - \beta_1)dw, \quad (8)$$

$$Vdb = \beta_1 Vdb + (1 - \beta_1)db, \quad (9)$$

$$Sdw = \beta_2 Sdw + (1 - \beta_2)dw^2, \text{ and} \quad (10)$$

$$Sdb = \beta_2 Sdb + (1 - \beta_2)db^2, \quad (11)$$

where Vdw , Vdb , Sdw , and Sdb are the derivative of the weights and bias, which is being computed in iteration or epoch, t . The initial values of Vdw_i , Vdb_i , Sdw_i , and Sdb_i are assigned to zero and then calculated for each weight. The calculated values of Vdw , Vdb , Sdw , and Sdb are then corrected using the power of the current epoch, t , described below

$$Vdw^{\text{corrected}} = \frac{Vdw}{1 - \beta_1^t}, \quad (12)$$

$$Vdb^{\text{corrected}} = \frac{Vdb}{1 - \beta_1^t}, \quad (13)$$

$$Sdw^{\text{corrected}} = \frac{Sdw}{1 - \beta_2^t}, \text{ and} \quad (14)$$

$$Sdb^{\text{corrected}} = \frac{Sdb}{1 - \beta_2^t}. \quad (15)$$

Each weight and bias will be updated using the equations below

$$w = w - \alpha \frac{V_{dw}^{\text{corrected}}}{\sqrt{S_{dw}^{\text{corrected}} + \epsilon}} \text{ and} \quad (16)$$

$$b = b - \alpha \frac{V_{db}^{\text{corrected}}}{\sqrt{S_{db}^{\text{corrected}} + \epsilon}}. \quad (17)$$

The learning rates (α) for Scenarios 1 and 2 were taken as 7.5×10^{-5} and 5.0×10^{-5} after fine-tuning using trial and error. The decay rates for both scenarios were $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 5.0 \times 10^{-6}$. The forward and backward propagation was implemented in a loop until the specified epoch was achieved, as shown in the DL process in Fig. 2. Note that every iteration of the optimisation loop comprises forward propagation, cost calculation, backward propagation, and weights updating. The epoch of the final DL run was 500 since this value had resulted in an optimum and stable loss curve with an acceptable learning period, completing the deep learning processes in less than three minutes for each scenario.

2.3 Sensitivity analysis

Sensitivity analyses help identify the independent influence of input variables on a targeted output. There are two types

of sensitivity analyses; local and global approaches. A local sensitivity analysis assesses the local impact of feature variations concentrated on the sensitivity in the proximate vicinity of a set of feature values [49]. On the other hand, a global sensitivity analysis quantifies the overall importance of the features and their interactions with the predicted results by implementing a comprehensive coverage of input values [50]. This study used a global sensitivity analysis approach by implementing the method by Saltelli [51] and Sobol [52, 53].

The bounds to generate the input features were specified as LL = 15–130, PL = 5–60, PI = 0–100, LS = 0–40, and %fines = 1–100 listed in Table 1, and then the scheme by Saltelli [51] was implemented. This was based on the range of values presented in Fig. 3b. Three indices were calculated. The first one was the first-order Sobol index (S_1), calculated as [52, 53]

$$S_1 = \frac{\text{var}(x_i)}{\text{var}(\hat{y})} = \frac{\text{var}(E(\hat{y}|x_i))}{\text{var}(\hat{y})}, \quad (18)$$

where $\text{var}(x_i)$ is the variance of a feature, $\text{var}(\hat{y})$ is the variance of the target output, I_{ss} , E denotes expectation, x_i is a feature, and \hat{y} is the target output, I_{ss} .

The term $E(\hat{y}|x_i)$ in Eq. 18 indicates the expected value of the output \hat{y} when feature x_i is fixed. The first-order Sobol index, S_1 , reflects the expected reduction in the variance of the model when feature x_i is not changing. Thus, S_1 measures the direct effect of each feature on the variance of the model. It is worth noting that the sum of all the calculated values of S_1 should be equal to or less than one. It is common to perform the calculation of S_1 , and the total Sobol sensitivity S_T , which includes the sensitivity of first-order effects and the sensitivity due to interactions between a feature X_i and all other features [54] given by

$$S_T = 1 - \frac{\text{var}(E(\hat{y}|x_{-i}))}{\text{var}(\hat{y})}, \quad (19)$$

where x_{-i} denoted the features except x_i , and the sum of all the calculated values of S_T is equal or greater than one.

If the values of S_T are substantially larger than the values of S_1 , then there are likely higher-order interactions occurring. Hence, it is worth calculating the second-order or higher-order sensitivity indices (e.g. S_2).

The second-order and higher-order sensitivity indices can similarly be expressed as

$$S_2 = \frac{\text{var}(x_i, x_j)}{\text{var}(\hat{y})}, \quad (20)$$

where $\text{var}(x_i, x_j)$ is the variance of features x_i and x_j . This calculates the amount of variance of \hat{y} explained by the interaction of features x_i and x_j .

3 Results and discussion

The results of the DL training, testing, and sensitivity analysis are discussed in the following sections.

3.1 Prediction of I_{ss} using deep learning

The DL process outlined in Fig. 3 was implemented to the randomly allocated dataset for training (118 and 81 data entries for Scenarios 1 and 2, respectively) and testing (51 and 35 data entries for Scenarios 1 and 2, respectively) listed in Table 1. The calculated loss values of the training and testing set using Eq. (7) against epochs are presented in Fig. 5. The loss values of both Scenarios 1 and 2 showed acceptable learning curves. This curve is a diagnostic tool for algorithms that learn from training datasets incrementally. The learning curves for both Scenarios 1 and 2 display a good fit that is negligibly experiencing underfitting or overfitting. This is characterised by training and testing loss values that decrease to the point of stability with a minimal gap between the two, as shown in Fig. 5. It is common to have a difference between the final loss values of the training and testing curves. Training curves having loss values less than testing curves are referred to as a “generalisation gap”. It can be observed that Scenario 1 (Fig. 5a) had a relatively wider gap between the training and testing loss values than Scenario 2 (Fig. 5b). This shows that Scenario 2, with features LL, PL, PI, LS, and fines, can give relatively better I_{ss} predictions with less overfitting due to generalisation.

The final performance training and testing of the model was assessed in terms of root mean squared error (RMSE) calculated as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}. \quad (21)$$

The RMSE indicates the average deviation of predictions from the measured values, with values closer to zero indicating better performance. The calculated RMSE for Scenario 1 was 1.26, whilst Scenario 2 had a lower RMSE of 0.90. This strengthens the authors’ inference that adding %fines as an input feature, even though this reduces the size of the dataset, can more reliably predict I_{ss} .

Further evaluation of the models of Scenarios 1 and 2 was carried out using an identity line or 1:1 line, as shown in Fig. 6. The 1:1 line has a slope of 1, forming a 45° angle. This line is used as a reference in a 2-dimensional scatter plot comparing two datasets that are expected to be alike under ideal conditions. When all the actual and predicted data points have equal values from the two datasets, the corresponding scatters fall along the 1:1 line [55]. Using the 1:1 line, there are two measurements we want to obtain

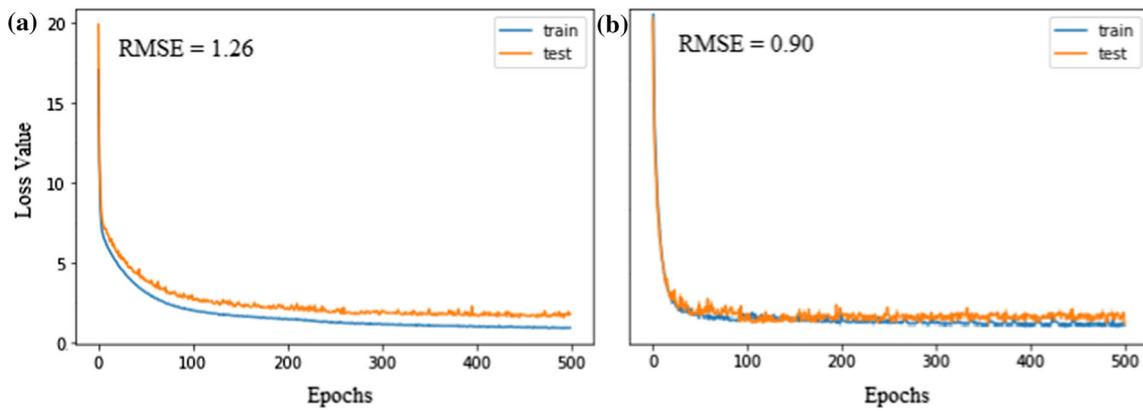


Fig. 5 Results and comparison of the calculated loss functions ($L(y, \hat{y})$) of training and testing sets for **a** Scenario 1 with features LL, PL, PI, and LS and **b** Scenario 2 with features LL, PL, PI, LS, and fines

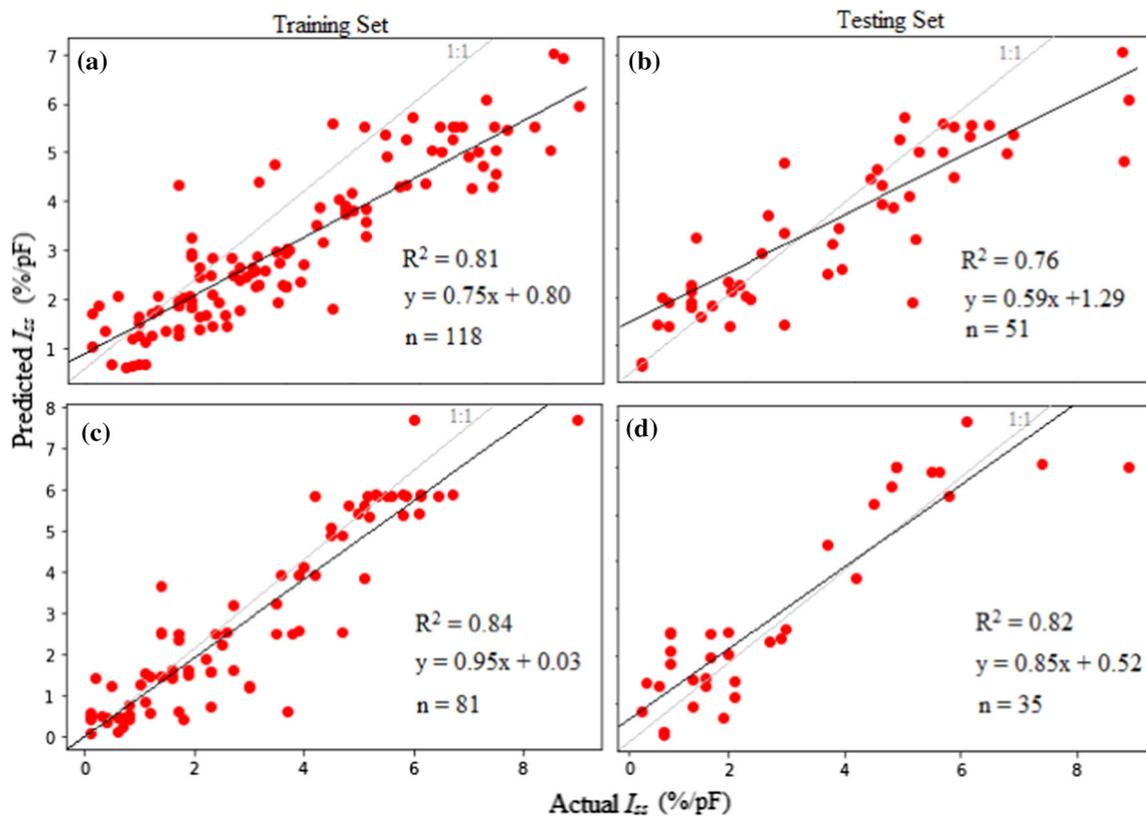


Fig. 6 Comparison between the predicted and actual values of I_{ss} of **a** the training set of Scenario 1, **b** the testing set of Scenario 1, **c** the training set of Scenario 2, and **d** the testing set of Scenario 2. The grey

line represents the 1:1 line and the black line represents the regression line (color figure online)

that reflect the reliability of the predictions of the models. The first measurement is the coefficient of determination, R^2 , calculated as

$$R^2 = 1 - \frac{RSS}{TSS} \tag{22}$$

where RSS is the sum of squares of residuals and TSS is the total sum of squares.

The second measure is the linear regression coefficient or slope that describes the relationship between the predicted and actual values. The values of R^2 and slope range between zero and one, with unity indicating a perfect fit.

The training set of Scenario 1 estimated R^2 to be 0.81 and the slope to be 0.75 (Fig. 6a), whilst the testing set of Scenario 1 resulted in a value of R^2 of 0.76 and a slope of 0.59 (Fig. 6b). These obtained correlations showed

improvements compared to most previous studies. For instance, Li et al. [9] reported that the correlation between I_{ss} and the Atterberg limits ranged from $R^2 = 0.20$ to $R^2 = 0.53$. Most of the performed studies up to date have concluded that I_{ss} and Atterberg limits are poorly correlated [11–14], except in the case of Jayasekera and Mohajerani [15] that found a relatively stronger correlation with R^2 values, which ranged from 0.85 to 0.91. However, the study of Jayasekera and Mohajerani [15] focused on a dataset with a low variance that limits the predictive ability of their model, with I_{ss} values that varied from 3.8 to 5.5.

The training set of Scenario 2 estimated R^2 to be 0.84 and the slope to be 0.95 (Fig. 6c), whilst the testing set of Scenario 2 resulted in a value of R^2 of 0.82 and a slope of 0.85 (Fig. 6d). The values of R^2 in the training and testing sets using the DL architecture and process implemented in

Scenario 2 were comparable with Earl [12], noting that Scenario 2 had a wider variance ($0.1 < I_{ss} < 9.0$). It can be observed from Fig. 6c that the slope is 0.95, which can be considered a strong correlation. However, due to the generalisation gap discussed earlier, the testing set commonly has lower accuracy than the training set. This holds in Fig. 7d, where the slope decreased to 0.85, which still shows a stronger correlation.

3.2 Sensitivity analysis

The bounds to generate the input features are specified in Table 1. The scheme by Saltelli [51] was implemented to generate the input features for predicting the values of I_{ss} . Sensitivity analysis was then performed (1) to assess the influence of the features on the targeted output, and (2) to

Fig. 7 Sobol indices for Scenario 1 showing **a** the total Sobol indices, **b** the first-order indices, **c** the second-order indices, and for Scenario 2 showing **d** the total Sobol indices, **e** the first-order indices, and **f** the second-order indices

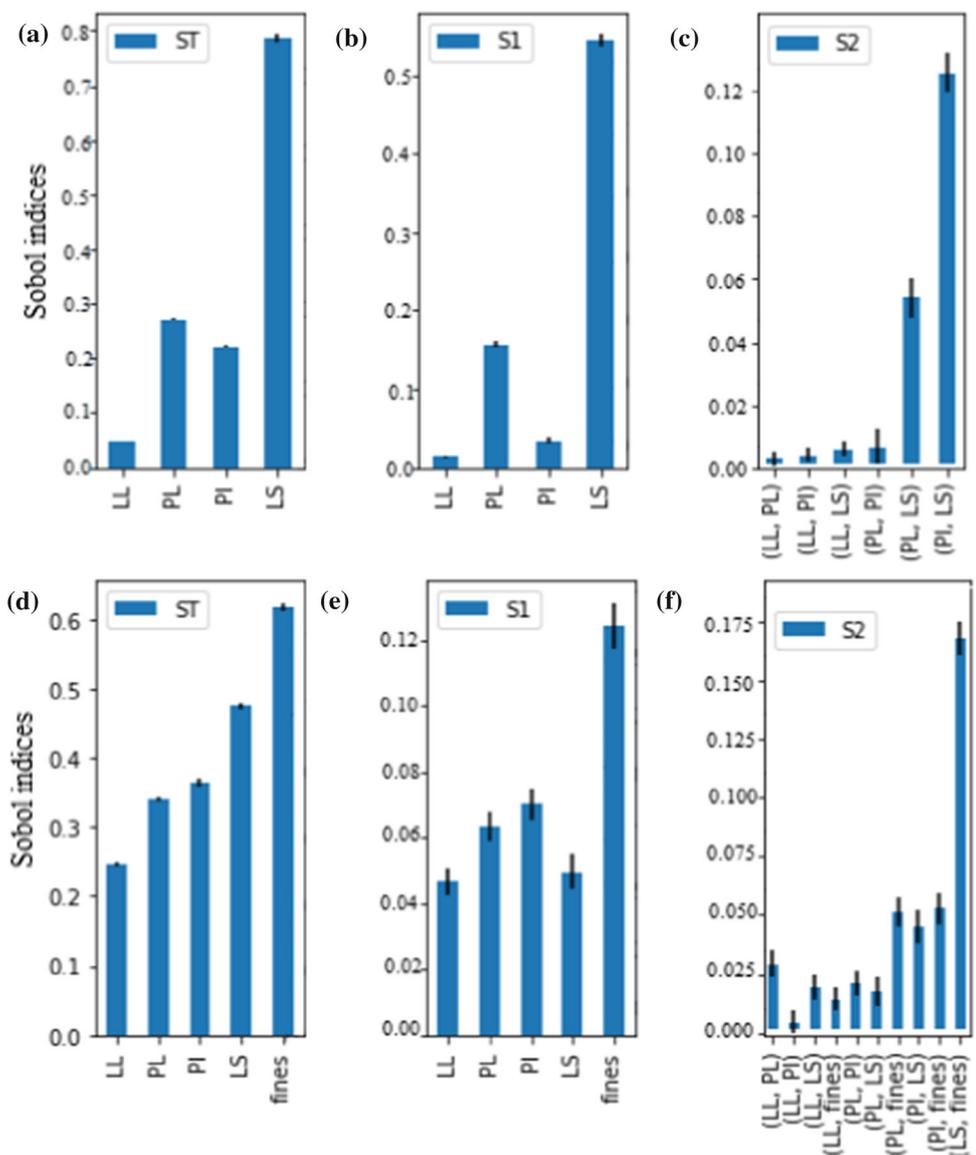


Table 3 Descriptive statistics of the generated input variables using the scheme by Saltelli [51] and the predicted results using DL for Scenario 1 (with features LL, PL, PI, and LS) and Scenario 2 (with features LL, PL, PI, LS, and fines)

	Scenario 1 ($n = 1,310,720$)					Scenario 2 ($n = 1,572,864$)					
	Generated input variables				Predicted results I_{ss} (%/pf)	Generated input variables				Predicted results	
	LL (%)	PL (%)	PI (%)	LS (%)		LL (%)	PL (%)	PI (%)	LS (%)	%fines (%)	I_{ss} (%/pf)
\bar{x}	72.5	32.5	50.0	20.0	5.2	72.5	32.5	50.0	20.0	50.5	2.5
s	33.2	15.9	28.9	11.5	3.3	33.2	15.9	28.9	11.5	28.6	1.4
Min	15.0	5.0	0.0	0.0	0.4	15.0	5.0	0.0	0.0	1.0	0.0
Q1 (25%)	43.7	18.7	25.0	10.0	2.6	43.7	18.7	25.0	10.0	25.7	1.6
Median (50%)	72.5	32.5	50.0	20.0	4.5	72.5	32.5	50.0	20.0	50.5	2.5
Q3 (75%)	101.2	46.2	75.0	30.0	7.0	101.2	46.2	75.0	30.0	75.2	3.3
Max	130.0	60.0	100.0	40.0	16.8	130.0	60.0	100.0	40.0	100.0	8.9

identify the relationship between input features and their influence on the target output. Descriptive statistics of the generated input variables using the scheme by Saltelli [51] and the predicted results using DL for Scenarios 1 and 2 are listed in Table 3. The generated input features of Scenario 1 were similar to the figures in Scenario 2. Interestingly, the predicted values of I_{ss} of Scenario 1 were observed to have higher values than Scenario 2, leading to higher calculated \bar{x} , s , minimum value, $Q1$, median, $Q3$, and maximum value. It is important to note that in Table 3, the maximum value of the predicted I_{ss} in Scenario 1 was 16.8, which is almost twice greater than the maximum value in the training set (9%/pF) presented in Fig. 3 and Table 1. On the other hand, the maximum value of the predicted I_{ss} in Scenario 2 was comparable to the training data (\approx 9%/pF). Thus, the range of predicted values of Scenario 2 is more practical and within the acceptable range.

The results of the global sensitivity analysis using Sobol [52, 53] for Scenario 1 and Scenario 2 are presented in Fig. 7. In Scenario 1, it can be observed in the first-order Sobol indices, S_1 , presented in Fig. 7b, that LS exhibited first-order sensitivities. This signifies that LS has the highest contribution of a single parameter to the output variance of I_{ss} . On the other hand, LL appears to have no first-order effects suggesting that it has a low contribution to the variation of the predicted values of I_{ss} . The values of the total-order Sobol index, S_T , were checked afterwards and are shown in Fig. 7a. If the values of S_T are markedly higher than those of S_1 , then there are likely higher-order interactions occurring. Higher-order interactions indicate that the fractional contribution of parameter interactions to the output variance exists. The values of S_T revealed (Fig. 7a) higher values than S_1 . Hence, the second-order

indices, S_2 , were calculated. PI and LS had the strongest feature interaction followed by PL and LS in Scenario 1, as shown in Fig. 7c. The remaining interactions can be considered insignificant.

In Scenario 2, it can be observed in S_1 presented in Fig. 7e that fines exhibited first-order sensitivities. This signifies that %fines has the highest contribution of a single parameter to the variance of I_{ss} . The other features, LL, PL, PI, and LS, appear to still have substantial first-order effects on the variation of the predicted values of I_{ss} . The values of S_1 in Scenario 2 (Fig. 7d) were more than four times lower than Scenario 1 (Fig. 7b). This reveals that the individual contribution of features in Scenario 2 is more distributed to other variables than Scenario 1. The values of S_T were then computed (Fig. 7d) and were found to be substantially larger than that of S_1 , hence, likely higher-order interactions are occurring. The interaction between LS and %fines had the largest values of S_2 , followed by PL-%fines, PI-LS, and PI-%fines, as shown in Fig. 7f.

4 Model comparison

The predictive accuracy of Scenario 1 was comparable to the models of Earl [12], Reynolds [13], and Li et al. [9]. The accuracy of the developed DL model was compared to the proposed models in previous studies. The comparison involved same seeded scenarios to predict the values of I_{ss} using the models listed in Table 4.

The developed DL neural network of Scenario 2 performed the best among the models, with the most desirable values of slope, R^2 , and RMSE. This indicates that Sce-

Table 4 Model comparison to predict values of I_{ss}

Model name	Model to predict I_{ss}	Slope, m	R^2	RMSE
Jayasekera and Mohajerani [15]—model 1	(PI + 40.3070)/17.5550	0.59	0.68	1.17
Jayasekera and Mohajerani [15]—model 2	(LL—8.5703)/15.7780	0.39	0.59	2.23
Earl [12]—model 1	0.0025LL ^{1.6817}	0.64	0.72	0.93
Earl [12]—model 2	0.0680PL ² —1.9830PL + 15.7190	3.42	0.34	11.00
Earl [12]—model 3	0.1048PI ^{0.7998}	0.28	0.58	1.25
Reynolds [13]—model 1	0.0745LL—1.6112	0.70	0.68	0.91
Reynolds [13]—model 2	0.0882PI—2.6370	0.60	0.59	2.25
Reynolds [13]—model 3	2.7910LS—1.4957	0.60	0.52	1.12
Li et al. [9]—model 1	0.0667LL—0.2860	0.62	0.68	1.48
Li et al. [9]—model 2	0.1269PL + 0.7508	0.32	0.29	2.05
Li et al. [9]—model 3	0.0793PI + 0.7954	0.54	0.59	1.51
Li et al. [9]—model 4	0.3115LS—1.4011	0.67	0.52	1.41
Scenario 1	DL using LL, PL, PI, and LS	0.59	0.76	1.26
Scenario 2 (proposed technique)	DL using LL, PL, PI, LS, and %fines	0.82	0.85	0.90

nario 2 predicted the most accurate values among the considered models in Table 4.

The models of Earl [12] and Reynolds [13] with LL as input had fairly acceptable values of slope, R^2 , and RMSE. The performance of the models and the values of R^2 conformed to the published work of Earl [12], Reynolds [13], and Li et al. [9] when applied to the compiled dataset used in this study (shown in Appendix). However, the models by Jayasekera and Mohajerani [15] underperformed when applied to the test dataset of this study. This may be due to the limited data and low variance of I_{ss} values used in their dataset and when applied to the compiled dataset in this study, the predictive range became inappropriate.

5 Conclusion

The soil parameter I_{ss} is widely used in Australia to determine the potential ground surface movement. However, determining I_{ss} takes a longer time to obtain results taking more than four days and involving hours of hands-on experimental work. This study developed an efficient method to estimate accurate values of I_{ss} through DL neural networks. This study proposed two scenarios; Scenario 1 involved the features LL, PL, PI, and LS whilst Scenario 2 added the input feature of %fines. The proposed

models were further investigated using sensitivity analysis to identify the relative influence of the input features on the targeted output, I_{ss} . This predictive DL model may significantly reduce the waiting period for the laboratory test results that are highly sensitive to instrumental conditions, skills and experience of the tester, and changing ambient conditions.

Results of implementing DL neural networks showed more reliable predictions in Scenario 2 (training: $R^2 = 0.84$ and slope = 0.95; testing $R^2 = 0.82$ and slope = 0.85) than Scenario 1 (training: $R^2 = 0.81$ and slope = 0.75; testing $R^2 = 0.76$ and slope = 0.59). These results suggested that adding a more relevant feature can be more beneficial than more data samples. Furthermore, the sensitivity analysis resulted in a more practical range of predicted values in Scenario 2, with %fines having the highest contribution to the variance of I_{ss} . The values of R^2 in the training and testing sets using the DL architecture and process implemented in Scenario 2 were considerably higher and had a wider variance than those of the previously conducted studies. This makes Scenario 2, the proposed model, around 10–65% more accurate than the models considered in this study for predicting I_{ss} . The developed DL neural network of Scenario 2 can then be used to estimate more accurate values of I_{ss} if an expedited result is required for design calculations.

Appendix 1

The dataset used for training and validation is presented below:

I_{ss}	LL	PL	PI	LS	%fines	Description	Resources
0.1	17	16	1	0.5	75	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.6	21	16	5	2.5	64	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.7	25	14	11	5	82	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.8	26	11	15	9	73	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.2	28	13	15	9	93	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.4	29	17	12	6	83	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.7	35	13	22	10.5	69	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.5	36	14	22	11.5	83	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.1	36	13	23	11	90	Pleistocene Quaternary Shepparton Formation	Earl [12]
2.3	36	13	23	11	77	Pleistocene Quaternary Shepparton Formation	Earl [12]
3	36	13	23	13	90	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.8	37	14	23	10.5	84	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.5	38	15	23	13	92	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.4	42	14	28	15	85	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.6	43	15	28	14.5	76	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.6	45	15	30	14	79	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.6	45	15	30	13.5	98	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.8	46	16	30	13	85	Pleistocene Quaternary Shepparton Formation	Earl [12]
2.1	46	15	31	14	86	Pleistocene Quaternary Shepparton Formation	Earl [12]
2.3	47	16	31	15.5	90	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.1	52	16	36	16	95	Pleistocene Quaternary Shepparton Formation	Earl [12]
3.5	53	17	36	13.5	96	Pleistocene Quaternary Shepparton Formation	Earl [12]
1	54	16	38	16.5	85	Pleistocene Quaternary Shepparton Formation	Earl [12]
2.3	54	16	38	17.5	93	Pleistocene Quaternary Shepparton Formation	Earl [12]
1.6	57	18	39	17	93	Pleistocene Quaternary Shepparton Formation	Earl [12]
3	59	16	43	17.5	93	Pleistocene Quaternary Shepparton Formation	Earl [12]
2.5	60	18	42	18	95	Pleistocene Quaternary Shepparton Formation	Earl [12]
4	62	20	42	15	96	Pleistocene Quaternary Shepparton Formation	Earl [12]
3.6	65	21	44	20	94	Pleistocene Quaternary Shepparton Formation	Earl [12]
0.4	28.2	14	14.2	8.6	77	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.9	31.8	18	13.8	8.2	60	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.1	32.4	5.6	26.8	8.6	66	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.6	32.4	17	15.4	9	69	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.7	32.4	16.8	15.6	8.6	74	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.3	34.4	13.2	21.2	9.8	62	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
3.7	35.8	16.6	19.2	10.6	54	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.1	36.6	18.2	18.4	10.4	61	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.6	42	19	23	13.2	63	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.2	43	13.6	29.4	14.2	67	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1	43.2	18	25.2	12.2	43	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.1	44	20	24	11.5	82	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
3	44	12.9	31.1	12.6	71	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
0.8	45	27	18	10	16	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.7	45.6	19	26.6	15.8	62	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.9	46.4	19.8	26.6	15	56	Queensland (Southeast, Central, and Northwest)	Reynolds [13]

(continued)

I_{ss}	LL	PL	PI	LS	%fines	Description	Resources
1.4	48	17	31	13.2	41	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
3.9	48	19.2	28.8	16.8	63	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.9	48.6	20	28.6	15.4	77	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.9	49	29.8	19.2	9.6	93	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.7	49.6	26	23.6	6.6	58	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.4	50.8	27	23.8	14.2	90	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.4	50.8	27	23.8	14.2	90	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.7	52	19.2	32.8	14.8	62	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1	53.8	29.8	24	13.2	73	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1	58.8	30.8	28	13.4	89	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.7	59	28.4	30.6	15	86	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.7	59	28.4	30.6	15	90	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1	60.8	28.4	32.4	16.4	82	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.4	60.8	28.4	32.4	16.4	82	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
3.8	61.8	28.8	33	13.2	89	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.6	62	24.4	37.6	19.4	52	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.4	63.4	32	31.4	16.2	59	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2	64.6	30	34.6	16.4	51	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
3.5	67	28	39	19	78	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.2	70.2	23.8	46.4	15.8	44	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.5	71.4	22.2	49.2	20.4	84	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.7	71.4	35.4	36	16.2	51	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.7	72.8	30	42.8	17.2	80	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.4	72.2	29.4	42.8	17.2	80	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.7	74.6	33.4	41.2	19	86	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.2	74.8	29	45.8	18	82	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
5.1	76	30	46	16	94	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
5.5	77	24	53	18.5	88	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.5	77.4	29.4	48	19	86	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
5	78.8	35	43.8	19.2	92	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
6.1	79.6	34.8	44.8	20	90	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.8	80	24	56	22	83	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.8	80	24	56	22	83	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
5.8	80.2	35.2	45	20.2	85	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
5.8	81	31.6	49.4	18.8	88	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
3.7	81.6	38.6	43	16.2	87	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.5	83	25	58	23.5	73	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
5.1	88	36	52	21.5	91	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
2.7	90	33	57	23	99	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.9	95	25	70	23	98	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
4.9	95	25	70	23	98	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
7.4	100	30	70	27	91	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
6	105	36	69	20	100	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
9	105	36	69	20	100	Queensland (Southeast, Central, and Northwest)	Reynolds [13]
1.47	39.94	19.43	20.51	11.2		Quaternary basalt	Zou [14]
6.16	65.26	28.51	36.75	20.2		Quaternary basalt	Zou [14]
2.7	48.2	24	24.2	14		Quaternary basalt	Zou [14]
3.55	58.61	22.18	36.43	17.1		Quaternary basalt	Zou [14]

(continued)

I_{ss}	LL	PL	PI	LS	%fines	Description	Resources
2.83	72.7	25.1	47.6	19.9		Quaternary basalt	Zou [14]
5.73	100	37.02	62.98	25.33		Quaternary basalt	Zou [14]
7.15	122	42	80	19.6		Quaternary basalt	Zou [14]
3.46	38.92	19.27	19.65	12		Quaternary basalt	Zou [14]
2.87	40.65	20.7	19.92	18		Quaternary basalt	Zou [14]
2.22	37.73	17.24	20.49	13.4		Quaternary basalt	Zou [14]
2.88	36.09	15.46	20.63	11.2		Quaternary basalt	Zou [14]
2.56	51.52	30.18	21.34	13.2		Quaternary basalt	Zou [14]
2.57	42.43	20.02	22.41	15.2		Quaternary basalt	Zou [14]
4.4	49	26.18	22.82	20		Quaternary basalt	Zou [14]
1.69	41.75	17.21	24.54	15		Quaternary basalt	Zou [14]
2.32	46.68	19.48	27.2	15.4		Quaternary basalt	Zou [14]
3.04	50.77	22.6	28.17	20.4		Quaternary basalt	Zou [14]
2.95	47.71	15.01	32.7	12.8		Quaternary basalt	Zou [14]
7.02	74.69	34.87	39.82	24		Quaternary basalt	Zou [14]
4.3	73.3	25.8	47.5	18.8		Quaternary basalt	Zou [14]
4.53	80.75	18.02	62.73	22.1		Quaternary basalt	Zou [14]
5.96	68.87	31.95	36.92	19.2		Quaternary basalt	Zou [14]
6.32	84.39	26.62	57.77	24.2		Quaternary basalt	Zou [14]
5.05	47.08	15.81	31.27	13.6		Quaternary basaltic residual	Zou [14]
1.59	66.29	18.66	47.63	18.3		Quaternary basaltic residual	Zou [14]
6.96	75.79	24.51	51.28	20.9		Quaternary basaltic residual	Zou [14]
6.88	70.28	21.74	48.54	18.4		Quaternary basaltic residual	Zou [14]
6.16	69.74	19.72	50.02	19.5		Quaternary basaltic residual	Zou [14]
6.77	80.52	22.37	58.15	23.2		Quaternary basaltic residual	Zou [14]
1.08	56.34	31.91	24.43	9.8		Tertiary basaltic	Zou [14]
3.28	49.99	29.74	20.25	12.5		Tertiary basaltic	Zou [14]
1.87	63.91	21.28	42.63	11.1		Tertiary basaltic residual	Zou [14]
4.21	56.92	17.42	39.5	16		Quaternary swamp and lagoon deposits	Zou [14]
2.42	64.17	20.57	43.6	16.1		Quaternary swamp and lagoon deposits	Zou [14]
1.91	36.01	12.82	23.19	17.9		Quaternary aged alluvium deposits	Zou [14]
1.57	37.6	24.91	32.69	10.4		Silurian sedimentary	Zou [14]
4.02	61.1	25.7	35.4	15.2		Silurian sedimentary	Zou [14]
1.75	54.79	19.27	35.52	12		Silurian sedimentary	Zou [14]
2.51	65.93	25.74	40.19	14.59		Silurian sedimentary clays	Zou [14]
2.55	54.96	27.16	27.8	12.7		Tertiary aged sedimentary deposits	Zou [14]
2.11	43.34	14.89	28.45	8.3		Tertiary aged sedimentary deposits	Zou [14]
3.06	69.8	23.5	46.3	15.5		Torquay group deposit	Zou [14]
3.23	41.53	16.86	24.67	13.2		Quaternary alluvium	Zou [14]
3.76	48.88	18.65	30.23	14.6		Quaternary alluvium	Zou [14]
3.9	66	25	41	18.5		Quaternary alluvium (Torquay)	Smith [44]
4.2	66	25	41	18.5		Quaternary alluvium (Torquay)	Smith [44]
1.4	88	38	50	21.5		Mornington volcanics	Smith [44]
1.6	53	22	31	12		Deutgam silt	Smith [44]
1.2	50	22	28	11		Deutgam silt	Smith [44]
1.9	53	23	30	13		Deutgam silt	Smith [44]
5.2	107	17	90	14.5		Newer volcanics (Footscray)	Smith [44]
6.1	107	32	75	22		Quaternary alluvium (Wyndham Vale)	Smith [44]

(continued)

I_{ss}	LL	PL	PI	LS	%fines	Description	Resources
0.9	31	14	17	7		Deutgam silt	Smith [44]
2	53	20	33	12		New volcanics (Wyndham Vale)	Smith [44]
0.8	33	15	18	7		Deutgam silt	Smith [44]
0.1	35	16	19	6.5		Deutgam silt	Smith [44]
8.9	73	25	48	18		Coode island silt	Smith [44]
5.51	76.4	22.6	53.8	18.9		Quaternary basalts	Karunaratne [43]
4.19	76.4	22.6	53.8	18.9		Quaternary basalts	Karunaratne [43]
5.56	76.4	22.6	53.8	18.9		Quaternary basalts	Karunaratne [43]
6.46	76.4	22.6	53.8	18.9		Quaternary basalts	Karunaratne [43]
6.12	76.4	22.6	53.8	18.9		Quaternary basalts	Karunaratne [43]
5.65	71.1	21.9	49.2	18.2		Quaternary basalts	Karunaratne [43]
5.3	71.1	21.9	49.2	18.2		Quaternary basalts	Karunaratne [43]
5.82	71.1	21.9	49.2	18.2		Quaternary basalts	Karunaratne [43]
6.72	71.1	21.9	49.2	18.2		Quaternary basalts	Karunaratne [43]
6.12	71.1	21.9	49.2	18.2		Quaternary basalts	Karunaratne [43]
5.33	76.6	23	53.6	17.1		Quaternary basalts	Karunaratne [43]
5.6	76.6	23	53.6	17.1		Quaternary basalts	Karunaratne [43]
5.88	76.6	23	53.6	17.1		Quaternary basalts	Karunaratne [43]
5.17	76.6	23	53.6	17.1		Quaternary basalts	Karunaratne [43]

Acknowledgements This work was funded in partnership with the Victorian State Government that the authors would like to acknowledge and thank.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. This work was funded in partnership with the Victorian State Government.

Availability of data and materials Data are available in Appendix 1.

Declarations

Conflict of interest The authors declare that there is no conflict of interest regarding the publication of this paper.

Code availability Python scripts are available at <https://github.com/bteodosio/soilshrinkindex>.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Fityus S, Cameron D, Walsh P (2005) The shrink swell test. *Geotech Test J GEOTECH Test J*. <https://doi.org/10.1520/GTJ12327>
- Shams MA, Shahin MA, Ismail MA (2018) Simulating the behaviour of reactive soils and slab foundations using hydro-mechanical finite element modelling incorporating soil suction and moisture changes. *Comput Geotech* 98:17–34
- Tran KM, Bui HH, Sánchez M, Kodikara J (2020) A DEM approach to study desiccation processes in slurry soils. *Comput Geotech* 120:103448
- Teodosio B, Baduge KSK, Mendis P (2020) Relationship between reactive soil movement and footing deflection: a coupled hydro-mechanical finite element modelling perspective. *Comput Geotech* 126:103720
- Teodosio B, Baduge KSK, Mendis P (2020) Simulating reactive soil and substructure interaction using a simplified hydro-mechanical finite element model dependent on soil saturation, suction and moisture-swelling relationship. *Comput Geotech* 119:103359
- Li J, Cameron DA, Ren G (2014) Case study and back analysis of a residential building damaged by expansive soils. *Comput Geotech* 56:89–99
- Miao L, Wang F, Cui Y, Shi S (2012) Hydraulic characteristics, strength of cyclic wetting-drying and constitutive model of expansive soils. In: 4th international conference actual problems. *Soils*, pp 303–322
- Jones LD, Jefferson I (2012) *Expansive Soils*. Institution of Civil Engineers
- Li J, Zou J, Bayetto P, Barker N (2016) Shrink-swell index database for Melbourne. *Aust Geomech J* 51:61–76
- Teodosio B, Kristombu Baduge KS, Mendis P (2021) A review and comparison of design methods for raft substructures on

- expansive soils. *J Build Eng* 41:102737. <https://doi.org/10.1016/j.jobe.2021.102737>
11. Young GS, Parmar M (1999) Shrink-swell testing in the Sydney region. In: Proceedings 8th ANZ conference on geomechanics. Hobart AGS, pp 221–5
 12. Earl D (2005) To determine if there is a correlation between the shrink-swell index and Atterberg limits for soils within the Shepparton Formation. University of Southern Queensland
 13. Reynolds PW (2013) Engineering correlations for the characterisation of reactive soil behaviour for use in road design. University of Southern Queensland
 14. Zou J (2015) Assessment of the reactivity of expansive soil in Melbourne metropolitan area. RMIT University
 15. Jayasekera S, Mohajerani A (2003) Some relationships between shrink-swell index, liquid limit, plasticity index, activity and free swell index. *Australian Geomechanics* 38:7
 16. Theodoridis S (2020) *Machine Learning—2nd Edition*. <https://www.elsevier.com/books/machine-learning/theodoridis/978-0-12-818803-3>. Accessed 12 June 2021
 17. Salehi H, Burgueño R (2018) Emerging artificial intelligence methods in structural engineering. *Eng Struct* 171:170–189. <https://doi.org/10.1016/j.engstruct.2018.05.084>
 18. Carbonell JG, Michalski RS, Mitchell TM (1983) An overview of machine learning. *Mach Learn* 3–23
 19. Abualigah L, Elaziz MA, Khasawneh AM, Alshinwan M, Ibrahim RA, Al-qaness MAA et al (2022) Meta-heuristic optimization algorithms for solving real-world mechanical engineering design problems: a comprehensive survey, applications, comparative analysis, and results. *Neural Comput Appl* 34:4081–4110. <https://doi.org/10.1007/s00521-021-06747-4>
 20. Dai Y, Khandelwal M, Qiu Y, Zhou J, Monjezi M, Yang P (2022) A hybrid metaheuristic approach using random forest and particle swarm optimization to study and evaluate backbreak in open-pit blasting. *Neural Comput Appl* 34:6273–6288. <https://doi.org/10.1007/s00521-021-06776-z>
 21. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323:533–536
 22. Xu Y, Zhou Y, Sekula P, Ding L (2021) Machine learning in construction: from shallow to deep learning. *Dev Built Environ* 6:100045. <https://doi.org/10.1016/j.dibe.2021.100045>
 23. Ghani S, Kumari S, Choudhary AK, Jha JN (2021) Experimental and computational response of strip footing resting on prestressed geotextile-reinforced industrial waste. *Innov Infrastruct Solut* 6:98. <https://doi.org/10.1007/s41062-021-00468-2>
 24. Akis E, Guven G, Lotfisdigh B (2022) Predictive models for mechanical properties of expanded polystyrene (EPS) geofoam using regression analysis and artificial neural networks. *Neural Comput Appl* 34:10845–10884. <https://doi.org/10.1007/s00521-022-07014-w>
 25. Soranzo E, Guardiani C, Wu W (2021) A soft computing approach to tunnel face stability in a probabilistic framework. *Acta Geotech*. <https://doi.org/10.1007/s11440-021-01240-7>
 26. Zhao S, Shadab Far M, Zhang D, Chen J, Huang H (2021) Deep learning-based classification and instance segmentation of leakage-area and scaling images of shield tunnel linings. *Struct Control Health Monit*. <https://doi.org/10.1002/stc.2732>
 27. Yang H-Q, Zhang L, Pan Q, Phoon K-K, Shen Z (2021) Bayesian estimation of spatially varying soil parameters with spatiotemporal monitoring data. *Acta Geotech* 16:263–278. <https://doi.org/10.1007/s11440-020-00991-z>
 28. Zhao H, Liu W, Shi P, Du J, Chen X (2021) Spatiotemporal deep learning approach on estimation of diaphragm wall deformation induced by excavation. *Acta Geotech*. <https://doi.org/10.1007/s11440-021-01264-z>
 29. Nayek PS, Gade M (2022) Artificial neural network-based fully data-driven models for prediction of newmark sliding displacement of slopes. *Neural Comput Appl* 34:9191–9203. <https://doi.org/10.1007/s00521-022-06945-8>
 30. Abbaszadeh SA (2016) Assessment and prediction of liquefaction potential using different artificial neural network models: a case study. *Geotech Geol Eng* 34:807–815. <https://doi.org/10.1007/s10706-016-0004-z>
 31. Lai Z, Chen Q (2019) Reconstructing granular particles from X-ray computed tomography using the TWS machine learning tool and the level set method. *Acta Geotech* 14:1–18. <https://doi.org/10.1007/s11440-018-0759-x>
 32. Cahyadi TA, Widodo LE, Syihab Z, Notosiswoyo S, Widijanto E (2017) Hydraulic conductivity modeling of fractured rock at Grasberg surface mine, Papua-Indonesia. *J Eng Technol Sci* 49
 33. Zhang P, Yin Z-Y, Jin Y-F, Liu X-F (2021) Modelling the mechanical behaviour of soils using machine learning algorithms with explicit formulations. *Acta Geotech*. <https://doi.org/10.1007/s11440-021-01170-4>
 34. Tran QA, Ho LS, Le HV, Prakash I, Pham BT (2022) Estimation of the undrained shear strength of sensitive clays using optimized inference intelligence system. *Neural Comput Appl* 34:7835–7849. <https://doi.org/10.1007/s00521-022-06891-5>
 35. Tran VQ, Dang VQ, Do HQ, Ho LS (2022) Investigation of ANN architecture for predicting residual strength of clay soil. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-022-07547-0>
 36. Cha Y-J, Choi W, Büyüköztürk O (2017) Deep learning-based crack damage detection using convolutional neural networks. *Comput-Aid Civ Infrastruct Eng* 32:361–378. <https://doi.org/10.1111/micc.12263>
 37. Rezazadeh Eidgahee D, Jahangir H, Solatifar N, Fakharian P, Rezaeemaneh M (2022) Data-driven estimation models of asphalt mixtures dynamic modulus using ANN, GP and combinatorial GMDH approaches. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-022-07382-3>
 38. Dimiduk DM, Holm EA, Niezgoda SR (2018) Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. *Integr Mater Manuf Innov* 7:157–172. <https://doi.org/10.1007/s40192-018-0117-8>
 39. Ebid AM (2021) 35 years of (AI) in geotechnical engineering: state of the art. *Geotech Geol Eng* 39:637–690. <https://doi.org/10.1007/s10706-020-01536-7>
 40. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444. <https://doi.org/10.1038/nature14539>
 41. Cui K, Jing X (2019) Research on prediction model of geotechnical parameters based on BP neural network. *Neural Comput Appl* 31:8205–8215. <https://doi.org/10.1007/s00521-018-3902-6>
 42. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
 43. Karunarathne A (2016) Investigation of expansive soil for design of light residential footings in Melbourne. Swinburne University of Technology
 44. Smith TW (2017) A correlation for the shrink swell index. AGS Vic. Chapter 2017 Symp
 45. Nguyen QH, Ly H-B, Ho LS, Al-Ansari N, Le HV, Tran VQ et al (2021) Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Math Probl Eng* 2021:e4832864. <https://doi.org/10.1155/2021/4832864>
 46. He K, Zhang X, Ren S, Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of IEEE international conference of computer vision, pp 1026–34
 47. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. *Int Conf for Machine Learning* 2010

48. Kingma DP, Ba J (2017) Adam: a method for stochastic optimization. ArXiv14126980 Cs
49. Zhou X, Lin H (2008) Local sensitivity analysis. In: Shekhar S, Xiong H (eds) *Encycl. GIS*, Boston, MA: Springer US, pp 616–616. https://doi.org/10.1007/978-0-387-35973-1_703
50. Saltelli A, Aleksankina K, Becker W, Fennell P, Ferretti F, Holst N et al (2019) Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. *Environ Model Softw* 114:29–39. <https://doi.org/10.1016/j.envsoft.2019.01.012>
51. Saltelli A (2002) Making best use of model evaluations to compute sensitivity indices. *Comput Phys Commun* 145:280–297
52. Sobol IM (1990) On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie* 2:112–118
53. Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comput Simul* 55:271–280. [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6)
54. Homma T, Saltelli A (1996) Importance measures in global sensitivity analysis of nonlinear models. *Reliab Eng Syst Saf* 52:1–17
55. Teodosio B, Pauwels VR, Loheide SP, Daly E (2017) Relationship between root water uptake and soil respiration: a modeling perspective. *J Geophys Res Biogeosci* 122:1954–1968

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.