Load-Settlement Response of A Footing Over Buried Conduit in A Sloping Terrain: A Numerical Experiment-Based Artificial Intelligent Approach

Muhammad Umer Arif Khan (umer.arif1@gmail.com)
Edith Cowan University - Joondalup Campus: Edith Cowan University https://orcid.org/0000-0002-3077-2875
Sanjay Kumar Shukla
Edith Cowan University - Joondalup Campus: Edith Cowan University
Muhammad Nouman Amjad Raja
Edith Cowan University - Joondalup Campus: Edith Cowan University

Research Article

Keywords: Buried-conduit, Slope, Artifical intelligence, Finite element modelling, Load-settlement behavior

Posted Date: November 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-172710/v1

Version of Record: A version of this preprint was published at Soft Computing on January 28th, 2022. See the published version at https://doi.org/10.1007/s00500-021-06628-x.

Load-settlement response of a footing over buried conduit in a sloping terrain: a numerical experiment-based artificial intelligent approach

Muhammad Umer Arif Khan[1], Sanjay Kumar Shukla[2], Muhammad Nouman Amjad Raja[3]

ABSTRACT

Settlement estimation of a footing located over a buried conduit in a sloping terrain is a challenging task for practicing civil/geotechnical engineers. In the recent past, the advent of machine learning technology has made many traditional approaches antiquated. This paper investigates the viability, development, implementation, and comprehensive comparison of five artificial intelligence-based machine learning models, namely, multi-layer perceptron (MLP), Gaussian processes regression (GPR), lazy *K*-Star (LKS), decision table (DT), and random forest (RF) to estimate the settlement of footing located over a buried conduit within a soil slope. The pertaining dataset of 3600 observations was obtained by conducting large-scale numerical simulations via the finite element modelling framework. After executing the feature selection technique that is correlation-based subset selection, the applied load, total unit weight of soil, constrained modulus of soil, slope angle ratio, hoop stiffness of conduit, bending stiffness of conduit, burial depth of conduit, and crest distance of footing were utilized as the influence parameters for estimating and forecasting the settlement. The predictive strength and accuracy of all models mentioned supra were evaluated using several well-established statistical indices such as Pearson's correlation coefficient (r), root mean square error (RMSE), Nash-Sutcliffe efficiency (NSE), scatter index (SI), and relative percentage difference (RPD). The results showed that among all the models employed in this study, the multi-layer perceptron model has shown better results with r, RMSE, NSE, SI, and RPD values of (0.977, 0.298, 0.937, 0.31, and 4.31) and (0.974, 0.323, 0.928, 0.44 and 3.75) for training and testing dataset, respectively. The sensitivity analysis revealed that all the selected parameters play an important role in determining the output value. However, the applied load, constrained modulus, unit weight, slope angle ratio, hoop stiffness have the highest strength with the relative importance of 18.4%, 16.3% and 15.3%, 13.8%, 11.4%, respectively. Finally, the model was translated into a functional relationship for easy implementation and can prove useful for practitioners and researchers in predicting the settlement of a footing located over a buried conduit in a sloping terrain.

Keywords: Buried-conduit; Slope; Artifical intelligence; Finite element modelling; Load-settlement behavior

Affiliations:

[1]Research Candidate, Discipline of Civil and Environmental Engineering, School of Engineering, Edith Cowan University, Perth, WA 6027, Australia; Lecturer, Mirpur University of Science and Technology, Mirpur, Azad Kashmir, Pakistan, E-mail address: mukhan3@our.ecu.edu.au; umer.arif1@gmail.com, ORCID number: 0000-0002-3077-2875

[2]Founding Research Group Leader, Geotechnical and Geoenvironmental Engineering Research Group, School of Engineering, Edith Cowan University, Perth, Australia; Adjunct Professor, Department of Civil Engineering, Delhi Technological University, Delhi, India, Email: s.shukla@ecu.edu.au; sanjaykshukla1@gmail.com; ORCID: 0000-0002-4685-5560

[3]Geotechnical and Geoenvironmental Engineering Research Group, School of Engineering, Edith Cowan University, Perth, Australia; Email: m.raja@ecu.edu.au; noumanamjad@live.com; ORCID: 0000-0001-7463-0601 (Corresponding Author).

1. Introduction

The tunneling and underground infrastructure is a salient feature of modern urbanization. The economic and safety benefits of the buried conduits have made them the most frequently used mode of utility conveyance. The scarcity of land to ever-increasing population growth has resulted in the construction activity over the buried infrastructure. The influence of the imposed loading on a buried conduit is always incorporated in its design and installation (Moser and Folkman 2001). The studies on the effect of the applied surface pressure on the soil-conduit interaction and the resulting stress distribution and structural response of the conduit can be found in the current literature (Dhar et al. 2004; Talesnick et al. 2012; Bryden et al. 2015; Robert et al. 2016; Wang et al. 2017; Al-Naddaf et al. 2019; Khan and Shukla 2021a). However, the research on the presence of the buried conduit on the settlement and bearing capacity of a surface footing is very limited. Srivastava et al. (2013) investigated the load-settlement response of a circular footing placed over a PVC conduit buried under the level ground. Using laboratory model tests, the load-settlement behavior and bearing capacity of the footing was analyzed in loose-medium (relative density = 50%) and very dense sands (relative density = 88%). The experimental results were also compared with the results obtained from finite element analysis of the same model. The results showed that in the case of the loose-medium dense sand, the induction of stiffer conduit material improved the load-settlement response of the footing. As a result, its bearing capacity increased by about 25%. Whereas, for the very dense sand, the presence of the flexible conduit reduced the bearing capacity of the overlying footing by approximately 8%. Therefore, it can be concluded that under the static load conditions, the relative density of the sand surrounding the buried conduit and the resulting relative stiffness with the conduit material governs the settlement and bearing capacity of the surface footing. Similarly, Bildik and Laman (2015, 2019) conducted laboratory model tests to analyze the effect of a buried PVC conduit on the load-settlement response and bearing capacity of an overlying strip footing. The study was conducted by varying the burial depth and the horizontal distance of the conduit from the footing. The settlement of the surface footing was measured by employing two deflection transduces instrumented on both sides of the surface footing. The results showed that the load-settlement behavior and bearing capacity of the footing improved significantly as the horizontal distance between the footing and the buried conduit was increased. Also, it was noted that as the buried conduit was moved away from the stress zone under the footing, the bearing capacity of the footing increased. At a burial depth of more than 4 times the conduit diameter, the buried conduit seized to impact the load-settlement behavior and bearing capacity of the surface footing. While the aforementioned studies investigated the effect of buried conduits on the load-settlement response and bearing capacity of footings located over the horizontal ground, only one study can be found in the literature that has analyzed the footing settlement in a sloping terrain. Khan and Shukla (2020) conducted laboratory model tests to investigate the settlement and bearing capacity of a strip surface footing located over a conduit buried within the soil slope. Using two linear variable displacement transducers (LVDTs) installed on both sides of the footing, the effects of un-plasticized polyvinyl chloride (PVC-U) conduits of diameters 80mm and 160mm were studied in detail. The shear failure mechanisms of the footing were analytically computed and illustrated to understand the resulting soil-conduit interaction. The study concluded that when the shear failure planes of the footing intersected with the buried conduit, its bearing capacity was reduced by about 40%. However, an increase in the burial depth of the conduit and the crest distance of the footing enhanced the distance between the buried conduit and failure planes of the footing, resulting in a decrease in the effect of buried conduit on the settlement and bearing capacity of the surface footing. Further, the sensitivity analysis categorized the burial depth of the conduit and the crest distance of the surface footing from the edge of the soil slope as the most influential parameters affecting the load-carrying behavior of the surface footing located over a conduit buried within a soil slope.

Summarizing, the limited number of related studies, as discussed above, have concluded that the load-settlement response and bearing capacity of footings is affected by the relative stiffness with the conduit material and the surrounding soil, burial depth of the conduit, and the crest distance of the conduit from the slope surface. However, the studies have only analyzed limited values of these influential parameters due to the experimental restraints. Furthermore, these experimental studies have been conducted on small-scale 1g laboratory models, which hinders the veracity of such studies due to the scale effect. While the small-scale model tests may explain the relevant mechanisms/ trends, the observed measurements may not reflect the actual field values (Sedran et al. 2001; Cerato and Lutenegger 2003). Additionally, the use of only one type of conduit material significantly limits the generalized use of related studies. To the authors' best knowledge, no study exists in the current literature that can be used for the direct estimation of the settlement of a strip footing located over a buried conduit with a soil slope.

Finite element modelling (FEM) can be used to solve complex geotechnical problems and achieve more accurate results (Khan and Shukla 2021b). However, the use of expensive software for FEM analysis significantly limits their application (Kim et al. 2012). In recent times, the use of machine learning techniques has been widely used in mapping the non-linear relationships between the input and output variables (e.g., Ahmadi et al. 2019; Yekani Motlagh et al. 2019; Aamir et al. 2020; Dorosti et al. 2020; Ghorbani et al. 2021; Kaloop et al. 2021). The novel metaheustarics algorithms are also developed for optimisation purposes in big-data analysis (Abualigah and Alkhrabsheh 2021; Abualigah et al. 2021a). Similarly, for geotechnical problems, the soft computing approaches are now commonly used for prediction purposes (e.g., Nguyen et al. 2019; Xiao and Zhao 2019; Bardhan et al. 2021; Kardani et al.

2021a, b; Khan et al. 2021; Raja and Shukla 2021a, b; Raja et al. 2021). The machine learning (ML) models that are based on large quantities of FEM data have also been developed to solve complex problems like soil-conduit interaction and settlement of foundations. Kim et al. (2012) employed FEM based artificial neural network (ANN) to predict deflections of buried corrugated conduits. The data collected from three-dimensional finite element modelling were used to develop a backpropagation (BP) neural network that examined the factors affecting the structural response of different corrugated conduits buried at various depths under the level ground. Shokouhi et al. (2013) used a FEM-ANN approach to develop an ANN model that could be used to predict the bending strains developed in conduits buried within a fault zone. Kardani et al. (2020) used the FEM-based data to successfully predict the factor of safety of a soil slope using the hybrid stacking ensemble machine learning modelling technique. The aspect of footing settlement has also been studied by using the FEM-ML approach. Moayedi and Hayati (2018) used large FEM data to develop a number of soft-computing models in order to predict the settlement of a strip surface footing located near a sandy soil slope. Similarly, Moayedi et al. (2020a) developed optimized neural networks such as differential evolution algorithm (DEA), adaptive neuro-fuzzy inference system (ANFIS) to predict the ultimate bearing capacity of a shallow footing on two-layered soil condition, utilizing FEM data. In this paper, an attempt has been made to predict the settlement of a surface footing located over a conduit buried within a soil slope, using various machine learning/intelligent modelling techniques, namely multi-layer perceptron (MLP), Gaussian processes regression (GPR), Lazy *K*-Star (LKS), decision table (DT), and random forest (RF). Using the finite element modelling, large-scale data were generated for the settlement of the footing located over the conduits of varying stiffness. The main objectives of this paper are as follows:

1) Development and assessment of five machine learning models such as to MLP, GPR, LKS, DT and RF for direct estimation of the settlement of a strip footing, located over a buried conduit with a soil slope.
2) Comprehensive analysis and comparison of these models for the same problem.
3) Feature selection to choose the most important parameters affecting the footing settlement.
4) Assess the robustness of the developed models and conduct the sensitivity analysis.
5) Develop and present a trackable functional ANN-based formula for direct estimation of the footing settlement locate over a buried conduit.

1.1. Significance of the Research

This study is useful in ensuring the stability of surface footings that are frequently located over tunnels and underground infrastructure in the current urban environment. Using extensive finite element modelling, it incorporates the effect of a large number of input parameters on the load-settlement response of a large-scale surface footing located over different types of buried conduits. The inclusion of numerous input parameters employed to define soil and the buried conduits and their complicated relationships with the output parameter results in highly complex geotechnical models. Further enhancing this complexity are the intricate correlations between the input parameters, where a change in one input parameter causes a change in another or more than one input parameters. Considering these convoluted relationships and the resulting rigorous finite element calculations, this study utilizes advanced machine learning/ intelligent modelling techniques to provide accurate and straightforward solutions to the complex soil-conduit interactions. The developed MLP-based formula can be used by the practicing engineers to directly estimate footing settlement when loaded over a conduit buried within a sloping terrain. The steps involved in the model developments can be summarized as follows:

(1) Data generation through Finite element modelling (FEM)

(2) Data pre-processing, feature validity, and data division into training and testing data

(3) Development and implementation of AI-based models

(4) Statistical analysis of the results and selection of best model

(5) Model robustness and sensitivity analysis

2. Material and methods

2.1. Finite element modelling

The commercial PLAXIS 2D software was used for data collection by conducting the finite element analysis of a large-scale model, as presented in Figure. 1. Using the Mohr-Coulomb model, the soil was modelled as per the field properties of the most common sandy soils, presented in Table 1 (Ghazavi and Eghbali 2008; Moayedi and Hayati 2018). The soils, numbered as one to five were differentiated in terms of their strength and stiffness parameters, namely total unit weight γ, elastic modulus E_s, friction angle ϕ, Poisson's ratio ν_s and dilation angle ψ , As suggested by Brinkgreve et al. (2018), the value of cohesion c was set as 0.3 to avoid any complications during software calculations. Also, the aspect of increasing soil stiffness with an increase in depth was simulated by employing $E_{inc} = 500$ kN/m^3. The stiffness parameters of the buried conduit were selected as per the properties defined by Elshimi and Moore (2013). Table 2 details the parameters of different types of conduits, presented in terms of their normal stiffness EA, flexural rigidity EI , and Poisson's ratio ν_c. The strength interaction parameter $R_{\mathrm{int}er}$ was selected as 0.8 to simulate the realistic frictional resistance between the buried conduit and the surrounding soil (Wadi et al. 2015). The pressure q was applied on a surface footing of width B, that was located centrally above the buried conduit of diameter B_c. The footing was modelled as a plate, having the stiffness properties as; EA = 5×10^6 kN/m and EI = 8.5×10^3 kNm^2/m (Brinkgreve et al. 2018). The standard model fixities and the default medium mesh size was used for conducting finite element simulations. In order to obtain the settlement of the footing located over a buried conduit with a soil slope, the crest distance of the footing e/B, the burial depth of the conduit z/B , and the slope angle i were varied.

2.2. Database collection, preprocessing and feature validity

A database of 3600 full-scale numerical simulations was generated by conducting extensive finite element modelling. As suggested by McGrath (1998), the problem related to the buried conduits can be described in terms of the constrained modulus of soil M_s, defined as,

$$M_s = \frac{E_s(1-\nu_s)}{(1+\nu_s)(1-2\nu_s)} \tag{1}$$

where E_s and ν_s represent the elastic modulus and the Poisson's ratio of the soil.

Similarly, the stiffness parameters of the conduit, namely, normal stiffness EA and flexural rigidity EI are usually normalized to incorporate the effect of conduit diameter B_c and wall thickness t. The resulting hoop stiffness PS_H (Mcgrath 1999) and bending stiffness PS_B (ASTM D2412, 2002) of the conduit are defined as,

$$PS_H = \frac{EA}{R} \tag{2}$$

$$PS_B = \frac{EI}{0.149R^3} \tag{3}$$

where R is the radius to centroid of the conduit wall.

The aspect of slope stability and the angle of repose of the granular soil is a function of its friction angle (Duncan and Wright 2005; Atkinson 2007). The graphical presentation of the complete dataset is illustrated in the form of box and whisker plots in Figure 2. Moreover, the statistical properties of the same are tabulated in Table 3. The dataset has been normalized between -1 to 1 before feeding it to ML algorithms.

$$x_{std} = 2\left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right) - 1 \tag{4}$$

where x, $x_{\min}$ and $x_{\max}$ present the observed, minimum and maximum values of the dataset, respectively.

In machine learning, dealing with high-dimensional data is a challenging task for scientists and researchers. Feature reduction is an important step that effectively omits the redundant data and chooses the most optimum combination of input parameters (Jie et al. 2017; Gao et al. 2019). For this study, correlation-based feature selection, abbreviated as the CFS method, was implemented in a Waikato environment for knowledge analysis (WEKA) using the multivariate filter. Initially proposed by Hall (1999), CFS combines the correlation measure for appropriate feature subset selection and heuristic strategies for the mode of search. Therefore, it evaluates the importance/correlation of individual variables with the output and the degree of redundancy between them. The results of the feature selection depict that among the most relatively important parameters as summarized in Table 3, the applied load (q), constrained modulus of soil (M_s), unit weight of soil (γ), slope-angle ratio (i/ϕ), hoop stiffness (PS_H), bending stiffness (PS_B), burial depth ratio of conduit (z/B), and crest distance of footing (e/B) has achieved the highest importance. Mathematically, the output, that is settlement of footing located over a buried conduit in soil (s/B %) can be expressed as follows:

$$s/B = f(q, M_s, \gamma, i/\phi, PS_H, PS_B, z/B, e/B) \tag{5}$$

Therefore, these input parameters are utilized for training the machine learning models described in the next section.

2.3. Theory of methods

The theory of the statistical concepts and the data-driven machine learning methods employed in this study to estimate the settlement of the footing located over a conduit buried within a soil slope are provided in this section. Moreover, 3060 samples were randomly earmarked in this study for training the MLP, GPR, LKS, DT, and RF models. Thereafter, the competency of each model was evaluated and validated against 540 samples. Additionally, the research scheme employed in this study is presented in Figure 3.

2.3.1. Multi-layer perceptron (MLP)

Multi-layer perceptron (MLP) is a neural network that has the ability to map adaptive non-linear relationships between the input dataset and the output targets, thus making it one of the most widely used machine learning techniques (Azadi et al. 2013; Gao et al. 2019; Moayedi et al. 2019). Figure 4 shows a feed-forward MLP network consisting of an input layer, a single hidden layer, and an output layer. Each layer consists of a varying number of neurons. The number of independent input parameters and the output target defines the number of neurons in the input and output layers, respectively. Whereas the number of neurons in the hidden layer depends upon the type and size of the problem (Ramezanian et al. 2019). The increase in the number of hidden neurons may enhance the prediction ability of the network but can also make the model computationally inconvenient and complex (Raja and Shukla 2020). As a thumb rule, the maximum number of hidden neurons should be limited to $2m+1$, where m presents the number of input parameters (Shahin 2010). In an MLP algorithm, a number of neurons are connected by associated weights. At each neuron, the data from the input layer x_i is multiplied by the associated weight w_{ik}. Thereafter, the bias vector λ_k is added to the summation of the weighted inputs to obtain V_k. Finally, the output of the processing neuron y_k is obtained by passing V_k through the sigmoidal activation function $g(.)$ (see Figure. 4). More detailed information about the MLP neural network can be found in the existing literature (Gurney 1997).

2.3.2. Gaussian processes regression (GPR)

Gaussian processes regression (GPR) uses a probabilistic approach and predicts through kernel functions that evaluate on the basis of the similarity between two data points. The GPR technique integrates a number of machine learning tasks, such as model training, parameter estimation, and uncertainty evaluation. This helps in reducing the subjectivity of the GPR results and makes them more interpretable. The GPR is based on a Gaussian process (GP), that works on the assumption of Gaussian priors for changed function values (Rasmussen 2006). A GP can be statistically presented as,

$$g(x) \sim GP(\mu(x), k(x, x')) \quad (6)$$

where $\mu(x)$ presents the mean and $k(x, x')$ presents the covariance function of g.

Any finite number of random variables in a GP have a joint multivariate Gaussian distribution (Suthar 2020). Assuming $g = [\hat{g}(x_i, w)]_{i=1}^{m}$ presents the model outputs in correspondence to the input dataset X,

$$\hat{g}(x_i, w) = \sum_{j=1}^{n} w_j \phi_j(x_i), \qquad i = 1, 2, ..., m \quad (7)$$

or simply, if $g = \phi w$, then the prior distribution of g is Gaussian

$$p(g|X, \theta) \sim N(0, K) \quad (8)$$

where ϕ is the design matrix.

In this study, various kernel functions namely, radial bias function (RBF), Pearson VII Universal kernel function (PUKF), polynomial kernel function are employed, and the best results are obtained via PUKF function.

2.3.3. Lazy *K*-star (LKS)

Lazy *K*-star (LKS) uses an instant base learning (IBL) classification system to generalize the training dataset. During the learning process, the learning algorithm spends most of its computation time for consultation, and learners do not operate until the system receives the query call (Webb 2011). Unlike other machine learning techniques, LKS algorithm does not predict from the instances in the training dataset but rather employs the nearest neighbor approach to provide a response from the data memory (Altman 1992). The LKS classifies a dataset by drawing a comparison with a pre-classified sample. By employing a distance function, the IBL adds up all the possible transitions of two instances and categorizes them into a simple class (Cleary and Trigg 1995). Thereafter, the generated classification function is used to provide new solutions. For example, new test data samples x are distributed to the most suitable class among the k closest information focuses y_i. The corresponding LKS formulation can be given as follows (Cleary and Trigg 1995; Gao et al. 2019):

$$K^*(y_i, x) = -\log P^*(y_i, x) \tag{9}$$

where, P^* is the probability function that presents the all possible transitions from instances x to y.

2.3.4. Decision table (DT)

Decision table (DT) can be used to organize logic in a manner that helps in easy analysis (Nanda et al. 2017). A DT consists of different sections, namely, condition and action stubs, and condition and action entries. While the “condition stub” presents the possible conditions or problems, the “action stud” illustrates potential actions or solutions. The condition and action entries are located across the corresponding stubs, in terms of rules and classes tabulated in columns and rows, respectively (Cragun and Steudel 1987). When provided with a new instance, a DT algorithm tries to find the match in the table (Kohavi 1995). It assists in testing a set of rules for conditions of completeness, redundancy, and ambiguity. The condition of completeness occurs when the rules in the table address all the possible combinations of logic. Redundancy is said to exist when more than one rule having the same actions is satisfied by the same logical conditions. Ambiguity exists when two or more different rules with different actions are satisfied by the same logical conditions (Cragun and Steudel, 1987). In comparison to the hierarchical structure of the decision tree technique, the simple straightforward architecture of DT is considered to be more stable for problem solutions (Gao et al. 2019).

2.3.5. Random forest (RF)

Random forest (RF) uses an ensemble-learning approach that employs numerous classification trees for solving regression and classification problems (Ho 1995; Gehrke 2011). The RF creates a grove of trees whose predictive relationship alters randomly. The average output of each tree is then provided as the output. In order to generate the forest, the user has to define some parameters, such as the variable splitting the nodes and the number of trees. The accuracy of the model is determined in terms of the forest population i.e., the number of trees. The generalization error (GE) is estimated unbiasedly during the computation of the RF model. The RF evaluates the increase in the GE error to estimate the significance of the predictive variables. A variable is said to have increased significance if the value of GE increases. Also, by employing the bootstrap aggregating technique, the RF model reduces the risk of overfitting and provides a more stable solution (Breiman 2001)

3. Model performance and assessment

After the development and implementation phase, the next most important step is the model assessment. For data-driven modelling, the accuracy is measured in terms of the following: (*i*) Statistical criteria, that is, “goodness of fit”; and (*ii*) Robustness and sensitivity. The former deals with model performance by evaluating it fit to the calibration data using several statistical criteria. In contrast, the latter is used to access its accuracy, reliability, and rationality according to the underlying physical behavior of the investigated system. A model can only be considered suitable if it makes accurate and realistic predictions over a wide range of data (Shahin et al. 2009). Therefore, for this study, the best ML model was selected based on these criteria.

3.1 Model performance based on statistical indices

For “goodness of fit”, five statistical indices namely, Pearson’s correlation coefficient (r), root mean square error (RMSE), Nash-Sutcliffe efficiency (NSE), scatter index (SI), and relative percentage difference (RPD) were used to access the accuracy of the developed models. Moreover, based on these criteria, a ranking system was developed by assigning the scores to the models in training and testing dataset. In may be noted that this ranking system was successfully applied in many previous (Gao et al. 2019; Moayedi et al. 2020a; Zhang et al. 2020). The mathematical forms of all the indices, namely, r, RMSE, NSC, SI, and RPD are given in Eqs (10-14)

$$r = \frac{\sum (s/B_{iobs} - \overline{s/B_{obs}}) \times (s/B_{ipre} - \overline{s/B_{pre}})}{\sqrt{\sum_{i=1}^{n} (s/B_{iobs} - \overline{s/B_{obs}})^2} \sqrt{\sum_{i=1}^{n} (s/B_{ipre} - \overline{s/B_{pre}})^2}} \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(s/B_{iobs} - s/B_{ipre})^2} \tag{11}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(s/B_{iobs} - s/B_{ipre})^2}{\sum_{i=1}^{n}(s/B_{iobs} - \overline{s/B_{obs}})^2} \tag{12}$$

$$SI = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(s/B_{iobs} - s/B_{ipre})^2}}{\overline{s/B_{obs}}} \tag{13}$$

$$RPD = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(s/B_{iobs} - \overline{s/B_{obs}}}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(s/B_{iobs} - s/B_{ipre})^2}} \tag{14}$$

where s/B_{iobs}, s/B_{ipre}, $\overline{s/B_{obs}}$, $\overline{s/B_{pre}}$ and n represent the *ith* observed value of settlement, *ith* predicted value of settlement, mean value of observed settlement, mean value of predicted settlement, and number of data samples, respectively. It is to be noted that the model is considered to be accurate if it has high r, NSE, and RPD values and low RMSE and SI values.

Table 4 reports the results of all the statistical parameters (*r*, RMSE, NSE, SI and RPD) for training dataset in MLP, GPR, LKS, DT and RF were (0.977, 0.298, 0.937, 0.31, and 4.31), (0.931, 0.5, 0.851, 0.43, and 3.67), (0.901, 0.536, 0.76, 0.73, and 2.31), (0.92, 0.491, 0.831, 0.74, and 2.53), and (0.981, 0.273, 0.933, 0.35, and 3.93), respectively. Similarly, for testing dataset, for the same parameters, the values were (0.974, 0.323, 0.928, 0.44, and 3.75), (0.905, 0.518, 0.817, 0.76, and 2.34), (0.876, 0.673, 0.691, 1.01, and 1.8), (0.87, 0.613, 0.743, 1.04, and 1.97), and (0.964, 0.349, 0.916, 0.52, and 3.46) respectively for MLP, GPR, LKS, DT and RF (Table 5). After reviewing both the training and the testing performance, MLP technique can be introduced as the most accurate model in determining the settlement of footing located over a conduit buried within a soil slope. Moreover, the performance of MLP is followed by

RF and GPR models and thus, can be considered as the second and third best models in the hierarchy. Also, the LKS and DT have shown rather poor predictive performance in comparison to their counterparts.

The combined performance of all the models in training and testing datasets is computed in Table 6. In this regard, a total rank is obtained by summing the partial scores given to the model based on the statistical performance indicators, that are r, RMSE, NSE, SI and RPD values (Tables 4 and 5). From the results, the supremacy of the MLP model can be established with the highest total ranking score of 48. The second-best performance is obtained by RF with a total score of 42. The total ranking scores for GPR, LKS, and DT were 28, 12, and 20, respectively. Furthermore, Figures 5a-5e depict the regression correlation coefficient between the observed and predicted values for all the prescient models in the testing dataset. It can be observed from the regression chart that the MLP model has achieved the highest R^2, that is, 0.948, in comparison to 0.931, 0.820, 0.770, and 0.756, respectively for RF, GPR, LKS, and DT. This also proves that the developed MLP model has outperformed all other ML models applied in the context of predicting the settlement of footing located over a conduit buried within a soil slope.

The predictive performance of all the models was also accessed via Taylor's diagram in Figure 6. Taylor's diagram is a useful graphical tool to illustrate the accuracy of the developed data-driven models on a single platform (Taylor 2001). The strength between the predicted and simulated field is evaluated on the basis of the combine effect of three statistical parameters, that are, centered RMSE, correlation coefficient and standard deviation (SD). In the given figure, the solid black lines depict the correlation coefficient, solid radial lines represent the standard deviation, and dotted radial lines show the centered RMSE -values in the simulated field. The reference model is shown by a black dot with the correlation coefficient of unity, measured SD of 1.21, and zero centered RMSE. It can be seen from the figure that the best performance is obtained by MLP model with the correlation coefficient of 0.974, SD of 1.33, and RMSE of 0.3168. The RF model has also shown a good correlation strength with coefficient of 0.965, RMSE of 0.347, but the spatial variability is low with an SD of 1.02 in reference to observed value. The other models such as GPR, LKS and DT have correlation coefficient (0.905, 0.876, and 0.870), centered RMSE (0.517, 0.673, and 0.613), and SD (1.04, 0.73, and 0.92), respectively. This depicts that these models are associated with high bias and have poor prediction strength compared to the MLP model. Therefore, to this point, it is admissible that the developed MLP model predicts the settlement of footing located over a conduit buried within a soil slope in a reliable and intelligent way. Additionally, the time consumed by each approach is shown in Figure 7 . It can be observed that apparently, the MLP and GPR approaches had consumed less time in comparison to other models.

3.2 Model robustness and sensitivity

In this section, a sensitivity analysis was carried out to investigate the reliability and robustness of the developed MLP model. For this, incremental stepwise sensitivity analysis, also known as one-at-time analysis was conducted to examine the robustness of the model. In this method, each variable is increased in a stepwise manner while other variables remain constant at their mean value. However, this is practically not possible if the variables have both independent and correlated effect, that is, the change in one variable cause an inherent change in another variable such as unit weight and constrained modulus of soil, and hoop and bending stiffness of conduit, for this study. Therefore, the combined effect is calculated in the sensitivity analysis for these variables, as illustrated in Figure 8. The results show that an increase in the footing settlement increases significantly with an increase in the applied loading. This is a very common effect that relates to the movement of the underlying soil particles that slide along the shear failure planes due to the downward motion of the loaded footing, hence providing more settlement space for the surface footing (Terzaghi 1943). The constrained modulus and unit weight of the soil are interrelated variables that correspond to the density and the resulting stiffness of the soil. The increased soil stiffness reflects an increase in the shear strength of soil that provides resistance to the footing settlement (Berardi and Lancellotta 1991; Mayne and Poulos 1999). The slope angle ratio symbolizes the aspect of slope stability of an unconfined granular material in terms of the state of failure of a soil at which the angle of repose equals its friction angle (Al-Hashemi and Al-Amoudi 2018; Miura et al. 1997). An increase in the slope angle ratio increases the slope instability and hence reduces the support available to the footing on the slope side, explained in terms of the asymmetric footing failure mechanism(Hovan 1985; Graham et al. 1988). Due to the increase in asymmetric footing failure with an increase in slope angle ratio, the soil along the unstable slope yields under load application, resulting in increased footing settlement (Keskin and Laman 2013; Dey et al. 2019). The hoop and bending stiffness of the conduit are the design parameters that are employed to classify a conduit as either rigid or flexible (Mcgrath 1999). When buried under a loaded footing, the relative stiffness of the conduit to the adjacent soil determines the footing settlement. An increase in relative stiffness increases the stiffness of the soil located between the buried conduit and the overlying footing, thereby reducing the footing settlement (Srivastava et al. 2013). The effect of the burial depth of the conduit on the footing settlement is related to the aspect of the intersection of the shearing failure plane of the loaded footing with the buried conduit. As the burial depth increases and the conduit is located below the shear failure places of the footing, it serves as a support and reduces the settlement of the overlying footing (Khan and Shukla 2020). The increase in the crest distance relates to the support available to the surface footing on the slope side, as discussed above. An increase in the crest distance reduces the asymmetric nature of the failure mechanism, allowing more support to the slope side of the footing and causing a decrease in the footing settlement (Dey et al. 2019). The trends illustrated in Figure 8 comprehensively prove that the developed MLP model network correctly predicts the underlying physical behavior of the investigated system according to the known knowledge pertaining to geotechnical engineering, and thus can be considered reliable and robust.

In order to find the importance of each variable affecting the settlement of footing located over a conduit buried within a soil slope, a sensitivity analysis was also conducted using the Garson's algorithm (Garson 1991). In the case of a single hidden layered network, this technique involves the deconstruction of the model weight connections. The algorithm is explained in the Appendix section for the MLP network with eight inputs, six hidden layer nodes, and one output node. From the results illustrated in Figure 9, it can be observed that the most important parameter for estimating the settlement is applied load with the relative importance of 18.4%, followed by the unit weight of soil and constrained modulus of soil with relative importance of 16.3% and 15.3%, respectively. The relative importance of other parameters such as slope-angle ratio, hoop stiffness, bending stiffness, burial depth, footing crest distance is 13.8%, 11.4%, 9.9%, 7.7%, and 6.9%, respectively.

4. MLP model formulation

In this section, the developed optimal MLP model was translated into a trackable equation for hand or spreadsheet calculations. The mathematical form of MLP is given as follows (Ghorbani et al. 2020):

$$y = g_{ho}(\lambda_o + \sum_{j=1}^{h} w_{ko} g_{ih}(\lambda_{hk} + \sum_{i=1}^{m} w_{ik} x_i)) \qquad (15)$$

where g_{ho} is the applied transfer between hidden-output layer, λ_o is the bias at output layer node, w_{ko} is the synaptic weight between node k of hidden layer and single output node, g_{ih} is the applied transfer function between input-hidden layer, λ_{hk} is the bias value for node k of hidden layer (k = 1 ,h), w_{ik} is the synaptic weight between input i and node k of hidden layer, and x_i is the *ith* input node (variable). The weights and biases of the network are summarized in Table 7.

In order to predict the settlement of footing located over a buried conduit with eight inputs ($q, M_s, \gamma, i/\phi, PS_H, PS_B, z/B, e/B$), the optimal MLP model can be formulated as follows:

$$(s/B)_p = ((s/B)_{np} + 1) \times ((s/B)_{\max} - (s/B)_{\min})/2 + (s/B)_{\min} \tag{16}$$

where $(s/B)_{np}$, $(s/B)_{\max}$, and $(s/B)_{\min}$ are the normalized settlement value, maximum values of the settlement, and minimum value of the settlement, respectively. The normalized settlement value can be estimated as follows:

$$(s/B)_{np} = \sum_{k=1}^{8} w_{ko} sig(\beta)_k + \lambda_o \tag{17}$$

$$\beta_k = w_{1k} q_n + w_{2k} Ms_n + w_{3k} \gamma_n + w_{4k} (i/\phi)_n + w_{5k} PS_{H\,n} + w_{6k} PS_{Bn} + w_{7k} (z/B)_n + w_{8k} (e/B)_n + \lambda_{kh} \tag{18}$$

where the subscript n denotes the normalized values of the corresponding input parameters. The mathematical form of sigmoid activation function is given in Eq. (19).

$$sig(x) = \frac{1}{1 + e^{(-\beta)}} \tag{19}$$

For easy comprehension, the design numerical example is also presented below.

Numerical example

The 1-m wide footing is located over a conduit buried within a soil slope at 1.5 m depth below the base of footing. The crest distance of the footing is 1.75 m. Other parameters, including the constrained modulus of soil, unit weight of soil, slope-angle ratio, hoop stiffness of pipe, and bending stiffness of pipe are 35000 kPa, 19.9 kN/m^3, 4071.2 kPa, 8.55 kPa, respectively. Estimate the settlement of the footing under the application of load (q) of 50 kPa, 100 kPa, and 150 kPa.

Solution:

Input parameters (x_i) = $\{\ q \quad M_s \quad \gamma \quad i/\phi \quad PS_H \quad PS_B \quad z/B \quad e/B\}$

$= \{\ 50 \quad 35000 \quad 19.9 \quad 0.45 \quad 4071.2 \quad 8.55 \quad 1.5 \quad 1.75\ \}$

Step 1:

Normalize the values using Eq. (4). The maximum and minimum values of all the parameters are mentioned in Table 3.

$$x_{in} = \{ -0.333 \quad -0.6598 \quad -0.143 \quad -0.10 \quad -1 \quad -1 \quad -0.5 \quad 0.1667\ \}$$

Step 2:

Estimate normalized *s/B* value using Eq. (17). For that, calculate β_k using Eq. (18) as follows:

It may be noted that all weights and biases of MLP network are given in Table 7.

$$\beta_{k1} = (-0.333\times -0.9407) + (-0.6598\times -0.7507) + (-0.143\times -0.0944) + (-0.1\times -0.0095) + (-1\times 0.0712) + (1\times -0.0281) + (0.5\times -0.0873) + (0.1667\times -0.0221) + (-3.617) = -2.796$$

Similarly,

$$\{\beta_{k2}, \beta_{k3}, \beta_{k4}, \beta_{k5}, \beta_{k6}\} = \{-5.053, -6.655, 9.681, -6.331, -3.374\}$$

Now using Eq. (16) estimate the normalized settlement value.

$$(s/B)_{np} = \sum\left((-0.549\times \frac{1}{1+e^{-(-2.796)}}) + (0.0463\times \frac{1}{1+e^{-(-5.053)}}) + \ldots + (-1.021\times \frac{1}{1+e^{-(-3.374)}}) + 0.6951\right) = -0.95084$$

Step 3:

De-normalise using Eq. (17)

$$(s/B)_p = ((-0.95084)+1)\times(25.5-0.063)/2+0.0633 = 0.688\%$$

Step 4:

The settlement (s) is given as:

$$s = (0.688\times 1)/100 = 0.00688\,\text{m} = 6.88\text{ mm}$$

Similarly, for the applied loads of 100 kPa and 150 kPa, the settlement values will be 10.41 mm and 16.36 mm, respectively.

For future purposes, the developed MLP model can be combined with newly developed metaheuristics (e.g., Mirjalili et al. 2014; Abualigah and Diabat 2021; Abualigah et al. 2021b, c)

5. Conclusions

Settlement estimation of the footing located over a buried conduit in a sloping terrain is a challenging task for civil engineers. A novel approach is presented in this study to predict this settlement. It involves generating the pertaining database using extensive large-scale numerical simulations. Thereafter, five machine learning models (MLP, GPR, LKS, DT, and RF) were developed and implemented to evaluate the feasibility of the investigated system. The following general conclusions can be drawn from the above discussion.

1. For settlement estimation, results of all the statistical parameters (r, RMSE, NSE, SI and RPD) for training dataset in MLP, GPR, LKS, DT and RF were (0.977, 0.298, 0.937, 0.31, and 4.31), (0.931, 0.5, 0.851, 0.43, and 3.67), (0.901, 0.536, 0.76, 0.73, and 2.31), (0.92, 0.491, 0.831, 0.74, and 2.53), and (0.981, 0.273, 0.933, 0.35, and 3.93), respectively. Similarly, for testing dataset, for the same parameters, the values were (0.974, 0.323, 0.928, 0.44, and 3.75), (0.905, 0.518, 0.817, 0.76, and 2.34), (0.876, 0.673, 0.691, 1.01, and 1.8), (0.87, 0.613, 0.743, 1.04, and 1.97), and (0.964, 0.349, 0.916, 0.52, and 3.46) respectively for MLP, GPR, LKS, DT and RF. This indicates the superior predictive performance of the MLP model in contrast to other models.

2. The MLP model has obtained the highest ranking score (total score = 48). The next best performance is achieved by RF model (total score = 42) followed by GPR (total score = 28). Therefore, RF and GPR can be introduced as second and third best models in estimating the settlement of footings located over buried pipes in sloping terrain.

3. DT and LKS showed subpar performance in predicting the settlement with the total score of 20 and 12, respectively.

4. Sensitivity analysis was conducted using Garson's algorithm to assess the strength of input variables in estimating the output (i.e., settlement). The results showed that the applied load ranked 1st with the relative importance of 18.4%, followed by unit weight of soil and constrained modulus of soil with relative importance of 16.3% and 15.3%, respectively. The relative importance of other parameters such as slope-angle ratio, hoop stiffness, bending stiffness, burial depth, footing crest distance is 13.8%, 11.4%, 9.9%, 7.7%, and 6.9%, respectively.

5. The combined predictive performance of all the model were assessed via Taylor's diagram. Based on the results, the standard deviation (SD) (1.33, 1.02, 1.04, 0.73 and 0.92), RMSE (0.3168, 0.374, 0.517, 0.673, and 0.613), and correlation coefficient (*r*) (0.974, 0.965, 0.905, 0.876, and 0.870), respectively estimated MLP, RF, GPR, LKS and DT, confirm the predictive strength of the developed MLP model.

6. Robustness analysis and generalisation ability check showed that the settlement of footing over buried conduit in a sloping terrain increases with the increase in applied load and slope-angle ratio. Whereas the increase in the hoop stiffness, bending stiffness, burial depth, and footing crest distance causes the decrease in the footing's settlement.

Most importantly, the developed MLP model network has been translated into a functional relationship for easy hand or spreadsheet calculations. It can prove useful in saving the computational cost associated with intensive numerical simulations.

Limitations and future works

Although a wide range of data is utilized to train and validate the developed models, the models can be further improved by incorporating more data in the future. Moreover, the future research will also be dedicated in exploiting the deep learning techniques and hybrid ensemble learning approach to further increase the reliability of artificial intelligence-based modelling techniques in predicting the load-settlement behavior of the footing resting on buried conduit within a sloping ground.

Acknowledgments

This research is jointly funded by Higher Education Commision, Pakistan and Edith Cowan University, Australia.

Authorship contribution

Muhammad Umer Arif Khan: Writing - review & editing, Finite element modelling, Problem conceptulization. **Sanjay Kumar Shukla:** Supervision, Technical input. **Muhammad Nouman Amjad Raja:** Writing - review & editing, Statistical analysis, Validation, Data interpretation.

Compliance with ethical standards:

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical approval: This work does not contain any studies with human participants or animals performed by any of the authors.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Data availability statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Appendix

Garson' Algorithm for sensitivity analysis

Garson (1991) proposed a sensitivity analysis for calculating the variable importance as follows:

1. Calculate G_{ik} by multiplying the absolute values of hidden-output weight with the absolute value of input-hidden weight of each input variable *j*, that is, $|w_{ko} \times w_{ik}|$. e.g. From table 7 (G_{11} = -0.9407× -0.5494 = 0.5168)

2. For each hidden neuron, divide G_{ik} by the sum of all the input variables to obtain Q_{ik}:

$$Q_{ik} = G_{ik} / \sum_{k=1}^{m} G_{ik} \tag{20}$$

e.g., (Q_{11} = 0.5168/ (0.5168 + 0.4124 + 0.0518 + 0.0052 + 0.0391 + 0.0154 + 0.0479 + 0.0121) = 0.4694)

3. For each input neuron, obtain F_k as the sum of Q_{ik}:

$$F_k = \sum_{i=1}^{n} Q_{ik} \tag{21}$$

e.g., (F_{11} = 0.4694 + 0.1822 + 0.1684 + 0.0612 + 0.1667 + 0.0588 = 1.1069)

4. Calculate the percentage relative importance (R_I) of each variable as follows:

$$R_I = \left(F_k / \sum_{k=1}^{m} F_j \right) \times 100 \tag{22}$$

e.g., $\left(R_I = 100 \times \left(1.1069 / (1.1069 + 0.9173 + 0.9783 + 0.0604 + 0.8315 + 0.6852 + 0.5980 + 0.4664 + 0.4161)\right)\right)$ = 18.449 %

Complete calculations for variable importance are given below:

G_{ik}	0.517	0.412	0.052	0.005	0.039	0.015	0.048	0.012
	0.026	0.027	0.009	0.023	0.011	0.032	0.006	0.007
	0.826	0.528	0.104	1.658	0.451	0.270	0.160	0.906
	3.125	0.735	1.869	6.808	18.606	14.331	2.847	2.680
	0.873	0.582	0.638	0.864	0.421	0.049	1.261	0.550
	0.086	0.167	1.005	0.043	0.050	0.014	0.076	0.016
					Q_{ik}			
	0.4694	0.3746	0.0471	0.0047	0.0355	0.0140	0.0436	0.0110
	0.1823	0.1946	0.0620	0.1608	0.0783	0.2290	0.0414	0.0517
	0.1685	0.1077	0.0213	0.3381	0.0919	0.0550	0.0327	0.1848
	0.0613	0.0144	0.0366	0.1335	0.3648	0.2810	0.0558	0.0525
	0.1667	0.1112	0.1217	0.1649	0.0804	0.0094	0.2407	0.1051
	0.0588	0.1149	0.6896	0.0295	0.0343	0.0096	0.0524	0.0109
F_k	1.1070	0.9173	0.9783	0.8315	0.6852	0.5980	0.4665	0.4161
R_I	18.45	15.29	16.31	13.86	11.42	9.97	7.77	6.94
	q	*Ms*	γ	i / ϕ	PS_H	PS_B	*z/B*	*e/B*

References:

Aamir M, Tolouei-Rad M, Vafadar A, Raja MNA, Giasin K (2020) Performance analysis of multi-spindle drilling of Al2024 with TiN and TiCN coated drills using experimental and artificial neural networks technique. Appl Sci 10(23):8633

Abualigah L, Alkhrabsheh M (2021) Amended hybrid multi-verse optimizer with genetic algorithm for solving task scheduling problem in cloud computing. J Supercomput: 1-26

Abualigah L, Diabat A (2021) Advances in Sine Cosine Algorithm: A comprehensive survey. Artif Intell Rev 54:2567–2608

Abualigah L, Diabat A, Elaziz MA (2021a) Intelligent workflow scheduling for Big Data applications in IoT cloud computing environments. Cluster Comput: 1-20

Abualigah L, Diabat A, Mirjalili S, Abd Elaziz M, Gandomi AH (2021b) The Arithmetic Optimization Algorithm. Comput Methods Appl Mech Eng 376:113609

Abualigah L, Yousri D, Abd Elaziz M, Ewees AA, Al-qaness MA, Gandomi AH (2021c) Aquila Optimizer: A novel meta-heuristic optimization algorithm. Comput Ind Eng 157:107250

Ahmadi M, Jafarzadeh-Ghoushchi S, Taghizadeh R, Sharifi A (2019) Presentation of a new hybrid approach for forecasting economic growth using artificial intelligence approaches. Neural Comput Appl 31:8661–8680

Al-Naddaf M, Han J, Xu C, Jawad S, Abdulrasool G (2019) Experimental Investigation of Soil Arching Mobilization and Degradation under Localized Surface Loading. J Geotech Geoenvironmental Eng 145(12):04019114

Altman NS (1992) An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. The American Statistician 46(3):175-185

ASTM D2412 (2002) Standard Test Method for Determination of External Loading Characteristics of Plastic Pipe by Parallel-Plate Loading 1. Annu B ASTM Stand 02:2–7

Atkinson J (2007) The Mechanics of Soils and Foundations.Classification of soils, Book Chapter 5,86-99. CRC Press

Azadi M, Pourakbar S, Kashfi A (2013) Assessment of optimum settlement of structure adjacent urban tunnel by using neural network methods. Tunn Undergr Sp Technol 37:1-9

Bardhan A, Samui P, Ghosh K, Gandomi AH, Bhattacharyya S (2021) ELM-based adaptive neuro swarm intelligence techniques for predicting the California bearing ratio of soils in soaked conditions. Appl Soft Comput 110:107595

Beakawi Al-Hashemi HM, Baghabra Al-Amoudi OS (2018) A review on the angle of repose of granular materials. Powder Technol 330:397–417

Berardi R, Lancellotta R (1991) Stiffness of granular soils from field performance. Geotechnique 41(1):149–157

Bildik S, Laman M (2015) Experimental investigation of the effects of pipe location on the bearing capacity. Geomech Eng 8(2):221–235

Bildik S, Laman M (2019) Experimental Investigation of Soil — Structure — Pipe Interaction. KSCE J Civ Eng 23(9):3753–3763

Breiman L (2001) Random forests. Mach Learn 45(1):5–32

Brinkgreve, R. B., Kumarswamy S, Swolfs WM, Foria F (2018) Plaxis 2D Technical Manual. Rotterdam

Bryden P, El Naggar H, Valsangkar A (2015) Soil-Structure Interaction of Very Flexible Pipes: Centrifuge and Numerical Investigations. Int J Geomech 15(6):04014091

Cerato AB, Lutenegger AJ (2003) Scale effects of shallow foundation bearing capacity on granular material. BGA Int Conf Found Innov Obs Des Pract:217–225

Cleary JG, Trigg LE (1995) K*: An Instance-based Learner Using an Entropic Distance Measure. Mach Learn Proc 1995:108–114

Cragun BJ, Steudel HJ (1987) A decision-table-based processor for checking completeness and consistency in rule-based expert systems. Int J Man Mach Stud 26(5):633–648

Dey A, Acharyya R, Alammyan A (2019) Bearing capacity and failure mechanism of shallow footings on unreinforced slopes: a state-of-the-art review. Int J Geotech Eng:1-14

Dhar AS, Moore ID, McGrath TJ (2004) Two-Dimensional Analyses of Thermoplastic Culvert Deformations and Strains. J Geotech Geoenvironmental Eng 130(2):199–208

Dorosti S, Jafarzadeh Ghoushchi S, Sobhrakhshankhah E, Ahmadi M, Sharifi A (2020) Application of gene expression programming and sensitivity analyses in analyzing effective parameters in gastric cancer tumor size and location. Soft Comput 24:9943–9964

Duncan JM, Wright SG (2005) Soil strength and slope stability. Changes 19–30

Elshimi TM, Moore ID (2013) Modeling the Effects of Backfilling and Soil Compaction beside Shallow Buried Pipes. J Pipeline Syst Eng Pract 4(4):04013004

Gao W, Alsarraf J, Moayedi H, Shahsavar A, Nguyen H (2019) Comprehensive preference learning and feature validity for designing energy-efficient residential buildings using machine learning paradigms. Appl Soft Comput J 84:105748

Garson DG (1991) Interpreting neural network connection weights. Artif Intell Expert 6:47–51

Gehrke J (2011) Classification and Regression Trees. Encycl Data Warehous Min 141–143

Ghazavi M, Eghbali AH (2008) A simple limit equilibrium approach for calculation of ultimate bearing capacity of shallow foundations on two-layered granular soils. Geotech Geol Eng 26(5):535–542

Ghorbani B, Arulrajah A, Narsilio G, Horpibulsuk S, Bo MW (2021) Thermal and mechanical properties of demolition wastes in geothermal pavements by experimental and machine learning techniques. Constr Build Mater 280:122499

Ghorbani B, Arulrajah A, Narsilio G, Horpibulsuk S, Bo MW (2020) Development of genetic-based models for predicting the resilient modulus of cohesive pavement subgrade soils. Soils Found 60(2):398–412

Graham J, Andrews M, Shields DH (1988) Stress characteristics for shallow footings in cohesionless slopes. Can Geotech J 25:238–249

Gurney K (1997) An Introduction to Neural Networks, 1st edn. UCL Press Limited, New York

Hall MA (1999) Correlation-based Feature Selection for Machine Learning. Doctoral Thesis, The University of Waikato, New Zealand

Ho TK (1995) Random decision forests. In: Proceedings of the International Conference on Document Analysis and Recognition: 278–282

Hovan JM (1985) Computation of bearing capacity and passive pressure coefficients in sand using stress-characteristics and critical state. MSc Thesis. University of Manitoba, Canada

Jie C, Jiawei L, Shulin W, Sheng Y (2017) Feature selection in machine learning: A new perspective. Neurocomputing 300:70-79

Kaloop MR, Bardhan A, Kardani N, Samui P, Hu JW, Ramzy A (2021) Novel application of adaptive swarm intelligence techniques coupled with adaptive network-based fuzzy inference system in predicting photovoltaic power. Renew Sustain Energy Rev 148:111315

Kardani N, Zhou A, Nazem M, Shen SL (2020) Improved prediction of slope stability using a hybrid stacking ensemble method based on finite element analysis and field data. J Rock Mech Geotech Eng 13(1):188-201

Kardani N, Bardhan A, Gupta S, Samui P, Nazem M, Zhang Y, Zhou A (2021a) Predicting permeability of tight carbonates using a hybrid machine learning approach of modified equilibrium optimizer and extreme learning machine. Acta Geotech:1-17

Kardani N, Bardhan A, Roy B, Samui P, Nazem M, Armaghani DJ, Zhou A (2021b) A novel improved Harris Hawks optimization algorithm coupled with ELM for predicting permeability of tight carbonates. Eng Comput:1-24

Keskin S, Laman M (2013) Model studies of bearing capacity of strip footing on sand slope. KSCE J Civ Eng 17(4):699–711

Khan MUA, Shukla SK (2020) Load–Settlement Response and Bearing Capacity of a Surface Footing Located Over a Conduit Buried Within a Soil Slope. Int J Geomech 20(10):04020173

Khan MUA, Shukla SK (2021a) Vertical load on a conduit buried under a sloping ground. Geomech Eng 24(6):599–610

Khan MUA, Shukla SK (2021b) Numerical investigation of the structural response of a conduit buried within a soil slope. Transp Geotech 30:100614

Khan MUA, Shukla SK, Raja MNA (2021) Soil–conduit interaction: an artificial intelligence application for reinforced concrete and corrugated steel conduits. Neural Comput Appl:1-25

Kim MK, Cho SH, Yun IJ, Won JH (2012) Three-dimensional responses of buried corrugated pipes and ANN-based method for predicting pipe deflections. Int J Numer Anal Methods Geomech 36(1):1–16

Kohavi R (1995) The power of decision tables. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics):174–189

Mayne PW, Poulos HG (1999) Approximate Displacement Influence Factors for Elastic Shallow Foundations. J Geotech Geoenvironmental Eng 125(6):453–460

Mcgrath TJ (1999) Calculating loads on buried culverts based on pipe hoop stiffness. Transp Res Rec 1656(1):73–79

McGrath TJ (1998) Replacing E′ with the constrained modulus in flexible pipe design. In: Proceedings of the Pipeline Division Conference:28–40

Mirjalili S, Mirjalili SM, Lewis A (2014) Grey Wolf Optimizer. Adv Eng Softw 69:46–61

Miura K, Maeda K, Toki S (1997) Method of Measurement for the Angle of Repose of Sands. Soils Found 37(2):89–96

Moayedi H, Aghel B, Foong LK, Bui DT (2020a) Feature validity during machine learning paradigms for predicting biodiesel purity. Fuel 262:116498

Moayedi H, Hayati S (2018) Modelling and optimization of ultimate bearing capacity of strip footing near a slope by soft computing methods. Appl Soft Comput J 66:208–219

Moayedi H, Moatamediyan A, Nguyen H, Bui XN, Bui DT, Rashid ASA (2020b) Prediction of ultimate bearing capacity through various novel evolutionary and neural network models. Eng Comput 36(2):671–687

Moayedi H, Nguyen H, Rashid ASA (2019) Novel metaheuristic classification approach in developing mathematical model-based solutions predicting failure in shallow footing. Eng Comput 37(1):223-230

Moser AP, Folkman S (2001) Buried Pipe Design. The McGraw-Hill Companies, New York

Nanda S, Zafari F, Decusatis C, Wedaa E, Yang B (2017) Predicting network attack patterns in SDN using machine learning approach. In: 2016 IEEE Conference on Network Function Virtualization and Software Defined Networks:167–172

Nguyen H, Mehrabi M, Kalantar B, Moayedi H, Abdullahi MAM (2019) Potential of hybrid evolutionary approaches for assessment of geo-hazard landslide susceptibility mapping. Geomatics, Nat Hazards Risk 10(1):1667–1693

Raja MNA, Shukla SK (2021a) Predicting the settlement of geosynthetic-reinforced soil foundations using evolutionary artificial intelligence technique. Geotext Geomembranes 49:1280–1293

Raja MNA, Shukla SK (2021b) Multivariate adaptive regression splines model for reinforced soil foundations. Geosynth Int 28(4):368–390

Raja MNA, Shukla SK (2020) An extreme learning machine model for geosynthetic-reinforced sandy soil foundations. Proc Inst Civ Eng - Geotech Eng:1–42

Raja MNA, Shukla SK, Khan MUA (2021) An intelligent approach for predicting the strength of geosynthetic-reinforced subgrade soil. Int J Pavement Eng:1-17

Ramezanian R, Peymanfar A, Ebrahimi SB (2019) An integrated framework of genetic network programming and multi-layer perceptron neural network for prediction of daily stock return: An application in Tehran stock exchange market. Appl Soft Comput J 82:105551

Rasmussen CE (2006) Gaussian Processes in Machine Learning. MIT press, Cambridge

Robert DJ, Soga K, O'Rourke TD, Sakanoue T (2016) Lateral Load-Displacement Behavior of Pipelines in Unsaturated Sands. J Geotech Geoenvironmental Eng 142(11):04016060

Sedran G, Stolle DFE, Horvath RG (2001) An investigation of scaling and dimensional analysis of axially loaded piles. Can Geotech J 38(3):530–541

Shahin MA (2010) Intelligent computing for modeling axial capacity of pile foundations. Can Geotech J 47(2):230–243

Shahin MA, Jaksa MB, Maier HR (2009) Recent Advances and Future Challenges for Artificial Neural Systems in Geotechnical Engineering Applications. Adv Artif Neural Syst 2009:1–9

Shokouhi SKS, Dolatshah A, Ghobakhloo E (2013) Seismic strain analysis of buried pipelines in a fault zone using hybrid FEM-ANN approach. Earthq Struct 5(4):417–438

Srivastava A, Goyal CR, Raghuvanshi A (2013) Load Settlement Response of Footing Placed over Buried Flexible Pipe through a Model Plate Load Test. Int J Geomech 13(4):477–481

Suthar M (2020) Applying several machine learning approaches for prediction of unconfined compressive strength of stabilized pond ashes. Neural Comput Appl 32(13):9019–9028

Talesnick ML, Xia HW, Moore ID (2011) Earth pressure measurements on buried HDPE pipe. Geotechnique 61(9):721-732

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res Atmos 106:7183–7192

Terzaghi K (1943) Theoretical Soil Mechanics. John Wiley and Sons, New York

Wadi A, Pettersson L, Karoumi R (2015) Flexible culverts in sloping terrain: Numerical simulation of soil loading effects. Eng Struct 101:111–124

Wang F, Han J, Corey R, Parsons RL, Sun X (2017) Numerical Modeling of Installation of Steel-Reinforced High-Density Polyethylene Pipes in Soil. J Geotech Geoenvironmental Eng 143(11):04017084

Webb GI (2011) Lazy Learning. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA.

Xiao F, Zhao Z (2019) Evaluation of equivalent hydraulic aperture (EHA) for rough rock fractures. Can Geotech J 56(10):1486–1501

Yekani Motlagh S, Sharifi A, Ahmadi M, Badfar H (2019) Presentation of new thermal conductivity expression for Al 2 O 3 –water and CuO –water nanofluids using gene expression programming (GEP). J Therm Anal Calorim 135(1):195–206

Zhang X, Nguyen H, Bui XN, Le HA, Nguyen-Thoi T, Moayedi H, Mahesh V (2020) Evaluating and Predicting the Stability of Roadways in Tunnelling and Underground Space Using Artificial Neural Network-Based Particle Swarm Optimization. Tunn Undergr Sp Technol 103:103517

Table 1. Properties of soils used in the finite element model

Soil type	Total unit weight, γ	Elastic modulus, E_s	Friction angle, ϕ	Poisson's ratio, ν_s	Dilation angle, ψ
	(kN/m^3)	(MPa)	(degree)	-	(degree)
S_1	19.0	17.5	30	0.333	3.4
S_2	19.9	25.0	33	0.313	5.8
S_3	20.5	35.0	36	0.291	8.0
S_4	20.9	50.0	39	0.270	10.0
S_5	21.1	65.0	42	0.249	11.5

Table 2. Properties of the different conduit materials used in the finite element model

Conduit material		Parameter					
		D_{inner}	t	B_c	EA	EI	ν_c
		m	mm	m	kN/m	kNm^2/m	-
Reinforced concrete	RC_1	2.0	190.5	2.38	5.7×10^6	1.7×10^4	0.3
	RC_2	2.0	100.0	2.2	3.0×10^6	2.5×10^3	0.3
Corrugated steel	CS_1	2.0	60.02	2.12	7.0×10^5	211.5	0.28
	CS_2	2.0	32.14	2.06	3.1×10^5	26.7	0.28
High density polyethylene	HDPE	2.0	63.26	2.13	4.2×10^3	1.4	0.46

Table 3. Statistical details of various input and output parameters

Parameter	Symbol	Min	Max	Mean	SD
Applied pressure (kPa)	q	25	100	62.5	27.95
Constrained modulus of soil (kPa)	M_s	26217	77855	49502.1	18645.9
Total unit weight of soil (kN/m^3)	γ	19	21.1	20.28	0.76
Dilation angle of soil (degrees)	ψ	0	11.5		
Hoop stiffness of conduit (kPa)	PS_H	4071	5204291	1808684	1970423
Bending stiffness of conduit (kPa)	PS_B	8.56	86841	20541.9	33591.1
Poisson's ratio of the conduit	ν_c	0.28	0.46		
Slope angle ratio	i/ϕ	0	1		
Burial depth of the conduit	z/B	1	3	2	0.82
Crest distance of the footing	e/B	0	3	1.5	1.11
Footing settlement (%)	s/B	25.5	0.06	0.706	1.42

SD: Standard deviation

Table 4: Performance and ranking of all the machine learning models in training dataset

Statistical indices	Network performances in training dataset				
	MLP	GPR	LKS	DT	RF
Pearson r	0.977	0.931	0.901	0.92	0.981
RMSE	0.298	0.5	0.536	0.491	0.273
NSE	0.937	0.851	0.76	0.831	0.933
SI	0.31	0.43	0.73	0.74	0.35
RPD	4.31	3.67	2.31	2.53	3.93
	Partial scores of the models				
	MLP	GPR	LKS	DT	RF
Pearson r	4	3	1	2	5
RMSE	5	2	1	3	4
NSE	5	3	1	2	4
SI	5	3	2	1	4
RPD	5	3	1	2	4
Total ranking score	24	14	6	10	21

Table 5: Performance and ranking of all the machine learning models in testing dataset

Statistical indices	Network performances in testing dataset				
	MLP	GPR	LKS	DT	RF
Pearson r	0.974	0.905	0.876	0.87	0.964
RMSE	0.323	0.518	0.673	0.613	0.349
NSE	0.928	0.817	0.691	0.743	0.916
SI	0.44	0.76	1.01	1.04	0.52
RPD	3.75	2.34	1.8	1.97	3.46
	Partial scores of the models				
	MLP	GPR	LKS	DT	RF
Pearson r	5	3	2	1	4
RMSE	5	3	1	2	4
NSE	5	3	1	2	4
SI	5	3	2	1	4
RPD	5	3	1	2	4
Total ranking score	25	15	7	8	20

Table 6: Final ranking of all the proposed machine learning models

Dataset	Statistical indices	Partial ranking scores				
		MLP	GPR	K-star	DT	RF
Training	Pearson r	4	3	1	2	5
	RMSE	4	2	1	3	5
	NSE	5	3	1	2	4
	SI	5	3	2	1	4
	RPD	5	3	1	2	4
Testing	Pearson r	5	3	1	2	4
	RMSE	5	2	1	3	4
	NSE	5	3	1	2	4
	SI	5	3	2	1	4
	RPD	5	3	1	2	4
Total ranking score		48	28	12	20	42
Final rank		1	3	5	4	2

Table 7: Weights and biases of the developed MLP network

Weights of input layer - hidden layer, w_{ik}								Hidden layer bias λ_k
1	2	3	4	5	6	7	8	
-0.9407	-0.7507	-0.0944	-0.0095	0.0712	-0.0281	-0.0873	-0.0221	-3.617
0.553	0.5903	0.1881	0.4877	0.2375	0.6948	-0.1255	0.157	-3.561
1.524	-0.9739	-0.1923	3.059	-0.8318	0.4979	-0.2954	-1.672	-6.715
-1.978	0.465	1.183	-4.309	11.776	-9.07	1.8021	1.696	12.391
0.93	-0.62	0.679	0.9196	0.4482	-0.0522	-1.3427	-0.586	-6.419
-0.0839	-0.164	0.9839	-0.0421	0.0489	0.0137	0.0747	0.0156	-3.277
Weights of hiden-ouput layer, w_{ko}								Output layer bias λ_o
	-0.549	0.0463	0.542	-1.58	-0.939	-1.021	-0.549	0.6951

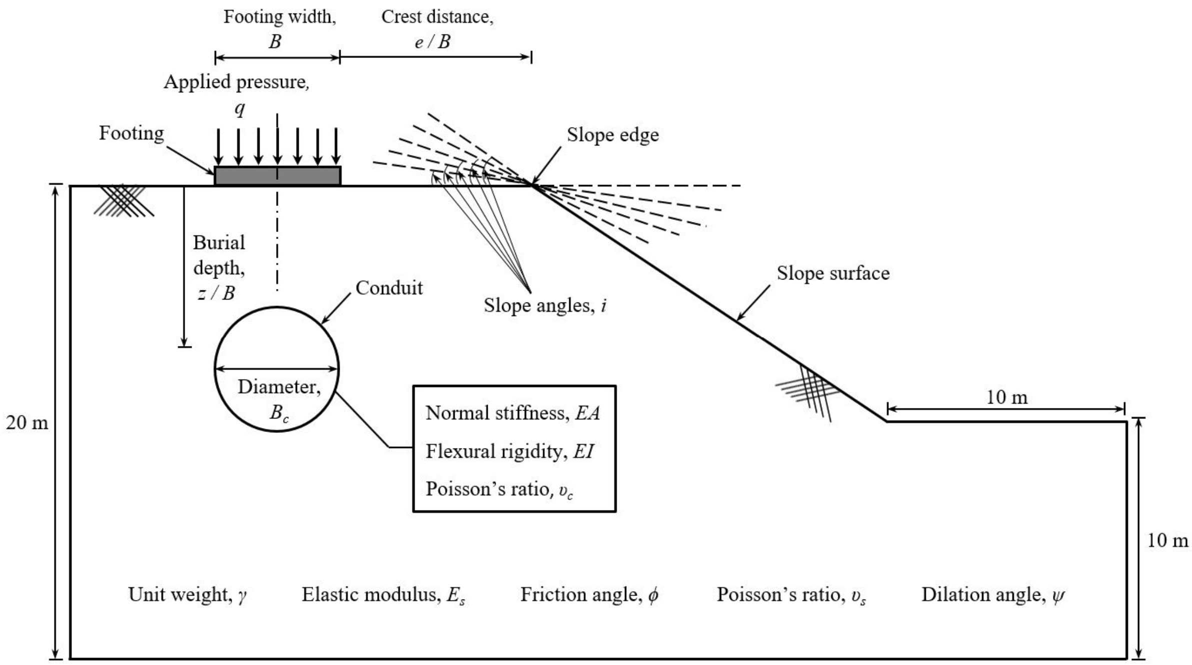


Fig. 1. Large-scale slope model used for the FEM analysis

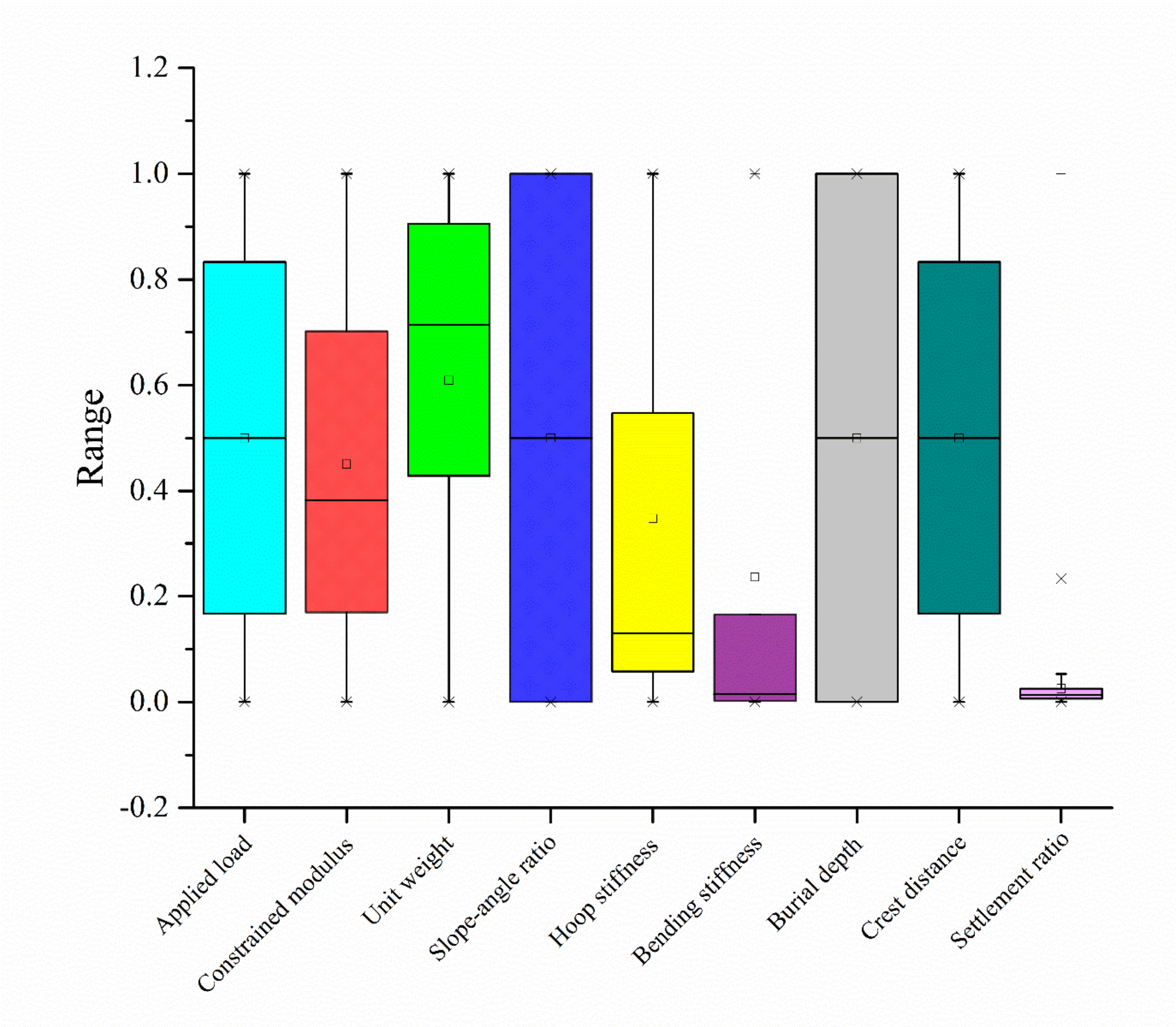


Fig. 2. Box and whisker plots of the dataset

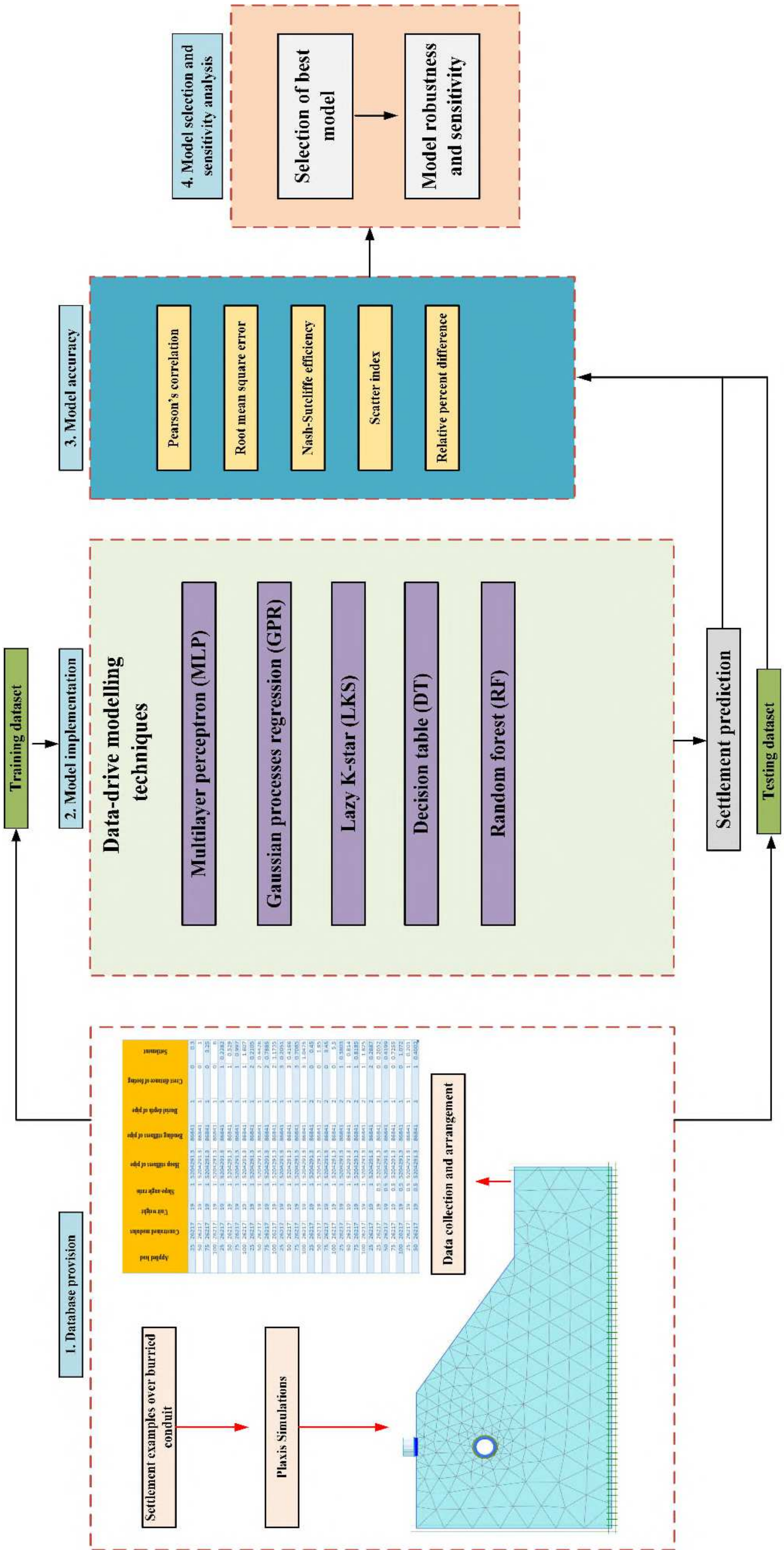


Fig. 3. Research framework employed in the present study

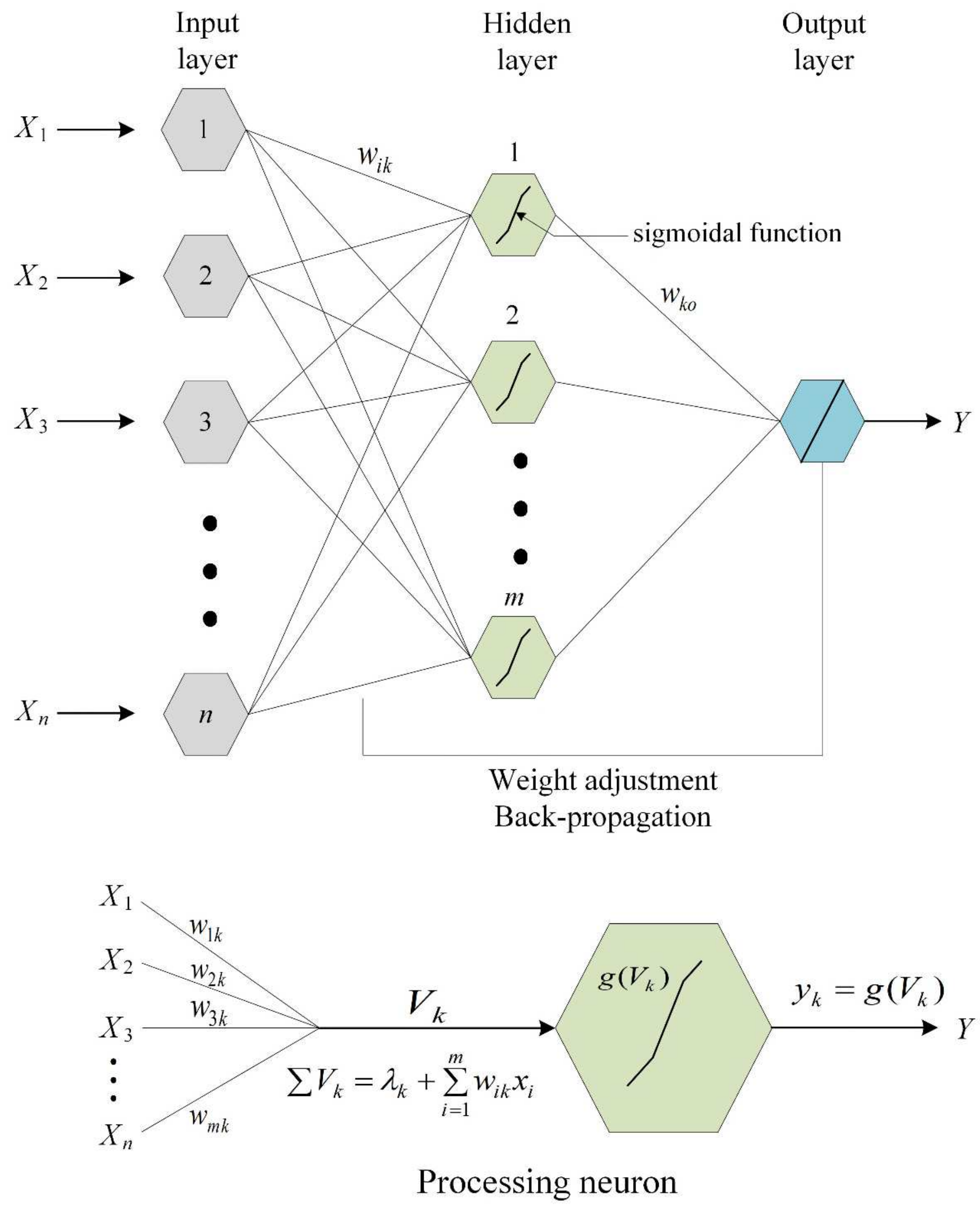
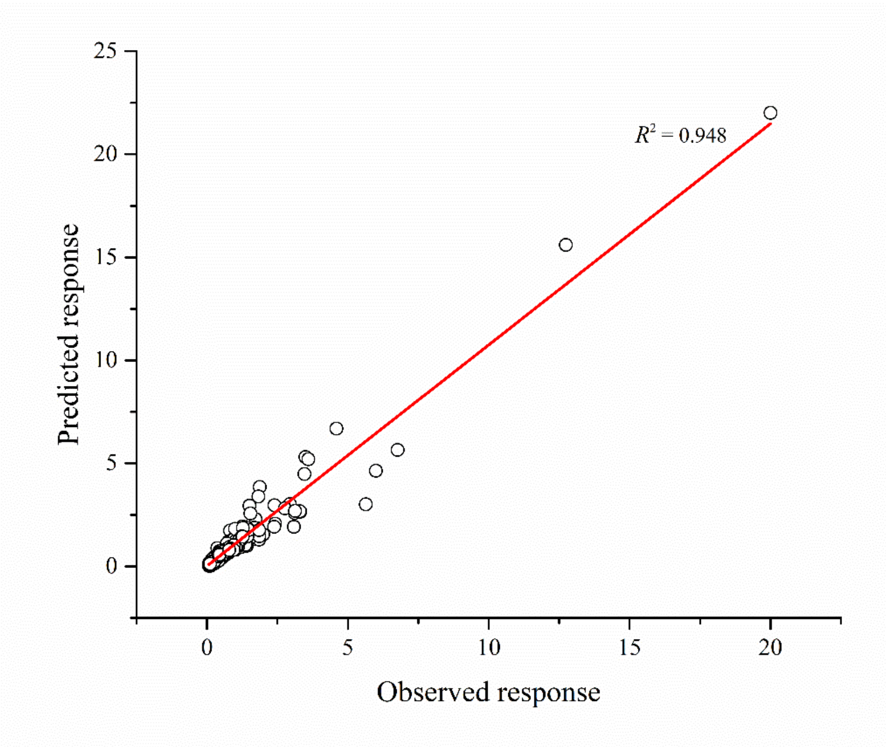
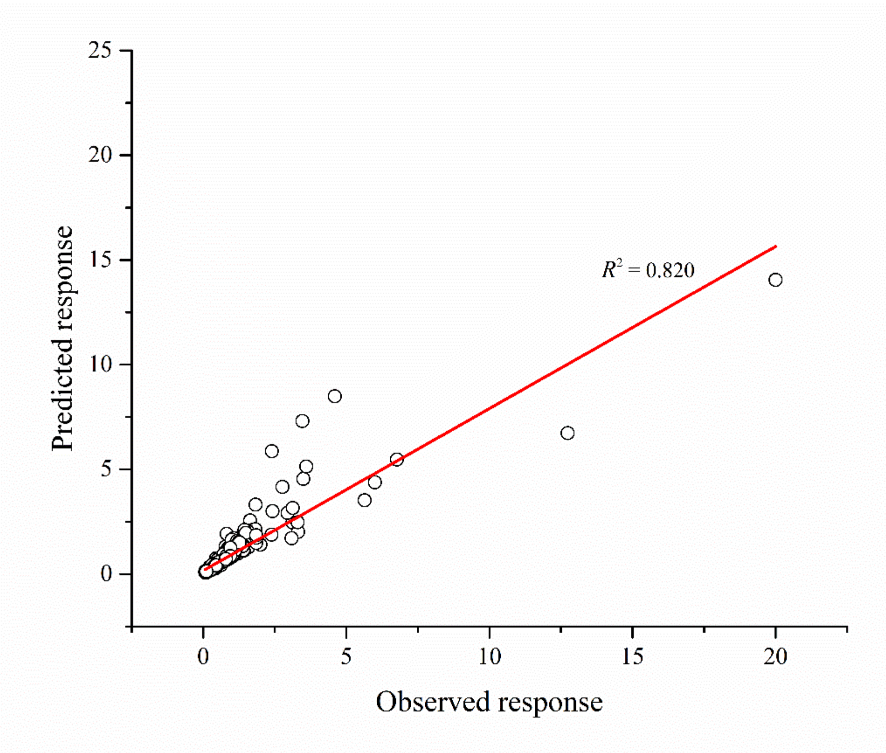


Fig. 4. Feed-forward MLP structure (adapted from Shahin, 2010)

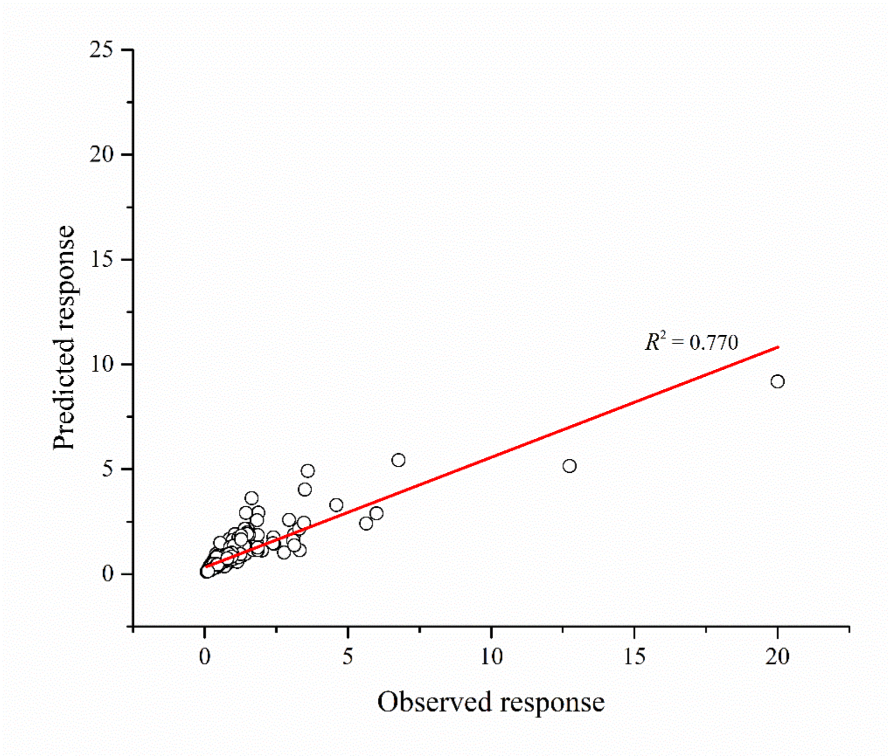


(a)

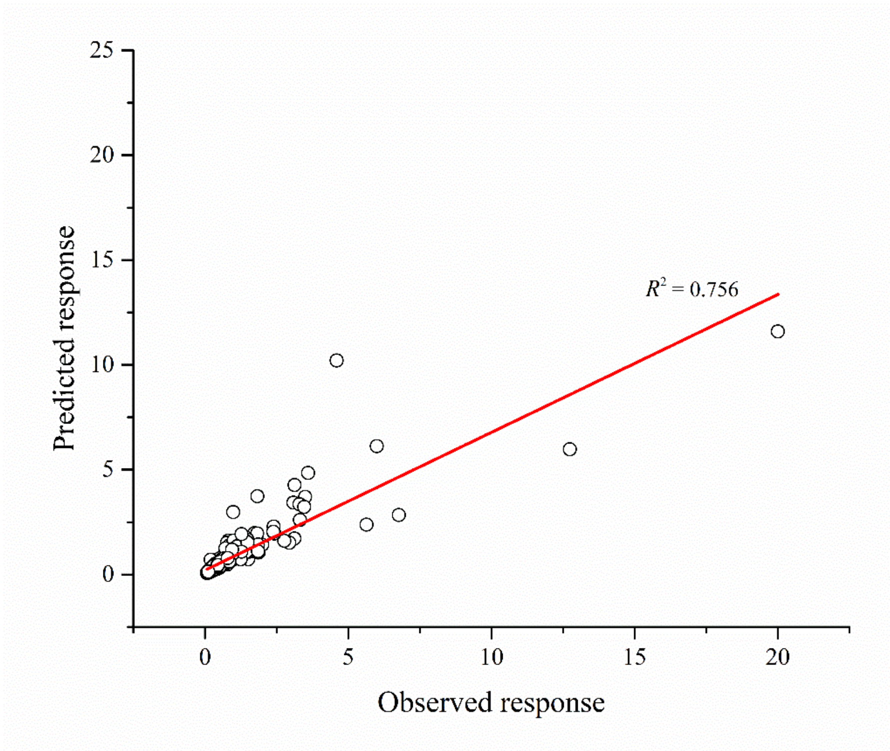


(b)

Fig. 5. Correlation between observed and predicted settlement values for: **(a)** MLP; **(b)** GPR; **(c)** LKS; **(d)** DT and **(e)** RF

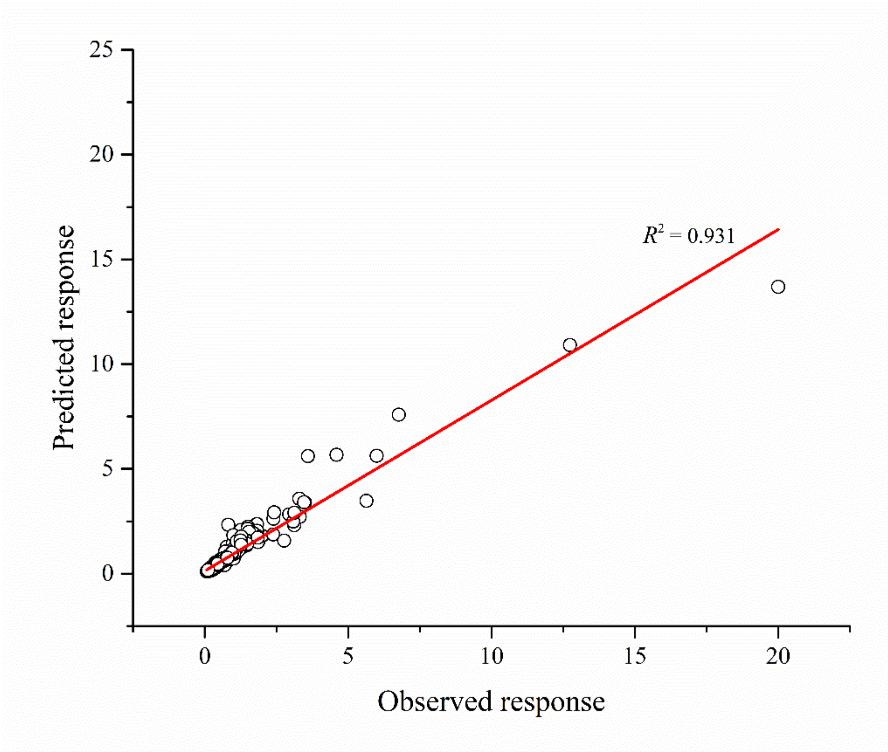


(c)



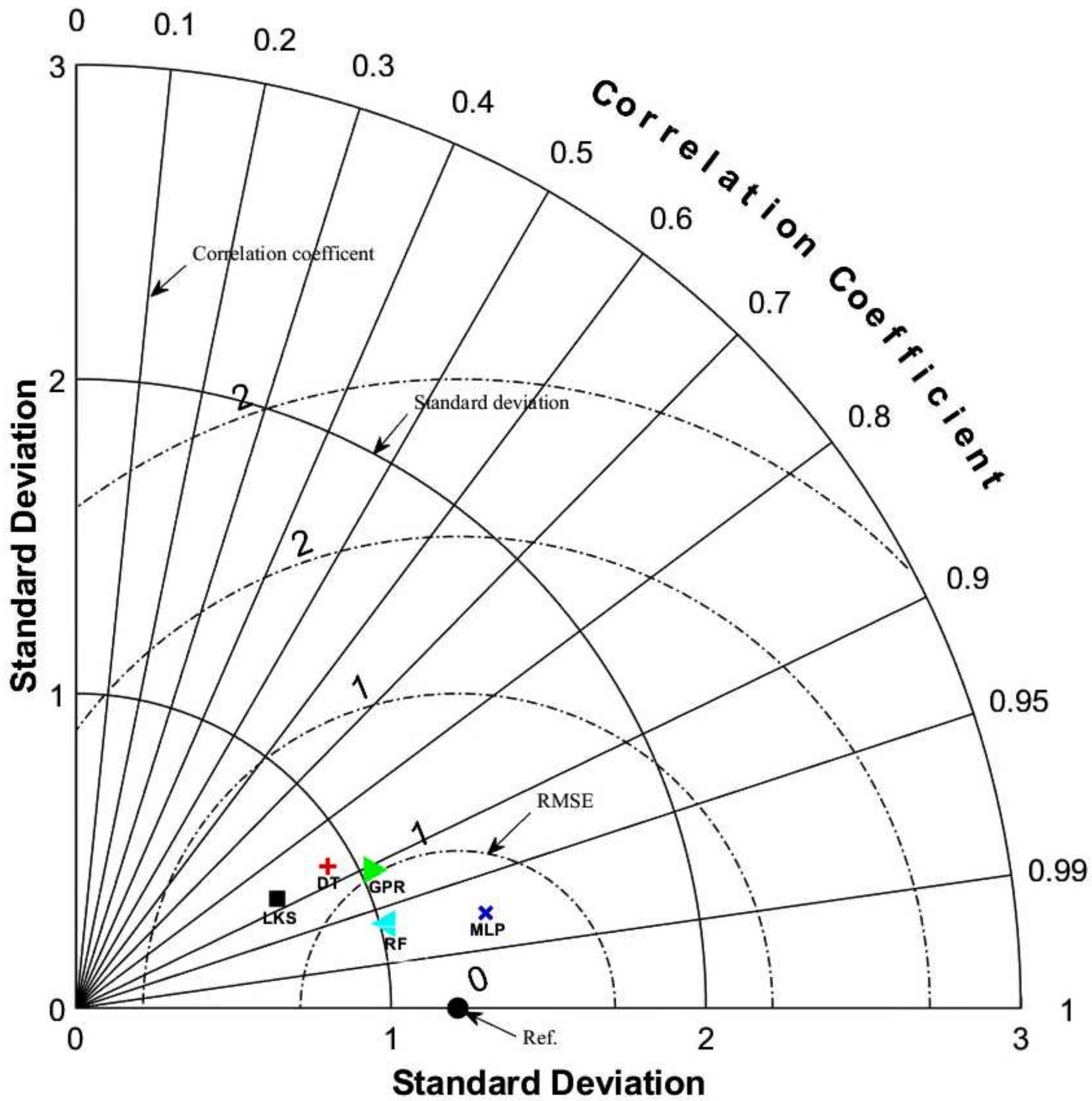
(d)

Fig. 5. continued



(e)

Fig. 5. continued



1098 **Fig. 6.** Taylor's diagram for all the data-driven modelling techniques

1099

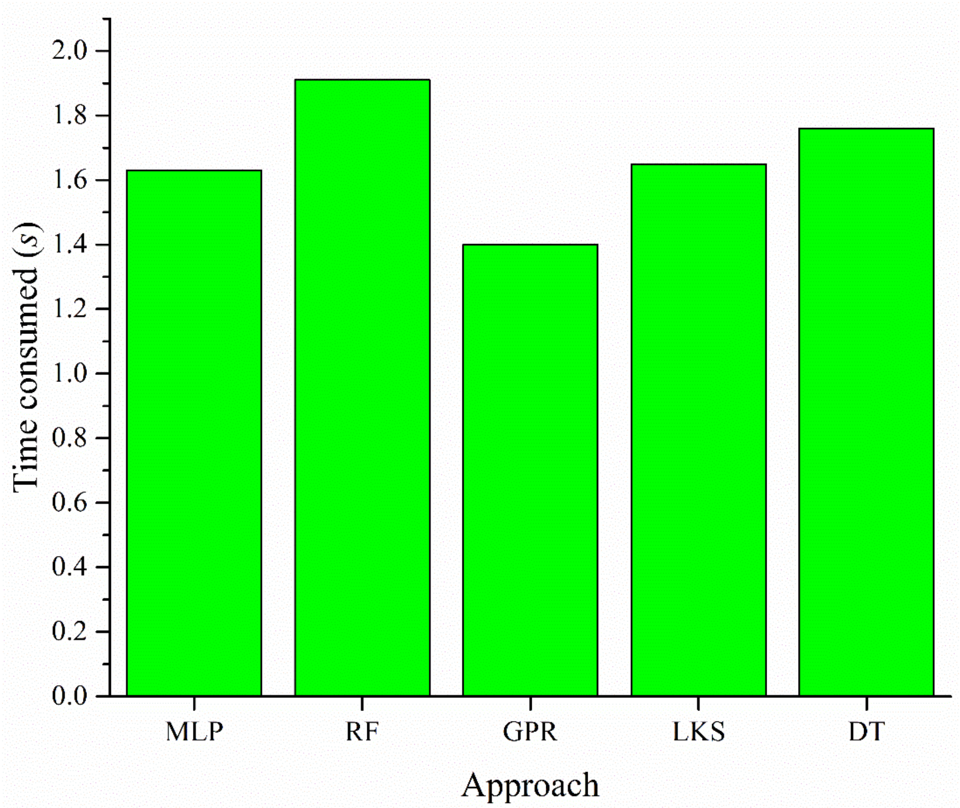


Fig.7: Time consumption of various approaches

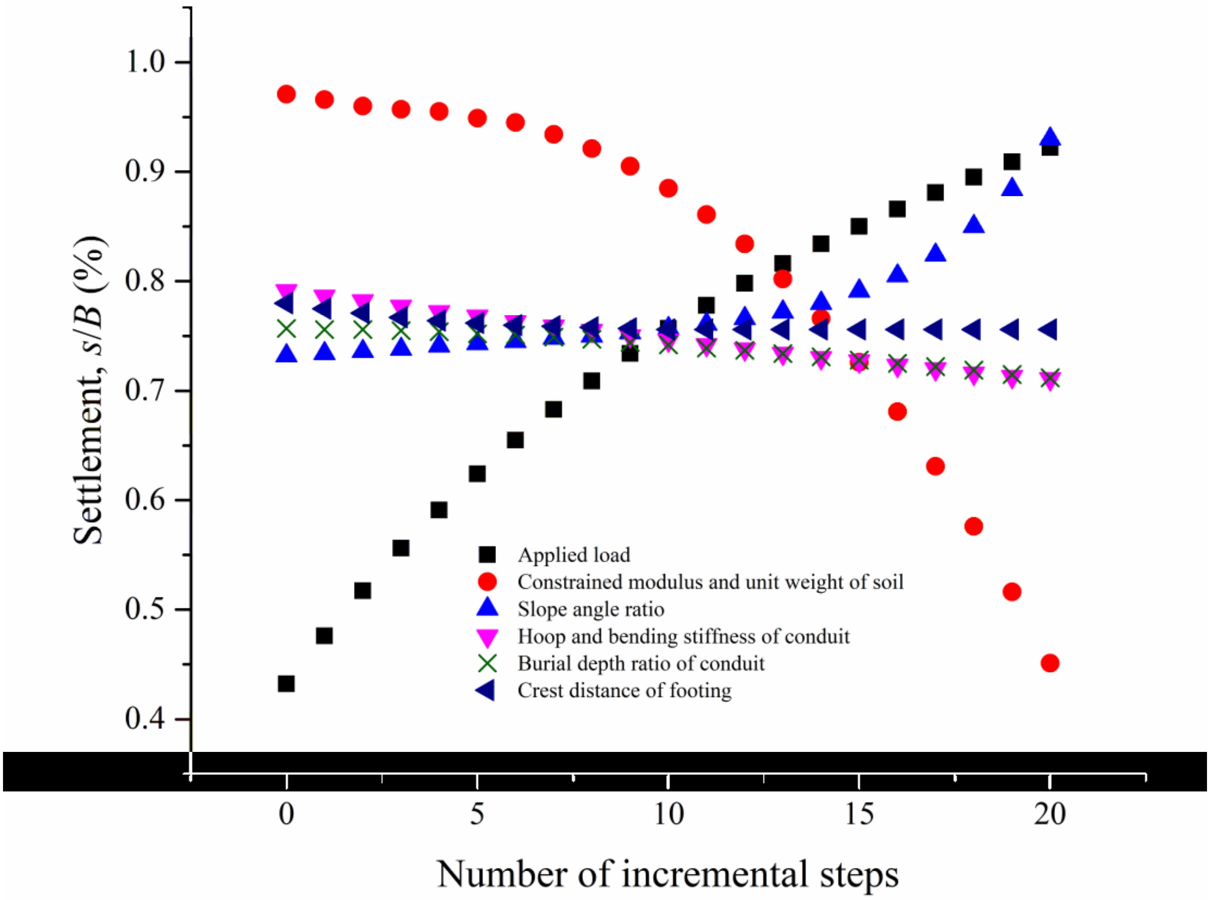


Fig. 8. Reliability and robustness analysis of the developed MLP model

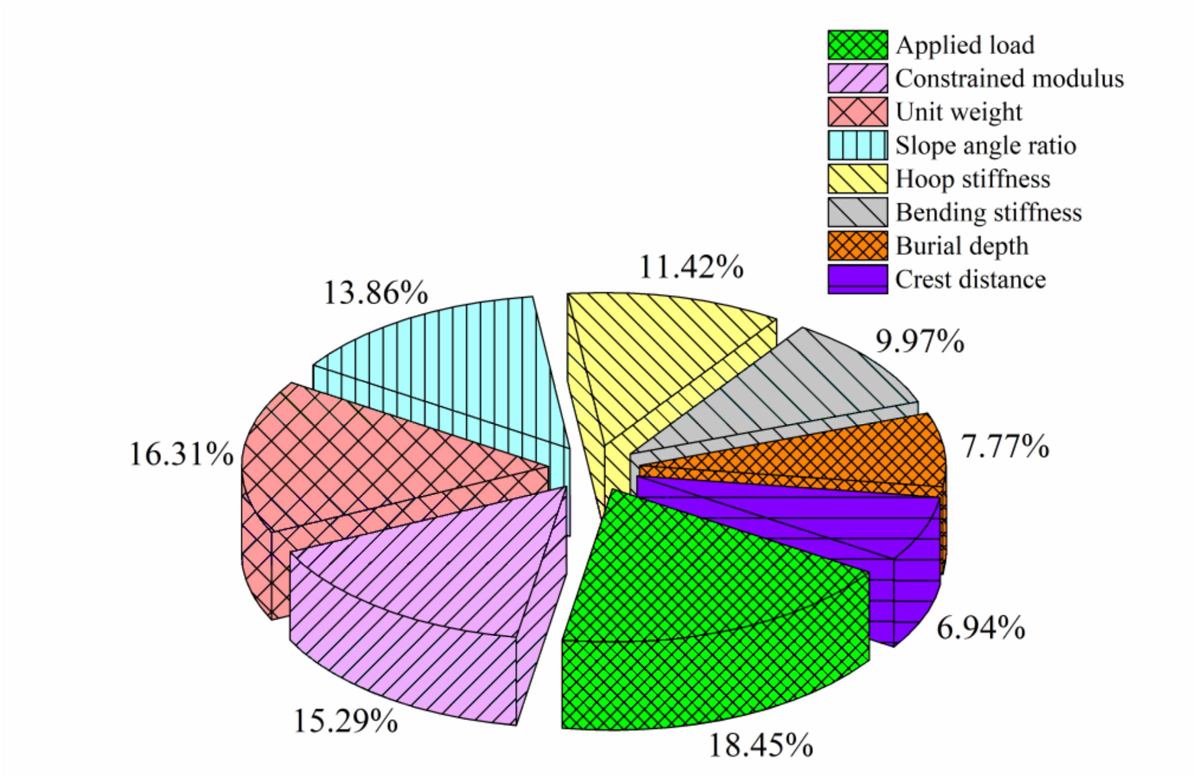


Fig. 9. Sensitivity analysis according to the Garson's algorithm of the MLP model