Computers and Geotechnics 66 (2015) 1–15

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo

Research Paper

Evolutionary metaheuristic intelligence to simulate tensile loads in reinforcement for geosynthetic-reinforced soil structures

Jui-Sheng Chou [a,*], Kuo-Hsin Yang [a], Jusieandra Pribadi Pampang [a], Anh-Duc Pham [a,b]

[a] Department of Civil and Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan
[b] Faculty of Project Management, The University of Danang, University of Science and Technology, Danang, Vietnam

ARTICLE INFO

Article history:
Received 23 July 2014
Received in revised form 31 December 2014
Accepted 1 January 2015

Keywords:
Reinforcement loads
Geosynthetic-reinforced soil structures
Machine learning
Optimisation
Evolutionary metaheuristic intelligence

ABSTRACT

The accurate estimation of reinforcement tensile loads is crucial for the evaluation of the internal stabilities of geosynthetic-reinforced soil (GRS) structures. This study developed an evolutionary metaheuristic intelligence model for efficiently and accurately estimating reinforcement loads. The proposed model improves the prediction capability of the firefly algorithm (FA) by integrating intelligent components, namely, a chaotic map, an adaptive inertia weight, and a Lévy flight. The enhanced FA is then used to optimise the hyperparameters for a least squares support vector regression model. The proposed model was validated using a database of 15 wall case studies (94 data points in total) *via* a cross-validation algorithm. The method was then compared with conventional prediction methods in terms of the accuracy for predicting the reinforcement tensile loads of GRS structures. The cross-validation results demonstrated that the proposed model has a superior accuracy and mean absolute percentage errors lower than 10%. Moreover, a comparison with the baseline models and empirical methods indicate that the evolutionary metaheuristic intelligence model provides a significant improvement in terms of the root mean square errors (by 63.61–92.30%). This study validates the effectiveness of the proposed model for predicting reinforcement tensile loads and its feasibility for facilitating early designs of GRS structures.

1. Introduction

Since the 1970s, the use of mechanical stabilised earth (MSE) structures has increased in various construction projects worldwide, especially in transportation projects such as bridge abutment, dam, highway, residential, retaining wall, and slope stabilisation projects. The preferential use of MSE walls reflects the multiple advantages of this structure, including the rapid construction, cost effectiveness, aesthetics, reliability, simple construction techniques, superior seismic performance, and a striking ability to withstand large deformations without structural distress, which account for the high desirability of MSE structures [1].

MSE walls can be categorised according to the composition of the reinforcement materials used. Geosynthetics (*e.g.*, geogrids and geotextiles) and metallic reinforcements (*e.g.*, metal strips, grids, and mats) are the most common types of reinforcement used in MSE walls. This study focuses specifically on MSE retaining structures that use geosynthetic reinforcements, *i.e.*, geosynthetic-reinforced soil (GRS) structures. Fig. 1 schematically illustrates a GRS structure.

The three primary agencies responsible for establishing recent MSE structure design specifications in North America are the American Association of State Highway and Transportation Officials (AASHTO) [2], the Federal Highway Administration (FHWA) [3], and the National Concrete Masonry Association (NCMA) [4]. In these design guidelines, the design of the MSE retaining structures is the result of a synergistic approach; the wall system is analysed for its internal, external, global, and seismic stability as well as deformability. The MSE structures must meet certain factors of safety, *FS*, in all failure models.

Several current design guidelines [2–6] use the earth pressure method to measure the stability of GRS structures. Fig. 1 illustrates how the method is used to predict the maximum mobilised reinforcement load, T_{max}, of each reinforcement layer of a GRS structure. The design rationale assumes that the tensile forces developed in the reinforcement are in local equilibrium with the lateral earth pressure generated in the MSE walls.

The FHWA design guidelines recommend using the following equation to predict the T_{max} of each reinforcement layer:

* Corresponding author. Tel.: +886 2 2737 6321; fax: +886 2 2737 6606.
E-mail addresses: jschou@mail.ntust.edu.tw (J.-S. Chou), khy@mail.ntust.edu.tw (K.-H. Yang), m10105810@mail.ntust.edu.tw (J.P. Pampang), d9905806@mail.ntust.edu.tw, paduc@dut.udn.vn (A.-D. Pham).

http://dx.doi.org/10.1016/j.compgeo.2015.01.001
0266-352X/

Nomenclature

AASHTO	American Association of State Highway and Transportation Officials
AI	artificial intelligence
ASTMD	American Society for Testing and Materials – D standards
CART	classification and regression tree
DM	data mining
FA	firefly algorithm
FEM	finite element method
FHWA	federal highway administration
GA	genetic algorithm
GP	genetic programming
GLR	generalised linear regression
GRS	geosynthetic-reinforced soil
LSSVM	least squares support vector machine
LSSVR	least squares support vector regression
MAE	mean absolute error (kN/m)
MAPE	mean absolute percentage error (%)
NCMA	National Concrete Masonry Association
PP	polypropylene
R	linear correlation coefficient
RBF	radial basis function
RMSE	root mean square error (kN/m)
SVM	support vector machine
SVR	support vector regression
SFA-LSSVR	smart firefly algorithm-least squares support vector regression
USCS	unified soil classification system
b	bias term of SVR
C	regularisation parameter
c	backfill cohesion (kN/m^2)
$g_j(x)$	set of nonlinear transformations from the input space in SVR
H	wall height (m)
J	reinforcement tensile stiffness (kN/m)
$L(x, \phi)$	loss function
k_r	lateral earth pressure coefficient inside a GRS structure (dimensionless)
K_a	active earth pressure coefficient (dimensionless)
K_0	at-rest earth pressure coefficient (dimensionless)
k_r/K_a	normalised lateral earth pressure coefficient (dimensionless)
q	uniformly distributed surcharge pressure (kN/m^2)
S_v	vertical spacing of a reinforcement layer (m)
T_{max}	maximum tensile load of each reinforcement layer (kN/m)
w	weight vector
y_k	sigmoid or logistic transfer function
Z_k	parameter of the chaotic map for the initial location
z	depth below the top of structure (m)
z_n	normalised depth (m)
ε_{max}	peak reinforcement tensile strain (%)
Φ	influence factor = $\Phi_g \times \Phi_{local} \times \Phi_{fs} \times \Phi_{fb}$ (dimensionless)
Φ_{fb}	facing batter factor (dimensionless)
Φ_{fs}	facing stiffness factor (dimensionless)
Φ_g	global stiffness factor (dimensionless)
Φ_{local}	local stiffness factor (dimensionless)
Φ_c	soil cohesion factor (dimensionless)
σ	sigma parameter for the RBF kernel
ω	wall face batter from vertical (°)
$\|\omega\|^2$	complexity term of the SVR
γ	unit weight of soil (kN/m^3)
β Lévy	parameter of the Lévy flight for movement
ϕ	soil friction angle (°)
ϕ_{tx}	soil peak friction angle under triaxial compression (°)
ϕ_{ps}	soil peak friction angle under plane strain (°)

$$T_{max} = \left(\frac{k_r}{K_a}\right) K_a(\gamma z + q) S_v \quad (1)$$

where T_{max} is the maximum reinforcement load of each reinforcement layer; k_r/K_a is normalised lateral earth pressure coefficient; K_a is theoretical Rankine (adopted by AASHTO and FHWA) or Coulomb (adopted by NCMA considering both the face inclination and the soil-facing friction) active earth pressure coefficient; γ is backfill unit weight; z is depth below the top of the structure, q is surcharge, S_v is tributary area (equivalent to the vertical spacing of the reinforcement when analyses are conducted per unit length of wall). The k_r/K_a depends on the type of reinforcement; for flexible MSE or GRS walls, the k_r/K_a has a value of 1.0 and remains constant across the depth of the wall.

Therefore, for flexible MSE or GRS walls, the horizontal movement that occurs during construction is sufficient for the soil to reach an active stress state and generate an active earth pressure. The forces on the reinforcement must be calculated to select a reinforcement with an adequate long-term strength (resistance to breakage), to determine the length required to resist pullout within a stable soil zone (resistance to pullout), and to calculate the required connection strength at the wall face (against connection failure). Hence, T_{max} must be calculated to estimate the internal stability of a GRS structure.

The prediction methods for the reinforcement loads within GRS structures currently in use for research and in practice can be categorised into two approaches: the force-equilibrium approach (*i.e.*, the use of the earth pressure method and limit equilibrium (LE) method) and the deformation-based approach (*i.e.*, the use of the K-stiffness method [7–9] and the finite element (FE) method).

Yang et al. [10] reviewed each prediction method and discussed their advantages and disadvantages for predicting reinforcement tensile loads. Yang et al. [10] also examined the accuracy of each method by comparing the predicted T_{max} with the measured T_{max} from a full-scale (3.6 m high) and carefully instrumented GRS wall. Yang et al. [10] further concluded that, out of all of the prediction methods, the earth pressure method based on the Rankine theory produced the largest overestimate of T_{max}. On average, the predicted loads for the GRS wall were three times greater than the measured values for the full-scale instrumented walls. The finite element method had the highest prediction accuracy under working stress conditions because it can capture the deformation (*i.e.*, the strain compatibility) in GRS structures.

In addition to the force-equilibrium and deformation-based approaches, the use of artificial intelligence (AI) is increasing, particularly for modelling the complex behaviour of geotechnical engineering materials and structures. Artificial intelligence has shown superior predictive ability compared to conventional methods [11]. Regarding geotechnical engineering, data mining (DM) approaches are a good alternative for establishing models with good prediction accuracy. According to the literature, applications of data mining in geotechnical engineering problems have been successfully accomplished and produced satisfactory results [11–13]. Based on these reports, this study applied several popular

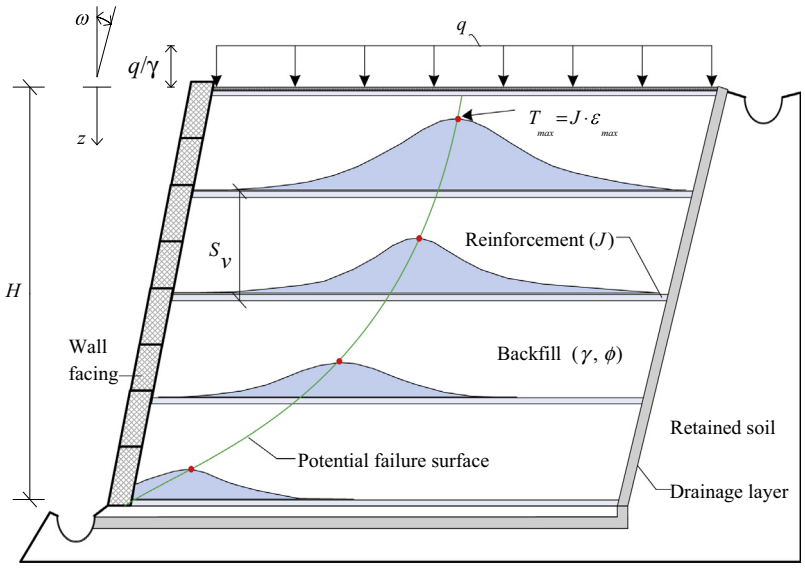


Fig. 1. Design model of a GRS structure.

numerical predictors from regressions-based models, *i.e.*, generalised linear regression (GLR), support vector regression (SVR), classification and regression tree (CART), in IBM SPSS Modeler [14] to demonstrate the excellence of DM.

In advanced data mining applications, evolutionary techniques can enhance prediction accuracy [15–17]. In this research, the proposed evolutionary metaheuristic intelligence technique (*i.e.*, smart firefly algorithm–based least squares SVR) is developed to predict reinforcement loads within GRS structures. The prediction model has three novel features: first, the smart components are employed to minimise the drawbacks of the standard firefly algorithm (FA); second, the improved FA, named the smart firefly algorithm (SFA), is used to optimise the least squares SVR (LSSVR) hyperparameters; finally, the optimised LSSVR is applied to estimate the reinforcement loads. The modelling systems are then validated using a k-fold cross-validation algorithm and a database of 15 wall case studies (94 data points in total) [8]. A synthesised index was developed to compare performance measures between the proposed model and those of conventional models.

The remainder of this paper is organised as follows. First, the related literature is reviewed, including studies of reinforcement-load prediction and a description of well-known predictive techniques that employ computational models suitable for geotechnical studies. The research methodology and experimental settings sections then describe the research methodology, evaluation methods, and experimental settings of the proposed models. The comparison of models and empirical methods section discusses the analytical results and compares the model performances, whilst the conclusions and recommendations section summarises the findings and recommends future studies.

2. Literature review

In the past decade, numerous studies have successfully applied AI within the field of geotechnical engineering. For example, the complex and uncertain behaviour of foundations (deep and shallow) in soils has encouraged several researchers to use AI techniques to discover and generalise structural patterns in geotechnical areas [18]. Recently, Abuel-Naga and Bouzza applied a numerical experiment-AI approach to develop empirical equations for predicting leakage rates [19]. Choi and Lee employed CART to develop an appropriate retaining wall selection systems [20]. Tinoco et al. employed SVM to predict the uniaxial compressive strength (UCS) of jet grouting and obtained a promising result [21]. Pal and Deswal used SVR to model the shear strengths of deep beams [22]. Samui and Karthikeyan predicted liquefaction by using least squares SVM (LSSVM) to find sufficient parameters to achieve the optimal result [23].

Another recent research direction in the field of geotechnical engineering is the use of advanced, hybrid AI models. Liquefaction during earthquakes is a very dangerous ground failure phenomenon that can severely damage civil engineering structures. However, although the mechanism of liquefaction is well known, methods of predicting liquefaction potential are very complex [24]. Therefore, a number of researchers have investigated the potential use of a hybrid optimisation model combined with baseline DM techniques, especially using regression division techniques, including SVR [12,13,21,25], CART [26,27] and GLR [28,29]. The excellent results obtained by these methods confirm that hybrid DM systems provide a valuable option for forecasting analyses.

Studies of the relationship between AI and retaining structures [11] were reviewed before selecting the input variables for the present research, but no studies have discussed the use of geosynthetic materials in retaining structures. Goh and Kulhawy proposed an SVR approach for solving the reliability assessment of the serviceability performance of braced retaining walls [30]. Kung et al. used the same approach with five input parameters to estimate the deflection in a diaphragm wall [31]. An SVR approach was also used to predict lateral earth pressures in rigid retaining walls and to calculate the total lateral thrust and its point of application to six parameters, and reported to be a powerful tool [32].

The literature in this research area indicates that SVR is amongst the most common and effective DM techniques. Despite its proven effectiveness for solving prediction problems, it still has major drawbacks. The accuracy of an SVR model depends on the SVR architecture. During the training process, the user may select a kernel function (*i.e.*, linear, polynomial, radial basis, or sigmoid) to enable the SVR to identify support vectors along the function surface. One of these functions is the Gaussian function, which maps the sample set from the input space into a high-dimensional feature space. The Gaussian function effectively represents the complex nonlinear relationship between the input and output

samples. The literature suggests that the Gaussian function is generally a reasonable first choice [33].

3. Methodology

3.1. Regression-based models

3.1.1. Classification and regression tree

The classification and regression tree (CART) is a decision tree method for constructing a classification or regression tree according to its dependent variable type, which may be categorical or numerical [34]. The same predictor field may be used several times across multiple tree levels. Surrogate splitting optimises the use of data with missing values. A CART has the flexibility to consider misclassification costs as it grows and to specify the prior probability distribution in a classification problem. Additionally, the logic rules of decision tree methods are markedly superior to those of other modelling techniques [35].

Depending on the target field, three impurity measurements can be used to define splits for CART models. For instance, Gini is typically applied in symbolic target fields, whilst the least-squared deviation method automatically selects continuous targets without explaining the selections. In a CART, the Gini index $g(t)$ at a node t is defined by Eq. (2)

$$g(t) = \sum_{j \neq i} p(j|t)p(i|t) \tag{2}$$

where i and j are target field categories

$$p(j|t) = \frac{p(j,t)}{p(t)};\ p(jt) = \frac{\pi(j)N_j(t)}{N_j};\ \text{and}\ p(t) = \sum_j p(j,t) \tag{3}$$

where $\pi(j)$ is the prior probability value for category j; $N_j(t)$ is the number of records in category j of node t; and N_j is the number of records of category j in the root node. Notably, when the Gini index is used to find the improvement after a split during tree growth, only node t records and root node records that have valid values for the split-predictor are used to compute $N_j(t)$ and N_j, respectively.

3.1.2. Generalised linear regression

The generalised linear regression (GLR) model developed by Nelder and Wedderburn [36] can be used to analyse different probability distributions (*i.e.*, normal, binomial, Poisson and gamma) for a dependent variable by using a link function as a mathematical model to specify the relationship between linear predictors and the mean distribution function. Compared to simple regression, the GLR model is more flexible and produces a more realistic relationship. The assumed distribution pattern of data points and the relationship between X and Y are defined in Eq. (4)

$$\eta = g(E(Y)) = X_i\beta_i + O, Y \sim F \tag{4}$$

where η is the linear predictor; O is an offset variable; X_i is an independent variable; β_i is the slope coefficient; and F is the distribution of Y.

The three components of the GLR model are an outcome variable Y with a specific random distribution with an expected value μ and a variance of σ^2 ($E(Y) = \mu$); a link function $g(\cdot)$ that connects the expected value (μ) of Y to transform predicted values of $\eta[\eta = g(\mu)]$; and a linear structural model.

3.1.3. Support vector regression

Support vector regression (SVR) first uses a fixed mapping procedure to map input onto an n-dimensional feature space [37]. Nonlinear kernel functions are then fitted to the high-dimensional feature space to increase the separation amongst the input data compared to the original input space. The linear model in the feature space, $f(x, w)$, can be expressed by Eq. (5).

$$f(x,w) = \sum_{j=1}^{n} w_j g_j(x) + b \tag{5}$$

where $g_j(x)$ is a set of nonlinear transformations from the input space; b is a bias term; and w denotes the weight vector estimated by minimising the regularised risk function that includes the empirical risk and a complexity term $\|\omega\|^2$. Moreover, estimation quality is measured by a loss function, $L(x, \phi)$ [38], defined as

$$L_\varepsilon = [y, f(x,\omega)] = \begin{cases} 0 & \text{if } |y - f(x,\omega)| \leqslant \varepsilon \\ |y - f(x,\omega)| & \text{otherwise} \end{cases} \tag{6}$$

The most interesting feature of the SVR is its use of an ε-insensitive loss function to compute a linear regression function for the new higher dimensional feature space whilst simultaneously decreasing the model complexity by minimising $\|\omega\|^2$. This function is introduced by including the nonnegative slack variables ξ_i and ξ_i^* for the data points outside the tube, where $i = 1, \ldots, n$ is used to identify training samples from the ε-insensitive zone. The SVR can thus be formulated as a minimised version of the following quadratic function:

$$\min_{\omega,b,\xi,\xi^*} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{7}$$

$$\text{subject to} \begin{cases} y_i - f(x_i,\omega) \leqslant \varepsilon + \xi_i^* \\ f(x_i,\omega) - y_i \leqslant \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geqslant 0, i = 1, \ldots, n \end{cases}$$

This optimisation problem can be transformed into a dual problem, which is solved by

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) \text{ subject to } 0 \leqslant \alpha_i^* \leqslant C, 0 \leqslant \alpha_i \leqslant C \tag{8}$$

where n is the number of support vectors. The kernel function is

$$K(x, x_i) = \sum_{i=1}^{n} g_i(x)g_i(x_i) \tag{9}$$

The training process for the SVR requires the selection of kernel functions (*i.e.*, linear, radial basis, polynomial, or sigmoid functions) to identify support vectors along the function surface. Default values of kernel parameters are highly dependent on their type and on the software implemented. Software applications typically treat SVR parameters as user-defined inputs [38]. Fig. 2 shows the general structure of the steps of the regression process.

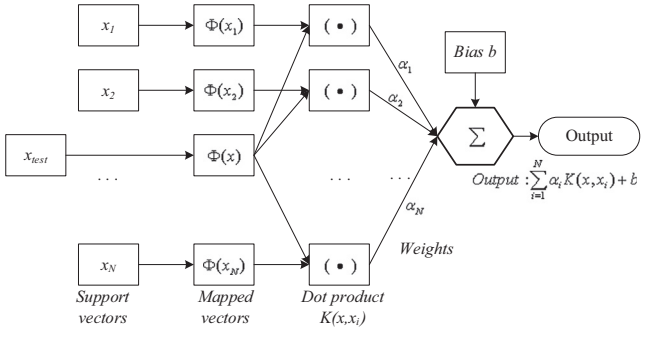


Fig. 2. Architecture of a regression machine constructed by a support vector algorithm.

The proposed regression model performs a least squares support vector regression (LSSVR) to minimise the computational cost [39]. The ε-SVR model requires the user to specify two parameters (C and ε) that are required for the model to perform various tasks. However, the hyperparameter ε can be eliminated when a least squares loss function is used instead of the original ε-insensitive loss function [40,41].

In the LSSVR for function estimation, given a training dataset $\{x_k, y_k\}_{k=1}^{N}$, the optimisation problem is formulated as

$$\min_{\omega,b,e} J(\omega, e) = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}C\sum_{k=1}^{N} e_k^2; \text{ subject to}$$
$$y_k = \langle \omega, \varphi(x_k)\rangle + b + e_k, \quad k = 1, \ldots N \tag{10}$$

where $e_k \in R$ are error variables and $C \geqslant 0$ is a regularisation constant that specifies the constant representing the trade-off between the empirical error and the flatness of the function.

The resulting LSSVR model for function estimation is

$$f(x) = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b \tag{11}$$

where α_k, b are the Lagrange multiplier and the "bias" term, respectively.

Selecting the kernel function is an important step in establishing the model. Recent studies suggest that the radial basis function (RBF) is useful for solving nonlinear problems [42–44]. In the LSSVR, two major hyperparameters must be set appropriately because of their large impact on the prediction accuracy: the regularisation parameter (C), which represents the tradeoff between the model complexity and training error; and the kernel parameter (σ), which controls the kernel width used to fit the training data. These tuning parameters must be set appropriately to obtain a prediction model that performs well [41]. Therefore, an RBF is used as follows to construct the kernel function in the LSSVR, where the kernel function is denoted as $K(x, x_k)$ and σ is the RBF kernel parameter.

$$K(x, x_k) = \exp\left(\frac{-\|x - x_k\|^2}{2\sigma^2}\right) \tag{12}$$

3.1.4. Ensemble regression models

The ensemble approach ranks a set of the above-described models based on their performances and then combines the best-performing models into an ensemble model [45]. The ensemble approach can be expressed mathematically as $g : \mathbb{R}^d \to \mathbb{R}$ with a d-dimensional predictor variable X and a one-dimensional response Y. In each procedure, a specified algorithm yields one estimated function, $\hat{g}(\cdot)$. The estimation by an ensemble-based function, $\hat{g}_{\text{en}}(\cdot)$, is obtained by the following linear combination of individual functions:

$$\hat{g}_{\text{en}}(\cdot) = \sum_{j=1}^{N} c_j \cdot \hat{g}(\cdot) \tag{13}$$

where c_j comprises the linear combination coefficients, which are simply the average values of the weights.

Generally, the predictions obtained by this approach are more accurate than predictions by experts or by conventional models [46]. In predictive models, generalisation can be enhanced by ensemble building or model averaging.

3.2. Evolutionary metaheuristic regression model

This section describes the novel smart firefly algorithm–based least squares support vector regression (SFA-LSSVR) method and its components. The objective was to use evolutionary metaheuristic intelligence to enhance the performance of the standard FA and to optimise the LSSVR parameters in a hybrid system. The novel smart firefly algorithm (SFA) uses three standard terms in a firefly algorithm (FA) (initial population, attractiveness coefficient, and movement). The terms are then adjusted by supplementary elements, including a chaotic map, an adaptive inertia weight, and a Lévy flight. This combination effectively optimises the LSSVR parameters in the proposed predictive technique.

3.2.1. Firefly algorithm

The firefly algorithm developed by Yang, which is one of the most successful swarm intelligence methods, is based on the flashing patterns and behaviour of tropical fireflies [47]. The literature is in agreement in that, for solving a number of optimisation problems, the FA is more efficient than existing algorithms, such as genetic algorithms and particle swarm optimisation [48,49]. For example, this stochastic, nature-inspired, and metaheuristic algorithm can find both the global optima and the local optima simultaneously and effectively.

In a minimisation problem, the brightness can simply be set as an inverse proportion of the value of the objective function. Other forms of brightness can be defined similarly to the fitness function in a genetic algorithm. In its simplest form, the light intensity $I(r)$ varies according to the inverse square law

$$I(r) = \frac{I_s}{r^2} \tag{14}$$

where I_s is the intensity at the source. For a given medium with a fixed light absorption coefficient γ, the light intensity I varies with the distance r, *i.e.*,

$$I = I_0 e^{-\gamma r} \tag{15}$$

where I_0 is the original light intensity. The combined effect of both the inverse square law and absorption can be approximated as the following Gaussian form

$$I(r) = I_0 e^{-\gamma r^2} \tag{16}$$

As the attractiveness of a firefly is proportional to the light intensity seen by adjacent fireflies, the attractiveness β of a firefly is defined as

$$\beta = \beta_0 e^{-\gamma r^2} \tag{17}$$

The movement of the ith firefly, when attracted to another more attractive (brighter) jth firefly, at x_i and x_j, respectively, is

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2}(x_j - x_i) + (rand - 0.5) \tag{18}$$

However, a continuing challenge is designing tuning parameters that improve convergence in the FA. Because these factors are problem-dependent, solutions for some problems require test runs to tune these values. The parameter settings for an FA are optimal if they balance exploitation (local search) and exploration (global search) [50]. To improve the balance between exploration and exploitation, the proposed model modifies the equation used in a conventional FA. The proposed algorithm mainly modifies the quality of the initial population and randomisation parameters. These parameters are then used to develop an elitist firefly algorithm by integrating smart components.

3.2.2. Smart components

3.2.2.1. Chaotic map. Recent theoretical advances and applications of nonlinear dynamics, especially in chaos theory, have attracted attention in numerous fields. One application of chaos theory is in optimisation algorithms [51]. Although different chaotic maps have been proposed as alternatives to pseudorandom sequences,

comparisons show that the Gauss/Mouse map is the best chaotic map for tuning the attractive parameters (β) in the original FA [52,53]. Moreover, the diversity of the initial population is essential for preventing premature phenomena when using an FA in an optimisation algorithm. Logistic mapping, which is the simplest chaotic mapping operator, provides more diversity compared to randomly selected initial populations and decreases the probability of premature occurrence [53]. Here, the above-described chaotic mapping operators are used instead of the random parameters used in the standard FA, as given by Eqs. (19) and (20).

$$\text{Gauss/Mouse map}: Z_{k+1} = \begin{cases} 0 & Z_k = 0 \\ 1/Z_k \bmod(1) & \text{otherwise} \end{cases} \tag{19}$$

$$\text{Logistic map}: Z_{k+1} = a \cdot Z_k(1 - Z_k) \tag{20}$$

where Z_k is the kth chaotic number; k indicates the iteration number; $Z \in (0, 1)$ and $Z_0 \notin \{0.0;0.25;0.50;0.75;1.0\}$; $a = 4$ is used for the experiments.

3.2.2.2. Adaptive inertia weight. The inertia weight has important effects on the convergence of the optimal solution to the optimal value and on the execution time of the simulation. The inertia weight determines the local and global exploration capabilities of a swarm algorithm. Because this study uses the adaptive inertia weight to tune the α parameter, distances are reduced to maintain α within a reasonable range [54].

$$\alpha = \alpha_0 \theta^t \tag{21}$$

where α_0 is the initial attractive coefficient; $0 < \theta < 1$ is the controlling parameter; and t is the number of iterations or time steps. In this implementation, the selected value for θ is 0.9, $t \in [0, T_{\max}]$; where $T_{\max}$ is the maximum number of iterations.

3.2.2.3. Lévy flight. Random walk theory plays a central role in modern metaheuristic algorithms and in stochastic optimisation. A Lévy flight is a random walk in which the step length is a Lévy distribution $L(s) \sim |s|^{-1-\beta}$; where $0 < \beta \leqslant 2$ is an index; s is a power-law distribution; and $L(s)$ is a Lévy distribution with an index β. Lévy flights are defined such that each jump, regardless of size, takes one unit of time [55]. A Lévy distribution is a distribution of a sum of N identically and independently distributed random variables for which the Fourier transformation takes the form of $F_N(k) = \exp(-N|k|^{\beta})$. The variant of such a power-law distribution is infinite for $0 < \beta \leqslant 2$. The following relationship has several special cases:

$$\sigma^2(t) \sim \begin{cases} t^2 & \text{if } 0 < \beta < 1; \\ t^{3-\beta} & \text{if } 1 < \beta < 2; \\ t & \text{if } \beta \geqslant 2 \end{cases} \tag{22}$$

A value of $\beta = 1$ results in a Cauchy distribution in which $\sigma^2(t) \sim t^2/\ln t$ when the target sites are randomly and sparsely distributed. A Lévy walk generates new solutions around the best solution obtained thus far, which increase the speed of the local search. Promising applications of the Lévy walk include optimisation and optimal searching. Eq. (18) is revised as shown below

$$x_i = x_i + \beta e^{-\gamma r_{ij}^2}(x_j - x_i) + \alpha \cdot sign[rand - 0.5] \otimes \text{Lévy} \tag{23}$$

where the second term is the attraction and the third term is the randomisation determined by Lévy flights for which α is the randomisation parameter. The product $\otimes$ indicates entry-wise multiplications [56]. The term $sign[rand - 0.5]$ with $rand \in [0, 1]$ provides a random sign or direction, whilst the random step length is drawn from a Lévy distribution.

3.2.3. Smart firefly algorithm for optimisation of the support vector regression

The regression model in this study uses an SVR with a least squares cost function (LSSVR) embedded in the MATLAB software. This study applied the radial basis function (RBF) kernel in the LSSVR because, in highly non-linear spaces, the RBF usually yields more promising results compared to other kernels, as reported in the literature [42]. For maximum accuracy in solving complex problems, the LSSVR parameters optimised in this study included (1) the regularisation parameter (C), which determines the trade-off cost between minimising the training error and minimising the model complexity, and (2) the sigma parameter (σ) of the RBF kernel function, which defines the non-linear mapping from the input space to a high-dimensional feature space.

The selection of the parameter optimisation mechanism is another critical issue in machine learning [43]. This issue is well recognised by scholars in widely varying disciplines. In practice, the identification of the best set of model parameters is an optimisation problem. Therefore, this study developed a fast and efficient advanced prediction model that combines smart FA and LSSVR (SFA-LSSVR) in a novel hybrid intelligence system.

Table 1
Numerical benchmark functions.

No.	Benchmark function	Hypothesis testing result		
		Mean of optimum	Standard deviation	p-value (%)
1	Bukin function No. 6 $f_1(x,y) = 100\sqrt{\lvert y - 0.01x^2\rvert} + 0.01\lvert x + 10\rvert$ $x = [-15;-5]$; $y = [-3,3]$; Minimum $f_2(-10,1) = 0$	6.94E−03	9.78E−03	47.99
2	Four peak function $f_2(x,y) = -e^{-(x-4)^2-(y-4)^2} + e^{-(x+4)^2-(y-4)^2} - 2[e^{-x^2-(y+4)^2} + 2e^{-x^2-y^2}]$ $x = [-5;5]$; $y = [-5,5]$; Minimum $f_3(0,0) = -4$	−4.00E+00	2.53E−06	37.50
3	Ackley's function $f_3(x,y) = -20e^{-0.2\sqrt{0.5*(x^2+y^2)}} - e^{0.5*(\cos(2\pi x)+\cos(2\pi y))} + 20 + e$ $x = [-5;5]$; $y = [-5,5]$; Minimum $f_4(0,0) = 0$ $x = [-600;600]$; $y = [-600,600]$; Minimum $f_7(100,100) = 0$	5.16E−03	1.15E−01	65.57
4	Rastrigin function $f_4(x,y) = (x^2 - 10\cos(2\pi x) + 10) + (y^2 - 10\cos(2\pi y) + 10)$ $x = [-5.12;5.12]$; $y = [-5.12,5.12]$; Minimum $f_8(0,0) = 0$	4.97E−02	2.17E−01	2.40
5	Rosenbrock function $f_5(x,y) = 100(y - x^2)^2 + (x - 1)^2$ $x = [-50;50]$; $y = [-50,50]$; Minimum $f_9(1,1) = 0$	1.82E−02	9.58E−02	6.03

The optimisation process (OP) is automated by using an SFA to simultaneously optimise all of the LSSVR parameters. The LSSVR mainly addresses learning and curve fitting, whereas the SFA optimises parameters C and σ to minimise prediction errors. The proposed hybrid algorithm was coded in MATLAB R2012a on a PC with a Pentium CORE 2 Quad processor, 2 GB of RAM, and Windows 7 OS. The fitness function of the SFA-LSSVR was as follows:

$$\text{Minimise}: f(m) = \text{RMSE}^{\text{OP}}_{\text{Training-data}} + \text{RMSE}^{\text{OP}}_{\text{Validation-data}} \quad (24)$$

Table 1 shows the benchmark functions [57] conventionally used to evaluate the performance of simulation experiments. The five described functions contain the global optimum and multiple local optima. The table also shows the values used for two dimensions of these functions: the range of the corresponding initial position of the firefly and the goal for each function performed by the algorithms. The analytical results demonstrate satisfactory solutions for these benchmark functions.

Fig. 3 is a flowchart of the SFA-LSSVR model construction, and the LSSVR calls the SFA as a subroutine for optimising its hyperparameters. Historical data were classified as training data, validation data, and test data. The training and validation data are used in the SFA optimisation process. In this process, the training data are used to build a set of models and the validation data are used to evaluate the performance of these models to identify the best model in the set. At the end of optimisation process, the best model with the optimal parameters is determined. Then, the test data set is used

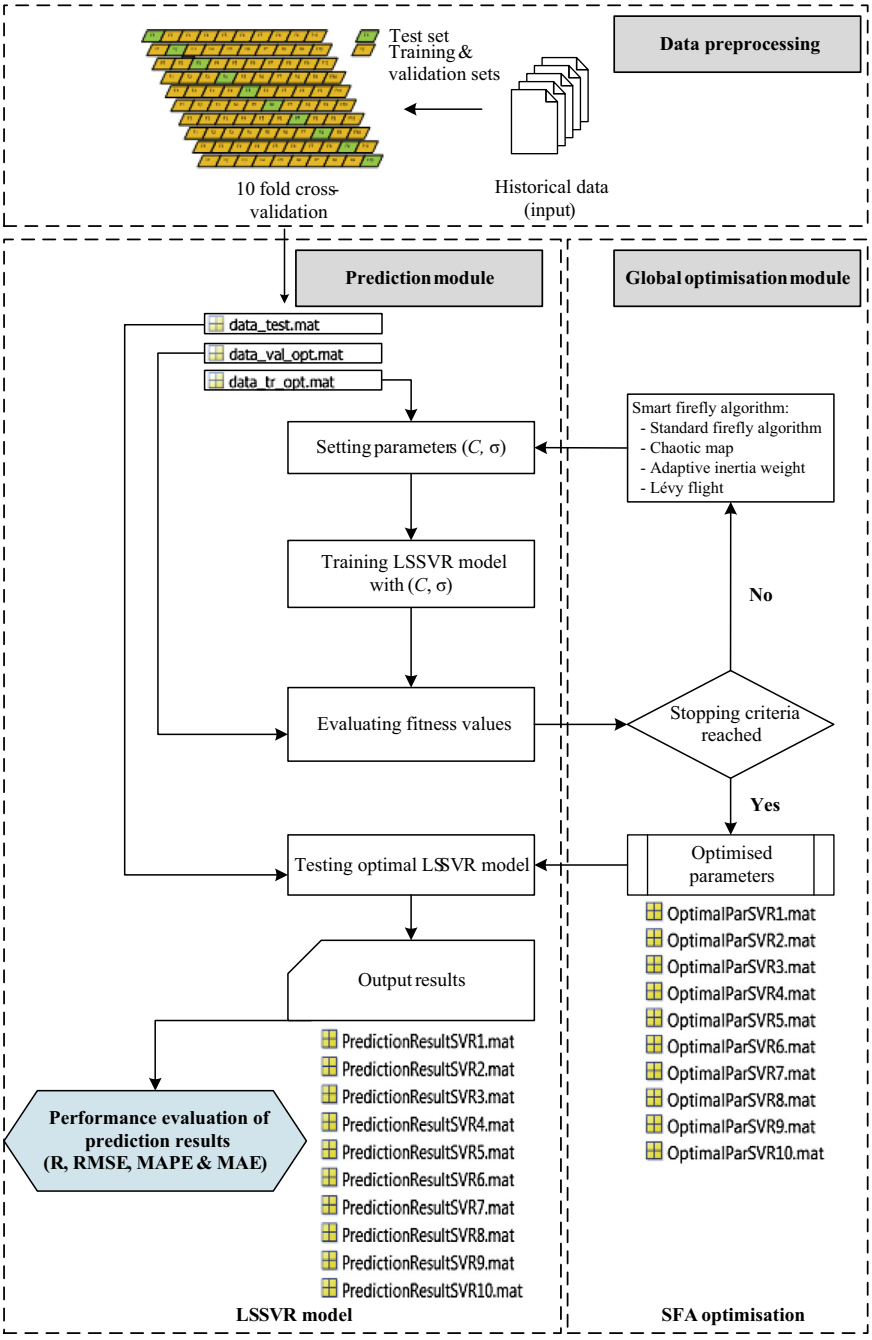


Fig. 3. Framework of the SFA-LSSVR model development.

to evaluate the performance of the optimised LSSVR model. Notably, to mitigate over-fitting, the hyperparameters were repetitively generated by a 10-fold cross validation method with an SFA, which is discussed below.

3.3. Prediction performance and evaluation methods

3.3.1. Cross validation

The k-fold cross-validation algorithm is widely used because it minimises the bias associated with random sampling of training and holdout data samples. Kohavi further confirmed that the 10-fold validation testing is optimal in terms of the computation time and variance [58]. Thus, the models in this work were assessed by a stratified 10-fold cross-validation approach. In this method, a fixed number of data samples are randomly categorised into mutually exclusive ten subsets. The model with the remaining nine data subsets is trained, and the validation subset is used to prove model accuracy. The algorithm accuracy is then expressed as the mean accuracy acquired for the ten validation rounds.

3.3.2. Performance measures

The following performance measures [59] were used to evaluate the prediction accuracy of each of the proposed models and prior works.

- Linear correlation coefficient (R)

$$R = \frac{n\sum y.y' - (\sum y)(\sum y')}{\sqrt{n(\sum y^2) - (\sum y)^2}\sqrt{n(\sum y'^2) - (\sum y')^2}} \tag{25}$$

where y' is the predicted value; y is the actual value; and n is the number of data samples.

- Root mean square error

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y' - y)^2} \tag{26}$$

- Mean absolute error

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y - y'| \tag{27}$$

- Mean absolute percentage error

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y - y'}{y}\right| \tag{28}$$

To provide a comprehensive performance measure, four statistical measures ($1 - R$, RMSE, MAE, and MAPE) were combined into the following synthesis index (SI) [59]:

$$\mathrm{SI} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{P_i - P_{\min,i}}{P_{\max,i} - P_{\min,i}}\right) \tag{29}$$

where m is the number of performance measures; and P_i is the ith performance measure. The SI is a relative performance measure for the single, ensemble and hybrid models. By normalising the prediction performance, the models can be compared with each other using the same baseline. The SI range is 0–1; an SI value close to 0 indicates a highly accurate predictive model.

4. Experimental settings

4.1. Data description and pre-processing

The efficacy of the prediction model was assessed using a database of 15 wall case studies, which contained 94 measurements of maximum reinforcement load [8]. The data were compiled from a portion of the database used to develop the K-stiffness method [7–9]. Three predictive techniques (*i.e.*, CART, GLR and SVR) were used as baseline models when using the experimental datasets for performance comparisons of the AI models. Table 2 lists the datasets obtained from various GRS wall case studies [8].

Table 3 describes the statistical attributes of the GRS structure dataset. The response/target (output) is the maximum reinforcement load ($T_{\max}$), and the predictor variables (input) are the remaining attributes.

4.2. Model construction

Table 4 shows the default settings of the modelling parameters for the experimental comparison of DM techniques to establish a baseline for the comparison of the outcome.

Both the proposed single and ensemble models were constructed by performing the following steps in the IBM SPSS Modeler software.

- Step 1: Input phase: add data to the source node based on the cross-validation algorithm.
- Step 2: Training model: use the numerical predictor node to train data.
- Step 3: Testing model: use the nugget node to validate testing data.
- Step 4: Combine models *via* the ensemble node.
- Step 5: Output phase: assess analytical results through table and analysis nodes.

Fig. 4 shows the modelling stream of these five steps. Analyses were performed using a cross-validation algorithm, *i.e.*, for each experimental model; one fold of the original dataset was used to evaluate the models and to identify the best model for each DM method. The training dataset (remaining folds) was used for model training in each DM method. Single models were then combined into an ensemble model.

The novel SFA-LSSVR algorithm was constructed and run in the MATLAB environment (R2012a version). The first stage of the analysis is optimisation, in which parameters C and σ are optimised for the RBF kernel in the LSSVR model. Because SFA-LSSVR is a complex algorithm, Fig. 5 shows the MATLAB flowchart and partial algorithms to illustrate the analytical process for this model.

The SFA-LSSVR is performed in the following steps:

- Step 1: Normalise the dataset. Subdivide the data into k subsets of training, validation, and test data.
- Step 2: Initialise the search parameters *via* the chaotic map operator (logistic map). For each stage, set the *initial population*, *number of generations*, and *boundaries of optimised parameters*.
- Step 3: Optimisation stage *via* SFA. (a) Generate the 'AlphaNew' function ($\alpha = \alpha_0 \cdot 0.9^t$); (b) use a chaotic Gauss/Mouse map to compute the 'BetaNew' function (β); (c) use the 'EvaluateSFA' function to evaluate the quality of the solution by the fitness function $f(m)$; (d) use the 'SortSFA' function to sort the firefly population by the fitness value and to select the best individual in the population; (e) use 'MoveSFA' to move the firefly toward more attractive individuals *via* Lévy flights in search space; (f) use 'MaxGEN' for a firefly search to generate the maximum number of fitness functions.
- Step 4: LSSVR function. Set the kernel (RBF) and cost function (*least squares*) parameters. Train the model with (C, σ) using the training data. Evaluate the fitness functions in terms of their RMSEs for both the training and validation sets.

Table 2
Sources of datasets reported in the literature [8].

Wall code number	Wall name	Wall geometry				Soil		Reinforcement		Number of data points
		Face batter angle ω[a] (°)	Wall height H (m)	Equivalent surcharge height q/γ[b] (m)	Facing type[c]	Unit weight γ (kN/m^3)	Peak friction angle ϕ_{ps} (°)	Stiffness $J_{2\%}$ (kN/m)	Measured maximum reinforcement load in wall T_{max} (kN/m)	
(GW5)	Tanque Verde Wall	0.0	4.7	0.0	1	19.6	53	340	1.12	3
(GW7)	Oslo, Norway, Section J & N	27	4.8	1.75	2	17	46	350	3.25	13
(GW8)	Algonquin Wall	0.0	6.1	0.0	1	20.4	43	500	3.80	5
(GW9)	Algonquin Wall	2.9	6.1	1.2	3	20.4	43	200	1.26	10
(GW10)	Algonquin Wall	0	5.9	0.0	2	20.4	43	180	3.50	5
(GW11)	RMC Geogrid Panel	0	2.9	0.6	2	17.6	55	105–110	2.79	4
(GW12)	RMC Full-Height Plywood Panel	0	3.0	0.0	1	18	55	550	2.48	8
(GW13)	RMC Incremental Panel	0	3.0	0.0	1	18	55	530	3.68	8
(GW14)	RMC Full-Height Aluminium Panel	0	3.0	0.0	1	18	55	90–95	1.67	8
(GW15)	RMC Incremental Aluminium Panel	0	3.0	0.0	1	18	55	87–93	1.87	8
(GW16)	Rainier Avenue wall	3	12.6	0.0	2	21.1	54	120–1000	6.19	8
(GW18)	Fredericton, New Brunswick Wall	0	6.1	0.0	1	24	45	500	2.50	2
(GW19)	St Remy Wall	0	6.4	0.0	1	16.4	39	7400	12.90	8
(GW20)	Vicenza, Italy Wall Section 1	0	4.0	0.0	2	21.1	57	300	4.43	2
(GW20)	Vicenza, Italy Wall Section 2	0	4.0	0.0	2	21.1	57	90–100	2.5	2

[a] A face batter angle $\omega = 0$ means the wall is in the vertical direction.
[b] A surcharge $q = 0$ means there is no surcharge.
[c] 3 Facing types: (1) panel face; (2) wrapped face and (3) block face.

Table 3
Statistical attributes of the GRS dataset.

Variable	Attribute	Notation	Unit	Min	Max
Input	Face batter angle	ω	(°)	0.00	27
	Wall height	H	(m)	2.85	12.6
	Equivalent surcharge height	q/γ	(m)	0.00	2.7
	Unit weight	γ	(kN/m^3)	16.4	24
	Peak friction angle	ϕ	(°)	31	57
	Facing type	–	–	1[a]	3[a]
	Reinforcement stiffness	J	(kN/m)	110	7400
	Reinforcement location (depth)	z	(m)	0.4	11.5
	Reinforcement vertical spacing	S_v	(m)	0.3	1.6
Output	Maximum reinforcement load	T_{max}	(kN/m)	0.04	12.9

[a] 3 Facing types: (1) concrete panel face; (2) wrapped around face and (3) modular block face.

- Step 5: Stopping condition. End the optimisation process if the generation number (MaxGEN) has achieved the maximum number of fitness functions $f(m)$ with the minimum RMSE values; otherwise, return to step 3.
- Step 6: Stop the looping condition when the parameters are optimised. Provisionally store these optimal sets of tuned parameters for use as a knowledge package for training new data patterns.
- Step 7: Output phase. Calculate the performance measures (*i.e.*, R, RMSE, MAE, MAPE, and SI) to find the prediction accuracy using the test set. Evaluate the post-process results and perform a final visualisation.

Finally, find the optimal set of tuned parameters based on the initial settings for the SFA-LSSVR algorithm, which were defined at the beginning of the analysis. For each parameter, Table 5 lists the initial values used to improve the model accuracy.

5. Comparison of models and conventional methods

Based on the experimental results, this section compares the performance of three single models, four ensemble models, and a hybrid technique. The performance measures reported in the literature for conventional models were used to evaluate the performance of the selected models.

Table 4
Default parameter settings of the baseline models.

Model	Parameter	Value
CART	Levels below root	5
	Mode	Simple
	Maximum surrogates	5
	Minimum change in impurity	0.0001
	Impurity measure for categorical targets	Gini
	Minimum records in parent branch (%)	2
	Minimum records in child branch (%)	1
GLR	Singularity tolerance	1.0E−4
	Probability entry	0.05
	Probability removal	0.1
	F value entry	3.84
	F value removal	2.71
SVR	Stopping criteria	1.0E−3
	Regularisation parameter (C)	10
	Regression precision (ε)	0.1
	Kernel type	RBF
	RBF sigma (σ)	0.1

5.1. Analytical results

Table 6 presents the performance measures for the various DM models. Eqs. (25)–(29) were used to evaluate the accuracy of the predictive DM techniques. This table summarises the cross-fold modelling performance for the dataset during the validation period. The SVR had the best results out of all of the single models. However, the overall performance (SI) for the ensemble regression models was no better than that for the SVR model.

Notably, the accuracy of the tensile load prediction is significantly improved by the proposed SFA-LSSVR model. The validation outcome (Table 6) of this hybrid model exhibited superior performance. The lowest MAPE was 9.55%, and the lowest RMSE was 0.27 kN/m; both values were superior to those of the other models. Fig. 6 presents a detailed comparison of the performance measures, demonstrating the superior performance of the SFA-LSSVR model. This hybrid model is characterised by an excellent performance in terms of the correlation between the predicted output and the

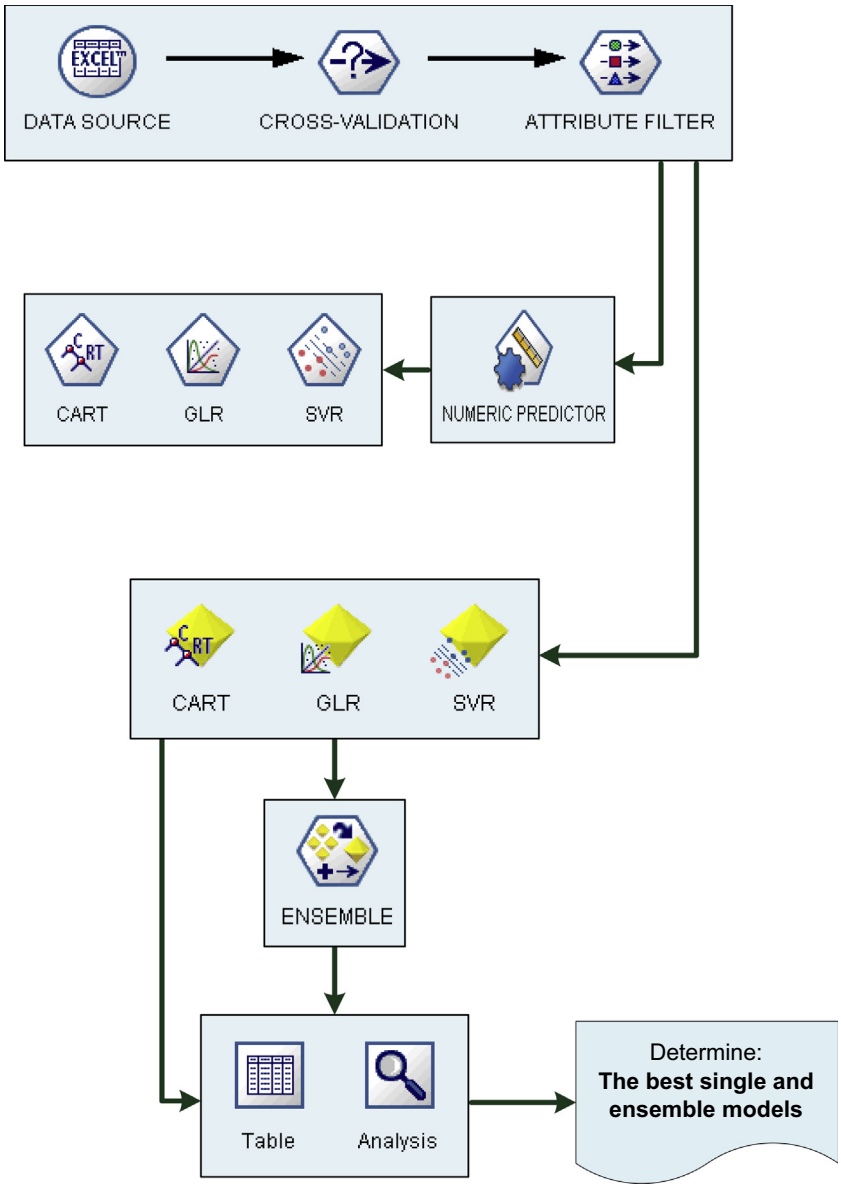


Fig. 4. Flowchart of the modelling stream for the baseline and ensemble models.

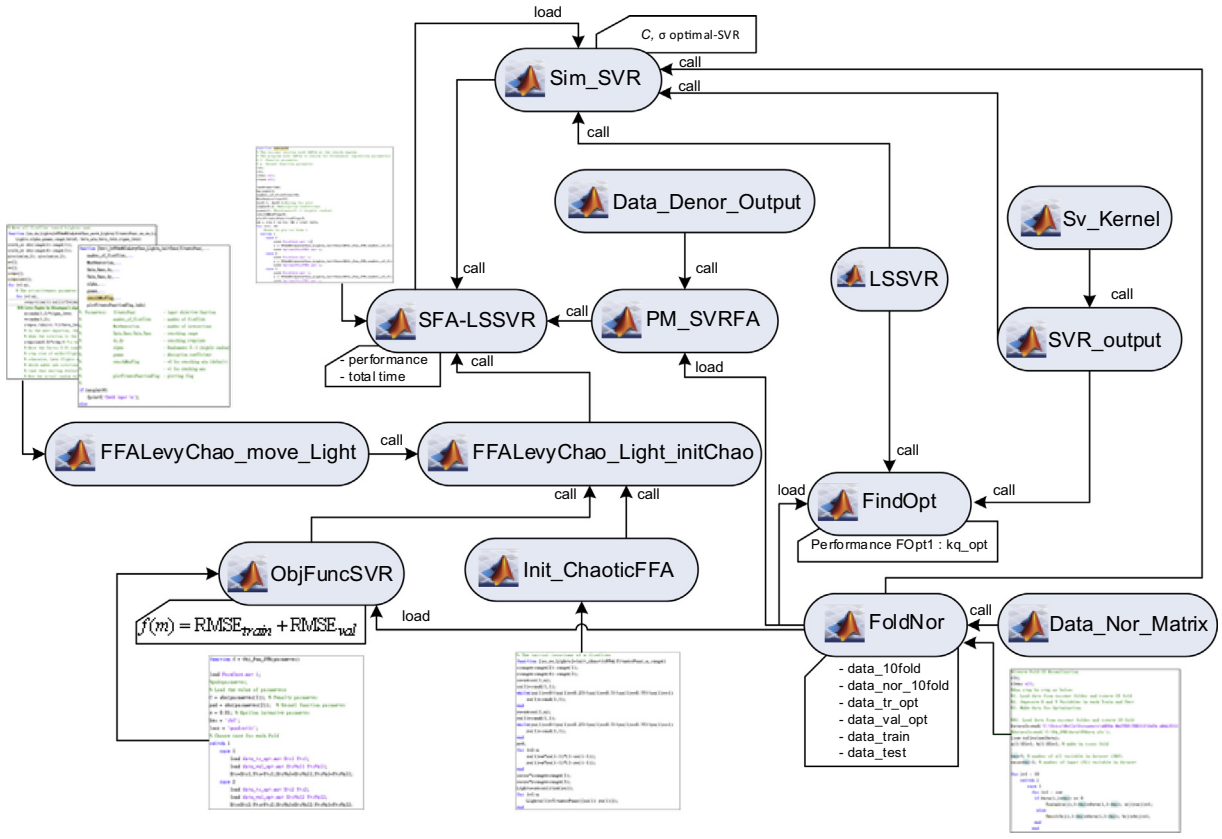


Fig. 5. The SFA-LSSVR scheme and partial algorithms.

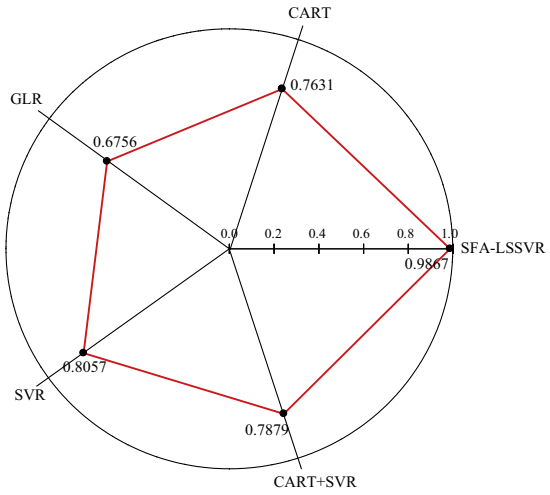
Table 5
Parameter settings for the SFA-LSSVR algorithm.

Algorithm/model	Parameter	Initial setting	Detail
SFA	Alpha	0.2	Randomness 0–1
	Gamma	1	Absorption coefficient
	Search max flag	0	Search maximum flag
	Plot fitness function flag	0	Plotting flag
	Number of fireflies	50	Number of fireflies
	Maximum generation	25	Number of interactions
	dx, *dy*	0.1	Searching step sizes
	$X_{\min}$, $X_{\max}$, $Y_{\min}$, $Y_{\max}$	–	Searching range
	Chaotic mapping operators	*rand*(0, 1)	Initial location
	Initial attractive coefficient	0.9	Control exploration
	β Lévy distribution	1.5	Initial coefficient
LSSVR (RBF)	Regularisation parameter (C)	–	Trade-off training error and model flatness
	Sigma parameter (σ)	–	Defines the non-linear mapping

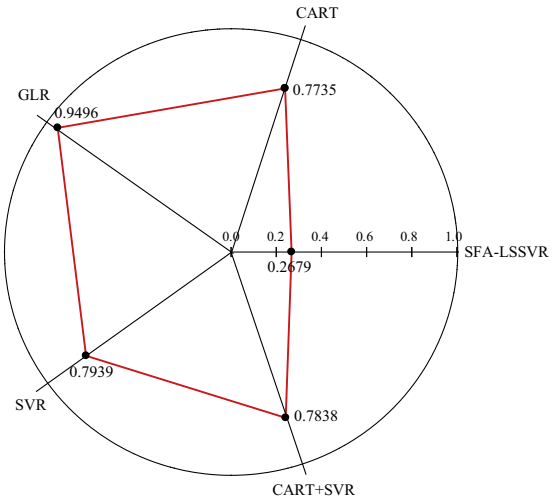
Table 6
Summary of the cross-fold modelling performance for the baseline, ensemble and hybrid models.

Category	Predictive model	Test dataset				
		R	RMSE (kN/m)	MAE (kN/m)	MAPE (%)	SI
Single	CART	0.76	**0.77**	0.67	41.71	0.81
	GLR	0.68	0.95	0.77	43.75	1.00
	SVR	**0.81**	0.79	**0.64**	**32.36**	**0.70 (2)**
Ensemble	CART + GLR	0.74	0.96	0.76	39.53	0.91
	CART + SVR	**0.79**	**0.78**	**0.64**	**34.13**	**0.72 (3)**
	GLR + SVR	0.73	0.95	0.74	40.88	0.92
	CART + GLR + SVR	0.77	0.91	0.73	40.10	0.86
Hybrid	SFA-LSSVR	**0.99**	**0.27**	**0.16**	**9.55**	**0.00 (1)**

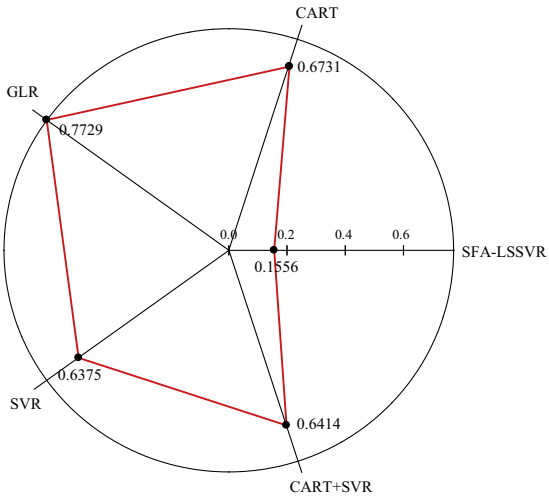
Note: The best performance measure in each category is given in bold and underlined; the overall performance ranking is indicated by an asterisk (·).



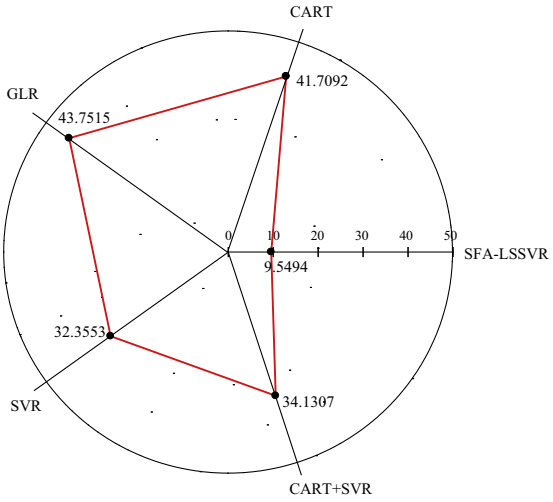
Average Correlation Coefficient (R)



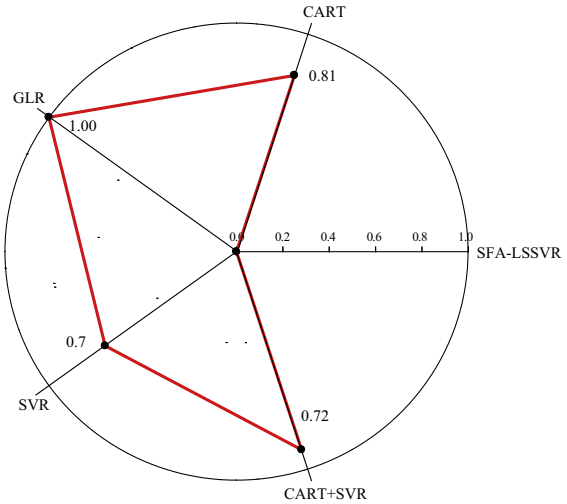
Average Root Mean Square Error (RMSE)



Average Mean Absolute Error (MAE)



Average Mean Absolute Percentage Error (MAPE)



Synthesis Index

Fig. 6. Modelling performance comparison.

Table 7
Optimal values of C and σ for ten folds.

SFA-LSSVR	C	σ
Fold 1	2.04E+10	9.82E+03
Fold 2	2.74E+11	1.00E+04
Fold 3	3.65E+10	9.80E+03
Fold 4	1.38E+10	9.99E+03
Fold 5	4.28E+10	9.69E+03
Fold 6	2.49E+12	8.84E+03
Fold 7	3.21E+13	3.51E+01
Fold 8	6.44E+10	1.00E+04
Fold 9	8.27E+10	1.00E+04
Fold 10	8.92E+09	4.96E+03

real values and the overall accuracy. These results confirm that the hybrid model is superior to the single and ensemble models.

It is further concluded that, without the inclusion of the fine-tuning parameters in the RBF kernel function, the SVR performs poorly, with an MAPE of 32.36%. That is, proper tuning in the parameter optimisation stage is essential when applying the SVR. For example, in the SFA-LSSVR model, the optimisation for the parameters C and σ can substantially improve the predictive capability of the LSSVR. After the defined maximum number of iterations (stopping criteria) is achieved, the firefly with the highest light intensity is defined as the best solution.

Table 7 provides the optimal parameters of C and σ for ten folds of this dataset. For predicting reinforcement loads in GRS structures, the ranges of C and σ obtained by the evolutionary model were 8.92E+9–3.21E+13 and 3.51E+1–1.00E+4, respectively, from the cross-fold validation. Therefore, for users who prefer not to employ an optimisation algorithm, the possible default setting for the LSSVR can be applied using the optimal range of values recommended in Table 7.

5.2. Discussion

After the validation and testing process, the evolutionary metaheuristic intelligence model developed to predict T_{max} was examined through a comparison of the T_{max} measured from a carefully instrumented full-scale GRS wall conducted by Bathurst et al. [60]. These new sample data for T_{max}, which were not included in the previous compiled database, were used to assess the accuracy of the proposed model in comparison with several conventional prediction methods currently in use for research and in practice to predict the reinforcement tensile loads within GRS structures. The GRS wall was 3.6 m high and constructed using six reinforcement layers at a spacing of $S_v = 0.6$ m and with a wrapped around facing slope of $\omega = 8°$. The backfill, named RMC (Royal Military College) sand, was a clean, uniform graded, beach sand classified as poor sand according to the USCS. The backfill soil had a unit weight of $\gamma = 16.7$ kN/m^3, a soil peak friction angle of $\phi_{tx} = 39°$ (according to triaxial compression tests), and a $\phi_{ps} = 42°$ (according to plane strain tests). The reinforcement was a polypropylene (PP) geogrid with a total length of 2.52 m, measured from the front wall face. The ultimate tensile strength was $T_{ult} = 13$ kN/m according to a wide-width strip tensile test (ASTM-D4595). Under these wall test conditions, the isochronous load–strain responses of the reinforcement at 1000 h was $T_{ult} = 7.7$ kN/m and the stiffness at a 2% strain was $J_{2\%} = 100$ kN/m.

After the construction of the wall was completed, uniform surcharges were applied in increments of 10 kPa to a final loading of 80 kPa. The wall was extensively instrumented to measure its performance at the end of the construction and during the staged, uniform surcharging; for instance, strain gauges and extensometers were attached to the reinforcements to measure the reinforcement strain in each layer. The measured maximum reinforcement strain of each reinforcement layer was then multiplied by the reinforcement secant stiffness ($T_{max} = J(\varepsilon) \cdot \varepsilon$), which was determined from the isochronous load–strain response at the same strain level, to estimate the reinforcement load. A total of 3 data points of T_{max} obtained from the third reinforcement layer under various surcharge levels (*i.e.*, after construction, traffic loading, and large loading) was selected for comparison. Notably, the third reinforcement layer had the highest load, *i.e.*, the maximum T_{max} of all the reinforcement layers. The accurate prediction of the maximum of the T_{max} value is required because this value is conventionally used to determine the reinforcement tensile strength against reinforcement breakage for the internal design of the GRS wall.

Table 8 shows the results for the performance measures for the new sample data and summarises the improvements achieved by the proposed model across the reinforcement loads datasets. Fig. 7 presents a performance comparison of the SFA-LSSVR model and prior empirical methods used for predicting the maximum reinforcement tensile load under three selected loading conditions, including no surcharge ($q = 0$ kPa), traffic loading ($q = 10$ kPa), and large-loading conditions ($q = 50$ kPa). The error bars in Fig. 7 represent ±1 standard deviation of the estimated T_{max} to account for the uncertainties of the strain measurements and the reinforcement stiffness interpretation. All of the T_{max} values for the current prediction methods were reported by Yang et al. [10]. Notably, there were no FE results reported by Yang et al. [10] for $q > 40$ kPa because of numerical instability (*i.e.*, the convergence problem) in the computation. Therefore, FE data are not available for large-loading conditions ($q = 50$ kPa).

The results in both Table 8 and Fig. 7 suggest that the SFA-LSSVR model is characterised by the best prediction performance compared to the individual models used in this study and the conventional methods reported in the literature. The SFA-LSSVR model is also capable of predicting T_{max} accurately under large-loading conditions, whilst the FE method failed to predict the T_{max} under such loading conditions. This comparison suggests the SFA-LSSVR is superior to the FE method for large-loading conditions. Regarding the comparison of the prediction performance of reinforcement loads presented in Table 8, the magnitude of the improvement attributed to the use of the SFA-LSSVR (in terms of the error rates, *i.e.*, RMSE) ranged from 63.61% to 92.30% and; thus, the SFA-LSSVR

Table 8
Performance measures and improvements achieved by the proposed model.

Best AI models and empirical methods	Performance measure			Improvement by the SFA-LSSVR (%)		
	RMSE (kN/m)	MAE (kN/m)	MAPE (%)	RMSE	MAE	MAPE
Rankine [10]	5.27	5.00	157.81	92.30	92.18	91.45
Coulomb [10]	2.79	2.64	82.98	85.48	85.17	83.74
Limit equilibrium [10]	1.12	0.81	21.44	63.74	51.62	37.08
K-stiffness [10]	1.11	1.05	32.91	63.61	62.72	59.02
SVR	2.47	0.73	18.25	83.58	46.19	26.07
CART + SVR	1.74	1.51	44.49	76.65	74.07	69.68
SFA-LSSVR	0.41	0.39	13.49	–	–	–

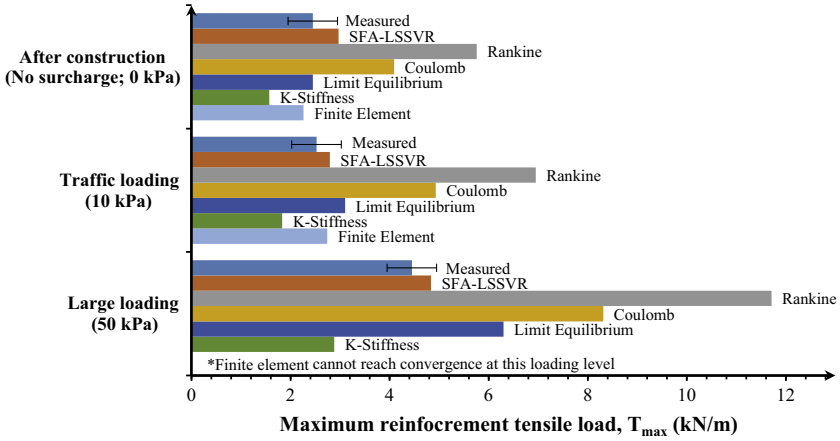


Fig. 7. SFA-LSSVR model and prior methods for predicting the T_{max} under three selected loading conditions.

performs better than reported for previously developed models and the other models adopted in this study. The superior values for the SFA-LSSVR confirm the efficacy and superiority of this evolutionary metaheuristic regression model.

6. Conclusions and recommendations

This study evaluated the use of single prediction models (CART, GLR, and SVR), their ensembles (CART + GLR, CART + SVR, GLR + SVR, and CART + GLR + SVR), and an evolutionary metaheuristic regression model (SFA-LSSVR) to analyse complex data for reinforcement loads in GRS structures. The evolutionary methodology developed in this study achieved a robust prediction model for reinforcement loads with a high prediction accuracy. The evaluated models were examined using a reliable database reported in the literature.

Experiments further demonstrated that the SFA optimiser can concurrently search and identify suitable parameters for the LSSVR. The purpose of the SFA is to optimise the learning process of the LSSVR to establish the regression function. This evolutionary model is effectively operated by using smart components (*i.e.*, a chaotic map, an adaptive inertia weight and a Lévy flight) to tune the standard FA parameters. The integration of the SFA and LSSVR enhances the efficiency of the machine learner for solving complex problems.

The accuracy of the proposed method was evaluated by comparing the measurements and prediction results to those for geotechnical empirical methods. The analytical results demonstrate that the SFA-LSSVR is superior and outperforms the other models (Table 6) in terms of the values of R (0.99), RMSE (0.27 kN/m), MAE (0.16 kN/m), and MAPE (9.55%). Notably, when including additional collected data samples, the SFA-LSSVR model was characterised by RMSE = 0.41 kN/m, MAE = 0.39 kN/m, and MAPE = 13.49% (Table 8). The RMSE, MAE, and MAPE of the proposed model were significantly better than those of the previous models (with maximum improvements of 92.30%, 92.18%, and 91.45%, respectively).

Moreover, compared with the other regression models, the proposed model provided a superior generalisation by applying a cross-validation. The proposed hybrid approach can capture and extract knowledge from historical data whenever the new data are provided. Because of its high accuracy and robust output, decision-makers can use the approach to forecast undesirable performance deviations or early warning of potential problems. The proposed hybrid metaheuristic model can be used as an alternative tool for obtaining fast and accurate predictions of the reinforcement loads in GRS structures.

The applicability of the developed model is limited to GRS structures with features within the ranges of the parameters for the database of the case histories compiled in this study (Tables 2 and 3). For example, the developed model can only be applied for GRS structures with a relatively simple geometry and backfilled with granular soil, but the model is not valid for GRS structures that having multi-tier configurations or are backfilled with fine-grained cohesive soil.

Although an excellent prediction accuracy was achieved in this study, several limitations can be considered as potential opportunities for future improvements. First, the fine-tunings of the parameters in each smart component of SFA were sub-optimal. A proper search algorithm is required to determine an appropriate range of the initial settings for the GRS data. Second, the computation time of the SFA was long and should be reduced by improving the efficiency of the method. Third, the SFA was validated *via* five benchmark functions in this study. Further research can calibrate a greater number of benchmark functions (over 50 functions) by expanding their dimensions to generalise the global optimisation capability of the SFA

Regarding the implementation and use of the SFA-LSSVR model practice, future studies can focus on the development of an evolutionary metaheuristic expert system with a Windows or browser interface to increase the ease of use by practitioners and to facilitate the application of the developed model to the design of GRS structures.

References

[1] Bilgin Ö, Mansour E. Effect of reinforcement type on the design reinforcement length of mechanically stabilized earth walls. Eng Struct 2014;59:663–73.
[2] AASHTO. LRFD bridge design specifications. Washington, DC: American Association of State Highway and Transportation Officials; 2007.
[3] Elias V, Christopher BR, Berg RR. Mechanically stabilized earth walls and reinforced soil slopes: design and construction guidelines. Washington, D.C., USA: National Highway Institute, Federal Highway Administration; 2001.
[4] NCMA. Design manual for segmental retaining walls. Herndon, Virginia, USA: National Concrete Masonry Association; 2010.
[5] AASHTO. Standard specifications for highway bridges. Washington, D.C., USA: American Association of State Highway and Transportation Officials; 2002.
[6] Berg RR, Christopher BR, Samtani N. Design and construction of mechanically stabilized earth walls and reinforced soil slopes. Washington, D.C., USA: National Highway Institute, Federal Highway Administration; 2009.
[7] Allen TM, Bathurst RJ, Holtz RD, Walters D, Lee WF. A new working stress method for prediction of reinforcement loads in geosynthetic walls. Canadian Geotech J 2003;40(5):976–94.
[8] Bathurst RJ, Miyata Y, Nernheim A, Allen AM. Refinement of K-stiffness Method for geosynthetic-reinforced soil walls. Geosynthetics Int. 2008:269–95
[9] Bathurst RJ, Allen TM, Walters DL. Reinforcement loads in geosynthetic walls and the case for a new working stress design method. Geotextiles Geomembr 2005;23(4):287–322.

[10] Yang KH, Utomo P, Liu TL. Evaluation of force-equilibrium and deformation-based design approaches for predicting reinforcement loads within geosynthetic-reinforced soil structures. J GeoEng 2013;8(2):41–54.
[11] Shahin MA. Artificial intelligence in geotechnical engineering: applications, modeling aspects, and future directions. Metaheuristics in water, geotechnical and transport engineering. Oxford: Elsevier; 2013. p. 169–204.
[12] Maalouf M, Khoury N, Laguros JG, Kumin H. Support vector regression to predict the performance of stabilized aggregate bases subject to wet–dry cycles. Int J Numer Anal Meth Geomech 2012;36(6):675–96.
[13] Kordjazi A, Pooya Nejad F, Jaksa MB. Prediction of ultimate axial load-carrying capacity of piles using a support vector machine based on CPT data. Comput Geotech 2014;55:91–102.
[14] IBM. SPSS Modeler. New York, USA: IBM Corporation; 2010.
[15] Chou JS, Cheng MY, Wu YW, Pham AD. Optimizing parameters of support vector machine using fast messy genetic algorithm for dispute classification. Expert Syst Appl 2014;41(8):3955–64.
[16] Zhang WG, Goh ATC. Multivariate adaptive regression splines for analysis of geotechnical engineering systems. Comput Geotech 2013;48:82–95.
[17] Cabalar AF, Cevik A, Gokceoglu C. Some applications of Adaptive Neuro-Fuzzy Inference System (ANFIS) in geotechnical engineering. Comput Geotech 2012;40:14–33.
[18] Javadi A, Rezania M. Applications of artificial intelligence and data mining techniques in soil modeling. Geomech Eng 2009;1(1):53–74.
[19] Abuel-Naga HM, Bouazza A. Numerical experiment-artificial intelligence approach to develop empirical equations for predicting leakage rates through GM/GCL composite liners. Geotextiles Geomembr 2014;42(3):236–45.
[20] Choi M, Lee G. Decision tree for selecting retaining wall systems based on logistic regression analysis. Automation Constr 2010;19(7):917–28.
[21] Tinoco J, Gomes Correia A, Cortez P. Support vector machines applied to uniaxial compressive strength prediction of jet grouting columns. Comput Geotech 2014;55:132–40.
[22] Pal M, Deswal S. Support vector regression based shear strength modelling of deep beams. Comput Struct 2011;89(13–14):1430–9.
[23] Samui P, Karthikeyan J. Determination of liquefaction susceptibility of soil: a least square support vector machine approach. Int J Numer Anal Meth Geomech 2013;37(9):1154–61.
[24] Samui P. Liquefaction prediction using support vector machine model based on cone penetration data. Front Struct Civil Eng 2013;7(1):72–82.
[25] Samui P, Lansivaara T, Bhatt M. Least square support vector machine applied to slope reliability analysis. Geotech Geological Eng 2013;31(4):1329–34.
[26] Samadi M, Jabbari E, Azamathulla HM. Assessment of M5 model tree and classification and regression trees for prediction of scour depth below free overfall spillways. Neural Comput Appl 2014;24(2):357–66.
[27] Alavi AH, Gandomi AH. A robust data mining approach for formulation of geotechnical engineering systems. Eng Comput 2011;28(3):242–74.
[28] Zhang J, Zhang LM, Huang HW. Evaluation of generalized linear models for soil liquefaction probability prediction. Environ Earth Sci 2013;68(7):1925–33.
[29] Chou JS, Lin C. Predicting disputes in public-private partnership projects: classification and ensemble models. J Comput Civil Eng 2013;27(1):51–60.
[30] Goh ATC, Kulhawy FH. Reliability assessment of serviceability performance of braced retaining walls using a neural network approach. Int J Numer Anal Meth Geomech 2005;29(6):627–42.
[31] Kung GTC, Hsiao ECL, Schuster M, Juang CH. A neural network approach to estimating deflection of diaphragm walls caused by excavation in clays. Comput Geotech 2007;34(5):385–96.
[32] Yildiz E, Ozyazicioglu MH, Ozkan MY. Lateral pressures on rigid retaining walls: a neural network approach. Gazi Univ J Sci 2010;23(2):201–10.
[33] Hsu SC, Chang CC, Lin CJ. A practical guide to support vector classification. Taipei, Taiwan: National Taiwan University; 2003.
[34] Breiman L, Olshen R, Stone C. Classification and regression trees. New York, USA: Chapman & Hall/CRC; 1984.
[35] SPSS. Clementine 12.0 Algorithm Guide Chicago, USA: Integral Solutions Limited; 2007.
[36] Nelder J, Wedderburn R. Generalized linear models. J Roy Stat Soc Series A (General) 1972;135(3):370–84.
[37] Lin CJ, Weng RC. Simple probabilistic predictions for support vector regression. Taipei, Taiwan: National Taiwan University; 2004.
[38] Smola A, Schölkopf B. A tutorial on support vector regression. Statistics Comput 2004;14(3):199–222.
[39] Wang H, Hu D. Comparison of SVM and LS-SVM for regression. In: International conference on neural networks and brain, 2005 ICNN&B '05; 2005. p. 279–83.
[40] Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. Neural Process Lett 1999;9(3):293–300.
[41] Suykens JAK, Gestel TV, Brabanter JD, Moor BD, Vandewalle J. Least squares support vector machines: world scientific; 2002.
[42] Kazem A, Sharifi E, Hussain FK, Saberi M, Hussain OK. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. Appl Soft Comput 2013;13(2):947–58.
[43] Fard AK, Samet H, Marzbani F. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. Expert Syst Appl 2014;41(13):6047–56.
[44] Sun LG, de Visser CC, Chu QP, Mulder JA. A novel online adaptive kernel method with kernel centers determined by a support vector regression approach. Neurocomputing 2014;124:111–9.
[45] Zhou ZH. Ensemble methods : foundations and algorithms. Boca Raton: CRC Press/Taylor & Francis; 2012.
[46] Fan C, Xiao F, Wang S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. Appl Energy 2014;127:1–10.
[47] Yang XS. Firefly algorithms for multimodal optimization. Stochastic algorithms: foundations and applications. Berlin, Germany: Springer; 2009. p. 169–78.
[48] Coelho LdS, Mariani VC. Improved firefly algorithm approach applied to chiller loading for energy conservation. Energy Build 2013;59:273–8.
[49] Khajehzadeh M, Taha M, Eslami M. A new hybrid firefly algorithm for foundation optimization. Natl Acad Sci Lett 2013;36(3):279–88.
[50] Yang XS. Firefly algorithm, Lévy flights and global optimization. In: Bramer M, Ellis R, Petridis M, editors. Research and development in intelligent systems XXVI. London: Springer; 2010. p. 209–18.
[51] Gandomi AH, Yang XS, Talatahari S, Alavi AH. Firefly algorithm with chaos. Commun Nonlinear Sci Numer Simul 2013;18(1):89–98.
[52] He D, He C, Jiang LG, Zhu HW, Hu GR. Chaotic characteristics of a one-dimensional iterative map with infinite collapses. IEEE Trans Circuits Syst I: Fund Theory Appl 2001;48(7):900–6.
[53] Hong WC, Dong Y, Chen LY, Wei SY. SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. Appl Soft Comput 2011;11(2):1881–90.
[54] Liu C, Tian Y, Zhang Q, Yuan J, Xue B. Adaptive firefly optimization algorithm based on stochastic inertia weight. In: Computational intelligence and design (ISCID), 2013 sixth international symposium on: IEEE; 2013. p. 334–7.
[55] Kamaruzaman AF, Zain AM, Yusuf SM, Udin A. Lévy flight algorithm for optimization problems-a literature review. Appl Mech Mater 2013;421:496–501.
[56] Yang XS, Deb S. Cuckoo search via levy flights. In: Nature & biologically inspired computing, 2009 NaBIC 2009 World Congress; 2009. p. 210–4.
[57] Surjanovic S, Bingham D. Virtual library of simulation experiments: test functions and datasets; 2013.
[58] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th international joint conference on Artificial intelligence, vol. 2. Montreal, Quebec, Canada: Morgan Kaufmann Publishers Inc.; 1995. p. 1137–43.
[59] Chou JS, Pham AD. Enhanced artificial intelligence for ensemble approach to predicting high performance concrete compressive strength. Constr Build Mater 2013;49:554–63.
[60] Bathurst RJ, Vlachopoulos N, Walters DL, Burgess PG, Allen TM. The influence of facing stiffness on the performance of two geosynthetic reinforced soil retaining walls. Canadian Geotech J 2006;43(12):1225–37.

Jui-Sheng (Rayson) Chou specialises in project analytics and engineering management. Chou received his BS and MS from the National Taiwan University and his Ph.D. in Construction Engineering and Project Management from The University of Texas at Austin. Currently, he is a Full Professor in the Department of Civil and Construction Engineering at National Taiwan University of Science and Technology (Taiwan Tech). He holds registered professional engineer licenses and serves on several professional committees. He has provided consulting services to a number of private and public engineering sectors. He is the author or co-author of hundreds of journal articles, conference papers, and technical reports related to civil engineering and management (CEM) and is a member of several international journal editorial boards. His main teaching and research interests are CEM issues related to data mining, knowledge discovery in databases, engineering data analytics, project behaviour management, decision, risk and reliability, simulation, cost management, and hazard mitigation in spatial planning practices.

Kuo-Hsin Yang is currently an Associate Professor in the Geotechnical Engineering Program at the National Taiwan University of Science and Technology (Taiwan Tech). He completed his Ph.D. degree at The University of Texas in Austin in 2009. Dr. Yang has numerous years of experience in research and practice in geotechnical and geosynthetics engineering and contributes considerably to this field in education and service.

Jusieandra Pribadi Pampang is a graduate research assistant at National Taiwan University of Science and Technology (Taiwan Tech) under the mentorship of Professor Jui-Sheng Chou. Mr. Jusieandra is earning his MS degree in the construction management program of Taiwan Tech.

Anh-Duc Pham earned his Ph.D. from the Department of Civil and Construction Engineering at National Taiwan University of Science and Technology (Taiwan Tech) under the mentorship of Professor Jui-Sheng Chou. Dr. Pham is currently the Dean of the Faculty of Project Management at The University of Danang, University of Science and Technology, Vietnam. His current research focuses on data mining applications in project management.