Soil Dynamics and Earthquake Engineering 132 (2020) 106097

Contents lists available at ScienceDirect

Soil Dynamics and Earthquake Engineering

journal homepage: http://www.elsevier.com/locate/soildyn

Multivariate adaptive regression splines model for prediction of the liquefaction-induced settlement of shallow foundations

Gang Zheng [a,b,c], Wenbin Zhang [a,b], Haizuo Zhou [a,b,c,*], Pengbo Yang [a,b]

[a] School of Civil Engineering, Tianjin University, Tianjin, 300072, China
[b] Key Laboratory of Coast Civil Structure Safety, Tianjin University, Ministry of Education, Tianjin, 300072, China
[c] State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin, 300072, China

ARTICLE INFO

Keywords:
Multivariate adaptive regression splines
Shallow foundation
Settlement
Liquefaction

ABSTRACT

Buildings with shallow foundation are vulnerable to liquefaction-induced settlement during an earthquake. Accurate settlement evaluation is an important step in earthquake damage mitigation. This paper presents a simplified approach to estimate the liquefaction-induced settlement of buildings with shallow foundations. The multivariate adaptive regression splines (MARS) algorithm is adopted. The validated finite difference method is used to produce artificial data that consider various properties of the soils, structures and ground motions. A relative importance analysis is conducted to quantify the effect of each input parameter and their coupled interactions on the liquefaction-induced settlement. The accuracy of the established model is demonstrated through centrifuge test results available in the literature.

1. Introduction

Shallow-founded buildings located on liquefiable deposits are vulnerable to liquefaction induced settlement during earthquakes, and extensive damage can be caused by the settlement. In recent decades, liquefaction-induced building settlement has resulted in enormous losses. During the 1990 Luzon earthquake, buildings in Dagupan city suffered large settlement and tilt as a result of liquefaction [1]. Similar damage occurred during the 1999 Kocaeli earthquake; the temporary loss of bearing capacity induced by the increasing pore water pressure led to building settlement and tilt [1]. During the 2011 Christchurch earthquake, excessive settlement, tilt and lateral movement due to soil liquefaction damaged 15,000 residential structures and affected commercial structures in the Central Business District [2].

Liquefaction-induced settlement has received considerable attention in academia. Physical modelling was used to study the response of foundations on liquefied ground and to provide insight into the mechanisms of liquefaction-induced settlement [3–6]. The influence of soil conditions, structural configurations and earthquake motions on liquefaction-induced settlement were experimentally evaluated. Yoshimi and Tokomatsu [7] used 1-g shaking table tests to investigate the effect of building width on settlement, and their findings were verified with observations collected during the 1964 Niigata earthquake. Generally, building width has a positive effect on building settlement [6]. Bertalot et al. [8] reported that foundation contact pressure significantly affects settlement. Dashti el al [4,9]. categorized the primary mechanisms and key parameters of liquefaction-induced settlement, and they reported that the seismic motions, soil conditions and structural configurations predominantly controlled the settlement mechanisms. Furthermore, nonlinear dynamic analyses were adopted to capture the main features of liquefaction-induced settlement [10–13]. The numerical simulation indicated that the soil conditions, such as the thickness of the dense crust (H_{crust}) and the thickness (HL) and relative density (Dr) of the liquefiable layer, are critical factors in liquefaction-induced settlement [12,14–17]. Additionally, the structural configurations, described by parameters including the width (W) and weight (Q) of the building, and the ground motions have an apparent effect on the liquefaction-induced settlement [3,4,6,8,9]. These aforementioned influencing factors are critical for evaluating liquefaction-induced settlement and are thus considered in this study.

Rational estimates of the liquefaction-induced shallow foundation settlement can help engineers take appropriate measures to reduce losses, which is of great importance in preliminary design in engineering practice. The relationship between liquefaction-induced settlement and related variables was analysed based on collected data from physical models, numerical analyses and field cases [7,8,11,16,18,19]. Liu and Dobry [6] developed an empirical design chart, using the building width

* Corresponding author. School of Civil Engineering, Tianjin University, Tianjin, 300072, China.
E-mail address: hzzhou@tju.edu.cn (H. Zhou).

https://doi.org/10.1016/j.soildyn.2020.106097
Received 31 August 2019; Received in revised form 25 December 2019; Accepted 12 February 2020
Available online 19 February 2020

Nomenclature

H_{crust}	thickness of nonliquefiable crust
HL	thickness of liquefiable layer
Dr	relative density
W	width of building
Q	contact pressure of foundation
G_0	shear modulus number
G	shear modulus
hp_0	contraction rate parameter
$(N_1)_{60}$	standard penetration count
P_a	atmospheric pressure
s_m	mean effective confining pressure
CRR	cyclic resistance ratio
PGA	peak ground acceleration
CAV	cumulative absolute velocity
S_c	measured settlement of centrifuge model
S_n	measured settlement of numerical model
S_p	predicted settlement of MARS model
m	mean value
s	standard deviation
P_{50}	50% cumulative probability
P_{90}	90% cumulative probability
R^2	coefficient of efficiency
GCV	generalized cross-validation
MAE	mean absolute error; and
RMSE	root mean square error

(W) normalized by the thickness of liquefied layers (HL) to predict the normalized liquefaction-induced settlement (S/HL). The influence of high building contact pressure on S/HL was further considered in Ref. [8]. Lu [19] developed a simplified method considering factors including the weight and width of the structure, conditions of the soil and magnitude of the earthquake. Shahir and Pak [11,20] proposed a practical relationship to estimate liquefaction-induced settlement on homogeneous or two-layered sand; this approach considered the density of the sand, weight and width of the building and peak ground acceleration (PGA). Bray and Macedo [16] proposed a procedure for estimating the liquefaction-induced settlement through a regression analysis, considering the influences of the structural configurations, soils and ground motions.

Generally, liquefaction-induced settlement involves a series of affecting factors, including ground motions, soil conditions and structure configurations. An approach that can describe the relationship between the complex input parameters and an output variable is necessary, and a quantitative analysis should be conducted to investigate the effect of each factor and their coupling effects on the liquefaction induced settlement.

This paper attempts to develop a prediction model for the liquefaction-induced settlement of buildings with shallow foundations using a multivariate adaptive regression splines (MARS) algorithm. The MARS model is based on an artificial data set, which is generated from a validated finite difference model. The effect of structure configurations, soil conditions and ground motions are investigated, and the contribution of each parameter is analysed. Validation of the presented model is performed through a generalized cross-validation procedure considering centrifuge test results available in the literature.

2. Numerical modelling

Finite difference method (FDM) analyses were performed using the Fast Lagrangian Analysis of Continua (FLAC version 8) software [21].

2.1. Model configurations

A suite of analyses is performed to study the effect of building configurations, soil conditions and ground motion characteristics on the performance of building settlement due to liquefaction. In natural deposits, soil generally becomes denser with increasing depth, and shallow foundations are always positioned on improved dense crust. Therefore, the soil profile is commonly idealized as three layers: a nonliquefiable dense crust of thickness H_{crust}, a liquefiable sand layer of thickness HL and relative density Dr, and a nonliquefiable base sand layer [3,4]. The baseline model is shown in Fig. 1. The height of the model is 26 m, while the water table is located at an elevation of −1 m [3,4,15]. The element height is set to 1 m to transmit frequencies up to approximately 30 Hz [15]. The range of parameter values examined in these analyses and the values input into the baseline case are given in Table 1. This type of 2D spectral statistical interpolation (SSI) analysis has been successfully employed in the back-analysis of field case studies [14] and centrifuge model tests 9,17,22. The seismic performance of 3D structural systems can be captured by this approach, as demonstrated in prior investigations [14,15].

A typical frame structure is composed of beam elements, and the reinforced mat foundation is defined by its width W and foundation contact pressure Q. The buried depth of the foundation is set to 1 m. As reported by previous studies [15–17,23], the liquefaction-induced

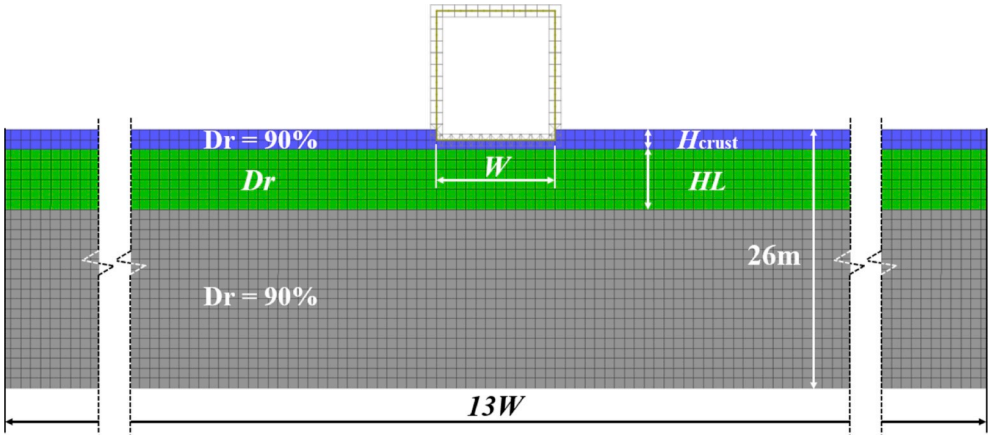

Fig. 1. Model configuration.

Table 1
Parameters considered in this study.

Parameter	Description	Value	Baseline case
Dr	Relative density (%)	30, 45, 60, 75	45
H_{crust}	Thickness of unliquefiable crust (m)	0, 2, 4, 6	2
HL	Thickness of liquefiable layer (m)	2, 3, 6, 9	6
Q	Contact pressure of foundation (kPa)	40, 80, 120, 180, 270, 320	80
W	Structure width (m)	6, 9, 12, 18	12

settlement is not sensitive to the use of interface elements. Thus, a rough interface at which the nodes of beam elements are connected with soil elements was assumed in this study. The density of the beam elements and foundation are altered to obtain the target building contact pressure. The flexural stiffness of the beam, column and foundation are 1280 MN*m^2, 1280 MN*m^2 and 2025 MN*m^2, respectively, to obtain the desired fixed-based structure period, which varies from 0.15 s to 0.95 s [15,24].

2.2. Constitutive models and boundary conditions

PM4Sand version 3 [25] in FLAC is used to model the cyclic response of saturated sand in this study. *PM4Sand* is a stress ratio-controlled, critical state-compatible bounding surface plasticity constitutive model developed for geotechnical earthquake engineering applications [25–27]. The development and calibration procedures for *PM4Sand* are described by Boulanger and Ziotopoulou [25]. There are three primary parameters in the *PM4Sand* model: the relative density (*Dr*), shear modulus number (G_0), and contraction rate parameter (hp_0). For a given $(N_1)_{60cs}$ value, the relative density could be computed as

$$D_r = \sqrt{\frac{(N_1)_{60cs}}{46}} \tag{1}$$

The shear modulus number is determined as

$$G_0 = 167\sqrt{46D_r^2 + 2.5} \tag{2}$$

The elastic shear modulus could be computed as

$$G = G_0 P_a \left(\frac{\sigma_m}{P_a}\right)^{0.5} \tag{3}$$

where P_a is the atmospheric pressure (101.3 kPa) and σ_m is the mean effective confining pressure. The last primary input parameter of the model is the contraction rate parameter (hp_0), which can be determined by element analyses to acquire the target cyclic resistance ratio (CRR) at 15 cycles [15,25]. The target cyclic resistance ratio is defined as a function of relative density by the simplified liquefaction procedures [28]. The 18 secondary model parameters are set to the default values that were calibrated in the prior literature [25] to simulate a generalized sand-like cyclic response. The hydraulic conductivities for the non-liquefiable layer and liquefiable layer are set to 0.001 cm/s and 0.004 cm/s, respectively [17]. The hydraulic conductivities remain constant during the analysis, as recommended by Macedo and Bray [15]. Liquefaction-induced building settlement is slightly sensitive to the value of hydraulic conductivity [15]. Thus, the hydraulic conductivity of sand is not included in the analysed parameters of the paper. As recommended by Boulanger [27], the free-field boundary condition [21] is adopted in this study. The distance from the building to the lateral boundaries of the model is set as six times the width of the building to mitigate the boundary effects.

2.3. Input ground motions

The ground motion data of six earthquake events are scaled to acquire the target PGAs of 0.1 g, 0.2 g, 0.3 g, 0.4 g, 0.5 g and 0.6 g. These records are then baseline-corrected with a linear function and filtered with a Butterworth lowpass filter with a cut-off frequency of 25 Hz. The selected earthquake records are listed in Table 2. Fig. 2 shows the acceleration time histories and normalized acceleration spectra of the six motions. Karimi el al [24,29]. compared several intensity measures for predicting structural performance, and the result showed that the cumulative absolute velocity (*CAV*) of the input motion correlated well with the settlement. Therefore, in this paper, *CAV* is selected as an optimum parameter for predicting foundation settlement and is defined as the integral of the absolute value of the acceleration time history over time:

$$CAV = \int_0^{t_d} |a(t)|dt \tag{4}$$

In this study, the *CAV* value of the input motion ranges from 138.5 cm/s to 1704.3 cm/s, covering a relatively wide range of earthquake intensities.

2.4. Validation of the FDM model

To demonstrate the accuracy of the established FDM model, the results of centrifuge tests conducted by Dashti et al. [4,17] were selected for comparison. The centrifuge tests T3-30, T3-50 and T6-30 included a liquefiable layer of Nevada sand with a prototype thickness of 3 m or 6 m and a relative density in the range of 30%–50%. The loose liquefiable layer was overlain by a 2-m thick layer of dense Monterey sand with a relative density of 85% and underlain by a dense Nevada sand layer with a thickness of 21 m or 18 m and a relative density of 90%. The building width and fixed-based period of prototype structure A were 6 m and 0.21 s, respectively, while the width and fixed-based period of structure B were 12 m and 0.26 s, respectively. The contact pressure of the structures was 80 kPa [4,17]. The input motions were based on the 1995 Kobe earthquake recorded by the Port Island down-hole array and included two scaled motions (the moderate Port Island motion and large Port Island motion). The permeabilities of the liquefiable sand layer and nonliquefiable sand layer were set to 0.004 cm/s and 0.001 cm/s, respectively [17]. A comparison between the numerical settlement and centrifuge results is shown in Fig. 3. The average error of the settlement between the simulation and centrifuge testing is approximately 13.4%, demonstrating that the established numerical model can accurately capture the building settlement.

3. MARS model for predicting liquefaction-induced settlement

The numerical model can accurately capture the settlement, as shown in the aforementioned study. The procedure to develop a MARS model for predicting liquefaction-induced settlement is described as follows:

1. The MARS methodology is briefly introduced.

Table 2
Information of earthquake records.

Earthquake Name	Year	Station Name	Magnitude	$V_{S,\ 30}$
"Kobe "	1995	"KJMA"	6.9	312.0
"Loma Prieta"	1989	"Coyote Lake Dam "	6.93	295.0
"Imperial Valley"	1979	"El Centro Array #11"	6.53	196.2
"Friuli"	1976	"Tolmezzo"	6.5	505.2
"Northridge"	1994	"Castaic - Old Ridge Route"	6.69	450.3
"Chi-Chi "	1999	"TCU045"	7.62	704.6

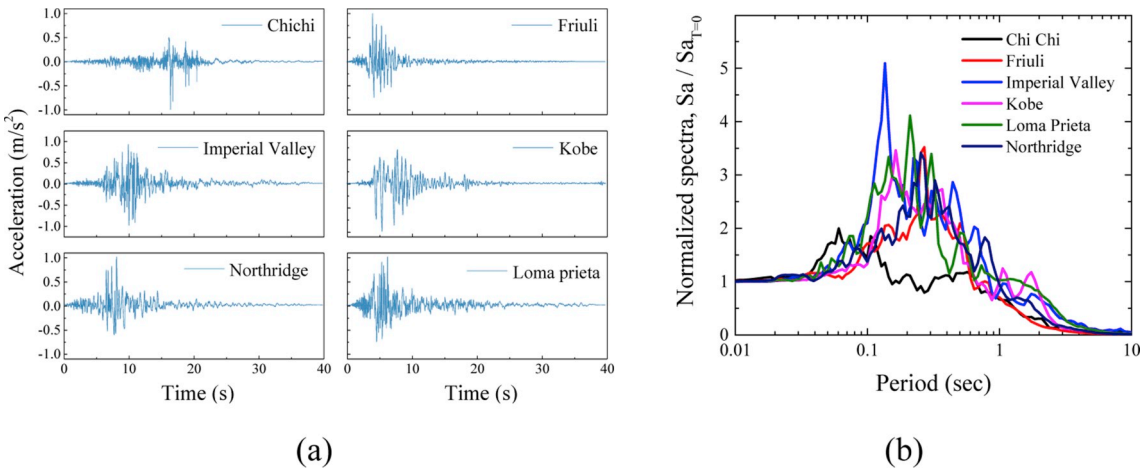

Fig. 2. Input ground motions (a) Acceleration time histories and (b) Normalized acceleration spectra.

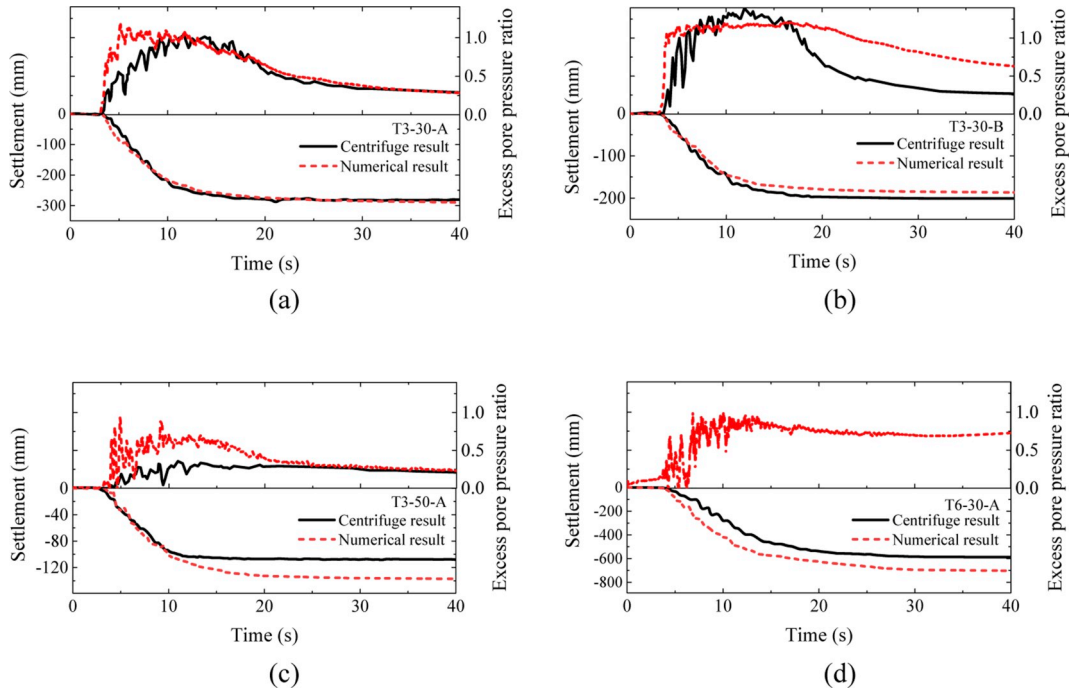

Fig. 3. Comparison between the results of the simulation and centrifuge testing: (a) structure A in T3-30 during moderate Port Island motion; (b) structure B in T3-30 during moderate Port Island motion; (c) structure A in T3-50 during moderate Port Island motion; and (d) structure A in T6-30 during large Port Island motion.

2. A substantial amount of artificial data, including a relatively wide range of input variables, are produced using the validated FDM model.
3. The MARS model, which includes six input parameters, is developed to evaluate the liquefaction-induced settlement.
4. The proposed model is validated using the artificial data.
5. The relative importance of the input parameters is estimated, and the coupled interactions of input values are analysed.

3.1. Methodology of MARS

MARS is a nonparametric and nonlinear statistical regression method developed by Friedman [30], and the underlying functional relationship between input and output variables can be revealed by this method without requiring any prior assumptions. The MARS model is able to identify optimal variable transformations and complex interactions between the output and high-dimensional inputs. In recent years, the MARS model has been widely applied in geotechnical engineering 31–34, including determining the undrained strength of clay [35], estimating the shaft resistance of piles 36,37, analyzing the slope stability [38,39] and estimating liquefaction induced deformations [40–43]. The end points of the segments (splines), which are called knots or nodes, mark the end of one piecewise set of data and the beginning of another. The resulting piecewise curves, known as basis functions (BFs), allow for bending, thresholds and other non-linear features [42,44,45]. Therefore, the MARS algorithm is efficient for exposing the complex interactions and the relationships between input variables and the output value.

The fitting of the MARS model includes forward and backward iterative approaches. The first step is to include the BFs by creating a forward stepwise model, and the locations of potential knots are identified to improve the performance of the model. Each subsequent step adapts a knot and its correlated pair of BFs, which produce the largest

decrease in the residual sum of squares error. The aim of this algorithm is to identify every possible predictor and potential knot location for every predictor. Generally, the BFs are introduced in some regions for each predictor. The expression of the global model that combines the BFs and interactions between them is as follows.

$$f(X) = \alpha_0 + \sum_{m=1}^{M} \alpha_m F_m(X) \tag{5}$$

where F_m is the mth BF, and it can be a spline function or the interaction of two or more splines functions; M is the number of BFs included in the model; α is a constant coefficient estimated by the least squares method; and $f(X)$ is the predicted value. During every subsequent step, the BF that produces the largest decrease in the training error is added into the model. This process continues until a predetermined maximum number of BFs is reached, but this might lead to overfitting. Therefore, a pruning phase is conducted to delete redundant BFs in the overfitted model.

The second step is the backward pruning algorithm: the redundant BFs that provide the smallest contributions are deleted by utilizing the generalized cross-validation (GCV) technique, which is less computationally expensive than other techniques. The GCV equation is a mean squared residual error divided by a penalty that depends on the complexity of the model. GCV can be expressed by Eq. (6).

$$GCV = \frac{\frac{1}{N}\sum_{i=1}^{N}[y_i - f(x_i)]^2}{\left[1 - \frac{M + d \times (M-1)/2}{N}\right]^2} \tag{6}$$

where M and N are the number of BFs and data points, respectively; d is the penalizing parameter; and $f(x_i)$ represents the value predicted by the MARS model. The term $(M-1)/2$ represents the number of knots; thus, the GCV value penalizes the number of BFs as well as the number of knots and serves to reduce the chance of overfitting. The penalizing parameter d is set to 3 as a default value, further recommendations on choosing the value of d are provided by Friedman [30]. At each pruning step, a BF is removed to minimize Eq. (6) until an optimal model is established. After the model is determined, an assessment of the relative importance of the input parameters can be produced by observing the GCV decrease when a parameter is eliminated from the model.

3.2. Data generation

To develop an accurate MARS model, an adequate database with systematically varying input variables is necessary [46]. A total of 1400 analyses were performed, approximately 75% of the total data (1050 datapoints) were randomly selected as the training dataset, while the remaining 25% of the total data (350 datapoints) are utilized for the model testing [47,48]. Table 3 shows the statistical parameters of the input variables and their ranges considered in the FDM model for the generation of artificial data. The ranges of the training data and testing data are equivalent, and the subsets for training and testing are statistically consistent.

Table 3
Statistical parameters of the training and testing data.

		H_{crust} (m)	HL (m)	W (m)	Dr (%)	Q (kPa)	CAV (cm/s)
Training data	Min.	0.00	2.00	6.00	30.00	40.00	138.53
	Max.	6.00	9.00	18.00	75.00	320.00	1704.32
	Avg.	2.63	4.65	8.68	53.07	133.59	784.28
	Std. dev.	1.54	2.78	2.95	16.28	98.73	380.03
Testing data	Min.	0.00	2.00	6.00	30.00	40.00	138.53
	Max.	6.00	9.00	18.00	75.00	320.00	1704.32
	Avg.	2.69	4.55	8.70	52.13	142.14	778.75
	Std. dev.	1.64	2.77	2.97	17.13	102.20	359.21

3.3. Model development

A MARS model is developed for estimating the liquefaction-induced settlement (S_p) as a function of six input parameters: thickness of the nonliquefiable crust layer (H_{crust}), thickness of the liquefiable layer (HL), relative density of the liquefiable layer (Dr), width (W) of the structure, contact pressure of the foundation (Q), and cumulative absolute velocity (CAV) of the input ground motions [12,14–17]. The GCV value shown in Eq. (6) is utilized to evaluate the optimal number of BFs for the model through backward iterative approaches. As shown in Fig. 4, a minimum GCV value is found with 38 BFs, and the model with the lowest GCV value is selected as the optimized model [40,49,50]. Thus, this number of BFs is optimal and is adopted to develop the MARS model. The MARS model and the BFs for predicting the liquefaction-induced settlement of shallow-founded buildings are listed in Table 4. As shown in this table, some BFs include two or three splines functions, that indicating the model is able to reveal interactions, and these interactions play an important role in the generation of an accurate model to estimate the liquefaction-induced settlement considering the influence of the structural configurations, soil conditions and ground motions.

3.4. Comparison with artificial data

The statistical performance, cumulative probability, arithmetic calculation, and ±30% accuracy level (in percentage using a histogram probability distribution of predicted value/numerical value), are utilized to evaluate the MARS model performance [49]. The predicted value versus numerical value, mean absolute error (MAE), root mean squared error (RMSE) and coefficient of determination (R^2) for the training and testing data are shown in Fig. 5. The MAE measures the variation of the error term by term and eliminates the large errors; the RMSE value gives more attention to large errors than small errors; and the R^2 value indicates the interpretation of the dependent variable by the independent variables. Generally, the error between the predicted settlement and the numerical settlement is within the range of ±30%, which shows that the MARS model has a reasonable accuracy. The R^2 values for the training and testing data are 0.951 and 0.946, respectively. This result shows that the MARS model can reasonably and generally predict building settlement. Thus, the MARS model is capable of predicting the liquefaction-induced settlement and can reveal the

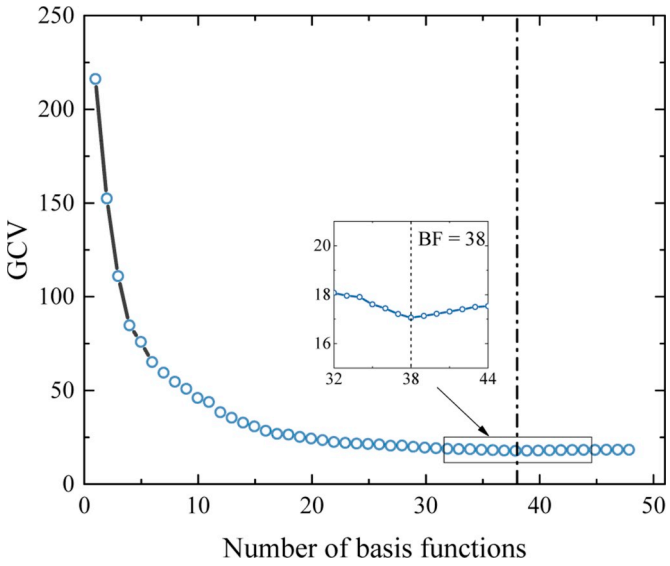

Fig. 4. The performance the MARS models versus the number of basis functions.

Table 4
Basis functions and corresponding equations of the MARS model for building settlement.

BF	Equation	BF	Equation
BF1	max (0, *Dr* −45)	BF20	max (0,6 -*HL*) * max (0,60 -*Dr*) * max (0,2 -H_{crust})
BF2	max (0,45 -*Dr*)	BF21	BF13 * max (0, *HL* -3)
BF3	max (0, *CAV* -1458.3)	BF22	BF13 * max (0,3 -*HL*)
BF4	max (0,1458.3 -*CAV*)	BF23	BF4 * max (0, *Dr* −45)
BF5	max (0, *Q* −80)	BF24	BF4 * max (0,45 -*Dr*)
BF6	max (0,80 -*Q*)	BF25	BF1 * max (0, *Q* −180)
BF7	max (0,45 -*Dr*) * max (0, *HL* -3) * max (0, *Q* −40)	BF26	BF1 * max (0,180 -*Q*)
BF8	max (0,45 -*Dr*) * max (0, *HL* -3) * max (0, H_{crust} −2)	BF27	max (0, H_{crust} −2)
BF9	max (0,45 -*Dr*) * max (0, *HL* -3) * max (0,2 -H_{crust})	BF28	max (0,2 -H_{crust})
BF10	max (0, *HL* -6)	BF29	BF25 * max (0, H_{crust} −2)
BF11	max (0,6 -*HL*)	BF30	BF24 * max (0, H_{crust} −3)
BF12	BF1 * max (0, *CAV* -1414.5)	BF31	BF1 * max (0, *HL* -6)
BF13	BF1 * max (0,1414.5 -*CAV*)	BF32	BF1 * max (0,6 -*HL*)
BF14	BF4 * max (0, *HL* -3)	BF33	BF27 * max (0, *HL* -3)
BF15	BF4 * max (0,3 -*HL*)	BF34	BF27 * max (0,3 -*HL*)
BF16	BF4 * max (0, *Q* −80)	BF35	max (0,3 -H_{crust})
BF17	BF4 * max (0,80 -*Q*)	BF36	BF27 * max (0, *Dr* −60)
BF18	max (0, W −9)	BF37	BF23 * max (0, *Q* −120)
BF19	max (0,6 -*HL*) * max (0,60 -*Dr*) * max (0, H_{crust} −2)	BF38	BF23 * max (0,120 -*Q*)

S_p = 43.282–2.4469*BF1 +0.65562*BF2 -0.15248*BF3 -0.019607*BF4 +0.073679*BF5 -0.34032*BF6 +0.0018038*BF7 -0.083799*BF8 +0.027917*BF9 +4.7678*BF10–6.4522*BF11 + 0.011123*BF12–0.021856*BF13–0.0041906*BF14 + 0.0057044*BF15–6.3622e-05*BF16 + 0.00020685*BF17–1.1875*BF18–0.045251*BF19 + 0.026675*BF20 + 0.00013386*BF21–0.00031323*BF22 + 0.022356*BF23–0.00026892*BF24–0.0012185*BF25 + 0.0025707*BF26 + 1.2193*BF27–10.216*BF28 + 0.00045645*BF29 + 0.00029312*BF30–0.11406*BF31 + 0.22738*BF32–0.18402*BF33 + 8.7351*BF34 + 8.9365*BF35 + 0.089612*BF36 + 1.0593e-06*BF37–2.2714e-06*BF38.

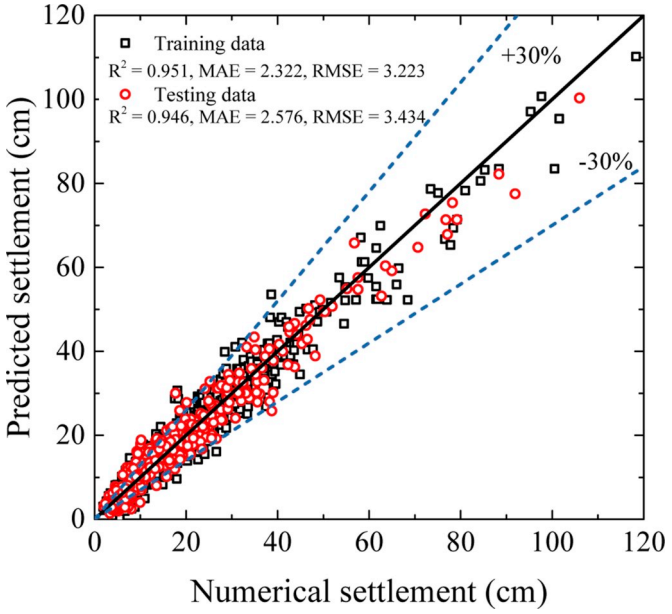

Fig. 5. Comparison between the target and MARS-predicted displacements.

complex relationships between the liquefaction-induced settlement and the input variables.

Some typical input variables and predicted results for the training and testing data are listed in Table 5 and Table 6, respectively, indicating that for most cases, the value predicted by the MARS model is relatively close to the numerical value.

Fig. 6 presents a histogram of the ratio of the predicted settlement (S_p) from the MARS model to the numerical settlement (S_n) results. Generally, most MARS-predicted settlement fell within ±30% of the numerical values. The mean value (μ) and standard deviation (σ) of the ratio S_p/S_n are selected as significant indicators to assess the MARS model's accuracy [48]. The mean values (μ) for the training data and testing data are 0.997 and 0.981, respectively, while the standard deviations (σ) are 0.174 and 0.163, respectively. Generally, the ratio of the MARS-predicted values to the numerical results is normally distributed, indicating that the error is distributed randomly.

The cumulative probability is important in the evaluation of the predictive model performance [47–49]. The ratios of the MARS-predicted settlement (S_p) to the numerical settlement (S_n) are arranged in ascending order, and the calculation of the cumulative probability is expressed as follows:

$$P = \frac{i}{n+1} \tag{7}$$

where n is the number of data points and i is the order number given to the S_p/S_n ratio.

Fig. 7 shows the cumulative probability (%) of S_p/S_n for the training and testing data for the MARS model. The values of the cumulative probabilities P_{50} and P_{90} for the MARS training data are 1.003 and 1.294, respectively, while those values for the MARS testing data are 0.995 and 1.348, respectively. The prediction of the MARS model is fair at a cumulative probability of 50%, while producing an overestimate at a cumulative probability of 90%.

3.5. Relative importance of input parameters

Fig. 8 shows the relative importance of the input variables considered in this MARS model. The relative importance is assessed using the increment of GCV value determined by eliminating the specified variables from the MARS model. The results show that the relative density (*Dr*) and thickness (*HL*) of the liquefiable layer and the cumulative absolute velocity (*CAV*) of the earthquake are the critically important parameters among the considered factors. The contact pressure (*Q*) and the structure width (*W*), as well as the thickness of the nonliquefiable crust (H_{crust}), also have an apparent influence on the settlement.

The tendency of predicted liquefaction-induced settlement considering a range of soil conditions and structure configurations is shown as follows. Fig. 9 shows the trends for liquefaction-induced building settlement affected by the thickness and the relative density of the liquefiable layer. As the picture depicts, the building settlement increases with the increasing liquefiable layer thickness, especially for a relative density lower than 50%. As expected, the magnitude of settlement significantly decreases with increasing relative density of the liquefiable layer, and the building settlement is less sensitive to the liquefiable layer thickness for a relative density greater than approximately 70%. Fig. 10 shows the effects of structural width and contact pressure on the

Table 5
Typical training data predicted by the MARS model.

H_{crust} (m)	*HL* (m)	*W* (m)	*Dr* (%)	*Q* (kPa)	*CAV* (cm/s)	S_n (cm)	S_p (cm)
6.00	9.00	6.00	75.00	180.00	1453.49	19.04	14.32
3.00	2.00	9.00	30.00	270.00	554.12	18.41	18.68
3.00	2.00	6.00	75.00	120.00	254.63	8.06	7.18
2.00	9.00	12.00	60.00	180.00	1018.48	37.94	28.85
3.00	2.00	9.00	60.00	40.00	1507.44	6.54	4.75
0.00	3.00	6.00	30.00	40.00	1018.48	20.12	24.57
2.00	6.00	6.00	45.00	80.00	1458.31	61.51	52.22
2.00	6.00	12.00	45.00	80.00	1527.75	37.81	38.06
6.00	3.00	6.00	30.00	320.00	909.59	28.55	25.92
2.00	6.00	18.00	45.00	80.00	763.88	21.13	19.19

Table 6
Typical testing data predicted by the MARS model.

H_{crust} (m)	HL (m)	W (m)	Dr (%)	Q (kPa)	CAV (cm/s)	S_n (cm)	S_p (cm)
3.00	2.00	6.00	60.00	270.00	848.71	22.52	19.22
0.00	9.00	6.00	30.00	180.00	475.85	57.57	54.74
2.00	3.00	12.00	30.00	320.00	848.71	33.00	33.10
3.00	2.00	6.00	75.00	40.00	682.21	3.12	4.51
6.00	3.00	6.00	75.00	320.00	909.59	23.28	22.77
2.00	3.00	12.00	30.00	180.00	1018.48	23.46	33.31
2.00	6.00	18.00	45.00	80.00	1527.75	25.21	30.94
3.00	2.00	9.00	45.00	270.00	277.07	7.44	8.01
3.00	2.00	12.00	45.00	120.00	1414.49	22.54	23.38
2.00	9.00	6.00	30.00	180.00	1018.48	88.36	82.20

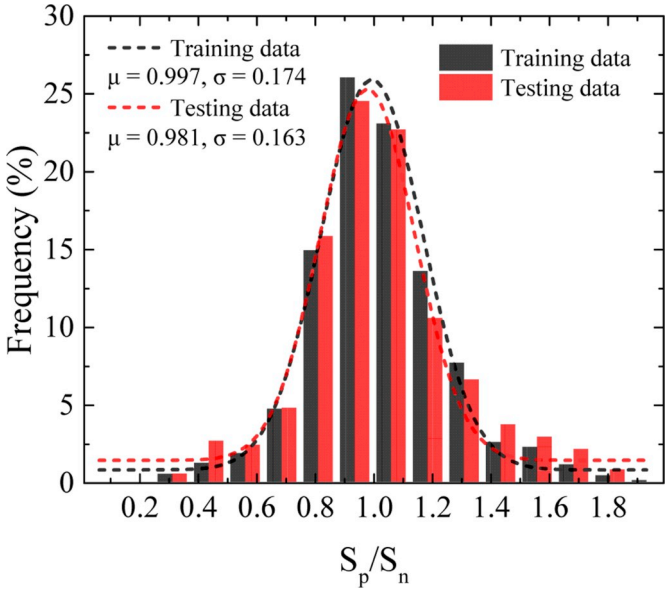

Fig. 6. Histogram of the relative errors of the MARS model.

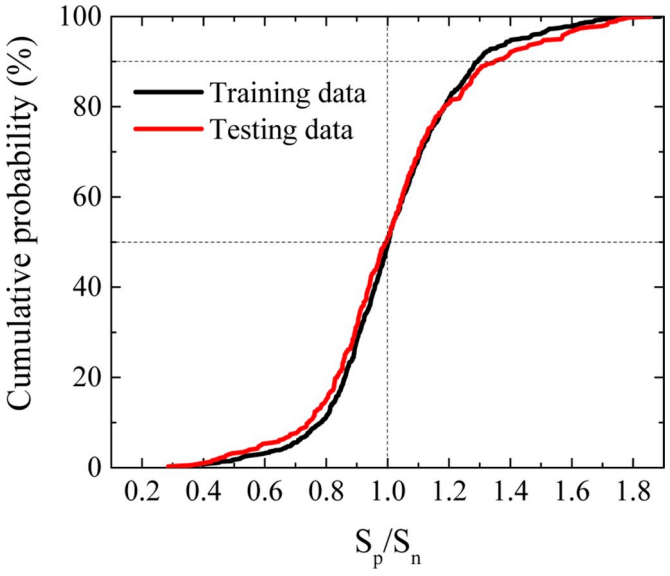

Fig. 7. Cumulative probability of S_p/S_n for the training and testing data for the MARS model.

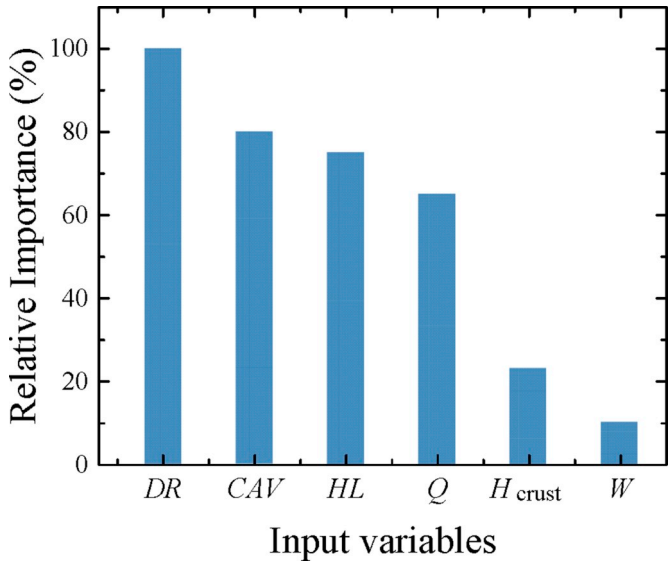

Fig. 8. Relative importance of the input variables selected in the MARS model.

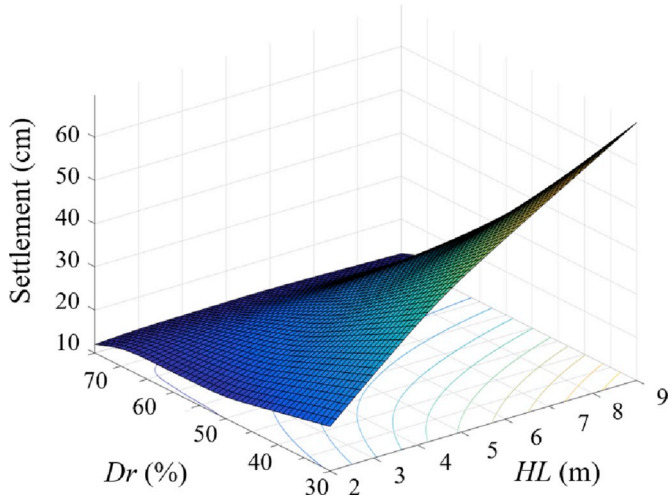

Fig. 9. Effect of relative density (Dr) and thickness (HL) of the liquefiable layer on foundation settlement.

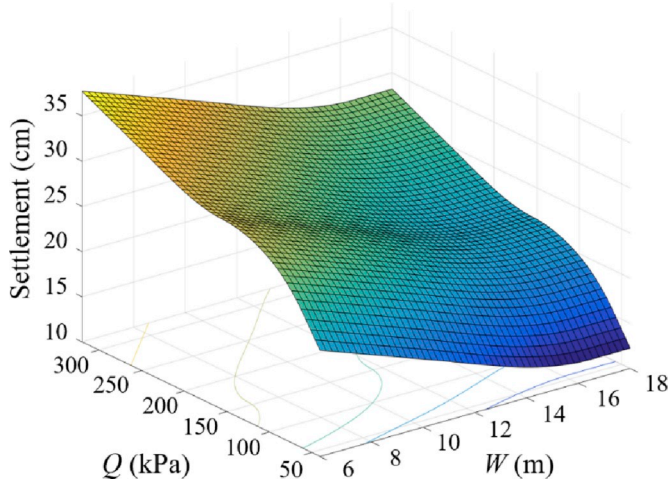

Fig. 10. Effect of the building width (W) and contact pressure (Q) on foundation settlement.

liquefaction-induced settlement. The building settlement increases greatly with a relatively low contact pressure ($Q < 80$ kPa), but the rate of increase decreases with the increasing contact pressure (80 kPa $< Q <$ 180 kPa). The rate of increase of building settlement keeps steady when the contact pressure is larger than 180 kPa. Additionally, the building settlement decreases moderately with increasing building width. These results coincide with those of previous studies [11,15,24].

4. Evaluation of the MARS model using centrifuge test data

As shown in Table 7, a group of 38 cases is collected from well-documented centrifuge tests to assess the applicability of the MARS model. Fig. 11 presents a comparison between the MARS model predictions and the centrifuge results from the prior literature. The MARS model predictions exhibit consistency with the centrifuge measurements; the error between the prediction and centrifuge test is within approximately ±30%, and the coefficient of determination (R^2) is 0.884. Thus, the developed MARS model generally agrees well with the centrifuge results, further demonstrating that it is capable of estimating the settlement of a shallow-founded building located on liquefiable layers.

5. Conclusions

This paper presents a MARS model to predict the liquefaction-induced settlement of buildings with shallow foundations. The accuracy of the presented MARS model is demonstrated by the results of different statistical parameters such as the mean average error (MAE), the root mean squared error (RMSE) and the coefficient of determination (R^2). A sensitivity analysis is carried out to assess the relative importance of the input parameters in the model. For the presented model, the relative density of the liquefiable layer is the most important parameter, the thickness of the liquefiable layer and the intensity of the earthquake

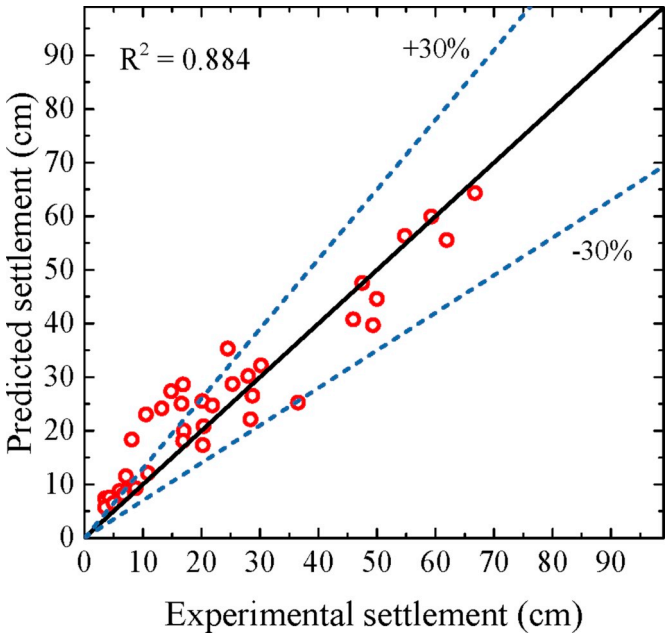

Fig. 11. MARS-predicted settlement versus settlement from centrifuge tests.

Table 7
Summary of centrifuge database from the available literature.

Test ID	H_{crust} (m)	*HL* (m)	*W* (m)	*H* (m)	*Dr* (%)	*Q* (kPa)	*CAV* (cm/s)	S_c (cm)	Reference
T6-30-03-A	2	6	6	2.12	40	80	405.83	8.0	Dashti et al. [51]
T6-30-03-B	2	6	12	2.12	40	80	405.83	9.5	
T6-30-03-C	2	6	6	4.2	40	130	405.83	10.5	
T6-30-04-A	2	6	6	2.12	40	80	1409.2	59.3	
T6-30-04-B	2	6	12	2.12	40	80	1409.2	54.8	
T6-30-04-C	2	6	6	4.2	40	130	1409.2	66.8	
T3-30-03-A	2	3	6	2.12	30	80	462.5	28.4	
T3-30-03-B	2	3	12	2.12	30	80	462.5	20.2	
T3-30-03-C	2	3	6	4.2	30	130	462.5	17.0	
T3-30-04-A	2	3	6	2.12	30	80	1461.3	50.0	
T3-30-04-B	2	3	12	2.12	30	80	1461.3	45.9	
T3-30-04-C	2	3	6	4.2	30	130	1461.3	47.5	
T3-50-02-A	2	3	6	2.12	50	80	404.43	10.8	
T3-50-04-A	2	3	6	2.12	50	80	1243.1	20.2	
T3-50-05-A	2	3	6	2.12	50	80	1566.9	16.9	
T4.5-50-09-A	1.9	4.5	6	2.12	50	65	556.18	20.0	Zupan et al. [51]
T4.5-50-10-A	1.9	4.5	6	2.12	50	65	1219.8	16.6	
T4.5-50-11-A	1.9	4.5	6	2.12	50	65	1297.7	25.3	
T4.6-40-05-A	1.7	4.6	6	2.12	40	65	575.9	20.4	Hayden et al. [51]
T4.6-40-05-K	1.7	4.6	6	2.42	40	177	575.9	28.7	
T4.6-40-08-A	1.7	4.6	6	2.12	40	65	1476.8	49.3	
T4.6-40-08-K	1.7	4.6	6	2.42	40	177	1476.8	61.9	
T3.9-50-03-J_E	2.8	3.9	7.5	10.5	50	179	491.8	7.1	Allmond and Kutter [51].
T3.9-50-07-J_E	2.8	3.9	7.5	10.5	50	179	1457.3	30.2	
T2.5-55-03-J	2.6	2.5	7.5	10.5	55	179	427.3	5.9	
T2.5-55-03-J_IM	2.6	2.5	7.5	16.4	55	269	427.3	8.8	
T2.5-55-06-J	2.6	2.5	7.5	10.5	55	179	438.6	4.3	
T2.5-55-06-J_IM	2.6	2.5	7.5	12	55	269	438.6	6.0	
T2.5-55-07-J	2.6	2.5	7.5	10.5	55	179	1522.7	21.9	
T2.5-55-07-J_IM	2.6	2.5	7.5	12	55	269	1522.7	28.0	
T2.5-55-11-J	2.6	2.5	7.5	10.5	55	179	1357.1	10.5	
T2.5-55-11-J_IM	2.6	2.5	7.5	12	55	269	1357.1	14.8	
T2.5-55-12-J	2.6	2.5	7.5	10.5	55	179	1307.5	8.1	
T2.5-55-12-J_IM	2.6	2.5	7.5	12	55	269	1307.5	13.2	
T2.3-70-16-J	2.5	2.3	7.5	10.5	70	179	415.8	3.6	NEES-2008-0565-Exp 610.4231/D37D2Q74Z [51].
T2.3-70-16-J_IM	2.5	2.3	7.5	12	70	269	415.8	5.0	
T2.3-70-21-J	2.5	2.3	7.5	10.5	70	179	1456.7	6.7	
T2.3-70-21-J_IM	2.5	2.3	7.5	12	70	269	1456.7	7.1	

(represented by *CAV*) significantly influence liquefaction-induced settlement, and the width and contact pressure of the structure have an apparent effect on the settlement. A wider and lighter building tends to exhibit less liquefaction-induced settlement; a reduction in liquefaction-induced settlement is also caused by increasing the relative density and decreasing the thickness of liquefiable soil. The established model is compared with a series of centrifuge test results, showing reasonable consistency, indicating that the model can be used to preliminarily estimate the liquefaction-induced settlement of buildings with shallow foundations.

Notably, the MARS model is based on data from numerical simulation. The continuum-based numerical model has difficulties in predicting the settlement induced by sand ejecta [16,24], but the nonlinear dynamic SSI analyses can identify the volumetric-induced and the deviatoric-induced building settlement [15]. Because the model is established by artificial data produced under simplified soil conditions, and idealized structure configurations, the developed model can be applied only to scenarios with similar soil and structural conditions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Gang Zheng: Resources, Conceptualization, Project administration, Funding acquisition. **Wenbin Zhang:** Software, Methodology, Writing - original draft. **Haizuo Zhou:** Supervision, Writing - review & editing, Methodology, Funding acquisition. **Pengbo Yang:** Visualization, Writing - review & editing.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (No. 51708405 and 41630641), the Project of self-dependent innovation funding (2019XZC-0027) and the Project of Tianjin Science and Technology Plan (No. 16YDLJSF00040). The authors appreciate the financial support.

References

[1] Sancio R, Bray JD, Durgunoglu T. Performance of buildings over liquefiable ground in Adapazari, Turkey. In: 13th world conf. On earthquake engineering, Canadian association for earthquake engineering, Vancouve, Canada; 2004.
[2] Green RA, Allen J, Wotherspoon L, Cubrinovski M, Bradley B, Bradshaw A, Cox B, Algie T. Performance of levees (Stopbanks) during the 4 September 2010 Mw 7.1 Darfield and 22 February 2011 Mw 6.2 Christchurch, New Zealand, earthquakes. Seismol Res Lett 2011;82:939–49.
[3] Dashti S, Bray JD, Pestana JM, Riemer M, Wilson D. Centrifuge testing to evaluate and mitigate liquefaction-induced building settlement mechanisms. J Geotech Geoenviron Eng 2010;136:918–29.
[4] Dashti S, Bray JD, Pestana JM, Riemer M, Wilson D. Mechanisms of seismically induced settlement of buildings with shallow foundations on liquefiable soil. J Geotech Geoenviron Eng 2010;136:151–64.
[5] Mehrzad B, Haddad A, Jafarian Y. Centrifuge and numerical models to investigate liquefaction-induced response of shallow foundations with different contact pressures. Int J Civ Eng 2016;14:117–31.
[6] Liu L, Dobry R. Seismic response of shallow foundation on liquefiable sand. J Geotech Geoenviron Eng 1997;123:557–67.
[7] Yoshimi Y, Tokimatsu K. Settlement of buildings on saturated sand during earthquakes. Soils Found 1977;17:23–38.
[8] Bertalot D, Brennan AJ, Villalobos FA. Influence of bearing pressure on liquefaction-induced settlement of shallow foundations. Geotechnique 2013;63:391–9.
[9] Bray JD, Dashti S. Liquefaction-induced building movements. Bull Earthq Eng 2014;12:1129–56.
[10] Popescu R, Prevost JH, Deodatis G, Chakrabortty P. Dynamics of nonlinear porous media with applications to soil liquefaction. Soil Dynam Earthq Eng 2006;26:648–65.
[11] Shahir H, Pak A. Estimating liquefaction-induced settlement of shallow foundations by numerical approach. Comput Geotech 2010;37:267–79.
[12] Karamitros DK, Bouckovalas GD, Chaloulos YK. Seismic settlements of shallow foundations on liquefiable soil with a clay crust. Soil Dynam Earthq Eng 2013;46:64–76.
[13] Karimi Z, Dashti S. Numerical and centrifuge modeling of seismic soil–foundation–structure interaction on liquefiable ground. J Geotech Geoenviron Eng 2016;142:04015061.
[14] Roberto L, Bray JD. Dynamic analyses of two buildings founded on liquefiable soils during the canterbury earthquake sequence. J Geotech Geoenviron Eng 2017;143:04017067.
[15] Macedo J, Bray JD. Key trends in liquefaction-induced building settlement. J Geotech Geoenviron Eng 2018;144:04018076.
[16] Bray JD, Macedo J. 6th Ishihara lecture: simplified procedure for estimating liquefaction-induced building settlement. Soil Dynam Earthq Eng 2017;102:215–31.
[17] Dashti S, Bray JD. Numerical simulation of building response on liquefiable sand. J Geotech Geoenviron Eng 2013;139:1235–49.
[18] Acacio AA, Kobayashi Y, Towhata I, Bautista RT, Ishihara K. Subsidence of building foundation resting upon liquefied subsoil : case studies and assessment. Soils Found 2001;41:111–28.
[19] Lu CW. A simplified calculation method for liquefaction-induced settlement of shallow foundation. J Earthq Eng 2017;21:1385–405.
[20] Ayoubi P, Pak A. Liquefaction-induced settlement of shallow foundations on two-layered subsoil strata. Soil Dynam Earthq Eng 2017;94:35–46.
[21] FLAC—Fast Lagrangian analysis of continua (version 8.0), Itasca Consulting Group, Inc, Minneapolis.
[22] Zhang W, Han L, Chen Z, Feng L, Ding X, Liu H. Model tests on seismic performance of double-box underground utility tunnel. Chin J Geotech Eng 2020;42:100–8. https://doi.org/10.11779/CJGE202001011.
[23] Dimitriadi VE, Bouckovalas GD, Papadimitriou AG. Seismic performance of strip foundations on liquefiable soils with a permeable crust. Soil Dynam Earthq Eng 2017;100:396–409.
[24] Karimi Z, Dashti S, Bullock Z, Porter K, Liel AB. Key predictors of structure settlement on liquefiable ground: a numerical parametric study. Soil Dynam Earthq Eng 2018;113:286–308.
[25] Boulanger RW, Ziotopoulou K. In: A sand plasticity model for earthquake engineering applications. Davis, CA: Center for Geotechnical Modeling, Dept. of Civil and Environmental Engineering, Univ. of California; 2015.
[26] Boulanger RW, Khosravi M, Khosravi A, Wilson DW. Remediation of liquefaction effects for an embankment using soil-cement walls: centrifuge and numerical modeling. Soil Dynam Earthq Eng 2018;114:38–50.
[27] Boulanger RW, Montgomery J. Nonlinear deformation analyses of an embankment dam on a spatially variable liquefiable deposit. Soil Dynam Earthq Eng 2016;91:222–33.
[28] Boulanger RW, Idriss IM. CPT-based liquefaction triggering procedure. J Geotech Geoenviron Eng 2016;142.
[29] Karimi Z, Dashti S. Ground motion intensity measures to evaluate II: the performance of shallow-founded structures on liquefiable ground. Earthq Spectra 2017;33:277–98.
[30] Friedman JH. Multivariate adaptive regression splines. Ann Stat 1991;19:1–67.
[31] Goh ATC, Zhang Y, Zhang R, Zhang W, Xiao Y. Evaluating stability of underground entry-type excavations using multivariate adaptive regression splines and logistic regression. Tunn Undergr Space Technol 2017;70:148–54.
[32] Goh ATC, Zhang W, Zhang Y, Xiao Y, Xiang Y. Determination of earth pressure balance tunnel-related maximum surface settlement: a multivariate adaptive regression splines approach. Bull Eng Geol Environ 2018;77:489–500.
[33] Zhang W, Zhang R, Wu C, Goh ATC, Lacasse S, Liu Z, et al. State-of-the-art review of soft computing applications in underground excavations. Geosci Front 2019. https://doi.org/10.1016/j.gsf.2019.12.003.
[34] Zheng G, He X, Zhou H, Yang X, Yu X, Zhao J. Prediction of the tunnel displacement induced by laterally adjacent excavations using multivariate adaptive regression splines. Acta Geotechnica 2020. https://doi.org/10.1007/s11440-020-00916-w.
[35] Samui P, Kurup P. Multivariate adaptive regression spline and least square support vector machine for prediction of undrained shear strength of clay. Int J Appl Metaheuristic Comput (IJAMC) 2012;3:33–42.
[36] Lashkari A. Prediction of the shaft resistance of nondisplacement piles in sand. Int J Numer Anal Methods GeoMech 2013;37:904–31.
[37] Zhang W, Wu C, Li Y, Wang L, Samui P. Assessment of pile drivability using random forest regression and multivariate adaptive regression splines. Georisk 2019. https://doi.org/10.1080/17499518.2019.1674340.
[38] Wang lin, Wu C, Gu X, Liu H, Mei G, Zhang W. Probabilistic stability analysis of earth dam slope under transient seepage using multivariate adaptive regression splines. Bull Eng Geol Environ 2020. https://doi.org/10.1007/s10064-020-01730-0.
[39] Wang L, Wu C, Li Y, Liu H, Zhang W, Chen X. Probabilistic Risk Assessment of unsaturated Slope Failure Considering Spatial Variability of Hydraulic Parameters. KSCE J Civ Eng 2019. https://doi.org/10.1007/s12205-019-0884-6.
[40] Zhang W, Goh ATC. Multivariate adaptive regression splines for analysis of geotechnical engineering systems. Comput Geotech 2013;48:82–95.
[41] Youd TL, Hansen CM, Bartlett SF. Revised multilinear regression equations for prediction of lateral spread displacement. J Geotech Geoenviron Eng 2002;128:1007–17.
[42] Zhang W, Goh ATC, Zhang Y, Chen Y, Xiao Y. Assessment of soil liquefaction based on capacity energy concept and multivariate adaptive regression splines. Eng Geol 2015;188:29–37.

[43] Zheng G, Yang P, Zhou H, Zeng C, Yang X, He X, Yu X. Evaluation of the earthquake induced uplift displacement of tunnels using multivariate adaptive regression splines. Comput Geotech 2019;113.
[44] Zhang W, Goh ATC. Regression models for estimating ultimate and serviceability limit states of underground rock caverns. Eng Geol 2015;188:68–76.
[45] Zhang W, Goh ATC. Multivariate adaptive regression splines and neural network models for prediction of pile drivability. Geosci Front 2016;7:45–52.
[46] Zheng G, Yang XY, Zhou HZ, Du YM, Sun JY, Yu XX. A simplified prediction method for evaluating tunnel displacement induced by laterally adjacent excavations. Comput Geotech 2018;95:119–28.
[47] Mohanty R, Das SK. Settlement of shallow foundations on cohesionless soils based on SPT value using multi-objective feature selection. Geotech Geol Eng 2018;36:3499–509.
[48] Suman S, Khan SZ, Das SK, Chand SK. Slope stability analysis using artificial intelligence techniques. Nat Hazards 2016;84:727–48.
[49] Das SK, Suman S. Prediction of lateral load capacity of pile in clay using multivariate adaptive regression spline and functional network. Arabian J Sci Eng 2015;40:1565–78.
[50] Zhang W, Zhang R, Wang W, Zhang F, Goh ATC. A Multivariate Adaptive Regression Splines model for determining horizontal wall deflection envelope for braced excavations in clays. Tunn Undergr Space Technol 2019;84:461–71.
[51] Allmond J, Kutter BL, Bray JD, Hayden CP. New database for foundation and ground performance in liquefaction experiments. Earthq Spectra 2015;31:2485–509.