TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Yin, X., Wang, H., Pisanò, F., Gavin, K., Askarinejad, A., & Zhou, H. (2023). Deep learning-based design model for suction caissons on clay. *Ocean Engineering*, *286*, Article 115542. https://doi.org/10.1016/j.oceaneng.2023.115542

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Ocean Engineering 286 (2023) 115542

Contents lists available at ScienceDirect

Ocean Engineering

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

Deep learning-based design model for suction caissons on clay

Xilin Yin [a], Huan Wang [a,c,*], Federico Pisanò [a], Ken Gavin [a], Amin Askarinejad [a], Hongpeng Zhou [b,**]

[a] Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, 2628 CN, The Netherlands
[b] Department of Computer Science, Faculty of Science and Engineering, The University of Manchetser, Manchetser, M13 9PL, United Kingdom
[c] Advanced Modelling Section, Norwegian Geotechnical Institute, Oslo, 0484, Norway

ARTICLE INFO

Keywords:
Caisson foundation
Load–displacement relationship
Deep learning
Numerical modelling

ABSTRACT

Predicting the non-linear loading response is the key to the design of suction caissons. This paper presents a systematic study to explore the applicability of deep learning techniques in foundation design. Firstly, a series of three-dimensional finite element simulations was performed, covering a wide range of embedment ratios and different loading directions, to provide training data for the deep neural network (DNN) model. Then, hyper-parameter tuning was performed and it is found that the basic Fully-Connected (FC) neural network model is sufficient to capture the non-linear response of suction caissons with excellent accuracy and robustness. Furthermore, the optimized FC neural network model was also successfully applied to a database of suction caissons in sand, demonstrating its broad applicability. By comparing three typical DNNs, i.e., FC, Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), it was observed that the FC neural network model excels over others in terms of simplicity, efficiency and accuracy. More importantly, by looking into the model's generalization performance, the FC neural network model can also identify the change in foundation failure mechanisms. This study demonstrates the DNN's powerful mapping ability and its potential for future use in offshore foundation design.

1. Introduction

Suction caissons have been applied successfully in the oil and gas industry for decades as supporting foundations or anchors (Randolph and Gourvenec, 2017; Byrne et al., 2002). Recently, the foundation is also used in offshore wind farms for both bottom-fixed and floating wind turbines. In offshore applications, due to complex environmental actions (i.e., wind, waves, and currents), the substructures of offshore infrastructures are subjected to three-dimensional loads. It is, therefore, critical for foundation design to understand its load–deflection response and develop accurate and efficient design approaches. Traditionally, foundation design is mainly focused on the ultimate bearing capacity. In this regard, the combined foundation capacity under complex vertical (V), horizontal (H) and moment (M) loads are normally represented by a failure envelope (Roscoe, 1956). Extensive studies have been conducted to investigate the failure envelope of the suction caisson foundation in drained sand and undrained clay (Bransby and Randolph, 1998; Bransby and Yun, 2009; Gourvenec and Barnett, 2011; Hung and Kim, 2014; Karapiperis and Gerolymos, 2014; Gerolymos et al., 2015; Vulpe, 2015; Mehravar et al., 2016). Many approximating expressions have been proposed to capture the VHM failure envelopes, as summarized in Table 1. It should be noted that the foundation deflection required to mobilize the bearing capacity is normally very large and exceeds the service limit condition, while the external loads on the offshore wind turbines are relatively small. Instead of the ultimate limit state, the foundation design is normally governed by the stiffness at small deflection (Byrne et al., 2002). Therefore, it is more important to accurately predict the foundation's non-linear load–deflection response.

Traditionally, the macro-element model is usually used to model the non-linear load–deflection response of a foundation under three-dimensional loads (Ibsen et al., 2014; Villalobos Jara, 2006; Byrne, 2000; Pisanò et al., 2016; Skau et al., 2018a; Yin et al., 2020). For example, Houlsby and Cassidy (2002), Zhang et al. (2014) and Wang et al. (2021) proposed macro-element models of spudcan for integrated dynamic analysis of the jack-up system; Salciarini and Tamagnini (2009) and Jin et al. (2019) proposed macro-element models for shallow foundations, and Li et al. (2016) proposed macro-element model for piles. For the suction caisson in undrained clay, Cassidy et al. (2006), Skau

* Corresponding author at: Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, 2628 CN, The Netherlands.
** Corresponding author.
E-mail addresses: yinxilin2020@163.com, hongpeng.zhou@manchester.ac.uk (X. Yin), h.wang-16@tudelft.nl (H. Wang), F.Pisano@tudelft.nl (F. Pisanò), K.G.Gavin@tudelft.nl (K. Gavin), A.Askarinejad@tudelft.nl (A. Askarinejad).

https://doi.org/10.1016/j.oceaneng.2023.115542
Received 7 April 2023; Received in revised form 28 July 2023; Accepted 1 August 2023
Available online 14 August 2023
0029-8018/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Table 1
Summary of existing approximating expressions for the VHM failure envelope.

Reference	Soil type	Proposed equation
Murff (1994)	Sand	$f=\left(\frac{V}{V_u}\right)^2-\left(1+\frac{V_t}{V_u}\right)\frac{V}{V_u}+\frac{V_t}{V_u}+\left[\left(\frac{H}{H_u}\right)^2+\left(\frac{M}{M_u}\right)^2\right]^{0.5}=0$
Bransby and Randolph (1998)	Clay	$f=\left(\left\lvert\frac{V}{V_u}\right\rvert\right)^{2.5}-\left(1-\left\lvert\frac{H}{H_u}\right\rvert\right)^{1/3}\left(1-\left\lvert\frac{M^*}{M_u}\right\rvert\right)+\frac{1}{2}\left(\left\lvert\frac{M^*}{M_u}\right\rvert\right)\left(\left\lvert\frac{H}{H_u}\right\rvert\right)^5$
Taiebat and Carter (2000)	Clay	$f=\left(\frac{V}{V_u}\right)^2+\left[\frac{M}{M_u}\left(1-\alpha_1\frac{HM}{H_u\lvert M\rvert}\right)\right]^2+\left\lvert\left(\frac{H}{H_u}\right)^3\right\rvert-1=0$
Gourvenec and Barnett (2011)	Clay	$f=\left(\frac{h}{h^*}\right)^\alpha+\left(\frac{m}{m^*}\right)^\alpha+2\beta\left(\frac{hm}{h^*m^*}\right)-1=0$
Hung and Kim (2014)	Clay	$f=\left(\frac{M}{M_u}\right)^2-\alpha\beta\lambda\left(\frac{H}{H_u}\right)\left(\frac{M}{M_u}\right)+\left(\frac{H}{H_u}\right)^2+\left(\frac{V}{V_u}\right)^2-1=0$
Karapiperis and Gerolymos (2014)	Clay	$f=\left(\frac{H}{H_u}\right)^2+\left(\frac{H}{H_u}\right)^2+m\left(\frac{H}{H_u}\right)\left(\frac{M}{M_u}\right)-1=0$
Vulpe et al. (2014)	Clay	$f=\left(\left\lvert\frac{h}{h^*}\right\rvert\right)^\alpha+\left(\frac{m}{m^*}\right)^\alpha+2\beta\frac{hm}{h^*m^*}-1=0$
Gerolymos et al. (2015)	Clay	$f=\left(\frac{H}{H_u}\right)^{n_1}+\left(\frac{M}{M_u}\right)^{n_2}+n_3\left(\frac{H}{H_u}\right)\left(\frac{M}{M_u}\right)-1=0$

Where: H_u, M_u, V_u are the ultimate bearing capacity; M^* is the moment calculated about a reference point; α_1 is factor that depends on the soil profile; $h=\left(H/H_{\text{ult}}\right), m=\left(M/M_{\text{ult}}\right)$; h^* and m^* are functions of $v=\left(V/V_{\text{ult}}\right)$; λ is the function of the shape of failure envelope; α, β are the fitting parameters of the embedment ratio; n_1, n_2, n_3 are the coefficients with respect to the embedment ratio

et al. (2018b) and Yin et al. (2020) also developed different macro-element models based on the traditional plasticity theory, hypoplastic theory or the multi-surface concept.

However, the flow rule and the hardening law in these models are strongly dependent on the geometrical configuration of the foundations (e.g., embedment ratio L/D, where D is foundation diameter and L is the foundation embedment length) (Zhang et al., 2014) and the geotechnical properties of the seabed (e.g., the stiffness and strength) (Cremer et al., 2002). Using one macro-element model to capture the non-linear deflection response in all three-dimensional directions is still very challenging. A different set of model parameters may be required for foundations of different geometric configurations and in different seabeds (Skau et al., 2018b). Alternatively, finite element (FE) modelling can explicitly model the soil-foundation system and predict the foundation response under complex loading. Benefiting from the advances in soil constitutive modelling and computational power, it becomes more common to directly model the foundation and soil as continuum bodies in the FE model (Jagota et al., 2013). However, FE analysis requires professional knowledge, making it less preferable in industry design (Szabo and Babuska, 2021; Houlsby, 2016). In addition, the computation efficiency of the three-dimensional (3D) FE modelling cannot satisfy the requirement of industry projects (Qu, 2004). For example, due to the iterative design process of the offshore wind turbine, extensive simulations must be conducted to design a typical foundation. It is unrealistic to perform all these simulations using a 3D FE model (Feng and Shen, 2017). It is, therefore, necessary to develop a model that inherits the accuracy and flexibility of the finite element model but is simpler and more efficient.

Recently, the powerful ability of deep learning (DL) technique to deal with non-linear regression problems has offered an alternative solution for foundation design (Reimers and Requena-Mesa, 2020). The DL technique does not need to pose any pre-assumption (compared with the pre-defined mathematical equations of yield surface, flow rule and hardening law in macro-element models) and has significant flexibility compared with traditional explicit design approaches, e.g., the macro-element model. Benefiting from its powerful non-linear mapping ability, the DL technique has been applied successfully in many geotechnical problems (Nejad and Jaksa, 2017; Shahin, 2014; Momeni et al., 2014; Tarawneh, 2013; Kuo et al., 2009). For instance, Zhang et al. (2020) successfully developed a surrogate model using the long short-term memory (LSTM) model to predict the load–deflection response of the suction caisson foundation in sand. However, it should be noted that the study was limited to the foundation behaviour in the two-dimensional H-M space.

In light of the above premises, this study aims to develop a DL-based design model to predict the non-linear response of suction caissons under three-dimensional loads. As the first exploratory study in this area, this paper is limited to the suction caissons in undrained clay under combined loads in the same plane (i.e., three degrees of vertical, horizontal and rotation movement in the same plane). A following study is undergoing to develop a DL-based model for caissons in sand and layered soil. A series of three-dimensional (3D) finite element simulations were performed first to provide the training database for the DL model. The response of the suction caissons with a wide range of aspect ratio (L/D) from 0.1 to 1 in both homogeneous (i.e., s_u is constant) and heterogeneous soil (i.e., s_u is linear with depth) was studied. For each foundation in each soil profile, 96 different displacement loading paths were simulated to obtain the foundation response in three-dimensional VHM space. The numerically generated data was then used to train the deep neural network (DNN) model, which could directly learn the mapping relationship between the foundation deflection and external loads from the raw numerical data. In this study, the FC neural network model is adopted to develop the surrogate foundation model. The developed FC neural network model with the optimized hyper-parameters was further applied to the database of suction caissons in sand (Zhang et al., 2020). Meanwhile, the robustness and generalizability of the trained DL-based surrogate model are thoroughly discussed. In particular, the evolution of the foundation failure mechanism was investigated by looking into the model generalization performance. An extra example application of this trained surrogate model is also provided to demonstrate the computational efficiency of the model. In the end, several typical neural network models, i.e., convolutional neural network (CNN) and LSTM models, are also compared to further evaluate the performance and applicability of the FC neural network model for suction caissons in clay.

2. Model training database

2.1. Numerical modelling

The finite element software Abaqus 6.14 was used in this study to simulate the behaviour of suction caisson foundations installed in clay seabed under combined loads (Systèmes, 2014). The influence of the geometrical configurations of the foundation and the properties of the soil on the shape and size of the envelope were systematically investigated. The suction caissons are thin-walled large-diameter steel cylinders, open-ended at the bottom and closed at the top, typically less than 20 m in foundation diameter (D) with an aspect ratio (L/D, L is the foundation embedment depth) typically less than 1 (Cassidy et al., 2006; Fu et al., 2020) and thickness ratio (D/t) ranging between 80–300 (Gourvenec and Cassidy, 2005). In this study, a fixed foundation diameter and wall thickness of 10 m and 0.1 m were adopted in all simulations, respectively. As all the computed results will be analysed after normalization with the foundation dimensions and soil untrained strength, it is believed the absolute value of the diameter will not affect the results (Gourvenec and Barnett, 2011). A total of 10 embedment depth-to-diameter ratios of $L/D = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]$

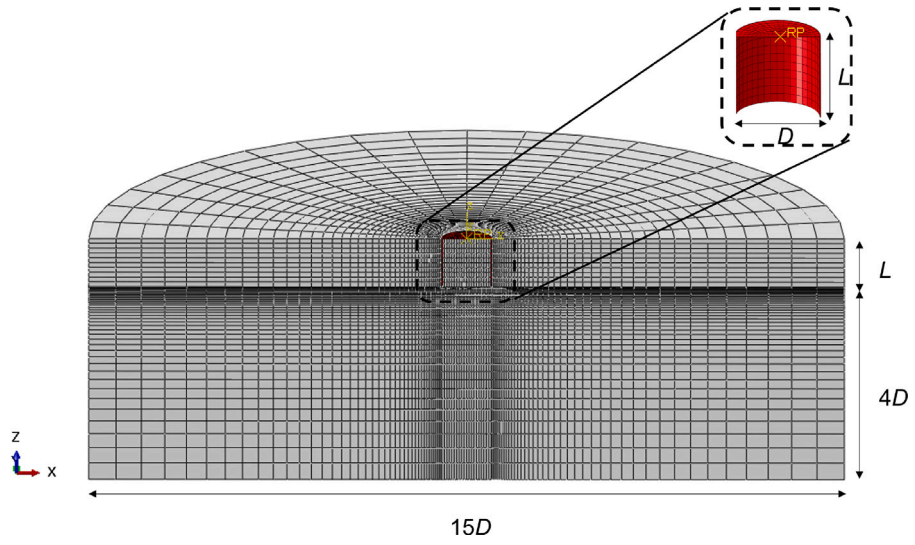


Fig. 1. A typical FE model mesh of caisson with a diameter of 10 m and aspect ratio of 1.

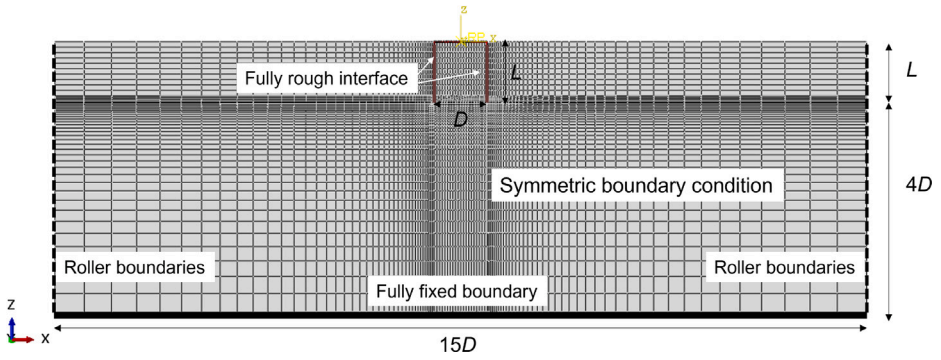


Fig. 2. Boundary conditions of the FE model.

were studied to cover the suction caisson foundations used in the field. The foundations were modelled through the well-established wished-in-place approach without considering the installation effect.

A typical mesh of the suction caisson FE model with a diameter of 10 m and an embedment depth of 10 m is presented in Fig. 1. By taking advantage of the symmetry of the problem, only half of the soil-foundation system was modelled to save computation time. Roller supports were applied around the mesh circumference while the base boundary was fully fixed, illustrated in Fig. 2. The model diameter and model depths are $15D$ and $4D$ beneath the foundation base, respectively. A fine mesh domain was constructed around the foundation skirts, while a coarser mesh domain was used in the far field to reduce computing expenditure. An additional simulation with a model twice the dimension and mesh density results in a change of 1% regarding the load–deflection response. A fully rough interface with no separation between the foundation and the soil was used in all simulations. The suction caisson and clay soil were modelled with eight-node linear strain brick elements with reduced integration (i.e., 'C3D8R' in Abaqus terminology) and hybrid eight-node linear strain brick elements (i.e., 'C3D8H' in Abaqus terminology) (Gourvenec and Barnett, 2011), respectively.

A linearly elastic and perfectly plastic material, obeying the Tresca yield criterion, was assumed for clay. The elastic response of clay is defined by the shear modulus and Poisson's ratio. The soil modulus G of clay is defined from the undrained strength (s_u) and equal to $500s_u$ (Hu and Randolph, 1998; Jeanjean et al., 2017). The effective unit weight of the soil was $\gamma' = 6$ kN/m^3. a Poisson's ratio ν of 0.495 and a dilation angle of 0.01° was used to simulate the undrained loading conditions. Two different shear strength profiles, i.e., a homogeneous shear strength ($s_u = 10$ kPa, simulating over-consolidated clay) and a linear increasing shear strength ($s_u = kz$ kPa, where z is depth below the ground surface and k is the strength increasing gradient with depth, simulating a normal-consolidated clay) were studied for each suction caisson. A fully elastic response was assumed for the suction caisson foundations. Young's modulus E and Poisson's ratio ν of steel were used for the foundations. Details of all the mechanical properties in FE modelling are summarized in Table 2.

In DNN model training, the embedment ratio L/D was used instead of the absolute values of D and L to accommodate as many combinations of feasible embedment depths as possible while ignoring the impact of the foundation dimensions. The chosen range of the L/D ratio was partitioned into 10 equal parts from 0 to 1, while embedment ratios larger than 1 and no-skirt foundations (embedment ratio = 0) were not taken into consideration. The same simulations were performed in both homogeneous and heterogeneous soil, respectively. The response force is continuously mobilized with increasing displacement and rotation in each direction until reaching the bearing capacity. As recommended by Butterfield et al. (1997), the sign system for displacements and loads described in this study uses right-handed axes and clockwise positive signs, as illustrated in Fig. 3.

To better illustrate the 3D loading direction, a spherical coordinate system is employed, which provides a better representation of the ellipsoidal surface compared to the traditional Cartesian coordinate system. After determining the three axes of the ellipsoid, each loading direction can be defined with only two parameters θ and ϕ, where θ is the positive angle to the x-axis when rotating counterclockwise and ϕ

Table 2
Mechanical properties in FE modelling.

Suction caisson	Outer diameter (D)	10 m
	Length (L)	1–10 m
Clay	Angle of internal friction (ϕ')	0.01°
	Angle of dilation (ψ)	0.01°
	Shear modulus to undrained shear strength ratio (G/s_u)	500
	Poisson's ratio (v_s)	0.495
	Shear strength in OC clay (s_u)	10 kPa
	Shear strength at mudline (s_{um})	0.01 kPa
	Strength increasing gradient with depth (k)	1
	Degree of heterogeneity (κ)	1000
Steel	Young's modulus (E_s)	$210 * 10^6$ kPa
	Poisson's ratio (v_s)	0.25

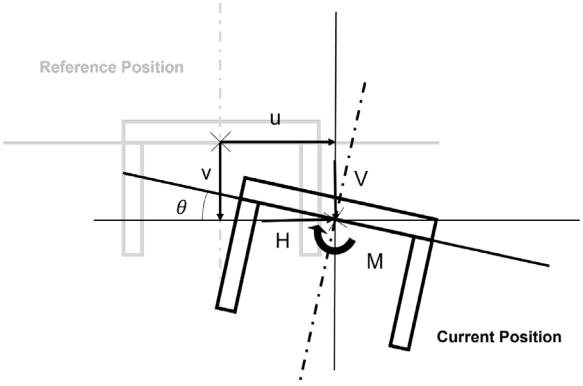


Fig. 3. Sign convention for loads and displacements.

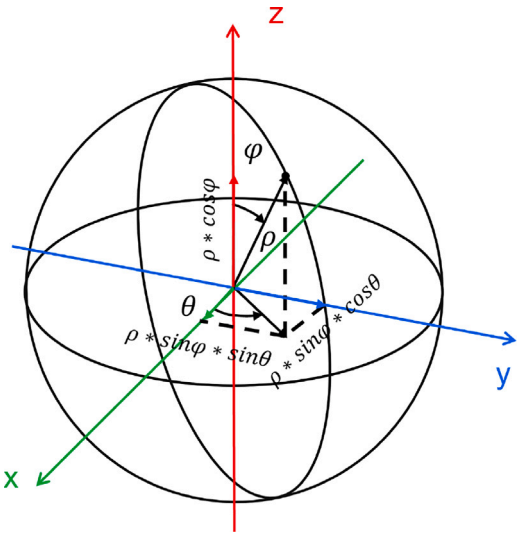


Fig. 4. Spherical coordinate system.

is the positive angle to the z-axis when rotating clockwise. For a better understanding, θ can be considered as the longitude of the earth and ϕ can be considered as the latitude of the earth (as shown in Fig. 4). In this study, only the compressing bearing capacity of the shallow foundation was investigated, not its pull-out resistance. After weighing the computational time against the size of the data set, 96 probe test directions were established for each embedment depth. This means that θ takes 15°, 45°, 75°, 105°, 135°, 165°, 195°, 225°, 255°, 285°, 315°, 345°, respectively (30-degree increments for a total of 12 angles), while ϕ takes 15°, 30°, 60°, 75°, 80°, 85°, 88°, 90°, respectively (a total of 8 angles). θ and ϕ should be equally spaced to cover the entire computation range. It should be noted, however, when the spherical coordinate system is converted to the Cartesian coordinate system, the z ($\rho * \cos\phi$) values are distributed unevenly. To better capture the failure pattern at smaller vertical displacements, the distribution of ϕ will change from sparse to dense after 80°.

The total 96 directions investigated in this study were plotted as a hemispheric envelope in Fig. 5. In each simulation, it was found that the first 100 data points already adequately represent the entire loading path with the remaining data points exhibiting a negligible variation. Therefore, the first 100 data points were intercepted in each direction, generating a total of 9600 data points for each foundation. Each data point set includes displacement information in three directions (v, u, θ), corresponding to three load components (V, H, M), a foundation configuration (L/D) and a degree of heterogeneity input (κ). The displacement information of the skirted foundation was written directly in the comma-separated values (CSV) files without any processing. The force information of the skirted foundation is dimensionless and can be represented by the bearing capacity factor ($N_{cV} = V/As_{u0}$, $N_{cH} = H/As_{u0}$, $N_{cM} = M/ADs_{u0}$, $A = \pi D^2/4$ is the foundation area). The foundation configuration information was represented by the embedment ratio (L/D), and the shear strength profile was represented by the soil strength heterogeneity index (κ). All data sets were saved in CSV files for quick access by Python. These CSV files contain a total

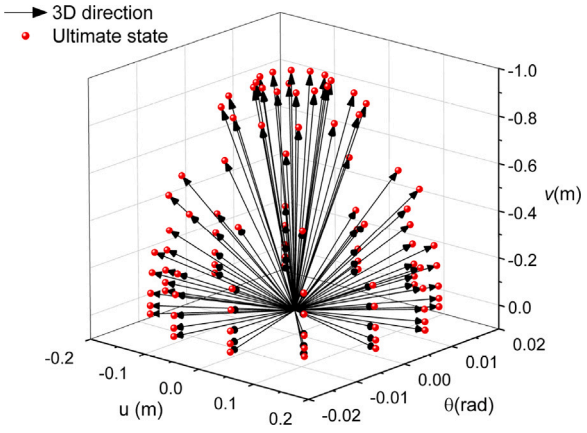
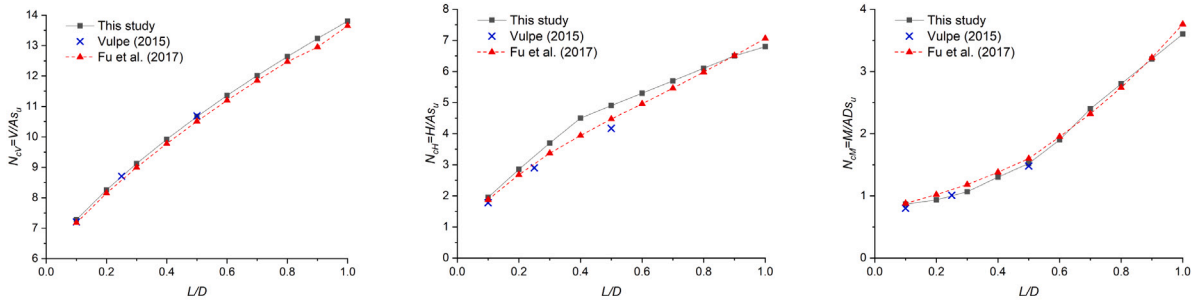


Fig. 5. Ultimate displacement probes in 96 directions (created by 8 latitude and 12 longitude combinations in a spherical coordinate system).

of 1920 sets of FE simulations, which took about 640 h to build. All the simulations and data post-processing were conducted automatically using a Python script. Both the FE simulation data and the Python script are available from the corresponding author upon request.

2.2. FE model validation

The FE model was first validated by comparing the computed ultimate bearing capacity with those in the literature. Fig. 6 shows



(a) Validation of vertical load (b) Validation of horizontal load (c) Validation of moment

Fig. 6. Ultimate bearing capacity as a function of embedment ratio.

the variation of three bearing capacity factors (i.e., N_{cV}, N_{cH}, N_{cM}) against the foundation embedment ratio L/D. The calculated bearing capacities in this paper and the analytical and numerical solutions calculated by Vulpe (2015) and Fu et al. (2017) are also shown in Fig. 6. It is clear that the computed results are consistent with the results from existing studies. The maximum relative differences in vertical, horizontal, and moment bearing capacities are less than 2.1%, 14.2%, and 9.6%, respectively, compared with the calculated from the proposed equations in Fu et al. (2017).

Fig. 7 presents the typical three-dimensional VHM failure envelopes of foundations in homogeneous soil as three 2-dimensional envelopes (i.e., H-M, V-H, V-M), respectively. As shown in the figure, both the size and shape of failure envelopes vary with the foundation embedment ratio. In particular, an elliptical shape of the failure envelope in the H-M space is observed, which is also well reported in existing studies (Bransby and Randolph, 1998; Bransby and Yun, 2009). The consistency of the computed results in this study with those reported in the literature demonstrated the reliability of the FE model in this study.

2.3. Data pre-processing

As the foundation displacement and rotation values are not dimensionless, different scales of the input and output parameters will affect the model performance. Therefore, the following normalizing equation was adopted to eliminate the size effect:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \left(\bar{x}_{\max} - \bar{x}_{\min}\right) + \bar{x}_{\min} \tag{1}$$

where $x_{\max}$ and $x_{\min}$ are the maximum and minimum of the parameter x; $\bar{x}_{\min}$ is the threshold to be scaled and $\bar{x}_{\max}$ is the upper boundary to be scaled. The upper boundary is typically set to 1, and the lower boundary is set to −1 in this paper.

This normalization operation was implemented through the Keras framework, an advanced deep learning library developed based on the Python programming language (Gulli and Pal, 2017). The subsequent programming and construction of DNN models are also based on the Keras framework. After normalization, the complete data is separated into three sub-datasets, consisting of 64% training data, 16% validation data, and 20% test data (Géron, 2022). Specifically, the DNN model will be trained on the training set, and the 20% validation dataset will be used to supervise the training process. Importantly, the 20% test dataset will not be used for training and will only be utilized for the purpose of evaluating the trained model. The research framework adopted to train a DNN model is illustrated in Fig. 8. The deep learning-based methodology will be described in more detail in Section 3.

3. Deep learning based methodology

3.1. Deep neural networks

In this paper, the fully-connected neural network is adopted (Hopfield, 1982), which is stacked by several adjacent layers (e.g., Fig. 9). All layers are composed of neurons to impart information. Specifically, layer 0 is referred to as the input layer, while the final layer and the intermediate layers are referred to as the output layer and the hidden layers, respectively. According to the universal approximation theorem (Hornik et al., 1989), a neural network can be treated as an ideal "universal" function to tackle the non-linear regression problem by approximating any input and output dataset. However, designing an appropriate model structure and learning model parameters poses a non-trivial challenge. The model's structure inherently plays a crucial role in determining its fitting ability. If the model is under-parameterized, it may lead to underfitting, whereas an over-parameterized model may give rise to the overfitting issue (Zhou et al., 2022; Brutzkus et al., 2017). Therefore, appropriate hyper-parameters related to model structure design (i.e., the number of hidden layers, and the number of neurons in each layer) should be selected. Besides, the process of learning model parameters in a neural network model is highly sensitive to the selection of training hyper-parameters (e.g., learning rate, batch size, training iterations), which can significantly impact the training process. The strategy for selecting hyper-parameters in model structure design and network training will be elaborated in the next section.

3.2. Hyper-parameter tuning and training details

The performance (e.g., convergence, training efficiency) of a model is highly dependent on the setting of the hyperparameters, including both model structure hyperparameters and model training hyperparameters. The model structure hyperparameters (e.g., the number of neurons in each layer) define the size and complexity of neural networks. The model training hyperparameters (e.g., learning rate) determine the prediction accuracy and the model convergence. A few examples of the hyperparameters are also illustrated in Fig. 10.

3.2.1. Hyper-parameters in model structure design

The model structure hyper-parameters are determined using the grid search method (Pontes et al., 2016), which is a classical and efficient method for hyperparameter screening. After preliminary experiments, it was found that a FC neural network model with only a single hidden layer was able to obtain good prediction performance. However, to further improve prediction accuracy, especially the stability and robustness of the prediction, more complex neural networks by increasing the number of hidden layers were investigated. In addition, as explained in Section 3.1, the balance between model complexity and model accuracy should also be considered during model training. Considering the factors above, a FC neural network model with two hidden layers is adopted in the end. In typical neural network models, to capture the non-linear response, the non-linear activation function will be introduced after each hidden layer. The typical activation functions include the sigmoid, rectified linear unit (ReLU), and hyperbolic tangent function (tanh) (Sharma et al., 2017). In this paper,

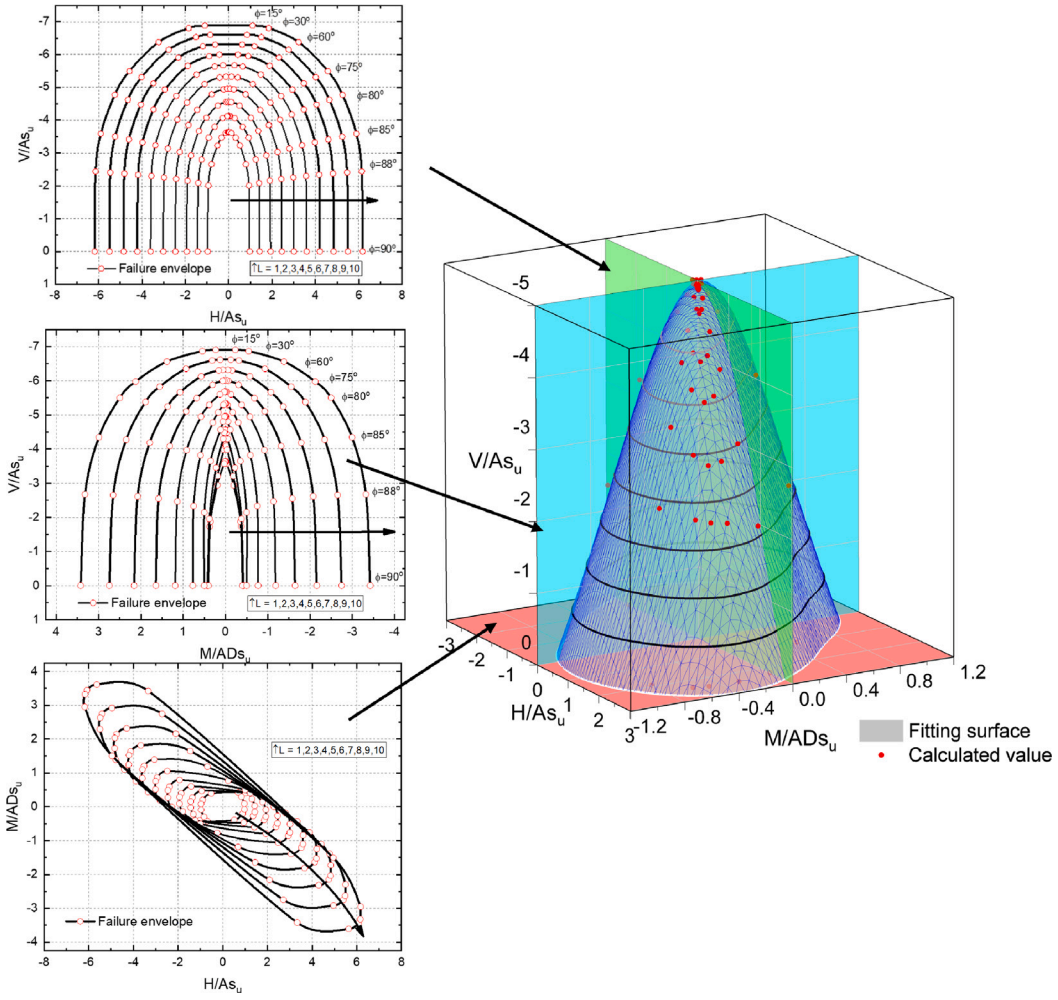


Fig. 7. 3D failure envelopes at L = 1–10 m calculated using FE analysis with slices (H-M slice in red, H-V slice in green and M-V slice in blue).

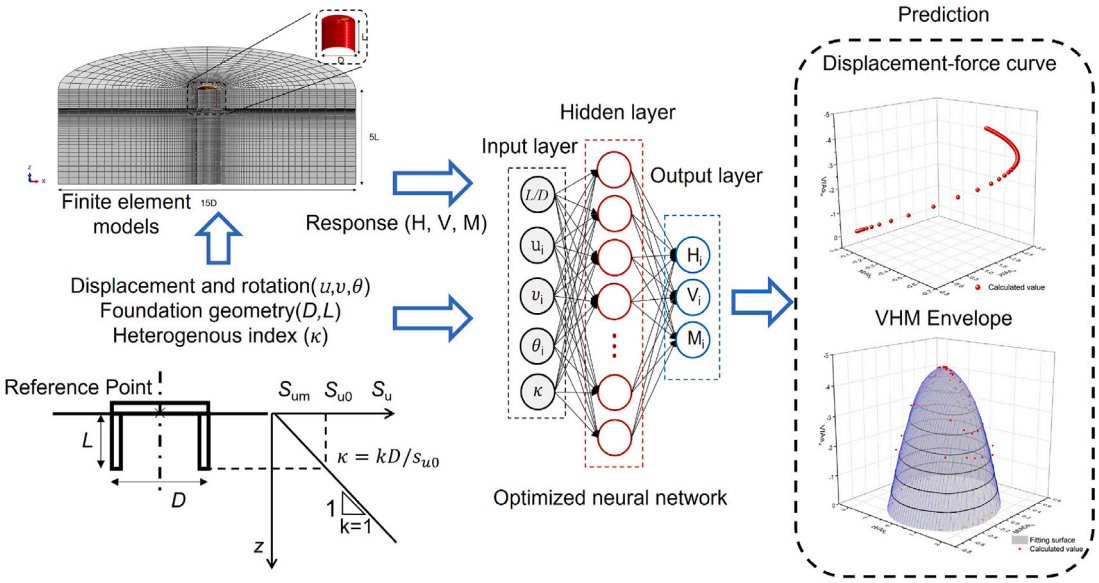


Fig. 8. Schematic process of hybrid FE modelling and deep learning algorithm surrogate modelling.

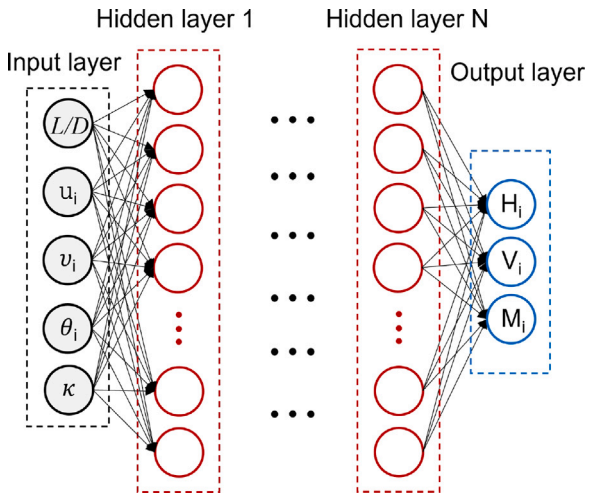


Fig. 9. Fully-connected neural network.

the ReLU (Nair and Hinton, 2010) activation function was adopted to avoid the gradient vanishing problem and accelerate the gradient descent convergence (Li and Yuan, 2017). Meanwhile, to maximize the performance of parallel computing on the GPU, the batchsize is often required to be a multiple of 8 (e.g., 32, 128) (Sanders and Kandrot, 2010). This paper investigates the influence of batchsize with different numbers, i.e., 128, 256, 512, 1024, and 2048. The selection of the number of neurons (i.e., 16, 32, 64, 128, and 256) in each layer is also analysed.

By selecting a different batchsize and the number of neurons, a total of 125 ($5 * 5 * 5$) sets of trials were implemented. Fig. 11 presents the minimum mean squared errors generated in each trail. Obviously, the larger the number of neurons and the smaller the batch size, the better the performance of the model. However, it should be noted that the continuous increase in the number of neurons does not contribute much to improving the prediction accuracy of the model. In addition, a small batch size may also reduce the generalization ability by trapping the model in a local optimum (Keskar et al., 2016). Therefore, by analysing the performance of the model in Fig. 11, the optimal batch size and the number of neurons in each layer were set as 128 and 256, respectively.

3.2.2. Hyper-parameters in training process

This section focuses on choosing the loss function and optimizer, as well as determining the learning rate and training epoch. The loss function in the regression problem is mainly divided into mean absolute error (MAE) loss and mean square error (MSE) loss, where MSE is more stable and accurate in the optimization process (Goodfellow et al., 2016). MSE calculates the mean of the squared discrepancies between the prediction and the target value. Larger errors could be punished more severely than smaller ones by square-rooting the error. Therefore, the MSE loss is used in this paper to produce a more precise result. Adam algorithm (adaptive moment estimation algorithm) is selected as the optimizer (Kingma and Ba, 2014), which is a combination of the momentum method and the RMSprop algorithm (an adaptive learning rate approach (Tieleman et al., 2012)). The Adam algorithm uses momentum as the direction of parameter update and alters the learning rate adaptively. A comparison experiment was performed to determine the initial learning rate. Fig. 12 compares the training process at different learning rates. It should be noted that the value of the training loss remains approximately stable at all learning rates and converges to a constant value within 50 epochs. The model could converge rapidly (less than 5 epochs) almost for all learning rates except for the learning rate of 0.0001. However, at learning rates of 0.01 and 0.005, the training process shows fluctuations in loss value.

Table 3
Hyper-parameters used for training the neural networks model.

	FC neural network	
Hyperparameter	Description	Value
Nh	Number of hidden layers	2
Nn	Number of nodes in the hidden layer	256,256
η	Learning rate in the optimizer	0.001
Batch_size	Number of training samples utilized in one iteration	128
Epoch	Number of iterations during training	50
validation_split	Proportion of validation set in total training	0.2

The loss curve could remain stable at learning rates of 0.001 and 0.002. After comparing with other learning rates, it was found that a learning rate of 0.001 can make the training process smoothly converge to a small loss faster. Therefore, the learning rate in model training was set to 0.001. The final hyperparameters utilized for FC neural network training are summarized in Table 3.

3.3. Evaluation metrics

In this study, the RMSE (Root Mean Squared Error) and R^2 (Coefficient of determination) are used as the primary evaluation metrics of the regression results, while the MAE (Mean Absolute Error) serves as the supplementary evaluation metric. RMSE is the root of MSE (Mean Squared Error) and can give the most intuitive prediction error. It is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)^2} \tag{2}$$

The goodness of fit measures the degree to which a regression line matches the observed values. The determination coefficient (R^2) is the statistical indicator of the goodness of fit. The greater the R^2, the better the model predicts the results, which is optimal at a value of 1. In contrast, when R^2 is closer to 0, the model fits the data less well. The R^2 could be calculated as follows:

$$R^2 = 1 - \frac{\sum_i \left(\hat{y}_i - y_i\right)^2}{\sum_i \left(\bar{y}_i - y_i\right)^2} \tag{3}$$

Considering that certain data points are 0, the Mean Absolute Percentage Error (MAPE) calculation would be biased. This metric divides the difference between the actual value and predicted value by the actual value. Consequently, the MAE is employed to assess the performance and describe errors between the predicted value and actual value:

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}\left|\left(y_i - \hat{y}_i\right)\right| \tag{4}$$

4. Model development and evaluation

4.1. Model prediction and applicability

In this section, the mechanical response prediction experiment was carried out first to verify if the FC neural network model can provide accurate predictions and accommodate the change in the soil profile. A total of 192000 data sets of homogeneous and heterogeneous soils (from Section 2) were fed to the FC neural network model. The soil strength heterogeneity index was introduced into the model input to differentiate between the two types of soils. As shown in Fig. 13, the three load components can be perfectly predicted by the trained FC neural network model. It demonstrates that even a very "shallow" FC neural network model of two layers can learn the intrinsic failure mechanisms of the caissons from raw data and effectively predict their non-linear mechanical responses under complex loading.

Similarly, Zhang et al. (2020) also successfully used deep learning algorithms to predict the behaviour of caisson in sand soils, although

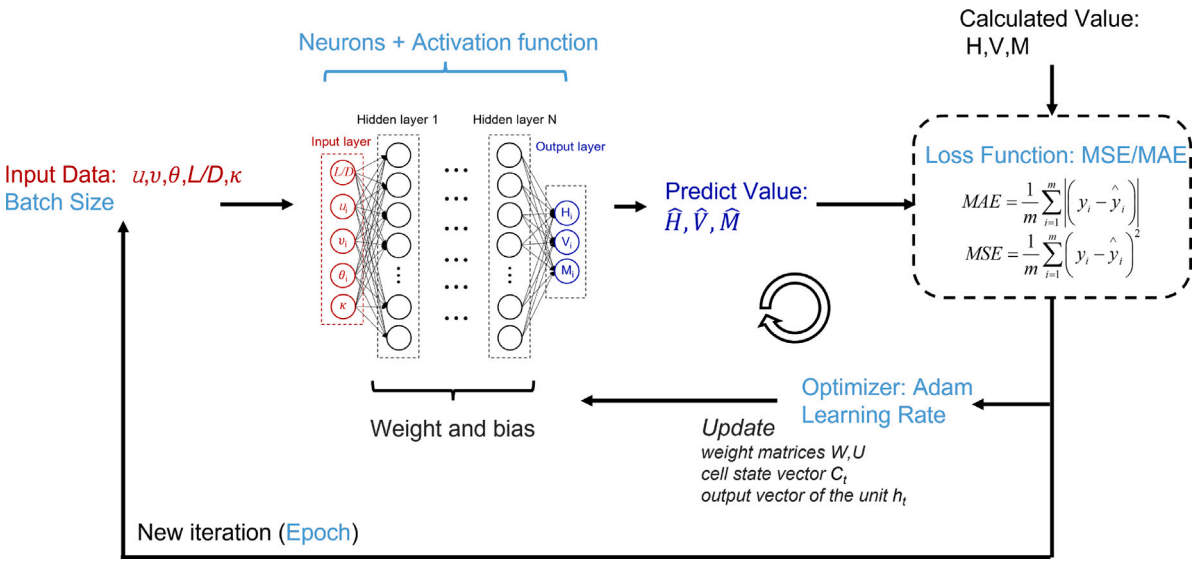


Fig. 10. DNN model's training process with hyperparameters marked in light blue, input parameters marked in red and output parameters marked in blue.

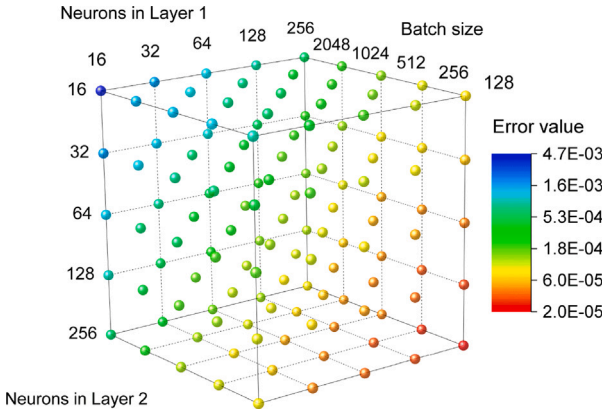


Fig. 11. The minimum loss of the model with different combinations of batchsize and neurons.

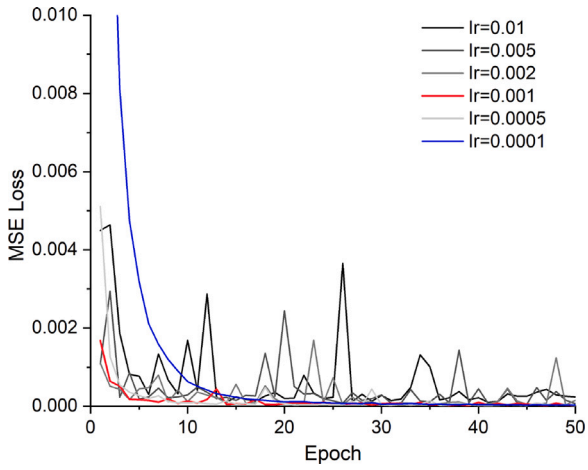
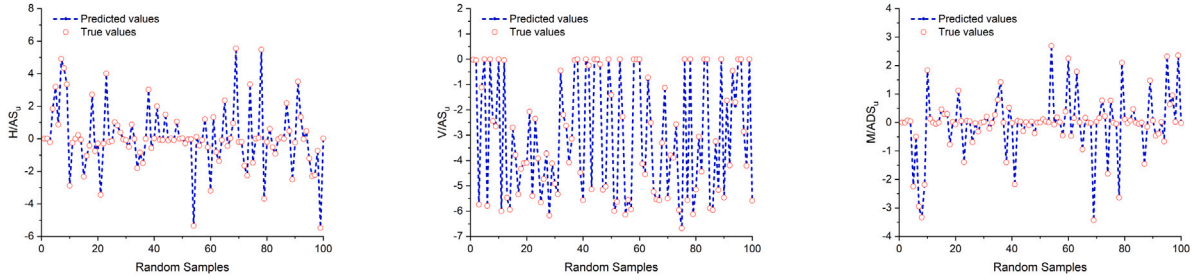


Fig. 12. Training loss for the DNN model with different learning rates.

the study was limited to the response in $H-M$ space. The training data was generated from advanced numerical modelling using smoothed particle hydrodynamics with the SIMSAND model. The database was validated by laboratory tests, physical model tests, and a caisson foundation field test. Instead of the FC neural network model, the model adopted in Zhang et al. (2020) was developed based on the LSTM, which could obtain a better prediction result in the time-series domain by incorporating history information (Hochreiter and Schmidhuber, 1997). To further test the application ability of the FC neural network model, the FC model adopted in this study is also tested against the database in Zhang et al. (2020). It should be noted that the training parameters of the FC neural network model (e.g., number of epoch and batch size) are changed to ensure that the training strategy is the same as in Zhang et al. (2020), while the same model hyperparameters as described in Section 3.2.1 were used. The FC neural network model was then retrained, using the training set, validation set and test set with the same data partitioning strategy as in Zhang et al. (2020). Despite using an identical training strategy, it should be noted that the trained model can be considered as a new model with the same model structure but different internal parameters. These weight and bias parameters stored in every neuron are fully renovated due to the input of a completely different dataset. This also explains why the same FC neural network model can fit two different soil response mechanisms. The training results are summarized in Table 4. As shown in the table, better prediction results were obtained using the FC neural network model. This is mainly attributed to its much simpler structure which is easy to train and the better prediction strategy. The essence of the prediction is aiming to reproduce every possible loading path in reality. When using different prediction strategies, the loading path can be considered either as a continuous line (temporal prediction) or as finite individual points (non-linear regression). While the LSTM model in Zhang et al. (2020) follows the temporal prediction strategy, the FC neural network in this study implements the nonlinear regression strategy and exhibits better performance. Therefore, although the LSTM could be very effective in dealing with complex temporal problems, its unique gating mechanism did not perform well in simple point-to-point regression problems (Hiransha et al., 2018). Instead, too much memory in LSTM inevitably leads to more parameters and a more complex model. This conclusion will be further elaborated in Section 4.4. In addition, it was observed that the prediction accuracy in sand decreased slightly compared to the predicted results in the clay dataset generated in Section 2. This is because the foundation response data in sand calculated using the smoothed particle hydrodynamics with the SIMSAND model is very noisy with relatively large fluctuation. As the MSE loss function will be affected by the noise, a prediction bias can be produced. Based on this extra test on the database of foundation in sand, it is clear that the FC neural network model is equipped with outstanding applicability and has excellent transfer learning capabilities.



(a) Prediction of horizontal load (b) Prediction of vertical load (c) Prediction of moment

Fig. 13. Prediction results for 100 random samples on H, V, and M, respectively, in homogeneous and heterogenous soil (RMSE: 0.024, R^2: 1.000, MAE: 0.016).

Table 4
Comparison of prediction errors of two models on sandy dataset.

Parameter	LSTM in Zhang et al. (2020)		FC neural network model	
	MAPE	NSE	MAPE	NSE
H	25.07%	0.93	20.20%	0.99
M	42.15%	0.92	37.16%	0.99

4.2. Model robustness

Fig. 13 has demonstrated that the FC neural network model can achieve excellent predictions. However, it should be noted that good predictions can come from the fortuitous selection of weights and biases, which helps model training and makes convergence very fast. In addition, it is also possible that the data were divided by chance, which put the challenging data in the training set, making the left data in the test set easily to be predicted. Both these issues are related to the model setting of random number seeds, which symbolize the randomness of the model. A random seed is a number used to initialize a pseudorandom number generator. The randomly generated number controlled by seed influences not just the model's initialization weights and bias parameters, but also the partitioning of the test and validation sets. Therefore, to evaluate whether the model has stable performance, the robustness of the model was investigated by repeating experiments with various random seeds (Madhyastha and Jain, 2019). The stability of the model was then studied (Blundell et al., 2015). For the tested 50 random number seeds, the corresponding MSE loss values for model training are shown below (Fig. 14). The blue line represents the mean of the fifty training runs, while the red line represents the standard deviation of each epoch. As shown in the figure, the training process converges rapidly and stabilizes after the fifth epoch. The error line (indicating the variance) reaches the highest at the twenty-fifth epoch, and the loss value at this moment swings by 0.0004, which is within an acceptable range. The box plots of H, V, and M predictions from these fifty experiments are represented by RMSE and R^2 (shown in Fig. 15). Clearly, the model can maintain a high prediction accuracy with the RMSE of predictions for the three forces being less than 0.012 and R^2 remaining above 0.9985. The prediction error distribution of the 50 replicate experiments demonstrates the high stability and robustness of the FC neural network model.

4.3. Model generalization ability

This study aims to train the model with existing data and produce a DNN model capable of predicting the response of any new embedment depth foundation under a unique combination of loads. It necessitates that the trained DNN model not only can predict with high accuracy and robustness but also has the ability to generalize. The generalization of the model can be demonstrated by mapping the failure envelope with a given embedment depth and soil strength profile only. The caisson foundation with $L = 3$ m in homogeneous soil is used as a showcase in

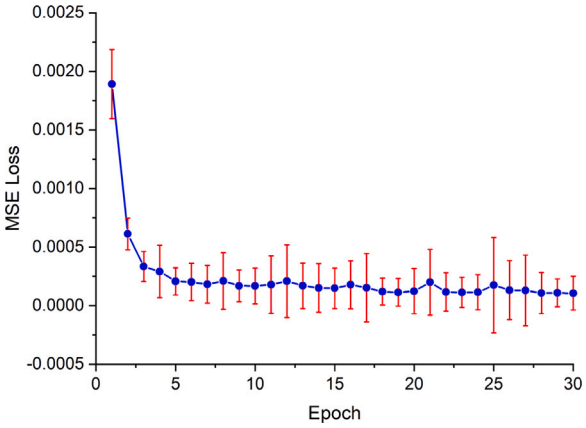
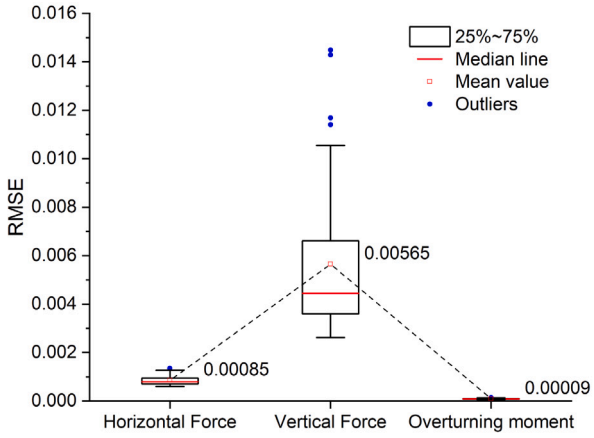


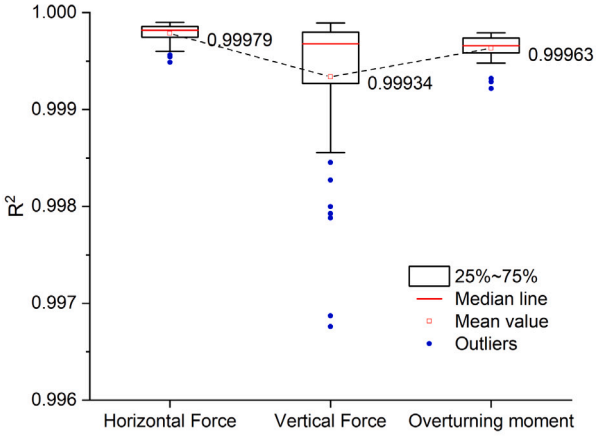
Fig. 14. The mean MSE loss under 50 random seeds.

Fig. 16. To better simulate the conditions in practice, data with $L = 3$ m is excluded from the training set of the model. The data points on each loading path were not intercepted to obtain the complete loading process. A total of 109,032 data points were obtained for the remaining nine embedment depths and were divided into 80% training set and 20% test set. After training, the interpolation prediction error on the test set was RMSE = 0.028, $R^2 = 1.000$, MAE = 0.018.

Once the model training was finished, the foundation mechanical response can be obtained by simply inputting the direction, displacement, or rotation. For the trained model, it is very free to set the case number of loading directions, for example, the 3000 used in this case study or even more. In fact, it will not cause any significant difference in the computation time. Following the spherical coordinate calculating equation, the displacement/rotation values at 3000 directions were determined and inputted into the trained model. The capacity envelope surface defined from the 3000 data from the trained model was compared with those obtained from 3D FE simulations of 96 directions in Fig. 16. The grey surface in the figure is the predicted envelope fitted by 3000 points; the true values essentially fall on the surface, indicating that the predictions are very accurate. The generalization (i.e., extrapolation test) errors at $L = 3$ m are RMSE = 0.025, $R^2 = 0.998$, MAE = 0.019. The accurate prediction demonstrates the excellent generalization ability of the FC model, which can predict the foundation behaviour through limited FE simulation data and avoid complex FE modelling. More importantly, the DL-based surrogate model can save significant computation time compared to the traditional FE method. Specifically, if this envelope with 3000 loading paths was simulated in a normal computer (Legion Y7000P), it would take more than 42 days. On the contrary, when using the DNN model, it takes only one second. In summary, the DNN-based surrogate model is more precise, adaptable, and efficient than the macro element model and 3D FE modelling.



(a) RMSE for H, V, M predictions



(b) R^2 for H, V, M predictions

Fig. 15. Distributions of H, V, M prediction error under 50 random seeds.

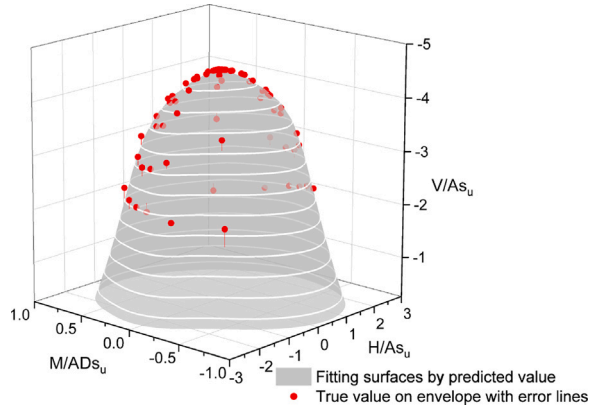


Fig. 16. Comparison of predicted envelope and true value at $L = 3$ m.

From a geotechnical perspective, the failure mechanism of the foundation varies with the embedment depth. As a result, the generalization ability at different embedment depths is very likely to vary as well. The foundation's failure mechanism should be continuous over the range of embedment depths, i.e., the failure mechanism at a specific embedment depth is close to the failure mechanisms at the two adjacent embedment

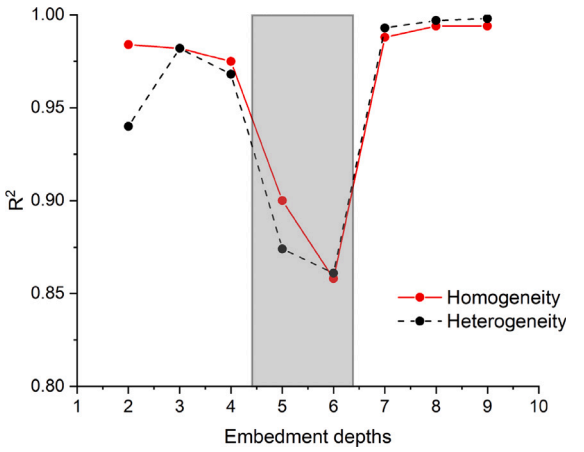
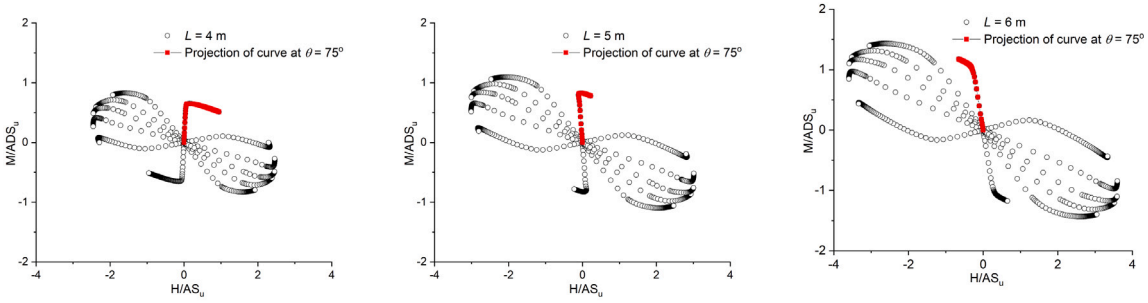


Fig. 17. Variation of R^2 at different embedment depths in homogeneous and heterogeneous soils.

depths. To reveal the evolution of the failure mechanism with foundation embedment depth, another experiment was designed to train the DNN model by utilizing data from two neighbouring embedment depths and predict the response at a specific depth in between (e.g., train the neural network using data from $L = 1$ m and $L = 3$ m and then predict the response at $L = 2$ m). Both homogeneous and heterogeneous soils were tested, and the results are shown in Fig. 17.

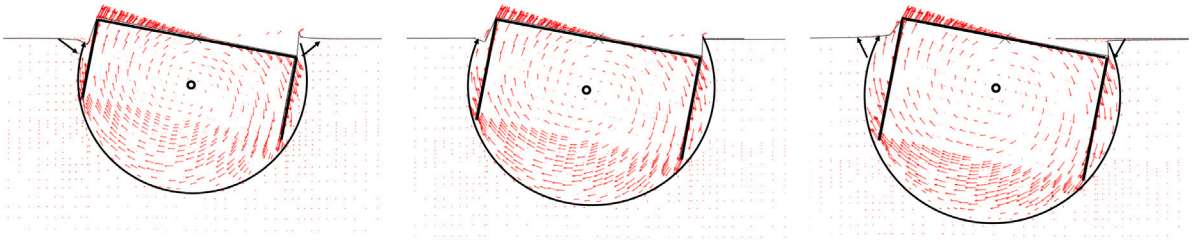
From Fig. 17, it is clear that the heterogeneous and homogeneous soils exhibit the same pattern of variation, with one poorly fitted position appearing at the embedment depth of around $5 - 6$ m. The excellent fit at an embedment depth of 4 m suggests that embedment depths of 3 m, 4 m and 5 m share the same failure mechanism. Similarly, the excellent predictions for embedment depths of 7 m indicate the same failure mechanism for embedment depths of 6 m, 7 m and 8 m. On the contrary, the inferior prediction between $L = 5$ m and $L = 6$ m indicates a transition point (i.e., a switch of the failure mechanism) between these two embedment depths. This clear change in the neural network's generalization ability highlights the corresponding variation of the foundation failure mechanism.

To further highlight the changes in mechanism, the combined H-M response of the foundations in different directions (Fig. 18) was investigated in more detail. As shown in Fig. 18, a clear transition of force path was identified between $L = 5$ m and $L = 6$ m for a loading direction θ of 75°. At $L = 4$ m, the $\theta = 75°$ loading path develops in a positive direction along the horizontal axis. In addition, the shape of the loading path changes significantly, and the foundation response is controlled by the scoop failure mechanism. As the load increases, the ultimate moment-bearing capacity is first reached. Then, the force path gradually develops towards the ultimate value of the horizontal force along the positive axis. At $L = 5$ m, the loading path is almost perpendicular to the transverse axis. After reaching the ultimate state of the moment, the path develops slightly along the positive horizontal axis. In contrast, for the foundation with $L = 6$ m, its loading path shifts along the negative direction of the transverse axis and keeps developing towards the negative ultimate horizontal bearing capacity after reaching the failure envelope. As the embedment depth steadily increases, the wedge caused by sliding vanishes at the mudline, and the displacement vector shifts from the right (at $L = 4$ m in Fig. 19(a)) to the left (at $L = 6$ m in Fig. 19(c)). Therefore, the shift in the general direction of the loading paths from the change of the failure mechanisms makes it more difficult for neural networks to predict. These observations suggest that the effect of the skirt geometry on the suction caisson failure mechanisms can be detected from the fluctuation of generalization ability. Therefore, the DNN model does not only capture the relationship between inputs and outputs adaptively but also mine the intrinsic patterns of data through the generalization ability.



(a) Data in H-M envelope at $L = 4$ m and the path in $\theta = 75°$ is marked by the red dot (b) Data in H-M envelope at $L = 5$ m and the path in $\theta = 75°$ is marked by the red dot (c) Data in H-M envelope at $L = 6$ m and the path in $\theta = 75°$ is marked by the red dot

Fig. 18. Comparison of data at three embedment depths.



(a) Mechanism in $\theta = 75°$ at $L = 4$ m (b) Mechanism in $\theta = 75°$ at $L = 5$ m (c) Mechanism in $\theta = 75°$ at $L = 6$ m

Fig. 19. Comparison of vector contour at three embedment depths.

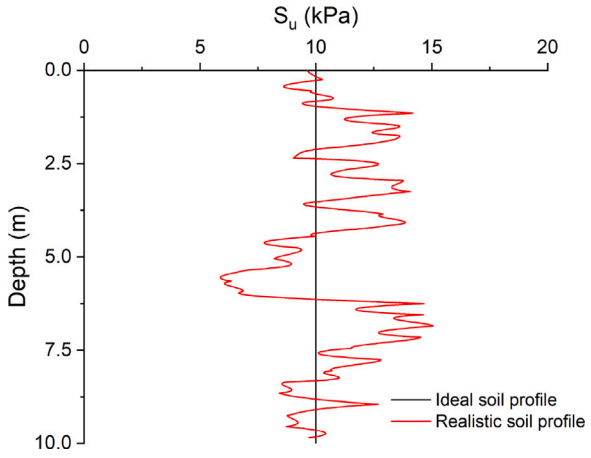


Fig. 20. Ideal versus realistic soil profile.

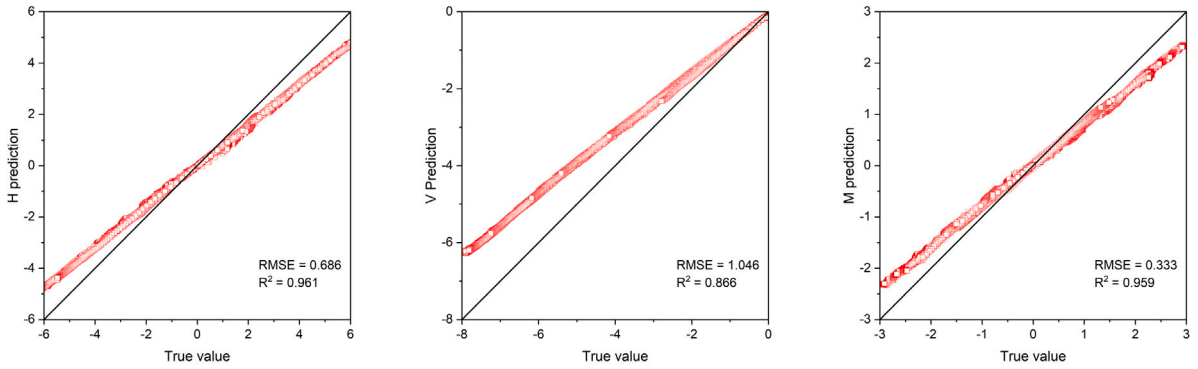
Table 5
Main hyper-parameter of LSTM model for comparison experiment.

	LSTM	
Hyper-parameter	Description	Value
Nh	Number of hidden layers	2
Nn	Number of nodes in the hidden layers	256
η	Learning rate in the optimizer	0.001
Batch_size	Number of training samples utilized in one iteration	128
Epoch	Number of iterations during training	50
Validation_split	Proportion of the validation set in total training	0.2
Input_length	Time length	1

4.4. Extended comparison with CNN and LSTM models

In the previous section, the FC neural network model was shown to be able to predict the caisson response with excellent performance. However, this is not enough to make a strong statement about the superiority of the FC model without horizontal model comparisons. This section will explore the performance of more complex DNN models on this specific problem of the caisson response prediction. Specifically, the predictions of the load–displacement relationship experiments were also carried out on another two prevailing models against the same dataset, i.e., the one-dimensional convolution neural network (1D-CNN) (Kiranyaz et al., 2021) and the LSTM model. Both the 1D-CNN and LSTM models are widely used in time-series problems by introducing convolutional algorithms and gating mechanisms respectively, and have the essential ability to achieve nonlinear regression prediction. The kernel size of 1D-CNN and the time length in LSTM share a similar function, which controls how many past time steps are taken into account in predicting the current response. In other words, they have the ability to consider the entire loading path when making predictions. So, for a fair comparison, these parameters are set to 1 to avoid introducing the temporal relationship between data. Other parameters in the 1D-CNN model (e.g. stride, padding and dilations) control the strategy of temporal input and will be defaulted to avoid cheating. In addition, the number of hidden layers and hidden neurons in these two models are the same as the setting in the FC neural network model (noted that filters in 1D-CNN are equivalent to the neurons). The detailed model hyper-parameters are summarized in Tables 5–6. Out of 96,000 sets of data from homogeneous soils, 76,800 sets were randomly selected to train these models and 19,200 sets were retained for testing.

The prediction results of the three neural networks on the test dataset are presented in Table 7, which compares all three evaluation matrices. In the table, all three models show excellent predictions, while the extremely small prediction loss demonstrates that the predicted result is close to the underlying values. However, compared to the other two neural network models, the FC neural network model has the simplest structure and the highest computational efficiency. The computational approach of convolution and the gating mechanism have proved effective in time-series prediction, but do not bring additional accuracy in non-linear regression, since additional activated parameters lead to a reduction in computational efficiency. Therefore, the FC



(a) Horizontal force prediction (b) Vertical force prediction (c) Moment prediction

Fig. 21. Comparison between mechanical prediction and simulation.

Table 6
Main hyper-parameter of 1D-CNN model for comparison experiment.

	1D-CNN	
Hyper-parameter	Description	Value
Nh	Number of hidden layers	2
Filters	Number of filters in the hidden layers	256
Kernel_size	Length of the convolution kernel	1
Strides	Step size of convolution kernel shift	1
Padding	Padding mode	valid
Dilations	Convolutional kernel dilation	[1]
η	Learning rate in the optimizer	0.001
Batch_size	Number of training samples utilized in one iteration	128
Epoch	Number of iterations during training	50
Validation_split	Proportion of the validation set in total training	0.2

Table 7
Prediction results of the three neural network models.

	FC	LSTM	1D-CNN
RMSE	0.031	0.045	0.049
R^2	1.000	0.999	0.999
MAE	0.021	0.024	0.028
Training time	135s	240s	161s

neural network model should be considered first in similar mechanical prediction problems.

5. Limitations and recommendations

In this study, two uniform clay seabeds with relatively ideal undrained shear strength profiles (i.e., constant and linear increase with depth) were studied. However, in the field, the soil profile may be layered with significant variation of the shear strength, as shown in Fig. 20. The realistic soil profile in the figure was generated by randomly using a mean value of 10 kPa. To highlight the influence of soil strength variation on the foundation response, a specific foundation with $L = 8$ m in realistic soil was modelled. It can be observed that the prediction performance decreases substantially in Fig. 21, where the total prediction errors of RMSE, R^2, and MAE are up to 0.760, 0.927, and 0.583, respectively. The predicted values are all lower than the computed from 3D FE model, implying an underestimation of the response. Therefore, a more complex nonuniform soil strength profile should be considered in the future. It is envisaged that in situ seabed characterization results, like CPT data, can be directly used as model training input to account for the spatial variation of soil properties.

In addition, it should also be noted that, in this study, the elastic-perfect plastic model with Tresca yield criterion was used for modelling the clay, which cannot accurately model the pre-yielding non-linearity of clay's behaviour. Therefore, in the future, advanced non-linear models such as the NGI-ADP (Grimstad et al., 2012), modified Cam-Clay (Matsuoka et al., 1999) or SaniClay (Dafalias et al., 2006) will be employed to capture the non-linear response of soil and foundation. The DNN models can then be refined with more accurate and realistic experimental data from physical testing, for example, centrifuge tests, field tests, etc.

Furthermore, this model is limited to shallow foundations with embedment depth ratios less than 1. When extrapolating the model to larger embedment depth ratio foundations in Fig. 22, the prediction accuracy gradually decreases as the embedment depth ratios are away from the training set range. The prediction accuracy of the model is acceptable ($R^2 > 0.96$) in the range of the embedment ratio less than 1.5. This implies that there is a change of failure mechanism for $L/D > 1.5$ and makes it challenging for the trained model to predict the foundation response. This demonstrates that the deep learning algorithm is highly dependent on the validity and scope of the training data. To further enhance the application of the model, more field data is required for the training set. However, obtaining comprehensive and practical training data is difficult, which limits the further evolution of the model. Therefore, it is worth studying how to compensate for the scarce data by improving the model's generalization ability. In addition, to highlight the influence of data amount, a new experiment was designed to explore the minimum data set required for model training. 10 random initialization and divisions are utilized to illustrate the robustness of the model under different data amounts. The results, as shown in Fig. 23, demonstrate that prediction accuracy decreases and becomes unstable at datasets less than 35,000. It should be noted that there is an optimum value of required data to ensure sufficient accuracy and robustness of the trained model, which is around 15,000 data sets in this study. Therefore, in future studies, some trial runs will be performed to find the optimum data amount to reduce the computation cost.

6. Conclusion

In this study, a FC neural network model has been developed to predict the mechanical response of the suction caisson in clay by simply inputting foundation configuration and soil profile. It was found that compared to the traditional general formulation, the FC neural network model is more accurate and flexible, without the limitation of the pre-assumptions in the conventional design models. More importantly, by testing the FC model against an independent dataset in Zhang et al. (2020), it was found that the FC neural network model can also well capture the response of suction caisson in sand. Even better performance of the FC model was observed than the original LSTM model used in Zhang et al. (2020). This implies that increasing the complexity of the DNN model does not necessarily improve the model performance. The robustness and generalization ability of the FC model was further evaluated, demonstrating that it possesses high reproductivity, high

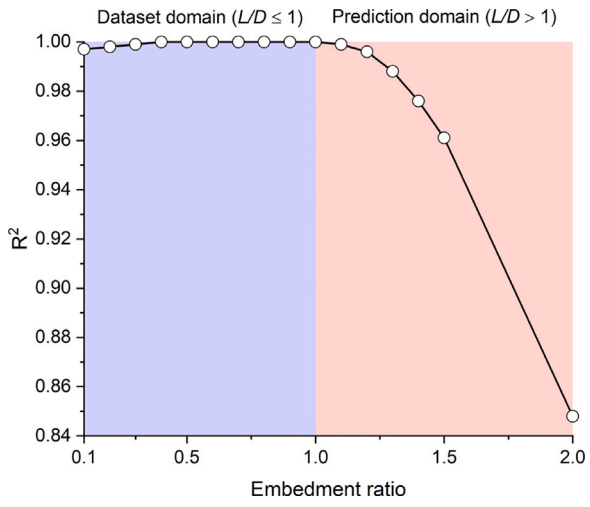


Fig. 22. Model prediction error for foundation responses with embedment ratios larger than 1.

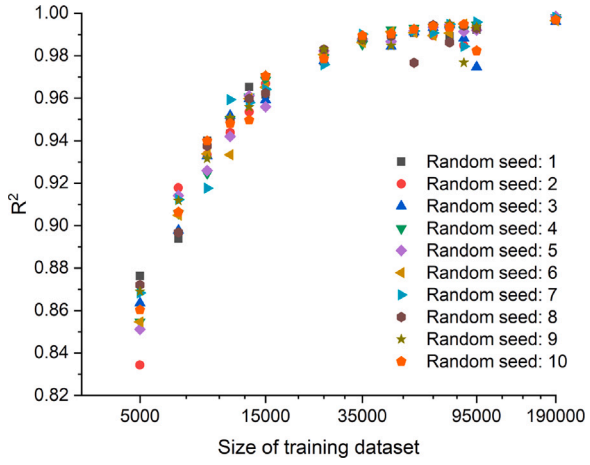


Fig. 23. The exploration of the minimal required data size.

stability, and good generalization ability. This suggests that even a very "shallow" FC neural network model can learn the intrinsic failure mechanisms of the caissons from raw data and predict their non-linear mechanical responses under complex three-dimensional loads effectively.

More importantly, this study also proves that the generalization ability analysis can also be employed to reveal the evolution of the intrinsic foundation failure mechanisms. This feasible strategy demonstrates that DNN not only has the ability to simulate the relationship between inputs and outputs adaptively, but also to mine the intrinsic patterns of data. This finding provides a new direction of exploration in the geotechnical field combined with deep learning techniques.

CRediT authorship contribution statement

Xilin Yin: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Huan Wang:** Conceptualization, Methodology, Resources, Supervision, Writing – review & editing. **Federico Pisanò:** Writing – review & editing, Supervision, Resources. **Ken Gavin:** Writing – review & editing, Supervision, Resources. **Amin Askarinejad:** Writing – review & editing, Supervision, Resources.

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.

Data availability

Data will be made available on request

References

Blundell, C., Cornebise, J., Kavukcuoglu, K., Wierstra, D., 2015. Weight uncertainty in neural network. In: International Conference on Machine Learning. PMLR, pp. 1613–1622.
Bransby, M., Randolph, M., 1998. Combined loading of skirted foundations. Géotechnique 48 (5), 637–655.
Bransby, M., Yun, G.-J., 2009. The undrained capacity of skirted strip foundations under combined loading. Géotechnique 59 (2), 115–125.
Brutzkus, A., Globerson, A., Malach, E., Shalev-Shwartz, S., 2017. SGD learns over-parameterized networks that provably generalize on linearly separable data. arXiv preprint arXiv:1710.10174.
Butterfield, R., Houlsby, G.T., Gottardi, G., 1997. Standardized sign conventions and notation for generally loaded foundations. Géotechnique 47 (5), 1051–1054.
Byrne, B.W., 2000. Investigations of suction caissons in dense sand (Ph.D. thesis). University of Oxford Oxford, UK.
Byrne, B., Houlsby, G., Martin, C., Fish, P., 2002. Suction caisson foundations for offshore wind turbines. Wind Eng. 26 (3), 145–155.
Cassidy, M., Randolph, M., Byrne, B., 2006. A plasticity model describing caisson behaviour in clay. Appl. Ocean Res. 28 (5), 345–358.
Cremer, C., Pecker, A., Davenne, L., 2002. Modelling of nonlinear dynamic behaviour of a shallow strip foundation with macro-element. J. Earthq. Eng. 6 (02), 175–211.
Dafalias, Y.F., Manzari, M.T., Papadimitriou, A.G., 2006. SANICLAY: Simple anisotropic clay plasticity model. Int. J. Numer. Anal. Methods Geomech. 30 (12), 1231–1257.
Feng, J., Shen, W.Z., 2017. Design optimization of offshore wind farms with multiple types of wind turbines. Appl. Energy 205, 1283–1297.
Fu, D., Gaudin, C., Tian, Y., Cassidy, M.J., Bienen, B., 2017. Uniaxial capacities of skirted circular foundations in clay. J. Geotech. Geoenviron. Eng. 143 (7), 04017022.
Fu, D., Zhang, Y., Yan, Y., 2020. Bearing capacity of a side-rounded suction caisson foundation under general loading in clay. Comput. Geotech. 123, 103543.
Gerolymos, N., Zafeirakos, A., Karapiperis, K., 2015. Generalized failure envelope for caisson foundations in cohesive soil: Static and dynamic loading. Soil Dyn. Earthq. Eng. 78, 154–174.
Géron, A., 2022. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.".
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.
Gourvenec, S., Barnett, S., 2011. Undrained failure envelope for skirted foundations under general loading. Géotechnique 61 (3), 263–270.
Gourvenec, S., Cassidy, M., 2005. Frontiers in Offshore Geotechnics: Proceedings of the International Symposium on Frontiers in Offshore Geotechnics. (IS-FOG 2005), 19-21 Sept 2005, Perth, WA, Australia, CRC Press.
Grimstad, G., Andresen, L., Jostad, H.P., 2012. NGI-ADP: Anisotropic shear strength model for clay. Int. J. Num. Anal. Methods Geomech. 36 (4), 483–497.
Gulli, A., Pal, S., 2017. Deep Learning with Keras. Packt Publishing Ltd.
Hiransha, M., Gopalakrishnan, E.A., Menon, V.K., Soman, K., 2018. NSE stock market prediction using deep-learning models. Procedia Comput. Sci. 132, 1351–1362.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.
Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. 79 (8), 2554–2558.
Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Netw. 2 (5), 359–366.
Houlsby, G., 2016. Interactions in offshore foundation design. Géotechnique 66 (10), 791–825.
Houlsby, G.T., Cassidy, M.J., 2002. A plasticity model for the behaviour of footings on sand under combined loading. Géotechnique 52 (2), 117–129.
Hu, Y., Randolph, M., 1998. A practical numerical approach for large deformation problems in soil. Int. J. Numer. Anal. Methods Geomech. 22 (5), 327–350.
Hung, L.C., Kim, S.-R., 2014. Evaluation of undrained bearing capacities of bucket foundations under combined loads. Mar. Georesources Geotechnol. 32 (1), 76–92.
Ibsen, L.B., Barari, A., Larsen, K.A., 2014. Adaptive plasticity model for bucket foundations. J. Eng. Mech. 140 (2), 361–373.
Jagota, V., Sethi, A.P.S., Kumar, K., 2013. Finite element method: An overview. Walailak J. Sci. Technol. (WJST) 10 (1), 1–8.
Jeanjean, P., Zhang, Y., Zakeri, A., Andersen, K., Gilbert, R., Senanayake, A., 2017. A framework for monotonic py curves in clays. In: Offshore Site Investigation Geotechnics 8th International Conference Proceeding, Vol. 108, no. 141. Society for Underwater Technology, pp. 108–141.

Jin, Z., Yin, Z.-Y., Kotronis, P., Li, Z., Tamagnini, C., 2019. A hypoplastic macroelement model for a caisson foundation in sand under monotonic and cyclic loadings. Mar. Struct. 66, 16–26.

Karapiperis, K., Gerolymos, N., 2014. Combined loading of caisson foundations in cohesive soil: Finite element versus Winkler modeling. Comput. Geotech. 56, 100–120.

Keskar, N.S., Mudigere, D., Nocedal, J., Smelyanskiy, M., Tang, P.T.P., 2016. On large-batch training for deep learning: Generalization gap and sharp minima. arXiv preprint arXiv:1609.04836.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J., 2021. 1D convolutional neural networks and applications: A survey. Mech. Syst. Signal Proc. 151, 107398.

Kuo, Y., Jaksa, M., Lyamin, A., Kaggwa, W., 2009. ANN-based model for predicting the bearing capacity of strip footing on multi-layered cohesive soil. Comput. Geotech. 36 (3), 503–516.

Li, Z., Kotronis, P., Escoffier, S., Tamagnini, C., 2016. A hypoplastic macroelement for single vertical piles in sand subject to three-dimensional loading conditions. Acta Geotech. 11 (2), 373–390.

Li, Y., Yuan, Y., 2017. Convergence analysis of two-layer neural networks with relu activation. Adv. Neural Inform. Process. Syst. 30.

Madhyastha, P., Jain, R., 2019. On model stability as a function of random seed. arXiv preprint arXiv:1909.10447.

Matsuoka, H., Yao, Y., Sun, D., 1999. The cam-clay models revised by the SMP criterion. Soils Found. 39 (1), 81–95.

Mehravar, M., Harireche, O., Faramarzi, A., 2016. Evaluation of undrained failure envelopes of caisson foundations under combined loading. Appl. Ocean Res. 59, 129–137.

Momeni, E., Nazir, R., Armaghani, D.J., Maizir, H., 2014. Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. Measurement 57, 122–131.

Murff, J., 1994. Limit analysis of multi-footing foundation systems. In: International Conference on Computer Methods and Advances in Geomechanics. pp. 233–244.

Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning. ICML-10, pp. 807–814.

Nejad, F.P., Jaksa, M.B., 2017. Load-settlement behavior modeling of single piles using artificial neural networks and CPT data. Comput. Geotech. 89, 9–21.

Pisanò, F., Flessati, L., Di Prisco, C., 2016. A macroelement framework for shallow foundations including changes in configuration. Géotechnique 66 (11), 910–926.

Pontes, F.J., Amorim, G., Balestrassi, P.P., Paiva, A., Ferreira, J.R., 2016. Design of experiments and focused grid search for neural network parameter optimization. Neurocomputing 186, 22–34.

Qu, Z.-Q., 2004. Model Order Reduction Techniques with Applications in Finite Element Analysis: With Applications in Finite Element Analysis. Springer Science & Business Media.

Randolph, M., Gourvenec, S., 2017. Offshore Geotechnical Engineering. CRC Press.

Reimers, C., Requena-Mesa, C., 2020. Deep learning–an opportunity and a challenge for geo-and astrophysics. In: Knowledge Discovery in Big Data from Astronomy and Earth Observation. Elsevier, pp. 251–265.

Roscoe, K., 1956. The stability of short pier foundations in sand. Br. Weld. J. 343–354.

Salciarini, D., Tamagnini, C., 2009. A hypoplastic macroelement model for shallow foundations under monotonic and cyclic loads. Acta Geotech. 4 (3), 163–176.

Sanders, J., Kandrot, E., 2010. CUDA By Example: An Introduction to General-Purpose GPU Programming. Addison-Wesley Professional.

Shahin, M.A., 2014. Load–settlement modeling of axially loaded drilled shafts using CPT-based recurrent neural networks. Int. J. Geomech. 14 (6), 06014012.

Sharma, S., Sharma, S., Athaiya, A., 2017. Activation functions in neural networks. Towards Data Sci. 6 (12), 310–316.

Skau, K.S., Chen, Y., Jostad, H.P., 2018a. A numerical study of capacity and stiffness of circular skirted foundations in clay subjected to combined static and cyclic general loading. Geotechnique 68 (3), 205–220.

Skau, K.S., Grimstad, G., Page, A.M., Eiksund, G.R., Jostad, H.P., 2018b. A macro-element for integrated time domain analyses representing bucket foundations for offshore wind turbines. Mar. Struct. 59, 158–178.

Systèmes, D., 2014. Abaqus 6.14—Abaqus analysis user's guide. In: Dassault Systèmes: Vélizy-Villacoublay. France.

Szabo, B., Babuska, I., 2021. Finite Element Analysis: Method, Verification and Validation. Method, Verification and Validation, Finite Element Analysis.

Taiebat, H., Carter, J., 2000. Numerical studies of the bearing capacity of shallow foundations on cohesive soil subjected to combined loading. Géotechnique 50 (4), 409–418.

Tarawneh, B., 2013. Pipe pile setup: Database and prediction model using artificial neural network. Soils Found. 53 (4), 607–615.

Tieleman, T., Hinton, G., et al., 2012. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural Netw. Mach. Learn. 4 (2), 26–31.

Villalobos Jara, F.A., 2006. Model testing of foundations for offshore wind turbines (Ph.D. thesis). University of Oxford.

Vulpe, C., 2015. Design method for the undrained capacity of skirted circular foundations under combined loading: Effect of deformable soil plug. Géotechnique 65 (8), 669–683.

Vulpe, C., Gourvenec, S., Power, M., 2014. A generalised failure envelope for undrained capacity of circular shallow foundations under general loading. Géotechnique Letters 4 (3), 187–196.

Wang, Y., Cassidy, M.J., Bienen, B., 2021. A macro-element model for predicting the combined load behaviour of spudcan foundations in clay overlying sand. Géotechnique 1–13.

Yin, Z.-Y., Teng, J.-C., Li, Z., Zheng, Y.-Y., 2020. Modelling of suction bucket foundation in clay: From finite element analyses to macro-elements. Ocean Eng. 210, 107577.

Zhang, Y., Cassidy, M.J., Bienen, B., 2014. A plasticity model for spudcan foundations in soft clay. Can. Geotech. J. 51 (6), 629–646.

Zhang, P., Yin, Z.-Y., Zheng, Y., Gao, F.-P., 2020. A LSTM surrogate modelling approach for caisson foundations. Ocean Eng. 204, 107263.

Zhou, H., Ibrahim, C., Zheng, W.X., Pan, W., 2022. Sparse Bayesian deep learning for dynamic system identification. Automatica 144, 110489.