An AI-based Model for Describing Cyclic Characteristics of Granular Materials

Pin ZHANG[1], Zhen-Yu YIN[1*], Yin-Fu JIN[1], Guan-Lin YE[2]

1 Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

2 Department of Civil Engineering, Shanghai Jiao Tong University, Shanghai 200030, China

* Corresponding author: Dr Zhen-Yu YIN, Tel: +852 3400 8470; Fax: +852 2334 6389; E-mail: zhenyu.yin@polyu.edu.hk; zhenyu.yin@gmail.com

Abstract: Modelling cyclic behaviour of granular soils under both drained and undrained conditions with a good performance is still a challenge. This study presents a new way of modelling the cyclic behaviour of granular materials using deep learning. To capture the continuous cyclic behaviour in time dimension, the long short-term-memory (LSTM) neural network is adopted, which is characterised by the prediction of sequential data, meaning that it provides a novel means of predicting the continuous behaviour of soils under various loading paths. Synthetic datasets of cyclic loading under drained and undrained conditions generated by an advanced soil constitutive model are first employed to explore an appropriate framework for the LSTM-based model. Then the LSTM-based model is used to estimate the cyclic behaviour of real sands, i.e., the Toyoura sand under the undrained condition and the Fontainebleau sand under both undrained and drained conditions. The estimates are compared with actual experimental results, which indicates that the LSTM-based model can simultaneously simulate the cyclic behaviour of sand under both drained and undrained conditions, i.e., (1) the cyclic mobility mechanism, the degradation of effective stress and large deformation under the undrained condition, and (2) shear strain accumulation and densification under the drained condition.

Keywords: sand, cyclic loading, deep learning, neural network, constitutive model, liquefaction

1. Introduction

Sand, as a granular material, has been extensively used in engineering practice, such as in the foundations of offshore wind turbines [1,2], airport tracks [3] and high-speed railways [4]. One common characteristic of these applications is that sand is subjected to various degrees of cyclic loading. The behaviour of sands under cyclic loading has been widely investigated by means of simple-shear, triaxial and resonant column tests under both drained and undrained conditions[5-10].

Researchers have proposed numerous constitutive models to describe the complex behaviour of sand. These can be categorised into five groups: (1) nonlinear hypoelastic models[11]; (2) incrementally nonlinear models[12,13]; (3) hypoplastic models[14,15]; (4) elastoplastic models[16-21]; and (5) micromechanics-based models [22,23]. However, conventional soil models are only applicable in limited cases because they depend heavily on a special mathematical formulation and certain assumptions. Up to now, modelling cyclic behaviour of soils under both drained and undrained conditions with a good performance is still a challenge.

Nowadays, data-based soil-constitutive models developed by machine learning (ML) algorithms have been gradually proposed because these algorithms are not constrained by assumptions or model classes, and their application scope can be extended with the increase in the datasets[24-26]. Moreover, ML-based model has capacity of capturing the mechanical behaviour of materials with the input of raw experimental tests, thereby the cumbersome tasks to develop constitutive model and identify parameters are dismissed [27-30]. Ghaboussi et al. [31] first proposed artificial neural network (ANN)-based constitutive models for simulating the behaviour of materials. Thereafter, numerous researchers have conducted works to develop ML-based soil-constitutive models [32-37]. The focus of these models has mainly been on the stress-strain behaviour of various soils under monotonic loading. Nevertheless, only limited research

works have been devoted to developing ML-based constitutive models for simulating soil behaviour under cyclic loading. For example, Basheer [38] proposed several methodologies based on ANN for modelling the cyclic behaviour of clay soils. This model was tested using synthetic data and experimental results, but the prediction error obviously increased with the uptick in the number of loading cycles, and the ability to extrapolate beyond the limits of the training set was deficient. Overall, published research works only investigated the behaviour of materials under several loading cycles using the ML-based constitutive model[39-41]. Soil behaviour occurring under numerous loading cycles, such as cyclic mobility, liquefaction, shear stain accumulation and densification, has not been investigated by the ML-based constitutive model. Furthermore, no ML-based constitutive model exists that can simultaneously simulate soil behaviour under drained and undrained conditions, although such a power would be much more significant for a model's generalisation ability. Therefore, the improvement in the prediction of soil behaviour under cyclic loading using the ML-based constitutive model is worthy of deep exploration.

Compared general neural networks, connections between hidden units can be established in recurrent neural networks (RNNs), allowing the latter to retain memories of recent events. Conventional RNNs exist gradients vanishing and exploding[42], a long short-term memory (LSTM) neural network as a variation of RNNs was thus developed to overcome this problem The LSTM algorithm has recently been used in practical engineering with time-series characteristics such as the prediction of long-term settlement of structures[43,44], hydro-mechanical responses of multi-permeability porous media[45] and structural seismic response[46]. Because soil behaviour under cyclic loading is a continuous process, the current stress-strain status depends on the soil behaviour at previous steps and also affects the soil behaviour at the later steps. The LSTM-based model may thus be far more appropriate than the current commonly used ANN-based constitutive model for simulating soil behaviour under cyclic loading.

In this study, an LSTM-based general-purpose constitutive model is developed for simultaneously simulating the cyclic behaviour of sand under drained and undrained conditions. Synthetic data, under both drained and undrained conditions generated by a traditional constitutive model, are first generated to explore an appropriate framework for the LSTM-based model. Thereafter, the LSTM-based model is employed to simulate the cyclic behaviour of Toyoura sand under the undrained condition and Fontainebleau sand under drained and undrained conditions, with comparisons drawn with the experimental data.

2. Modelling soil cyclic behaviour using LSTM

2.1 Long short-term memory neural networks

RNNs contain cyclic connections that make them powerful tools for modelling sequential data. In an RNN model, each unit is associated with the others in hidden layers at different time steps; the previous information stored at previous hidden layers is thus applied to the current output [47] (see Fig. 1(a)). However, traditional RNNs that are used to learn what to put in short-term memory take too much time and sometimes even do not work at all, especially when minimum time lags between inputs and corresponding signals are long [42]. LSTM, as a variation of RNNs, has been first proposed to solve complex, long-time-span tasks that have never been solved by traditional RNNs. Based on this, LSTM and its variants such as GRU and bi-directional LSTM have been extensively used in engineering domain [43,48,49]. Herein, LSTM, as the standard algorithm of this class of algorithms, has been adopted in this study to investigate its feasibility in modelling soil cyclic behaviour.

To learn long-term dependence, a memory cell is added to LSTM (see Fig. 1(b)). The cell consists of three gates: forget, input and output gates, which are actually nonlinear functions. Herein, x_i denotes the

input values at time t, and h_{t-1} denotes the hidden output at time t–1. The computation of the memory cell at time t starts from the forget gate f_t, which decides whether to retain or forget the previous state c_{t-1} (Eq. [1]). The input gate i_t decides whether to update the state of the LSTM using the current input (Eq. [2]). $\tilde{c}_t$ is the candidate of input to be stored (Eq. [3]). The output gate o_t decides whether to pass on the hidden state h_t to the next iteration (Eq. [4]). The new state c_t stored in the LSTM is the sum of the new gated input $\tilde{c}_t$ and the previous gated state c_{t-1} (Eq. [5]). Finally, the hidden output at time t can be obtained by Eq. [6].

$$\mathbf{f}_j^t = \sigma\left(\left[\mathbf{W}_f\mathbf{x}^t\right]_j + \left[\mathbf{U}_f\mathbf{h}^{t-1}\right]_j + \left[\mathbf{b}_f\right]_j\right) \quad \mathbf{f}^t = \sigma\left(\mathbf{W}_f\mathbf{x}^t + \mathbf{U}_f\mathbf{h}^{t-1} + \mathbf{b}_f\right) \tag{1}$$

$$\mathbf{i}_j^t = \sigma\left(\left[\mathbf{W}_i\mathbf{x}^t\right]_j + \left[\mathbf{U}_i\mathbf{h}^{t-1}\right]_j + \left[\mathbf{b}_i\right]_j\right) \quad \mathbf{i}^t = \sigma\left(\mathbf{W}_i\mathbf{x}^t + \mathbf{U}_i\mathbf{h}^{t-1} + \mathbf{b}_i\right) \tag{2}$$

$$\tilde{\mathbf{c}}_j^t = tanh\left(\left[\mathbf{W}_c\mathbf{x}^t\right]_j + \left[\mathbf{U}_c\mathbf{h}^{t-1}\right]_j + \left[\mathbf{b}_c\right]_j\right) \quad \tilde{\mathbf{c}}^t = tanh\left(\mathbf{W}_c\mathbf{x}^t + \mathbf{U}_c\mathbf{h}^{t-1} + \mathbf{b}_c\right) \tag{3}$$

$$\mathbf{o}_j^t = \sigma\left(\left[\mathbf{W}_o\mathbf{x}^t\right]_j + \left[\mathbf{U}_o\mathbf{h}^{t-1}\right]_j + \left[\mathbf{b}_o\right]_j\right) \quad \mathbf{o}^t = \sigma\left(\mathbf{W}_o\mathbf{x}^t + \mathbf{U}_o\mathbf{h}^{t-1} + \mathbf{b}_o\right) \tag{4}$$

$$\mathbf{c}^t = \mathbf{f}^t \odot \mathbf{c}^{t-1} + \mathbf{i}^t \odot \tilde{\mathbf{c}}^t \tag{5}$$

$$\mathbf{h}^t = \mathbf{o}^t \odot tanh\left(\mathbf{c}^t\right) \tag{6}$$

where σ and *tanh* ~~are the~~ = activation functions *sigmoid* and *tanh*; $\mathbf{W}_f$, $\mathbf{W}_i$, $\mathbf{W}_c$, $\mathbf{W}_o$ = weight matrix connecting input and hidden layers. $\mathbf{U}_f$, $\mathbf{U}_i$, $\mathbf{U}_c$, $\mathbf{U}_o$ = weight matrix connecting adjacent hidden layers in the time sequence dimension; $\mathbf{b}_f$, $\mathbf{b}_i$, $\mathbf{b}_c$, $\mathbf{b}_o$ = biases matrix; e = element-wise product. Note that this section merely introduces the principle of the LSTM algorithm. The actual activation functions used in this study need to be fine-tuned according to the practical issue. *tanh* and *ReLU* are two commonly used activation functions. Because the combination of two functions can overcome their own deficiencies[50], they are thus adopted in this study with detailed introduced in Appendix A.

2.2 Modelling strategy

Herein, a strategy is proposed for marking the cyclic loading process. As shown in Fig. 2, each cyclic loading process is divided into four segments. Fig. 2(a) presents the labels used in a typical strain-control cyclic loading mode. The first label number donates the number of cyclic loading. The second number donates strain direction, where 0 donates positive strain and 1 donates negative strain. The third number donates the evolution of strain, where 0 donates the increase and 1 donates the decrease in strain. The whole cyclic loading process can be clearly recorded in the database in this way; for example, 1, 0, 0 donates the increase in the first positive strain, and 1, 0, 1 donates the decrease in the first positive strain. Fig. 2(b) presents the labels used in a typical stress-control cyclic loading mode. The meaning of each number is similar to those in the strain-control strategy except the second number, where 0 donates positive stress and 1 donates negative stress. Therefore 1, 0, 0 donates the increase in the first positive stress, and 1, 0, 1 donates the decrease in the first positive stress.

Fig. 3 illustrates the framework of the proposed LSTM-based model in this study. Two LSTM-based models are developed for simulating strain-control and stress-control soil behaviour. Note that each model proposed in this study can simultaneously simulate soil cyclic behaviour under drained and undrained conditions and can also simulate the behaviour under each condition separately. An incremental simulation strategy is employed; that is, the output variables at the *i*th step are used as the input variables at the (*i*+1)th step. At the *i*th step, the output variables of two LSTM-based models are expressed by:

$$\text{Strain control: } \left(q^{i+1}, p'^{i+1}, \varepsilon_v^{i+1}\right) = f_1\left(L1, L2, L3, e_0, m, \varepsilon_d^{i+1}, q^i, p'^i, \varepsilon_v^i\right) \tag{7}$$

$$\text{Stress control: } \left(p'^{i+1}, \varepsilon_v^{i+1}, \varepsilon_d^{i+1}\right) = f_2\left(L1, L2, L3, e_0, m, q^{i+1}, p'^i, \varepsilon_v^i, \varepsilon_d^i\right) \tag{8}$$

where q^{i+1}, p'^{i+1}, $\varepsilon i+1$ *v* and $\varepsilon i+1$ *d* = four output variables of model at the (*i*+1)th step: deviatoric stress,

effective mean stress, volumetric strain and axial strain; q^i, p'^i, $\varepsilon i\ v$ and $\varepsilon i\ d$ = output variables of model at the ith step, and they are also the input variables of the model at the (i+1)th step; for εi+1 d in Eq. [7] and q^{i+1} in Eq. [8], both control the strain-control and stress-control cyclic loading process and increase in a fixed interval. $L1$, $L2$, $L3$ = three labels for marking the cyclic loading process as mentioned earlier; e_0 = initial void ratio; m = code for controlling experiment types, 1 represents drained condition and 0 represents undrained condition; f_1, f_2, = formulations of LSTM-based models for describing strain-control and stress-control soil behaviour, respectively. Herein, the LSTM-based model for describing strain-control soil behaviour is introduced clearly to demonstrate the intrinsic mechanism of such models. For the strain-control soil behaviour, the p' holds constant under the drained condition, and εv holds constant under the undrained condition. Such constant variables can be deleted when training an LSTM-based model. Therefore, the output variables are q^{i+1} and εi+1 v under the drained condition, and the input variables are L_1, L_2, L_3, e_0, m (1), εi+1 d, q^i and $\varepsilon i\ v$. For the undrained condition, q^{i+1} and p'^{i+1} are output variables, and L_1, L_2, L_3, e_0, m (0), εi+1 d, q^i and p'^i are input variables. Both drained and undrained data are integrated to train the LSTM-based model, and the number of input and output variables is eight and two, respectively. The first five fixed input variables are related to soil properties and the loading process, and the last three input variables need to be updated at each step. The LSTM-based model for describing stress-control soil behaviour remains similar as long as the output variables are replaced.

It should be noted that all datasets used in the LSTM-based model have been normalised into an interval (–1, 1) using Eq. [9] because doing so can eliminate the effect of different magnitudes of input variables on the model's performance and can clearly reduce computational costs:

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \left(\bar{x}_{\max} - \bar{x}_{\min} \right) + \bar{x}_{\min} \tag{9}$$

where x = actual value of input variables, $x_{\max}$, $x_{\min}$ = maximum and minimum values of input variables; $\bar{x}_{\min} = -1$; $\bar{x}_{\max} = 1$. As a consensus within the field of machine learning, it is very reasonable that 80% of datasets are used to train model and the remaining 20% are used to test model. Similarly, it is important to assign an appropriate proportion of experiments as the training and testing sets for modelling cyclic behaviour of granular materials. Therefore, to guarantee the generalization ability of the proposed LSTM based model, the proportion of testing sets is preferably not more than 20%.

The deep learning tool Keras, based on the Python programming language, which uses TensorFlow as a backend, was employed to establish the LSTM-based model. All results were acquired on a computer with Intel Core i7-7700K 4.2 GHz CPU and 16 GB memory.

2.3 Evaluation indicators

Two commonly used evaluation indicators – mean absolute error (MAE) and mean absolute percentage error (MAPE) – were used to evaluate the performance of ML models in this study. The combination of MAE and MAPE helps overcome the deficiencies of both, so that both can be used extensively to evaluate model performance [51,52]. Low values of these two indicators show that a model has excellent performance. The expression of MAE and MAPE can be obtained by

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |r_i - p_i| \tag{10}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{r_i - p_i}{r_i} \right| \times 100\% \tag{11}$$

where r = actual output value; p = predicted output value; n = total number of datasets.

3. Performance of LSTM–based constitutive models on the synthetic data

3.1 Synthetic data generated by a conventional constitutive model

To investigate the performance of the LSTM-based models, a constitutive model proposed by Yin et al. [13] for describing the cyclic behaviour of sand under drained and undrained conditions was used in this study (see Eq. [12]–[13]). The model was developed using a theoretical function rather than being directly based on the experimental data to eliminate the interference of experimental and measurement errors related to the mapping capability of LSTM. Moreover, the experimental data tend to be limited and insufficient for exploring the capability of LSTM, whereas the data can be generated infinitely by a theoretical function:

$$\dot{\varepsilon}_v = M_{pt}\left|\dot{\varepsilon}_d\right| - \frac{q}{p'}\dot{\varepsilon}_d \tag{12}$$

$$\dot{q} = 3G\left[\dot{\varepsilon}_d - \left(\frac{q}{p'M_p}\right)\left|\dot{\varepsilon}_d\right|\right] \tag{13}$$

where $\dot{\varepsilon}_v$ = volumetric strain increment; $\dot{q}$ = incremental deviatoric stress; $\dot{\varepsilon}_d$ = deviatoric stain increment; q = deviatoric stress; p' = mean effective stress; M_p = stress ratio; M_{pt} = phase-transformation stress ratio. The formulations for determining G, M_p and M_{pt} can be gained by referring to Yin et al. [13], taking the nonlinear elasticity, critical state and nonlinear stress dilatancy into consideration. The relevant parameters in this constitutive model are presented in Table 1.

The cyclic behaviour of sand under drained and undrained conditions was investigated in this study. Stress-control and strain-control loading modes were taken into consideration in both drained and undrained conditions. Therefore, four cases were conducted to validate the LSTM-based models: drained condition under stress-control and strain-control cyclic loading modes, and undrained condition under

stress-control and strain-control cyclic loading modes. In each case, a total of 16 simulation tests under cyclic loading were generated using Eq. [12]–[13] to develop the LSTM-based models. The researchers used 13 tests (e_0 = 0.600, 0.625, 0.650, 0.675, 0.700, 0.725, 0.750, 0.800, 0.825, 0.850, 0.875, 0.900 and 0.925) to train the LSTM-based models, and the remaining three tests (e_0 = 0.575, 0.775 and 0.950) were used to test the models. Tests with e_0 = 0.575 and 0.950 were used to test the extrapolation capability of the LSTM-based model, and the test with e_0 = 0.775 was used to test the interpolation capability of the model.

3.2 Determination of parameters of LSTM model

Hyper-parameters optimisation is a key part of deep learning. The optimum parameters in the LSTM-based models for describing two cyclic loading modes are summarised in Table 2. Other parameters that are not present in this table were set as the default values in Keras (learning rate, regularization and dropout). In this study, the architecture, epoch number and activation functions that control the topology and representational capacity of LSTM were hard to tune to find a fixed pattern for different problems; therefore these three parameters were determined based on the grid search method[53]. The batch size was identical to the number of datasets in each test, therefore LSTM was fed by a set of intact experimental test results at each training round, which was beneficial in terms of allowing LSTM to capture the relationships between input and output parameters. Ten-fold cross-validation (validation split = 0.1) was adopted to avoid the overfitting issue and enhance the robustness of the LSTM-based model. Meanwhile, the loss function curves of the training and testing sets were extracted as feedback information to examine the potential underfitting and overfitting problems. The optimiser Adaptive Moment Estimation (*Adam*) is the commonly used optimiser for training a deep and complex LSTM neural network[54] based on its efficiency and effectiveness; therefore it was used in this study. The

learning rate that controls the magnitude of update of weights at each step is an important hyper-parameter in *Adam*[55]. It is generally set as the default value 0.001 and is reasonable for most problems to find the optimum weights of deep learning algorithms, but this parameter needs to be finely tuned for a complex problem with numerous saddle points[56].

The training of LSTM based model consists of three phases: training, validation and testing, as presented in Fig. 4. The database is first divided into training and testing sets. The allocation of training and testing sets has a great impact on the model performance, because the model trained by different training sets exists great difference. *k*-fold cross-validation method is thus proposed to reduce the effect of random allocation on the model performance, among which the 10-fold cross-validation is widely used. It should be noted that the 10-fold cross-validation process only uses the training set, in which the training set is divided into 10 subsets. At each training epoch, nine subsets are randomly selected to train the model, and its performance is evaluated by the remaining one subset. In other words, at each training epoch, a total of 10 sub-models are trained and their prediction errors on the individually remaining subset are obtained. The current weighs and biases that represent the model performance are evaluated by the mean error on the remaining ten subsets (i.e. one subset for one sub-model). After the training process terminated, the optimum trained model with optimum weights and biases is obtained, and meanwhile the testing set is used to examine the performance of the trained model.

Mean square error (MSE) is defined as the loss function, as shown in Eq. [14]. ~~Ten-fold cross-validation method is used to enhance the robustness of proposed model, in which the original training sets are divided into ten subsets. Nine subsets are used to train model and one remaining subset is used to validate the model's performance.~~ The values of weights and biases in LSTM are updated to eliminate the loss value on the validation sets. Meanwhile the loss values on both training and testing sets

can be obtained, as presented in Fig. 5, which are used to evaluate the training performance. If the loss values on both training and testing sets converge to low values and keep constant, the training process would be terminated. It can be seen from Fig. 5 that all loss values were virtually close to zero as the epoch number reaches to 300. The simultaneous decrease in the loss function values for the training and testing sets indicates that the underfitting problem did not exist in the LSTM-based model. Meanwhile, the unchangeable convergence values of the testing set with increasing epochs indicate that overfitting problems were effectively eliminated:

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}\left(y_p-y_a\right)^2 \tag{14}$$

where n = total number of datasets; y_p = predicted values; y_a = actual values.

3.3 Simulation results

This section presents the simulation results for the testing sets using the LSTM-based model. The predicted results for the training sets show perfect agreement with the actual results, which are therefore presented as supplementary information for brevity.

3.3.1 Strain-control drained and undrained triaxial tests

Fig. 6 presents the predicted strain-control results of drained triaxial tests, compared with the actual results. It can be observed that the predicted stress ratio–axial strain and volumetric strain–axial strain relationships showed excellent agreement with the actual relationships. The LSTM-based model exhibited great extrapolation (see Figs. 6(a) and 6(c)) and interpolation capability (see Fig. 6(b)) in predicting the stress ratio–axial strain relationship, whereas the interpolation performance (see Fig. 6(e)) outperformed the extrapolation performance (see Figs. 6(d) and 6(f)) in predicting the volumetric strain–axial strain relationship. It can be seen from Figs. 6(d–f) that the LSTM-based model can accurately capture the

contractive behaviour for loose soils and dilatant behaviour for dense soils.

Fig. 7 presents the predicted strain-control results of undrained triaxial tests, compared with the actual results. The initial confining pressure was 300 kPa, and the axial strain was maintained at a fixed amplitude (–1% ~ 1%). To eliminate the scale of input variables on the model performance, p' and q were divided by initial effective confining pressure during the model training phase. The predicted deviatoric stress–effective mean stress and deviatoric stress–axial strain relationships displayed great agreement with the actual relationships for the interpolated and extrapolated testing sets. For loose sand (see Figs. 7(b-c)), the LSTM-based model can accurately capture the degradation of effective mean stress as the increase in the number of loading cycles. Meanwhile, a looser sand has great liquefaction potential and can also be identified by the LSTM-based model.

3.3.2 Stress-control drained and undrained triaxial tests

The predicted stress-control results of drained triaxial tests are presented in Fig. 8. The predicted stress ratio–axial strain and volumetric strain–axial strain relationships using the LSTM-based model showed excellent agreement with the actual relationships. This model can recognise the dilatant behaviour for dense soils (see Fig. 8(d)), contractive behaviour for loose soils (see Fig. 8(f)) and the phase transformation from soil dilation to contraction (see Fig. 8(e)).

The predicted stress-control results under the undrained condition are shown in Fig. 9. The initial effective confining pressure was 500 kPa, and the deviatoric stress was maintained at a fixed amplitude (–20 ~ 120 kPa). Similar to the training strategy of the strain-control constitutive model, p' and q were also divided by initial effective confining pressure. The LSTM-based model could accurately capture the deviatoric stress–effective mean stress and deviatoric stress–axial strain relationships for the interpolated and extrapolated testing sets. Moreover, for dense sand, a cyclic mobility phenomenon could be observed

by the model (see Figs. 9(a) and 9(d)). Medium sand also exhibited a cyclic mobility with larger deformation (see Figs. 9(b) and 9(e)) and almost reached a liquefied state in fewer loading cycles than did the dense sand. Large deformation could be directly observed in fewer loading cycles for loose sand (see Figs. 9(c) and 9(f)). Therefore, the LSTM-based model is capable of predicting the dependence of large deformation and liquefaction potential on the relative density of sands.

4. Performance of LSTM–based constitutive models on experimental data

4.1 Undrained cyclic triaxial tests for Toyoura sand

4.1.1 Data source

To investigate the reliability of LSTM-based models in simulating actual sand behaviour, a series of stress-control undrained cyclic triaxial tests on Toyoura sand conducted by Ye et al. [57] were used to validate the models proposed in this study. These experimental results clearly described cyclic mobility under cyclic loading, and had sufficient data collected from the flow liquefaction process. The accurate description of cyclic mobility and flow liquefaction is vitally important in practical geotechnical engineering, such as in the evaluation of the shaft capacity of piles[58,59]. Researchers have perennially been preoccupied with developing advanced and constitutive models to describe the cyclic mobility effects and liquefaction of sands. Wichtmann et al. [10] recently comprehensively compared the performance of three constitutive models (hypoplasticity[15] vs. SANISAND[16] vs. ISA[60]) based on monotonic and cyclic tests on fine sand. The deviation between the predicted and experimental results could be clearly observed, and several predicted stress–strain relationships even differed entirely from the experimental results. In reality, the prediction accuracy of traditional soil models is limited, because they depend heavily on the special mathematical formulation types and prescribed assumptions. Therefore, it is worth investigating

the performance of the LSTM-based model in simulating the cyclic behaviour of sands.

Herein, five experimental tests with a total of 23,270 datasets from Ye et al. [57] were selected as the training sets, and one experimental test with a total of 3,839 datasets was selected as the testing set. The corresponding experimental test datasets are attached at Appendix B. The cyclic loading conditions and relevant parameters are presented in Table 3. Not that it is regret only undrained test data available on this sand.

4.1.2 Simulation strategy

The simulation strategy for Toyoura sand is similar to that in the framework presented in Fig. 3. Seven input variables are taken into consideration, and two output variables include mean effective stress and axial strain. The fixed parameters for each test, including the relative density D_r, the average deviatoric stress q_{ave} and the amplitude of the cyclic deviatoric stress q_{cyc} are selected, because these three variables significantly affect the cyclic behaviour of sands. The remaining four input variables are updated at each step. The LSTM-based model for describing the cyclic behaviour of Toyoura sand under the undrained condition at the *i*th step is given by:

$$\left(p'^{i+1}, \varepsilon_1^{i+1}\right) = f\left(D_r, q_{ave}, q_{cyc}, N_{cyc}^{i+1}, q^{i+1}, p'^{i}, \varepsilon_1^{i}\right) \tag{15}$$

where p'^{i+1}, εi+1 1, *Ni+1 cyc* and q^{i+1} = mean effective stress, axial strain, the number of loading cycles and deviatoric stress at the (*i*+1)th step, where *Ni*+1 *cyc* is similar to the labels for classifying the loading stage in Eqs. [7]–[8]; q^{i+1} is selected as an input parameter, because these are stress-control cyclic triaxial tests. In each test, deviatoric stress q is imposed continuously until liquefaction occurs, and meanwhile the N_{cyc} reaches the maximum value. p'^{i} and εi 1 = the predicted mean effective stress and axial strain at the (*i*–1)th step. f = the LSTM-based model for Toyoura sand. It should note that parameter m presented in Eqs. [7]–[8] for controlling experiment types is discarded, because only undrained results were

included in this experimental test.

To eliminate the effect of the scale of input variables on the model performance, q_{ave}, q_{cyc}, q and p' are divided by the corresponding initial effective normal stress p'_0 in each test during the model training phase. Meanwhile all input and output variables are mapped into the same interval (–1, 1) using Eq. [9].

4.1.3 Results

Fig. 10 shows the predicted mean effective stress p^i and axial strain ε_1 using the LSTM-based model for the training and testing sets, compared with the actual results. The corresponding parameter values of the constitutive model are presented in Table 4, where the methods for determining these parameters are consistent with the results presented in section 3.2. For the training set, MAE and MAPE values were closer to zero, showing perfect agreement with the actual results. Predicted p' and ε_1 for the testing sets using the LSTM-based model also showed great agreement with the experimental results, and all datasets were close to the line with the slope of 1. The largest error occurred when predicting the p' in the testing set, where the value of p' ranged from 50 to 70 kPa, and the LSTM model slightly underestimated the magnitude of mean effective stress degradation.

To clearly reveal the prediction performance of the LSTM-based model, the predicted evolution of effective normal stress p^i and axial strain ε_1 for testing set B2 is presented in Fig. 11. The simulated results for the training sets are presented as supplementary information for brevity. It can be observed from Fig. 10(a) that the mean effective stress decreased to 50 kPa and 70 kPa in the actual and predicted results, respectively, when liquefaction occurred. The ultimate mean effective stress in both actual and predicted results decreased to zero. As shown in Figs. 11(c) and (d), the predicted evolution of the axial strain and mean effective stress degradation exhibited perfect agreement with the experimental results. In particular, it can be seen from Fig. 11(b) that the largest axial strain caused by flow liquefaction could be

accurately captured by the LSTM-based model. Overall, the main sand behaviour under cyclic loading, including cyclic mobility (see Fig. 11(a)), large deformation induced by flow liquefaction (see Figs. 11(b-c)) and effective normal stress degradation could be captured (see Fig. 11(d)) by the LSTM-based model. Such factors indicate the reliability of application of the LSTM model to simulate the cyclic behaviour of Toyoura sand.

4.2 Undrained and drained cyclic simple shear tests for Fontainebleau sand

4.2.1 Data source

To investigate the performance of LSTM-based models on the simultaneous simulation of soil cyclic behaviour under drained and undrained conditions, a series of stress-control drained and undrained simple shear experimental tests on Fontainebleau sand under cyclic loading, conducted by Wu et al. [61], were used to validate the proposed models.

Herein, 24 experimental tests (18 undrained tests and 6 drained tests) with a total of 283,726 datasets from Wu et al. [61] were selected as the training sets, and 2 experimental tests with a total of 15,000 datasets were selected as the testing sets. The corresponding experimental test datasets are attached at Appendix B. Detailed parameters of these experimental results are presented in Table 5.

4.2.2 Simulation strategy

Compared with the LSTM-based model for Toyoura sand, the corresponding model for Fontainebleau sand has an additional parameter m for controlling drained or undrained conditions. Therefore, the four fixed parameters for each test include the initial void ratio e_0, the average sheer stress τ_{ave}, the cyclic shear stress amplitude τ_{cyc} and the code m. The remaining input variables are updated at each step. The LSTM-based model for describing the cyclic behaviour of Fontainebleau sand under drained and undrained conditions at the ith step is given by:

$$\left(\sigma_n^{'i+1}, \gamma^{i+1}, e^{i+1}\right) = f\left(e_0, \tau_{\text{ave}}, \tau_{\text{cyc}}, m, N_{\text{cyc}}^{i+1}, \tau^{i+1}, \sigma_n^{'i}, \gamma^i, e^i\right) \quad (16)$$

where $\sigma'i+1\ n$, γ^{i+1}, e^{i+1}, τ^{i+1} and $Ni+1\ cyc$ = the effective normal stress, shear strain, void ratio, number of loading cycles and shear stress at the (i+1)th step; $\sigma'i\ n$, γ^i and e^i = the predicted effective normal stress, shear strain and void ratio at the (i–1)th step. f = the LSTM-based model for Fontainebleau sand. τ_{ave}, τ_{cyc}, τ and $\sigma'i\ n$ are also divided by the corresponding initial effective normal stress $\sigma'i\ n$ in each test during the model training phase. Meanwhile all input and output variables are mapped into the same interval (–1, 1) using Eq. [9]. Note that the $\sigma'i\ n$ holds constant under the drained condition, and e holds constant under the undrained condition. Such constant variables can be discarded when training the LSTM-based model. Therefore, the output variables are γ^{i+1} and e^{i+1} under the drained condition, and the input variables are e_0, τ_{ave}, τ_{cyc}, m (1), τ^{i+1}, γ^i and e^i. For the undrained condition, $\sigma'i+1\ n$ and γ^{i+1} are the output variables, and e_0, τ_{ave}, τ_{cyc}, m (0), τ^{i+1}, $\sigma'i\ n$ and γ^i are the input variables. Both drained and undrained data are integrated to train the LSTM-based model, and the number of input and output variables is eight and two, respectively.

4.2.3 Results

Fig. 12 shows the predicted effective normal stress $\sigma'\ n$, shear strain γ and void ratio e for the training and testing sets using the LSTM-based model, compared with the actual results. The corresponding values of parameters in said model are presented in Table 6, where the methods for determining these parameters are consistent with the results presented in section 3.2. In this study, considering the complexity of simultaneous simulation of soil cyclic behaviour under drained and undrained conditions, a cyclical learning rate for training a deep learning based model proposed by Smith [56] were used . The learning rate (the default value is 0.001) of the LSTM-based model first increased from 0.0002 to 0.002 within 10 epochs and then decreased from 0.002 to 0.0002 within 10 epochs; thus each round included 20 epochs. Such a method effectively enhances the global optimization capability of the LSTM-based model. For the

training sets, unlike the predicted results of Toyoura sand, MAE and MAPE values were not equal to zero. This indicates that the simultaneous simulation of cyclic behaviour of sands under drained and undrained conditions is much more challenging than simulation under the single drained condition. Overall, the predicted $\sigma' n$, γ and e for the training and testing sets using the LSTM-based model were in great agreement with the experimental results.

To clearly reveal the performance of the LSTM-based model in simultaneously modelling soil cyclic behaviour under drained and undrained conditions, the predicted evolution of shear strain γ and void ratio e for the testing set CD3 (drained) is presented in Fig. 13, and the predicted evolution of effective normal stress $\sigma' n$ and shear strain γ for the testing set C34 (undrained) is presented in Fig. 14. Meanwhile, the simulated results for the training sets are presented as supplementary information for brevity. Comparing the predicted and measured results for the drained condition in Figs. 13(a–b), the shear accumulation and densification trends could be identified by the LSTM-based model. Meanwhile the predicted evolution of shear accumulation and densification as the loading cycles increased showed great agreement with the measured results, as presented in Figs. 13(c–d). The ultimate measured shear strain ranged from 0.01% to 0.58%, and the predicted values ranged from 0.07% to 0.60%. The ultimate measured and predicted void ratios were 0.562 and 0.560, respectively. The excellent agreement between measured and predicted results indicates that the cyclic behaviour of Fontainebleau sand under the drained condition can be captured by the LSTM-based model.

It can be seen from Fig. 14 that the number of loading cycles reached 100 as the liquefaction occurred. It can be observed that the LSTM-based model could perfectly predict the change in the effective normal stress and shear strain at the first 80 loading cycles. Then, the predicted effective normal stress and shear strain slightly deviated from the measured results because of the development of

liquefaction. The measured effective normal stress ultimately reached a residual value of 52 kPa from the initial 416 kPa, and the shear strain reached 3.5%. The predicted residual effective normal stress and the shear strain using the LSTM-based model were 70 kPa and 2.7%, respectively. The deviation between predicted and measured residual effective normal stress and shear strain can be primarily attributed to inadequate datasets collected after liquefaction, which remarkably affect the performance of the LSTM-based model in predicting sand behaviour near to the liquefaction. Overall, the instability corresponding to a cyclic mobility mechanism and the degradation of effective normal stress could be captured by the LSTM-based model. All results indicate that the LSTM-based model is reliable for simultaneously modelling the cyclic behaviour of Fontainebleau sand under drained and undrained conditions.

5. Conclusions

As pioneer, a LSTM-based general-purpose constitutive model was proposed for simultaneously describing the cyclic behaviour of sands under drained and undrained conditions. Synthetic data under both drained and undrained conditions, generated by an advanced constitutive model, were first generated to explore an appropriate framework for the LSTM-based model. The LSTM-based models were thereafter employed to simulate the cyclic behaviour of Toyoura sand under the undrained condition and Fontainebleau sand under undrained and drained conditions, compared with the experimental data.

For building a LSTM-based model, the input variables can be generally categorized into four categories: fixed parameters related to soil properties and loading condition, labels for marking the number of loading cycles, stress or strain increment in stress or strain-control strategy and predicted outputs at the previous step. Then, an incremental strategy is used to train the LSTM-based model; that is,

the predicted outputs at the current step are used as the input variables at the next step.

Compared with the synthetic data, the LSTM-based model can simultaneously predict the cyclic behaviour of sand under drained and undrained conditions including the dilatant behaviour for dense soils, contractive behaviour for loose soils and the phase transformation from soil dilation to contraction. Compared with the actual experimental results, the instability corresponding to a cyclic mobility mechanism, the degradation of effective normal stress and flow liquefaction under the undrained condition, the shear strain accumulation, and densification under the drained condition can be accurately predicted by the LSTM-based model. Such results indicate that LSTM is a novel and effective means of exploring soil-constitutive models. As the increasing collection of datasets, future study will continuously enhance and explore the performance of the LSTM based model such as development a uniform model to capture the behaviour of different materials.

Acknowledgements

This research was financially supported by the Research Grants Council (RGC) of Hong Kong Special Administrative Region Government (HKSARG) of China (Grant No.: 15209119, PolyU R5037-18F) and the National Natural Science Foundation of China (Grant No. 41727802, 41630633).

Appendix A. Introduction of activation functions

$$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{17}$$

$$ReLU(x) = \begin{cases} x, & x > 0 \\ 0, & x \le 0 \end{cases} \tag{18}$$

The activation function is an important part of neural networks. A neural network without an activation

function is merely akin to a linear regression model. *tanh* and *ReLU* are commonly used activation functions in various neural networks (see Fig. 15). The saturated activation function *tanh* is a continuous function and maps the input variables into the range of (0, 1) and (–1, 1), but gradients of this activation function can vanish or explode, and the computational cost is huge. The non-saturated activation function *ReLU* effectively overcomes these two issues, but the issue with this activation function is that the neuron is not activated if the input variable is a large negative value. It should be noted that there is no direct rule for scenario-based activation function selection. The decision for selecting the most appropriate activation function is completely reliant on the properties of the problem.

Appendix B. Datasets used in this study

The datasets for Toyoura sand under the undrained condition, and Fontainebleau sand under undrained and drained conditions, can be downloaded from the following links, respectively:

https://www.researchgate.net/publication/336606047_Data_for_Toyoura_sand_under_undrained_cyclic_loading?_sg=F8zxKEOa3U-bjqUJViEuScrdyfnElRCB79xD-rtEK6Dx56tiVw7nFltb6pMUGG8laBihTSk-qHsu3aVjXQO55tNyLJGWOkIfUUxlubM.2WrOaumOEzcRVbGAwx3Y49NOe2MRIz-ZOF1bytf6e1k2Y3zugitNk4JCudg0uneh0vthqDSaspFs7y10UT45GQ

https://www.researchgate.net/publication/336612973_Data_for_Fontainebleau_sand_under_undrained_and_drained_cyclic_loading?_sg=WJt9uvyvX26O0DlduF1rRcHBWVu1N-_NwoxeZ90ut9ECrshmxxMAJscmqK5_wPSuNOY0nSn0pOB89aZbrXSiM93Hr9o1xSGxQ4iZxwUY.F8TAK83V00npgZhfaB4Ay9V_ZfF5V8zHMBW-mmrDwlzzCTabuSts0uT-d62KmeTahN0RNt6y_ybr2rNVV4SZzQ

References

1. Jin Z, Yin Z-Y, Kotronis P, Li Z, Tamagnini C. A hypoplastic macroelement model for a caisson foundation in sand under monotonic and cyclic loadings. *Mar Struct.* 2019;66:16-26.
2. Zhu F, Bienen B, O'Loughlin C, Morgan N, Cassidy MJ. The response of suction caissons to multidirectional lateral cyclic loading in sand over clay. *Ocean Eng.* 2018;170:43-54.
3. Shogaki T, Nakano Y, Yoshizu T. Estimating in situ dynamic strength and deformation properties of Niigata East Port sand deposits. *Soils Found.* 2017;57(6):947-964.
4. Esmaeili M, Aela P, Hosseini A. Experimental assessment of cyclic behavior of sand-fouled ballast mixed with tire derived aggregates. *Soil Dyn Earthq Eng.* 2017;98:1-11.
5. Pradhan TBS, Tatsuoka F, Sato Y. Experimental stress-dilatancy relations of sand subjected to cyclic loading. *Soils Found.* 1989;29(1):45-64.
6. LóPez-Querol S, Coop MR. Drained cyclic behaviour of loose Dogs Bay sand. *Géotechnique.* 2012;62(4):281-289.
7. Georgiannou VN, Tsomokos A, Stavrou K. Monotonic and cyclic behaviour of sand under torsional loading. *Géotechnique.* 2008;58(2):113-124.
8. Wijewickreme D, Sriskandakumar S, Byrne P. Cyclic loading response of loose air-pluviated Fraser River sand for validation of numerical models simulating centrifuge tests. *Can Geotech J.* 2005;42(2):550-561.
9. Hu X, Zhang Y, Guo L, et al. Cyclic behavior of saturated soft clay under stress path with bidirectional shear stresses. *Soil Dyn Earthq Eng.* 2018;104:319-328.
10. Wichtmann T, Fuentes W, Triantafyllidis T. Inspection of three sophisticated constitutive models based on monotonic and cyclic tests on fine sand: Hypoplasticity vs. Sanisand vs. ISA. *Soil Dyn Earthq Eng.* 2019;124:172-183.
11. Duncan JM, Chang CY. Nonlinear analysis of stress and strain in soils. *Journal of Soil Mechanics and Foundations Engineering Division.* 1970;96(5):1629-1653.
12. Darve F, Labanieh S. Incremental constitutive law for sands and clays: simulations of monitonic and cyclic tests. *Int J Numer Anal Met.* 1982;6:243-275
13. Yin Z-Y, Wu Z-X, Hicher P-Y. Modeling monotonic and cyclic behavior of granular materials by exponential constitutive function. *J Eng Mech-ASCE.* 2018;144(4).
14. Niemunis A, Herle I. Hypoplastic model for cohesionless soils with elastic strain range. *Mechanics of Cohesive Frictional Materials.* 1997;2:279-299.
15. Wu W, Bauer E, Kolymbas D. Hypoplastic constitutive model with critical state for granular materials. *Mech Mater.* 1996;23:45-69.
16. Taiebat M, Dafalias YF. SANISAND: Simple anisotropic sand plasticity model. *Int J Numer Anal Met.* 2008;32(8):915-948.
17. Yao YP, Sun DA, Matsuoka H. A unified constitutive model for both clay and sand with hardening parameter independent on stress path. *Comput Geotech.* 2008;35(2):210-222.
18. Jin Y, Yin Z, Zhang D, Huang H. Unified modeling of the monotonic and cyclic behaviors of sand and clay. *Acta Mech Solida Sin.* 2015;28(2):111-132.
19. Ye B, Ye G, Zhang F. Numerical modeling of changes in anisotropy during liquefaction using a generalized constitutive model. *Comput Geotech.* 2012;42:62-72.
20. Zhang F, Ye B, Noda T, Nakano M, Nakai K. Explanation of cyclic mobility of soils: approach by stress-induced anisotropy. *Soils Found.* 2007(47):635-648.

21. Zhang F, Ye B, GL. Y. Unified description of sand behavior. *Frontiers of Architecture and Civil Engineering in China.* 2011;5(2):121-150.
22. Yin Z-Y, Chang CS. Microstructural modelling of stress-dependent behaviour of clay. *Int J Solids Struct.* 2009;46(6):1373-1388.
23. Yin Z-Y, Chang CS, Hicher P-Y. Micromechanical modelling for effect of inherent anisotropy on cyclic behaviour of sand. *Int J Solids Struct.* 2010;47(14-15):1933-1951.
24. Zhang P, Chen RP, Wu HN. Real-time analysis and regulation of EPB shield steering using Random Forest. *Automat Constr.* 2019;106:102860.
25. Chen RP, Zhang P, Kang X, Zhong ZQ, Liu Y, Wu HN. Prediction of maximum surface settlement caused by EPB shield tunneling with ANN methods. *Soils Found.* 2019;59:284–295.
26. Zhang P, Yin Z-Y, Jin Y-F, Chan THT. A novel hybrid surrogate intelligent model for creep index prediction based on particle swarm optimization and random forest. *Eng Geol.* 2020;265:105328.
27. Yin Z-Y, Jin Y-F, Shen JS, Hicher P-Y. Optimization techniques for identifying soil parameters in geotechnical engineering: Comparative study and enhancement. *Int J Numer Anal Met.* 2018;42(1):70-94.
28. Yin Z-Y, Jin Y-F, Shen S-L, Huang H-W. An efficient optimization method for identifying parameters of soft structured clay by an enhanced genetic algorithm and elastic–viscoplastic model. *Acta Geotech.* 2016;12(4):849-867.
29. Jin Y-F, Yin Z-Y, Shen S-L, Hicher P-Y. Selection of sand models and identification of parameters using an enhanced genetic algorithm. *Int J Numer Anal Met.* 2016;40(8):1219-1240.
30. Jin YF, Yin ZY, Wu ZX, Zhou WH. Identifying parameters of easily crushable sand and application to offshore pile driving. *Ocean Eng.* 2018;154:416-429.
31. Ghaboussi J, Garret J, Wu X. Knowledge-based modeling of material behavior with neural networks. *J Eng Mech-ASCE.* 1991;117(1):132-153.
32. Habibagahi G, Bamdad A. A neural network framework for mechanical behavior of unsaturated soils. *Can Geotech J.* 2003;40(3):684-693.
33. Penumadu D, Zhao RD. Triaxial compression behavior of sand and gravel using artificial neural networks (ANN). *Comput Geotech.* 1999;24:207-230.
34. He S, Li J. Modeling nonlinear elastic behavior of reinforced soil using artificial neural networks. *Appl Soft Comput.* 2009;9(3):954-961.
35. Rashidian V, Hassanlourad M. Application of an artificial neural network for modeling the mechanical behavior of carbonate soils. *Int J Geomech.* 2014;14(1):142-150.
36. Romo MP, Garcí a SR, Mendoza MJ, Taboada‐Urtuzuá stegui V. Recurrent and constructive‐algorithm networks for sand behavior modeling. *Int J Geomech.* 2001;1(4):371-387.
37. Kohestani VR, Hassanlourad M. Modeling the mechanical behavior of carbonate sands using artificial neural networks and support vector machines. *Int J Geomech.* 2016;16(1):04015038.
38. Basheer IA. Selection of methodology for neural network modeling of constitutive hystereses behavior of soils. *Comput-Aided Civ Inf.* 2000;15:440–458.
39. Brewick PT, Masri SF, Carboni B, Lacarbonara W. Data-Based Nonlinear Identification and Constitutive Modeling of Hysteresis in NiTiNOL and Steel Strands. *J Eng Mech-ASCE.* 2016;142(12).
40. Yun GJ, Ghaboussi J, Elnashai AS. A new neural network-based model for hysteretic behavior of materials. *International Journal for Numerical Methods in Engineering.* 2008;73(4):447-469.

41. Yun GJ, Ghaboussi J, Elnashai AS. Self-learning simulation method for inverse nonlinear modeling of cyclic behavior of connections. *Comput Method Appl M.* 2008;197(33-40):2836-2857.
42. Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Comput.* 1997;9(8):1735–1780.
43. Yang B, Yin K, Lacasse S, Liu Z. Time series analysis and long short-term memory neural network to predict landslide displacement. *Landslides.* 2019.
44. Xu S, Niu R. Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. *Comput Geosci-UK.* 2018;111:87-96.
45. Wang K, Sun W. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. *Comput Method Appl M.* 2018;334:337-380.
46. Zhang R, Chen Z, Chen S, Zheng J, Büyüköztürk O, Sun H. Deep long short-term memory networks for nonlinear structural seismic response prediction. *Comput Struct.* 2019;220:55-68.
47. Xu SL, Niu RQ. Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. *Comput Geosci-UK.* 2018;111:87-96.
48. Wang J, Yan J, Li C, Gao RX, Zhao R. Deep heterogeneous GRU model for predictive analytics in smart manufacturing: Application to tool wear prediction. *Computers in Industry.* 2019;111:1-14.
49. Zhou J, Xu W. End-to-end learning of semantic role labeling using recurrent neural networks. Paper presented at: International joint conference on Natural Language Processing2015; Beijing, China.
50. Orimoloye LO, Sung M-C, Ma T, Johnson JEV. Comparing the effectiveness of deep feedforward neural networks and shallow architectures for predicting stock price indices. *Expert Syst Appl.* 2020;139.
51. Zhang P. A novel feature selection method based on global sensitivity analysis with application in machine learning-based prediction model. *Appl Soft Comput.* 2019;85:105859.
52. Chen RP, Zhang P, Wu HN, Wang ZT, Zhong ZQ. Prediction of shield tunneling-induced ground settlement using machine learning techniques. *Front Struct Civ Eng.* 2019;13(6):1363–1378.
53. Goodfellow I, Bengio Y, Courville A. *Deep learning.* MIT; 2016.
54. Ruder S. An overview of gradient descent optimization. 2016:blog: http://sebastianruder.com/optimizing-gradient-descent/index.html.
55. Kingma DP. Adam: a method for stochastic optimization. International Conference on Learning Representations; 2015; San Diego, CA.
56. Smith LN. Cyclical learning rates for training neural networks. IEEE Winter Conference on Applications of Computer Vision (WACV); 2017; Santa Rosa, California.
57. Ye B, Ni X, Ye G, Huang Y, Lu P. Prediction of the initial point of the last cycle in undrained cyclic triaxial tests on flow liquefaction. *Soil Dyn Earthq Eng.* 2019;120:12-22.
58. Andersen KH. Bearing capacity under cyclic loading — offshore, along the coast, and on land. The 21st Bjerrum Lecture presented in Oslo, 23 November 2007. *Can Geotech J.* 2009;46(5):513-535.
59. Gavin K, Igoe D, Doherty P. Piles for offshore wind turbines: a state-of-the-art review. *Proceedings of the Institution of Civil Engineers - Geotechnical Engineering.* 2011;164(4):245-256.
60. Fuentes W, Tafili M, Triantafyllidis T. An ISA-plasticity-based model for viscous and non-viscous clays. *Acta Geotech.* 2017.
61. Wu Z-X, Dano C, Hicher P-Y, Yin Z-Y. Estimating normal effective stress degradation in sand under undrained simple shear condition. *European Journal of Environmental and Civil Engineering.*

2018:1-20.

Table

Table 1 Values of parameters in the model after Yin et al (2018)

Parameter	G_0/ kPa	υ	N	φ_c/ °	e_{c0}	Λ	ξ	d	p'/ kPa
Value	3000	0.3	0.67	31.2	0.937	0.022	0.71	2	98

Table 2 Values of parameters in LSTM-based model

Parameter	Strain-control	Stress-control
Time_step	1	1
Architecture	90–40–2	90–30–2
Epoch	600	750
Batch_size	1111	400
Validation_split	0.1	0.1
Activation function	*ReLU*, *tanh, linear*	*ReLU*, *tanh, linear*
Optimizer	*Adam*	*Adam*

Remark: activation functions in each layer are as followings: 90(*ReLU*)–40(*tanh*)–2(*linear*); 90(*ReLU*)–60(*tanh*)–70(*tanh*)–2(*linear*).

Table 3 Adopted laboratory tests on Toyoura sand

	Test No.	Loading type	Dr	p_0'	q_{ave}	q_{cyc}
Training sets	B1	Symmetrical loading	10%	150	0	23
	B3	Symmetrical loading	30%	150	0	32
	B4	Non-symm stress reversal	10%	150	9	23
	B5	Non-symm stress reversal	20%	150	9	27
	B6	Non-symm stress reversal	30%	150	9	32
Testing set	B2	Symmetrical loading	20%	150	0	27

Remark: Initial void ratio corresponding to initial mean effective stress p'_0.

Table 4 Values of parameters in LSTM-based model for Toyoura sand

Parameter	Toyoura sand
Time_step	1
Architecture	80–80–80–2
Epoch	100
Batch_size	4500
Validation_split	0.1
Activation function	*tanh, ReLU, linear*
Optimizer	*Adam*

Remark: activation functions in each layer are as followings: 80(*tanh*)–80(*ReLU*)–80(*ReLU*)–2(*linear*)

Table 5 Adopted laboratory tests on Fontainebleau sand

	Test No.	e_0	σ'_{n0}	τ_{ave}	τ_{cyc}	N_L
Training set	CU13	0.658	104	0	5.2	18
	CU14	0.664	208	0	5.2	88
	CU17	0.663	312	0	15.6	33
	CU18	0.662	416	0	10.4	366
	CU19	0.660	416	0	10.4	300
	CU21	0.663	416	0	20.8	62
	CU22	0.656	416	0	20.8	62
	CU23	0.655	416	0	31.2	10
	CU27	0.654	416	5.2	20.8	57
	CU28	0.669	416	10.4	10.4	368
	CU29	0.663	416	10.4	20.8	19
	CU31	0.641	416	10.4	20.8	56
	CU32	0.653	416	20.8	10.4	240
	CU33	0.641	416	20.8	10.4	310
	CU35	0.666	416	20.8	20.8	45
	CU36	0.640	416	20.8	41.6	7
	CU37	0.654	416	41.6	10.4	330
	CU38	0.664	416	41.6	20.8	19
	CD1	0.668	416	41.6	0	/
	CD2	0.628	416	41.6	0	/
	CD4	0.539	416	41.6	0	/
	CD5	0.623	52	5.2	0	/
	CD6	0.627	104	10.4	0	/
	CD7	0.619	208	20.8	0	/
Testing set	CU34	0.635	416	20.8	20.8	100
	CD3	0.568	416	41.6	0	/

Remark: Initial void ratio corresponding to initial effective normal stress σ'_{n0}; N_L - number of cycles to liquefaction.

Table 6 Values of parameters in LSTM-based model for Fontainebleau sand

Parameter	Value
Time_step	1
Architecture	60–50–2
Epoch	100
Batch_size	4000
Validation_split	0.1
Activation function	*tanh, linear*
Optimizer	*Adam*
Learning rate	0.0002–0.002

Remark: activation functions in each layer are as followings: 60(*tanh*)–50(*tanh*)–2(*linear*)

Figure caption

Fig. 1 Schematic view of ML algorithm: (a) LSTM; (b) memory cell

Fig. 2 A hysteresis loop with labels used in classifying the loading stage: (a) strain-control cyclic loading; (b) stress-control cyclic loading

Fig. 3 Framework of LSTM-based model

Fig. 4 Framework of training LSTM-based model

Fig. 5 Convergence curve of LSTM-based model

Fig. 6 Comparison between predicted and actual strain-control results using LSTM-based model for drained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

Fig. 7 Comparison between predicted and actual strain-control results using LSTM-based model for undrained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

Fig. 8 Comparison between predicted and actual stress-control results using LSTM-based model for drained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

Fig. 9 Comparison between predicted and actual stress-control results using LSTM-based model for undrained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

Fig. 10 Comparison between predicted and experimental results of Toyoura sand under undrained condition: (a) mean effective stress p'; (b) axial strain ε_1

Fig. 11 Cyclic response of Toyoura sand under undrained condition for the testing set: (a) p'–q; (b) ε_1–q; (c) N_{cyc}–ε_1; (d) N_{cyc} –p'

Fig. 12 Comparison between predicted and experimental results of Fontainebleau sand under undrained and drained conditions: (a) effective normal stress $\sigma_n^{'}$; (b) shear strain γ; (c) void ratio e

Fig. 13 Cyclic response of Fontainebleau sand under drained condition for the testing set: (a) τ–γ; (b) e–γ; (c) γ–N_{cyc}; (d) e–N_{cyc}

Fig. 14 Cyclic response of Fontainebleau sand under undrained condition for the testing set: (a) τ–$\sigma_n^{'}$; (b) τ–γ; (c) γ–N_{cyc}; (d) $\sigma_n^{'}$–N_{cyc}

Fig. 15 Curves of five activation functions

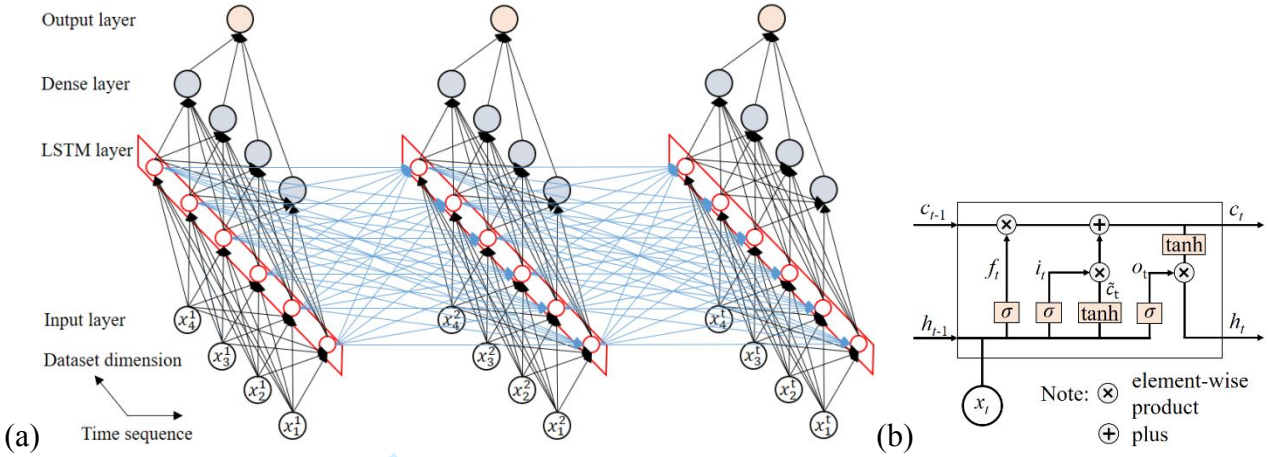

Fig. 1 Schematic view of ML algorithm: (a) LSTM; (b) memory cell

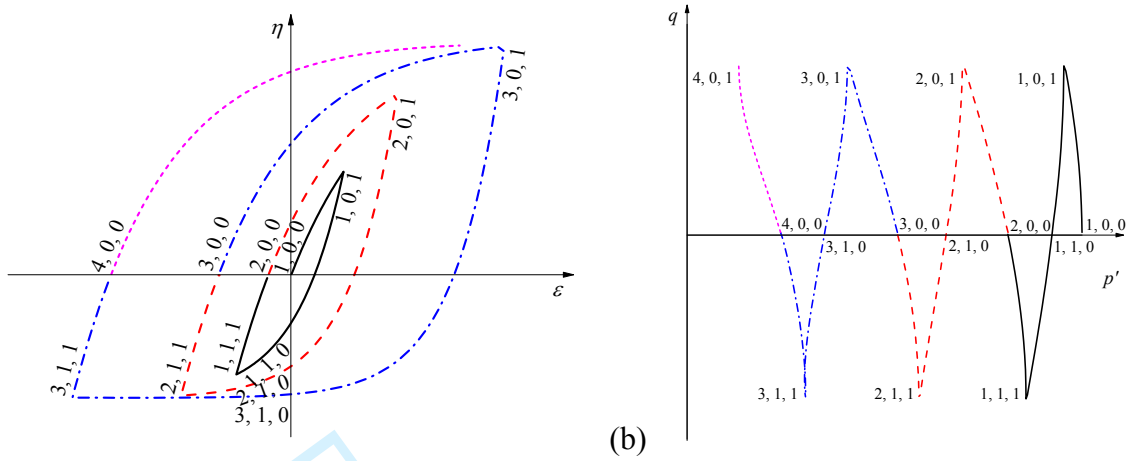

Fig. 2 A hysteresis loop with labels used in classifying the loading stage: (a) strain-control cyclic loading; (b) stress-control cyclic loading

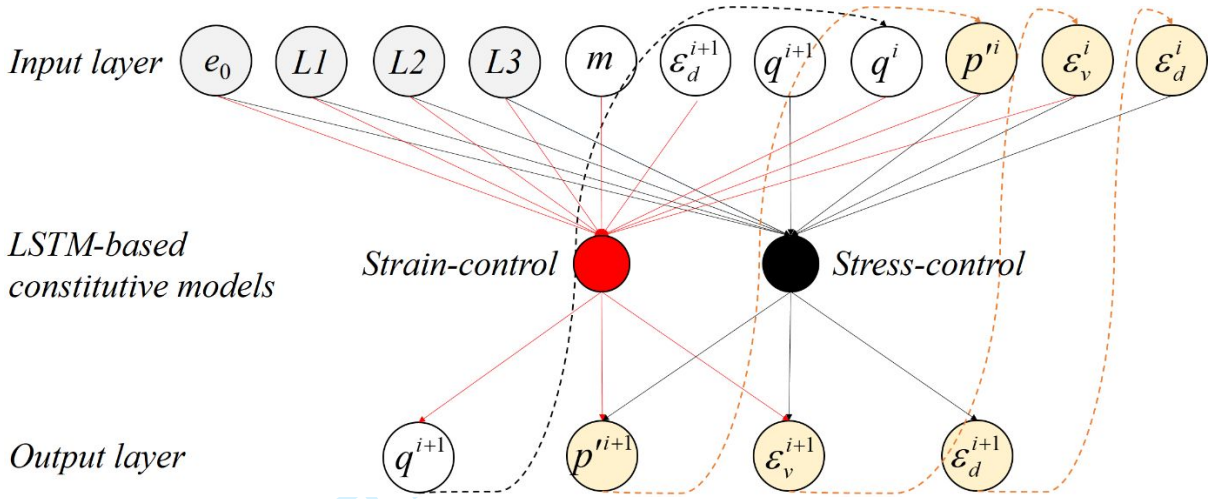

Fig. 3 Framework of LSTM-based model

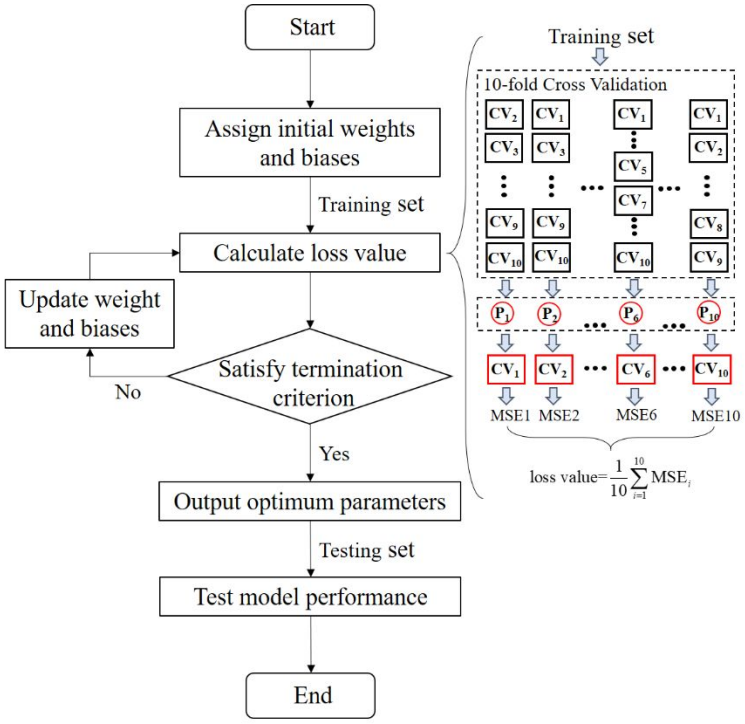

Fig. 4 Framework of training LSTM-based model

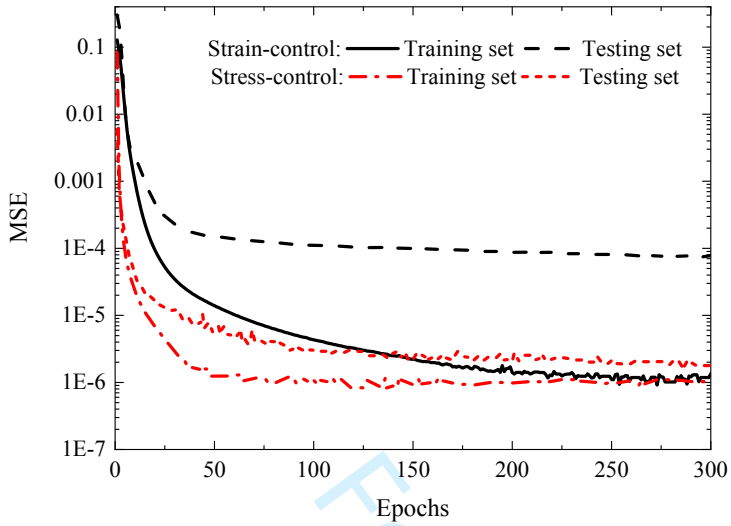

Fig. 5 Convergence curve of LSTM-based model

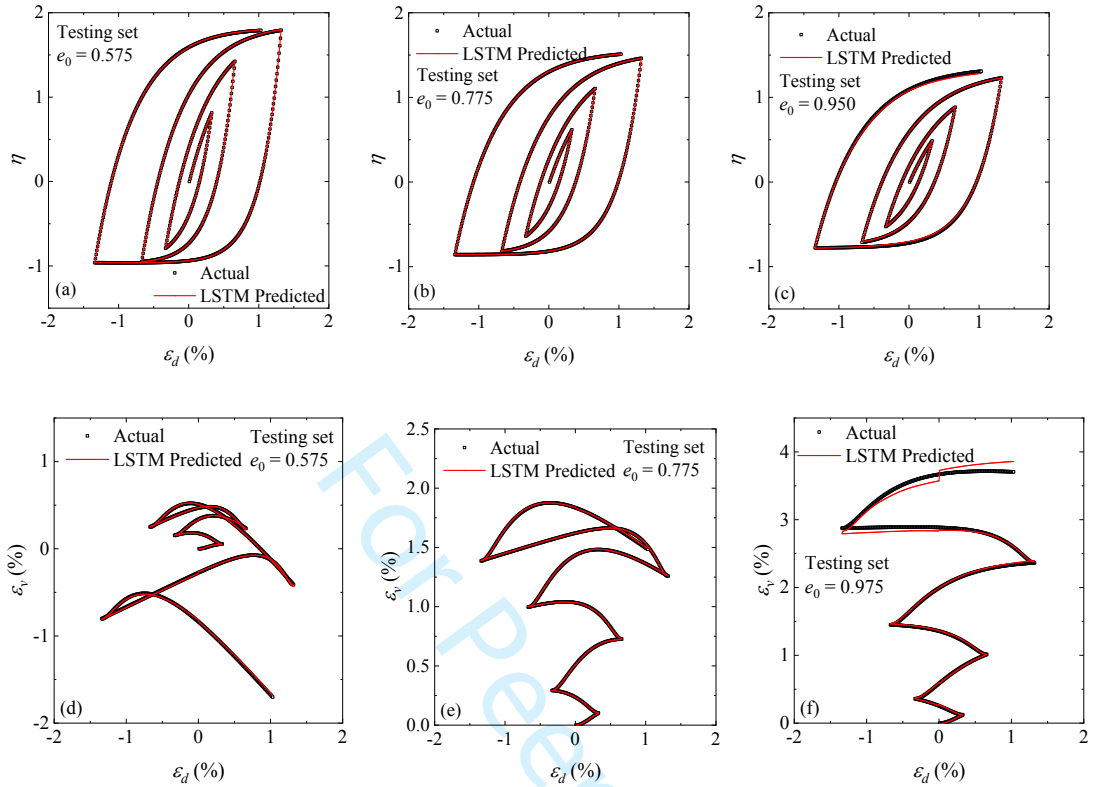

Fig. 6 Comparison between predicted and actual strain-control results using LSTM-based model for drained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

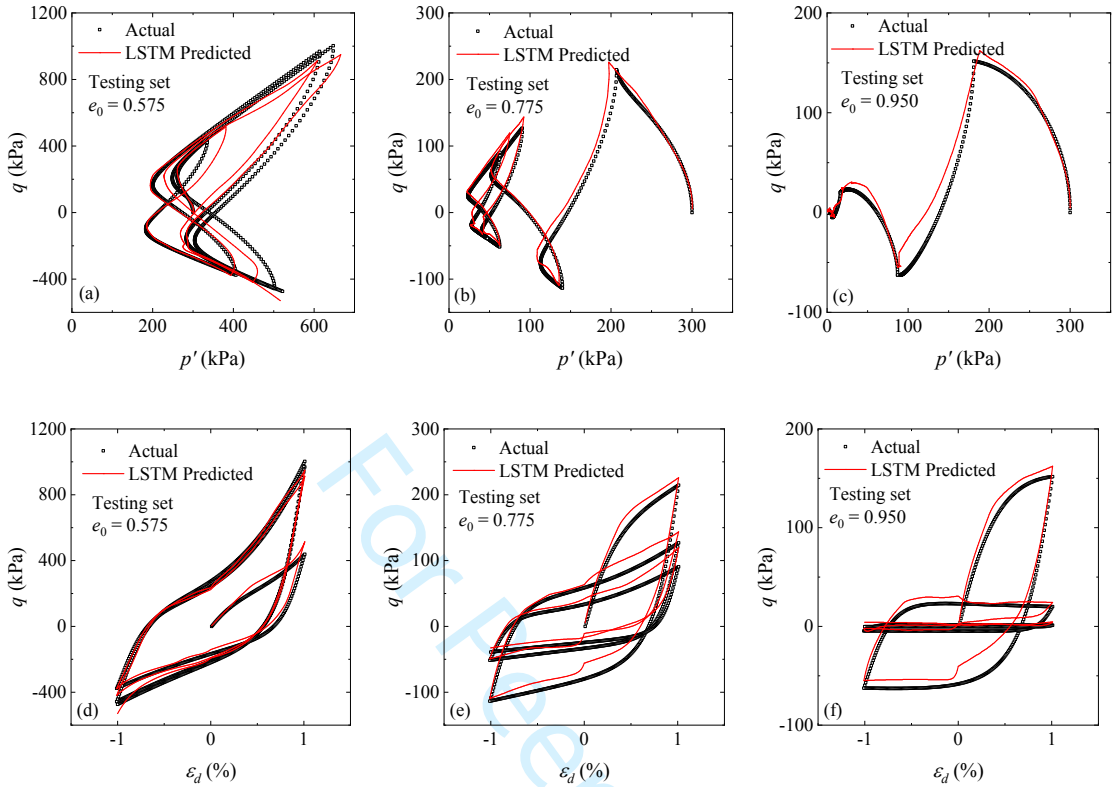

Fig. 7 Comparison between predicted and actual strain-control results using LSTM-based model for undrained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

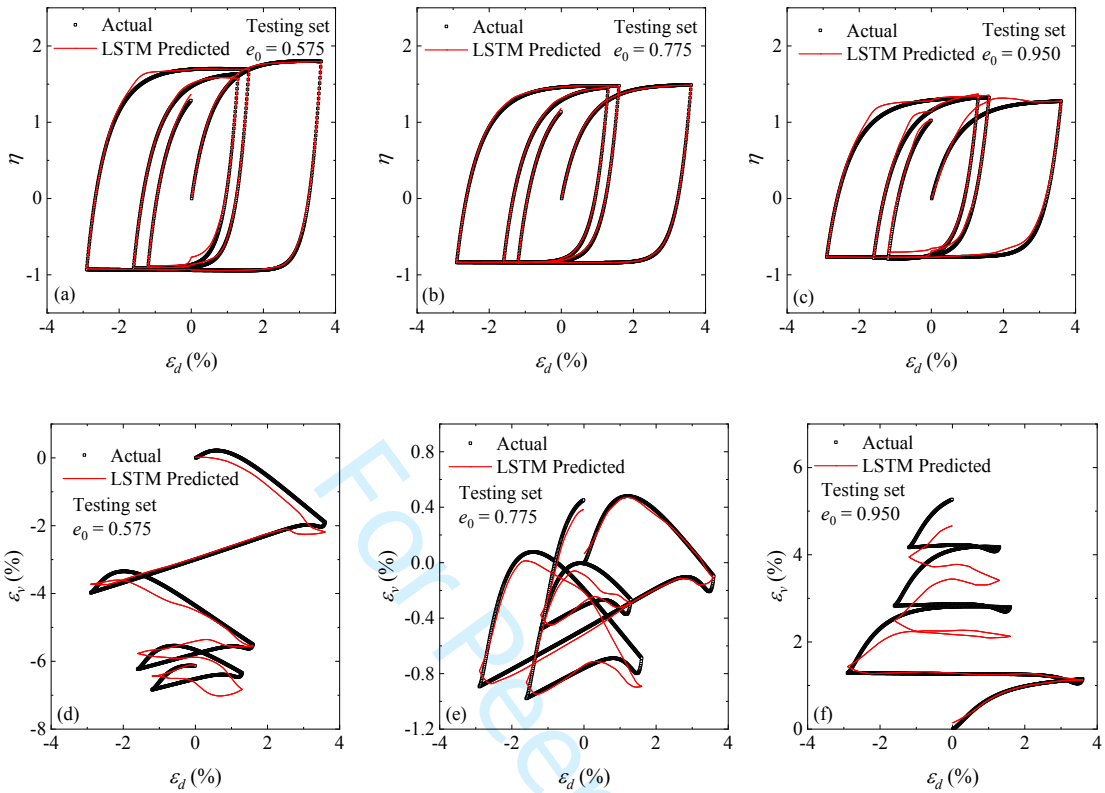

Fig. 8 Comparison between predicted and actual stress-control results using LSTM-based model for drained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

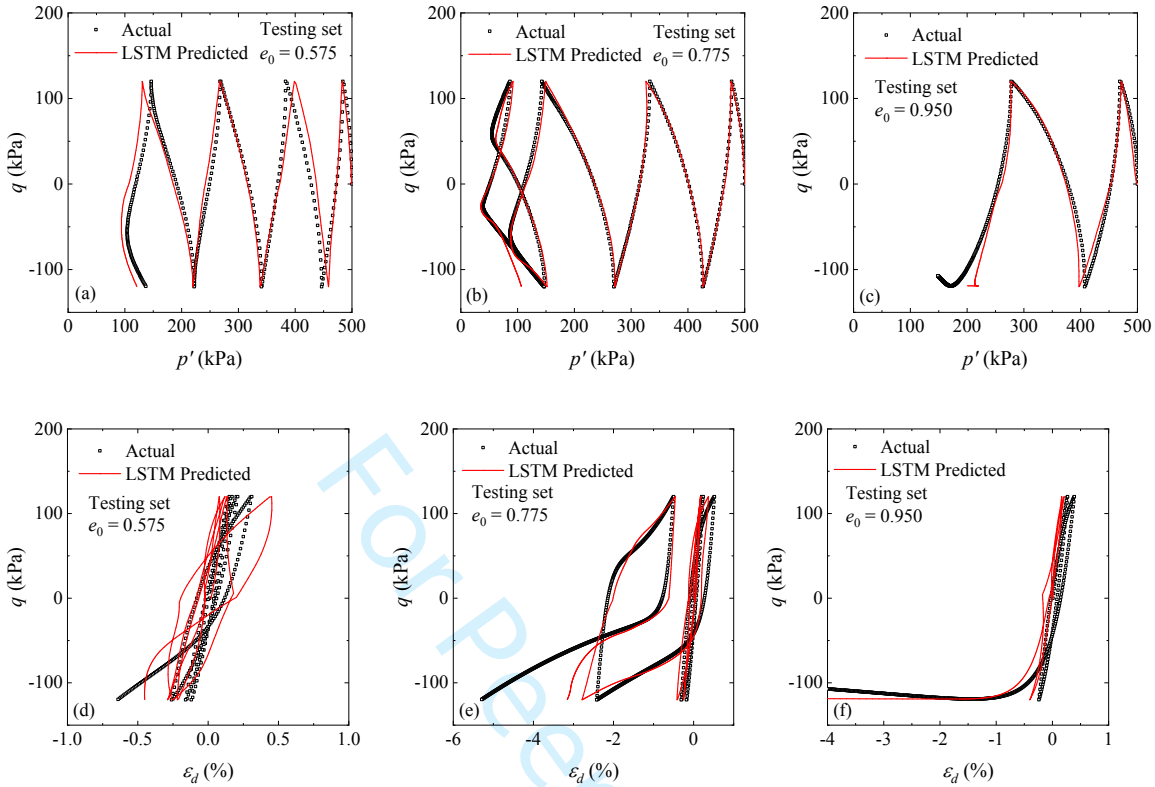

Fig. 9 Comparison between predicted and actual stress-control results using LSTM-based model for undrained triaxial tests: (a–c) stress ratio; (d–f) volumetric strain

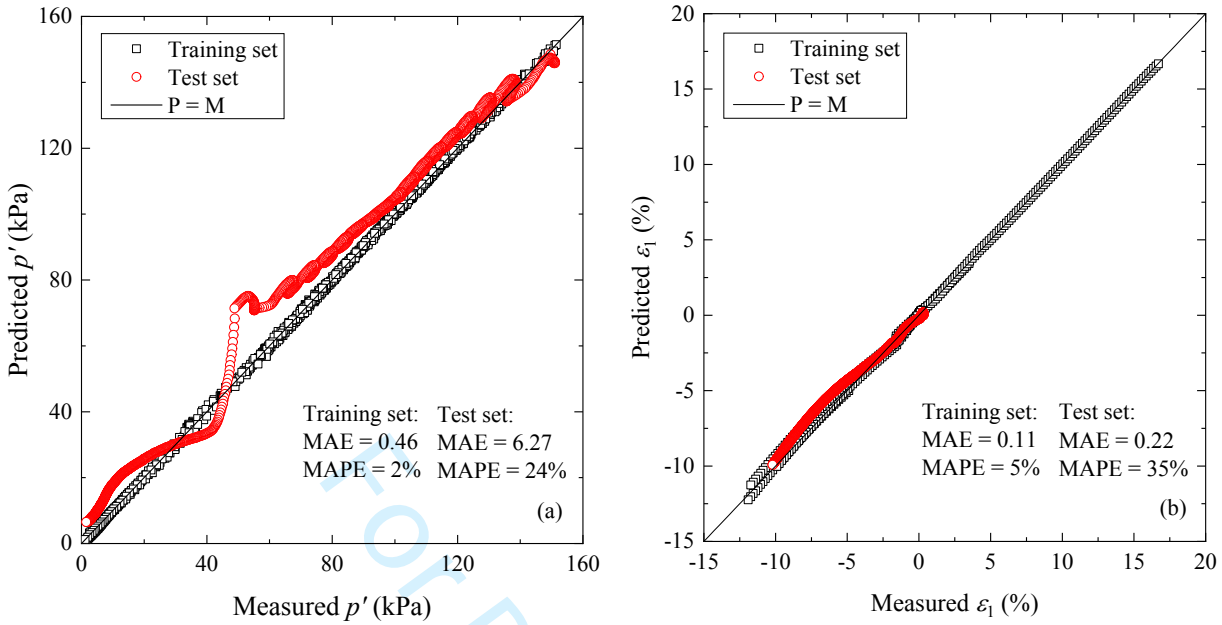

Fig. 10 Comparison between predicted and experimental results of Toyoura sand under undrained condition: (a) mean effective stress p'; (b) axial strain ε_1

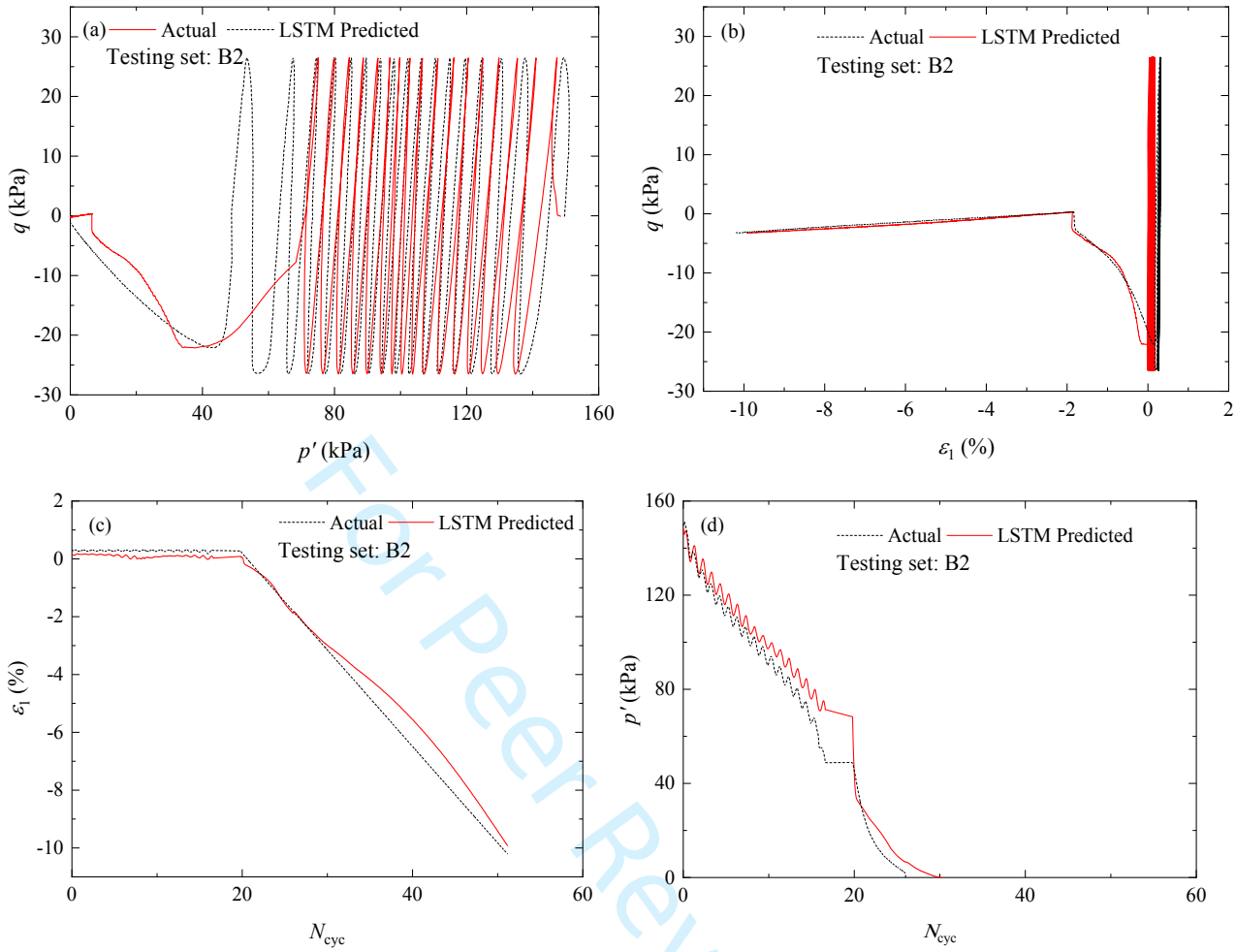

Fig. 11 Cyclic response of Toyoura sand under undrained condition for the testing set: (a) p'–q; (b) ε_1–q; (c) N_{cyc}–ε_1; (d) N_{cyc} –p'

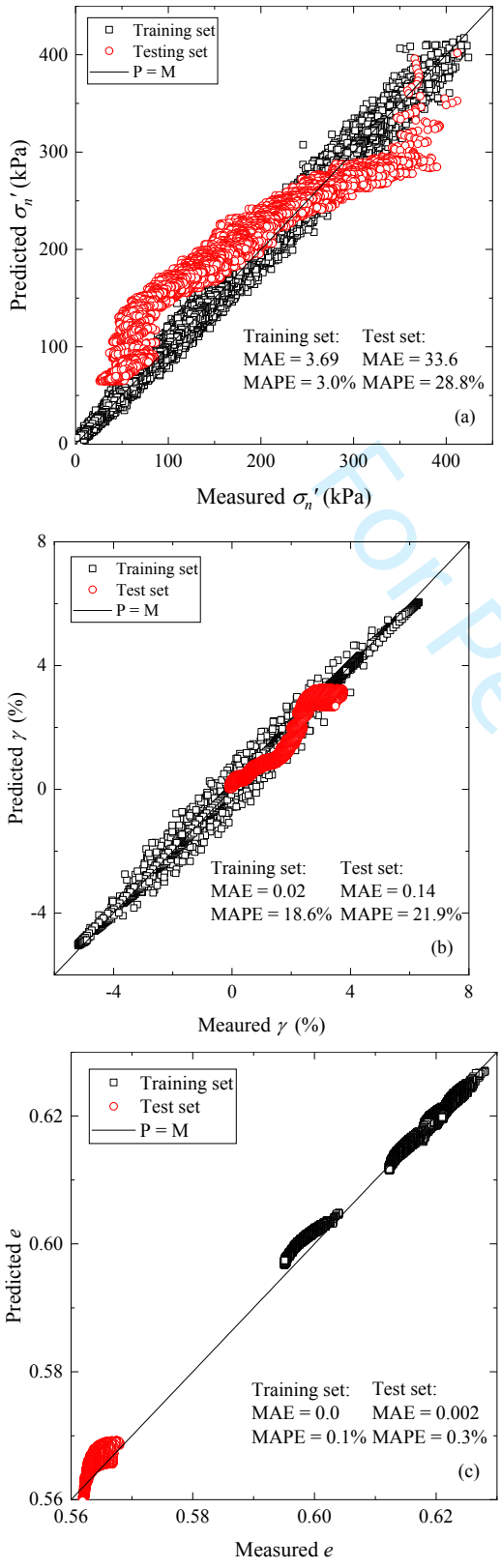

Fig.12 Comparison between predicted and experimental results of Fontainebleau sand under undrained and drained conditions: (a) effective normal stress σ_n'; (b) shear strain γ; (c) void ratio e

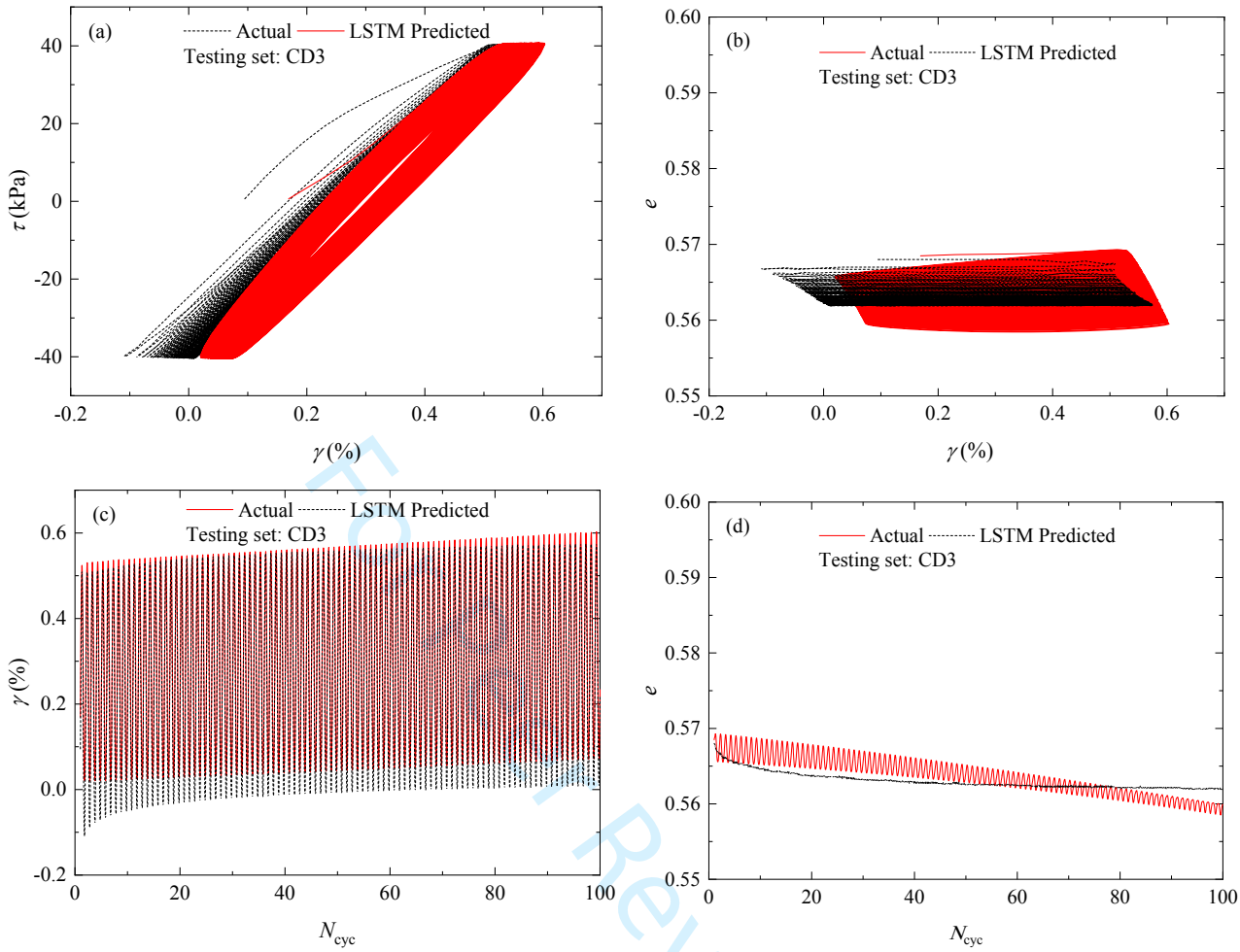

Fig. 13 Cyclic response of Fontainebleau sand under drained condition for the testing set: (a) τ–γ; (b) e–γ; (c) γ–N_{cyc}; (d) e–N_{cyc}

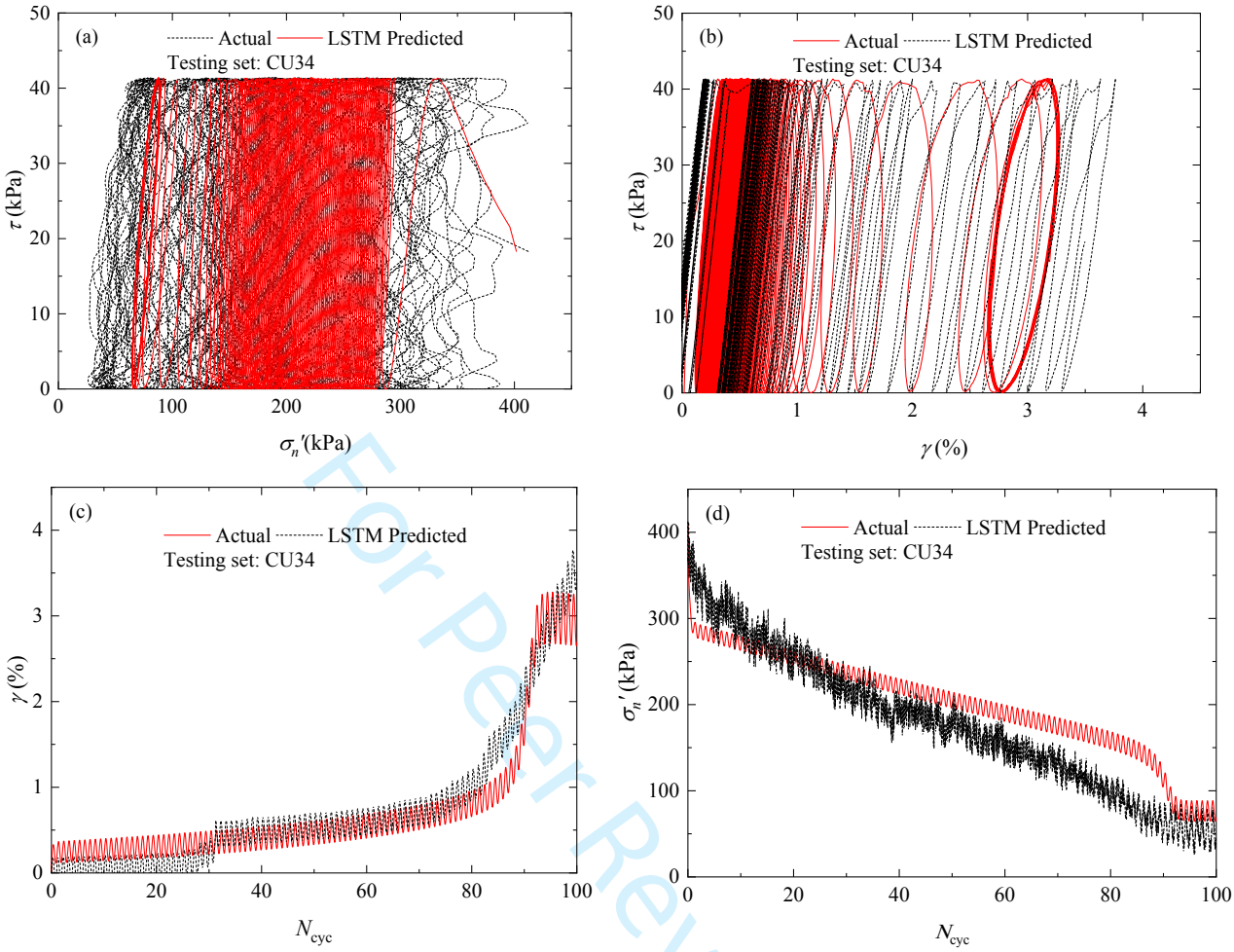

Fig. 14 Cyclic response of Fontainebleau sand under undrained condition for the testing set: (a) τ–$\sigma_n^{'}$; (b) τ–γ; (c) γ–N_{cyc}; (d) $\sigma_n^{'}$–N_{cyc}

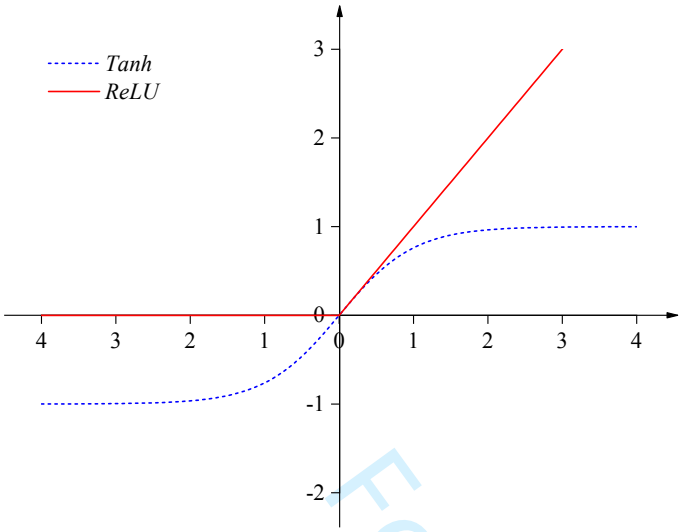

Fig. 15 Curves of two activation functions