

Evaluation and Future Prospects of Data-Driven Intelligence-Based Framework for Predicting Cyclic Behavior of Reconstituted Sand

Kaushik Jas; Amalesh Jana; G. R. Dodagoudar

Accessibility Disclaimer:

For a more accessible version of this document, please submit an accessibility request form through the Montana State University Library website.

RESEARCH ARTICLE

Evaluation and Future Prospects of Data-Driven Intelligence-Based Framework for Predicting Cyclic Behavior of Reconstituted Sand

Kaushik Jas^{1,2}  | Amalesh Jana³  | G. R. Dodagoudar⁴ 

¹Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu, India | ²Hokkaido University, Sapporo, Japan | ³Civil Engineering Department, Montana State University, Bozeman, Montana, USA | ⁴Computational Geomechanics Laboratory, Department of Civil Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu, India

Correspondence: Amalesh Jana (amalesh.jana@montana.edu)

Received: 14 September 2024 | Revised: 26 December 2024 | Accepted: 27 December 2024

Funding: This study was supported by the computational geomechanics laboratory and HPCE laboratory of the Indian Institute of Technology Madras, Chennai, India. Ministry of Education, Govt. of India provided funds in the form of Prime Minister's Research Fellowship.

Keywords: dynamic soil properties | excess pore pressure ratio | gated recurrent unit (GRU) | long short-term memory (LSTM) | shear strain | uniform cyclic loading

ABSTRACT

Most of the robust artificial intelligence (AI)-based constitutive models are developed with synthetic datasets generated from traditional constitutive models. Therefore, they fundamentally rely on the traditional constitutive models rather than laboratory test results. Also, their potential use within geotechnical engineering communities is limited due to the unavailability of datasets along with the model code files. In this study, the data-driven constitutive models are developed using only laboratory test databases and deep learning (DL) techniques. The laboratory database was prepared by conducting cyclic direct simple shear (CDSS) tests on reconstituted sand, that is, PDX sand. The stacked long short-term memory (LSTM) network and its variants are considered for developing the predictive models of the shear strain (γ [%]) and excess pore pressure ratio (r_u) time histories. The suitable input parameters (IPs) are selected based on the physics behind the generation of r_u and γ (%) of the liquefiable sands. The predicted responses of γ (%) and r_u agree well in most cases and are used to predict the dynamic soil properties of the PDX sand. The same modeling framework is extended for other sand and compared with existing AI-based constitutive models to verify its practical applicability. In summary, it is observed that though the trained models predicted the time histories of r_u and γ reasonably well; however, they struggled to predict the hysteresis loops at higher cycles. Therefore, more research is needed to verify and enhance the predictability of existing AI-based models in the future before using them in practice for simulating cyclic response.

1 | Introduction

Catastrophic failures have been attributed to earthquake-induced soil liquefaction which in turn poses a significant societal challenge. Several methods have been developed to evaluate the liquefaction triggering potential and its consequences over the past several decades. The most widely used approaches are

the simplified stress-based methods and their useful alternatives, that is, the strain-based methods [1–7]. Robust determination of constitutive model parameters is essential for predicting the soil response under dynamic loading. Laboratory studies have provided vital insight into the cyclic response of soils, including the critical parameters such as the development of shear strain and excess pore pressure time histories. However, in many cases,

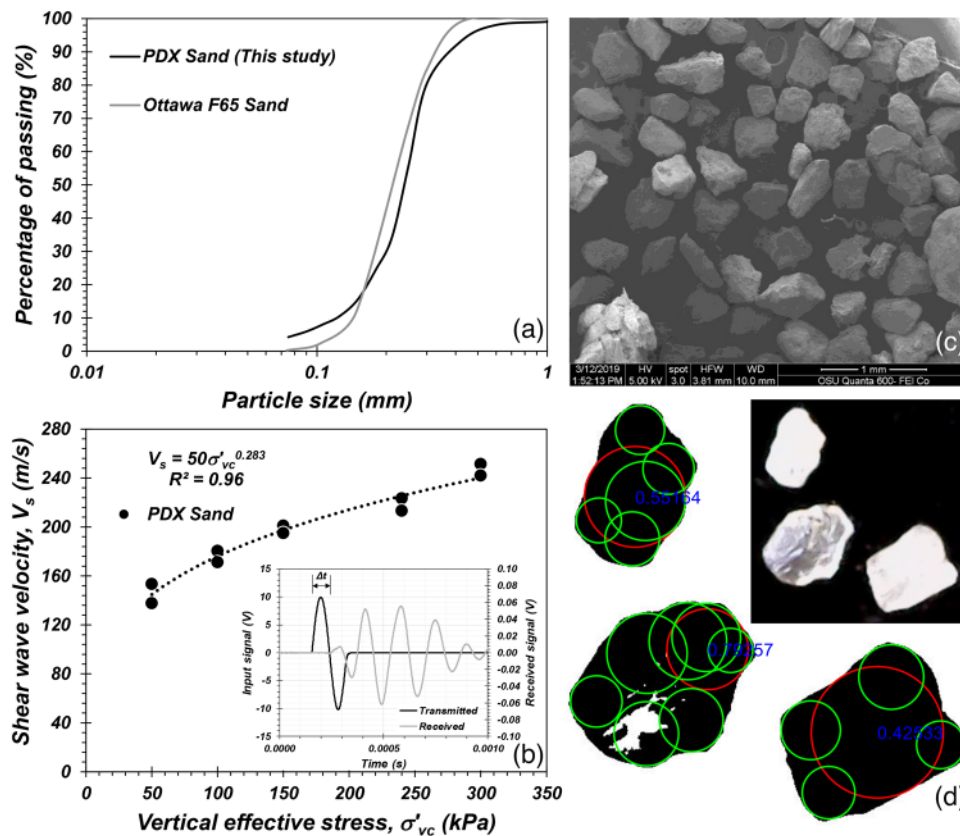


FIGURE 1 | (a) Grain size distribution of PDX sand and Ottawa F65 sand; (b) variation of shear wave velocity with vertical effective stress determined for PDX sand; (c) SEM images of PDX sand; (d) particle shape analysis of PDX sand.

the lack of advanced laboratory devices impedes the researchers and practitioners from obtaining such important dynamic properties. The researchers have developed over the years many advanced constitutive models for soils which capture the cyclic and dynamic response realistically [8–11]. These models are extensively used in practice for nonlinear site response and deformation analyses [12–15]. These models are very powerful tools to predict the cyclic response of sands and require many input model parameters, which are estimated/assumed in many instances in the absence of high-quality advanced cyclic test.

An alternative to traditional constitutive models, researchers have developed data-driven constitutive models for different soils using machine learning (ML) techniques, such as traditional NN [16–20], evolutionary polynomial regression (EPR) [21–23], support vector machine [18, 24], and recurrent neural network (RNN) [25, 26]. Most of these studies are based on simplified ML training strategies and are unsuitable for time history analysis. To alleviate this issue, researchers have tried to model the sequential relationships of stress-strain time history response using long short-term memory (LSTM)-based algorithms [27–31]. Hochreiter and Schmidhuber [32] constructed the LSTM algorithm (a variant of RNN) to resolve the major shortcomings of traditional RNN techniques. Zhang et al. [28] developed one-step ahead stacked LSTM-based models using the synthetic datasets obtained from advanced constitutive models for simulating the cyclic behavior of Toyoura and Fontainebleau sands. Of late, an enhanced LSTM model was developed by Zhang et al. [29] to simulate the stress-strain relationships using the synthetic datasets of various

soil types obtained from numerical modeling experiments. Recently, Choi and Kumar [27] attempted to predict the excess pore pressure (r_u) response of Nevada sands using laboratory test data of the same sand subjected to nonuniform cyclic loading. It is evident from their work that the LSTM-based models can capture the shielding effect [33] in r_u , suggesting the suitability and capability of these artificial intelligence (AI)-based models. Guan and Yang [31] developed a hybrid LSTM model using the synthetic dataset generated by a constitutive model for predicting the monotonic and cyclic responses of the sands. Finally, the hybrid model was tested for both synthetic datasets obtained from the constitutive model and laboratory database of Karlsruhe fine sand. However, these LSTM-based models cannot give satisfactory results for higher number of loading cycles. Thus, Meng and Pei [30] developed an unrolled sequence-to-sequence (seq2seq) model for solving the long-term dependence associated with longer sequence responses of the Ottawa F65 sands. They used multistep-ahead RNNs instead of existing one-step-ahead LSTM-based models. The gated recurrent unit (GRU) cells are used instead of the traditional LSTM cells to construct the model. The model was trained with limited amount of experimental data, which was improved using the cubic spline interpolation, that is, the combination of experimental and synthetic datasets. Overall, most of these existing models are developed using synthetic datasets generated using the traditional constitutive models, which inherently depend on the traditional constitutive models rather than actual laboratory test results. In addition, their potential use within geotechnical engineering communities is inadequate because of the unavailability of training datasets with

TABLE 1 | Summary of grain size characteristics of sands.

Boring/sample	Fines content,		USCS classification
	C_u^a	C_c^b	
PDX sand	2.18	1.28	4
Ottawa F65 sand	1.73	0.95	0

^aCoefficient of uniformity.

^bCoefficient of curvature.

details, model or code files, and architectures of the models. In recent times, the availability of powerful computational systems and the accumulation of large databases provide the opportunity to develop robust AI-based models. There are additional advantages of using an AI-based model such as (i) self-learning from the datasets only, (ii) ease in updating the trained models using the new datasets that improve further predictability, and (iii) generalization ability and extrapolation capability [34–40].

In this study, the data-driven constitutive models for poorly graded reconstituted sand are developed using the laboratory tests database only. The primary objectives of the study are (i) to conduct laboratory tests on a poorly graded reconstituted sand under uniform cyclic loading for establishing training and testing datasets for data-driven constitutive models, (ii) to develop corresponding data-driven constitutive models for predicting the shear strain and excess pore pressure ratio time histories using deep learning (DL) techniques and laboratory datasets, (iii) to compare and evaluate the performance of the developed data-driven constitutive models with the responses from the laboratory tests, (iv) to compare the same framework for data-driven con-

stitutive models with the existing models and other sands, and (v) to highlight the potential (strengths and limitations), and prospects of the framework for data-driven constitutive models. The subsequent sections provide more details of the laboratory tests performed, the mathematical background of the selected DL techniques, the construction of data-driven constitutive models, and the performance assessment of the developed constitutive models.

2 | Laboratory Investigations of the PDX Sand

2.1 | Material Characterization

The cyclic behavior of sand is influenced by several features, such as relative density, soil fabric, stress anisotropy, overconsolidation ratio (OCR), and other factors [41–44]. In this study, the laboratory investigations were conducted on sand samples obtained by split-spoon samplers. The index tests, advanced imaging, and cyclic strength testing were performed. The reconstituted samples were subjected to stress-controlled and constant-volume cyclic direct simple shear (CDSS) testing. The split-spoon samples were obtained from the medium-dense sand layer at a test location in Portland, Oregon, and were referred to as PDX sand in this study for simplicity. The particle morphology (Figure 1a,c,d) and small strain shear wave velocity, V_s (Figure 1b) of the sand were determined. Figure 1a presents the grain size distributions of the sand, indicating a median grain diameter, D_{50} , as 0.26 mm and Table 1 summarizes the grain size characteristics. The sand is generally classified as poorly graded fine sand (SP) with fine content, FC , 4%. The minimum and maximum void ratios are $e_{min} = 0.65$

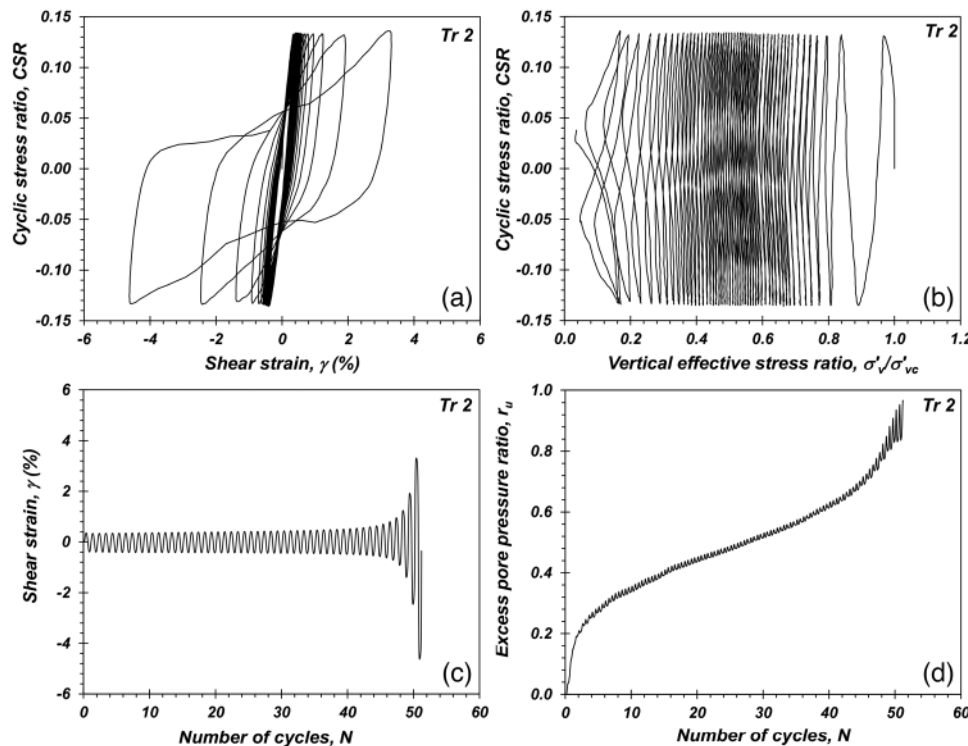


FIGURE 2 | Example of stress-controlled cyclic DSS test results of PDX sand (a) shear stress-shear strain hysteresis; (b) effective stress path; (c) accumulation of shear strain with the number of loading cycles; (d) generation of excess pore pressure ratio with number of loading cycles. [Tr = Training dataset]

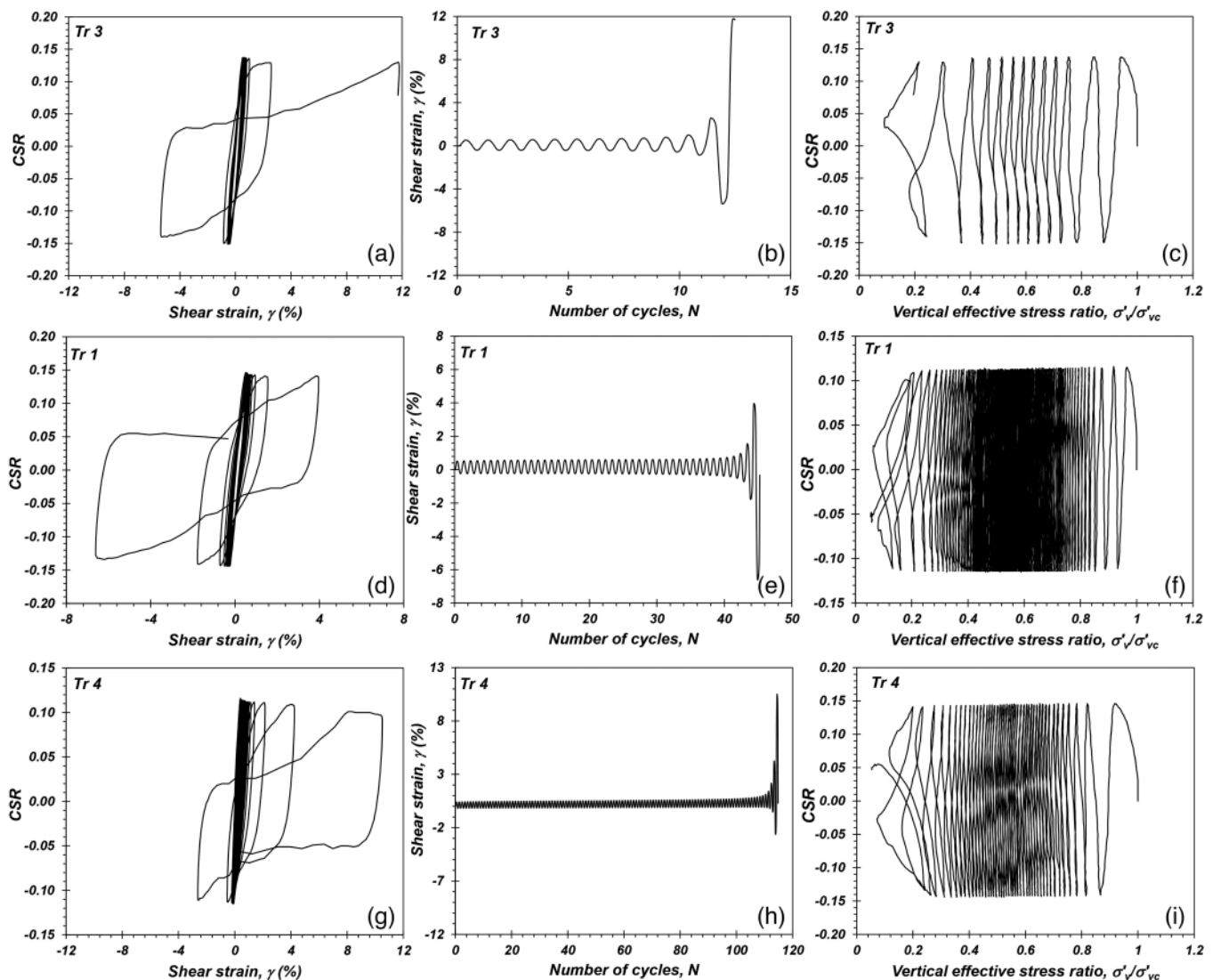


FIGURE 3 | Various examples of stress strain hysteresis (a, d, and g), shear strain accumulation with number of loading cycles (b, e, and h), and effective stress paths (c, f, and i) of PDX sand at different CSR for a vertical effective stress of 100 kPa.

and $e_{max} = 0.95$, respectively. Figure 1c depicts the image of sand grains obtained using scanning electron microscopy (SEM). High-resolution microscopy images of the sand particles were captured, ensuring no physical interaction occurred between any two particles. These images were taken to minimize the errors in predicting the shape parameters. An appropriate background color is selected that would allow for decent visualization of the sand particles. The photos were subsequently transformed into binary format after eliminating the noise using the Image processing toolbox of MATLAB. They were then examined using the MATLAB code developed by Zheng and Hryciw [45] to determine the roundness, two-dimensional (2D) sphericity, and regularity. Figure 1d shows the schematic depiction of the procedure displaying raw images before and after the image analysis. Thirty images were taken, and then particle shapes were analyzed for the sand.

Zheng and Hryciw [45] established methods using computational geometry to facilitate the estimation of soil particle sphericity and roundness. The area and perimeter of the grains can be readily

calculated using the image-processing software from MathWorks in MATLAB. The outline of a particle is discretized to calculate the lowest circumscribing circle, the maximum inscribed circle, and the particle's length and width (Figure S1). The smallest number of outer points that will enclose all others is determined, and the points most distant from one another delineate the circumference of a trial circle. If a point is located outside this circle, the new trial circle is redrawn by incorporating the third point, resulting in the minimum circumscribing circle. The largest inscribed circle is determined using a Euclidean map, with additional information provided in Zheng and Hryciw [45]. The length and breadth are determined by devising imaginary rectangular boxes that enclose the outermost points of the particle, with the box having the largest single dimension defining both the length and width. Moreover, the computation of roundness necessitates the fitting of circles into the different grain corners, requiring high-quality photos to prevent pixelation effects that could lead to inconsistent and erroneous roundness estimates. This holds for other shape descriptions as well. In this study, high-quality images were taken to reduce this uncertainty.

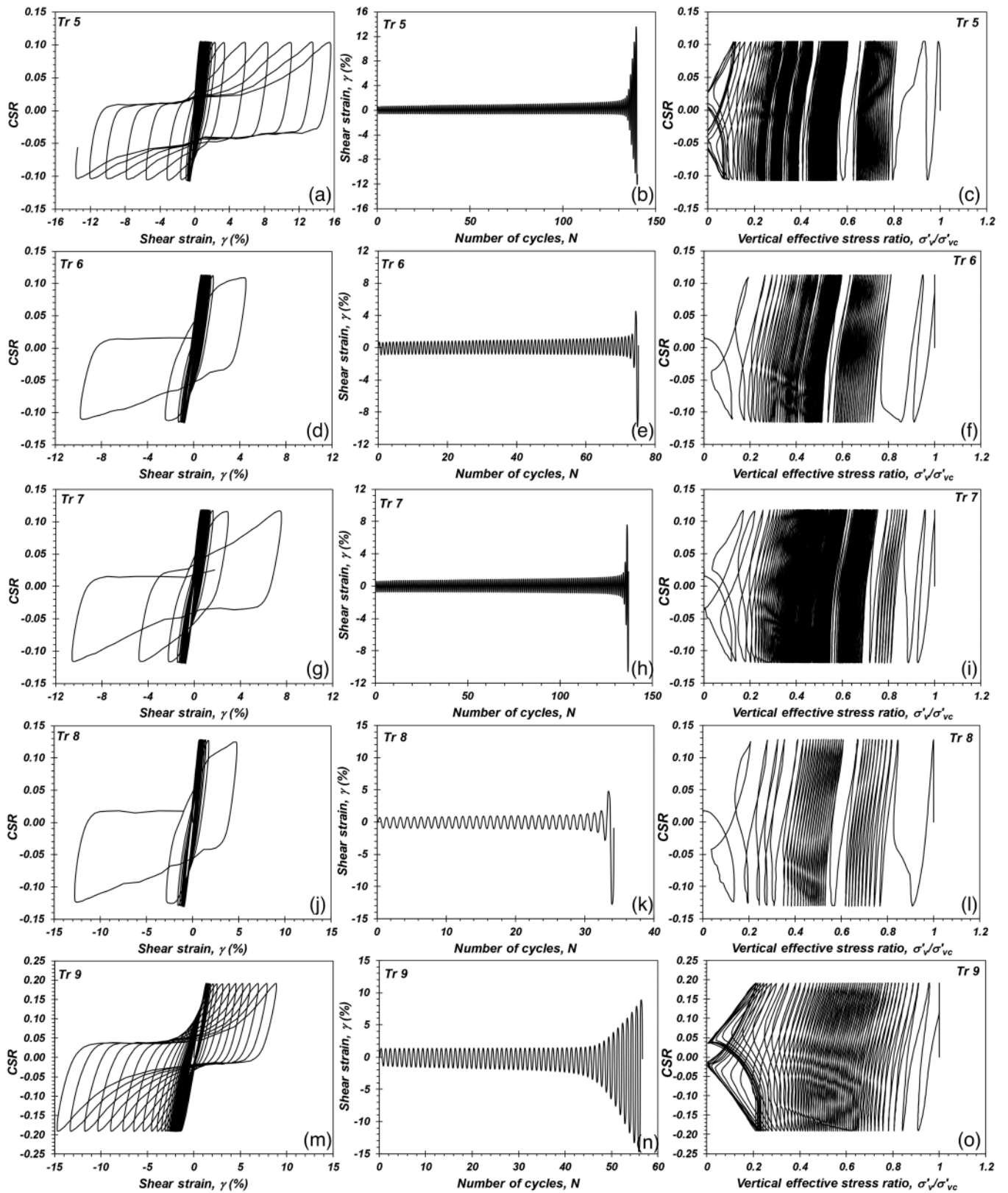


FIGURE 4 | Various example of stress strain hysteresis (a, d, g, j, and m), shear strain accumulation with number of loading cycles (b, e, h, k, and n), and effective stress paths (c, f, i, l, and o) of PDX sand at different CSR for a vertical effective stress of 240 kPa.

Roundness (R) and sphericity (S) are the two most widely recognized characteristics used to describe the particle form. The roundness is defined as the ratio of the average radius of curvature of the corners of a particle to the radius of the largest circle that can fit inside the particle. The sphericity is defined as the ratio of the surface area of a sphere with the same volume as that of the particle to the surface area of the particle. Based on the analysis, the average R and S for the sands were 0.65 and 0.7, respectively, indicating that the sand particles are rounded to subrounded in characteristics [46].

2.2 | Small Strain Shear Wave Velocity

Dry sand was air-pluviated in the membrane-lined direct simple shear (DSS) rings and compacted to the average medium relative density, D_r , of 48% and 53% for two different vertical effective stresses. The typical diameter and height of the sand specimens were 20 and 72 mm, respectively. Later on, the small strain shear wave velocity, V_s , was measured using bender elements. Following the velocity measurement, sinusoidal shear wave pulses were transmitted from the bottom of the specimen to the top, where it was received [47, 48]. Figure 1b (inset) presents typical transmitted and received signals and the corresponding delay in travel time. A variation of the measured V_s of the sand with vertical effective stress is shown in Figure 1b. The average V_s of the sand at 100 kPa ($\sim D_r = 46\%$) and 240 kPa ($\sim D_r = 53\%$) vertical effective stresses were 184 and 235 m/s, respectively, and the associated cyclic resistance ratio was estimated to be 0.21 and 0.23, respectively for the corresponding V_s .

2.3 | CDSS Tests

Constant-volume CDSS tests were carried out on reconstituted sand specimens. The GCTS testing system [49] makes use of servo-controlling shear and normal load actuators during the execution of the CDSS tests [50–54]. These actuators are controlled by a microcontroller, which is embedded into the product as well as CATS software [49]. The height of the usual cylindrical DSS specimen was 20 mm, and its diameter was 72 mm. A latex membrane was utilized for enclosing the specimen, and seven brass metal rings with a thickness of 3 mm were utilized to confine the specimens against lateral deformations during the consolidation process. The diameter of the inner rings was 72.9 mm, which was slightly larger than the diameter of the specimen to provide room for the latex membrane encapsulation. Under conditions of zero lateral strain, sand samples were consolidated to achieve a particular vertical effective stress (σ'_{vc}). Figure 2 presents the CDSS loading response of the sand specimens consolidated to σ'_{vc} of 100 kPa. The relative density of the specimen was $\sim 48\%$. This specimen was subjected to uniform sinusoidal loading with a frequency of 0.1 Hz and a cyclic stress ratio ($CSR = \tau_{cyc} / \sigma'_{vc}$) of 0.134, where stress τ_{cyc} is cyclic shear stress. During the first loading cycle, the specimen underwent a peak shear strain of 0.39% and gained an equal residual excess pore pressure ratio, denoted by $r_{u,r}$, of 12%. After 50 cycles, the specimen reached a peak shear strain of 3%, which corresponded to a $r_{u,r}$ of 95%. Over the course of 52 cycles, the specimen reached a peak shear strain of approximately 4.55%, corresponding to a r_u of 97%. Figure 3 presents the additional examples of shear stress-shear strain

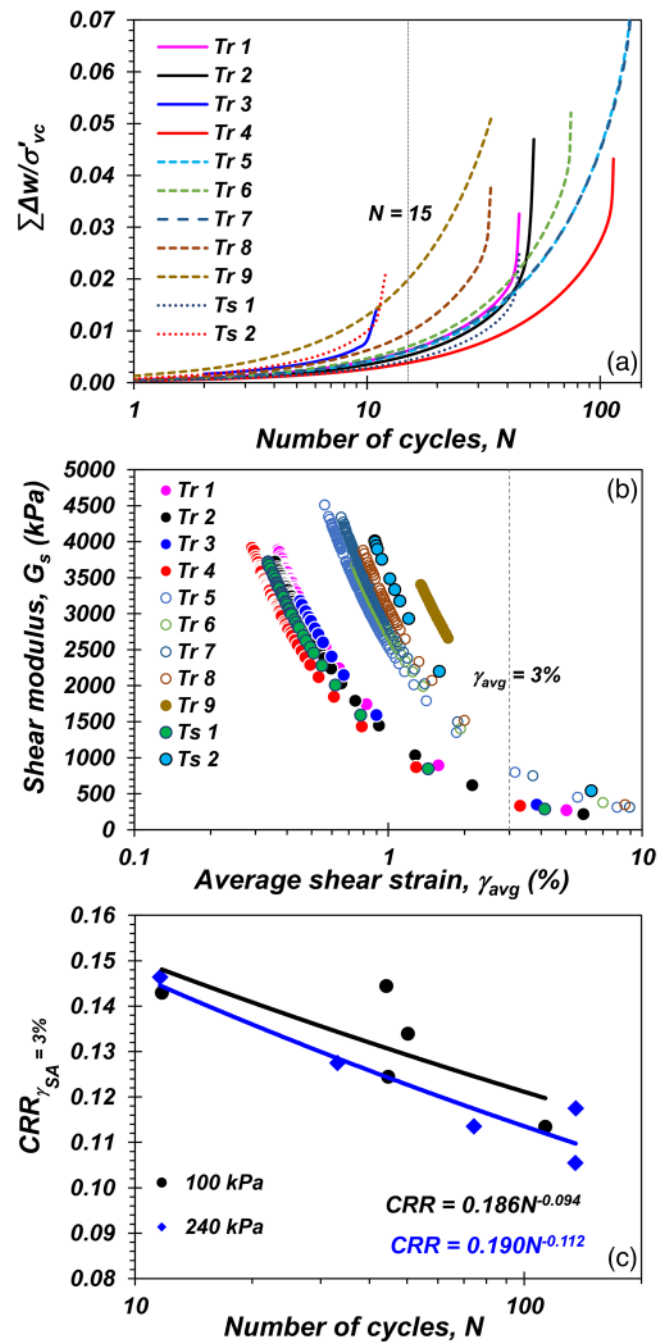


FIGURE 5 | Variation of (a) cumulative dissipated energy with number of loading cycles; (b) normalized shear modulus with the average shear strain for eleven different specimens; (c) comparison of cyclic resistance curves for specimens consolidated to vertical effective stresses of 100 kPa and 240 kPa. [Tr = Training dataset; Ts = Testing dataset].

hysteretic responses, accumulation of shear strain and effective stress paths of various specimens performed using different CSRs at the same or similar relative density ($D_r \sim 46\%$) and consolidated to $\sigma'_{vc} = 100$ kPa. Specimens that were subjected to higher CSR reached 3% single-amplitude shear strain in fewer number of cycles than those subjected to lower CSR, as anticipated from the previous cyclic testing on the reconstituted and natural sand specimens [55, 56]. Figure 4 presents the additional test results obtained at σ'_{vc} of 240 kPa and $D_r \sim 53\%$. Figure 4m–o present the cyclic response of dense specimens ($D_r \sim 80\%$), indicating the

TABLE 2 | Summary of the cyclic response of the PDX sand determined from stress controlled cyclic DSS tests.

Sample	CSR	σ'_{vc} (kPa)	D_r	$N_{\gamma=3\%}$	N_{max}	γ_{max} (%)	$r_{u,max}$	$r_{u,N=15}$
Tr 1	0.144	100	51	44.2	45.2	6.53	0.95	0.40
Tr 2	0.134	100	47	50.23	52.0	4.55	0.96	0.40
Tr 3	0.143	100	48	11.72	12.5	11.55	0.90	NA
Tr 4	0.113	100	38	113.2	114.8	10.46	0.95	0.27
Tr 5	0.106	240	52	135.2	143.2	13.54	1.00	0.31
Tr 6	0.114	240	55	74.2	75.2	9.788	1.00	0.37
Tr 7	0.118	240	52	135.6	137.8	10.37	1.00	0.34
Tr 8	0.128	240	54	33.1	34.2	12.65	1.00	0.48
Tr 9	0.192	240	80	47.0	56.6	14.38	1.0	0.38
Ts 1	0.125	100	53	44.7	45.0	5.45	0.94	0.37
Ts 2	0.146	240	48	11.6	12.7	14.92	1.00	NA

Abbreviations: NA, not applicable; Tr, training dataset; Ts, testing dataset.

cyclic mobility type response. The accumulation of shear strain is gradual, and the specimen shows a dilatant response. Figure 5a shows the variation of normalized cumulative dissipated energy ($\sum \Delta w / \sigma'_{vc}$) with the number of loading cycles for the sand, where Δw is the area of the hysteresis loop. At first, the hysteresis loops have similar sizes, and the correlation between the normalized dissipated energy ($\Delta w / \sigma'_{vc}$) and the number of cycles is linear (Figure S2). As the loading cycles progress, the hysteresis loops increase in size due to a rapid increase in the r_u . The relationship between the $\Delta w / \sigma'_{vc}$ and N approaches an exponentially growing pattern (Figure S2). Note that the threshold value of $\Delta w / \sigma'_{vc}$ ranges from 0.0008 to 0.001 and a rapid change in $\Delta w / \sigma'_{vc}$ triggers the liquefaction. This number is slightly lower than that reported for Monterey sand [57]. The variation of secant shear modulus with average shear strain is shown in Figure 5b, indicating a higher nonlinearity at larger shear strain. Figure 5c depicts the variation of CSR with the number of cycles required to reach 3% peak shear strain for specimens consolidated to $\sigma'_{vc} = 100$ kPa and 240 kPa. A liquefaction triggering curve as a power law fits the stress-controlled test data, including two testing data (Table 2). The power law model is given by Idriss and Boulanger [55]:

$$CRR = aN^{-b} \quad (1)$$

where N is the number of uniform loading cycles to reach 3% shear strain, and a and b are the fitted coefficient and exponent, respectively. The cyclic resistance of the sand showed a reduction with an increase in $\sigma'_{vc} = 100$ –240 kPa. The CRR values at $N = 15$ cycles for the above two σ'_{vc} values correspond to an M_w of 7.5 earthquake are respectively 0.144 and 0.140, which are lower than that derived from the V_s values (0.21 and 0.23).

Liquefaction could trigger loose to medium-dense coarse-grained soil depending on the soil resistance and earthquake-induced shear stress. In general, if the earthquake-induced shear stress exceeds the soil capacity, then liquefaction could trigger, and the surface manifestation in terms of sand boils, deformation, or cracks can be seen on the ground. The basic framework of the

Simplified Method was initially proposed by Seed and Idriss [58] and then modified by several researchers to evaluate liquefaction triggering (Figure S3a). In this study, CDSS tests were conducted on three different sands, such as PDX sand (with relative density ranging from 38% to 80% with an average value of 53%), Ottawa-65 sand (with relative density ranging from 66% to 76% with an average value of 71%), and Nevada sand (with relative density ranging from 39 to 89% with an average value of 56%) [59, 60] with a broad range of relative density, have been used to train the DL models and the range is marked on the liquefaction resistance plot with relative density (Figure S3b). Nonetheless, future efforts are necessary to train the DL model on CDSS test results on very loose to loose sands.

3 | Development of DL Models for Predicting Cyclic Response of PDX Sand

3.1 | Mathematical Background of LSTM and Its Variants

RNN is a suitable DL method over the conventional NN for modeling sequential data. This is possible because of the cyclic connection topology adopted in the RNN. Thus, it is appropriate for modeling stress-strain relationships of various soils under cyclic loading. On the other hand, gradient boosting methods (GBMs) and conventional NNs are suitable for classification and regression analysis. Therefore, the GBMs and conventional NN are appropriate for classifying the sites of different soil types as liquefiable or non-liquefiable with probabilistic values [61]. However, the traditional RNN has two major limitations: (i) exploding gradients problem for the long sequential dataset and (ii) the learning efficiency of the hidden layers decreases from end to front. Thus, the RNN is mainly suitable for keeping short-term history. In the literature, the LSTM, a special type of RNN, was developed by Hochreiter and Schmidhuber [32] to overcome the limitations associated with the traditional RNN. This is made possible by the inclusion of a “memory cell” in the LSTM NN (Figure 6a). This cell learns the long-term dependence through its three types of gates: forget, input, and output gates. The outputs

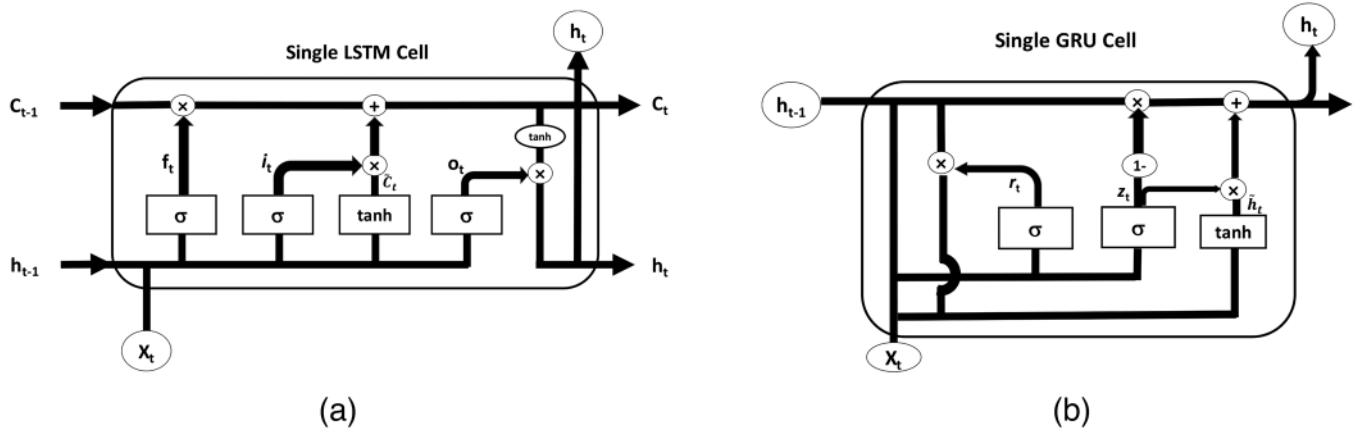


FIGURE 6 | Basic framework of LSTM and GRU cells.

associated with these three gates are determined as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2a)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2b)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2c)$$

where x_t is the current input; h_{t-1} is the hidden state value in the previous iteration; subscripts t and $(t-1)$ represent the present and prior iterations, respectively; b and W are the biases and weights, respectively; the subscripts f , i , and o represent the forget, input, and output gates; σ is the sigmoid function ($= 1/[1 + e^{-x}]$). The sigmoid function is used to decide which information should be stored ($\sigma = 1$) and which should be excluded entirely ($\sigma = 0$). The current or t^{th} hidden state value is updated using the updated cell state (C_t) value and output from the output gate (o_t) as follows:

$$h_t = o_t \odot \tanh(C_t) \quad (2d)$$

The cell state is the principal element in the LSTM cell. The current cell state (C_t) value is determined by

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2e)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (2f)$$

where $\tilde{C}_t$ is the new candidate value; $\tanh$ is the activation function; $\odot$ is the Hadamard/element-wise product.

Cho et al. [62] developed a simplified variant of the LSTM, which is named GRU. The complexity associated with the LSTM architecture is solved by merging input and forget gates (Figure 6b) thereby resulting in only two gates in the GRU cell: update gate (z_t) and reset gate (r_t). Thus, the total number of parameters is reduced, which makes the GRU faster and computationally inexpensive. Due to fewer parameters, the GRU is less prone to overfitting, however it may face difficulties in long-sequence predictions. The mathematical operations behind these two gates in the GRU are described by the following sets of

equations:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (3a)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (3b)$$

The current hidden state (h_t) is updated in the update gate using the prior hidden state (h_{t-1}) and the candidate's hidden state ($\tilde{h}_t$) is obtained as follows:

$$\tilde{h}_t = \tanh(W \cdot [r_t \odot h_{t-1}, x_t]) \quad (3c)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t * \tilde{h}_t \quad (3d)$$

3.2 | Data Pre-Processing, and Training, Tuning and Testing of LSTM-Based Models

The current study aims to predict the cyclic response (e.g., γ (%) and r_u time histories) of PDX sand using DL algorithms. The cyclic response, including the development of pore pressure generation, accumulation of shear strain, shear modulus degradation, cumulative energy dissipation, and cyclic resistance was determined using a series of stress-controlled tests at two different vertical effective stresses. The databases obtained from these CDSS laboratory tests (Table 2) have been considered for training and testing of the DL models. Baseline correction with zero padding is performed for these datasets (Table 2) before training and testing the LSTM algorithm. Baseline correction ensures that the data are centered around a standard reference by removing offsets or biases in the datasets. The zero padding strategy helps to adjust the sequence length by adding zeros at the start of the datasets. This baseline correction with zero padding combination is effective for tasks like time-series forecasting. The combination ensures the learning of the LSTM model from corrected and standardized input data. In addition, truncation is performed at the end, that is, after the failure of the soil sample, of each dataset based on the laboratory response of shear strain time history. Truncation reduces unnecessary computational costs and focuses only on the meaningful portion of the dataset. All these data pre-processing steps can improve the accuracy and efficiency of the LSTM model training and testing, which leads to more

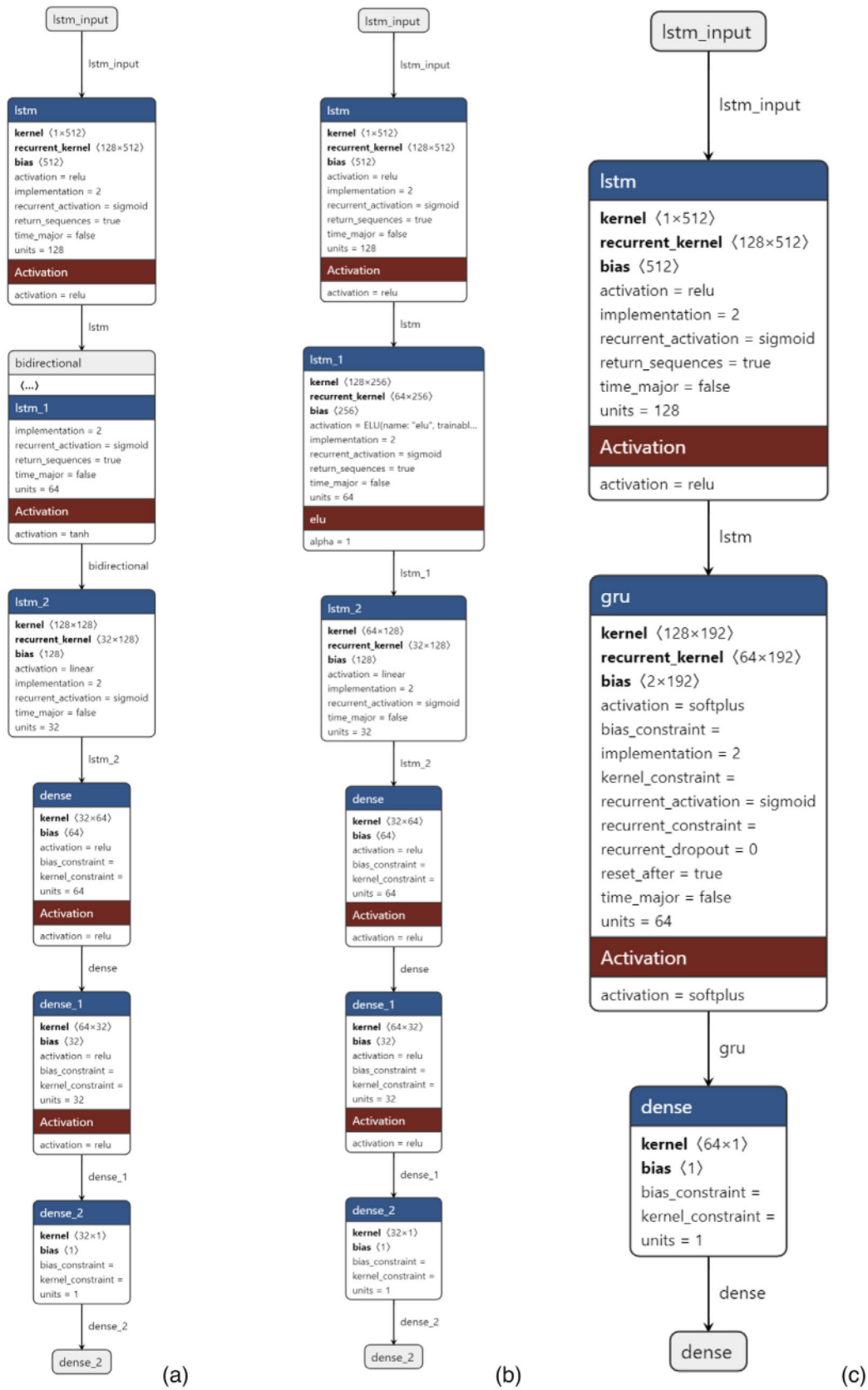


FIGURE 7 | Adopted architectures of LSTM-based models for PDX sand: (a) shear strain model for Trial 1; (b) shear strain model for Trial 2; (c) r_u model for Trials 1 and 2.

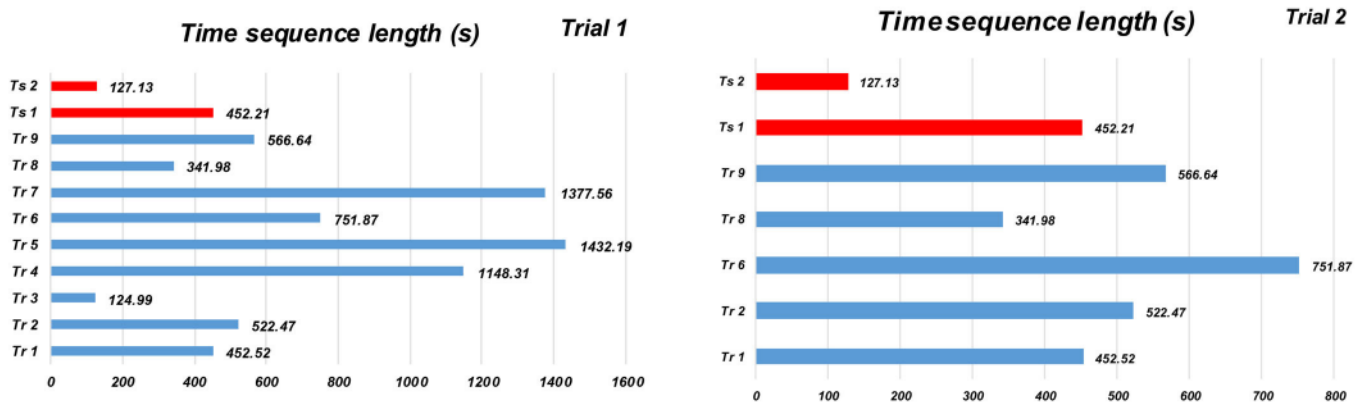


FIGURE 8 | Training and testing dataset analyses for each trials of the PDX sand models. [Tr = Training dataset; Ts = Testing dataset].

reliable predictions. During the training and testing phases, the nine training datasets and the two testing datasets (Table 2) are all available in separate Excel files. In the present study, slightly different training, validation, and testing approaches are incorporated to construct LSTM-based models. Each of the nine datasets is fed to the models one at a time during the training process. Thus, during LSTM model training in the Tensorflow DL framework, the *time_step* is assigned as 1, which ensures the *sequence length* is effectively 1 [28]. This training strategy of the LSTM model helps to understand the target outputs of each training dataset based on the current inputs of each dataset, that is, one-to-one mapping. Finally, the trained models are tested on two separate/unseen/independent testing datasets. This methodology benefits the LSTM models by allowing them to recognize the responses from CDSS tests one at a time, and then update the weights and biases of the models accordingly. Assume all the nine datasets are merged in a single Excel file and divided into training, validation, and testing sets at 80:10:10. In that case, the models can learn better for only some selective portions of the combined dataset during the training and validation processes. Therefore, the separate loading/feeding approach supports the models to understand the whole sequence of responses in a better manner. The input parameters (IPs) needed for predicting the time history of shear strain (γ [%]) are: shear stress time history, initial vertical effective stress, coefficient of uniformity (C_u), relative density (D_r), and effective particle size (D_{10}). The predicted γ (%) time history is considered as one of the IPs in addition to other IPs for predicting the r_u time history, that is, both the models are connected type. The DL algorithms, such as the LSTM and its variants (e.g., bidirectional LSTM [BiLSTM], and GRU), are considered for predicting the γ (%) and r_u time histories.

Before training the LSTM-based models, researchers generally follow crucial pre-processing steps such as data normalization and standardization. Conversely, this step can exhibit very small or large predicted output compared to experimental output if outliers are present, that is rapid fluctuation in experimental results. As the LSTM-based model utilized gradient-based optimization techniques, it can be sensitive to normalizing and standardizing the input and output parameters. MaxAbs scaling is a data pre-processing technique used to scale feature values in a dataset. It scales the data ($x_{scaled} = \frac{x}{\max(|x|)}$, where x is the actual value of the feature) to the range $(-1, 1)$ by dividing each

feature by the maximum absolute value of that feature. This scaling method has these advantages for the LSTM modeling task: (i) preserves the sparsity of selective features that ensures the original structure of features remains intact, (ii) ensures that the feature magnitudes are normalized without altering their distribution, (iii) maintains the sign of the actual data, which is essential for selective features where the sign carries significant information, and (iv) computationally efficient and aligns with the sensitivity of LSTM to feature magnitudes. Therefore, MaxAbs scaling is selected over other scaling methods because of these advantages and descriptive statistics of the features in the present study. In the present study, MaxAbs scaling is performed only for a few selective features (relative density, r_u , etc.) based on the descriptive statistics (minimum, mean, mode, standard deviation, maximum, etc.) of input and output features.

Finding a suitable group of hyperparameters is vital for constructing robust LSTM-based models. In this study, the best group of hyperparameters is not evaluated exclusively by means of the existing tuners such as AutoML [63, 64], and so forth, on the validation or testing datasets. These tuners can produce better responses of the models specifically for validation or testing sets instead of the more generalized performance of the models. Thus, the activation functions, number of neurons, and other hyperparameters in the architectures for the LSTM-based networks are selected in the present study based on multiple approaches, such as mainly focused on these four criteria (i) available expert domain knowledge from useful and relevant existing studies, (ii) manual tuning experiments for a wide range of variations in hyperparameters, (iii) optimization algorithms, and (iv) self-judgment. The four ways can enhance the extrapolation capability, generalizability, and reduce the risk of overfitting for the trained models. Figure 7 represents the adapted architecture for the developed LSTM-based models. The Dell Precision 3630 Workstation is used for training and testing of the LSTM-based models. The Keras LSTM package [65] is used in this study to train all the LSTM models in Python platform.

The LSTM-based models are initially (i.e., Trial 1) trained using nine datasets and tested using two separate testing datasets having shorter and longer time sequence lengths. Figure 8 shows the time sequence length (in seconds) for these datasets. From

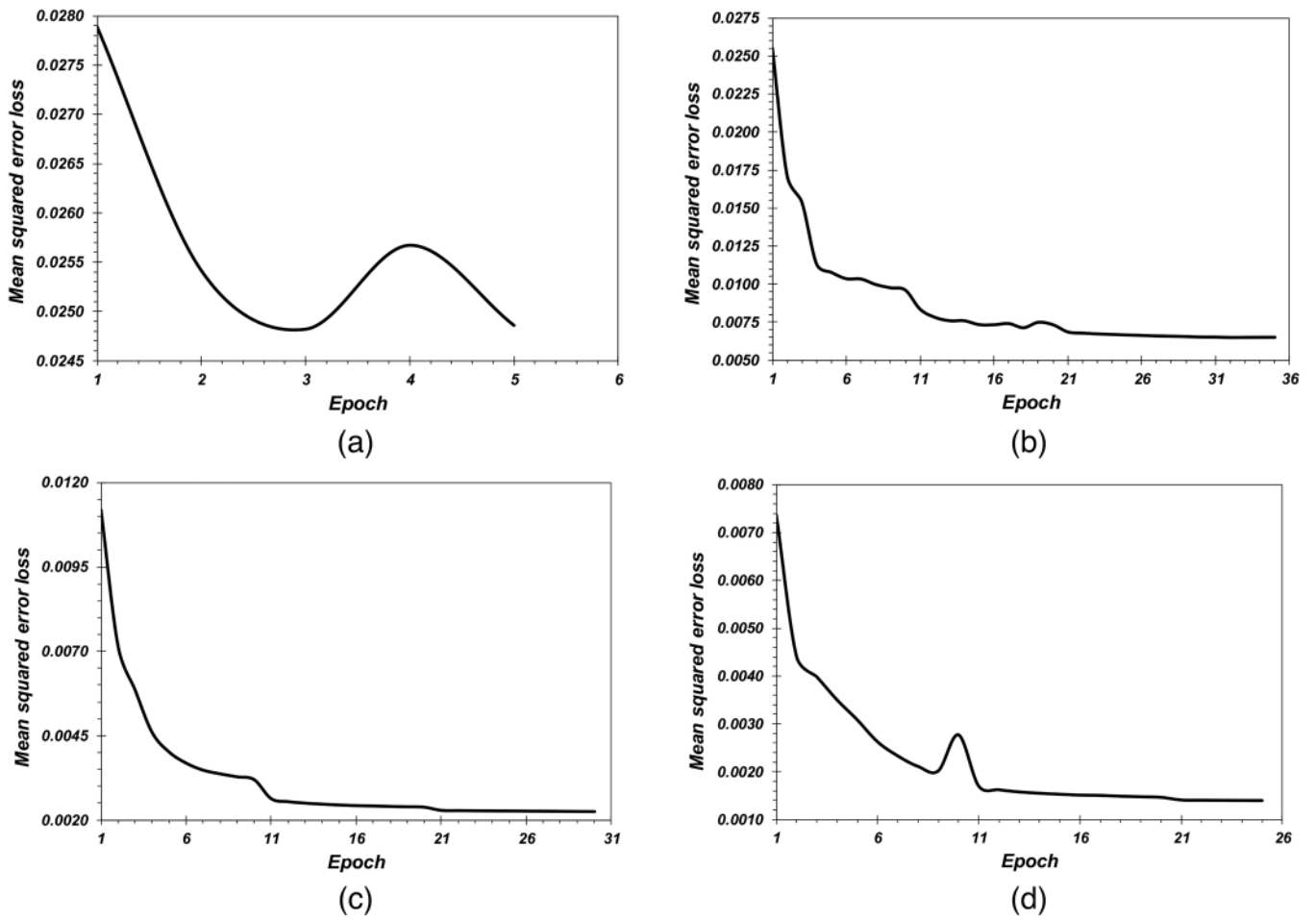


FIGURE 9 | Loss curves of LSTM-based models of PDX sand for all trials: (a) shear strain model for Trial 1; (b) shear strain model for Trial 2; (c) r_u model for Trial 1; (d) r_u model for Trial 2.

these figures, it is clear that a few of these datasets have a wide range of variation in time sequence length compared to the remaining datasets. Therefore, the variations can cause sudden jumps in prediction during the training and testing of the models. During the training and testing, a few datasets have been excluded for each trial to understand the effect of the number of training datasets for each range of time sequence length in the predictions. The mean squared error (MSE) evaluation metric is used during the training and testing of the models. The MSE metric is defined as

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4)$$

where y_i is the actual value and $\hat{y}_i$ is the predicted value for the i^{th} observation, and n is the total number of observations. During the training of the model, the best set of hyperparameters (number of LSTM layers, units, activation functions, etc.) is obtained by minimizing the MSE loss. The MSE losses for each trial during training of the γ (%) and r_u models are shown in Figure 9. An overall decreasing trend in the MSE loss is observed as the number of epochs is increased. However, sudden jumps (increase in MSE loss) are observed for certain epochs due to model parameters, such as learning rate, batch size, weights, biases,

and so forth. After tuning the hyperparameters and excluding a few datasets, the best trial is obtained (i.e., Trial 2). The training performance of the shear strain and r_u models for the best trial (Trial 2) are depicted in Figures 10 and 11. The testing performance of the shear strain models and r_u models corresponding to both trials (Trial 1 and Trial 2) are depicted in Figures 12 and 13. The training performances of both the γ (%) and r_u models for the initial trial (Trial 1) are depicted in Figures S4 and S5, respectively (see the Supporting Material).

Figure 14a,b compare the stress-strain hysteresis of the first ($N = 1$) and last ($N = 45$) cycles between the experimental data and DL-based simulation. Although the LSTM-based model can predict the general trend of the hysteresis loop; however, the accuracy of the model reduces for higher cycles. Nevertheless, the LSTM-based models satisfactorily captured the variation of maximum excess pore pressure ratio, $r_{u, \max}$ and average shear strain, γ_{avg} in each cycle (Figure 14c). The difference in the predicted response at higher cycles is attributed to the rapid change in the stress-strain response close to liquefaction, which the LSTM model is unable to predict. Similarly, a variation in the LSTM-based predicted shear modulus, G with γ_{avg} , follows a similar trend as observed in the experimental response.

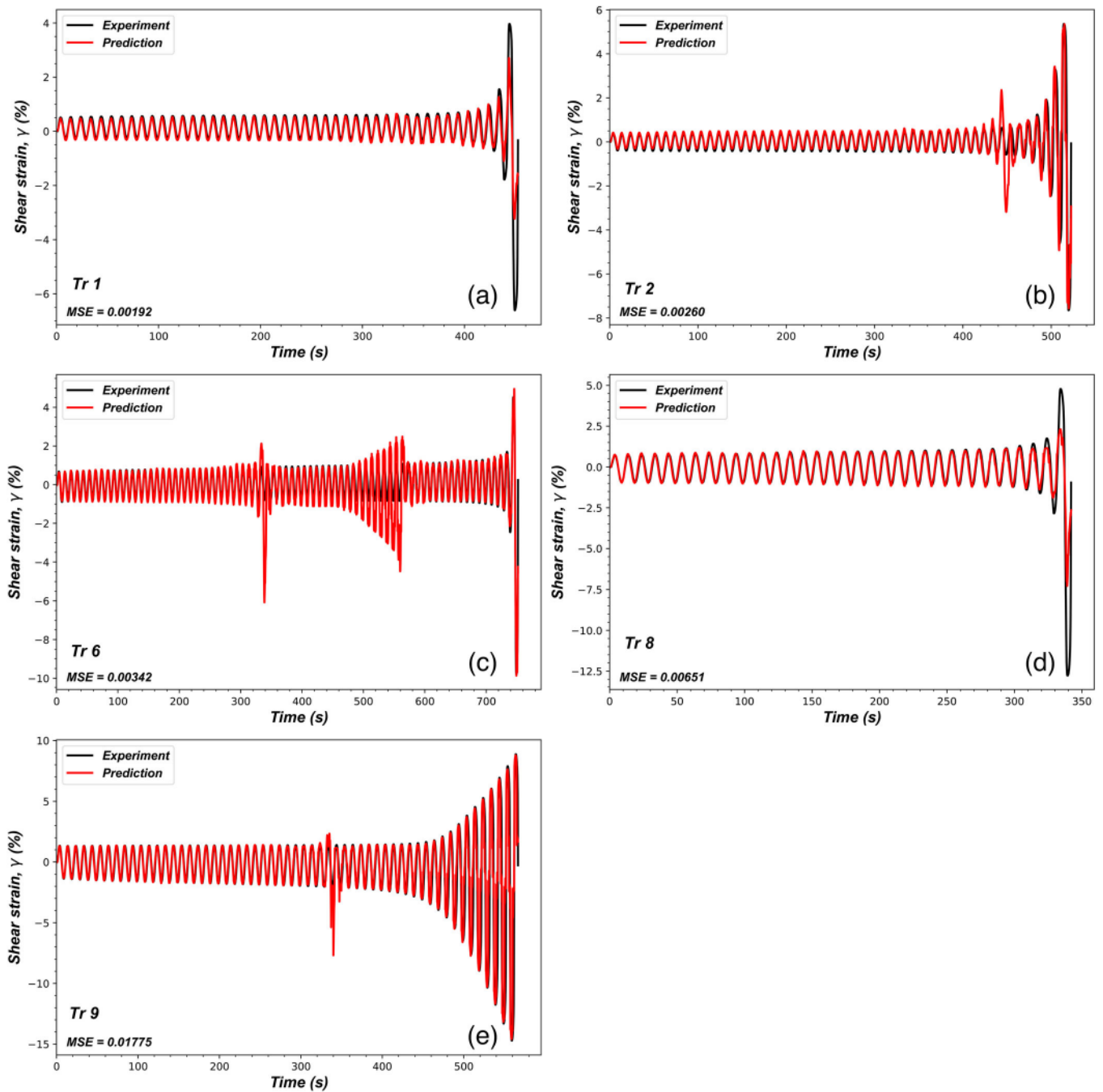


FIGURE 10 | Training performance for shear strain model of PDX sand corresponding to best trial.

3.3 | Interpretation of Predictive Models and the Predicted Cyclic Response

3.3.1 | Analysis of Modeling Strategies

In the present study, the stacked LSTM layers are adopted for modeling the time histories of shear strain (γ [%]) and excess pore pressure ratio (r_u) of the PDX sand. The BiLSTM in the second layer of γ (%) model is unable to give satisfactory results for the training and testing performances. During the training, the BiLSTM layer processes the sequential data both in the forward and backward directions, which is helpful for predicting the longer sequence response. However, the same training process

of BiLSTM violates causality in time series prediction, such as the γ (%) and r_u responses. In cyclic soil behavior, future states cannot physically influence past states because the stress-strain response is path-dependent. Using future information for prediction cannot align with physical principles and could result in a non-causal model [29, 66]. Therefore, in the final γ (%) model architecture (Figure 7b), three stacked unidirectional LSTM layers are considered.

More details of the architecture and hyperparameters of the γ (%) model are also given in the Table S1 and Figure 7. The batch size and initial learning rate values are 64 and 0.001, respectively. More detailed meaning of all the tuned hyperparameters

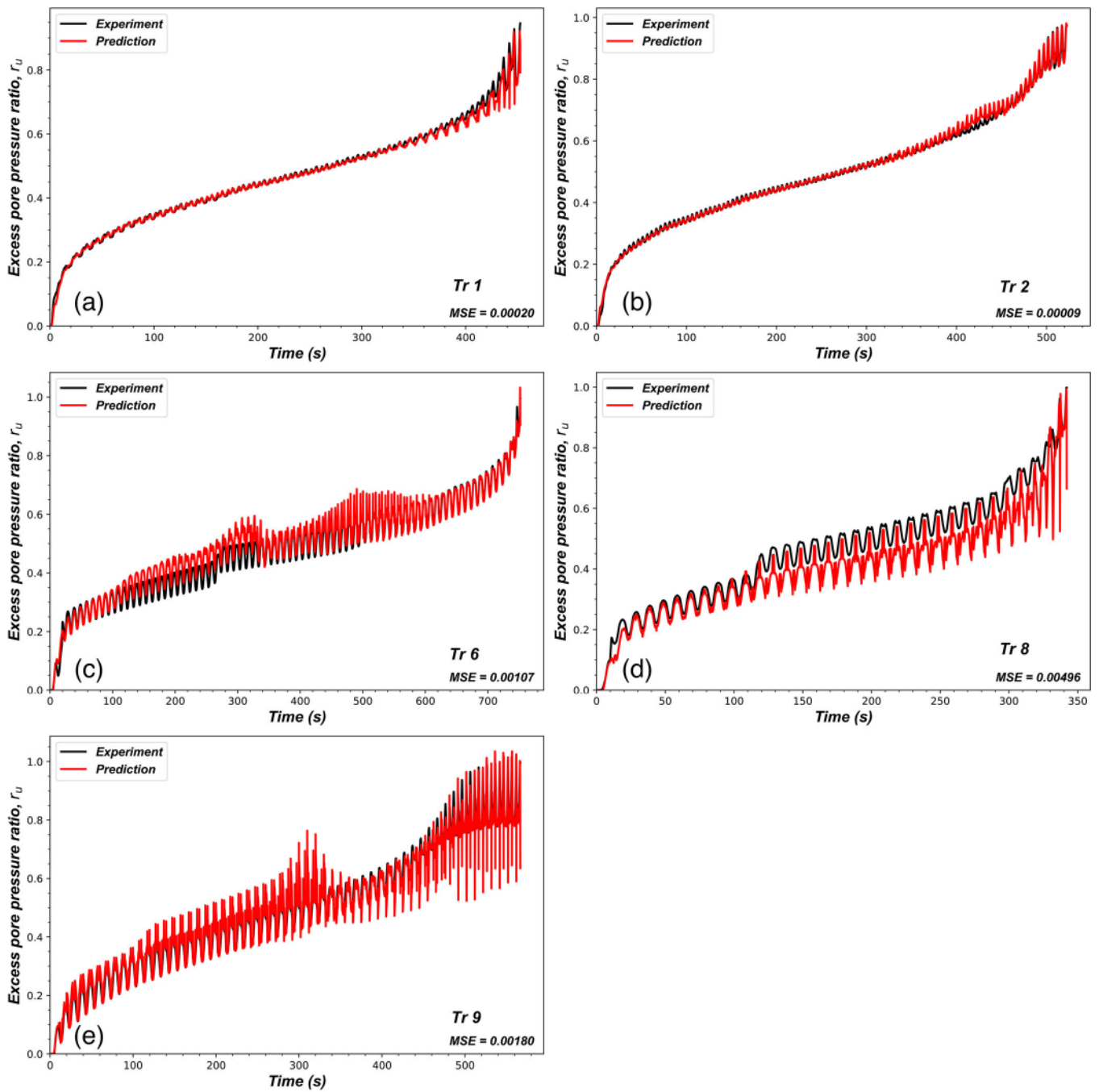


FIGURE 11 | Training performance for r_u model of PDX sand corresponding to best trial.

considered in the study can be found in the Keras LSTM package [65]. The adaptive moment estimation (*Adam*) [67] optimizer is considered for the compilation of the model. This optimizer incorporates both the root mean square propagation (RMSProp) [68] and adaptive gradient algorithms (AdaGrad) [69]. The batch size, initial learning rate and optimizer are the same for all the remaining models developed in this study. The gradient clipping in the first LSTM layer is considered to prevent the exploding gradient problem. It helps update the model parameters effectively and improves convergence during model training. In the study, the gradient clipping is incorporated in the first LSTM layer by applying the *MaxNorm* constraint (with a threshold of 3.0) to both the kernel and recurrent weights. By

imposing a maximum L2 norm of 3.0 on these gradients, the model experiences more consistent and controlled parameter updates, contributing to enhanced numerical stability in the training process of the LSTM network. In addition, a learning rate scheduling strategy has been implemented to adjust the learning rate dynamically while compiling the LSTM models. The initial learning rate is set to 0.001, and a *LearningRateScheduler* is employed to reduce it by a factor of 10 after every 10 epochs. This approach allows for a gradual decay in the learning rate using a step-wise decay function, and facilitates better convergence and fine-tuning as the optimization progresses. The *Adam* optimizer, initialized with the specified learning rate, is utilized for minimizing the MSE during the training phase.

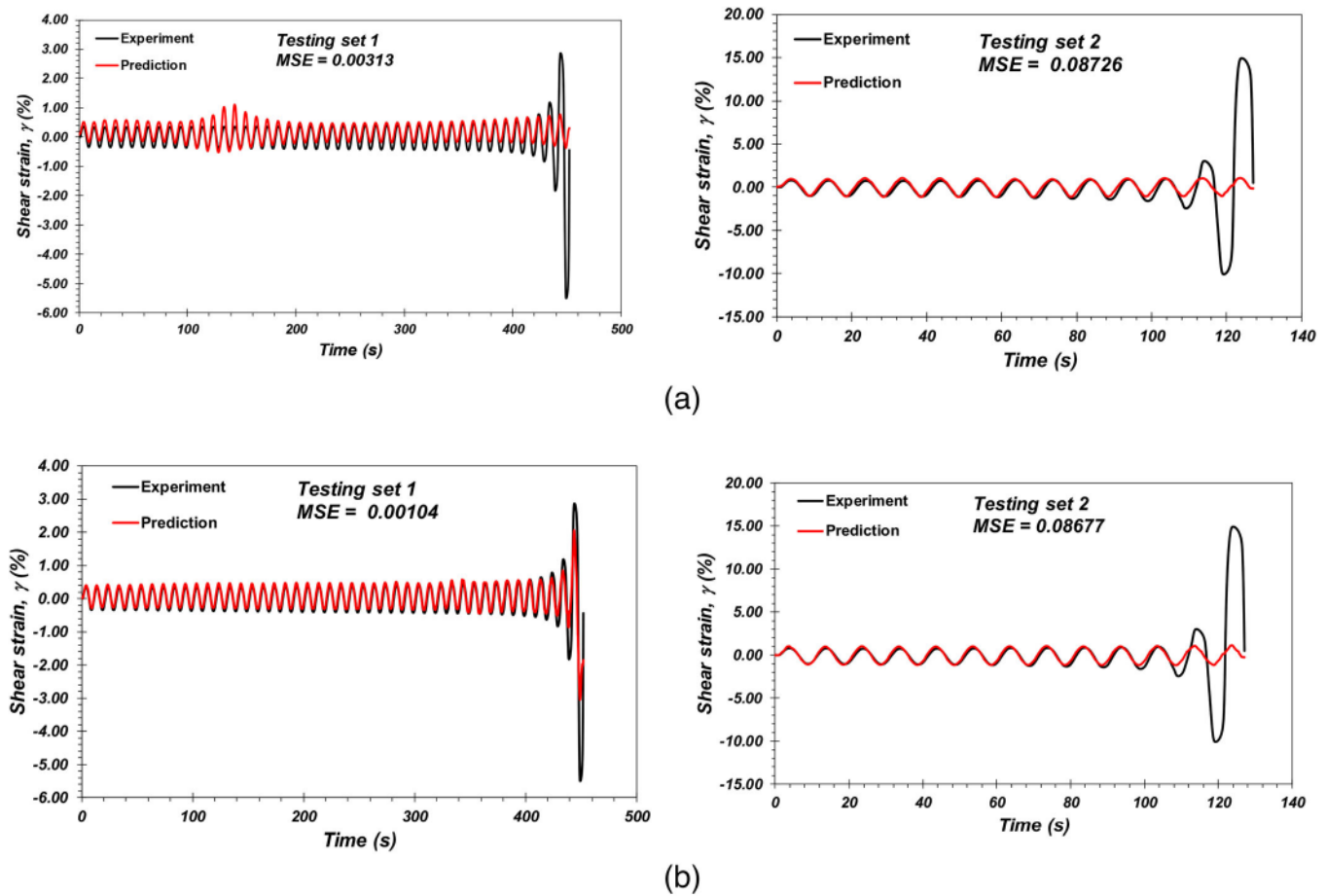


FIGURE 12 | Testing performance of shear strain models for PDX sand corresponding to both the trials: (a) Trial 1; (b) Trial 2.

3.3.2 | Choice of Activation Functions and Cycle-Dependent Analysis

The rectified linear unit (ReLU) activation function is efficient and computationally inexpensive because of its simpler structure. It handles sparse activations, which leads to faster training. On the other hand, the unbounded positive output can lead to gradient instability in cyclic loading, and zero gradients for negative values may cause information loss in stress reversal. The dead neuron problem is particularly critical for cyclic loading where negative values are physically meaningful [29, 70]. Zhang et al. [29, 71] addressed the problems associated with the ReLU activation function to model the stress-strain behavior of soils and highlighted the suitability of the tanh activation function. Shen et al. [70] proposed a novel tanhLU activation function to combine the strengths of tanh and ReLU activation functions. The proposed tanhLU avoids the shortcomings of the tanh and ReLU activation functions. The exponential linear unit (ELU) behaves similarly to ReLU for positive values but smooths out negative ones and prevents dying neurons. This smoothing facilitates better gradient flow, particularly for cyclic data involving negative values [28, 71, 72]. Therefore, for the shear strain model in the present study, the ReLU and ELU activation functions are considered at the first two LSTM layers, respectively, and the ReLU in the two dense layers. These two activation functions help to avoid the vanishing gradient problem in longer sequences. ReLU in the first LSTM layer efficiently captures sparse activations, while

ELU handles negative inputs in the next LSTM layer. In this way, this ReLU-ELU combination helps to prevent the dying neurons problem and improve gradient flow. ReLU efficiently extracts key features, while ELU enhances transitions for smoother and more accurate predictions. Thus, the architecture balances simplicity and performance, which makes it well-suited for modeling the cyclic responses of soil.

Additionally, the performance of the trained shear strain model in the present study (Model A2) is compared with a shear strain model having a single LSTM layer (Model C) with a *tanh* activation function. In a single-layer LSTM, the tanh activation function is used to update the cell state, while sigmoid is used to operate the input, forget, and output gates (Figure 6a). The rest hyperparameters of this unidirectional single LSTM layer shear strain model (Model C) are the same as that of the trained model (Trial 02) in the present study. The loss curve during the training of Model C is given in Figure S6. It is observed that there are sudden significant jumps in the loss curve, and the loss curve is also faced to converge within the same number of epochs. However, the Model A2 converges smoothly (Figure 9b) without any significant sudden peaks. It can be due to the architecture of Model A2 and gradient clipping in the first LSTM layer. The performance of Model C is also tested on separate datasets, that is, testing sets 01 and 02 (Figure S7). The MSE values for these testing sets 01 and 02 are 0.00305 and 0.08678, respectively. These values are relatively higher compared to Model A2. In addition,

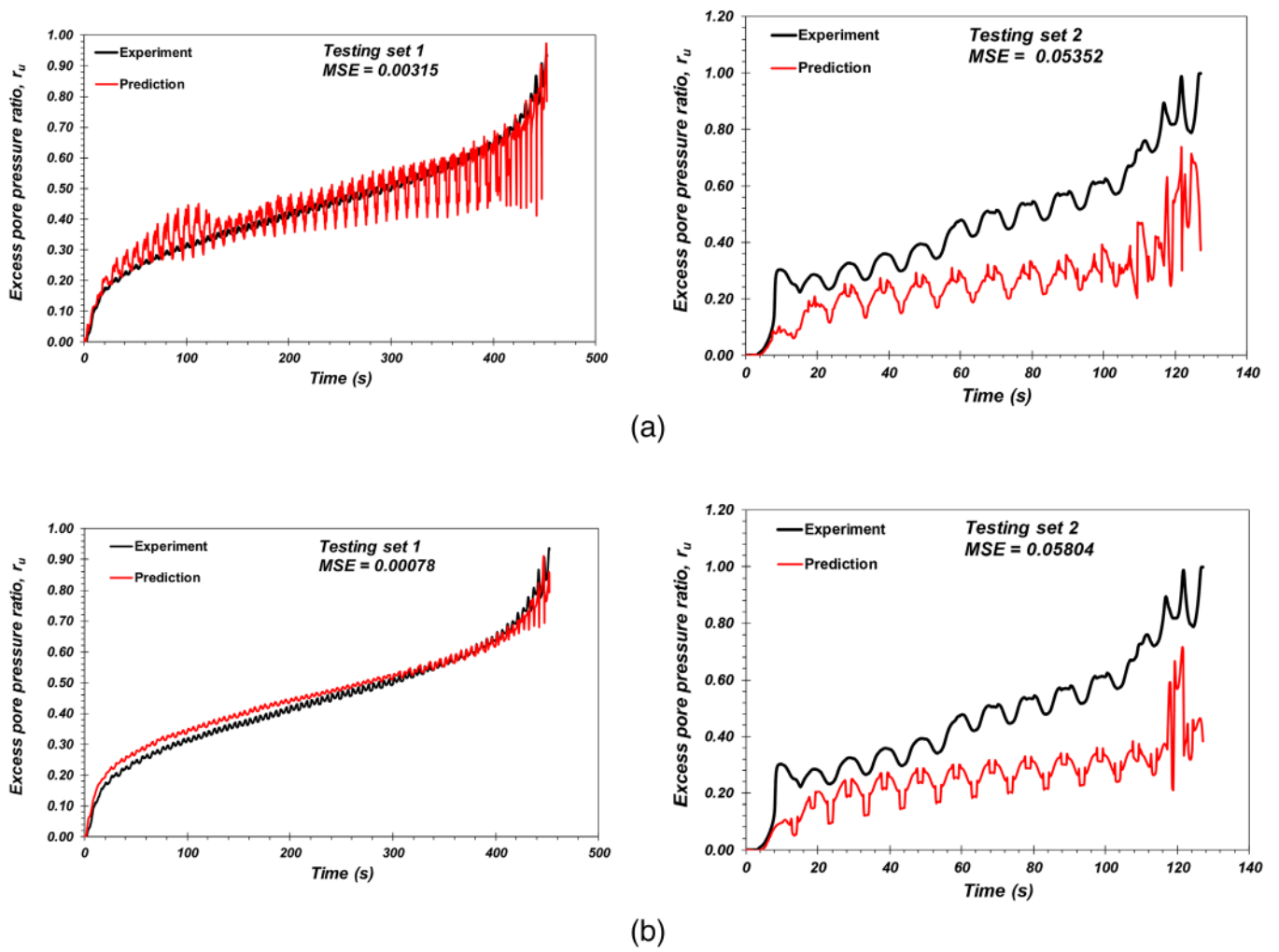


FIGURE 13 | Testing performance of r_u models for PDX sand corresponding to both the trials: (a) Trial 1; (b) Trial 2.

cycle-by-cycle analysis of Models A2 and C is conducted to understand the neuron activation statistics (Figure 15a,b). The average shear strain versus number of cycles for both models and the testing datasets are calculated and plotted (Figures 14c and S8). The scale-invariant metric, such as relative mean squared error (REMSE), is suitable for evaluating the performance of the models for datasets with a more significant fraction of small-magnitude data [29, 73]. Thus, in the present study, first, relative squared error (RSE) is computed for each cycle and then based on RSE values of each cycle, the REMSE is calculated for each case (Figure 15). The REMSE is defined as:

$$REMSE_{Overall} = \frac{1}{n} \sum_{i=1}^n \left(\frac{\hat{y}_i - y_i}{y'_i} \right)^2 \quad (5)$$

$$RSE_{(for\ i^{th}\ cycle)} = \left(\frac{\hat{y}_i - y_i}{y'_i} \right)^2 \quad (6)$$

where y'_i is the substitute of y_i , and i is the cycle number. If $y_i = 0$, then $y'_i = 1$; otherwise $y'_i = y_i$. This substitution helps to avoid division by zero in cases where $y_i = 0$. In the case of the testing set 1 (CSR = 0.125), the RSE values are close to zero for Model A2 (ReLU-ELU) except for the last five cycles (Figure 15a). On the other hand, the RSE values are comparatively higher

for Model C (tanh) w.r.t. Model A2 (ReLU-ELU). Also, in the case of higher cycles, such as 44th and 45th, the RSE values by Model A2 are significantly lower w.r.t. Model C. Only, in the case of 40–43 cycles, the RSE values are either almost the same or slightly higher for Model A2 w.r.t. Model C. REMSE for Model A2 is comparatively lower w.r.t. Model C, that is, overall better performance for Model A2. Thus, it can be concluded that Model A2 performs relatively better than Model C in the case of lower as well as higher cycles for testing set 1. In the case of the testing set 2 (CSR = 0.146), the RSE values are significantly lower (Figure 15b) for Model A2 (ReLU-ELU) w.r.t. Model C (tanh) up to the eighth cycle. Also, in the case of intermediate (9 and 10) and higher cycles (11 and 12), the RSE values are either almost the same or slightly higher for Model A2 w.r.t. Model C. REMSE for Model A2 is comparatively lower w.r.t. Model C, that is, overall better performance for Model A2. Thus, it can be concluded that Model A2 performs relatively better than Model C in the case of lower as well as intermediate and higher cycles for testing set 2.

The laboratory/experimental responses of r_u are always positive throughout the time history. The unbounded positive nature of the ReLU activation function also makes it suitable for modeling r_u response during the input shear stress reversal also. Recently, a few studies [28, 71, 72] have shown the advantages of ReLU in

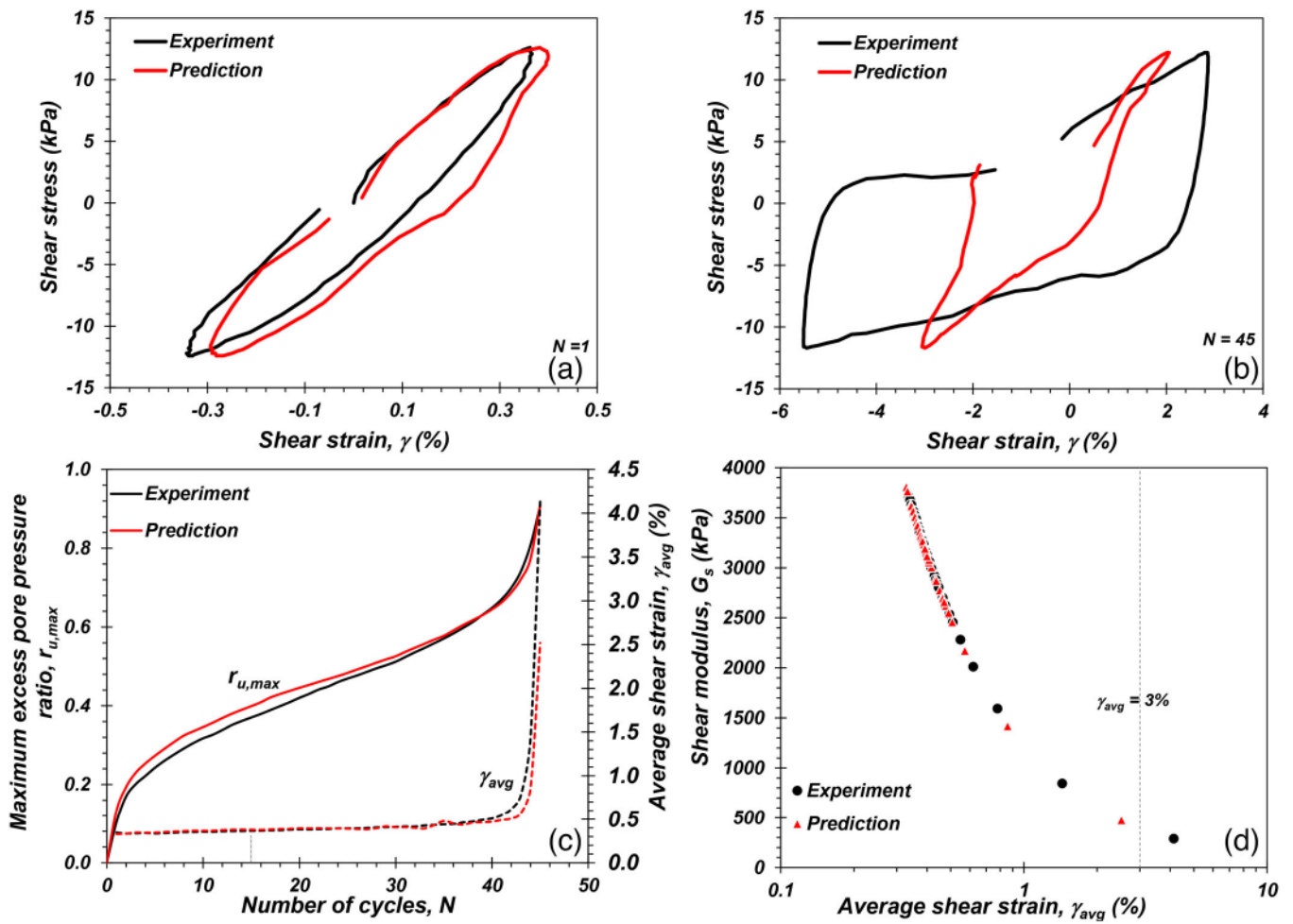


FIGURE 14 | Comparison of cyclic response between experimental data (Ts 1) and LSTM-based prediction: (a) shear stress and shear strain hysteresis loop during first cycle of loading; (b) shear stress and shear strain hysteresis loop during last cycle of loading; (c) development of maximum excess pore pressure ratio and average shear strain accumulation with number of loading cycles; (d) variation of shear modulus with average shear strain.

modeling the r_u time histories of various soils. Thus, the ReLU activation function is considered in the present study for the first LSTM layer of the r_u model because of these advantages. The second layer of the r_u model is the GRU, and the corresponding activation function is *softplus*. The *softplus* activation function in the second layer (GRU layer) further smoothens the output from the ReLU activation function in the first layer (LSTM). The word “plus” in *softplus* indicates its non-negative nature and has the following properties: smooth, non-negative, differentiable, and monotonically increasing. Therefore, this is a good option after the ReLU activation function for predicting r_u response. The GRU layer has the following advantages: (i) simpler architecture, (ii) fewer parameters, (iii) generalizability, (iv) higher training speed, and so forth [62]. Thus, the GRU layer after the traditional LSTM layer helps to generalize the prediction from the previous traditional LSTM layer and avoids overfitting. The cycle-dependent performance of the r_u model corresponding to the best/final trial (Trial 02) is evaluated (Figures 14c and S8a) in terms of RSE and REMSE and represented in terms of $r_{u,max}$. Figure 15d and f shows the cycle-dependent RSE and overall REMSE values for $r_{u,max}$. In the case of the testing set 1 (CSR = 0.125), the RSE value at the first cycle is a little higher, after that, it rapidly decreases, and then from the 25th cycle onwards, it is close to zero. Thus, the r_u

model can also perform better for testing set 1 at higher cycles (longer sequences). In the case of the testing set 2 (CSR = 0.146), the RSE value at the first cycle is slightly higher, after that, it rapidly decreases for cycles 2–5, and then again slightly increasing trend for cycles 6–11. However, at the 12th cycle, it rapidly drops to a lower RSE value; therefore, the model performs reasonably well at this higher cycle. The overall REMSE values for testing sets 1 and 2 are 0.006526 and 0.150619, respectively. Thus, from the RSE and REMSE values, it can be concluded that the overall cycle-dependent performance of the r_u model is reasonably well for the testing sets after the first cycle, intermediate cycles, and higher cycles. In addition, the cycle-dependent performance of the γ (%) and r_u models at two different stress levels (CSR = 0.125 and CSR = 0.146) provides a reasonable outcome. However, there is a need for further studies to verify the performance of the γ (%) and r_u models at very high-stress levels (CSR > 0.19).

3.3.3 | Ablation Study for Design Choices of Architecture

An ablation study is also performed for the γ (%) model to determine suitable architectural decisions, such as the number of

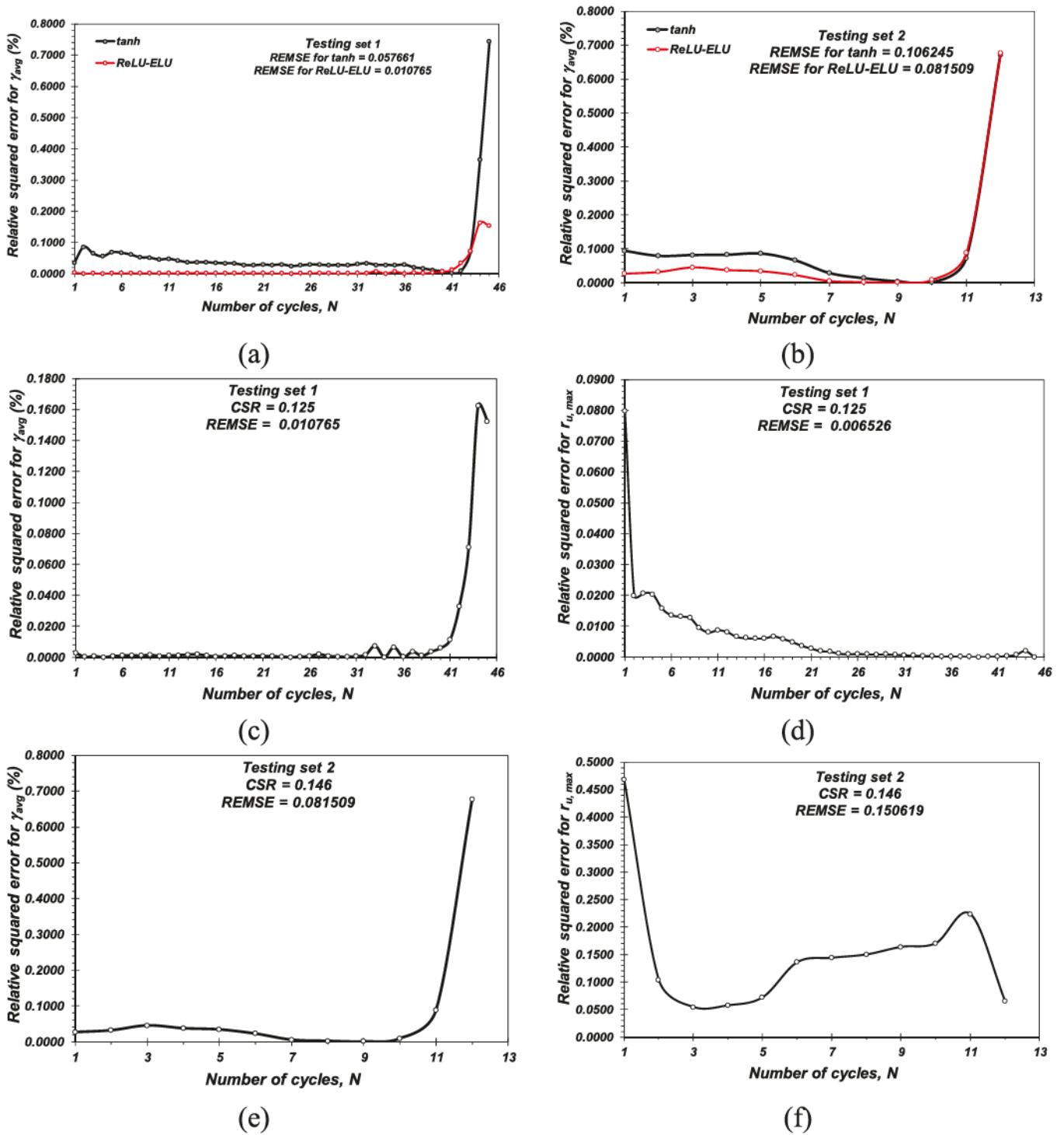


FIGURE 15 | Cycle-dependent performance of the LSTM-based models on both the testing sets having different stress levels: (a) and (b) two γ (%) models having different configurations of activation functions; (c), (d), (e), and (f) γ (%) and r_u models corresponding to best/final trial (Trial 02).

LSTM layers, the number of neurons (units), and the activation functions. Nine configurations (configs.) are considered for the ablation study by varying the number of LSTM layers, units in each LSTM layer, and the activation functions in each LSTM layer. The rest of the hyperparameters are considered as the default values for the γ (%) model. The performance of each configuration is evaluated on the testing set 01 in terms of MSE and other scale-invariant metrics such as NRMSE and SMAPE (Table S2). A horizontal bar chart (Figure 16) is prepared based on MSE values

for comparative analysis of these nine configurations, along with the initial trial (Trial 1) and the final/best trial (Trial 2). From the results, it is observed that the minimum and the maximum MSE values are 0.0028 and 0.00452, respectively. The minimum and the maximum MSE values correspond to the single LSTM layer with 64 units (config. 2), and three LSTM layers of each 64 units (config. 8), respectively. The MSE value (= 0.00285) corresponds to the config. 4 (two LSTM layers of each 32 units and ReLU-tanh activation functions) are almost close to the minimum value. The

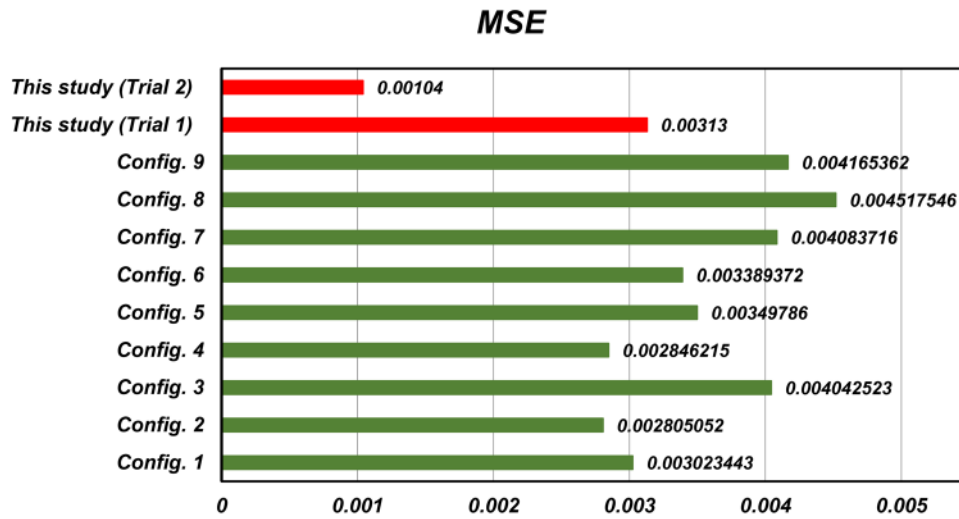


FIGURE 16 | Comparison of ablation study results for nine configurations with the initial trial (Trial 1), and the final adopted configurations (Trial 2) in the present study for γ (%) model.

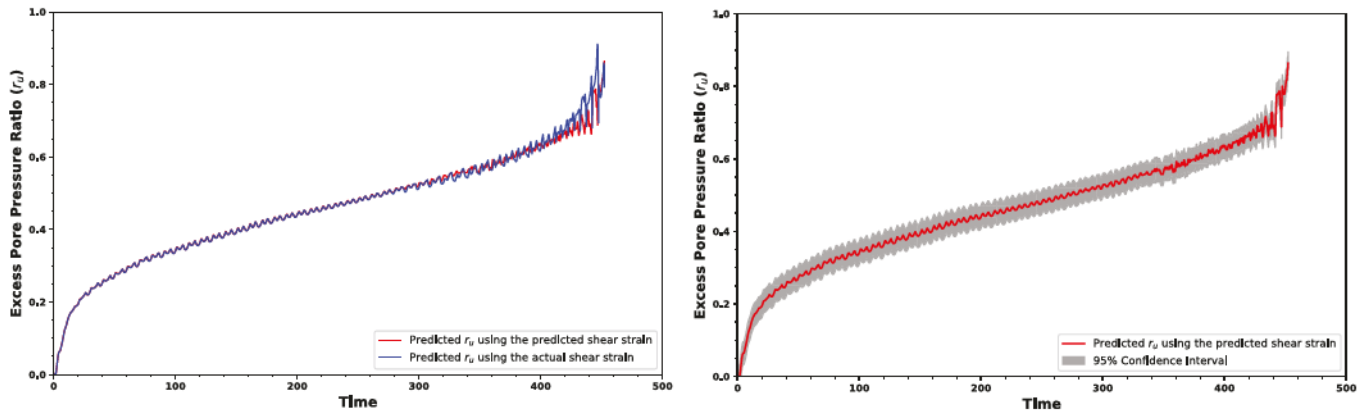


FIGURE 17 | (a) Comparison of r_u predictions using laboratory (actual) and predicted responses of γ (%); (b) The 95% CI plot of r_u using predicted γ (%) response. [CI = confidence interval].

MSE values of the remaining configurations fall within this range with slight variations. However, the MSE value (= 0.00104) for the γ (%) model (Trial 2) is significantly less w.r.t. minimum MSE value corresponding to config. 2 (Figure 16). Also, the MSE value corresponding to Trial 1 (initial trial) is less compared to some other configurations (3; 5–9). Therefore, it can be concluded from the ablation study that the architectural choices (Figure 7b; Tables S1 and S2) for Trial 02 (final/best trial) are suitable design choices out of all the configurations. It is feasible due to the different units and activation functions for each LSTM layer, and the suitable combination of other hyperparameters.

3.3.4 | Uncertainty Propagation in the Sequential Prediction Approach

A thorough error propagation analysis is conducted to validate the physical intuition of the two-step sequential prediction approach of the LSTM models (shear strain $\rightarrow$ PWP). The final trained r_u model (Trial 2) is tested on testing set 1 separately using laboratory and predicted responses of γ (%) (Figure 17a). The MSE value for these responses of r_u using laboratory and predicted

responses of γ (%) are 0.000782 and 0.001095, respectively. The addition of confidence intervals (CIs) in the predictions is helpful for understanding the uncertainty propagation in the sequential approach [74]. The 95% CIs between these two predictions are calculated (Figure 17b) to understand uncertainty propagation in the sequential approach. It is observed that the up to 300 s, r_u predictions using predicted γ (%) match well with the r_u predictions using laboratory γ (%) (Figure 17a). After that, the r_u predictions using predicted γ (%) are slightly over-predicted and, at the end, under-predicted (Figure 17a) w.r.t. r_u predictions using laboratory γ (%). It is observed that the width of the CI is dynamic (Figure 17b) for the r_u time history predictions using predicted γ (%) time history. It indicates that the model effectively captures varying levels of uncertainty over time, which is essential for real-world time series applications. The CI width is very narrow at the initial and final portions of the r_u time history, indicating high confidence in the predictions during these periods. It can be due to the reduced variability and complexity in input training data (e.g., predicted γ [%]) during these periods. In contrast, the intermediate portion of the r_u time history exhibits an almost constant and moderate CI width, which suggests a stable/consistent but slightly moderate level of uncertainty

compared to the initial and final stages of the time history. The intermediate portion could correspond to periods of more complex input training data (γ [%]) but well-captured patterns by the r_u time history model.

3.3.5 | Analysis of Loss Curves

The plots of MSE loss versus epoch are essential for assessing the training progress of the DL models. They provide insights into how well the model learns over epochs and help identify the issues: underfitting, overfitting, sensitivity to the training datasets, and so forth. Figure 9 shows the training performance of the γ (%) and r_u models in terms of MSE loss versus epoch. The γ (%) model for Trial 1 (named Model A1) is trained for five epochs only (Figure 9a). The MSE loss decreases from 0.02789 to 0.02486. However, there is a sudden jump in the MSE loss at the fourth epoch. This can be due to the following reasons: (i) improper model architecture and hyperparameters and (ii) sensitivity to the training datasets. On the other hand, the γ (%) model for Trial 2 (named Model A2) shows a smoother descent in the MSE loss curve (Figure 9b). The MSE loss decreases from 0.02552 to 0.00649 after the model has been trained for 35 epochs. The nature of this curve indicates stable and consistent learning through epochs. The training performance of the γ (%) model is improved over the trials because of the exclusion of sensitive training datasets and tuning of the hyperparameters. The r_u model for Trial 1 (named Model B1) is trained for 30 epochs, and the corresponding MSE loss decreases from 0.011201 to 0.00224 (Figure 9c). Instead, the r_u model for Trial 2 (named Model B2) is trained for 25 epochs and the corresponding MSE loss decreases from 0.00736 to 0.00139 (Figure 9d). Though there is a sudden jump at the tenth epoch for both trials, the overall trials indicate effective training and convergence. This indicates that the suitable architecture and the associated hyperparameters are considered for the r_u models in the case of both trials. However, Model B2 shows better training performance than Model B1 because a few sensitive training datasets are excluded during the model training. The removal of these datasets makes the final five training datasets well-balanced in terms of *time sequence lengths*. Therefore, these balanced datasets increase the predictive ability of the models corresponding to the best trial (Trial 2).

3.3.6 | Training and Testing Performance Analysis

The γ (%) and r_u models are tested on two independent/separate testing datasets. The MSE values are used to evaluate their predictive power on unseen/independent testing datasets. From the initial to best trial (Trial 2), the MSE value decreases from 0.00313 to 0.00104 in the case of the shear strain model for the first testing set (Figure 12). In the case of the second testing set, the MSE value decreases from 0.08726 to 0.08677 (Figure 12). The shear strain model corresponding to the best trial predicts the 3% single-amplitude shear strain [55] for the first testing set at 449.02 s ($N = 44.9$); however, w.r.t. laboratory response, it has occurred at 448.08 s ($N = 44.8$). In the case of the second testing set, the shear strain model corresponding to the best trial predicts the shear strain of 1.04236% instead of the laboratory response of 3% shear strain observed at 113.89 s. For the r_u model (Figure 13), the

MSE value corresponding to the first testing set decreases from 0.00315 to 0.00078 over the trials. On the other hand, the MSE value slightly increases from 0.05352 to 0.05804 for the second testing set. These lower MSE losses for the testing sets indicate the superior generalization performance of the models. The detailed analyses of the testing and training performance of these models are given in Tables 3 and S3.

3.3.7 | Analysis and Effect of the Imbalance Problem

In the present study, the DL models are developed considering a purely data-driven framework. Therefore, the training and testing performances of the DL models are influenced by the constraints associated with the training datasets. One such constraint is the imbalance of the training datasets. The nine training datasets do not have the proper proportion in terms of low, medium, and high *time sequence length* (Figure 8). The datasets are classified in this study as low, medium, and high if the *time sequence length* is less than 340 s, within 340–760 s, and more than 1000 s, respectively. The proportion of low, medium, and high *time sequence length* are one (Tr 3), five (Tr 1, 2, 6, 8, 9), and three (Tr 4, 5, 7), respectively. Therefore, it shows the imbalance problem in the training datasets corresponding to the initial trial. An attempt has been made to improve the predictive ability of the models' corresponding to the initial trial by solving the imbalance problem. Therefore, some of the datasets are excluded (Figure 8) for the best trial (Trial 2). The five training datasets have proper proportion in *time sequence length*. Thus, the DL models predict the responses reasonably well for the first testing set compared to the second testing set for the best trial (Tables 3 and S3). Hence, in future, the training datasets from the laboratory tests can be prepared based on this strategy for developing more robust DL models.

3.4 | Comparison of the Proposed LSTM-Based Framework with the Existing Models

In this section, a comparative analysis of the proposed LSTM-based modeling framework is made with the existing LSTM-based models. The proposed LSTM-based models are compared with the existing LSTM-based models of Zhang et al. [28], Meng and Pei [30], and Jas et al. [59]. For this purpose, the same LSTM-based framework is retrained with the database [75] of Ottawa F65 sands. The training and testing datasets are the same as those of Jas et al. [59]. Figure 18 shows the relative comparison of the responses of this study w.r.t. Jas et al. [59]. The responses of Zhang et al. [28] and Meng and Pei [30] are represented in Figure 19. The responses obtained from this study are plotted w.r.t. number of cycles (N) instead of time. Also, instead of r_u , the vertical effective stress (σ'_v) is plotted in Figure 18b,d. This will help to compare all the plots of this study (Figures 18 and 19) with the existing studies.

The trained γ (%) model of Ottawa F65 sand predicted the overall trend of the shear strain time history reasonably well, as observed in the experiment (Figure 18a). The MSE value decreased from 3.8334 (Figure 18c) to 3.0980 (Figure 18a) w.r.t. Jas et al. [59]. Thus, the overall performance of the γ (%) model is improved by 19.18 % in the present study for Ottawa F65 sand. The predicted σ'_v response of Ottawa F65 sand is somewhat variable in the initial

TABLE 3 | Detailed analysis of the testing performance of the LSTM-based models of PDX sand.

a) Detailed analysis of the testing performance of the shear strain models of PDX sand	
Shear strain model, testing set - 1, trial number 1	
<i>Observation</i>	<i>Explanation</i>
1. A sudden jump in predicted shear strain is observed in between 100 and 200 s.	1. This shear strain prediction after 100 s is influenced by (Figure S4) Tr 3. The same is observed during the training process for training datasets 1, 2, and 3.
2. The sudden jump in actual shear strain after 400 s is not predicted.	2. The total number of samples in the training datasets for this kind of jump around/after 400 s is limited as compared to other samples in the training datasets.
Shear strain model, testing set - 1, trial number 2	
<i>Observation</i>	<i>Explanation</i>
1. No such sudden jump in shear strain is observed in between 100 and 200 s.	1. Because Tr 3 is removed for this trial. Thus, the model is trained properly within this range.
2. The model is able to predict the sudden jump in actual shear strain after 400 s.	2. Because in this trial, some of the datasets are removed. Thus, the number of samples for this kind of sudden jump at the end is more compared to Trial 1. Figure 10 shows how the model is trained properly to predict the sudden jump in shear strain at the end of most datasets.
Shear strain model, testing set - 2, trial number 1	
<i>Observation</i>	<i>Explanation</i>
The model captures the pattern precisely up to 110 s but not the sudden jump in actual shear strain after 110 s.	As in the training datasets, the number of samples within this range (0–110 s) are more compared to the number of samples within the sudden jump region at end of time history.
Shear strain model, testing set - 2, trial number 2	
<i>Observation</i>	<i>Explanation</i>
The model predicts the pattern almost the same as in the case of Trial 1.	The explanations are the same as in Trial 1. However, tuning of the hyperparameters may improve the model performance slightly.
b) Detailed analysis of the testing performance of the excess pore pressure ratio models of PDX sand	
Excess pore pressure ratio model, testing set - 1, trial number 1	
<i>Observation</i>	<i>Explanation</i>
Though in some regions, the model under or over-predicted the response frequently, however, it captures the loading and unloading pattern reasonably well. In addition, like experiment data it restricts the predicted maximum value of r_u to 1.	The model is neither underfitted nor overfitted. Therefore, it shows the generalized performance based on the training using the nine training datasets (Figure S5).
Excess pore pressure ratio model, testing set - 1, trial number 2	
<i>Observation</i>	<i>Explanation</i>
1. Though there are slightly over predictions up to 300 s, however, the response is smoother. Overall, the model captures the loading and unloading pattern more precisely.	The model is trained using five training datasets. Out of them, two training datasets (Tr 1 and Tr 2) have similar patterns (Figure 11) in r_u response compared to Ts 1. Hence, the trained model can capture the pattern precisely for the independent/unseen testing set.
2. Predicted all the peak values satisfactorily after 300 s.	
3. Restrict the predicted maximum value of r_u to 1.	

(Continues)

TABLE 3 | (Continued)

Excess pore pressure ratio model, testing set - 2, trial number 1	
Observation	Explanation
1. The model captures the loading and unloading pattern precisely, however, it underpredicted the pattern throughout the time history. 2. On the other hand, it is able to predict a certain jump in r_u after 105 s.	1. The model is neither underfitted nor overfitted; therefore, exhibited a better generalization performance. 2. The model is able to predict the sudden jump because of the sudden jump in the input shear strain time history.
Excess pore pressure ratio model, testing set - 2, trial number 2	
Observation	Explanation
The predicted response is almost the same as Trail 1, but relatively smoother.	The explanations are the same as in Trial 1. However, tuning of the hyperparameters may enhance the model performance slightly.

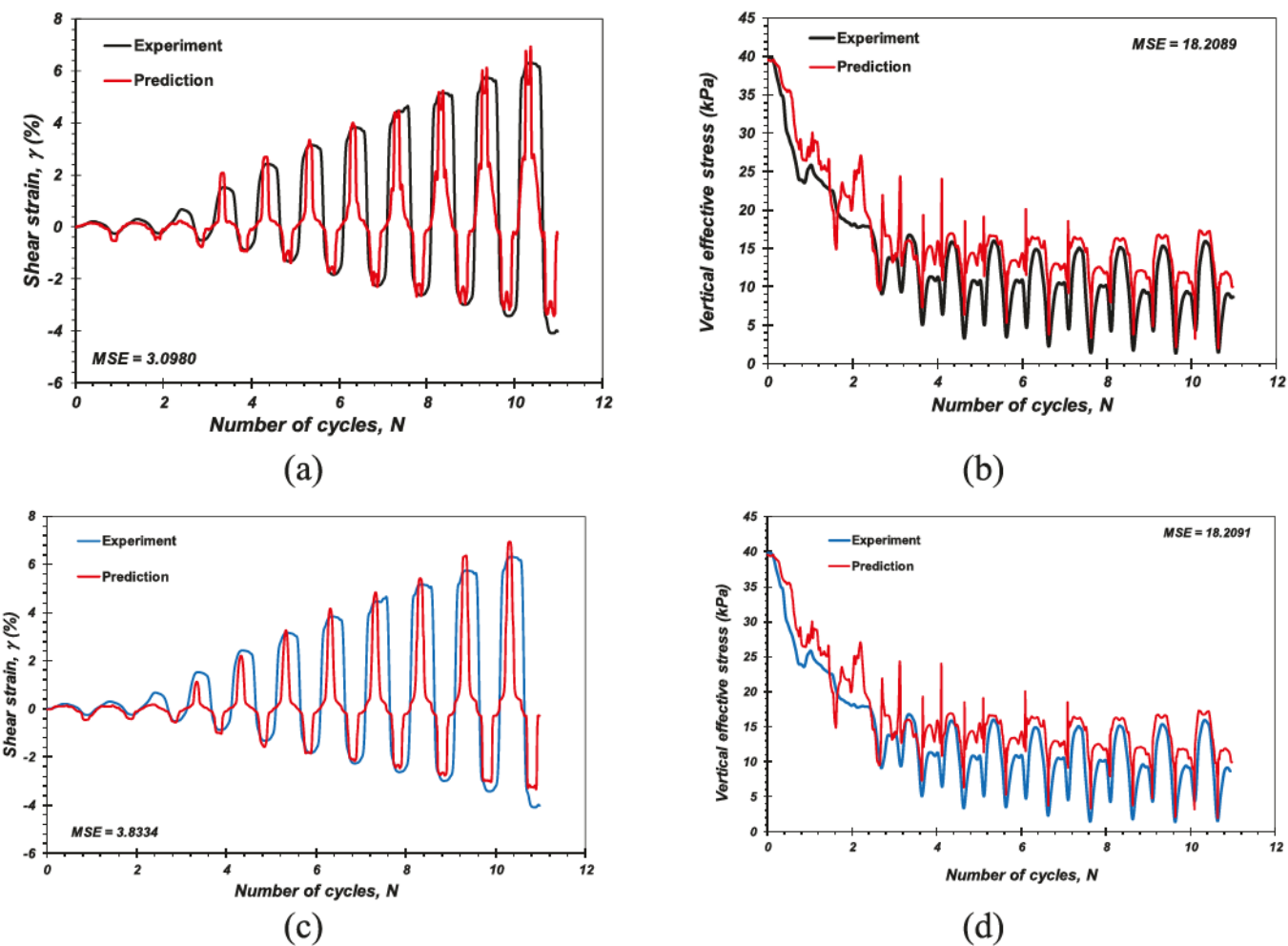


FIGURE 18 | Relative comparison of the responses obtained from two models for Ottawa F65 sand: (a) shear strain response of this study; (b) vertical effective stress response of this study; (c) shear strain response of Jas et al. [59]; (d) vertical effective stress response of Jas et al. [59].

few cycles, and after that, the response matches satisfactorily with the experimental results right up to the end of the time history (Figure 18b). The MSE value is decreased from 18.2091 (Figure 18d) to 18.2089 (Figure 16b) w.r.t. Jas et al. [59]. Thus, the

overall performance of the r_u model is slightly improved in the present study for Ottawa F65 sand. Zhang et al. [28] predicted the responses (Figure 19a,b) of Fontainebleau sand reasonably well compared to the responses obtained from cyclic simple shear

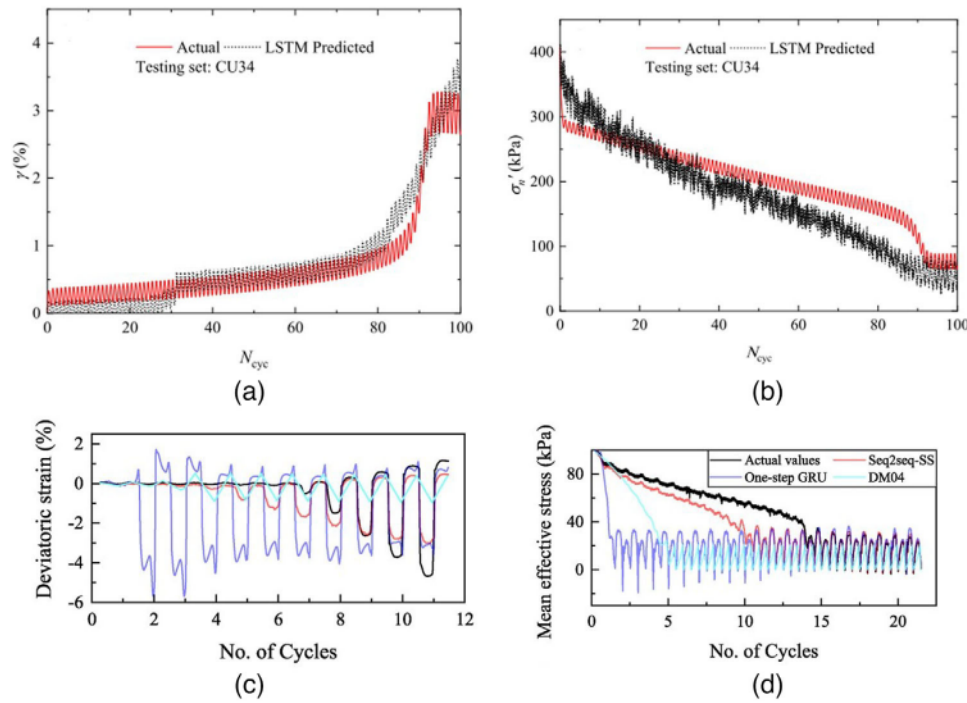


FIGURE 19 | Relative comparison of the responses for different sands from recent studies: (a) shear strain response of Fontainebleau sand by Zhang et al. [28]; (b) vertical effective stress response of Fontainebleau sand by Zhang et al. [28]; (c) deviatoric strain response of Ottawa F65 by Meng and Pei [30]; (d) mean effective stress of Ottawa F65 by Meng and Pei [30].

test. The LSTM-based models of Zhang et al. [28] are trained using the synthetic datasets obtained from advanced constitutive models. Up to 30 s, the γ (%) model underpredicted the response, after that, it is overpredicted (Figure 19a). However, the model can predict the sudden jump in γ (%) after the 80 s. The LSTM model predicted σ'_v response (Figure 19b) is under or overpredicted significantly w.r.t. the laboratory response; however, the overall trend is captured by the LSTM model. Meng and Pei [30] developed a GRU-based seq2seq using cyclic triaxial test databases of Ottawa F65 sand. The database of experimental tests is improved using the cubic spline interpolation. Thus, the training datasets are the combination of experimental datasets and synthetic datasets. Figure 19c,d shows the responses obtained from this study and the response obtained from DM04 model. The legend of both the plots is the same. Finally, they show a satisfactory prediction compared to the DM04 model. However, the trained model is still unable to capture the experimental response satisfactorily.

Therefore, from Figures 18 and 19, it is clear that the model trained using synthetic datasets can simulate the response satisfactorily. Thus, they are unable to provide more insights than the traditional constitutive model. On the other hand, the LSTM-based models trained using only laboratory datasets have a certain potential to predict the experimental response satisfactorily. However, in the future, these models can be used as alternative tools in practice after further minimizing their constraints and limitations. In summary, these models predict the generalized response through the training process from the training datasets. The responses obtained from these LSTM-based models are not truly comparable because of the following reasons: (i) unavailability of training datasets with details, model

or code files, and architectures of the models, (ii) testing on different datasets, (iii) different laboratory tests, equipment and sands, and (iv) uncertainties associated during the laboratory testing. Therefore, actual comparison is possible only when these constraints can be solved.

4 | Conclusions

This study presents both the experimental insight and data-driven AI-based constitutive models of poorly graded reconstituted sand, that is, PDX sand. The following conclusions can be made from this study:

1. The experimental results on PDX sand resulted in $CRR_{N=15}$ of 0.146 and 0.136 for two different σ'_{vc} of 100 kPa and 240 kPa, respectively, which are lower than the V_s -based $CRR_{Mw=7.5}$ (0.21 and 0.23). The threshold value of $\Delta w/\sigma'_{vc}$ for the PDX sand ranges from 0.0008 to 0.001, at this state a rapid change in $\Delta w/\sigma'_{vc}$ occurs which in turn triggers the liquefaction. However, this ratio is slightly lower than that reported in the literature [57].
2. The AI-based shear strain (γ [%]) model trained using the PDX sand dataset predicts the responses reasonably well in the case of the number of γ (%) cycles for the unseen/independent testing sets. The AI-based excess pore pressure ratio (r_u) model trained using the same dataset predicts the r_u response well and closely follows the laboratory-based response for uniform sinusoidal loading. The differences in the predicted and actual responses are

due to the number of training datasets and the pattern of input-output time histories.

3. The accumulation of r_u is gradual; thus, the model satisfactorily captured the general trend in the response. The γ (%) close to liquefaction is highly nonlinear and rapidly varying when the r_u reaches close to unity. The models are trained using fewer training samples in the dataset in the region near the rapidly varying γ (%), and hence, these models face difficulties during prediction.
4. The LSTM and its variants have an inherent “forgetting problem” for longer sequences during the training and testing phases. Thus, specific developments in algorithmic structures are necessary to overcome this problem. The models cannot capture the hysteresis loop close to liquefaction because the number of data points in that region is smaller due to the rapid rise in strain accumulation compared to the other parts of time histories.

There exists inherent uncertainty with the laboratory databases where each test varies from one to another in response to soil fabric, anisotropy, and relative density, even when the tests are performed with the utmost care. Hence, the AI-based models face difficulties while capturing these uncertainties, as in the case of existing constitutive models. Undoubtedly, in future, these novel AI-based models developed herein could be trained and tuned with a large number of global datasets to predict the cyclic response of various soils around the world more realistically. Then, these unified AI-based models could be used as alternative tools beside the existing constitutive models. It is important to note that the developed and existing AI-based models are trained only for uniform and nonuniform sinusoidal loading; hence, at this moment, they are not able to predict the cyclic responses of the soils under transient loading. The work on this specific case is currently underway.

Acknowledgments

The authors would like to thank the support provided by the computational geomechanics laboratory and HPCE laboratory of Indian Institute of Technology Madras, Chennai, India and express their gratitude to the Ministry of Education, Govt. of India for providing the funding in the form of Prime Minister's Research Fellowship. The authors sincerely thank Professor Armin W. Stuedlein for his valuable assistance in acquiring the experimental data on the PDX sand. The authors would like to thank the editors and two anonymous reviewers for their valuable suggestions, which have improved the final version of the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Some or all data, models, or code generated or used during the study will be available in a repository online (<https://github.com/Kaushikjas10/CyclicResponse-PDX-LSTM-KJAJGRD.git>) in accordance with funder data retention policies.

References

1. R. W. Boulanger and I. M. Idriss, “Closure to “Liquefaction Susceptibility Criteria for Silts and Clays” by Ross W. Boulanger and I. M. Idriss,” *Journal of Geotechnical and Geoenvironmental Engineering* 134, no. 7 (2008): 1027–1028, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2008\)134:7\(1027\)](https://doi.org/10.1061/(ASCE)1090-0241(2008)134:7(1027)).
2. R. W. Boulanger and I. M. Idriss, *CPT and SPT Based Liquefaction Triggering Procedures* (Davis, USA: Center for Geotechnical Modeling, Department of Civil and Environmental Engg., University of California, 2014): 1–134. Report No.: UCD/CGM-14.
3. K. O. Cetin, R. B. Seed, R. E. Kayen, et al., “SPT-Based Probabilistic and Deterministic Assessment of Seismic Soil Liquefaction Triggering Hazard,” *Soil Dynamics and Earthquake Engineering* 115, no. Dec (2018): 698–709, <https://doi.org/10.1016/j.soildyn.2018.09.012>.
4. R. Dobry, R. S. Ladd, F. Y. Yokel, R. M. Chung, and D. Powell, *Prediction of Pore Water Pressure Buildup and Liquefaction of Sands during Earthquakes by the Cyclic Strain Method*, vol. 138 (Gaithersburg, MD: National Bureau of Standards, 1982), 150.
5. R. E. S. Moss, R. B. Seed, R. E. Kayen, J. P. Stewart, A. Der Kiureghian, and K. O. Cetin, “CPT-Based Probabilistic and Deterministic Assessment of in Situ Seismic Soil Liquefaction Potential,” *Journal of Geotechnical and Geoenvironmental Engineering* 132, no. 8 (2006): 1032–1051, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2006\)132:8\(1032\)](https://doi.org/10.1061/(ASCE)1090-0241(2006)132:8(1032)).
6. H. B. Seed, K. Tokimatsu, L. F. Harder, and R. M. Chung, “Influence of SPT Procedures in Soil and Liquefaction Resistance Evaluations,” *Journal of Geotechnical Engineering* 111, no. 12 (1985): 1425–1445, [https://doi.org/10.1061/\(ASCE\)0733-9410\(1985\)111:12\(1425\)](https://doi.org/10.1061/(ASCE)0733-9410(1985)111:12(1425)).
7. T. L. Youd and I. M. Idriss, “Liquefaction Resistance of Soils: Summary Report from the 1996 NCEER and the 1998 NCEER/NSF Workshops on the Evaluation of Liquefaction Resistance of Soils,” *Journal of Geotechnical and Geoenvironmental Engineering* 127, no. 4 (2001): 297–313, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2001\)127:4\(297\)](https://doi.org/10.1061/(ASCE)1090-0241(2001)127:4(297)).
8. M. H. Beaty and P. M. Byrne, UBCSAND constitutive model: Version 904aR. Documentation Report: UBCSAND Constitutive Model on Itasca UDM Web Site; February 2011 (2011), <http://www.itsasca-udm.com/pages/UBCSand.html>.
9. R. W. Boulanger and K. Ziotopoulou, *PM4Sand (version 3.1): A Sand Plasticity Model for Earthquake Engineering Applications* (Davis, CA: Center for Geotechnical Modeling, University of California, 2017). Report No. UCD/CGM-17/01.
10. Y. F. Dafalias and M. T. Manzari, “Simple Plasticity Sand Model for Accounting for Fabric Change Effects,” *Journal of Engineering Mechanics* 130, no. 6 (2004): 622–634, [https://doi.org/10.1061/\(ASCE\)0733-9399\(2004\)130:6\(622\)](https://doi.org/10.1061/(ASCE)0733-9399(2004)130:6(622)).
11. A. Elgamal, Z. Yang, and E. Parra, “Computational Modeling of Cyclic Mobility and Post-Liquefaction Site Response,” *Soil Dynamics and Earthquake Engineering* 22 (2002): 259–271, [https://doi.org/10.1016/S0267-7261\(02\)00022-2](https://doi.org/10.1016/S0267-7261(02)00022-2).
12. D. Basu, J. Montgomery, and A. W. Stuedlein, “Postliquefaction Reconsolidation Settlement of a Soil Deposit Considering Spatially Variable Properties and Ground Motion Variability,” *Journal of Geotechnical and Geoenvironmental Engineering* 150, no. 3 (2024): 04024001, <https://doi.org/10.1061/JGGEFK.GTENG-11768>.
13. M. H. Beaty, “Application of UBCSAND to the LEAP Centrifuge Experiments,” *Soil Dynamics and Earthquake Engineering* 104 (2018): 143–153, <https://doi.org/10.1016/j.soildyn.2017.10.006>.
14. M. Khosravi and S. Zaregarizi, “Liquefaction Potential and Sediment Ejecta Manifestation of Thinly Interbedded Sands and Fine-Grained Soils: Palinurus Road Site in Christchurch Subjected to 2010–2011 Canterbury Earthquake Sequence,” *Journal of Geotechnical and Geoenvironmental Engineering* 150, no. 5 (2024): 04024026, <https://doi.org/10.1061/JGGEFK.GTENG-11616>.

15. M. T. Manzari, M. El Ghoraiy, M. Zeghal, et al., "LEAP-2017 Simulation Exercise: Calibration of Constitutive Models and Simulation of the Element Tests," in *Model Tests and Numerical Simulations of Liquefaction and Lateral Spreading: LEAP-UCD-2017*, eds. B. Kutter, M. Manzari, and M. Zeghal (Springer International Publishing, 2020), 165–185, https://doi.org/10.1007/978-3-030-22818-7_9.
16. Y. M. A. Hashash, Q. Fu, J. Ghaboussi, P. V. Lade, and C. Saucier, "Inverse Analysis-Based Interpretation of Sand Behavior from Triaxial Compression Tests Subjected to Full End Restraint," *Canadian Geotechnical Journal* 46, no. 7 (2009): 768–791, <https://doi.org/10.1139/T09-015>.
17. A. Johari, A. A. Javadi, and G. Habibagahi, "Modelling the Mechanical Behaviour of Unsaturated Soils Using a Genetic Algorithm-Based Neural Network," *Computers and Geotechnics* 38, no. 1 (2011): 2–13, <https://doi.org/10.1016/j.compgeo.2010.08.011>.
18. V. R. Kohestani and M. Hassanlourad, "Modeling the Mechanical Behavior of Carbonate Sands Using Artificial Neural Networks and Support Vector Machines," *International Journal of Geomechanics* 16, no. 1 (2016): 04015038, [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000509](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000509).
19. S. Lbibb and M. T. Manzari, "Stress-Strain Behavior of Ottawa Sand in Cyclic Direct Simple Shear and Modeling of Cyclic Strength Using Artificial Neural Networks," *Soil Dynamics and Earthquake Engineering* 164 (2023): 107585, <https://doi.org/10.1016/j.soildyn.2022.107585>.
20. P. Lin, R. Ratnam, H. Sankari, and A. Garg, "Mechanism of Microstructural Variation under Cyclic Shearing of Shantou Marine Clay: Experimental Investigation and Model Development," *Geotechnical and Geological Engineering* 37, no. 5 (2019): 4163–4210, <https://doi.org/10.1007/s10706-019-00900-6>.
21. A. A. Javadi and M. Rezaia, "Applications of Artificial Intelligence and Data Mining Techniques in Soil Modeling," *Geomechanics and Engineering* 1, no. 1 (2009): 53–74, <https://doi.org/10.12989/gae.2009.1.1.053>.
22. A. A. Javadi, A. Faramarzi, and A. Ahangar-Asr, "Analysis of Behaviour of Soils under Cyclic Loading Using EPR-Based Finite Element Method," *Finite Elements in Analysis and Design* 58 (2012): 53–65, <https://doi.org/10.1016/j.finel.2012.04.005>.
23. A. Nassr, M. Esmaeili-Palak, H. Katebi, and A. Javadi, "A New Approach to Modeling the Behavior of Frozen Soils," *Engineering Geology* 246 (2018): 82–90, <https://doi.org/10.1016/j.enggeo.2018.09.018>.
24. H. Zhao, Z. Huang, and Z. Zou, "Simulating the Stress-Strain Relationship of Geomaterials by Support Vector Machine," *Mathematical Problems in Engineering* 2014 (2014): 1–7, <https://doi.org/10.1155/2014/482672>.
25. J.-H. Zhu, M. M. Zaman, and S. A. Anderson, "Modelling of Shearing Behaviour of a Residual Soil with Recurrent Neural Network," *International Journal for Numerical and Analytical Methods in Geomechanics* 22, no. 8 (1998): 671–687, [https://doi.org/10.1002/\(SICI\)1096-9853\(199808\)22:8<3C671::AID-NAG939>3E3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1096-9853(199808)22:8<3C671::AID-NAG939>3E3.0.CO;2-Y).
26. J.-H. Zhu, M. M. Zaman, and S. A. Anderson, "Modeling of Soil Behavior with a Recurrent Neural Network," *Canadian Geotechnical Journal* 35, no. 5 (1998): 858–872, <https://doi.org/10.1139/t98-042>.
27. Y. Choi and K. Kumar, "A Machine Learning Approach to Predicting Pore Pressure Response in Liquefiable Sands under Cyclic Loading," in *Geo-Congress 2023* (2023), 202–210, <https://doi.org/10.1061/9780784484692.021>.
28. P. Zhang, Z. Y. Yin, Y. F. Jin, and G. L. Ye, "An AI-Based Model for Describing Cyclic Characteristics of Granular Materials," *International Journal for Numerical and Analytical Methods in Geomechanics* 44, no. 9 (2020): 1315–1335, <https://doi.org/10.1002/nag.3063>.
29. N. Zhang, A. Zhou, Y. F. Jin, Z. Y. Yin, and S. L. Shen, "An Enhanced Deep Learning Method for Accurate and Robust Modelling of Soil Stress–Strain Response," *Acta Geotechnica* 18 (2023): 4405–4427, <https://doi.org/10.1007/s11440-023-01813-8>.
30. F. Meng and H. Pei, "Cyclic Shear Stress-Strain Prediction of Saturated Sand Based on the Unrolled Seq2Seq Model and Scheduled Sampling," *Soil Dynamics and Earthquake Engineering* 165 (2023): 107665, <https://doi.org/10.1016/j.soildyn.2022.107665>.
31. Q. Z. Guan and Z. X. Yang, "Hybrid Deep Learning Model for Prediction of Monotonic and Cyclic Responses of Sand," *Acta Geotechnica* 18, no. 3 (2022): 1447–1461, <https://doi.org/10.1007/s11440-022-01656-9>.
32. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation* 9, no. 8 (1997): 1735–1780, <https://doi.org/10.1162/neco.1997.9.8.1735>.
33. S. Sideras, *Evolutionary Intensity Measures for More Accurate and Informative Evaluation of Liquefaction Triggering* (PhD Dissertation, Seattle, WA: University of Washington, 2019).
34. K. Jas and G. R. Dodagoudar, "Explainable Machine Learning Model for Liquefaction Potential Assessment of Soils Using XGBoost-SHAP," *Soil Dynamics and Earthquake Engineering* 165 (2023): 107662, <https://doi.org/10.1016/j.soildyn.2022.107662>.
35. K. Jas and G. R. Dodagoudar, "Liquefaction Potential Assessment of Soils Using Machine Learning Techniques: A State-of-the-Art Review from 1994–2021," *International Journal of Geomechanics* 23, no. 7 (2023): 03123002, <https://doi.org/10.1061/IJGNAL.GMENG-7788>.
36. K. Jas, S. Mangalathu, and G. R. Dodagoudar, "Evaluation and Analysis of Liquefaction Potential of Gravelly Soils Using Explainable Probabilistic Machine Learning Model," *Computers and Geotechnics* 167 (2024): 106051, <https://doi.org/10.1016/j.compgeo.2023.106051>.
37. M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning* (Cambridge: MIT Press, 2018).
38. P. Ni, S. Mangalathu, and Y. Yi, "Fragility Analysis of Continuous Pipelines Subjected to Transverse Permanent Ground Deformation," *Soils Foundation* 58, no. 6 (2018): 1400–1413, <https://doi.org/10.1016/j.sandf.2018.08.002>.
39. J. Macedo and L. Vergaray, "State-Based Assessment of Cyclic Liquefaction Manifestation," *Soil Dynamics and Earthquake Engineering* 179 (2024): 108520, <https://doi.org/10.1016/j.soildyn.2024.108520>.
40. P. Zhang, Z. Y. Yin, and Y. F. Jin, "State-of-the-art Review of Machine Learning Applications in Constitutive Modeling of Soils," *Archives of Computational Methods in Engineering* 28, no. 5 (2021): 3661–3686, <https://doi.org/10.1007/s11831-020-09524-z>.
41. R. S. Ladd, "Specimen Preparation and Cyclic Stability of Sands," *Journal of the Geotechnical Engineering Division* 103, no. 6 (1977): 13014, <https://doi.org/10.1061/AJGEB6.0000435>.
42. T. Park and M. L. Silver 1975. *Dynamic Soil Properties Required to Predict the Dynamic Behavior of Elevated Transportation Structures*. U.S. Dept. of Transportation Report DOT-TST-75-44.
43. H. B. Seed, "Soil Liquefaction and Cyclic Mobility Evaluation for Level Ground during Earthquakes," *Journal of the Geotechnical Engineering Division* 105, no. 2 (1979): 201–255, <https://doi.org/10.1061/AJGEB6.0000768>.
44. J. P. Mulilis, *The Effects of Method of Sample Preparation on the Cyclic Stress-Strain Behaviour of Sands* (Technical Report, Univ. of California at Berkeley, 1975), 75.
45. J. Zheng and R. D. Hryciw, "Traditional Soil Particle Sphericity, Roundness and Surface Roughness by Computational Geometry," *Géotechnique* 65, no. 6 (2015): 494–506, <https://doi.org/10.1680/geot.14.P.192>.
46. M. A. Maroof, A. Mahboubi, A. Noorzad, and Y. Safi, "A New Approach to Particle Shape Classification of Granular Materials," *Transportation Geotechnics* 22 (2020): 100296, <https://doi.org/10.1016/j.trgeo.2019.100296>.

47. R. Kayen, R. E. S. Moss, E. M. Thompson, et al., "Shear-Wave Velocity-Based Probabilistic and Deterministic Assessment of Seismic Soil Liquefaction Potential," *Journal of Geotechnical and Geoenvironmental Engineering* 139, no. 3 (2013): 407–419, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0000743](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000743).
48. J. S. Lee and J. C. Santamarina, "Bender Elements: Performance and Signal Interpretation," *Journal of Geotechnical and Geoenvironmental Engineering* 131, no. 9 (2005): 1063–1070, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2005\)131:9\(1063\)](https://doi.org/10.1061/(ASCE)1090-0241(2005)131:9(1063)).
49. GCTS. GCTS Testing Systems (2024), <https://www.gcts.com/about-gcts/>, <https://www.gcts.com/wp-content/uploads/2022/03/ePVC.pdf>.
50. A. Dadashiserej, A. Jana, A. W. Stuedlein, and T. M. Evans, "Effect of Strain History on the Monotonic and Cyclic Response of Natural and Reconstituted Silts," *Soil Dynamics and Earthquake Engineering* 160 (2022): 107329, <https://doi.org/10.1016/j.soildyn.2022.107329>.
51. A. Dadashiserej, A. Jana, Z. Xu, et al., "Dynamic Response of a Low Plasticity Silt Deposit: Comparison of in-Situ and Laboratory Responses," *Canadian Geotechnical Journal* 61, no. 7 (2023): 1399–1417, <https://doi.org/10.1139/cgj-2022-0579>.
52. A. Jana and A. W. Stuedlein, "Dynamic, in Situ, Nonlinear-Inelastic Response and Post-Cyclic Strength of a Plastic Silt Deposit," *Canadian Geotechnical Journal* 59, no. 1 (2022): 111–128, <https://doi.org/10.1139/cgj-2020-0652>.
53. A. Jana and A. W. Stuedlein, "Monotonic, Cyclic and Post-Cyclic Response of an Alluvial Plastic Silt Deposit," *Journal of Geotechnical and Geoenvironmental Engineering* 147, no. 12 (2021): 04020174, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0002462](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002462).
54. A. W. Stuedlein, A. Dadashiserej, A. Jana, and T. M. Evans, "Liquefaction Susceptibility and Cyclic Response of Intact Nonplastic and Plastic Silts," *Journal of Geotechnical and Geoenvironmental Engineering* 149, no. 1 (2023): 04022125, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0002935](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002935).
55. I. M. Idriss and R. W. Boulanger, *Soil Liquefaction during Earthquakes. Monograph MNO-12* (Oakland, CA: Earthquake Engineering Research Institute, 2008).
56. Y. P. Vaid and S. Sivathayalan, "Static and Cyclic Liquefaction Potential of Fraser Delta Sand in Simple Shear and Triaxial Tests," *Canadian Geotechnical Journal* 33, no. 2 (1996): 281–289, <https://doi.org/10.1139/t96-007>.
57. K. J. Ulmer, R. A. Green, and A. Rodriguez-Marek, "Recommended b-Value for Computing Number of Equivalent Stress Cycles and Magnitude Scaling Factors for Simplified Liquefaction Triggering Evaluation Procedures," *Journal of Geotechnical and Geoenvironmental Engineering* 148, no. 12 (2022): 04022113, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0002926](https://doi.org/10.1061/(ASCE)GT.1943-5606.0002926).
58. H. B. Seed and I. M. Idriss, "Simplified Procedure for Evaluating Soil Liquefaction Potential," *Journal of the Soil Mechanics and Foundations Division* 97, no. 9 (1971): 1249–1273, <https://doi.org/10.1061/jsfeaq.0001662>.
59. K. Jas, A. Jana, and G. R. Dodagoudar, "Prediction of Shear Strain and Excess Pore Water Pressure Response in Liquefiable Sands under Cyclic Loading Using Deep Learning Model," *Japanese Geotechnical Society Special Publication* 10 (2024): 1729–1734, <https://doi.org/10.3208/jgssp.v10.OS-35-05>.
60. W. S. Kwan, S. S. Sideras, S. L. Kramer, and C. El Mohtar, "Experimental Database of Cyclic Simple Shear Tests under Transient Loadings," *Earthquake Spectra* 33, no. 3 (2017): 1219–1239, <https://doi.org/10.1193/093016eqs167dp>.
61. K. Jas and G. R. Dodagoudar, "SPT- and V_S -Based Probabilistic Evaluation of Soil Liquefaction Potential Using Explainable Boosting Machine," in *Proceedings of the 18th World Conference on Earthquake Engineering*, WCEE 2024, June 30 - July 5, 2024 (Milan, Italy, 2024).
62. K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio 2014. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. *arXiv preprint arXiv:1409.1259*, <https://doi.org/10.48550/arXiv.1409.1259>.
63. C. Wang, Q. Wu, X. Liu, and L. Quintanilla, "Automated Machine Learning & Tuning with FLAML," in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, (2022), 4828–4829, <https://doi.org/10.1145/3534678.3542636>.
64. T. O'Malley, E. Bursztein, et al., "KerasTuner," (2019), <https://github.com/keras-team/keras-tuner>.
65. F. Chollet, et al. 2015. Keras. Available at, <https://keras.io>.
66. J. N. Fuhg, G. Anantha Padmanabha, N. Bouklas, et al., "A Review on Data-Driven Constitutive Laws for Solids," *Archives of Computational Methods in Engineering* (2024): 1–43, <https://doi.org/10.1007/s11831-024-10196-2>.
67. D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization (2014). *arXiv preprint arXiv:1412.6980*, <https://doi.org/10.48550/arXiv.1412.6980>.
68. T. Tieleman and G. Hinton, "Lecture 6.5-RMSProp: Divide the Gradient by a Running Average of Its Recent Magnitude," *COURSERA: Neural Networks for Machine Learning* 4 (2012): 26–31.
69. J. Duchi, E. Hazan, and Y. Singer, "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization," *Journal of Machine Learning Research* 12 (2011): 2121–2159, <http://jmlr.org/papers/v12/duchilla.html>.
70. S. L. Shen, N. Zhang, A. Zhou, and Z. Y. Yin, "Enhancement of Neural Networks with an Alternative Activation Function tanhLU," *Expert Systems with Applications* 199 (2022): 117181, <https://doi.org/10.1016/j.eswa.2022.117181>.
71. N. Zhang, S. L. Shen, A. Zhou, and Y. F. Jin, "Application of LSTM Approach for Modelling Stress-Strain Behaviour of Soil," *Applied Soft Computing* 100 (2021): 106959, <https://doi.org/10.1016/j.asoc.2020.106959>.
72. P. Zhang, Z. Y. Yin, Y. F. Jin, and B. Sheil, "Physics-Constrained Hierarchical Data-Driven Modelling Framework for Complex Path-Dependent Behaviour of Soils," *International Journal for Numerical and Analytical Methods in Geomechanics* 46, no. 10 (2022): 1831–1850, <https://doi.org/10.1002/nag.3370>.
73. N. Zhang, S. L. Shen, A. Zhou, and Y. S. Xu, "Investigation on Performance of Neural Networks Using Quadratic Relative Error Cost Function," *IEEE Access* 7 (2019): 106642–106652, <https://doi.org/10.1109/ACCESS.2019.2930520>.
74. K. Q. Li, Z. Y. Yin, N. Zhang, and Y. Liu, "A Data-Driven Method to Model Stress-Strain Behaviour of Frozen Soil Considering Uncertainty," *Cold Regions Science and Technology* 213 (2023): 103906, <https://doi.org/10.1016/j.coldregions.2023.103906>.
75. M. ElGhoriayb and M. Manzari, "LEAP-2018 - Stress-Strain Response of Ottawa F65 Sand in Cyclic Simple Shear," *LEAP-2018 - Stress-strain Response of Ottawa F65 Sand in Cyclic Simple Shear* (DesignSafe-CI, 2018), <https://doi.org/10.17603/DS2HX3H>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.