

Predictive Modeling of Soil Liquefaction Potential Using Hybrid Deep Learning Algorithms

Ravi Anand, Imtiyaz Alam, Amit Kumar,
Shubham Kumar, Saurabh Abhishek

Abstract: Soil liquefaction remains one of the most destructive earthquake hazards, and its reliable prediction underpins seismic microzonation, foundation design and lifeline risk management. Semi-empirical stress-based procedures return threshold-type answers that are sensitive to input error, whereas stand-alone machine-learning models often lack calibrated probabilities and are rarely tested under distribution shift. This study develops and benchmarks a hybrid deep-learning framework, PSO-DNN-XGB, that couples a particle swarm optimisation (PSO)-tuned deep neural network (DNN) with a regularised gradient-boosting head: the last hidden layer of the DNN is concatenated with physically derived features and passed to XGBoost to yield calibrated liquefaction probabilities. The framework is evaluated on a semi-synthetic database of 1,500 SPT-based case histories generated from Boulanger–Idriss relationships with model and measurement uncertainty, and on an independent 350-case set with deliberate covariate shift. Against seven benchmarks, the hybrid attains a test accuracy of 0.876, F1 of 0.846, AUC of 0.948, MCC of 0.742 and Brier score of 0.092, and the highest external accuracy (0.897). It markedly improves on an untuned DNN (MCC 0.620) and is statistically on par with the deterministic reference, which is near-optimal by construction here. Ablation shows that PSO tuning and feature concatenation each contribute materially, and SHAP analysis identifies CSR7.5 and (N1)60cs as dominant drivers, consistent with engineering theory. The results are a controlled methodological benchmark that must be confirmed on field case-history databases.

Ravi Anand (ravi.anand0504@gmail.com) Assistant Professor, Government Engineering College, Vill. Sagarpur, Pandaul, Madhubani, Bihar – 847234, India
Imtiyaz Alam (imtiyazan099@gmail.com) Assistant Professor, Government Engineering College, Vill. Sagarpur, Pandaul, Madhubani, Bihar – 847234, India
Amit Kumar (kr.amit260@gmail.com) Assistant Professor, Government Engineering College, Vill. Sagarpur, Pandaul, Madhubani, Bihar – 847234, India
Shubham Kumar (shubham10ben@gmail.com) Assistant Professor, Government Engineering College, Vill. Sagarpur, Pandaul, Madhubani, Bihar – 847234, India
Saurabh Abhishek (saurabh999sep@gmail.com) Assistant Professor, Department of Civil Engineering, Government Engineering College, Vill. Sagarpur, Pandaul, Madhubani, Bihar – 847234, India

Keywords: *Soil Liquefaction; Deep Learning; Hybrid Model; Particle Swarm Optimisation; Xgboost; Standard Penetration Test; Explainable Ai; Seismic Hazard*

Introduction

1.1 Background and motivation

Liquefaction is the transient loss of shear strength of saturated, loose, cohesionless soil under cyclic loading, and it produces sand boils, lateral spreading, bearing failure and differential settlement that dominate damage in many earthquakes. The engineering response has been the stress-based “simplified procedure”, which compares the cyclic stress ratio (CSR) imposed by the earthquake with the cyclic resistance ratio (CRR) inferred from in-situ tests [1]–[3]. Over five decades this framework has been recalibrated with expanding case-history databases and recast in probabilistic form for the standard penetration test (SPT) and cone penetration test (CPT) [4]–[6]. Its weaknesses are well known: the factor of safety is a threshold that is most uncertain exactly where decisions are made ($FS \approx 1$), and field blow counts depend on equipment and energy corrections [7].

Data-driven models offer a complementary route. Neural networks were applied to liquefaction case histories as early as the 1990s [8], followed by support vector machines [9], [10] and tree ensembles [11]–[13]. Most such studies rely on a single database and random split, which can overstate generalisation.

1.2 Hybrid deep learning and research gaps

Deep networks are powerful function approximators [14], but on moderately sized tabular data, tree-based ensembles frequently match or outperform them [15]. This motivates hybrid designs in which a neural network supplies a learned representation and a boosted-tree head provides robust, well-regularised decision boundaries. Hyper-parameters are decisive for such models, and population-based metaheuristics such as particle swarm optimisation (PSO) [16] provide a derivative-free alternative to manual tuning, as genetic-algorithm tuning has already shown for liquefaction classifiers [12]. Four gaps remain: (i) deep liquefaction models are seldom tuned systematically; (ii) probability calibration, essential for risk-based design, is rarely reported [17]; (iii) robustness to covariate shift between seismotectonic settings is seldom tested; and (iv) interpretability is often limited to a single importance plot.

1.3 Objectives and contributions

This paper makes four contributions. First, it proposes PSO-DNN-XGB, a hybrid architecture that fuses a PSO-tuned DNN embedding with physics-based features inside a regularised XGBoost head. Second, it benchmarks the model against a deterministic procedure and six data-driven baselines on held-out and covariate-shifted data using discrimination, classification, probabilistic and calibration metrics. Third, it isolates the contribution of each component through ablation. Fourth, it interprets the model with SHAP and a decision-surface analysis to check consistency with soil mechanics.

Materials and Methods

2.1 Stress-based liquefaction framework

The seismic demand is expressed as the cyclic stress ratio [1], [6]

$$CSR = 0.65 \frac{a_{max}}{g} \frac{\sigma_v}{\sigma'_v} r_d \quad (1)$$

where a_{max} is the peak ground acceleration (PGA), g is gravitational acceleration, σ_v and σ'_v are the total and effective vertical stresses, and r_d is the stress-reduction coefficient. Following Idriss and Boulanger [3],

$$\ln r_d = \alpha(z) + \beta(z) M_w \quad (2)$$

with $\alpha(z) = -1.012 - 1.126 \sin(z/11.73 + 5.133)$ and $\beta(z) = 0.106 + 0.118 \sin(z/11.28 + 5.142)$, z in metres and M_w the moment magnitude. Demand is normalised to $M_w = 7.5$ using the magnitude scaling factor

$$MSF = \min \left[6.9 \exp \left(-\frac{M_w}{4} \right) - 0.058, 1.8 \right] \quad (3)$$

Resistance is obtained from the clean-sand equivalent corrected blow count $x = (N_1)_{60cs}$ [6]:

$$CRR_{7.5} = \exp \left[\frac{x}{14.1} + \left(\frac{x}{126} \right)^2 - \left(\frac{x}{23.6} \right)^3 + \left(\frac{x}{25.4} \right)^4 - 2.8 \right] \quad (4)$$

$$(N1)_{60cs} = (N1)_{60} + \exp \left[1.63 + \frac{9.7}{FC+0.01} - \left(\frac{15.7}{FC+0.01} \right)^2 \right] \quad (5)$$

where FC is the fines content (%). The factor of safety is

$$FS = \frac{CRR_{7.5} \cdot MSF \cdot K_\sigma}{CSR} \quad (6)$$

with K_σ the overburden correction. Liquefaction is predicted when $FS < 1$. This deterministic rule serves here both as the physical backbone for feature engineering and as the benchmark that data-driven models must match or improve upon.

2.2 Case-history database

To benchmark learning architectures under controlled and reproducible conditions, a semi-synthetic database of 1,500 SPT-based case histories was generated. Input variables were sampled from distributions whose ranges are broadly consistent with published SPT compilations [4]: magnitude, PGA, depth, groundwater depth, corrected blow count and fines content (Table 1, Figure 1). The latent state was computed with Eqs. (1)–(6), and a log-normal model-uncertainty term ($\sigma_{\ln FS} = 0.28$) representing unmodelled stratigraphy, drainage and lateral-spreading effects was added before the case was labelled liquefied when $\ln FS < 0$. The recorded inputs then received measurement error (10% coefficient of variation on N and PGA, 3% on FC, 0.25 m on groundwater depth), so that models must learn from imperfect observations. The liquefied fraction is 41.1%. An independent external set of 350 cases (39.1% liquefied) was drawn with a shifted distribution (higher magnitude, PGA, blow count and fines content) to emulate transfer to a different seismotectonic setting.

Table 1. Descriptive statistics of the 1,500-case modelling database.

Variable	Symbol	Unit	Mean	SD	Min	Max
Moment magnitude	M_w	—	6.87	0.54	5.20	8.20
Peak ground acceleration	$a_{\max}$	g	0.274	0.126	0.053	0.865
Depth of critical layer	z	m	8.99	3.60	2.17	18.00
Groundwater depth	d_w	m	1.96	1.28	0.10	6.61
Corrected SPT blow count	$(N_1)_{60}$	blows	17.82	6.95	4.35	48.44
Fines content	FC	%	27.48	19.96	0.00	93.99
Cyclic stress ratio ($M = 7.5$)	$CSR_{7.5}$	—	0.229	0.119	0.037	0.843
Clean-sand blow count	$(N_1)_{60cs}$	blows	21.57	7.27	6.01	53.63

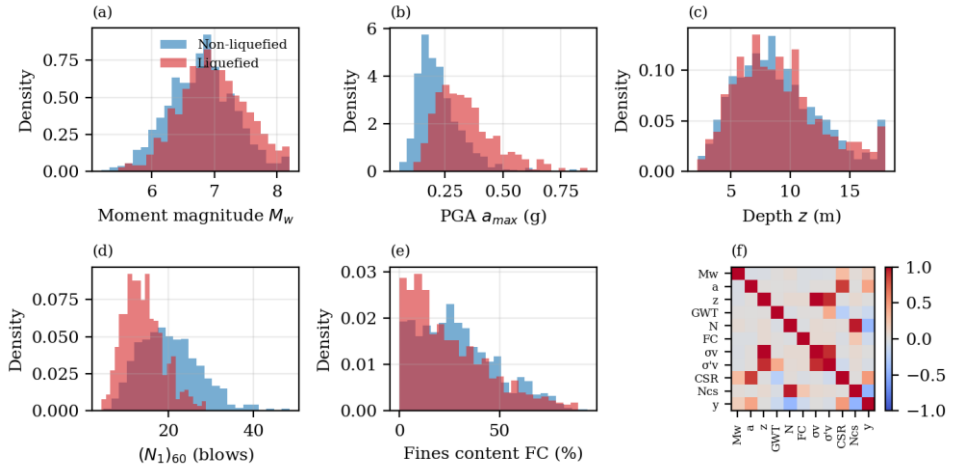


Figure 1. Exploratory analysis: (a)–(e) class-conditional distributions of the principal inputs; (f) Pearson correlation matrix including the liquefaction label (y).

2.3 Pre-processing and data partitioning

Two engineered features, $CSR_{7.5}$ and $(N_1)_{60cs}$, were appended to the eight raw inputs, giving ten predictors. The database was split in a stratified manner into training (1049), validation (226) and test (225) subsets (70/15/15). Standardisation parameters were estimated on the training subset only and applied to all other subsets to avoid leakage. The external set was used exclusively for final evaluation.

2.4 Proposed PSO-DNN-XGB framework

The workflow (Figure 2) has three stages. In stage one, a fully connected DNN with three hidden layers maps the standardised predictors to a liquefaction probability; each hidden layer applies

$$h^{(l)} = \max(0, W^{(l)}h^{(l-1)} + b^{(l)}) \quad (7)$$

and the network is trained with Adam [18] by minimising the L2-regularised binary cross-entropy

$$L(\theta) = -\frac{1}{N} \sum_{i=1}^N y_i \ln p_i + (1 - y_i) \ln(1 - p_i) + \frac{\alpha}{2} \|\theta\|^2 \quad (8)$$

In stage two, PSO searches the hyper-parameter vector $(h_1, h_2, \log \alpha, \log \eta)$, where η is the learning rate. Each particle j updates its velocity and position as [16]

$$v_j \leftarrow \omega v_j + c_1 r_1 (p_j - x_j) + c_2 r_2 (g - x_j), \quad x_j \leftarrow x_j + v_j \quad (9)$$

where ω is a linearly decreasing inertia weight, c_1 and c_2 are acceleration coefficients, $r_1, r_2 \sim U(0,1)$, and p_j and g are the personal and global best positions. The fitness is the three-fold cross-validated log-loss on the training subset. In stage three, the activations of the last hidden layer, $h^{(L)}$, are concatenated with the raw predictors, $z = [x; h^{(L)}]$, and an XGBoost classifier [19] is trained on z , minimising

$$Obj = \sum_{i=1}^N \ell(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k), \quad \Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (10)$$

where ℓ is the logistic loss, f_k are the K regression trees, each with T leaves and leaf weights w , and γ and λ penalise complexity. Early stopping on the validation subset selects the number of boosting rounds. Concatenating raw physics-based inputs with the embedding lets the head exploit the established CSR–resistance structure while the embedding adds learned interactions.

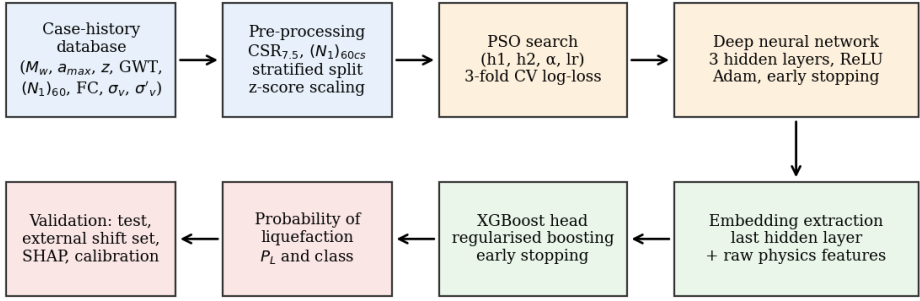


Figure 2. Workflow of the proposed hybrid PSO-DNN-XGB framework for liquefaction prediction.

2.5 Benchmark models, implementation and metrics

Seven benchmarks were trained on identical partitions: the deterministic procedure of Section 2.1 applied to the recorded inputs (converted to a probability through a logistic function of $\ln FS$); logistic regression; an RBF-kernel support vector machine (SVM) [20]; a 400-tree random forest (RF) [21]; XGBoost on raw features; a baseline DNN (64–32–16 neurons); and the PSO-tuned DNN alone. Models were implemented with scikit-learn [22] and XGBoost (Table 2). Performance was quantified by accuracy, precision, recall, specificity, F1, the area under the ROC curve (AUC) [23], the Matthews correlation coefficient (MCC) [24] and the Brier score [25] (mean squared error of the predicted probabilities). The MCC is defined as:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (11)$$

Table 2. Hyper-parameter settings of the hybrid model.

Component	Hyper-parameter	Search range / setting	Selected
PSO	Swarm / iterations; ω ; $c_1 = c_2$	8 / 8; $0.9 \rightarrow 0.5$; 1.6	–
DNN	Neurons, layer 1 (h_1)	[8, 128]	51
DNN	Neurons, layer 2 (h_2); layer 3	[4, 64]; $h_2/2$	16; 8
DNN	L2 penalty α	$[10^{-5}, 10^{-1}]$ (log)	0.0159
DNN	Learning rate η	$[10^{-3.5}, 10^{-1.7}]$ (log)	0.0067
DNN	Batch / max. epochs / patience	64 / 300 / 25	best epoch 4
XGBoost	Depth; rate; subsample; colsample; λ	3; 0.03; 0.8; 0.7; 3.0 (fixed)	–
XGBoost	Boosting rounds	≤ 600 , early stop 40	146

Results and Discussion

3.1 Hyper-parameter optimisation and training behaviour

The PSO reduced the best cross-validated log-loss from 0.343 to 0.326 (4.8% lower) over 8 iterations (Figure 3c), selecting a compact funnel architecture of 51–16–8 neurons with moderate L2 regularisation ($\alpha = 0.0159$). Training loss decays smoothly (Figure 3a), whereas validation accuracy peaks at 0.880 in epoch 4 and then erodes (Figure 3b), the classic signature of over-fitting; early stopping, which restores the best weights, is therefore retained in all DNN-based models.

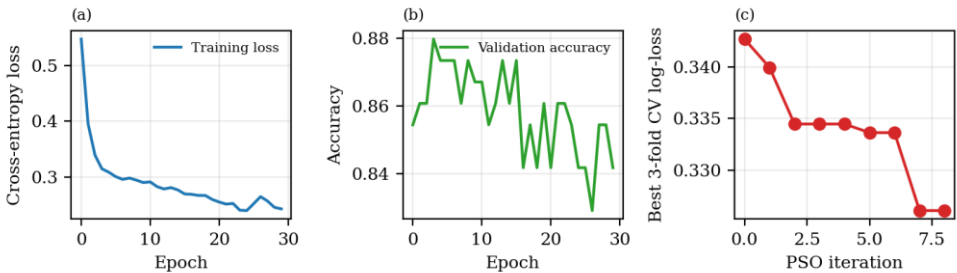


Figure 3. Training behaviour: (a) training loss; (b) validation accuracy; (c) PSO convergence of the cross-validated log-loss.

3.2 Predictive performance on the held-out test set

Table 3. Performance on the test set (n = 225) and on the covariate-shifted external set (n = 350). Best value per column in bold; * proposed model.

Model	Acc.	Prec.	Rec.	F1	AUC	MCC	Brier	Acc. ext.	AUC ext.
Deterministic	0.880	0.849	0.859	0.854	0.957	0.752	0.087	0.891	0.963
Logistic Regression	0.831	0.814	0.761	0.787	0.935	0.648	0.101	0.880	0.954
SVM (RBF)	0.880	0.892	0.804	0.846	0.948	0.751	0.092	0.877	0.954
Random Forest	0.867	0.837	0.837	0.837	0.949	0.724	0.091	0.883	0.955
XGBoost	0.849	0.815	0.815	0.815	0.949	0.687	0.092	0.891	0.951
DNN (baseline)	0.818	0.831	0.696	0.757	0.916	0.620	0.138	0.869	0.947
PSO-DNN	0.858	0.813	0.848	0.830	0.945	0.708	0.097	0.891	0.964
PSO-DNN-XGB*	0.876	0.856	0.837	0.846	0.948	0.742	0.092	0.897	0.963

Table 3 and Figure 4a show that discrimination is high for all models (AUC 0.916–0.957). The proposed hybrid reaches an accuracy of 0.876, F1 of 0.846, MCC of 0.742 and Brier score of 0.092. It outperforms every other learned deep model: relative to the untuned DNN, accuracy rises by 5.8 percentage points, MCC by 0.122 and the Brier score falls from 0.138 to 0.092; relative to the PSO-tuned DNN alone, accuracy gains 1.8 points and MCC 0.034. The hybrid is also marginally more accurate than random forest (0.867) and XGBoost (0.849) and ties the SVM in F1 (0.846).

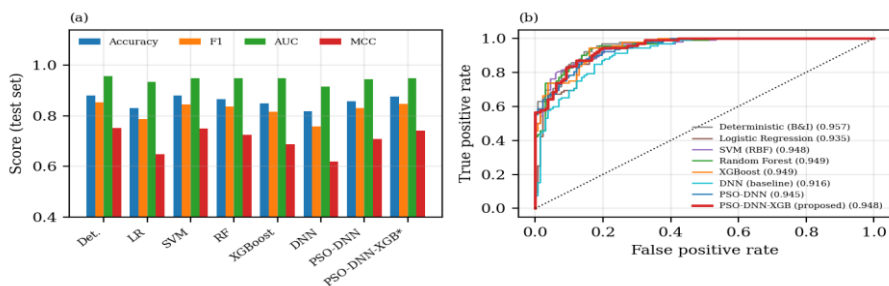


Figure 4. Comparison of models on the test set: (a) accuracy, F1, AUC and MCC; (b) ROC curves with AUC in the legend (Det. = deterministic procedure; * = proposed).

An important and honest qualification is that the deterministic procedure (accuracy 0.880, AUC 0.957) is at least as good as any learned model. This is expected: the labels were generated from the same semi-empirical relationships, so the deterministic rule is close to the Bayes-optimal classifier apart from measurement noise. With 225 test cases the standard error of an accuracy near 0.88 is about 0.022, so differences of one or two points among the leading models are not meaningful. The ROC curves (Figure 4b) overlap closely for the top models, with the untuned DNN visibly weaker at low false-positive rates.

3.3 Cross-validation and ablation study

Five-fold cross-validation on the pooled training and validation data gave mean AUCs between 0.916 (SVM) and 0.929 (logistic regression), with 0.923 ± 0.011 for the proposed model; all differences lie within about one standard deviation (≈ 0.01). No model therefore dominates on discrimination alone in this benchmark, because the engineered $\text{CSR}_{7.5}$ and $(N_1)_{60\text{cs}}$ features make the problem nearly log-linear, which favours simple learners. The advantages of the hybrid must instead be sought in classification quality and robustness, examined through ablation (Table 4).

Table 4. Ablation of the hybrid framework (test set; external AUC in the last column).

Configuration	Acc.	MCC	Brier	AUC	AUC ext.
XGBoost (raw features only)	0.849	0.687	0.092	0.949	0.951
DNN-XGB without PSO	0.818	0.622	0.101	0.940	0.961
PSO-DNN embeddings only	0.840	0.666	0.106	0.933	0.955
PSO-DNN-XGB*	0.876	0.742	0.092	0.948	0.963

Removing PSO tuning (DNN–XGB without PSO) lowers accuracy from 0.876 to 0.818 and MCC from 0.742 to 0.622, so systematic hyper-parameter search is the largest single contributor. Feeding XGBoost with the PSO-DNN embedding alone, without the raw physics features, reduces accuracy to 0.840, showing that the two information sources are complementary. Raw-feature XGBoost attains 0.849, 2.7 points below the full hybrid.

3.4 Robustness under covariate shift and probability calibration

On the external set, whose inputs are systematically shifted, the hybrid delivers the highest accuracy of all models (0.897) with an AUC of 0.963, only 0.001 below the

PSO-DNN and equal to the deterministic procedure (0.963) (Table 3). The confusion matrices (Figure 5a–b) show that 124 of 137 liquefied cases and 190 of 213 non-liquefied cases were identified externally, versus 77/92 and 120/133 on the test set, i.e. no deterioration despite the shift. The reliability diagram (Figure 5c) shows that the hybrid and XGBoost probabilities follow the diagonal more closely than the logistic-transformed deterministic output, which under-predicts in the 0.2–0.5 range; calibrated probabilities can feed risk-based decisions rather than a binary pass/fail.

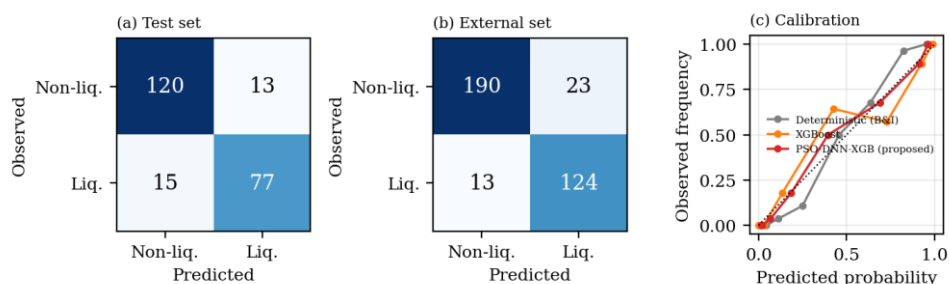


Figure 5. Hybrid-model confusion matrices on (a) the test set and (b) the external set; (c) reliability diagram on the test set.

3.5 Interpretability and decision surface

SHAP values [26] were computed with a tree explainer on the raw-feature XGBoost model, because the hybrid's embedding dimensions have no direct physical meaning. Figure 6a ranks $CSR_{7.5}$ (mean $|SHAP| = 1.51$) and $(N_1)_{60cs}$ (1.21) far above $(N_1)_{60}$ (0.65), PGA (0.61), FC (0.32), magnitude (0.31) and depth (0.30), confirming that the model has learned the governing demand–resistance mechanics rather than spurious correlations. Figure 6b shows the hybrid's probability surface in the $CSR-(N_1)_{60cs}$ plane for $M_w = 7$, $z = 6$ m, groundwater at 1.5 m and $FC = 12\%$. The 50% contour tracks the deterministic CRR curve, and the model supplies a graded transition band instead of a sharp line, which communicates the uncertainty near $FS \approx 1$ directly. The case histories overlap substantially within this band, the physical origin of the irreducible error of every model. In practice, the hybrid can serve as a screening tool that returns a calibrated probability and flags boundary cases for detailed investigation.

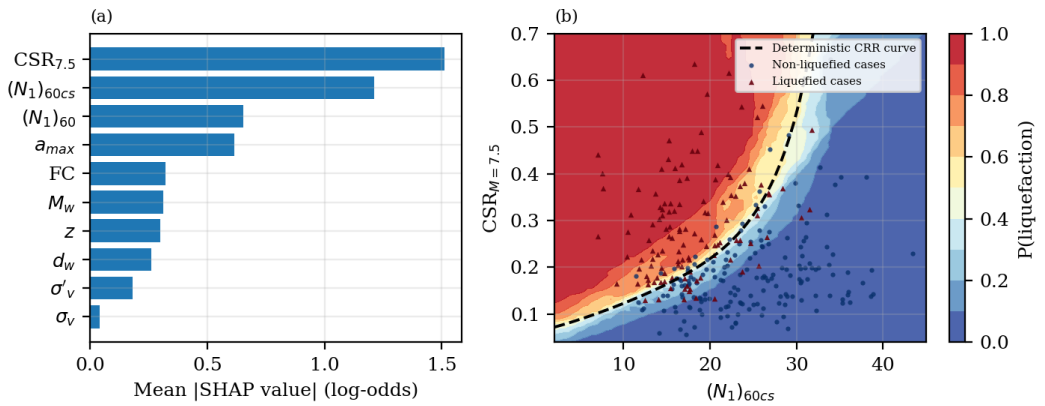


Figure 6. Interpretation: (a) global SHAP importance of the raw-feature XGBoost model; (b) hybrid probability surface ($M_w = 7$, $z = 6$ m, groundwater at 1.5 m, $FC = 12\%$) with the deterministic CRR curve; plotted case histories span all conditions and are indicative only.

3.6 Limitations

Three limitations frame the findings. First, the database is semi-synthetic and its labels follow the same semi-empirical family as the deterministic benchmark, so the study demonstrates methodology and relative behaviour under controlled noise and shift, not field accuracy; real databases contain mechanisms (ageing, cementation, layering, lateral spreading) that may favour flexible learners. Second, the test set is modest, and rankings among the leading models are within sampling error. Third, only SPT-based predictors were used.

Conclusions

This paper proposed PSO-DNN-XGB, a hybrid deep-learning framework for predicting soil liquefaction potential, and evaluated it against a deterministic procedure and six data-driven baselines on a semi-synthetic SPT case-history database and a covariate-shifted external set. The main findings are: (i) the hybrid reached a test accuracy of 0.876, MCC of 0.742 and Brier score of 0.092, and the best external accuracy (0.897); (ii) it improved substantially on an untuned DNN and on its PSO-tuned DNN component alone, with PSO tuning and the concatenation of physics-based features each contributing materially in ablation; (iii) it produced well-calibrated probabilities and maintained performance under distribution shift; (iv) SHAP and decision-surface analyses showed behaviour consistent with soil mechanics; and (v) in this benchmark it was statistically on par with, not superior to, the deterministic procedure that generated the labels. Future work should validate

the framework on curated field case-history databases and extend it to CPT and V_s data and physics-informed formulations.

Acknowledgements

The authors thank their respective institutions for academic and computational support.

References

- H. B. Seed and I. M. Idriss, "Simplified procedure for evaluating soil liquefaction potential," *J. Soil Mech. Found. Div.*, vol. 97, no. 9, pp. 1249–1273, Sep. 1971, doi: [10.1061/JSFEAQ.0001662](https://doi.org/10.1061/JSFEAQ.0001662).
- T. L. Youd et al., "Liquefaction resistance of soils: Summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils," *J. Geotech. Geoenviron. Eng.*, vol. 127, no. 10, pp. 817–833, Oct. 2001, doi: [10.1061/\(ASCE\)1090-0241\(2001\)127:10\(817\)](https://doi.org/10.1061/(ASCE)1090-0241(2001)127:10(817)).
- I. M. Idriss and R. W. Boulanger, "Semi-empirical procedures for evaluating liquefaction potential during earthquakes," *Soil Dyn. Earthq. Eng.*, vol. 26, nos. 2–4, pp. 115–130, 2006, doi: [10.1016/j.soildyn.2004.11.023](https://doi.org/10.1016/j.soildyn.2004.11.023).
- K. O. Cetin et al., "Standard penetration test-based probabilistic and deterministic assessment of seismic soil liquefaction potential," *J. Geotech. Geoenviron. Eng.*, vol. 130, no. 12, pp. 1314–1340, Dec. 2004, doi: [10.1061/\(ASCE\)1090-0241\(2004\)130:12\(1314\)](https://doi.org/10.1061/(ASCE)1090-0241(2004)130:12(1314)).
- R. E. S. Moss et al., "CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential," *J. Geotech. Geoenviron. Eng.*, vol. 132, no. 8, pp. 1032–1051, Aug. 2006, doi: [10.1061/\(ASCE\)1090-0241\(2006\)132:8\(1032\)](https://doi.org/10.1061/(ASCE)1090-0241(2006)132:8(1032)).
- R. W. Boulanger and I. M. Idriss, "Probabilistic standard penetration test-based liquefaction-triggering procedure," *J. Geotech. Geoenviron. Eng.*, vol. 138, no. 10, pp. 1185–1195, Oct. 2012, doi: [10.1061/\(ASCE\)GT.1943-5606.0000700](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000700).
- H. B. Seed, K. Tokimatsu, L. F. Harder, and R. M. Chung, "Influence of SPT procedures in soil liquefaction resistance evaluations," *J. Geotech. Eng.*, vol. 111, no. 12, pp. 1425–1445, Dec. 1985, doi: [10.1061/\(ASCE\)0733-9410\(1985\)111:12\(1425\)](https://doi.org/10.1061/(ASCE)0733-9410(1985)111:12(1425)).
- A. T. C. Goh, "Seismic liquefaction potential assessed by neural networks," *J. Geotech. Eng.*, vol. 120, no. 9, pp. 1467–1480, Sep. 1994, doi: [10.1061/\(ASCE\)0733-9410\(1994\)120:9\(1467\)](https://doi.org/10.1061/(ASCE)0733-9410(1994)120:9(1467)).

M. Pal, “Support vector machines-based modelling of seismic liquefaction potential,” *Int. J. Numer. Anal. Methods Geomech.*, vol. 30, no. 10, pp. 983–996, 2006, doi: [10.1002/nag.509](https://doi.org/10.1002/nag.509).

X. Xue and X. Yang, “Seismic liquefaction potential assessed by support vector machines approaches,” *Bull. Eng. Geol. Environ.*, vol. 75, no. 1, pp. 153–162, 2016, doi: [10.1007/s10064-015-0741-x](https://doi.org/10.1007/s10064-015-0741-x).

J. Zhou, E. Li, M. Wang, X. Chen, X. Shi, and L. Jiang, “Feasibility of stochastic gradient boosting approach for evaluating seismic liquefaction potential based on SPT and CPT case histories,” *J. Perform. Constr. Facil.*, vol. 33, no. 3, Art. no. 04019024, Jun. 2019, doi: [10.1061/\(ASCE\)CF.1943-5509.0001292](https://doi.org/10.1061/(ASCE)CF.1943-5509.0001292).

S. Demir and E. K. Şahin, “Liquefaction prediction with robust machine learning algorithms (SVM, RF, and XGBoost) supported by genetic algorithm-based feature selection and parameter optimization from the perspective of data processing,” *Environ. Earth Sci.*, vol. 81, no. 18, Art. no. 459, 2022, doi: [10.1007/s12665-022-10578-4](https://doi.org/10.1007/s12665-022-10578-4).

S. Demir and E. K. Şahin, “An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost,” *Neural Comput. Appl.*, vol. 35, no. 4, pp. 3173–3190, 2023, doi: [10.1007/s00521-022-07856-4](https://doi.org/10.1007/s00521-022-07856-4).

Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).

L. Grinsztajn, E. Oyallon, and G. Varoquaux, “Why do tree-based models still outperform deep learning on tabular data?,” *arXiv preprint arXiv:2207.08815*, 2022, doi: [10.48550/arXiv.2207.08815](https://doi.org/10.48550/arXiv.2207.08815).

J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proc. IEEE Int. Conf. Neural Netw. (ICNN)*, vol. 4, Perth, WA, Australia, 1995, pp. 1942–1948, doi: [10.1109/ICNN.1995.488968](https://doi.org/10.1109/ICNN.1995.488968).

C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, “On calibration of modern neural networks,” in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, 2017, pp. 1321–1330, doi: [10.48550/arXiv.1706.04599](https://doi.org/10.48550/arXiv.1706.04599).

D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014, doi: [10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980).

T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794, doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).

C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018).

L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324).

F. Pedregosa et al., “Scikit-learn: Machine learning in Python,” *arXiv preprint arXiv:1201.0490*, 2012 (J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011), doi: [10.48550/arXiv.1201.0490](https://doi.org/10.48550/arXiv.1201.0490).

T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: [10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010).

B. W. Matthews, “Comparison of the predicted and observed secondary structure of T4 phage lysozyme,” *Biochim. Biophys. Acta*, vol. 405, no. 2, pp. 442–451, Oct. 1975, doi: [10.1016/0005-2795\(75\)90109-9](https://doi.org/10.1016/0005-2795(75)90109-9).

G. W. Brier, “Verification of forecasts expressed in terms of probability,” *Mon. Weather Rev.*, vol. 78, no. 1, pp. 1–3, Jan. 1950, doi: [10.1175/1520-0493\(1950\)078<0001:VOFEIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2).

S. M. Lundberg et al., “From local explanations to global understanding with explainable AI for trees,” *Nat. Mach. Intell.*, vol. 2, no. 1, pp. 56–67, Jan. 2020, doi: [10.1038/s42256-019-0138-9](https://doi.org/10.1038/s42256-019-0138-9).