Bridging Geomechanics and Machine Learning with Physics-Informed Neural Surrogates for Triaxial Soil Testing

Minhaz Chowdhury, Azmine Toushik Wasi, Md Manjurul Ahsan
Department of {Petroleum and Mining Engineering, Industrial and Production Engineering},
Shahjalal University of Science and Technology
Department of Industrial and Systems Engineering, University of Oklahoma
minhaz40@student.sust.edu, azminetoushik.wasi@gmail.com, ahsan@ou.edu

Abstract

Modeling the time-dependent responses of geotechnical materials under triaxial loading poses a dual challenge: capturing strongly nonlinear constitutive behavior while mitigating the influence of experimental noise. We present a two-model learning benchmark that jointly predicts *Displacement*, *Load*, and *Deviator Strain* from elapsed time, comparing a transparent LinearRegressor baseline with a *Physics-Informed Neural Network* (PINN). The PINN encodes two physically grounded priors, (i) monotonic displacement progression and (ii) non-negative incremental work as differentiable penalty terms embedded directly in the training objective. This design ensures physically admissible trajectories without constraining the network's capacity to model nonlinear temporal patterns. The pipeline incorporates precise preprocessing time normalization, feature alignment, z-score standardization and a fixed train–test split for reproducible benchmarking. Across all target channels, the PINN achieves substantial gains in mean absolute error and R^2, with *Deviator Strain* showing the largest improvement due to its inherently nonlinear dynamics. All evaluations are reported in denormalized physical units to preserve engineering interpretability. Results confirm that integrating minimal, interpretable physics priors into neural predictors significantly improves fidelity in time-series modeling of laboratory geomechanics, offering a scalable, domain-adaptable framework for triaxial testing and related applications.

1 Introduction

Predicting the coupled responses of geomaterials under controlled loading remains a central challenge in rock and soil mechanics. This difficulty arises from the nonlinear nature of constitutive laws [28], intrinsic material heterogeneity [25, 35, 39], and pervasive measurement noise [32, 17, 31]. Data-driven models offer flexibility in capturing complex dependencies [37], yet they often fail to generalize beyond the training regime [6] and may produce outputs that violate known physical principles [21]. Physics-informed learning provides a principled alternative by embedding domain knowledge into machine learning architectures [32], thereby enhancing both extrapolation and interpretability [9].

In this work, we investigate a **minimal-feature, physics-informed framework** for predicting *Displacement*, *Load*, and *Deviator Strain* in quasi-static triaxial tests using elapsed time (sec) as the sole input variable [3, 23]. Two predictive paradigms are considered. The first is a standard linear regression baseline, which captures dominant monotonic trends but is limited in representing nonlinearities [19]. The second is a compact *Physics-Informed Neural Network* (PINN) that augments a feed-forward architecture with two soft physics constraints: (i) displacement must be non-decreasing, and (ii) incremental work must remain non-negative [8, 13, 27]. These constraints are encoded as differentiable penalty terms within the loss function and jointly optimized with MSE [24]. Our central hypothesis is that embedding such constraints suppresses non-physical predictions while allowing the model to capture essential nonlinear behaviors [16], thus improving predictive accuracy without sacrificing transparency [34]. To ensure reproducibility and fair comparison, both models are trained under an identical pipeline involving time conversion [36], channel selection [1], z-score normalization [30], and fixed train–test splits [4]. This study bridges machine learning and geomechanics by showing that even minimal yet physically meaningful priors can enable high-fidelity, interpretable predictions of time-dependent material responses.

2 Methodology

2.1 Data and Preprocessing

The dataset, originating from open-source controlled triaxial compression tests on Kaggle, includes timestamped sequences of **Displacement (mm)**, **Load (N)**, and **Deviator Strain**, characterizing soil stress–strain behavior.

Preprocessing involved transforming timestamps to elapsed seconds, excluding incomplete records, and standardizing target channels using the training subset's mean and variance. To prevent data leakage, test set predictions were denormalized using these same stored training statistics.

A fixed random seed generated a single 80/20 train-test partition to mitigate stochastic effects and ensure identical, non-shared data splits for training and evaluation. This identical preprocessed dataset was supplied to all competing models, ensuring a fair and reproducible comparison. [14, 18, 7]

2.2 Model Architectures

Two modeling paradigms were examined to evaluate predictive performance across contrasting levels of model complexity.

(i) Baseline: Linear Regression. As a reference point, we employed a standard linear regression model, which maps the scalar input variable (time t) to a three-dimensional output vector representing the target channels [12]. This baseline serves primarily to capture monotonic temporal trends, thereby highlighting the limitations of purely linear approaches in representing the nonlinear stress–strain responses typical of geomechanical systems [4].

(ii) Physics-Informed Neural Network (PINN). To capture richer dynamics, a compact physics-informed neural network (PINN) was implemented. The network comprises a fully connected feed-forward architecture that accepts the scalar input t and produces the three target responses. Nonlinear activation functions were introduced in the hidden layers to accommodate temporal curvature and more complex dynamics [29]. The depth and width of the architecture were deliberately constrained to remain within a shallow-to-moderate range, reflecting the balance between expressive capacity and the relatively limited size of experimental datasets in geomechanics [32, 38].

A key feature of the PINN lies in the incorporation of domain knowledge through physics-based regularization. Specifically, two constraint terms derived from geomechanical principles were embedded into the loss function as differentiable penalties. This design enables the network to simultaneously optimize predictive accuracy and adherence to physical plausibility, thereby enhancing generalizability while preserving interpretability [20].

2.3 Physics constraints

The physics-informed component augments the standard mean-squared error (MSE) objective with two penalty terms [5]:

1. **Monotonic Displacement Constraint.** Displacement $\delta(t)$ is physically required to be non-decreasing over time under quasi-static loading [10, 28]. This is enforced via:

$$\mathcal{L}_{\text{mono}} = \frac{1}{N-1} \sum_{i=1}^{N-1} \mathrm{ReLU}\big(- \Delta\delta_i \big) \tag{1}$$

where $\Delta\delta_i = \delta_{i+1} - \delta_i$ is the discrete increment between successive predictions [2].

2. **Non-Negative Incremental Work Constraint.** Incremental mechanical work, approximated via trapezoidal integration of predicted load over displacement [33], must remain non-negative [26]:

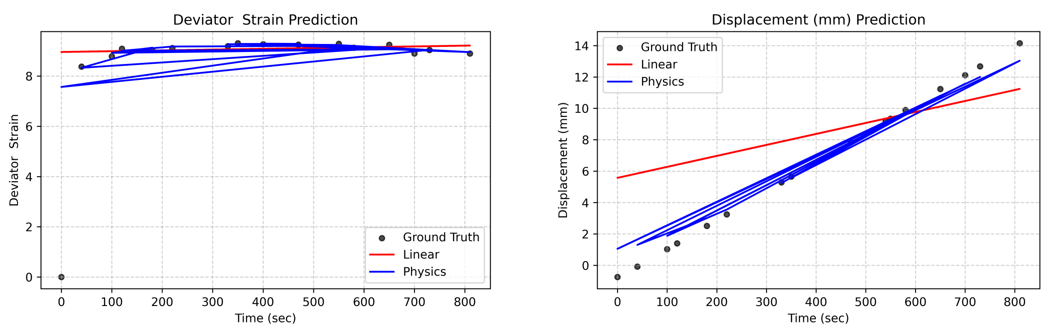
$$w_i \approx \frac{F_i + F_{i+1}}{2} \cdot \Delta\delta_i ; \mathcal{L}_{\text{work}} = \frac{1}{N-1} \sum_{i=1}^{N-1} \mathrm{ReLU}(-w_i) \tag{2}$$

The total PINN loss is: $\mathcal{L}_{\text{PINN}} = \mathcal{L}_{\text{MSE}} + \lambda_1 \mathcal{L}_{\text{mono}} + \lambda_2 \mathcal{L}_{\text{work}}$, where λ_1 and λ_2 are tunable penalty weights, selected to balance fidelity to the observed data against adherence to physical constraints. Both penalties are fully differentiable, preserving end-to-end trainability with standard backpropagation [15].

2.4 Training, Evaluation and Visualization

The baseline linear regressor was trained using the Adam [22] optimizer (lr $= 10^{-2}$) for 3000 epochs. The PINN was trained using Adam (lr $= 10^{-3}$) for up to 5000 epochs, with early stopping based on validation loss to mitigate overfitting and accommodate fluctuations in the composite loss landscape. Training was monitored via periodic logging of loss components and metrics. PINN loss convergence typically exhibited two phases: initial rapid MSE reduction, followed by gradual refinement driven by the physics penalties, leading to a stable low-loss regime.

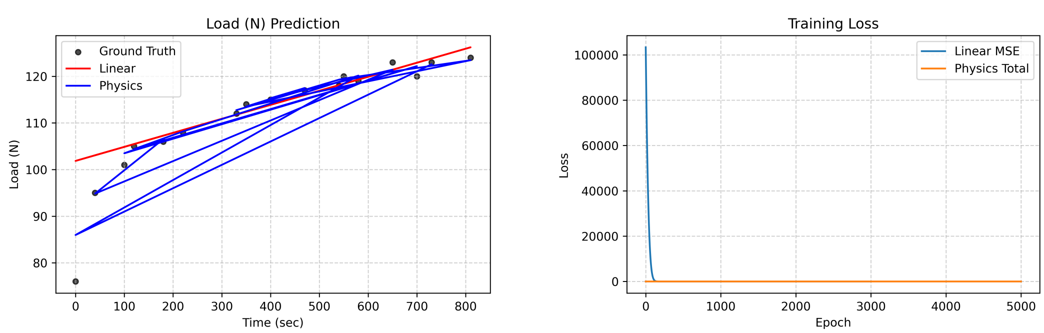
Figures 1a–2a present model predictions versus ground truth for all three target channels. The PINN consistently reproduces local curvature and fluctuation patterns absent in the linear baseline [11]. Figure 2b contrasts training loss trajectories, highlighting the stabilizing effect of the physics constraints. Aggregate performance metrics are summarized in Table 1 and visualized in Figure 3, showing consistent gains in both MAE and R^2 across channels, with the largest improvements observed for *Deviator Strain*.



(a) Prediction: Deviator Strain

(b) Prediction: Displacement (mm)

Figure 1: Prediction results for (a) Deviator Strain and (b) Displacement (mm).



(a) Prediction: Load

(b) Training Loss

Figure 2: (a) Prediction results for Load and (b) Training Loss for Linear vs Physic Informed Model.

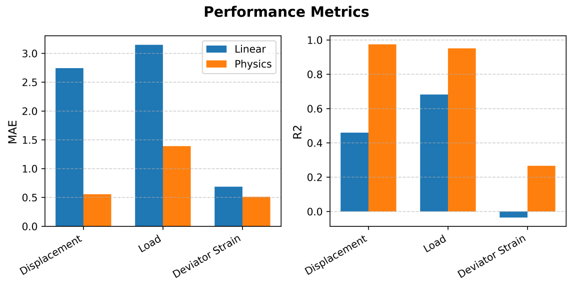


Figure 3: Performance

Table 1: Regression test metrics

Model	MAE	R^2	Max Error
Linear Displacement	2.74	0.46	6.32
Phys-NN Displacement	**0.56**	**0.97**	**1.80**
Linear Load	3.15	0.68	25.86
Phys-NN Load	**1.39**	**0.95**	**9.97**
Linear Deviator Strain	0.69	-0.04	8.96
Phys-NN Deviator Strain	**0.51**	**0.27**	**7.57**

3 Results and Discussion

The comparative evaluation demonstrates a consistent advantage of the physics-informed neural network (PINN) over the baseline linear regressor across all response channels. The linear model converges rapidly during early epochs, reflecting its ability to capture a dominant temporal trend. However, its predictions lack fidelity to the nuanced curvature and localized fluctuations present in the experimental data. In contrast, the PINN maintains a lower composite loss throughout training and achieves closer alignment with ground truth by leveraging dual physics priors that enforce displacement monotonicity and non-negative incremental work. These constraints not only suppress non-physical artifacts but also enable the network to represent nonlinear temporal behavior more effectively.

Channel-specific analyses further highlight the improvements achieved by the PINN. For *Deviator Strain*, the ground-truth trajectory exhibits nonlinear fluctuations that the linear model largely smooths out, whereas the PINN reproduces these variations with higher fidelity. For *Displacement*, the PINN tracks the ground-truth envelope more closely, particularly in regions of curvature and inflection, where the linear model tends to underestimate the response. Similarly, for *Load*, the PINN demonstrates superior ability to capture localized fluctuations absent in the baseline predictions.

Quantitative results corroborate these observations. Table 1 summarizes mean absolute error (MAE) and coefficient of determination (R^2) for all channels. The PINN achieves substantial reductions in MAE and consistent increases in R^2, with the largest relative improvements observed for *Deviator Strain* and *Displacement*. These results validate the central hypothesis that incorporating interpretable physics priors into a predictor enhances predictive accuracy while ensuring physical plausibility.

The visual evidence is consistent with the numerical findings. Figures 1a–2a illustrate the improved alignment of PINN predictions with experimental trajectories across all channels. Training dynamics (Figure 2b) demonstrate the stability of the physics-informed optimization process, while the comparative performance summary (Figure 3) highlights the systematic gains achieved over the baseline. Together, these results underscore the value of embedding minimal yet meaningful physics constraints in data-driven models for time-dependent geotechnical responses.

4 Concluding Remarks

In this study, we demonstrate that combining a physics-informed neural network (PINN) with a linear baseline provides a robust framework for predicting time-dependent geotechnical responses from triaxial test data. Three contributions stand out. First, a transparent data-processing workflow ensures reproducibility through explicit normalization and denormalization. Second, a two-penalty physics-informed objective enforces monotonic displacement and non-negative incremental work, embedding physical plausibility into the learning process. Third, a comprehensive evaluation across *Displacement*, *Load*, and *Deviator Strain* shows that the PINN consistently outperforms the baseline in MAE and R^2, with the greatest gains for *Deviator Strain*, where nonlinear effects dominate.

Empirical results and visual analyses confirm that physics priors enhance accuracy, stabilize training, and improve trajectory fidelity. The framework remains interpretable and computationally tractable, supporting application to other laboratory-scale geotechnical datasets.

Future work could incorporate additional priors from energy conservation or constitutive modeling, extend inputs beyond elapsed time, and test across diverse geomaterials and loading conditions. Such extensions would clarify the broader role of physics-informed learning in linking experimental geomechanics with predictive modeling.

References

[1] Turky Alotaiby, Fathi E Abd El-Samie, Saleh A Alshebeili, and Ishtiaq Ahmad. A review of channel selection algorithms for eeg signal processing. *EURASIP Journal on Advances in Signal Processing*, 2015(1):66, 2015.

[2] Brandon Amos and J Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks. In *International conference on machine learning*, pages 136–145. PMLR, 2017.

[3] Alan Wilfred Bishop and David John Henkel. The measurement of soil properties in the triaxial test. 1962.

[4] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

[5] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[6] Satrajit Chatterjee and Piotr Zielinski. On the generalization mystery in deep learning. *arXiv preprint arXiv:2203.10036*, 2022.

[7] Minhaz Chowdhury, Shreoshi Roy Shrima, and Md Shofiqul Islam. Comparative analysis of deep learning architectures for multi-class mineral classification: a study using efficientnet and resnet models. *Earth Science Informatics*, 18(3):1–24, 2025.

[8] Bernard D Coleman, Bernard D Coleman, Bernard D Coleman, Etats-Unis Chimiste, and Bernard D Coleman. *Thermodynamics of materials with memory*. Springer, 1972.

[9] Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.

[10] Chandra S Desai. Constitutive modeling using the disturbed state concept. *Continuum models for materials with microstructure*, 1993.

[11] Ronald DeVore, Boris Hanin, and Guergana Petrova. Neural network approximation. *Acta Numerica*, 30:327–444, 2021.

[12] Harris Drucker, Christopher J Burges, Linda Kaufman, Alex Smola, and Vladimir Vapnik. Support vector regression machines. *Advances in neural information processing systems*, 9, 1996.

[13] David R Gaskell and David E Laughlin. *Introduction to the Thermodynamics of Materials*. CRC press, 2024.

[14] Gene H Golub and James M Ortega. *Scientific computing: an introduction with parallel computing*. Elsevier, 2014.

[15] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

[16] Somdatta Goswami, Cosmin Anitescu, Souvik Chakraborty, and Timon Rabczuk. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture. *Theoretical and Applied Fracture Mechanics*, 106:102447, 2020.

[17] Wolfgang R Habel and Katerina Krebber. Fiber-optic sensor applications in civil and geotechnical engineering. *Photonic sensors*, 1(3):268–280, 2011.

[18] ECF Hamman, DJ du Plooy, and JM Seery. Data management and geotechnical models. In *Deep Mining 2017: Proceedings of the Eighth International Conference on Deep and High Stress Mining*, pages 461–487. Australian Centre for Geomechanics, 2017.

[19] Trevor Hastie. The elements of statistical learning: data mining, inference, and prediction, 2009.

[20] Amad Hussen, Tanveer Alam Munshi, Minhaz Chowdhury, Labiba Nusrat Jahan, Abu Bakker Siddique, and Mahamudul Hashan. Evaluating reservoir permeability from core data: Leveraging boosting techniques and ann for heterogeneous reservoirs. *Applied Computing and Geosciences*, page 100247, 2025.

[21] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[22] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[23] Poul V Lade. *Triaxial testing of soils*. John Wiley & Sons, 2016.

[24] Isaac E Lagaris, Aristidis Likas, and Dimitrios I Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 9(5):987–1000, 1998.

[25] Hengxing Lan, C Derek Martin, and Bo Hu. Effect of heterogeneity of brittle rock on micromechanical extensile behavior during compression loading. *Journal of Geophysical Research: Solid Earth*, 115(B1), 2010.

[26] Erastus H Lee. Elastic-plastic deformation at finite strains. 1969.

[27] Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3):218–229, 2021.

[28] Teruo Nakai. *Constitutive modeling of geomaterials: principles and applications*. CRC press, 2012.

[29] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. On the difficulty of training recurrent neural networks. In *International conference on machine learning*, pages 1310–1318. Pmlr, 2013.

[30] SGOPAL Patro and Kishore Kumar Sahu. Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*, 2015.

[31] Hua-Fu Pei, Jun Teng, Jian-Hua Yin, and Rui Chen. A review of previous studies on the applications of optical fiber sensors in geotechnical health monitoring. *Measurement*, 58:207–214, 2014.

[32] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

[33] Junuthula Narasimha Reddy. An introduction to the finite element method. *New York*, 27(14), 1993.

[34] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, 1(5):206–215, 2019.

[35] CA Tang, H Liu, PKK Lee, Yi Tsui, and LGa Tham. Numerical studies of the influence of microstructure on rock failure in uniaxial compression—part i: effect of heterogeneity. *International Journal of Rock Mechanics and Mining Sciences*, 37(4):555–569, 2000.

[36] Fumio Tatsuoka, Masanori Ishihara, BENEDETTO HERVE DI, and Reiko Kuwano. Time-dependent shear deformation characteristics of geomaterials and their simulation. *Soils and foundations*, 42(2):103–129, 2002.

[37] Jiawei Xie, Jinsong Huang, Cheng Zeng, Shui-Hua Jiang, and Nathan Podlich. Systematic literature review on data-driven models for predictive maintenance of railway track: Implications in geotechnical engineering. *Geosciences*, 10(11):425, 2020.

[38] Pengcheng Xu, Xiaobo Ji, Minjie Li, and Wencong Lu. Small data machine learning in materials science. *npj Computational Materials*, 9(1):42, 2023.

[39] TH Yang, LG Tham, CA Tang, ZZ Liang, and Y Tsui. Influence of heterogeneity of mechanical properties on hydraulic fracturing in permeable rocks. *Rock mechanics and rock engineering*, 37(4):251–275, 2004.