

Physics-informed LSTM for forecasting laboratory fault instability and its uncertainty analysis

JianHuan Liu^{1,2}, JiaJun zou¹, HuiRu Lei², Hao Chen³, JianYe Chen^{2*}

1 School of Mathematics and Computer Sciences, Nanchang University, Nanchang 330031, China

2 State Key Laboratory of Earthquake Dynamics and Forecasting, Institute of Geology, China Earthquake Administration, Beijing 100029, China

3 Northwest Electric Power Design Institute Co., Ltd. of China Power Engineering Consulting Group, Xian 710075, China

Correspondence to: J. Y. Chen, jychen@ies.ac.cn

Key Points:

- A physics-informed LSTM is proposed for forecasting laboratory fault shear stress.
- Embedded elastic coupling constraints suppress non-physical forecasts and reduce uncertainty.
- The elastic coupling constraint boosts forecasting accuracy under data-scarce conditions.

Abstract: Reliable forecasting of fault shear stress evolution is critical for understanding fault instability and mitigating seismic hazards. However, most laboratory earthquake studies focus on inferring current fault states rather than forecasting future shear stress evolution. We propose a novel physics-informed long short-term memory (PI-LSTM) network for fault shear stress forecasting in laboratory seismic cycles. It takes ultrasonic wave speed and spectral amplitude from acoustic emission monitoring as inputs and embeds the elastic coupling relationship between the fault and the host rock into the LSTM framework to ensure mechanical consistency. We compare the PI-LSTM with a purely data-driven LSTM using double-direct-shear experimental data under four training data ratios (70%, 50%, 30%, and 10%) and quantify forecasting

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as doi: [10.26464/epp2026080](https://doi.org/10.26464/epp2026080).

This article is protected by copyright. All rights reserved

uncertainty via Monte Carlo Dropout. Results show that the PI-LSTM performs comparably in data-rich scenarios but significantly outperforms the baseline under data-limited conditions. The physical constraint effectively shrinks the solution space, suppresses non-physical predictions, and reduces uncertainty, as reflected by narrower 95% confidence intervals (CIs) and lower differential entropy; empirical coverage analysis further confirms the improved reliability of the physically constrained forecasts. Overall, this work transforms a black-box model into a physically interpretable forecaster and offers a promising pathway toward transferring lab-trained models to natural fault settings where observational data are scarce.

Keywords: laboratory fault-state forecasting; physics-informed neural network; uncertainty quantification

1. Introduction

Reliable forecasting of fault instability remains a grand scientific challenge with profound societal implications, as destructive earthquakes pose persistent threats to communities and critical infrastructure worldwide (Beroza et al., 2021; Allen and Stogaitis, 2022). Despite decades of research, short-term earthquake forecasting in nature remains elusive, largely owing to the inaccessibility of deep seismogenic fault zones, scarce in situ observations, and long recurrence intervals of tectonic earthquakes spanning decades to centuries (Rouet-Leduc et al., 2019; Pritchard et al., 2020; Wang K et al., 2021). Laboratory earthquake experiments, which simulate frictional stick-slip failure under controlled conditions, provide invaluable approaches for unraveling the physical mechanisms of fault instability and validating predictive models (Marone, 1998; Bolton et al., 2020; Dong P et al., 2022, 2023; Huang et al., 2026). These experiments reproduce the seismic cycle, capture variations in elastic properties and precursory signals, and enable systematic testing of forecasting approaches under well-constrained mechanical conditions (Scholz, 1968; Shreedharan et al., 2021; Yamashita et al., 2021).

In recent years, machine learning (ML) and deep learning (DL) have revolutionized laboratory earthquake prediction by leveraging acoustic emission (AE) signals and seismicity catalogs to infer fault states (Rouet-Leduc et al., 2017, 2018; Lubbers et al., 2018; Corbi et al., 2019; Hulbert et al., 2020; Shokouhi et al., 2021; Laurenti et al., 2022; Borate et al., 2023; Karimpouli et al., 2023; Liu JH et al., 2026). Data-driven models, including random forests, convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and temporal convolutional networks (TCNs), have achieved high accuracy in estimating time-to-failure (TTF), instantaneous shear stress, and fault friction (Johnson et al., 2021; Norisugi et al., 2025; Wang CL et al., 2025). These advances confirm that precursory signals encoded in fault-zone geophysical observations carry critical information about impending slip failure.

However, critical limitations persist in current laboratory earthquake prediction research. First, the vast majority of studies focus on inferring the remaining time to the next failure event (TTF) or estimating instantaneous fault stress at the current time step, rather than forecasting the evolution of shear stress into the future (Rouet-Leduc et al., 2017; Jaspersion et al., 2021; Laurenti et al., 2022). Forecasting stress evolution is essential to characterize the preparatory phase of fault instability and to quantify the timing and magnitude of impending slip, yet this capability remains underdeveloped. Second, purely data-driven models rely heavily on large volumes of training data, which are impractical to obtain for natural faults; under small-sample conditions, these models suffer from poor generalization and produce physically inconsistent predictions that violate fundamental fault mechanics (Borate et al., 2023; Wang CL et al., 2025). Third, few studies quantify the uncertainty of forecasting outputs, leaving the reliability of predictive results unassessed.

Physics-informed neural networks (PINNs) offer a transformative solution to these limitations by embedding physical laws into the loss function of DL models, ensuring that predictions adhere to mechanical consistency while reducing dependence on labeled data (Raissi et al., 2019; Karniadakis et al., 2021; Borate et al., 2023). Recent applications of PINNs to laboratory earthquakes have improved prediction accuracy and transferability under data-scarce

conditions, but existing PINN frameworks focus on inferring the remaining time from the current moment to the next slip instability, lacking the ability to continuously predict the evolution of future fault stress states (Borate et al., 2023). To date, no study has developed a physics-informed LSTM (PI-LSTM) tailored for forecasting fault shear stress while simultaneously quantifying forecasting uncertainty.

Here, we present a novel PI-LSTM model for forecasting the evolution of fault shear stress during laboratory seismic cycles. The model takes AE wave speed and spectral amplitude as inputs and integrates the elastic coupling relationship between the fault and surrounding host rock as a strict physical constraint within the LSTM-based (Hochreiter and Schmidhuber, 1997) forecasting framework. This design ensures physically consistent forecasting of future stress states while retaining the strong sequential modeling capability of LSTM. We systematically compare the forecasting performance and generalization ability of the PI-LSTM against a purely data-driven LSTM under both data-rich and data-limited conditions. We further conduct a quantitative uncertainty analysis to demonstrate that embedding physical constraints effectively reduces forecasting uncertainty and enhances prediction reliability.

2. Data and Method

2.1 Double-direct-shear (DDS) experiments

This study uses the publicly available laboratory friction dataset from double-direct-shear (DDS) experiments (Borate et al., 2023), which provides synchronized mechanical and ultrasonic measurements across regular and irregular stick-slip cycles, making it ideal for validating the PI-LSTM forecasting model. The DDS experiment (p5270) employs three blocks of Westerly granite to form two parallel fault surfaces, each with an area of $5 \times 5 \text{ cm}^2$. Fine quartz powder with a layer thickness $< 200 \text{ }\mu\text{m}$ is uniformly distributed on the fault surfaces to simulate natural fault gouge. A constant normal stress of 10 MPa is applied, and shear loading is imposed at a steady rate of $8.9 \text{ }\mu\text{m/s}$.

Mechanical signals, including shear stress, normal stress, fault displacement, and slip rate, are recorded continuously at 10 kHz during the experiments. Concurrently, active-source

ultrasonic monitoring is performed using a pair of 500 kHz P-wave piezoelectric transducers; ultrasonic pulses are transmitted every 1 ms, and full waveforms are sampled at 25 MHz. Two key ultrasonic features are extracted from the raw waveforms using established processing workflows: (1) wave speed (v_i), derived by cross-correlating each waveform with a reference signal to quantify time-of-flight variations; and (2) spectral amplitude (A_i), calculated as the Fourier-transformed amplitude at 400 kHz (near the transducer center frequency) within a fixed time window. These ultrasonic features, paired with the mechanical measurements (Figure 1), form a high-quality multivariate time-series dataset covering hundreds of stick-slip cycles, including both regular and irregular fault slip behaviors.

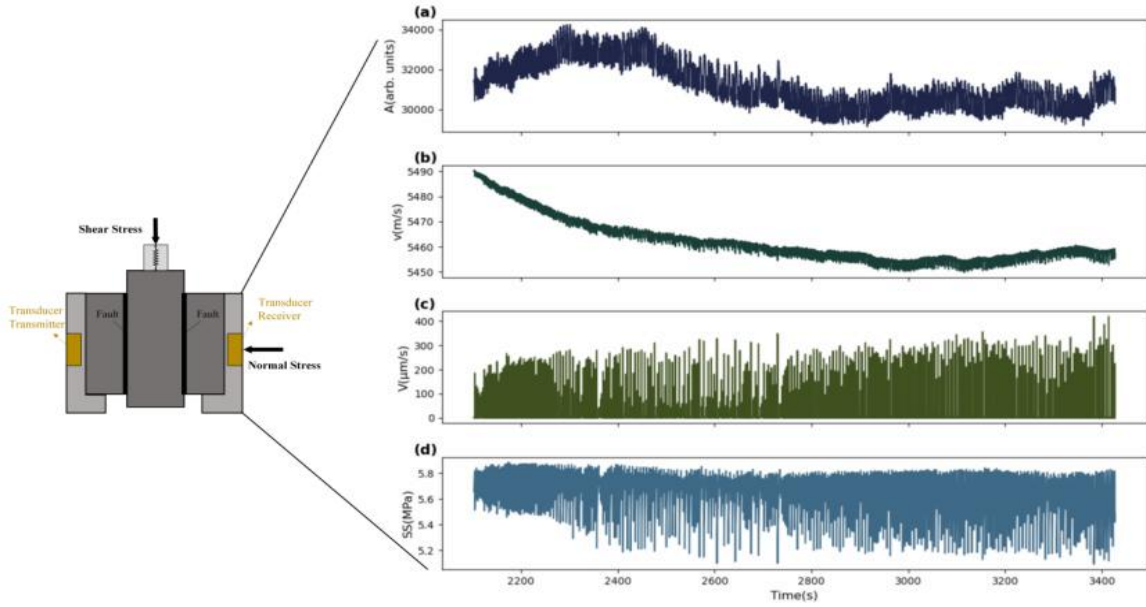


Figure 1. Double-direct-shear (DDS) loading configuration. Input features: (a) ultrasonic wave speed (v_i) and (b) spectral amplitude (A_i) extracted from active-source ultrasonic monitoring. The corresponding mechanical measurements, including slip rate (c) and shear stress (d), are from experiment p5270.

2.2 Data preprocessing

To systematically evaluate model performance, the dataset is first divided into training, validation, and test sets with fixed validation (10%) and test (20%) ratios, while the training set size is varied across four levels: 70%, 50%, 30%, and 10% of the total data (Figure 2). A

sliding-window strategy is then employed to transform the continuous frictional time series into labeled samples for supervised learning. A historical observation window of 3 s and a forecasting horizon of 3 s are specified, enabling the model to take the preceding 3 s of ultrasonic features as input to forecast the subsequent 3-s sequences of shear stress (and slip rate for the PI-LSTM model). The 3-s window length is selected because it approximately matches the average duration of a single seismic cycle in the dataset, corresponding to 300 time steps after resampling the original data. The window advances by one time step per iteration; this overlapping windowing scheme augments the dataset size while preserving the inherent temporal dependencies of fault dynamic behavior. Min-max normalization is applied to the input to standardize the data range to $[0, 1]$. To avoid data leakage, the normalization scaling parameters (minimum and maximum values) are computed exclusively on the training set, and the same parameters are used to scale the validation and test sets.

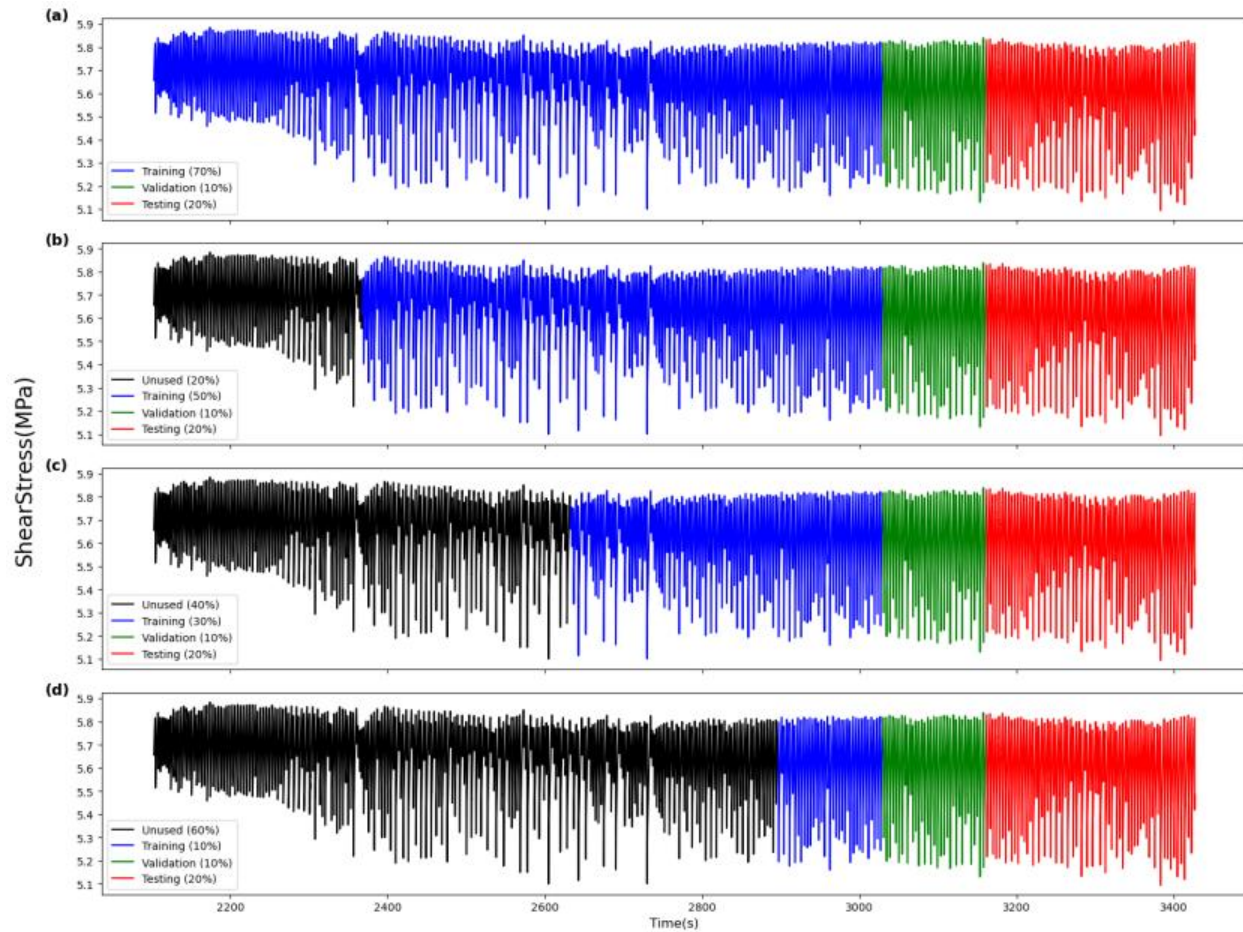


Figure 2. Different data partitioning strategies for the shear stress. (a) 70% training–10% validation–20% test split. (b–d) Custom partitioning schemes with varying proportions of unused raw data. Blue: training set; green: validation set; red: test set; black: unused data.

2.3 Data-driven baseline model

An LSTM network is constructed as the baseline data-driven model for shear stress forecasting (Figure 3), optimized to capture long-range temporal dependencies in the ultrasonic feature sequences. The model architecture comprises a unidirectional LSTM encoder and a multi-layer fully connected (FC) regression head, designed for end-to-end mapping from historical ultrasonic observations to future shear stress evolution.

The LSTM encoder is configured with 64 hidden units to extract high-dimensional temporal representations from the input sequential features. The hidden state vector at the final time step of the LSTM output is selected as the global temporal feature, summarizing the integrated dynamic state of the fault within the 300-step observation window. A dropout layer is appended after the LSTM module to mitigate overfitting and enhance model generalization. The regression head consists of two FC layers with ReLU activation functions, which compress the feature dimensions sequentially from 64 to 32 and then to 16, reducing parameter complexity while preserving nonlinear dynamic information. A final linear output layer generates the forecasted shear stress values for the 300-step forecasting horizon.

For model training, the mean squared error (MSE) is used as the loss function to quantify the discrepancy between forecasted and true shear stress. The model is optimized with the Adam optimizer (initial learning rate = 1×10^{-3}) and L_2 regularization (weight decay = 1×10^{-4}) to suppress overfitting. An early stopping strategy (patience = 20 epochs) is implemented: training terminates if the validation loss does not improve for 20 consecutive epochs, and the model weights with the minimum validation loss are saved for testing.

2.4 Physics-informed LSTM (PI-LSTM) model

The PI-LSTM model shares an identical network architecture, hyperparameters, and training configuration with the data-driven baseline, with two critical extensions: an additional slip-rate forecasting branch and a differentiable physical constraint module that enforces fault mechanical consistency (Figure 3). The slip-rate forecasting branch is explicitly added to compute the physical residual term, linking the model outputs to the underlying fault physics.

The physical constraint embedded in the model is the elastic coupling relationship between the fault and surrounding host rock, derived from a single-degree-of-freedom spring-slider system abstraction of the DDS experimental setup. The governing equation for friction coefficient (μ) evolution is defined as:

$$\frac{d\mu}{dt} = \frac{d}{dt} \left(\frac{\tau_i}{\sigma} \right) = \frac{k}{\sigma} (V_l - V_i) \quad (1)$$

where $\mu = \frac{\tau_i}{\sigma}$ is the friction coefficient, k is the total system stiffness, τ_i is the shear stress, V_i is the predicted slip rate, and σ (normal stress) and V_l (loading velocity) are known experimental constants.

A differentiable gradient operator module is designed to impose this physical constraint. Taking the forecasted shear stress ($\hat{\tau}_i$) and slip rate ($\hat{V}_i$) sequences as inputs, the module approximates the temporal derivative via backward difference. The physical parameters σ , V_l , and k are treated as trainable parameters, with a Softplus activation function applied to ensure non-negativity and avoid physically invalid values. The physical residual loss (L_{phy}) is computed as the MSE of the deviation from the elastic coupling law:

$$L_{\text{phy}} = \frac{1}{N} \sum_{i=1}^N \left[\frac{d}{dt} \left(\frac{\hat{\tau}_i}{\sigma} \right) - \frac{k}{\sigma} (V_l - \hat{V}_i) \right]^2 \quad (2)$$

where N is the number of time steps in the sequence.

The total loss function of the PI-LSTM model is a weighted combination of three MSE-based loss components: shear stress forecast loss (L_{ss}), slip rate forecast loss (L_{v}), and physical residual loss (L_{phy}):

$$L_{\text{total}} = \lambda_{\text{ss}} L_{\text{ss}} + \lambda_{\text{v}} L_{\text{v}} + \lambda_{\text{phy}} L_{\text{phy}} \quad (3)$$

where λ_{ss} , λ_v and λ_{phy} are weighting coefficients for the shear stress, slip rate and physical residual losses, respectively. All other training settings (optimizer, learning rate, batch size, early stopping) are identical to the data-driven model to ensure a fair comparison.

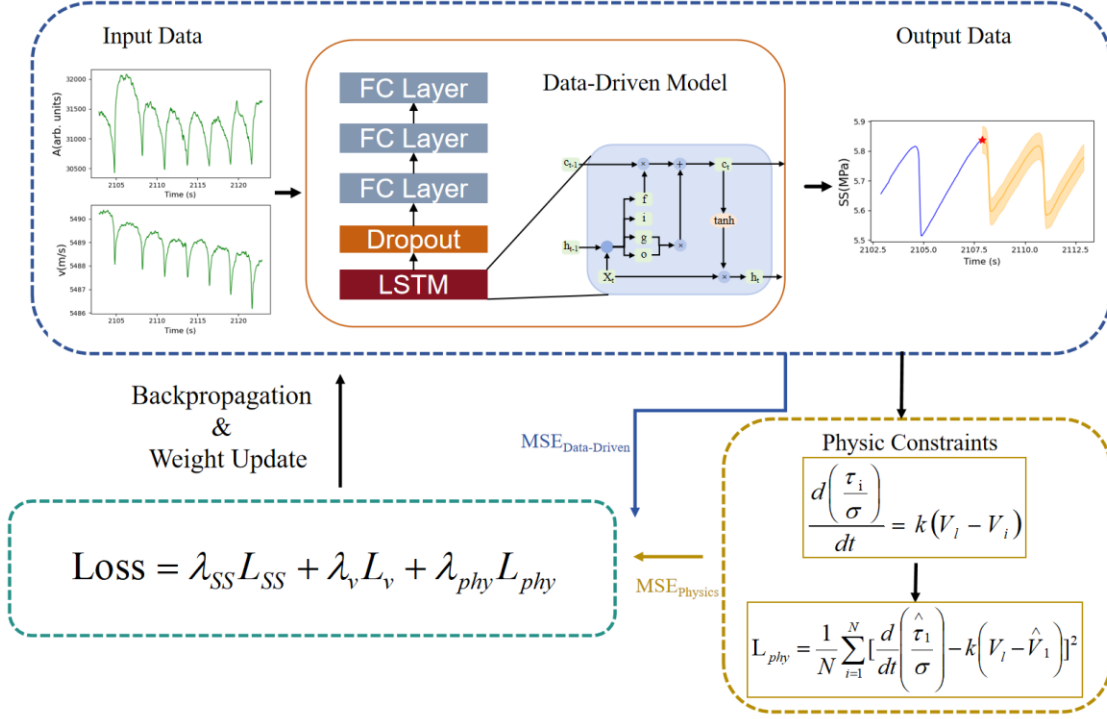


Figure 3. Schematic of the data-driven LSTM model and the PI-LSTM model with embedded physical constraints. The data-driven model uses an LSTM encoder and fully connected (FC) layers for shear stress forecasting; the PI-LSTM model adds a slip-rate forecasting branch and a physical loss term to enforce elastic coupling between the fault and host rock.

2.5 Monte Carlo Dropout Uncertainty Analysis

Monte Carlo Dropout is employed to quantify the epistemic uncertainty of the data-driven and PI-LSTM models in shear stress forecasting. This Bayesian approximation method keeps dropout layers active during inference, generating a set of stochastic forecasts via repeated forward passes to approximate sampling from the model's posterior predictive distribution.

For a given input historical sequence X , the model performs T stochastic forward passes to obtain a set of forecasted shear stress samples $\hat{\tau}_1, \hat{\tau}_2, \dots, \hat{\tau}_T$. The forecasted mean ($\mu(X)$) and variance ($\sigma^2(X)$) of the shear stress are calculated as:

$$\mu(X) = \frac{1}{T} \sum_{t=1}^T \hat{\tau}_t \quad (4)$$

$$\sigma^2(X) = \frac{1}{T-1} \sum_{t=1}^T (\hat{\tau}_t - \mu(X))^2 \quad (5)$$

where $\mu(X)$ is the expected model forecast, and $\sigma^2(X)$ measures the dispersion of forecasts, quantifying the model's epistemic uncertainty.

To prevent model overfitting and enable Bayesian approximate inference via Monte Carlo Dropout, we place a Dropout layer with a dropout rate of 0.2 after the LSTM encoder. The Dropout layer remains activated during model inference, and the predictive mean and variance are obtained through multiple stochastic forward passes.

Taking both computational efficiency and estimation stability into account, we adopt $T = 200$ stochastic samples to compute the predictive mean, predictive variance, differential entropy, and 95% confidence interval. Sensitivity analysis (Figure 4) shows that both the differential entropy and the width of the 95% confidence interval have essentially converged when T reaches 200. Further increasing T exerts a negligible effect on the uncertainty estimation results, so $T = 200$ strikes a favorable balance between computational cost and estimation accuracy.

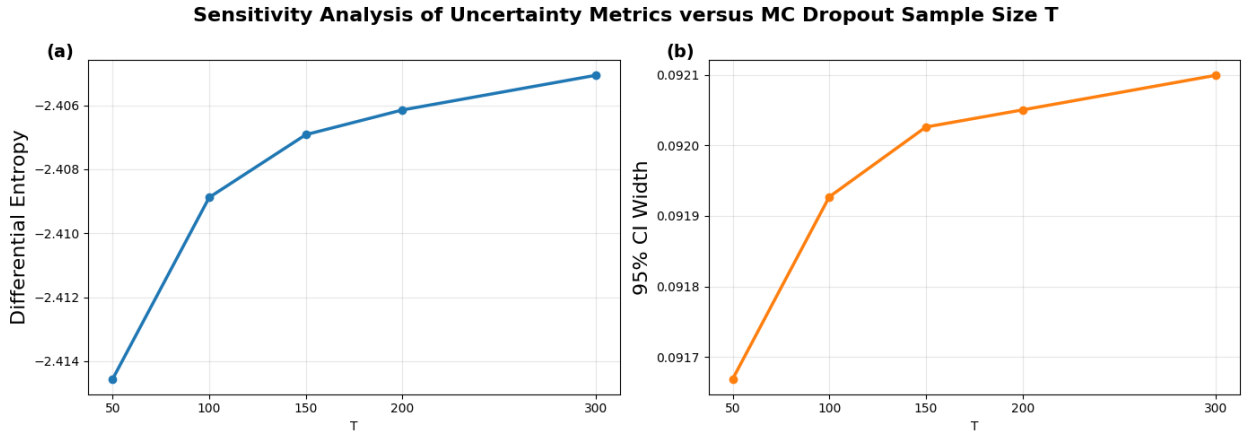


Figure 4. Variation of differential entropy and 95% CI width with T for the PI-LSTM model (10% training data).

Under a Gaussian approximation of the forecasting distribution, a 95% confidence interval (CI) is constructed to visualize the reliable forecast range:

$$95\% \text{ CI} = [\mu(X) - 1.96\sigma(X), \mu(X) + 1.96\sigma(X)] \quad (6)$$

where $\mu(X)$ and $\sigma(X)$ represent the predictive mean and predictive standard deviation estimated from multiple stochastic forward passes using Monte Carlo Dropout, respectively. The coefficient 1.96 corresponds to the critical value of the standard normal distribution for a two-sided 95% confidence interval. Under the assumption that the predictive distribution is approximately Gaussian, the 95% confidence interval is computed as $\mu(X) \pm 1.96\sigma(X)$.

To fully characterize uncertainty, two metrics are computed: the width of the 95% CI and differential entropy. The correlation between these uncertainty metrics and the R^2 score (forecasting accuracy) is analyzed to validate the reliability of uncertainty estimates. A negative correlation between uncertainty and accuracy confirms that MC Dropout effectively identifies low-confidence forecast regions, providing a quantitative basis for fault slip risk assessment and model reliability analysis.

3. Results

3.1 Forecasting Performance of Data-Driven and PI-LSTM Models

We evaluate the forecasting performance of the data-driven LSTM baseline and the PI-LSTM using active-source ultrasonic monitoring signal features (wave speed v_i and spectral amplitude A_i) as inputs to forecast the future evolution of shear stress τ_i . For the PI-LSTM, slip rate V_i is additionally forecasted to enforce the elastic coupling constraint between the fault and host rock. Model performance is quantified using the coefficient of determination (R^2) and root-mean-square error (RMSE) across training, validation, and test sets, with four training data fractions (70%, 50%, 30%, 10%) to mimic data-limited scenarios relevant to field geophysical monitoring.

When training data are abundant ($\geq 30\%$ of the full dataset), the data-driven and PI-LSTM models yield near-identical RMSE values (Table 1). This similarity arises because, with sufficient training data, data-driven models can fully capture the complex statistical relationship between AE signals and shear stress, so the additional introduction of physical constraints yields only marginal benefits. Conversely, since the PI-LSTM model optimizes both prediction error and physical consistency simultaneously, its optimization objective is no longer solely to minimize prediction error but to strike a trade-off between prediction accuracy and physical consistency.

As a result, the degrees of freedom of the model are constrained to some extent, leading to a slight increase in RMSE and a decrease in R^2 under the 50% training data condition. This phenomenon indicates that physical constraints essentially act as a form of physical regularization, which narrows the feasible solution space by introducing inductive bias. When data are abundant, such additional constraints may slightly limit the model's capacity to fit the complex experimental dynamics; by contrast, under data-scarce conditions, they can effectively mitigate overfitting and improve generalization. Furthermore, the elastic coupling relationship adopted in this study is derived from a simplified single-degree-of-freedom spring-slider model, which cannot fully describe complex physical processes such as grain rearrangement, local stress heterogeneity, and fault gouge evolution. For this reason, this approximate physical constraint may introduce minor bias when data are sufficient.

However, as the training data fraction drops to 30% and below, a pronounced performance gap emerges: the PI-LSTM maintains high R^2 and low RMSE on the test set, while the data-driven model suffers severe degradation in generalization (Figure 5, Table 1). At 10% training data, the PI-LSTM reduces test RMSE by 20.6% and increases R^2 by 11.9% relative to the data-driven model, demonstrating the critical role of physical constraints in small-sample learning.

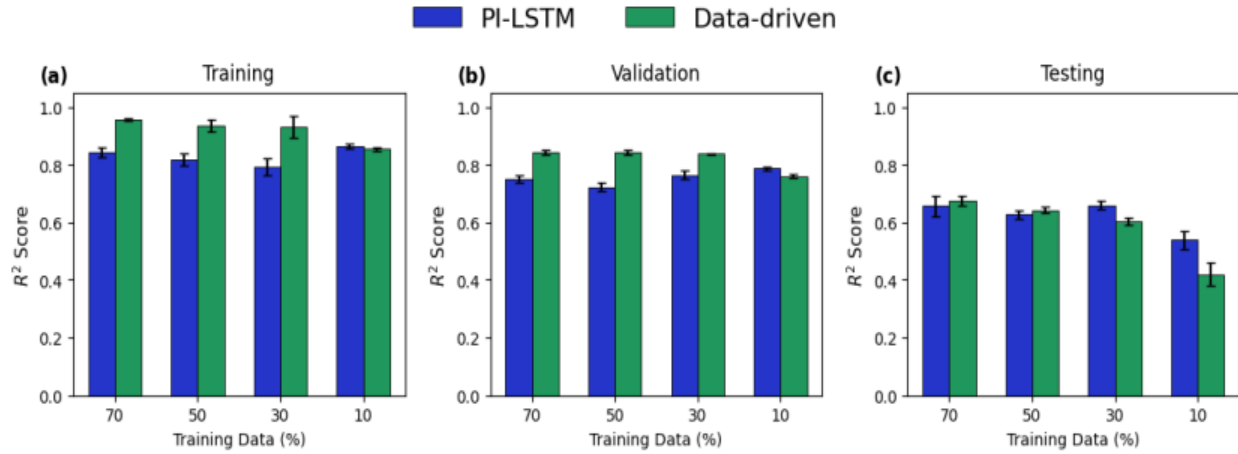


Figure 5. Performance of data-driven and PI-LSTM models across different training data fractions. R^2 scores for (a) training, (b) validation, and (c) test sets are shown for training data fractions of 70%, 50%, 30%, and 10%. Error bars denote one standard error across three random initializations.

Model	Data Split (train-val-test)	RMSE	
		Data-Driven	PI-LSTM
1	70%-10%-20%	0.11741	0.11352
2	50%-10%-20%	0.11129	0.11375
3	30%-10%-20%	0.12447	0.11776
4	10%-10%-20%	0.15981	0.12686

Table 1 RMSE of data-driven and PI-LSTM models for different train-validation-test splits. All values are computed on normalized shear stress data for the test set.

This performance divergence stems from the physical regularization imposed by the fault-host rock elastic coupling equation, which restricts the model's output to physically consistent trajectories of friction coefficient, shear stress, and slip rate, shrinking the feasible solution space and suppressing non-physical predictions. Scatter density plots of predicted vs. true shear stress (Figure 6) confirm this effect: at 10% training data, the data-driven model produces widely

scattered predictions that deviate sharply from the 1:1 line, while the PI-LSTM retains tight clustering around the perfect prediction line. Figure 7 further reveals that the PI-LSTM accurately tracks transient, abrupt stress drops during stick-slip events—regions where the data-driven model exhibits instability. The narrower error bars of the PI-LSTM in Figure 7 also indicate improved forecasting stability. Collectively, these results show that embedding frictional physics into the loss function transforms the LSTM from a purely statistical fitter to a physically consistent forecaster when observational data are scarce.

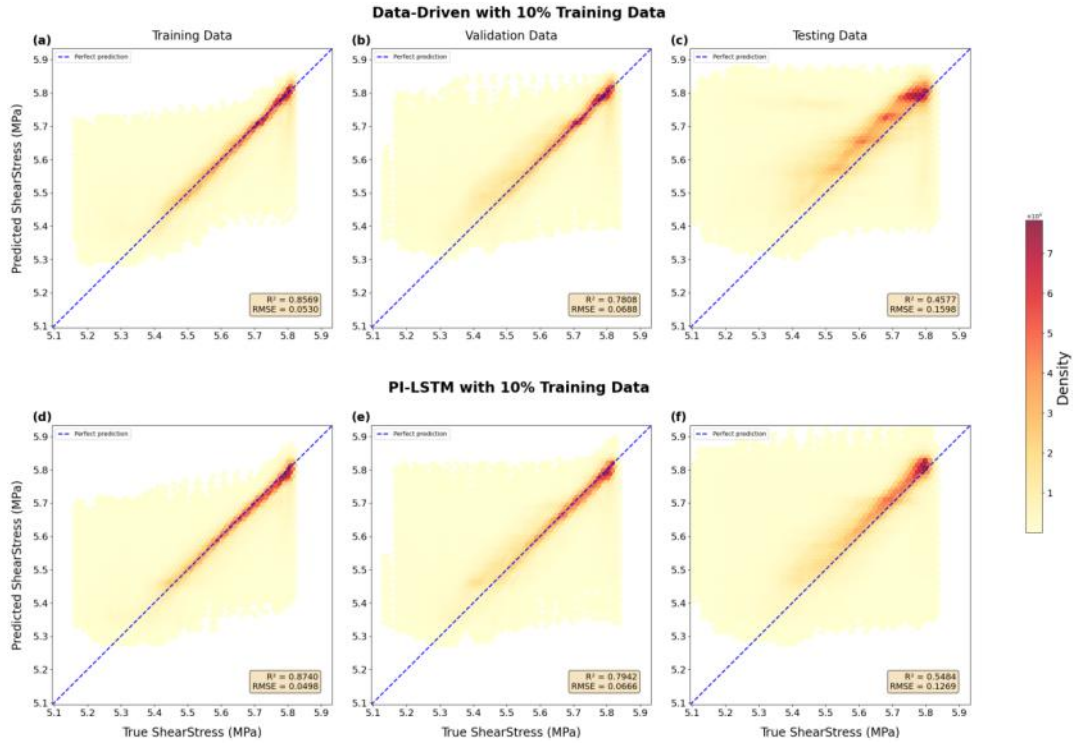


Figure 6. Scatter density plots of predicted vs. true shear stress at 10% training data. Data-driven model predictions on (a) training, (b) validation, and (c) test sets; PI-LSTM predictions on (d) training, (e) validation, and (f) test sets.

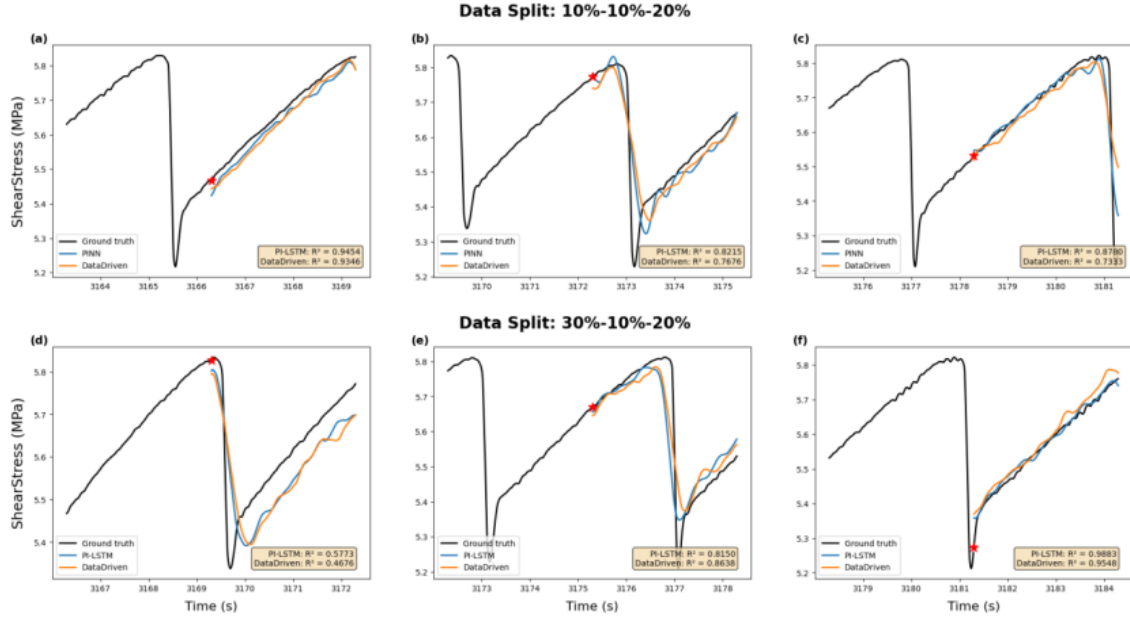


Figure 7. Shear stress forecasts for representative test samples. Forecasting results for data-driven and PI-LSTM models with 10% training data (a–c) and 30% training data (d–f). Black lines represent ground truth, while colored lines indicate model forecasts.

3.2 Uncertainty Analysis via Monte Carlo Dropout

We quantify the epistemic uncertainty of shear stress forecasts using Monte Carlo Dropout (MC Dropout) (Gal and Ghahramani, 2016), which approximates Bayesian inference by enabling dropout layers during inference to sample from the model's posterior predictive distribution. This analysis characterizes prediction reliability and links uncertainty to physical consistency. We evaluate uncertainty using two metrics: differential entropy (a measure of distribution spread) and the width of the 95% confidence interval (CI), and we validate their correlation with forecasting accuracy (R^2).

For the PI-LSTM at 10% training data, differential entropy and 95% CI width both show significant negative correlations with R^2 (Figure 8), with correlation coefficients of -0.3469 and -0.2974 , respectively. This confirms that higher uncertainty (larger entropy, wider CI) corresponds to lower prediction accuracy, validating that MC Dropout reliably identifies low-confidence forecasts. A direct comparison of uncertainty distributions for test data (Figure 9) reveals that the PI-LSTM yields a lower mean differential entropy (-2.4061 vs. -2.3398) and narrower mean 95%

CI width (0.0921 vs. 0.0970) than the data-driven model, indicating tighter, more reliable forecast distributions. Figure 10 illustrates the forecasted shear stress for the test data at different time windows, showing that the PI-LSTM maintains narrower, well-constrained 95% CIs across entire stick-slip cycles compared with the data-driven approach. The reduction in uncertainty from the PI-LSTM arises directly from the physical constraint, which prunes unphysical solutions and stabilizes the model's parameter space. In the data-driven model, sparse training data allows the network to explore implausible stress-slip trajectories, increasing distribution variance; in contrast, the elastic coupling constraint forces the model to adhere to the fundamental dynamics of fault friction, concentrating forecasts around physically meaningful trajectories.

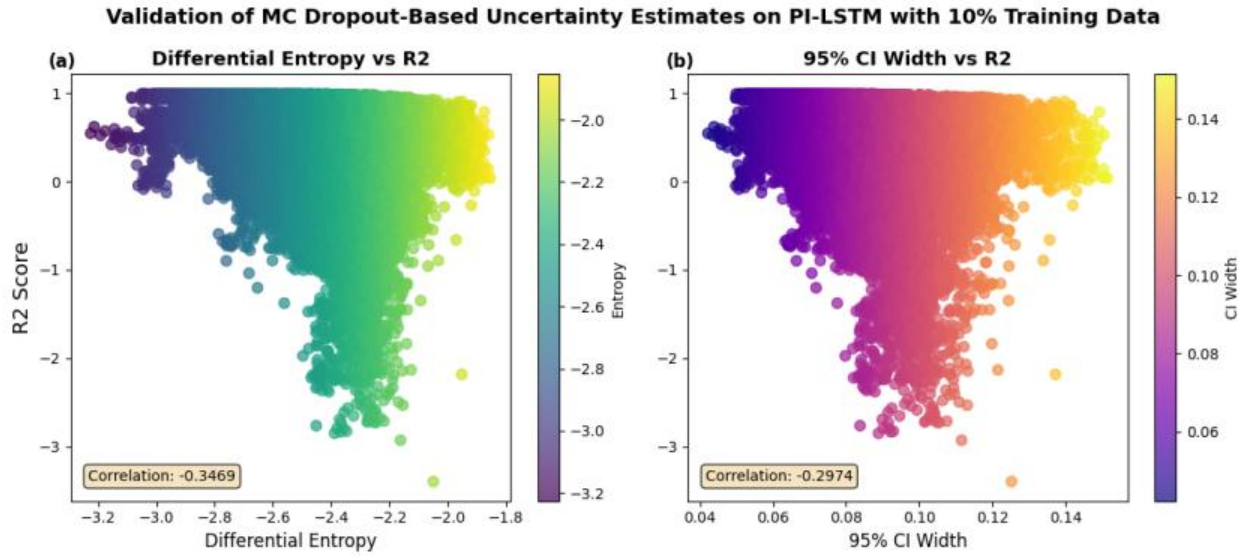


Figure 8. Validation of MC Dropout-derived uncertainty estimates for the PI-LSTM at 10% training data. (a) Correlation between sample-wise differential entropy and R^2 ; (b) correlation between sample-wise 95% CI width and R^2 .



Figure 9. Uncertainty distributions at 10% training data. (a–b) Differential entropy and 95% CI width for the data-driven model; (c–d) differential entropy and 95% CI width for the PI-LSTM. Mean values are annotated for each distribution.

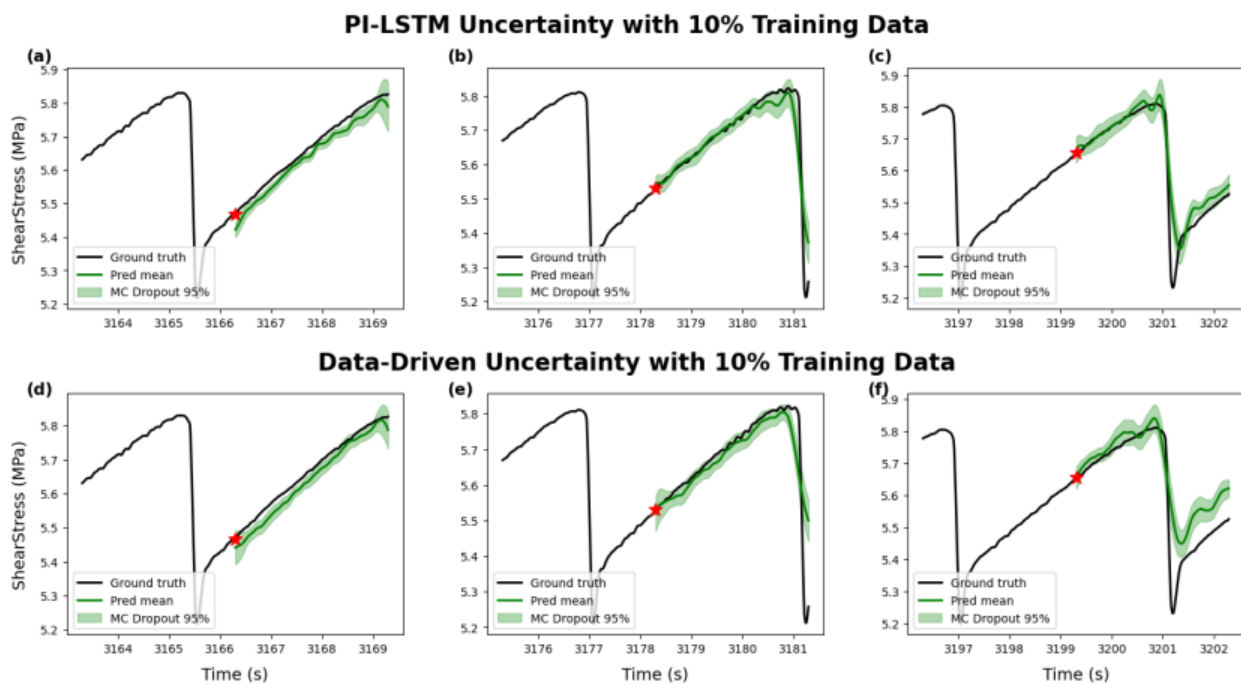


Figure 10. Uncertainty bounds for shear stress forecasts at 10% training data. (a–c) PI-LSTM forecasts with mean (green line) and 95% CI (green shaded area); (d–f) data-driven model forecasts with mean (blue line) and 95% CI (blue shaded area). Black lines represent ground truth.

To further evaluate the calibration performance of Monte Carlo Dropout-based uncertainty estimation, we introduce the Empirical Coverage Rate (ECR) and the Average 95% Confidence Interval Width as evaluation metrics. The empirical coverage rate refers to the proportion of true values in the test set that fall within the predicted 95% confidence interval and is used to measure the calibration performance of prediction intervals; the average 95% confidence interval width is adopted to assess the sharpness of prediction intervals.

Figure 11 presents the calibration results of the two models under different training data fractions. The empirical coverage rates of both models are lower than the theoretical nominal level of 95%, indicating that the prediction intervals constructed via Monte Carlo Dropout still exhibit a certain degree of under-calibration. However, under the same training data conditions, the PI-LSTM consistently achieves higher empirical coverage rates: it increases from 0.610 to 0.718 (a 17.7% improvement) with 10% training data, and rises from 0.782 to 0.825 with 30% training data, demonstrating that physical constraints can improve the ability of prediction intervals to cover true values.

Meanwhile, the PI-LSTM yields narrower prediction intervals. Its average 95% confidence interval width decreases from 0.064 to 0.051 under the 10% training data condition, and drops from 0.050 to 0.043 under the 30% training data condition. This indicates that the PI-LSTM can maintain higher prediction sharpness while improving coverage, achieving a better balance between forecast calibration and sharpness. Overall, although the 95% confidence intervals obtained via MC Dropout have not yet reached the theoretical nominal coverage, the PI-LSTM can provide more reliable uncertainty estimation with higher empirical coverage and narrower prediction intervals, further verifying the positive effect of physical constraints on enhancing prediction credibility and stability.

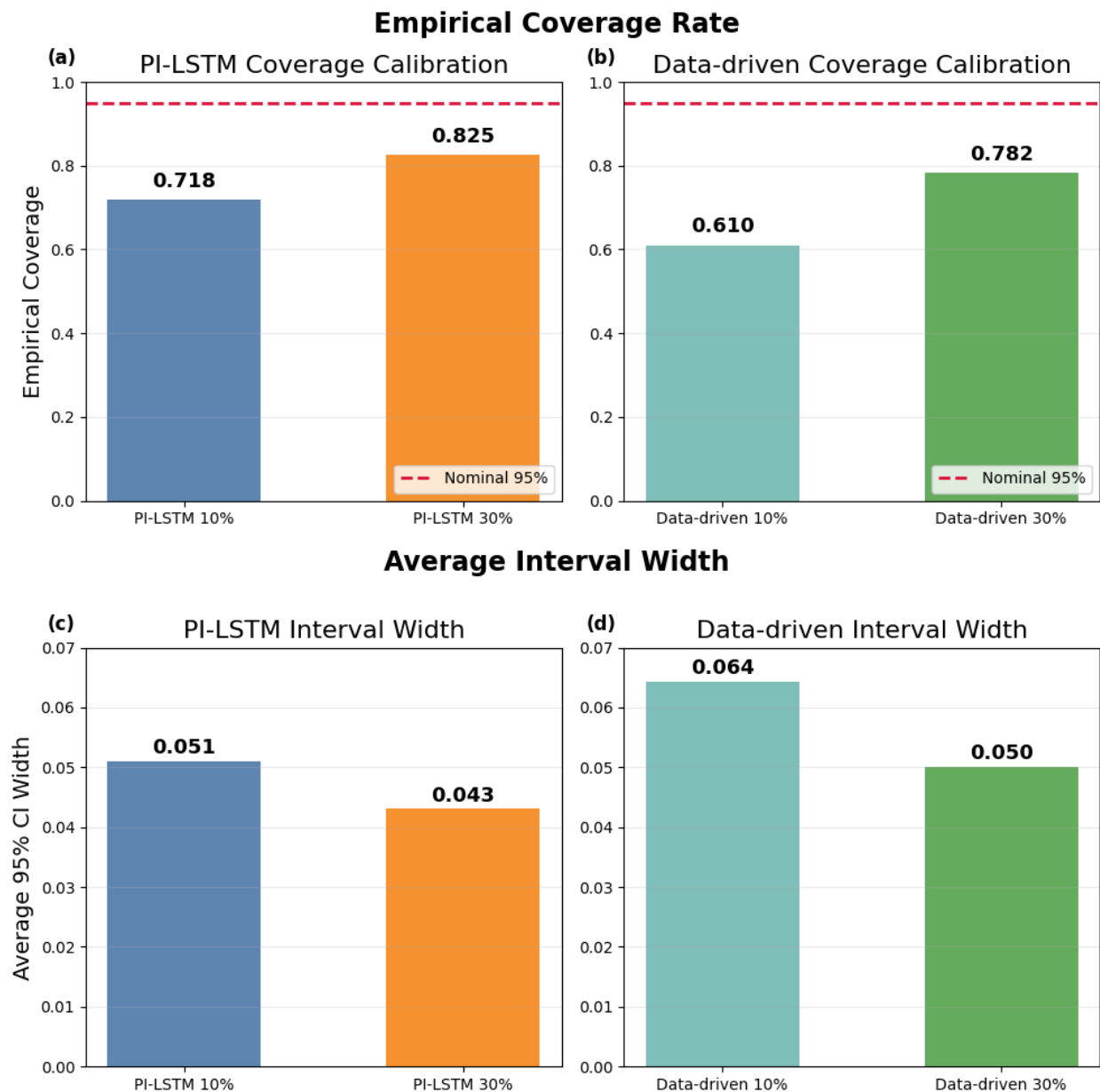


Figure 11. Empirical coverage rate and average interval width from uncertainty analysis for the PI-LSTM and data-driven models under 10% and 30% training data conditions. (a) Empirical coverage rate of the PI-LSTM; (b) empirical coverage rate of the data-driven model; (c) average 95% confidence interval width of the PI-LSTM; (d) average 95% confidence interval width of the data-driven model.

4. Discussion

Our work has important implications for scaling laboratory-derived machine learning models to natural tectonic faults. Natural earthquakes suffer from extreme data scarcity, with recurrence intervals of decades to centuries and limited direct observations, as current geophysical and geodetic methods remain constrained in their ability to precisely resolve in-situ fault stress and slip deficits. The PI-LSTM's strong small-sample performance and uncertainty quantification provide a pathway to transfer lab-trained models to field data via transfer learning, by fine-tuning physically constrained models on limited field observations.

This lab-to-field upscaling offers two potential applications. First, by embedding physical laws into the network, our framework can further constrain rate-and-state friction (RSF) laws and alternative slip laws from sparse data. Because these empirical models are fundamental to simulating earthquake cycles, extracting accurate state variables through our PI-LSTM provides vital constraints on the frictional behavior of natural faults. Second, the model can be adapted to forecast in-situ fault stress drops. While tectonic scaling remains challenging, our approach is valuable for seismic hazard monitoring in unconventional oil and gas exploration. In these settings, microseismicity closely mirrors laboratory stick-slip dynamics, allowing time-series data from active-source monitoring to be coupled with models of stress perturbation or fluid-injection-induced seismicity.

Future work will test cross-experiment generalization across biaxial and triaxial loading setups, mirroring the diversity of natural fault zones. Limitations of the current study include residual forecasting errors during abrupt stress drops, which can be alleviated by adopting quasi-dynamic equations of motion to impose targeted constraints on fault instability states.

5. Conclusions

In this study, we developed a physics-informed long short-term memory (PI-LSTM) framework for forecasting fault shear stress evolution during laboratory seismic cycles. By integrating ultrasonic wave speed and spectral amplitude with the elastic coupling relationship

between the fault and surrounding host rock, the model enables physically consistent forecasting of fault stress dynamics. Forecasting uncertainty was further quantified using Monte Carlo Dropout. Compared with a purely data-driven LSTM, the PI-LSTM achieves similar performance when training data are abundant but exhibits superior generalization under data-limited conditions. The embedded physical constraint effectively reduces the feasible solution space, suppresses non-physical stress trajectories, and improves forecasting robustness, particularly during rapid stress-drop phases associated with stick-slip events. In addition, the PI-LSTM yields lower differential entropy and narrower 95% confidence intervals, and the empirical coverage analysis confirms that it achieves higher coverage with narrower intervals, demonstrating more reliable and better-calibrated probabilistic forecasts. Overall, the proposed framework demonstrates that incorporating fault physics into deep learning can improve both forecasting accuracy and reliability under sparse-data conditions. By unifying physical constraints and uncertainty quantification within a single architecture, the PI-LSTM transforms a conventional black-box predictor into a physically interpretable forecasting model, providing a promising foundation for broader geophysical forecasting applications where observations are limited but governing physical principles are well understood.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China - Joint Fund for Earthquake Science (Grant No. U2539202), the Jiangxi Provincial Natural Science Foundation for Youth (Grant No. 20252BAC200599), and the Central Public-interest Scientific Institution Basal Research Fund (Grant No. IGCEA20240330). We are grateful to the research group for making the experimental dataset publicly available, which enabled the validation of our proposed approach. We also extend our sincere appreciation to the reviewers and the responsible editors for their rigorous and constructive comments, which have significantly enhanced the quality of this manuscript.

References

- Allen, R. M., and Stogaitis, M. (2022). Global growth of earthquake early warning. *Science*, 375(6582), 717-718.
- Beroza, G. C., Segou, M., and Mostafa M. S. (2021). Machine learning and earthquake forecasting—next steps. *Nat. Commun.*, 12(1), 4761.
- Bolton, D. C., Shreedharan, S., Rivière, J., and Marone, C. (2020). Acoustic energy release during the laboratory seismic cycle: Insights on laboratory earthquake precursors and prediction. *J. Geophys. Res.: Solid Earth*, 125(8), e2019JB018975.
- Borate, P., Rivière, J., Marone, C., Mali, A., Kifer, D., and Shokouhi, P. (2023). Using a physics-informed neural network and fault zone acoustic monitoring to predict lab earthquakes. *Nat. Commun.*, 14(1), 3693.
- Corbi, F., Sandri, L., Bedford, J., Funiciello, F., Brizzi, S., Rosenau, M., and Lallemand, S. (2019). Machine learning can predict the timing and size of analog earthquakes. *Geophys. Res. Lett.*, 46(3), 1303-1311.
- Dong, P., Chen, R., Xia, K. W., Yao, W., Peng, Z. G., and Elsworth, D. (2023). Earthquake delay and rupture velocity in near-field dynamic triggering dictated by stress-controlled nucleation. *Seismol. Res. Lett.*, 94(2A), 913-924.
- Dong, P., Xia, K. W., Xu, Y., Elsworth, D., and Ampuero, J. P. (2023). Laboratory earthquakes decipher control and stability of rupture speeds. *Nat. Commun.*, 14(1), 2427.
- Gal, Y., and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In international conference on machine learning, 1050-1059, PMLR.
- Huang, Y., Shen, M., Shen, F. M., Wang, L. F., Liu, P. X., and Ma, S. L. (2026). Machine learning reveals early preseismic signals of laboratory earthquakes. *Earth Planet. Sci. Lett.*, 687, 120091.
- Hulbert, C., Rouet-Leduc, B., Jolivet, R., and Johnson, P. A. (2020). An exponential build-up in seismic energy suggests a months-long nucleation of slow slip in Cascadia. *Nat. Commun.*, 11(1), 4139.

- Jaspersen, H., Bolton, D. C., Johnson, P., Guyer, R., Marone, C., and de Hoop, M. V. (2021). Attention network forecasts time-to-failure in laboratory shear experiments. *J. Geophys. Res.: Solid Earth*, 126(11), e2021JB022195.
- Johnson, P. A., Rouet-Leduc, B., Pyrak-Nolte, L. J., Beroza, G. C., Marone, C. J., Hulbert, C., and Reade, W. (2021). Laboratory earthquake forecasting: A machine learning competition. *Proc. Natl. Acad. Sci.*, 118(5), e2011362118.
- Karimpouli, S., Caus, D., Grover, H., Martínez-Garzón, P., Bohnhoff, M., Beroza, G. C., and Kwiatek, G. (2023). Explainable machine learning for labquake prediction using catalog-driven features. *Earth Planet. Sci. Lett.*, 622, 118383.
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. (2021). Physics-informed machine learning. *Nat. Rev. Phys.*, 3(6), 422-440.
- Laurenti, L., Tinti, E., Galasso, F., Franco, L., and Marone, C. (2022). Deep learning for laboratory earthquake prediction and autoregressive forecasting of fault zone stress. *Earth Planet. Sci. Lett.*, 598, 117825.
- Liu, J. H., Chen, J. Y., Chen, H., Zhang, L., Ma, X., Zhang, M., and Mai, T. L. (2026). Ensemble learning for inferring the time to failure of laboratory faults. *Chin. J. Geophys.*, 69(4), 1634-1644.
- Lubbers, N., Bolton, D. C., Mohd-Yusof, J., Marone, C., Barros, K., and Johnson, P. A. (2018). Earthquake catalog-based machine learning identification of laboratory fault states and the effects of magnitude of completeness. *Geophys. Res. Lett.*, 45(24), 13-269.
- Marone, C. (1998). Laboratory-derived friction laws and their application to seismic faulting. *Annu. Rev. Earth Planet*, 26(1), 643-696.
- Norisugi, R., Kaneko, Y., and Rouet-Leduc, B. (2025). Machine learning predicts meter-scale laboratory earthquakes. *Nat. Commun.*, 16(1), 9593.
- Pritchard, M. E., Allen, R. M., Becker, T. W., Behn, M. D., Brodsky, E. E., Bürgmann, R., and Vincent, H. (2020). New opportunities to study earthquake precursors. *Seismol. Res. Lett.*, 91(5), 2444-2447.

- Raissi, M., Perdikaris, P., and Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.*, 378, 686-707.
- Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., and Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes. *Geophys. Res. Lett.*, 44(18), 9276-9282.
- Rouet-Leduc, B., Hulbert, C., Bolton, D. C., Ren, C. X., Riviere, J., Marone, C., and Johnson, P. A. (2018). Estimating fault friction from seismic signals in the laboratory. *Geophys. Res. Lett.*, 45(3), 1321-1329.
- Rouet-Leduc, B., Hulbert, C., and Johnson, P. A. (2019). Continuous chatter of the Cascadia subduction zone revealed by machine learning. *Nat. Geosci.*, 12(1), 75-79.
- Schmidhuber, J., and Hochreiter, S. (1997). Long short-term memory. *Neural Comput.*, 9(8), 1735-1780.
- Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bull. Seismol. Soc. Am.*, 58(1), 399-415.
- Shokouhi, P., Girkar, V., Rivière, J., Shreedharan, S., Marone, C., Giles, C. L., and Kifer, D. (2021). Deep learning can predict laboratory quakes from active source seismic data. *Geophys. Res. Lett.*, 48(12), e2021GL093187.
- Shreedharan, S., Bolton, D. C., Rivière, J., and Marone, C. (2021). Machine learning predicts the timing and shear stress evolution of lab earthquakes using active seismic monitoring of fault zone processes. *J. Geophys. Res.: Solid Earth*, 126(7), e2020JB021588.
- Wang, C. L., Xia, K. W., Yao, W., and Marone, C. (2025). Generalizable deep learning models for predicting laboratory earthquakes. *Commun. Earth Environ.*, 6(1), 219.
- Wang, K., Johnson, C. W., Bennett, K. C., and Johnson, P. A. (2021). Predicting fault slip via transfer learning. *Nat. Commun.*, 12(1), 7319.
- Yamashita, F., Fukuyama, E., Xu, S., Kawakata, H., Mizoguchi, K., and Takizawa, S. (2021). Two end-member earthquake preparations illuminated by foreshock activity on a meter-scale laboratory fault. *Nat. Commun.*, 12(1), 4302.