https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

A dynamic informed deep learning method for future estimation of laboratory stick-slip

Enjiang Yue[1,2], Mengjiao Qin[3], Linshu Hu[1,2], Sensen Wu[1,2], and Zhenhong Du[1,2*]

[1]School of Earth Sciences, Zhejiang University, Hangzhou 310058, China.
[2]Zhejiang Provincial Key Laboratory of Geographic Information Science, Hangzhou 310058, China.
[3]School of Safety Science and Emergency Management, Wuhan University of Technology, Wuhan, 430081, China.

Correspondence to: Zhenhong Du (duzhenhong@zju.edu.cn)

Abstract. Fault activities modelling holds vital importance for earthquake monitoring, risk management, and early alert. Studies on laboratory earthquakes are instrumental in the modelling of natural fault ruptures and in enhancing our grasp of natural earthquake dynamics. Recently, deep learning methods have been proven effective in predicting instantaneous fault stress in laboratory settings and slow slip events on Earth. However, these methods have struggled to conduct steady future prediction lacking grasping of the complex dynamics of highly nonlinear laboratory fault slip systems. Addressing this, we introduce the Hankel Koopman Auto-encoder (HKAE), a novel method inspired by dynamical system theories. HKAE performs dynamic modelling of laboratory fault system and provides a continuous estimation of the future state of the system. It has been deployed on experiments with different slip behaviours and shows superior ability to predict shear stress variation during a slip cycle and also slip activities in longer-term seismic cycles. The HKAE model surpasses conventional time series prediction deep learning methods, showing superior statistical evaluation metrics like RMSE and R^2 with two prediction horizons. Meanwhile, we find that the HKAE can model the slip dynamics better than purely statistical modelling, as evidenced by its more accurate modelling of the slip timing, slip cycle intervals and its ability to summarize the quasi-periodic dynamics as an operator from a small number of samples to generate more robust beyond-horizon prediction. The capability of HKAE to decompose and model complex temporal dynamics highlights its potential in and sparse-observed geophysical system with quasi-periodic characteristics like natural fault activities.

1 Introduction

Modelling fault activities plays a crucial role in understanding the pattern of seismic activities, monitoring and even predicting the occurrence of earthquakes, and estimating seismic hazards. Laboratory earthquake studies contribute to modelling natural fault ruptures and enhancing our understanding of natural earthquakes (Johnson et al., 2021). It indicates the similar mechanism between slow and fast slips (Hulbert et al., 2019), and inspire to extract fault physical property changes from a large number of earthquake records (Rouet-Leduc et al., 2019). Machine learning has been proven effective in extracting information about the rupture behaviour of laboratory earthquakes from acoustic emission signals for instantaneous prediction. Rouet-Leduc et al. (2017) firstly found that random forest can accurately predict the time-to-failure using acoustic emissions. Subsequently,

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024


Geoscientific Model Development Discussions

stress variation, a crucial physical feature of faults, has been identified and evaluated from acoustic emissions using XGBoost, enabling further analysis of the acoustic signals (Rouet-Leduc et al., 2018). Lubbers et al. (2018) found that the event catalogue, which is more available in natural earthquake, can also predict the transient fault mechanism. Active source seismic data is also a valid data source to predict the instantaneous fault behaviour. Jasperson et al. (2019) and Karimpouli et al. (2023) discussed the prediction methods like traditional machine learning methods, neural networks and explainable machine learning methods. An assessment of the transferability across diverse experiments and simulations was conducted, highlighting its critical role in applying laboratory methods to in-field models (Wang et al., 2021; Borate et al., 2023).

While most studies focus on instantaneous prediction, several have explored future prediction. The state-of-the-art sequence modelling architecture, Transformer, has shown promise in extracting future friction information from continuous acoustic emission signals (Wang et al., 2022). The model's attention score reveals that the closer the friction is to the rupture moment, the stronger the indication of stress drop in seismic records. Laurenti et al. (2022) discovered that laboratory fault zone stress can be inferred autoregressively. Additionally, spatial dimensions have been introduced for the autoregressive spatiotemporal prediction of surface velocity fields during laboratory fault slips (Mastella et al., 2022). Although these studies underscore potential for inferring future behaviour of fault slips, they confront challenges in modelling stability and future lead prediction owing to the intricate dynamics of laboratory fault slip systems. Gualandi et al. (2023) proposed that the laboratory earthquake cycle exhibits characteristics of a low-dimensional system, with an average dimension similar to natural slow earthquakes. Lyapunov exponent analysis reveals the predictability within a certain period, albeit with deterministic and stochastic chaotic behaviours, which are challenging to model using machine learning methods designed from traditional statistical knowledge.

Physics-informed machine learning methods have been proven to be a framework for geoscientific applications (Degen et al., 2023) like glacier modelling (Riel et al., 2021), ocean modelling (Hammoud et al., 2022) and also solid-earth (Okazaki et al., 2022). It introduces domain prior knowledge, the key factor in geoscientific analysis while leveraging the benefits of machine learning. Recent advancements in dynamic theory, led by Koopman theory (Koopman, 1931), have shown efficacy in integrating dynamical insights within a data-driven framework, yielding results more aligned with dynamic situations (Karniadakis et al., 2021). Various methods based on Koopman theory have been acknowledged as powerful for modelling and deciphering complex nonlinear dynamical systems (Brunton et al., 2022), such as in fluid mechanics (Brunton et al., 2020), and have found applications in geophysical fields, including climate (Li et al., 2020; Froyland et al., 2021), ocean variability (Franzke et al., 2022), and electromagnetic fields (Brunton et al., 2017; Lintner et al., 2023).

Given the complicated dynamics of laboratory fault slip systems, we propose a deep learning method imbued with Koopman theory. Instead of focusing on directly learning a mapping from historical data to future predictions, our model envisions the future prediction as the continuous evolution of laboratory fault slip systems. Laboratory fault systems with different slip behaviours under different prediction horizons are adapted to evaluate the effectiveness of HKAE modelling.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024


Geoscientific Model Development Discussions

2 Materials, methods and models

2.1 Laboratory stick-slip data

Our study incorporates two categories of data - one drawn from laboratory experiments carried out with biaxial shear equipment, and the other derived via numerical simulation. The experimental data from the biaxial shear equipment comes from the PSU laboratory (Laurenti et al., 2022). Different shear materials are situated between the two plates to which positive pressure and shear force are exerted from each side, and the equipment is used to record the changes in system properties such as pressure recorded during the shear process (Figure 1a-b). This results in time series data recorded at a temporal sampling rate of 0.001s. We derive the numerical simulation data from a modified rate-and-state friction (RSF) model (Gualandi et al., 2023). Both types of data record several mechanical variables, but here we mainly focus on the variation of shear stress because of its direct indication of the onset of laboratory fault slip. Experiment 4581 (Exp. 4581) and Experiment 5198 (Exp. 5198) demonstrate quasi-periodic slow and fast slip behaviours, respectively, while Experiment 4679 (Exp. 4679) contains a switch between two types of slip behaviours (Figure 1c). We used the simple simulation of the data as an additional validation (Figure 1d), which shows a robust switch between two slip behaviours with stress variation.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions
Open Access
EGU

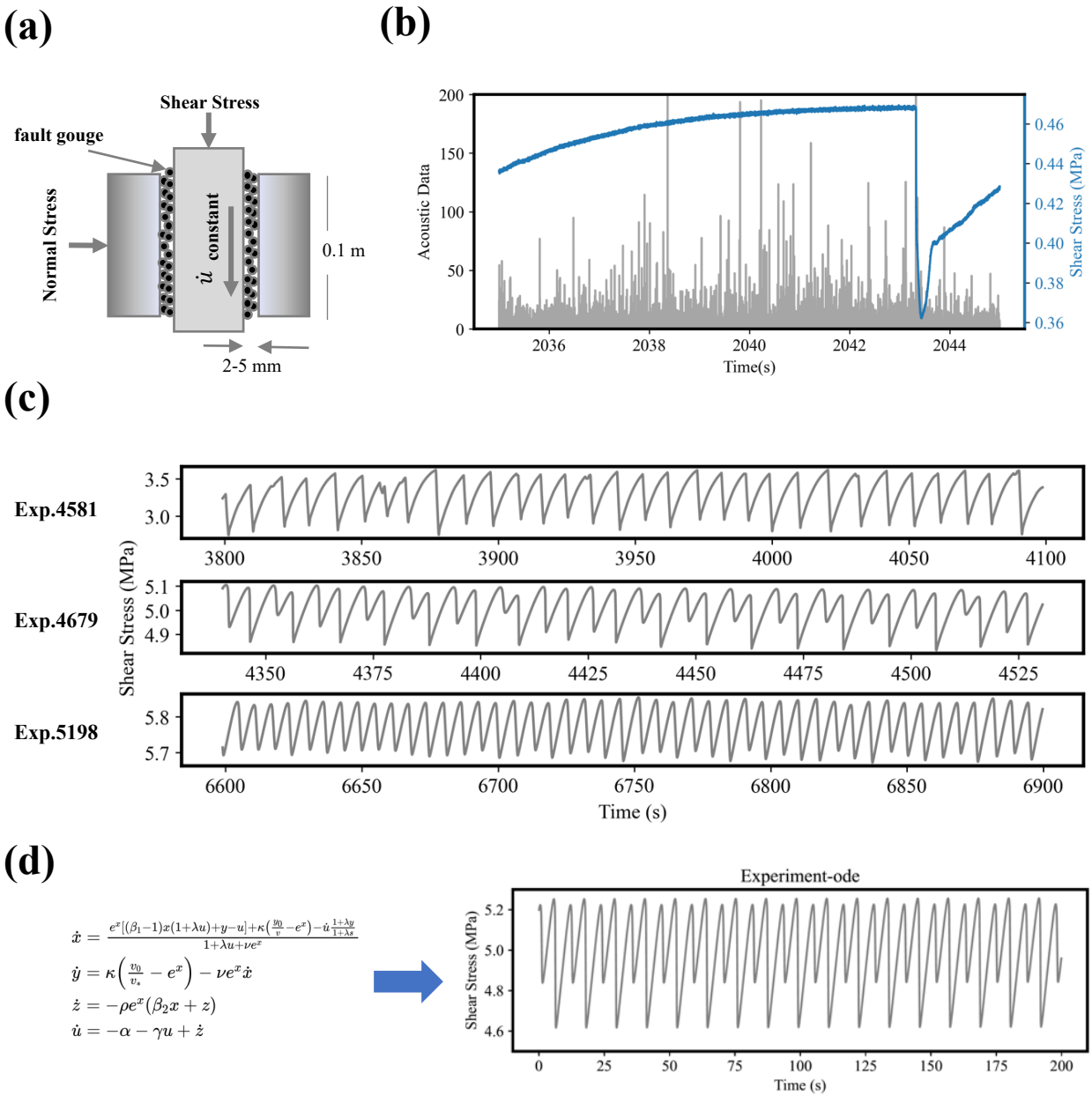


Figure 1: (a) The laboratory fault slip experiment setting and (b) recorded data, acoustic emissions record in gray and shear stress time series in blue. (c) Three modelled experiments with different slip behaviours. (d) Simulated shear stress.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

2.2 Dynamical System Theories

2.2.1 Koopman Theory

Laboratory earthquake can be conceptualized as being governed by a dynamical system, and shear stress can be regarded as a measurement of this system, fits Eq. (1)

$$x_{t+1} = F(x_t) \tag{1}$$

x_t is the shear stress measured from laboratory slip system, F presents the governing function of the system.

Koopman theory is a mathematical theoretical framework. It states that all finite-dimensional nonlinear systems can evolve in an alternative space through the mapping g of the infinite-dimensional Koopman operator K. The Koopman operator on the transformed space can be used directly to perform the linear evolution of the system state as Eq. (2)-(3) and Figure 2 show.

$$Kg(x_t) = g\big(F(x_t)\big) = g(x_{t+1}) \tag{2}$$

$$x_{t+1} = g^{-1}Kg(x_t) \tag{3}$$

Owing to the linear properties of the Koopman operator, linear methods like spectral decomposition can be employed on the operator for enhanced analysis, prediction, and control. The dimension of the the learned approximate Koopman operator indicates the dynamic modes needed to describe the dynamical process, which can be decomposed as follows:

$$K = V\Lambda V^{-1} \tag{4}$$

Where $V = [v_1, v_2, \dots, v_k]$ is the eigen vectors of K, $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_k]$ is the eigenvalues of K. Each eigenvalue describes the strength and oscillatory properties of its corresponding dynamical component:

$$b_j^k = b_0^k e^{\frac{j}{\Delta t} log \lambda_i} \tag{5}$$

Here $\boldsymbol{b^k} = [b_1^k, b_2^k, \dots, b_N^k]$ represents the temporal evolution of k^{th} dynamic modes.

Provided we're aware of the current state, we can infer the system's future behaviour incrementally using mapping function g and linear operator K, and the dynamic characteristic can be explored through the eigen decomposition of K, for example, to explore what are the main driving components in the evolution of control systems, what is the pattern of its growth, and so on.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

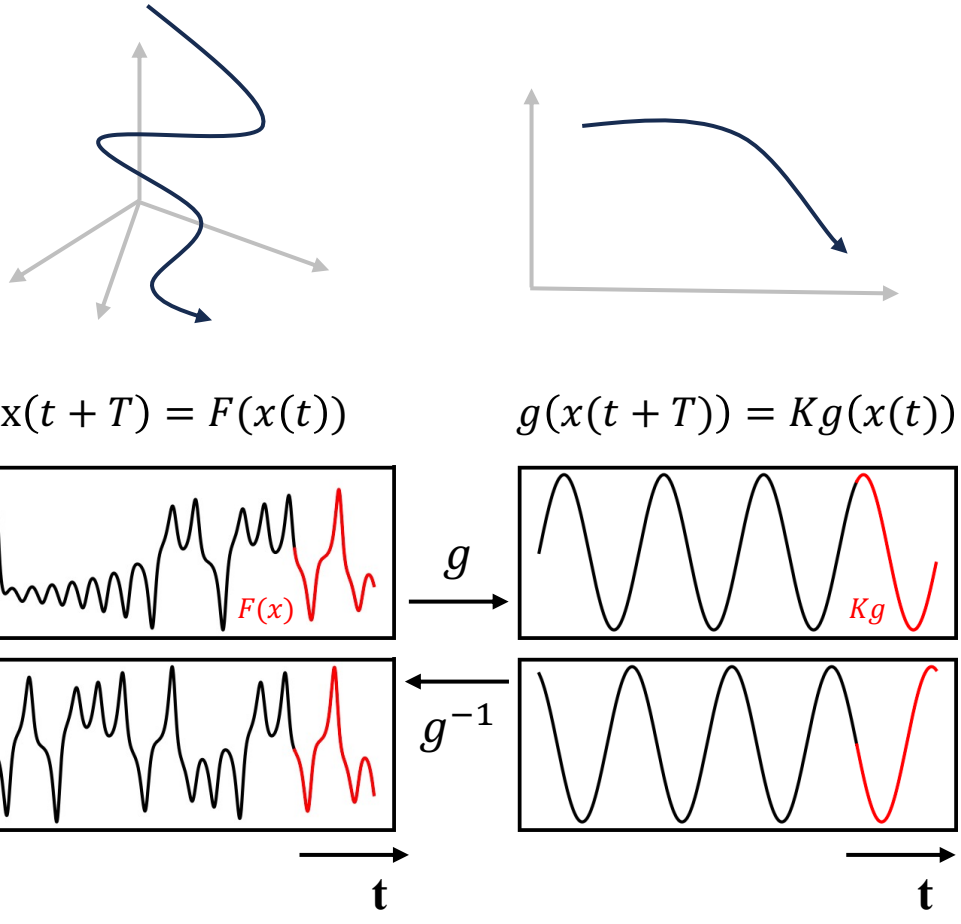


Figure 2: Koopman Theory.

2.2.2 Delay Embedding Theory

While Koopman theory best accommodates states or majority observations rich in dynamical system information, the shear stress in this context merely provides partial observation of the laboratory slip system. Here we introduce delay embedding theory to reconstruct the system behaviour. Delay embedding theory supposes that topological reconstruction of the attractors from the original high-dimensional system, also known as phase space reconstruction, can be performed using only the observed univariate long time series (Takens, 1981). We define $\boldsymbol{h}$ as the embedded variable, taking $H = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_i]$ as the input. The embedding process is described in Eq. (6) with the parameters, the embedded dimension d and delay time τ. The delay time τ usually takes 1 in most situations (Brunton et al., 2017).

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

$$\mathbf{H} = \begin{bmatrix} x_1 & x_2 & \cdots & x_i \\ x_{1+\tau} & x_{2+\tau} & \cdots & x_{i+\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1+(d-1)\tau} & x_{2+(d-1)\tau} & \cdots & x_{i+(d-1)\tau} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_i \\ x_2 & x_3 & \cdots & x_{i+\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_d & x_{d+1} & \cdots & x_{i+(d-1)} \end{bmatrix} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_i] \quad (6)$$

Here, we aim to utilize historical shear stress observations to evolve future states or to discern the relationship between historical shear stress $(x_{t-K}, x_{t-K+1}, \dots, x_t)$ and future states $(x_{t+1}, \dots, x_{t+L-1}, x_{t+L})$. Based on the delay embedding, the questions lie in determining the relationship from $[\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_i]$ to $[\mathbf{h}_{i+1}, \mathbf{h}_{i+2}, \dots, \mathbf{h}_{i+L}]$, which can be deconstructed into the mapping function g and operator K.

As shown in Figure 3, which represents the process of delayed embedding. To verify the retention of the original system's phase space topological relations following the delayed embedding, we often resort to the singular value decomposition to identify the system's three primary components, and use the corresponding singular vectors to open the space to represent the system evolution. Taking the Lorentz system as an example, Figure 3a represents the tensor space using the state variables of the original system. Observation yields merely a single scalar series, as depicted in Figure 3c. Figure 3b shows the first three singular vectors that result from the delayed embedding, which is diffeomorphic with the system represented by Figure 3a, i.e., they are considered to represent the same dynamical system behaviour. This approach has been shown to be effective in enhancing the feature dimensions from the data, and it has also been shown to be important in governing equation extraction (Bakarji et al., 2023) and modelling with Dynamic Mode Decomposition (DMD) (Avila and Mezic, 2020).

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

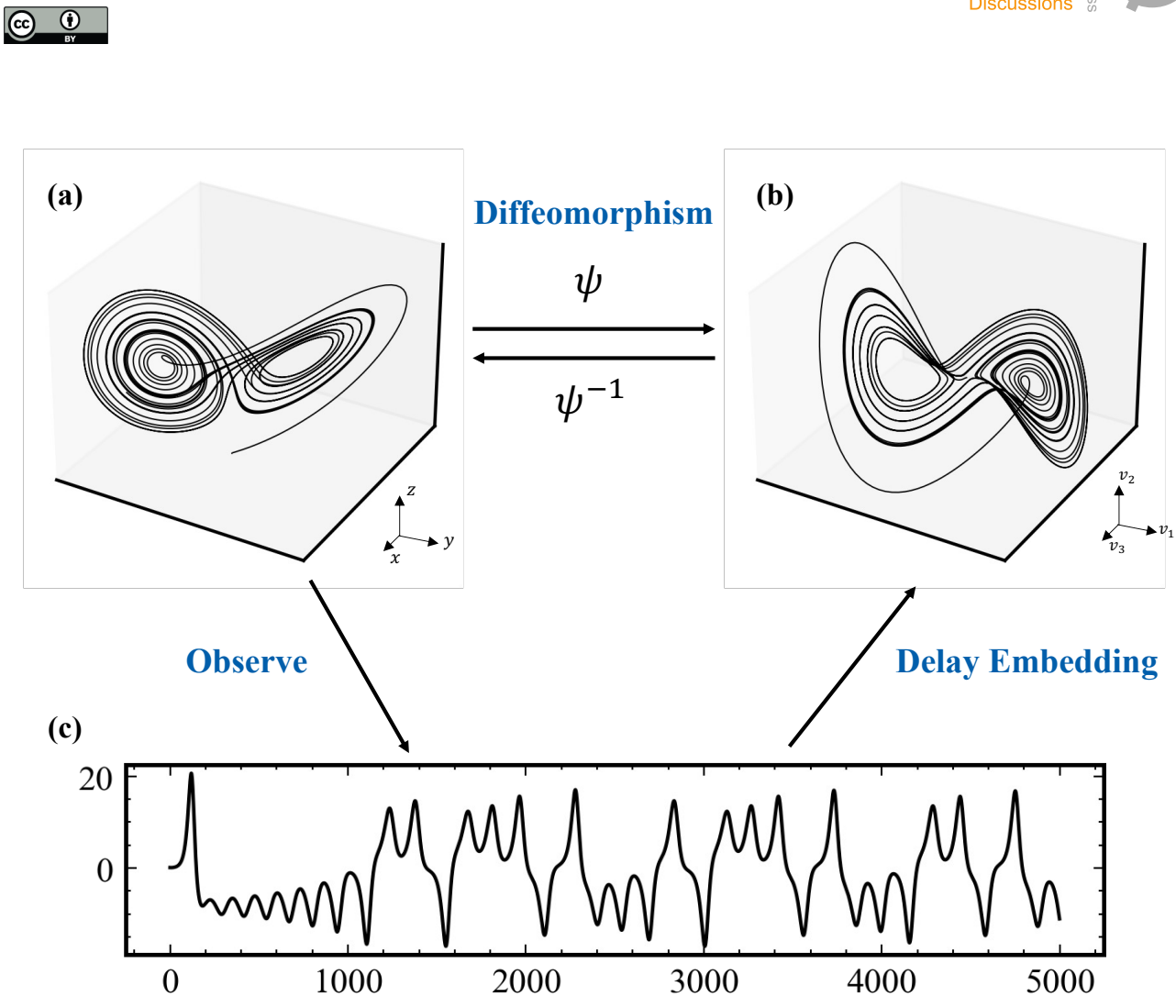


Figure 3: Delay embedding theory. (a) Representation of Lorenz system behavior using original state $[x, y, z]$. (b) Representation of Lorenz system behavior using singular vectors $[v_1, v_2, v_3]$ from SVD result of delay embedded observation x. (c) Single Observation of Lorenz system.

2.3 Architecture of Hankel Koopman Autoencoder (HKAE)

Here we propose the model named Hankel Koopman Autoencoder (HKAE), synthesizing the inductive bias of dynamical system theories and deep learning nonlinear fitting ability. This model encompasses three key modules (Figure 4):

(1) Delay Embedding Module: Employing delay embedding theory, the shear stress time series $(x_{t-d}, x_{t-d+1}, \ldots, x_t)$ are reconstructed in phase space to obtain their Hankel matrix $H = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_i]$ in this module.

(2) Mapping Learning Module: The powerful nonlinear fitting capabilities of deep learning have been demonstrated to effectively learn the mapping, approximating an optimal Koopman operator (Takeishi et al., 2017; Lusch et al., 2018; Azencot

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

et al., 2020) . Here, we utilize an encoder-decoder backbone incorporating a 3-layer Multi-layer Perceptron (MLP) to learn the mapping between phase space and Koopman invariant subspaces.

(3) Koopman Evolution Module: In this module, the Koopman operator is represented as a layer of neurons, consisting solely of weights and devoid of bias. Following the encoding process, the system is mapped into Koopman invariant subspaces, where the Koopman operator is applied to facilitate system evolution. Multi-step prediction is achieved through the iterative application of the same set of operators corresponding to the predefined prediction steps. The decoded evolution results remain in phase space. Subsequently, a re-embed process is applied to derive the predicted shear stress, implemented by selecting the final value of each evolved result.

The loss of this model includes 2 parts, reconstruction and evolution loss, shown in Eq. (7)-(9). The reconstruction loss is set to minimize the loss that occurred during the mapping process, while the evolution loss is to minimize the loss of linear evolution achieved by the Koopman operator.

$$\varepsilon_{reconstruct} = \frac{1}{2n}\sum_{i=1}^{n} \| h_i - \hat{h_i} \|_2^2 \tag{7}$$

$$\varepsilon_{evolution} = \frac{1}{2nh}\sum_{j=1}^{h}\sum_{i=1}^{n} \| h_{i+j} - \tilde{h}_{i+j} \|_2^2 \tag{8}$$

$$Loss = \varepsilon_{reconstruct} + \lambda\varepsilon_{evolution} \tag{9}$$

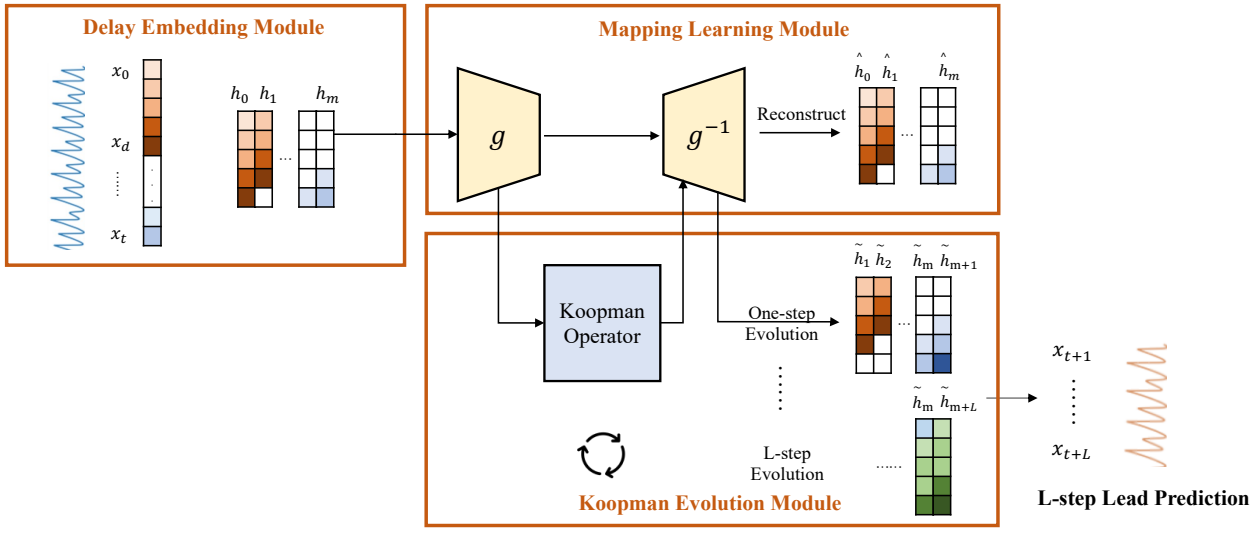


Figure 4: Architecture of HKAE.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

2.4 Time series prediction deep learning models in statistical perspective

2.4.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a widely used deep learning model for sequence modelling to address the challenge of vanishing and exploding gradients in long sequence data (Hochreiter and Schmidhuber, 1997). It has become a powerful tool in temporal modelling in earth sciences. The core of LSTM is the LSTM cell. A LSTM cell consists of a cell state and a set of gates (input, forget and output gates). The input gate determines the new input information to add to cell state while the forget gate choose the information to drop in the cell state. The output gate controls how cell state is mapped to the output. The structure of a LSTM cell is shown in Figure 5a. The equation expressions of three gates are shown as follows:

$$i_t = \sigma(\omega_i[h_{t-1}, x_t] + b_i) \quad (10)$$

$$f_t = \sigma\left(\omega_f[h_{t-1}, x_t] + b_f\right) \quad (11)$$

$$o_t = \sigma(\omega_o[h_{t-1}, x_t] + b_o) \quad (12)$$

2.4.2 Temporal Convolutional Network (TCN)

Temporal Convolutional Network (TCN) is a sequence modelling method inspired by convolutional operations widely used in the image processing (Bai et al., 2018). It has been regarded as the state-of-the-art model in laboratory fault flip modelling (Laurenti et al., 2022). The core of TCN is a series of sequence convolution and pooling blocks. In general, the features of input sequence are extracted through dilated convolution with sliding window. Pooling layer is adapted to dimension reduction. Weight norm, drop out are used to improve the robustness of model. The structure of TCN model and TCN block is shown in Figure 5.

2.4.3 Multi-layer Perceptron (MLP)

Except for two sequence modelling method, we also test the simple MLP model. Here we take the historical sequence as input feature and predict sequence as output, as shown in Figure 5. In this way, the MLP is equivalent to performing a 1*1 convolution directly in the time dimension, which is equivalent to a TCN model with dilation set to 1.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development
Discussions
Open Access
EGU

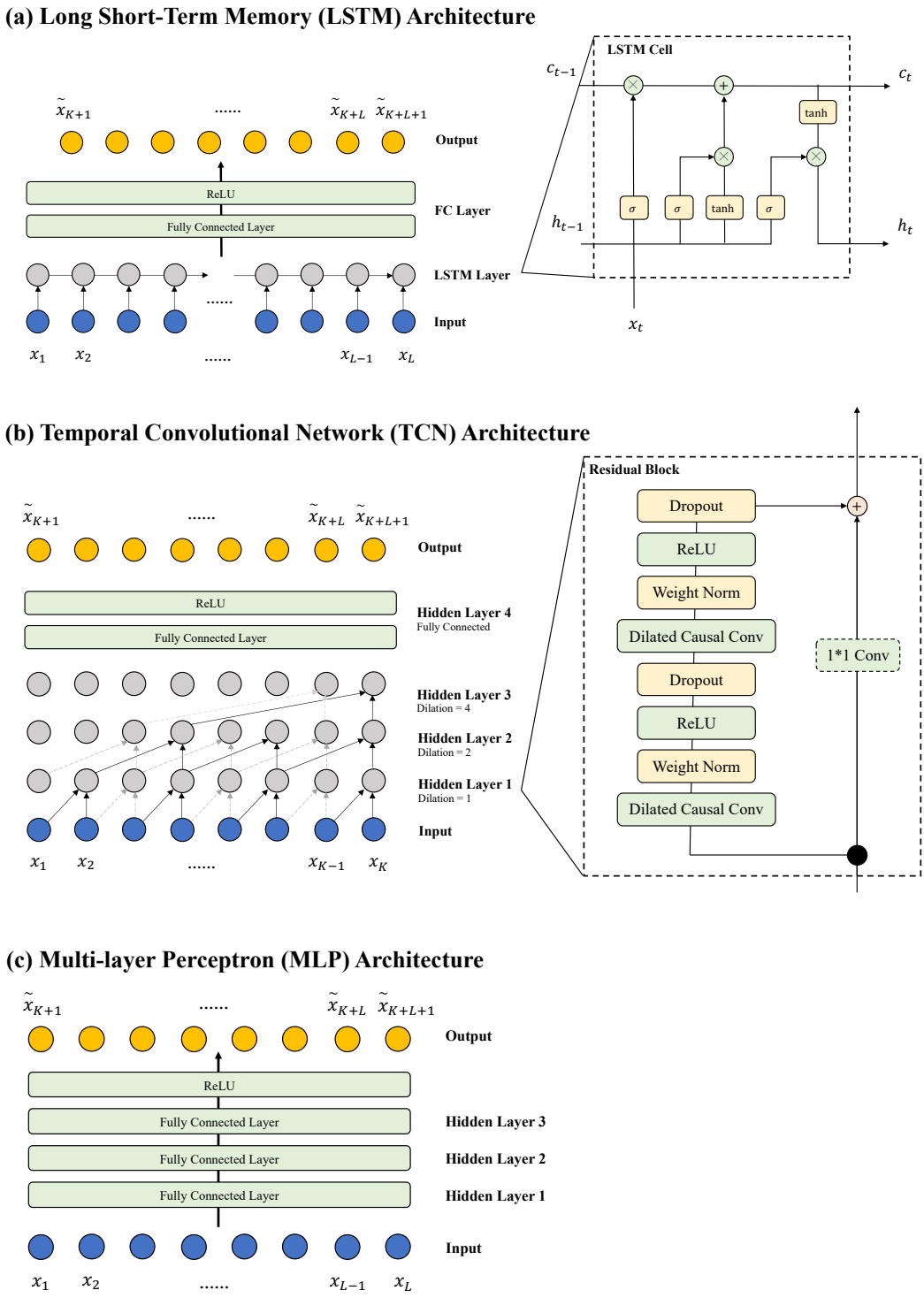


Figure 5: Architecture of comparative deep learning models

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

2.5 Model training and prediction

In our experiments, we utilized three datasets with distinct slip features and one numerical simulation result to assess the robustness of our model. We employed a unified time sampling rate of 0.1s. A ratio of 50%, 10%, and 40% was applied to each dataset for training, validation, and testing, respectively. The sliding window method was employed to generate data sequences for model training and inference. For the input dataset $[x_1, x_2 \ldots, x_N]$, we sample historical K steps to predict future L steps. The sliding window prediction can be presented by:

$$(x_i, \ldots, x_{i+K}) \rightarrow (\tilde{x}_{i+K+1}, \ldots, \tilde{x}_{i+K+L}) \tag{13}$$

Where $i = (1, \ldots, N - K - L + 1)$ represents number of times the window slides.

The HKAE model's hyperparameters are divided into three main categories, corresponding to the three model modules. The first set comprises the embedding dimension and delay time in the Delay Embedding Module. The optimal embedding dimension of 100 and delay time of 1 were confirmed by integrating the optimized parameter confirmation method proposed by (Cao, 1997) with the actual prediction results and metrics. The number of MLP layers in the Mapping Learning Module was set to 3, based on previous works and results. Lastly, the dimension of the Koopman Operator Module was set to 10, based on the data and performance. We adopted a batch size of 64, L2 loss as the loss function, and Adam as the optimization algorithm. Weight decay and gradient crip skills are adapted to improve the performance. Given the single-step evolutionary nature of the Koopman operator, in order to keep it robust on future leading predictions by learning the dynamics, we use a multi-step trick on error estimation during model training (Eq. (8)).

For the comparative models, LSTM and MLP obey the similar setting in Laurenti's work. The detailed key parameters setting is listed in Table 1. To achieve multi-step prediction, here we add a Fully Connected Layer as decoder in our LSTM, TCN and MLP models, to map the extracted features to prediction window. Temporal Bundling (TB) is adapted to reduce the rate of error propagation through reducing the model calls (Brandstetter et al., 2022).

Table 1: Key parameters adapted in the experiments for different models.

LSTM	TCN	MLP	HKAE
Input_size: 1	Input_size: 1	Input_size: K	Embedded Dimension: K
Num_layer: 3	Num_channel: [64, 256]	Hidden_size: [64, 256]	Delay Time: 1
Hidden_size: 128	Kernel_size: 3	Output_size: L	Bottleneck: 10
Output_size: L	Drop_out: 0.1		Hidden_size: 16*a, a = 5
	Output_size: L		

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

2.6 Evaluation metrics

We evaluate the prediction result mainly using R^2 and Rooted Mean Square Error (RMSE) metrics. As Eq. (13) shows, the predictions are conducted in a certain time window, and the window keep sliding to the end of this lead situation. Then the lead time grows to adapt new round of sliding predictions. Here we calculate the difference between the predicted results and the ground truth corresponding to each prediction leading step case.

For prediction lead steps $j = (1, \dots, L)$, we compute the R_j^2 and $RMSE_j$ as follows:

$$R_j^2 = 1 - \frac{\sum_{i=0}^{N-K-L+1} (\widetilde{x_i^j} - x_i^j)^2}{\sum_{i=0}^{N-K-L+1} (\overline{x^j} - x_i^j)^2} \tag{14}$$

$$RMSE_j = \sqrt{\frac{1}{N-K-L+1} \sum_{i=0}^{N-K-L+1} (\widetilde{x_i^j} - x_i^j)^2} \tag{15}$$

Where j is the prediction lead step, i is the number of slide windows. R_j^2 and $RMSE_j$ represents the R^2 and $RMSE$ for j^{th} lead predictions. $\tilde{x}$, $\bar{x}$, x are respectively the predictions, mean and ground truth of x.

To assess the efficacy of temporal dynamics modelling, we utilized mean period statistics commonly employed in laboratory earthquake system analysis (Veedu et al., 2020). We pick the peaks of shear stress and compute the interval of peaks ΔT as the period of shear stress variation respectively for predictions and ground truth. The R^2 metric is used to test whether the stress variation period of predictions fit well. The slip interval ΔT is computed based on the algorithm from (Gualandi et al., 2023), according to the peaks of shear stress series. The peaks are found based on *find_peak* function from *scipy*.

$$\Delta T_i = \Delta t * \left(slip_{center_{index_{i+1}}} - slip_{center_{index_i}} \right) \tag{16}$$

$$slip_{center_{index_i}} = stress_{top_{index_i}} + argmin\left(\left| x_{stress_{top_{index_i}}}, \dots, x_{stress_{bottom_{index_{i+1}}}} \right| \right) \tag{17}$$

The $slip\_center\_index$ represents the central moment of a slip behaviour, derived from $stress\_top\_index$, the positive peaks of shear stress series, and $stress\_bottom\_index$, the negative peaks of shear stress series.

Then we compute the $R_{slip_j}^2$ for modelled slip intervals for different lead steps:

$$R_{slip_j}^2 = 1 - \sum_{i=0}^{N_{slip}} (\widetilde{\Delta T_i^j} - \frac{\Delta T_i^j)^2}{\sum_{i=0}^{N_{slip}} (\overline{\Delta T^j} - \Delta T_i^j)^2} \tag{18}$$

Where j is the prediction lead step, i is the number of slips in the j^{th} predictions. $\widetilde{\Delta T}$, $\overline{\Delta T}$, ΔT are respectively the predictions, mean and ground truth of ΔT.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

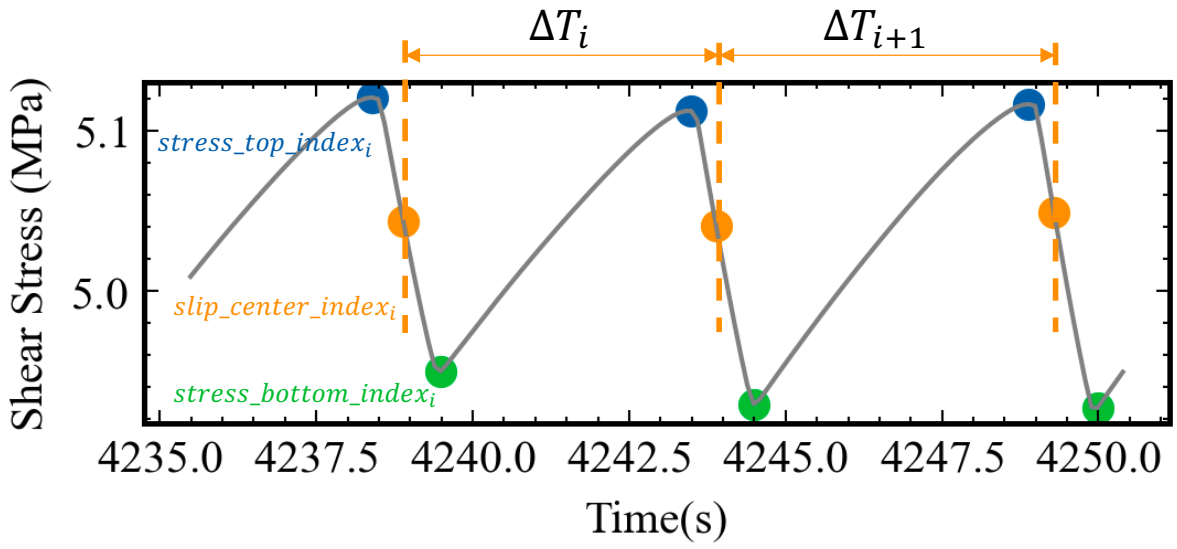


Figure 6: Calculation of slip interval ΔT.

3 Results

3.1 Evaluation of dynamic modelling process

We conduct multi-step prediction experiments on three types of shear stress records from laboratory seismic experiments, each exhibiting different rupture characteristics: fast rupture (Exp. 4581), alternating fast and slow rupture (Exp. 4679), and slow slip (Exp. 5179). These experiments illustrate the robustness of our model across various slip modes. To validate the effectiveness of the dynamic modelling, here we employ Singular Value Decomposition (SVD) and Eigen Decomposition (ED) to check the data status during the model processing.

Firstly, we adapt SVD on the result after delay embedding, utilizing three dominant right singular vector modes as coordinates to represent the system's behaviour (Figure 7b). Using simple ODE simulation data as an example, it is evident that the system evolves along a “two-cycle-like” stable trajectory. Each cycle represents a slip process with varying stress drop intensities, and the system alternates between these two states. The attractors of Exp. 4581 and 5198 do not exhibit concentrated traces but rather more complex trajectories, indicative of the quasi-periodic behaviour of slow and fast slips. For the Exp. 4679, the attractor displays a more significantly disordered pattern, yet cycles around two main traces, illustrating the bifurcation process (Veedu et al., 2020). The structure of the attractors is preserved following the nonlinear mapping by the MLP encoder, as illustrated in Figure 7c, indicating that the system dynamics are retained post-mapping from the observed space to the Koopman subspace.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024


Geoscientific Model Development Discussions

We apply ED to the learned Koopman operator (Eq. (4)) and get the decomposed complex eigen value (Table 2), representing various dynamic modes characterized by distinct amplitudes and periods. We illustrate these modes within a unit circle, which represents the attraction set. As Figure 7d shows, not all modes are near the unit circle. This may be caused by a single observation dimension, or it may be caused by an unstable stick-slip system where the data does not exactly follow the attractor trajectory during the experiment (Jasperson et al., 2021). The dynamic modes represented by eigenvalues are discriminated by a narrow threshold (0.01) of the distance to unit circle. Eigenvalues that exceed the threshold are considered unstable modes if they are large, stable modes if they are small, and neutral if they are within the threshold.

All three experiments decomposed and obtained stable modes (red points, modes close to the unit circle,). Modes with about 10s and 9s for the quasi-periodic fast slip in Exp. 4581 and slow slip in Exp. 5198. While dual modes with about 6s and 10s for the bifurcations of slips in Exp. 4679 are also modelled successfully. There are modes with amplitudes close to 1 and zero frequency in Exp. 4581 and 5198, but the amplitude of the zero-frequency mode in Exp. 4679 with slip switch is much lower, which reflects the stronger time-varying component in this slip system.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

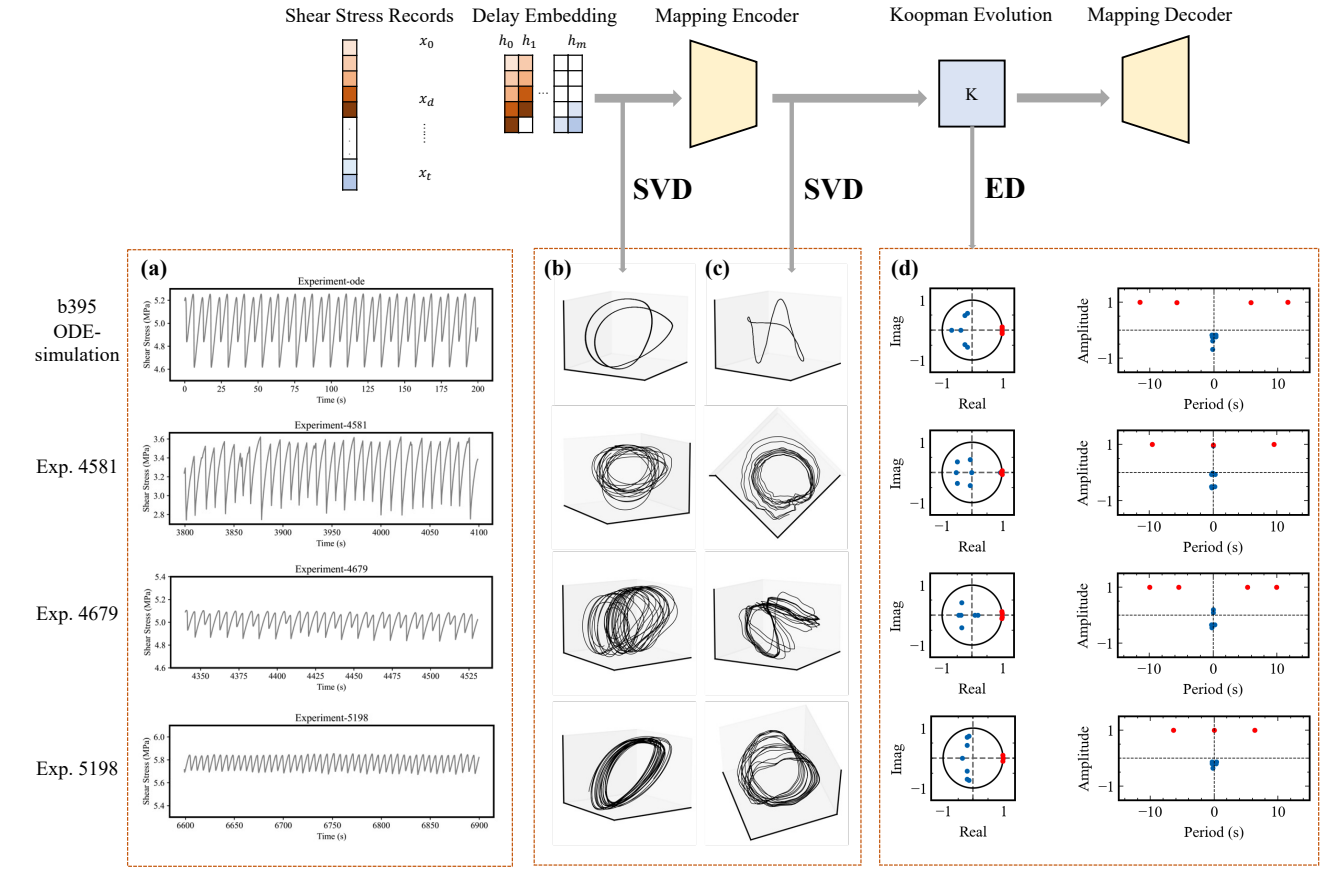


Figure 7: Dynamic modelling steps and interpretation with HKAE. The rows represent the different modelled experimental data. The columns represent the dynamic modelling of the data through the different steps of the model. (a) Shear stress time series from Ordinary Differentiation Equation (ODE) Simulation and three lab experiments. (b)-(c) Attractor structure after delay embedding and HKAE-encoder, using the coordinates expanded by the first three right singular vector modes (Text S3 in Supplementary Information S1). The black line illustrates the system evolution under three-dimensional projections. (d) Eigen Decomposition of learned Koopman Operator. Left in (d) illustrates the real and image part of the complex eigen value. Right in (d) convert the value to amplitude and period format.

Table 2: Learned Koopman operator eigen modes. Only show the eigenvalue with positive periods due to the conjugation.

	Exp. 4679		Exp. 4581		Exp. 5198	
# Eigenvalue	Period(s)	Amplitude	Period(s)	Amplitude	Period(s)	Amplitude
1	9.970778	0.999401	9.548719	1.001883	6.388905	0.994914
2	5.378387	0.993912	0.294214	-0.071703	0.349487	-0.13397
3	0.258224	-0.34971	0.239758	-0.500562	0.329522	-0.20161
4	inf	0.091451	inf	0.967017	0.276048	-0.195414
5	inf	0.193768	inf	0.94443	inf	1.00072

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024


Geoscientific Model Development Discussions

3.2 Evaluation of different slip behaviours with different models

We test two different experimental settings to validate the modelling and prediction capabilities pertaining to slip behaviours. The first one focusses on the stress variation. Employing the historical 100 steps (spanning 10s, which encompasses a complete slip cycle of stress raise and release), we predict the subsequent 30 steps of shear stress (lasting 3s, long enough to include a stress raise or drop process). Then we extend the input to 20s and prediction horizon to 10s, to test the skills on seismic cycle modelling. We compare our results under 2 experiment settings with 3 comparative deep learning time series prediction methods.

Figure 8 illustrates the evaluation metrics varies with prediction lead steps for the experiment data. First, the slow slip represented by Exp. 5198, with its relatively gentle stress changes and quasi-periodic characteristics, is not too difficult from the perspective of time series prediction, so both LSTM and TCN outperform HKAE in terms of R^2 and RMSE. But from the results of our slip interval ΔT modelling metrics, HKAE clearly shows better results. On the other hand, for the other two data with more complex slip behaviour, Exp. 4679 with alternating fast and slow slips and Exp. 4581 with predominantly fast ruptures, the traditional deep learning methods also appear to have poorer statistical metrics than the HKAE. While it can be observed that the R^2 of HKAE keeps a steady trend although the initial prediction gets a higher miss, showing the ability of dynamic extraction and modelling. And from the results of slip interval modelling, on the two experiments with more slip characteristics, the HKAE also shows robust results with the growth of the prediction lead.

In addition, we can find that in the lead prediction of the three experiments, the metric will decay more rapidly in accuracy before 1s lead, while the decay slows down after 1s, especially for the traditional machine learning method. Under the "20s-10s" experimental setting (Figure 8a), we believe that the lead prediction can cover 1-2 complete seismic cycles (Laurenti et al., 2022), at which point the statistical period of the slip interval will be longer and more representative. Compared with the experimental settings of "10s-3s" (Figure 8b), predicting the results for the next 10s, the slip interval scores of the HKAE predictions get higher than other methods, which reflects the superiority of HKAE in modeling seismic dynamics.

Furthermore, unlike the decrease in accuracy with an increasing number of leading steps evident in stress value predictions, the metrics for slip intervals do not demonstrate a steady decline. This could be attributed to the periodicity of the slip activities. Moreover, the capacity of the HKAE model to model slip behaviour more robustly indicates its strength in capturing the long-term trend in the data.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

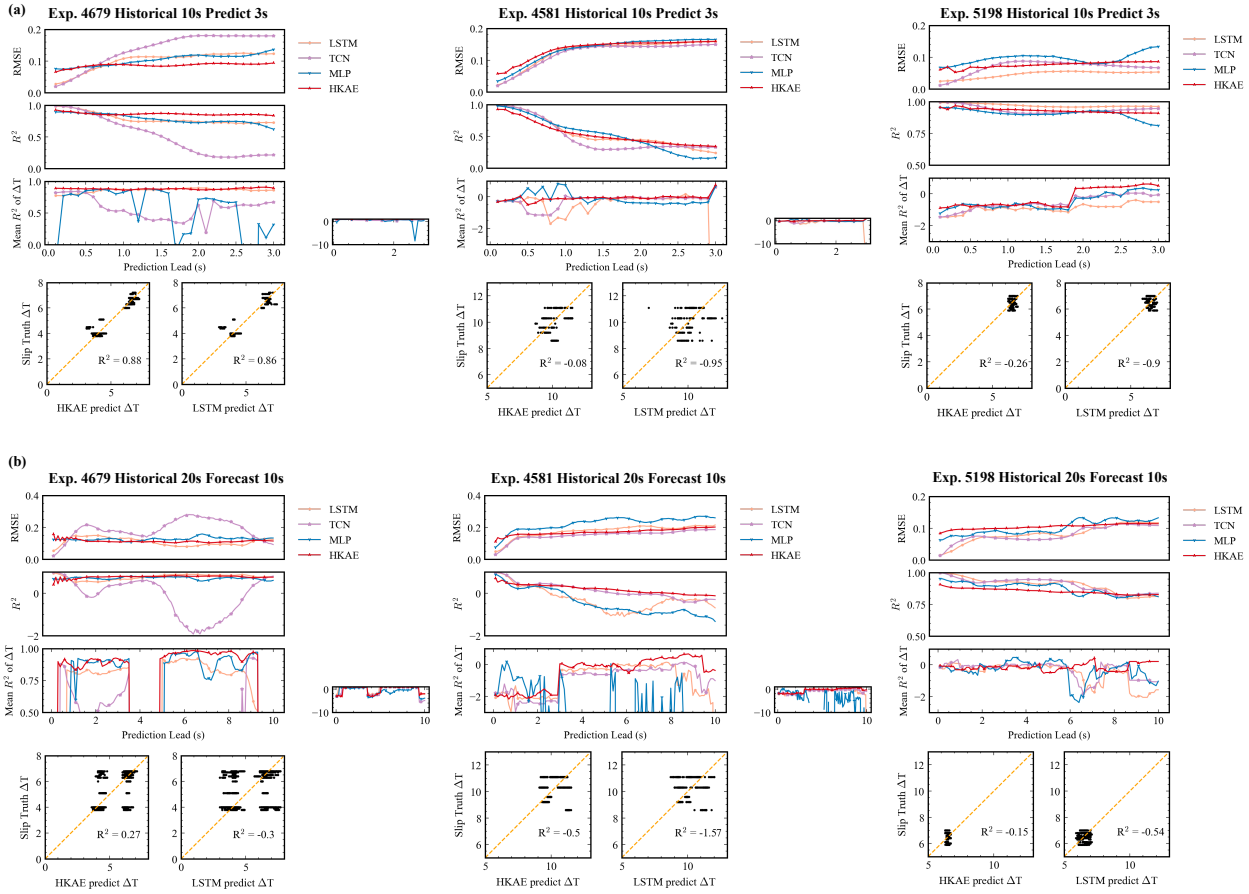


Figure 8: Lead prediction experiments. (a) 3s lead prediction using historical 10s shear stress. (b) 10s lead prediction using historical 20s shear stress.

Besides the evaluation metrics of multi-step lead predictions, we further validate the predicted shear stress under the setting of "10s-3s" to check the accuracy of the predicted stress. Figure 9 illustrates the final lead predictions and future prediction for the test set from three experiments. We pick the final lead predictions to shows the stability of lead predictions. We choose two sections for each experiment where the stress indicates typical stress raise-drop during slip behaviours. Our primary focus is on the model's accuracy regarding the timing of stress changes and the magnitude of stress drops.

Regarding the timing of slip occurrences, Exp. 5198 intuitively exhibits the best prediction, as demonstrated in Figure 9c, f. However, Exp. 4581 shows a less accurate prediction, with a notable mismatch in the slip cycle and an "early release" of stresses (Figure 9a, d). This discrepancy might be attributed to the absence of creep information, as the rapid stress rupture in Exp. 4581 occurs immediately at rupture onset, in contrast to the other two experiments where accelerated stress attenuation precedes rapid stress drop. Despite successfully modelling the rate of stress release, Exp. 4581 tends to prematurely estimate the rupture onset, unlike Exp. 4679, which has a comparable stress release rate. Exp. 4679 (Figure 9b, e) delivers the most

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

fluctuant prediction results. During the stress raise phase, the model struggles to accurately predict the rate of increase or the correct trajectory. This might be due to the strong nonlinear components in the system, hindering the model's ability to learn the mapping and operator from solely shear stress inputs. Alternatively, the presence of time-variant dynamics or multiple invariant subspaces in the theoretical linear evolving space could challenge the single operator's capacity to depict evolution accurately (Lan and Mezić, 2013). In terms of modelling the magnitude of stress drop during slips, all three experiments tend to underestimate stress drop values when making lead predictions. Notably, Exp. 4679 and Exp. 5198, which include slow slips, provide more accurate predictions of maximum stress values during slips, though they tend to overestimate the minimum value after stress release. Exp. 4581, on the other hand, predicts a narrower range of stress variation, underestimating both the maximum and minimum values. These prediction details suggest that the error in stress change values is also linked to the misjudgement of slip onset timing. If a stress-drop or stress-raise is predicted prematurely, the stress conversion occurs before reaching the threshold value. Comparing to the prediction result of LSTM, HKAE shows more accurate modeling result. In exp. 4581, the HKAE predictions indicate similar stress drop rate with ground truth (Figure 9a), while LSTM results predict a slower rate (Figure 9d). For exp. 4679, the HKAE predict a more accurate moment of stress release.

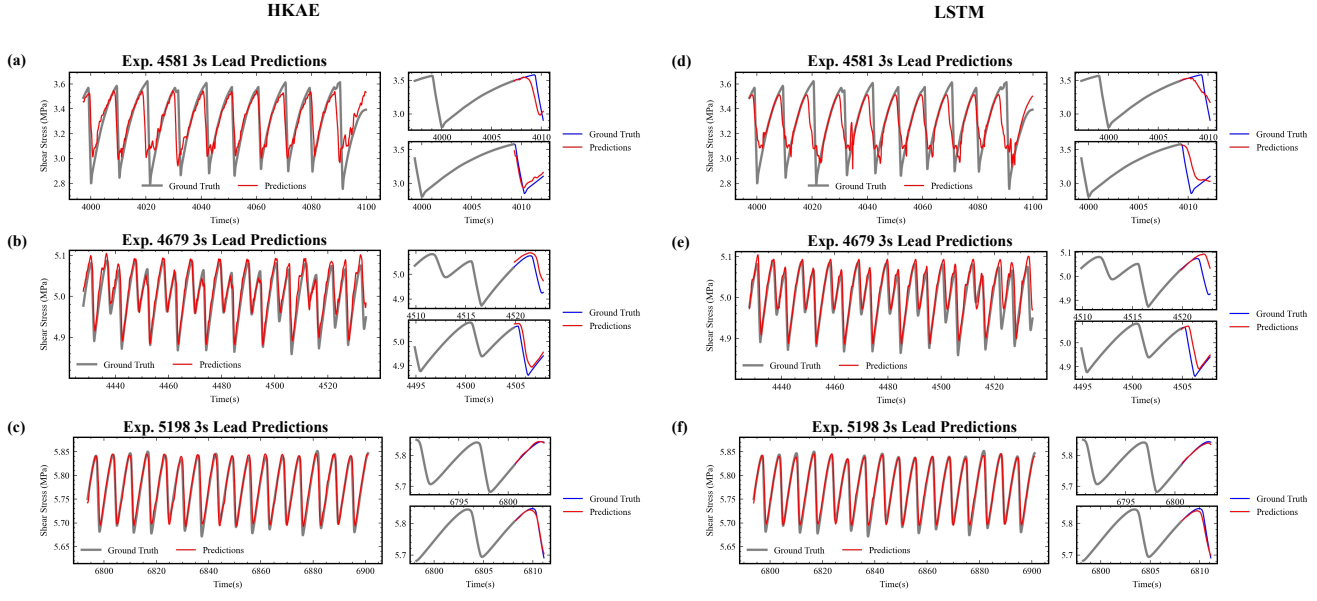
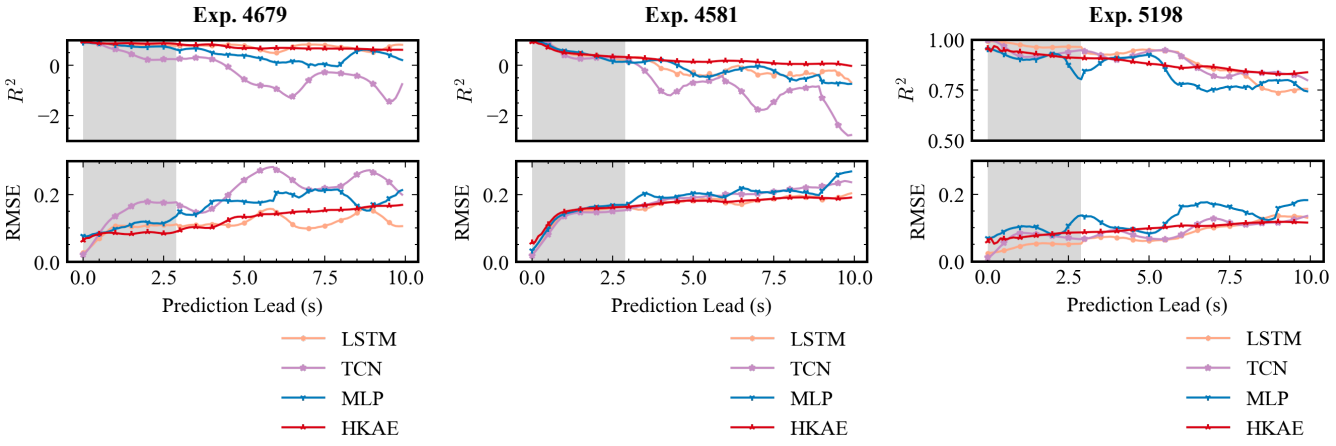


Figure 9: Future 3 seconds multi-step prediction of shear stress in 3 experiments, taking the historical 10s shear stress as input. (a)-(c) HKAE predicted results. (d)-(f) LSTM predicted results.

3.3 Zero-shot prediction out of the model trained horizon

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions



335 **Figure 10: Zero-shot prediction results of models out of the model trained horizon on different experiments. Gray section indicates the model trained prediction horizon.**

To further validate our model's efficacy in modelling the dynamics of the slip process, we conducted a zero-shot test. Specifically, we incorporate 10s of historical stress data during the training phase and predict 3s of stress change in the future, but during the inference phase, we extended the prediction horizon to 10s. This means that the model was only trained to 340 perform the prediction for the next 3s while the data samples recorded beyond the initial 3s, which weren't included during the training phase, are now required to anticipate changes over the subsequent 10s. This necessitates greater generalization capability from the model. We executed this manner of testing across all three experimental datasets. The process is depicted in Figure 10, where the grey area is the predicted horizon on which the model was trained. HKAE appears to maintain a stable prediction score for the unplanned following 7s. In Exp 5198, HKAE starts to outperform the other models in the leading 345 prediction after 6s. And in Exp. 4679 and Exp. 4581, two more challenging experiments, HKAE maintains the leading prediction, and all other prediction methods show more obvious accuracy decay.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Geoscientific Model Development Discussions

4 Discussion

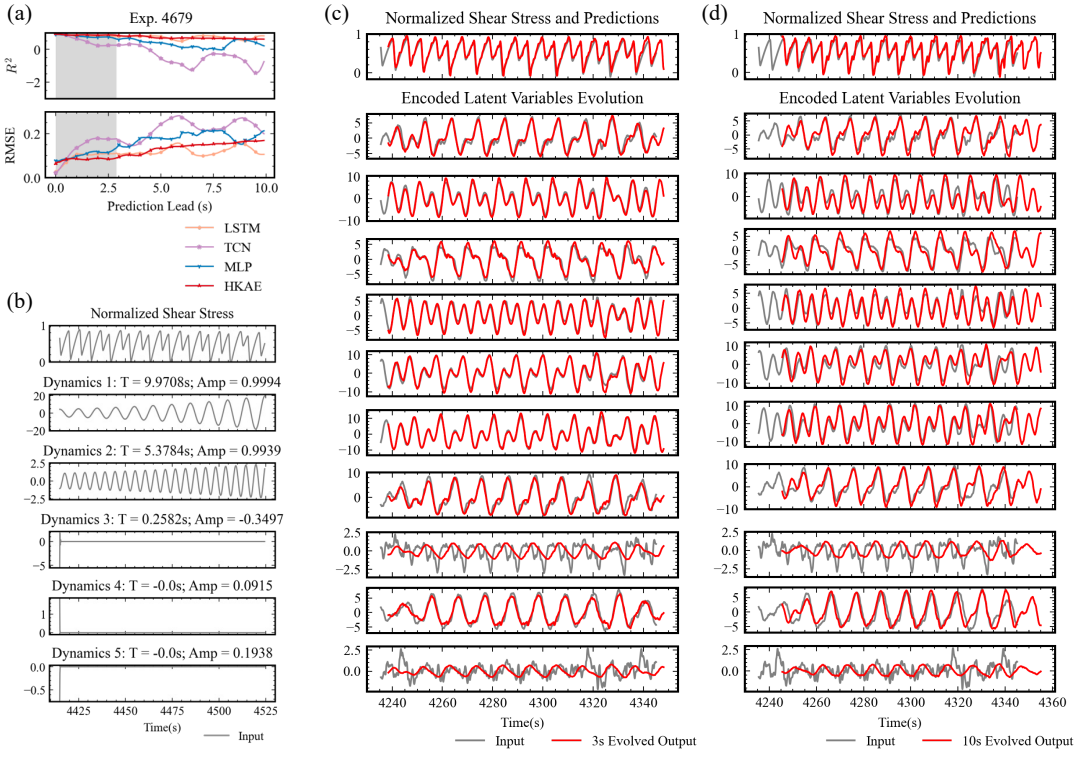


Figure 11: Zero-shot test for Exp. 4581. (a) evaluation metrics. (b) HKAE modelled dynamic modes. (c) 3s lead evolution of encoded latent variables. (d)10s lead evolution of encoded latent variables.

Modelling and predicting the behaviour of fault slip is crucial for understanding natural earthquakes. Studying natural earthquakes still presents various challenges like indirect observation, sampling history of shorter fault activity cycles etc. (Herrera et al., 2022), while laboratory settings provide new sights in an easier, more controllable and observable manner. Machine learning methods have been effective in accurately predicting instantaneous slip behaviour based on near-term acoustic emissions (Rouet-Leduc et al., 2017; Shokouhi et al., 2021; Borate et al., 2023). However, attempts to forecast future behaviour have encountered temporal limits due to the high nonlinearity of the laboratory fault slip system.

To address these questions, informed by dynamical system theories, we pioneered a dynamic informed method HKAE, to predict the future shear stress of laboratory fault slips. The HKAE model is designed based on the delay embedding theory and Koopman theory, leverages the nonlinear fitting capabilities of neural networks and the system perspective of dynamic theories. The superior performance of HKAE in various timeseries prediction methods and different slip mode experiments may come out that the HKAE model is able to recognize the main and stable dynamics of laboratory fault slips, rather than define it as a statistical time series prediction task.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024


Geoscientific Model Development Discussions

Here we further discuss the dynamic modelling skills of HKAE. We implemented the zero-shot approach beyond the model's trained horizon, as outlined in Section 3.3. As indicated by (Laurenti et al., 2022), while conventional deep learning methods falter in untrained prediction horizons, particularly for fast slips as in Exp. 4581 and slip switch in Exp. 4679, HKAE maintains a relatively steady trend. We carry out additional dynamic analysis through the latent variable evolution and decomposed Koopman dynamic modes.

Figure 11 illustrates the dynamics unravelled from the learned Koopman Operator. Two identified stable dynamics demonstrate steady oscillation, whereas the other three appear relatively disordered. The latent variables (Figure 11c-d) represent the components ready to evolve linearly after the mapping of encoder. Evolving these components for 3s using the Koopman operator reveals that the stabilized oscillatory variables can be estimated relatively well (Figure 11c). Upon expanding the evolution window to 10 seconds, these variables continue to evolve relatively stably (Figure 11d), which means the learned linear operator with two stable dynamics is able to generally model the slip process. However, the components that are insufficiently linearized result in deteriorating predictions over time. These disordered signals might stem from inadequate mapping; however, they more likely originate from the meta-stable and time-varying dynamics of the laboratory fault slip system (Jasperson et al., 2021).

Significantly, the model anticipates future shear stress without requiring any additional information or labels. Considering the laboratory environment and the nature of the earthquake system, integrating external data such as historical acoustic emissions or other measurable laboratory observations as forcing factors could enhance the generalizability of the model. Implementing the HAVOK framework (Brunton et al., 2017) and local dynamic modelling strategy (Liu et al., 2023) might further enhance the predictability of the laboratory rupture system with continuous spectrum.

In regard to the field application, we propose that applying this approach to long-term or modelling of seismic activity would be of great interest because of the ability to learn information about the dynamics of the system behind it from fewer samples of data. The ability to model dynamical processes as mappings and operators with a small amount of data would be more advantageous for both future long-term predictions and historical activity reconstructions, which are both important questions in seismic research. As for short-term modelling or alert, we believe that HKAE should also perform well on a continuous proxy for seismicity (Tong et al., 2023).

5 Conclusion

Drawing upon delay embedding and Koopman theories, we have proposed a dynamic informed machine learning method, the Hankel Koopman Auto-Encoder (HKAE), to achieve robust future predictions of complex laboratory earthquake systems. This model demonstrates superior performance in terms of prediction accuracy and system dynamic modelling compared to other common deep learning time series prediction methods. To the best of our knowledge, this is the first instance of predicting laboratory fault shear stress from a dynamical system perspective. Furthermore, HKAE highlights the potential for simplifying

and decomposing complex geophysical systems with a data-driven approach but combining a dynamical system perspective. It also delivers a new approach for modelling, predicting and understanding for the complex geophysical system, especially for those with partial and sparse observations, showing great potential in seismicity of fault physical mechanism monitoring and modelling.

400

Author contribution

Enjiang Yue: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Formal analysis. **Mengjiao Qin:** Writing – review & editing, Conceptualization, Supervision. **Linshu Hu:** Methodology, Writing – review & editing. **Sensen Wu:** Supervision, Writing – review & editing. **Zhenhong Du:** Supervision, Writing – review & editing.

405

Acknowledgments

This research was supported by the National Key Research and Development Program of China (grant 2021YFB3900900), the National Natural Science Foundation of China (42306213) and the Deep-time Digital Earth (DDE) Big Science Program. The manuscript was polished by GPT-4 (https://openai.com/gpt-4).

410

Competing interests

The authors declare that they have no conflict of interest.

Code and data availability

415 The laboratory shear stress data is obtained from https://github.com/lauralaurenti/DNN-earthquake-prediction-forecasting and http://psudata.s3-website.us-east-2.amazonaws.com/ (Last access at 2024-03-21). The ODE simulation data and code is obtained from Open Science Framework (OSF) (at DOI 10.17605/OSF.IO/9DQH7) and https://github.com/Geolandi/labquakesde (Last access at 2024-03-21). The code to run and analyze the results of the experiments are available online (https://zenodo.org/records/10846361).

420

References

Avila, AllanM. and Mezic, I.: Data-driven analysis and forecasting of highway traffic dynamics., Nature Communications, https://doi.org/10.1038/s41467-020-15582-5, 2020.

Azencot, O., Erichson, N. B., Lin, V., and Mahoney, MichaelW.: Forecasting Sequential Data using Consistent Koopman
425 Autoencoders, International Conference on Machine Learning. PMLR, 2020.

Bai, S., Kolter, J. Z., and Koltun, V.: An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling, arXiv: Learning, 2018.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Bakarji, J., Champion, K., Kutz, J. N., and Brunton, StevenL.: Discovering Governing Equations from Partial Measurements with Deep Delay Autoencoders, Proceedings of the Royal Society A, https://doi.org/10.1098/rspa.2023.0422, 2023

Borate, P., Rivière, J., Marone, C., Mali, A., Kifer, D., and Shokouhi, P.: Using a physics-informed neural network and fault zone acoustic monitoring to predict lab earthquakes, Nature Communications, https://doi.org/10.1038/s41467-023-39377-6, 2023.

Brandstetter, J., Worrall, D., and Welling, M.: Message Passing Neural PDE Solvers, arXiv: Learning, 2022.

Brunton, S. L., Brunton, B. W., Proctor, J. L., Kaiser, E., and Kutz, J. N.: Chaos as an Intermittently Forced Linear System, Nature Communications, https://doi.org/10.1038/s41467-017-00030-8, 2017.

Brunton, S. L., Noack, B. R., and Koumoutsakos, P.: Machine Learning for Fluid Mechanics, Annual Review of Fluid Mechanics, 477–508, https://doi.org/10.1146/annurev-fluid-010719-060214, 2020.

Brunton, S. L., Budišić, M., Kaiser, E., and Kutz, J. N.: Modern Koopman Theory for Dynamical Systems., SIAM Review, 229–340, https://doi.org/10.1137/21m1401243, 2022.

Cao, L.: Practical method for determining the minimum embedding dimension of a scalar time series, Physica D: Nonlinear Phenomena, 43–50, https://doi.org/10.1016/s0167-2789(97)00118-8, 1997.

Degen, D., Caviedes Voullième, D., Buiter, S., Hendriks Franssen, H.-J., Vereecken, H., González-Nicolás, A., and Wellmann, F.: Perspectives of Physics-Based Machine Learning for Geoscientific Applications Governed by Partial Differential Equations, https://doi.org/10.5194/gmd-2022-309, 2023.

Franzke, C., Gugole, F., and Juricke, S.: Systematic multi-scale decomposition of ocean variability using machine learning, Chaos: An Interdisciplinary Journal of Nonlinear Science, https://doi.org/10.1063/5.0090064, 2022.

Froyland, G., Giannakis, D., Lintner, BenjaminR., Pike, M., and Slawinska, J.: Spectral analysis of climate dynamics with operator-theoretic approaches, Nature Communications, 2021.

Gualandi, A., Faranda, D., Marone, C., Cocco, M., Mengaldo, G., and Bendick, R.: Deterministic and stochastic chaos characterize laboratory earthquakes, Earth and Planetary Science Letters, https://doi.org/10.1016/j.epsl.2023.117995, 2022.

Hammoud, M. A. E. R., Titi, E. S., Hoteit, I., and Knio, O.: CDAnet: A Physics-Informed Deep Neural Network for Downscaling Fluid Flows, Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2022ms003051, 2022.

Herrera, V., Rossello, E., Orgeira, M., Arioni, L., Soon, W., Velasco, G., Cruz, L., Zúñiga, E., and Vera, C.: Long-Term Forecasting of Strong Earthquakes in North America, South America, Japan, Southern China and Northern India With Machine Learning, Frontiers in Earth Science, https://doi.org/10.3389/feart.2022.905792, 2022.

Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Computation, 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735, 1997.

Hulbert, C., Rouet-Leduc, B., Johnson, P. A., Ren, C. X., Rivière, J., Bolton, D. C., and Marone, C.: Similarity of fast and slow earthquakes illuminated by machine learning, Nature Geoscience, 69–74, https://doi.org/10.1038/s41561-018-0272-8, 2019.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Jasperson, H., Bolton, DavidC., Johnson, PaulA., Guyer, RobertA., Marone, C., and Hoop, MaartenV. de: Attention network forecasts time-to-failure in laboratory shear experiments, Journal of Geophysical Research: Solid Earth, https://doi.org/10.1029/2021JB022195, 2021.

Johnson, P. A., Rouet-Leduc, B., Pyrak-Nolte, L. J., Beroza, G. C., Marone, C. J., Hulbert, C., Howard, A., Singer, P., Gordeev, D., Karaflos, D., Levinson, C. J., Pfeiffer, P., Puk, K. M., and Reade, W.: Laboratory earthquake forecasting: A machine learning competition, Proceedings of the National Academy of Sciences, https://doi.org/10.1073/pnas.2011362118, 2021.

Karimpouli, S., Caus, D., Grover, H., Martínez-Garzón, P., Bohnhoff, M., Beroza, G., Dresen, G., Goebel, T., Weigel, T., Kwiatek, G., and Bendick, R.: Explainable machine learning for labquake prediction using catalog-driven features, Earth and Planetary Science Letters, https://doi.org/10.1016/j.epsl.2023.118383, 2023.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L.: Physics-informed machine learning, Nature Reviews Physics, 422–440, https://doi.org/10.1038/s42254-021-00314-5, 2021.

Koopman, B. O.: Hamiltonian Systems and Transformation in Hilbert Space, Proceedings of the National Academy of Sciences, 315–318, https://doi.org/10.1073/pnas.17.5.315, 1931.

Lan, Y. and Mezić, I.: Linearization in the large of nonlinear systems and Koopman operator spectrum, Physica D: Nonlinear Phenomena, 42–53, https://doi.org/10.1016/j.physd.2012.08.017, 2013.

Laurenti, L., Tinti, E., Galasso, F., Franco, L., and Marone, C.: Deep learning for laboratory earthquake prediction and autoregressive forecasting of fault zone stress, Earth and Planetary Science Letters, https://doi.org/10.1016/j.epsl.2022.117825,2022.

Li, L., Zuo, H., Chen, B., Yang, G., Ma, M., and Longxiang, D.: Revealing the Dynamical Transition of Anisotropy Behind the HOST by Koopman Analysis, Geophysical Research Letters, https://doi.org/10.1029/2020GL091123, 2020.

Lintner, B. R., Giannakis, D., Pike, M., and Slawinska, J.: Identification of the Madden–Julian Oscillation With Data-Driven Koopman Spectral Analysis, Geophysical Research Letters, https://doi.org/10.1029/2023gl102743, 2023.

Liu, Y., Li, C., Wang, J., and Long, M.: Koopa: Learning Non-stationary Time Series Dynamics with Koopman Predictors, arXiv: Learning, 2023.

Lubbers, N., Bolton, D. C., Mohd-Yusof, J., Marone, C., Barros, K., and Johnson, P. A.: Earthquake catalog-based machine learning identification of laboratory fault states and the effects of magnitude of completeness., Geophysical Research Letters, https://doi.org/10.1029/2018gl079712, 2018.

Lusch, B., Kutz, J. N., and Brunton, S. L.: Deep learning for universal linear embeddings of nonlinear dynamics., Nature Communications, https://doi.org/10.1038/s41467-018-07210-0, 2018.

Mastella, G., Corbi, F., Bedford, J., Funiciello, F., and Rosenau, M.: Forecasting surface velocity fields associated with laboratory seismic cycles using Deep Learning, Geophysical Research Letters, https://doi.org/10.1029/2022GL099632, 2022.

Okazaki, T., Hirahara, K., and Ueda, N.: Physics-Informed Deep Learning Approach for Modeling Crustal Deformation, , https://doi.org/10.21203/rs.3.rs-1576456/v1, 2022.

Riel, B., Minchew, B., and Bischoff, T.: Data-Driven Inference of the Mechanics of Slip Along Glacier Beds Using Physics-Informed Neural Networks: Case Study on Rutford Ice Stream, Antarctica, Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2021ms002621, 2021.

https://doi.org/10.5194/gmd-2024-46
Preprint. Discussion started: 22 March 2024

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., and Johnson, P. A.: Machine Learning Predicts Laboratory Earthquakes, Geophysical Research Letters, 9276–9282, https://doi.org/10.1002/2017gl074677, 2017.

Rouet-Leduc, B., Hulbert, C., Bolton, D. C., Ren, C. X., Riviere, J., Marone, C., Guyer, R. A., and Johnson, P. A.: Estimating Fault Friction from Seismic Signals in the Laboratory, Geophysical Research Letters, 1321–1329, https://doi.org/10.1002/2017gl076708, 2018.

Rouet-Leduc, B., Hulbert, C., and Johnson, P. A.: Continuous chatter of the Cascadia subduction zone revealed by machine learning, Nature Geoscience, 75–79, https://doi.org/10.1038/s41561-018-0274-6, 2019.

Shokouhi, P., Girkar, V., Rivière, J., Shreedharan, S., Marone, C., Giles, C. L., and Kifer, D.: Deep Learning Can Predict Laboratory Quakes From Active Source Seismic Data, Geophysical Research Letters, https://doi.org/10.1029/2021gl093187, 2021.

Takeishi, N., Kawahara, Y., and Yairi, T.: Learning Koopman Invariant Subspaces for Dynamic Mode Decomposition, Advances in neural information processing systems, 2017.

Takens, F.: Detecting strange attractors in turbulence, in: Lecture Notes in Mathematics,Dynamical Systems and Turbulence, Warwick 1980, 366–381, https://doi.org/10.1007/bfb0091924, 1981.

Tong, Y., Hong, R., Zhang, Z., Aihara, K., Chen, P., Liu, R., Chen, L., and Kelemen, P.: Earthquake alerting based on spatial geodetic data by spatiotemporal information transformation learning, Proceedings of the National Academy of Sciences, https://doi.org/10.1073/pnas.230227512, 2023.

Veedu, D., Giorgetti, C., Scuderi, MarcoM., Barbot, S., Marone, C., and Collettini, C.: Bifurcations at the Stability Transition of Earthquake Faulting, Geophysical Research Letters, https://doi.org/10.1029/2020GL087985, 2020.

Wang, K., Johnson, C. W., Bennett, K. C., and Johnson, P. A.: Predicting fault slip via transfer learning., Nature Communications, https://doi.org/10.1038/s41467-021-27553-5, 2021.

Wang, K., Johnson, C., Bennett, K., and Johnson, P.: Predicting Future Laboratory Fault Friction Through Deep Learning Transformer Models, Geophysical Research Letters, https://doi.org/10.1029/2022GL098233, 2022.