Danu Caus [1,2], Harsh Grover [1,2], Thomas H. Goebel [4], Grzegorz Kwiatek [3], and Tobias Weigel [1,2]

Predicting **fault stress level** and stress drop using seismo-mechanical and statistical features derived from Acoustic Emissions recorded during laboratory stick-slip friction experiments and **assessing feature importance** via the derived models

TL;DR

Predict rock pressure using deep learning and assess feature importance using the trained model

1 DKRZ DEUTSCHES KLIMARECHENZENTRUM
2 HELMHOLTZ AI | ARTIFICIAL INTELLIGENCE COOPERATION UNIT
3 GFZ Helmholtz-Zentrum Potsdam
4 THE UNIVERSITY OF MEMPHIS Center for Earthquake Research and Information

Motivation

- ➢ The scientific community lacks realistic datasets that give detailed insight into processes leading to occurrence of large seismic events.

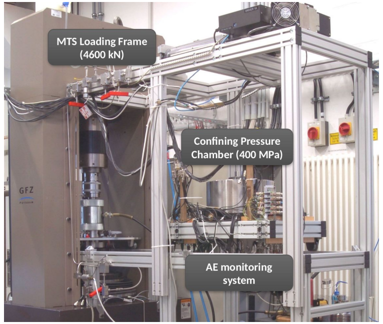
- ➢ Laboratory stick-slip experiments are a direct analog of a seismic cycle.

- ➢ Applied AI methods forecasting earthquakes in laboratories focus on performance at the expense of understanding intrinsic seismo-mechanical processes leading to the nucleation.

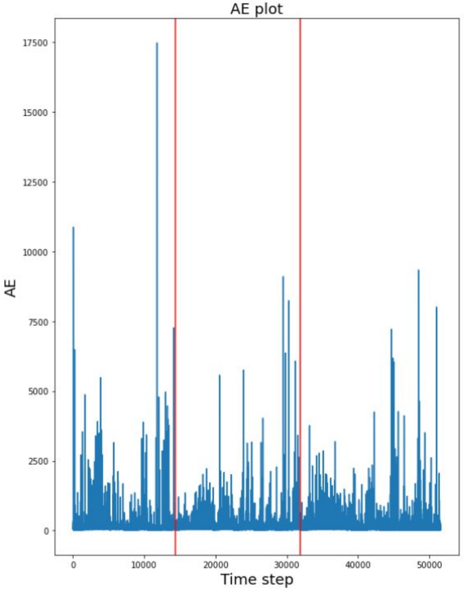
- ➢ ML/AI forecasting techniques applied in the lab were applied to smooth faults. What about the rough and heterogeneous faults?

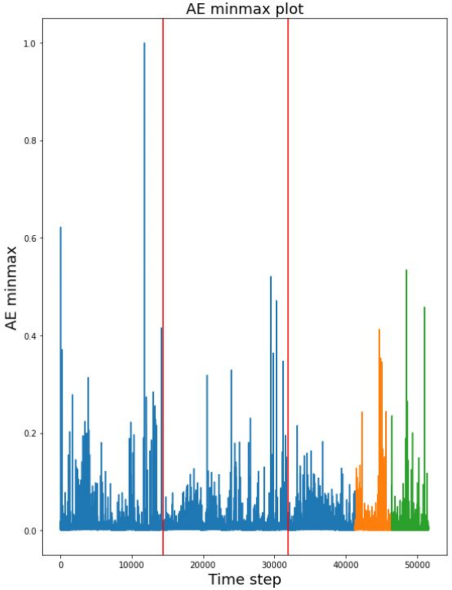
- ➢ Recent and ongoing development of AI explainability methods allows us to gain insight into the best features as seen through the eyes of a neural network.

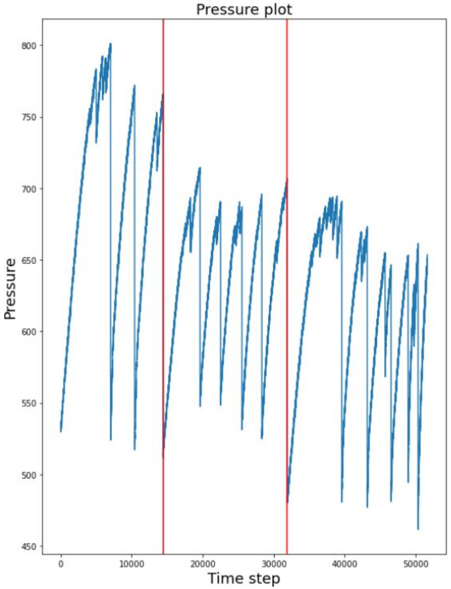
Dataset

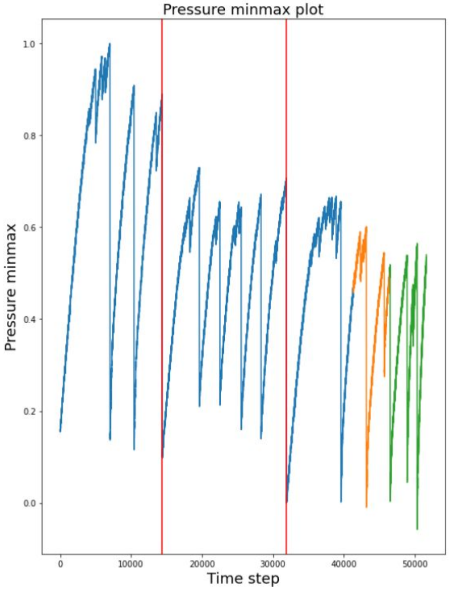


- ➢ We conducted 3 laboratory triaxial stick-slip friction experiments on rough pre-fractured Westerly Granite samples. The experiments were associated with AE activity and geomechanical parameters (**axial pressure**) monitoring .
- ➢ We derived **domain scientific features** from the acquired data to be used as input for our model.
- ➢ We merged the data from all experiments and then split the dataset into 80/10/10 **train/val/test splits**.









Compute meaningful features using **acoustic emissions/AE**

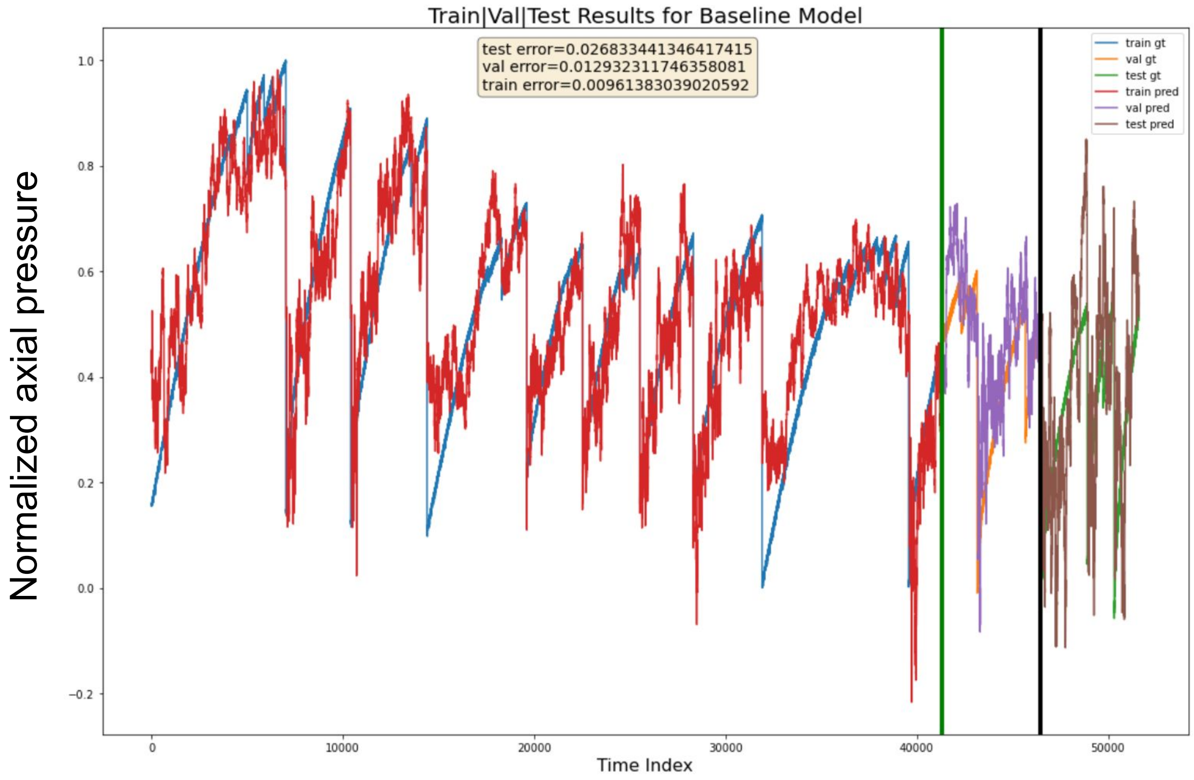
Predict **pressure at time t+1** using past time features

Feature list

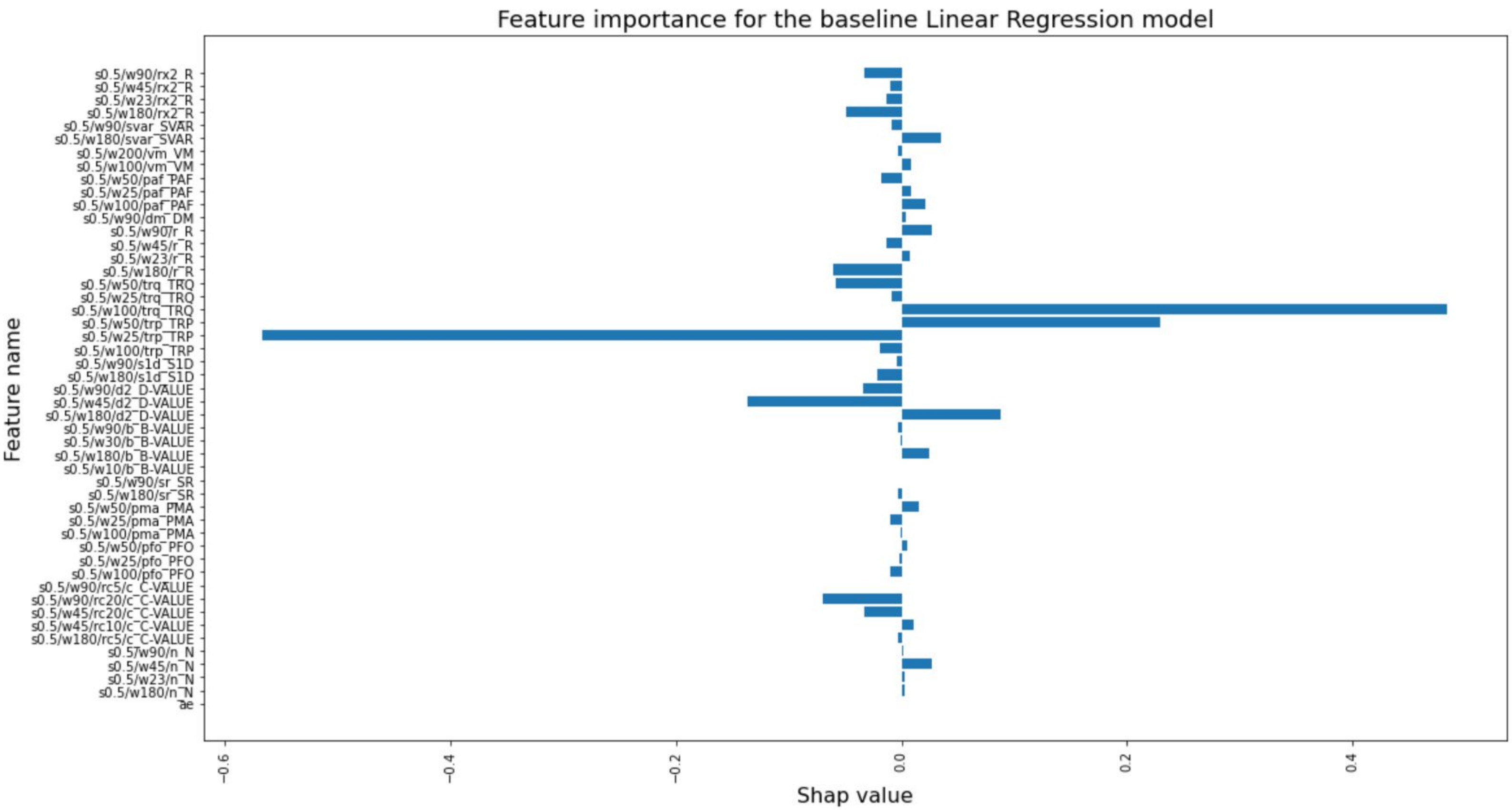
- Our features are **computed from Acoustic Emissions** / AE.
- They are **physically explainable**, not waveform-based simple statistical parameters.
- They **describe damage and stress** evolution in the sample rock.
- They are computed for different spatial and temporal scales / using **different window sizes**.
- **Examples**: Gutenberg-Richter **b-value**, Clustering coefficient **c-value**, Fractal dimension **d-value**, **and many more**.

Baseline model performance

- ➢ We used the derived features as input and the pressure as output => **regression problem**.
- ➢ Running a baseline linear regression model suggested **a more powerful non-linear model is needed.**



Sneak peek at feature importance using the SHAP algorithm



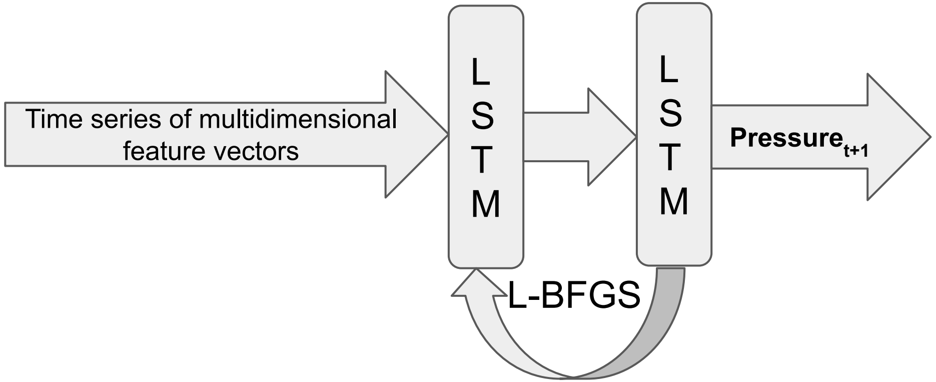
The neural networks will reshape this figure and balance the distribution in favor of other features as well.

Neural network architecture and training technique

- 2 LSTM Cells

- 64 neurons per layer

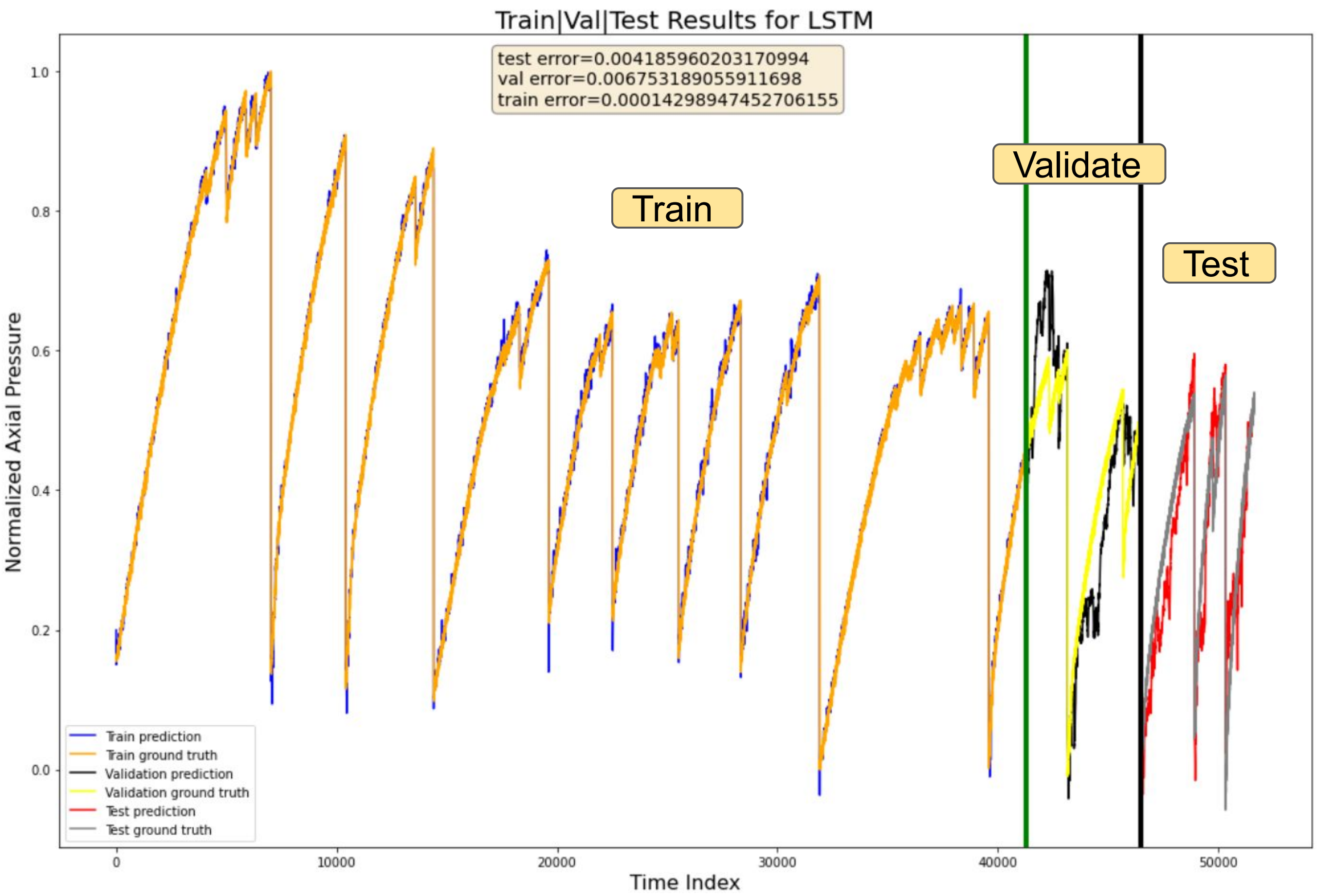
- 0.9 learning rate



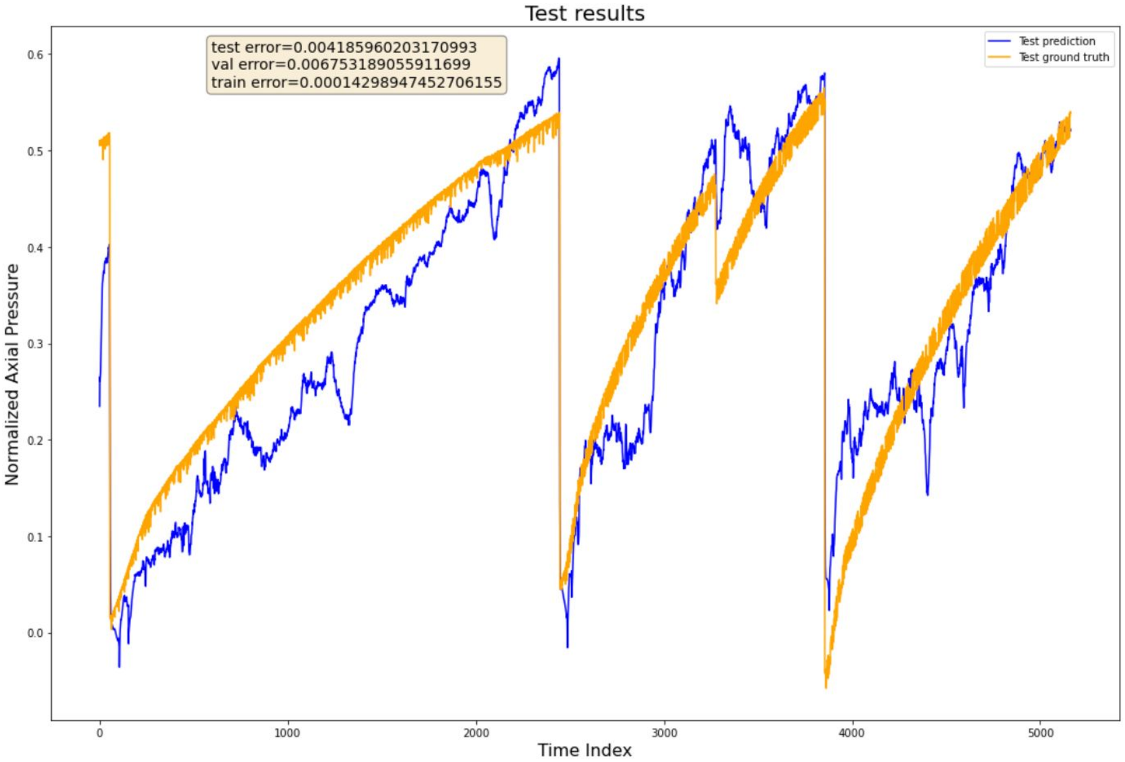
- We used a **quasi-Newton** method for training, namely **Limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm** / **L-BFGS** optimization. This allowed us to achieve better results than with the ADAM optimizer.

- We deliberately **did not impose a specific sequence length** / window size on the LSTM, resetting h and c variables after the entire training set history was seen.

LSTM model performance



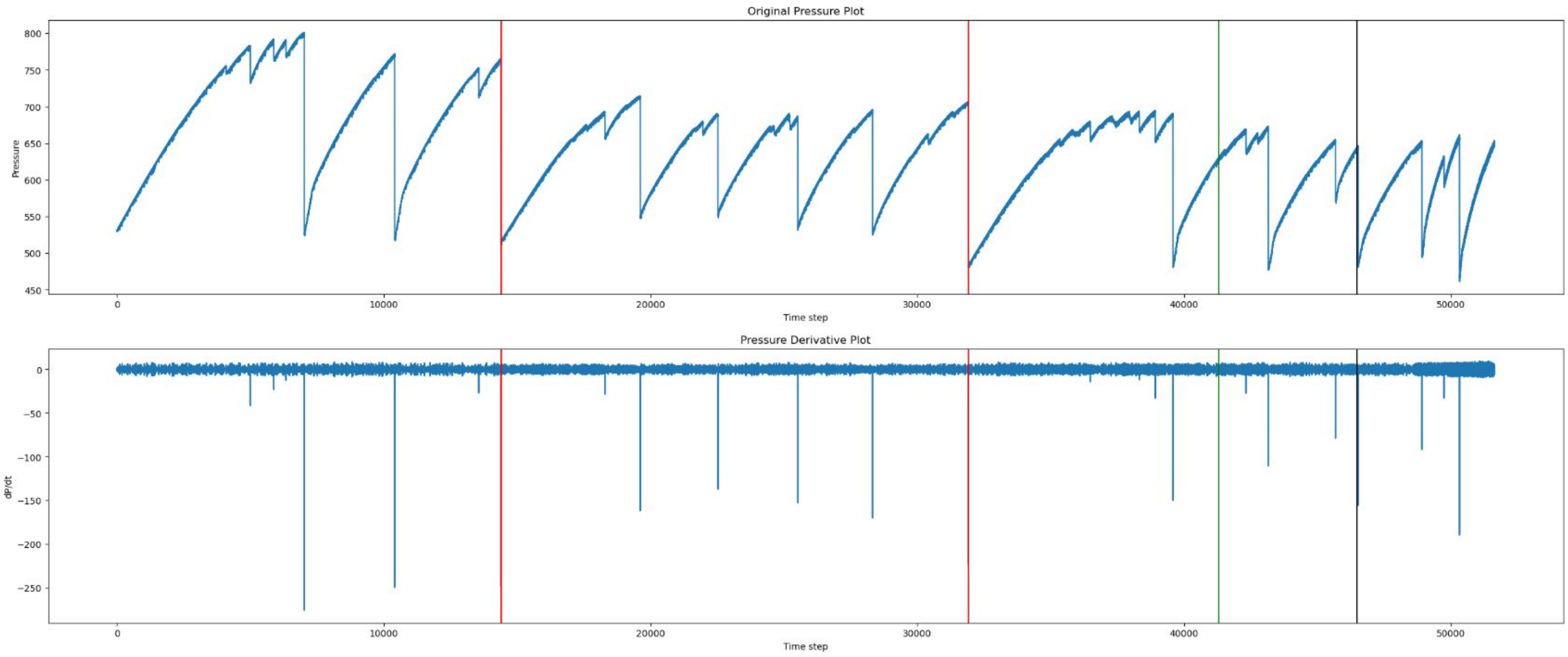
Test dataset performance



Takeaways

- Basic parts of the seismic cycle are modelled well.
- Model cannot capture the rough faults very well.

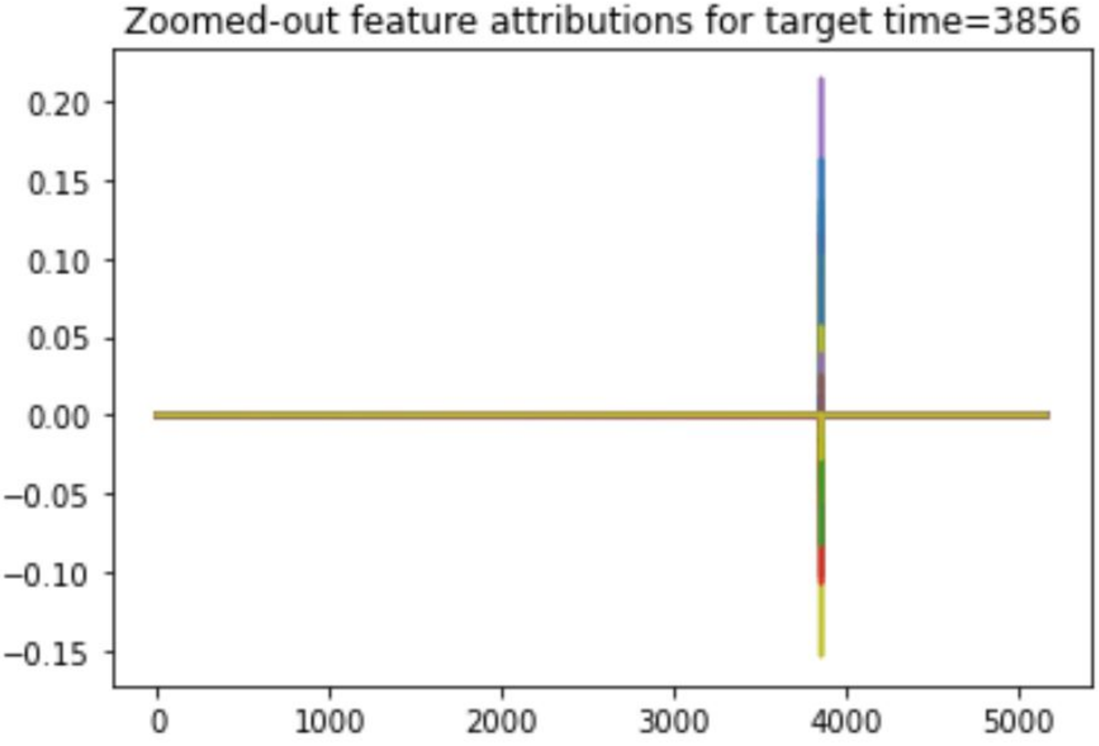
Focusing attention on pressure drop locations



We focus on pressure drop locations **initially** and use **Integrated Gradients** method to find feature attributions.

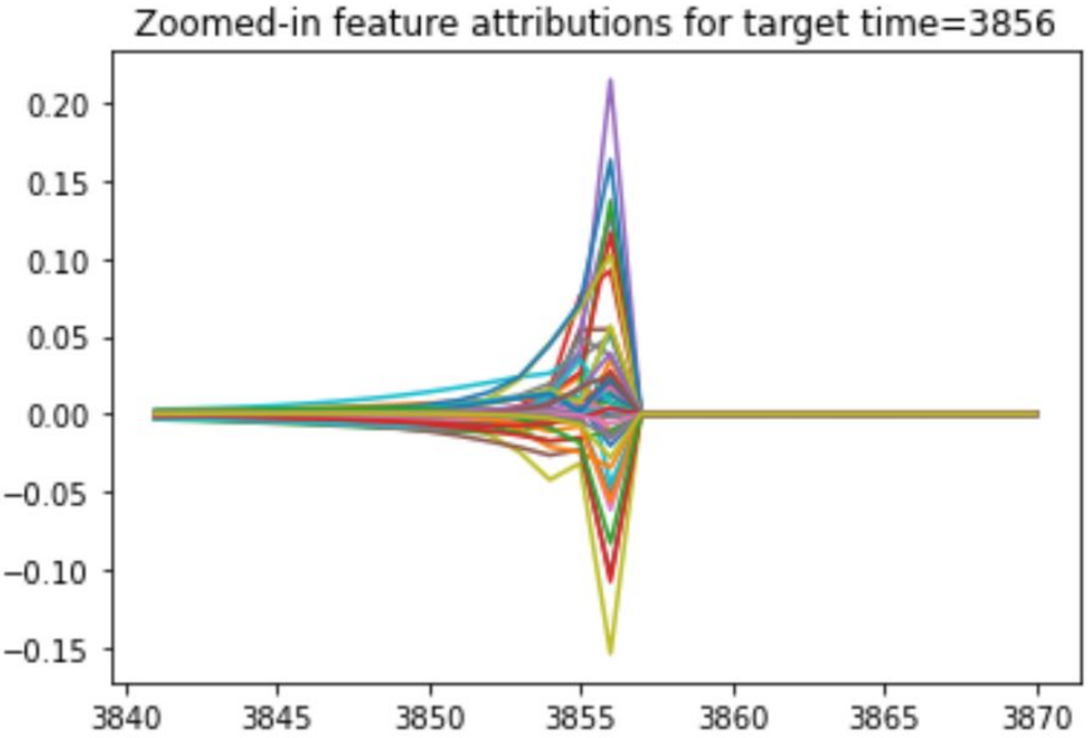
[Sundararajan, M., Taly, A. and Yan, Q., 2017, July. Axiomatic attribution for deep networks. In *International conference on machine learning*]

Seeing the bigger picture



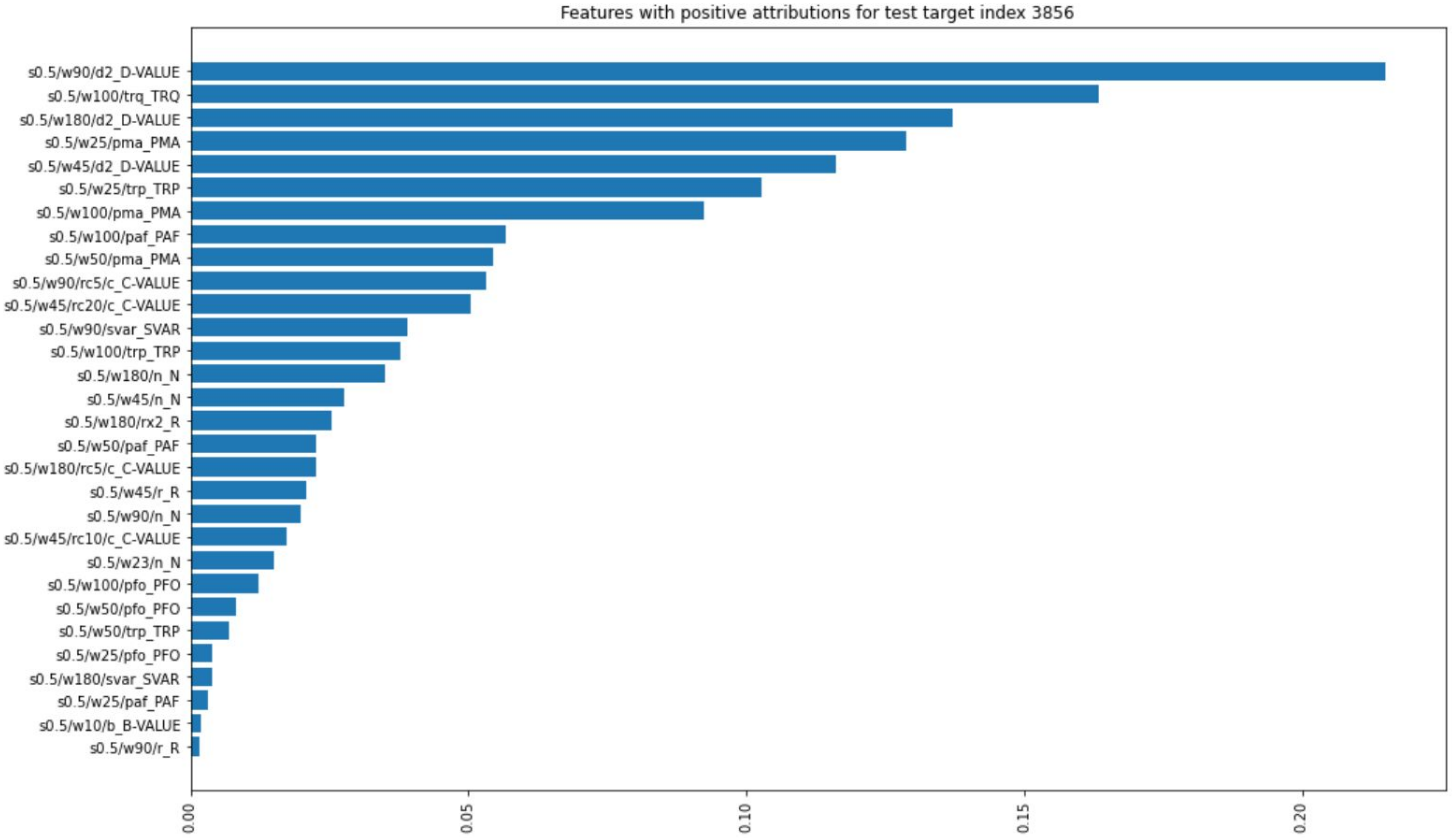
Takeaway: The LSTM learns from data points not too far back in the history, but remember that our **data points already encode** some of the **history** already with various window sizes.

Zooming in on the local history



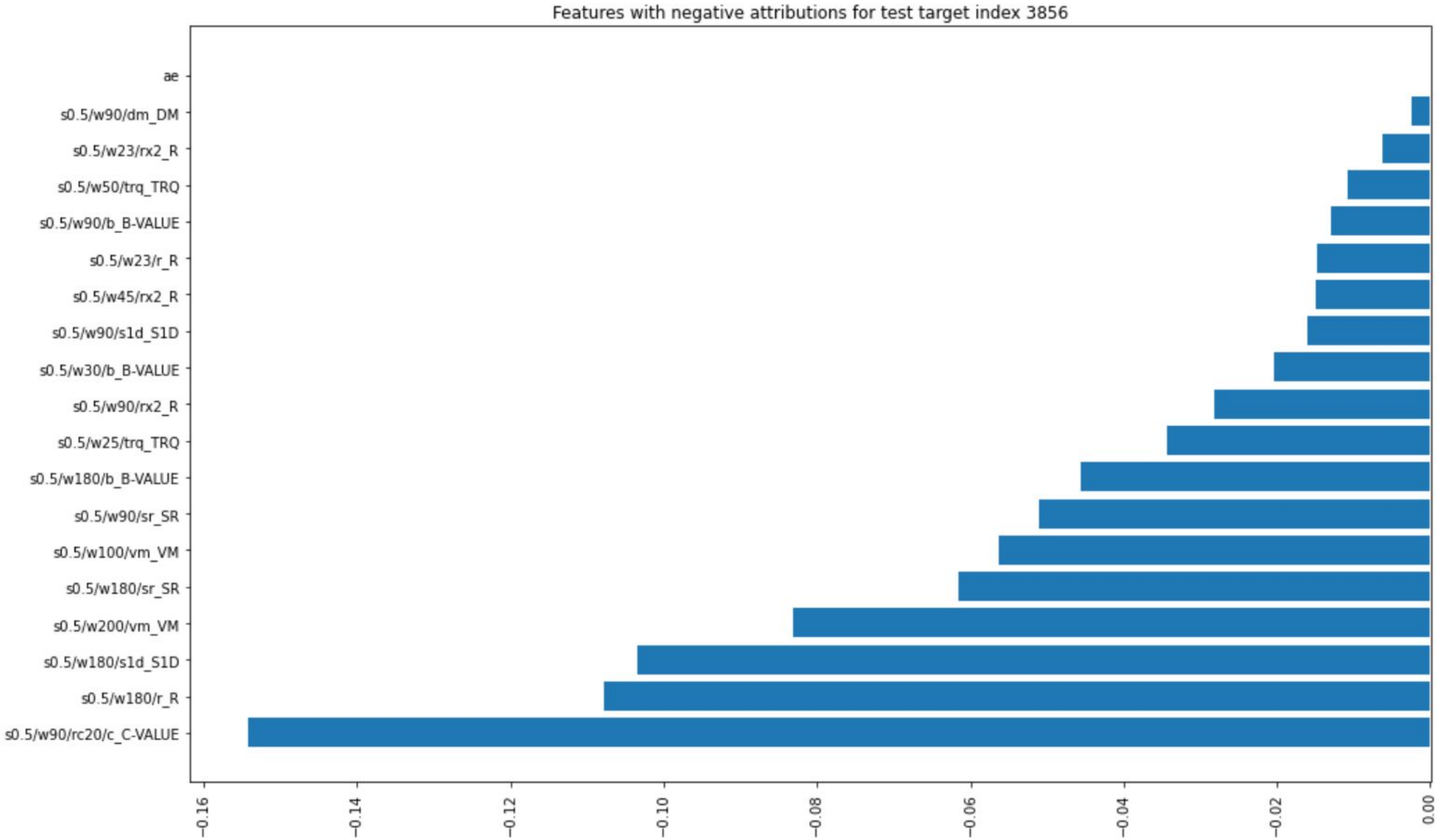
Takeaway: The LSTM considers roughly **10 past data points** to make one pressure prediction at time t+1 and some data points are **computed using 100 seconds** of the process.

Positively correlated features



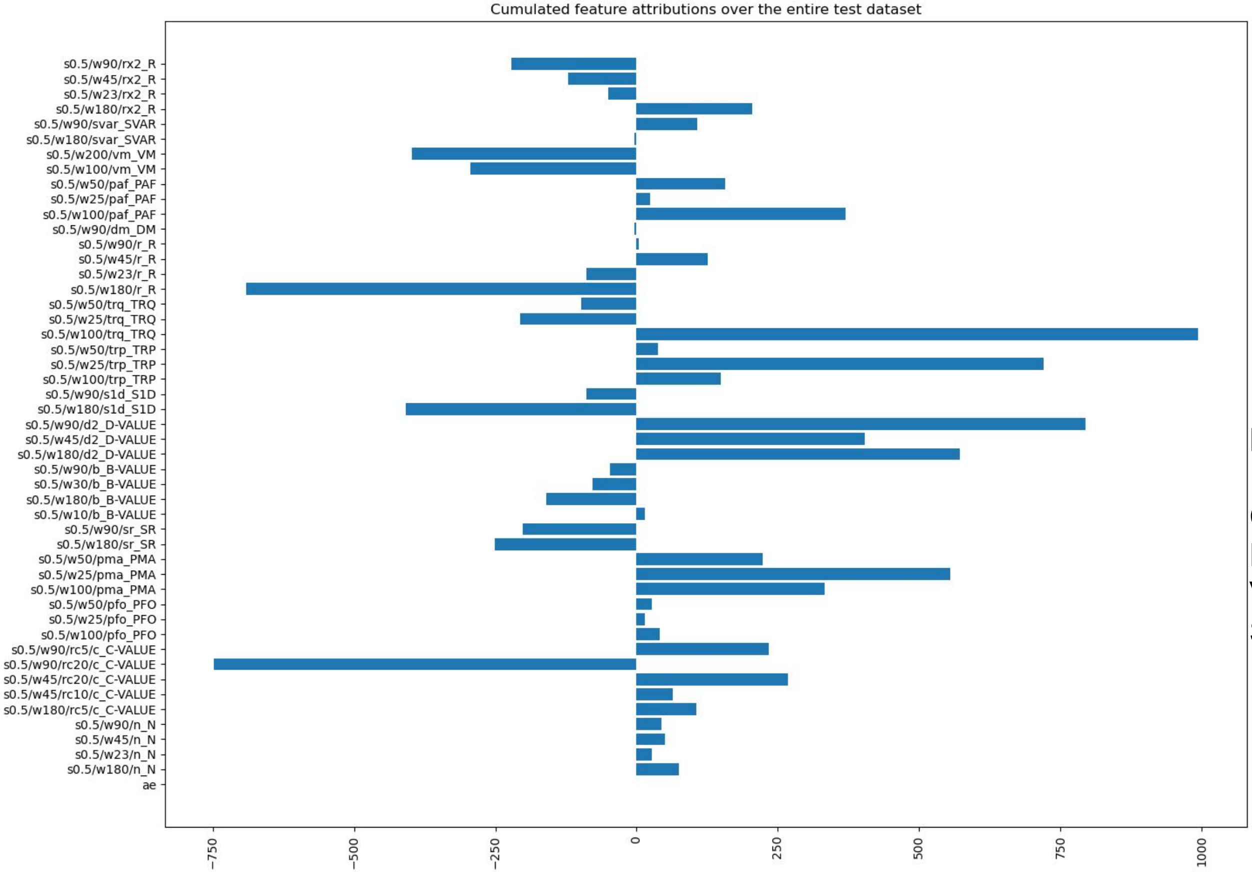
D-value (geometry of AE seismicity clouds) has an important positive effect on making correct predictions.

Negatively correlated features



C-value (clustering coefficient) is inversely correlated with axial pressure prediction.

Feature importance after integrating over the entire test dataset



Positive Correlation Features

TRQ (measure of interevent time & distances)
TRP (spatio-temporal localization of AEs)
D (AE hypocentral distribution)
PMA (proportion of AE mainshocks)

Negative Correlation Features

C (spatial clustering of events)
R (inter-event time ratio, temporal clustering)
VM (local focal mechanism variability)
S1D (average dip of maximum principal stress)

Contributions

- We created a realistic dataset from three laboratory stick-slip friction experiments.
- We successfully trained a deep LSTM model using this dataset.
- We analyzed feature importance using the best deep learning model.
- We ranked and selected the salient features for domain scientists to focus on in subsequent work. These features can be further developed and/or improved.