AI foundation models for complex tectonic applications

Christopher W Johnson @ LANL in EES-17

With collaborators: Paul Johnson, Robert Guyer, Laura Laurenti, Chris Marone

AI For Earth Sciences Workshop, Santa Fe, NM

LA-UR-25-29182

Basic Energy Sciences (BES)

1. Labquake and tectonic applications overview

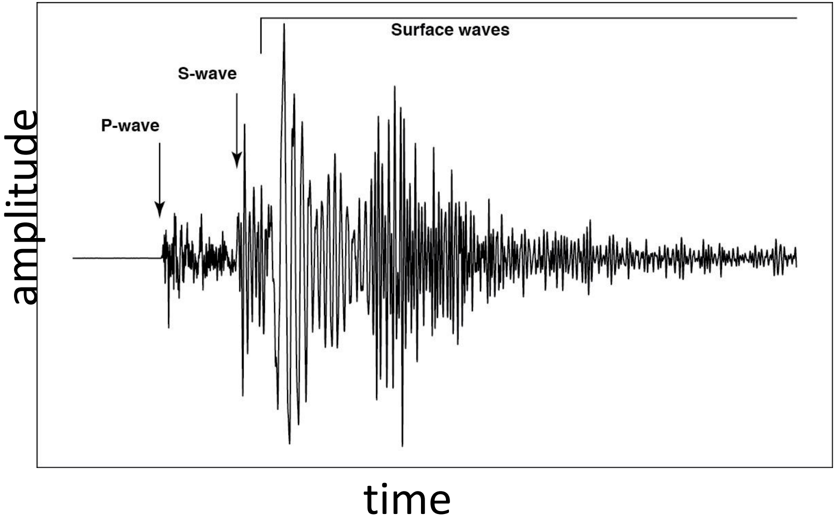
1. **Labquake and tectonic applications overview**

2. Transfer Learning to Self-Supervised Learning
   a) Lessons from the lab
   b) Automatic Speech Recognition for Ground Displacement Predictions

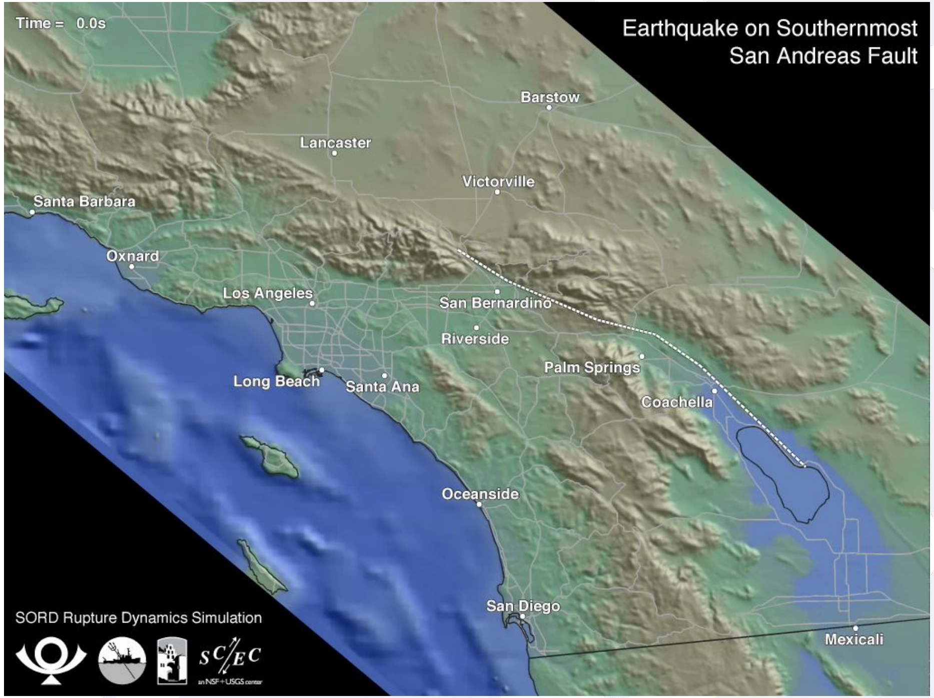
3. Labquake Task Specific Application
   a) PINN Applied to Rate-and-State friction

4. Phase Detection Task Specific Application
   a) Training and Validation
   b) Base Model Contribution

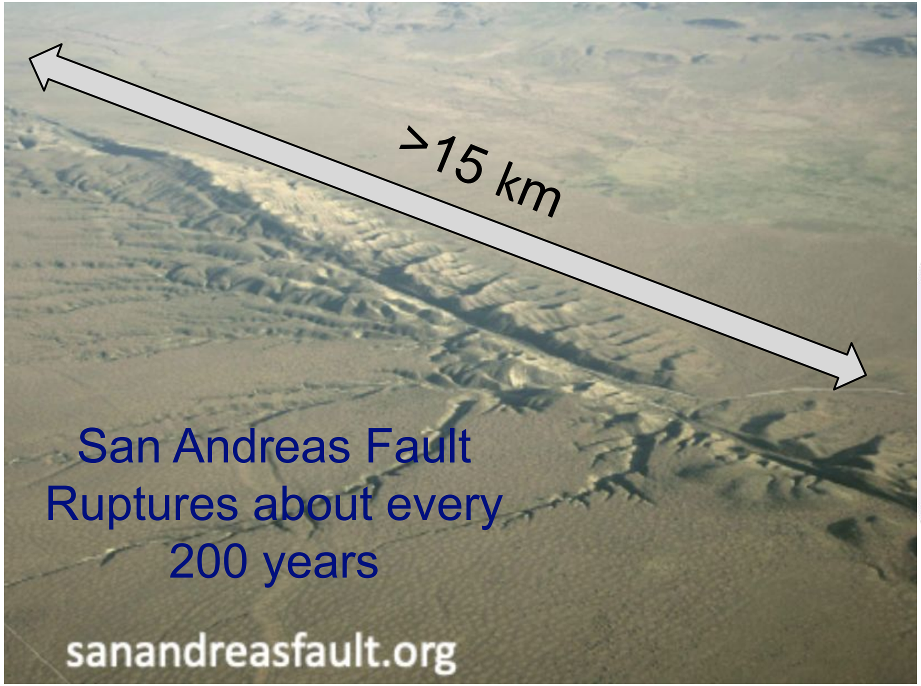
Seismologists study earthquake faults using ground motions recorded by seismometers



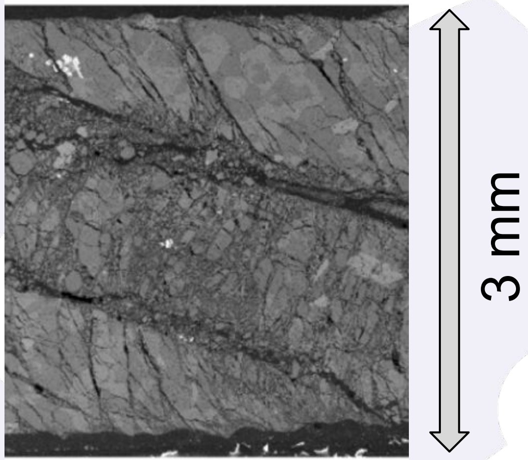
Waveforms contain a lot of information about the source (geometry, magnitude, kinematics), propagation (velocity structure) and site effects.



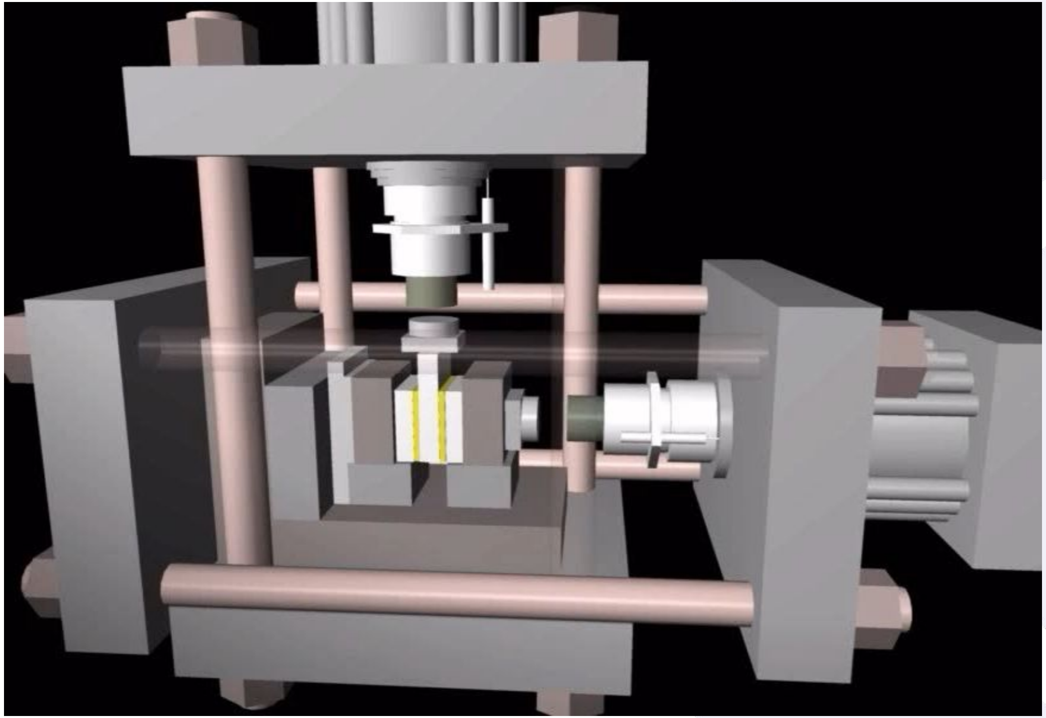
Mechanics of faults is studied in the field and lab, with very different scales



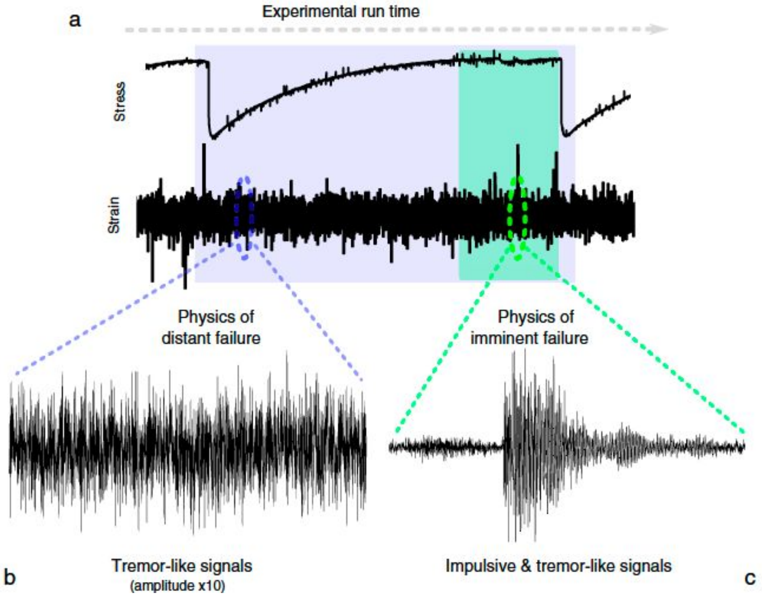
Laboratory Fault



Double direct shearing produce high-fidelity physical measurements of the lab seismic cycle

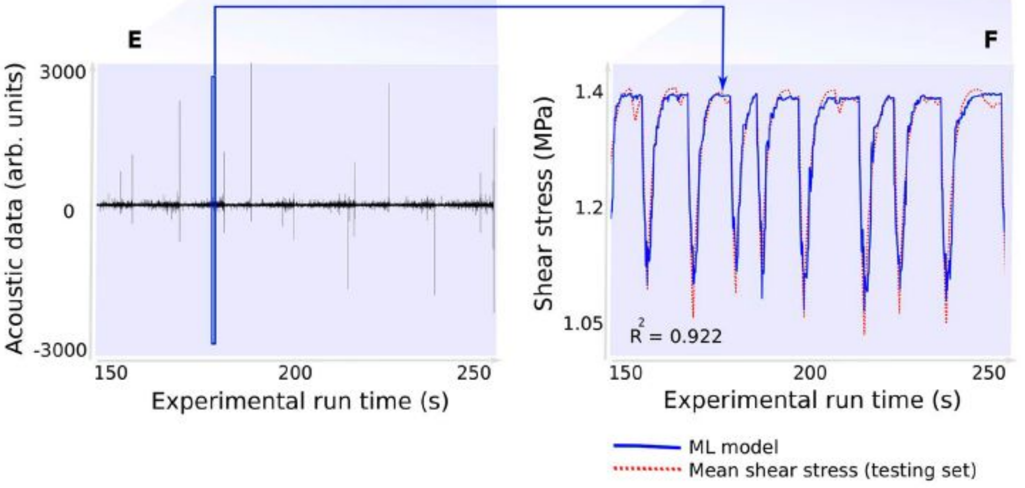


Laboratory signal characteristics are mapped to stress on the fault



Rouet-Leduc et al., 2017, 2018

Regression of shear stress from acoustic emission



Motivated a Kaggle competition with >5000 participants and many publications

Machine learning applied to laboratory data for tracking the evolution of mechanical properties

Geophysical Research Letters
RESEARCH LETTER
10.1029/2018GL081251

Machine Learning Can Predict the Timing and Size of Analog Earthquakes

F. Corbi[1], L. Sandri[2], J. Bedford[3], F. Funiciello[1], S. Brizzi[1,4], M. Rosenau[3], and S. Lallemand[5]

Key Points:
- We simulate multiple seismic cycles in a laboratory-scale subduction

Laboratory earthquake forec
learning competition

Paul A. Johnson[a,1,2], Bertrand Rouet-Leduc[a,1], Laura J. Pyrak
Chris J. Marone[f,g], Claudia Hulbert[h], Addison Howard[i], Philip
Dimosthenis Karaflos[k,3], Corey J. Levinson[l,3], Pascal Pfei

Edited by David A. Weitz, Harvard University, Cambridge, MA, and appr

Geophysical Research Letters
RESEARCH LETTER
10.1029/2022GL098233

Predicting Future Laboratory Fault Friction Through Deep Learning Transformer Models

Kun Wang and Christopher W. Johnson contributed equally to this work.
Kun Wang[1,2], Christopher W. Johnson[1], Kane C. Bennett[1], and Paul A. Johnson[1]

Geophysical Research Letters
RESEARCH LETTER
10.1002/2017GL076708

Estimating Fault Friction From Seismic Signals in the Laboratory

Key Points:
- Machine learning models can discern the frictional state of a laboratory fault from the statistical characteristics

Bertrand Rouet-Leduc[1], Claudia Hulbert[1], David C. Bolton[2], Christopher X. Ren[3], Jacques Riviere[2,4], Chris Marone[2], Robert A. Guyer[1], and Paul A. Johnson[1]

Probing Seismogenic Faults with Machine Learning

1st Paul A. Johnson
Geophysics Group
Los Alamos National Laboratory
Los Alamos, New Mexico, USA
paj@lanl.gov

2nd Christopher W. Johnson
Geophysics Group
Los Alamos National Laboratory
Los Alamos, New Mexico, USA
cwj@lanl.gov

Geophysical Research Letters
RESEARCH LETTER
10.1002/2017GL076708
Estimating Fault Friction From Seismic Signals in the Laboratory
Key Points:
- Machine learning models can discern the frictional state of a laboratory fault from the statistical characteristics of the seismic signal
Bertrand Rouet-Leduc[1], Claudia Hulbert[1], David C. Bolton[2], Christopher X. Ren[3], Jacques Riviere[2,4], Chris Marone[2], Robert A. Guyer[1], and Paul A. Johnson[1]

nature communications

Geophysical Research Letters
RESEARCH LETTER
10.1029/2022GL098233
Predicting Future Laboratory Fault Friction Through Deep Learning Transformer Models
tt[1], and Paul A. Johnson[1]

Characterizing Acoustic Signals and Searching for Precursors during the Laboratory Seismic Cycle Using Unsupervised Machine Learning

by David C. Bolton, Parisa Shokouhi, Bertrand Rouet-Leduc, Claudia Hulbert, Jacques Rivière, Chris Marone, and Paul A. Johnson

Geophysical Research Letters
RESEARCH LETTER
10.1029/2024GL108288

Seismic Features Predict Ground Motions During Repeating Caldera Collapse Sequence

Christopher W. Johnson[1] and Paul A. Johnson[1]
[1]Los Alamos National Laboratory, Los Alamos, NM, USA

Key Points:
- Features extracted from 30 s of continuous seismic waveforms contain information about contemporaneous

Kun Wang[1,2], Christopher W. Johnson[1], Kane C. Bennett[1] & Paul A. Johnson[1]

ing and Shear Stress
; Active Seismic

Rivière[3], and Chris Marone[1,4]

nature communications
Article
https://doi.org/10.1038/s41467-023-39377-6

Using a physics-informed neural network and fault zone acoustic monitoring to predict lab earthquakes

Received: 8 December 2022
Accepted: 7 June 2023

Prabhav Borate[1], Jacques Rivière[1], Chris Marone[2,3], Ankur Mali[4], Daniel Kifer[5] & Parisa Shokouhi[1]

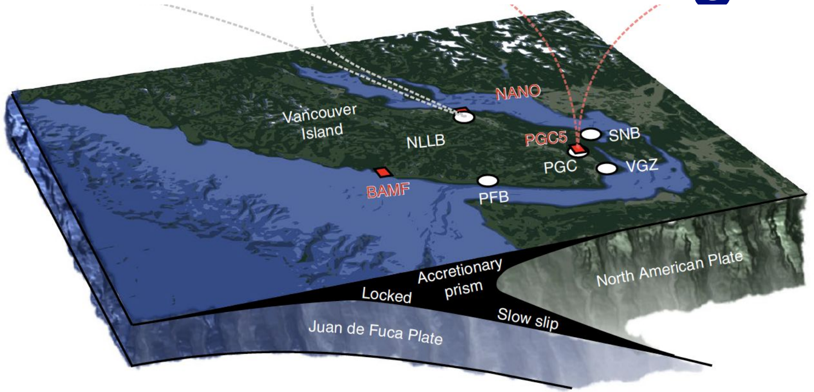
Check for updates

OPEN

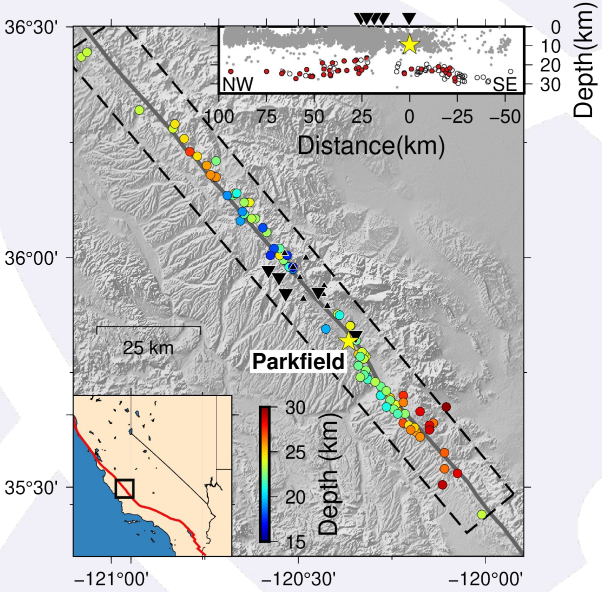
Physics informed neural network can retrieve rate and state friction parameters from acoustic monitoring of laboratory stick-slip experiments

Prabhav Borate[1], Jacques Rivière[1], Samson Marty[2], Chris Marone[3,4], Daniel Kifer[5] & Parisa Shokouhi[1]

Lab machine learning techniques applied to continuous seismic signal in a fault zone



Cascadia subduction zone and central San Andreas fault studied for slowly occurring tectonic motions



Rouet-Leduc et al., 2019
Johnson & Johnson 2021

Los Alamos NATIONAL LABORATORY

2. Transfer Learning to Self-Supervised Learning

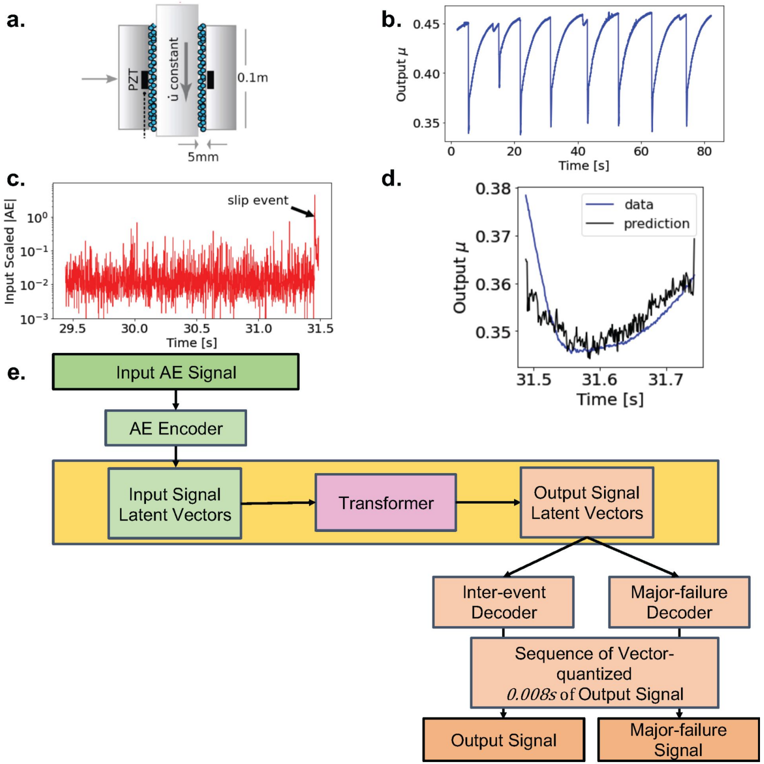
1. Labquake and tectonic applications overview

2. **Transfer Learning to Self-Supervised Learning**
   a) **Lessons from the lab**
   b) **Automatic Speech Recognition for Ground Displacement Predictions**

3. Labquake Task Specific Application
   a) PINN Applied to Rate-and-State friction

4. Phase Detection Task Specific Application
   a) Training and Validation
   b) Base Model Contribution

Forecasting "labquake" using transformer model with a discrete window of history data

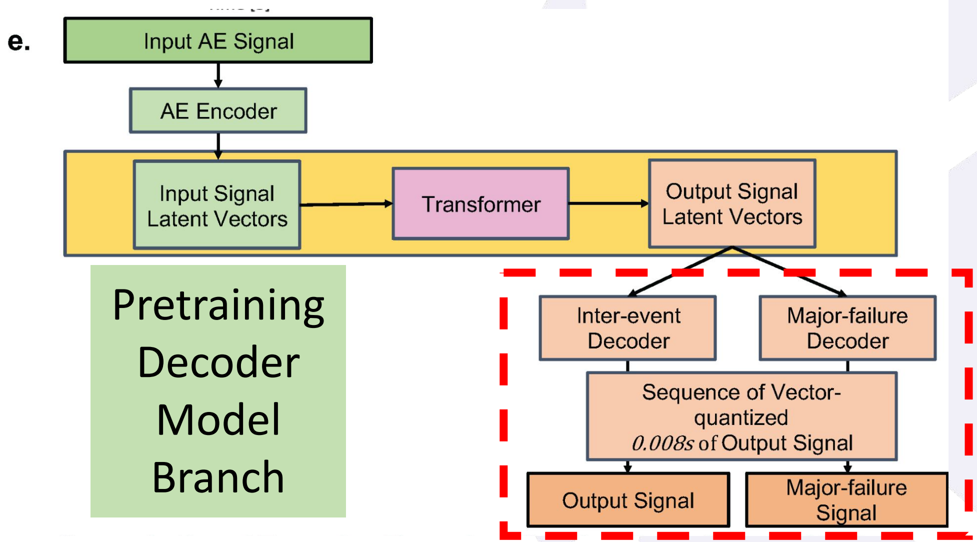


Convolutional Encoder-Decoder (CED) – Transformer Model

- (a) Lab setup
- (b) Lab friction measurements
- (c) Input acoustic emission
- (d) Output window
- (e) CED transformer model

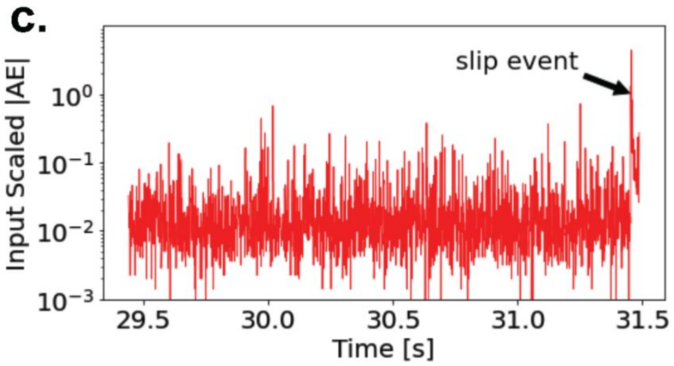
Wang et al. 2022

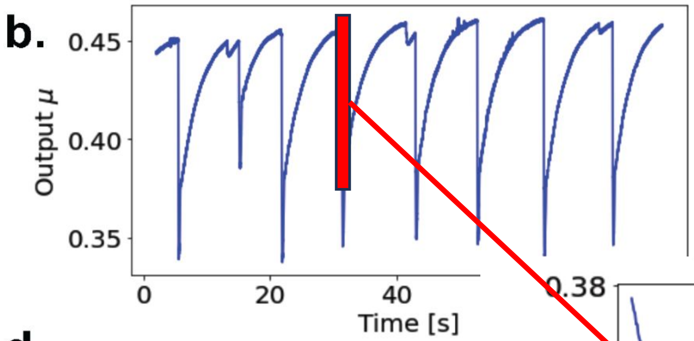
Convolutional encoder decoder model design with pretraining and multiple outputs

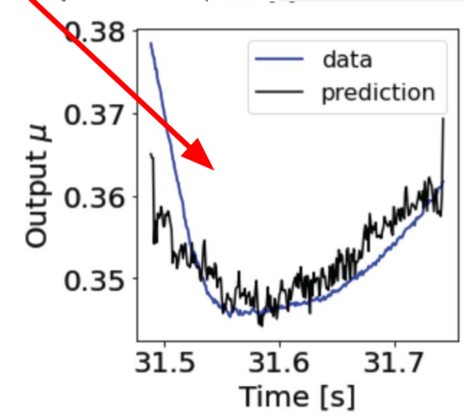


Wang et al. 2022

Forecasting major slip events in laboratory data using AE history as input



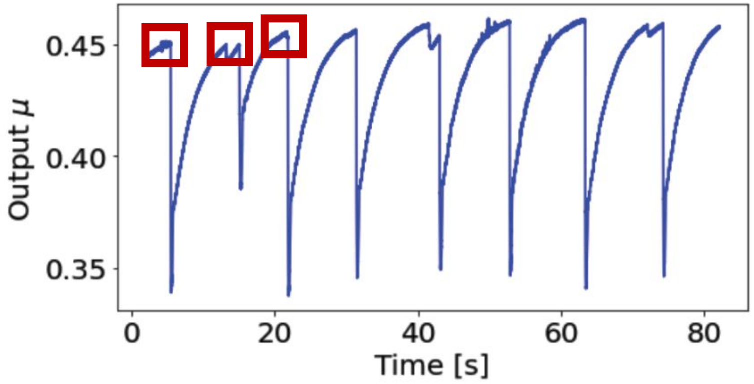


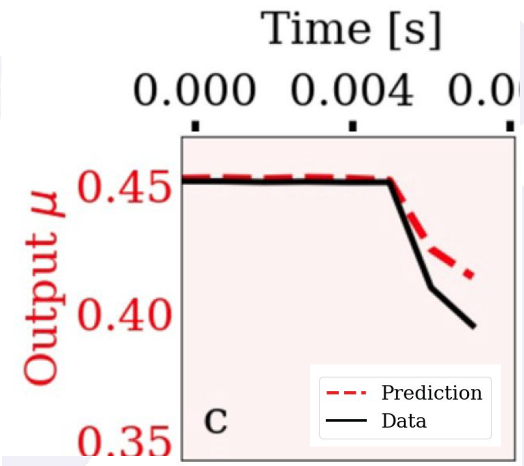


Map information in acoustic emission to slip event

Wang et al. 2022

Forecasting major slip events in laboratory data using AE history as input



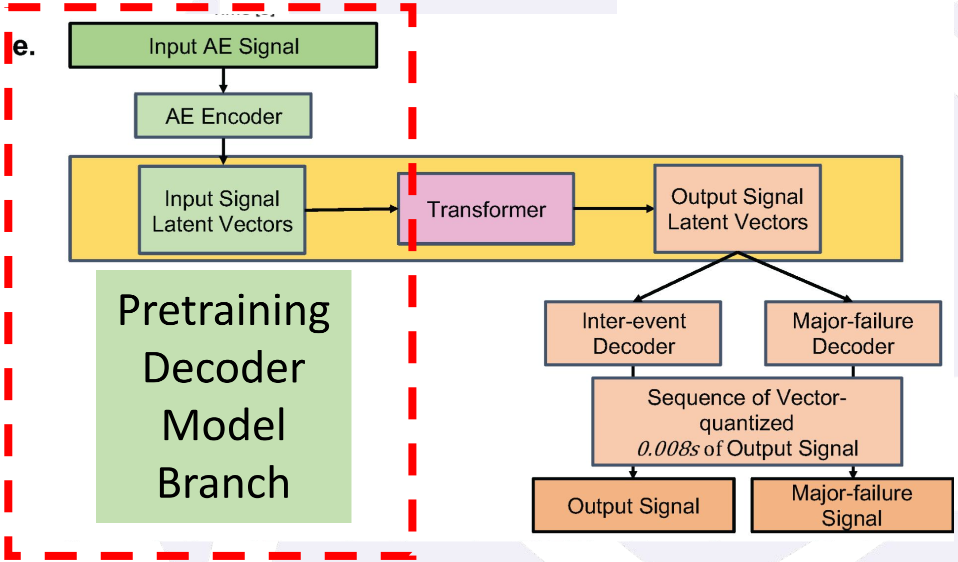


Wang et al. 2022

Pretraining is the key to this convolutional encoder decoder model with multiple outputs

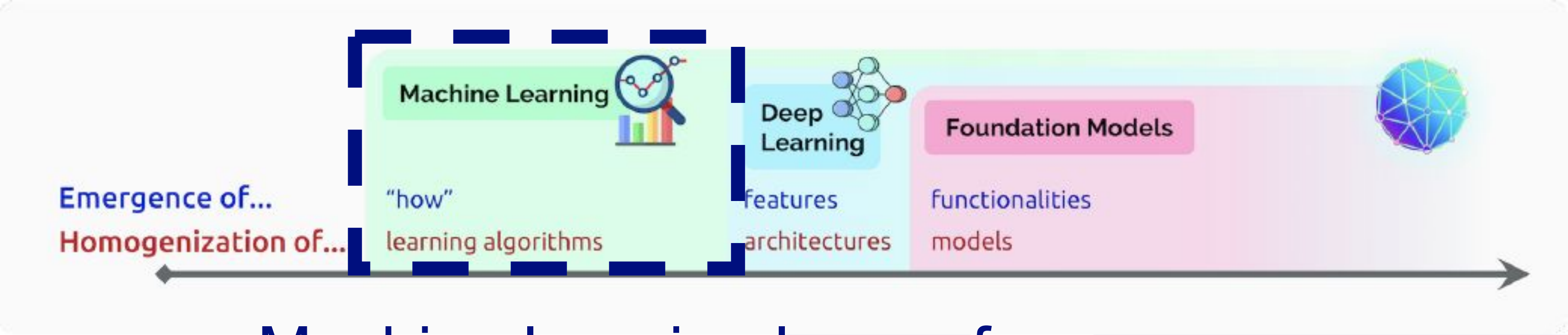
Encoder generalizes inputs and applying transfer learning for specific tasks

Model design utilized supervised learning



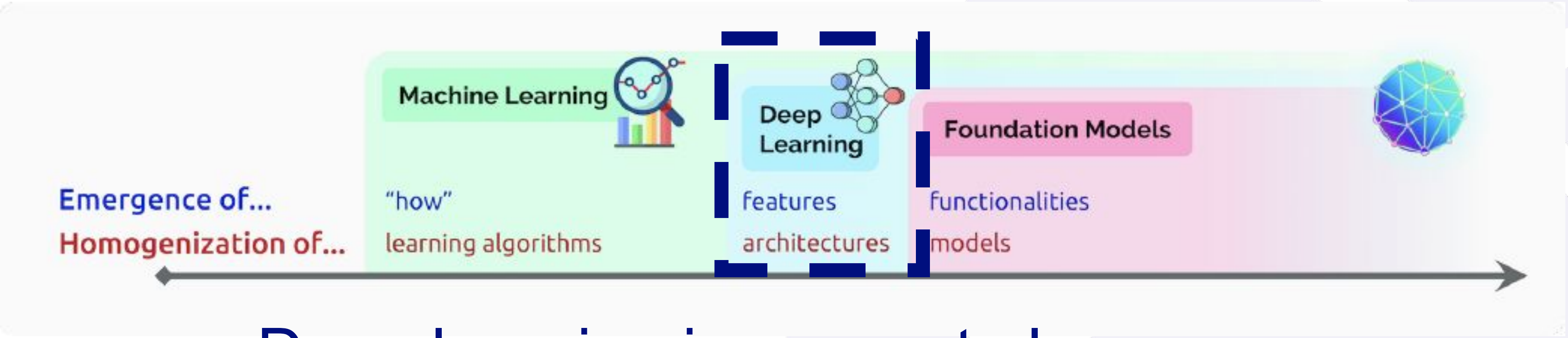
Wang et al. 2022

Machine learning to deep learning and now foundation models for next generation



Machine learning learns from historical data to make predictions "How" to solve the task emerges from the data

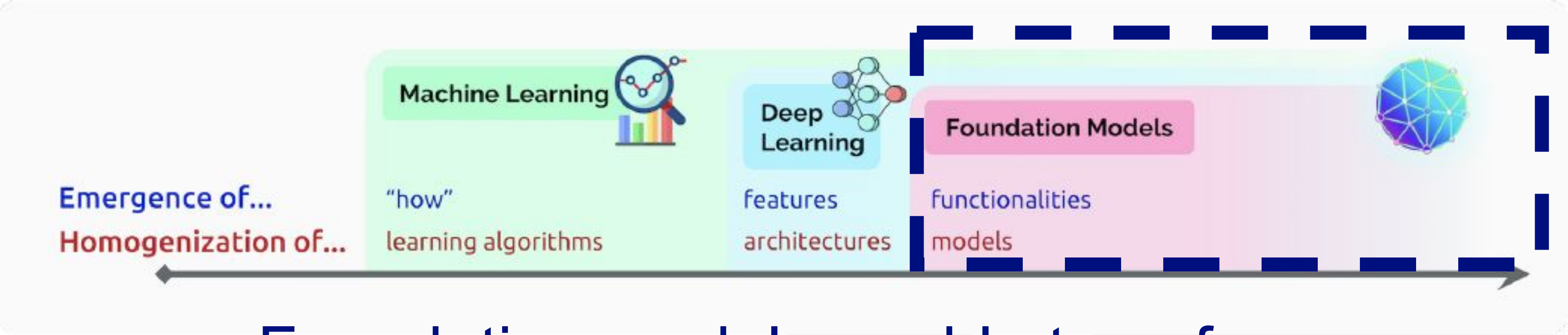
Machine learning to deep learning and now foundation models for next generation



Deep learning incorporate large neural networks. Larger data sets and GPUs allowed higher-level feature extraction from the data

Bommasani et al. 2021

Machine learning to deep learning and now foundation models for next generation

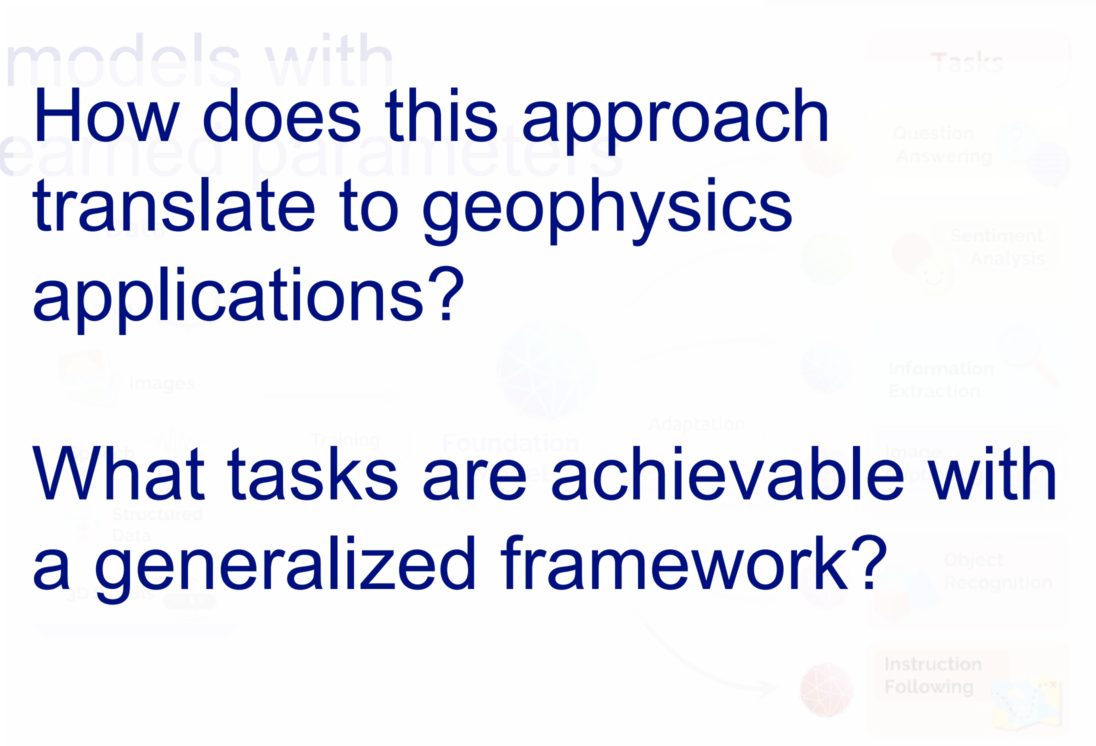


Foundation models enable transfer learning and scale makes these powerful. Pretraining achieved with self-supervised learning

Bommasani et al. 2021

Foundation models for multi-modal data adapted to downstream tasks

Very large models with billions of learned parameters

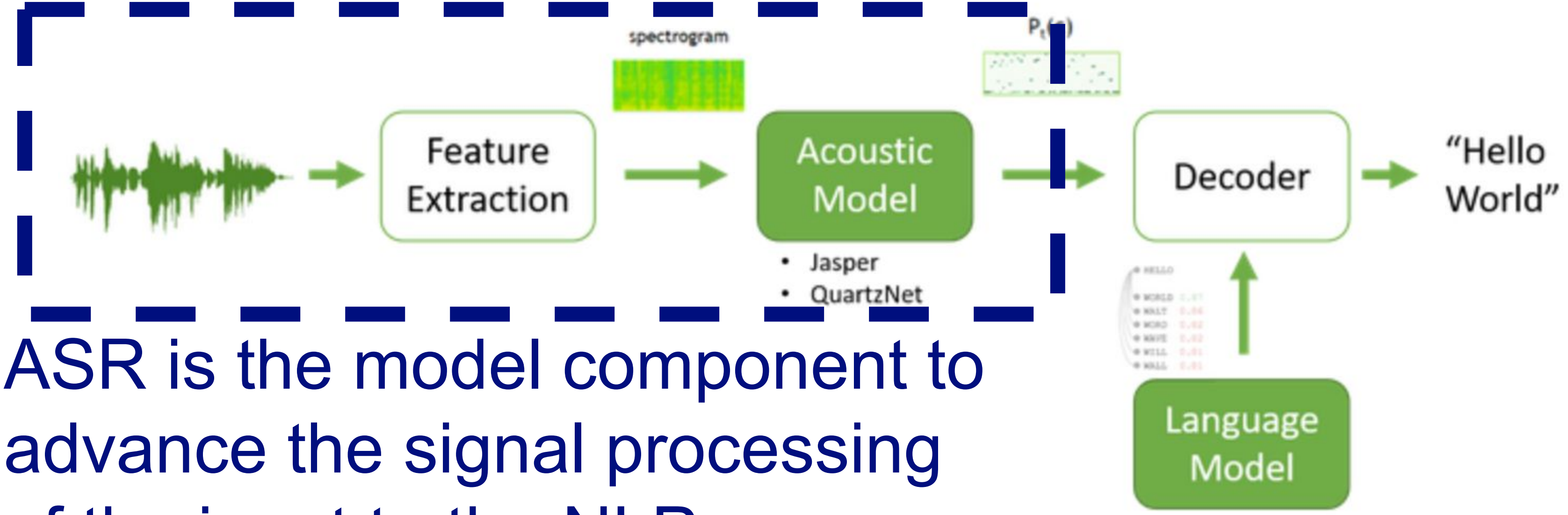


How does this approach translate to geophysics applications?

What tasks are achievable with a generalized framework?

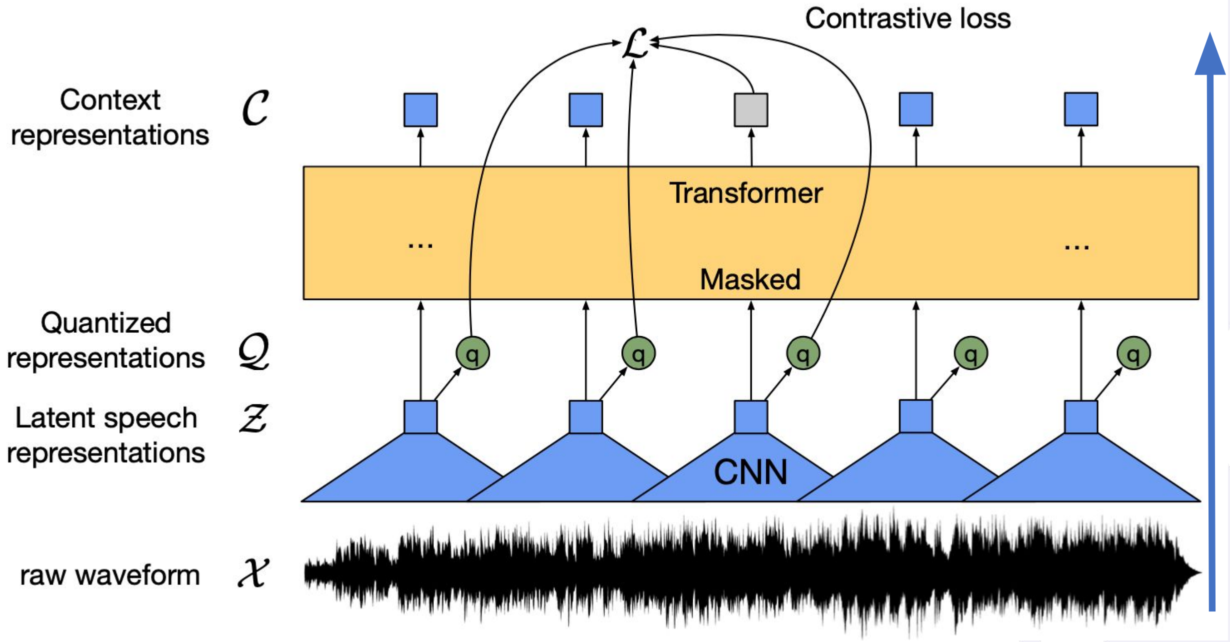
Bommasani et al. 2021

Automatic speech recognition is related to natural language processing



ASR is the model component to advance the signal processing of the input to the NLP

Wav2Vec-2.0 Automatic speech recognition applied to seismic waveforms



Loss on constructed context vectors

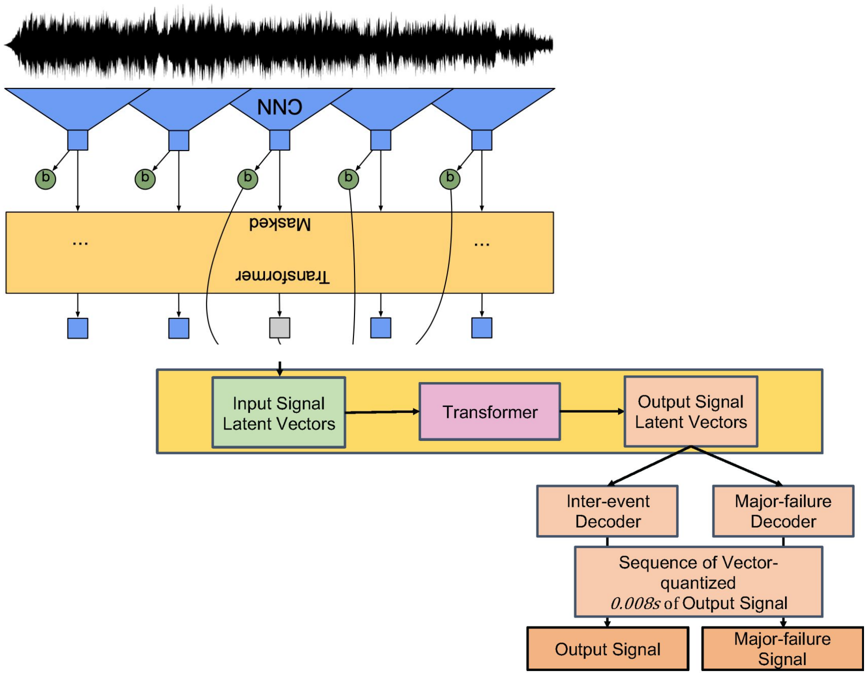
Mask latent vectors in transformer

CNN encoding

Input waveforms

Baevski et al. 2020

Forecasting "labquake" using transformer model with a discrete window of history data



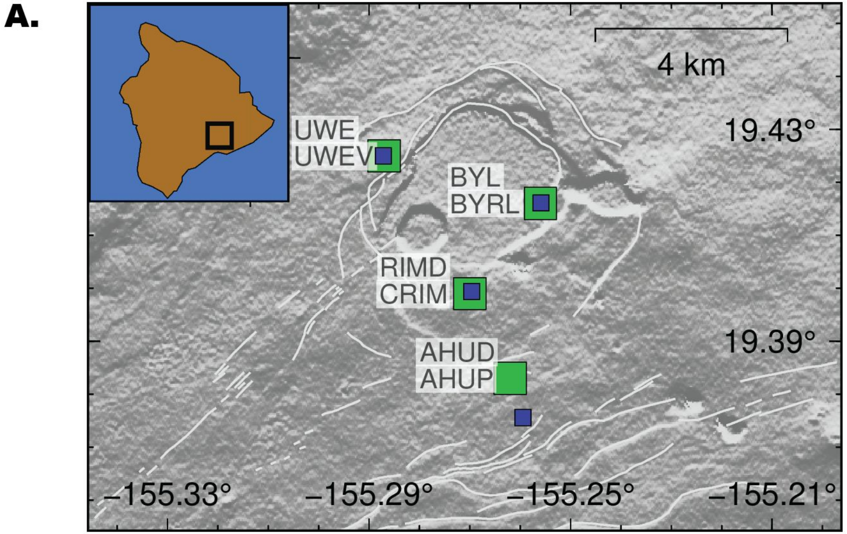
Convolutional Encoder-Decoder (CED) – Transformer Model

Replace encoder with ASR encoder for application to tectonic environment

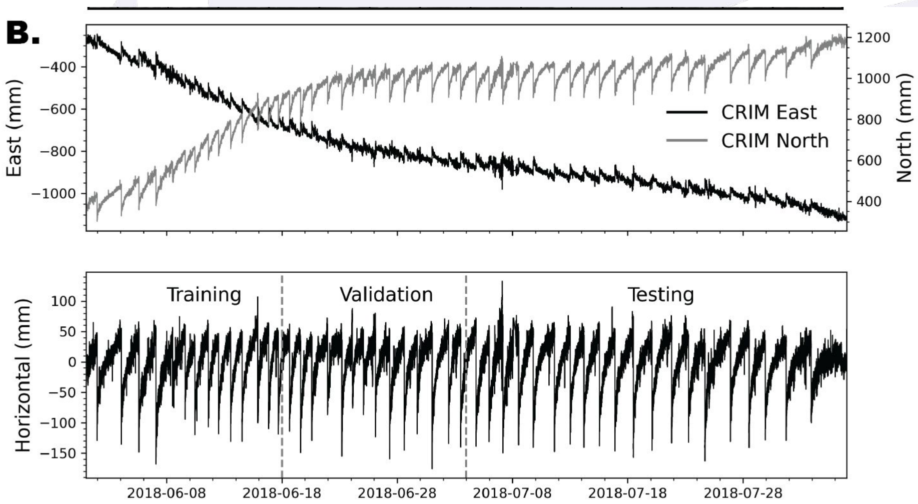
Fine-tune model with 1-week of waveforms
Encode 3 channels separately and concat hidden space

Caldera collapse at Kilauea in 2018 resulted in >50 M~5 earthquakes in 2 months

Seismometer and GNSS

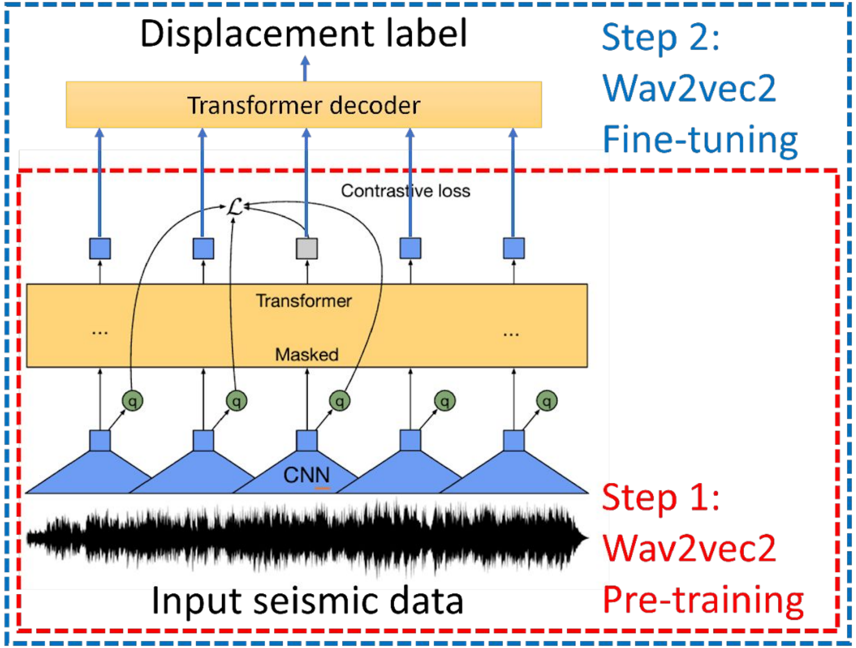


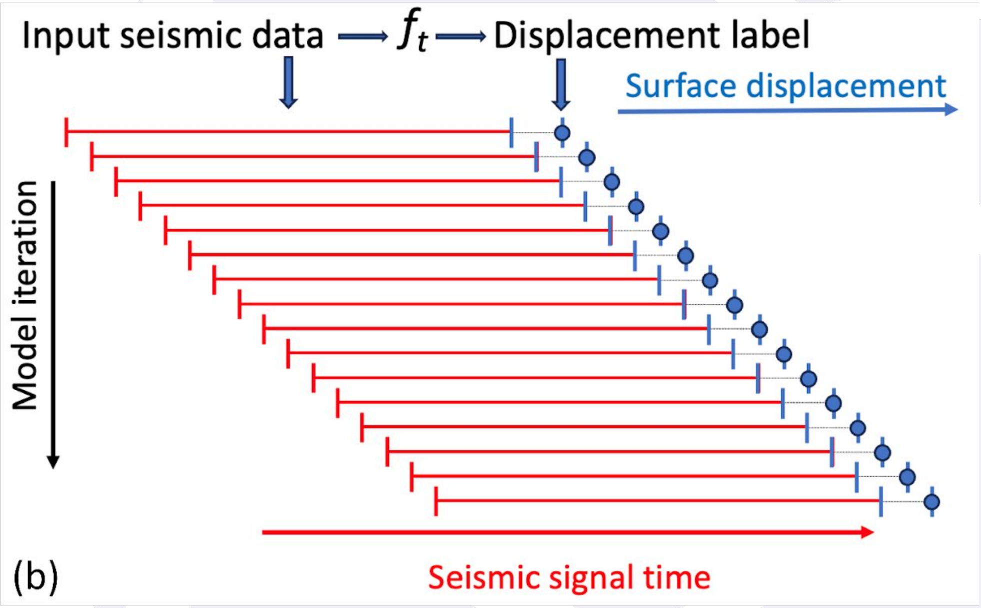
High-rate GNSS



Johnson et al. 2025

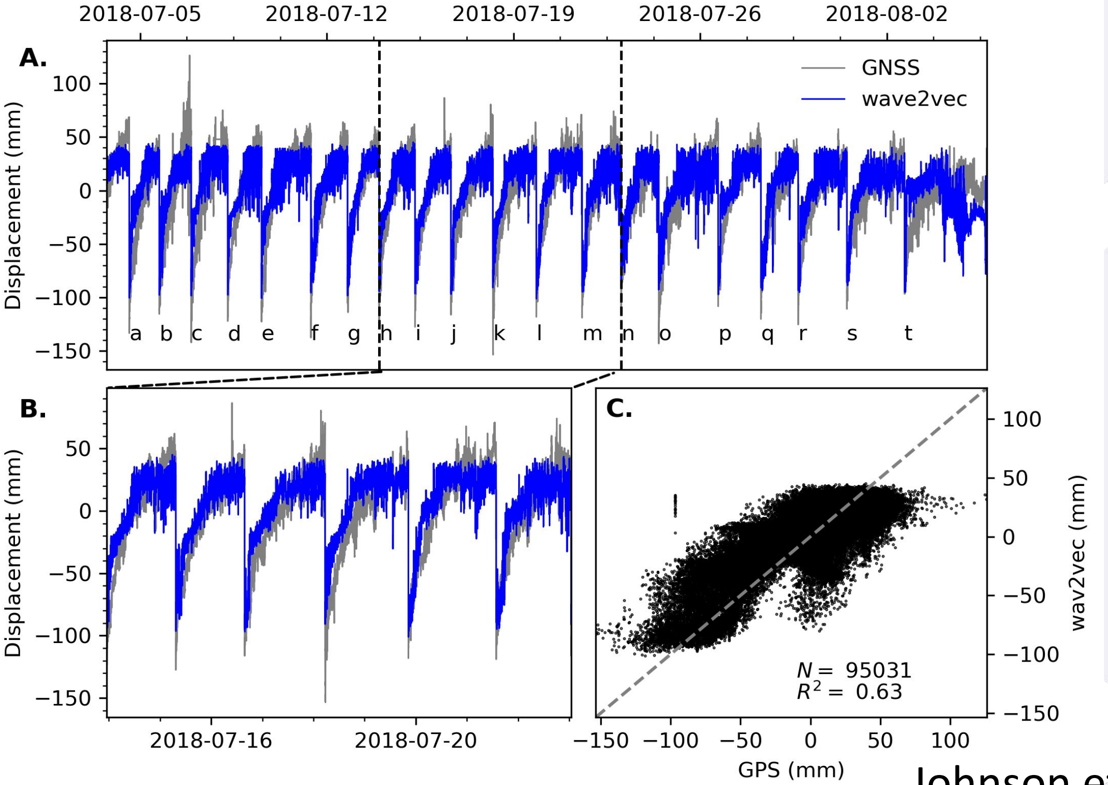
Wav2Vec-2.0 ARS trained to predict ground motions and collapse events

a)


(b)


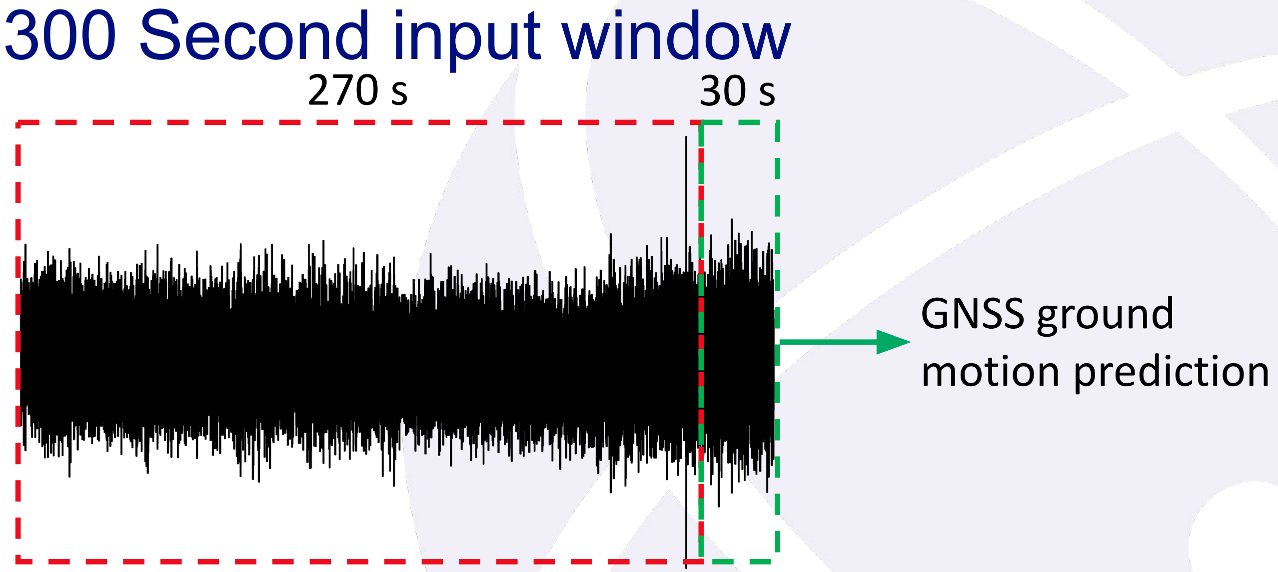
Johnson et al. 2025

Wav2Vec-2.0 ARS trained to predict ground motions and collapse events

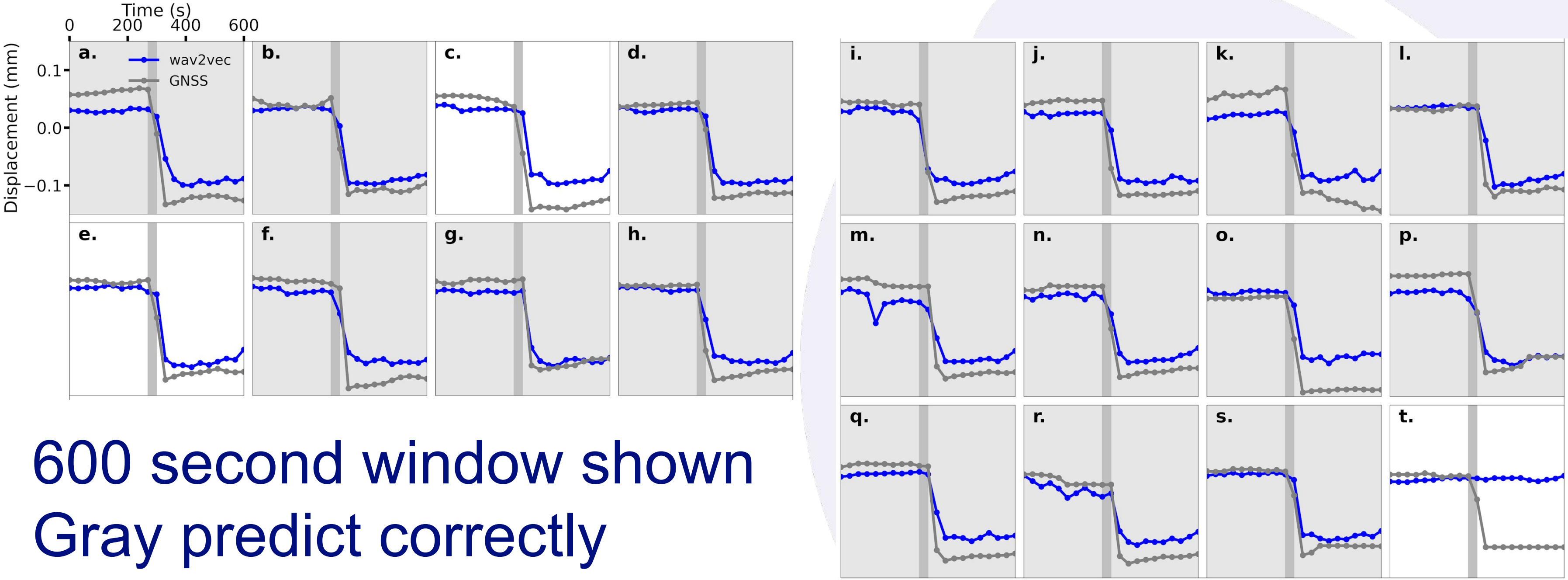


Johnson et al. 2025

Wav2Vec-2.0 ARS trained to predict ground motions and collapse events



Wav2Vec-2.0 predictions for the instantaneous onset of slip for 20 collapse events



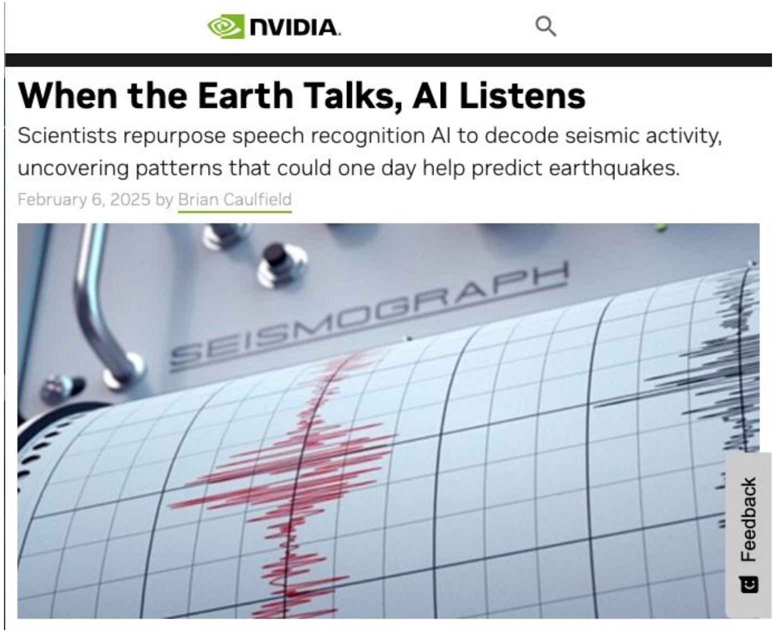
600 second window shown
Gray predict correctly

First time event predictions shown

Johnson et al. 2025

Los Alamos NATIONAL LABORATORY

Pretrained Wav2Vec2 model shows great potential for encoding seismic data


NVIDIA

When the Earth Talks, AI Listens

Scientists repurpose speech recognition AI to decode seismic activity, uncovering patterns that could one day help predict earthquakes.

February 6, 2025 by Brian Caulfield

Open questions:

1/ 1-chan vs 3-chan inputs

2/ Sample rate, seismic (~100 Hz) vs audio (~16kHz)

3/ Training on "quiet" data

3. Labquake Task Specific Application

1. Labquake and tectonic applications overview

2. Transfer Learning to Self-Supervised Learning
   a) Lessons from the lab
   b) Automatic Speech Recognition for Ground Displacement Predictions

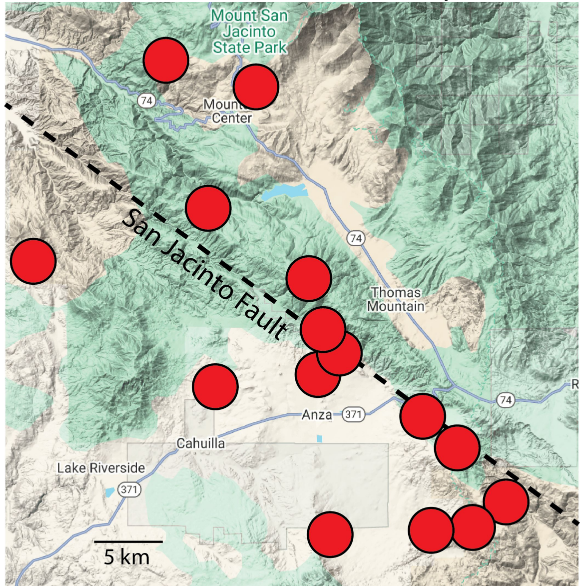
3. **Labquake Task Specific Application**
   a) **PINN Applied to Rate-and-State friction**

4. Phase Detection Task Specific Application
   a) Training and Validation
   b) Base Model Contribution

Reconfigure model for 3-channel seismic waveform inputs and train on continuous data

ANZA network

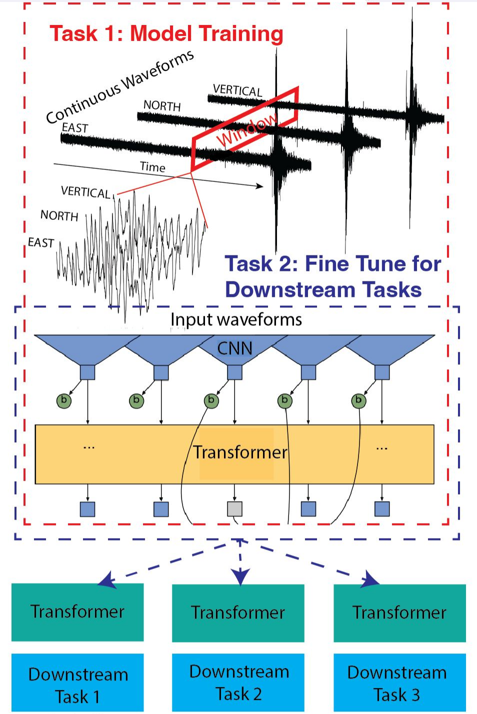
San Jacinto Fault Zone, SoCal



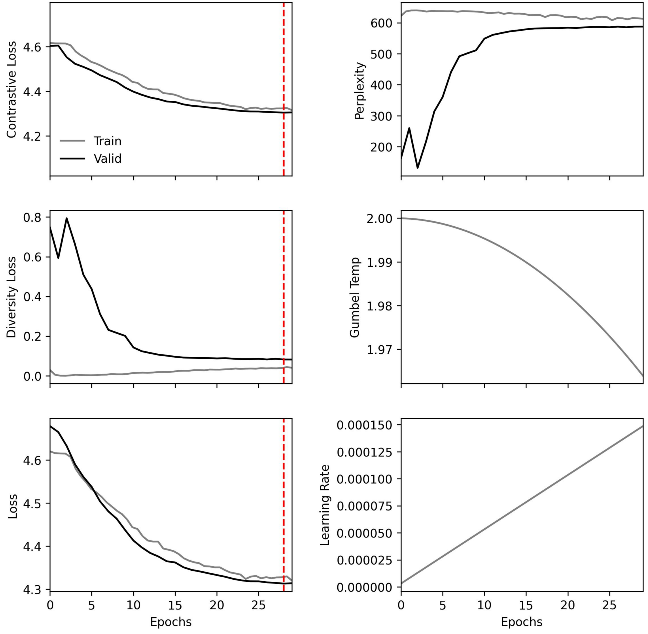
Task 1:
Train model with 5 years of continuous seismic waveforms

~8Tb of data

~25M waveforms

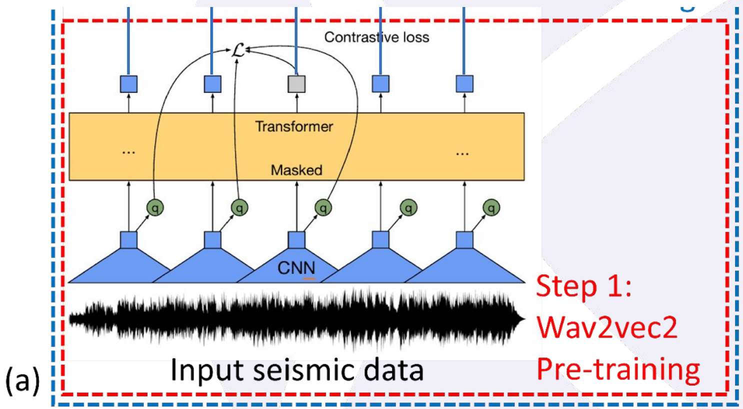


Reconfigure model for 3-channel seismic waveform inputs and train on continuous data



Waveforms are input to model

Model restarts are applied to continue training

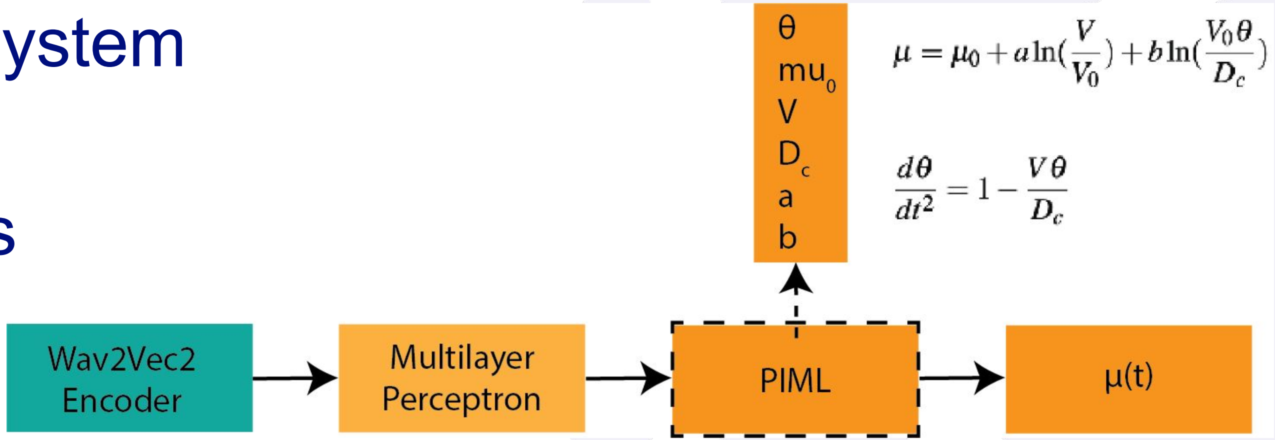


Trained model used for specific tasks

Physics Informed Machine Learning (PIML) using encoded waveforms

Lab experiment with granite-on-granite stick-slip system

Input waveforms to frozen encoder



Johnson, et al. in-prep

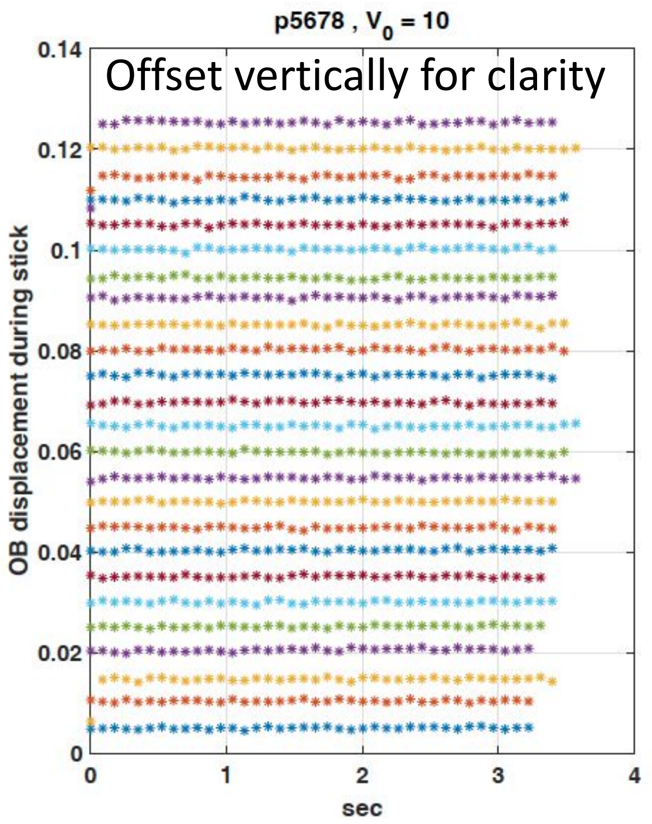
Continuous waveforms from locked system in laboratory experiment

Granite-Granite Blocks

10 μm/s loading velocity

Onboard displacement shown

No measurable motion during loading cycle

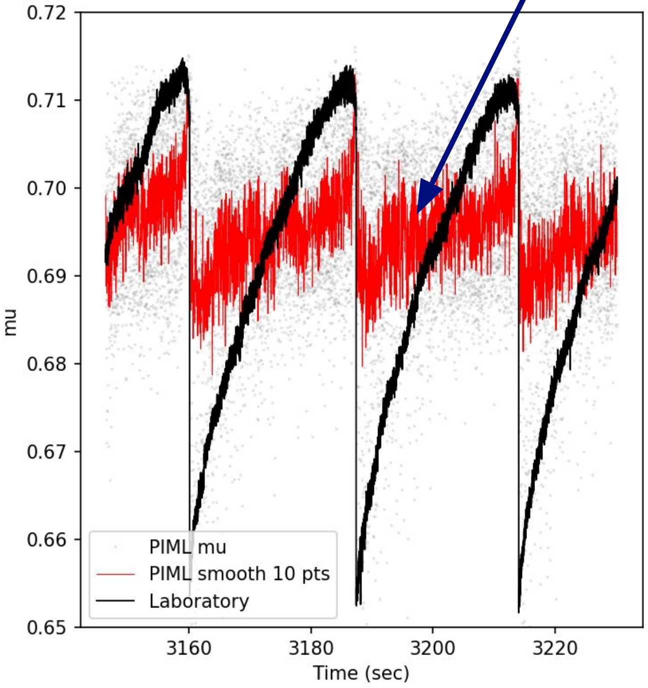


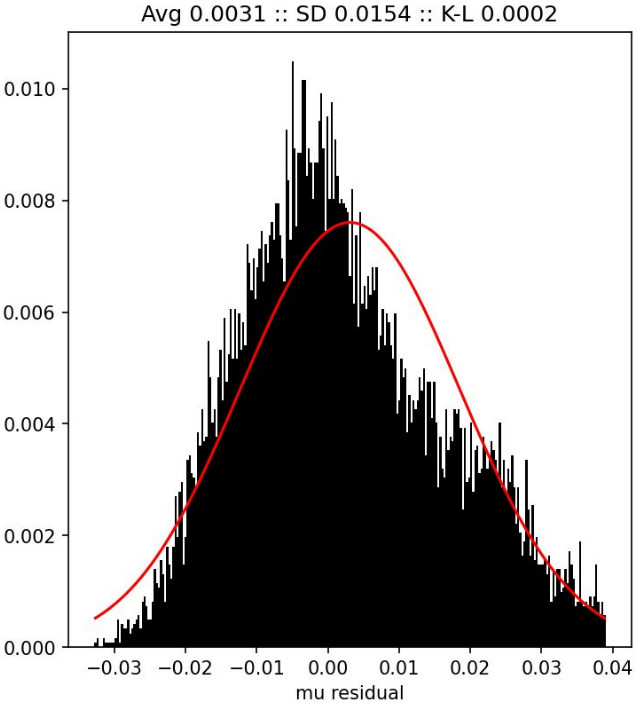
Solving rate-and-state friction in neural network

Model solved parameters are in expected range

Tracks evolution of loading, not perfectly but good

a = 0.012
b = 0.009
mu_0 = 0.65
V_0 = $10^{-3.0}$
V = $10^{-1.4}$
Dc = $10^{-1.8}$





Los Alamos NATIONAL LABORATORY

4. Phase Detection Task Specific Application

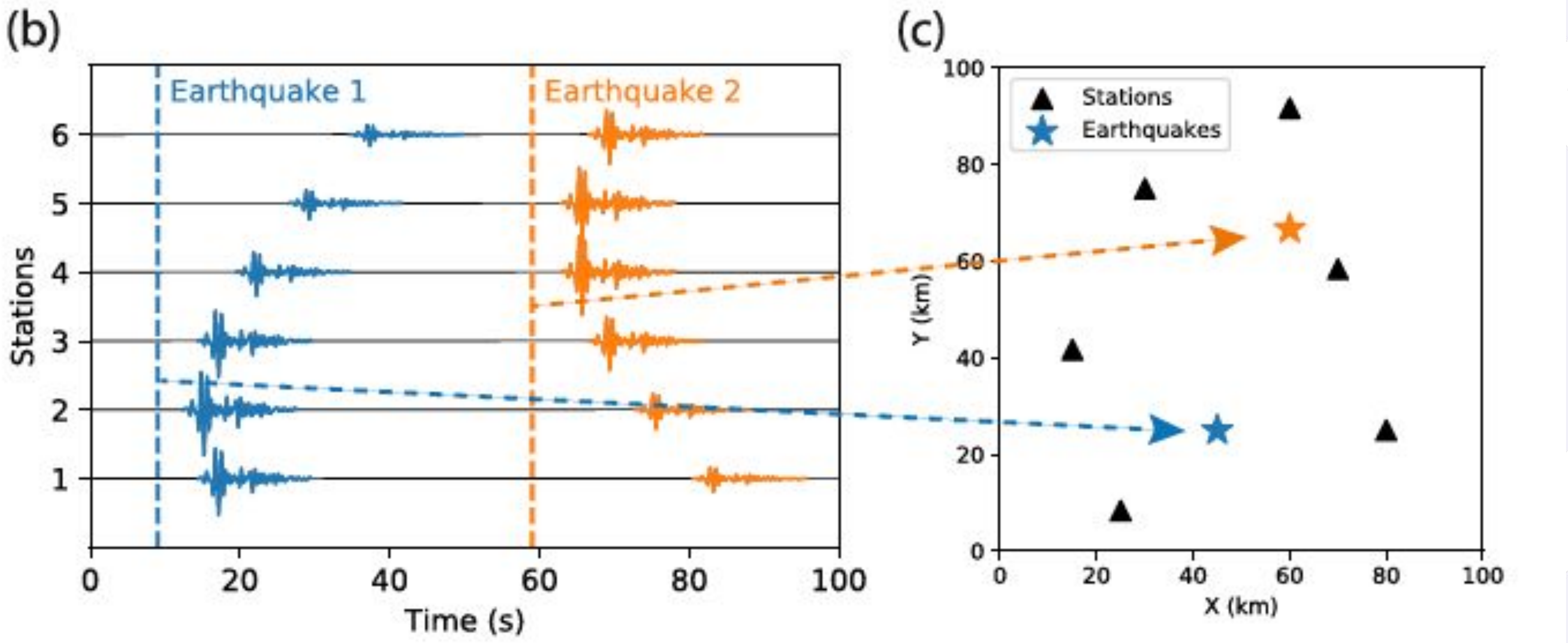
1. Labquake and tectonic applications overview

2. Transfer Learning to Self-Supervised Learning
   a) Lessons from the lab
   b) Automatic Speech Recognition for Ground Displacement Predictions

3. Labquake Task Specific Application
   a) PINN Applied to Rate-and-State friction

4. **Phase Detection Task Specific Application**
   a) **Training and Validation**
   b) **Base Model Contribution**

Earthquake phase detection with wav2vec2 encoded waveforms



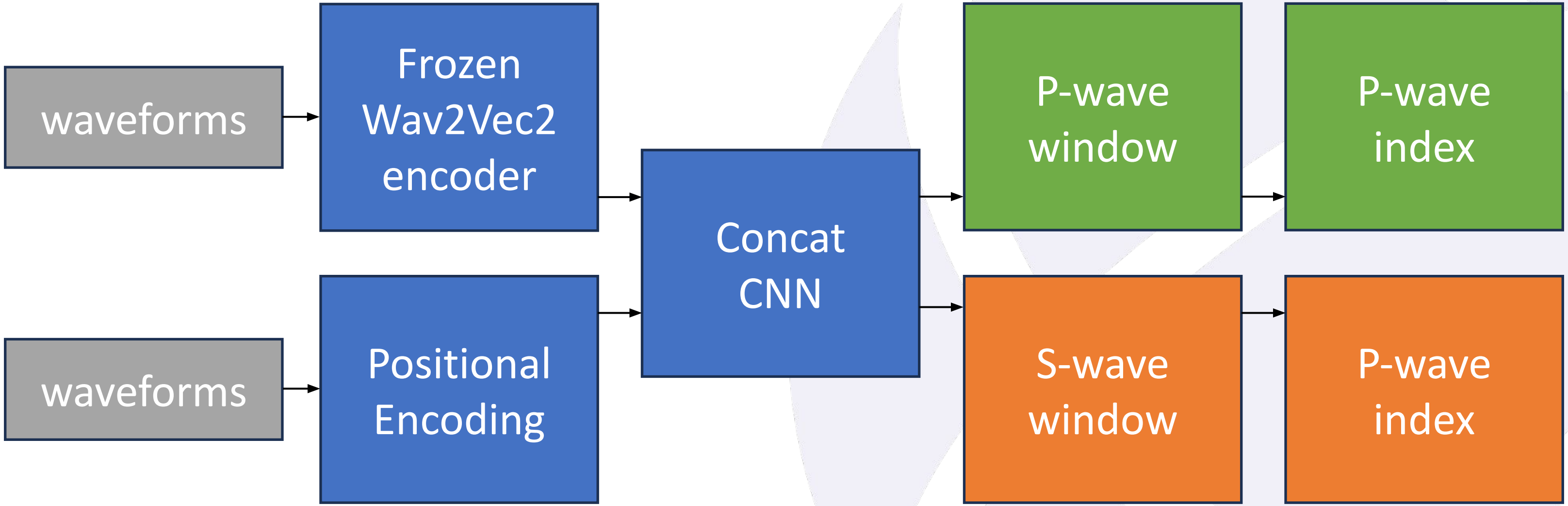
From Zhu et al. 2021

Earthquake phase detection needs labeled phase arrival times for training

Organize 618,000 waveforms from SCEDC with P/S arrivals
429 stations from 2000-2019
3 minutes (18,000 pts)

Labeled data sets are available, e.g. STEAD or INSTANCE, and used for testing

Phase Detection Task Specific Application

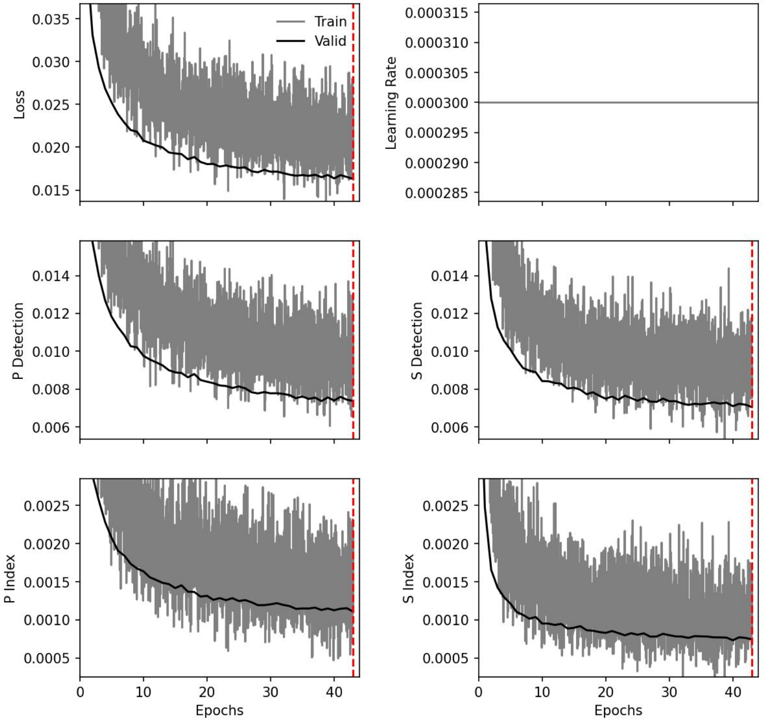


~440k trainable parameters

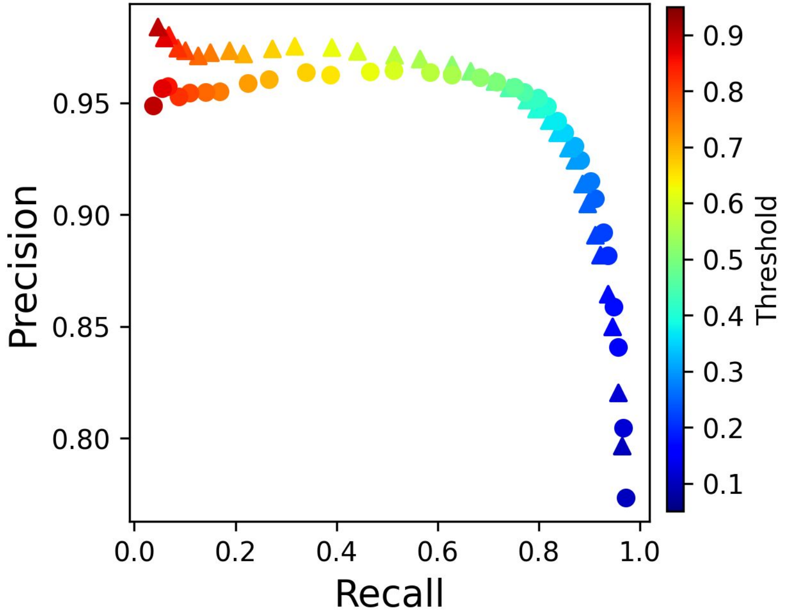
Earthquake phase detection with wav2vec2 encoded waveforms

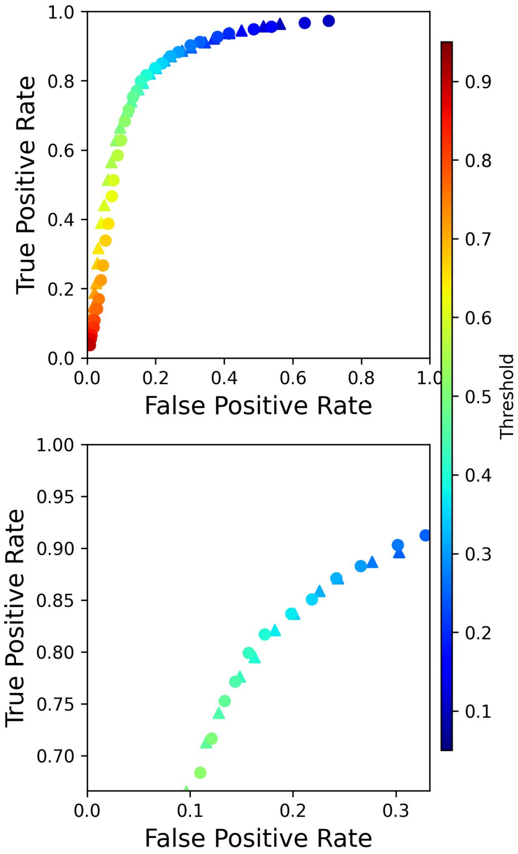
300k training example
50k validation examples

Train detection model 45 epochs



STEAD metrics for every 1.28 second window containing an arrival

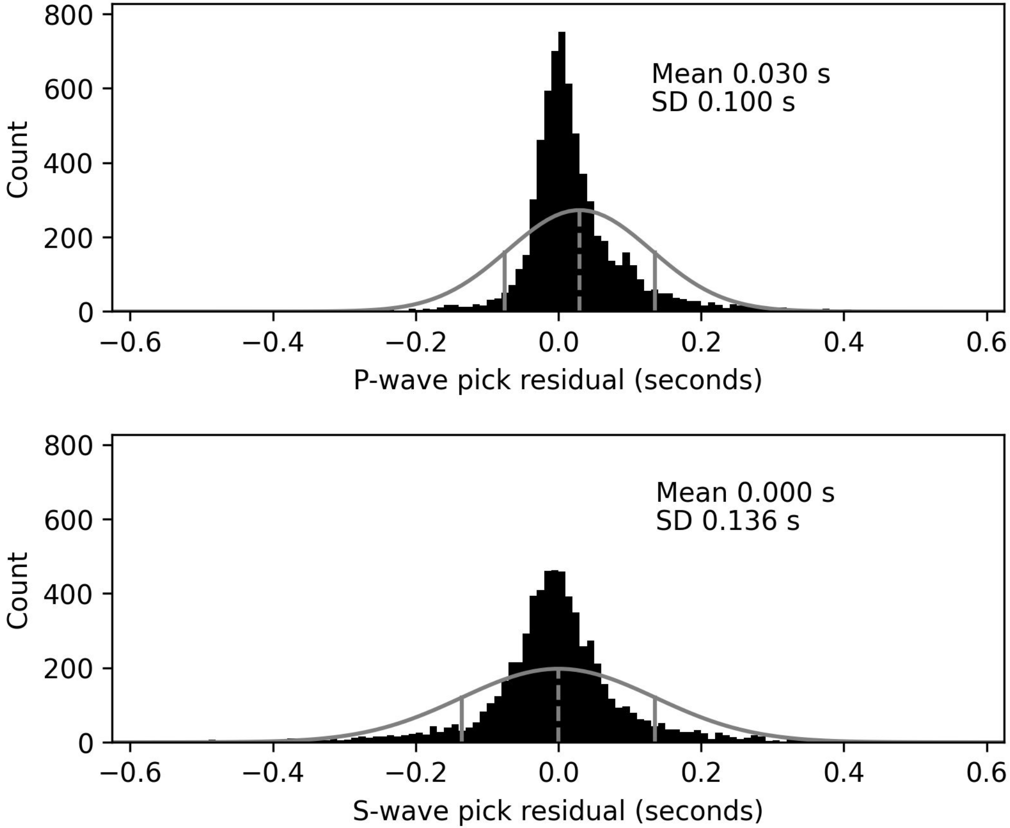




Los Alamos NATIONAL LABORATORY

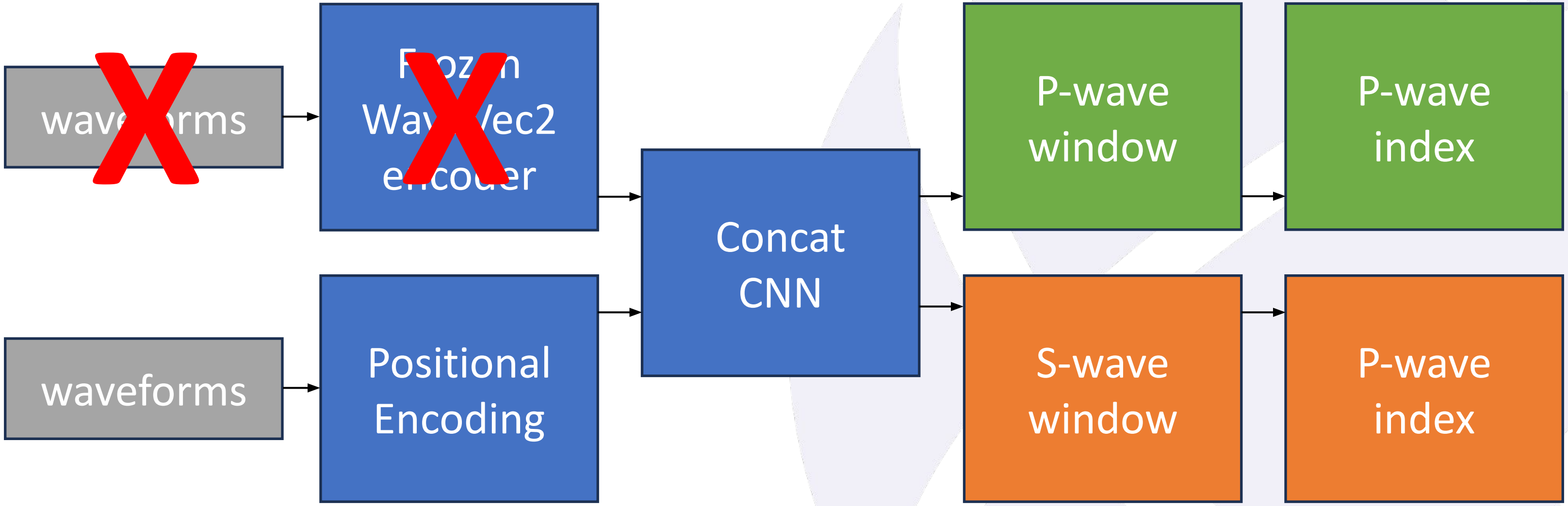
Johnson, et al. in-prep

P and S arrival time residuals for 0.35 threshold



Phase Detection Task Specific Application

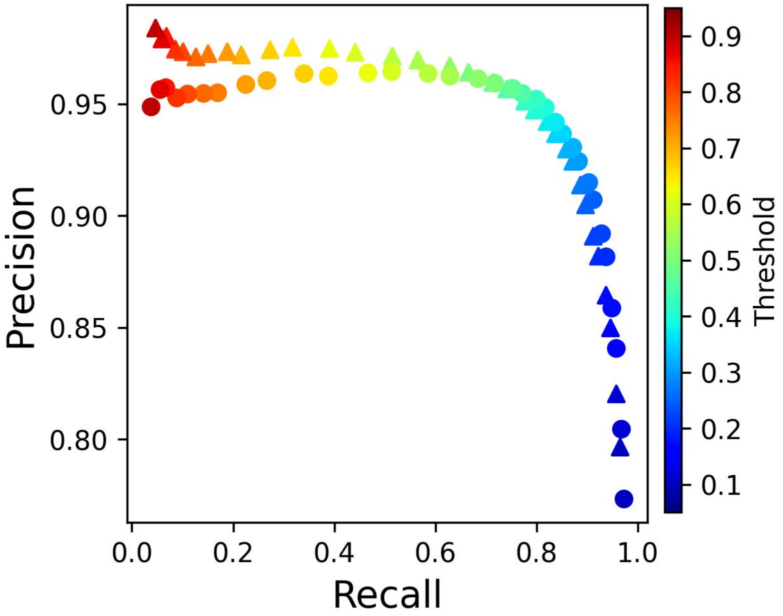
What is gained from pretraining?



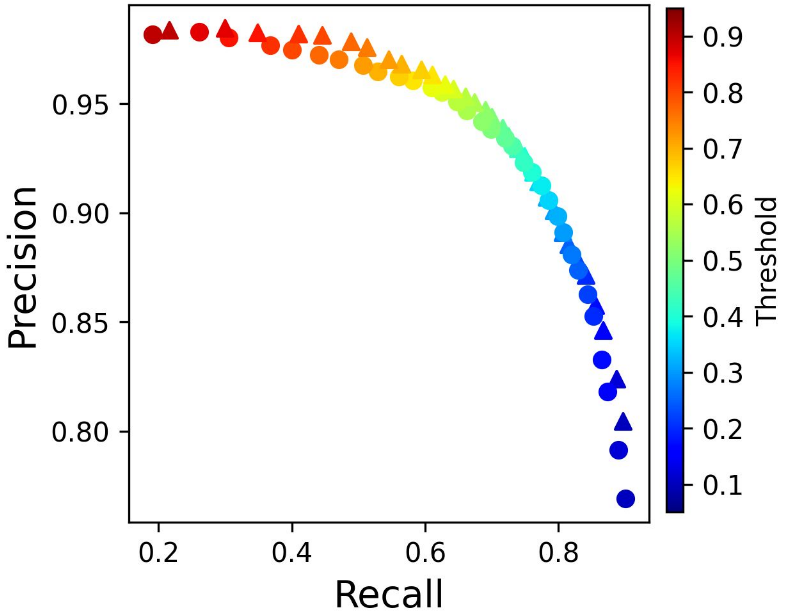
~440k trainable parameters

STEAD metrics for every 1.28 second window containing an arrival

Wav2Vec2

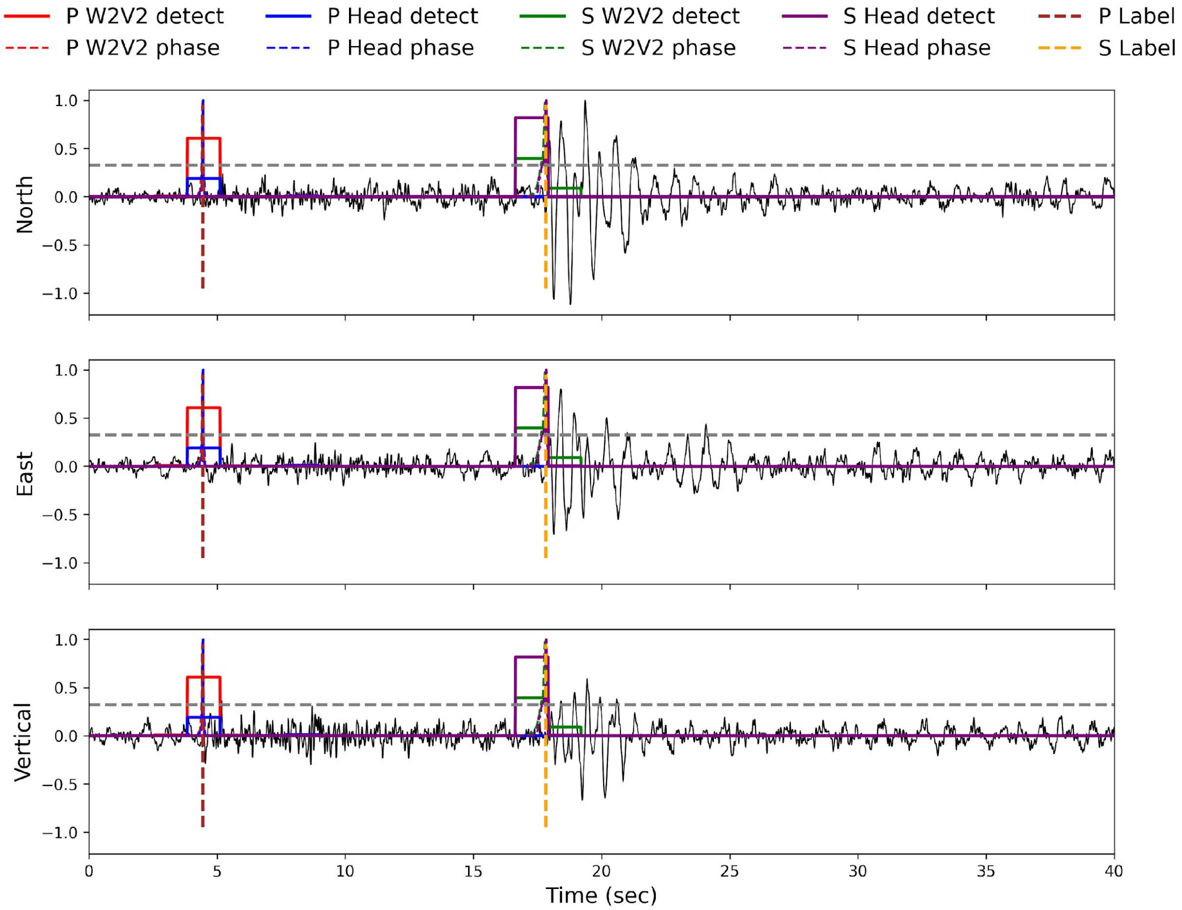


No encoding to detection model

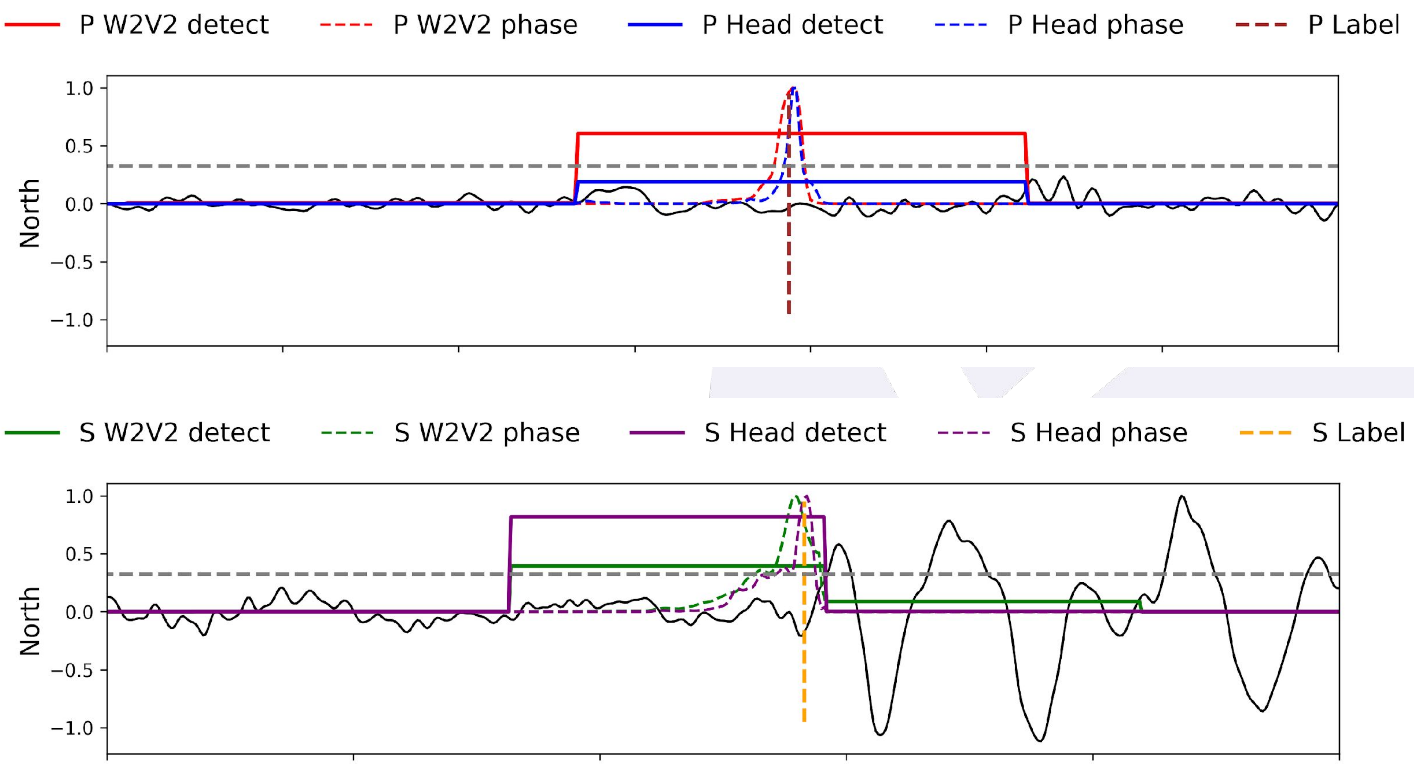


Johnson, et al. in-prep

Example where models perform nearly identical

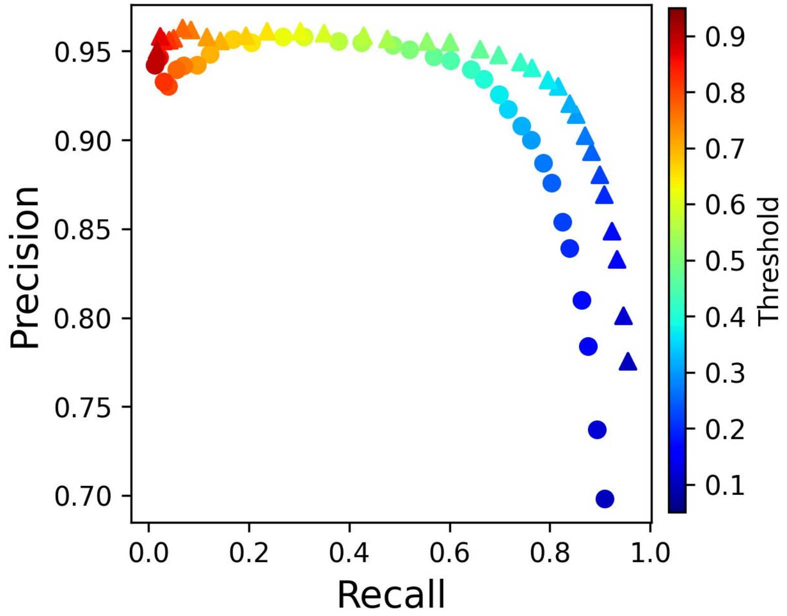


Example where models perform nearly identical

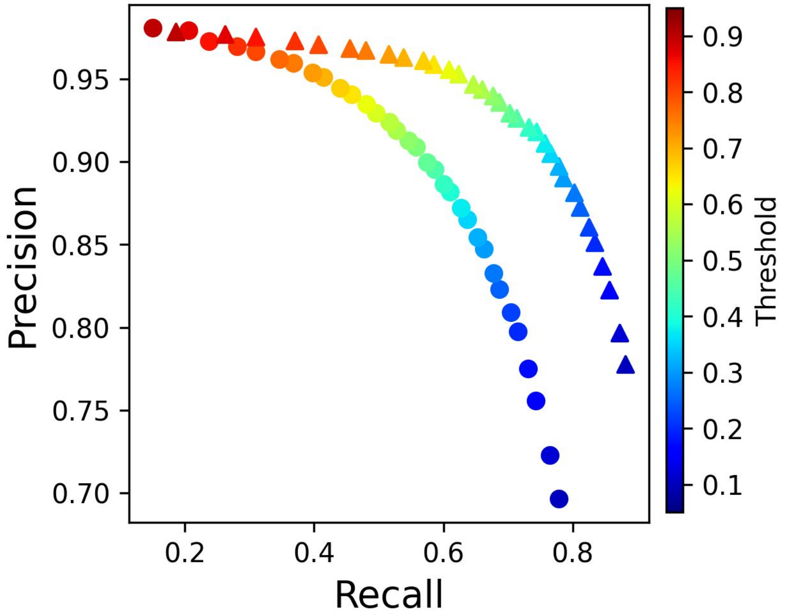


Apply augmentation to lower SNR and Wav2Vec2 maintains performance

Wav2Vec2



No encoding to detection model



Johnson, et al. in-prep

Summary

- Components of foundation models are applicable to geophysical time series analysis
- Audio speech recognition encoder adaptable to seismic data with or without modifications
- Mapping continuous seismic waveforms to Physics Informed Machine Learning providing insight to what the model is extracting
- Foundation model encoders applicable to seismic data processing, e.g., phase picks, etc., are showing progress

Questions?

All work shown is in collaboration with:
Paul Johnson
Robert Guyer
Laura Laurenti
Chris Marone

Wang, K., C.W. Johnson, K.C. Bennett, and P.A. Johnson, *Predicting Future Laboratory Fault Friction Through Deep Learning Transformer Models*. Geo. Res. Letters, 2022.

Wang, K., C.W. Johnson, K.C. Bennett, and P.A. Johnson, *Predicting fault slip via transfer learning*. Nature Comm., 2021. 12(1): p. 7319.

Johnson, C.W. and P.A. Johnson, *Learning the Low Frequency Earthquake Activity on the Central San Andreas Fault*. Geo. Res. Letters, 2021.

Johnson, C.W. and P.A. Johnson, *Seismic Features Predict Ground Motions During Repeating Caldera Collapse Sequence*. Geo. Res. Letters, 2024.

Johnson, C.W., K. Wang, and P.A. Johnson, *Automatic speech recognition predicts contemporaneous earthquake fault displacement*. Nature Comm., 2025.

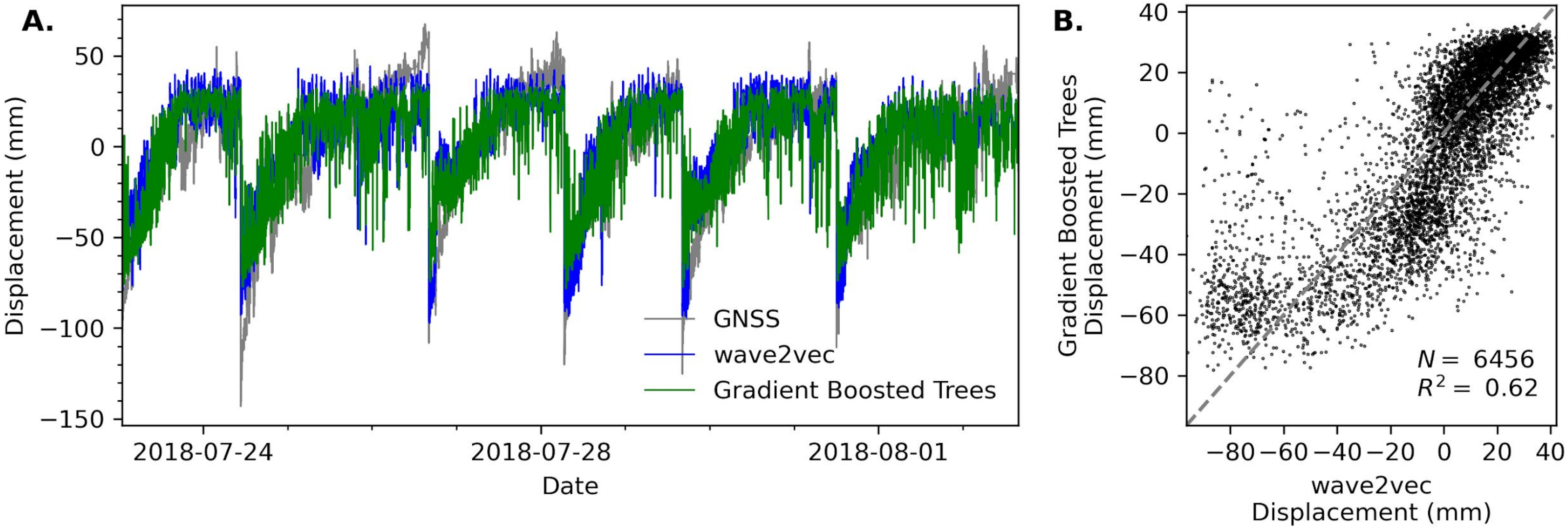
U.S. DEPARTMENT OF ENERGY | Office of Science

Basic Energy Sciences (BES)

Automatic speech recognition is related to Natural Language Processing

Goal of NLP is to "understand" the language and generate response

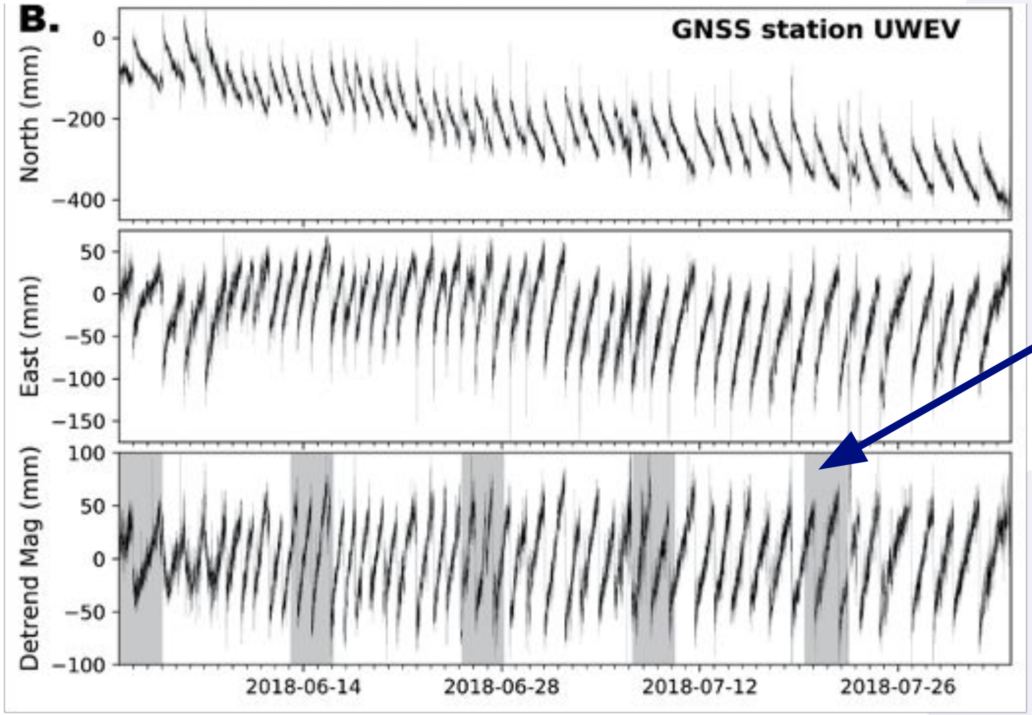
Compare with shallow architecture model with engineered features



Green is GBT and results are good, but …

Johnson & Johnson, 2024
Johnson et al. 2025

Gradient boosted tress unable to handle properly predict nonstationary data



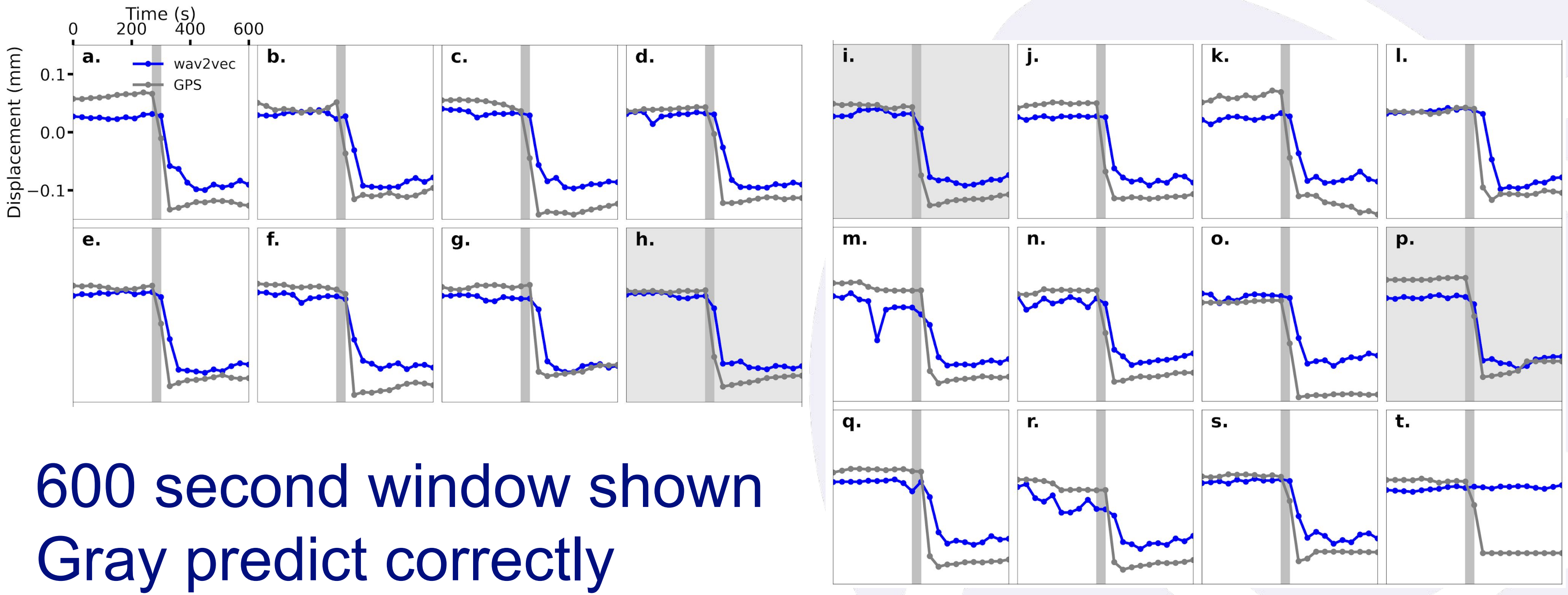
Remove all impulsive signals

Block sample for train-val-test splits

Tree based methods do not generalize

Johnson & Johnson, 2024

Reformulate the data for predictions of future onset of slip for 20 collapse events



600 second window shown

Gray predict correctly

Not fully successful but 3 were correct

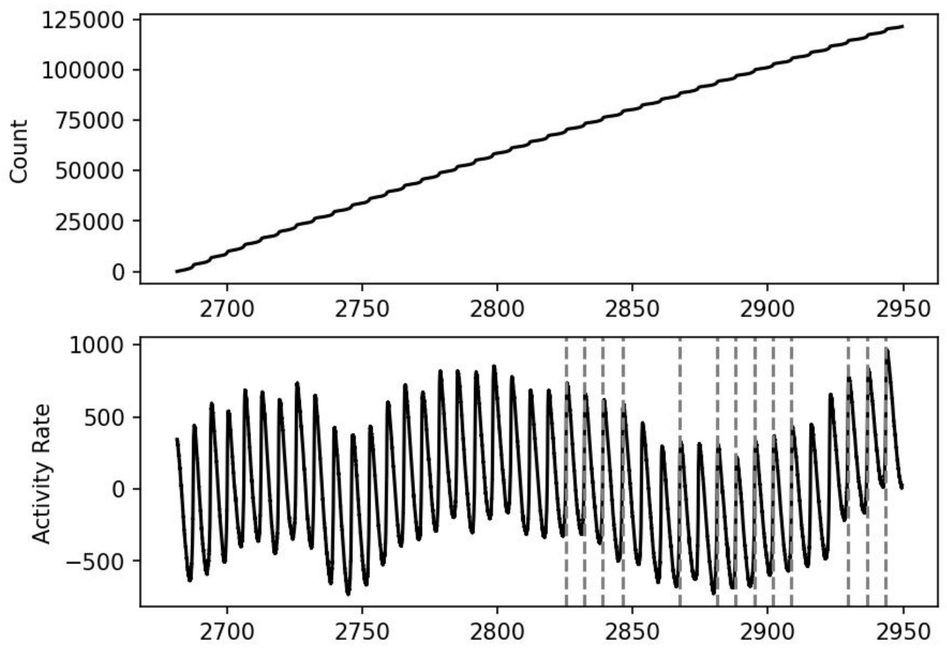
Johnson et al. 2025

Continuous waveforms from locked system in laboratory experiment

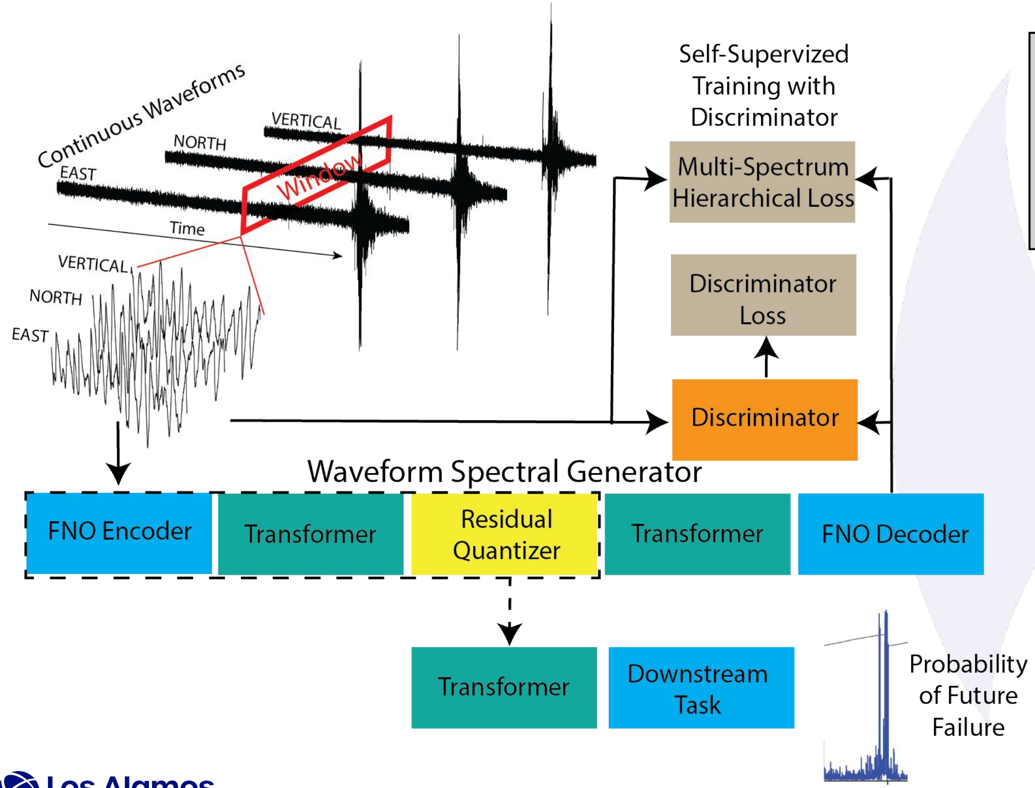
Acoustic emission datable throughout loading cycle

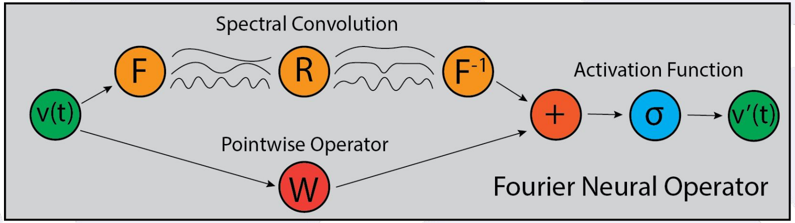
Near constant rates

~125,000 events detected



Redesign encoder-decoder model for spectral feature extraction





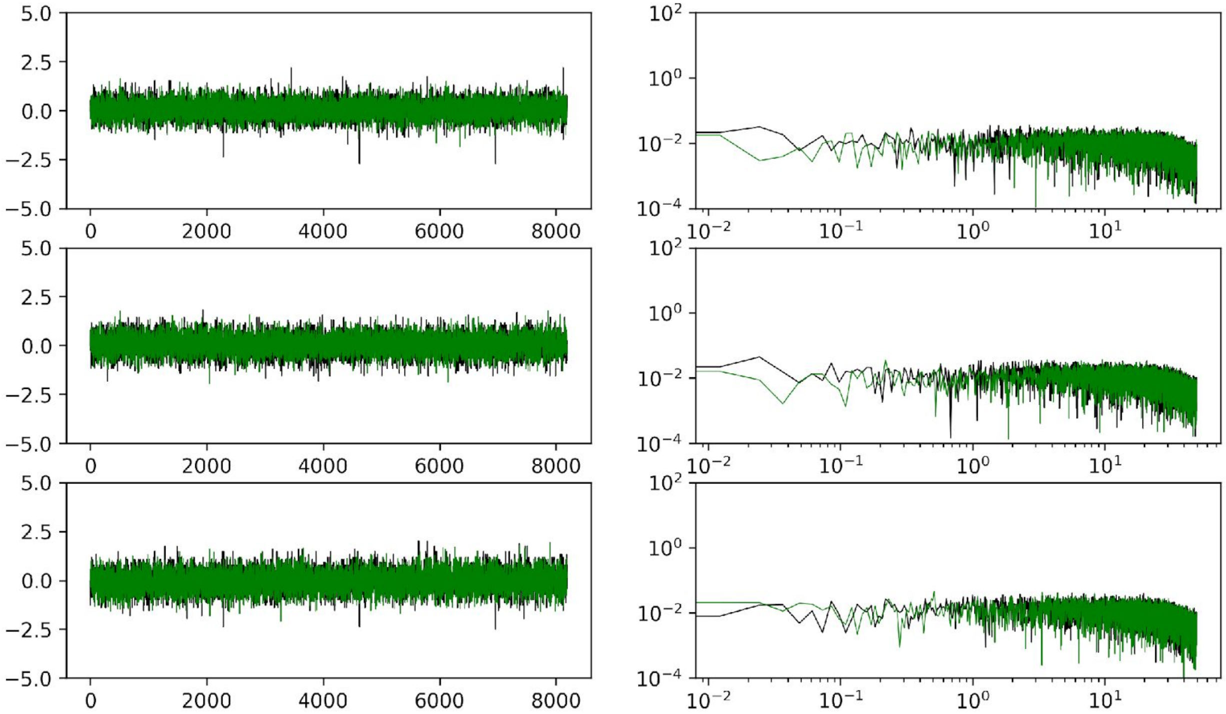
Replace CNN filters in time domain
Operate on waveforms in spectral domain

Johnson, et al. in-prep

Encoder-decoder training with FNO

Input – black

Reconstructed - green

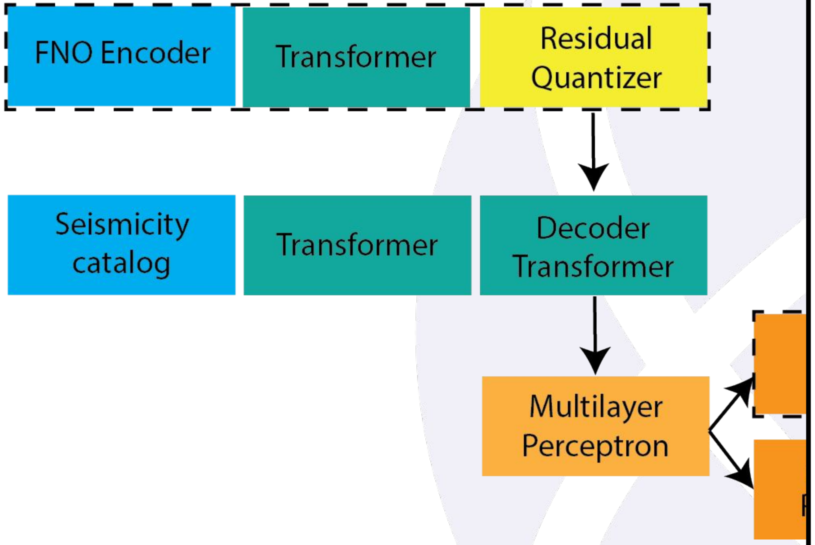


Physics Informed Machine Learning (PIML) using encoded waveforms

Attention is all you need
Vaswani et al., 2017

Input waveforms and catalog

Lab experiment with granite on granite locked stick-slip system



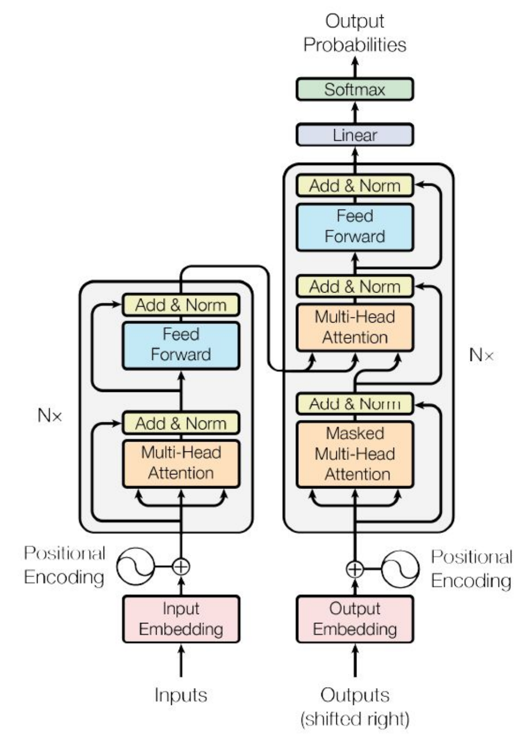
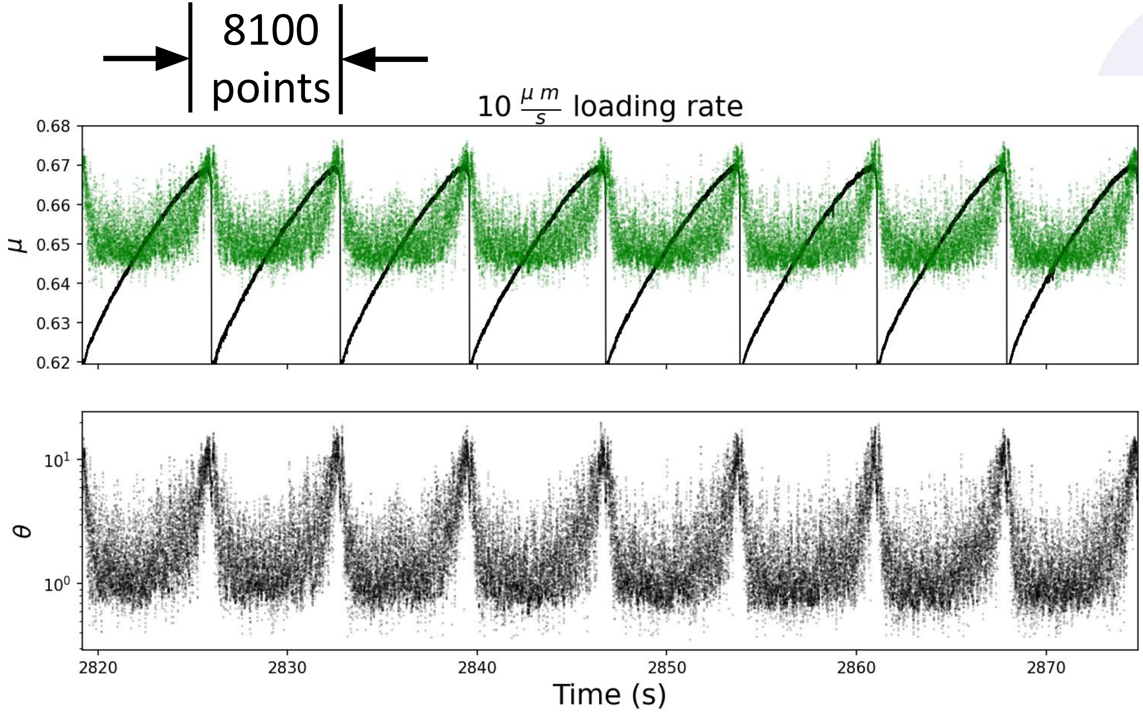


Figure 1: The Transformer - model architecture.

Johnson, et al. in-prep

Continuous waveforms with transformer model inform PIML neural network



$$\mu = \mu_0 + a\ln\left(\frac{V}{V_0}\right) + b\ln\left(\frac{V_0\theta}{D_c}\right)$$

Learned Model parameters

a = 5.1e-03

b = 9.4e-03

μ_0 = 0.64

V = 1.1e-04

Dc = 1.9e-05

Waveform is 8192 in length to resolve one ϕ and μ

Johnson, et al. in-prep

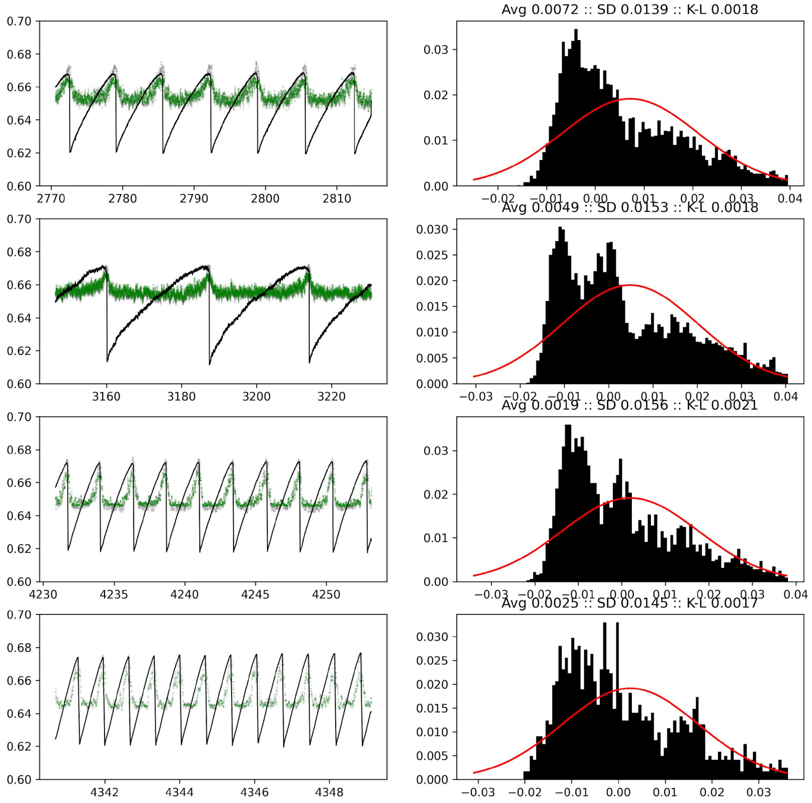
Consistent results with varying loading velocity

10 μm/s

3 μm/s

30 μm/s

100 μm/s



Locked granite-granite

Load point displacement at 0.0m during loading

Learning failure time from weak acoustic emission

Unable to resolve early loading period