Feature-based Magnitude Estimates for Small, Nearby Earthquakes in the Yellowstone Volcanic Region

Alysha Armstrong, Ben Baker, Keith Koper

UNIVERSITY OF UTAH SEISMOGRAPH STATIONS

INTRO

- Conventional magnitude methods can fail or be prohibitively time consuming during periods of high seismicity rates, such as the many earthquake swarms in Yellowstone.
- We introduce a machine learning method that uses features derived from short-duration waveform segments of individual phase arrivals and event source parameters to predict local magnitude (M_L).

KEY POINTS

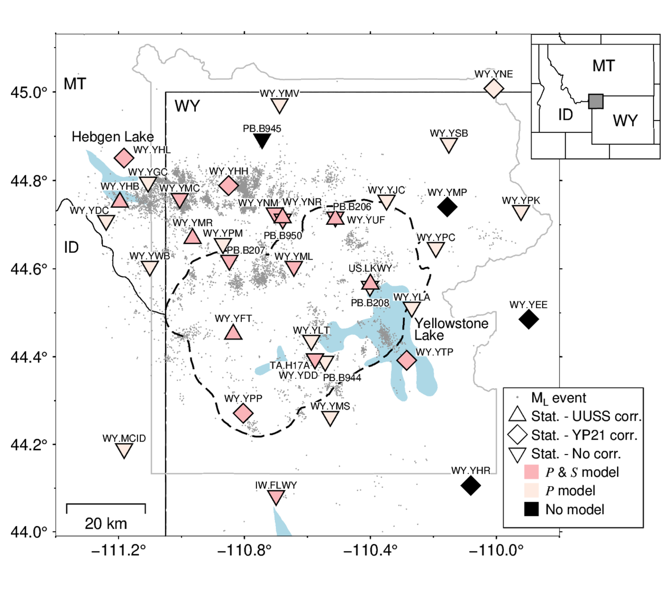
Our approach:

- accurately estimates M_L for ~0–3.5
- works for events with small temporal separation
- does not require 3C broadband stations for M_L
- is easily interrogated, updated, and modified
- maintains consistency with the UUSS catalog
- requires the event source location
- can take advantage of the abundant phase arrivals, particularly *S*, available in deep-learning enhanced catalogs

DATA

- **Model targets**: High-quality M_L values
- **Model inputs**: Features derived from 0.95–1.40 s of pre-arrival noise and 2.55–3.60 s of post-arrival signal.
  - Separate feature datasets for *P* and *S* arrivals
  - Start with 38 frequency-domain, 4 time-domain, and 3 location-based candidate features:

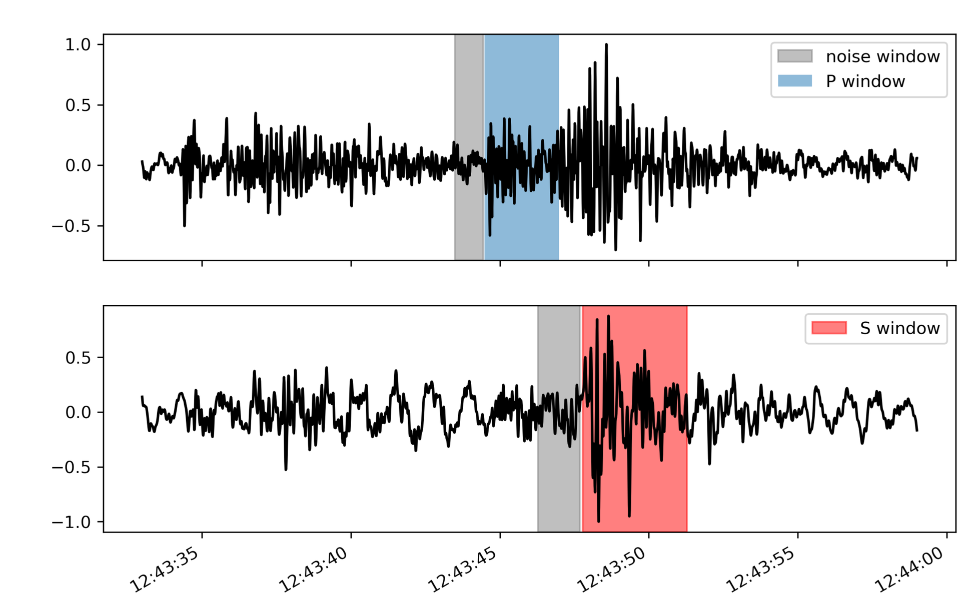
Name	Abbreviation	Equation or explanation	Transform	Type
Amplitude Ratio	ratio [freq.]	The ratio of the average signal and average noise at the specified corner frequency (freq.) between 1–18 Hz	$\log_{10}$	Time/Freq.
Average Amplitude	amp. [freq.]	The average signal at the specified corner frequency between 1–18 Hz	$\log_{10}$	Time/Freq.
Signal Dominant Frequency	sig. dom. freq.	The dominant frequency in Hz of the phase arrival	$\log_{10}$	Freq.
Signal Dominant Amplitude	sig. dom. amp.	The maximum amplitude of the signal dominant frequency	$\log_{10}$	Freq.
Signal Maximum Amplitude	sig. max. amp.	The difference of the maximum signal amplitude and the minimum signal amplitude	$\log_{10}$	Time
Noise Maximum Amplitude	noise max. amp.	The difference of the maximum amplitude and the minimum amplitude in the noise window	$\log_{10}$	Time
Signal Variance	sig. var.	The variance of the signal time series from zero	$\log_{10}$	Time
Noise Variance	noise var.	The variance of the noise time series from zero	$\log_{10}$	Time
Source-receiver Distance	distance	The distance from the event epicenter to the receiver in km	$\log_{10}$	Event
Source-receiver Back Azimuth	back az.	The distance from the receiver to the event epicenter in degrees. If using a linear model, the sine is used	sine (if linear model)	Event
Source Depth	depth	The depth of the event in km relative to sea-level	-	Event



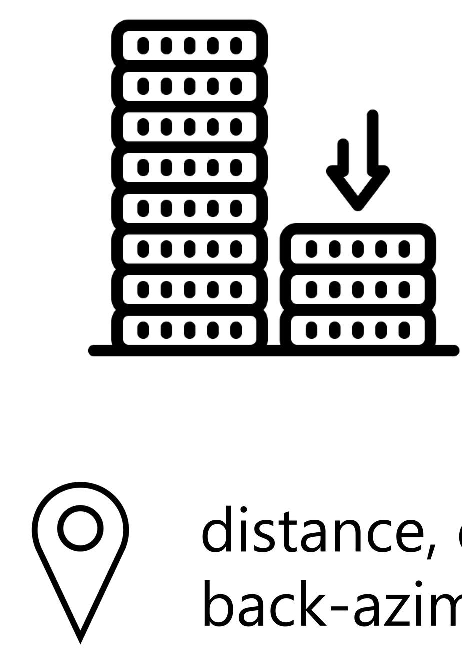
Using the **earthquake location** & **4 waveform features** we can accurately compute event magnitude during periods of **elevated seismicity**.

Overview:

1. We **extract time and frequency candidate features** from **short-duration** waveform segments around individual *P* and *S* arrivals.



2.

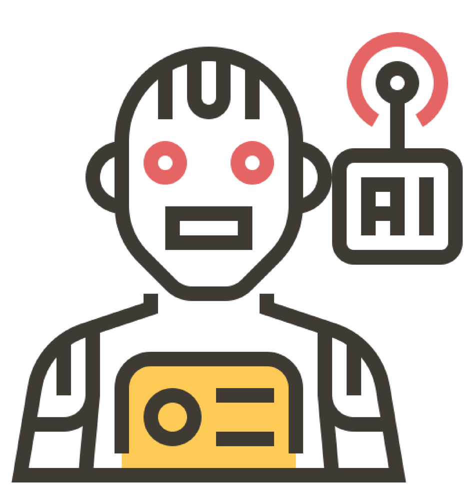


We use our generalizable approach to **reduce the number of features** from 45 to 7:

distance, depth, back-azimuth

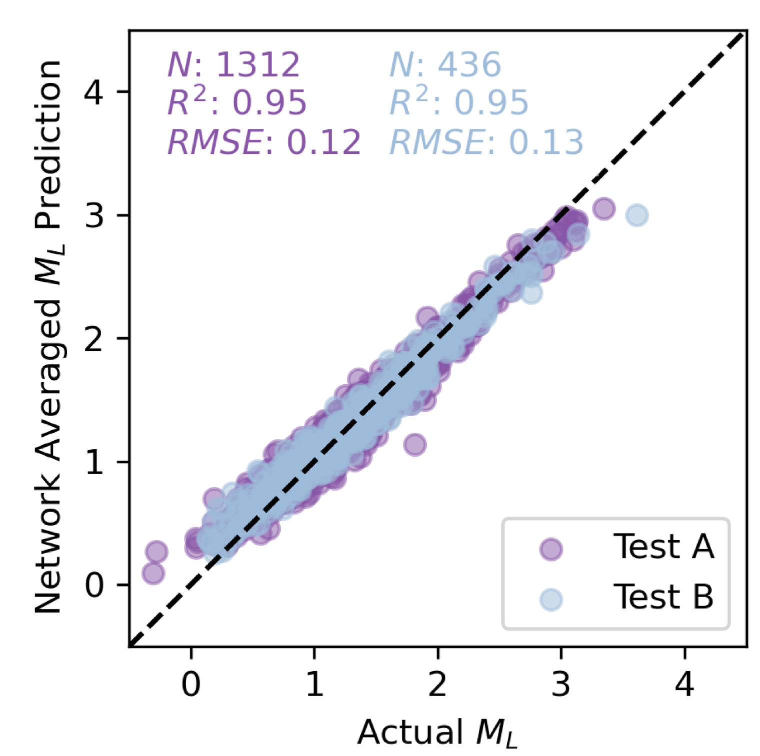
amplitude/SNR proxies: 3 signal & 1 noise

3.



We **train a machine learning model** for each stations to **predict local magnitude** using the 7 selected features.

4.



We average all model predictions to create network magnitudes, which **are generally within ~0.13 mu** of the actual magnitudes.

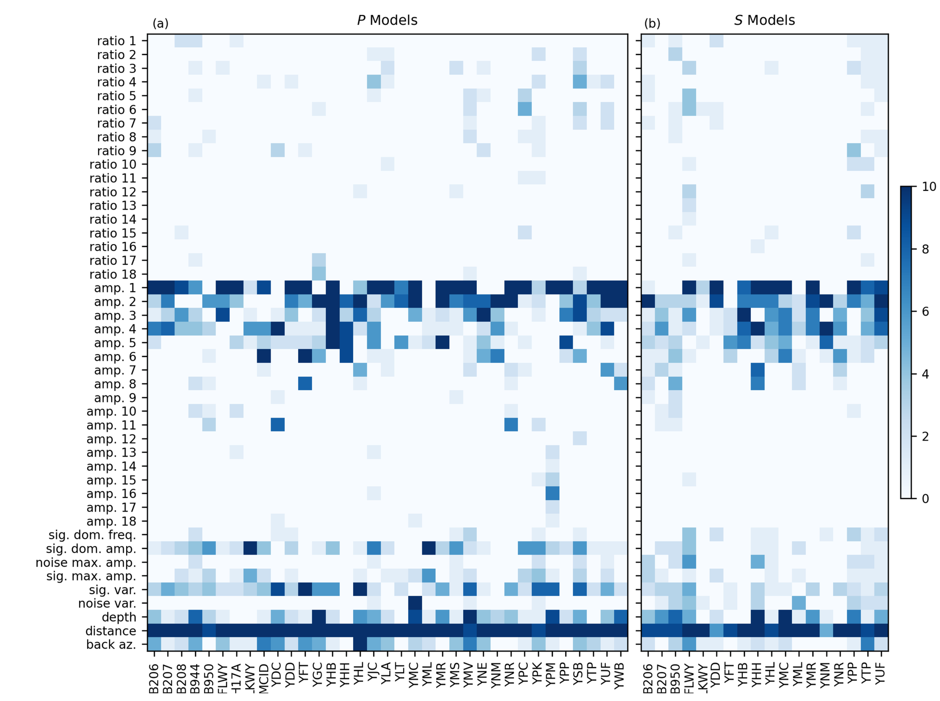
METHODS

Training and Testing Datasets

- Use an 80:20 split of all events in the *P*- & *S*-feature catalogs occurring before 1/1/23 for the training set and **testing set A**.
- Use features computed from events during 1/1/23–1/1/24 as **testing set B**.
- 35 stations in the *P* dataset & 18 in the *S* dataset
- *P* stations have ≥ 300 & *S* stations have ≥ 150 training examples

Recursive Feature Elimination Algorithm (RFEA)

- While our candidate features are designed to be related to magnitude, many of them are redundant (i.e., highly correlated) or may be uninformative.
- We use a RFEA that both simplifies and improves the predictive performance of the machine-learning models and limits feature selection bias.
- Our RFEA is a two-step process, in which we first identify the most important features at each station and then **select a common feature set for all stations**.
- Selected waveform features:
  - *P:* amplitude at 1 & 2 Hz, signal & noise variance
  - *S:* amplitude at 1, 2 & 4 Hz, noise variance



Support Vector Machines (SVM)

- We use an SVM with radial basis function kernel to learn a mapping from the features to M_L
- We train one model per station-phase pair
  - *P*: 35 models, *S*: 18 models
- SVMs are used in our RFEA

RESULTS