

Application of UMAP and XGBoost to Pg and Lg Ratios to Identify Moroccan Mining vs. Earthquake Events

K. Goddard¹, S. Badrane², M. Menzhi², C. Saikia¹

1. Air Force Technical Application Center

2. Centre National pour la Recherche Scientifique et Technique



INTRODUCTION AND MAIN RESULTS

We selected 231 mining and 143 earthquake events recorded on seismic station MD31 from the CNRST bulletin. These waveforms were used to construct raw vertical component $\log_{10}(\text{RMS})$ Pg/Lg ratios and used to train an XGBoost machine learning model. When higher frequency ratios are used a 99% accuracy is achieved.

The best performing model was ran on two other data sets:

- 525 unlabeled events from the mining region
- 45 labeled EXPs and EQKs recorded on seismic station AVE

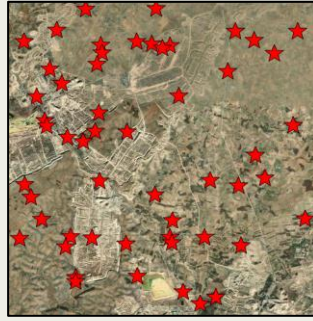
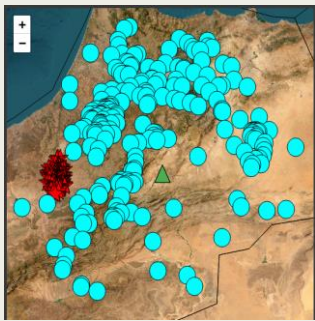


K. Goddard, S. Badrane, M. Menzhi, C. Saikia

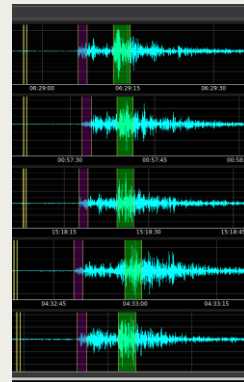
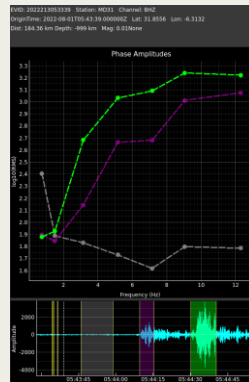
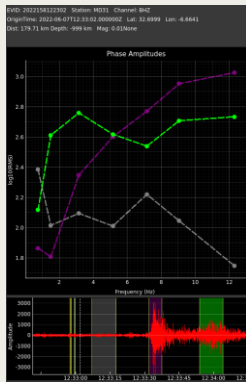
P3.5-408

Event and Feature Selection

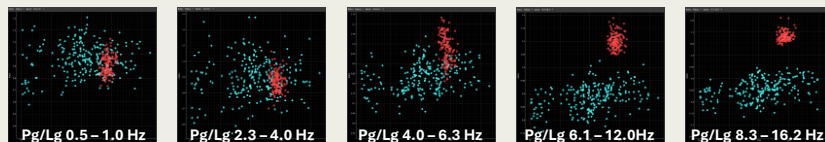
Using the event origin times and locations from the CNRST bulletin (2014 – 2023), 143 explosion and 231 earthquake raw vertical component log10(rms) Pg/Lg ratios were calculated for events recorded on MD31.



Measuring Amplitudes



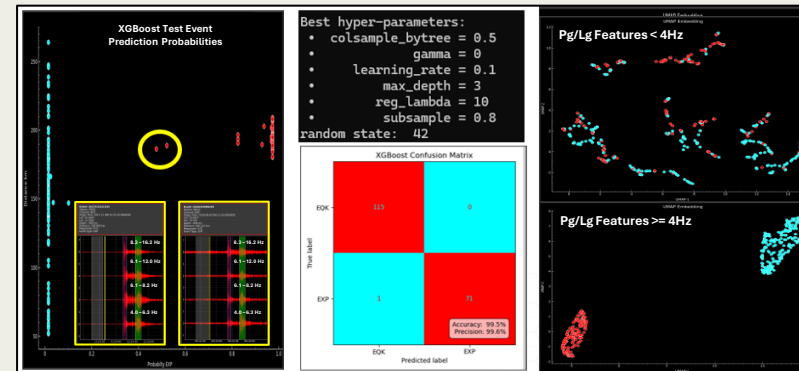
Ratio Features



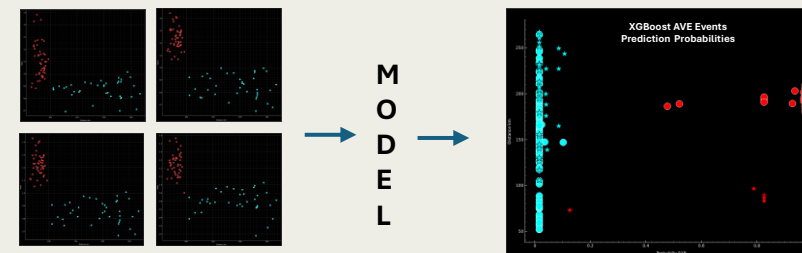
UMAP Clustering and Model Training

Unlabeled high-frequency ratios (>4 Hz) were analyzed with UMAP, revealing clear separation between mining and non-mining events.

A 50/50 train-test split was used to train an XGBoost model, with hyperparameters optimized via grid search. The final model achieved 99% accuracy and precision, with high confidence for nearly all events; only two had ~50% probability.



To test transferability, the model was applied to 43 explosion and 38 earthquake events from a different station (AVE), correctly classifying all. Red and blue stars show predicted probabilities for each event.



Earthquakes in the Mining Region?

525 unlabeled MD31 mining-region events were ran through the trained model. While most were confidently classified as mining, several had $\geq 90\%$ probability of being earthquakes. 6–12 Hz waveform analysis of these events showed earthquake-like signatures, suggesting the presence of earthquakes in the mining region. This finding highlights the model's potential to uncover previously unrecognized seismic sources within mining-dominated regions, providing valuable insight for refining event catalogs and improving regional seismic monitoring.

