SAND2018-13461C

Sandia National Laboratories

National Nuclear Security Administration

Classification of Local Seismic Events in the Utah Region: A Comparison of Amplitude Ratio Methods with a Machine Learning Approach

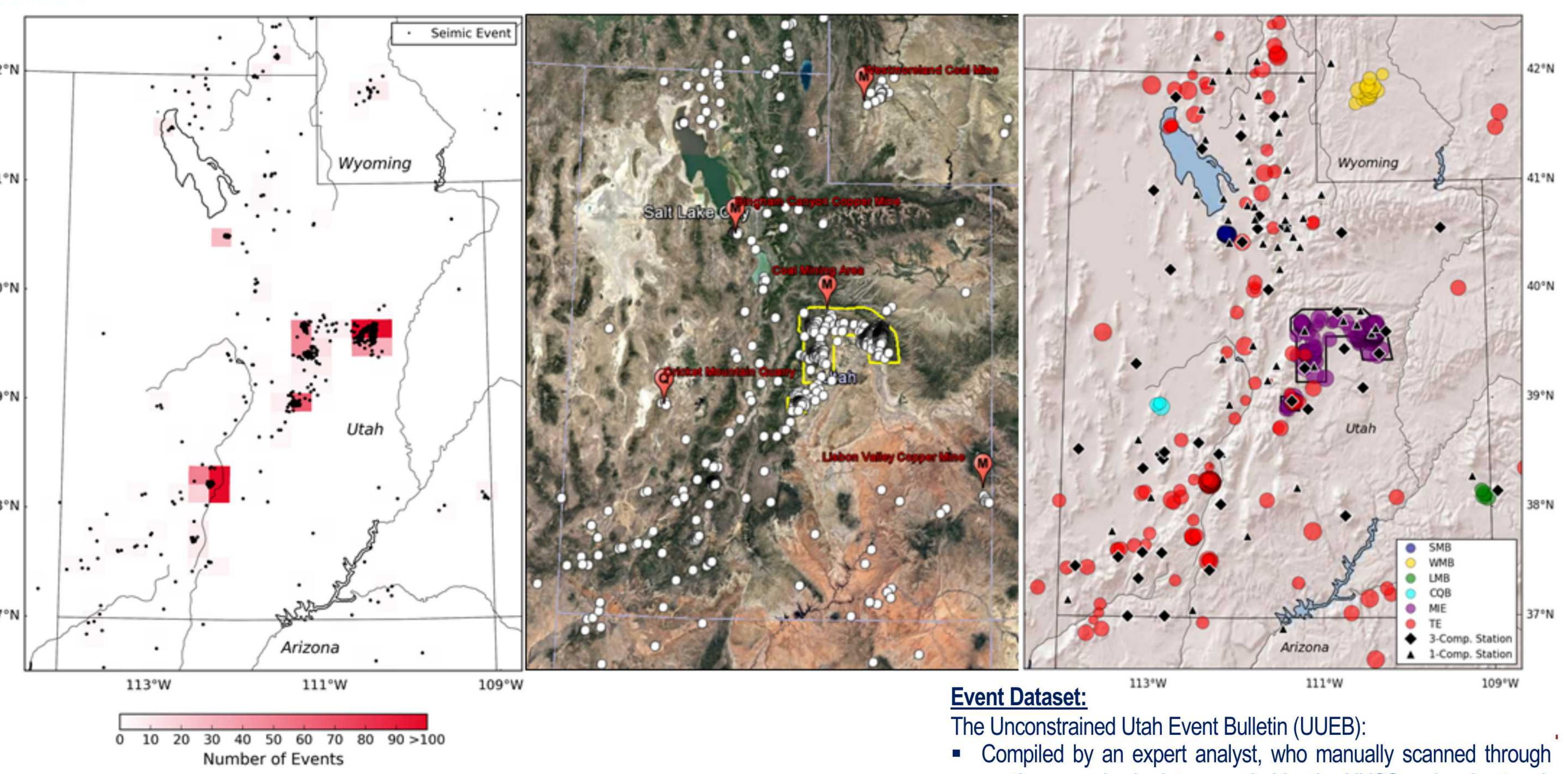
Lisa Linville[1], Rigobert Tibi[1], Christopher Young[1], and Ronald Brogan[2] (*Email: rtibi@sandia.gov*)

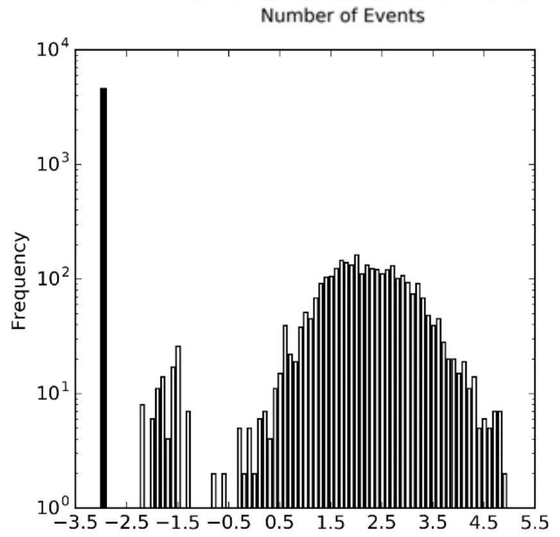
[1]) *Sandia National Laboratories;* [2]) *ENSCO, Inc., Falls Church, VA*

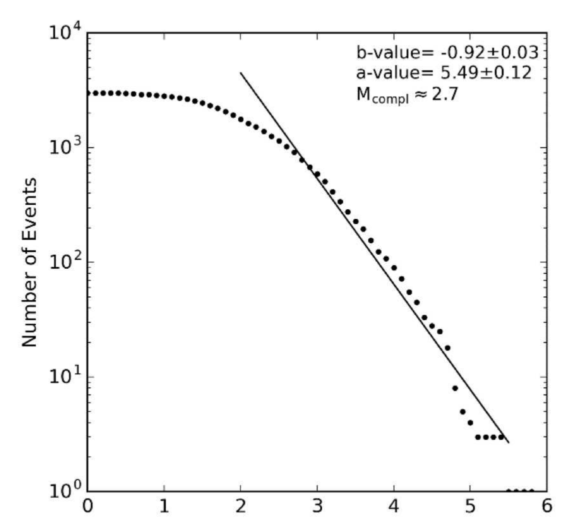
OVERVIEW

The capability to discriminate low-magnitude earthquakes from low-yield anthropogenic sources, both detectable only at local distances, is of increasing interest to the event monitoring community. We used a dataset of seismic events in Utah recorded during a 14-day period (1–14 January 2011) by the University of Utah Seismic Stations (UUSS) network to perform a comparative study of event classification at local-scale using amplitude ratio (AR) methods and a machine learning (ML) approach. We compare the AR and ML methodologies using a broad set of criteria and conclude that a major advantage to machine learning methods is their robustness to low signal-to-noise ratio (SNR) data, allowing them to classify significantly smaller events.

DATA







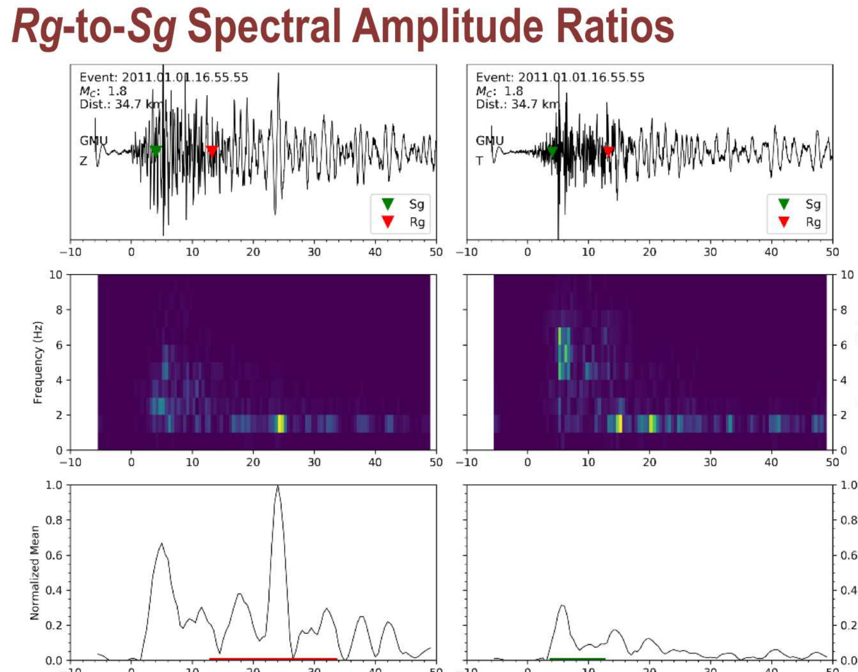
Event Dataset:

The Unconstrained Utah Event Bulletin (UUEB):

- Compiled by an expert analyst, who manually scanned through continuous seismic data recorded by the UUSS regional network from 1 to 14 January 2011.
- Consists of 7,377 event within or around the state of Utah with M_C of ~2.0 and lower up to 5.8. Data for ~58% of the events have insufficient SNR. Hence their magnitudes could not be estimated. These events are not included in the calculation of the magnitude of completeness.
- Six event categories identified based on the event locations:
  - 1,040 tectonic earthquakes (TEs) located mostly within the ISB
  - 6,286 mining-induced events (MIEs) in coal mining regions
  - 27 ripple-fired mining blasts from Bingham Canyon mine (SMBs)
  - 5 mining blasts from Lisbon Valley mine (LMBs)
  - 16 mining blasts from Westmoreland mine (WMBs)
  - 3 quarry blasts from the Cricket Mountain quarry (CQBs)

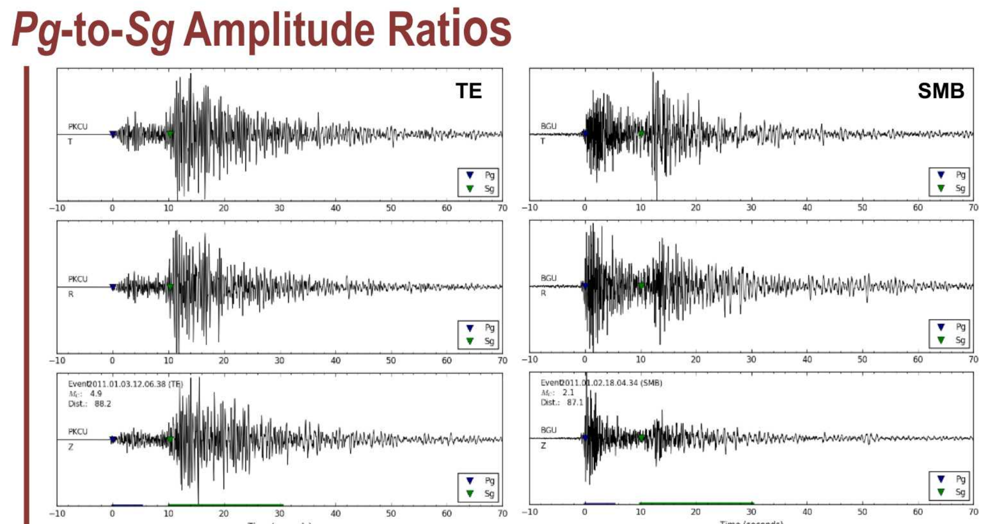
METHODS

Rg-to-*Sg* Spectral Amplitude Ratios



From the Z-component spectrogram, a mean time series is generated for the frequency range of 0.5–2 Hz. A similar time series is generated from the T-component spectrogram for the frequency range of 0.5–8 Hz. The time series are then smoothed. The maximum amplitude in the *Rg* window (*max*$\{Rg_{[0.5-2Hz]}\}$) of the Z time series is estimated. The same is done in the *Sg* window of the T time series to obtain (*max*$\{Sg_{[0.5-8Hz]}\}$). These amplitudes are then corrected for the propagation effects. The reported *Rg*-to-*Sg* ratios consists of the corrected *max*$\{Rg_{[0.5-2Hz]}\}$ divided by the corrected *max*$\{Sg_{[0.5-8Hz]}\}$.

Pg-to-*Sg* Amplitude Ratios



For each frequency in the range of 1–15 Hz in 1-Hz increment: $A_{P_g} = (P_Z^2+P_R^2)^{1/2}$, where P_Z and P_R are rms of the amplitudes in the *Pg* window on the Z and R components, respectively; $A_{S_g} = (S_Z^2+S_R^2+S_T^2)^{1/2}$, whereby S_Z, S_R, and S_T are the rms of the amplitudes in the *Sg* window on the Z, R and T components, respectively. A_{P_g} and A_{S_g} are then corrected for the propagation effects. The reported *Pg*-to-*Sg* ratios are obtained by dividing the corrected A_{P_g} by the corrected A_{S_g}.

To allow sufficient separation and avoid interferences between the phases, only data recorded at distances from 25 to 150 km were used for both *Rg*-to-*Sg* and *Pg*-to-*Sg* ratios.

The authors acknowledge the support of the National Nuclear Security Administration Office of Defense Nuclear Nonproliferation Research and Development for funding this work.
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC,
a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Multivariate Quadratic Discriminant Analysis

Assumptions:

- $f_j(\boldsymbol{r})$ with $j = 1, 2, \ldots, N$ is the probability density associated with the *j*-th population, whereby $\boldsymbol{r} = (r_1, r_2, r_3, r_4, r_5)^T$ is the vector of amplitude ratios. The variables r_1 to r_4 are *Pg*-to-*Sg* ratios at 4 discrete frequencies, while r_5 represents the *Rg*-to-*Sg* ratio. The characteristics of $f_j(\boldsymbol{r})$ is inferred from the learning set.
- Amplitude ratios of the training sets are normally distributed within each population
- Events in the training sets are correctly classified.
- The *N* populations have equal prior probability of occurrence ($\pi = 1/N$).
- The *N* populations have equal misclassification costs

The quadratic discriminant function, $D_j(r)$, is define as:

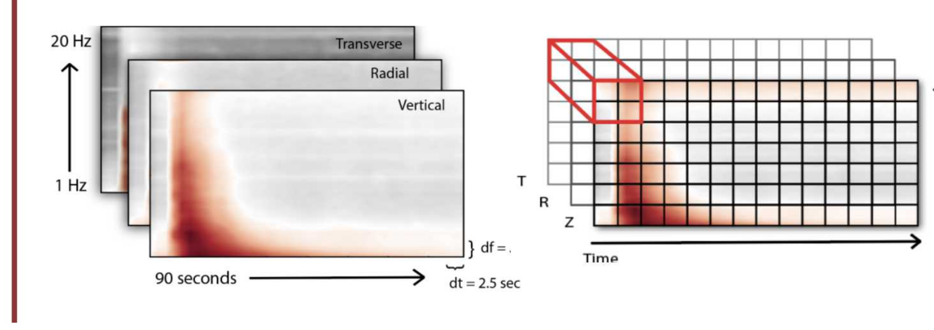
$$D_j(\boldsymbol{r}) = -\tfrac{1}{2}\left(\boldsymbol{r}-\mu_j\right)^T \boldsymbol{S}_j^{-1}\left(\boldsymbol{r}-\mu_j\right) - \tfrac{1}{2}\ln\left|\boldsymbol{S}_j\right| + \ln\left(\tfrac{1}{N}\right),$$

where μ_j and $\boldsymbol{S}_j$ are the ratio vector mean and ratio covariance matrix of the learning set for the *j*-th population, respectively.
An event of interest with the ratio vector $\boldsymbol{x}$ is allocated to the *l*-th population if $D_l(\mathrm{x})$ is the largest of the score values $D_1(\mathrm{x}), D_2(\mathrm{x}), \ldots, D_N(\mathrm{x})$.

For $N = 2$ (two populations *p* and *q*), the event of interest is classified as *p*-type if the score difference $d_{pq} = D_p(\mathrm{x}) - D_q(\mathrm{x}) > 0$, and as *q* type if $d_{pq} < 0$.

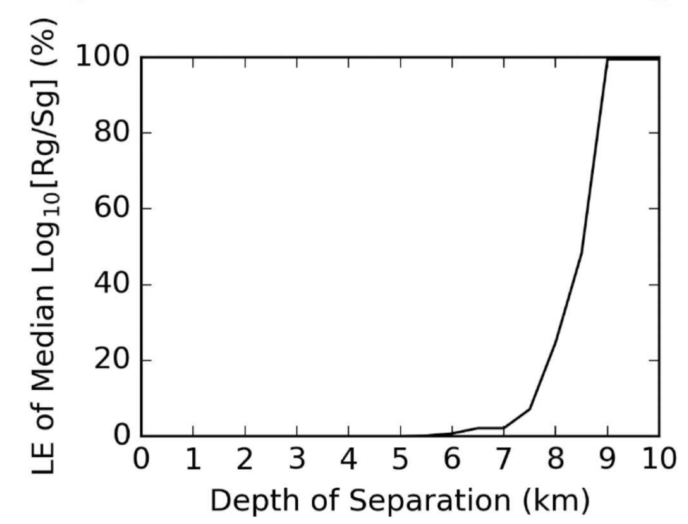
Spectrogram-Based Machine Learning Method

- Input are 90-sec long spectrograms (starting 10 sec before first *P* arrivals) from 3-component (when available) or single local to near-regional stations
- Models were built using 5 years (2012–2016) of labeled UUSS events, except for the MIEs for which we used the current dataset
- We used the (VGG11) convolutional neural network (CNN) architecture, that involves between 64 and 512 filters for each layer and includes batch normalization and max pooling with Relu non-linearity
- Our hyperparameter optimization relied on extensive testing from previous studies that relied on the UUSS dataset
- We used 10 fold cross validation and early stopping
- Our models trained between 20 and 80 epochs



RESULTS

Depth Discrimination Based on *Rg*-to-*Sg* Ratios



Likelihood of equality (LE, inferred from Mood's median tests) of the median *Rg*-to-*Sg* ratios as function of the depth of separation between shallow and deeper TEs.

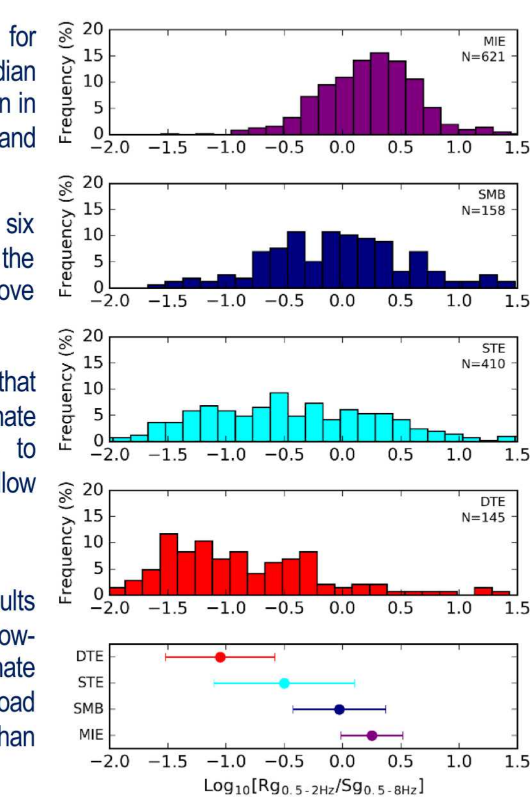
For the depths of separation of 6 km or less, the LE values are below 2%, implying that for any of those depths of separation shallow and deeper TEs are statistically different

Histograms of the estimated *Rg*-to-*Sg* ratios for each population. The median and one-median absolute deviation values of the ratios are shown in the bottom plot. The division of TEs into STEs and DTEs occur along the depth of 5 km.

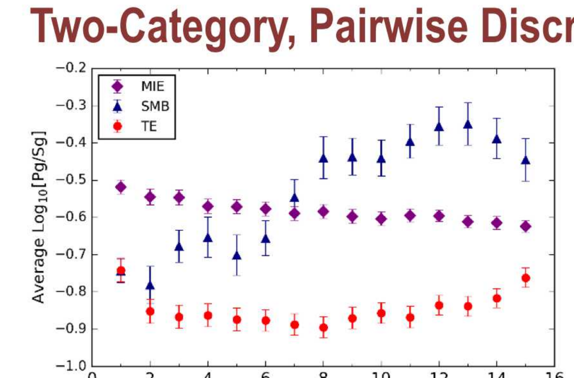
Mood's median tests suggest that for the six possible population pairs the median ratios of the members are different at confidence levels above 99%.

This provides further argument for the notion that using *Rg*-to-*Sg* ratios one can not only discriminate shallow from deeper events, but is able to discriminate among different populations of shallow events.

This finding is also consistent with the results reported by Kolaj (2018), according to which low-frequency *Sg*-to-*Rg* ratios successfully discriminate low-magnitude shallow earthquakes and road construction blasts recorded at distances less than 50 km in New Brunswick, Canada.

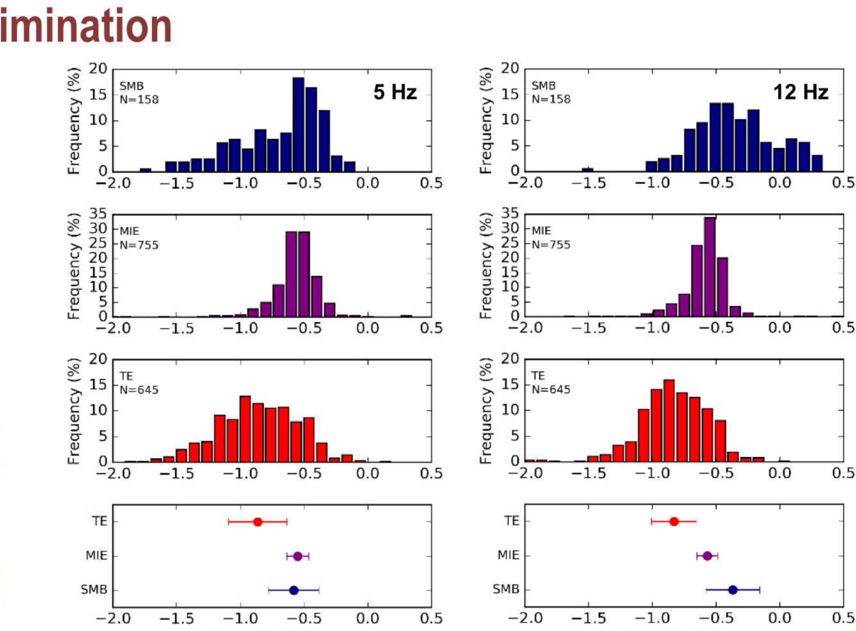


Two-Category, Pairwise Discrimination

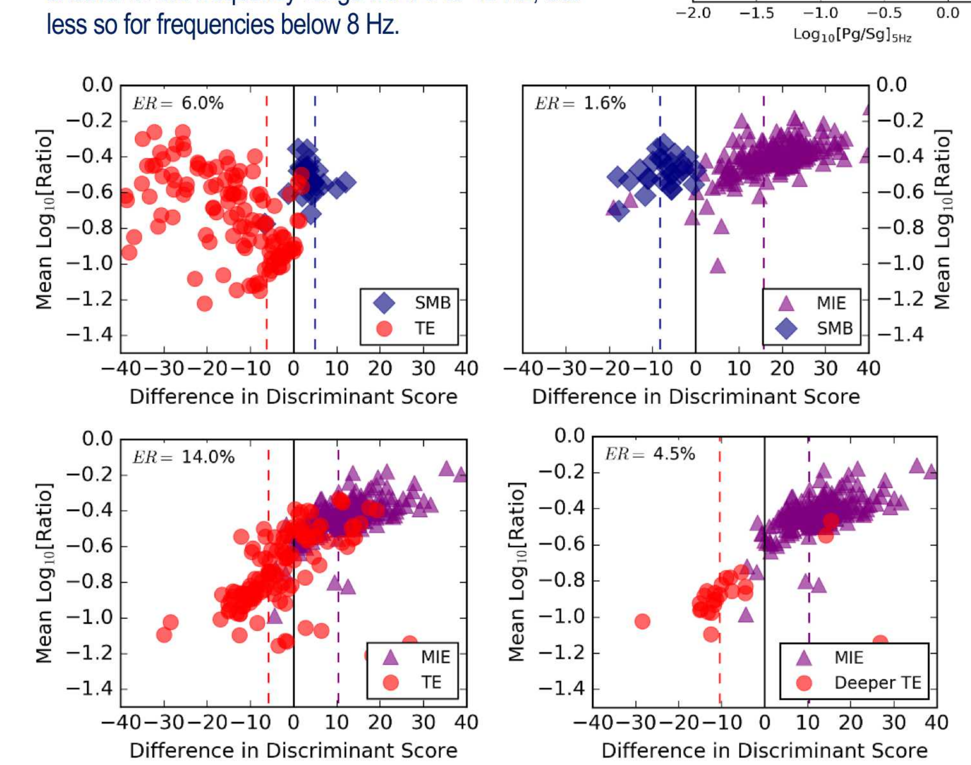


Average *Pg*-to-*Sg* ratios and associated two-standard deviations as estimated for each discrete frequency from 1 to 15 Hz for each population.

The three populations are well separated from one another in the frequency range from 8 to 15 Hz, but less so for frequencies below 8 Hz.



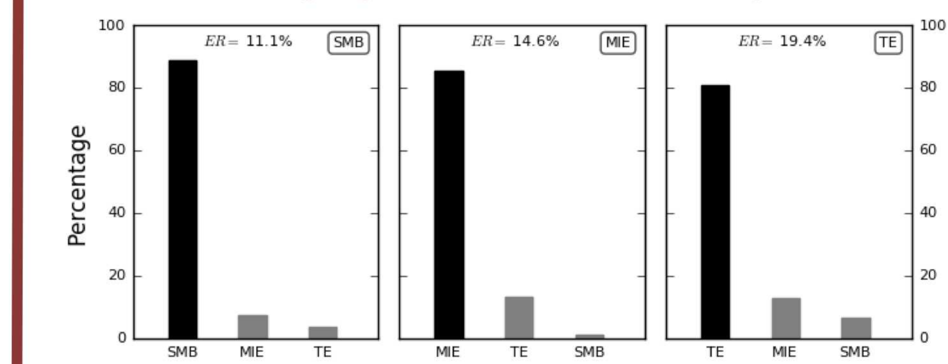
Mood's median tests suggest that at 12 Hz the populations are all statistically different. At 5 Hz, the SMBs are different from the TEs, and the MIEs from the TEs; but, there is a ~7% likelihood that the SMBs and MIEs are the same.



The QDFs were tailored for each population pair. The optimum set of frequencies for the *Pg*-to-*Sg* ratios were selected for each pair using a trial-and-error process, and corresponds to the set associated with the lowest misclassification rate. The optimum frequency sets were 1, 2, 10, 15 Hz; 1, 2, 6, 8 Hz; and 6, 8, 14, 15 Hz for the pairs MIE/SMB, SMB/TE, and MIE/TE, respectively.

Difference in discriminant scores for each of the population pairs. The discriminant scores were calculated by jointly exploiting Pg-to-Sg and Rg-to-Sg ratios in the multivariate QDFs.
With an error rate (ER) of only 1.6%, the QDF performs best for the population pair MIE/SMB, followed by the pair SMB/TE (ER = 6%). MIEs and TEs share some characteristics, as reflected by the ER value of 14%. This is particularly the case between the MIEs and shallow (depth ≤ 5 km), as implied by the fact that the ER value drops to 4.5% if only deeper TEs are considered.

Three-Category Classification Using the AR Method



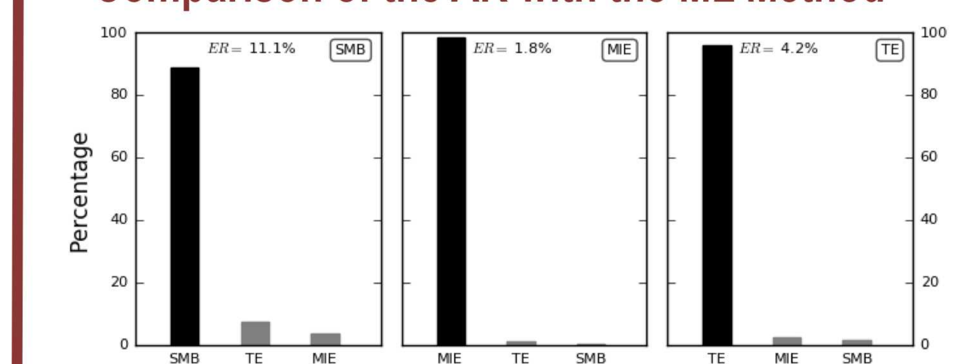
		Actual Class		
		SMB	MIE	TE
Predicted Class	SMB	24	3	8
	MIE	2	187	16
	TE	1	29	100

Confusion matrix: The diagonal elements indicate the number of events classified correctly (true positives).

All three populations are considered simultaneously in the QDF approach.

- The optimum frequency set for *Pg*-to-*Sg* ratios is 1, 2, 7, and 8 Hz.
- On average about 85% of the events are classified correctly.
- The population of SMBs is associated with the lowest error rate (ER = 11.1%) followed by the MIEs (14.6%). The TE group exhibits the highest error rate of any population (19.4%).
- The majority of the misclassified TEs are assigned to the MIE group. Similarly, most of the misclassified MIEs are assigned to the TE population. This is consistent with the outcome of the two-category classification discussed above, which suggests that the two populations have some properties in common, which make it difficult to discriminate between them.

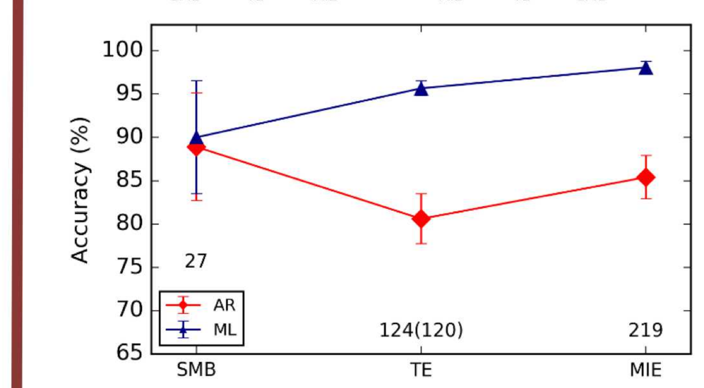
Comparison of the AR with the ML Method



3-category classification using the ML approach.

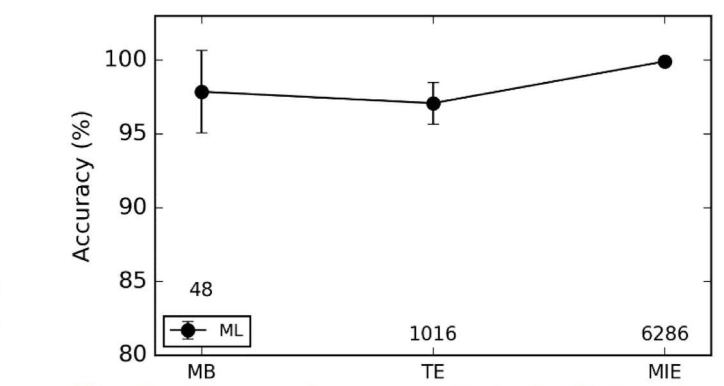
		Actual Class		
		SMB	MIE	TE*
Predicted Class	SMB	24	1	2
	MIE	1	215	3
	TE	2	3	115

Confusion matrix for the ML method. The reported values represent averages from 10 prediction runs.
* Note that in this case only 120 TEs were investigated, compared with 124 TEs for the AR approach.



- For each classification method and each population, the standard deviations estimated from the 10 prediction runs are indicated.
- The high variability in classification success rates (accuracy) for the SMB group is mainly the result of the limited number of events in that group.

- For the SMB population the performance of the AR method is comparable with that of the ML approach within the margin of errors.
- For the TE and MIE groups, however, the latter method outperforms the former by an average of about 14% in terms of success rate.



Using the ML approach, we were able to classify the entire UUEB dataset, including the abundant, extremely low-magnitude events. In this case, the SMBs, WMBs, LMBs, and CQBs were grouped together to form a class called "mining blast" (MB). We achieved success rates of ~97–99%.

CONCLUSIONS

	AR Method	ML Method
Advantages	▪ Require only limited number of learning events for each category ▪ Fast computation ▪ Physical basis well understood (i.e., energy partitioning between specific phases)	▪ Achieve high accuracy (≥ 90%) ▪ Can classify extremely low-magnitude events ▪ Fast to apply the model once it has been built
Limitations	▪ Achieve moderate accuracy (~80–90%) ▪ Require high-quality data	▪ Require large dataset (100s–1000s) of labeled training events for each category ▪ Computationally intensive to build the model ▪ Lack of insight into the physical basis

- ❑ Using the same set of events that were well-recorded at local distances by the UUSS network, a traditional method of event classification in which amplitude ratios are exploited in the QDFs is compared with a spectrogram-based machine learning approach.
- ❑ With the AR method, we achieve classification success rates of about 80–90%.
- ❑ The ML approach uses convolutional neural network models to classify the populations, and achieves success rates of 90% and higher.
- ❑ The ML method is more robust to low SNR data, allowing it to classify extremely low-magnitude events.
- ❑ The complex algorithms involved in the ML method are able to expose and exploit characteristics that are specific to each event population to the level that the traditional AR approach cannot, allowing ML to achieve higher accuracies and classify significantly smaller events.
- ❑ The lack of insight into the physical basis for ML classifications has traditionally been a major reason why people have been reluctant to use ML methods. SNL intends to work on this problem.

REFERENCES

- Kolaj, M. (2018). Discrimination between low-magnitude shallow earthquakes and road construction blasts near Big Salmon River, New Brunswick, Canada, *Seismol. Res. Lett.*, doi: 10.1785/0220180077.
- Linville, L., R. Brogan, C. Young, and K. A. Aur (2018). Global to local high-resolution event catalogs for algorithm testing and source studies. Submitted to *Bull. Seismol. Soc. Am.*
- Linville, L., T. Draelos, K. L. Pankow, and C. J. Young (2018). Deep learning models augment analyst decisions for event discrimination. Submitted to *Geophys. Res. Lett.*
- Tibi, R., K. D. Koper, K. L. Pankow, and C. J. Young (2018a). Depth discrimination using *Rg*-to-*Sg* spectral amplitude ratios for seismic events in Utah recorded at local distances, *Bull. Seismol. Soc. Am.* **108,** 1355–1368, doi: 10.1785/0120170257.
- Tibi, R., K. D. Koper, K. L. Pankow, and C. J. Young (2018b). Discrimination of anthropogenic events and tectonic earthquakes in Utah using a quadratic discriminant function approach with local distance amplitude ratios, *Bull. Seismol. Soc. Am.* **108,** 2788–2800, doi: 10.1785/0120180024.