SnT 2023

CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

HOFBURG PALACE - Vienna and Online

19 TO 23 JUNE

Testing and Design of Discriminants for Local Seismic Events Recorded During the Redmond Salt Mine Monitoring Experiment in Utah

Rigobert Tibi[1], Nathan Downey[1], and Ronald Brogan[2]

1) Sandia National Laboratories; 2) ENSCO, Inc.

O3.5-197

Presentation Date: 20 June 2023

Acknowledgments: Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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-NA-0003525. This Low Yield Nuclear Monitoring (LYNM) research was funded by the National Nuclear Security Administration, Defense Nuclear Nonproliferation Research and Development (NNSA DNN R&D). The authors acknowledge important interdisciplinary collaboration with scientists and engineers from LANL, LLNL, NNSS, PNNL, and SNL.

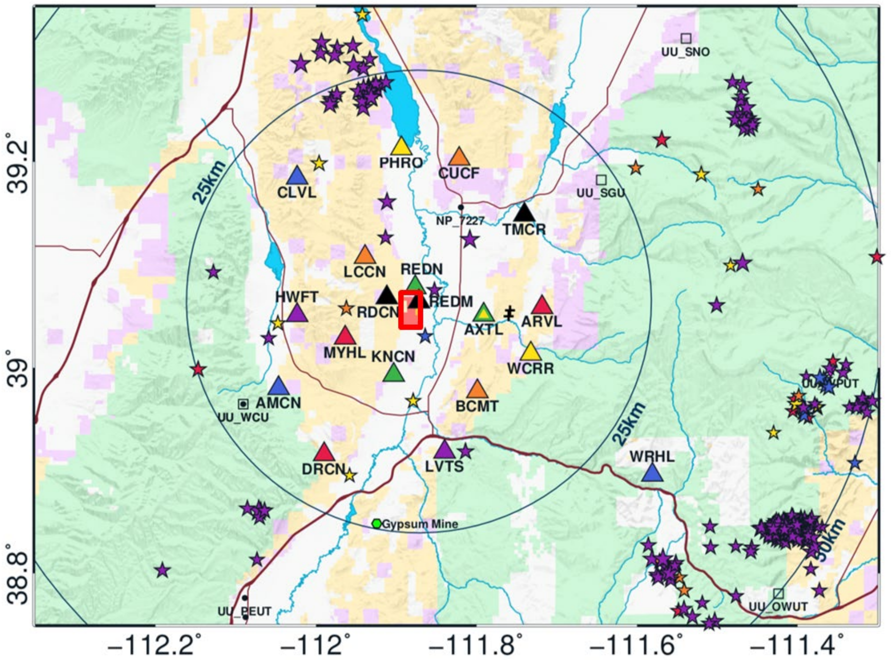
SAND2023-03407C

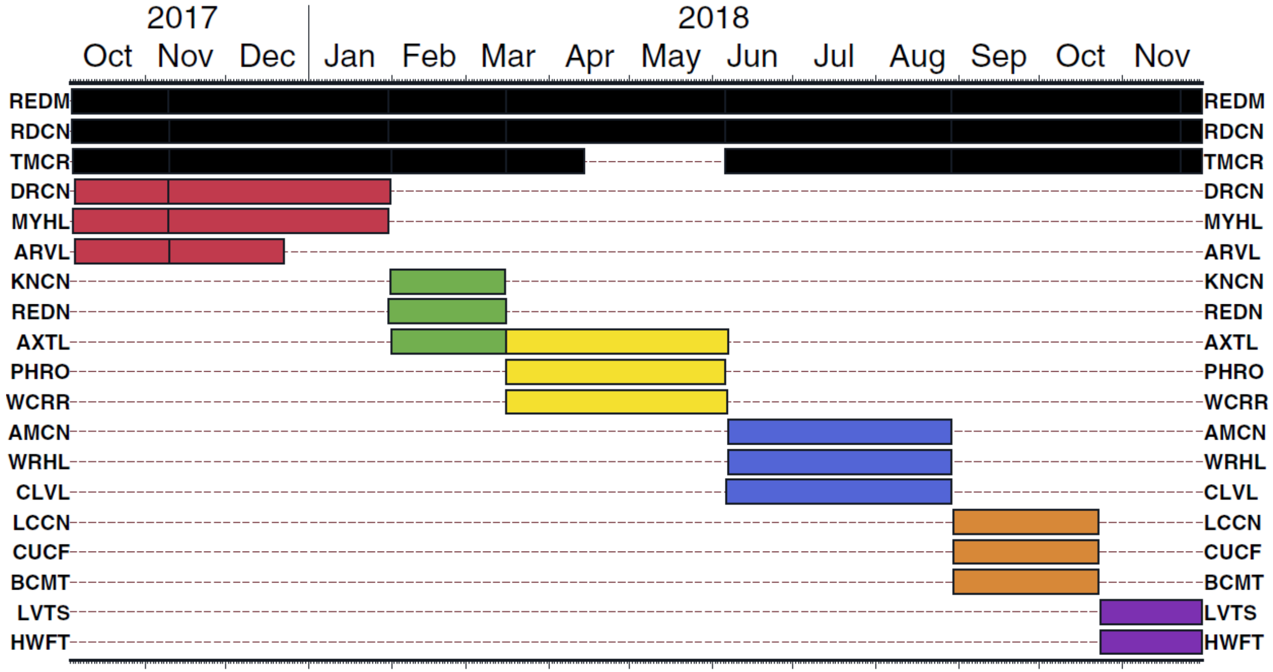
Sandia National Laboratories

NNSA National Nuclear Security Administration

INTRODUCTION

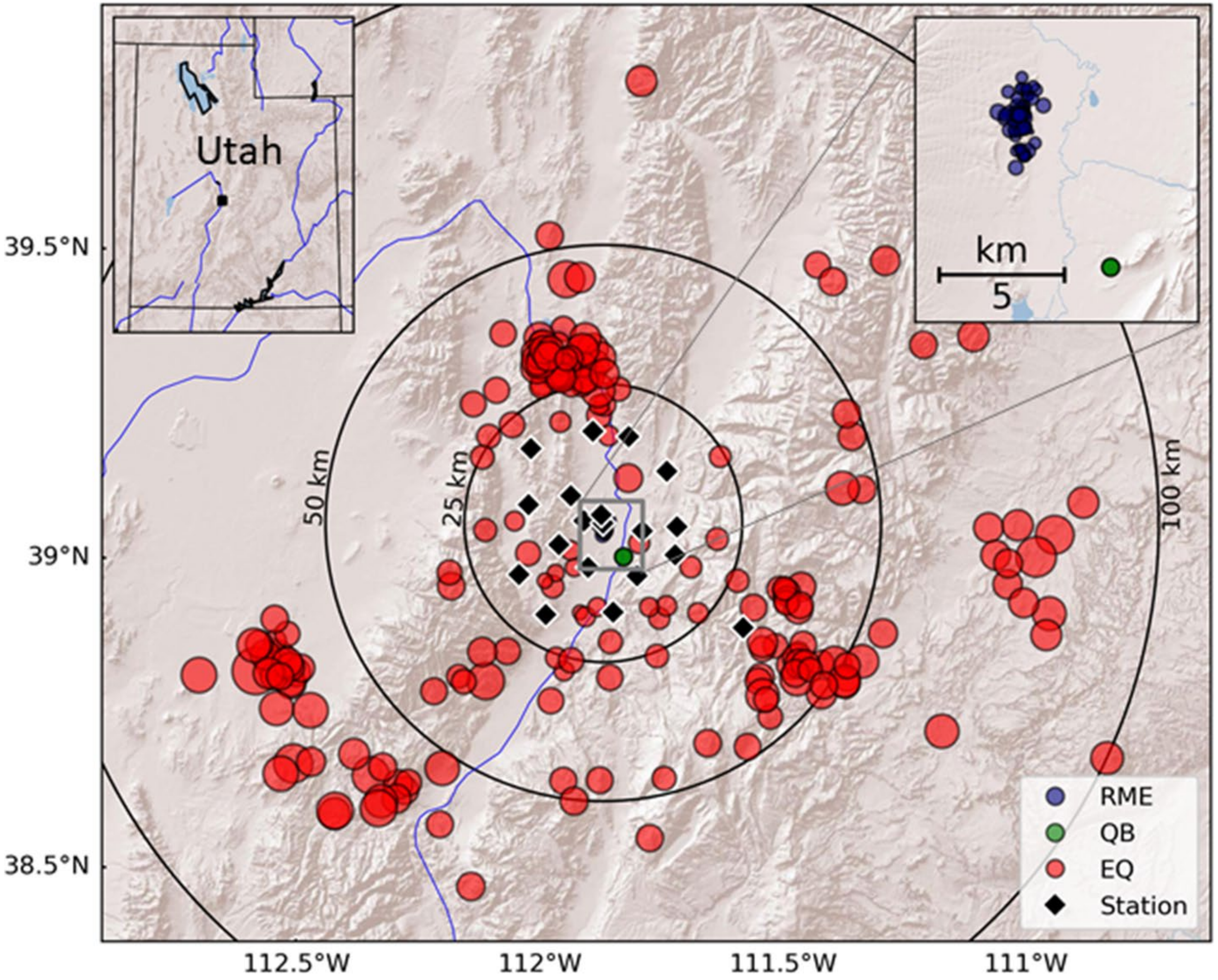
- The Redmond Salt Mine Monitoring Experiment in Utah was designed to record seismoacoustic data at distances less than 50 km for algorithm testing and development.
- During the experiment from October 2017 to July 2019, six broadband seismic stations were operating at a time, with three of them having fixed locations for the duration, while the three other stations were moved to different locations every one-and-half to two-and-half months.
- Redmond Salt Mine operations consist of night-time underground blasting several times per week.
- The mine is located within a belt of active seismicity, allowing for easy comparison of the mining blasts with tectonic earthquakes.



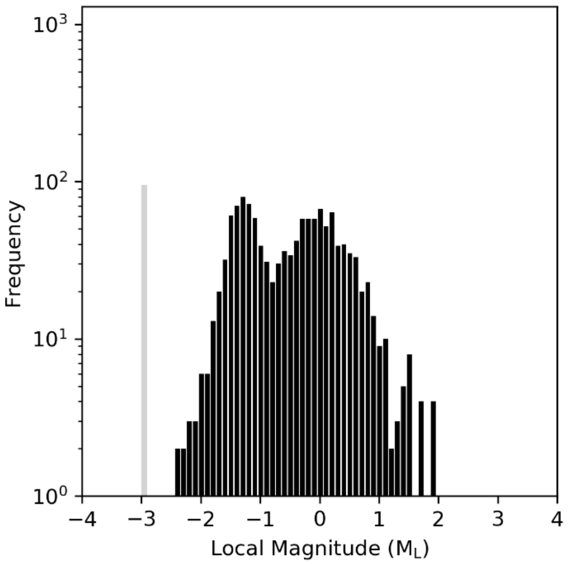


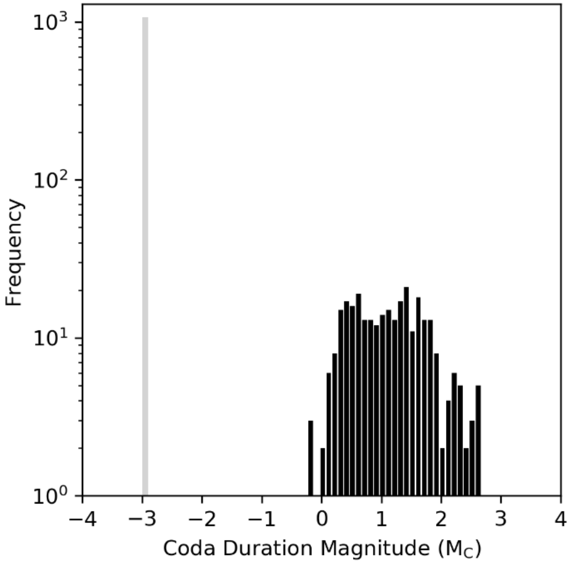
DATA

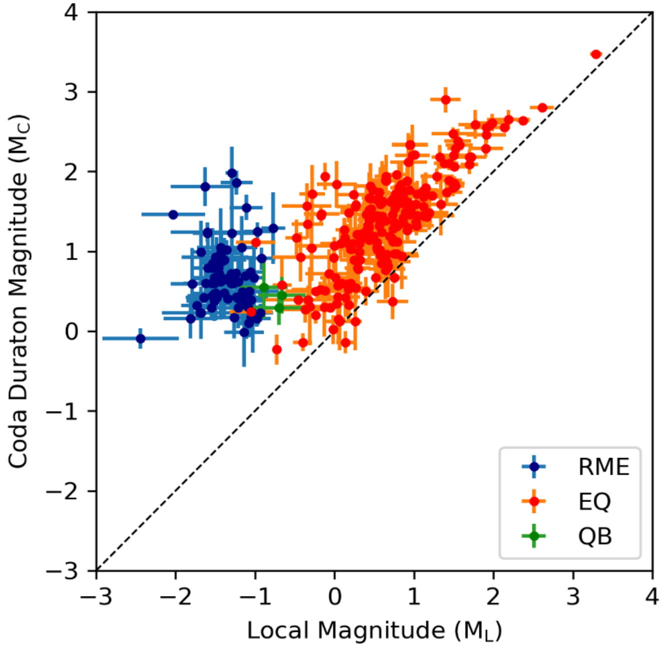
- We built 1373 events. For 284 of the events, both M_L and the coda duration magnitude (M_C) are well constrained.
- Based on the event locations and the signal onset characters, this subset was divided into three populations:
  - 75 blasts from the Redmond Salt Mine (**RMEs**),
  - 206 tectonic earthquakes (**EQs**), and
  - 3 blasts (**QBs**) from a mine/quarry located about 8 km from the Redmond Mine.
- **We used the subset of events to test and design discriminants that are effective at local distances.**

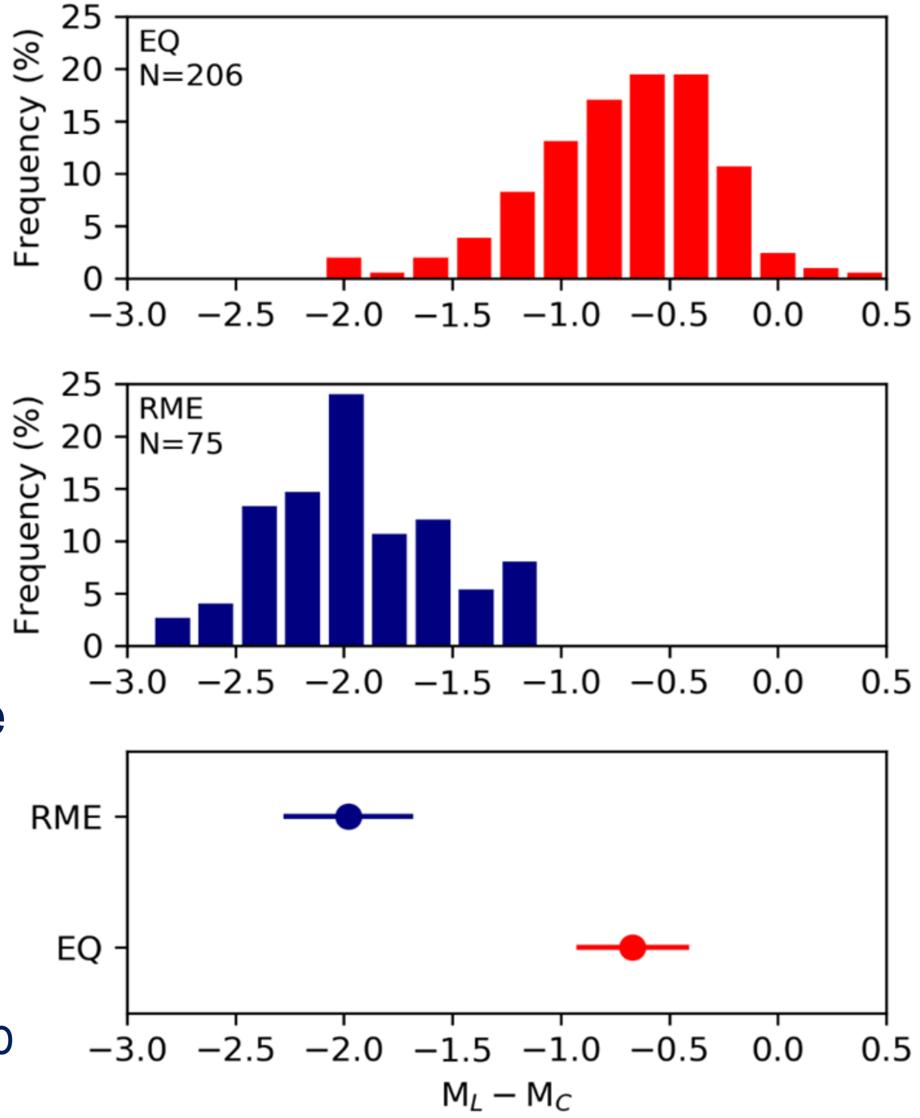


MAGNITUDE-BASED DISCRIMINATION



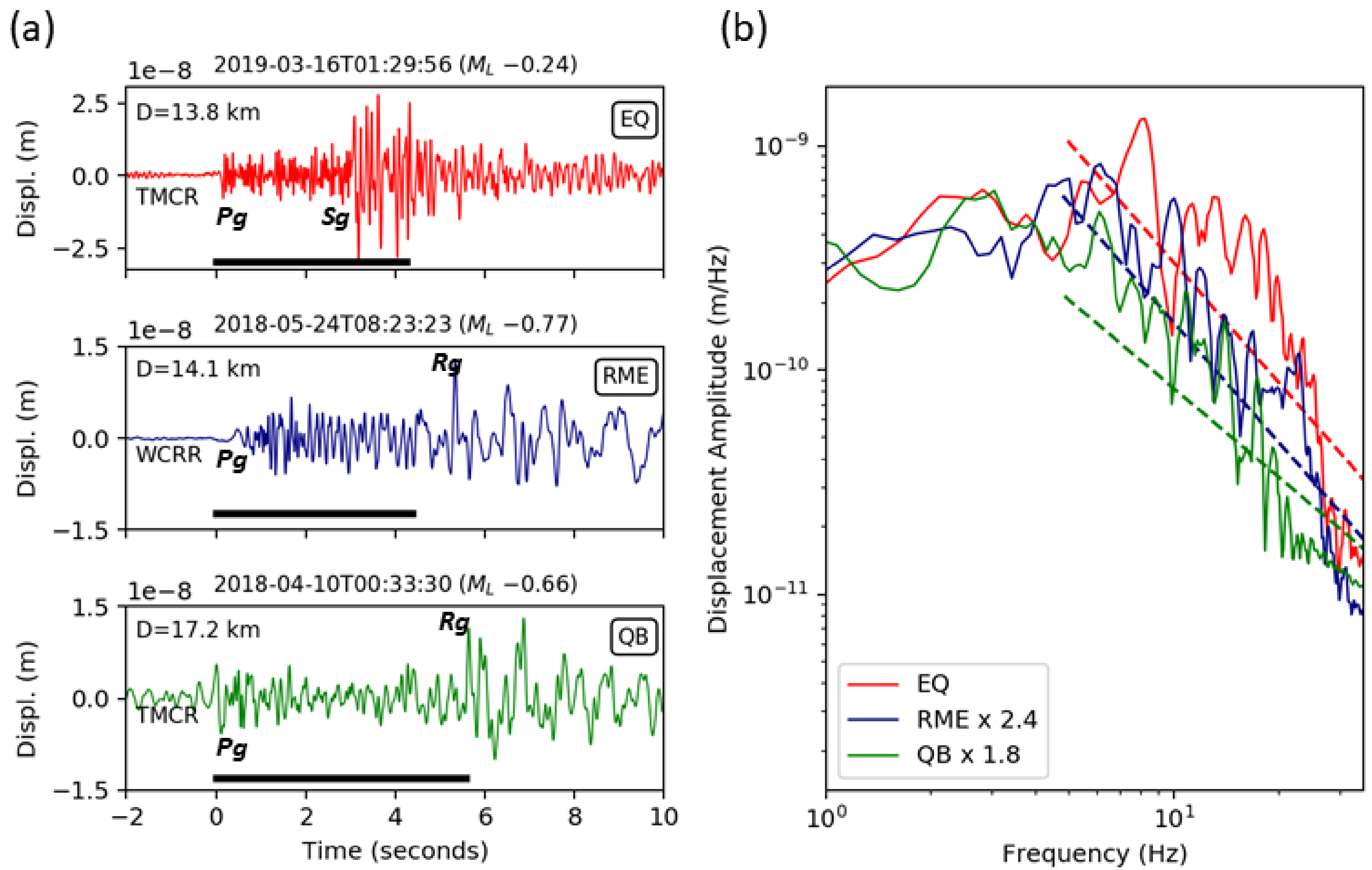






- Data points for the mining events are significantly off the 1:1 line, due to enhanced coda for the shallow mining events (Koper et al., 2016).
- Chi-square tests suggest that the two populations are statistically distinct.

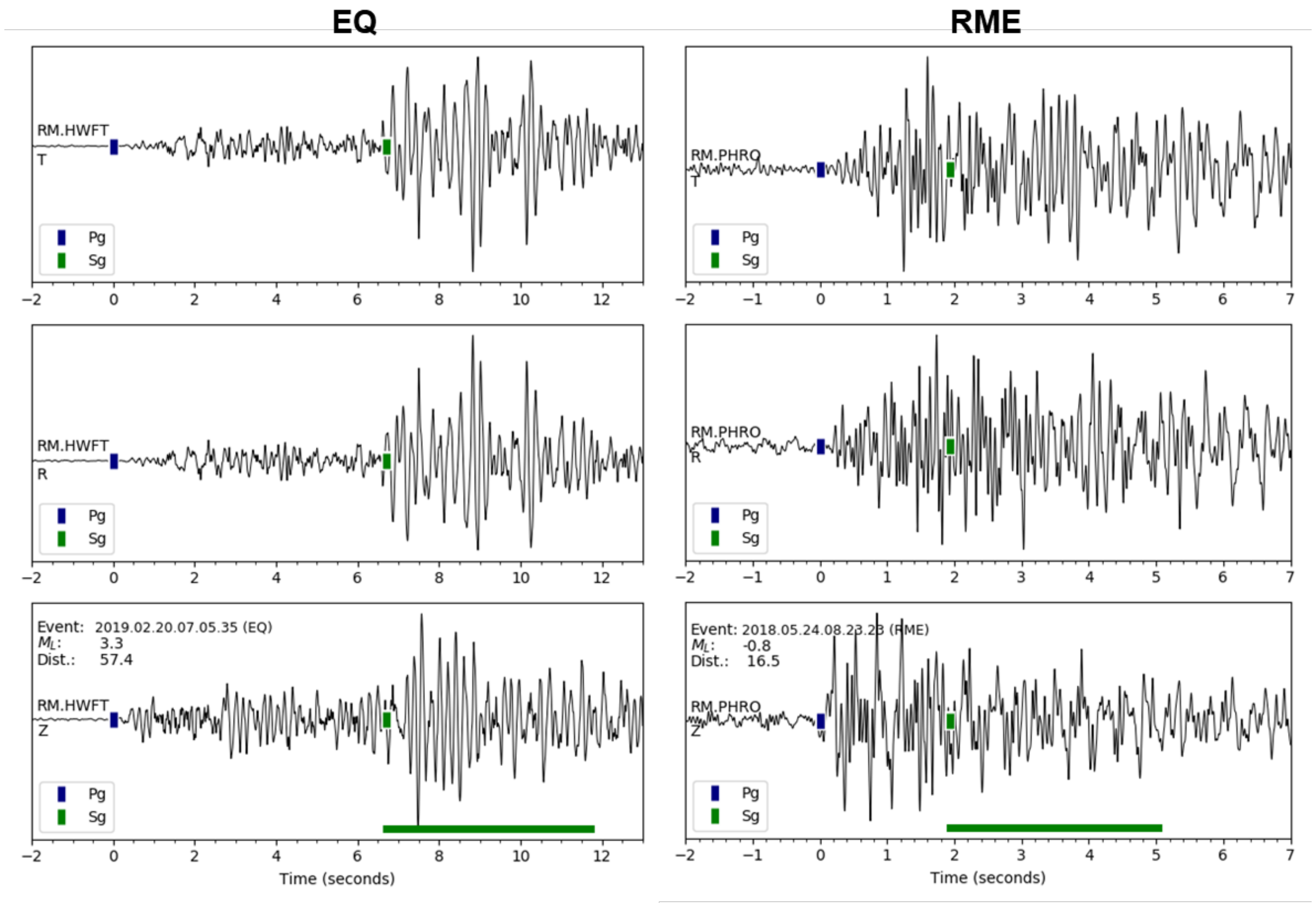
EXAMPLE SEISMOGRAMS AND SPECTRA



- Differing excitations of *Pg*, *Sg*, and *Rg* between the events
- Steeper fall-off slope for EQ spectrum, indicating a relative deficiency in high-frequency energy
- Dissimilarities between event classes constitute the base for discriminants investigated

DISCRIMINATION BASED ON AMPLITUDE RATIOS

Measurement of Sg Phase Amplitudes (for Low Frequency Sg to High Frequency Sg Discriminant)



- The root mean square (RMS) of *Sg* amplitudes are measured in the four frequency bands 5 –10 Hz, 10–15 Hz, 15–20 Hz, and 20–30 Hz.
- As a result of the Parseval's theorem, the RMS in the time domain is equivalent to the RMS in the frequency domain.

DISCRIMINATION BASED ON AMPLITUDE RATIOS

Separation of Source Excitation, Propagation, and Site Terms

- The recorded amplitude A_{ij} for an event *i* recorded at a station *j* is expressed as in Equation 1.

$$\log A_{ij}(f) = \log EXC_i(f) + \log SITE_j(f) + \log G(r_{ij}, f) \quad (1)$$

  $EXC_i(f)$: Source excitation term for source *i*; $SITE_j(f)$: Site term for station *j*;

  $G(r_{ij}, f)$: Distance-correction term (combined effect of geometrical spreading and attenuation); r_{ij}: Distance from source *i* to station *j*.

- We parameterized the distance-correction term using a piecewise linear function (Yazd, 1993; Kintner et al., 2020). We defined series of nodes, r_k, in 5-km increment over the source-distance range of 5–100 km ($k = 1, 2, \ldots, 20$).
- Distance-correction term, $G(r_{ij}, f)$, approximated using linear interpolation:

$$G(r_{ij}, f) = G(r_k, f) + \frac{(r_{ij}-r_k)}{(r_{k+1}-r_k)}(G(r_{k+1}, f) - G(r_k, f)) \quad (2)$$

- Linearized equation for an observation at r_{ij} between r_k and r_{k+1}:

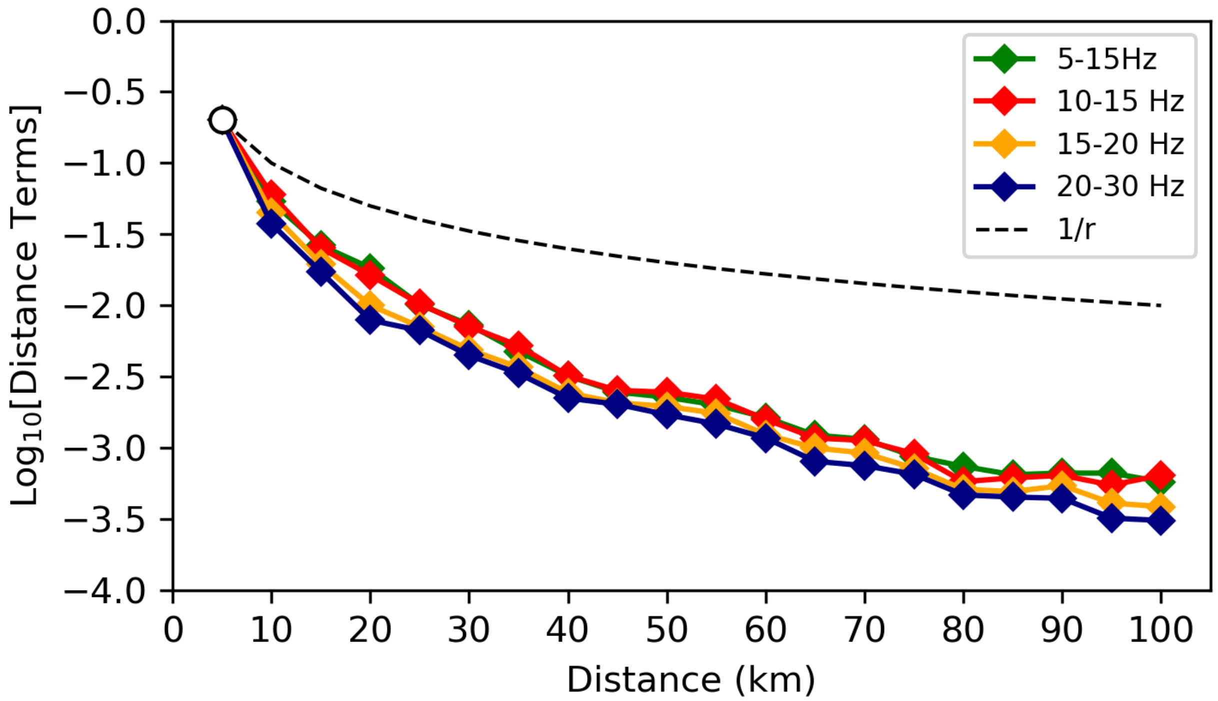
$$\log A_{ij}(f) = \log EXC_i(f) + \log SITE_j(f) + q\log G(r_k, f) + p\log G(r_{k+1}, f), \quad (3)$$

where $p = \frac{(r_{ij}-r_k)}{(r_{k+1}-r_k)}$ and $q = 1 - p$

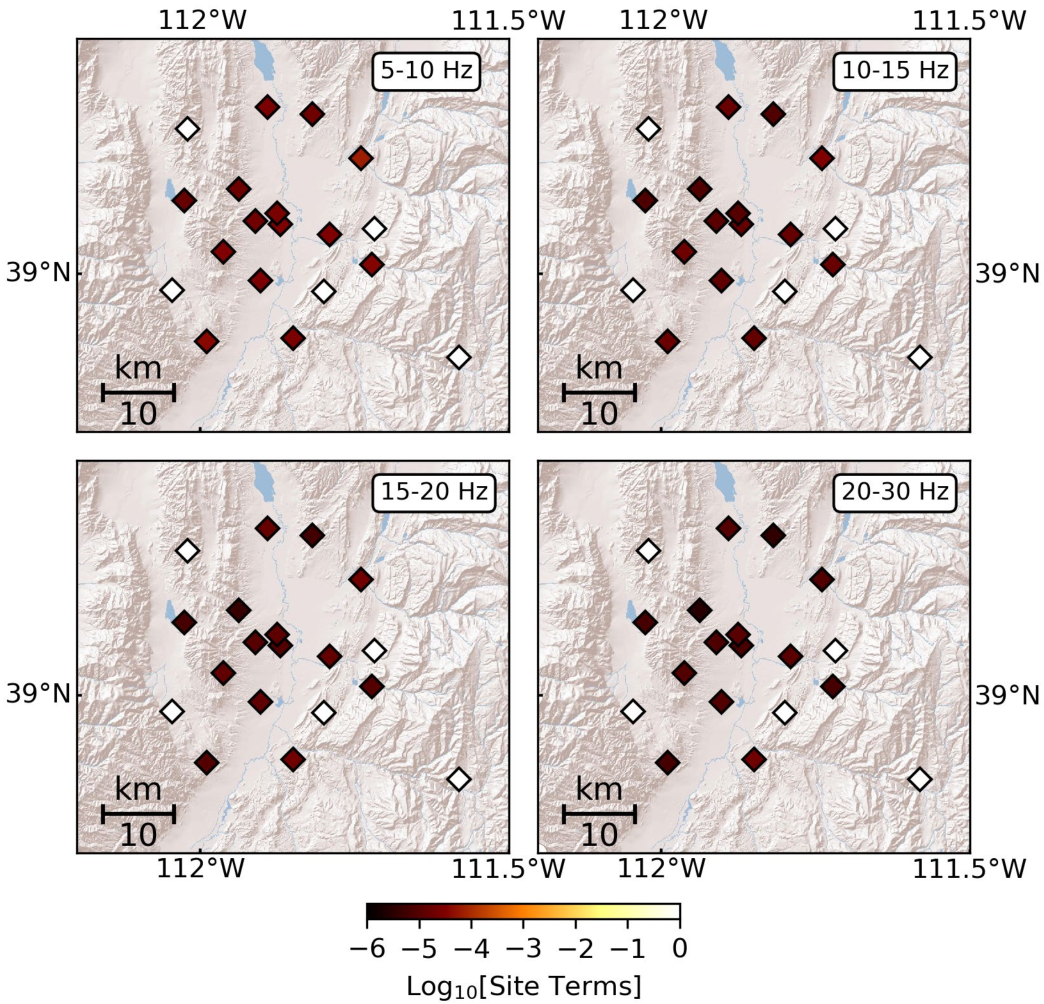
  - Matrix form: $\boldsymbol{a} = \boldsymbol{Km}$ (4)
    - ***a***: Array of logarithm of amplitude measurements
    - ***m***: Array of logarithm of model parameters (source excitation terms, station site terms, and distance-correction terms at nodes 5, 10, 15, …, 100 km)
      - 276 sources, 19 stations, and 20 distance nodes.
- <u>Constraint:</u> For r_k = 5 km, $G(r_k, f) = 1/r_k$ (i.e., only geometrical spreading, no attenuation)
- $\boldsymbol{m}$ is estimated using the SVD technique.

DISCRIMINATION BASED ON AMPLITUDE RATIOS

Propagation and Site Terms



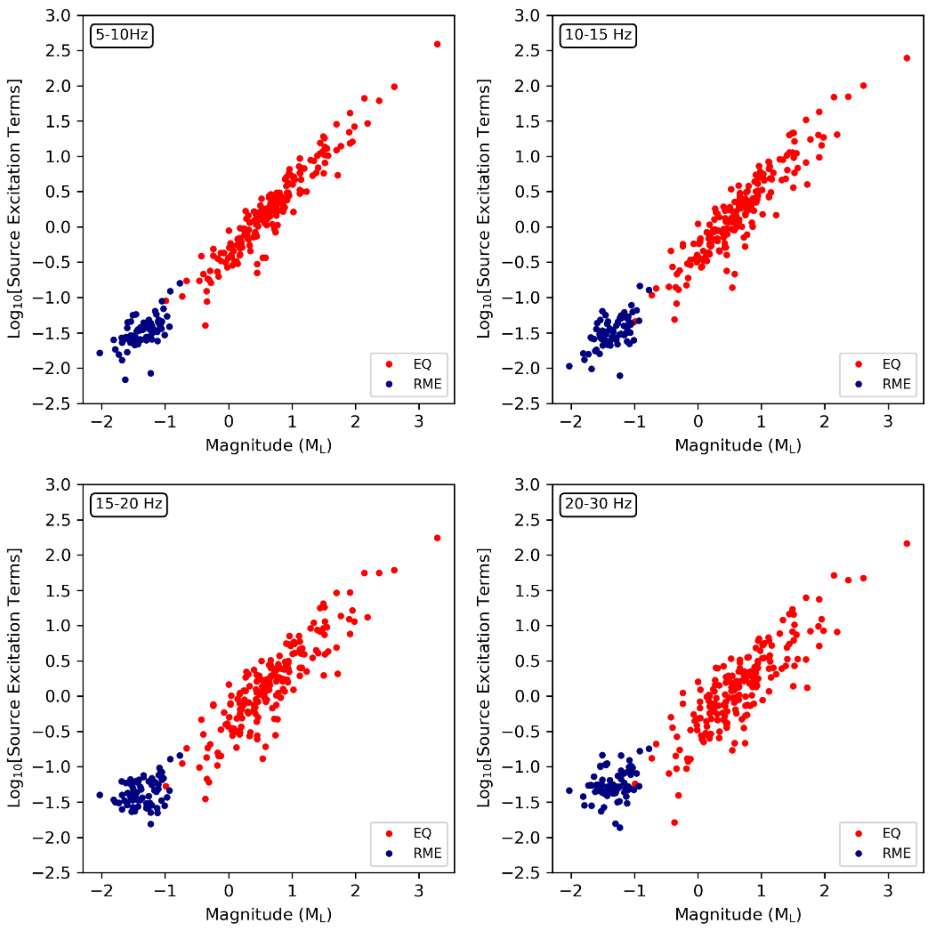
As would be expected for a frequency-dependent attenuation model, high-frequency *Sg* signals are slightly more attenuated than low frequencies.



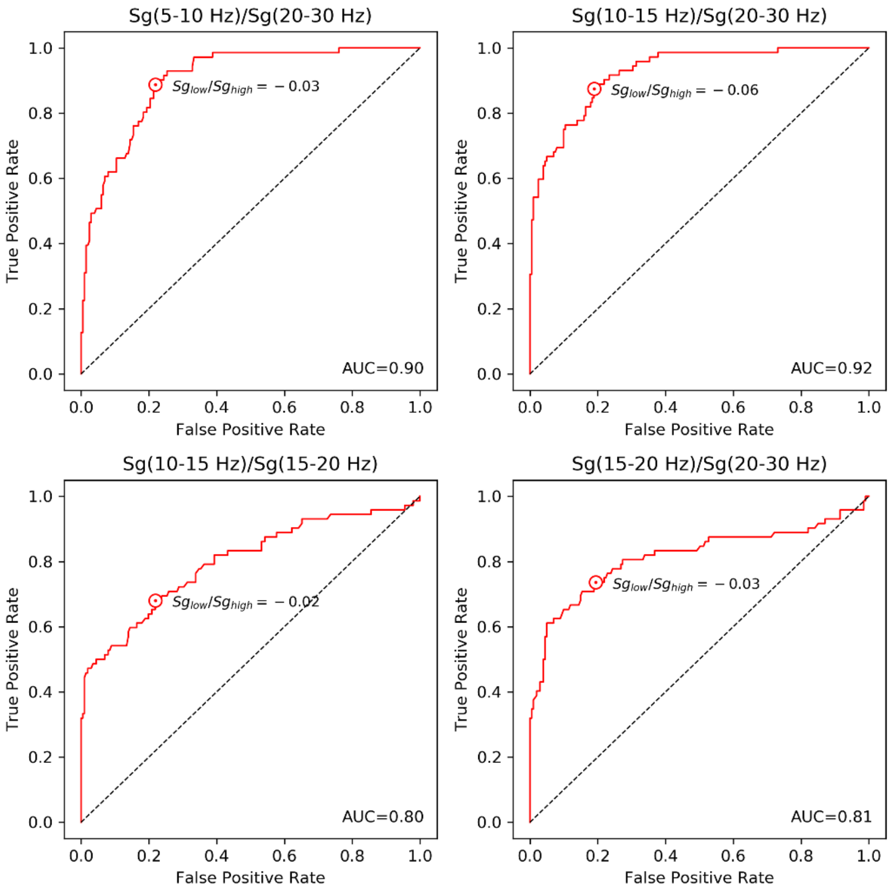
The observed damping at most stations is likely due to locations in basins.

DISCRIMINATION BASED ON AMPLITUDE RATIOS

Source Terms and Low Frequency Sg to High Frequency Sg Discriminant

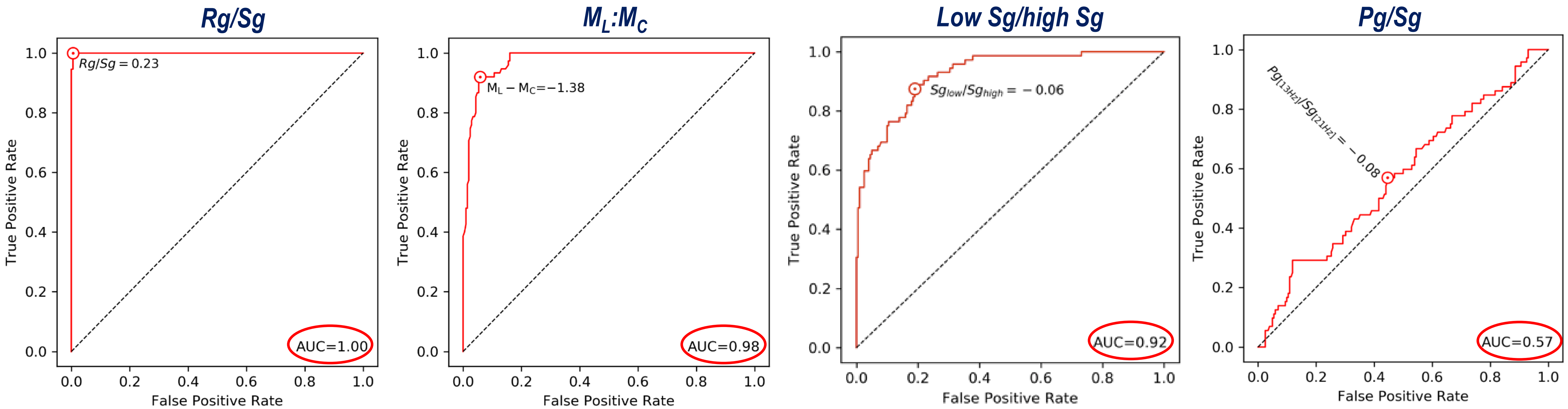


The degree of scattering of the data points increases with increasing frequency, likely due to decreasing SNR.



ROC curve for RMEs vs. EQs. The AUCs range from 0.80 to 0.92, with the highest value associated with $Sg_{[10-15Hz]}/Sg_{[20-30Hz]}$.

DISCRIMINANT COMPARISON



$$LM_{min} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\Delta/2} e^{-x^2/2}\, dx$$

	Discriminant			
	$Rg_{[0.5-2\,Hz]}/Sg_{[0.5-8\,Hz]}$	$M_L - M_C$	$Sg_{[10-15Hz]}/Sg_{[20-30Hz]}$	Pg/Sg
Mahalanobis Distance (Δ^2)	11.8	4.0	1.95	0.03
Minimum Likelihood of Misclassification (LM_{min} in %)	4.3	16.0	24.3	46.4

MULTIVARIATE DISCRIMINANT ANALYSIS

GOAL: Improve discriminant power by combining 2 or more (uncorrelated) classifiers

- A multivariate quadratic discriminant function (QDF), $\boldsymbol{D}(\boldsymbol{v})$, is defined as:

$$\boldsymbol{D}(\boldsymbol{v}) = \boldsymbol{v}^{T}\boldsymbol{A}\boldsymbol{v} + \boldsymbol{B}\boldsymbol{v} + k, \qquad (5)$$

where

$$\boldsymbol{A} = -\frac{1}{2}\left(\boldsymbol{S}_{re}^{-1} - \boldsymbol{S}_{eq}^{-1}\right) \qquad (6)$$

$$\boldsymbol{B} = \mu_{ne}^{T}\boldsymbol{S}_{re}^{-1} - \mu_{eq}^{T}\boldsymbol{S}_{eq}^{-1} \qquad (7)$$

$$k = -\frac{1}{2}\left[\ln\left(\frac{|S_{re}|}{|S_{eq}|}\right) + \left(\mu_{re}^{T}\boldsymbol{S}_{re}^{-1}\mu_{re} - \mu_{eq}^{T}\boldsymbol{S}_{eq}^{-1}\mu_{eq}\right)\right] \qquad (8)$$

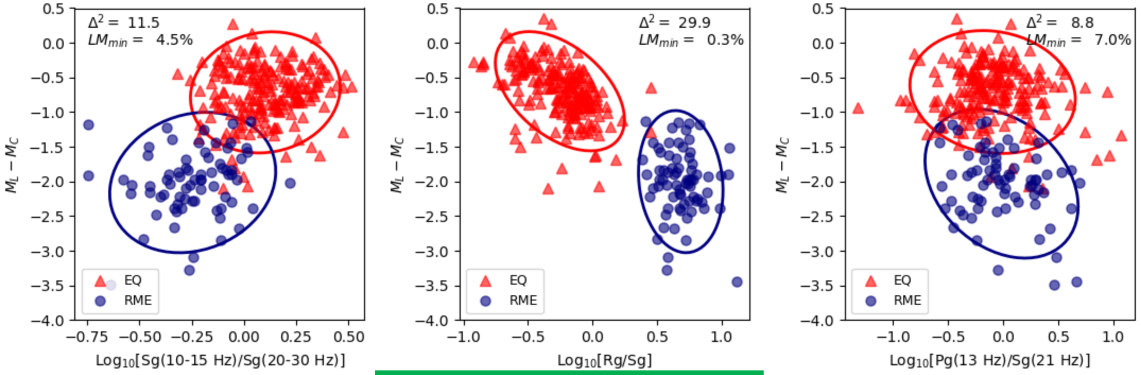
$\boldsymbol{v} = (v_1, \dots, v_n)^T$ is the *n*-dimensional vector of $M_L - M_C$ and/or amplitudes ratios; $\boldsymbol{S}_{re}$ and $\boldsymbol{S}_{eq}$ are the covariance matrices of the parameters; μ_{re} and μ_{eq} the vector means of the parameters for the learning events for the populations *re* and *eq*, respectively.

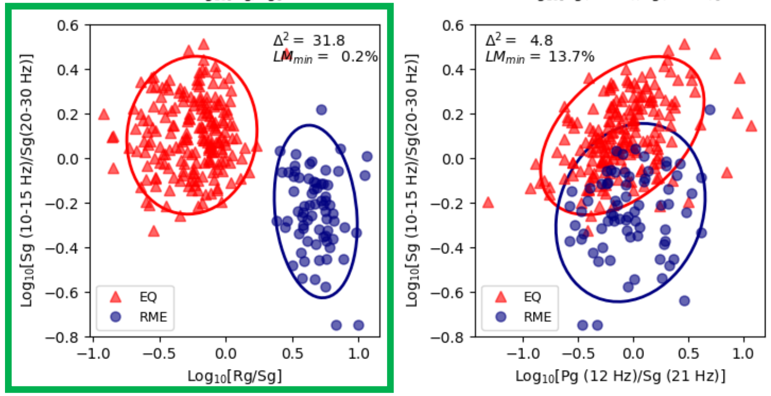
- An event of interest with the parameter vector $\boldsymbol{x} = (x_1, \dots, x_n)^T$ is classified as *re*-type if $\boldsymbol{D}(\boldsymbol{x})$ is positive, and *eq*-type if $\boldsymbol{D}(\boldsymbol{x})$ is negative, with $\boldsymbol{D} = \boldsymbol{0}$ representing the classification line.

MULTIVARIATE DISCRIMINANT ANALYSIS

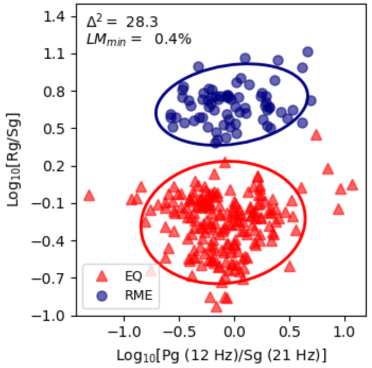
Combination	LM_{min} (%)	LM Improvement (%)
$M_L - M_C$ and low Sg/high Sg	4.5	-0.2
$M_L - M_C$ and Rg/Sg	0.3	4
$M_L - M_C$ and Pg/Sg	7.0	-2.7
low Sg/high Sg and Rg /Sg	0.2	4.1
low Sg/high Sg and Pg/Sg	13.7	-9.4
Rg/Sg and Pg/Sg	0.4	3.9

LM improvements are relative to the best performing single classifier, *Rg/Sg.*



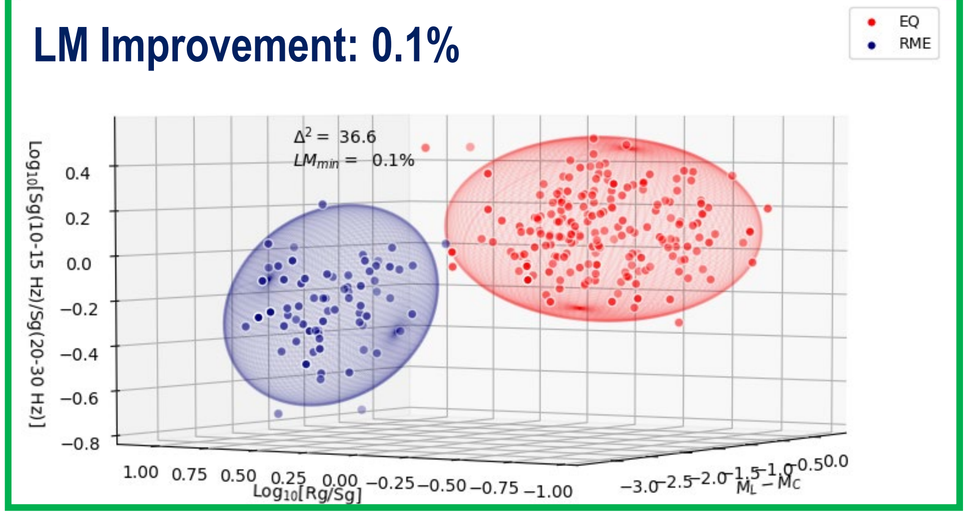


Ellipses outline the 90% confidence regions.

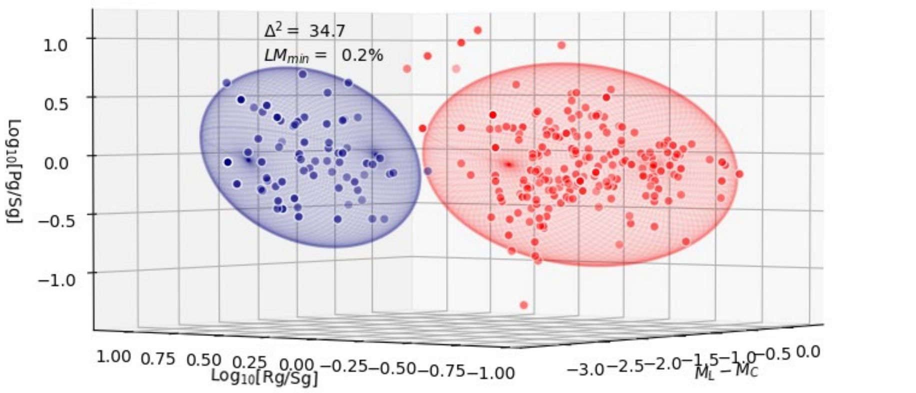


MULTIVARIATE DISCRIMINANT ANALYSIS

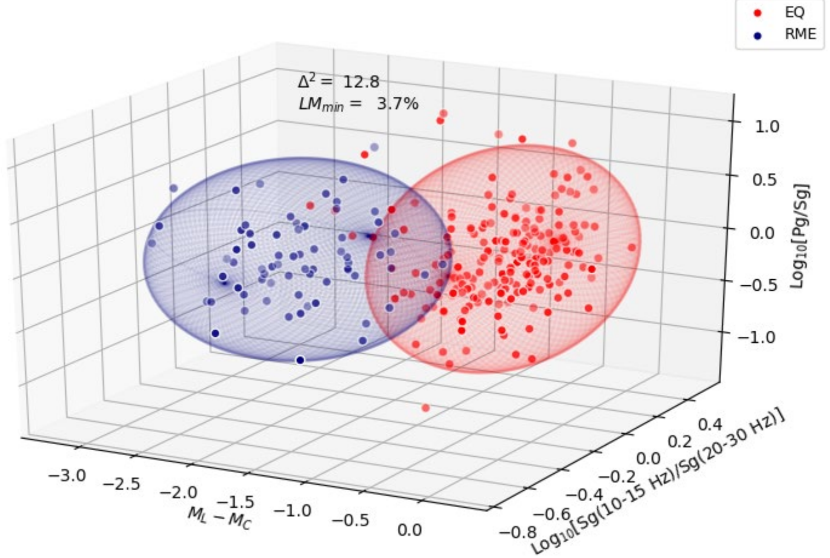
LM Improvement: 0.1%



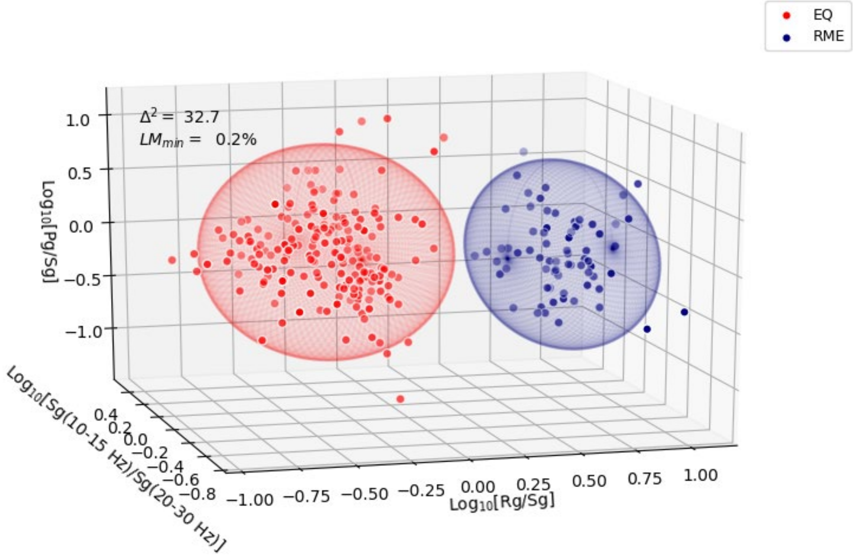
LM Improvement: 0%



LM Improvement: -1.7%



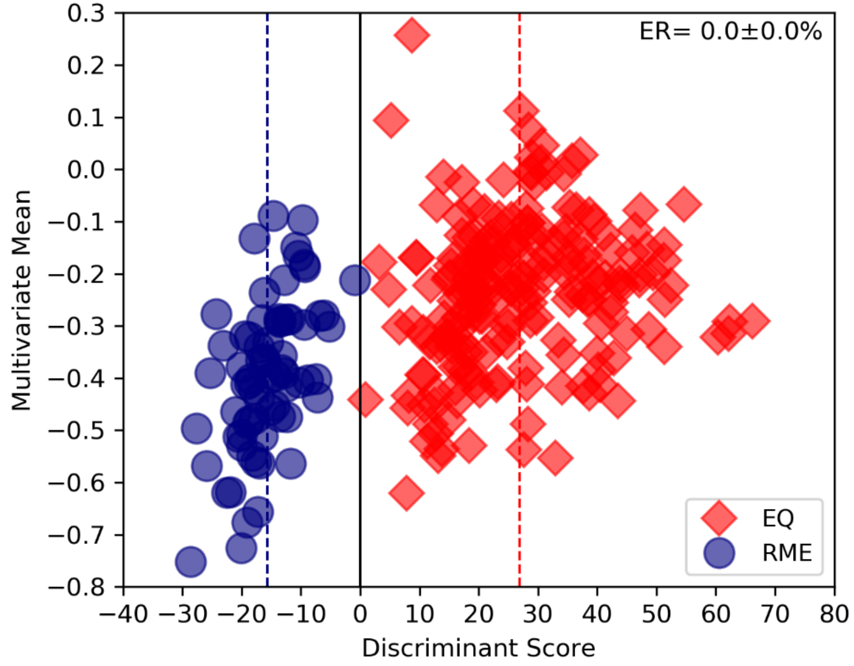
LM Improvement: 0%



CONCLUSIONS

- We tested and designed several classifiers to separate the population of mining blasts from the group of earthquakes recorded at local distances. The classifiers consist of M_L–M_C, low frequency *Sg to* high frequency *Sg*, *Pg/Sg, and Rg/Sg ratios,* and different combinations of 2 or more of these classifiers.

- While the areas under the receiver operating characteristic curve (AUC) of 0.92–1.0 for M_L–M_C, *low Sg/high Sg*, and *Rg/Sg* indicate that these discriminants are very effective, the AUC of only 0.57 suggests that *Pg/Sg* is only slightly better than a random classifier.



- Among the individual classifiers, *Rg/Sg*, which is a depth discriminant, shows the lowest likelihood of misclassification (LM_{min} = 4.3%) for the populations.

- To improve the discriminant power, we combined 2 and more of the discriminants by performing multivariate discriminant analyses. For the bivariate classifier, the combination of *Rg/Sg* with *low Sg/high Sg* provides the largest improvement (4.1% or LM_{min} = 0.2%) over the best single discriminant, while for the best performing tri-variate classifier this improvement is 4.2% (LM_{min} = 0.1%).