

Defense Nuclear Nonproliferation Research & Development

Nuclear Explosion Monitoring Program Review

NEM 2019

Ground-based Nuclear Detonation Detection (GNDD): Waveform Signal Analysis Research & Development

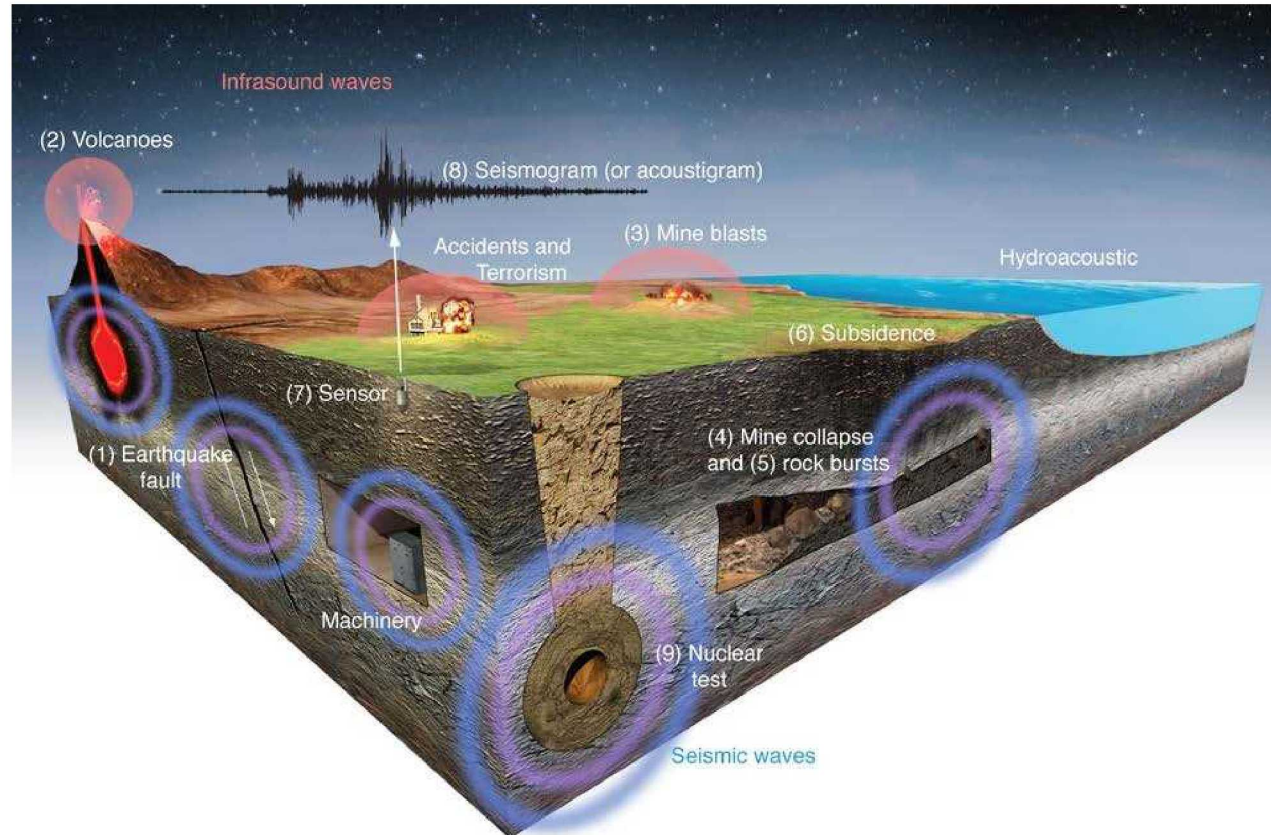
John Merchant

Sandia National Laboratories

March 19, 2019

Signal Analysis focuses on signal enhancement, signal detection, parameter measurement, signal association, event location, magnitude/yield determination, and event identification.

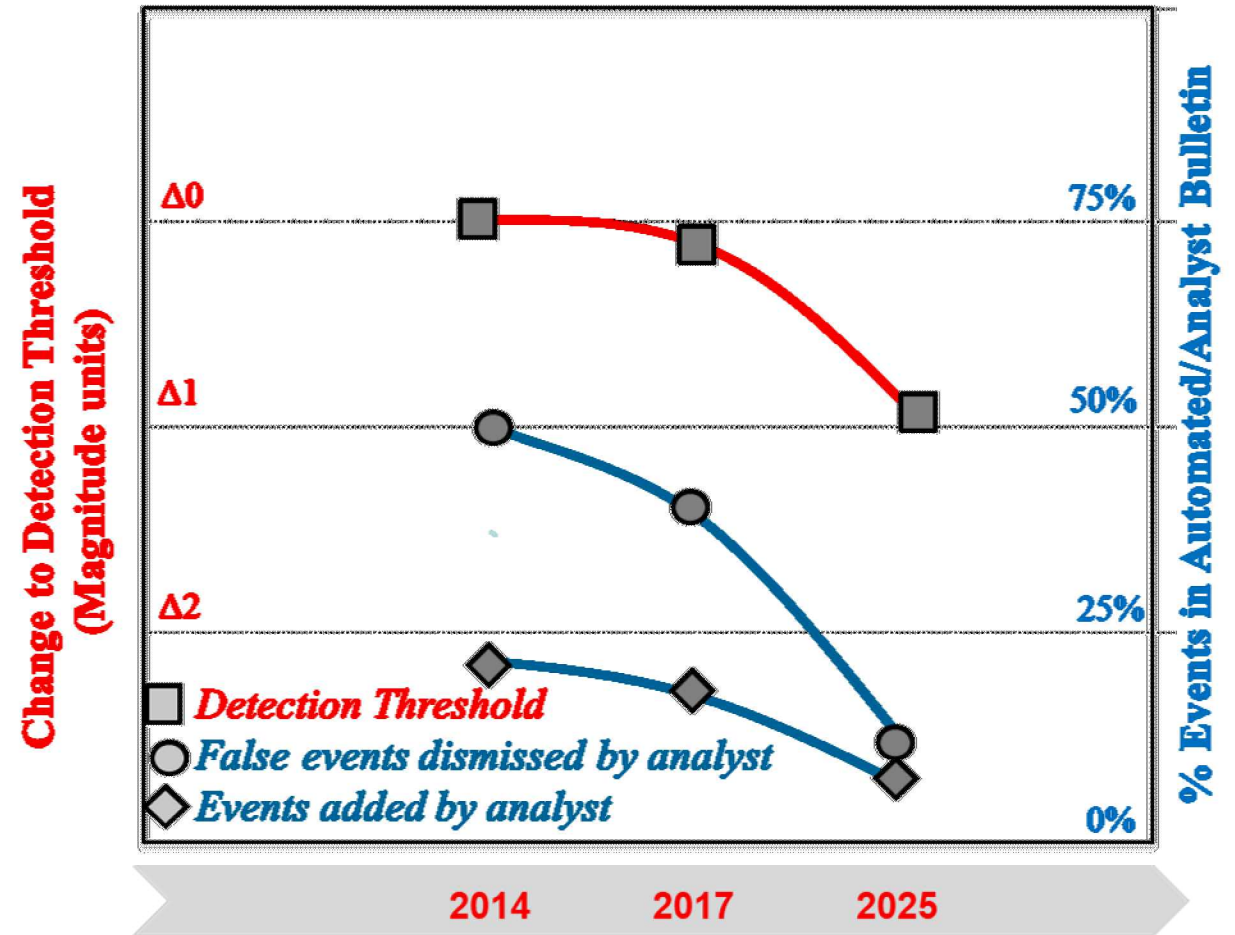
Advances in Signal Analysis support improvements to the detection thresholds and reduced analyst workload.



GNDD Signal Analysis Metric

The major R&D for Signal Analysis has evolved as monitoring practices changed from relying primarily on teleseismic distances (>2000km) to include regional distances (<1500km).

Waveform Signal Analysis Metric: Lower Event Detection Thresholds while Decreasing Analyst Workload

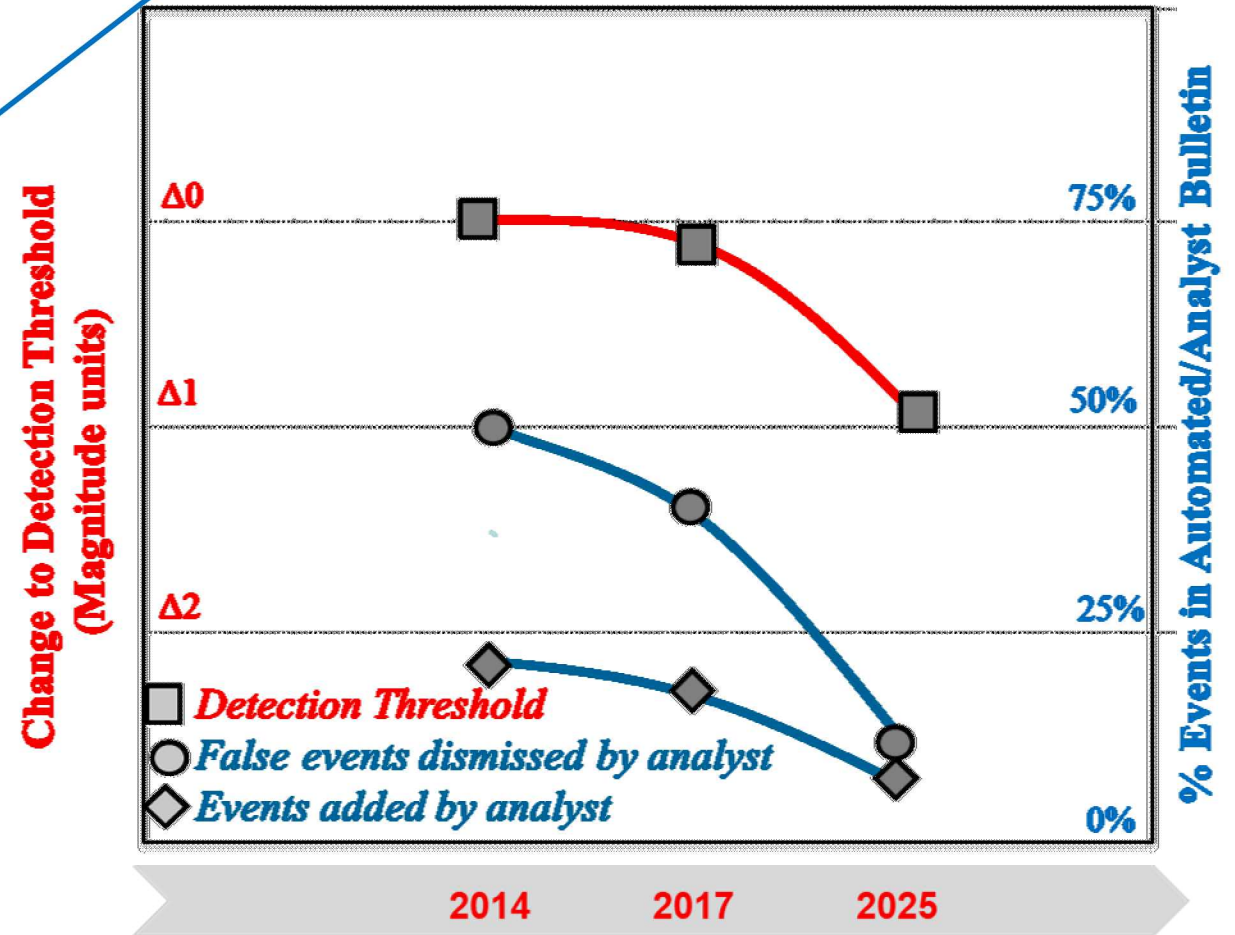


GNDD Signal Analysis Metric

The major R&D for Signal Analysis has evolved as monitoring practices changed from relying primarily on teleseismic distances (>2000km) to include regional distances (<1500km).

The overarching metric for Signal Analysis is to improve the ability to detect events of interest while decreasing workload for the human analyst.

Waveform Signal Analysis Metric: Lower Event Detection Thresholds while Decreasing Analyst Workload

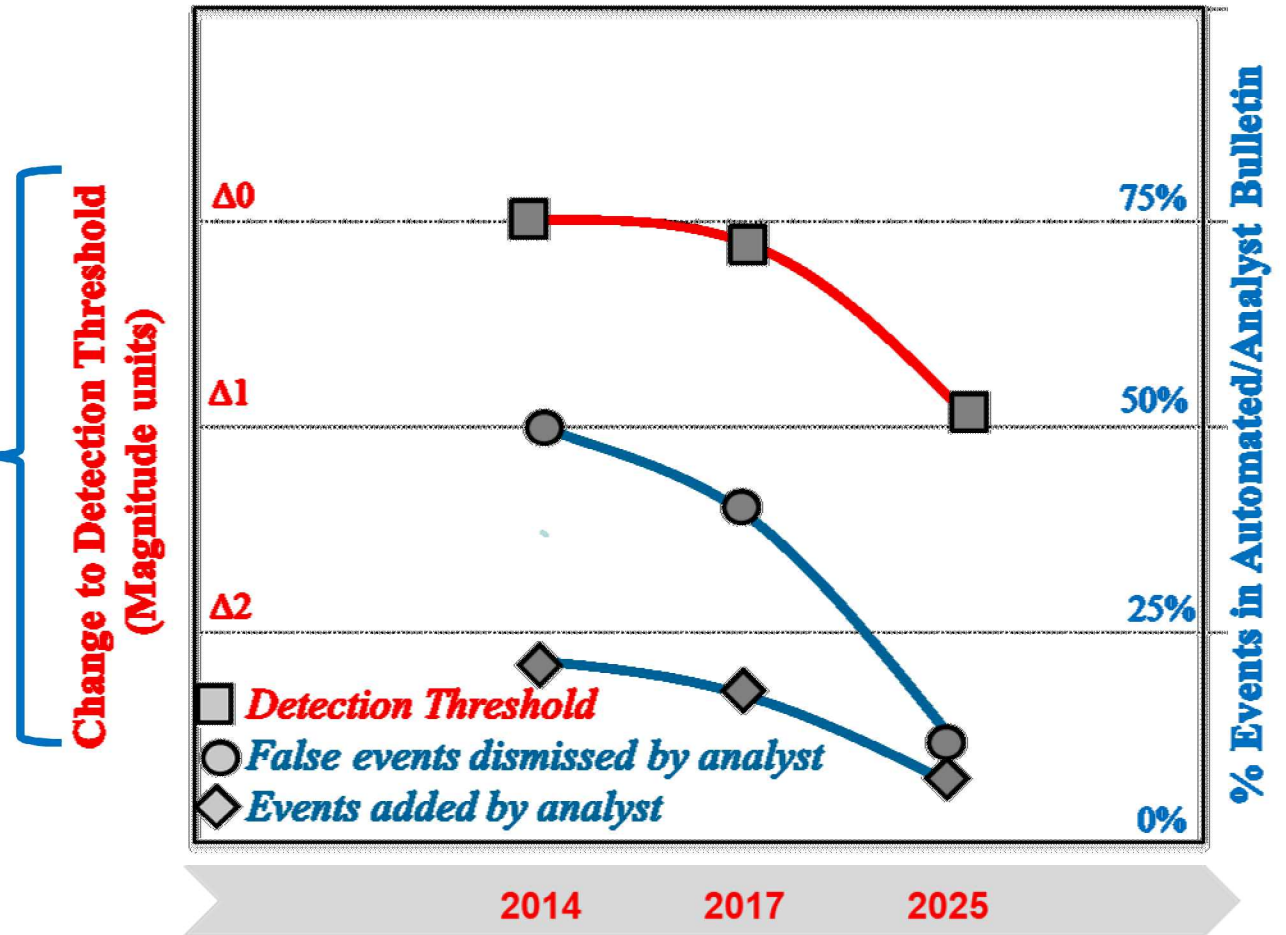


GNDD Signal Analysis Metric

The major R&D for Signal Analysis has evolved as monitoring practices changed from relying primarily on teleseismic distances (>2000km) to include regional distances (<1500km).

One of the ways to achieving this metric is the lowering of the minimum detectable threshold

Waveform Signal Analysis Metric: Lower Event Detection Thresholds while Decreasing Analyst Workload

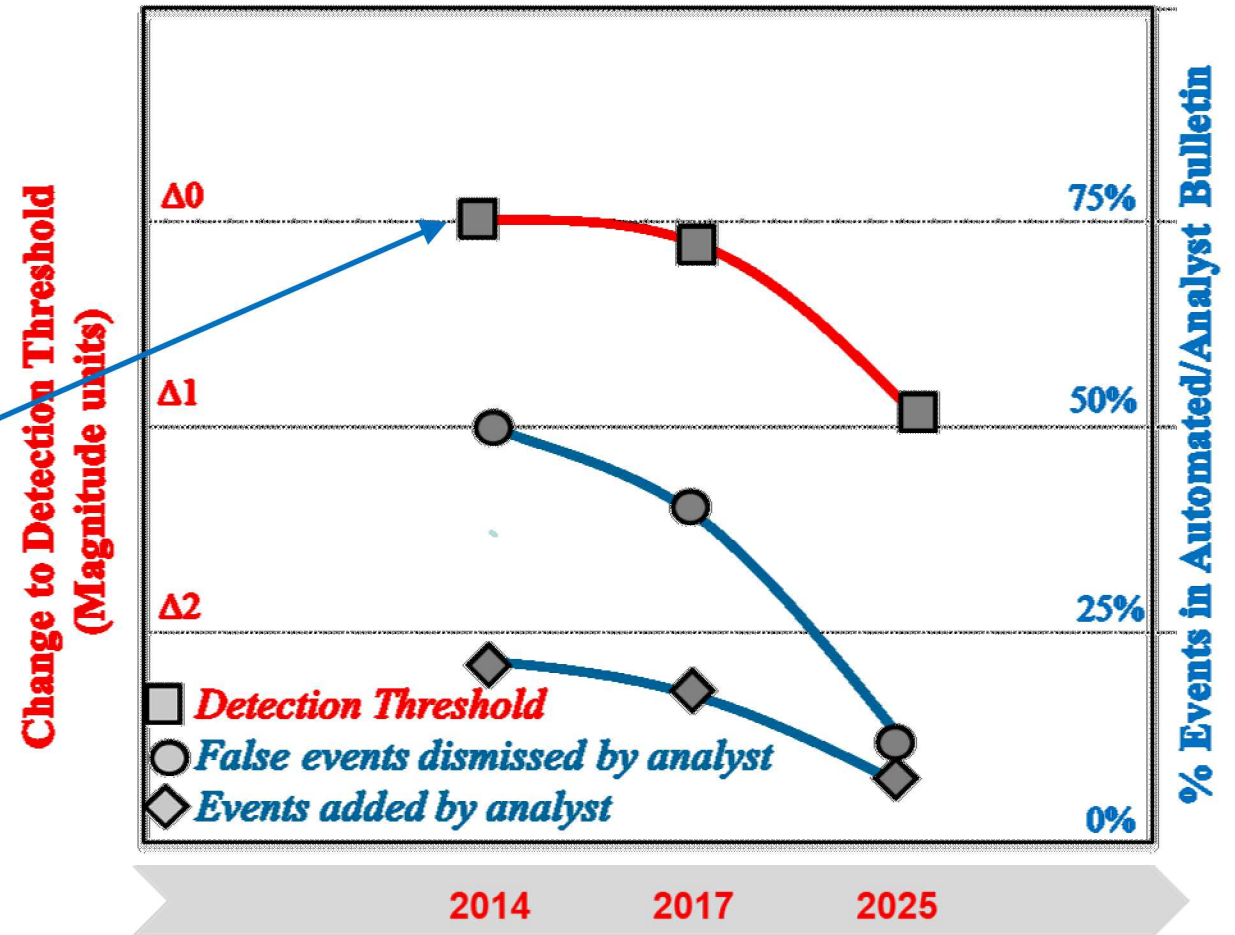


GNDD Signal Analysis Metric

The major R&D for Signal Analysis has evolved as monitoring practices changed from relying primarily on teleseismic distances (>2000km) to include regional distances (<1500km).

Advances in algorithms used to process the waveform time series, such as utilizing data archives to identify signatures of interest, are expected to result in improvements to the detection threshold.

Waveform Signal Analysis Metric: Lower Event Detection Thresholds while Decreasing Analyst Workload

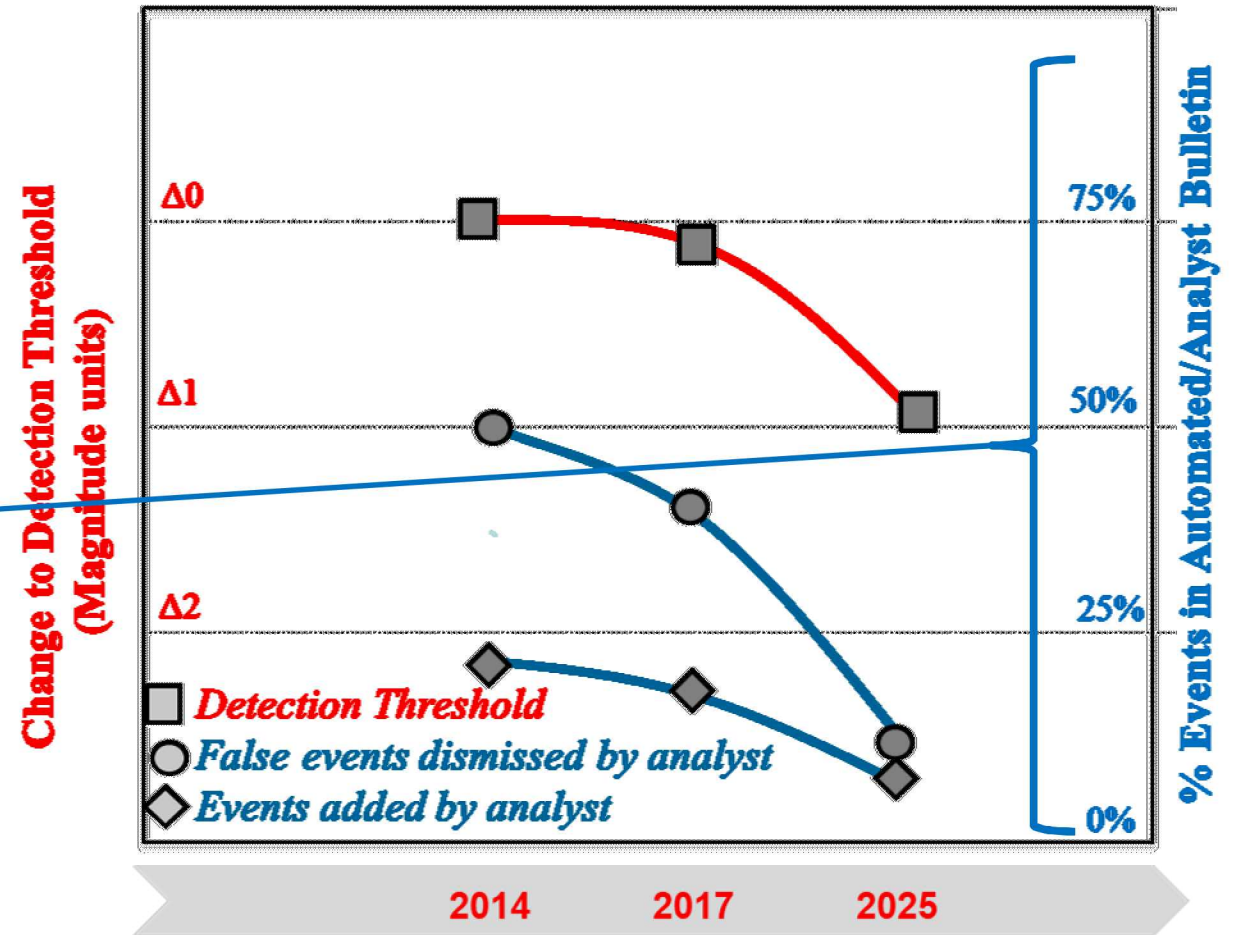


GNDD Signal Analysis Metric

The major R&D for Signal Analysis has evolved as monitoring practices changed from relying primarily on teleseismic distances (>2000km) to include regional distances (<1500km).

A way to decrease the Analyst workload is to increase the number of events that are correctly built by the automated processing system, using signal detections.

Waveform Signal Analysis Metric: Lower Event Detection Thresholds while Decreasing Analyst Workload



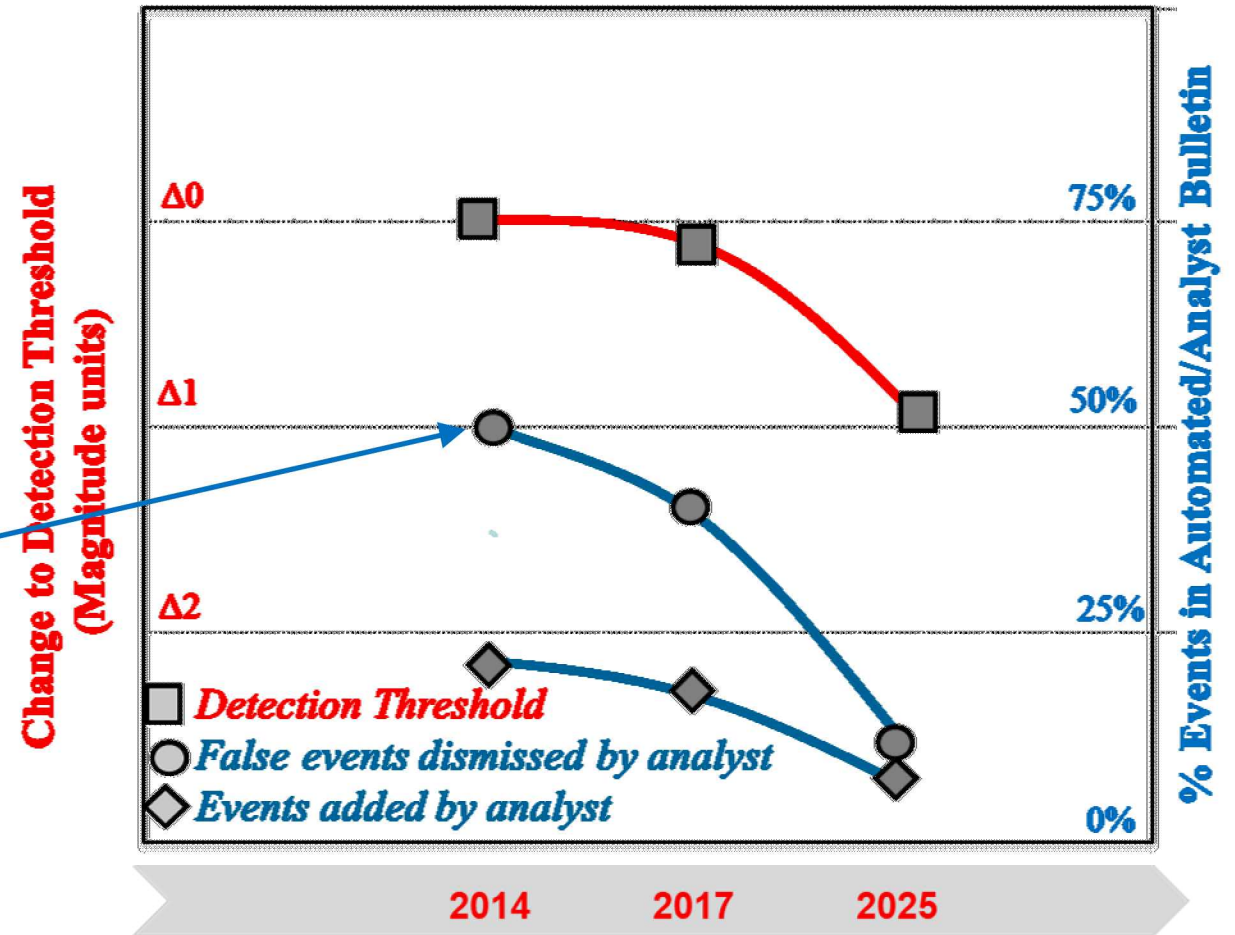
GNDD Signal Analysis Metric

The major R&D for Signal Analysis has evolved as monitoring practices changed from relying primarily on teleseismic distances (>2000km) to include regional distances (<1500km).

False events built by automated systems must be reviewed and discarded by human analysts, requiring significant time.

Automated bulletins may contain up to 50% false events.

Waveform Signal Analysis Metric: Lower Event Detection Thresholds while Decreasing Analyst Workload



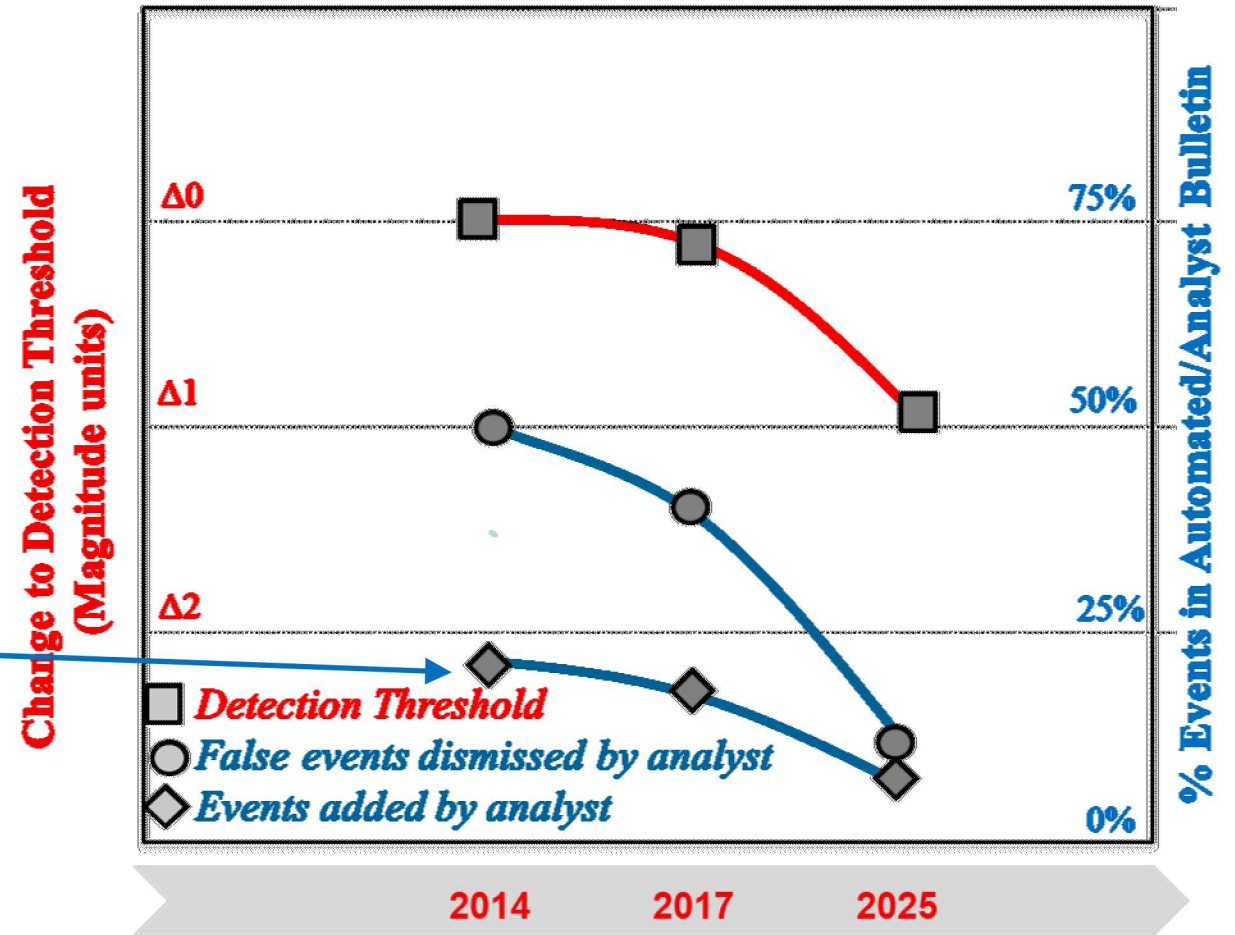
GNDD Signal Analysis Metric

The major R&D for Signal Analysis has evolved as monitoring practices changed from relying primarily on teleseismic distances (>2000km) to include regional distances (<1500km).

True events, missed by automated systems, must be manually built, also requiring significant resources.

Automated bulletins miss roughly 25% of the events included in final, reviewed bulletins

Waveform Signal Analysis Metric: Lower Event Detection Thresholds while Decreasing Analyst Workload



Generalized Likelihood Ratio Test detection using a dynamic correlation processor (DCP)

[lower detection threshold]

A DCP deliberately constructed to exhaustively sample the signal space around a target consists of M simultaneous correlators, one for each template. These correlators operate simultaneously and, for any particular trigger exceeding the threshold, the DCP picks the one with the maximum sample correlation.

This maximizes the ratio:

$$\frac{(\mathbf{u}_\theta^T \mathbf{r})^2}{\mathbf{r}^T \mathbf{r}}$$

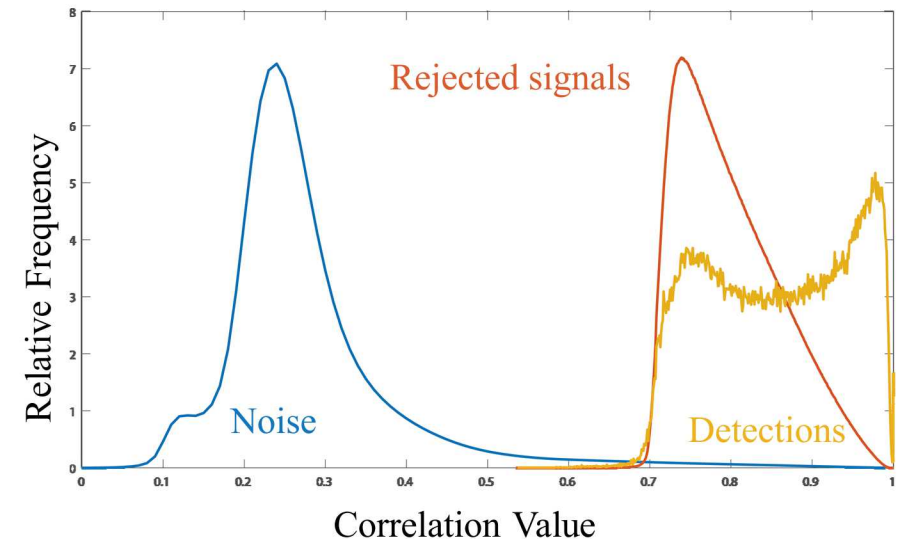
This, in turn can be shown to maximize the likelihood ratio:

$$\hat{\Lambda}(\mathbf{r}) = \frac{\max_{A, \sigma, \theta} p_1(\mathbf{r})}{\max_{\sigma} p_0(\mathbf{r})}$$

This approach is called the Generalized Likelihood Ratio Test (GLRT), and although it does not guarantee a theoretically optimum detection rule, experience has shown it usually produces a very good rule.

We apply this to data recorded by station ROA in the Book Cliffs mining region of Eastern Utah. The time interval of the experiment is 1 year (2010).

Preliminary results indicate the DCP was very successful in rejecting non-target signals.



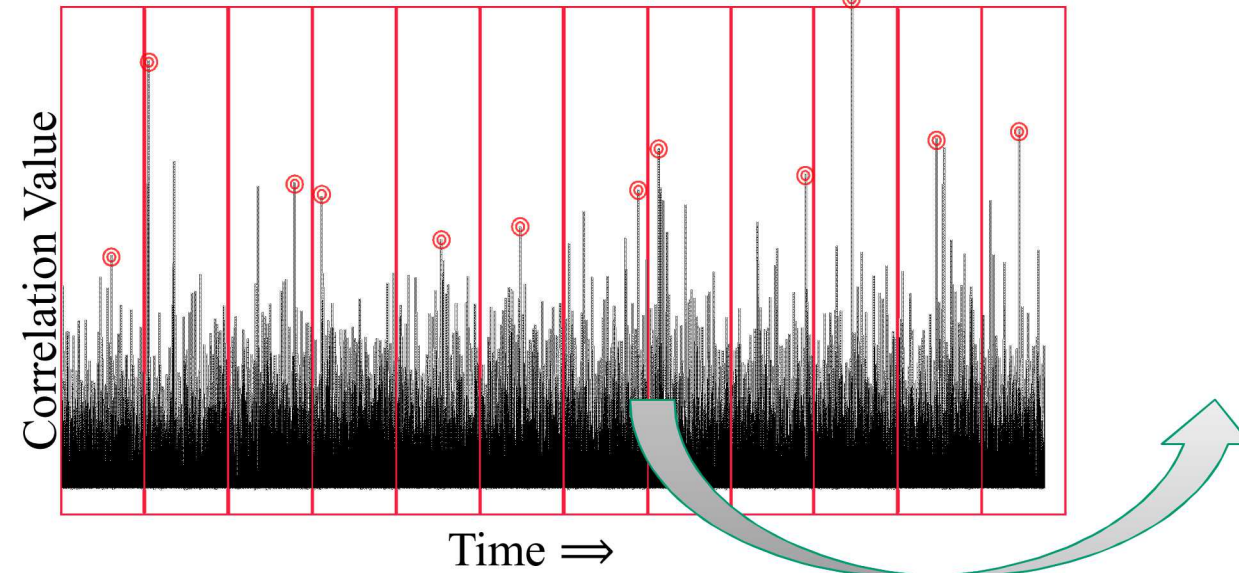
Automated threshold update algorithm in dynamic correlation processor (DCP)

[lower detection threshold]

(IRP 1.46)

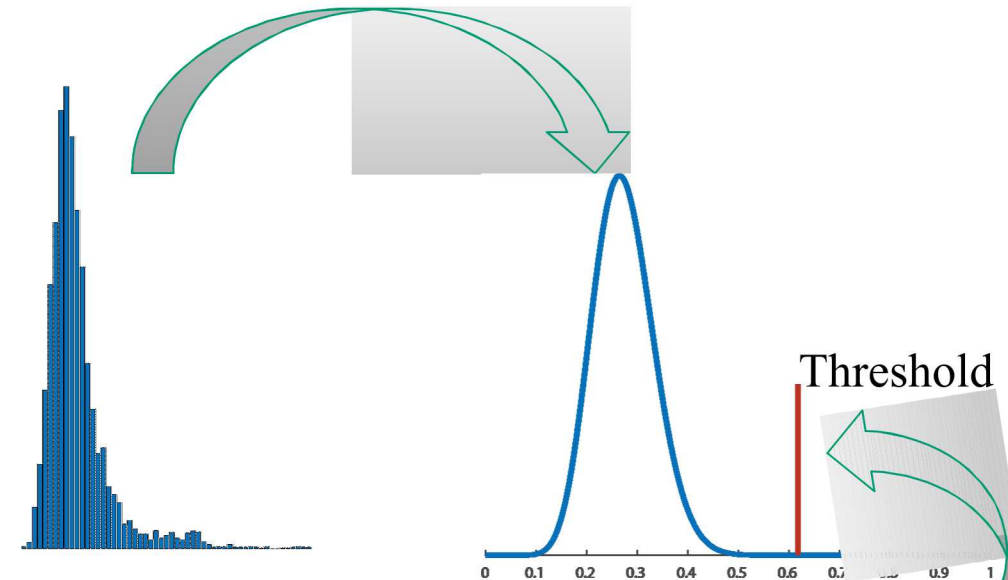
DCP processes data in blocks.

Snapshot of detection statistic for one detector



Accumulate from each block, the max
value of the detection statistic.

Periodically, estimate parameters (α, β) of
Beta distribution

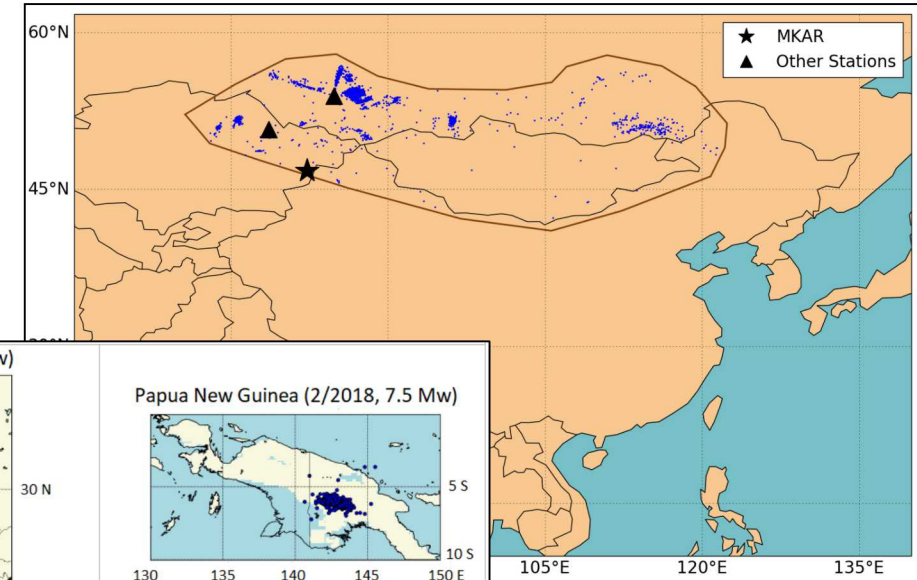


Estimate threshold for a fixed
false-alarm rate
Using inverse CDF (α, β).

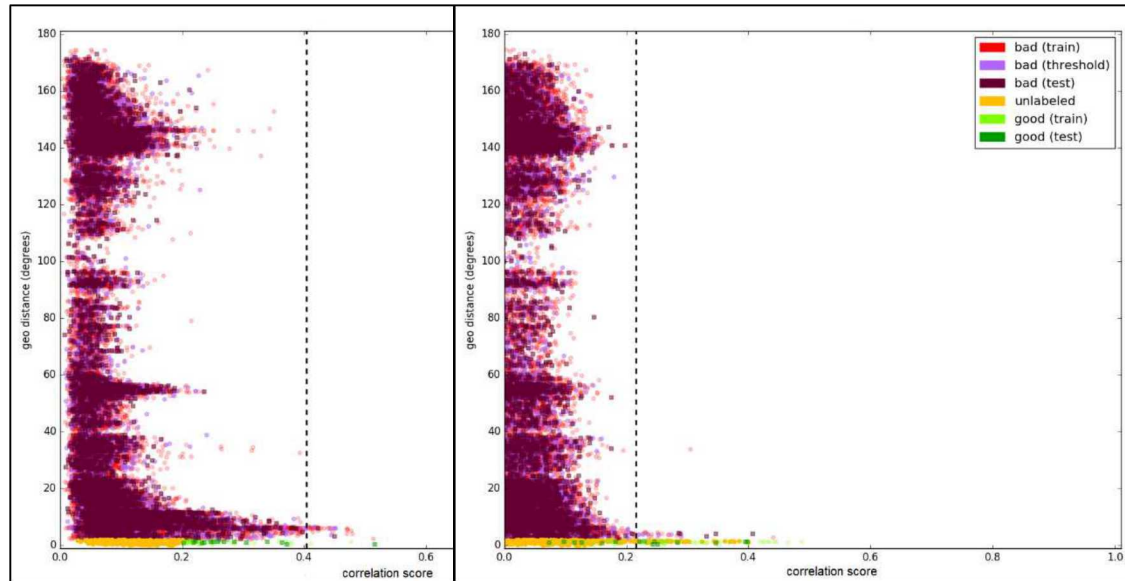
[lower detection threshold and false alarms]

Waveform Correlation Analysis to improve detection:

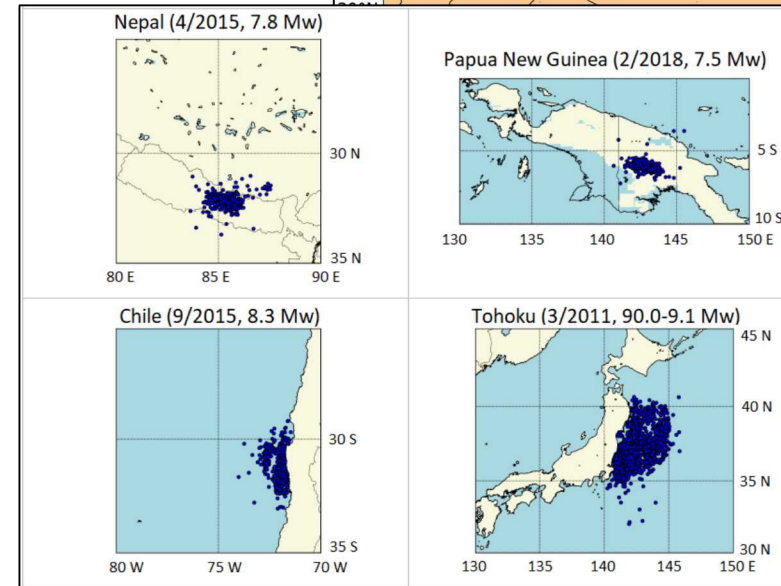
- **Alternative Null Hypothesis Correlator (ANCorr)** improves false alarm rate
- **Regional Waveform Correlation Study in Mongolia**
(Collaborative with Columbia University)
- **IMS Aftershock Sequence study**
(Collaborative with NORSAR and CTBTO).



Mongolia Regional Study



ANCORR Improved Detection Threshold

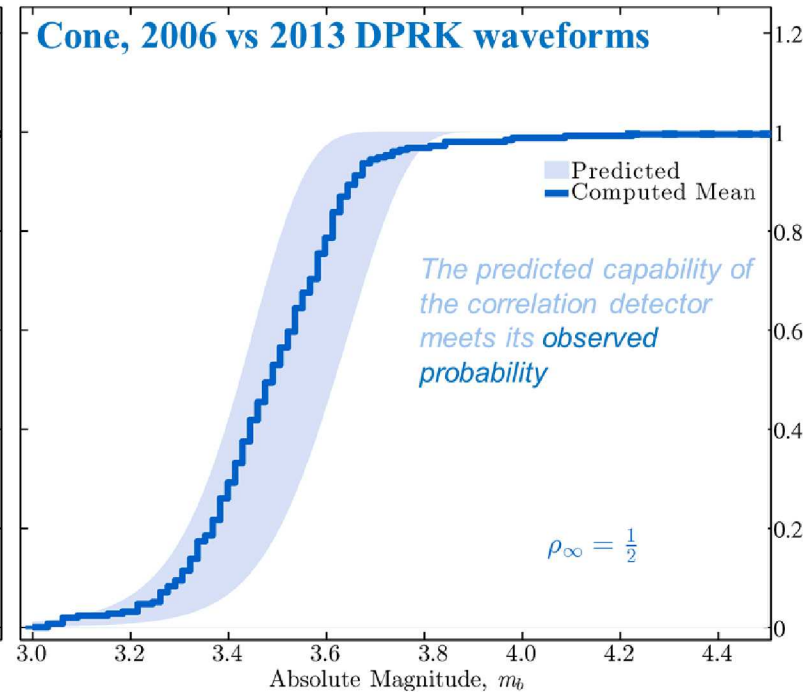
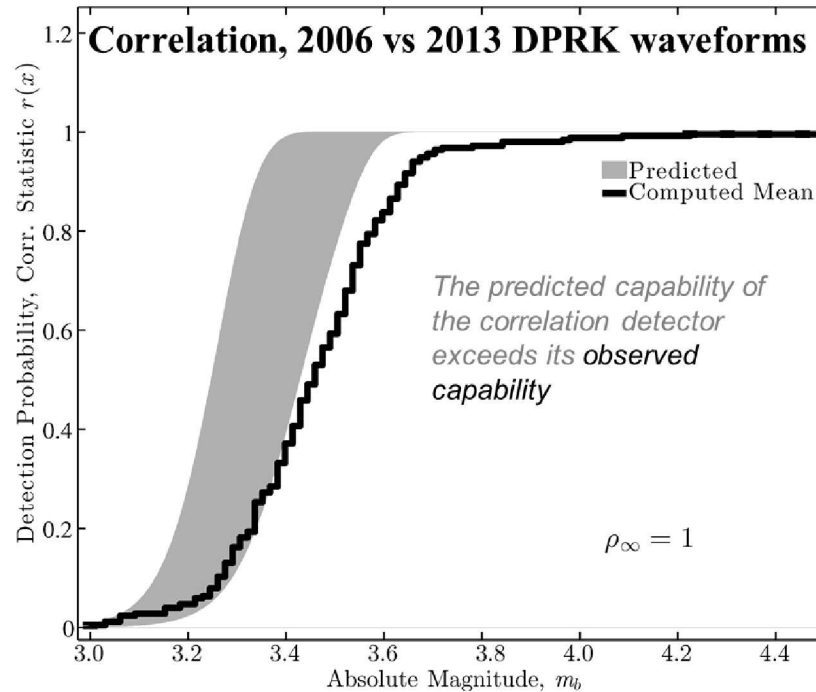
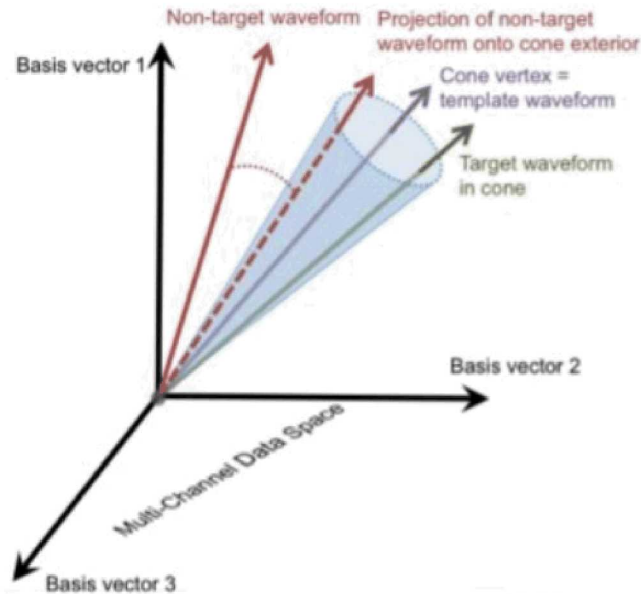


Aftershock Sequence Study

Cone Detector Algorithm

[lower detection threshold and false alarms]

- Correlation detectors match waveform shapes. Cone Detectors implicitly model target waveforms to partially match templates. They can achieve their predicted capability, with fewer false alarms.
- The cone statistics show less variability between waveform detections, compared to the correlation detector. All waveforms are more similar within the cone.



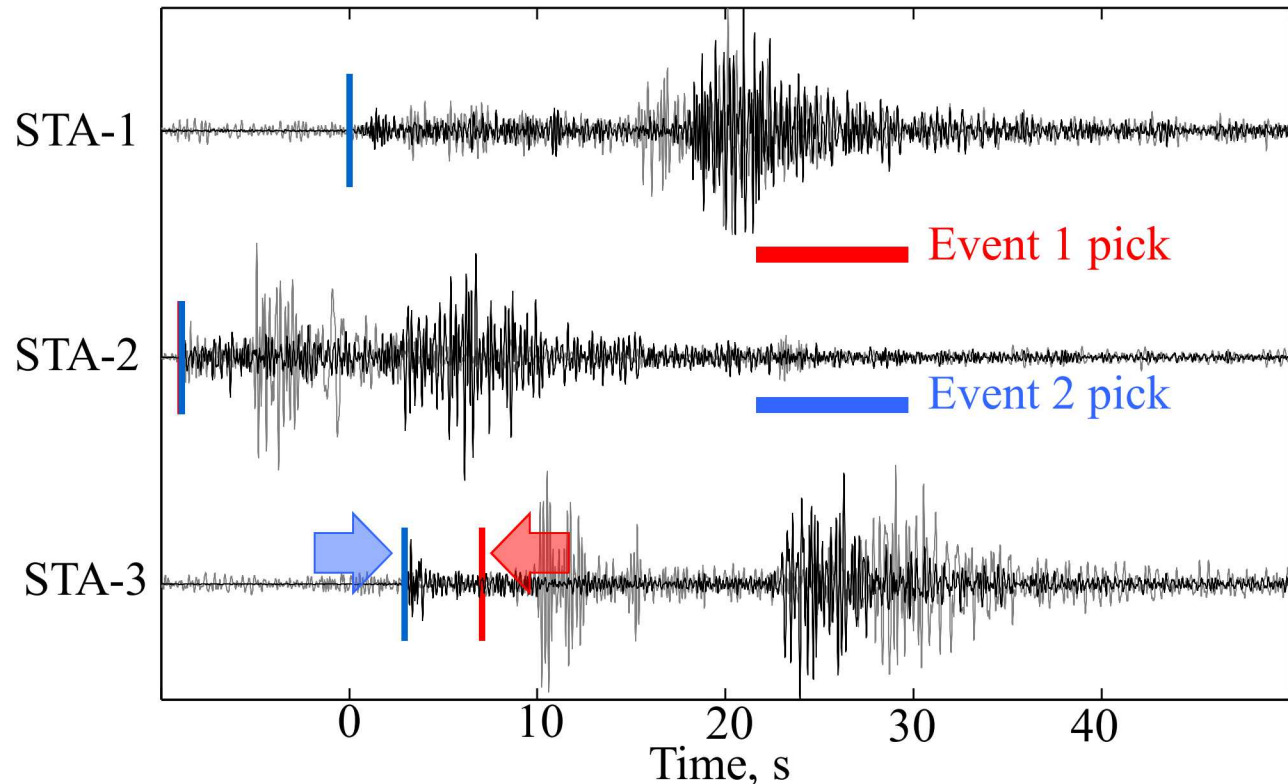
Carmichael, Joshua D., **A Waveform Detector that Targets Template-Decorrelated Signals and Achieves its Predicted Performance: Demonstration with IMS Data (Part I)**, *Bulletin of the Seismological Society of America*, doi: 10.1785/0120160047

Waveform Association with Catalogue Events

[decreasing analyst workload]

Associate with “Template Waveforms” that are consistent with moveout pattern

- 1 Process waveform data recorded on each station with a detector
- 2 Compare inter-station detection time differences with those for reference event, as computed by same detector
- 3 Associate events if inter-station double time differences match for the reference, within error tolerance
- 4 Result: About 30 events per day fit criteria that locate within 50 km of a reference event and were **not otherwise cataloged, from ROI**

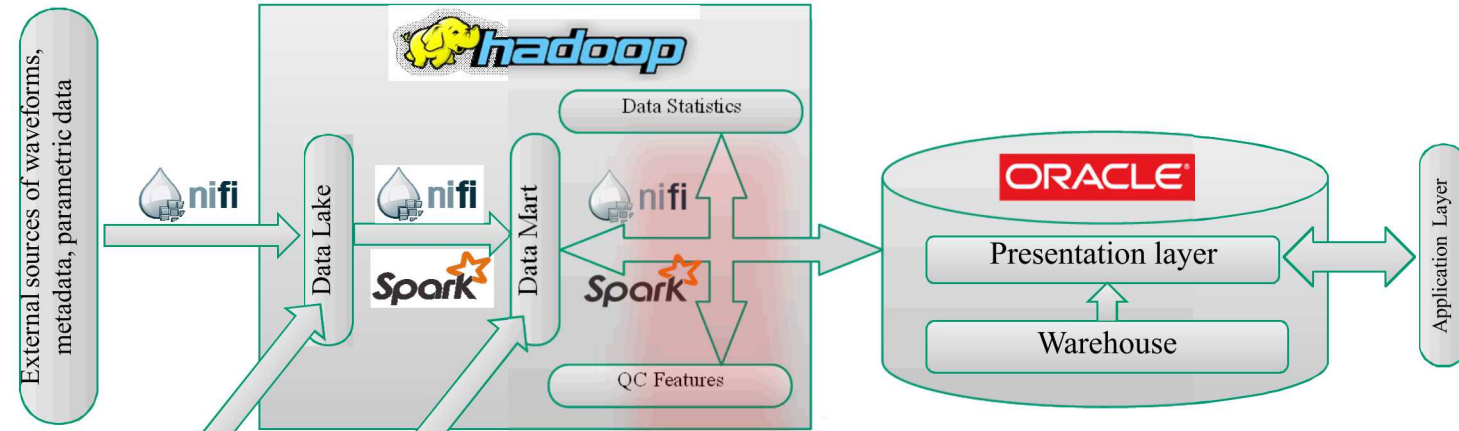


*Reference seismograms for sources within a region of interest (ROI). The associator fuses detection times with manual phase picks. The double difference in phase-arrivals/pick-times from the ROI will be similar for the target source. This reference, double difference in phase arrivals is an **association constraint**.*

Advanced Processing Redesign

[decreasing analyst workload]

- Extended keys for seismic entities accommodate complexity of modern data
- Data mart layer fully normalized
- Improved origin ranking and merging strategies
- Automated ingestion from FDSN sources
- Dynamic generation of presentation layer

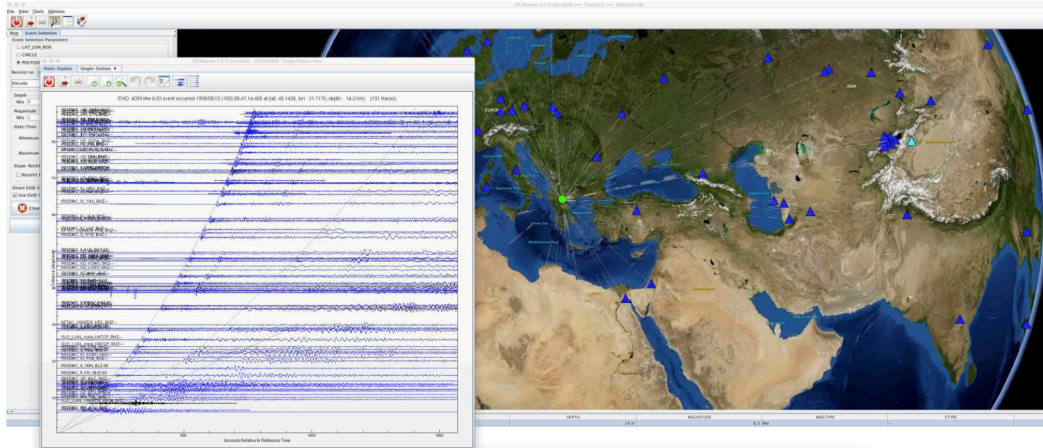


All ingested data are retained unaltered.

consistency transformations (e.g. format, null values, units, etc.)

Data analytics plays a key role in generation of the presentation layer.

Database browsing application ported



Status of bulletin processing:

FDSN WS-event loading code is in testing and NiFi processors are under development.

Status of metadata processing:

ISC metadata loader is fully operational

FDSN WS-station loading code is in testing

Status of waveform processing:

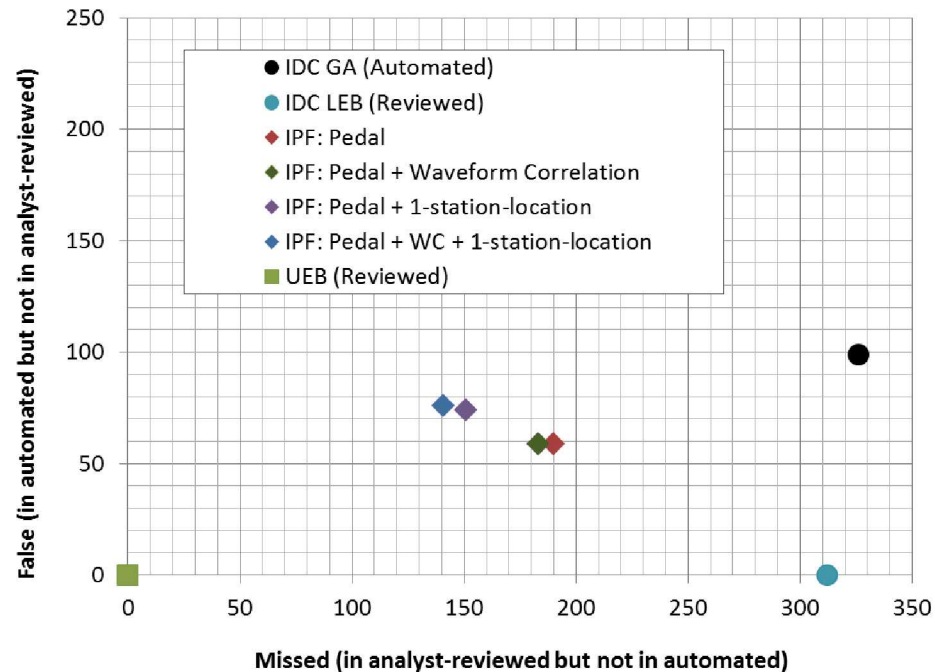
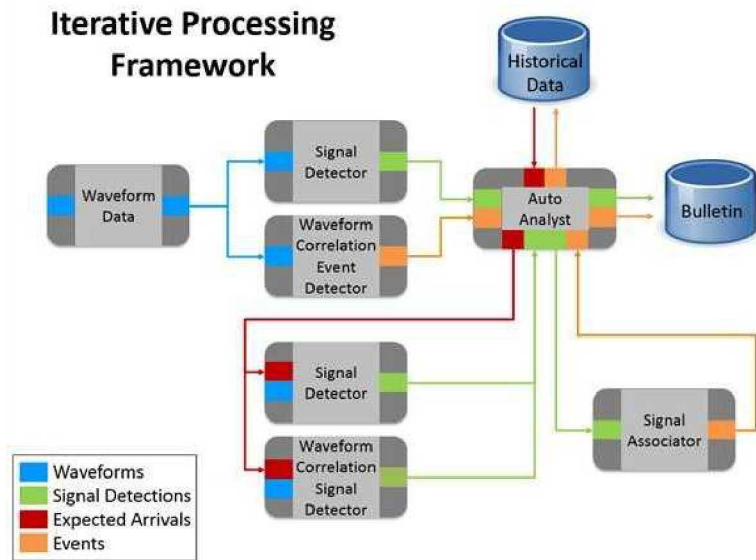
FDSN waveform loading code core module has been developed

Iterative Processing Framework Report

[lower detection threshold and decreasing analyst workload]

Identify human analyst behaviors and implement those action taken in automated bulletin generation to improve performance.

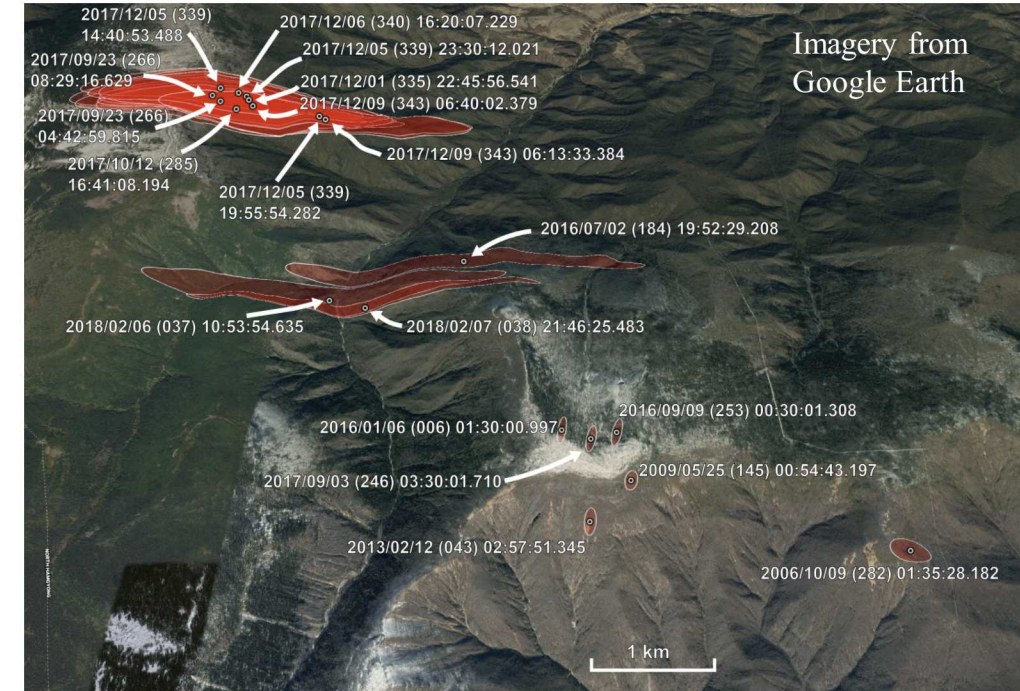
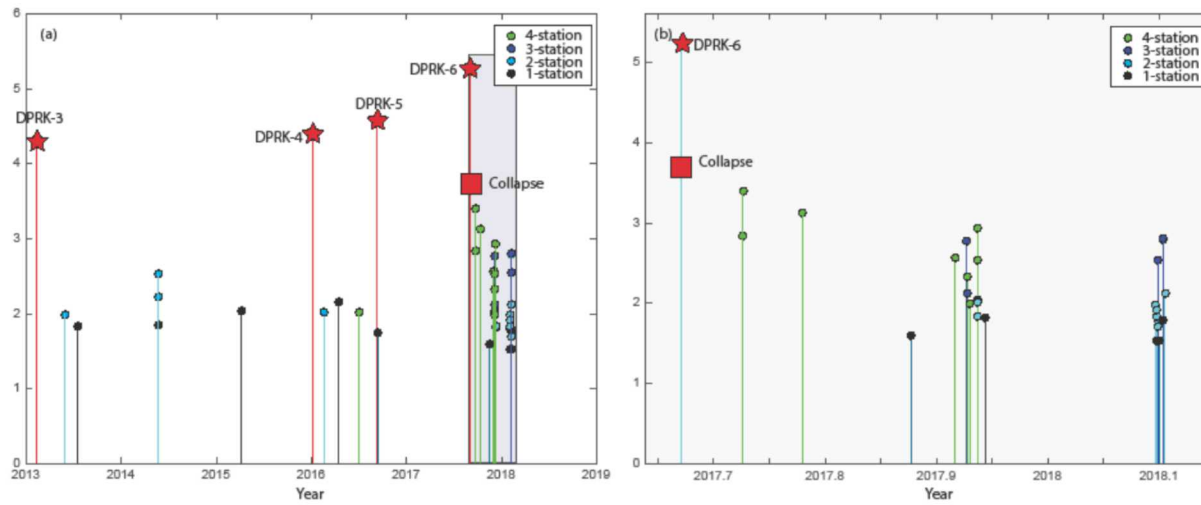
- Use historical data to guide processing.
- Dynamically adjust parameters to look for features that should be present but weren't.



- Improve detection rates while reducing human analyst time required (missed true and found false events)

Signals from post DPRK-6 earthquakes recorded at 4 stations were used to build subspace detectors which were then used to scan the available data for previously unobserved events near the test site. More than 30 additional template-matched events were detected. Thirteen of the detected events have been located. These events form two clusters, both to the northwest of the test site.

Prior to DPRK-3, no events were observed from either cluster, but since DPRK-3 there has been at least one event per year.



Seismological Research Letters ; 89 (6): 2094–2112. doi: <https://doi.org/10.1785/0220180127>

Ground Truth Catalogue Development

[decreasing analyst workload]

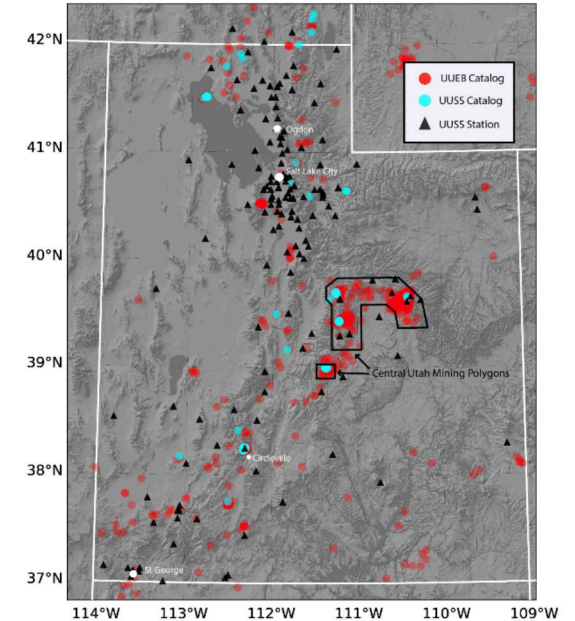
Developed 2 two-week “GT” reference catalogues to support algorithm development.

– **Unconstrained Global Event Bulletin (UGEB)**

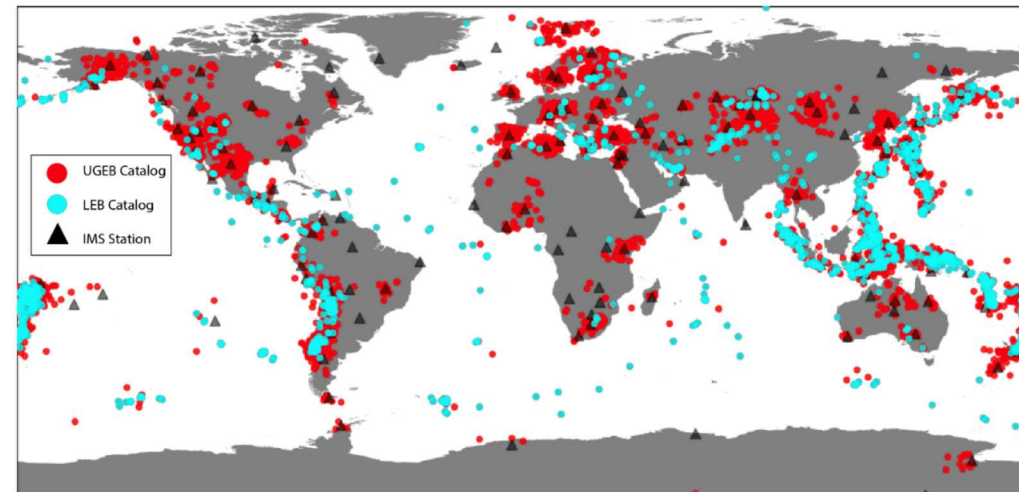
- May 15-28, 2010, IMS primary and auxiliary networks.
- 11,378 events (10 months of analyst time)

– **Unconstrained Utah Event Bulletin (UUEB)**

- Jan 1-14, 2011, UUSS network
- 7,889 events (6 months of analyst time)



Upcoming SRL Paper with datasets available.



Algorithm for Amplitude Quality Control using Midnight Noise

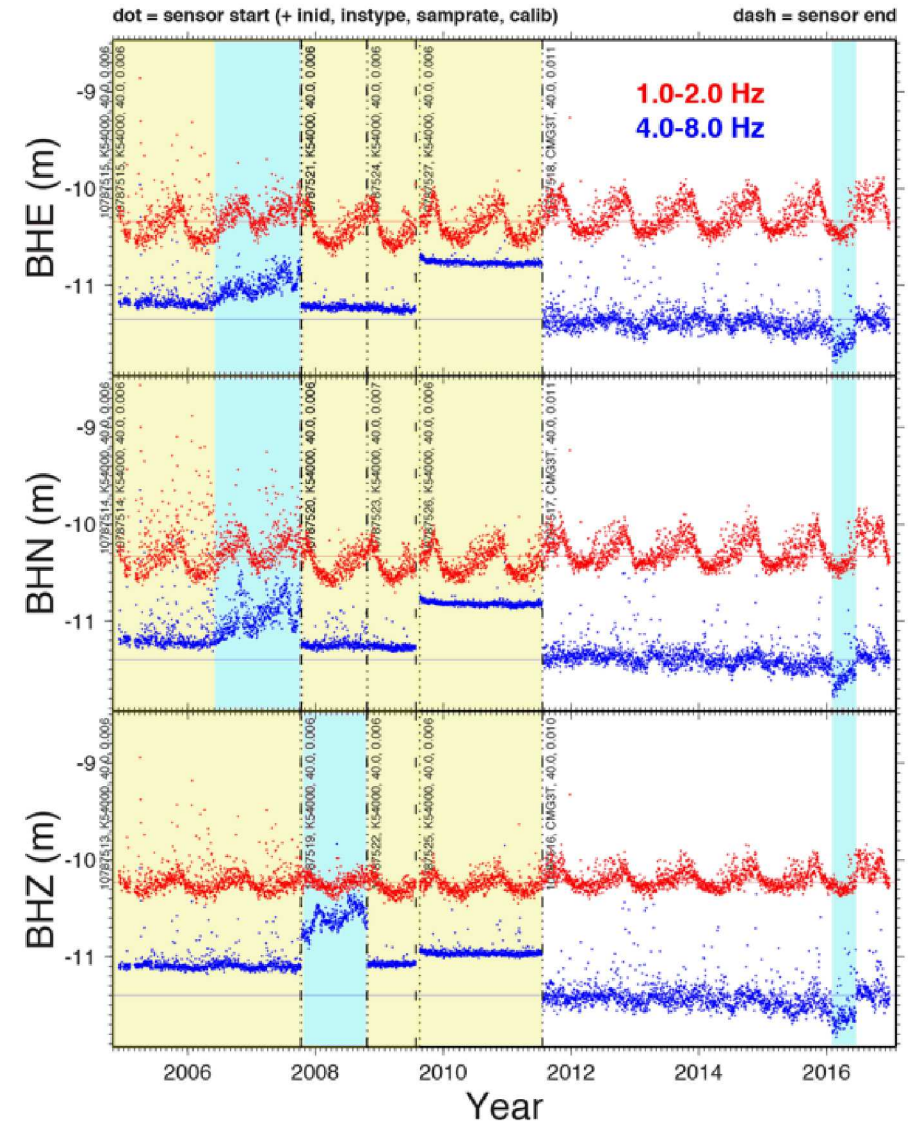
[decreasing analyst workload]

Developing algorithms to perform quality control using noise levels at night, when background is lower and is be more consistent day-to-day.

- Problems indicate an issue with the station response.
- Screen suspected data from further research analysis (amplitude and coda corrections)
- Alert data center and station operators to fix issues.

Analysis Process:

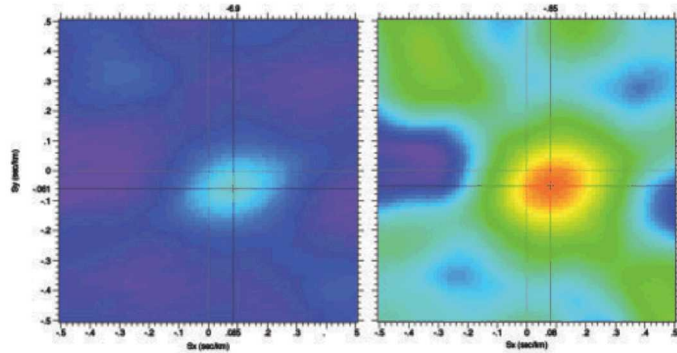
- Global coverage, focus on Asia (in house continuous data)
- Time domain RMS, band-passed traces, also use IRIS MUSTANG noise mode
- 0.25×10^9 noise measurements, 5 high-frequency bands, 16,000 stations
- 21,000 manually reviewed channels, 15,000 suspect intervals
- Sensor start/end marked
- Suspect intervals chosen manually
 - Level changes (yellow)
 - Drift (cyan)



Array Quality Control Report

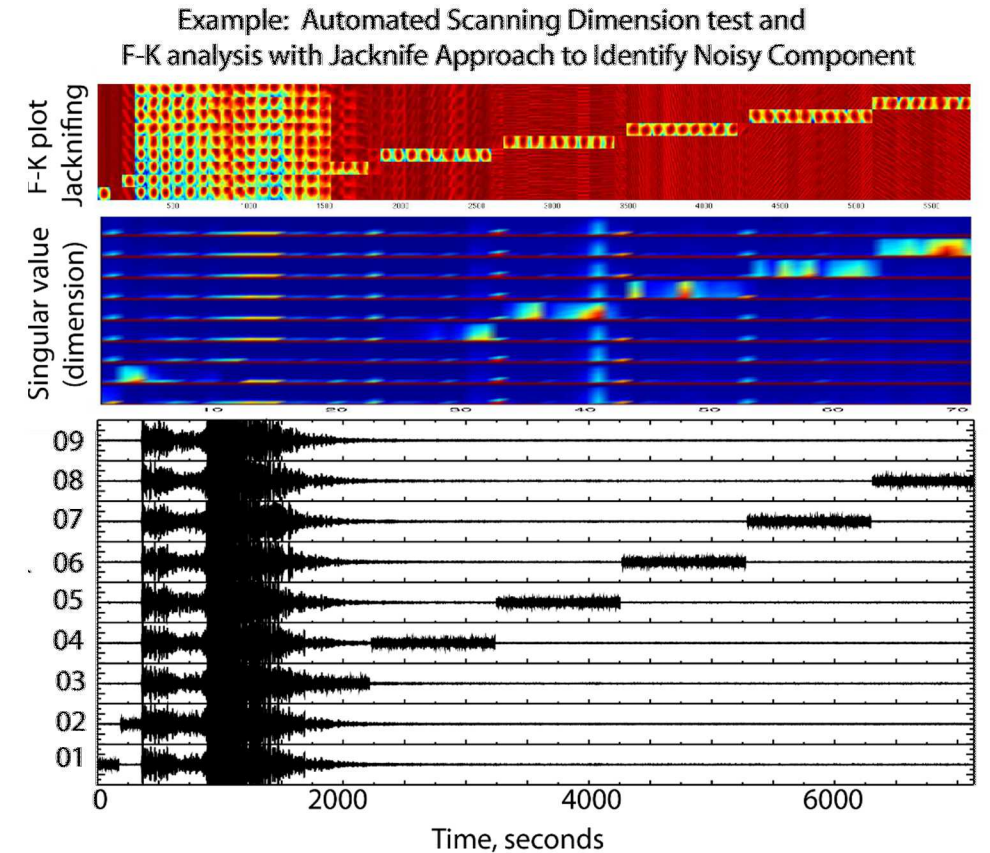
[decreasing analyst workload]

The goal of this work is to build an **automated algorithm that can perform quality checking on array components** to flag those channels that may hinder the usefulness of the array for a particular research objective.



FK plot for small Wenshuan aftershock observed at Makanchi Array, with (left) and without (right) noisy channels.

- Examine the SVD function to find time periods of bad, noisy or missing array elements so that an automated analysis system might flag and/or discard them.
- In particular, **automated beamforming or F-K analysis can be seriously degraded**, and the identification of an intermittently bad channel can require time-consuming human intervention.
- Build an array QC tool that performs a **dimensional estimate on the fly** to identify bad channels prior to automated F-K analysis.
- Intermittent noise of sufficiently high amplitude will not be possible to represent within the “normal” data dimension and **will thus increase the dimension of the array**.
- This **bad channel may be undetected during ordinary operations** if it is intermittent.

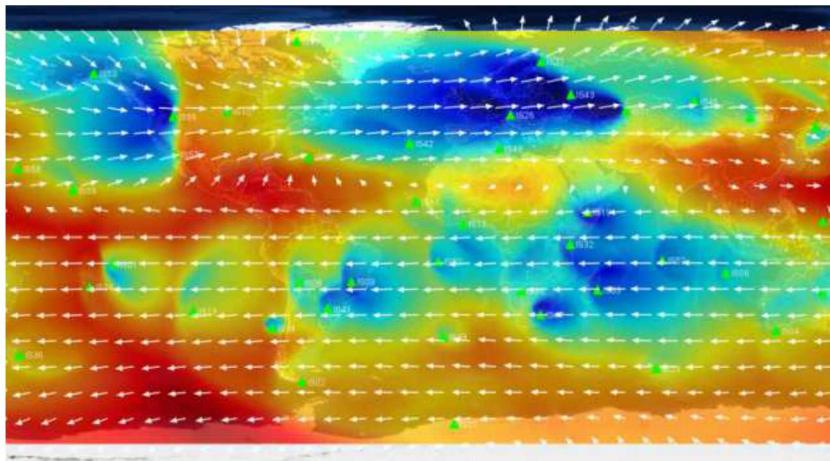


Simulation tool to assess network detection performance

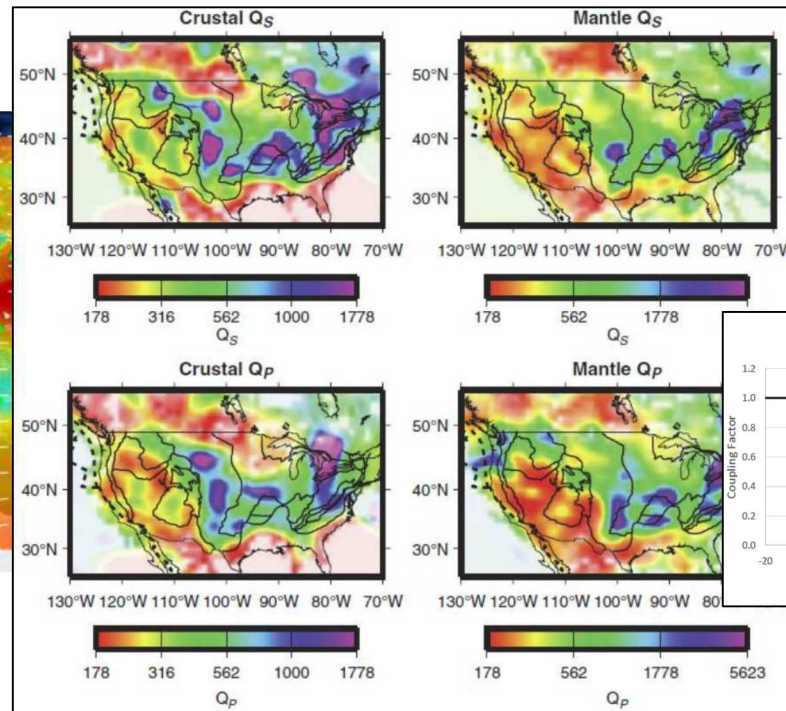
- Performed Seismic, Infrasound, and Hydroacoustic detection simulations

What's New:

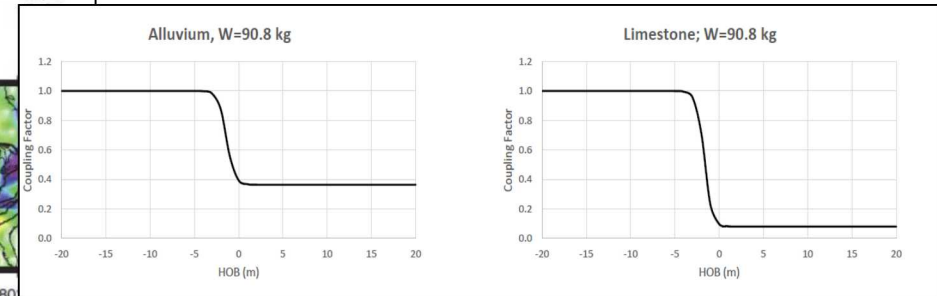
- Added G2S and ECMWF wind models to infrasound simulations
- Added 2D Q regional attenuation models to seismic simulations
- Added Height of Burst / Depth of Burial correction to seismic simulations



Example Infrasound IMS Simulation with wind



US Q Attenuation
Pasyanos (2013)



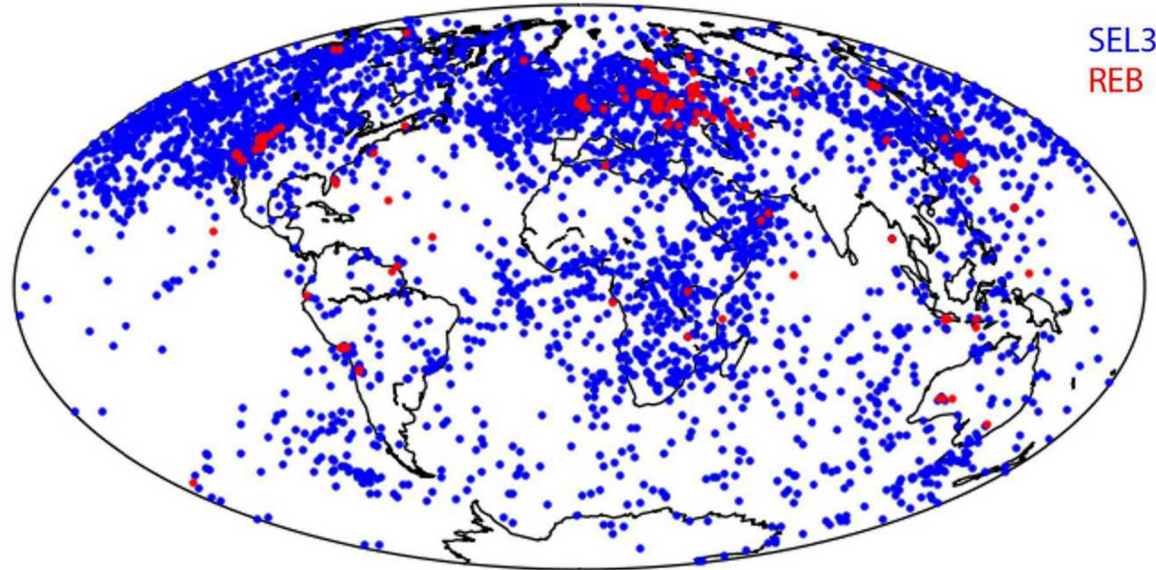
Depth of Burial / Height of Burst Correction
Pasyanos & Ford (2015)

Report on IMS Infrasound Detection

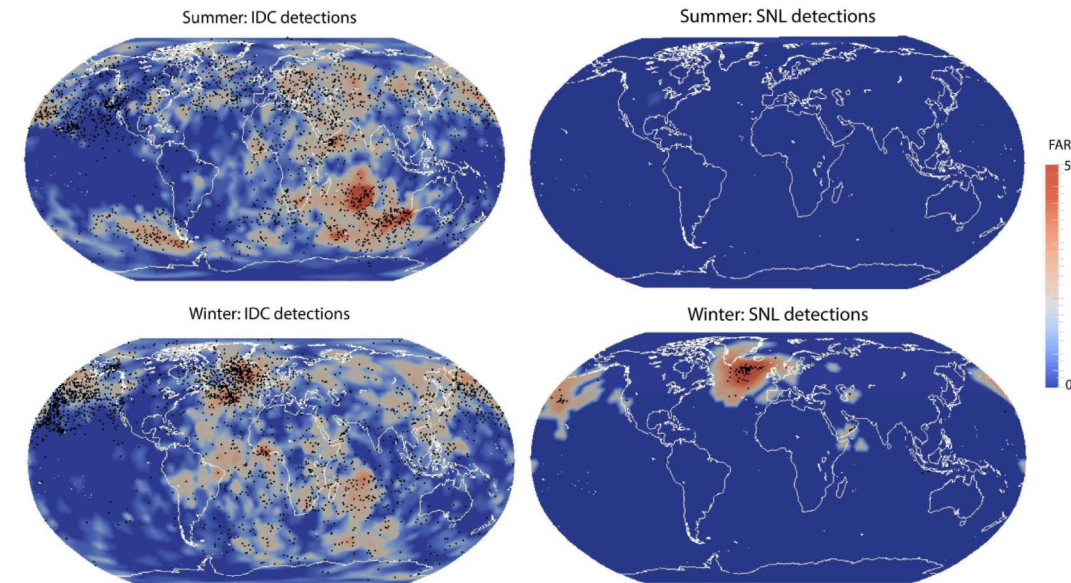
[lower detection threshold and decreasing analyst workload]

Development of algorithms to improve infrasound detection

- IMS Infrasound network is sufficient for 2-station detection requirements. However, far too many false alarms due to weather induced microbaroms:

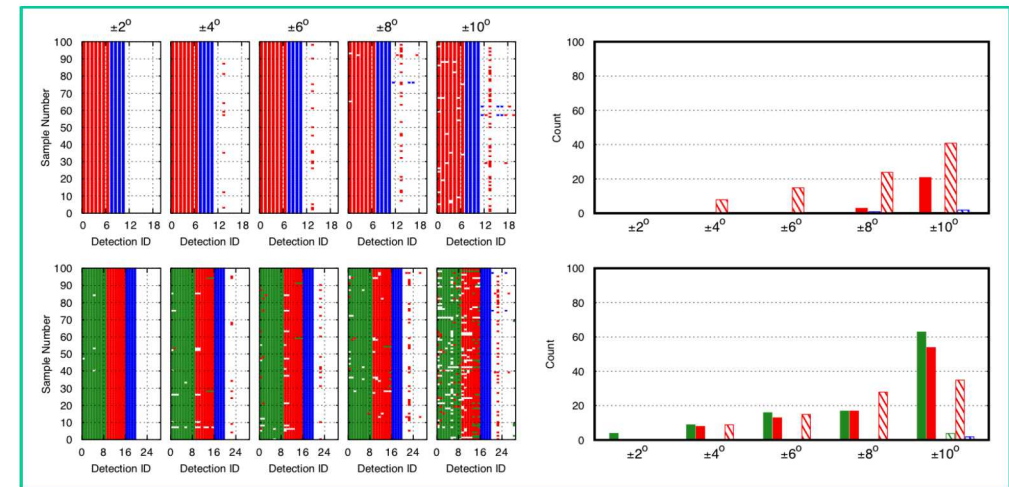
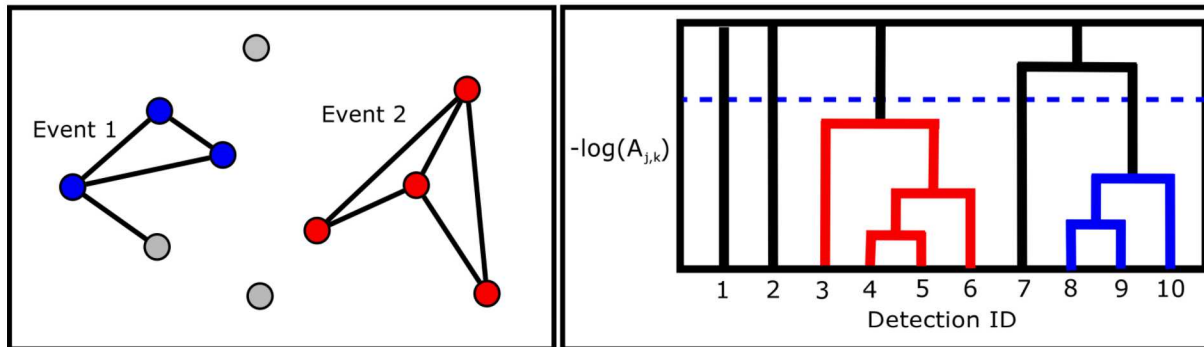


- Developed multi-feature detection algorithm (semblance, stable back azimuth, kurtosis, etc.) that maintains detections while screening nuisance microbaroms.



IRP 7.38: Python Geophysics Suite (PyGS)

- A pair-based, joint-likelihood, Bayesian approach to association has been investigated and shows promise for identifying infrasonic events
 - **Joint-likelihood analysis definition of linkages between detection pairs and enables more robust clustering analysis**
 - **Implemented in InfraPy/PyGS**
 - **Overview will be published in Geophysical Journal International**



- Method has been applied to a synthetic data set to quantify performance
- Also been applied to infrasound data in the western US with promising results

Graphical User Interface for Interactive Infrasound Signal Analysis

(IRP 7.38)

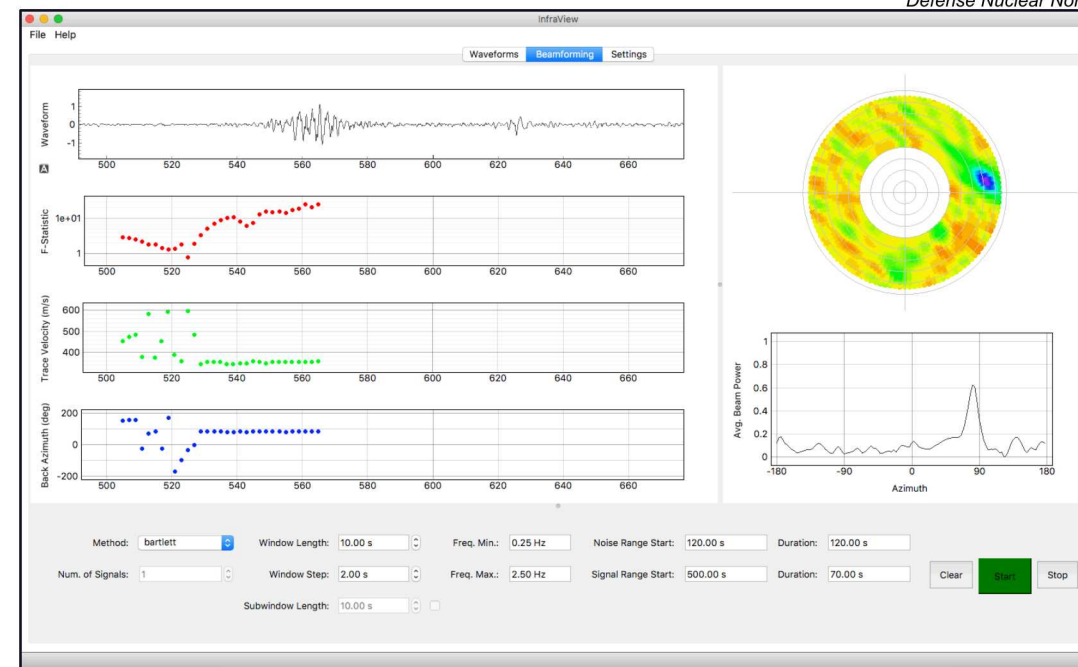
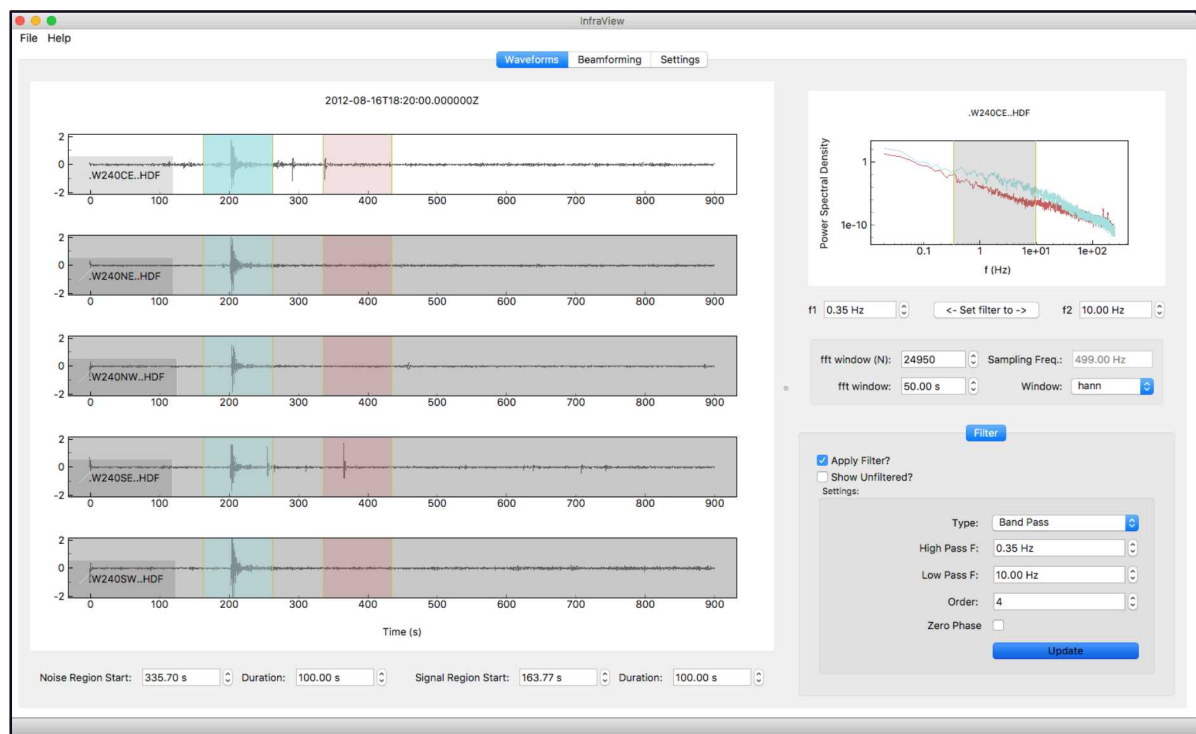


[decreasing analyst workload]



Python Geophysics Suite (PyGS)

- A Graphical User Interface (GUI) has been developed for the InfraPy software toolkit
- Enables interactive evaluation of timeseries



- Beamforming capabilities are currently implemented
- Association and localization analyses are planned for inclusion

January 29, 2019

“Very creative and innovative! Panel liked what they saw.”

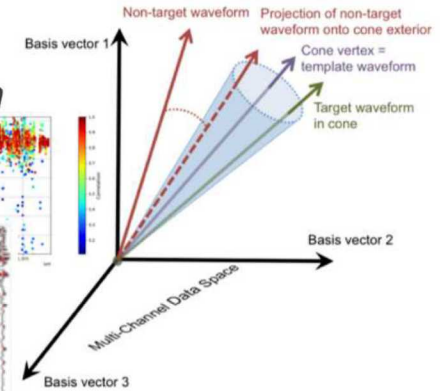
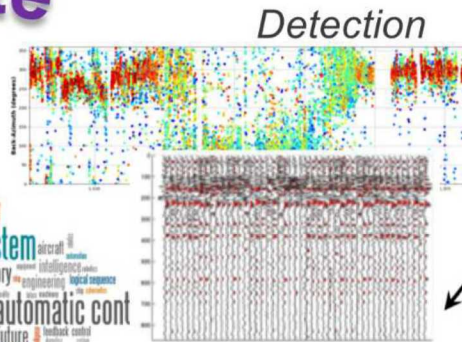
“Panel recognized the fact that the lab demonstrated world-class expertise in infrasound.”

“Demonstrated knowledge base and R&D base that were very credible.”

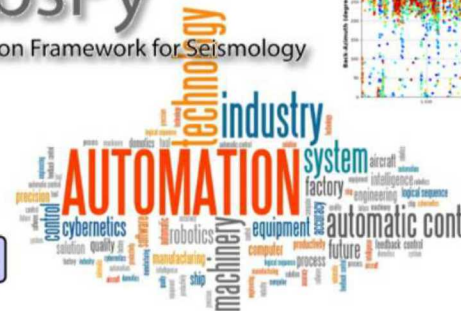
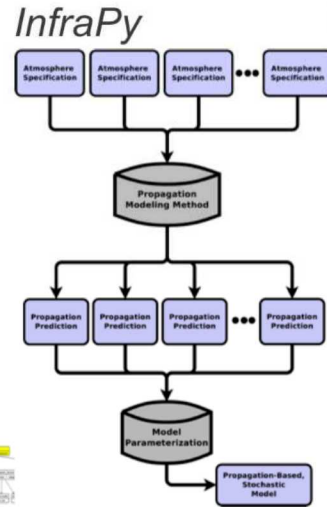
PyGS: Python Geophysics Suite



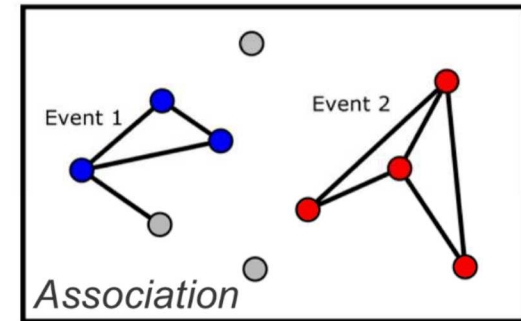
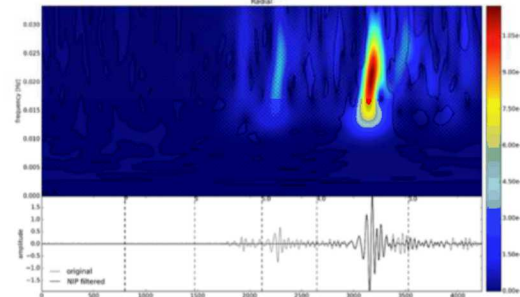
ObsPy
A Python Framework for Seismology



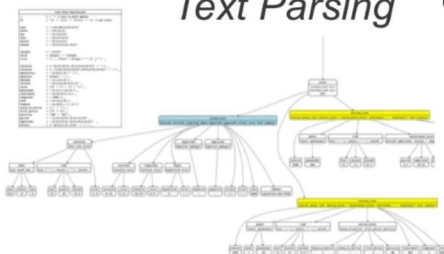
Cloud Computing



Advanced Signal Processing



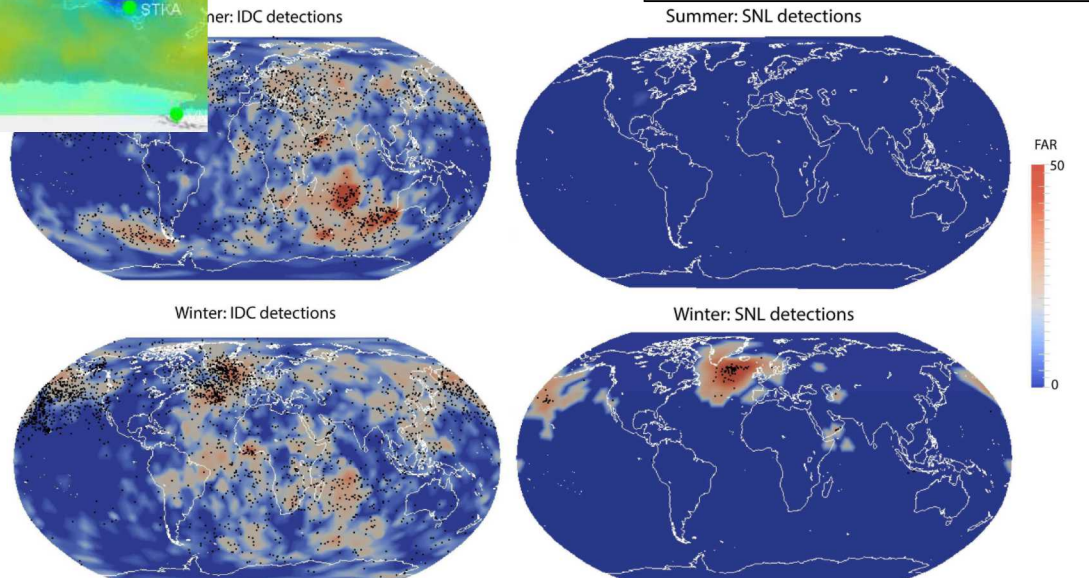
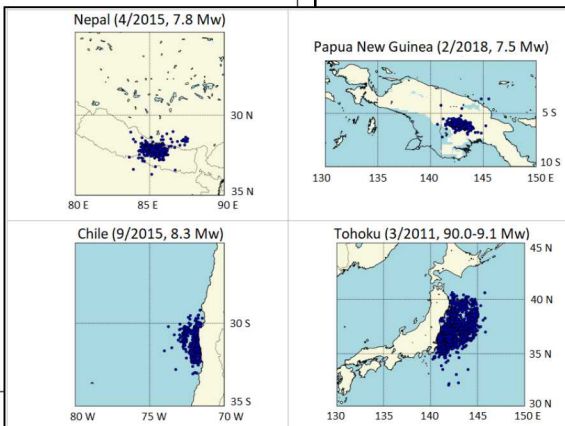
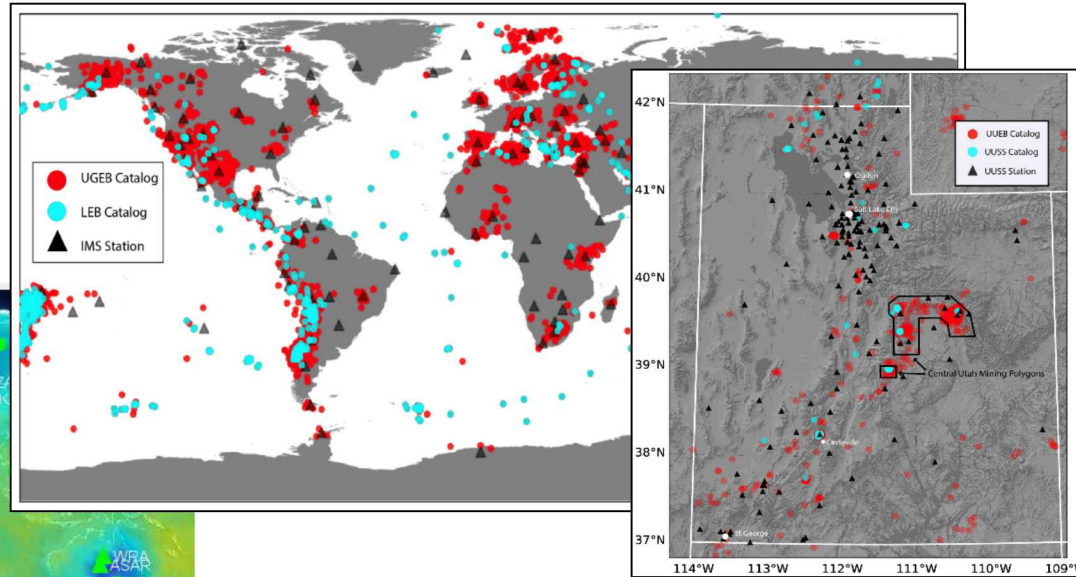
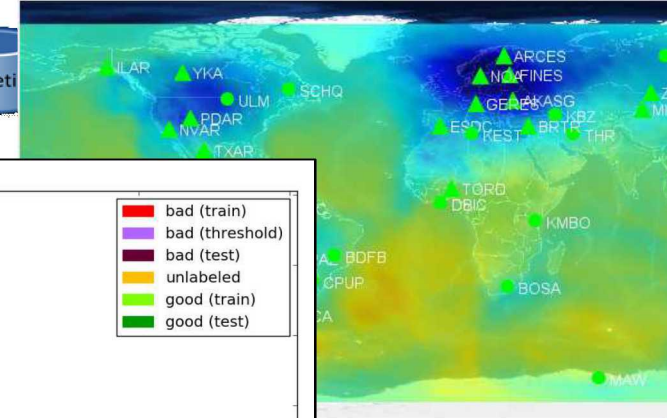
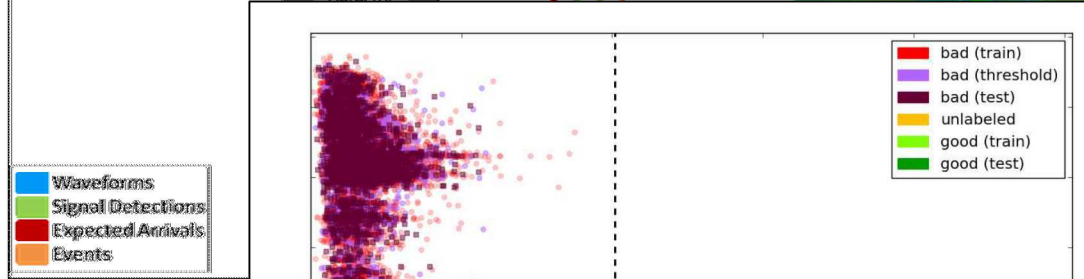
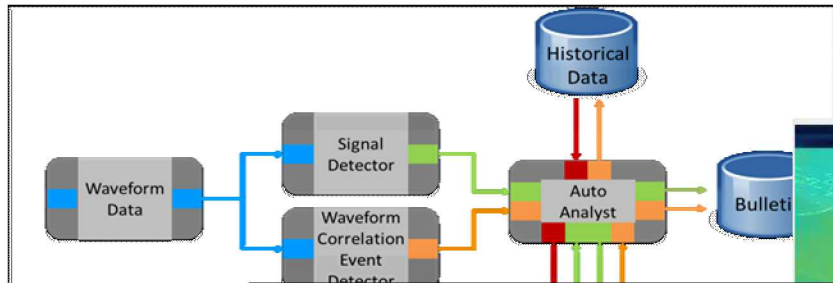
Text Parsing



Other Research & Development

SNL GNDD Signal Analysis Schubert Review

February 19, 2019 “very impressed with the technical work”
“great job communicating the message”

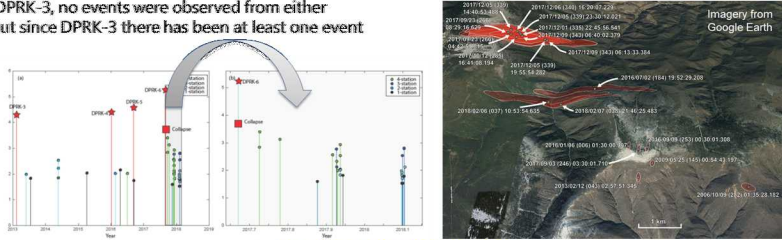


March 05, 2019

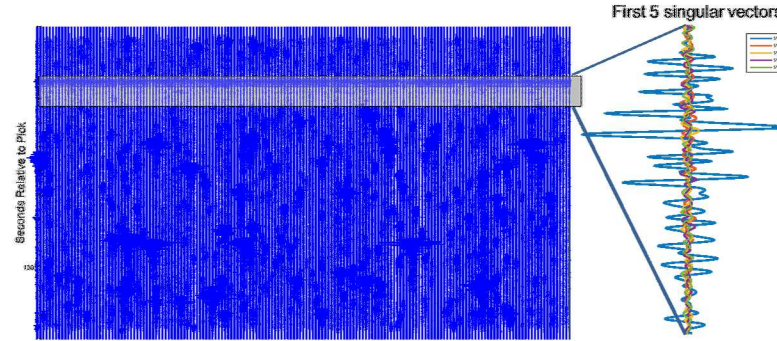
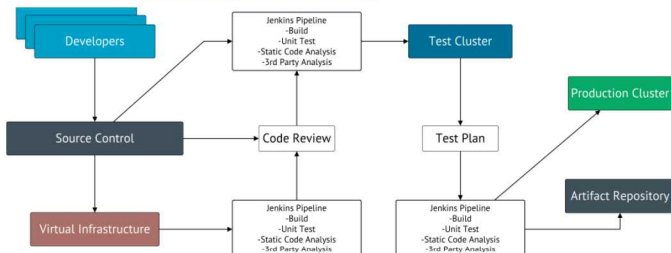
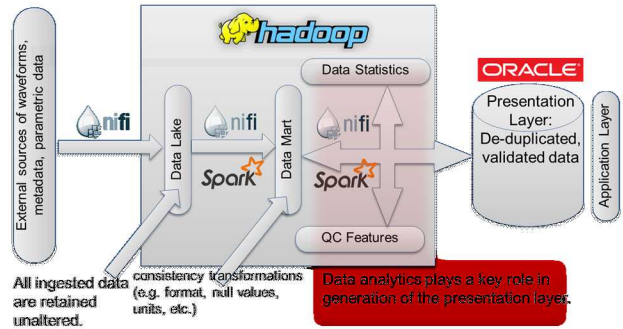
Searching for induced seismicity at Punggye-ri Nuclear Test Site using subspace detectors

Signals from post DPRK-6 earthquakes recorded at 4 stations were used to build subspace detectors which were then used to scan the available data for previously unobserved events near the test site. More than 30 additional template-matched events were detected. Thirteen of the detected events have been located. These events form two clusters, both to the northwest of the test site.

Prior to DPRK-3, no events were observed from either cluster, but since DPRK-3 there has been at least one event per year.



Seismological Research Letters ; 89 (6): 2094–2112. doi: <https://doi.org/10.1785/0220180127>



$$TBP \equiv \sqrt{\sigma_t^2 \sigma_\Omega^2}$$

$$t_0 = \frac{\int_{t_s}^{t_e} t [x(t)]^2 dt}{\int_{t_s}^{t_e} [x(t)]^2 dt}$$

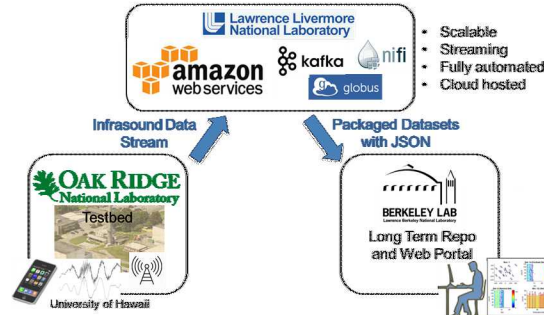
$$\sigma_t^2 = \frac{\int_{t_s}^{t_e} (t - t_0)^2 [x(t)]^2 dt}{\int_{t_s}^{t_e} [x(t)]^2 dt}$$

$$\Omega_0 = \frac{\int_0^{\Omega_{Ny}} \Omega [\hat{x}(\Omega)]^2 d\Omega}{\int_0^{\Omega_{Ny}} [\hat{x}(\Omega)]^2 d\Omega}$$

$$\sigma_\Omega^2 = \frac{\int_0^{\Omega_{Ny}} (\Omega - \Omega_0)^2 [\hat{x}(\Omega)]^2 d\Omega}{\int_0^{\Omega_{Ny}} [\hat{x}(\Omega)]^2 d\Omega}$$

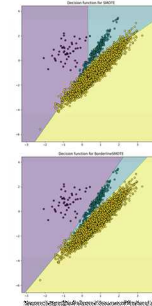
In these equations, $\hat{x}(\Omega)$ is the Fourier transform of $x(t)$ and Ω_{Ny} is the Nyquist frequency.

Cloud Deployment

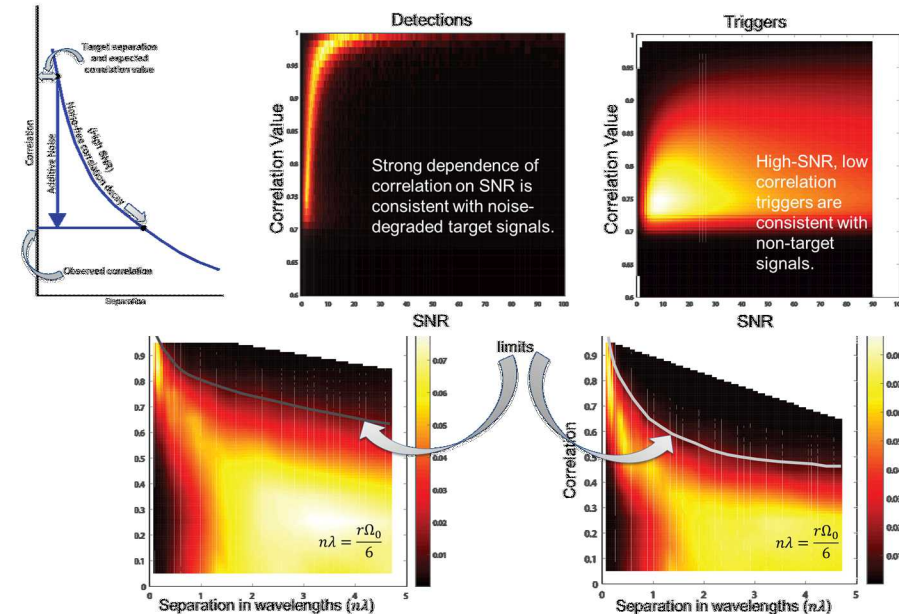


Handling Class Imbalance

- Labeling data is a time consuming, tedious, error-prone task that requires SME time.
- We started out with 659,407 artifact instances and 43,891 valid signal instances in the dataset before performing a train-test split.
- We used Borderline Synthetic Minority Over-sampling Technique (BorderlineSMOTE) (Han et al., 2005) to create a balanced dataset.
 - Generated synthetic data points near the decision boundary.
 - Increases true-positive rate
 - Better than random over-sampling



SNR-dependence of correlation value as a metric



- Blom, P., G. Euler, O. Marcillo, and F. Dannemann-Dugick, 2019, **Evaluation of a pair-based, joint-likelihood association approach for infrasound analysis**. *Geophysical Journal International*, in review.
- Carmichael, J. D., 2019, **Narrowband signals recorded near a moulin that are not moulin tremor: A cautionary short note**, *Annals of Glaciology*, Under Review. Manuscript ID AOG-79-0146.
- Carmichael, J. D., and R. J. Nemzek, 2019, **Fused multi-physics waveforms improve event detection: Predictive capability and performance**, *Geophysical Journal International*, in review.
- Dodge, D.A. 2018, Searching for induced seismicity at Punggye-ri Nuclear Test Site using subspace detectors, *Seism. Res. Lett.*, doi: 10.1785/0220180127
- Linville, L., Brogan C., Young, C., Aur, A., 2019, **Global and local scale high-resolution event catalogs for algorithm testing**. *Seismological Research Letters*, in review.
- MacCarthy, J., O. Marcillo, and C. Trabant, 2019, **Putting the cloud to work for seismology**, EOS, in press.
- Marcillo, O. E., and J. Carmichael, 2018, **The detection of wind-turbine noise in seismic records**, *Seismological Research Letters*, doi:10.1785/0220170271.
- Morton, E. A., S. M. Bilek, and C. A. Rowe, 2018, **Newly detected earthquakes in the Cascadia subduction zone linked to seamount subduction and deformed upper plate**, *Geology*, doi:10.1130/G45354.1.
- Park, J., I.-Y. Che, B. Stump, C. Hayward, F. Dannemann, S. Jeong, K. Kwong, S. McComas, H. R. Oldham, M. M. Scales, and V. Wright, 2018, **Characteristics of infrasound signals from North Korean underground nuclear explosions on 2016 January 6 and September 9**, *Geophysical Journal International*, 214, no. 3, 1865-1885, doi:10.1093/gji/ggy252.
- Sundermier, A., et al, 2018, **Alternate Null Hypothesis Correlation: A New Approach to Automatic Seismic Event Detection**. *Bulletin of the Seism. Soc. Of America*, Vol 108, No. 6, pp. 3528-3547, December 2018, doi:10.1785/0120180074
- Tibi, R., Ballard, S., et all, 2019, **Detection improvement of an iterative processing framework**. *Geophysical Journal International*, in progress.
- Walter, W. R., D. A. Dodge, G. Ichinose, S. C. Myers, M. E. Pasyanos and S. R. Ford, 2018, Body-wave methods of distinguishing between explosions, collapses and earthquakes – application to recent events in North Korea, *Seism. Res. Lett.*, doi: 10.1785/0220180128