Extending Waveform Correlation Techniques to Broad Regional Monitoring: Processing 3 Years of Data at 2 Stations

Megan Slinkard[1*], Stephen Heck[1], Dorthe Carr[1], Christopher Young[1] , Paul Richards[2] , Natalya Mikhailova[3] | *Sandia National Laboratories*[1] , *Lamont-Doherty Earth Observatory*[2] , *KNDC*[3]

SAND2013-2322C

National Nuclear Security Administration

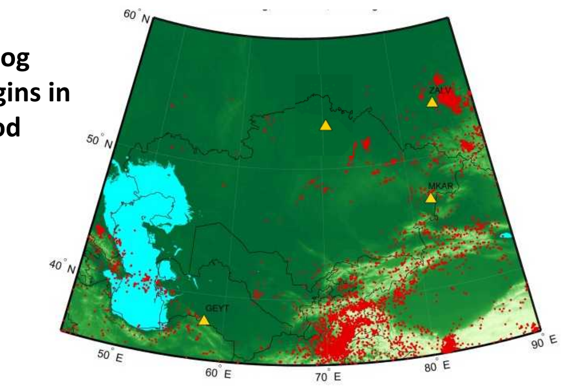
INTRODUCTION

Waveform correlation techniques have garnered increasing attention in the last few years, as their value in detecting and classifying repeated events has been demonstrated again and again. In this research, we show the potential in extending waveform correlation techniques to broad regional monitoring for the benefit of nuclear monitoring. We use data from the CTBTO's International Monitoring System, which consists of a sparse network of stations that must monitor the entire globe.

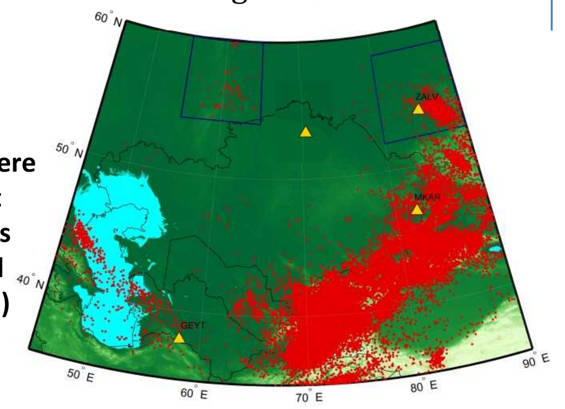
MOTIVATION AND OBJECTIVES

Comparing the CTBTO's LEB catalog to a regional catalog from Kazakhstan which covers central Asia, we note the potential for waveform correlation to enhance the completeness of the LEB catalog.

The LEB catalog had 8015 origins in a 3 year period

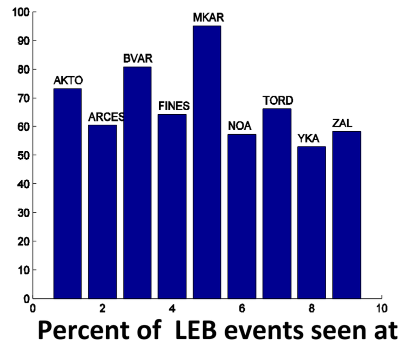


The KZ regional catalog had over 45000, AFTER mining events were removed (except for mining events in the two boxed regions in Russia)



STATION SELECTION

For this phase of the project, we focused on comparing the results of waveform correlation at 2 stations. To determine which LMS network stations to use for our study, we calculated for each station the percent of events in the LEB catalog (in 2008, in our region) which had Lg arrivals. MKAR and BVAR had the highest number of observations (figure 2) . We perform correlation on each element of MKAR's 9 element array, and used 9 elements of BVAR.



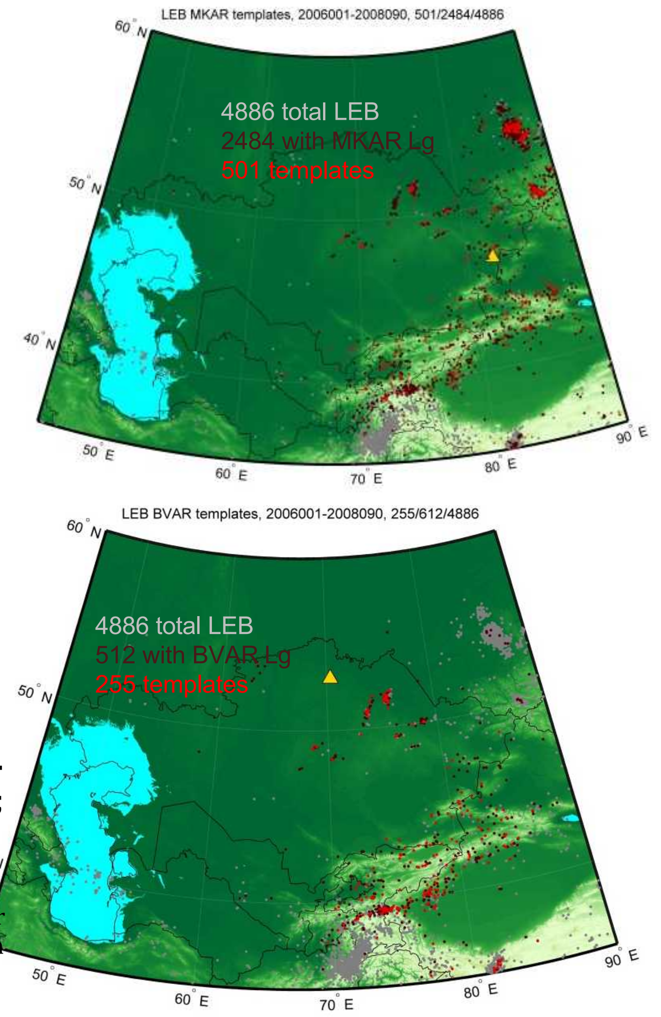
Percent of LEB events seen at several IMS stations

TEMPLATE SELECTION

Template selection is a critical aspect of a well functioning waveform correlation system. The first question to ask is which phase of an arrival to use as the template. In previously presented work we demonstrated that for this region of the world Lg arrivals found the most quality matches, so we made our templates from the Lg arrivals.

Templates should be clean, clear waveforms, without noise artifacts which will trigger detections on noise segments. We screen templates in a two step process; first we require a strong arrival with STA/LTA > 3; second we screen the histograms of correlation values generated using each template. Time-reversed templates are used to obtain the characteristics of the template in noise (the forward and reversed templates produce identical histograms in noise); in this manner we can select a suitable correlation threshold for each template. Lastly, we cluster the templates and keep only 1 representative template from each cluster.

Query LEB origin and arrival tables for events that meet time, station, and lat/lon requirements → Use STA/LTA to verify observable arrival at station → Verify histogram of correlation values meets histogram criteria (screens noisy templates) → Select correlation threshold (template-specific)



512 MKAR templates

255 BVAR templates

82 used by both

121 orids or equivalents in both

Template libraries were developed independently at each station. Only 82 templates were created from the same orid at each station; an additional 39 template equivalents* were identified.

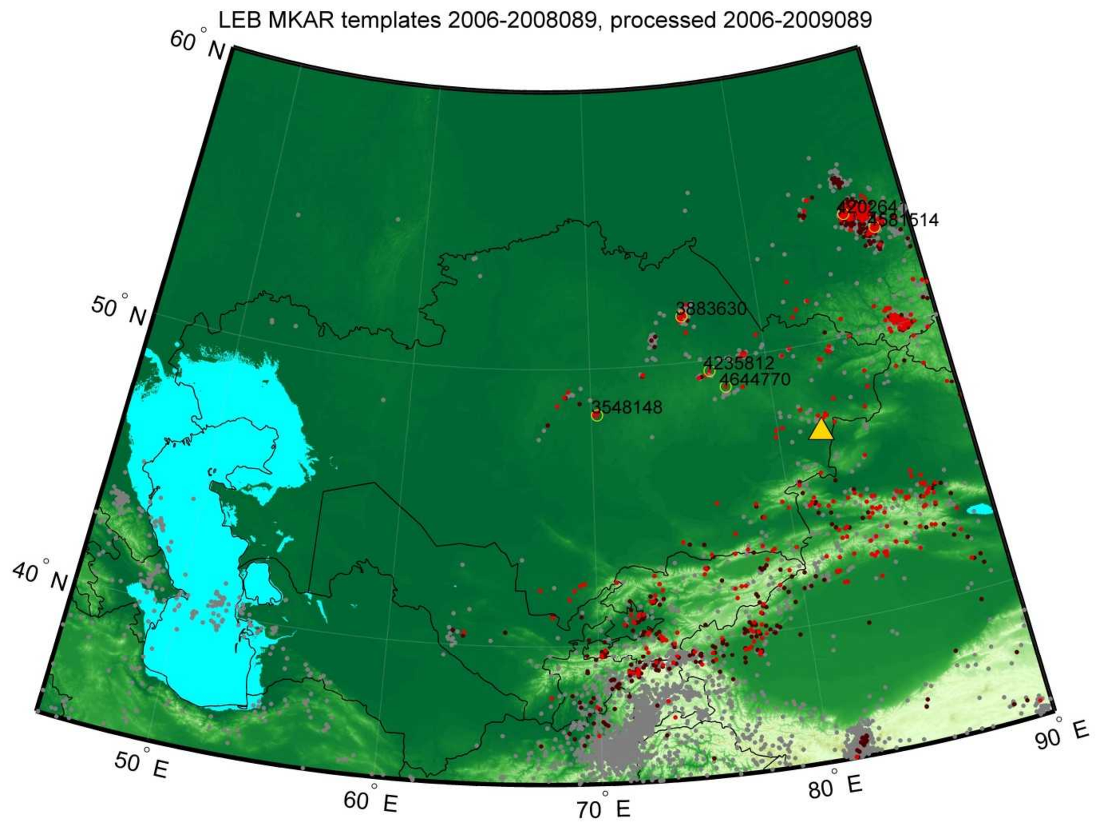
*When orid X is a template at MKAR, but orid X's waveform is a detection of orid Y's template at BVAR, we say X and Y represent that same family of events; this can happen, for example, if X and Y are similar events, and X was noisy at BVAR and therefore not included in the template library).

RESULTS

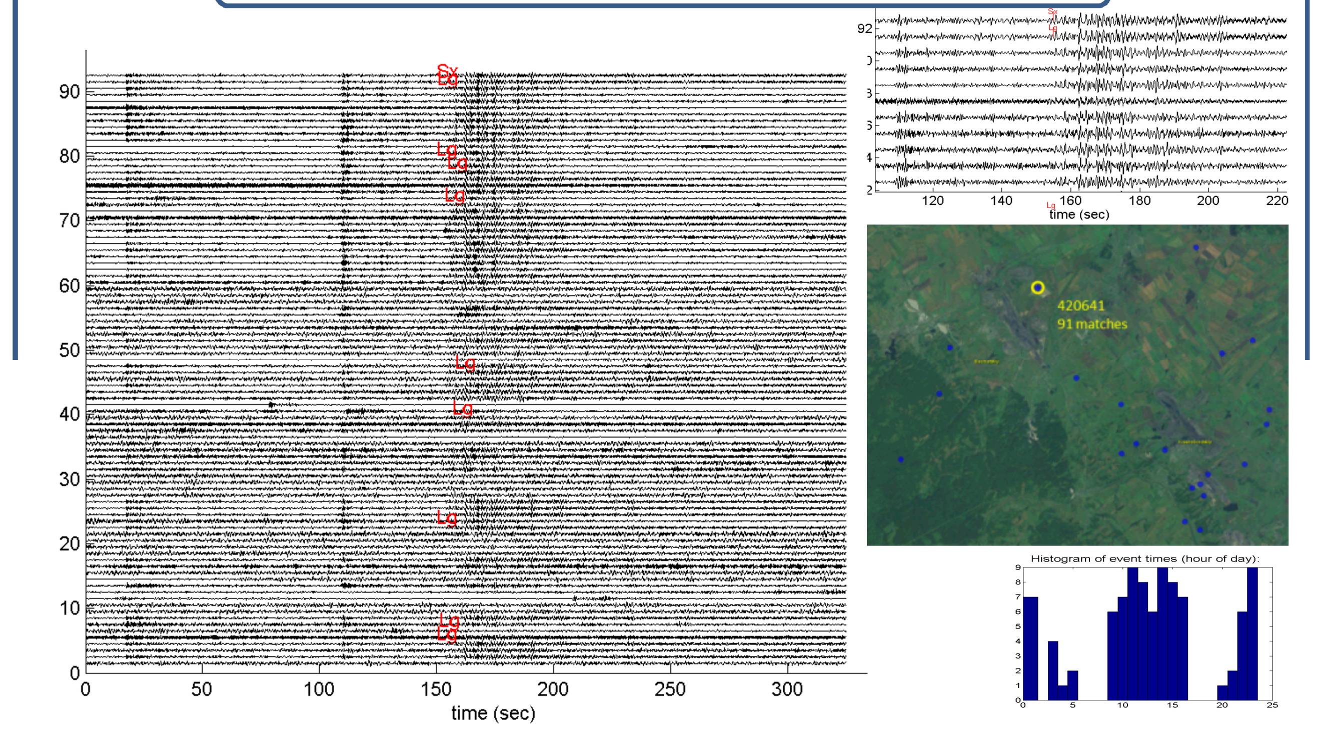
Dates WCD Ran (Templates formed: 2006/1/1 – 2008/3/30; Raw data processed: 2006/1/1 – 2009/3/30)	1/1/2006 – 4/1/2009 (3.25 years) The first 2.25 years overlapped with the period used to make the template library; the last year was processed to study the valued of using archival data for templates.
Stations used	MKAR, BVAR
Array elements used	9
Templates: acquisition dates	1/1/2006 - 3/30/2008 **2.25 years** (Included swarm in March 2008)
Templates: lat/lon box	lat : 35- 60; lon : 45- 90

Typical families of similar events are plotted, showing waveform plots and a histogram of the time of day at which the events occurred (to help distinguish mining families from earthquake families). In the waveform plots the top waveform is the template; below it are detected events, sorted by correlation value. The first detection is always the template finding itself; this serves as a nice sanity check, and is not counted in our detection statistics.

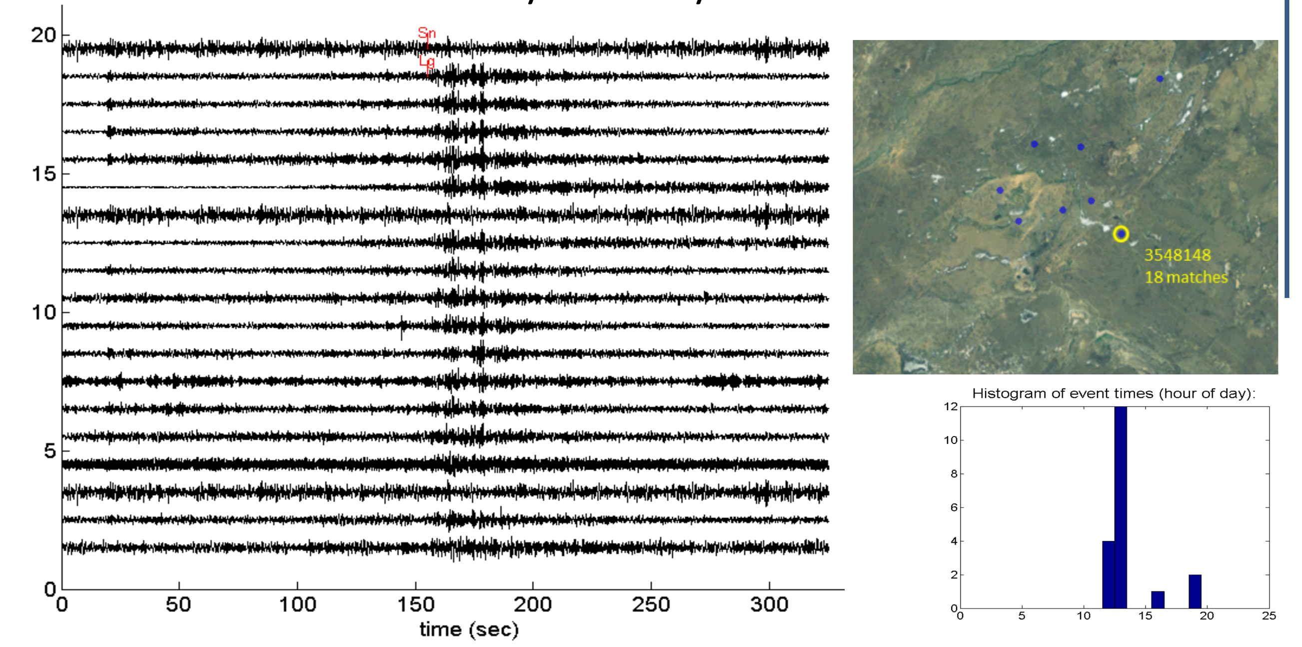
Geographic Distribution of Waveform families



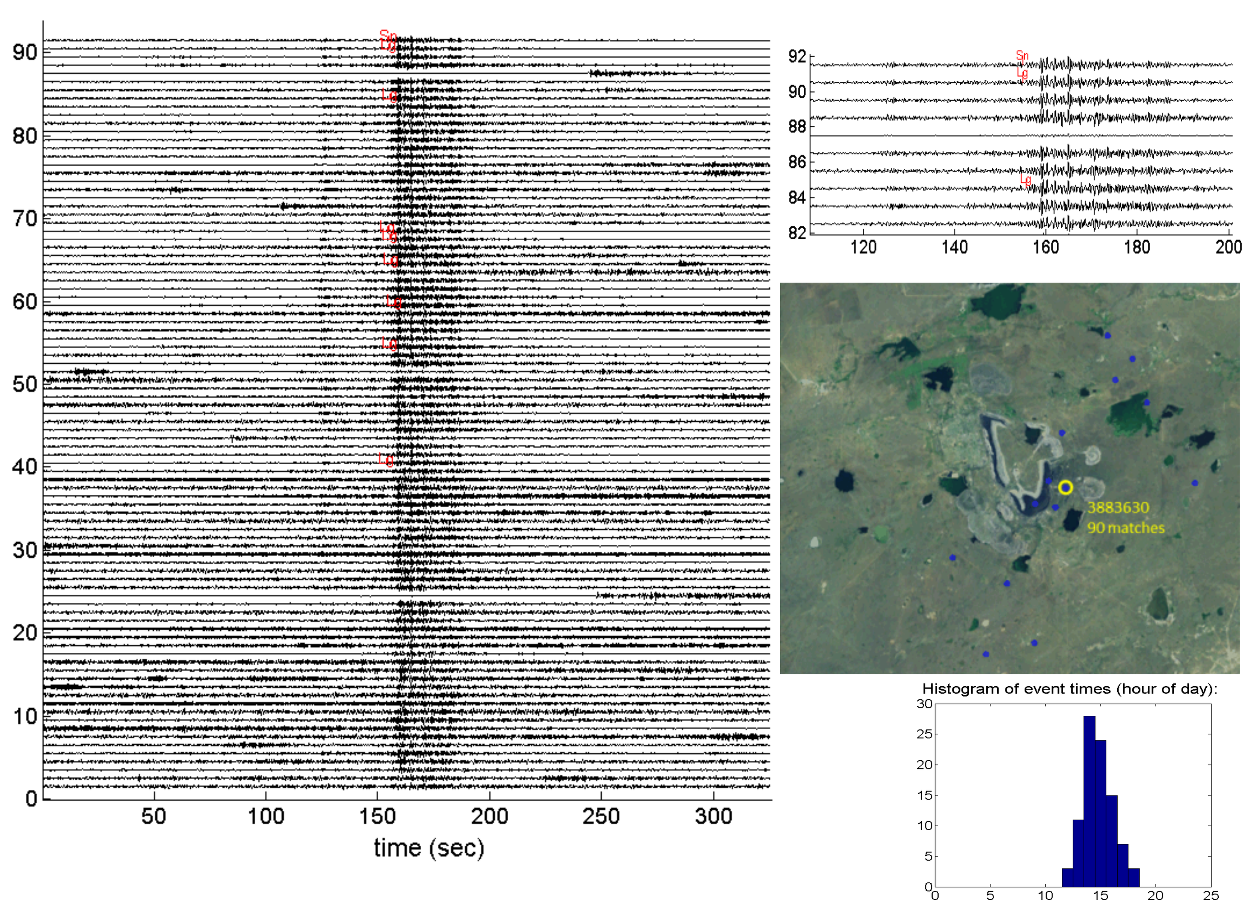
Family found by Orid 4202641



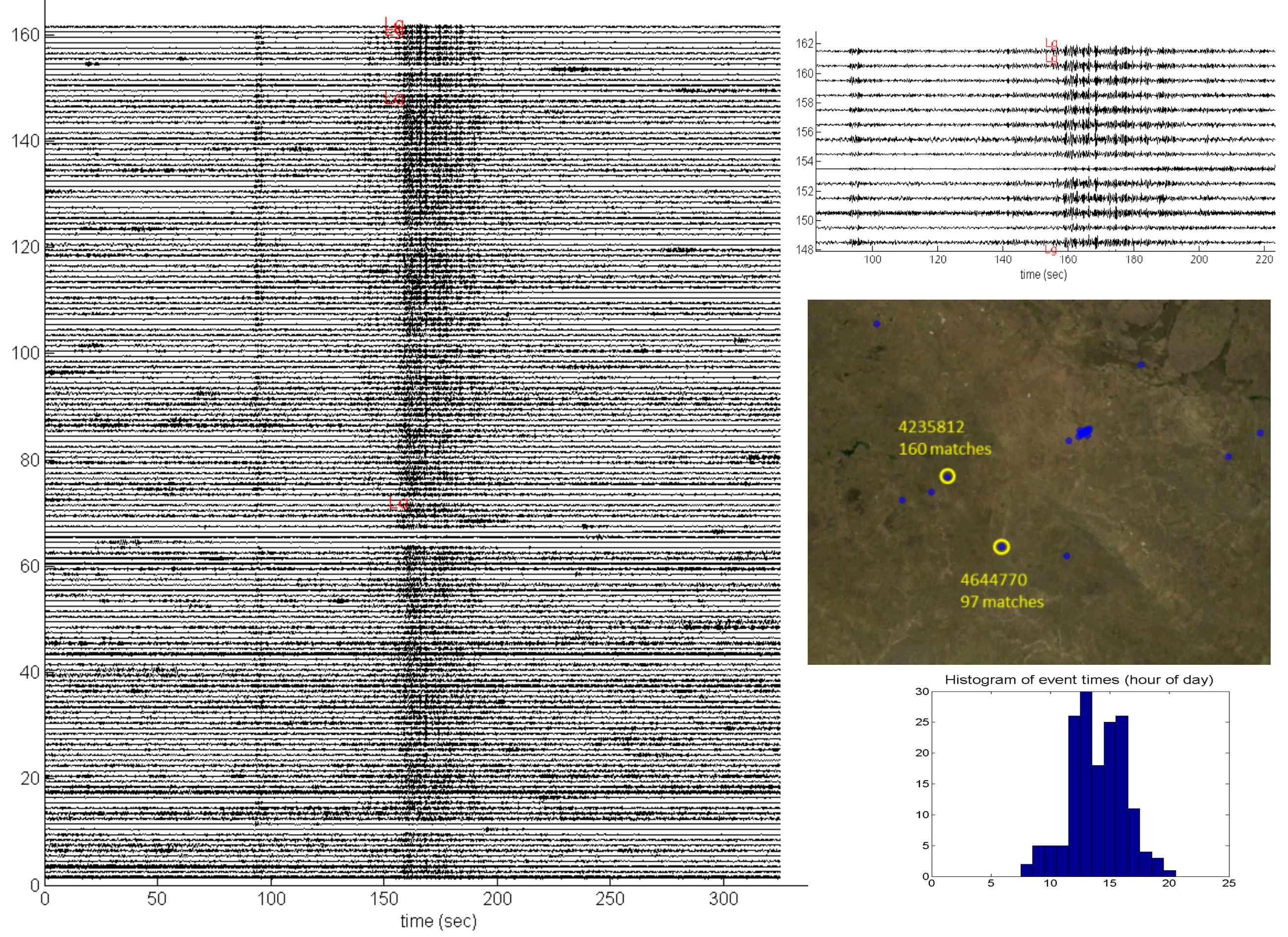
Family found by 3548148



Family found by Orid 3883630



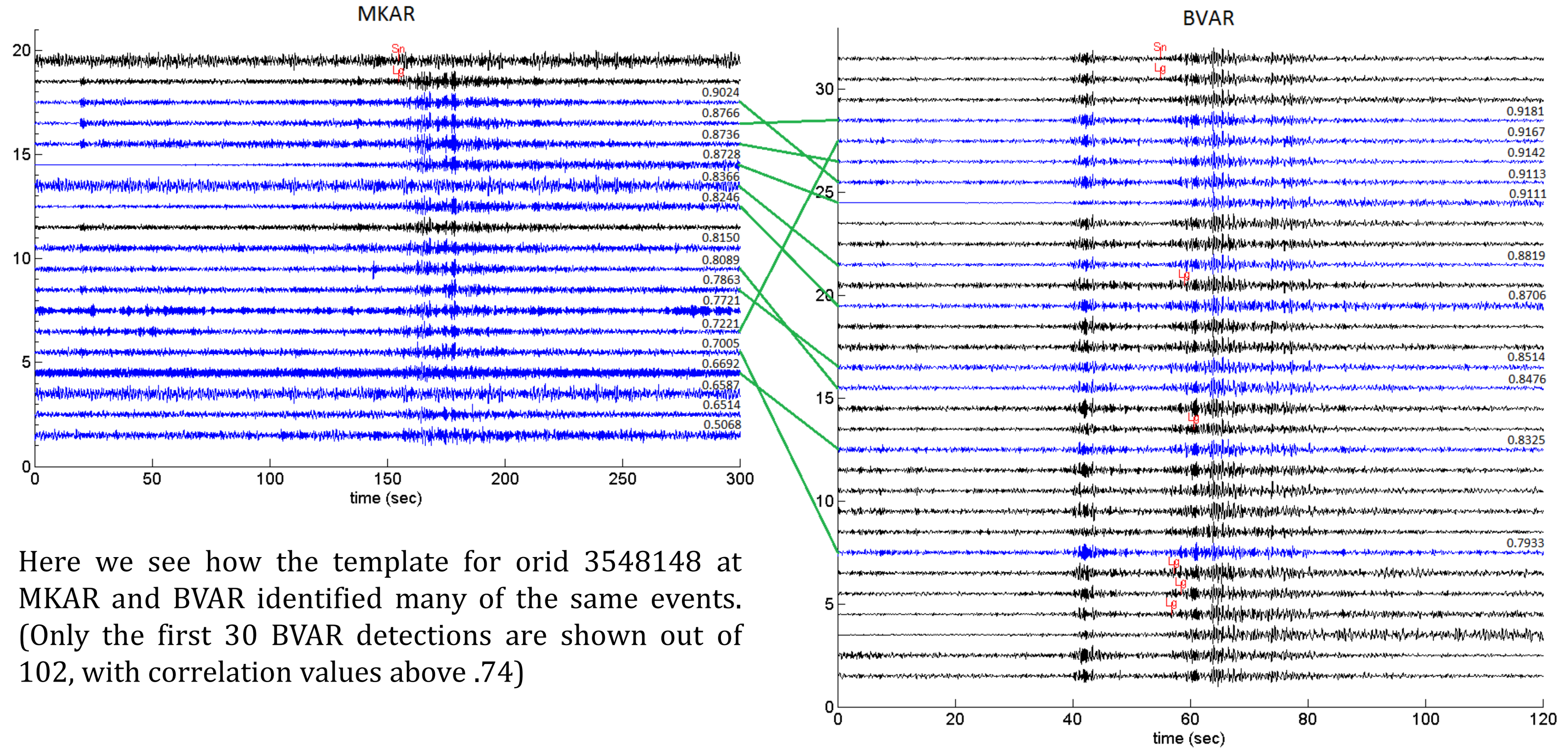
Family found by Orid 4235812



RESULTS SUMMARY

Detections can be confirmed by 1) seeing if there is an origin in a catalog at the calculated origin time, or 2) seeing if another station also made a detection at the same time. More debatable is to count as confirmed events with extremely high correlation values (> .7), even if they are only seen at one station.

	Number of Detections	Detections confirmed in the LEB	Detections confirmed in the KZ Catalog*	Detections confirmed in either catalog	Detections confirmed by 2 station detection	Confirmed by high correlation	Confirmed in some manner
MKAR	4863 (from 354 templates)	308 (6.3%)	450 (9.2%)	639 (13.1%)	700 (14.4%) (563 not in LEB)	1153 (23.7%) (774 not in LEB or 2 station)	1763 (36%)
BVAR	2706 (from 212 templates)	241 (8.9%)			700 (26%) (563 not in LEB)	438 (16%) (90 not in LEB or 2 station)	1009 (37%)



Here we see how the template for orid 3548148 at MKAR and BVAR identified many of the same events. (Only the first 30 BVAR detections are shown out of 102, with correlation values above .74)

SUMMARY

We performed waveform correlation on 3 years of data for stations MKAR and BVAR in central Asia. Our computational abilities allowed us to perform this analysis in 2.5 days. By using two stations to validate waveform correlation detections, we were able to approximately double the number of confirmed detections compared to the LEB; this illustrates the value of waveform correlation in improving catalog completeness.

FUTURE WORK

Expand template library for MKAR
- Use more years of LEB data to develop template library.
- Could develop a library based on the KZ catalog events.

Expand to additional IMS stations
- Lower detection threshold – using more stations should mean a more even detection threshold across the region.
- More robust events – detecting events at more than 1 station is an additional method to corroborate events, especially events with a low SNR.

Expand monitoring region
- Our ultimate goal is global monitoring.

Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

*** meslink@sandia.gov** **// Ground-Based Nuclear Explosion Monitoring Research and Development**