T3.7-P6

Automatic identification of repeated industrial seismicity in the Reviewed Event Bulletin

Dmitry Bobrov[2], Ivan Kitov[1], Mikhail Rozhkov[1], Matthew Yedlin[2]

[1]Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization, [2]University of British Columbia

Abstract

The Reviewed Event Bulletin (REB) produced by the International data centre (IDC) of the Comprehensive nuclear-test-ban Treaty Organization (CTBTO) includes thousands of industrial explosions detected by the primary seismic network of the International monitoring system every year. According to the CTBTO's monitoring mandate, these events represent background noise, which demands human and computer resources for interactive and automatic processing. It can also increase the overall detection threshold. Most of these blasts are repeated events generating similar signals. Using an extended set of signals from the Aitik and Kiruna mines in Sweden measured at the closest IMS array stations ARCES, FINES, NOA, and HFS we develop a method of automatic event formation and identification based on waveform cross correlation (WCC). Real-time automatic processing includes signal detection and characterization, local association of the detected arrivals with seismic events, relative location and magnitude estimation. A prototype pipeline is currently tested at the IDC. We present select results of detection, relative location and mine identification obtained since January 1, 2017. Worldwide, there are tens of mines with hundreds of events per year, which are present in the REB. Overall, the developed method can reduce the overall IDC analysts' workload by several percentage points saving time and resources for comprehensive monitoring. Our method can be applied to smaller blasts detected only by local and near-regional seismic networks.

Conclusion

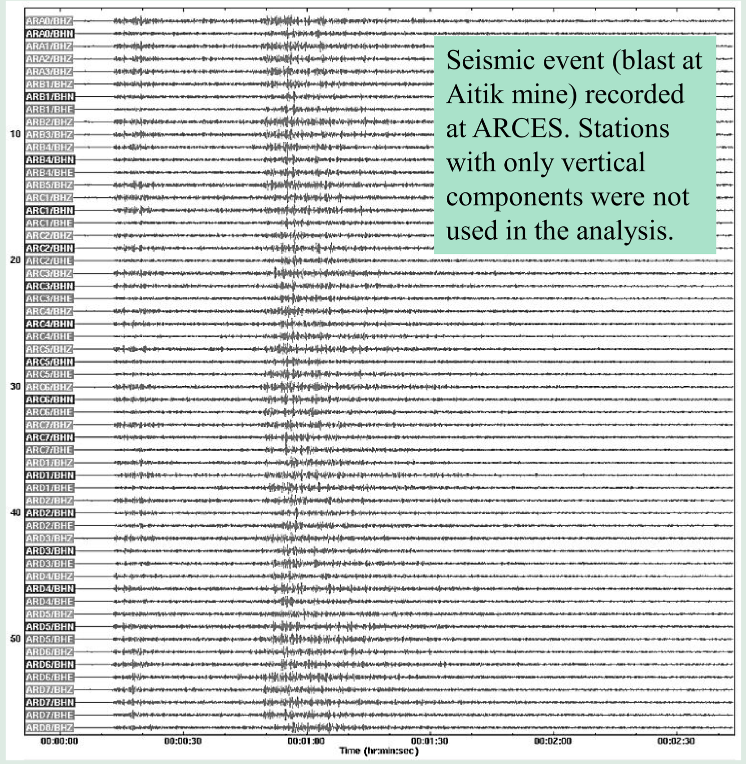
Continuous detection of signals using cross waveform correlation with several master events is likely the most reliable method of detection and creation of seismic events. It allows to actually distinguish between blasts conducted at the Aitik and Kiruna. The results of detection, relative location and mine identification obtained since January 1, 2017 demonstrate that our method confirms the existing REB events and finds many additional REB-compatible events. Overall, the obtained results validate detection and association procedures related to the WCC method.

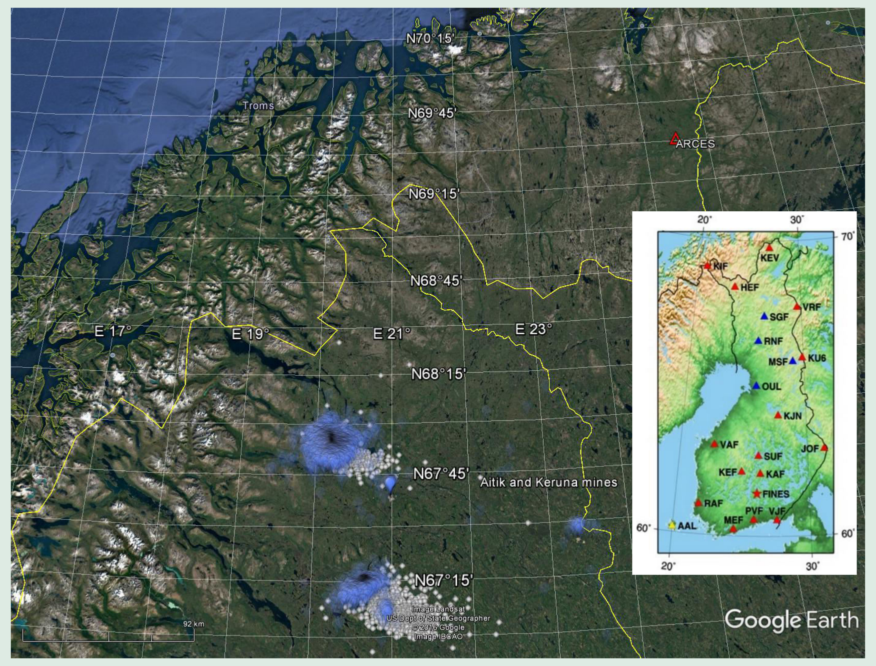
We have tested a design of multidimensional (high order, HO) master event template for further seismic event location based on cross correlation. The HO template design has to represent a multidimensional seismic 3-component array multidimensional array in a computer semantic sense and to apply the dimensionality reduction technique. The array seismogram was considered as a 3-order tensor so the training data set turned to 4-order tensor. The following methods were tested: (1) high-order SVD (HOSVD), (2) multidimensional discrete cosine transform (MD DCT), (3) two dimensional SVD (2D SVD), and (4) tensor interpolation.

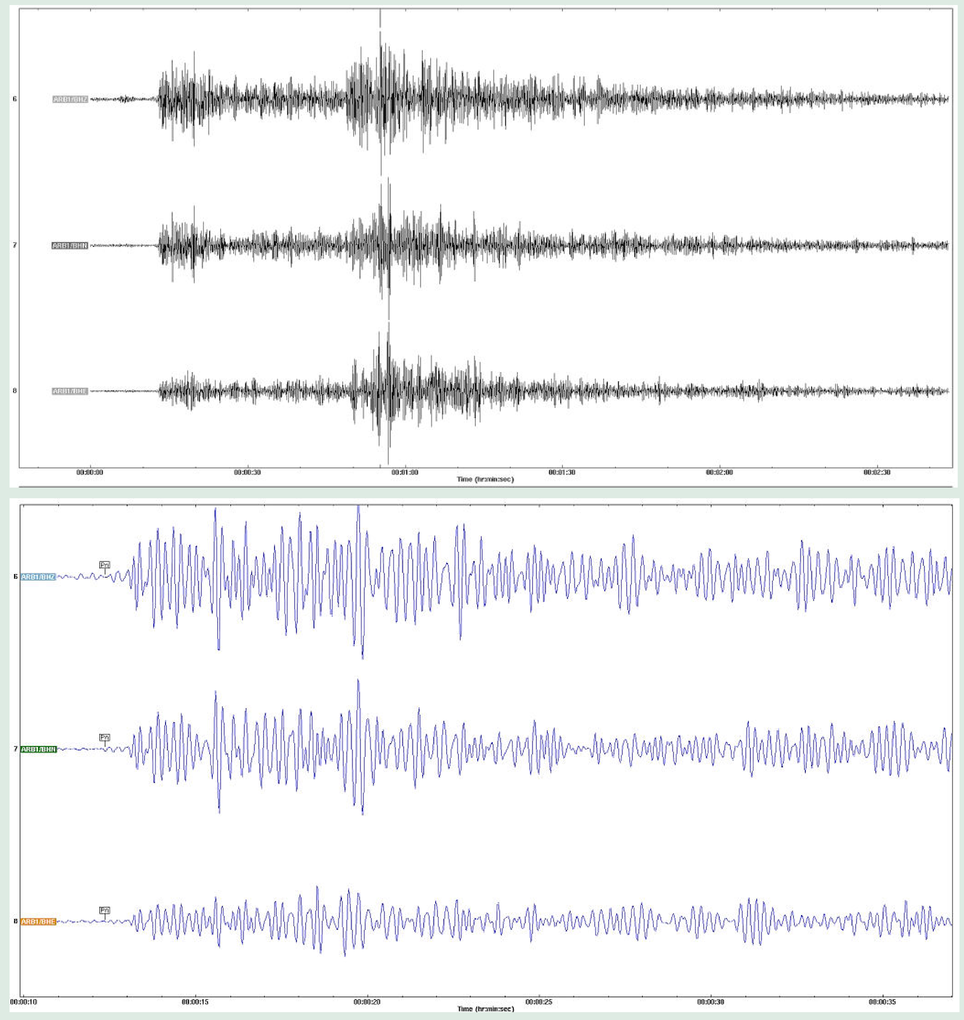
1. Aitik and Kiruna mines seismicity

We have conducted a study on template selection for further location of seismic events by the waveform cross-correlation method using regional quarry blast data recorded at 3-component IMS seismic array ARCES. One of reasons for this study was the difference in location conducted by the IDC and presented in the REB, and the location results presented by the International Seismological Center, ISC, based on data from the seismographic network of the University of Helsinki (UH, inlay on a Google Earth map below). The blast clusters by the IDC are white balls for both mines; the UH results are blue clouds.

In total, 122 seismic events from the Aitik (copper) and Kiruna (iron) mines were studied at 19 3-component stations of ARCES.







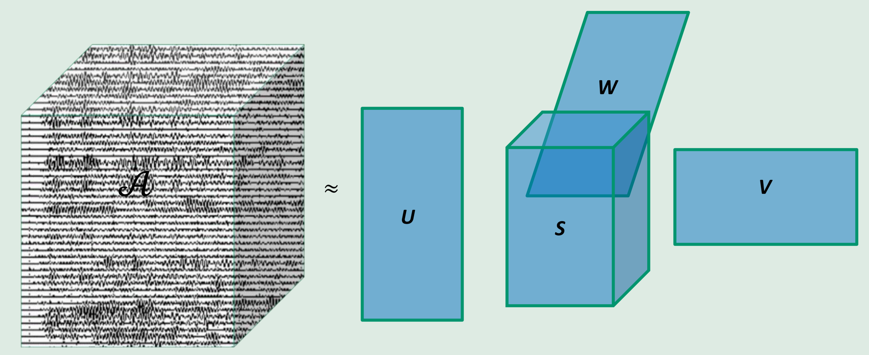
3-component records at a single ARCES station. Bandpass filter is 3-6 Hz. 160 (upper figure) and 27 (lower) second time windows presented.

2. Tensor Approach to Seismic Array Data Processing

Introduction. Here, we introduce an approach to construct master event templates for further cross-correlation-based location with data recorded at multichannel seismic installations, such as 3-C seismic arrays of the International Monitoring System (IMS) of the CTBTO. Dealing with the tensor representation of seismic wavefields can simplify in certain sense the multidimensional approach to data processing, in particular, to the data set which was the same as used for the single component processing. Multichannel data corresponding to a seismic event from the Aitik and Kiruna quarries in Sweden can be rearranged as a 3-mode tensor, where first mode is time, or sample number, the second mode is station, or sensor number, and the third mode is the direction of ground motion (Z, N and E). Then, a complete test data set would consist of a 4-mode tensor with the event number corresponding to the 4th dimension. Considering a 3-component seismic array as a multitude of observations with a tensor description (not the tensor field in general sense), the corresponding data tensor, formally, can be regarded as a tensor product of 3 vector spaces, each with its own coordinate system. Then we could apply tensor operations to the data recorded by such arrays gaining certain benefits from utilizing joint volumetric (sensor) and spatial (array) information. Further dimensionality reduction of tensor data produces a basis for the multidimensional waveform templates. Note that a first-mode, or first-order tensor is a vector, a second-order tensor is a matrix, and tensors of higher orders are higher-order tensors.

General approach. Traditional approaches to finding lower dimensional representations of tensor data include flattening the data and applying matrix factorizations such as principal components analysis (PCA) or employing tensor decompositions such as the CANDECOMP/PARAFAC (canonical polyadic decomposition with parallel factor analysis) and Tucker decompositions, which may be regarded as a more flexible PARAFAC model. Tucker decomposition, which we use in this work, decomposes a tensor into a set of matrices and one core tensor. Then the eigenimages can be extracted for resizing the input tensor to lower dimensions. There are more approaches to the multimodal dimensionality reduction we explored in this study, such as the Multidimensional Discrete Fourier Transform (DCT) mostly used in image processing (JPEG, for instance), 2D SVD (based on low rank approximation of the matrix), and tensor interpolation (for example, Hotz, et al, 2010, Tensor Field Reconstruction Based on Eigenvector and Eigenvalue Interpolation). With this, we make an accent on the Tucker tensor decomposition made with the alternating least squares (ALS) method.

Math formalism. Following standard multilinear algebra, any tensor can be expressed as the product $\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \dots \times_N U^{(N)}$, where $U^{(n)} = (U_1^{(n)} U_2^{(n)} \dots U_{I_n}^{(n)})$ is an orthogonal $(I_n \times I_n)$ matrix (H. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos (2006), Multilinear principal component analysis for tensor objects for classification). A visual representation of this decomposition in the third-order case is shown on next figure:



A matrix representation of this decomposition can be obtained by unfolding $\mathcal{A}$ and $\mathcal{S}$ as:

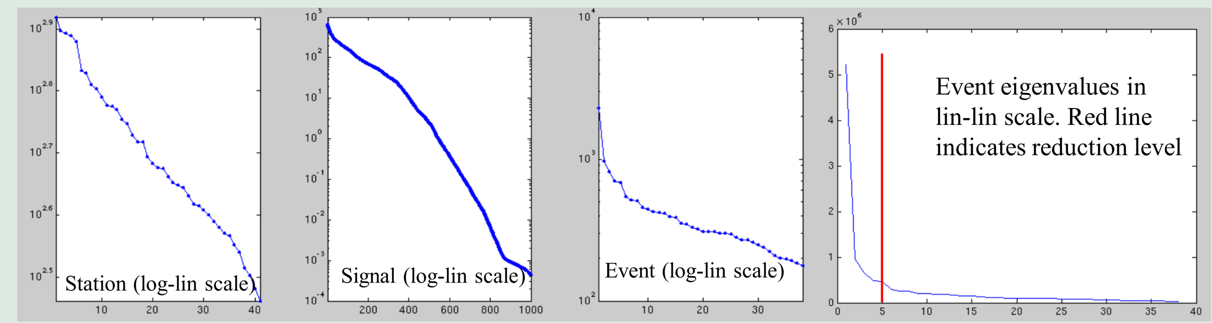
$$\mathbf{A}_{(n)} = \mathbf{U}^{(n)} \cdot \mathbf{S}_{(n)} \cdot (\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \dots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \dots \otimes \mathbf{U}^{(n-1)})^T$$

where $\otimes$ denotes the Kronecker product and $\mathcal{S}$ is a core tensor of size $R_1 \times R_2 \times \cdot \cdot \cdot \times R_N$. The decomposition can also be written as

$$\mathcal{A} = \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \dots \sum_{i_N=1}^{I_N} \mathcal{S}(i_1 i_2, \dots, i_N) \times \mathbf{u}_{i_1}^{(1)} \circ \mathbf{u}_{i_2}^{(2)} \circ \dots \circ \mathbf{u}_{i_N}^{(N)}$$

i.e., any tensor can be written as a linear combination of $I_1 \times I_2 \times \dots \times I_N$ rank-1 tensors. This decomposition is used in the following to formulate a multilinear projection for dimensionality reduction.

Following basic principles of the 1D PCA, we approach to this projection with analyzing the core tensor $\mathcal{S}$. It's eigenvalues (for station array NORSAR) corresponding to the dimensions (station; signal; event) are represented on a figure below. Only 5 eigenvalues are meaningful, so dimensionality reduction can be performed to the decomposed tensors order of 5.



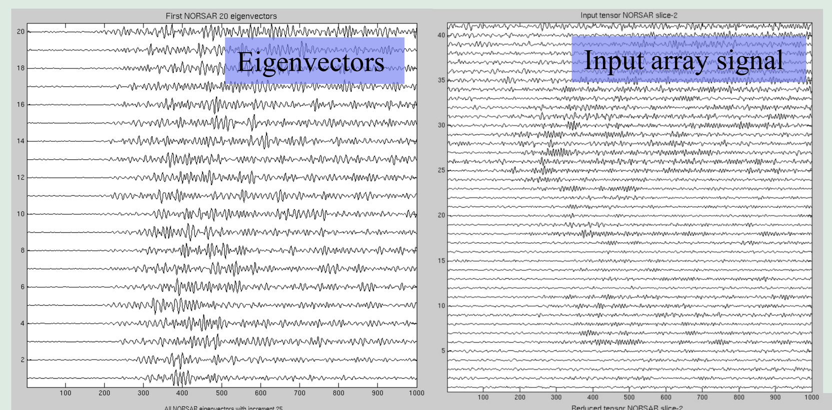
We perform a dimensionality reduction through the truncation of U and $\mathcal{S}$ terms and building the restored tensor $\mathcal{A}_{red}$ through the block term decomposition (BTD) which approximates a tensor by a sum of low multilinear rank terms.

$$\mathcal{A}_{red} = \sum_{r=1}^{R} \mathcal{S}^{(r)} \bullet_1 U^{(r,1)} \bullet_2 U^{(r,2)} \bullet_3 \dots \bullet_1 \boldsymbol{N} U^{(r,N)}$$

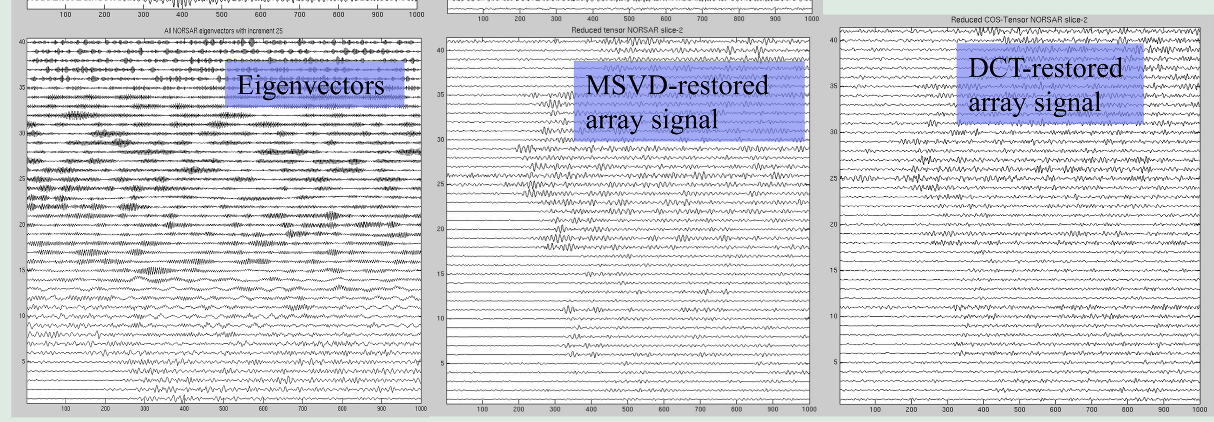
Multimodal dimensionality reduction with other methods. The multidimensional discrete (MD) cosine transform (DCT-II and DCT-III for inverse) is popular compression structures for MPEG-4, H.264, and HEVC (high efficiency video coding), and is accepted as the best suboptimal transformation since its performance is very close to that of the statistically optimal Karhunen-Loeve transform

$$X(k_1, k_2, \dots, k_r = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \dots \sum_{n_r=0}^{N_r-1} x(n_1, n_2, \dots, n_r) \cdot \cos \frac{\pi(2n_1+1)k_1}{2N_1} \dots \cos \frac{\pi(2n_r+1)k_r}{2N_r},$$

where $k_i = 0, 1, \dots, N_i$ -1 and $i = 1, 2, \dots, r$. Inverse truncated MD DCT returns the reduced tensor array with the required number of components used for cross correlation template construction.

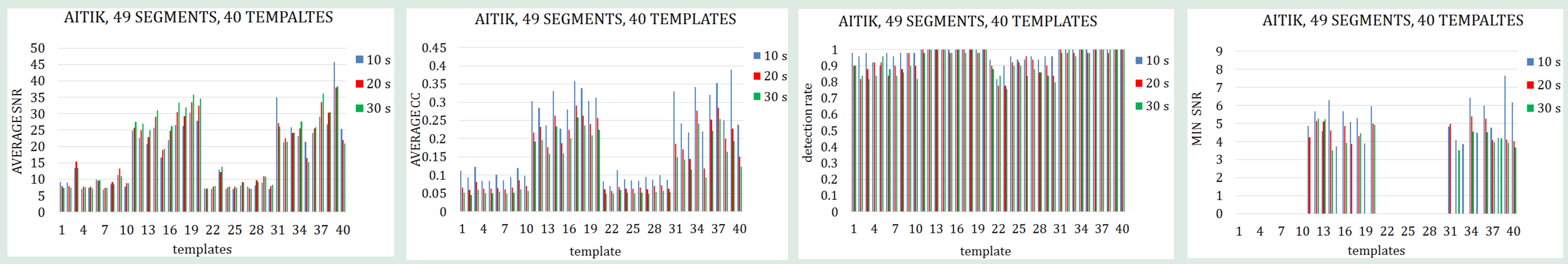


After back-projection, the tensor slices can be used as templates for further cross-correlation techniques. Left: vertical eigenvectors of NORSAR array. Center: second slice (out of 39 total number of signals) of input (up) and restored tensor (down) after projection back to signal space. Notice significant noise suppression after restoring. Down-right: reduction with DCT.

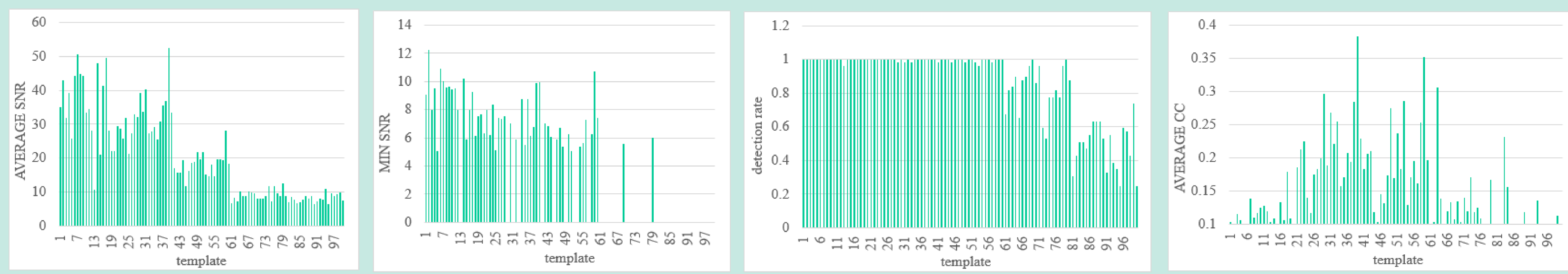


3. Tensor Templates Test

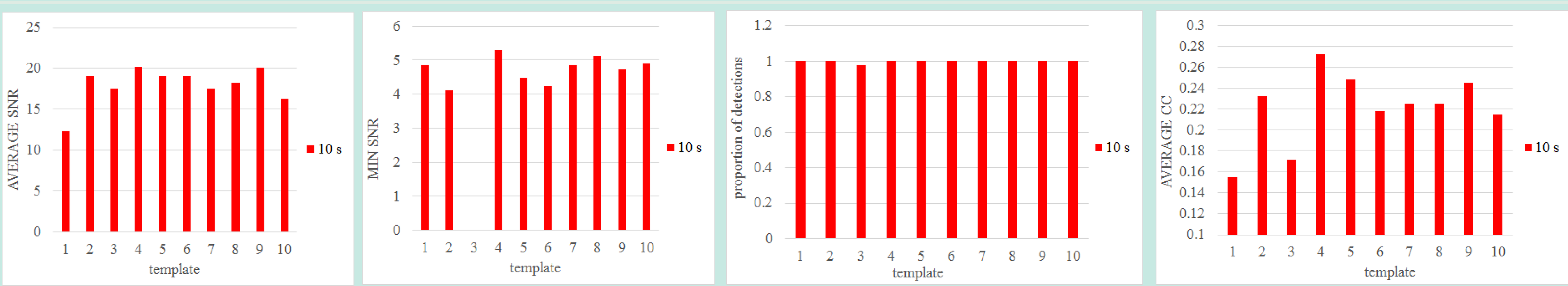
We have conducted a number of tests with the obtained reduced tensor components in order to evaluate performance of the cross correlation detector with these components used as templates. Complete sets of components for the 3C array were produced with 4 methods: tensor interpolation, DCT, 2DSVD and HOSVD. In this study, the 3D templates were reduced to 1D vectorized case ([Z,NS,WE]) and we applied a well established system of tests. Selected results are presented on figures below. The first test was based on cross correlation (CC) of the developed templates with continuous waveforms measured from the set of 122 events and determining the detection rate based on SNR threshold (see poster S51A-2758 for details, also Bobrov, et al. 2012): the percentage of detections having SNR_{CC}>3.5. Then, we tested 46 events not included into the training set of 122. In this test, the ten first reduced tensor components were tested for detection, i.e. 40 components were tested altogether as presented on figures below: (1) 2D SVD, (2) MD DCT, (3) HOSVD, and (4) tensor interpolation. Different template lengths were used from 10 to 30 seconds. The best results were produced by the MD DCT and interpolation templates for average SNR; the minimum SNR_{CC} over all tested signals for a given template, which has to be above 3.5; and the average CC. The overall difference in detection rates is not large. Similar tests were also carried out with the eigenimages (instead of reduced back-projections) produced for the HOSVD (PC cell array in TPCA algorithm). It was found that the algorithm destroys the proper channel alignment in sensor triads and move-outs related to different stations of the array and the test results were not impressive. In case of single 3-C station, it does not make any difference since all the channels in a training set are aligned by default and there no is need to keep the move-outs so the regular SVD/PCA case works fine.



Left figure: 4 methods, 10 templates in each method, variable time window length



Left figure: 2 methods, 10 templates in each method, 5 filter bands: 2-4 Hz, 4-8 Hz, 3-6 Hz, 6-12 Hz, 8-16 Hz.



Left figure: HOSVD tests non-standard truncated reconstruction of the reduced tensor set

After all methods were tested with different template lengths and filters, we have taken the window (10 seconds) and the filter (3-6 Hz), which work the best. Then we returned to the high order tensor decomposition and tried our own truncated reconstruction of the reduced tensor set. The detection rate was the same as for the best methods above, and the smallest SNR_{CC} was even larger (4.2 against 3.5). More work has to be conducted to find the optimal multidimensional template design.

4. Continuous detection, association and relative location

Association and relative location

For all valid arrivals, which are found with a given master event, origin times, OT*ij*, are calculated. The empirical travel times from the master event to the relevant primary stations, TT*ij*, are subtracted from the arrival times, AT*ij*.

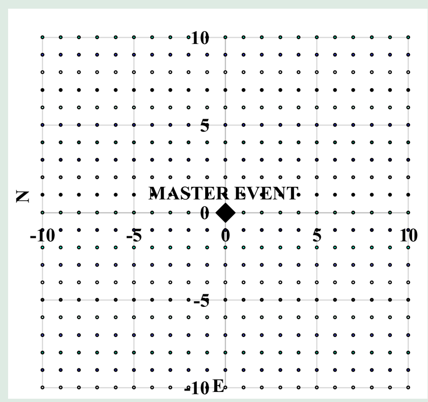
OT*ij* = AT*ij* – TT*ij*

where *i* is arrival index at station *j*.

TT*ij* = TT*j* **!**

Empirical travel times from a master event to seismic stations are characterized by **ZERO** modelling errors and very low measurement errors. These conditions allow extremely accurate *relative* location.

Relative location grid



$dt_k = \mathbf{S} \cdot \mathbf{d_k}$ - travel time correction

$OT^k_{ij} = AT_{ij} - TT_j + dt_{jk}$ – corrected origin time

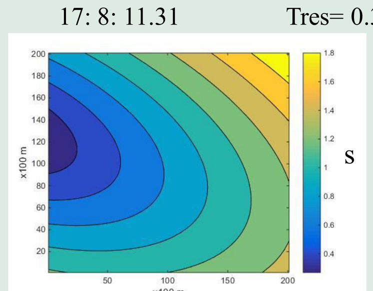
k nodes, rectangular of circle;
grid size from 1 to 100 km;
spacing from meters to 10-15 km
Average OT and RMS OT residual are calculated in each node

Ground truth events for Aitik mine

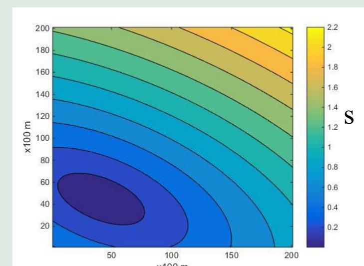
There are 97 GT events confirmed by local infrasound measurements, from which 33 (see table below) are missing in the REB. All were found by cross correlation at ARCES, FINES and NOA

Date	Time
2015055	18:04
2015023	17:55
2015013	18:27
2014343	18:08
2014336	18:03
2014313	17:46
2014289	17:02
2014268	15:48
2014258	17:06
2014245	17:01
2014241	17:02
2014238	17:05
2014217	17:08
2014206	17:00
2014183	17:01
2014177	17:04
2014146	17:02
2014135	17:01
2014120	17:01
2014105	17:05
2014100	17:04
2014098	17:06
2014086	18:08
2014071	18:03
2014063	18:01
2014045	15:51
2014016	18:02
2014014	18:19
2013339	18:00
2013332	18:02
2013329	18:08
2013304	18:01
2013290	17:06

DATE 2013290; ORIGIN TIME 17: 6: 23.20
FINES 17: 7: 52.85 Tres=-0.308
ARCES 17: 7: 10.66 Tres=-0.039
NOA 17: 8: 11.31 Tres= 0.348

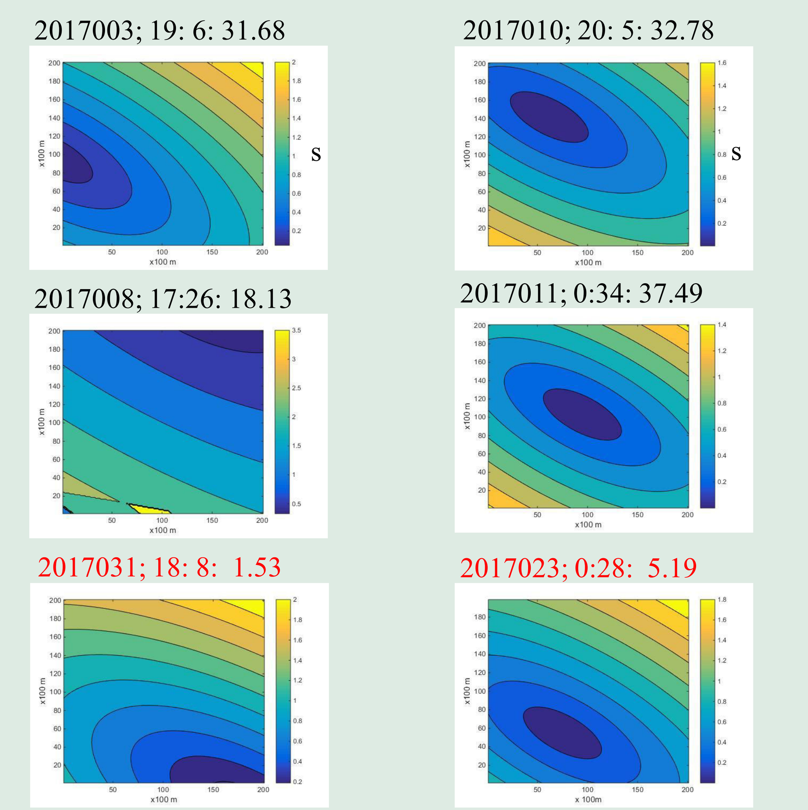


DATE 2015055; ORIGIN TIME 18: 4: 2.58
ARCES 18: 4: 50.04 Tres=-0.000
FINES 18: 5: 33.87 Tres= 0.003
NOA 18: 5: 51.01 Tres=-0.002



Continuous association and location of Aitik and Kiruna blasts

Continuous detection with cross correlation uses 14 master events: 7 from Aitik mine and 7 from Kiruna. The WCC method finds all REB (black date and time) events and many events not in the REB (red date and time)



Preparatory Commission for the Comprehensive Nuclear-Test Ban Treaty Organization,
Provisional Technical Secretariat, Vienna International Centre, P.O. Box 1200, A-1400 Vienna,
Austria. E-mail: mikhail.rozhkov@ctbto.org