

Blind Source Separation of Seismic Events with Independent Component Analysis: CTBT related exercise Mikhail Rozhkov¹, and Ivan Kitov² (1 - International Data Centre, CTBTO, 2 - Institute of Geosphere Dynamics, RAS)

Blind Source Separation (BSS) methods used in signal recovery applications are attractive for they use minimal a priori information about the signals they are dealing with. Homomorphic deconvolution and cepstrum estimation are probably the only methods used in certain extent in CTBT applications that can be attributed to the given branch of technology. However Expert Technical Analysis (ETA) conducted in CTBTO to improve the estimated values for the standard signal and event parameters according to the Protocol to the CTBT may face problems which cannot be resolved with certified CTBTO applications and may demand specific techniques not presently used. The problem to be considered within the ETA framework is the unambiguous separation of signals with close arrival times. Here, we examine two scenarios of interest: (1) separation of two almost co-located explosions conducted within fractions of seconds, and (2) extraction of explosion signals merged with wave-trains from strong earthquake. The importance of resolving the problem related to case 1 is connected with the correct explosion yield estimation. Case 2 is a well-known scenario of conducting clandestine nuclear tests. While the first case can be approached somehow with the means of cepstral methods, the second case can hardly be resolved with the conventional methods implemented at the International Data Centre, especially if the signals have close slowness and azimuth. Independent Component Analysis (in its FastICA implementation) implying non-Gaussianity of the underlying processes signal's mixture is a blind source separation method that we apply to resolve the mentioned above problems. We have tested this technique with synthetic waveforms, seismic data from DPRK explosions and mining blasts conducted within East-European platform as well as with signals from strong teleseismic events (Sumatra, April 2012 Mw=8.6, and Tohoku, March 2011 Mw=9.0 earthquakes). The data was recorded by seismic arrays of the International Monitoring System of CTBTO and by small-aperture seismic array Mikhnevo (MHVAR) operated by the Institute of Geosphere Dynamics, Russian Academy of Sciences. Our approach demonstrated a good ability of separation of seismic sources with very close origin times and locations (hundreds of meters), and/or having close arrival times (fractions of seconds), and recovering their waveforms from the mixture. Perspectives and limitations of the method are discussed.

DPRK 2006 and 2013 nuclear tests mixtures, recorded at the MHVAR array. Sampling rate is 200 Hz.

400	
2	~~~
	\sim
	M
/////////////////////////////////	m
	ma
	~~~
·	m
······································	~~~~
······	~~~
	~~~
· ····································	~
WWAMMIN MARKAN	~~~
1000 2000 3000 4000 5000 6000 7000 8000 9000	10000

	for manite of the second second with the second of the second s	MMMm
	hamming warman war and a second a	MMM
	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	nna
	man	www
	w	vvv
	m	m
		m
and where the second se		www
www.www.www.www.www.www.www.www.www.ww	mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm	vw
man man man and the second and the second and the second s	m	m
mannessen and the second	mm	m
	man man man and the second sec	www

DPRK 2013 (top) and 2006 records at MHVAR array, Moscow region. Mixtures on the right: 3 (top) and 0.3 seconds signal separation

mmminuperimeters and more than the second seco	Mo
man M	M.
································	~A
	w
	~~~
www.www.www.www.www.www.www.www.www.ww	~~
	~-
	w
a www.www.www.www.www.www.www.www.www.ww	~~
······································	~~
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	~~~~
	~
1000 2000 3000 4000 5000 6000 7000 8000 9000	10000
for my han many many many and many my many my	14.~
mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm	M
······································	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
	~~
	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
men and the second s	~~
	~~~~
	~~~~
	~~~~
	~~~
	~~~~



## Independent Component Analysis



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



inal (left) and Recovered (right) signals for different time separation Correlation coefficients (CC) is between restored and original ls are in a range of (0.84, 0.94). CC's variability is due to instability mbiguity of the ICA procedures.

The density function of the Laplace distribution, which is a typical supergaussian distribution. For comparison, the Gaussian density is given by a dashed line. Both densities are

> Hyvärinen, A. (1999). Survey on independent component analysis. Neural Computing Surveys, 2:94–

> Hyvärinen, A. and Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. Neural Computation, 9(7):1483– 1492.

The FastICA learning rule finds a direction, i.e. a unit vector W such that the projection  $\mathbf{W}^T \mathbf{X}$  maximizes nongaussianity. Nongaussianity is here measured by the approximation of **negentropy**  $J(W^T X)$ . The FastICA is based on a fixed-point iteration scheme for finding a maximum of the non-gaussianity of  $\mathbf{W}^T \mathbf{X}$ , see (Hyvärinen and Oja, 1997; Hyvärinen, 1999). It can be also derived as an approximative Newton iteration (Hyvärinen, 1999).

Conclusion. We have investigated a few feasible scenarios for different ranges from recording stations. Data from teleseismic (IMS energy and order of components (permutation)) we consider this avoiding comprehensive explosion test monitoring with hiding array MKAR) and regional (small aperture array MHVAR operated approach so far as a tool for extended interactive analysis, and more specific signals from one seismic event on the background of by Russian Academy of Sciences) seismic stations was processed work to be done to introduce it in automatic processing pipeline. another. Independent Component Analysis appears to be a promising showing applicability of the method for different types of equipment Further plans of using the ICA in seismic monitoring is improving way in seismic signals separation taking into account some and sampling rate (40 Hz at MKAR vs 200 Hz at MHVAR). Also, the robustness of this approach, enforcing homomorphic limitations of this approach. ICA in seismology was performed good results were achieved for recovery of nuclear explosion buried deconvolution with the ICA applied in quefrency domain, and mainly for the seismic exploration data, separating a non-gaussian in a wave-train of strong earthquakes (MKAR array example). ICA introducing the multi-station association independent components reflectivity sequence from the minimum phase sounding signal. In shows good results when other approaches fail, for example, association for improved seismic event location. We also work on many seismic applications ICA showed better separating separation with FK-analysis based on different slowness and development of formal criteria exposing the presence of several performance then SVD-based methods and classical PCA. We have azimuths, or cepstral methods for echo. However ICA performance composing signals in the observed data. successfully applied ICA methods to separation of signals from is proportional to the azimuthal gap between the composing signals, earthquakes, quarry blasts, and nuclear explosions conducted at slowness and sampling rate. Due to the ICA's ambiguity (polarity,

	12	12 marine many many many many many many many many
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	In the second of the second	n www.www.www.www.www.www.www. 10 ANMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM
	10 Why Wind Wind Wind Wind Wind Wind Wind Wind	10 ANAMAMIN MANAMANAMANAMANA
	to be allow to show a solubly service we have allow as a service of solution and solution and	
		Manuhamananapanana
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	· ····································	o arminimum with more and more and and a second and as
		Amprovemment and a second
mmm MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM	· ····································	o flat man for a second of the
		Win Min Manus M
		+ fruhmmont who was the work of the second
		www.www.lahan.www.www.hun
month and the second of the second se	= ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	s flown flow many many many many many many many many
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~		1000 1500
aparation results. Original signals, sanarated	300 1000 1000 2000 2000 3000 3000	Separation results.
(CA) and PCs of original signals	Quarry #1+ Quarry#1 signals mixture	(ICA) and PCs of
	a have been and the second of	1
	12	•
100 200 300 400 500 600 700 800 900 1000 Single Array Element: Event#2		1 100 200 300
	to be as the sector of an also as a sector of a	
100 200 300 400 500 600 700 800 900 1000 First Independent Component		4
i i i i i i i i i i i i i i i i i i i	s and a series of a few and a series and a series and a series full with the most day of a few to be the few a	-4 100 200 300 4
- AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		
100 200 300 400 500 600 700 800 900 1000 First Principle Component Event #1	4 h and more a man and a non and an a low of a non an an an an and and and and and and an	-4 10 200 300
	and a second a second a second a second a second a second a	0
100 200 300 400 500 600 700 800 900 1000	And all an all more and a well bloch and the second and the second and a second and and a second all and a second all and the second seco	-2 100 200 300
	www.www.www.www.www.www.www.www.www.ww	
	500 1000 1500 2000 2500 3000 3500 4000	-2 100 200 300
		A second seco

Quarry #1	Quarry #2	Quarry #2	Qu
	22	12	12 month have a man man man man
man man man man man man man man and man han han han han han han han han han h	Man Market and Market		hunder the second secon
mound of the second	10	10 mmmmy happy and many happy and ha	10 ANAMMANA WINDOW WWWW
warman warm		mmmmy way was a second way was a second way and a second way was a second way way and a second way way a second	Manahan Manana Mananana
annon an annon an annon ann ann ann ann	· ····································	· ····································	o www.www.www.www.www.www.
Man Marken Ma	MANNA MAN	mmm My Mar man	por many many
mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm	· ····································		offer mary www.
mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm		man and all the for the second stand of the second stand the second stand sta	Man
Ammenter and the second and the second secon			+ from monowing and the monowing a
manner man man man man MMMMMMMMMMMMMMMMMMMMMMM		monor when the second of the second	manyamalatan
e an	2 mm Manufann m MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM	2 mmmmm man man man man man man man man m	2 4 marthall walk with the walk when the second
Man Marken Ma			Monworken-Whenwork
500 1000 1500 2000 2500 3000 3500 4000 Quarry #1+ Quarry #2 signals mixture	Separation results. Original signals, separated (ICA) and PCs of original signals	Quarry #1+ Quarry#1 signals mixture	Separation results. (ICA) and PCs of
12		12 marshin way	
amman Manan Mana	-1 100 200 300 400 500 600 700 800 900 1000	www.www.www.www.www.www.www.www.www.ww	-1 100 200 300
		10 & Manumenter and the second se	
mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmm		warman	-1 100 200 300
o		a-manumumumumumumumumumumumumumumumumumumu	
Mar Man Mar		www.www.www.www.www.www.www.www.www.ww	-2 -4 100 200 300
s a manufacture was a second with a second was a second a second with a second was a second was a second with a second was a second w	4 Last Independent Component	o warmen w	
			-2 -4 100 200 300
A Arman Market	2 First Principle Component Event #1	· frances manufacture and a second manufacture and a second secon	
man man man man and man and man and man and and and and and and and and and a			-2 <u>100 200 300</u>
z	2 100 200 300 400 500 600 700 800 900 1000 First Principle Component Event #2	s ar when a man and a philling and a	2 Fir
menon was well and the second se		www.www.www.www.www.www.WWWW/WWWWWWWWWW	
Comparing: ICA com	ponents (restored signals) w	ith first principle compone	nt of the arı

Separating two regional earthquakes (left to right): EQ1, $\Delta = 8^{\circ}$; EQ2, $\Delta = 5^{\circ}$; Mixture; Decomposed, 400 sec; Decomposed, 50 sec.

	MMMANANCESCHARMENERGENERGENERGENERGENERGENERGENERGENE
	effetially/familier-instrumenter-inforce-ie-i-
	nfhefahlandersen skolanteren atornaliseren er
	the fill of the participation
	la fan skip of her her fan
	114/14-warentherestertenen auto-anarramentenen .
	nality of the state of the stat
	and the second
	will will will be the full in month to all on a south on more
These (servanisses) 22:00:00 22:00:00 22:00:00 22:00:00 22:00:00 1000 2000 3000 4000 5000 6000	1 1 1 7000 8000

Separating events with close azimuths from the mix: DPRK-2013, Feb-2013 Mb=4.9 (scaled up to Mb=7) test and Tohoku, March 2011 Mw=9.0 EQ.

Earthquake waveforms at MKAR IMS array

	-
9	- an an and for the and the state of the state and the state of the st
8	
7	
ò	
5	
5	
2	
l	

•
,
٥
s
•
·
2
1 1

Earthquake waveforms at MKAR array (top) and mix with DDDV 2012 test

	and mix with DFKK-2015 test	
9		1500
8		1000
7		1000 -
6		500
5		
4		0
3		-500
2		
1		-1000
	1000 2000 3000 4000 5000 6000 7000	
9		
9		20
9		20 15
9 8 7 6		20 15 10
9 8 7 6 5		20 15 - · 10 - · 5 - ·
9 8 7 6 5		20
9 8 7 6 5 4 3		20 15 10 5 5 5
9 8 7 6 5 4 3		20 15 10 5 -5 -10
9 8 7 6 4 3 3 1		20 15 10 5 -5 -10 15

Disclaimer

The views expressed on this poster are those of the authors and do not necessary reflect the views of the CTBTO Preparatory Commission and Institute of Geosphere Dynamics, RAS



Quarry blast signals separation



DPRK-2013 test at MKAR array

	A mix of both events
1000 2000	



A case of separation of events with different back-azimuths from the mix: DPRK-2013, Feb-2013 Mb=4.9 (scaled up to Mb=7) test and Sumatra, April 2012 Mw=8.6 earthquake Recovered test signal (top) and first PC Repeatable mix of the synthetic DPRK-2013 array seismogram

3 for my for my for my for my for my for my for the fo

7 for me for the for the former former former for the former former

6 from the form the method of the formation of the format

4 for the for the second for the former former for the former former for the former former former for the former former former for the former former

3 for an form for a form for any property of the second for the second for the second and the second as a s

2 to for the providence of the second of the

1 from from from the market of the part of

with the EQ array seismogram, MKAR IMS array





in different time scales, and 1st PC of the original



preparatory commission for the comprehensive nuclear-test-ban treaty organization

uarry #1

and and a second the second of the second seco mmmmmmllimlilling mstormthellymbershippinghamman on www.www. manana and and a second and a MANAMAMAMANA MANAMANA MANAMANA

Original signals, separated original signals

ay seismogram

Separation of signals from different quarries. Solid lines indicate input waveforms for the actual example (mixture, and decomposition on the right). Dash line indicates typical waveform used to produce similar signal separation result. In first case primary phase of Q#1 event fall onto the secondary phase of Q#2 event making them visually inseparable. In second case all arrivals of both events fall on the same time producing visually a new event. A figure with separation results consists of: (1) a pair of input signals (from same channels) composing the mixture, (2) separated signals, and (3) first principle components of the array seismograms corresponding to each signal.

We present the PCs to emphasize that the restored signals inherit the property of the whole array, not just some single component, since the best similarity is provided between the first PCs of the input array and the best Independent Component. The black vertical rounded lines indicate corresponding signals (omitted on right figure). Correlation coefficients between the original and restored signals are in a range (0.84, 0.92).

ncomp 1	: Regional EQ, d	elta=500km				
WUU	allouldation	na dan baharan sa	and the second second second			
undihan.	all distantic marries					
40	UU 51	uuu (5000 7	000	8000	9000
IGA	i xi (nestoreu sij	yna)	!	!	1	
an lat	la de la composición	1 billenta				-
ime M	alah yang kali yang k	walke sales billion	NH-WANNAMAN	al water the street of the str	and the second state of the second	-
						1
40	00 5	000 6	5000 7	000	8000	9000
incomp	1: Regional EQ,	delta=800				_
						_
mtulh		with with he had	ilin Maliniahow	hiddlewn Marson ar	www.	_
	a na shi dina sa shi	a na naffe difta.	uli radianal a			_
40	00 5	000 0	8000 7	000	8000	9000
ICA	x2 (Restored sid	unal)	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	000	0000	5000
			1	!	1	٦
بالمصاب	ad tabat mardala		an an dan an ar			1
ad a second s	talbuardis and the second	and a subject of the	hit. Manual hits and a state of the state of	lihilika na ditavani	ndellandellanger (norder ogene	1
40	00 5	000 0	2000 7	000	9000	9000



Recovered EX signal (bottom) and first PC Recovered EQ signal (top) and first PC

_×	10 ⁶		Third Inc	dependent Cor	nponent			
2			1					
5								
ĩ								
5		dial - data -		All de la				
0				M MARA	ulikilkilki	hilling	or a state of the	k
5	dam total at	() fa han a definition in		a la contrava	Al de l'an	dike ta e wale.	a addit hale a sec	·····-
1		1 " 1		· PI 1 *				
5								
20	1000	2000	3000	4000	5000	6000	7000	8000
20	1000	2000	3000	4000	5000	6000	7000	8000
20	1000	2000	3000	4000	5000	i 6000	7000	8000
8	1000	2000	3000 First F	4000 Principle Comp	onent	6000	7000	8000
8	1000	2000	3000 First F	4000 Principle Comp	5000 onent	i 6000	7000	8000
8	1000	2000	3000 First F	4000 Principle Comp	5000 onent	i 6000	7000	8000
2 0 8 - 4 -	1000		3000 First F	4000 Principle Comp	onent	j 8000	7000	8000
8			3000 First F	4000	onent	і 6000		0008
8 6 4 0	1000	2000	3000 First F	Principle Comp		oooa Mundikalaan	7000	8000
8 6 4 2 2	1000		3000 First F	Principle Comp			7000	8000

EQ and synth-DPRK-2013 mix, restored explosion



Reliability of the ICA estimates is illustrated by the clusters separation (top). Two connected separated clusters 1 and 2 correspond to first two (out of 9) ICs (bottom figure), representing the input explosion repeatable sequence. Cluster #5 is also well separated and represents the EQ. Other clusters are mostly EQ+EX related patterns. (ICASSO software (Johan Himberg and Aapo Hyvärinen, 2003)

		Independent components (ranked according to I _q)	
		-flerne flerne	**
ł	2 -	-fle-an fle-an fle-an fle-an fle-an fle-an fleman parts and an ser fleman f	a
;	5 -	fr for the free from the free the free the free the free of the fr	4
	-	en e	Ŵ
Laluen	3 =		-
1	3 -	to the	~
;	-	m man flage and flage and flage into flage into flage the second on the	~
-	+		H
5	•	an and a stand of the stand of	'n
	0	, 1000 2000 3000 4000 5000 6000 7000 6000 9000 10 Sample ♥	1000

International Data Centre http://www.ctbto.org