

# Recovery of the largest aftershock sequences using waveform cross correlation

Ivan Kitov and Christos Saragiotis

Poster No. P3.5-354

CTBTO



The waveform cross-correlation (WCC) method - based on the similarity of neighbouring seismic sources and using a fixed station set - is an alternative method well suited to the nature of the aftershock sequences, i.e. similar repeating events

Shallow earthquakes with  $M_w > 8-9$  usually create extensive aftershock activity with thousands of events with  $m_b > 3.5$ . This is a tremendous challenge for IDC automatic and interactive processing, with the Tohoku-2011 earthquake being a prime example. The SHI IDC processing is fine-tuned to find the smallest events (with fixed human resources), and thousands of similar events within a very short time stresses the system. The waveform cross-correlation (WCC) method - based on the similarity of neighbouring seismic sources and using a fixed station set - is an alternative method well suited to the nature of the aftershock sequences, i.e. similar repeating

events within a confined area measured by seismic stations of the International Monitoring System. The IDC has been testing the performance of a prototype WCC-based processing pipeline. There is a feature very specific to the IDC: the resulting cross-correlation bulletin (XSEL) contains only the events matching the full set of event definition criteria applied to the Reviewed Event Bulletin. Here we present the testing results as obtained from four different aftershocks sequences. For these sequences we show the effect of varying defining parameters of the prototype WCC pipeline and two sets of master events – the REB and automatic SEL3

- Principal features of aftershock sequences of catastrophic earthquakes: Tohoku as the greatest challenge
- Waveform cross correlation (WCC) processing to build preliminary cross correlation standard event list (XSEL)
- Master Events
- Tentative results of four aftershock sequences recovery
- Tuning of CC-detection, Local Association, and Conflict Resolution parameters
- Conclusion



## Four aftershock sequences with different best seismic station sets are studied

### Area setting

sequence	start	end	lat, deg	lon, deg
Tohoku	2011070	2011076	30N - 46N	134E-148E
Nepal	2015115	2915121	25N - 30N	83E-88E
Chile	2015259	2015265	28S - 34S	67W-75W
PNG	2018056	2018062	1S - 11S	136E-148E

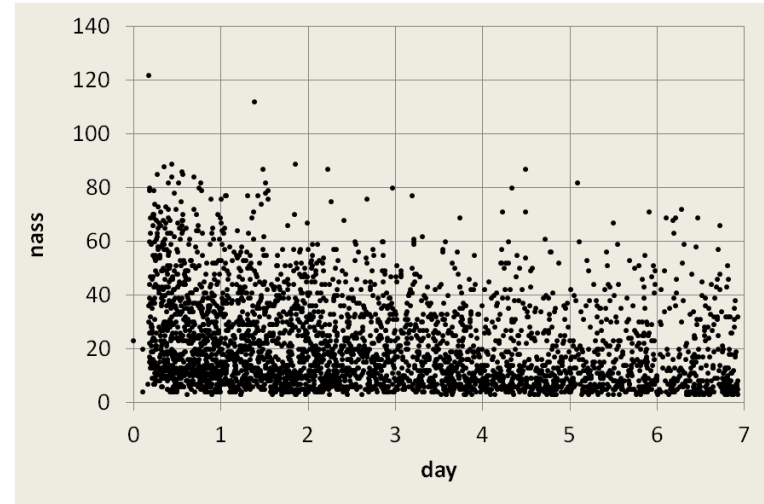
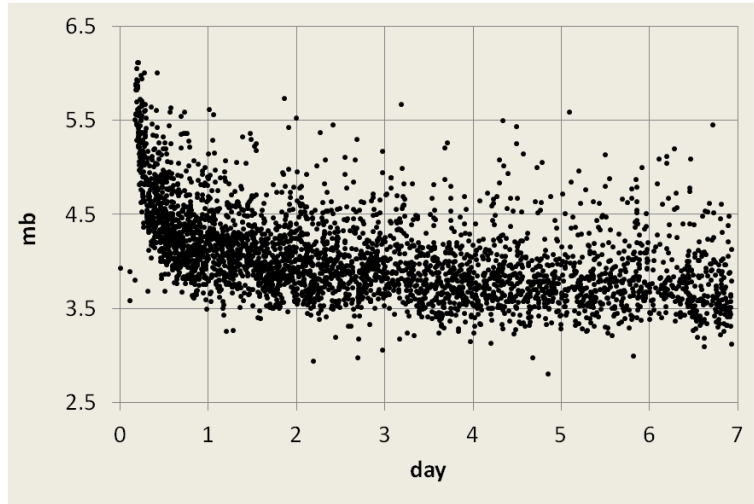
Total area of the aftershock zones can be larger than 100,000 sq.km. The aperture is larger than 600 km. Seismic stations at shorter distances have varying azimuth to aftershocks. The master event approach needs to cover the whole area for efficient processing.

### Main shock IDC solution

IDC	Origin	lat, deg	long, deg	mb	D, km	nass
Tohoku	05:46:20	38.435	142.520	5.51	0.0	122
Nepal	06:11:24	28.159	84.703	5.98	0.0	74
Chile	22:54:29	-31.517	-31.517	5.86	0.0	82
PNG	17:44:40	-6.044	142.800	5.85	0.0	63

All four events were shallow. Body wave magnitude scale is saturated near 6.0 and does not represent the size of these earthquakes well. The number of stations in the zones of existing seismic waves (excluding the shadow zone between 100 and 115 degrees) varies over the studied events.

## Tohoku REB : 2011070-2011076



Body wave magnitudes of the REB events during the first day are much higher, and then the average magnitude decreases. The same effect, but less prominent, is observed in the number of associated phases. The biggest events generate extremely large number of secondary phases, which mask primary phases from smaller events, lead to multiple (**combinatorial explosion**) misassociations into invalid event hypotheses in SEL3, and also create noise coherent with valid signals, *i.e.* negative conditions for waveform cross correlation.

## To compare results of different procedures it is necessary to develop metrics

1. How many SEL3/XSEL events were built?
2. How many SEL3 events match REB by *evid* and how many were rejected?
3. How many SEL3/XSEL events match REB by 2+ phases within 10 s?
4. What in the distribution of SEL3/REB and XSEL/REB distances?
5. What is the distribution of SEL3/REB and XSEL/REB depth difference?
6. How many arrivals associated with SEL3 are in the REB by *arid*?
7. What is the rate of correct interpretation (*e.g.* N-phases renamed) of SEL3 phases associated with the REB by *arid*?
8. What is the distribution of arrival time residuals?
9. ...

**Main question: how can we reduce the analyst workload related to the largest sequences and how can we estimate the improvement?**

***This is a multi-dimensional task lacking quantitative estimate***



Tohoku	day	REB	nass (P)	SEL3	by evid	SEL3		
						nass	by arid	REB only
1	2011070	743	14375	887	540	14272	6182	203
2	2011071	714	12044	946	559	14082	5952	155
3	2011072	574	8781	712	440	19715	4694	134
4	2011073	444	6544	603	354	7477	3289	90
5	2011074	392	4978	522	303	5954	2585	89
6	2011075	321	4157	481	257	4923	2163	64
7	2011076	322	3829	477	241	5182	2139	81
Total		3510	54708	4628	2694	71605	27004	816

Nepal	day	REB	nass (P)	SEL3	by evid	SEL3		
						nass	by arid	REB only
1	2015115	211	2620	262	132	2882	1411	79
2	2015116	83	935	239	58	978	494	25
3	2015117	49	417	154	31	351	189	18
4	2015118	26	180	174	17	146	84	9
5	2015119	22	215	175	18	220	116	4
6	2015120	16	109	183	7	78	39	9
7	2015121	13	124	182	11	126	70	2
Total		420	4600	1369	274	4781	2403	146

Chile	day	REB	nass (P)	SEL3	by evid	SEL3		
						nass	by arid	REB only
1	2015259	36	406	220	13	396	162	23
2	2015260	312	3385	391	184	3556	1531	128
3	2015261	119	1154	256	73	1280	513	46
4	2015262	82	994	233	58	1142	516	24
5	2015263	64	605	195	36	722	314	28
6	2015264	48	571	210	37	880	347	11
7	2015265	37	364	211	30	461	213	7
Total		698	7479	1716	431	8437	3596	267

PNG	day	REB	nass (P)	SEL3	by evid	SEL3		
						nass	by arid	REB only
1	2018056	122	1326	243	80	1397	589	42
2	2018057	195	1633	284	133	2102	810	62
3	2018058	90	826	250	69	1020	451	21
4	2018059	71	605	225	53	946	357	18
5	2018060	53	317	195	40	426	168	13
6	2018061	56	466	222	44	706	278	12
7	2018062	33	223	171	27	320	113	6
Total		620	5396	1590	446	6917	2766	174

The Tohoku sequence is the most difficult challenge to automatic and interactive processing. From 23% (Tohoku) to 38% (Chile) of REB events were added by analysts from scratch. This work also included a huge number of newly associated phases. Thousands of event hypotheses were rejected.



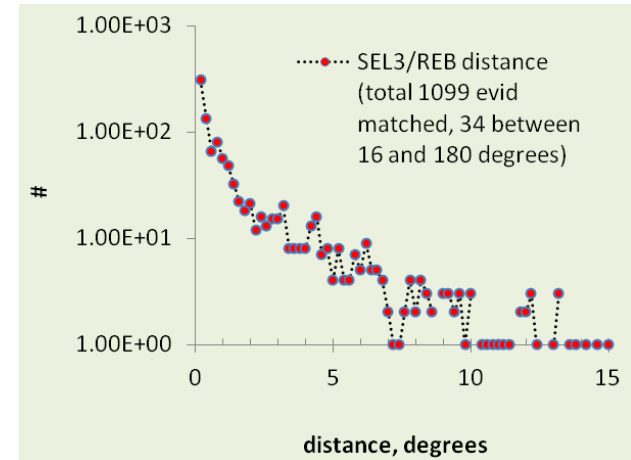
## SEL3 problems for the largest aftershock sequences

1. Extremely high detection rate close to saturation
2. However, too high detection threshold – many actual REB detections missing
3. Large distance between consecutive arrivals with similar characteristics – appropriate detections suppressed/rejected
4. High rate of phase misinterpretation due to similarity in characteristics
5. High rate of phase misassociation into false events
6. As a result, combinatorial explosion in the number of false event hypotheses (especially for scanner)
7. Extremely high event hypotheses rate close to saturation

***Too many poor event hypotheses (rejection workload), not enough good event hypotheses (creation from scratch workload)***

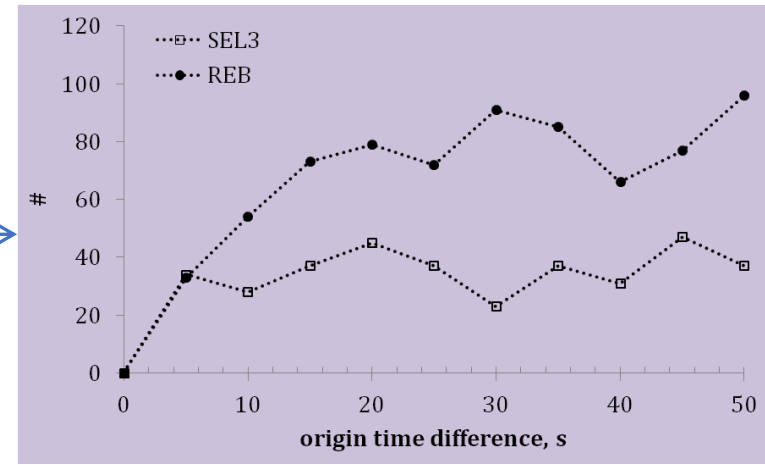
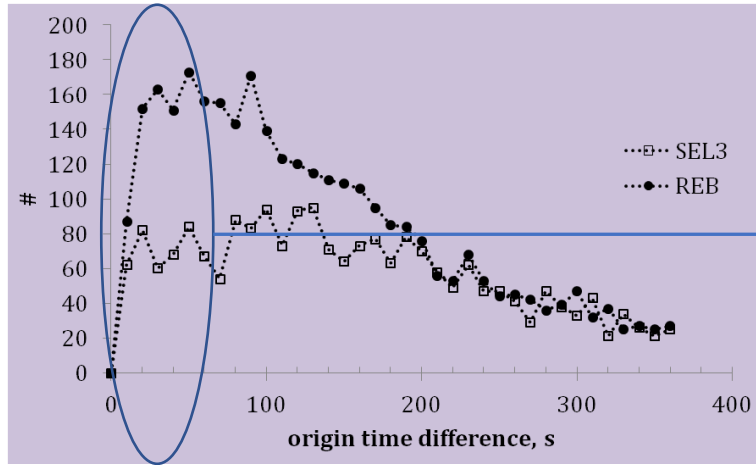


SEL3 events – blue circles  
REB events - black circles



SEL3 events (matching REB by *evid*) are distributed over the globe and the REB events are within a relatively small area. The distance between the SEL3 and REB events varies from several km to 153 degrees. There are 201 SEL3 events beyond 4 degrees and 34 in the range between 15 and 153 degrees.

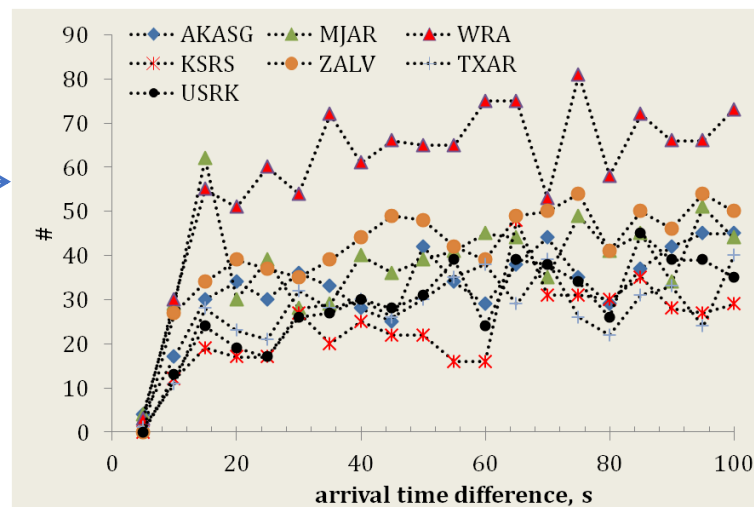
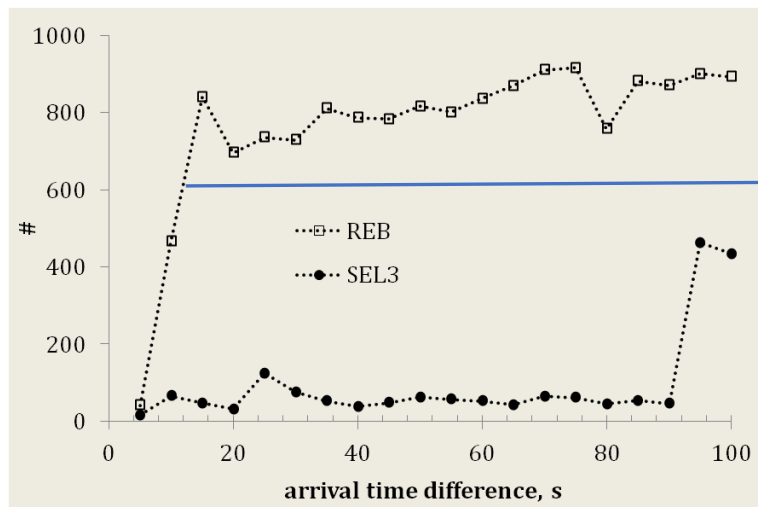
## Frequency distribution of time differences between **origin times** of sequential events



Because of very high rate of events and signals, the SEL3 misses many events with short difference between origin times. This is likely due to the rule prohibiting similar P-wave arrivals within 90 s, and the signals from the aftershock area are similar. This rule suppresses detection of valid signals from events close in time and space: subsequent arrivals with shorter time difference at a given station are highly likely tx instead of P. The REB events are added by analysts from scratch.



## Frequency distribution of time difference between arrival times of sequential arrivals at IMS stations



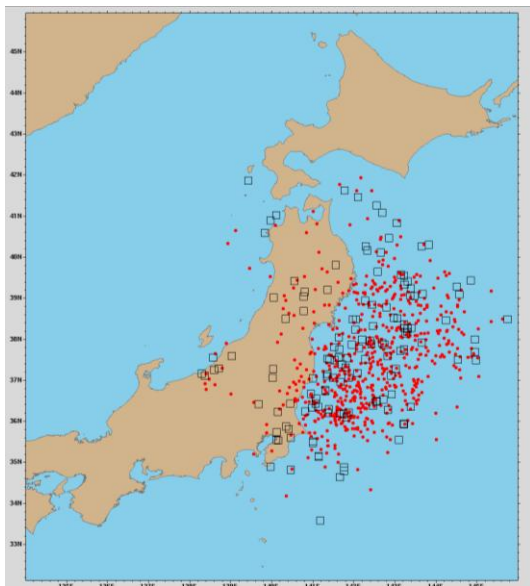
- The REB events have large number of arrivals at the same station within 20 s (likely added by analysts)
- The SEL3 has a significant leap in arrival time difference around 90 s (DFX rule)
- Small time difference between arrivals in a challenge for any association and conflict resolution process



## Aftershock sequence recovery: general approach

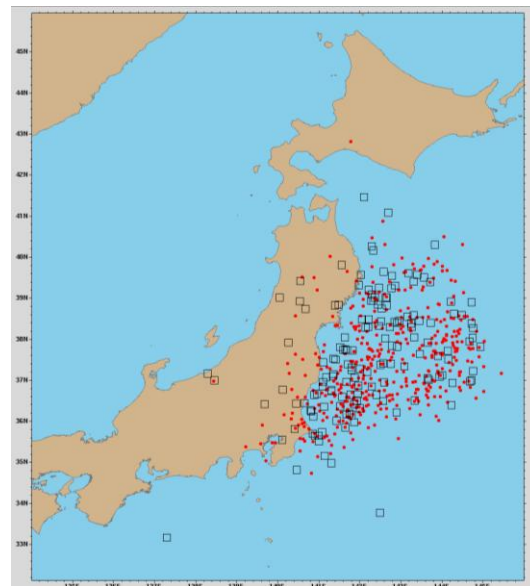
1. Use waveform cross correlation (WCC) with master events from the REB/SEL3 and signals at primary IMS stations as waveform templates
2. Use WCC programs of the XSEL prototype pipeline
3. Use Local Association and Conflict Resolution programs as in routine prototype XSEL production
4. Use only the REB masters events from the past: more than 2 days before the data day (48 hours rule for the REB production). Expedite REB creation is also considered.
5. It is allowed to use SEL3 master events from the same day without interactive review
6. An REB event is matched by an event in the XSEL when 2 or more phases are within **10 s** from each other at two or more different stations (statistics with 1 match is also helpful)
7. Compare the number of matched events: the rate of matched events is a priority
8. Calculate total number of events in the XSEL: balance the analysts workload with that expected with the SEL3 and scanner seed events for interactive analysis (*to be determined*)
9. Estimate the trade off between the rate of matched and additional XSEL events depending on Local Association and Conflict Resolution parameters
10. Evaluate the quality of the REB events using Spot Check Tool

## REB and historical master events 2011070



Red circles – 743 REB events, black squares – 138 master events. Masters – historical REB

## REB and (fresh REB) master events 2011073



Red circles – 444 REB events, black squares – 150 master events. Some masters – REB from 2011070

		#		XSEL/	REB	%REB	Quality	QC-SC REB	% SC REB	LA REB	% LA REB
Tohoku	REB	masters	XSEL	REB	match	match	check, SC	match	match	match	match
2011070	743	138	987	1.33	549	0.74	866	689	0.93	635	0.85
2011071	714	193	1470	2.06	611	0.86	815	656	0.92	675	0.95
2011072	574	203	1242	2.16	505	0.88	657	563	0.98	548	0.95
2011073	444	150	981	2.21	413	0.93	556	437	0.98	431	0.97
2011074	392	214	762	1.94	356	0.91	466	389	0.99	373	0.95
2011075	321	159	551	1.72	284	0.88	368	316	0.98	302	0.94
2011076	322	214	611	1.90	274	0.85	356	308	0.96	287	0.89

The Tohoku aftershock sequence is the largest by the number and daily rate of REB events. Overall success (REB matching) rate is much larger than in the SEL3 (from 74% to 93%). On 2011070, the automatic Spot Check Tool confirmed 689 from 743 events (93%). Therefore, 54 events might be revisited with interactive (spot check) analysis to re-confirm.

For the days 3 through 7, the QC confirmation rate was around 98%. During LA processing, before CR, there were 635 REB events matched by preliminary events hypotheses (85%). The success rate of LA match is high for all days except the first one, likely because of lower ambient noise level. When the fresh REB events from the

same sequence were used on days 4 to 7, the success rate increased to 93%, with the total number of XSEL events two times larger when in the REB.

Overall, the matching rate and the excess of XSEL events demonstrate good performance.



Nepal	REB	# masters	XSEL	XSEL/ REB	REB match	%REB match	Quality check, SC	QC-SC REB match	% SC REB match	LA REB match	% LA REB match
2015115	211	55	332	1.57	170	0.81	200	185	0.88	179	0.85
2015116	83	62	126	1.52	75	0.90	92	78	0.94	77	0.93
2015117	49	62	68	1.39	44	0.90	51	49	1.00	44	0.90
2015118	26	24	44	1.69	25	0.96	27	26	1.00	26	1.00
2015119	22	35	59	2.68	21	0.95	22	22	1.00	22	1.00
2015120	16	44	42	2.63	15	0.94	17	14	0.88	15	0.94
2015121	13	45	54	4.15	13	1.00	13	13	1.00	13	1.00

Chile	REB	# masters	XSEL	XSEL/ REB	REB match	%REB match	Quality check, SC	QC-SC REB match	% SC REB match	LA REB match	% LA REB match
2015259	36	41	54	1.50	25	0.69	36	32	0.89	29	0.81
2015260	312	87	482	1.54	257	0.82	332	290	0.93	276	0.88
2015261	119	102	169	1.42	104	0.87	129	110	0.92	107	0.90
2015262	82	59	102	1.24	72	0.88	90	78	0.95	73	0.89
2015263	64	134	77	1.20	44	0.69	69	58	0.91	45	0.70
2015264	48	95	77	1.60	42	0.88	49	44	0.92	42	0.88
2015265	37	36	50	1.35	32	0.86	35	35	0.95	32	0.86



PNG	REB	# masters	XSEL	XSEL/ REB	REB match	%REB match	Quality check, SC	QC-SC REB match	% SC REB match	LA REB match	% LA REB match
2018056	122	60	128	1.05	99	0.81	105	103	0.84	102	0.84
2018057	195	83	243	1.25	170	0.87	176	178	0.91	172	0.88
2018058	90	64	124	1.38	76	0.84	93	89	0.99	77	0.86
2018059	71	38	109	1.54	67	0.94	69	69	0.97	67	0.94
2018060	53	66	104	1.96	50	0.94	53	53	1.00	50	0.94
2018061	56	78	110	1.96	53	0.95	55	55	0.98	53	0.95
2018062	33	27	70	2.12	31	0.94	33	33	1.00	31	0.94

The overall success (REB matching) rate is high for this aftershocks sequence (from 81% to 95%), which is not a challenge by the number of events per hour and total number. The highest rate was on 2018056 and this day is characterized by the lowermost rate of matched events of 81%. The automatic Spot Check Tool

confirmed only 103 from 122 events for this day. Therefore, 19 events might to be revisited with interactive analysis. During LA processing, there were 102 REB events matched by preliminary events hypotheses.

When the REB events from the same sequence were used on days 4 to 7, the success rate increased to >94%,

with the total number of XSEL events two times larger when in the REB. For the first 3 days, the number of XSEL events was only by ~ 25% to 35% larger.

Overall, the matching rate and the excess of XSEL events are close to the most efficient performance.

Tohoku	REB	# masters	XSEL	XSEL/REB	REB match	%REB match	LA REB match	% LA REB match
Standard LA and CR parameters								
2011070	743	138	987	1.33	549	0.74	635	0.85
2011071	714	193	1470	2.06	611	0.86	675	0.95
2011072	574	203	1242	2.16	505	0.88	548	0.95
Lowered thresholds for LA and CR parameters								
2011070	743	138	1208	1.63	587	0.79	709	0.95
2011071	714	193	1423	1.99	623	0.87	696	0.97
2011072	574	203	1341	2.34	504	0.88	555	0.97
Increased thresholds for LA and CR parameters								
2011070	743	138	725	0.98	530	0.71	654	0.88
2011071	714	193	758	1.06	544	0.76	657	0.92
2011072	574	203	792	1.38	479	0.83	547	0.95

We have changed the following LA and CR parameters:

1. Area of LA from 60 km to 300 km in radius;
2. Travel (origin) time residual from 3 s to 5 s;
3. Window for CR from 10 s to 60 s;
4. The max SNRcc in the event – from 5.5 to 6.5;
5. Minimum weight from 20 to 24;
6. Minimum sum of SNRcc in a 3-station event from 14 to 17.

Optimal setting of LA and CR parameters is a challenge for the largest aftershock sequences. The first 3 days of the Tohoku sequence is the best example of potential tuning benefits. The effect of increasing or decreasing thresholds can be counterintuitive when the flux of detections and origin time is saturated and XSEL event hypotheses compete with many rivals.

**Disclaimer:** The views expressed on this poster are those of the author and do not necessarily reflect the view of the CTBTO

## Further tuning LA and CR: Tohoku most intensive days

1. Change STA = 0.8 s for more effective (CC-) noise suppression
2. Vary CC-detection thresholds (from 2.4 to 3.4)
3. Filter CC-detections with standard SNR
4. Vary LA parameters – grid, thresholds in the EDC
5. Vary Conflict Resolution parameters, *e.g.* time window, event magnitude
6. Select various sets of masters for various scenarios of larger aftershock processing
7. Compare with REB (and SEL3) for optimal set of parameters

## Alternative strategies to reduce the IDC analyst workload and corresponding sets of master events

1. Historical REB as master events (as many as needed, up to 400, tested before)
2. Best historical REB as master events selected by CC-procedure (126)
3. SEL3 biggest events (122)
4. SEL3 events distributed over regular (global) grid (102)
5. REB events built from the set of SEL3 events regularly distributed over the grid (102)
6. REB events near regular grid (65)
7. REB events built from the biggest SEL3 events (159)
8. **Benchmark:** Best REB for a given day as master events selected by CC-procedure (126)



## Further tuning detection, LA and CR: Tohoku 2011070. SEL3 events as masters

MASTERS	CASE	Threshold	LA CR Setting	#REB (6-24)	#XSEL	#REB/#XSEL	#REB Match	Match XSEL/ REB	LA Match	LA Match
SEL3	GRID	3.2	REGULAR	734	647	0.88	553	0.75	717	0.98
SEL3	GRID	3.0	REGULAR	734	658	0.90	568	0.77	725	0.99
SEL3	GRID	2.8	REGULAR	734	684	0.93	582	0.79	726	0.99
SEL3	GRID	2.4	REGULAR	734	839	1.14	602	0.82	732	1.00
SEL3	BIGGEST	3.2	REGULAR	734	671	0.91	555	0.76	723	0.99
SEL3	BIGGEST	3	REGULAR	734	692	0.94	571	0.78	727	0.99
SEL3	GRID	3.2	LOWER	734	738	1.01	561	0.76	722	0.98
SEL3	GRID	3.0	LOWER	734	778	1.06	560	0.76	731	1.00
SEL3	GRID	2.8	LOWER	734	841	1.15	588	0.80	731	1.00
SEL3	GRID	2.4	LOWER	734	990	1.35	609	0.83	734	1.00
SEL3	BIGGEST	3.2	LOWER	734	773	1.05	575	0.78	726	0.99
SEL3	BIGGEST	3.0	LOWER	734	825	1.12	577	0.79	731	1.00
SEL3	BIGGEST	2.4	LOWER	734	951	1.30	603	0.82	733	1.00
SEL3	GRID	2.4	LOWEST	734	1321	1.80	639	0.87		
SEL3	BIGGEST	2.4	LOWEST	734	1317	1.79	637	0.87		

## Further tuning detection, LA and CR: Tohoku 2011070. Various sets of past REB events as masters

MASTERS	CASE	Threshold	LA CR Setting	#REB (6-24)	#XSEL	#REB/#XSEL	#REB Match	Match XSEL/ REB	LA Match	LA Match
REB	BEFORE	3	REGULAR	734	836	1.14	584	0.80	731	1.00
REB	BEFORE	2.8	REGULAR	734	808	1.10	587	0.80	730	0.99
REB	BEFORE	2.4	REGULAR	734	1131	1.54	619	0.84	733	1.00
REB	FROM_SEL3_GRID	2.4	REGULAR	734	926	1.26	618	0.84	733	1.00
REB	FROM_SEL3_GRID	3	REGULAR	734	737	1.00	573	0.78	729	0.99
REB	GRID	2.4	REGULAR	734	893	1.22	611	0.83	734	1.00
REB	GRID	2.6	REGULAR	734	819	1.12	597	0.81	732	1.00
REB	GRID	2.8	REGULAR	734	712	0.97	588	0.80		
REB	GRID	3	REGULAR	734	705	0.96	583	0.79		
REB	GRID	3.2	REGULAR	734	660	0.90	569	0.78		
REB	GRID	3.4	REGULAR	734	639	0.87	559	0.76		
REB	BIGGEST	2.6	REGULAR	734	901	1.23	608	0.83	733	1.00
REB	BIGGEST	3	REGULAR	734	773	1.05	587	0.80	733	1.00
REB	BEST	2.4	REGULAR	734	941	1.28	611	0.83	731	1.00
REB	BEST	2.8	REGULAR	734	796	1.08	592	0.81	733	1.00
REB	BEST	3	REGULAR	734	752	1.02	578	0.79	733	1.00

## Further tuning detection, LA and CR: Tohoku 2011070. Various sets of past REB events as masters

MASTERS	CASE	Threshold	LA CR Setting	#REB (6-24)	#XSEL	#REB/#XSEL	#REB Match	Match XSEL/ REB	LA Match	LA Match
REB	BEFORE	3.0	LOWER	734	1018	1.39	589	0.80		
REB	BEFORE	2.8	LOWER	734	1003	1.37	589	0.80		
REB	BEFORE	2.4	LOWER	734	1423	1.94	629	0.86	734	1.00
REB	FROM_SEL3_GRID	2.4	LOWER	734	1110	1.51	623	0.85	734	1.00
REB	FROM_SEL3_GRID	3.0	LOWER	734	874	1.19	582	0.79	732	1.00
REB	GRID	2.4	LOWER	734	1089	1.48	618	0.84	734	1.00
REB	GRID	2.6	LOWER	734	972	1.32	611	0.83	734	1.00
REB	GRID	2.8	LOWER	734	929	1.27	579	0.79	731	1.00
REB	GRID	3.0	LOWER	734	824	1.12	578	0.79	729	0.99
REB	GRID	3.2	LOWER	734	807	1.10	591	0.81	725	0.99
REB	GRID	3.4	LOWER	734	693	0.94	559	0.76	711	0.97
REB	BIGGEST	2.6	LOWER	734	1062	1.45	618	0.84	725	0.99
REB	BIGGEST	3.0	LOWER	734	906	1.23	598	0.81	734	1.00
REB	BEST	2.4	LOWER	734	1113	1.52	616	0.84	732	1.00
REB	BEST	2.8	LOWER	734	968	1.32	611	0.83	733	1.00
REB	BEST	3.0	LOWER	734	877	1.19	586	0.80	734	1.00
REB	BEFORE	3.0	LOWEST	734	1559	2.12	624	0.85		
REB	FROM_SEL3_GRID	2.4	LOWEST	734	1523	2.07	632	0.86		
REB	GRID	2.4	LOWEST	734	1378	1.88	645	0.88		
REB	BIGGEST	2.6	LOWEST	734	1342	1.83	626	0.85		
REB	BEST	2.4	LOWEST	734	1551	2.11	639	0.87		

**Disclaimer:** The views expressed on this poster are those of the author and do not necessarily reflect the view of the CTBTO

- Waveform cross correlation (WCC) is used to ***detect signals and build aftershock*** events in four large aftershock sequences with the area up to 100,000 km<sup>2</sup>
- This aftershock study uses routine XSEL processing. Historical master events from the REB are allowed as well as contemporary (not reviewed) SEL3 events. The choice can be made automatically.
- Large number of master events (up to 500) was used for the largest region of the Tohoku earthquake aftershocks
- Strategies with ***expedite*** creation of a reasonable number of REB events are considered
- For the largest aftershock sequences, the preliminary estimates of the XSEL-REB daily match rate varies 70% to 100%
- The XSEL is closer to the REB than the SEL3 and suggest less work (***must be quantitatively estimated***) on event creation/confirmation and event rejection (*e.g.*, scanner)
- The use of the best SEL3 events is almost ***as efficient as*** the REB masters
- The distribution of master events over a regular grid is likely **the best choice**
- Further improvements are possible without large additional computations. The learning curve is still steep.