

SnT 2023

CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

HOFBURG PALACE - Vienna and Online

19 TO 23 JUNE



CTBTO
PREPARATORY COMMISSION

COMPREHENSIVE
NUCLEAR-TEST-BAN
TREATY ORGANIZATION

Analysis and Insights on Manually Added Events by IDC Analysts in the Reviewed Event Process

Haijun WANG, Paulina BITTNER, Marcela VILLARROEL,
Christos SARAGIOTIS, Wolfgang SOMMERER, Alexander
POPLAVSKIY, Gerard RAMBOLAMANANA, Gerhard GRAHAM,
Josep VILA CODINA, Thierry HERITIER
CTBTO

PTS-423_WANG

Presentation Date: 22 June 2023

1. IDC Waveform Automatic Data Processing and Interactive Analysis
2. Identifying Events Added by IDC Analysts in the IDC Database
3. Statistic and Parameters Analysis of Added Events
 - 3.1 The System Developed for Parameters Analysis of Added Events
 - 3.2 Statistic and Parameters Analysis
 - 3.3 Comparing bulletins SEL3 and LEB
 - 3.4 In-depth Analysis of a particular Example
 - 3.5 Network detection threshold estimation
4. Discussion and Conclusion

❑ SHI data analysis procedure

▪ Automatic data processing(bulletin)

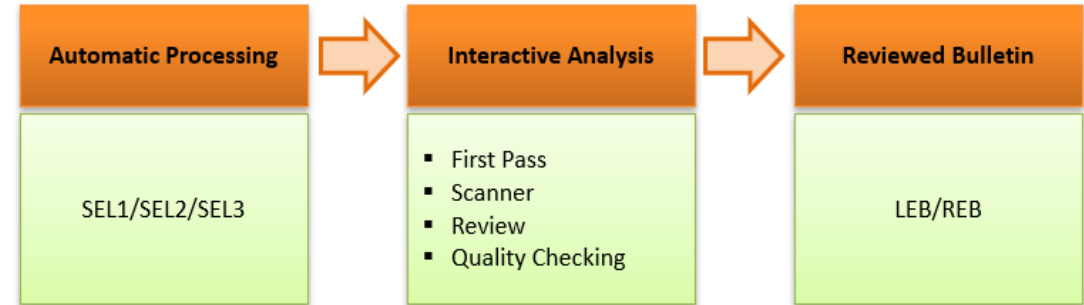
- SEL1 -- 1h after data have been processed in pipeline
- SEL2 -- 4h after data have been processed in pipeline
- SEL3 -- 6h after data have been processed in pipeline

▪ Interactive Analysis

- First Pass: Based on SEL3 events to reject bogus events, improve the quality of real events, save events which meet event definition criteria as LEB.
- Scanner: to build events not produced in SEL3.
- Review: Check quality of LEB events.
- BullQC: Check quality and consistence of LEB events.

▪ Reviewed bulletin

- LEB (3 seismic stations, or 2 infrasound stations, or 2 hydroacoustic stations).
- REB(3 primary stations and weight > 4.6).



IDC Waveform Data Processing Workflow

- ✓ SEL: Standard Event List
- ✓ LEB: Late Event Bulletin
- ✓ REB: Reviewed Event Bulletin

□ Definition of Added Event

- Waveform data from IMS stations are processed automatically in the IDC system. Detections from the same event recorded at different IMS stations are associated together for event location, and those location results that meet Event Definition Criteria in the final stage of data automatic processing are saved as SEL3 events. Then IDC analysts work interactively based on SEL3 events as input to reject bogus events, modify real events, add events missed in SEL3, etc. The results is LEB, that includes events added during interactive analysis. These events are named Added Events.

□ Identifying Manually Added Events

- Each event in IDC database has a unique identification number EVID. Although event parameters might be updated during processing and analysis, EVID is kept unchanged. Manually added events will be assigned a unique EVID automatically after being stored into the database. By comparing EVID of SEL3 and LEB, these manually added events can then be easily identified.

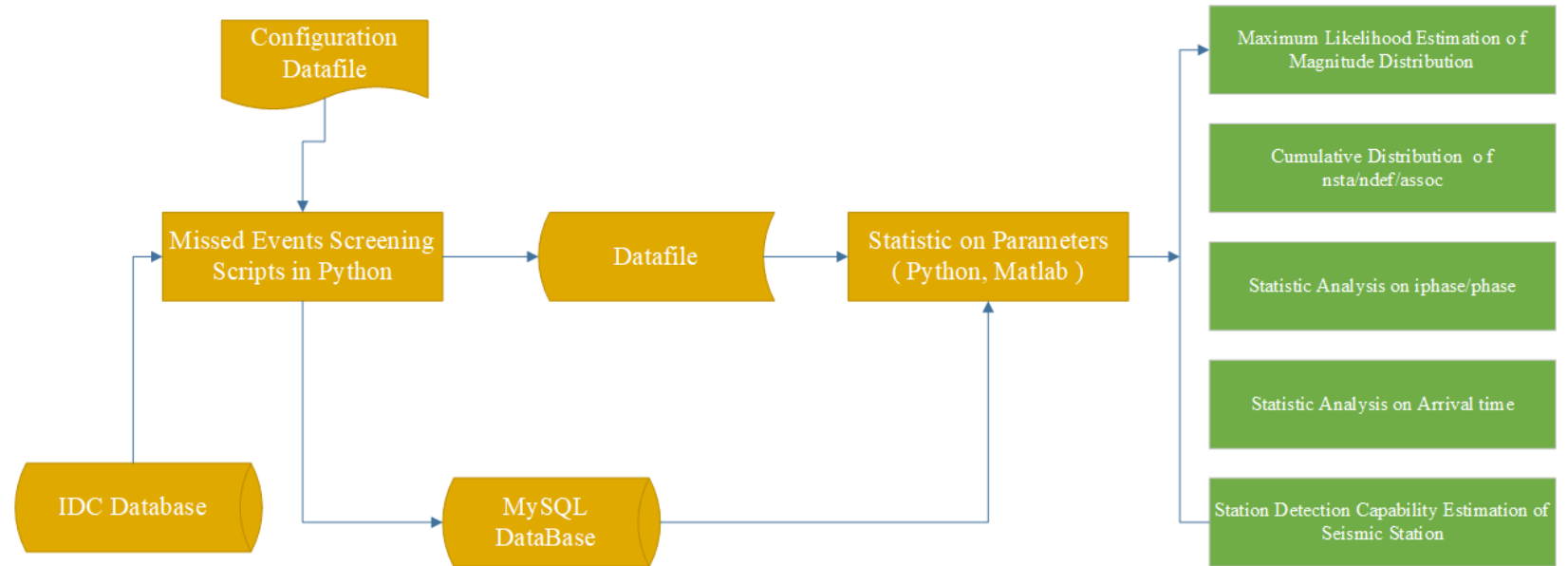
3.1 The System Developed for Parameters Analysis of Added Events

➤ Components and main functions

- Separate MySQL database for storage of added events.
- Anaconda Python platform.
- Oracle database connection.
- MySQL database connection, query and output.
- Identifying added events.
- Parameters statistic and graph.
- Estimation of seismic station detection capability.

➤ Programme language

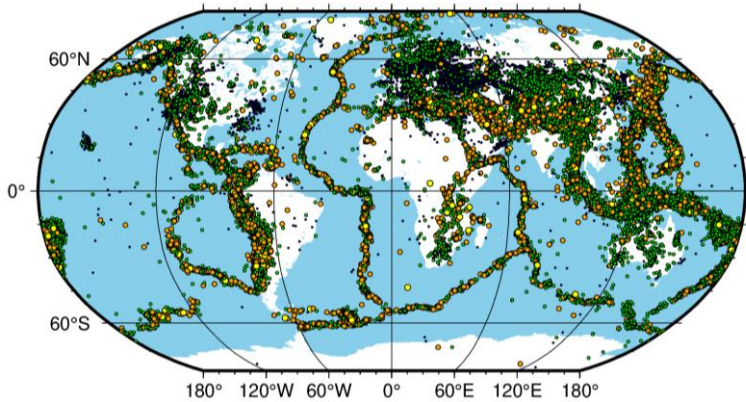
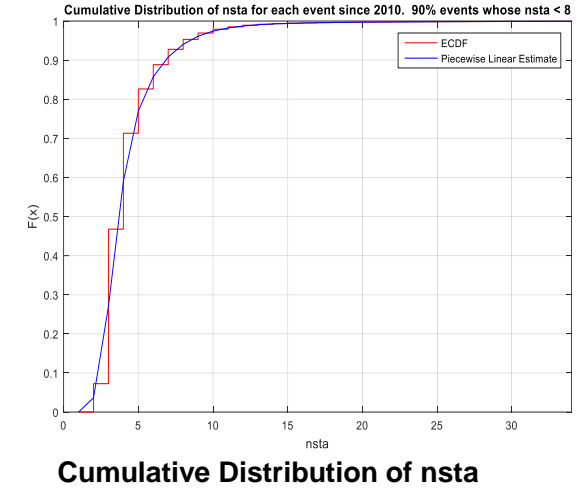
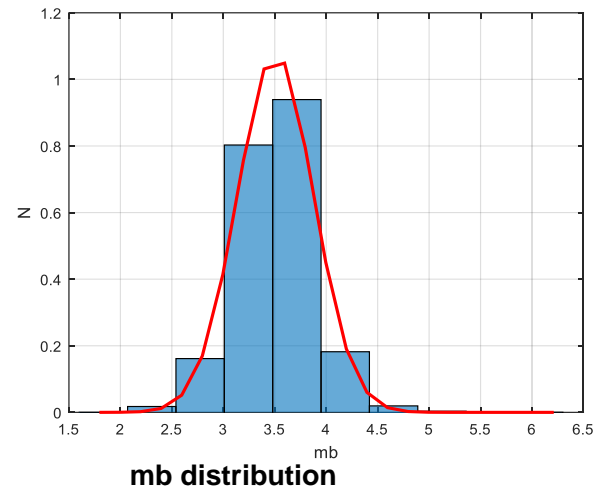
- Python, Matlab, SQL



The system developed for statistical analysis of added events in IDC database

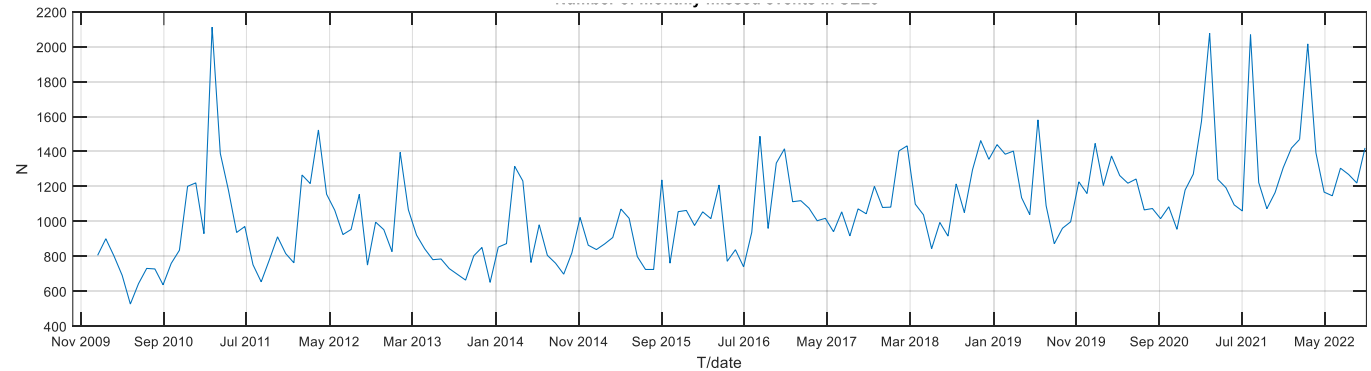
3.2 Statistics and Parameters Analysis

- Events in IDC SEL3 and IDC LEB from January 2010 to November 2022 were selected for this study. 167966 added events were found by comparing EVIDs in both bulletins, which represent 27% of LEB.
- Magnitude (mb) distribution with the center magnitude around mb=3.5.
- The number of contributing stations is, for 90% of added events, less than 8. Most of added events are small events.
- Statistically the daily number of added events is around 36.



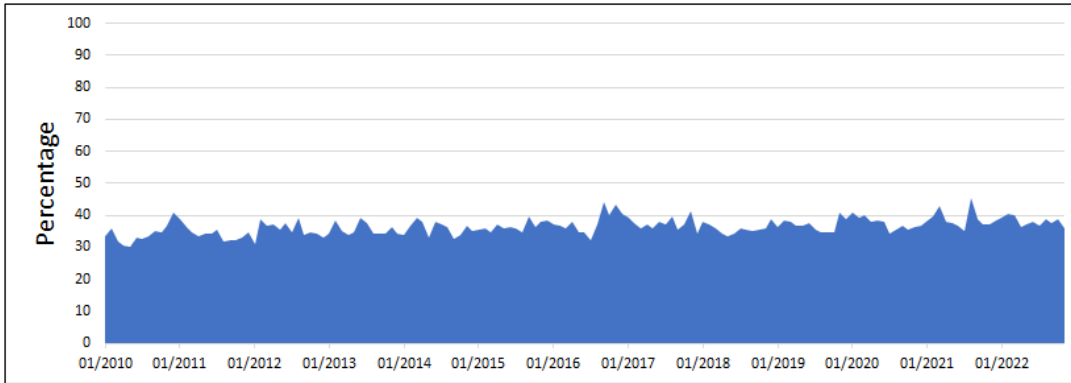
Added events since Jan. 2010 to Nov.2022.

- nsta<3
- 3<nsta<6
- 6<nsta<12
- nsta>12

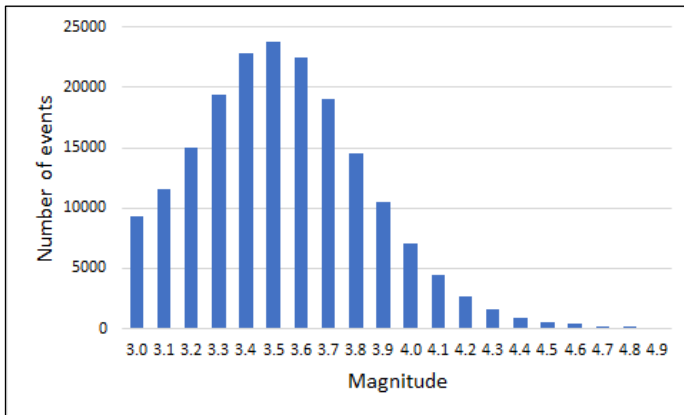


Number of added events over time (monthly aggregated)

3.3 Comparing bulletins SEL3 and LEB

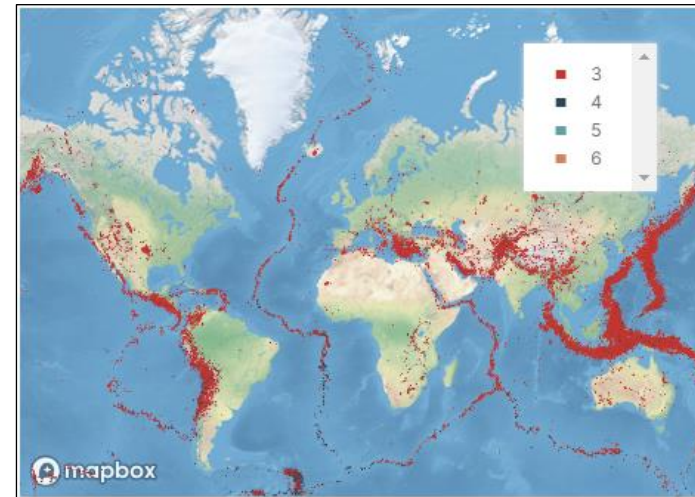


Evolution of the percentage of extra events over time



Absolute distribution of the extra events per magnitude

- **BulCMP** : A software package to compare two different bulletins (reference and assessed), using probabilistic approach. It is used to associate events present in both bulletins (**matched**), identify events present only in the reference bulletin (**missed**) and events that are present only in the assessed bulletin (**extra**).
- Using SEL3 as a reference and LEB as assessed, and running BulCMP for the time period selected for this study, the number of extra events (**actually events added in the LEB by the IDC analysts**) represents the 36% of number of LEB events.
- The majority of these events added by the analysts are small-magnitude events, and its magnitude distribution is centered around $m_b=3.5$.
- These findings are consistent with the results obtained by comparing EVID.



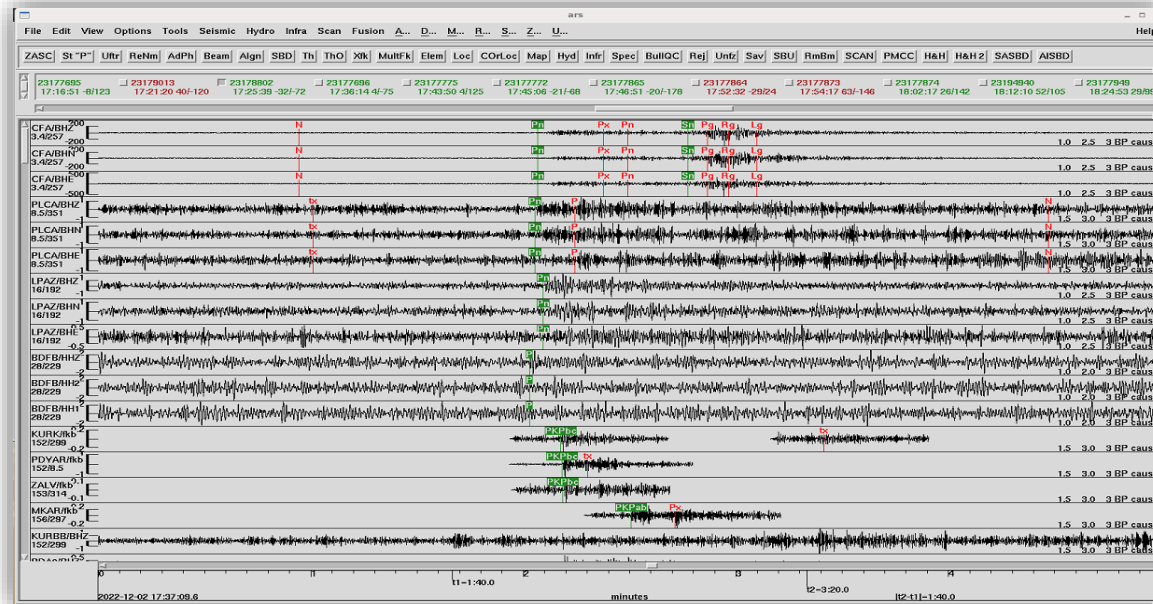
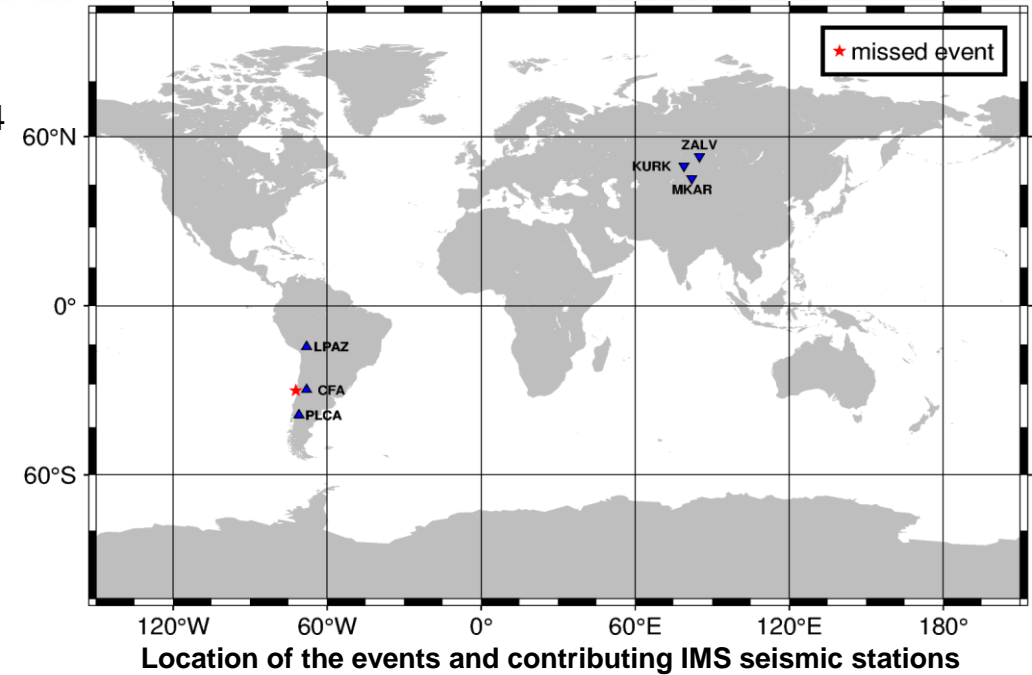
Map of the location of extra events

3.4 In-depth Analysis of a particular Example

The focal parameters of this event in REB are: Lat -32.30, Lon -72.19, z=0, mb=4.4

- This event was not produced in IDC SEL3.
- Event was detected by 8 IMS seismic stations.
- First arrivals are emergent with lower signal to noise ratio at nearby stations, while detections at teleseismic stations have high signal to noise ratio.

An estimation of the detection probability of this particular event was done using six contributing seismic stations.



Associated detections of this event as recorded in IMS seismic stations

Date	Time	Err	RMS	Latitude	Longitude	Smaj	Smin	Az	Depth	Err	Ndef	Nsta	Gap	mdist	Mdist	Qual	Author
2022/12/02	17:25:39.07	1.17	0.99	-32.3091	-72.1946	27.2	21.1	155	0.0f		9	8	183	3.44	155.72	m	uk IDC_LEB
Magnitude	Err	Nsta	Author	OrigID													
ML	3.7	0.1	3	IDC_LEB	23194553												
mb	4.4	0.3	1	IDC_LEB	23194553												
mbtmp	4.0	0.2	4	IDC_LEB	23194553												
Sta	Dist	EvAz	Phase	Time	TRes	Azim	AzRes	Slow	SRes	Def	SNR	Amp	Per	Qual	Magnitude	ArrID	
CFA	3.44	79.2	Sn	17:27:16.825	0.6	144.3	-113	19.2	-5.5	T	7.3	13.7	0.33	a		175615266	
H0351	5.77	253.1	T	17:33:47.047	48.2	82.9	6.2	75.1								175614351	
H03N1	5.78	256.8	T	17:33:36.880	44.1	86.6	6.2	73.6								175614349	
H03S2	5.79	253.0	T	17:33:34.605	35.8	82.9	6.3	75.1								175614355	
H03S3	5.79	253.2	T	17:33:43.820	45.0	82.9	6.1	75.1								175614353	
H03N2	5.80	256.9	T	17:33:41.431	48.6	175.2	94.6	73.6								175614350	
PLCA	0.51	171.5	Pn	17:27:43.450	0.4	324.4	-26.1	5.0	-8.7	T	2.1	0.1	0.33	a		175715177	
LPAZ	16.37	14.0	Pn	17:29:33.000	2.1	178.4	-13.8	5.6	-7.2	T	6.4	0.1	0.33	a		175614121	
BDFB	27.53	58.6	P	17:31:25.840	-1.9	222.3	-6.4	15.0	6.1	T	3.6	1.2	0.62	a		175617130	
H1152	125.65	270.4	T	20:03:43.019	58.9	117.7	0.8	75.2			7.1	2.6	1.03	a		175617128	
H1153	125.66	270.4	T	20:03:46.259	62.1	117.7	0.8	75.2			7.6			a		175617132	
H1151	125.66	270.5	T	20:03:46.371	62.2	117.7	0.8	75.2			10.0			a		175617130	
KURK	151.71	41.1	PKPbc	17:45:35.050	0.1	299.9	0.7	3.1	0.7	T	6.1	0.7	0.68	a		175615302	
PDYAR	152.47	354.9	PKPbc	17:45:36.300	-0.2	351.8	-16.8	2.0	-0.3	T	12.7	0.8	0.45	a		175614485	
ZALV	152.85	30.4	PKPbc	17:45:37.300	-0.2	299.1	-14.5	1.5	-0.8	T	4.5	0.6	0.43	a		175614510	
MKAR	155.72	46.0	PKPbc	17:46:00.800	0.1	291.8	-5.7	3.9	-0.4	T	4.3	0.1	0.35	a		175614478	

IDC REB solution for this particular event

3.4 In-depth Analysis of a particular Example

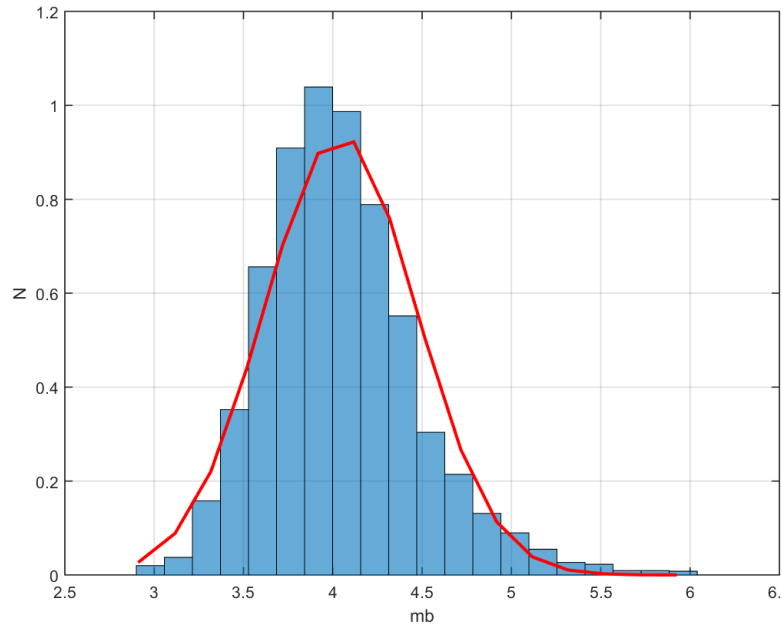
Catalogue of events in the area where the event was located

- 4845 IDC LEB were selected from January 2010 to November 2022 with distance of 2.6 degree to the added event. Magnitude of those IDC LEB events is from mb2.9 to mb6.0, average magnitude is around mb4.0

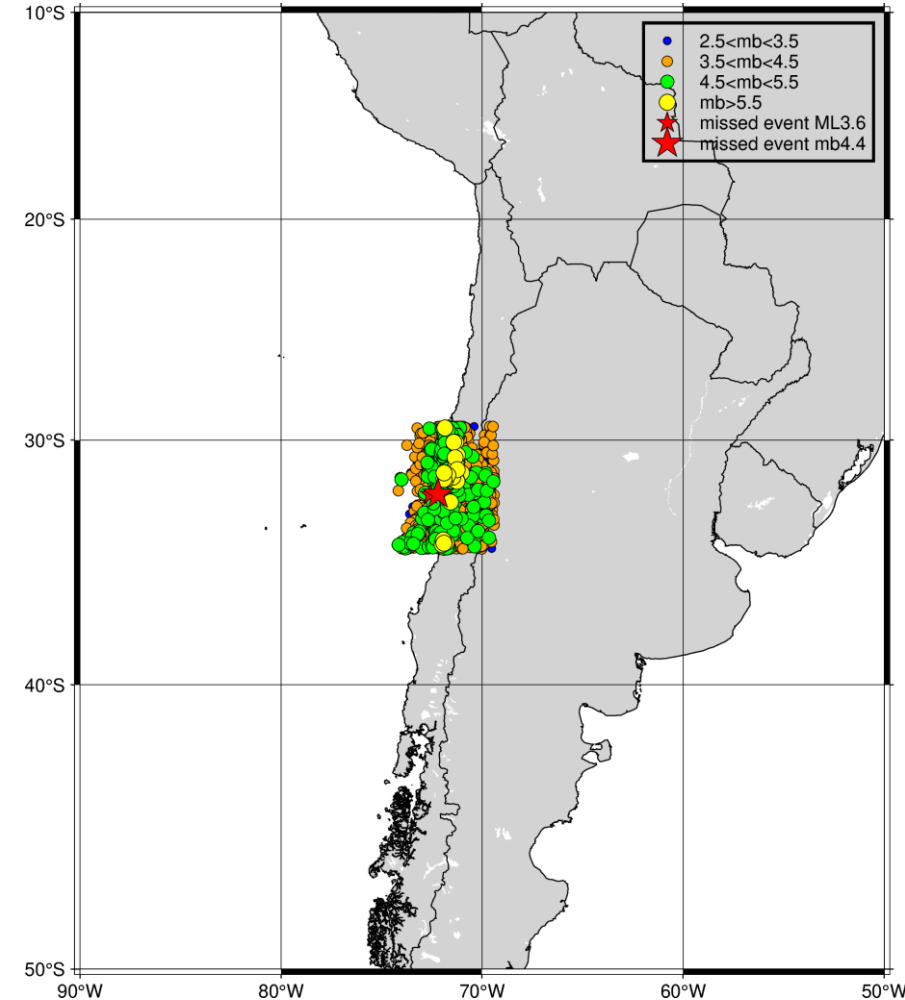
Event detection probabilities

$$P_D(M, L) = \frac{d}{d + n}$$

- d the number of events detected on station for events of magnitude M .
- n the number of events undetected on station for events of magnitude M .
- M Magnitude.
- L Epicentre distance.



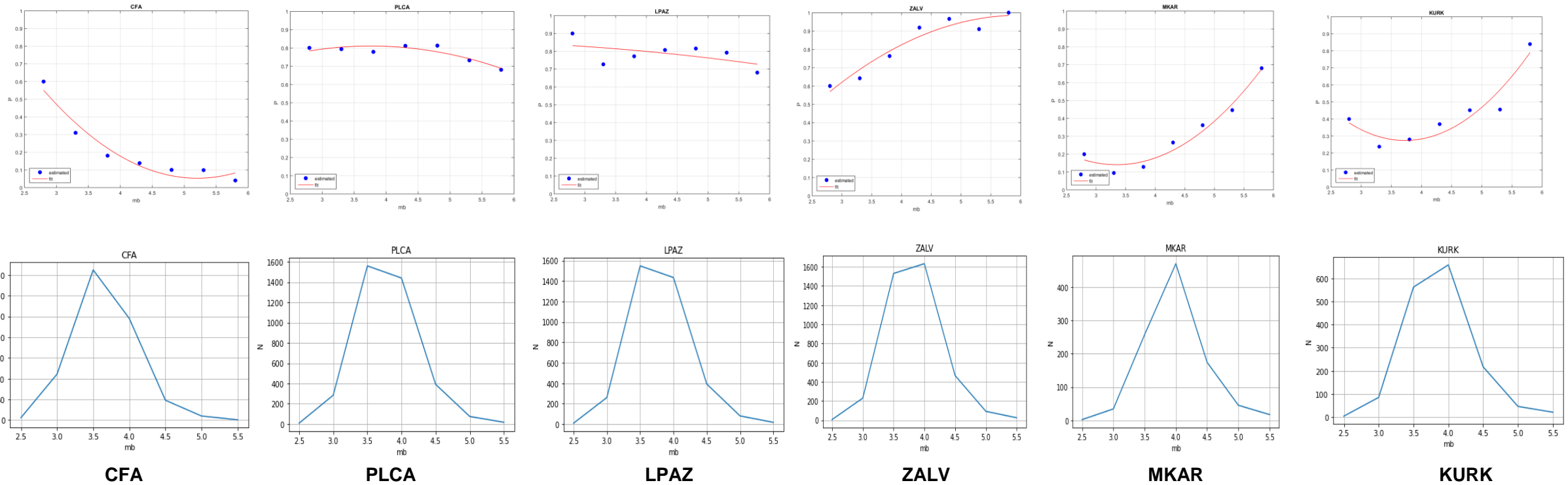
Magnitude distribution of IDC LEB events used for estimation of detection capability



Location of IDC LEB events used for estimation of the detection capability

3.4 In-depth Analysis of a particular Example

- Detection probabilities estimated for CFA, PLCA, LPAZ, MKAR, ZALV, KURK were shown below. Those stations with highest detection probability are three primary stations **PLCA, LPAZ, and ZALV**, for which the value of detection probability is above 60%. CFA the closest IMS station to the added event, as well as MKAR, KURK have lower detection probabilities.



3.4 In-depth Analysis of a particular Example

- Event information in SEL1. This event was built in SEL1, but disappeared in SEL2 and SEL3. Associated stations in SEL1 is LPAZ, PLCA, ZALV, which is consistent to the analysed result of detection probability.

ORID	TIME	LAT	LON	STA	ARTIME	IPHASE
23177692	1670001946	-32.092037	-73.648169	LPAZ	221202 17:29:37	P
23177692	1670001946	-32.092037	-73.648169	PLCA	221202 17:27:55	P
23177692	1670001946	-32.092037	-73.648169	PDYAR	221202 17:45:36	P
23177692	1670001946	-32.092037	-73.648169	PDYAR	221202 17:45:43	tx
23177692	1670001946	-32.092037	-73.648169	ZALV	221202 17:45:38	P

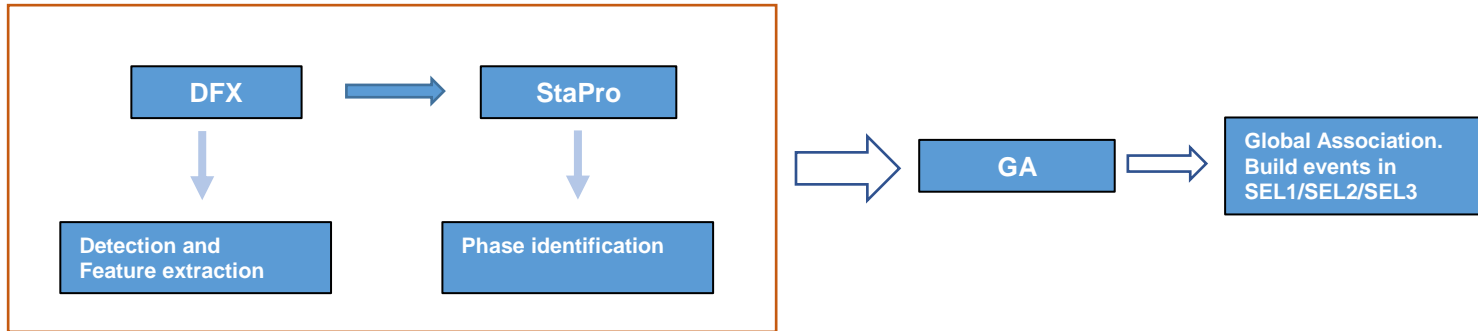
Event information in SEL1

- Arrival information. This event was detected on associated stations below, phase name of detection given in automatic processing system is not correct.

STA	DELTA	IPHASE	PHASE	AZIMUTH	AZRES	SLOW	SLORES	TIMERES
CFA	3.4400914	Pn	Pn	249.14415	-8.0090922	12.399997	-1.3496234	.51332125
PLCA	8.5101743	P	Pn	324.40564	-26.14703	5.0359	-8.680504	.37471732
LPAZ	16.37179	P	Pn	178.44967	-13.842955	5.5805555	-7.2141938	2.1312637
BDFB	27.528759	N	P	222.27369	-6.354456	15.02527	6.062388	-1.882708
KURK	151.70836	P	PKPbc	299.85443	.74212886	3.1092833	.70736703	.14973715
PDYAR	152.47328	P	PKPbc	351.76095	-16.752959	2.032932	-.29793386	-.16056869
ZALV	152.8536	P	PKPbc	299.0546	-14.496026	1.5264291	-.77151616	-.18316341
MKAR	155.72096	Pn	PKPab	291.80141	-5.6889977	3.881131	-.41070908	.099591063

Part of detection parameters on associated stations for this example

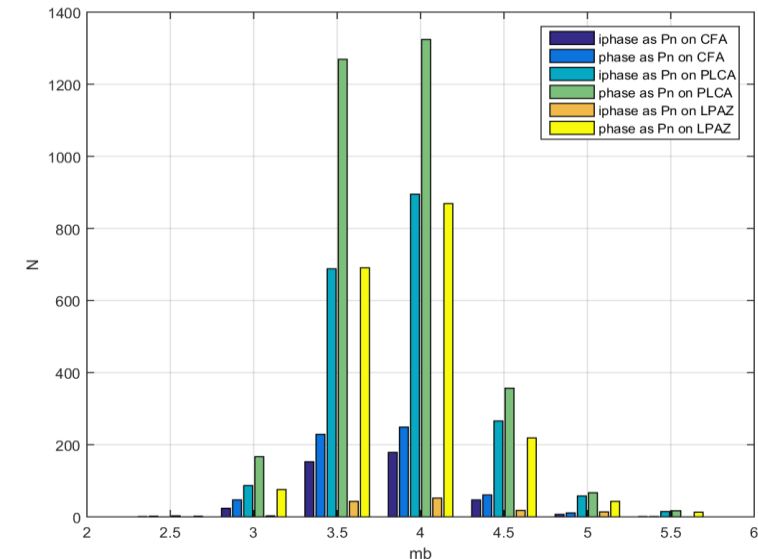
3.4 In-depth Analysis of a particular Example



- ✓ This event was detected on associated stations.
- ✓ Incorrect phase name for all detections.
- ✓ Built in SEL1, disappeared in SEL2/SEL3.

Table 1. Statistic of phase identification on CFA,PLCA,LPAZ

mb	CFA		PLCA		LPAZ	
	iphase as Pn	phase as Pn	iphase as Pn	phase as Pn	iphase as Pn	phase as Pn
2.5-3.0	1	2	0	3	0	2
3.0-3.5	24	47	87	167	3	76
3.5-4.0	153	229	688	1269	43	691
4.0-4.5	179	249	895	1324	52	869
4.5-5.0	47	61	266	357	18	219
5.0-5.5	7	11	58	67	14	43
5.5-6.0	1	1	15	17	0	13
Total	412	600	2009	3204	130	1913
Ratio(iphase/phase)	0.69		0.63		0.07	

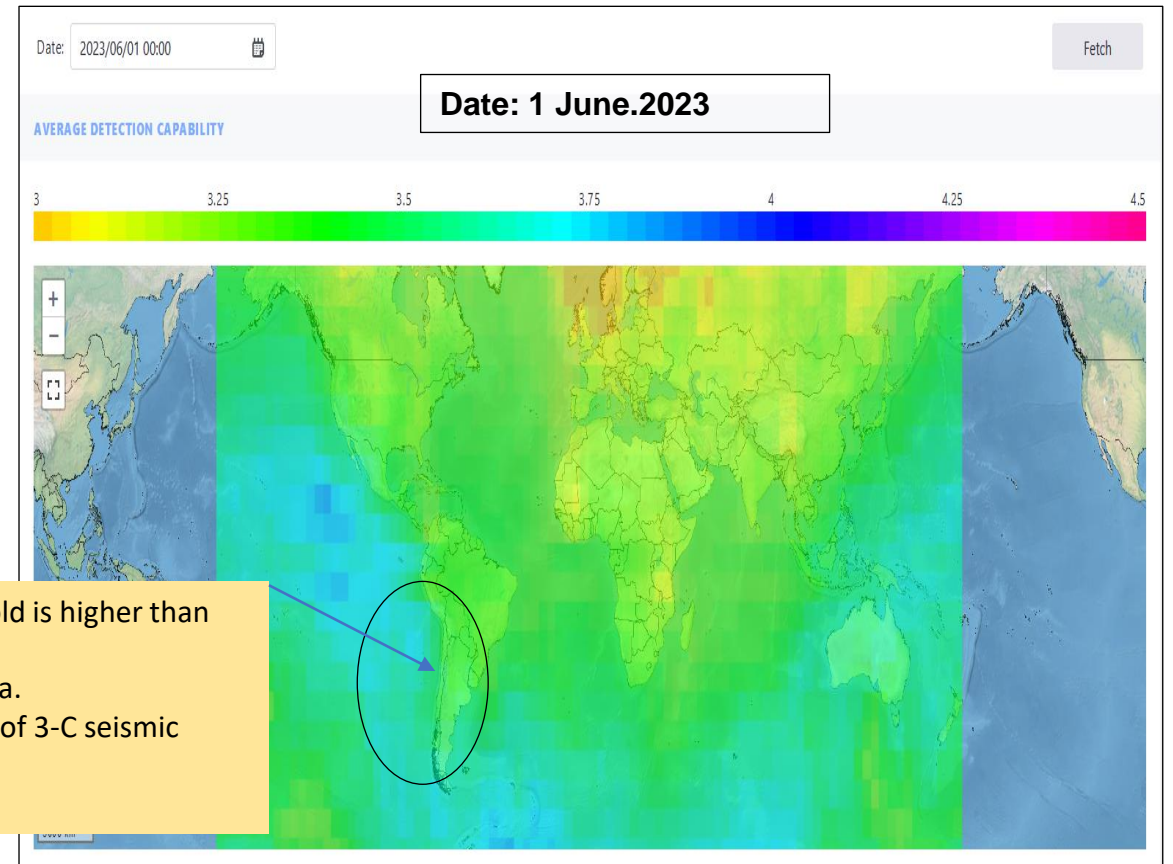
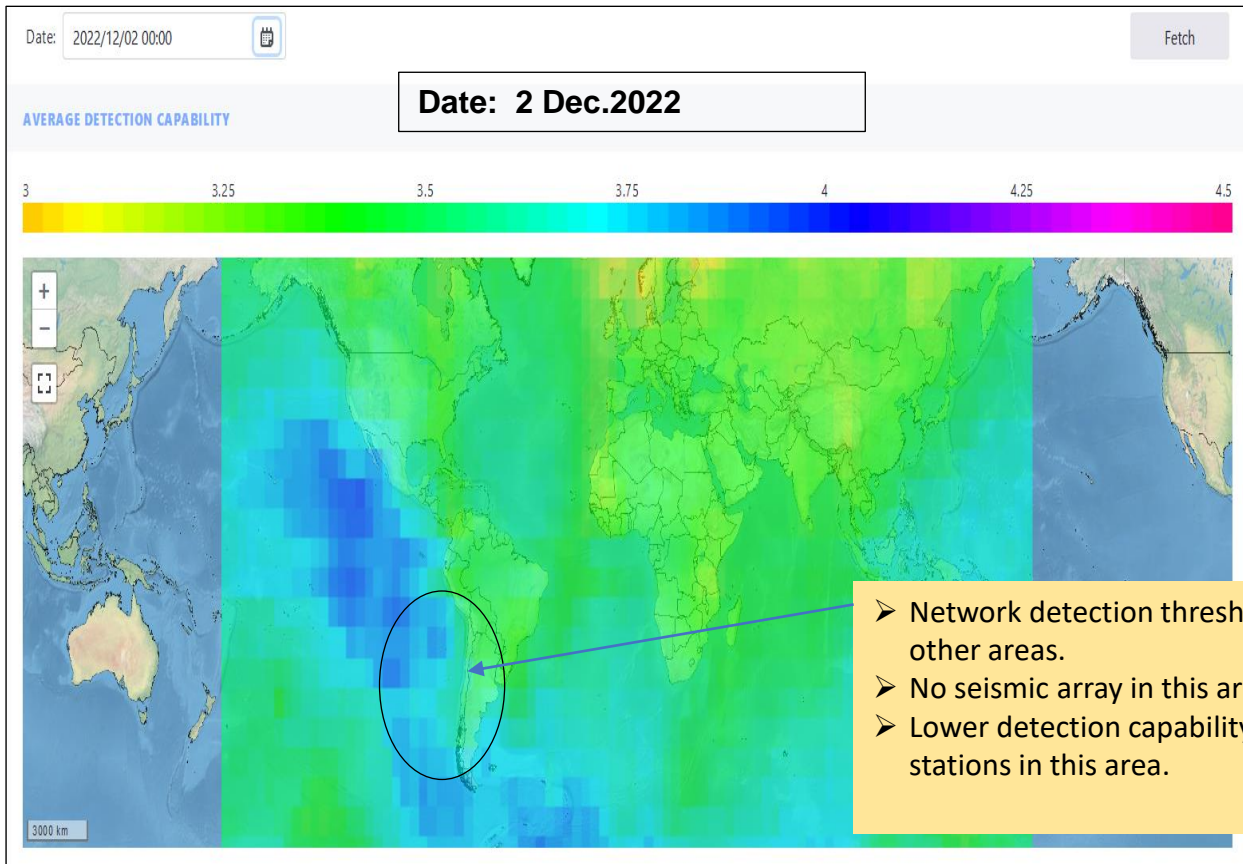


Statistic of phase identification

- iphase -- name of arrival in IDC SEL3. phase -- name of arrival in IDC LEB

3.5 Network detection threshold estimation

✓ Reference:
Threshold Monitoring Subsystem/IDC Document



Network detection threshold estimation

4. Discussion and Conclusion (1/2)

- The added event rate of the IDC interactive waveform analysis is approximately 27% of LEB, with an average of 36 added events per day.
 - The distribution of magnitude of the added seismic events is centered around $m_b=3.5$
 - 50% of the added events have less than 4 contributing stations
 - Most of the added events are infrasound events.

- Comparing SEL3 and LEB bulletins by using BulCMP results in a percentage of added events around 36% of LEB.
 - The distribution of magnitudes is centered around $m_b=3.5$, which is consistent to that obtained by checking the EVID.

- Obviously, all those manually added events are a contribution from IDC analysts, to maintain high quality of the IDC reviewed bulletin.

4. Discussion and Conclusion (2/2)

- An in-depth analysis of one particular manually added event of magnitude $m_b=4.4$ located in the South American region has been done
 - Eight IMS seismic stations (including 6 primary seismic) and one hydroacoustic detected the event, with a network GAP of 183° .
 - The phase associated detections on each station were generated by the automatic processing.
 - Although 8 seismic stations detected this event, only one station has the correct phase name identified by DFX.

- An automatic solution for this event is present in SEL1
 - Four primary seismic contributing stations (PLCA, LPAZ, PDYAR, ZALV) and magnitude $M_L=3.8$. The network GAP in SEL1 is 187° .

- No automatic solution is present in SEL2 and SEL3.

- The network detection threshold of the epicentral area where the event is located is significantly lower than in other regions worldwide. This is probably due to a lack of seismic arrays with good event detection capability in that area. The detection capability of the three-component seismic stations in this area is not as high as with arrays would be.

Thank you.