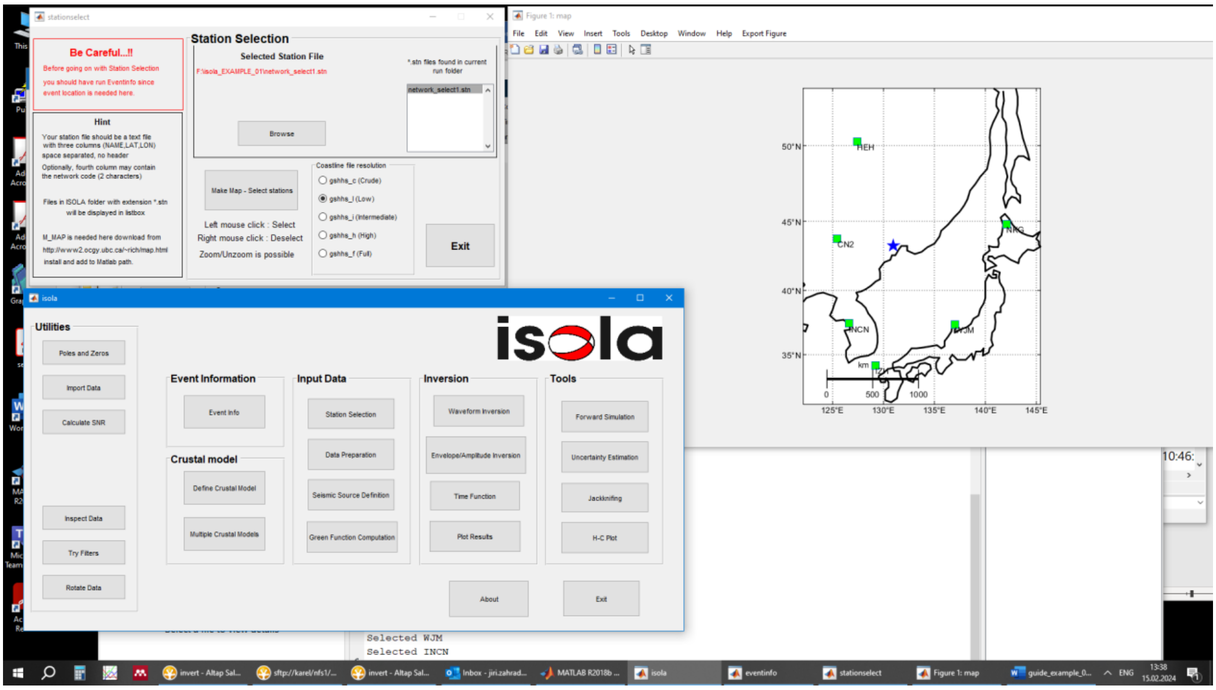
Example 01
Introductory; to be studied in the beginning.
China deep-focus Mw 6.9 earthquake.
Non-DC components.

It is assumed that Isola2024 is already installed on your computer (the same procedure as in previous versions, e.g. Isola2022, or older, https://geo.mff.cuni.cz/~jz/for_Brasilia2020/). Isola2024 is a version with more sophisticated covariance matrices (‘cova’) believed to provide more reliable moment tensors and better uncertainty assessment.

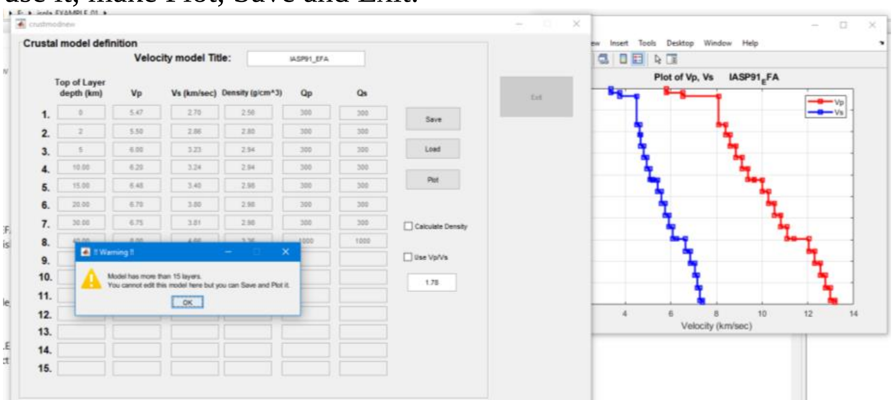
Save the content of isola_EXAMPLE_01 in a safe place. It will be rewritten during your replication tests. The spare copy will enable a comparison of the former and replicated results.

In Matlab, go to isola_EXAMPLE_01; this is now your working folder.
Type ‘isola’, and the main menu of Isola GUI appears.
Make yourself the below-listed actions, previously taken by the authors of the example. You should obtain the same results as in the stored copy. Later you may try to change some options to have your own results.

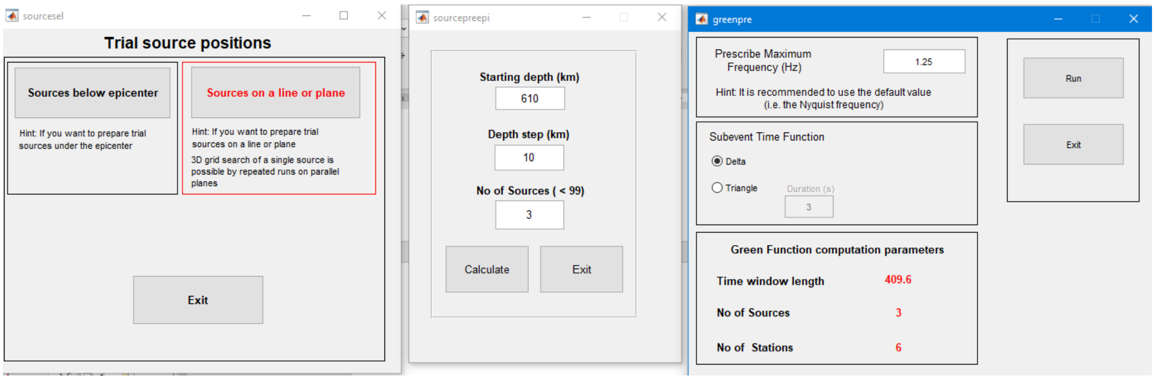
Preparation.
We refer to the individual tools of the main menu of the Isola GUI.
<u>Event Info</u>. The selected event is: 20110510, Lon 130.957, Lat 43.282; the time window length should be TWL = 409.6 s. Save. It implies waveform sampling with dt=TWL/1024=0.4 s. Data are saved in files event.isl and duration.isl.
<u>Station Selection</u>. Shown are the used stations, previously defined in network_select1.stn (the file in the root folder). Select all, and exit. The stations appear in allstat.dat file in invert folder.



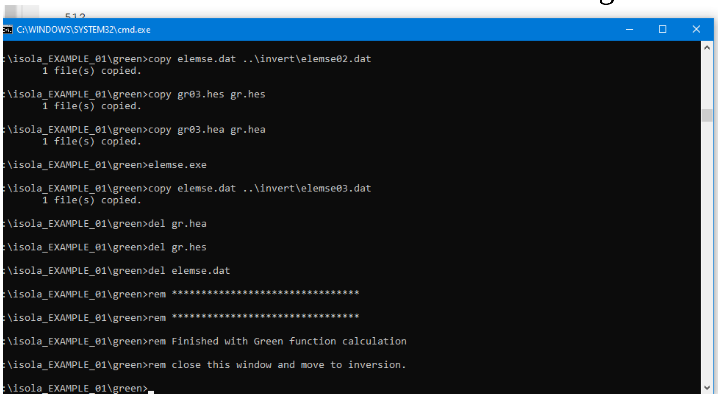
[Data Preparation]. No need to do anything; the instrumentally corrected waveforms *raw.dat are prepared in the invert folder. These files contain ground motion velocity in a 4-column format (time, N, E, Z). Time 0 is the origin time. For the sampling step, dt, see above in EventInfo. The same files have been also stored (as backup) in folder data_Isola_raw.
Define Crustal Model. Load model from file crustal_iasp_EFA.cru from the root. A warning appears that the file cannot be shown numerically (more than 15 layers). Despite it, you can use it, make Plot, Save and Exit.



Seismic Source Definition. Three trial positions of this example have been chosen below the epicenter, starting at depth 610, with a depth step of 10 km. Calculate and exit. The trial sources now appear as src*.dat files in the folder green.
Green function calculation. GF is calculated up to (default) Nyquist frequency 1.25 Hz, assuming delta-function moment-rate (default); 3 trial sources, 6 stations. Run. The calculation can be monitored in the system command window ('black window'). Results are in green folder (gr*hea, gr*hes) and the invert folder (elemse*).

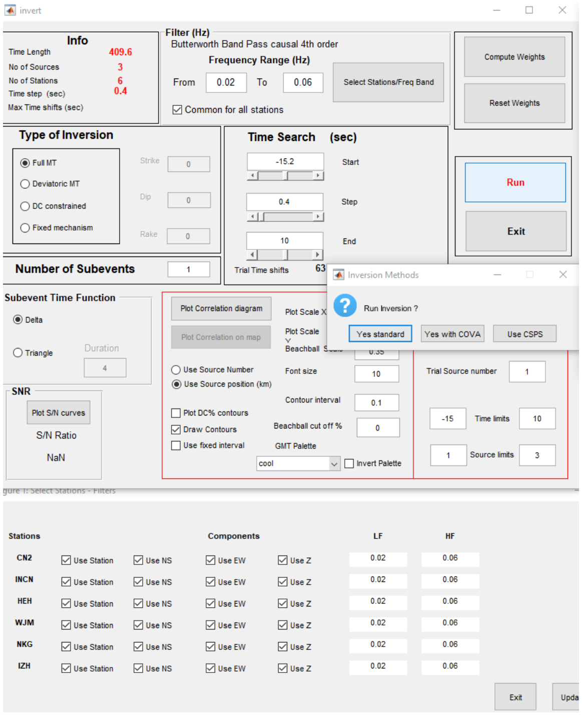


The GF calculation should finish with a message like this:



Inversion no-cova (iteration 0).

Waveform Inversion tool. Frequency range 0.02-0.06 Hz, common for all stations. Check this option in SelectStation/FreqBand button. Ignore the (infrequently needed) buttons Compute Weights and Reset Weights. Chose the Type of inversion: Full MT. Time search: start at -15 s, step 0.4 s (=dt), end at 10 s.

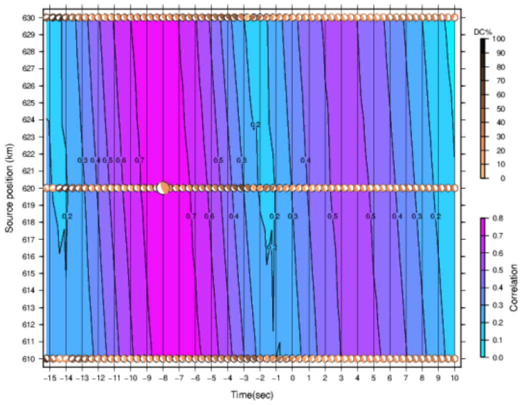


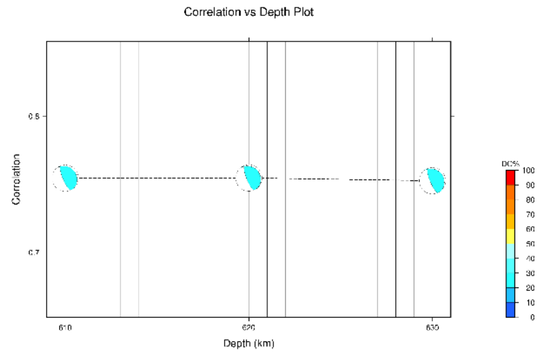
Run: select the option Standard (without covariance matrix). In your version, this left button may have already a different name. It runs Fortran code isola_nocova.exe. The output files are in the invert folder. The richly commented file is inv1.dat. In this example, we are mainly interested in the non-double-couple (non-DC) components, in particular the low VOL (=ISO) and large (negative) CLVD components of moment tensor (MT).

```
moment (Nm):  2.2317415E+17
   moment magnitude:   5.5
             VOL % :  11.8
              DC % :  24.2
            CLVD % : -64.0
 strike,dip,rake:       158        75        96
 strike,dip,rake:       314        15        67
varred=  0.5694842    (= VR, variance reduction)
```

Plotting: In the invert window it is useful to select Plot Correlation diagram [with the option Use Source Position (km)]. The plot appears in the folder output as *corr01.png.

Plot Results tool can be used to produce other plots. Try Plot Correlation vs. Source, and Plot Solution Summary, these are automatically saved as *inv1.png and *best.png files in the folder output.





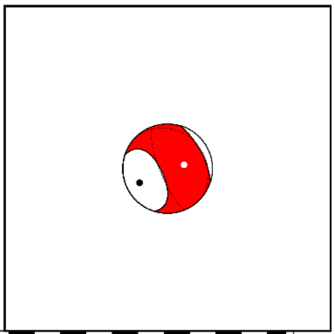
MOMENT TENSOR SOLUTION

HYPOCENTER LOCATION (WOXIN)

Origin time 20110510 15:26:08.80
Lat 43.282 Lon 130.957 Depth 564.5

CENTROID

Trial source number : 2 (Fixed Epicenter inversion)
Centroid Lat (N)43.282 Lon (E)130.957
Centroid Depth (km) : 620
Centroid time : -8 (sec) relative to origin time



Moment (Nm) : 2.232e+17 Mw : 5.53
Inversion Type:Full
VOL% :11.8
DC% :24.2
CLVD% :-64

	SNR	CN	FMVAR	STVAR
Var.red.:(for stations used in inversion):0.57	NaN	5.3	4±3	0.14
Var.red.(for all stations) :				

Strike	Dip	Rake
158	75	96
Strike	**Dip**	**Rake**
315	16	67

P-axis	Azimuth	Plunge
	243	30
T-axis	**Azimuth**	**Plunge**
	77	59

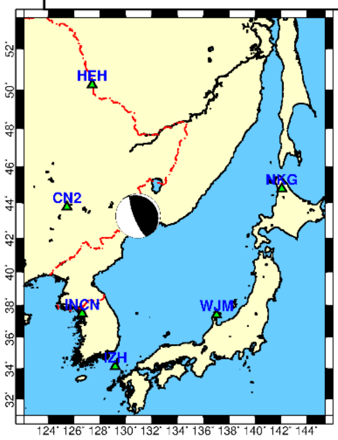
Frequency band used in inversion (Hz)
0.02 - 0.06

Stations-Components Used-Distance

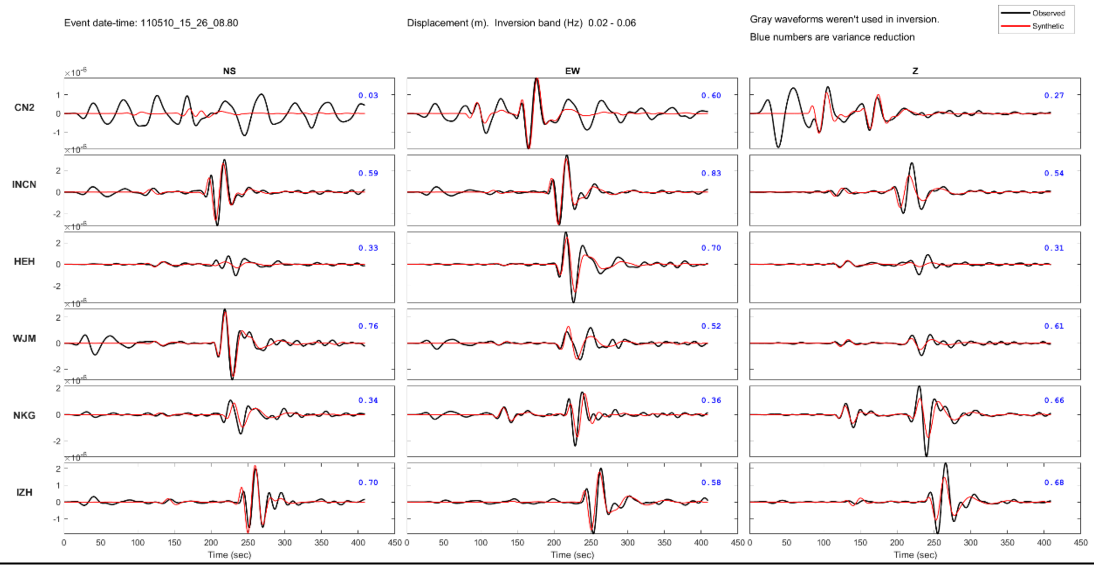
	NS	EW	Z	D(km)
CN2	+	+	+	448
INCN	+	+	+	742
HEH	+	+	+	821
WJM	+	+	+	831
NKG	+	+	+	907
IZH	+	+	+	1027

Mrr	Mtt	Mpp
82.029	66.384	-68.435
Mrt	**Mrp**	**Mtp**
74.460	-162.679	99.441

Exponent (Nm): 15



Plot waveforms using the option Plot Real-Synthetics, while not using the offered normalization. The N, E, and Z components are shown from the left to the right. When the plot appears, go to its top menu and select Export Figure/Convert to PNG format. Waveforms (= filtered displacements) are saved as *wave.png in folder output. For this run, the results have been copied also in a spare sub-folder of invert 10_no-cova_0.02-0.06.



Inversion with cova (iteration 1); output files in folder invert will be re-written!
Waveform Inversion. Keep the options as above, except the option under Run. In Run, select the inversion with covariance matrix; three types are available, select the cova matrix called Toeplitz; in previous texts called 'FULL'. This step must be preceded by inversion 'no-cova', so a warning appears. It is because the data covariance matrix is calculated from residual waveforms, i.e. the difference between observed waveforms and synthetic waveforms from the no-cova run, i.e., res = obs-syn(0), where 0 stands for the 0-th iteration.
Results appear in folder invert, and, for your convenience, have been copied into a subfolder of the folder invert: 11_cova_fullNNEEZZ_0.02-0.06. The most important result is again the file inv1.dat. You can see that centroid position, centroid time, strike/dip/rake angles, and moment are practically the same as in no-cova, **however, the VOL, DC, and CLVD are slightly different**, and ***should*** be better than in the no-cova regime, although variance reduction (VR) is typically somewhat lower in the calculation with cova. The lower VR here should be advantage! If data include error, waveforms should not be fitted perfectly because in that way we may fit the error (e.g. noise). With cove, such an 'overfit' should be prevented.

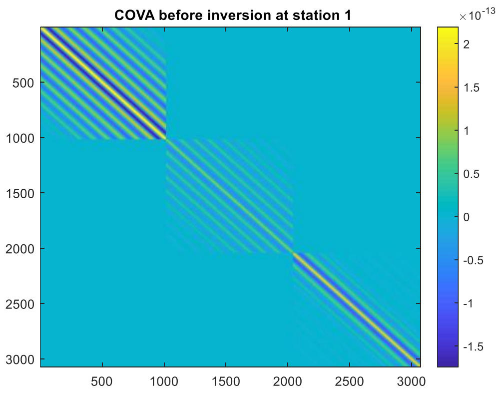
Whether the centroid position, the MT, and its uncertainty assessment obtained with cova is indeed better or not, or when it is better, is a sensitive issue, is still a topic of current research. Simply speaking, the results are better if the above "res" (residual) is a good representation of true error, comprising noise, inaccuracy of velocity model, etc. In an academic case, when the noise is white and the velocity model is reliable in the investigated frequency range, there is no need for cova; the usual Isola – e.g. Isola2022 was based on such a simplifying assumption, thus the uncertainty estimate of MT and its position was approximately controlled by station distribution.

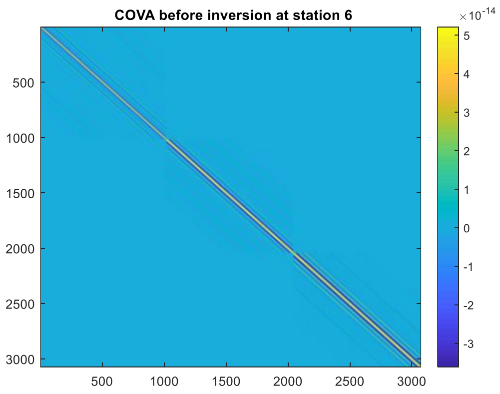
Here, with cova, we obtain (see inv1.dat):
moment (Nm): 2.2127637E+17
moment magnitude: 5.5
VOL % :   5.1 (in no-ova regime we had 11.8 %)
DC % : 30.1
CLVD % : -64.7
strike,dip,rake: 160   76   97
strike,dip,rake: 311   14   62
varred= 0.5661213 (=VR; in the no-cova regime we had 0.569 a possible "overfit")

At the end of the run, user can see Matlab command window which compares variance reduction values obtained with cova in two forms: varred_classical = 0.56612, varred_standardized = 0.69026. Both are calculated as 1 – norm(obs-syn)^2/norm(obs)^2, i.e., using L2 norm. Obs and syn and filtered displacements. VR_classical makes use of normal observed and synthetic waveforms; this value also appears above as varred in inv1.dat. VR_standardized makes use of obs and syn multiplied by the upper triangular matrix from Choleski factorization of the inverted covariance matrix (eqs. 13-15 of Vackar et al., 2017). Typically, varred_standardized should be higher than varred_classical, and below I try to explain why. Interested user can also study the corresponding code prep_NORM.m in isola_GUI_2024.

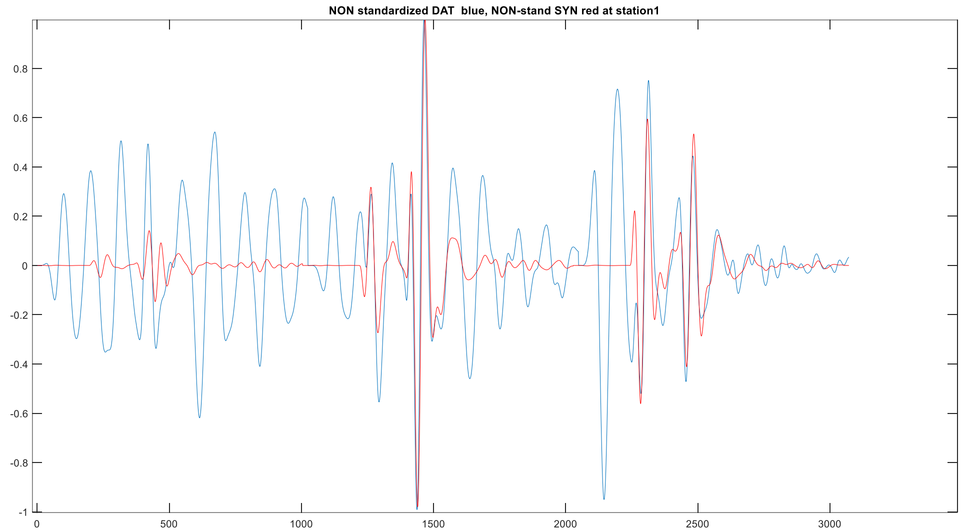
The inversion with cova produces many plots from prep_NORM.m [in later versions, their number will be probably significantly reduced, here I used the numerous plots to enable understanding of the method]. After their inspection, the plots can be easily removed by typing ‘close all’ in the Matlab command window. Examples of the plots were saved in subfolder 11_cova_fullNNEEZZ_0.02-0.06 with the following denotation: cova_st* is a covariance matrix of station *, row100_st* is the autocovariance function of data residuals (N, E, Z components) calculated without and with standardization, seis1,2,3_st* is another similar comparison in time domain, and spec_st* in spectral domain.

Related to the low-frequency noise well visible in the above-shown seismograms at station 1 (CN2), we observe the long-range correlation in cova matrix for station 1 (correlation up to large lags, large non-diagonal elements), and the less correlated noise (more diagonally dominated) matrices in station 6. From the top left to the bottom right, we plot the N, E, and Z components. Note also different amplitudes in color scales. Very roughly, the larger the diagonal Cd value, the lower is the diagonal value of the inverted matrix (Cd^{-1}, or Cdinv), implying lower weight in the inversion.

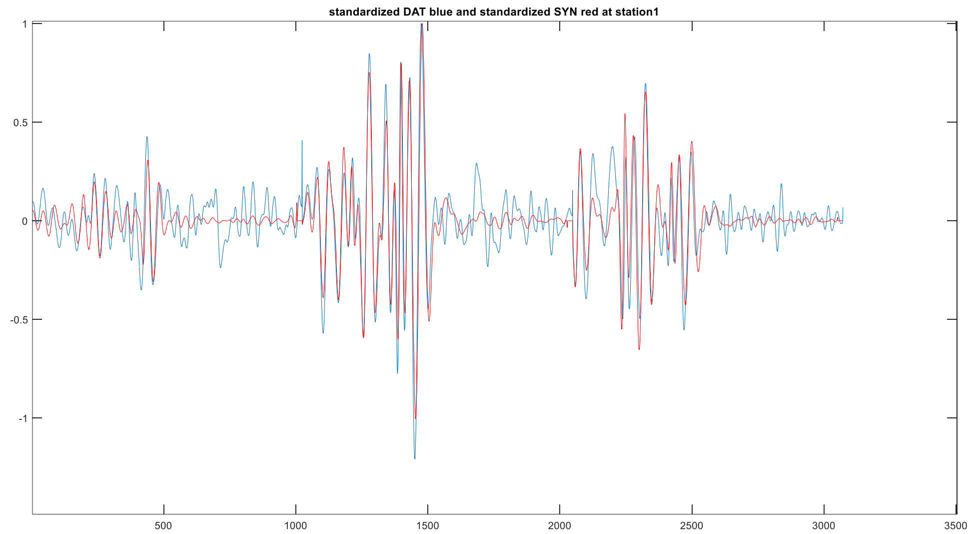




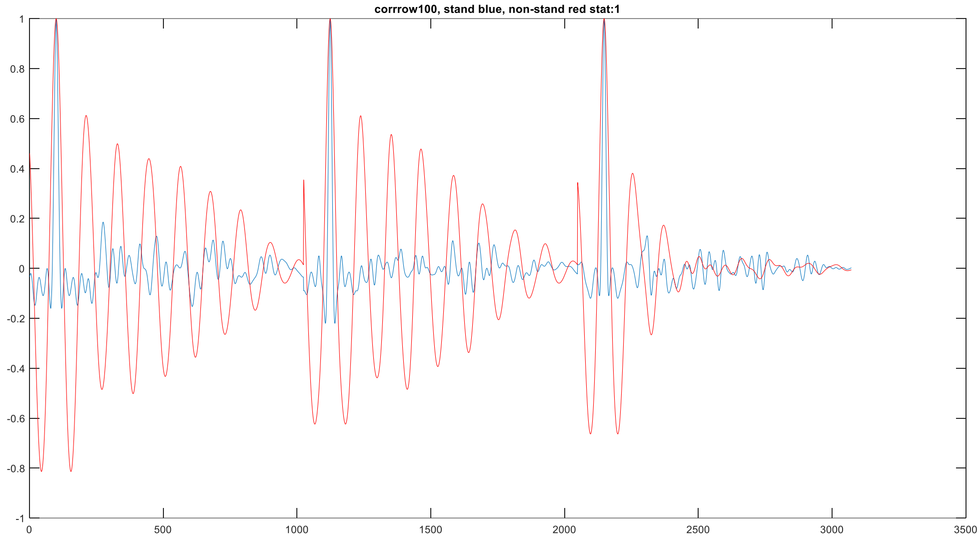
Therefore, without cova, the low-frequency noise produced a poor waveform fit in station 1 (see N, E, and Z shown from left to right); perhaps only E was somehow fitted, but the fit might have been misleading because we fitted well the signal plus the noise; this is the dangerous “overfitting” – one of the reasons why we should prefer the cova use.



The cova method implicitly, or internally improves the fit. Users cannot see it in standard (non-standardized) waveform plots! It can be only seen in so-called standardized waveforms (see the above explanation of VR_standardized), produced by prep_NORM.

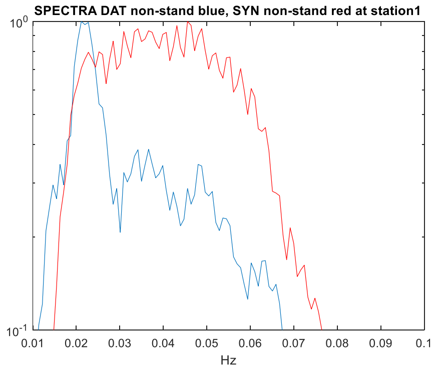


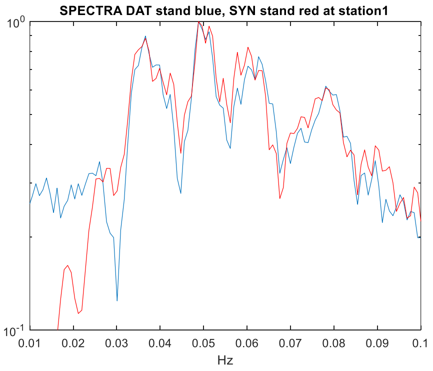
The improvement occurs because the cova application is equivalent to suppression of the long-range data correlation. With cova, i.e. with standardization, the autocovariance functions of the N, E, and Z data residuals, plotted below, are considerably narrower, or more 'delta-like'. Optimally, the uncorrelated error would have autocorrelation of the residuals in the form of a delta function.



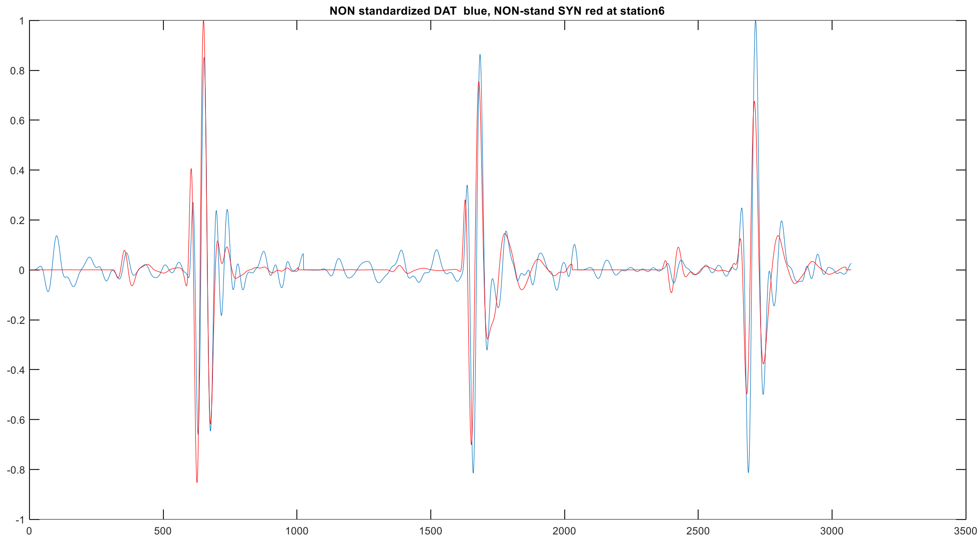
Below we demonstrate the same effect in the spectral domain for station 1 (the spectrum being calculated from all three components). Application of cova (standardization) causes "whitening" of spectra, i.e. making them closer to a constant. The cova application is equivalent to converting data error to white (uncorrelated) Gaussian noise. Since both data and synthetics are "whitened", they have a better chance of fitting each other. In the below-presented example, we can observe a considerably improved fit between standardized data and synthetics at f > 0.03 Hz.

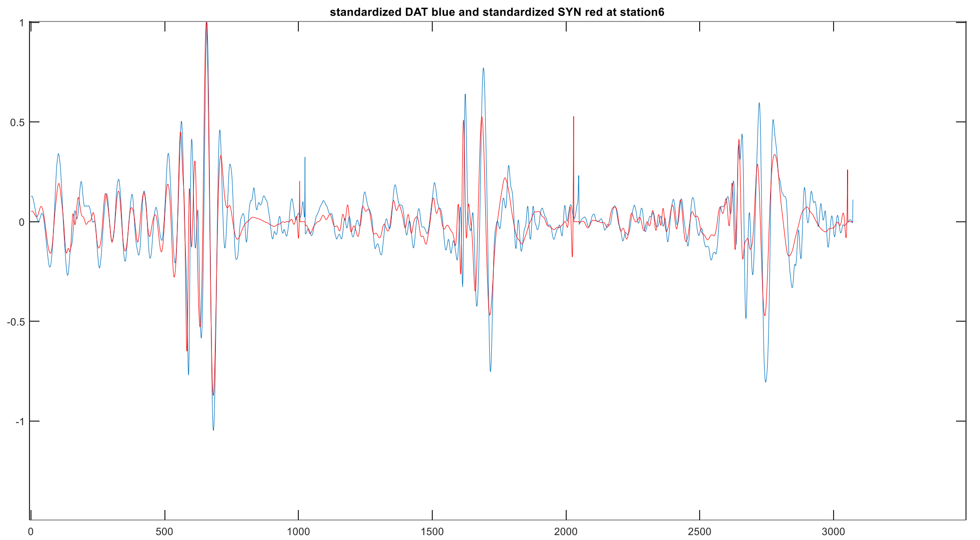
Sometimes, in literature, we read that cova tends to suppress noise. It needs to be understood with caution. The noise is not truly removed. The correlated noise at f < 0.03 Hz is *effectively* suppressed in *a relative* sense because, in this example, the high-frequency spectrum is enhanced, while the spectrum as a whole is flattened. The standardized waveforms are more broadband than the normal (non-standardized) waveforms. A "standardized" seismogram is an abstract entity that is useful just for illustration of how the inversion with cova works. We assume knowledge of real data error (represented by res= obs-syn(0)), find out how this error is correlated, and use this information to invert obs into new syn, syn(1), in such a way that the new residual res = obs-syn(1) error is less correlated (optimally uncorrelated).

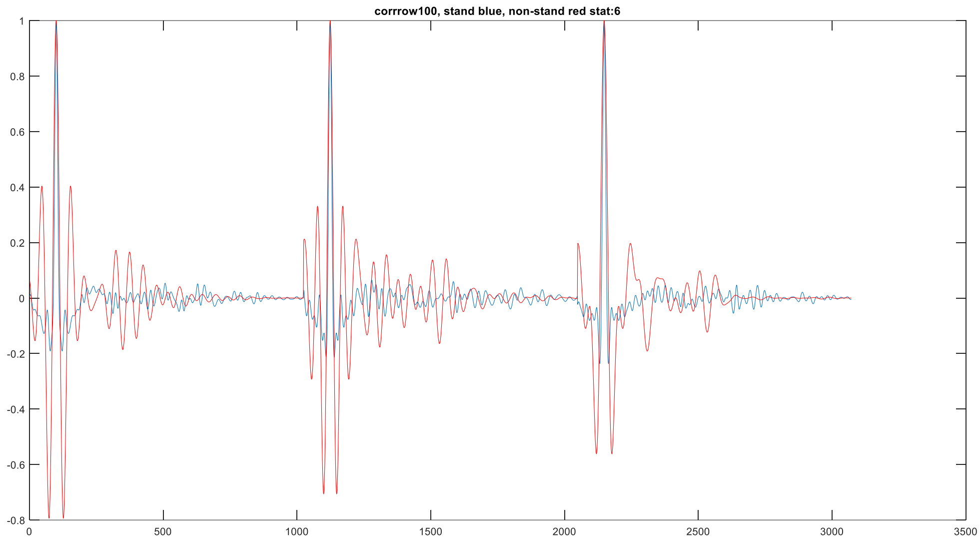


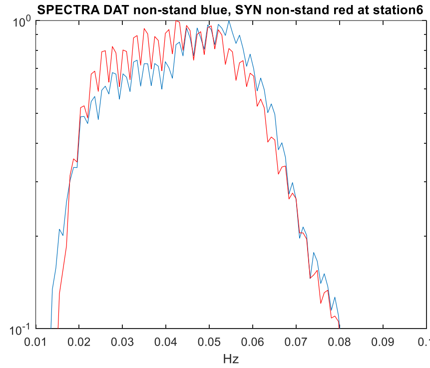


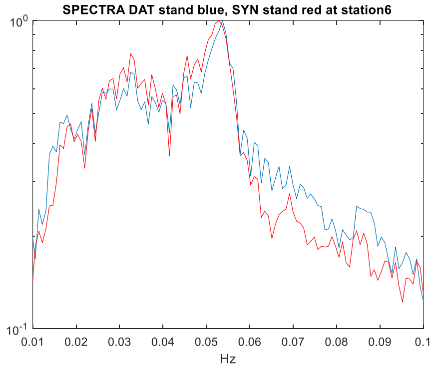
In station 6, where the low-frequency noise is much lower, we observe smaller effects than in station 1. The obs and syn waveform fit well in station 6 even without standardization because data error in this station was less correlated, closer to white noise, and the matrix was closer to diagonal:


NON standardized DAT blue, NON-stand SYN red at station6
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
500
1000
1500
2000
2500
3000
3500


standardized DAT blue and standardized SYN red at station6
1
0.5
0
-0.5
-1
0
500
1000
1500
2000
2500
3000



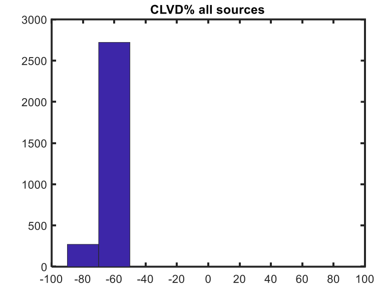


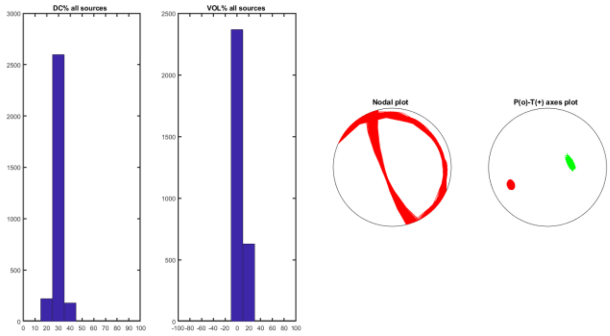


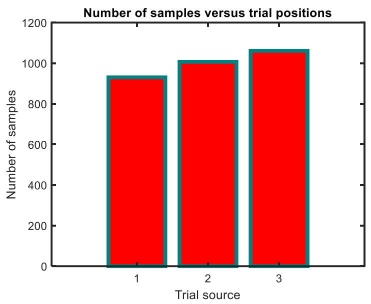
Uncertainty

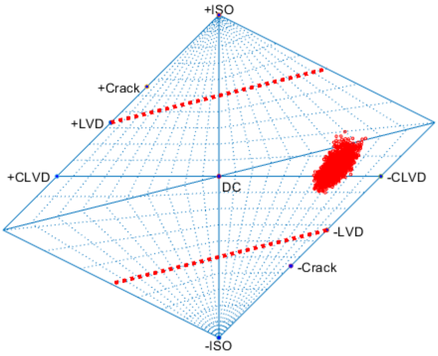
After iteration 1, we return to the main menu of Isola and choose the tool <u>Uncertainty Estimation</u>. Use the Full MT option. Press Calculate, be patient, it can take time, which is proportional to the number of trial sources; then choose Plot MT, then PDF of Trial Positions [formerly called Samples per Source], and finally the Source-Type Plot. Examples of the plots, manually saved, are shown in the spare subfolder 11_cova_fullNNEEZZ_0.02-0.06 as uncplot1-4.emf. Numeric results appear in newunc folder, e.g. files CLVD_ISO_DC_Mo_MW_eq6.dat and newunc_log.dat. In this example, the most important is just the source-type (Hudson's) plot, uncplot4.emf, visualizing the departure of the earthquake from pure DC source.

As shown below, we obtain low uncertainty of nodal planes and P/T axes, and also low uncertainty of DC, CLVD, VOL. It is because we have a relatively low noise (except station 1), good azimuthal distribution of the stations, and a satisfactory velocity model. The "red" histogram of trial source positions has a little (if any) depth preference of the largest tested source depth. The Hudson's (diamond) source-type plot indicates a significant deviation from the 100% DC (The DC solution would lie in the center of the plot), i.e., we obtain a large (negative) CLVD. The VOL component is small or zero (the zero VOL is in the horizontal axis of the plot). If this were true research not just a tutorial example, we would like to confirm the result with more stations, at least eight with uniform azimuthal distribution. We present this example because we indeed obtained analogous results with many (~ 20) stations elsewhere.







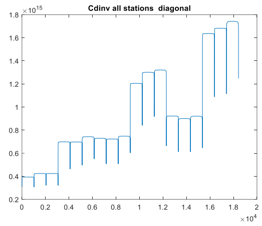


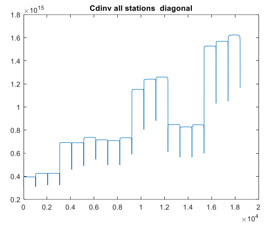
Further iterations

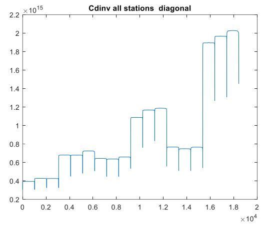
The user can return to the Waveform Inversion tool, and calculate another iteration with the same cova matrix type. In this case, the covariance matrix is calculated from obs-syn(1), where 1 stands for the first iteration. The updated result appears in the invert folder, rewriting previously created files. I saved the result in 11_cova_fullNNEEZZ_0.02-0.06, subfolder iteration 2 and copied a part of inv1.dat here:

moment (Nm): 2.2268115E+17 (in iteration 1 it was 2.212E+17)
moment magnitude: 5.5
VOL % :   0.4 (in iteration 1 it was 5.1%)
DC % : 33.0
CLVD % : -66.6
strike,dip,rake: 161   77   97
strike,dip,rake: 309   14   59
varred= 0.5613401   (in iteration 1 it was 0.5661213)

Here we plot numeric values of diagonal terms of Cdinv, which can be interpreted as weights implicitly applied by cova. Changes in Cdinv should decrease from one iteration to the other (convergence should be observed). Here we compare Cdinv for iterations 2, 4, and 8:



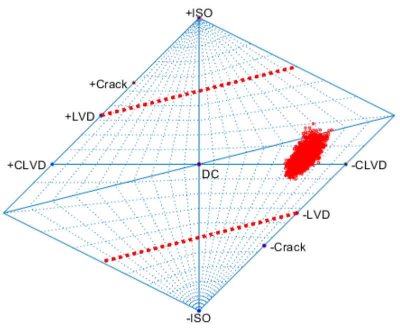


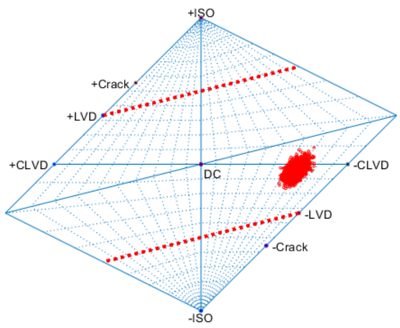


MTs from the iterations should converge. Included in 11_cova_fullNNEEZZ_0.02-0.06 are subfolders for 8 iterations; following inv1.dat files you can see that VOL% was 11.8 without cova, while the iterations 1-8 provided VOL%= 5.1, 0.4, -2.6, -4.0, -4.8, -3.2, -3.2, and -3.0. A small (generally non-monotonous) variation of the moment and a decrease of the varred value is typical when iterating with cova, and it would need more research.

Iteration 1, 2, ……………………………………………..8
VOL%= 5.1, 0.4, -2.6, -4.0, -4.8, -3.2, -3.2, and -3.0
mom= 2.212, 2.226, 2.247, 2.245, 2.241, 2.202, 2.201, 2.198
VR =0.566, 0.561, 0.557, 0.555, 0.554, 0.552, 0.554, 0.555

Without the need of your own repetition of all 8 iterations, you can compare the plots uncplot4.emf in subfolders of iteration 1 and iteration 8; you will find a small decrease of VOL. We conclude that this data indicates VOL converging at ~ -3%. See iteration 1 and 8:





Remark: We made many iterations because of our interest in VOL. In more routine applications, we can probably recommend just 1 or 2 iterations because it seems that most changes occur between iterations 0-3.

Warning

The inversion without covariance matrix (or with a single-diagonal-constant matrix in normal Isola – no_cova) implicitly assumes that data error is uncorrelated. However, looking at the above examples, we see strongly correlated residuals (large non-diagonal terms), mainly at station 1. This error correlation should be suppressed because the least-squares MT inversion theoretically assumes Gaussian data errors. Inversion with cova indeed makes such a suppression. Nevertheless, we cannot be sure that the covariance matrix used here (i.e., the matrix derived from the residuals) contains all information about data error. Moreover, the same data residuals can be used to generate another type of covariance matrix, e.g. the matrix including cross-component correlations, and such matrix gives again a slightly different result. Moreover, variations in the results occur when changing the used stations, velocity model, and frequency range. For example, in 0.02-0.05 Hz, using the same 6 stations as here, VOL% converged to -2%. In this sense, the application of cova, although seemingly mandatory, introduces certain non-uniqueness. And the true uncertainty can be much greater than seen in uncertainty plots (!).

Is this too pessimistic? A more optimistic view is this one: Station 1 (CN2) had a noisy record; we have intentionally selected such an example. If using a normal Isola with another (less noisy) frequency band for this station, or replacing CN2 by another non-noisy China station, the need for the covariance matrix is reduced. In a research of 6 similar deep-focus events, processed at many stations, we obtained low VOL (< 10%) and similar large CLVD without cova, too.

For readers interested in details of Inversion with cova: This type of Isola inversion consists of three automatically chained actions in Waveform Inversion tool:

a) When choosing cova type Toeplitz, Matlab runs the code prep_FULL.m. It creates covariance matrices of the individual stations, uses them to form the block-matrix of all stations (Cd), calculates the inverted matrix (Cdinv), and writes Cdinv into a binary file hinv.bin in folder invert.

b) Fortran code isola_cova.exe reads hinv.bin and uses it to invert waveform for MT; running can be seen in the system command ('black') window. Results appear in folder invert and should represent a better approximation of true MT than those without the cova.

c) Postprocessing occurs, again in Matlab, using hinv.bin in the code prep_NORM.m. It produces some plots useful for understanding how the inversion works, e.g. what is the difference between so-called non-standardized and standardized seismograms and between their spectra. In the Matlab command window it gives also variance reduction for the non-standardized and standardized seismograms as a measure of the usefulness of the cova method. Another very useful measure of the usefulness are the plots (see above "corrow100") depicting autocovariance functions of data residuals before and after standardization, and their analogy in spectral domain ("Spectra DAT…SYN").