Test Report: EEGSourceLocalizer Software

Leila Mousapour

December 21, 2020

1 Revision History

Date	Version	Notes
19 Dec, 2020	1.0	First version of the VnV report

2 Symbols, Abbreviations and Acronyms

symbol	description
T	Test
SRS	Software Requirements Specification
VnV	Verification and Validation
VnVR	Verification and Validation Report
MG	Module Guide
MIS	Module Interface Specification
EEGSourceLocalizer	Electroencephalogram Source Localization Software
EEG	Electroencephalogram
SL	Source Localization

For completion, please see SRS Documentation at https://github.com/LeilaMousapour/Brain-Computer-Interface-/blob/master/docs/SRS/SRS.pdf

Contents

List of Tables

List of Figures

This document is a report on the results of testing based on VnVP for EEGSourceLocalizer software. Detailed descriptions of the test cases can be found in the VnV plan documents. For each designed test, the data for the test case and the automated test script (if applicable) can be found in `https://github.com/LeilaMousapour/Brain-Computer-Interface-/tree/master/src` under test folder.

3 Functional Requirements Evaluation

3.1 Static testing

3.1.1 Code walk-through

As mentioned in the design documents, this software is built on Fieldtrip toolbox which is an open source toolbox in MATLAB that offers source localization. The official tutorials of this toolbox are mainly designed and written for MEG data and are not directly applicable to EEG data. Thus, it is worth mentioning that the main problems with this software are due to the base scientific knowledge of the source localization problem in general and LCMV in particular and in these tests, the focus is not the pure programming challenges such as memory efficiency, style etc. . Therefore, the user has to take the challenge understanding every step of the SL problem, including the mathematical details of the forward and inverse problems profoundly in order to modifying all the steps for EEG data.

Even in the widely-used Brainstorm application, which the user needs to interact with it through a GUI and it is more easier than coding up every step, the user need to have very good knowledge of the SL problem and the algorithm they are using in order to obtain meaningful results. Therefore, all the tests, including the static test of code walk-through are designed in a way to build confidence that mathematics of the problem is properly defined and implemented as different configuration, specifically for EEG data. Thus, this walk-through was designed in a more thorough way to go over the mathematical basics and checking the code accordingly.

Several sessions were held to walk the undergraduate student, Yar Mohammad Al Dabagh who working on the project through the code. The student and the developer both were assigned the bellow steps:

1. Read the LCMV base paper thoroughly and the necessary chapter of the source localization book Veen et al. (1997)

2. Read and follow the Fieldtrip tutorial on "Localization of oscillatory sources" in depth

3. The developer walked the student through every module of the software and explained how the implementation and different configurations matches the mathematics presented in the LCMV paper and why the chosen configurations are appropriate for EEG signal.

4. The coding style and module design of the software was discussed.

5. The quality of the comments was checked.

6. The simplicity of the main (control) module was checked.

As a result of this test, several issues are revealed:

1. There was not enough understanding of the value that is calculated as "pow" in the output of the LCMC Fieldtrip function which will be plotted in the end. Thus, we could not interpret the result of the plots very well.

2. The final source activity plots are basically plotting the -averaged- activity over the time period of the trial. It should be discussed and further looked into that how valid it is to look at and interpret the averaged activity of the sources as it is huge assumption.

3. The source space of the sLORETA algorithm is not the whole brain and is only the cortical surface (a thin layer on the surface of the brain).

3.2 Dynamic testing

3.2.1 test-input

When the input EEG data is not meeting the condition (in-range amplitude and frequency), EEGSourceLocalizer will show an error. This test has been passed.

3.2.2 test-rank

When the input EEG data is not full rank, EEGSourceLocalizer will show an error. This test has been passed.

3.2.3 test-SimulatedEEG

In this test case, the input data (EEG) was manufactured based on the known solution (the location and strength of the sources). Recall, the sources are modeled as current dipoles. This test was done in 2 different configuration. First, the simplest case which is the EEG signal generated from one single source was tested. Next, to check the performance of the software in the presence of several sources with different strengths and orientations, the EEG signal generated from 3 separate dipoles was tested in the software.

In both tests, first the head model and the forward model (lead filed) were calculated based on the template ICBM 152 MRI. Then, using the ft_dipolesimulation() function, by setting the configurations, the EEG signals were generated. Finally, a normal noise was with relative amplitude of 1/100000 was added to the signal. These EEG signals were plotted on the 2D electrode locations to confirm that the simulated EEG is as expected.

Additionally, this test was automated. After calculation of the sources time series and power, the maximum activity was found and the location of the source point corresponding tho the maximum activity was obtained. Afterwards, the distance between the distance between the location of the dipole that EEG is generated from and the location of the maximum activity in the software output is calculated and compared with the resolution of the head model grid (which is 1 cm as it is configured in the software code). If the distance is less than the grid resolution, the test will pass which in this case, it successfully passed.
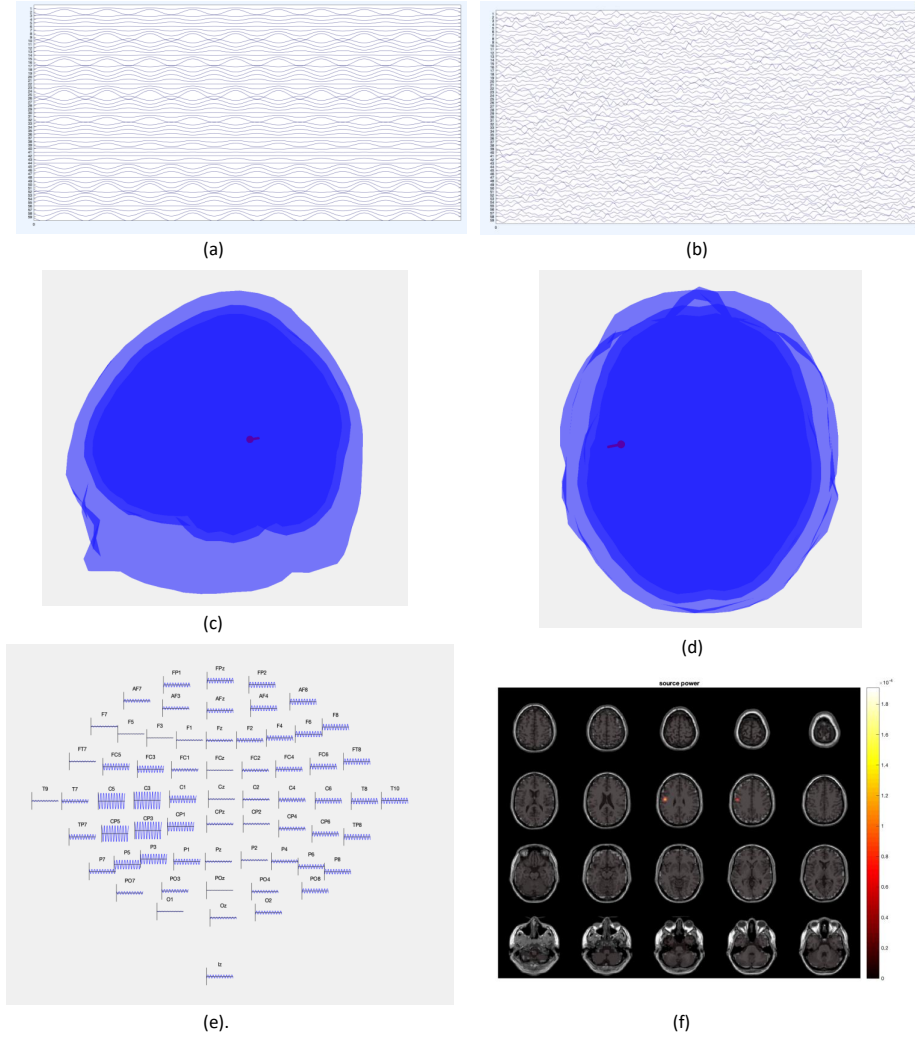
Figure 1: The result of EEG simulation test. (a) simulated EEG from one dipole on all channels. (b) simulated EEG plus noise. (c,d) the pre-defined location of the dipole that generates the EEG. (e) a 2D map of the signal at each channel. (f) the result of the software for localization of the simulated EEG

(a)

(b)

(b) (d)

(e)

(f)
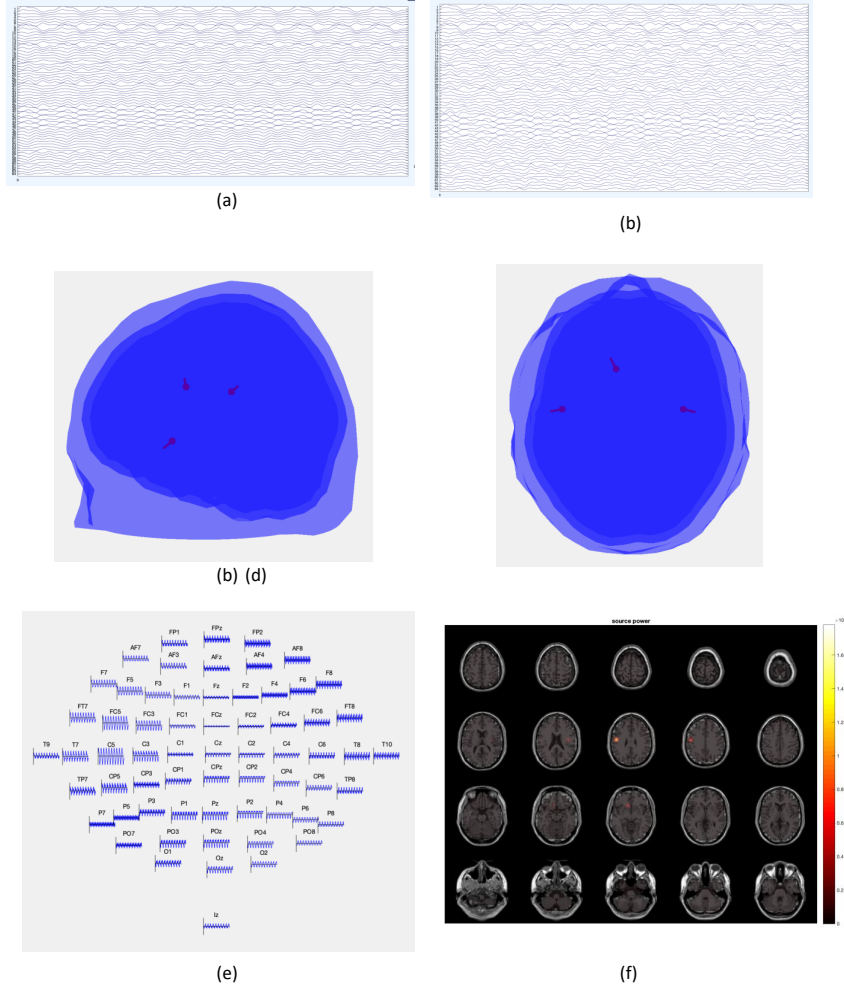
Figure 2: The result of EEG simulation test. (a) simulated EEG from three dipoles on all channels. (b) simulated EEG plus noise. (c,d) the pre-defined location of the dipole that generates the EEG. (e) a 2D map of the signal at each channel. (f) the result of the software for localization of the simulated EEG

It is worth mentioning and recording that in the test case of 3 dipoles, when the frequency of the dipoles are not different, EEGSourceLocalizer is not able to localize 3 separate sources and it merges the sources. It is one of the shortcoming of the LCMV algorithms in general and it is due to the mathematics not the implementation or code.

5

3.2.4 test-BrainstormPsudoOracle

Brainstorm application is a software for a variety of MEG/EEG signal processing including source localization. This application is controlled via a GUI and there are online tutorials on how to work with it. In this test, the BCI iv dataset is used. This dataset comprises of he data from 5 subjects which have performed 2 kind of moor imagery tasks over the period of 4 seconds. The goal is to locate the sources active in each of these 2 classes of motor imagery and classify them.

This test aimed to compare the numerical results of the sources time series extracted from Brainstorm (as a pseudo oracle) as the output result of the EEGSourceLocalizer. However, after learning how to work with Brainstorm and obtaining the result, I found out that this software uses "OpenMEEG" to create the head model while I use "Dipoli" method for head model generation which makes the coordinate systems incompatible. Resolving this issue took plenty of time without any success so far. Therefore, in this test we only show the manual visual verification based on the source activity maps. As we can see based on the maps, the maps does not seem very similar. This problem should be look into further more.
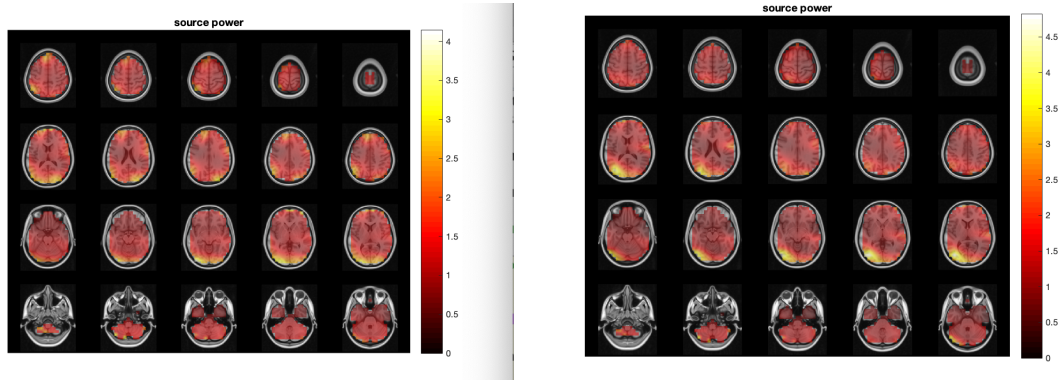

Figure 3: The source power of right hand motor imagery is shown on the left and the source power of left hand motor imagery is shown on the right
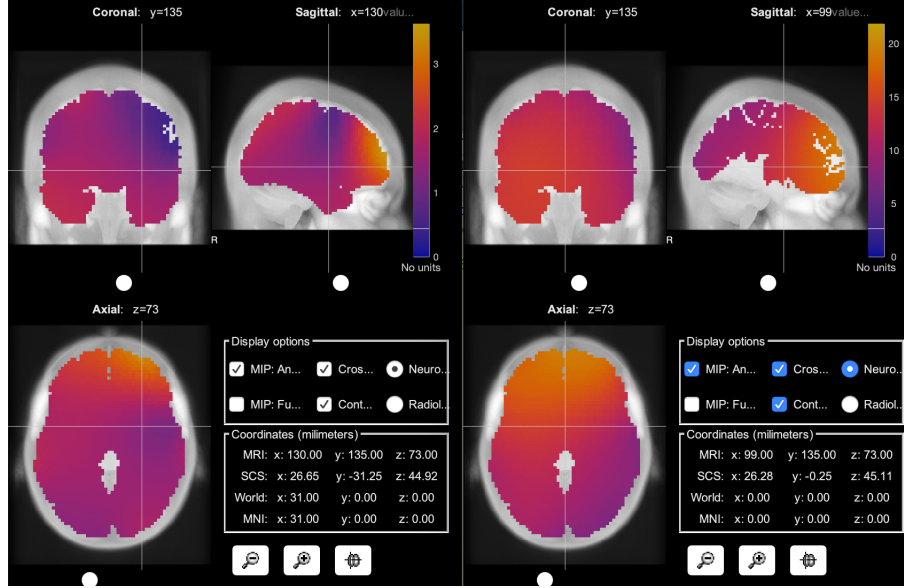
Figure 4: The source power of right hand motor imagery is shown on the left and the source power of left hand motor imagery is shown on the right

3.2.5 test-GroundTruthData

As mentioned in the VnVP, "Simultaneous human intracerebral stimulation and HD-EEG, ground-truth for source localization methods" data set that comprises EEG recorded electrical activity originating from precisely known locations inside the brain of living humans Mikulan et al. (2020) can be used as a solid test case for any source localization procedure. Seven subjects participated in the study and a total of 61 sessions were obtained. Intracranial shafts were implanted using a robotic assistant and single-pulse biphasic currents lasting 0.5 ms were delivered at intensities ranging from 0.1 to 5 mA. EEG signals were recorded from 256 channels and sampled at 8000 Hz. The location of the intracranial electrodes was assessed registering the post-implant CT to the pre-implant MRI by means of the FLIRT software tool and these locations are provided in the dataset in MIN coordinate system. Contact positions were plotted on a flatmap of the MNI152 template by projecting each contact's coordinates to the closest vertex of the brain surface reconstruction. A copy of the paper is included in the corresponding test folder
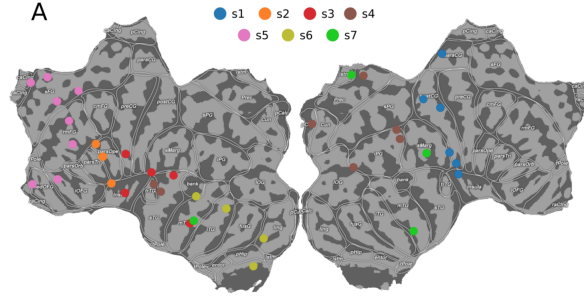
Figure 5: Flatmap of stimulation sites by subject Mikulan et al. (2020)

In order to test this dataset, one run from each subject was chosen and the result is reported. The data in the dataset is provided in .npy format and has been read and transformed into .mat format by a python script. Also, the location of the implanted electrodes are provided. For each subject, the source localization was performed using the software. Recall that the activity of sources are estimated in time, thus it is important where in time we are looking at the source activity. After the presentation in the class (when I explained I could not get results from this data), the code walk-through brought out the point that -averaging- over time is a very big assumption. Also, a testing this data in the Brainstorm application, confirmed this point that we should look are the activity of the sources around the moment of the spike which is only 1ms and if we average over the whole 2500ms of the data, this activity will fade away. Therefore, in the automated test script in the "testGroundTruthData" folder, first the the time point in which the spike happens is located by finding the index of the maximum amplitude of the signal. Then, the activity of all the sources in that particular time point is obtained and the location of the source point with the maximum activity is found. The norm distance of this location and the actual location of the electrode is calculated and compared against the threshold of 30 mm. This threshold is chosen as based on the result in the paper, localization error varied between 1 20mm for the 3 algorithms they tested the data on. Thus, this software would pass the test if it can localize the source by the penalty of 10mm.

The result of this test on one run of every subject is reported in table 1 and compared with the result of the paper. Also, for demonstration purposes, the source activity plot and the actual and estimated location of the source

8

Table 1: The result of the ground truth dataset mm

Sbj	Run	Software error	Paper min error	Pass
sbj1	01	15.39	5	Pass
sbj2	01	81.33	6	Not Pass
sbj3	01	37.92	6	Not Pass
sbj4	01	53.98	4	Not Pass
sbj5	01	18.48	6	Pass
sbj6	01	42.10	5	Not Pass
sbj7	01	66.30	6	Not Pass

inside the head model is included for one of the subjects (subject5-run01). As we can see, EEGSourceLocalizer has the ability of locating the source within an acceptable error range for subjects 1 and 5 and for other subjects it is a relatively big error which shows the need for more parameter tuning in the EEGSourceLocalizer. It is important to note that the min localization error reported in the paper is the result of extensive hyperparameter tuning. A total of 4800 solutions were calculated for each session and then the best solution across all montages and parameter's configurations was computed and reported in the paper.
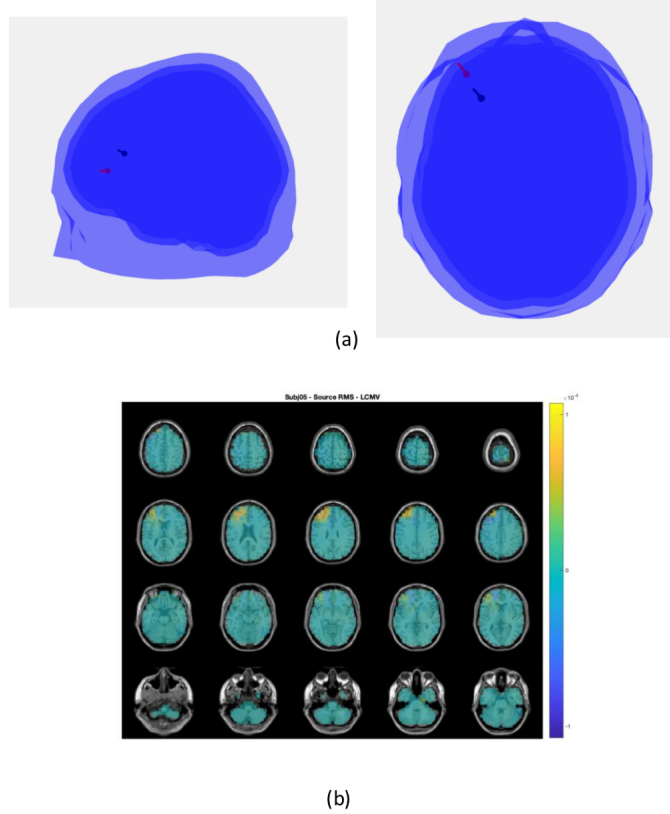
(a)

(b)

Figure 6: (a) The estimated dipole location is shown in black and the actual location of the implanted electrode is shown in blue. (b) The activity of the sources at the moment of the spike

It is very much worth mention that during this test I learnt a very important information about the way the ft_sourceanalysis() produces the results. This important point is that the results of source localization are in the in NM or CTF space which are 2 coordinate systems. However, I have always been comparing the locations of the sources with locations in the MNI coordinations system which is a huge mistake. Automation of this test, which required comparing the locations of the implanted intracranial electrodes with the result the software produces made me realize this longterm mistake!

4   Nonfunctional Requirements Evaluation

4.1   Accuracy

As explained in the VnVP and as you could see from the testings, due to the nature of the problem, testing the accuracy is very hard for an arbitrary EEG data as we do not know the true solution to compare our results with. The ground truth dataset is a way of testing the accuracy of the software which the results are reported in section 3.2.5.

5   Unit Testing

Still working on it.

6   Changes Due to Testing

Each of the tests mentioned in previous sections revealed different errors or lack of understanding that is explained in the descriptions of the test report.

Additionally, EEGSourceLocalizer is far from perfect and these modifications and test should be applied on this software in the near future:

- Automate the code to find the best setting and configuration of LCMV appropriate to the data

- The result of the LCMV algorithm should be similar to other SL algorithms and should be tested against them.

- sLORETA algorithm must be added to the software as stated in the SRS.

7   Trace to Requirements

The following table demonstrate the connection between the tests and requirements in SRS.

	R1	R2	R3	R4	R5	NR1
test-input	X					
test-rank		X				
test-SimulatedEEG			X	X		
test-BrainstormPsudoOracle				X	X	X
test-GroundTruthData				X	X	X

Table 2: Traceability Matrix Showing the Connections Between Tests and Functional and Nonfunctional System Requirements

8 Trace to Modules

The following table demonstrate the connection between the modules and tests.

Test	M2	M3	M4	M5	M6	M7	M8
code walkthrough 3.1.1	X	X	X	X	X	X	X
test-input 3.2.1	X						
test-rank 3.2.2	X						
test-SimulatedEEG 3.2.3	X	X	X	X	X	X	X
test-BrainstormPsudoOracle 3.2.4	X	X	X	X	X	X	X
test-GroundTruthData 3.2.5	X	X	X	X	X	X	X

Table 3: Traceability Matrix showing the connections between unit tests and modules

References

Ezequiel Mikulan, Simone Russo, Sara Parmigiani, Simone Sarasso, Flavia Maria Zauli, Annalisa Rubino, Pietro Avanzini, Anna Cattani, Alberto Sorrentino, Steve Gibbs, Francesco Cardinale, Ivana Sartori, Lino Nobili, Marcello Massimini, and Andrea Pigorini. Simultaneous human intracerebral stimulation and HD-EEG, ground-truth for source

localization methods. *Scientific Data*, 7(1):127, 2020. ISSN 2052-4463. doi: 10.1038/s41597-020-0467-x. URL https://doi.org/10.1038/s41597-020-0467-x.

B.d. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44(9): 867–880, 1997. doi: 10.1109/10.623056.