

User Guide

Version 2.3.0

Contents

[What is HAPPE? 4](#_Toc95395260)

[What does HAPPE require? 5](#_Toc95395261)

[How to get HAPPE 6](#_Toc95395262)

[Setting Up HAPPE: 7](#_Toc95395263)

[Adjust Java Heap Settings 7](#_Toc95395264)

[Set MATLAB Paths 8](#_Toc95395265)

[Configure EEGLAB Preferences 9](#_Toc95395266)

[How to Run HAPPE v2\_1 10](#_Toc95395267)

[HAPPE Pipeline Guides 11](#_Toc95395268)

[HAPINNES Resting State 11](#_Toc95395269)

[Following Command Line Prompts 11](#_Toc95395270)

[HAPPE Task-Related 22](#_Toc95395271)

[Following Command Line Prompts 22](#_Toc95395272)

[HAPPE ERPs (HAPPE+ER) 34](#_Toc95395273)

[Following Command Line Prompts 34](#_Toc95395274)

[HAPPILEE Resting-State 48](#_Toc95395275)

[Following Command Line Prompts 48](#_Toc95395276)

[HAPPILEE Task-Related 59](#_Toc95395277)

[Following Command Line Prompts 59](#_Toc95395278)

[HAPPILEE ERPs 69](#_Toc95395279)

[Following Command Line Prompts 69](#_Toc95395280)

[HAPPE+ER’s generateERPs Add-On 80](#_Toc95395281)

[Run generateERPs 80](#_Toc95395282)

[Following Command Line Prompts: 80](#_Toc95395283)

[Data Quality and Pipeline Quality Outputs 84](#_Toc95395284)

[Tracking Data Quality 84](#_Toc95395285)

[Data Quality Measures 84](#_Toc95395286)

[Pipeline Quality Measures 84](#_Toc95395287)

[Using Outputs to Exclude Files 84](#_Toc95395288)

[Data Quality Measures 85](#_Toc95395289)

[Pipeline Quality Measures 85](#_Toc95395290)

[Variables as Co-Variates and Reporting Measures 86](#_Toc95395291)

[Data Quality Measures 86](#_Toc95395292)

[Pipeline Quality Measures 86](#_Toc95395293)

[Optimized Code Examples 87](#_Toc95395294)

[HAPPE 2.0 87](#_Toc95395295)

[HAPPILEE 87](#_Toc95395296)

[Bad Channel Rejection 87](#_Toc95395297)

[Wavelet Thresholding 87](#_Toc95395298)

[Segment Rejection 87](#_Toc95395299)

[HAPPE+ER 88](#_Toc95395300)

[Filtering to Specified Frequencies 88](#_Toc95395301)

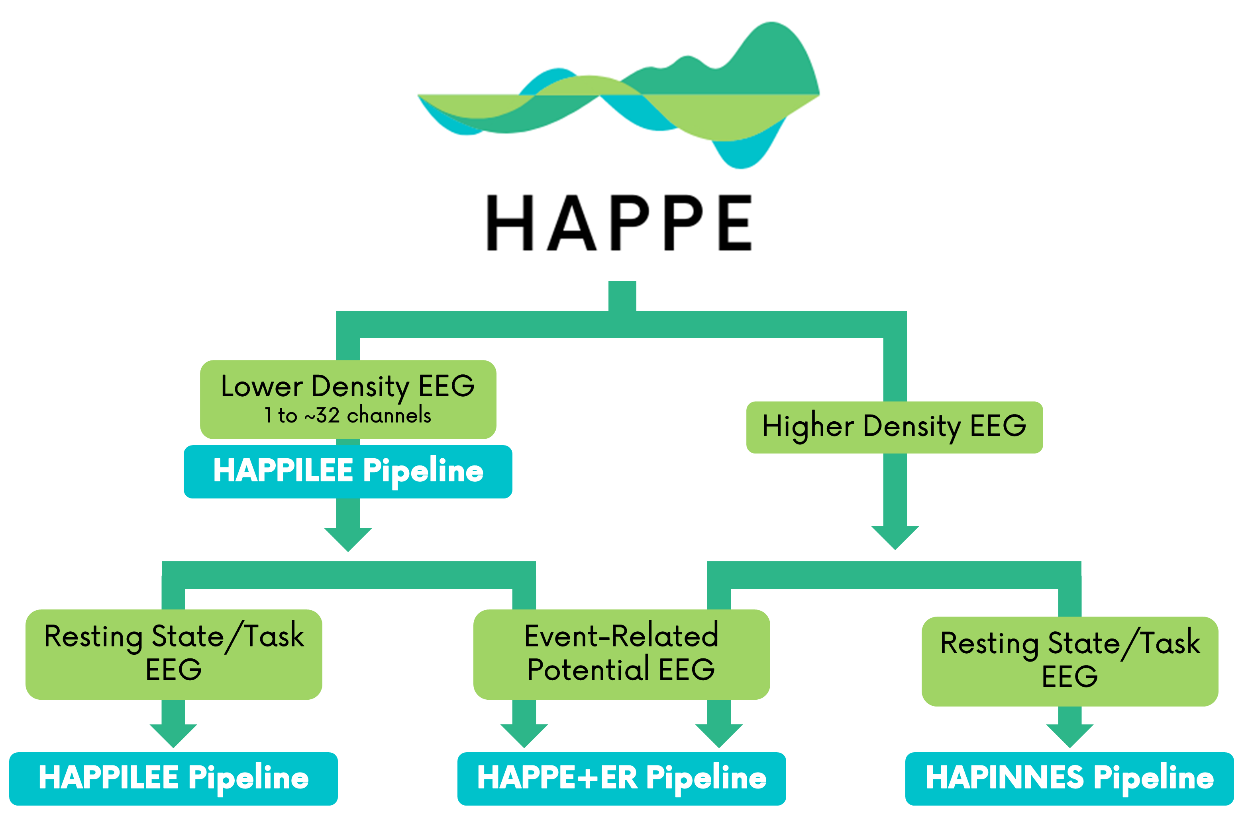
[Wavelet Thresholding 88](#_Toc95395302)

[Segmenting 88](#_Toc95395303)

[Split EEG by Tags 88](#_Toc95395304)

# What is HAPPE?

* HAPPE is a software for taking unprocessed EEG data and automatically processing it in preparation for analysis. It is a combination of the HAPPE 2.0, HAPPE+ER, and HAPPILEE pipelines.
  + HAPPE 2.0: citation pending
  + HAPPE+ER: Monachino, A.D., Lopez, K.L., Pierce, L.J., Gabard-Durnam, L.J. (in press) The HAPPE plus Event-Related (HAPPE+ER) Software: A Standardized Processing Pipeline for Event-Related Potential Analyses
  + HAPPILEE: Lopez, K.L., Monachino, A.D., Morales, S., Leach, S.C., Bowers, M.E., Gabard-Durnam, L.J. (in press) HAPPILEE: The Harvard Automated Processing Pipeline in Low Electrode Encephalography, a standardized software for low density EEG and ERP data.
* Translates recent advances in adult EEG processing to developmental data context.
* Implements wavelet-thresholding approaches for EEG artifact removal.
* Agnostic to the program for running analyses afterwards. Compatible with:
  + BEAPP
  + EEGLAB
  + MATLAB
  + Anything that can read a .txt file of EEG data as an input
* Includes processing reports with data and pipeline quality metrics for assessing data and performance, reporting in manuscripts, and setting quality thresholds for removing data from further analysis.



# What does HAPPE require?

* MATLAB version 2019b or newer
  + NOTE: this version is not compatible with NetStation version 4.5, so you cannot run HAPPE on the same computer as NetStation 4.5.
* The following MATLAB toolboxes (all should be standard for academic licenses):
  + Signal Processing Toolbox
  + Optimization Toolbox
  + Statistics Toolbox
  + Wavelet Toolbox
* The following free software (all included in the HAPPE download):
  + EEGLAB
  + CleanLine EEGLAB plugin
  + MARA EEGLAB plugin
  + FASTER functions
* The HAPPE scripts (all included in the HAPPE download)

# How to get HAPPE

Download HAPPE from the following link: <https://github.com/PINE-Lab/HAPPE>

# Setting Up HAPPE:

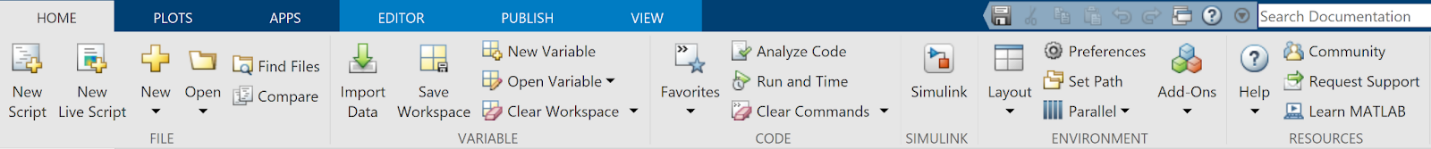
If you do not follow these steps, HAPPE will not run properly.

## Rescue EEGLAB Files from Quarantine (MACS ONLY)

1. If open, close MATLAB.
2. Open the terminal.
3. Type the following, replacing FILEPATH with the full path to the folder holding HAPPE and ignoring line breaks:  
   sudo xattr -r -d com.apple.quarantine FILEPATH/HAPPE/Packages/eeglab2022.0  
   Press ‘return’ to execute the command.  
   Example for FILEPATH: **/Users/lgd/Desktop/**HAPPE/Packages/eeglab2022.0
4. When prompted, enter your user password.
5. Type the following, replacing FILEPATH with the full path to the folder holding HAPPE and ignoring line breaks:  
   sudo find FILEPATH/HAPPE/Packages/eeglab2022.0 -name \\*.mexmaci64 -exec spctl --add {} \;  
   Press ‘return’ to execute the command.
6. Close the terminal.
7. Re-open MATLAB.

## Confirm Toolbox Installation

1. In MATLAB under the HOME tab, select the drop-down arrow under “Add-Ons.”



1. In the drop-down menu, select “Manage Add-Ons.”

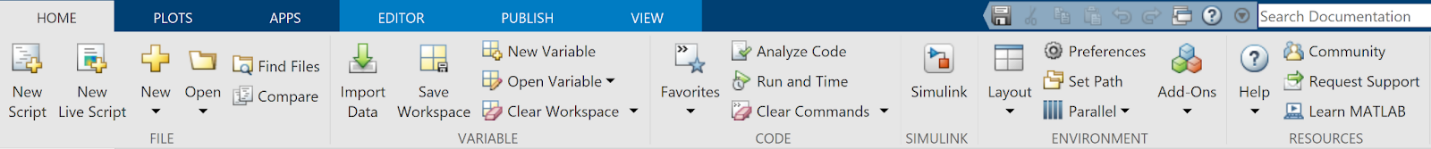
Graphical user interface, text, application, chat or text message

Description automatically generated

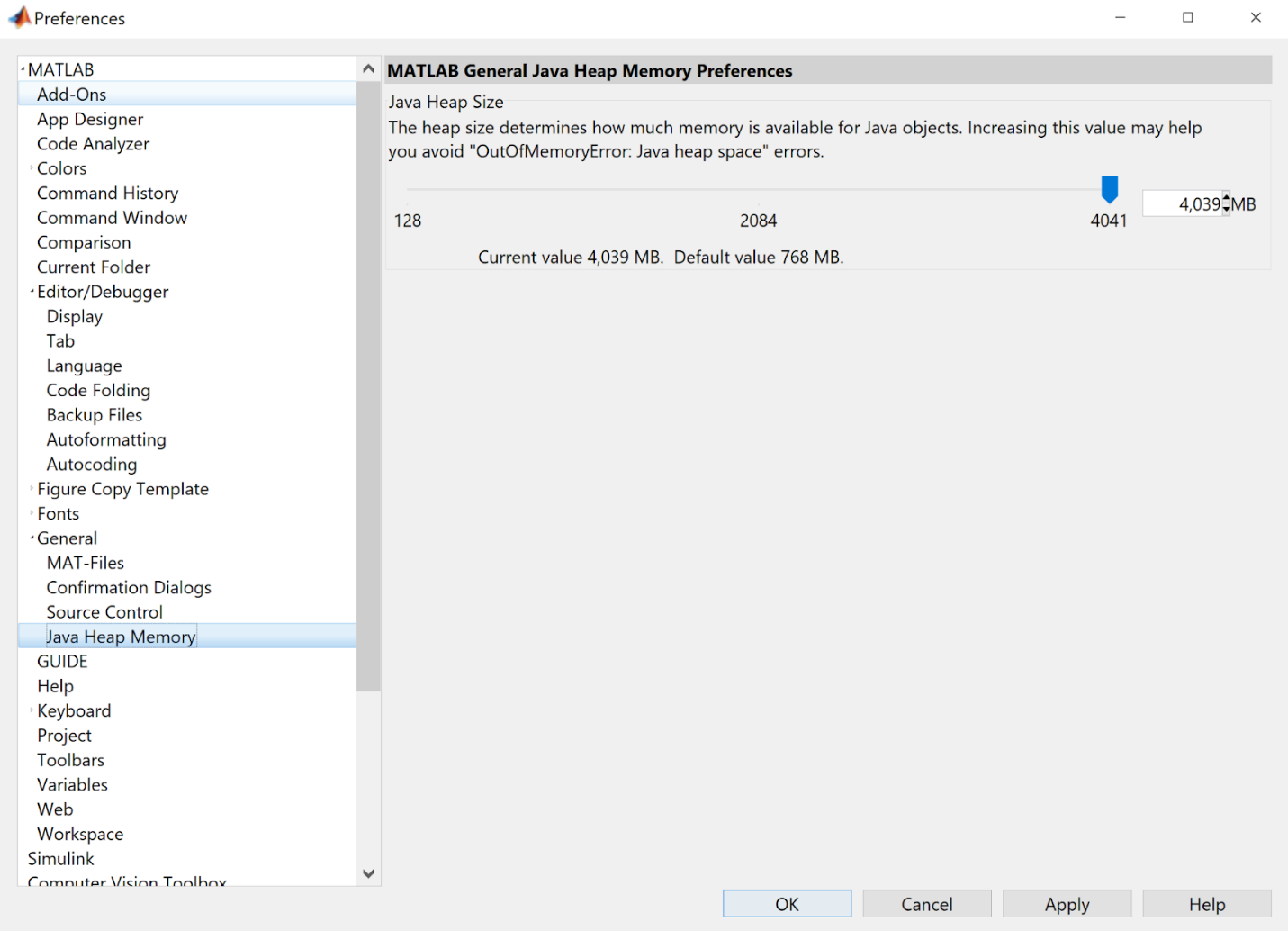
1. Check the list of Installed Add-Ons for the following toolboxes:
   1. Wavelet Toolbox
   2. Signal Processing Toolbox
   3. Statistics and Machine Learning Toolbox
   4. Optimization Toolbox
2. If any of the listed toolboxes are not included in the list, click “Get Add-Ons” in the top right corner of the Add-On Manager. For each missing toolbox do the following:
   1. In the Add-On Explorer, filter to official toolboxes using the MathWorks Source filter.
   2. Search for the toolbox in the list and click on its name.
   3. If needed, click the Sign-In to Install button. Otherwise, simply click install.
   4. Follow the install launcher prompts through to completion. MATLAB may close and re-open during this process.
3. Confirm in the Add-On Manager that all toolboxes have been installed.

## Adjust Java Heap Settings

1. In MATLAB under the HOME tab, select “Preferences.”



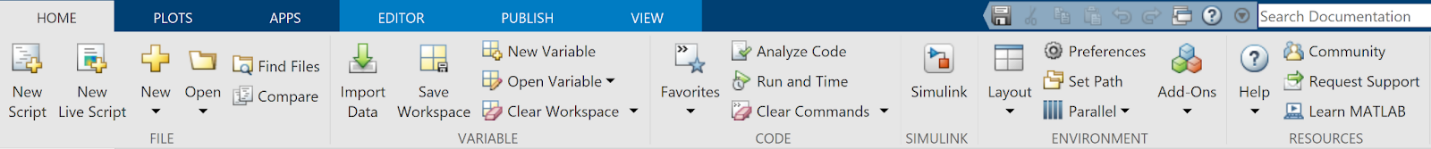
1. In the left-hand sidebar, under General, select “Java Heap Memory.”
2. The default value for the Java Heap is 768 MB. You will want to increase this to the maximum by either dragging the blue arrow all the way to the right or using the text box to the right of the bar.
3. Select “Apply” then “OK.”



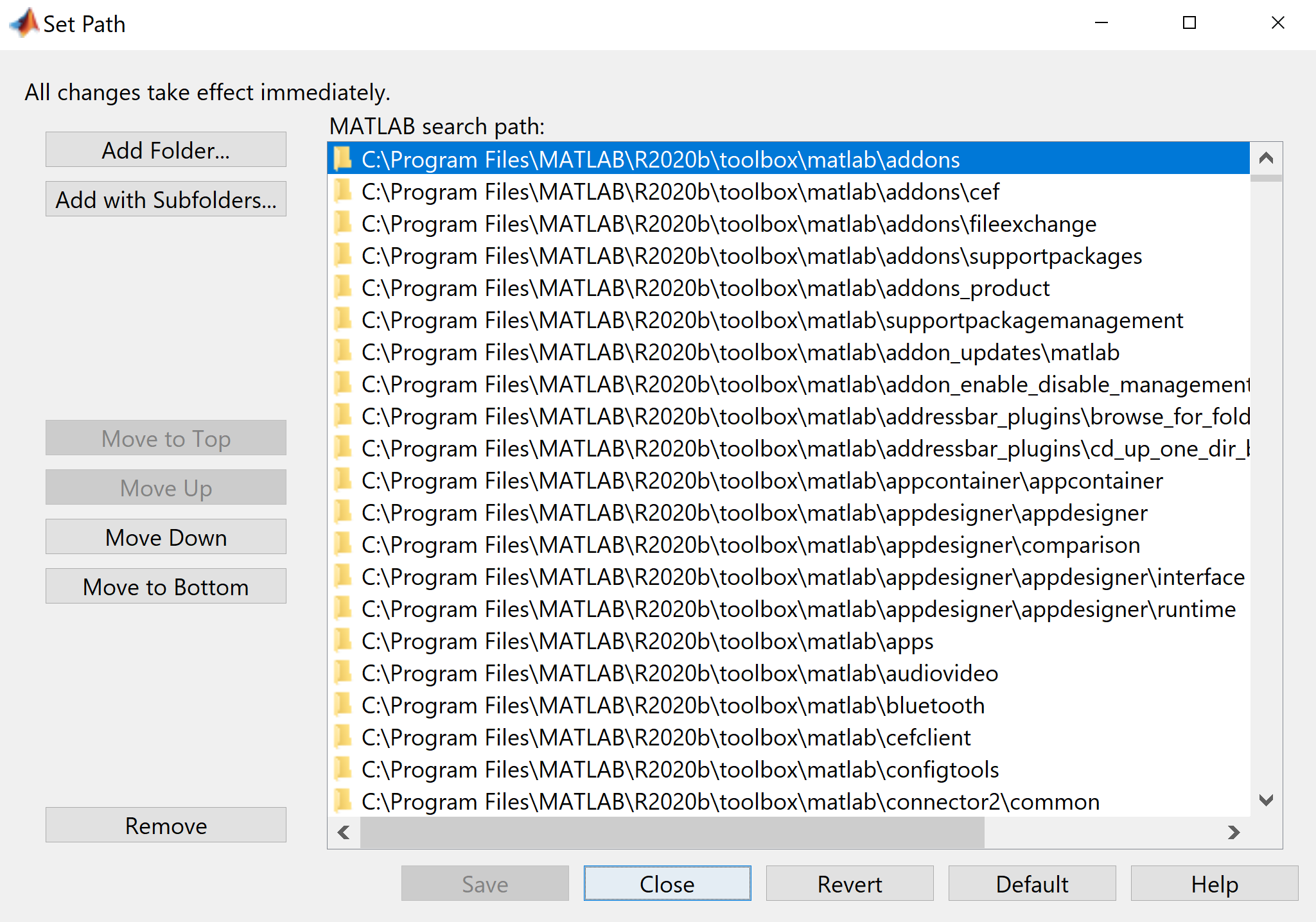
## Set MATLAB Paths

We recommend adding HAPPE and certain subfolders to your path.

1. Under the HOME tab in MATLAB, select “Set Path.”

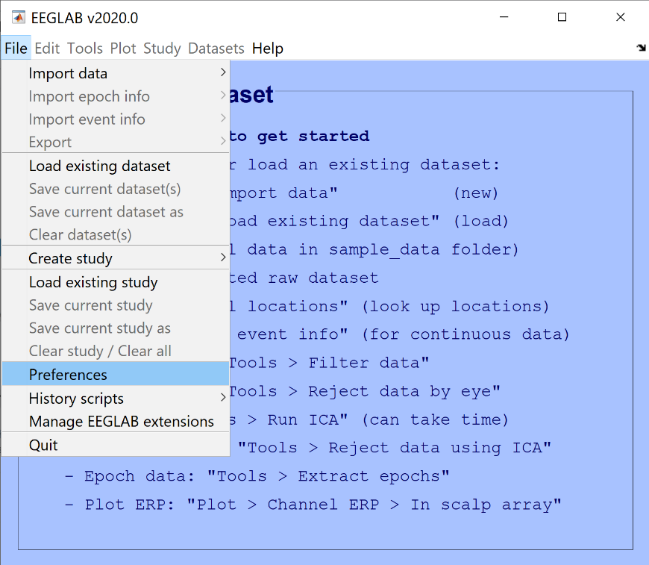


1. In the Set Path window that opens, use the Add Folder button to add the following folders (Please note that if you are using a Mac, \ will be replaced with /):
   * HAPPE
   * HAPPE\scripts
   * HAPPE\scripts\pipeline\_scripts
   * HAPPE\scripts\UI\_scripts
   * HAPPE\acquisition\_layout\_information
   * HAPPE\Packages
   * HAPPE\Packages\eeglab2022.0
   * HAPPE\add-ons
   * HAPPE\add-ons\generate
   * HAPPE\add-ons\generate\scripts
   * HAPPE\add-ons\validate
   * HAPPE\add-ons\validate\etc
2. Click the “Save” button.
3. Click the “Close” button.

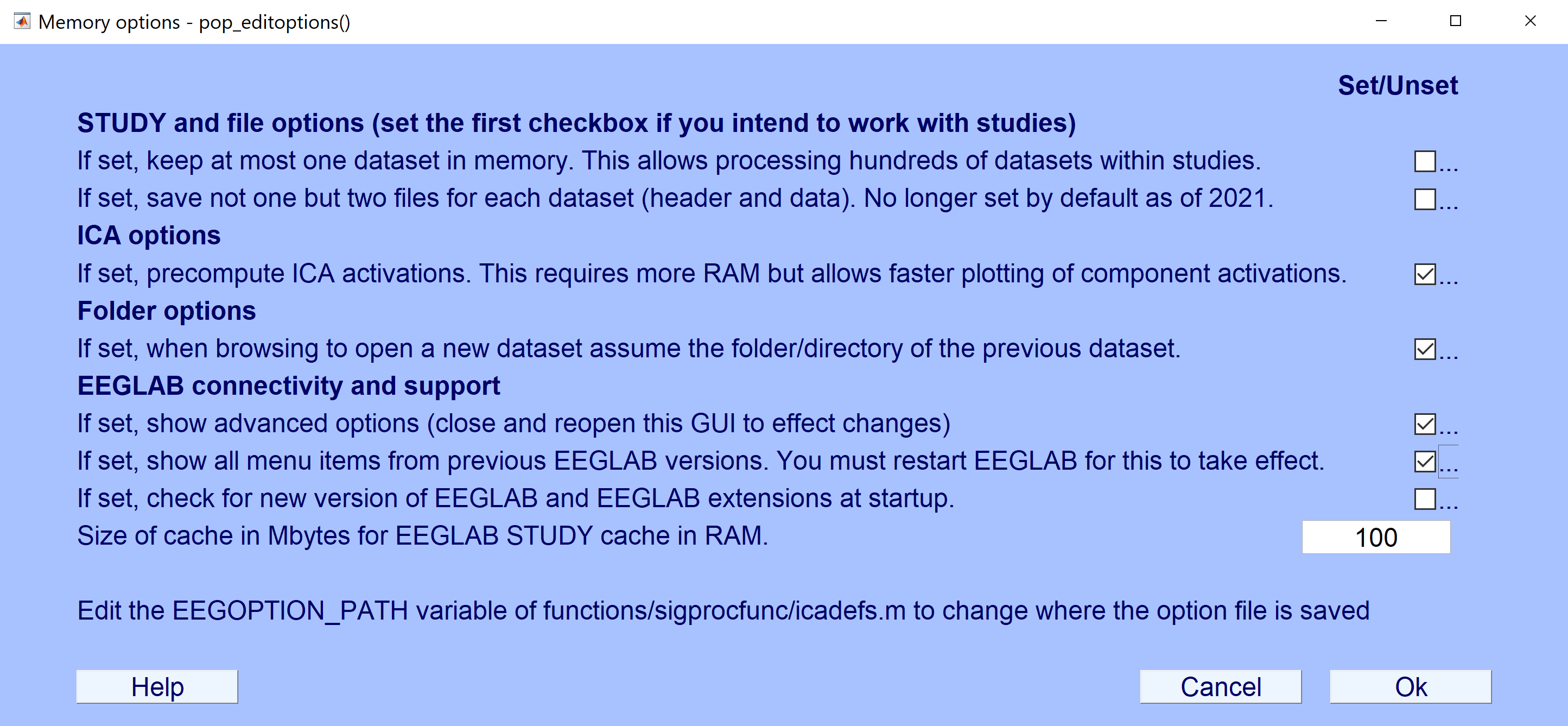


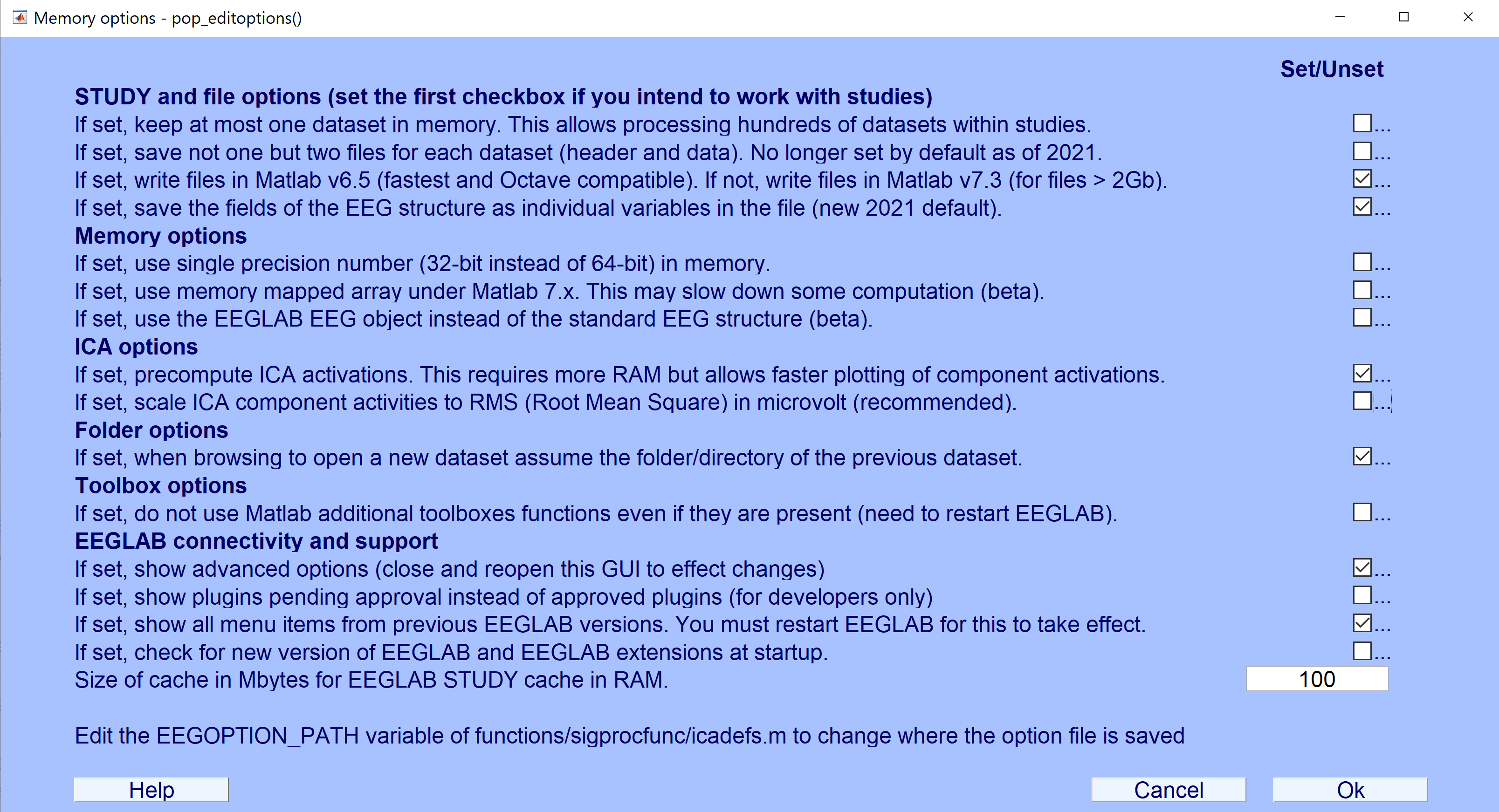
## Configure EEGLAB Preferences

1. Launch EEGLAB by entering eeglab into the command window.
2. In the File menu in the toolbar at the top of the window, select Preferences.

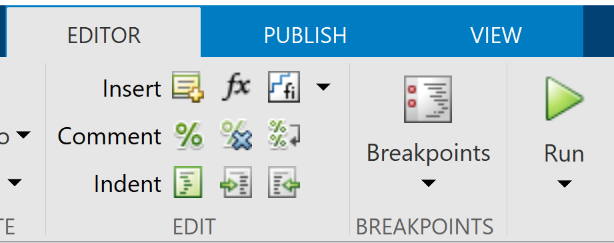


1. Click the checkboxes to enable advanced options and to show all menu items from previous EEGLAB versions. Click OK.



1. Close EEGLAB. Re-Open EEGLAB and the Preferences menu using steps 1 and 2, above.
2. Click the checkboxes so your EEGLAB preferences match those shown below. When done, click OK. 
3. Close EEGLAB.

# How to Run HAPPE

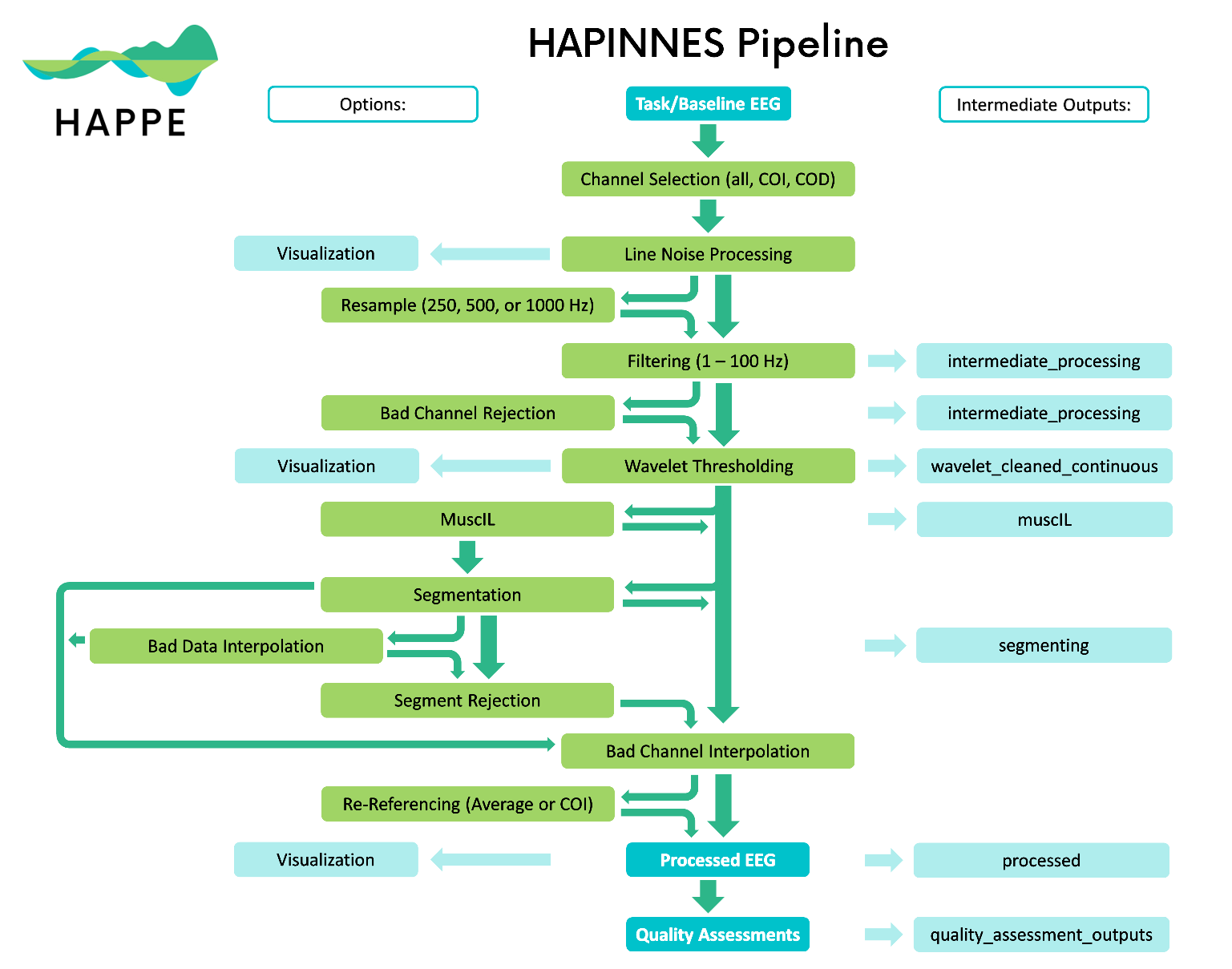
1. Navigate to the main HAPPE folder in your file browser.
2. Open HAPPE\_v2\_3.m in MATLAB.
3. In the Editor tab, hit “Run.”
4. Follow the prompts in the command window of MATLAB
   * For detailed instructions regarding the prompts, see relevant sections below.

**OR**

1. Type HAPPE\_v3 in the Command Window and hit your newline key (enter – Windows; return – Mac)
2. Follow the prompts in the command window of MATLAB
   * For detailed instructions regarding the prompts, see relevant sections below.

# HAPPE Pipeline Guides

## HAPINNES Resting State



### Following Command Line Prompts

Prompts as they appear in the command window are written in a monospace font, like this, followed by a brief description of what to enter, with an example. Sometimes a certain choice will result in a different set of prompts, in which case they will be indented under a heading that describes the choice resulting in the branch in directions.

Enter the path to the folder containing the dataset(s):

Input the folder, including the full path, where the raw data is located. If running HAPPE for the first time on this data, the folder should ONLY include the raw files you wish to process; no other folders or files should live inside that folder. If reprocessing data, all outputs (folders and documents) should be in the folder in the same file structure as they were created during the original HAPPE run in addition to the raw data.

**Example (Mac):** /Users/laurelg-d/Desktop/Data Folder

**Example (PC):** C:\Users\laurelg-d\Documents\Data Folder

Select an option:

raw = Run on raw data from the start

reprocess = Run on HAPPE-processed data starting post-artifact reduction

If this is your first time processing the raw data file(s), input raw. If you wish to re-run raw data that has previously been processed, starting at the MuscIL or segmentation step, input reprocess. Depending on your choice, you may be asked to follow additional prompts (see below).

If reprocessing existing data:

Files, such as processed data and quality metrics may already exist for this dataset.

overwrite = Overwrite existing files

new = Save new files

To keep both the already existing files and the newly created files, input new. If you do not want to keep previous files for this dataset, input overwrite.

If saving new files:

Use a default or custom label for processed data?

default = Default name (\_rerun\_dd-mm-yyyy)

custom = Create your own file label, starting with an

underscore

To create your own save name for the processed data, input custom. You will then be prompted to enter your custom label. This label should start with an underscore (“\_”) to encourage best naming practices. If you do not want to create a custom label, you can use the default (“\_rerun\_dd-mm-yyyy”), using the current date) by entering default. In either case, this will save the files as the original raw file name with the label added. Caution: if you already have data with this label, HAPPE will overwrite the existing files.

Load pre-existing set of input parameters? [Y/N]

If you have already set parameters through HAPPE for either this or another dataset, you can choose to load these parameters instead of entering them from scratch. To use a previously saved set, enter Y (case insensitive); otherwise, enter N (case insensitive).

If loading pre-existing parameters:

Enter your parameter file, including the full path and file extension:

Enter the full name of the parameter file including the path and file extension, examples below. By default, parameters are saved in a folder named **input\_parameters**. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/Data Folder/input\_parameters/

inputParameters\_dd-mm-yyyy.mat

**Example (PC):** C:\Users\lgd\Documents\Data Folder\input\_parameters\

inputParameters\_dd-mm-yyyy.mat

If running a different version of HAPPE than the parameters were saved with:

These parameters were saved using a different version of HAPPE and ay be incompatible with the running pipeline.

Proceed anyway? [Y/N]

If your parameters were saved using a different version of HAPPE, there is a chance that they do not have the full set of parameters needed to properly run through the current version. If you are certain that you have all the correct parameters or want to try to run anyway, enter Y (case insensitive). However, be warned that there is a high probability that HAPPE will error out. If you do not want to run potentially incompatible parameters, enter N (case insensitive) and HAPPE will end the current run via error.

Change an existing parameter? [Y/N]

If any of the loaded parameters do not match what you want for the current run, you can change them at this step by entering Y (case insensitive). Otherwise, enter N (case insensitive) to continue the run.

If creating a new parameter set:

Low density data? [Y/N]

For HAPPE, low density data contains 1 to ~32 channels.

For high density data, always enter N (case insensitive).

Data type:

rest = Resting-state/baseline EEG

task = Task-related EEG

Input rest (case insensitive) for HAPINESS with resting-state/baseline EEG data.

File format:

1 = .mat (MATLAB file)

2 = .raw (Net Station simple binary)

3 = .set (EEGLAB format)

4 = .cdt (Neuroscan format)

5 = .mff (EGI format)

Enter the file format of your raw data. If you are uncertain, you can confirm by looking at the file extension of your raw data. Depending on your choice, you may need to follow a different set of prompts.

**Example:** 3

Acquisition layout type:

1 = EGI Geodesic Sensor Net

2 = EGI HydroCel Geodesic Sensor Net

3 = Neuroscan Quik-Cap

4 = Other

Select the acquisition layout used to collect your data. If your layout is not specifically listed, select 4 for other. Depending on your choice, you may need to follow a different set of prompts.

**Example:** 2

At this point, depending on your choice, you may be presented with a list of channel numbers associated with the selected acquisition layout and file format that HAPPE currently supports.

Number of channels:

Enter the number of channels/electrodes included in the layout. Make sure to use the number from the layout (e.g., 128 for a 128-channel EGI net), even if the data itself does not use or include all electrodes.

**Example:** 128

If your data is in .mat format:

What best describes your data?

net station = Data exported using Net Station with additional

information

matrix = MATLAB matrix of the data

For data exported in .mat format from Net Station, enter net station. When loaded into MATLAB, it should include information beyond the data matrix, such as the sampling rate. Otherwise, if the data is a MATLAB matrix of data without any additional information, enter matrix. Depending on your choice, you may need to follow a different set of prompts.

If your data is exported from Net Station:

Enter the potential EEG variable names, one at a time.

Press enter/return between each entry.

When you are finished, enter “done” (without quotations).

NOTE: variable names containing “segment” may cause issues.

For Net Station .mat files, enter the potential names of the variable that consists of the EEG data matrix. Between each entry, press your newline key (enter – Windows, return – Mac). When you are done entering potential variable names, enter done. If you don’t have any ideas of what your variable name could be, try loading the Net Station .mat file into an empty MATLAB workspace and checking the variables.

**Example:** Category\_1

If your data is a MATLAB matrix:

If your channel locations file is not included in HAPPE’s acquisition layouts:

Do you have a channel locations file for your data? [Y/N]

NOTE: A list of supported file types can be found in the HAPPE User Guide.

Unless HAPPE includes the channel locations for your acquisition layout by default, you will likely need to provide them for MATLAB matrixes in .mat format. Accepted channel locations file formats are those accepted by EEGLAB (Delorme & Makieg, 2004) and include:

1. .loc, .locs, .eloc – EEG polar coordinates
2. .ced – EEGLAB with polar, cartesian, and spherical
3. .sph – MATLAB spherical coordinates
4. .elc – Cartesian 3-D from EETrack
5. .elp – Polhemus Cartesian coordinates
6. .elp – BESA spherical coordinates
7. .xyz – MATLAB/EEGLAB Cartesian coordinates
8. .asc, .dat – Neuroscan Cartesian polar coordinates
9. .mat – Brainstorm channel locations
10. .sfp – BESA/EGI xyz Cartesian coordinates

If you do have a channel locations file,

Enter the name of the channel locations file, including the full path and file extension:

Enter the full path to where the channel locations file exists, followed by the name of the file, including its extension. The channel locations file does not need to be in a specific location on your computer or on an external drive, only that it is accessible.

**Example (Mac):** /Users/lgd/Desktop/

chanlocs.loc

**Example (Windows):** C:\Users\lgd\Documents\

chanlocs.loc

If you do not have a channel locations file,

HAPPE functionality is limited without channel locations. Continue anyway? [Y/N]

If you don’t have a channel locations file, you will still be able to run your data through HAPPE, granted with a limited set of steps. Without channel locations, you cannot filter to channels of interest, perform bad channel detection, interpolate bad channels, or re-reference your data. If this is acceptable, you can continue running HAPPE by entering Y (case insensitive). Otherwise, you can enter N (case insensitive) and HAPPE will stop running by throwing an error.

Do all your files share the same sampling rate? [Y/N]

If all your files share the same sampling rate, input Y (case insensitive). Otherwise, input N (case insensitive). Depending on your input, you may have to follow different prompts.

If your files do share a sampling rate:

Sampling rate:

Enter the sampling rate of all the files as an integer.

**Example:** 250

If your files do not share a sampling rate:

Enter the name of the file containing the sampling rates for each data file, including the path and file extension:

See the HAPPE User Guide for how this table should be formatted.

Enter the name of the file, including the full path and the file’s extension. The file does not have to be located in any particular location on your computer or on an external drive; it only has to be accessible. Additionally, the file could be in any format so long as it is readable as a table by MATLAB, but we recommend .csv or an Excel file.

An example for formatting this file is shown below. The first column should contain the file name while the second column should contain the sampling rate in Hz.

|  |  |
| --- | --- |
| samplefilename1.mat | 250 |
| samplefilename2.mat | 500 |

**Example (Mac):** /Users/lgd/Desktop/sampRate.csv

**Example (Windows):** C:\Users\lgd\Documents\

sampRate.csv

If your data is in .mff format:

Do you have any additional type fields besides “code”? [Y/N]

This requirement is a holdover from older EGI code. In most cases, there are no additional type field requirements and you can enter N. If you know any additional type fields, input Y. If you are uncertain, refer to your system.

If you do have additional type fields:

Enter your type field names, one at a time. When you have finished entering type fields, enter “done” (without quotations).

Enter the type fields you wish to include one at a time. When you are done entering type fields, enter done. There is no need to add code to the list, it will be included automatically.

**Example:** datafield

If your file has channel locations included or provided:

Select channels of interest:

all = Select all channels in the data

coi = Select a user-specified subset of channels

To process all possible channels included in each EEG file, enter all. If you wish to only process a subset of channels, enter coi.

If only processing a subset of channels:

Choose an option for entering channels:

include = Include ONLY the entered channel names

exclude = Include every channel EXCEPT the entered

channel names

To enter your channel subset, you have the option to select the channels of interest or the channels of disinterest. If you select include all the channel names that you enter in the next prompt will be the only channels included in processing. Otherwise, by selecting exclude, you will instead process only the channels *not* entered in the next prompt. It may be easier to choose to include or exclude based on which list of channels is shorter.

Enter channels, including the preceding letter, one at a time.

Press enter/return between each entry.

Examples:

E17

M1

When you have entered all channels, input “done” (without quotations).

Enter the channels you wish to include or exclude, depending on your previous choice, one at a time. You should enter the channel name exactly as it appears in the data, case sensitive. This means you should include the preceding letter, if applicable (example: E17). If you have questions about your channel names, refer to your acquisition layout. Don’t use quotations around the channel name. Between each entry, press your newline key (enter for Windows, return for Mac). When you are finished entering all channels in your list of inclusion/exclusion, enter done.

Frequency of electrical (line) noise in Hz:

USA data probably = 60; Otherwise, probably = 50

Enter the frequency of electrical noise at the site of data collection. For locations within the United States, the frequency is likely 60 Hz, and outside is likely 50 Hz. We’ve included a link to a list of frequencies by country, but you should independently confirm that you have the correct frequency: <https://www.oaktreeproducts.com/img/product/description/List%20of%20Worldwide%20AC%20Voltages.pdf>

Are there any additional frequencies (e.g., harmonics) to reduce? [Y/N]

In some cases, there may be prominent line noise at a frequency other than the electrical noise frequency. Often, but not always, these are the harmonics of the electrical frequency. If there is enough noise at other frequencies that it affects the data, you can choose to reduce line noise at those frequencies by entering Y, case insensitive. Otherwise, enter N, case insensitive.

If reducing line noise in additional frequencies:

Enter the frequencies, one at a time, to pass through noise reduction.

When finished entering frequencies, enter “done” (without quotations).

Enter the frequencies you want to reduce, aside from the electrical noise frequency, that you want to reduce, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering frequencies, enter done.

**Example:** 25

Line noise reduction method:

cleanline = Use Tim Mullen’s CleanLine

notch = Use a notch filter (COMING SOON)

For reducing line noise, you have two options: using Tim Mullen’s CleanLine EEGLAB or using a notch filter. At this time, only the CleanLine option is available, so entering notch will end the HAPPE run via an error.

If using CleanLine:

Use legacy or default line noise reduction?

default – Default method optimized in HAPPE v2

legacy – Method from HAPPE v1 (NOT RECOMMENDED)

The legacy line noise reduction method uses the CleanLine plugin for EEGLAB from a far less effective version. For this reason, we strongly recommend using the default to process your data as it uses an improved version of CleanLine to reduce line noise. If you want to use the legacy method for comparison analyses or to finish existing data processing, enter legacy. For all other situations, you should enter default.

Resample data? [Y/N]

If you want to resample your data, enter Y. This might occur if you have different sampling rates across files and want them to be the same. Otherwise, enter N to not resample your data. Both options are case insensitive.

If resampling your data:

HAPPE supports resampling to 250, 500, and 1000 Hz.

Resample frequency:

Currently, HAPPE restricts resampling to the following values: 250, 500, or 1000 Hz. Choose one of these frequencies by entering the value.

**Example:** 500

If your file has channel locations included or provided:

Perform bad channel detection? [Y/N]

If you want remove channels with high impedances, damage to the electrodes, insufficient scalp contact, and/or excessive movement or electromyographic (EMG) artifact throughout the recording from subsequent analyses, enter Y, case insensitive. To include all your channels of interest regardless of quality, enter N, case insensitive.

If rejecting bad channels:

Bad channel detection method:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED).

You can choose to run either the default or the legacy version of bad channel detection. The legacy version was used in the original HAPPE software, is outdates, and we no longer support it, thus it is not recommended. Despite this, to select this option, enter legacy. The default method uses a new method of bad channel detection using the Clean Rawdata function with optimized criteria. To use this optimized method, enter default.

Method of wavelet thresholding:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED)

The new wavelet method uses a soft empirical Bayesian level-dependent threshold for the wavelets. Wavelet family is coiflet (level 4). This method has been optimized on EEG data (see Lopez et al., in press; Monachino et al., in press). To use this method, enter default. The legacy method of wavelet thresholding uses a soft, global threshold for the wavelets. The wavelet family is coiflet (level 5) and the threshold multiplier is used to remove more high frequency noise. This method has not been optimized, so it is not recommended, but if you wish to use it, enter legacy.

Use ICLabel to reduce remaining muscle artifact in your data? [Y/N]

NOTE: This will drastically increase processing time. Recommended for files with significant muscle artifact.

Sometimes there may be a high level of electromyographic (EMG) artifact in the data that remains despite wavelet thresholding. In this case, you can use ICA to remove the remaining artifact. HAPPE uses ICLabel to classify independent components and to reject all artifacts labeled as muscle with at least 25% probability. Using this is only recommended when you know there is EMG artifact in the data that wavelet thresholding doesn’t handle as ICA takes a lot of time and this step can exponentially increase processing time. If you want to use this option, called MuscIL (for muscle ICLabel), enter Y (case insensitive). Otherwise, enter N (case insensitive).

Segment data? [Y/N]

To segment your data, enter Y. Otherwise, enter N. If you choose to segment your data, you may be asked to answer additional prompts to properly segment your data.

If segmenting your data:

Segment length, in SECONDS:

Enter the desired length for your segments, in seconds. Use a positive, non-zero number.

**Example:** 2

If channel locations are included or provided and the data is segmented:

Interpolate the specific channels data determined to be artifact/bad within each segment? [Y/N]

This option enables evaluation of channels within epochs to determine whether there is bad data present. Channels flagged with bad data for that segment will then have their data interpolated only for that segment, using FASTER (Nolan et al., 2010). If you wish to interpolate bad channels within segments, enter Y. Otherwise, enter N.

Perform segment rejection? [Y/N]

Users can choose to reject segments that are determined to still have artifact remaining after waveleting and MuscIL (if selected). Criteria for rejection includes a choice of joint-probability criteria, amplitude-based criteria, or a combination of the two. If you would like to reject segments, enter Y. Otherwise, enter N.

If performing segment rejection:

Choose a method of segment rejection:

amplitude = Amplitude criteria only

similarity = Segment similarity only

both = Both amplitude criteria and segment similarity

The first option is using amplitude criteria only: amplitude-based criteria sets a minimum and maximum signal amplitude as the artifact threshold, with segments being removed when their amplitude exceeds either threshold. To run this option alone, enter amplitude. The second option is segment similarity criteria, which considers how likely a segment’s activity is to be artifact-laden given the activity of other segments for that channel and other channels’ activity for the same segment and removes outlier segments. The assumption is that artifact segments should be rare relative to the rest of the data at this point in the processing stream. To run this option alone, enter similarity. You can also choose to run both criteria by entering both. Depending on your choice, you may need to follow additional prompts.

If using amplitude criteria or both criteria:

Minimum signal amplitude to use as the artifact threshold:

Enter the minimum signal amplitude used for segment rejection.

**Example:** -150

Maximum signal amplitude to use as the artifact threshold:

Enter the maximum signal amplitude used for segment rejection.

**Example:** 150

Use all channels or a region of interest for segment rejection?

all = All channels

roi = Region of interest

If you plan to analyze all user-specified channels in your dataset, enter all. If you have a specific region of interest that you will be analyzing enter roi. Depending on your choice you may need to answer additional prompts.

If rejecting channels based on a region of interest:

Enter the channels in the ROI, one at a time.

When you have finished entering all channels, enter “done” (without quotations).

Enter the channels that you wish to include in your region of interest one at a time. Press your newline key (enter for Windows, return for Mac) between each entry. If you have any questions about your channel names, refer to your acquisition layout. As with entering your initial channels of interest, enter the names exactly as they appear in the dataset and without any quotations. When you have finished entering channels, enter done.

**Example:** O1

If channel locations are included or provided:

Re-reference data? [Y/N]

If you want to re-reference your data, input Y. Otherwise, enter N.

If re-referencing:

Does your data contain a flatline or all zero reference channel? [Y/N]

Some datasets include a reference channel that is a flat line; in other words, each data point for that channel is zero. If your data includes such a channel, enter Y. Otherwise, enter N.

If your data includes a flatline reference channel:

Enter reference channel ID:

If unknown, press enter/return.

Enter the channel ID of your flatline reference channel exactly as it appears in the data. If you don’t enter the exact name, it can prevent HAPPE from processing correctly.

**Example:** Cz

Re-referencing type:

average = Average re-referencing

subset = Re-reference to a channel or subset of channels

rest = Re-reference using infinity with REST (Yao, 2001)

HAPPE offers three options for re-referencing: average re-referencing, re-referencing to a subset, or using REST. To use average re-referencing, input average. To re-reference to a single channel or to a subset of channels, enter subset. To re-reference to infinity using REST, enter rest. Depending on your choice, you may need to answer additional prompts.

If re-referencing to a subset of channels:

Enter channel/subset of channels to re-reference to, one at a time.

When you have entered all channels, input “done” (without quotations).

Enter the channels that you wish to re-reference to one at a time. As with entering your initial channels of interest, the names should be entered exactly as they appear in the data. If you have questions about your channel names, refer to your acquisition layout. Between each channel entry, press your newline key (enter for Windows, return for Mac). When you finish entering channels, enter done.

**Example:** M1

If re-referencing using REST:

Reference type:

average = Average reference

fixed = Fixed reference

mastoid = Ipsilateral and/or contralateral mastoid

Select whether your original reference is average, fixed (as in a single electrode), or mastoid (ipsilateral and/or contralateral mastoid). If you have questions about which option to use for your dataset, please refer to the REST documentation.

If your reference selection is mastoid:

Enter list of left channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the left side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T7

Enter list of right channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the right side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T8

Load or calculate leadfield?

load = Load existing leadfield

calculate = Calculate new leadfield

If you have an existing leadfield that you want to use, enter load. Otherwise, if you don’t have a leadfield, you can have one calculated as part of the REST process by entering calculate. Depending on your choice, you may need to answer additional prompts.

If calculating the leadfield and reference selection is mastoid:

Enter XYZ coordinates for the left reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the left reference (often this is M1) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-37.1 89.8 -68.3]

Enter XYZ coordinates for the right reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the right reference (often this is M2) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-36.9 -88.8 -69.2]

If loading an existing leadfield:

Enter the name of the leadfield file, including path and file extension:

Enter the full name of the leadfield file including the path and file extension, examples below. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/

leadfield.dat

**Example (Windows):** C:\Users\lgd\Documents\

leadfield.dat

Format to save processed data:

1 = .txt file (electrodes as columns, time as rows) – Choose this for

ERP timeseries

2 = .mat file (MATLAB format)

3 = .set file (EEGLAB format)

Select your preferred format in which to save your processed data. For resting state data, you will likely select .mat (2) or .set (3) format.

Run HAPPE with visualizations? [Y/N]

HAPPE has the option of creating visualizations as intermediate methods of validating the processing stream on your data. By choosing Y (case insensitive), you can run HAPPE in the semi-automated setting with several visualizations for every file. We recommend running this mode on a couple of files to validate your parameter settings. Once you have optimized your parameters for your data, you can run your files in batch without interruption by entering N (case insensitive).

If visualizations are enabled:

Minimum value for power spectrum figure:

Enter the minimum frequency to plot for the power spectrum figure.

**Example:** 1

Maximum value for power spectrum figure:

Enter the maximum frequency to plot for the power spectrum figure.

**Example:** 45

Enter the frequencies, one at time, to generate spatial topoplots for:

When you have entered all frequencies, input “done” (without quotations).

Enter any frequencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the power spectrum for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the frequencies you wish to plot, enter done.

**Example:** 10

Please check your parameters before continuing.

At this point, your parameters are listed out in the command window.

Are the above parameters correct? [Y/N]

Go through the parameters listed in the command window and check to see that your parameters are correct. If so, enter Y (case insensitive). Otherwise, if your parameters are not correct, enter N (case insensitive) to change them.

If changing or correcting a parameter:

Parameter to change:

HAPPE will list out the parameters you can change.

Choose a parameter to change by entering the name of the parameter exactly as it appears in the above prompt (excluding newline breaks). You will be prompted with the original command to change the parameter. Depending on your data and parameter set, you may be presented with different options. You can change as many parameters as you want but must change each parameter one at a time. Some selections may require you to answer more than one prompt. When you finish entering parameters, enter done.

**Example:** segment rejection

If you changed an existing parameter set or created a new set:

Parameter file save name:

default = Default name (inputParameters\_dd-mm-yyyy.mat)

custom = Create a custom file name

To create your own save name for the parameter set, input custom. You will then be prompted to enter your custom save name. If you do not want to create a custom label, you can use the default name, inputParameters\_dd-mm-yyyy.mat (using the current date), by entering default.

If using a custom name:

File name (Do not include .mat):

Enter a name for your saved parameter set. Name the file the same way you would if you were using your file explorer; no need to include the .mat extension as it will be automatically added.

**Example:** baselineProcessing\_studyname

If a file with the parameter set name exists:

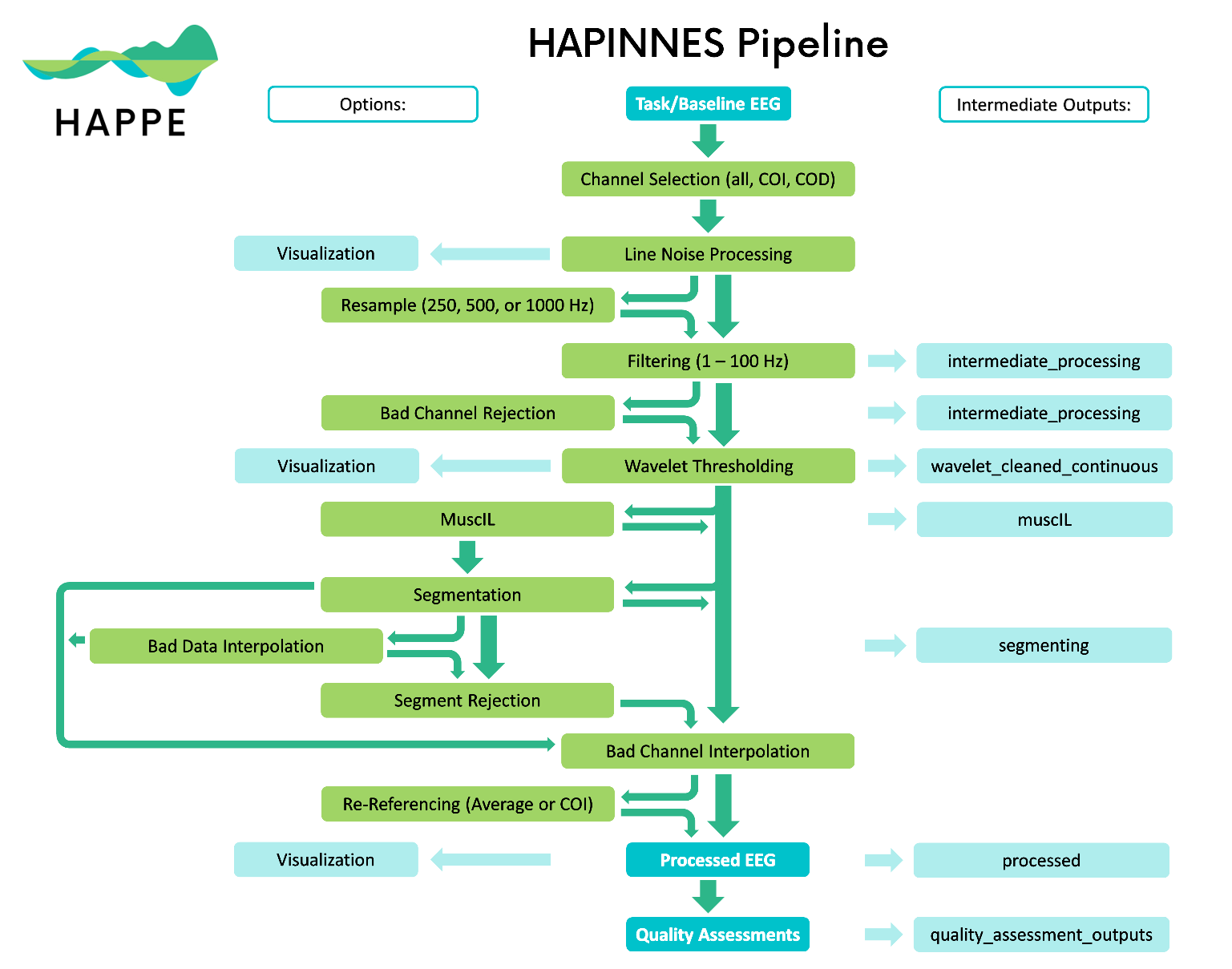
A set of input parameters with this name already exists.

overwrite = Overwrite the existing file

new = Save a new file with a different name

Entering multiple parameters sets on the same day or using the same name for a parameter set can result in multiple files vying for the same file name. The MATLAB default is to overwrite existing files, and if you want to replace the existing file with the new file, you can do so by entering overwrite. However, if you do not want to overwrite the file, enter new.

## HAPINNES Task-Related



### Following Command Line Prompts

Prompts as they appear in the command window are written in a monospace font, like this, followed by a brief description of what to enter, with an example. Sometimes a certain choice will result in a different set of prompts, in which case they will be indented under a heading that describes the choice resulting in the branch in directions.

Enter the path to the folder containing the dataset(s):

Input the folder, including the full path, where the raw data is located. If running HAPPE for the first time on this data, the folder should ONLY include the raw files you wish to process; no other folders or files should live inside that folder. If reprocessing data, all outputs (folders and documents) should be in the folder in the same file structure as they were created during the original HAPPE run in addition to the raw data.

**Example (Mac):** /Users/laurelg-d/Desktop/Data Folder

**Example (PC):** C:\Users\laurelg-d\Documents\Data Folder

Select an option:

raw = Run on raw data from the start

reprocess = Run on HAPPE-processed data starting post-artifact reduction

If this is your first time processing the raw data file(s), input raw. If you wish to re-run raw data that has previously been processed, starting at the MuscIL or segmentation step, input reprocess. Depending on your choice, you may be asked to follow additional prompts (see below).

If reprocessing existing data:

Files, such as processed data and quality metrics may already exist for this dataset.

overwrite = Overwrite existing files

new = Save new files

To keep both the already existing files and the newly created files, input new. If you do not want to keep previous files for this dataset, input overwrite.

If saving new files:

Use a default or custom label for processed data?

default = Default name (\_rerun\_dd-mm-yyyy)

custom = Create your own file label, starting with an

underscore

To create your own save name for the processed data, input custom. You will then be prompted to enter your custom label. This label should start with an underscore (“\_”) to encourage best naming practices. If you do not want to create a custom label, you can use the default (“\_rerun\_dd-mm-yyyy”), using the current date) by entering default. In either case, this will save the files as the original raw file name with the label added. Caution: if you already have data with this label, HAPPE will overwrite the existing files.

Load pre-existing set of input parameters? [Y/N]

If you have already set parameters through HAPPE for either this or another dataset that support the current dataset, you can choose to load these parameters instead of entering them from scratch. To use a previously saved set, enter Y (case insensitive); otherwise, enter N (case insensitive).

If loading pre-existing parameters:

Enter your parameter file, including the full path and file extension:

Enter the full name of the parameter file including the path and file extension, examples below. By default, parameters are saved in a folder named **input\_parameters**. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/Data Folder/input\_parameters/

inputParameters\_dd-mm-yyyy.mat

**Example (PC):** C:\Users\lgd\Documents\Data Folder\input\_parameters\

inputParameters\_dd-mm-yyyy.mat

If running a different version of HAPPE than the parameters were saved with:

These parameters were saved using a different version of HAPPE and ay be incompatible with the running pipeline.

Proceed anyway? [Y/N]

If your parameters were saved using a different version of HAPPE, there is a chance that they do not have the full set of parameters needed to properly run through the current version. If you are certain that you have all the correct parameters or want to try to run anyway, enter Y (case insensitive). However, be warned that there is a high probability that HAPPE will error out. If you do not want to run potentially incompatible parameters, enter N (case insensitive) and HAPPE will end the current run via error.

Change an existing parameter? [Y/N]

If any of the loaded parameters do not match what you want for the current run, you can change them at this step by entering Y (case insensitive). Otherwise, enter N (case insensitive) to continue the run.

If creating a new parameter set:

Low density data? [Y/N]

For HAPPE, low density data contains 1 to ~32 channels.

For high density data, always enter N (case insensitive).

Data type:

rest = Resting-state/baseline EEG

task = Task-related EEG

Input task (case insensitive) for HAPINESS with task-related, non-ERP EEG data.

Performing event-related potential (ERP) analysis? [Y/N]

Input N (case insensitive) for HAPINESS with task-related, non-ERP EEG data.

Enter the task onset tags, one at a time, pressing enter/return between each entry. When you have entered all tags, input “done” (without quotations).

Your data should include tags that indicate the onset of each trial in the dataset. Any stimuli/event/onset tags that EEGLAB can detect can be used in HAPPE. To verify that your tags are readable, load a data file into EEGLAB using the GUI and verify that your events are present. For detailed guidance on loading data into EEGLAB’s GUI and checking/importing events, we recommend the EEGLAB tutorial “Importing events and channel locations” found on YouTube and a written tutorial found here: <https://eeglab.org/tutorials/04_Import/Importing_Event_Epoch_Info.html>.

If you entered more than one onset tag:

Do multiple onset tags belong to a single condition? [Y/N]

Example: “happy\_face” and “sad\_face” belong to “faces”.

Sometimes, you may have multiple tags for different types of trials within a single condition. For example, you may have two conditions: faces and objects. Within the faces condition you have trials marked differently for happy faces, sad faces, and neutral faces. This function allows you combine the different kinds of trials for the faces condition into a single EEG to be output alongside the individual trial EEGs and the full EEG after processing. You do not need to enter all the tags within a condition of you do not want to, for example, only including happy faces and sad faces in your faces condition. If you have conditions and wish to have trials considered this way, enter Y. Otherwise, if you do not have conditions or do not wish to examine your data on the level of conditions, enter N.

If multiple tags belong to a single condition:

Enter the conditions and included onset tags:

Enter each condition as a list, each element separated by a blank space, with the condition name as the first item in the list. Press enter/return between entries. When you have entered all conditions, input “done” (without quotations).

Example: faces happy\_face sad\_face angry\_face

For each condition you must enter the condition name and all the onset tags associated with that condition. This should be entered as a list, starting with the condition name, which you may choose at this time. The onset tags should be entered as they appear in your data. Separate each item in the list using only a blank space, no commas. If you include a tag name not included in your original set of onset tags (see above), it may not be included or processed correctly, so ensure that you enter all onset tags when completing that step. You do not need to enter all the tags within a condition of you do not want to. For example, you may only include happy faces and sad faces in your faces condition when the condition may also have neutral faces. When you have completed this list, press your newline key to submit the condition and to enter a new one. When you finish entering conditions, enter done.

**Example:** condition tag1 tag2

File format:

1 = .mat (MATLAB file)

2 = .raw (Net Station simple binary)

3 = .set (EEGLAB format)

4 = .cdt (Neuroscan format)

5 = .mff (EGI format)

This is the file format of your raw data. Depending on your choice, you may need to follow a different set of prompts.

**Example:** 2

Acquisition layout net type:

1 – EGI Geodesic Sensor Net

2 – EGI HydroCel Geodesic Sensor Net

3 – Neuroscan Quik-Cap

4 – Other

Select the type of acquisition layout used to collect the data. Different choices may result in different prompts.

**Example:** 2

Number of channels:

The number of channels included in the layout. Make sure to use the number from the layout, even if your data does not use or include all electrodes.

**Example:** 128

If your data is in .mat format:

What best describes your data?

net station = Data exported using Net Station with additional

information

matrix = MATLAB matrix of the data

For data exported in .mat format from Net Station, enter net station. When loaded into MATLAB, it should include information beyond the data matrix, such as the sampling rate. Otherwise, if the data is a MATLAB matrix of data without any additional information, enter matrix. Depending on your choice, you may need to follow a different set of prompts.

If your data is exported from Net Station:

Enter the potential EEG variable names, one at a time.

Press enter/return between each entry.

When you are finished, enter “done” (without quotations).

NOTE: variable names containing “segment” may cause issues.

For Net Station .mat files, enter the potential names of the variable that consists of the EEG data matrix. Between each entry, press your newline key (enter – Windows, return – Mac). When you are done entering potential variable names, enter done. If you don’t have any ideas of what your variable name could be, try loading the Net Station .mat file into an empty MATLAB workspace and checking the variables.

**Example:** Category\_1

If your data is a MATLAB matrix:

If your channel locations file is not included in HAPPE’s acquisition layouts:

Do you have a channel locations file for your data? [Y/N]

NOTE: A list of supported file types can be found in the HAPPE User Guide.

Unless HAPPE includes the channel locations for your acquisition layout by default, you will likely need to provide them for MATLAB matrixes in .mat format. Accepted channel locations file formats are those accepted by EEGLAB (Delorme & Makieg, 2004) and include:

1. .loc, .locs, .eloc – EEG polar coordinates
2. .ced – EEGLAB with polar, cartesian, and spherical
3. .sph – MATLAB spherical coordinates
4. .elc – Cartesian 3-D from EETrack
5. .elp – Polhemus Cartesian coordinates
6. .elp – BESA spherical coordinates
7. .xyz – MATLAB/EEGLAB Cartesian coordinates
8. .asc, .dat – Neuroscan Cartesian polar coordinates
9. .mat – Brainstorm channel locations
10. .sfp – BESA/EGI xyz Cartesian coordinates

If you do have a channel locations file,

Enter the name of the channel locations file, including the full path and file extension:

Enter the full path to where the channel locations file exists, followed by the name of the file, including its extension. The channel locations file does not need to be in a specific location on your computer or on an external drive, only that it is accessible.

**Example (Mac):** /Users/lgd/Desktop/

chanlocs.loc

**Example (Windows):** C:\Users\lgd\Documents\

chanlocs.loc

If you do not have a channel locations file,

HAPPE functionality is limited without channel locations. Continue anyway? [Y/N]

If you don’t have a channel locations file, you will still be able to run your data through HAPPE, granted with a limited set of steps. Without channel locations, you cannot filter to channels of interest, perform bad channel detection, interpolate bad channels, or re-reference your data. If this is acceptable, you can continue running HAPPE by entering Y (case insensitive). Otherwise, you can enter N (case insensitive) and HAPPE will stop running by throwing an error.

Do all your files share the same sampling rate? [Y/N]

If all your files share the same sampling rate, input Y (case insensitive). Otherwise, input N (case insensitive). Depending on your input, you may have to follow different prompts.

If your files do share a sampling rate:

Sampling rate:

Enter the sampling rate of all the files as an integer.

**Example:** 250

If your files do not share a sampling rate:

Enter the name of the file containing the sampling rates for each data file, including the path and file extension:

See the HAPPE User Guide for how this table should be formatted.

Enter the name of the file, including the full path and the file’s extension. The file does not have to be in any particular location on your computer or on an external drive; it only must be accessible. Additionally, the file could be in any format so long as it is readable as a table by MATLAB, but we recommend .csv or an Excel file.

An example for formatting this file is shown below. The first column should contain the file name while the second column should contain the sampling rate in Hz.

|  |  |
| --- | --- |
| samplefilename1.mat | 250 |
| samplefilename2.mat | 500 |

**Example (Mac):** /Users/lgd/Desktop/sampRate.csv

**Example (Windows):** C:\Users\lgd\Documents\

sampRate.csv

Path to .txt files containing task event info:

For .mat files, you will need to provide .txt files containing your event information. Enter the full path to the folder holding these files, including the name of the folder. This folder does not need to be on the same path as your data.

**Example (Mac):** /Users/lgd/Desktop/Task InfoFolder

**Example (PC):** C:\Users\lgd\Documents\Task Info Folder

If your data is in .mff format:

Do you have any additional type fields besides “code”? [Y/N]

This requirement is a holdover from older EGI code. In most cases, there are no additional type field requirements and you can enter N. If you know any additional type fields, input Y. If you are uncertain, refer to your system.

If you do have additional type fields:

Enter your type field names, one at a time. When you have finished entering type fields, enter “done” (without quotations).

Enter the type fields you wish to include one at a time. When you are done entering type fields, enter done. There is no need to add code to the list, it will be included automatically.

**Example:** datafield

If your file has channel locations included or provided:

Select channels of interest:

all = Select all channels in the data

coi = Select a user-specified subset of channels

To process all possible channels included in each EEG file, enter all. If you wish to only process a subset of channels, enter coi.

If only processing a subset of channels:

Choose an option for entering channels:

include = Include ONLY the entered channel names

exclude = Include every channel EXCEPT the entered

channel names

To enter your channel subset, you have the option to select the channels of interest or the channels of disinterest. If you select include all the channel names that you enter in the next prompt will be the only channels included in processing. Otherwise, by selecting exclude, you will instead process only the channels *not* entered in the next prompt. It may be easier to choose to include or exclude based on which list of channels is shorter.

Enter channels, including the preceding letter, one at a time.

Press enter/return between each entry.

Examples:

E17

M1

When you have entered all channels, input “done” (without quotations).

Enter the channels you wish to include or exclude, depending on your previous choice, one at a time. You should enter the channel name exactly as it appears in the data, case sensitive. This means you should include the preceding letter, if applicable (example: E17). If you have questions about your channel names, refer to your acquisition layout. Don’t use quotations around the channel name. Between each entry, press your newline key (enter for Windows, return for Mac). When you are finished entering all channels in your list of inclusion/exclusion, enter done.

Frequency of electrical (line) noise in Hz:

USA data probably = 60; Otherwise, probably = 50

Enter the frequency of electrical noise at the site of data collection. For locations within the United States, the frequency is likely 60 Hz, and outside is likely 50 Hz. We’ve included a link to a list of frequencies by country, but you should independently confirm that you have the correct frequency: <https://www.oaktreeproducts.com/img/product/description/List%20of%20Worldwide%20AC%20Voltages.pdf>

Are there any additional frequencies (e.g., harmonics) to reduce? [Y/N]

In some cases, there may be prominent line noise at a frequency other than the electrical noise frequency. Often, but not always, these are the harmonics of the electrical frequency. If there is enough noise at other frequencies that it affects the data, you can choose to reduce line noise at those frequencies by entering Y, case insensitive. Otherwise, enter N, case insensitive.

If reducing line noise in additional frequencies:

Enter the frequencies, one at a time, to pass through noise reduction.

When finished entering frequencies, enter “done” (without quotations).

Enter the frequencies you want to reduce, aside from the electrical noise frequency, that you want to reduce, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering frequencies, enter done.

**Example:** 25

Line noise reduction method:

cleanline = Use Tim Mullen’s CleanLine

notch = Use a notch filter (COMING SOON)

For reducing line noise, you have two options: using Tim Mullen’s CleanLine EEGLAB or using a notch filter. At this time, only the CleanLine option is available, so entering notch will end the HAPPE run via an error.

If using CleanLine:

Use legacy or default line noise reduction?

default – Default method optimized in HAPPE v2

legacy – Method from HAPPE v1 (NOT RECOMMENDED)

The legacy line noise reduction method uses the CleanLine plugin for EEGLAB from a far less effective version. For this reason, we strongly recommend using the default to process your data as it uses an improved version of CleanLine to reduce line noise. If you want to use the legacy method for comparison analyses or to finish existing data processing, enter legacy. For all other situations, you should enter default.

Resample data? [Y/N]

If you want to resample your data, enter Y. This might occur if you have different sampling rates across files and want them to be the same. Otherwise, enter N to not resample your data. Both options are case insensitive.

If resampling your data:

HAPPE supports resampling to 250, 500, and 1000 Hz.

Resample frequency:

Currently, HAPPE restricts resampling to the following values: 250, 500, or 1000 Hz. Choose one of these frequencies by entering the value.

**Example:** 500

If your file has channel locations included or provided:

Perform bad channel detection? [Y/N]

If you want remove channels with high impedances, damage to the electrodes, insufficient scalp contact, and/or excessive movement or electromyographic (EMG) artifact throughout the recording from subsequent analyses, enter Y, case insensitive. To include all your channels of interest regardless of quality, enter N, case insensitive.

If rejecting bad channels:

Bad channel detection method:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED).

You can choose to run either the default or the legacy version of bad channel detection. The legacy version was used in the original HAPPE software, is outdates, and we no longer support it, thus it is not recommended. Despite this, to select this option, enter legacy. The default method uses a new method of bad channel detection using the Clean Rawdata function with optimized criteria. To use this optimized method, enter default.

Method of wavelet thresholding:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED)

The new wavelet method uses a soft empirical Bayesian level-dependent threshold for the wavelets. Wavelet family is coiflet (level 4). This method has been optimized on EEG data (see Lopez et al., in press; Monachino et al., in press). To use this method, enter default. The legacy method of wavelet thresholding uses a soft, global threshold for the wavelets. The wavelet family is coiflet (level 5) and the threshold multiplier is used to remove more high frequency noise. This method has not been optimized, so it is not recommended, but if you wish to use it, enter legacy.

Use ICLabel to reduce remaining muscle artifact in your data? [Y/N]

NOTE: This will drastically increase processing time. Recommended for files with significant muscle artifact.

Sometimes there may be a high level of electromyographic (EMG) artifact in the data that remains despite wavelet thresholding. In this case, you can use ICA to remove the remaining artifact. HAPPE uses ICLabel to classify independent components and to reject all artifacts labeled as muscle with at least 25% probability. Using this is only recommended when you know there is EMG artifact in the data that wavelet thresholding doesn’t handle as ICA takes a lot of time and this step can exponentially increase processing time. If you want to use this option, called MuscIL (for muscle ICLabel), enter Y (case insensitive). Otherwise, enter N (case insensitive).

Segment data? [Y/N]

To segment your data, enter Y. Otherwise, enter N. If you choose to segment your data, you may be asked to answer additional prompts to properly segment your data.

If segmenting your data:

Segment start, in MILLISECONDS, relative to stimulus onset:

Example: -100

Enter the number of milliseconds present in the trial prior to your stimulus onset. If you have no baseline period within your trial, this may be 0. Otherwise, this is a negative number. For example, if you have a trial with a 100-millisecond baseline, you will enter -100.

Segment end, in MILLISECONDS, relative to stimulus onset:

Enter the number of milliseconds present in the trial after stimulus onset. This should always be a number greater than 0. For example, if you have a trial that runs for 500 milliseconds after the stimulus onset or trial start, you will enter 500.

**Example:** 500

If channel locations are included or provided and the data is segmented:

Interpolate the specific channels data determined to be artifact/bad within each segment? [Y/N]

This option enables evaluation of channels within epochs to determine whether there is bad data present. Channels flagged with bad data for that segment will then have their data interpolated only for that segment, using FASTER (Nolan et al., 2010). If you wish to interpolate bad channels within segments, enter Y. Otherwise, enter N.

Perform segment rejection? [Y/N]

Users can choose to reject segments that are determined to still have artifact remaining after waveleting and MuscIL (if selected). Criteria for rejection includes a choice of joint-probability criteria, amplitude-based criteria, or a combination of the two. If you would like to reject segments, enter Y. Otherwise, enter N.

If performing segment rejection:

Choose a method of segment rejection:

amplitude = Amplitude criteria only

similarity = Segment similarity only

both = Both amplitude criteria and segment similarity

The first option is using amplitude criteria only: amplitude-based criteria sets a minimum and maximum signal amplitude as the artifact threshold, with segments being removed when their amplitude exceeds either threshold. To run this option alone, enter amplitude. The second option is segment similarity criteria, which considers how likely a segment’s activity is to be artifact-laden given the activity of other segments for that channel and other channels’ activity for the same segment and removes outlier segments. The assumption is that artifact segments should be rare relative to the rest of the data at this point in the processing stream. To run this option alone, enter similarity. You can also choose to run both criteria by entering both. Depending on your choice, you may need to follow additional prompts.

If using amplitude criteria or both criteria:

Minimum signal amplitude to use as the artifact threshold:

Enter the minimum signal amplitude used for segment rejection.

**Example:** -150

Maximum signal amplitude to use as the artifact threshold:

Enter the maximum signal amplitude used for segment rejection.

**Example:** 150

Use all channels or a region of interest for segment rejection?

all = All channels

roi = Region of interest

If you plan to analyze all user-specified channels in your dataset, enter all. If you have a specific region of interest that you will be analyzing enter roi. Depending on your choice you may need to answer additional prompts.

If rejecting channels based on a region of interest:

Enter the channels in the ROI, one at a time.

When you have finished entering all channels, enter “done” (without quotations).

Enter the channels that you wish to include in your region of interest one at a time. Press your newline key (enter for Windows, return for Mac) between each entry. If you have any questions about your channel names, refer to your acquisition layout. As with entering your initial channels of interest, enter the names exactly as they appear in the dataset and without any quotations. When you have finished entering channels, enter done.

**Example:** O1

Use pre-selected “usable” trials to restrict analysis? [Y/N]

If you have previously determined which trials in your dataset are usable, you can choose to filter out the “unusable” trials using this option. THIS OPTION DOES NOT CURRENTLY FUNCTION!

If channel locations are included or provided:

Re-reference data? [Y/N]

If you want to re-reference your data, input Y. Otherwise, enter N.

If re-referencing:

Does your data contain a flatline or all zero reference channel? [Y/N]

Some datasets include a reference channel that is a flat line; in other words, each data point for that channel is zero. If your data includes such a channel, enter Y. Otherwise, enter N.

If your data includes a flatline reference channel:

Enter reference channel ID:

If unknown, press enter/return.

Enter the channel ID of your flatline reference channel exactly as it appears in the data. If you don’t enter the exact name, it can prevent HAPPE from processing correctly.

**Example:** Cz

Re-referencing type:

average = Average re-referencing

subset = Re-reference to a channel or subset of channels

rest = Re-reference using infinity with REST (Yao, 2001)

HAPPE offers three options for re-referencing: average re-referencing, re-referencing to a subset, or using REST. To use average re-referencing, input average. To re-reference to a single channel or to a subset of channels, enter subset. To re-reference to infinity using REST, enter rest. Depending on your choice, you may need to answer additional prompts.

If re-referencing to a subset of channels:

Enter channel/subset of channels to re-reference to, one at a time.

When you have entered all channels, input “done” (without quotations).

Enter the channels that you wish to re-reference to one at a time. As with entering your initial channels of interest, the names should be entered exactly as they appear in the data. If you have questions about your channel names, refer to your acquisition layout. Between each channel entry, press your newline key (enter for Windows, return for Mac). When you finish entering channels, enter done.

**Example:** M1

If re-referencing using REST:

Reference type:

average = Average reference

fixed = Fixed reference

mastoid = Ipsilateral and/or contralateral mastoid

Select whether your original reference is average, fixed (as in a single electrode), or mastoid (ipsilateral and/or contralateral mastoid). If you have questions about which option to use for your dataset, please refer to the REST documentation.

If your reference selection is mastoid:

Enter list of left channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the left side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T7

Enter list of right channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the right side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T8

Load or calculate leadfield?

load = Load existing leadfield

calculate = Calculate new leadfield

If you have an existing leadfield that you want to use, enter load. Otherwise, if you don’t have a leadfield, you can have one calculated as part of the REST process by entering calculate. Depending on your choice, you may need to answer additional prompts.

If calculating the leadfield and reference selection is mastoid:

Enter XYZ coordinates for the left reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the left reference (often this is M1) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-37.1 89.8 -68.3]

Enter XYZ coordinates for the right reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the right reference (often this is M2) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-36.9 -88.8 -69.2]

If loading an existing leadfield:

Enter the name of the leadfield file, including path and file extension:

Enter the full name of the leadfield file including the path and file extension, examples below. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/

leadfield.dat

**Example (Windows):** C:\Users\lgd\Documents\

leadfield.dat

Format to save processed data:

1 = .txt file (electrodes as columns, time as rows) – Choose this for

ERP timeseries

2 = .mat file (MATLAB format)

3 = .set file (EEGLAB format)

Select your preferred format in which to save your processed data. For continuous task data, you will likely select .mat (2) or .set (3) format, but you may also select .txt (1).

Run HAPPE with visualizations? [Y/N]

HAPPE has the option of creating visualizations as intermediate methods of validating the processing stream on your data. By choosing Y (case insensitive), you can run HAPPE in the semi-automated setting with several visualizations for every file. We recommend running this mode on a couple of files to validate your parameter settings. Once you have optimized your parameters for your data, you can run your files in batch without interruption by entering N (case insensitive).

If visualizations are enabled:

Minimum value for power spectrum figure:

Enter the minimum frequency to plot for the power spectrum figure.

**Example:** 1

Maximum value for power spectrum figure:

Enter the maximum frequency to plot for the power spectrum figure.

**Example:** 45

Enter the frequencies, one at time, to generate spatial topoplots for:

When you have entered all frequencies, input “done” (without quotations).

Enter any frequencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the power spectrum for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the frequencies you wish to plot, enter done.

**Example:** 10

Please check your parameters before continuing.

At this point, your parameters are listed out in the command window.

Are the above parameters correct? [Y/N]

Go through the parameters listed in the command window and check to see that your parameters are correct. If so, enter Y (case insensitive). Otherwise, if your parameters are not correct, enter N (case insensitive) to change them.

If changing or correcting a parameter:

Parameter to change:

HAPPE will list out the parameters you can change.

Choose a parameter to change by entering the name of the parameter exactly as it appears in the above prompt (excluding newline breaks). You will be prompted with the original command to change the parameter. Depending on your data and parameter set, you may be presented with different options. You can change as many parameters as you want but must change each parameter one at a time. Some selections may require you to answer more than one prompt. When you finish entering parameters, enter done.

**Example:** segment rejection

If you changed an existing parameter set or created a new set:

Parameter file save name:

default = Default name (inputParameters\_dd-mm-yyyy.mat)

custom = Create a custom file name

To create your own save name for the parameter set, input custom. You will then be prompted to enter your custom save name. If you do not want to create a custom label, you can use the default name, inputParameters\_dd-mm-yyyy.mat (using the current date), by entering default.

If using a custom name:

File name (Do not include .mat):

Enter a name for your saved parameter set. Name the file the same way you would if you were using your file explorer; no need to include the .mat extension as it will be automatically added.

**Example:** taskProcessing\_studyname

If a file with the parameter set name exists:

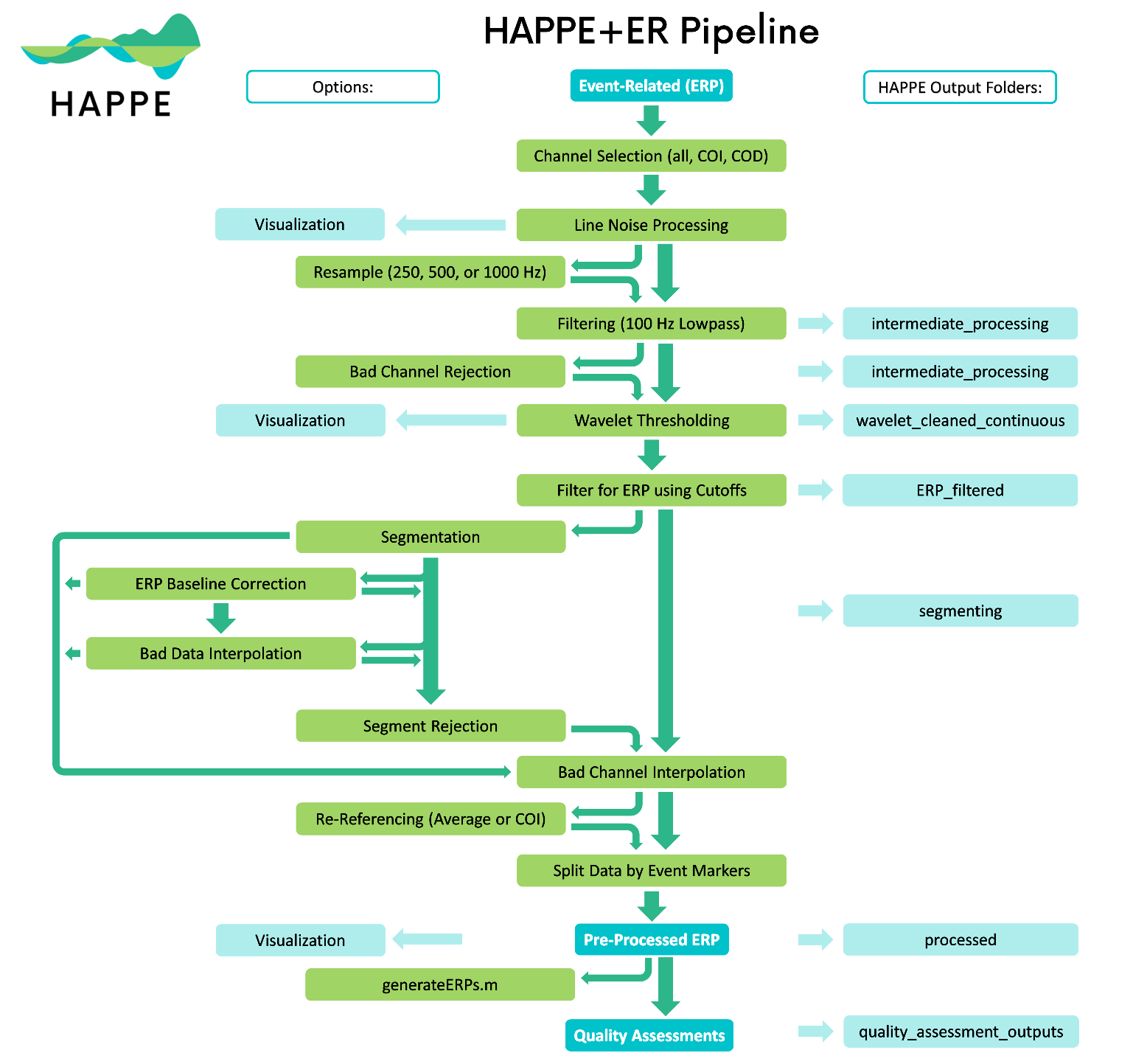
A set of input parameters with this name already exists.

overwrite = Overwrite the existing file

new = Save a new file with a different name

Entering multiple parameters sets on the same day or using the same name for a parameter set can result in multiple files vying for the same file name. The MATLAB default is to overwrite existing files, and if you want to replace the existing file with the new file, you can do so by entering overwrite. However, if you do not want to overwrite the file, enter new.

## HAPPE+ER



### Following Command Line Prompts

Prompts as they appear in the command window are written in a monospace font, like this, followed by a brief description of what to enter, with an example. Sometimes a certain choice will result in a different set of prompts, in which case they will be indented under a heading that describes the choice resulting in the branch in directions.

Enter the path to the folder containing the dataset(s):

Input the folder, including the full path, where the raw data is located. If running HAPPE for the first time on this data, the folder should ONLY include the raw files you wish to process; no other folders or files should live inside that folder. If reprocessing data, all outputs (folders and documents) should be in the folder in the same file structure as they were created during the original HAPPE run in addition to the raw data.

**Example (Mac):** /Users/laurelg-d/Desktop/Data Folder

**Example (PC):** C:\Users\laurelg-d\Documents\Data Folder

Select an option:

raw = Run on raw data from the start

reprocess = Run on HAPPE-processed data starting post-artifact reduction

If this is your first time processing the raw data file(s), input raw. If you wish to re-run raw data that has previously been processed, starting at the MuscIL or segmentation step, input reprocess. Depending on your choice, you may be asked to follow additional prompts (see below).

If reprocessing existing data:

Files, such as processed data and quality metrics may already exist for this dataset.

overwrite = Overwrite existing files

new = Save new files

To keep both the already existing files and the newly created files, input new. If you do not want to keep previous files for this dataset, input overwrite.

If saving new files:

Use a default or custom label for processed data?

default = Default name (\_rerun\_dd-mm-yyyy)

custom = Create your own file label, starting with an

underscore

To create your own save name for the processed data, input custom. You will then be prompted to enter your custom label. This label should start with an underscore (“\_”) to encourage best naming practices. If you do not want to create a custom label, you can use the default (“\_rerun\_dd-mm-yyyy”), using the current date) by entering default. In either case, this will save the files as the original raw file name with the label added. Caution: if you already have data with this label, HAPPE will overwrite the existing files.

Load pre-existing set of input parameters? [Y/N]

If you have already set parameters through HAPPE for either this or another dataset that support the current dataset, you can choose to load these parameters instead of entering them from scratch. To use a previously saved set, enter Y (case insensitive); otherwise, enter N (case insensitive).

If loading pre-existing parameters:

Enter your parameter file, including the full path and file extension:

Enter the full name of the parameter file including the path and file extension, examples below. By default, parameters are saved in a folder named **input\_parameters**. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/Data Folder/input\_parameters/

inputParameters\_dd-mm-yyyy.mat

**Example (PC):** C:\Users\lgd\Documents\Data Folder\input\_parameters\

inputParameters\_dd-mm-yyyy.mat

If running a different version of HAPPE than the parameters were saved with:

These parameters were saved using a different version of HAPPE and ay be incompatible with the running pipeline.

Proceed anyway? [Y/N]

If your parameters were saved using a different version of HAPPE, there is a chance that they do not have the full set of parameters needed to properly run through the current version. If you are certain that you have all the correct parameters or want to try to run anyway, enter Y (case insensitive). However, be warned that there is a high probability that HAPPE will error out. If you do not want to run potentially incompatible parameters, enter N (case insensitive) and HAPPE will end the current run via error.

Change an existing parameter? [Y/N]

If any of the loaded parameters do not match what you want for the current run, you can change them at this step by entering Y (case insensitive). Otherwise, enter N (case insensitive) to continue the run.

If creating a new parameter set:

Low density data? [Y/N]

For HAPPE, low density data contains 1 to ~32 channels.

For high density data, always enter N (case insensitive).

Data type:

rest = Resting-state/baseline EEG

task = Task-related EEG

Input task (case insensitive) for HAPINESS with task-related, non-ERP EEG data.

Performing event-related potential (ERP) analysis? [Y/N]

Input Y (case insensitive) to run the HAPPE+ER pipeline on ERP EEG data.

Enter the task onset tags, one at a time, pressing enter/return between each entry. When you have entered all tags, input “done” (without quotations).

Your data should include tags that indicate the onset of each trial in the dataset. Any stimuli/event/onset tags that EEGLAB can detect can be used in HAPPE. To verify that your tags are readable, load a data file into EEGLAB using the GUI and verify that your events are present. For detailed guidance on loading data into EEGLAB’s GUI and checking/importing events, we recommend the EEGLAB tutorial “Importing events and channel locations” found on YouTube and a written tutorial found here: <https://eeglab.org/tutorials/04_Import/Importing_Event_Epoch_Info.html>.

If you entered more than one onset tag:

Do multiple onset tags belong to a single condition? [Y/N]

Example: “happy\_face” and “sad\_face” belong to “faces”.

Sometimes, you may have multiple tags for different types of trials within a single condition. For example, you may have two conditions: faces and objects. Within the faces condition you have trials marked differently for happy faces, sad faces, and neutral faces. This function allows you combine the different kinds of trials for the faces condition into a single EEG to be output alongside the individual trial EEGs and the full EEG after processing. You do not need to enter all the tags within a condition of you do not want to, for example, only including happy faces and sad faces in your faces condition. If you have conditions and wish to have trials considered this way, enter Y. Otherwise, if you do not have conditions or do not wish to examine your data on the level of conditions, enter N.

If multiple tags belong to a single condition:

Enter the conditions and included onset tags:

Enter each condition as a list, each element separated by a blank space, with the condition name as the first item in the list. Press enter/return between entries. When you have entered all conditions, input “done” (without quotations).

Example: faces happy\_face sad\_face angry\_face

For each condition you must enter the condition name and all the onset tags associated with that condition. This should be entered as a list, starting with the condition name, which you may choose at this time. The onset tags should be entered as they appear in your data. Separate each item in the list using only a blank space, no commas. If you include a tag name not included in your original set of onset tags (see above), it may not be included or processed correctly, so ensure that you enter all onset tags when completing that step. You do not need to enter all the tags within a condition of you do not want to. For example, you may only include happy faces and sad faces in your faces condition when the condition may also have neutral faces. When you have completed this list, press your newline key to submit the condition and to enter a new one. When you finish entering conditions, enter done.

**Example:** condition tag1 tag2

Enter low-pass filter, in Hz:

Common low-pass filter is 30 – 45 Hz.

Enter the low-pass filter at which to filter your data for ERP analyses.

**Example:** 35

Enter high-pass filter, in Hz:

Common high-pass filter is 0.1 – 0.3 Hz.

Enter the high-pass filter at which to filter your data for ERP analyses.

**Example:** 0.1

File format:

1 = .mat (MATLAB file)

2 = .raw (Net Station simple binary)

3 = .set (EEGLAB format)

4 = .cdt (Neuroscan format)

5 = .mff (EGI format)

This is the file format of your raw data. Depending on your choice, you may need to follow a different set of prompts.

**Example:** 2

Acquisition layout net type:

1 – EGI Geodesic Sensor Net

2 – EGI HydroCel Geodesic Sensor Net

3 – Neuroscan Quik-Cap

4 – Other

Select the type of acquisition layout used to collect the data. Different choices may result in different prompts.

**Example:** 2

Number of channels:

The number of channels included in the layout. Make sure to use the number from the layout, even if your data does not use or include all electrodes.

**Example:** 128

If your data is in .mat format:

What best describes your data?

net station = Data exported using Net Station with additional

information

matrix = MATLAB matrix of the data

For data exported in .mat format from Net Station, enter net station. When loaded into MATLAB, it should include information beyond the data matrix, such as the sampling rate. Otherwise, if the data is a MATLAB matrix of data without any additional information, enter matrix. Depending on your choice, you may need to follow a different set of prompts.

If your data is exported from Net Station:

Enter the potential EEG variable names, one at a time.

Press enter/return between each entry.

When you are finished, enter “done” (without quotations).

NOTE: variable names containing “segment” may cause issues.

For Net Station .mat files, enter the potential names of the variable that consists of the EEG data matrix. Between each entry, press your newline key (enter – Windows, return – Mac). When you are done entering potential variable names, enter done. If you don’t have any ideas of what your variable name could be, try loading the Net Station .mat file into an empty MATLAB workspace and checking the variables.

**Example:** Category\_1

If your data is a MATLAB matrix:

If your channel locations file is not included in HAPPE’s acquisition layouts:

Do you have a channel locations file for your data? [Y/N]

NOTE: A list of supported file types can be found in the HAPPE User Guide.

Unless HAPPE includes the channel locations for your acquisition layout by default, you will likely need to provide them for MATLAB matrixes in .mat format. Accepted channel locations file formats are those accepted by EEGLAB (Delorme & Makieg, 2004) and include:

1. .loc, .locs, .eloc – EEG polar coordinates
2. .ced – EEGLAB with polar, cartesian, and spherical
3. .sph – MATLAB spherical coordinates
4. .elc – Cartesian 3-D from EETrack
5. .elp – Polhemus Cartesian coordinates
6. .elp – BESA spherical coordinates
7. .xyz – MATLAB/EEGLAB Cartesian coordinates
8. .asc, .dat – Neuroscan Cartesian polar coordinates
9. .mat – Brainstorm channel locations
10. .sfp – BESA/EGI xyz Cartesian coordinates

If you do have a channel locations file,

Enter the name of the channel locations file, including the full path and file extension:

Enter the full path to where the channel locations file exists, followed by the name of the file, including its extension. The channel locations file does not need to be in a specific location on your computer or on an external drive, only that it is accessible.

**Example (Mac):** /Users/lgd/Desktop/

chanlocs.loc

**Example (Windows):** C:\Users\lgd\Documents\

chanlocs.loc

If you do not have a channel locations file,

HAPPE functionality is limited without channel locations. Continue anyway? [Y/N]

If you don’t have a channel locations file, you will still be able to run your data through HAPPE, granted with a limited set of steps. Without channel locations, you cannot filter to channels of interest, perform bad channel detection, interpolate bad channels, or re-reference your data. If this is acceptable, you can continue running HAPPE by entering Y (case insensitive). Otherwise, you can enter N (case insensitive) and HAPPE will stop running by throwing an error.

Do all your files share the same sampling rate? [Y/N]

If all your files share the same sampling rate, input Y (case insensitive). Otherwise, input N (case insensitive). Depending on your input, you may have to follow different prompts.

If your files do share a sampling rate:

Sampling rate:

Enter the sampling rate of all the files as an integer.

**Example:** 250

If your files do not share a sampling rate:

Enter the name of the file containing the sampling rates for each data file, including the path and file extension:

See the HAPPE User Guide for how this table should be formatted.

Enter the name of the file, including the full path and the file’s extension. The file does not have to be in any particular location on your computer or on an external drive; it only must be accessible. Additionally, the file could be in any format so long as it is readable as a table by MATLAB, but we recommend .csv or an Excel file.

An example for formatting this file is shown below. The first column should contain the file name while the second column should contain the sampling rate in Hz.

|  |  |
| --- | --- |
| samplefilename1.mat | 250 |
| samplefilename2.mat | 500 |

**Example (Mac):** /Users/lgd/Desktop/sampRate.csv

**Example (Windows):** C:\Users\lgd\Documents\

sampRate.csv

Path to .txt files containing task event info:

For .mat files, you will need to provide .txt files containing your event information. Enter the full path to the folder holding these files, including the name of the folder. This folder does not need to be on the same path as your data.

**Example (Mac):** /Users/lgd/Desktop/Task InfoFolder

**Example (PC):** C:\Users\lgd\Documents\Task Info Folder

If your data is in .mff format:

Do you have any additional type fields besides “code”? [Y/N]

This requirement is a holdover from older EGI code. In most cases, there are no additional type field requirements and you can enter N. If you know any additional type fields, input Y. If you are uncertain, refer to your system.

If you do have additional type fields:

Enter your type field names, one at a time. When you have finished entering type fields, enter “done” (without quotations).

Enter the type fields you wish to include one at a time. When you are done entering type fields, enter done. There is no need to add code to the list, it will be included automatically.

**Example:** datafield

If your file has channel locations included or provided:

Select channels of interest:

all = Select all channels in the data

coi = Select a user-specified subset of channels

To process all possible channels included in each EEG file, enter all. If you wish to only process a subset of channels, enter coi.

If only processing a subset of channels:

Choose an option for entering channels:

include = Include ONLY the entered channel names

exclude = Include every channel EXCEPT the entered

channel names

To enter your channel subset, you have the option to select the channels of interest or the channels of disinterest. If you select include all the channel names that you enter in the next prompt will be the only channels included in processing. Otherwise, by selecting exclude, you will instead process only the channels *not* entered in the next prompt. It may be easier to choose to include or exclude based on which list of channels is shorter.

Enter channels, including the preceding letter, one at a time.

Press enter/return between each entry.

Examples:

E17

M1

When you have entered all channels, input “done” (without quotations).

Enter the channels you wish to include or exclude, depending on your previous choice, one at a time. You should enter the channel name exactly as it appears in the data, case sensitive. This means you should include the preceding letter, if applicable (example: E17). If you have questions about your channel names, refer to your acquisition layout. Don’t use quotations around the channel name. Between each entry, press your newline key (enter for Windows, return for Mac). When you are finished entering all channels in your list of inclusion/exclusion, enter done.

Frequency of electrical (line) noise in Hz:

USA data probably = 60; Otherwise, probably = 50

Enter the frequency of electrical noise at the site of data collection. For locations within the United States, the frequency is likely 60 Hz, and outside is likely 50 Hz. We’ve included a link to a list of frequencies by country, but you should independently confirm that you have the correct frequency: <https://www.oaktreeproducts.com/img/product/description/List%20of%20Worldwide%20AC%20Voltages.pdf>

Are there any additional frequencies (e.g., harmonics) to reduce? [Y/N]

In some cases, there may be prominent line noise at a frequency other than the electrical noise frequency. Often, but not always, these are the harmonics of the electrical frequency. If there is enough noise at other frequencies that it affects the data, you can choose to reduce line noise at those frequencies by entering Y, case insensitive. Otherwise, enter N, case insensitive.

If reducing line noise in additional frequencies:

Enter the frequencies, one at a time, to pass through noise reduction.

When finished entering frequencies, enter “done” (without quotations).

Enter the frequencies you want to reduce, aside from the electrical noise frequency, that you want to reduce, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering frequencies, enter done.

**Example:** 25

Line noise reduction method:

cleanline = Use Tim Mullen’s CleanLine

notch = Use a notch filter (COMING SOON)

For reducing line noise, you have two options: using Tim Mullen’s CleanLine EEGLAB or using a notch filter. At this time, only the CleanLine option is available, so entering notch will end the HAPPE run via an error.

If using CleanLine:

Use legacy or default line noise reduction?

default – Default method optimized in HAPPE v2

legacy – Method from HAPPE v1 (NOT RECOMMENDED)

The legacy line noise reduction method uses the CleanLine plugin for EEGLAB from a far less effective version. For this reason, we strongly recommend using the default to process your data as it uses an improved version of CleanLine to reduce line noise. If you want to use the legacy method for comparison analyses or to finish existing data processing, enter legacy. For all other situations, you should enter default.

Resample data? [Y/N]

If you want to resample your data, enter Y. This might occur if you have different sampling rates across files and want them to be the same. Otherwise, enter N to not resample your data. Both options are case insensitive.

If resampling your data:

HAPPE supports resampling to 250, 500, and 1000 Hz.

Resample frequency:

Currently, HAPPE restricts resampling to the following values: 250, 500, or 1000 Hz. Choose one of these frequencies by entering the value.

**Example:** 500

Choose a filter:

fir = Hamming windowed sinc FIR filter (EEGLAB’s standard filter)

butter = IRR butterworth filter (ERPLAB’s standard filter)

You may choose to filter your ERP data to your ERP cutoffs using one of two options. The FIR filter is what is used for all non-ERP designs in HAPPE and is EEGLAB’s standard filter. To use this option, enter fir. However, you may also instead choose to filter using a butterworth filter. In this case, HAPPE will use ERPLAB’s IRR butterworth filter by entering butter.

If your file has channel locations included or provided:

Perform bad channel detection? [Y/N]

If you want remove channels with high impedances, damage to the electrodes, insufficient scalp contact, and/or excessive movement or electromyographic (EMG) artifact throughout the recording from subsequent analyses, enter Y, case insensitive. To include all your channels of interest regardless of quality, enter N, case insensitive.

If rejecting bad channels:

Bad channel detection method:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED).

You can choose to run either the default or the legacy version of bad channel detection. The legacy version was used in the original HAPPE software, is outdates, and we no longer support it, thus it is not recommended. Despite this, to select this option, enter legacy. The default method uses a new method of bad channel detection using the Clean Rawdata function with optimized criteria. To use this optimized method, enter default.

Method of wavelet thresholding:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED)

The new wavelet method uses an empirical Bayesian level-dependent threshold for the wavelets. Whether this is a soft or a hard threshold can be determined by the user in following prompts. Wavelet family is coiflet (level 4). This method has been optimized on EEG data (see Lopez et al., in press; Monachino et al., in press). To use this method, enter default. The legacy method of wavelet thresholding uses a soft, global threshold for the wavelets. The wavelet family is coiflet (level 5) and the threshold multiplier is used to remove more high frequency noise. This method has not been optimized, so it is not recommended, but if you wish to use it, enter legacy.

If using the default method for wavelet thresholding:

Threshold rule for wavelet thresholding:

soft – Use a soft threshold

hard – Use a hard threshold

HAPPE+ER allows the user to choose whether to use a hard or soft threshold for wavelet thresholding in ERP designs. Using a soft threshold slightly preserves ERP amplitude but may keep slightly more artifact in the data. Using a hard threshold removes more artifact but at the slight cost of ERP amplitude. For more information about the difference, refer to the HAPPE+ER manuscript. To use the soft threshold, enter soft. Otherwise, enter hard.

Segment data? [Y/N]

To segment your data, enter Y. Otherwise, enter N. If you choose to segment your data, you may be asked to answer additional prompts to properly segment your data.

If segmenting your data:

Segment start, in MILLISECONDS, relative to stimulus onset:

Example: -100

Enter the number of milliseconds present in the trial prior to your stimulus onset. If you have no baseline period within your trial, this may be 0. Otherwise, this is a negative number. For example, if you have a trial with a 100-millisecond baseline, you will enter -100.

Segment end, in MILLISECONDS, relative to stimulus onset:

Enter the number of milliseconds present in the trial after stimulus onset. This should always be a number greater than 0. For example, if you have a trial that runs for 500 milliseconds after the stimulus onset or trial start, you will enter 500.

**Example:** 500

Offset delay, in MILLISECONDS, between stimulus initiation and presentation:

NOTE: Please enter the total offset (combined system and task-specific offsets).

Enter the total offset delay between when the stimulus is initiated and when it is presented. This number should include the system offset as well as any task-specific offsets and be entered in milliseconds. Currently, HAPPE only supports a single offset value.

**Example:** 20

Perform baseline correction (by subtraction)? [Y/N]

If you would like to perform baseline correction on your data, input Y. Otherwise, input N. If you choose to perform baseline correction, you may be shown additional prompts.

If performing baseline correction:

Enter, in MILLISECONDS, where the baseline segment begins:

Example: -100

The number of milliseconds prior to the event tag where the baseline period starts. Entering a negative number will start the baseline period that number of milliseconds before the event tag. Entering 0 starts the segment at the event tag.

Enter, in MILLISECONDS, where the baseline segment ends:

NOTE: 0 indicates stimulus onset.

The number of milliseconds in the segment, relative to the onset tag, where the baseline period ends. Entering 0 will end the baseline period at stimulus onset.

If channel locations are included or provided and the data is segmented:

Interpolate the specific channels data determined to be artifact/bad within each segment? [Y/N]

This option enables evaluation of channels within epochs to determine whether there is bad data present. Channels flagged with bad data for that segment will then have their data interpolated only for that segment, using FASTER (Nolan et al., 2010). If you wish to interpolate bad channels within segments, enter Y. Otherwise, enter N.

Perform segment rejection? [Y/N]

Users can choose to reject segments that are determined to still have artifact remaining after waveleting and MuscIL (if selected). Criteria for rejection includes a choice of joint-probability criteria, amplitude-based criteria, or a combination of the two. If you would like to reject segments, enter Y. Otherwise, enter N.

If performing segment rejection:

Choose a method of segment rejection:

amplitude = Amplitude criteria only

similarity = Segment similarity only

both = Both amplitude criteria and segment similarity

The first option is using amplitude criteria only: amplitude-based criteria sets a minimum and maximum signal amplitude as the artifact threshold, with segments being removed when their amplitude exceeds either threshold. To run this option alone, enter amplitude. The second option is segment similarity criteria, which considers how likely a segment’s activity is to be artifact-laden given the activity of other segments for that channel and other channels’ activity for the same segment and removes outlier segments. The assumption is that artifact segments should be rare relative to the rest of the data at this point in the processing stream. To run this option alone, enter similarity. You can also choose to run both criteria by entering both. Depending on your choice, you may need to follow additional prompts.

If using amplitude criteria or both criteria:

Minimum signal amplitude to use as the artifact threshold:

Enter the minimum signal amplitude used for segment rejection.

**Example:** -150

Maximum signal amplitude to use as the artifact threshold:

Enter the maximum signal amplitude used for segment rejection.

**Example:** 150

Use all channels or a region of interest for segment rejection?

all = All channels

roi = Region of interest

If you plan to analyze all user-specified channels in your dataset, enter all. If you have a specific region of interest that you will be analyzing enter roi. Depending on your choice you may need to answer additional prompts.

If rejecting channels based on a region of interest:

Enter the channels in the ROI, one at a time.

When you have finished entering all channels, enter “done” (without quotations).

Enter the channels that you wish to include in your region of interest one at a time. Press your newline key (enter for Windows, return for Mac) between each entry. If you have any questions about your channel names, refer to your acquisition layout. As with entering your initial channels of interest, enter the names exactly as they appear in the dataset and without any quotations. When you have finished entering channels, enter done.

**Example:** O1

Use pre-selected “usable” trials to restrict analysis? [Y/N]

If you have previously determined which trials in your dataset are usable, you can choose to filter out the “unusable” trials using this option. THIS OPTION DOES NOT CURRENTLY FUNCTION!

If channel locations are included or provided:

Re-reference data? [Y/N]

If you want to re-reference your data, input Y. Otherwise, enter N.

If re-referencing:

Does your data contain a flatline or all zero reference channel? [Y/N]

Some datasets include a reference channel that is a flat line; in other words, each data point for that channel is zero. If your data includes such a channel, enter Y. Otherwise, enter N.

If your data includes a flatline reference channel:

Enter reference channel ID:

If unknown, press enter/return.

Enter the channel ID of your flatline reference channel exactly as it appears in the data. If you don’t enter the exact name, it can prevent HAPPE from processing correctly.

**Example:** Cz

Re-referencing type:

average = Average re-referencing

subset = Re-reference to a channel or subset of channels

rest = Re-reference using infinity with REST (Yao, 2001)

HAPPE offers three options for re-referencing: average re-referencing, re-referencing to a subset, or using REST. To use average re-referencing, input average. To re-reference to a single channel or to a subset of channels, enter subset. To re-reference to infinity using REST, enter rest. Depending on your choice, you may need to answer additional prompts.

If re-referencing to a subset of channels:

Enter channel/subset of channels to re-reference to, one at a time.

When you have entered all channels, input “done” (without quotations).

Enter the channels that you wish to re-reference to one at a time. As with entering your initial channels of interest, the names should be entered exactly as they appear in the data. If you have questions about your channel names, refer to your acquisition layout. Between each channel entry, press your newline key (enter for Windows, return for Mac). When you finish entering channels, enter done.

**Example:** M1

If re-referencing using REST:

Reference type:

average = Average reference

fixed = Fixed reference

mastoid = Ipsilateral and/or contralateral mastoid

Select whether your original reference is average, fixed (as in a single electrode), or mastoid (ipsilateral and/or contralateral mastoid). If you have questions about which option to use for your dataset, please refer to the REST documentation.

If your reference selection is mastoid:

Enter list of left channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the left side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T7

Enter list of right channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the right side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T8

Load or calculate leadfield?

load = Load existing leadfield

calculate = Calculate new leadfield

If you have an existing leadfield that you want to use, enter load. Otherwise, if you don’t have a leadfield, you can have one calculated as part of the REST process by entering calculate. Depending on your choice, you may need to answer additional prompts.

If calculating the leadfield and reference selection is mastoid:

Enter XYZ coordinates for the left reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the left reference (often this is M1) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-37.1 89.8 -68.3]

Enter XYZ coordinates for the right reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the right reference (often this is M2) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-36.9 -88.8 -69.2]

If loading an existing leadfield:

Enter the name of the leadfield file, including path and file extension:

Enter the full name of the leadfield file including the path and file extension, examples below. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/

leadfield.dat

**Example (Windows):** C:\Users\lgd\Documents\

leadfield.dat

Format to save processed data:

1 = .txt file (electrodes as columns, time as rows) – Choose this for

ERP timeseries

2 = .mat file (MATLAB format)

3 = .set file (EEGLAB format)

For processing ERPs, we strongly recommend entering 1, which will provide .txt files of the timeseries as well as .set versions of the EEG. However, if you wish, you may also select either 2 or 3.

If visualizations are enabled:

Minimum value for power spectrum figure:

Enter the minimum frequency to plot for the power spectrum figure.

**Example:** 1

Maximum value for power spectrum figure:

Enter the maximum frequency to plot for the power spectrum figure.

**Example:** 45

Enter the frequencies, one at time, to generate spatial topoplots for:

When you have entered all frequencies, input “done” (without quotations).

Enter any frequencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the power spectrum for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the frequencies you wish to plot, enter done.

**Example:** 10

Start time, in MILLISECONDS, for the ERP timeseries figure:

Enter the start time, in milliseconds, to plot for the ERP timeseries figure.

**Example:** 0

End time, in MILLISECONDS, for the ERP timeseries figure:

NOTE: This should end 1+ millisecond(s) before your segmentation parameter ends (e.g., 299 for 300).

Enter the end time, in milliseconds, to plot for the ERP timeseries figure. Make sure that this ends before your segmentation parameters ends. For example, if your segment ends at 300 milliseconds, we recommend setting your value at 298 or 299.

**Example:** 498

Enter the latencies, one at a time, to generate spatial topoplots for:

When you have entered all latencies, input “done” (without quotations).

Enter any latencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the ERP timeseries for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the latencies you wish to plot, enter done.

**Example:** 100

Please check your parameters before continuing.

At this point, your parameters are listed out in the command window.

Are the above parameters correct? [Y/N]

Go through the parameters listed in the command window and check to see that your parameters are correct. If so, enter Y (case insensitive). Otherwise, if your parameters are not correct, enter N (case insensitive) to change them.

If changing or correcting a parameter:

Parameter to change:

HAPPE will list out the parameters you can change.

Choose a parameter to change by entering the name of the parameter exactly as it appears in the above prompt (excluding newline breaks). You will be prompted with the original command to change the parameter. Depending on your data and parameter set, you may be presented with different options. You can change as many parameters as you want but must change each parameter one at a time. Some selections may require you to answer more than one prompt. When you finish entering parameters, enter done.

**Example:** segment rejection

If you changed an existing parameter set or created a new set:

Parameter file save name:

default = Default name (inputParameters\_dd-mm-yyyy.mat)

custom = Create a custom file name

To create your own save name for the parameter set, input custom. You will then be prompted to enter your custom save name. If you do not want to create a custom label, you can use the default name, inputParameters\_dd-mm-yyyy.mat (using the current date), by entering default.

If using a custom name:

File name (Do not include .mat):

Enter a name for your saved parameter set. Name the file the same way you would if you were using your file explorer; no need to include the .mat extension as it will be automatically added.

**Example:** erpProcessing\_studyname

If a file with the parameter set name exists:

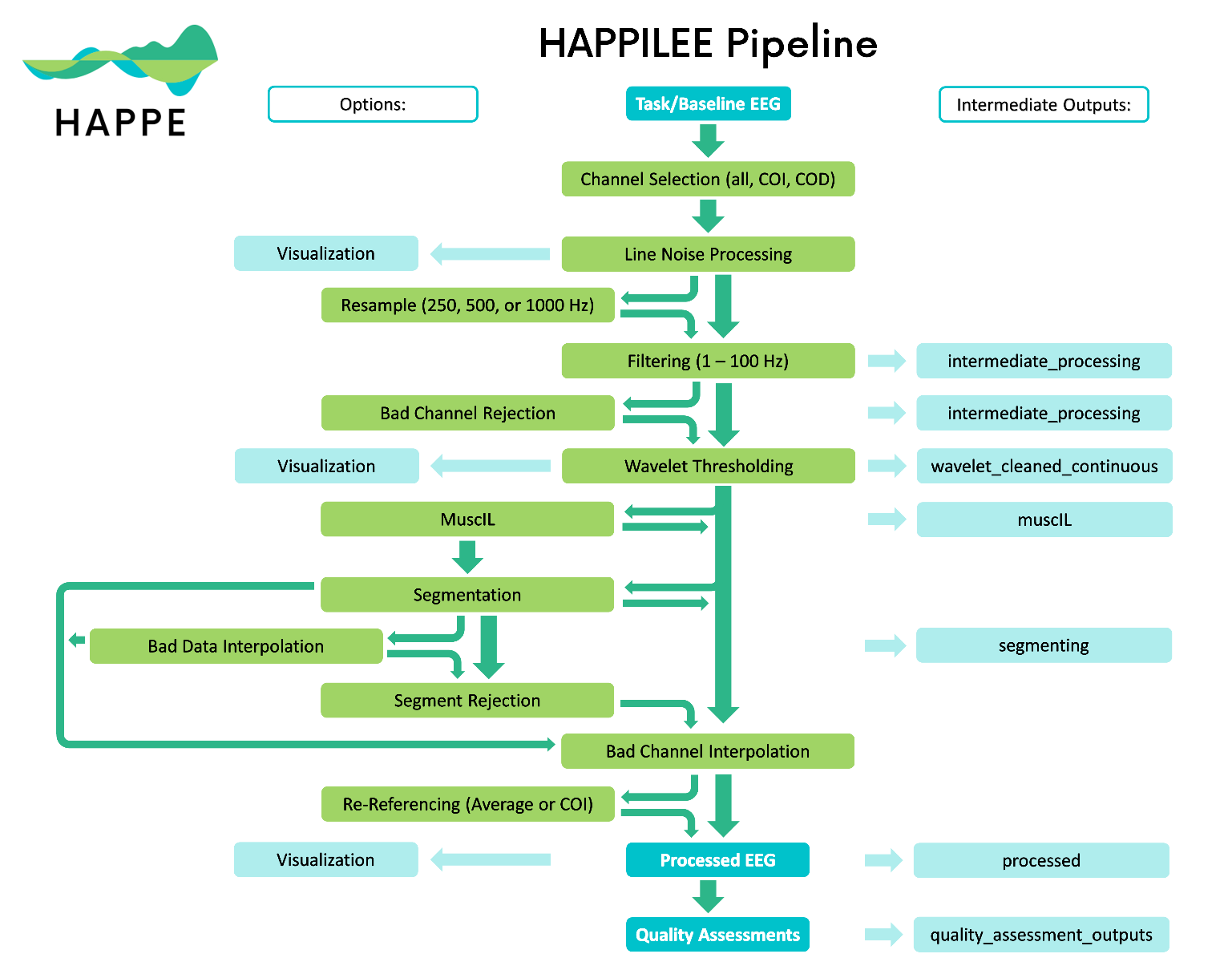
A set of input parameters with this name already exists.

overwrite = Overwrite the existing file

new = Save a new file with a different name

Entering multiple parameters sets on the same day or using the same name for a parameter set can result in multiple files vying for the same file name. The MATLAB default is to overwrite existing files, and if you want to replace the existing file with the new file, you can do so by entering overwrite. However, if you do not want to overwrite the file, enter new.

## HAPPILEE Resting-State



### Following Command Line Prompts

Prompts as they appear in the command window are written in a monospace font, like this, followed by a brief description of what to enter, with an example. Sometimes a certain choice will result in a different set of prompts, in which case they will be indented under a heading that describes the choice resulting in the branch in directions.

Enter the path to the folder containing the dataset(s):

Input the folder, including the full path, where the raw data is located. If running HAPPE for the first time on this data, the folder should ONLY include the raw files you wish to process; no other folders or files should live inside that folder. If reprocessing data, all outputs (folders and documents) should be in the folder in the same file structure as they were created during the original HAPPE run in addition to the raw data.

**Example (Mac):** /Users/laurelg-d/Desktop/Data Folder

**Example (PC):** C:\Users\laurelg-d\Documents\Data Folder

Select an option:

raw = Run on raw data from the start

reprocess = Run on HAPPE-processed data starting post-artifact reduction

If this is your first time processing the raw data file(s), input raw. If you wish to re-run raw data that has previously been processed, starting at the MuscIL or segmentation step, input reprocess. Depending on your choice, you may be asked to follow additional prompts (see below).

If reprocessing existing data:

Files, such as processed data and quality metrics may already exist for this dataset.

overwrite = Overwrite existing files

new = Save new files

To keep both the already existing files and the newly created files, input new. If you do not want to keep previous files for this dataset, input overwrite.

If saving new files:

Use a default or custom label for processed data?

default = Default name (\_rerun\_dd-mm-yyyy)

custom = Create your own file label, starting with an

underscore

To create your own save name for the processed data, input custom. You will then be prompted to enter your custom label. This label should start with an underscore (“\_”) to encourage best naming practices. If you do not want to create a custom label, you can use the default (“\_rerun\_dd-mm-yyyy”), using the current date) by entering default. In either case, this will save the files as the original raw file name with the label added. Caution: if you already have data with this label, HAPPE will overwrite the existing files.

Load pre-existing set of input parameters? [Y/N]

If you have already set parameters through HAPPE for either this or another dataset that support the current dataset, you can choose to load these parameters instead of entering them from scratch. To use a previously saved set, enter Y (case insensitive); otherwise, enter N (case insensitive).

If loading pre-existing parameters:

Enter your parameter file, including the full path and file extension:

Enter the full name of the parameter file including the path and file extension, examples below. By default, parameters are saved in a folder named **input\_parameters**. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/Data Folder/input\_parameters/

inputParameters\_dd-mm-yyyy.mat

**Example (PC):** C:\Users\lgd\Documents\Data Folder\input\_parameters\

inputParameters\_dd-mm-yyyy.mat

If running a different version of HAPPE than the parameters were saved with:

These parameters were saved using a different version of HAPPE and ay be incompatible with the running pipeline.

Proceed anyway? [Y/N]

If your parameters were saved using a different version of HAPPE, there is a chance that they do not have the full set of parameters needed to properly run through the current version. If you are certain that you have all the correct parameters or want to try to run anyway, enter Y (case insensitive). However, be warned that there is a high probability that HAPPE will error out. If you do not want to run potentially incompatible parameters, enter N (case insensitive) and HAPPE will end the current run via error.

Change an existing parameter? [Y/N]

If any of the loaded parameters do not match what you want for the current run, you can change them at this step by entering Y (case insensitive). Otherwise, enter N (case insensitive) to continue the run.

If creating a new parameter set:

Low density data? [Y/N]

For HAPPE, low density data contains 1 to ~32 channels.

For low density data, always enter Y (case insensitive).

Data type:

rest = Resting-state/baseline EEG

task = Task-related EEG

Input rest (case insensitive) for HAPPILEE with resting-state/baseline EEG data.

File format:

1 = .mat (MATLAB file)

2 = .raw (Net Station simple binary)

3 = .set (EEGLAB format)

4 = .cdt (Neuroscan format)

5 = .mff (EGI format)

Enter the file format of your raw data. If you are uncertain, you can confirm by looking at the file extension of your raw data. Depending on your choice, you may need to follow a different set of prompts. HAPPE does not currently support low-density data in formats other than .mat and .set.

**Example:** 3

Acquisition layout type:

1 = EGI Geodesic Sensor Net

2 = EGI HydroCel Geodesic Sensor Net

3 = Neuroscan Quik-Cap

4 = Other

Select the acquisition layout used to collect your data. If your layout is not specifically listed, select 4 for other. Depending on your choice, you may need to follow a different set of prompts.

**Example:** 2

At this point, depending on your choice, you may be presented with a list of channel numbers associated with the selected acquisition layout and file format that HAPPE currently supports.

Number of channels:

Enter the number of channels/electrodes included in the layout. Make sure to use the number from the layout (e.g., 16 for a 16-channel net), even if the data itself does not use or include all electrodes.

**Example:** 16

If your data is in .mat format:

What best describes your data?

net station = Data exported using Net Station with additional

information

matrix = MATLAB matrix of the data

For data exported in .mat format from Net Station, enter net station. When loaded into MATLAB, it should include information beyond the data matrix, such as the sampling rate. Otherwise, if the data is a MATLAB matrix of data without any additional information, enter matrix. Depending on your choice, you may need to follow a different set of prompts.

If your data is exported from Net Station:

Enter the potential EEG variable names, one at a time.

Press enter/return between each entry.

When you are finished, enter “done” (without quotations).

NOTE: variable names containing “segment” may cause issues.

For Net Station .mat files, enter the potential names of the variable that consists of the EEG data matrix. Between each entry, press your newline key (enter – Windows, return – Mac). When you are done entering potential variable names, enter done. If you don’t have any ideas of what your variable name could be, try loading the Net Station .mat file into an empty MATLAB workspace and checking the variables.

**Example:** Category\_1

If your data is a MATLAB matrix:

If your channel locations file is not included in HAPPE’s acquisition layouts:

Do you have a channel locations file for your data? [Y/N]

NOTE: A list of supported file types can be found in the HAPPE User Guide.

Unless HAPPE includes the channel locations for your acquisition layout by default, you will likely need to provide them for MATLAB matrixes in .mat format. Accepted channel locations file formats are those accepted by EEGLAB (Delorme & Makieg, 2004) and include:

1. .loc, .locs, .eloc – EEG polar coordinates
2. .ced – EEGLAB with polar, cartesian, and spherical
3. .sph – MATLAB spherical coordinates
4. .elc – Cartesian 3-D from EETrack
5. .elp – Polhemus Cartesian coordinates
6. .elp – BESA spherical coordinates
7. .xyz – MATLAB/EEGLAB Cartesian coordinates
8. .asc, .dat – Neuroscan Cartesian polar coordinates
9. .mat – Brainstorm channel locations
10. .sfp – BESA/EGI xyz Cartesian coordinates

If you do have a channel locations file,

Enter the name of the channel locations file, including the full path and file extension:

Enter the full path to where the channel locations file exists, followed by the name of the file, including its extension. The channel locations file does not need to be in a specific location on your computer or on an external drive, only that it is accessible.

**Example (Mac):** /Users/lgd/Desktop/

chanlocs.loc

**Example (Windows):** C:\Users\lgd\Documents\

chanlocs.loc

If you do not have a channel locations file,

HAPPE functionality is limited without channel locations. Continue anyway? [Y/N]

If you don’t have a channel locations file, you will still be able to run your data through HAPPE, granted with a limited set of steps. Without channel locations, you cannot filter to channels of interest, perform bad channel detection, interpolate bad channels, or re-reference your data. If this is acceptable, you can continue running HAPPE by entering Y (case insensitive). Otherwise, you can enter N (case insensitive) and HAPPE will stop running by throwing an error.

Do all your files share the same sampling rate? [Y/N]

If all your files share the same sampling rate, input Y (case insensitive). Otherwise, input N (case insensitive). Depending on your input, you may have to follow different prompts.

If your files do share a sampling rate:

Sampling rate:

Enter the sampling rate of all the files as an integer.

**Example:** 250

If your files do not share a sampling rate:

Enter the name of the file containing the sampling rates for each data file, including the path and file extension:

See the HAPPE User Guide for how this table should be formatted.

Enter the name of the file, including the full path and the file’s extension. The file does not have to be located in any particular location on your computer or on an external drive; it only has to be accessible. Additionally, the file could be in any format so long as it is readable as a table by MATLAB, but we recommend .csv or an Excel file.

An example for formatting this file is shown below. The first column should contain the file name while the second column should contain the sampling rate in Hz.

|  |  |
| --- | --- |
| samplefilename1.mat | 250 |
| samplefilename2.mat | 500 |

**Example (Mac):** /Users/lgd/Desktop/sampRate.csv

**Example (Windows):** C:\Users\lgd\Documents\

sampRate.csv

If your file has channel locations included or provided:

Select channels of interest:

all = Select all channels in the data

coi = Select a user-specified subset of channels

To process all possible channels included in each EEG file, enter all. If you wish to only process a subset of channels, enter coi.

If only processing a subset of channels:

Choose an option for entering channels:

include = Include ONLY the entered channel names

exclude = Include every channel EXCEPT the entered

channel names

To enter your channel subset, you have the option to select the channels of interest or the channels of disinterest. If you select include all the channel names that you enter in the next prompt will be the only channels included in processing. Otherwise, by selecting exclude, you will instead process only the channels *not* entered in the next prompt. It may be easier to choose to include or exclude based on which list of channels is shorter.

Enter channels, including the preceding letter, one at a time.

Press enter/return between each entry.

Examples:

E17

M1

When you have entered all channels, input “done” (without quotations).

Enter the channels you wish to include or exclude, depending on your previous choice, one at a time. You should enter the channel name exactly as it appears in the data, case sensitive. This means you should include the preceding letter, if applicable (example: E17). If you have questions about your channel names, refer to your acquisition layout. Don’t use quotations around the channel name. Between each entry, press your newline key (enter for Windows, return for Mac). When you are finished entering all channels in your list of inclusion/exclusion, enter done.

Frequency of electrical (line) noise in Hz:

USA data probably = 60; Otherwise, probably = 50

Enter the frequency of electrical noise at the site of data collection. For locations within the United States, the frequency is likely 60 Hz, and outside is likely 50 Hz. We’ve included a link to a list of frequencies by country, but you should independently confirm that you have the correct frequency: <https://www.oaktreeproducts.com/img/product/description/List%20of%20Worldwide%20AC%20Voltages.pdf>

Are there any additional frequencies (e.g., harmonics) to reduce? [Y/N]

In some cases, there may be prominent line noise at a frequency other than the electrical noise frequency. Often, but not always, these are the harmonics of the electrical frequency. If there is enough noise at other frequencies that it affects the data, you can choose to reduce line noise at those frequencies by entering Y, case insensitive. Otherwise, enter N, case insensitive.

If reducing line noise in additional frequencies:

Enter the frequencies, one at a time, to pass through noise reduction.

When finished entering frequencies, enter “done” (without quotations).

Enter the frequencies you want to reduce, aside from the electrical noise frequency, that you want to reduce, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering frequencies, enter done.

**Example:** 25

Line noise reduction method:

cleanline = Use Tim Mullen’s CleanLine

notch = Use a notch filter (COMING SOON)

For reducing line noise, you have two options: using Tim Mullen’s CleanLine EEGLAB or using a notch filter. At this time, only the CleanLine option is available, so entering notch will end the HAPPE run via an error.

If using CleanLine:

Use legacy or default line noise reduction?

default – Default method optimized in HAPPE v2

legacy – Method from HAPPE v1 (NOT RECOMMENDED)

The legacy line noise reduction method uses the CleanLine plugin for EEGLAB from a far less effective version. For this reason, we strongly recommend using the default to process your data as it uses an improved version of CleanLine to reduce line noise. If you want to use the legacy method for comparison analyses or to finish existing data processing, enter legacy. For all other situations, you should enter default.

Resample data? [Y/N]

If you want to resample your data, enter Y. This might occur if you have different sampling rates across files and want them to be the same. Otherwise, enter N to not resample your data. Both options are case insensitive.

If resampling your data:

HAPPE supports resampling to 250, 500, and 1000 Hz.

Resample frequency:

Currently, HAPPE restricts resampling to the following values: 250, 500, or 1000 Hz. Choose one of these frequencies by entering the value.

**Example:** 500

If your file has channel locations included or provided:

Perform bad channel detection? [Y/N]

If you want remove channels with high impedances, damage to the electrodes, insufficient scalp contact, and/or excessive movement or electromyographic (EMG) artifact throughout the recording from subsequent analyses, enter Y, case insensitive. To include all your channels of interest regardless of quality, enter N, case insensitive.

If rejecting bad channels:

Bad channel detection method:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED).

You can choose to run either the default or the legacy version of bad channel detection. The legacy version was used in the original HAPPE software, is outdates, and we no longer support it, thus it is not recommended. Despite this, to select this option, enter legacy. The default method uses a new method of bad channel detection using the Clean Rawdata function with optimized criteria. To use this optimized method, enter default.

Method of wavelet thresholding:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED)

The new wavelet method uses a soft empirical Bayesian level-dependent threshold for the wavelets. Wavelet family is coiflet (level 4). This method has been optimized on EEG data (see Lopez et al., in press; Monachino et al., in press). To use this method, enter default. The legacy method of wavelet thresholding uses a soft, global threshold for the wavelets. The wavelet family is coiflet (level 5) and the threshold multiplier is used to remove more high frequency noise. This method has not been optimized, so it is not recommended, but if you wish to use it, enter legacy.

Use ICLabel to reduce remaining muscle artifact in your data? [Y/N]

NOTE: This will drastically increase processing time. Recommended for files with significant muscle artifact.

Sometimes there may be a high level of electromyographic (EMG) artifact in the data that remains despite wavelet thresholding. In this case, you can use ICA to remove the remaining artifact. HAPPE uses ICLabel to classify independent components and to reject all artifacts labeled as muscle with at least 25% probability. Using this is only recommended when you know there is EMG artifact in the data that wavelet thresholding doesn’t handle as ICA takes a lot of time and this step can exponentially increase processing time. If you want to use this option, called MuscIL (for muscle ICLabel), enter Y (case insensitive). Otherwise, enter N (case insensitive).

Segment data? [Y/N]

To segment your data, enter Y. Otherwise, enter N. If you choose to segment your data, you may be asked to answer additional prompts to properly segment your data.

If segmenting your data:

Segment length, in SECONDS:

Enter the desired length for your segments, in seconds. Use a positive, non-zero number.

**Example:** 2

If channel locations are included or provided and the data is segmented:

Interpolate the specific channels data determined to be artifact/bad within each segment? [Y/N]

This option enables evaluation of channels within epochs to determine whether there is bad data present. Channels flagged with bad data for that segment will then have their data interpolated only for that segment, using FASTER (Nolan et al., 2010). If you wish to interpolate bad channels within segments, enter Y. Otherwise, enter N.

Perform segment rejection? [Y/N]

Users can choose to reject segments that are determined to still have artifact remaining after waveleting and MuscIL (if selected). Criteria for rejection includes a choice of joint-probability criteria, amplitude-based criteria, or a combination of the two. If you would like to reject segments, enter Y. Otherwise, enter N.

If performing segment rejection:

Choose a method of segment rejection:

amplitude = Amplitude criteria only

similarity = Segment similarity only

both = Both amplitude criteria and segment similarity

The first option is using amplitude criteria only: amplitude-based criteria sets a minimum and maximum signal amplitude as the artifact threshold, with segments being removed when their amplitude exceeds either threshold. To run this option alone, enter amplitude. The second option is segment similarity criteria, which considers how likely a segment’s activity is to be artifact-laden given the activity of other segments for that channel and other channels’ activity for the same segment and removes outlier segments. The assumption is that artifact segments should be rare relative to the rest of the data at this point in the processing stream. To run this option alone, enter similarity. You can also choose to run both criteria by entering both. Depending on your choice, you may need to follow additional prompts.

If using amplitude criteria or both criteria:

Minimum signal amplitude to use as the artifact threshold:

Enter the minimum signal amplitude used for segment rejection.

**Example:** -150

Maximum signal amplitude to use as the artifact threshold:

Enter the maximum signal amplitude used for segment rejection.

**Example:** 150

Use all channels or a region of interest for segment rejection?

all = All channels

roi = Region of interest

If you plan to analyze all user-specified channels in your dataset, enter all. If you have a specific region of interest that you will be analyzing enter roi. Depending on your choice you may need to answer additional prompts.

If rejecting channels based on a region of interest:

Enter the channels in the ROI, one at a time.

When you have finished entering all channels, enter “done” (without quotations).

Enter the channels that you wish to include in your region of interest one at a time. Press your newline key (enter for Windows, return for Mac) between each entry. If you have any questions about your channel names, refer to your acquisition layout. As with entering your initial channels of interest, enter the names exactly as they appear in the dataset and without any quotations. When you have finished entering channels, enter done.

**Example:** O1

If channel locations are included or provided:

Re-reference data? [Y/N]

If you want to re-reference your data, input Y. Otherwise, enter N.

If re-referencing:

Does your data contain a flatline or all zero reference channel? [Y/N]

Some datasets include a reference channel that is a flat line; in other words, each data point for that channel is zero. If your data includes such a channel, enter Y. Otherwise, enter N.

If your data includes a flatline reference channel:

Enter reference channel ID:

If unknown, press enter/return.

Enter the channel ID of your flatline reference channel exactly as it appears in the data. If you don’t enter the exact name, it can prevent HAPPE from processing correctly.

**Example:** Cz

Re-referencing type:

average = Average re-referencing

subset = Re-reference to a channel or subset of channels

rest = Re-reference using infinity with REST (Yao, 2001)

HAPPE offers three options for re-referencing: average re-referencing, re-referencing to a subset, or using REST. To use average re-referencing, input average. To re-reference to a single channel or to a subset of channels, enter subset. To re-reference to infinity using REST, enter rest. Depending on your choice, you may need to answer additional prompts.

If re-referencing to a subset of channels:

Enter channel/subset of channels to re-reference to, one at a time.

When you have entered all channels, input “done” (without quotations).

Enter the channels that you wish to re-reference to one at a time. As with entering your initial channels of interest, the names should be entered exactly as they appear in the data. If you have questions about your channel names, refer to your acquisition layout. Between each channel entry, press your newline key (enter for Windows, return for Mac). When you finish entering channels, enter done.

**Example:** M1

If re-referencing using REST:

Reference type:

average = Average reference

fixed = Fixed reference

mastoid = Ipsilateral and/or contralateral mastoid

Select whether your original reference is average, fixed (as in a single electrode), or mastoid (ipsilateral and/or contralateral mastoid). If you have questions about which option to use for your dataset, please refer to the REST documentation.

If your reference selection is mastoid:

Enter list of left channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the left side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T7

Enter list of right channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the right side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T8

Load or calculate leadfield?

load = Load existing leadfield

calculate = Calculate new leadfield

If you have an existing leadfield that you want to use, enter load. Otherwise, if you don’t have a leadfield, you can have one calculated as part of the REST process by entering calculate. Depending on your choice, you may need to answer additional prompts.

If calculating the leadfield and reference selection is mastoid:

Enter XYZ coordinates for the left reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the left reference (often this is M1) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-37.1 89.8 -68.3]

Enter XYZ coordinates for the right reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the right reference (often this is M2) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-36.9 -88.8 -69.2]

If loading an existing leadfield:

Enter the name of the leadfield file, including path and file extension:

Enter the full name of the leadfield file including the path and file extension, examples below. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/

leadfield.dat

**Example (Windows):** C:\Users\lgd\Documents\

leadfield.dat

Format to save processed data:

1 = .txt file (electrodes as columns, time as rows) – Choose this for

ERP timeseries

2 = .mat file (MATLAB format)

3 = .set file (EEGLAB format)

Select your preferred format in which to save your processed data. For resting state data, you will likely select .mat (2) or .set (3) format.

Run HAPPE with visualizations? [Y/N]

HAPPE has the option of creating visualizations as intermediate methods of validating the processing stream on your data. By choosing Y (case insensitive), you can run HAPPE in the semi-automated setting with several visualizations for every file. We recommend running this mode on a couple of files to validate your parameter settings. Once you have optimized your parameters for your data, you can run your files in batch without interruption by entering N (case insensitive).

If visualizations are enabled:

Minimum value for power spectrum figure:

Enter the minimum frequency to plot for the power spectrum figure.

**Example:** 1

Maximum value for power spectrum figure:

Enter the maximum frequency to plot for the power spectrum figure.

**Example:** 45

Enter the frequencies, one at time, to generate spatial topoplots for:

When you have entered all frequencies, input “done” (without quotations).

Enter any frequencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the power spectrum for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the frequencies you wish to plot, enter done.

**Example:** 10

Please check your parameters before continuing.

At this point, your parameters are listed out in the command window.

Are the above parameters correct? [Y/N]

Go through the parameters listed in the command window and check to see that your parameters are correct. If so, enter Y (case insensitive). Otherwise, if your parameters are not correct, enter N (case insensitive) to change them.

If changing or correcting a parameter:

Parameter to change:

HAPPE will list out the parameters you can change.

Choose a parameter to change by entering the name of the parameter exactly as it appears in the above prompt (excluding newline breaks). You will be prompted with the original command to change the parameter. Depending on your data and parameter set, you may be presented with different options. You can change as many parameters as you want but must change each parameter one at a time. Some selections may require you to answer more than one prompt. When you finish entering parameters, enter done.

**Example:** segment rejection

If you changed an existing parameter set or created a new set:

Parameter file save name:

default = Default name (inputParameters\_dd-mm-yyyy.mat)

custom = Create a custom file name

To create your own save name for the parameter set, input custom. You will then be prompted to enter your custom save name. If you do not want to create a custom label, you can use the default name, inputParameters\_dd-mm-yyyy.mat (using the current date), by entering default.

If using a custom name:

File name (Do not include .mat):

Enter a name for your saved parameter set. Name the file the same way you would if you were using your file explorer; no need to include the .mat extension as it will be automatically added.

**Example:** baselineProcessing\_studyname

If a file with the parameter set name exists:

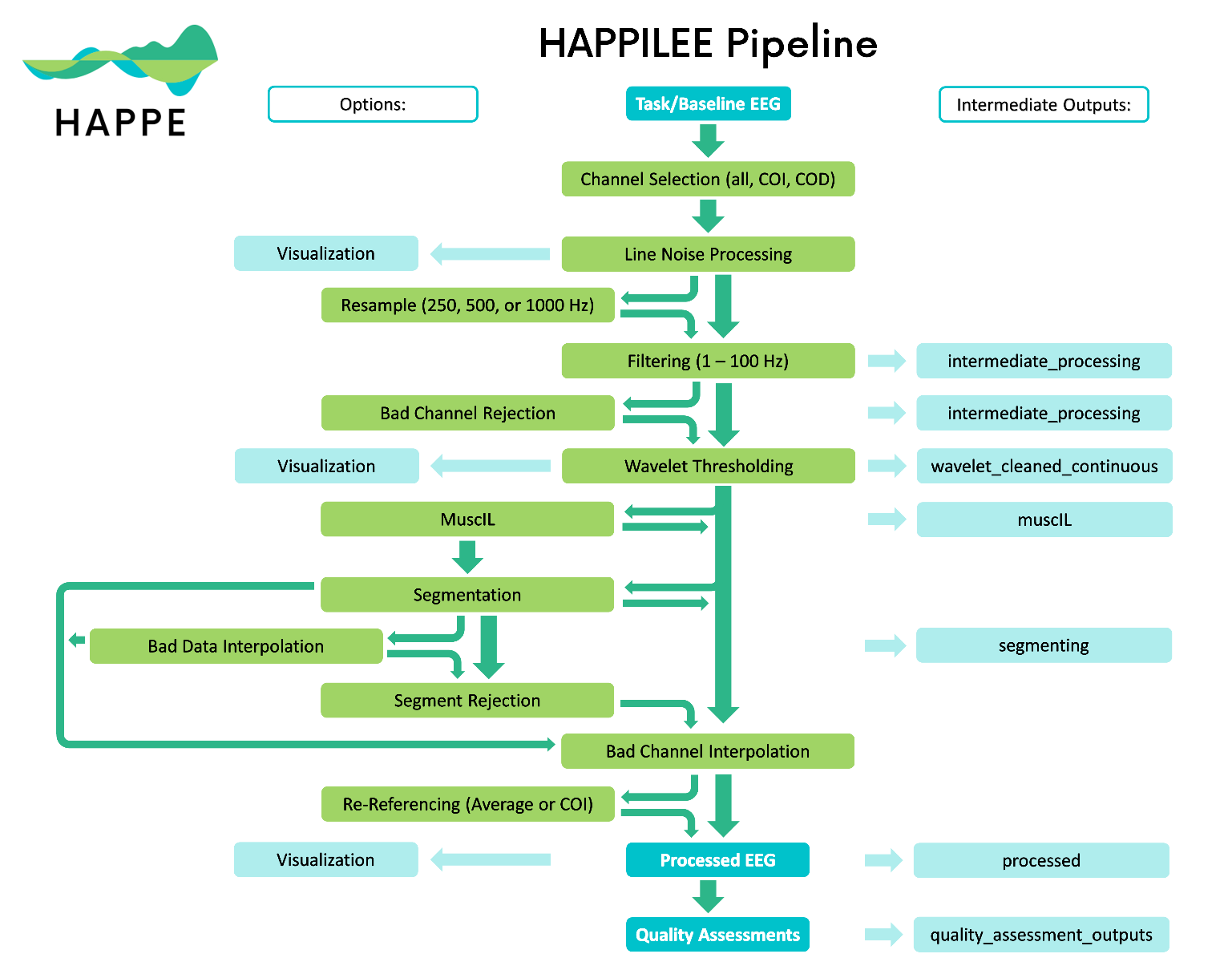
A set of input parameters with this name already exists.

overwrite = Overwrite the existing file

new = Save a new file with a different name

Entering multiple parameters sets on the same day or using the same name for a parameter set can result in multiple files vying for the same file name. The MATLAB default is to overwrite existing files, and if you want to replace the existing file with the new file, you can do so by entering overwrite. However, if you do not want to overwrite the file, enter new.

## HAPPILEE Task-Related



### Following Command Line Prompts

Prompts as they appear in the command window are written in a monospace font, like this, followed by a brief description of what to enter, with an example. Sometimes a certain choice will result in a different set of prompts, in which case they will be indented under a heading that describes the choice resulting in the branch in directions.

Enter the path to the folder containing the dataset(s):

Input the folder, including the full path, where the raw data is located. If running HAPPE for the first time on this data, the folder should ONLY include the raw files you wish to process; no other folders or files should live inside that folder. If reprocessing data, all outputs (folders and documents) should be in the folder in the same file structure as they were created during the original HAPPE run in addition to the raw data.

**Example (Mac):** /Users/laurelg-d/Desktop/Data Folder

**Example (PC):** C:\Users\laurelg-d\Documents\Data Folder

Select an option:

raw = Run on raw data from the start

reprocess = Run on HAPPE-processed data starting post-artifact reduction

If this is your first time processing the raw data file(s), input raw. If you wish to re-run raw data that has previously been processed, starting at the MuscIL or segmentation step, input reprocess. Depending on your choice, you may be asked to follow additional prompts (see below).

If reprocessing existing data:

Files, such as processed data and quality metrics may already exist for this dataset.

overwrite = Overwrite existing files

new = Save new files

To keep both the already existing files and the newly created files, input new. If you do not want to keep previous files for this dataset, input overwrite.

If saving new files:

Use a default or custom label for processed data?

default = Default name (\_rerun\_dd-mm-yyyy)

custom = Create your own file label, starting with an

underscore

To create your own save name for the processed data, input custom. You will then be prompted to enter your custom label. This label should start with an underscore (“\_”) to encourage best naming practices. If you do not want to create a custom label, you can use the default (“\_rerun\_dd-mm-yyyy”), using the current date) by entering default. In either case, this will save the files as the original raw file name with the label added. Caution: if you already have data with this label, HAPPE will overwrite the existing files.

Load pre-existing set of input parameters? [Y/N]

If you have already set parameters through HAPPE for either this or another dataset that support the current dataset, you can choose to load these parameters instead of entering them from scratch. To use a previously saved set, enter Y (case insensitive); otherwise, enter N (case insensitive).

If loading pre-existing parameters:

Enter your parameter file, including the full path and file extension:

Enter the full name of the parameter file including the path and file extension, examples below. By default, parameters are saved in a folder named **input\_parameters**. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/Data Folder/input\_parameters/

inputParameters\_dd-mm-yyyy.mat

**Example (PC):** C:\Users\lgd\Documents\Data Folder\input\_parameters\

inputParameters\_dd-mm-yyyy.mat

If running a different version of HAPPE than the parameters were saved with:

These parameters were saved using a different version of HAPPE and ay be incompatible with the running pipeline.

Proceed anyway? [Y/N]

If your parameters were saved using a different version of HAPPE, there is a chance that they do not have the full set of parameters needed to properly run through the current version. If you are certain that you have all the correct parameters or want to try to run anyway, enter Y (case insensitive). However, be warned that there is a high probability that HAPPE will error out. If you do not want to run potentially incompatible parameters, enter N (case insensitive) and HAPPE will end the current run via error.

Change an existing parameter? [Y/N]

If any of the loaded parameters do not match what you want for the current run, you can change them at this step by entering Y (case insensitive). Otherwise, enter N (case insensitive) to continue the run.

If creating a new parameter set:

Low density data? [Y/N]

For HAPPE, low density data contains 1 to ~32 channels.

For low density data, always enter Y (case insensitive).

Data type:

rest = Resting-state/baseline EEG

task = Task-related EEG

Input task (case insensitive) for HAPPILEE with task-related, non-ERP EEG data.

Performing event-related potential (ERP) analysis? [Y/N]

Input N (case insensitive) for HAPPILEE with task-related, non-ERP EEG data.

Enter the task onset tags, one at a time, pressing enter/return between each entry. When you have entered all tags, input “done” (without quotations).

Your data should include tags that indicate the onset of each trial in the dataset. Any stimuli/event/onset tags that EEGLAB can detect can be used in HAPPE. To verify that your tags are readable, load a data file into EEGLAB using the GUI and verify that your events are present. For detailed guidance on loading data into EEGLAB’s GUI and checking/importing events, we recommend the EEGLAB tutorial “Importing events and channel locations” found on YouTube and a written tutorial found here: <https://eeglab.org/tutorials/04_Import/Importing_Event_Epoch_Info.html>.

If you entered more than one onset tag:

Do multiple onset tags belong to a single condition? [Y/N]

Example: “happy\_face” and “sad\_face” belong to “faces”.

Sometimes, you may have multiple tags for different types of trials within a single condition. For example, you may have two conditions: faces and objects. Within the faces condition you have trials marked differently for happy faces, sad faces, and neutral faces. This function allows you combine the different kinds of trials for the faces condition into a single EEG to be output alongside the individual trial EEGs and the full EEG after processing. You do not need to enter all the tags within a condition of you do not want to, for example, only including happy faces and sad faces in your faces condition. If you have conditions and wish to have trials considered this way, enter Y. Otherwise, if you do not have conditions or do not wish to examine your data on the level of conditions, enter N.

If multiple tags belong to a single condition:

Enter the conditions and included onset tags:

Enter each condition as a list, each element separated by a blank space, with the condition name as the first item in the list. Press enter/return between entries. When you have entered all conditions, input “done” (without quotations).

Example: faces happy\_face sad\_face angry\_face

For each condition you must enter the condition name and all the onset tags associated with that condition. This should be entered as a list, starting with the condition name, which you may choose at this time. The onset tags should be entered as they appear in your data. Separate each item in the list using only a blank space, no commas. If you include a tag name not included in your original set of onset tags (see above), it may not be included or processed correctly, so ensure that you enter all onset tags when completing that step. You do not need to enter all the tags within a condition of you do not want to. For example, you may only include happy faces and sad faces in your faces condition when the condition may also have neutral faces. When you have completed this list, press your newline key to submit the condition and to enter a new one. When you finish entering conditions, enter done.

**Example:** condition tag1 tag2

File format:

1 = .mat (MATLAB file)

2 = .raw (Net Station simple binary)

3 = .set (EEGLAB format)

4 = .cdt (Neuroscan format)

5 = .mff (EGI format)

This is the file format of your raw data. Depending on your choice, you may need to follow a different set of prompts. HAPPE does not currently support low-density data in formats other than .mat and .set.

**Example:** 3

Acquisition layout net type:

1 – EGI Geodesic Sensor Net

2 – EGI HydroCel Geodesic Sensor Net

3 – Neuroscan Quik-Cap

4 – Other

Select the type of acquisition layout used to collect the data. Different choices may result in different prompts.

**Example:** 4

Number of channels:

The number of channels included in the layout. Make sure to use the number from the layout, even if your data does not use or include all electrodes.

**Example:** 16

If your data is in .mat format:

What best describes your data?

net station = Data exported using Net Station with additional

information

matrix = MATLAB matrix of the data

For data exported in .mat format from Net Station, enter net station. When loaded into MATLAB, it should include information beyond the data matrix, such as the sampling rate. Otherwise, if the data is a MATLAB matrix of data without any additional information, enter matrix. Depending on your choice, you may need to follow a different set of prompts.

If your data is exported from Net Station:

Enter the potential EEG variable names, one at a time.

Press enter/return between each entry.

When you are finished, enter “done” (without quotations).

NOTE: variable names containing “segment” may cause issues.

For Net Station .mat files, enter the potential names of the variable that consists of the EEG data matrix. Between each entry, press your newline key (enter – Windows, return – Mac). When you are done entering potential variable names, enter done. If you don’t have any ideas of what your variable name could be, try loading the Net Station .mat file into an empty MATLAB workspace and checking the variables.

**Example:** Category\_1

If your data is a MATLAB matrix:

If your channel locations file is not included in HAPPE’s acquisition layouts:

Do you have a channel locations file for your data? [Y/N]

NOTE: A list of supported file types can be found in the HAPPE User Guide.

Unless HAPPE includes the channel locations for your acquisition layout by default, you will likely need to provide them for MATLAB matrixes in .mat format. Accepted channel locations file formats are those accepted by EEGLAB (Delorme & Makieg, 2004) and include:

1. .loc, .locs, .eloc – EEG polar coordinates
2. .ced – EEGLAB with polar, cartesian, and spherical
3. .sph – MATLAB spherical coordinates
4. .elc – Cartesian 3-D from EETrack
5. .elp – Polhemus Cartesian coordinates
6. .elp – BESA spherical coordinates
7. .xyz – MATLAB/EEGLAB Cartesian coordinates
8. .asc, .dat – Neuroscan Cartesian polar coordinates
9. .mat – Brainstorm channel locations
10. .sfp – BESA/EGI xyz Cartesian coordinates

If you do have a channel locations file,

Enter the name of the channel locations file, including the full path and file extension:

Enter the full path to where the channel locations file exists, followed by the name of the file, including its extension. The channel locations file does not need to be in a specific location on your computer or on an external drive, only that it is accessible.

**Example (Mac):** /Users/lgd/Desktop/

chanlocs.loc

**Example (Windows):** C:\Users\lgd\Documents\

chanlocs.loc

If you do not have a channel locations file,

HAPPE functionality is limited without channel locations. Continue anyway? [Y/N]

If you don’t have a channel locations file, you will still be able to run your data through HAPPE, granted with a limited set of steps. Without channel locations, you cannot filter to channels of interest, perform bad channel detection, interpolate bad channels, or re-reference your data. If this is acceptable, you can continue running HAPPE by entering Y (case insensitive). Otherwise, you can enter N (case insensitive) and HAPPE will stop running by throwing an error.

Do all your files share the same sampling rate? [Y/N]

If all your files share the same sampling rate, input Y (case insensitive). Otherwise, input N (case insensitive). Depending on your input, you may have to follow different prompts.

If your files do share a sampling rate:

Sampling rate:

Enter the sampling rate of all the files as an integer.

**Example:** 250

If your files do not share a sampling rate:

Enter the name of the file containing the sampling rates for each data file, including the path and file extension:

See the HAPPE User Guide for how this table should be formatted.

Enter the name of the file, including the full path and the file’s extension. The file does not have to be in any particular location on your computer or on an external drive; it only must be accessible. Additionally, the file could be in any format so long as it is readable as a table by MATLAB, but we recommend .csv or an Excel file.

An example for formatting this file is shown below. The first column should contain the file name while the second column should contain the sampling rate in Hz.

|  |  |
| --- | --- |
| samplefilename1.mat | 250 |
| samplefilename2.mat | 500 |

**Example (Mac):** /Users/lgd/Desktop/sampRate.csv

**Example (Windows):** C:\Users\lgd\Documents\

sampRate.csv

Path to .txt files containing task event info:

For .mat files, you will need to provide .txt files containing your event information. Enter the full path to the folder holding these files, including the name of the folder. This folder does not need to be on the same path as your data.

**Example (Mac):** /Users/lgd/Desktop/Task InfoFolder

**Example (PC):** C:\Users\lgd\Documents\Task Info Folder

If your file has channel locations included or provided:

Select channels of interest:

all = Select all channels in the data

coi = Select a user-specified subset of channels

To process all possible channels included in each EEG file, enter all. If you wish to only process a subset of channels, enter coi.

If only processing a subset of channels:

Choose an option for entering channels:

include = Include ONLY the entered channel names

exclude = Include every channel EXCEPT the entered

channel names

To enter your channel subset, you have the option to select the channels of interest or the channels of disinterest. If you select include all the channel names that you enter in the next prompt will be the only channels included in processing. Otherwise, by selecting exclude, you will instead process only the channels *not* entered in the next prompt. It may be easier to choose to include or exclude based on which list of channels is shorter.

Enter channels, including the preceding letter, one at a time.

Press enter/return between each entry.

Examples:

E17

M1

When you have entered all channels, input “done” (without quotations).

Enter the channels you wish to include or exclude, depending on your previous choice, one at a time. You should enter the channel name exactly as it appears in the data, case sensitive. This means you should include the preceding letter, if applicable (example: E17). If you have questions about your channel names, refer to your acquisition layout. Don’t use quotations around the channel name. Between each entry, press your newline key (enter for Windows, return for Mac). When you are finished entering all channels in your list of inclusion/exclusion, enter done.

Frequency of electrical (line) noise in Hz:

USA data probably = 60; Otherwise, probably = 50

Enter the frequency of electrical noise at the site of data collection. For locations within the United States, the frequency is likely 60 Hz, and outside is likely 50 Hz. We’ve included a link to a list of frequencies by country, but you should independently confirm that you have the correct frequency: <https://www.oaktreeproducts.com/img/product/description/List%20of%20Worldwide%20AC%20Voltages.pdf>

Are there any additional frequencies (e.g., harmonics) to reduce? [Y/N]

In some cases, there may be prominent line noise at a frequency other than the electrical noise frequency. Often, but not always, these are the harmonics of the electrical frequency. If there is enough noise at other frequencies that it affects the data, you can choose to reduce line noise at those frequencies by entering Y, case insensitive. Otherwise, enter N, case insensitive.

If reducing line noise in additional frequencies:

Enter the frequencies, one at a time, to pass through noise reduction.

When finished entering frequencies, enter “done” (without quotations).

Enter the frequencies you want to reduce, aside from the electrical noise frequency, that you want to reduce, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering frequencies, enter done.

**Example:** 25

Line noise reduction method:

cleanline = Use Tim Mullen’s CleanLine

notch = Use a notch filter (COMING SOON)

For reducing line noise, you have two options: using Tim Mullen’s CleanLine EEGLAB or using a notch filter. At this time, only the CleanLine option is available, so entering notch will end the HAPPE run via an error.

If using CleanLine:

Use legacy or default line noise reduction?

default – Default method optimized in HAPPE v2

legacy – Method from HAPPE v1 (NOT RECOMMENDED)

The legacy line noise reduction method uses the CleanLine plugin for EEGLAB from a far less effective version. For this reason, we strongly recommend using the default to process your data as it uses an improved version of CleanLine to reduce line noise. If you want to use the legacy method for comparison analyses or to finish existing data processing, enter legacy. For all other situations, you should enter default.

Resample data? [Y/N]

If you want to resample your data, enter Y. This might occur if you have different sampling rates across files and want them to be the same. Otherwise, enter N to not resample your data. Both options are case insensitive.

If resampling your data:

HAPPE supports resampling to 250, 500, and 1000 Hz.

Resample frequency:

Currently, HAPPE restricts resampling to the following values: 250, 500, or 1000 Hz. Choose one of these frequencies by entering the value.

**Example:** 500

If your file has channel locations included or provided:

Perform bad channel detection? [Y/N]

If you want remove channels with high impedances, damage to the electrodes, insufficient scalp contact, and/or excessive movement or electromyographic (EMG) artifact throughout the recording from subsequent analyses, enter Y, case insensitive. To include all your channels of interest regardless of quality, enter N, case insensitive.

If rejecting bad channels:

Bad channel detection method:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED).

You can choose to run either the default or the legacy version of bad channel detection. The legacy version was used in the original HAPPE software, is outdates, and we no longer support it, thus it is not recommended. Despite this, to select this option, enter legacy. The default method uses a new method of bad channel detection using the Clean Rawdata function with optimized criteria. To use this optimized method, enter default.

Method of wavelet thresholding:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED)

The new wavelet method uses a soft empirical Bayesian level-dependent threshold for the wavelets. Wavelet family is coiflet (level 4). This method has been optimized on EEG data (see Lopez et al., in press; Monachino et al., in press). To use this method, enter default. The legacy method of wavelet thresholding uses a soft, global threshold for the wavelets. The wavelet family is coiflet (level 5) and the threshold multiplier is used to remove more high frequency noise. This method has not been optimized, so it is not recommended, but if you wish to use it, enter legacy.

Use ICLabel to reduce remaining muscle artifact in your data? [Y/N]

NOTE: This will drastically increase processing time. Recommended for files with significant muscle artifact.

Sometimes there may be a high level of electromyographic (EMG) artifact in the data that remains despite wavelet thresholding. In this case, you can use ICA to remove the remaining artifact. HAPPE uses ICLabel to classify independent components and to reject all artifacts labeled as muscle with at least 25% probability. Using this is only recommended when you know there is EMG artifact in the data that wavelet thresholding doesn’t handle as ICA takes a lot of time and this step can exponentially increase processing time. If you want to use this option, called MuscIL (for muscle ICLabel), enter Y (case insensitive). Otherwise, enter N (case insensitive).

Segment data? [Y/N]

To segment your data, enter Y. Otherwise, enter N. If you choose to segment your data, you may be asked to answer additional prompts to properly segment your data.

If segmenting your data:

Segment start, in MILLISECONDS, relative to stimulus onset:

Example: -100

Enter the number of milliseconds present in the trial prior to your stimulus onset. If you have no baseline period within your trial, this may be 0. Otherwise, this is a negative number. For example, if you have a trial with a 100-millisecond baseline, you will enter -100.

Segment end, in MILLISECONDS, relative to stimulus onset:

Enter the number of milliseconds present in the trial after stimulus onset. This should always be a number greater than 0. For example, if you have a trial that runs for 500 milliseconds after the stimulus onset or trial start, you will enter 500.

**Example:** 500

If channel locations are included or provided and the data is segmented:

Interpolate the specific channels data determined to be artifact/bad within each segment? [Y/N]

This option enables evaluation of channels within epochs to determine whether there is bad data present. Channels flagged with bad data for that segment will then have their data interpolated only for that segment, using FASTER (Nolan et al., 2010). If you wish to interpolate bad channels within segments, enter Y. Otherwise, enter N.

Perform segment rejection? [Y/N]

Users can choose to reject segments that are determined to still have artifact remaining after waveleting and MuscIL (if selected). Criteria for rejection includes a choice of joint-probability criteria, amplitude-based criteria, or a combination of the two. If you would like to reject segments, enter Y. Otherwise, enter N.

If performing segment rejection:

Choose a method of segment rejection:

amplitude = Amplitude criteria only

similarity = Segment similarity only

both = Both amplitude criteria and segment similarity

The first option is using amplitude criteria only: amplitude-based criteria sets a minimum and maximum signal amplitude as the artifact threshold, with segments being removed when their amplitude exceeds either threshold. To run this option alone, enter amplitude. The second option is segment similarity criteria, which considers how likely a segment’s activity is to be artifact-laden given the activity of other segments for that channel and other channels’ activity for the same segment and removes outlier segments. The assumption is that artifact segments should be rare relative to the rest of the data at this point in the processing stream. To run this option alone, enter similarity. You can also choose to run both criteria by entering both. Depending on your choice, you may need to follow additional prompts.

If using amplitude criteria or both criteria:

Minimum signal amplitude to use as the artifact threshold:

Enter the minimum signal amplitude used for segment rejection.

**Example:** -150

Maximum signal amplitude to use as the artifact threshold:

Enter the maximum signal amplitude used for segment rejection.

**Example:** 150

Use all channels or a region of interest for segment rejection?

all = All channels

roi = Region of interest

If you plan to analyze all user-specified channels in your dataset, enter all. If you have a specific region of interest that you will be analyzing enter roi. Depending on your choice you may need to answer additional prompts.

If rejecting channels based on a region of interest:

Enter the channels in the ROI, one at a time.

When you have finished entering all channels, enter “done” (without quotations).

Enter the channels that you wish to include in your region of interest one at a time. Press your newline key (enter for Windows, return for Mac) between each entry. If you have any questions about your channel names, refer to your acquisition layout. As with entering your initial channels of interest, enter the names exactly as they appear in the dataset and without any quotations. When you have finished entering channels, enter done.

**Example:** O1

Use pre-selected “usable” trials to restrict analysis? [Y/N]

If you have previously determined which trials in your dataset are usable, you can choose to filter out the “unusable” trials using this option. THIS OPTION DOES NOT CURRENTLY FUNCTION!

If channel locations are included or provided:

Re-reference data? [Y/N]

If you want to re-reference your data, input Y. Otherwise, enter N.

If re-referencing:

Does your data contain a flatline or all zero reference channel? [Y/N]

Some datasets include a reference channel that is a flat line; in other words, each data point for that channel is zero. If your data includes such a channel, enter Y. Otherwise, enter N.

If your data includes a flatline reference channel:

Enter reference channel ID:

If unknown, press enter/return.

Enter the channel ID of your flatline reference channel exactly as it appears in the data. If you don’t enter the exact name, it can prevent HAPPE from processing correctly.

**Example:** Cz

Re-referencing type:

average = Average re-referencing

subset = Re-reference to a channel or subset of channels

rest = Re-reference using infinity with REST (Yao, 2001)

HAPPE offers three options for re-referencing: average re-referencing, re-referencing to a subset, or using REST. To use average re-referencing, input average. To re-reference to a single channel or to a subset of channels, enter subset. To re-reference to infinity using REST, enter rest. Depending on your choice, you may need to answer additional prompts.

If re-referencing to a subset of channels:

Enter channel/subset of channels to re-reference to, one at a time.

When you have entered all channels, input “done” (without quotations).

Enter the channels that you wish to re-reference to one at a time. As with entering your initial channels of interest, the names should be entered exactly as they appear in the data. If you have questions about your channel names, refer to your acquisition layout. Between each channel entry, press your newline key (enter for Windows, return for Mac). When you finish entering channels, enter done.

**Example:** M1

If re-referencing using REST:

Reference type:

average = Average reference

fixed = Fixed reference

mastoid = Ipsilateral and/or contralateral mastoid

Select whether your original reference is average, fixed (as in a single electrode), or mastoid (ipsilateral and/or contralateral mastoid). If you have questions about which option to use for your dataset, please refer to the REST documentation.

If your reference selection is mastoid:

Enter list of left channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the left side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T7

Enter list of right channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the right side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T8

Load or calculate leadfield?

load = Load existing leadfield

calculate = Calculate new leadfield

If you have an existing leadfield that you want to use, enter load. Otherwise, if you don’t have a leadfield, you can have one calculated as part of the REST process by entering calculate. Depending on your choice, you may need to answer additional prompts.

If calculating the leadfield and reference selection is mastoid:

Enter XYZ coordinates for the left reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the left reference (often this is M1) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-37.1 89.8 -68.3]

Enter XYZ coordinates for the right reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the right reference (often this is M2) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-36.9 -88.8 -69.2]

If loading an existing leadfield:

Enter the name of the leadfield file, including path and file extension:

Enter the full name of the leadfield file including the path and file extension, examples below. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/

leadfield.dat

**Example (Windows):** C:\Users\lgd\Documents\

leadfield.dat

Format to save processed data:

1 = .txt file (electrodes as columns, time as rows) – Choose this for

ERP timeseries

2 = .mat file (MATLAB format)

3 = .set file (EEGLAB format)

Select your preferred format in which to save your processed data. For continuous task data, you will likely select .mat (2) or .set (3) format, but you may also select .txt (1).

Run HAPPE with visualizations? [Y/N]

HAPPE has the option of creating visualizations as intermediate methods of validating the processing stream on your data. By choosing Y (case insensitive), you can run HAPPE in the semi-automated setting with several visualizations for every file. We recommend running this mode on a couple of files to validate your parameter settings. Once you have optimized your parameters for your data, you can run your files in batch without interruption by entering N (case insensitive).

If visualizations are enabled:

Minimum value for power spectrum figure:

Enter the minimum frequency to plot for the power spectrum figure.

**Example:** 1

Maximum value for power spectrum figure:

Enter the maximum frequency to plot for the power spectrum figure.

**Example:** 45

Enter the frequencies, one at time, to generate spatial topoplots for:

When you have entered all frequencies, input “done” (without quotations).

Enter any frequencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the power spectrum for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the frequencies you wish to plot, enter done.

**Example:** 10

Please check your parameters before continuing.

At this point, your parameters are listed out in the command window.

Are the above parameters correct? [Y/N]

Go through the parameters listed in the command window and check to see that your parameters are correct. If so, enter Y (case insensitive). Otherwise, if your parameters are not correct, enter N (case insensitive) to change them.

If changing or correcting a parameter:

Parameter to change:

HAPPE will list out the parameters you can change.

Choose a parameter to change by entering the name of the parameter exactly as it appears in the above prompt (excluding newline breaks). You will be prompted with the original command to change the parameter. Depending on your data and parameter set, you may be presented with different options. You can change as many parameters as you want but must change each parameter one at a time. Some selections may require you to answer more than one prompt. When you finish entering parameters, enter done.

**Example:** segment rejection

If you changed an existing parameter set or created a new set:

Parameter file save name:

default = Default name (inputParameters\_dd-mm-yyyy.mat)

custom = Create a custom file name

To create your own save name for the parameter set, input custom. You will then be prompted to enter your custom save name. If you do not want to create a custom label, you can use the default name, inputParameters\_dd-mm-yyyy.mat (using the current date), by entering default.

If using a custom name:

File name (Do not include .mat):

Enter a name for your saved parameter set. Name the file the same way you would if you were using your file explorer; no need to include the .mat extension as it will be automatically added.

**Example:** taskProcessing\_studyname

If a file with the parameter set name exists:

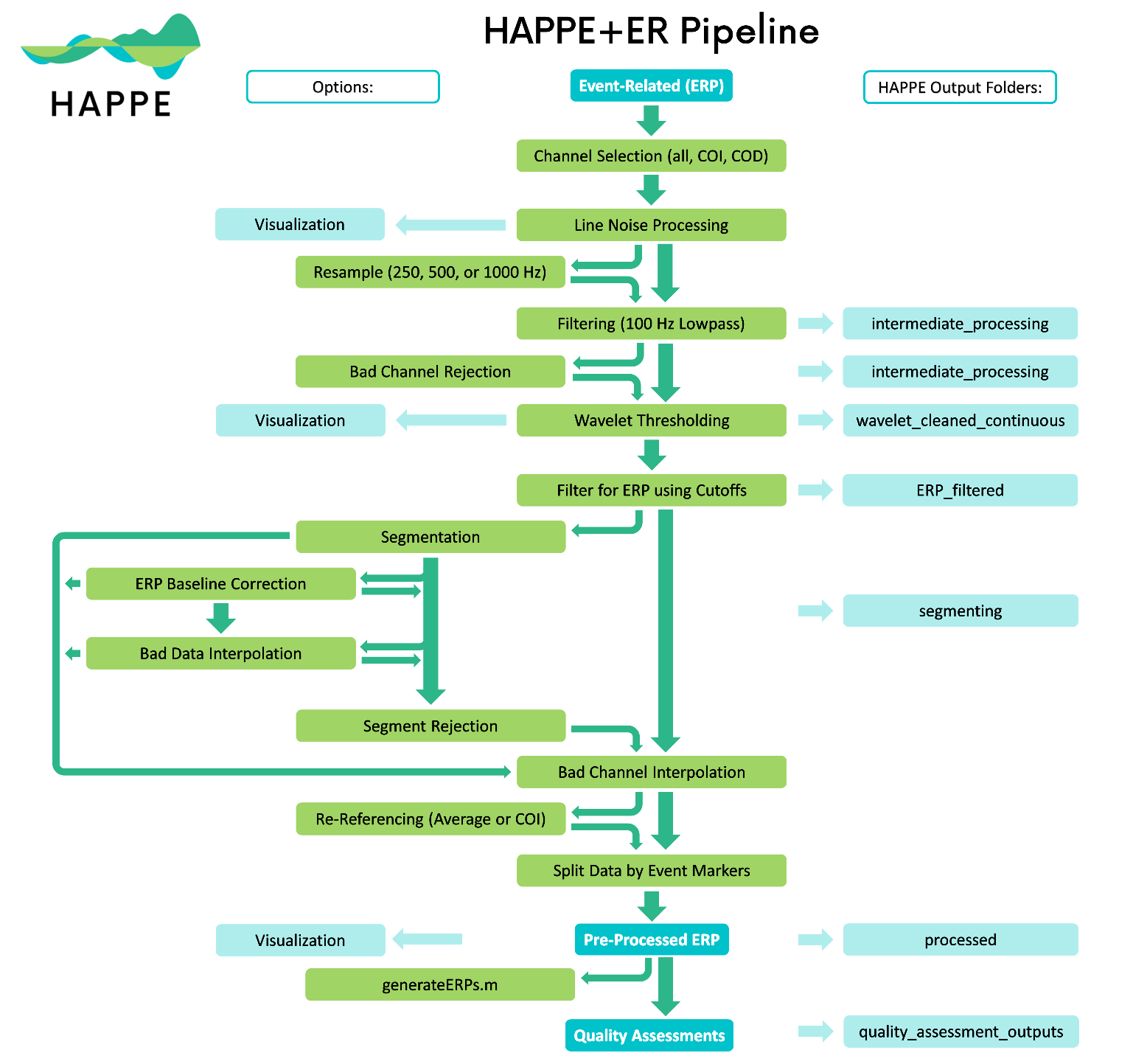
A set of input parameters with this name already exists.

overwrite = Overwrite the existing file

new = Save a new file with a different name

Entering multiple parameters sets on the same day or using the same name for a parameter set can result in multiple files vying for the same file name. The MATLAB default is to overwrite existing files, and if you want to replace the existing file with the new file, you can do so by entering overwrite. However, if you do not want to overwrite the file, enter new.

## HAPPILEE & HAPPE+ER



### Following Command Line Prompts

Prompts as they appear in the command window are written in a monospace font, like this, followed by a brief description of what to enter, with an example. Sometimes a certain choice will result in a different set of prompts, in which case they will be indented under a heading that describes the choice resulting in the branch in directions.

Enter the path to the folder containing the dataset(s):

Input the folder, including the full path, where the raw data is located. If running HAPPE for the first time on this data, the folder should ONLY include the raw files you wish to process; no other folders or files should live inside that folder. If reprocessing data, all outputs (folders and documents) should be in the folder in the same file structure as they were created during the original HAPPE run in addition to the raw data.

**Example (Mac):** /Users/laurelg-d/Desktop/Data Folder

**Example (PC):** C:\Users\laurelg-d\Documents\Data Folder

Select an option:

raw = Run on raw data from the start

reprocess = Run on HAPPE-processed data starting post-artifact reduction

If this is your first time processing the raw data file(s), input raw. If you wish to re-run raw data that has previously been processed, starting at the MuscIL or segmentation step, input reprocess. Depending on your choice, you may be asked to follow additional prompts (see below).

If reprocessing existing data:

Files, such as processed data and quality metrics may already exist for this dataset.

overwrite = Overwrite existing files

new = Save new files

To keep both the already existing files and the newly created files, input new. If you do not want to keep previous files for this dataset, input overwrite.

If saving new files:

Use a default or custom label for processed data?

default = Default name (\_rerun\_dd-mm-yyyy)

custom = Create your own file label, starting with an

underscore

To create your own save name for the processed data, input custom. You will then be prompted to enter your custom label. This label should start with an underscore (“\_”) to encourage best naming practices. If you do not want to create a custom label, you can use the default (“\_rerun\_dd-mm-yyyy”), using the current date) by entering default. In either case, this will save the files as the original raw file name with the label added. Caution: if you already have data with this label, HAPPE will overwrite the existing files.

Load pre-existing set of input parameters? [Y/N]

If you have already set parameters through HAPPE for either this or another dataset that support the current dataset, you can choose to load these parameters instead of entering them from scratch. To use a previously saved set, enter Y (case insensitive); otherwise, enter N (case insensitive).

If loading pre-existing parameters:

Enter your parameter file, including the full path and file extension:

Enter the full name of the parameter file including the path and file extension, examples below. By default, parameters are saved in a folder named **input\_parameters**. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/Data Folder/input\_parameters/

inputParameters\_dd-mm-yyyy.mat

**Example (PC):** C:\Users\lgd\Documents\Data Folder\input\_parameters\

inputParameters\_dd-mm-yyyy.mat

If running a different version of HAPPE than the parameters were saved with:

These parameters were saved using a different version of HAPPE and ay be incompatible with the running pipeline.

Proceed anyway? [Y/N]

If your parameters were saved using a different version of HAPPE, there is a chance that they do not have the full set of parameters needed to properly run through the current version. If you are certain that you have all the correct parameters or want to try to run anyway, enter Y (case insensitive). However, be warned that there is a high probability that HAPPE will error out. If you do not want to run potentially incompatible parameters, enter N (case insensitive) and HAPPE will end the current run via error.

Change an existing parameter? [Y/N]

If any of the loaded parameters do not match what you want for the current run, you can change them at this step by entering Y (case insensitive). Otherwise, enter N (case insensitive) to continue the run.

If creating a new parameter set:

Low density data? [Y/N]

For HAPPE, low density data contains 1 to ~32 channels.

For low density data, always enter Y (case insensitive).

Data type:

rest = Resting-state/baseline EEG

task = Task-related EEG

Input task (case insensitive) for HAPPILLE with task-related data.

Performing event-related potential (ERP) analysis? [Y/N]

Input Y (case insensitive) to run the HAPPE+ER pipeline on ERP EEG data.

Enter the task onset tags, one at a time, pressing enter/return between each entry. When you have entered all tags, input “done” (without quotations).

Your data should include tags that indicate the onset of each trial in the dataset. Any stimuli/event/onset tags that EEGLAB can detect can be used in HAPPE. To verify that your tags are readable, load a data file into EEGLAB using the GUI and verify that your events are present. For detailed guidance on loading data into EEGLAB’s GUI and checking/importing events, we recommend the EEGLAB tutorial “Importing events and channel locations” found on YouTube and a written tutorial found here: <https://eeglab.org/tutorials/04_Import/Importing_Event_Epoch_Info.html>.

If you entered more than one onset tag:

Do multiple onset tags belong to a single condition? [Y/N]

Example: “happy\_face” and “sad\_face” belong to “faces”.

Sometimes, you may have multiple tags for different types of trials within a single condition. For example, you may have two conditions: faces and objects. Within the faces condition you have trials marked differently for happy faces, sad faces, and neutral faces. This function allows you combine the different kinds of trials for the faces condition into a single EEG to be output alongside the individual trial EEGs and the full EEG after processing. You do not need to enter all the tags within a condition of you do not want to, for example, only including happy faces and sad faces in your faces condition. If you have conditions and wish to have trials considered this way, enter Y. Otherwise, if you do not have conditions or do not wish to examine your data on the level of conditions, enter N.

If multiple tags belong to a single condition:

Enter the conditions and included onset tags:

Enter each condition as a list, each element separated by a blank space, with the condition name as the first item in the list. Press enter/return between entries. When you have entered all conditions, input “done” (without quotations).

Example: faces happy\_face sad\_face angry\_face

For each condition you must enter the condition name and all the onset tags associated with that condition. This should be entered as a list, starting with the condition name, which you may choose at this time. The onset tags should be entered as they appear in your data. Separate each item in the list using only a blank space, no commas. If you include a tag name not included in your original set of onset tags (see above), it may not be included or processed correctly, so ensure that you enter all onset tags when completing that step. You do not need to enter all the tags within a condition of you do not want to. For example, you may only include happy faces and sad faces in your faces condition when the condition may also have neutral faces. When you have completed this list, press your newline key to submit the condition and to enter a new one. When you finish entering conditions, enter done.

**Example:** condition tag1 tag2

Enter low-pass filter, in Hz:

Common low-pass filter is 30 – 45 Hz.

Enter the low-pass filter at which to filter your data for ERP analyses.

**Example:** 35

Enter high-pass filter, in Hz:

Common high-pass filter is 0.1 – 0.3 Hz.

Enter the high-pass filter at which to filter your data for ERP analyses.

**Example:** 0.1

File format:

1 = .mat (MATLAB file)

2 = .raw (Net Station simple binary)

3 = .set (EEGLAB format)

4 = .cdt (Neuroscan format)

5 = .mff (EGI format)

This is the file format of your raw data. Depending on your choice, you may need to follow a different set of prompts. HAPPE does not currently support formats other than .mat and .set for low-density data.

**Example:** 3

Acquisition layout net type:

1 – EGI Geodesic Sensor Net

2 – EGI HydroCel Geodesic Sensor Net

3 – Neuroscan Quik-Cap

4 – Other

Select the type of acquisition layout used to collect the data. Different choices may result in different prompts.

**Example:** 4

Number of channels:

The number of channels included in the layout. Make sure to use the number from the layout, even if your data does not use or include all electrodes.

**Example:** 16

If your data is in .mat format:

What best describes your data?

net station = Data exported using Net Station with additional

information

matrix = MATLAB matrix of the data

For data exported in .mat format from Net Station, enter net station. When loaded into MATLAB, it should include information beyond the data matrix, such as the sampling rate. Otherwise, if the data is a MATLAB matrix of data without any additional information, enter matrix. Depending on your choice, you may need to follow a different set of prompts.

If your data is exported from Net Station:

Enter the potential EEG variable names, one at a time.

Press enter/return between each entry.

When you are finished, enter “done” (without quotations).

NOTE: variable names containing “segment” may cause issues.

For Net Station .mat files, enter the potential names of the variable that consists of the EEG data matrix. Between each entry, press your newline key (enter – Windows, return – Mac). When you are done entering potential variable names, enter done. If you don’t have any ideas of what your variable name could be, try loading the Net Station .mat file into an empty MATLAB workspace and checking the variables.

**Example:** Category\_1

If your data is a MATLAB matrix:

If your channel locations file is not included in HAPPE’s acquisition layouts:

Do you have a channel locations file for your data? [Y/N]

NOTE: A list of supported file types can be found in the HAPPE User Guide.

Unless HAPPE includes the channel locations for your acquisition layout by default, you will likely need to provide them for MATLAB matrixes in .mat format. Accepted channel locations file formats are those accepted by EEGLAB (Delorme & Makieg, 2004) and include:

1. .loc, .locs, .eloc – EEG polar coordinates
2. .ced – EEGLAB with polar, cartesian, and spherical
3. .sph – MATLAB spherical coordinates
4. .elc – Cartesian 3-D from EETrack
5. .elp – Polhemus Cartesian coordinates
6. .elp – BESA spherical coordinates
7. .xyz – MATLAB/EEGLAB Cartesian coordinates
8. .asc, .dat – Neuroscan Cartesian polar coordinates
9. .mat – Brainstorm channel locations
10. .sfp – BESA/EGI xyz Cartesian coordinates

If you do have a channel locations file,

Enter the name of the channel locations file, including the full path and file extension:

Enter the full path to where the channel locations file exists, followed by the name of the file, including its extension. The channel locations file does not need to be in a specific location on your computer or on an external drive, only that it is accessible.

**Example (Mac):** /Users/lgd/Desktop/

chanlocs.loc

**Example (Windows):** C:\Users\lgd\Documents\

chanlocs.loc

If you do not have a channel locations file,

HAPPE functionality is limited without channel locations. Continue anyway? [Y/N]

If you don’t have a channel locations file, you will still be able to run your data through HAPPE, granted with a limited set of steps. Without channel locations, you cannot filter to channels of interest, perform bad channel detection, interpolate bad channels, or re-reference your data. If this is acceptable, you can continue running HAPPE by entering Y (case insensitive). Otherwise, you can enter N (case insensitive) and HAPPE will stop running by throwing an error.

Do all your files share the same sampling rate? [Y/N]

If all your files share the same sampling rate, input Y (case insensitive). Otherwise, input N (case insensitive). Depending on your input, you may have to follow different prompts.

If your files do share a sampling rate:

Sampling rate:

Enter the sampling rate of all the files as an integer.

**Example:** 250

If your files do not share a sampling rate:

Enter the name of the file containing the sampling rates for each data file, including the path and file extension:

See the HAPPE User Guide for how this table should be formatted.

Enter the name of the file, including the full path and the file’s extension. The file does not have to be in any particular location on your computer or on an external drive; it only must be accessible. Additionally, the file could be in any format so long as it is readable as a table by MATLAB, but we recommend .csv or an Excel file.

An example for formatting this file is shown below. The first column should contain the file name while the second column should contain the sampling rate in Hz.

|  |  |
| --- | --- |
| samplefilename1.mat | 250 |
| samplefilename2.mat | 500 |

**Example (Mac):** /Users/lgd/Desktop/sampRate.csv

**Example (Windows):** C:\Users\lgd\Documents\

sampRate.csv

Path to .txt files containing task event info:

For .mat files, you will need to provide .txt files containing your event information. Enter the full path to the folder holding these files, including the name of the folder. This folder does not need to be on the same path as your data.

**Example (Mac):** /Users/lgd/Desktop/Task InfoFolder

**Example (PC):** C:\Users\lgd\Documents\Task Info Folder

If your file has channel locations included or provided:

Select channels of interest:

all = Select all channels in the data

coi = Select a user-specified subset of channels

To process all possible channels included in each EEG file, enter all. If you wish to only process a subset of channels, enter coi.

If only processing a subset of channels:

Choose an option for entering channels:

include = Include ONLY the entered channel names

exclude = Include every channel EXCEPT the entered

channel names

To enter your channel subset, you have the option to select the channels of interest or the channels of disinterest. If you select include all the channel names that you enter in the next prompt will be the only channels included in processing. Otherwise, by selecting exclude, you will instead process only the channels *not* entered in the next prompt. It may be easier to choose to include or exclude based on which list of channels is shorter.

Enter channels, including the preceding letter, one at a time.

Press enter/return between each entry.

Examples:

E17

M1

When you have entered all channels, input “done” (without quotations).

Enter the channels you wish to include or exclude, depending on your previous choice, one at a time. You should enter the channel name exactly as it appears in the data, case sensitive. This means you should include the preceding letter, if applicable (example: E17). If you have questions about your channel names, refer to your acquisition layout. Don’t use quotations around the channel name. Between each entry, press your newline key (enter for Windows, return for Mac). When you are finished entering all channels in your list of inclusion/exclusion, enter done.

Frequency of electrical (line) noise in Hz:

USA data probably = 60; Otherwise, probably = 50

Enter the frequency of electrical noise at the site of data collection. For locations within the United States, the frequency is likely 60 Hz, and outside is likely 50 Hz. We’ve included a link to a list of frequencies by country, but you should independently confirm that you have the correct frequency: <https://www.oaktreeproducts.com/img/product/description/List%20of%20Worldwide%20AC%20Voltages.pdf>

Are there any additional frequencies (e.g., harmonics) to reduce? [Y/N]

In some cases, there may be prominent line noise at a frequency other than the electrical noise frequency. Often, but not always, these are the harmonics of the electrical frequency. If there is enough noise at other frequencies that it affects the data, you can choose to reduce line noise at those frequencies by entering Y, case insensitive. Otherwise, enter N, case insensitive.

If reducing line noise in additional frequencies:

Enter the frequencies, one at a time, to pass through noise reduction.

When finished entering frequencies, enter “done” (without quotations).

Enter the frequencies you want to reduce, aside from the electrical noise frequency, that you want to reduce, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering frequencies, enter done.

**Example:** 25

Line noise reduction method:

cleanline = Use Tim Mullen’s CleanLine

notch = Use a notch filter (COMING SOON)

For reducing line noise, you have two options: using Tim Mullen’s CleanLine EEGLAB or using a notch filter. At this time, only the CleanLine option is available, so entering notch will end the HAPPE run via an error.

If using CleanLine:

Use legacy or default line noise reduction?

default – Default method optimized in HAPPE v2

legacy – Method from HAPPE v1 (NOT RECOMMENDED)

The legacy line noise reduction method uses the CleanLine plugin for EEGLAB from a far less effective version. For this reason, we strongly recommend using the default to process your data as it uses an improved version of CleanLine to reduce line noise. If you want to use the legacy method for comparison analyses or to finish existing data processing, enter legacy. For all other situations, you should enter default.

Resample data? [Y/N]

If you want to resample your data, enter Y. This might occur if you have different sampling rates across files and want them to be the same. Otherwise, enter N to not resample your data. Both options are case insensitive.

If resampling your data:

HAPPE supports resampling to 250, 500, and 1000 Hz.

Resample frequency:

Currently, HAPPE restricts resampling to the following values: 250, 500, or 1000 Hz. Choose one of these frequencies by entering the value.

**Example:** 500

Choose a filter:

fir = Hamming windowed sinc FIR filter (EEGLAB’s standard filter)

butter = IRR butterworth filter (ERPLAB’s standard filter)

You may choose to filter your ERP data to your ERP cutoffs using one of two options. The FIR filter is what is used for all non-ERP designs in HAPPE and is EEGLAB’s standard filter. To use this option, enter fir. However, you may also instead choose to filter using a butterworth filter. In this case, HAPPE will use ERPLAB’s IRR butterworth filter by entering butter.

If your file has channel locations included or provided:

Perform bad channel detection? [Y/N]

If you want remove channels with high impedances, damage to the electrodes, insufficient scalp contact, and/or excessive movement or electromyographic (EMG) artifact throughout the recording from subsequent analyses, enter Y, case insensitive. To include all your channels of interest regardless of quality, enter N, case insensitive.

If rejecting bad channels:

Bad channel detection method:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED).

You can choose to run either the default or the legacy version of bad channel detection. The legacy version was used in the original HAPPE software, is outdates, and we no longer support it, thus it is not recommended. Despite this, to select this option, enter legacy. The default method uses a new method of bad channel detection using the Clean Rawdata function with optimized criteria. To use this optimized method, enter default.

Method of wavelet thresholding:

default = Default method optimized in HAPPE v2.

legacy = Method from HAPPE v1 (NOT RECOMMENDED)

The new wavelet method uses an empirical Bayesian level-dependent threshold for the wavelets. Whether this is a soft or a hard threshold can be determined by the user in following prompts. Wavelet family is coiflet (level 4). This method has been optimized on EEG data (see Lopez et al., in press; Monachino et al., in press). To use this method, enter default. The legacy method of wavelet thresholding uses a soft, global threshold for the wavelets. The wavelet family is coiflet (level 5) and the threshold multiplier is used to remove more high frequency noise. This method has not been optimized, so it is not recommended, but if you wish to use it, enter legacy.

If using the default method for wavelet thresholding:

Threshold rule for wavelet thresholding:

soft – Use a soft threshold

hard – Use a hard threshold

HAPPE+ER allows the user to choose whether to use a hard or soft threshold for wavelet thresholding in ERP designs. Using a soft threshold slightly preserves ERP amplitude but may keep slightly more artifact in the data. Using a hard threshold removes more artifact but at the slight cost of ERP amplitude. For more information about the difference, refer to the HAPPE+ER manuscript. To use the soft threshold, enter soft. Otherwise, enter hard.

Segment data? [Y/N]

To segment your data, enter Y. Otherwise, enter N. If you choose to segment your data, you may be asked to answer additional prompts to properly segment your data.

If segmenting your data:

Segment start, in MILLISECONDS, relative to stimulus onset:

Example: -100

Enter the number of milliseconds present in the trial prior to your stimulus onset. If you have no baseline period within your trial, this may be 0. Otherwise, this is a negative number. For example, if you have a trial with a 100-millisecond baseline, you will enter -100.

Segment end, in MILLISECONDS, relative to stimulus onset:

Enter the number of milliseconds present in the trial after stimulus onset. This should always be a number greater than 0. For example, if you have a trial that runs for 500 milliseconds after the stimulus onset or trial start, you will enter 500.

**Example:** 500

Offset delay, in MILLISECONDS, between stimulus initiation and presentation:

NOTE: Please enter the total offset (combined system and task-specific offsets).

Enter the total offset delay between when the stimulus is initiated and when it is presented. This number should include the system offset as well as any task-specific offsets and be entered in milliseconds. Currently, HAPPE only supports a single offset value.

**Example:** 20

Perform baseline correction (by subtraction)? [Y/N]

If you would like to perform baseline correction on your data, input Y. Otherwise, input N. If you choose to perform baseline correction, you may be shown additional prompts.

If performing baseline correction:

Enter, in MILLISECONDS, where the baseline segment begins:

Example: -100

The number of milliseconds prior to the event tag where the baseline period starts. Entering a negative number will start the baseline period that number of milliseconds before the event tag. Entering 0 starts the segment at the event tag.

Enter, in MILLISECONDS, where the baseline segment ends:

NOTE: 0 indicates stimulus onset.

The number of milliseconds in the segment, relative to the onset tag, where the baseline period ends. Entering 0 will end the baseline period at stimulus onset.

If channel locations are included or provided and the data is segmented:

Interpolate the specific channels data determined to be artifact/bad within each segment? [Y/N]

This option enables evaluation of channels within epochs to determine whether there is bad data present. Channels flagged with bad data for that segment will then have their data interpolated only for that segment, using FASTER (Nolan et al., 2010). If you wish to interpolate bad channels within segments, enter Y. Otherwise, enter N.

Perform segment rejection? [Y/N]

Users can choose to reject segments that are determined to still have artifact remaining after waveleting and MuscIL (if selected). Criteria for rejection includes a choice of joint-probability criteria, amplitude-based criteria, or a combination of the two. If you would like to reject segments, enter Y. Otherwise, enter N.

If performing segment rejection:

Choose a method of segment rejection:

amplitude = Amplitude criteria only

similarity = Segment similarity only

both = Both amplitude criteria and segment similarity

The first option is using amplitude criteria only: amplitude-based criteria sets a minimum and maximum signal amplitude as the artifact threshold, with segments being removed when their amplitude exceeds either threshold. To run this option alone, enter amplitude. The second option is segment similarity criteria, which considers how likely a segment’s activity is to be artifact-laden given the activity of other segments for that channel and other channels’ activity for the same segment and removes outlier segments. The assumption is that artifact segments should be rare relative to the rest of the data at this point in the processing stream. To run this option alone, enter similarity. You can also choose to run both criteria by entering both. Depending on your choice, you may need to follow additional prompts.

If using amplitude criteria or both criteria:

Minimum signal amplitude to use as the artifact threshold:

Enter the minimum signal amplitude used for segment rejection.

**Example:** -150

Maximum signal amplitude to use as the artifact threshold:

Enter the maximum signal amplitude used for segment rejection.

**Example:** 150

Use all channels or a region of interest for segment rejection?

all = All channels

roi = Region of interest

If you plan to analyze all user-specified channels in your dataset, enter all. If you have a specific region of interest that you will be analyzing enter roi. Depending on your choice you may need to answer additional prompts.

If rejecting channels based on a region of interest:

Enter the channels in the ROI, one at a time.

When you have finished entering all channels, enter “done” (without quotations).

Enter the channels that you wish to include in your region of interest one at a time. Press your newline key (enter for Windows, return for Mac) between each entry. If you have any questions about your channel names, refer to your acquisition layout. As with entering your initial channels of interest, enter the names exactly as they appear in the dataset and without any quotations. When you have finished entering channels, enter done.

**Example:** O1

Use pre-selected “usable” trials to restrict analysis? [Y/N]

If you have previously determined which trials in your dataset are usable, you can choose to filter out the “unusable” trials using this option. THIS OPTION DOES NOT CURRENTLY FUNCTION!

If channel locations are included or provided:

Re-reference data? [Y/N]

If you want to re-reference your data, input Y. Otherwise, enter N.

If re-referencing:

Does your data contain a flatline or all zero reference channel? [Y/N]

Some datasets include a reference channel that is a flat line; in other words, each data point for that channel is zero. If your data includes such a channel, enter Y. Otherwise, enter N.

If your data includes a flatline reference channel:

Enter reference channel ID:

If unknown, press enter/return.

Enter the channel ID of your flatline reference channel exactly as it appears in the data. If you don’t enter the exact name, it can prevent HAPPE from processing correctly.

**Example:** Cz

Re-referencing type:

average = Average re-referencing

subset = Re-reference to a channel or subset of channels

rest = Re-reference using infinity with REST (Yao, 2001)

HAPPE offers three options for re-referencing: average re-referencing, re-referencing to a subset, or using REST. To use average re-referencing, input average. To re-reference to a single channel or to a subset of channels, enter subset. To re-reference to infinity using REST, enter rest. Depending on your choice, you may need to answer additional prompts.

If re-referencing to a subset of channels:

Enter channel/subset of channels to re-reference to, one at a time.

When you have entered all channels, input “done” (without quotations).

Enter the channels that you wish to re-reference to one at a time. As with entering your initial channels of interest, the names should be entered exactly as they appear in the data. If you have questions about your channel names, refer to your acquisition layout. Between each channel entry, press your newline key (enter for Windows, return for Mac). When you finish entering channels, enter done.

**Example:** M1

If re-referencing using REST:

Reference type:

average = Average reference

fixed = Fixed reference

mastoid = Ipsilateral and/or contralateral mastoid

Select whether your original reference is average, fixed (as in a single electrode), or mastoid (ipsilateral and/or contralateral mastoid). If you have questions about which option to use for your dataset, please refer to the REST documentation.

If your reference selection is mastoid:

Enter list of left channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the left side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T7

Enter list of right channels, one at a time.

When finished entering channels, enter “done” (without quotations).

Enter the list of the channels associated with the right side of the head, one at a time, pressing your newline key (enter for Windows, return for Mac) between each entry. These names should be entered exactly how they appear in the data. If you are uncertain about your channel names or placement, refer to your acquisition layout.

**Example:** T8

Load or calculate leadfield?

load = Load existing leadfield

calculate = Calculate new leadfield

If you have an existing leadfield that you want to use, enter load. Otherwise, if you don’t have a leadfield, you can have one calculated as part of the REST process by entering calculate. Depending on your choice, you may need to answer additional prompts.

If calculating the leadfield and reference selection is mastoid:

Enter XYZ coordinates for the left reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the left reference (often this is M1) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-37.1 89.8 -68.3]

Enter XYZ coordinates for the right reference:

Enter in this format: [x y z] (including brackets), then press enter/return.

Enter the XYZ coordinates for the right reference (often this is M2) electrode. These coordinates should be entered in a very specific format for HAPPE to re-reference properly: [x y z]. Each letter should be the corresponding coordinate value, and the square brackets need to be included. When properly formatted, press your newline key to submit the input. If you are uncertain about the coordinate values for your reference, refer to your acquisition layout.

**Example:** [-36.9 -88.8 -69.2]

If loading an existing leadfield:

Enter the name of the leadfield file, including path and file extension:

Enter the full name of the leadfield file including the path and file extension, examples below. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/

leadfield.dat

**Example (Windows):** C:\Users\lgd\Documents\

leadfield.dat

Format to save processed data:

1 = .txt file (electrodes as columns, time as rows) – Choose this for

ERP timeseries

2 = .mat file (MATLAB format)

3 = .set file (EEGLAB format)

For processing ERPs, we strongly recommend entering 1, which will provide .txt files of the timeseries as well as .set versions of the EEG. However, if you wish, you may also select either 2 or 3.

If visualizations are enabled:

Minimum value for power spectrum figure:

Enter the minimum frequency to plot for the power spectrum figure.

**Example:** 1

Maximum value for power spectrum figure:

Enter the maximum frequency to plot for the power spectrum figure.

**Example:** 45

Enter the frequencies, one at time, to generate spatial topoplots for:

When you have entered all frequencies, input “done” (without quotations).

Enter any frequencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the power spectrum for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the frequencies you wish to plot, enter done.

**Example:** 10

Start time, in MILLISECONDS, for the ERP timeseries figure:

Enter the start time, in milliseconds, to plot for the ERP timeseries figure.

**Example:** 0

End time, in MILLISECONDS, for the ERP timeseries figure:

NOTE: This should end 1+ millisecond(s) before your segmentation parameter ends (e.g., 299 for 300).

Enter the end time, in milliseconds, to plot for the ERP timeseries figure. Make sure that this ends before your segmentation parameters ends. For example, if your segment ends at 300 milliseconds, we recommend setting your value at 298 or 299.

**Example:** 498

Enter the latencies, one at a time, to generate spatial topoplots for:

When you have entered all latencies, input “done” (without quotations).

Enter any latencies for which you would like topoplots (spatial distribution of power in that frequency across the scalp) made. This image will appear in the same figure as the ERP timeseries for that file. Enter as a number, pressing your newline key (enter for Windows, return for Mac) between each entry. When you have finished entering the latencies you wish to plot, enter done.

**Example:** 100

Please check your parameters before continuing.

At this point, your parameters are listed out in the command window.

Are the above parameters correct? [Y/N]

Go through the parameters listed in the command window and check to see that your parameters are correct. If so, enter Y (case insensitive). Otherwise, if your parameters are not correct, enter N (case insensitive) to change them.

If changing or correcting a parameter:

Parameter to change:

HAPPE will list out the parameters you can change.

Choose a parameter to change by entering the name of the parameter exactly as it appears in the above prompt (excluding newline breaks). You will be prompted with the original command to change the parameter. Depending on your data and parameter set, you may be presented with different options. You can change as many parameters as you want but must change each parameter one at a time. Some selections may require you to answer more than one prompt. When you finish entering parameters, enter done.

**Example:** segment rejection

If you changed an existing parameter set or created a new set:

Parameter file save name:

default = Default name (inputParameters\_dd-mm-yyyy.mat)

custom = Create a custom file name

To create your own save name for the parameter set, input custom. You will then be prompted to enter your custom save name. If you do not want to create a custom label, you can use the default name, inputParameters\_dd-mm-yyyy.mat (using the current date), by entering default.

If using a custom name:

File name (Do not include .mat):

Enter a name for your saved parameter set. Name the file the same way you would if you were using your file explorer; no need to include the .mat extension as it will be automatically added.

**Example:** erpProcessing\_studyname

If a file with the parameter set name exists:

A set of input parameters with this name already exists.

overwrite = Overwrite the existing file

new = Save a new file with a different name

Entering multiple parameters sets on the same day or using the same name for a parameter set can result in multiple files vying for the same file name. The MATLAB default is to overwrite existing files, and if you want to replace the existing file with the new file, you can do so by entering overwrite. However, if you do not want to overwrite the file, enter new.

# HAPPE+ER’s generateERPs Add-On

## Run generateERPs

1. If ignoring bad channels, create the bad channels spreadsheet.
   * For detailed instructions, read through the following command line prompts section.
2. Navigate to the main HAPPE folder in your file browser.
3. Open generateERPs.m in MATLAB.
4. In the Editor tab, hit “Run.”
5. Follow the prompts in the command window of MATLAB
   * For detailed instructions regarding the prompts, see below.

## Following Command Line Prompts:

Prompts as they appear in the command window are written in a monospace font, like this, followed by a brief description of what to enter, with an example. Sometimes a certain choice will result in a different set of prompts, in which case they will be indented under a heading that describes the choice resulting in the branch in directions.

Enter the path to the folder containing the processed dataset(s):

Input the folder, including the full path, where the processed .txt data is located. All outputs (folders and documents) should be in the same file structure as they were created during the original HAPPE run.

**Example (Mac):** /Users/laurelg-d/Desktop/Data Folder/5 - processed

**Example (PC):** C:\Users\laurelg-d\Documents\Data Folder\5 – processed

Load pre-existing set of input parameters? [Y/N]

If you have already set parameters through generateERPs for either this or another dataset, you can choose to load these parameters instead of entering them from scratch. To use a previously saved set, enter Y (case insensitive); otherwise, enter N (case insensitive).

If loading pre-existing parameters:

Enter your parameter file, including the full path and file extension:

Enter the full name of the parameter file including the path and file extension, examples below. By default, parameters are saved in a folder named **input\_parameters**. The loaded file does not need to be on the same path as the datasets you are currently running.

**Example (Mac):** /Users/lgd/Desktop/Data Folder/input\_parameters/

genERPs\_parameters\_dd-mm-yyyy.mat

**Example (PC):** C:\Users\lgd\Documents\Data Folder\input\_parameters\

genERPs\_parameters\_dd-mm-yyyy.mat

Change an existing parameter? [Y/N]

If any of the loaded parameters do not match what you want for the current run, you can change them at this step by entering Y (case insensitive). Otherwise, enter N (case insensitive) to continue the run.

If creating a new parameter set:

Trial type:

average = Average over trials (AveOverTrials.txt)

individual = Individual trials (IndivTrial.txt)

When HAPPE outputs .txt files, it creates two different types per file: one that contains an average of trials, and another that contains all the trials individually. You can determine which is which by the tag included in the file name. If you want to examine only the average over trials, enter average. Otherwise if you want to examine the individual trials in each file, enter individual.

If running on individual trials:

Choose your export format:

sheets = A single Excel file with multiple sheets

files = Multiple .csv files

You have two options for exporting your data when running individual files. First, you can export your data in a single Microsoft Excel file comprised of multiple sheets by entering sheets. Alternatively, you can export your data in multiple .csv files by entering files. In either case, the individual sheet or file will either contain the results for one subject or the results for one calculated measure – determined by the user in a later prompt.

Select channels of interest:

all = Select all channels in the data

coi = Select a user-specified subset of channels

To process all possible channels included in each EEG file, enter all. If you wish to only process a subset of channels, enter coi.

If only processing a subset of channels:

Choose an option for entering channels:

include = Include ONLY the entered channel names

exclude = Include every channel EXCEPT the entered

channel names

To enter your channel subset, you have the option to select the channels of interest or the channels of disinterest. If you select include all the channel names that you enter in the next prompt will be the only channels included in processing. Otherwise, by selecting exclude, you will instead process only the channels *not* entered in the next prompt. It may be easier to choose to include or exclude based on which list of channels is shorter.

Enter channels, including the preceding letter, one at a time.

Press enter/return between each entry.

Examples:

E17

M1

When you have entered all channels, input “done” (without quotations).

Enter the channels you wish to include or exclude, depending on your previous choice, one at a time. You should enter the channel name exactly as it appears in the data, case sensitive. This means you should include the preceding letter, if applicable (example: E17). If you have questions about your channel names, refer to your acquisition layout. Don’t use quotations around the channel name. Between each entry, press your newline key (enter for Windows, return for Mac). When you are finished entering all channels in your list of inclusion/exclusion, enter done.

Include bad channels in calculating ERP?

include = Keep bad channels

exclude = Remove bad channels

During HAPPE+ER’s processing, bad channels may have been detected in the dataset. If this is true, you can choose whether to include the interpolated bad channels or to exclude them from the ERP timeseries. If you decide to exclude the bad channels, enter exclude; otherwise, enter include. Depending on your choice you may need to answer additional prompts.

If removing bad channels:

Enter the file containing the bad channels, including the complete path.

Refer to the HAPPE User Guide for instructions on creating this file and an example.

Prior to running the generateERPs script, you should make a spreadsheet (.xlsx, .csv, etc.) containing a list of the files and the bad channels associated with each file. The easiest way to do this is to copy the first column (labeled “Row”) and the “Bad\_Chan\_IDs” column from the HAPPE\_dataQC output created at the time the data was processed and create a new spreadsheet using those. Make sure to remove any rows for files you are not examining.

|  |  |
| --- | --- |
| Data File Name | Bad Channels |
| file01.raw | F3 F4 T6 CZ E31 |
| file02.raw | T6 O1 O2 P3 P4 CZ E6 |

**Example:**

Enter the path to where the spreadsheet is saved followed by a slash (forward-slash or backslash depends on your OS) and the name of the spreadsheet, including the file extension (e.g., .csv). It does not matter where the file is stored, only that it is accessible.

**Example (Mac):** /Users/lgd/Documents/badChannels.csv

**Example (Windows):** C:\Users\lgd\Documents\badChannels.csv

Calculate ERP values? [Y/N]

This script has the capability to calculate values commonly measured from ERPs, including peak amplitudes and their latencies, mean amplitudes, and area under the curve. If interested in calculating these values on your data, enter Y (case insensitive); otherwise enter N (case insensitive). Depending on your choice, you may need to answer additional prompts.

If calculating values:

Enter latency windows of interest with anticipated peak:

Enter each window as two consecutive numbers followed by “max” or “min” (without quotations).

Press Enter/Return between entries.

When you have entered all windows, input “done” (without quotations).

Example: 10 100 max

Enter your latency periods as two consecutive numbers, with the first number representing the starting latency for your window and the second as the ending latency for your window. Try to use latencies that are included in your data or the script will correct your boundaries to the closest existing latency. Windows cannot include a negative latency value.

For each window you must also specify whether you anticipate a maximum or a minimum in the provided window. You can do so by including max or min, respectively, following the two numbers representing the window’s boundaries. If you want to look for both a maximum and a minimum within the same window, you must enter the latency window twice and alternate specifying max and min.

**Example:** 50 90 min

Choose a method for calculating mean amplitude:

windows = Restrict calculations to the user-specified

latency window(s)

zeros = Calculate using points where the amplitude is 0

both = Calculate by both windows and zeros

You can choose whether to calculate the mean amplitude using the windows you specified as bound, using zero crossings present in the dataset as bound, or using both methods. Choosing windows restricts the mean amplitude calculations to only the latency windows you specified previously. Choosing zeros allows the script to find the zero-crossings in the dataset and use those points as boundaries to create new windows in which to calculate the mean amplitude.

Choose a method for calculating area under the curve:

windows = Restrict calculations to the user-specified

latency window(s)

zeros = Calculate using points where the amplitude is 0

both = Calculate by both windows and zeros

You can choose whether to calculate area under the curve using the windows you specified as bounds, using zero crossings present in the dataset as bounds, or using both methods. Choosing windows restricts all area under the curve calculations to only the latency windows you specified previously. Choosing zeros allows the script to find the zero-crossings in the dataset and use those points as boundaries to create new windows in which to calculate the area under the curve.

If running on individual trials:

Choose an option:

subjects = Rows as trials, columns as values split

by subject

values = Rows as subjects, columns as trials split

by value

NOTE: See HAPPE User Guide for examples.

To output calculated values in the correct manner, you have to choose how to divide the data. We currently allow you to split either by subjects or by values. Splitting by subjects will output the data for each subject as its own sheet or csv (depending on your previous choice) with each sheet/csv comprised of the trials as rows and the columns as the calculated values. To use this option, enter subjects. Splitting the data by values will output the data for each calculated value as its own sheet or csv (depending on your previous choice) with each sheet/csv comprised of the subjects as rows and the columns as trials. To better comprehend the format, we have included a visual example below.

**By Subjects:** 

**By Values:** 

Please check your parameters before continuing.

At this point, your parameters are listed out in the command window.

Are the above parameters correct? [Y/N]

Go through the parameters listed in the command window and check to see that your parameters are correct. If so, enter Y (case insensitive). Otherwise, if your parameters are not correct, enter N (case insensitive) to change them.

If changing or correcting a parameter:

Parameter to change: ave/indiv trials, channels of interest, bad channel inclusion, calculating values.

Enter “done” (without quotations) when finished changing parameters.

Choose a parameter to change by entering the name of the parameter exactly as it appears in the above prompt (excluding newline breaks). You will be prompted with the original command to change the parameter. You can change as many parameters as you want but must change each parameter one at a time. Some selections may require you to answer more than one prompt. When you finish entering parameters, enter done.

**Example:** channels of interest

If you changed an existing parameter set or created a new set:

Parameter file save name:

default = Default name (genERP-parameters\_dd-mm-yyyy.mat)

custom = Create a custom file name

To create your own save name for the parameter set, input custom. You will then be prompted to enter your custom save name. If you do not want to create a custom label, you can use the default name, genERP\_parameters\_dd-mm-yyyy.mat (using the current date), by entering default.

If using a custom name:

File name (Do not include .mat):

Enter a name for your saved parameter set. Name the file the same way you would if you were using your file explorer; no need to include the .mat extension as it will be automatically added.

**Example:** VEP\_studyname

Enter the suffix used for this dataset, including stimulus tag (if applicable).

If no extension beyond \_AveOverTrials/\_IndivTrial, press enter/return.

Files created by HAPPE as a default are named *filename\_processed\_AveOverTrials* or *filename\_processed\_IndivTrial*. In the case of a re-run, they may have additional text following “AveOverTrials” or “IndivTrial”, for example, *filename\_processed\_AveOverTrials\_rerun2021*. If there is additional text, enter all the text that follows “AveOverTrials”/“IndivTrial”. In this example, you would enter \_rerun2021. If there is no additional text, you can simply hit your newline key - “enter” on Windows keyboards and “return” on Mac keyboards. Regardless of additional text, you do not need to include “.txt”.

**Example:** \_VEP1

# Data Quality and Pipeline Quality Outputs

We include this information to be of use in decided which data files to include in your analyses. Please note that the example thresholds we give are **not a global standard** for all data, simply recommendations.

## Tracking Data Quality

The two quality outputs can be used to examine data fidelity during collection and to track quality over time. This is useful for identifying issues that arise on the data collection end, including but not limited to failing electrodes on a net and the presence of a cell phone within the testing environment against protocol. Below, how to use these measures is detailed.

### Data Quality Measures

File Length:

Percent Good Channels:

* Look at this value across files. If this value is consistently low, it may indicate an issue. Firstly, there may be technical issues with the cap – including that it is no longer functioning correctly. Alternatively, this may indicate that there needs to be training in properly placing the cap, debriding the scalp, and ensuring proper connection between electrodes and the scalp.

Interpolated Channel IDs:

* Look at the names of the interpolated channels throughout the files. If there is a channel name that is consistently listed, it is worth checking out the cap to see if the electrode is damaged or otherwise not functioning correctly.

Number of Segments Post Segment Rejection:

### Pipeline Quality Measures

R Pre/Post Line-Noise Reduction:

* Look at the r value for your line-noise frequency (50 or 60 Hz). If there is a file that has a significantly lower value than the others, for example a 0.3 when most files are around 0.8-0.9, that indicates there was an additional source of line noise in the testing environment. This usually occurs when a cell phone is brought into the room or if proper shielding procedures are not followed. If there are many instances of low values, it may indicate a need to address protection against line-noise in the testing protocol.

## Using Outputs to Exclude Files

A primary use of the two assessment outputs generated as part of a HAPPE run is to examine the quality of your data and the performance of the HAPPE software on your dataset. You can use the measures to determine which individual files meet the standards needed for further analysis, removing those that do not qualify. Below, how each measure can be used is detailed, including recommended (but not global) thresholds you can use in your quality check.

### Data Quality Measures

File Length: *Does your file have enough data?*

* Remove files that do not have enough data. This could be an insufficient length of time for a particular paradigm, or an insufficient amount of time needed to conduct your analyses. As this threshold varies wildly depending on your experiment and your analyses, we are not able to make a recommendation for this measure.

Percent Good Channels Selected: *Do you have enough channels?*

* A threshold for this measure varies based on the experiment, the analyses, and the density of your acquisition layout. However, we recommend that at least 70% of your channels of interest are good.

Interpolated Channel IDs: *Are your channels of interest good?*

* The interpolated channels are channels that are marked as ‘bad’. If you have less than 2 channels remaining in your channels of interest or your region of interest after removing the listed bad channels, the file should be excluded.

Channels Interpolated for Each Segment: *How much data has been interpolated?*

* If you have chosen to interpolate your channels, this lists the channels in each segment that have been interpolated. We recommend having no more than 30% of your data interpolated for an individual file, but this threshold can be set at your discretion.

Number of Segments Post Segment Rejection: *Do you have enough trials/segments?*

* For ERP data we recommend having 20 segments or more. This should also apply to the number of segments for each stimulus tag/condition when you have multiple stimulus tags/conditions.
* For data with multiple tags/conditions, we suggest removing files that do not include all your conditions or an insufficient number of segments for any one condition, excluding cases in which the file should not have a stimulus tag/condition present in accordance with your study design.
* For other paradigms…

### Pipeline Quality Measures

R Pre/Post Line-Noise Removal: *Does HAPPE effectively reduce line-noise in your data?*

* Look at the value for all frequencies and for each frequency.
* Values closer to 1 are better. It is okay if the r value at your line noise frequency (either 50 or 60 Hz) is low if the values within 2 Hz of that frequency are close to 1.

RMSE

* Do not use for ERP data.

MAE

* Do not use for ERP data.

SNR and PeakSNR

* Do not use for ERP data.

R Pre/Post Wavelet Thresholding: *Does HAPPE effectively apply wavelet-thresholding on your data?*

* Look at the value for all frequencies and for each frequency included in your post-processed data. For example, if you filter your ERP data at 30 Hz, you need not consider the 45 and 75 Hz values.
* Values closer to one are better, but not as strictly as with line noise. Exclude files where r for any relevant frequency is lower than 0.2. For a more stringent threshold, exclude files where r for any relevant frequency is lower than 0.5.

## Variables as Co-Variates and Reporting Measures

A goal of the HAPPE software is to encourage standardization and to enable reporting and comparison across studies. Below, we describe how each measure can be included as a co-variate and/or a reporting measure in manuscripts.

### Data Quality Measures

### Pipeline Quality Measures

R Pre/Post Wavelet Thresholding

* Can be used as a continuous co-variate.
* Report the average value per study group and across all groups.

# Optimized Code Examples

## HAPPE 2.0

## HAPPILEE

### Bad Channel Rejection

EEG = pop\_clean\_rawdata(inEEG, 'FlatlineCriterion', 5, 'ChannelCriterion', ...

.1, 'LineNoiseCriterion', 20, 'Highpass', 'off', 'BurstCriterion', ...

'off', 'WindowCriterion', 'off', 'BurstRejection', 'off', 'Distance', ... 'Euclidian') ;

EEG = pop\_clean\_rawdata(EEG, 'FlatlineCriterion', 'off', 'ChannelCriterion', ...

.7, 'LineNoiseCriterion', 2.5, 'Highpass', 'off', 'BurstCriterion', ...

'off', 'WindowCriterion', 'off', 'BurstRejection', 'off', 'Distance', ...

'Euclidian') ;

EEG = pop\_rejchan(EEG, 'elec', [1:EEG.nbchan], 'threshold', [-2.75 2.75], ...

'norm', 'on', 'measure', 'spec', 'freqrange', [1 100]) ;

### Wavelet Thresholding

wavLevel = 10 ;

ThresholdRule = 'Hard' ;

artifacts = wdenoise(reshape(EEG.data, size(EEG.data, 1), [])', wavLevel, ...

'Wavelet', 'coif4', 'DenoisingMethod', 'Bayes', 'ThresholdRule', ...

ThresholdRule, 'NoiseEstimate', 'LevelDependent')' ;

### Segment Rejection

% APPLY SIMILARITY CRITERIA:

if strcmpi(params.seg\_rej\_method, 'similarity') || strcmpi(params.seg\_rej\_method, 'both')

if params.lowDensity; num = 2; else; num = 3 ; end

if ~params.ROI\_channels\_only

EEG = pop\_jointprob(EEG, 1, [1:EEG.nbchan], num, num, ... params.vis.enabled, 0, params.vis.enabled, [], params.vis.enabled) ;

else

EEG = pop\_jointprob(EEG, 1, [ROI\_indices\_in\_selected\_chanlocs]', ... num, num, params.vis.enabled, 0, params.vis.enabled, [], ... params.vis.enabled) ;

end

end

## HAPPE+ER

### Filtering to Specified Frequencies

EEG = pop\_eegfiltnew(EEG, [], 100, [], 0, [], 0) ;

% START CODE FOR OTHER PROCESSING STEPS

...

% END CODE FOR OTHER PROCESSING STEPS

% If the user indicated to use ERPLAB's butterworth

% filter, filter using the butterFilt function (adapted

% from ERPLAB's functions - see butterFilt

% documentation for more information). Currently only

% available for ERP paradigms.

if params.filt.butter

EEG = butterFilt(EEG, params.paradigm.ERP.lowpass, ...

params.paradigm.ERP.highpass) ;

% If the user indicated to use EEGLAB's FIR filter,

% filter using the EEGLAB function.

else

EEG = pop\_eegfiltnew(EEG, params.paradigm.ERP.highpass, ...

params.paradigm.ERP.lowpass, [], 0, [], 0) ;

end

### Wavelet Thresholding

if EEG.srate > 500; wavLevel = 13 ;

elseif EEG.srate > 250 && EEG.srate <= 500; wavLevel= 12 ;

elseif EEG.srate <= 250; wavLevel = 11 ;

end

ThresholdRule = 'Soft' ;

artifacts = wdenoise(reshape(EEG.data, size(EEG.data, 1), [])', wavLevel, ...

'Wavelet', 'coif4', 'DenoisingMethod', 'Bayes', 'ThresholdRule', ...

ThresholdRule, 'NoiseEstimate', 'LevelDependent')' ;

### Segmenting

if params.ERP\_analysis

samples\_offset = srate \* params.task\_offset/1000 ;

for i = 1:size(EEG.event, 2)

EEG.event(i).latency = EEG.event(i).latency + samples\_offset ;

end

end

EEG = pop\_epoch(EEG, params.task\_onset\_tags, params.task\_segment\_start ...

params.task\_segment\_end], 'verbose', 'no', 'epochinfo', 'yes') ;

### Split EEG by Tags

eeg\_byTags = [] ;

for i=1:length(params.task\_onset\_tags)

try

eeg\_byTags = [eeg\_byTags pop\_selectevent(EEG, 'type', ...

params.task\_onset\_tags{i})] ;

dataQC\_erp{current\_file, i\*2} = length(eeg\_byTags(i).epoch) ;

catch

dataQC\_erp{current\_file, i\*2} = 'ERROR' ;

fprintf('No instances of tag %s appear in this file', ...

params.task\_onset\_tags{i}) ;

end

end