**TITLE:**

An open-source EEGLAB plugin to jointly analyze EEG and cardiovascular (ECG/PPG) signals.

**AUTHORS AND AFFILIATIONS:**

Cédric Cannard1,2, Helané Wahbeh2, Arnaud Delorme1,2,3

1Centre de Recherche Cerveau et Cognition (CerCo), CNRS, Toulouse III University, France

2Institute of Noetic Sciences (IONS), Novato, California, USA.

3Swartz Center of Computational Neuroscience (SCCN), INC, UCSD, La Jolla, USA

Corresponding Author:

Cedric Cannard

ccannard@noetic.org

Tel: (510) 326-1734

Email Addresses of Co-authors**:**

Helané Wahbeh hwahbeh@noetic.org

Arnaud Delorme adelorme@ucsd.edu

**KEYWORDS:** *(6 minimum, 12 maximum)*

open-source, EEGLAB, EEG, ECG, PPG, HRV, feature-analysis, heartbeat-event potentials/oscillations (HEP/HEO), heart component removal.

**SUMMARY:** *(10-word minimum, 50-word maximum)*

The BrainBeats toolbox is an open-source EEGLAB plugin designed to jointly analyze EEG and cardiovascular (ECG/PPG) signals. It offers three main protocols: heartbeat-evoked potentials assessment, feature-based analysis, and heart artifact extraction from EEG. This will aid researchers and clinicians in studying brain-heart interactions with enhanced reproducibility and accessibility.

**ABSTRACT:** *(150-word minimum, 300-word maximum)*

The link between the cortical and cardiovascular systems is garnering increased attention due to its potential to offer valuable insights into brain and heart function coupling. Current joint analysis methodologies often employ invasive (PET) or high-cost (MEG, fMRI) neuroimaging methods. In contrast, EEG and ECG/PPG are non-invasive, cost-effective, and mobile alternatives for monitoring brain and cardiovascular activity. This facilitates data collection from more extensive and diverse samples in real-world settings. Nonetheless, analyzing these biosignals is challenging due to their noisy, nonstationary, and nonlinear nature (Klonowski 2009). While there is a history of independent EEG and heart rate variability (HRV) data investigation, the processing and statistical methods utilized suffer from a lack of consensus among scholars and clinicians, with essential steps often overlooked in publications or requiring complex programming skills, thereby hindering reproducibility.

Furthermore, a notable gap in this field has been the absence of standardized feature extraction and analysis, which can enhance the efficacy of clinical diagnostics (e.g., qEEG) and improve the robustness of machine learning models to interpret brain-cardiovascular interactions. To address these issues, we introduce the BrainBeats toolbox, an open-source EEGLAB plugin offering a comprehensive suite of signal processing and feature-extraction functions. This toolbox adheres to the latest guidelines in its default parameters and incorporates three primary method protocols: 1) Heartbeat-evoked potentials (HEP) and oscillations (HEO); 2) Feature-based analysis for identifying associations and differences between EEG and HRV features (time, frequency, and nonlinear domains); and 3) Automated extraction of heart artifacts from EEG signals through independent component analysis (ICA). Accompanied by sample data and step-by-step guidance, BrainBeats should ensure user-friendly engagement and reproducibility. This open-source toolbox should present an asset for researchers and clinicians investigating brain-heart interplays and can be effortlessly adapted and extended to cater to distinct research requirements.

**INTRODUCTION:** *(150-word minimum, 1500-word maximum, 2-paragraph minimum)*

The purpose of this method is to ease the investigation of relationships between the brain and the heart by facilitating the joint analysis of electroencephalography (EEG) and cardiovascular signals, namely electrocardiography (ECG) and photoplethysmography (PPG) while implanting the latest guidelines from experts in this field. This tool addresses limitations from existing methods and is made open source to facilitate accessibility and reproducibility in the area. The proposed toolbox should serve as a valuable resource for researchers and clinicians interested in removing cardiac artifacts from EEG signals, extracting features from EEG and ECG/PPG signals, or studying the relationship between brain and cardiovascular activity. Ultimately, this toolbox aims to pave the way for more in-depth investigations into the complex interplay between the brain and heart systems.

For a long time, the reductionist approach has dominated scientific inquiry in human physiology and cognition. This approach involved dissecting complex bodily and mental processes into smaller, more manageable components, allowing researchers to focus on individual systems in isolation. This strategy arose due to the immense challenge of studying the intricate and interconnected nature of the human body and mind (von Bertalanffy, 1968). Reductionism has been instrumental in understanding individual subsystems in isolation, such as elucidating the role of ion channels and action potentials for neural (Hodgkin & Huxley, 1952) and cardiac (Bean et al., 1984) communication. However, a significant gap remains in understanding how these isolated systems interact on a larger spatial and temporal scale. Thanks to recent technological advances, a more integrative approach is gaining interest. Innovations in hardware, signal processing, data storage, and computing capabilities have enabled researchers to collect and analyze electrophysiological signals from different bodily systems simultaneously (e.g., heart, brain, skin, etc.; Jiang et al. 2019; Insel et al., 2017; Kappenman & Luck, 2010). This integrative – or multimodal – approach will pave the way for a more comprehensive understanding of human physiology, cognition, pathology, and consciousness (Bunge, 2003; Fuchs, 2018; Thayer & Lane, 2009). Consequently, this multimodal approach is now considered an essential component of modern research, complementing traditional reductionist methods and offering novel insights into the synergistic mechanisms governing the human body and mind.

**Heart-brain research: which measures?**

Studying the intricate relationship between the brain and the heart can yield valuable insights into the underlying physiology and anatomy of the human body, ultimately leading to the development of novel diagnostic and therapeutic tools. The relationship between the heart and the brain has been studied via neuroimaging methods such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET). Using these tools, researchers highlighted some brain regions associated with cardiovascular control (e.g., manipulation of heart rate and blood pressure; Critchley 2000; Napadow 2008), showed the influence of heart rate on the BOLD signal (Chang et al., 2009; Shmueli 2007), or identified potential brain-body pathways contributing to coronary heart disease (i.e., stress-evoked blood pressure; Gianaros 2009).

While these studies have significantly advanced our understanding of the complex interplay between the central nervous system (CNS) and cardiovascular function, these neuroimaging techniques are expensive, have limited availability, and are confined to controlled laboratory settings, which restricts their practicality for real-world and large-scale applications. In contrast, EEG and ECG/PPG are more affordable and portable tools that offer the potential for studying brain-heart interactions in more diverse settings and populations or over longer periods, providing new opportunities for investigating the dynamic relationship between brain and heart function. ECG measures the electrical signals generated by the heart when it contracts and relaxes via electrodes placed on the skin (usually on the chest, arms, or legs). PPG measures blood volume changes in the microvascular tissues using a light source (e.g., LED) and a photodetector placed on the skin (commonly on a fingertip, earlobe, or forehead). Since blood absorbs more light than the surrounding tissue, the PPG signal can be used to estimate blood flow and pulse rate. Both methods provide valuable information about cardiovascular function but serve different purposes and offer distinct data types. As such, the use of EEG and ECG/PPG holds great promise for advancing our understanding of the physiological, cognitive, and emotional processes underlying brain-heart interactions and their implications for human health and well-being. Like ECG, EEG records the electrical fields generated by the synchronized activity of thousands of cortical neurons by placing electrodes on the scalp.

**The two approaches for jointly analyzing these signals**

There are two main approaches to studying interactions between EEG and cardiovascular signals:

1. The heartbeat-evoked potentials (HEP) for the time domain (i.e., ERP) and heartbeat-evoked oscillations (HEO) for the time-frequency domain (i.e., event-related spectral perturbations). This approach examines how the brain processes cardiovascular activity with millisecond accuracy and requires that both time series are time-locked, the heartbeats be marked with events in the EEG signals, and the heart signal removed. This approach is the most popular (Schandry 1981; Couto 2015; Al 2020; Antonacci 2022; Banellis and Cruse 2020; Baranauskas 2020; Billeci 2021; Bogdány 2022; Canida-Rivera 2021; Chang 2022; Fuseda 2021; Immanuel 2014; Kumral 2022; Lechinger 2015; Montoya 1993; Park 2021; Pollatos 2004; Schmitz 2021; Zwienenberg 2023).
2. Feature-based: this approach extracts EEG and heart-rate variability (HRV) features from continuous signals and examines associations between them. This has been done with ECG (Thayer et al. 2012; Mather 2018; Kemp 2013) and PPG to a lesser extent (Khosrow-Khavar 2014). This approach presents promising applications by capturing both the state-related and trait-related variables. Note that, for both EEG and cardiovascular signals, the longer the recording, the more dominant the trait variable (Shaffer Ginserbg; Coan and Allen 2003; Hagemann 2005). Thus the applications depend on the recording parameters. Feature-based analyses (including qEEG or feature-based machine learning) are growing in interest for both clinicians and scholars, providing new objective metrics for making early predictions/forecasting of the development of mental and neurological disorders, of treatment-response, or of relapse (Widge et a. 2019; Olbrich 2013; Sheikhani 2012; Yasin 2021; Kumar 2014; Acharya 2012; Adeli 2010; de Aguiar Neto 2019; Cannard 2021). This approach is compelling with large datasets and real-world settings, which can be more easily obtained thanks to the recent innovations in wearable neurotechnology (Cannard et al., 2020).

**The advantages of alternative methods**

While many tools exist to process cardiovascular and EEG, and HRV signals independently from one another, none is currently available for jointly analyzing them. Furthermore, the tools available to process cardiovascular signals require expensive license purchase, do not allow processing large datasets in batch via command line, have proprietary algorithms that limit reproducibility, or require advanced programming skills by not providing a graphical user interface (GUI). To our knowledge, three open-source MATLAB toolboxes exist to support HEP analysis with a GUI. The *ecg-kit* toolbox ([REF](https://github.com/marianux/ecg-kit); Demski and Soria, 2016), the *HEPLAB* EEGLAB plugin ([REF](https://github.com/perakakis/HEPLAB)), and the *CARE-rCortex* EEGLAB plugin ([Grosselin et al. 2018](https://www.sciencedirect.com/science/article/abs/pii/S0165027018302474)). While *HEPLAB* and *ecg-kit* facilitate HEP analysis by detecting heartbeats and marking them in the EEG signals, they do not provide statistical tools for analysis and are limited to the time domain (i.e., ERP). The *CARE-rCortex* pluginaddressed these issues by supporting ECG and respiratory signals, time-frequency domain analysis, statistics, and advanced baseline normalization and correction methods adapted to HEP analysis. However, it uses the Bonferroni method for multiple comparison correction, which is too conservative for EEG analysis (Ref), and it does not offer command-line access for processing and analyzing large datasets. Furthermore, recent studies recommend against baseline correction methods, as they reduce the signal-to-noise ratio (SNR) and are "statistically unnecessary and undesirable" ([Alday 2019](https://onlinelibrary.wiley.com/doi/abs/10.1111/psyp.13451); Delorme 2023).

To address these limitations, we introduce the *BrainBeats* Matlab toolbox, implemented as an open-source EEGLAB plugin designed to process and analyze EEG and ECG/PPG signals jointly. It incorporates the following advantages over previous methods:

I) An easy-to-use GUI (for non-programmers) and command line mode (for programmers aiming to perform automated processing and analysis on large datasets using more advanced parameters).

II) Implementation of validated algorithms, parameters, and guidelines for processing cardiovascular signals, such as detecting R peaks, interpolating RR artifacts, and computing HRV metrics (e.g., implanting guidelines for windowing, resampling, normalization, etc.; Vest et al. 2019; Schaffer Ginsberg 2017; Pham 2021). This is important because Vest et al. (2018) demonstrated how modest differences in these processing steps can lead to divergent results, contributing to the lack of reproducibility and clinical applicability of HRV metrics.

III) Implementation of validated algorithms, parameters, and guidelines for processing EEG signals, including re-referencing to infinity when at least 30 channels are available (Yao 2001, 2005; 2017; Dong 2017; Candia-Rivera 2021), removal of abnormal channels and artifacts (Mullen 2015; Delorme 2004), interpolation of removed channels, optimized parameters for ICA decomposition and classification of independent components (Ablin et al. 2017 and 2018; Frank 2022; Kim et al. 2023; Pion-Tonachini 2019), guidelines filtering windowing and parameters (Winkler et 2015; Widmann; Smith et al. 2017; Gross et al. 2013; Delorme 2023). Note: users can also use the toolbox with processed data and turn off processing steps.

IV) Heartbeat-evoked potentials (HEP, i.e., time domain) and oscillations (HEO; event-related spectral perturbations with wavelet or FFT methods, and inter-trial coherence are available). Advanced statistics, including hierarchical linear modeling, which accounts well for within and between-subjects variance, with weighted least square optimization to downweigh remaining artifactual trials (WLS; Pernet 2022). Nonparametric statistics include permutation statistics and spatiotemporal corrections for multiple comparisons (Maris and Oostenveld 2007; Pernet XX). Note: statistical analyses can also be performed in the independent component domain.

V) Supports the joint extraction of EEG and HRV features for the first time. These features can be analyzed separately (direct differences between groups/conditions, or associations with third variables) and jointly (same but from correlation coefficients). See above for applications.

VI) The toolbox provides various data visualizations to inspect signals at various necessary processing steps and outputs at the subject level (see Representative results for illustration).

**Information to help readers decide whether the method described is appropriate for them**

This toolbox is appropriate for any researcher or clinician with EEG and ECG/PPG data. The plugin does not yet support importing EEG and ECG/PPG signals from separate files (although this feature will be available soon). The toolbox is appropriate for anyone aiming to perform HEP/HEO analysis, extract EEG and/or HRV features with standardized methods, or simply remove heart artifacts from EEG signals.

**PROTOCOL:** *(1-page minimum, 10-page maximum)*

1. **Method 1: Heartbeat-evoked potentials (HEP)**
   1. Load data into EEGLAB: *File > Load existing dataset > select "sample\_data1.set"*. See **Figure 1.1.** Note: this dataset was collected with a "wet" Biosemi system during a mind-wandering (shortened to 5 minutes to facilitate replication of the following steps) with 64 EEG channels and 2 ECG channels.

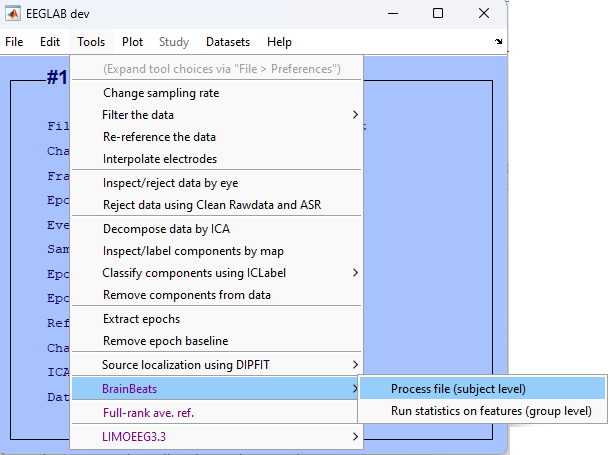


Figure 1.1. Main EEGLAB menu to launch BrainBeats' general user interface (GUI) to select processing parameters.

* 1. Open the general user interface (GUI) to select parameters: Tools > BrainBeats > Process file (subject level).
  2. Select "Heartbeat-evoked potentials HEP" as analysis to run, "ECG" as heart data type, click on the button to display the list of channels to select the ECG channels labeled "ECG1" and "ECG2" (or type the channel labels directly in the text box next to the button. Select "Shape-preserving piecewise cubic interpolation" as the cleaning method of RR artifacts and clean EEG data. Select "Clean EEG" to process the EEG data, plot and save outputs, and click "Ok" to launch. See the overview in **Figure 1.2**. All steps are automated (see **Representative Results**).

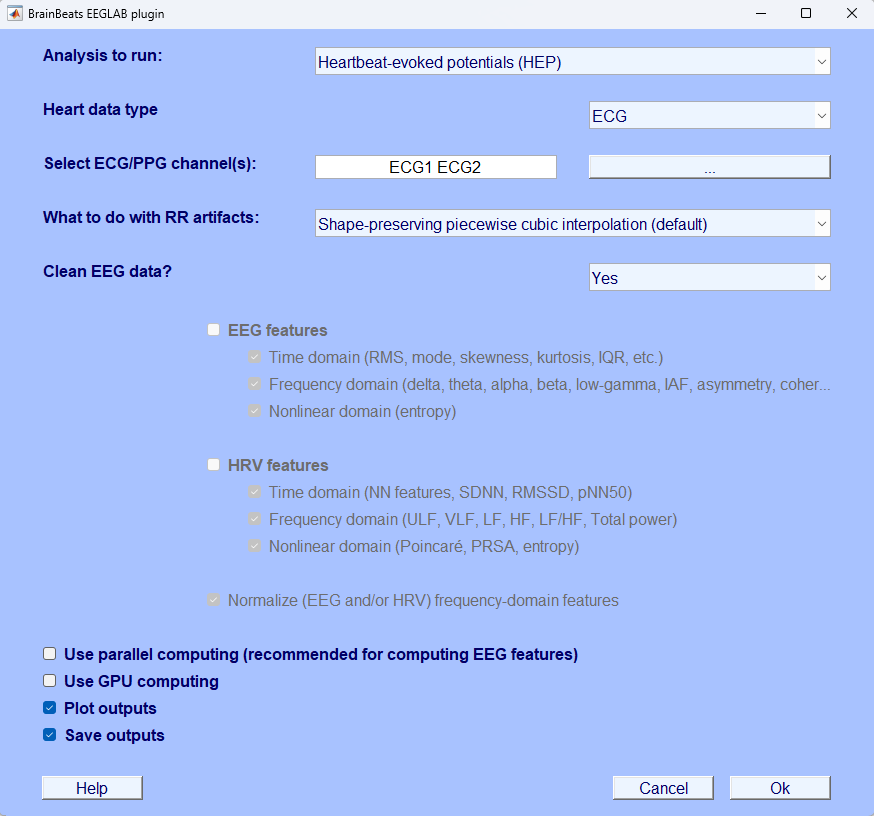


Figure 1.2. Using the GUI to select parameters for preparing data for heartbeat-evoked potentials (HEP) analysis.

Advanced users can perform all the above steps with the following command lines:

eeglab; close;

mainDir = fileparts(which('eegplugin\_BrainBeats.m')); cd(mainDir);

EEG = pop\_loadset('filename','sample\_data1.set','filepath',fullfile(mainDir, 'sample\_data'));

EEG = brainbeats\_process(EEG,'analysis','hep','heart\_signal','ECG', ...

'heart\_channels',{'ECG1' 'ECG2'},'clean\_rr','pchip','clean\_eeg',true, ...

'parpool',false,'gpu',false,'vis',true,'save',true);

1. **METHOD 2: Extract EEG and HRV features from continuous data**
   1. Load the dataset and launch the BrainBeats GUI as for METHOD 1. Select "Extract EEG & HRV features from continuous data", the same parameters for ECG channels, RR interpolation method, and "clean EEG". Notice that the EEG and HRV feature fields are now unlocked. Check both to extract EEG and HRV features. All domains (time, frequency, nonlinear) are set by default. Check "Use parallel computing" to increase computation speed. "Plot and save outputs" are set by default. Click "Ok" to launch.

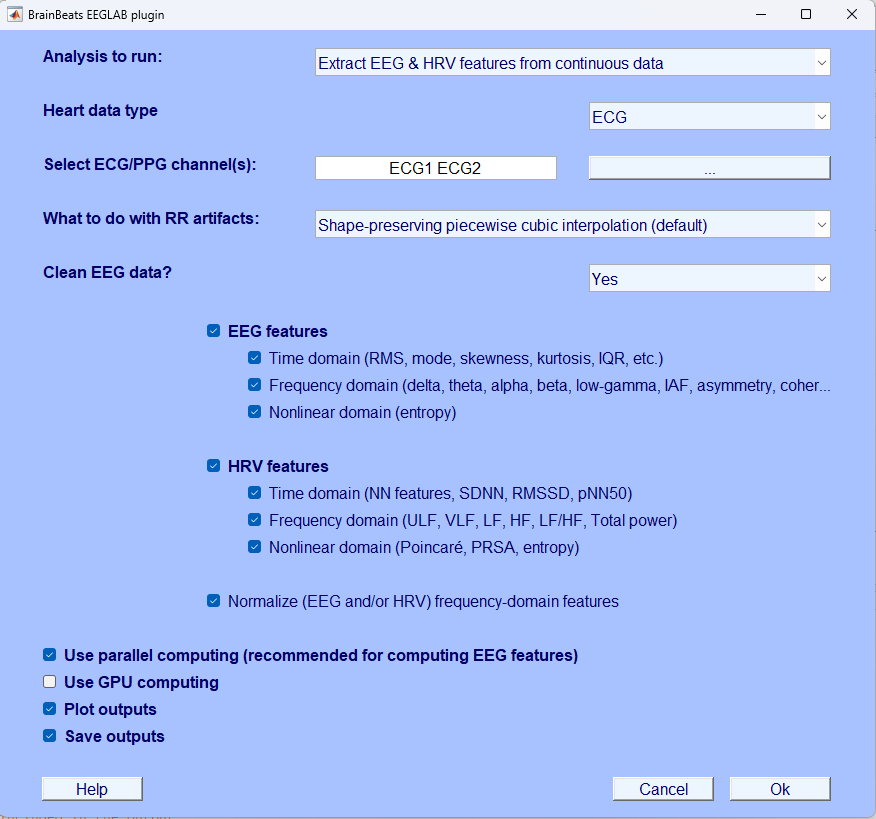


Figure 2.1. BrainBeats GUI to select parameters for extracting HRV and EEG features from continuous data.

Advanced users can perform all steps above with the following command:

eeglab; close;

EEG = pop\_loadset('filename','sample\_data2.set','filepath',fullfile(dataDir, 'sample\_data'));

mainDir = fileparts(which('eegplugin\_BrainBeats.m')); cd(mainDir);

[~, Features] = brainbeats\_process(EEG,'analysis','features','heart\_signal','ECG', 'heart\_channels',{'ECG1' 'ECG2'}, 'clean\_rr','pchip','clean\_eeg',true,'norm',true,...

'eeg\_features', {'time' 'frequency' 'nonlinear'}, ...

'hrv\_features', {'time' 'frequency' 'nonlinear'}, ...

'gpu',false,'parpool',true,'save',true,'vis',true);

1. **Method 3: Remove heart components from EEG signals.**
   1. Load "sample\_data2.set" containing 3 EEG channels and one ECG channel, collected with a "dry" Muse wearable headset.

Open the same BrainBeats GUI (see **Step 1.1**. and **Figure 1.1.**) and set the analysis type as "Remove heart components from EEG signals", and select the ECG channel from the list of channels. Select "No (already processed)" in the "Clean EEG?" field since this sample file was already preprocessed. Set "Plot outputs" and "Save outputs" options (set by default). Click "Ok" to launch.

These steps can be run automatically over many files using the following command lines:

eeglab; close; mainDir = fileparts(which('pop\_BrainBeats.m')); cd(mainDir);

EEG = pop\_loadset('filename','sample\_data2.set','filepath',dataDir);

EEG = brainbeats\_process(EEG,'analysis','rm\_heart','heart\_signal','ECG', ...

'heart\_channels',{'ECG'},'clean\_eeg',false,'save',true,'vis',true);

As for Method 1, users can perform statistical analyses using the EEGLAB STUDY mode. Users can perform basic statistical analyses in EEGLAB (e.g., permutation statistics, FDR-correction), or advanced hierarchical linear modeling with weighted-least square (WLS) optimization, to better account for within and between subject variance, using the LIMO-EEG plugin. Furthermore, the plugin provides advanced corrections for type 1 error (e.g., max likelihood or spatiotemporal cluster corrections).

**REPRESENTATIVE RESULTS:**

**METHOD 1**

BrainBeats first separates the ECG from EEG data to process them separately. EEG data are bandpass filtered using a zero-phase non-causal finite impulse response (FIR) filter to remove low-frequency drifts and high-frequency noise (highpass cutoff frequency = 1 Hz, low-pass cutoff = 45; transition bandwidth = 0.5 Hz). If the data have at least 30 channels and have not been referenced, BrainBeats re-references them to infinity/REST (Yao [2001](https://pubmed.ncbi.nlm.nih.gov/11761077/), [200](https://pubmed.ncbi.nlm.nih.gov/15798293/)5; [2017](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5331115/); [Dong 2017](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5670162/); Candia-Rivera 2021 ). Then, BrainBeats uses the *clean\_rawdata* plugin to automatically detect and remove abnormal EEG channels (flatlinecriterion = 5; ChannelCriterion = .8; LineNoiseCriterion = 5). Removed channels are plotted (**Figure 1.3, in red**) and interpolated using spherical splines (Perrin et al., 1989). CAUTION: These default parameters are implemented for the best performance for this sample dataset. In most cases, we recommend users clean their datasets before launching BrainBeats to tune parameters for their datasets. For example, abnormal channels may not be reliably detected on low-density EEG montages using this default method.

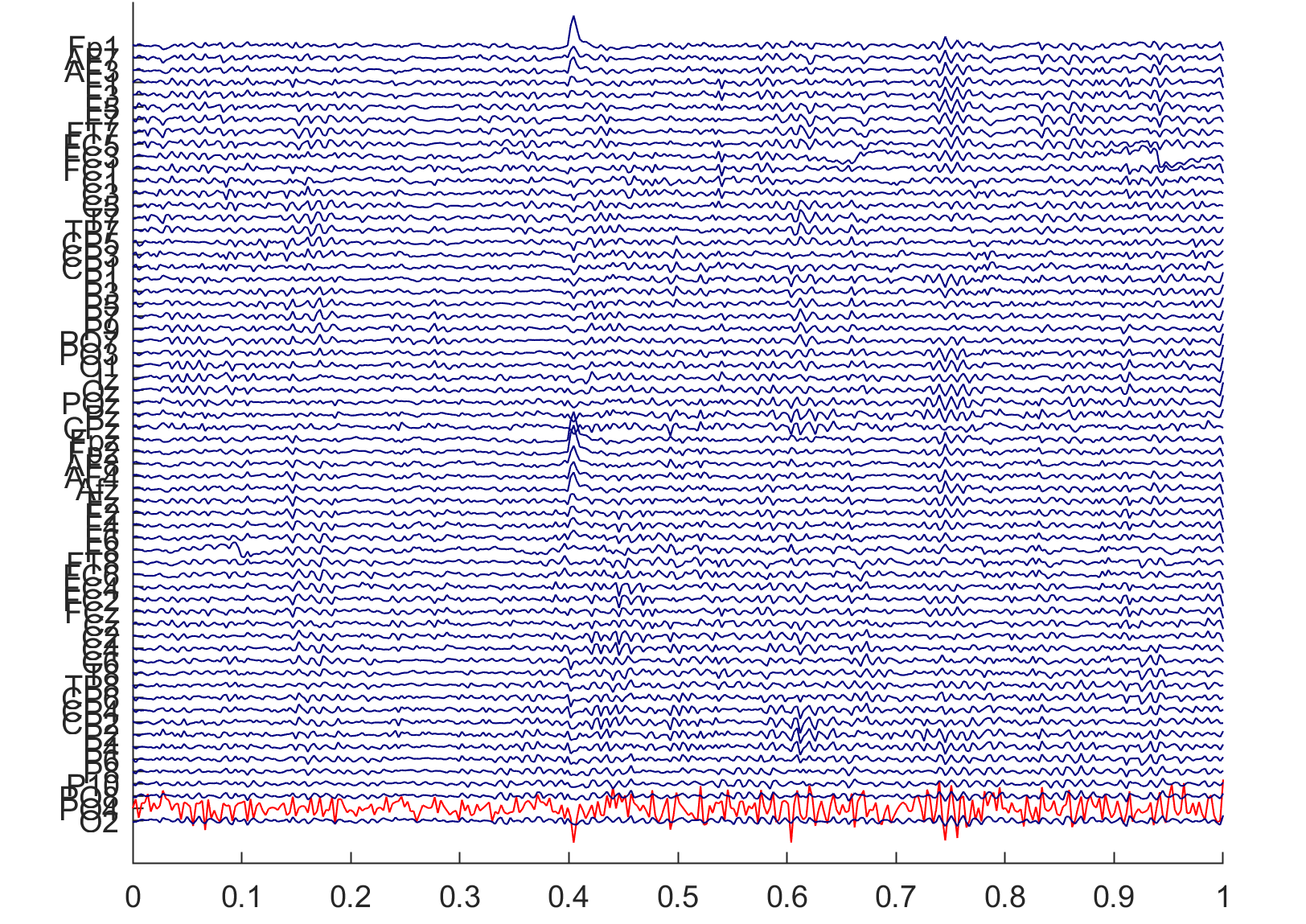


Figure 1.3. Abnormal EEG channels are automatically detected and removed by BrainBeats (in red).

Next, BrainBeats detects R-peaks (from QRS complexes) on non-overlapping windows disregarding initial signal artifacts. The signal undergoes bandpass filtering and Pan–Tompkins (P&T) method, including differentiation, squaring, integration, and smoothing ([ref](https://ieeexplore.ieee.org/document/4122029)). The P&T energy threshold is estimated to avoid disruption from large bumps, and if the RR interval variability exceeds 1.5 times the median, it conducts a "search back" for missed peaks. The R-peak's mean sign over 30 seconds determines the QRS complex sign, ensuring consistent detection. QRS detection and search back are based on an energy threshold defined by the signal's sample rate and smoothed ECG values. Each segment's peak sign is ascertained, and peak points are refined through a refractory period check, also managing flatline conditions. The output consists of RR intervals, heart rate (HR), RR interval timestamps, filtered signal, R-peak indices, peak sign, and the estimated P&T energy threshold. The code (*get\_RR* function) is based on methods developed by Behar et al. (2014); Johnson et al. (2014).

Next, BrainBeats identifies abnormal RR intervals (e.g., too closely or broadly spaced, physiological and non-physiological artifacts), flag them, and interpolates them using the shape-reserving piecewise cubic method (default) to obtain the normal-to-normal (NN) intervals. Spikes within RR intervals are detected using a forward-backward search. Users can choose another interpolation method or remove them if desired. When several ECG channels are present, these steps are performed on each of them, and the electrode with the least number of RR artifacts is selected for the following steps. The filtered ECG signal, identified R-peaks, NN intervals, and interpolated artifacts (from the best electrode) are plotted (see **Figure 1.4.**). Note that users can scroll through the R-peaks more closely using a scroll bar (30-s windows). This code is adapted from the validated algorithms developed for the Physionet Cardiovascular Signal toolbox (Clifford 2002; Vest et al. 2018). See Vest et al. (2018) for validation and performance comparison with other state-of-the-art algorithms.

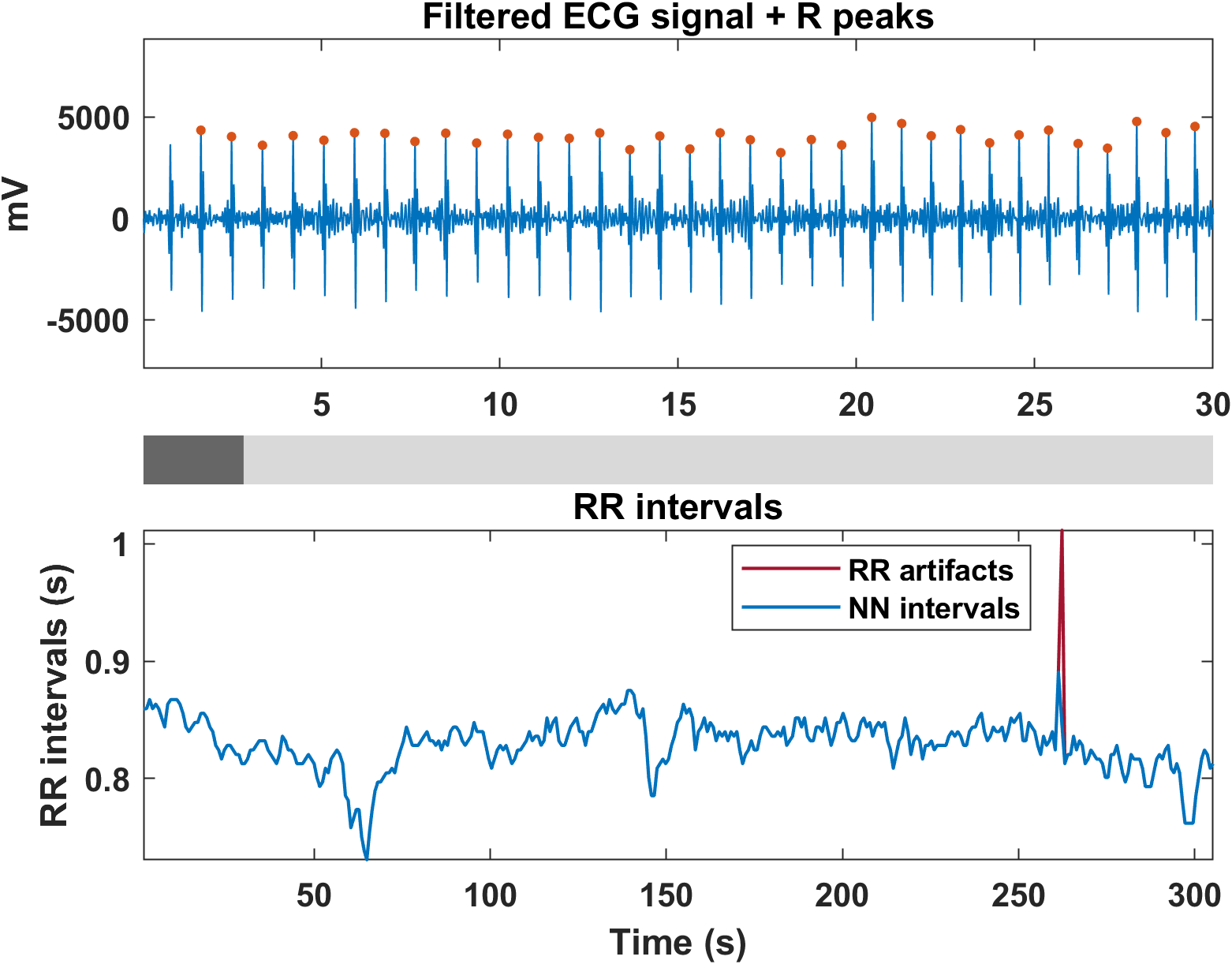


Figure 1.4. Filtered and smoothed ECG signal with identified R-peaks (top panel). Normal-to-normal (NN) intervals with the interpolated RR artifacts (bottom panel).

Next BrainBeats adds R-peak markers to the EEG signals, calculates inter-beat-intervals (IBIs), and removes trials with IBIs less than 550 ms (following recommendations by Candia-Rivera et al., 2021) and Park & Blanke, 2019). The *run\_HEP* function removes outlier trials (detected by MATLAB's *isoutlier* function, *'grubbs'* method) and generates a histogram of the resulting IBIs with fitted normal density (see Figure 1.5.). To determine the minimum epoch size cutoff following R-peak events, the 5th percentile of the IBI data is calculated (i.e., the value below which 5% of the IBI falls, displayed as a dashed red line in the histogram). This epoch size maximizes the number of epochs and limits overlapping epochs.

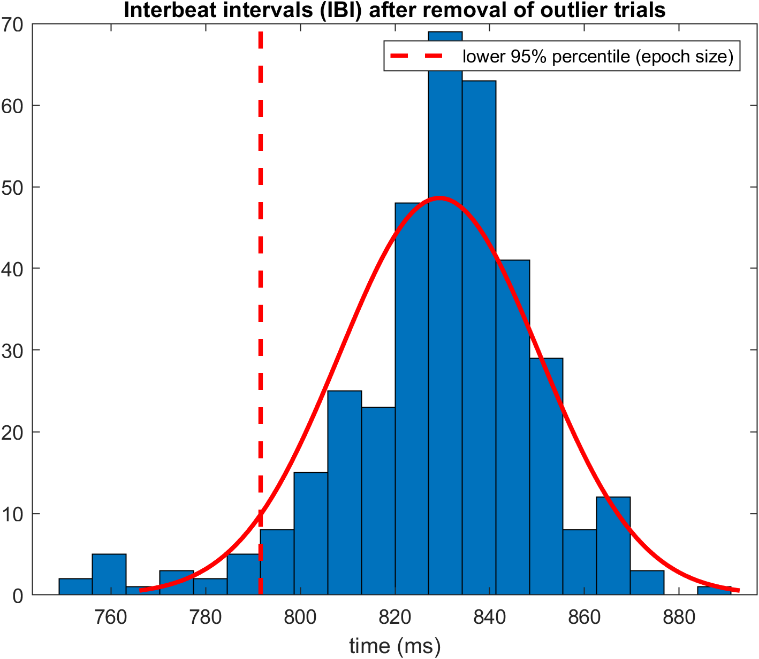


Figure 1.5. Histogram of the interbeat intervals (IBI) with fitted normal density (red line) and the 5% percentile (red dashed line) indicating the minimum cutoff value at which the EEG data are epoched.

BrainBeats then epochs the EEG data from -200 ms to + 5% percentile of IBI and eliminates abnormal trials (by extracting RMS and signal-to-noise ratio features for each epoch and identifying outliers) and eye/muscle components with ICA and ICLabel (at least 90% classification confidence for ocular components and 95% confidence for muscular components; see **Figure 1.6.**). These flagged components are then automatically subtracted from the EEG signals while preserving relevant brain signals. Note: if the Picard plugin is already installed in EEGLAB, it is used by default for fast computation of ICA, otherwise, the Infomax algorithm is used (REF). Effective data rank is calculated prior to running ICA, and PCA-dimension reduction is applied when the data are rank-deficient to avoid ghost IC artifacts; Kim et al., 2023). The final output of processed HEP data is plotted for final inspection (see **Figure 1.7.**) and saved in the same directory as the original file loaded by the user (same name with "\_HEP" at the end). Note: it is recommended to follow the BIDS for better organization, replication, and performing statistics at the group level (EEGLAB STUDY). Users can pause before processing the next file (next condition or participant).

A picture containing colorfulness, aqua, circle, turquoise

Description automatically generated

Figure 1.6. Indpendent component automatically classified and removed from HEP data to extract ocular artifacts without removing relevant brain signals.

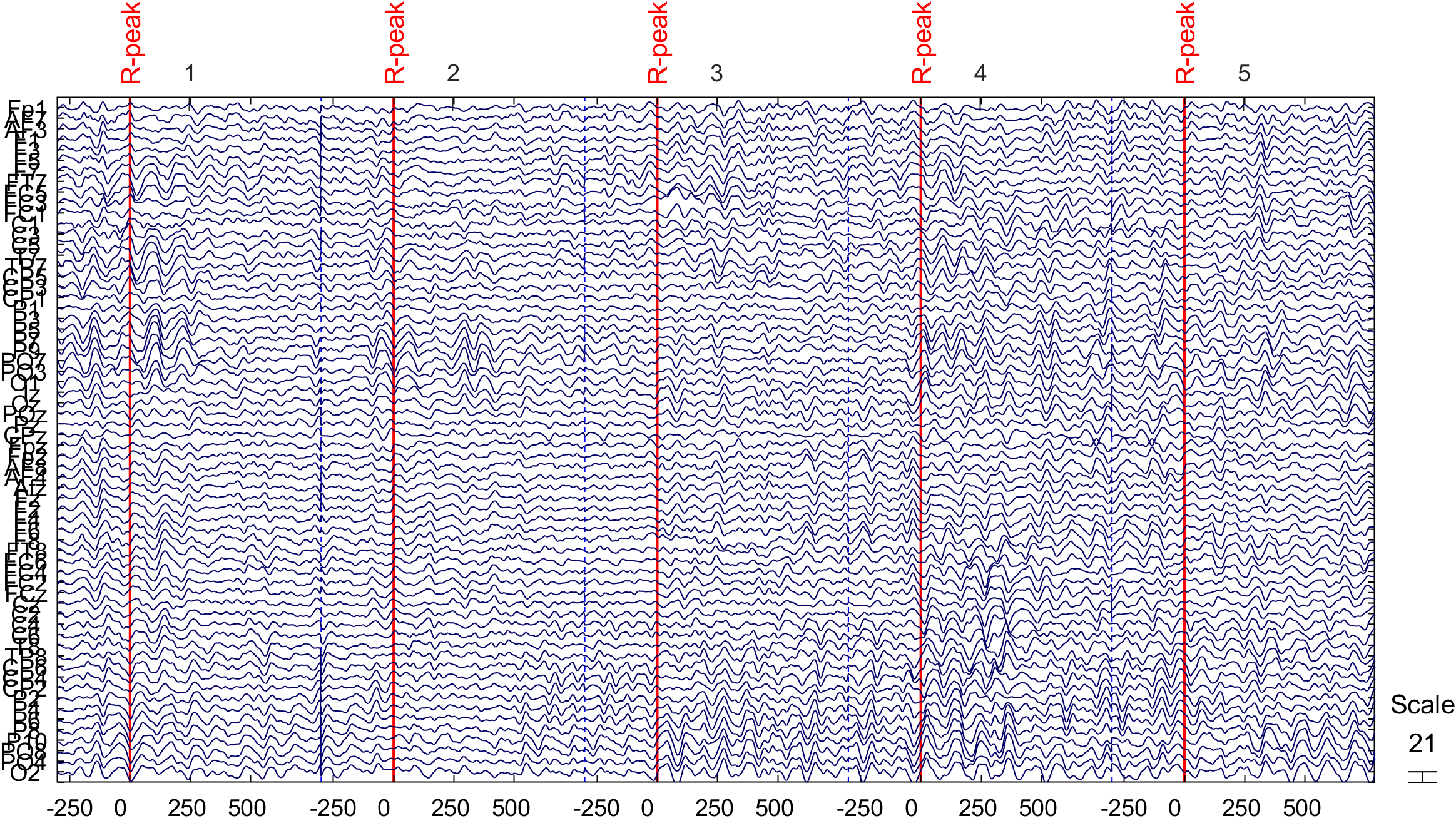


Figure 1.7. Final, cleaned EEG data, ready for HEP analysis.

The HEP averaged across all epochs are plotted as a scalp topography allowing users to inspect each electrode more closely (see Figure 1.8.), superimposed with scalp topography in the region of interest (200-500 ms after heartbeat; **Figure 1.9. top panel**; see Candia-Rivera et al. 2021, and Park & Blanke 2019), and the HEP evolution over time (**Figure 1.9. bottom panel**).

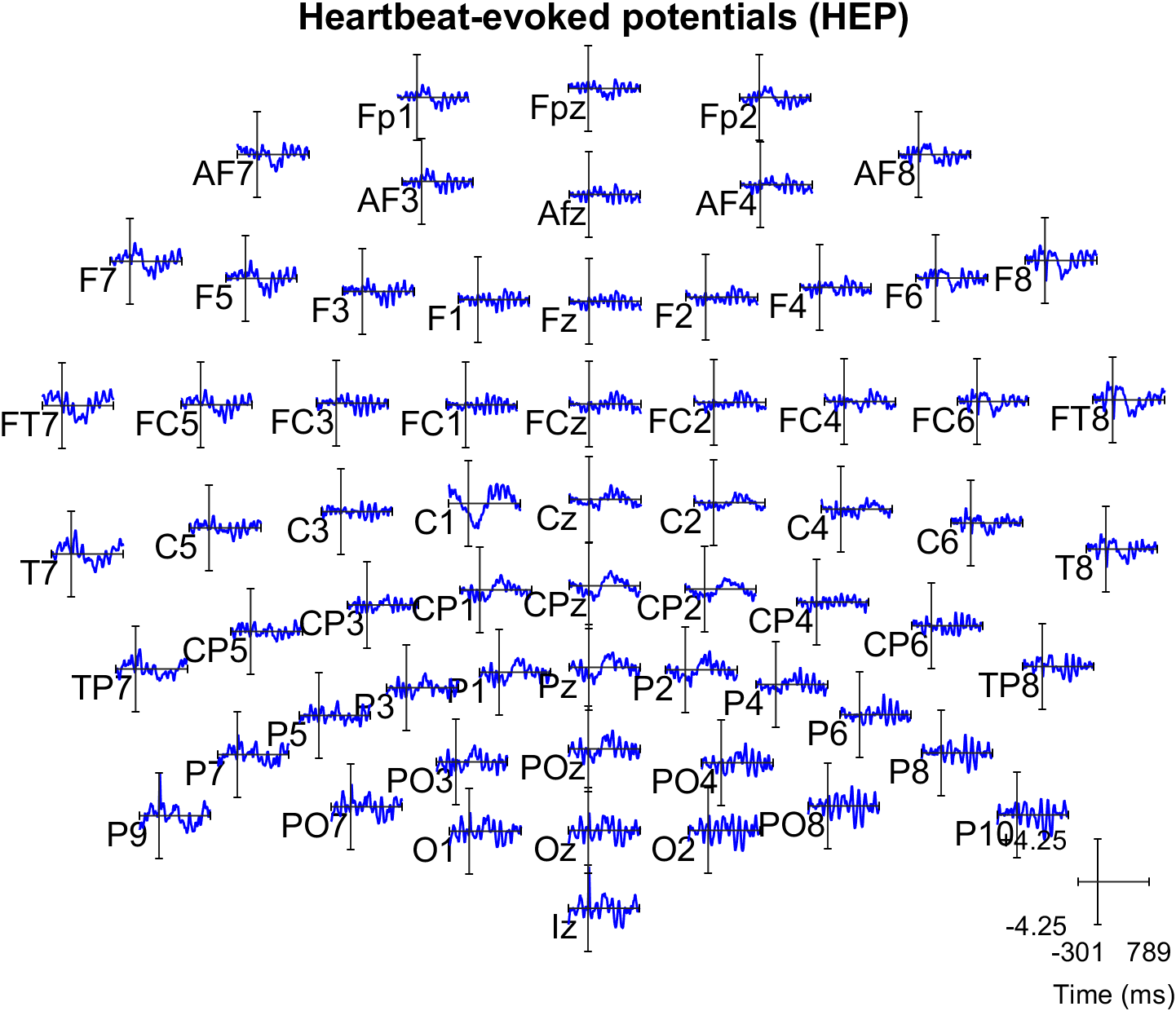


Figure 1.8. HEP averaged across trials for each electrode. Users can click on electrodes of interest for closer inspection.

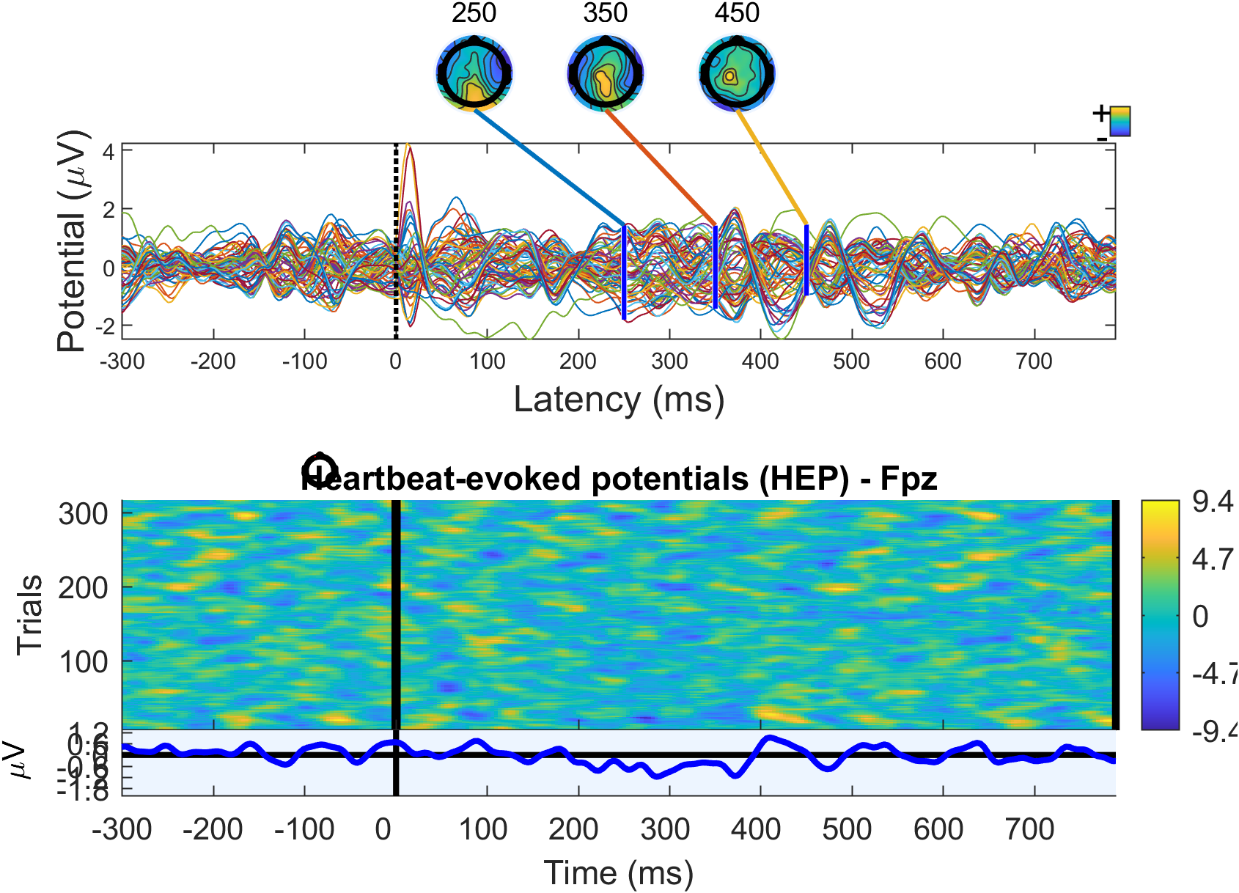
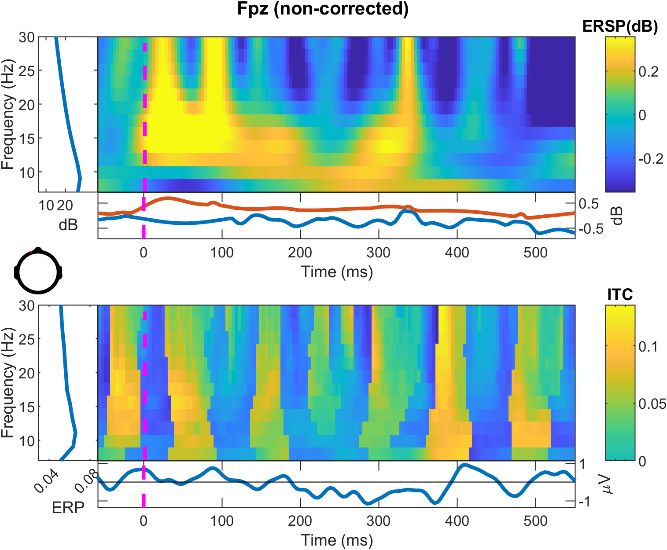
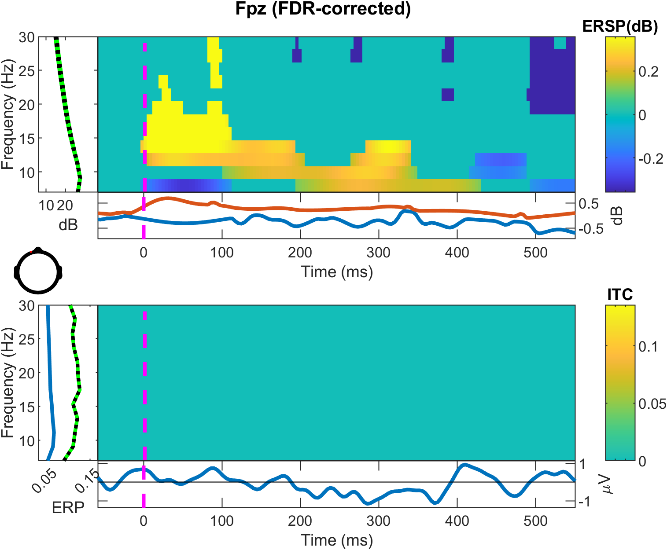


Figure 1.9. Top: HEP averaged across trials, all electrodes superimposed in the time domain, with scalp topographies showing amplitude distribution in the period of interest (200-500 ms after heartbeat). Bottom: HEP evolution over time (each "trial" corresponds to a heartbeat).

The heartbeat evoked oscillations (HEO, or event-related spectral perturbations; ERSP; **Figure 1.10. top panels**) and intertrial coherence (ITC; **Figure 1.10. bottom panels**) are computed for the frontal electrode Fpz using wavelet ([3 0.8] cycles; pre-event baseline removal; pad ratio of 2) for frequencies 7-30 Hz. The same plot is generated after applying permutation statistics and FDR correction to control for multiple comparisons (type 1 error or family-wise error). After correction (**Figure 1.10. Right**), a significant HEO effect is observed on the sample dataset in the beta band during the QRS complex (within 100 ms after R-peak) and in the alpha band during the period of interest (i.e., 200-500 ms, consistent with previous findings; REF). No effect is observed in the ITC data after correction. Note: these plots and results are for replication and illustration only. Low frequency cannot be estimated due to the short epoch size between the heartbeats.

**** 

When all files are processed, users may import them into an EEGLAB STUDY to compute HEP (i.e., ERP) and HEO (i.e., ERSP and ITC) data on the whole group and run statistics at the group level. We recommend perform hierarchical linear modeling using the LIMO plugin (REF). A full tutorial is available at <https://github.com/LIMO-EEG-Toolbox/limo_tools/wiki>

METHOD 2

RR and NN series obtained as in METHOD 1 (The same **Figures 1.3.** and **1.4.** pop-up). However, EEG data are cleaned differently since EEG and ECG signals do not need to be time-locked with ms accuracy to extract features (unlike for HEP/HEO analysis). Here, instead of removing abnormal epochs, BrainBeats first removes segments with large EEG artifacts using artifact subspace reconstruction (ASR; SD criterion = 30; REF; **Figure 2.1.**). Then, Infomax ICA (or the Picard algorithm if installed) is performed with PCA-dimension reduction to account for effective data rank, and ocular and muscular components are classified and subtracted from the signal using ICLabel (as in Method 1, but on continuous data). One ocular component is removed with 98.6% accuracy.

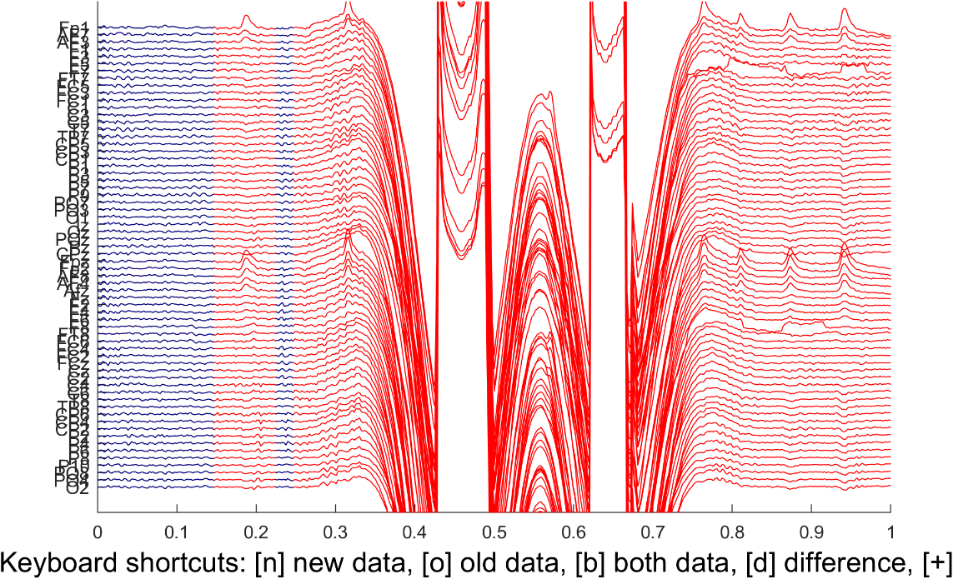


Figure 2.1. EEG artifacts are removed automatically from the continuous EEG data using artifact subspace reconstruction (ASR).

Then, HRV and EEG features are extracted in the time (NN), frequency, and nonlinear domains. See **Table 1** for a list of the features currently available. The Lombscargle-periodogram is the default method to compute frequency HRV as it does not require interpolation or resampling unlike the Welch or FFT methods, better preserving original information. The normalized periodogram is used by default to scale total power to the time series' variance, better dealing with non-uniformly sampled data (common in NN data), and facilitating comparison of results across different participants. When normalization is set in the GUI (default), each band power is divided by the total power to provide a more intuitive measure of the relative contribution of each band on overall power (Shaffer and Ginsberg, 2017). Note: the Welch and FFT methods, and the required resampling step are also available through command line. EEG spectral power is also normalized to decibels (dB) and to total power (same reasons as for HRV power). Additionally, an algorithm is used to detect individualized frequency bounds and compute band power on the individualized bands (see Corcoran et al. 2018).

Table 1. HRV & EEG features estimated with the BrainBeats plugin

|  |  |  |
| --- | --- | --- |
|  | **HRV** | **EEG** |
| **Time** | NN statistics (mean, mode, variance, skewness, kurtosis, interquartile range), SDNN, RMSSD, pNN50. | Amplitude statistics (RMS, variance, skewness, kurtosis, interquartile range) |
| **Frequency** | Spectral power in the ultra-low frequency (ULF), very-low frequency (VLF), low-frequency (LF), and high-frequency (HF) bands, and total power. | Power spectra (all frequencies), mean band power (delta, theta, alpha, beta, low-gamma), individual alpha frequency (IAF), alpha asymmetry (all pairs), EEG coherence (between all non-neighbor pairs) |
| **Nonlinear** | Poincaré (SD1, SD2, SD1/SD2 ratio), fuzzy entropy, multiscale fuzzy entropy, phase-rectified signal averaging (PRSA) | Fuzzy entropy, multiscale fuzzy entropy (all scale factors, mean, SD, peak scale factor, and area under the curve) |

When features are computed, a plot displays the power spectral density (PSD) and multiscale fuzzy entropy (MFE) estimated on the NN series (**Figure 2.2., left**), and on the EEG data (the average across all electrodes is used for illustration; **Figure 2.2., Right**). For EEG MFE, each scale factor is bandpass filtered by default, following guidelines to reduce spectral bias in fine time scales (presumed to indicate fast dynamics) by broadband spectral power (dominated by low-frequency contributions, see Kosciessa et al. 2020 and Grandy et al. 2016 for more detail). As a result, the frequency bounds of each scale factors are provided. **Note**: EEG entropy measures are very computation-heavy. Thus, 20 scale factors are set by default, and when EEG signals are longer than 5,000 samples, they are resampled (or decimated when the factor is not an integer) to 90 Hz (i.e., corresponding to a Nyquist frequency of 45 Hz, to match our default low-pass filter). Furthermore, parallel and/or GPU computing to further accelerate the process. Original code for entropy was developed by Azami and Escudero (2016). Next, the EEG scalp topographies are displayed for each frequency band (**Figure 2.3. top and middle rows**), as well as for the individual alpha frequency (IAF; **Figure 2.3. bottom left**) and fuzzy entropy (**Figure 2.3. bottom right**).

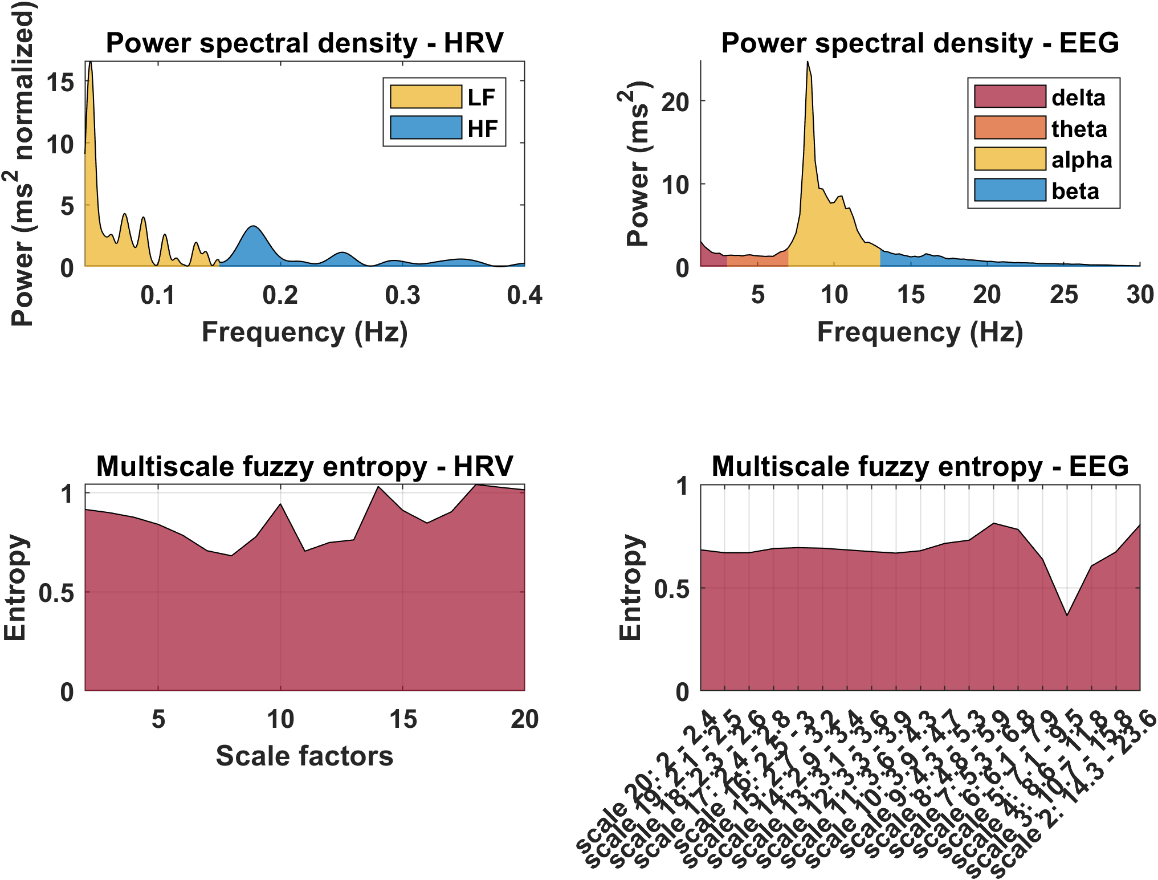
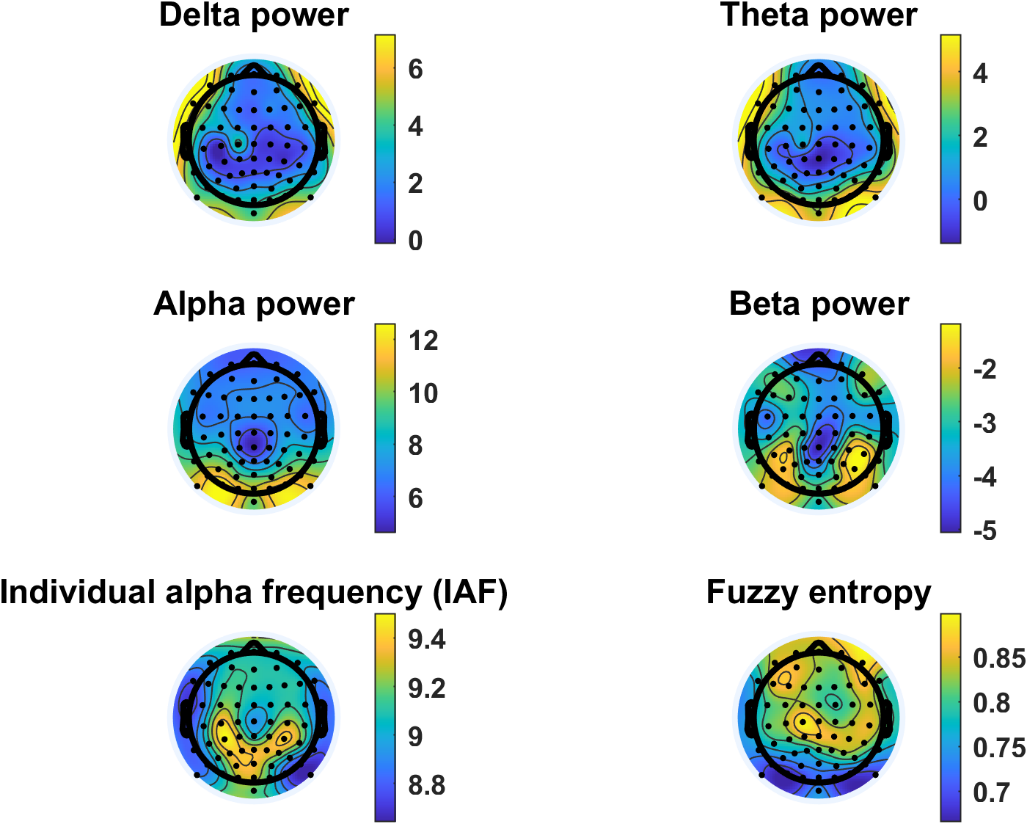


Figure 2.2. Power spectral density (PSD; top panel) and multiscale fuzzy entropy (MFE' bottom panels) features estimated from NN intervals (left panel) and EEG data (right panel). Note: calculated on the *sample\_data1.set*, corresponding to 5 min of mindwandering eyes closed.



**Figure 2.3.** EEG scalp topographies showing the distribution of mean power spectral density (PSD) for the delta (top left), theta (top right), alpha (middle left), and beta (middle right) frequency bands. The individual alpha frequency (IAF) is plotted in the bottom left corner, and the fuzzy entropy distribution in the bottom right corner. Note: from *sample\_data1.set*, corresponding of 5 min of mindwandering eyes closed.

**METHOD 3**

BrainBeats runs Infomax ICA (or Picard algorithm if installed) using PCA dimension-reduction to account for effective data rank ([Ref](https://www.frontiersin.org/articles/10.3389/frsip.2023.1064138/full?&utm_source=Email_to_authors_&utm_medium=Email&utm_content=T1_11.5e1_author&utm_campaign=Email_publication&field=&journalName=Frontiers_in_Signal_Processing&id=1064138)). Next, the heart component(s) is automatically classified and removed with 95.3% confidence, using the ICLabel plugin. The ECG channel(s) is included in this process to increase ICA's source separation performance and chances to separate heart components from the EEG signals. Users can click on the component to inspect its properties in more detail (see **Figure 3.1. left**). The final output is displayed, showing the heart components removed from the EEG signals in red (**Figure 3.1. right**). Notes: The ECG signal is included here only for visualization purposes, but is otherwise removed from the output dataset. When no heart components are detected, users are informed in MATLAB's command window, and the program ends.

Graphical user interface, application

Description automatically generated A picture containing text, font, handwriting, line

Description automatically generated

**Figure 3.1.** **Left**: Heart independent component classified by ICLabel and its properties. **Right**: EEG signals after removing the heart artifacts (in red). The ECG signal is included here for visualization purposes, but is otherwise removed from the final output dataset.

**DISCUSSION:** *(3-6 paragraphs)*

**Critical steps in the protocol**

Critical steps prior to launching Brainbeats include: having MATLAB and EELAB installed, installing the BrainBeats plugin, installing a few EEGLAB plugins (for Methods 1 and 3), possessing a dataset that includes both EEG and cardiovascular (ECG or PPG) signals and importing with appropriate EEGLAB plugin, knowing the name of the cardiovascular channel(s) to select from the list, selecting initial parameters in the main GUI. Note that, data importation cannot be automated as various plugins are required to be installed to account for the many different EEG data formats that exist (e.g., .edf, .bdf, .vhdr, etc.).

Warnings and error messages are implemented at various places in the toolbox to help users understand why they may encounter issues (e.g., file length being too short for calculating a reliable measure of ultra-low frequency HRV, signal quality being too low for any reliable analysis, missing electrode locations etc.).

Note that for advanced users, each function contains a brief description of what it does and what the parameters are (and default, recommended parameters), which can be displayed in the command window by calling "help function\_name".

**Limitations of the method**

Entropy features are particularly promising for capturing complex, bidirectional interactions between cardiovascular, subcortical, and cortical systems that may be hidden in nonlinear feedback loop dynamics (Costa et al., 2005; Klonowski 2009). However, they are very computationally heavy and can take a very long time to compute on EEG signals. While improvements were made (parallel computing, GPU computing, signal downsampling/decimation), further accelerations are required, if possible. While statistics are provided for the Method 1 and 3 (at the subject and group levels using the EEGLAB STUDY mode), no statistical analysis is currently provided for Method 2, to assess interactions between HRV and EEG features. This is because a lot of flexibility must be implemented to account for the various analyses users may want to perform (e.g., statistics, machine learning). However, users may use the Method 2 to extract features and easily perform statistics with any standard statistical software.

Users cannot import EEG and ECG/PPG data from separate files yet (although this feature will be available soon).

**Significance of the method with respect to existing methods**

Overall, BrainBeats provides state-of-the-art signal processing techniques and statistical methods for both EEG and cardiovascular signals, that are superior to currently available tools. Method 1 can be performed using advanced hierarchical linear modeling and statistical testing, as well as analyzing brain-evoked response by heartbeats using both ERP, ERSP, and ITC measures, as well as channel (scalp signals from each electrode) and independent component (signals that are source-separated with ICA) domains. Method 2 does not have a pre-existing alternative currently. Method 3 allows quick and easy removal of heart components from the EEG signals, in a fully automated manner. While this already possible in EEGLAB, it requires users to perform a series of steps and choice of parameters (e.g., highpass filtering the signals, running ICA, running ICLabel, tuning parameters, subtracting the heart components from the EEG signals, and removing the ECG channels) that can easily lead to errors (e.g., ghost ICs; Kim et al 2023).

Additionally, the toolbox implements computing performance improvements to accelerate the estimation of EEG features (mainly multiscale entropy measures), including both GPU and parallel computing. Note that these options are only as beneficial as the users' hardware (i.e., graphic card and number of processors and threads).

**Future directions**

The toolbox will continue to be modified and improved by the authors in the long term, to implement the latest guidelines and recommendations by experts in the field, and fixing any errors that may arise.

Currently, users can extract features of interest and perform statistical analyses using their own methods (correlations, regressions, machine learning). However, BrainBeats will provide the following statistical methods for feature analysis soon:

1. Feature selection for HRV and EEG features using random forest (separately), to remove redundant and correlated features and reduce multiple comparisons.
2. Skipped correlations (Pearson or Spearman depending on data distribution and variance) will be performed across HRV and EEG features (Pernet XXX), with FDR-correction for multiple comparisons to control for type 1 error.
3. Next, since correlation coefficients are non-normally distributed, they will be transformed using the Fisher r-to-z method to meet the normal distribution assumption of t-tests. Nonparametric tests (e.g., Wilcoxon signed-rank test and Wilcoxon rank-sum test) test for differences in central tendency and are not well-suited for differences in correlation coefficients, unlike t-tests. However, they will be used if N < 30.
4. If distributions are still significantly skewed, Yuen t-tests will be applied (using 20% trimmed means) to address the issue. If conditions/groups have different variance, a Welch t-test will be used.
5. If users do not wish to compare conditions or groups, but investigate associations between features and a third variable, or predict a response variable (e.g., age, sex, diagnosis, etc.), they will be able to do so using correlation matrices, simple/multiple linear regressions. Corrections for multiple comparisons will be implemented.

For HEP/HEO analysis, the short interbeat intervals (IBI) that occur naturally (~600-1000 ms) lead to short EEG epochs. While this is not an issue for HEP analysis in the time domain, it presents limitations for HEO analysis. Time-frequency decomposition requires data to extend up to 3 cycles in the lowest frequency of interest beyond the window of interest. For HEO, the window of interest being 200-500 ms, one would require for example an additional 600 ms before and after thw window (i.e., -400 to 900 ms) for examining frequencies as low as 5 Hz. This is strongly recommended for obtaining correct time and frequency resolution, and for avoiding edge effects. Thus, one would end up rejecting most trials since they would not have the required length (especially if one wanted to examine frequencies as low as 1 Hz an additional 3 s before and after window of interest). Hence, future developments will implement the ”reflection” method, which mirrors the signal from the window of interest (i.e. backward version of the signal) before and after to expand the available window. This provides smooth transitions and moves the edge effects further away. The mirrored sections are then removed.

Finally, other features and methods will be added to assess interactions between EEG and cardiovascular signals. New methods will include for example EEG-ECG coherence, the "time-resolved directional brain/heart interplay measurement" (Catrambone et al. 2019), or classification of HEP or feature data using machine learning (e.g., decision trees, random forest, naïve bayes, SVM, KNN, long short-term memory networks, etc.).

**ACKNOWLEDGMENTS:**

Our research was supported in part by the Institute of Noetic Sciences.

**DISCLOSURES:**

The authors have nothing to disclose.

**REFERENCES:**

In-Text Formatting:

* The corresponding reference numbers should appear as superscripts after the appropriate statement(s) in the text (before punctuation but after closed parenthesis). The references should be numbered in order of appearance.
* Multiple references should be separated by commas, or a dash for inclusive numbers: example2,5 refers to references 2 and 5, while example2-5 refers to references 2 through 5.
* Personal communications, unpublished data, and conference abstracts can be cited parenthetically in the text with author last name, initials, and year.
* Footnotes should not be used. Grant details and personal acknowledgments should not be cited as a numbered reference (but included in the Acknowledgments section).

*Citation Formatting:* (in order)

* Last name, first and middle initials (if available). List ALL authors. If there are six or more authors, list the first author and then "*et al."*.
* Include article, book, or chapter titles. Titles of books should be italicized, whereas article and chapter titles should not.
* Article titles should start with capital letters and end with periods, and should appear exactly as they were published in the original work with no abbreviations or truncations.
* Write journal names in italics. The city or country where a journal is located can be included in parenthesis with the journal name. For books or other works, a publisher name, city, and country can be included.
* Write volume numbers in bold, followed by a space, issue number (in parentheses), a comma, and then a range of page numbers (start page – last page). A single page number or digital object identifier [doi] can be substituted for a range of page numbers.
* Provide the year published (in parentheses), followed by a period.
* "Accepted" or "In Press" can be listed after the title or journal name and before the year. Manuscripts that are in preparation or under review should not be listed.

*Citation Examples:*

Bedford, C.D., Harris, R.N., Howd, R.A., Goff, D.A., Koolpe, G.A. Quaternary salts of 2-[(hydroxyimino)methyl]imidazole. *Journal of Medicinal Chemistry*. **32** (2), 493-503 (1998).

Kioh, L.G. *et al.* *Physical Treatment in Psychiatry*. Blackwell Scientific Pubs. Boston (1988).