Integrating Data Across Oscillatory Power Bands Predicts the Epileptogenic Zone: the Frequency Range Explorer Epileptogenic Zone (FREEZ) Identification Algorithm

Authors: Sean O'Leary[1,†], Anne-Cecile Lesage[1,†], Liliana Camarillo-Rodriguez[1], Oliver Zhou[1], Diosely Silveira[2], Jiefei Wang[3], Sameer A. Sheth[4], Michael S. Beauchamp[5], Zhengjia Wang[5], John F. Magnotti[5], Patrick J. Karas[1,6]

Abstract

Drug-resistant focal epilepsy affects over 25 million people globally. Identification and removal of the seizure onset site, termed the epileptogenic zone (EZ), is the most successful treatment to stop seizures in these people. Implanting electrodes into the brain to record intracranial electroencephalography (iEEG) is the gold standard test for identifying the EZ. But identification of the EZ with iEEG remains challenging in many cases. We developed a novel methodology using a stacked random forest ensemble machine learning model integrating power across delta, theta, alpha, beta, gamma, and high gamma frequencies over time to identify the EZ. In patients who were seizure-free after surgery, electrodes within the EZ showed significantly higher area under the curve for mean power over time in the first 20 seconds after a seizure compared to electrodes outside the EZ in the alpha ($p = 0.0272$), beta ($p = 0.0263$), gamma ($p = 0.0013$), and high gamma ($p = 0.0086$) ranges. Additionally, electrodes within the EZ in patients that became seizure-free after surgery had significantly higher AUC compared to electrodes marked within the EZ in patients who did not become seizure-free after surgery in the gamma ($p = 0.0145$) and high gamma ($p = 0.0024$) power ranges, but not delta ($p=0.8678$), theta ($p=0.4956$), alpha ($p=0.1320$) or beta ($p=0.1624$) ranges. A machine learning model was trained to predict electrodes within the EZ. Leave-one-out patient cross validation of the machine learning model yielded a 95.7% positive predictive value and 99.1% specificity for identifying electrodes within the epileptogenic zone, and 95.2% accuracy for predicting seizure outcome based on a planned resection. We implemented this algorithm into the open-source software tool "Reproducible Analysis and Visualization of iEEG" (RAVE) to enable users to reproduce our results and implement this methodology with new datasets, creating a software module titled FREEZ. The software facilitates quantification of the spectral power changes during seizures, including displaying time-frequency spectrograms and projecting results across patient-specific 3D brain maps. Users can also adjust parameters for visualizing multiple frequency ranges from various time regions around seizure onsets in a web-browser-based interface.

Affiliations:
[1]Department of Neurosurgery, University of Texas Medical Branch, Galveston, TX, USA
[2]Department of Neurology, University of Texas Medical Branch, Galveston, TX, USA
[3]Department of Biostatistics & Data Science, University of Texas Medical Branch, Galveston, TX, USA

[4]Department of Neurosurgery, Baylor College of Medicine, Houston, TX, USA
[5]Department of Neurosurgery, University of Pennsylvania, Philadelphia, PA, USA
[6]Department of Neurosurgery, University of Texas MD Anderson Cancer Center, Houston, TX, USA

[†]These authors contributed equally to this work

Corresponding Author's name and current institution:

Patrick Karas
301 University Blvd, JSA 9.300
Galveston, TX 77555-0517
pjkaras@utmb.edu

Running Title: Frequency Range Explorer Epileptogenic Zone (FREEZ) Identification Algorithm

Key Words: drug-resistant epilepsy, epileptogenic zone localization, iEEG power analysis, multi-taper, FREEZ, RAVE

Introduction

Drug-resistant epilepsy affects approximately one third of epilepsy patients and is associated with increased mortality rates (1–3). Surgery is the most effective treatment for these patients, offering potential seizure freedom rates of 58% to 78% following resection of the epileptogenic zone (EZ) (4). Historically, the EZ has been defined as the site of origin and primary organization of a seizure (5,6). When non-invasive diagnostic exams do not conclusively identify the EZ, intracranial electrodes (iEEG) are implanted and the EZ is identified by an epileptologist interpreting anatamo-clinical-electrical correlations of the iEEG recordings (7). Epileptiform discharges are visually identified from iEEG recordings and correlated with their anatomic location and the patient's clinical features. Surgical outcomes (seizure freedom versus continuation of seizures) indicate whether the EZ was correctly identified and removed, serving as a post-hoc definition of this region. Using surgical outcomes as validation, various computational epileptogenic zone identification algorithms have been developed to help clinicians identify the EZ. Computational algorithms seek to distill a subset of features from iEEG recordings to identify electrical EZ biomarkers.

A majority of computational algorithms are based on analysis of iEEG power spectra at seizure onset and early in the seizure time course. These methods rely on different frequency patterns that have been observed clinically (8,9). For example, high power in the beta frequency range has been found to correlate to the EZ (10,11). In another method, Bartolomei *et al.* (12) developed the *epileptogenic index* which assesses the timing of changes in the ratio of energy between fast (beta and gamma range) and slow (alpha and theta range) power spectra. Grinenko *et al.* (13) correlate the association of pre-ictal broadband spiking, early narrow-band fast activity and broad low frequency suppression to define a spectral EZ *fingerprint*. Izumi *et al.* (14) identified that the cooccurrence of ictal direct current shift (infraslow power below delta band)

and high frequency oscillations was a reliable marker of the EZ. These studies suggest a complex interplay of neural activity across many frequency bands during the early ictal period. However, a more unified method to consider all frequency patterns correlating with EZ localization is missing.

To tackle this issue, we ran a spectral spatiotemporal analysis across multiple frequency ranges (delta, alpha, theta, beta, gamma, high-gamma). While spectral power analysis has been performed using limited time-frequency patterns to benchmark performance of other computational EZ algorithms (10), we went further by unifying frequency range results with an ensemble learning model designed to combine weak predictors into a single more robust model (15). Multiple random forest models were trained on the correlation between the EZ and spatiotemporal spectral power in each frequency range prior to inclusion in an ensemble model. Ensemble modeling is a well-known methodology designed to combine diverse base models to decrease prediction error. Stacking, also referred to as super learning (16), is a type of ensemble approach consisting of training a second level "meta-learner" to obtain a performant combination of the base learners through the computation of optimal weights.

A critical hurdle facing adoption of computational EZ localization algorithms is lack of validation, largely stemming from the difficulty of carefully critiquing and reproducing the methods. To address this hurdle and make our algorithm easily available, we also created an opensource software package titled "Frequency Range Explorer of the Epileptogenic Zone" (FREEZ) (17) available for free download: https://zenodo.org/records/11440403. FREEZ is coded in R and R Shiny, enabling graphical user interface via any web browser, and created using the **R**eproducible **A**nalysis and **V**isualization of intracranial **E**EG (RAVE) (18) platform. Combined with powerful preprocessing, electrode localization, and data visualization tools from RAVE, the FREEZ software enables spatiotemporal visualization of the mean relative power by across electrodes overlaid on patient brain models.

FREEZ implements our machine learning models for both *electrode classification* and *post-operative seizure freedom prediction* based on changes in the spectral power around the seizure onset zone. FREEZ also allows for multi-patient analysis through script-based pipelining to generate results and provide statistical analysis and prediction using ensemble machine learning (15) stacking models built on multi-frequency range analysis.

Material and Methods

Patient Data and Selection

The Fragility multi-center dataset (10), available in part publicly on OpenNeuro (19), encompasses data from 35 patients distributed across four surgical epilepsy centers. The iEEG data is in the iEEG-BIDS format (20) with signal data is in BrainVision Core Data Format (Brain Products GmbH, Gilching, Germany). Clinicians at each clinical facility labeled electrodes within the EZ. Electrode names and electrode classification were available, but electrode

localization was not provided for all patients. For six patients of the patients we were able to retrieve patient-specific MRI imaging and electrode localization from an overlapping dataset (21) (subjects PT01, PT2, PT3, PT6, PT14, PT15). The dataset features an event file with markers annotated by clinicians to indicate the timing of seizure onset/offset clinically and electrographically. These seizures are marked according to standard procedures within the clinical epilepsy workflow. Additional details about the dataset are available (10).

The following inclusion criteria were used: (1) patient underwent resection or laser ablation surgery guided by intracranial EEG (ECoG or sEEG); (2) had at least one seizure recorded during intracranial monitoring, allowing for confirmatory testing of selected channels; (3) their recordings were at frequencies of 500 Hz or higher, suitable for spectral frequency analysis up to 150 Hz; and (4) detailed surgical outcomes, including Engel scores, were available. As in the study by Li *et al.* in which this data was originally published (10), we assume for the validation of our EZ localization tools that electrodes annotated SOZ electrodes are inside the resected area for all patients and that they are a good estimate of the EZ in case of seizure freedom outcome.

iEEG data preprocessing in RAVE

Raw signal was preprocessed in RAVE 2.0 (18). Data was epoched for seizure onset time and baseline signal. Electrodes with artifacts were removed from the analysis. Data was further notch filtered and common average referenced. These steps were performed using the RAVE 2.0 graphical user interface. Additional details about iEEG data preprocessing in RAVE 2.0 can be found in the RAVE 2.0 tutorial (https://openwetware.org/wiki/RAVE:ravepreprocess). Electrode localization was performed in the original dataset (21).

Spectral spatiotemporal power analysis

For each electrode, we computed the mean spectral power for six frequency bands from the raw signal: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), gamma (30-90 Hz), and high gamma (90–200 Hz) (Fig. 1A and 1B). Calculations were performed using the multi-taper Fourier transform (22) method and computed using the multi-taper R library (23) (step size 0.5s, window length 2.5s). These computations were performed across two epochs, the ictal interval (first 20 seconds after seizure onset) and the baseline period (10 second time interval starting 30 before seizure onset). Additional details about the multi-taper method and specific parameters are available in the *Supplementary Methods*. Results of the mean spectral power were then displayed overlayed on the reconstructed electrode implantation (Fig. 1C). This was compared to the clinician labels (Fig. 1D).

Statistical analysis of correlation between spectral power and EZ

To analyze power change at seizure onset, mean spectral power output from multi-taper transform was exported for each electrode in .csv format for each of the six frequency ranges. The resulting time-frequency matrices were grouped based on electrode labels as EZ *vs.* non-EZ. For each timepoint within the analysis window, average power and standard error were

computed for grouped EZ electrodes and grouped non-EZ electrodes at every frequency range. Electrode labels (EZ vs. non-EZ) were correlated to surgical outcome (seizure free vs. not seizure free) by calculating area under the curve (AUC) for these averaged power timeseries at every frequency range (Fig. 1C). To evaluate significant differences in AUC for each category, pairwise two-sample t-tests were conducted. A Bonferroni correction was applied to all p-values to adjust for multiple comparisons (24 comparisons).

We computed the distribution of power across frequency bands in EZ *vs.* non EZ electrodes (using clinician labels) to predict post-operative seizure freedom outcomes (10). For each frequency band, we ranked the electrode power into deciles from 10-100 over the analysis time window for EZ and non-EZ electrode populations separately. Decile plots were created from this analysis (Fig. 1F).

Ensemble machine learning method for multi-frequency analysis

Resulting power across all frequency bands was used to train two separate ensemble machine learning algorithms. One algorithm was trained to identify electrodes as within or not within the EZ (*electrode classification*). A second algorithm was trained to predict post-operative seizure freedom based on removal of preselected electrodes (*patient outcome prediction*).

ELECTRODE CLASSIFICATION

To determine the spectral features that could distinguish between electrodes labeled as EZ *vs.* not EZ we used a stacked random forest ensemble (SRFE). This method builds a model to identify various interactions between spectral patterns present in the training data that correlate with the EZ.

To create the SRFE, we first trained six unique random forest models, one for each frequency band. Individual frequency band models were trained with 100 decision trees using the seizure-free patient cohort with clinician-determined EZ electrode labeling as ground truth. We then used the R H2O package (24) to compute the optimal coefficients (second-learning) to combine the six random forest models into the SRFE. The H2O R Library was chosen for ensemble learning training as it allows for packaged implementation, testing, and scaling of machine learning pipelines with a wide variety of models integrated together. The hyperparameters for the SRFE included the selection of a Generalized Linear Model (GLM) as the metalearner algorithm (26,27). Alpha is the regularization distribution parameter, we used the default parameter alpha equal to 0 (Ridge regression) to prevent overfitting and enhance generalization (26,27). Lambda is the regularization strength parameter, we enabled lambda search to find the optimal parameter (26,27). The SRFE model was then tested using leave-one-out cross validation. To convert prediction probabilities to binary EZ labels (EZ *vs.* non-EZ) we tuned the model to prioritize specificity over sensitivity to avoid false positive EZ labeling (reduce the potential for unnecessary tissue resection). This tuning was done by thresholding the model prediction probabilities to maximize true positives while maintaining positive predictive value (PPV) ≥ 0.95. Majority voting ($\geq 50\%$) across different seizures for each patient was then used to select final electrode EZ labeling (13).

We additionally compared the SRFE method with two other machine learning methods using variations of random forest models. In one model, we trained a “simple random forest” on features consisting of all unique analysis time windows for all six frequency ranges. Second, we trained a unique random forest on each frequency range, as performed in the SRFE, but combined them by averaging their prediction results (probability that an electrode is in the EZ) without using a SRFE model “averaged random forest”. Both random forest models were tuned similarly to the tuning of the SRFE model described above. Each model’s performance (simple random forest, averaged random forest, and SRFE) was evaluated with the following metrics: accuracy, specificity, sensitivity, PPV, and negative predictive value (NPV) to compare the predicted outcome with the ground truth electrode labels. For testing of all models, leave-one-out patient cross validation was employed, with majority voting (≥ 50%) across each seizure to select final predictions.

PATIENT OUTCOME PREDICTION

Using the deciles calculated across frequency bands for EZ vs. non-EZ electrodes (Fig. 1F), we first normalized these decile values in each time window between 1 and 10. Normalization was done for EZ and non-EZ electrode populations separately within each time window. The ratio of EZ over non-EZ power in each respective decile at every time window was then calculated for input into a surgical outcome SRFE model. These ratios were first used to train six random forest models consisting of 100 decision trees, one for each frequency band. Following this, an SRFE model was constructed incorporating the six-frequency range random forest models. Again, SRFE hyperparameters included a GLM as the metalearner algorithm, now with the lambda parameter set to 0.1 and the alpha value set to 0.5 (specifying a combination of LASSO and Ridge regression), to prevent over-regularization of the already normalized features (26,27). The SRFE model was tuned by selecting a threshold to maximize model accuracy. For all cross-validation procedures, a leave-one-out patient cross-validation method was used, along with majority voting (≥ 50%) across multiple seizures for each patient to finalize predictions.

FREEZ module: a toolbox for visualization of correlation analysis between spectral features and EZ

The different processes described in the previous subsections have been integrated and are freely available through the FREEZ module GUI (Frequency Range Explorer Epileptogenic Zone) for visualization after processing and reproducibility of this study results. FREEZ is integrated as a RAVE module written in R programming language, HTML (hypertext markup language), CSS (cascading style sheets), JavaScript, and shell (28). RAVE integration allows users to run on laptops, desktops, or in the cloud, with seamless integration between platforms as all module interaction is performed through a web browser (28). Because all computing in RAVE can be performed on any server or user machine, interaction with the module does not require independent computation (28). iEEG data set have to be preprocessed in RAVE 2.0 as described previously for further processing in FREEZ. Figure 2 shows the workflow of the FREEZ module. Please refer to supplementary methods for more details about the FREEZ module GUI. Finally, FREEZ allows to pipeline the whole workflow (Fig. 2) through a script in

R. The script uses the advanced pipelining functions of RAVE 2.0. Please refer to *Supplementary Methods* for detailed module description and pseudo-code.

Code and data Availability

Users can explore the FREEZ module GUI and results on patient PT01 through a web app. They can also download a FREEZ module docker (17) with which they can run FREEZ on PT01 like in the web app but also on their patient data after preprocessing. The source repository for the module and the multi-patient pipelining script are accessible on githubFREEZ with installation guidelines for the Windows, MacOS and Linux platforms on. Please visit our lab web page https://openwetware.org/wiki/Karas_Lab:FREEZ_Module for more details on FREEZ module tutorials (wiki page, YouTube video), the web app address and updates. Users can reach out for help about installation and software use in a dedicated Slack support channel.

Results

21 patients were included in the study. Mean age at epilepsy onset was 19.5 ± 12 years (range: 13-59), and the mean age at the time of surgery was 36.39 ± 12.13 years (range: 13-59). The mean epilepsy duration before surgery was 16.89 years (range: 4-41). The cohort included 9 males, 9 females, and 3 patients with unspecified sex. All patients underwent ECoG, with an average of 3.05 ± 0.84 ictal events recorded per patient (range: 1-4). Surgical targets were predominantly in the temporal lobe (N = 17), followed by the frontal lobe (N = 2) and the parietal lobe (N = 2). Post-surgery, 14 patients achieved seizure freedom (Engel I), while 7 did not. Among the non-seizure-free patients, 3 were classified as Engel II, 1 as Engel III, and 3 as Engel IV [see *Supplementary Table 1*].

High power in alpha, beta, gamma and high-gamma bands correlates with the EZ

We first inspected the heatmaps for individual patients and sorted the heatmaps be EZ vs non-EZ electrodes. We saw that there was higher power for EZ vs non-EZ electrodes. To determine if this was the case across patients, we calculated the auc for mean power across patients for all frequency bands.

In the patient population that became seizure free after surgery, we compared the AUC of the mean power over time in the first 20 seconds after seizure onset for EZ *vs.* non-EZ electrodes. AUC of mean power for EZ electrodes was significantly higher compared to non-EZ electrodes in alpha, beta, gamma, and high-gamma bands (Fig. 3. alpha 3.3 ± 5.8 e+10 EZ *vs.* 0.2 ± 0.4 e+10 non-EZ, p = 0.0272; beta 1.4 e+10 ± 2.4 e+10 *vs.* 0.05 ± 0.10 e+10, p = 0.0263; gamma 6.9 ± 9.2 e+08*vs.* 3.6 ± 4.9 e+07, p = 0.0013; and high gamma 2.4 ± 3.2 e+07 *vs.* 3.6 ± 10 e+06, p = 0.0086), but not for delta and theta bands (delta 9.3 ± 22.0 e+10 *vs.* 4.9 ± 11.0 e+10, p = 1.0; theta 6.0 ± 11.0 e+10 *vs.* 0.6 ± 1.0 e+10, p = 0.0955). This result suggests that increases in alpha, beta, gamma, and high-gamma mean power immediately after seizure onset correlates with the EZ.

We then performed the same analysis in the patient population that did not become seizure-free after surgery. While there were similar trends, the AUC of mean power for EZ vs. non-EZ electrodes did not reach statistical significance after correcting for multiple comparisons (Sup. Table 2). This result can be interpreted in different ways. Perhaps the EZ was incorrectly identified with visual inspection. Alternatively, less increase in mean power could be a result of poor sampling (e.g. incorrect pre-implantation hypothesis the EZ location led to the EZ being poorly sampled or not sampled). Finally, less increase in mean power could be due to a physiologic or genetic difference in the underlying tissue that is less amenable to surgical resection.

Additionally, we compared AUC of mean power for EZ electrodes in patients who were seizure-free *vs.* not seizure-free after surgery (Fig. 3). EZ electrodes in seizure free patients had significantly higher AUC of mean power compared to in non-seizure free patients in the gamma (6.9 ± 9.2 e+08 seizure free *vs.* 1.4 ± 1.7 e+08 non-seizure free, p = 0.0145) and high gamma (2.4 ± 3.2 e+07 *vs.* 2.1 ± 2.4 e+06, p = 0.0024) bands, but not delta, theta, alpha, or beta bands (delta 9.3 ± 22.0 e+10*vs.* 0.56 ± 4.2 e+10, p = 0.8678 ; theta 6.0 ± 11.0 e+10 *vs.* 1.6 ± 2.7 e+10, p = 0.4956; alpha 3.3 ± 5.8 e+10 *vs.* 0.64 ± 0.86 e+10, p = 0.1320; and beta 1.4 ± 2.4 e+10 *vs.* 0.29 ± 0.31 e+10, p =0.1624) [Supp. Table 3]**.** These results may highlight that the magnitude of the power difference is important, and could reflect differences in EZ sampling or underlying physiology that may affect seizure-freedom outcomes.

Finally we compared AUC of mean power for non-EZ electrodes in patients who became seizure-free vs. not seizure-free after surgery. There were no significant differences between these groups (Sup. Table 3).

Combining Spectral Power Bands with Machine Learning Improves Electrode Classification

With the above results suggesting that multiple different spectral power bands correlate strongly with the EZ, we hypothesized that combining results across all power bands would improve EZ localization. We therefore trained an SRFE machine learning model on the post-operative seizure-free cohort to integrate features across all power bands and identify electrodes within the EZ. We justified the use of an SRFE model by determining that the predictions of individual base models across power bands are not highly correlated, suggesting an ensemble model would combine features effectively (Sup. Fig. 1A). Additionally, the SRFE outperformed both the simple random forest and averaged random forest models (Sup. Fig. 2). A 89.56% threshold on the prediction probability output of the SRFE model was identified from leave-one-out cross validation to result in the greatest number of true positives while maintaining PPV ≥ 95% (Fig. 4A). In this cohort there were a total of 951 electrodes across 14 patients. 202 (21.2%) were marked EZ by an epileptologist. The SRFE yielded 22 true positive predictions within the EZ and one false positive (Fig. 4B). These true positive electrodes were identified in 8 of the 14 (57%) seizure-free patients (PT01, PT02, PT08, PT13, PT17, UMMC002, UMMC005, JH105); the model did not identify any EZ electrodes in 6 patients. The SRFE model achieved an

accuracy of 81.0%, with a PPV of 95.7%, NPV of 80.6%, sensitivity of 10.9%, and specificity of 99.9%.

To gain insight on the contribution of each power band to the SRFE model's EZ classification, we extracted the GLM meta-learner's optimized frequency-specific model coefficients. A larger base model coefficient indicates greater impact of that power band on the final SRFE model, identifying which of the six frequency bands most strongly influence EZ determination. The beta power importance coefficient was highest (0.66 ± 0.07), followed by high gamma (0.53 ± 0.07), delta (0.51 ± 0.06), alpha (0.41 ± 0.07), gamma (0.38 ± 0.05), and theta range (0.18 ± 0.07) (Fig. 4C).

The SRFE model was then applied to the patient population that did not become seizure free after surgery (n=7). We hypothesized that the model would identify EZ electrodes that had not been originally identified or removed. The model identified 3 out of 75 electrodes that had been labeled by clinicians as EZ and removed, and 4 electrodes that were not labeled as EZ and not removed (Fig. 4D). These electrodes were identified across 3 of the 7 non-seizure free patients.

Finally, we examined the temporal evolution of SRFE optimal weights by analyzing the variable importance derived from frequency-specific random forest models. Each model applied variable importance metrics to distinct frequency ranges across the time windows, allowing us to visualize the significance of each frequency band over a 20-second period (Fig. 4E). From seizure onset (0s) to 7.5s, we observed an initial phase dominated by delta activity, transitioning to a period of high gamma predominance between 5 and 7.5s. This was followed by a second phase, characterized by sustained gamma predominance from 7.5s to 15s. The third and final phase, spanning 15s to 20s, was marked by increased prominence in the theta, alpha, and beta frequency ranges, along with a resurgence in delta activity. We then correlated these findings with the averaged time-frequency plot for the 22 EZ electrodes identified by the SRFE model, revealing that periods of high power in the time-frequency plots were temporally aligned with the random forest variable importance metrics (Fig. 4F).

Combining Spectral Power Bands also Predicts Post-Operative Seizure Outcome

The second SRFE was trained using electrodes power deciles across spectral bands (Fig. 5A) to predict patient outcomes. Similar to the EZ identification SRFE, the Pearson correlation coefficients between power band predictions revealed low correlation justifying the choice of ensemble model (Sup. Fig 1B). The SRFE was also compared to predictions constrained to each frequency band individually. The SRFE performed better on patient outcome prediction compared to any individual band RF (Fig. 5B). A 74.2% threshold for the SRFE model was identified from leave-one-out cross validation to result in the greatest number of true positives while maintaining PPV ≥ 95%. For the population of 21 patients, the SRFE model yielded 14 true positives, 6 true negatives, 1 false positive, and 0 false negatives (Fig. 5C). Overall, the

SRFE model achieved an accuracy of 95.2%, with a PPV of 93.3%, NPV of 100%, sensitivity of 100%, specificity of 85.7%, and AUC of 0.91.

To gain insight on the contribution of each power band to the SRFE model's seizure outcome prediction, we extracted the model's optimized frequency-specific random forest model coefficients. The beta power importance coefficient was highest (0.29 ± 0.07), followed by alpha (0.18 ± 0.15), high gamma (0.13 ± 0.15), theta (0.10 ± 0.13), and gamma (0.09 ± 11). Delta was determined to have a negative coefficient value (-0.10 ± 0.12) indicating an inverse correlation between the delta power decile feature and seizure outcome prediction utility.

FREEZ Software for Data Analysis and Visualization

To aid transparency and replicability, the algorithms used in this project were incorporated into the RAVE software platform as the FREEZ module. Created with R Shiny, this module has a graphical user interface via and web browser. The software displays all analyses illustrated in Fig. 1. The software also enables pipelining functionalities for multi-patient analyses.

FREEZ enables interactive analysis of data to examine results of the SRFE models. Fig. 6 illustrates visualizations of the electrode classification SRFE results for four patients. It shows the heatmap results in the beta range. The heatmap displays the electrode numbers and the time in seconds, respectively on the y and x axes. Electrodes marked as EZ are highlighted in blue and stacked on the heatmap bottom. Cross marks highlight electrodes that have been classified as EZ by the SRFE model from the multi-patient analysis. ***Figure 6A*** shows PT08, who was seizure-free, and had all predicted EZ by the SRFE model resected. We can see on ***Figure 6B*** another visualization of electrode classification result on PT15, a seizure free patient, with high gamma power outside of the resected EZ. As such, our SRFE model selected electrodes both within and outside of resected tissue, resulting in the one false positive result. PT10 demonstrates the predictive power of the SRFE model on a non-seizure free patient, with the expected EZ determined by the model to be outside of the tissue that was resected. This can be seen by the EZ electrodes on the heatmap bottom that do not correlate well with high relative power in the area predicted by the SRFE model in ***Figure 6C.*** Finally, PT06 represents a patient where potentially sparse sampling resulted in relative spectral power not being predictive of which electrodes should be removed. High relative power is within the removed electrodes and correlates with the SRFE model predictions as seen in ***Figure 6D***, however the patient was non-seizure free postoperatively.

Discussion

<u>Multi-Patient Analysis</u>

Our spectral statistical findings on this multi-patient study indicate that specific frequency ranges, particularly suppression in the lower frequencies including delta and theta, and high power in gamma and high gamma, play critical roles in identifying the EZ. These results

agree with previous studies computing metrics of this spectral pattern as the Epileptogenic index (12) the Fingerprint method (13) and HFO (29,30). The distribution of frequency band feature importance for the SRFE model indicates that the beta, delta and high gamma frequency range are significant predictor which is confirmed by several studies (10,14,21). Our study confirms complex interplay between spectral features in the EZ. Thus, it is important to consider all frequency bands features for EZ classification generalization on diverse patient data set.

Our multi-patient results suggests that SRFE using mean power of IEEG signals in six different frequency ranges (delta, alpha, theta, beta, gamma and high gamma) is a relevant tool to provide a unified exploration of the correlation between spectral features and EZ localization to automatically classify electrodes. This SRFE performance results encourage further experiments including results based on other features than spectral features like network analysis (10), phase locked high gamma (31) and source localization (11).

Most of the studies describe features correlating with the EZ in case of seizure freedom but evaluate their methods on patient outcome prediction. However, the performance assessment of automatic electrode classification is important to assist clinician in identifying the EZ. Our ensemble model performance (95.7% PPV) for electrode classification compares well with the support vector machine trained on fingerprint features (90% PPV) (32). However, direct comparison is not possible since patient data is unavailable.

Our SFRE results on patient outcome prediction (AUC score 0.91) compare well with similar performance as the Manifold Random Forest (33) trained on the fragility method spatiotemporal results (AUC score 0.89). Their individual Manifold RF on each frequency range performed better than ours, however, in contrast with them we use a classic RF. Moreover, a direct comparison is not possible since seventy patients from the Cleveland clinic included in their study are missing in the Fragility Data Set (10).

The weak random forest models based on individual frequency ranges are weighting the mean power in each time window. Reviewing the frequency-dependent random forest weights, we observed 3 periods of post-ictal frequency content. Period 1) delta and high-gamma, period 2) gamma, and period 3) theta, alpha, and beta. This pattern reflects previously observed time-frequency signatures of the EZ, namely the Fingerprint described by Grinenko *et al* (32).

FREEZ module

The FREEZ software developed for this study provides a user-friendly, web-based tool designed to visualize and detect changes in spectral power associated with the onset of a seizure. The tool is tailored to be accessible to both researchers and clinicians, offering a unified platform for informed decision-making during EZ localization. Developed as a module in R and seamlessly integrated into RAVE, an NIH-supported platform, this tool displays a spatiotemporal spectral analysis, statistics about power feature correlation with the EZ and a template brain with a drop-down menu. The 3D brain map allows users to view both the location of electrodes within the brain and those that exhibit biomarker properties.

This tool encompasses a range of valuable features to enhance the user experience. Firstly, it offers a user-friendly interface. Our web-based tool prioritizes usability, ensuring that users, even without programming backgrounds, can easily navigate and utilize its functionalities. Secondly, the tool is seamlessly integrated into the RAVE platform. Our module streamlines the user experience, eliminating the need to switch between multiple platforms for preprocessing, analysis, and result visualization. Further, users can access and implement other RAVE features. Additionally, recognizing the variability inherent in epilepsy data, we provide users with the flexibility to modify settings and parameters as necessary. This adaptability enhances the tool versatility across diverse epilepsy datasets. Lastly, the module goes beyond calculating and displaying results, but enables real-time dynamic adjustments to parameters, facilitating exploration of data. This feature fosters a more dynamic and personalized analysis experience for users, allowing them to tailor their analysis in real-time.

Furthermore, the open-source nature of our software aims to democratize access to advanced epileptic seizure analysis tools and result reproducibility, potentially bridging the gap in epilepsy surgery referrals and facilitating more informed clinical decisions. Through a user-friendly, web-based interface, our tool not only simplifies the analysis process for researchers and clinicians alike but also provides a flexible and dynamic platform for the exploration of epileptic seizure data. Finally with providing web app and docker image version of our module, we allow users to run and explore our software without the need of downloading software or patient data. This allows the user to interact with results presented in this study and demonstrates reproducibility,

Within the current framework of managing patients with drug-resistant epilepsy, financial impediments are evident, as access to epilepsy surgery is constrained due to the necessity for referrals to a restricted pool of highly specialized physicians (34–36). By making our open-source software freely available, we hope to address this challenge by allowing for it to be leveraged as a quantitative tool to assist with surgical decision making.

Limitations of this study

For fragility (37) patients, the SOZ contacts are an estimation of the resect contacts and the ground truth EZ. Post-surgical T1 MRI was not available. Because XYZ locations were available for only six patients in the fragility (37) multicenter patients, the exploration of the 3D brain plot functionality for correlation between EZ and MSA heatmap results was limited.

The results of this study are limited to twenty-one patients. We plan in the future to test our module on larger patient number thanks to a retrospective patient study. We are currently collecting retrospective iEEG data at our institution with patient specific imaging allowing accurate electrode localization (preoperative T1 MRI, CT post implantation) and accurate mapping of the resected area. This data will allow exploration of 3D brain mapping for EZ localization. However, the tool is freely available to researchers to perform similar study on independent data set than the one used in this paper. Thus, we encourage further testing of our tool.

Acknowledgements

We thank Katrina Reichardt for assistance with data acquisition.

Funding

This research was funded by the American Society for Stereotactic and Functional Neurosurgery Research Fellowship.

Competing interests

The authors have no personal or institutional interest with regard to the authorship and/or publication of this manuscript.

References

1. Perucca E, Perucca P, White HS, Wirrell EC. Drug resistance in epilepsy. The Lancet Neurology. 2023 Aug;22(8):723–34.

2. Kwan P, Arzimanoglou A, Berg AT, Brodie MJ, Allen Hauser W, Mathern G, et al. Definition of drug resistant epilepsy: Consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies: Definition of Drug Resistant Epilepsy. Epilepsia. 2009 Nov 3;51(6):1069–77.

3. Rheims S, Sperling MR, Ryvlin P. Drug resistant epilepsy and mortality—Why and when do neuromodulation and epilepsy surgery reduce overall mortality. Epilepsia. 2022 Dec;63(12):3020–36.

4. Lamberink HJ, Otte WM, Blümcke I, Braun KPJ, Aichholzer M, Amorim I, et al. Seizure outcome and use of antiepileptic drugs after epilepsy surgery according to histopathological diagnosis: a retrospective multicentre cohort study. The Lancet Neurology. 2020 Sep;19(9):748–57.

5. Jehi L. The Epileptogenic Zone: Concept and Definition. Epilepsy Curr. 2018 Feb;18(1):12–6.

6. MacDonald DB, Simon MV, Nuwer MR. Neurophysiology during epilepsy surgery. In: Handbook of Clinical Neurology [Internet]. Elsevier; 2022 [cited 2023 Jun 5]. p. 103–21. Available from: https://linkinghub.elsevier.com/retrieve/pii/B978012819826100017X

7. Gnatkovsky V, De Curtis M, Pastori C, Cardinale F, Lo Russo G, Mai R, et al. Biomarkers of epileptogenic zone defined by quantified stereo-EEG analysis. Epilepsia. 2014 Feb;55(2):296–305.

8. Perucca P, Dubeau F, Gotman J. Intracranial electroencephalographic seizure-onset patterns: effect of underlying pathology. Brain. 2014 Jan;137(1):183–96.

9. Lagarde S, Buzori S, Trebuchon A, Carron R, Scavarda D, Milh M, et al. The repertoire of seizure onset patterns in human focal epilepsies: Determinants and prognostic values. Epilepsia. 2019 Jan;60(1):85–95.

10. Li A, Huynh C, Fitzgerald Z, Cajigas I, Brusko D, Jagid J, et al. Neural fragility as an EEG marker of the seizure onset zone. Nat Neurosci. 2021 Oct;24(10):1465–74.

11. Diamond JM, Diamond BE, Trotta MS, Dembny K, Inati SK, Zaghloul KA. Travelling waves reveal a dynamic seizure source in human focal epilepsy. Brain. 2021 Jul 28;144(6):1751–63.

12. Bartolomei F, Chauvel P, Wendling F. Epileptogenicity of brain structures in human temporal lobe epilepsy: a quantified study from intracerebral EEG. Brain. 2008 Jul;131(7):1818–30.

13. Grinenko O, Li J, Mosher JC, Wang IZ, Bulacio JC, Gonzalez-Martinez J, et al. A fingerprint of the epileptogenic zone in human epilepsies. Brain. 2018 Jan 1;141(1):117–31.

14. Izumi M, Ikeda A. Focal ictal direct current shifts by a time constant of 2 seconds were clinically useful for resective epilepsy surgery. Epilepsia. 2023;

15. Dong X, Yu Z, Cao W, Shi Y, Ma Q. A survey on ensemble learning. Front Comput Sci. 2020 Apr;14(2):241–58.

16. Van Der Laan MJ, Polley EC, Hubbard AE. Super Learner. Statistical Applications in Genetics and Molecular Biology [Internet]. 2007 Jan 16 [cited 2024 Jun 12];6(1). Available from: https://www.degruyter.com/document/doi/10.2202/1544-6115.1309/html

17. O'Leary S, Lesage AC, Camarillo-Rodriguez L, Zhou O, Diosely S, Wang J, et al. Frequency Range Explorer of the Epileptogenic Zone (FREEZ) [Internet]. Zenodo; 2024 [cited 2024 Jun 13]. Available from: https://zenodo.org/doi/10.5281/zenodo.11440403

18. Magnotti JF, Wang Z, Beauchamp MS. RAVE: Comprehensive open-source software for reproducible analysis and visualization of intracranial EEG data. NeuroImage. 2020 Dec;223:117341.

19. Markiewicz CJ, Gorgolewski KJ, Feingold F, Blair R, Halchenko YO, Miller E, et al. The OpenNeuro resource for sharing of neuroscience data. Kahnt T, Baker CI, Dosenbach N, Hawrylycz MJ, Svoboda K, editors. eLife. 2021 Oct 18;10:e71774.

20. Holdgraf C, Appelhoff S, Bickel S, Bouchard K, D'Ambrosio S, David O, et al. iEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology. Sci Data. 2019 Jun 25;6(1):102.

21. Diamond JM, Withers CP, Chapeton JI, Rahman S, Inati SK, Zaghloul KA. Interictal discharges in the human brain are travelling waves arising from an epileptogenic source. Brain. 2023 May 2;146(5):1903–15.

22. Thomson. Spectrum estimation and harmonic analysis. In: Proc IEEE. IEEE; 1982. p. 1055–96.

23. Prerau MJ, Brown RE, Bianchi MT, Ellenbogen JM, Purdon PL. Sleep Neurophysiological Dynamics Through the Lens of Multitaper Spectral Analysis. Physiology. 2017 Jan;32(1):60–92.

24. Frida T. h2o: R Interface for the “H2O” Scalable Machine Learning Platform. 2023.

25. Tomas Nykodym, Tom Kraljevic, Amy Wang, Wendy Wong. Generalized Linear Modeling with H2O [Internet]. 2022. Available from: https://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets/GLMBooklet.pdf.

26. Mark Landry. Machine Learning with R and H2O [Internet]. 2024. Available from: https://docs.h2o.ai/h2o/latest-stable/h2o-docs/booklets/RBooklet.pdf.

27. Magnotti JF, Wang Z, Beauchamp MS. RAVE: Comprehensive open-source software for reproducible analysis and visualization of intracranial EEG data. NeuroImage. 2020 Dec;223:117341.

28. Balach J, Havel T, Janča R, Ježdík P, Čmejla R, Kršek P, et al. 53. Methods of high frequency oscillations detection: Advantages and disadvantages. Clinical Neurophysiology. 2015 Mar;126(3):e49.

29. Burnos S, Frauscher B, Zelmann R, Haegelen C, Sarnthein J, Gotman J. The morphology of high frequency oscillations (HFO) does not improve delineating the epileptogenic zone. Clin Neurophysiol. 2016 Apr;127(4):2140–8.

30. Weiss SA, Lemesiou A, Connors R, Banks GP, McKhann GM, Goodman RR, et al. Seizure localization using ictal phase-locked high gamma: A retrospective surgical outcome study. Neurology. 2015 Jun 9;84(23):2320–8.

31. Grinenko O, Li J, Mosher JC, Wang IZ, Bulacio JC, Gonzalez-Martinez J, et al. A fingerprint of the epileptogenic zone in human epilepsies. Brain. 2018 Jan 1;141(1):117–31.

32. Li A, Ronnan P. Manifold Forests: Closing the Gap on Neural Networks.

33. Khoo A, De Tisi J, Mannan S, O’Keeffe AG, Sander JW, Duncan JS. Reasons for not having epilepsy surgery. Epilepsia. 2021 Dec;62(12):2909–19.

34. Engel J. What can we do for people with drug-resistant epilepsy?: The 2016 Wartenberg Lecture. Neurology. 2016 Dec 6;87(23):2483–9.

35. Steinbrenner M, Kowski AB, Holtkamp M. Referral to evaluation for epilepsy surgery: Reluctance by epileptologists and patients. Epilepsia. 2019 Feb;60(2):211–9.

36. Li A, Huynh C, Fitzgerald Z, Cajigas I, Brusko D, Jagid J, et al. Neural fragility as an EEG marker of the seizure onset zone. Nat Neurosci. 2021 Oct;24(10):1465–74.

37. Fischl B. FreeSurfer. NeuroImage. 2012 Aug;62(2):774–81.

Figure/Table Legends:

Figure 1: FREEZ module results on PT01. Plots are shown generated from the module with the ictal intracranial EEG signal (1A), multitaper power result over time per electrode (1B). 3D brain projection showing marked SOZ channels (1C), as well as the projected multitaper result over time (1D) in the beta frequency range [13 – 30 Hz]. Average power over time line plot (1E), and time plot of the 10th through 100th power percentiles (in steps of 10) for EZ and non-EZ channels (1F), with data from -5 to 10s around seizure onset, which is marked by a green dotted line.

Figure 2: FREEZ module workflow. Data and user defined inputs shown in blue, intermediary processes in green, module visualization options in pink, and outputs in orange.

Figure 3: Area under the curve (AUC) for mean power over time from 0 to 20s following seizure onset. Seizure free and non-seizure free patients separated by marked EZ and non EZ channels in the delta (3A), theta (3B), alpha (3C), beta (3D), gamma (3E), and high gamma (3F) ranges. Significant ($P < 0.05$ following Bonferroni correction) pairwise comparisons marked with *.

Figure 4: Machine learning results from the stacked random forest ensemble (SRFE). (4A) Positive predictive value (PPV) vs. true positives (TP) for changing threshold values (shown in blue text), with maximum true positives with PPV ≥ 95% marked by horizontal and vertical dotted red lines, annotated with threshold, TP, false positives (FP), and PPV. (4B) Results of leave-one-out patient cross-validation. (4C) Distribution of feature importance for the SRFE showing meta-learner coefficient for each base frequency band model. (4D) Results of EZ predictions applied to non-seizure free patients showing distribution between resected and non-resected tissue, juxtaposed against seizure-free patients. (4E) Distribution of feature importance over time for random forest models trained on each frequency range. (4F) Average time-frequency plot from seizure onset to 20s following seizure onset for the 22 electrodes identified as EZ by the SRFE model.

Figure 5: Results from patient outcome predictions are presented in four sections: 5A shows the statistics over time plots for a seizure free and non-seizure free patient in the beta frequency range [13 – 30 Hz]. 5B displays the Area Under the Curve (AUC) for individual frequency bands calculated using random forests, as well as the AUC for the final stacked random forest ensemble (SRFE) along with accuracy, positive predictive value (PPV), and negative predictive value at the threshold cutoff with maximum accuracy. 5C presents the results

from the leave-one-out patient cross validation at the maximum accuracy cutoff, which includes true positives, false positives, true negatives, and false negatives from the SRFE model. Finally, 5D shows the distribution of feature importance for the SRFE showing meta-learner coefficient for each base frequency band model.

Figure 6: Patient multitaper heatmap results in the beta (13 – 30 Hz) range. EZ channels (marked blue) are organized at the bottom of the heatmap and delineated with a blue box. Predicted EZ channels by the random forest are highlighted with an asterisk (*). (6A) PT08, who was seizure-free, had all predicted EZ resected. (6B) PT15, also seizure-free, had one predicted EZ removed, and one not removed. (6C) PT10, who was not seizure-free, had predicted EZ not removed. (6D) PT06, also not seizure-free, had the predicted EZ removed.

Supplementary Table 1: Characteristics of the included patients. Details include demographics, neural recording, surgical procedure, and outcome

Supplementary Table 2: Area under the curve (AUC) average power for included patients in a time window from seizure onset to 20s following seizure onset. Data is mean baselined to a time window -30s to -20s before seizure onset. Reported P-Values for pairwise comparison between seizure condition EZ AUC and seizure condition not EZ AUC. $P < 0.05$ is significant following Bonferroni correction for multiple P-Value testing. * indicates significance.

Supplementary Table 3: Area under the curve (AUC) average power for included patients in a time window from seizure onset to 20s following seizure onset. Data is mean baselined to a time window -30s to -20s before seizure onset. Reported P-Values for pairwise comparison between seizure free AUC and not seizure free AUC. $P < 0.05$ is significant following Bonferroni correction for multiple P-Value testing. * indicates significance.

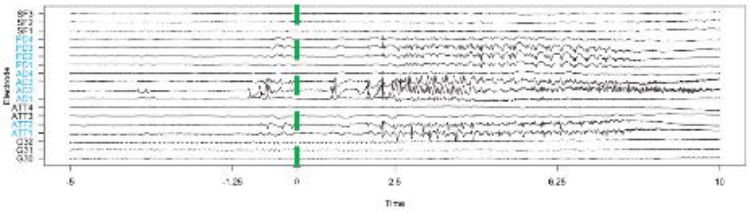
Supplementary Figure 1: Pearson correlation coefficients for machine learning outcome predictions. Correlations between power band predictions for electrode outcome predictions (A) and seizure outcome predictions (B).

Supplementary Figure 2: Machine learning results on electrode selection for multiple models tested machine learning models. Positive predictive value (PPV) (y-axis) vs. true positives (x-axis) for three tested machine learning models on time-frequency features across six frequency ranges (delta, theta, alpha, beta, gamma, high gamma). (2A) Simple random forest model, (2B) frequency-specific random forests with stacked predictions, and (2C) stacked random forest ensemble (SRFE). Blue numbers represent applied thresholds to the output model probabilities. Dotted red lines mark the greatest number of true positives achieved with a PPV ≥ 95% considered the optimal threshold, with intersecting horizontal and vertical dotted red lines specifying the optimal threshold, as well as the number of true positives (TP), false positives (FP), and PPV at that optimal threshold.

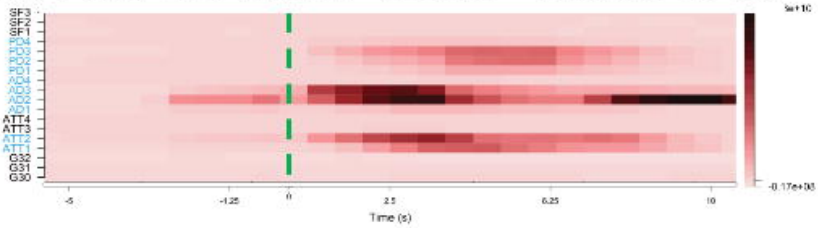
Supplementary Figure 3: Selection panels provided as user interface options in the FREEZ module. Basic parameter selection options include frequency range, window size, step size, and time-half bandwidth product (1A), while advanced options cover the number of DPSS tapers, minimum allowable NFFT size, taper weights, detrend options, and parallel computing options (1B). Analysis options allow for selection of baselining condition and duration (1C), and analysis

frequency and time ranges, with default buttons for predefined ranges in delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), gamma (30-90 Hz), and high gamma (90+ Hz) (1D). Plotting options include selection of normalization method, color map, decibel units toggle, definition of seizure onset zone and resect electrodes (colored blue and purple respectively, with overlap in green), grouping of SOZ/Resect channels at the bottom of the map, and toggle for electrode labels (1E). Machine learning options include prediction capabilities and visualization on heatmaps (1F).

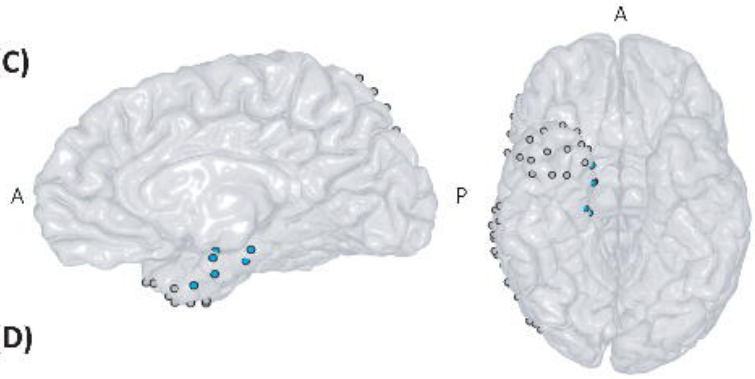
(A) Signal Over Time Per Electrode



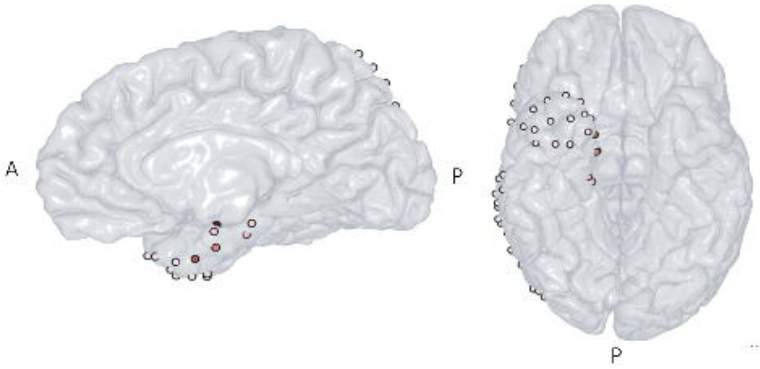
(B) Power Over Time Per Electrode Beta [13 - 30 Hz]



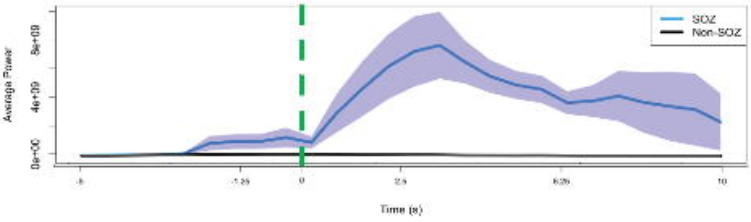
(C)



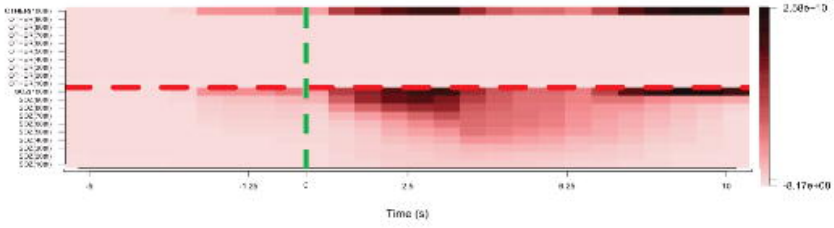
(D)



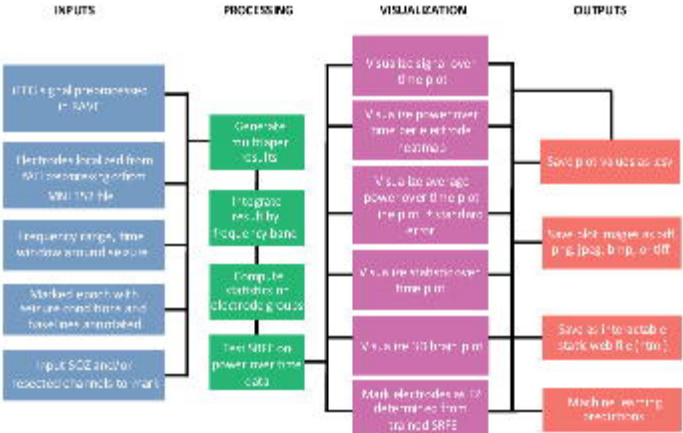
(E) Average Power Over Time Plot Beta [13 - 30 Hz]

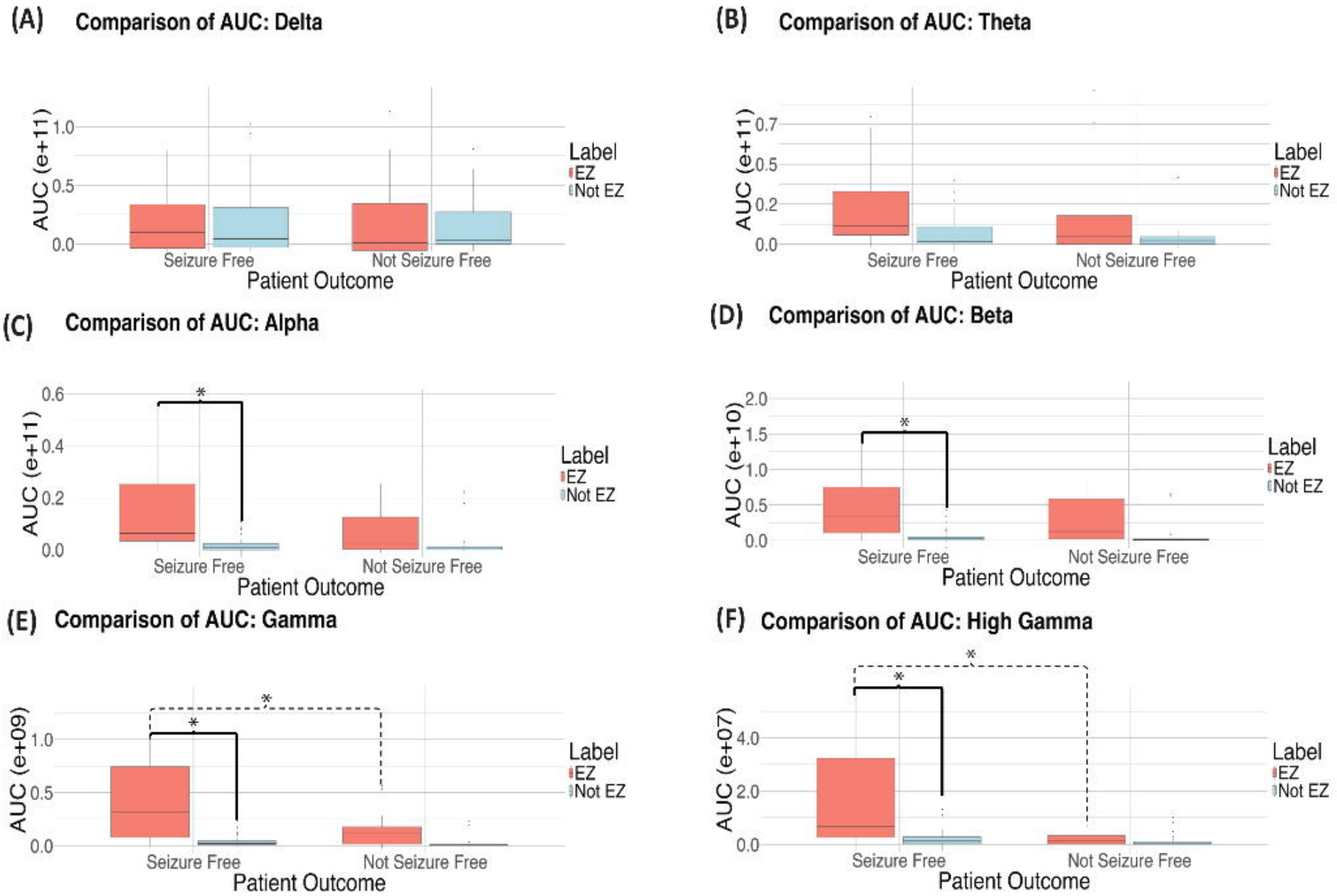


(F) Decile Over Time Plot Beta [13 - 30 Hz]

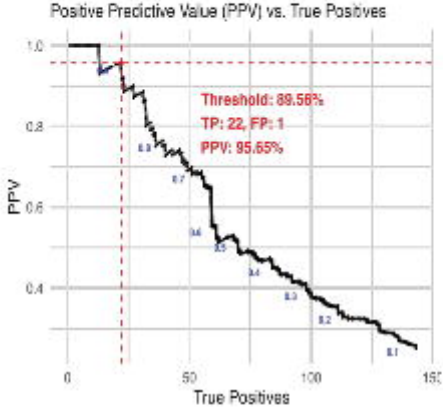


FREEZ Module Workflow




(A) Comparison of AUC: Delta
AUC (e+11)
1.0
0.5
0.0
Seizure Free
Not Seizure Free
Patient Outcome
Label
EZ
Not EZ
(B) Comparison of AUC: Theta
AUC (e+11)
0.7
0.5
0.2
0.0
Seizure Free
Not Seizure Free
Patient Outcome
Label
EZ
Not EZ
(C) Comparison of AUC: Alpha
AUC (e+11)
0.6
0.4
0.2
0.0
*
Seizure Free
Not Seizure Free
Patient Outcome
Label
EZ
Not EZ
(D) Comparison of AUC: Beta
AUC (e+10)
2.0
1.5
1.0
0.5
0.0
*
Seizure Free
Not Seizure Free
Patient Outcome
Label
EZ
Not EZ
(E) Comparison of AUC: Gamma
AUC (e+09)
1.0
0.5
0.0
*
*
Seizure Free
Not Seizure Free
Patient Outcome
Label
EZ
Not EZ
(F) Comparison of AUC: High Gamma
AUC (e+07)
4.0
2.0
0.0
*
*
Seizure Free
Not Seizure Free
Patient Outcome
Label
EZ
Not EZ

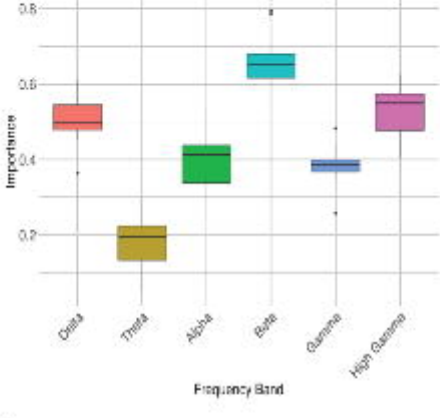
(A) Stacked Random Forest Ensemble:



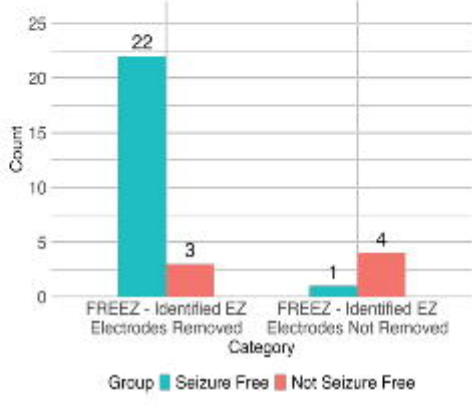
(B) Seizure Free Patients - LOOCV Results

	EZ	Not EZ
Predicted EZ	22	1
Predicted Not EZ	180	748

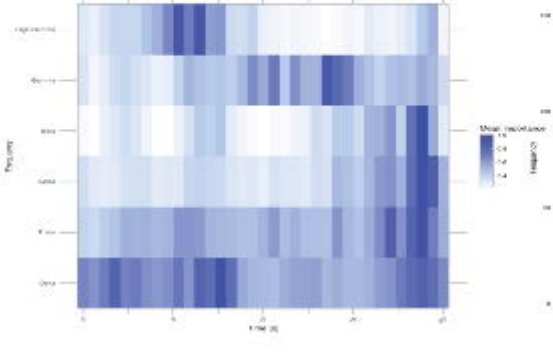
(C) Distribution of Feature Importance Across Folds



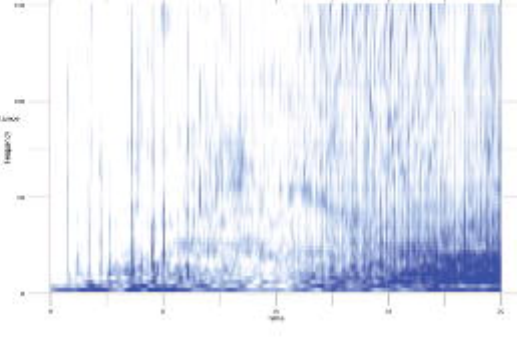
(D) EZ Predictions for Seizure-Free and Not Seizure-Free Patients



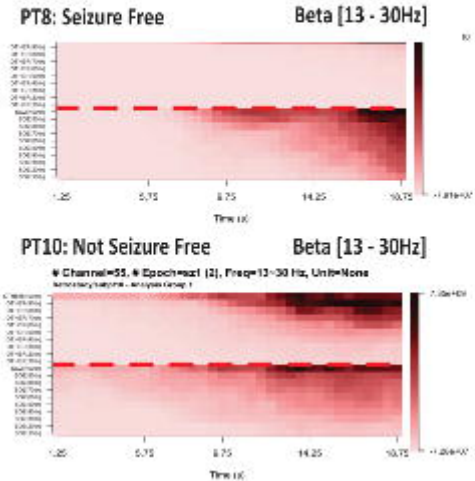
(E) Time Importance per Frequency Across Folds



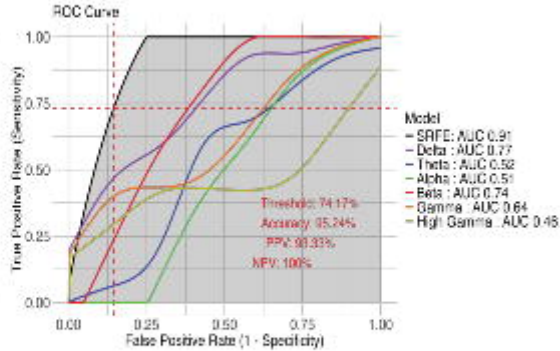
(F) Time-Frequency Plot SRFE 22 EZ Electrodes Averaged



(A) Decile Over Time Plots



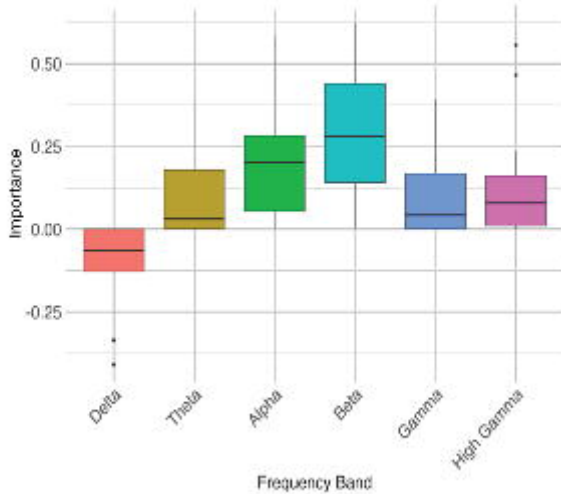
(B) Frequency Band and SRFE Patient Outcome - AUC

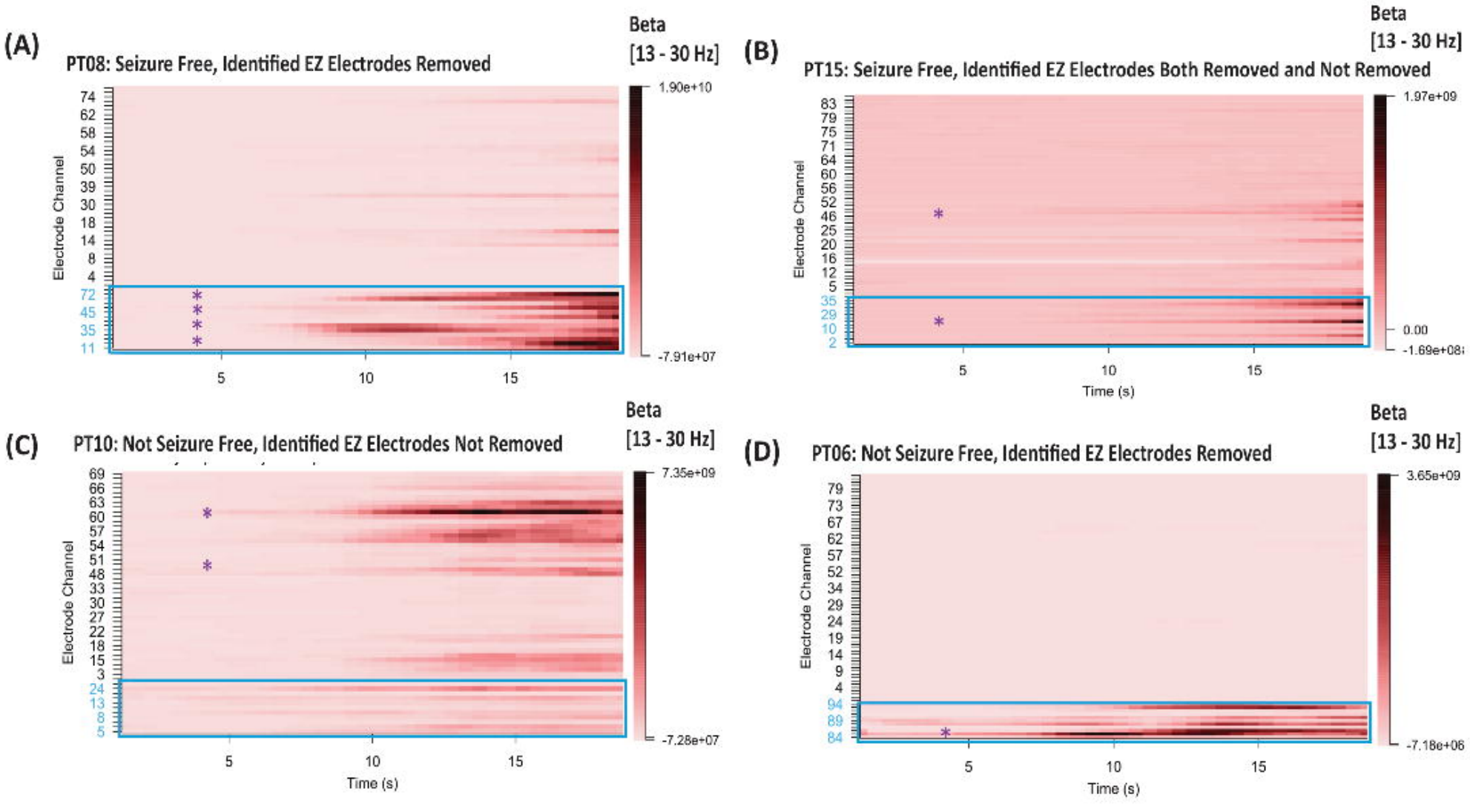


(C) Patient Outcome LOOCV SRFE Results

	Seizure Free	Not Seizure Free
Predicted Seizure Free	14	0
Predicted Not Seizure Free	1	6

(D) Distribution of Feature Importance Across Folds




(A)
PT08: Seizure Free, Identified EZ Electrodes Removed
Beta
[13 - 30 Hz]
Electrode Channel
74
62
58
54
50
39
30
18
14
8
4
72
45
35
11
1.90e+10
-7.91e+07
5
10
15
(B)
PT15: Seizure Free, Identified EZ Electrodes Both Removed and Not Removed
Beta
[13 - 30 Hz]
Electrode Channel
83
79
75
71
64
60
56
52
46
25
20
16
12
5
35
29
10
2
1.97e+09
0.00
-1.69e+08
5
10
15
Time (s)
(C)
PT10: Not Seizure Free, Identified EZ Electrodes Not Removed
Beta
[13 - 30 Hz]
Electrode Channel
69
66
63
60
57
54
51
48
33
30
27
22
18
15
3
24
13
8
5
7.35e+09
-7.28e+07
5
10
15
Time (s)
(D)
PT06: Not Seizure Free, Identified EZ Electrodes Removed
Beta
[13 - 30 Hz]
Electrode Channel
79
73
67
62
57
52
34
29
24
19
14
9
4
94
89
84
3.65e+09
-7.18e+06
5
10
15
Time (s)