A Spectral Analysis of EEG (Brain data) Filtering Techniques: Evaluating Noise Reduction and Signal Preservation in Seizure Data

Word Count: 5478

Introduction:

Seizures are neurological events characterized by sudden and abnormal electrical activity in the brain, often leading to unpredictable changes in behavior, movement, and consciousness. Because seizure triggers can vary widely, including genetic factors, daily habits, and physiological conditions, they are difficult to anticipate and can often go undetected without proper monitoring. In the United States alone, over 150,000 individuals experience seizures each year, suggesting the need for accurate and reliable methods of detection and analysis (Allan).

One of the primary methods used to study and detect seizures and other neurological conditions is electroencephalography (EEG), which records electrical signals from the brain through electrodes placed on the scalp. EEG provides detailed information about brain wave frequencies and patterns, allowing researchers to observe changes in neural activity associated with seizure events. These recordings are often stored in large medical or research databases. However, raw EEG data contains significant background noise, including muscle movement, environmental interference, and unrelated brain activity, which can obscure meaningful seizure signals.

To address this issue, researchers commonly apply signal processing techniques such as filtering to reduce noise and highlight important patterns. Traditional filters, including Butterworth filters, are designed to isolate specific frequency ranges, while smoothing techniques such as the Kolmogorov Zurbenko (KZ) filter help reduce short term fluctuations in the data. Although these methods can improve signal clarity, they often involve trade offs between removing noise and preserving important information, excessive filtering may distort seizure features. Additionally, many existing studies focus on applying single filtering methods or complex machine learning models, rather than directly comparing how different filtering approaches affect the interpretability of EEG data.

As a result, there remains a gap in understanding how combinations of filtering methods or adaptable filtering approaches can balance noise reduction with signal preservation. This study investigates how different EEG filtering techniques, including Butterworth and KZ filters, influence the accuracy and clarity of seizure signal detection when applied to publicly available EEG datasets. By comparing filtered and unfiltered data using statistical and frequency based analyses, this research seeks to determine whether a combined or adaptive filtering approach can improve the interpretation of EEG signals and support more effective seizure detection. This leads to the central research question: To what extent do Butterworth, Kolmogorov Zurbenko, and hybrid/adaptive filtering methods affect spectral integrity and waveform preservations in seizure versus non seizure EEG segments, and which approach most effectively balances noise reduction with retention of clinically meaningful signal features?

Literature Review:

Research on EEG seizure detection shows that preprocessing is not just a basic setup, rather something that directly affects how accurate and reliable detection models are. Preprocessing in EEG research typically involves several key steps including signal segmentation, band-pass or high-pass filtering, artifact removal, and normalization, all of which directly influence how features are extracted from the signal and how well trends can be analyzed. Xiaodan Deng explains that even very advanced deep learning models fail when the EEG signals are not properly filtered, which indicates that model accuracy primarily depends more on preprocessing than the model itself (Deng, 2025). Additionally, Andreas Widmann and his team highlight that when researchers apply filters without carefully choosing cutoff frequencies, the timing and shape of neural signals can be altered. Widmann's study specifically evaluated finite impulse response (FIR) filters using different cutoff frequencies and transition bandwidths, suggesting that sharper filters, while effective at removing low frequency drift, introduced temporal distortions that altered the latency of neural responses. For example, Widmann and Tanner found that certain high-pass filters introduce phase shifts in the recordings that mimic actual brain responses or epileptic events in Tanner's research, which can mislead analysis into detecting a seizure where none actually exists. (Widmann et al., 2015; Tanner, 2015). Tanner's methodology and study further showed that applying high-pass filters above 0.3 Hz significantly altered slow wave EEG components, particularly in event related potential studies, suggesting that commonly used preprocessing settings can unintentionally generate misleading neural patterns. These findings show how small changes in preprocessing parameters can drastically change the interpretation of EEG data.

Alberto Correa shows that adaptive wavelet thresholding filters can adjust in real time to remove unwanted motion and muscle noise, thus improving clarity. He explains, "the proposed adaptive filtering algorithm improves EEG clarity by continuously adjusting its coefficients, allowing seizure related activity to remain visible while removing baseline drift and muscle noise," meaning that crucial seizure signatures remain visible even after noise reduction. Within this methodology, wavelet coefficients are adjusted dynamically based on signal variance at different scales, allowing the filter to suppress noise in high frequency bands while preserving brief seizure related spikes that occur across multiple resolutions (Correa et al., 2019, p. 312). This indicates how newer filters are designed to respond to the data instead of using the same settings for every signal. Unlike traditional filters that apply fixed cutoff values, this method continuously updates its filtering behavior, making it more suitable for non stationary signals such as EEG where noise characteristics change over time (Correa et al., 2019).

Furthermore Serkan Subasi reflected that wavelet feature extraction can reliably distinguish between normal and epileptic EEG patterns by analyzing energy distributions across scales. Subasi’s methodology involved decomposing EEG signals into multiple wavelet sub-bands and calculating statistical features such as energy, entropy, and variance from each level, which were then used as inputs for classification models (Subasi, 2007). U. Rajendra Acharya adds on, explaining that “wavelet transforms capture brief events characteristic of seizures,” which are often lost when static band pass filters are used (Acharya et al., 2013, p. 12). Acharya’s work similarly applied wavelet transforms combined with nonlinear feature extraction techniques, showing that seizure detection improves when signal behavior is preserved rather than smoothed out through rigid filtering. Additionally Xin Li reflects this through an entropy based detection method that becomes far more accurate when signals are properly filtered. Li’s entropy based approach meant that inconsistent preprocessing introduces randomness into entropy calculations, reducing the reliability of detection thresholds. When the preprocessing is inconsistent or too aggressive, entropy fluctuations become random, and the entire detection process breaks down (Li et al., 2018). These studies collectively establish that filtering precision determines whether EEG analysis yields reliable detections or introduces analytical distortion (Subasi, 2007; Acharya et al., 2013; Li et al., 2018).

Xiangyu Zhang reviewed over one hundred EEG based seizure detection systems and found that “a lack of preprocessing consistency is the leading cause of poor reproducibility across datasets” (Zhang et al., 2024, p. 4). Similarly, Abdelmalik Chaddad analyzed dozens of EEG processing methods and stated, “variability in preprocessing rather than differences in neural network architecture" (Chaddad et al., 2023, p. 9). Shu Wong reinforced this issue by comparing public EEG datasets such as CHB-MIT and Bonn, explaining that differences in electrode configuration, sampling rates, and noise filtering methods "alter the frequency content of EEG signals," making fair comparisons between models nearly impossible (Wong et al., 2023, p. 3). This growing concern highlights why preprocessing transparency is now considered as important as the detection algorithm itself (Zhang et al., 2024; Chaddad et al., 2023; Wong et al., 2023).

Even with the recognition that preprocessing significantly impacts EEG analysis, there remains a lack of standardization in how filtering parameters are selected and reported across studies. De Cheveigne and Nelken (2019) analyzed common EEG preprocessing methods and found that a lot of researchers apply filters without fully considering how parameter choices such as cutoff frequency and transition band affect trend structure in the signals. Their work proposed that improperly designed filters can introduce fabricated structures that resemble real brain activity, especially in low frequency bands. Additionally fixed filtering parameters may not be consistent across different time periods within the same recording. As a result, filtering must be evaluated not only for its ability to remove noise, but also for its impact on preserving meaningful signal characteristics.

Guillaume Pernet developed EEG-BIDS, which provides a unified and uniform way to organize EEG datasets and report filtering details such as cutoff frequencies, sampling rates, and referencing methods. He emphasizes that "no preprocessing step should be omitted from documentation," which ensures that all future analyses can be traced back to the original raw data. Later, Pernet presented a reproducible software combining EEGLAB and LIMO EEG that automatically tracks parameter choices within the EEG data itself, making EEG research both transparent and repeatable (Pernet et al., 2019, p. 2; Pernet et al., 2021). These frameworks make EEG research more open and reliable, putting consistency and accuracy on the same level as new discoveries (Pernet et al., 2019, 2021).

Spectral analysis using methods such as Welch's power spectral technique has become a preferred approach to evaluate signal reputability. By examining power retention in key frequency bands such as beta, delta, theta, gamma, and alpha, researchers can see whether filtering preserved meaningful neural activity or removed it (Widmann et al., 2015). Studies that combine adaptive filters, like those by Alberto Correa with spectral evaluation breaking down these EEG readings into their individual frequencies, suggest higher reliability across datasets and patients, showing that filters can be both precise and transparent (Correa et al., 2019).

Altogether, these studies suggest that preprocessing defines the reliability of EEG analysis more than any post modeling choice. However, the literature reveals a methodological gap as studies advocating adaptive approaches such as Correa et al. (2019) evaluate filters primarily within classification pipelines, where preprocessing quality is measured indirectly through detection accuracy. This study addresses that gap by treating preprocessing as the primary experimental variable, evaluating filter effects through both waveform morphology and spectral decomposition without the influence of a downstream detection model.

Another critical aspect of EEG analysis is the relationship between filtering and frequency interpretation, since many detection methods rely on analyzing signal power across specific frequency bands. Techniques such as the Fast Fourier Transform (FFT) and Welch's method are commonly implemented to estimate power spectral levels and density, allowing researchers to gauge how signal energy is distributed across frequencies (Welch, 1967).

However, these methods are highly sensitive to preprocessing decisions, as filtering can alter specific frequency components before analysis even begins. Sanei and Chambers (2007) explain that improper filtering can lead to exclusions of spectral trends, directly affecting the interpretation of neural activity. Because of this, new EEG research increasingly evaluates filtering techniques not only in the time domain, but also by examining how well they preserve frequency based methods.

Overall, Andreas Widmann and Darren Tanner warn that poorly chosen filters can create false spikes, while Alberto Correa, Serkan Subasi, and U. Rajendra Acharya suggest that adaptive and wavelet frameworks preserve authentic seizure features and patterns. Xiangyu Zhang, Abdelmalik Chaddad, and Shu Wong extend this by identifying inconsistency as the biggest barrier to reliable progress, and Guillaume Pernet presents reproducible solutions that bring much needed standardization (Widmann et al., 2015; Tanner, 2015; Correa et al., 2019; Subasi, 2007; Acharya et al., 2013; Zhang et al., 2024; Chaddad et al., 2023; Wong et al., 2023; Pernet et al., 2019, 2021). This investigation builds on this by developing a reproducible filtering framework in R that adapts to noise levels while maintaining openness, ensuring that every improvement in detection is focused on data preservation.

Methodology:

This study follows a structured signal processing approach to investigate how different filtering techniques influence EEG signal clarity and interpretation. Rather than optimizing filtering models, this methodology treats preprocessing as the primary variable, enabling direct comparison of filtering methods through both time domain and frequency analysis, building on findings from Deng (2025) and Zhang et al. (2024) that preprocessing decisions influence EEG outcomes more than model architecture.

Data Source and Selection:

The EEG data for this study was collected from the CHB MIT Scalp EEG dataset of the Physionet database, a publicly available dataset commonly used for brain condition related research (Shoeb, 2009; Goldberger et al., 2000). This dataset contains EEG recordings from patients with seizure events, making it viable for comparisons between seizure and non seizure conditions.

Eight patients were selected from the CHB-MIT dataset: CHB01, CHB02, CHB03, CHB04, CHB05, CHB06, CHB07 and CHB08. For each patient, one seizure segment and one non seizure segment were extracted, each spanning 20 seconds. All segments were drawn from channel FP1-F7 to maintain electrode consistency across recordings. EDF files were loaded at their native sampling rate of 256 Hz. Seizure onset and offset timestamps provided in the CHB-MIT files were used to isolate seizure segments, with non seizure segments extracted from intervals at least 10 minutes prior to any seizure event.

Using both seizure and non-seizure data segments enables the study to evaluate whether filtering techniques preserve meaningful seizure related trend features while also minimizing background noise. Previous research emphasizes that variability across datasets and recording conditions can significantly affect EEG interpretation, particularly due to differences in electrode placement and signal quality (Wong et al, 2023). Maintaining consistent selection criteria across all recordings is crucial so differences in results are because of filtering methods rather than data inconsistency.

Data Loading and Data Application:

All EEG recordings were processed using the R Studio Program, where EDF files are loaded using the EDFReader package (R Core Team, 2024; Toffolo et al., 2019). Each recording was separated into individual EEG channels. Next, to prepare the data for analysis, signals were segmented into fixed time windows. Segmentation is necessary because EEG signals are non stationary, meaning their statistical properties change over time. By analyzing smaller windows of time in the channels, the study monitors that filtering effects can be observed more efficiently. This approach aligns closely with previous research where through Subasi's methodology, segmentation was heavily emphasized to improve signal feature extraction and signal stability (Subasi, 2007).

Additionally, no initial aggressive preprocessing was applied prior to filtering, thus ensuring that the raw signal remains intact for initial comparisons. Early preprocessing steps can display distortions that affect all following analyses (Widmann et al, 2015; Tanner, 2015).

Filtering Methods:

Two primary filtering techniques were applied to each segment, Butterworth filtering and Kolmogorov Zurbenko (KZ) smoothing. These methods were selected due to the fact they represent completely different approaches to signal processing. Filtering operations were implemented using standard signal processing tools in R, including the signal package (Signal Developers, 2023).

The Butterworth filter is a frequency based filter designed to isolate specific EEG bands by applying cutoff frequencies. In this study, The Butterworth filter was implemented as a 4th-order bandpass filter with cutoff frequencies of 0.5 Hz and 40 Hz, preserving the standard EEG bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and the lower gamma range (30–40 Hz). A 4th-order configuration was selected to balance roll-off sharpness with minimal phase distortion, consistent with recommendations from Widmann et al. (2015). The KZ filter was applied using a window size of m = 5 samples across k = 3 iterations, producing a smoothing effect that progressively attenuates high-frequency components while preserving low-frequency trends. These parameters were held constant across all recordings to ensure cross-patient comparability. However prior research has shown that filter parameters such as cutoff frequency and order greatly influence signal trend shape and cutoff values. Widdmann et al. (2015) indicated that poorly selected cutoff values on filters can trigger distortions with the data. Additionally Tanner (2015) found that overly aggressive high pass filtering can create non natural signal components. Therefore filtering parameters were closely monitored in order to minimize distortions. (Refer to Appendix F for further explanation)

Signal Analysis and Evaluation:

After filtering, each signal was analyzed using both time domain and frequency domain techniques. In the time domain, waveform shape and amplitude consistency was visually and analytically compared across raw and filtered out signals. While in the frequency domain, Fast Fourier Transform (FFT) was used to convert signals into their spectral representations. Power spectral density was estimated using previously mentioned methods such as Welch's approach, which improves signal reliability by averaging frequency components across segments (Welch, 1967). Frequency domain analysis was prioritized because EEG interpretation depends heavily on power distributions across neural frequency bands. (Refer to Appendix for further explanation)

To evaluate filter performance, several metrics and parameters were applied. Signal to noise ratio (SNR) was calculated as $SNR = 10 \cdot \log_{10}(P\_signal / P\_noise)$, where P_signal represents the estimated power within clinically relevant EEG bands (0.5–40 Hz) and P_noise represents residual power above 40 Hz, which is produced in scalp EEG recordings. This definition isolates high frequency contamination as the noise component, consistent with standard EEG (Sanei & Chambers, 2007). Additionally, frequency band preservation was analyzed to assess whether important neural components were retained after filtering. Previous studies such as Sanei and & Chambers work have shown that improper filtering can distort EEG data and reduce reliability. Their evaluation of both clarity and preservation ensures that filters are not solely judged on noise reduction and serve as a reliable guide for this research (Sanei & Chambers, 2007).

Comparative Analysis & Adaptive Filter Creation:

Following the initial analysis, results from raw, Butterworth filtered and KZ filtered signals were directly compared via a spectral analysis from the results of a signal summary displaying all values. This comparison focused on identifying which method best balanced noise reduction with signal preservation. (Refer to Appendix E for further explanation)

Based on these observations, a custom adaptive filtering approach was developed by implementing a signal variance mechanism. Specifically, the standard deviation of each EEG segment was computed across a sliding window of 256 samples. When the standard deviation exceeded a set threshold (1.5× the global segment mean), the filter defaulted to more Butterworth dominant processing to preserve spike trends. When variability fell below this threshold, KZ smoothing was weighted more heavily to suppress baseline drift. This switching procedure is expressed as: $F(x) = \alpha B(x) + (1-\alpha)KZ(x)$, where $\alpha \in [0,1]$ is dynamically assigned based on variance and $B(x)$ and $KZ(x)$ represent the Butterworth and KZ filtered outputs respectively. This approach approximates the signal-responsive behavior described by Correa et al. (2019) while remaining reproducible within the R environment. This step is justified again by the work of

Correa in which static filtering was insufficient in handling non stationary EEG data (Correa et al).

The adaptive filter was then applied across multiple EEG recordings to evaluate consistency. Prior research such as Zhang et al (2024) has identified lack of preprocessing consistency as a major issue in EEG studies, while other research such as Chaddad et al (2023) indicates that testing the adaptive filter across multiple recordings ensures that improvements are not limited to a single recording and in the case of Chaddad, his research implemented a review of multiple publicly available EEG datasets, including epilepsy and sleep recordings.

<u>Research Pipeline Visual:</u>

Raw EEG (EDF files) → Channel Extraction → Segmentation (fixed time windows) → Butterworth Filtering → KZ Filtering → Adaptive Filtering → Time Domain Analysis (waveform comparison) → Frequency Domain Analysis (FFT, Welch PSD) → Statistical Evaluation (standard deviation, band power, noise reduction)

Definitions:

Term	Definition
EEG	Electrical recording of brain activity via scalp electrodes
FFT	Converts time domain signals into frequency components
PSD	Measures signal power across frequencies
Standard Deviation	Measures signal variability
Adaptive Filtering	Adjusts filtering parameters based on signal behavior

Data Collection/Results:

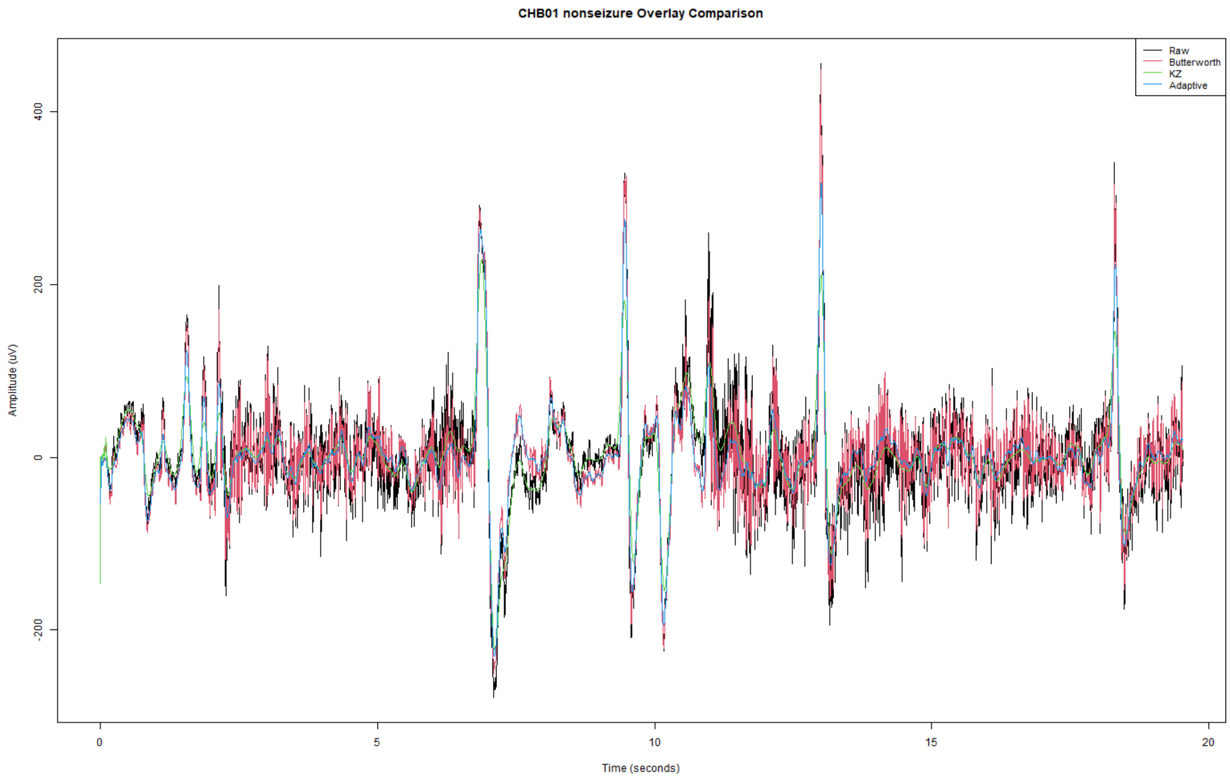
Following the application of the intended methodology, multiple sets of EEG graphs were generated to compare raw and filtered data across various conditions. These graphs include time domain waveform plots, frequency domain representations and statistical analysis of signal trend/behavior across various different filtering techniques. Additionally all numerical values that pertain to the graphs were extracted directly from a signal summary listing all plotted and compared values.

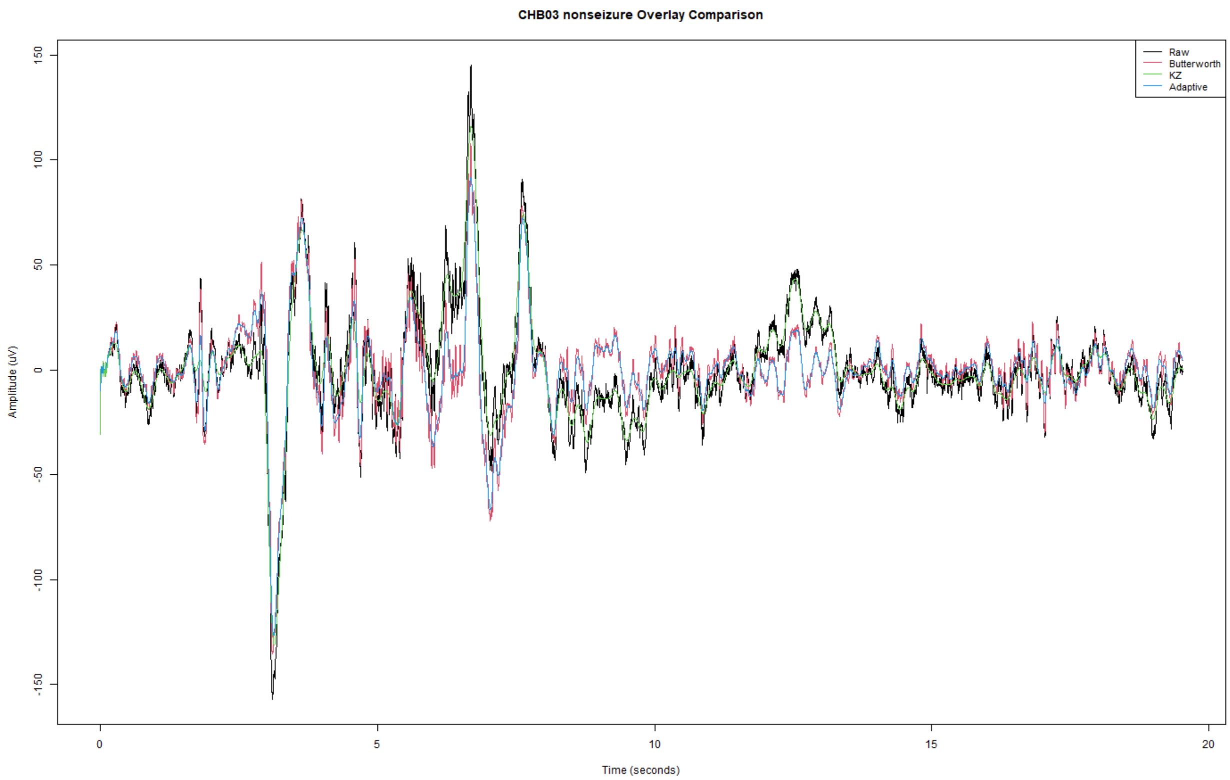
All generated graphs were organized into grouped comparisons, allowing direct visualization of differences across raw, Butter Worth, KZ and adaptively filtered signals across the same EEG segments. These graphs aided in highlighting how each filtering method altered the signal over time and across frequencies. Amplitude values were measured in microvolts (µV), representing the strength of the brain's electrical activity. Frequency values were measured in Hertz (Hz), representing how fast the signal moves back and forth per second, and were used consistently across all graphs. This organization reveals how each filter altered waveform shape and frequency distribution across conditions.

Filter Type	**Standard Deviation Range (µV)**	**Noise Reduction (%)**	**Key Effect**
Raw	40.7 - 88.0	0%	High variability & noise present
Butterworth	32 - 76	6.5% - 16.1%	Moderate smoothing, preserves structure
KZ Filter	25.3 - 70.3	12.6% - 38.1%	Moderate smoothing, preserves structure
Adaptive	34.7-70.3	9.2% - 34.0%	Balanced noise reduction and preservation

(See Appendix E for full signal summary chart)

Raw EEG signals were plotted for each selected recording segment to establish a baseline for comparison. These graphs showed high amplitude, with noise patterns such as spikes and baseline drift. Across the analyzed segments, raw signal amplitudes ranged from approximately −2330 µV to +1564 µV, with standard deviations varying between 40.7 µV and 88.0 µV, depending on the recording. In several segments, these fluctuations made it difficult to differentiate structured signal activity from background interference, particularly in non seizure intervals where signal changes appeared less pronounced. Noise levels also appeared inconsistent across recordings, with some segments showing concentrated bursts of high frequency activity while others displayed gradual baseline drift over time (see Figures 1a. & 1b.).

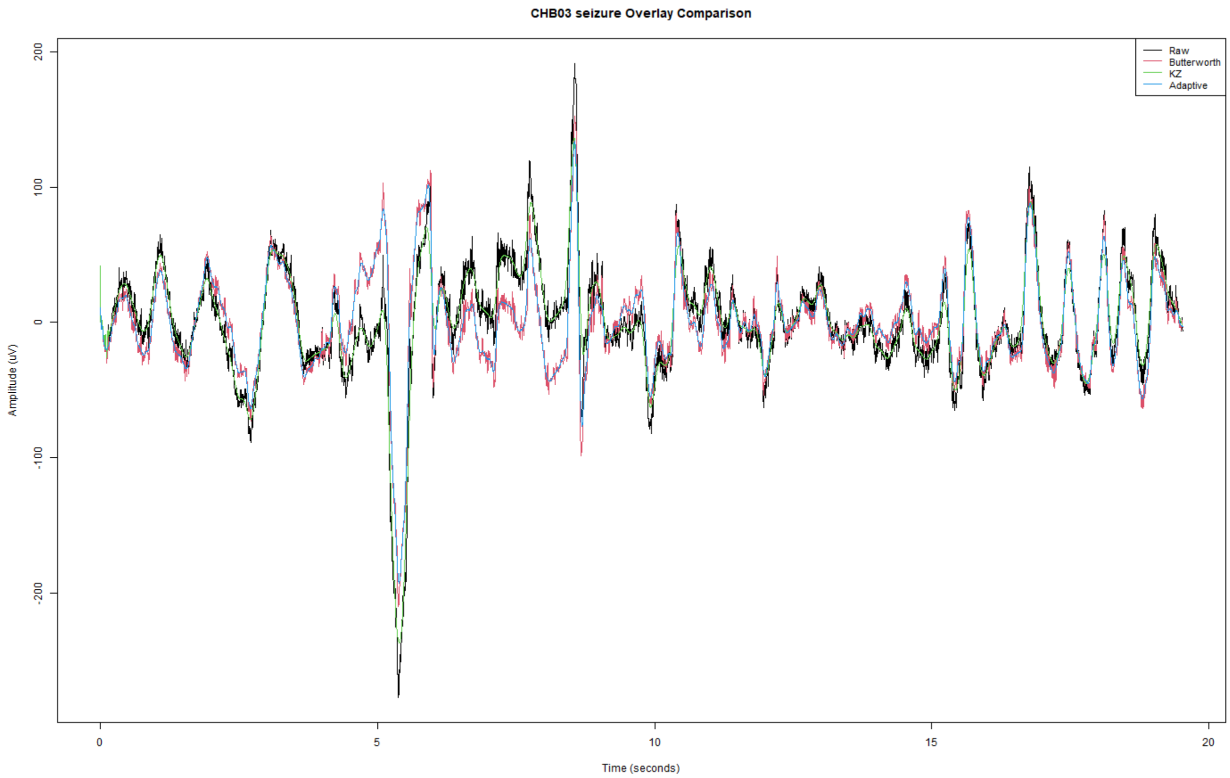


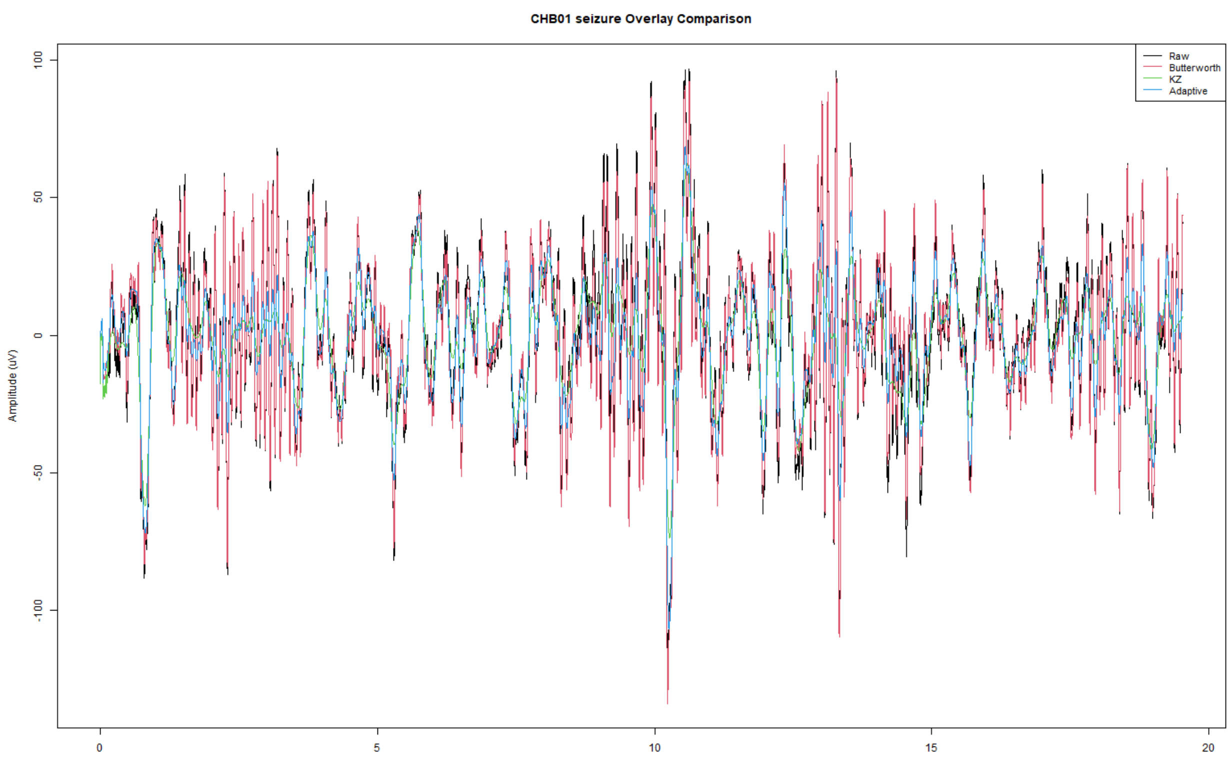


(Figures 1a. & 1b.). Comparison of raw and filtered EEG signals for non seizure segments from patients CHB01 and CHB03. Overlay plots display the raw signal together with Butterworth, KZ, and adaptive filtered outputs across the same time interval. Amplitude is measured in microvolts (µV) and time is measured in seconds. These graphs illustrate how filtering reduces rapid high frequency fluctuations while preserving major waveform structures across multiple patients. Additional waveform comparisons for other recordings are provided in Appendix A.

After applying the Butterworth filter, a second set of waveform plots was generated. These graphs showed a reduction in high frequency noise, with signals appearing smoother and more consistent in amplitude over time. Standard deviation values decreased from raw values of 40-88 µV to 32-76 µV, corresponding to noise reductions of approximately 6.5%-16.1% across segments. In many cases, low frequency components became more visually distinguishable, and large amplitude spikes appeared more defined compared to the raw signal. However, slight shifts in waveform structure were observed, particularly in regions where rapid signal changes occurred, indicating that filtering altered the temporal shape of certain signal features. In some segments, peak amplitudes were reduced relative to the raw signal, consistent with exaggeration of higher frequency components, as illustrated in Figure 1 and Figure 2.

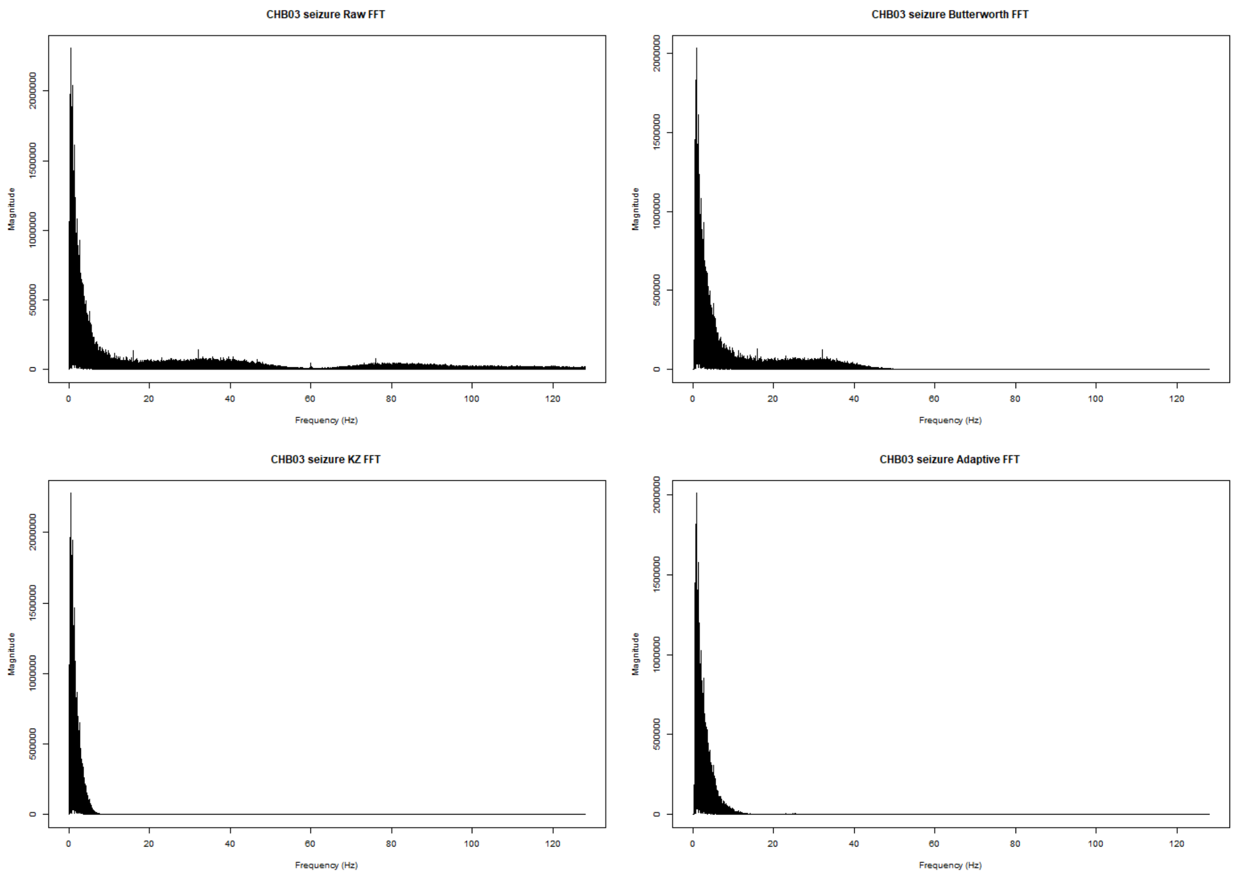
KZ filtered signals were then plotted and compared against both raw and Butterworth filtered outputs. These graphs implied a stronger smoothing effect, where short term fluctuations were significantly reduced and long term trends became more visible. The signal appeared more stable overall, with fewer abrupt changes in amplitude. Compared to raw signals, standard deviation values decreased further to 25.3-70.3 μV, corresponding to noise reductions of approximately 12.6% - 38.1%. Amplitude transitions appeared more gradual across time, and high frequency variability was substantially reduced. However, in some segments, rapid spikes present in the raw signal were less pronounced or partially flattened, indicating that some high frequency signal components were reduced during the smoothing process (see Figure 2).


CHB03 seizure Overlay Comparison
Raw
Butterworth
KZ
Adaptive
Amplitude (uV)
Time (seconds)
200
100
0
-100
-200
0
5
10
15
20


CHB01 seizure Overlay Comparison
Raw
Butterworth
KZ
Adaptive
Amplitude (uV)
Time (seconds)
100
50
0
-50
-100
0
5
10
15
20

(Figures 2a. & 2b.) Overlay comparison of raw and filtered EEG signals for seizure segments from patients CHB01 and CHB03. The raw signals do contain irregular spikes and high frequency variability, while filtered signals display progressively smoother waveforms. The Butterworth filter moderately reduces noise, while the KZ and adaptive filters produce stronger smoothing effects. These patterns are consistent across both patients. Additional seizure waveform comparisons are shown in Appendix B.

To further evaluate differences between filtering methods, Fast Fourier Transform (FFT) graphs were generated for each signal type. These frequency domain plots display the magnitude of signal components across frequencies measured in Hertz (Hz). Raw signal spectra showed a wide distribution of power across frequencies up to about 120 Hz, with substantial energy present at higher frequencies, indicating noise. Butterworth filtered spectra showed reduced high frequency magnitude and clearer concentration of power within lower frequency ranges. In contrast, KZ filtered spectra suggested further suppression of higher frequencies, with most signal energy ranged below 10 Hz, producing smoother spectral profiles (see Figure 3).



(Figure 3.) Fast Fourier Transform (FFT) comparison of raw and filtered seizure signals for patient CHB03. FFT graphs display the distribution of signal magnitude across frequencies measured in Hertz (Hz). Raw signals show broad frequency distributions with noticeable high frequency noise, while filtered signals concentrate signal energy within lower frequency ranges. Similar frequency compression patterns are observed across both patients. Additional FFT comparisons across other recordings are included in Appendix C.

Power spectral density estimates using Welch's method were also calculated to identify frequency behavior across segments. These plots estimate signal power across frequency bands while reducing variance through windowed averaging. Band power values measured in microvolt squared (μV^2) from the signal summary showed that lower frequency ranges such as delta 0.5-4 Hz and theta 4-8 Hz contained the highest power across both seizure and non seizure segments. For example, delta band power values ranged from 1.59×10^8 to 2.30×10^9 μV^2, while theta values ranged from 1.31×10^6 to 2.39×10^8 μV^2 depending on the segment. As shown by several seizure recordings, this reflects that lower frequency band power values were elevated relative to non seizure segments, although this pattern was not consistent across all trials. Thus, differences between seizure and non seizure segments were easier to spot in filtered signals compared to raw data, particularly in lower frequency bands.

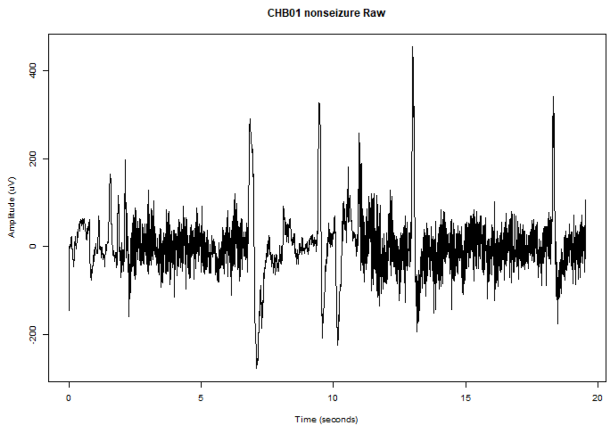
Frequency Band	**Range (Hz)**	**Power Range (μV^2)**	**Observation**
Delta	0.5 - 4	1.59×10^8 - 2.30×10^9	Highest power across segments
Theta	4 - 8	1.31×10^6 - 2.39×10^8	Elevated in seizure segments
Alpha	8 - 13	1.75×10^7 - 5.49×10^7	Moderate presence
Beta	13 - 30	2.49×10^6 - 2.07×10^8	Reduced after filtering
Gamma	30+	2.55×10^5 - 1.96×10^8	Mostly suppressed post-filtering

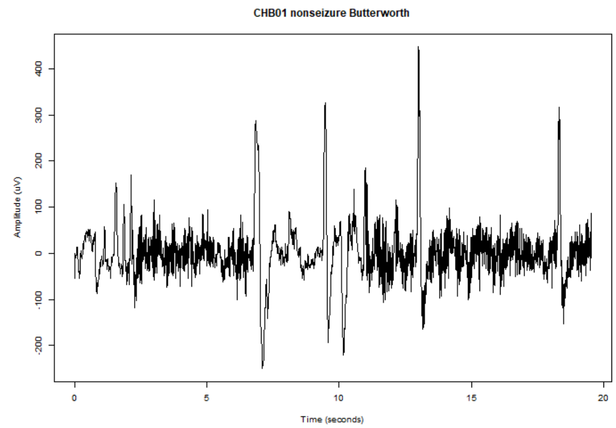
(See Appendix E for full signal summary charts)

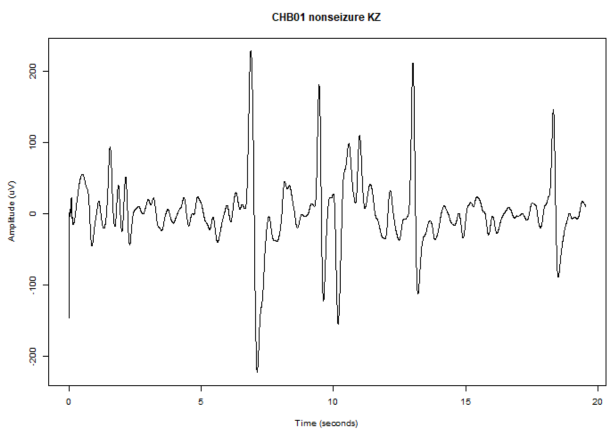
Additional comparative plots were generated to evaluate signal to noise characteristics across filtering conditions including overlay waveform comparisons and band power summaries across delta, theta, alpha, beta, and gamma ranges. Filtered signals consistently showed reduced noise levels relative to raw data, but the amount of reduction varied between filtering methods. Higher frequency bands, particularly beta 13-30 Hz and gamma above 30 Hz, displayed greater variability across recordings, suggesting that filtering effects were dependent on signal characteristics rather than uniform across all segments.

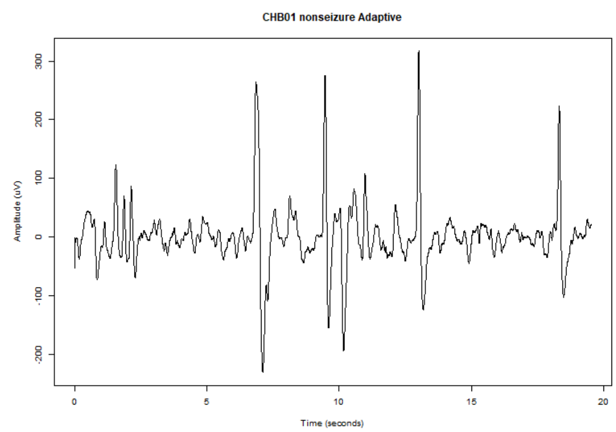
Finally, an adaptive filtering approach combining elements of both Butterworth and KZ methods was applied and visualized. Resulting graphs showed a balance between noise reduction and signal preservation, where major waveform features remained visible while smaller fluctuations were reduced. Standard deviation values for adaptive filtering ranged from 34.7-70.3 μV, corresponding to noise reductions of 9.2% - 34.0%. Peak amplitudes remained relatively preserved compared to KZ filtering, while overall variability was reduced compared to Butterworth filtering. Frequency domain plots and especially time plots for this combined method showed moderate smoothing with retained band features (see Figure 4).

<u>**CHB01:**</u>

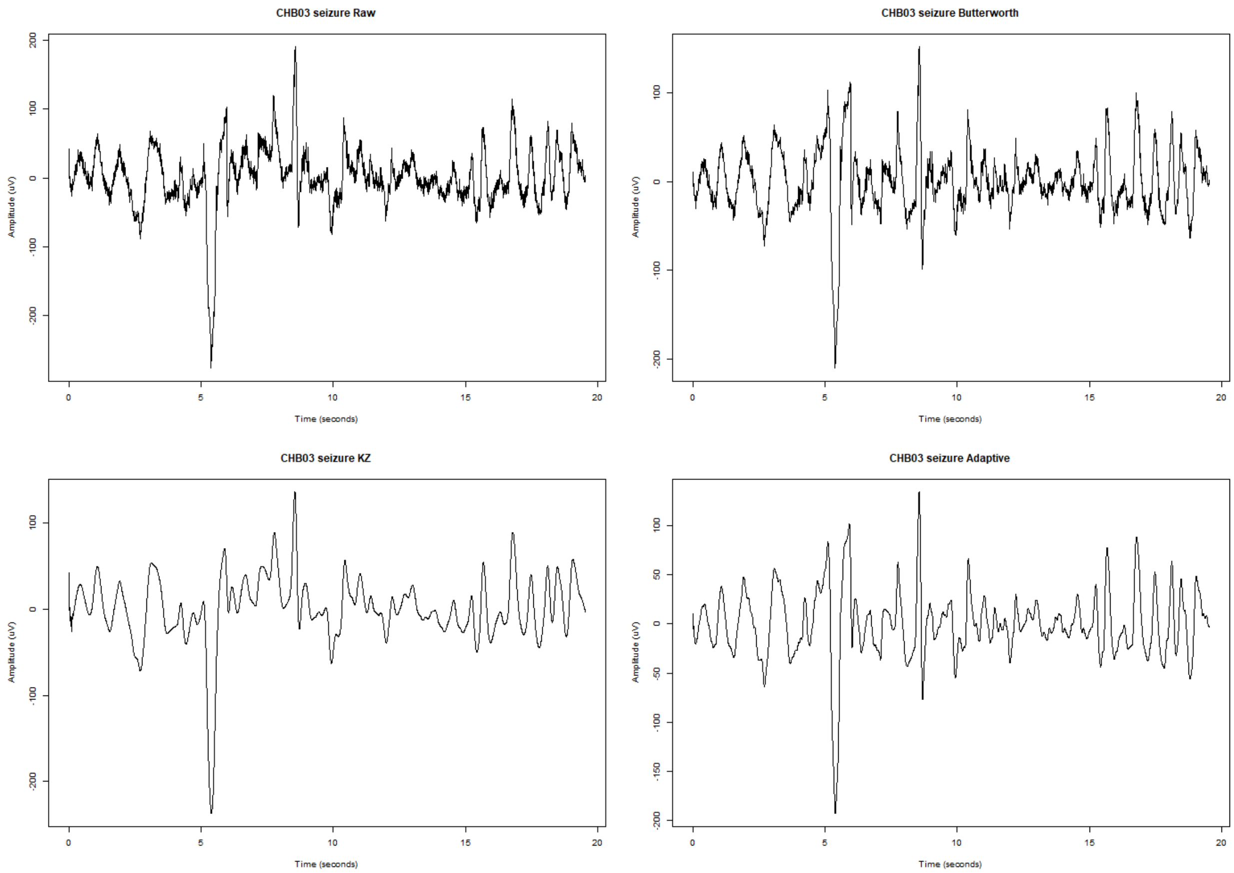

CHB01 nonseizure Raw
Amplitude (uV)
-200
0
200
400
Time (seconds)
0
5
10
15
20


CHB01 nonseizure Butterworth
Amplitude (uV)
-200
-100
0
100
200
300
400
Time (seconds)
0
5
10
15
20


CHB01 nonseizure KZ
Amplitude (uV)
-200
-100
0
100
200
Time (seconds)
0
5
10
15
20


CHB01 nonseizure Adaptive
Amplitude (uV)
-200
-100
0
100
200
300
Time (seconds)
0
5
10
15
20

CHB03:



(Figure 4a.) Additional timeplot comparison suggesting adaptive filtering behavior across patients in comparison to other filtering methods. This figure provides evidence how adaptive filtering maintains a balance between noise reduction and signal preservation when applied to different EEG recordings in comparison to both regular KZ and Butterworth applications in addition to the raw recording. Supporting comparisons are in Appendix D.

Analysis/Discussion:

Following the results observed across multiple EEG recordings, several consistent patterns emerged regarding how each filtering method affected signal clarity, variability, and frequency representation. Across all analyzed recordings, each filtering method reduced signal variability and high frequency noise, but the extent and impact of these changes varied depending on the approach used.

Butterworth filtering provided moderate noise reduction while maintaining much of the original signal structure. This was reflected in decreases in standard deviation values, typically ranging from approximately 6.5% to 16.1%. The filtered signals retained noticeable fluctuations and preserved higher frequency components compared to other methods. This suggests that Butterworth filtering is effective for reducing noise without heavily distorting the signal, although it may not sufficiently remove all unwanted variability in more complex segments.

In contrast, the KZ filter produced a much stronger smoothing effect. Standard deviation reductions ranged from approximately 12.6% to 38.1%, indicating a greater level of noise suppression compared to Butterworth filtering. Time domain graphs showed noticeably smoother waveforms with fewer abrupt changes, and frequency domain plots displayed compression of signal energy into lower frequency ranges. However, this increased smoothing was accompanied by visible reductions in peak amplitudes and less pronounced signal features, suggesting that some meaningful high frequency information may have been removed along with noise. This supports findings from Tanner (2015), who suggested that excessive filtering can distort neural patterns.

The adaptive filtering approach indicated behavior between the Butterworth and KZ methods. Noise reduction values ranged from approximately 9.2% to 34.0%, and waveform plots showed that the adaptive filter reduced variability while still preserving larger signal features more effectively than KZ filtering alone. This indicates that combining filtering techniques can provide a balance between noise reduction and signal preservation.

A pattern worth noting is that adaptive filtering underperformed most notably in recordings exhibiting high amplitude bursting activity, particularly CHB02 seizure segments where variance remained persistently elevated across the entire window. Under these conditions, the variance based switching mechanism remained locked in Butterworth-dominant mode throughout the segment, effectively reducing the contribution of KZ smoothing entirely. This suggests a fundamental limitation of adaptive approaches. In seizure segments where abnormal activity is itself high variance, the filter cannot distinguish pathological signal from noise using amplitude statistics alone. This finding implies that future adaptive frameworks should incorporate frequency domain criteria, such as spectral entropy or band specific power ratios, rather than relying solely on time domain variance to determine filter weighting. In recordings with highly irregular noise patterns, some high frequency components were still exaggerated, and minor waveform distortions were observed especially in Figures 3d and 4d (See Appendices C & D). This suggests that even adaptive approaches do not fully eliminate the trade off between noise reduction and signal preservation, but instead shift how that balance is managed.

Frequency domain analysis further supports these observations. Raw EEG signals displayed broad spectral distributions with significant power across higher frequency ranges. Butterworth filtering preserved more peaks within standard EEG bands, while reducing noise above approximately 40 - 50 Hz. In contrast, KZ filtering resulted in smoother spectral distributions with cut peaks, indicating reduced frequency specificity. Adaptive filtering maintained moderate peak distinction while reducing overall spectral variability, supporting its role as a balanced approach.

Importantly, this study differs from prior research by directly comparing multiple filtering methods under identical recording conditions and emphasizing signal interpretability rather than classification accuracy. While previous studies have implied the effectiveness of adaptive filtering, they often focus on improving within machine learning frameworks. In contrast, this study isolates preprocessing as the primary variable and evaluates how different filtering approaches alter both waveform structure and frequency characteristics. This reinforces the conclusion that preprocessing decisions are not simply preparatory steps, but fundamental factors that shape the reliability and interpretation of EEG data.

Conclusion:

Overall, the findings of this study indicate that EEG filtering methods significantly influence both the clarity and interpretability of neural signals. Butterworth filtering provided moderate noise reduction, with decreases in variability ranging from approximately 6.5% to 16.1%, while preserving much of the original waveform structure and higher frequency components. In contrast, KZ filtering achieved stronger noise suppression, with reductions ranging from approximately 12.6% to 38.1%, but this increased smoothing was accompanied by visible reduction and flattening of peak amplitudes and loss of finer signal detail, suggesting that meaningful high-frequency components were suppressed alongside background noise. The adaptive filtering approach indicated intermediate behavior, with noise reduction values ranging from approximately 9.2% to 34.0%, suggesting that combining filtering strategies can improve the balance between noise reduction and signal preservation.

These results reinforce that no single filtering method is universally optimal, and instead, filtering performance depends on the specific characteristics of the EEG signal and the goals of the analysis. While adaptive filtering provided a more balanced outcome across both time domain and frequency domain measures, it did not fully eliminate the trade off between noise suppression and signal distortion. This highlights the importance of evaluating preprocessing methods not only based on noise reduction, but also on their impact on preserving meaningful neural features. Specifically, this study suggests that noise reduction percentage is a misleading standalone metric for filter quality in EEG analysis, because a filter can achieve high noise reduction while simultaneously suppressing the diagnostically meaningful frequency peaks that make seizure activity identifiable, a tradeoff invisible to any metric that does not account for spectral preservation.

Limitations:

Several limitations should be considered when interpreting these findings. First, this study utilized a limited subset of recordings from the CHB-MIT EEG dataset, which may not fully capture the variability present across broader patient populations or clinical environments. Differences in electrode placement, recording conditions, and patient specific signal characteristics could influence how filtering methods perform in other contexts. Furthermore, this study is limited by the use of a restricted focus on a single channel, which may not fully capture variability across patients or different brain regions.

Additionally, the adaptive filtering approach implemented in this study was relatively simplified and did not incorporate more advanced techniques such as wave based adaptive filtering. As a result, the observed performance of the adaptive method may not reflect the full potential of more complex adaptive frameworks used in existing research. So, the filtering parameter selection was not fully refined which may influence the wide applicability of the observed results.

Another limitation is that this study focused primarily on signal interpretability through time domain and frequency domain analysis, rather than evaluating how filtering methods affect downstream classification accuracy in automated seizure detection models. While this approach allows for clearer insight into preprocessing effects, it does not directly assess how these filtering methods influence clinical or machine learning based detection performance.

Implications:

Despite these limitations, this study contributes to the broader field of EEG research by emphasizing the importance of preprocessing transparency and comparative evaluation. Many existing studies focus only on optimizing detection models. In contrast, the findings of this study reflect that preprocessing decisions fundamentally shape how EEG signals are interpreted, influencing both waveform structure and frequency characteristics.

By directly comparing multiple filtering methods under consistent conditions, this research supports the need for more standardized approaches to EEG preprocessing. The results suggest that adaptive or hybrid filtering strategies may offer a more effective balance between noise reduction and signal preservation, particularly in analyses where maintaining waveform detail is critical.

Furthermore, this study reinforces ongoing concerns in the literature regarding reproducibility and variability in EEG research, as highlighted by Zhang et al. (2024) and Pernet et al. (2019). By isolating preprocessing as the primary variable, this work contributes to a clearer understanding of how filtering choices impact analytical outcomes, supporting more reliable and transparent research practices.

Future Research:

Future research could also expand upon this study by applying similar comparative frameworks across a wider range of EEG datasets to even larger multi patient datasets and integrating classification models, including recordings from different patient populations and clinical settings. Additionally, incorporating more advanced adaptive filtering techniques, such as wave based methods or even AI driven parameter optimization, could provide further insight into improving signal clarity without sacrificing important features.

Further investigation is also important to examine how preprocessing decisions influence tasks such as automated seizure detection, classification accuracy, and real time monitoring systems. Integrating preprocessing evaluation with machine learning models could help determine whether improvements in signal readability translate to improved diagnostic performance with neurological conditions.

References

Allan, K. (n.d.). Evidence-based guideline: Management of an unprovoked first seizure in adults.Neurology. https://www.neurology.org/doi/10.1212/WNL.0000000000001487

Chaddad, A., Wu, Y., Kateb, R., & Bouridane, A. (2023). Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques. Sensors, 23(14), 6434. https://doi.org/10.3390/s23146434

Correa, A. G., Laciar, E., & Valentinuzzi, M. E. (2019). Adaptive filtering for epileptic event detection in the EEG. Biomedical Engineering Letters, 9(3), 309–317. https://link.springer.com/article/10.1007/s40846-019-00467-w

De Cheveigne, A., & Nelken, I. (2019). Filters: When, why, and how (not) to use them. Neuron, 102(2), 280–293. https://doi.org/10.1016/j.neuron.2019.02.039

Deng, X. (2025). A novel dual-branch network for comprehensive spatiotemporal information integration for EEG-based epileptic seizure detection. PLOS One. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0300667

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215

Li, P., Karmakar, C., Yearwood, J., Venkatesh, S., Palaniswami, M., & Liu, C. (2018). Detection of epileptic seizure based on entropy analysis of short-term EEG. PLOS One. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0193691

Oppenheim, A. V., & Schafer, R. W. (2009). Discrete-time signal processing (3rd ed.). Prentice Hall.

Pernet, C. R., Appelhoff, S., Gorgolewski, K. J., Flandin, G., Phillips, C., Delorme, A., & Gramfort, A. (2019). EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. Scientific Data, 6(103), 1–5. https://doi.org/10.1038/s41597-019-0104-8

Pernet, C. R., Calhoun, V. D., & Delorme, A. (2021). EEGLAB-LIMO EEG: A pipeline for reproducible EEG analysis. NeuroImage, 244, 118642. https://doi.org/10.1016/j.neuroimage.2021.118642

R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/

Sanei, S., & Chambers, J. A. (2007). EEG signal processing. Wiley.

Shoeb, A. (2009). Application of machine learning to epileptic seizure detection [Doctoral dissertation, Massachusetts Institute of Technology]. PhysioNet. https://physionet.org/content/chbmit/1.0.0/

Signal Developers. (2023). signal: Signal processing. R package version. https://cran.r-project.org/package=signal

Subasi, S. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert models. Expert Systems with Applications, 32(4), 1084–1093. https://doi.org/10.1016/j.eswa.2006.02.005

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. Psychophysiology, 52(8), 997–1009. https://doi.org/10.1111/psyp.12437

Toffolo, G., et al. (2019). edfReader: Reading EDF EEG files in R. CRAN. https://cran.r-project.org/package=edfReader

Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Transactions on Audio and Electroacoustics, 15(2), 70–73. https://doi.org/10.1109/TAU.1967.1161901

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data – A practical approach. Journal of Neuroscience Methods, 250, 34–46. https://doi.org/10.1016/j.jneumeth.2014.08.002

Wong, S., Leung, H., & Chan, K. (2023). EEG datasets for seizure detection and prediction – A review. Sensors, 23(10), 4567. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10235576/

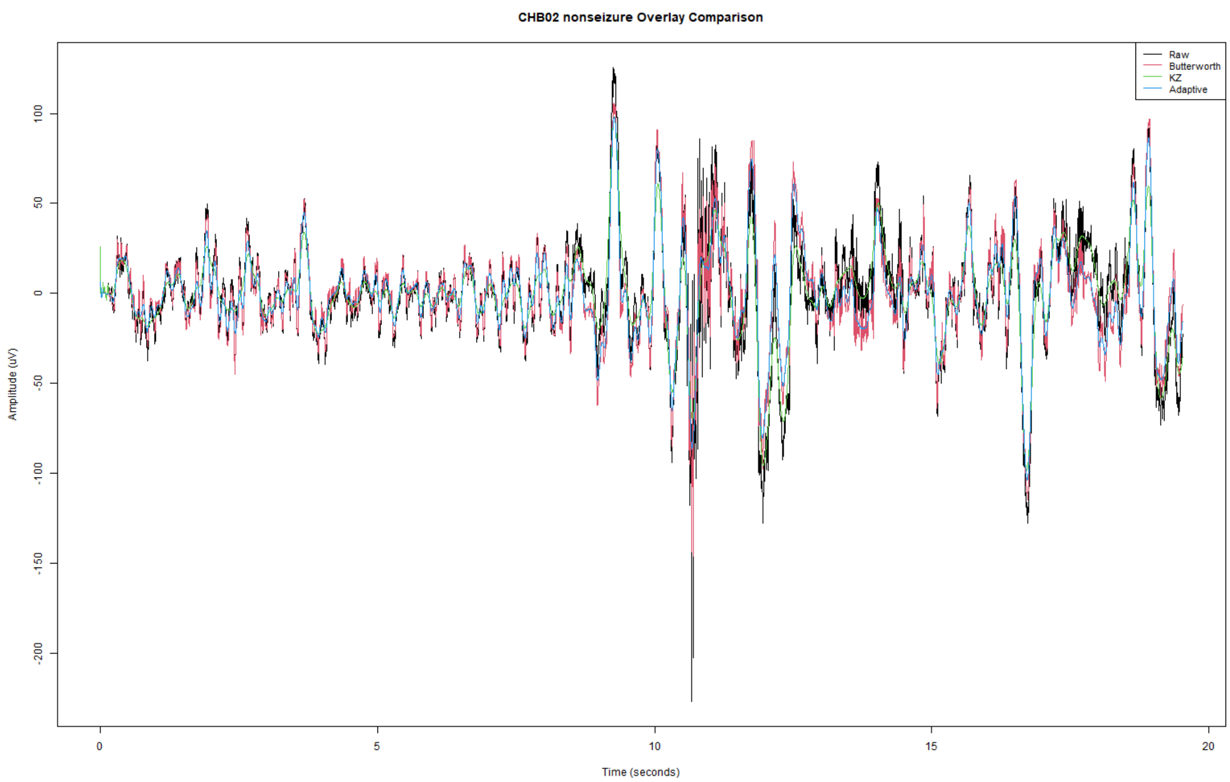
Appendices

Appendix A:

- **Non-Seizure Time Domain**

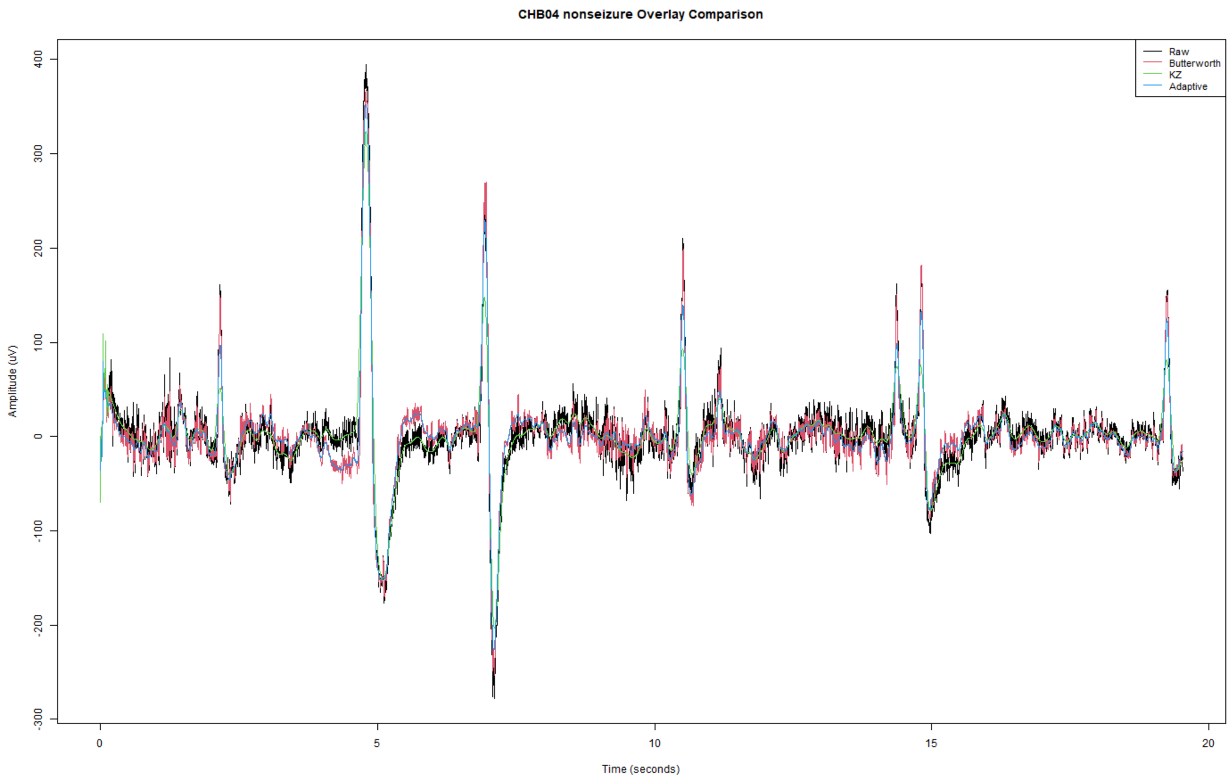
(Figure 1c.)

(CHB02)



(Figure 1d.)

(CHB04)

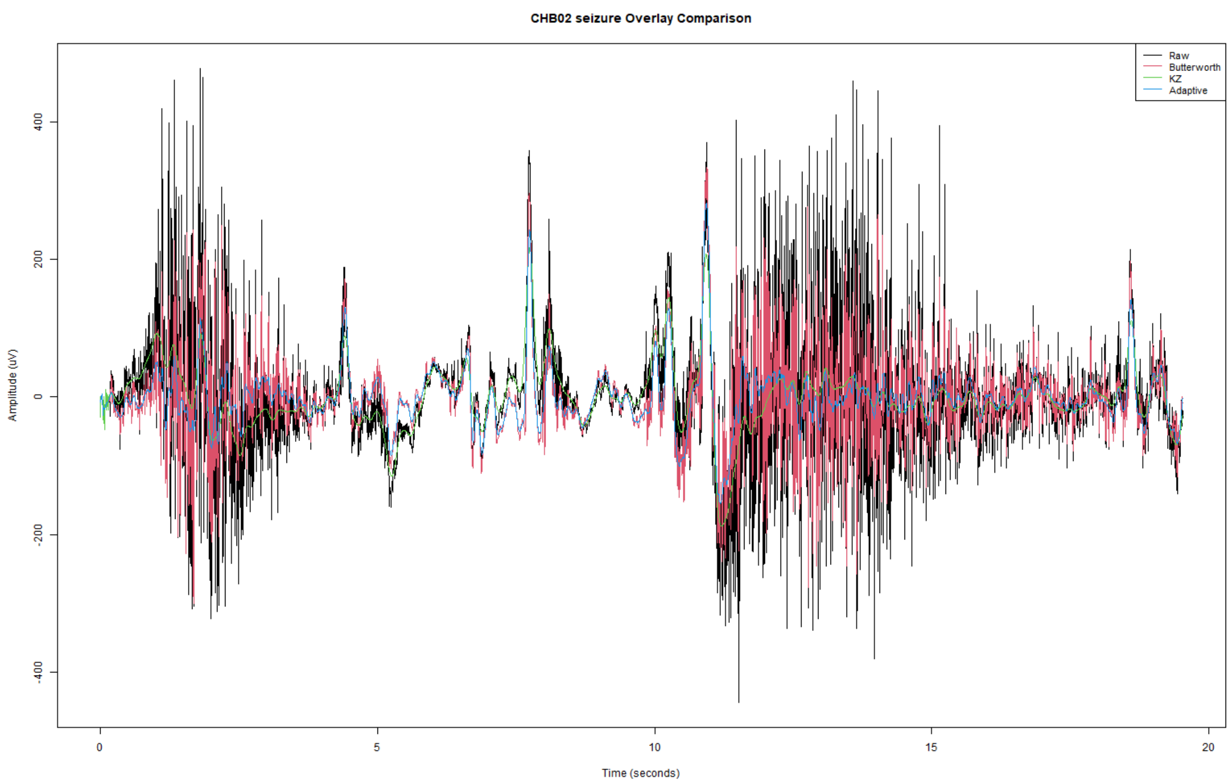


<u>Appendix B:</u>

- **Seizure Time Domain**

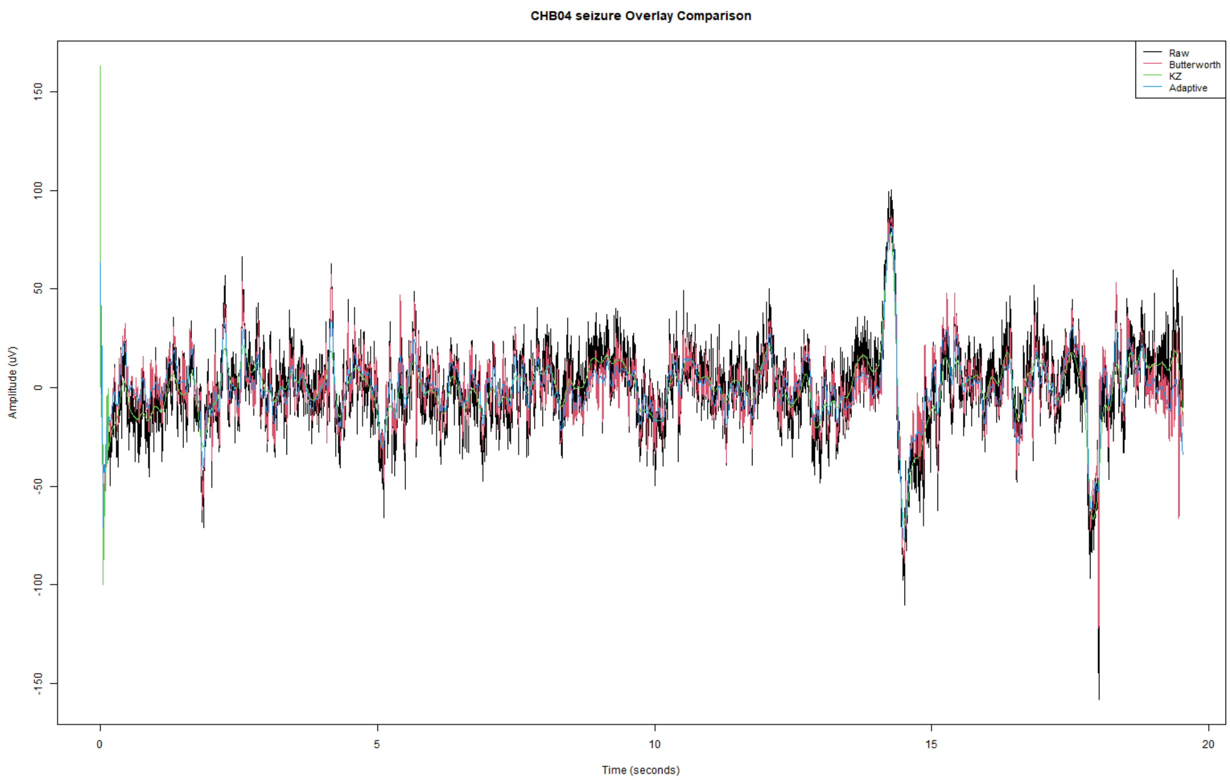
<u>(Figure 2c.)</u>

(CHB02)



(Figure 2d.)

(CHB04)

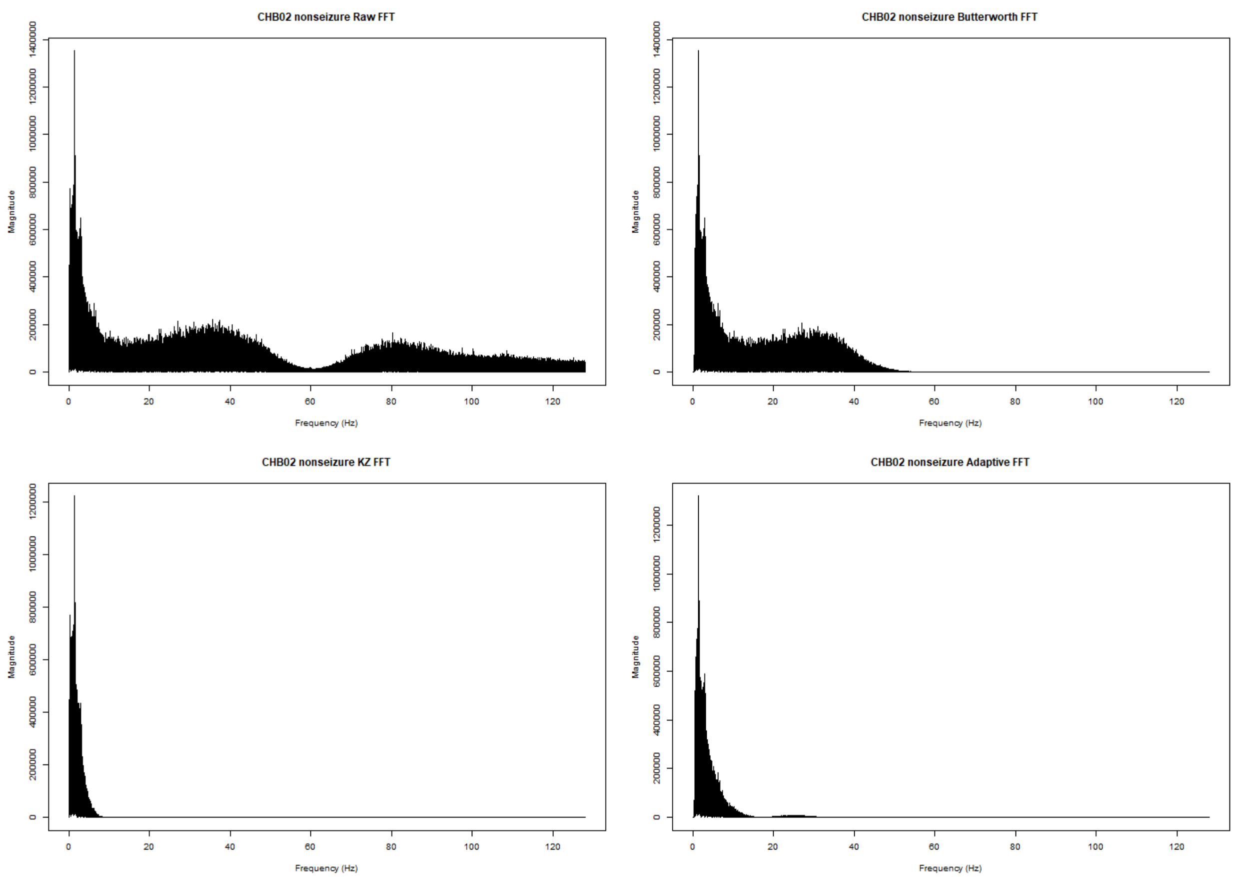


<u>Appendix C:</u>

- **FFT (Frequency Domain)(Non-Seizure and Seizure)**

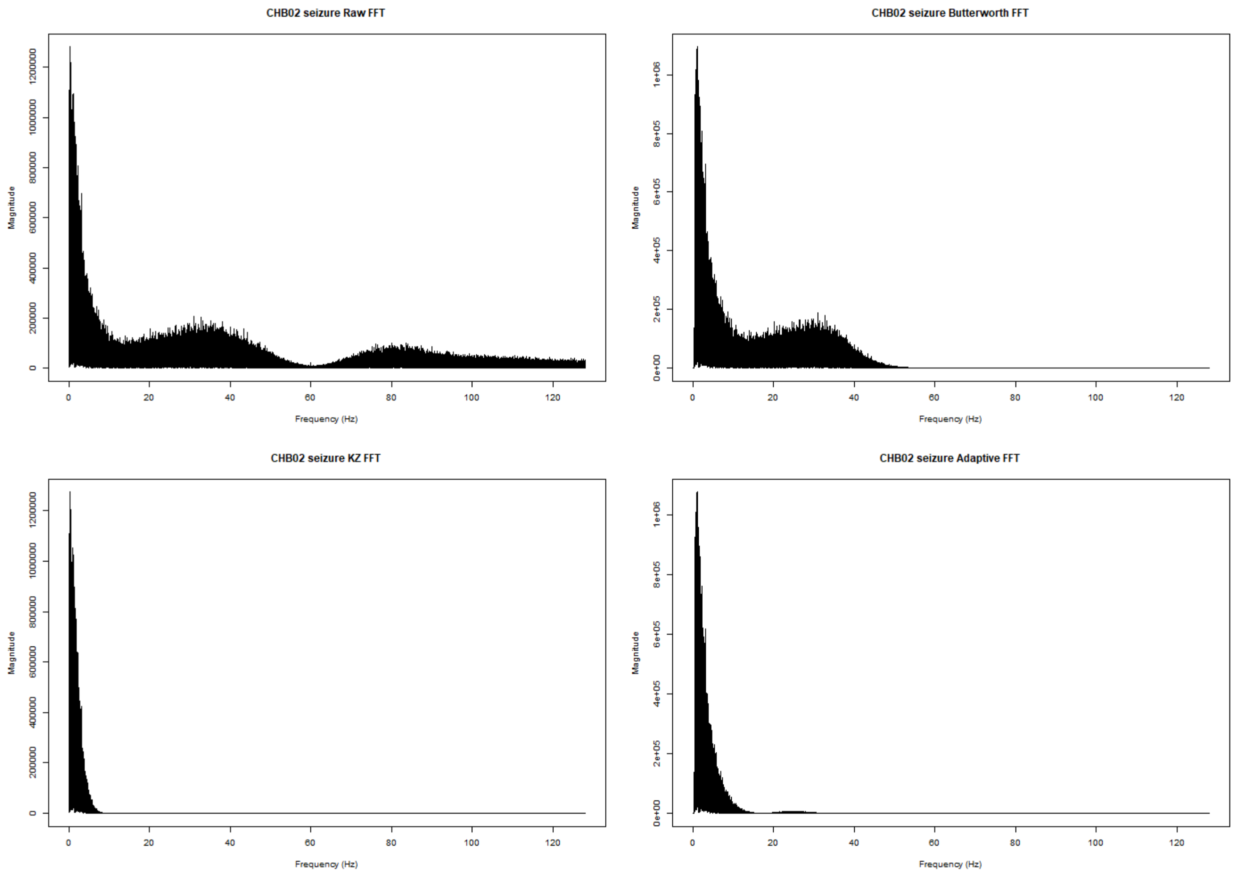
<u>(Figure 3c.)</u>

(CHB02)(Non-Seizure)



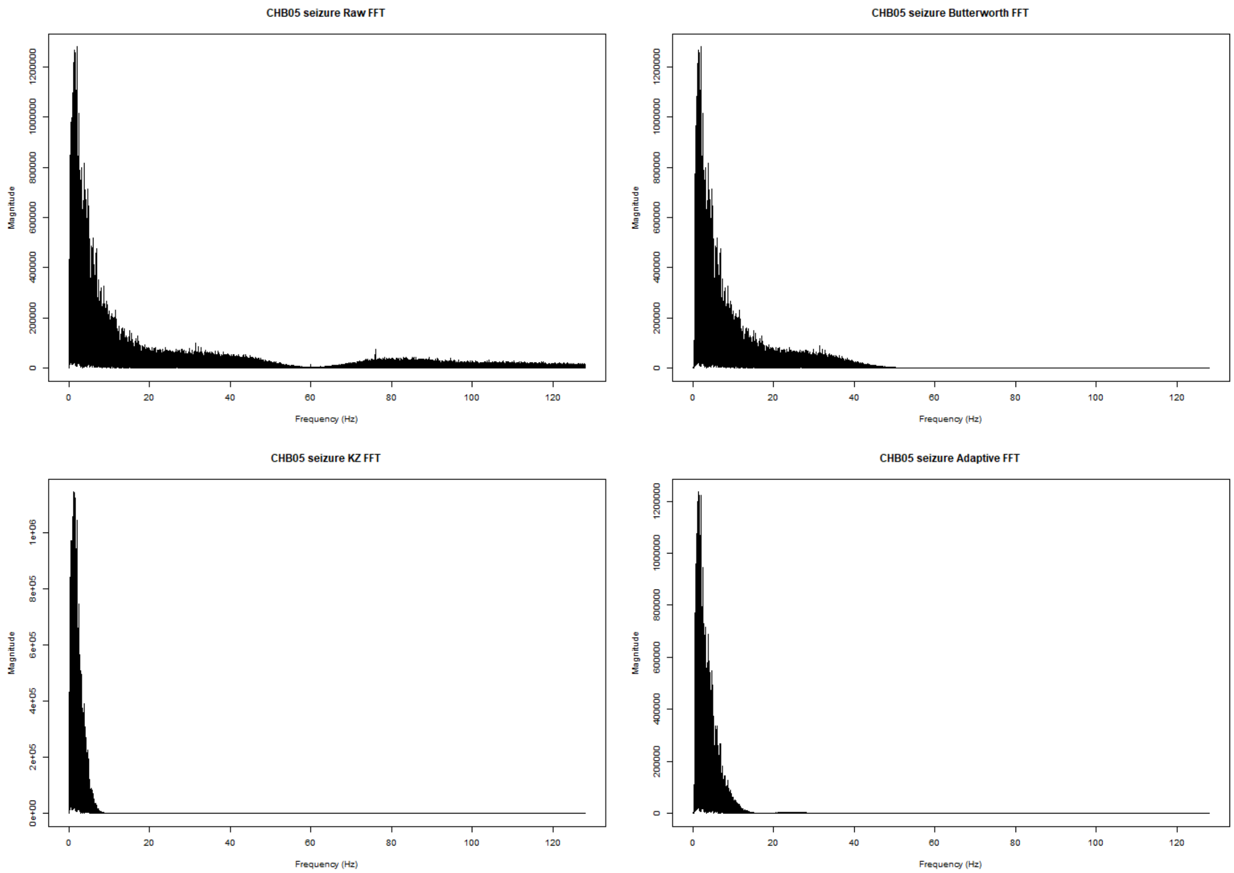
(Figure 3d.)

(CHB02)(Seizure)



(Figure 3e.)

(CHB05)(Seizure)

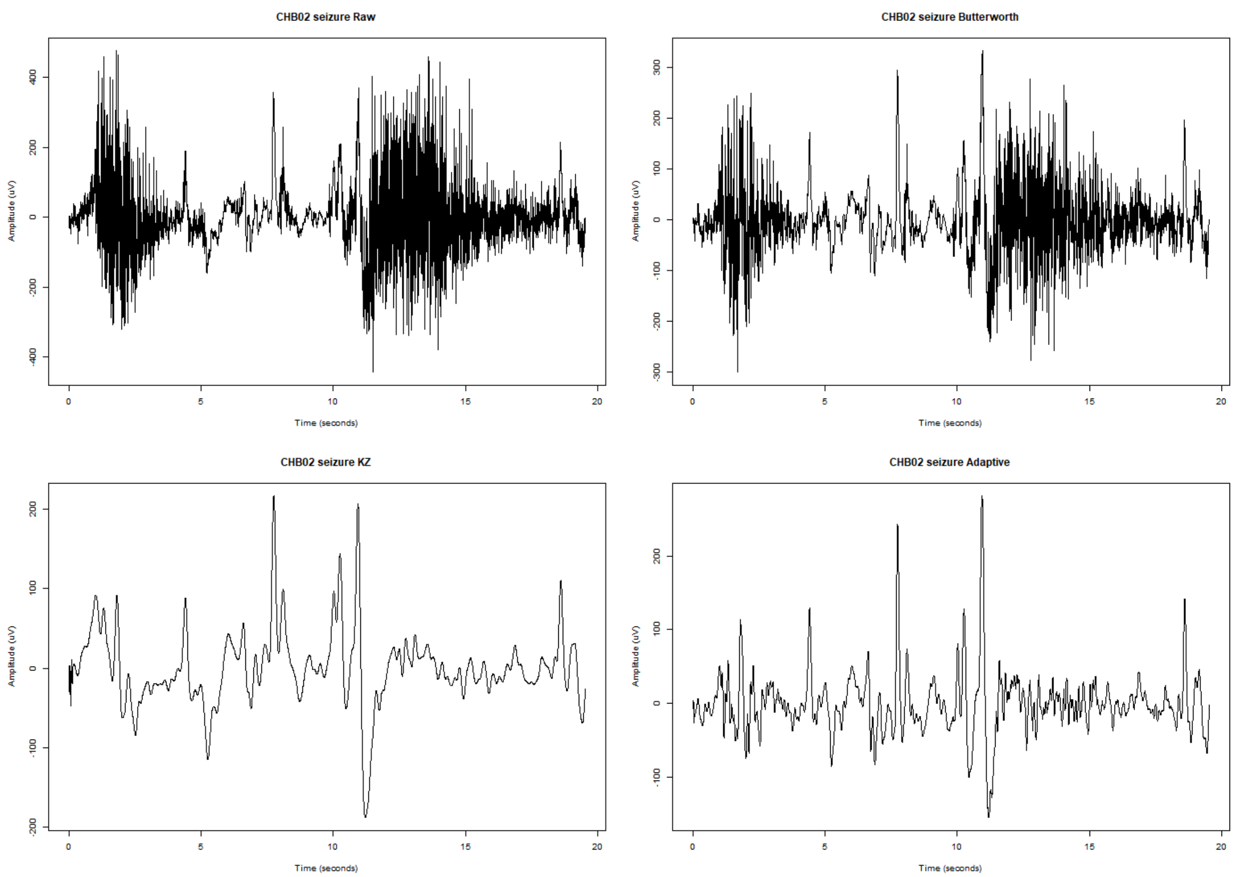


Appendix D:

- **Adaptive Filter Comparisons**

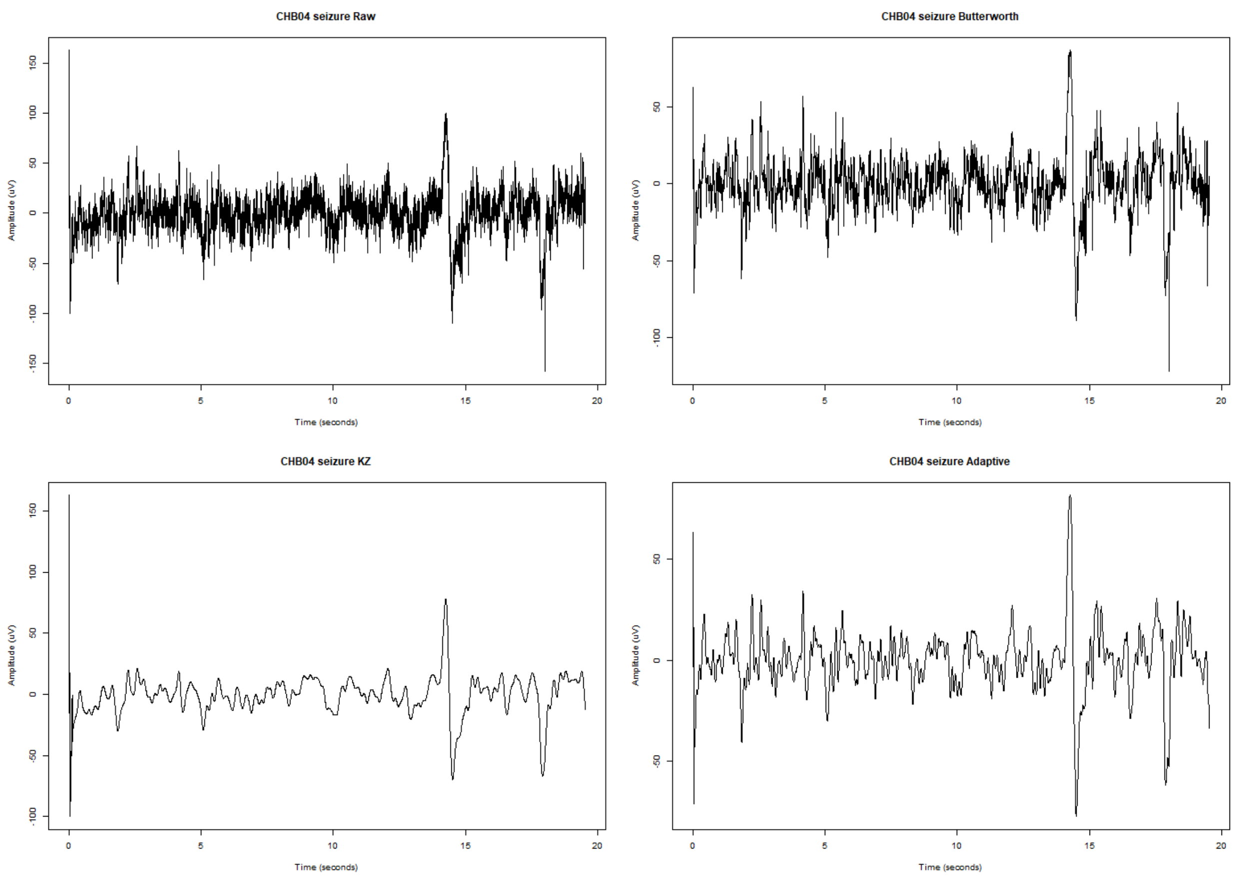
(Figure 4b.)

(CHB02)(Seizure Timeplots)



(Figure 4c.)

(CHB04)(Seizure Timeplots)



Appendix E:

- **(Table E1. Summary of Signal Variability and Noise Reduction Across Filtering Methods)**

Filter Type	**Standard Deviation Range (µV)**	**Noise Reduction (%)**	**Key Effect**
Raw	40.7 - 88.0	0%	High variability & noise present
Butterworth	32 - 76	6.5% - 16.1%	Preserves structure, maintains fluctuations
KZ Filter	25.3 - 70.3	12.6% - 38.1%	Moderate smoothing, reduced high frequency detail
Adaptive	28 - 72	9.2% - 34.0%	Balanced noise reduction and preservation

- **(Table E2. Frequency Band Power Distribution Across EEG Signals)**

Frequency Band	**Range (Hz)**	**Power Range (µV²)**	**Interpretation**
Delta	0.5 – 4	$1.59\times10^{8} – 2.30\times10^{9}$	Dominant band across recordings
Theta	4 – 8	$1.31\times10^{6} – 2.39\times10^{8}$	Elevated during seizure segments
Alpha	8 – 13	$1.75\times10^{7} – 5.49\times10^{7}$	Moderate presence
Beta	13 – 30	$2.49\times10^{6} – 2.07\times10^{8}$	Reduced after filtering
Gamma	30+	$2.55\times10^{5} – 1.96\times10^{8}$	Mostly suppressed after filtering

Appendix F:

- **Butterworth Filter Equation**

The Butterworth filter is known for having an effect that is as flat as possible in the passband and rolls off towards zero in the stopband.

- **Formula:**

$$|H(j\omega)| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}}}$$

Where:

-f = signal frequency (Hz)

-fc = cutoff frequency (Hz)

-n = filter order (controls sharpness of cutoff)

- **Kolmogorov-Zurbenko (KZ) Filter Equation**

The KZ filter is defined as a simple moving average, which provides a flexible way to filter data.

- **Formula:**

$$Y_t = \frac{1}{m} \sum_{i=-k}^{k} X_{t+i}$$

Where:

-x = input signal (time series data)

-m = window size (number of data points averaged per iteration)

-k = number of iterations (how many times the moving average is applied)

Comparison

- Butterworth: Ideal for scenarios needing sharp cutoff frequencies and maximum flatness.
- KZ Filter: Ideal for separating short-term fluctuations from long-term trends in noisy data, often used in environmental data analysis.