4-29-2015

Early Detection of Seizure With a Sequential Analysis Approach

Xin Wan
University of Pennsylvania

Early Detection of Seizure With a Sequential Analysis Approach

Keywords
epilespy

Disciplines
Business

Early detection of seizure with a sequential analysis approach

Xin Wan

April 29, 2015

Advisor: Professor Abba Krieger

I. Introduction

 A. Epilepsy

Epilepsy has been one of the most prevalent diseases worldwide. It is the second most common neurological disorder after stroke, effecting roughly 0.6-0.8% of the global population (Hauser et al., 1996). Therefore, controlling major negative consequences of epilepsy is of great interest and importance to many researchers and clinicians.

Traditionally, there are two main ways of controlling epilepsy: anti-convulsive medication and resective surgery. However, these methods do not provide a complete solution to the problem. For about 25% of epilepsy patients, seizures cannot be controlled by these traditional methods. (Mormann et al., 2007) This shortcoming has motivated new research into trying to find alternative methods. One such approach, which has gained momentum recently, is through implantable devices. Micro-devices are implanted in the patient's brain with the aim of predicting and preventing seizures through intervening by releasing a tiny dose of medicine or a tiny electric shock. As the New York Times article documented, this method functions as a "pacemaker for the brain" – "…research has shown that seizures start with a tiny spark of activity and that they take hours to build to surge". This approach is motivated by the hope that characteristic features can be extracted from continuous EEG recordings and can be used to predict seizure. (Mormann et al., 2007)

The algorithm with which seizures can be predicted from EEG data is the main concern of this paper. We test a seizure prediction algorithm developed by Krieger and Pollak on a period of EEG recording of a dog which had a seizure during the recording, and found that the algorithm is promising because it detects a clear peak about 2.5 hours prior to seizure. We further tested this algorithm on six other recordings from dogs without seizure, and found mixed results. In some datasets there are peaks of about the same magnitude as when there are seizures thereby producing false positives. A more detailed description of the data can be found in section II. A description of the algorithm is presented in section III.

 B. Detecting Algorithms

Historically, many different algorithms have been proposed by scientists in order to address this issue. Mormann et al. (2007) provides us with a good review of the development of these algorithms: in early stages, focus was directed to the pre-ictal period, the period that is known to be followed by an onset. For example, Rogowski et al. (1981) and Salant et al. (1998) used autoregressive modelling and identified pre-ictal changes in the modelled parameters up to 6

seconds prior to seizure onset; Le Van Quyen et al. (1999, 2000, 2001a) found that dynamical similarity decreases before seizures in both intracranial and scalp EEG recordings. However, these studies focus on the pre-ictal period and thus lack specificity – we never know when we are in the pre-ictal period in advance.

Several studies were thereafter conducted to address this issue. Cohen et al. (2002) showed from selected examples of five patients that, before seizures, similarity measure drops more frequently than during inter-ictal periods. Mormann et al. (2000) found that some changes in phase synchronization between distinct brain areas before seizures cannot be found in exemplary seizure-free recordings.

While these studies were promising, skepticism rose soon afterwards when a number of studies found that the previous results were not reproducible. For example, studies done by De Clercq et al. (2003) and Winterhalder et al. (2003) questioned the reliability of the results reported for the similarity index (Le Van Quyen et al., 2001a). Many other previous results (Litt et al., 2001; Lehnertz and Elger, 1998; Iasemidis et al., 1990; Martinerie et al., 1998) were challenged or found to be not reproducible. (Maiwald et al., 2004; Harrison et al., 2005a; Aschenbrenner-Scheibe et al., 2003; Harrison et al., 2005b; McSharry et al., 2003). Collectively, these studies pointed to the problem of applying highly optimized algorithms to small, selected data sets, because the results cannot be reproduced on unselected, larger dataset, which is the real challenge that we need to overcome in order to predict seizure in real-time from continuous EEG recordings.

As technologies advanced at the turn of the millennium and mass storage capacity became available, testing on complete, unselected pre-surgical monitoring datasets became possible. In 2005, the First International Collaborative Workshop on Seizure Prediction (Lehnertz and Litt, 2005) held in 2002 yielded a series of studies. D'Alessandro et al. (2005), Esteller et al., (2005), Harrison et al. (2005a) and Mormann et al. (2005) all showed a poor performance of univariate measures. On the other hand, Iasemidis et al. (2005), Le Van Quyen et al. (2005) and Mormann et al. (2005) showed better performances of bi- and multi-variate measures. However, the studies up to this point were still retrospective in nature. In real life, an algorithm can only be clinically beneficial if it can raise a warning prior to seizure, rather than give a summary afterwards. Hence, seizure-prediction algorithms in a prospective manner are the necessary next step.

Several studies have been conducted to test various prospective prediction algorithms. However, some are not sensitive or specific enough for clinical implementation (Iasemidis et al., 2003; D'Alessandro et al., 2005) while others are inconclusive (Mormann et al., 2006b; Chaovalitwongse et al., 2006; see also Winterhalder et al., 2006).

A special motivation of our research is a paper by Pearce et al. (2013). High Frequency Oscillations (HFO) data from patients were divided into four time epochs and three clusters (ripples, fast ripples, and mixed events). This study observed patient-specific changes in relative rates of ripples, fast ripples, and mixed frequency events. These changes in relative rate occurred in pre- and post-ictal periods up to thirty minutes before and after seizures. Evidence also suggested that the distribution of HFOs during these different time periods varied greatly

between individual patients. Hence, we know that at least in retrospect, EEG data behaves differently in time periods that are close to seizure relative to time periods long before seizure. This provided hope that algorithms to predict seizures are   plausible. The question is then how to create these algorithms prospectively.

Our research tests a new algorithm which predicts seizure prospectively from unselected, continuous EEG recordings. A detailed description of the algorithm can be found in Section III.

II.   Data

The data for this research are EEG brainwave filtered at certain frequencies. The EEG is usually recorded by electrodes placed on a patient's scalp; in our experiment, the EEG is recorded from dogs. These data are collected over time. There are 16 electrodes placed at different locations in the brain (16 different channels). Measurements are recorded every ten seconds hence we have time-spatial data of fine granularity. Features of the data are extracted (e.g., the amplitude of the brainwave or the power of the spectrum) – at each location, in our data we extract four features, namely the power of the spectrum filtered and four different frequencies. The technique of extracting features from raw recordings is out of the scope of this paper. Therefore we will take the extracted features as given. Hence, at any given point in time, we have 64 data points, coming from 16 channels each having 4 features (frequencies). The graph below illustrates the process of EEG recording, with a seizure illustrated in plotted EEG in Figure 1B.
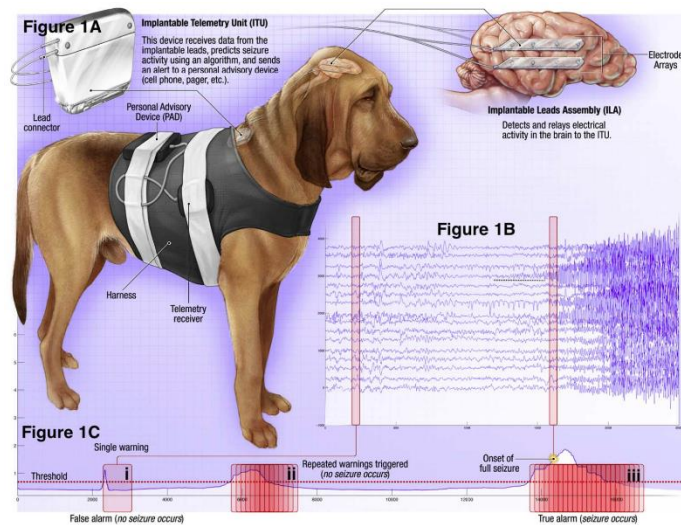
Source: Litt Lab, University of Pennsylvania

Specifically, we test our algorithm on 7 different datasets. Each dataset has 64 columns. The first 4 columns are channel 1, feature 1-4 in order, and columns 5-8 are channel 2, feature 1-4 in order, and so on. Column 61-64 would then be channel 16, feature 1-4 in order.

The datasets varies in numbers of rows. Each rows represents a point in time – the sequences are recorded at intervals of 10 seconds. Therefore, if Row 1 is time 0, then Row 2 is 10 seconds after, and so on. Hence the length of the recording would be 10 times the number of rows in seconds. For example, Row 15 Column 12 of the first dataset would be feature 4 of channel 3 at 150 seconds after time 0.

The first dataset is 1620 by 64, with a seizure occurring at observation 1441 (row 1441). The rest of the datasets are 7200 by 64 each, recorded from another dog without the occurrence of seizure during recording. We will test the algorithm on the dataset with seizure to assess its predictive power, and then test it on the other datasets to assess the possibility of false positives (alarms not followed by an actual seizure).

III.   Algorithm

Our algorithm is inspired by the sequential analysis method commonly used in statistical process control. The main tool is a Shiryaev-Roberts statistic, calculated from the sequence of EEG recordings. The statistic, generally speaking, gives us the likelihood that a change point has occurred prior to a given point in time. The change can be in terms of mean or variance. Specific to our analysis, we only test the change of mean. Our null hypothesis would be that, under normal conditions, when there is no seizure, the distribution of EEG data should be close to identically distributed, or at least somehow consistent in time. Thus, when we identify a change in one of the parameters (the mean in our case), we can raise an alarm that a seizure might be imminent.

For any one of the dataset we have, we use the following algorithm to process and obtain the series of Shiryaev-Roberts statistics over time:

1. First we take log of the data. If any row contains any missing value or NA after logging, we throw out the entire row. Data should be positive and thus logged data should have valid values. Only two rows are missing in the dataset with seizure;
2. Then we arrange logged data by feature, and take the first principal component of each feature (logged) over the sixteen channels. For convenience, we take the mean of the 16 channels for each feature as a surrogate for the first principal component. Therefore, we end up having four time series, one sequence for each feature. To illustrate this, for each row, we take the mean of column 1, 5, 9, 13…61 as the first principal component of feature 1, and that of 2, 6, 14… 62 for feature 2, etc. Hence for any point in time, we have four data points, one for each feature. We plot the first principal component for each feature.
3. We run an AR-3 process on a period of each of the sequences to train our model. The length of the training period can be adjusted; in our model, we run the AR-3 on the first 720 data points. Coefficients of the fitted model is saved. It has been found empirically that these pre-preparation processes organize our data so that they are close to normal IID.
4. We use the fitted coefficients to predict the rest of the observations, and obtain the residual at each point in time. For example, in the first dataset, we will predict the next 878 observations from the 721st time point onwards, and subtract the predicted values from the actual values to obtain the residuals. For the first 720 observations, we can get residuals from our fit. This process is separately applied to each feature's sequence. Therefore, we obtain four sequences of residuals, one for each feature. We eliminate the first three points in time, as our AR process has a lag of 3. We plot the residuals here.

5. We apply the statistic to the sequences of residuals, we plot the sequence of statistics in time. The formula for the statistic is:

$$R_n = \sum_{k=b}^{n} \prod_{g \in G} \Omega(k, n, g),$$

where b = 1 as a default.

Hence, R_n will be a sequence in time. G is the set of features, and in our case, we calculate R_n for each feature, so there is no multiplication to be made. Ω values can be calculated with the following formula:

$$\Omega(k, n, g) = \frac{1}{\eta^{n-k+1}} \left(\frac{\sum_{i=1}^{n} Z_{ig}^2}{\sum_{i=1}^{k-1} Z_{ig}^2 + \frac{1}{\eta^2} \sum_{i=k}^{n} Z_{ig}^2} \right)^{\frac{n}{2}},$$

where Z_{ig} denotes the ith residual of feature g (g can be 1, 2, 3, or 4), and $1/\eta^2$ is taken to be ½ in our case. Often, a prior is placed on $1/\eta^2$, often taken to be $\Gamma(\alpha, \beta)$, with typical values of the parameters of $\alpha = 10$ and $\beta = 20$.

6. The series of Rn is eventually plotted as the output of this algorithm. The hope is that when a seizure is about to occur, the Rn values will show an abnormal peak that can be distinguished from normal conditions.

This algorithm, if proven to be significantly better than chance, will be an advancement over existing algorithms. Its main advantage is it predicts a seizure in a prospective manner, and we can thus build a close-loop device based on it. More detailed discussion regarding the methods of implantable devices can be found in Section V.

IV.   Results

The algorithm is first applied to the dataset with seizure. First we show the first principal components of each feature.

means for the 3's feature

means for the 4's feature
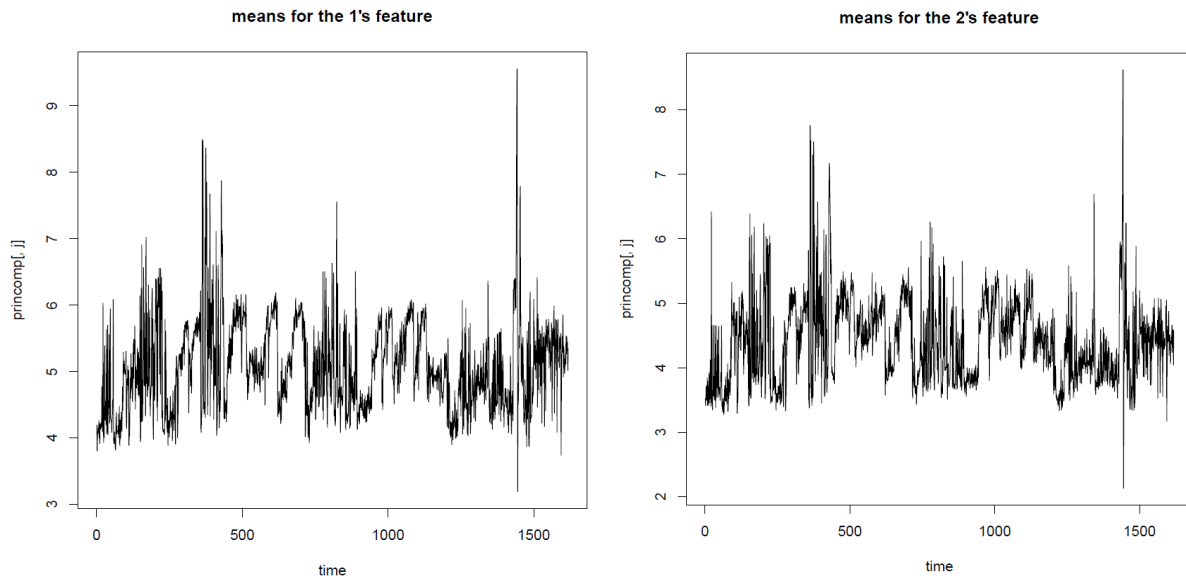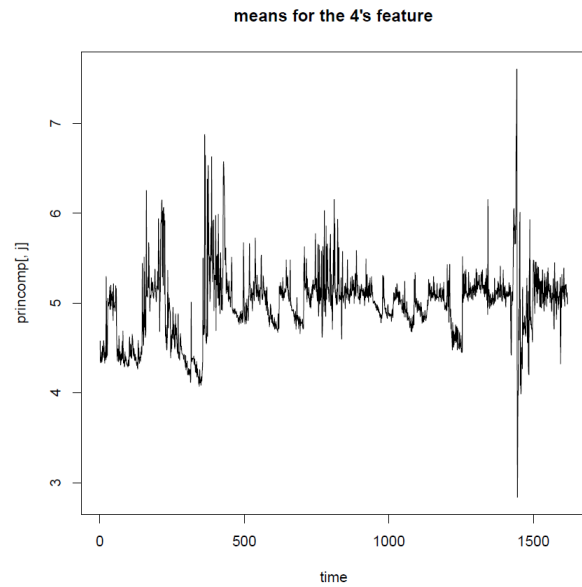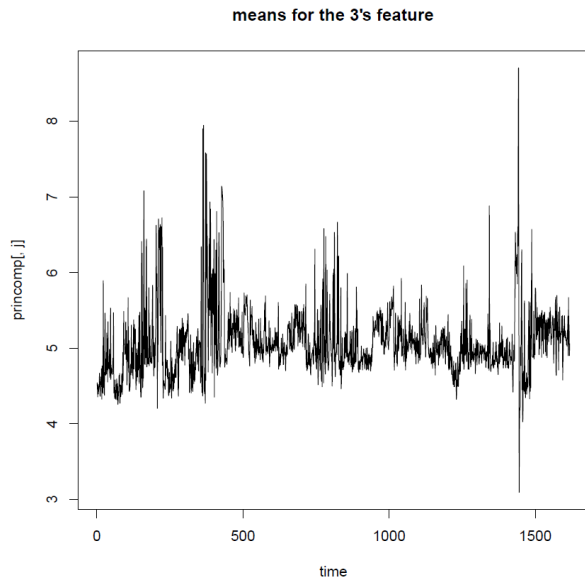
It is very hard to tell any pattern from these graphs, as the data are very noisy. It would not be reasonable to base prediction decisions on these data. We need something more sophisticated. Hence, we apply the AR-3 process to these data and obtain the residuals for them.

residuals for the 1 's feature

residuals for the 2 's feature

residuals for the 3 's feature   residuals for the 4 's feature
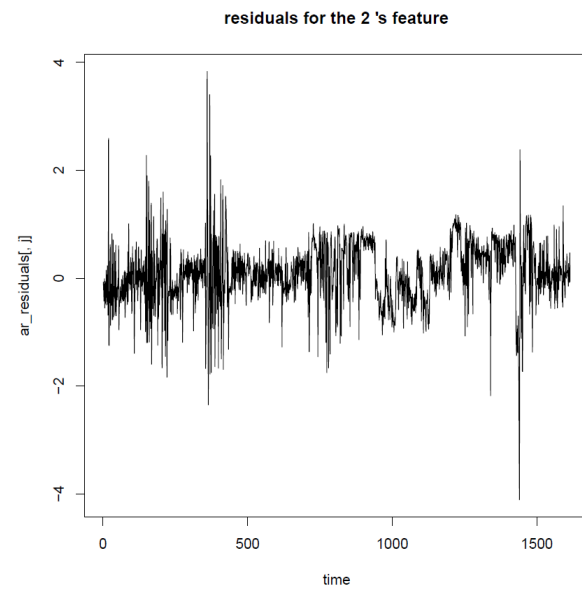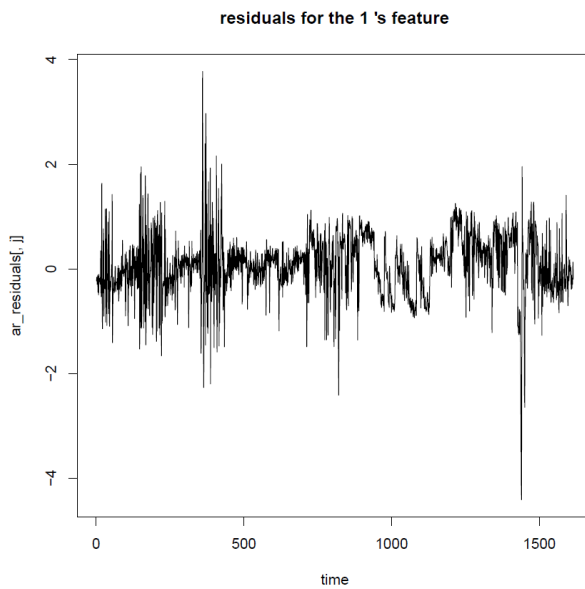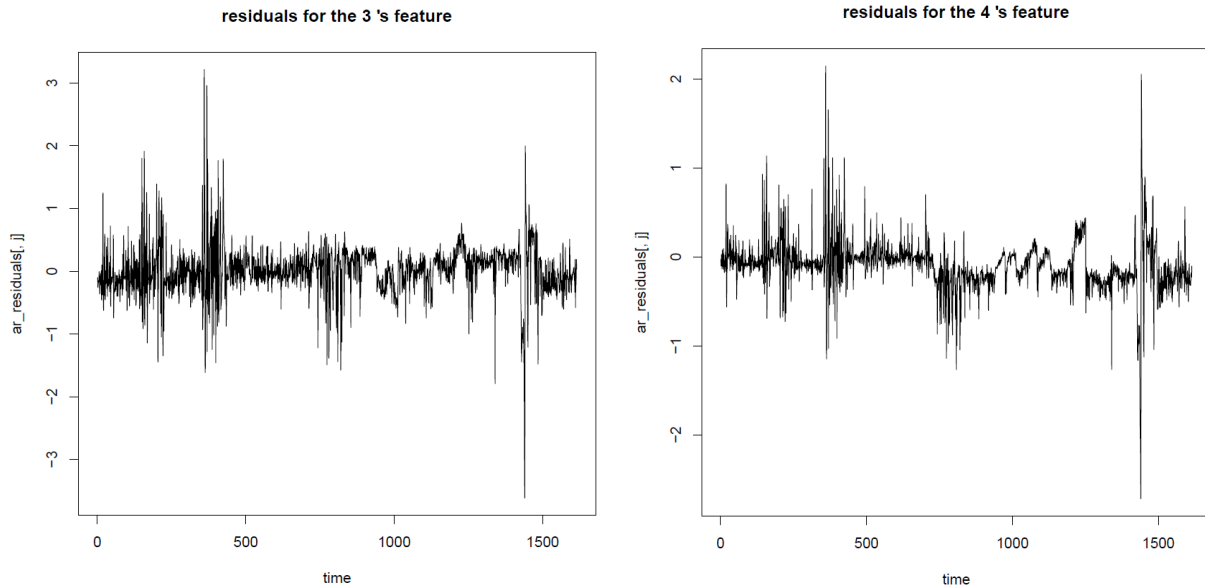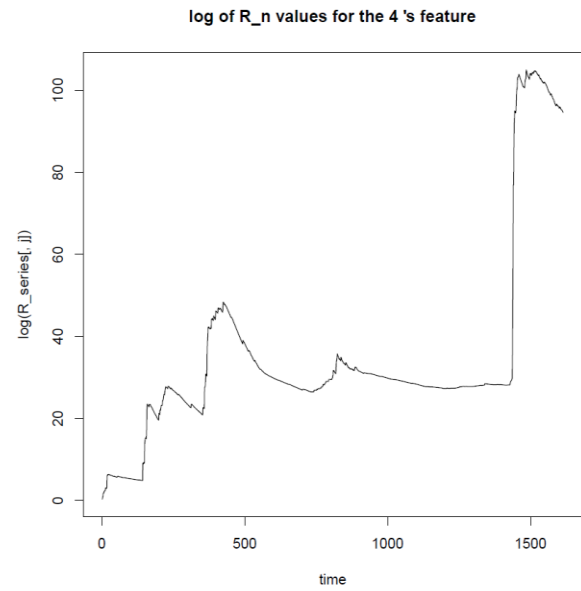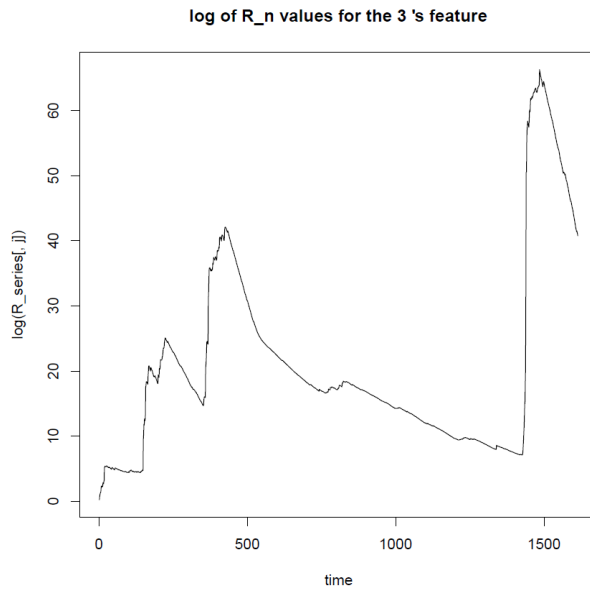
Compared to the first principal components, the residual values demonstrate smaller variances. However although slightly better, the residuals are still noisy enough to prevent reasonable prediction of seizure. Hence, we need a mechanism to smooth out the "normal" part of the data, and only amplify the changing part of it. Also, we need an approach that will show small changes overtime.

The Shiryaev-Roberts (SR) statistic provides us with such a tool (Pollak, 2009). It is designed to detect changepoints in a sequence of observations whose baseline distribution might change to a new one, once the process goes "out of control". The formula for the statistic is introduced in Section III, step 5. Due to limited scope of this research, the details of this statistic will not be discussed. It suffices to know that our specific formula is designed to find a potential change in the mean of the underlying distribution. Below is the log of the series of SR statistics we obtained from the residuals for each feature.

log of R_n values for the 1 's feature

log of R_n values for the 2 's feature

log of R_n values for the 3 's feature

log of R_n values for the 4 's feature
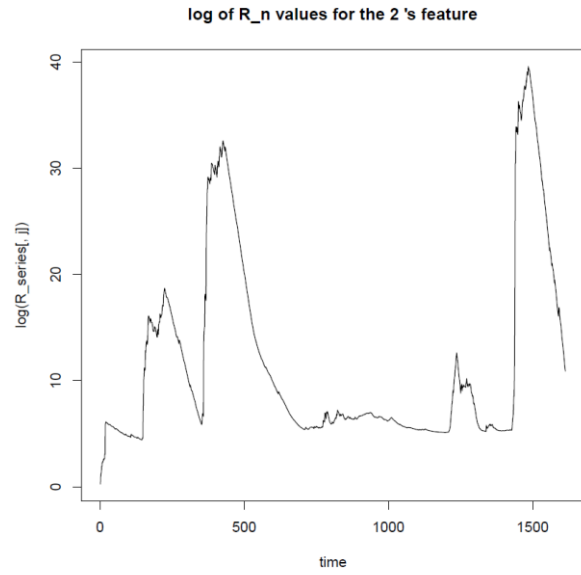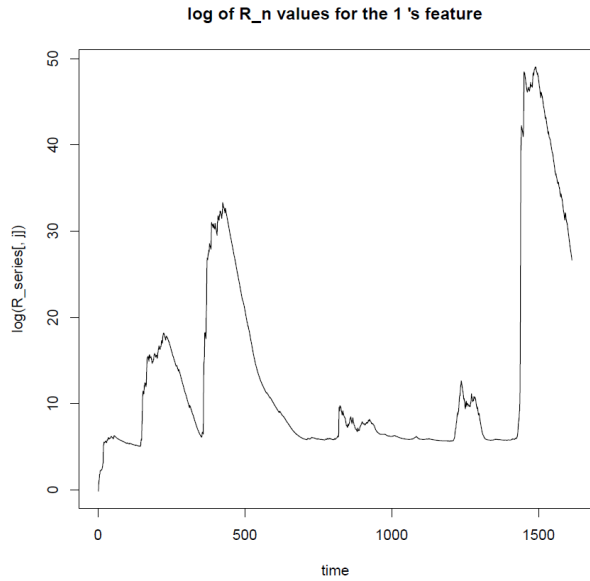
As we can see clearly from the data, there is a high peak around the time of seizure onset – at observation 1440 in the sequence. More importantly, there is another peak that occurs in all four features prior to the seizure, around observation 440 in the sequence. This can be seen as a potential precursor to seizure, although much more testing is needed. If this is indeed a signal of immenent seizure, then we would be able to raise an alarm 1000 observations in advance, which is almost 3 hours in advance. The algorithm can then be programmed into a chip to intervene and hence prevent the subsequent seizure.

log of first 1020 R_n values for the 1 's feature

log of first 1020 R_n values for the 2 's feature

log of first 1020 R_n values for the 3 's feature

## log of first 1020 R_n values for the 4 's feature
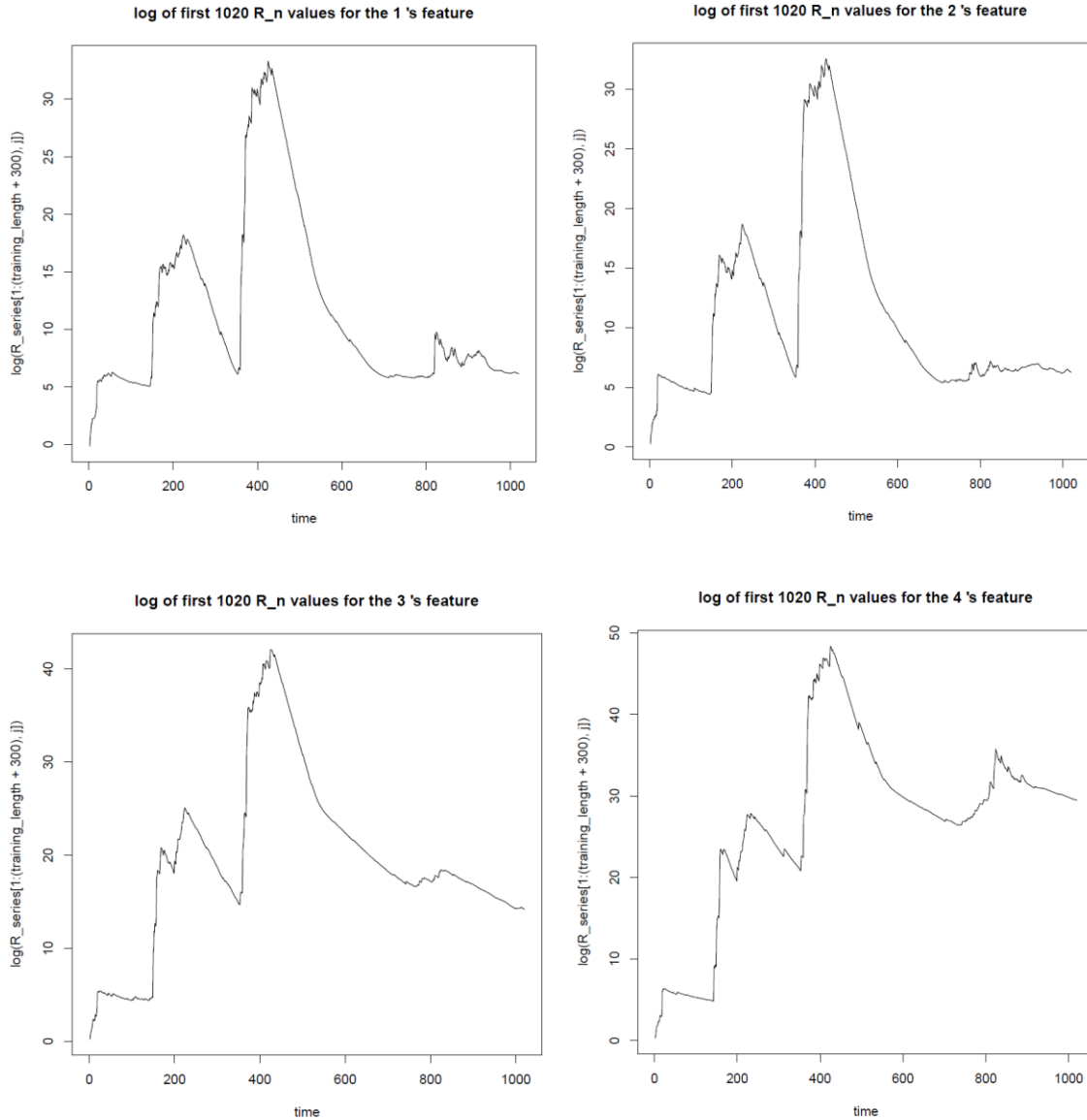
Here we attached the same sequences, although we focus on the first 1020 values for each sequence, which is 300 terms after the training period, so we can better see the details. It can be seen that there is no immediate peak occurring at observation 721, signaling that the training length might be sufficiently long. However, there is a consistent but small peak around 800. This might be due to low out-of-sample predicting power of the fitted model, or specific to the dataset. We need further testing on other datasets to decide.

After applying the algorithm to the dataset with seizure, we apply the same algorithm to datasets without seizures. The goal is to compare the performance of this algorithm on a control sample to investigate the probability of false positives – the probability of raising false alarms when no seizure is going to occur. Appendix II has graphs of the log of SR statistics from six different EEG recordings without seizure occurrences. Here we only include the following two for illustration purposes.

The graph to the left is the log of SR statistics for the 3rd feature in the second control data sample. As we can see, during the 7200 recorded data (about 20 hours of recording), the series remained below 20. If we refer back to the series from the dataset with seizure, we could have set a threshold at 30, and therefore this algorithm would give an alarm when there will be seizure, and not give an alarm according to the control dataset.

The graph to the right illustrates an occasion of false positives – the log of SR statistics goes all the way to 200, even higher than the one we calculated for the actual seizure onset – this will certainly lead to an alarm being raised, but we know that no seizure symptoms occurred in the actual clinical observations.
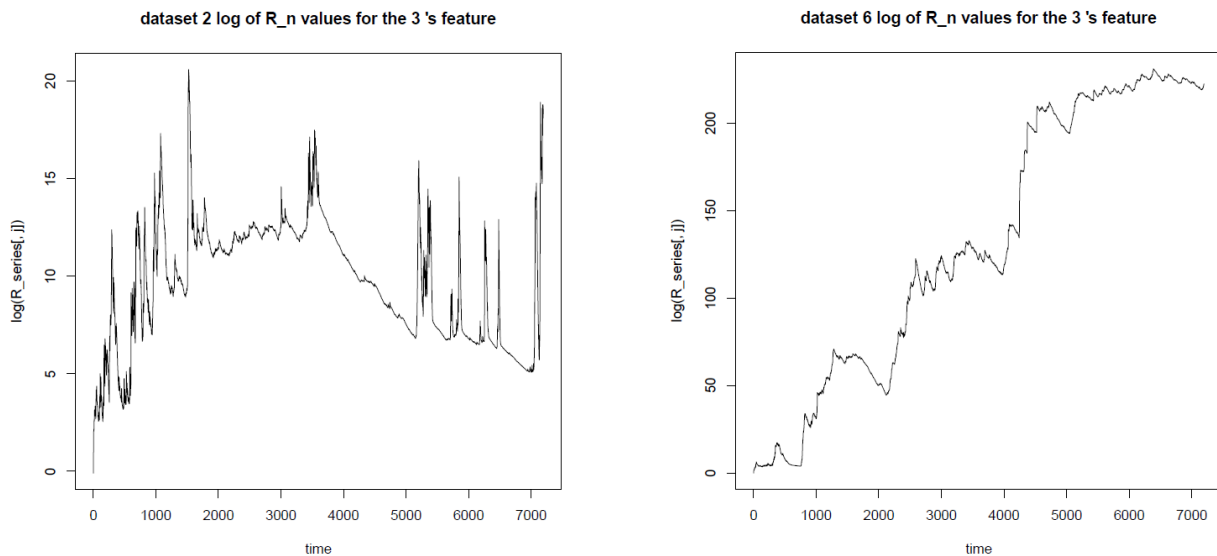
Many cases in Appendix II show false positives. This result is less optimal than previous results obtained by Professor Krieger and Professor Pollak in their analysis, and several factors that could have contributed to it are included in Section V: Discussion.

V.   Discussion

1.   The Algorithm

We demonstrated the potential benefit of this algorithm by showing that more naïve methods would not have satisfactory predicting capabilities that can be used in real life applications. Specifically, we examined the possibility of using the first principal component, or the residuals of AR-3 process on the first principal component, and found that both would not work well enough.

Similar issues are often encountered in the realm of statistical process control, into which we can look for ideas. Indeed, we borrowed the idea of Shiryaev-Roberts statistics. We showed that the SR statistic can give us much better understanding of the EEG recordings by smoothing out noise during normal conditions and amplifying changes that occur. This method effectively detects the change in some parameters in the transformed EEG data sequenced over time, thus making the prediction of seizures possible if we have knowledge about what kind of characteristic changes are signals of imminent seizures.

From the analysis that we have done, we found that with the graphs from the dataset with seizure, we can manually pick a cutoff line for each feature – for example, 25, 20, 25, 40 for feature 1, 2, 3, 4 respectively. We can hypothesize that, in normal conditions, the log of statistics we get won't go past this threshold; they increase beyond this point suggests strongly that a seizure is coming. Hence, we can set the algorithm to raise an alarm at this threshold – whenever the algorithm calculates a log of SR statistic above the threshold. However, by testing this algorithm on datasets without seizure, we found that this threshold can be passed even in control data samples – meaning that we are faced with the problem of false positives. This will be discussed in detail below. Moreover, some consistent patterns are displayed such as that the statistics began to pick after around 100 observations after the training period ends, which might be related to training sufficiency, discussed below.

2. Training length and sufficiency

We are tasked with finding the optimal training length for the algorithm. A longer training length will theoretically yield more sufficient training and better predicting power, but it also gives less time left for the algorithm to gather data before the next seizure. In our case, we used the first 720 data points as the training period. In the dataset with seizure, the "precursor" was found at observation 440, within the training period. This raises the question of in-sample vs. out-of-sample prediction.

We also tested training the algorithm on the first 180 data points, and graphs of the log of SR statistics are attached. As we can see, there is not a clear peak prior to seizure that can be used as the alarm threshold, potentially due to insufficient training on the short time periods.

3. The issue of false positives

False positives occur when the algorithm raises an alarm based on real data, but no seizure occurs afterwards. We are concerned with false positives, but the degree of our concern depends largely on the consequences of false positives. As introduced earlier, the ultimate realization of this algorithm will be a close-loop, implantable device capable of intervention and "pace-making". Thus, the device will "unnecessarily" intervene by sending out electronic signals to deflect the potential upcoming seizure   but no such seizure actually occurred. . If the intervention will not have serious negative impact on the patients, then we will not worry as much – in fact, we might just let the device constantly "intervene" to prevent seizure altogether, since there is no harm anyway. However, if the intervention can have significant side effects – for example, an electric shock or release of medicine might cause uncomfortable feeling or dysfunction of the brain – then we need to be really cautious, and try to reduce false positives to a minimal level. After all, it is a cost-benefit analysis – will the harms avoided by effective interventions outweigh the harm caused by all interventions, whether true or false?

From our testing on the dataset from quiet periods – during which no seizure occurred – we found the problem of false positives is still present in our algorithm, despite the promise that it showed in the dataset with seizure. Feature 3 performed in general better than other features, but still we observed some cases of false positives. In other features we saw more false positives. This result is less promising than the previous analysis done by Krieger and Pollak, in which false positives occurred in only a small fraction of all the tests. I hypothesize that two simplifications in my algorithm could have contributed to the difference.

First, I did not assume a prior distribution on eta, the constant used in calculating the SR statistics. Krieger and Pollak assumed a gamma distribution for $1/eta^2$, and used Monte Carlo simulation to compute the expected value of the statistics. For simplification, I simply fixed $1/eta^2$ to be the expected value of the gamma distribution -- 0.5 -- and proceeded with it throughout. Second, for each feature, Krieger and Pollak used the principal component function in JMP software to find the first and second principal components, while I only took the mean of each feature's 16 locations to be the first principal component. By cross-checking the processed dataset, I found the two methods to have slightly different results.

It should be noted that the data with seizure and data without seizure are on different dogs. Perhaps it is necessary to calibrate the SR statistic for each dog separately, as suggested by evidence from Pearce et al. (2013). One artifact of the data is that the dog data that was used without seizure tends to produce higher SR values and if we had seizure data for that dog the peak might be much higher.

4.   Future research directions

There are some direct extensions of our research that we were not included due to limited scope of the paper. , The main ones are simple modification of the algorithm.

First, our algorithm does not have to be applied to all channels at once. In fact, running our algorithm on a selected set of channels can bring two benefits. It could potentially refine our algorithm and improve the predictive performance of it; it could also help to locate the area of the brain that is responsible for seizure, and increase the precision of surgery. As we know, in many cases seizures are focal, that is only caused by a certain location of the brain. Channels that are in our near the focal area might have better predictive power of the occurrence of seizures.

Second, more principal components could possibly improve the performance of the algorithm. As documented by Krieger and Pollak, the inclusion of the second principal component improved the performance of the algorithm in their testing. We have to be cautious about including too many principal components, however, as this increases computing burden quickly. Furthermore, since currently our alarm is raised within the training period, there could potentially be risk of over fitting.

Last but not least, researchers could consider processes more complicated than AR-3 to capture the variation of the principal components. From the graphs we plotted of the first principal components, there seems to be seasonal variations in the data, and therefore we could consider adding a seasonal component to the AR model to improve it. However, this observation is casual and subject to further statistical testing.

VI.   Conclusion

In this paper, we tested the seizure prediction algorithm proposed by Krieger and Pollak on datasets from dogs. The testing on the dataset with seizure shows that this algorithm is promising, because it successfully captured the change in distribution of the data more than two hours before seizure onset, signaling that it is possible to predict seizure in a prospective manner. However, the testing on the control datasets without seizure shows the possibility of false positives. Specifically, my algorithm produced significantly more false positives than the previous analysis done by Krieger and Pollak, and I hypothesized that two simplifications in my algorithm might have contributed to the difference. Finally, I

discussed three possible extensions of this algorithm which can be easily tested in the future, namely channel selection, more principal components, and more comprehensive processes.

VII.   Notes

Appendix I contains the graphs plotted for the dataset with seizure; Appendix II contains the graphs plotted for the datasets without seizure; Appendix III contains the R-code used for testing.

In fact, because different dataset has slightly different structure, there are two versions of the code. One is for dataset with seizure, and one is for the ones without.

References

Aschenbrenner-Scheibe, R., et al. "How well can epileptic seizures be predicted? An evaluation of a nonlinear method." *Brain* 126.12 (2003): 2616-2626.

Chaovalitwongse, Wanpracha Art, Oleg A. Prokopyev, and Panos M. Pardalos. "Electroencephalogram (EEG) time series classification: Applications in epilepsy." *Annals of Operations Research* 148.1 (2006): 227-250.

Cohen, Ivan, et al. "On the origin of interictal activity in human temporal lobe epilepsy in vitro." *Science* 298.5597 (2002): 1418-1421.

D'Alessandro, Maryann, et al. "A multi-feature and multi-channel univariate selection process for seizure prediction." *Clinical neurophysiology* 116.3 (2005): 506-516.

Esteller, Rosana, et al. "Continuous energy variation during the seizure cycle: towards an on-line accumulated energy." *Clinical Neurophysiology* 116.3 (2005): 517-526.

Harrison, Mary Ann F., Mark G. Frei, and Ivan Osorio. "Accumulated energy revisited." *Clinical Neurophysiology* 116.3 (2005): 527-531.

Hauser, W. Allen, John F. Annegers, and Walter A. Rocca. "Descriptive epidemiology of epilepsy: contributions of population-based studies from Rochester, Minnesota." *Mayo Clinic Proceedings*. Vol. 71. No. 6. Elsevier, 1996.

Iasemidis, L. D., et al. "Long-term prospective on-line real-time seizure prediction." *Clinical Neurophysiology* 116.3 (2005): 532-544.

Iasemidis, Leon D. "Epileptic seizure prediction and control." *Biomedical Engineering, IEEE Transactions on* 50.5 (2003): 549-558.

Iasemidis, Leonidas D., et al. "Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures." *Brain topography* 2.3 (1990): 187-201.

Le Van Quyen, Michel, et al. "Preictal state identification by synchronization changes in long-term intracranial EEG recordings." *Clinical Neurophysiology* 116.3 (2005): 559-568.

Le Van Quyen, Michel, et al. "Anticipating epileptic seizures in real time by a non-linear analysis of similarity between EEG recordings." *Neuroreport* 10.10 (1999): 2149-2155.

Le Van Quyen, M., et al. "Spatio-temporal characterizations of non-linear changes in intracranial activities prior to human temporal lobe seizures." *European Journal of Neuroscience* 12.6 (2000): 2124-2134.

Le Van Quyen, Michel, et al. "Anticipation of epileptic seizures from standard EEG recordings." *The Lancet* 357.9251 (2001): 183-188.

Le Van Quyen, Michel, et al. "Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony." *Journal of neuroscience methods* 111.2 (2001): 83-98.

Mormann, Florian, et al. "Seizure prediction: the long and winding road." *Brain* 130.2 (2007): 314-333.

Mormann, Florian, et al. "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients." *Physica D: Nonlinear Phenomena* 144.3 (2000): 358-369.

Maiwald, Thomas, and Matthias Winterhalder. "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic." (2003).

Litt B, Esteller R, Echauz J, D'Alessandro M, Shor R, Henry T, et al. Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. Neuron 2001; 30: 51–64.

Lehnertz, Klaus, and Christian E. Elger. "Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity."*Physical Review Letters* 80.22 (1998): 5019.

Lehnertz, Klaus, and Brian Litt. "The first international collaborative workshop on seizure prediction: summary and data description." *Clinical neurophysiology*116.3 (2005): 493-505.

Mormann, Florian, et al. "On the predictability of epileptic seizures." *Clinical neurophysiology* 116.3 (2005): 569-587.

Mormann, Florian, Christian E. Elger, and Klaus Lehnertz. "Seizure anticipation: from algorithms to clinical practice." *Current opinion in neurology*19.2 (2006): 187-193.

Martinerie, J., et al. "Epileptic seizures can be anticipated by non-linear analysis." *Nature medicine* 4.10 (1998): 1173-1176.

Maiwald, Thomas, et al. "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic." *Physica D: nonlinear phenomena* 194.3 (2004): 357-368.

McSharry, Patrick E., Leonard A. Smith, and Lionel Tarassenko. "Prediction of epileptic seizures: are nonlinear methods relevant?." *Nature medicine* 9.3 (2003): 241-242.

Pearce, Allison, et al. "Temporal changes of neocortical high-frequency oscillations in epilepsy." *Journal of neurophysiology* 110.5 (2013): 1167-1179.

Pollak, M. (2009). The Shiryaev–Roberts changepoint detection procedure in retrospect-theory and practice. In *Proceedings of the 2nd International Workshop in Sequential Methodologies*.

Rogowski, Z., I. Gath, and E. Bental. "On the prediction of epileptic seizures."*Biological cybernetics* 42.1 (1981): 9-15.

Salant, Y., I. Gath, and O. Henriksen. "Prediction of epileptic seizures from two-channel EEG." *Medical and Biological Engineering and Computing* 36.5 (1998): 549-556.

Winterhalder, Matthias, et al. "Spatio-temporal patient–individual assessment of synchronization changes for epileptic seizure prediction." *Clinical neurophysiology* 117.11 (2006): 2399-2413.