5-2019

Spectrogram Analysis of Blood Pressure on Neonates with Hypoxic Ischemic Encephalopathy (HIE)

Tianrui Zhu
College of William & Mary

Spectrogram Analysis of Blood Pressure on
Neonates with
Hypoxic Ischemic Encephalopathy (HIE)

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Science in Mathematics from
The College of William and Mary

by

Tianrui Zhu

Accepted for Honors
(Honors)

John Delos, Director

Josh Erlich

Junping Shi

Williamsburg, VA
May 2, 2019

Spectrogram Analysis of Blood Pressure on Neonates with Hypoxic Ischemic Encephalopathy (HIE)

Tianrui Zhu

College of William & Mary

Abstract

This paper explores the correlation between blood pressure fluctuations and the severity of Hypoxic Ischemic Encephalopathy (HIE). Via isolating particular frequency bands of blood pressure signal and calculating their power, the researcher was able to discern patients with different HIE severity and verified the assumption that less fluctuation in blood pressure is correlated with more severe brain injury. Furthermore, the researcher attempted to construct a prediction algorithm based on power within certain frequency bands of the blood pressure signal.

Keywords: *Blood pressure, HIE, Fourier Transform, Kernel Density Estimation*, *prediction*

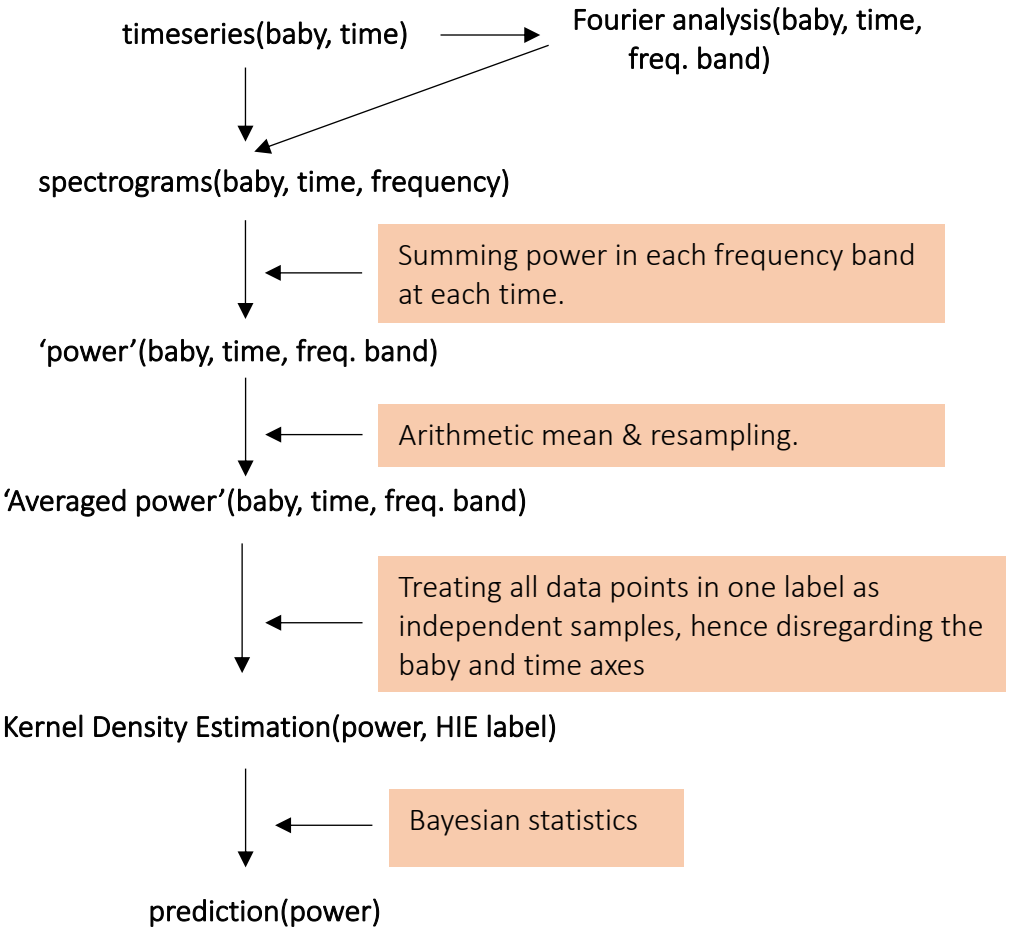
Acknowledgement

I would first of all like to thank Professor John Delos for his supervision and advising throughout the past two years. I would also like to thank Doctor Karen Fairchild, Professor Abigail Flower, Professor Douglas Lake, Doctor Zachery Vesoulis, and many others at the University of Virginia who took part in the project – without them this project would not have existed. I would like to thank Professor Daniel Vasiliu for his valuable insight and suggestions. Last but not least, I would like to extend my most sincere appreciate to my family and friends, and especially my parents, who all helped me a lot throughout the process.

This research is partly funded by the 2018 Charles Center Summer Honors Research Scholarship.

Table of Contents

Flowchart


timeseries(baby, time)
Fourier analysis(baby, time, freq. band)
spectrograms(baby, time, frequency)
Summing power in each frequency band at each time.
'power'(baby, time, freq. band)
Arithmetic mean & resampling.
'Averaged power'(baby, time, freq. band)
Treating all data points in one label as independent samples, hence disregarding the baby and time axes
Kernel Density Estimation(power, HIE label)
Bayesian statistics
prediction(power)

I. Introduction

Hypoxic Ischemic Encephalopathy (HIE) is a subtype of Neonatal Encephalopathy (NE), a clinical condition in which a neonate's brain malfunctions. Although lacking a universal definition, HIE can be broadly described as a condition where the brain experiences a reduced amount of blood supply or a diminished amount of blood perfusion (Volpe, 2001). Other definitions exist, and in fact in many cases the terms HIE and NE are used interchangeably for various reasons. For our purposes, it is sufficient to define HIE as neonatal brain injury related to insufficient blood flow and reduced oxygen supply.

Reported incident rates of HIE varies greatly across nations and time, with lower estimations around 1.0 per 1000 live babies (Badawi, et al., 1998) and higher estimations around 8.0 per 1000 live babies (Thornberg, Thiringer, Odeback, & Milsom, 1995; Hull & Dodd, 1992). The highest incidence rate reported was approximately 26 per 1000 live babies based on a single hospital observation in Nigeria (Airede, 1991). This discrepancy in estimation may be partially explained by the lack of common standards as well as varying ways of measuring HIE, which is not uncommon in this field of neonatology. For example, no two of the previously mentioned four studies used the same criteria or measurement in determining HIE, and while some studies were population-based estimation, others were hospital based, which usually results in higher numbers. Other potential confounds exist such as time of measurement and other parameters. Since HIE is a progressive disease, the time of measurement as well as the gestational age of a baby greatly influence the outcome. Currently, the best estimation given by the American College of Obstetricians and

Gynecologists and the Academy of Pediatrics puts HIE incidence rate at 1.9 – 3.8 per 1000 live babies (Hankins & Speer, 2003).

As expected, blood flow is critical to neonatal brain development. Any prolonged disruption in the delivery of glucose or oxygen may severely damage brain function, resulting in a hypoxia event. A variety of clinical conditions may result in a hypoxia, including placental abruption, prolapse of umbilical cord, and uterine rupture (Douglas-Escobar & Weiss, 2015). A moderate hypoxia event will likely prompt the brain to shut down its anterior circulation to redirect and maintain sufficient perfusion for more critical regions such as the brain stem, cerebellum, basal ganglia, and other regions supporting basic life functions (Douglas-Escobar & Weiss, 2015; Harteman, et al., 2013). As a result, regions supplied by the anterior circulation such as the cerebral cortex will suffer from a sequence of injuries. In extreme events where hypoxia is severe and sudden, the brain might fail to adequately redirect blood perfusion to critical regions, resulting in damage to the basal ganglia, thalami, or even eventually causing premature death (Douglas-Escobar & Weiss, 2015).

Historically, the severity of HIE is categorized by the Sarnat Staging System. The Sarnat Staging system is an observation-based system which comprises components such as muscle tone or pupil reaction. An infant suspected of HIE may be classified into normal, Label I: Mild, Label II: Moderate, or Label III: Severe according to the Sarnat Staging System. Most babies being diagnosed as Label III: Severe HIE suffer from long-term neurological disorders if not treated properly and in time.

As one might have realized, the Sarnat Staging System is arguably neither very objective nor accurate. Thanks to technological advances, better ways of clinical classifying

HIE severity such as Electroencephalography (EEG) analysis or Magnetic Resonance Imaging (MRI) analysis have been developed. However, due to various constraints, such measurements usually cannot be administered until some time after birth. EEG analysis is usually done approximately 24 hours, 48 hours, or 72 hours into an infant's life. MRIs are usually done around the same time.

While MRI is often considered by the medical field as the gold standard, executing MRI on infants is particularly hard. Furthermore, although MRI is thought to be the safest screening method developed to date, its safety on preterm or near-term infants has never been examined (Stokowski, 2005). Modern day MRI machines are predominantly designed for adults and are often large and noisy. Expecting an infant to cooperate throughout an MRI scan will be unreasonable. Besides the practical difficulty in putting an infant through an MRI scan, interpretation of MRI results is largely subjective to the examiner's experience. In fact, it is not uncommon for different medical practitioners with different backgrounds to disagree and come up with contrasting interpretation of the same MRI image. In our research, for example, initial MRI readings were later determined by the same group of physicians at UVa Health as rather inaccurate and required second readings. Moreover, developing a common rubric to interpret infant MRI collected at different locations is exceedingly difficult due to sheer number of confounds present. Incorporating MRI information from other institutions requires sending data back and forth so that one person could do all the readings to ensure consistency.

EEG, on the other hand, is considered easier to examine and depends less on the interpretation of individuals. Examining wave forms at different frequencies gives clinicians a way to estimate the severity of brain damage of an infant. However, it has been argued that

compared to MRI, EEG is less accurate for predicting long term outcome. More importantly, unlike MRI which is a direct measurement and pinpoints specific regions of damage, EEG only measures brain electrical activity and is thus still indirect.

In conclusion, we found that most current ways of screening for HIE either require subjective interpretation or are hard to execute on infants. We were therefore asked to explore alternative methods to objectively classify HIE severity using regularly-collected data available from bed-side monitors. Ideally, such models will provide predictions with reasonable accuracy for physicians before any of the above-mentioned elaborate measurements are available.

This is not the first study in which researchers attempted to classify brain damage based on solely bed-side monitor data. Previous research has established that heart rate variability may be negatively correlated with the severity of various diseases among infants such as sepsis through spectrogram analysis (Aboab, et al., 2008). The correlation between certain pathologies and other bed-side monitor readings has also been examined. Continuous blood pressure, however, is often available yet overlooked. In particular, to the best of our knowledge, the relationship between blood pressure and HIE severity has never been addressed. Blood pressure in Neonatal Intensive Care Unit (NICU) infants is usually collected by sensors implanted in umbilical or other arteries. Preliminary visual inspection suggests that blood pressure or its variability might actually be reflective of brain damage. Naturally, we hypothesized that similar to how heart rate fluctuation correlates with sepsis, blood pressure fluctuations may provide a measure of HIE severity. Furthermore, we also picked up surprising examples of suspicious pseudo-periodic oscillation (see Figure I - A) in blood pressure among some neonates. Some of these cases imitate sinusoidal waves with

periodicity of approximately 30 minutes, others are pseudo-periodic yet heavily right skewed and have lower frequencies. These cases are few and far between, yet the patterns are inherently interesting since clear-cut oscillations are seldom seen in infants. Moreover, such patterns are seldom reported and their clinical significance has never been examined. We are thus also interested in examining whether these periods of oscillation may be indicative of underlying diseases.

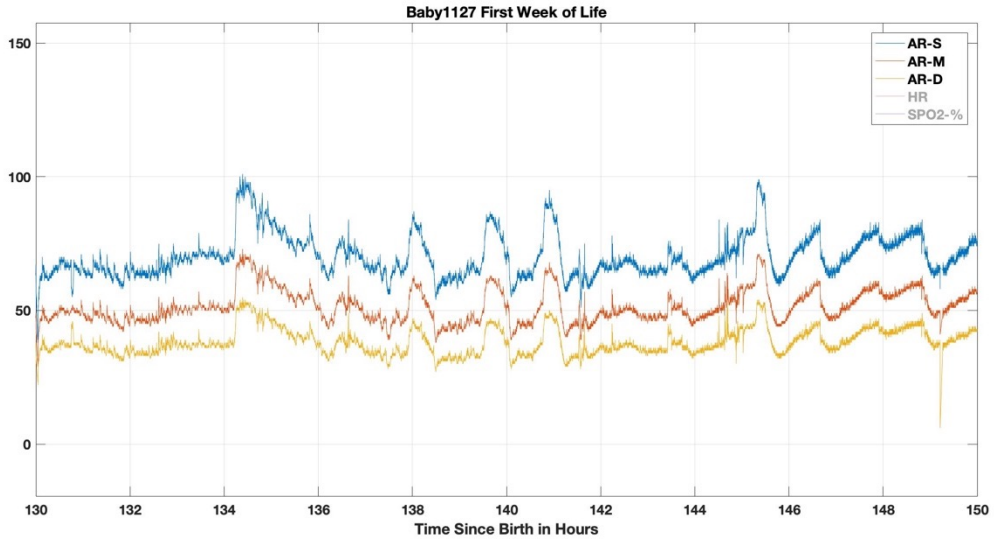


Figure I - A. Example of pseudo-periodicity in blood pressure. The horizontal axis denotes hours since birth and the vertical axis denotes blood pressure reading in mmHg (systolic, mean and diastolic). Interesting cyclical patterns emerge around hour 140.

II. Overview of Dataset

Our database consists of a variety of electronic signals collected from monitoring devices in Neonatal Intensive Care Unit (NICU) at the University of Virginia Health System (UVa Health) between 2009 to date. All data presented are anonymous and labeled only by NeoID, a pseudo-identification number which can only be correlated to the respective individual medical records through information held confidential at UVa Health. This work has been approved by research compliance committees at UVa and William & Mary. The particular dataset used in this study consists of 78 full-term babies (NeoID range: 1127 ~ 5562) who had some period of blood-pressure data recorded within the first week of life. All babies in this dataset were term or near-term infants who were suspected of suffering from HIE, and hence brought to the NICU. It is worth noting that apart from HIE, many of these babies also suffered from other pathologies, which makes it harder to make deterministic conclusions as we progress.

We identified two sets of blood pressure data with three signals per set in the database: *sig 12*/*sig 57*, *sig 13*/*sig 58*, *sig 14*/*sig 59*, and corresponding time signals *T_sig 12*/*T_sig 57*, *T_sig 13*/*T_sig 58*, *T_sig 14*/*T_sig 59*. *sig12*/*sig57* corresponds to systolic blood pressure, *sig 13*/*sig 58* corresponds to diastolic and *sig 14*/*sig 59* corresponds to mean blood pressure. Systolic and diastolic blood pressures are direct measurements while mean blood pressure is computed according to clinical convention by taking a weighted arithmetic mean between systolic and diastolic blood pressure. A typical ratio would be:

$$M_{bp} = \frac{2}{3} D_{bp} + \frac{1}{3} S_{bp}$$

where M_{bp} represents mean blood pressure, D_{bp} represents diastolic blood pressure and S_{bp} represents systolic blood pressure. All blood pressure signals are originally presented with a sampling frequency of 0.5Hz (see Figure II-A).

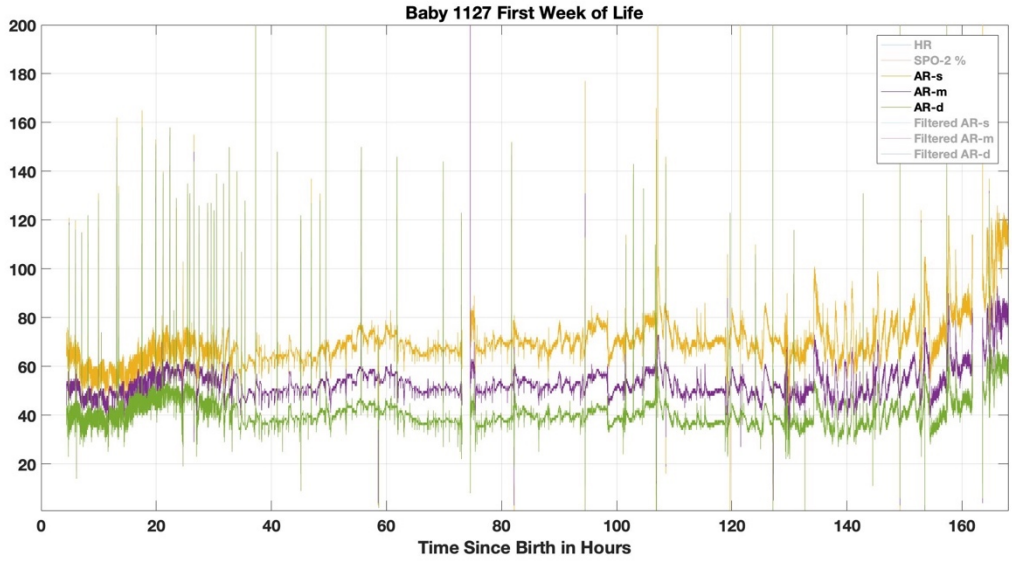


Figure II - A. Example of raw blood pressure data. The horizontal axis is hours since birth and the vertical axis is blood pressure reading in mmHg. Note the amount of unrealistic fluctuations that exist in this data.

As for HIE labeling, the dataset includes both binary labels (Label 0: Normal/mild HIE, Label 1: Moderate/severe HIE) and conventional Sarnat-like three-label classification (Label 1: Normal/mild, Label 2: Moderate, Label 3: Severe) based on 24 hour EEG(EEG-24), 48 hour EEG(EEG-48), 72 hour EEG(EEG-72), 24 hour MRI(MRI-24), and 48 hour MRI(MRI-48). In this particular study, we choose Sarnat-like three-label EEG-24 classification as "ground truth".

a. **Pre-processing.** As discussed above, the particular subset of data we utilized consists of data ranging from NeoID 1127 to NeoID 5562. This particular subset can be further divided into two parts in terms of how data is presented.

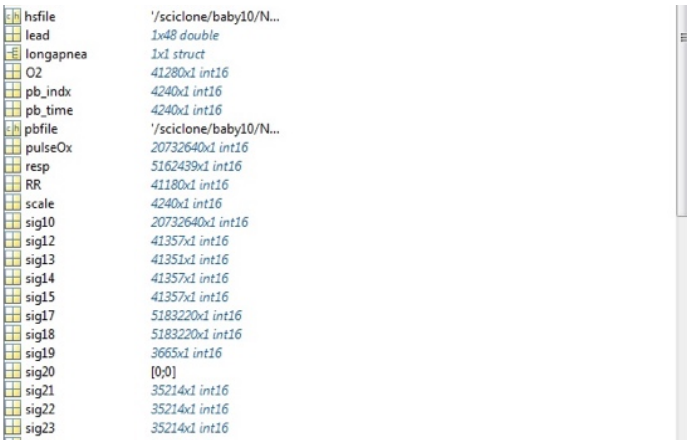


Figure II - B. Example of the contents of a .mat file. This file contains information for bed 11 day 0027.

i. ***NeoID ≤4019*:** Data is stored as .mat files (see Figure II - B) in an external hard drive and indexed by bed-day rather than NeoID. Each .mat file consists of all data collected by the bed-side monitor at that particular bed on that day. In cases where multiple babies occupied the same bed throughout a day, the .mat file still stores all data as one file without indicating a switch in baby. Additional information including birth time and initial bed location information is provided in a separate .mat file to align the above-mentioned bed-day data with NeoID. Effort is needed to extract data from these bed-day files and convert them into variables indexed by NeoID for further processing. As is expected for any medical catalog, availability of all the above-mentioned measurements vary widely across babies and time.

ii. ***NeoID >4019:*** Data is stored as .mat files in a secured online database. These data have mostly been preprocessed; they contain only blood pressure and other critical signals. The baby-bed-day matching which was necessary for infants with NeoID ≤ 4019 had already been finished by collaborators at UVa Health. Hence, these data are already indexed by NeoID, and minimum pre-processing is necessary for these data.

III. Filtering

Clinical data is usually messy, and our dataset does not deviate from this norm. Artifacts in blood pressure data can arise for a variety of reasons, including flushing, motional artifacts, body position, and also other medical interventions such as drug administration. In our particular dataset, three particular types of artifacts are prominent:

a. **Flushing.** Flushing refers to the act of disconnecting the measurement device to release gas that has accumulated throughout the system. Artifacts caused by flushing usually exhibit unique patterns where all three blood pressure readings spike and/or converge (see Figure III - A).

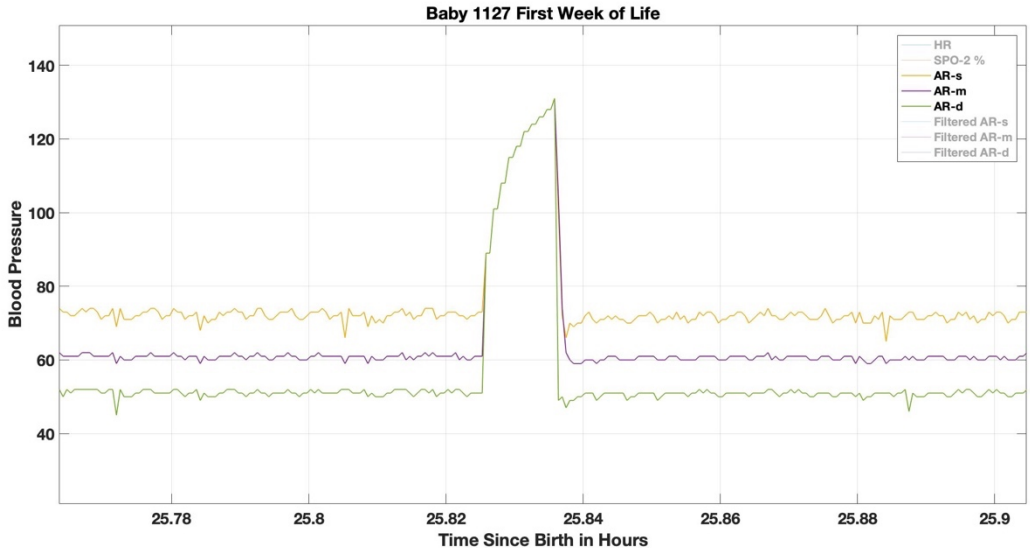


Figure III - A. Example of possible flushing. The horizontal axis denotes hours since birth and vertical axis denotes blood pressure readings in mmHg. Note the sudden spike and convergence around hour 25.83.

The first part of our filtering algorithm deals with this type of artifact. First, we define time-series S_{bp}, D_{bp}, M_{bp} that correspond to systolic, diastolic and mean blood pressure as usual. The i^{th} point in these time series can be referenced by $S_{bp}(i)$, $D_{bp}(i)$, $M_{bp}(i)$

respectively. For this part of the filtering, we assume no time discrepancy between three signals which means $S_{bp}(i)$, $D_{bp}(i)$, $M_{bp}(i)$ should share a common time-stamp *t*. (Time discrepancies are handled through another part of the filtering process.) In this filtering step, we eliminate points where the differences between the three blood pressures signals are smaller than a selected threshold $\alpha \geq 0$, which can be represented as:

$$\begin{cases} S_{\text{cleaned}} = \{S_{bp}(i) | \left|S_{bp}(i) - M_{bp}(i)\right| \geq \alpha \vee |M_{bp}(i) - D_{bp}(i)| \geq \alpha\} \\ M_{\text{cleaned}} = \{M_{bp}(i) | \left|S_{bp}(i) - M_{bp}(i)\right| \geq \alpha \vee |M_{bp}(i) - D_{bp}(i)| \geq \alpha\} \\ D_{cleaned} = \{D_{bp}(i) | \left|S_{bp}(i) - M_{bp}(i)\right| \geq \alpha \vee |M_{bp}(i) - D_{bp}(i)| \geq \alpha\} \end{cases}$$

By adjusting α, the researcher has total control over the filtering criteria and can tighten or loosen the criteria as desired.

For this particular study, α was set to 3 mmHg.

b. **Gaps.** Gaps in data are another type of artifact prominent in this dataset. Gaps can arise from multiple reasons such as the infant being transferred to another location, detachment from measurement device, or any other medical operation in which a blood pressure measurement is not available (see Figure III - B).

To remove gaps in the dataset, the researcher applied linear interpolation where applicable. In particular, the researcher utilized the Matlab built-in *fillmissing()* function, and in particular, parameters were set to:

[x_new] = fillmissing(…,'linear');

Upon completion, additional visual inspection was implemented to ensure reasonable interpolation results.

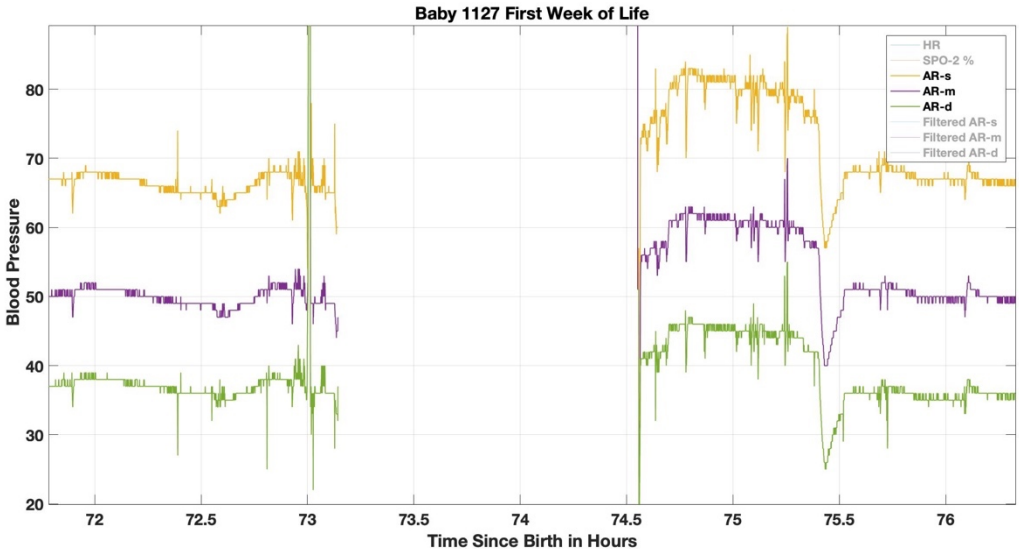


Figure III - B. Example of a substantial gap in data. The horizontal axis denotes hours since birth and the vertical axis denotes blood pressure readings in mmHg. Note the substantial gap starting around hour 73 and lasting for over an hour.

c. **Random Fluctuations.** A third prominent artifact in this dataset is random fluctuation. This could again be caused by different reasons, including movement of the baby, medication, inspection of the baby, or other reasons not known to us. Typical characteristics of such fluctuations: (a) a sudden rise or drop in one or all of the blood pressure signals, (b) one or all of the blood pressure readings drastically undershooting or overshooting the normal range, (c) sudden discrepancy between two consecutive data points of a single signal, (d) prolonged period of wild oscillation (see Figure III - C).

Carrying on the notations assumed in 3.1 Flushing, we define a second threshold variable τ. Then we filtered these artifacts by the following procedure:

$$\begin{cases} \mathrm{S}_{\mathrm{cleaned}} = \{S_{bp}(i)|\ |S_{bp}(i+1) - S_{bp}(i)| \le \tau \vee 20 \le S_{bp}(i) \le 100\} \\ \mathrm{M}_{\mathrm{cleaned}} = \{M_{bp}(i)|\ |M_{bp}(i+1) - M_{bp}(i)| \le \tau \vee 20 \le M_{bp}(i) \le 100\} \\ D_{cleaned} = \{D_{bp}(i)|\ |D_{bp}(i+1) - D_{bp}(i)| \le \tau \vee 20 \le D_{bp}(i) \le 100\} \end{cases}$$

- Remove the second data point if two consecutive data points in one signal differ by a number larger than τ;

- Remove data points substantially over or under the normal range, which is defined as between 20 ~ 100 for all blood pressure signals.

It is worth noting an additional protection mechanism was in place that will override this particular section of the filtering algorithm if triggered. If the gap between two consecutive data points is longer than or equal to 10 seconds, the above-mentioned filtering criteria is not imposed to avoid excessive filtering.

In this study, τ was set to be 5.

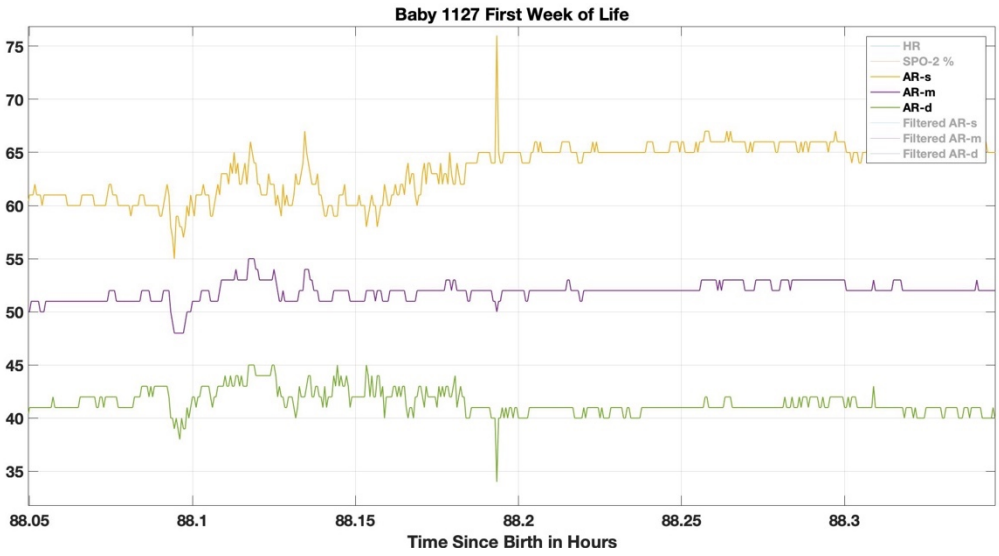


Figure III - C. Example of random fluctuations. The horizontal axis denotes hours since birth and the vertical axis denotes blood pressure readings in mmHg. Note the unrealistic readings in systolic and diastolic blood pressure around hour 88.2.

d. **Misalignment.** Misalignment refers to cases where some or all three signals do not start at the same time. This phenomenon is especially prominent when the signal picks up after a prolonged gap. Conceptually, such time discrepancies are reasonable since systolic pressure and diastolic pressure alternate. At any particular moment, the heart is either in systolic or diastolic state, and in theory one cannot measure systolic and diastolic blood pressure simultaneously. Additionally, since mean blood pressure is computed by taking

a weighted mean of systolic and diastolic pressure, it is reasonable for mean blood pressure to be lagging even more. This discrepancy, however, proves to be problematic for further analysis. To tackle this issue, we resampled all signals to 1Hz frequency with universal time stamp:

$$T_{universal} = \{t | t \in T\}.$$

e. **Other general rules.** Other general rules are observed and imposed to ensure reasonable outcomes throughout the filtering processing. Some of these rules include:

- No negative data-points should be allowed;
- The order $S_{bp}(i) \geq M_{bp}(i) \geq D_{bp}(i)$ should always hold.

The abovementioned rules of filtering are imposed upon the data from all available babies, and filtered blood-pressure signals are stored as both variables in workspace as well as .fig Matlab figure files for future examination and retrieval (see Figure III - D).

Visual inspection of all the filtered blood pressure signals made us exclude 24 babies whom either had insufficient data (less than 2 days), substantial periods of missing data, or who had blood pressure signal only after Day 3. Babies with no corresponding HIE labels were also not included. This left us with a total of 54 babies (Label 1: n = 25, Label 2: n = 11, Label 3: n = 18), which is the dataset we based our analysis on.

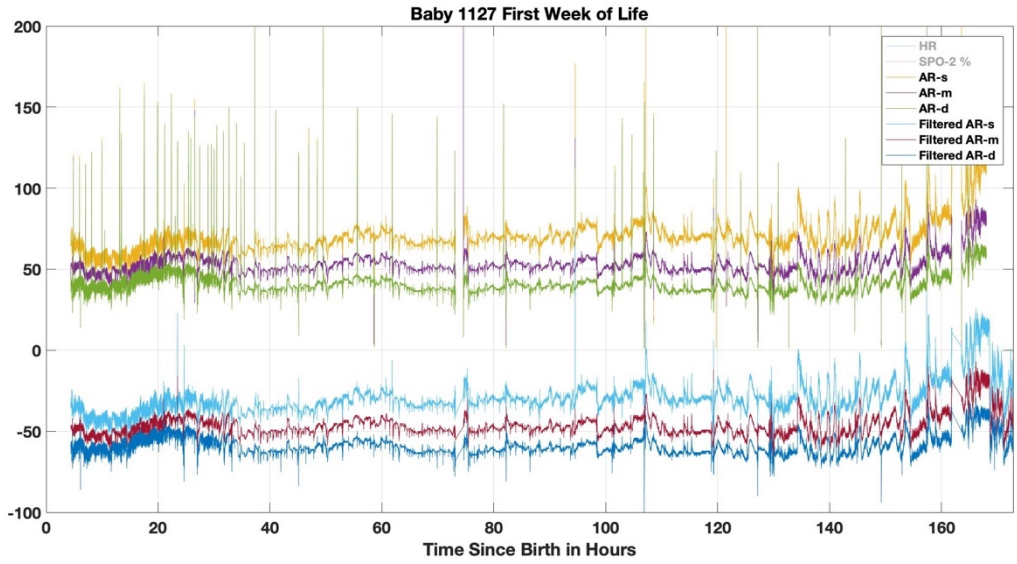


Figure III - D. Example of filtering result. The horizontal axis denotes hours since birth and the vertical axis denotes blood pressure readings in mmHg. The top three signals are the unfiltered raw blood pressure data. The bottom three signals are the corresponding filtered blood pressure minus 100 mmHg. Note the degree to which filtering has helped, albeit some apparent artifacts still exist.

IV. Fourier Analysis

A previous study by Zachery Vesoulis and others has established correlation between blood pressure fluctuation and effectiveness of inotrope treatment (Vesoulis, Mathur, & Mcpherson, 2017). In that study, Vesoulis et al. approached the data by breaking down blood pressure into four continuous frequency bands which correlate to different underlying physiological phenomena. As we were also interested in pseudo-periodic behaviors of blood pressure and also hypothesized a correlation between blood pressure fluctuation and brain damage, we adopt a similar approach. Since our data was approximately week-long and we were also interested in hour-long fluctuations, we added additional frequency bands to analyze fluctuations with lower frequencies. To be specific, we broke down our data into the following 3 sets which constitutes a total of 10 continuous frequency bands. These sets and bands will be repeatedly utilized throughout this paper:

- Set 1, identified by Vesoulis et al. (Vesoulis, Mathur, & Mcpherson, 2017):
    - Band A: 6 ~ 17 seconds, corresponding to myogenic waves;
    - Band B: 17 ~ 51 seconds, corresponding to neurogenic waves;
    - Band C: 51 ~ 105 seconds, corresponding to endothelial NO-dependent waves;
    - Band D: 105 ~ 300 seconds, roughly corresponding to endothelial waves;
- Set 2:
    - Band E: 5 ~ 10 minutes;
    - Band F: 10 ~ 20 minutes;
    - Band G: 20 ~ 30 minutes;

- Set 3:
    - Band H: 0.5 ~ 1 hour;
    - Band I: 1 ~ 2 hours;
    - Band J: 2 ~ 3 hours.

To begin, we sought to analyze how each frequency band contributes to the overall signal by isolating each frequency band. Two different methods of isolating signal in each frequency bands were tried and tested.

a. **Hard Thresholding:** A hard threshold essentially imitates a band-pass filter with a vertical cliff and no tail (also known as a brick-wall filter). To achieve this desired performance, we built our own algorithm based on Matlab's built-in FFT function *fft()*.

According to Matlab documentation (MATLAB, 2006), the *fft()* function computes a discrete Fourier transform based on the standard fast Fourier transform algorithm. Suppose we have a vector/array of time-series data X, then *Y = fft(X)* computes

$$Y(k) = \sum_{j=1}^{n} X(j) e^{-\frac{2\pi i}{n}(j-1)(k-1)}$$

where Y(k) is the frequency domain representation of the respective input vector/array.

After retrieving the frequency domain representation of the input data, we may impose a hard threshold by the following:

- Identify frequencies that match with the upper-bound and lower-bound of the desired frequency band to construct an interval of interest. It is worth noting that more often than not the frequencies at which coefficients are calculated through discrete Fourier transform do not exactly coincide with the desired upper or lower boundaries, so a closest frequency is chosen.

- Annihilate the signal at all frequencies that are not in the interval of interest by replacing their Fourier coefficients with zeroes. Preserve all data within the interval of interest for band reconstruction.

After imposing a hard threshold, we may reconstruct the blood-pressure within this particular frequency band by utilizing the Matlab built-in inverse Fast Fourier Transform function, *ifft()*. Again according to Matlab documentation (MATLAB, 2006), the *ifft()* function calculates the following: Suppose we have frequency domain information Y, then the corresponding time-series data is computed by:

$$X(j) = \frac{1}{n}\sum_{k=1}^{n} Y(k)e^{-\frac{2\pi i}{n}-(j-1)(k-1)}$$

After examining this methodology, we discovered that imposing a hard threshold could potentially result in undesired behaviors such as oscillations and other artifacts in the reconstructed data due to the nature of a brick-wall filter. Hence, we decided that a soft threshold would be a better option.

b. **Soft Thresholding:** A soft threshold is essentially a band-pass filter with a non-vertical cliff and some tail across the entire frequency range. To impose a soft threshold, we utilized the Matlab built-in zero-phase filtering *filtfilt()* function, which filters a time-series data according to a provided filter. To impose a filtfilt() function, one must first construct a filter using the *designfilt()* function. In this particular application, we chose a Butterworth Infinite Impulse Response filter (IIR) rather than a Finite Impulse Response filter (FIR) to achieve faster computation speed. Specifically, we defined the half-power frequency of the filter to be the upper and lower boundaries of the frequency band of

interest and a filter order of around five to avoid the ringing oscillations resulting from a brick-wall filter.

After constructing the desired filter, we imposed the filter on our time-series blood-pressure data to isolate the desired frequency band.

Below is an example of the reconstructed blood pressure (see Figure IV - A):

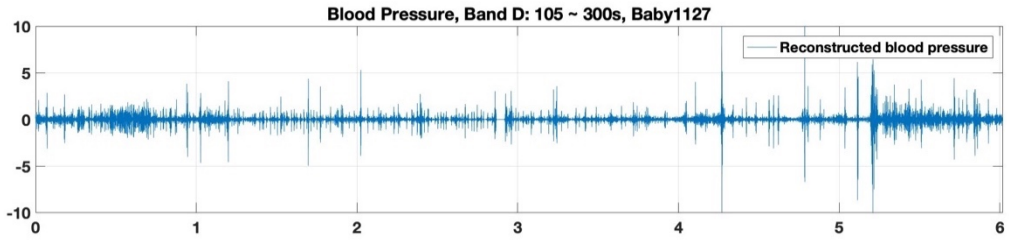


Figure IV - A. Example of reconstructed blood pressure within a frequency band. The horizontal axis denotes days since start of data, and the vertical axis denotes fluctuation of blood pressure.

After identifying soft-thresholding as the way to isolate a particular frequency band, we isolated blood pressure for all ten frequency bands (Band A to J) for each of the 54 babies for whom we had data and a corresponding HIE label. Visual examination of the reconstructed blood pressures suggests particularly interesting changes in Band D: 105 ~ 300 seconds, which we decided to further examine with spectrogram analysis.

V. Spectrogram & Power Analysis

After identifying frequencies of interest via Fourier transform, we decided it would be beneficial to further analyze the patterns of blood-pressure fluctuations by computing their power via spectrogram.

a. **Spectrogram across all Frequencies.** For our purpose, we once again utilized the Matlab built-in *spectrogram()* function. The *spectrogram()* function calculates the power spectrum density (PSD) of each input segment at different frequencies using short-term Fourier transform similar to what is done in Chapter IV. To demonstrate, the *Spectrogram()* function can be called with the following:

[s, f, t, ps, fc, tc] = spectrogram(x, …, fs)

where x is the input time-series (usually a segment of the entire time-series), fs indicates the sample rate of the input signal. On the output side of the function, we are particularly interested in the variables *ps* and *f* since they correspond to the power spectral density and the exact frequencies at which PSD coefficients are calculated. Below is an example of *f* (see Figure V - A) and *ps* (see Figure V - B):

2	4.8828e-04
3	9.7656e-04
4	0.0015
5	0.0020
6	0.0024
7	0.0029
8	0.0034
9	0.0039
10	0.0044
11	0.0049
12	0.0054
13	0.0059
14	0.0063
15	0.0068
16	0.0073
17	0.0078
18	0.0083
19	0.0088
20	0.0093
21	0.0098
22	0.0103
23	0.0107

Figure V - A. Example of frequencies in Hz at which power is calculated.

1	2	3	4	5	6	7	8	9	10	11	12	13	
2.0006e+03	5.1462e+03	7.0450e+03	5.6481e+03	3.0779e+03	1.3348e+03	510.8379	161.2183	9.0678	50.8141	300.2092	637.2893	904.6803	1
2.5405e+03	5.6482e+03	6.4589e+03	4.9268e+03	2.5868e+03	986.6403	415.8412	203.6676	109.3042	150.9430	331.8963	581.5751	696.6053	
677.0415	736.4896	465.2194	228.2976	254.7961	35.3394	16.0847	12.5044	25.8572	21.2810	32.6079	44.0761	14.7891	
7.7341	28.7493	75.8351	47.4680	49.4502	67.8120	6.6992	9.3458	1.3204	1.2063	0.8525	0.8431	9.5436	
84.7583	47.5672	42.6285	41.6620	33.8091	44.0343	46.8673	27.9821	11.5239	2.5664	0.2123	1.0088	9.0876	
36.7224	1.9918	27.8750	5.8448	9.6585	4.6634	21.5063	50.9564	26.9753	6.8446	3.0379	3.2393	3.0268	
35.2733	42.6282	63.1294	94.4564	77.1832	52.7037	55.4870	55.2223	22.5665	0.8512	2.3488	2.8802	1.1698	
28.7358	2.9451	106.1181	204.4121	151.1734	60.4759	43.2520	34.0901	14.0951	2.9267	2.4424	5.0014	6.6848	
53.0098	88.4173	174.9984	201.5076	78.6281	2.3305	5.3937	10.4946	4.8297	0.5908	1.7863	5.4531	2.8378	
33.7177	52.3501	145.4932	124.3960	30.1157	7.2521	5.0241	1.9911	0.1611	1.8522	5.3142	5.3356	2.2422	
36.9643	38.5261	119.8318	97.1323	14.4706	1.9487	1.7095	1.2426	0.9210	0.7080	6.2408	10.5666	3.9319	
78.6444	99.9212	106.2530	94.1907	43.7513	5.6878	1.7246	2.9322	1.2438	0.8658	6.6639	17.0191	19.1892	
31.1381	7.4352	31.8679	34.0426	24.7863	19.5389	9.5018	2.7192	1.6121	0.9870	2.9267	13.8171	24.5500	
5.3516	9.9148	21.9350	20.8068	9.0552	35.1998	18.5498	2.4414	5.7761	5.1105	6.3246	2.8725	5.2528	
5.4531	17.9503	63.9798	72.2236	56.8489	41.8026	43.5667	30.7188	19.1833	12.6395	9.4383	5.7040	7.7251	
37.8541	77.6278	89.4204	56.8925	16.2788	8.2879	16.6163	33.2790	19.1017	1.3775	5.4309	18.4910	12.0959	
31.3595	52.7475	57.4585	16.0989	0.1239	3.6173	5.3442	4.9691	14.0042	13.5670	9.7814	12.7914	16.3399	
14.5864	0.1268	30.6222	56.4337	37.2755	34.2497	17.0627	13.4634	22.7216	11.4230	0.1113	4.0223	6.0435	
4.1262	17.2123	58.9490	99.5420	77.6291	41.7172	34.5546	34.9838	37.9341	27.2577	10.6791	2.3248	2.3842	
7.3616	33.9995	86.7640	82.9560	23.6272	2.4623	13.6872	3.4276	17.2473	49.8546	41.0178	16.6207	1.7174	
25.9466	49.7892	48.6179	18.2532	16.9155	24.3369	22.0499	13.4818	9.6211	27.5862	55.8822	39.1359	16.1690	
14.7376	4.2392	20.5101	20.9139	22.1813	69.1472	65.0358	28.7716	4.2956	22.9832	30.2592	19.4878	14.5113	
6.1490	7.9229	23.1594	63.9437	86.0517	58.5095	42.2003	29.0672	25.5084	26.3120	13.5428	0.7072	3.1479	
3.0026	7.3271	54.1509	71.6563	31.8640	12.3355	0.6214	8.2710	2.0488	7.6243	26.1042	20.3401	7.4059	

Figure V - B. Example of power spectrum density matrix. Each column represents one segment in which power is calculated.

As demonstrated, *ps* is an a × b matrix with each column indicating the power in a particular window and each row indicating a particular frequency. It is worth noting that since Fourier transform operates by transforming time domain information into frequency domain information, technically *f* and *tc* cannot be retrieved at the same time. *tc* in this calculation actually represents the middle time-stamp of the input blood pressure segment. To be specific, we performed windowing to generate each segment, which is discussed later.

One thing worth noting is that the Matlab documentation of the *spectrogram()* function (Matlab, 2006) is not clear. In particular, it did not explain how it calculates the underlying Fourier transform and glossed over the details of what is exactly is calculated to get the resulting power spectral density. To achieve a better understanding of how PSD is actually calculated in the *spectrogram()* function since this is arguably the most important part of this project, we decided compose a naïve indigenous power spectral function based on simple *fft()* to test the concept. In essence, we adopted the discrete

version of the following formula under the assumption that the total power in our signal is finite and hence the Parseval theorem is applicable:

$$psd = \int_{-\infty}^{\infty} |x(f)|^2 df$$

where

$$x(f) = \int_{-\infty}^{\infty} e^{-2\pi i f t}\, x(t) dt$$

By comparing results calculated by the built-in spectrogram() function and our naïve function, we arrived at the conclusion that although slight differences exist in the exact frequencies that were chosen to calculate the PSD, the results are generally in line. Hence, we are cleared to proceed by implementing the built-in function as it is generally computationally faster (see Figure V – C for example of a spectrogram).

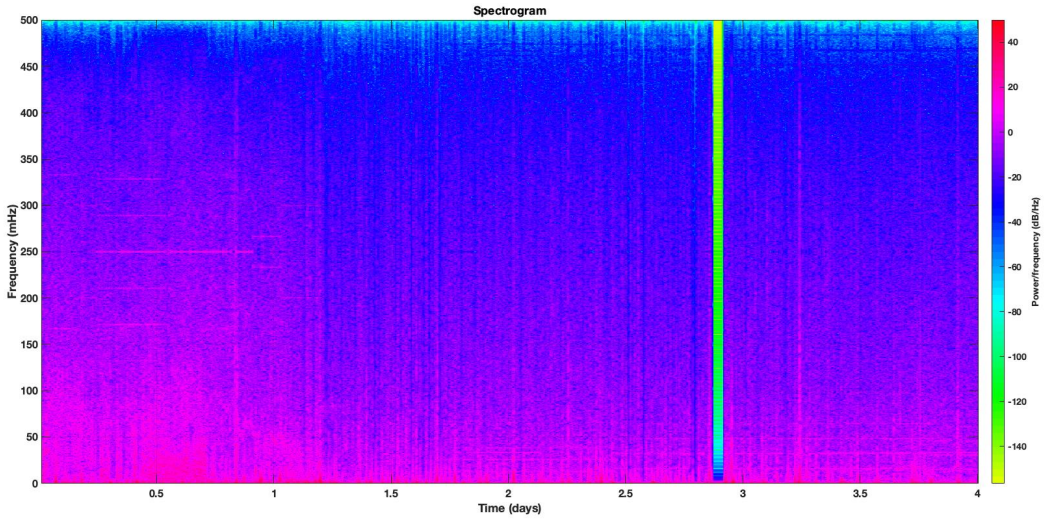


Figure V - C. Example of a full-scale spectrogram. The horizontal axis denotes time since start of data and the vertical axis denotes frequency. This spectrogram is done with a Hann window, which will be discussed below.

To ensure that the PSD calculated is not dramatically influenced by the spikes and remaining artifacts in the blood pressure signal, we decided to impose windowing with

overlaps. We imposed a Hann window (see Figure V - D) on each segment of blood pressure input. A Hann window has the following characteristics:

$$w(n) = \frac{1}{2}(1 - cos\ (\frac{2\pi n}{N}))$$

where w(n) are the corresponding coefficients of the window with size N. The size of the Hann window will be discussed below.

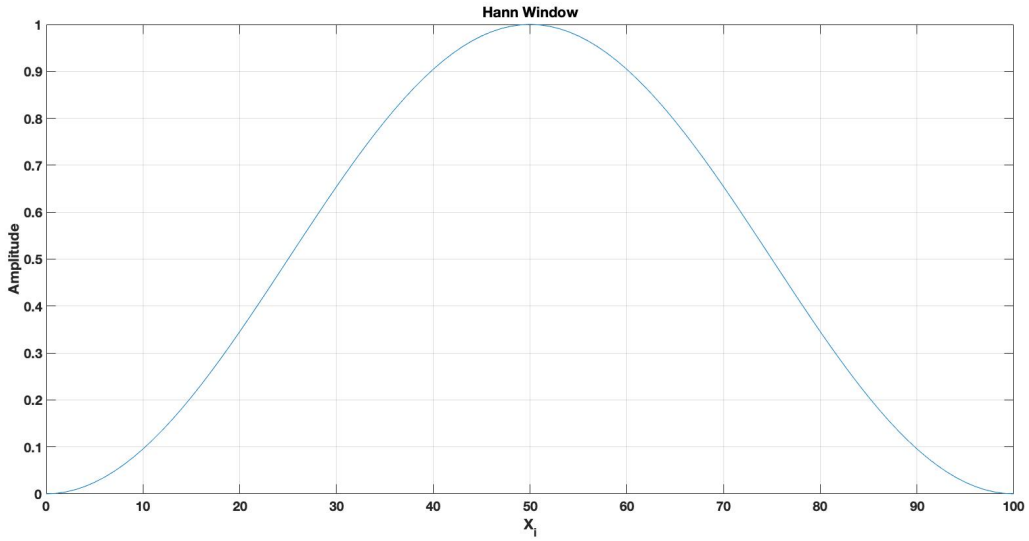


Figure V - D. Example of a Hann window of size 100.

b. **Power within each Frequency Band.** Although we were primarily interested in fluctuations in Band D: 105 ~ 300 seconds, we decided to calculate the power spectrum for all bands identified in Chapter IV. This requires our window length to vary according to the frequency band in question since overly-large or small windows might have negative effects such as masking or spiking on the outcome. For each frequency Set identified in Chapter IV, we defined corresponding window length:

$$l_a \cong 10 * minfreq(\{set_a : Band_\alpha, Band_\beta, Band_\gamma \ldots\})$$

To be specific, the window for each 'Set' is defined as:

- Window size and overlaps for Set 1: 30-minute window moving 5 minutes ahead each;
- Window size and overlaps for Set 2: 5-hour window moving 5 minutes ahead each;
- Window size and overlaps for Set 3: 12-hour window moving 30 minutes ahead each;

We calculated three different PSD for each baby according to the abovementioned three sets of window length and overlaps, with the results labeled *PS_1, PS_2,* and *PS_3*. Note that at this moment, *PS_1, PS_2,* and *PS_3* are three complete power spectra with different window lengths and they each include data for all 10 frequencies. To isolate each frequency band, the following procedure is needed:

- Determine which 'Set' the frequency band in question belongs to and find the corresponding *PS* data with correct window length;
- Identify the upper and lower boundaries u and l in terms of Hz of the frequency band in question;
- Find the corresponding upper and lower frequencies in frequency variable *f* and establish a row interval. Similar to Chapter IV-a, the closest approximate is used when an exact match cannot be achieved;
- Collapse the row interval by summing the power within the interval.

$$ps_{sum}(1{:}\,t) = sum(ps(u{:}\,l, t))$$

The resulting 1 × b matrix will contain the sum power within a frequency band at each time stamp. The corresponding time stamp remains to be *tc*.

With this method, we generated, for each baby, 10 power spectra which correspond to the 10 frequency bands identified. Plotting these power spectra with *area()* function plots gives us results similar to the following (see Figure V – E,F,G):

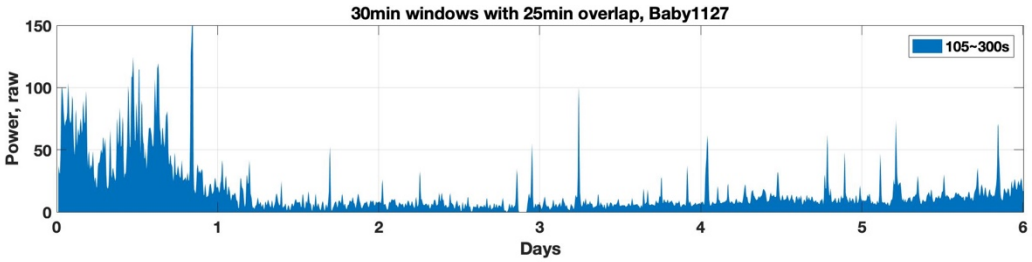


Figure V - E. Example of raw power in Band D: 105 ~ 300s.

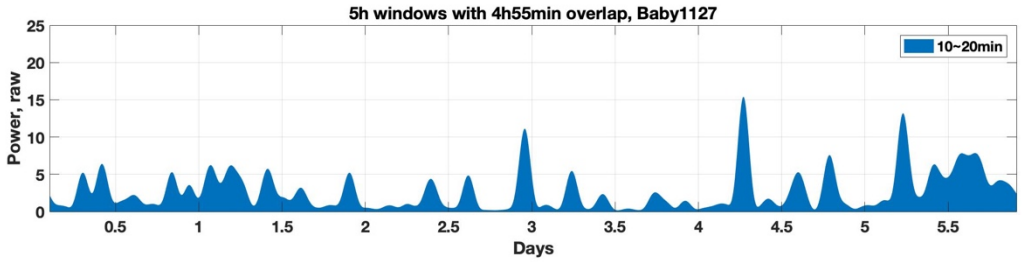


Figure V - F. Example of raw power in Band F: 10 ~ 20min.

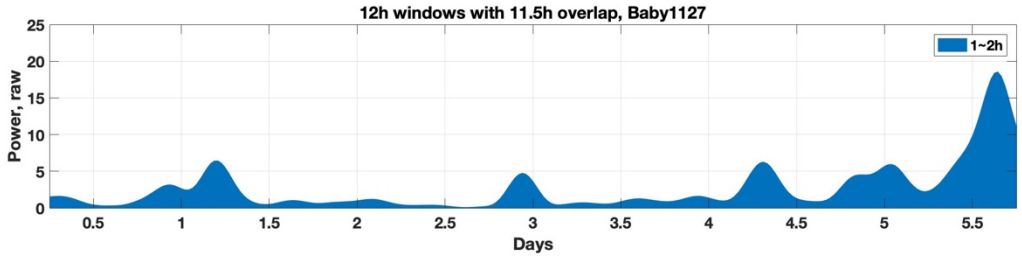


Figure V - G. Example of raw power in Band I: 1 ~ 2h

c. **Averaging & Resampling Power.** After visually inspecting the power spectra generated and HIE severity labels for some babies, we concluded that there are preliminary trends supporting our initial hypothesis stating that the degree of fluctuation is negatively correlated with brain damage severity. In particular, we observe stark differences in

degree of fluctuation in Band D, which we demonstrated with the following two examples (see Figure V – H, I):

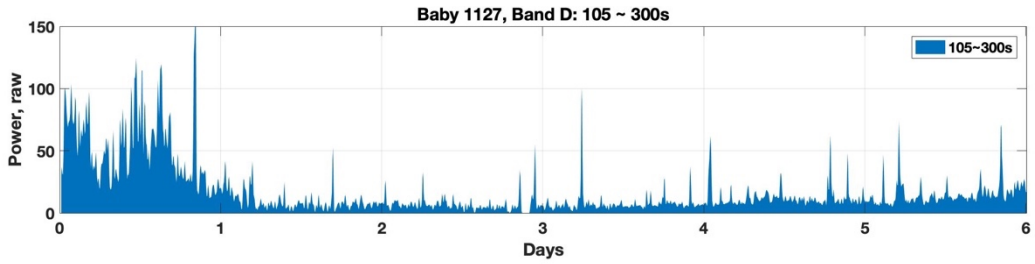


Figure V - H. Band D 105 ~ 300s raw power for Baby NeoID 1127. This baby is classified as Label 3 – Severe HIE.

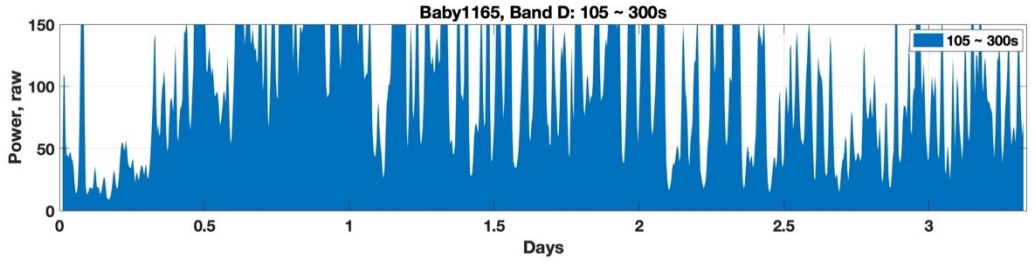


Figure V – I. Band D 105 ~ 300s raw power for Baby NeoID 1165. This baby is classified as Label 1 – Normal/Mild.

To better display the differences between Group 1 and Group 3, we further process the power spectra by smoothing out sudden spikes and grooves. Smoothing was carried out by taking the arithmetic mean of the power spectrum in each band across a six-hour window moving one hour ahead in each step. The resulting curve will be referred to as "raw average power" (see Figure V - J). This technique helps us trace the silhouette of the power spectrum without being distracted by brief and sudden changes.

To compare the difference in power in different babies, we need to harmonize the time stamps. The Matlab built-in *resample()* algorithm was used to linearly resample the *raw average power* at each hour mark within the first week, resulting in a 1×168 one-

dimensional array for each baby. This array will be referred to as the *average power* of each baby and will be utilized in the clustering process discussed in the next part.

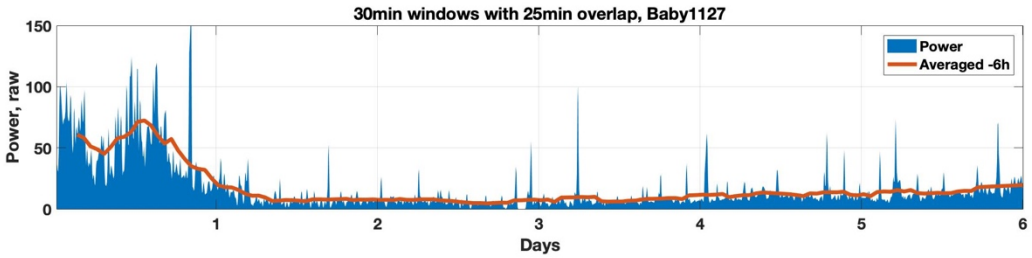


Figure V – J. Example of raw averaged power in Band D. Note that the orange raw averaged power line traces the outline of the raw power but leaves out details such as the sudden spikes.

VI. Clustering, Kernel Density Estimation & Predictions

Tasked with the mission to construct a prediction algorithm, we first test if *averaged power* can be clustered according to HIE severity. Encouraged by promising visual inspection results suggesting a significant difference in power in Band D (105 ~ 300s), we decided that our main focus will be to attempt to cluster Band D *averaged power* against HIE severity.

a. **Clustering.** Recall that under the Sarnat Staging System, HIE severity can be divided into three different categories: Label 1 – Normal/mild, Label 2 – Moderate, and Label 3 – Severe. We started by simply plotting *averaged powers* against their severity label according to the Sarnat Staging-like EEG-24 labeling supplied by UVa Health (see Figure VI - A).

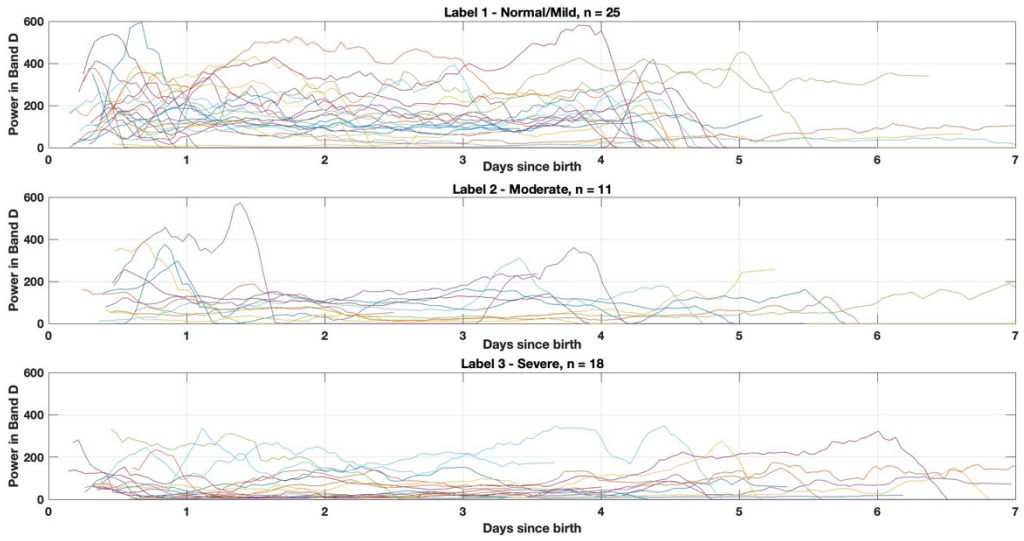


Figure VI - A. Averaged power plotted against HIE severity labels. Note the dramatic difference between Label 1 and Label 3

As seen in the above figure, it is tempting to conclude that Label 1 – Normal/mild and Label 3 – Severe are two different clusters. Label 2 – Moderate seems to exhibit patterns of both Label 1 and Label 3 and is thus hard to interpret.

An obvious drawback to the above figure is that the vertical lines which signal the end of the *averaged power* data as well as the random fluctuations across the week made it hard to identify the exact boundaries or patterns for clustering. To better discern the three groups from each other, we utilize a box plot to demonstrate the general trend while leaving out the individual fluctuations. In particular, the Matlab built-in *boxplot()* function is called, and a box plot is generated at each hour mark showing the spread of *averaged power* within each group for the entire first week (see Figure VI - B).

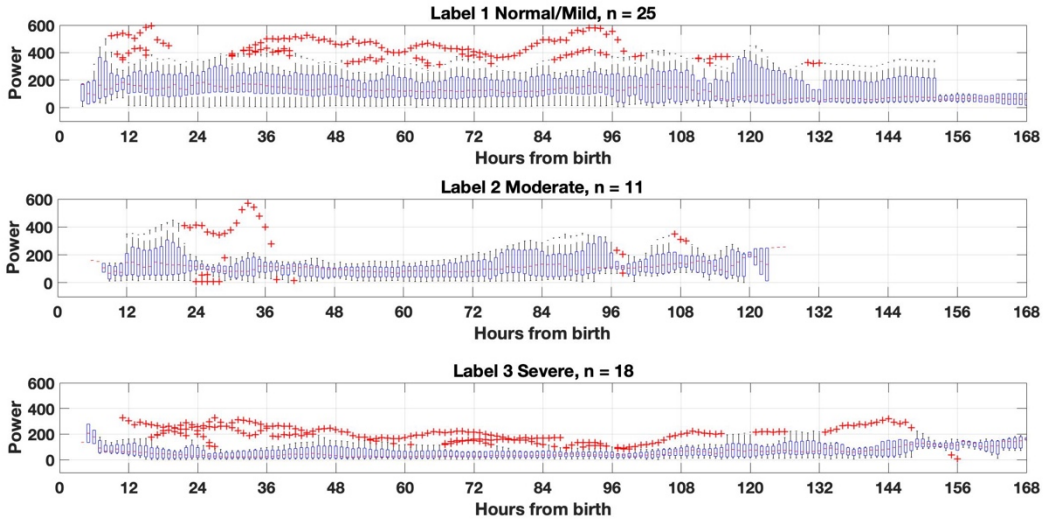


Figure VI - B. Box plot for averaged power in each Label throughout the first week. Red crosses indicate outliers identified by Matlab; red dashes in the box indicate the median value, each box indicate the 25th ~ 75th quartile range, and the whiskers signify the range of valid data-points.

With the box plot above, we were able to verify our hypothesis that there is a clear separation between Label 1 and Label 3. Without assuming normality, additional ranked-sum tests further supported this claim. In particular, we identified two interesting patterns which we would like to further investigate and could be valuable for constructing a prediction algorithm:

- The difference in trend in Label 1 vs. Label 3 before the 8th data point (5th ~ 11th hour averaged power): While *averaged power* exhibits a rise and fall pattern near the start of the first day in Label 1, *averaged power* consistently drops with no obvious rise in Label 3.
- The difference in absolute value in Label 1 vs. Label 3: While *average powe*r in Label 1 has a median of approximately 125 across the entire week, *average power* in Label 3 has a median of only approximately 25. The contrast is especially obvious towards the end of the 24th data point (21st ~ 27th hour averaged power).

To verify the whether the trend identified in the first point of interest could be helpful, we carried out further calculations. In particular, we calculated:

$$ratio = \frac{p(t_1)}{p(t_2)}, t_1 < t_2 \leq 8$$

where p(t) is the averaged power for each baby. We concluded that although the trend might be valid, the exact location at which the *averaged power* rises and falls differ substantially across babies and thus a simple ratio calculation will not be useful. Due to time constraints, we decided to table this route and moved on to examine whether we could base our predictions on the second point of interest.

b. **Kernel Density Estimation (KDE).** One common technique that is frequently used to discern samples collected from hypothetically different distribution is to perform a Kernel Density Estimation. A KDE on our data will be particularly beneficial since it will give us a more intuitive understanding of how data with different Labels are distributed. To construct a KDE, we once again relied on the Matlab built-in *ksdensity()* function. According to Matlab documentation (MATLAB, 2006), *ksdensity()* estimates the

probability density function(pdf) of a random variable. In particular, *ksdensity()* calculates:

$$p_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

where $x_i = [x_1, x_2, x_3, \dots, x_n]$ is the input sample, K(x) is the kernel function, and h represents the corresponding bandwidth. For our purpose, we chose to fit our data with an Epanechnikov distribution, which can be calculated by:

$$K(u) = \frac{3}{4}(1 - u^2), |u| < 1,$$

and K(u) = 0 otherwise.

Before we proceed to fit our data with Epanechnikov kernels, some pre-processing is needed. For each severity Label, we collected *averaged power* for all babies within that group, resulting in a $n \times 168$ matrix, with n being the number of babies with that label. We then reshaped the abovementioned matrix into a $1 \times m$ one-dimensional array and removed NaNs and 0s which we originally used to signal missing data. We then considered each data point in this $1 \times m$ array as an independent data point, hence collapsing the time and baby dimensions of our data. After this procedure, we were able to fit the abovementioned Epanechnikov kernel onto our data by simply calling the *ksdensity()* function (see Figure VI – C).

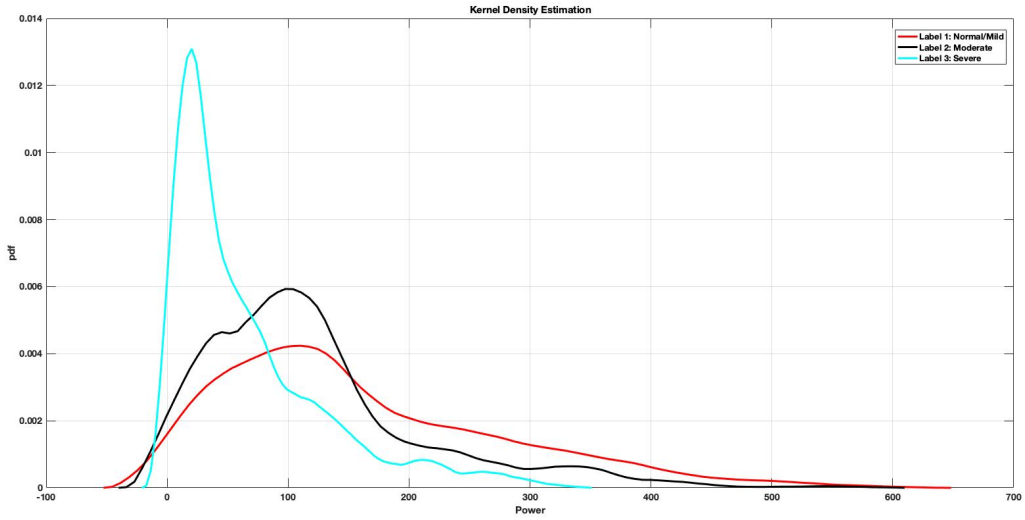


Figure VI - C. Kernel Density Estimation based on the entire week's data. The horizonal axis represents averaged window and the vertical axis represents corresponding probability.

The KDE analysis once again supports our previous conclusion that most data points in Label 1 are distributed around 125, most data points in Label 3 are distributed around 25 and data points in Label 2 exhibit properties of both Label 1 and 3.

c. **Prediction.** Based on these probability density functions, we were able to construct the probability that a given observed power will correspond with a particular label. In particular, according to Bayes theorem we may compute:

$$p(label\ \alpha|power) = \frac{p(power|label\ \alpha)p(label\ \alpha)}{p(power)}$$

After consulting medical practitioners, we were informed that a prediction based on data collected within approximately the first 24 hours will be most useful since EEG and MRI examinations are usually only carried out after the 24th hour mark and very few biomarkers have been identified to reliably predict HIE severity well within the first 24 hours of birth. Hence, we decided that rather than utilizing the entire week's data, it would be more useful to construct density distributions using only data around the 24th

hour mark and construct a prediction based on these pdfs only. To be specific, we examined the following two periods:

- $18^{th} \sim 24^{th}$ data point, or $15^{th} \sim 27^{th}$ hour due to averaging;
- $24^{th} \sim 30^{th}$ data point, or $21^{st} \sim 33^{rd}$ hour due to averaging.

Adopting the same steps we used to construct KDE for the entire week's data, the resulting KDE based on the $18^{th} \sim 24^{th}$ data point is (see Figure VI - D):

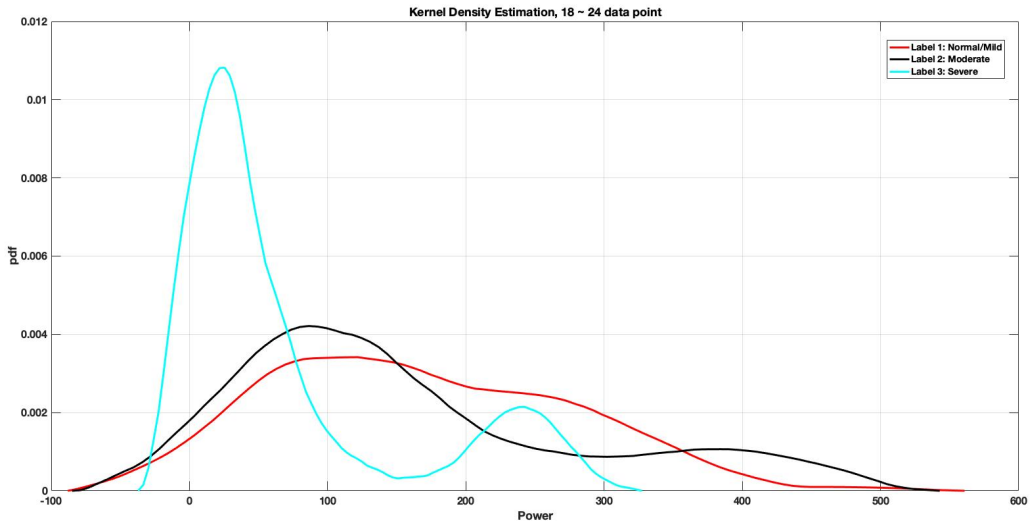


Figure VI - D. Kernel Density Estimation based on the $18^{th} \sim 24^{th}$ data point. The horizonal axis represents raw power and the vertical axis represent the corresponding probability density. The general trend follows our expectation.

Combining this KDE and the abovementioned Bayes theorem allows us to predict the probability a given power will fall into each Label within this period of time (see Figure VI - E).

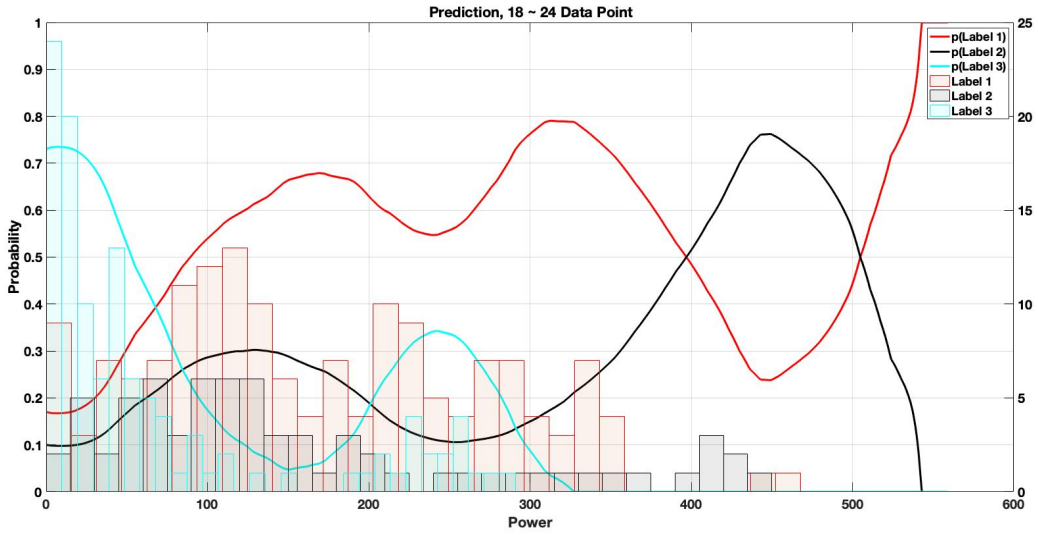


Figure VI – E. Prediction functions plotted above histogram based on 18th ~ 24th data point. The left vertical axis denotes the probability a power lies in a particular group, and the right vertical axis denotes number of instances each power is observed in each label.

We overlaid a histogram representing data distribution within this period over the prediction curves to emphasize the difference in strength of our prediction at different powers. As expected, a prediction based on more data (higher histogram value) is less susceptible to artifacts and random fluctuations than a prediction based on relatively less data. Bearing in mind this difference, we concluded that predictions at power above 300 are less meaningful compared to predictions at power 0 ~ 300. We were able to further conclude that by reading off the figure, 80% of observations with an *averaged power* of 25 came from babies with Label 3, and very few observations with an *averaged power* of 125 came from babies with Label 3.

We then constructed KDE for the 24th ~ 30th data point. With the exact same methodology, we were able to achieve the following KDE (see Figure VI - F):

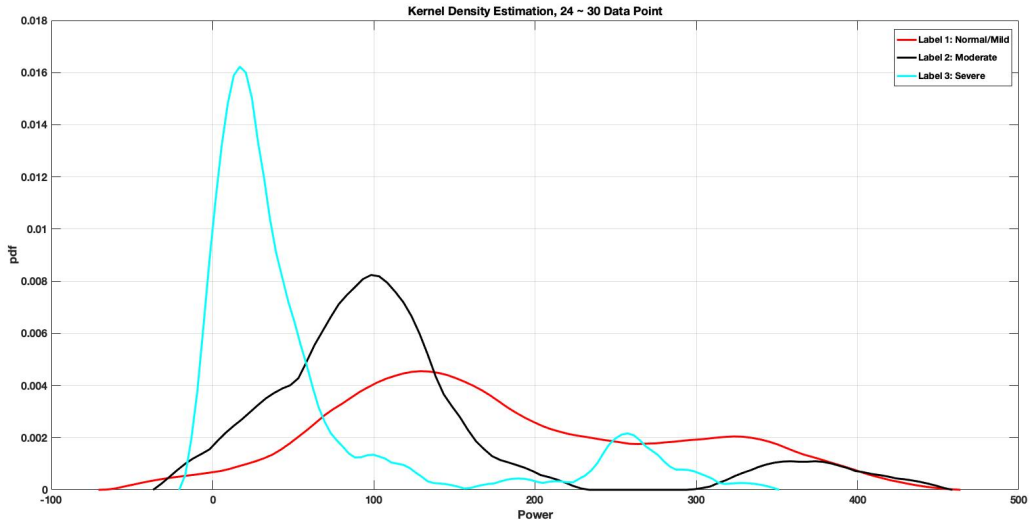


Figure VI – F. Kernel Density Estimation based on the 24th ~ 30th data point. The horizonal axis represents raw power and the vertical axis represent the corresponding probability density.

And the corresponding prediction (see Figure VI – G):

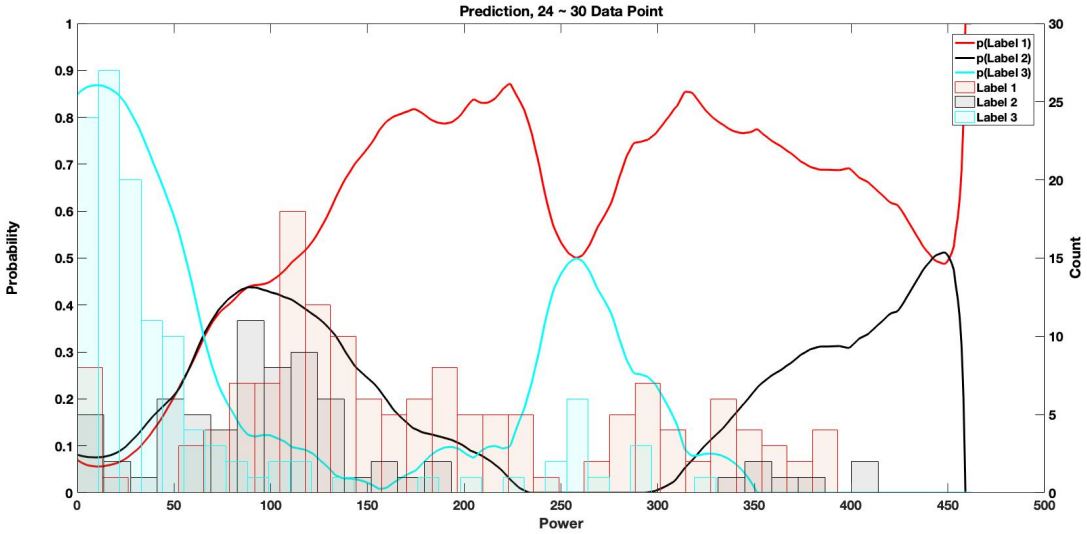


Figure VI – G. Prediction functions plotted above histogram based on 24th ~ 30th data point. The left vertical axis denotes the probability a power lies in a particular group, and the right vertical axis denotes number of instances.

Again, by reading of the figure, a baby with an averaged blood pressure power of 25 within this period has a near 90% probability of being in Label 3, and a baby with a blood pressure power of above 150 is very unlikely to belong to Label 3.

VII. Discussion & Future Work

With this study, we were able to conclude via Fourier Transform and spectrogram analysis that substantial difference in blood pressure fluctuation/power exists among babies with different HIE severity. We established a negative correlation between blood pressure fluctuation/power and HIE severity, correlating babies with more severe HIE to a smaller degree of blood pressure fluctuation/power. Furthermore, we were able to construct a primitive prediction algorithm that will predict how likely an observed averaged blood pressure power will fall into each Label based on Bayesian statistics techniques such as Kernel Density Estimation.

Several limitations to this project exist. First of all, similar to any other research purely based on bed-side monitor data, we cannot rule out the possibility that the abovementioned patterns are partly or completely caused by confounds such as different medications administered to different groups. We were aware of this potential drawback throughout the project and repeatedly confirmed with physicians at UVA Health who collected these data that to the best of their knowledge these differences are not caused by treatment. One way to ensure the validity of these differences would be to pore through individual medical records to verify the exact treatments administered. This process, however, must be left to our clinician partners. Another potential limitation is that our prediction is based on the Bayesian assumption that the ratio of Label 1 vs. Label 2 vs. Label 3 in our sample is representative of the population. The accuracy of this assumption can only be verified by future studies. Concerning the details of this study, work may be done to refine the various windows used in this study as they are mostly empirical rather than based on theory.

To predict HIE severity based on our prediction curves, one must calculate *averaged power* exactly the way we did. We would argue that while the negative correlation between degree of blood pressure fluctuation and HIE severity is most likely valid, our particular *averaged power* measurement is hard to interpret and additional work can be done to simplify the calculation while retaining its essence.

Also, once again due to the time constraint, we were not able to look into how some patterns of pseudo-periodicity of blood pressure we originally identified may be indicative of various pathologies, which was one of our two original motivations. This part will also be left for future studies.

Additionally, future studies may want to explore the predictiveness of the other frequency bands we identified but were not able to examine due to time constraint. A long shot will be to analyze whether incorporating other bed-side monitor data such as heart rate with blood pressure will yield even more accurate predictions. Also, since this study is purely data driven, it offers no physiological explanation whatsoever. It would very beneficial to explore the underlying physiological mechanisms that prompted the difference in blood pressure fluctuation/power we've seen. Finally, one would always want to explore the possibility of training a machine learning model such as Long-short Term Memory or Convolutional Neural Network based on this dataset given additional data.

VIII. References

Aboab, J., Polito, A., Orlikowski, D., Sharshar, T., Castel, M., & Annane, D. (2008, May). Hydrocortisone effects on cardiovascular variability in septic shock: a spectral analysis approach. *Critical Care Medicine, 36*(5), 1481-1486.

Airede, A. (1991). Birth asphyxia and hypoxic-ischaemic encephalopathy: incidence and severity. *Annals of tropical paediatrics, 11*, 331-335.

Badawi, N., Kurinczuk, J., Keogh, J., Alessandri, L., O'Sullivan, F., Burton, P., . . . Stanley, F. (1998). Intrapartum risk factors for newborn encephalopathy: the Western Australian case-control study. *BMJ, 317*(7172), 1554-1558.

Douglas-Escobar, M., & Weiss, M. D. (2015). Hypoxic-Ischemic Encephalopathy: A Review for the Clinician. *JAMA Pediatrics, 169*(4), 397-403.

Hankins, G., & Speer, M. (2003). Defining the pathogenesis and pathophysiology of neonatal encephalopathy and cerebral palsy. *Obstetrics & Gynecology, 102*(3), 628-636.

Harteman, J. C., Nikkel, P. G., Benders, M. J., Kwee, A., Groenendaal, F., & Vries, L. S. (2013, 10). Placental Pathology in Full-Term Infants with Hypoxic-Ischemic Neonatal Encephalopathy and Association with Magnetic Resonance Imaging Pattern of Brain Injury. *The Journal of Pediatrics, 163*(4), 968-995.

Hull, J., & Dodd, K. (1992). Falling incidence of hypoxic-ischaemic encephalopathy in term infants. *BJOG: An International Journal of Obstetrics, 99*(5), 386-391.

MATLAB. (2006). *fft*. Retrieved 04 2019, from Matlab Documentation: https://www.mathworks.com/help/matlab/ref/fft.html

MATLAB. (2006). *Kernel Distribution*. Retrieved from Documentation: https://www.mathworks.com/help/stats/kernel-distribution.html#btwp98t

Matlab. (2006). *Spectrogram()*. Retrieved from Documentations: https://www.mathworks.com/help/signal/ref/spectrogram.html#bultmx7-f

Stokowski, L. (2005). Ensuring safety for infants undergoing magnetic resonance imaging. *Advances in Neonatal Care, 5*(1), 14-27.

Thornberg, E., Thiringer, K., Odeback, A., & Milsom, I. (1995, 8). Birth asphyxia: incidence, clinical course and outcome in a Swedish population. *Acta Pediatrica, 84*(8), 927-932.

Vesoulis, Z., Mathur, A., & Mcpherson, C. (2017). Low-frequency blood pressure oscillations and inotrope treatment failure in premature infants. *Journal of Applied Physiology, 123*, 55-61.

Volpe, J. J. (2001). *Neurology of the Newborn.* Philadelphia: WB Saunders Company.