`

Optimization of Biosensor Waveform Preprocessing

Siemens Healthineers

Report Week 5

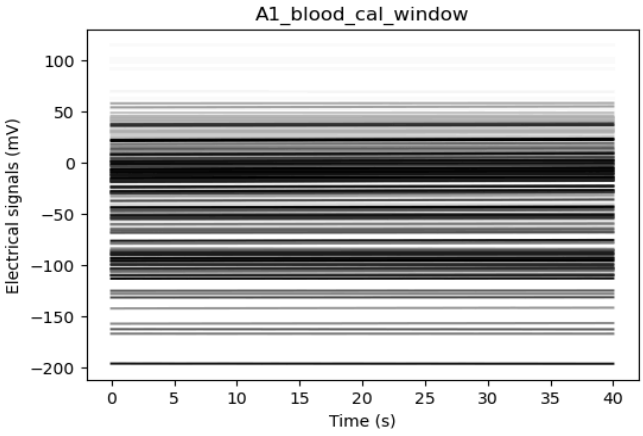
Dylan Longert
Jinxin (Jessie) Wang
Nan Tang
Nayeli Montiel Rodríguez
Han (Eden) Chen

`

1. Scale Approaches Comparison

We conduct a scaling comparison as a preliminary step before employing Functional Principal Component Analysis (FPCA). This approach is prompted by our uncertainty regarding the centering parameter in the 'FPCA()' function of the 'scikit-fda' library. Furthermore, during our recent meeting with Siemens, there was interest in interpreting the components following data scaling to ensure comparable ranges between both systems. Consequently, we explore four different scaling methods as preprocessing steps and subsequently apply FPCA to the preprocessed data to evaluate differences.

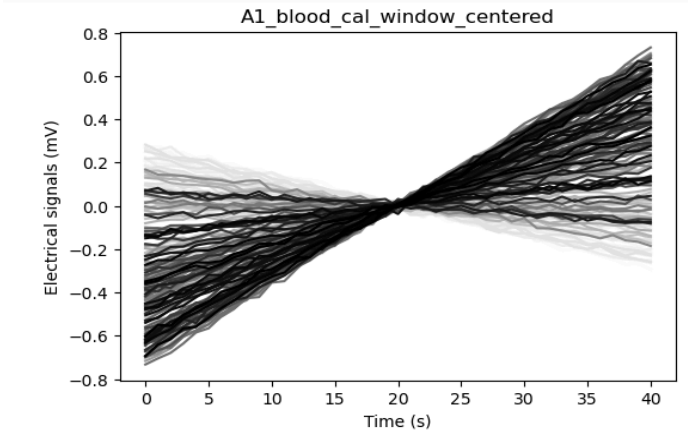
Baseline: Raw Data



The raw data is compressed by waveforms with positive and negative values at different levels. For instance, the figure shows waveforms in System 1 - Sensor A - Blood fluids post-extraction from the calibration window. The range of the measurements in millivolts goes from -200 to 100.

Figure 1. Raw waveforms.

Row-Centered



Standardizing data by subtracting the row mean, calculated for each waveform labeled by TestID, maintains the original shape of the waveforms while ensuring comparability between System 1 and System 2. Additionally, the variance captured by the first component is over 98%. However, the interpretation of the components may be challenging.

Figure 2. Waveforms row-wise centered.

`

For instance, consider the following case where the mean function has a positive trend while component 1 has a negative trend.

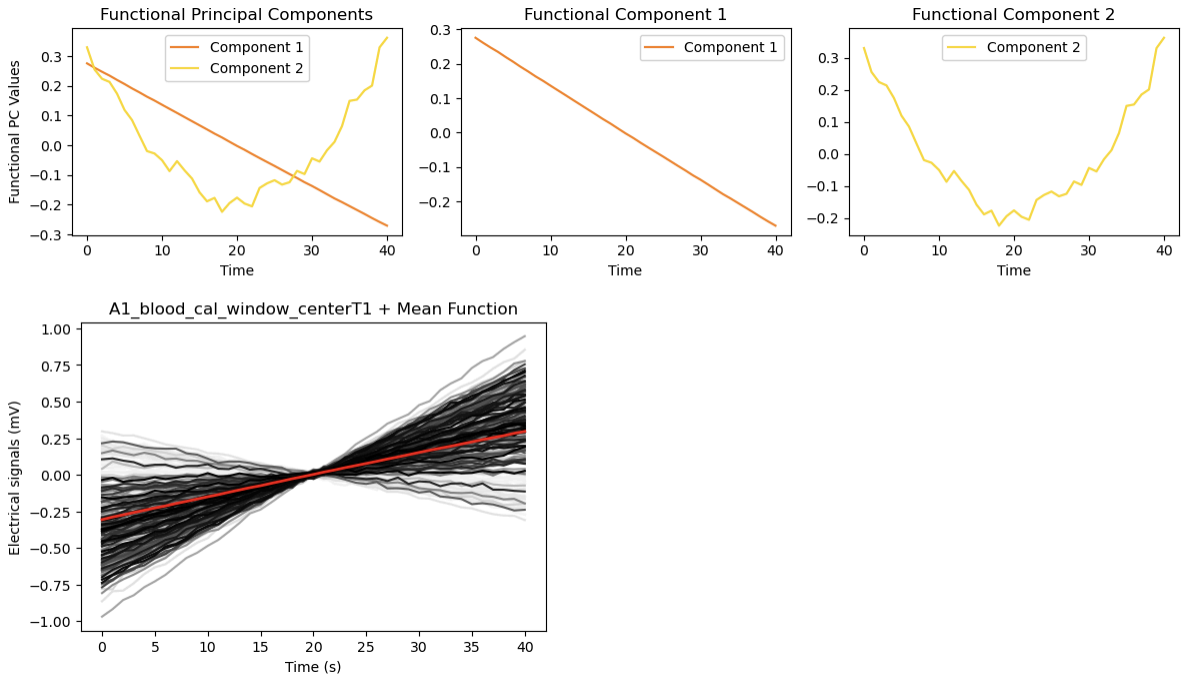
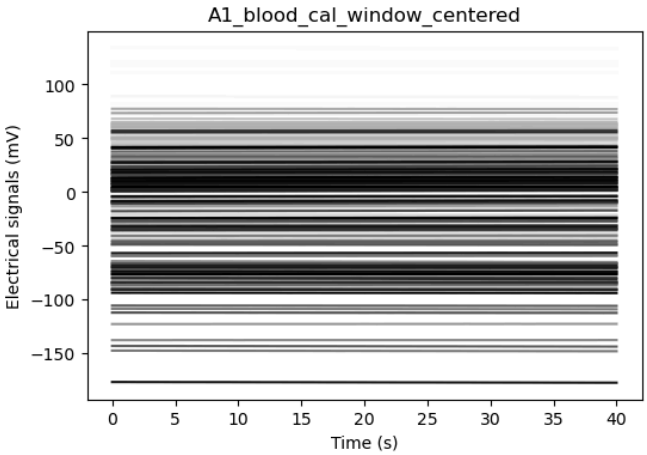


Figure 3. Principal Components after the row-wise centering.

Column-Centered



Centering the data by subtracting the column mean (the mean of test IDs for each time stamp) parallelly moves the centered data and shrinks the range of the sensor values while maintaining its shape and variation. The variance captured by the first component is over 99%. However, the FPCA scores for each binned feature applied with this centered data don't have any interesting or meaningful findings which can be interpreted.

Figure 4. Waveforms column-wise centered.

`

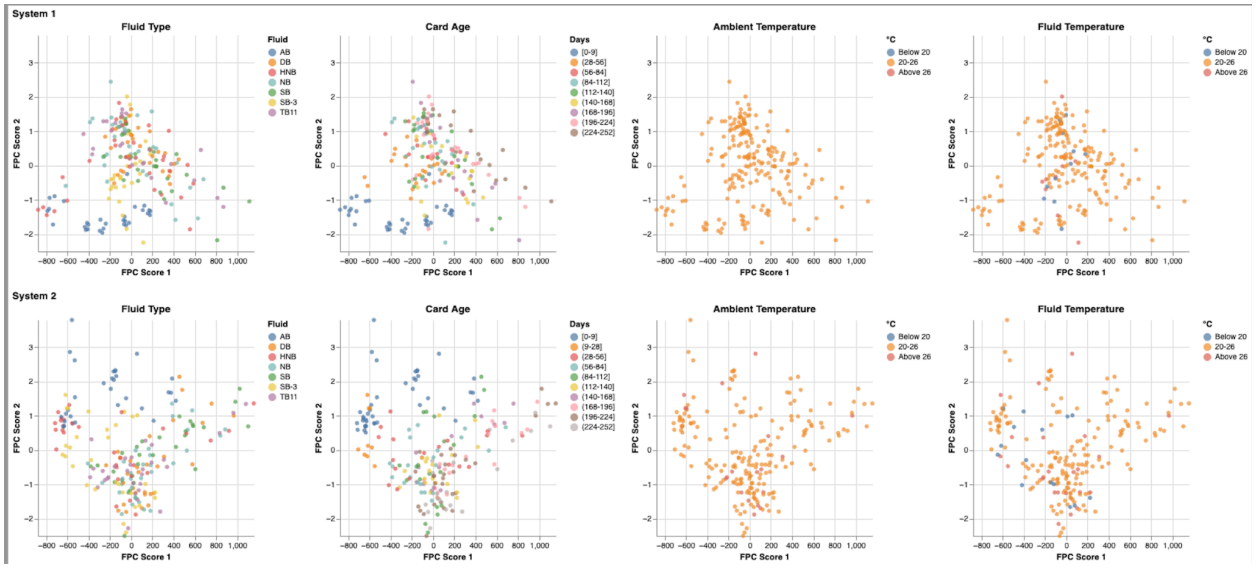
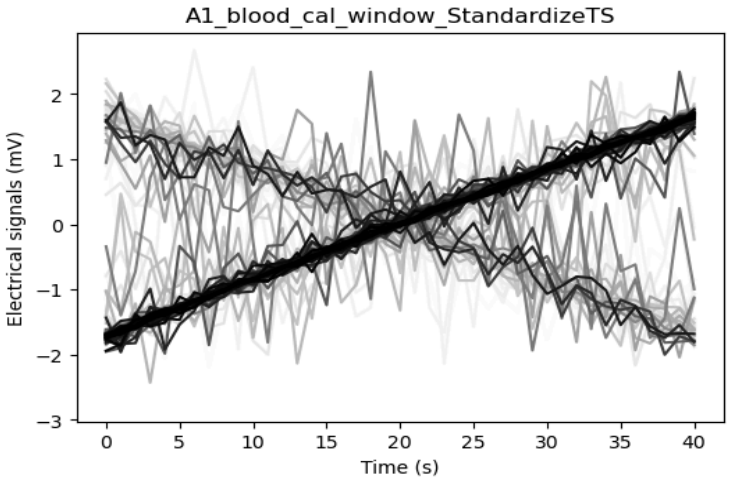


Figure 5. Scatterplots from the PC scores after column-wise centering.

Row-Standardized



Standardizing data along rows(timestamps) can keep the original shape of waveforms and eliminate the impact of scales in all waveforms. However, PC1 calculated based on this kind of scaling data fails to explain the most of variance in data, especially for Sensor B (less than 80% in most cases).

Figure 6. Waveforms row-wise standardization
after subtracting the mean and dividing by the standard deviation.

For Sensor A, System 1 and System 2 are very similar in PC1 values; For Sensor B sample Window, the distribution of aqueous types in PC1 and PC2 shows some patterns: Eurotrol L5 is different from other aqueous.

`

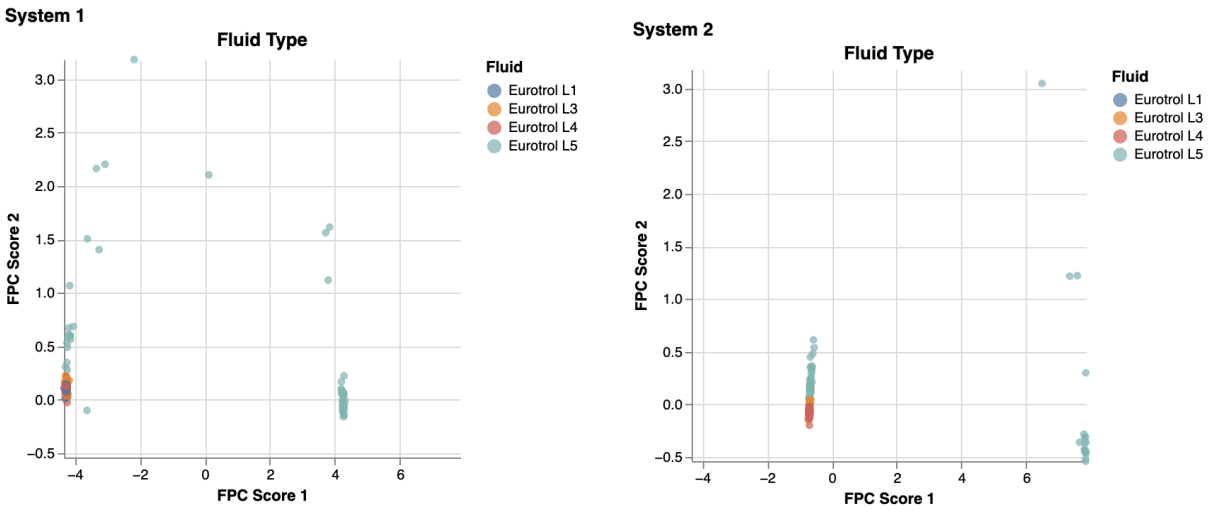
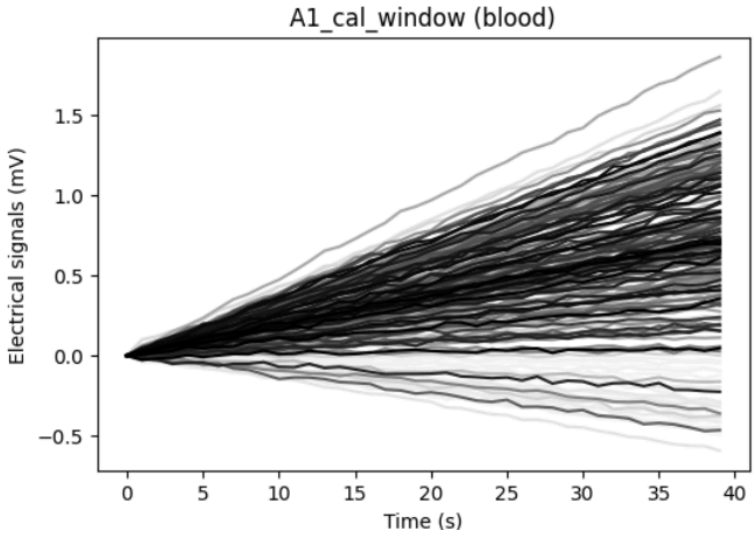


Figure 7. Scatterplots from the PC scores after row-wise standardization.

Zero-Aligned



Aligning data with the same starting points indeed provides a clearer representation of waveform shapes by minimizing the influence of magnitude differences. This approach can enhance the interpretability of the data, making it easier to compare and analyze waveforms. While aligning by waveform center may capture slightly more information (99.99% versus 99.87% for Component 1).

Figure 8. Align the same start points.

As shown in Figure 6, System 2 displays significant groupings, while System 1 does not. This observation contrasts with previous results when different centering methods were used.

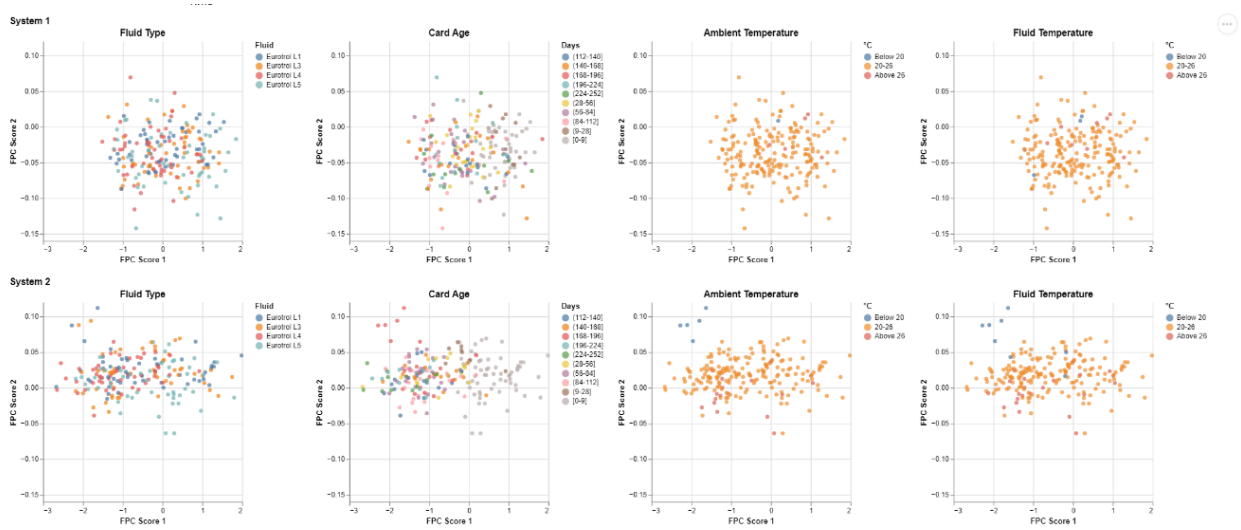


Figure 9. Scatterplots from the PC scores after zero aligning.

2. Simple Linear Regression on Component 1

As Table 1 shows, for any kind of window and sensor, the aqueous samples exhibit the same slope values in both system 1 and system 2. However, for blood samples, the slopes are the same only in the sensor B calibration window, while in other cases, the slopes are nearly opposite.

On the other hand, for the same sensor and window, the magnitude of the slopes appears to be the same across different systems.

Table 1: The slope of Component 1 in System 1 & 2 (Before aligning the same start point)

	Window	Slope 1	Slope 2
0	A_aqueous_cal_window	-0.007052	-0.007054
1	A_blood_cal_window	0.007061	-0.007074
2	B_aqueous_cal_window	-0.002035	-0.002000
3	B_blood_cal_window	-0.002042	-0.002016
4	A_aqueous_sample_window	-0.014822	-0.014594
5	A_blood_sample_window	-0.014746	0.014603
6	B_aqueous_sample_window	-0.020267	-0.020199
7	B_blood_sample_window	0.020699	-0.020958

`

3.Slope on raw data

For sensor A, the slopes within the calibration window for System 1 and System 2 across bins are close. The slopes across bins within the sample window are greater for System 2 than for System 1. For sensor B, the slopes within both the calibration and sample window across bins are also greater for System 2 than System 1.

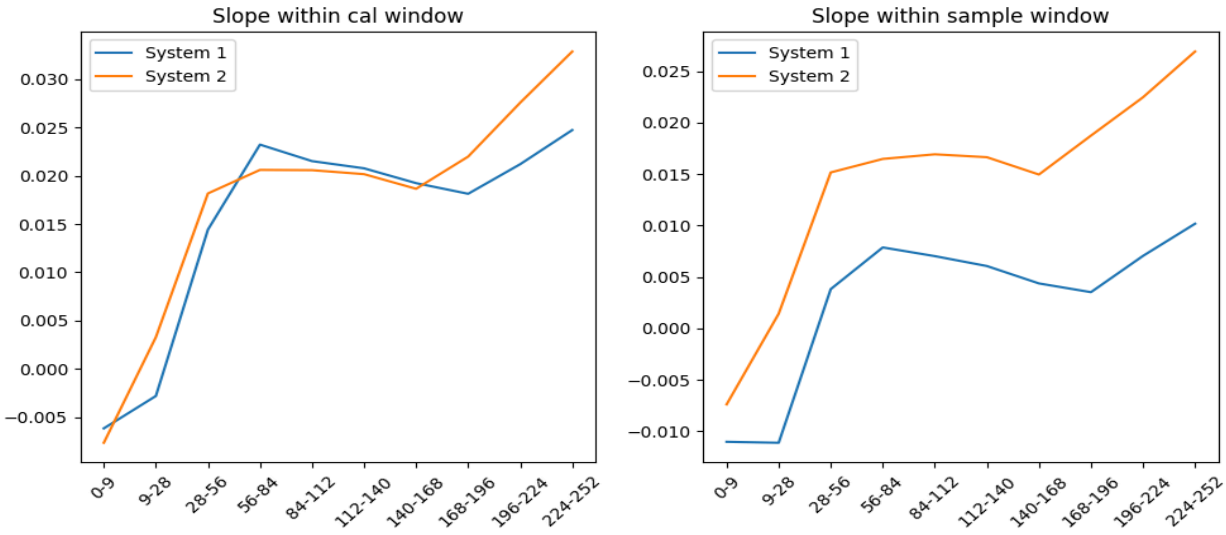


Figure 10. Sensor A slope of the raw sensor data binned by card age with samples aggregated by mean

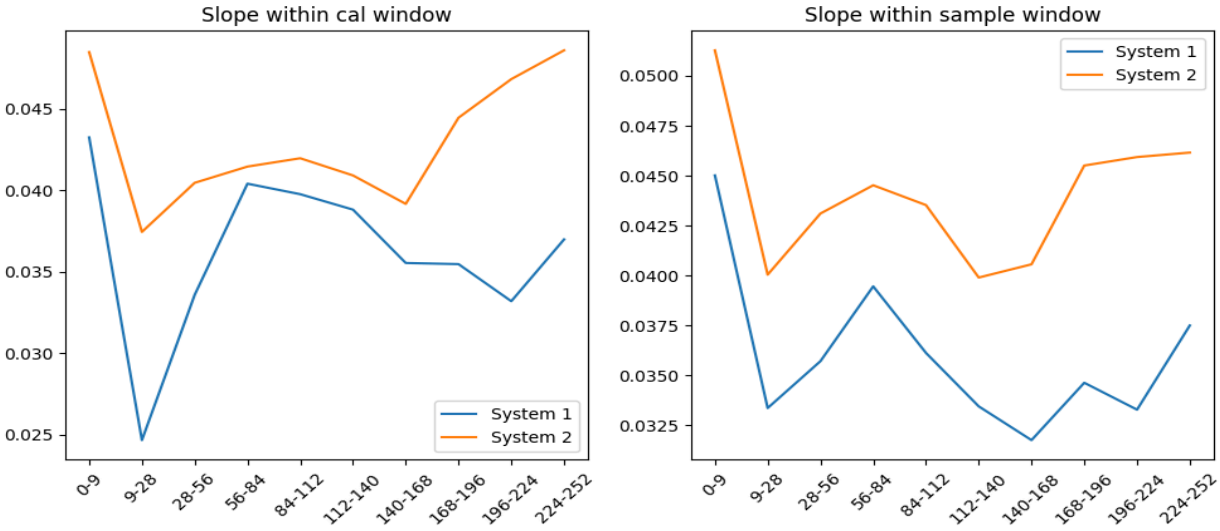


Figure 11. Sensor B slope of the raw sensor data binned by card age with samples aggregated by mean

`

4. Analysis of variance

Eager to quantify the degree of flatness between the waveforms of the two systems, we decided to compare their variances. In signal processing, a system with lower variance tends to produce smoother and flatter waveforms, as discussed in [1].

We employed the data after the window extraction (Table 2) and verified the following key points:

- **Missing values:** The data does not have missing values post-extraction from the windows of interest.
- **Equal length:** All the time series within a group (system, sensor, fluid) have the same length.
- **Outliers:** We removed the unsuccessful tests run before the window extraction.
- Aligning to the same level: We decided to align all the time series to start from zero to facilitate the comparison across systems, Figure 12 (B).
- **Differentiation:** We differentiated the time series by subtracting one lag to remove trend, Figure 12 (C).
- **Stationarity:** We used the Augmented Dickey-Fuller Test to assess stationarity: We label every waveform (TestID) by adding two columns, the first one with the p-value and the second one indicating if the waveform is or not stationary.

  H0: The time series is non-stationary.
  This means that the time series has some form of trend, and its statistical properties (like mean and variance) change over time.
  H1: The time series is stationary.
  This means that the time series does not have a trend, and its statistical properties (like mean and variance) remain constant over time.

  P-value interpretation:

  ❖ If the p-value is < alpha: Reject the null hypothesis. This suggests that the time series is stationary.

  ❖ If the p-value is > alpha: Fail to reject the null hypothesis. This suggests that the time series is non-stationary.
  Being alpha the level of significance equal to 0.05.

- **Proportion of stationary waveforms versus non-stationary waveforms**: We identify the compute number of non-stationary waveforms compared to stationary ones, Figure 13.

`

- **Removal of non-stationary**: We removed the non-stationary waveforms before the analysis.
- **Comparison of variances**: By direct wise row computation and by bootstrapping.

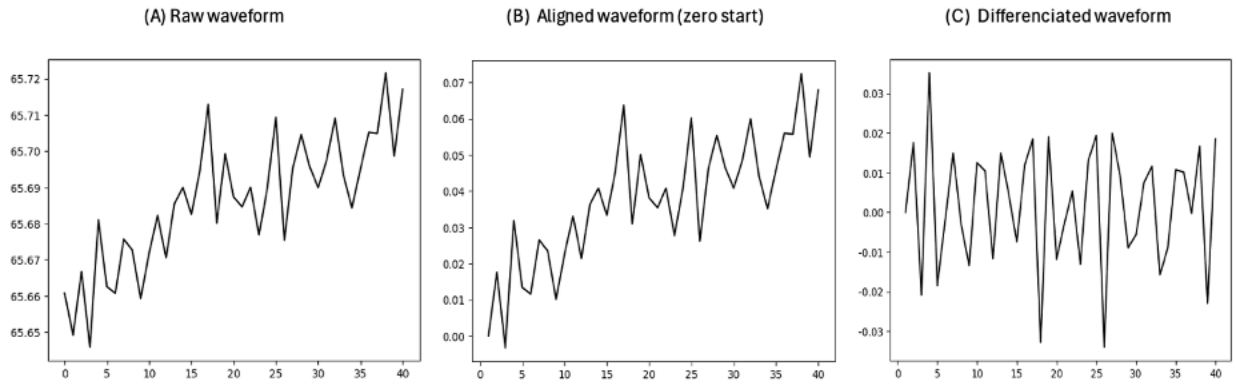


Figure 12. Visualization of a waveform from System 1 Sensor A Blood Fluids extracted from the Calibration Window. (A) The original waveform, displaying values around 65 mV (millivolts) with an observable uptrend. (B) The same waveform, aligned to start from zero. The shape and uptrend are preserved. (C) The differentiated waveform, where the trend has been removed, resulting in fluctuations around zero.

(A) Raw Data

		Cal Window		Sample Window	
System	Sensor	Aqueous	Blood	Aqueous	Blood
1	A	1949 rows, 41 columns	1433 rows, 41 columns	1949 rows, 26 columns	1433 rows, 26 columns
2	A	4892 rows, 41 columns	2851 rows, 41 columns	4892 rows, 26 columns	2851 rows, 26 columns
1	B	1944 rows, 91 columns	1431 rows, 91 columns	1944 rows, 21 columns	1431 rows, 21 columns
2	B	4893 rows, 91 columns	2852 rows, 91 columns	4893 rows, 21 columns	2852 rows, 21 columns

(B) Stationary Time Series

		Cal Window		Sample Window	
System	Sensor	Aqueous	Blood	Aqueous	Blood
1	A	1839 rows, 40 columns	1361 rows, 40 columns	1330 rows, 25 columns	985 rows, 25 columns
2	A	4527 rows, 40 columns	2600 rows, 40 columns	3334 rows, 25 columns	1882 rows, 25 columns
1	B	1780 rows, 90 columns	1288 rows, 90 columns	859 rows, 20 columns	665 rows, 20 columns
2	B	3448 rows, 90 columns	1999 rows, 90 columns	2046 rows, 20 columns	1207 rows, 20 columns

(C) Non-Stationary Time Series

		Cal Window		Sample Window	
System	Sensor	Aqueous	Blood	Aqueous	Blood
1	A	110 rows, 40 columns	72 rows, 40 columns	619 rows, 25 columns	448 rows, 25 columns
2	A	365 rows, 40 columns	251 rows, 40 columns	1558 rows, 25 columns	969 rows, 25 columns
1	B	164 rows, 90 columns	143 rows, 90 columns	1085 rows, 20 columns	766 rows, 20 columns
2	B	1445 rows, 90 columns	853 rows, 90 columns	2847 rows, 20 columns	1645 rows, 20 columns

Figure 13. Dimensionality of waveform groups considered in the analysis of variance. (A) The original waveform shapes post-window extraction. (B) Shapes of stationary

`

waveforms post-preprocessing (zero alignment and differencing). (C) Shapes of non-stationary waveforms post-preprocessing (zero alignment and differencing). The sum of the number of rows in (B) and (C) gives the total number of waveforms in (A). The values in orange highlight a moderate number of non-stationary waveforms when compared with the stationary ones. The values in red highlight a large number of non-stationary waveforms when compared with the stationary ones

Analysis of variances method 1

We computed the variances for each time series within system 1 and system 2, performing row-wise calculations. Subsequently, we compare these variances between the two groups to assess if there are significant differences using Levene's test. The null hypothesis of this test posits that the variances across groups are equal. The test statistic is derived from the absolute deviations of the observations from their respective group means, Figure 14.

(A) Calibration window

```
A1_blood_cal_window vs A2_blood_cal_window

Mean Variance for System 1: 0.00021000504839200157
Mean Variance for System 2: 8.323891907578327e-05
Levene's test statistic: 239.47869533627448
p-value: 1.7108684772224785e-52
There is a significant difference in variances.

B1_blood_cal_window vs B2_blood_cal_window

Mean Variance for System 1: 0.00022125160318817668
Mean Variance for System 2: 9.952619689370222e-05
Levene's test statistic: 9.232078698145637
p-value: 0.0023968288362989534
There is a significant difference in variances.

A1_aqueous_cal_window vs A2_aqueous_cal_window

Mean Variance for System 1: 0.00020579766759835013
Mean Variance for System 2: 7.802873311509115e-05
Levene's test statistic: 281.21127033322205
p-value: 8.544573106001218e-62
There is a significant difference in variances.

B1_aqueous_cal_window vs B2_aqueous_cal_window

Mean Variance for System 1: 0.00022125809964721043
Mean Variance for System 2: 0.00010729698653908612
Levene's test statistic: 0.7527909205665924
p-value: 0.38563388241693763
There is no significant difference in variances.
```

(B) Sample window

```
A1_blood_sample_window vs A2_blood_sample_window

Mean Variance for System 1: 0.00019893377239714572
Mean Variance for System 2: 6.770967792449386e-05
Levene's test statistic: 417.8884567685221
p-value: 8.204096122886265e-87
There is a significant difference in variances.

B1_blood_sample_window vs B2_blood_sample_window

Mean Variance for System 1: 0.0002829478426988557
Mean Variance for System 2: 0.00018505024948763736
Levene's test statistic: 135.9964020322602
p-value: 2.1899850675927717e-30
There is a significant difference in variances.

A1_aqueous_sample_window vs A2_aqueous_sample_window

Mean Variance for System 1: 0.0002264548459440104
Mean Variance for System 2: 0.00011901440435175032
Levene's test statistic: 11.121953071049457
p-value: 0.0008597712005880781
There is a significant difference in variances.

B1_aqueous_sample_window vs B2_aqueous_sample_window

Mean Variance for System 1: 0.00029686726257582526
Mean Variance for System 2: 0.00020690917092400888
Levene's test statistic: 26.088324935015425
p-value: 3.472443220851734e-07
There is a significant difference in variances.
```

Figure 14. (A) Results of the comparison of variances from system 1 versus system 2 in the calibration window for sensors A and B. (B) Results of the comparison of variances from system 1 versus system 2 in the calibration window for sensors A and B.

`

Pros

- We identify that there are differences among the variances of the two systems.

Cons

- Does not provide any measure of uncertainty around these estimates.
- Single point estimate for the mean variance of each system, without information on the stability or variability of this estimate.
- Assumes that the sample mean variance is a good estimator of the population variance. However, this assumption may not hold if the data is not normally distributed or if the sample size is small.

Analysis of variances method 2

We conduct a bootstrap variance comparison between two sets of waveforms (system 1 and system 2). By iteratively resampling from each set, computing the variance for each resampled set of time series, and then calculating the mean and standard deviation of the variances for both systems. Additionally, we compute the 95% confidence intervals for the variances and perform Levene's test to compare the variances of the bootstrap samples, Figure 15. Finally, we display box plots (Figures 16 and 17) to visually compare the distribution of variances between the two systems based on the bootstrap samples.

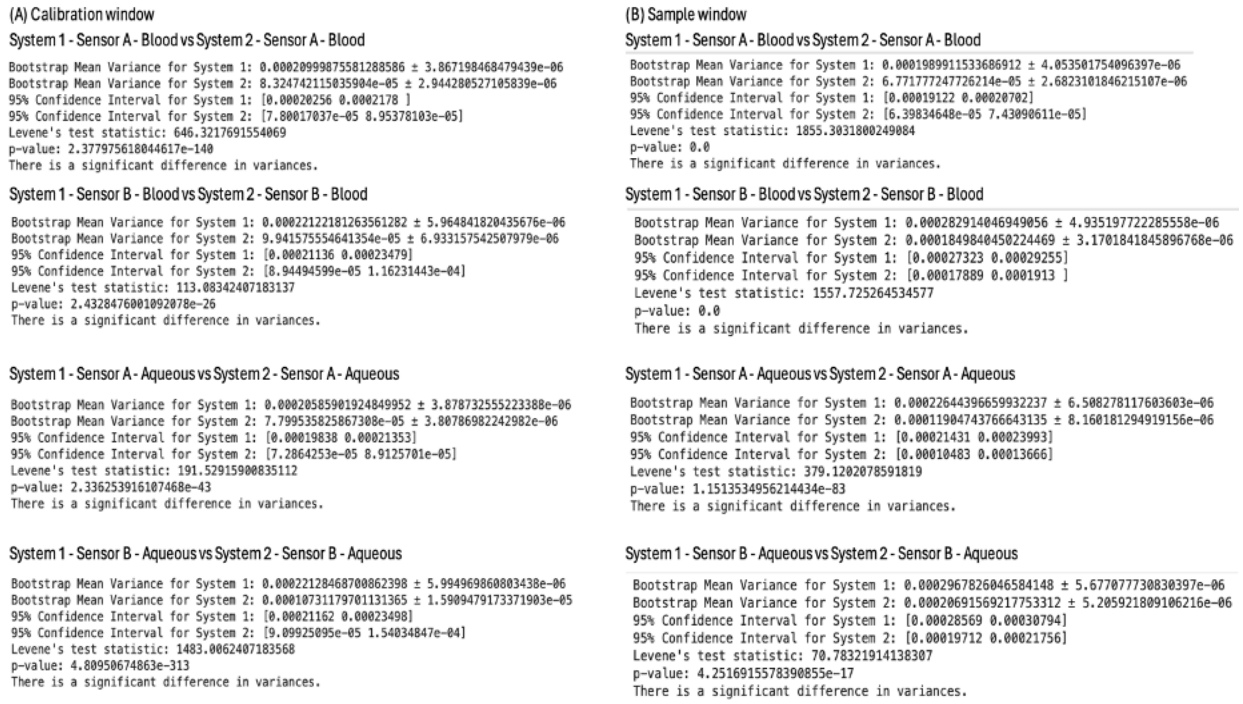


Figure 15. (A) Results of the comparison of variances from system 1 versus system 2 in

`

the calibration window for sensors A and B. (B) Results of the comparison of variances from system 1 versus system 2 in the calibration window for sensors A and B.

Pros

- Generate confidence intervals for the mean variance, giving us a sense of the variability and reliability of your estimates because they provide insights into how the variance might change with different samples, making the results more robust.
- Bootstrapping is a non-parametric method that does not rely on assumptions about the underlying distribution of the data or small sample sizes.

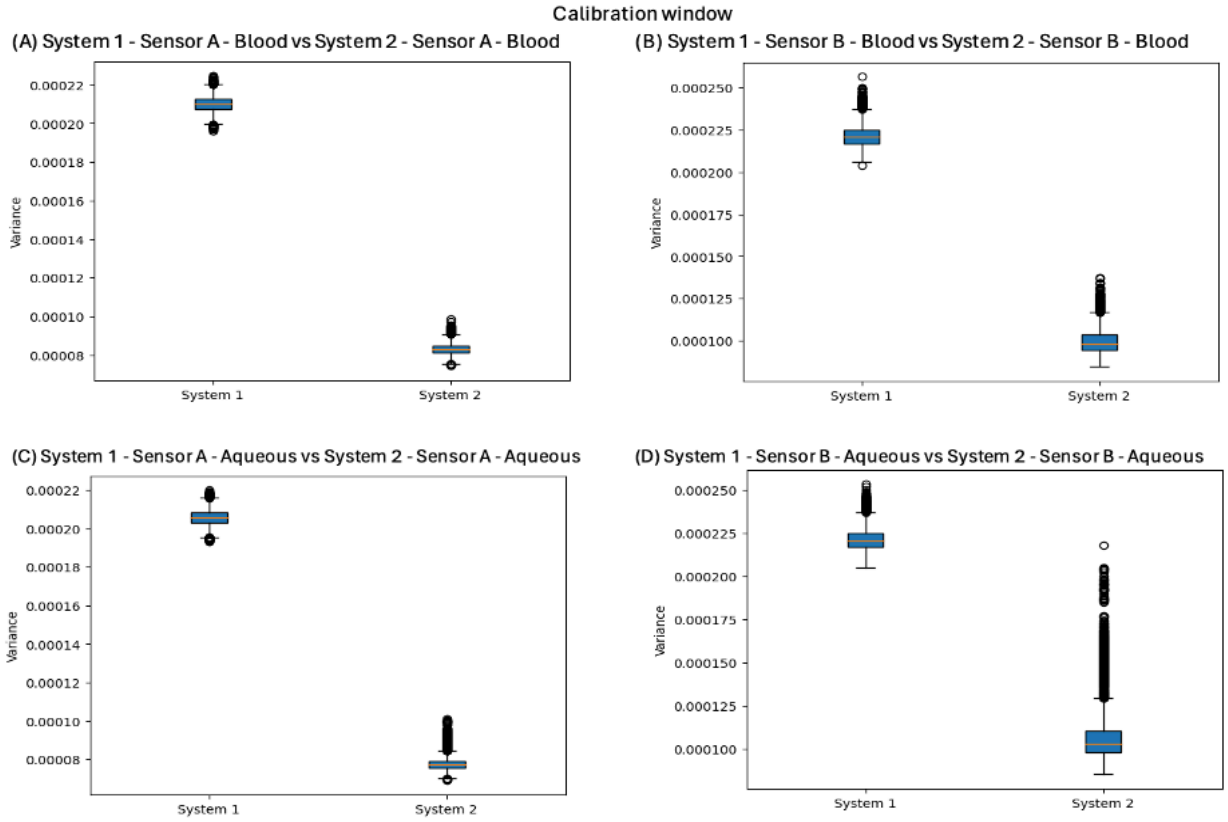


Figure 16. Variance distribution from the waveforms in the calibration window.

`

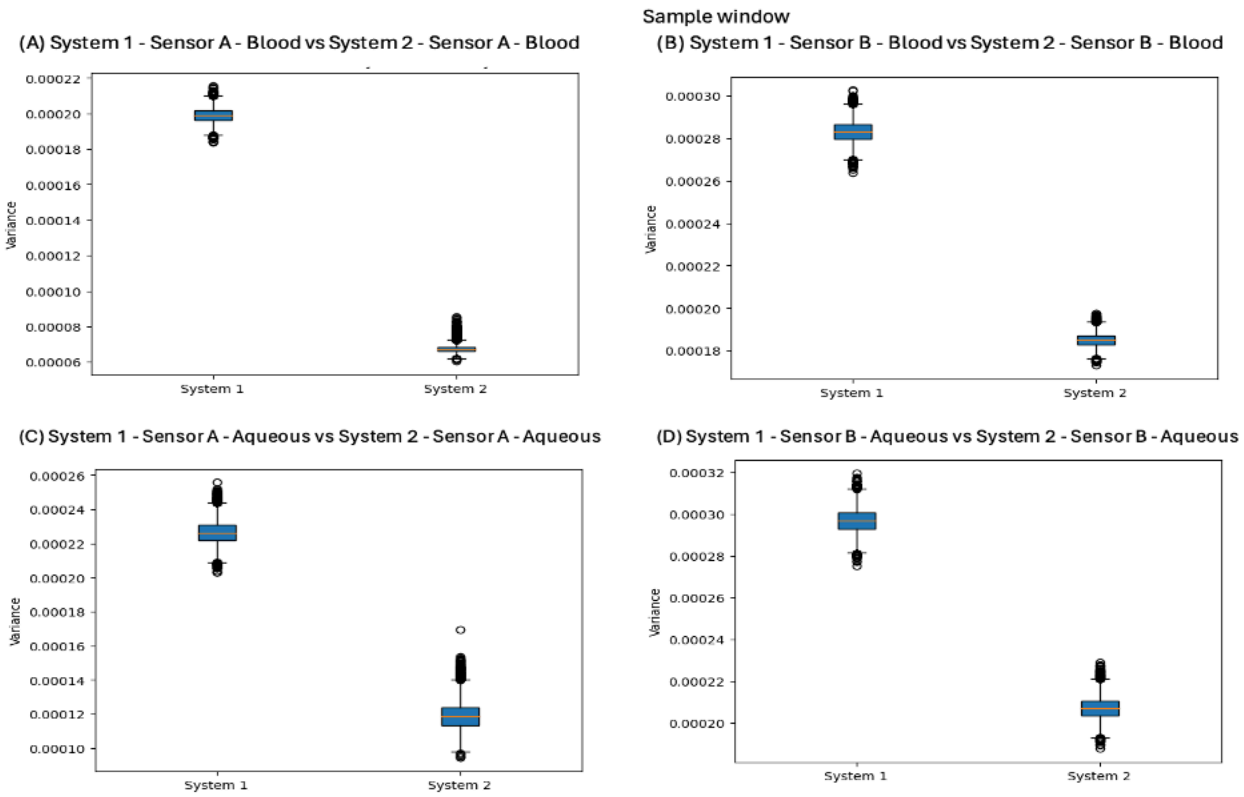


Figure 17. Variance distribution from the waveforms in the sample window.

Overall, the analysis of variance highlights differences among the two systems. However, there are many challenges in this type of analysis. First, after aligning to zero and differentiating the waveforms there are a large number of time series which are not stationary. Secondly, the assumption of independence within observations does not hold, because the data is autocorrelated. Finally, applying the comparison on large datasets of waveform can be computationally expensive. Therefore, we look for your feedback about other ways to quantify the degree of flatness of the waveforms from both systems to assess their accuracy, and ultimately decide if the window placement in system 2 is adequate or needs to be optimized.

5. Next Steps

Since we have abandoned the clustering methods for characterization as per client's suggestions, we will focus on the Functional Data Analysis for the waveforms characterizations and propose to define some quantification metrics for identifying the differences between two systems. Before we delve into the characterization differences quantification task, we will make some changes for the way we balance the data and center the data. We will pick the most effective and reasonable approach of preprocessing and apply the preprocessed data for the Functional PCA to characterize the waveform differences. So far, we've balanced the data only by the fluid types. However, clients are more interested in the analysis for the fluid temperature, therefore

`

we are planning to balance the data also by the fluid temperature by controlling the main category of fluid types: aqueous or blood. By doing so, we hope to provide more interesting findings for the impacts of fluid temperature on the Functional PCA and scores to satisfy client's expectation.

Once we decide the optimal approach to preprocess and balance the data. We will try to evaluate the flatness by obtaining the slopes of the FPCA components curve and the aggregated mean function of raw data for the system comparisons. Consistent results or findings from those may help us to summarize the differences between two systems. Besides, we will also perform bootstraps to obtain confidence intervals for the coefficients of the regressed FPCA components for both systems, and to see if there are any overlaps between them. If there is, which means system 1 and system 2 have some similarities in waveforms, which can provide the degree of the similarities in a more straightforward and accessible way.

Finally, we are looking for guidance to evaluate the analysis on the raw data such as slope calculation on the raw data in section 4 and variance analysis section 5 of this weekly report are feasible to quantify the differences among systems and the degree of flatness of the waveforms.

References

[1] J. G. Proakis and D. G. Manolakis, Digital Signal Processing: Principles, Algorithms, and Applications, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 1996.