

# **Siemens Biosensor Waveform Characterization and Window Optimization**

Siemens Healthineers

MDS Capstone Project: Final Report

Dylan Longert

Jinxin (Jessie) Wang

Nan Tang

Nayeli Montiel Rodríguez

Han (Eden) Chen

June 25th, 2024

## Table of Contents

<b>1. Executive Summary</b>	<b>1</b>
<b>2. Introduction</b>	<b>2</b>
<b>3. Background and Related Work</b>	<b>3</b>
<b>4. Data</b>	<b>4</b>
4.1. Exploratory Data Analysis	4
4.1.1. Time Series Data	4
4.1.2. Additional Features	5
4.1.2.1. Fluid Type	6
4.1.2.2. Card Age	7
4.1.2.3. Temperature	7
4.1.3. Current Window Limits	8
4.2. Preprocessing Steps	9
4.2.1. General Steps	9
4.2.2. Specific Steps	12
<b>5. Tools, Methodology, Techniques</b>	<b>13</b>
5.1. Tools	13
5.2. Methodology and Techniques	13
5.2.1. Window Flatness Evaluation	13
5.2.2. Functional Data Analysis	14
5.2.3. Window Optimization	15
<b>6. Analysis and Interpretation</b>	<b>17</b>
6.1. Main task: Waveform Characterization	17
6.1.1. Pipeline 1: Window Flatness Evaluation	17
6.1.2. Pipeline 2: Functional Data Analysis	19
6.2. Secondary task: Window Optimization	26
6.2.1. New Window Boundaries	26
6.2.2. Evaluation and Comparison on Pipeline 1	29
6.2.3. Evaluation and Comparison on Pipeline 2	31
<b>7. Conclusions</b>	<b>32</b>
<b>Appendices</b>	<b>34</b>
Appendix A. Methods and Techniques	34
Appendix B. Plots	37
<b>References</b>	<b>56</b>

## 1. Executive Summary

**Title:** Siemens Biosensor Waveform Characterization and Window Optimization

**Team Members:** Dylan Longert, Han (Eden) Chen, Nayeli Montiel Rodríguez, Nan Tang, and Jinxin (Jessie) Wang.

**Capstone Partner:** Siemens Healthineers, led by Raegan Allan, Research Engineer

**Background and Focus:** The epoc® Blood Analysis System, developed by Siemens Healthineers, is essential for accurate and rapid blood analysis in medical settings. This project compares two generations of this system: system 1 (current) and system 2 (under development). The primary goal is to analyze and optimize the biosensor waveform data from both systems, focusing on window flatness evaluation, functional data analysis (FDA), and window optimization.

### Summary of Methods:

#### 1. Waveform Characterization:

- **Window Flatness Evaluation:** Assessed the flatness of the calibration and sample windows by calculating slope coefficients for various attributes (fluid temperature, fluid type, and card age). Differences in slope between system 1 and system 2 were quantified using two-tailed t-tests.
- **Functional Data Analysis (FDA):** Used Functional Principal Component Analysis (FPCA) to compare the main modes of variation in waveform data produced from each system by using the first principal component (FPC1) that captures the majority of the data's variability and characterizing the waveforms using the principal component scores. Additionally, Functional Regression was used to assess the influence of external factors on the waveforms, such as the fluid temperature.

#### 2. Window Optimization:

- Optimized the calibration and sample windows for system 2 by adjusting delimit values to minimize slope differences between system 1 and system 2. This involved looping through possible delimit values and comparing the results to the previously defined delimit values.

### Summary of Results:

- **Waveform Characterization:** System 2 consistently showed higher slopes in sensor readings relative to system 1. FPCA revealed that the primary mode of variation captured over 99% of the data's variability. In most cases, system 1 is more stable than system 2.
- **Window Optimization:** The new optimized window limits for system 2 resulted in minimized slope differences, especially for fluid types Eurotrol L4, NB, and SB-3, and test cards aged 56 to 84 days.

**Conclusions:** We completed the two tasks required by Siemens: waveform characterization of sensor data in system 1 and system 2, and window optimization for system 2. The developed pipelines provide robust methodologies for characterizing sensor data and evaluating window flatness. The optimized windows for system 2 may contribute as a starting point for Siemens to compare other methods or support new findings in the future.

## 2. Introduction

Biosensor technologies play a pivotal role in the medical and pharmaceutical sectors, offering swift and precise measurements of specific biomarkers. Of particular interest are potentiometric biosensors [1]: they convert the biochemical interactions of a reference sample with a target analyte into an electrical signal. Notably, potentiometric biosensors excel at isolating analytes of interest from other substances during the biochemical process. The epoc® Blood Analysis System [2] is a portable blood analyzer comprising three components: Host, Reader, and Test Card with a biosensors module. The epoc Host, as a mobile computer equipped with a well-designed software application, communicates with the epoc Reader that reads and measures electrical signals from the Test Card with the calibration fluid and the sensor module (multiple sensors built-in) to receive potentiometric sensor data and processes this data to calculate analytical values and then displays the test result.



*Figure 1. The epoc® Blood Analysis System (system 1)*

The calibration fluid rehydrates the sensors, such that each sensor can detect a base value of electrical signal before every blood test, which is sealed in each test card for being released after insertion into the reader. Once the sensor module has been calibrated by detecting base values for the electrical voltages, the sample fluid or blood sample will be injected into the test card and flow through the sensor to provide a new sensor value. The difference in the sensor values between these two windows can be used to calculate the concentration for each sensor, which can be measured as biosensor waveforms via time in seconds.

System 1 is the device currently in use by Siemens and system 2 is the next generation with a new epoc Host, a different card reader, and a different Test Card. Tests being performed on these two systems mainly involve three sensor reading periods: calibration, bubble detection, and sample.

Since system 2 is under development and will be upgraded based on system 1, Siemens wants to investigate how the received biosensor data reacted differently between the two systems. Therefore, in this report, we aimed to answer two research questions for biosensor waveform analysis. The primary question is how the biosensor waveform characteristics differ between two types of systems of epoch® Blood Analysis built by Siemens Healthineers (explained as above). To answer the primary research question, we developed two different pipelines: Pipeline 1: Flatness Evaluation of Aggregated Mean Function, and Pipeline 2: Functional Data Analysis (FDA) [3]. The secondary question is how to optimize the window placement for the new generation to improve the precision and accuracy of blood analysis. For the secondary task, we optimized windows by minimizing the differences in the regressions slope of aggregated mean function between system 1 and system 2.

This report will first introduce the literature that informed the FDA techniques used to address the main research question. It will then describe the datasets provided by the Siemens research team through Exploratory Data Analysis. Next, the report will detail the preprocessing steps applied to the raw data. Following this, we will outline the methodology for each task, present the results, and offer an interpretation, which will be explained and visualized. Finally, we will conclude with the main findings based on the analysis, address the limitations, and discuss future directions for the study.

### **3. Background and Related Work**

Despite their widespread use in medical applications, there is limited research specifically addressing their use in blood analysis. This gap presents a challenge for our project, which aims to compare the electrical signals produced by two generations of the epoch® Blood Analysis System by Siemens Healthineers.

While specific studies on potentiometric biosensors for blood analysis are scarce, there is extensive literature on their application in other medical contexts. For instance, in the dairy industry, these biosensors are employed to measure urea concentration in cow urine. They also contribute to early detection of Alzheimer's disease by identifying amyloid  $\beta$ -42 [4]. The lack of focused research on blood analysis necessitates exploring methodologies from other domains, such as data collected from electrical signals of the nervous system.

Recent studies by Corson Areshenkov et al.[5] and Zhu and Wei [6] highlight advanced techniques in this field. Areshenkov et al. employed Functional Principal Component Analysis (FPCA) [7] to simplify complex sensorimotor network data into a concise format. Similarly, Zhu and Wei utilized Poisson Functional Principal Component Analysis to analyze neural data, providing intuitive visualizations and elucidating neural patterns. These studies provide the motivation to apply the Functional Data Analysis framework, based on Ramsay and Silverman, particularly for characterizing waveforms and comparing the principal modes of variation between the two generations of the epoch® Blood Analysis System.

Further work, such as Functional Regression, is included as supplementary analysis for identifying differences among the blood systems based on different feature effects. Functional Regression, as reviewed by Morris [8] and Wang et al. [9], plays a crucial role in extending functional data analysis beyond principal component analysis. It focuses on modeling relationships between functional predictors and responses, allowing for flexible modeling of complex data structures. Morris provides an extensive overview, discussing methodologies, applications, and computational aspects of functional regression. Wang et al. further elaborate on the theoretical foundations and practical applications, emphasizing its utility in various scientific domains. For a foundational understanding of Functional Data Analysis, including Functional Regression, additional insights can be gained here<sup>1</sup>.

In addressing the secondary objective of window optimization, our methodology is underpinned by the delineated constraints detailed in Section 4.1.3 and informed by collaborative insights provided by the Siemens research team.

## 4. Data

### 4.1. Exploratory Data Analysis

#### 4.1.1. Time Series Data

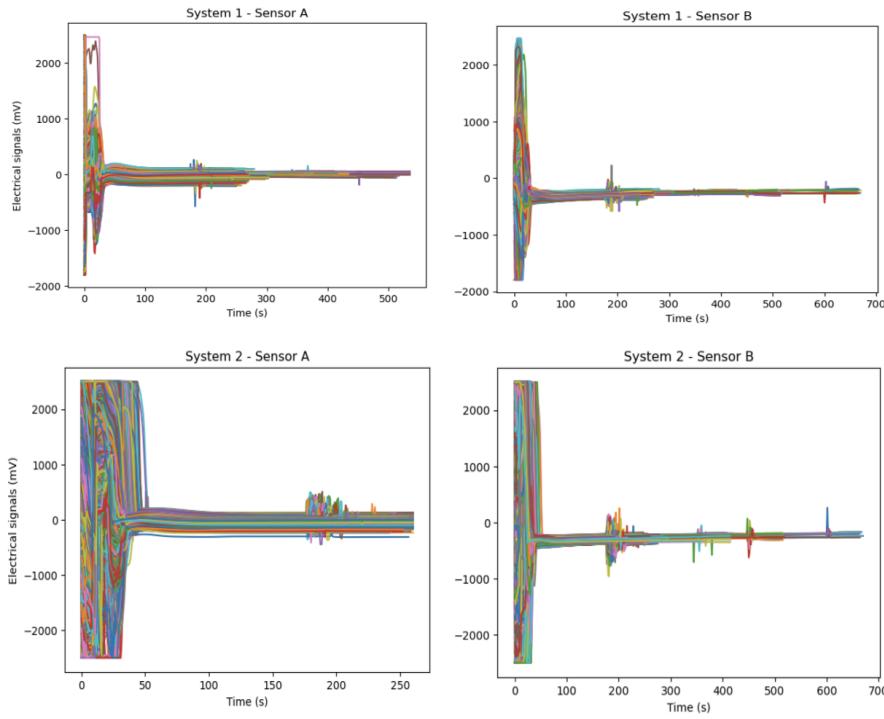
The time series datasets contain multiple time-series signals, referred to as waveforms or readings. There are four datasets total, consisting of system 1 sensor A, system 1 sensor B, system 2 sensor A, and system 2 sensor B. System 1 includes a total of 7,039 time series, with 3,524 for sensor A and 3,515 for sensor B. System 2 comprises 15,561 time series, with 7,781 for sensor A and 7,780 for sensor B. The raw time series data for both systems and sensors vary in length, as indicated by the shorter durations in seconds on the x-axis. Additionally, the electrical signals measured in millivolts (mV) in the y-axis exhibit both positive and negative values (Figure 2), as visualized in Table 1. The first column represents time in seconds, and each subsequent column contains the target variable - the electrical signal values produced by the potentiometric sensors in millivolts (mV).

**Table 1. Structure of time series data files**

Time (s)	TestID_1	TestID_2	TestID_3	TestID_4
0.2	-1797.2	1744.735	-1797.314	-231.9601
0.4	-1797.224	1749.621	-1797.341	-240.4039
0.6	-1797.221	1756.121	-1797.333	-248.0002
0.8	-1797.198	1762.368	-1797.349	-254.0748

---

<sup>1</sup> YouTube video "What is functional data analysis?" (available at  What is functional data analysis? )



**Figure 2. Time series of electrical signal values from the potentiometric sensors in millivolts (mV) for system 1 (current) and system 2 (new) for two different sensors (A and B).**

#### 4.1.2. Additional Features

The file with additional information on the time series contains 18 columns (Table 2), each providing supplementary information for each unique test run. These columns include one date type variable (DateTime), one string variable (Lot), eight categorical variables (System, Sensor, FluidType, and five return codes derived from the readings), and eight numeric variables (TestID, CardNumber, AgeofCardInDaysAtTimeOfTest, ReaderSerialNumber, SampleDetectTime, BubbleDetectTime, AmbientTemperature, and FluidTemperature). For the purposes of the analysis in this project, the following subsections focus on the description of fluid type, card age, and temperature (both fluid and ambient temperatures are considered together), which are our features of interest.

**Table 2. Additional Features Description**

Name of the variable	Description
TestID	Unique identifier of the test run created at the time of the test.
FluidType	Sample injected into the card.
DateTime	Date and time that the test was run.
Lot	Card lot. All cards built in one batch on one day.
CardNumber	Number that identifies the card that was used for running the test.
AgeOfCardInDaysAtTimeOfTest	Card age (in days) at the time of the test. Cards expire after 168 days.
ReaderSerialNumber	Unique identifier of the reader that ran the card.
SampleDetectTime	Time (in seconds) from the start of the test to when the sample fluid front is detected.
BubbleDetectTime	Time (in seconds) from the start of the test to when the bubble between calibration fluid and sample fluid is detected.
System	Identifies the system that runs the test.
Sensor	A or B.
AmbientTemperature	Temperature of the entire system (card, reader, host, and fluid) if Fluid Temperature is not available.
Fluid Temperature	Temperature of the sample fluid.
CardReturnCode, ReturnValue, Results Error Code, Device Return Code, Bge Test Return Code	Determine whether or not the test was successful.

#### 4.1.2.1. Fluid Type

The fluid samples injected onto the card are divided into two categories: aqueous and blood. The aqueous samples are further subdivided into Eurotrol L1, Eurotrol L3, Eurotrol L4, and Eurotrol L5, and blood samples are further subdivided into NB, AB, HNB, TB11, DB, SB, and SB-3. The number of tests run provided from each fluid type varies significantly between systems, being the minimum number of tests on SB-3 with 30 for system 1 and 36 for system 2 (Table 3).

**Table 3. Number of tests run by fluid type after merging the time series with their additional features by TestID**

Fluid Type	Fluid Name	System 1		System 2		Total
		Sensor A	Sensor B	Sensor A	Sensor B	
Aqueous	Eurotrol L1	617	617	1,495	1,494	4,223
	Eurotrol L3	602	602	1,629	1,629	4,462
	Eurotrol L4	317	312	525	525	1,679
	Eurotrol L5	423	423	1,263	1,263	3,372
Blood	NB	594	593	1,008	1,008	3,203
	AB	232	232	728	728	1,920
	HNB	189	189	402	402	1,182
	TB11	175	174	225	225	799
	DB	117	117	210	210	654
	SB	111	109	260	260	740
	SB-3	30	30	36	36	132
Total		3,407	3,398	7,781	7,780	22,366

#### 4.1.2.2. Card Age

Card age at the time of the test in system 1 and system 2, which employed sensor A and sensor B, was tested on cards ranging from 5 to 240 days old. Additionally, The cards for system 1 had a median age of 92 days, while the cards for system 2 had a median card age of 68 days.

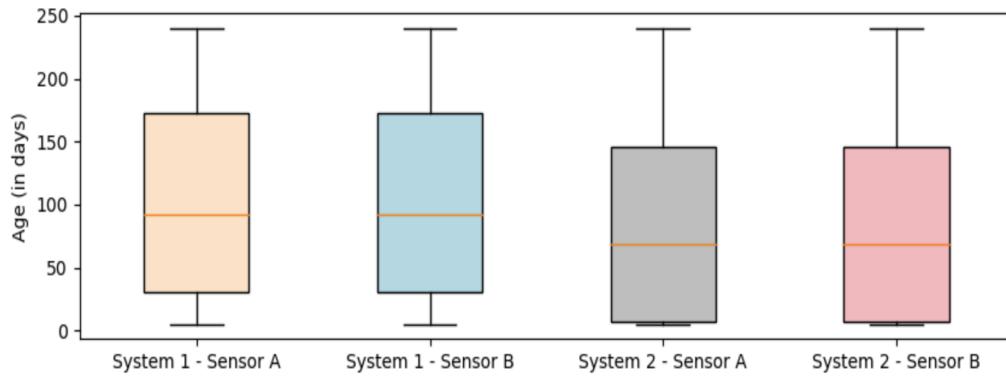
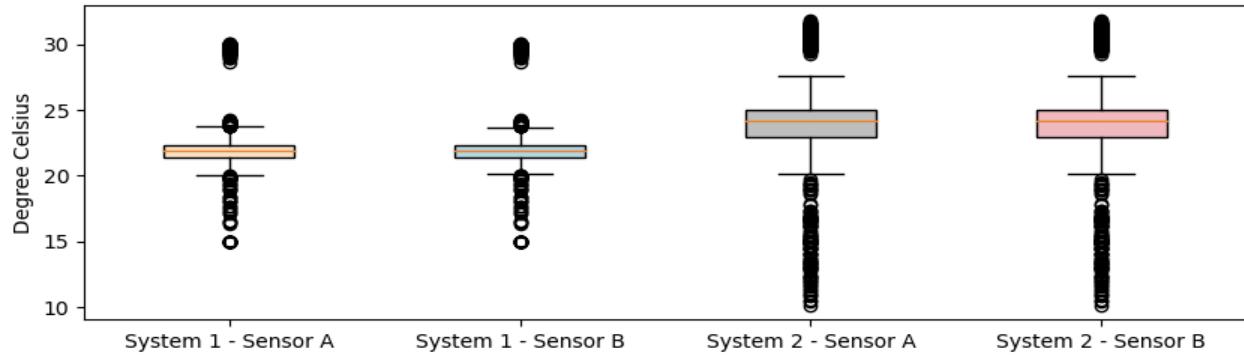


Figure 3. Distribution of card age in days for system 1 (current) and system 2 (new) for two different sensors (A and B).

#### 4.1.2.3. Temperature

Ambient temperature was provided for all systems and sensors. However, fluid temperature was recorded for only 119 waveforms in system 1 and 194 waveforms in system 2. Therefore, following the Siemens research team's recommendations, ambient temperature was used to represent the entire system temperature when fluid temperature data was unavailable. After filling the missing fluid temperature values with ambient temperature, the median fluid temperature was approximately 21.90°C for system 1 and 24.15°C for system 2 for Sensors A and B (Figure 4).



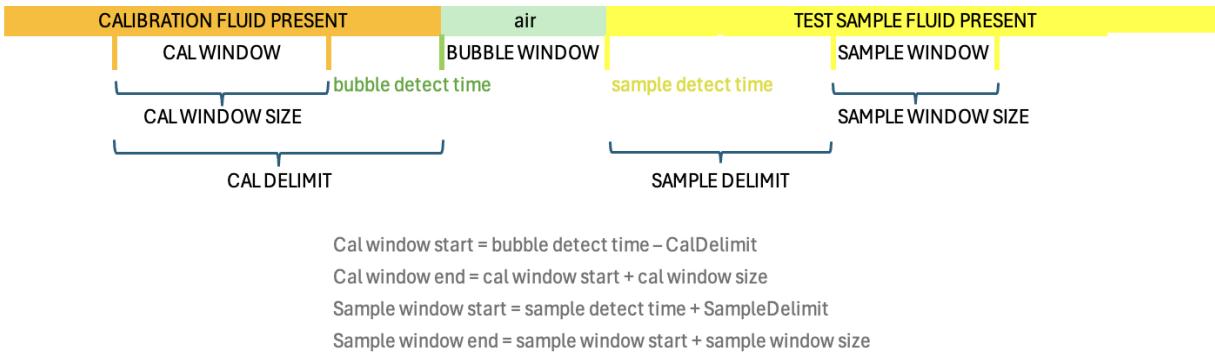
**Figure 4. Distribution of fluid temperature in degree Celsius after filling NA values with ambient temperature, for system 1 (current) and system 2 (new) for two different sensors (A and B).**

#### 4.1.3. Current Window Limits

The fixed values for the boundaries currently used by Siemens Healthineers were provided to determine the *calibration* window and *sample* window. These windows are considered representatives of the Calibration Fluid Present and Test Sample Fluid Present, respectively. Specifically, different values are applied to readings from sensor A, sensor B with blood fluids, and sensor B with aqueous fluids. The details of the window definitions from each sensor time series are as follows:

- **SampleDetectTime:** The time length from the start of the test to when the sample fluid front is detected (in seconds).
- **BubbleDetectTime:** The time length from the start of the test to when the bubble between calibration fluid and sample fluid is detected (in seconds).
- **CalDelimit:** The time length from the Calibration Start timestamp to Bubble Detect timestamp (in seconds).
- **SampleDelimit:** The time length from the Sample Detect timestamp to Sample Start timestamp (in seconds).
- **CalWindowSize:** The time length of calibration window (in seconds).
- **SampleWindowSize:** The time length of the sample window (in seconds).

We focused on the characteristics of the response variable, specifically within the calibration window and test sample window. The differences and the degree of flatness of the curves within these windows are expected to provide stable sensor values for the Siemens team to calculate their analyte of interest.



**Figure 5. Current window partitioning.** “**Current Fluid Present**” corresponds to the period when the calibration fluid flows through the sensor, “**Air**” corresponds to the period when the bubble is being detected after sample injection, and “**Test Sample Fluid Present**” corresponds to the period when the sample is being detected.

## 4.2. Preprocessing Steps

### 4.2.1. General Steps

Data preprocessing was performed separately on the time series data and the additional features data. For the time series data, the waveforms were transposed so that each one was identified by TestID in one row and the timestamps were in the columns. For the additional features data, we added a new column called ‘Fluid\_Temperature\_Filled’. This column preserved the fluid temperature values and was filled with the ambient temperature if the fluid temperature was not available.

In this dataset, we relabeled the values ‘system 2A’ and ‘system 2B’ to ‘system 2’ in the ‘System’ column, which identifies the waveforms from the new system. We also categorized key attributes of interest as follows: fluid type (Blood and Aqueous fluids), card age (0-9, 9-28, 28-56, 56-84, 84-112, 112-140, 140-168, 168-196, 196-224, 224-256 days), and fluid temperature (below 20°C, between 20°C and 25°C, and over 25°C). These categorized columns (‘FluidTypeBin’, ‘CardAgeBin’, ‘FluidTempBin’) were added to the additional features data. Finally, we filtered the column ‘Return code’ to get only the waveforms with values ‘Success’ or ‘UnderReportableRange’, both of which correspond to successful tests.

The next step involved merging the time series and additional feature datasets for each System (1 and 2) and Sensor (A and B). The merged data was cleaned by removing 126 waveforms that corresponded to unmatched tests run.

From the merged datasets, we computed the start and end values for both the calibration and sample windows. These calculations were based on the provided starting points and delimiters for both windows (Figure 5). As a result, four new columns (cal window start, cal window end, sample window start, and sample window end) were added to the merged datasets (Table 4).

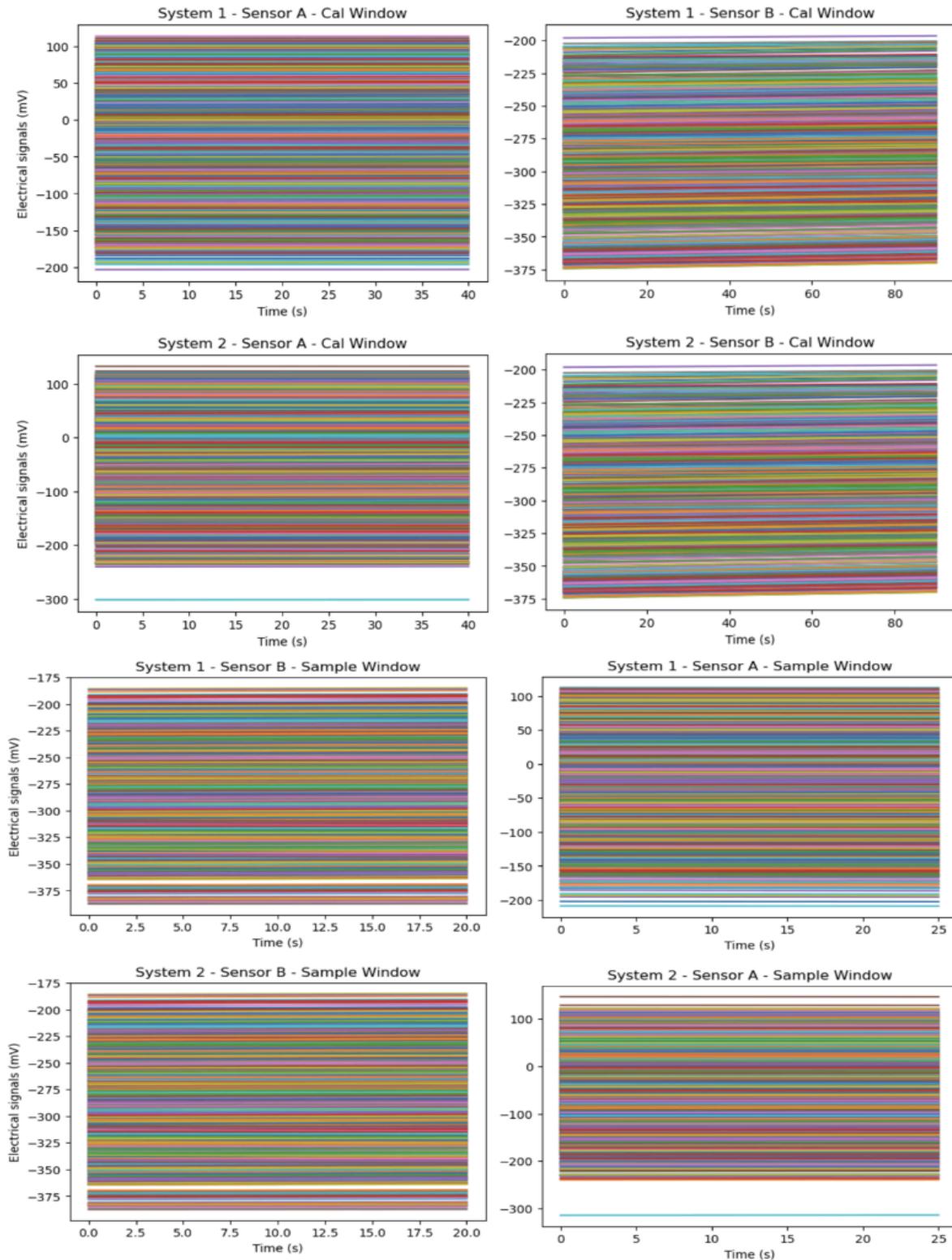
**Table 4. Subset of the merged data showing the calibration and sample window start and end points.**

TestID	Sample Detect Time	Bubble Detect Time	Calibration Window Start	Calibration Window End	Sample Window Start	Sample Window End
TestID_1	185.4	184.4	173.4	181.4	200.4	205.4
TestID_2	185.8	185.0	174.0	182.0	200.8	205.8
TestID_3	186.8	186.8	175.0	183.0	201.8	206.8
TestID_4	185.8	185.0	176.6	184.6	203.4	208.4

Next, we extracted the calibration and sample window data for each time series identified by TestID. This extraction was done by extracting timestamps greater than or equal to the window start and less than or equal to the window end for both the calibration and sample windows. Figure 6 illustrates the shape of the waveforms after extraction, while Table 5 details the number of waveforms (rows) and the number of timestamps (columns) within each dataset following the window extraction.

**Table 5. Dimensionality of the data after window extraction.**

Sensor	System	Window	Rows	Columns
A	1	Calibration	3368	41
A	2	Calibration	7743	41
B	1	Calibration	3373	91
B	2	Calibration	7743	91
A	1	Sample	3368	26
A	2	Sample	7743	26
B	1	Sample	3373	21
B	2	Sample	7741	21

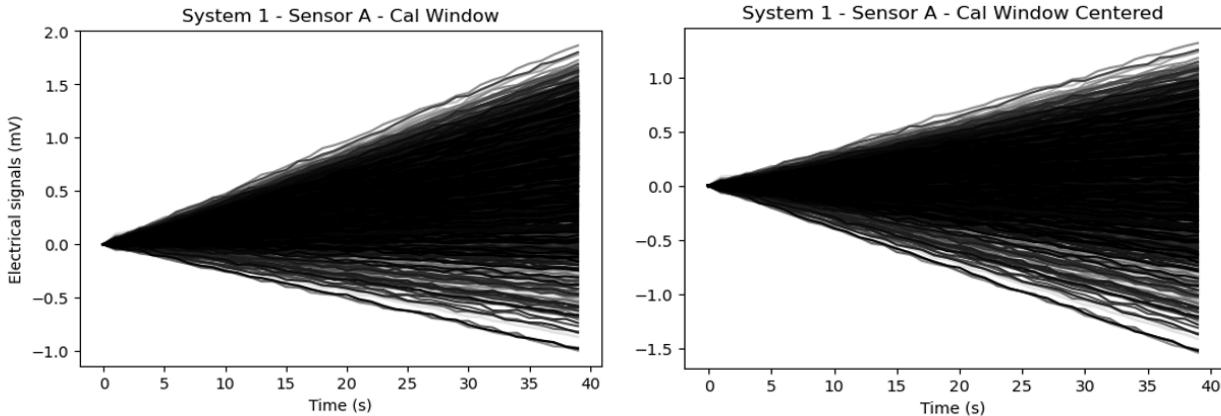


**Figure 6. Waveforms from the calibration and sample windows after the extraction for system 1 (current) and system 2 (new) for two different sensors (A and B).**

#### 4.2.2. Specific Steps

For the Functional Data Analysis methods, specifically under the Functional Principal Components, we considered the data after window extraction starting at zero and then centered by subtracting the column-wise means. Additionally, we balanced the data by attribute of interest and bin.

- **Zero-alignment:** The windowed data was aligned to zero by subtracting the values in the first column (timestamp 0) from all other columns. This adjustment standardizes the starting point of all waveforms, preserving their overall shape. As a result, we can effectively visualize the waveforms (Figure 7) as a trending time series, with some displaying upward trends and others showing downward trends.
- **Centering by column:** The FDA object's centering parameter subtracts the column mean from each data point within that column [10]. This process adjusts the waveforms so that the mean (known as the mean function in FDA) of each column becomes zero. This normalization aids in the comparison and analysis of trends or anomalies over time, enhancing clarity in identifying relative changes or patterns. Figure 7 illustrates the windowed data starting at zero (left) and centered (right), shifting from -1 to 2 in the aligned data to -1.5 to 1 in the centered data.



**Figure 7. Windowed data after zero-alignment (left) and subsequent centering by column (right) for system 1 (current) sensor A in the calibration window.**

- **Balancing:** The window data for system 2 has roughly twice the number of time series compared to system 1 (Table 5). Additionally, when examining the number of time series per attribute binned ('FluidTypeBin', 'CardAgeBin', 'FluidTempBin') the class imbalances became more apparent. To address this, TestIDs were randomly sampled without replacement, with the sample size for each class set to the minimum count of samples from the bins. Resulting in 30, 142, and 72 time series values per bin for fluid type, card age, and fluid temperature respectively.

## 5. Tools, Methodology, Techniques

### 5.1. Tools

All data wrangling, preprocessing, and analyses were completed in Python v3.11.0. Furthermore, the *scikit-fda* library [11] was used to run Functional Data Analysis. We implemented Functional Principal Component Analysis on a grid representation of the data, where the grid consists of the number of points compressed by the time series extracted from the original data. Additionally, we applied Functional Regression on a basis to model the relationships between a grid representation of the data as functional responses and scalar predictors. For window optimization, we employed the module *linear\_model.LinearRegression* from the Python *scikit-learn* library [12] to conduct the simple linear regression on sensor window data, where the regression coefficient can be regarded as the slope of window data.

### 5.2. Methodology and Techniques

We evaluated the slope of different bins using 2-tailed t-tests to determine if the observed differences in the slopes of system 1 and system 2 were statistically significant. This provided valuable insights into the performance variations between the systems.

Next, we applied Functional Principal Component Analysis (FPCA), to reduce dimensionality and characterize the waveforms by their PCA scores. This method captured key variations and represented the complex waveform data in a clear manner, which was especially valuable to the Siemens research team.

Both above techniques were applied to answer our primary research question, which is our main task. Additionally, to optimize window placement, we calculated the range of window locations and visualized the differences in slopes to find the minimum difference between system 1 and system 2. We then used the optimized window limits in the pipelines from the main task and compared them against the results from the original window limits.

#### 5.2.1. Window Flatness Evaluation

Our goal was to compare the slopes of the waveforms of system 1 and system 2 with the effect of three attributes: card age, fluid temperature, and fluid type. First, we grouped each of the attributes into smaller bins: 10 for card age, 3 for fluid temperature, and 11 for fluid type. Following this, we aggregated the timestamps within each bin by calculating the mean across all wavelength samples. Finally, using the aggregated data, we calculated slope coefficients for each of the attributes. The difference in the mean of system 1 and system 2 was quantified and assessed with a 2-tailed t-test. The overall pipeline is illustrated in Figure 8.

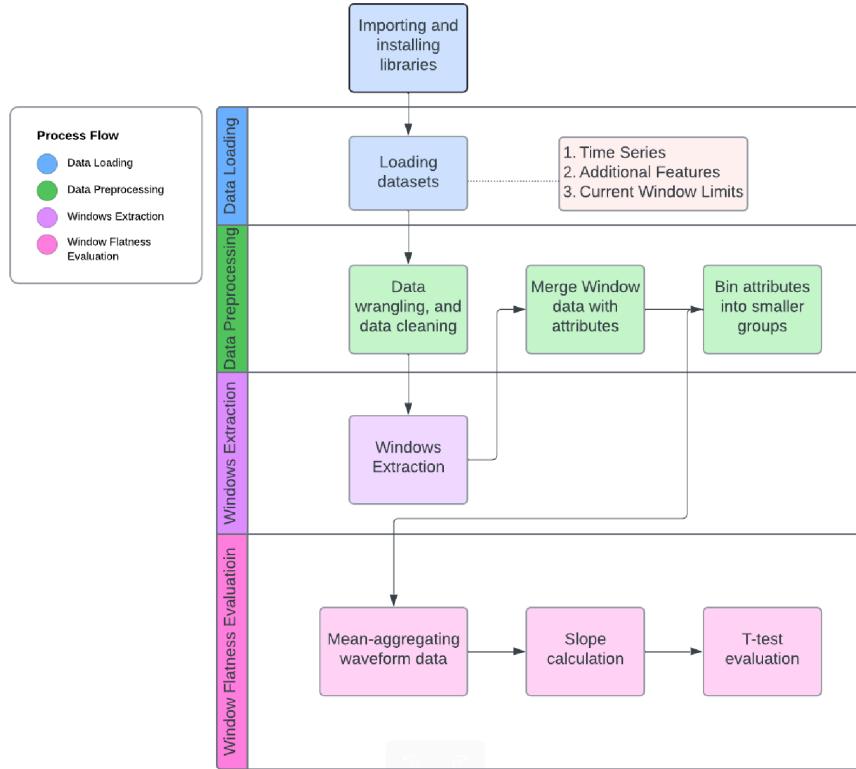


Figure 8. **Pipeline 1 - Window Flatness Evaluation**

### 5.2.2. Functional Data Analysis

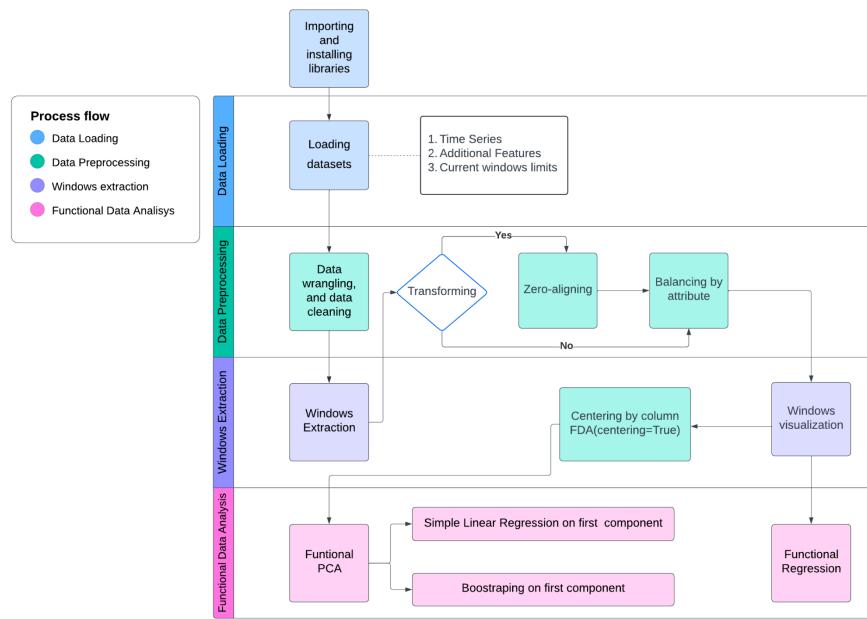
Based on the Functional Data Analysis framework, we employed two methods. Firstly, we used Functional Principal Component Analysis (FPCA) to extract key features from the thousands of time-series data to create a low-dimensional representation of the patterns in multiple time series. FPCA identifies eigenfunctions (analogous to principal components) that capture the dominant patterns of variability in the functional data. For more details, refer to Appendix A. Secondly, we employed Functional Regression to investigate the influence of the attributes fluid temperature, card age, and fluid type. These steps were executed on the data grouped by attribute binned.

First, we visualized the first two components for both system 1 and system 2. This visualization helped identify the dominant patterns of variability in each system, providing a clear comparison of how the two systems behave.

Next, we narrowed our focus to the first component, which captured the majority of the variation. We used simple linear regression to calculate the slopes of the first component for both systems. These slopes provided a quantitative measure of the rate of change in the primary mode of variation, offering further insight into the dynamics of each system. To assess the stability of the first component, we reran FPCA

while bootstrapping the window-binned data. From this, we calculated a 95% confidence level for the mean of the first functional principal component.

In addition to FPCA, we applied Functional Regression to the individual predictors, fluid temperature, and card age with the entire balanced time series in different systems and sensors as response variables. The coefficient plots derived from these regression analyses were compared to one another. The overall pipeline is visualized in Figure 9.



**Figure 9. Pipeline 2 - Functional Data Analysis**

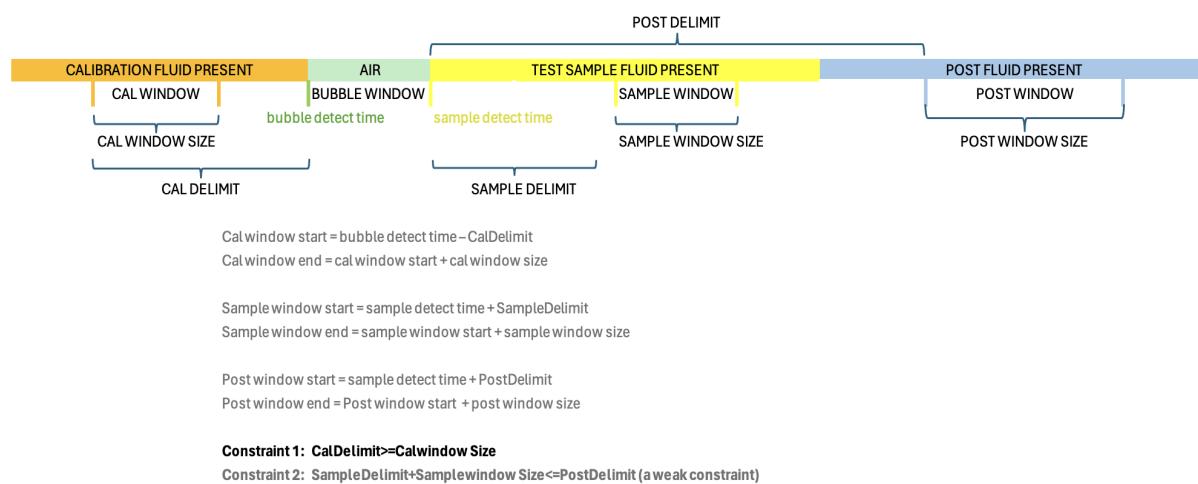
### 5.2.3. Window Optimization

We created a customized function that minimizes the regression slopes between system 1 and system 2 while keeping moving the current windows within certain limits. We adopted this approach because the Siemens research team expected us to optimize system 2 windows by adjusting the flatness level to be closely similar to the current windows for system 1.

Since the window limits like detect time and window size are constants, only the delimit variable can be adjusted (Figure 10). By inputting different values for the delimit variable when calling our custom function, the specific window can be moved to different positions within certain limits. We were able to find the optimal delimit value that minimized the slope differences, and the optimal delimit value can be used to calculate the optimal window position based on the calculation equations.

First, we calculated the column-averaged sensor values of the calibration and sample window for sensor A and sensor B in system 1. Next, we calculated the slope of sensor calibration/sample window data for sensor A and sensor B in system 2 using simple linear regression, iterating over each possible CalDelimit and SampleDelimit value within the available range. Then we computed the difference between these slopes for each CalDelimit and SampleDelimit value and plotted the difference values against these values. Finally, we identified the optimal solution, which is the minimum absolute value of the difference between system 1 and system 2, and marked it with a different color in the line plot. Additionally, for better comparison, we also marked the point corresponding to the current CalDelimit and SampleDelimit values of system 1.

We initially looped through both CalDelimit and SampleDelimit from 0 to 100 seconds, and we found that both minimums were found further behind the current windows not before them, which was also consistent with the client's expectation. Therefore, we tried moving the calibration window further in time (closer to the bubble detect time, i.e. a smaller value for CalDelimit) and the sample window further in time (further from the sample detect time, i.e. a larger value for SampleDelimit) by looping through delimit values within certain ranges to avoid overlapping windows. While the ranges are continuous with 0.2-second intervals, the current sample window slightly overlaps the post window for the sensor B blood fluid attribute. Hence, we only focused on the post window end constraint.



	System 1			System 2		Window Size		
	CalDelimit	SampleDelimit	PostDelimit	CalDelimit	SampleDelimit	Calwindow	SampleWindow	PostWindow
<b>Sensor A</b>	11	15	25	8~13	13~20	8	5	5
<b>Sensor B-Blood</b>		24	27.5		24~34			
<b>Sensor B-Aqueous</b>	20	30	37.5	18~22	28~34	18	4	5

**Figure 10. Window Limits Summary for Window Optimization**

## 6. Analysis and Interpretation

### 6.1. Main task: Waveform Characterization

Each approach described above was summarized into 2 pipelines, which can be used for future analyses by the Siemens research team.

#### 6.1.1. Pipeline 1: Window Flatness Evaluation

Since the results for all Sensors and attributes are identical, we will only be presenting sensor A and B data with the Fluid Temperature attribute. See Appendix A for all other attributes

As can be seen in Figure 11 and Table 6, while the slope pattern within sensor A's calibration window is difficult to discern, the slope of system 2 is consistently higher than system 1 within the sample window across all bins. Within sensor B, the slope of system 2 is consistently greater than system 1 in both the calibration and sample window (Figure 12; Table 7), thereby indicating that the slopes for each system are indeed different across all bins. For both sensors A and B, the t-test is significant across all bins ( $p < 0.001$ ), indicating that the slopes for systems 1 and 2 are significantly different from each other. This thereby accomplishes our primary goal of describing differences between system 1 and system 2.

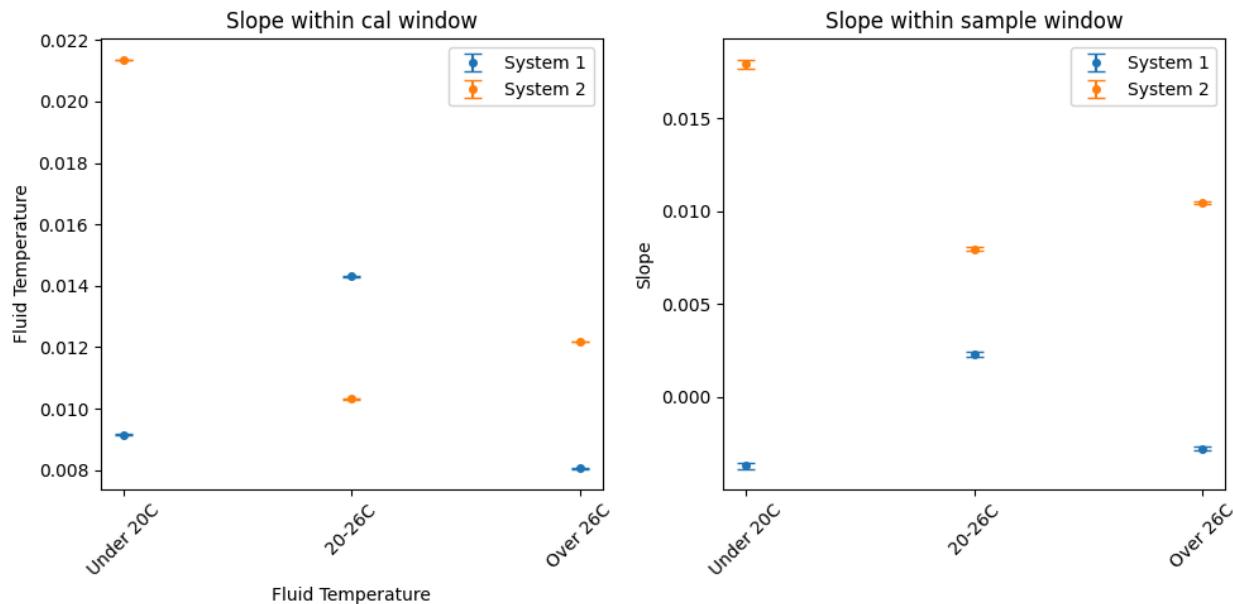
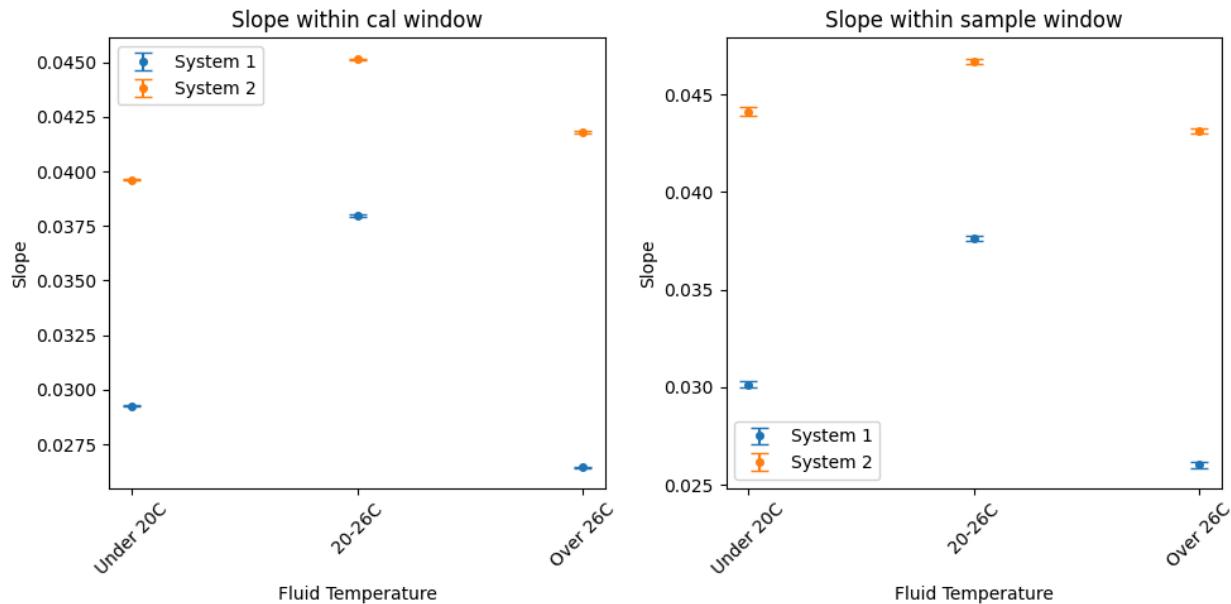


Figure 11. Slope of the waveforms for system 1 calibration window (left) and system 2 sample window (right) sensor A against fluid temperature. X-axis labels represent fluid temperature bins.

**Table 6. T-statistic and p-value comparing the Slope of system 1 and system 2 sensor A, binned by fluid temperature. The first two columns represent the calibration window and the latter two columns represent the sample window.**

	t-statistic-cal	p-value-cal	t-statistic-sample	p-value-sample
Under 20C	-534.027736	0.0	-69.820302	0.0
20-26C	131.610849	0.0	-30.199707	0.0
Over 26C	-256.310852	0.0	-115.363388	0.0



**Figure 12. Slope of the waveforms for system 1 calibration window (left) and system 2 sample window (right) sensor B against fluid temperature. X-axis labels represent fluid temperature bins.**

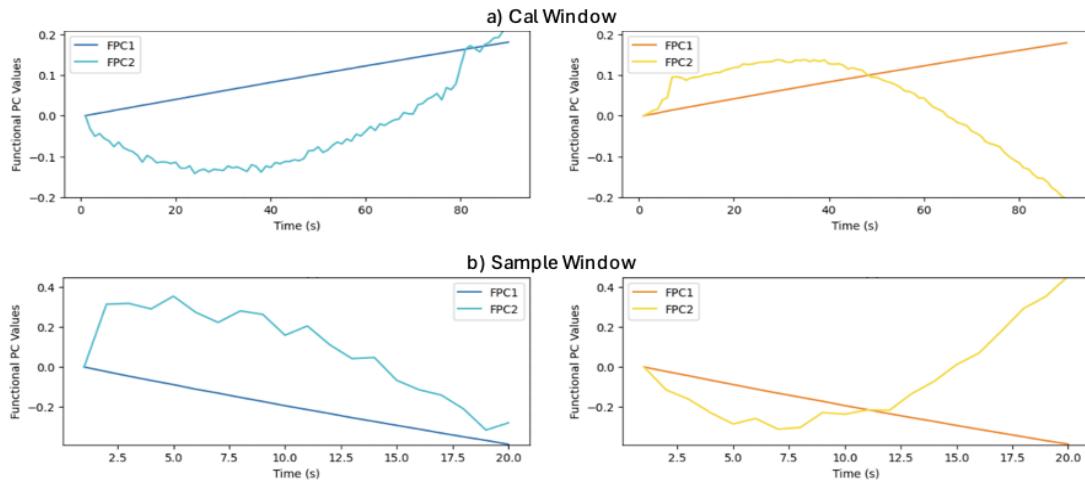
**Table 7. T-statistic and p-value comparing the Slope of system 1 and system 2 sensor B, binned by fluid temperature. The first two columns represent the calibration window and the latter two columns represent the sample window.**

	t-statistic-cal	p-value-cal	t-statistic-sample	p-value-sample
Under 20C	-231.683017	0.0	-52.154815	0.0
20-26C	-113.494405	0.0	-48.087302	0.0
Over 26C	-317.134318	0.0	-89.934663	0.0

### 6.1.2. Pipeline 2: Functional Data Analysis

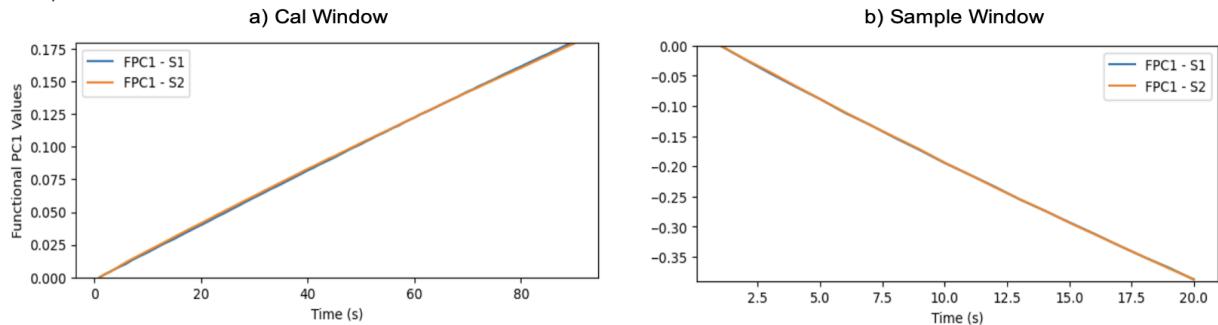
Below we present a summary of the results of the Functional Data Analysis approach on one random sample taken from system 1 and system 2, sensor A and sensor B performed separately on the three attributes of interest categorized per bin. For more details, refer to Appendix A.

The first component captures over 99% of the variability of the window data, while the remaining components capture around 1%. We start by visualizing the first two components to understand the behavior of the main modes of variation. We presented the results for system 2 sensor B (Figure 13). See Appendix B for the remaining systems and sensors. Our analysis indicates that the first component (FPC1) represents variations around the mean function, while the remaining components appear to represent minor variations or noise. Consequently, we selected the first component and retained it for analyzing and comparing the differences between the two systems.



**Figure 13. First Component (FPC1) and Second Component (FPC2) for system 1 (left) and system 2 (right) using sensor B, with window data balanced by fluid temperature: (a) Calibration window and (b) Sample window.**

Figure 14 shows a line chart used to visually assess the trend of component 1. Additionally, Table 8 summarizes our findings on both sensors and systems for the three attributes. Interestingly, when balanced by fluid temperature the first component has the same trend for all except sensor A calibration window. The same opposite pattern occurs when comparing the trend of component 1 under the windows data balanced by fluid type. In contrast, card age preserves the same trend in the main mode of variation across all cases.



**Figure 14. Comparison of the First Component (FPC1) for two systems (system 1: S1 and system 2: S2) using sensor B, with windows data balanced by fluid temperature: (a) Calibration window and (b) Sample window.**

**Table 8. First Component trend comparison between system 1 and system 2.**

Sensor	Window	Fluid Temperature	Fluid Type	Card Age
A	Calibration	Opposite trend	Negative trend	Negative trend
A	Sample	Positive trend	Opposite trend	Negative trend
B	Calibration	Positive trend	Opposite trend	Positive trend
B	Sample	Negative trend	Negative trend	Negative trend

### Comparison by Simple Linear Regression

While we can compare patterns of the components in different systems by visual inspection, we still need to find a statistically robust method to quantify the differences, especially for the slope of the first component, before applying a hypothesis test to determine statistical significance.

The absolute value of the first component is almost identical not only in the two systems but also across different features within the same window and sensor (Table 9).

Additionally, as indicated by the underlined cells in Table 9, there are three groups of data that exhibit opposite slopes of the same magnitude: (1) the fluid temperature attribute of sensor A's calibration window, and the fluid type attribute of (2) sensor A's sample window and (3) sensor B's calibration window. While some remaining groups have similar slope values, they are still significantly different ( $p<0.1$ ).

**Table 9: Comparison of the slope of the first component in system 1/ system 2.** (A. The underlined content indicates different patterns in systems. B. \*\*\* Significant at the  $p<0.01$  level; \*\* Significant at the  $p<0.05$  level; \* Significant at the  $p<0.1$ )

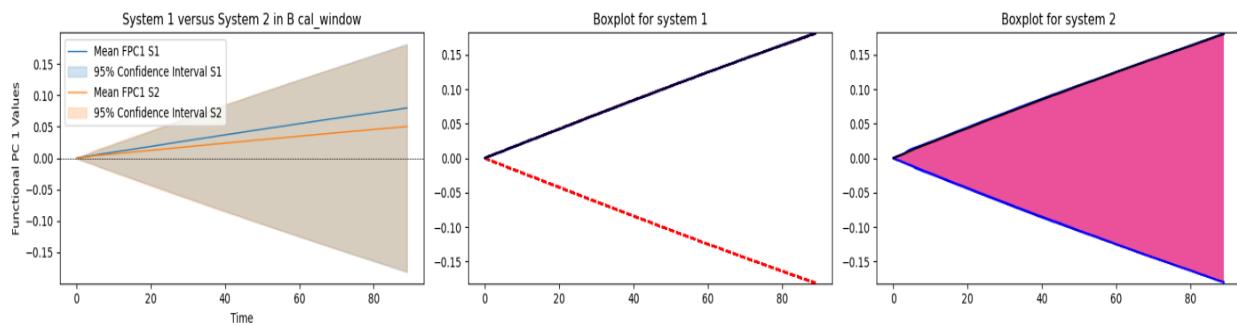
Sensor	Window	Fluid	Fluid Type	Card Age
Temperature				
A	Calibration	<u>0.007072</u> / <u>-0.007055</u> ***	-0.007058 / -0.007065 ***	-0.007061 / -0.007069
A	Sample	0.014846 / 0.014447 ***	<u>-0.014793</u> / <u>0.014669</u> ***	-0.014793 / -0.014652 ***
B	Calibration	0.002036 / 0.002007 ***	<u>0.002047</u> / <u>-0.002023</u> ***	0.00203 / 0.002015 **
B	Sample	-0.020308 / -0.020375	-0.020301 / -0.020309	-0.020283 / -0.020312

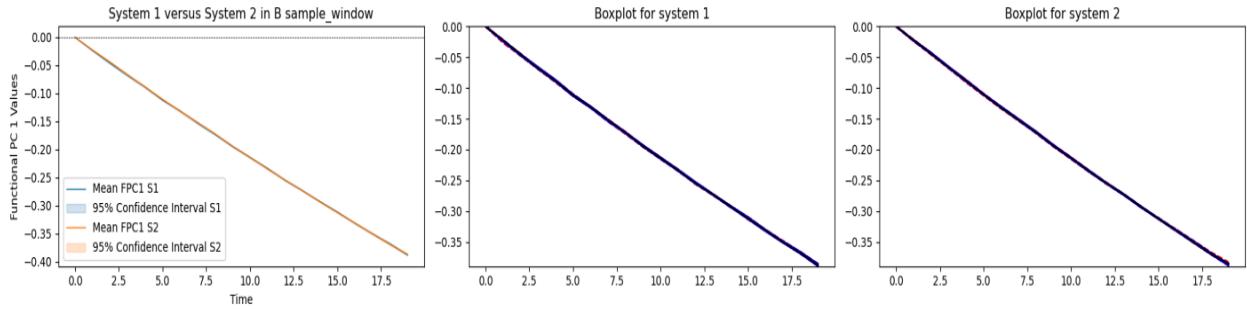
### Bootstrap and boxplots

Bootstrapping was utilized to validate the robustness of the FPCA results. Here, we present the results for sensor B focusing on Fluid Temperature. For further details, refer to the results displayed in Appendix B.

Within the Calibration Window, both systems exhibit similar behaviors in the first component, showing predominantly positive trends indicated by the slope of the Mean of FPC1 (Figure 15). However, the distribution of FPC1 appears more consistent in system 1 compared to system 2 (Figure 15). Fewer instances exhibit downward trends, and negative slopes are considered outliers in system 1 (marked by red dashed lines). Conversely, approximately 25% of the first components display negative patterns in system 2 (as indicated by the blue fence and pink center envelope).

We applied the same methodology to the sample window, and both systems exhibited a consistent downward trend without any fluctuations, conforming to our client's expectation that system 2 shows no difference from system 1 (Figure 15).





**Figure 15. Confidence interval and boxplots for the first component in the sensor B (top) Calibration Window (bottom) Sample window, focusing on Fluid Temperature [13].** The left panel figures represent the 95% confidence interval: the black line represents  $y=0$ , and the blue and orange lines represent the mean of all the first components of the resampled data in system 1 and system 2, respectively. The central and right panels visualize the box plots for the first component: the black line represents the median, the pink band represents the central envelope, the blue lines represent the maximum non-outlying envelope (quartiles) and the red dashed lines represent outliers.

Systems 1 and 2 exhibit similar patterns specifically for Card Age, after balancing this feature, underscoring its importance (Table 10). However, despite the overall pattern similarity between the two systems, there are still some fluctuations, indicating potential for window optimization (see section 6.2).

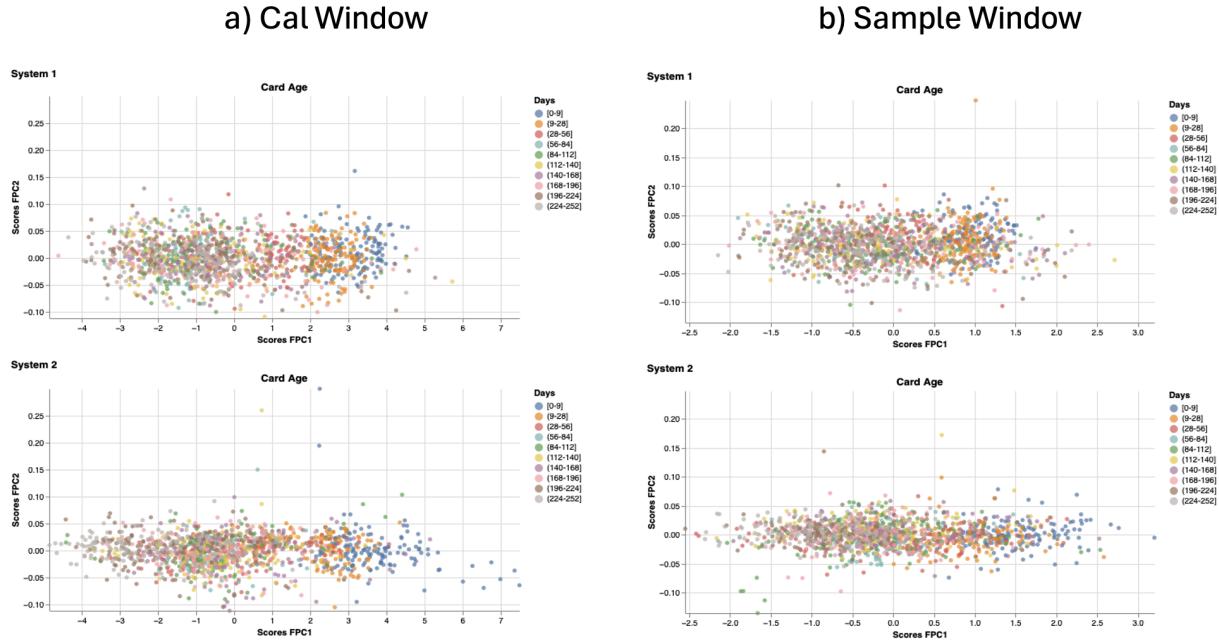
After applying the bootstrap to sensor B's sample window, the two systems are visually identical, regardless of the features. Conversely, in sensor A's sample window for Fluid Type, the two systems consistently exhibit opposite trends. This suggests that the location of the sample window for sensor B is accurate, but it still requires some adjustments for sensor A.

**Table 10: Comparison of the trend and stability of the first component in two systems.** (A. The slopes of the two systems are the same and without any fluctuation. B. The **Blood content** indicates a strong conclusion. C. The underlined content indicates different patterns in systems)

Sensor	Window	Fluid Temperature	Fluid Type	Card Age
A	Calibration	<u>Opposite trend with outliers</u>	Same trend with outliers	Same trend with outliers
A	Sample	Same trend with outliers	<u>Opposite trend without outliers</u>	Same trend with outliers
B	Calibration	Same trend with larger fluctuations in System 2	<u>Opposite trend with outliers</u>	Same trend with outliers
B	Sample	<b>Same trend without outliers</b>	<b>Same trend without outliers</b>	<b>Same trend without outliers</b>

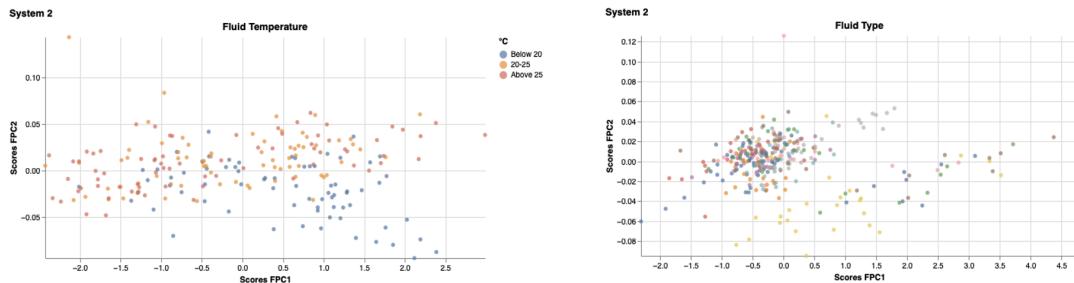
## Scores

Cards aged between 0 and 9 days tended to cluster within the calibration window for both systems and similarly within the sample window for both systems, sensor A. The same pattern appears on system 2 sensor A cal window (Figure 16).



**Figure 16. Scores of the first two components are binned by Card Age for system 1 (top) and system 2 (bottom) using sensor A in the (a) Calibration window and (b) Sample window.**

Waveforms below 20 degrees Celsius displayed negative scores across both principal components for system 2 sensor B. Moreover, fluid type samples injected with Eurotol L5 seem to be clustered (Figure 17).

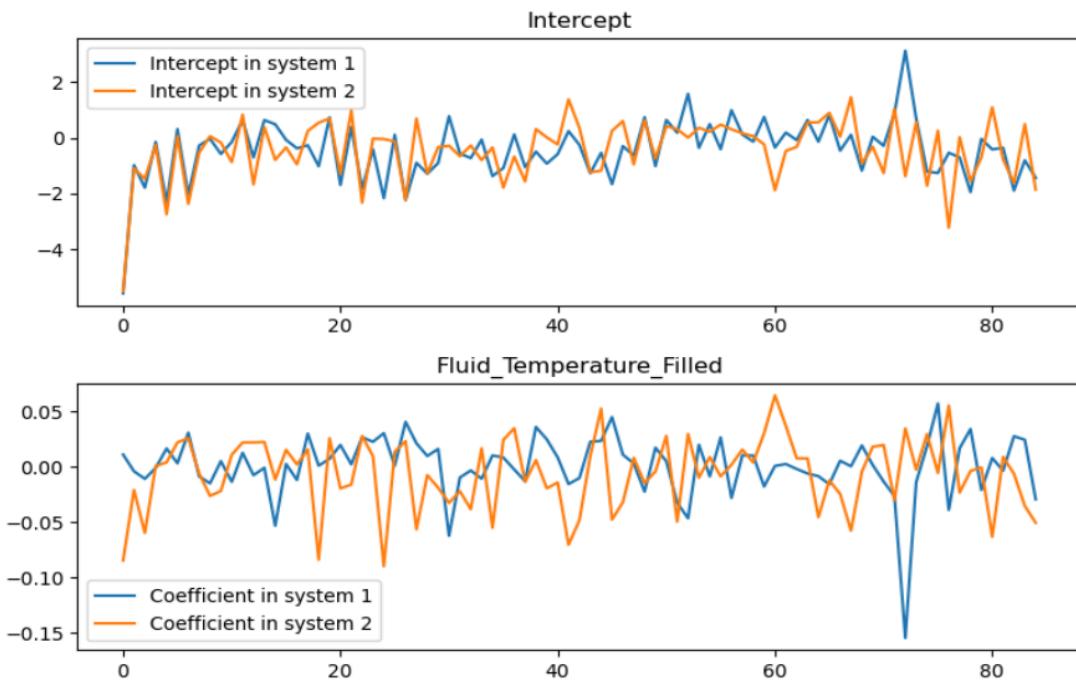


**Figure 17. Scores of the first two components for system 2 using sensor B in the sample window, with data balanced by fluid temperature (left) and fluid type (right).**

## Functional Regression

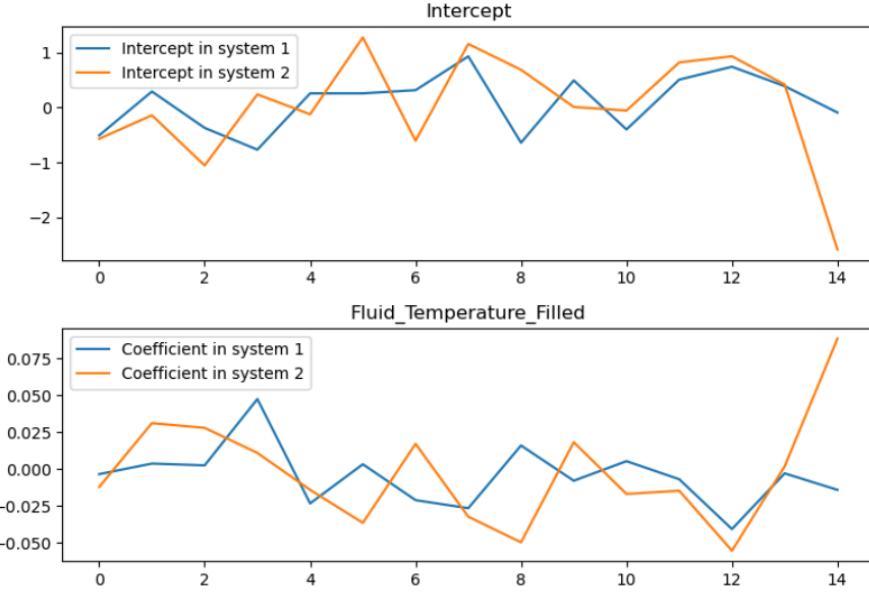
Figure 18 shows the truncated time stamps from 1 to 85 to make the plot-trend clear and focused. This truncation is necessary because the average magnitude of most coefficients is near zero, while the coefficient at the 0 timestamp is 10 times larger than those with a power of 1. Additionally, the coefficients at timestamps 86 to 90 have magnitudes more than 10 times greater. This truncation applies to all the coefficient plots as well.

The intercept displays a similar pattern across different systems, particularly in the first few time stamps (Figure 18). The fluid temperature has a small but varying effect on different systems, with the effect being negative most of the time. With the exception of timestamps around 73 seconds, system 1 is generally more stable than system 2.



**Figure 18. The coefficients from Functional Regression in the Cal window in sensor B, focusing on Fluid Temperature.** Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 85 in a sequence of 90 in total.

The coefficient of Fluid Temperature has different effects on the two systems (Figure 19). System 2 demonstrates greater negative effects compared to system 1, exemplified by timestamps 5, 8, and 12. In the last few timestamps, the coefficient in system 2 changes more rapidly than in system 1.



**Figure 19. Functional Regression coefficients within sensor B's Sample window, focusing on Fluid Temperature.** Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 15 in a sequence of 20 in total.

In sensor A, most cases in system 1 and system 2 exhibit a similar pattern, while system 2, with more fluctuations, indicates the possibility for improvement (Table 11). Furthermore, both calibration and sample windows show that these windows in system 2 sensor B, reacting more slowly, should be moved closer to the bubble window.

**Table 11: The Coefficients of Different Features Across Different Systems and Sensors.**

Sensor	Window	Fluid Temperature	Card Age
A	Calibration	Similar pattern with more fluctuations in system 2	Almost the same pattern with all negative coefficients in both systems
A	Sample	system 1 is stable around y=0; system 2 fluctuates largely	Almost same pattern with more fluctuations in system 2
B	Calibration	system 1 is more stable, except for one timestamp	system 1 is more stable
B	Sample	system 1 is more stable	system 1 is more stable

## 6.2. Secondary task: Window Optimization

### 6.2.1. New Window Boundaries

The optimized calibration and sample windows for sensor A are visualized below in Figure 20. Our results indicate that a CalDelimit value of 9.6 and a SampleDelimit value of 17.2 led to the greatest improvements to the calibration and sample windows.

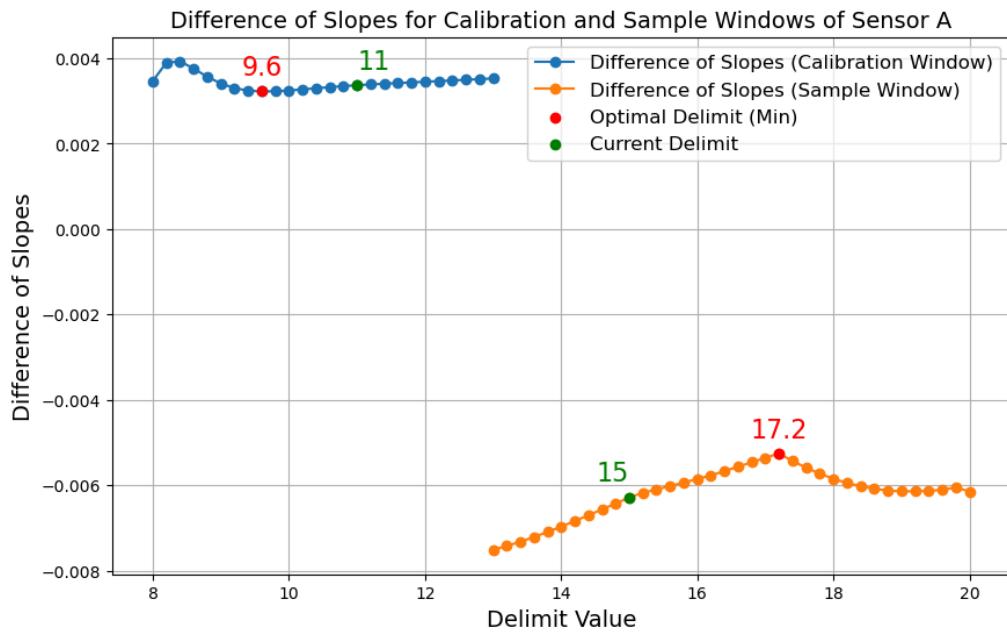
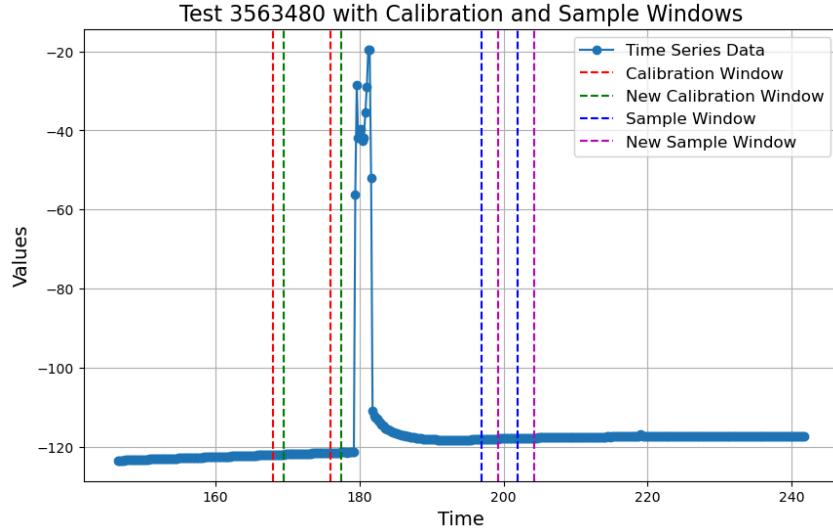


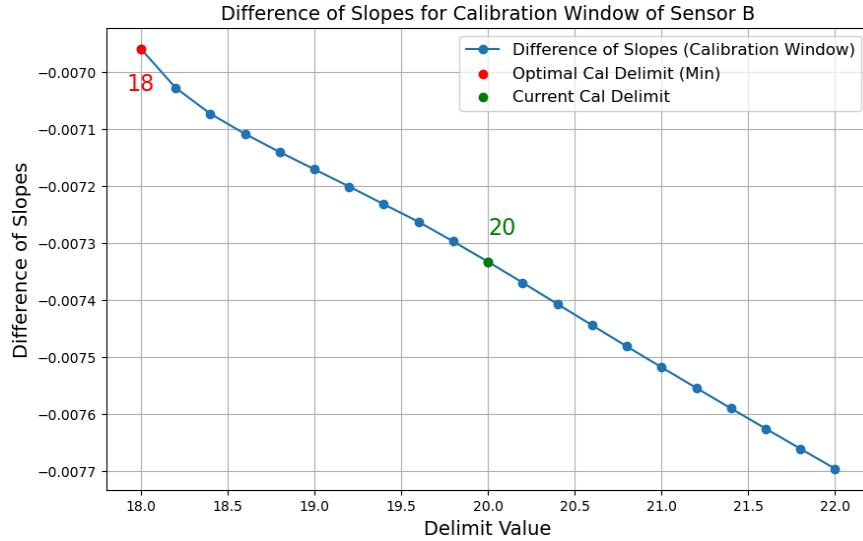
Figure 20. **Difference of slope for Calibration and Sample Windows of sensor A**

Using a random test ID from system 2 sensor A, we generated a plot with the current and new window for both calibration and sample periods of system 2 with the newly obtained delimits applied. Our results indicate that the optimized calibration window shifted towards the bubble window (Figure 21).



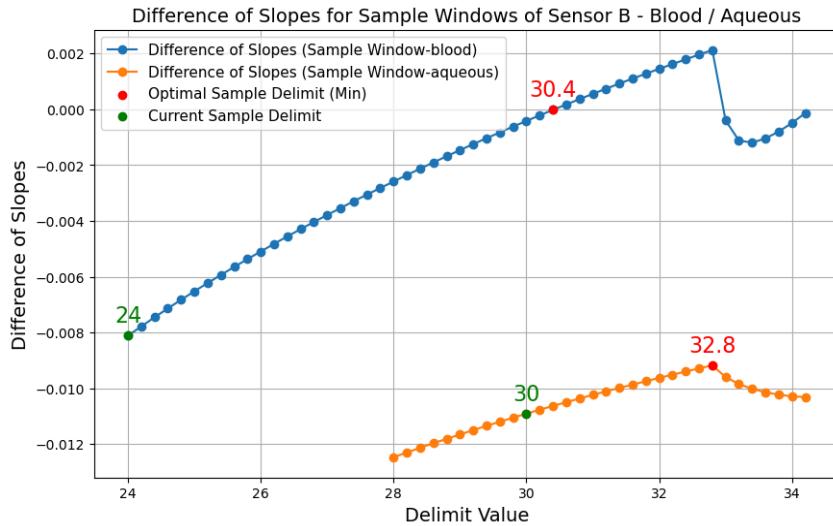
**Figure 21. Waveform of a Random Test of sensor A with New and Current Windows**

We separately visualized sensor B's calibration and sample window because sensor B has two different SampleDelimits for blood and aqueous fluid types. Our results indicate an optimized calibration window for sensor B with a CalDelimit value of 18 (Figure 22). Sensor B's optimized calibration window shifted closer to the bubble window, indicating the curve gradually flattened until detecting the air bubble for sensors A and B.



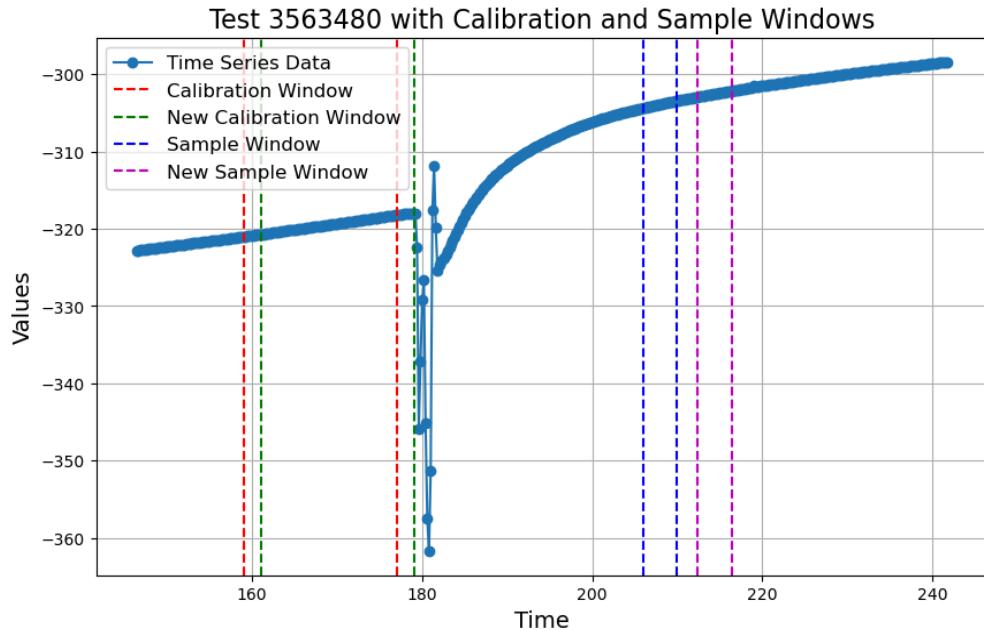
**Figure 22. Difference of Slopes for Calibration Window of sensor B**

We also obtained the improved sample window by taking a larger SampleDelimit value, which is 30.4 for blood and 32.8 for aqueous (Figure 23).



**Figure 23. Difference of Slopes for Sample Window of sensor B**

We also provided a plot with the current window and the new window for both calibration and sample periods for a randomly selected test ID of sensor B, system 2 with the newly obtained delimits applied, as shown in Figure 24.



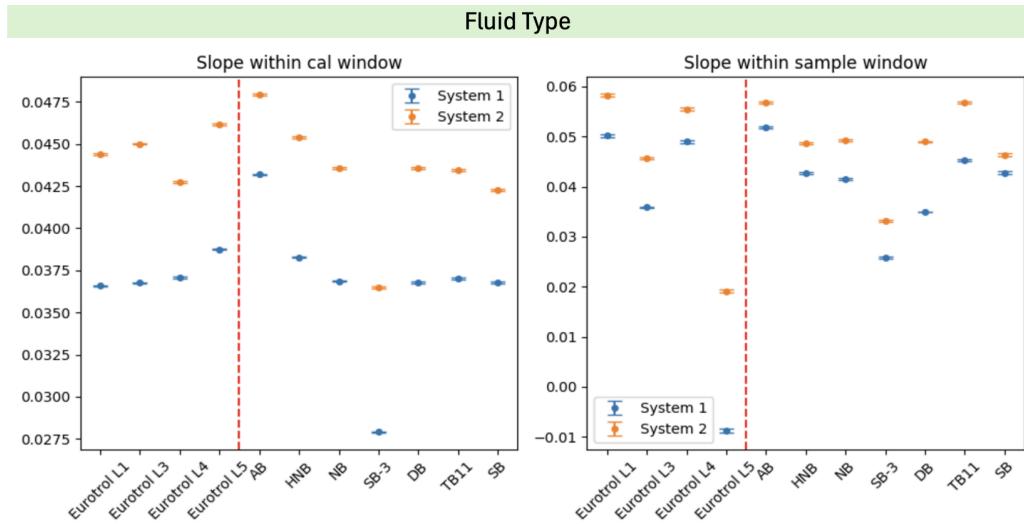
**Figure 24. Waveform of a Random Test of sensor B with New and Current Windows**

Our results indicate the optimal solutions were found at specific delimit values, and they each shifted the windows further in time. This suggests those improved windows for system 2 whose flatness levels are closest to the current ones for system 1. The new calibration and sample window for both sensors have been optimized by imitating the flatness level of system 1.

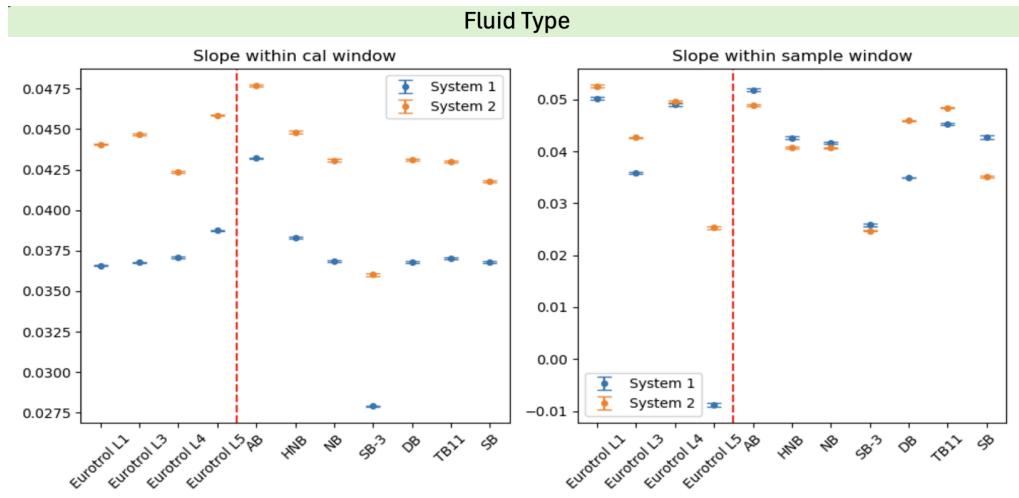
### 6.2.2. Evaluation and Comparison on Pipeline 1

Our custom window optimization function is designed to find the minimum differences between the regression slopes of the aggregate mean function between system 1 and system 2. To further explore the optimized solutions, we applied the new limits to the Flatness Evaluation Pipeline to assess the degree to which the system 2 regression slopes approach the regression slopes for system 1. For each window, the regression slope of the aggregated mean function for system 2's calibration and sample window approaches those for system 1. However, given that the slope differences between the two systems are minuscule, they can be discounted from the visualization, especially for the sensor A slope comparison of binning fluid temperature. However, the decrease in slope differences for sensor B is quite obvious, especially for the binning fluid type and test card age (Figures 25 and 26).

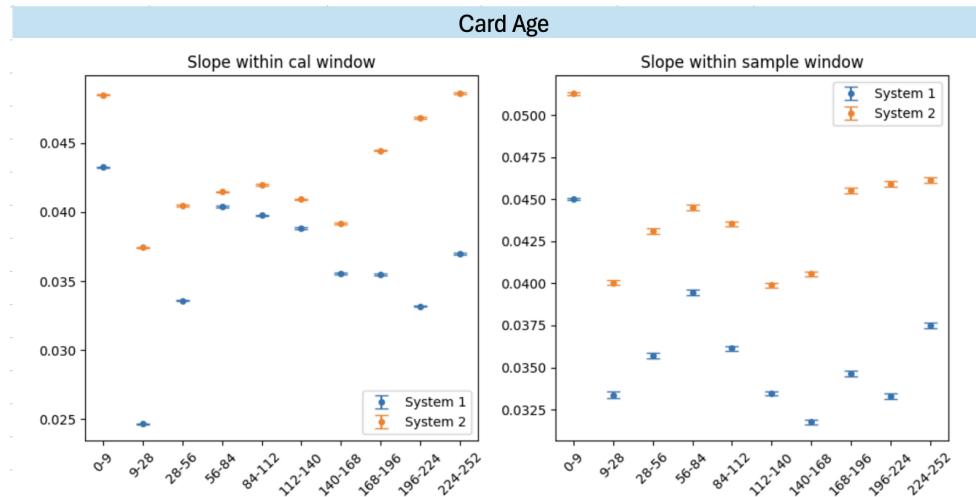
Sensor B's calibration small window changes are consistent with the results we obtained in section 6.2.1 (Figure 27). This is likely due to the smaller distance between the optimized calibration window and the bubble window. Sensor B's sample window does have a substantial difference in slopes after applying the new window limits. Our results also indicate that the new window works best with the Eurotrol L4 (aqueous), NB (Blood), and SB-3 (Blood) fluids because the slope of system 2 closely resembles the slope of system 1. Finally, the card age bin with the minimum slope differences is 56 to 84 days old. The Siemens research team may find these findings valuable for informing their future work on window optimization.

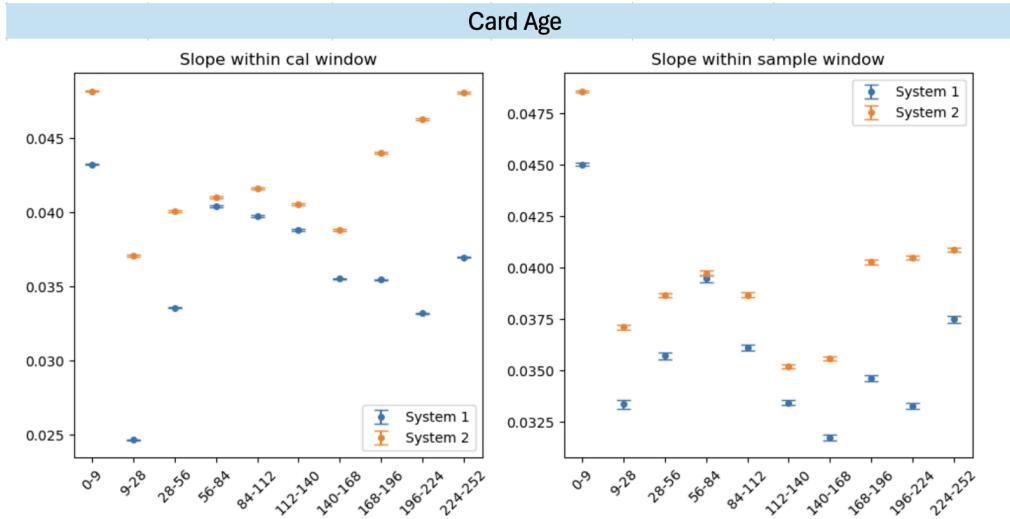


**Figure 25. sensor B Aggregated Mean Function Slope Comparison for Fluid Type Bins between Current Windows.**



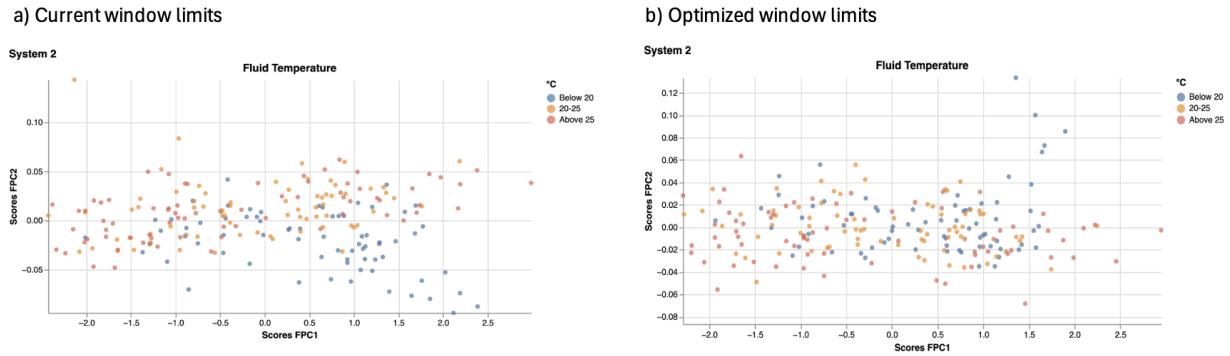
**Figure 26. sensor B Aggregated Mean Function Slope Comparison for Fluid Type Bins Optimized Windows.**





**Figure 27. Aggregated Mean Function Slope Comparison for Card Age Bins between Current Windows (Upper Plots) and Optimized Windows (Lower Plots).**

### 6.2.3. Evaluation and Comparison on Pipeline 2



**Figure 28. sensor A FPCA Score Comparison of system 2 Sample Window between Current window (left) and Optimized Window (right).**

The optimized calibration and sample windows have little overall impact on the FPCA scores. However, we found one for sensor A that has a relatively obvious change of FPC1 score patterns is the bin of Below 20°C with fluid temperature balanced (Figure 28). From the current window of system 2's FPC1 scores, we can find the scores for 50% of "Below 20°C" are scattered in the area with a score range from 0.5 to 2.5. After applying the optimized window limits, they get more clustered towards the range from 0.5 to 1.5, suggesting the scores of half "Below 20°C" tend to be smaller when we extract the sample window with the optimized limits. PCA scores evaluate how much the component would contribute to the variation of the response. Smaller scores of the "Below 20 bin" on the sensor value may imply that the contribution of tests with low fluid temperature can be shrunk on the sensor value with the newly extracted window. We

believe it would be valuable for the Siemens' research team to delve deeper into how optimizing the window relates to the low fluid temperature observed in the tests.

## 7. Conclusions

Through our capstone project, we accomplished two tasks: (1) we characterized the waveforms of system 1 and system 2, and (2) we optimized system 2's calibration and sample window. For task (1), we developed two pipelines. The overall workflows for each pipeline are designed to be easily applied to the characterization of other sensors in the future.

The first pipeline calculates and compares the window flatness in both systems using raw sensor data, with results qualitatively validated through visualizations and quantitatively validated through 2-tailed t-tests. Our analyses demonstrate that system 2's slope is consistently greater than system 1's slope, irrespective of the sensor.

The second pipeline involves functional data analysis, including FPCA and functional regression, on attribute-balanced sensor data (fluid temperature, fluid type, and binned card age) to summarize the window data characteristics in both systems. From FPCA we were able to identify differences in the trend between the two systems. When balanced by fluid temperature and fluid type, FPC1 has the same trend for all cases except sensor A's calibration window. In contrast, card age preserves the same trend in the main mode of variation across all cases.

For task (2), we experimented with different delimit values in system 2 to minimize the slope differences between system 2 and system 1. We successfully identified optimal delimit values, which are shown in the 4th and 5th columns of Table 12.

**Table 12. Delimit Values of system 1 (current) and system 2 (optimal)**

	System 1		System 2	
	CalDelimit	SampleDelimit	CalDelimit	SampleDelimit
<b>Sensor A</b>	11	15	9.6	17.2
<b>Sensor B-Blood</b>		24		30.4
<b>Sensor B-Aqueous</b>	20	30	18.0	32.8

Upon obtaining the new delimits of system 2, we tested and compared the new and old optimized windows using the pipelines generated from the first task. While most slope differences are negligible for pipeline 1, sensor B showed notable decreases in slope differences, especially for fluid type and card age. Specifically, fluid types Eurotrol L4, NB, and SB-3, and test cards aged 56 to 84 days, show minimal slope differences, which could be valuable for further investigation by Siemens. For pipeline 2, the new optimized windows demonstrated little impact, except for sensor A in the "Below 20 °C" fluid temperature bin, where scores increasingly clustered between 0.5 and 1.5 post window optimization. This suggests that lower fluid temperature tests contribute less to sensor value variation with the optimized window.

Due to the time restriction, there still exist some limitations and some aspects that can be improved in the future. Future directions for this project can include the following:

- **Expanding sensor types:** Apply the developed pipelines to additional sensors in both systems to further validate and enhance the characterization and optimization processes.
- **Characterization of the post window:** Include a detailed characterization of the sensor data in the post window period.
- **Involvement of other test attributes:** Expand the analysis to include additional test attributes like humidity, fluid flow rate, etc.
- **Advanced window optimization:** Optimize the calibration, sample, and post window by changing both the window size and delimits.

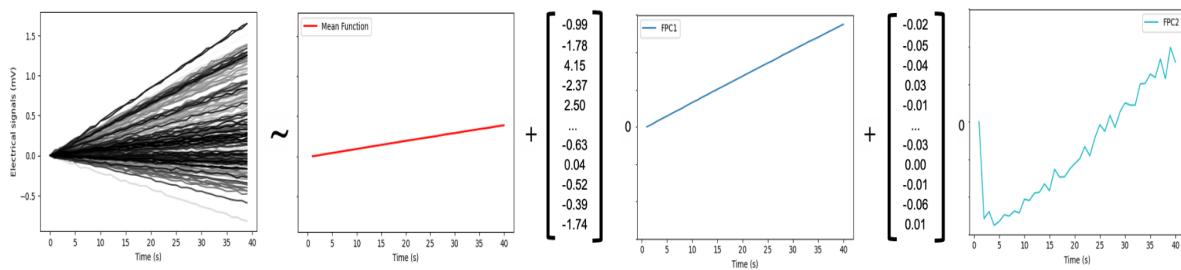
## Appendices

### Appendix A. Methods and Techniques

#### Functional Principal Components

Functional principal component analysis (FPCA) is a statistical method used to analyze data represented as functions rather than vectors. FPCA extends principal component analysis (PCA) to functional data, providing a way to identify the dominant modes of variation throughout the visualization of eigenfunctions [7]. The Functional principal components show the patterns of variation in the data that differ from the mean function, which is computed as the column means across all functions.

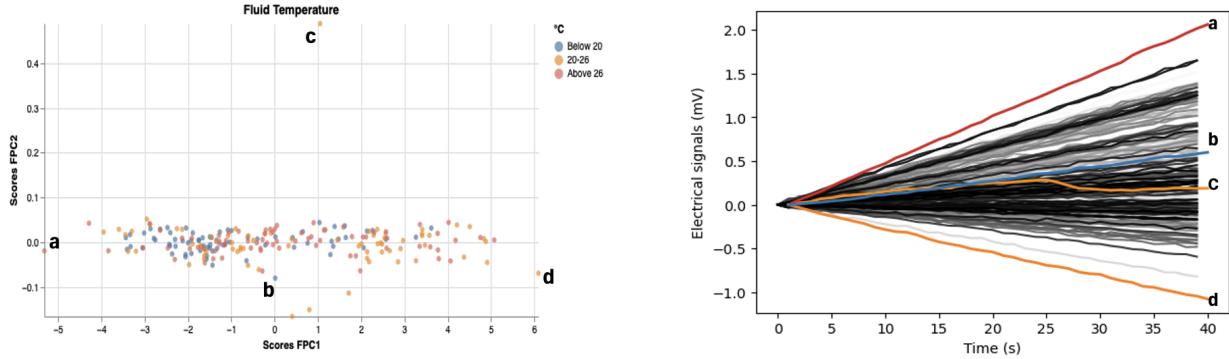
FPCA decomposes the waveforms into a mean component and functional principal components, each weighted by their contribution to the waveform. The FPCs and their weights (scores) together capture the waveform fluctuations [5]. Figure A1 is a representation of the waveform decomposition for the zero-aligned data.



*Figure A1. Waveform decomposition for the zero-aligned data [6].*

The Principal Component scores (PCscores) are numerical representations of each observation's contribution to the principal components extracted through functional PCA. These scores indicate the degree to which each observation aligns with the principal components and provides insights into the underlying patterns or structures within the data. Refer to Ramsay and Silverman [3] for more details about PCA and eigenanalysis.

The scores for the dataset balanced by fluid temperature within the cal window are displayed in Figure A2. It is evident that for the zero-aligned data representation, the outlier labeled as "c" declines at timestamp 25. Large negative scores, denoted with the label "a" correspond to waveforms with an evident upward trend, while large positive scores, denoted with the label "d" correspond to waveforms with a noticeable downward trend. Scores around 0 on the y-axis are close to the mean function, denoted with the label "b".



**Figure A2. Scatterplot of the scores from the first two components, with FPC1 scores on the y-axis and FPC2 scores on the x-axis. The dataset balanced by fluid temperature in the cal window is used, with zero-aligned centered data. Each point represents a waveform and is color-coded by binned temperature.** The letters a, b, c, and d indicate specific points that correspond to the original waveforms in the second column plots.

### Bootstrap

Bootstrap is a statistical method used to estimate the uncertainty of sample statistics by resampling with replacement from the original data. It helps in assessing parameters, testing hypotheses, and validating models without strict distributional assumptions, making it versatile and widely applicable across various statistical analyses.

In this study, we employed undersampling bootstrap, where we selected the bin with the minimum data volume as the reference for sampling. Samples were drawn with replacement from other bins to equalize the data volumes across all bins. After each sampling iteration, Functional Principal Component Analysis (FPCA) was conducted to obtain the first principal component. A total of 100 bootstrap resamples, resulted in one hundred samples of the first principal components for subsequent boxplot preparation.

### Functional boxplot

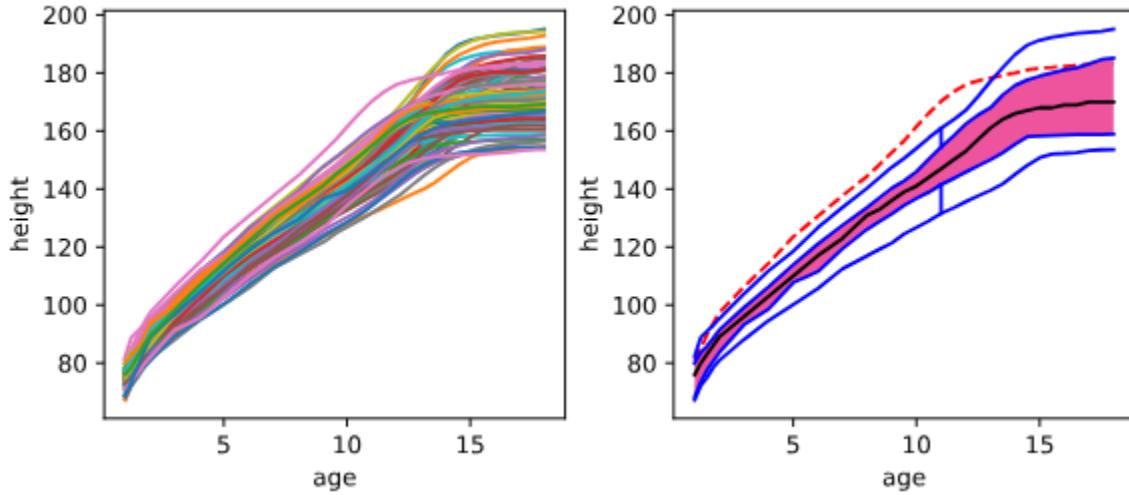
The functional boxplot, introduced by Sun and Genton (2011) [14], extends the traditional univariate boxplot to handle functional data. Key components include:

- Functional Median: The deepest curve in the dataset, representing central tendency.
- Central Envelope: Encloses the deepest 50% of observations around the median.
- Maximum Non-Outlying Envelope: Defined by scaling the central envelope by a constant factor (default is 1.5), marking the boundary for outliers.

Within Python's *scikit-fda* package, the *Boxplot* class generates functional boxplots, by setting the specific parameters we can identify outliers. Key features include customizable depth measures and centered

bands that cover user-defined fractions of the deepest observations. We focused on utilizing Modified Band Depth (MBD), an integrated depth measure that enhances statistical robustness by considering both the number of bands containing a function and the time spent within each band [15].

Functional boxplots effectively visualize central tendency, variability, and potential outliers in functional data, enhancing exploratory data analysis in functional data analysis (FDA). Figure A3 illustrates an application demonstrating these capabilities, showcasing the versatility of the package in analyzing and visualizing functional datasets.



**Figure A3: Functional boxplots of the Berkeley Growth Study dataset. Left: The original curves. Right: Standard functional boxplot with the black line representing the functional median, the central envelope is displayed as a pink band around the median, blue whiskers and their fences mark the maximum non-outlying envelope, and outliers are represented as red dashed lines.**

### Functional Regression

Functional response regression involves modeling functional responses on a set of scalar predictors, which is called Functional linear models with functional responses, or functional response regression[8][9]. Given a sample of functional responses  $Y_i(t_j)$  where  $i = 1, \dots, N$  and  $j = 1, \dots, N$ , and scalar predictors  $X_{ia}$ , where  $a = 1, \dots, p$  a general linear functional response regression model is expressed as follows:

$$Y_i(t_j) = \beta_0(t_j) + \sum_{a=1}^p X_{ia} \beta_a(t_j) + \varepsilon_i(t_j), \quad (1)$$

where:

- $Y_i(t_j)$  represents the functional response for the  $i$ -th individual at time  $t_j$ .
- $\beta_0(t_j)$  is the intercept function at time  $t_j$ .
- $\beta_a(t_j)$  are the coefficient functions for the  $a$ -th predictor at time  $t_j$ .
- $\varepsilon_i(t_j)$  denotes the error term for the  $i$ -th individual at time  $t_j$ , representing the curve-to-curve residual error deviations. These errors are often assumed to be independently and identically distributed (iid) with mean zero and follow a Gaussian distribution, with covariance  $S(t_i, t_j)$ .

## Appendix B. Plots

### 4.2.2 Windowed data after zero-alignment and subsequent centering by column

#### 6.1.1.1 Window Flatness Evaluation

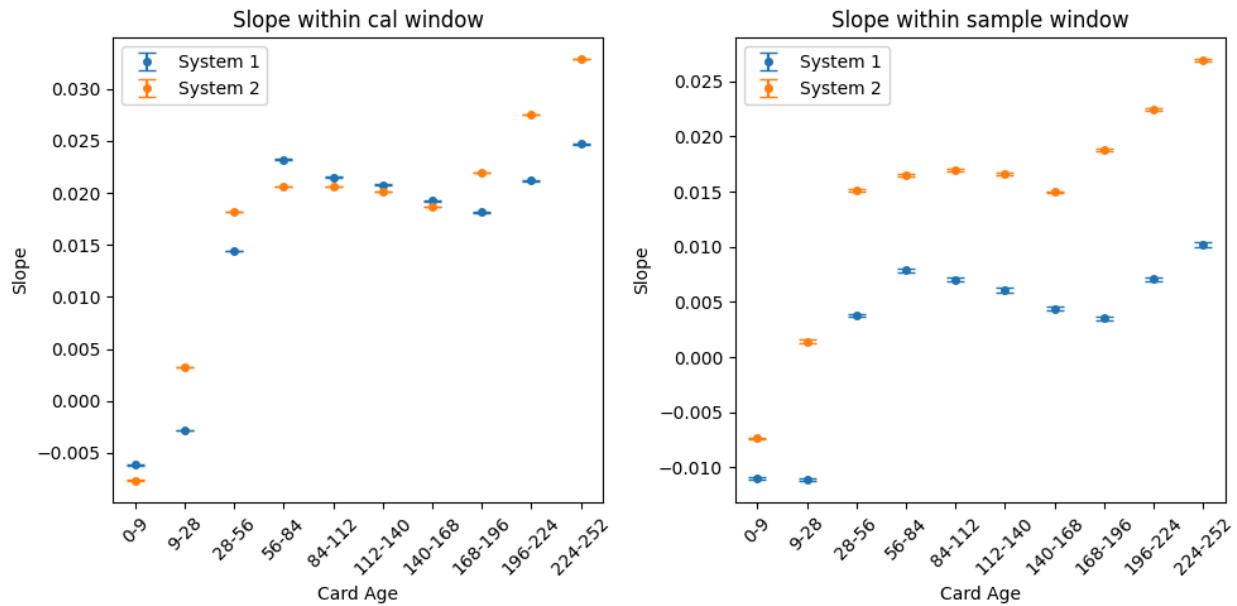
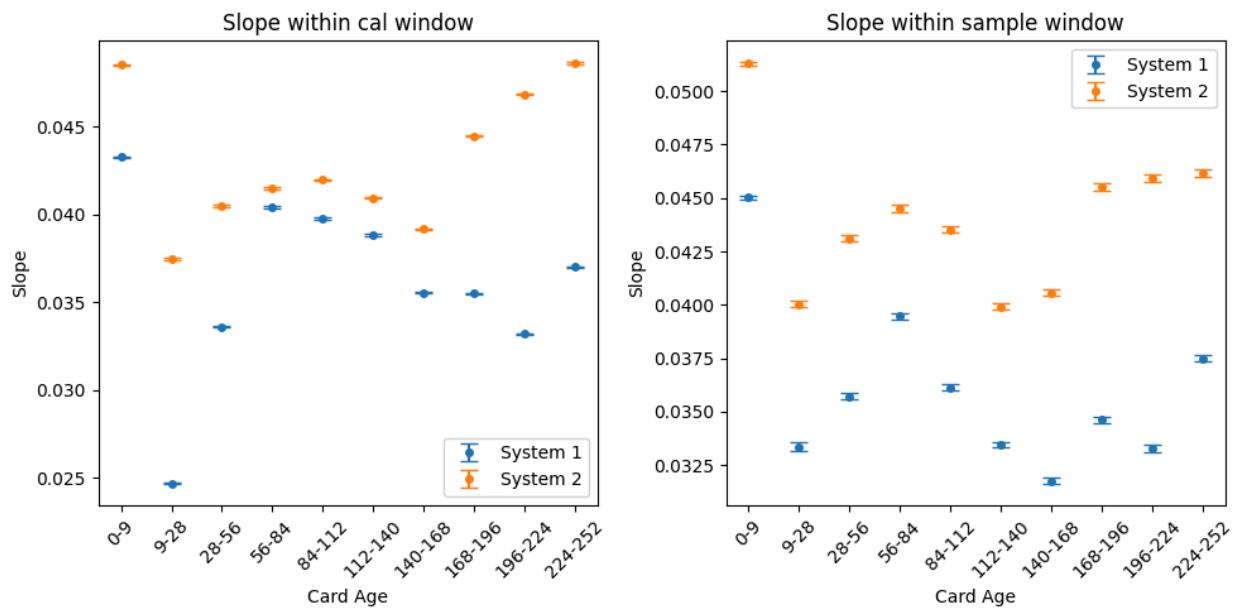


Figure A4. Slope of the waveforms for system 1 calibration window (left) and system 2 sample window (right) sensor A against card age. X-axis labels represent card age bins.

**Table A1. T-statistic and p-value comparing the Slope of system 1 and system 2 sensor A, binned by card age. The first two columns represent the calibration window and the latter two columns represent the sample window.**

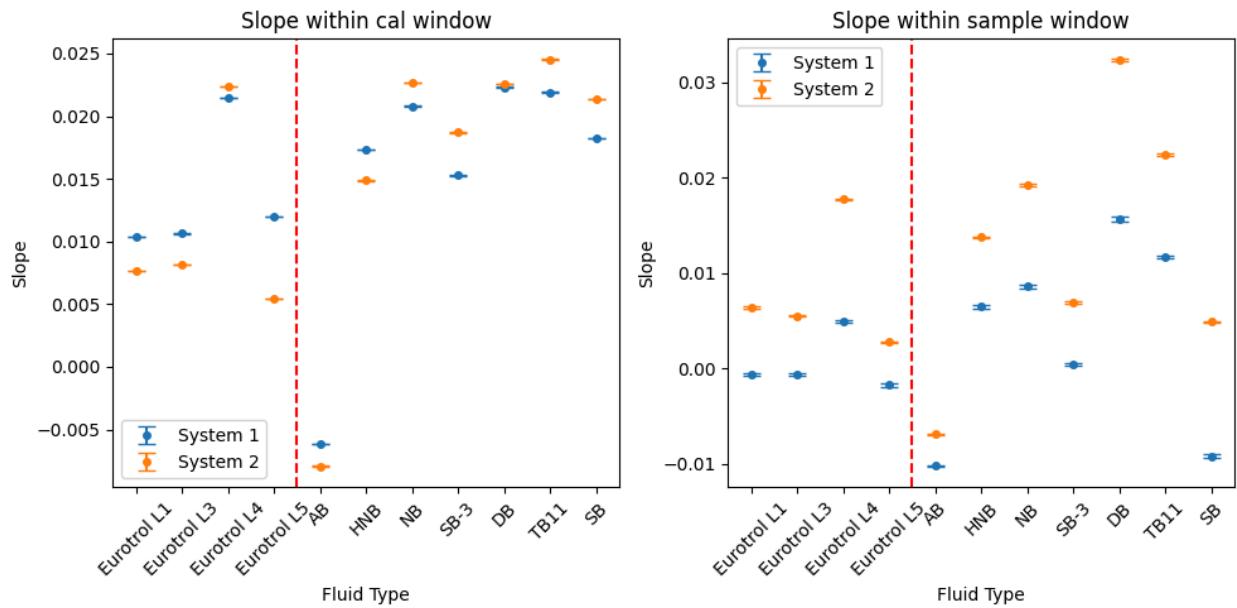
	t-statistic-cal	p-value-cal	t-statistic-sample	p-value-sample
0-9	41.372843	0.0	-31.150710	0.0
9-28	-361.855274	0.0	-72.445685	0.0
28-56	-173.716596	0.0	-62.487666	0.0
56-84	66.729533	0.0	-38.391019	0.0
84-112	26.060094	0.0	-45.516297	0.0
112-140	15.782989	0.0	-50.740842	0.0
140-168	14.334316	0.0	-49.307080	0.0
168-196	-86.602288	0.0	-68.267739	0.0
196-224	-143.865271	0.0	-69.061333	0.0
224-252	-155.984248	0.0	-72.677117	0.0



**Figure A5. Slope of the waveforms for system 1 calibration window (left) and system 2 sample window (right) sensor B against card age. X-axis labels represent card age bins.**

**Table A2. T-statistic and p-value comparing the Slope of system 1 and system 2 sensor B, binned by card age. The first two columns represent the calibration window and the latter two columns represent the sample window.**

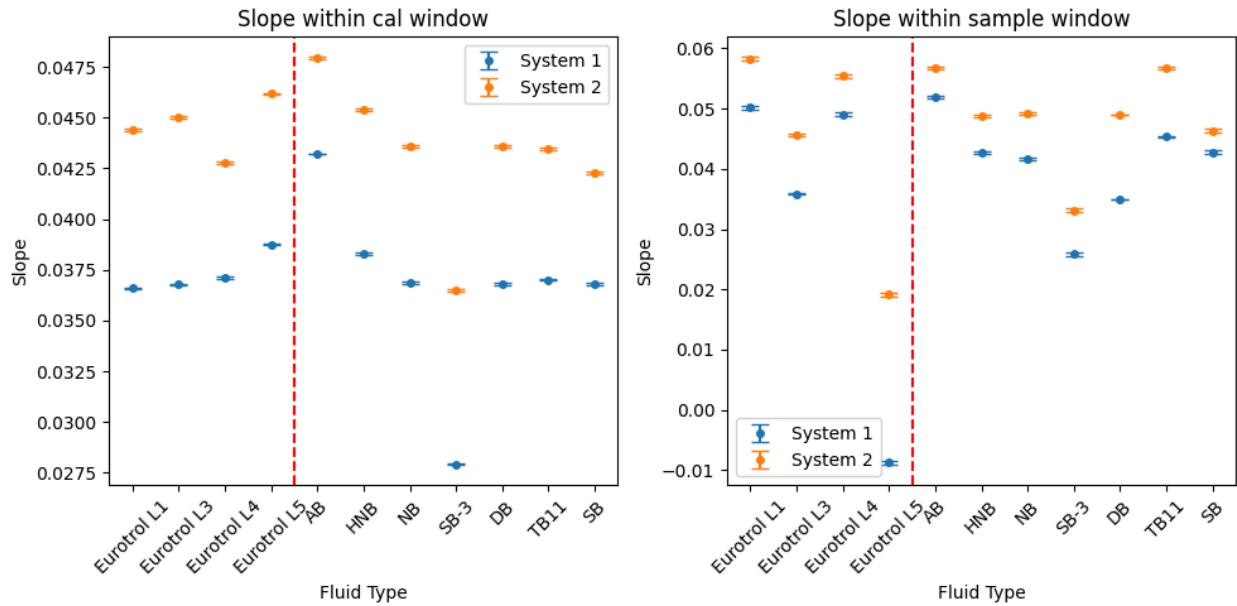
	t-statistic-cal	p-value-cal	t-statistic-sample	p-value-sample
0-9	-111.365390	0.0	-53.852949	0.0
9-28	-267.120423	0.0	-27.882981	0.0
28-56	-123.227264	0.0	-34.342816	0.0
56-84	-14.308112	0.0	-20.140767	0.0
84-112	-33.840359	0.0	-34.454003	0.0
112-140	-31.266411	0.0	-32.960107	0.0
140-168	-58.204522	0.0	-43.582285	0.0
168-196	-121.069388	0.0	-47.298521	0.0
196-224	-182.233730	0.0	-54.762169	0.0
224-252	-138.746938	0.0	-37.959927	0.0



**Figure A6. Slope of the waveforms for system 1 calibration window (left) and system 2 sample window (right) sensor A against fluid type. X-axis labels represent fluid types, where those left of the red line are aqueous fluids and those to the right of the red line are blood fluids.**

**Table A3. T-statistic and p-value comparing the Slope of system 1 and system 2 sensor A, binned by fluid type. The first two columns represent the calibration window and the latter two columns represent the sample window.**

	t-statistic-cal	p-value-cal	t-statistic-sample	p-value-sample
Eurotrol L1	120.648803	0.000000e+00	-38.175330	0.0
Eurotrol L3	99.080785	0.000000e+00	-31.865951	0.0
Eurotrol L4	-25.201529	0.000000e+00	-63.163997	0.0
Eurotrol L5	210.711821	0.000000e+00	-22.393849	0.0
AB	38.737782	0.000000e+00	-31.428874	0.0
HNB	63.064327	0.000000e+00	-37.120288	0.0
NB	-48.206279	0.000000e+00	-53.138222	0.0
SB-3	-89.805244	0.000000e+00	-35.442008	0.0
DB	-7.151951	3.249729e-10	-54.585209	0.0
TB11	-53.365664	0.000000e+00	-68.303467	0.0
SB	-69.254306	0.000000e+00	-60.123294	0.0

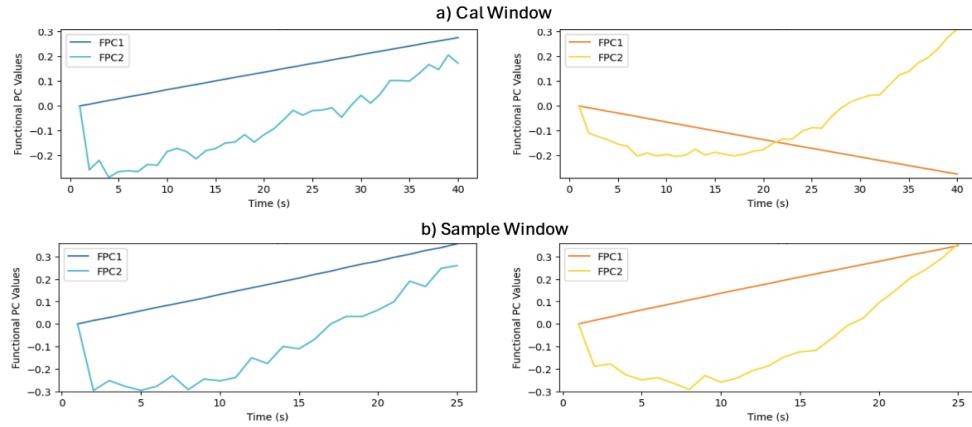


**Figure A7. Slope of the waveforms for system 1 calibration window (left) and system 2 sample window (right) sensor B against fluid type. X-axis labels represent fluid types, where those left of the red line are aqueous fluids and those to the right of the red line are blood fluids.**

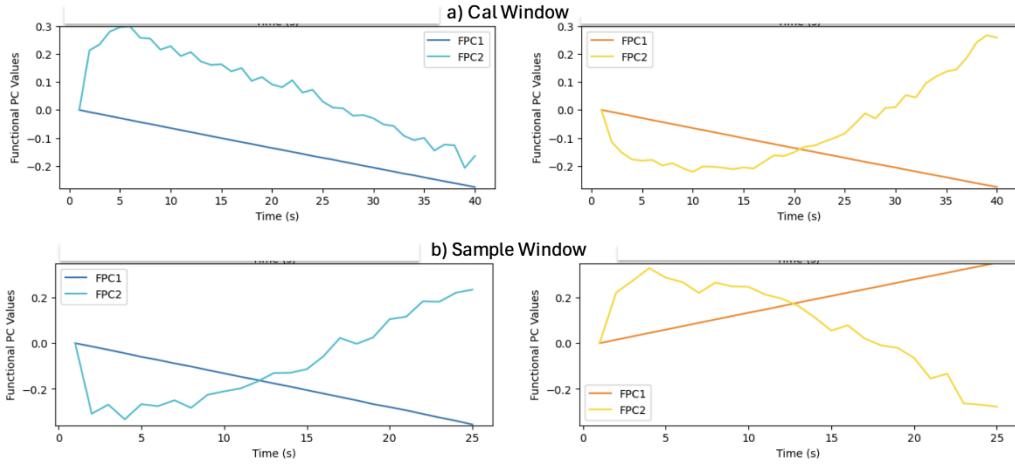
**Table A4. T-statistic and p-value comparing the Slope of system 1 and system 2 sensor B, binned by fluid type. The first two columns represent the calibration window and the latter two columns represent the sample window.**

	t-statistic-cal	p-value-cal	t-statistic-sample	p-value-sample
Eurotrol L1	-139.256939	0.0	-19.874456	0.000000e+00
Eurotrol L3	-142.659890	0.0	-49.285859	0.000000e+00
Eurotrol L4	-78.515222	0.0	-15.089769	0.000000e+00
Eurotrol L5	-119.187257	0.0	-58.905381	0.000000e+00
AB	-117.943766	0.0	-17.336125	0.000000e+00
HNB	-107.139560	0.0	-21.331796	0.000000e+00
NB	-93.591519	0.0	-26.201890	0.000000e+00
SB-3	-160.000107	0.0	-21.706098	0.000000e+00
DB	-91.034062	0.0	-170.294684	0.000000e+00
TB11	-82.631120	0.0	-42.634646	0.000000e+00
SB	-71.465266	0.0	-8.067198	4.509062e-10

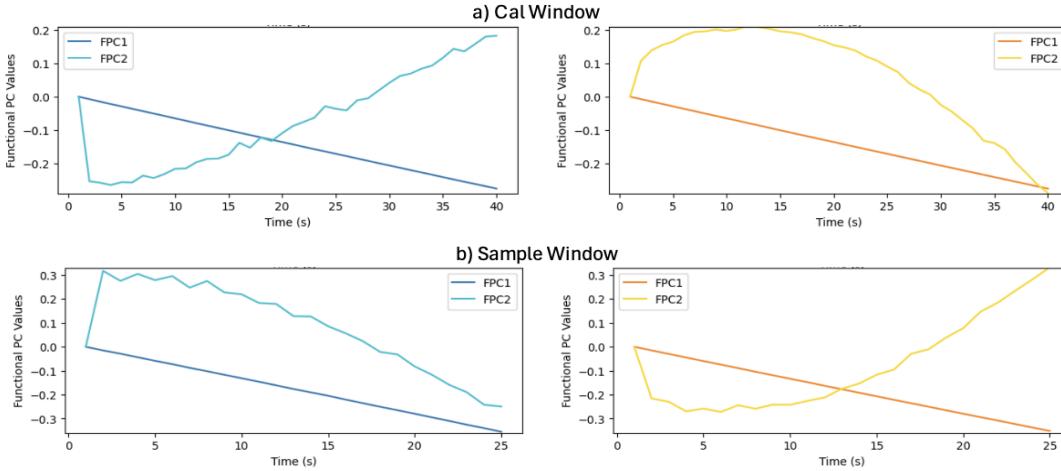
#### 6.1.2.1 First Component (FPC1) and Second Component (FPC2) for system 1 and system 2



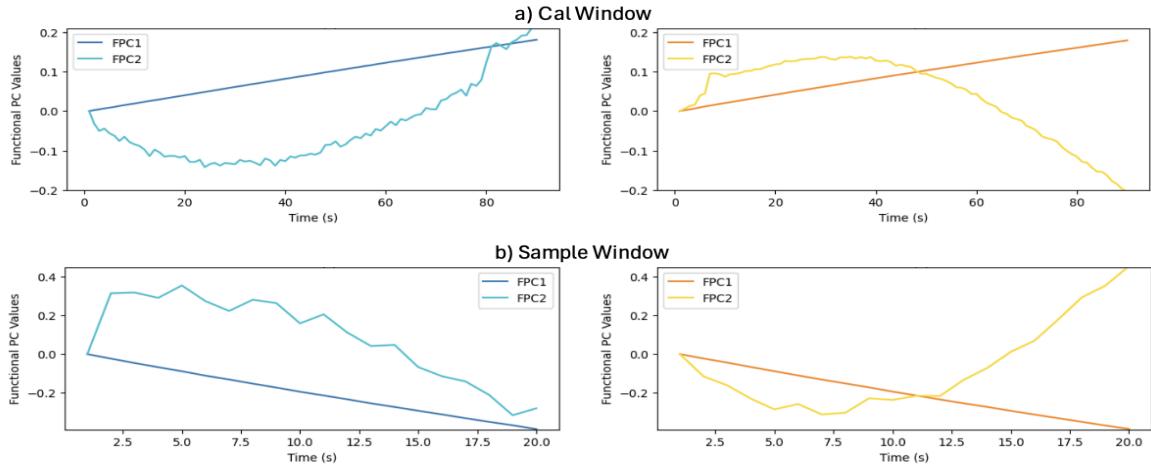
**Figure A8. First Component (FPC1) and Second Component (FPC2) for system 1 (left) and system 2 (right) using sensor A, with window data balanced by Fluid Temperature: (a) Calibration window and (b) Sample window.**



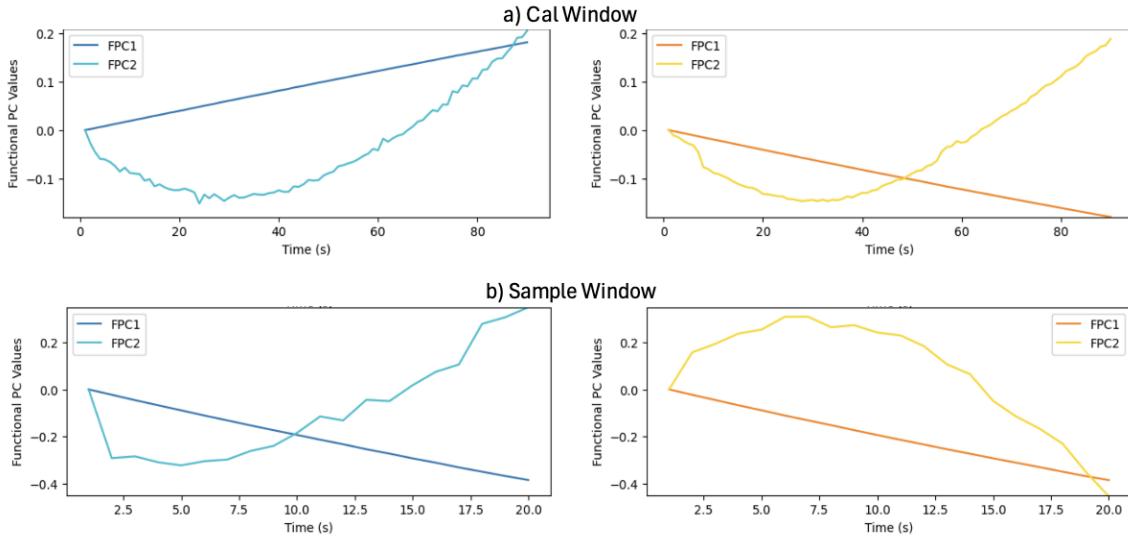
**Figure A9. First Component (FPC1) and Second Component (FPC2) for system 1 (left) and system 2 (right) using sensor A, with window data balanced by Fluid Type: (a) Calibration window and (b) Sample window.**



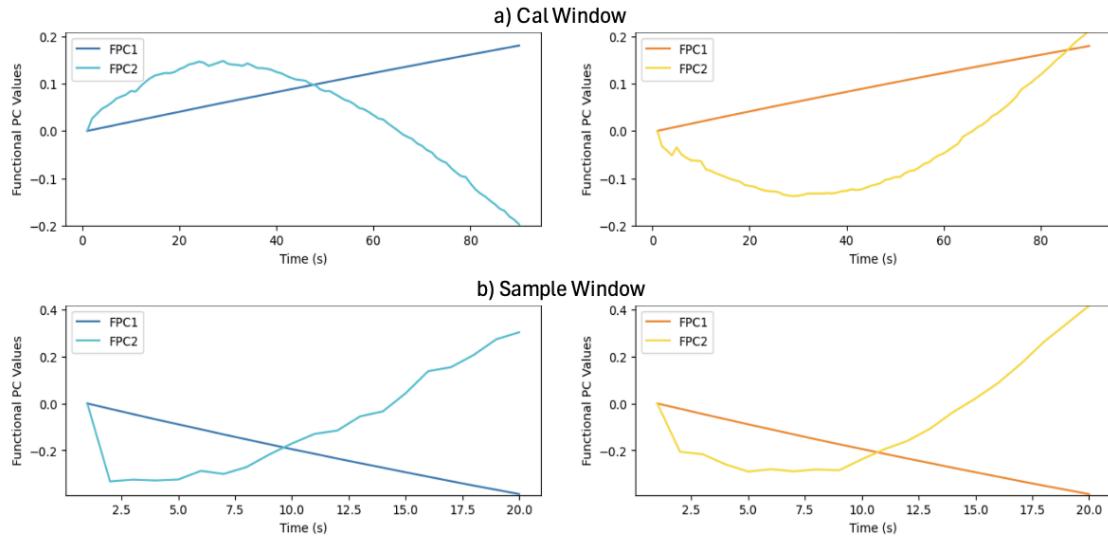
**Figure A10. First Component (FPC1) and Second Component (FPC2) for system 1 (left) and system 2 (right) using sensor A, with window data balanced by Card Age: (a) Calibration window and (b) Sample window.**



**Figure A11. First Component (FPC1) and Second Component (FPC2) for system 1 (left) and system 2 (right) using sensor B, with window data balanced by Fluid Temperature: (a) Calibration window and (b) Sample window.**

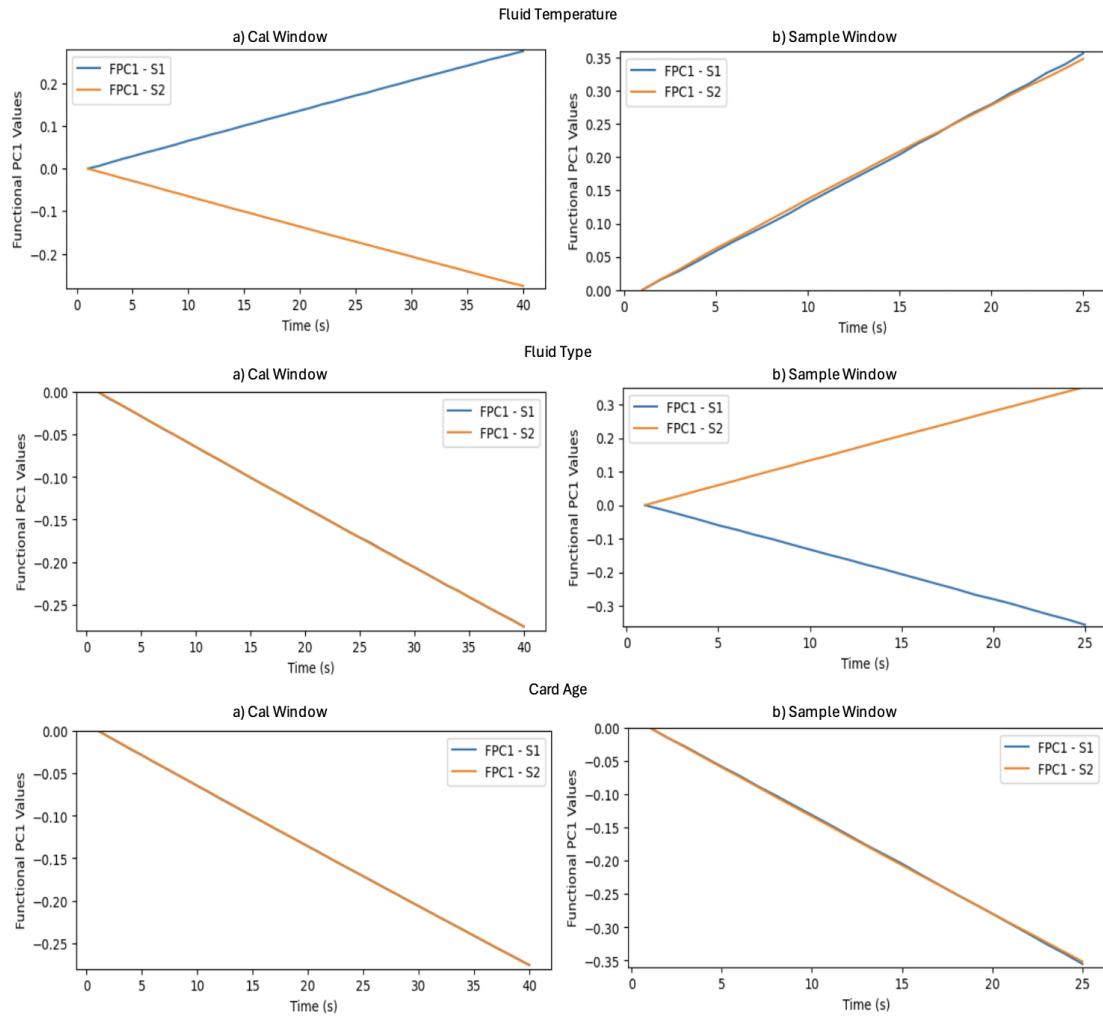


**Figure A12. First Component (FPC1) and Second Component (FPC2) for system 1 (left) and system 2 (right) using sensor B, with window data balanced by Fluid Type: (a) Calibration window and (b) Sample window.**

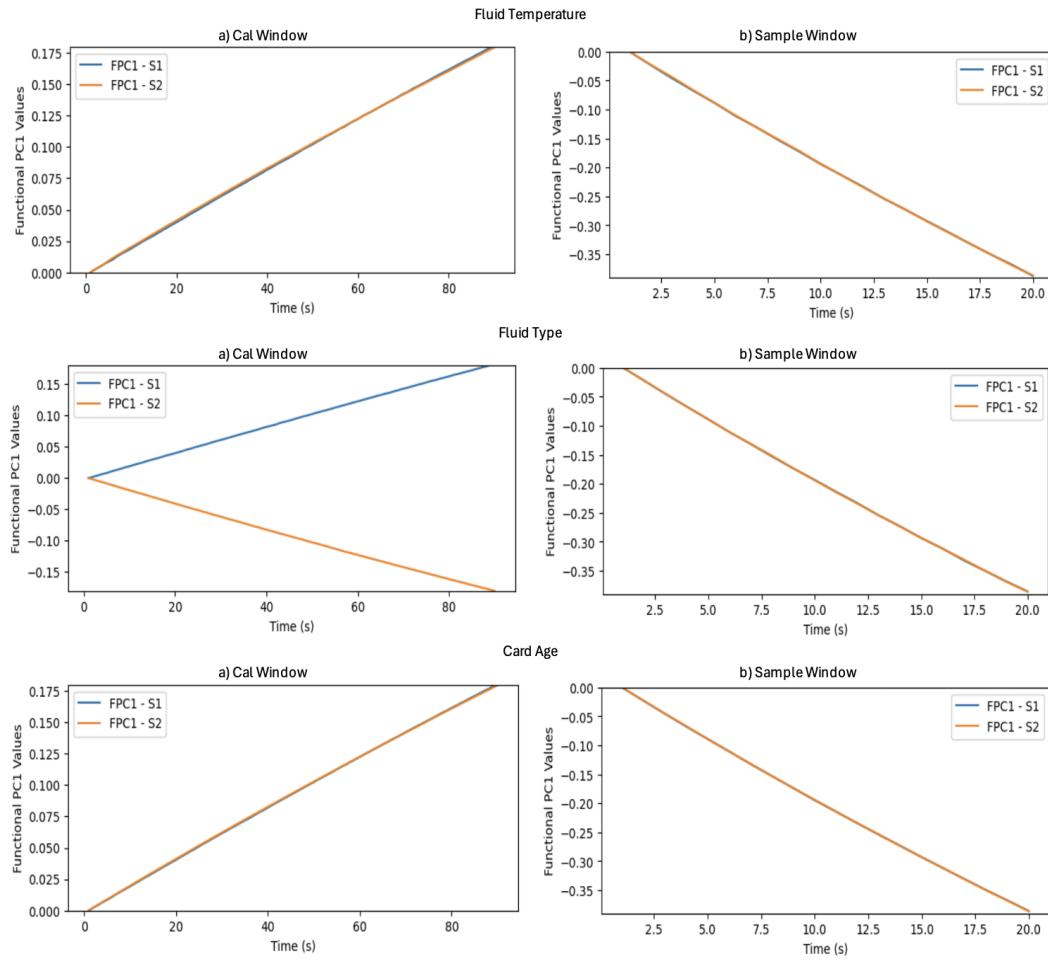


**Figure A13. First Component (FPC1) and Second Component (FPC2) for system 1 (left) and system 2 (right) using sensor B, with window data balanced by Card Age: (a) Calibration window and (b) Sample window.**

### 6.1.2.2 Comparison of the First Component (FPC1) for two systems



**Figure A14. Comparison of the First Component (FPC1) for two systems (system 1: S1 and system 2: S2) using sensor A, with windows data balanced by attribute: Fluid Temperature (top), Fluid Type (middle), and Card Age (bottom) in the (a) Calibration window and (b) Sample window.**



**Figure A15. Comparison of the First Component (FPC1) for two systems (system 1: S1 and system 2: S2) using sensor B, with windows data balanced by attribute: Fluid Temperature (top), Fluid Type (middle), and Card Age (bottom) in the (a) Calibration window and (b) Sample window.**

### 6.1.2.3 Scores of the first two components for system 2 using sensor B in the sample window

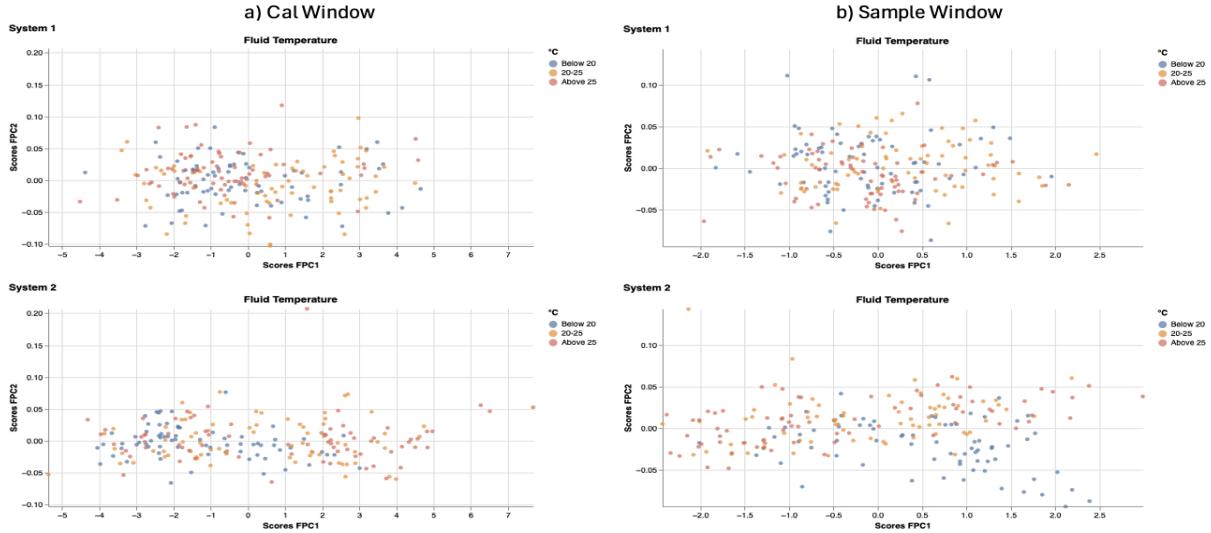


Figure A16. Scores of the first two components are binned by Fluid Temperature for system 1 (top) and system 2 (bottom) using sensor A in the (a) Calibration window and (b) Sample window.

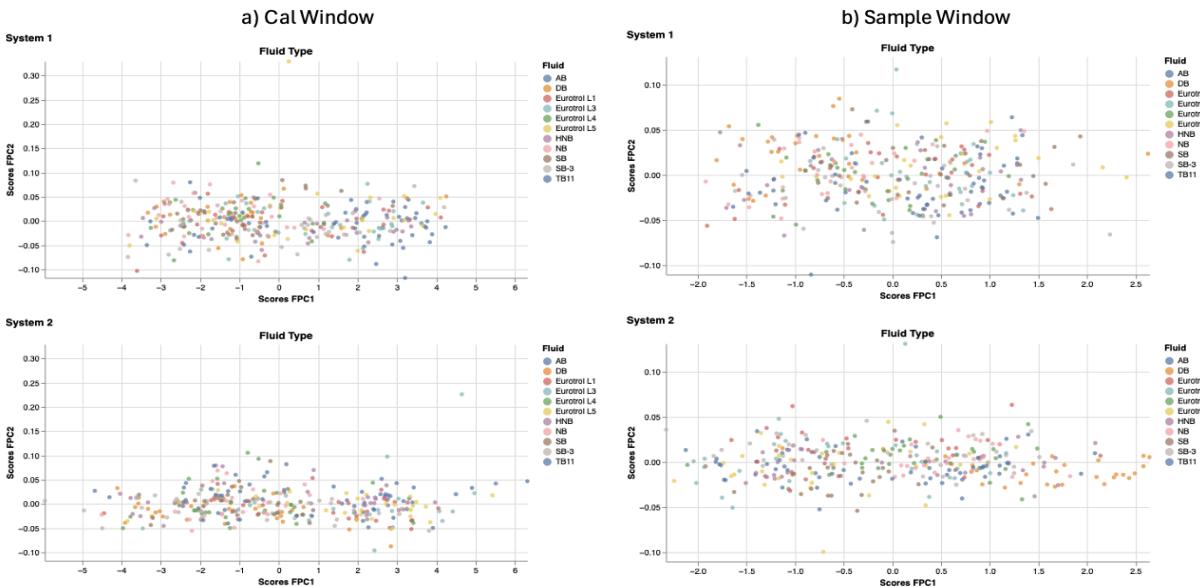
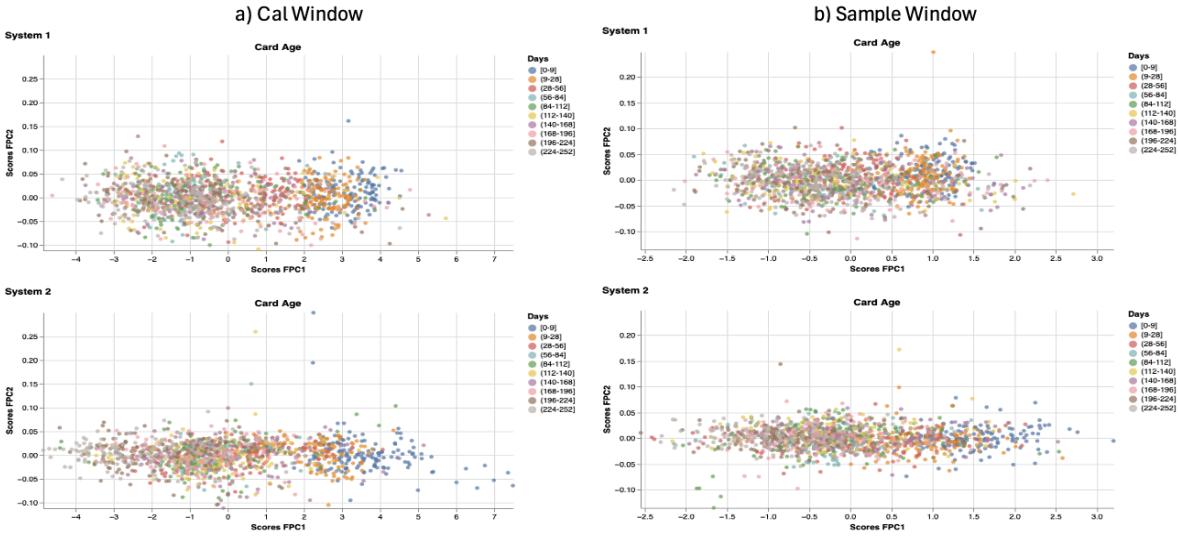
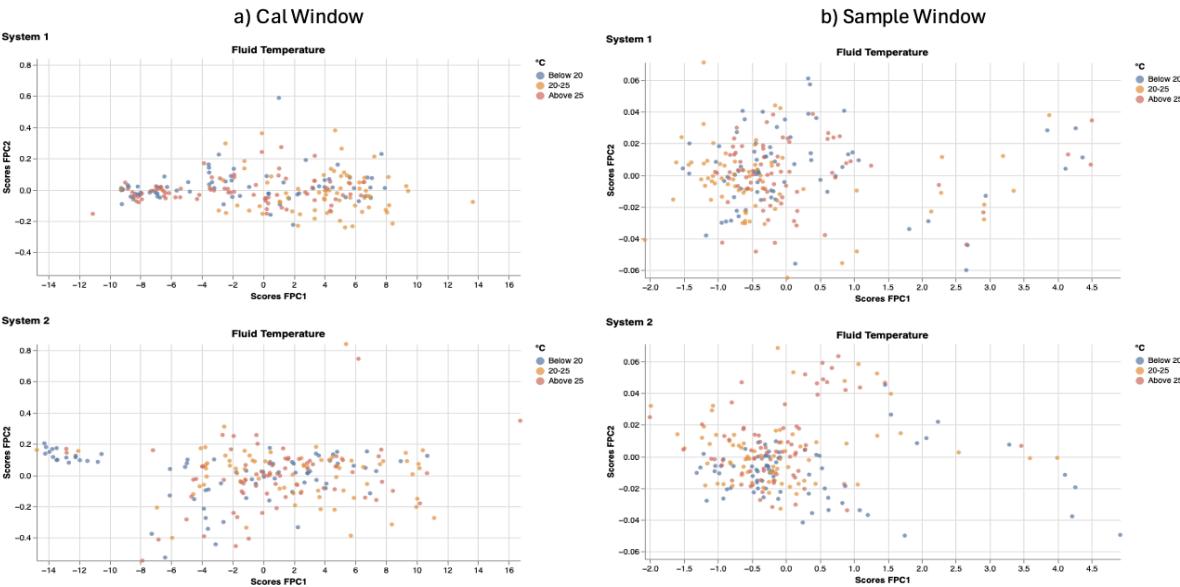


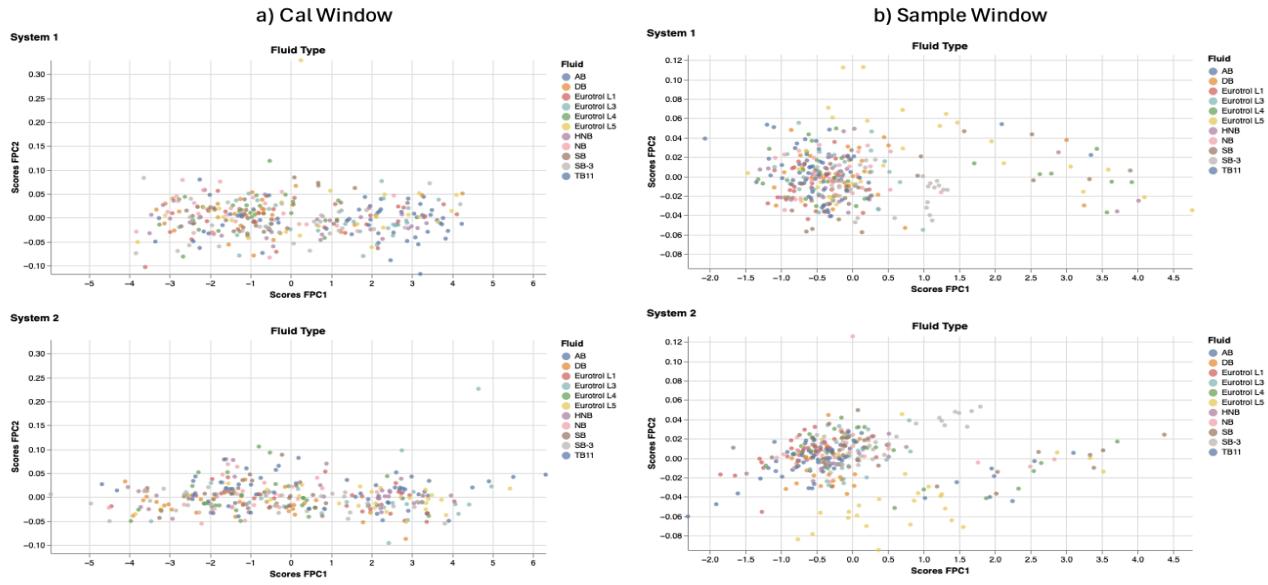
Figure A17. Scores of the first two components are binned by Fluid Type for system 1 (top) and system 2 (bottom) using sensor A in the (a) Calibration window and (b) Sample window.



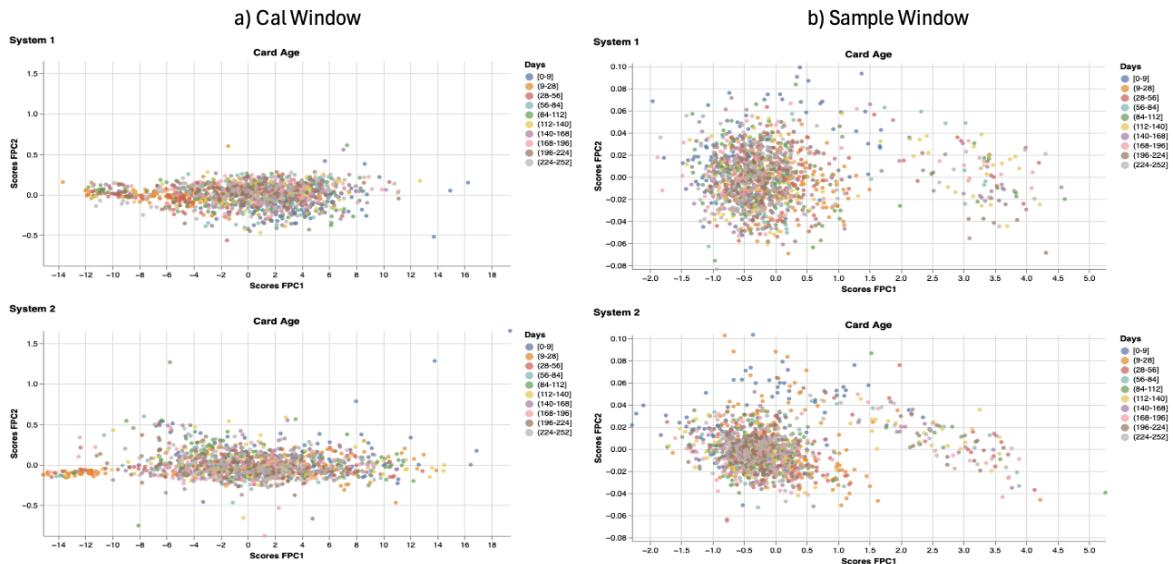
**Figure A18. Scores of the first two components are binned by Card Age for system 1 (top) and system 2 (bottom) using sensor A in the (a) Calibration window and (b) Sample window.**



**Figure A19. Scores of the first two components are binned by Fluid Temperature for system 1 (top) and system 2 (bottom) using sensor B in the (a) Calibration window and (b) Sample window.**

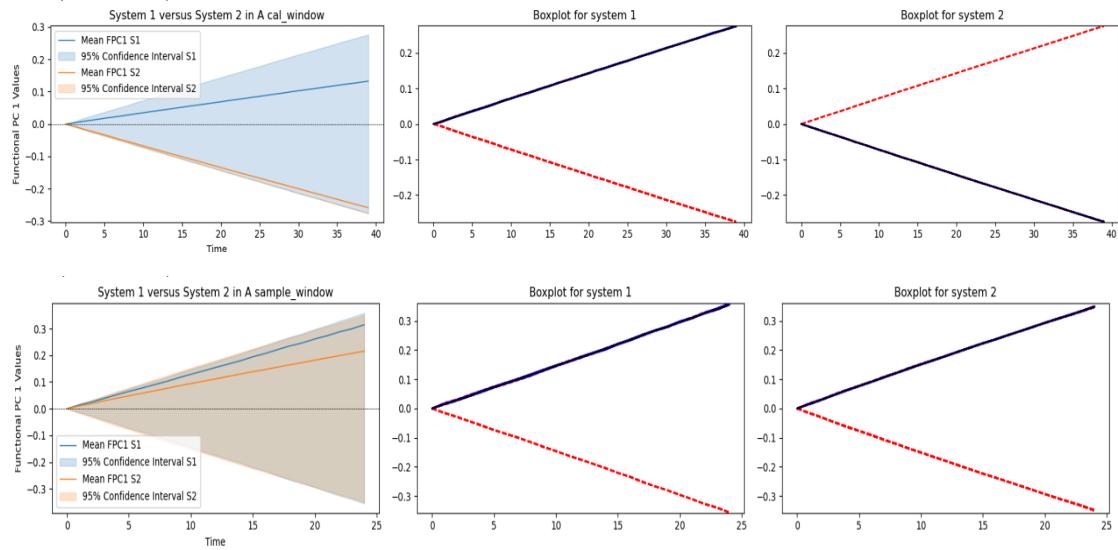


**Figure A20.** Scores of the first two components are binned by Fluid Type for system 1 (top) and system 2 (bottom) using sensor B in the (a) Calibration window and (b) Sample window.

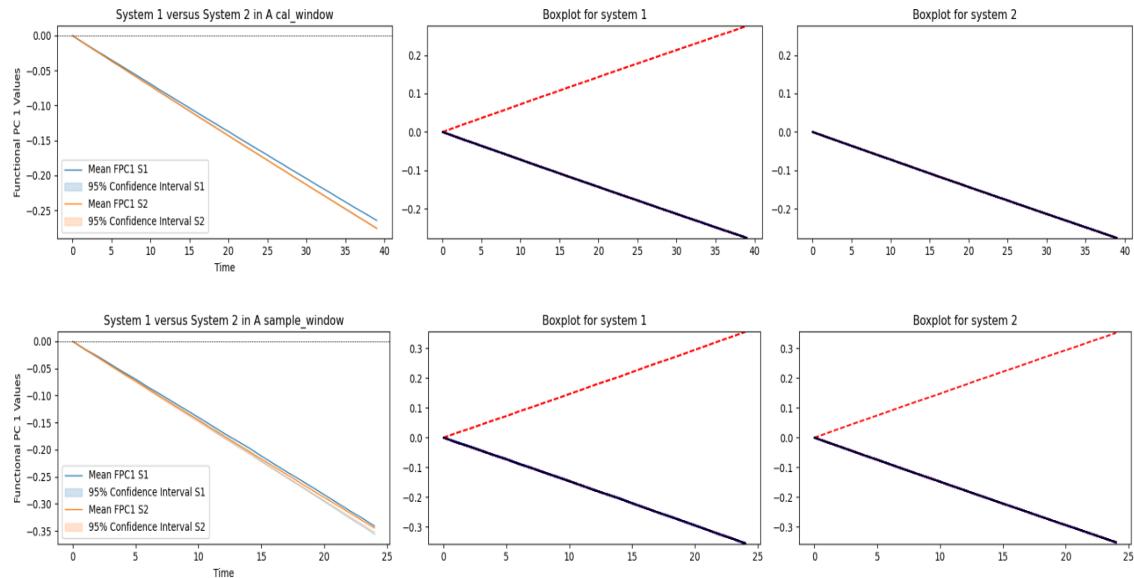


**Figure A21.** Scores of the first two components are binned by Card Age for system 1 (top) and system 2 (bottom) using sensor B in the (a) Calibration window and (b) Sample window.

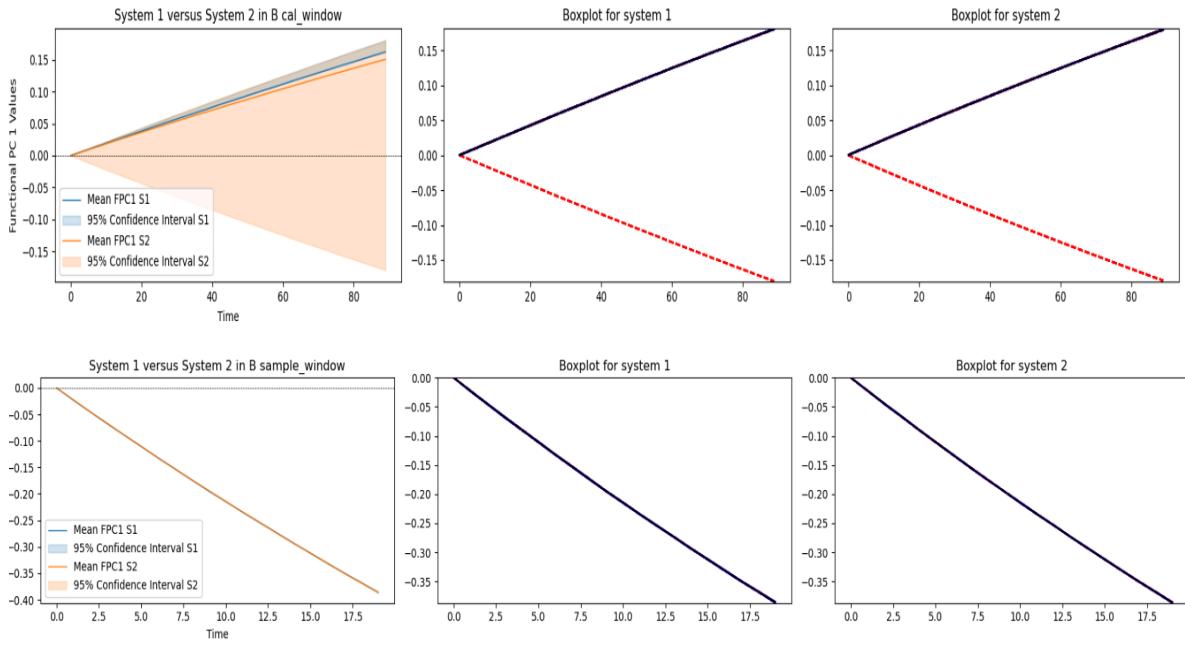
#### 6.1.2.4 Confidence interval and boxplots for the first component



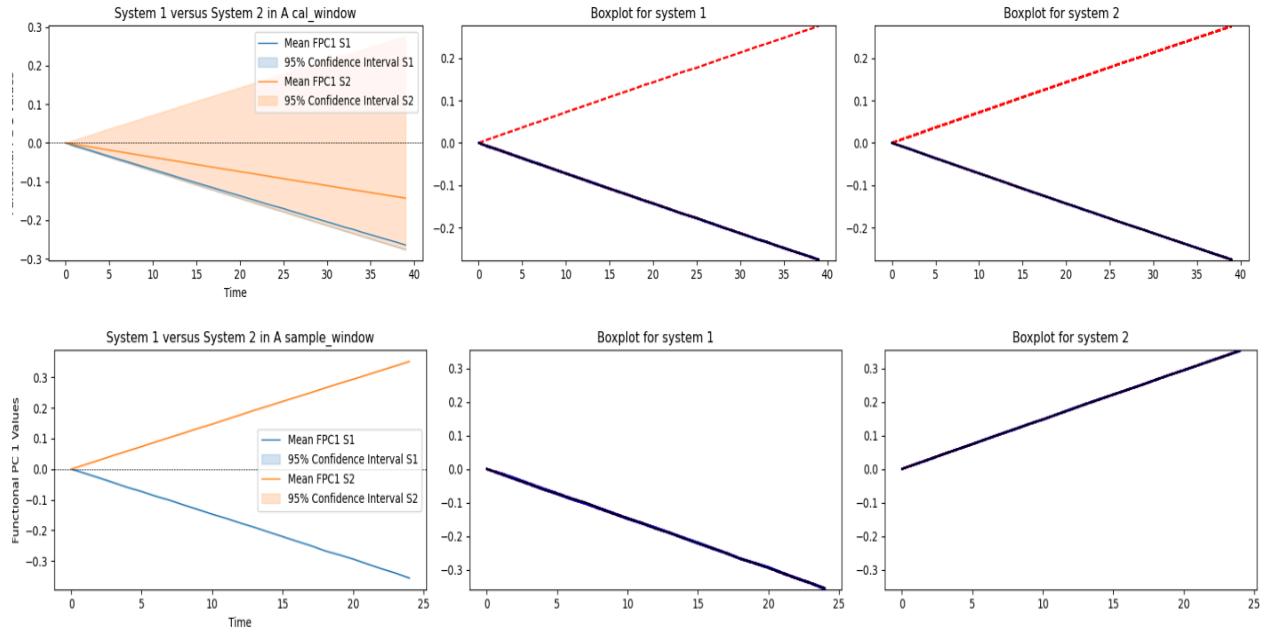
**Figure A22. Confidence interval and boxplots for the first component in the sensor A (top)  
Calibration Window (bottom) Sample window, focusing on Fluid Temperature.**



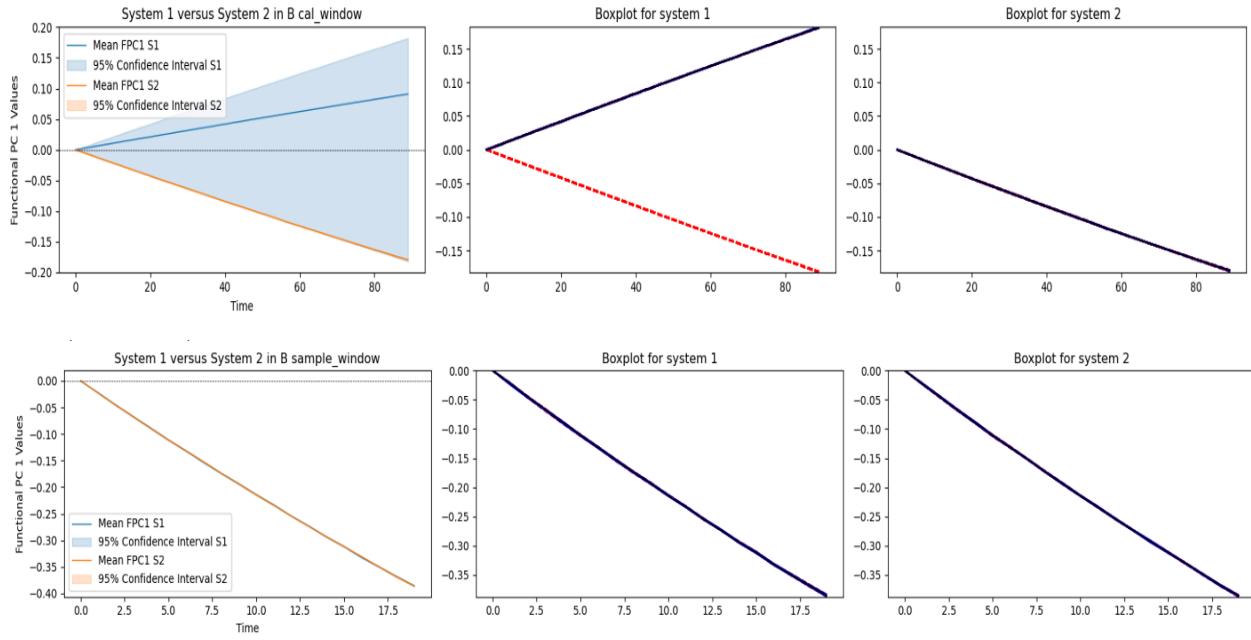
**Figure A23. Confidence interval and boxplots for the first component in the sensor A (top)  
Calibration Window (bottom) Sample window, focusing on Card Age.**



**Figure A24. Confidence interval and boxplots for the first component in the sensor B (top)**  
**Calibration Window (bottom) Sample window, focusing on Card Age.**

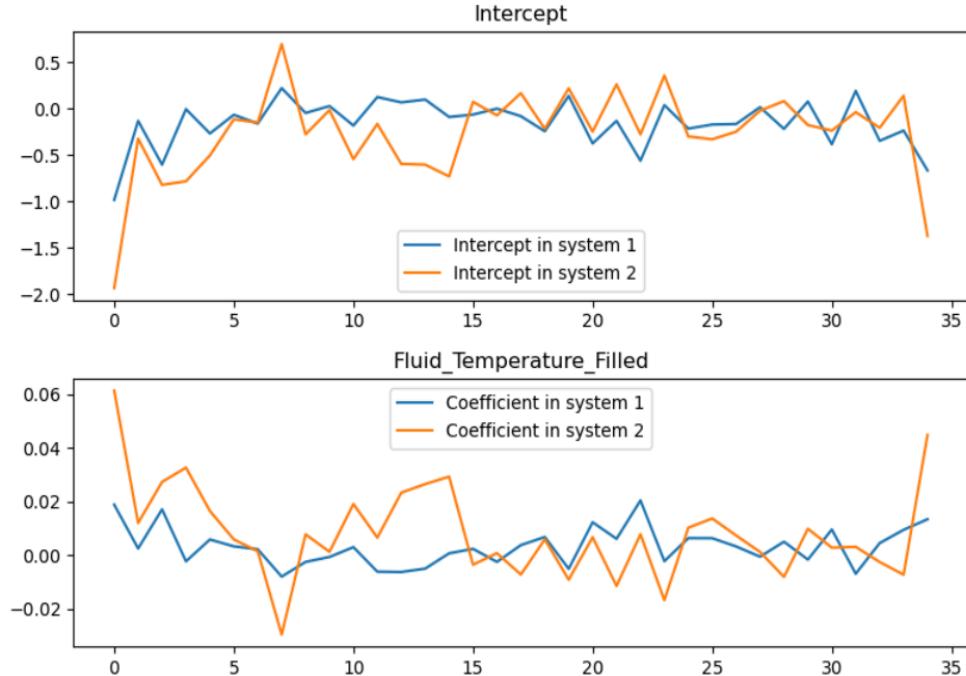


**Figure A25. Confidence interval and boxplots for the first component in the sensor A (top)**  
**Calibration Window (bottom) Sample window, focusing on Fluid Type.**

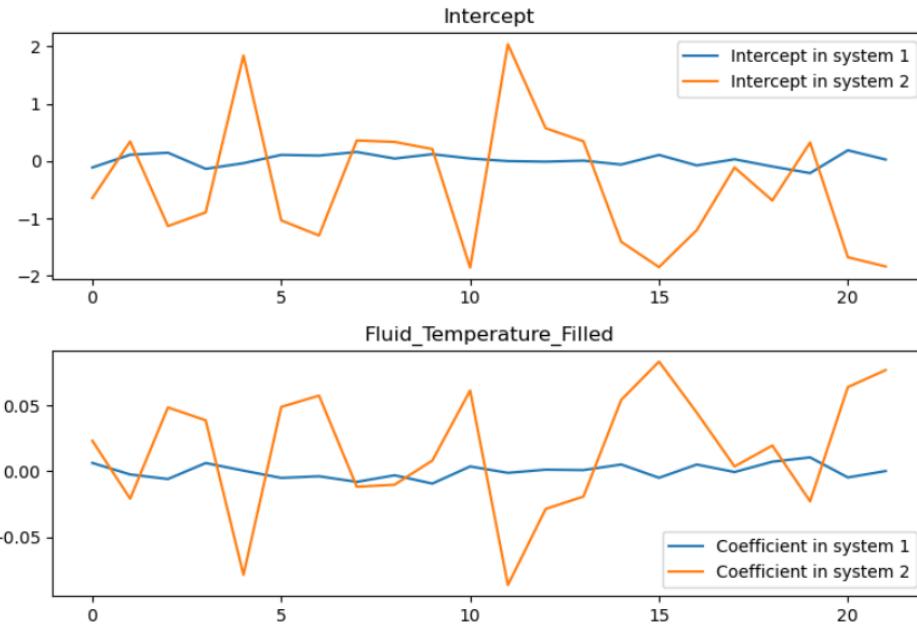


**Figure A26. Confidence interval and boxplots for the first component in the sensor B (top)  
Calibration Window (bottom) Sample window, focusing on Fluid Type.**

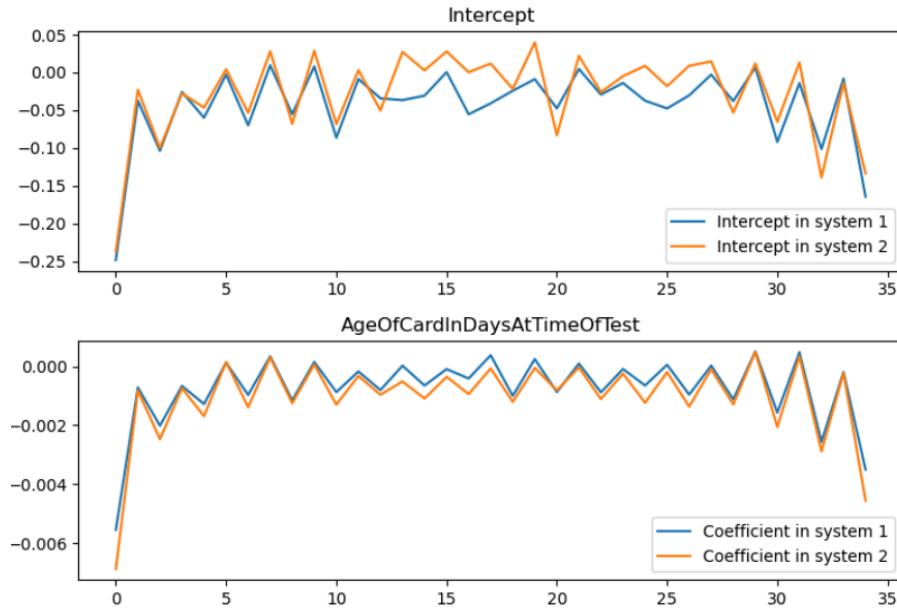
#### 6.1.2.5 The coefficients come from Functional Regression



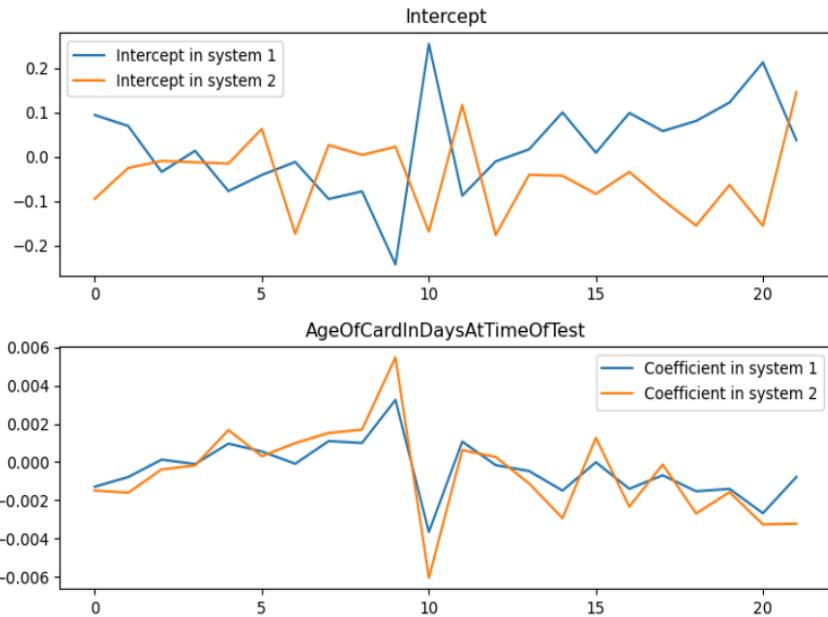
**Figure A27. The coefficients from Functional Regression in the Cal window in sensor A, focusing on Fluid Temperature.** Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 36 in a sequence of 40 total.



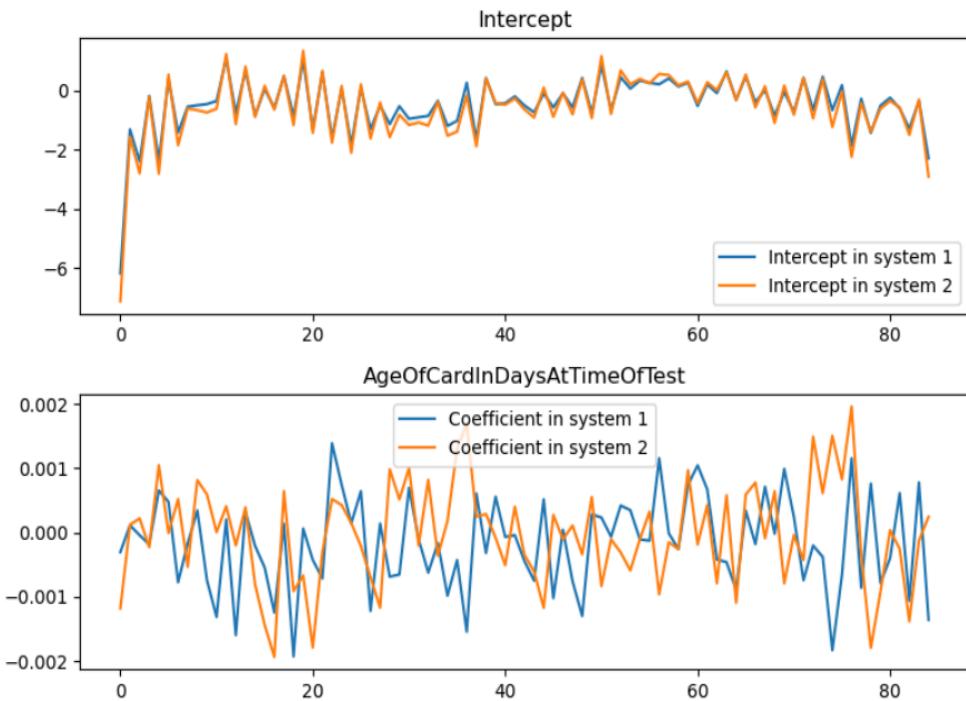
**Figure A28.** The coefficients from **Functional Regression in the Sample window in sensor A, focusing on Fluid Temperature**. Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 23 in a sequence of 25 in total.



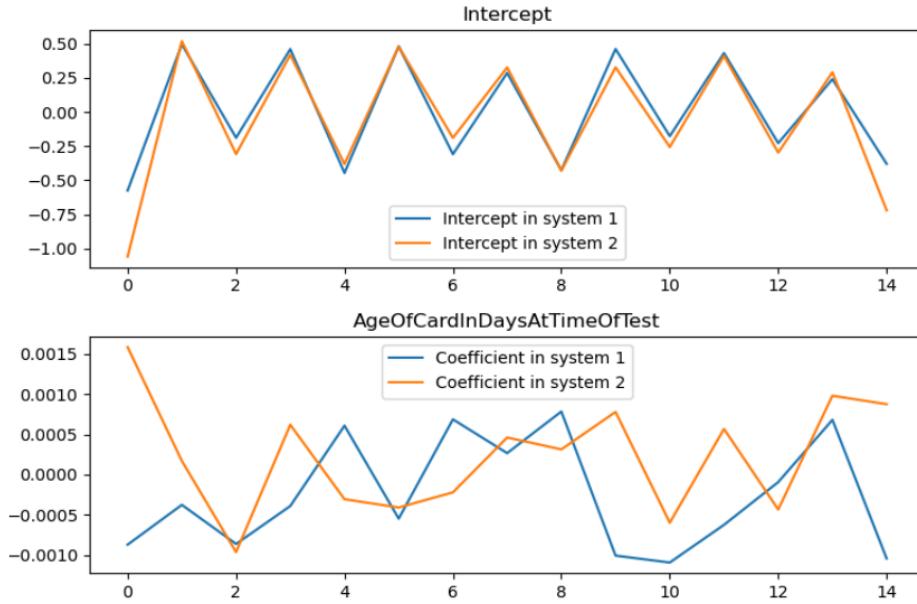
**Figure A29.** The coefficients from **Functional Regression in the Cal window in sensor A, focusing on Card Age**. Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 36 in a sequence of 40 total.



**Figure A30.** The coefficients from Functional Regression in the Sample window in sensor A, focusing on Card Age. Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 23 in a sequence of 25 total.



**Figure A31.** The coefficients from Functional Regression in the Cal window in sensor B, focusing on Card Age. Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 86 in a sequence of 90 in total.



**Figure A32. The coefficients from Functional Regression in the Sample window in sensor B, focusing on Card Age.** Based on the significantly different magnitudes of the coefficients, the plotted timestamps are 16 in a sequence of 20 in total.

## References

- [1] U. B. Trivedi et al., "Potentiometric biosensor for urea determination in milk," *Sensors and Actuators B: Chemical*, vol. 140, no. 1, pp. 260-266, 2009.
- [2] Epochal Inc., "System Manual with Epoch NXS Host," Siemens Healthcare Diagnostics Inc., Ottawa, ON, Canada, 2023. [Online]. Available: <https://siemens-healthineers.com/epoch>
- [3] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, 2nd ed., Springer Series in Statistics. Springer, 2005.
- [4] S. C. Ribeiro et al., "Potentiometric biosensor Based on artificial antibodies for an Alzheimer biomarker detection," *Applied Sciences*, vol. 12, no. 7, art. 3625, 2022.
- [5] C. Areshenkov, D. J. Gale, D. Standage, J. Y. Nashed, J. R. Flanagan, and J. P. Gallivan, "Neural excursions from manifold structure explain patterns of learning during human sensorimotor adaptation," *eLife*, vol. 11, art. e74591, 2022. [Online]. Available: <https://doi.org/10.7554/eLife.74591>
- [6] R. J. B. Zhu and X. X. Wei, "Unsupervised approach to decomposing neural tuning variability," *Nature Communications*, vol. 14, art. 2298, 2023. [Online]. Available: <https://doi.org/10.1038/s41467-023-37982-z>
- [7] "Functional principal component analysis," *Wikipedia, The Free Encyclopedia*, [https://en.wikipedia.org/wiki/Functional\\_principal\\_component\\_analysis](https://en.wikipedia.org/wiki/Functional_principal_component_analysis) [Accessed May 19, 2024].
- [8] J. S. Morris, "Functional regression," *Annual Review of Statistics and Its Application*, vol. 2, pp. 321-359, 2015.
- [9] J. L. Wang, J. M. Chiou, and H. G. Müller, "Functional data analysis," *Annual Review of Statistics and Its Application*, vol. 3, pp. 257-295, 2016.
- [10] GAA-UAM, "skfda/preprocessing/dim\_reduction/\_fPCA.py," [Online]. Available: [https://github.com/GAA-UAM/scikit-fda/blob/develop/skfda/preprocessing/dim\\_reduction/\\_fPCA.py](https://github.com/GAA-UAM/scikit-fda/blob/develop/skfda/preprocessing/dim_reduction/_fPCA.py). [Accessed: June 1, 2024].
- [11] "Functional Principal Component Analysis (FPCA)," *scikit-fda documentation*, [https://fda.readthedocs.io/en/latest/modules/preprocessing/autosummary/skfda.preprocessing.dim\\_reduction.FPCA.html](https://fda.readthedocs.io/en/latest/modules/preprocessing/autosummary/skfda.preprocessing.dim_reduction.FPCA.html) [Accessed May 19, 2024].
- [12] "LinearRegression," *scikit-learn documentation*, [https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LinearRegression.html#linearregression](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html#linearregression) [Accessed June 6, 2024].

- [13] C. Ramos-Carreño, J. L. Torrecilla, M. Carbajo-Berrocal, et al., "scikit-fda," *arXiv preprint*, arXiv:2211.02566, 2022.
- [14] Y. Sun and M. G. Genton, "Functional boxplots," *Journal of Computational and Graphical Statistics*, vol. 20, no. 2, pp. 316-334, 2011.
- [15] S. Nagy, I. Gijbels, M. Omelka, et al., "Integrated depth for functional data: statistical properties and consistency," *ESAIM: Probability and Statistics*, vol. 20, pp. 95-130, 2016.