Mass Spectrometer Interface

A Desktop Application for Reading Instrument Data

Cousins Photosynthesis Lab in the School of Biological Sciences at WSU



**Team Linnaea Borealis**

****

Kyler Kupp, Erik Holtrop

**TABLE OF CONTENTS**

[I. Introduction 2](#_Toc167133396)

[II. System Requirements Specification 3](#_Toc368571065)

[II.1. Functional Requirements 3](#_Toc1678139455)

[II.1.1 Calculating Concentrations (Module 1) 3](#_Toc919943970)

[II.1.2 Analyzing Enzyme Activity (Module 3) 4](#_Toc1262166227)

[II.1.3 Data Conversion (Module 4) 4](#_Toc206537085)

[II.1.4 Analyzing Exchanges in Leaves (Module 5) 4](#_Toc1062457675)

[II.2. Non-Functional Requirements 5](#_Toc221716674)

[II.3 Use Cases 6](#_Toc1391703108)

[II.4 User Stories 10](#_Toc1409903311)

[II.5 Tracibility Matrix 12](#_Toc625041687)

[III. System Evolution 12](#_Toc1001219433)

[IV. Glossary 13](#_Toc1669912293)

[V. References 13](#_Toc1176085628)

# Introduction

When plants breathe, they take carbon dioxide (CO2) out of the air and replace it with oxygen (O2). Determining what affects plants’ respiration rate, or their breathing rate, is incredibly valuable data. These factors point backwards in time, reflecting causes for evolutionary trends, and forwards in time, providing opportunities to improve agriculture. We can use a mass spectrometer to measure this breathing rate. The mass spectrometer measures the amount of CO2 and O2 in the air, and so the respiration rate is the change in each of those.

The Cousins Photosynthesis Lab in the School of Biological Sciences at Washington State University uses one of these mass spectrometers. These instruments are complicated devices, requiring complex calculations for calibration. The lab uses proprietary software from the mass spectrometer’s manufacturer, but that software outputs massive amounts of data over the course of a multi-hour lab, most of which isn’t needed. This problem has been partially solved with the creation of a Python desktop application, but this application is not perfect. Our task is to improve this application. This application currently faces small bugs, and only works for one instrument. The application is also in process of a UI upgrade. There’s also a few non-spectrometer instruments in the lab that are provide similar data, that would be easier to use if their data-streams were combined.

# System Requirements Specification

This section outlines the key functional and non-functional requirements, use cases, user stories, and the traceability matrix for the Mass Spectrometer Interface system. The system’s requirements aim to improve the accuracy of calculations, resolve existing bugs, and introduce new data analysis features. Through these improvements, the system will offer more streamlined data analysis and enhance the lab's research capabilities.

## Functional Requirements

The functional requirements outlined in this section detail the essential and desirable capabilities that the system must provide to meet the client's needs. Each requirement specifies a key functionality, ranging from data calculations and visualization to data manipulation and analysis. These functionalities are organized by module, with each module representing a standalone program. Module 2 has been omitted, as it is already fully functional in the existing solution.

### II.1.1 Calculating Concentrations (Module 1)

|  |  |
| --- | --- |
| **Functional**  **Requirement** | **[FR-1] Calculate Atom Ratio** |
| Description | The system must allow the user to calculate the ratio of bicarbonate to carbon dioxide. The result of the calculation must display the ratio in decimal format. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-2] Center Mean Bars** |
| Description | The system must visually present mean bars in the center of the graph window. These vertical bars must be movable by the user so that a segment of data can be selected to calculate the mean for. |
| Source | Client |
| Priority | Level 1 (Desirable) |

### II.1.2 Analyzing Enzyme Activity (Module 3)

|  |  |
| --- | --- |
| **Functional**  **Requirement** | **[FR-3] Plot Atom Percentage** |
| Description | The system must be able to plot the percentage of carbon dioxide with a mass of 49 out of all present molecules. The graph must be normalized using natural log. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-4] Plot Derivative** |
| Description | The system must be able to plot the derivative of the natural log of the percentage of carbon dioxide with mass 49. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-5] Copy Table** |
| Description | The system must allow the user to copy the contents of the mean value table. |
| Source | Client |
| Priority | Level 1 (Desirable) |
| **Functional**  **Requirement** | **[FR-6] Plot Datasets Consecutively** |
| Description | The system must be able plot data from one acquisition folder, stop plotting, and then allow the user to select a new acquisition folder to begin plotting from. |
| Source | Client |
| Priority | Level 1 (Desirable) |
| **Functional**  **Requirement** | **[FR-7] Record Sample Name** |
| Description | The system must take user input and record the name of the sample to the mean table in the same row as its corresponding mean value. |
| Source | Client |
| Priority | Level 1 (Desirable) |

### II.1.3 Data Conversion (Module 4)

|  |  |
| --- | --- |
| **Functional**  **Requirement** | **[FR-8] Accept Data** |
| Description | The system must accept data from the EZ-Tap listener device attached to the secondary mass spectrometer through the USB port. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-8] Identify Port** |
| Description | The system must identify which port the EZ-Tap is plugged into and intelligently select it for listening. |
| Source | Client |
| Priority | Level 1 (Desirable) |
| **Functional**  **Requirement** | **[FR-15] Format Data** |
| Description | The system must convert data from the secondary mass spectrometer into a specific format so that it can be analyzed using existing applications. The data must be converted from hexadecimal to decimal format and be resorted into the appropriate column order. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-16] Split Data** |
| Description | The system must split the data into many different csv files to match the data file structure used by the primary mass spectrometer. |
| Source | Client |
| Priority | Level 0 (Essential) |

### II.1.4 Analyzing Exchanges in Leaves (Module 5)

|  |  |
| --- | --- |
| **Functional**  **Requirement** | **[FR-9] Select Data Streams** |
| Description | The system must be able to select file sources to pull data from for three different instruments. These instruments include a LI-COR Leaf-gas Exchange System, a tunable diode laser (TDL), and a Picarro that measures different oxygen isotopes. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-10] Plot Data Streams** |
| Description | The system must plot each of the three data streams on one graph. The graph must be able to plot data dynamically in near-real time. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-11] Manipulate Scale** |
| Description | Users must be able to manipulate the scale of each graph. The user should be able to edit the lower and upper bound of both the horizontal and vertical axis. |
| Source | Client |
| Priority | Level 1 (Desirable) |
| **Functional**  **Requirement** | **[FR-12] Change Plotting Speed** |
| Description | The user must be able to change the speed at which the data is plotted as well as pause the plotting of data. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-13] Mean Bars** |
| Description | Two vertical bars must be included in the graph of the application. The user should be able to move the bars to encapsulate a desired segment of data from which the mean can be calculated. |
| Source | Client |
| Priority | Level 0 (Essential) |
| **Functional**  **Requirement** | **[FR-14] Calculate Mean** |
| Description | The application must provide the ability to calculate the mean value between the two mean bars. |
| Source | Client |
| Priority | Level 0 (Essential) |

## Non-Functional Requirements

Our project exists in the context of academic plant biology research, and thus has several ideals and values it strives to represent and uphold. These ideals don’t necessarily describe what the program *does*, they rather describe how it *is*. This section clarifies the non-functional requirements which we’ll use to guide our design and development to uphold these ideals and values.

**Modern Visuals:**

The system shall look like a modern 2020’s program. PyQt5’s baseline UI elements meets this standard

**Python For Extendibility:**

The system shall use Python since it is one of the most common languages among Biology students, including those that use the Cousins Photosynthesis Lab. This will enable extension and maintenance.

**Documentation For Extendibility:**

The system shall have documentation for code that’s thorough enough to allow college senior skill level software engineers to extend and maintain it.

**Documentation For Non-Technical Users:**

The system shall have documentation to enable non-technical users to make full use of it. For example, an explanation of how to create executables of each module.

**Accuracy:**

The system shall accurately calculate and portray data. The instruments have their own levels of accuracy, so maintenance of significant figures should be enough.

**Compatibility:**

The system shall use the established data format, which is versatile enough to work with other devices.

**Non-Destructive:**

The system shall not overwrite or otherwise destroy any data.

## II.3 Use Cases

The use cases below describe common scenarios of user interactions with the Mass Spectrometer Interface system, illustrating how various functional requirements are applied in specific situations. These use cases provide a clear understanding of how users will engage with the system's features to achieve their goals. The proposed use cases are visually represented in the use case diagram shown in Figure 1.

  
Figure 1: Use Case Diagram

**Use Case 1: Calculate Bicarbonate/CO2 Ratio**

|  |  |
| --- | --- |
| Use Case | Calculate Bicarbonate/CO2 Ratio |
| Actors: | Lab researcher |
| Pre-condition: | File path to data folder has been selected, and bicarbonate and CO2 have both been calibrated. |
| Post-condition | Bicarbonate/CO2 Ratio is displayed in decimal format. |
| Main Flow | * User selects data acquisition folder. * Starts plotting data. * Selects data segment. * Calibrates bicarbonate and CO2. * Selects calculate Bicarbonate/CO2 button. * Bicarbonate/CO2 ratio is displayed. |
| Alternative Flow | * If bicarbonate or CO2 have not been calibrated, display an error message and retry. |
| Related Requirements | FR-1: Calculate Atom Ratio  FR-2: Center Mean Bars |

**Use Case 2: Plot Derivatives**

|  |  |
| --- | --- |
| Use Case | Plot Derivatives |
| Actors: | Lab researcher |
| Pre-condition: | File path to data folder has been selected. |
| Post-condition | The Atom Percent and its derivative are fully plotted on their respective graphs. Alert user data has run out. |
| Main Flow | * User selects data acquisition folder. * Starts plotting data. * Data is plotted to each graph in real-time. |
| Alternative Flow | * Invalid data in acquisition folder. * User is prompted to select new file path. |
| Related Requirements | FR-3: Plot Atom Percentage  FR-4: Plot Derivative |

**Use Case 3: Convert Data**

|  |  |
| --- | --- |
| Use Case | Convert Data |
| Actors: | Lab researcher |
| Pre-condition: | File path to data folder has been selected. File path to output folder has also been selected. |
| Post-condition | The output folder has been populated with formatted data and the user has been notified by popup that conversion has finished. |
| Main Flow | * User selects data acquisition folder. * Selects output folder. * Starts conversion. * Output file is populated with formatted data. * User is notified that conversion has finished. |
| Alternative Flow | * Invalid data is encountered. * Error message is displayed. * User may select a new input file. |
| Related Requirements | FR-8: Format Data |

**Use Case 4: Select Input Files**

|  |  |
| --- | --- |
| Use Case | Select Input Files |
| Actors: | Lab researcher |
| Pre-condition: | User has opened program. |
| Post-condition | File path has been set. Confirmation message is displayed. |
| Main Flow | * User selects data acquisition folder. * Confirmation message is displayed. |
| Alternative Flow | * Invalid input file is encountered. * Error message is displayed. * User may select a new input file. |
| Related Requirements | FR-9: Select Data Streams |

**Use Case 5: View Graph**

|  |  |
| --- | --- |
| Use Case | View Graph |
| Actors: | Lab researcher |
| Pre-condition: | Input file has been selected. Plotting has been started. |
| Post-condition | Graphs are filled by plotted data. User is notified that end of data source has been reached. |
| Main Flow | * User selects data acquisition folder. * Starts plotting * Graphs are populated * User is notified that end of data source has been reached. |
| Alternative Flow | * Invalid input file is encountered. * Error message is displayed. * User may try plotting again. |
| Related Requirements | FR-10: Plot Data Streams |

**Use Case 6: Manipulate Graph Scale**

|  |  |
| --- | --- |
| Use Case | Manipulate Graph Scale |
| Actors: | Lab researcher |
| Pre-condition: | Program is open |
| Post-condition | Graphs are resized to desired scale. |
| Main Flow | * User hovers over graph * Scrolls up or down to zoom in or out * Graph scales appropriately |
| Alternative Flow | * Maximum or minimum scale is reached * Zooming is limited by maximum size. |
| Related Requirements | FR-11: Manipulate Scale |

**Use Case 7: Change Plotting Speed**

|  |  |
| --- | --- |
| Use Case | Change Plotting Speed |
| Actors: | Lab researcher |
| Pre-condition: | Input file has been selected and data stream has not run out |
| Post-condition | Speed at which data is plotted is changed. |
| Main Flow | * User selects input folder. * Begins plotting. * Sets plotting speed. * Speed at which data is plotted to graph changes appropriately. |
| Alternative Flow | * Data stream runs out. * User is notified and must start over plotting. |
| Related Requirements | FR-12: Change Plotting Speed |

**Use Case 8: Select Data Points**

|  |  |
| --- | --- |
| Use Case | Select Data Points |
| Actors: | Lab researcher |
| Pre-condition: | Input file has been selected and data has been plotted to graph. |
| Post-condition | Vertical bars display the bounds of selected data. |
| Main Flow | * User selects input folder. * Begins plotting. * Selects Mean Bar option * Adjusts Mean Bars’ location |
| Alternative Flow | * If data has not been plotted, the Mean Bars will still appear but contain no data yet. |
| Related Requirements | FR-13: Mean Bars |

**Use Case 9: Calculate Mean**

|  |  |
| --- | --- |
| Use Case | Calculate Mean |
| Actors: | Lab researcher |
| Pre-condition: | Input file has been selected and data has been plotted to graph. Mean bars have been selected and adjusted to the desired location. |
| Post-condition | Mean value within the selected data segment is displayed. |
| Main Flow | * User selects input folder. * Begins plotting. * Selects Mean Bar option. * Adjusts Mean Bars’ location. * Selects calculate mean. * Mean value is displayed. |
| Alternative Flow | * If empty data segment is selected, user is prompted to select a new segment. |
| Related Requirements | FR-14: Calculate Mean |

## II.4 User Stories

The following user stories outline key tasks that users can perform within the system, highlighting their goals and the reasons behind them. Each story provides a clear and concise description of what the user aims to achieve, along with the expected system behavior, ensuring that user needs are directly addressed and aligned with the system's functionality.

**User Story US1: Calculate Bicarbonate/CO2 Ratio**

As a Lab Researcher, I need to calculate the ratio of bicarbonate to carbon dioxide so that I can log the given output.

Feature: Calculate Bicarbonate/CO2 Ratio

Scenario: Lab researcher calculates bicarbonate/CO2 ratio

Given the mass spectrometer data file has data for HCO-3 and CO2  
AND CO2 and HCO-3 have been calibrated using the plotted data  
When I click the button "BiCarb/CO2"  
Then the correct ratio of BiCarb/CO2 will be computed and displayed.

**User Story US2: Plot Derivatives**

As a Lab Researcher, I need the first and second derivatives of percent CO2 49 So that I can analyze and log it.

Feature: Plot Derivatives

Scenario: Lab researcher plots Atom Percent and its derivative for CO2 data

Given the mass spectrometer data file has data on the CO2 49, 47 & 45 masses.  
When I run module 3 (using the "start" or "unpause" button)  
Then two graphs will show the first and second derivatives of percent CO2 49 respectively.

**User Story US3: Convert Data**

As a Lab Researcher, I need modules 1-3 to work for the second mass spectrometer so that I can analyze its data the same way.

Feature: Convert Data

Scenario: Lab researcher converts data to be compatible with modules 1-3

Given Instrument B (the second mass spectrometer) is outputting data, or has outputted data  
WhenI select the Instrument B option on module 1/2/3  
And I select the Instrument B datastream (likely a directory)  
Then module 1/2/3 functions as normal, including all use cases for module 1/2/3.

**User Story US4: Datastream Combining**

As a Data Researcher, I need data from the LI-COR Leaf-gas Exchange System, Tunable Diode Laser, and Picarro consolidated into one or more spreadsheets so that I can analyze the data more efficiently.

Feature: Datastream Combining

Scenario: Lab researcher combines three streams of data

Given I have data from all three instruments  
When I run module 5  
Then one or more spreadsheets collectively containing all the data, collated by time, is created.

**User Story US5: Data Isolation**

As a Data Researcher, I need to isolate portions of data so that I can view and analyze the most important parts of a multi-hour lab experimentation session.

Feature: Data Isolation

Scenario: Lab researcher isolates data

Given the System has collected a non-trivial amount of data from the mass spectrometer  
When I drag the left and right edges of my selection to a portion of the data graph  
Then only data from that portion of the graph will be analyzed in the calculation dashboard.

## II.5 Tracibility Matrix

The table below links functional requirements to their corresponding use cases and user stories. This mapping ensures that every requirement is fully addressed and connected to relevant user interactions, providing a clear line of traceability throughout the system.

|  |  |  |  |
| --- | --- | --- | --- |
| Functional Requirement | Use Case | User Story | Priority |
| [FR-1] Calculate Atom Ratio | UC-1: Calculate Bicarbonate/CO2 Ratio | US5: Data Isolation | Level 0 |
| [FR-2] Center Mean Bars | UC-1: Calculate Bicarbonate/CO2 Ratio | US1: Calculate Bicarbonate/CO2 Ratio | Level 1 |
| [FR-3] Plot Atom Percentage | UC-2: Plot Derivatives | US2: Plot Derivatives | Level 0 |
| [FR-4] Plot Derivative | UC-2: Plot Derivatives | US2: Plot Derivatives | Level 0 |
| [FR-8] Format Data | UC-3: Convert Data | US3: Convert Data | Level 0 |
| [FR-9] Select Data Streams | UC-4: Select Input Files | US4: Datastream Combining | Level 0 |
| [FR-10] Plot Data Streams | UC-5: View Graph | US4: Datastream Combining | Level 0 |
| [FR-11] Manipulate Scale | UC-6: Manipulate Graph Scale | US4: Datastream Combining | Level 1 |
| [FR-12] Change Plotting Speed | UC-7: Change Plotting Speed | US4: Datastream Combining | Level 0 |
| [FR-13] Mean Bars | UC-8: Select Data Points | US5: Data Isolation | Level 0 |
| [FR-14] Calculate Mean | UC-9: Calculate Mean | US4: Datastream Combining | Level 0 |

# System Evolution

The Cousins Lab’s relationship with WSU places it in a position to enlist Computer Science students to work on this codebase each year as part of their capstone. We see this in the project’s history, having two teams as previous maintainers. The sponsor liaison, Dr. Cousins, has voiced an intention to have software created for many of the instruments in the lab. With these considerations, we should make our software compatible with different machines if possible, and make it maintainable and extendable by software engineers with the skill and education level of a college senior.

We can see what this looks like from the issues and opportunities presented by previous teams. For instance, the modules written from the lab’s primary Mass Spectrometer, may be usable for the second one, provided we write software to reformat its data stream. In another case, some of the existing code lacks basic files to enable maintenance, like a requirements.txt for relevant Python libraries or the context files for executable creation.

# Glossary

**Mass spectrometer:** an apparatus for separating isotopes, molecules, and molecular fragments according to mass. The sample is vaporized and ionized, and the ions are accelerated in an electric field and deflected by a magnetic field into a curved trajectory that gives a distinctive mass spectrum.

**UI**: User Interface. The means by which the user and application interact.

# References