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

# 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.

# Background and Related Work

To assess the current leaders in the same field as our project, we must first clearly define that field. For the purposes of this document, we have refined the definition to "data analysis tools for mass spectrometry." Based on this scope, our research identifies two distinct and exemplary products that are highly successful within this domain.

The first state-of-the-art leader in the domain is Thermo Fisher Scientific with their proprietary mass spectrometry software, GC-MS. This program interfaces directly with the mass spectrometer to collect and visualize data. GC-MS provides users with a detailed view of incoming high-frequency data. However, it lacks functionality for analyzing smaller segments of data. Our project addresses this gap by enabling users to select small partitions of data in which they can analyze. This allows users to accurately calibrate gases such as oxygen, carbon dioxide in buffer, and carbon dioxide in hydrogen chloride. The Cousins Photosynthesis Lab in the School of Biological Sciences at Washington State currently uses GC-MS for data collection. This data is then exported and imported into a python desktop application that our project focuses on improving and extending. The current python application can only import data that is formatted in a particular way. Our goal is to extend the accepted file formats to be compatible with other mass spectrometers.

The second notable contemporary in the field of mass spectrometry is OpenMS (2007), a flexible open-source software platform used for data analysis. OpenMS provides a set of 185 tools and workflows for mass spectrometric data processing. These tools can be accessed by integrating the OpenMS library. OpenMS provides a comprehensive suite of mass spectrometry tools specifically for liquid chromatography mass spectrometry (LC-MS). However, it does not support gas chromatography mass spectrometry (GC-MS). Our project focuses specifically on GC-MS.

In Summary, there are several notable existing analysis tools that fall within the definition of our project domain. However, the tools’ functionality does not allow fined-tuned analysis of GC-MS. Our project aims to match the detailed analysis OpenMS provides for LC-MS, but for GC-MS. It will use data from Thermo Fisher Scientific’s GC-MS software but allow calibrations and computations to be performed on subsections of data. The project should also import data from multiple mass spectrometer devices.

In order to complete this project, we will need to learn a few new technical skills. The primary language that will be utilized in this project is Python. The team already has experience using Python, but user interface will require the Python binding PyQt5. The mastery of this GUI toolkit will be necessary for providing an easy-to-use interface for non-technical users. Familiarity with the NumPy framework will also be paramount for managing multi-dimensional arrays of concentration data. Mastering these skills will be necessary to produce a satisfactory product.

# Project Overview

Describe your project problem in detail and summarize the project objectives.

Provide a detailed description of the goals and desired outcomes. You should discuss the intended outcomes of the project with your mentor and summarize them here.

This section should be detailed enough to show that your team has a clear understanding of the project objectives and outcomes. Please plan on having 1+ page text on “Project Overview”.

# Client and Stakeholder Identification and Preferences

Identify your clients and stakeholders (In CptS 421/423 your industry sponsor and mentor is your primary client, however many stakeholders also exist beyond the sponsoring company, mentor and co-mentor. A stakeholder is a person or role that is affected by the system that you will develop in some way, so even a description of any expected future users should be included.)

Briefly describe the needs and preferences of your clients and stakeholders.

Our client is the Cousins Photosynthesis Lab. This is a lab maintained by Washington State University, primarily through Dr. Asaph Cousins, a professor at WSU’s School of Biological Sciences. In this regard, the university is our client, and Dr. Cousins is their liaison. WSU’s primary interest is efficient and impactful research.

The most prominent class of stakeholders are the users of this lab. This includes Dr. Cousins, but also graduate students and postdoctoral researchers both currently and in the future. These researchers are our primary user class. They need our product to be usable, extendable, and powerful. Their experience with coding is variable but generally limited, and this has presented conflicts with previous iterations of this product. These clients could potentially benefit greatly from instructive usage manuals.

Another class of stakeholders is the plant biology academic community. The data processed by our product will likely directly or indirectly affect research created at the Photosynthesis Lab. This makes it incredibly important to ensure that our product preserves the accuracy of the data outputted by the lab’s mass spectrometers.

# Glossary

**WSU**: Washington State University

**GC-MS**: Gas Chromatography Mass Spectrometry

**LC-MS**: Liquid Chromatography Mass Spectrometry

# References

Cite your references here. Please follow the IEEE citation standard for your references. (http://www.ieee.org/documents/ieeecitationref.pdf)

For the papers you cite give the authors, the title of the article, the journal name, journal volume number, date of publication and inclusive page numbers. Giving only the URL for the journal is not appropriate. For the websites, give the title, author (if applicable) and the website URL.

[1]“Mass Spectrometry Software | Thermo Fisher Scientific - UK,” *Thermofisher.com*, 2024. <https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/mass-spectrometry-software.html> (accessed Sep. 23, 2024).‌

[1]H. L. Röst *et al.*, “OpenMS: a flexible open-source software platform for mass spectrometry data analysis,” *Nature Methods*, vol. 13, no. 9, pp. 741–748, Sep. 2016, doi: <https://doi.org/10.1038/nmeth.3959>.

‌