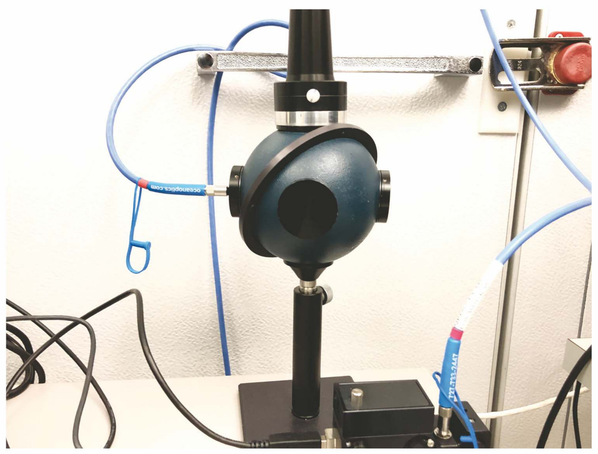
Analysis and Validation of MicroPEM Data Files and Filter Optical Density Measurements

Mukil Guruparan(1); Jon Thornburg, PhD(2); Ryan Chartier, MS(2); Michelle McCombs(2)
(1)UNC-Chapel Hill, Chapel Hill, NC; (2)RTI International, Research Triangle Park, NC

Introduction

The MicroPEM is a personal exposure monitor with a filter that collects particulate matter (PM) from the air around the person who wears it. It is used worldwide to study and understand various environments and how they affect people's health. The MicroPEM has two important components. First is the MicroPEM data file, which contains real-time PM concentration, temperature reading, humidity, and other instrument performance parameters. These parameters must be manually analyzed using a complex graph. This process is not ideal because it wastes time and requires a researcher experienced with the data files. Second is the filter, which undergoes optical density (OD) analysis to determine the PM chemical composition.

The OD system (shown below) works by transmitting high-powered light through the filter and detecting the absorbance at specific wavelengths using a highly sensitive spectrometer. The high sensitivity gives a low minimum detection limit and a high maximum detection but increases the number of erroneous readings. Thus, researchers needed a method to automatically detect these incorrect measurements and report them to the user.



To develop this method, I created programs in Python that facilitate interaction with the data from both processes, especially for our colleagues who work in other countries. The MicroPEM program creates text files that summarize and validate data files from the MicroPEMs, and the OD program validates OD data and alerts the user if the data looks abnormal, and if a certain filter requires additional measurements.

Objectives of Programs

MicroPEM

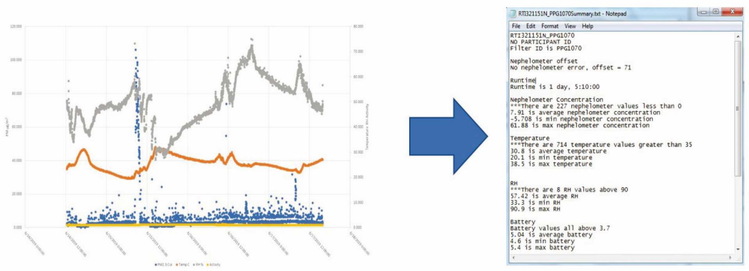
- Summarize each parameter of the MicroPEM data file, including minimum, maximum, and average
- Validate each parameter of the program, and alert the user if any parameters need to be reported as problematic
- Calculate runtime, which can be challenging due to lack of standardization between files
- Ensure the program is easy to use, for those in different countries and those who don't have experience with Python

OD

- Validate PreOD and PostOD, the two measurements that are made on each filter, independently
- Alert the user if any filters need to be remeasured as the user is taking measurements
- Find PreOD and PostOD locations in each file, which has been challenging due to lack of standardization

Results

The MicroPEM program converts the data that were previously displayed in graphs (shown below on the left) into a text file (shown below on the right). This text file summarizes the data and includes flags for the user if any values have to be further evaluated. The program accurately finds runtime of any MicroPEM, using the date/time module in Python, and it is compatible with different file types to avoid any issues that may arise. Python's win32gui module (1), which creates an interface allowing the user to select the appropriate folder, makes this program easy to use.



The OD program works similarly but validates the data as the user is making measurements. It successfully validates PreOD and PostOD data, comparing the two for accuracy. It also finds exactly where the two different datasets are located in the Excel file by looping through all of the columns using openpyxl (2), so the lack of standardization isn't an issue. Finally, the program is completed by alerting the user to all the measured filters with abnormal data (i.e., outside the accepted range).

Conclusions and Impact

The MicroPEM program can help researchers who have not used a MicroPEM file to understand the data, which then allows them to continue to the next steps in the research. After researchers receive the data, these programs can serve as the initial screening step to quickly validate the data so that it can be analyzed further.

The OD program works similarly as an initial screening for OD data, so that the user knows whether measurements need to be made again, rather than after later analysis. This step can increase efficiency in the overall process and improve overall data quality by identifying technician or instrumentation errors quickly and effectively. A higher percentage of valid data increases the statistical significance of the research findings, thereby having a greater impact on public health.

References

1. Golden, T. (n.d.) Browse for a folder. *Tim Golden's Python Stuff.* http://timgolden.me.uk/python/win32_how_do_i/browse-for-a-folder.html
2. Sweigart, A. (2015). *Automate the Boring Stuff with Python.* San Francisco: No Starch Press. Web. 2 July 2019.

Presenting Author: Mukil Guruparan
School: UNC-Chapel Hill
E-mail: internships@rti.org
RTI International: 3040 E. Cornwallis Road
Research Triangle Park, NC 27709

Presented at: 11th Annual RTI Internship Showcase
August 12, 2019 | 1:00 pm–4:00 pm
Horizon Building
William M. Moore Jr. Collaboration Center