R "SHINY" GUI DEVELOPMENT FOR URANIUM ISOTOPIC ANALYSIS WITH MATRIX-ASSISTED IONIZATION MASS SPECTROMETRY

Elizabeth D. LaBone, Kyle M. Samperton, Danielle R. Mannion; Savannah River National Laboratory, Aiken, SC, United States

Abstract

The international nuclear safeguards community continues to seek rapid, accurate, and precise characterization capabilities for the in-field measurement of uranium isotopic compositions in nuclear facilities. Mass spectrometry (MS) is considered the "gold standard" for analysis of relatively long-lived actinides such as uranium (U) and plutonium; however, conventional MS analysis often requires time consuming sample preparation and complex analytical methodologies that are difficult to perform in-field or in-facility. Matrix assisted ionization (MAI) is a novel ambient ionization MS technique (i.e., MAI-MS) that potentially addresses these challenges due to the relative simplicity of the ionization phenomenon and ruggedness of ambient MS instrumentation. Savannah River National Laboratory (SRNL, USA) has demonstrated this technique for nanogram-level $^{235}U/^{238}U$ isotope ratio measurements within seconds, with percent-level analytical uncertainties capable of discriminating depleted, natural, and low-enriched uranium. Current experimental work on developing MAI methods for uranium isotopic analysis has been enabled by parallel development of a comprehensive MAI-MS data analysis suite at SRNL. Development of this bespoke data analysis software was necessary because commercially available ambient MS software is poorly suited for uranium isotope ratio measurement. The effort leverages the power of R, a popular open-source programming language, and Shiny, an R package providing tools for graphical user interface (GUI) and web interface coding. This software allows researchers without any programming experience to harness and utilize R's considerable data analysis/visualization power.

1. Introduction

Environmental sampling at nuclear facilities is one of the key safeguards measures for the detection of undeclared material or activities since early 1999.[1] Swipe sampling, as part of the International Atomic Energy Agency (IAEA) safeguards activities inside of nuclear facilities, is a process by which inspectors collect samples that may contain trace amounts of nuclear material. These samples are then transferred under strict chain of custody to the IAEA's Environmental Sample Laboratory (ESL) and Network of Analytical Labs (NWAL), which are specialized laboratories with dedicated environmental sample analysis capabilities. The IAEA and others in the safeguards community are highly interested in the timely detection of highly enriched uranium (HEU) production in low enriched uranium (LEU) enrichment plants.[2, 3] Such a capability would be transformational, enabling uranium isotopic analysis of trace samples on location for an immediate determination of uranium enrichment levels, potentially allowing faster information transfer to and decision making by the IAEA board of governors. Rapid screening of environmental samples could also allow inspectors to collect additional samples from suspect surfaces/areas and identify high priority sample locations within a safeguarded facility, thus increasing the likelihood of identifying undeclared activities.

Atmospheric pressure ionization mass spectrometry (API-MS) is a broad family of MS techniques utilizing ionization mechanisms that occur at or near ambient conditions.[4] These mechanisms exploit low-energy chemical interactions to generate ions, resulting in "soft" ionization that produces little to no chemical fragmentation of the analytes. Upwards of 40 unique API-MS techniques have been reported in the literature since the 1980s.[5] One of the most well-known variants of API-MS is the Nobel Prize-winning electrospray ionization (ESI) technique, which is now the standard ionization method for protein MS analysis.[6] API-MS is an attractive option for MS in-facility as its ion sources and instrumentation are relatively simple and robust when compared to more traditional MS methods[7] (e.g., thermal ionization (TIMS); multicollector–inductively coupled plasma (MC-ICP-MS)). One promising API-MS technique that eliminates the need for a conventional ion source is matrix-assisted ionization (MAI). First described in 2013, MAI is one of the newest API-MS techniques.[8] It was observed that ions were generated when a sample was mixed with a chemical called 3-nitrobenzonitrile (3-NBN) and placed under vacuum with no application of heat, electrons, high voltage, or lasers.[9] Early data suggested that triboluminescence may play a role in ion generation; however, some contrary data has subsequently been reported.[10, 11] Despite the limited mechanistic understanding of MAI, it demonstrates pico-to-femtogram detection capabilities for a number of organic analytes in complex matrixes, including tissue,[12] blood,[13] tea,[14] and urine[15]. SRNL was the first to perform uranium analysis using this method, demonstrating nanogram-level detection limits with analysis times <10 seconds, all with virtually no optimization.[16] For sample preparation, the sample is simply mixed with the matrix chemical 3-NBN and solvent (methanol). Then approximately 1-5 µL is directly aspirated into the atmospheric pressure orifice of an ambient MS. The aspiration into the vacuum creates a burst of ions that are separated and analyzed by the mass analyzer.

Recent efforts at SRNL have focused on examining MAI method development for uranium isotopic determination using a JEOL AccuTOF time-of-flight MS. One major problem faced by the research team was the lack of appropriate API-MS commercial software for targeted isotopic data analysis, since API-MS methods are traditionally used to analyze organics. In our search for alternative data analysis software, we found that data management and reductions tools for isotopic analysis are not commonly available even for established methods like TIMS and MC-ICP-MS. This being the case, the SRNL research team opted to develop an in-house data analysis capability using the R language and statistical programming environment. R is a powerful, popular, open-source data analysis tool[17] that is used across a diverse array of scientific and technical disciplines for data reduction, analysis, and visualization. Its impressive "base R" functionality can be augmented by freely downloading any of the >19,000 user-contributed packages on the online Comprehensive R Archive Network (CRAN) repository.[18] One such package is Shiny, which provides tools for convenient development and customization of "point-and-click" GUIs within the R environment.[19] Integrating R-based data analytics with Shiny-enabled data interactivity thus proffers an attractive route for any analyst to take advantage of R's statistical and computational power.

2. Structure of the Shiny GUI for MAI

The Shiny GUI was built as an R package using the Golem framework,[20] which includes documentation, testing functions, and code organization. A blank Golem project is set up to include folders for documentation, R code, data, and testing. As with any R package, for each function the Golem framework encourages the use of documentation and *vignettes* (i.e., short explanatory articles on the package and its functions, thus lowering the burden-of-entry for new users). These vignettes can be called from within the R programming environment. The testing folder contains files with code to evaluate the package's major functions and will indicate if changes to a function produce unexpected results. The basic structure of a Shiny GUI includes user interface (UI) and Server sections. The UI section is where the structure of the GUI is defined, while the Server is where calculations and reactive components are defined. The Golem framework has separate files for the UI
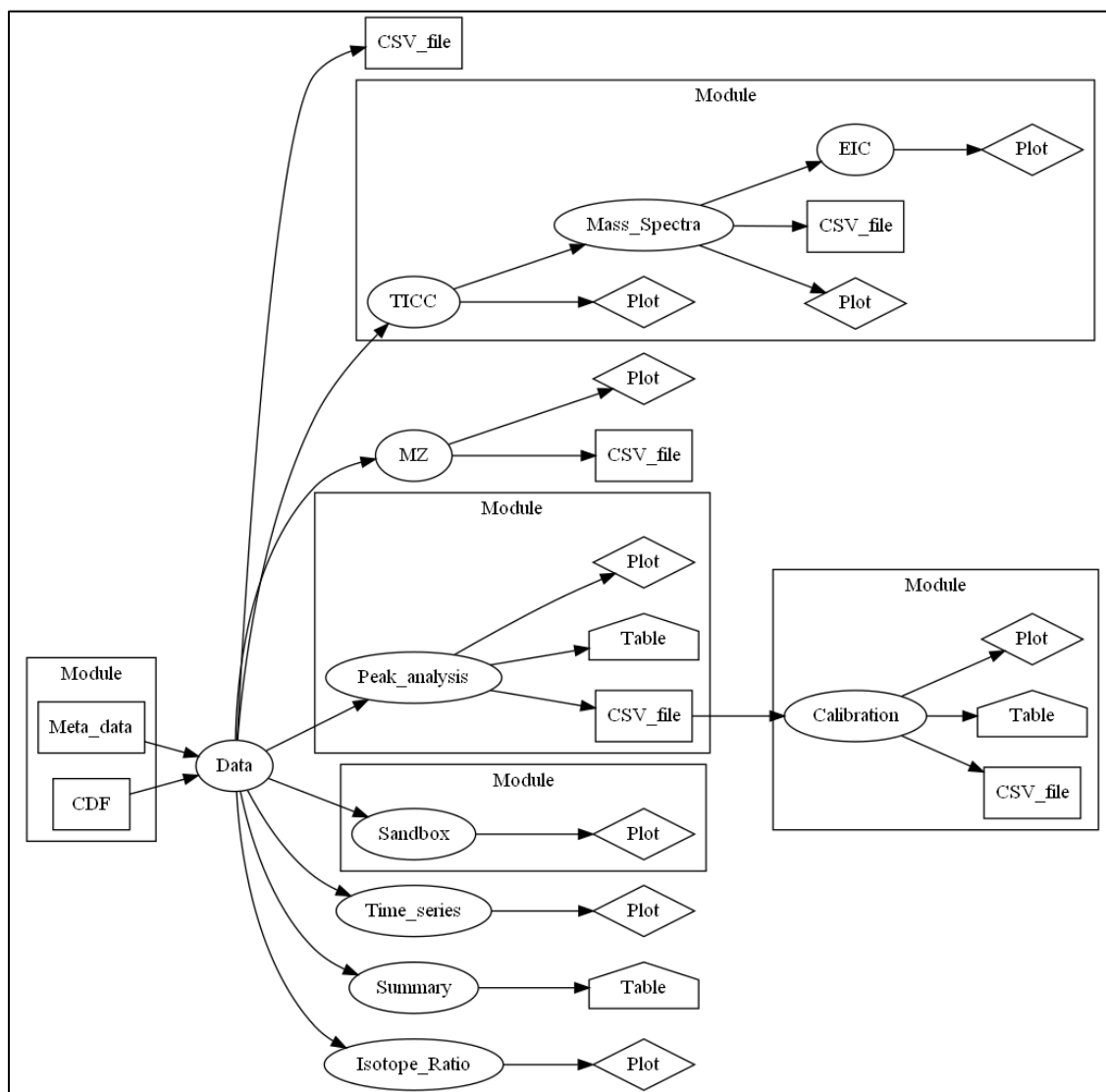

Figure 1. Data flow of the MAI Shiny GUI application. Shape in the above schematic indicates object type, including data file (small rectangle); analysis topic (oval); plot (diamond); table (pentagon); and module (enclosing rectangle).

and Server components. It encourages breaking down long sections of code into modules to facilitate code organization, readability, and maintenance. Modules are split-off sections of code that contain a UI and Server section and can be called from the main UI and Server; they are useful for organizing long sections of code by subject. For the MAI GUI, the code for the chromatogram, peak analysis, calibration, sandbox, and file handling sections are separated into dedicated modules.
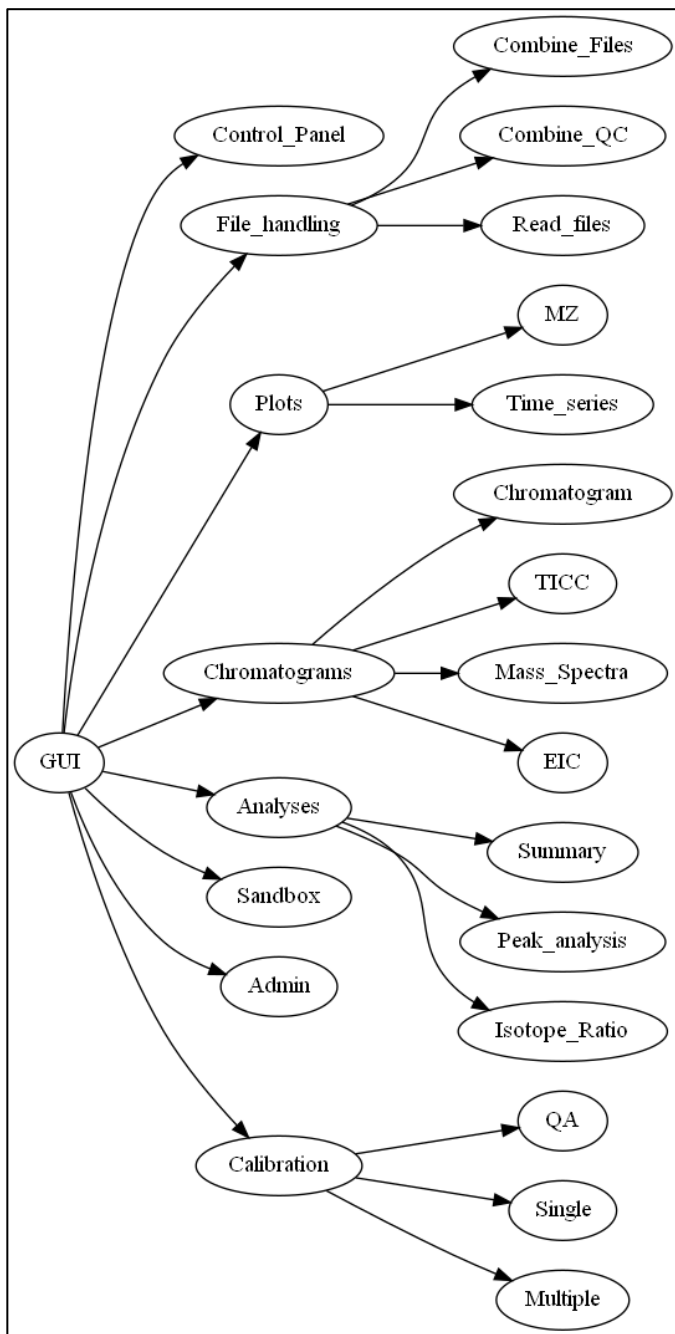
The GUI handles Common Data Format (CDF) and comma-separated value (CSV) format data for analysis. The AccuTOF MS software outputs data in CDF format, which is a non-human readable format used to efficiently store multi-dimensional data (https://cdf.gsfc.nasa.gov/). Within the GUI, the CDF data can be analyzed alone or in combination with a corresponding meta-data file that is populated by the analyst running the MS. The meta-data file is a CSV file with information such as the time when samples are introduced to the instrument, sample name, element of interest, concentration (if known), etc. The meta-data also contains information on whether the analyst running the MS considers a particular sample to have run successfully (true) or not (false), enabling exclusion of unsuccessful sample data from subsequent analysis. The CDF and meta-data file can be combined into one integrated CDF file. Data is output from the GUI in CSV format for use either in external programs or further in the GUI analysis pipeline.

The Shiny GUI can be conceptually approached from two directions: data flow (Figure 1) and structure (Figure 2). Combining both approaches allows for a more complete overview of the GUI. Elements of the GUI that do not handle data (e.g., Admin tab, Control Panel) show up in the structure diagram (Figure 2), but not the data flow diagram (Figure 1). Meanwhile, the structure diagram does not show dependencies between different tabs, whereas the data flow diagram does.

Figure 2. GUI structure of the MAI application.

A diagram based on *how data flows through the GUI* is useful for planning what data is required, what outputs will be produced, and the dependencies of each analysis. For example, the Time_series analysis requires only the input data; it interacts with no other analyses and outputs a plot. In contrast, the Calibration analysis requires that the data first go through Peak_analysis before it can be used. The data flow diagram also is a useful stage to define what is contained in each module. By comparison, a diagram based on *the structure of the GUI* is useful for showing the layout of the application. The Shiny GUI consists of a series of tabs and sub-tabs that allow the user to work through different analyses. Sub-tabs break up tabs by subject/step and keep any one tab from being too crowded.

3. Shiny App Demonstration

Upon initiating the GUI, the first tab the user sees is the Control Panel (Figure 3), wherein the user can make GUI-wide selections. The user 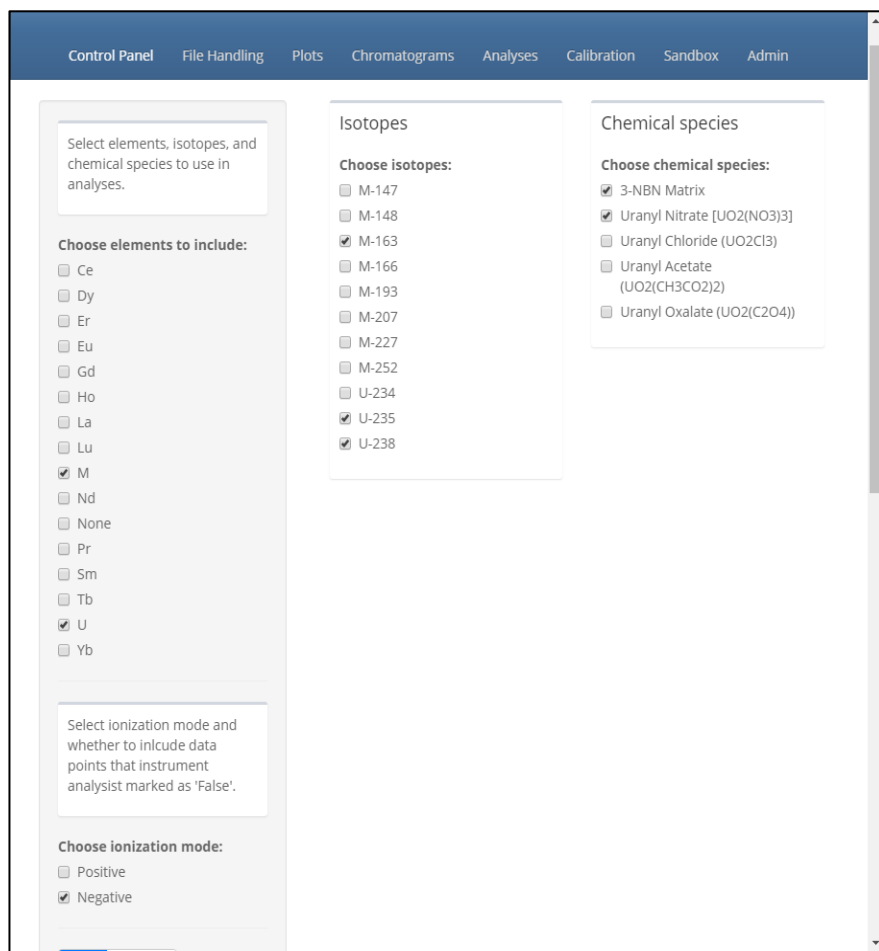can choose which chemical species to populate drop down menus later in the GUI by selecting those elements, isotopes, ionization states, and chemical species of interest. This decreases the time needed to scroll through exhaustively long, auto-populated drop-down menus. Not all tabs require the user to select chemical species (e.g., Sandbox). The user can select whether to include data points labeled 'false' in the meta-data. Plot options, such as theme, size, and file format, can also be selected within the Control Panel. The ability to set file format (e.g., PNG, PDF, EPS) is particularly useful if the user needs to prepare figures for a manuscript or presentation.

Figure 3. Control Panel tab within the MAI Shiny GUI, showing options for elements-/isotopes-of-interest; ionization mode; and chemical species.

The next tab is the File Handling tab, where files can be read in and combined. The user works through the Combine-Files and Combine-QC tabs the first time they upload a CDF file with an associated meta-data file. After the data is combined into one CDF, the file can be read into the GUI in the Read Files

tab. Files with or without meta-data can be read in on the Read Files tab, viewed, and downloaded as CSV files.

The user can then progress to one of the tabs for analyses after progressing through the Control Panel and File Handling tabs. The Sandbox, Plots, 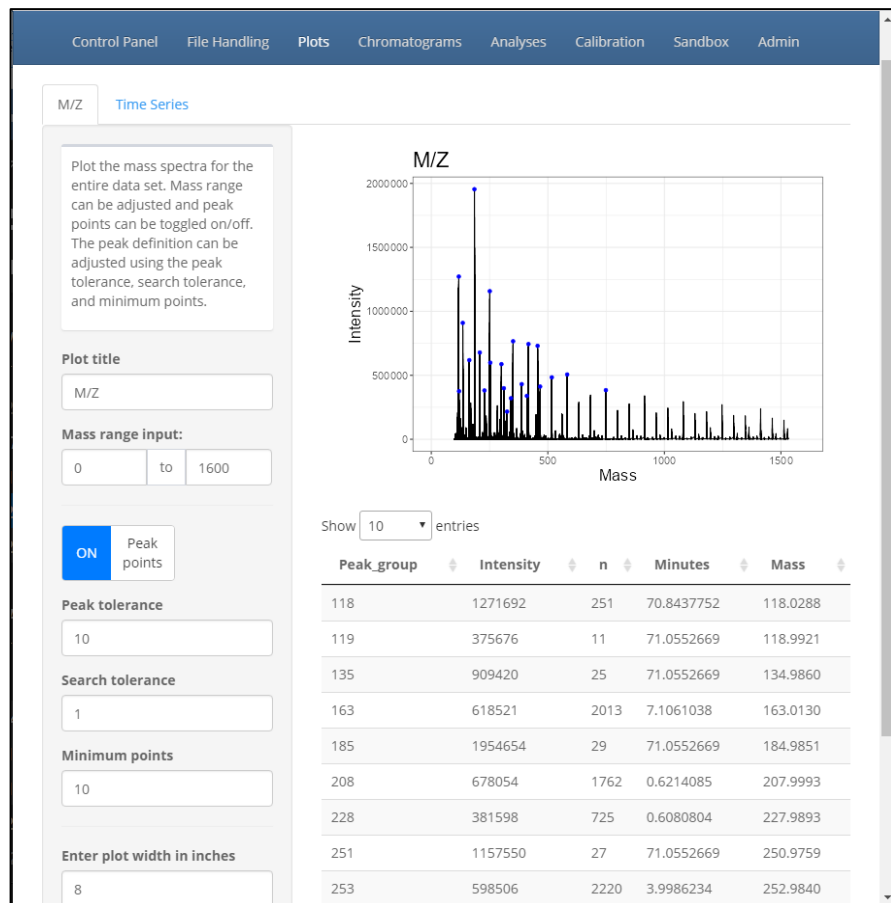and Chromatogram tabs enable graphical analyses of the data. The Peak Analysis and Calibration tabs are part of their own workflow of producing calibrated data and plots. The Admin tab contains fields to edit "global" GUI variables that are considered fixed in most instances and thus usually left unchanged (e.g., the minimum data points to define a peak for Peak Analysis), unless the user has both a complete understanding of their use and a compelling reason to adjust a given value.

The largely standalone Sandbox tab allows users to explore data graphically, either with or without meta-data. Within this tab the user can filter the data using a table below the plot by defining ranges for continuous variables and selecting desired categorical variables. Several graph styles (e.g., scatter, line, histogram) are available for selection. Color can be applied to the plot as either one color or based on continuous/categorical variables.

Figure 4. Plots tab showing example mass-to-charge ratio (M/Z) spectra.

The Plots tab contains sub-tabs to plot M/Z (i.e., mass-to-charge ratio) and times series. The M/Z sub-tab plots the mass spectra for the entire data set (Figure 4). Peaks are identified and marked with a blue dot that can be toggled on/off. The identified peaks are listed in a table below the plot with the intensity, mass, time, and number of points for each peak. The Time Series sub-tab plots a time series of a user-selected isotope from either a drop-down menu or manually entered M/Z value. The time series can be displayed as either a scatter or line graph. Both the M/Z and Time Series plots can be exported in the format chosen in the Control Panel and in dimensions input to Width/Height fields on the respective sub-tab. The M/Z peak table can be downloaded as a CSV.

The Chromatogram tab allows the user to plot the total ion count chromatogram (TICC) and select masses and times to plot the mass spectra and extracted ion chromatogram (EIC). The Chromatogram sub-tab plots the TICC and the MS for the entire time of analysis. The user can use their cursor interactively to draw a rectangle around a peak in the TICC and select a time interval; the mass spectra is then re-plotted using the selected time interval. The user can then click a peak in the MS to select a mass to plot the EIC. The TICC, mass spectra, and EIC each have their own sub-tab where the plots can be customized (e.g., edit title, change x-axis limits, etc.) and saved. The mass spectra sub-tab also includes a table with a line for each peak that includes the intensity, number of points, mass, time, and a proposed chemical species identity based on mass-matching to a chemical species reference table in the software.

The Analyses tab includes sub-tabs for Peak Analysis, Summary, and Isotope Ratio Plot. The Peak Analysis sub-tab is part of a larger workflow within the Calibration tab. The Summary tab displays the summary statistics (min, 1st quartile, median, mean, 3rd quartile, and max) for the time, mass, and intensity in the data set. Isotope Ratio Plot graphs the intensities of two user-selected isotopes as a scatter plot. A linear trend line and linear regression equation are also displayed on the plot.

The Peak Analysis sub-tab (Figure 5) and the Calibration tab (Figure 6) are part of a workflow to identify specific sample injections and apply an isotope ratio calibration correction to raw, measured data. As shown in Figure 5, the software is capable of automatically identifying discrete analytical aliquots within a single CDF data file's time-series and assigning corresponding

Figure 5. Peak Analysis sub-tab, showing an example of peak-finding logic applied to automatically identify discrete analytical aliquots within a continuous, multi-modal time-series in a single CDF data file.

information from an associated meta-data file. In the Peak Analysis sub-tab the user selects isotope ratios of interest; the resulting isotope ratio calculation intervals are plotted as vertical bars over a

scatter plot of the raw intensity data. Resulting time-series of isotope ratios of interests with corresponding uncertainties are plotted (Fig. 5), with the ratios and uncertainties summarized in a results table. The identified peaks, ratios, uncertainties, and associated meta-data can be saved to a CSV file.

The resulting CSV files saved from the Peak Analysis sub-tab can be used in the Calibration tab for isotopic calibration and data visualization. This is the first tab where meta-data is mandatory, as known primary and secondary standard (QC) concentration information is required for binning and measurement repeatability/uncertainty calculations. The QA (Quality Assurance) sub-tab is where the user uploads the Peak Analysis CSV file for a primary standard and samples/QC. The data is displayed as plots and tables, where "greyed-out" points/rows indicate samples labeled 'false' by the analyst operating the instrument that will not be used in calculations. The user can select points/rows to remove/re-add by simply clicking on them. The data can be saved to CSV once QA is complete.
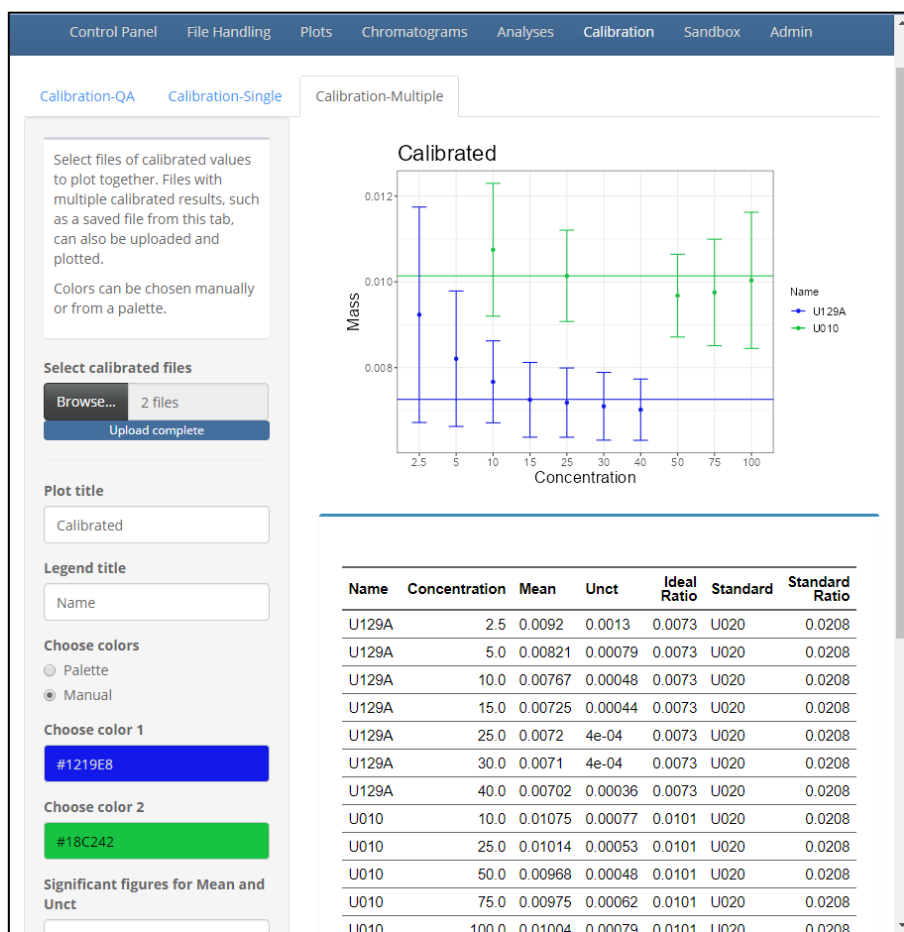

Figure 6. Calibration Multiple sub-tab, showing $^{235}U/^{238}U$ results of NBL CRM C129A (U129A in data table/plot) and U010 calibrated against U020.

The Calibration-Single sub-tab is where $^{235}U/^{238}U$ isotopic calibration occurs using a primary standard and a sample/QC. The user first uploads the QA-ed files from the QA sub-tab and selects or enters an ideal ratio, or certificate value, for each file. The ideal/certified ratio for the primary standard is used in the calibration calculation and the ideal/certified ratio for the sample/QC is only used for plotting and reference purposes. The primary standard and sample/QC data is binned by concentration value and displayed on a plot as points with $^{235}U/^{238}U$ uncertainties generated by the data reduction algorithm. A table displaying the name, concentration, mean, and uncertainty for each sample concentration of the two data sets is generated. The user can then press the 'Calibrate' button to run the $^{235}U/^{238}U$ calibration algorithm. A new plot and table of the calibrated

sample/QC data is displayed below the noncalibrated table. The plots and tables can all be downloaded by the user. The calibrated data can also be downloaded as a CSV for use in the Calibration-Multiple sub-tab.

The Multiple sub-tab allows the user to overlay multiple calibrated $^{235}U/^{238}U$ data sets on one plot and display the results in a single table. The user can upload several files that were saved from the Single sub-tab or upload files previously saved from the Multiple sub-tab. The data is plotted, displayed in a table, and can be downloaded as a CSV with all uploaded data sets included. Having the ability to save the data and re-upload it later allows for more easily reproduced figures.

6. Conclusion

This work demonstrates the successful construction and implementation of a data analysis package using the open-source R statistical language and programming environment for uranium MS data analysis/visualization. The package was built using end-user-driven requirements/feedback, enabling reproducible and rigorous data analysis with point-and-click simplicity by utilizing a GUI design. While this software was created specifically for MAI, the presented approach is broadly applicable to any mass spectrometry method. Future refinement of the MAI R package will include software quality assurance and implementation of a more refined isotopic calibration algorithm, with the power of this approach being that new capabilities/functionality can be incorporated relatively quickly into the existing programmatic framework.

7. Acknowledgements

5. References

1. Donohue, D., Strengthening IAEA safeguards through environmental sampling and analysis. *Journal of Alloys and Compounds* **1998**, 271, 11-18.
2. International Atomic Energy Agency (IAEA), *STR-375* **2013**.
3. International Atomic Energy Agency (IAEA), *STR-399* **2022**.
4. Huang, M.-Z.; Yuan, C.-H.; Cheng, S.-C.; Cho, Y.-T.; Shiea, J., Ambient ionization mass spectrometry. *Annual review of analytical chemistry* **2010**, 3, 43-65.
5. Feider, C. L.; Krieger, A.; DeHoog, R. J.; Eberlin, L. S., Ambient ionization mass spectrometry: Recent developments and applications. *Analytical Cheystry.* **2019**, 91 (7), 4266-4290.
6. Weston, D. J., Ambient ionization mass spectrometry: current understanding of mechanistic theory; analytical performance and application areas. *Analyst* **2010**, 135 (4), 661-668.
7. Snyder, D. T.; Pulliam, C. J.; Ouyang, Z.; Cooks, R. G., Miniature and fieldable mass spectrometers: recent advances. *Analytical Chemistry* **2016**, 88 (1), 2-29.
8. Trimpin, S., " Magic" ionization mass spectrometry. *Journal of The American Society for Mass Spectrometry* **2015**, 27 (1), 4-21.

9. Trimpin, S.; Inutan, E. D., Matrix assisted ionization in vacuum, a sensitive and widely applicable ionization method for mass spectrometry. *Journal of The American Society for Mass Spectrometry* **2013**, 24 (5), 722-732.

10. Trimpin, S.; Lu, I.-C.; Rauschenbach, S.; Hoang, K.; Wang, B.; Chubatyi, N. D.; Zhang, W.-J.; Inutan, E. D.; Pophristic, M.; Sidorenko, A., Spontaneous charge separation and sublimation processes are ubiquitous in nature and in ionization processes in mass spectrometry. *Journal of The American Society for Mass Spectrometry* **2017**, 29 (2), 304-315.

11. Trimpin, S.; Lutomski, C. A.; El-Baba, T. J.; Woodall, D. W.; Foley, C. D.; Manly, C. D.; Wang, B.; Liu, C. W.; Harless, B. M.; Kumar, R., Magic matrices for ionization in mass spectrometry. *International Journal of Mass Spectrometry* **2015**, 377, 532-545.

12. Wang, B.; Dearring, C. L.; Wager-Miller, J.; Mackie, K.; Trimpin, S., Drug detection and quantification directly from tissue using novel ionization methods for mass spectrometry. *European Journal of Mass Spectrometry* **2015**, 21 (3), 201-210.

13. Inutan, E. D.; Trimpin, S., Matrix assisted ionization vacuum (MAIV), a new ionization method for biological materials analysis using mass spectrometry. *Molecular & Cellular* Proteomics **2013**, 12 (3), 792-796.

14. Cody, R. B., Ambient Profiling of Phenolic Content in Tea Infusions by Matrix-Assisted Ionization in Vacuum. *Journal of The American Society for Mass Spectrometry* **2018**, 29 (8), 1594-1600.

15. Woodall, D. W.; Wang, B.; Inutan, E. D.; Narayan, S. B.; Trimpin, S., High-throughput characterization of small and large molecules using only a matrix and the vacuum of a mass spectrometer. *Analytical Chemistry* **2015**, 87 (9), 4667-4674.

16. Mannion, D. R.; Mannion, J. M.; Kuhne, W. W.; Wellons, M. S., Matrix-Assisted Ionization of Molecular Uranium Species. *Journal of the American Society for Mass Spectrometry* **2020**, 32 (1), 8-13.

17. R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria **2022**, URL https://www.R-project.org/.

18. Comprehensive R Archive Network (CRAN), URL https://cran.r-project.org/web/packages/.

19. Winston, C.; Joe, C.; Allaire, J. J.; Carson, S.; Barret, S.; Yihui, X.; Jeff, A.; Jonathan, M.; Alan, D.; Barbara, B., shiny: Web Application Framework for R, R package version 1.7.2 **2022**, URL https://cran.r-project.org/package=shiny.

20. Colin, F.; Vincent, G.; Sébastien, R.; Cervan, G., golem: A Framework for Robust Shiny Applications, R package version 0.3.5 **2022**, URL https://cran.r-project.org/package=golem.