

User Manual for

MQ

An open source tool for mass spectrometry data analysis

Developed by
Avinash Ghanate

Conceptualization by
Dr. Venkateswarlu Panchagnula
Principal investigator

Contents

1. Introduction
2. Installation
 - 2.1. *System prerequisites*
 - 2.2. *Installation procedure*
3. Salient features of MQ
4. Supported file formats
 - 4.1. *File conversion tool*
 - 4.2. *steps to convert raw data*
5. Description of various modules and instruction for data analysis
 - 5.1. *Spectrum viewer*
 - 5.2. *Isotopic confirmation*
 - 5.3. *Quan-Calibration*
 - 5.4. *Quan-Prediction*
 - 5.5. *Relative quantification*
 - 5.6. *Database query*
 - 5.7. *Multivariate analysis-PCA*
6. Acknowledgments

1. Introduction

MQ is software tool designed with an aim to provide a single platform for data analysis of mass spectrometry data from chromatography and direct ionization sources. It is a Java based open source project which relies on input files in generic formats such as ASCII and mzXML. A dedicated data conversion script has been provided along with the software that allows users to convert the data from various instrument/manufacture specific data formats into generic formats. The software contains various modules dedicated to perform specific data analysis steps.

The user guide details the various modules in the MQ algorithm, aided by step by step data analysis for test data set.

2. Installation

2.1. System prerequisite

- Compatible operating systems: Windows XP/7 or later, Linux, Mac OS
- Operation of MQ requires Java Runtime Environment (JRE) pre-installed on computer. The latest version of JRE can be downloaded and installed from [here](#). Detailed instructions to install and configure JRE are provided [here](#).

2.2. Installation procedure

2.2.1. Download the MQ.zip from [here](#) and save it in working directory.

2.2.2. Unzip the software package.

2.2.3. Open the folder and execute jar file named “mq.jar”. This will open up the graphical user interface (GUI) of MQ.

3. Salient features of MQ:

- 3.1. Open source/generic file formats are supported (ASCII/mzXML)
- 3.2. Accompanied with data conversion tool which enable users to convert data from various manufacturers into compatible open source formats mentioned above. An additional module “MSAvg” that transforms chromatographic MS data into average compiled spectra file.
- 3.3. Support data processing for LC/GC-MS with data transformation or direct ionization-MS based data.
- 3.4. Spectrum viewer module with simple interface and configurable parameters for optimizing peak detection based upon the quality of acquired data. Allows users to label peaks, save the spectra in image format, copy list of detected peaks and corresponding intensities and other visual changes.
- 3.5. Confirmation of detected ions based upon natural isotopic abundance distribution.
- 3.6. Ability to perform absolute/relative quantification from mass spectrometry data. Users can generate different calibration models and use them to quantify multiple analytes in single processing step.
- 3.7. Ability to query m/z from available databases.

3.8. Untargeted multivariate analysis module offers the principle component analysis (PCA) of MS data.

4. Supported file formats

MQ accepts ASCII and mzXML file formats of data having only MS spectral profile. Hence, conversion of native data formats and their transformation in average profile is necessary to be carried out. This operation is carried out using two modules. One of the module “msConvert” has been adopted from ProteoWizard that generates MS level 1 data in ASCII format and the output data is further processed using module “MSAvg” to generate an average MS profile data.

4.1. File conversion tools

- ProteoWizard software package (Download link [here](#))
- MSAverage (Download link here)

4.2. Steps to convert raw data

4.2.1. Download the .msi file for proteowizard from the above mentioned link and install the software package.

4.2.2. Download MSAvg file and save in working directory.

4.2.3. Through the installed program menu, search “MSConvert” and run the file.

4.2.4. You will be able to see the following interface on successful execution of previous steps.

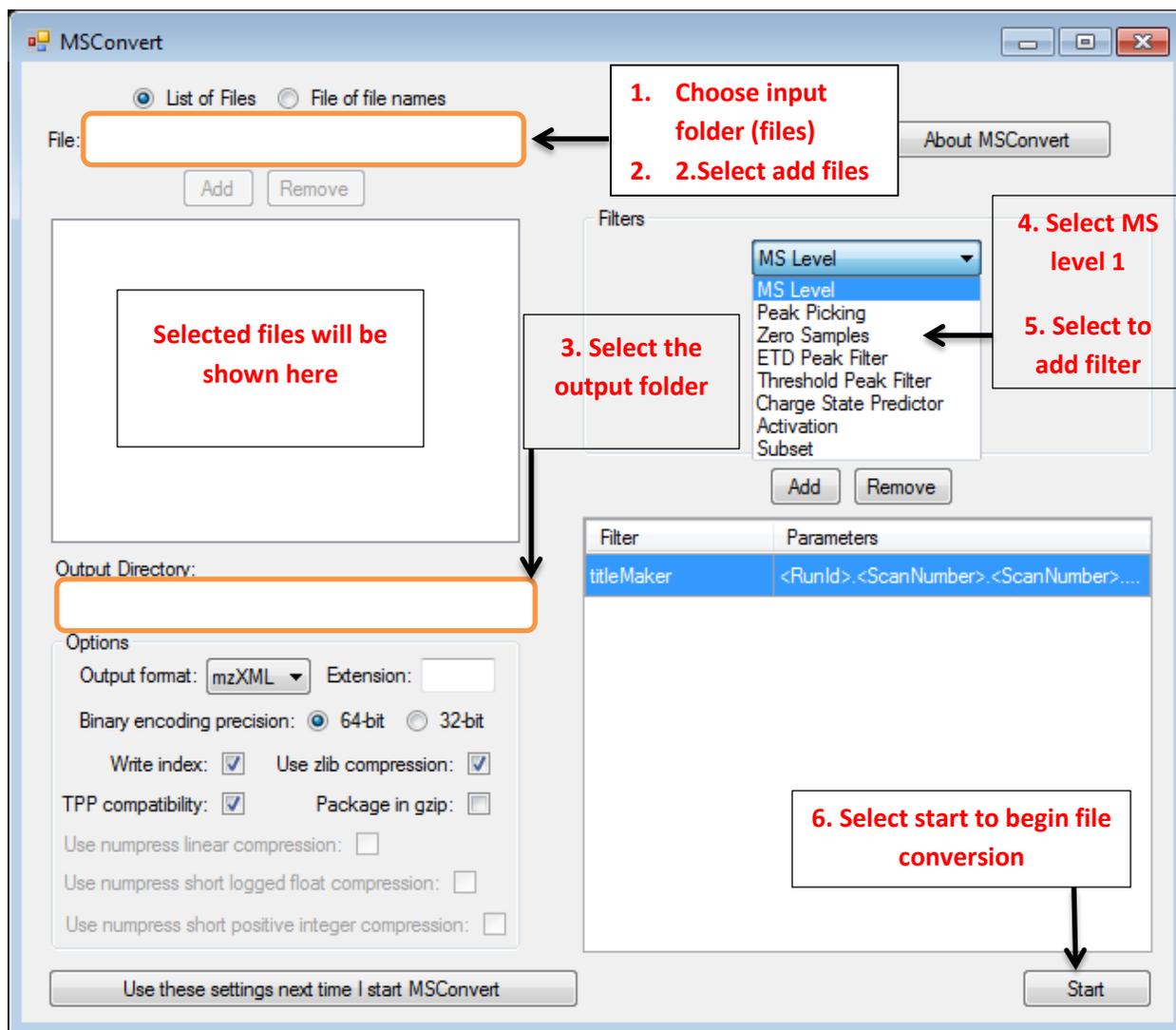


Figure. 1. Screenshot of MSConvert module (Proteowizard) depicting the steps in file conversion

4.2.5. After obtaining the first level of file conversion. Use the output file for further file conversion (averaging) through MSAvg.

4.2.6. The output files from MSAvg module can be used for data processing using MQ.

5. Description of various modules and instruction for data analysis

5.1. Spectrum viewer

5.1.1. User interface

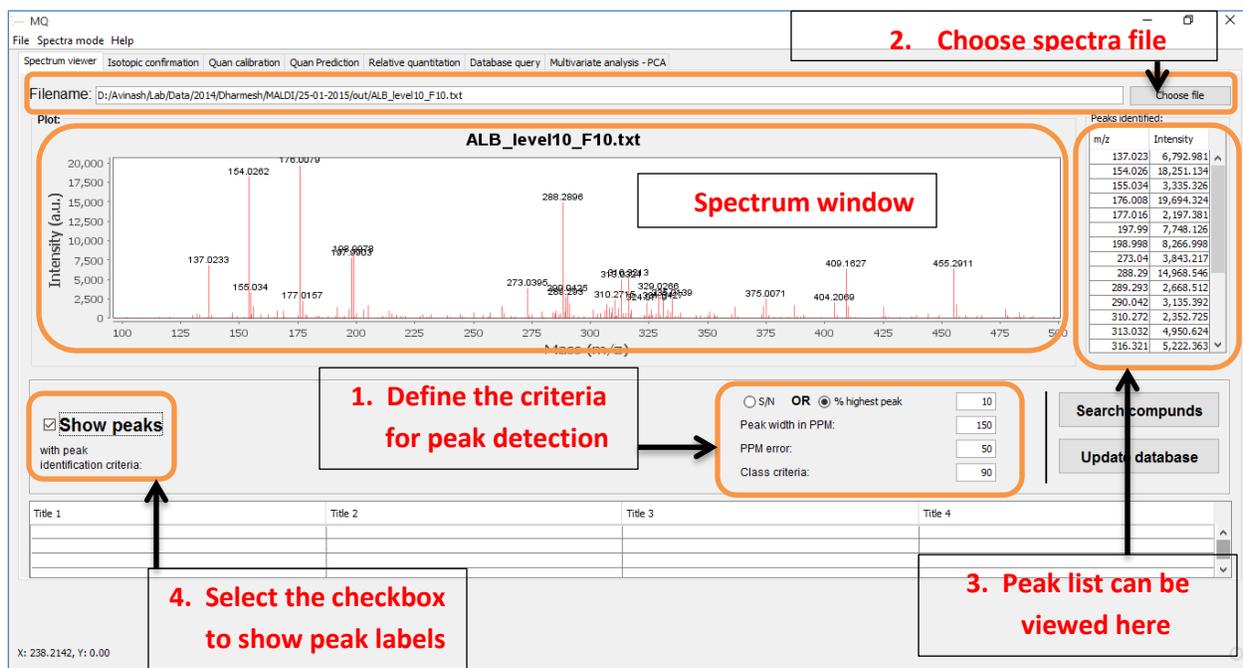


Figure. 2. Interface of spectrum viewer module

5.1.2. Steps to view spectra file

5.1.2.1. To view a spectra file in mzXML/ASCII format chose “open file” to select a particular spectra file from a directory and load the file.

5.1.2.2. The spectra will be displayed in the spectra window. User can zoom in to the spectra by pressing “left click and dragging” it across the region of interest. Vice-versa the zoom-out function is carried out by pressing “right click and dragging towards left”.

- 5.1.2.3. To view the list of peaks, define the peak detection criteria such as S/N or % intensity to the highest peak, peak width, PPM error. These parameters can vary according to the resolution of the instrument on which the data was originally acquired.
- 5.1.2.4. After defining the peak detection criteria, select show peaks to label the peaks. The list of detected peaks according to the peak detection parameters is displayed as a list on the right side of spectrum window. If peak(s) are not detected through initial criteria, users can tweak the values in order to include or exclude a peak.
- 5.1.2.5. Spectrum viewer also allows user to scan through the database to identify and annotate particular peak(s) of interest in automated fashion. User can create and save their own databases or can use already available databases for peak annotation.

5.1.3. Modifying display properties

- 5.1.3.1. Right click in the spectrum window and select “properties” from the drop down menu.
- 5.1.3.2. Click on the X/Y-axis tab to modify various plot properties such as font, font size, colour and axis name.

5.2. Isotopic confirmation

5.2.1. User interface

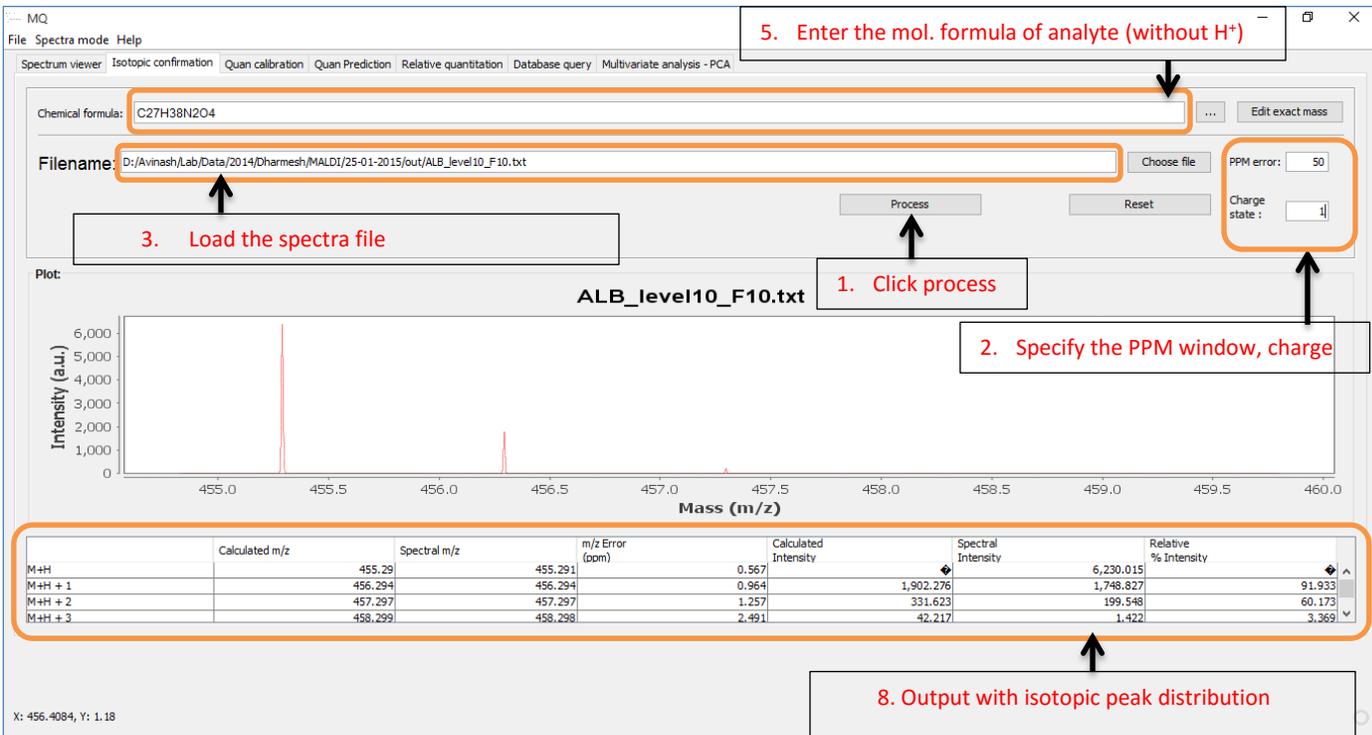


Figure 2. Interface of isotopic confirmation module

5.2.2. Steps for validating isotopic pattern

5.2.2.1. Enter the chemical formula of the analyte of interest with the adduct ion. For protonated species enter molecular formula considering $[M+H^+]$ and for sodiated adduct ion with $[M+Na^+]$.

5.2.2.2. An option to add/edit the exact mass of elements with respective abundances for isotopic distribution is also provided in the window. User can click on the “Edit exact mass” tab and enter elements with exact mass values along with

isotopic abundance values (natural or doped) according to the need of the analysis.

5.2.2.3. Chose a spectra file and specify the PPM window and charge state of the molecule.

5.2.2.4. Click process for calculating isotopic pattern and matching index with observed peak cluster.

5.2.2.5. The output will contain the isotopic pattern of the peaks in the form of spectra as well the intensity table of the peaks. The output table contains the details about theoretical peak intensities, observed peak intensities and relative % error.

5.3. Quan-calibration

This module has been designed to generate calibration models of various analytes through standard spectra files. Multiple numbers of analytes can be processed simultaneously. User can choose amongst the most prominent adduct or adducts with best regression values.

5.3.1. User interface

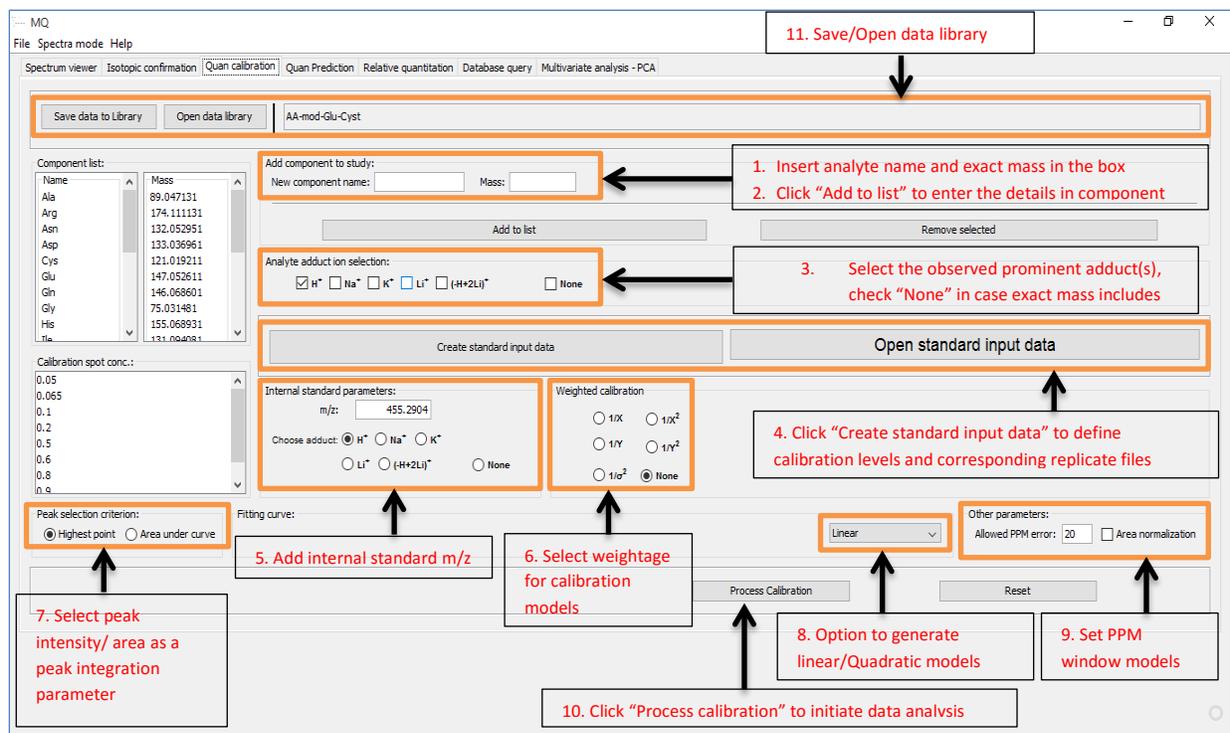


Figure. 3(a). Interface of relative quantification module

5.3.2. Steps to generate calibration models

- 5.3.2.1. Add component name (Analyte name) and corresponding exact mass in the "Add component window". User can remove the component/analyte by selecting it through component window and clicking on "Remove component" tab.
- 5.3.2.2. Set internal standard "m/z" and adduct(s) of choice.
- 5.3.2.3. Select prominent adduct ion for standard component and internal standard or select "None" if the exact mass provided includes adduct ion mass.
- 5.3.2.4. Click "Create standard input data" tab and add levels of metabolites and select the associated standard spectra files. Add the calibration levels and save the file.

- 5.3.2.5. Select other parameters such as calibration models (weighted/non-weighted), peak intensity/area, Linear/quadratic fit and allowed PPM error.
- 5.3.2.6. Click “Process calibration” tab to initiate data processing.
- 5.3.2.7. The final output windows will be displayed after successful processing of data. The screenshots of representative output windows are shown below in figure 3(b) and 3(c).

Calibration Results

Calibration result Calibration Plot

Results:

Component	Adduct	Slope	Intersection	Regression
Ala	H+	0.46326288078053196	0.5357359420294722	0.975889148459437
Arg	H+	3.609956085586572	0.020963467294169036	0.9988872605924438
Asn	H+	0.3093958978768705	0.015673758569060094	0.9899734764486221
Asp	H+	0.5606938746543794	0.04366985224457365	0.9929978371116506
Cys	H+	0.002009258901191342	0.012825954387933432	0.07362477407369417
Glu	H+	0.5143115227173904	0.029266535804113224	0.9947631017266387
Gln	H+	10.591901960999884	0.6210750692125271	0.9945306972607091
Gly	H+	1.4100875575035858	0.13215880835147925	0.9939181484814388
His	H+	1.7426689846572905	0.02956017512854231	0.9984142107736068
Ile	H+	51.7701438541396	2.537348116790194	0.9955021538734495
Lys	H+	5.685857850456548	0.059078534626060275	0.9985587373511886
Met	H+	2.571512177884135	0.09692963800052395	0.9925077291923502
Phe	H+	7.8288885377881385	0.5752704424050215	0.990837331994899
Pro	H+	1.5558284193556364	2.844286578576503	0.9271118369531559
Ser	H+	1.4731872496811809	0.16719285669331807	0.9906348064200614
Thr	H+	4.998423758447952	0.38845282921893576	0.9913943779977689
Trp	H+	0.5825475675449886	0.03812740284659409	0.9897945486927755
Tyr	H+	2.6847958038339925	0.18211962073030763	0.9922853048095062
Val	H+	20.511868199025727	7.229434492947114	0.9906925709232839
Cyst	H+	0.5563969428464783	0.007371919568161904	0.9973926250722742
GlucNa	H+	0.027117222457360035	-0.0013927600394728807	0.9854585417067553

Output of Quan-calibration module. User can Copy/paste the content of this table.

Figure. 3(b). Output with calibration models of analytes (result table).

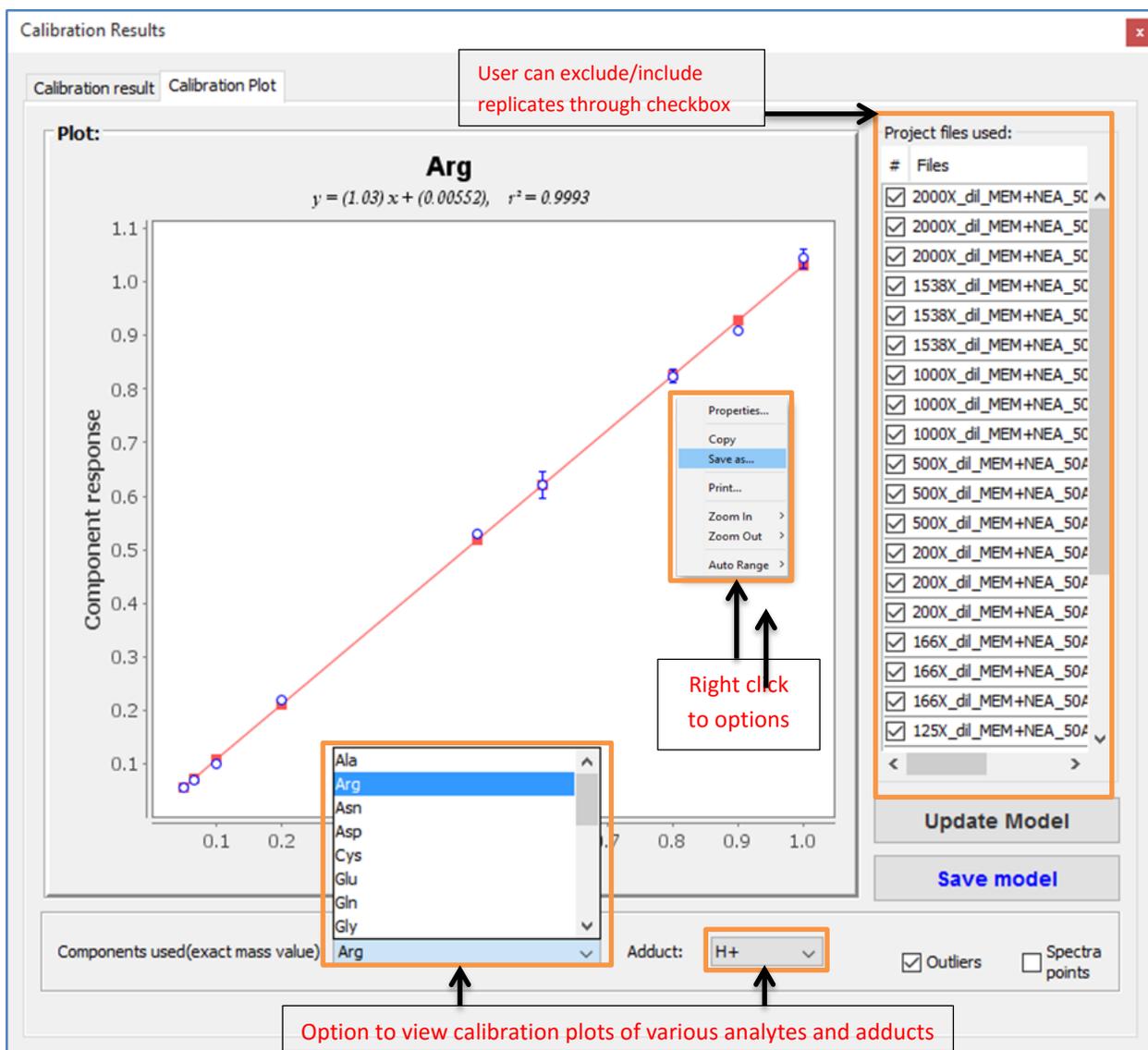


Figure. 3(C). Output with calibration models of analytes (calibration plots with replicate details).

5.3.2.8. Save the calibration model in the appropriate directory.

5.4. Quan-prediction

5.4.1. User interface

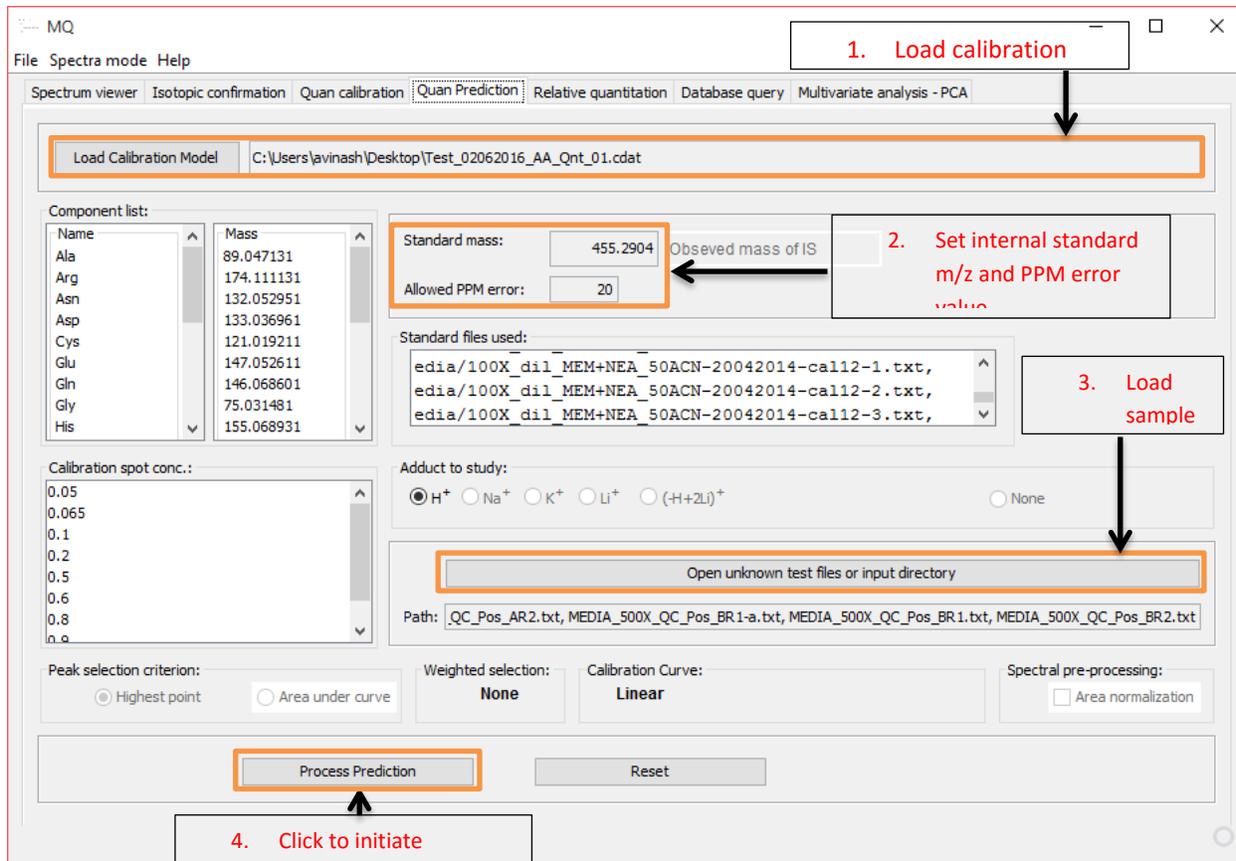


Figure. 3(C). Interface of Quan-prediction module

5.4.2. Steps to perform quantitation of samples

- 5.4.2.1. Load the previously created calibration model. After successful loading of model component list present in the model file will be displayed in the component window.
- 5.4.2.2. Enter the internal standard m/z , allowed PPM window value as well as potential adduct.
- 5.4.2.3. Load the sample spectra files through “Open unknown test files or input directory” tab.

5.4.2.4. Click “Process prediction” tab.

5.5. Relative Quantification

5.5.1. User interface

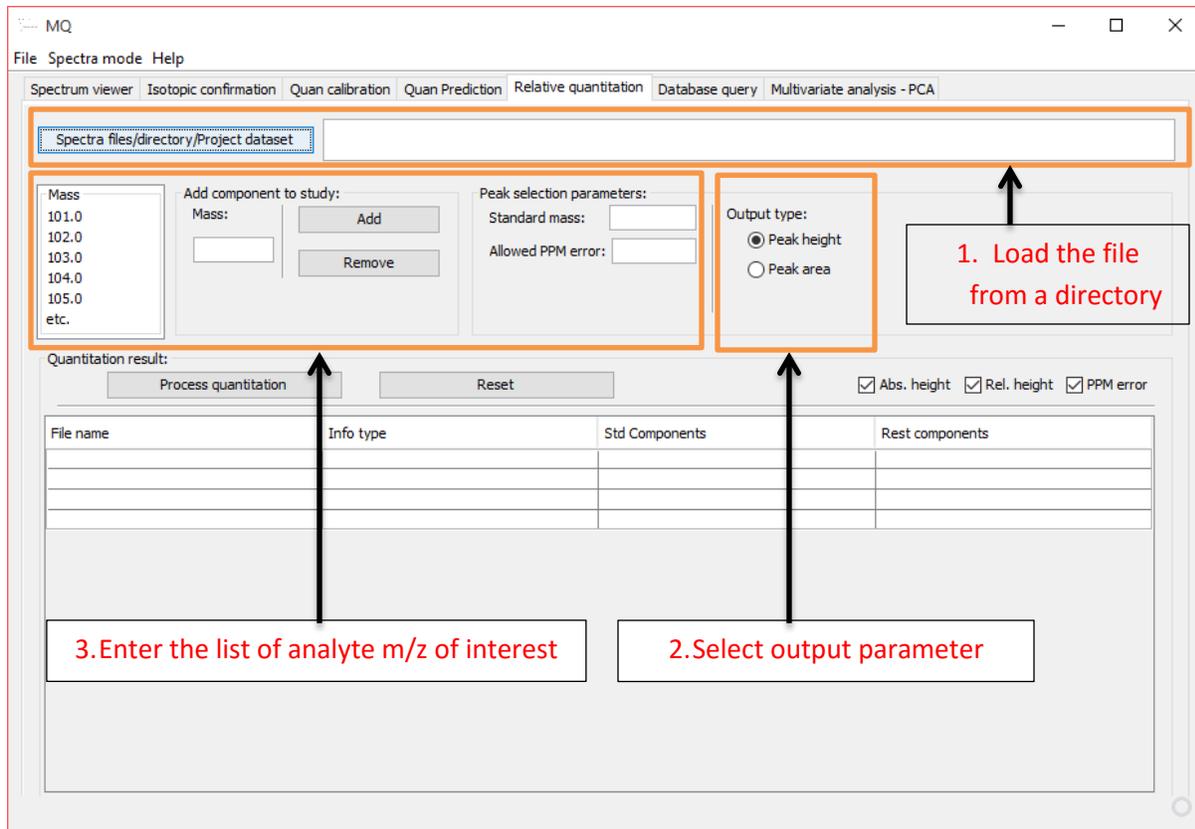


Figure. 5. Interface of relative quantification module

5.5.2. Steps for relative quantification

5.5.2.1. Load list of files/directory/Project dataset for which relative quantification has to be performed.

5.5.2.2. Enter the m/z values in the component section. A list of m/z can be copied into component window to add multiple m/z at a time.

5.5.2.3. Enter m/z of internal standard in standard mass window.

- 5.5.2.4. Specify PPM window as per the accuracy and resolution of MS data set.
- 5.5.2.5. User can choose either peak intensities or peak area based output.
- 5.5.2.6. Initiate the relative quantification by selecting process quantitation tab.
- 5.5.2.7. After completion of the data processing the output will appear in the result table. The result contains file-name wise details of detected m/z , absolute intensity/area and their normalized response with respect to internal standard. The result table can be copied into excel-sheet for further usage.

5.6. Database query

5.6.1. User interface

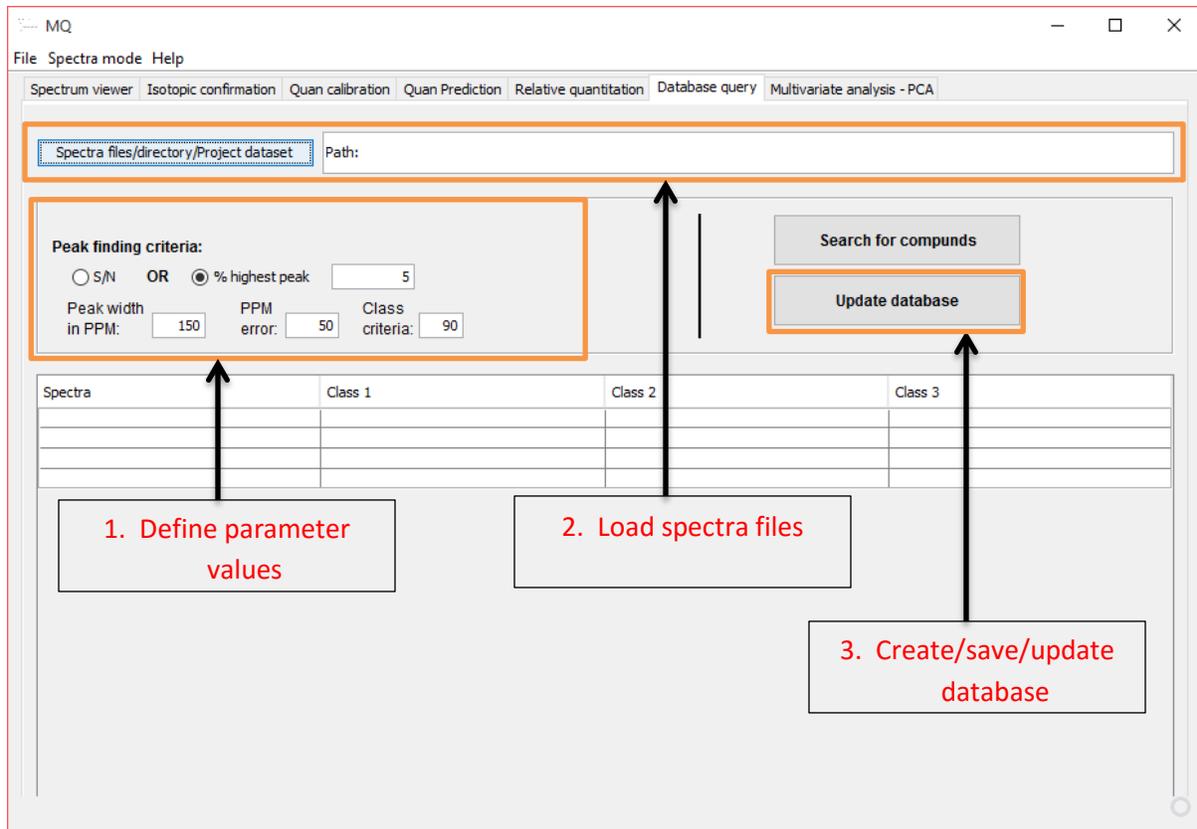


Figure. 6. Interface of relative database query module

5.6.2. Steps for peak annotation through database search

5.6.2.1. Through this module user can create and save databases or call the available databases for peak annotation of the spectral files/project data set.

5.6.2.2. Load list of files/directory/Project dataset for which peak annotation has to be performed.

5.6.2.3. Specify the values for peak finding criteria.

- 5.6.2.4. Click “search for compounds” tab to initiate the query.
- 5.6.2.5. The output is provided in the result table below with list of file names and corresponding m/z hits obtained against the query.
- 5.6.2.6. User can copy this output table into excel-sheet for any further data processing.

5.7. Multivariate analysis – PCA

This module allows users to carry out unsupervised multivariate analysis of MS data. It can be used for various applications such as grouping the data sets, identifying outliers and identifying trends of variation within a set of data. MQ uses principle component analysis (PCA) as a multivariate analysis tool for MS data.

5.7.1. User interface

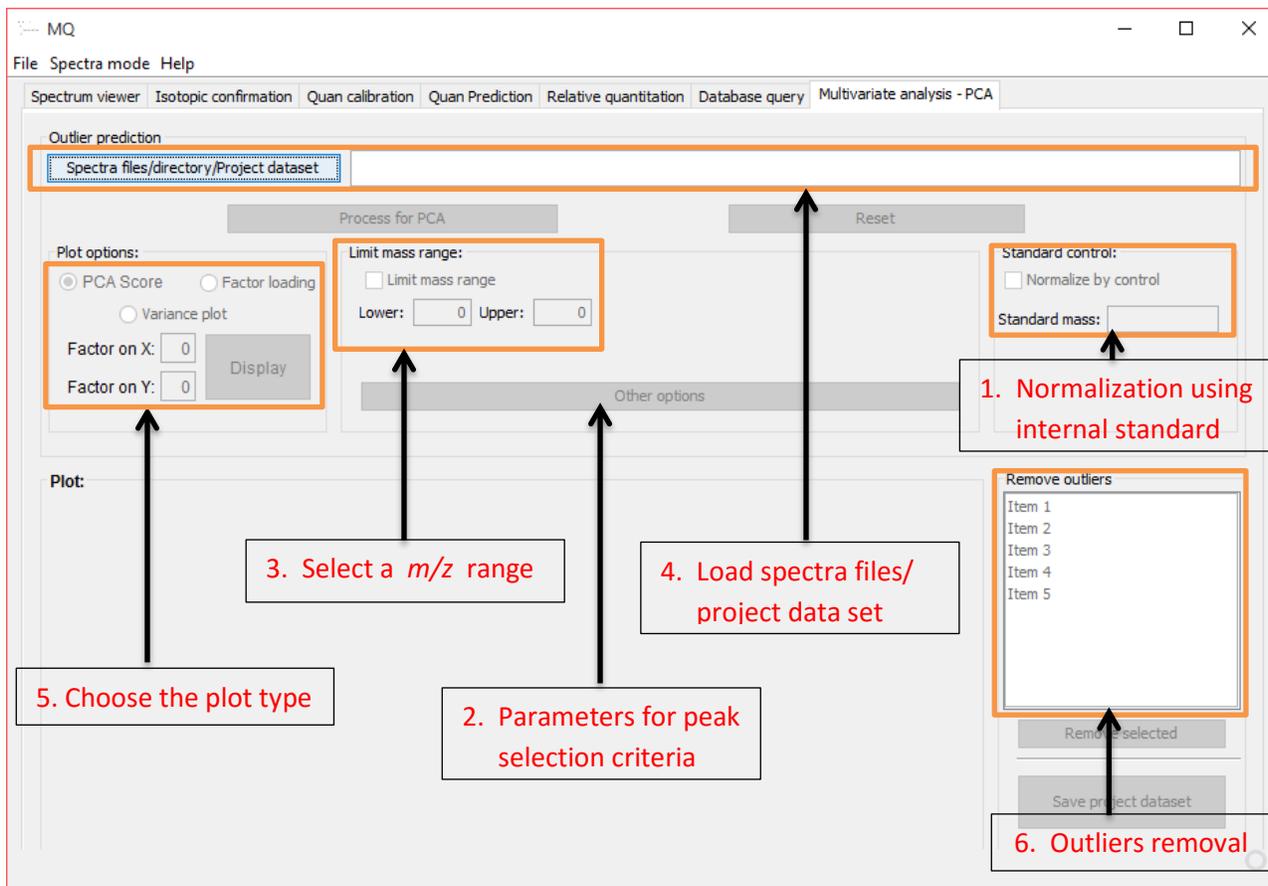


Figure. 7. Interface of relative multivariate analysis module

5.7.2. Steps for performing multivariate PCA analysis

- 5.7.2.1. Load the list of spectra files/project dataset by clicking on “Spectra files/directory/project dataset” tab.
- 5.7.2.2. Users can choose the range of m/z within which PCA analysis is desired. To include the criteria of m/z , select “Limit mass range” option. Subsequently set the lower and upper m/z limits.
- 5.7.2.3. Click on “Other option” to specify peak selection criteria such as Peak Width, S/N or cut off % relative to base peak. Additionally, user can select m/z list from available databases for PCA analysis instead of using all the peaks from the spectra.
- 5.7.2.4. Data normalization prior to PCA can be carried out using internal standard (IS). Specify the m/z of IS along with adduct ion.
- 5.7.2.5. User can also choose between different types of plots, such as PC score, factor loading and variance plot. As a default “PCA score plot” has been selected.
- 5.7.2.6. Users can update list of files through “outlier removal” option. If project dataset was used as an input for list of files, this updated file list can be saved as project dataset entry.

5.8. dacdp

