How to use the “CMS Workflow” perspective in OPENChrom

OPENChrom is an open source software package that is primarily used to analyze and visualize chromatography and mass spectrometry measurements. It can be found at <https://openchrom.net/>.

The CMS Workflow perspective adds to OPENChrom the ability to decompose a series of complicated mass spectra, obtained by sampling a gas mixture over time, into the concentrations of the individual components found in the mixture for each spectrum. The CMS Workflow perspective is primarily intended for use in quantitative analysis of gas mixtures by mass spectrometry. The decomposition is accomplished by solving a set of linear equations to minimize the sum of squares error between a set of component library mass spectra and the unknown mass spectrum. This decomposition is done for each spectrum making up the unknown series. The decomposition result can be presented as either the partial pressure (in pressure units) or mol fraction (dimensionless) of each library component that is present in the unknown spectrum. These results are displayed in the form of graphs presented on screen and also made available as text tables which can be copy/pasted into another program, such as Microsoft Excel, for further analysis. Both the composition results and the residual spectrum are made available as tables.

To begin, both the unknown spectra series and the known library component spectra need to be put into CMS format. The CMS format is text based and a detailed specification is given in a separate document.

Once files are in CMS format, start OPENChrom. At the startup screen, go to the <Windows> pull-down menu, select <Perspective Switcher>, scroll down and select the <CMS Workflow> perspective, and click OK.

The CMS Workflow perspective shows 6 views or windows. The borders between views can be click-dragged to enlarge individual views as needed during the workflow. There are 3 views related to library editing in the upper half of the screen. These library editing views, while offering some utility at the present time, are still in development and will not be described further at this time.

There are 3 views in the lower half of the screen. These views are where the mass spectra decomposition is set up and where the results are displayed.

The leftmost view has a tab labeled “Library Spectra”. The Library Spectra view is used to load component library files (in CMS format) and select which components from the library should be considered during the decomposition.

The middle view has a tab labeled “Signals” and a tab labeled “Console”. The Signals view is used to load and display the total ion current of a target unknown spectra, initiate a decomposition, and display the residual ion current remaining after decomposition has been performed. The Console view holds a table of both the composition results and the residual ion current results in a tab separated text format. These tables can be copy/pasted into another program if needed for further analysis, viewing, or reporting.

The rightmost view has a tab labeled “Compositions” and a tab labeled “Residual Correlations”. The Compositions view is used to display the calculated composition results after a decomposition has been performed. The Residual Correlations view is used to display a plot of how well the residual ion current spectra correlate with library components which were not included in the decomposition. This can be helpful in identifying library components which appear to be present in the unknown spectra and should be considered in the decomposition.

To begin, select the Signals view and load the target unknown spectra file by clicking on the leftmost of the three buttons located at the top of the view. This button can be identified by hovering the mouse cursor over the button until the text “Select the \*.cms spectra.” appears. Browse to the target unknown spectra file, select it, and click Open. The spectra file will load and the fully qualified file name will appear in the text box to the left. The graph will display a plot of total ion current (obtained by summing the ion currents for each ion mass present) as a function of elapsed time in seconds. If elapsed time is not present in the file data, the plot will show scan number instead of elapsed time. If scan number is preferred to be shown instead of elapsed time, de-selecting the “Use ETimes” checkbox will show plots as a function of scan number and not elapsed time.

If desired, the Start Time/Scan and End Time/Scan can be adjusted to define a sub-set of the unknown spectra file. Just click on the up and down arrows that are associated with the spinner controls to the right. The Start and End positions are shown by vertical lines in the graph and the numerical values for Start and End ETimes and Scan Number are shown in the text boxes to the left of the spinners.

Next, select the Library Spectra view and load the component library file by clicking on the button locate at the top right of the view. This button can be identified by hovering the mouse cursor over the button until the text “Select the \*.cms library file.” appears. Browse to the library file, select it, and click Open. The Library file will load and the fully qualified file name will appear in the text box to the left. A list of the library components by name will appear below.

Select the individual library components to be considered during the decomposition by clicking on the name of each library component to be considered. When selected, the component name changes to show an “\*” at the beginning of the name.

After all the library components have been selected, go back to the Signals view and click on the Decompose button, the middle of the three buttons located at the top of the view. This button can be identified by hovering the mouse cursor over the button until the text “Decompose the selected CMS spectra.” appears. Clicking on the Decompose button begins the decomposition process.

When the decomposition process completes, the signals graph is updated to add a plot of residual total ion current. Residual ion current is calculated by subtracting the ion current which is due to the calculated composition from the original ion currents. The difference between the original ion current and the residual ion current represents how much of the ion current has been accounted for by the calculated composition. If the residual ion current is near zero, then the calculated composition closely predicts the measurement. If the residual ion current is near the original ion current, then the calculated composition is not predicting the measurement and it is likely that additional components are present in the spectra. To reduce this discrepancy, select additional library components as appropriate and repeat the decomposition.

When the residual ion current is near zero, it can be hard to see in a linear plot. By checking the “Use Offset Log Scale” checkbox, the Y-axis is converted to log scale and a small offset is added so that any negative or zero signal values are still displayed on the graph. The value of the small offset is also shown as a horizontal line on the graph and its numerical value is shown in the legend. The “Use Offset Log Scale” checkbox only affects how results are displayed in the signals plot, the text tables in the Console view are unchanged.

Normally, a weighted least squares fit is performed where the errors are weighted by multiplying by 1/sqrt(abs(signal)). This weighting is done to increase the significance of small signals so that they do not get lost in comparison with large signals. Error weighted least squares is necessary when the signals have a high dynamic range over several orders of magnitude and have constant relative error. A non-weighted least squares fit can be performed by unchecking the “Use Weighted Error” checkbox. But I have not seen any value when doing this so the “Use Weighted Error” checkbox may be eliminated in a future revision.

Sometimes a mass spectrometer can have a small offset signal present. This offset signal represents an error and biases the ion current measurement either positive or negative relative to the actual ion current. This can result in ion current measurements that are either too low, even negative, or too high compared with the actual ion current. By checking the “Apply Zero Correction” checkbox, an attempt is made to compensate for offset signal errors using the following approach. For every spectrum in the spectra file, the ion current measurements at m/e values 5, 6, and 7 (if present) and the ion current measurement that has the median value is subtracted from all the ion current measurements in the spectrum. For most gas mass spectra, ions are absent at masses 5, 6, and 7 so that the actual ion current for these masses is zero and the measurement represents the offset signal. The median value is selected in an attempt to avoid choosing an anomalous signal offset.

When the decomposition process completes, the graphs in the “Compositions” and “Residual Correlations” views update.

The Compositions view shows the calculated composition as a function of either time or scan number as selected by the “Use ETimes” checkbox in the Signals view. The composition units can be changed between Partial Pressure, Mol Fraction, and Library Fraction by checking the appropriate checkbox. For Partial Pressure units, the compositions are presented as the partial pressure of each component using the pressure units that were given in the SPUNITS field in the unknown spectra file. When the source pressure varies between scans, it can be useful to display composition in mol fraction where the component partial pressures are divided by the source pressure. When source pressure information is lacking, then compositions are displayed in units of Library Fraction. Library Fraction units are dimensionless and represent the number by which the ion currents in the library component spectrum are multiplied to account for their contribution to the ion current calculated in the solution. It is this calculated ion current which gets subtracted from the original spectrum ion current to generate the residual ion current.

When the composition results are near zero, it can be hard to see in a linear plot. By checking the “Use Offset Log Scale” checkbox, the Y-axis is converted to log scale and a small offset is added so that any negative or zero signal values are still displayed on the graph. The value of the small offset is also shown as a horizontal line on the graph and its numerical value is shown in the legend. The “Use Offset Log Scale” checkbox only affects how results are displayed in the compositions plot, the text tables in the Console view are unchanged.

The text box located below the “Use Offset Log Scale” checkbox can be used to change the small offset when needed to better show the compositions. Just change the default value to something slightly lower or higher and the graph will update automatically. Keep adjusting the offset value until you are satisfied with the result. Un-checking the “Use Offset Log Scale” checkbox will return to the default linear Y-axis and checking the “Use Offset Log Scale” checkbox again will restore the default offset.

The Residual Correlations view shows the calculated correlation coefficient for library components which were not part of the decomposition with the residual ion current. The correlation coefficient ranges from +1 to -1. A value near +1 indicates a high degree of correlation and suggests that this component may be present in the residual spectrum. To test this possibility, just include the component in the Library Spectra view and run the decomposition again. The Residual Correlations tends to show false positives, particularly when library components are similar. The analyst will need to consider the likelihood of components actually being present when judging whether or not to include a component in the decomposition.

Components are selected for display in the Residual Correlations view depending on whether or not their correlation value has ever been among the topmost values for any scan. For each scan, the correlation values are sorted from highest to lowest. The component with the highest correlation value becomes the top 1 component for that scan. The component with the next highest correlation value becomes the top 2 component, and so forth. By default, only the components which have been ranked as a top 1 component in at least one scan are displayed. This can be changed by adjusting the spinner labeled “Show Top:” so that the top 2, 3, 4, and so forth components are displayed also. The graph rapidly becomes too crowded to be useful, but occasional some insight can be gained by knowing which components are almost significant enough to be displayed by default.