

GCWerks GC and GC/MS data

Copyright © 2016 GC Soft, Inc. All rights reserved.
GCWerks is a registered trademark of GC Soft, Inc.

Updated 27 February 2016

Contents

1 Introduction	1
1.1 Supported instruments and formats	1
1.2 Features	1
2 Stored chromatograms	3
3 Meta data	4
3.1 Inlet port plumbing (ports log)	4
3.2 Run types	4
3.3 Standards file	4
3.4 Flags	5
4 Identifying peaks	6
5 Integrating peaks	8
5.1 Integration parameters	10
5.2 Peak detection	10
5.3 Detection options	11
5.4 Baseline options	12
5.5 Peak options	13
5.6 Baseline detrending	14
6 Viewing results	15
6.1 Viewing all run types	15
6.2 Exporting data	15
6.3 Viewing the chromatogram corresponding to each point	16
6.4 Graphical flagging	17
6.5 Viewing filtered or flagged data	18
6.6 Changing plot scales	18
7 Calculating results	19
7.1 Y-axis variables	20

1 Introduction

GCWerks for GC and GC/MS are part of the GCWerks product family for greenhouse gas measurements, with the same software interface for GC, GC/MS and CRDS measurements.

This document provides an overview of GCWerks (*version 3.2*) GC and GC/MS data processing capabilities, whether importing existing data or acquiring data from the instrument. For more details and up-to-date documentation see <http://gcwerks.com>.



Figure 1: Advanced Global Atmospheric Gases Experiment (AGAGE) station at Ragged Point, Barbados.

1.1 Supported instruments and formats

GCWerks can import ChemStation data (or directly acquire data) from the Agilent 6890 and 7890 GC, and the 5973, 5975 and 5977 MSD. For GC or GC/MS data stored in other formats, GCWerks has the capability to import data in a generic ASCII format.

1.2 Features

- Rapid review of chromatograms and time-series results for many years of data.
- Easy batch integration of years of data, with integration parameters saved for each time period.
- Integrator accurate and precise over a wide range of peak sizes, requiring minimal changes in parameters over time.

- Determines results by both peak area and height, allowing the operator to report either and to compare them as a diagnostic.
- Extensive diagnostic plots to assess instrument and integrator performance.
- Calculates results based on periodic standard measurements.
- Nonlinearity and other diagnostics
- Graphical data flagging
- Flexible exporting of results
- A single multi-user Linux PC can be shared by a group for remote-desktop access to GCWerks data for many instruments.

2 Stored chromatograms

Chromatograms imported (from ChemStation files) or acquired directly by GCWerks from the instrument are stored in the same format, as compressed binary files, with one file for each chromatogram. For multi-detector GC data, there is one file for each detector. The figure below shows a GC/MS chromatogram with the list of chromatogram names displayed. Files are named by GMT date and time, with additional information appended to identify the sample (such as the inlet port number and run type):

yymmdd.hhmm.port - with inlet port number

yymmdd.hhmm.type.port - with run type and inlet port number

For chromatograms imported from ChemStation, there is typically sample identifying information in the "sample_name" field. If this information is consistent, it may be mapped into a form useful in GCWerks (see <http://gcwerks.com/index.php/documentation/identifying-samples>).



Figure 2: GC/MS chromatogram showing total ion (bottom) and individual SIM ions (above). The list of chromatograms is displayed.

3 Meta data

3.1 Inlet port plumbing (ports log)

For measurements at field sites, where the sample plumbed to each inlet port (air, standard, etc.) is fixed for long periods, the the ports log describes what is on each inlet for each time period. Lines are added to the ports log as calibration tanks or other changes are made, keeping a complete record of the measurements.

#	Date	Time	Port	Tank	Regulator	Type
#	-----	----	----	-----	-----	----
110101	0000	0000	1	90m	-	air
110101	0000	0000	2	90m	-	air
110101	0000	0000	3	50m	-	air
#						
110101	0000	0000	4	JB03331	-	cal
110101	0000	0000	5	JB03203	-	std
#						
130601	0000	0000	5	JB03205	-	std

3.2 Run types

The 'Type' column specifies the run type. Common types are:

- **air:** *ambient air*
- **std:** *periodic standard measurement used to correct for instrumental drift*
- **cal:** *additional calibration tanks, spanning a range of concentrations*
- **zero:** *zero gas (may be run as type tank also)*
- **blank:** *a system blank measurement*
- **tank:** *tank being calibrated*

The 'std' type is used to calibrate the data, with time-interpolation to correct for instrumental drift. Additional calibration tanks (type 'cal'), spanning a range of concentrations, may be used to characterize instrumental nonlinearity. GCWerks only requires a 'std' type to calculate results.

3.3 Standards file

Calibration tank concentrations are specified in the standards file. The header defines the columns in the file.

tank	co2	ch4
#		
JB03331	393.392	1858.17
JB03203	393.169	1871.4

3.4 *Flags*

Section 6.4 describes graphical data flagging, which is used to flag occasional outliers or groups of outliers.

For flagging very long time periods, months or years, it is better to enter the flag manually in the “range flags file”. This specifies the date range, inlet port number and species:

```
#
#date  time  date  time  port  gas  flag
#-----  ----  -----  ----  ----  ---  ----
120101 0000  120607 2000   4    all  F
```

This can also be useful for ongoing problems, by setting the ending date into the future, so that newly acquired data are automatically flagged (for the specified inlet port) until the problem is resolved.

4 Identifying peaks

The peak ID table is saved for each time period, normally several months or longer unless retention times change quickly. The times of saved peak ID files are displayed on the small lower plot in the figure below. The blue lines show all the saved peak ID files, while the orange line shows the peak ID table currently being edited. Click to the right of any line on the lower plot to view the corresponding peak ID table.

Display the peak ID table with **View > peakid windows**. Add peaks to the table by clicking on a peak and selecting **new peak id window** from the plot menu (right mouse button). Then fill in the information on the list at left. Times may be adjusted graphically by dragging the peak ID window with the middle mouse button. The figure below shows an example GC/MS peak ID table, including target and qualifier ions for each peak. For multi-detector GC, there is a separate peak ID table for each channel.

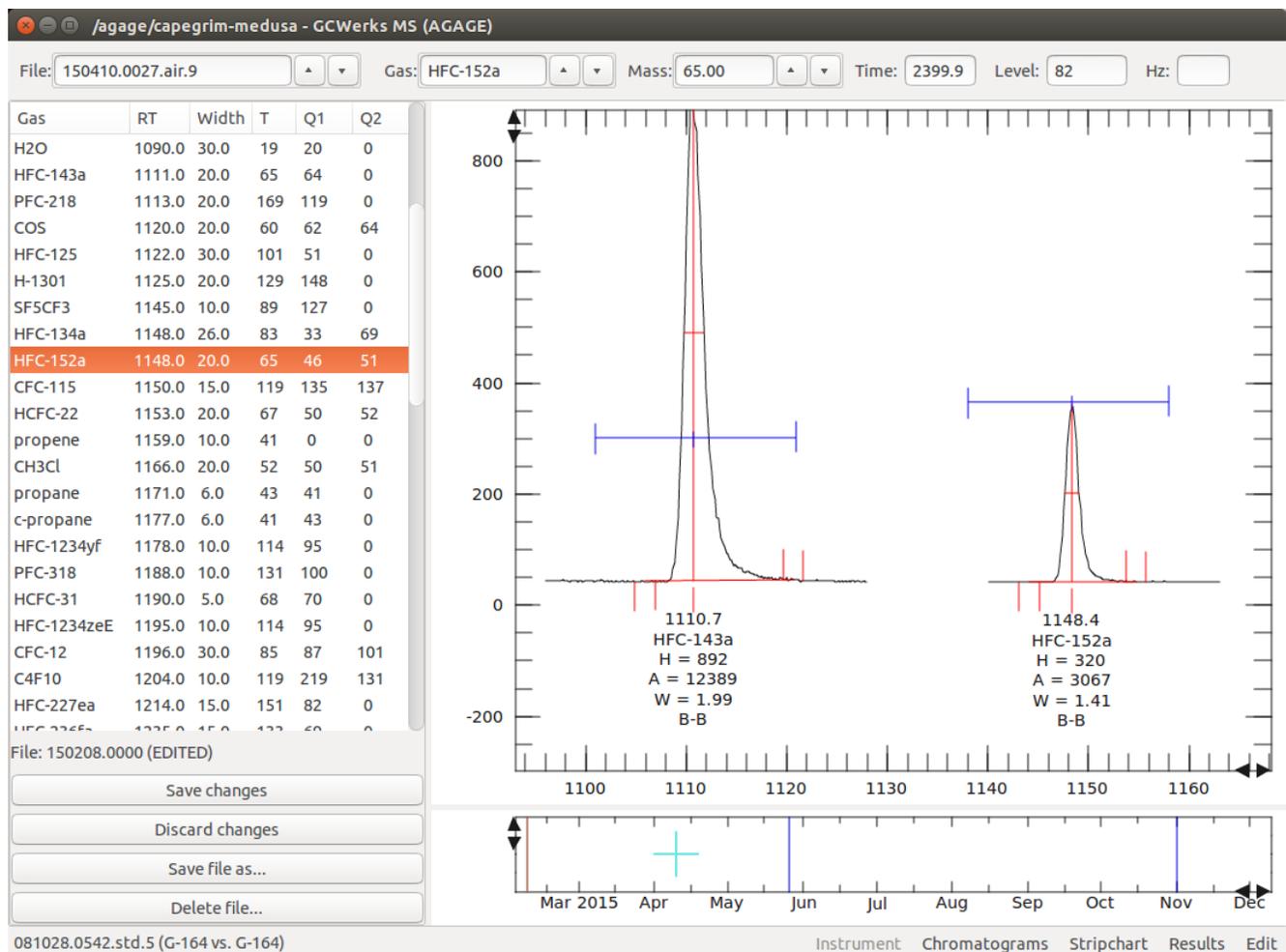


Figure 3: Peak ID table (left) saved as file 150208.0000 (orange line lower plot). This file applies from 150208.0000 until the date of the next peak ID file (blue lines lower plot).

To create a new peak ID file (when retention times change) press **Save file as...** This will copy the last peak ID file, creating a new one which applies starting at the current chromatogram.

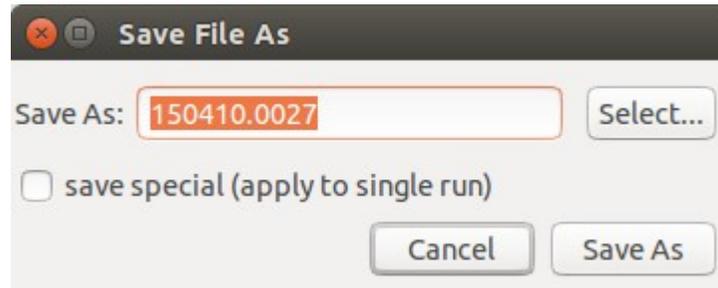


Figure 4: Creating a new peak ID file.

To save a peak ID file which applies only to a single run, click **save special (apply to single run)**. This allows creating special peak ID files for occasional runs where retention times shift, without having to create too many peak ID files.

5 Integrating peaks

Integration parameters saved for each time period, normally several months or longer unless the chromatography change quickly. The times of integration parameter files are displayed on the small lower plot in the figure below. The blue lines show all the saved integration parameter files, while the orange line shows the integration parameters currently being edited. Click to the right of any line on the lower plot to view the corresponding integration parameters. (For GC/MS, the grey lines show integration parameters which apply to different masses.)

Display the integration parameters with **View > integration parameters**. Parameters may be edited on the list at left, or added graphically on the chromatogram by clicking on the right mouse button and selecting **new paramter**. Times may be adjusted graphically by dragging the parameters with the middle mouse button. For GC/MS, integration parameters are set for each mass (the figure below shows parameters for m/z 65). For multi-detector GC, there are separate integration parameters for each channel.

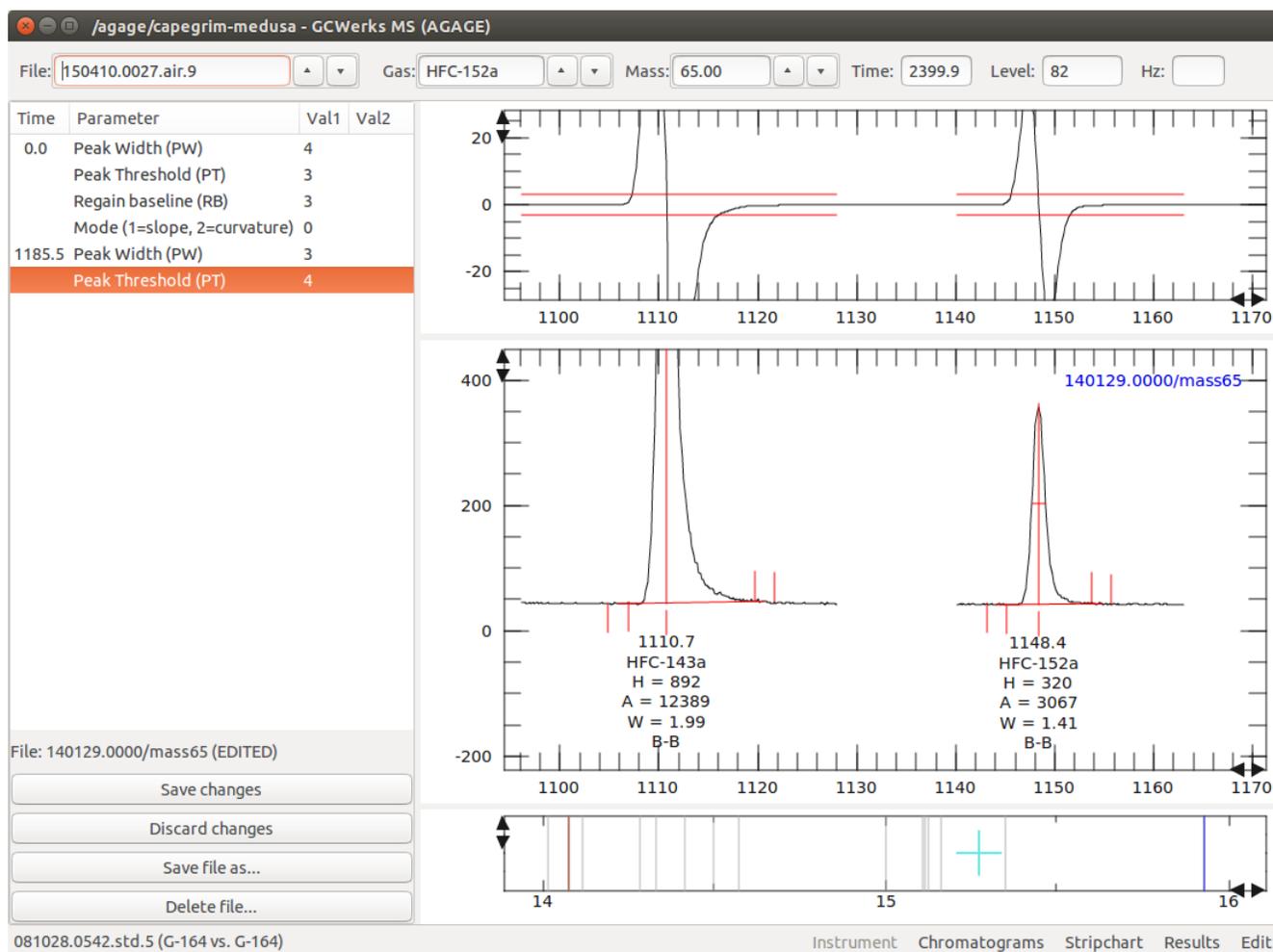


Figure 5: Integration parameters applying to m/z 65. The slope (first derivative) is plotted above.

To create a new integration parameters file press **Save file as...** This will copy the last integration parameters file, creating a new one which applies starting at the current chromatogram and for the current mass (the mass can be changed by editing the “Save As:” file name).

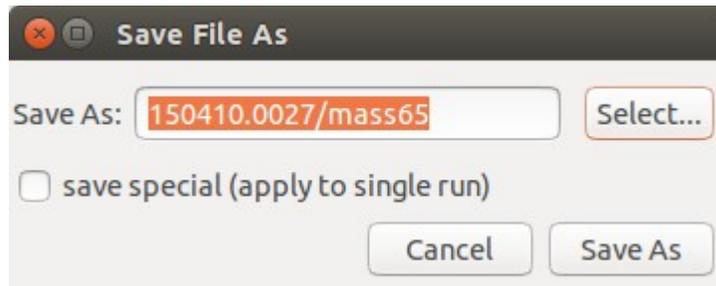


Figure 6: Creating a new parameters file.

To save an integration parameters file which applies only to a single run, click **save special (apply to single run)**. This allows creating special integration parameters file for occasional runs, without having to create too many integration parameters files.

5.1 Integration parameters

The integration parameters are divided in categories (Peak detection, Baseline options, etc), as they are grouped on the **new parameters** menu. Each parameter can take either one or two values. Most take only one value, unless otherwise specified. These are referred to as **N1** and **N2** below.

5.2 Peak detection

Peak detection parameters affect which peaks are detected. The most important of these are PW and PT. *Set PW and PT before setting any other parameters. Many quantities in the integrator scale to PW.*

Mode

The mode controls the type of peak detection: mode 1 = slope, mode 2 = curvature. Detection by curvature works better for shoulder peaks (especially shoulders high on the parent peak), or for sloping or curved baselines. The mode may be set differently for each integration file. (There is a mode 0 for a historical version of the integrator, done by slope. This should not be used for new projects.)

Peak Width (PW)

Peak width at half height in seconds. Set this to the approximate width at half-height of peaks in the chromatogram. PW determines the smoothing for the slope or curvature. Small peaks much narrower than PW are smoothed out as noise. *Increasing PW reduces noise in the slope plot, but makes it more difficult to detect small fused peaks.*

Default: PW 10 (GC), PW 2 (MS)

Peak Threshold (PT)

Peak slope or curvature detection threshold (red lines on slope plot). *Set PT to at least 1.5 to 2 times the noise in the slope plot.*

Optional second value 'a' or 'auto' allows the integrator to automatically determine the threshold. In this case the first value specifies a multiple of auto threshold ('PT 2 auto' means use 2 times the auto threshold). This is especially useful for small peaks where baseline noise is variable, and for blanks. *For the auto threshold to be accurately determined, at least one-third of the chromatogram must be baseline (without peaks).*

Default: PT 10

Blank Threshold (BT)

If specified, used in place of PT *for run type blank*. This is typically specified as a multiple of the auto threshold. A typical blank threshold is 'BT 2 auto', and the minimum auto threshold should be 'BT 1.5 auto'.

Default when set: BT 2 auto

Negative Threshold (NT)

If specified, used in place of PT for the negative slope threshold. It is turned off with 'NT 0'. *(This is rarely used. Better to detect peaks by curvature rather than have a different positive and negative peak threshold.)*

5.3 Detection options

Disable integration (DI)

Disables new peak detection, but finishes integrating any peak already detected. (1=on, 0=off)

Fused factor (FF)

Peaks are fused if the time between *the end of the first peak* and the *start of the second peak* is less than **$N1 \times PW$** . When FF is set to -1, no fused peaks are found. If the optional second value is 'c' or 'crest' then the peak crests must be within **$N1 \times PW$** seconds to be fused.

Default: FF 0.75 (if FF not set or set to 0, this default is used). FF is used differently for mode 0 and only takes value 0 or 1.

Disable fused (DF)

Disable fused peak detection. Results in a single peak in place of fused peaks (so a fused group would be a single peak). (1=on, 0=off)

Disable shoulder (DS)

Disable shoulder peak detection. Shoulders become part of the parent peak. (1=on, 0=off)

Note: Currently for mode 2 (curvature), shoulder and fused peaks are not distinguished.

Tailing rider filter (TF)

N1 specifies the minimum peak height of 'rider' peaks on a tailing parent peak. This is used to filter small (and variable) peaks or noise on tailing peaks, to prevent early termination of the peak. The height is measured from the tailing baseline at the start of the rider peak. *The parent peak height must be at least $4 \times N1$. The rider peak must be fused to the parent to be filtered.*

5.4 Baseline options

Baseline options affect how the baseline is drawn between the start and the end of a peak or group of fused peaks. Two types of 'baseline points' referred to below are 'fused' or 'resolved'. A group of fused peaks has resolved points at the start and end, with fused points in between.

Note: The peak starting and ending baseline levels are determined by averaging $PW/4$ of the baseline ($PW/2$ for MS). Fused points use one-half this width.

Valley-to-valley (VV)

Draw valley-to-valley baselines between fused points. (1=on, 0=off)

Forward horizontal (FH)

Draw a forward horizontal baseline *from the starting resolved point*. (1=on, 0=off)

Backward horizontal (BH)

Draw a backward horizontal baseline *from the ending resolved point*. (1=on, 0=off)

Minimum horizontal baseline (MH)

Draw a forward horizontal or backward horizontal baseline from the minimum level of the starting or ending resolved baseline point. (1=on, 0=off)

Horizontal baseline (HH)

Draw a horizontal baseline using the mean level of the starting and ending resolved baseline points. (1=on, 0=off)

Minimum baseline (MB)

If **N1** is not 0, sets the peak start at the minimum baseline level within **N1 x PW** before the peak. If **N2** is not 0 sets the peak end at the minimum baseline level within **N2 x PW** after the peak.

Baseline width (BW)

If set, the width of baseline to average for the peak starting and ending levels is **N1 x PW/4**. If **N2** is set it is used for the ending level.

Default: BW 1 (GC), BW 2 (MS)

5.5 Peak options

Most peak options take two values, **N1** and **N2**. The first 4 of these options change the time of the *detected* peak start or peak end.

Tangent skim (TS)

If **N1** is not 0, finds the baseline to be tangent to the chromatogram at the peak start and end. The tangent start is found within **N1 x PW** before the peak, and the tangent end found within **N2 x PW** after the peak. This does not work well for very noisy baselines, but can be improved by increasing *BW*. *Tangent skim works for a concave up chromatogram baseline (tangent points will not be found where the chromatogram baseline is concave down).*

Default when set: TS 0.5 2 (MS), TS 2 2 (GC)

Regain baseline (RB)

This option is useful for tailing peaks, where the end of the peak is found high on the peak tail. The peak end is moved forward until the baseline level falls to within **N1** counts of the starting *resolved* baseline level.

Peak start (PS)

Move peak start back **N1** seconds from the *detected peak start*. If the optional second value is 'c' or 'crest' this is relative to the *peak crest*.

Peak end (PE)

Move peak end forward **N1** seconds from the *detected peak end*. If the optional second value is 'c' or 'crest' this is relative to the *peak crest*.

Manual peak start (MS)

Manually set a new peak to start at the time of the MS parameter. **N1** specifies the width of the peak (at the baseline). *This manual integration should be used with caution, in rare cases when small noisy peaks cannot be detected using the peak threshold.*

Peak fit (PF)

If set, a Gaussian peak fit is used to determine *peak height* for sparsely-spaced MS data. The top portion of the peak is fit, with **N1** specifying the fraction of the way up the peak to begin the fit, and **N2** specifying the fraction of the way down the peak to end the fit. This is used to determine *both* the peak height and the retention time.

Default: PF 0.4 0.8 (MS), PF 0 (off, for GC)

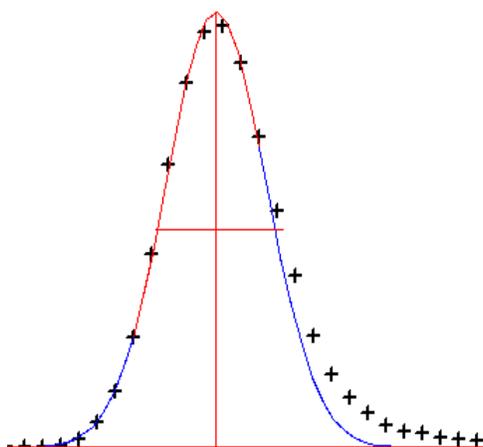


Figure 7: Example PF 0.4 0.8

5.6 Baseline detrending

Baseline detrending is typically used to detrend chromatograms with large baseline slope, where small peaks are not properly detected by slope. It is generally better to use *mode 2* to detect peaks by curvature, rather than detrending.

Baseline point (BP)

Specifies a point on the baseline to use for detrending. Detrending is done by connecting these points with line segments whenever 2 or more points are set.

6 Viewing results

Integration results (peak areas, heights, widths, etc.) and calculated results may be viewed in a number of ways. Press on the **Results** button on the lower right to view results.

6.1 Viewing all run types

Select run type “all” to display data for all run types (air, std, cal, tank, etc.) simultaneously in different colors. The figure below shows peak areas for air and std measurements.

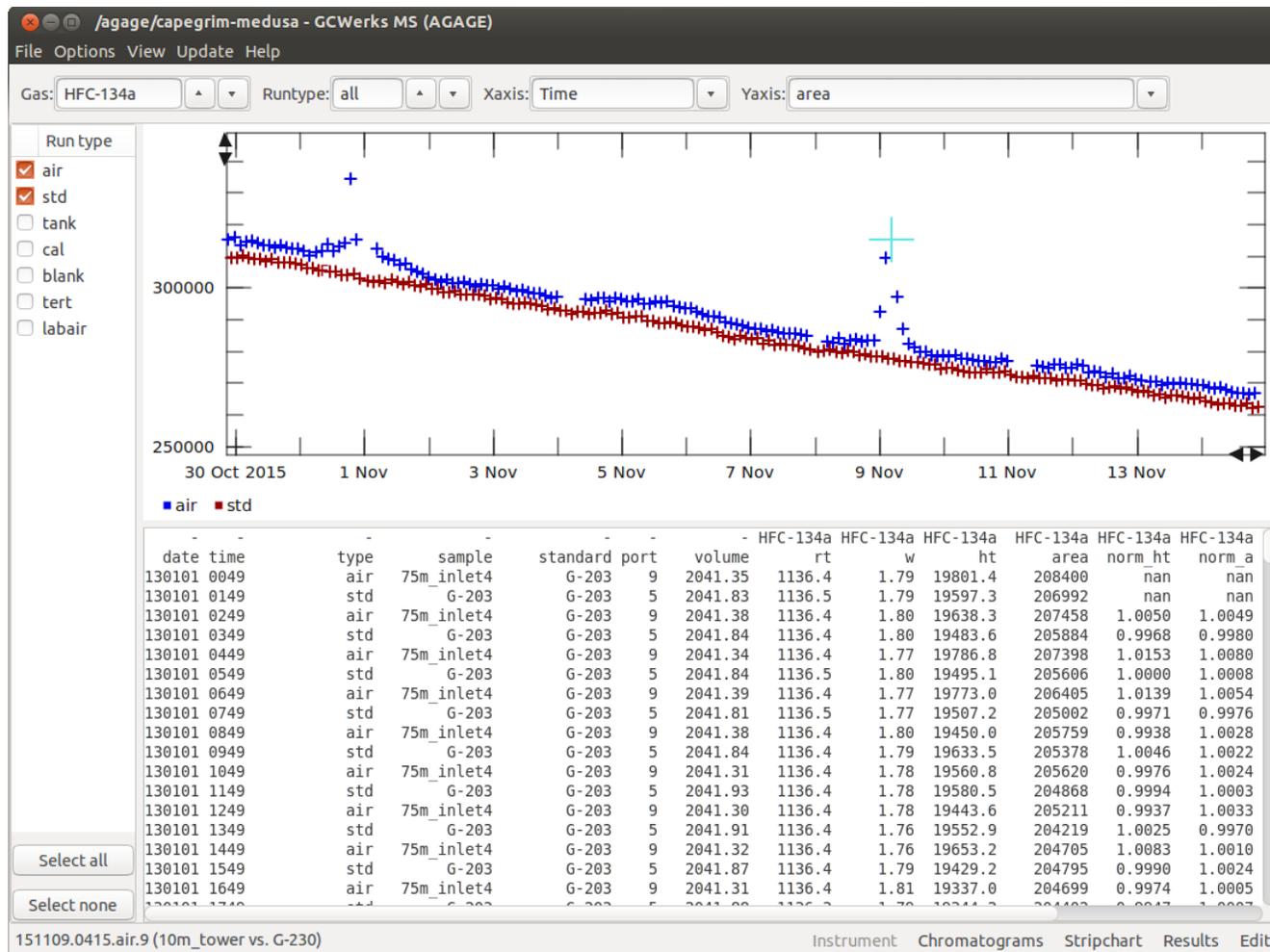


Figure 8: Plot of peak areas (above) for types air and std, and corresponding report (below).

6.2 Exporting data

The report corresponding to the plotted data is displayed with **View > report**. The listed data may be saved to a file with **File > Saveas....** To automate exporting or for more control over exported columns and data, see *exporting results* on the GCWerks website documentation.

6.3 Viewing the chromatogram corresponding to each point

Display the chromatogram below the time series plot with **View > chromatogram**. Click on any point on the time series to see the corresponding chromatogram. Hold down the middle mouse button and drag the cursor across the time series to review chromatograms quickly.

The figure below shows calculated concentrations for run type 'air'. When viewing a single run type (air, std, etc.) the color changes when the inlet port changes or when the std tank used to calculate the results changes.

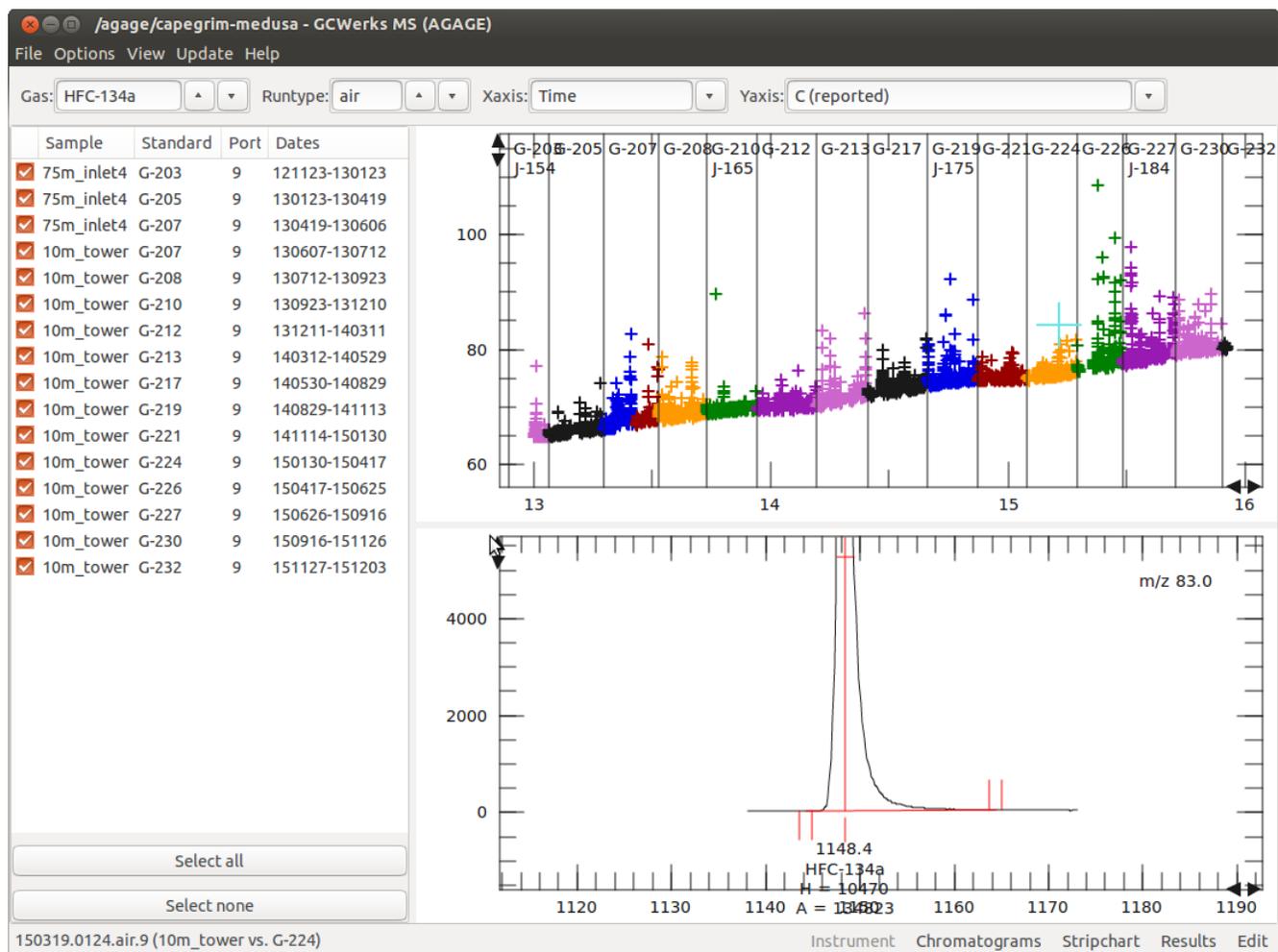


Figure 9: Calculated concentrations (above) and chromatogram (below) corresponding to the cursor in the top plot.

6.4 Graphical flagging

To select a single point to flag, hold down the SHIFT key and click on the point. Click on the point again to un-select it. To flag a group of points, hold down the SHIFT key and use the zoom box as shown in the figure below. To clear all selected points, press **clear**.

The default flag is the 'F' flag, which flags a point as a flyer to exclude from the calculated results. The 'U' flag may be used to *unflag* a point. Press **apply flags** to flag the currently selected points for the displayed gas.

After applying flags in this way, the the **update** menu to calculated results with the new flags applied.

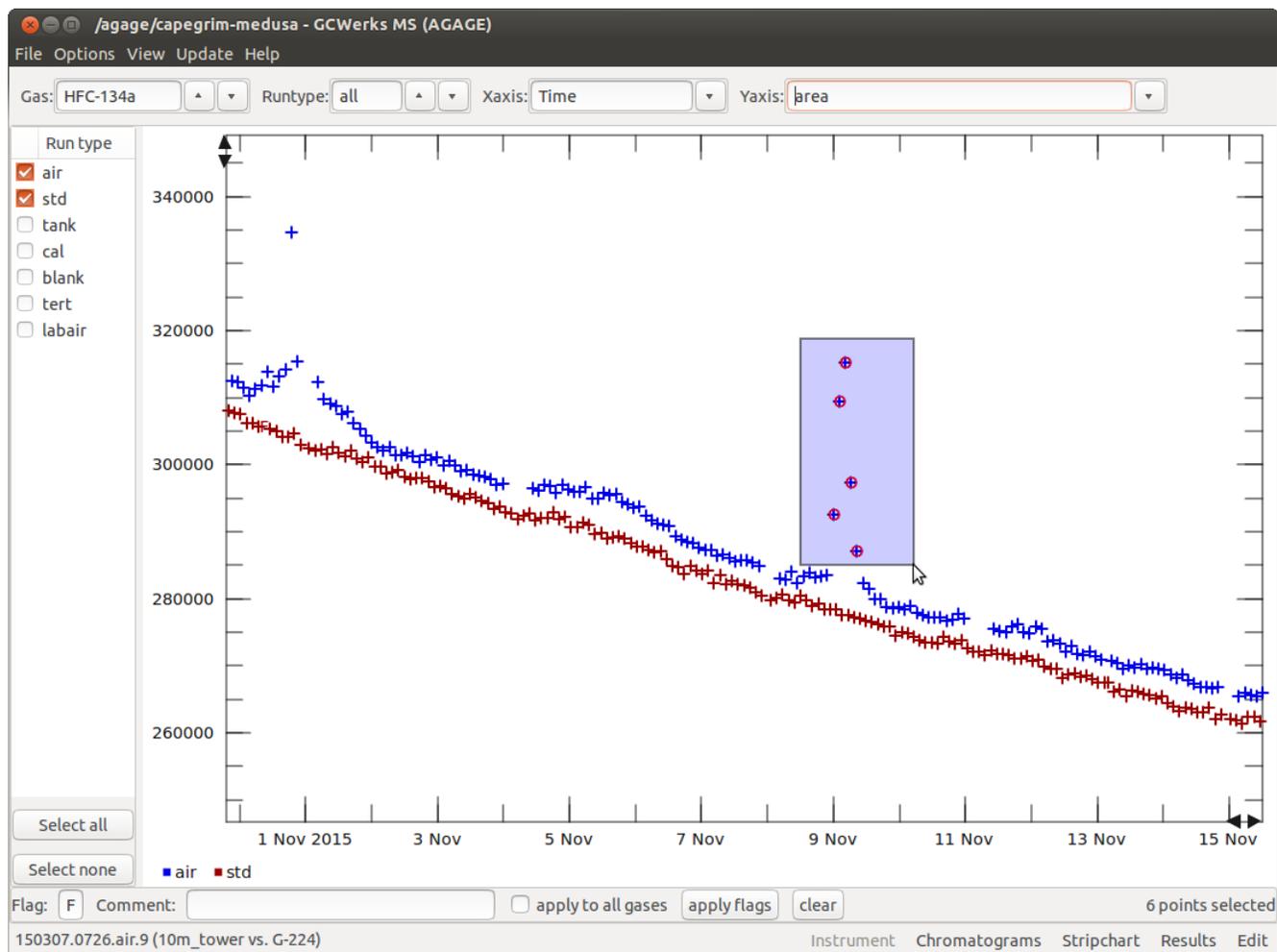


Figure 10: SHIF-zoom box to select points to flag graphically.

6.5 Viewing filtered or flagged data

Two options control whether flagged or filtered data are displayed and included in the self-scaling:

Options > show flagged points ('F')

options > show filtered points ('*')

6.6 Changing plot scales

Zoom into a region by dragging the cursor on the time-series plot (left mouse button). To zoom in on the time-axis only, allowing the y-axis to self-scale, select **x-zoom only** from the plot menu (right mouse button). This is useful as the y-axis will always scale properly for any displayed time period, gas, or y-axis variable.

The left/right and up/down arrow buttons on the axes may also be used to change the the plot scales. Press the left mouse button on an arrow to move the view. Press the right mouse button to compress or stretch the scale.

7 Calculating results

The **update** menu may be used to select a time period to batch integrate and calculate. Results may also be updated from a script which runs the `gcupdate` (integrate) and `gccalc` (calculate) programs (see <http://gcwerks.com>).

The figure below shows integrated peak areas (top) and calculated sample/standard area ratios (normalized area) below. The normalized area is determined by time interpolation to the peak area of the adjacent std runs. This is determined for the std runs also, as a measure of precision. In the figure, the std is measured every other run, but GCWerks can handle different spacing, for example: `std air air air std air air air std`.

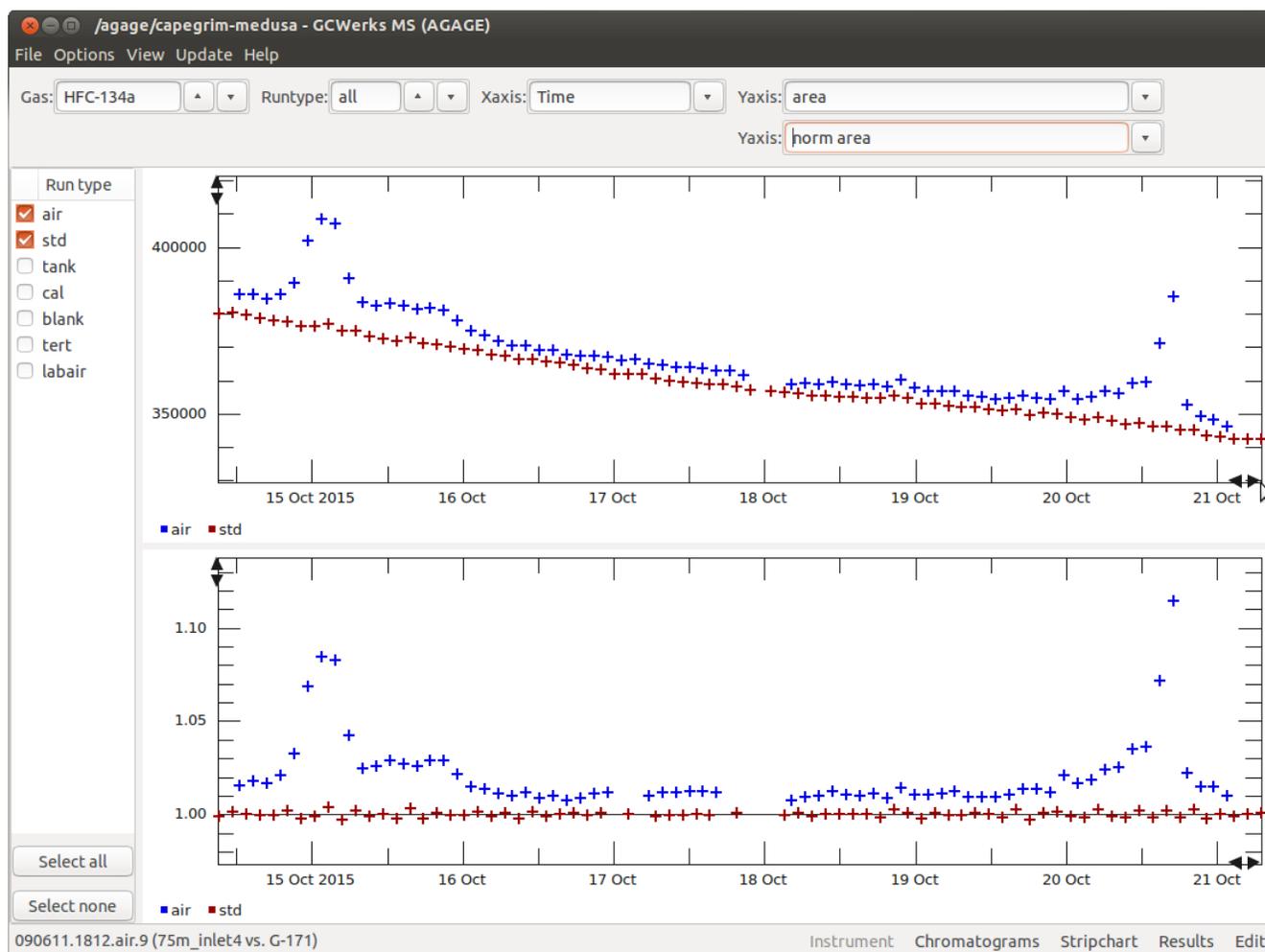


Figure 11: Peak areas for air and std (top) and normalized area (bottom).

7.1 Y-axis variables

Any of the integrated or calculated results may be plotted on the y-axis. These include:

Integration results:

- **area**
- **height**
- **width**
- **skew**
- **retention time (rt)**

Drift-corrected integration results (ratio to the time-interpolated adjacent std runs):

- **norm area**
- **norm height**
- **norm width**
- **norm height / norm area**

For GC/MS, qualifier / target ratios may be plotted (for Q1 or Q2):

- **Q1 / target rt**
- **Q1 / target area**
- **Q1 / target height**
- **Q1 / target width**

A future update will add Q1/target normalized area and Q1/target normalized height, so the drift-corrected (calibrated) results may be compared for qualifier and target.

Calculated results:

- **C (reported)**
- **C (area)**
- **C (height)**