Manual Addendum

DIFFRAC.EVA V7.0

This document contains information and corrections, which were not available at time of print.

Contents

Import Scan Behavior (V2.1 and later versions)

The import of multiple measurement files works differently in DIFFRAC.EVA V2.1 and later versions than in former versions. An import of multiple measurement files at once using the Import from Files/Database dialog results in the following situation, dependent on the content of the brml/raw files:

Type of the imported files	Resulting scan lists in DIFFRAC.EVA V2.0	Resulting scan lists in DIFFRAC.EVA V2.1 and later
Files containing only single scans	- One scan list containing all scans which were measured with the same scan axis - Multiple scan lists if multiple scan axes were used	- The same behavior as in DIFFRAC.EVA V2.0
At least one file contains multiple scans	- One scan list for every multiple scan file - One scan list for all single scan files with the same scan axis	- One scan list for all scans with the same scan axis regardless of if the files contained single or multiple scans

The import into an already open document reacts sensitive to the selection in the tree. A subsequent import while a scan list or a child of the scan list is selected inserts the imported scans into the scan list if the scan axis matches. If another node is selected (e.g. the Document node), the import command will create new scan lists.

Additions and Corrections to the Manuals (V4.0)

User Manual

D x by and Tune Cell Commands are also Performed on the Replaced or Cloned Patterns (V4.0)

Starting with V4.0, the *D x by* and *Tune Cell* commands will be performed on the replaced or cloned patterns, if the original patterns had the same commands applied.

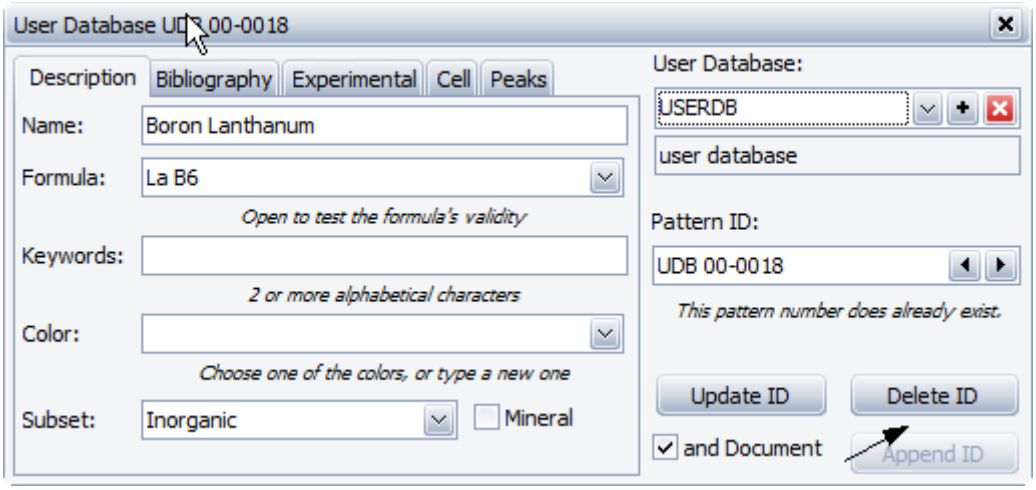
DIFFRAC.EVA User Database Conversion if Variable Slits Were Used (V4.0)

Until DIFFRAC.EVA V4.0, user patterns created on scans displayed for variable slit, either measured or with simulated variable slits, have wrong relative intensities.

This bug has been corrected with DIFFRAC.EVA V4.0, but it cannot be fixed for existing wrong user patterns. The solution is to delete such patterns and to recreate them. This is a quick operation. Here is the detailed procedure:

(1) Import the scan corresponding to such a user pattern. Use the tool **Search by number** on this scan, select the user database in which this pattern has been stored and enter its pattern number.

(2) Once this pattern is displayed, launch the **User database** tool on it, and then click “Delete ID” from this tool.



(3) At this step, you may import the DIF that has been used to create this pattern if saved. If not, use the peak search to create a new DIF from scratch, and then launch the “User database” tool. Select the same database. The pattern ID (number) which will be proposed is most likely the one of the previously deleted pattern. Create it with this ID or another one, if you prefer.

Repeat this procedure for all user patterns with wrong relative intensities.

Cluster Analysis Manual

Potential Problems (page 12)

The analysis with multiple datasets is continued without warning if the dataset number does not match.

Label Extraction Options (page 22)

Some examples for the label extraction with regular expressions follow:

Example 1

Original labels:

```
FORM A1.raw
FORM A2.raw
FORM B3.raw
FORM C4.raw
```

if filtered by RE pattern: [A-C].+

```
A1.raw
A2.raw
B3.raw
C4.raw
```

if filtered by RE pattern: [^.]+

```
FORM A1
FORM A2
FORM B3
FORM C4
```

if filtered by RE pattern: [A-C].

```
A1
A2
B3
C4
```

if filtered by RE pattern: [^B]+B.+

```
FORM B3.raw
```

Example 2

Original labels:

```
Indonisia_Ore_5per_scale_B.raw
Indonisia_Ore_10per_scale_B.raw
Indonisia_Ore_20per_scale_B.raw
Indonisia_Ore_40per_scale_B.raw
```

if filtered by RE pattern: [0-9]+per

```
5per
10per
20per
40per
```

if filtered by RE pattern: [0-9]+[^.]+

```
5per_scale_B
10per_scale_B
20per_scale_B
40per_scale_B
```

Scan Colors in Cluster Analysis Views

The scan colors in Cluster Analysis may be different from the colors which are used in 1D Views. This is due to an independent coloring algorithm used for Cluster Analysis.

If the scans are transferred from a scan list into a Cluster Analysis set (*e.g.* by using the *Run Cluster Analysis* command) then the colors stay the same.

Mouse Action Missing in the Description for the 3D Views

Ctrl + Mouse drag changes the sphere's size.

Scree Plot

The cluster analysis has 15 different ways of estimating the number of clusters. Two of these are based on principal components analysis (PCA). One PCA method uses the non-transformed matrix, the other method uses the transformed matrix. The amorphous samples are included. The PCA results using the transformed matrix are used to generate the Scree Plot, so this plot ONLY relates to the PCA method not to the final estimate of the number of clusters. It's included because it's a standard statistical method.
As a result, the number of clusters shown in the Scree Plot may differ from the number of clusters in the text output.

Additions and Corrections to the Manuals (V4.1)

User Manual

Properties Page

The Properties page has three new settings for scans in the group *Wavelength*:

Anode, ka1, ka2, ka2 Ratio, kb and *Wavelength for display:*

If these properties are set any new scan loaded from text files which does not contain an anode material in the header will be filled with these values.

This is useful e.g., for synchrotron measurements. If the anode material is set to a character string which is not an Element symbol the entered wavelengths are not overridden.

New and Changed Scan Properties

New properties

If the detector is a LYNXEYE XE-T the new property “Detector Profile” is displayed.

The new property “Air Scatter Slit” shows whether an air scatter slit was used.
The new property “Air Scatter Slit Mode” shows whether the mode was automatic or fixed.

Cluster Analysis Manual

The computer code for generating the 3-d plots (the PCA and MMDS representations of the data) has been completely re-written.

DIFFRAC.EVA V4.1 generates 3-d plots that are more accurate, and scale better as the number of samples increases. As a consequence, the 3D plots may look different than in the manual.

There is also increased parallelization of the code for large data sets.

Additions and Corrections to the Manuals (V4.2)

Cluster Analysis Manual

Cell Display

The behavior of the cell display is different, depending on if a clustering was carried out without or with reference samples.

Without Reference Samples

The cell display visualizes the cluster membership. Consequently, it will be updated after the cut line was moved.

With reference Samples

The cell display visualizes the calculated composition. Moving the cut line will not affect this.

Run Cluster Analysis Dialog

The option “Hide results similar to references for cell display” does not affect views besides the cell display and the dendrogram. The affected results are only suppressed in the cell display. The dendrogram moves these results to the side.

Additions and Corrections to the Manuals (V4.3)

User Manual

Creating a Database Filter (Addendum to Chapter 7.3.1.2)

The algorithm for determining the sub-files from the search database has been improved in V4.3.

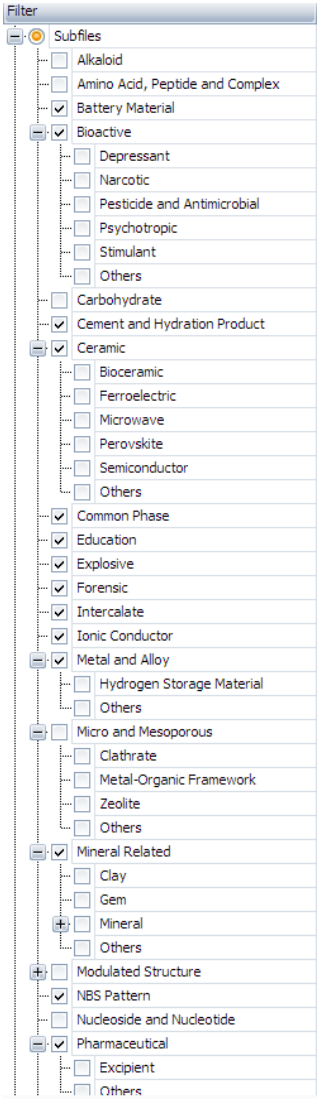
Using the categories stored in the database, a sub-file tree is automatically generated.

A new radio button was introduced to quickly switch between

- Full Files: Inorganic, organic, and dual, the latter belonging to both inorganic and organic according to the database entries. If all three are selected, the full database is searched.
- Sub-Files: The former sub-files are now grouped with parent-child dependencies. “Mineral related” and other sub-files belong to this sub.

The checks are inclusive. If only the radio button for sub-files is selected - the full files are thus ignored - and no sub-files are selected below, no patterns will be found. At least one full file or sub file must be selected to find patterns.

The current parent-child relationships from the PDF-4 database Release 2018 are displayed on the following picture:



Automatic Views (Addendum to chapter 4.5.3)

Starting with version 4.3, a feature called “Automatic Views” was introduced. It means that depending on the data selection corresponding views are displayed automatically in the “Automatic Views Panel”.

Automatic views cannot be printed. If printing is required, a double click on the Automatic View’s tab creates a corresponding persistent view, which can be printed in the usual manner.

If the Automatic Views Panel is not displayed by default intentionally (switched off in the View menu), it will be displayed automatically for holding the results from a pattern search. After the search tool has been closed manually, the Automatic Views Panel will be closed as well, but automatically.

On computers with low memory, the automatic views may cause memory full exceptions. If this happens, please check the entry “Disable some Automatic Views to save memory” in the Settings dialog on the General page.

Performing a Peak Fit (Chapter 7.7)

This chapter is not valid for version 4.3.

Recording a Workflow for User Guidance (Chapter 13)

This chapter is not valid for version 4.3.

Exceptions in 21 CFR Part 11 Mode

The following commands invalidate the controlled state of a document and will disable the save (to database) command.

- Import measurements from file
- Import XRF results from file
- Import frame integration cursors from file

Appendix: Computing Areas (Addendum to Appendix 17.9)

The left and right limits are always assumed to be background points, and a straight background is computed between the two limits. This is done, because EVA’s curved background is somewhat sensitive to outlier points, thus the reproducibility is better this way.

Let n0 be the index of the first point of interest in the range, n1 the index of the last one. Let nB = (n1 – n0 – 1) / 10 or 3, whichever is less; the left background is defined by

$xbl = (\sum_{i=n0-nB}^{n0+nB} m_i.X)$ / $(2 * nB + 1)$, $ybl = (\sum_{i=n0-nB}^{n0+nB} m_i.R\ /m_i.)$ / $(2 * nB + 1)$

where $m_i = i^{th}$ measured point in range. X = X coordinate (usually 2θ) and R = Y/T = rate (Cps). The same calculation is carried out for the right background. The actual range of left or right backgrounds is also restricted in case in case the ROI is at the beginning or at the end of the measurement range. The net rate at any point k in the region of interest is thus $N_k = m_k.R - (ybl + \frac{m_k.X - xbl}{xbr - xbl} * (ybr - ybl))$

The **raw area** is the integral of $m_k.R$ over the ROI computed by the trapeze method, using a linear interpolation for the end points, since the limits of the ROI generally do not fall on exact coordinates of measured points.

The **net area** is the integral of N_k over the ROI computed by the same integration method.

The **center of gravity** is (integral of $N_k * m_k.X$) / (net area), using the same integration method.

The traditional trapeze method is used, because other optimized integration techniques work only for functions that can be calculated at any point and most of the time implicitly require that the function be continuous or even continuously derivable. The measurement is given by a table and entails a random counting statistics, thus its continuity is not known, not to speak of the derivatives, which are dominated by the noise. The difference between trapezes and other methods is a second order difference, which is small compared to the statistical noise, and the number of points is so small (a few

tens or hundreds at maximum) that the addition involved in the trapeze method is actually faster than the more complicated calculations needed by other methods. The method does not assume either constant step size or constant time.

The **raw height** is obtained by fitting a parabola across all the points with a net height > 0.75 net maximum

The **net height** is raw height – interpolated background at the calculated peak position.

The **integral width** is (net area) / (net height)

The **FWHM** is obtained by linearly fitting the two points at 0.5 net height, and subtracting the left value from the right value.

The **chord center** is the average of the two half height points.

The **Scherrer crystallite size** is derived from the Scherrer formula using either the FWHM or the integral width.

Additions and Corrections to the Manuals (V5.0)

User Manual

Quantitative Result Export to the Database

During a save to database operation, quantitative results that are contained in the data are exported to the compound table automatically.

Patterns have an additional property “DB Compound Name” which can be freely configured to represent the compound in the Results Manager.

Performing a Peak Fit (Chapter 7.7)

Refresh button removed

Pattern fit added

Additions and Corrections to the Manuals (V5.1)

User Manual

Drag and Drop in the Data Tree (Chapter 5.8)

It is possible to select a group of similar objects and use them for drag and drop operations.

Two different icons are displayed during drag and drop operations depending on the drop result: the straight green arrow to the right means that the dragged element will be added to the end of the destination list. The curved blue arrow means that the element will be inserted above the current line.

Additions and Corrections to the Manuals (V6.0)

EVA Settings

There is a new check at the bottom of the Settings Dialog “Save layout in document” which allows storing the layout of EVA’s panels (Data Tree, Properties, etc.) in the document. If this is selected, opening a document will restore the panel layout to the version which was active when the document was saved.

Chemical Balance: Comparison with a Chemical Analysis

The XRF database import was extended to allow importing XRF data from recent SPECTRA*plus* databases (*.accdb) and from SPECTRA.ELEMENTS databases.

SPECTRA*plus* databases

EVA V6 32bit can read XRF data from SPECTRA*plus* databases in MDB and ACCDB formats.

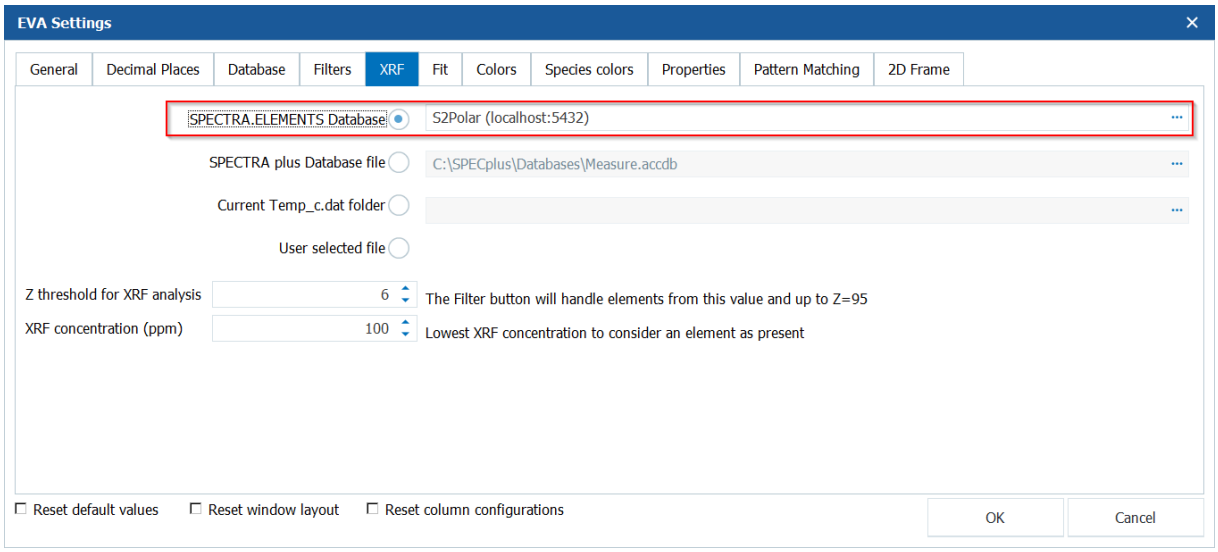
To read the data, the 32bit or 64bit Microsoft Access driver from 2010 or 2016 (for ACCDB) must be installed.

EVA V6 64bit can read XRF data from SPECTRA*plus* databases in ACCDB format only. To read the data, the 64bit Microsoft Access driver from 2016 must be installed.

Remark: An installed Microsoft Office limits the installation of the database drivers to its installed variant (either 32bit or 64bit).

SPECTRA.ELEMENTS databases

EVA V6 can read XRF data from SPECTRA.ELEMENTS databases. To configure a SPECTRA.ELEMENTS database, use the new entry in the Settings dialog (1st line):

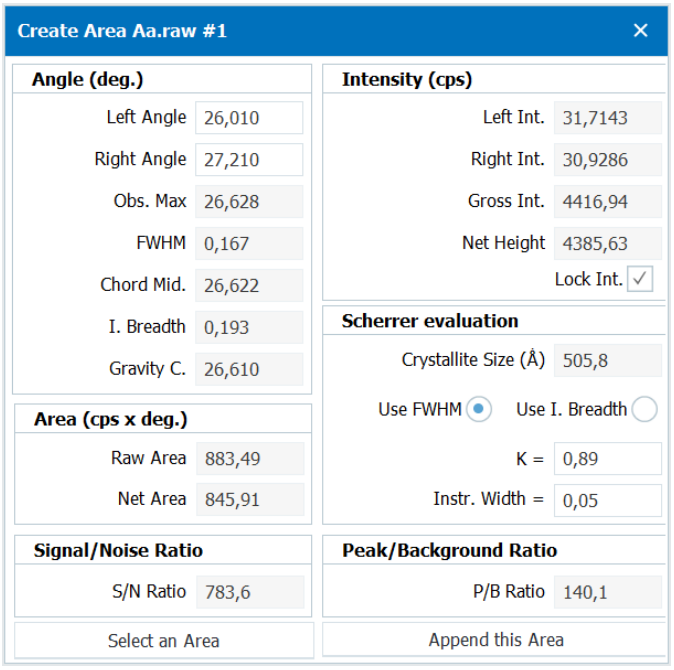


Aberrant Tool

After removing aberrant points from a scan, the automatic background curvature and the result of peak search may be modified. This is because both parameters depend on the calculated bandwidth of the signal. The aberrant points typically resemble a high intensity high frequency noise, artificially inflating the frequency content of the scan. This results in too low curvature and too many “ghost” peaks. In most cases the elimination of the aberrant data improves the fit of the background and the peak search. The automatic search/match algorithm uses the background and peaks and thus depends on the calculated signal bandwidth too.

Area Tool

The area too was extended with the peak-to-background ratio and the signal-to-noise ratio:



The peak-to-background ratio is:

```
Math.Abs(NetHeight / YP)
```

Where YP is the background height below the peak position.

The signal-to-noise ratio is calculated for a one-sigma confidence interval:

```
Math.Abs(NetHeight) / (0.5 * (Math.Sqrt(Math.Abs(Y0)) +
Math.Sqrt(Math.Abs(Y1))))
```

Y0 and Y1 are the area’s left and right background positions. The confidence interval can be configured in the area’s properties:

Crystallite Size	50,7 nm
S/N Ratio	719,9
Sigma (S/N Ratio)	1
P/B Ratio	118,3

Automatic Search/Match

The FOM (Figure Of Merit) Value

The FOM is a measure of the likelihood a phase is present, *kind of*. In 'traditional' search, the results are sorted by decreasing FOM, but in automatic mode, the phases with best FOM are evaluated for their contribution to the diagram, and it is possible that a phase with a comparatively low FOM contributes more to the observed intensity than one with a higher FOM, this does happen in many quite often and is no cause for concern.

Compiling the Reference Databases using the DSRD Compiler

Installing and Updating the COD (Crystallography Open database)

BRUKER provides a pre-compiled full COD database on the installation media and for download from Brukersupport.com. At the time of writing, the complete COD has a size of about 21 GB and contains more than 474,000 entries.

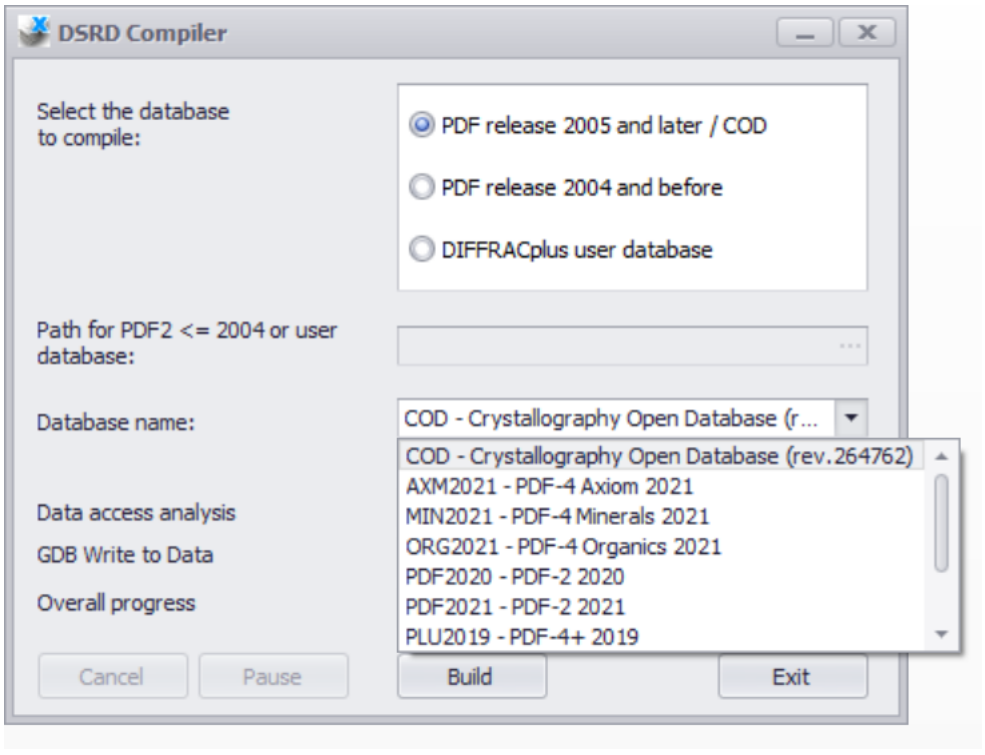
Installing the COD from installation media

The COD files are in the COD folder on the USB stick. There are five files whose names begin with "BrukerAXS.COD-", followed by a version number and a specifier like "Data", "Index", "Search", "Duplicates" or "Cif" and the extension "GDB". These files must be copied to the DSRD6 folder. In case this folder was not changed during installation, the batch file "InstallCOD.cmd" can be used for installation.

EVA will recognize the COD after the next start.

Updating the COD with the DSRD Compiler

The DSRD compiler can be used for an in-place update of the installed COD. When a COD is recognised the database selection list contains an entry for the COD. Its revision is mentioned in brackets:



By clicking the “Build” button the COD update process is started. In a first step, the version information is downloaded from the COD web site. This may take several minutes, depending on the connection speed. In a second step, the versions are analyzed, and it is determined, which entries needs to be updated. The final step adds all new and modified entries to the compiled COD database. Depending on the database’s age and the computer’s calculation power, it may take less than an hour until several hours to update the COD. Mind that an internet connection is required during updating.

After update the new database has a higher version number than the original database.

Remark 1:
The time to update the COD is dependent on the age of the installed COD. While a recent COD may be updated in minutes, a several months old version may take hours to update. The first update step is always a download of the full COD’s phase information (about 300MB). The time to download contributes to the full update time.

Remark 2:

Do not try to update if the file with the specification “Cif” in the file name is missing. In that case an update may succeed but will take several days while the computer is blocked for other tasks.

COD Minerals

A small database subset, the COD Minerals, is installed with EVA. No extra download or compilation is required. At the time of writing, it contains more than 13800 entries.

SQUALL

Introduction

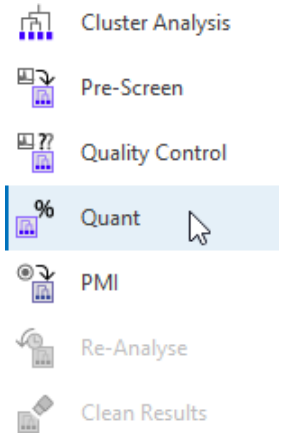
SQUALL (**S**emi-**QU**antitative Analysis for **ALL** patterns) is a quantitative analysis program for use in analyzing the contents of mixtures using full-profile PXRD data in the DIFFRAC.EVA software package since license level 5. It also works equally well with IR, NIR, Raman and ED-XRF data. It is simple to use, fast and uses few computing resources. It has two modes of operation that are interlinked:

- Qualitative analysis: it searches a collection of full profile 1-d reference patterns (scans) and eliminates them in turn to get a subset of reflections that describe the mixture contents. It then performs semi-quantitative analysis.
- Semi-quantitative analysis: it produces estimates of the components of mixtures.

SQUALL aims at achieving a best match of the sum of appropriately scaled reference patterns versus a sample pattern. This methodology offers several advantages:

- Being a full-pattern method, it is less sensitive to preferred orientation
- SQUALL makes no difference between crystalline and amorphous references
- Peak overlap of the sample pattern is intrinsically treated by the individually measured references
- As SQUALL references are per-se measured they do not require an analytical model for the calculation of the sub-profiles of the constituents of the mixture. This may make SQUALL attractive for analyzing materials with only partly or even unknown crystal structure, or materials that are notorious for disorder such as soil or clays, that are hard to model otherwise.

Throughout this manual the term SQUALL is synonymously used for Quantitative Pattern Matching. SQUALL is part of the Pattern Matching module in EVA and accessed via the Quant option in the Tools submenu of the EVA Data Tree:



SQUALL is semi-quantitative in the sense that it uses full measured profiles, but it is not a replacement for quantitative multi-phase analysis using the TOPAS Rietveld software, although it can give comparable results. In its basic form it needs only measured profile data for the samples and references, the maximum and minimum number of components expected in the mixture and the minimum % for any component. Several methods are provided to convert the scale of the individual reference patterns into wt-%. Some of those methods like Reference Intensity Ratios (RIR) or the

Direct Derivation Method (DDM) are also used in the literature in the context of semi-quantitative analysis.

In many situations, especially in complicated mixtures, SQUALL may find more than one possible solution, and these will be ranked according to figures of merit that evaluate the agreement between observed and calculated profiles. Alternatively, a unique solution can be forced for cases where the phase composition is well known from previous species- (a.k.a. phase-) identification, or in the DIFFRAC.EVA terminology a previous SEARCH/MATCH.

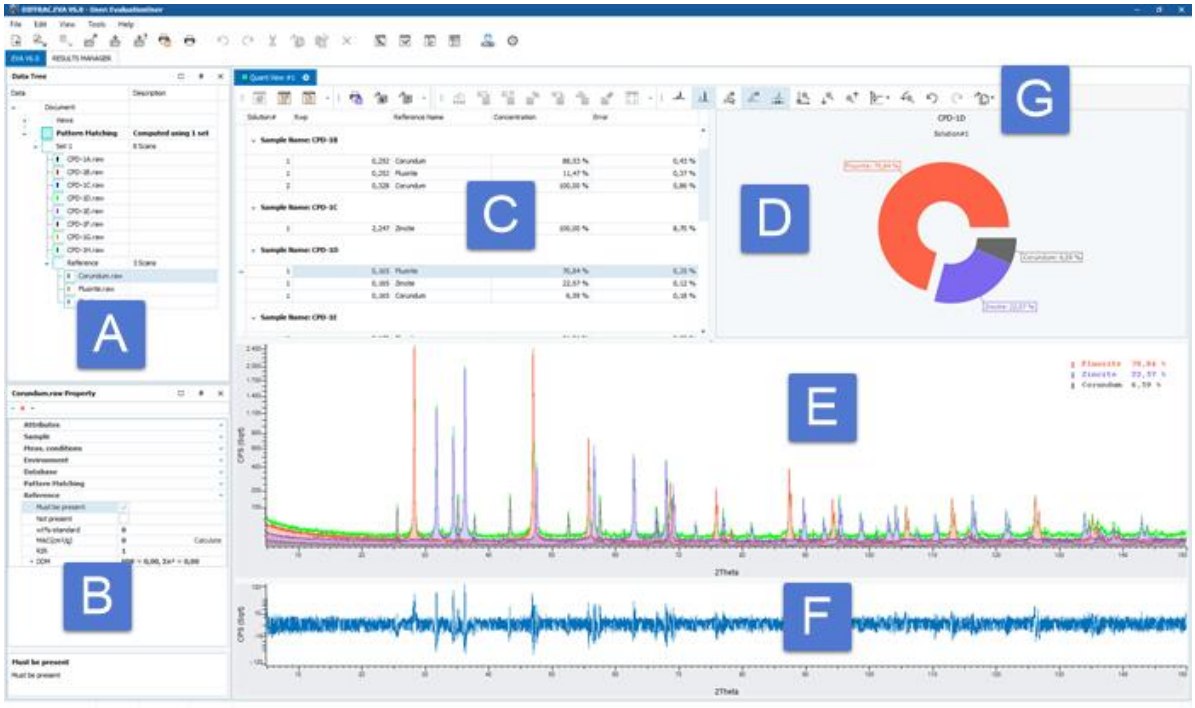
SQUALL can be performed in three different modes:

1. Expert mode SQUALL within the GUI of DIFFRAC.EVA. This is useful for the interactive evaluation of data but at the same time mandatory for establishing EVA projects that run SQUALL semi- or non-interactive
2. Operator mode SQUALL. This uses SNAPQUANT, an interface with reduced complexity. It allows reading EVA project files (that define a PMI method), loading and evaluating unknowns
3. online in a scripted version

All three modes are available in file-based operation, as well as using the instrument database of the Bruker XRD, starting with version DIFFRAC.MEASUREMENT CENTER V8.6. The latter also stores SQUALL results in the database and uses the Results Manager to display the identified material, the correlation, as well as a chart of the unknown materials scan and the matching reference scan.

Recommended Screen Layout – Working in Expert Mode

The below screenshot shows a DIFFRAC.EVA screen layout that is convenient to use for SQUALL and contains a minimal set of panels required. For users who do not like to work with toolbar icons or context sensitive menus in the data tree it is advised to additionally open the Data Command panel. The arrangement of panels is called a Layout. Layouts can be Saved and Loaded from the Tools menu of the DIFFRAC.EVA framework. While it is a matter of personal taste, use behavior and available computer screen hardware, we found individual Layouts to be useful for the different analytical task to be performed with DIFFRAC.EVA.



A. Data Tree Panel with scans (patterns) loaded to Set 1 and its Reference section.

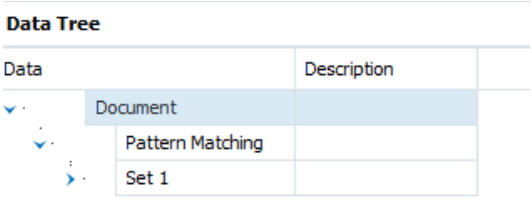
B. Data Property Panel. This is mostly required for changing scan colors, and to set the properties of the references. Note, only the reference section is displayed, while all other sections are collapsed for better clarity.
   Items C to G refer to the so-called Quant view that shows the results of a SQUALL analysis
C. Results table. It gives a numeric overview over the number of solutions, the species identified, concentrations with e.s.d. and the agreement parameter Rwp.
D. A graphical representation of the concentrations for the solution selected in item C. The colors of the donut match the color settings of the reference scans
E. Scan view presenting the actual sample selected either in the data tree (A) or the results table (C), Reference scans scaled according to their concentration in the sample, a calculated sum of all contributing references. The actual appearance, content and axis scaling depends on the settings in the toolbar G and the zoom / pan state in panel E.
F. Difference view of scan and sum of scaled references
G. Toolbar with Icon groups for creating views, manipulating the Quant view, or general tools

Further customization of the workspace is possible by

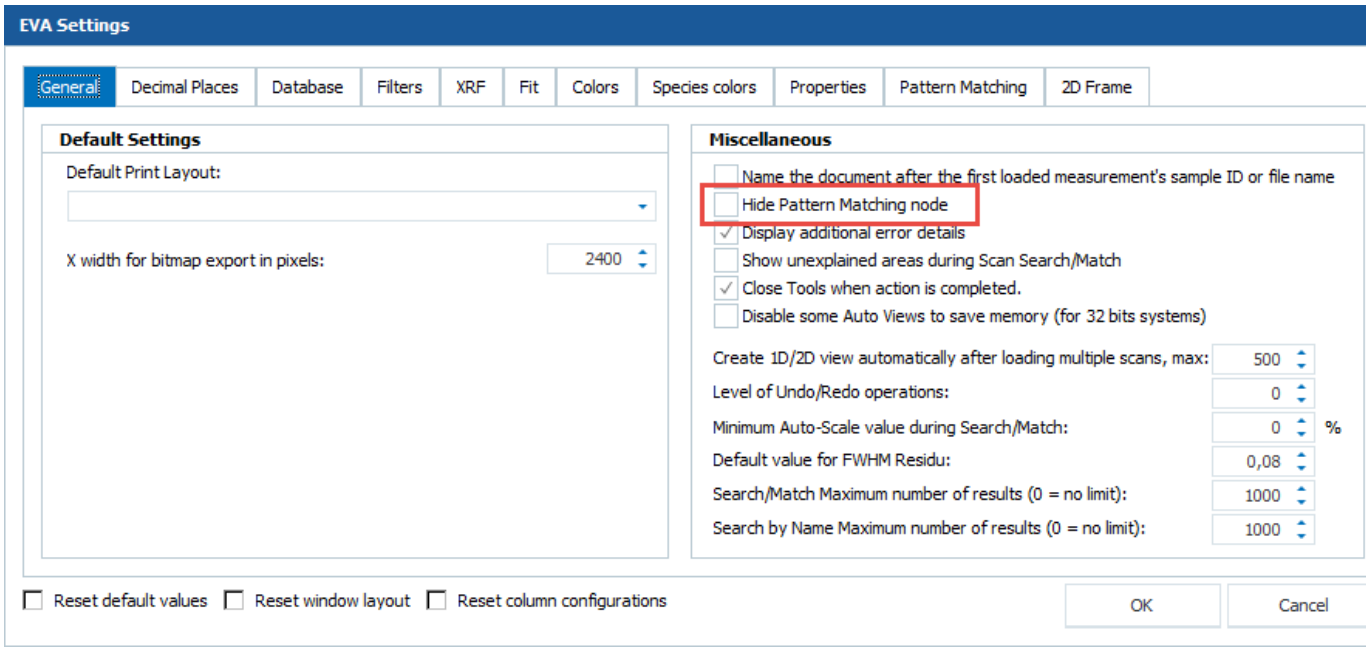
- Different docking patterns of the panels (see DIFFRAC.EVA User Manual)
- Removal of icons, relocation, or reconfiguration of toolbars (see DIFFRAC.EVA User Manual)
- Moving the splitter bars between the panels C, D, E and F to adjust their size.

Loading the data

In the *Data Tree* pane, you will see

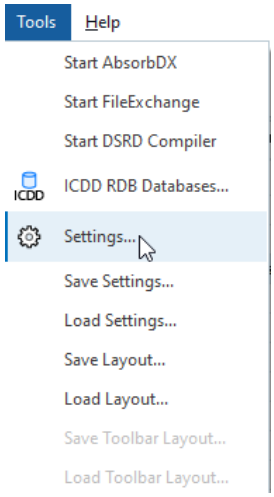


In case no Pattern Matching node is shown in the data tree check whether it was disabled in the general settings dialogue.

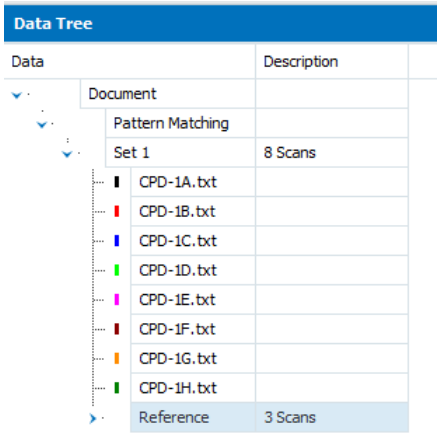


This is opened via the wheel icon in the main toolbar 

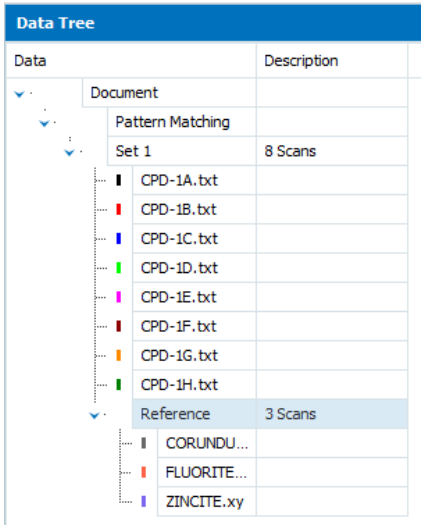
or the Tools menu:



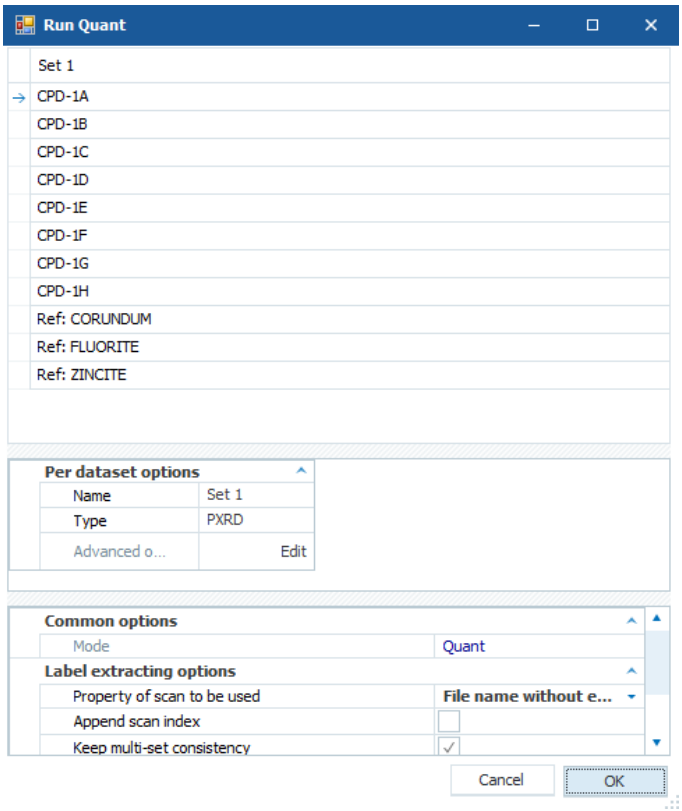
Load the sample files (those to be analyzed) under Set 1 in the usual way:



Now load the references in the same way *e.g.*

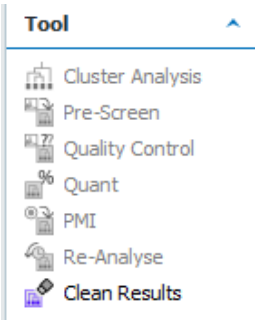


If you want to run SQUALL with the default settings, click on *Quant* in the Tool menu of the Pattern Matching node. The *Run Quant* GUI appears

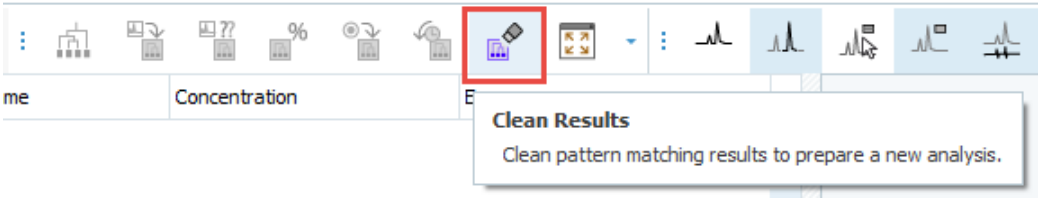


Click on OK and the program will run.

One run will analyze all sample patterns that are loaded to the Set. However, if you are performing multiple runs you will need to clean the results between runs. The same applies if you are opening a previously saved EVA file. This operation is found under the *Tool*: menu of the Command panel, or the Tool section of the Pattern Matching node in the data tree,

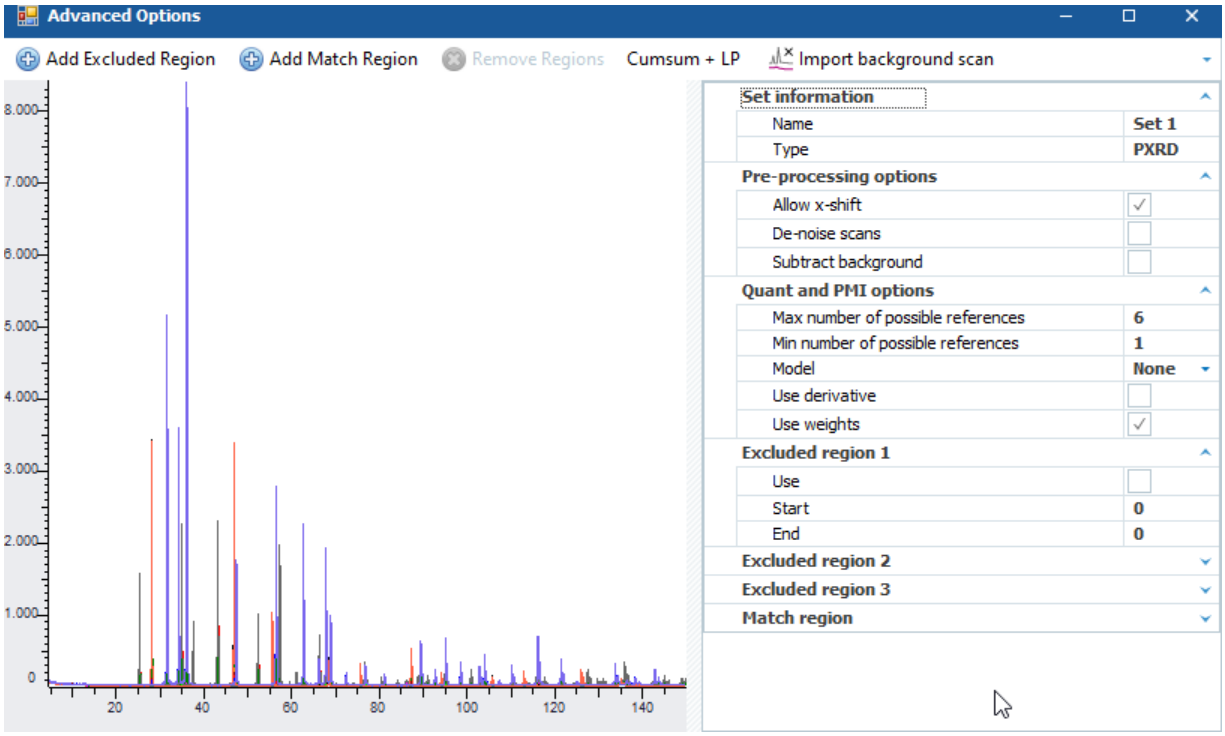


or use the toolbar icon



Program options

If you wish to select other options for SQUALL, then click Edit in the *Per dataset option* pane of the Run Quant GUI. The *Advanced Options* GUI will appear:



Data pre-processing

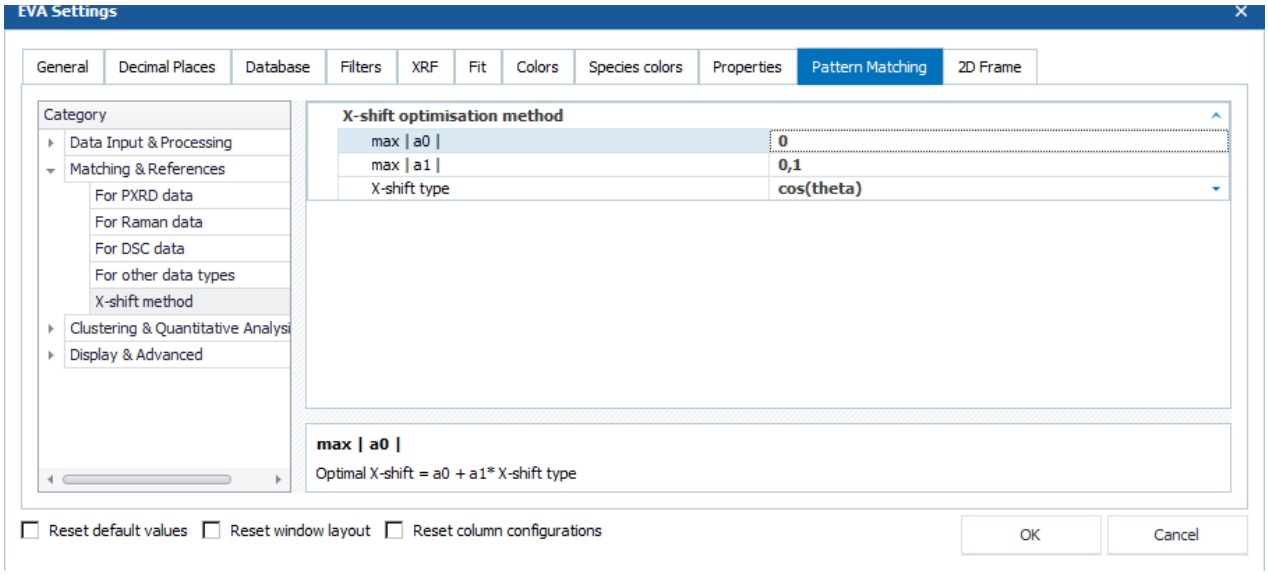
There are several user-decide pre-processing steps but also some automatic pre-processing is applied:

- Reference and sample scans are converted to CPS scale. That allows using scans measured at different counting statistics. We would advise to measure references at higher counting statistics than samples in order to reduce the pileup of noise for the calculated sum of references
- Reference and sample scans are truncated to cover a common X-range. Patterns not matching a common range are rejected
- Scans measured at different step width are rebinned

We will now describe each user-defined option in detail.

X-shift

For PXRD data SQUALL can correct for 2θ shifts in the references to optimally match the sample(s). X-shifts may be caused by different zero-point offsets of the XRD used for measuring the samples or references, sample height issues introduced in sample preparation, or the transparency of low absorbing material. Accordingly, three options for the 2θ shift are offered in the settings menu



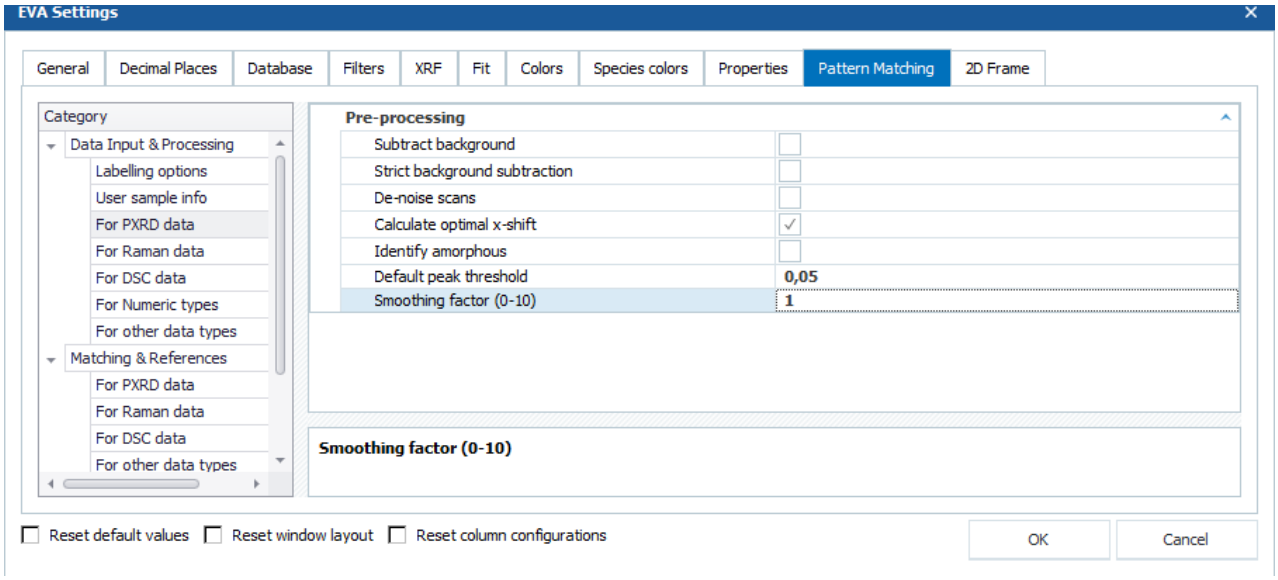
a_0	Linear shift which acts as a zero-point correction. The optimum value of a_0 is computed.
$a_1\cos\theta$	Corrects for sample height. The optimum value of a_1 is computed.
$a_1\sin 2\theta$	Corrects for sample transparency. The optimum value of a_1 is computed.

Note, these variables are highly correlated and only one of them can be computed at any time. SQUALL does this using the full pattern profile and a grid search method. The default maximum values of the shifts are 0.5° for a_0 and 0.3°a_1. These can be changed using the above shown *Settings/Eva Settings* Graphical User Interface (GUI).

To use shifts in the analysis you must tick the box labelled *Allow x-shift* in the Advanced Options GUI. Shifting is currently only possible using PXRD data.

Denoise scans

Noisy scans can be smoothed using Daubechies wavelets. The degree of smoothing is configured in the EVA Settings GUI as shown below



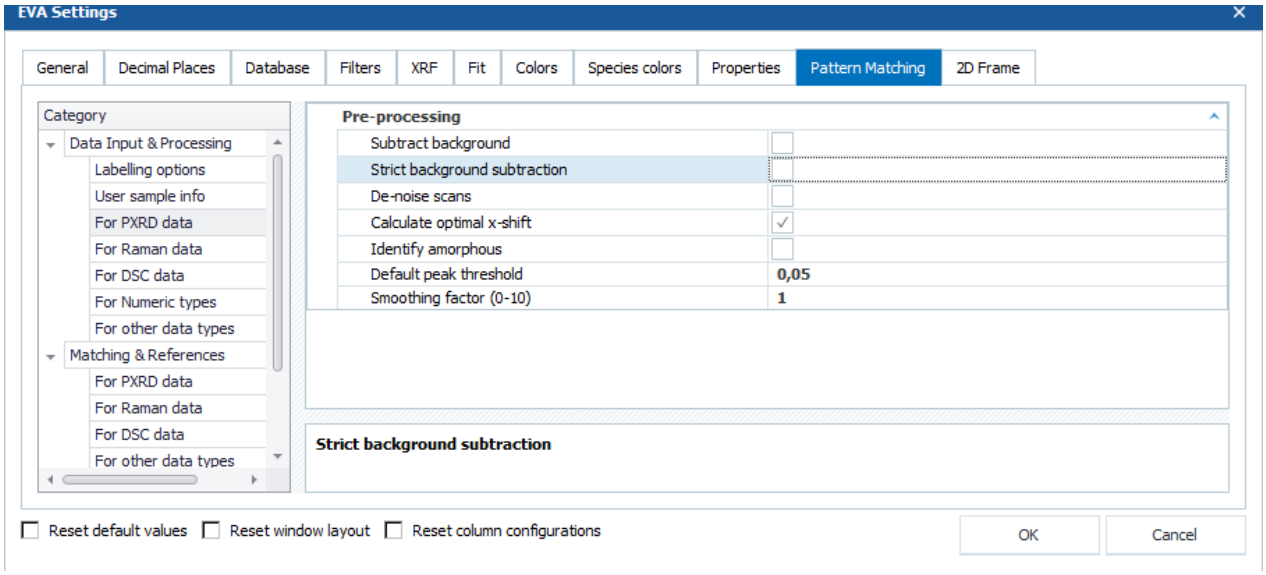
The Smoothing Factor (0-10) can be set here. One is the default and 10 the maximum value. The higher the number the greater the degree of smoothing. In general smoothing should have only a small effect on the results of SQUALL unless the data are very noisy. In that case it is advised to rather collect better data with improved counting statistics instead to risk distortions of the data by mathematical corrections.

Subtract background

Background subtraction is important, particularly for an accurate quantitative analysis of well crystalline material. If not removed, the background of each reference may contribute to a different extent to the calculated sum of intensities.

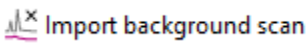
SQUALL may use different background removal options:

- **Polynomial fitting** in either a regular or strict mode. The latter uses smaller 2θ intervals and reduces the background more drastically than the regular method. In general, the regular subtraction method works best. The regular vs. strict mode is selected in the EVA settings dialogue while the Subtract background in this dialogue is a pre-set that can be disabled in the Advanced Options GUI.

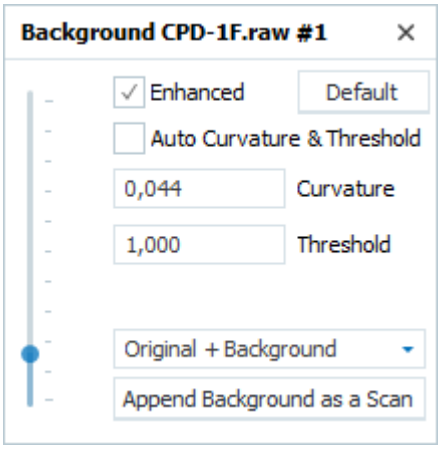


Note: If you are analyzing samples with a high degree of amorphous content, do not subtract the polynomial background – it will remove a significant portion of the amorphous signal.

- **Experimental background.** Alternatively, a well crystallized sample with a small number of peaks and ideally not showing fluorescence can be used to experimentally determine a "measured instrument background". Use the EVA program to determine such a background, safe to file and then use the option Import background scan which is found on the top menu of the *Advanced Options* GUI:

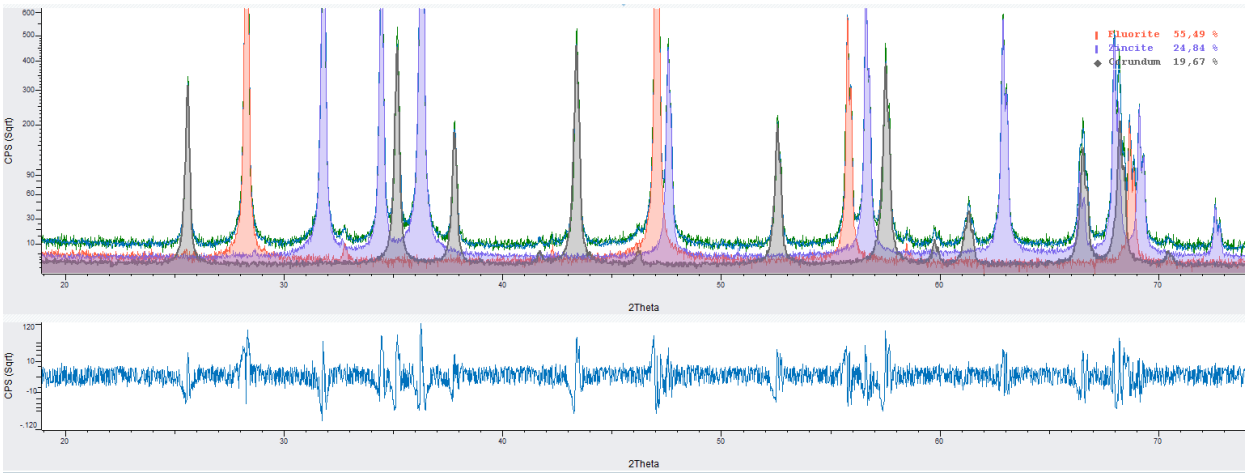


- **Manually subtracted background**. Good experience was made using background subtraction in EVA, applying the extended model at manually adjusted low curvature, saving the background subtracted scans and using those for SQUALL

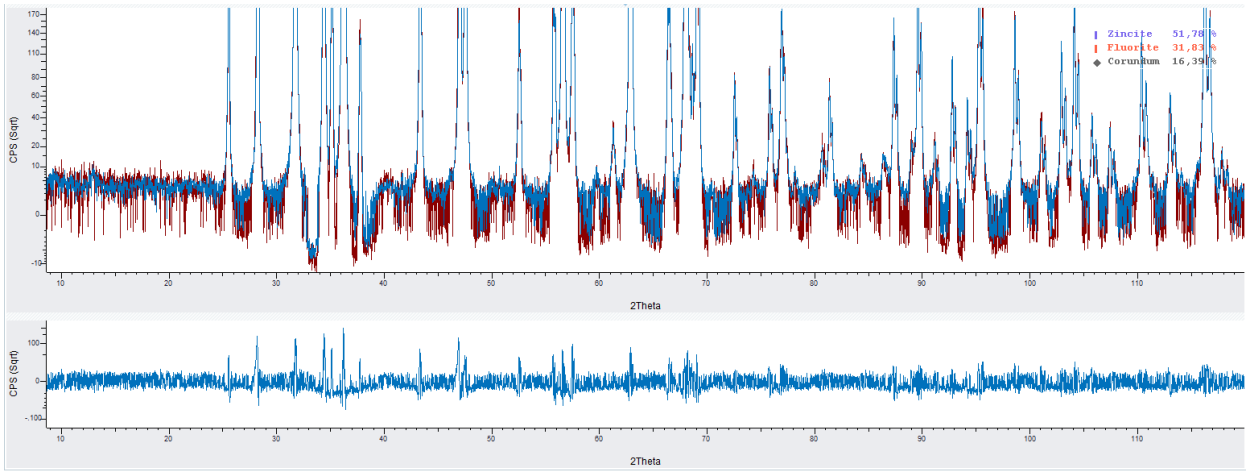


Note, this approach is no feasible for on-line data evaluation in SnapQUANT as it does not approach the newly measured samples. A combination of polynomial fitting together with manually background corrected references may work well.

Visual inspection of the background correction in the scan and difference view is highly recommended. For a successful background correction, the resulting difference curve should oscillate around zero. Looking at the below example no background correction is performed, a nice difference curve is shown. The calculated curve is matching the sample scan (green). However, references show a different base level each. That means, the scale factors also take that into account, and this may introduce a systematic error to the concentrations of minor phases



The next chart shows a poor example of polynomial subtraction in strict mode. Individual reference scans are not shown for clarity. Again, the calculated curve seems to nicely fit the measured sample scan. However, there are also oscillations towards negative intensities. They are typical for overfitting with polynomials, they distorted peak tails. It cannot be controlled how such behavior of a major phase may affect the scaling of a minor phase, thus again introducing systematic errors in the quantification.



Other Quant Options

Max number of possible references

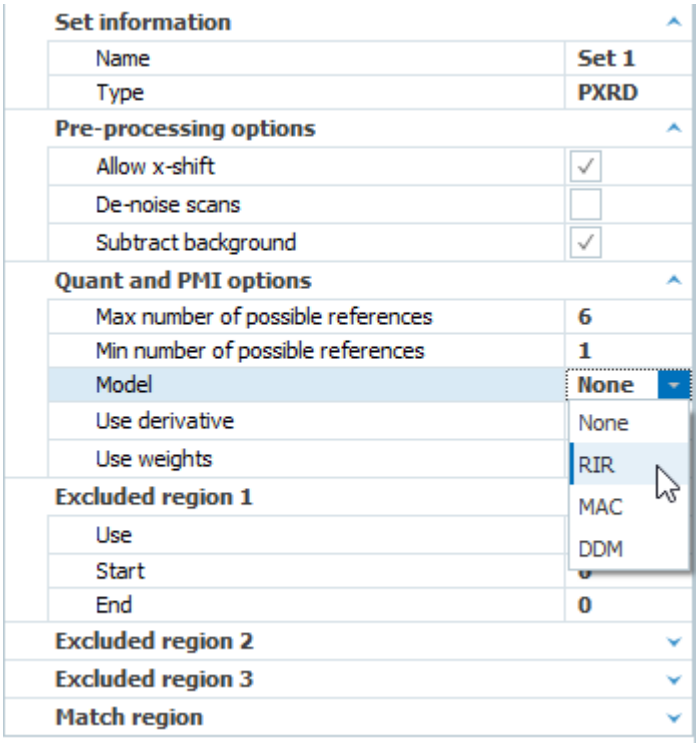
This entry is used to specify the maximum possible number of phases in the mixture. It can be as large as needed, but it helps SQUALL if it is set to a realistic figure.

Min number of possible references

This entry is used to specify the minimum possible number of phases in the mixture. It can be as small as needed, but it helps SQUALL if it is set to a realistic figure.

Model

Model is used to specify how the conversion form scale-% to wt-% is performed. Currently, there are up to four options, depending which ones are configured for the respective references, they can be selected from the drop-down list:



- **None.** This corresponds to scale% and may be sufficient for cases where the scattering power of the references is the same
- **Reference Intensity Ratios (RIR).** This is a standard PXRD technique for semi-quantitative analysis. Usually, it is based on the so-called I/Ic ratio, the intensity of the strongest peak of the analyte to the strongest peak of corundum in a 50/50 wt% mixture. This factor can either be determined experimentally, taken from the databases or calculated from the crystal structure directly. It is not necessary to use corundum, it could be any reference if the set of RIR values of the references is determined self-consistently. That also implies to use the same method for all references to determine the intensities. Since SQUALL relies on scale factors of profiles it deems consistent to use peak amplitudes for the determination of the RIRs.

  To enter the RID values left-click on each reference in turn, the *Property* panel for the reference becomes active. Enter the RIR value. You do this once for each reference.

- **Direct Derivation Method (DDM).** This is a relatively new technique. (See H. Toraya. *J. Appl. Cryst.* (2016) **49**, 1508-1516.) For each reference it requires:

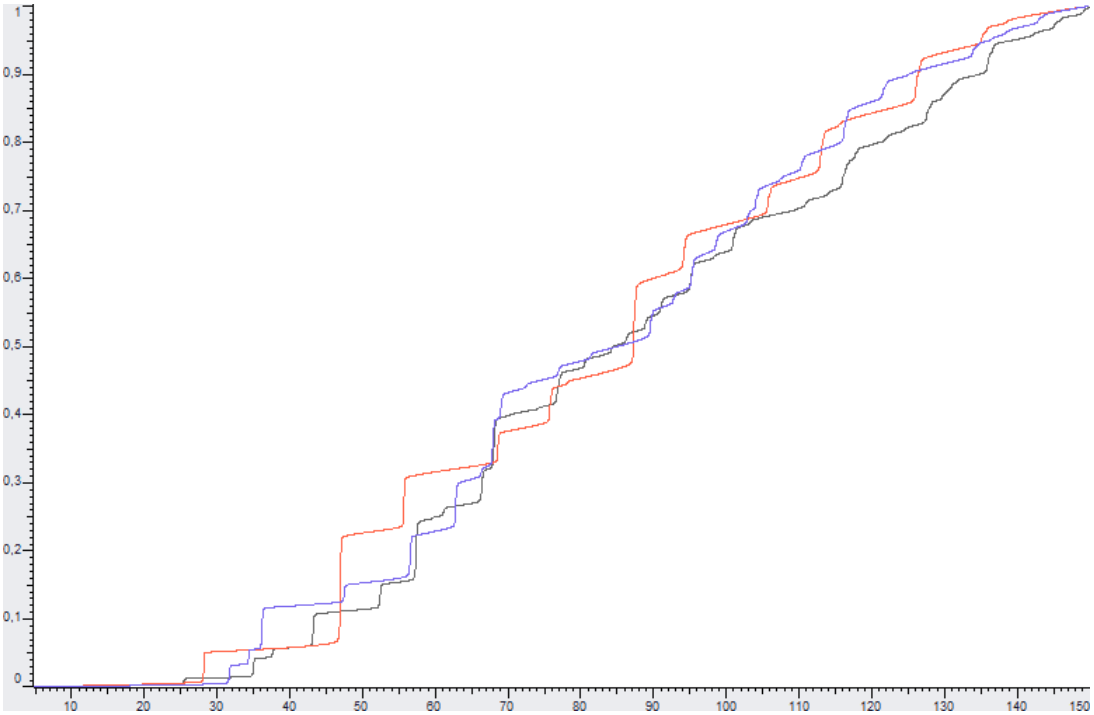
  1. The chemical formula molecular weight (not the unit cell molecular weight).

  2. The sum of the number of electrons squared (Σn^2) for each atom in the chemical formula.

     Example: for corundum (Al_2O_3) the molecular weight is 101.84 and $\Sigma n^2 = 542$. To enter these values left-click on each reference in turn, the *Property* pane will appear. Click on the *DDM* option and enter the MW and Σn^2 values. Keep in mind the sum of electrons squared changes for neutral or ionic states. It may be worth trying this.

  3. DDM further uses a “total scattering factor” of the reference that is the sum over all Polarization- and Lorentz-corrected step intensities of the reference. This value is automatically determined from the reference pattern by SQUALL. Obviously, the sum over all step intensities is influenced by the step width or number of steps (SQUALL takes care of it), the background intensity of the scan (subtraction is mandatory for DDM) and the scan range. Since the number of peaks in a given scan interval depends on the crystal symmetry and the wavelength a scan (or evaluation-) range should be chosen where all the references do not show major variation of the cumulated intensity with the diffraction angle.

     This can be checked in the Cumsum+LP chart of the Advanced Options window.

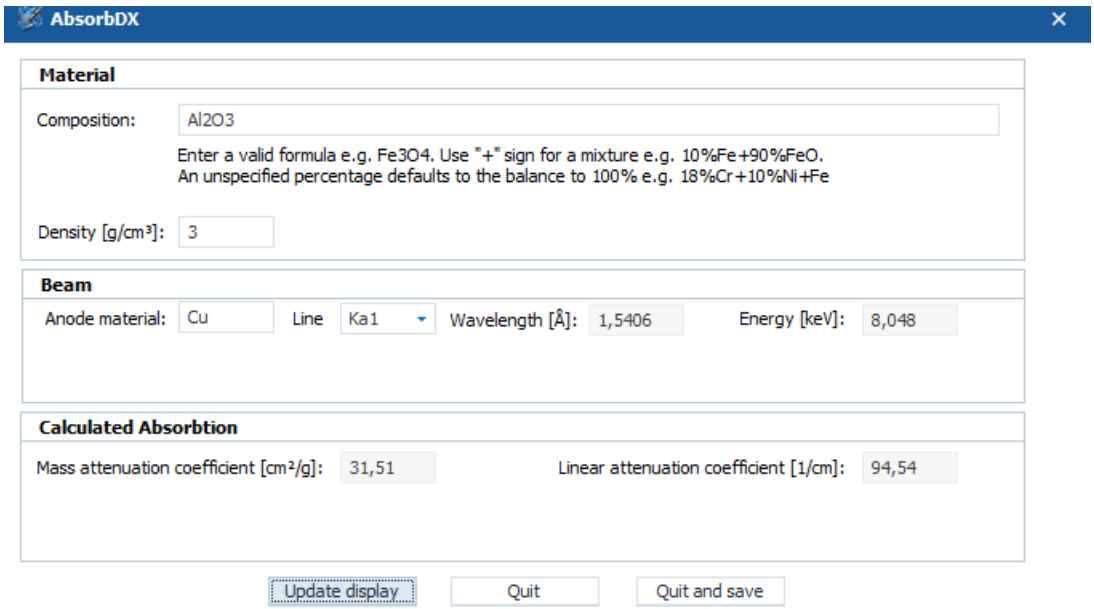


It displays the cumulated PL-corrected intensity of the references as a function of the diffraction angle, normalized for the highest diffraction angle. The steps indicate the presence of strong diffraction peaks. Toward higher angles the curves tend to overlap or at least to show no crossover, as the maximum number of peaks per phase levels out and the scattering power decreases with sin(theta)/lambda. If excluded regions are chosen, they should be placed such that crossover between the cumulated intensity curves is avoided.

- **Mass Absorption Coefficients (MAC or μ*)** MACs can also be used. The approximations used in applying this technique in SQUALL make it the least reliable of the three techniques. For details on MAC see Gilmore, C.J., Barr, G., and Dong, W. Clustering and visualization of powder-diffraction data. International Tables for Crystallography (2018) Vol. H. Chapter 3.8. pp. 325-343.

  To enter the MACs left click on each reference in turn, the Property pane will appear. Click on the MAC(cm^2/g) option and enter the relevant values. Alternatively, a calculator is provided.

  Click on the *Calculate* option and GU appears:



The entries here should be self-evident. The program is described in detail in section 16 of the DIFFRAC.SUITE user manual.

Use derivative

The x-shift method uses the full pattern. If the peaks are broad, it is sometimes preferable to use the 1st derivative data instead and this is the purpose of this option. For most data the effect is very small.

Use weights

In the final stages of the analysis, SQUALL uses ridge regression which is particularly useful to overcome problems with multi-collinearity which are a feature of mixture analysis. This can be weighted using standard counting statistics methods used in PXRD or unweighted. The effect is generally small.

Defining which regions to use

Quite often there are regions of 2θ (or cm^{-1} for spectroscopic data) which the user wishes to exclude from the analysis, typically small angle and high angle regions. SQUALL provides two ways of doing this.

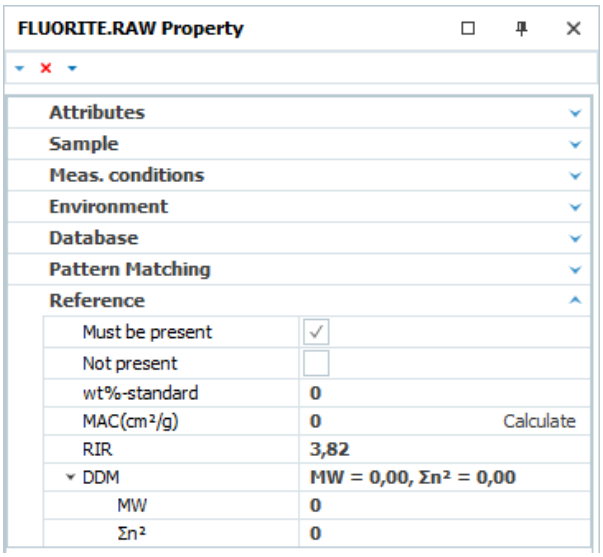
- **Excluded regions** These are regions that are to be excluded. Up to three regions can be defined. This is done either by clicking on the *Add Excluded Region* at the top left-hand side of the *Advanced Options* GUI and dragging the cursor or by entering the values in the *Excluded region 1/ Excluded region 2/ Excluded region 3* boxes.
- **Match Region** This option works in the opposite way: the match region is used in the calculation and the other regions excluded. It is possible to have an excluded region inside a match region.

You can remove excluded and match regions using the *Remove Region*s button in the GUI.

Reference specific options

The references have in addition to the above-described quantification methods several associated options which can be very useful. They are accessed by clicking on the relevant reference in the data tree and shown in the Data Property panel:

- **Must be present** This reference must be present in the mixture. Note that if there is no evidence for the reference to be present a warning message is produced
- **Not present** This reference is explicitly excluded from the mixture. This option can be useful with multiple runs of SQUALL where the components are uncertain.
- **Wt%-standard** Fixes the weight % value. In contrast to semi-quantitative analysis in DIFFRAC.EVA this is not used for the determination of an additional amorphous phase



Program Output

The main output pane is the Quant View as pictured earlier. In the top left scrollable window, the results are displayed for each sample. The first column is the solution number. For this data set there is only one possible solution found by SQUALL, so the entry is always 1, but if there are multiple solutions, they are output in order of increasing R_{wp} value with the numbering 1, 2, *etc.* The 2nd column gives the R_{wp} value, then the %concentration, and then the estimated error. The doughnut shaped diagram at the top right pictures the concentration distribution.

The Quant Toolbar

What is displayed in the scan graphics pane depends on the options selected from the Quant Toolbar buttons at the top of the window.

Going from left to right these buttons in the Quant Toolbar set the following options

Summed: Shows the calculated pattern derived from the summation of all the components.

All Components: Shows all the reference components individually.

Selected Component: Shows the currently selected reference component highlightes in the view legend while hovering the mouse over the reference scans

Legend: Display the legend of references with the % contents.

Residual: Opens/closes a graphics pane showing the residual plot when the calculated pattern is subtracted from the observed.

Base Zero: Sets the bottom of the working area to zero intensity.

Fit Bottom: To adjust the bottom of the working area in order to always fit the lowest measured point available in the zoomed area

Fit Top: Adjusts the top of the working area in order to fit the highest measured point available in the zoomed area

Linear/Square Root: Selects a linear or square root intensity scale for the scan view and the residual.

The View Toolbar

Buttons at that toolbar facilities for creating and viewing the analysis results.

Print quant table Brings up a print preview of the results in tabular form with print options to customize the print, save or send in various formats including PDF, HTML and various text- and spreadsheet-processing file formats

Copy as Bitmap copies the window as a bitmap for pasting into word processing documents.

Copy as metafile copies the above in window metafile format

The Create View Toolbar

From left to right the respective buttons open the Quant View (default), a Log View, and a Report Writer View.

The Log File View contains ASCII text providing a multitude of numeric details that is not shown in the main concentration table whereas the Report contains the main results together with essential run information and several references.

The Tools Toolbar

This is the general Tools collection for all Pattern Matching applications. The reader is referred to the Cluster Analysis Manual for the greyed-out items.

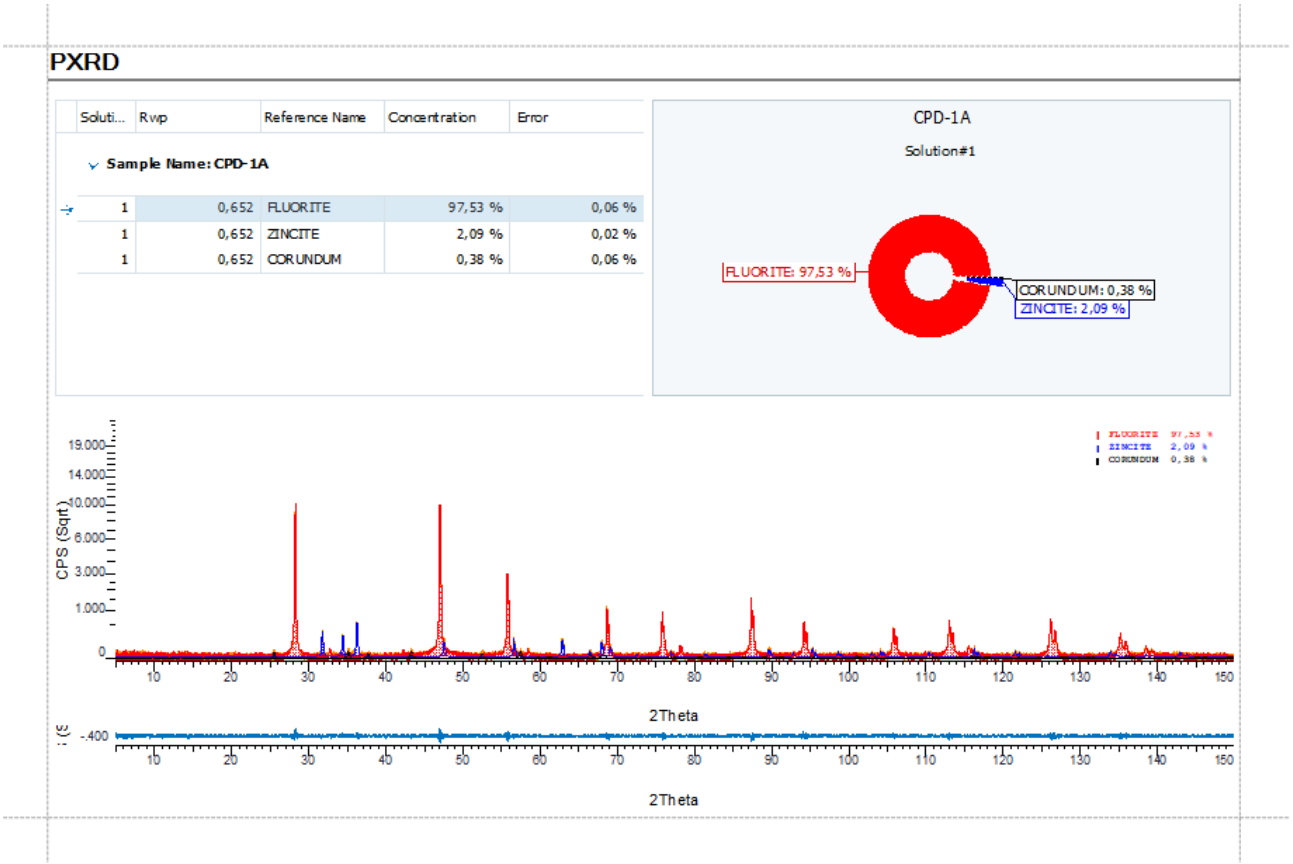
The two icons at the right-hand side are:

Clean Results: to clean the results table for a new pattern matching analysis

Full Screen: Toggle the application in full-screen mode to make maximal use of the available monitor space

General printing

The Print Preview or Print icons in the main program toolbar allow printing via the usual EVA print facilities. For SQUALL analysis a combined view per sample is created that contains the respective phase list and concentrations, the concentration diagram, and the scan view (and residual view) according to the chosen view settings.



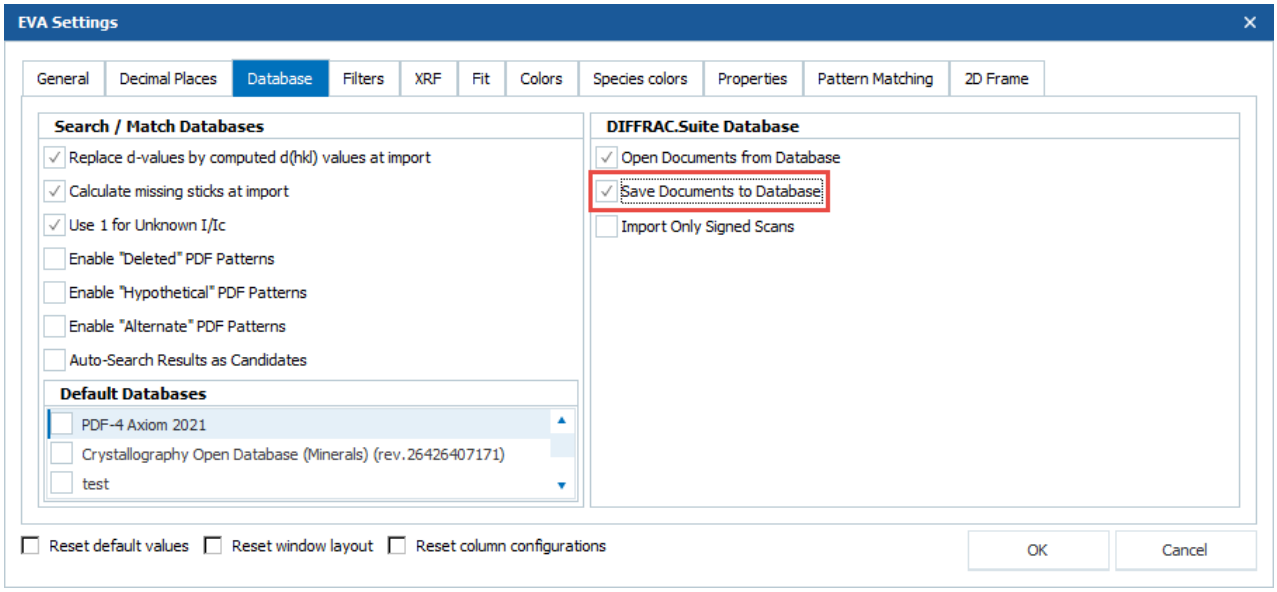
Saving SQUALL results to the Database

Results of a SQUALL analysis for the individual solutions suggested by the software are

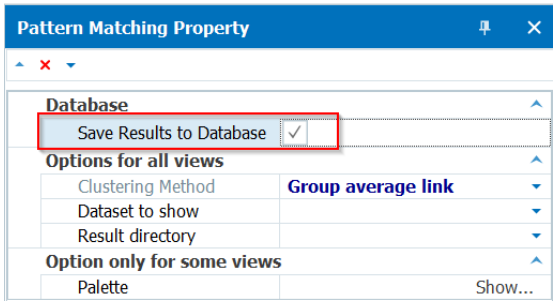
- the possible reference phases
- the corresponding Rwp value
- the concentrations including estimated standard deviation

Those results, together with the corresponding reference scans can be saved to the DIFFRAC.SUITE database for later reporting using the Results Manager module of the DIFFRAC.MEASUREMENT CENTER.

The export of results to the database is only possible while running EVA in database mode and requires enabling results writing in the general settings.



The Pattern Matching node's property "Save to database" must have been checked prior to saving the EVA file to allow storing the above-mentioned results like concentrations:



Operator mode SQUALL

To run a SQUALL analysis in operator mode an EVA project file needs to be established first. Use the Expert Mode in the EVA GUI to load references, prepare the preprocessing and quantification options. Save the EVA project file either to the file system or the database.

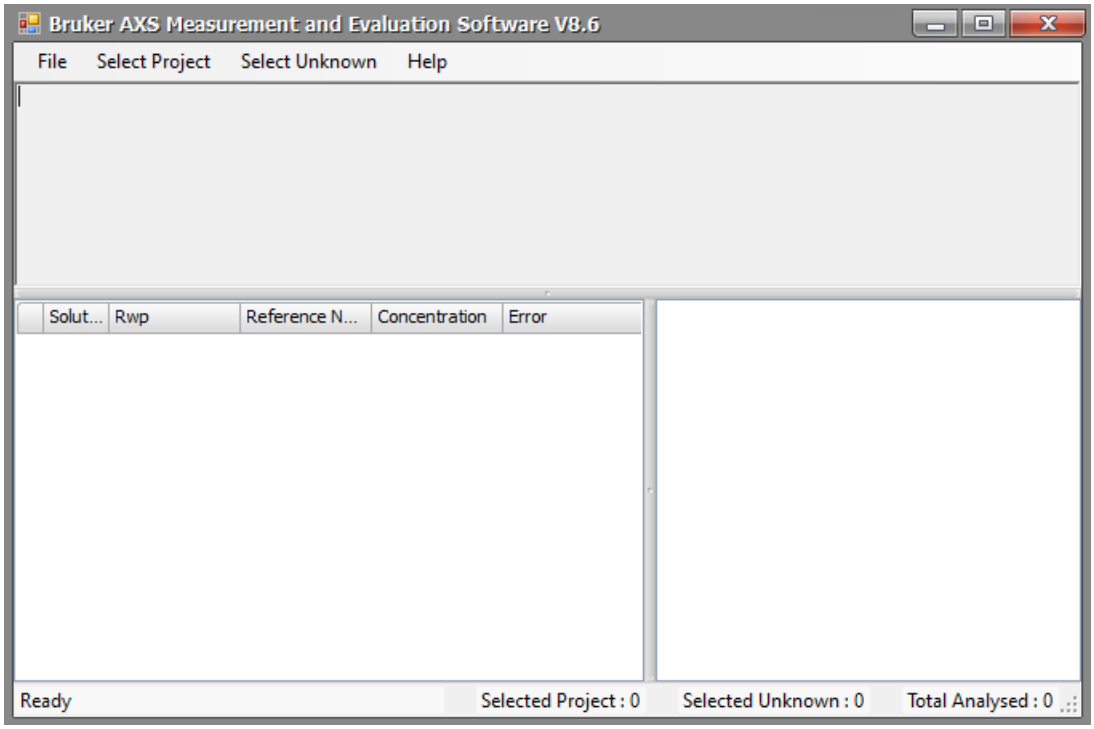
Note, the evaluation in the GUI needs to be performed once with a sample file prior to saving to properly initialize all processing steps.

Open the program SnapQuant.exe from the DIFFRAC.EVA V6.0 group in the Windows start menu.

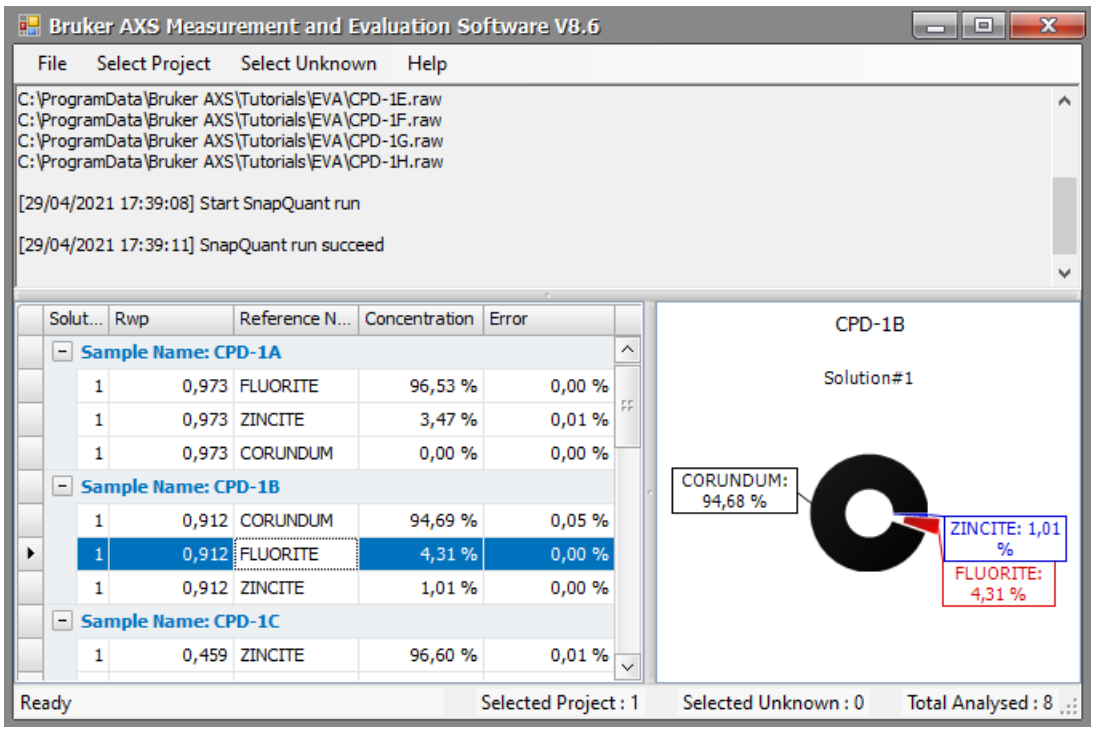
Use

- Select Project to load an EVA SQUALL project from either the file system or the database
- Select Unknown to load the measurements to be identified from either the file system or the database

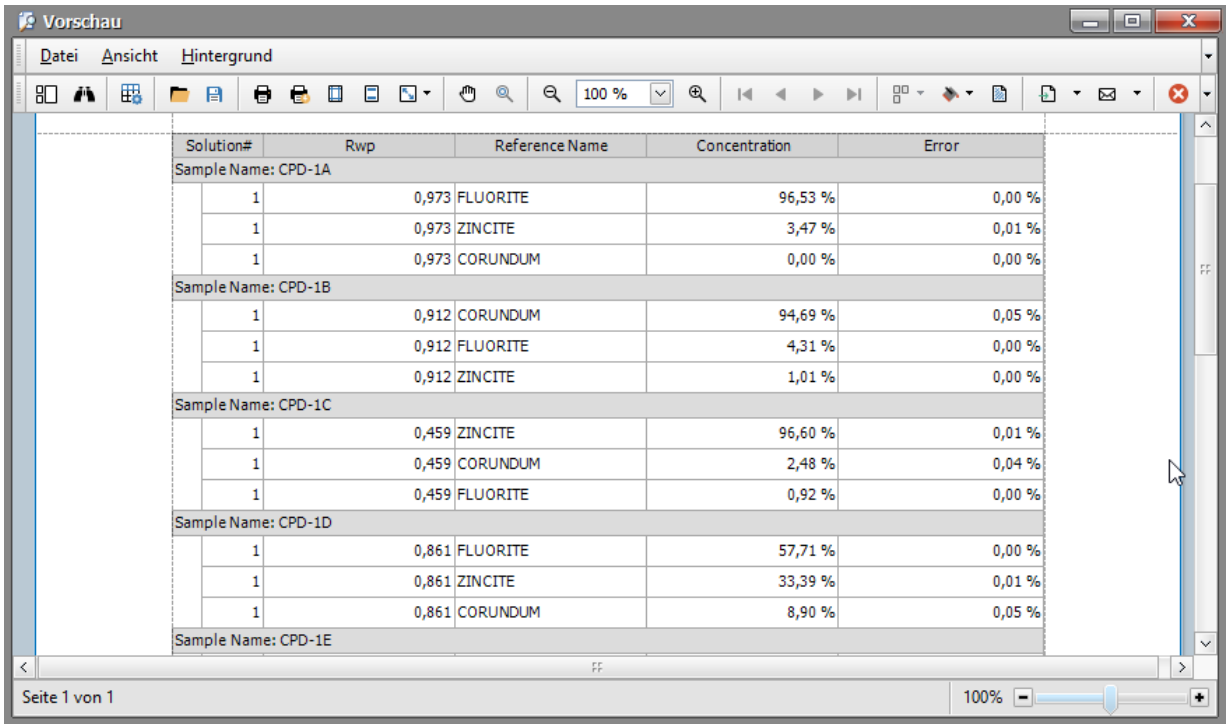
- Execute the evaluation by pressing the Run button inside the File menu.



The upper part of the SnapQuant window contains a log of the actions. The SQUALL results are shown in the lower part of the screen in tabular and graphical form.



A print preview of the concentration table can be accessed via the File menu. It allows customization of the report, export to spreadsheet and text processing software.



Solution#	Rwp	Reference Name	Concentration	Error
Sample Name: CPD-1A				
1	0,973	FLUORITE	96,53 %	0,00 %
1	0,973	ZINCITE	3,47 %	0,01 %
1	0,973	CORUNDUM	0,00 %	0,00 %
Sample Name: CPD-1B				
1	0,912	CORUNDUM	94,69 %	0,05 %
1	0,912	FLUORITE	4,31 %	0,00 %
1	0,912	ZINCITE	1,01 %	0,00 %
Sample Name: CPD-1C				
1	0,459	ZINCITE	96,60 %	0,01 %
1	0,459	CORUNDUM	2,48 %	0,04 %
1	0,459	FLUORITE	0,92 %	0,00 %
Sample Name: CPD-1D				
1	0,861	FLUORITE	57,71 %	0,00 %
1	0,861	ZINCITE	33,39 %	0,01 %
1	0,861	CORUNDUM	8,90 %	0,05 %
Sample Name: CPD-1E				

Prior to reporting further samples can be added to the analysis, or the results table can be cleaned after the analysis prior to further evaluations.

On-line SQUALL

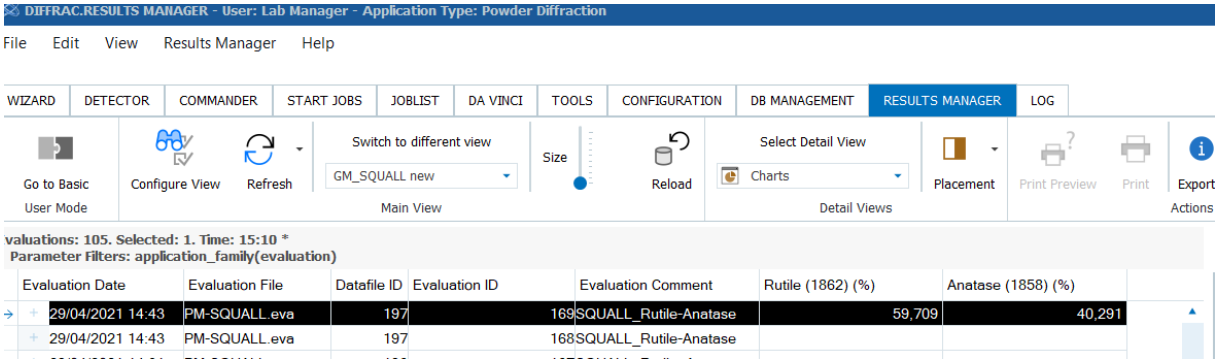
Prerequisite

The EVA project file with the reference samples must be saved in the database as described before.

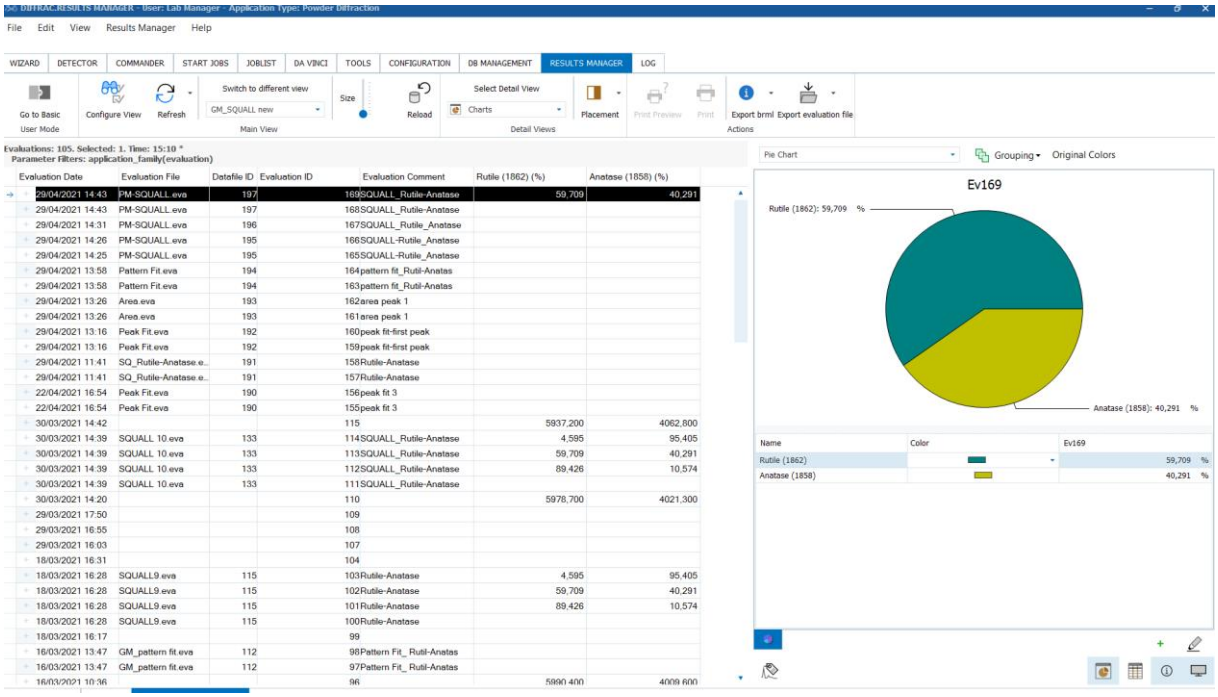
Perform the on-line analysis

While entering the required sample data in the Start Jobs plugin of the measurement software the EVA file must be given in the respective column and the evaluation script “SnapQuant” must be selected.

After the measurement is finished, the SQUALL on-line analysis is carried out without user intervention and the results are written to the database. The results can be inspected with Results Manager. Depending on the selected view, the results appear in a table,



as graphical representation, or both:



PMI: Positive Materials Identification

Introduction

Positive Materials Identification (PMI) is a new tool starting from DIFFRAC.EVA V6 to perform Pattern Matching of XRD scans. While in quantitative analysis (Quant) we attempt to analyse the contents of a mixture; in Positive Materials Identification (PMI) the question is simpler: I have a unknown sample and a set of reference scans of different materials – which reference best matches my unknown sample? Note, in the context of PMI a material does not need to be a pure phase. A material can also be a mixture of phases with reasonably well-defined composition.

The PMI program identifies materials by finding the reference with the highest correlation with the sample; if it is above a given threshold (which is set by the user) we can say that we have made a positive identification of the material. Of course, we are not limited to a single sample: the number of samples is virtually unlimited.

PMI can be performed in three different modes:

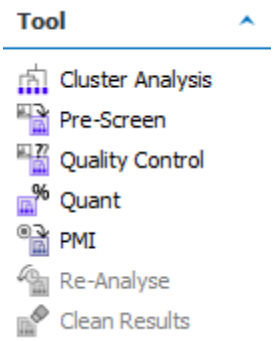
1. Expert mode PMI within the GUI of DIFFRAC.EVA. This is useful for the interactive evaluation of data but at the same time mandatory for establishing EVA projects that run PMI semi- or non-interactive
2. Operator mode PMI. This uses SNAPQUANT, an interface with reduced complexity. It allows reading EVA project files (that define a PMI method), loading and evaluating unknowns
3. Online in a scripted version

All three modes are available in file-based operation, as well as using the instrument database of the Bruker XRD, starting with version DIFFRAC.MEASUREMENT CENTER V8.6. The latter also stores PMI results in the database and uses the Results Manager to display the identified material, the correlation, as well as a chart of the unknown materials scan and the matching reference scan.

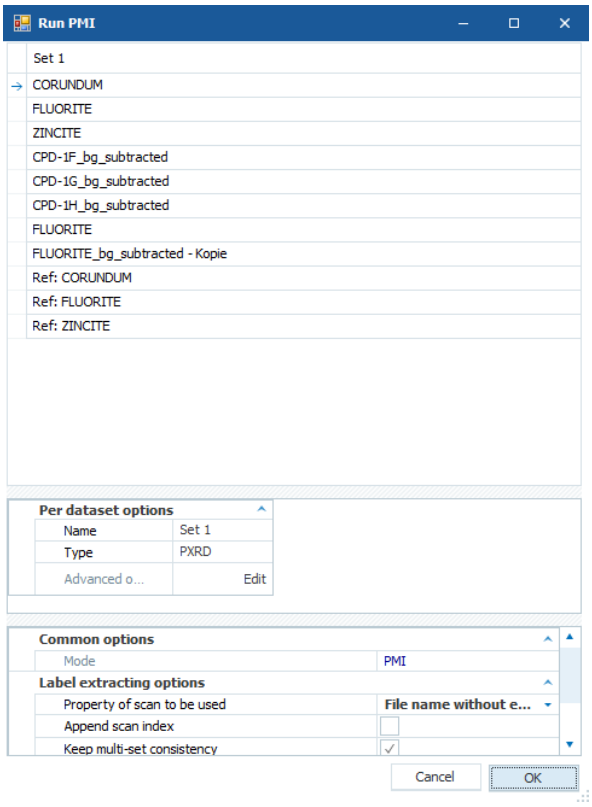
Interactively running PMI

Here is an example using data from CPD-1 (to be found in the EVA tutorial directory): a set of mixtures and pure phases of Corundum, Fluorite and Zincite. We load the unknown samples and reference scans for Pattern Matching by creating a new document and filling the Set 1 and Reference nodes. Set 1 must contain at least one scan of an unknown, whereas References hosts all scans that shall be used for the identification.

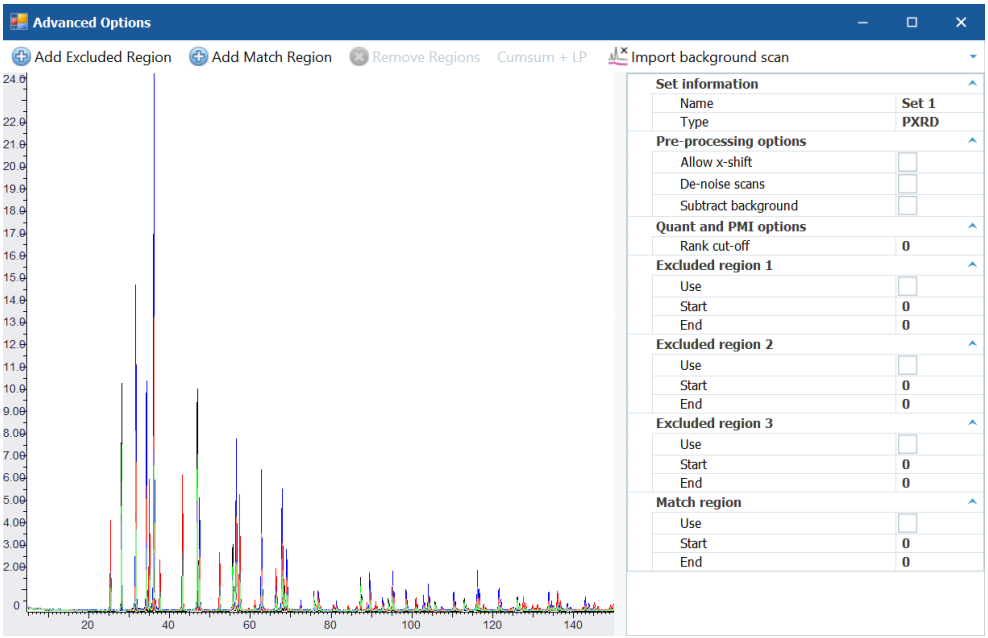
Choose the *PMI* option from the Tools context menu of the Pattern Matching node



A *Run PMI* GUI appears:



For this run click *Edit*. The Advanced Options GUI appears:



All these options have been discussed in the SQUALL documentation except the *Rank cut-off*.

It is also important not to ignore background and x-shift options.

The rank cut-off parameter

The cut-off parameter is the key to the success of PMI. It is best determined by using samples with known composition and varying the cut-off to give the optimum results. Let the rank cut-off = r, and the correlation coefficient between the sample and a given reference be ρ, then

ρ ≥ r: the sample is deemed to have been **positively identified**.

ρ < r – 0.2: the sample is **unidentified.**

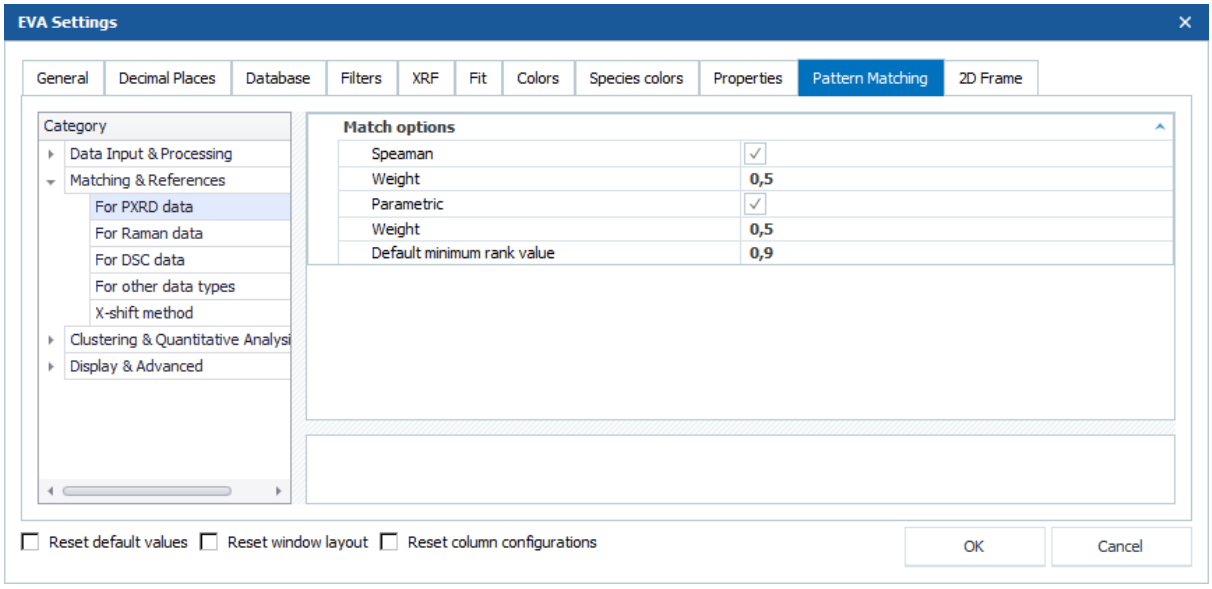
In the log file any sample with r < ρ < r – 0.2 is flagged as **possibly identified.**

How PMI works

PMI is entirely based on correlation coefficients which are fully described in the SQUALL section of the manual. Three coefficients are used on the full-profile data:

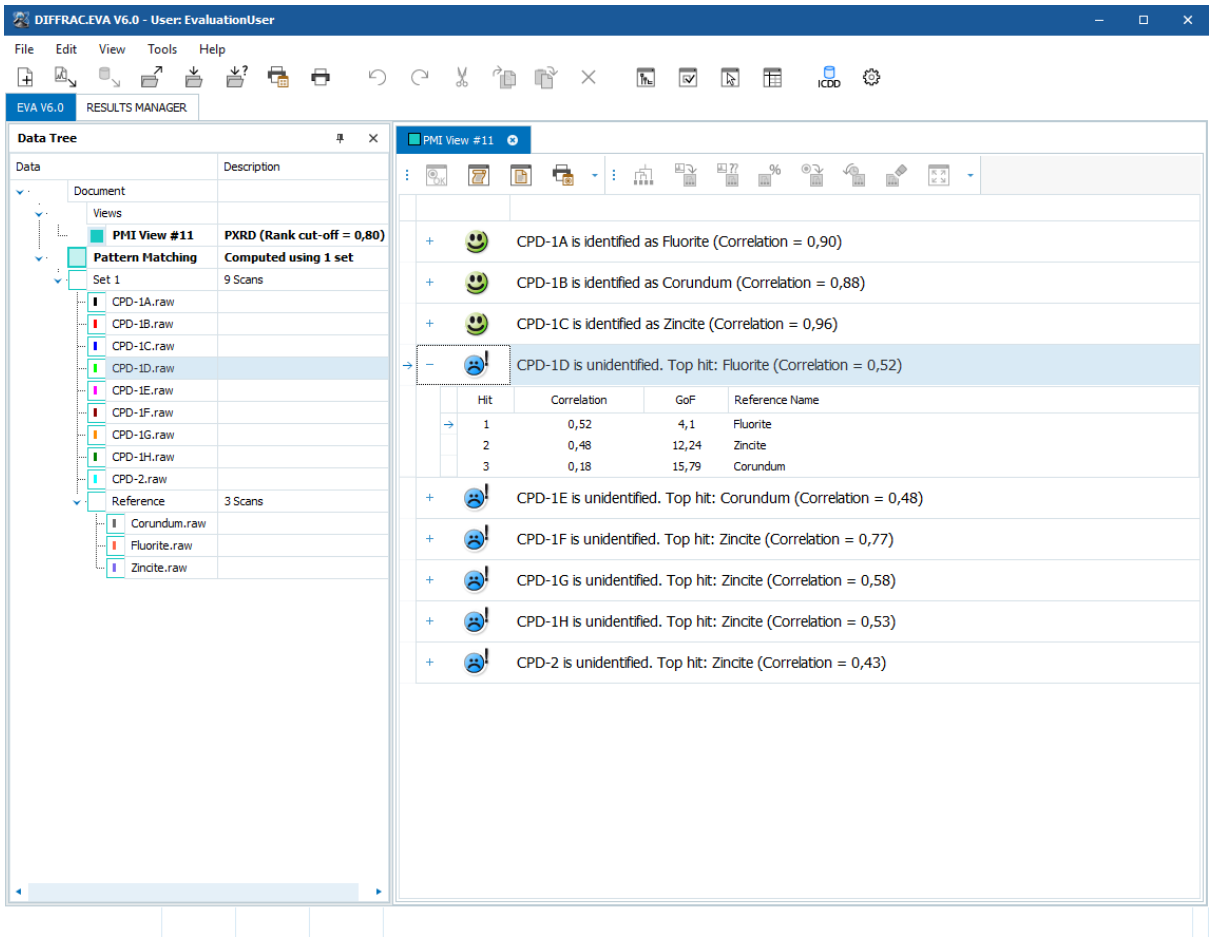
Spearman's rank order coefficient, ρ.
Pearson's r.
Kendal's τ.

They are averaged using Fischer transforms to give a combined correlation coefficient that is compared to the cut-off. A global default value of the rank cut-off is set in the EVA settings dialog



The PMI view

The results appear in the *PMI View* window. With the above-described example data, a cut-off of 0.8 and no data pre-processing, the windows look like this:



The PMI view presents a quick graphical overview of identified and unidentified scans. For each scan the name of the closest matching reference is provided, together with the combined correlation coefficient.

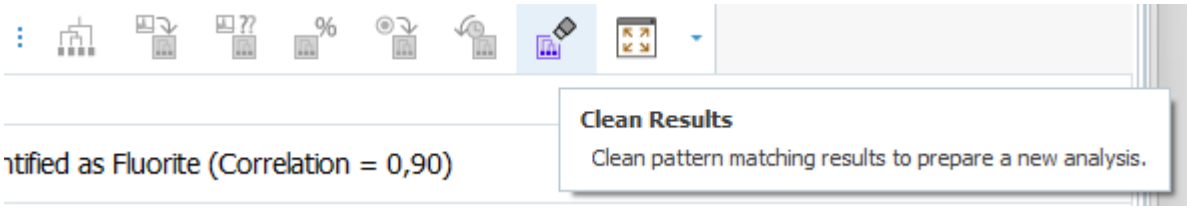
Each entry in the PMI view may be expanded to see a list of the three closest matches. This can be helpful in method development to identify similarities to references that possibly may be undetected otherwise and need adjustment of the rank cut-off.

The view toolbar provides access to log file and an automated report. The above example furthermore shows the icon for the print-preview of the report that was manually added to that group for convenience

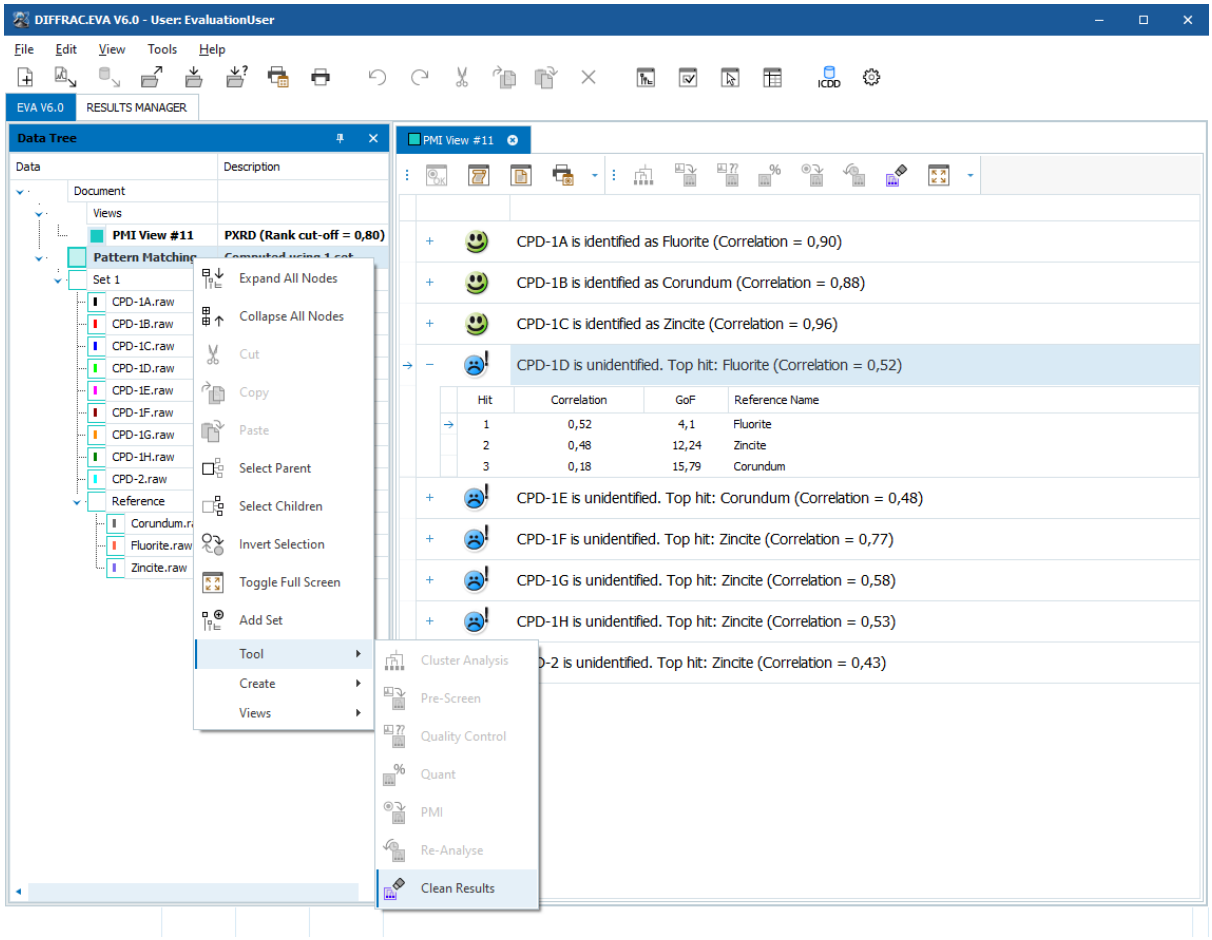
For customization of toolbars, we refer to the EVA manual.

Creating an EVA project file

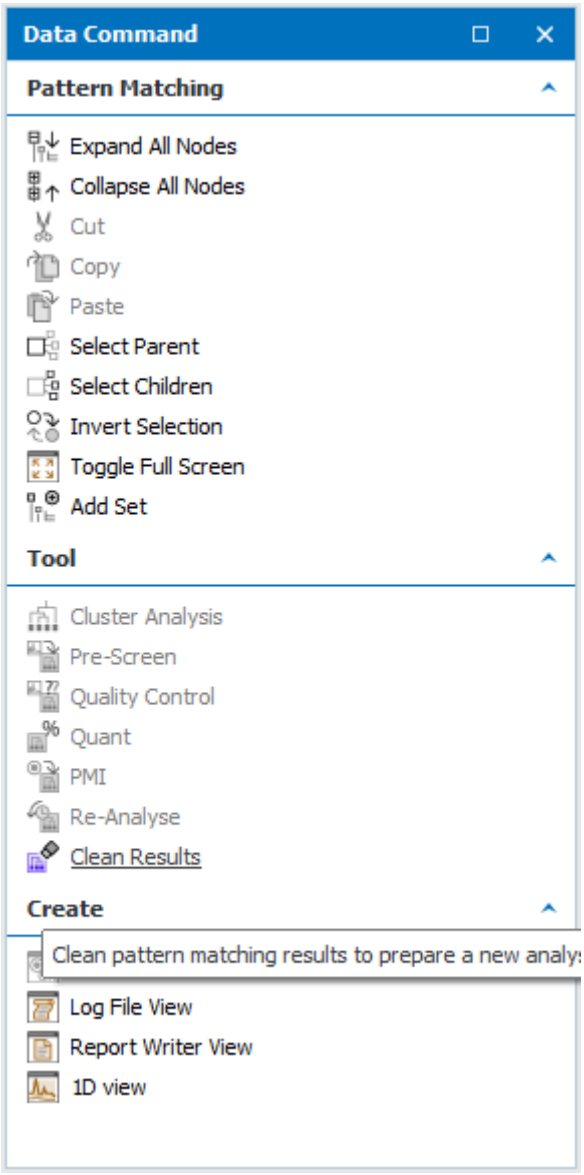
For creating an EVA project, it is mandatory to run a PMI for one unknown sample at least. It is however advised to test several samples in order to tune the correlation cut-off. To do so, the current PMI run must be prepared for a new evaluation. Highlight the pattern matching node in the data tree, use the toolbar icon



Analogously, a right mouse-click on the pattern matching node gives access to the Clean Results command under the Tools section



or use the Data Command panel (F4 function key).

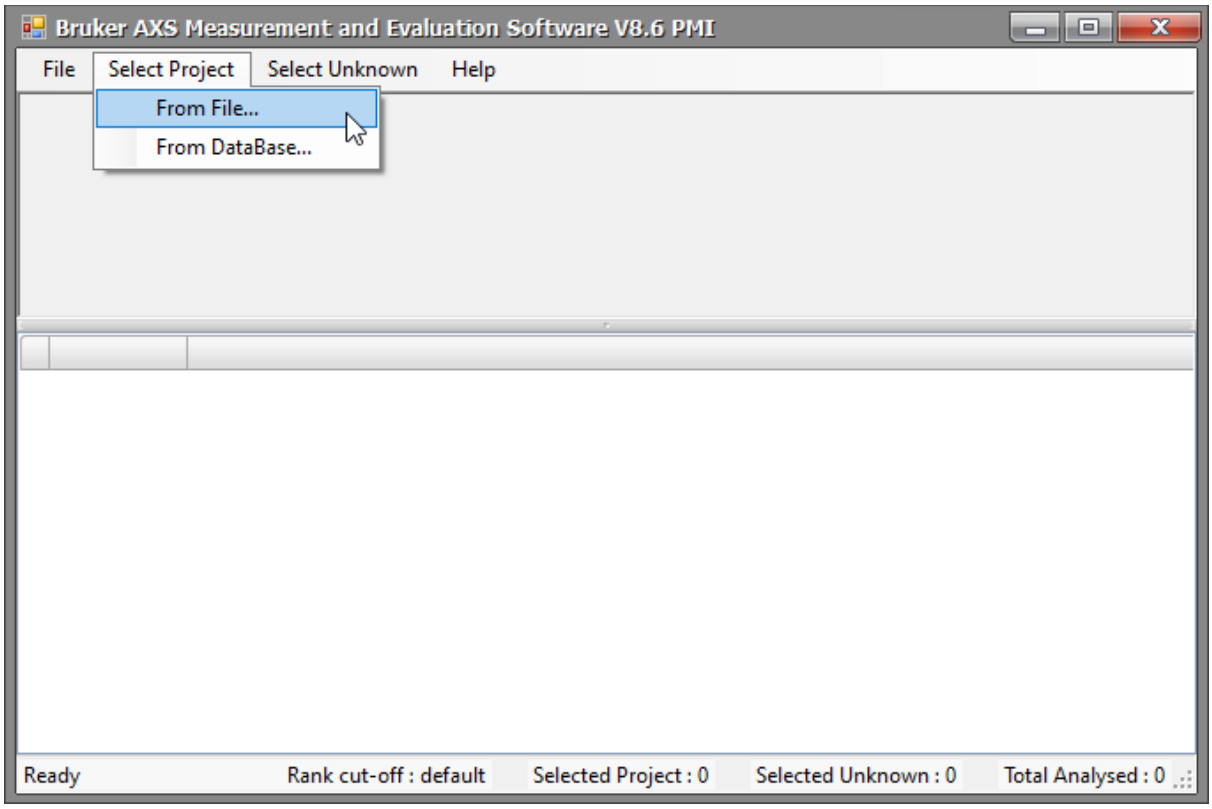


After adjusting the correlation cut-off and the necessary pre-processing steps re-run the project and save the EVA file, either to the database or the file system, and close EVA.

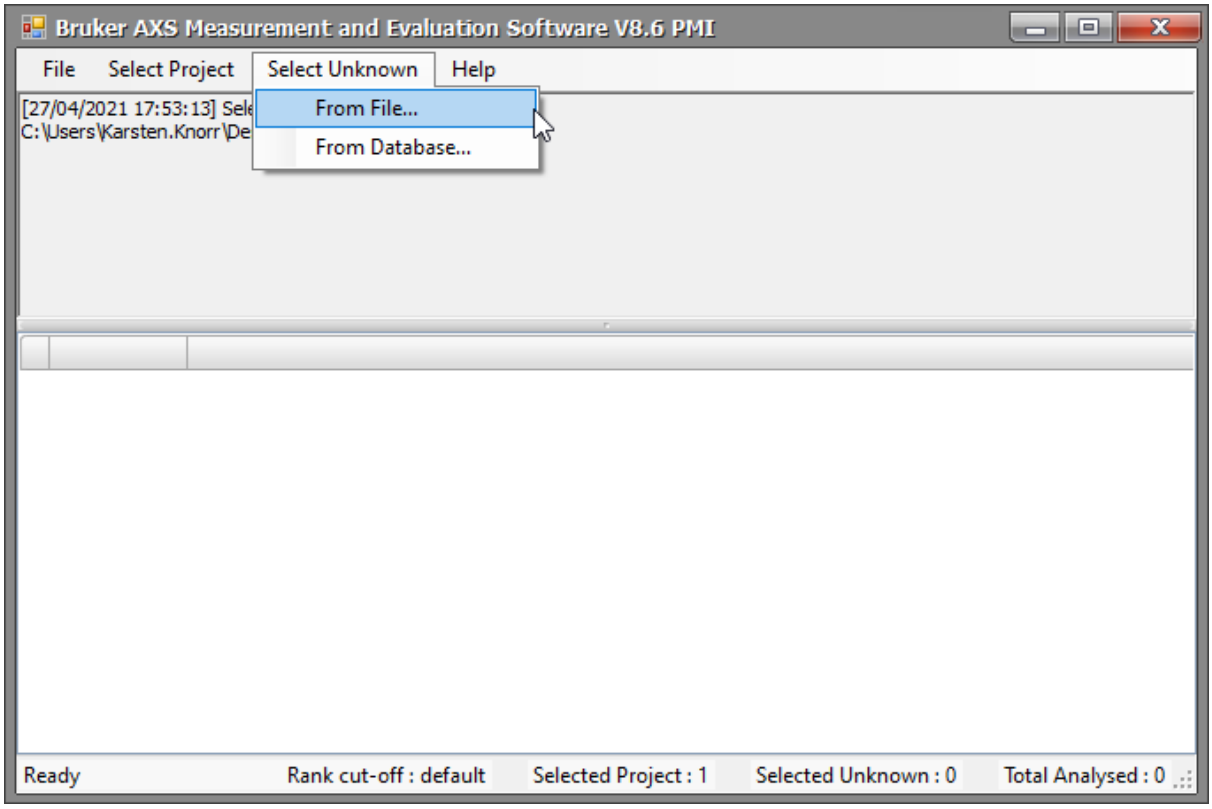
Operator mode PMI

Open the program PMI from the DIFFRAC.EVA V6.0 group in the Windows start menu. Use

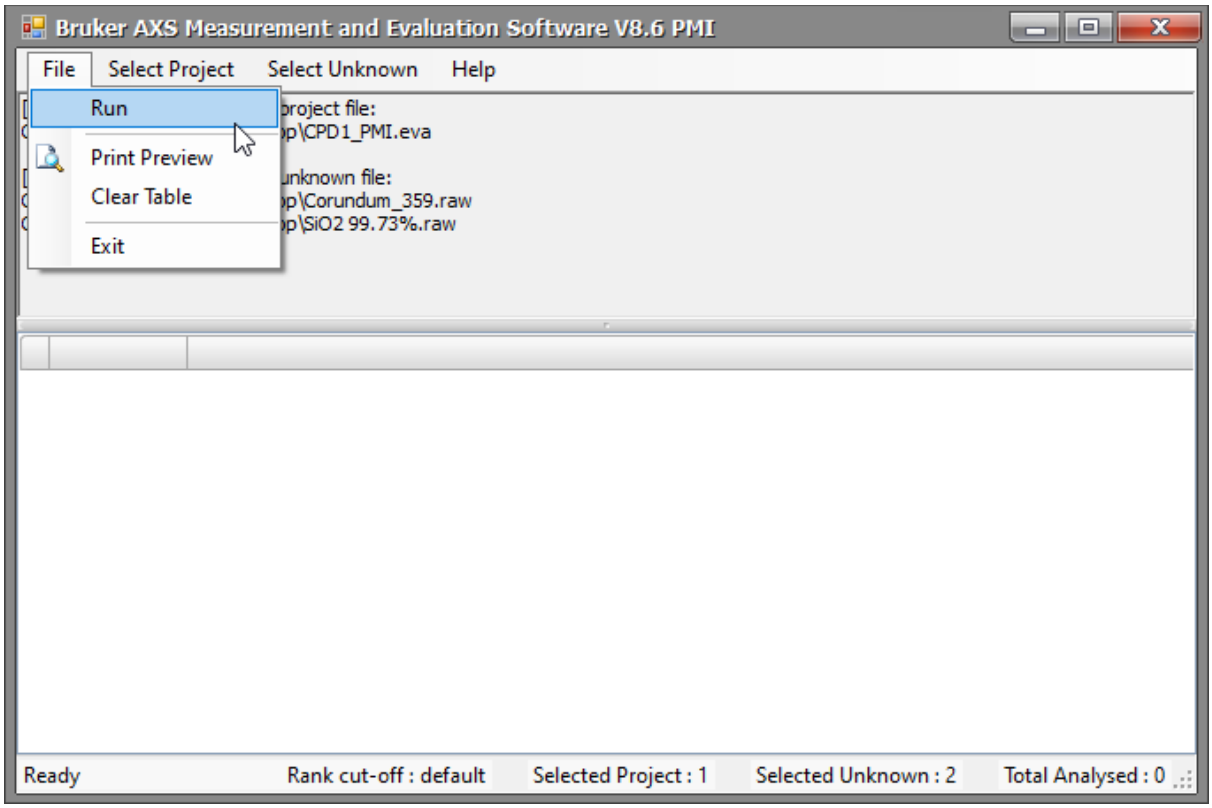
- Select Project to load an EVA PMI project from either the file system or the database



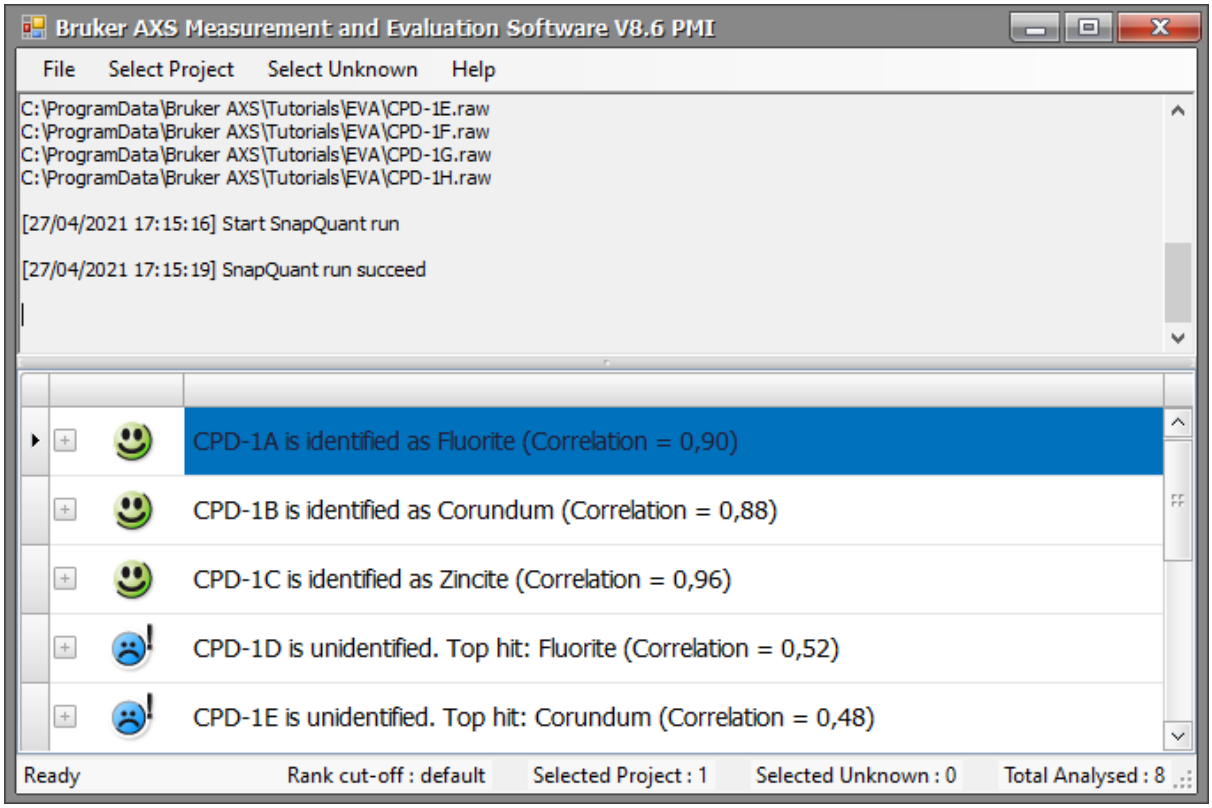
- Select Unknown to load the measurements to be identified from either the file system or the database



- Execute the evaluation by pressing the Run button inside the File menu.



The upper part of the PMI window contains a log file of the actions. The PMI results are shown in the lower part of the PMI.exe screen. The individual identification results may be expanded the same way as within the PMI view of the EVA framework. A print preview of the results can be accessed via the File menu.



Prior to reporting further samples can be added to the analysis, or the results table can be cleaned after the analysis prior to further evaluations.

On-line PMI

The on-line PMI analysis is controlled by the SnapQuantPMI script. When a measurement is set up in Start Jobs, the previously prepared EVA file must be chosen and “SnapQuantPMI” must be selected as evaluation script.

The PMI analysis is performed automatically after the measurement is finished. The result is displayed within the Results Manager.

Additions and Corrections to the Manuals (V7.0)

Peak and Pattern Fit

The peak and pattern fit code has been revamped to improve the pattern fit capabilities.

Peak Profile Functions

The Pearson-VII and split Pearson-VII functions have been added because they have a better convergence behavior than the Voigt functions in some circumstances.

The formerly available (split) pseudo-Voigt function has been replaced by the (split) Voigt function.

Peak Fit

Peak lists are now the place where the peak profile function is determined.

If is important to create a peak list if individual peak fitting is required.

A DIF which can be created from the peak search tool is treated as a pattern with the peaks having a fixed position relative to each other, whereas they are allowed to move relative to each other when they are members of a peak list.

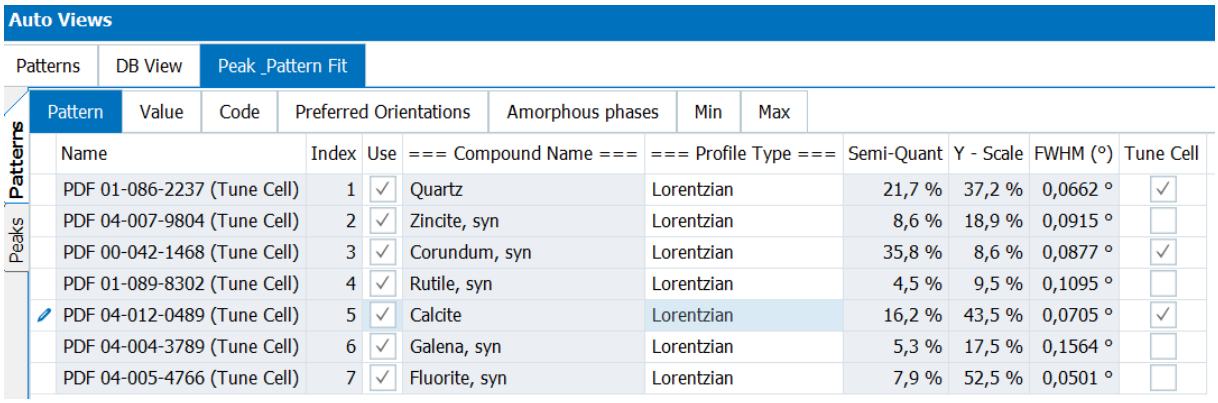
Pattern Fit

The fitting algorithm has been rewritten and improved to allows refining more pattern parameters:

- Lattice Parameters
- Preferred Orientation
- Amorphous content

Lattice Parameters

The lattice parameter fit is activated with the “Tune Cell” check box in the automatic Pattern table which is displayed during the fit:

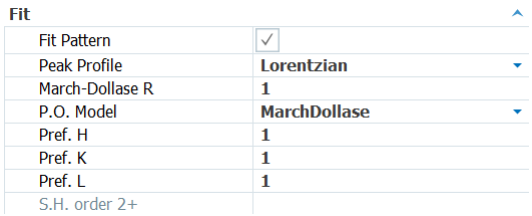


Name	Index	Use	=== Compound Name ===	=== Profile Type ===	Semi-Quant	Y - Scale	FWHM (°)	Tune Cell
PDF 01-086-2237 (Tune Cell)	1	✓	Quartz	Lorentzian	21,7 %	37,2 %	0,0662 °	✓
PDF 04-007-9804 (Tune Cell)	2	✓	Zincite, syn	Lorentzian	8,6 %	18,9 %	0,0915 °	
PDF 00-042-1468 (Tune Cell)	3	✓	Corundum, syn	Lorentzian	35,8 %	8,6 %	0,0877 °	✓
PDF 01-089-8302 (Tune Cell)	4	✓	Rutile, syn	Lorentzian	4,5 %	9,5 %	0,1095 °	
PDF 04-012-0489 (Tune Cell)	5	✓	Calcite	Lorentzian	16,2 %	43,5 %	0,0705 °	✓
PDF 04-004-3789 (Tune Cell)	6	✓	Galena, syn	Lorentzian	5,3 %	17,5 %	0,1564 °	
PDF 04-005-4766 (Tune Cell)	7	✓	Fluorite, syn	Lorentzian	7,9 %	52,5 %	0,0501 °	

Alternatively, the “Tune Cell” pattern property can be used.

Preferred Orientation

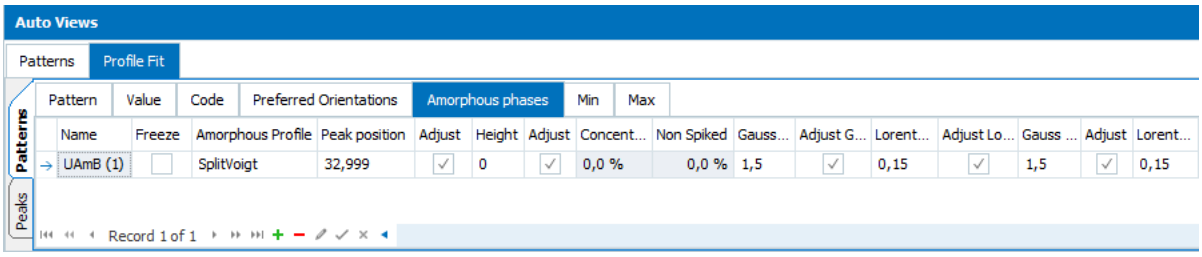
The preferred orientation fit is controlled by new pattern properties:



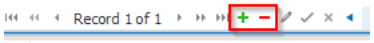
The March-Dollase model and spherical harmonics up to the 6th degree are available. A Pawley fit without line height constraints is also an option.

Amorphous Content

One or more amorphous “phases” can be added with the buttons on the “Amorphous Phases” tab.

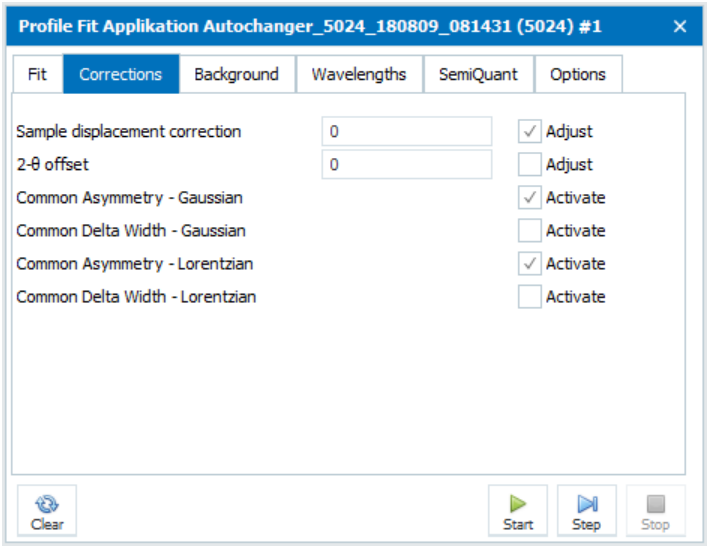


Amorphous entries can be added and removed with the control at the bottom of the Amorphous phases tab:

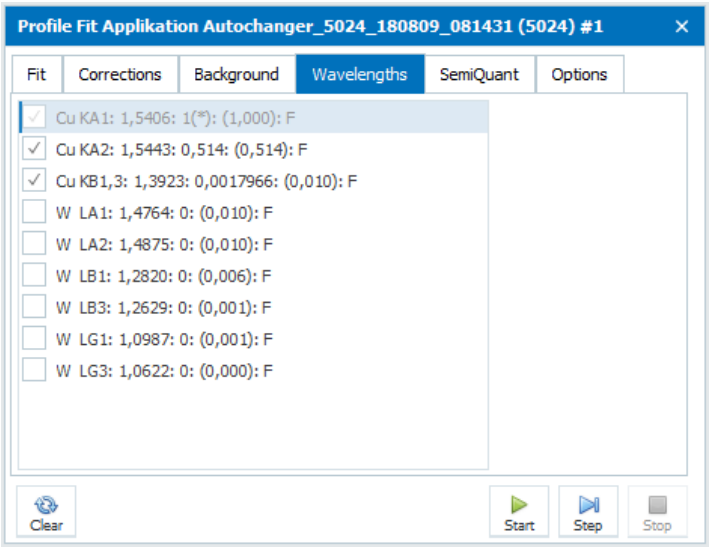


Profile Fit Tool

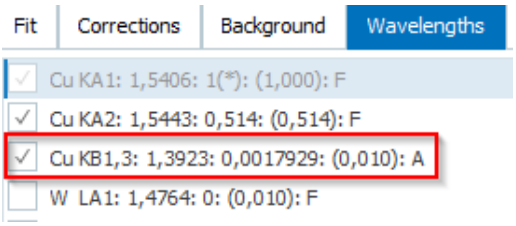
The profile fit tool has a new Corrections tab to control various fit parameters:



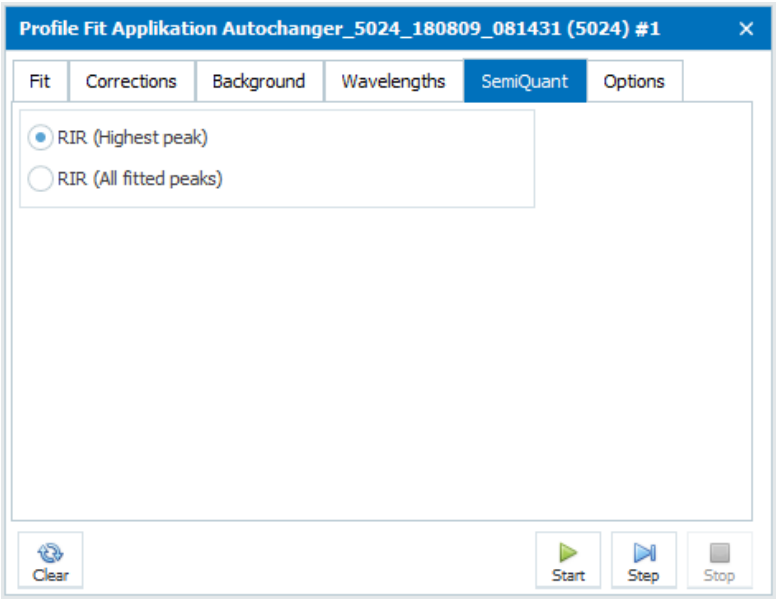
The new Wavelengths tab controls which wavelengths are used in the fit:



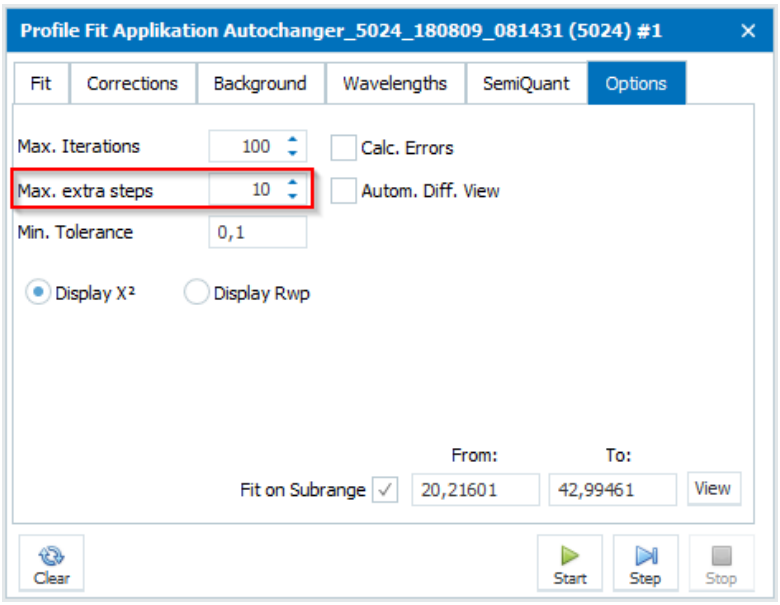
Repeated clicks on a wavelength entry allow switching from a fixed line height ratio to Ka1, marked by an “F”, to a refinable ratio, marked by an “A”:



Finally, there is a new SemiQuant tab to select how the semiquantitative analysis is carried out:



Note: The max. number of iterations on the Options tab (default: 100) is only an approximate value and may be exceeded by some iterations. To limit these extra steps, enter a value. The default is 10:



A “Clear” button at the bottom left allows a reset of the fit results in case of error.

Parameter Limits and ESD (Estimated Standard Deviation)

The Profile Fit Auto View displays the calculated parameters as tables.

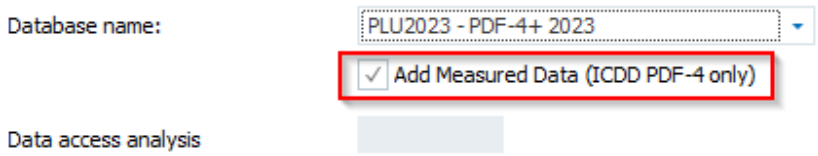
If a parameter has reached a limit or has another condition it is marked as follows:

- “>>”: the upper limit was reached, no esd was calculated
- “<<”: the lower limit was reached, no esd was calculated
- “(<->)”: the value is not significant because it is smaller than the calculated esd
- “(*)”: fixed value, no esd was calculated
- “?”: the esd could not be calculated

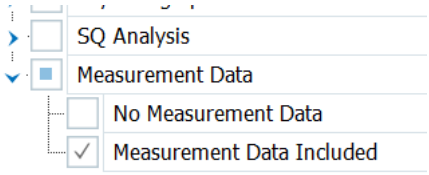
Measured Data in PDF-4 Databases

The DSRD Compiler collects all raw data which are part of the PDF-4 and places them in the precompiled files for quick access.

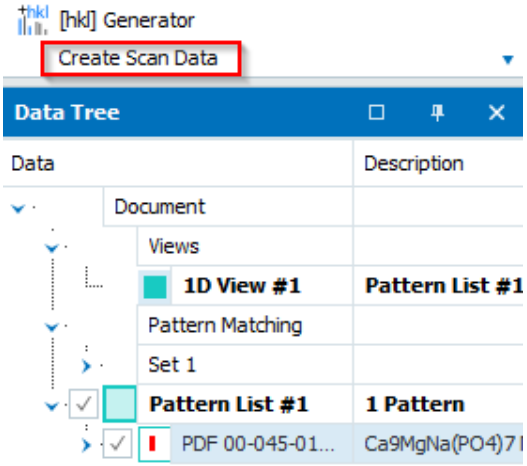
Check the marked box in the compiler to include the data:



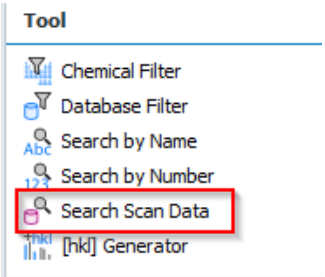
A new database filter allows extracting all patterns during a database search which have the original raw data attached:



If such a pattern is added to the document, it will have a new command “Tool | Create Scan Data” to import its raw data:



The new direct search for scans is available on the document and the scan list levels:

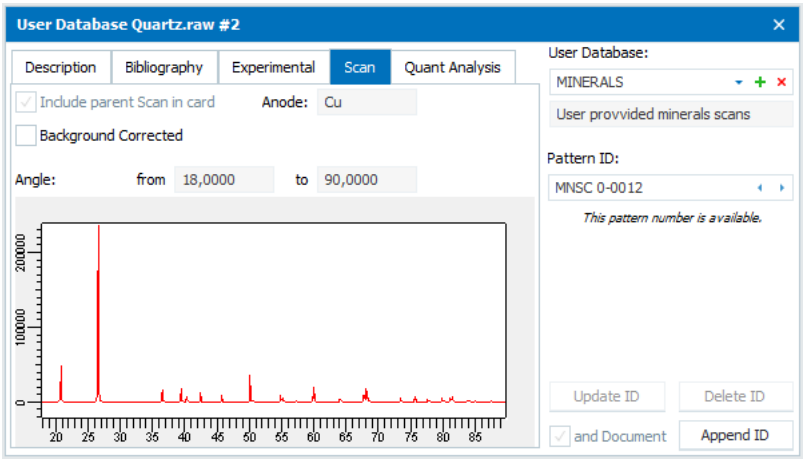


This command is also available for Pattern Matching Sets and References.

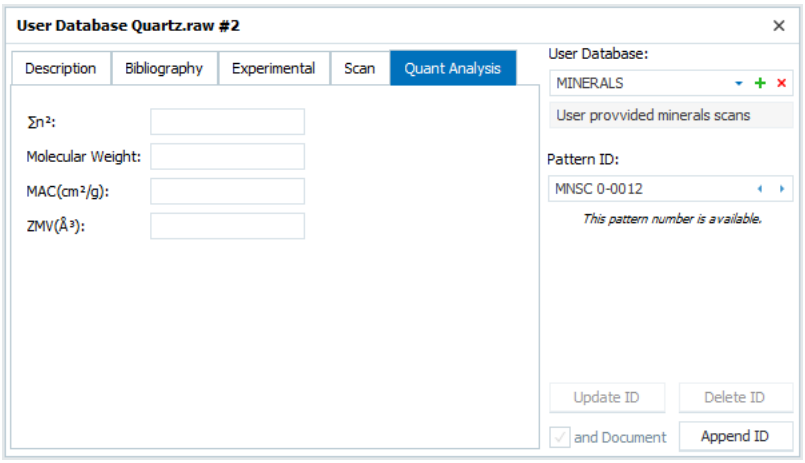
Measured Data in User Databases

The capabilities of user databases have been extended to include measured data. Consequently, user databases can be utilized to store reference measurements or other measurements of interest.

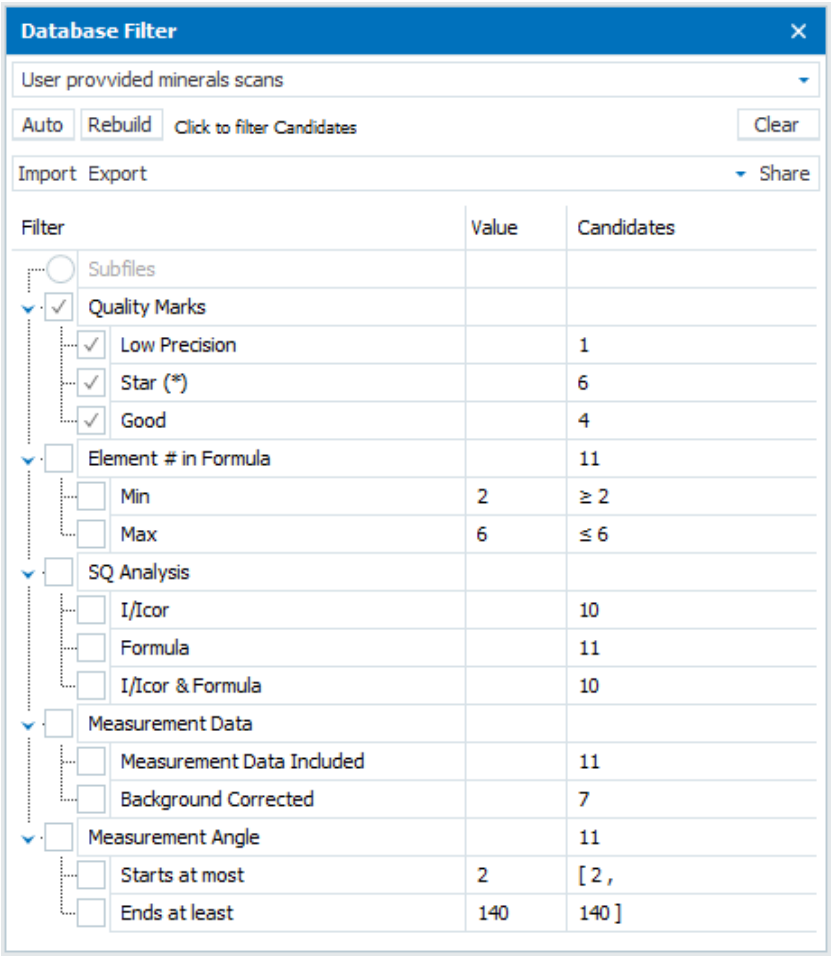
To add a scan to a user database, employ the newly available “Tool | User Database” command on the scan. The user Database tool opens with some adaptations for scans:



The scan can be marked as “Background Corrected” and on the “Quant Analysis” tab some extra properties can be entered that are useful for quantitative calculations:



While searching in user databases that contains scans, filtering is possible, e.g.:



The filters are calculated dynamically depending on the provided data.

Search/Match

Search/Match: Consideration of Kβ Peaks

The automatic search/match algorithm takes Kβ peaks into account if the Kβ/Kα ratio is defined in the scan properties.

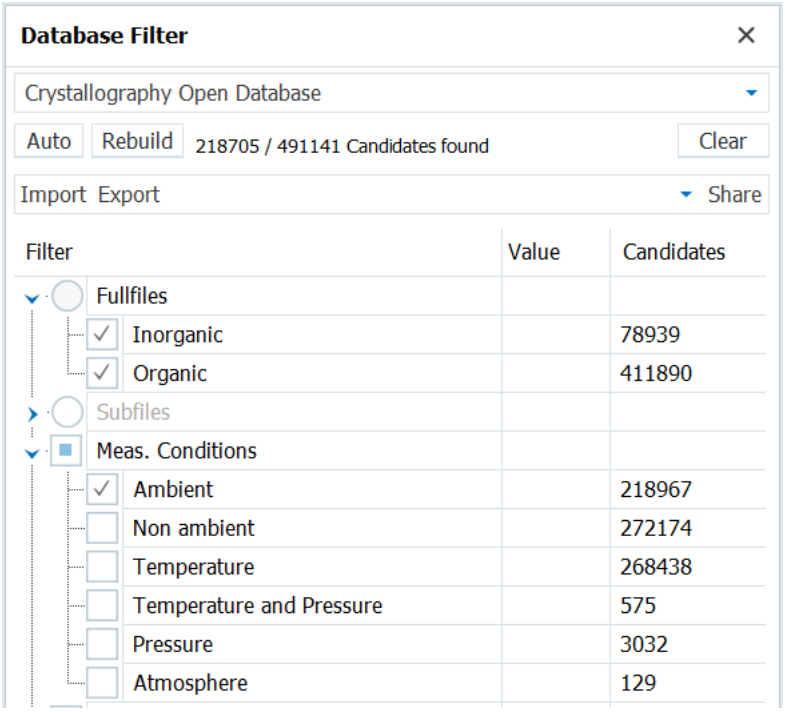
Search List

If an automatic search was performed and the setting “Auto-Search Results as Candidates” was selected, the likely candidates are marked green in the first column.

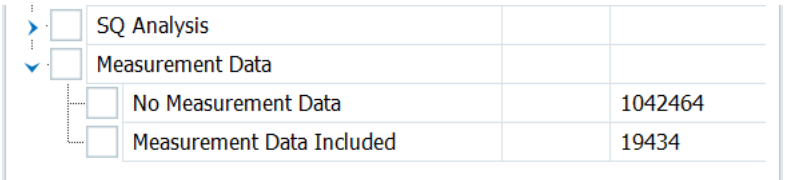
A new button below the search list allows exporting all search results as TOPAS structure files.

Database Filter: Measurement Conditions in PDF and COD databases

The database filter contains several measurement conditions for non-ambient measurements:



A new filter is available to filer patterns which have their measurement data attached in the PDF database:



PDF Processing

Introduction

Pair distribution function (PDF) analysis is an analytical technique that can provide local structural information from disordered materials. A PDF represents the number of times a specific distance occurs in a structure, weighted by the masses of the atoms at that distance. This information is encoded in both the Bragg diffraction and diffuse scattering signals. To obtain an experimental PDF a so-called “total scattering” measurement is made, which captures both of these signals. The raw diffraction measurement is then mathematically transformed into the PDF by the following process: several corrections are applied, including absorption, polarization, Compton scattering, background scattering, and X-ray fluorescence; the data is normalized to an absolute scale; and finally, a Fourier transformation is applied to generate the real-space PDF.

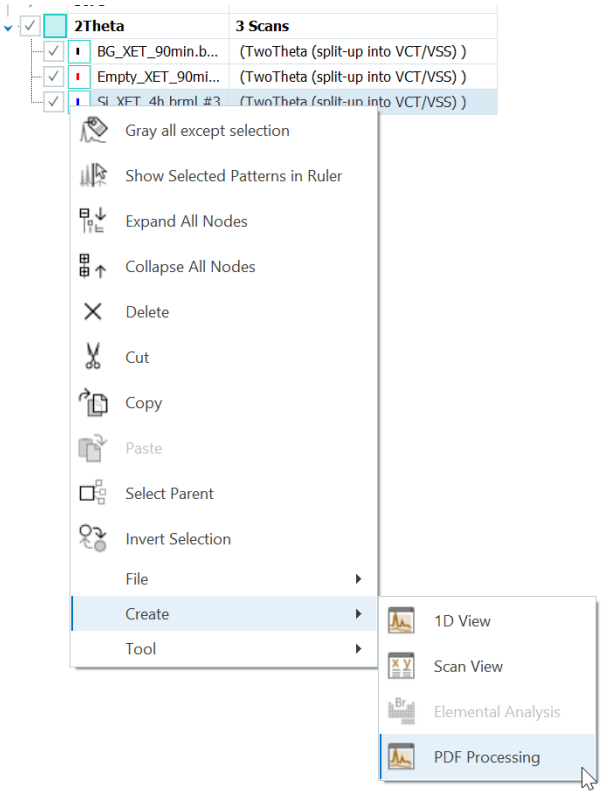
Starting from license level 7, the DIFFRAC.EVA software package offers the functionality to obtain the atomic PDF from angle-dispersive X-ray diffraction experiments. This feature supports both laboratory X-ray data and synchrotron X-ray data. The implementation utilizes code sourced from the pyTSRedX software package, developed by Olivier Masson from the University of Limoges. For more detailed information, please refer to the pyTSRedX manual, which provides further insights into the underlying algorithms and procedures.

PDF data reduction example

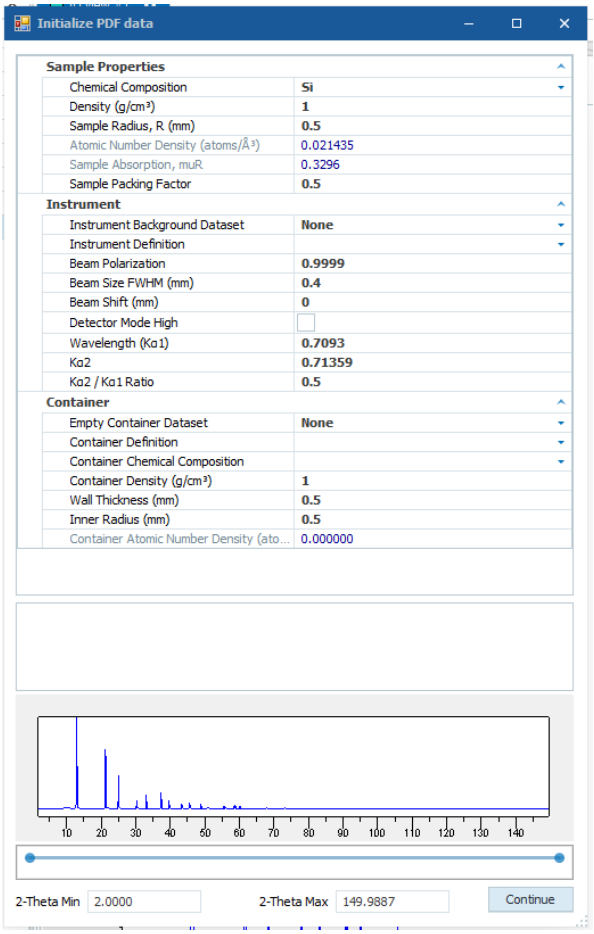
In this example, we will create an experimental PDF from raw diffraction data of silicon powder. The data was collected on a D8 ADVANCE diffractometer in capillary geometry, equipped with Mo sealed tube anode, focusing Goebel mirror optic, and a LYNXEYE XE-T detector in high count rate mode.

First, load the 3 datasets “Si_XET_4h.brml”, “Empty_XET_90min.brml”, and “BG_XET_90min.brml” from the folder “C:\ProgramData\Bruker AXS\Tutorials\PDF Processing\Si Powder”. These are the sample data, empty capillary data, and instrument background data, respectively.

Next, select the silicon dataset in the Data Tree, and then choose *PDF Processing* from the *Create* context menu. This can be accessed either via the **Data Command** window or by right-clicking on the dataset of interest.

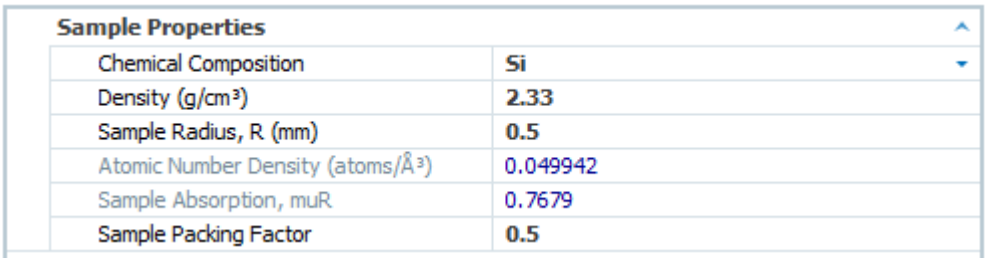


A new window called *Initialize PDF data* will appear, where all the information required to run the processing will be entered.



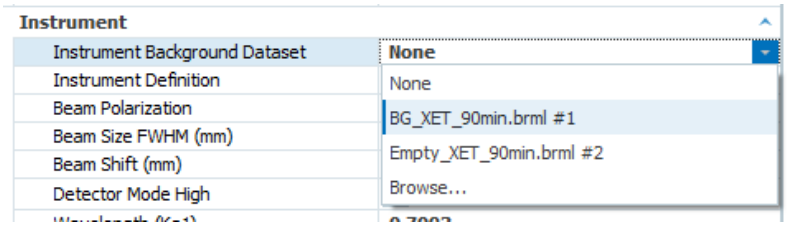
Some information is read from the scan and added automatically, if available (e.g., wavelength). The program will also attempt to read a chemical formula from the file name if one is included. In this example, “Si” was read before the first underscore and was added to the *Chemical Composition* under *Sample Properties*.

There are 3 main sections in the window: **Sample Properties**, **Instrument**, and **Container**. Sample Properties contains the information necessary to describe the sample for processing. At a minimum, the user should enter the *Chemical Composition* and the *Density* of the sample. If this is not known exactly, a good approximation is also sufficient. The evaluation will not run unless a valid chemical composition is entered.

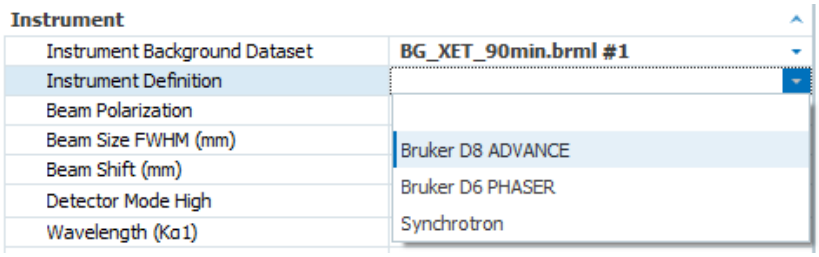


In this example, the sample is silicon powder measured in a 1.0 mm diameter glass capillary. The density of silicon is 2.33 g/cm^3, and the sample radius (0.5 mm) is half the capillary diameter. The *Atomic Number Density* and *Sample Absorption* are then automatically calculated based on the information in the first three fields (note: these fields are not directly editable). The *Sample Packing Factor* describes how efficiently the sample material is packed into the capillary. In general, an estimate of this value is sufficient, but it can also be determined by weighing the capillary before and after the sample has been filled. For powders, this is typically between 0.4 and 0.6, while for liquids this can be assumed to be 1.

In the **Instrument** section, all relevant information about the instrument used for the experiment is entered. The *Instrument Background Dataset* contains the scattered intensity measured from the empty environment (i.e., without sample or container), which is predominantly air scattering. This dataset can be useful for absorption calculation when sample absorption is high but is generally not needed. If this dataset has already been loaded into EVA, it can be selected in the dropdown menu; otherwise, the data can be loaded by choosing *Browse…* from the dropdown menu. For this example, select *BG_XET_90min.brml*.

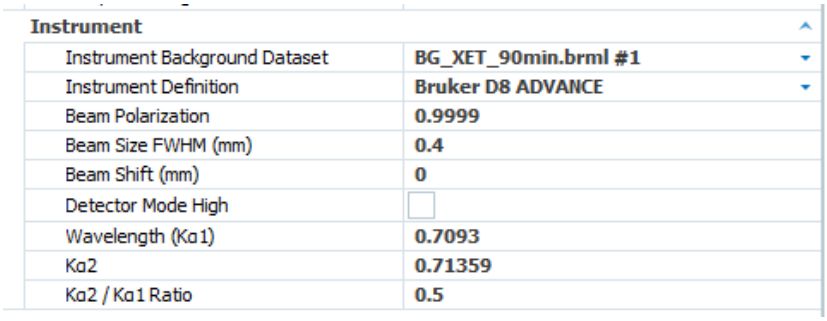


Other instrument information, like polarization, beam size, beam shift, and detector mode, can be stored in an *Instrument Definition*. Some instrument definitions are provided with EVA; these can be modified or more can be added by editing the file **Instrument.json** in the folder C:\ProgramData\Bruker AXS\DSRD6. The appropriate instrument definition can be selected using the dropdown menu. For this example, choose *Bruker D8 ADVANCE*.

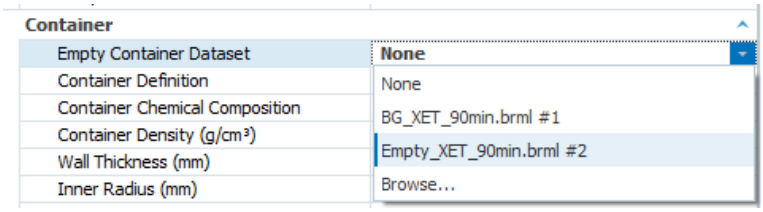


Wavelength is the X-ray wavelength used in the experiment and is used to convert the data into Q-space. For laboratory data this is the Ka1 wavelength. *Wavelength*, *Ka2*, and *Ka2/Ka1 Ratio* are loaded automatically when the anode information can be read from the data. In this example, the data was collected with a Mo anode. For monochromatic data, *Ka2* and *Ka2/Ka1 Ratio* can be set to 0.

Once all data has been entered, the Instrument section should look like this:

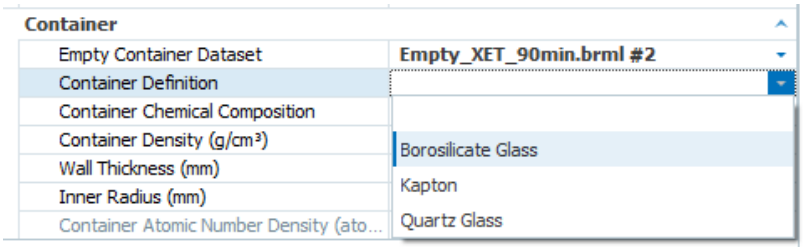


The **Container** section includes all the information about the sample container. The *Empty Container Dataset* is the measurement of the empty capillary. If this dataset has already been loaded into EVA, it can be selected in the dropdown menu; otherwise, the data can be loaded by choosing *Browse…* from the dropdown menu:

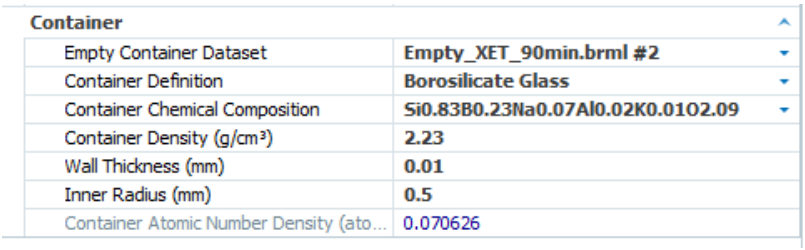


The most common capillary materials have been pre-defined in *Container Definition*. As with the instrument, container definitions are included in the file **Container.json** in C:\ProgramData\Bruker

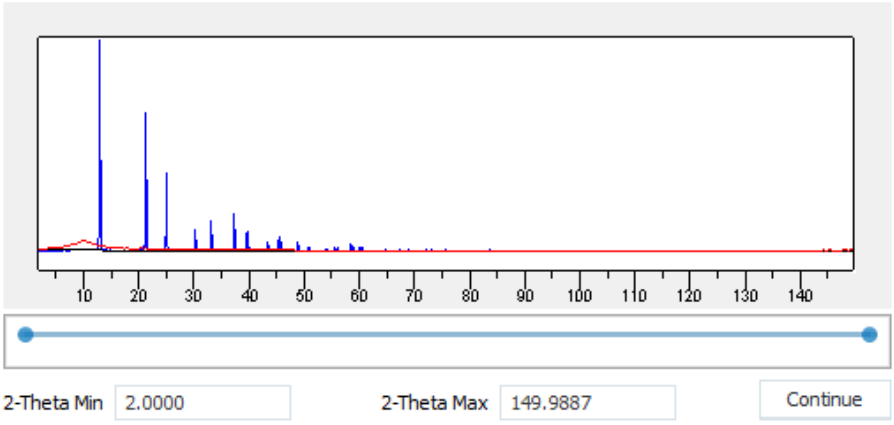
AXS\DSRD6, which can be edited as required. This example was measured in a 1.0 mm glass capillary, so choose *Borosilicate Glass* for the *Container Definition*:



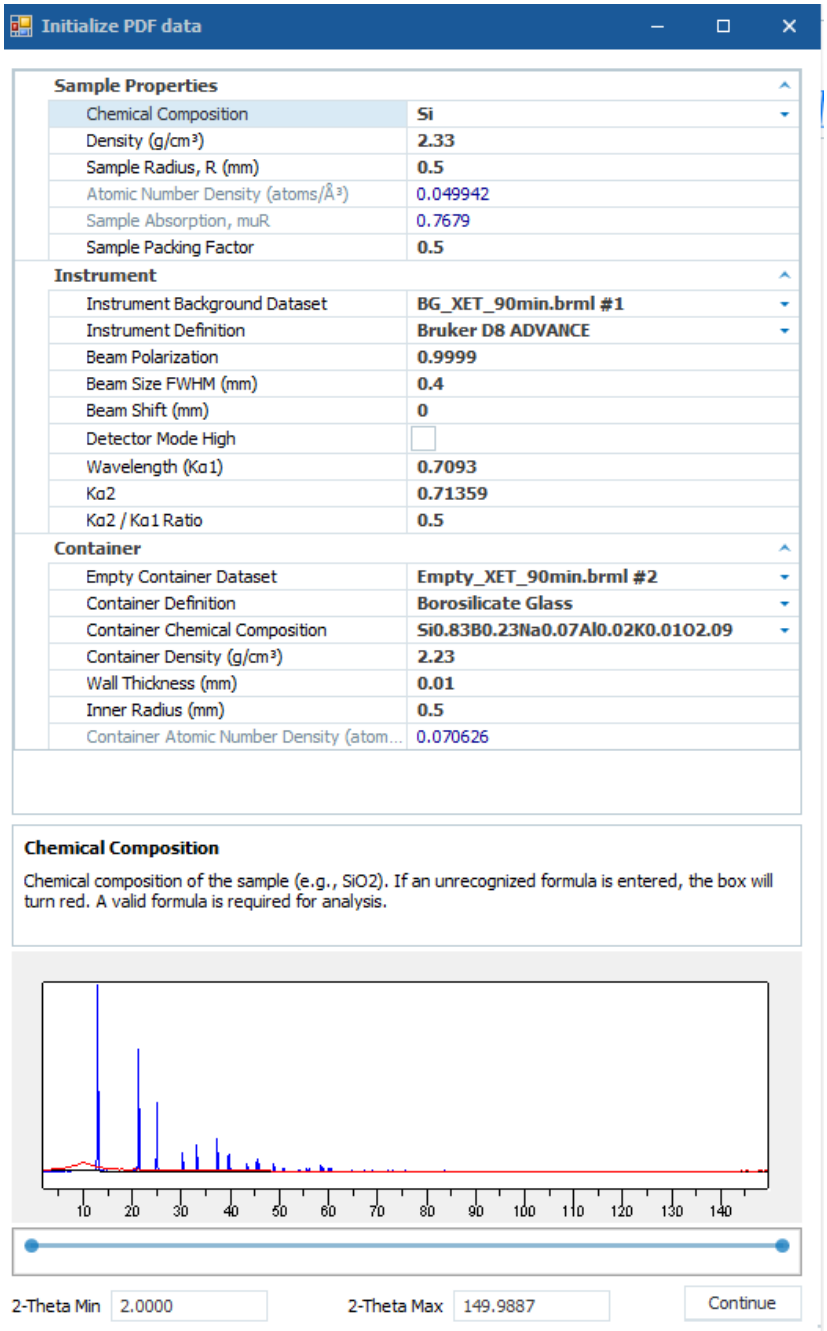
The Container properties will be automatically filled. The default inner radius of 0.5 is correct for this example.



At the bottom of the Initialize PDF data window is a small graph showing the data, a slider, and two additional parameters: *2-Theta Min* and *2-Theta Max*. The data range to use for the processing can be defined here; for example, any part of the data outside of this range will be ignored.

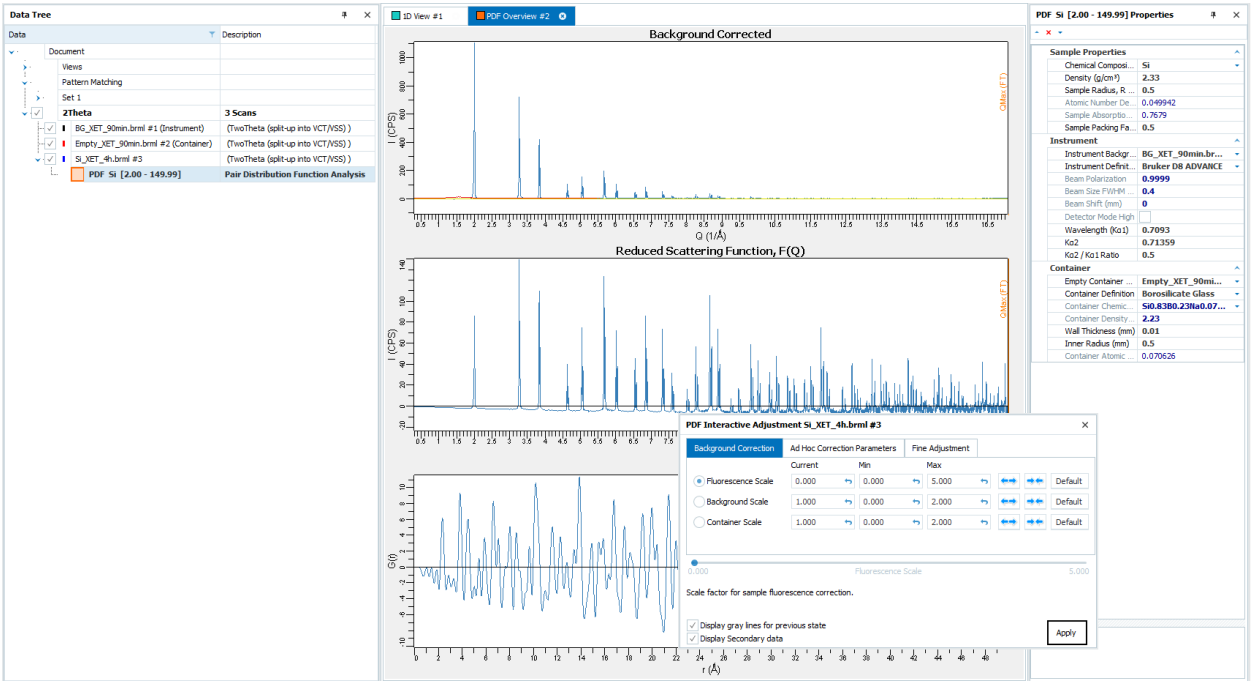


The *Initialize PDF data* window should look like the screenshot below. Once all the parameters have been entered correctly, click "*Continue*".



PDF Interactive Adjustment

PDF data reduction will run using the parameters from the *Initialize PDF data* window (initial calculation will take a few seconds), and a new view is created called **PDF Overview**, which contains three separate plots. An additional window is displayed, called **PDF Interactive Adjustment**.



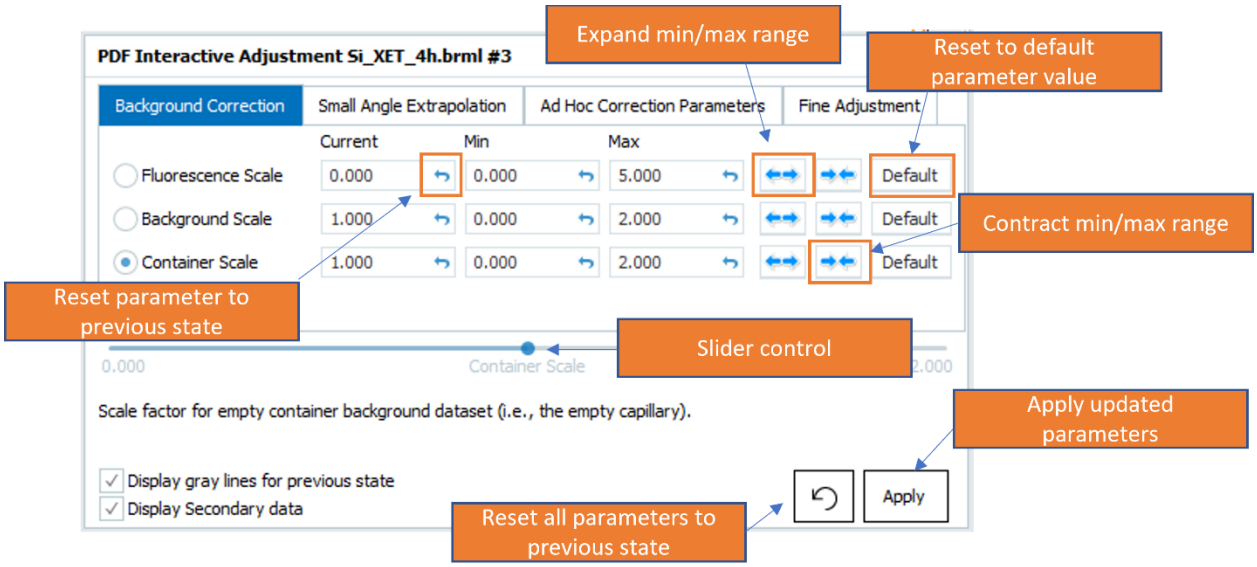
The plot at the top displays the background corrected data. This view will change depending on which tab is selected in the PDF Interactive Adjustment window.

The middle plot is the **Reduced Structure Function**, or **F(Q)**. This is the final form of the data after all the correction and normalization steps have been done and contains the structural information of the sample. This plot is important, as it is the first visual check that the data has been properly collected and that the counting statistics in the raw data are sufficient to produce a high-quality PDF.

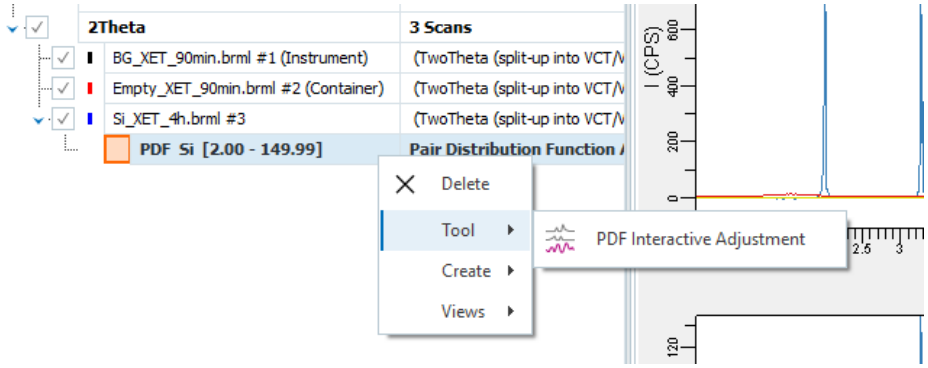
The bottom plot is the **PDF**, or **G(r)**. This is the sine Fourier transform of the reduced structure function.

The PDF Interactive Adjustment window

This window is where parameters can be adjusted by either using the slider control or by entering the desired value into the box next to the parameter under the column marked *Current*. The plots will automatically update when changing a parameter. To save a change permanently, click *Apply*. A brief description of the user interface components is shown below.

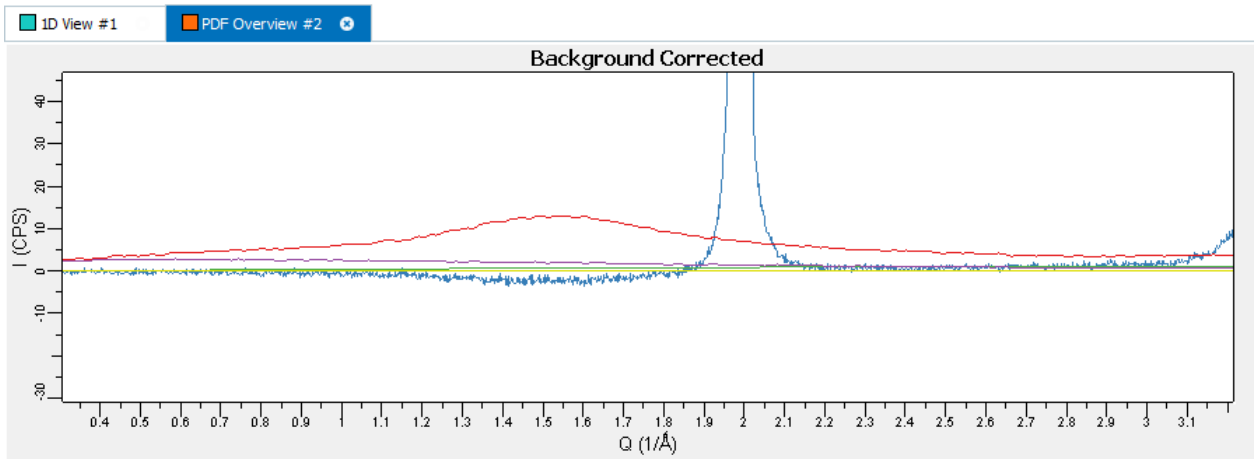


The window can be minimized by double clicking on the top banner or closed at any time. To reopen the window, select *PDF Interactive Adjustment* from the *Tool* context menu, by right-clicking on the PDF entry in the Data Tree.

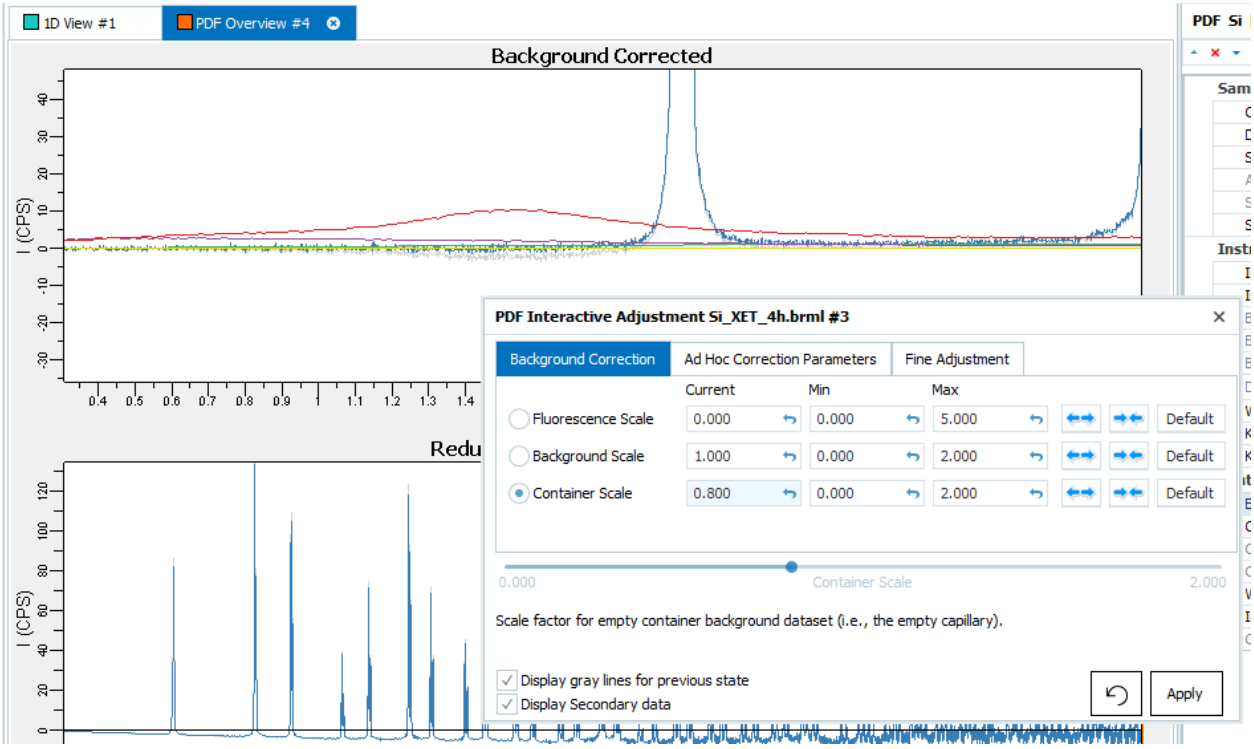


Background Correction

Select the **Background Correction** tab in the PDF Interactive Adjustment window and zoom in on the top plot around the first peak. There are several datasets displayed in this plot, including the empty capillary data (red curve), the instrument background (purple), and the background corrected sample data (blue). The additional curves can be turned off by deselecting *Display Secondary data* checkbox.

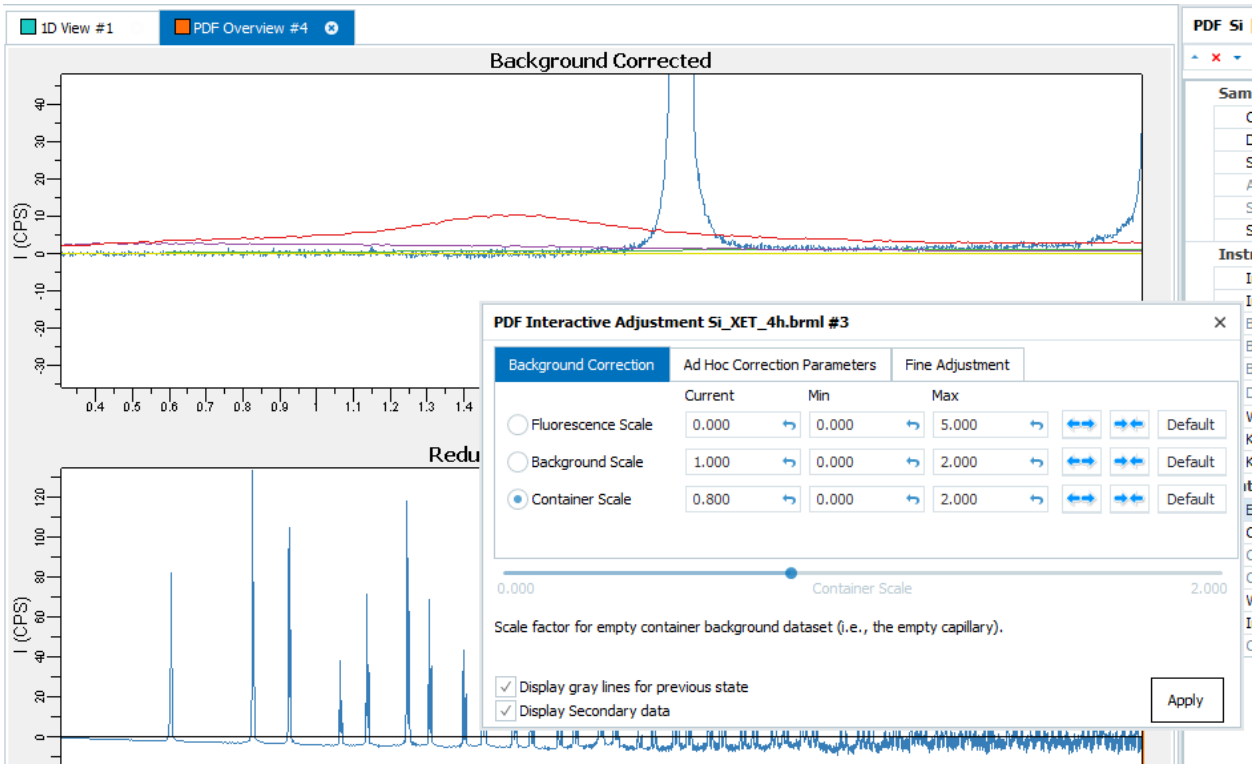


Change the scale factor of the empty capillary data to approximately 0.8 by adjusting the parameter *Container Scale* using the slider control in the PDF Interactive Adjustment window.



The plots will update using the new parameters. The previous state of the data is displayed as a faint gray line for visual reference. This can be turned off by deselecting the box *Display gray lines for previous state* checkbox.

To save the changes and discard the previous state, click *Apply* in the lower right corner of the PDF Interactive Adjustment window. All plots will update, and the parameters will be stored. To go back to the previous state, click the *Reset* button next to Apply.

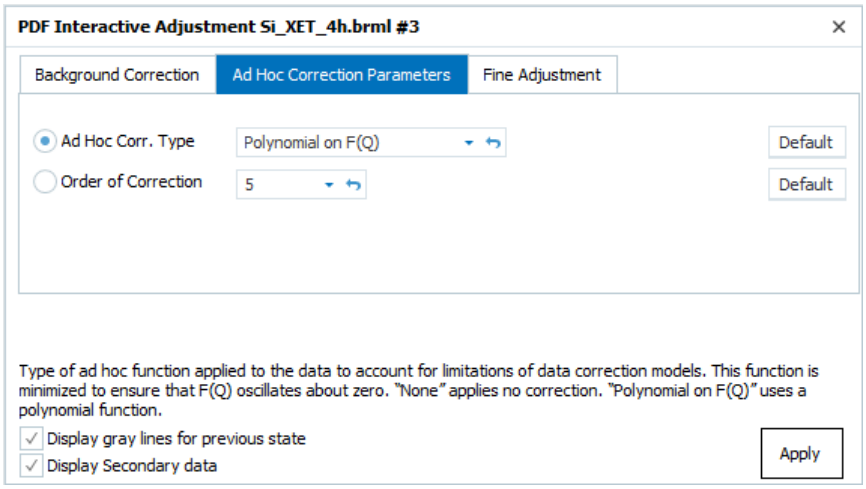


Ad Hoc Correction

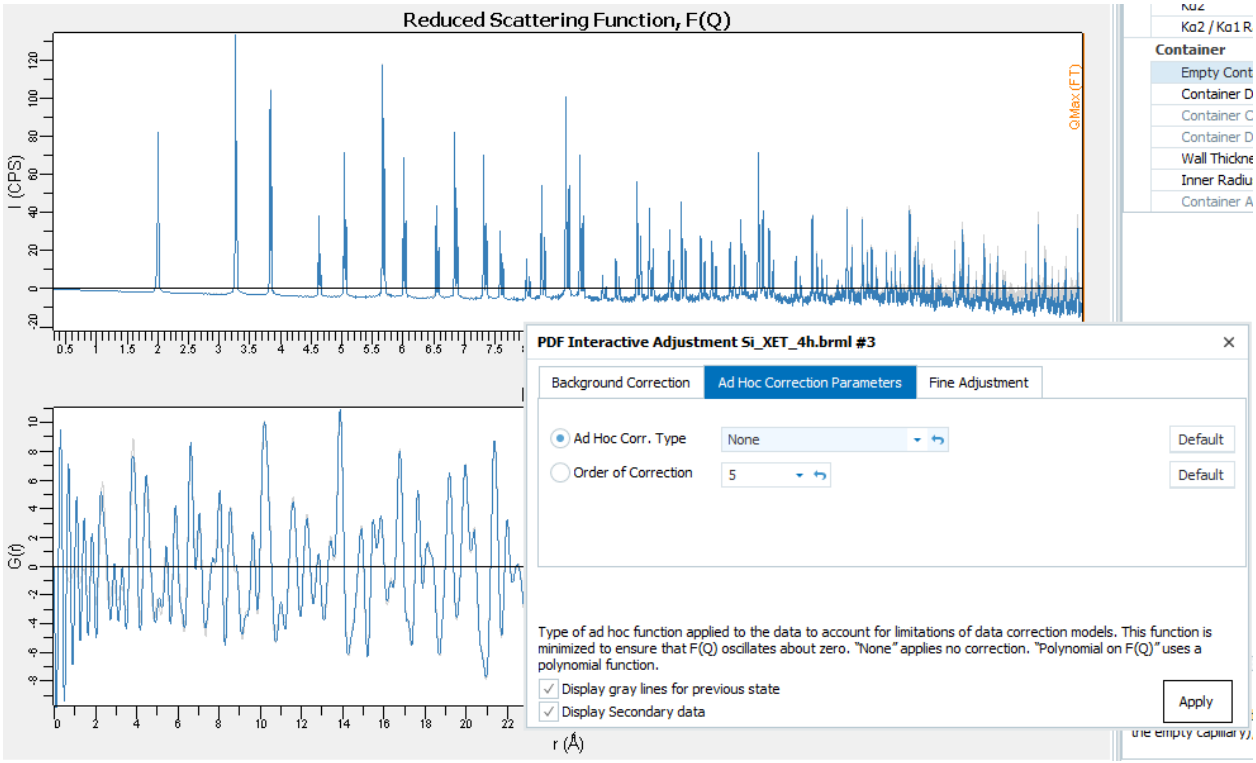
Despite all the care that one takes in processing the data, the corrections are always incomplete to some degree. An ad hoc correction has been implemented in DIFFRAC.EVA that will account for any limitations in the data correction models, without influencing the structural information in the PDF.

The expected behavior of F(Q) is well known. Ideally, F(Q) will oscillate around zero and go to zero at both $Q = 0$ and $Q = Q_{Max}$. The correction type *Polynomial on F(Q)* applies a polynomial fit to F(Q) which minimizes the difference between F(Q) and zero over the full Q range.

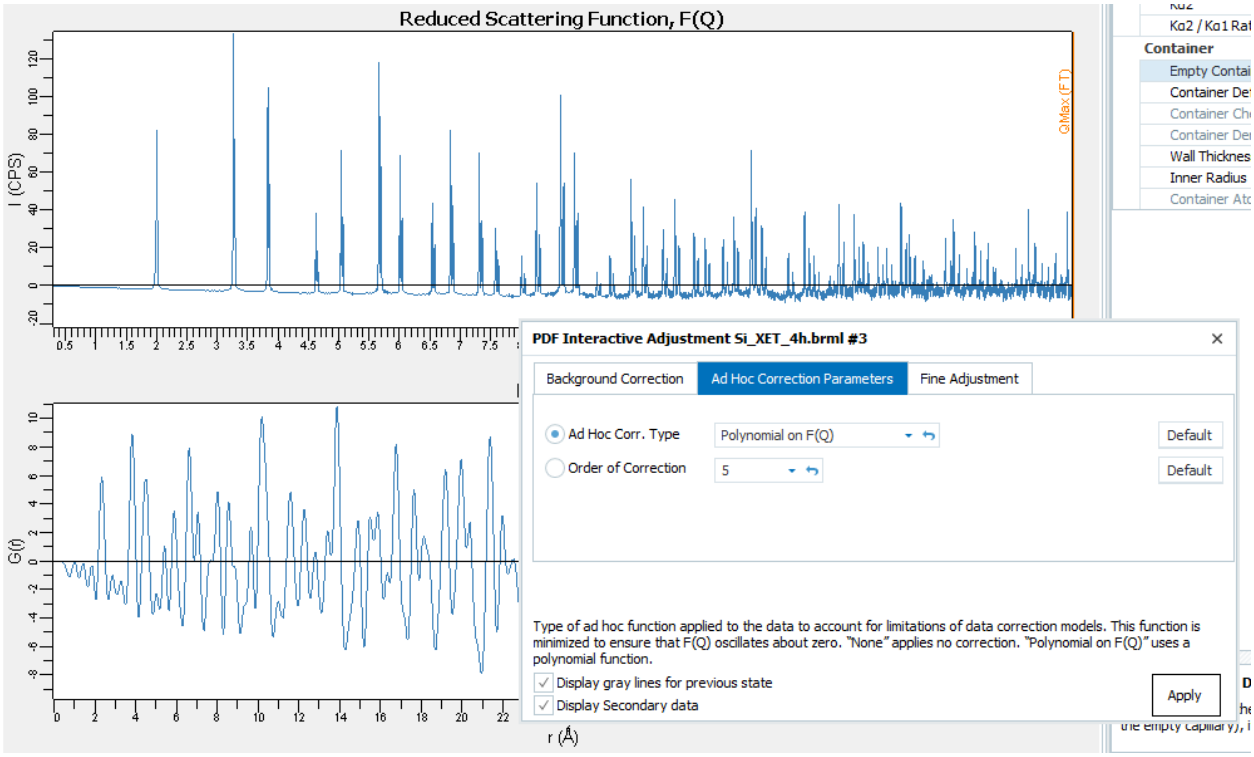
By default, this correction is applied to the data using a 5th order polynomial. The order of the polynomial can be set in the PDF Interactive Adjustment window, under the tab **Ad Hoc Correction Parameters**.



This correction can be turned off altogether, by selecting the option *None* from the dropdown menu next to *Ad Hoc Corr. Type*.



The effect is that F(Q) no longer goes to zero at high Q, which adds large ripples to the PDF at low r. These ripples are an artefact of the processing, contain no structural information and interfere with the low r peaks in the PDF. Changing the Ad Hoc Corr. Type back to *Polynomial on F(Q)* forces F(Q) to oscillate around zero and removes the spurious ripples in the PDF.

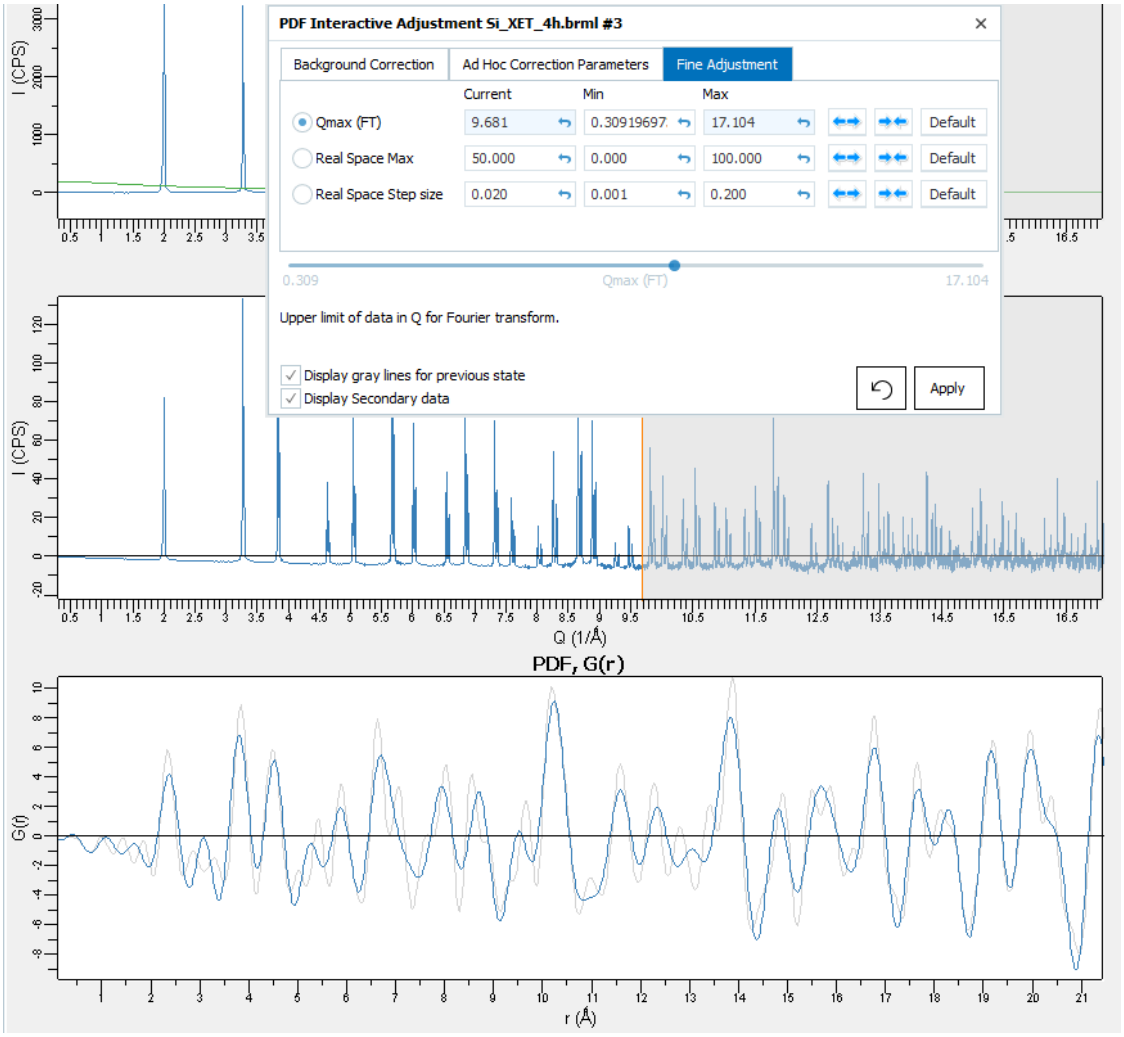


The polynomial order can be tuned by adjusting the *Order of Correction* parameter. The best choice for this value is the smallest number (i.e., lowest order) that leads to a well-behaved F(Q) (see above for definition of well-behaved F(Q)). Typically, an order between 3 and 5 is sufficient for most samples. For this example, use an *Order of Correction* of 5 and click *Apply*.

Fine Adjustment

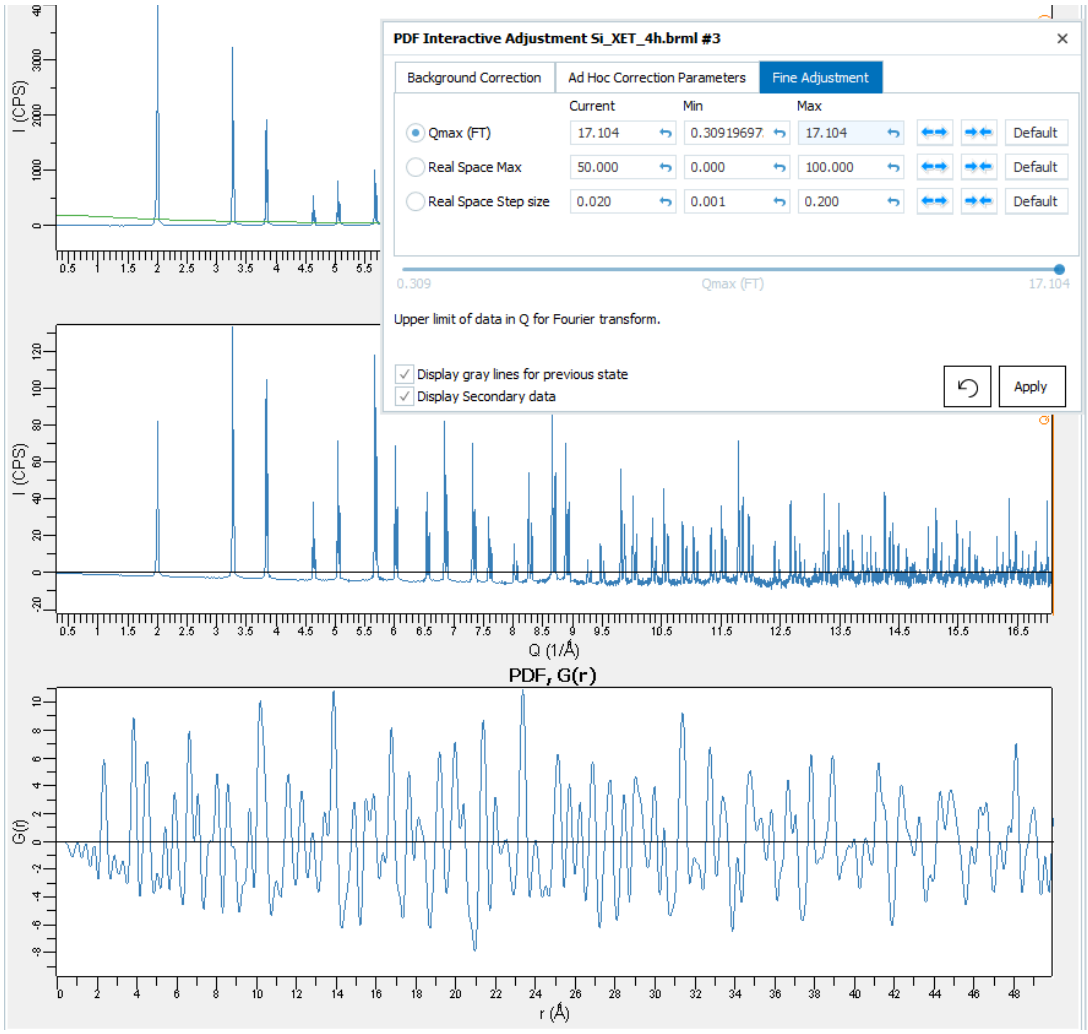
In the **Fine Adjustment** tab, there are several parameters for making fine adjustments to the PDF curve.

Qmax (FT) defines the upper limit in *Q* of the data used to produce the PDF by the Fourier transform. It is quite common for data at high Q to be noisier than that at low Q – this noise can manifest itself as additional peaks in the PDF, which contain no structural information. To prevent this from happening a Qmax (FT) less than the Q_{Max} of the data can be chosen to exclude the noisier data. In this example, the default values work quite well, and no further adjustment is required. However, feel free to try changing Qmax (FT) to see the effect on the PDF. For example, adjusting Qmax (FT) to 9.7 1/Å has a significant effect on the PDF, in particular its r-space resolution.

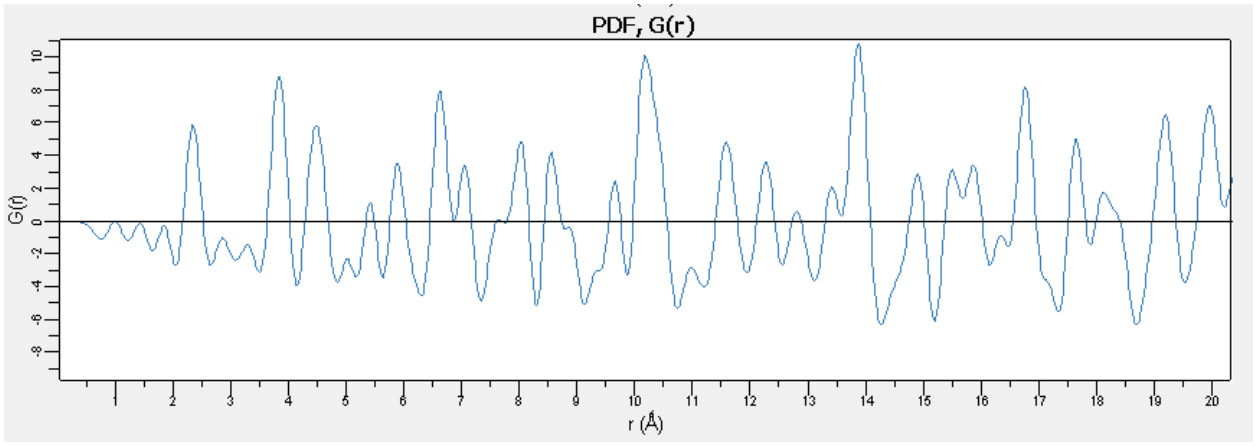


The PDF is calculated by default from 0 to 50 Å, with a step size of 0.02 Å. These settings can be adjusted by changing the parameters *Real Space Max* and *Real Space Step Size*, respectively.

To complete this tutorial, if you have changed any of the parameters in this window, click the *Reset* button to go back to the previous state, and then click *Apply*. The experimental PDF of silicon has been successfully created! The results can be now saved as an EVA document for future evaluation.

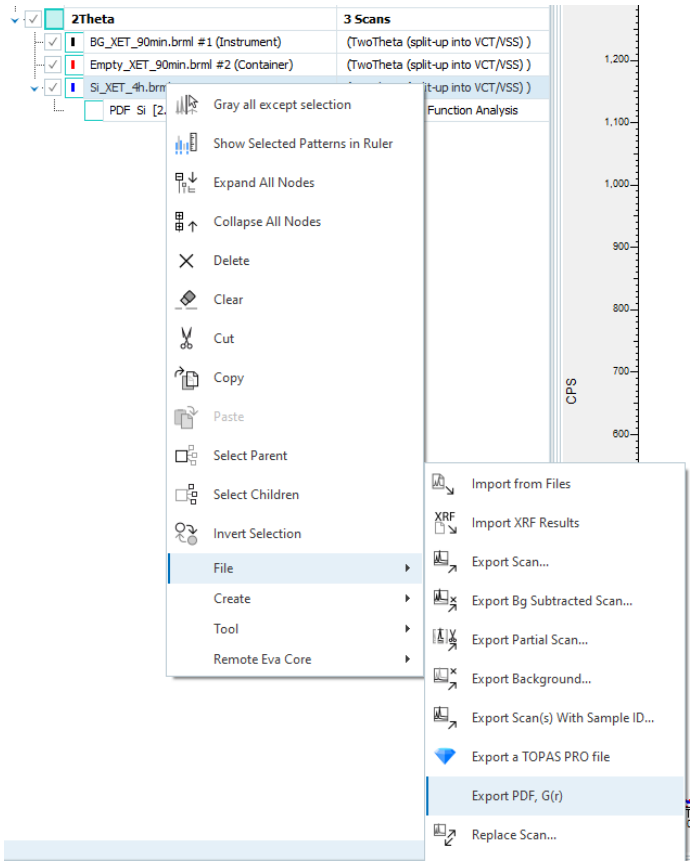


In the end, the results should look like the screenshot below. Based on the structure of silicon, the first peaks in the PDF should be expected at r = 2.32 Å, 3.84 Å, 4.50 Å, 5.43 Å, and 5.92 Å.



Exporting the PDF

To export the PDF curve, select the sample data in the Data Tree (in this example, “Si_XET_4h.brml”) by right clicking on it, and then select *Export PDF, G(r)* from the *File* context menu. The plot will be exported as an .xy file which can be read in any other analysis program, like DIFFRAC.TOPAS, for further structure analysis.



Additional tutorial examples

There are additional datasets in the tutorial folder, including a synchrotron dataset of Ni powder measured at the beamline P02.1, PETRA III, DESY and a dataset of amorphous silica measured on a D8 ADVANCE. All the information required for processing is included in a text file with the datasets.

Modifying Instrument.json and Container.json files

Additional instrument or container definitions can be added by modifying the files *Instrument.json* and *Container.json*, respectively. The files can be found in the folder C:\ProgramData\Bruker AXS\DSRD6.

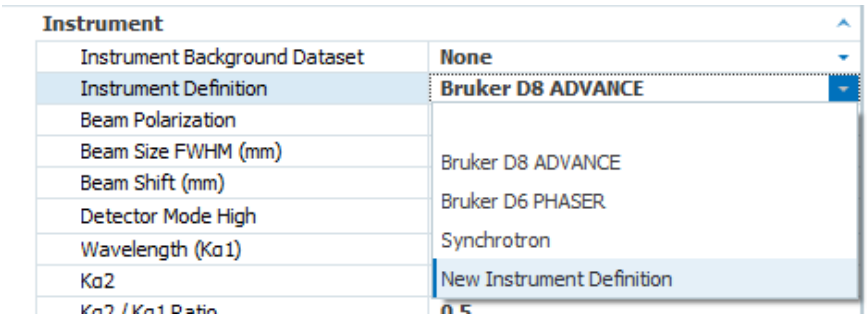
To modify the file, open *Instrument.json* in a text editor (Notepad++ is used here).

```
1  {
2      "instrument01": {
3          "name": "Bruker D8 ADVANCE",
4          "beam_fwhm": 0.4,
5          "beam_shift": 0,
6          "wavelength": 0.7093,  // should be overridden by value from inp
7          "polarization": 0.9999,
8          "detector_mode": "low"  // should be overridden by value from in
9      },
10     "instrument02": {
11         "name": "Bruker D6 PHASER",
12         "beam_fwhm": 0.5,
13         "beam_shift": 0,
14         "wavelength": 0.7093,
15         "polarization": 0.9999,
16         "detector_mode": "low"
17     },
18     "instrument03": {
19         "name": "Synchrotron",
20         "beam_fwhm": 1,
21         "beam_shift": 0,
22         "wavelength": 0.4,
23         "polarization": 0.001,
24         "detector_mode": "low"
25     },
26 }
```

To add an instrument definition, copy the lines 18 to 25. Add an empty new line between line 25 and 26 and paste the copied lines here. Make sure there is still a closing } at the end of file – Notepad++ will indicate if there are any errors in the file. Rename the new entry as “instrument04” (line 26 below) and modify the parameters for the new instrument definition as desired.

```
18     "instrument03": {
19         "name": "Synchrotron",
20         "beam_fwhm": 1,
21         "beam_shift": 0,
22         "wavelength": 0.4,
23         "polarization": 0.001,
24         "detector_mode": "low"
25     },
26     "instrument04": {
27         "name": "New Instrument Definition",
28         "beam_fwhm": 1,
29         "beam_shift": 0,
30         "wavelength": 0.4,
31         "polarization": 0.001,
32         "detector_mode": "low"
33     },
34 }
```

The next time PDF Processing is run, the new instrument definition will appear in the dropdown menu of *Instrument Definition* in the **Initialize PDF data** window.

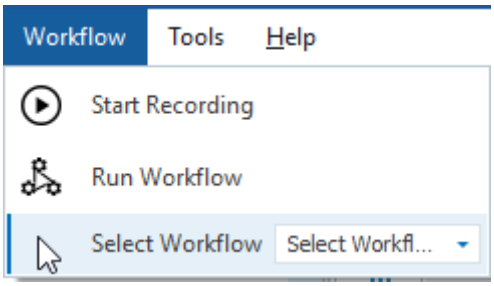


Container.json can be modified in the same way.

Workflows

EVA V7.0 allows recording and replaying lists of commands, called workflows. Workflows are collections of commands which were previously recorded and can be re-played at any time.

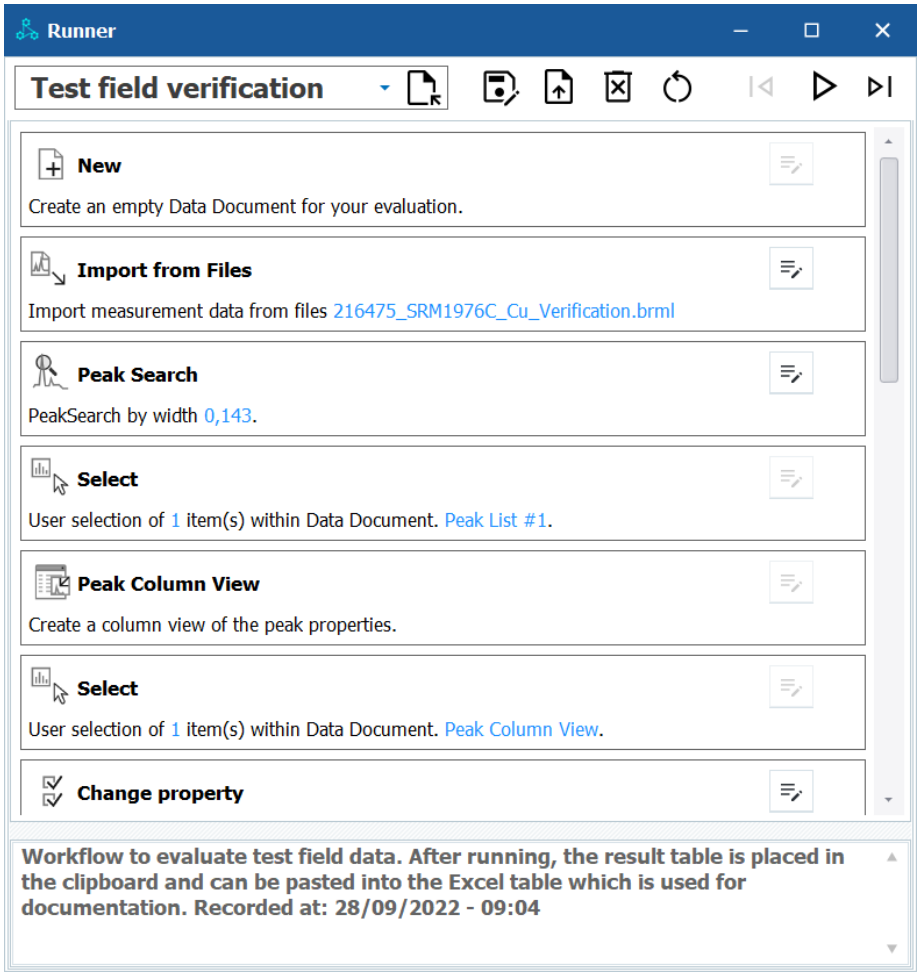
New Workflow menu commands and corresponding main toolbar buttons indicate the availability of workflows.



A rich subset of more than 110 EVA commands can be recorded in workflows. While a recording is active, only recordable commands are enabled. The recording icon changes during recording. Each time a command is recorded a status message is displayed.

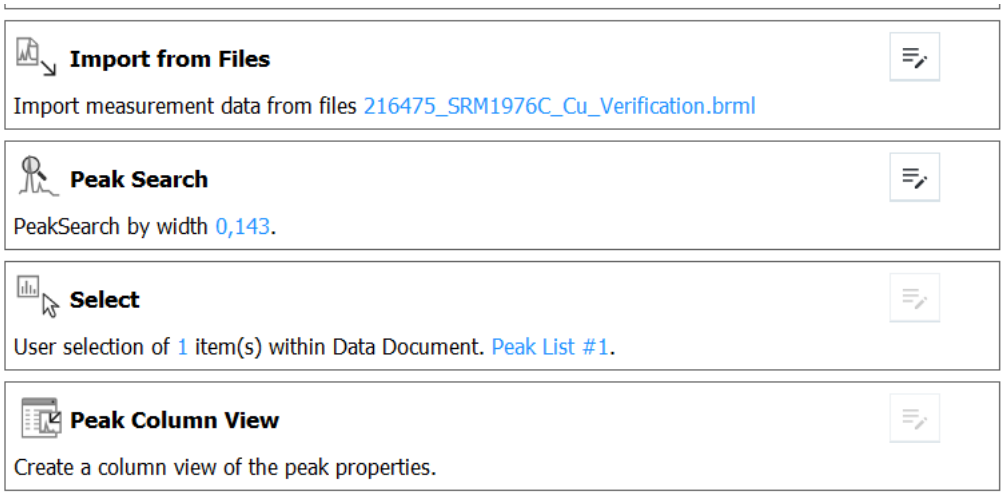
The Workflow Runner Tool

The Workflow Runner tool is the interface to control all workflow-related tasks.



The toolbar on top of the Workflow Runner contains the workflow selector on the left side and a collection of buttons to control the workflow on the right side.

The command list below the toolbar is populated with the workflow commands after a workflow was selected.



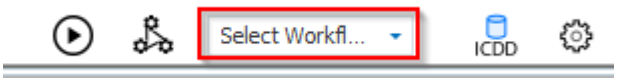
Command parameters may be changed in a pre-recorded workflow using the edit button on the right side of a command: 

Below the command list is a guidance control that displays information about the workflow or the command under the mouse cursor:

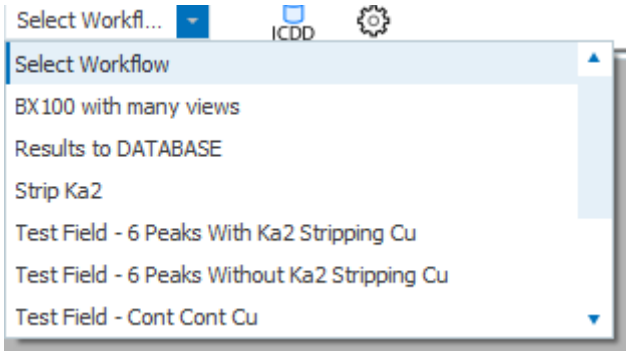
The workflows which are shown in the runner control are context sensitive. Only workflows that start with a command applicable to the current selection are displayed and allowed to start.

The Workflow Instant Runner Tool

The toolbar contains a special drop-down list that starts the workflow instantly after being selected. This list contains all available workflows:

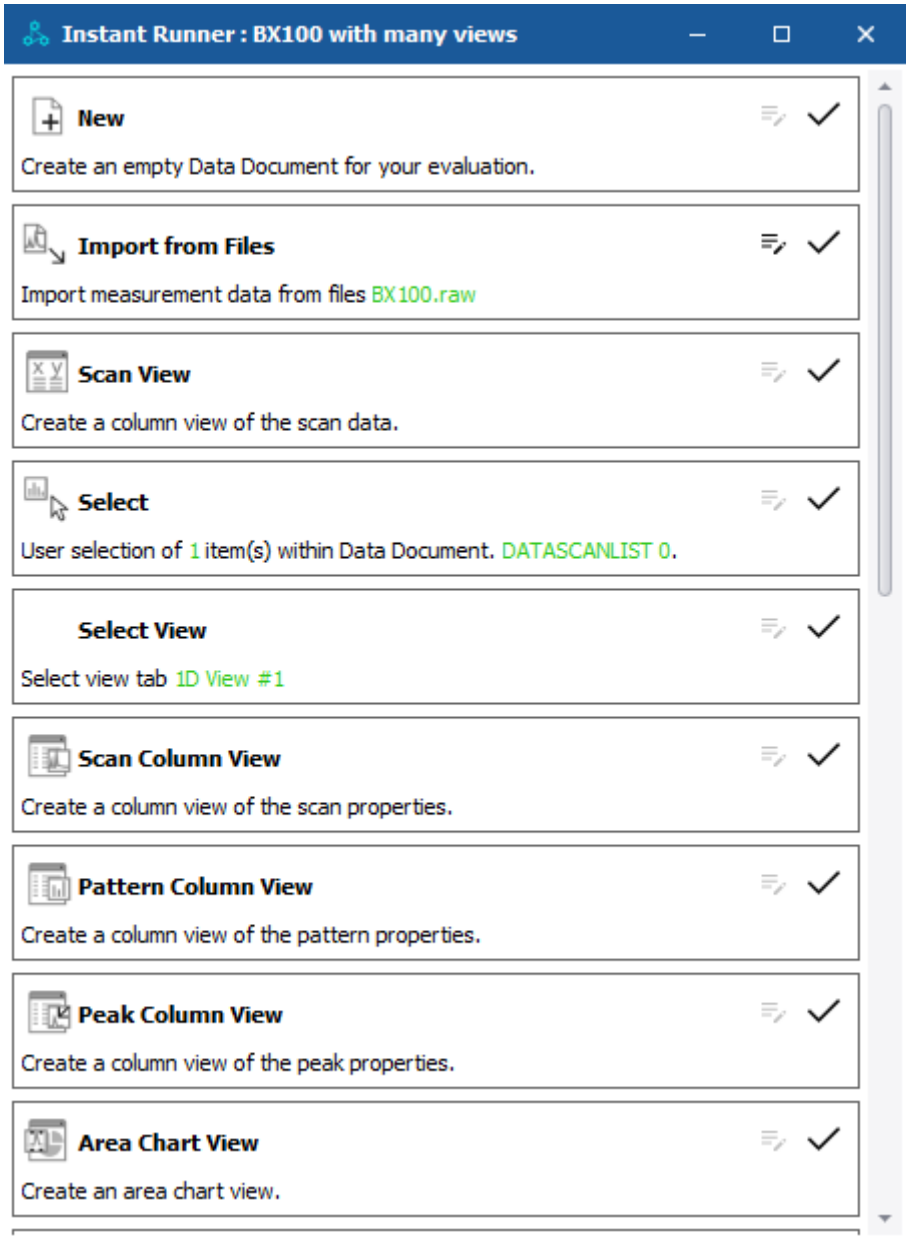


A click on the arrow on the list's right side opens the drop-down with all available workflows:

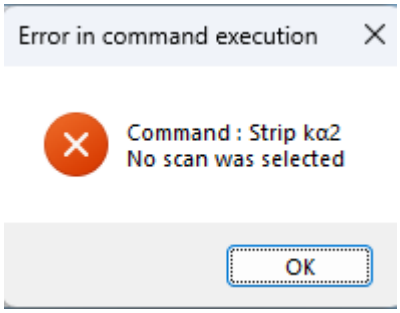


The workflow selection is not context sensitive and therefore the full list of available workflows is always displayed.

A click on one of the workflows starts the *Instant Runner* which executes the workflow immediately:

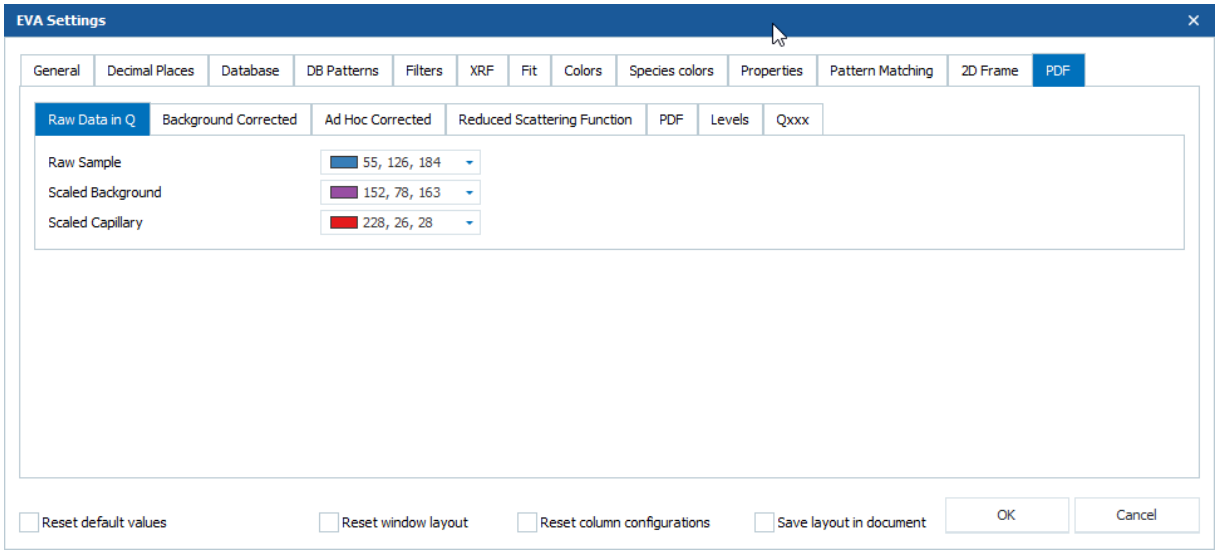


If an unsuitable workflow is selected or the conditions to execute a workflow are not met, an error message is displayed, e.g.:

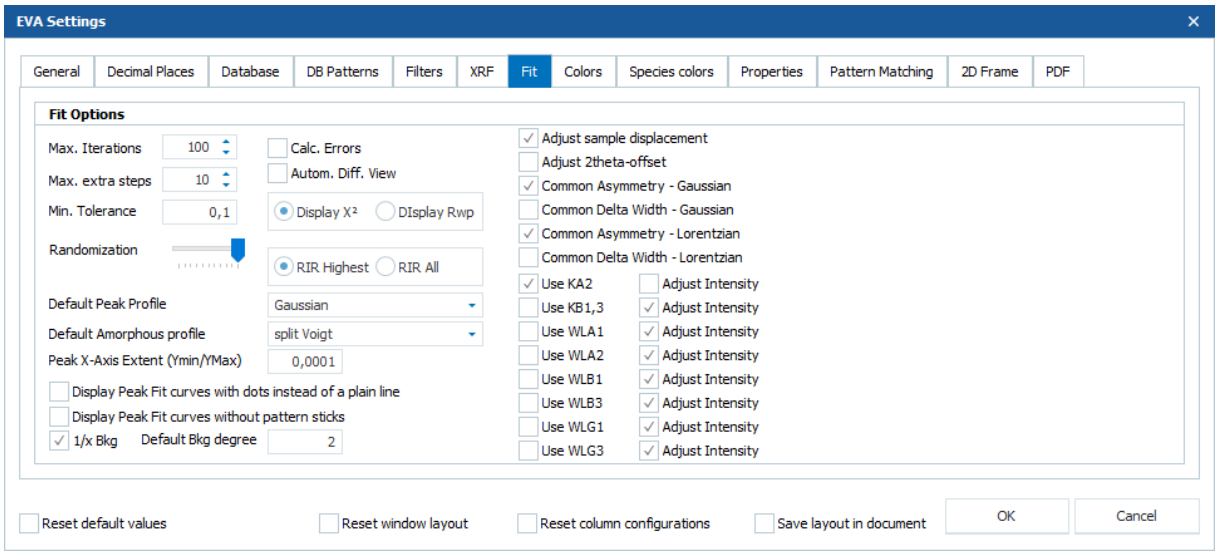


EVA Settings

There is a new PDF tab in the EVA Settings. This tab contains all globally configurable properties for the PDF processing.



The Fit tab was extended to accommodate the new profile fit features:



User Manual Appendix:

17.1 DIFFRAC.EVA V7.0 and License Level

DIFFRAC.EVA V7.0 has two features which were introduced with V7.0 and require license level 7:

- Pair Distribution Function (PDF) Processing
- Workflows