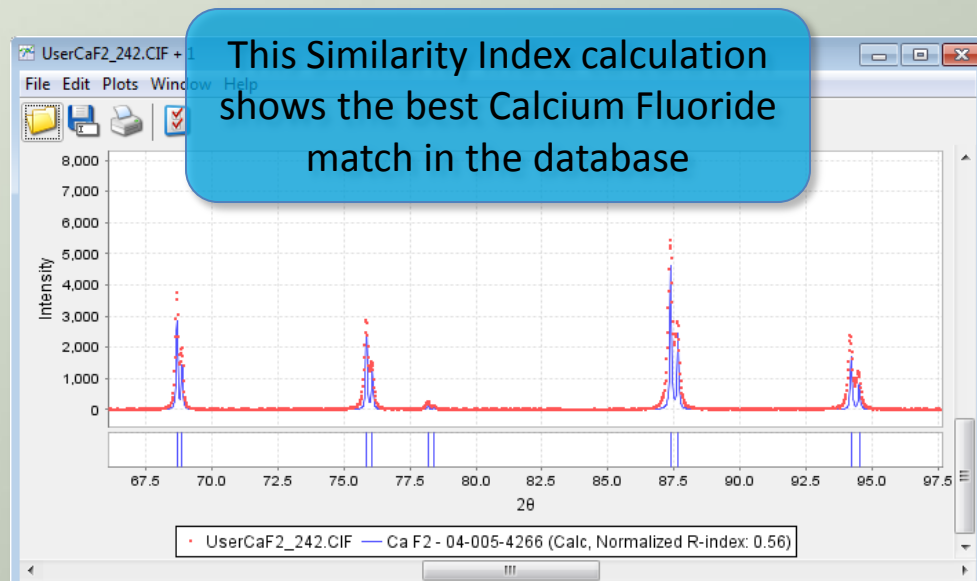


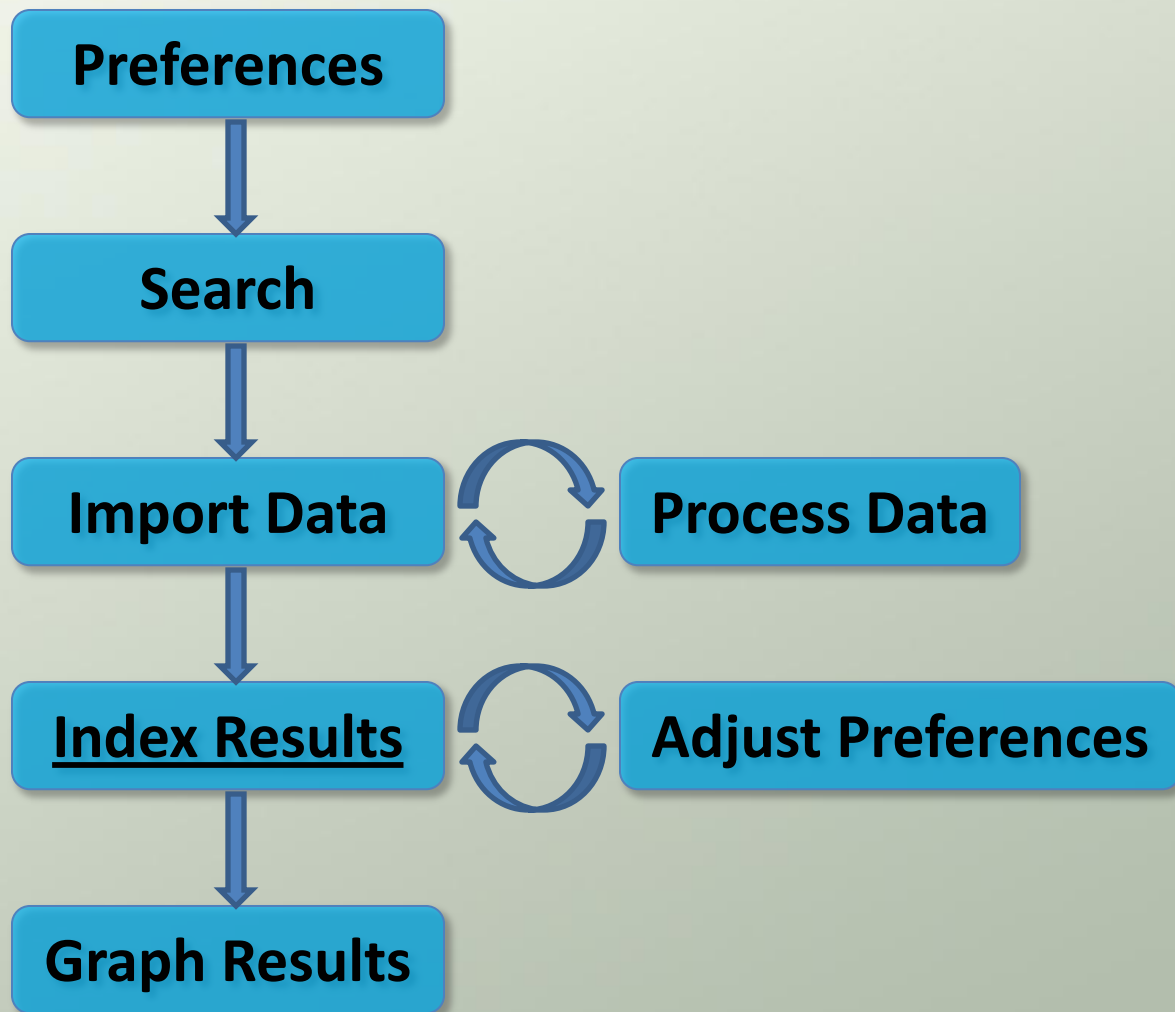
Use Similarity Indexes, Integral Index

What is the Similarity Index*?

- The Similarity Index is a numerical value based on total pattern analysis of how well imported experimental data compares to full simulated patterns in the ICDD PDF-4 database.
- Because the Similarity Index uses every point of the pattern scan for comparison (instead of just the peak positions), it is best used for only single phase experiments.
- There are numerous preferences available to customize the Similarity Index.
- Applications of the Similarity Index can be applied to both the Results form and the Sleeve+ form.

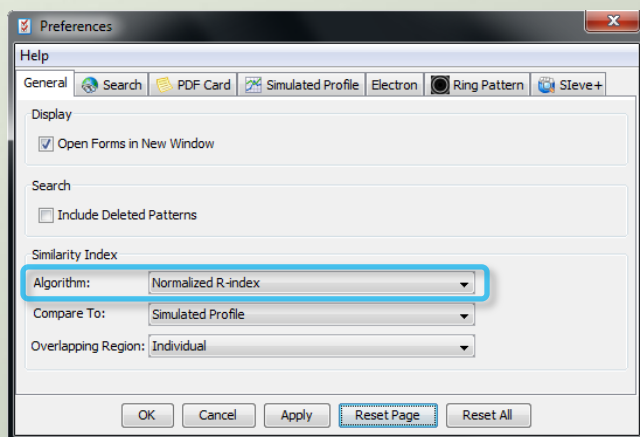


The Similarity Index Process



Preferences

- The first step in calculating Similarity Indexes is to set your *Similarity Index Preferences*. These preferences control how the Similarity Indexes are calculated. (Use the *Edit Menu > Preferences > General tab*)
- Algorithm:** This is the algorithm used in calculating the Similarity Indexes.
 - Normalized R-index:* This algorithm (normalized by J. Faber) is based on original algorithms found in [1] and [2]:
 - Integral Index:* This algorithm is based on the Similarity Index found in [1]:
 - It is generally recommended to keep the default algorithm set to *Normalized R-index*, as this is the most accurate. However, the *Integral Index* algorithm is still available for legacy purposes.



Normalized R-index

$$S_r = \sum_{i=1}^n \left| \frac{I_i^{\text{count}}}{\sum_{j=1}^n I_j^{\text{count}}} - \frac{I_i^{\text{calc}}}{\sum_{j=1}^n I_j^{\text{calc}}} \right|$$

Integral Index

$$S_{\text{int}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\sum_{j=1}^n I_j^{\text{calc}}}{\sum_{j=1}^n I_j^{\text{calc}}} - \frac{\sum_{j=1}^n I_j^{\text{count}}}{\sum_{j=1}^n I_j^{\text{count}}} \right|$$

S_r : Similarity Index using the Normalized R-index algorithm

S_{int} : Similarity Index using the Integral Index algorithm

n : Number of data points being compared

I^{count} : Intensity of current data point in experimental data

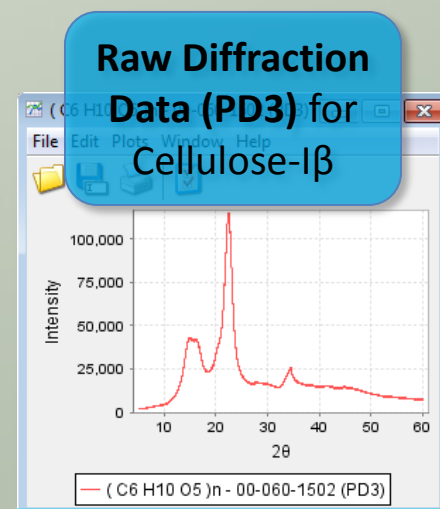
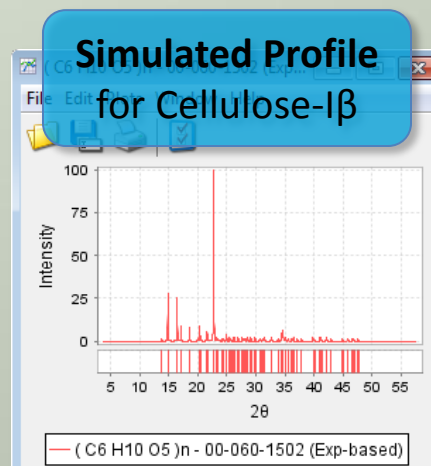
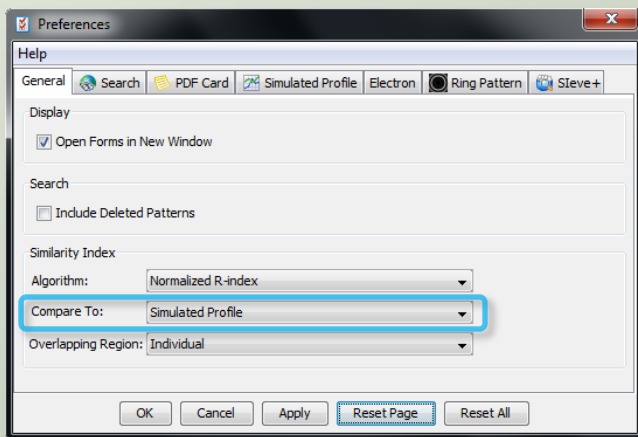
I^{calc} : Intensity of current data point in simulated/PD3 pattern

[1] Hofmann, D. & Kuleshova, L. (2005). J. Appl. Cryst. **38**, 861-866.

[2] Karfunkel, H.R., Rohde, B., Leusen, F.J.J., Gganitz, R.J. & Rihs, G. (1993). J. Comput. Chem. **14**, 1125.

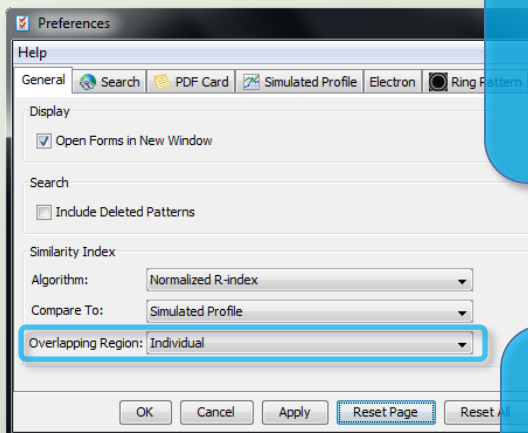
Preferences

- **Compare To:** The type of pattern the user's experimental data is compared to.
 - *Simulated Profile:* This compares the user's experimental data to diffraction profiles simulated from the d-Spacings and intensities. The settings used to simulate these profiles can be changed in the Simulated Profile Preferences form. (See slide 7)
 - *Raw Diffraction Data (PD3):* This compares the user's experimental data to PD3 Patterns (high quality raw experimental data). The wavelength of the PD3 patterns will be converted internally to match the user's experimental data. Because there are a low number of PD3 patterns currently in the database, the simulated profile will be used if a PD3 pattern does not exist for a particular entry.

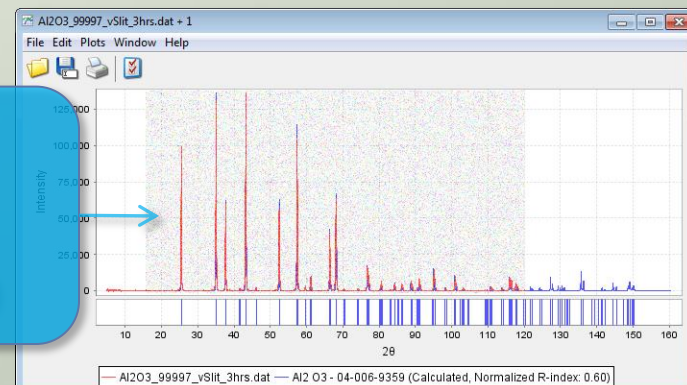


Preferences

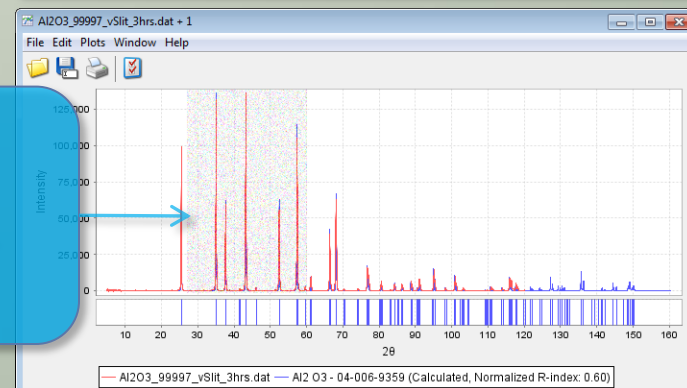
- **Overlapping Region:** The common 2θ region that is used in calculating the Similarity Indexes.
 - *Individual:* The common 2θ region between the user's experimental data and the current pattern in the search results.
 - *Set-wide:* The common 2θ region between the user's experimental data and all patterns in the search results. If this region does not exist, try reducing the number of patterns.



Individual Overlapping Region:
In this example, the common (compared) 2θ region is $15.53^\circ - 119.96^\circ$ (over the current pattern)

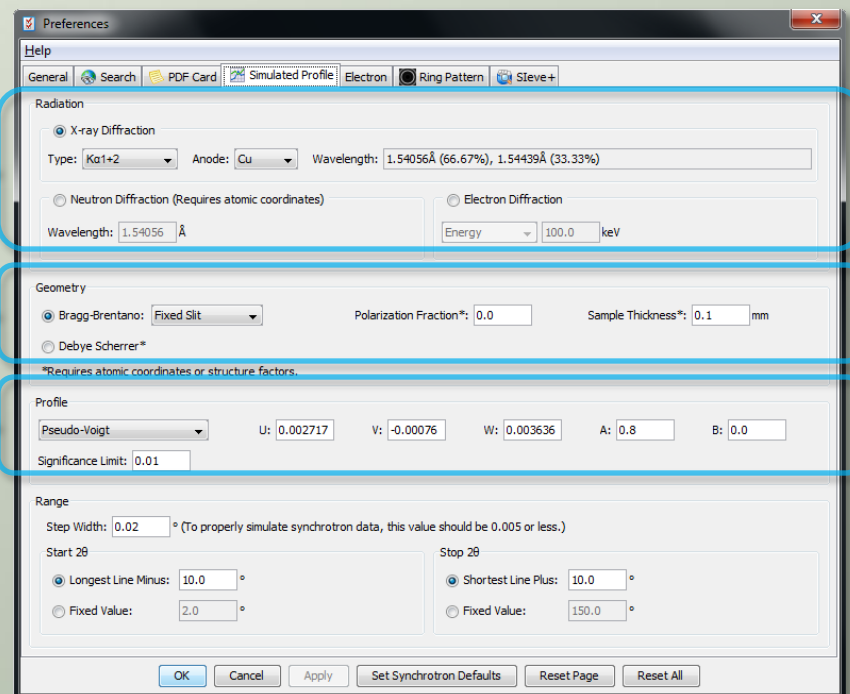


Set-wide Overlapping Region:
In this example, the common (compared) 2θ region is $27.65^\circ - 60.14^\circ$ (over all 129 patterns)



Preferences

- The next step in calculating Similarity Indexes is to set your Simulated Profile Preferences. These preferences control how the simulated profiles are calculated, which are used for comparison to your experimental data. (Use the *Edit Menu > Preferences > Simulated Profile tab*)
- The following preferences are used for simulating profiles:
 - Radiation**
 - Geometry**
 - Profile**
 - Since the 2θ range is handled automatically by the **Overlapping Region** preference, the **Range** preferences are ignored.



Preferences

Help

General Search PDF Card **Simulated Profile** Electron Ring Pattern Sieve+

Radiation

☒ X-ray Diffraction

Type: Anode: Wavelength:

☐ Neutron Diffraction (Requires atomic coordinates)

☐ Electron Diffraction

Wavelength: Å Energy: keV

Geometry

☒ Bragg-Brentano: Polarization Fraction*: Sample Thickness*: mm

☐ Debye Scherrer*

*Requires atomic coordinates or structure factors.

Profile

U: V: W: A: B:

Significance Limit:

Range

Step Width: * (To properly simulate synchrotron data, this value should be 0.005 or less.)

Start 2θ

☒ Longest Line Minus: °

☐ Fixed Value: °

Stop 2θ

☒ Shortest Line Plus: °

☐ Fixed Value: °

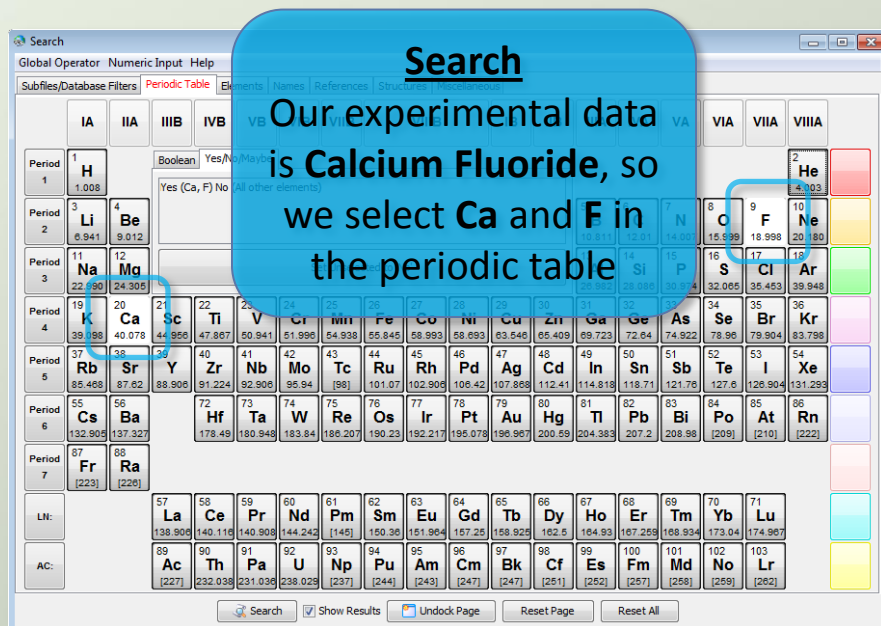
OK Cancel Apply Set Synchrotron Defaults Reset Page Reset All

Search

- The next step in calculating Similarity Indexes is to create search criteria based on known information about your experimental data. Some common examples of known information are:
 - Subfile/subclass
 - Elemental analysis
 - Compound/common/mineral name
 - Strong/long lines (d-Spacings)
- The more strict the search criteria, the better. Because calculating Similarity Indexes is time consuming, a small and accurate set of search results is recommended.
- This search criteria will generate a specialized list of entries (search results) from the database that will then be compared to your experimental data.

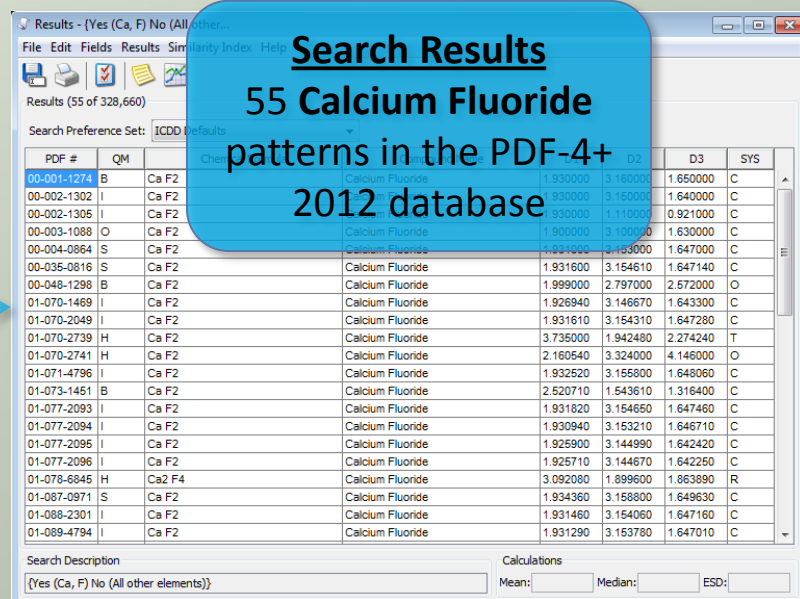
Search

Our experimental data is Calcium Fluoride, so we select Ca and F in the periodic table



Search Results

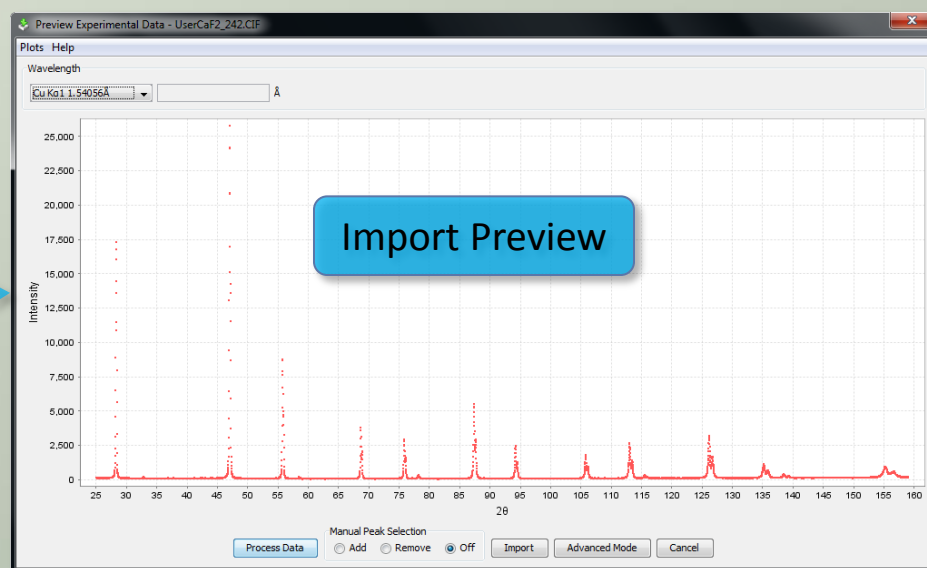
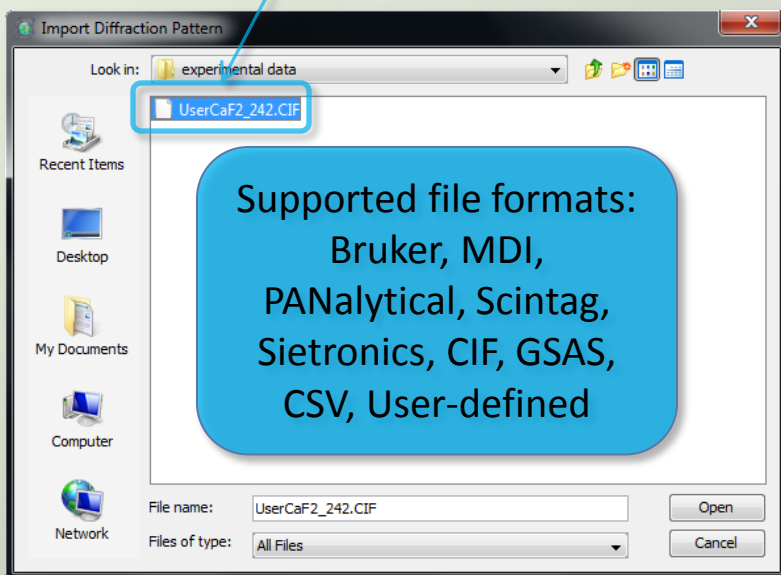
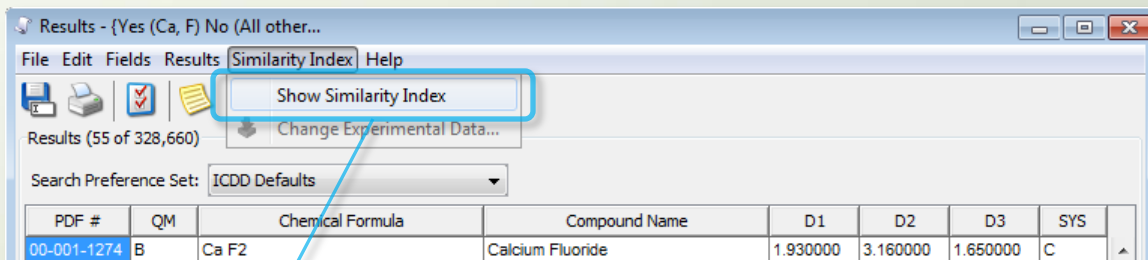
55 Calcium Fluoride patterns in the PDF-4+ 2012 database



PDF #	QM	Chemical	D1	D2	D3	SYS
00-001-1274	B	Ca F2	1.930000	3.160000	1.660000	C
00-002-1302	I	Ca F2	1.930000	3.160000	1.640000	C
00-002-1305	I	Ca F2	1.930000	3.160000	0.921000	C
00-003-1088	O	Ca F2	1.930000	3.160000	1.630000	C
00-004-0864	S	Ca F2	1.930000	3.160000	1.647000	C
00-035-0816	S	Ca F2	1.931600	3.154610	1.647140	C
00-048-1298	B	Ca F2	1.999000	2.797000	2.672000	O
01-070-1469	I	Ca F2	1.926940	3.146670	1.643300	C
01-070-2049	I	Ca F2	1.931610	3.154310	1.647280	C
01-070-2739	H	Ca F2	3.735000	1.942480	2.274240	T
01-070-2741	I	Ca F2	2.160540	3.324000	4.146000	O
01-071-4796	I	Ca F2	1.932520	3.155800	1.648060	C
01-073-1451	B	Ca F2	2.520710	1.543610	1.316400	C
01-077-2093	I	Ca F2	1.931820	3.154650	1.647460	C
01-077-2094	I	Ca F2	1.930940	3.153210	1.646710	C
01-077-2095	I	Ca F2	1.925900	3.144990	1.642420	C
01-077-2096	I	Ca F2	1.925710	3.144670	1.642250	C
01-078-6845	H	Ca2 F4	3.092080	1.899600	1.863890	R
01-087-0971	S	Ca F2	1.934360	3.156800	1.649630	C
01-088-2301	I	Ca F2	1.931460	3.154060	1.647160	C
01-088-4794	I	Ca F2	1.931290	3.153780	1.647010	C

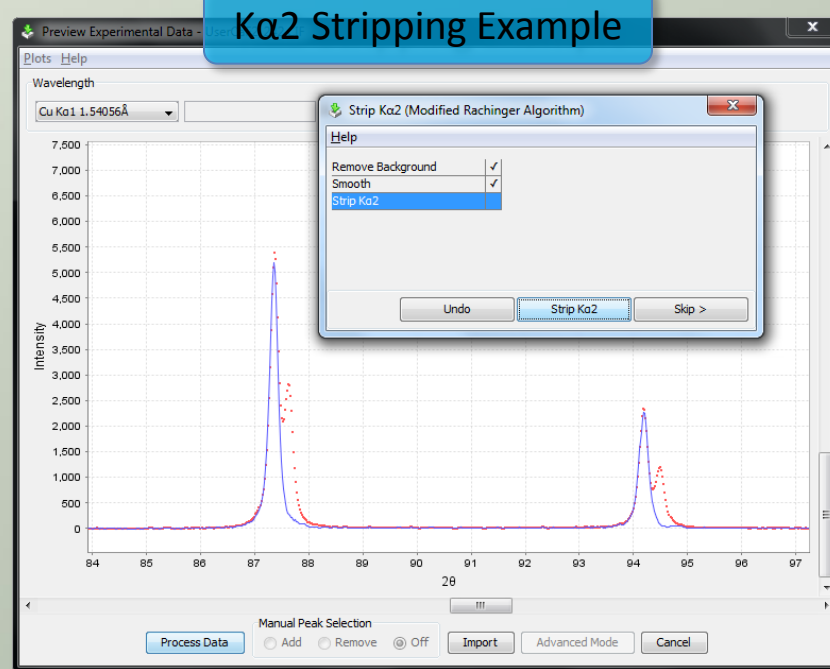
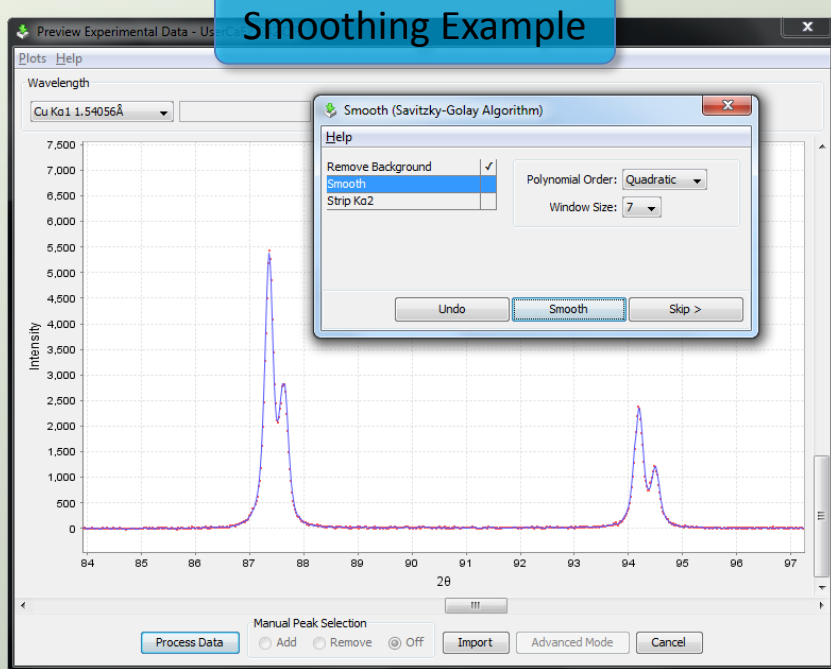
Import Data

- Before we can calculate Similarity Indexes for the search results, we need to first import our experimental data. (Use the *Similarity Index Menu > Show Similarity Index*)



Process Data

- Before the experimental data is imported, you can optionally process the experimental data to better match the simulated profiles:
 - Remove Background:** If your experimental data has background intensity, it is recommended to remove this as the simulated profiles do not use a background function.
 - Smooth:** If your experimental data is noisy, it is recommended to smooth it to better match the “clean” simulated profiles.
 - Strip K α 2:** If your experimental data includes K α 2 radiation, you can strip this out. However, you must also set the radiation type to K α 1 in the Simulated Profile Preferences. (See slide 7)



Index Results

- Once the experimental data is imported, a Similarity Index will be automatically calculated for each entry in the search results. This is done by:
 - Creating a simulated profile “on-the-fly” using the Simulated Profile Preferences or by reading the raw PD3 pattern from the database. (If applicable; see slide 5.)
 - Formulating a numerical value based on the *fit quality between the experimental data and this current compared pattern*. The lower the Similarity Index number, the better the match. (Zero is a theoretical perfect match.)

Click column to sort

Best match (lowest number):
PDF # 04-005-4266

Similarity Indexes for
Calcium Fluoride patterns

2 θ range used for
comparison is shown in
parenthesis

Results - {Yes (Ca, F) No (All other...}

File Edit Fields Results Similarity Index Help

Results (55 of 328,660)

Search Preference Set: ICDD Defaults

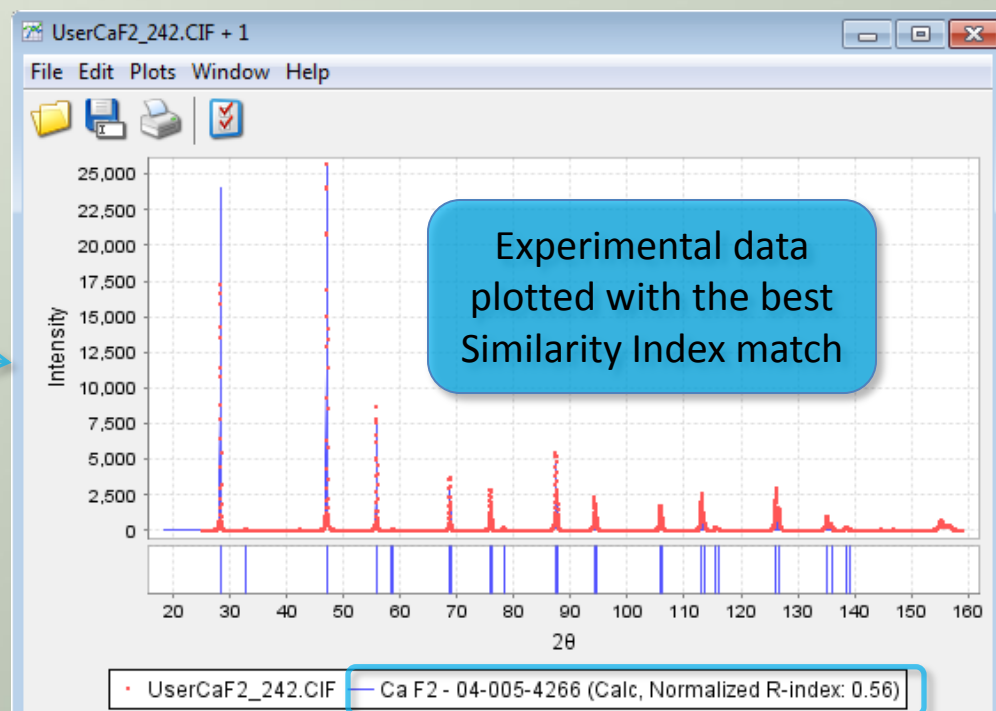
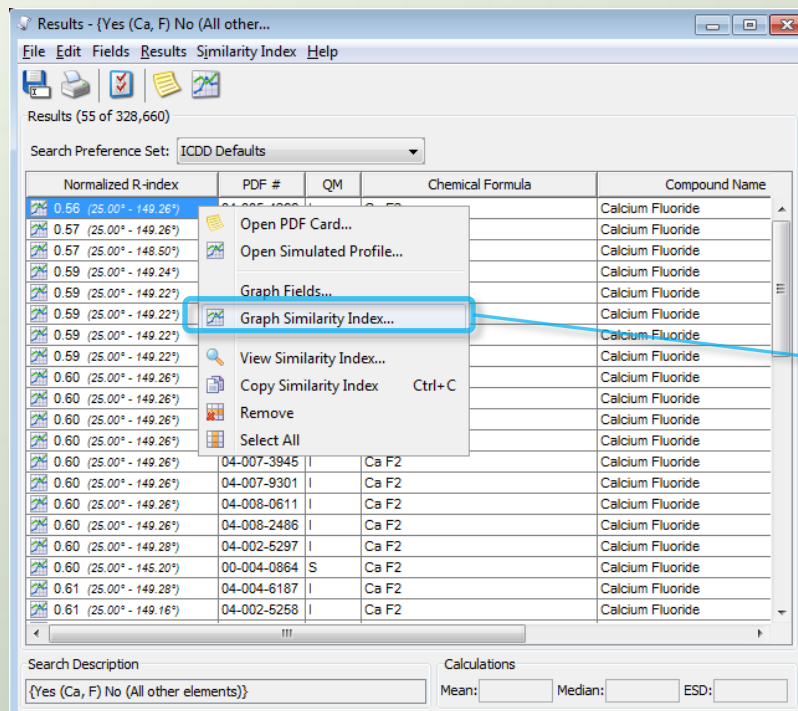
Normalized R-index	PDF #	QM	Chemical Formula	Compound Name	D1	
0.56 (25.00° - 149.26°)	04-005-4266	I	Ca F2	Calcium Fluoride	1.931460	3.
0.57 (25.00° - 149.26°)	01-088-2301	I	Ca F2	Calcium Fluoride	1.931460	3.
0.57 (25.00° - 148.50°)	00-035-0816	S	Ca F2	Calcium Fluoride	1.931600	3.
0.59 (25.00° - 149.24°)	01-070-2049	I	Ca F2	Calcium Fluoride	1.931610	3.
0.59 (25.00° - 149.22°)	04-006-2690	I	Ca F2	Calcium Fluoride	1.931730	3.
0.59 (25.00° - 149.22°)	01-077-2093	I	Ca F2	Calcium Fluoride	1.931820	3.
0.59 (25.00° - 149.22°)	04-002-4952	I	Ca F2	Calcium Fluoride	1.931820	3.
0.59 (25.00° - 149.22°)	04-002-4993	I	Ca F2	Calcium Fluoride	1.931820	3.
0.60 (25.00° - 149.26°)	04-002-2191	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.26°)	04-002-4443	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.26°)	04-002-5199	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.26°)	04-006-4980	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.26°)	04-007-3945	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.26°)	04-007-9301	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.26°)	04-008-0611	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.26°)	04-008-2486	I	Ca F2	Calcium Fluoride	1.931460	3.
0.60 (25.00° - 149.28°)	04-002-5297	I	Ca F2	Calcium Fluoride	1.931430	3.
0.60 (25.00° - 145.20°)	00-004-0864	S	Ca F2	Calcium Fluoride	1.931000	3.
0.61 (25.00° - 149.28°)	04-004-6187	I	Ca F2	Calcium Fluoride	1.931320	3.
0.61 (25.00° - 149.16°)	04-002-5258	I	Ca F2	Calcium Fluoride	1.932170	3.

Search Description: {Yes (Ca, F) No (All other elements)}

Calculations: Mean: Median: ESD:

Graph Results

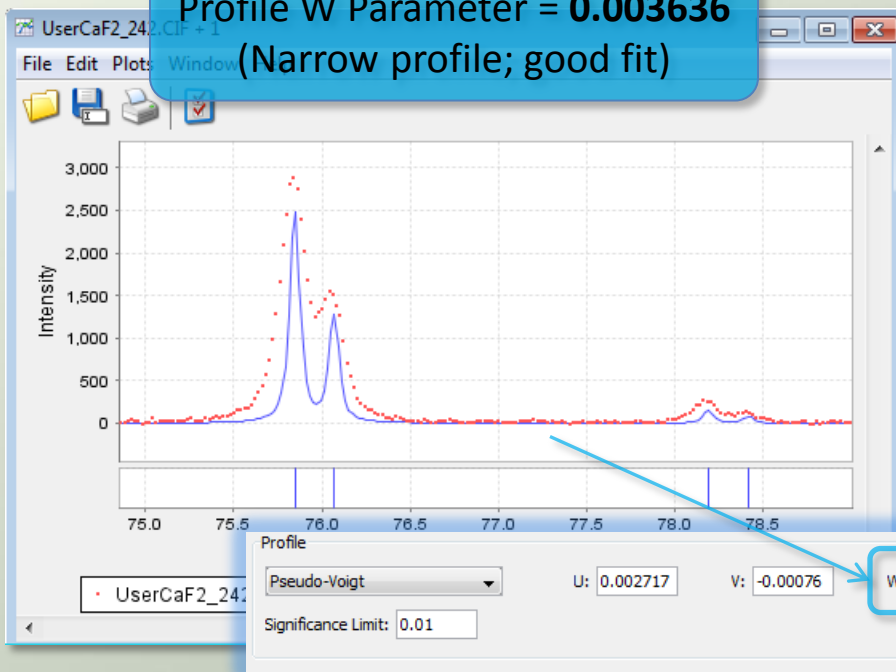
- While the Similarity Indexes themselves numerically describe the fit quality, it is also beneficial to visually see the fit quality for yourself.
- Select one or more entries in the search results and use the right-click menu item *Graph Similarity Index*.



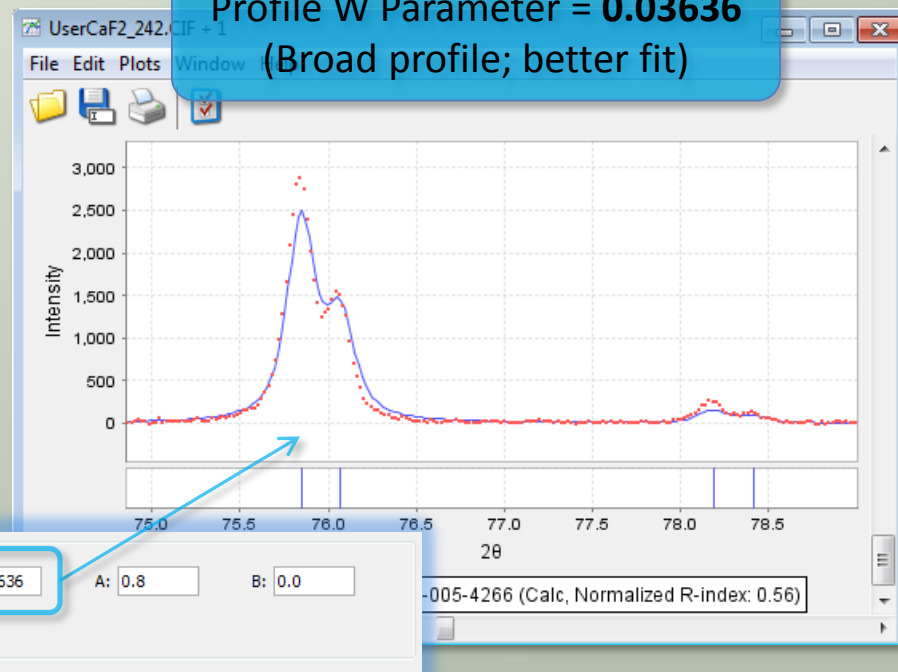
Graph Results

- Using the mouse, you can zoom in to better see the fit quality.
- In this example, we can see that the experimental data profile (red) is somewhat broader than the simulated profile (blue).
- Try readjusting the Simulated Profile Preferences to get a better fit. (See slide 7)
- To print the graph, use the menu item *File > Print*.

Profile W Parameter = **0.003636**
(Narrow profile; good fit)



Profile W Parameter = **0.03636**
(Broad profile; better fit)



Readjust Preferences

- After readjusting the Similarity Index Preferences and/or the Simulated Profile Preferences, the Similarity Indexes will automatically recalculate to new values.
- To print the search results, use the menu item *File > Print*.

Similarity Indexes with Profile
W Parameter = **0.003636**
(Good indexes)

Results - (Yes (Ca, F) No (All other elements))

File Edit Fields Results Search

Results (55 of 328,660)

Search Preference Set: ICDD Defaults

Normalized R-index	PDF #	QM	Chemical Formula	Compound Name
0.56 (25.00° - 149.26°)	04-005-4266	I	Ca F2	Calcium Fluoride
0.57 (25.00° - 149.26°)	01-088-2301	I	Ca F2	Calcium Fluoride
0.57 (25.00° - 148.50°)	00-035-0816	S	Ca F2	Calcium Fluoride
0.59 (25.00° - 149.24°)	01-070-2049	I	Ca F2	Calcium Fluoride
0.59 (25.00° - 149.22°)	04-006-2590	I	Ca F2	Calcium Fluoride
0.59 (25.00° - 149.22°)	01-077-2093	I	Ca F2	Calcium Fluoride
0.59 (25.00° - 149.22°)	04-002-4952	I	Ca F2	Calcium Fluoride
0.59 (25.00° - 149.22°)	04-002-4993	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-002-2191	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-002-4443	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-002-5199	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-006-4980	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-007-3945	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-007-9301	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-008-0611	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.26°)	04-008-2486	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 149.28°)	04-002-5297	I	Ca F2	Calcium Fluoride
0.60 (25.00° - 145.20°)	00-004-0864	S	Ca F2	Calcium Fluoride
0.61 (25.00° - 149.28°)	04-004-6187	I	Ca F2	Calcium Fluoride

Similarity Indexes with Profile
W Parameter = **0.03636**
(Better indexes)

Results - (Yes (Ca, F) No (All other elements))

File Edit Fields Results Search

Results (55 of 328,660)

Search Preference Set: ICDD Defaults

Normalized R-index	PDF #	QM	Chemical Formula	Compound Name
0.44 (25.00° - 149.26°)	04-005-4266	I	Ca F2	Calcium Fluoride
0.45 (25.00° - 149.26°)	01-088-2301	I	Ca F2	Calcium Fluoride
0.49 (25.00° - 148.50°)	00-035-0816	S	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.24°)	01-070-2049	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.22°)	04-006-2590	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.22°)	01-077-2093	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.22°)	04-002-4952	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.22°)	04-002-4993	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-002-2191	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-002-4443	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-002-5199	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-006-4980	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-007-3945	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-007-9301	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-008-0611	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.26°)	04-008-2486	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.28°)	04-002-5297	I	Ca F2	Calcium Fluoride
0.50 (25.00° - 145.20°)	00-004-0864	S	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.28°)	04-004-6187	I	Ca F2	Calcium Fluoride
0.51 (25.00° - 149.16°)	04-002-5258	I	Ca F2	Calcium Fluoride

Profile

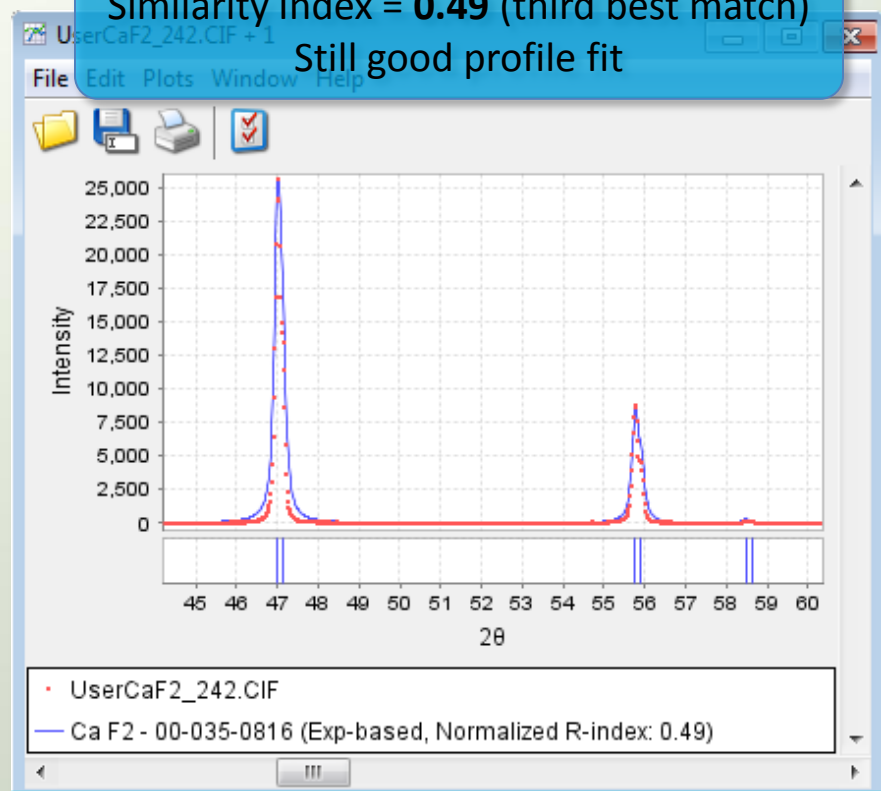
Pseudo-Voigt

Significance Limit: 0.01

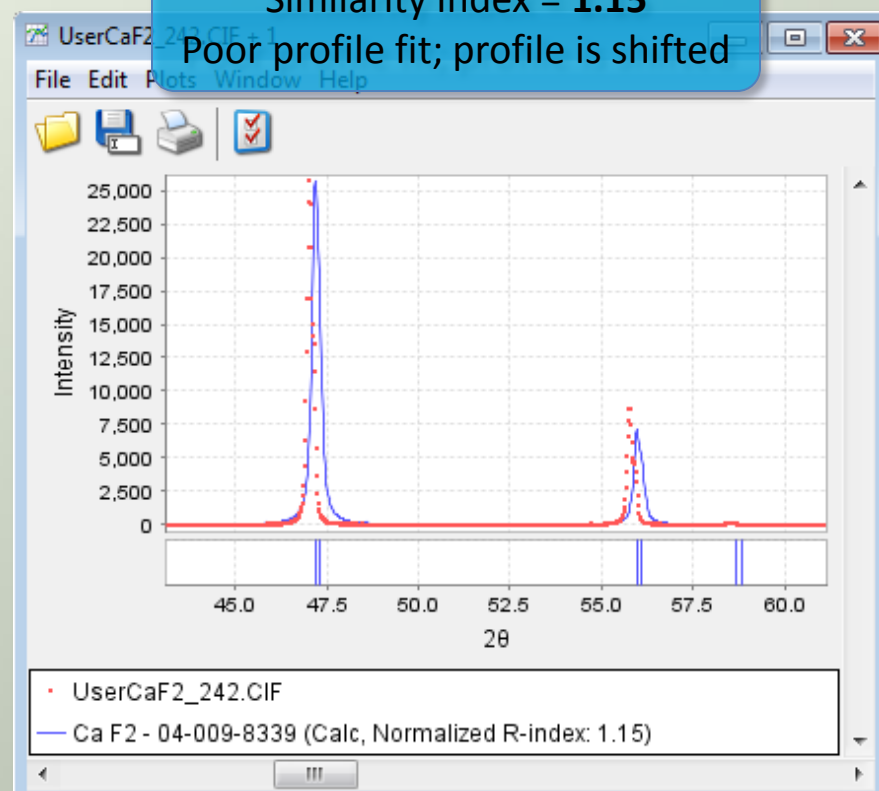
U: 0.002717 V: -0.00076 W: 0.03636 A: 0.8 B: 0.0

Graph Results

Similarity Index = **0.49** (third best match)
Still good profile fit

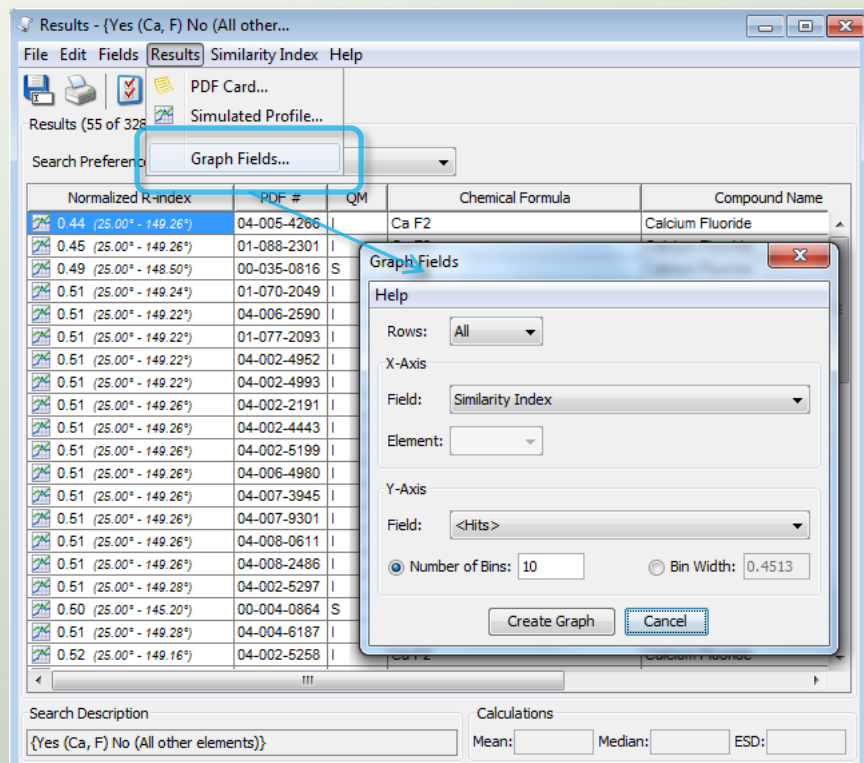


Similarity Index = **1.15**
Poor profile fit; profile is shifted



Graph Results

- To display a histogram of the Similarity Indexes, use the menu item *Results > Graph Fields*.
- This will show the statistical distribution of Similarity Index values for the entire search results.



Results - {Yes (Ca, F) No (All other...}

File Edit Fields Results Similarity Index Help

PDF Card...
Simulated Profile...

Results (55 of 328)

Search Preference: Graph Fields...

Normalized R-index	PDF #	QM	Chemical Formula	Compound Name
0.44 (25.00° - 149.26°)	04-005-4266	I	Ca F2	Calcium Fluoride
0.45 (25.00° - 149.26°)	01-088-2301	I		
0.49 (25.00° - 148.50°)	00-035-0816	S		
0.51 (25.00° - 149.24°)	01-070-2049	I		
0.51 (25.00° - 149.22°)	04-006-2590	I		
0.51 (25.00° - 149.22°)	01-077-2093	I		
0.51 (25.00° - 149.22°)	04-002-4952	I		
0.51 (25.00° - 149.22°)	04-002-4993	I		
0.51 (25.00° - 149.26°)	04-002-2191	I		
0.51 (25.00° - 149.26°)	04-002-4443	I		
0.51 (25.00° - 149.26°)	04-002-5199	I		
0.51 (25.00° - 149.26°)	04-006-4980	I		
0.51 (25.00° - 149.26°)	04-007-3945	I		
0.51 (25.00° - 149.26°)	04-007-9301	I		
0.51 (25.00° - 149.26°)	04-008-0611	I		
0.51 (25.00° - 149.26°)	04-008-2486	I		
0.51 (25.00° - 149.28°)	04-002-5297	I		
0.50 (25.00° - 145.20°)	00-004-0864	S		
0.51 (25.00° - 149.28°)	04-004-6187	I		
0.52 (25.00° - 149.16°)	04-002-5258	I		

Graph Fields

Help

Rows: All

X-Axis

Field: Similarity Index

Element:

Y-Axis

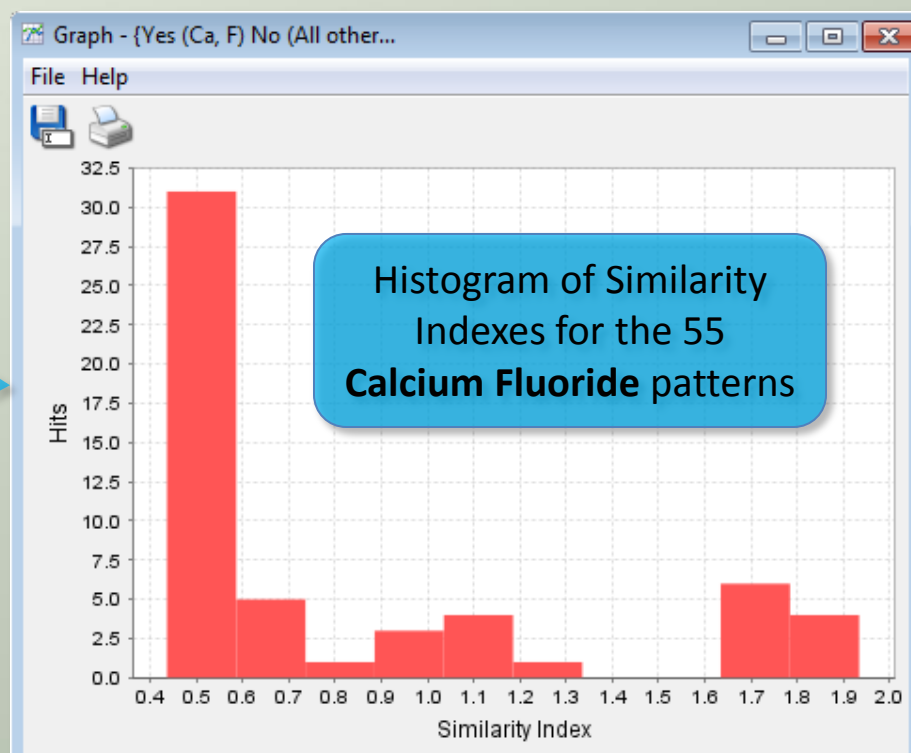
Field: <Hits>

☒ Number of Bins: 10 ☐ Bin Width: 0.4513

Create Graph Cancel

Search Description: {Yes (Ca, F) No (All other elements)}

Calculations: Mean: Median: ESD:





Thank you for viewing our tutorial.
Additional tutorials are available at the ICDD website.

www.icdd.com

International Centre for Diffraction Data

12 Campus Boulevard

Newtown Square, PA 19073

Phone: 610.325.9814

Toll Free Number in US & Canada: 866.378.0331

Fax: 610.325.9823