

## CIF applications. XII. Inspecting Rietveld fits from pdCIF: *pdCIFplot*

Brian H. Toby

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

CIF applications. XII. Inspecting Rietveld fits from  
pdCIF: *pdCIFplot*

Brian H. Toby

Received 6 February 2003

Accepted 29 July 2003

NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8562, USA. Correspondence e-mail: brian.toby@nist.gov

A description is given of the *pdCIFplot* program. This program is used to inspect the quality of fits obtained from Rietveld refinements where the results are communicated in pdCIF files. The program runs on almost all contemporary operating systems.

## 1. Introduction

Rietveld refinement is commonly used to obtain crystallographic information from powder diffraction data (Rietveld, 1969). Applications of the Rietveld technique include structure determination, phase quantification, lattice parameter determination, and characterization of strain and crystallite size. The Rietveld technique fits structural parameters to powder diffraction data, but also requires that experimental parameters such as the peak profile and background be fit. The only statistically relevant error indices,  $\chi^2$  and  $R_{wp}$  and related values, reflect a composite of how well the profile, the background and the structural parameters have been fit, but may not indicate that the refinement is providing a reliable crystallographic determination (Young, 1995). The author has seen many Rietveld refinements where poor structural fits are obtained, yet  $R$  factors were low. This can occur when the data have large uncertainties, due to inadequate counting time or high background levels. Likewise, with very high resolution and excellent counting statistics, excellent structural models may still result in very high  $R$  factors due to small imperfections in peak profiles that cannot be accurately modeled. Since these deviations are very large compared with the experimental uncertainties, they make a large contribution to the error indices. The IUCr Commission on Powder Diffraction has stated that the quality of Rietveld refinements is best judged *via* a graphical examination of the observed and calculated diffraction data, in addition to examining the plausibility of the structural results (McCusker *et al.*, 1999).

The crystallographic information file (CIF) provides a mechanism for communicating crystallographic results (Hall *et al.*, 1991). The CIF has been extended to encompass powder diffraction applications (pdCIF) and thus provides a standardized way to communicate observed and calculated Rietveld patterns, as well as resulting structural information (Toby, 2003a). The *pdCIFplot* program, which will be described in this article, has been created for graphical examination of Rietveld refinement results, particularly for editorial review of manuscripts. Rietveld results, including the observed and computed diffraction data, can be exported from the *GSAS* software package (Larson & Von Dreele, 1986; Toby *et al.*, 2003). It is hoped that this capability will soon be available from all Rietveld software.

## 2. pdCIF data loops

Powder diffraction data are collected using a variety of techniques. This is exemplified by Table 1, which lists the pdCIF data items that can be used as the dependent axis for powder diffraction data collection. Note that `_pd_proc_` data items in this table are derived

from the dependent variable using calibration information and are not actually experimental values.

For complete display of Rietveld results, four quantities are needed: the observed data, the uncertainties in the observed data, the observed, estimated or fitted background, and the intensity values calculated from the fitted structural model. Again, multiple CIF data items may be used to communicate these values, so data names are listed in Table 2. For some experiments, the as-collected diffraction

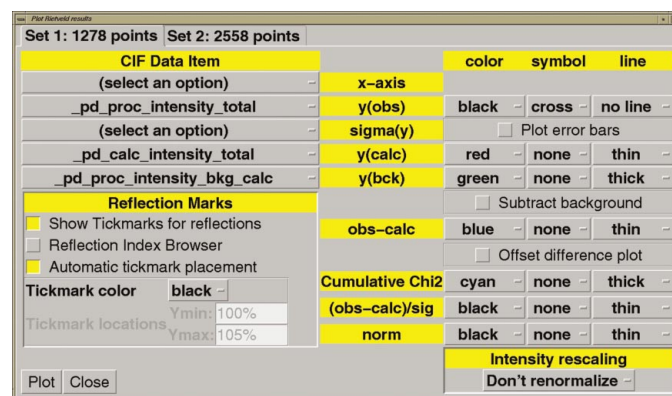


Figure 1

The Rietveld plotting window, where data items to be plotted, as well as plot options, are selected.

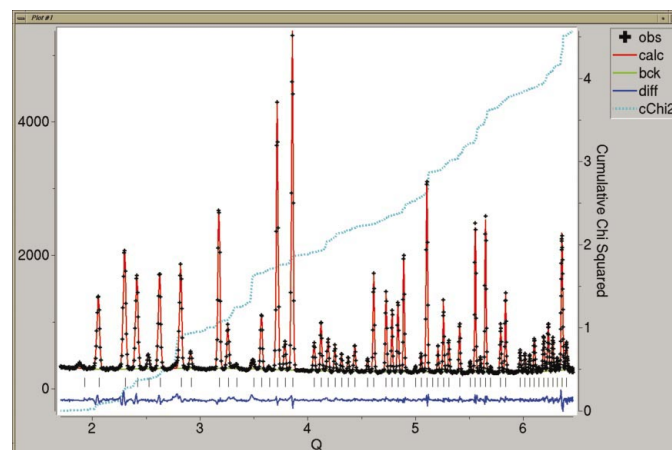


Figure 2

A sample plot from *pdCIFplot*. Note that the computer mouse can be used to enlarge ('zoom in on') any section of the figure.

**Table 1**  
pdCIF data items used for recording the powder diffraction dependent variable.

<code>_pd_meas_2theta_range_†</code>	Uncorrected $2\theta$ values with constant steps
<code>_pd_meas_2theta_scan</code>	Uncorrected $2\theta$ values, which may not have constant steps
<code>_pd_proc_2theta_range_†</code>	Calibration corrected $2\theta$ values with constant steps
<code>_pd_proc_2theta_corrected</code>	Calibration corrected $2\theta$ values, which may not have constant steps
<code>_pd_meas_time_of_flight</code>	Time for time-of-flight neutron diffraction measurements
<code>_pd_meas_position</code>	Linear detector position
<code>_pd_proc_energy_incident</code>	X-ray energy for energy-dispersive measurements
<code>_pd_proc_wavelength</code>	X-ray or neutron wavelength, when not constant
<code>_pd_proc_d_spacing</code>	$d$ spacing corresponding to an intensity value
<code>_pd_proc_recip_len_q</code>	Momentum transfer ( $Q = 4\pi \sin \theta/\lambda$ ) for an intensity value

† The data names indicated as `_pd_XXXX_2theta_range_` actually correspond to three CIF data items, `_pd_XXXX_2theta_range_min`, `_pd_XXXX_2theta_range_max` and `_pd_XXXX_2theta_range_inc`, which define a range of equally spaced values.

data (e.g. `_pd_meas_counts_total`, etc.) need to be scaled to account for factors such as variation of counting time, `_pd_meas_step_count_time`, or incident intensity, `_pd_meas_counts_monitor`. However, the processed intensity values (e.g. `_pd_proc_intensity_total`) should be on the same scale as the calculated values (e.g. `_pd_calc_intensity_total`). If data are averaged or rebinned, the number of observed and processed diffraction data points will differ. This will require two data loops, one for the `_pd_meas_` values and another for the `_pd_proc_` and `_pd_calc_` values.

### 3. The *pdCIFplot* program

The *pdCIFplot* program is supplied with the name of a CIF file, which is then parsed using the *CIFEDIT* CIF parser (Toby, 2003b). The program then lists the CIF data blocks in the file so that a block that contains powder diffraction data can be selected. Once a block has been selected, the data can be plotted with either the custom plotting menu or the Rietveld plotting menu. The CIF contents can also be viewed as ASCII text or with the hierarchical viewer implemented in the *CIFEDIT* program.

#### 3.1. Custom plotting menu

The custom plotting menu allows pdCIF data to be scaled and manipulated in unusual ways. Users can choose the window where the results will be displayed, so that results can be superimposed at will. This option was created as a precursor to development of the Riet-

veld plotting menu. It is not expected to be of value to most users, but was retained as it may have occasional utility.

#### 3.2. Rietveld plotting menu

The Rietveld plotting menu is shown in Fig. 1. The 'notebook tabs' in the upper left-hand corner of this menu select between sets of data, where data are grouped by the number of points in the loop. This is done as all plots are made against the same dependent variable ( $x$  axis). Thus, items with differing numbers of points cannot be plotted together.

**3.2.1. Data items selection.** Located directly below the notebook tabs, on the upper left-hand side of this plotting menu, are found 'menu buttons' that select the data items to be plotted. Note that, depending on the CIF contents, there may be no appropriate data items for a particular category, in which case the menu button will be disabled and will be labeled in gray 'not available'. If there is only one option available, the option is selected. Otherwise, the button is labeled 'select an option' until a selection is made from the pull-down menu.

The top button offers choices for the  $x$  axis. The choices include all the data items listed in Table 1 that are found in the CIF. Values computed from these data items may also be offered. For example, plotting in  $d$  space or  $Q$  is possible when both  $2\theta$  values and a wavelength value are reported. Directly below the  $x$  axis button are four menu buttons corresponding to each of the four columns in Table 2.

After selections have been made for all five menu buttons, and the Plot button (lower left-hand corner) is pressed, the *pdCIFplot* program creates three plot windows. The most complex of these windows is shown in Fig. 2. Note that fewer windows and less complex plots are created when some data items are unavailable or unselected.

The functions in the plot window in Fig. 2 are the observed data [ $y(\text{obs})$ ], the calculated data [ $y(\text{calc})$ ], the background values [ $y(\text{bck})$ ], the difference between the observed and calculated data [ $\text{obs} - \text{calc}$ ] and the cumulative  $\chi^2$  function. The cumulative  $\chi^2$  function is the running sum for the first  $i$  data points,  $\sum_i [y(\text{obs}) - y(\text{calc})]^2 / (i\sigma^2)$ . This function was developed by David (2003) to show the sections of the fit with the worst weighted agreement. The second plot window shows the weighted difference between the observed and calculated data [ $(\text{obs} - \text{calc})/\sigma$ ]. The third plot window displays a normalization constant [ $y(\text{norm})$ ] [ $y(\text{norm})_i = y(\text{obs})_i / \sigma_i^2$ ]. The reported intensity, multiplied by  $y(\text{norm})$ , gives the number of counts that would duplicate the ratio of intensity to standard uncertainty.

**3.2.2. Plot options.** The appearance of the plots is adjusted *via* the buttons on the right-hand side of the plot menu. Most of these buttons select the color, symbol and line type for each function. Selecting the symbol type as 'none' and the line type to 'no line'

**Table 2**  
pdCIF data items used for powder diffraction intensity values.

Observed intensities $y(\text{obs})$	Uncertainty values $\sigma_{y(\text{obs})}$	Background intensity $y(\text{bck})$	Calculated intensities $y(\text{calc})$
<code>_pd_meas_counts_total</code>	<code>_pd_meas_counts_total†</code>	<code>_pd_meas_counts_background</code>	<code>_pd_calc_intensity_net</code>
<code>_pd_meas_intensity_total</code>	<code>_pd_meas_intensity_total‡</code>	<code>_pd_meas_counts_container</code>	<code>_pd_calc_intensity_total</code>
<code>_pd_proc_intensity_total</code>	<code>_pd_proc_intensity_total‡</code>	<code>_pd_meas_intensity_background</code>	
<code>_pd_proc_intensity_net</code>	<code>_pd_proc_intensity_net‡</code>	<code>_pd_meas_intensity_container</code>	
	<code>_pd_proc_ls_weight§</code>	<code>_pd_proc_intensity_bkg_calc</code>	
		<code>_pd_proc_intensity_bkg_fix</code>	

† Standard uncertainty is the square-root of the counts for this data item. ‡ Standard uncertainty is directly reported for these data items. § Standard uncertainty is one over the square-root of the weights for this data item.

causes that function to be omitted from the display. Four options are also available in this side of the plot menu:

(a) Selection of the 'Plot error bars' option causes the standard uncertainties for the observed data to be displayed as vertical lines superimposed on each of the data points.

(b) When the 'Subtract background' option is selected, two functions are plotted,  $y(\text{obs}) - y(\text{bck})$  and  $y(\text{calc}) - y(\text{bck})$ , rather than the three  $y(\text{obs})$ ,  $y(\text{calc})$  and  $y(\text{bck})$  functions that are plotted otherwise.

(c) The  $\text{obs} - \text{calc}$  function is plotted on the same scale as the  $y(\text{obs})$ ,  $y(\text{calc})$  and  $y(\text{bck})$  values. Normally, no offset is applied to the  $\text{obs} - \text{calc}$  function, e.g. a point is plotted at zero when  $y(\text{obs}) = y(\text{bck})$ . However, when the 'Offset difference plot' option is selected, a positive or negative offset is applied to the  $\text{obs} - \text{calc}$  values to move this function just below the  $y(\text{obs})$ ,  $y(\text{calc})$  and  $y(\text{bck})$  functions. Note that the 'Offset difference plot' option is always invoked when automatic reflection-mark placement is used (see below).

(d) The 'Intensity rescaling' menu button in the lower right-hand corner changes the display of intensity values. When this option is set to 'Don't renormalize', values are displayed as found in the CIF. However, when standard uncertainties are present, different scaling options can be applied based on  $y(\text{norm})$ . When this option is set to 'Renormalize to counts', the  $y(\text{obs})$ ,  $y(\text{calc})$ ,  $y(\text{bck})$  and  $\text{obs} - \text{calc}$  functions are directly multiplied by  $y(\text{norm})$ . This gives a reviewer a feeling for the statistical significance of the data. For some instruments, for example where multiple detectors are used or where the incident beam intensity changes rapidly, the  $y(\text{norm})$  values will have discontinuities; when the rescaling option is set to 'Smooth and renormalize', the  $y(\text{norm})$  values are smoothed before they are applied. If the rescaling option is set to 'Average and renormalize', all the  $y(\text{norm})$  values are averaged and a single value is applied to all intensities.

**3.2.3. Reflection positions.** When analysing Rietveld results, it is often valuable to note reflection positions. This is available in *pdCIFplot* as an option using the 'Reflection Marks' submenu in the lower left-hand corner of the Rietveld plot menu. When reflection marks are selected, tick marks, i.e. vertical lines, are drawn at the position of every reflection found in the CIF reflection table (stored via the `_refln_*` data items). For reflection positions to be displayed, data item `_refln_d_spacing` must appear in the loop and *pdCIFplot* must be able to convert  $d$  spaces into the selected  $x$ -axis units.

Reflection tick marks can be placed at a specified height for each phase, or the height can be determined automatically by the program. When specified manually, the height units can be the actual intensity units or can be expressed as a percentage of the intensity scale by including a percent sign (%). Note that when automatic reflection placement is used, the 'Offset difference plot' option is invoked automatically.

The 'reflection index browser' mode is used to connect tick marks to reflection indices. If the mouse is placed over a tick mark when this mode is selected, the reflection position and  $hkl$  values are placed in a

separate window. A button can then be used to label the position visually.

#### 4. Program availability and installation

The *pdCIFplot* program is not subject to copyright. The source code is freely distributed for use, modification and redistribution. It can be downloaded from a NIST website as part of the NCTR CIFtools distributions (<ftp://ftp.ncnr.nist.gov/pub/cryst/cif>). Software installation is discussed in the software documentation (<http://www.ncnr.nist.gov/xtal/software/cif/pdCIFplot.html>). Microsoft Windows (9x/ME and NT/2000/XP) users will probably not need to read installation instructions; downloading and executing a single self-installing program places all files needed by the software, including the *pdCIFplot* and *CIFEDIT* source code, onto the user's hard disk and creates required shortcuts. Use on Unix computers requires the installation of a system-specific version of the Tcl/Tk and BLT packages. The source code for these packages is freely available and precompiled binary versions can be found for many Unix platforms. The program has also been tested on the Macintosh OS X platform, but installation documentation cannot be provided.

The author would like to thank Dr Judith Stalick for useful discussions on intensity rescaling and Dr Stalick, Dr Diane Toby and Brian McMahon for reviewing an initial version of this article. Certain commercial products are identified in this report in order to describe the documented software adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that these products are necessarily the best available for the purpose.

#### References

- David, W. I. F. (2003). *J. Res. Natl Inst. Stand. Technol.* **108**. In the press.
- Hall, S. R., Allen, F. H. & Brown, I. D. (1991). *Acta Cryst.* **A47**, 655–685.
- Larson, A. C. & Von Dreele, R. B. (1986). *General Structure Analysis System (GSAS)*. Report LAUR 86-748. Los Alamos National Laboratory, New Mexico, USA.
- McCusker, L. B., Von Dreele, R. B., Cox, D. E., Louër, D. & Scardi, P. (1999). *J. Appl. Cryst.* **32**, 36–50.
- Rietveld, H. M. (1969). *J. Appl. Cryst.* **2**, 65–71.
- Toby, B. H. (2003a). *International Tables for Crystallography*, Vol. G, *Definition and Exchange of Crystallographic Data*, edited by S. R. Hall & B. McMahon, ch. 3.3, *Classification and Use of Powder Diffraction Data*. In the press.
- Toby, B. H. (2003b). *J. Appl. Cryst.* **36**, 1288–1289.
- Toby, B. H., Von Dreele, R. B. & Larson, A. C. (2003). *J. Appl. Cryst.* **36**, 1290–1294.
- Young, R. A. (1995). Editor. *The Rietveld Method, International Union of Crystallography Monographs on Crystallography*. Oxford: International Union of Crystallography/Oxford University Press.