**Tutorial for batch data processing**

Lei Wang

**1. Datasets used for this tutorial: Ibuprofen**

3D ED datasets of Ibuprofen were collected on a ***JEOL 2100*** transmission electron microscope (TEM) with a ***Timepix*** camera. The acceleration voltage of TEM was *200 kV* and the targeting crystal was continually tilted with an oscillation range of 0.58. Data was collected at *cryogenic* temperature using a Gatan 914 cryo-holder. The software ***Instamatic*** was used for data collection. Every 20th frame was used to image the crystal for re-centering.

Data downloads from Zenodo :

<https://zenodo.org/records/10567517>

**2. Installation**

Download the code package from **Gitlab** or **our box folder** and save it in one folder.

Please check that you have all the scripts as shown in Figure 1.

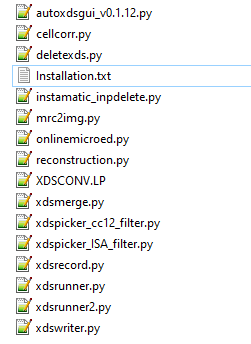


Figure 1 Script list for GUI

The GUI is written in Python and uses the same environment as your XDS. Before running the GUI, several Python packages should be installed. Run the following command lines in the same environment as your XDS. Make sure you have updated wsl to version 2 or have X-service installed if you have wsl 1.

*sudo apt-get update*

*sudo apt-get upgrade*

*sudo apt-get install python3-pip*

*sudo apt-get install python3-tk*

*pip3 install pandas*

*pip3 install mrcfile*

*pip3 install fabio*

*pip3 install openpyxl*

*sudo apt install libreoffice #This one is not necessary. It is just for displaying excel in Linux.*

To open the GUI, run the following command line in your bash (as you run XDSGUI):

*python3 autoxdsgui\_v0.1.12.py*

(current version)

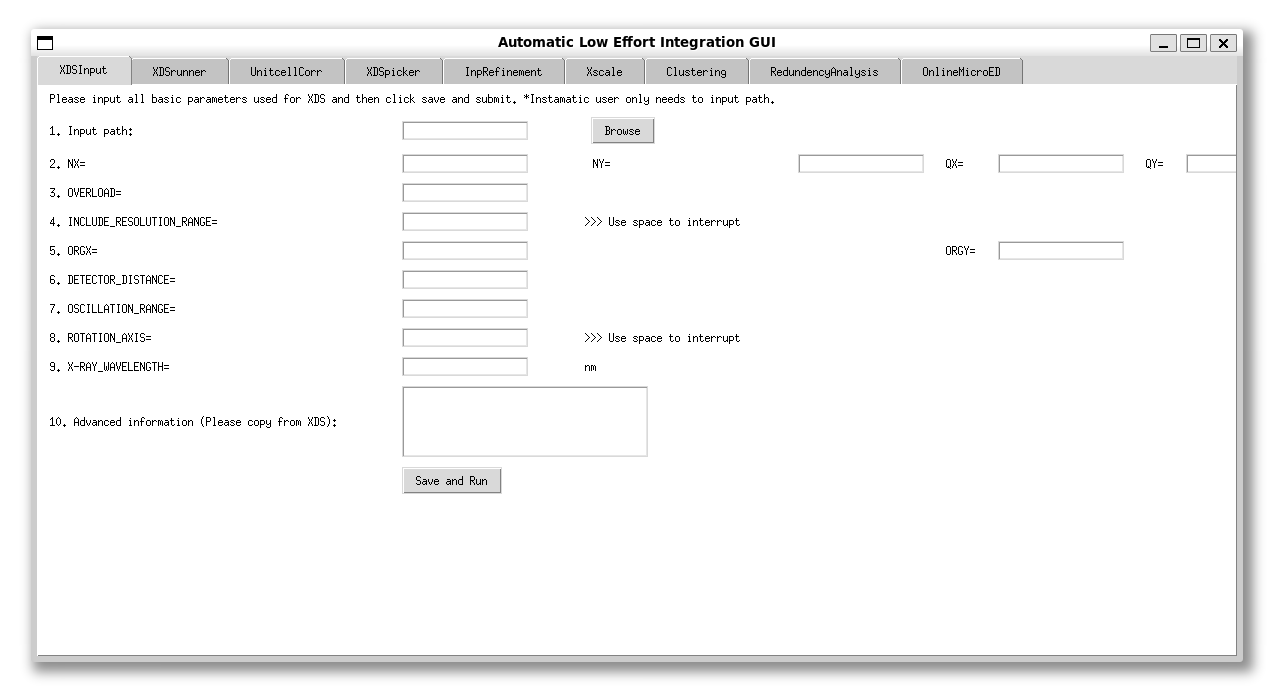


Figure 2 User interface of AutoGUI

**3. Batch data processing**

24 ibuprofen datasets are used as the tutorial data. Not all of them will be used for final structure determination.

*3.1 XDSInput*

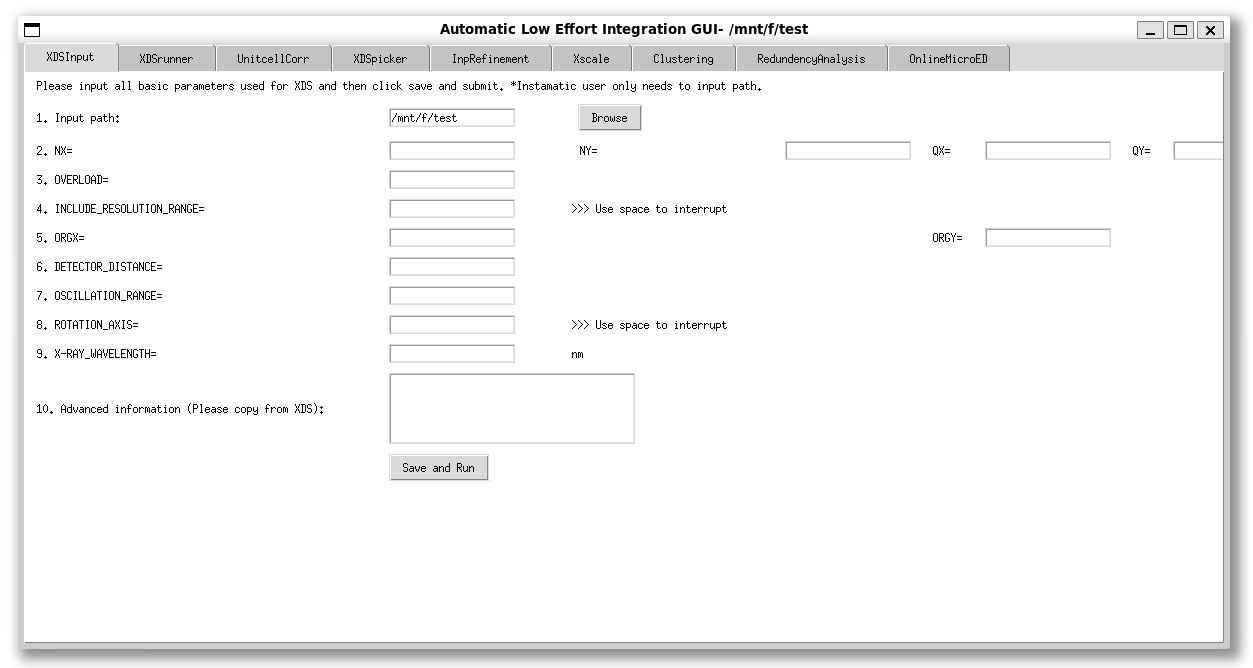


Figure 3 Basic input parameters for XDS

Since the data is collected using Instamatic, only the input path should be selected. As shown in Figure 3,

*Click “Save and Run”;*

It is noted that the title of the GUI will also change.

*3.2 XDSrunner*

Since the data is collected using Instamatic, we only need to

*click “Instamatic xdsinp”;*

Then,

*click “Yes”;*

The GUI will modify the xdsinp generated by Instamatic slightly and make it readable in this GUI.

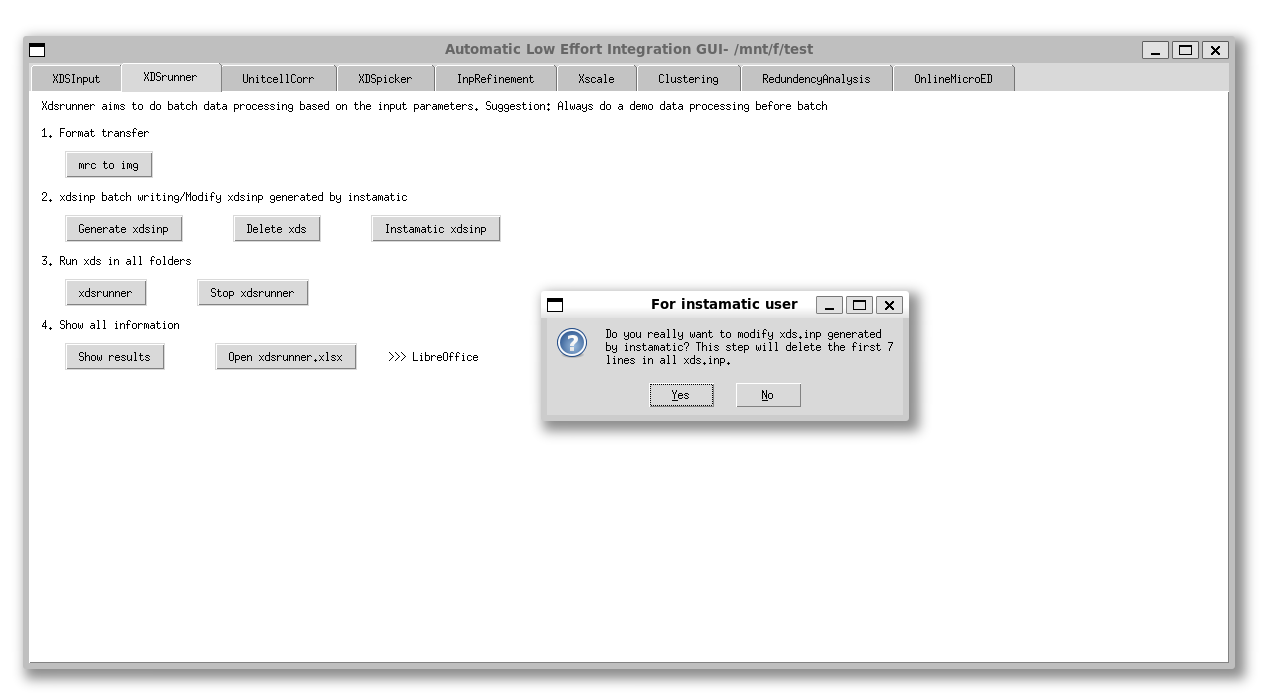


Figure 3 The entry for instamatic users in XDSrunner page

*Click “xdsrunner”*

Xdsrunner will run all xds.inp in the target folder, including all subfolders and summarize all general information from xds outputs. Data processing speed depends on the performance of your computer. Both the animation and bash window can tell when the processing will finish.

When the data processing finishes,

*click “Show results”;*

in step 4 to display the result table. Editing the Excel table is also possible by

*clicking “Open xdsrunner.xlsx”;*

(Need to install LibreOffice or just open it in windows).

As shown in Figure 4, there is one ice crystal. You can delete it or remove it later in XDSpicker.

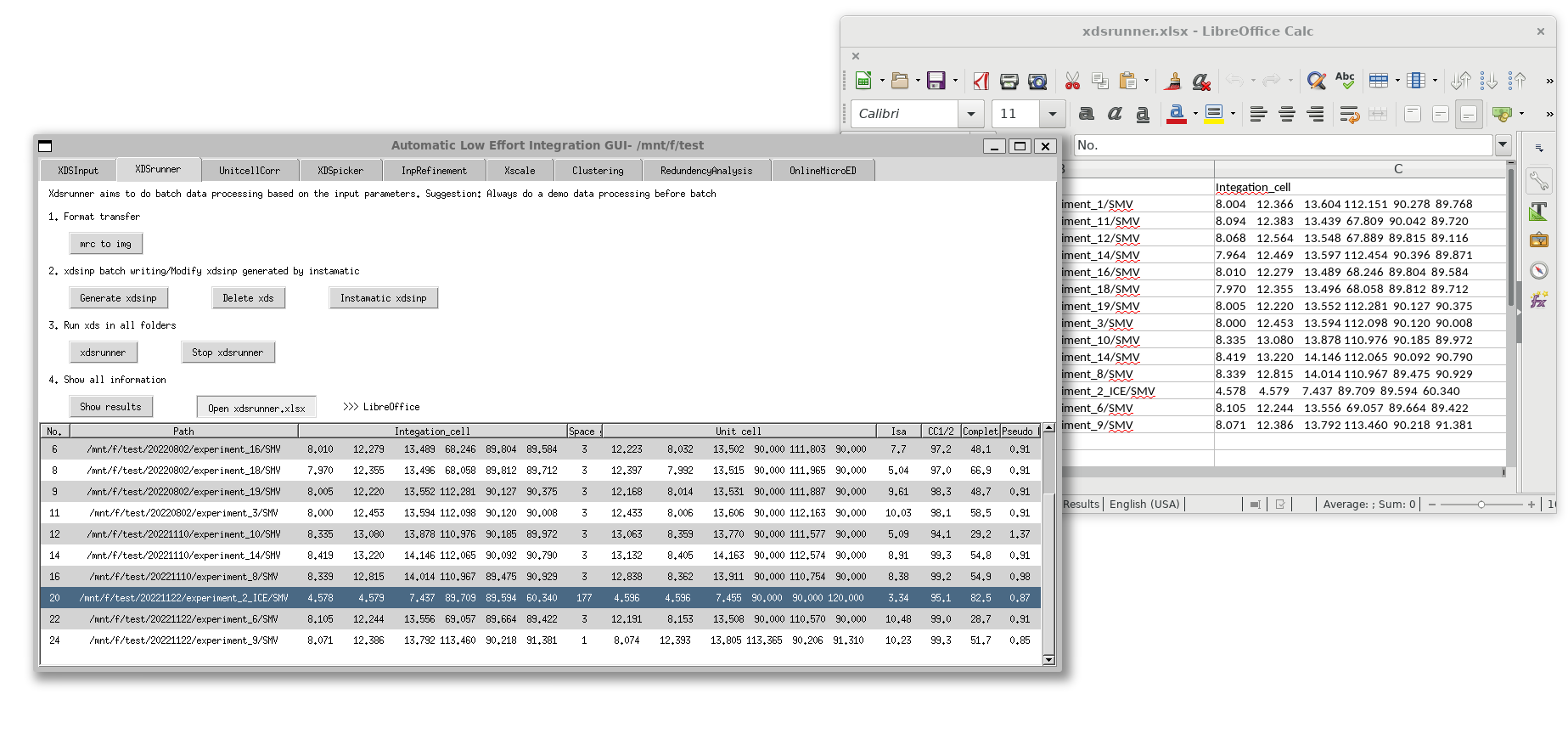


Figure 4 The result display after running xdsrunner

*3.3 UnitcellCorr*

Ibuprofen is beam-sensitive. Data merging is necessary to solve this structure. On this page, unit cell and space group corrections are first carried out. Starting unit cell parameters should be given and XDS will refine the result based on the starting cell. Please enter this information to make sure we can get the same result.

**Space group: 4**

**Unit cell: 12.370 8.024 13.537 90.000 111.733 90.000**

To do xds correction, please click the following buttons:

*Click Save;*

*Click Update cell information;*

Once all inp files are modified, you can:

*Click xdsrunner;*

Click *Show Results* or *Open xdsrunner2* to see the result.

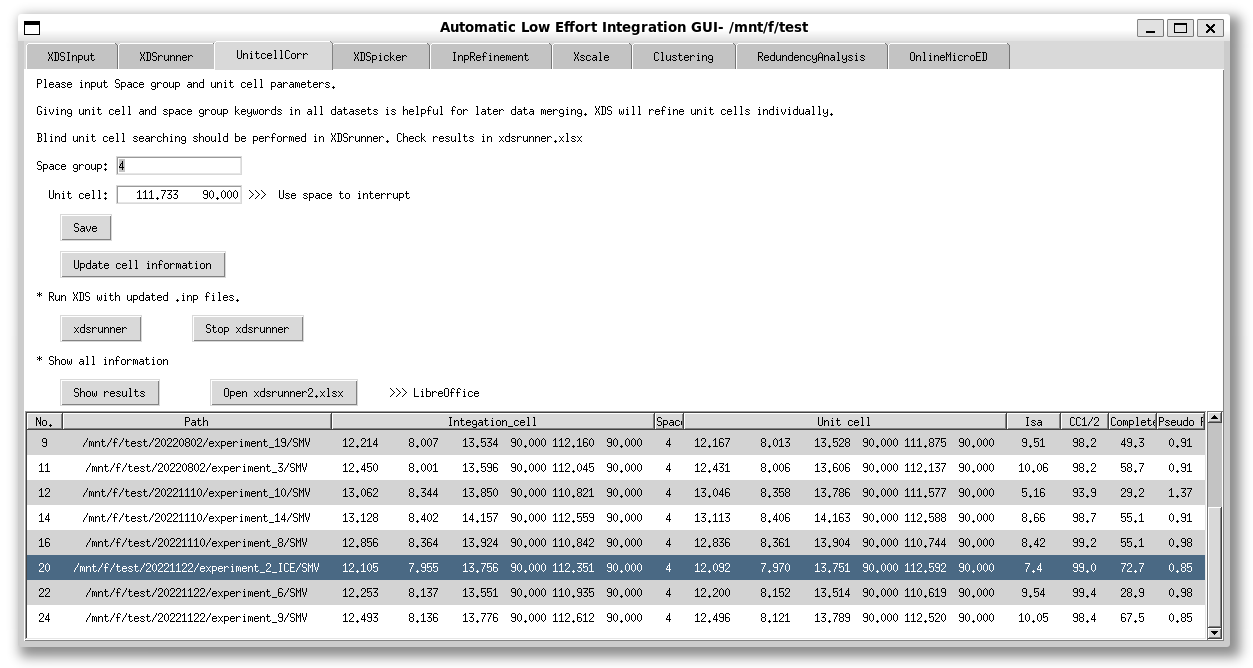


Figure 5 XDS correction result

Surprisingly, the ice crystal can also be processed with the correct space group and unit cell. However, if you check the data in detail, you will find the I/sigma is lower than 1 and the indexing is wrong. Here we can:

*Click Open xdsrunner2.xlsx*

And delete this data from the excel as shown below.

*Select* *“Use Microsoft Excel 2007-2013 xml format”*.

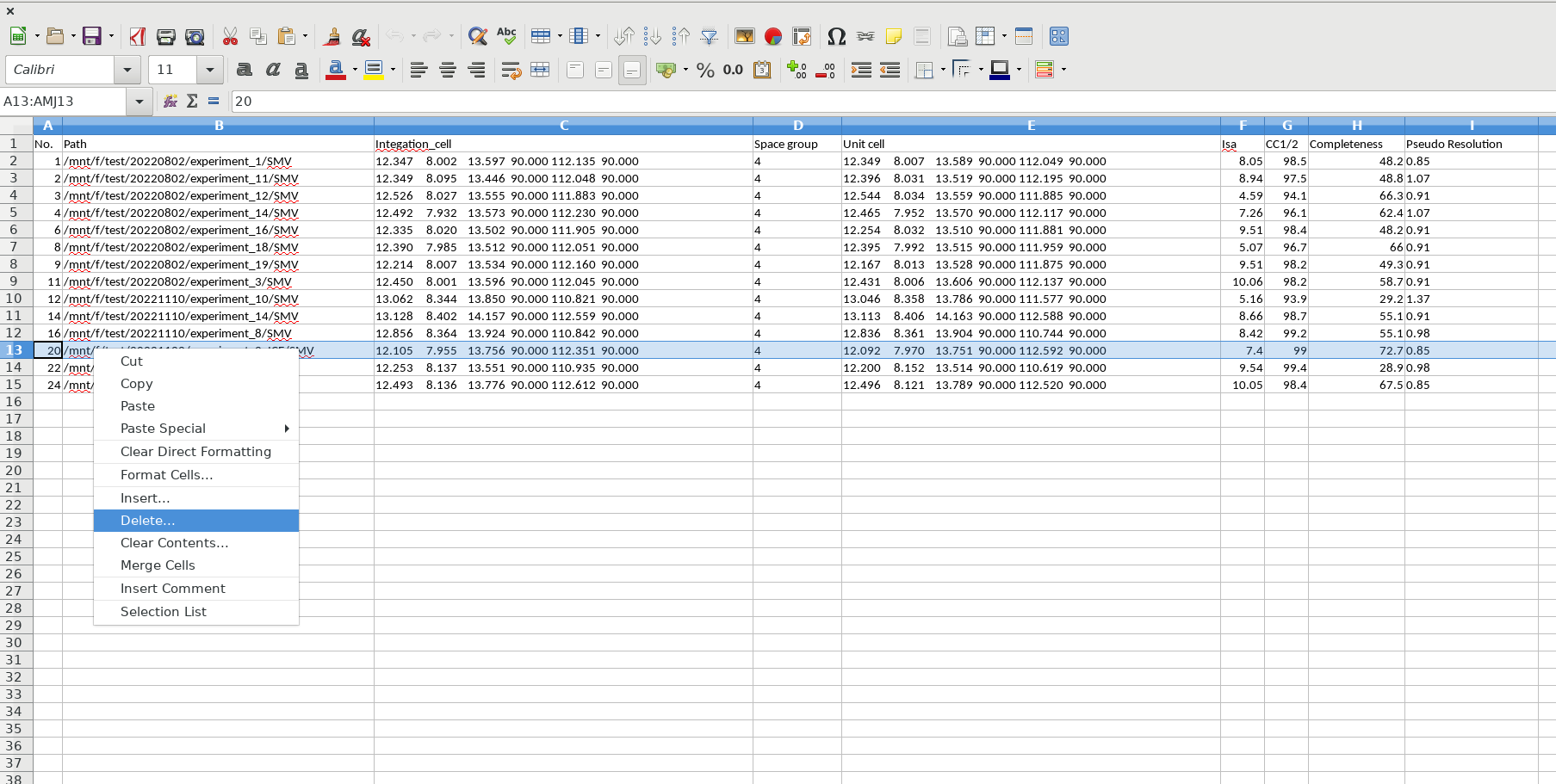
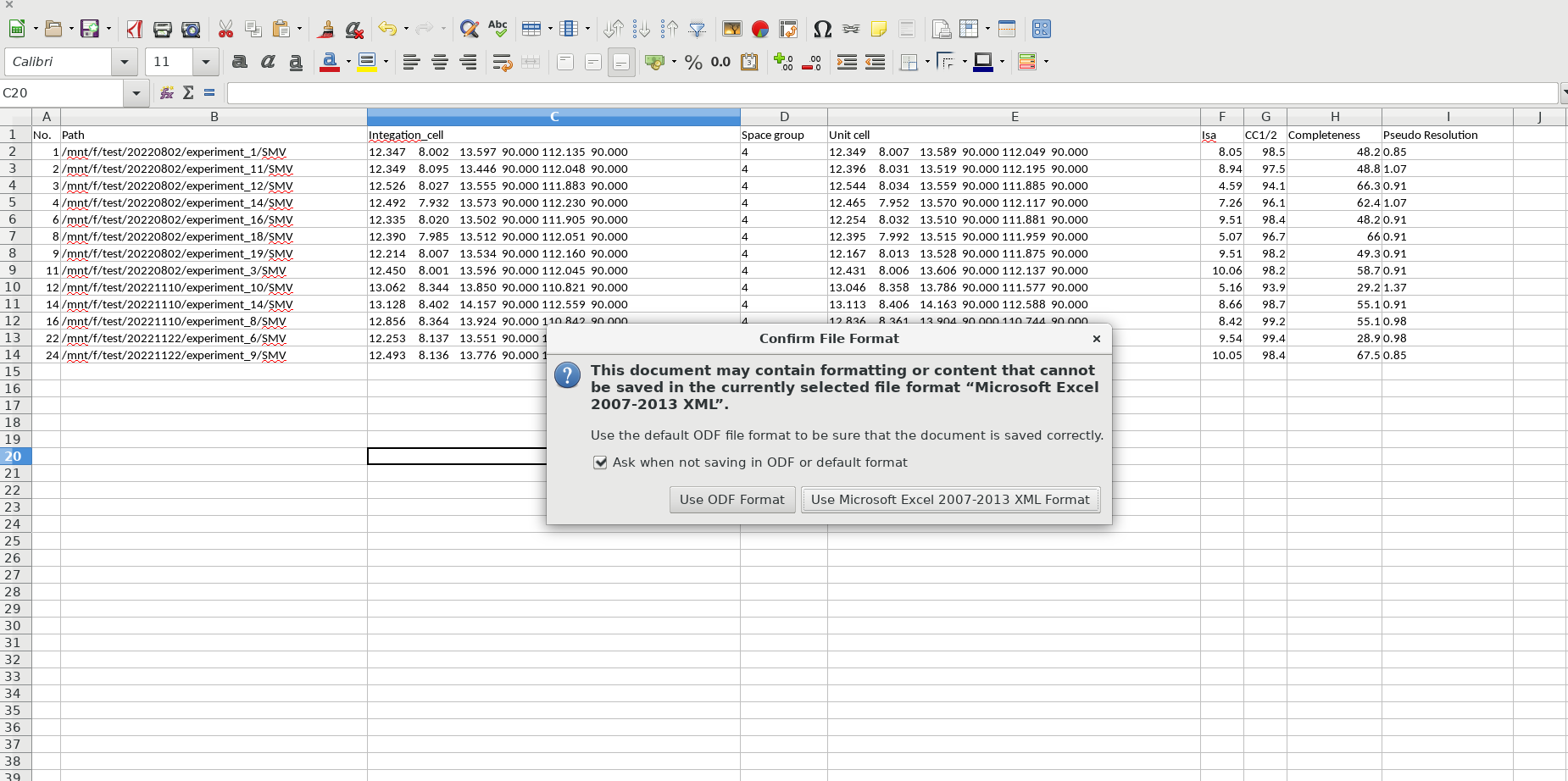
 

Figure 6 Delete the ice data and save the result

*3.4 XDSpicker*

“In this step, we select datasets used for data merging. You can always select data manually by making xdspicker.xlsx (Copy xdsrunner2.xlsx, rename it as xdspicker.xlsx and choose datasets you want).”

To pick datasets automatically, two simple filters are available right now.

In this tutorial, the cc1/2 filter is employed.

Use data with CC1/2 > **95**, and

*Click Run cc1/2 filter*

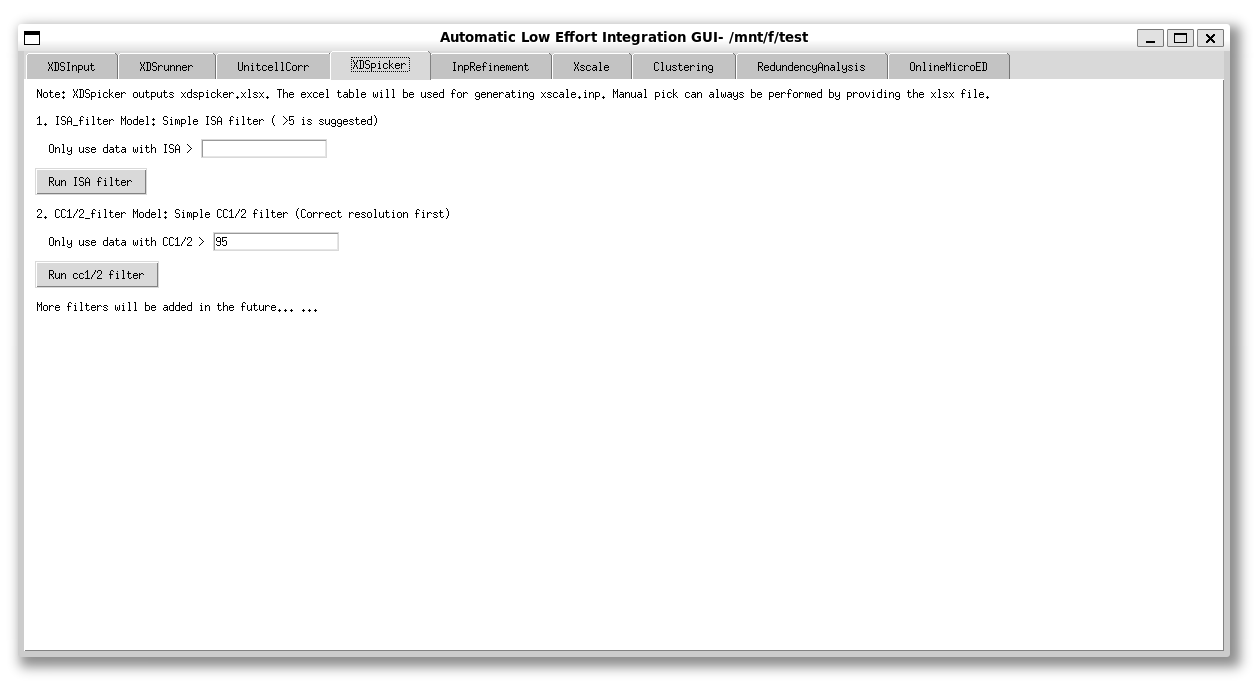


Figure 7 CC1/2 filter for selecting data

*3.4 Xscale*

*Click Merge Data;*

*Click Show result;*

Here the merging result can be checked.

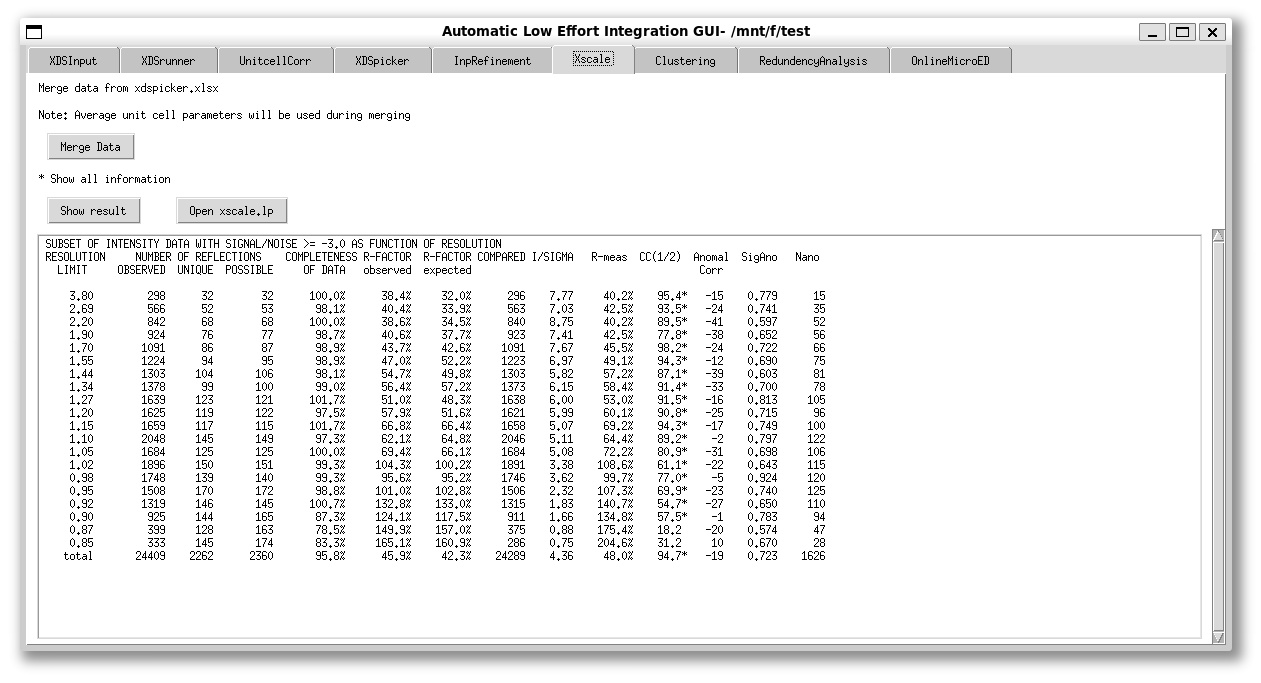
**

Figure 8 One click xscale

After this step, we can find the all.HKL file in the merge folder.

The merging data is not extremely good but can be used for structure refinement.