INSTRUCTIONS FOR THE PHASE CLUSTERING AND IDENTIFICATION USING THE PLUGINS FOR "IMAGEJ"

Beat Münch *

Empa Materials Science and Technology, 8600 Dübendorf, Switzerland

Address:

Empa
Laboratory for Concrete & Construction Chemistry
Beat Münch
Überlandstrasse 129
CH-8600 Dübendorf
Switzerland

Tel. +0041-44-823 49 68
Fax. +0041-44-823 40 35

beat.muench@empa.ch

* Corresponding author: beat.muench@empa.ch

This article describes the steps needed for the clustering and phase identification of elemental maps as described in [1]. The required data is a set of elemental map images assessed by EDX mappings conducted on XRD. The calculations are performed by using some of the *xlib_.jar* plugins of *ImageJ* or *Fiji*. The required ImageJ library **"xlib_.jar"**, some text documents required for the plugins, the data of the test example referenced here, as well as this manual can be freely downloaded from

ftp://ftp.empa.ch/pub/empa/outgoing/BeatsRamsch/lib/

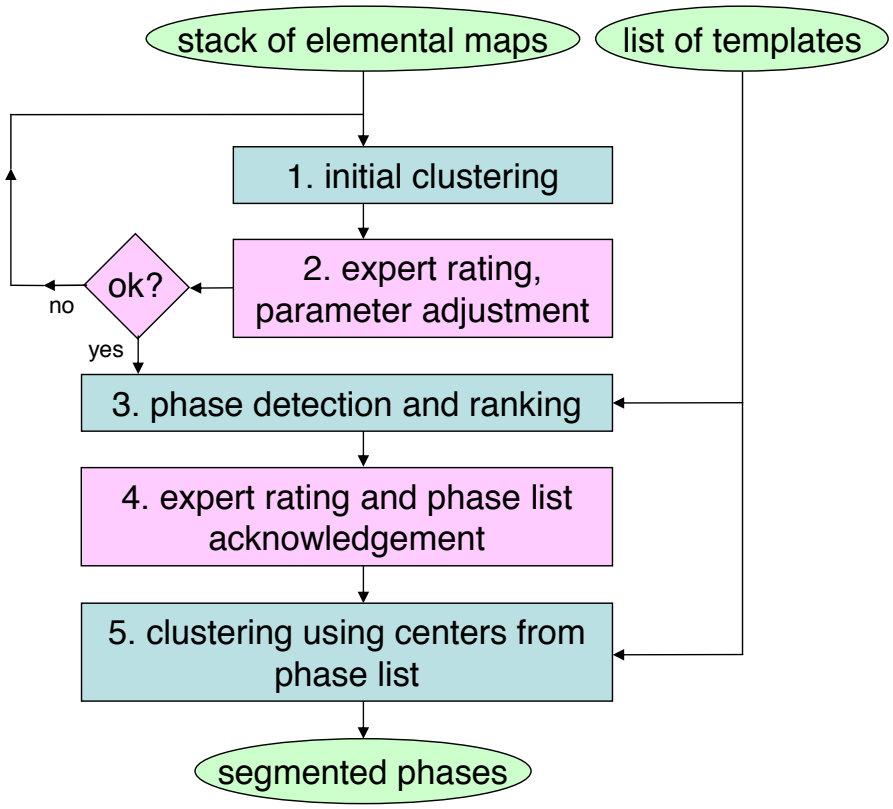

Fig. 1: *Concept of the phase segmentation and clustering. The input consists of images from elemental maps and a list of possible templates, the output of a mask containing multiple segmented phases. The processing steps shaded in light blue are running automatically without any user interaction. The steps requiring some expert decisions and parameter adjustment are shaded in pink.*

1. Initial clustering

1. launch ImageJ.
2. drag and drop to ImageJ all images containing the elemental maps.
3. choose the map image with the lowest atomic number (the image for *F* in this example).
4. run the clustering plugin ***"Plugins⟶Beat⟶Cluster Image"*** on it. Under "*Source for clustering*" in the lower part of the plugin window, the path of the chosen image should appear.
5. select the ***"Clustering type"*** (K-means in this example) and set the required parameters, e.g. the ***"Number of clusters"*** (26 in this example).
6. optional constraints can be provided by changing one or more of the numeric values in the entries ***"Total minimal distance to closest neighbour cluster"***, ***"Single minimal distance to closest***

neighbour cluster", ***"Minimal number of pixels in a cluster"***, and ***"Maximally allowed ratio of surface pixels in percent"***. The last two constraints prove to be especially useful. In the current example, ***"Total minimal distance to closest neighbour cluster"*** was set to 5 and ***"Maximally allowed ratio of surface pixels in percent"*** to 77. Constraints should only be given if it is not possible to obtain desirable clustering results without any of them.

7. from the ***"Provide information about clusters"*** pulldown menu, select the last option ***"text, Gaussian graph with std.dev. bars, separated"*** to enable a result text window as well as pseudo EDX plots to be generated. The graphic may be written to PNG or to EPS by inserting an appropriate file name in the text box ***"Plot graphics to this file (extension .png or .eps)"***.
8. check the check box ***"Format readable for 'Closest To Cluster'"*** in order to obtain a text format which is readable by the program ***"Closest Cluster"***.
9. enter the list of measured chemical elements into the text window ***"Captures for plot of x-axis (optional)"*** by separating them with spaces. Don't forget the first element which is given by ***"Source for clustering"***.
10. by choosing ***"Additional sources"***, add all maps in the order of the atomic number. Selecting an image by the pulldown menu for the first time will add it to the sources text field, selecting it for the second time will remove it again. By checking ***"Add all opened sources"***, the entire list of images will be inserted to the sources text field. However, the order of the images then probably might not follow the order of the atomic number. As a remedy, it is possible to edit the text by copy-paste operations. Yet, it is important to uncheck the ***"Add all opened sources"*** text field before running the clustering, because checking this flag causes the images to be loaded in the original order when running the plugin. To prohibit the loss of the edited text after unchecking the flag, it can be copied and then be pasted after unchecking the flag.
11. if desired, a binary mask for the regions where clustering should run can be provided by a mask image (values 0: unmasked, 255: masked). Insert the path to the text box ***"Mask for clustering"***.
12. user interface see Fig.2. Now run the clustering process by pressing the ***"OK"*** button. The clustering image and a graphic window with pseudo EDX plots will result, as well as a text window containing the mean values for all clusters and map images.

2. Expert rating, parameter adjustment

13. once you are satisfied with the results after visual examination and appropriate parameter adjustment, save the contents of the resulting text window into a text file, e.g. by using ***"File⟶Save As..."*** menu in the text window (in the current example to *"Mess_Gestein_02C.txt"*, included in the download folder). For all clusters, it contains the mean values of all included elemental maps.
14. prepare a text list of possible template phases that eventually might exist in the measured sample. The format of the text file is explained in step 19. It is useful to start with a preferably complete list of possible template phases, e.g. of all minerals in the case of the rock sample. It is suggested to exclude all template phases that are impossible to appear according to the pre-knowledge about the sample. This facilitates the subsequent interactive selection of considered template phases after the ranking. Excluding a template phase can be easily achieved by setting a hash ("#") at the beginning of the line.

3. Phase detection and ranking

15. run the phase detection and ranking plugin ***"Plugins⟶Beat⟶Closest Cluster"***.
16. from the ***"Find what"*** pulldown menu, select ***"Closest substances from chemical pattern list"***.
17. add into the text box ***"Molar weights file: path to text file containing a list of molar weights of all elements"*** the path to the text file containing the molar weights for all elements. The following text format is required:

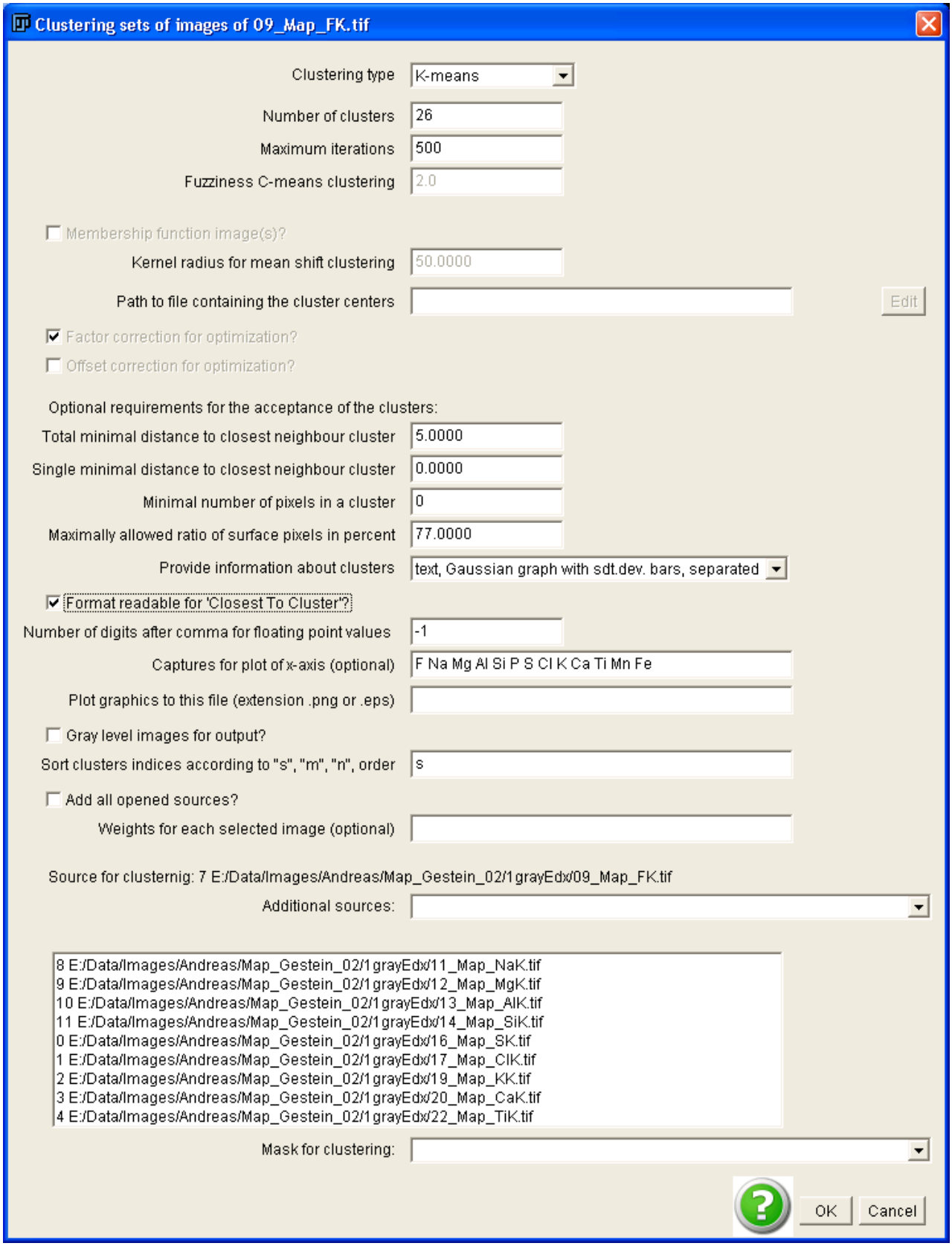

*Fig. 2: User interface for the initial clustering (**"Plugins⟶Beat⟶Cluster Image"**).*

```
20      Ca      Calcium         40.0784
21      Sc      Scandium        44.9559
22      Ti      Titan           47.8671
```

A suitable text file called *"xMolekulargewicht.txt"* is included in the download folder. The path to the text file can either be directly copied to the text box, or the respective file can be chosen from a file selector by pressing the ***"Edit"*** button to the right of the text box. A text editor containing the file will appear. Just close the text editor window by pressing the "X" to the top right window corner if no changes should be achieved.

18. add a comma separated list of all used elements into the text box ***"Considered chemical elements (space separated)"***, e.g. *"F Na Mg Al Si P S Cl K Ca Ti Mn Fe"*.
19. add into the text box ***"Chemical pattern file: path to text file containing a list of chemical pattern phases"*** the path to the text file containing a list of possible chemical labels together with its chemical formula in each line. Giving more than one chemical formula for a label is also possible. Lines starting with a hash are not being considered in the rating. Thus, instead of removing a line and loosing its information, quickly inserting a hash is a useful option. An example text file where some lines are excluded and where *"muskovite"* and *"biotite"* may be defined by more than one possible chemical formula is given here:

```
muskovite               Si6.05Al1.95Al3.86Fe0.09Fe0.00Mn0.01Mg0.02Li0.26Na0.17K1.72F0.39(OH)3.68
muskovite               Si6.60Al1.40Al2.96Ti0.09Fe0.67Fe0.09Na0.17K1.95(OH)3.92
#biotite                Si7.55Al0.45Al0.18Ti0.01Fe2.43Fe0.32Mn0.00Mg1.11Na0.04K1.67(OH)4.00
biotite                 Si5.72Al2.28Al0.56Ti0.08Fe0.30Fe0.28Mg4.78Na0.03K1.75F0.28(OH)3.95
biotite                 Si5.68Al2.32Al0.34Ti0.17Fe1.93Fe0.03Mn2.91MgCa0.12Na0.06K1.51(OH)4.00
biotite                 Si5.50Al2.50Al0.61Ti0.34Fe0.42Fe2.87Mn0.16Mg1.40Na0.13K1.89F0.46(OH)3.18
#biotite                Si6.64Al1.4Al2.6Ti0.02Fe0.08Fe0.36Mn0.05Mg0.02Li2.4Ca0.01Na0.16K1.6F1.7(OH)2.5
eastonite               KMg2Al3Si2O10(OH)2
illite                  (K,H2O)Al3Si3O10(OH)2
#ilmenite               FeTiO3
```

20. add into the text box ***"Pattern data file: path to text file containing a list of data values"*** the path to the text file generated from the initial clustering containing the mean values of all included elemental maps. The first line of the text file must be a list of the included elements starting with a hash. For each cluster, a line with the mean values follows. An example text file:

```
# F Na Mg Al Si P S Cl K Ca Ti Mn Fe
#          F      Na     Mg     Al     Si     P      S      Cl     K      Ca     Ti     Mn     Fe
pix86917   0.35   10.88  1.12   21.69  59.97  0.84   0.66   0.65   1.97   1.43   0.27   0.03   0.14
pix56183   0.82   2.18   3.15   28.06  44.58  0.88   0.69   0.74   13.74  1.11   0.62   0.09   3.34
```

21. optionally add into the numeric box ***"Maximal error distance from the first ranked candidate"*** the maximal distance the regarded candidates might disagree from the first ranked candidate. If lower than 0, no maximal distance is requested.
22. optionally add into the numeric box ***"Number of first templates considered as candidates"*** the maximal count of the listed candidates. If lower or equal 0, no maximal count is provided.
23. check the check box ***"Instead of molar weights, consider atomic ratio only"***.
24. check the check box ***"Factor correction prior to optimization (default and recommended)"***. A linear factor for fitting the EDX characteristics to the modelled characteristics of some given templates can be physically justified due to varying material densities.

25. the check box ***"Offset correction prior to optimization (not default and questionable)"*** is unchecked by default. Checking yields slightly better fits. However, there is no physical justification of using an offset for improving the fits.
26. optionally check the check box ***"Display plots"*** such that the corresponding model EDX plots are released for all first ranked candidates.
27. user interface see Fig.3. Now run the phase detection process by pressing the ***"OK"*** button.

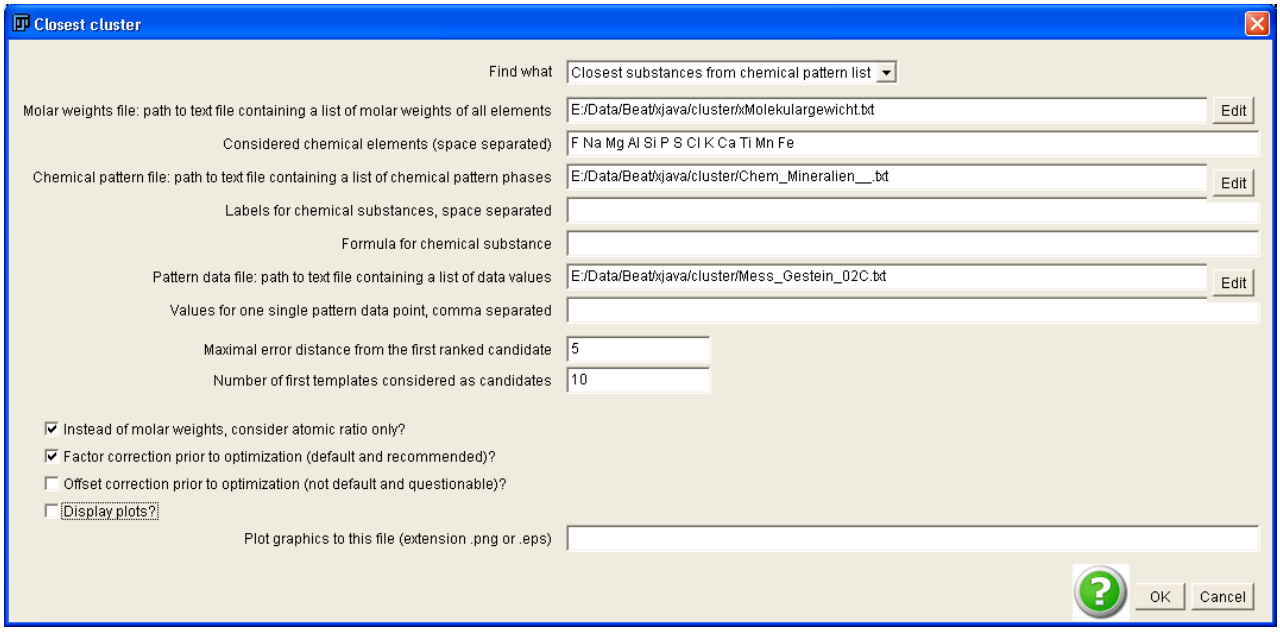

Fig. 3: *User interface for the phase detection and ranking (**"Plugins⟶Beat⟶Closest Cluster"**).*

4. Expert rating and Phase list acknowledgement

28. the former process results in a list of possible chemical candidates for each phase, ranked by the distances to the model values. Reviewing this list by a domain expert is required. Thereby, a collection of all involved phases must be ultimatively decided.

5. Clustering using centers from phase list

29. again, run the phase detection and ranking plugin ***"Plugins⟶Beat⟶Closest Cluster"***.
30. from the ***"Find what"*** pulldown menu, select ***"Chemical patterns from substance list"***.
31. as in step 17, add into the text box ***"Molar weights file: path to text file containing a list of molar weights of all elements"*** the path to the text file containing the molar weights.
32. as in step 18, add a comma separated list of all used elements into the text box ***"Considered chemical elements (space separated)"***, e.g. *"F Na Mg Al Si P S Cl K Ca Ti Mn Fe"*.
33. as in step 19, add into the text box ***"Chemical pattern file: path to text file containing a list of chemical pattern phases"*** the path to the text file containing the list of chemical labels together with their chemical formula.

34. add into the text box ***"Labels for chemical substances, space separated"*** the labels of the final collection of all involved phases from step 28. The labels must be separated by spaces.
35. as in step 26, you might optionally check the check box ***"Display plots"*** such that the corresponding model EDX plots are released for all first ranked candidates.
36. user interface see Fig.4. As in step 27, now run the pattern generation process by pressing the ***"OK"*** button.

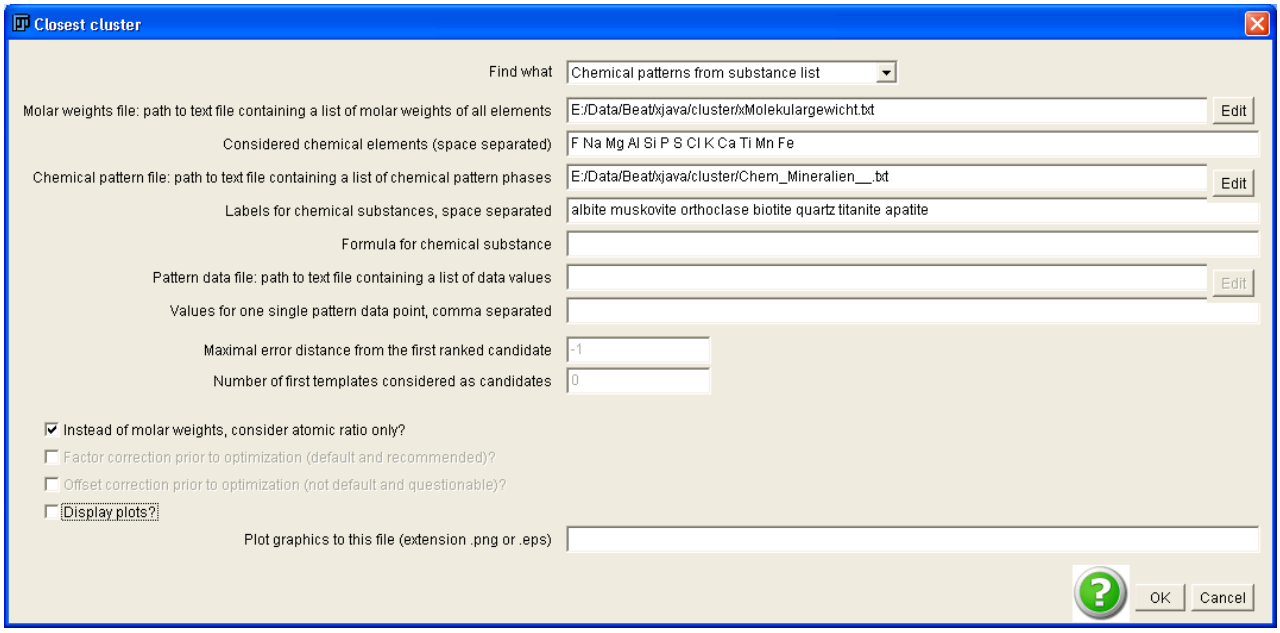

Fig. 4: *User interface for the generation of the clustering patterns (**"Plugins⟶Beat⟶Closest Cluster"**).*

37. save the contents of the resulting text window into a text file, e.g. by using ***"File⟶Save As..."*** menu in the text window (in the current example to *"Model_AlbMusKaliBioQuaTitApaB.txt"*, included in the download folder).
38. again, run the clustering plugin ***"Plugins⟶Beat⟶Cluster Image"*** and make sure that the correct sources are provided as explained in steps 4 and step 10.
39. choose ***"Predefined centers"*** from the pulldown menu ***"Clustering type"***.
40. add into the text box ***"Path to file containing the cluster centers"*** the path to the text file generated above (in the current example to *"Model_AlbMusKaliBioQuaTitApaB.txt"*, included in the download folder). This can be achieved either be directly copying the path to the text box, or by pressing the ***"Edit"*** button to the right of the text box and by choosing the file from the file selector.
41. as in step 24, check the check box ***"Factor correction for optimization (default and recommended)"***.
42. as in step 25, the check box ***"Offset correction for optimization (not default and questionable)"*** is unchecked by default and checking is optional.
43. as in step 7 from the ***"Provide information about clusters"*** pulldown menu select the last option ***"text, Gaussian graph with std.dev. bars, separated"*** to enable a results text window as well as pseudo EDX plots to be generated. Make sure the check box ***"Format readable for 'Closest To Cluster'"*** is unchecked.

44. as in step 11 if desired, a binary mask for the regions where clustering should run can be provided by a mask image (values 0: unmasked, 255: masked). Insert the path to the text box ***"Mask for clustering"***.
45. user interface see Fig.5. As in step 12, now run the clustering process by pressing the ***"OK"*** button. The final clustering result and a graphic window with pseudo EDX plots will result, as well as a text window containing the cluster mean values.

References

1. B.Münch, L.Martin, A.Leemann, *"Segmentation of elemental EDS maps by means of multiple clustering combined with phase identification"*, Journal of Microscopy, Vol.000 Issue 00, pp1–16, 2015.

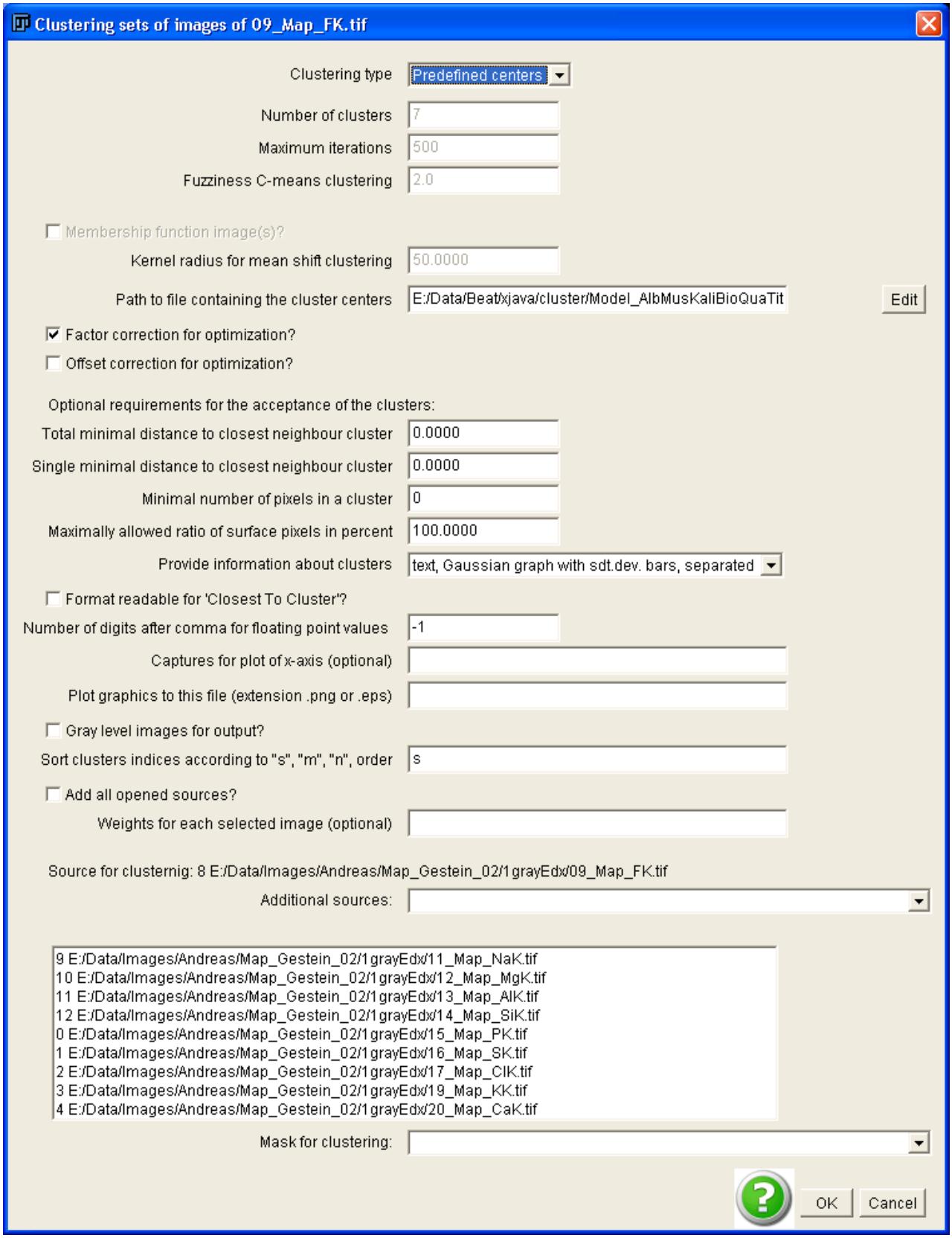

Fig. 5: *User interface for the final clustering by using the clustering pattern (**"Plugins⟶Beat⟶Cluster Image"**).*