

## Interpreting and Validating Results

PolySNAP incorporates many tools that can help the user to interpret the results.

### Cell Display:

The cell display represents the samples as discs in an array (see Figure 1.a). The discs are coloured according to the group they have been assigned to.

Figure 1.a - a cell display

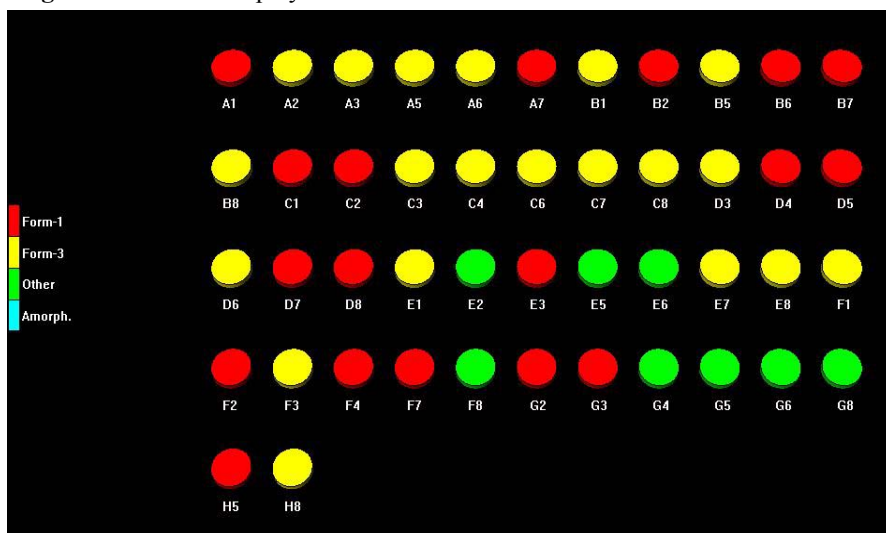
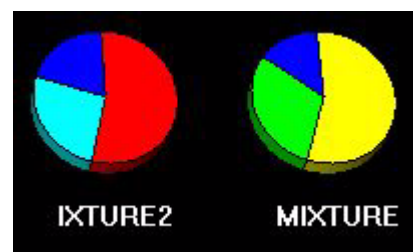


Figure 1.b - a mixture sample pie chart



The portions that the cell have been divided up into represent the relative porportion of that known phase thought to be in these mixtures

When known phases have been provided the cell display represents the actual reference sample that the samples are believed to match. Samples that are believed to be mixtures are represented as pie charts which indicate the relative porportions of each known phase thought to be in the mixture (see Figure 1.b).

### Dendrogram:

The dendrogram is at the heart of PolySNAP. There is a seperate quick guide explaining the features of the dendrogram and how it can be used in interpretation. Generally a dendrogram that exhibits clear distinct clusters suggests a well-behaved dataset where there is real scope to sub-divide the data even further.

### 3D plots:

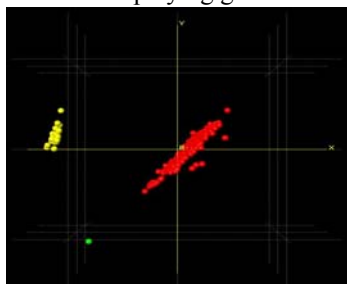
Each pattern is plotted as a sphere in a 3D box. The closer together two patterns are on this plot, the more similar they are. Similar samples tend to clump together. A good cluster in the 3D plot should therefore have all of its spheres close together in the plot (see Figure 2.a). Ideally, the cluster should not be diffuse. A 3D plot where a cluster is large and diffuse, or contains spheres which lie relatively far outside the main body of the group may indicate that the cut-level needs to be lowered. A 3D plot where the spheres do not

---

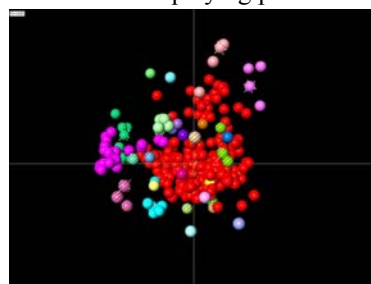
appear to separate into any distinct groups may suggest that there may not be any structure in the dataset (see Figure 2.b).

---

**Figure 2.a** - a 3D Plot displaying good clustering



**Figure 2.b** - a 3D Plot displaying poor clustering



It is useful to switch between the dendrogram and the 3D plot to set the cut-level, as ideally the clustering in the two displays should be consistent with each other - *i.e.* separate clusters should correspond with the groupings in the other display. 3D Plot colours update as the dendrogram cut-level is altered to assist with this.

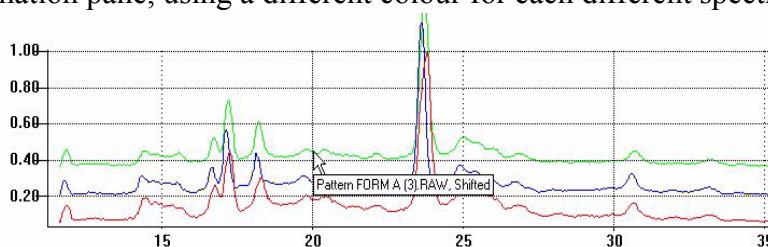
### Plot comparison:

*PolySNAP* contains two different 3D plot methods, the MMDS (Metric Multi-Dimensional Scaling) plot and the PCA (Principal Component Analysis) plot. The plot displays look similar but use different methods of calculation, and the positions of the spheres are different. Ideally, the clustering should be consistent to a reasonable degree. This is a further indication of a good set of results. When one of the methods produces significantly better results than the other two asterisks are shown on its tab header.



### Overlaying spectra:

By selecting multiple samples at the same time the spectra for these samples are superimposed on each other in the sample information pane, using a different colour for each different spectra



This is a useful tool for comparing the patterns directly. The spectra can be offset so they are displayed clearly and labels can be displayed by holding the mouse cursor over one of the spectra as shown above.

### Other features:

There are several different methods for validating the data, including silhouettes which give a quick overview of how well-defined a cluster is and allows easy identification of outliers. As well as validation there is also a logfile that records all of the numeric output from the the import, processing and analysis carried out on the data. For more details see the full program manual.

---