

HOW TO COMBINE PXRD, RAMAN AND OTHER 1-D DATA IN HIGH THROUGHPUT POLYMORPH/SALT/CO-CRYSTAL STUDIES

Christopher Gilmore, Gordon Barr and Wei Dong, WestCHEM, Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, UK.

Powder X-ray diffraction is usually considered the gold standard of fingerprinting methods in high throughput studies but other techniques such as Raman, near infra-red, and terahertz spectroscopy can have a major role to play also. The new combined X-ray - Raman instrument from Bruker (the D8 SCREENLAB) is the first in what may well be many new instruments from many manufacturers that work with multiple radiation sources.

We have established a methodology and an associated computer program, POLYSNAP, that is designed for high throughput PXRD studies [1-4], and we are now adapting the techniques in this software to process multiple data types. The problem arises as to how best to combine such disparate data with quite different properties. For example, PPXRD usually shows large differences between different polymorphs, but in Raman spectroscopy the differences are subtle and related to such things as H-bonding.

The PolySNAP methodology compares each pattern to every other on a point by point basis using the entire measured data set to generate a correlation matrix. It is possible to weight each individual matrix and thus combine them to get an overall matrix, but how are the weights to be estimated?

We use a dynamic weighting scheme [5]. Let \mathbf{D}_k be the squared distance matrix of dimension $(n \times n)$ for n data sets and K data types *e.g.* if you have PXRD + Raman and 20 samples then $n=20$ and $K=2$. We want a *group average matrix* \mathbf{G} that describes our data. We specify a set of diagonal weight matrices, \mathbf{W}_k , We now need to minimise:

$$\sum_{k=1}^K \|\mathbf{B}_k - \mathbf{G}\mathbf{W}_k^2\mathbf{G}'\| \quad (1)$$

where

$$\mathbf{B}_k = -\frac{1}{2}(\mathbf{I} - \mathbf{N})\mathbf{D}_k(\mathbf{I} - \mathbf{N}) \quad (2)$$

and

$$\mathbf{N} = \mathbf{1}\mathbf{1}' / n \quad (3)$$

$\mathbf{1}$ is the unit matrix. Solving equation (1) to get \mathbf{G} and \mathbf{W} is required.

Examples using combined Raman and PXRD data to study polymorph searching for pharmaceutical compounds will be presented.

1. Gilmore, C.J., Barr, G & Paisley, J. (2004). *J. Appl. Cryst.* **37**, 231-242.
2. Barr, G., Dong, W. & Gilmore, C.J. (2004). *J. Appl. Cryst.* **37**, 243-252.
3. Barr, G., Gilmore, C.J. & Paisley, J. (2004). *J. Appl. Cryst.* **37**, 231-242.
4. Barr, G., Dong, W. & Gilmore, C.J. (2004). *J. Appl. Cryst.* **37**, 658-664.
5. Carroll, J.D. & Chang, J.J. (1970). *Psychometria* **35**, 283-319.