Easy Automation & More Accurate Analysis with HighScore(Plus) V4.9

T. Degen, E. Bron, M. Sadki

Malvern Panalytical B.V., Lelyweg 1, 7602 EA, Almelo, The Netherlands

thomas.degen@malvernpanalytical.com

Creating an automated XRPD analysis is often hampered by either a complex GUI (setting it up is difficult) or because the analysis itself is too complicated and requires decisions, loops or other non-linear elements.

In our HighScore(Plus) V4.9 [1] release we have solved both obstacles at once by providing a graphical Flowchart alike design & execution Interface, that can contain decision steps as well as any number of loops. The automation batch is simply put together by dragging and connecting action and decision step boxes.
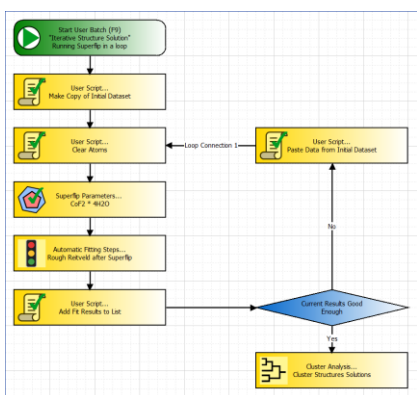

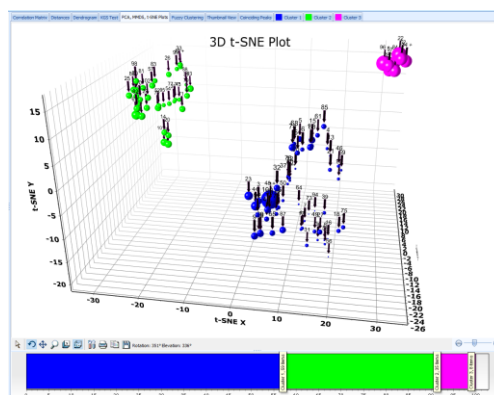**Figure 1.** Flowchart like automation of structure solution with Superflip [7].

Figure 2. t-SNE [2] analysis of resulting solution space.

In addition, supervised and unsupervised learning features are greatly improved, by adding the very popular t-SNE method [2] to cluster (neighbourhood) analysis, and by enhancing the automatic optimization of pre-processings [3] and automated variable selection [4] to the PLSR [5,6] implementation. The cross validation is sped up by a factor of 100 by using multi-threading and algorithm optimizations. All in all, these additions allow to create better, more accurate predictive models in a much shorter time.

[1] T. Degen, M. Sadki, E. Bron, U. König & G. Nénert, *The HighScore Suite*, *Powder Diffr*. Vol. **29**, (2014), Supplement S2, 13-18.

[2] L.J.P. van der Maaten *and* G.E. Hinton, *Visualizing High-Dimensional Data Using t-SNE*. *Journal of Machine Learning Research* 2008, **9**, 2579-2605.

[3] Jan Gerretzen, Ewa Szymańska*, Jeroen J. Jansen, Jacob Bart, Henk-Jan van Manen, Edwin R. van den Heuvel, Lutgarde M. C. Buydens, Simple and Effective Way for Data Preprocessing Selection Based on Design of Experiments, Anal. Chem.* (2015), **87**, 12096-12103.

[4] Loann David Denis Desboulets, *A Review on Variable Selection in Regression Analysis*, *Econometrics* 2018, 6(4), 45

[5] Wold, Herman, (1966). *Estimation of principal components and related models by iterative least squares, in P.R*. Krishnaiaah, *Multivariate Analysis.* New York: Academic Press. (1966), pp. 391-420.

[6] S. de Jong, SIMPLS: *An alternative approach to partial least squares regression, Chemometrics Intell. Lab. Syst.*, (1993), **18(3)**, 251-263.

[7] L. Palatinus and G. Chapuis, Superflip - *a computer program for the solution of crystal structures by charge flipping in arbitrary dimensions*, J. *Appl. Cryst.* (2007), 40, 786-790.

Keywords: PLSR; t-SNE; Cluster Analysis; XRPD;