

WSC-2: Barnaul & Belokurikha lecture abstracts (Saturday 01.03)

Kim H. Esbensen, prof.
Ålborg University Esbjerg, Denmark
www.acacsrge.aue.auc.dk

1. "*Chemometric applications in science and technology – part 1*"
The use of PCA and PLS-regression is illustrated with several applications and case histories from science and technology.
2. "*Chemometric applications in science and technology – part 2*"
The use of PCA and PLS-regression is illustrated with several applications and case histories from science and technology.

Belokurikha lectures:

1. "*Relationships between sampling, analysis and chemometrics*" (Sunday 02.03)
The interrelationships to be illustrated in this lecture can be viewed from the perspective of a broad-view "total analytical process", which presently has one critical missing link within chemometrics: **sampling**. Analytical chemists concern themselves with high accuracy/precision analysis, while chemometricians concern themselves with the subsequent data analysis. Truly there is no end to the successes of both these disciplines – but hardly anyone from these communities pays much attention to the issues surrounding the *ultimate data quality* arising from *sampling*. As all data originate from a sampling process, and because the inherent sampling errors involved in any industrial, natural or laboratory sampling process typically reach two orders of magnitude (or more) larger than even the worst analytical errors, the total data analysis quality (which includes not only accuracy and precision, but also model relevance and model-errors a.o.) is de facto out of control (*sic*). What to do with this situation? – The situation is, thankfully, far from hopeless. This lectures focuses on the new holistic view necessary to master the situation delineated.
2. "*Proper validation – the most important thing in calibration*" (Wednesday 05.03) ← N.B.
Based on the insights gained during WSC-2, this concluding lecture will draw together the necessary requirements for mastering the critical issue surrounding validation: The master example of validation of chemometric prediction models will be used in order to foster a comprehensive understanding of the complexity of multivariate calibration/validation. It will be demonstrated how the very popular *scheme* of cross-validation fails on all accounts, as soon as the issues regarding the missing understanding of the underlying sampling issues are brought out in the open. Proper validation constitutes a fundamental principle to be applied judiciously on any data analysis operation, rather than a mere technique. In this context, scientific logic demands the use of test set validation - the optimal validation procedure should ideally always include doing the initial sampling correctly. But there are admittedly also situation in which a bona fide test set cannot be obtained, and this is the only situation where cross-validation may seem acceptable. Even in this situation there are ample pitfalls of ill-informed use hereof. This lecture deals with the principles of proper validation, focusing on **WHEN**, **HOW** and with what **NECESSARY JUSTIFICATION** one should make use of test set – and cross-validation respectively.