

## Sixth International Symposium “Modern Methods of Data Analysis”

DOI: 10.1134/S1061934809080188

The latest symposium on chemometrics “Modern Methods of Data Analysis” (WSC-6) was held in Kazan in February 2008; it opened the second step of the Drushbametris international program [1]. The increased number of participants was indicative of the interest in the series of WSC conferences from both Russian and foreign scientists. The working language of the symposium traditionally was English. It was organized by the Kazan State University of Architecture and Building Construction and Russian Chemometrics Society with the assistance of the Vernadsky Institute of Geochemistry and Analytical Chemistry of the Russian Academy of Sciences and Aalborg University (Denmark). The WSC-6 symposium was supported by the Russian Foundation for Basic Research (grant no. 08-03-06002).

The symposium was preceded by a two-day school on chemometrics held in Russian for students, graduate students, and young scientists. More than 60 listeners became acquainted with the basic concepts and approaches used to process the results of experiments and also obtained primary practical skills in the analysis of experimental data. The most active participants, representatives of the Dubna University, were awarded with a prize, the Unscrambler software [2] and a monograph [3].

The symposium took place on February 18–22 in the Glubokoe Ozero tourist camp near Kazan. The general approach to the analysis and interpretation of big data files was considered in the lecture by Professor K. Esbensen (Denmark) under the name “Chemometric data analysis - a holistic process with many elements.” The approach started with an experimental setup and design, the application of the sampling theory, data accumulation and the selection of variables and finished with the construction of a mathematical model and tests of its quality and predictive ability. Nowadays scientists are developing methods that allow the processing of complex hybrid experiments, as reported by A. Jankevich (Latvia) in the lecture “Metabolomic studies of experimental and clinical diabetic urine samples by  $^1\text{H}$  NMR and LC/MS spectroscopy,” or are building models for simultaneous processing of diverse experiments, as reported by F. Marini (Italy) in the lecture “Coupling of spectroscopic, thermoanalytic and chemometric methods for the resolution of enantiomeric mixtures.” One of the possible approaches used

in the cases when one and the same sample/phenomenon was studied in various experiments, was considered in the lectures by D. Routledge (France) “Outer Product Analysis (OPA) as a method to study the relations between sets of variables measured on the same individuals,” in which the efficiency of theoretical methods was checked by examples of the processing of experimental data of IR or near IR spectroscopy together with the data of X-ray diffraction or nuclear magnetic resonance. The advantage of the OPA approach is the possibility of processing expanded data sets by classical projective methods and also accounting for their structural multimodality, for example, by the PARAFAC method. The sophistication of computational procedures in working with multimodal data was the subject of D. Savost'yanov's (Russia) lecture on “Fast computation of CANDECOMP-PARAFAC and Tucker decompositions.”

Purely classical analytical applications were considered in the works on the use of multidimensional calibration in atomic emission spectrometry (S. Pavlov and E. Shabanova, Russia), uncertainty estimation in experiments on solubility determinations (M. Paakkunainen, Finland), on the use of PLS-regression for determining the composition of complex mixture using spectrophotometry (I. Vlasova, Russia), and on the revelation of falsified drugs by near IR spectrometry (O. Rodionova, Russia). Several works were devoted to the application of sensors—both chemical (S. Morzhukhina, Russia) “The data processing method of multisensor system for natural water monitoring” and biological (A. Reshetilov, Russia) “Elements of pattern recognition theory in the analysis of biosensor signals”). It was shown that the use of the multivariate approach can significantly improve the efficiency of such data processing.

Modern lines in chemometrics were presented in two important directions. The first is associated with the use of chemometric methods for the analysis of images and hyperspectral data. This direction at the symposium was presented in S. Kucheryavskii's (Denmark) lecture on “Image analysis in chemometrics,” and in I. Belyaev's lecture on applied research entitled “Estimation of age in forensic medicine using multivariate data analysis.” The second direction is the analytical control of industrial processes (Process Analytical Technology, PAT). This direction most vividly demonstrates the trends and prospects of the development

of the general approach uniting real-time physical and chemical experiments and mathematical methods of multivariate data analysis. Theoretical aspects were considered in the lecture by A. de Juan (Spain) on "Advanced process modeling with Multivariate Curve Resolution." Both the hardware and software required for solving PAT problems were considered in the lecture by A. Bogomolov (Germany) "PAT-enabling hardware and software." Particular solutions of applied problems in this field were proposed in the lectures on the control of biogas formation processes (M. Madsen, Denmark) and on the control of the granulation process in the pharmaceutical industry (S. Matero, Finland).

In the traditional mathematical modeling direction, I would like to note the works on the description and study of molecular structures by their spectra. The lecture by M. Elyashberg (Russia) on "Chemometrical methods in expert systems for the elucidation of the molecular structure" described the general concept of the construction of expert systems for revealing molecular structures by their spectra and the systems of axioms from which structural formulas are deduced. The author presented the Structure Elucidator system capable of identifying large organic molecules (more than 100 skeleton atoms) by 2D NMR spectra. The presentation by K. Blinov (Russia) on "Prediction of NMR Chemical Shifts, A Chemometrical Approach," which was devoted to the elaboration of a method for the prediction of chemical shifts, continued this topic. The prediction of molecular structures by IR spectra as applied to a new class of polymeric molecules, dendrimers, was the subject of the lecture by S. Furer (Russia) entitled "Molecular structure and IR spectra of phosphorus dendrimers by DFT calculations." The lecture by V. Dement'ev (Russia) on "Computer methods for evaluation of the confidence intervals of molecular modeling prediction" was devoted to the estimation of prediction errors in standardless molecular analysis and general philosophical questions of nature and the laws of error distribution.

The mathematical apparatus of chemometrics is being developed following the current requirements of growing increasingly complicated experimental bases and more and more diverse applications. The new methods for the assessment of the determinate zone used in classification problems, PAT, and estimations of calibration overshoots were considered in the work by A. Pomerantsev (Russia) on "Critical levels in projection techniques." New approaches to the construction of

models of complex multistep processes were presented by A. Höskuldsson (Denmark) in his lecture entitled "Multi-block and Path Methods—The next generation of chemometric methods."

The competition of young scientists was won by Kirill Blinov, who won an annual subscription to the journal *Chemometrics and Intelligent Laboratory Systems*. The authors presenting posters participated in a separate competition; its winner was N. Ryumina (Samara, Russia) with his work on "IR spectroscopy and PCA applied for polymer sorting in domestic waste."

The symposium was attended by more than 100 researchers from 13 countries: Russia, Germany, Denmark, Italy, Spain, Poland, Romania, the United States, Turkey, France, Finland, and Southern Africa. There were many young people, students, and graduate students. The wide range of topics discussed at the symposium, and the variety of participants' specialties (chemists, physicists, mathematicians, engineers) are indicative of the vivid interest in chemometrics and the necessity of the application of its methods. Participants noted the usefulness of such conferences, which allow them to discuss interdisciplinary approaches and methods developed at the intersections of various scientific disciplines.

Abstracts of papers were published in a separate edition [4]; the lectures and oral presentations can be found on the conference web site [5]. The materials of the symposium will be published in a special issue of the journal *Chemom. Intell. Lab. Syst.* Edited by A. Pomerantsev and R. Tauler, as was the case in the issue devoted to the previous symposium WSC-5 [6].

## REFERENCES

1. <http://rcs.chph.ras.ru/drushba.htm>
2. The Unscrambler, <http://www.camo.com/rt/Products/Unscrambler/unscrambler.html>
3. *Esbensen K.H.* Multivariate Data Analysis – In Practice 4-th Ed., CAMO, 2000.
4. VIth International Symposium "Modern Multivariate Data analysis", February 18–22, 2008, Kazan, KGASU, 2008, 68 p.
5. <http://chemometrics.ru/drushbametrics/wsc6/>
6. *Chemom. Intell. Lab. Syst.* 88. № 1. 2007.

A.L. Pomerantsev