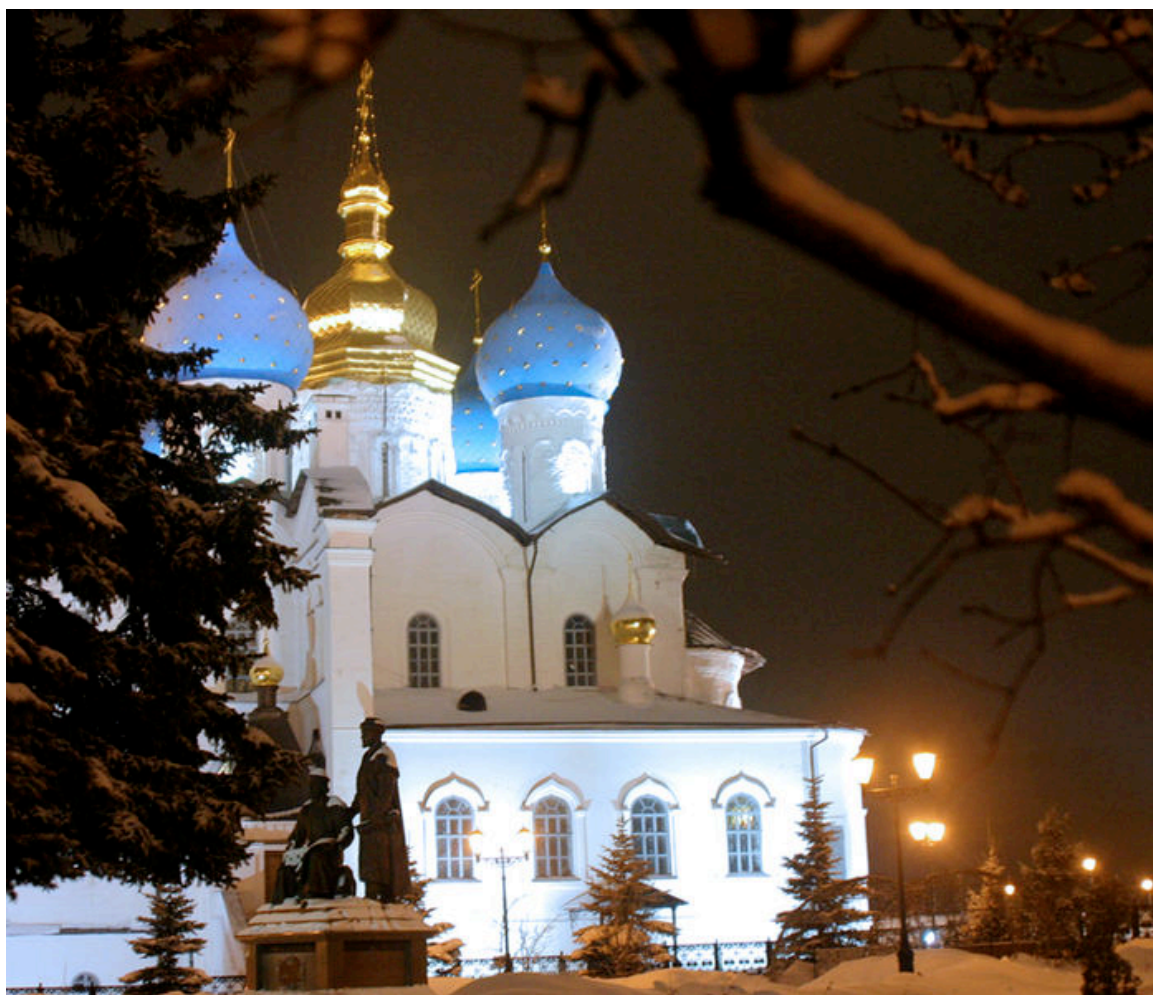


Drushbameetrics Project

Sixth Winter Symposium on Chemometrics

Modern Methods of Data Analysis



Russia, Kazan, February 18–22, 2008

Казанский государственный архитектурно-строительный университет
Научный Совет по аналитической химии РАН
Академия наук республики Татарстан
Российское хемометрическое общество
Университет Ольборг (ACABS group)
Мэрия города Казань

при поддержке Российского фонда фундаментальных исследований



Шестой международный симпозиум

Современные методы анализа многомерных данных

Организационный комитет

Сопредседатели

член-корр. РАН Грибов Л.А. (ГЕОХИ РАН, Россия)

д.т.н. Сулейманов А.М. (КГАСУ, Россия)

проф. Эсбенсен К. (ACABS, Дания)

Ученый секретарь

Старовойтова И.А. (КГАСУ, Россия)

Члены оргкомитета

проф. Лыгина Т.З. (ЦНИИ Геолнеруд, Россия)

д.ф.-м.н. Померанцев А.Л. (ИХФ РАН, Россия)

проф. Рейникайнен С.-П. (LUT, Финляндия)

к.ф.-м.н. Родионова О.Е. (РХО, Россия)

проф. Таулер Р. (CSIC, Испания)

к.ф.-м.н. Кучерявский С.В. (ACABS, Дания)

Адрес Организационного комитета

КГАСУ, ул. Зеленая 1, Казань, 420043

<http://www.chemometrics.ru/wsc6/>

e-mail: wsc6@chemometrics.ru

Thanks

The WSC-6 organizers and participants wish express greatest appreciation to the following conference sponsors for their valuable economic and friendly help:

Russian Foundation for Basic Research

Russian Academy of Sciences

Academy of Sciences of the Republic of Tatarstan

Kazan State University of Architecture and Engineering

ACABS Research Group

Institute of Geology of Industrial Minerals

Perkin Elmer Inc.

Elsevier B.V.

J&M Analytische Mess- und Regeltechnik GmbH

Bruker Optics Inc.

CAMO ASA

The Kazan mayoralty

Finally, we are grateful to all the WSC-6 attendees, lecturers, accompanying persons, and visitors for their interest to the conference.

See you again at the next WSC-7 conference!



Sponsors



**RUSSIAN
FOUNDATION
FOR BASIC
RESEARCH**



**Russian
Academy
of Sciences**

**Academy of Sciences
of the Republic of Tatarstan**



**Kazan State University
of Architecture and
Engineering**



ACABS
Research Group



*Building Insights
Breaking Boundaries*



Information

Location and accommodation

The attendees will live in the comfortable single and double rooms. All campus bungalows have halls for evening discussions. Russian banya (sauna) and cross-country ski will be available.

Conference sessions will be held at the second floor of the main building.

Meals

All meals as well as the banquet will be served at the dining room at the ground floor of the main building.

Scores & Loadings

Traditional "Scores and Loadings" meeting will be organized in the old-fashioned, self service Kostroma style.

Communication

The main two Russian cellular networks, Beeline and MTS, have a proper coverage in the campus and around. **The Internet can be accessed from administrative building during the coffee-breaks.**

Manuscript submission instructions

All participants are invited to submit a full-length manuscript for inclusion in a special issue, devoted to the 6th Winter Symposium on Chemometrics (WSC-6), of the journal **Chemometrics and Intelligent Laboratory Systems**. All manuscripts will be subject to the usual refereeing procedure.

To submit:

- Go to: <http://ees.elsevier.com/chemolab>
- Click on the "Submit Paper" option from the top menu
- Enter your user name and password (first time users will have to register)
- After entering the titles of your MS, please select **WSC-6** as the 'Article Type'
- Select **R. Tauler** as Editor
- Follow the remaining step-by-step instructions to submitting your paper

Points to note:

- The deadline to submit your MS to the WSC-6 special issue is **31 March 2008**
- When preparing your MS please follow the journal's Guide for Authors
- In your cover letter, please mention that your MS is intended for WSC-6 special issue

If you experience any technical problems during the submission process, please contact Elsevier's Author Support: authorsupport@elsevier.com.

If you have any questions on the submission and reviewing process, please contact the Guest Editor Alexey Pomerantsev: forecast@chph.ras.ru or alexey.pomerantsev@gmail.com

Young chemometrician award

Elsevier Company instituted award for the best oral or poster presentation of a young scientist at WSC-6. This award consists of a free paper subscription to Chemometrics and Intelligent Laboratory Systems for the duration of one year.

Best student award

CAMO Company instituted award to the best student at WSC-6 School (pre-courses). This award consists of a free one year license of the Unscrambler software and the book Multivariate Data Analysis - In Practice, written by Kim H. Esbensen

Best poster presentation award

CAMO Company instituted award for the best poster presentation at WSC-6. This award has three degrees.

The first place prize consists of a free one year license of the Unscrambler software and the book Multivariate Data Analysis - In Practice, written by Kim H. Esbensen.

The second place prize consists of a free registration at the CAMO Training Courses and the book Multivariate Data Analysis - In Practice, written by Kim H. Esbensen.

The third place prize consists of the book Multivariate Data Analysis - In Practice, written by Kim H. Esbensen.

Money

You may exchange EURs and US dollars to Russian rubles (RUR) in Kazan banks, at the railway station, or in the airport. The organizing committee could also change the reasonable amount of currency.

Excursion

On Friday, February 22 the participants will have a chance to visit one of the most attractive places in Kazan – Kazan Kremlin.

Accompanying persons program

During the Symposium several excursions will be organized for the accompanying persons.

- The National Museum of the Republic of Tatarstan
- Konstantin Vasiliev Memorial Museum
- The Raifa Holy Virgin Monastery
- The City-Island Sviyazhsk

Miscellaneous

The conference official language is English.

Everyone is encouraged to have his/her badge attached, both during the symposium sessions and social activities.

Useful Phone Numbers

Alfred Sulejmanov, *symposium chair* +7 (960) 032-66-58 (cellular)

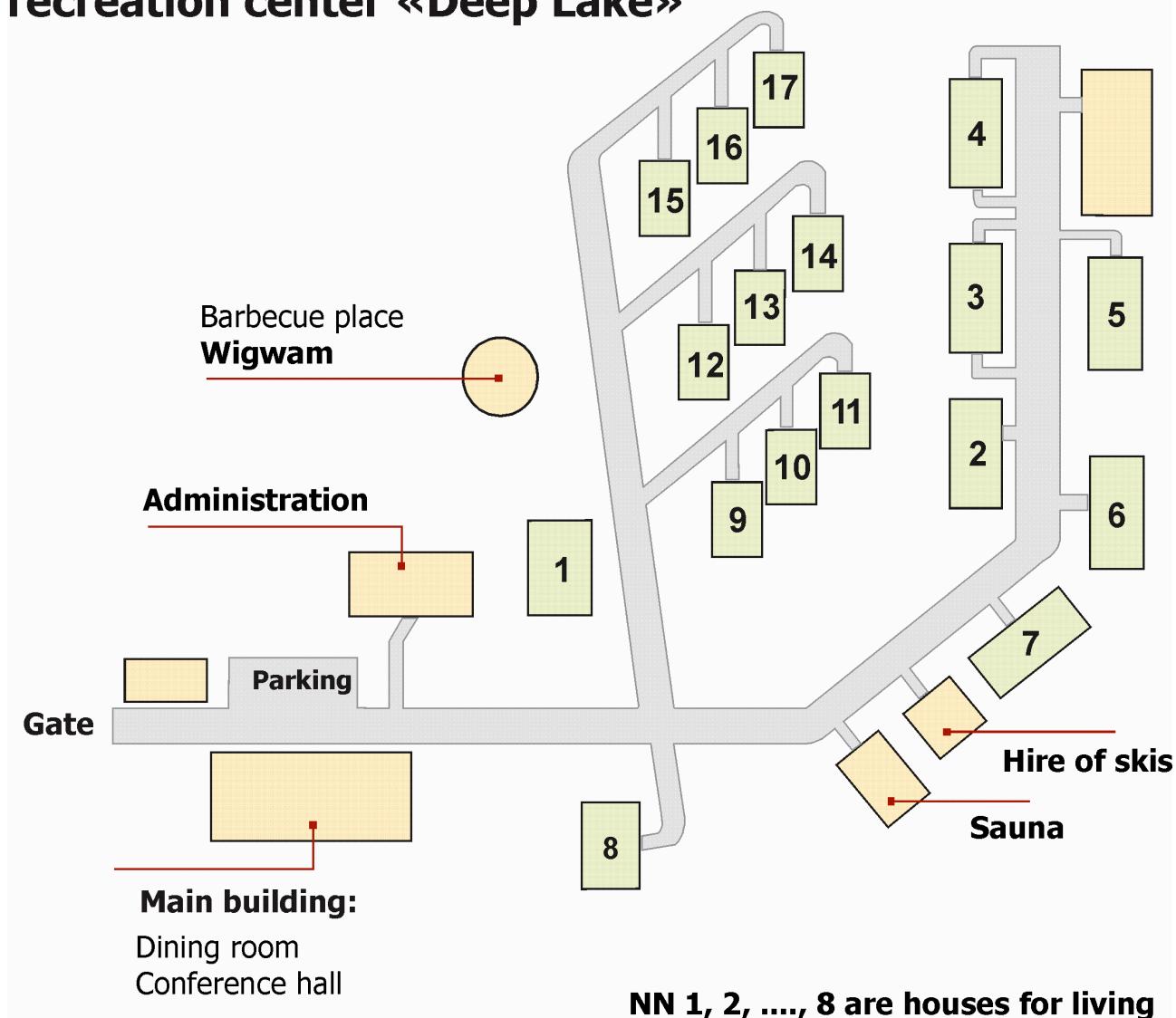
Irina Starovoitova, *symposium secretary* +7 (917) 392-93-02 (cellular)

Oxana Rodionova, *org. committee* +7 (903) 150-48-55 (cellular)

Alexey Pomerantsev, *org. committee* +7 (903) 561-82-27 (cellular)

“Deep Lake” campus +7 (843) 557-87-47 (local)

Scheme of the recreation center «Deep Lake»



Timetable

Monday, February 18, 2007

09:00–12:30	Registration
13:00–14:00	Lunch
14:00–14:30	Opening
Session 1	Chair: Alexey Pomerantsev
14:30–15:30	L1 <i>Kim Esbensen</i> Chemometric data analysis — a holistic process with many elements
15:30–16:00	T1 <i>Alfred Sulejmanov</i> Chemometric in Civil Engineering
16:00–16:30	Coffee break
Session 2	Chair: Kim Esbensen
16:30–17:00	T2 <i>Alexey Pomerantsev</i> Critical levels in projection techniques
17:00–17:30	T3 <i>Oxana Rodionova</i> Recognition of the 'high quality forgeries' among the medicines
17:30–18:00	T4 <i>Federico Marni</i> Coupling of spectroscopic, thermoanalytic and chemometric methods for the resolution of enantiomeric mixtures
18:00–18:30	T5 <i>Alisa Rudnitskaya</i> Instrumental measurements of beer taste attributes using an electronic tongue
18:30–19:00	Free time
19:00–20:00	Dinner
20:00–00:00	Scores & Loadings

Tuesday, February 19, 2007

08:00–09:30	Breakfast
Session 3	Chair: Alfred Sulejmanov
09:30–10:30	L2 <i>Anna de Juan</i> Advanced process modeling with Multivariate Curve Resolution
10:30–11:00	T6 <i>Leon Rusinov</i> Diagnostics of complex chemical processes on the base of PCA and fuzzy neural networks
11:00–11:30	Coffee break
11:00–13:00	Free time (skiing)
13:00–14:00	Lunch
Session 4	Chair: Victor Furer
14:00–15:00	L3 <i>Sergey Kucheryavskiy</i> Image Analysis in Chemometrics
15:00–15:30	T7 <i>Ivan Belyaev</i> Estimation of age in forensic medicine using multivariate data analysis
15:30–16:00	T8 <i>Andris Jankevics</i> Metabolomic studies of experimental and clinical diabetic urine samples by 1H NMR and LC/MS spectroscopy
16:00–16:30	Coffee break
Session 5	Chair: Anna de Juan
16:30–17:00	T9 <i>Dmitriy Savostyanov</i> Fast computation of CANDECOMP-PARAFAC and Tucker decompositions
17:00–17:30	T10 <i>Svetlana Sokolova</i> Intelligent module of stock market simulator
17:30–18:00	T11 <i>Vasily Dementiev</i> Computer methods for evaluation of the confidence intervals of molecular modeling prediction
18:00–18:30	T12 <i>Anatoly Reshetilov</i> Elements of pattern recognition theory in the analysis of biosensor signals
18:30–19:00	Free time
19:00–20:00	Dinner
20:00–00:00	Scores & Loadings

Wednesday, February 20, 2007

08:00–09:30	Breakfast
Session 6	Chair: Sergey Kucheryavskiy
09:30–10:30	L4 <i>Douglas Rutledge</i> Outer Product Analysis (OPA) as a method to study the relations between sets of variables measured on the same individuals
10:30–11:00	T13 <i>Andrey Bogomolov</i> PAT-enabling hard- and software
11:00–11:30	Coffee break
11:00–13:00	Free time (skiing)
13:00–14:00	Lunch
Session 7	Chair: Michail Elyashberg
14:00–15:00	L5 <i>Victor Furer</i> Molecular structure and IR spectra of phosphorus dendrimers by DFT calculations
15:00–15:30	T14 <i>Alexey Potapov</i> The study of dispersed systems using NIR spectroscopy and PLS regression technique
15:30–16:00	T15 <i>Yaroslav Sokovikov</i> The discriminating pharmaceutical raw materials on warehouse by FT-NIR spectrometer Spectrum 100 (PerkinElmer) with Fiber Optic Accessory
16:00–16:30	Coffee break
Session 8	Chair: Anatoly Reshetilov
16:30–18:30	Poster Session
18:30–19:00	Free time
19:00–20:00	Dinner
20:00–00:00	Scores & Loadings

Thursday, February 21, 2007

08:00–09:30	Breakfast
Session 9	Chair: Agnar Höskuldsson
09:30–10:30	L6 <i>Michail Elyashberg</i> Chemometrical methods in expert systems for the molecular structure elucidation
10:30–11:00	T16 <i>Kirill Blinov</i> Prediction of NMR Chemical Shifts. A Chemometrical Approach
11:00–11:30	Coffee break
11:00–13:00	Free time (skiing)
13:00–14:00	Lunch
Session 10	Chair: Douglas Rutledge
14:00–15:00	L7 <i>Agnar Höskuldsson</i> Multi-block and Path Methods The next generation of chemometric methods
15:00–15:30	T17 <i>Sergey Pavlov</i> Mathematical methods and algorithms used in the integrated program ARDES for automatic processing of atomic-emission spectra
15:30–16:00	T18 <i>Costel Sarbu</i> Classical and fuzzy principal component analysis of environmental samples concerning the pollution with heavy metals
16:00–16:30	Coffee break
Session 11	Chair: Oxana Rodionova
16:30–17:00	T19 <i>Yuri Kalambet</i> POPLC technology: best separation in one step. Present state and perspective
17:00–17:30	T20 <i>Sylvia Paul</i> Chemometrics in South Africa: applications, challenges, and the development of the SA chemometrics society
17:30–18:00	Discussion
18:00–18:30	Closing
18:30–20:00	Free time
20:00–00:00	Banquet

Friday, February 22, 2007

08:00–09:30

Breakfast

09:30–14:00

Excursion



Abstracts

Lectures

L01. Chemometric data analysis — a holistic process with many elements

K. Esbensen

Aalborg University Esbjerg, Esbjerg, Denmark

Chemometric data analysis has a natural focus on the statistical/data analytical methods applied, or developed. While this may seem understandable, there is not always a complementary emphasis on the data analysis objective, nor of data quality. A holistic understanding of chemometric data analysis must (at least) include the following elements:

1. Is data generation (observation) active (controllable) or passive (WYSIWYG)?
2. If active, is use made of an appropriate experimental design?
3. If passive, is appropriate consideration made to representativity, Theory of Sampling (TOS)?
4. Are the variables appropriate for the data analysis objective/method — e.g. what constitute proper criteria for problem-specific variable selection?
5. Does the realized matrix (objects/variables) support the data analysis method?
6. Is the realized data quality commensurate with the specific data analysis method requirements?

It is advantageous to conceptualize of chemometric data analysis as a process starting with a thorough specification of the data analysis objective in the relevant problem context (from which follows an optimal choice of method), followed by experimental design and/or representative sampling before chemical, physical (other) analysis generates the data. Proper consideration of all such pre-data analysis issues is necessary to increase the potential of meaningful/useful data analysis results. Indeed, most often the only guarantee for interpretable results lies already in the pre-data stages. Finally: in holistic data analysis the crucial role of proper validation reigns supreme, without which no scientific merit at all.

L02. Advanced process modeling with Multivariate Curve Resolution

¹*A. de Juan*, ²*R. Tauler*

¹*Universitat de Barcelona, Barcelona, Spain*

²*Institute of Chemistry and Environmental Research, CSIC, Barcelona, Spain*

Multivariate curve resolution is a powerful tool to model complex processes. The underlying bilinear model of curve resolution, which expresses the raw data collected during process monitoring as the composition-weighted sum of pure component signal contributions, adapts to the description of processes collected with a variety of spectroscopic techniques and with partially or completely unknown mechanisms.

Curve resolution can easily handle multitechnique and/or multiexperiment data structures. The analysis of these data sets provides a very robust structural information to characterize the compounds involved in a process (multitechnique monitoring) or can allow for the simultaneous study of a set of designed experiments that can allow for the detection of minor or intermediate compounds or for the assessment of the effect of different inducing agents or process control variables in the process mechanism. Multiset arrangements are also the key strategy to tackle problems of rank-deficiency, i.e., resolution of compounds with identical spectral shapes in a particular technique or identical concentration profile shapes in certain experimental conditions.

Introducing external information related to properties of the signal recorded or to the shape of the process profiles, when available, is a capability of CR methods. Whereas soft-modeling constraints (non-negativity, closure, ...) are applied since long in process modeling, process profiles can also be fitted according to a hard model, if the mechanism of the process is partially or totally known. Unlike pure hard-modeling approaches, where the total mechanism of the process should be known and a global model is needed in multiexperiment analysis, hard-modeling constraints can fit only some of the compounds contributing to the signal recorded during a process and can handle sets of experiments with a non-common model or structures with model-free and model-based experiments. All these possibilities will be shown with real examples.

L03. Image analysis in chemometrics

S. Kucheryavskiy

Aalborg University Esbjerg, Denmark

Digital image processing is a powerful research tool that is employed for discovering both the properties and state of different objects by analyzing of their structure and appearance. In spite of the fact that basic methods and algorithms have been developed in 70-80th of previous century, the area is still open for the researchers. Moreover, even the meaning of 'digital image' term has undergone a very thorough revision since that time, as now the term includes multispectral, hyperspectral, etc images. Such area expansion stimulates researchers to new findings in processing, analysis and classification of images in science and engineering.

The presentation gives an overview of image processing and analysis methods, used in chemometrics. These methods are divided on several parts, depending on the type of investigated image. Thus, for the ordinary images, in which every pixel is represented by one (intensity) or three (basic colors) values, it is very important first to extract useful information for a posterior analysis. In this case, the transform based methods: wavelet analysis (that transforms image from a spatial to the frequency-spatial domain) and Angle Measure Technique (that transforms image to the scale domain) are efficient. Several examples of using these methods for features extraction will be shown.

In the multi- and hypespectral images each pixel contains enough information (from tens up to hundreds corresponding values) to become an independent object of analysis. A huge amount of data (e.g., the 512x512 pixel image with 100 channels gives a data matrix with 262144 objects and 100 variables) does not allow applying the traditional methods for analysis. In such case it is very helpful.to use a score space instead of a variable space. Two methods that implement this idea, namely Multivariate Image Analysis (MIA) and Multivariate Image Regression (MIR) are overviewed.

L04. Outer Product Analysis (OPA) as a method to study the relations between sets of variables measured on the same individuals

D. Rutledge

Laboratoire de Chimie Analytique, AgroParisTech, Paris, France

Let \mathbf{X} and \mathbf{Y} be two groups of variables containing p and q variables respectively, measured on the same N individuals.

The analysis of the relations between the two tables \mathbf{X} and \mathbf{Y} is often carried out by Tucker analysis [1]. This analysis is based on the singular values decomposition of the matrix of covariances between the two groups of variables ($1/n \mathbf{X}^T \mathbf{Y}$). This decomposition of the covariance matrix corresponds to the successive search of couples of variables ($\mathbf{T}_H = \mathbf{X} \mathbf{A}_H \mathbf{U}_H = \mathbf{Y} \mathbf{B}_H$) such that the covariance between \mathbf{T}_H et \mathbf{U}_H is maximum, under the constraints that the axes \mathbf{A}_H are orthogonal as are the axes \mathbf{B}_H .

Recently, Barros *et al.* [2] proposed an alternative procedure based on the analysis of a three-way table generated by the outer product between two tables \mathbf{X} and \mathbf{Y} . The three entries of this cube are respectively the individuals, the variables of \mathbf{X} and the variables of \mathbf{Y} . The analysis this three-way table can be carried out by unfolding followed by standard multivariate methods such as *PCA*, *ICA*, *PLS* etc., or directly by multiway methods such as *PARAFAC*.

Compared to the often-used Tucker analysis for the study of the relations between two tables \mathbf{X} and \mathbf{Y} , the first interesting aspect of OPA, is that the Tucker method is in fact *a compromise analysis (uniform average) of the cube of outer products between X and Y*. The matrix of covariance between the two groups of variables \mathbf{X} and \mathbf{Y} is equal to the average of the cube of outer products along the direction of the individuals I (averages by columns).

In an explicit way:

$$\frac{1}{n} \sum_{i=1}^n \mathbf{Z}_i = \frac{1}{n} \mathbf{X}^T \mathbf{Y}$$

The second interesting aspect is that with OPA it is possible to study relations among several tables, \mathbf{W} , \mathbf{X} and \mathbf{Y} , by calculating multiple Outer Products to generate a hyper-cube.

The most interesting property of OPA is that it is possible to not only unfold the outer products and apply classical multivariate methods, but it is also possible to

retain the cubic structure and apply multiway methods such as *PARAFAC* which are already well-known for the analysis of cubic data.

In this presentation, Outer Product Analysis will be compared to other techniques such as Tucker [1], Generalised 2D-correlation spectroscopy [3], and PLS2 regression [4] using examples taken from many fields including Time-Domain NMR, Mid- and Near Infrared spectroscopy, X-ray diffraction and NMR spectroscopy.

References

1. LR Tucker, 1958, *Psychometrika*, **23**, 111 -136.
7. A.S. Barros, M. Safar, M.F. Devaux, P. Robert, D. Bertrand, D.N. Rutledge, 1997, *Applied Spectroscopy*, **51** (9) 1384-1393
8. I Noda, A.E. Dowrey, C. Marcott, G.M. Story, Y. Osaki, 2000, *Applied Spectroscopy*, **54**, 236A-248A.
9. R. Lew and S. T. Balke, 1993, *Applied Spectroscopy*, **47**, 1747-1750

L05. Molecular structure and IR spectra of phosphorus dendrimers by DFT calculations

V. Furer

Kazan State University of Architecture and Engineering, Kazan, Russia

Keywords: phosphorus-containing dendrimers, IR spectra, normal vibrations

The main features of vibrational spectra of starburst dendrimers have been analysed for the first time. Their spectral pattern, in general, is determined by the ratio of a number of terminal groups to a number of repeating units. This ratio trends to $mr - 1$ (mr – branching functionality of repeating unit), and become the constant, when generation number of the starburst dendrimer increases higher than 3-5. IR and Raman spectra of the twelve generations of the phosphorus-containing dendrimers are represented and interpreted on the basis of the calculation of frequencies of the normal vibrations and band intensities in the IR spectra of terminated by dangling methyl groups "molecules", which are the fragments of dendrimer molecule. Tailored spectra of these fragments then compared with experimental spectra and satisfactory similarity have been obtained. Experimental spectra of generations higher than 4 are very close similar, according theoretical approach. The results can be used for the analysis of the chemical and physical transformations in starburst dendrimers.

L06. Chemometrical methods in expert systems for the molecular structure elucidation

M. Elyashberg

Advanced Chemistry Development (ACD) Inc., Moscow Russia

Keywords: expert system, molecular structure elucidation, 1D and 2D NMR spectra, structure generation, spectrum prediction

Conventionally, two following branches may be distinguished in chemometrics: quantitative and qualitative ones. Quantitative chemometrics is based on continuous mathematics, where mathematical statistics plays the leading role. The major areas in this field include multivariate calibration, pattern recognition, mathematical mixture resolution, etc. The main attention is focused on quantitative aspects of chemical analysis. Qualitative chemometrics is based on discrete mathematics, mainly on mathematical logic, graph theory and combinatorial analysis. These methods allow realizing the basic stages of an analytical procedure which is commonly called Computer-Aided Structure Elucidation (CASE).

This procedure is successfully applied to the structure elucidation of unknown (newly separated or synthesized) organic compounds from their MS, NMR and optical spectra, that is for qualitative analysis. To automate the qualitative analysis, CASE expert systems were developed. Methodology of CASE expert systems includes the following main stages: 1) spectral parameters determination in molecular spectra of different nature; 2) selection of possible molecular fragments on the basis of spectral data; 3) generation of all molecular structures which met spectral data and additional constraints, the found fragments and molecular formula being used for this goal; 4) selection of the most probable structure by means of spectrum prediction for all candidate structures included into the output file; 5) determination of 3D model and relative stereochemistry of the preferable structure.

In our presentation, we consider how the methods intrinsic for both branches of chemometrics are combined within CASE expert systems. Important role of quantitative chemometrics and its impact on the entire CASE strategy is discussed and illustrated by series of examples taken from analytical practice. Expert system ACD/Structure Elucidator is used as an example of the most advanced CASE expert system.

L07. Multi-block and Path Methods – The next generation of chemometric methods

A. Höskuldsson

Centre for Advanced Data Analysis, Kgs Lyngby, Denmark

When data are collected for modelling tasks there is usually available information that should be included in the modelling task. E.g., data from a process plant is basically of multi-block type. Each sub-process generates separate data. Similarly, organisational diagrams may suggest sub-divisions of data. Here are presented methods where data are organised in a directional network of data blocks. The data blocks are arranged in directional paths. Each data block can lead to one or more data blocks. It is assumed that there are given a collection of input data blocks. Each of them is supposed to describe one or more intermediate data blocks. The output data blocks are those that are at the end of the paths and have no succeeding data blocks. The optimisation procedure finds weights for the input data blocks such that the size of the total loadings for the output data blocks is maximised. When the optimal weight vectors have been determined, the score and loading vectors for the data blocks in the network are determined. Appropriate adjustment of the data blocks is carried out at each step. Regression coefficients are computed for each data block that show how the data block is estimated by data blocks that lead to it. Methods of standard regression analysis are extended to present methods. Three types of 'strengths' of relationship are computed for each set of two connected data blocks. First is the strength in the path, second the strength where only the data blocks leading to the last one is use and third if only the two are considered. Cross-validation and other standard methods of linear and non-linear regression are carried out in a similar manner.

In industry processes are organised in different ways. It can be useful to model the processes in the way they are carried out. By proper alignment of sub-processes, overall model can be specified. There can be several useful path models during the process, where the data blocks in a path are the ones that are actual or important at given stages of the process. An important aspect of these methods is that we get score vectors for each data block. Thus we can study the role of each data block in the network.

Talks

T01. Chemometric in Civil Engineering

A. Sulejmanov

Kazan State University of Architecture and Engineering, Kazan, Russia

Keywords: optimization of construction composite material, degradation, PLS models

Multivariate approach (MVA) for data processing has shown its great efficiency in various applications both in chemistry and biology. The complexity of objects and problems in civil engineering also requires modern MVA for data analysis. An engineer is increasingly faced with need to use mathematical and statistical methods in everyday practice. MVA in general and chemometric methods in particular may essentially enhance the capability of newer, mostly multicomponent, methods. Applications of various chemometric techniques should give possibility for efficient investigations in different areas such as

- monitoring of historical buildings with non-intrusive measurements (image analysis, acoustic and ultrasonic probing);
- estimation of heterogeneity in concrete slabs, identification of the substrate defects on painted slates, etc. (novel methods of texture analysis);
- ecological analysis of pollution; (PCA, PLS and N-way methods);
- analysis of aging and degradation of various materials (multivariate projection methods);
- quantitative structure—property relationship (PCA);

First attempts of chemometric applications in the KSUAE are presented in this conference. They are: optimization of construction composite material by nonlinear PCR and study of aging and degradation of PVC films with the help of soft models. The latter problem is presented in details.

The samples of PVC soft jackets of various thickness, colour, and light absorbers were exposed for the accelerated aging. PCA helps to reveal main properties of initial material that influence on degradation and correlation between the mechanical and colour characteristics in the course of aging. PLS models show the possibility to substitute a time consuming mechanical testing by quick and cheap colour analysis.

T02. Critical levels in projection techniques

A. Pomerantsev

Institute of Chemical Physics RAS, Moscow, Russia

Keywords: PCA, SIMCA, leverage distribution, residual variance distribution, type I error

The critical levels in projection are used intensively in chemometrics; particularly, in the following areas. The first one is SIMCA method; a popular chemometric tool for the supervised pattern recognition. The second area is the multivariate statistical process control (MSPC) that employs critical levels for tracing the behavior of a process. The third application is the outlier detection in the multivariate calibration with the influence plot as a specific tool.

After projecting initial data onto a score subspace, each sample can be presented as the sum of the vector that lies in the subspace (a projection) and transversal vector (a residual). The lengths of these vectors characterize the sample position with respect to the model (subspace). They are termed as the score distance (SD, aka leverage) and the orthogonal distance (OD, aka residual variance), correspondingly. The SDs and ODs obtained for the known class members constitute two samplings, which represent the population. By exploring these samplings the critical membership levels can be established. Further, when a new candidate object x is considered it can be projected onto the model subspace, and its own SD and OD values can be compared with the critical levels to make a decision on the membership of the class.

In such a context, several statistical problems are of vital importance. The first is the form of the SD and OD distributions. Moreover, in each specific case, the distribution parameters are to be evaluated using the training data set. It is further important to set up the rules that reveal the extremes and outliers in the data. At last, the acceptance area in the SD-OD plot should be defined for a given type I error. Such issues have already been discussed in the numerous publications but they are still topical.

This research aims at an open theoretical discussion on these topics. Several real-world examples will be shown at presentation T03, which is closely connected with this one

T03. Recognition of the 'high quality forgeries' among the medicines

O. Rodionova

Institute of Chemical Physics RAS, Moscow, Russia

Keywords: NIR spectroscopy, SIMCA, control limits

Counterfeit drugs could be of different type, such as placebo, the medicines with lower concentration of active substances, the drugs that do not contain the proper concentrations or contain a wrong type of excipients, etc. The most difficult for revealing are 'the high quality fakes', which have a proper composition but produced by the underground manufactures with violation of technological regulations. We consider that methods that are based only on determination of an active substance do not suffice for our purposes due to a wide range of the counterfeit products. As NIR measurements carry information regarding not only chemical but also physical phenomena, NIR spectroscopy was chosen as an instrument. A well-known method of soft independent modeling of class analogy (SIMCA) with modified control limits was applied for the mathematical data processing.

Investigation of fifteen different types of medicines has shown that application of NIR measurements together with chemometric data processing is an effective approach. Proper discrimination should be free from human failures. For these purposes:

- Each medical product should be carefully investigated on the batch-to-batch variability.
- Selection of spectral region should be done individually for each type of medicine. The spectral region may essentially influence on the final classification results.
- It is equal important to recognize the forgeries as well as to avoid a misclassification for the genuine samples. Application of reliable acceptance limits is of great importance.

Several real-world examples illustrate the presentation.

T04. Coupling of spectroscopic, thermoanalytic and chemometric methods for the resolution of enantiomeric mixtures

F. Marini, R. Bucci, I. Ginevro, A. L. Magri

University of Rome, Rome, Italy

Keywords: IR spectroscopy, DSC, enantiomeric excess, chemometrics

Enantiomers are isomers that are the mirror image of one another and that, except for the property of rotating the plane-polarized light in opposite direction, share common physical and chemical properties. However, quite often it can happen that their pharmacological properties are significantly different so that one enantiomer is active while the other one doesn't produce any effect or, in some cases, can result even noxious. It is evident, therefore, that the possibility of determining the quantity of individual enantiomers in mixtures is of capital importance, especially in the quality control of drugs.

In this communication, we will describe the successful combination of IR spectroscopy, DSC and chemometrics to determine the enantiomeric excess in pharmaceutical preparations by means of an example of application to a pharmaceutical active principle (ketoprofen).

T05. Instrumental measurements of beer taste attributes using an electronic tongue

¹A. Rudnitskaya, ¹E. Polshin, ¹D. Kirsanov, ²K. Beullens, ²J. Lammertyn, ²B. Nicolai, ¹A. Legin

¹Laboratory of Chemical Sensors, Chemistry Department, St. Petersburg University, Russia

² Department of Agrotechnique and Economics, Catholic University of Leuven, Heverlee, Belgium

The human tongue remains one of the most effective instruments for identification and characterization of taste qualities of foods and sensory panels remain an integral part of the quality control in the food industry. However, sensory panels are notorious for being expensive and time-consuming to train and analysis they perform for being slow and expensive. Certain characteristics are difficult to taste because of their lingering nature and carryover from one sample to the next when tasting is done in quick succession. This causes that the rate and effectiveness of tasting panels are impaired. Due to these considerations significant efforts were directed into application of the fast instrumental techniques for the taste evaluation.

One of the promising approaches is multisensor systems or electronic tongues (ET). Several successful applications of the ETs to the taste assessment were reported including correlation between instrument output and perceived bitterness in various foodstuffs and pharmaceuticals as well as other flavour characteristics.

The aim of the present study was application of the ET based on potentiometric chemical sensors to the assessment of various aspects of taste and flavour of Belgian beer. Fifty sorts of beer were measured using an electronic tongue based on potentiometric chemical sensors and evaluated by the trained sensory panel.

Different aspects of the data processing were addressed. One of the issues is related to the nature of sensorial data namely to the difference between individual tasters in the sensory panel. Even trained tasters differ in sensitivity and estimates they produce are to some degree subjective. Therefore, calculating arithmetic mean does not seem an appropriate approach. An established technique for calculating a consensus mean of sensory assessors is general procrustes analysis. It allows obtaining sensor panel average estimates that are more relevant with respect to the samples than mean values. In the present study general procrustes analysis was used for the calculating mean of sensorial assessors of beer. A combination of general procrustes analysis with partial least square regression was used for relating sensory and instrumental i.e. electronic tongue data. Canonical correlation analysis was employed as well for correlating those two data sets.

T06. Diagnostics of complex chemical processes on the base of PCA and fuzzy neural networks

L. Rusinov, I. Rudakova, O. Remizova V. Kurkina

Saint-Petersburg Technological Institute, Saint-Petersburg, Russia

Keywords: diagnostics, chemical process faults, PCA, neural

PCA method is widely applied to detection of faults in the state of chemical processes. At the same time this method does not allow simply to identify the reasons of these faults especially when they were caused by the change of many variables at once. For this purpose the application of neural networks is perspective. However, faults are characterized by the illegible description and by fuzzy values of variables that determine these faults. In this report the use of fuzzy neural networks is discussed for solve the problem. Efficiency of the suggested method is shown on the example of the process of hydrocarbons pyrolysis.

T07. Estimation of age in forensic medicine using multivariate data analysis

¹*I. Belyaev*, ²*S. Kucheryavskiy*

¹*Altai State University, Barnaul, Russia*

²*Aalborg University Esbjerg, ACABS group, Esbjerg, Denmark*

Keywords: forensic medicine, age estimation, digital image analysis,

Estimation of age of victims' remains is one of the most important problems in forensic medicine. Usually this estimation is based on analysis of bones of skeleton, first of all because of their stability to putrefying transformation. Furthermore bones have different age-related information features, first of all caused by osteoporosis process. The existent methods use photolorimetric definition indexes of light transmission and absorption for sponge tissue. These methods do not give stable results and require additional routine operations and calculations.

In the present work, a new method of remains' age estimation is proposed. The method is based on multivariate analysis of digital images of bone sections. Images were acquired by digital camera in visible light. Two different approaches (wavelet analysis and Angle Measure Technique) were used for extraction of problem related features from the images. Based on these features, several PLS models for each specific age period were established. The age of unknown sample was estimated by applying all models and choosing one with the smallest deviation value. The results of experiments include comparison between different feature extraction approaches as well as between different bone types used for imagery.

T08. Metabolomic studies of experimental and clinical diabetic urine samples by ¹H NMR and LC/MS spectroscopy

A. Jankevics, M. Dambrova, E. Liepinsh, E. Liepinsh, S. Grinberga, O. Pugovics

Latvian Institute of Organic Synthesis, Riga, Latvia

Keywords: Goto-Kakizaki rat, diabetes, PLS-DA, PCA, ¹H-NMR, LC/MS, urine, metabolomics

In recent years, metabolomics has developed to an accepted and valuable diagnostic tool. The substantial improvements of analytical hardware allow routine analysis of biofluids in a high throughput manner and metabolic fingerprinting of metabolites. The obtained fingerprints may then be used to for early diagnosis, e. g., distinguishing between healthy and diseased individuals.

Nuclear magnetic resonance (NMR) spectroscopy and liquid chromatography/mass spectrometry (LC/MS) based approaches are two primary analytical methods of choice for conducting metabolomic measurements. ¹H-NMR spectroscopy requires minor sample pretreatment and measurement data can be presented vector-wise (i.e. one matrix for each data set). Even though LC/MS allows achieving higher sensitivity, the obtained results are a matrix for each data point, i.e., the retention time and mass spectra serving as two dimensions.

The goal of this study was to compare the performance of some common data processing methods for both NMR and LC/MS analytical methods. The data sets were obtained in non-obese experimental model of Type II diabetes in Goto-Kakizaki (G-K) rats. G-K rat exhibits metabolic, hormonal and vascular disorders similar to those of human diabetes, manifested as a fasting hyperglycemia, impaired secretion of insulin in response to glucose both in vivo and in isolated pancreatic cells, hepatic and peripheral insulin resistance. In present study, the urine samples were obtained from 30 G-K male rats (12 weeks, 290 g at the beginning of experiment) and 10 age matched male Wistar rats within 8 week period.

Following raw data pretreatment approaches were used:

- LC/MS data were processed by freely available software packages *MZmine* and *xcms* package from Bioconductor software project. Although both packages perform well, some simple data treatment methods are discussed for LC/MS data analysis.
- ¹H-NMR data were processed by freely available *HiRes* software from Hatch MR Research Center.

Several packages of statistical language R (pls, caret) and PLStoolbox for Matlab® were used to construct models for principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). Both accuracy and sensitivity for classification models were used as quality criteria of data processing. The results of PLS-DA statistical model show that both NMR and LC/MS analytical methods discriminate between G-K rats and control Wistar animals. In addition, PLS-DA model allows the differentiation of metabolic profiles related to animal age.

LC/MS method requires very accurate data acquisition control and the choice of data preprocessing parameters is the critical step of the procedure. Moreover, non-volatile sample components cause certain instability of ionization source.

Identification of metabolites related to differences between both sample groups might be more complicated if only LC/MS method is applied.

Although ^1H NMR approach requires no special sample preparation and allows easy identification of metabolites, many peaks in spectra are overlapped and signals of minor compounds are suppressed by those of native compounds in urine sample and water.

One can conclude that both analytical approaches yield similar results, however the ^1H NMR procedure is more robust and gives more information about the identity of metabolites.

The developed analytical approach was tested with pilot samples of human urine and good separation between healthy volunteers and diabetes patients was observed.

T09. Fast computation of CANDECOMP-PARAFAC and Tucker decompositions

D. Savostyanov

Institute of Numerical Mathematics RAS, Moscow, Russia

Keywords: CANDECOMP-PARAFAC, Tucker decomposition, nonlinear

Multi-way data structures are natural in chemometrics, and the most popular models for analysis of multi-way data, i.e. the Tucker and CANDECOMP-PARAFAC models, are widely used in this field. Many applications have shown that data analysis gains in robustness and information outcome by deriving meaningful parameters from multi-linear models of measured data. We propose very fast and robust method to approximate given data array by Tucker and PARAFAC structured models.

We consider Tucker-like approximations with an $r \times r \times r$ core tensor for three-dimensional $n \times n \times n$ arrays in the case of $r \ll n$ and possibly very large n (up to 10^4 – 10^6). As the approximation contains only $O(rn + r^3)$ parameters, it is natural to ask if it can be computed using only a small amount of entries of the given array. A similar question for matrices (two-dimensional tensors) was asked and positively answered 10 years ago in [1]. We extend the positive answer to the case of three-dimensional tensors. More specifically, it is shown that if the tensor admits a good Tucker approximation for some (small) rank r , then this approximation can be computed using only $O(nr)$ entries with $O(nr^3)$ complexity.

But the existence is good, but the algorithm is better. We also propose an algorithm to compute such approximation, based on the same ideas as incomplete cross approximation method, proposed for matrices in [2]. It successively computes rows, columns and "fibers" of 3D array $A = [a_{ijk}]$ (and hence is named "3D-cross method", or simply "Cross-3D") using a neat adaptive technique. However, the good implementation of the method (which is even better than algorithm) requires to solve a number of matrix problems. We have found a bunch of them, rather interesting and new. For example, given a factors low-rank matrix $A=UV^T$ (A is $n \times n$ and U, V are $n \times r$) how to find maximum modulus element of A in $O(nr^a)$ flops? (and how to provide the least a ?) This and some other questions are answered using the maximum-volume principle [3].

We are happy to notice, that the implementation of maximum-volume principle in 3D-cross method boosts its speed up to hundreds times.

Finally, we demonstrate the applications of our methods to different kinds of problems and show, how one can pack 128 petabytes of raw data into 114 megabytes of structured data with relative accuracy 10^{-5} at 30 minutes on 2GHz personal workstation.

References

1. S.A. Goreinov, E.E. Tyrtysnikov, N.L. Zamarashkin, *Linear Algebra Appl.* **261**, 1-21 (1997).
2. E.E. Tyrtysnikov, *Computing*, **4**, 367-380 (2000).
3. S.A. Goreinov, E.E. Tyrtysnikov, *Contemporary Mathematics*, **208**, 47-51 (2001).

T10. Intelligent module of stock market simulator

S. Sokolova

Saint-Petersburg State University of Aerospace Instrument Making, Saint-Petersburg, Russia

Keywords: interval OLAP-cube, immunocomputing, multivariable interval data

This paper provides a further development of mathematical models that use an Immunocomputing approach to predicting the stock market dynamics with interval input vectors. The application of interval rules and interval mathematics is then used to develop the computing techniques of supervised and unsupervised learning, and

classification and formation of the stock risk index in the context of interval input vectors. To illustrate this, the model example is provided.

T11. Computer methods for evaluation of the confidence intervals of molecular modeling prediction

V. Dementiev

Vernadsky Institute of Geochemistry and Analytical Chemistry, Moscow, Russia

Keywords: computer modeling, prediction error, supercomputer, confidence intervals

The report applies to such a branch of analytical chemistry, and of chemometrics as well, as spectroscopic analysis of mixed materials using no natural or artificial standards (etalons) [1]. In this rather new method the experimental spectrum of the mixture of assumed components is compared with theoretical ones of varied initial concentrations. The confidence level of the results depends on authenticity of molecular models and on the precision of their parameterization. Let us assume that the first is proved by the experience of inverse spectral problems solving for individuals in the mixture. Then accuracy of resulting theoretical spectrum is formed by precision of molecular parameters and by the calculation process details. At present we have no effective theoretical means for evaluation of above mentioned accuracy. It may well be that two or three ideas below will help.

1. Errors in molecular parameterization as any measurement errors are limited ones. It follows from the ROxY law [2]. (Here ROxY stands for Rodionova Oxana Yeugenievna).
2. As a consequence of ROxY law, the variances of predicted spectrum of the mixture are limited ones as well. So, the results of No-Standards molecular spectral analysis should always be valid ones.
3. The confidence intervals of molecular modeling prediction may be evaluated by means of direct computational formation with varied molecular parameters. The statistical properties of components of the parameters hazard vector are up to molecular model author, but they all might be taken into account using high-performance supercomputer technique in the regime of parallel calculations. The proposed method as Bootstrap method generalization allows clear visual representation of statistical properties of the final molecular modeling results.

References

1. L.A. Gribov, V.I. Baranov, M.E. Eliashberg. *No-Standards molecular spectral analysis. Theoretical principles*. Moscow, Editorial URSS, 2002, (In Russian)
2. Ox.Ye. Rodionova. Dr. Sci. Theses, Moscow, 2007.

T12. Elements of pattern recognition theory in the analysis of biosensor signals

¹A. Reshetilov, ¹A. Lobanov, ²O. Ponamoreva, ¹T. Reshetilova, ²V. Alferov

¹G. K. Skryabin Institute of Biochemistry and Physiology of Microorganisms RAS, Pushchino, Russia

²Tula State University, Tula, Russia

Keywords: amperometric microbial sensor, ethanol, glucose, xylose, selectivity, chemometrics, artificial neural networks

Detection and determination of concentrations of various substances in samples is one of the most relevant problems in analytical chemistry. The key components of biological liquids are determined for medical procedures; the concentrations of pollutants are determined at environmental monitoring; and the concentrations of substrates, products, and intermediate metabolites have to be controlled for efficient transformation of substances in biotechnological production.

Conventional methods of detection (mass spectrometry, chromatography, modifications of physicochemical analysis) are comparatively expensive, require complex equipment, and do not give results in the real time mode. An effective solution of the problem may be application of biosensors, in particular, microbial ones. At the same time, microbial sensors have low selectivity, i.e., sensitive to an ample quantity of substances. Hence, in some cases the range of their application is significantly limited. This presentation considers the general character and peculiarities of signal generation by microbial and enzyme biosensors, with an attempt to outline the most typical problems of using this analytical tool. Besides, the specific problem of recognition is considered. It has been shown possible to identify the components of a mixture containing glucose, xylose, and ethanol by sensors based on microbial cells. Cluster analysis and artificial neural networks were used for experimental data processing. The measuring system was represented by three amperometric microbial sensors with immobilized cells of *Gluconobacter oxydans*, *Hansenula polymorpha*, and *Escherichia coli*. At the analysis of 39 control samples, 37 samples were recognized correctly and two samples were recognized falsely (the

ethanol + xylose mixture was identified as ethanol only). It was shown that the processing by artificial neural networks needs a considerable data level for network training (with less than 70% of experimental data used for training, sample recognition was incorrect).

Finally, the possibility of combining the efforts of specialists in pattern recognition and developers of biosensors for joint solution of the problems of detection quality improvement is considered.

T13. PAT-enabling hard- and software

A. Bogomolov

J&M GmbH, Aalen, Germany

Keywords: PAT, NIR spectroscopy, Lighthouse Probe, software

Industrial process monitoring and control in the pharmaceutical industry, commonly referred to as Process Analytical Technology (PAT), is a front-end application area for scientific and technical innovations. Optical, and specifically, Near Infrared (NIR) spectroscopy and multivariate data analysis are important components of the PAT technology.

J&M has an essential expertise in the development and practical application of special spectroscopic equipment for PAT. One of recent developments of the company is the Lighthouse Probe™ (LHP), a NIR probe (sonde) for the on-line monitoring of granulation, drying, blending, and similar processes is presented.

Data analysis along with the corresponding software is another inherent side of any PAT solution. Modern PAT practice creates a demand for a new software concept to replace the conventional expert-oriented approach. Some ideas on the new-generation process chemometrics software are discussed.

T14. The study of dispersed systems using NIR spectroscopy and PLS regression technique

A. Potapov, A. Kalinin, S. Sadovsky

Institute of Spectroscopy of RAS, Troitsk, Russia

Keywords: spectroscopy, dispersed systems, PLS regression

The modified PLS-regression technique has been used for analysis of NIR-spectra of the number of dispersed systems with water as dispersed phase or dispersion medium. These systems are prevalent in nature and are of special interest

of physics, biology and chemistry as well as of various industries. Investigated systems were hydrated reversed micelles, water-in-crude oil emulsions, liquid and powdered milk, and water-ethanol-gasoline mixtures. The sizes of dispersed phase particles ranged from nanometers to micrometers. The purpose of our investigations was to create a method of NIR spectra treatment aimed to data extraction not only about the components maintenance of the dispersed system but else about its structure features. We used laser correlation spectroscopy of scattered light in addition to absorption and diffuse reflection spectroscopy techniques. This method allowed us to determine size and shape of dispersed phase particles and its size distribution in solution. Besides several calibration models were performed to study the application feasibility of two portable devices designed in the Institute of Spectroscopy; namely, of a grating spectrophotometer equipped with a Si linear sensor and of a diffuse reflection spectrometer based on InGaAr linear sensor and fiber-optical probe, as the tools for biotechnological product inspection. Some other models were used to discriminate sample temperature influence on a reliability of components maintenance prediction in the powder mixtures.

T15. The discriminating pharmaceutical raw materials on warehouse by FT-NIR spectrometer Spectrum 100 (PerkinElmer) with Fiber Optic Accessory

Y. Sokovikov

SchelTec AG, Moscow, Russia

Keywords: FT-NIR spectroscopy, pharmaceutical raw materials, SIMCA, fiber optic accessory

This presentation describes the use of FT-NIR spectroscopy to discriminate raw materials on warehouse through plastic packing. These materials were sampled using a diffuse reflectance fiber optic probe. Also are discussed problems which arise at construction and use of models on base SIMCA.

T16. Prediction of NMR Chemical Shifts. A Chemometrical Approach

K. Blinov

ACD Inc., Moscow, Russia

Keywords: NMR spectra, NMR Spectrum prediction, PLS, Neural

Nuclear magnetic resonance (NMR) spectroscopy plays a key role in determination of unknown chemical structures (Structure Elucidation). Generally, a

signal in a carbon NMR spectrum corresponds to the carbon atom in a chemical structure. Main characteristic of a signal in NMR spectrum is its chemical shift, which provides a lot of information about the corresponding carbon atom and its neighborhood. The chemical shift of a carbon atom depends on its neighborhood and so can be predicted based on this information. The best way to choose correct structure from a number of structures corresponding to spectral data is to compare the predicted experimental values with the experimental ones. The smaller difference means the corresponding structure. Today, this procedure is widely used in Computer-Aided Structure Elucidation (CASE).

All modern methods to predict a chemical shift can be considered of two types depending on the way of calculating a shift they based on. One of the techniques means that each neighbor atom has its own "increment" value (which depends both on the atom type and the distance to a target atom). The chemical shift of a target atom determined by this technique is the sum of the values of all neighbor atoms. The other way to perform this is to find some atoms with the similar neighbors from database and average their values to assign the obtained shift to a target atom. The first method is fast, but inaccurate, while the second one is usually accurate, but slow. As the result, the usage of both methods is limited in CASE systems.

In our work we tried to develop a fast and still accurate method of chemical shift prediction applicable for CASE. As we had a large amount of experimental data available (near 1.5 million of measured carbon chemical shifts), it seems reasonable to use the fast "increment" method and apply chemometrics methods to extract necessary parameters for this method from experimental data. PLS regression has been used for process experimental data. This method usually produces excellent results when the correct model used. The main goal of our work was to develop an appropriate model of structure representation for prediction of chemical shifts. The models were developed using knowledge about "the nature" of chemical shifts. Several models have been checked and the best one has been kept. Thus, chemometric allows development of a novel method of chemical shift prediction, which outperforms known methods and can be successfully used in CASE.

T17. Mathematical methods and algorithms used in the integrated program ARDES for automatic processing of atomic-emission spectra

S. Pavlov, E. Shabanova, I. Vasilyeva

Institute of Geochemistry SB RAS, Irkutsk, Russia

The integrated program ARDES has been developed for computer processing atomic-emission spectra as a visual interpretation. The atomic-emission analysis with visual interpretation allows a simultaneous determination of major- and trace elements with accounting matrix effects, spectral overlaps and large-range concentrations of the analyte. These requirements for computer processing are fulfilled using the multivariate analysis. The multivariate calibration models are constructed by such methods as OLS, PCR, and PLS. Primary information for regression models are reference and experimental data sets. The information is grouped into two tables. The analyte wavelengths are placed in the column headers, while the row headers contain the RSM names. At the intersection point of rows and columns of the first table there are certified concentrations of analyte, and in the second table there are analytical parameters of spectral lines. The major elements influencing on an intensity of spectral lines are selected by the PCA method. The wavelength and analytical parameters of spectral lines of interfering elements are included in the data for accounting spectral overlaps. The range of analyte concentrations can be broadened with RSM collection with various certified element concentrations from detection limits to some tens percent.

Application of multivariate methods demands a learning procedure and checking the calibration with the test samples. The back-propagation method with minimization of evaluation functions of RMSEC, RMSEP was used to select the best calibration model. The integrated program includes assessment of analytical result accuracy required for analytical works to be implemented.

Remarkably, that the combination of different methods, construction of multivariate calibration model, selection of the optimum calibration model and assessment of analytical results accuracy in the same program provides to exchange visual interpretation by computer processing.

T18. Classical and fuzzy principal component analysis of environmental samples concerning the pollution with heavy metals

C. Sarbu

Babes-Bolyai University, Cluj-Napoca, Romania

Keywords: PCA, Fuzzy principal components analysis, robust methods, soil samples

Principal component analysis (PCA) is a favorite tool in environmetrics for data compression and information extraction. However, it is well-known that PCA, as with any other multivariate statistical method, is sensitive to outliers, missing data, and poor linear correlation between variables due to poorly distributed variables. As a result data transformations have a large impact upon PCA. One of the most powerful approach to improve PCA appears to be the fuzzification of the matrix data, thus diminishing the influence of the outliers. In this work, we apply two robust fuzzy PCA algorithms (FPCA). The efficiency of the new algorithms is illustrated on data sets concerning the pollution of soil with heavy metals in north of Romania and the former East Germany. Considering, for example, a model with two components for Romania data set, FPCA-1 accounts for 77.16 % of the total variance and PCA accounts only for 57.42 %. Much more, PCA showed only a partial separation of the variables and scores, whereas a much sharper differentiation of the variables and scores is observed when FPCA algorithms were applied.

T19. POPLC technology: best separation in one step. Present state and perspective

¹*Y. Kalambet, ²P. Kozmin*

¹*Ampersand Ltd., Moscow, Russia*

²*Shemyakin Institute of Bioorganic Chemistry, Moscow, Russia*

Keywords: POPLC chromatography optimization separation

POPLC abbreviation stands for Phase OPTimised Liquid Chromatography, new approach to chromatographic separation. This approach is based on computer-assisted development of the column, optimised to particular separation. The idea of the approach is in construction of columns, composed of segments of different lengths, filled with stationary phases of different selectivity. Retention time for every target component can be measured on test columns of each type, and software calculates predicted times for the column, composed of any combination of segments. User has to enter system limitations (maximum run time, maximum length of the column, required resolution) and software looks for one of two segment

combinations, that could be of interest to the customer: "Best" combination is the one where best resolution, that can be achieved within limited time and column length is found, "Optimal" combination stays for column composition where required resolution of all peak pairs is achieved within minimum time.

Currently this approach is positioned by the manufacturer (Bischoff Chromatography) as an alternative to gradient separation in chromatography. Besides, approach of construction of the column, optimized for particular separation can be productive in wide number of fields, such as selection of optimal solvent composition for isocratic separation, optimisation of gradient separations. Those areas will require application of multiple chemometrics techniques, and can be an interesting field for application of efforts of chemometric community.

T20. Chemometrics in South Africa: applications, challenges, and the development of the SA chemometrics society

S. Paul

University of South Africa, Pretoria, South Africa

The agricultural and food sciences, mining and manufacturing industries, and some academic studies in South Africa benefit from the application of Chemometrics. These applications were initially performed by groups which did not know much about one another; however, as the interest in and the use of multivariate data analysis grew, and since there seemed to be a common connection in the form of advisory experts and software from the Scandinavian countries, the members of the groups started to meet and interact. The result was the founding of the SA Chemometrics Society, which has its main and administrative centre in the Cape Province (Stellenbosch), and also a dynamic branch in the Gauteng Province (Pretoria/Johannesburg). Interactions, collaboration, and common activities are being developed.

Examples of the activities of various sections, illustrating achievements in agricultural as well industrial applications, will be shown. The aims, problems and challenges to apply Chemometrics successfully, and the plans to accomplish this, will be outlined.

Posters

P01. Application of method PLS in the spectrophotometric analysis for simultaneous definitions of all substances in 2-6 plural-component mixes

I. Vlasova, E. Masjakova, A. Shelpakova

Omsk State University, Omsk, Russia

The low selectivity connected with imposing of absorption spectra is the basic lack of the spectrophotometric analysis which leads to increase in an error of definition of components at increase in their number.

The different regression methods are used for increase of accuracy of the analysis. For example, the method of Multiple Linear Regressions (MLR) can be applied to processing results of the analysis. However method MLR can give failure if multicollinearity is noted, i.e. presence of the internal, latent communications between variables. In this case more exact results can be received by a method of Projection on Latent Structures (PLS), based on construction of multivariate graduation.

This work is devoted to the application of method PLS in the spectrophotometric analysis of plural-component mixes. The absorption spectra of the given mixes received on the spectrophotometer SF-2000 in the quartz cavity a cm thick over the range 220-350 nm. 2-3-component mixes of pharmaceutical substances and 4-6-component mixes of water-soluble vitamins group B are used as model mixes. The qualitative composition and blending ratio are identical to some medical products and vitaminous second helping for birdseed.

Constructions of multivariate graduation were fulfilled with the use of absorption spectra of the mixes which were obtained with step of 0.4 and 5 nm. We have constructed and have checked up about 80 models for 2-6-component mixes with the purpose of a finding of optimum conditions at execution of experiment. In the models the number of mixes in the training set and number of principal components were varied.

Dependence of definition's errors of all components from the number of mixes in a training set and from the volume of an introduced spectral array is studied for all investigated mixes. We have shown that there is some optimum number of mixes in the training set. The error of definition of all components is increasing if we reduce

the number of mixes in a set and essentially it does not change if this number increases. The formula connecting an optimum training set with blend composition is deduced from this regularity. It is also established, that the error of definition depends on the volume of the introduced spectral array. The best results turn out at lead-in of greater number of lengths of waves, i.e. with step of 0.4 nm. Correlations of number of the principal component with blend composition have not been discovered. Models which gave the least errors of definition of all components were chosen for the analysis.

Method PLS is applied to the analysis of the model and real mixes containing 2-6 components. In all cases the error of definition of individual components amounted 0.5-2 % and the variation coefficient was below 1.5 % ($n = 3$, $P = 0.95$).

P02. On the interesting property of Maxwell-Boltzmann distribution got with data mining. Some cross-disciplinary applications

V. Prelov

Mechanical Engineering Research Institute of RAS, Moscow, Russia

Keywords: Maxwell-Boltzmann distribution, Data Mining

We present a result of analytical processing done for huge data arrays (300 mlns of records), mostly, of econometrical and geophysical nature. It was noticed, first, and verified and proved, then, the existence of the asymptotically stable limit value for the new parameter named "the field efficiency". We succeeded in explaining this non-trivial phenomenon and getting an explicit expression and numerical value for our constant. The thermodynamical approach made it possible to establish the relations between Maxwell-Boltzmann distribution and our parameter. Then, by means of number-theoretic methods we found our constant has the cross-disciplinary nature. Wide spectrum of possible and desired applications of this result discussed in the work. New found property shows real opportunities of OLAP and Data Mining.

P03. Application of near infrared spectroscopy, acoustic chemometrics, and process sampling for on-line monitoring and control of biogas processes

M. Madsen, J. B. Holm-Nielsen, K. Esbensen

Aalborg University Esbjerg, Esbjerg, Denmark

Biogas plants represent versatile biological processing plants that can be implemented for various purposes. For the purpose of energy production, biogas plants are capable of treating many different types of organic wastes, energy crops, and agricultural residues aiming at producing renewable energy (biogas) and organic fertiliser (digestate) for crop cultivation. In the context of wastewater treatment, biogas technology can be applied for removing persistent organic pollutants and thus secure the water environment. In order to be able to operate the biogas process optimally, reliable, fast, and extensive monitoring is needed. Otherwise the process might be imbalanced leading to severe economic losses.

The feasibility of using near infrared spectroscopy (NIR) for monitoring important process intermediates (volatile fatty acids and ammonia) suitable for advanced control of biogas processes is reported along with acoustic chemometrics (AC), which is an emerging technology that still needs well-documented feasibility studies.

In this study, a recurrent pump loop was mounted on the side of a 2400 m³ biogas reactor at the centralised biogas LinkoGas A.m.b.A., Lintrup, Denmark. The sampling strategy was developed in accordance with the Theory of Sampling (TOS) implying that incremental samples were taken from an upwards flowing stream yielding representative composite reference samples.

A commercial near infrared reflection probe was installed in the loop and spectra were acquired with an ABB Bomem MB160 FT-NIR spectrometer equipped with a sensitive InGaAs detector. A piezo-electric accelerometer type 4396 (Brüel and Kjær, Denmark) was mounted on a bend on the loop facilitating acquisition of acoustic spectra. Chemical reference analyses were performed in order to quantify the concentrations of volatile fatty acids (acetic and propanoic acid), total solids, volatile solids, and ammonia. All parameters were modelled using the PLS-1 algorithm. NIR spectra were MSC pre-treated to remove scatter effects from suspended particles and fibres.

Useful regression models were not accomplished using the natural variation in the biogas reactor. Addition of e.g. glycerol or pressure sterilised food-waste is necessary to manipulate the volatile fatty acid level and obtain a broader span in the calibration data. Further research and develop work is needed in order to optimise the optical interface to the process and automating the sampling procedure.

Acknowledgements

This research work was supported by LinkoGas A.m.b.A. and Aalborg University.

References

1. Holm-Nielsen, J.H., Andree, H., Lindorfer, H., and Esbensen, K.H. *J. Near Infrared Spectroscopy*, **15**, 123-135, (2007)
2. Holm-Nielsen, J.H., Lomborg, C.J., Oleskowicz-Popiel, P., and Esbensen, K.H., *Biotechnology and Bioengineering*, (in print)

P04. The application of Multivariate Data Analysis at the evaluation of the macroeconomic factors' dominance over the transaction costs of vertically-integrated oil companies

O. Klescheva

Kazan State Financial Economic Institute, Kazan, Russia

Keywords: Multivariate Data Analysis, macroeconomic factors, transaction cost, vertically-integrated oil company, projection methods

Nowadays projection methods are widely adopted in economics. Applying some projection methods to the analysis of macroeconomic factors' dominance over the quantity of the transaction costs of vertically-integrated oil companies, we made use of great number of various descriptions, including descriptions of political, economic and ecological circumstances. The results of the application of Multivariate Data Analysis at the evaluation of the macroeconomic factors' dominance over the transaction costs of vertically-integrated oil companies are presented in the work.

P05. Multivariate standard addition method solved by net analytesignal calculations

B. Hemmateenejad, S. Yousefinejad

Shiraz University, Shiraz, Iran

Keywords: standard addition, multivariate calibration, net analyte

Interferences are a common and severe problem that may render a chemical analysis invalid. In general, interferences can be divided into two classes: Direct and indirect. Indirect interferences or sample matrix effects include all the chemical and physical interferences that do not contribute directly to the measured signal, but affect the signal produced by the analyte of interest. Direct or spectral interferences are those which arise when a sensor is not completely specific for the analyte and are quite common in most spectroscopic methods of analysis.

The univariate standard addition method (SAM) only discovers the indirect interferences problem. When direct interferences are also present, the generalized standard addition method (GSAM) can be applied. This method is the generalization of standard addition in multivariate analysis. However, because of the complexity of the calculation, the number of literature reports using GSAM is rare.

In this study, a simple multivariate standard addition method using net analyte signal (NAS) calculations has been proposed. NAS is a part of mixture spectrum that is unique for the analyte and is orthogonal to the space of interferences. By the use of NAS calculations, we derived a linear calibration graph of the norm of NAS vs. the concentration of the standards added. In this way, the multivariate SAM was converted to a univariate model, similar to what it is obtained in conventional SAM. This proposed method is so simple and we can also calculate figures of merit simultaneously. The method was validated by the analysis of simulated data as well as spectrophotometric analysis of indicators binary mixtures with severe matrix effect as experimental model, and relative errors lower than 5% were obtained in most cases.

P06. Spectrophotometric monitoring of drugs photodegradation by multivariate curve resolution methods

¹*B. Hemmateenejad, ²K. Javidnia, ²R. Miri and ²M. Saeidi-Boroujeni*

¹*Shiraz University, Shiraz, Iran*

²*Shiraz University of Medical Sciences, Shiraz, Iran*

Keywords: drug, photodegradation, spectrophotometry, multivariate curve

Drug stability research is critical in pharmaceutical studies as the increased degradation decreases the potency of the drug and can create compounds with undesirable pharmacological effects. An increasing number of drugs belonging to different therapeutic classes (calcium channel blockers, non-steroidal anti-inflammatory drugs, chemotherapeutic agents, diuretics, benzodiazepines, beta-

blockers, etc.) were found to be photolabile. The most of the methods used for monitoring drug were different chromatographic methods, which are difficult to operate and use expensive instruments.

On the other hand, spectrophotometric methods are in general simple, sensitive and very suitable for studying chemical reactions in solutions. The spectral overlapping, as the major problem in almost all of the spectrochemical methods, can be overcome utilizing different chemometrics methods. For example, spectral curve deconvolution or multivariate curve resolution (MCR) methods are chemometrics techniques concerning with the extraction of the pure spectra and concentration profiles of the components in a chemical reaction preceded in an evolutionary process.

In our continuing investigations on the application of different MCR methods in drug photodegradation studies, in this work we applied a collection of different MCR methods including Kubista method, ITTFA, MCR-ALS and combined hard model MCR-ALS to study the photodegradation kinetics of nitrendipine and felodipine (as calcium channel blocker agents) and nimesulide (as COX-2 inhibitor anti-inflammatory drug). By application of factor analysis and evolving factor analysis to the data matrices of absorbance spectra recorded at different lighting times, one photodegradation product was detected for nitrendipine and felodipine, and two photodegradation products were indicated for nimesulide. By using different MCR methods, the optimum concentration profiles and pure spectra of reactant and product(s) were calculated. The results showed that in the case of nitrendipine and felodipine the photodegradation reaction was zero-order and changed to a first-order manner when the concentration of product exceeded than that of reactant. On the other hand, a two step consecutive kinetic of the form of A B C were obtained for nimesulide, where both step obey first-order kinetic. By fitting the resulted concentration profile of each drug to the kinetic models, the respective rate constants were calculated.

P07. The data processing method of multisensor system for natural water monitoring

*S. Morzhukhina, P. Gladyshev, E. Osadchii, A. Chulkova, N. Shastova
State university "Dubna", Dubna, Russia*

Keywords: data processing, multisensor system, monitoring, natural

It is impossible to construct universal monitoring the big number of components because of the economic reasons. Therefore at the first stage of monitoring expediently definition of the integrated and generalized parameters which totally characterize the general danger of various pollution. Only after finding of pollution and toxicity on integrated parameters there is a necessity of expensive definitions of pollution levels on individual components. It is offered to use for primary monitoring multisensor system including 12 sensor: measurements of the temperature, the dissolved oxygen, electroconductivity, electrodes of measurement pH, Eh, concentration Ca^{2+} , NH_4^+ , NO_3^- , Na^+ , Cl^- , $\text{Ca}^{2+}+\text{Mg}^{2+}$ measurements.

For simplicity of perception received primary analytical information the method of mathematical visualization of data with formation of quality (or pollution) water a graphic image is used. For reception of a graphic image the simplest technique of generalization for multisensor information by means of the modified petalled diagrams is chosen. As analytical signals for various sensors have different interval of measured values, the algorithm of graphic image construction of quality provides normalization for all signals within the limits of chosen for each signal of concentration interval. Maximum-permissible concentration or other statistically proved norms of measured or calculated parameters are chosen as norm. The processing algorithm also considered that for some parameters are regulated by normative documents not only maximum, but also the minimum limit of a parameter value.

The chosen parameters and methods of their processing allow making a multiplane estimation of quality of water — from an ecological estimation on separate parameters of water quality till their graphic integrated image.

P08. The mechanism and kinetic model for enhancement of the initiating activity of three component photoinitiator systems in solid polymers

E. Khavina

Institute of Chemical Physics RAS, Moscow, Russia

Keywords: kinetic model, polymer photochemistry, photooxidation, photoinitiating systems

Experimental results are presented which suggest a common character of the synergism phenomenon for photoinitiator systems composed of an aromatic ketone, a halomethylaromatic compound, and an aliphatic or aromatic amine. The influence of

the nature of components of such systems, their concentration ratio, polymer structure, and illumination conditions on the synergistic effect was considered.

A kinetic model of the phenomenon taking into account the mechanism of action of individual components: the influence of amine on the quantum yield of primary radical pairs and the enhancement of radical escape from the cage by the action of the halomethylaromatic compound, was derived. The model fairly well describes experimental data and allows both the scale of the phenomenon and the optimal ratio between the components of photoinitiator systems to be predicted.

P09. Application of nonlinear PCR for optimization of the hybrid binder used in construction materials

¹*I. Starovoitova, ¹V. Khozin, ¹L. Abdrachmanova, ²O. Rodionova, ²A. Pomerantsev*

¹*Kazan State University of Architecture and Engineering, Kazan, Russia*

²*Institute of Chemical Physics RAS, Moscow, Russia*

Keywords: PCA, nonlinear PCR, optimization of hybrid binder for construction composite material

The present research is aimed at optimization of a hybrid binder formulation that includes organic and inorganic oligomers. Inorganic component is a water solution of sodium silicate (liquid glass=LG) and organic additive is polyisocyanate (=PIC) that reacts with moisture forming a hardened composite. It is known that such the formulations yield different binders having valuable practical features.

The following features (input variables) constitute the investigated formulation: SiO₂/Na₂O weight ratio (x_1), LG density (x_2), water contents in LG (x_3), PIC contents (x_4), and LG contents (x_5). The optimization is performed with respect to twelve output quality characteristics such as elasticity (y_1), hardness (y_2), heat-resistance (y_3), etc.

The data were obtained at fractional factorial experiment with 27 formulations. The PCA of **X** block (27x5) explains 99% of data variation with two PCs. However, the direct PCR (as well as PLS2) modeling reveals strong nonlinear relations between **X** and **Y** blocks (27x12).

Therefore, calibration modeling is done as a two-step procedure. At first, PCA is applied to the **X**- block for variable reduction. Then nonlinear PCR approach is used. Namely, each response variable y is modeled using a quadratic equation –

$$y = b_{00} + b_{01}t_1 + b_{02}t_2 + b_{11}t_1^2 + b_{12}t_1t_2 + b_{22}t_2^2,$$

where \mathbf{t}_1 and \mathbf{t}_2 are the PCA score vectors, and b_{ij} are the regression coefficients.

Each model predicts a particular quality characteristic y as a function of scores \mathbf{t}_1 and \mathbf{t}_2 with appropriate accuracy. The input variables reduction enabled us to choose an optimal binder formulation that meets the predefined quality requirements. The prediction has been confirmed by the auditorial experiments.

P10. Block-based approach to modelling of granulated fertilizers' quality

J. Kohonen

Lappeenranta University of Technology, Lappeenranta, Finland

Keywords: fertilization industry, granulation, priority PLS, multi-block PLS

Fertilizer manufacturing is a customer-driven industry, where the quality of a product is a key factor in order to survive the competition. However, measuring the most important feature with granulated fertilizers, flowability, is tedious, time-consuming and thus expensive.

Flowability can be defined through testing the flow rate with e.g., seed drill. Besides the chemical composition, flowability can be considered as one of the most important characteristics. There are numerous factors affecting the flowability of a granulated fertilizer, most of them related to the crystallization process. Particle size distribution of the granulated product has to be within the customer specification, but is also highly significant factor, especially in the presence of moisture. This can lead to Ostwald ripening, where small particles dissolve while the large ones grow. Another difficulty is agglomeration and aggregation of granules. Chemical composition affects the crystallization process and is significant for particle shape and various physical properties.

The present approach is based on priority regression and multi-block PLS. The data is measured from the final product and is divided in blocks between physical properties, such as granule hardness and roundness, chemical composition and particle size distribution. The goals are to find a reliable model for flowability using this data and to find the most important variables.

P11. A study of spatial structure of the Chernobyl ^{137}Cs contamination field in natural forest landscape

E. Korobova

Vernadsky Institute of Geochemistry and Analytical Chemistry, Moscow, Russia

Keywords: spatial structure, ^{137}Cs contamination, spatial analysis, radioecology, landscape geochemistry

^{137}Cs was used as a tracer to test the hypothesis of the regular secondary ^{137}Cs redistribution in natural landscapes (soil and moss cover). The initial deposition within the study area (70x100 m) was suggested to be uniform. Field gamma -spectrometry was performed in grid (5 m for the whole area; 1 m for 5x5 m plot; 0,2 m for 1x1 m plot) with the help of gamma-spectrometer VIOLINIST III (USA) equipped with scintillated detector. Vertical distribution of radiocesium was studied by core sampling to the depth of 40 cm. Moss samples collected from the area 15x15 cm in 10 m grid. Laboratory determination of ^{137}Cs in soil and moss samples was performed with the help of Canberra gamma-spectrometer (HP Ge detector).

Field gamma-spectrometry revealed a system of ^{137}Cs polycentric anomalies both in soil and moss cover. Spatial structure can be followed on different scales down to 20-150 cm and is reflected in the form of distribution histogram. ^{137}Cs field in mosses although regular did not corresponded to that in the soil completely and probably reflects peculiarities of radionuclide fixation and uptake. The absence of significant erosion in woodlands allow suggest that the observed structure result from peculiarities of the secondary redistribution of atmospheric contamination due to water mass migration related to meso- and microscale relief features . Performed study proved that spatial distribution of particular element or compound can be studied in the environment as a type of geofield with regular structure and presents an original object for spatial investigation and analysis.

The authors are grateful to Dr. Linnik for topographic map and his earlier data on ^{137}Cs contamination of the site.

P12. Genetic algorithm as a variable selection procedure for the simulation of ¹³C nuclear magnetic resonance spectra of flavonoids derivative using multiple linear regression

R. Ghavami

University of Kurdistan, Sanandaj, Iran

Keywords: flavonoids, NMR, simulation, genetic algorithm, QSPR

In order to accurately simulate ¹³C NMR spectra of hydroxy, polyhydroxy and methoxy substituted flavonoids a Quantitative Structure-Property Relationship (QSPR) model, relating atom-based calculated descriptors to ¹³C NMR chemical shifts (ppm, TMS = 0), is developed. A dataset consisting of 50 flavonoids derivatives was employed for the present analysis. A set of 417 topological, geometrical, and electronic descriptors representing various structural characteristics was calculated and separate multilinear QSPR models were developed between each carbon atom of flavonoids and the calculated descriptors. Genetic algorithm (GA) and multiple linear regression analysis (MLRA) were used to select the descriptors and to generate the correlation models. Analysis of the results revealed a correlation coefficient and Root Mean Square Error (RMSE) of 0.998 and 1.42 ppm, respectively, for the prediction set.

P13. QSPR Models for predicting the environmentally relevant physicochemical properties of some organoiodine compounds from molecular topology using genetic-MLR and genetic-PLS

R. Ghavami

University of Kurdistan, Sanandaj, Iran

Keywords: environmentally relevant physicochemical properties, QSPR, organoiodine compounds

A Quantitative structure-property relationship (QSPR) analysis had been applied to a set of organoiodine compounds which were of special interest because of their roles on environmental samples. Semi-empirical quantum chemical calculations at AM1 level was used to find the optimum 3D geometry of the studied molecules. Modeling of the octanol-water partition coefficient, aqueous solubility, air solubility, enthalpy of vaporization and enthalpy of solution of 43 organic iodides compounds as a function of their theoretically-derived descriptors (i.e. topological indices) was established by means of multiple linear regression (MLR) and partial least squares (PLS) regression methods based on genetic algorithm for descriptor selection. The

stability and validity of models were tested by cross-validation technique and by prediction of the response values for the external prediction set. The average percent relative errors of prediction for MLR and PLS models were 2.94 and 0.36, respectively, which indicated the success of the developed QSPR models in modeling the physicochemical properties of the organoiodine compounds.

P14. Solving classical statistical problems using non-statistical methods: analysis of variance and time trends modeling

S. Zhilin

Altai State University, Barnaul, Russia

Keywords: analysis of variance, time trends, fitting experimental data, interval error

If the error in linear regression model is assumed bounded or, in other words, is represented by interval, then the set theoretic estimation approach may be used for regression building instead or along with traditional statistical methods. This idea was independently realized in almost the same form by several authors (see [1] and references therein). A lot of techniques for building and analyzing of regression model under interval error are proposed but their further development is actual as usual.

It is well known [2] that classical regression analysis tools could be used to solve problems of variance analysis and time trends by introducing "dummy" variables in regression models. Using simple data sets from [2] we show how this technique could be used for building and analyzing regression under interval error as well. In particular we consider problems of variance analysis and two variants of problem of taking in account time trends: when the point of trends intersection is known and unknown.

References

1. Zhilin S.I. *Chemom. Intell. Lab. Syst.*, **88** (1), 60-68 (2007).
2. Draper N.R., Smith H. *Applied Regression Analysis*, Wiley, 1981.

P15. Tree-electrode system of divided cells. Problem solving of drifts in voltammetric responses

A. Sidelnikov, F Kudasheva, V. Maistrenko, R. Zilberg, D. Bikmeev
Bashkir State University, Ufa, Russia

Keywords: voltammetric electronic tongue, PCA, drift of responses

A survey of publications devoted to voltammetric electronic tongue shows that there is a time drift of responses in electrochemical measurements taken in succession. A similar phenomenon is observed during the reduction of nitrocompounds in the system of divided cells (SDC). It is impossible to get a stable PCA-models in the time it takes for SIMCA-classification: the same samples are often erroneously classified as basically different. In order to resolve the problem we investigated voltammograms of p-, o-nitroanilines and nitrophenols received every 5 minutes within two hours. It was found the samples obtained in final time were not successfully identified as initial time samples by SIMCA.

For decreasing time-drift of samples in multivariate spaces we suggest to use tree-electrode (or more) PCA-models that including voltammograms of three (or more) nitrocompounds instead one. It was shown the tree-electrode PCA-models were more stable in time than mono-electrode ones. It is expedient to identify a solution placed into divided cell by using tree-electrode SDC. Practically, the tree-electrode PCA-model of a solution is multivariate nature-print of it. The visual nature-prints of various “Vodka” were obtained with use of PCA. It was shown the modified SDC makes it possible to identify samples of vodka with an accuracy exceeding 80 percent.

P16. Global Collaborative Knowledge System — is it possible?

E. Milosz

*High School of Enterprise and Administration, Lublin University of Technology,
Lublin, Poland*

Keywords: knowledge systems, knowledge management,

Global Collaborative Knowledge Systems (GCKS) base on worldwide collaboration in knowledge acquisition, content creation with permanent and fast feedback, and more or less democracy of participants. Knowledge acquisition in GCKS is massively and decentralise due to the big scale of the Internet and special futures of software tools. Author discusses advantages and disadvantages of GCKS.

P17. Data Mining as a Modern Method of Data Analysis

M. Milosz

Lublin University of Technology, Lublin, Poland

Keywords: Data Mining

Some remarks and examples of using the data mining methods to data analysis are presented in this paper.

P18. New results in applications of p-adic pseudo-differential equations to the protein dynamics.

¹V. Avetisov, ¹A. Bikulov, ²A. Zubarev

¹*Institute of Chemical Physics RAS, Moscow, Russia*

²*State Academy of Railways, Samara, Russia*

Keywords: models for biopolymer dynamics, p-adic mathematical physics, theory of random processes.

On the First International conference on p-adic mathematical physics, we pointed out that the protein dynamics can be described by the p-adic pseudo-differential equation of ultrametric diffusion [1]. We suggested an ultrametric model for the ligand rebinding kinetics of myoglobine and demonstrated good agreement with experimental data. Our new application of p-adic pseudo-differential equations to the protein dynamics is related to the phenomenon of spectral diffusion in globular proteins. Spectral diffusion is a peculiar random process propagated on a "frequency line", which is observed by measurements of the absorption frequency of a marker injected into the protein macromolecule. Two distinguish feature are inherent in spectral diffusion: anomalously slow widening of the spectral diffusion kernel and aging effect. We present a model of spectral diffusion in proteins based on ultrametric description of the protein dynamics and exhibit excellent agreement with experimental data in this case too. These results support an idea that proteins are macromolecular structures with ultrametric order.

References

1. Avetisov V.A., Bikulov A.Kh., Osipov V.Al. *J.Phys.A Math. Gen.* (2003) **36**, 4239.

P19. Dissolution test of drug release: Uncertainty measurements

M. Paakkunainen

Lappeenranta University of Technology, Lappeenranta, Finland

Keywords: sampling error, sampling theory, uncertainty, drug release, dissolution

Drug dissolution is important in pharmaceutical research. In this case study, Pierre Gy's sampling theory was applied, when the minimum standard deviation of

the drug release procedure is estimated. In Pierre Gy's sampling theory all the aspects of sampling are analyzed, such as material properties as well as the design of the sampling equipment. In this case study the error components assumed to affect to the global (total) estimation error are: error when sampling is assumed to be ideal, analytical error and error arising from the heterogeneous samples.

In the studies mixtures of few drug compound and excipient are examined. From this lot a primary sample is taken and tabletted. Before the samples are tabletted the homogeneity of the mixture is tested (with UV-measurements). After tableting the dissolution is estimated at different time points (from 5 min to several hours). From the UV-measurements the drug release is calculated and reported as percentage of dissolved drug.

For the drug mixtures studied the FSE seems to be for most of the cases quite small, lower than 1.0 %. FSE is the minimum error of the process. It is the error component that cannot be eliminated. It can be estimated based of the physical properties of the sample. Analytical error of the dissolution procedure was found to be about 2.0 %, but it clearly depends on the drug compound studied. The results of the homogeneity tests show that the primary powder mixture is heterogeneous, as the results of the two homogeneity replicated differs from the target drug concentration.

The total uncertainty of the dissolution procedure can be clearly improved with optimization.

P20. Quantitative assessment of glucose, methanol and ethanol contents in their mixture using biosensors and neuronet technologies

V. Arlyapov

Tula State University, Tula, Russia

Keywords: multicomponent analysis, biosensor, microbial sensor, artificial neuronets (ANN), immobilization of bacteria, glucose, ethanol, methanol

The methods of production of thin films containing immobilized *G. oxydans* and *P. angusta* cells have been developed. It has been shown possible to obtain bioreceptor elements with different sensitivities to glucose and ethanol depending on the method of immobilization. Systems of free microbial sensors were created for selective estimation of glucose, methanol and ethanol contents in the mixture of these compounds. The technology of artificial neuronets (ANN) applied for experimental data processing allowed for selective analysis in the concentration range of 0.16 to

5.00 mM by each of the substrates. It has been shown that the data processing by neuronets gives a relative error of concentration detection within 2-33% with the best net structure

P21. A neural network approach for classification of *Helicobacter Pylori* based on national strains

B. Karlik

Fatih University, Istanbul, Turkey

Keywords: national strain, neural network, backpropagation algorithm

The studies for sequencing the genome of *Helicobacter pylori* are completed and several strains are known as now. The genome of the strain "26695" consists of about 1.7 million base pairs, with some 1550 genes. The two sequenced strains show large genetic differences, with up to 6% of the nucleotides differing. Study of the *H. pylori* genome is centered on attempts to understand pathogenesis, the ability of this organism to cause disease. However, the purpose of this project is about classification of *Helicobacter pylori*. As it is stated above there are several strains of the genome and national strains are a class of these strains. The national strains are based on the nation of the host human. The cause of the difference on national strains is the difference of the cultural activities. If the national strain of a random *H. pylori* bacteria is known, the host human's nation can also be known approximately. The classification of DNA is difficult because the sequence of DNA differs from among all individuals. As a highly developed technique of classification artificial neural networks can be used to solve this problem. The aim of this study is to classify *Helicobacter pylori* according to National strain using artificial neural networks which has backpropagation learning algorithm.

P22. Acoustic emission usability in determining granule water content and granule size during fluidized bed granulation

¹*S. Matero, ¹S. Poutiainen, ¹J. Leskinen, ¹M. Hakulinen, ¹M. Lahtela-Kakkonen, ¹J. Ketolainen, ¹R. Lappalainen, ¹K. Järvinen, ²S. Reinikainen, ¹Ai Poso*

¹*University of Kuopio, Finland*

²*Lappeenranta University of Technology, Finland*

Keywords: PAT, on-line control

Fluidized bed granulation is a complex and multivariate like any other process [1] and it must be handled in a multivariate way. Until recently granulation has been

monitored by recording a few process parameters separately, e.g. in - and outlet air humidity and temperature [2]. Separate parameters carry information of the process based on earlier experiments using the same process conditions, and the end-point of granulation can be estimated using experience that has led to a desired end product. However, separate process variables do not provide a comprehensive picture of the granulation process. Furthermore such operating methods ultimately based on users' experience, are not in agreement with the high-quality product manufacture proposed in FDA's PAT initiative. Thus, further development of sophisticated controlling and monitoring systems is needed.

Binder content of the fluidized bed affects the granule growth rate, size and eventually, yield. An even distribution of the aqueous binder produces narrow granule size distribution, while over wetting the mass can lead to a bed collapse. Thus, the granulation process can be followed by tracing the water content of the fluidized mass. A common technique for water content measurement is Karl Fischer titration, but it requires the removal of the sample from the process line, and obtaining statistically reliable results can be cumbersome. Near-infrared spectroscopy is a capable tool in determining water content and granule size during granulation [3], but collecting representative spectra can be problematic due to the probe or window contamination. Passive acoustic emission is a technology used in a wide variety of chemical engineering processes [4]. The elastic properties of materials depend on moisture content, affecting the acoustic emissions caused by particle impact or friction. In addition, the dependence of water content on process parameters is also valuable for moisture prediction.

Our study aimed to develop models for real-time water content and granule size determination during granulation. For the models, we combined the information from the process parameters and acoustic emission data using PLS regression. The AE frequency spectra were subdivided and averaged into 32 segments in order to simplify the data for modelling purposes. The models represented enable the granule water content to be tracked throughout the granulation process and also the granule size determination during fluidization. The relative humidity of ambient air is crucial for determining the granule moisture and it is of importance to be able to stabilize its effect on the model. The data analysis were carried out using PLS_toolbox [5] version 4.0.

References

1. Kourti T. *Crit. Rev. Anal. Chem.* 2006, **36**, 257-278.

2. Rambali B. et al. *Int. J. Pharm.* 2001, **220**, 149-160.
3. Rantanen J. et al. *Chemom. Intell. Lab. Syst.* 2001, **56**, 51-58
4. Boyd J. *Chem. Eng. Sci.* 2001, **56**, 1749-1767
5. PLS_toolbox4.0, Eigenvector Research Inc., <http://software.eigenvector.com/#pls>
[11 January 2008]

P23. Multivariate calibrations in atomic emission spectrometry for analysis of solid geological samples

E. Shabanova, S. Pavlov, I. Vasilyeva

Institute of Geochemistry SB RAS, Irkutsk, Russia

Keywords: atomic emission spectrometry with arc discharge, solid

The atomic emission spectrometry with arc discharge for analysis of the solid geological samples determines simultaneously up to 60 major and trace elements. The modern spectrometers provide multichannel spectra recoding and measuring of analytical signal set from atom of each element. The "difficulties" of this analytical method includes the lack of any physical model describing a dependence of intensity of spectral line on analyte content with sufficient precision in the presence of significant variations of sample compositions. The usage of multivariate information calibration model takes into account the influence of substance evaporation process, transfer of vapor and atom exciting in plasma, construction features of device, which causes matrix effects and spectral overlaps.

We study ways of constructing data tables for train and test sets and also conditions of applying multivariate calibrations (OLS, PCR, PLS). The choice of more informative characteristics (i.e. some analyte lines, macro compounds and interferences; their contents in reference standard materials) involved in table structure is based on experience of visual spectral interpretation, analysis of principal components and cluster analysis of spectral data. The chemometric selecting type of calibration model in given analyte concentration range for a group of petrologic compositions. The optimization criterion is to minimize simultaneous an error of analyte determination in train and test samplings.

The application of multivariate calibration reduces significant errors of direct atomic emission analysis results of Ag, As, Sb, Cu, Zn, P, B, Mn and other elements in rock and decay for geochemical prospecting of gold and base metal deposits.

P24. Chemometric functions in Excel

A. Pomerantsev, O. Rodionova

Semenov Institute of Chemical Physics RAS, Moscow, Russia

Keywords: Add-In, Worksheet function, PCA, PLS

Chemometrics uses a very large variety of software, special chemometric and general mathematical packages or various environments as Matlab or VBA. As a result, to make first steps for a student or analyst it is necessary to obtain some special software and to acquaint with it.

To make the chemometric start quick and easy we propose to design the basic projection methods as worksheet functions in Excel, a most widely-spread data handling environment. In this case all calculations are carried out in the open Excel books. Moreover all regular Excel capacities can be applied for additional calculations, graphical presentations, export and import of data and results, customizing individual templates, etc.

Excel 2007 gives additional incentive to this idea as now very large arrays (1,048,576 rows by 16,384 columns) can be input and processed directly in the worksheets.

We have designed the core functions for the PCA/PLS decompositions and ensured that calculations are performed very quickly even for rather large data sets (200 samples by 4500 variables). These functions are programmed in C++ language and linked to Excel as an Add-In tool named Chemometrics.

The goal of this presentation is to discuss the benefit of such software and find out what additional functions should be included in Chemometrics Add-In to make its application simpler and user-friendly.

P25. Two atypical problems, which can be solved by projection methods

O. Karpukhin

Semenov Institute of Chemical Physics RAS, Moscow, Russia

Keywords: Projection methods, sets of polymer chains, information compression

1. Elaboration of methods for quantitative characterization of polymer chain configurations.

The PCA method was applied to the sets of various computer simulated polymer chains (PC). There were established the regularities in changes of scores and loadings depending on PC simulation algorithms and dynamics of change in the PC sets. The perspective of development of these methods and their application for the PC configuration characterization is discussed.

2. Opportunities of information compression by projection methods.

Projective methods enable to reveal and to assess quantitatively the distinct features of multivariate data sets, which are similar in some sense. The basic opportunity of information compression and reproducibility in similar data sets is demonstrated. Losses of information and other problems of signal restoration for different degrees of compression are analyzed on the example of a smoothed white noise signal.

P26. IR spectroscopy and PCA applied for polymer sorting in the domestic waste

N. Ryumina, E. Mikhailov

Samara State Technical University, Samara, Russia

Keywords: classification, SIMCA, recycling

Polymers constitute only 10-15% of the domestic refuse mass. They, however, could give some 60% of the reclamation profit. The conventional sorting methods based on a visual inspection fail to reveal all valuable polymer components.

This work aims at developing of a chemometric based method that employs IR spectroscopy for polymers identification. Samples of polypropylene (PP), polystyrene (PS), polyethylene (PE) and polyethylene terephthalate (PET) refuse were collected at the landfills. The samples differ in composition, colour, fouling factor, and appearance. IR spectra were obtained on AVATAR 360 FT-IR instrument in the range 400- 4000 cm^{-1} .

Five PCA models were constructed for the data: the general one for all polymers together, and four individual models for each polymer. SIMCA method was then applied for identification of a new polymer samples

P27. Metric on a set results chemical analysis

B. Chuprina

Institute of Geology of Industrial Minerals, Kazan, Russia

Chemometrics is based on the Euclid metric with element of a distance:

$$ds^2 = \sum_{i=1}^n dx^2$$

Parameters of a distribution (mean, variance, etc.) is calculated in this metric. However, Euclid metric, alike Euclid geometry, is first it, but not the only it. A chemical component x_i may be $0 \leq x_i \leq 1$, i. e.

$$x_i \in B = [x_i \in R, \|x\|^2 \leq 1]$$

With element of a distance

$$ds^2 = \frac{4 \sum_{i=1}^n dx_i^2}{\left(1 - \sum_{i=1}^n x_i^2\right)^2}$$

Of course, the value of the mean, variance, etc. will be distinguish from Euclid value.

List of participants

Arlyapov, Vyacheslav	Tula State University Researcher Tula, Russia	gwinbleidd@rambler.ru
Belyaev, Ivan	Altai State University Ph.D. student Barnaul, Russia	+7 (3852) 364185 iab@asu.ru
Bikmееv, Denis	Bashkir State University Student Ufa, Russia	limm@list.ru
Bikulov, Albert	Institute of Chemical Physics RAS Chief researcher Moscow, Russia	+7 (495) 9397227 bikulov1903@rambler.ru
Blinov, Kirill	ACD Inc. Leading Expert Moscow, Russia	+7 (495) 4382153 kirill.blinov@gmail.com
Bogomolov, Andrey	J&M GmbH Project Manager Aalen, Germany	+49 736192810 bogomolov@j-m.de
Bykov, Dmitry	Samara State Technical University Head of the Chair of Chemical Technology and Industrial Ecology Samara, Russia	promeco@sstu.smr.ru
Chudova, Julia	Moscow State University Student Moscow, Russia	+7 (495) 9393031
Chulkova, Alena	State university “Dubna” Student Dubna, Russia	+7 (4962) 22803
Chuprina, Boris	Institute of Geology of Industrial Minerals Researcher Kazan, Russia	+7 (843) 2363873 chuprina-41@mail.ru
de Juan, Anna	Universitat de Barcelona Associate professor Barcelona, Spain	+34 (93) 4034445 anna.dejuan@ub.edu

Dementiev, Vasily	Vernadsky Institute of Geochemistry and Analytical Chemistry Leading Researcher Moscow, Russia	+7 (495) 9395223 d_vasily@mail.ru
Elyashberg, Michail	ACD Inc. Leading Expert Moscow, Russia	+7 (495) 4382153 elyas@acdlabs.ru
Esbensen, Kim	Aalborg University Esbjerg Professor Esbjerg, Denmark	+45 7912 7688 kes@aaue.dk
Furer, Victor	Kazan State University of Architecture and Engineering Professor Kazan, Russia	+7 (8432) 104737 furer@ksaba.ru
Ghavami, Raof	University of Kurdistan Assistant Professor Sanandaj, Iran	+98 8716624204 rghavami@uok.ac.ir
Hemmatenejad, Bahram	Shiraz University Assistant Professor Shiraz, Iran	+98 9177096402 hemmatb@sums.ac.ir
Höskuldsson, Agnaar	Centre for Advanced Data Analysis Professor Kgs Lyngby, Denmark	ah@predict.ws
Ibyatov, Ravil	Kazan State Agrarian University Professor Kazan, Russia	+7 (8432) 5104677 r.ibjatov@mail.ru
Ivanova, Mariya	Institute of Geology of Industrial Minerals Engineer Kazan, Russia	i.marry@mail.ru
Jankevics, Andris	Latvian Institute of Organic Synthesis Student Riga, Latvia	+371 26308876 andza@osi.lv
Kalambet, Yuri	Ampersand Ltd. Director general Moscow, Russia	kalambet@ampersand.ru

Kalinin, Andrey	Institute of Spectroscopy RAS Senior researcher Troitsk, Russia	+7 (495) 3340235 kalinin@isan.troitsk.ru
Karlik, Bekir	Fatih University Professor Istanbul, Turkey	bkarlik@fatih.edu.tr
Karpukhin, Oleg	Institute of Chemical Physics RAS Laboratory Manager Moscow, Russia	kon1935@mail.ru
Khaliullin, Marat	Kazan State University of Architecture and Engineering Associate professor Kazan, Russia	+7 (843) 5104727 snxm2@bk.ru
Khalymbadga, Inna	Institute of Geology of Industrial Minerals Researcher Kazan, Russia	khalym@bk.ru
Khavina, Elizaveta	Institute of Chemical Physics RAS Researcher Moscow, Russia	+7 (495) 9397314 kelizabeth@yandex.ru
Khodyreva, Nadezhda	Saint-Petersburg State Technological University of Plant Polymers Vice Rector for International Relations Saint-Petersburg, Russia	+7 (812) 7865809 zsv@gturp.spb.ru
Klescheva, Olga	Kazan State Financial Economic Institute Ph.D. student Kazan, Russia	+7 (950) 3252285 Olgaklescheva@mail.ru
Kohonen, Jarno	Lappeenranta University of Technology Ph.D. student Lappeenranta, Finland	+358 56212152 jarno.kohonen@lut.fi
Korobova, Elena	Vernadsky Institute of Geochemistry and Analytical Chemistry Science Secretary Moscow, Russia	+7 (495) 1372484 Korobova@geokhi.ru

Kucheryavskaya, Elena	Accompanying person Esbjerg, Denmark	
Kucheryavskiy, Sergey	Aalborg University Esbjerg Assistant Professor Esbjerg, Denmark	+45 79127707 svk@aaue.dk
Kuptsov, Albert	Bruker Ltd. IR applications leading expert Moscow, Russia	+7 (495) 5029006 akuptsov@bruker.ru
Latypov, Valeriy	Ufa State Petroleum Technological University Head of "Building constructions" department Ufa, Russia	stexpert@mail.ru
Latypova, Tatiana	Ufa State Petroleum Technological University Deputy head of "Building constructions" department Ufa, Russia	stexpert@mail.ru
Loginova, Evgeniya	Moscow State University Student Moscow, Russia	Eureka.mail@gmail.com
Lygina, Talia	Institute of Geology of Industrial Minerals Professor Kazan, Russia	+7 (843) 236-5373 talialygina@gmail.com
Madsen, Michael	Aalborg University Esbjerg Ph.D. student Esbjerg, Denmark	+45 65504165 mima@bio.sdu.dk
Marini, Federico	University of Rome Post. Doc. Rome, Italy	+39 06 49913856 fmmonet@hotmail.com
Marks, Christofer	Academic Scientist Richmond, USA	laurelandchris@verizon.net
Matero, Sanni	University of Kuopio Ph.D. student Kuopio, Finland	Sanni.Matero@uku.fi

Mikhailov, Evgeniy	Samara State Technical University Ph.D. student Samara, Russia	evgeniystat@gmail.com
Milosz, Marek	Lublin University of Technology Doctor Lublin, Poland	+48 601334934 marekm@pluton.pol.lublin.pl
Milosz, Elzbieta	High School of Enterprise and Administration, Lublin Doctor Lublin, Poland	+48 601150035 milosz@antenor.pol.lublin.pl
Morzhukhina, Svetlana	State university “Dubna” Head of chemistry department Dubna, Russia	8 (49621) 22803 msv@uni-dubna.ru
Osokin, Alexander	Metall Company LLP Project Manager Moscow, Russia	osokin@mtzd.ru
Osokin, Anton	Moscow State University Student Moscow, Russia	
Osokina, Tatiana	Moscow, Russia	
Paakkunainen, Maaret	Lappeenranta University of Technology Ph.D. student Lappeenranta, Finland	maaret.paakkunainen@lut.fi
Paul, Sylvia	University of South Africa Professor Pretoria, South Africa	+27 124298010 paulso@unisa.ac.za
Pavlov, Sergey	Institute of Geochemistry SB RAS Researcher Irkutsk, Russia	+7 (3952) 425837 roland@bk.ru
Pomerantsev, Alexey	Institute of Chemical Physics RAS Leading Researcher Moscow, Russia	+7 (495) 9397483 forecast@chph.ras.ru
Popov, Alexander	ESRF Scientist Grenoble, France	+33 476882633 apopov@esrf.fr

Popova, Lidiya	Accompanying person Grenoble, France	
Potapov, Alexey	Institute of Spectroscopy of RAS Researcher Troitsk, Russia	+7 (495) 3340239 potapov@isan.troitsk.ru
Prelov, Vladimir	Mechanical Engineering Research Institute of RAS Researcher Moscow, Russia	+7 (499) 1357771 prelov@iitp.ru
Rakhimova, Nailya	Kazan State University of Architecture and Engineering Associate professor Kazan, Russia	+7 (8432) 5104727 rahimova.07@list.ru
Reinikainen, Satu-Pia	Lappeenranta University of Technology Doctor Lappeenranta, Finland	+358 407301851 satu-pia.reinikainen@lut.fi
Reshetilov, Anatoly	Institute of Biochemistry and Physiology of Microorganisms Head of Laboratory Pushchino, Russia	+7 (495) 9563370 anatol@ibpm.pushchino.ru
Reshetilova, Tatiana	Institute of Biochemistry and Physiology of Microorganisms Scientific Secretary Pushchino, Russia	+7 (495) 9563370 rta@ibpm.pushchino.ru
Rodionova, Oxana	Institute of Chemical Physics RAS Senior researcher Moscow, Russia	+7 (495) 9397483 rcs@chph.ras.ru
Rudnitskaya, Alisa	Saint-Petersburg State University Researcher Saint-Petersburg, Russia	alisa.rudnitskaya@gmail.com
Rusinov, Leon	Saint-Petersburg technological institute Professor Saint-Petersburg, Russia	+7 (812) 4957453 rusinov@ztel.spb.ru
Rutledge, Douglas	UMR INRA/AgroParisTech Professor Paris, France	+33 144081648 rutledge@agroparistech.fr

Ryumina, Natalia	Samara State Technical University Ph.D. student Samara, Russia	nryumina@gmail.com
Samigullin, Rais	Institute of Geology of Industrial Minerals Researcher Kazan, Russia	rais_s@rambler.ru
Sarbu, Costel	Babes-Bolyai University Associate professor Cluj-Napoca, Romania	+40 264593833 costelsrb@yahoo.co.uk
Savostyanov, Dmitry	Institute of Numerical Mathematics RAS Researcher Moscow, Russia	+7 (495) 9383764 drraug@gmail.com
Shabanova, Elena	Institute of Geochemistry SB RAS Researcher Irkutsk, Russia	+7 (3952) 425837 shev@igc.irk.ru
Shastova, Nataliya	State university "Dubna" Student Dubna, Russia	+7 (4962) 22803
Shireeva, Dilyara	Kazan State University of Architecture and Engineering Senior Lecture Kazan, Russia	+7 (843) 2728009 sulejmanov@ksaba.ru
Sokolova, Svetlana	Saint-Petersburg State University of Aerospace Instrument Making Professor Saint-Petersburg, Russia	+7 (812) 3931657 sokolova_sv@mail.ru
Sokovikov, Yaroslav	SchelTec AG Support specialist Moscow, Russia	+7 (495) 9358888 Yaroslav.Sokovikov@scheltec.ru
Starovoitova, Irina	Kazan State University of Architecture and Engineering Ph.D. student Kazan, Russia	irina-starovoitova@yandex.ru

Sulejmanov, Alfred	Kazan State University of Architecture and Engineering Vice Rector Kazan, Russia	+7 (8432) 728009 alfred-sulejmanov@yandex.ru
Valiev, Ashraf	Institute of Geology of Industrial Minerals Researcher Kazan, Russia	ashik_ashik@yandex.ru
Vasilyeva, Irina	Institute of Geochemistry SB RAS Researcher Irkutsk, Russia	+7 (3952) 425837 vasira@igc.irk.ru
Vdovin, Evgeniy	Kazan State University of Architecture and Engineering Associate professor Kazan, Russia	+7 (8432) 5104677 vdovin007@mail.ru
Vlasova, Irina	Omsk State University Associate professor Omsk, Russia	+7 (3812) 642485 vlasova@univer.omsk.su
Zhilin, Sergey	Altai State University Associate professor Barnaul, Russia	sergei@asu.ru
Zilberg, Rufina	Bashkir State University Ph.D. student Ufa, Russia	artsid@ufamail.ru