Fourteenth Winter Symposium on Chemometrics

Modern Methods of Data Analysis

Armenia, Tsaghkadzor, February 26–March 1, 2024

SCIENTIFIC COMMITTEE

Prof. Alejandro C. Olivieri, Department of Analytical Chemistry, Faculty of Biochemical and Pharmaceutical Sciences, National University of Rosario, Argentina

Dr. Paolo Oliveri, Analytical Chemistry and Chemometrics Department of Pharmacy (DIFAR) University of Genova, Italy

Dr. Frederico Luis Felipe Soares, Federal University of Paraná, Brazil

Dr. Hadi Parastar, Department of Chemistry, Sharif University of Technology Tehran, Iran

ORGANIZING COMMITTEE

Co-chairmen
Prof. Dmitry Kirsanov, Institute of Chemistry, St. Petersburg University, Russia

Prof. Levon Agasy Tavadyan, National Academy of Sciences of Armenia, Armenia

Secretary
Dr. Anastasiia Surkova, Samara State Technical University, Samara, Russia

http://wsc14.com   e-mail: wsc14info@gmail.com

Thanks

The WSC-14 organizers and participants wish to express their greatest appreciation to the following conference sponsors for their valuable economic and friendly help:

Finally, we are grateful to all the WSC-11 attendees, lecturers, accompanying persons, and visitors for their interest in the conference.

Useful information

The conference will be held in Ararat Resort Hotel Tsaghkadzor (Armenia, 2310, Kotayk Marz, Tsaghkadzor, Mher Mkrtchyan 54). The hotel is 4.1 km from Tsaghkadzor City Centre. Capital Yerevan is in 54 km, and Zvartnots International Airport in 67 km from the hotel. Besides the official conference transfer, you can reach the capital by The Yerevan Airport Express buses N201 departing every half hour from 7 am until 10 pm (ticket costs 300 AMD (under $1), then take a taxi to Tsaghkadzor. There is no Metro Station in the Airport. If not the Express Bus, you will need to take a taxi (Yandex Taxi, ggTaxi through mobile applications or have prearranged transportation (20,000 AMD for 1 Pax by Sedan, 25,000 AMD for 6-7 Paxes by Viano). Ararat Resort Tsaghkadzor complex is located in the most picturesque corner of Tsaghkadzor particularly in the heart of the forest of pine trees, within a 15-minute walk to the Monastery Kecharis, and 10-minute drive to Tsaghkadzor Ropeway (for more details https://www.ropeway.am/index-en.php). Hotel provides free transfer to the Ropeway at 10:00/11:00/12:00/13:00/14:00/15:00 and vice versa at 12:20/13:20/14:20/15:20/16:20/17:20 (should be reserved with the reception in advance)

Conference and activities

Conference sessions will be held in the Conference Hall (3rd Floor). The official conference language is English. Everyone is encouraged to have his/her badge attached, both during the symposium sessions and social activities.

In their free time, participants can try various winter activities: skiing, skating, snowboarding (rental services are available at the place). In addition, SPA (swimming pool, gym (10:00 – 22:00)), sauna, jacuzzi (10:00 – 12:00 for women, 14:00 – 20:00 for men and women, after 20:00 1 hour 4000 AMD for a pax) and billiard (1 hour 4000 AMD) are also available. Don't forget your swimming suit!

Meals

All meals will be served at "Ararat" restaurant (2nd floor). The banquet will take place at Panorama Restaurant Tsaghkadzor (10-minute drive from the Hotel. The Transfer will be provided).

Scores & Loadings

Traditional "Scores and Loadings" meetings will be held in Lobby Bar (3rd floor).

Communication

The main cellular networks Ucom & Vivacell have proper coverage around the hotel. Wi-Fi Connection throughout the hotel is available. Password: araratresort

Money

Armenia's national currency is the Armenian Dram (AMD). The trade in foreign currency is prohibited. You can exchange most common currencies in the Airport, Yerevan and Tsaghkadzor's banks, supermarkets and exchange shops throughout the cities. Hotel and other establishments accept credit cards. There is also an ATM machine in the hotel area.

Excursion

Friday morning all participants are invited to have an excursion to Lake Sevan + Sevanavank Monastery (built in 874 AD) and enjoy walking City Tour in Yerevan. Lake Sevan, called "The Jewel of Armenia", is the biggest freshwater lake in the Caucasus region. It is 1900 meters above sea level. Lake Sevan is considered to be a top visited destination for the tourists. During Yerevan City Tour you will see and enjoy the impressive architecture and cultural heritage of the city, which is more than 2800 years old (older than Rome) (duration of the tour nearly 6 hours).

Useful mobile Apps

Hike Armenia, ggTaxi, Yandex Taxi, Rate.am, Visit Yerevan, Tomsarkgh.am

Useful Phone Numbers

Anastasiia Surkova, *conference secretary*   +7 (927) 745-49-32
Dmitry Kirsanov, *organizing committee*   +7 (921) 333-12-46
Mir photonics LLC   + 374 96 522 863
Ararat Resort Reception   +374 60 388 388, +374 55 504 504

Conference program

Monday, February 26, 2024

14:00–15:00	**Registration and opening**
Session 1	**Chair: Janos Elek**
15:00–15:30	**T1** *Yuri Kalambet* Fuzzy Locally Optimal Polynomial Filter – why Optimal, Fuzzy and Polynomial?
15:30–16:00	**T2** *Anastasiia Shuba* Multidimensional patterns of parameters of gas sensors with polycomposite coatings for assessing microbiological parameters of raw milk
16:00–16:30	**T3** *Dmitry Kirsanov* Advancing X-ray fluorescence analysis with chemometric tools
16:30–17:00	**Coffee break**
17:00–19:00	**Free time**
19:00–20:00	**Dinner**
20:00–00:00	**Scores & Loadings**

Tuesday, February 27, 2024

08:00–09:00	**Breakfast**
Session 2	**Chair: Alexey Pomerantsev**
09:00–10:00	**L1** *Stefan Tsakovski* Chemometrics in soil quality assessment
10:00–10:30	**T4** *Ekaterina Boichenko* Near-infrared spectroscopy and chemometrics: a promising combination for real-time and nondestructive classification of urinary stones
10:30–11:00	**T5** *Andrey Stavrianidi* Novel QSRR and deep learning approaches for retention time modeling in RP and HILIC conditions
11:00–11:30	**Coffee break**
11:30–14:00	**Free time & skiing time**
Session 3	**Chair: Stefan Tsakovski**
14:00–15:00	**L2** *Alexey Pomerantsev* New concepts and methods originated from SIMCA
15:00–15:30	**T6** *Janos Elek* A New Hope in the Analysis of Novel Foods in Rat Diet for Toxicology Studies: Transitioning from NIR to Reflectance UV Spectroscopy
15:30–16:00	**T7** *Natalia A. Burmistrova* Multivariate Data Analysis of 2D Fluorescence Spectra in Monitoring of Imprinted Proteins Purification
16:00–16:30	**Coffee break**
Session 4	**Chair: Oxana Rodionova**
16:30–17:00	**T8** *Andrey Samokhin* Comparison and optimization of mass spectral library search algorithms
17:00–17:30	**T9** *Anastasiia Surkova* Aquaphotomics and NIR spectroscopy: from basics to practical application
17:30–18:30	**Workshop: DD-SIMCA for web**
18:30–19:00	**Free time**
19:00–20:00	**Dinner**

20:00–00:00	Scores & Loadings

Wednesday, February 28, 2024

08:00–09:00	Breakfast
Session 6	Chair: Xihui Bian
09:00–10:00	**L3** *Hadi Parastar* Integration of Hyperspectral Imaging with Chemometrics to solve food authentication challenges
10:00–10:30	**T10** *Andrey Bogomolov* Spectroscopy without spectra?!
10:30–11:00	**T11** *Ekaterina Yuskina* Chemometric processing of the data from sensor device for contactless chemical analysis based on high-frequency inductance coil
11:00–11:30	Coffee break
11:30–14:00	Free time & skiing time
Session 7	Chair: Andrey Bogomolov
14:00–15:00	**L4** *Oxana Rodionova* Chemometric support for precision and personalized medicine
15:00–15:30	**T12** *Mate Csontos* Needle in a heystack
15:30–16:00	**T13** *Nikolai Sushkov* Laser-based analytical techniques and multiblock chemometrics for the analysis of marine zooplankton
16:00–16:30	Coffee break
Session 8	Chair: Douglas N. Rutledge
16:30–19:00	Poster Session
19:00–20:00	Dinner
20:00–00:00	Scores & Loadings

Thursday, February 29, 2024

08:00–09:00	**Breakfast**
Session 9	**Chair: Hadi Parastar**
09:00–10:00	**L5** *Douglas N. Rutledge* Chemometrics and Agriculture : a long lasting collaboration
10:00–10:30	**T14** *Mikhail Khrisanfov* A new approach to finding potentially erroneous entries in a database using machine learning
10:30–11:00	**T15** *Alexey Tretyakov* Feasibility study of X-ray fluorescence spectra and chemometrics in dating XVIII century glass samples
11:00–11:30	**Coffee break**
Session 10	**Chair: Dmitry Kirsanov**
11:30–12:30	**L6** *Xihui Bian* Advanced chemometric methods for spectral analysis of complex samples
12:30–13:00	**T16** *Dmitriy Matyushin* Non-target analysis by gas chromatography-mass spectrometry using machine learning and deep learning
13:00–13:30	**Closing**
13:30–19:00	**Free time & skiing time**
19:00–00:00	**Banquet**

Friday, March 1, 2024

08:00–09:00	**Breakfast**
09:00–10:00	**Check out**
10:00–15:00	**Bus to Yerevan and excursion**

Abstracts

L01. Chemometrics in soil quality assessment

S. Tsakovski, V. Simeonov

Group of Chemometrics and Environmetrics, Chair of Analytical Chemistry, Faculty of Chemistry and Pharmacy, Sofia University "St. Kliment Ohridski", Sofia, Bulgaria

Soil quality can be defined as the ability of the soil to perform functions for its intended use. Soil quality evaluations are based on soil indicator (attributes) measurements, which reflect the inherent soil properties. Human management and natural disturbances can lead to significant changes in soil properties, which require a soil monitoring network, including the proper selection of an indicator data set related to reliable soil quality assessment.

Along with the selection and assessment of an optimal set of soil indicators, such a complex and multivariate task as soil quality assessment requires adequate data treatment. The application of appropriate multivariate statistical methods, like Cluster Analysis (CA), Principal Components Analysis (PCA), Sel-Organizing Maps (SOM), Partial Least Squares – Discrimination Analysis (PLS-DA) etc., can not only reveal hidden interactions between different soil indicators but also identify the factors affecting soil quality, including natural processes and anthropogenic pressures.

The aim of the current lecture is to present different ways in which chemometric approaches could be applied in soil quality assessment, depending on the available set of soil indicators and the goal of the analysis. Special attention will be paid to the application of chemometric approaches in studies assessing the bioaccessibility and bioavailability of potential toxic elements. The lecture will present examples concerning:

- pollution soil assessment in mining and smelter regions;
- soil quality assessment based on national soil quality monitoring network data.

L02. New concepts and methods originated from SIMCA

A.L. Pomerantsev, O.Ye. Rodionova

Federal Research Center for Chemical Physics RAS, Moscow, Russia

SIMCA is a well-known method that uses two characteristics: score and orthogonal distances (SD and OD) to develop a decision rule in the frame of class modelling, also known as one-class classification [1]. Now it is in great demand, mainly for solving authentication problems in different areas. In this lecture, we

will not discuss SIMCA itself, but, on the contrary, we will turn to several new methods that directly follow from this concept.

The first issue to discuss is the concept of Cumulative Analytical Signal (CAS), which develops the idea of characteristic distances and introduces new characteristics as the sum of old ones [2]. CAS gives birth to a series of new methods (SIMCA grandchildren) such as outlier detection [4], representative subset selection [5], limit of detection estimation [3], multi-block data fusion [6], etc.

The second unseal application of SIMCA methodology is the precise and personalized medicine (PPM) that solves the problems of similarity/novelty, evaluation of curing results, selection of individual treatment, etc. [7]. Details of the PPM approach are provided in the lecture of Prof. Rodionova.

Another SIMCA generation is Procrustes Cross-Validation (PCV) [8], which gives rise to a new concept of validation based on a pseudo-test set generated from the results of conventional cross-validation and the CAS rules. The main advantage is that the PCV set exists in reality as a set of numbers, and not as a virtual procedure that provides us with the final results of modelling only [9].

In the lecture, we discuss all these methods, focusing on ideas and concepts, and not on the mathematical apparatus, which can be found in the references.

References

[1] A.L. Pomerantsev, O.Ye. Rodionova, *J. Chemom.* **34** (2020) e3250.
[2] A.L. Pomerantsev, O.Ye Rodionova, *Trends. Anal. Chem*. **143** (2021) 116372.
[3] A.L. Pomerantsev, D.N. Vtyurina, et al, *Microchem. J.* **195** (2023) 109490.
[4] O.Ye. Rodionova, A.L. Pomerantsev, *Anal. Chem.* 92 (2020) 2656–2664.
[5] A.L. Pomerantsev, O.Ye. Rodionova, *Microchem. J.* **190** (2023) 108654.
[6] O. Rodionova, A. Pomerantsev, *Anal. Chim. Acta*. **1265** (2023) 341328.
[7] O. Rodionova, N. Kurysheva, et al, *Anal. Chim. Acta*. **1250** (2023) 340958.
[8] S. Kucheryavskiy, S. Zhilin, et al, *Anal. Chem.* **92** (2020) 11842–11850.
[9] S. Kucheryavskiy, O.Rodionova, et al, *Anal. Chim. Acta*. **1255** (2023) 341096.

L03. Integration of Hyperspectral Imaging with Chemometrics to Solve Food Authentication Challenges

H. Parastar

Department of Chemistry, Sharif University of Technology, Tehran, Iran

Hyperspectral imaging (HSI) is a recently developed technique which is a rapid, reagent less and non-destructive technique with different applications in

various fields such as chemistry, pharmaceutical science, remote sensing, agricultural science, forensic science and also in food science [1]. By capturing detailed information across numerous wavelengths, HSI can detect subtle differences or irregularities in food composition, quality and authenticity. HSI has been already utilized as a potent technique for food authenticity in different matrices [2].

Chemometrics plays a crucial role in processing the vast amount of spectral information obtained from HSI. Multivariate chemometric techniques have been proposed in the last decade for the classification and regression of three-way HSI data [3]. In this regard, principal component analysis (PCA) is the best option for dimension reduction and visualizing the pattern of these data sets. Additionally, multivariate curve resolution-alternating least squares (MCR-ALS) and independent component analysis (ICA) have been proposed for multivariate resolution and to extract the pure spatial and spectral profiles of existing components in HSI datasets. Furthermore, other methods such as support vector machine (SVM), extreme learning machine (ELM), partial least squares regression (PLSR) and partial least squares-discriminant analysis (PLS-DA) have also been used in previous studies for identification and quantification of food fraud.

Due to the vast applicability of chemometrics in food authentication, the objective of this work is to introduce the idea of chemometrics for HSI data analysis in the context of food authentication in a simple way using illustrative examples.

It is concluded that chemometrics combined with HSI can open a new window in solving food authentication problems in a fast, non-destructive and reliable way.

References

[1] Y. Liu, H. Pu, D. Sun, *Trends Food Sci. Technol.* **69** (2017) 25–35.
[2] F.S. Hashemi-Nasab, H. Parastar, *Food Chem.* **393** (2022) 133450.
[3] A. de Juan, R. Tauler, R. Dyson, C. Marcolli, M. Rault, M. Maeder, *Trends Anal. Chem.* **23** (2004) 70–79.

L04. Chemometric support for precision and personalized medicine

O.Ye. Rodionova, A.L. Pomerantsev

Federal Research Center for Chemical Physics RAS, Moscow, Russia

The modern approach for the choice of medical treatment is personalized medicine, which involves the use of genetic or other biomarker information to

make a decision about the treatment of a particular patient. A related term is "precision medicine", which expands on patient information to include anatomical, environmental, and lifestyle factors. We view personalized or precision medicine as a typical problem of multivariate data analysis. The following aspects are discussed.

Novelty and similarity are closely related concepts and they are certainly worthy of exploring given their numerous applications in medicine and healthcare. Several important principles are considered. The first one concerns proper complexity of the model. The second point is interpretability, the important issue that is often overlooked. DD-SIMCA as a probabilistic method, which is suitable for analysis of small data sets is used for solving novelty-similarity problem.

A new concept of a generalized characteristic of the success of a treatment as proximity of subjects to the Healthy group is presented. For this purpose full distance calculated in the DD-SIMCA model based on Healthy group as a target class is employed.

The methodology for the personalized choice of an effective method of medical intervention is considered. It is based on principal component analysis and variable selection.

Three real-world examples are presented: (1) the comparison of morphological and spectral detection of abnormal spermatozoa [1]; (2) analysis of two groups of patients who participate in the evaluation of the effectiveness of the treatment of ophthalmology disease [2]; medical diagnostics of coronary heart disease [3].

References

[1] R.V. Nazarenko, A.V. Irzhak, A.L. Pomerantsev, O.Ye. Rodionova, *Chemom. Intell. Lab. Syst.* **182** (2018) 172–179.

[2] N.I. Kurysheva, O.Ye. Rodionova, A.L. Pomerantsev, G.A. Sharova, *Biomed. Signal Process. Control.* **85** (2023) 104884

[3] M. Kelly, R. Longjohn, K. Nottingham, The UCI Machine Learning Repository, https://archive.ics.uci.edu

L05. Chemometrics and Agriculture: a long lasting collaboration

D.N. Rutledge [1,2]

[1] *Muséum National d'Histoire Naturelle, Paris, France*

[2] *Faculté de Pharmacie, Université Paris-Saclay, Orsay, France*

The development of chemometrics could be seen as an example of Darwinian co-evolution where instrumentation and chemometric methods evolve together under the selective pressure of industrial applications.

The term 'chemometrics' was coined by Svante Wold (Professor of organic chemistry at Umeå University, Sweden) in a grant application in 1971 [1].

However, one could consider that the first application of a sort of proto-chemometrics was presented by Karl Norris in 1963 at the International Symposium on Humidity and Moisture [2] where he demonstrated the direct determination of the moisture content of grains and seeds using Near Infrared Spectroscopy.

Since then, the instrumentation used in agricultural quality control has evolved enormously from simple filter-based NIR spectrometers to Fourier-Transform NIR and MIR, Raman, GC- and LC-Mass, Hyperspectral Imaging, 3D Fluorescence and even NMR.

In parallel, chemometric methods have evolved from simple linear regressions to multivariate methods like PCA, ICA, MCR-ALS and PLS, as well as multi-way and multi-block methods.

This co-evolution has occurred because, as Lutgarde Buydens said: "Chemometricians verify the validity of their models" [3]. They must do this because their models are of value to industry.

Examples of agricultural applications will be presented.

References

[1] S. Wold, *Chemom. Intell. Lab. Syst.* **30**(1) (1995) 109–115.

[2] K.H. Norris, J.R. Hart, *Proceedings of the 1963 International Symposium on Humidity and Moisture*, Reinhold Publ Corp, New York, N. Y., "Principles and Methods of Measuring Moisture in Liquids and Solids," **4** (1965), 19–25.

[3] https://theanalyticalscientist.com/techniques-tools/gurus-of-chemometrics.

L06. Advanced chemometric methods for spectral analysis of complex samples

X. Bian [1,2]*, P. Mpango [1], H. Bao [1], R. Wang [1], P. Liu [1]

[1]School of Chemical Engineering and Technology, Tiangong University, Tianjin, P. R. China

[2]NMPA Key Laboratory for Technology Research and Evaluation of Drug Products, Shandong University, Jinan, China

Spectral analysis combined with chemometrics has drawn increasing attention for rapid and nondestructive analysis of complex samples. Outlier detection, data grouping, spectral preprocessing, variable selection, modeling methods are the main contents of chemometric research. Our group focuses on ensemble modeling, representative sample selection, swarm intelligence optimization for variable selection and spectral denoising in recent years.

For ensemble modeling, empirical mode decomposition (EMD) [1] and variational mode decomposition (VMD) weight multi-scale SVR were proposed for quantitative analysis of edible blend oil and Chinese medicine, respectively. In addition, double ensemble models based on Lasso and Monte Carlo resampling were proposed for quantitative analysis of corn, edible blend and blood samples. VMD unfolded PLS was proposed for ultraviolet-visible spectral analysis of edible oil blend, fuel oil and aqueous samples. These methods all demonstrated superiorities in accuracy and stability compared with single models.

For spectral denoising, HHT combined with F-test was proposed firstly for denoising of X-ray diffraction (XRD) and X-ray photoelectron spectrum (XPS). EMD combined t-test was then developed for XRD spectral denoising of catalysts. For solving the end effect and mode mixing of EMD, VMD was introduced for Raman spectral denoising. Finally, a piecewise mirror extension local mean decomposition method was proposed for near-infrared spectral denoising with uneven noise [2]. The feasibility of the proposed method is validated by visualized, signal-to-noise ratio and modeling results.

For variable selection, swarm intelligence (SI) algorithms have attracted more and more attention due to their simple structure, efficiency and powerful ability. Our group introduced grey wolf optimizer (GWO), whale optimization algorithm (WOA) and butterfly optimization algorithm (BOA) for spectral variable selection of complex samples [3]. We also summarized 94 SI algorithms and classified them into six classes based on their behavior.

References

[1] X.Y. Wu, X.H. Bian, E. Lin, et al. *Food Chem.* **342** (2021) 128245.

[2] M.X. Ling, X.H. Bian, S.S. Wang, et al. *Chemom. Intell. Lab. Syst.* **230** (2022) 10465.

[3] X.H. Bian, R.L. Zhang, P. Liu, et al. *Spectrochim. Acta A.* **284** (2023) 121788.

T01. Fuzzy Locally Optimal Polynomial Filter – why Optimal, Fuzzy and Polynomial?

Y. Kalambet

Ampersand Ltd., Moscow, Russia

Locally optimal filter (LOF) concept is based on minimization of confidence interval of the approximation. This filter makes multitude of local approximations of the data array using different models, calculating both approximated values and their confidence intervals, and selects an approximation with the lowest confidence interval. LOF provides good approximation of the baseline signal without disturbing peak shape. In the case of approximations, based on polynomials, filter is named Locally Optimal Polynomial (LOPF). Fuzzy locally optimal polynomial filter (FLOPF) approximation improves LOPF, accounting for a multitude of estimates and confidence interval of every estimate made by LOPF. FLOPF-filtered signal is smoother than that by LOPF, provides higher S/N values in extreme cases of low signal, has computational complexity comparable to LOPF and higher robustness. All models, used for approximation, are using a limited set of measurements in the vicinity of approximated point, so fine tuning of algorithm parameters included investigation of smoothing accuracy with respect to parameters of model construction. Smoothing results for chromatographic peaks are compared to smoothing by popular linear filters including optimal Matched filter, Whittaker smoother and LOPF. Residual relative random error was estimated for area and height, FLOPF algorithm managed to outperform all other smoothers for area; FLOPF relative random error of height was behind that of Matched filter, but without 40% systematic distortion intrinsic to the Matched filter.

T02. Multidimensional patterns of parameters of gas sensors with polycomposite coatings for assessing microbiological parameters of raw milk

A. Shuba [1], T. Kuchmenko [1], E. Anokhina [2], E. Bogdanova [3], R. Umarkhanov [1]

[1]Deparment Physical and Analytical Chemistry, Voronezh State University of Engineering Technologies, Voronezh, Russia

[2]Laboratory of metagenomic and food biotechnology, Voronezh State University of Engineering Technologies, Voronezh, Russia

[3]*Department of Food Production from Animal Raw Materials, Voronezh State University of Engineering Technologies, Voronezh, Russia*

Currently, the issue of microbiological safety of food products using simple and inexpensive methods is relevant. The use of methods that measure changes in the optical, physicochemical properties or changes in the gas phase makes it possible to obtain a large array of data in which to identify variables associated with changes in microbiological indicators in the samples. When using more complex detection devices, the processing of the received data inevitably becomes more complicated. It was previously shown that changes in the gas phase above a sample of complex composition are associated, among other things, with changes in microbiological parameters in the sample. Firstly, the indicator of such changes is the qualitative composition of the gas phase. Previously, it was also shown that the measurement of the qualitative composition of gas mixtures can be carried out using calculated parameters (sorption efficiency parameters) of an array of sensors with varying, but not absolute, selectivity to vapors of substances. It has been shown that using sensors with composite coatings consisting of several sorbents, it is possible to obtain information about the composition of a gas mixture comparable with an array of several sensors [1]. Additional consideration of the kinetic sorption of volatile compounds on a composite coating allows one to determine parameters that connecting with the qualitative composition of the gas mixture [2-3]. At the same time, a more detailed study of the output curves of the sensors to calculate additional parameters that increase the reliability prediction is a relevant task.

The purpose of the work is to determine significant indicators at the output curves of gas sensors with composite coatings to calculate additional parameters that increase the reliability of the assessment of microbiological indicators of raw milk samples.

We studied the gas composition of a raw milk sample (no more than 4 hours after milking) with an array of 8 piezoelectric sensors with composite coatings, their physicochemical and microbiological indicators in accordance with GOST. The use of projection approaches to select the most significant indicators in the kinetic output curves of sensors allows us to obtain appropriate calculated parameters that reduce the prediction error in the regression models for the total microbiological contamination of milk. It is shown that additional calculated parameters also increase the accuracy of models for predicting the content of yeast and moulds in raw milk. The combination of the proposed

parameters and sorption efficiency parameters leads to a decrease in the predicting error, both in regression and discriminant models. Taking into account the results of predicting microbiological indicators, multidimensional patterns were constructed from the entire set of parameters, with the help of which it is possible to differentiate raw milk samples by the presence of microorganisms, yeasts and moulds above the norm.

Acknowledgment

References

[1] A. Shuba, T. Kuchmenko, R. Umarkhanov, *Sensors* **22** (2022) 8529.

[2] A. Shuba, E. Anokhina, R. Umarkhanov, E. Bogdanova, I. Burakova, *Eng. Proc.* **48** (2023) 21.

[3] A. Shuba, E. Anokhina, R. Umarkhanov, E. Bogdanova, *Eng. Proc.* **48** (2023) 34.

T03. Advancing X-ray fluorescence analysis with chemometric tools

V. Panchuk [1,2], V. Semenov [1,2], D. Kirsanov [1]

[1] *Institute of Chemistry, St. Petersburg University, St. Petersburg, Russia*

[2] *Institute for Analytical Instrumentation RAS, St. Petersburg, Russia*

X-ray fluorescence analysis (XRF) is a popular method, widely employed in routine laboratory practice for the qualitative and quantitative determination of elements in a wide variety of samples in numerous fields, such as materials science, geology, mineralogy, environmental studies, archaeology and others. The popularity of the method is due to its advantages: non-destructiveness, possibility to simultaneously determine a large number of elements in samples with different aggregate states, minimal sample preparation. Further development of the XRF method is associated with three main directions: 1) improvement of instrumental base (hardware); 2) development of new analytical techniques; 3) application of new methods of spectral data processing. The last direction seems to be the most promising, as it does not require significant investments and does not lead to increased labor costs in the analysis. While chemometric tools have already gained a broad popularity in optical spectroscopy, chromatography, mass spectrometry, in XRF they are still applied rather limitedly. Meanwhile, such approaches allow both to solve the tasks atypical for XRF (e.g. determination of light elements and physical characteristics of samples), as well as increase the accuracy of quantitative

elemental analysis (e.g., in objects with strong matrix effects) in traditional XRF applications. The talk will present the recent developments and advantages of applying chemometric tools to XRF data [1-5] and will highlight the prospects of further development of this research direction.

Acknowledgment

References

[1] Z. Selivanovs, V. Panchuk, D. Kirsanov, *Appl. Sciences* **13** (2023) 5415.

[2] V. Panchuk, Y. Petrov, V. Semenov, D. Kirsanov, *Anal. Chim. Acta* **1239** (2023) 340694.

[3] S. Aidene, M. Khaydukova, S. Savinov, V. Semenov, D. Kirsanov, V. Panchuk, *Spectrochim. Acta B* **193** (2022) 106452.

[4] S. Aidene, M. Khaydukova, G. Pashkova, V. Chubarov, S. Savinov, V. Semenov, D. Kirsanov, V. Panchuk, *Spectrochim. Acta B* **185** (2021) 106310.

[5] S. Aidene, V. Semenov, D. Kirsanov, D. Kirsanov, V. Panchuk, *Measurement* **172** (2021) 108888.

T04. Near-infrared spectroscopy and chemometrics: a promising combination for real-time and nondestructive classification of urinary stones

E. Boichenko[1], A. Reznichenko [1], M. Paronnikov [2], D. Kirsanov [1]

[1] Department of Analytical Chemistry, Institute of Chemistry, Saint Petersburg State University, Saint Petersburg, Russia

[2] Department of Urology, S. M. Kirov Military Medical Academy, Saint Petersburg, Russia

Growth of a solid concretion of mineral salts in the urinary tract is a disease called urolithiasis. This pathology often requires surgery, typically a laser destruction of the stone with X-ray guidance, or a lithotripsy. Though this kind of surgery is minimally invasive and generally well-tolerated by a patient, the absence of a real-time intraoperative method to determine chemical composition of a stone before its removal hinders further development of treatment for urolithiasis. According to the guidelines of European Association of Urology, the composition of a urinary calculus is one of the factors, which define tactics in the surgical and conservative treatment: settings of the laser beam during the lithotripsy, supporting diet for prevention of disease

recurrence, etc. Currently, qualitative analysis of a urinary stone is performed by X-ray phase methods post-operatively (and often in commercial laboratories), which causes a significant delay in making clinical decisions.

Near-infrared spectroscopy (NIRS) is a promising method for classification of urinary stones by their main component: oxalates, urates, apatites, brushites, etc., since the vibrational overtones of these molecules can be observed in the NIR region. NIRS is also suitable for real-time applications, because sample preparation is not required and a wide range of portable spectrometers and optical fibers is available. Certainly, NIRS is well-known for broad overlapping peaks and high absorption of water, which should be taken into account due to the permanent presence of saline (0,9% NaCl solution) in the operating field. However, multivariate methods of spectral data analysis allow extracting clinically relevant information from the NIR spectra of urinary stones, measured under different conditions, as shown in our preliminary feasibility study [1].

In this work, the results of analysis of 172 real urinary calculi, removed by laser lithotripsy, are presented. The measurements were performed using a portable NIR spectrometer (Avantes, the Netherlands) within 939–1799 nm and a fiber optic probe 1.8 mm in diameter (Endoprobe 7NIR + 1NIR, Optofiber, Russia). Each stone was measured in three different points with further averaging of the spectra. The experiment consisted of several steps: firstly, the NIR spectra were registered under ambient atmosphere (on air), then the stones were soaked in saline or in real urine for two hours before NIRS measurement. The comparison of several combinations of data preprocessing and multivariate analysis (linear discriminant analysis, support vector machines with various kernel functions), as well as the results of intraoperative approbation of the proposed method, are also reported.

Acknowledgment

References

[1] E. Boichenko, M. Paronnikov, D. Kirsanov, *Eng. Proc*. **48**(1) (2023) 64.

T05. Novel QSRR and deep learning approaches for retention time modeling in RP and HILIC conditions

A. Stavrianidi [1,2], D. Obradović [3], E. Fedorova [1], O. Shpigun [2], S. Lazović [3], A. Buryak [1]

¹A.N. Frumkin Institute of Physical Chemistry and Electrochemistry, Russian Academy of Sciences, Moscow, Russia
²Chemistry Department, Lomonosov Moscow State University, Moscow, Russia
³Institute of Physics Belgrade, National Institute of the Republic of Serbia, Belgrade, Serbia

The lack of analytical standards is a bottleneck for comprehensive evaluation and quality control in pharmaceutical industry and related fields. Pharmaceutical substances are complex mixtures containing structurally related impurities. Due to the variety of chromatographic phases, finding the optimal separation conditions can also became a costly challenge. Therefore, compound retention modeling can be a way to the better understanding of chromatographic retention mechanisms, which helps to predict and compare separation parameters and selectivity for such systems. Retention time (RT) prediction also provides additional support for peak annotation in LC-MS metabolomics and pharmacodynamics studies where individual standards are not always available.

There are numerous ways to construct such models in Quantitative Structure Retention Relationship (QSRR) and there are a lot of available web-applications and software which allow computing the basic physico-chemical properties and other molecular features and performing machine learning (ML). In our studies, we employed nine different descriptor calculation tools combined with different feature selection methods (genetic algorithm (GA), stepwise, Boruta) and various ML modeling approaches (support vector machine (SVM), multiple linear regression (MLR), random forest (RF), XGBoost) in order to provide a reliable molecular interpretation of chromatographic retention behavior in different conditions (RP, HILIC, non-aqueous HILIC and supercritical fluid chromatography (SFC). GA showed the most suitable performance as descriptor selection method, while in most cases SVM modeling gave satisfactory predictive characteristics. ChemoPy/RF approach demonstrated the best elution order predictivity.

We also conducted a comparative study on the use of various NN architectures in combination with various features (fingerprints and SMILES strings) for RT prediction in RP LC using the largest data set of standard compounds (METLIN SMRT). Extended-connectivity fingerprints (ECFPs) were shown as the most appropriate input features for processing with the NN with a linear stack of fully connected layers. The MAE and MedAE for cross-validation were equal to 34.7 s

and 18.7 s, respectively. The generalization ability of the constructed model (1D CNN) for the small data sets was demonstrated with a transfer learning approach. Thus, we believe that the results of these studies can be used as a solid basis for retention modeling of small drug-like molecules in different LC conditions.

Acknowledgment

T06. A new hope in the analysis of novel foods in rat diet for toxicology studies: transitioning from NIR to reflectance UV Spectroscopy

J. Elek [1], M. Csontos [1,2], I. Szálkai[1]
[1] Science Port Ltd. Debrecen, Hungary
[2] University of Debrecen, Debrecen, Hungary

In a galaxy not so far away, at the forefront of culinary exploration, numerous factions have sought approval from the esteemed European Food Safety Authority (EFSA) for their novel food creations. These culinary innovations, derived from the farthest reaches of the galaxy, present unique challenges in ensuring their safety for consumption across the cosmos.

In the bustling Science Port Ltd., nestled among the stars, our skilled scientists, equipped with the knowledge of ancient and modern techniques, have embarked on a critical mission. The complex, multi-component nature of these interstellar foods, brimming with biological mysteries, defies the traditional methods of analysis, such as the chromatographic techniques of old, particularly in dietary studies of the diverse species of the galaxy. First, we've met the Near Infrared (NIR) Spectroscopy – a method as swift and non-intrusive as a Jedi's mind trick. This technique, was a beacon of hope and the entrance into this rapidly expanding market, and has been the chosen path for its ability to analyze without destruction, providing insights at the speed of light.

However, the tides are changing. The innovative minds at Science Port Ltd. have now set their sights on the advanced Reflectance UV Spectroscopy, harnessing the power of an integrating sphere UV spectrophotometer. This new ally in our

quest offers a robust and sensitive approach, eclipsing the older methods with its superior capabilities. The essence of this transition lies in the unique powers of Reflectance UV Spectroscopy – its immunity to the interference of water absorption, akin to a starship's shields against cosmic rays, and its heightened sensitivity, especially for compounds that absorb in the UV spectrum. The integrating sphere, a marvel of our technological prowess, allows for diffuse reflectance measurements, capturing the full spectrum of these enigmatic solid samples.

In an age where simplicity is often outmatched by complexity of deep learning algorithms, our scientists have embraced the wisdom of ancient chemometric techniques, such as principal component analysis (PCA) and partial least squares regression (PLSR). These methods, like navigational pathways through the stars, enable us to decode and interpret the complex data tapestries woven by UV spectroscopy with unparalleled precision. Our latest endeavor presents a comparative star chart of NIR and Reflectance UV Spectroscopy in their quest to assess the safety of these novel foods, designed for the diverse diets of species across the galaxy. Can our findings, illuminated by the integrating sphere UV spectrophotometer, demonstrate a leap in sensitivity and specificity?

We stand at the dawn of a new era in toxicological evaluations, armed with our advanced instrumentation and data analysis tools. This convergence of science and technology illuminates a path forward for the safe consumption of novel foods, ensuring that all beings, from the deserts of Tatooine to the forests of Endor, can enjoy these cosmic culinary delights safely.

T07. Multivariate Data Analysis of 2D Fluorescence Spectra in Monitoring of Imprinted Proteins Purification

N.A. Burmistrova, P.M. Ilicheva, K.Y. Presnyakov, P.S. Pidenko
Institute of Chemistry, Saratov State University, Saratov, Russia

Molecular imprinting is a widespread technique to create synthetic recognition receptors based on template-shaped cavities in polymer matrices with predetermined selectivity and high affinity. The formation of binding sites on the surface of proteins as a result of modification of biomolecules in the presence of templates is one of the modern trends of contemporary chemical analysis. Biocompatible imprinted proteins (IPs, proteins with molecular imprints, bioimprinted proteins) are a promising tool for solving various problems of analytical chemistry [1].

The purification is an important stage of the IPs synthesis that determines the sorption and analytical properties of analytical systems. However, the spectral bands of the protein and the template may overlap, thus, making it difficult to use UV and IR spectroscopy to monitor the IPs purification degree. Fluorescence makes it possible to reduce the sensitivity of detection and record the signal at multiple excitation wavelengths, thus, it is a more suitable method. Fluorescence spectra can be detected at excitation wavelengths over a wide range (3D) or at characteristic wavelengths of the protein and template (2D). However, the recording of 2D and 3D spectra leads to a large amount of data that may prove to be difficult to interpret correctly. Standard approaches to data processing are not suitable for this task. Previously, we demonstrated the possibility of visualizing excitation-emission matrices (3D fluorescence) in the form of heat maps.

This work presents an example of fluorescence data analysis obtained as a results of monitoring the process of purification of IPs obtained in varied conditions. The imprinting of bovine serum albumin (BSA) in the presence of 4-hydroxycoumarin were performed with varied protonation conditions.

We used the principal component analysis (PCA) and PCA models to study the fluorescence data obtained with excitation wavelengths corresponding to BSA and template, as well as 2D fluorescence data.

The analysis of PCA score plots revealed the clustering of objects before and after purification of IPs. Moreover, the position of objects after purification was dependent on the conditions of BSA imprinting. Decomposition of multimodal fluorescence data allowed to identify patterns and was used to select optimal conditions of IPs performance. This approach can be used as a basis for automating the decision-making process based on analysis of fluorescence data.

Acknowledgement

References

[1] P.S. Pidenko, K.Y. Presnyakov, D.D. Drozd, et al. *J. Anal. Chem.* **78** (2023) 1146–1151.

T08. Comparison and optimization of mass spectral library search algorithms

A.S. Samokhin [1], M.D. Khrisanfov [1,2], D.D. Matyushin [2]

1 Chemistry Department, Lomonosov Moscow State University, Moscow, Russia
2 A.N. Frumkin Institute of Physical Chemistry and Electrochemistry, Russian Academy of Sciences, Moscow, Russia

Distance-based algorithms have been typically used to perform the library search in spectroscopy and mass spectrometry. The dot product seems to be the most popular one. Nevertheless, Euclidean or Manhattan distances are also applied. In mass spectrometry similarity index is often calculated as a sum of several terms or a product of several factors. Moreover, algorithms based on machine learning [1] are the trend of recent years (however they are not considered in the present communication). The selection of a particular library search algorithm depends on an analytical method in use. For example, in mass spectrometry, different algorithms are used for electron ionization mass spectra, tandem mass spectra of small organic molecules, or tandem mass spectra of peptides. In the present communication, we limit the scope of claims to searching against electron ionization mass spectral databases using algorithms based on traditional programming.

One of the most efficient library search algorithms implemented in commercial software products is the Identity (also known as Composite) algorithm [2]. In this case, the similarity index is calculated as a weighted average of two terms (dot product and a term taking into account ratios of adjacent mass spectral peaks). This algorithm is described in both the scientific article [3] and the MS Search (NIST) software manual [4]. Nevertheless, according to our knowledge no one has been able to exactly reconstruct the Identity algorithm, despite a large number of attempts. We also had made several attempts but in vain. As a result, an API described in the MS Search manual has been the only way to use the Identity algorithm in third-party products for two decades. Recently we have made another attempt to reveal the algorithm. This time we iteratively considered several groups of artificial mass spectra. We started with trivial spectra containing only two peaks and gradually increased their complexity. Luckily we were able to puzzle out how the similarity index is actually calculated and got identical results using the MS Search software and our implementation of the Identity algorithm. It inspired us to perform a new round of comparison/optimization of known library search algorithms taking into account new findings.

Acknowledgement

References

[1] D.D. Matyushin, A.Y. Sholokhova, A.K. Buryak, *Anal. Chem.* **92** (2020) 11818–11825.

[2] A. Samokhin, K. Sotnezova, V. Lashin, I. Revelsky, *J. Mass Spectrom.* **50** (2015) 820–825.

[3] S.E. Stein, D.R. Scott, *J. Am. Soc. Mass Spectrom.* **5** (1994) 859–866.

[4] Mass Spectral Libraries and NIST Mass Spectral Search Program (Version 2.4). User's Guide. National Institute of Standards and Technology; (2020).

T09. Aquaphotomics and NIR spectroscopy: from basics to practical application

A. Surkova [1], *E. Boichenko* [2]

[1] *Samara State Technical University, Samara, Russia*

[2] *Institute of Chemistry, Saint Petersburg State University, Saint Petersburg, Russia*

Water is one of the main components of various biological systems. However, scientists are usually focused on detecting single biological molecules (proteins, nucleotides, etc.) during the study of biological systems, considering water only as a medium. Thus, the aqueous structure of many biological systems remains incompletely investigated, while the water matrix of a sample, i.e., different conformations of water molecules, can be a valuable source of information itself. Near infrared (NIR) spectroscopy is a rapid and accurate method for nondestructive and noninvasive analysis of biological objects. NIR spectra contain rich information about the structure of water, especially in the 1st overtone region of the OH stretching band (1300–1600 nm), where many water absorbance bands are recognized. A new promising scientific discipline – aquaphotomics, proposed by Prof. Tsenkova in 2005 [1], is studying the molecular structure of water by NIR spectra of biological objects. Aquaphotomics studies aqueous and biological systems as holistic systems and considers a combination of water absorption bands rather than a single one. Based on numerous studies in this field, a spectral database of water absorption bands with their physicochemical interpretation, called the global aquaphotome, was created. This database consists of more than 700 characterized water absorption bands in the region of 700–1900 nm and it is constantly being enriched by new studies of various objects, which contributes to further

development of the scientific basis of aquaphotomics and understanding of the role of water in biological systems. Aquaphotomics can be used for prediction, diagnosis and advanced understanding of the biology, chemistry and physics of biological and aquatic systems.

The application of chemometrics methods to extract information of water molecular structure from NIR spectra is a necessary part of the aquaphotomics data analysis. This study focuses on the methodological aspects of aquaphotomics, from required equipment to data processing. Aquaphotomics data analysis will be demonstrated on several datasets. One of them is devoted to the study of biological fluids of patients with diagnosed cancer in different locations before and after cancer surgery. A second example of the application of aquaphotomics will be presented for the detection of malignant tumor sites in a mouse model of carcinogenesis.

References

[1] R. Tsenkova, J. Munćan, B. Pollner, Z. Kovacs. *Front. Chem.* **6** (2018) 363.

T10. Spectroscopy without spectra?!

A. Bogomolov [1,2], E. Ignatiev [1]

[1] *Samara State Technical University, Samara, Russia*

[2] *art photonics GmbH, Berlin, Germany*

In the modern world, the number and complexity of chemically analyzed samples is growing rapidly. There is an increasing need for real-time analysis performed outside the chemical laboratory in-line or in the field conditions. The development of optical multisensor systems (OMS) [1] provides an answer to these new challenges in chemical analysis. OMS is a low-selectivity optical analyzer specialized for a particular application. It typically includes only a few highly optimized light sources, such as light-emitting diodes (LEDs) or optical filters covering relatively broad wavelength ranges and a simple photometric detector that captures their averaged signals.

The selection of optimal spectral intervals is therefore a critical step in the development of OMS. As shown in our recent paper [2] using simulated spectra of a two-component mixture, calibration models for such optimized intervals often provide an even more accurate prediction than the underlying full-spectrum data. Moreover, the best intervals for calibration tend to be broad and overlapping (Fig. 1).
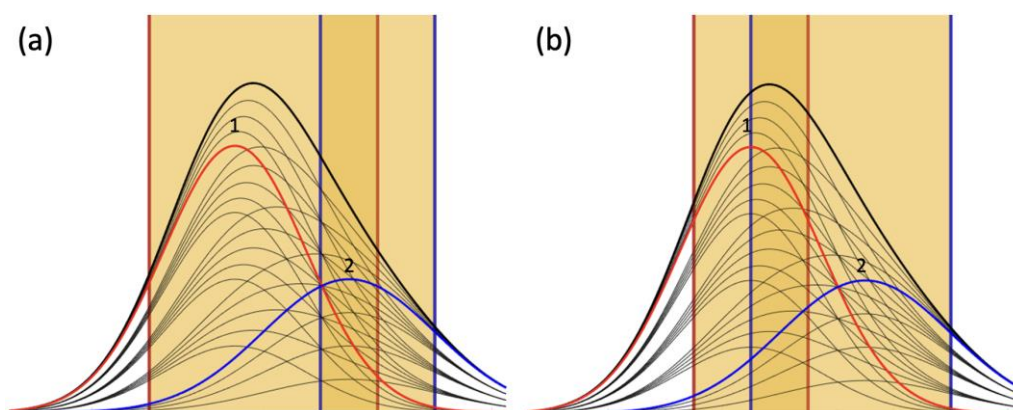
Figure 1. The best two intervals for calibration for: (a) component 1 (red curve) and component 2 (blue curve).

The advantages of using broad overlapping intervals as a viable alternative to the traditional high-resolution spectroscopy are discussed. The data generated by new spectroscopic analyzers of this type differs dramatically from conventional spectral variables represented by a sequence of hundreds of narrow spectral intervals. This may require a revision of the modern software and algorithms for data visualization and analysis.

References

[1] A. Bogomolov, *Sensors.* **21** (2021) 3541.

[2] A. Bogomolov, A. Evseeva, E. Ignatiev, V. Korneev, *TrAC.* **160** (2023) 116950.

T11. Chemometric processing of the data from sensor device for contactless chemical analysis based on high-frequency inductance coil

E. Yuskina [1], M. Khaydukova [2], V. Semenov [1,3], V. Panchuk [1,3], D. Kirsanov [1]

[1]Institute of Chemistry, St. Petersburg University, St. Petersburg, Russia

[2]Laboratory of Alternative Antimicrobial Biopreparations, World-Class Research Center "Center for Personalized Medicine", FSBSI Institute of Experimental Medicine, St. Petersburg, Russia

[3] Institute for Analytical Instrumentation RAS, St. Petersburg, Russia

An important task in modern analytical chemistry is the development of cost-effective and simple sensor devices for analyzing complex samples in real time. Moreover, there is a preference for contactless methods that enable analysis without the need for sampling and sample preparation. This study is focused on developing and investigating the analytical capabilities of a low-cost contactless sensor device. The device is based on an inductance coil connected to a high-frequency electric field generator. When a sample is placed inside the coil, it becomes a core of the inductor, which alters the characteristics of the electric

signal passing through the coil. A receiver connected to the coil detects and records these alterations. Unlike the non-contact conductometers of the mid-20th century, which were only able to operate at a single alternating current (AC) frequency, the proposed device, due to its modern electronic components, can measure responses across a spectrum of frequencies ranging from 1 to 114 MHz. The obtained signal is a complex function of the sample's conductivity, polarizability, and capacitive properties. Analyzing this signal using chemometric methods yields qualitative and quantitative information regarding the sample's composition.

This sensor device can be employed to differentiate between substances with distinct physical and chemical properties, ascertain the concentration of ions in aqueous solutions, determine the key characteristics of complex mixtures (e.g., fat content in milk, alcohol concentration in aqueous-alcohol mixtures), and differentiate between cultural media containing diverse bacteria and cancer cells [1]. Furthermore, the device can be utilized to obtain the growth curves of bacteria and cells. The presentation will focus on the application of various chemometric tools for processing the data from high-frequency inductance coil sensor.

Acknowledgement

References

[1] E. Yuskina, N. Makarov, M. Khaydukova, T. Filatenkova, O. Shamova, V. Panchuk, D. Kirsanov, *Anal. Chem.* **94**(35) (2022) 11978–11982.

T12. Needle in a haystack

M. Csontos [1,2], J. Elek [1]

[1] Science Port Ltd. Debrecen, Hungary

[2] University of Debrecen, Debrecen, Hungary

The registration and toxicological testing of the next generation of the food additives and dietary supplements – so called novel foods – bring new challenges from the analytical point of view as well. A large number of these materials belong to the group of UVCB materials: Unknown or Variable Composition, Complex Reaction product or Biological Origin1. These multi-component test items may contain a variety of proteins, fibres, fat compounds, and many other types of constituents. Not only the separation of these components can be strenuous with the classical analytical methods, but also the

selection and targeted analysis of the main components. In this presentation I will demonstrate the development and validation of a near-infrared method combined with multivariate data analysis. The method was validated for the quantitation of a multicomponent biological novel food in a quite complex matrix: semi-synthetic rodent diet.

References

[1] S.D. Dimitrov, D.G. Georgieva, T.S. Pavlov, Y.H. Karakolev, P.G. Karamertzanis, M. Rasenberg, O.G. Mekenyan, *Environ Toxicol Chem.* **34** (2015) 2450–2462.

T13. Laser-based analytical techniques and multiblock chemometrics for the analysis of marine zooplankton

N. Sushkov [1], N. Lobus [2], A. Kéri [3], Á. Bélteki [3], G. Galbács [3], T. Labutin [1]

[1] Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia

[2] Timiryazev Institute of Plant Physiology, Russian Academy of Sciences, Moscow, Russia

[3] Faculty of Science and Informatics, University of Szeged, Szeged, Hungary

This study deals with the analysis of marine zooplankton. Some species of small crustaceans are known to accumulate considerable amounts of lithium, uranium and other elements. At present, the reasons and mechanisms behind this phaenomenon are unclear. Laser-based techniques make it possible to characterise the chemical composition of animals with minimal sample destruction in order to provide basis for the elucidation of biochemical mechanisms behind the element enrichment. In our previous studies [1,2], we used full-length spectra obtained by laser-induced breakdown spectroscopy (LIBS) and Raman spectroscopy to explore the correlations between elemental and molecular composition of *Calanoida* and *Euphausiacea* crustaceans, arrow worms, and sea snails. Lithium content has been found to correlate with that of pigments and amino acids.

As a follow-up research, we have considered elemental signals obtained by inductively coupled plasma mass spectrometry with laser ablation (LA-ICP-MS). This technique enabled the detection of additional elements which were hard to detect with LIBS (such as S, As, U, etc.). The experiments were carried out using the Applied Spectra J200 Tandem LA-LIBS spectrometer (266 nm ablating laser) and the Agilent 7700 Series ICP-MS spectrometer. Raman measurements were done using the Thermo Scientific DXR Raman Microscope (780 nm laser) in the range of 45–3500 cm^{-1}. In all spectra, baseline was removed, and prominent

peaks were integrated. Data blocks were Frobenius-normalised. The final matrix contained around 2700 mass-spectral and Raman signals originating from 29 zooplankton samples.

Conventional PCA decomposition seemed inefficient for the exploration of the data, because the information was dispersed across too many principal components. Thus, we opted for multiblock algorithms, namely, consensus PCA (CPCA) and ComDim–PCA. For the moment, ComDim–PCA results are the easiest to interpret. One of the common dimensions (CDs) contains the signals of Li and pigments together, which supports the earlier findings. The same CD also contains the U signal. Zinc is partially associated with lithium, while the other studied elements (B, Na, Mg, Al, Si, P, S, Cl, K, Ca, Fe, Cu, As, Sr, and Ba) are associated with fats and definitely go to a separate CD. Thus, a special physiological role may exist for Li, U, and Zn.

Acknowledgement

References

[1] N. Sushkov, G. Galbács, K. Fintor, N. Lobus, T. Labutin, *Analyst.* **147** (2022) 3248–3257.

[2] N. Sushkov, G. Galbács, P. Janovszky, N. Lobus, T. Labutin, *Sensors.* **22** (2022) 8234.

T14. A new approach to finding potentially erroneous entries in a database using machine learning

M.D. Khrisanfov [1,2], _D.D. Matyushin [2]_, _A.S. Samokhin [1]_

[1] Chemistry Department, Lomonosov Moscow State University, Moscow, Russia

[2] A.N. Frumkin Institute of Physical Chemistry and Electrochemistry, Russian Academy of Sciences, Moscow, Russia

The NIST Retention Index (RI) database is one the most widely used sources of retention indices. Its latest version contains more than 300 000 entries (combinations of compound-stationary phase-RI value) from various sources: experimental articles, personal communication, etc. According to our estimates about 80% of the compounds from both NIST17 RI and NIST20 RI databases

have only one RI value per stationary phase, which makes searching for erroneous values with statistical methods impossible. Manual inspection of entries also does not seem viable due to the size of the database.

We designed a two-step approach to find potentially erroneous retention indices. The first step was to use five predictive models to obtain predicted RI values for the whole database. The models and their respective data preprocessing pipelines were heavily inspired by the following article [1]. We used 5-fold cross-validation to train each of the five predictive models. However, validation sets were used only for obtaining predicted RI values. We did all the tuning of hyperparameters and architectures for the models beforehand. At the end of the step, we had five predicted RI values for each compound-stationary phase-RI combination.

The second step was to compare these five predicted values against the experimental ones. For each of the models we considered 5% RI values with the biggest difference between the predicted value and the experimental one to be outliers and marked them with a "yellow card". This way each of the entries got from zero to five "yellow cards". We wanted to stay cautious and deemed only entries with 5 "yellow cards" to be potentially erroneous.

There are two main reasons for an entry to get 5 "yellow cards". Either the compound is too unique for its RI value to be adequately predicted or there is an erroneous RI value in the database. In the former case the predictions of different models would be likely inconsistent: the standard deviation would be big, because prediction errors of the models are not fully correlated. In the latter case the predictions would likely align better and the standard deviation would be much smaller. We discovered that the median and the maximum of the distribution of the standard deviations of all five predicted values in the group with 5 "yellow cards" was significantly smaller than even in the group with 1 "yellow card". Overall, we were able to detect 2093 outlier entries in the NIST 17 RI database, 566 of those were corrected or removed by the developers in NIST 20 RI.

Acknowledgement

References

[1] D.D. Matyushin, A.K. Buryak, *IEEE Access.* **8** (2020) 223140–223155.

T15. Feasibility study of X-ray fluorescence spectra and chemometrics in dating XVIII century glass samples

A. Tretyakov [1], S. Sirro [2], D. Kirsanov [1]
[1]Institute of Chemistry, St. Petersburg University, St. Petersburg, Russia
[2] The State Russian Museum, St. Petersburg, Russia

The 18th century was an era of change and instability in Russia, with several emperors succeeding each other on the throne. The exquisite glassware, made individually for each of these rulers, serves as a testament to their rule and greatness. In modern museum practice, the dating of this glassware is important for the structuring of collections and their correct chronological classification.

Various dating methods are used to determine the age and origin of these objects. Glass can be dated by means of epigraphy (analysis of inscriptions and graphemes), stylistic analysis, and an increasing number of instrumental methods of analysis. The most popular method of analysis for a number of parameters is X-ray fluorescence (XRF). It is a non-destructive method that is widely used because it is fast and accurate and can be carried out with commercially available equipment.

This study is an attempt to investigate the dating potential of mathematical processing of XRF spectra applied to 18th century imperial glassware samples. The research was carried out on 247 glass samples. In the first stage of the research, principal component analysis (PCA) was used to reduce the dimensionality and extract the main elements determining the belonging of samples to a certain time period. In the second phase of the study, PLS-DA and NCA (neighborhood component analysis) methods were used to analyse the data more precisely, extract key features and create a classification model. The details of the work will be discussed in the presentation

T16. Non-target analysis by gas chromatography-mass spectrometry using machine learning and deep learning

D.D. Matyushin [1], A.Yu. Sholokhova [1]
[1] A.N. Frumkin Institute of Physical Chemistry and Electrochemistry, RAS, Moscow, Russia

Chromatography-mass spectrometry can be used for non-target analysis of complex mixtures without their preparative separation. The mass spectrum of fragment ions provides valuable information about the structure of the unknown, but this information is usually not enough for unambiguous structure

determination. Chromatographic retention provides additional information that is somewhat orthogonal to the spectral data. This work is devoted to several aspects of using machine learning and deep learning for the determination of unknowns in complex mixtures using chromatography-mass spectrometry.

Prediction of gas chromatographic retention indices for candidate molecules during non-target analysis and then further comparison with observed retention can considerably reduce the candidate space. In this work, we compare several methods of machine learning for this task: deep learning, support vector regression, gradient boosting. Previously developed by our team deep learning-based software was also considered. The developed in this work improved atom-wise features for graph neural networks allow us to obtain record-breaking accuracy. It is generally believed that deep learning (deep neural networks) is the most accurate and versatile method for this task. In this work, we show that the combination of k-means clustering for partitioning the space of possible structures and support vector regressions (for each of the clusters) achieves accuracy comparable to deep learning and even surpassing deep learning for many classes of chemical structures.

The next task to be considered is the machine learning ranking. We previously demonstrated that using deep learning to compare observed and reference spectra using machine learning outperforms previously used methods in a library search task. In this work, we consider the problem of selecting a candidate that best matches the observed data. We use deep learning for prediction of a so-called "molecular fingerprint" (the list of substructures present) from a spectrum, for prediction of a spectrum from a structure, and for overall comparison of data obtained from the experiment with data calculated from a structure. We use previously developed approaches for prediction and demonstrate that machine learning-driven comparison outperforms simpler approaches.

Finally, we demonstrate the application of machine learning methods to a real-world example: the determination of structures of new transformation products of unsymmetrical dimethylhydrazine. Being introduced to the environment, this rocket propellant forms an enormous variety of potentially toxic compounds that mostly are not investigated and absent in databases. It was shown that previously reported approaches (that do not use preparative separation and standard samples) often result in wrong structures. The combination of several machine learning methods and several methods of chromatography-mass

spectrometry allows a relatively reliable prediction to be achieved in many cases.

Acknowledgment

P01. Paint It Blue: identification of synthetic and natural ultramarine via NIR spectroscopy and chemometrics

M. Khaydukova [1], E. Boichenko [2], I. Andreev [3], S. Sirro [3], D. Kirsanov [2]

[1] Laboratory of Alternative Antimicrobial Biopreparations, World-Class Research Center "Center for Personalized Medicine", FSBSI Institute of Experimental Medicine, Saint Petersburg, Russia

[2] Department of Analytical Chemistry, Institute of Chemistry, St Petersburg University, Saint Petersburg, Russia

[3] The State Russian Museum & ITMO University, Saint Petersburg, Russia

Physico-chemical examination of paintings, especially characterization of pigments, is a challenging, but aesthetically pleasing area of research. Certainly, the development of nondestructive analytical methods is preferable due to the uniqueness and value of art objects under study. Identification of synthetic and natural pigments is one of the ways of dating of paintings, since the production of synthetic pigments started in the 1850s, and each painter has their own preferences in dyes. However, current methods imply taking of small probes of the pigments: for example, highly sensitive HPLC techniques [1] or IR spectroscopy in the middle spectral region. Energy-dispersive X-ray spectroscopy is a noninvasive method, but its application is limited to the determination of elemental composition of the sample. Its sensitivity is not sufficient for classification of natural and synthetic dyes, since the differences between them are subtle [2].

Near-infrared spectroscopy (NIR) with chemometric data analysis is an interesting approach that allows using fiber probes to avoid the sample preparation step and is suitable for classification of multicomponent objects. This approach has shown promising results in various fields, for example, as a nondestructive method of counterfeit detection [3].

In this study, we report the results of a classification of different dyes based on their NIR spectra, measured on 14 samples of ultramarine pigments (9 natural and 5 synthetic). Three replicates were prepared for each paint with different

binders: whole egg, yolk and linseed oil. The NIR measurements were performed by a fiber optic probe 1.8 nm in diameter (Optofiber, Russia), connected to the portable NIR spectrometer (Avantes, Netherlands) within the 939–1799 nm spectral range. The results of PCA and other methods of multivariate spectral data analysis will be presented on the poster. It is shown that the proposed approach allows classification of the paints according to the binder and to the origin of the ultramarine pigment.

References

[1] M. Trojanowicz, J. Orska-Gawryś, I. Surowiec, B, et al, *Studies in Conservation*. **49**(1) (2004) 115–130.

[2] M. Bacci, C. Cucci, E. Del Federico, A. Ienco, et al. *Vib. Spectrosc.* **49**(1) (2009) 80–83.

[3] O. Rodionova, A. Titova, N. Demkin, K. Balyklova, A. Pomerantsev, *Talanta.* **195** (2019) 662–667.

P02. Optical and Electrochemical sensor arrays for Non-Steroidal Anti-Inflammatory Drags Assessment

L. Lvova, R. Paolesse

Department of Chemical Sciences and Technologies, University "Tor Vergata", Rome, Italy

Non-steroidal anti-inflammatory drugs (NSAIDs) are used against fever and other symptoms of colds and flu, to reduce inflammation, headache, muscular pain, and others [1]. Among NSDAIDs ketoprofen (KF) having less negative side effects is the most widely used, and it is a main component of many pharmaceutical compositions both for adults and children. The COVID-19 pandemic of last years has significantly increased the overall consumption of NSAIDs and KF in particular. This made the anti-inflammatory painkillers the emerging environmental pollutants. Also, NSAIDs are toxic for biota and able to accumulate in tissues thus causing serious damages [2], and their assessment in environment and careful concentrations screening in pharmaceutical compositions is an important and challenging analytical task. The standard analytical methods for NSAIDs assessment are chromatography [3], and spectrophotometry [4], also chemical sensors application were reported. [5-7]. In this contribution we summary our last achievements on development of potentiometric and optical sensor array for detection of KF and its recognition among other, similar in structure and in pharmaceutical properties, non-

steroidal pain-killers, as for instance naproxen, NS, and ibuprofen, IB. The acyclic squaramide receptors [8] were exploited as membranoactive components for potentiometric ion-selective electrodes with PVC solvent polymeric membranes, while the open-chain polypyridine ligands bearing pyrene fluorogenic groups able to signal selective analyte binding through the enhancement of the emission [9] were used for optical sensors development. The importance of a transduction method, the selection of an appropriate method of image analysis for data extraction for optical output, and construction of calibration models and pattern recognition procedure choice will be discussed. Finally, the applications of the developed optical and potentiometric sensor arrays for NSAIDs assessment in real samples will be illustrated.

References

[1] P.M. Brooks, R.O. Day, *N. Engl. Med.* **324** (1991) 1716.

[2] I. Nishi, T. Kawakami, S. Onodera, *J. Env. Sci. Health.* **50** (2015) 1108.

[3] Y. Fan, Y.-Q. Feng, S.-L. Da, Z.-H. Wang, *Talanta.* **65** (2005) 111.

[4] M. El-Sadek, S. El-Adl, M. Afiou-Kull, S.M. Sabb, *Talanta.* **40** (1993) 585.

[5] J. Lenik, J. Nieszporek, *Sens. Act. B.* **255** (2018) 2282.

[6] D.A. Ahmed, M.K. Abd El-Rahman, H. M. Lotfy, S. A. Weshahy, *J. Electrochem. Soc.* **165** (2018) H999.

[7] H.M. Essam, Y.F. Bassuoni, E.S. Elzanfaly, H.E.-S. Zaazaa, K.M. Kelani, *Microchem. J.* **159** (2020) 105286.

[8] S. Farotto, G. Picci, J. Milia, C. Caltagirone, V. Lippolis, M.C. Aragoni, C.Di Natale, R.Paolesse, L. Lvova, *ACS Sens.* **8** (2023) 3225.

[9] L. Conti, C. Giorgi, B. Valtancoli, P. Paoli, P. Rossi, A. Marchionni, E. Faggi, A. Bencini, *Chem. Plus. Chem.* **85** (2020) 659.

P03. Inductance coil and chemometrics – the way to contactless chemical analysis

E. Yuskina [1], N. Makarov [1], M. Khaydukova [2], T. Filatenkova [2], O. Shamova [2], V. Semenov [1,3], V. Panchuk [1,3], D. Kirsanov [1]

[1] Institute of Chemistry, St. Petersburg University, St. Petersburg, Russia

[2] Laboratory of Alternative Antimicrobial Biopreparations, World-Class Research Center "Center for Personalized Medicine", FSBSI Institute of Experimental Medicine, St. Petersburg, Russia

[3] Institute for Analytical Instrumentation RAS, St. Petersburg, Russia

Modern analytical chemistry is focused on the development of analytical methods suitable for analysis of real complex objects in field conditions preferably in a non-contact mode. Recently we have proposed [1] a novel type of sensor based on an inductance coil that fulfil these requirements. The device operates in the alternating current frequency range from 2 to 112 MHz, and it can be considered as a simplified version of dielectric spectroscopy – the method which has already proved its potential in analysis of complex samples using chemometric data processing. An inductance coil is connected to a high-frequency electric field generator and the sample in a plastic tube is placed inside the coil, thus becoming the core of the inductor. The sample changes the properties of the electrical signal passing through the coil. A receiver connected to the coil records the response spectrum, which depends in a complex way on the conductivity, dielectric constant, polarizability of the sample, as well as its magnetic and capacitive properties. Since there are numerous factors contributing to the signal, the obtained spectra can be processed by chemometric algorithms to obtain qualitative and quantitative information about the samples.

The presentation will describe the applications of the sensor explored so far: quantification of ethanol in its aqueous mixtures, quantification of several inorganic ions in aqueous solutions; determination of physical properties of liquids (dipole moment); quantification of integral quality parameters (% of fat in milk); and recognition of complex biological media containing different bacterial and cell cultures.

Acknowledgment

References

[1]   E. Yuskina, N. Makarov, M. Khaydukova, T. Filatenkova, O. Shamova, V. Panchuk, D. Kirsanov, *Anal. Chem.* **94** (2022) 11978–11982.

P04. Chemometrics can convert Mössbauer spectrometer into X-ray fluorescence one

V. Panchuk [1,2], V. Semenov [1,2], M. Grigoriev [1], D. Kirsanov [1]
[1]Institute of Chemistry, St. Petersburg University, St. Petersburg, Russia
[2]Institute for Analytical Instrumentation RAS, St. Petersburg, Russia

Mössbauer spectroscopy is a popular tool for analysis of chemical states of resonant atoms and phases in solids. It has found a broad applicability in

material science, medical chemistry, geology, biology, mineralogy. Quite often the information from Mössbauer spectroscopy has to be accompanied with the data on elemental composition of samples, thus, additional instrumental methods have to be involved in the study. One of the most popular options for elemental analysis in these studies is X-ray fluorescence (XRF) spectrometry, since it allows direct quantification of elements in solids without sample pretreatment. Noteworthy, both Mössbauer and XRF spectrometry employ excitation electromagnetic radiation in similar energy ranges around 10–100 keV. The radiation detection components of both types of spectrometers are also similar, the most widely spread are semiconductor detectors, scintillation detectors and proportional gaseous counters. However, commercial Mössbauer spectrometers usually use scintillation or proportional counters as detectors that have higher registration efficiency than semiconductor detectors, while EDXRF spectrometers use semiconductor detectors with higher resolution. All these circumstances and similarities made us hypothesize that Mössbauer spectrometer can be in principle employed for direct registration of EDXRF spectra. It is important, however, that scintillation or proportional detectors from standard Mössbauer setup will yield rather poor spectral resolution for XRF measurements. This issue can be at least in part compensated by using multivariate data processing and chemometric algorithms.

This study aims to check the validity of our hypothesis on the possibility of using Mössbauer spectrometer for elemental analysis. In order to prove this thesis, we first have simulated spectral XRF data from steel samples with good (150 eV) and poor (2000 eV) spectral resolution and compared the accuracy in model element quantification. Second, we have studied real steel samples using both Mössbauer spectrometer and commercial EDXRF spectrometer. The obtained spectral data were processed with multivariate regression modeling to obtain target element content. The results of Ti, Cr, Mn, Ni, Nb and Mo quantification in real steel samples obtained from Mössbauer spectrometer were compared with those obtained with commercial energy dispersive X-ray fluorescence (EDXRF) spectrometer. In case of Ti, Cr and Ni the accuracy attained with Mössbauer spectrometer was similar to that of an ordinary commercial EDXRF device.

Acknowledgment

P05. Spectroscopic analysis of somatic cells in milk in the region of 400-1000 nm

J. Kostyuchenko, A. Bogomolov

Samara State Technical University, Samara, Russia

One of the most important microbiological parameters of milk quality is somatic cell content. Somatic cells come from various tissues and organs. These cells are always present in milk, but higher somatic cell content may indicate the presence of mastitis and other inflammatory processes. Therefore, it is important to determine excess of somatic cells in a timely manner. Currently, somatic cells are determined by various methods that involve expensive equipment, reagents and highly qualified staff.

Optical spectroscopy in the mid infrared (IR) region is widely used in the dairy industry. Although this method allows the detection of somatic cells, it is too expensive for general farming. Visible and short-wave near IR (Vis/SW-NIR) spectroscopy in the range of 400–1100 nm can be an affordable alternative. It is a flexible and less expensive technique that does not require extensive sample preparation and is therefore suitable for rapid milk analysis.

In this work, the analytical potential of Vis/SW-NIR spectroscopy for somatic cell detection was explored. Data analysis has shown the presence of spectral signals in the Vis region, which are supposedly related to somatic cell absorption. Application of principal component analysis (PCA) and partial least-squares discriminant analysis (PLS-DA) algorithms has proved that Vis/SW-NIR spectroscopy can be used to detect somatic cells at practically relevant quantities. The proposed method is suitable for an express-analysis of somatic cell count and can be used as a development basis for an optical multisensor system.

P06. Chemometrics-assisted determination of some beta-lactam antibiotics in their mixtures

R. Mursalov, E. Ketrush, D. Silaev, T. Rusanova

Saratov State University, Saratov, Russia

In medications, a combination of two or more active pharmaceutical ingredients (API) is often used to enhance the pharmacological effect and ease of using combined preparations. The time-consuming and complex preliminary

separation of APIs, which are part of combined preparations and have similar physico-chemical properties, makes it difficult to conduct rapid analysis during quality control in the laboratory. Chemometric algorithms, which are currently promising, rapid and do not require much effort, are an alternative to time-consuming procedures associated with the preliminary separation of the components of the mixture and the use of expensive equipment in the analysis. The purpose of this work was to evaluate the possibility of applying a combination of the spectrophotometry with chemometric algorithms (partial least squares, PLS) for the quantitative determination of certain β-lactam antibiotics in their mixtures.

The objects of the study were selected antibiotics of the penicillin (Amoxicillin, Amox) and cephalosporin (Cefuroxime (Cefur), Ceftriaxone (Ceft), Cefotaxime (Ctax), Cefazolin (Cef)) series. Five binary mixtures Cefur-Cef, Ceft-Ctax, Cefur-Ctax, Ceft-Cef and Cefur-Amox have been studied. A series of 24 model solutions with different ratios of antibiotics were prepared for each binary mixture. The electronic absorption spectra were recorded using a Shimadzu UV-1800 spectrophotometer in the range of 220–350 nm with a registration step of 1 nm. The determination of antibiotics was carried out by two variants of the partial least squares method: PLS-1 and PLS-2. The number of mixtures in the training and test sets was selected by the method of cross-checking. Based on the values of the standard errors of prediction, the optimal numbers of latent variables are selected. Verification of the accuracy of the determination was carried out by the "added-found" method.

It was found that application of the PLS1 and PLS2 chemometric methods for the "cephalosporin : cephalosporin" systems provides lower values of standard deviations compared with the "cephalosporin : penicillin" systems.

Thus, the possibility of spectrophotometric determination of the studied antibiotics in their binary mixtures using the PLS chemometric algorithm is shown. It was found that all the studied models give high values of correlation coefficients and tangents of the slope angle of the "measured-predicted" dependencies close to one, which allows us to recommend the proposed method for practical use [1].

References

[1] E. Ketrush, R. Mursalov, D. Silaev, T. Rusanova, *Izvestiya of Saratov University. Chemistry. Biology. Ecology.* **23**(4) (2023) 392–403.

P07. Chemometric Analysis of ATR-FTIR Spectra of Paintwork Materials and Coatings

N.A. Burmistrova, K.T. Musina, S.P. Kurchatkin

Institute of Chemistry, Saratov State University, Saratov, Russia

Paintwork materials and coatings are the most common pieces of physical evidence used in the investigation of road traffic accidents. The purpose of an expert study during an accident investigation is to establish the identity of paintwork coatings samples taken from the suspected vehicle and the substance layer from injured vehicle.

Multilayer and multicomponent composition of paintwork materials and coatings as well as a limited amount of material available for testing significantly complicates the task. In addition, the tested materials usually lack individual characteristics due to the uniform vehicle batch production and unknown manufacturing data for paints and varnishes. Also, the greatest difficulty of such analysis comes from taking a sample only from the top layer without penetrating and taking a part of the base color enamel.

Attenuated total reflectance (ATR) using sampling methodology for Fourier transform infrared (FTIR) spectroscopy is a powerful tool of forensic examination. ATR-FTIR spectroscopy allows for quick and easy measurement of a huge variety of samples, including liquids, solids, powders, semisolids, and pastes [1-3]. ATR-FTIR does not require sample preparation and allows trace analysis of car paints. However, visual interpretation and comparison of IR spectra is a complex task and requires highly qualified specialists. Moreover, this method is unsuitable in the case of spectral analysis of a mixture of coatings having similar or identical composition.

Chemometric analysis of ATR-FTIR data allows to expand the capabilities of forensic examination of paintwork materials and paints. Chemometrics provide a statistical estimation of how similar or dissimilar the paintwork materials are to each other in the composition. At the same time, the reliability and reproducibility of results are improved, while the personal impact of expert's measurements is minimized, thus, significantly increasing the reliability of the expert's conclusions.

The aim of this paper is identification of hidden patterns in ATR-FTIR spectral data of decorative varnishes (top layers of paintwork coatings) of cars based on principal component analysis (PCA). The PCA model for model samples (n=48) was built and clustering of paintwork coatings with different composition has been obtained.

References

[1] E. Mistek-Morabito, I.K. Lednev, *Anal. Bioanal. Chem.* **413** (2021) 2513–2522.
[2] C. Nimi, R. Chophi, R. Singh, *Forensic Chem.* **32** (2023) 100461.
[3] A. Sharma, R. Verma, R. Kumar, R. Chauhan, V. Sharma, *Microchem. J.* **159** (2020) 105504.

P08. Classification of malignant and benign formations on mouse skin tissues using spectral data and machine learning

N. Ginoyan [1], E. Boichenko [2], D. Kirsanov [2]
[1] MIR photonics LLC, Armenia, Yerevan
[2] Department of Analytical Chemistry, Institute of Chemistry, Saint Petersburg State University, Saint Petersburg, Russia

Skin cancer is one of the most common and deadly forms of cancer, affecting millions of people worldwide. Early and accurate diagnosis of skin cancer is crucial for effective treatment and prevention of mortality. However, the current methods of diagnosis, such as clinical examination and histopathological analysis, are often subjective, time-consuming, invasive, and costly. Therefore, there is a need for developing easy-to-use and noninvasive tools for skin cancer detection and classification in real time.

Machine learning (ML) and deep learning (DL) are powerful techniques that can recognize patterns, features, and anomalies in images. In recent years, many researchers have applied these techniques to the problem of skin cancer detection [1], using various types of images, such as photographs, dermoscopic images, sonographic images, confocal microscopy images, and multispectral images. These techniques can achieve high levels of accuracy, sensitivity, and specificity, comparable or superior to human experts, and can provide additional information, such as lesion segmentation and localization. It is also important to include various types of biosamples in the training set for ML and DL models, because real tumor sites are surrounded by different concomitant conditions such as inflammation, benign hyperplasia, etc.

In our study, we processed the spectral data in near-infrared region (NIR), collected from 59 paraffin blocks containing fragments of mouse skin. Each block underwent histological examination, and the spectra were captured at the points corresponding to inflammation, papilloma (benign formations), and normal skin without pathological signs. In addition, a spectral set of Ehrlich carcinoma samples, located on mouse skin (10 paraffin blocks), was included in the total dataset. The four distinct classes include normal skin, tumor, inflammation, and papilloma (112 spectra). The spectral data were also transformed using mel-spectrogrtam transformation, and mel-frequency cepstrum coefficients were calculated to decrease the dimensions of data points. We used three methods – Partial Least Squares Discriminant Analysis (PLS-DA), Neighborhood Component Analysis (NCA), and a Convolutional Neural Network (CNN) – which demonstrated satisfactory results in both spectral and image-based classifications. We compared three strategies of classification: (1) multilevel (classification of all classes at the same time, (2) binary classification of a specific class versus others, (3) binary classification between two classes. For the multilevel classification, the results of NCA and CNN were comparable (70% accuracy). For binary classifications, the results for all three methods were consistent (accuracy varied from 84 to 98%). Therefore, the combination of NIR spectroscopy and ML/DL techniques is a promising and express method of skin cancer diagnosis without sophisticated equipment.

References

[1] A. Esteva, B. Kuprel, R. Novoa, J. Ko, S. Swetter, J. Blau, S. Thrun, *Nature.* **542** (2017) 115–118.

P09. A flow photometric cell for in-line analysis of the biomass content of the microalgae *Nannochloropsis sp*.

E. Bazarnov, A. Bogomolov
Department of Analytical and Physical Chemistry, Samara State Technical University, Samara, Russia

In-line process monitoring is increasingly used in biotechnological production. Modern biotechnology is used in different areas, including the production of biofuels [1]. Microalgae are currently the most promising renewable raw materials for biofuel. The technology of biofuel production significantly differs from the technology of production of traditional fossil fuels. Unlike fossil raw, the biomass chemical composition can be influenced by the conditions of its

cultivation process [2]. The process of cultivation of microorganisms consists in accumulation of their biomass and metabolic products under the conditions of nutrient medium. Optimization of the cultivation process allows to influence on the selectivity of biotechnological proceed and obtain the maximum content of target chemical compounds in biomass composition, but it requires an in-line analytical control of its main parameters [3]. One of the critical indicators of biotechnological processes is the biomass content in the reactor. In photosynthesis-based processes the biomass is represented by microalgae or bacteria. The stages of photosynthesis take from 10^{-5} to 10^3 seconds, and cell division stages and biomass accumulation can be measured in minutes [4]. So, in such cases, traditional biomass determination methods, such as cell counting chambers (Goryaev chamber, Fuchs-Rosenthal chamber), are too time-consuming and labor-intensive, and are hardly suitable for in-line control. This research work solved the task of the in-line determination of the biomass in the microalgae *Nannochloropsis sp*. cultivation using an optical sensor.

A photometric flow cell for the in-line determination of the biomass content in a microalgae *Nannochloropsis sp.* cultivation has been constructed, integrated into the photobioreactor and tested. The biomass content in photometric measurements does not generally obey Beer's law, and biomass cells are able not only to absorb but also to scatter the light radiation. Nevertheless, the results of tests show the linear dependence of absorbance on the biomass content. This result can be explained by an assumption that light scattering effects observed by the flow cell also have a linear character in the considered concentration range.

The results obtained have shown the linear dependence of the detected optical density on the biomass content. Therefore, the concept of the in-line determination of the biomass content in similar cultivation processes using a photometric sensor based on a single laser light source 650 nm has been proven [5].

References

[1] U. Muhammad, I. Shamsuddin, A. Danjuma, S. Ru, U. Dembo, *Petrol. Sci. Eng.* **2** (2018) 44–49.

[2] N. Mishra, P. Mishra, E. Gupta, P. Singh, *J. Microbiol. Biotech. Food Sci.* **12** (2023) e3632.

[3] P. Benner, S. Effenberger, L. Franzgrote, T. Kurzrock-Wolf, K. Kress, D. Weuster-Botz, D. *Biochem. Eng. J.* **164** (2020) 107749.

[4] L.-Q. Dong, K. Niu, S.-L. Cong, *Chem. Phys. Let.* **432** (2006) 286–290.

[5] E. Bazarnov, A. Bogomolov, *Engineering Proceedings*, **48** (2023) 23.

P10. Multidimensional Calibration on Overlapping Spectral Intervals

E. Ignatiev [1], A. Bogomolov [1, 2]

[1] Samara State Technical University, Samara, Russia

[2] art photonics GmbH, Berlin, Germany

This article explores the use of overlapping spectral intervals in the development of optical multisensor systems (OMS) [1]. The objective is to investigate whether this approach enables efficient analysis of mixtures and identification of substances contained within them.

The research involved the examination of synthetic two-component data and real mixtures of Co(II), Ni(II), and Cu(II). A full factorial experiment was employed for the synthetic data, while a diagonal experimental design based on a Latin hypercube was used for the real mixtures. Sixteen equidistant wavelength points were selected within the visible range. All possible sets of intervals were created on these points, with the number of intervals equal to the number of components in the mixtures. Models were built for each interval set using the partial least squares method (PLS), based on the averaged spectral values of substances within those intervals. The efficiency analysis of the constructed model was conducted using the obtained R^2cv values.

The analysis of the constructed models revealed that a significant number of interval sets produced better results compared to the expert-based approach. Furthermore, there were interval sets for which the models showed better performance than analyzing the full spectrum. The most accurate results were obtained from models constructed using overlapping intervals.

Creating OMS on overlapping spectral intervals proves to be highly effective. Exhaustive determination of the most efficient intervals through complete enumeration provides a coarse estimation, particularly with a low number of points, and becomes computationally intensive with a large number of points. Therefore, modification of the interval selection algorithm is required.

References

[1] A. Bogomolov, *Sensors*. **21** (2021) 3541.

P11. Diagnosis of halitosis by the multisensor gas analyzer Aramos-7

I. Jahatspanian

Joint-Stock Company "NPO "Pribor", St. Petersburg, Russia

Chronic tonsillitis ranks first as the cause of bad breath (halitosis). The basis of halitosis is the presence of volatile compounds with an unpleasant smell in the exhaled air from the oral cavity. Most often, these are volatile sulfur compounds (hydrogen sulfide, methyl mercaptan, dimethyl sulfide, organic acids (acetic, propionic), amines, and volatile aromatic compounds (indole, skatol, cadaverine). Most of these volatile compounds are produced by bacteria during the anaerobic metabolism of food protein residues in the oral cavity. Sulfur compounds, particularly hydrogen sulfide, are poorly dissolved in saliva; non-dissolved hydrogen sulfide evaporates and exhales, and dissolved ionized hydrogen sulfide reduces the redox potential of saliva and participates in the anaerobic breakdown of protein. Few clinics are involved in the diagnosis and treatment of halitosis. Despite the fact that bad breath is a fairly frequent complaint, patients do not always know which specialist to contact. Assistance for halitosis, which is associated with chronic tonsillitis is provided by an otorhinolaryngologist after a comprehensive examination of the patient, including halimetry in addition to standard methods. The organoleptic method for diagnosing halitosis is based on the sense of smell and is a subjective one. The method requires preparation in the form of exclusion of taking antibiotics, certain types of food, abstinence from smoking, rinses and cosmetics, which often requires repeated visits. An instrumental method for evaluating halitosis was proposed by Linus Pauling in 1971. He used gas-liquid distribution chromatography. Subsequently, sensor devices that convert information about gases into an electrical signal were used to assess halitosis. The latest development for evaluating halitosis is an "electronic nose" that can recognize complex gas mixtures. "Electronic noses" use odour-sensitive gas sensors.
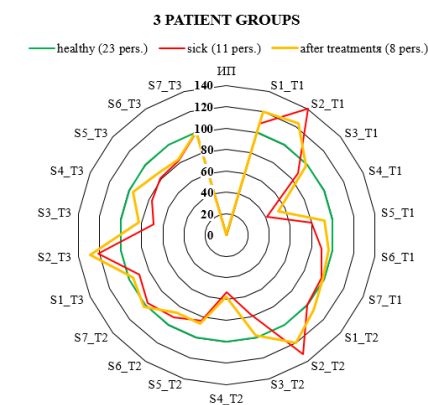
ARAMOS -7 (ARoma Analyzer by Metal Oxide Sensors):
7 sensors at 3 temperatures = 21 information indicators

5 classifiers:
- k-Nearest Neighbors (kNN)
- Linear Discriminant Analysis (LDA)
- Random Forest Classifier (RFC)
- Logistic Regression (LR)
- Support Vector Machine (SVM)

CLINICAL SENSITIVITY AND CLINICAL SPECIFICITY > 85%

P12. Multivariate calibration transfer methods in multisensor potentiometry

M. Saveliev, D. Kirsanov
Institute of Chemistry, St. Petersburg University, St. Petersburg, Russia

In analytical practice, the problem of deterioration of the predictive ability of calibration models, in particular multivariate models, often occurs. This can be caused by various factors: the changes in the chemical composition of samples, that were not taken into account in the original calibration model, the instrumental drift, etc. This problem can be solved by building a new calibration model that accurately describes the system under the influence of current factors, but often it is very time and resource consuming or impossible at all. For example, in the quantitative determination of lanthanides in the reprocessing of spent nuclear fuel, it is impossible to calibrate the multisensor array on real samples due to numerous safety limitations. An effective solution to this problem can be the application of calibration transfer methods. They make it possible to correct the calibration model based on the data of several measurements under new conditions with the influence of previously unaccounted interfering factors without the need for complete recalibration. In the field of infrared spectroscopy the methods of dynamic orthogonal projection (DOP) [1], its version without a teacher (uDOP) [2], domain-invariant partial least square regression (diPLS) [3] and domain-Invariant Iterative Partial Least Squares (DIPALS) algorithm [4] were successfully applied for such purposes. The potential of these methods for other instrumental techniques, e.g. for multisensor potentiometry remains unexplored.

This presentation will give the results of the application of calibration transfer methods in the determination of the content of Nd^{3+}, Eu^{3+} and Er^{3+} cations in nitric acid solutions using an array of cross-sensitive potentiometric sensors. Barium and cesium nitrate at a constant concentration of 10-3 M were added to standard lanthanide nitrate solutions to simulate the occurrence of interfering factors compared to the original calibration model. The results of application of transfer methods show that their use allows to reduce RMSEP values and to adapt the calibration model for lanthanide determination under new conditions. The optimal number of measurements for each transfer method is also selected and their comparative characterization is given.

References

[1] M. Zeaiter, J.M. Roger, V. Bellon-Maurel, *Chemom. Intell. Lab. Syst.* **80** (2006) 227–235.

[2] V. Fonseca Diaz, J.-M. Roger, W. Saeys, *Anal. Chim. Acta.* **1225** (2022) 340154.

[3] R. Nikzad-Langerodi, W. Zellinger, E. Lughofer, S. Saminger-Platz, *Anal. Chem.* **90** (2018) 6693–6701.

[4] R. Nikzad-Langerodi, W. Zellinger, S. Saminger-Platz, B.A. Moser, *Knowl. Based Syst.* **210** (2020) 106447.

P13. Getting insights into chromatographic properties of HILIC and mixed-mode homemade stationary phases with PCA and clustering

A. Shemiakina[1], N. Chikurova[1], A. Gorbovskaia[1], A. Uzhel[1], A. Chernobrovkina[1]
[1]Lomonosov Moscow State University, Moscow, Russia

Stationary phase type plays a key role in retention mechanism in hydrophilic interaction liquid chromatography (HILIC). Development of their wider variety is relevant for obtaining alternative selectivity and better separation. Among the most popular stationary phases are silica-based ones functionalized with diverse polar groups and mixed-mode adsorbents able to work in two or three HPLC modes (HILIC, reversed-phase HPLC, and ion chromatography).

Nowadays there is no unified procedure for characterization and comparison of HILIC columns. Characterization can be performed by means of special tests (e.g. the Tanaka test [1]) in which several types of selectivity are observed. Alternatively, comparison of columns can be carried out by obtaining retention factors for compounds with different properties such as amino acids, sugars, water-soluble vitamins, etc. Only a few works applied chemometric approaches such as principal component analysis (PCA) or cluster analysis to the retention

factors or selectivity data for the comparison and grouping of stationary phases used in HILIC according to their chromatographic properties [2,3]. Chemometric-assisted approaches could also be useful for comparing homemade adsorbents with commercially available ones [2], for confirming successful substrate modification after the synthesis [2], and for evaluating retention mechanism [4].

In this work, 70 homemade columns including amide-, zwitterionic-, and polymer-bonded silica-based and polystyrene-divinylbenzene-based mixed-mode stationary phases were characterized in terms of the Tanaka HILIC test (7 selectivity parameters). Also, the retention factors for 33 polar compounds including sugars, nucleobases, nucleosides, carboxylic acids, and water-soluble vitamins were obtained. The same chromatographic experiments were conducted for commercially available amino phases. Additionally, the literature data for commercially available columns with the same types of functionalities were added in the dataset.

Principal component analysis was used to obtain two separate sets of linearly uncorrelated dimensions based on Tanaka test values and the other 33 retention factors, respectively. Several approaches to clustering were applied to reveal the groups of columns with similar characteristics. Some outlier columns both with polymer and monomer functional layers were detected. They were located far from their assumed groups despite the hypothesis that columns with similar chemical composition should exhibit similar properties. The results of the clustering and the visual representations of the dataset were also compared between the Tanaka test-based and retention factors-based ones.

The results of the PCA and clustering allowed us to get some insights into the similarities in properties between the groups of stationary phases of different functional layer structure, reveal the correlations between the retention of some polar compounds, and select the analytes useful for grouping of columns and characterization. The conclusions may contribute to developing a more precise algorithm for HILIC columns studying and comparison, which is useful for a more accurate column choice for the particular analytical task.

Acknowledgment

References

[1] Y. Kawachi, T. Ikegami, H. Takubo, Y. Ikegami, M. Miyamoto, N. Tanaka, *J. Chromatogr. A.* **1218** (2011) 5903–5919.

[2] A. Gargano, T. Leek, W. Lindner, M. Lämmerhofer, *J. Chromatogr. A.* **1317** (2013) 12–21.

[3] M. Grübner, A. Dunkel, F. Steiner, T. Hofmann, *Anal. Chem.* **93** (2021) 12565–12573.

[4] N. Dinh, T. Jonsson, K. Irgum, *J. Chromatogr. A.* **1218** (2011) 5880–5891.

P14. Various aspects of quantitative structure-retention relationships for ionic liquids-based stationary phases

A.Yu. Sholokhova, D.D. Matyushin

A.N. Frumkin Institute of Physical Chemistry and Electrochemistry, RAS, Moscow, Russia

Quantitative structure retention relationships (QSRR) link the structure of a molecule and a value representing chromatographic retention, e.g., a retention index. The retention value is represented as a function (usually linear) of a number of molecular descriptors (MD), i.e., quantities characterizing the structure of the molecule. In this case, the most relevant MD are selected and an equation is constructed that relates them to retention. QSRR can be used for retention prediction that can be further used in chemical analysis. The other use of QSRR is the investigation of retention mechanism and stationary phase (SP) using a set of selected MD and their relative importance [1]. However, the reliability of such outputs is highly questionable. The authors of such works often draw some conclusions about retention by repeating the MD selection procedure once for a fixed set of molecules [2] or by comparing different SP using different sets of molecules for each of them [1]. We considered the issue of reproducibility of various methods for selecting MD for QSRR and studied how a relatively small alteration of a data set changes the relative importance of MD and the set of selected MD. Several methods of MD selection were considered: stepwise selection, L1-regularized regression, Boruta algorithm, genetic algorithm and others. It was shown that the results strongly depend on the used method, and a small change of the data set (for example, the random removal of ~10% of molecules from the set) severely affects the results.

Ionic liquids (IL) are widely used as SP in gas chromatography due to their high polarity and high thermal stability. There are many works devoted to their retention behavior, however, there were no relatively large and diverse data

sets of retention times or indices of molecules of various classes for IL-based SP. We created and published the first such data sets (120–160 molecules) for three imidazolium-based IL. These data sets were used in the QSRR study described above and can be used in further works on QSRR and machine learning. We also developed retention index prediction models (both deep learning-based and an interpretable linear model) for polyethylene glycol used as SP. These models were trained using a large data set (NIST 17 database) and are quite accurate. We demonstrated that the retention index predicted for polyethylene glycol can be used as MD for QSRR for IL-based SP. The use of this MD in conjunction with other MD significantly improves the accuracy of QSRR.

Open source software (https://github.com/mtshn/chereshnya) with a graphical user interface was created to perform such research. Several methods for selecting MD were implemented, and MD of various types were calculated.

References

[1] J. Yan, D. Cao, F. Guo, L. Zhang, M. He, J. Huang, Q. Xu, Y. Liang, *J. Chrom A.* **1223** (2012) 118–125.

[2] C. Rojas, P. Duchowicz, P. Tripaldi, R. Diez, *J. Chrom A.* **1422** (2015) 277–288.

P15. Potentiometric sensor arrays based on fluoropolymer membranes modified with polyaniline for the sulfonamide drug analysis

A. Yelnikova [1], A. Parshina [1], P. Yurova [2], I. Stenina [3], O. Bobreshova [1], A. Yaroslavtsev [3]

[1] Voronezh State University, Voronezh, Russia

[2] National Research University Higher School of Economics (HSE University), Moscow, Russia

[3] Kurnakov Institute of General and Inorganic Chemistry RAS, Moscow, Russia

The drug quality control is one of the tasks solving by multisensory systems. The use of cross sensitive sensors in combination with chemometrics is advantageous for the analysis of combination drugs or drugs prone to degradation. Multidimensional regression analysis is well-known for the quantitative analysis of complex media, novel approaches are also studied [1].

Potentiometric multisensory systems were developed for the analysis of sulfacetamide (SAA) drugs, including ones subjected to UV-degradation with

formation of sulfanilamide (SA), and combination drugs of sulfamethoxazole (SMX) and trimethoprim (TMP). Perfluorosulfonic acid (PFSA) membranes modified with polyaniline (PANI) were used as the sensitive materials of the sensors with inner reference solutions [2,3]. Some PFSA/PANI membranes were obtained by in situ oxidative polymerization. The varied factors were the order of the membrane treatment with the solutions of the monomer and oxidant, the precursor concentrations, and the conditions for the composite membrane hydrothermal treatment. The PFSA/PANI membranes were also prepared by casting procedure from solutions containing specified amount of both polymers. The calibration of the sensors was performed in SAA+NaOH (pH 5.76–11.19) and SAA+SA+NaOH (pH 4.76–10.70) solutions with the component concentrations ranged from $1.0 \cdot 10^{-4}$ to $1.0 \cdot 10^{-2}$ M, as well as in SMX+TMP (pH 4.53–6.84) solutions with the component concentrations ranged from $1.0 \cdot 10^{-5}$ to $1.0 \cdot 10^{-3}$ M. In the test solutions, SAA and SMX were in the anionic forms, SA was in the uncharged form, while TMP was both in the form of cations and molecules. The concentration matrix of the calibration solutions was chosen taking into account the pairwise correlation and multicollinearity of the varied factors. The developed multisensory systems were tested in the analysis of the Sulfacyl sodium-SOLOpharm eye drops (Grotex, Russia), including ones subjected to the forced UV-degradation, and the Biseptol tablets (Adamed Pharma, Poland).

It was revealed that the potentiometric sensors' sensitivity to sulfonamide anions increased with increasing diffusion permeability of the membranes and dopant concentration in their surface layer. Hydrothermal treatment of the PFSA/PANI membranes changed the conditions for the sorption of ions of the drug substances depending on their size and hydrophilicity. The prepared PFSA/PANI membranes characterized by sufficiently different distribution of the cross sensitivity to the analytes. That allowed us to choose sensors with low correlation between their responses for the multisensory system organization. The relative errors and relative standard deviations of the determination of the active ingredients, their degradation products, and excipients were no greater than 5% and 8%, respectively. The re-estimated analytical characteristics of the sensors used in the multisensory systems did not statistically change after 1 year of their use due to the selected regeneration and storage conditions. The reference method for the drug analysis was spectrophotometry in combination with multidimensional regression analysis. The last was used because of lack of

discrimination of the analyte absorbance peaks and significant difference in their intensities.

References

[1]   R. Belugina, A. Senchikhina, S.Volkov, A. et al. *Anal. Methods.* **14** (2022) 3064–3070.

[2]   A. Parshina, A. Yelnikova, T. Titova, et al, *Polymers.* **14** (2022) 2545.

[3]   A. Parshina, A. Yelnikova, T. Kolganova, et al, *Membranes*. **13** (2023) 311.

P16. Antiradical/Antioxidant Activity of Folic Acid. A DFT Study

Z.H. Manukyan

A.B. Nalbandyan Institute of Chemical Physics, NAS Republic of Armenia, Yerevan

The study of the antiradical ability of the multifunctional water-soluble vitamin of the B group (vitamin B9) – folic acid (FA) (Figure) is an urgent task.

The main aim of the research is to investigate the antioxidant potency of FA and to identify the reaction sites responsible for its antioxidant activity at the molecular level. The obtained data gives necessary information to improve pharmaceutical forms of FA and to increase the efficiency of medical therapy.
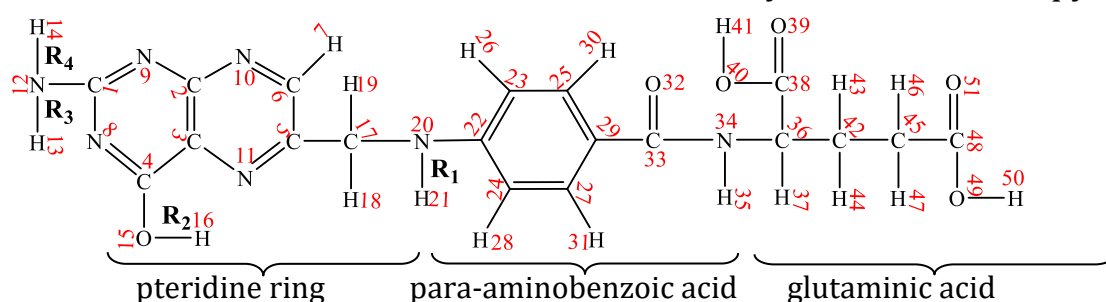

Figure. Chemical structure of folic acid (FA) with numbered atom sites.

In the present work, reaction sites of FA responsible for its antiradical/antioxidant activity were determined by quantum chemical calculations in vacuum.

Phenolic and aromatic amines antioxidants scavenge free radicals via phenolic O-H and aromatic amine N-H groups mainly through three distinct mechanisms yielding the same products: One-step transfer of a hydrogen atom: HAT (Hydrogen Atom Transfer); A sequential two-step process: SPLET (Sequential Proton-Loss Electron-Transfer); A single electron transfer followed by proton transfer: SET-PT (Single Electron Transfer-Proton Transfer). The thermodynamics of the mechanisms of free radical scavenging reactions by O-H

and N-H bonds of FA are largely determined by the enthalpies (H) of the processes:

(1) $BDE = H(A^\bullet) + H(H^\bullet) - H(AH)$ BDE (bond dissociation enthalpy)

(2) $PA = H(A^-) + H(H^+) - H(AH)$ PA (phenoxide and amide anions proton affinity)

(3) $ETE = H(A^\bullet) + H(e^-) - H(A^-)$ ETE (electron transfer enthalpy)

(4) $IP = H(AH^{+\bullet}) + H(e^-) - H(AH)$ IP (ionization potential)

(5) $PDE = H(A^\bullet) + H(H^+) - H(AH^{+\bullet})$ PDE (proton dissociation enthalpy)

The reactions enthalpies were calculated using the hybrid methods B3LYP and M06-2X of the density functional theory (DFT) (6-31G(d,p) basis set). The obtained parameters for three mechanisms are given in the table, where R1, R2, R3 and R4 are the potential antiradical reaction sites of FA.

Table. Thermochemical parameters for aromatic three antioxidant mechanisms for the free radical scavenging reaction by phenolic O-H and amine N-H bonds of the folic acid (FA) molecule.

Reaction Sites of FA	HAT		SET-PT				SPLET			
	BDE (kcal/mol)		IP (kcal/mol)		PDE (kcal/mol)		PA (kcal/mol)		ETE (kcal/mol)	
	B3LYP	M06-2X	B3LYP	M06-2X	B3LYP	M06-2X	B3LYP	M06-2X	B3LYP	M06-2X
R1	86.38	91.09	154.22	166.25	246.84	237.25	352.94	350.24	48.10	53.25
R2	95.62	102.44			256.07	248.60	327.16	330.98	83.12	83.86
R3	101.84	106.39			262.30	252.55	348.18	351.44	68.32	67.35
R4	102.87	107.49			263.33	253.65	349.88	353.45	67.65	66.44

An analysis of the data in the table shows that the energy efficiency of free radical scavenging reaction by phenolic and aromatic amine bonds increases in the order:

R4 < R3 < R2 < R1.

The obtained data confirm the antiradical properties of FA. The results show that FA represents as a powerful antioxidant via the HAT mechanism. Meanwhile, high BDE values for the amine N-H group of the pteridine ring (R3 and R4) from 101 to 107 kcal/mol indicate almost absence of antiradical activity via the HAT mechanism. High ionization potentials (IP) and proton affinities (PA) proves that a free radical scavenging reaction by the SET-PT and SPLET mechanisms in gas phase is unlikely.

P17. Universal Calibration Models in X-ray Fluorescence Analysis

Z. Selivanovs, D. Kirsanov, V. Panchuk
Institute of Chemistry, Saint Petersburg University, Russia

Quantitative determination of elements in X-ray fluorescence analysis (XRF) is complicated by the influence of matrix effects. Their type, intensity and contribution to the spectra are very different for samples of different chemical composition. Two main approaches are used to solve this problem: the method of fundamental parameters and methods based on the construction of regression models. The fundamental parameter method has relatively low accuracy and is often considered as a semi-quantitative analysis method. The second group of methods to account for matrix effects (intensity correction, concentration correction, standard-background method, etc.) requires the construction of a calibration model within one specific type of object. This significantly increases the amount of actions needed to be done for XRF when objects of different composition are analyzed, since each type of object requires its own separate calibration model. Recently, a new method of accounting for matrix effects in XRF has been proposed, which is based on a combination of the method of intensity correction and the partial least squares regression [1]. One of the advantages of this method is the possibility of taking into account an unlimited number of correction factors, which, theoretically, can give the possibility of constructing a single calibration for different types of samples. The presentation will show the results of studying this possibility on the example of building universal calibration models for the determination of a number of elements in samples of ores and steels. The effectiveness of the proposed approach is compared with the fundamental parameter method. In addition, the results of application of nonlinear multivariate regression method for construction of universal calibration models KRLS (Kernel-Regularized Least Squares) will be presented [2].

Acknowledgment

References

[1] S. Aidene, M. Khaydukova, S. Savinov, V. Semenov, D. Kirsanov, V. Panchuk, *Spectrochim. Acta B* **193** (2022) 106452.
[2] Selivanovs, Z.; Panchuk, V.; Kirsanov, D. *Appl. Sci.* **13** (2023) 5415.

Participants

Bian Xihui
Tiangong University
Professor
Tianjin, China
bianxihui@163.com

Bogomolov Andrey
Samara State Technical University
Head of department
Samara, Russia
a.bogomolov@mail.ru

Boichenko Ekaterina
Saint Petersburg State University
Senior researcher
Saint Petersburg, Russia
ekaterina.boichenko@inbox.ru

Burmistrova Natalia
Saratov State University
Professor
Saratov, Russia
naburmistrova@mail.ru

Csontos Máté
Science Port Kft.
Research scientist
Debrecen, Hungary
info@scienceport.hu

Elek   János
Science Port Kft.
General director
Debrecen, Hungary
elek@scienceport.hu

Ginoyan Narek
MIR photonics LLC
Student
Erevan, Republic of Armenia
narek.ginoyan.1000@gmail.com

Jahatspanian Igor
Joint-Stock Company "NPO "Pribor"
Leading engineer
Saint Petersburg, Russia
drjie@mail.ru

Kalambet Yuri
Ampersand Ltd.
General director
Moscow, Russia
kalambet@ampersand.ru

Khaydukova Maria
FSBSI Institute of Experimental Medicine
Senior researcher
Saint Petersburg, Russia
khaydukova.m@gmail.com

Khrisanfov Mikhail
Lomonosov Moscow State Univertsity
Research scientist
Moscow, Russia
khrisanfovmike@gmail.com

Kirsanov Dmitry
Saint Petersburg State University
Professor
Saint Petersburg, Russia
d.kirsanov@gmail.com

Lvova Larisa
University "Tor Vergata"
Professor
Rome, Italy
Larisa.Lvova@uniroma2.it

Manukyan Zaruhi
A.B. Nalbandyan Institute of Chemical
Physics of NAS RA
Research scientist
Erevan, Republic of Armenia
zara@ichph.sci.am

Matyushin Dmitriy
A.N. Frumkin Institute of Physical Chemistry
and Electrochemistry, RAS
Research scientist
Moscow, Russia
dm.matiushin@mail.ru

Panchuk Vitaly
Saint Petersburg State University
Research scientist
Saint Petersburg, Russia
vitpan@mail.ru

Parastar Hadi
Sharif University of Technology
Associate professor
Tehran, Iran
h.parastar@sharif.edu

Pomerantsev Alexey
Semenov Federal Research Center for
Chemical Physics
Principal researcher
Moscow, Russia
forecast@chph.ras.ru

Rodionova Oxana
Semenov Federal Research Center for
Chemical Physics
Leading researcher
Moscow, Russia
oksana@chph.ras.ru

Rusanova Tatiana
Saratov State University
Professor
Saratov, Russia
tatyanarys@yandex.ru

Rutledge Douglas
Muséum National d'Histoire Naturelle
Professor Emiritus
France
Douglas.Neil.Rutledge@gmail.com

Samokhin Andrey
Lomonosov Moscow State
Research scientist
Moscow, Russia
andrey.s.samokhin@gmail.com

Saveliev Mikhail
Saint Petersburg State University
Student
Saint Petersburg, Russia
mishasavelev@list.ru

Selivanovs Zahars
Saint Petersburg State University
Student
Saint Petersburg, Russia
zahar.selivanov.rzg@gmail.com

Shemiakina Anna
Lomonosov Moscow State
PhD student
Moscow, Russia
a.o.shemyakina@yandex.ru

Sholokhova   Anastasia
A.N. Frumkin Institute of Physical Chemistry
and Electrochemistry RAS
Research scientist
Moscow, Russia
shonastya@yandex.ru

Shuba Anastasiia
Voronezh State University of Engineering
Technologies
Associate professor
Voronezh, Russia
shuba@vsuet.ru

Stavrianidi   Andrey
Chemistry Department, Lomonosov Moscow
State University
Associate professor
Moscow, Russia
stavrianidi.andrey@gmail.com

Surkova Anastasiia
Samara State Technical University
Associate professor
Samara, Russia
melenteva-anastasija@rambler.ru

Sushkov Nikolai
Lomonosov Moscow State University
Leading specialist
Moscow, Russia
nikolaisushkov@yandex.ru

Tretyakov Alexey
Saint Petersburg State University
Student
Saint Petersburg, Russia
hazzi5055@gmail.com

Tsakovski Stefan
University of Sofia
Professor
Sofia, Bulgaria
STsakovski@chem.uni-sofia.bg

Yelnikova Anastasia
Voronezh State University
Research scientist
Voronezh, Russia
anastasia_elnikova@outlook.com

Yuskina Ekaterina
Saint Petersburg State University
Student
Saint Petersburg, Russia
yuskina.k@yandex.ru