

ШЕСТОЙ МЕЖДУНАРОДНЫЙ СИМПОЗИУМ “СОВРЕМЕННЫЕ МЕТОДЫ АНАЛИЗА МНОГОМЕРНЫХ ДАННЫХ”

Очередной симпозиум по хеометрике “Современные методы анализа многомерных данных” (WSC-6), прошедший в Казани, в феврале 2008 года, открыл второй этап международной программы Дружбаметрика [1]. Увеличившееся число участников подтвердило интерес к серии WSC-конференций как со стороны российских, так и зарубежных ученых. Традиционно симпозиум проходил на английском языке. Мероприятие проводилось Казанским государственным архитектурно-строительным университетом и Российским хеометрическим обществом, при участии ГЕОХИ РАН им. В.И. Вернадского и университета Ольборг, Эсбьерг (Дания). Российский фонд фундаментальных исследований поддержал проведение WSC-6 (Грант № 08-03-06002).

Симпозиум предваряла двухдневная школа по хеометрике, проводимая на русском языке для студентов, аспирантов и молодых ученых. Более 60-ти слушателей познакомились с основными понятиями и подходами, применяемыми при обработке результатов современных экспериментов, а так же получили первичные практические навыки анализа экспериментальных данных. Наиболее активные участники, представители университета “Дубна”, были награждены призом: программой Unscrambler [2], а так же книгой [3].

Непосредственно симпозиум проходил с 18–22 февраля на турбазе “Глубокое озеро” в пригороде Казани. Общий подход к анализу и интерпретации больших массивов данных был представлен лекцией проф. К. Эсбенсена (Дания) “Chemometric data analysis – a holistic process with many elements”. Такой подход начинается с постановки и планирования эксперимента, применения теории пробоотбора, накопления данных и отбора переменных, а заканчивается построением математической модели с проверкой ее качества и предсказательной способности. В настоящее время разрабатываются методы, позволяющие обрабатывать сложные гибридные эксперименты, как это было представлено в докладе А. Янкевича (Латвия) “Metabolomic studies of experimental and clinical diabetic urine samples by ^1H NMR and LC/MS spectroscopy”, либо строить модели для совместной обработки разнородных экспериментов, как в докладе Ф. Марини (Италия) “Coupling of spectroscopic, thermoanalytic and chemometric methods for the resolution of enantiomeric mixtures”. Один из возможных подходов, используемых в случае, когда один и тот же образец/явление изучался с

помощью различных экспериментов, был представлен в лекции проф. Д. Рутледжа (Франция) “Outer Product Analysis (OPA) as a method to study the relations between sets of variables measured on the same individuals”, в которой эффективность теоретических приемов проверялась на примерах обработки экспериментальных данных, таких как ИК или ближняя ИК-спектроскопия, совместно с рентгеновской дифракцией или ЯМР-спектроскопией. Преимуществом ОРА подхода является возможность обработки развернутых данных классическими проекционными методами, а также учет многомерности структуры, например, методом PARAFAC. Усовершенствованию вычислительных процедур при работе с многомерными данными был посвящен доклад Д. Савостьянова (Россия) “Fast computation of CANDECOMP-PARAFAC and Tucker decompositions”

К чисто классическим аналитическим приложениям можно отнести работы, посвященные применению многомерной калибровки в атомно-эмиссионной спектрометрии (С. Павлов, Е. Шабанова, Россия), оценке неопределенности в экспериментах на растворимость (М. Пааккунайнен, Финляндия), применению ПЛС-регрессии для определения состава сложной смеси с помощью спектрофотометрии (И. Власова, Россия), распознаванию фальшивых лекарств с помощью БИК-спектроскопии (О. Родионова, Россия). Были представлены работы посвященные применению сенсоров: химических (С. Моржухина, Россия) “The data processing method of multisensor system for natural water monitoring” и биологических (А. Решетилов, Россия) “Elements of pattern recognition theory in the analysis of biosensor signals”). Показано, что использование многомерного подхода может существенно повысить эффективность обработки таких данных.

Современные тенденции в хеометрике были представлены в двух актуальных направлениях. Первое связано с использованием хеометрических методов для анализа изображений и гиперспектральных данных. На конференции это направление было представлено лекцией С. Кучерявского (Дания) “Image analysis in chemometrics”, и прикладной работой И. Беляева “Estimation of age in forensic medicine using multivariate data analysis”. Второе направление – это аналитический контроль процессов (Process Analytical Technology – PAT). В PAT наиболее ярко проявились тенденции и перспективы развития общего подхода, объеди-

няющего физико-химические эксперименты, проводимые в режиме реального времени, с математическими методами анализа многомерных данных. Теоретические аспекты были представлены в лекции проф. А. де Хуан (Испания) "Advanced process modeling with Multivariate Curve Resolution". Аппаратные и программные средства для решения задачи в области аналитического контроля процессов были представлены в докладе А. Богомоллова (Германия) "PAT-enabling hard- and software". Решению частных прикладных задач в этой области были посвящены доклады по контролю процессов образования биогаза (М. Мадсен, Дания) и контролю процесса грануляции в фармацевтической промышленности (С. Матеро, Финляндия).

В рамках традиционного направления математического моделирования можно выделить работы по описанию и изучению структур молекул по их спектральным данным. В лекции проф. М. Эляшберга (Россия) "Chemometrical methods in expert systems for the molecular structure elucidation" была представлена общая идеология построения экспертных систем для выявления структуры молекул по их спектрам, сформулированы системы "аксиом", из которых выводятся структурные формулы. Была представлена система Structure Elucidator, способная идентифицировать крупные органические молекулы (более 100 скелетных атомов) по 2М ЯМР спектрам. Продолжением этой темы был доклад К. Блинова (Россия) "Prediction of NMR Chemical Shifts. A Chemometrical Approach", посвященный разработке метода для предсказания химических сдвигов. Предсказанию молекулярной структуры по ИК-спектрам в приложении к новому классу полимерных молекул – дендримерам, была посвящена лекция проф. В. Фурера (Россия) "Molecular structure and IR spectra of phosphorus dendrimers by DFT calculations". Оценке погрешностей прогнозирования в безэталонном молекулярном анализе, а так же общеприкладным вопросам природы и законам распределения погрешностей посвящен доклад проф. В. Деметьева (Россия) "Computer methods for evaluation of the confidence intervals of molecular modeling prediction". Математический аппарат хеометрики развивается, откликаясь на современные требования усложняющейся экспериментальной базы и все более разнообразные области применения. Новые методы оценки областей принятия решений, используемых в задачах классификации, PAT, оценки выбросов при калибровке были представлены в работе А. Померанцева (Россия) "Critical levels in

projection techniques". Новые подходы при построении моделей сложных, многостадийных процессов представил проф. А. Хоскюльдсен (Дания) в лекции "Multi-block and Path Methods – The next generation of chemometric methods".

Победителем в конкурсе молодых ученых был признан Кирилл Блинов, получивший годовую подписку на журнал "Chemometrics and Intelligent Laboratory Systems". Авторы, представившие стендовые доклады, участвовали в отдельном конкурсе, победителем которого стала работа Н. Рюминой (Самара, Россия) "IR spectroscopy and PCA applied for polymer sorting in the domestic waste".

В симпозиуме приняло участие более 100 человек из 13 стран: России, Германии, Дании, Италии, Испании, Польши, Румынии, США, Турции, Франции, Финляндии и Южной Африки. Было много молодых людей: студентов и аспирантов. Широта тематики, представленной на симпозиуме, и разнообразие специальностей участников (химики, физики, математики, инженеры) подчеркивают живой интерес к хеометрике и необходимость в применении ее методов. Участники отмечали полезность таких конференций, позволяющих обсуждать междисциплинарные подходы и методы, разрабатываемые на стыке различных научных дисциплин.

Тезисы докладов изданы отдельным сборником [4], а презентации лекций и устных докладов размещены на сайте конференции [5]. По материалам симпозиума будет издан специальный выпуск журнала Chemom. Intell. Lab. Syst. (IF 2.450, Elsevier) под редакцией А. Померанцева и Р. Таулера, аналогично номеру, посвященному предыдущему симпозиуму WSC-5 [6].

СПИСОК ЛИТЕРАТУРЫ

1. <http://rcs.chph.ras.ru/drushba.htm>.
2. The Unscrambler, <http://www.camo.com/rt/Products/Unscrambler/unscrambler.html>
3. *Esbensen K.H.* Multivariate Data Analysis – In Practice 4-th Ed., CAMO, 2000.
4. 6-ой международный симпозиум "Современные методы анализа многомерных данных", 18–22 февраля 2008 г., Казань, КГАСУ, 2008, 68 с.
5. <http://chemometrics.ru/drushbametrics/wsc6/>
6. Chemom. Intell. Lab. Syst. 88. № 1. 2007.

А.Л. Померанцев