Technical Programme
Fundamentals of multivariate data analysis
for chemical and biological sensing

Introduction
The School explores the nature of data generated by existing chemical sensors and analytical instrumentation, such as mass spectrometers, optical spectrometers and time of flight ion mass spectrometers, as well as an introduction to more recent protein and DNA sequencing chips.

We will cover such concepts as data exploration, feature extraction and feature selection techniques currently used. We will also look at linear and non-linear methods, supervised and unsupervised techniques, and parametric and non-parametric techniques. The School consists of lectures in the morning and the laboratories in the evening during which attendees can apply the techniques on standard databases.

The School is ideal for anyone with an interest in data processing and new to the field; for example, PhD students, researchers, technologists and industrialists. The data analysis does not need a high performance computer and can be carried out on a basic laptop or personal computer running the windows operating system.

For further information visit: www.olfactionsociety.org/winterschool.html

Data sets
The laboratory sessions will employ four realistic data sets, chosen to illustrate the types of multivariate techniques available and required by the differing nature of the sensing methods. They will be presented in a standard XML format and will be available to other researchers from the ISOCS website after the course.

I. Chemical sensor array, simple components, static model, linearly separable classes.
II. Mass spectra data, simple components, static model, linearly separable.
III. Chemical sensor array, complex data-set, static model, non-linear class separation so needs ANN.
IV. Chemical sensor array, simple components, time-series with significant drift, linear filters applied (e.g. auto-regression with moving average).

School Technical Director:
Prof Julian Gardner  University of Warwick, UK
Dr Udo Weimar  University of Tübingen, Germany
Prof Krishna Persaud  University of Manchester, UK
Prof Agustin Gutierrez  University of Barcelona, Spain
Dr Jan Mitrovics  JLM Innovation, Germany

Lecturers:
Prof Julian Gardner  University of Warwick, UK
Dr Udo Weimar  University of Tübingen, Germany
Prof Krishna Persaud  University of Manchester, UK
Prof Agustin Gutierrez  University of Barcelona, Spain
Dr Jan Mitrovics  JLM Innovation, Germany

Laboratory supervisors:
Dr Zoltan Racz  University of Manchester, UK
Dr Jan Mitrovics  JLM Innovation, Germany

School organizer:
Dr Monika Kwoka  University of Tübingen, Germany
Schedule

Sunday 1st February
15:00 – 18:00  Registration and laptop preparation
18:00 – 20:00  Welcome Reception
20:00       Dinner

Monday 2nd February
Introduction to sensors and data handling

Morning Session: Theory
08:30 – 09:30  Introduction to chemical sensors and analytical instrumentation
               Udo Weimar
09:30 – 09:50  Coffee break
09:50 – 10:50  Basic statistical analysis of data from a single sensor and initial pre
               processing
               Jan Mitrovics
10:50 – 11:50  Introduction to data handling and data formats suitable for data analysis
               Jan Mitrovics
12:00       Lunch

Evening Session: Computer Lab
17:30 – 19:30  Introduction to data formats and data analysis software using database I.
20:00       Dinner

Tuesday 3rd February
Exploratory data analysis and linear multivariate techniques

Morning Session: Theory
08:30 – 09:30  Signal and data pre-processing: experimental design
               Krishna Persaud
09:30 – 09:50  Coffee break
09:50 – 10:50  Linear regression methods: multi linear regression and partial least squares
               Udo Weimar
10:50 – 11:50  Linear methods for multivariate data analysis: principal component analysis
               and linear discriminant analysis
               Udo Weimar
12:00       Signal and data pre-processing: experimental design

Evening Session: Computer Lab
17:30 – 19:30  Exploration of database II using linear methods
20:00       Dinner
Wednesday 4th February

Non-linear multivariate analysis of chemical sensor data

Morning Session: Theory

08:30 – 09:30 Non-linear statistical methods
Krishna Persaud

09:30 – 09:50 Coffee break

09:50 – 11:50 Introduction to data handling and data formats suitable for data analysis
Krishna Persaud

12:00 Lunch

Evening Session: Computer Lab

17:30 – 19:30 Exploration of database III using non-linear methods

Thursday 5th February

Time series analysis of data from chemical sensors

Morning Session: Theory

08:30 – 09:30 Problems of sensor drift, time-varying parameters and aging
Julian Gardner

09:30 – 09:50 Coffee break

09:50 – 10:50 Analysis of time-dependent data-sets: linear models (filters)
Julian Gardner

10:50 – 11:50 Analysis of time-dependent data-sets: non-linear models (with thermally modulated sensors analysed by artificial neural networks and discrete wavelet transforms)
Agustin Gutierrez

12:00 Lunch

Evening Session: Computer Lab

17:30 – 19:30 Exploration of time-varying databases I and IV, sensor data and ion mobility spectra

Friday 6th February

Large data-sets and bio-inspired techniques

Morning Session: Theory

08:30 – 09:30 Feature selection techniques on large sensor arrays
Agustin Gutierrez

09:30 – 09:50 Coffee break

09:50 – 10:50 Biologically-inspired signal processing
Agustin Gutierrez

10:50 – 11:50 Challenges in the field of chemical and biological sensing: the odour segmentation problem, healthcare and security applications
Julian Gardner

11:50 – 12:00 Concluding remarks and farewell

12:00 Lunch