2019 Copenhagen School of Chemometrics

World-class teaching...
...in a World-class environment

COPENHAGEN
6th May – 7th June, 2019

# CSC – 2019

# COPENHAGEN SCHOOL OF CHEMOMETRICS

## May 6th – June 7th, 2019

## Main responsible: José Manuel Amigo

2019 Copenhagen School of Chemometrics

World-class teaching...
...in a World-class environment
COPENHAGEN
6th May – 7th June, 2019

# Description of the course:

**Chemometrics Analytical Technologies (CAT – www.models.life.ku.dk)** is more than happy to announce the new edition of the PhD course:

# Copenhagen School of Chemometrics - CSC

**CSC-2019** is a five-week school designed to be an introduction to different key aspects of advanced data analysis in different brands of Science (Chemistry, physics, environmental, political economics, etc). **CSC-2019** addresses to BSc, MSc, PhD students/post-docs, associate professors, etc. who want to acquire or refresh basic knowledge on multivariate data analysis from different disciplines. **CSC-2019** also addresses to researchers working in industry or research laboratories who want to implement multivariate data analysis in their daily research environment.

**CSC-2019** is thought to be an intensive school. Therefore, **CSC-2019** will be held in five weeks structured in different seminars/workshops where the students will be more than welcome to work with their own data together with high-qualified teachers. Moreover, anyone can choose the seminars to participate, not having a minimum of seminars request.

# Targets of CSC-2019:

**CSC-2019** aims at being a platform for:
**- Learning basic and advanced data analysis methods: CSC-2019** is specifically designed for researchers who want to start using data analysis in their routine work.

**- Sharing knowledge and interchange ideas between students covering different scientific backgrounds:** One of the key points of **CSC-2019** is the interaction between the students to discuss issues and troubleshooting always within the framework of scientific data analysis and performance. The reports days and the workshop will offer the opportunity to the students to discuss, share and improve their main issues and initiatives in a professional environment.

**- Meeting world-wide recognized experts of Multivariate Data Analysis in an open discussion forum environment**: **CSC-2019** will count on teachers that are well-recognized experts on chemometrics and multivariate data analysis in their respective fields. This, at the same time, will offer the possibility of opening new collaborative frameworks between students and teachers.

**- Flexibility in the seminars and ECTS credits:** The students can choose to attend the seminars which they consider more relevant for their research. There is no a minimum of seminars that the student must attend. Also, they will have the opportunity to deepen into any multivariate method. In total, attending the whole school, one student can obtain **12 ECTS credits**.

# Timetable, topics and lecturers:

**The timetable and topics for CSC-2019 are:**

| May | | | | | | |
|---|---|---|---|---|---|---|
| **Monday** | **Tuesday** | **Wednesday** | **Thursday** | **Friday** | **SAT** | **SUN** |
| 6<br>BASICS | 7<br>BASICS | 8<br>BASICS | 9<br>BASICS | 10<br>BASICS | 11 | 12 |
| 13<br>ALGEBRA | 14<br>ALGEBRA | 15<br>DoE | 16<br>DoE | **17**<br>DoE | 18 | 19 |
| 20<br>CLASS | 21<br>CLASS | 22<br>CLASS | 23<br>VARSEL | 24<br>VARSEL | 25 | 26 |
| 27<br>NONLIN | 28<br>NONLIN | 29<br>NONLIN | **30**<br>--- | 31<br>NONLIN | | |

| June | | | | | | |
|---|---|---|---|---|---|---|
| **Monday** | **Tuesday** | **Wednesday** | **Thursday** | **Friday** | **SAT** | **SUN** |
| | | | | | 1 | 2 |
| 3<br>CR | 4<br>CR | **5**<br>CR | 6<br>CR | 7<br>GLUE | 8 | **9** |

All seminars run from 9 a.m. to 5 p.m. with a lunch break. Organization of the breaks and the final timetable for each seminar will be under the responsibility of the corresponding teacher/lecturer.

**Seminars, lecturers and ECTS points:**

| | Acronym | Name | Teacher | ECTS |
|---|---|---|---|---|
| 1 | **BASICS** | Basic Chemometrics | José M. Amigo | 3 |
| 2 | **ALGEBRA** | Lineal Algebra | Morten A. Rasmussen | 1 |
| 3 | **DoE** | Design of Experiments | Riccardo Leardi | 1.5 |
| 4 | **CLASS** | Linear Classification | Davide Ballabio | 1.5 |
| 5 | **VARSEL** | Variable selection methods | Rasmus Bro | 1 |
| 6 | **NONLIN** | Nonlinear Modelling | Federico Marini | 2 |
| 7 | **CR** | Curve Resolution | Anna de Juan /José M. Amigo | 2 |
| BT | **GLUE** | How NOT to make Chemometrics | J.M. Amigo/R. Bro | 0 |

# For PhD students:

Each seminar accounts for **1, 1.5 or 2 ECTS** (see previous list). All PhD students who want to obtain the corresponding ECTS credits from the seminars will be obliged:

1) To attend the corresponding seminars.

2) To deliver the corresponding reports (see below). The length, content and the final delivering date of the reports will be **specified by the corresponding lecturer**.

# Optional "Introduction to Matlab for Multivariate Data Analysis" course offered:

Matlab is one of the main software packages that will be used in the **CSC-2019**. Therefore, attending some suggestions from the students of previous editions, we have decided to merge our PhD course "Introduction to Matlab for Multivariate Data Analysis" with **CSC-2019**.

**"Introduction to Matlab for Multivariate Data Analysis"** will be held from **25th, 26th, 29th, 30th of April and 2nd and 3rd of May, 2019,** and the students will be able to obtain 3 ECTS credits. This PhD course is totally independent from the **CSC–2019**.

If any student of the **CSC–2019** is interested in attending the "Introduction to Matlab for Multivariate Data Analysis" course, please, write to José Manuel Amigo (jmar@life.ku.dk – jmar@food.ku.dk).

# Detailed description for each seminar of CSC-2019:

## 1.- BASICS
## Basic Chemometrics

Principal Component Analysis has become the most powerful and versatile tool for exploring data tables in Analytical Sciences. Here we present a course to show the main benefits and drawbacks of PCA when it is used for different kind of analytical data: Spectroscopy, environmental assessment, sensory, experiments performance, chromatography, etc. Moreover, preprocessing of different type of data will be also addressed in the seminar as a prerequisite for having the optimal possibility for exploring the data.

If PCA is the keystone of pattern recognition methods, PLS is the keystone of multivariate calibration methods. This seminar will give a general overview of different multivariate calibration strategies and will focus in Partial Least Squares regression.

**Previous knowledge needed:** None
**Software:** Matlab and PLS-Toolbox. A full demo available at: http://www.eigenvector.com/
**Teacher:** Dr. José Manuel Amigo
**ECTS:** 3

## 2.- ALGEBRA
## Lineal algebra for multivariate data analysis

Linear Algebra is the basis of the Multivariate Data Analysis. Therefore, it is important to give it a central position in any school about multivariate data analysis. This seminar will be a clear introduction to the linear algebra behind the multivariate models that will be presented in the school.

**Previous knowledge needed:** Basic Chemometrics knowledge is advisable.
**Software:** Matlab.
**Teacher:** Dr. Morten A. Rasmussen
**ECTS:** 1

## 3.- DoE
## Introduction to Design of Experiments

DoE. Design of Experiments (Riccardo) The basic theory and practice of Design of Experiments is revisited. The aim is to give a practitioner idea or reminder about the main features and uses of DoE. Several real examples in very different fields will be shown. The seminar will be based on teaching hours and some guided exercises.

**Previous knowledge needed:** Very basic statistics.
**Software:** Free R-based software: http://gruppochemiometria.it/gruppo-lavoro-r-in-chemiometria.html
**Teacher:** Prof. Riccardo Leardi
**ECTS:** 1.5

## 4. CLASS
## Linear Classification

The seminar is focused on the theory and practice of linear classification tools (SIMCA and PLS-DA). The seminar will be based on teaching hours with guided exercises and practical sessions with real cases.

**Previous knowledge needed:** Basic knowledge of PCA.
**Software:** Matlab and free Classification Toolbox: http://michem.disat.unimib.it/chm/download/classificationinfo.htm
**Teacher:** Dr. Davide Ballabio
**ECTS:** 1.5

## 5.- VARSEL
## Variables selection methods

This seminar aims at revisiting the most important variable selection methods for regression and classification purposes with the aim at improving the performance of the models. The emphasis will be on practical applications, and what methods could be applied to which problem. There will also be hints as to what methods are good, and which ones to stay away from.

**Previous knowledge needed:** Basic knowledge of multivariate regression methods.
**Software:** Matlab and PLS-Toolbox. A full demo available at: http://www.eigenvector.com/
**Teacher:** Prof. Rasmus Bro
**ECTS:** 1

## 6.- NONLIN
## Non-linear modelling

This module aims at providing a basic introduction to the techniques which may be used in all those situations when a linear relation is not enough to provide accurate results (e.g. due to the presence of multiple sources of variability). In this respect, the most important aspects of data modeling will be considered (exploratory analysis, classification and calibration). Topics such as kernel and dissimilarity-based approached (including support vector machines), local modeling (kNN and locally weighted regression/classification) and artificial neural networks will be covered.

**Previous knowledge needed:** Basic knowledge of PCA and PLS
**Software:** Matlab (plus in-house routines, will be provided).
**Teacher:** Dr. Federico Marini
**ECTS:** 2

## 7. CR
## Curve Resolution. Multivariate Curve Resolution and Multiway analysis.

Curve resolution techniques are gaining importance in modeling of different analytical data types. Among them, we can distinguish between the ones working in matrices (Multivariate Curve Resolution) and the ones working in data cubes (Parallel Factor Analysis). They have widely demonstrated their usefulness in kinetic modeling, solving problems in chromatographic data (peak resolution/deconvolution) and hyperspectral images. This seminar will offer a general overview of curve resolution methodologies and will focus in two of them: 1) Multivariate curve resolution: The two first days, given by Prof. Anna de Juan and 2) PARAFAC: The two second days, given by Dr. José Manuel Amigo.

**Previous knowledge needed:** Basic knowledge of PCA.

**Software:** Matlab and free MCR-ALS Toolbox: http://www.mcrals.info/ and PLS-Toolbox. A full demo available at: http://www.eigenvector.com/
**Teacher:** Prof. Anna de Juan and Dr. José Manuel Amigo
**ECTS:** 1

**BT.- GLUE**
**How NOT to make chemometrics**

In this seminar we will take a very close look at all the most common mistakes that even experienced people will do when doing multivariate analysis. We will cover exploration, calibration, interpretation, visualization and many other subjects. And always with a focus on what is the most common problem as well as a sounder alternative.

**Previous knowledge needed:** Basic knowledge of Chemometrics
**Software:** Matlab will be used for demonstrations. Software and data will be provided.
**Teacher:** Dr. José Manuel Amigo and Prof. Rasmus Bro
**ECTS:** 0

# Location and subscription:

**Location: CSC-2019** will be held by the department of food sciences of the University of Copenhagen, Denmark. The classrooms will be announced soon.

**Subscription:**
The persons interested in joining the CSC-2019 will have to subscribe at:
https://phdcourses.ku.dk

After accessing to the course catalogue, the student will have to sign up for the seminars. There are two ways:
- If the student wants to make the whole school, please, follow the link to **Copenhagen School of Chemometrics - the complete course (12 ECTS)**
- If, on the contrary, the student wants to join specific seminars, follow the links to each individual seminar.

For further information, just send an e-mail to Dr. José Manuel Amigo (jmar@food.ku.dk / jmar@life.ku.dk).

**Note:** Please, send us an e-mail after registration (jmar@life.ku.dk / jmar@food.ku.dk)

# Fees:

**Academia:**
- 800 DKK per credit
- 7000 DKK the whole school (5 weeks)

**Industry/Companies:**
- 1600 DKK per credit
- 15000 DKK the whole school (5 weeks)

**There will be a fee of 3000 Danish Krone for not coming when enrolled. All payments with no exception will be charged in Danish Krone (DKK).**

# Best poster award:

In order to encourage communication between students and with the teachers, the students will be able to bring a poster and present his/her work in a flash presentation in the corresponding seminar. Students are invited to bring the poster no matter which seminars they are attending.

The poster can also be a poster presented in another conference.

The poster size should be A1 (594 x 841 mm).

Once the posters are evaluated, the last day of CSC-2019 the winner of the "best poster award" will be announced and the price given. **The price will consist of a diploma and the return of the CSC-2019 fees.**

# Important information:

- Read careful the description for each seminar and the needs regarding software. Unless strictly needed, we will **NOT** provide with laptops or software packages. Therefore, bring your own laptops with the required software already installed.

- We will provide with office material (notebooks, folders, USB, pens, etc.)

- There will be **free WI-FI internet** connection.

- We can issue parking permission. But communicate in advance.

- We can issue letters of participation. But communicate in advance.

- **Lunch is not included.** Nevertheless, there will be free coffee and refreshment for the coffee breaks.