

# Introduction to Process Analytics using Multivariate Methods

DAY 1 – DAY 2 (10-11 June 2019)

The participants will be introduced to modern day multivariate data analytics methods through lectures and hands-on workshops. The syllabus is geared towards general concepts on latent variable modeling (LVM) theory and advanced topics on the analysis of specific data scenarios (e.g. batch data, image analysis and chemometrics). LVM is a data-driven modeling technique particularly useful to understand processes where acquired data is: abundant, complex, correlated and noisy. Basic knowledge of statistics, linear algebra and geometry are helpful to fully understand the concepts of this course.

Course Instructor: Dr Salvador Garcia-Munoz.

*The course will take place in Lecture Theatre 2, room 203, ACEX Building, Imperial College London, South Kensington Campus, London SW7 2AZ.*



*Dr Salvador Garcia-Munoz is a Visiting Professor at Imperial College London, with +20 years of experience in the implementation of systems engineering tools to industrial problems. He works for the pharmaceutical R&D sector leading the application of digital design tools for the development of new products and accelerated process design. He is an active member of AIChE, a founder of the Systems Based Pharmaceuticals Alliance and associate editor for Chemical Engineering Research and Design. His research in multivariate modeling spans from industrial applications to the development of new methods and algorithms to analyze complex datasets common in contemporary industrial scenarios*

These training courses have been organised by CPSE, a multi-institutional research centre of world-class departments at Imperial College London and University College London. CPSE was inaugurated in 1989 by Professor Roger W.H. Sargent, the founding Director of the Centre from 1989 to his retirement in 1992. Since then, CPSE has continued the legacy of Professor Sargent and remains a Centre of Excellence. CPSE academics come from multi-disciplinary backgrounds, including: chemical engineering, mathematics, physics and chemistry and are international leaders in their fields.

# Advanced Applications of Process Analytics using Multivariate Methods

DAY 3 – DAY 4 (12-13 June 2019)

This course will explore advanced applications of the Latent Variable Modeling (LVM). The syllabus is geared towards more advanced topics such as the analysis of batch data, process and product design, multivariate image and texture analysis and chemometrics. Knowledge of multivariate methods is required to fully understand the concepts covered in this course.

Course Instructor: Dr Salvador Garcia-Munoz.

*The course will take place in Lecture Theatre 2, room 203, ACEX Building, Imperial College London, South Kensington Campus, London SW7 2AZ.*

## Registration Fee (for each course):

- £480 for industry participants  
(20% discount for CPSE Consortium companies)
- £240 for non-CPSE researchers
- £80 for non-CPSE students
- Free for CPSE students and researchers

*If you register for both courses a 30% discount will be applied on the Advanced course.*

**To register please email Dr Cristina Bertulli at [c.bertulli@imperial.ac.uk](mailto:c.bertulli@imperial.ac.uk)**

**Cancellations:** cancellations must be received in writing before or on **3<sup>rd</sup> June 2019**. After this date cancellations will be subject to an administration charge of £50.00. Substitutions may be made at any time, whilst a valid place is held. The organizer cannot accept liability for costs incurred in the event of a course having to be cancelled as a result of circumstances beyond its reasonable control. If you would like to discuss the course, please. contact Dr Cristina Bertulli [c.bertulli@imperial.ac.uk](mailto:c.bertulli@imperial.ac.uk)

# Introduction to Process Analytics using Multivariate Methods

## Detailed Course Summary

### DAY 1—10 June 2019, 9am - 5pm

#### Principal Components Analysis (PCA)

- 1.1. Projection Mechanisms
- 1.2. The dot Product and coordinate transformation of data
- 1.3. The objective function behind PCA
- 1.4. Solving the parameter estimation problem
- 1.5. The PCA Model and its basic parameters
  - 1.5.1. Scores, Loadings, R2
  - 1.5.2. Contribution plots
- WORKSHOP 1 and 2
- 1.6. The PCA Model and its diagnostics
  - 1.6.1. SPE, Hotelling's T2
  - 1.6.2. Cross-validation
- 1.7. Outlier detection
- 1.8. Handling of missing data
  - 1.8.1. The NIPALS algorithm
  - 1.8.2. Missing data in model building
  - 1.8.3. Missing data in model evaluation
- WORKSHOP 3
- 1.9 Multivariate Statistical Process Control
  - WORKSHOP 4
- 1.10 Establishment of multivariate specifications for raw materials

### DAY 2—11 June 2019, 9am - 5pm

#### 2. Projection to Latent Structures (PLS)

- 2.1. Fundamentals
  - 2.1.1. From PCA to PCR to PLS
  - 2.1.2. The PLS model, objective function and parameter estimation solutions
  - 2.1.3. PLS Parameters and Diagnostics
- WORKSHOP 5
- 2.1.4. The unique features of the PLS model
  - WORKSHOP 6
- 2.2 Multi-Block PLS approaches
  - 2.2.1 MBPLS
    - 2.2.1.1 Conventional algorithm and equivalences to block scaled PLS
- WORKSHOP 7
- 2.2.2 LPLS – Algorithm, application, objective function.
- 2.2.3 JYPLS – Algorithm, application, objective function
- 2.2.3 JRPLS – Algorithm, application, objective function.
- WORKSHOP 8
- 2.2.4 TPLS – Algorithm, applications.
  - WORKSHOP 9

#### 3. Practical Issues in Data Based Modelling

- 3.1. Handling of Missing Samples
- 3.2. Modelling Systems under Feedback
- 3.3. Non-linear PLS
  - WORKSHOP 10

# Advanced Applications of Process Analytics using Multivariate

## Detailed Course Summary

### DAY 3—12 June 2019, 9am - 5pm

#### WORKSHOP

#### **4. Analysis of Batch Data**

- 4.1 Unfolding options for 3D data structures and consequences
- 4.2 Equivalences between multiway PCA and Adaptive Moving Horizon Estimators
- 4.3 Alignment of batch trajectories

#### WORKSHOP 9

- 4.4 Time varying diagnostics
- 4.5 MSPC for Batch Processes

#### WORKSHOP 10

#### **5. Process and Product Design : PLS Model Inversion**

- 5.1 The inverse of a PLS model
- 5.2 Rank considerations and null-spaces, hard constraints vs soft-constraints.

#### WORKSHOP 11

- 5.3 Reformulation of the inverse problem using Non-linear Optimization
- 5.4 Constraints and formulations of the inverse problem.

#### WORKSHOP 12

### DAY 4—13 June 2019, 9am - 5pm

#### **6. Multivariate Image Analysis**

- 6.1 PCA kernel solutions to specially shaped data
- 6.2 Scores and masks
- 6.3 Image classification techniques
- 6.4 Image regression – The covariance method

#### WORKSHOP 13

#### **7. Multivariate Texture Analysis**

- 7.1 Wavelet Decomposition
- 7.2 Multivariate Decomposition of Textural Features

#### **8. Considerations in Chemometrics**

*The rest of the day (if any time left) will be spent working with the students on their projects*