

SCIENTIFIC AND METHOD MODULES

Module name	Chemometrics applied to spectroscopic data
Number	2019-B6
Aims	<ul style="list-style-type: none"> • To develop an understanding of the application of chemometrics to spectroscopic data analysis • To gain a basic overview of how to process spectral data for multivariate analysis • Learn how to apply Principal Component Analysis, Partial Least Squares Discriminant Analysis and Partial Least Squares Regression to spectroscopic data using <i>The Unscrambler</i>
Contents	<ul style="list-style-type: none"> • Understand that the Universe is multivariate and applications of multivariate statistics • Explore types of data used in chemometrics Near-IR, Raman and IR, qualitative • Distinguish between classification (PCA) and regression modelling • Explain the difference between objects, variables including continuous and discrete variables • Describe ordinal, binary and category, dummy variables • Input data into Unscrambler • Plot line plots, bar plots, matrix plots • Identify noise and variance in spectra • Explain PCA and how it is used for data analysis • Understand and interpret scores and loadings plots • Recognise the difference between calibration, validation and test set • Explain “Explained Variance” • Define PCs, scores, loadings, residuals, variance and the PCA model equation • Know the pre-processing approaches for spectral data (baseline correction, derivatives, reduce average, mean centring, smoothing, normalisation, standard normal variate, Multiplicative scatter correction) • Be able to detect outliers • Partial Least Squares Discriminant Analysis • Partial Least Squares Regression Analysis
Type	Two-day block course

Date (month/year)	October 15 (12 – 8 pm), October 16 (8 – 10 am & 12:30 – 5 pm) 2019
Time	see schedule next page
Work load	15 hours presence/ 45 hours self-study
Examination	Written exam with multiple choice questions
Credit points	2
Responsible scientists	Bayden Wood, Philip Heraud, Monash, Australia
Guest lecturers	Bayden Wood, Philip Heraud, Monash, Australia
Industrial partners	None
Recommendations for literature	Multivariate Data Analysis 6 th Edition, K. H. Esbensen & B. Swarbrick, 2018, CAMO Analytics

PRELIMINARY TIME SCHEDULE

Time	Lecturer	Title
Tuesday, October 15th 2019, Location: PC-Pool TA		
12:00 Welcome		
12:15	Wood	The World is multivariate
	Heraud	Introduction to multivariate data analysis (Unscrambler Exercises 1 and 2 data input and plotting spectra)
15:00 <i>Coffee break</i>		
15:30-18:00	Wood / Heraud	Spectral pre-processing
	Wood / Heraud	Principal Component Analysis
	Wood / Heraud	Exercises 3 and 4 (Preprocessing data and PCA)
Wednesday, October 16th 2019, PLS-DA and PLS-R, Location: PC-Pool TA		
8:30	Wood / Heraud	Tute on Infrared and Raman spectroscopy
10:00 <i>Break / lunch</i>		
12:00	Wood / Heraud	PLS-DA (Exercise 5-6)
14:00 <i>Coffee break</i>		
14:30	Wood / Heraud	PLS-R (Exercises 7-8)
16:30 <i>Coffee break</i>		
16:30-18:30 Scientific discussions on students data		

Status 28.08.2019