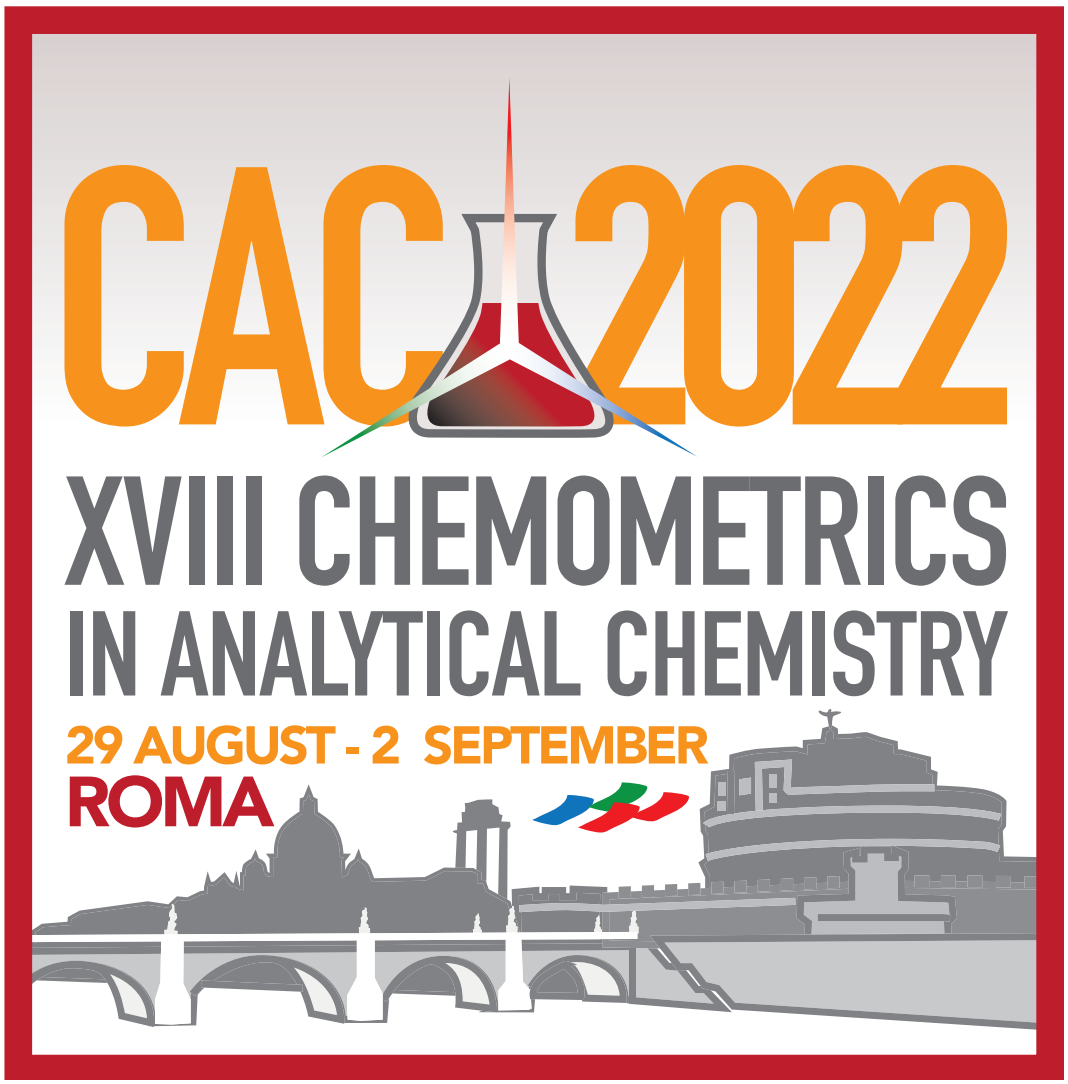


BOOK OF ABSTRACTS

CAC Permanent Committee

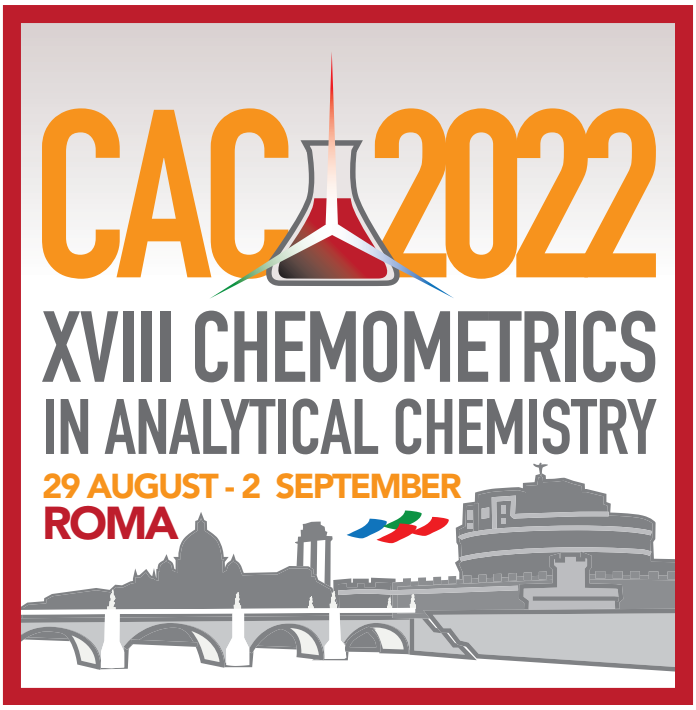
- **Ricard Boqué** *Universitat Rovira i Virgili (Spain)*
- **Lutgarde Buydens** *Radboud University Nijmegen (The Netherlands)*
- **Márcia M.C. Ferreira** *Universidade Estadual de Campinas (Brazil)*
- **Michele Forina** *Retiree University of Genova (Italy)*
- **Neal B. Gallagher** *Eigenvector Research (Inc. (United States)*
- **Philip Hopke** *Clarkson University (United States)*
- **Jose Cardoso Menezes** *Instituto Superior Tecnico (Portugal)*
- **Jean-Michel Roger** *INRAE (France)*
- **Sarah C. Rutan** *Virginia Commonwealth University (United States)*
- **Romà Tauler** *IDAEA – CSIC (Spain)*
- **Bernard Vandeginste** *Retiree Unilever Research and Development (The Netherlands)*
- **Yvan Vander Heyden** *Vrije Universiteit Brussel (Belgium)*
- **Pierre Van Espen** *Universiteit Antwerpen (Belgium)*
- **Peter Wentzell** *Dalhousie University (Canada)*
- **Barry M. Wise** *Eigenvector Research Inc. (United States)*

Scientific Committee

- **José Manuel Amigo** *University of Basque Country (Spain)*
- **Claudia Beleites** *Chemometrix GmbH (Germany)*
- **Anna De Juan** *University of Barcelona (Spain)*
- **Hector Goicoechea** *Universidad Nacional del Litoral (Argentina)*
- **Aoife Gowen** *University College Dublin (Ireland)*
- **Peter Harrington** *Ohio University (USA)*
- **John Kalivas** Idaho State University (USA)
- **Mohsen Kompany-Zareh** *IASBS (Iran)*
- **Hongmei Lu** *Central South University (China)*
- **Marena Manley** *Stellenbosch University (South Africa)*
- **Fernanda Pimentel** *Federal University of Pernambuco (Brazil)*
- **Åsmund Rinnan** *University of Copenhagen (Denmark)*
- **Oxana Rodionova** *Semenov Institute of Chemical Physics (Russia)*
- **Jean-Michel Roger** *INRAE (France)*
- **Serge Rudaz** *University of Geneva (Switzerland)*
- **Francesco Savorani** *Polytechnic University of Turin (Italy)*
- **Age Smilde** *University of Amsterdam (The Netherlands)*
- **Agnieszka Smolinska** *Maastricht University (The Netherlands)*
- **Beata Walczak** *University of Silesia (Poland)*
- **Peter Wentzell** *Dalhousie University (Canada)*

Organizing Committee

- **Federico Marini** *University of Rome La Sapienza (Italy)* Conference Co-Chair
- **Cyril Ruckebusch** *University of Lille (France)* Conference Co-Chair
- **Marina Cocchi** *University of Modena and Reggio Emilia (Italy)*
- **Ludovic Duponchel** *University of Lille (France)*
- **Raffaele Vitale** *University of Lille (France)*
- **Alessandra Biancolillo** *University of L'Aquila (Italy)*
- **Olivier Devos** *University of Lille (France)*
- **Paolo Oliveri** *University of Genova (Italy)*



CONFERENCE PROGRAM

PRELIMINARY PROGRAM

Monday 29 August 2022

10:00-19:00 **Registration of the participants**

14:30-15:00 **Conference opening**
15:00-16:00 **PL1: Harald Martens - Human-interpretable Machine Learning with an Eye for Causalities: *Making sense of modern measurement streams inside and outside chemistry***
16:00-16:30 **Coffee break & Poster session**
16:30-18:10 **Contributed Session I: Big Data and Machine Learning**
16:30-16:50 OL1: Davide Ballabio - Enhancing LC-MS/MS Spectral Searching With Multi-Task Neural Networks And Molecular Fingerprints
16:50-17:10 OL2: Gabriel Vivo-Truyols - On the use of Bayesian statistics for (big) data analysis: automation for both qualitative and quantitative chemometrics
17:10-17:30 OL3: Michael Sorochan Armstrong - Chemometrics with Amazon Web Services (AWS)
17:30-17:50 OL4: Jeroen Jansen - Process economy, efficiency and sustainability go hand in hand, how chemometrics can build a greener industry
17:50-18:10 OL5: Priyanka Kumari - QSRR for small pharmaceutical compounds in RPLC: A Machine learning approach
18:10-20:10 **Welcoming cocktail**

Tuesday 30 August 2022

09:00-10:00 **PL2: Roy Goodacre - Lessons from large-scale metabolic phenotyping**
10:00-11:00 **Contributed Session II: Omics/ASCA and related methods I**
10:00-10:20 OL6: Albert Menéndez Pedriza - Comparison of mid-level fusion strategies for the multi-omic analysis of toxicological data
10:20-10:40 OL7: Andrés Martínez Bilesio - Metabolomics-guided insights on Bariatric Surgery: a longitudinal chemometrics approach over 1H NMR spectra from serum samples
10:40-11:00 OL8: Sarah Malek - Evaluation of mid-infrared spectra of serum and synovial fluid in predicting early post-traumatic osteoarthritis in an equine model
11:00-11:30 **Coffee break & Poster session**
11:30-13:10 **Contributed Session III: Applications I**
11:30-11:50 OL9: Zuzana Małyjurek - Class-Modelling - Optimization, Validation and Application
11:50-12:10 OL10: Agnieszka Martyna - Likelihood ratio in forensic discrimination/classification tasks
12:10-12:30 OL11: Martina Foschi - Supervised and Unsupervised Chemometric Methods to deal with saffron aging and its Quality Control
12:30-12:50 OL12: Lorenzo Strani - Real time prediction of ABS properties through multiblock and local regression methods
12:50-13:10 OL13: Sebastian Orth - Spectral imaging - pre-harvest malting barley germination classification with sequential orthogonalised multiblock data fusion methodologies
13:10-14:30 **Lunch & Poster session**

14:30-15:00 **KN1: Mathias Sawall - On the ambiguity underlying the spectral recovery problem and its analysis by the area of feasible solutions**
15:00-16:00 **Contributed Session IV: Theory & algorithms I**
15:00-15:20 OL14: Sergey Kucheryavskiy - Procrustes cross-validation of multivariate regression models
15:20-15:40 OL15: Stephan Seifert - Opening the random forest black box with Surrogate Minimal Depth
15:40-16:00 OL16: Oxana Rodionova - Expansion of the DD-SIMCA concept
16:00-16:30 **Coffee break & Poster session**
16:30-18:10 **Contributed Session V: Multi-block/Multi-way/Multi-set I**
16:30-16:50 OL17: Paul-Albert Schneide - Speeding up PARAFAC2 dramatically
16:50-17:10 OL18: Isabelle Viegas - Coupled factorization of fluorescence data of proteins and quantum dots to assess their conjugation process
17:10-17:30 OL19: Oksana Mykhalevych - New tools for designing food ingredients structures
17:30-17:50 OL20: Maria Cairoli - Monitoring pollution pathways in river water by predictive path modelling using untargeted GC-MS measurements
17:50-18:10 OL21: Ivan Krylov - Fluorescence and scattering model estimation

Wednesday 31 August 2022

09:00-10:00 **PL3: Ingrid Måge - Industrial bioprocessing – an amusement park for chemometricians and analytical chemists**
10:00-11:00 **Contributed Session VI: Multi-block/Multi-way/Multi-set II**
10:00-10:20 OL22: Jean-Michel Roger - N-CovSel, a new strategy for feature selection in N-way data
10:20-10:40 OL23: Mahdiyeh Ghaffari - Using Multi-Block Non-Negative Matrix Factorization for Multi-layer Plastic Sorting
10:40-11:00 OL24: Paul Gemperline - Combining ASCA and Tucker3 models to explain high-dimensional data
11:00-11:30 **Coffee break & Poster session**
11:30-13:10 **Contributed Session VII: Applications II**
11:30-11:50 OL25: Sabina Licen - Data fusion based on self-organizing map algorithm for the integration of different source/frequency instrumental data and spot sampling contextualization for environmental monitoring
11:50-12:10 OL26: Vicky Caponigro - Application of different chemometric approaches for MALDI-MSI data set of heterogeneous tissues. Case study: parotid tumour
12:10-12:30 OL27: Ewa Szymanska - Comprehensive chemometric strategy for the high-throughput screening of in-line spectroscopic sensors for milk composition traits
12:30-12:50 OL28: Maxime Ryckewaert - Combining hyperspectral imaging data with climate data to predict physiological variables of grapevine plants
12:50-13:10 OL29: Eleni Ioannidi - Using ATR FT-IR and MCR as a method to understand the crystal state of chocolates tempered under different conditions
13:10-14:30 **Lunch & Poster session**

14:30-15:00 **KN2: Hadi Parastar - Integration of handheld spectrometers and chemometrics for food authentication**
15:00-16:00 **Contributed Session VIII: Spectroscopy & Imaging I**
15:00-15:20 OL30: Cristina Malegori - HSI-NIR and chemometrics for the quantification of collagen in bones: how chemical mapping can help in preserving archeological finds
15:20-15:40 OL31: Manuela Mancini - Spectroscopy and chemometrics for sorting waste wood material according to the best-suited application
15:40-20:30 **Social activity (tours will start from the agreed meeting points at 17:00)**

Thursday 1 September 2022

09:00-10:00 **PL4: Romà Tauler - Bilinear model factor decomposition: a general mixture analysis tool**
10:00-11:00 **Contributed Session IX: Curve Resolution**
10:00-10:20 OL32: Martina Beese - An active constraint approach to identify essential spectral information in noisy data
10:20-10:40 OL33: Laureen Coic - A phasor view of Multivariate Curve Resolution
10:40-11:00 OL34: Anna De Juan - Trilinearity in Multivariate Curve Resolution: hybrid modeling and missing data
11:00-11:30 **Coffee break & Poster session**
11:30-13:10 **Contributed Session X: Spectroscopy & Imaging II**
11:30-11:50 OL35: Florent Abdelghafour - Combining spectral and spatial features extracted from hyperspectral images: Application on the detection of scab disease
11:50-12:10 OL36: Rodrigo Rocha de Oliveira - 2-D wavelet image decomposition and Multivariate Statistical Process Control for blending end-point detection
12:10-12:30 OL37: Valeria Tafintseva - Modelling and preprocessing of sparse infrared spectra
12:30-12:50 OL38: Nicola Cavallini - Tracing the identity of mountain product Parmigiano Reggiano PDO cheese using 1H-NMR spectroscopy and multivariate data analysis
12:50-13:10 OL39: Paolo Oliveri - A combined chemometric strategy for a non-destructive age estimation of biological fluid stains
13:10-14:30 **Lunch & Poster session**
14:30-15:00 **KN3: Maria Cruz Ortiz - Analytical Quality by Design using a computational approach for the inversion of a PLS model**
15:00-16:00 **Contributed Session XI: Omics/ASCA and related methods II**
15:00-15:20 OL40: Miguel De Figuereido - Rebalanced ASCA (RASCA) to handle unbalanced multifactorial designs
15:20-15:40 OL41: Michel Thiel - LMWiRe: an R package for Linear Modeling of Wide Responses based on ASCA family of methods
15:40-16:00 OL42: Claudia Beleites - An Experimental Design Perspective on Cross-Validation
16:00-16:30 **Coffee break & Poster session**
16:30-17:50 **Contributed Session XII: Spectroscopy & Imaging III**
16:30-16:50 OL43: Siewert Hugelier - Quantifying the Tau protein aggregation degradation process by classification of super-resolution fluorescence microscopy localizations
16:50-17:10 OL44: Erik Tengstrand - Calibration transfer of Near-Infrared and Raman models without using transfer samples

17:10-17:30 OL45: Alisa Rudnitskaya - Characterization of microplastics from marine organisms using near infrared hyperspectral imaging
17:30-17:50 OL46: Jose Luis Aleixandre-Tudo - Spectral evaluation of fresh grapevine organs using self-organizing maps (SOM)
17:50-18:30 Awards Ceremony (Elsevier Chemometrics and Intelligent Laboratory Systems Award & Lifetime Achiement Award)
20:15-01:00 Social dinner

Friday 2 September 2022

09:00-11:00 Contributed Session XIII: Theory & algorithms II
09:00-09:20 OL47: Nematollah Omidikia - Infrared Ion Spectroscopy Peak Matching using Peak Annotation Technique
09:20-09:40 OL48: Sergio Oller Moreno - Peak matching across Gas Chromatography-Ion Mobility Spectrometry samples
09:40-10:00 OL49: Wouter Saeys - Multivariate monitoring and update strategies for calibration models
10:00-10:20 OL50: Sean Rozinski - What's UMAP Doing Anyway?
10:20-10:40 OL51: Ramin Nikzad-Langerodi - Does it Transfer? Assessing model generalization in domain adaptation with data fusion
10:40-11:00 OL52: Benjamin Mahieu - New developments around the VIP index
11:00-11:30 Coffee break & Poster session
11:30-12:50 Contributed Session XIV: Applications III
11:30-11:50 OL53: Dmitry Kirsanov - Chemometrics in spent nuclear fuel reprocessing
11:50-12:10 OL54: Joscha Christmann - Monitoring of fermentation processes by gas chromatography-ion mobility spectrometry (GC-IMS)
12:10-12:30 OL55: Martín Bravo - Development of an analytical platform for the identification of Fusarium circinatum in culture media, using VIS-NIR spectroscopy and chemometric methods
12:30-12:50 OL56: Tim Offermans - Retrospective Quality by Design r(QbD) using Historical Process Data and Design of Experiments
12:50-13:30 Conference Closing

TRACING THE IDENTITY OF MOUNTAIN PRODUCT PARMIGIANO REGGIANO PDO CHEESE USING ^{1}H-NMR SPECTROSCOPY AND MULTIVARIATE DATA ANALYSIS

N. Cavallini[1], P.P. Becchi[2], C. Durante[2], L. Strani, V. Pizzamiglio[3], S. Michelini[3], F. Savorani[1], M. Cocchi[2]

[1]*Department of Applied Science and Technology, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italia*

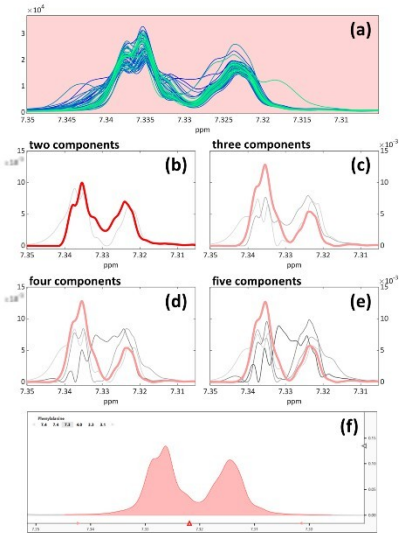
[2]*Dipartimento di Scienze Chimiche e Geologiche, Università di Modena e Reggio Emilia, Via Campi 103, 41125 Modena, Italia*

[3]*Consorzio Formaggio Parmigiano Reggiano, via Kennedy 50, 42124, Reggio Emilia, Italia*

nicola.cavallini@polito.it

Parmigiano Reggiano cheese is one of the most appreciated and famous Italian cheeses and even if its name is usually associated with a single product idea, different varieties can be found. In particular, the product "Prodotto di Montagna - Progetto Qualità Consorzio ("Mountain Product - Consortium Quality Project") represents a quality denomination for Parmigiano Reggiano PDO (Protected Designation of Origin) cheese that must comply with rather strict rules about its aging, geographical origin and the cow feed and breeding.

In this scenario, there has been an increasing request from both dairy farmers and consortia to protect the authenticity of Mountain Product Parmigiano Reggiano PDO from analogues, and to promote it as a higher quality product. To this aim, comprehensive analytical techniques can provide objective quality and identity assessments and proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy can be used as a tool for metabolic fingerprinting of milk and its derivatives [1-2]. NMR spectroscopy can provide huge amounts of information directly related to many metabolites with a single analytical run, and it can be therefore used for the identification of sugars, small organic acids, vitamins, nucleotides, and aromatic compounds. Due to the NMR signals complexity, multivariate data analysis is needed to interpret and extract information from this type of data, generally leading to optimal results in food characterization and authenticity assessments [3–4].



In this study the metabolic profile of "Mountain Product - Consortium Quality Project" and conventional Parmigiano Reggiano PDO samples were analyzed by means of ^{1}H-NMR spectroscopy, with the aim of finding information useful to distinguish the two denominations. To this aim, two different data analysis approaches were employed: the full spectra dataset (i.e., without any compression) was compared with a "features dataset" obtained by applying Multivariate Curve Resolution (MCR, [5]) to carefully defined small intervals. The extracted features were compared and matched with literature and reference libraries to obtain a putative identification of the resolved compounds/metabolites.

Figure 1 – Chemical identification of MCR-resolved components (b–e) by comparison with a reference library (f).

References

[1] P. Scano, E. Cusano, P. Caboni, R. Consonni, *Int. Dairy J.*, **90** (2019) 56–67

[2] Celso F. Balthazar, Jonas T. Guimarães, Ramon S. Rocha, Tatiana C. Pimentel, Roberto P.C. Neto, Maria Inês B. Tavares, Juliana S. Graça, Elenilson G. Alves Filho, Mônica Q. Freitas, Erick A. Esmerino, Daniel Granato, Sueli Rodrigues, Renata S.L. Raices, Marcia C. Silva, Anderson S. Sant'Ana, Adriano G. Cruz, *Trends Food Sci. Technol.*, **108** (2021) 84-91

[3] N. Cavallini, F. Savorani, R. Bro, M. Cocchi, *Molecules*, **26**(5) (2021) 1472

[4] L. Laghi, G. Picone, F. Capozzi, *TrAC*, **59** (2014) 93–102

[5] A. De Juan, J. Jaumot, R. Tauler, *Anal. Methods*, **6** (2014) 4964–4976

The NIR side of lentil

N. Cavallini[1], A. Giraudo[1], M. Sozzi[1], E. Cazzaniga[1], F. Geobaldo[1], F. Savorani[1]
[1]*Department of Applied Science and Technology (DISAT), Polytechnic of Turin, Turin, Italy*
nicola.cavallini@polito.it

The interest towards bioeconomy concepts has been considerably growing during the last years and, in particular, the development of sustainable and renewable bio-based technologies for food production is becoming increasingly important and studied. One of the most interesting applications of bioeconomy in the "food" area is the use of enzymes for the modification of food materials [1], to improve safety and to optimize the overall treatment processes.

In this perspective, the present study was focused on two processes for treating lentil flour: extraction and hydrolyzation, aimed at making protein available in solution. For both processes, an initial extraction phase with $Ca(OH_2)$ at controlled pH = 8 and fixed temperature of 60 °C was performed. Regarding only the hydrolyzation process, 0.2 % of protease enzyme was also added. The two processes were then carried out with two experimental runs each, differing by stirring rate (60 rpm and 120 rpm). A total of 32 samples of the processed solutions were collected at fixed time points in the range 0–300 minutes and stored frozen upon spectroscopic analysis. All samples were then analysed with a FT-NIR spectrometer (MPA by Bruker) and a Visible spectrometer (Carey by Agilent).

The acquired data were imported under MATLAB environment to undergo data quality assessment aimed at removing clear outliers and at choosing the proper preprocessing, specific for each dataset. The NIR data were preprocessed using standard normal variate (SNV) and mean centering, while the Visible data were only mean centered. The noisy regions with also low variance were removed from both datasets.

The datasets were explored by means of PCA, with the aim of obtaining information related to the type of process (extraction only *vs* extraction + hydrolyzation) and its evolution in time.

Curiously, the NIR data provided less clear information compared to the Visible data, with which interesting trends could be more easily identified. For this reason, a low-level data fusion approach was also put in place, by directly joining the two spectral datasets, after proper preprocessing and after the application of block scaling to give both blocks the same variance. The detected trends were confirmed, suggesting that both datasets provided useful information that could be efficiently combined and extracted.

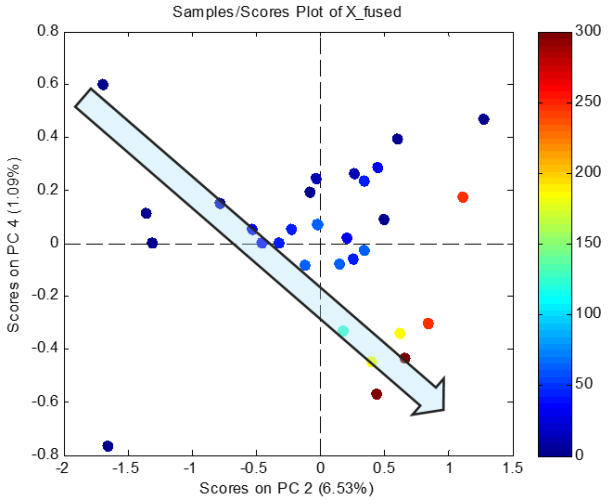


Figure 1 – Interesting time trend detected in the Visible dataset PCA.

These results suggest that Visible spectroscopy could be very useful for monitoring the extraction and hydrolyzation processes, while the application of NIR spectroscopy to this specific process and experimental setup would probably need further investigations.

References

[1] O. L. Tavano, Journal of Molecular Catalysis B: Enzymatic, 2013, 90, 1-11.