



HAL
open science

Editorial: Variable selection in chemometrics

Raffaele Vitale, Jean-Michel J. -M. Roger

► **To cite this version:**

Raffaele Vitale, Jean-Michel J. -M. Roger. Editorial: Variable selection in chemometrics. *Frontiers in Analytical Science*, 2023, 3, pp.1186952. 10.3389/frans.2023.1186952. hal-04541804

HAL Id: hal-04541804

<https://hal.inrae.fr/hal-04541804>

Submitted on 11 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



OPEN ACCESS

EDITED AND REVIEWED BY

Frank Westad,
Norwegian University of Science and
Technology, Norway

*CORRESPONDENCE

Raffaele Vitale,
✉ raffaele.vitale@univ-lille.fr

SPECIALTY SECTION

This article was submitted to
Chemometrics, a section of the journal
Frontiers in Analytical Science

RECEIVED 15 March 2023

ACCEPTED 31 March 2023

PUBLISHED 11 April 2023

CITATION

Vitale R and Roger J-M (2023), Editorial:
Variable selection in chemometrics.
Front. Anal. Sci. 3:1186952.
doi: 10.3389/frans.2023.1186952

COPYRIGHT

© 2023 Vitale and Roger. This is an open-
access article distributed under the terms
of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Editorial: Variable selection in chemometrics

Raffaele Vitale^{1*} and Jean-Michel Roger^{2,3}

¹Centre National de la Recherche Scientifique (CNRS), LASIRE (UMR 8516), Laboratoire Avancé de Spectroscopie pour les Interactions, la Réactivité et l'Environnement, Université de Lille, Lille, France, ²Technologies et Méthodes pour les Agricultures de Demain - Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (ITAP-INRAE), Institut Agro, Université de Montpellier, Montpellier, France, ³ChemHouse Research Group, Université de Montpellier, Montpellier, France

KEYWORDS

chemometrics, variable selection, spectrometry, sensors, analytical chemistry

Editorial on the Research Topic Variable selection in chemometrics

In statistics, the term *variable selection* relates to a wide variety of computational approaches that have been commonly exploited all over the years in an attempt to solve the so-called *curse of dimensionality*, which usually hampers the use of most classical data analysis methodologies in many diverse practical situations typical of the modern Big Data era. Conversely, in chemometrics, variable selection has not played the same historical role owing to the fact that basically all the techniques conceived and developed in this particular domain directly allow complex sets of high-dimensional instrumental measurements to be processed without the need of preliminarily compressing them and reducing their size. However, recent technological advances—impacting especially the field of analytical chemistry—are bringing this subject back into the *chemometric spotlight* under slightly different perspectives. For example, today's spectroscopic platforms are constantly being adapted and/or miniaturised so as to enable also on-line or on-site applications and, in such circumstances, the design of simple, efficient, low-cost sensors spanning reduced, but informative ranges/subsets of wavelength channels are essential to guarantee a satisfactory characterisation performance. On another hand, mass spectrometers as well as hyperspectral cameras can nowadays deliver tons of GBs of data in very short times and *via* one-step sampling procedures. Often, though, only a smaller portion of such data actually encode the most useful and meaningful information required for tackling tasks like multivariate classification or linear unmixing¹.

In the light of all this, this Research Topic was originally proposed with the aim of gathering contributions from scientists working in different disciplines, but sharing common interests in the conceptualisation and utilisation of algorithmic strategies for variable selection. Two of such contributions have reviewed and compared some of the most popular existing tools devised for this purpose: [Armstrong et al.](#) focused specifically on variable selection when discrimination problems are dealt with, while [Westad and Marini](#) assessed to what extent the variables retained by means of these tools are relevant when it comes to interpreting predictive-type models. [Latchoumane et al.](#) have introduced the use of *N-CovSel* for the extraction of key features from front-face

1 For example, filtering out non-informative (noisy) regions of the recorded signal profiles may actually improve the performance of predictive models constructed on the measurements at hand.

fluorescence spectra recorded for differentiating healthy and diseased fruit products. Finally, [López-Fornieles et al.](#) have extended its algorithmic scheme for handling multispectral satellite image time series registered for remote sensing-based crop and plant growth monitoring.

Overall, as far as the editors are concerned, this Research Topic has undoubtedly permitted to stress the importance and relevance that variable selection can have in both basic and applied research scenarios.

Author contributions

RV wrote the first draft of the editorial. J-MR contributed to its revision. All the authors approved it for submission.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.