



25^{èmes} rencontres HélioSPIR
Montpellier, 11 & 12 juin 2024

Résumés des communications



	<p>Association HélioSPIR <i>Réseau de spectroscopie proche infrarouge</i> www.heliospir.net</p>
---	--

HélioSPIR est l'association francophone dédiée à la spectrométrie dans le proche infrarouge.

HélioSPIR a vocation à fédérer les scientifiques et les utilisateurs de la technologie SPIR au sein d'un réseau et à promouvoir l'utilisation de la spectroscopie proche infrarouge. Fondée en 2004 autour de la communauté scientifique d'Agropolis à Montpellier, l'association dépasse maintenant les contours de la région Occitanie et de l'hexagone. C'est un pôle de compétences à dimension internationale dans le domaine de la spectroscopie proche infrarouge.

HélioSPIR organise chaque année une ou deux sessions de rencontres scientifiques. C'est un moment privilégié d'échanges autour de diverses thématiques autour de la spectroscopie proche infrarouge et de découverte des derniers travaux de la communauté. C'est également l'occasion de découvrir ou redécouvrir les équipements de spectroscopie et d'imagerie hyperspectrales des principaux fabricants du secteur.

Président : G. Chaix ; adjoint : J.-M. Roger

Secrétaire : V. Rossard ; adjointe : A. Cambou

Trésorier : C. Fontange ; adjoint : R. Cinier

Conseil d'administration : V. Baeten, D. Bastianelli, S. Beaumont, A. Cambou, G. Chaix, R. Cinier, M. Ecarnot, C. Fontange, A. Herrero-Langreo, M. Loudiyi, S. Mas-Garcia, T. Ricour, J.M. Roger, V. Rossard, S. Roussel

Comment citer ce document

HélioSPIR, 2024. Résumés des communications présentées aux 25èmes rencontres HélioSPIR, Montpellier (France), 11-12 juin 2024. D. Bastianelli, G. Chaix, Eds. Association HélioSPIR, Montpellier (France), 40p. DOI : 10.19182/agritrop/00228

Comment citer un résumé particulier

Auteur1, Auteur2... Auteur n, 2024. Titre du résumé. In : HélioSPIR, 2024. Résumés des communications présentées aux 25èmes rencontres HélioSPIR, Montpellier (France), 11-12 juin 2024. D Bastianelli, G Chaix, Eds. (DOI : 10.19182/agritrop/00228), Association Héliospir, Montpellier (France). Numéro de page.



Publié sous licence *Creative Commons* CC-BY

Jchemo: Chemometrics and machine learning on high-dimensional data with Julia

Matthieu LESNOFF

¹UMR SELMET, CIRAD, 34398 Montpellier – France
Email : Matthieu.lesnoff@cirad.fr

Mots-clefs : Chimiométrie, Machine learning, Package, Langage Julia

Julia (<https://julialang.org>) is a programming language designed for high performance. It is an open source project made available under the MIT license. The language tries to tackle the “two-language problem” referring to the fact that many scientific codes are prototyped in a slow but flexible language (to test an idea quickly) but then have to be moved to a faster (e.g. C++) but less flexible language for practical applications. Julia allows fast computations with simple and easily readable coding. Works on Julia began in 2009. Julia's syntax is now considered stable, since version 1.0 in 2018 (actual version June 2024: 1.10.4), with many registered available packages and a very active users' forum (<https://discourse.julialang.org>).

The proposed poster will present Jchemo [1] (<https://github.com/mlesnoff/Jchemo.jl>), a Julia package (tool-box) dedicated to chemometrics and machine learning in general.

- **Why did I decide to switch in 2021 from the language R to Julia for my chemometrics works?** Trying to run a PLSR (25 LVs) with $n = 1e6$ samples and $p = 500$ variables with my R function crashed systematically my working session (with a I9 Intel processor). With the same computer and function but written in Julia, the computation took 8 seconds.
- **Why did I choose Julia compared to Matlab?** Since Julia is free.

Jchemo was initially dedicated to partial least squares regression (PLSR) and discrimination (PLSDA) models and their extensions, in particular locally weighted PLS models (kNN-LWPLS-R & -DA). The package has then been expanded to various dimension reduction and regression/discrimination models.

Beside usual chemometrics methods (signal preprocessing, PCA, PLS etc.), multi-block methods are available for dimension reduction (e.g. MBPCA, ComDim, rCCA, etc.) and regression/discrimination (MBPLS, ROSAPLS, SOPLS, etc.). Various ridge and sparse models are proposed as well as many nonlinear models useful for modeling heterogeneous data (kernel latent variables/ridge, kNN, RF, SVM). The syntax of Jchemo is very consistent between all the functions **and therefore can be learned and used by non-specialists of programming.**

[1] M. Lesnoff. Jchemo: Chemometrics and machine learning on high-dimensional data with Julia. 2021, <https://github.com/mlesnoff/Jchemo>. UMR SELMET, Univ Montpellier, CIRAD, INRA, Institut Agro, Montpellier, France