Advanced clustering to integrate spatial socio-ecological components in decision-making: a case study on sustainable agriculture

Stéphane Bourrelly  
*University of Lyon III, UMR 5600 (EVS) CNRS, stephane.bourrelly@univ-lyon3.fr*

Antonino Marvuglia  
*Luxembourg Institute of Science and Technology (LIST), antonino.marvuglia@list.lu*

Ian Vázquez-Rowe  
*Pontificia Universidad Católica del Perú, ian.vazquez@pucp.pe*

Follow this and additional works at: [https://scholarsarchive.byu.edu/iemssconference](https://scholarsarchive.byu.edu/iemssconference)

Part of the Civil Engineering Commons, Data Storage Systems Commons, Environmental Engineering Commons, Hydraulic Engineering Commons, and the Other Civil and Environmental Engineering Commons

Bourrelly, Stéphane; Marvuglia, Antonino; and Vázquez-Rowe, Ian, "Advanced clustering to integrate spatial socio-ecological components in decision-making: a case study on sustainable agriculture" (2016). *International Congress on Environmental Modelling and Software and Software*. 44.  
[https://scholarsarchive.byu.edu/iemssconference/2016/Stream-B/44](https://scholarsarchive.byu.edu/iemssconference/2016/Stream-B/44)

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Advanced clustering to integrate spatial socio-ecological components in decision-making: a case study on sustainable agriculture

Stéphane Bourrelly, Antonino Marvuglia, Ian Vázquez-Rowe

*University of Lyon III, UMR 5600 (EVS) CNRS (stephane.bourrelly@univ-lyon3.fr)
Luxembourg Institute of Science and Technology (LIST) (antonino.marvuglia@list.lu)
Pontificia Universidad Católica del Perú, Peru (ian.vazquez@pucp.pe)

Abstract: Accounting for the direct and indirect role of ecosystem services (ES) in supporting human activities is a priority among governmental targets to develop strategies of impact mitigation with the mutual effect of increasing human well-being. Although a great deal of research is currently being done on ES assessment, the volume of available data and the complexity of mechanisms make the decision-making process difficult. We highlight the strengths of unsupervised clustering for handling geographic multidimensional datasets. We explain why the outputs, both in qualitative and quantitative form, can help stakeholders to conceive sustainable policies. At the same time, we point out the fact that hierarchical clustering algorithms are currently still the most used in decision-making, although they rarely provide optimal results. Moreover, the spatial connectivity between the studied administrative areas is never taken into account, even though many phenomena occurring in a spatial unit influence the neighbouring areas. To illustrate these statements we present a case study in Luxembourg, which focuses on the chemical emissions of agricultural fertilisers and the socio-demographic outcomes. We present an ongoing R package combining hierarchical clustering and dynamic reallocation algorithms to improve the significance of results. It also integrates spatial statistics and advanced clustering methods to take into account the interconnectivity of spatial units. In addition, the algorithm computes the usual statistical criteria to assess the clustering quality. Unfortunately, statistical criteria never guarantee the true representation of the real world; and basing decisions on unrepresentative outputs might induce unfavourable consequences. We present a function using a variable selection strategy so to identify the best clustering through a trade-off between statistical criteria and decisional composite indicators - measuring the socio-ecological disfavour and the overall fertiliser impacts on ecosystems.

Keywords: Spatial clustering; machine learning; sustainable agriculture; decision-making; lattice data.