Title:

ELSA: A new entropy-based statistic to measure local and global spatial structure

Authors & affiliations:

Babak Naimi[1], Nicholas A.S. Hamm[2], Thomas A. Groen[3], Andrew K. Skidmore[3,4], Albertus G. Toxopeus[3], Sara Alibakhshi[5]

[1] Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, The Netherlands
[2] Geospatial and Geohazards Research Group, University of Nottingham, China
[3] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, The Netherlands
[4] Department of Environmental Science, Macquarie University, Australia
[5] Department of Built Environment, Aalto University, Finland

Abstract:

Introduction: There are several local indicators of spatial association (LISA) that allow exploration of local patterns in spatial data. Despite numerous situations where categorical variables are encountered, few attempts have been devoted to the development of methods to explore the local spatial pattern in categorical data. To our knowledge, there is no indicator of local spatial association that can be used for both continuous and categorical data at the same time. We introduce ELSA which can be used for exploring and testing local spatial association for continuous and categorical variables. We introduce the entrogram for exploring global spatial structure. We provide the R-package elsa for making these computations

Methods: ELSA at site i is defined as ($E_i = E_{ai} \times E_{ci}$) and is calculated within a local window. E_{ai} summarizes the attribute dissimilarity between site i and the neighbouring sites. E_{ci} is the Shannon entropy normalized by the number of categories in the entire dataset. For categorical data the dissimilarity measure can account for hierarchical classification structures. For continuous data, the data are first binned using an iterative approach that balances information loss against the number of classes. A non-parametric bootstrap randomization approach is used to test the spatial association against a null distribution. Finally ELSA is extended to the global context using the entrogram.

Results: ELSA is evaluated for three scenarios: (1) categorical data with no hierarchical structure, (2) categorical data with a hierarchical structure and (3) continuous data. We use both synthetic and actual datasets (CORINE land cover map for Spain). These demonstratedthe functioning of ELSA in comparison to other LISA's. ELSA gives consistent results, also where there is global autocorrelation. Finally we demonstrate the entrogram.

Discussion: Our research demonstrates the applicability of ELSA and the consistency of the results.

Naimi et al. (2019) *Spatial Statistics*, 29, 66-88.