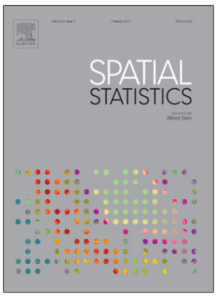
Accepted Manuscript



Including spatial correlation in structural equation modelling of soil properties

Marcos E. Angelini, Gerard B.M. Heuvelink

PII: S2211-6753(17)30297-X
DOI: https://doi.org/10.1016/j.spasta.2018.04.003
Reference: SPASTA 293

To appear in: *Spatial Statistics*

Received date : 17 November 2017
Accepted date : 6 April 2018

Please cite this article as: Angelini M.E., Heuvelink G.B.M., Including spatial correlation in structural equation modelling of soil properties. *Spatial Statistics* (2018), https://doi.org/10.1016/j.spasta.2018.04.003

Including spatial correlation in structural equation modelling of soil properties

Marcos E. Angelini[a,b,c,*], Gerard B.M. Heuvelink[b,c]

[a]*Instituto de Suelos, Instituto Nacional de Tecnología Agropecuária (INTA), N. Repetto y Los Reseros s/n 1686, Hurlingham, Provincia de Buenos Aires, Argentina*
[b]*ISRIC-World Soil Information, P.O. Box 353, 6700 AJ Wageningen, The Netherlands*
[c]*Soil Geography and Landscape group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands*

Abstract

Digital soil mapping techniques usually take an entirely data-driven approach and model soil properties individually and layer by layer, without consideration of interactions. In previous studies we implemented a structural equation modelling (SEM) approach to include pedological knowledge and between-properties and between-layer interactions in the mapping process. However, it typically does not consider spatial correlation. Our goal was to extend SEM by accounting for residual spatial correlation using a geostatistical approach. We assumed second-order stationary and estimated the semivariogram parameters, together with the usual SEM parameters, using maximum likelihood estimation. Spatial prediction was done using regression kriging. The methodology is applied to mapping cation exchange capacity, clay content and soil organic carbon for three soil horizons in a 150 100-km^2 study area in the Great Plains of the United States. The calibration process included all parameters used in `lavaan`, a SEM software, plus two extra parameters to model residual spatial correlation. The residuals showed substantial spatial correlation, which indicates that including spatial correlation yields more accurate predictions. We also compared the standard SEM and the spatial SEM approaches in terms of SEM model coefficients. Differences were substantial but none of the coefficients changed sign. Presence of residual spatial correlation suggests that some of the causal factors that explain soil variation were not captured by the set of covariates. Thus, it is worthwhile to search for additional covariates leaving only unstructured residual noise, but provided that this is not achieved, it is beneficial to include residual spatial correlation in mapping using SEM.

Keywords: Digital soil mapping, `lavaan`, Pedometrics, Regression kriging
2017 MSC: 00-01, 99-00

*Corresponding author
Email addresses: `angelini75@gmail.com` (Marcos E. Angelini), `gerard.heuvelink@wur.nl` (Gerard B.M. Heuvelink)