

## CLUSTER ANALYSIS OF SPATIALLY DEPENDENT DATA

1PAWLOWSKY-GLAHN, V., 1JIMÉNEZ-GONZÁLEZ, S. and 1EGOZCUE, J. J.  
1Universidad Politécnica de Cataluña, ETSECCPB, Departamento de Matemática  
Aplicada III, Barcelona, Spain.

Cluster analysis is a useful technique to detect different populations in a data set. When data are spatially dependent, it is necessary to apply techniques which take into account both the similarity of samples in the variable space and their spatial dependence. This can be done using a modified version of Ward's method based on a spatial extension of the Mahalanobis distance, which incorporates an estimate of the matrix of cross-covariances for each spatial distance considered. But the estimation of the cross-covariance matrix as a function of distance assumes data to belong to the same population.

To solve this problem, an iterative approach is used. It is based on the assumption that the groups have different spatial means but identical patterns of spatial variability and that the variables considered are stationary within each group. This is equivalent to assume that the semivariograms and cross-semivariograms are the same within each group. Thus, at each iteration the number of assumed clusters is increased and the semivariograms and cross-semivariograms are recomputed using only pairs of samples from same groups until the function stabilises. Case studies have been performed to illustrate the approach using simulated and real granulometric data. Results, compared to those obtained with the traditional method of Ward, are less spherical in variable space and spatially more homogeneous; the price is a larger variability in variable space.