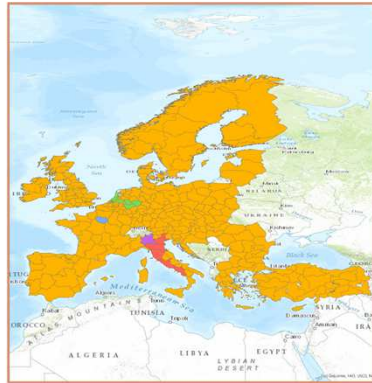# An evaluation of spatially-constrained regionalization routines

Mark V. Janikas[1], Juan Carlos Duque[2] and Liang-Huan Chin[1]
1. Environmental Systems Research Institute (esri) 2. Universidad EAFIT, RiSE Group

ARISEL

Objective Function= 1955.62

SKATER
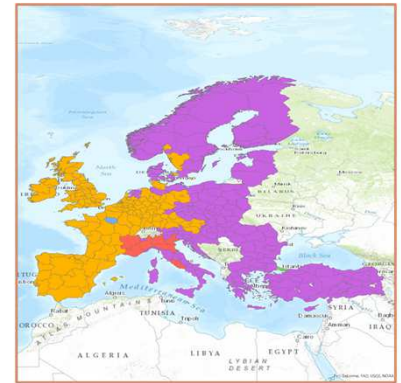

Objective Function= 1426.12

Abstract

This paper focuses on spatially-constrained regionalization routines, where we are interested in discretizing features in such a manner that minimizes heterogeneity within groups while honoring a set of contiguity constraints. A variety of algorithms are considered from two distinct software packages: **ArcGIS** and **ClusterPy**. A set of experiments are provided that distinguish the relative quality of the algorithms contained within the software as they relate to varying degrees of spatial dependence and alternative proportions of group/sample sizes. The methodologies include a breadth of advanced heuristics that build regions from a set of seeds as well as alternative flavors of a technique that relies on dividing minimum spanning trees. Attention is given to both the power and performance of the methods in order to guide applied researchers in empirical analyses.
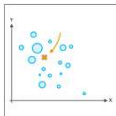
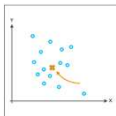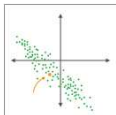## What is Central?

Mean Center: Calculate the Geographic Mean Center for each existing cluster and then start the search at the closest node.
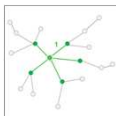
Weighted Mean Center: Weighting the Mean Center by the data being analyzed can result in starting nodes that are distinct from their current group.
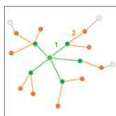
Data Center: Finds the node in an existing partition that is most like its centroid.
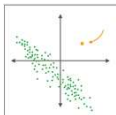
Data Outlier: Identifies the location in a group that differs most from its centroid.

Link Centrality: Begins search from the node in each cluster that has the most links as defined by the neighborhood structure.
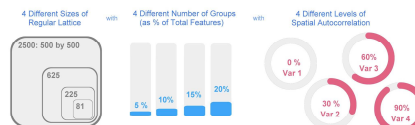
Second Order Link Centrality: Accounts for Second Order contiguous neighborhood links.
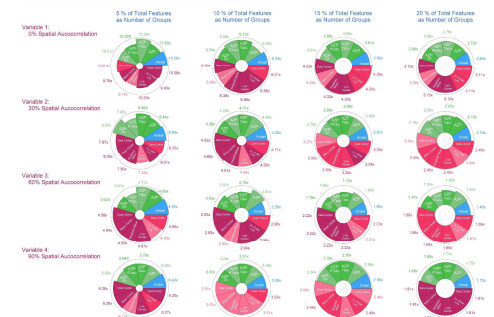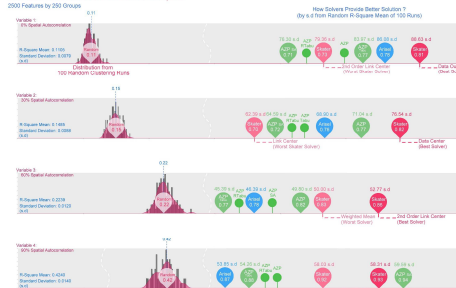
Simulation Design

Overall

Detail Example

The Software