Abstract

Conventional methods for cartographic classification are often solely based on underlying attributable values. There are numerous algorithms to determine the resulting classes, such as Jenks Optimized classification, but they do not account for the spatial patterns that are inherent to spatial data. This can cause a visual disruption of areas that would normally be considered a cluster, thus making the overall message of a map harder to grasp.

With a method called "Autocorrelation-Based Regioclassification" TRAUN & LOIDL (2012) introduced an alternative approach that takes spatial properties into account and classifies data values in respect to their statistical and spatial properties. My work builds upon their method and shows how their approach has been implemented for ArcMap using ArcObjects in C#.

The main objectives of the Add-in are to (a) decrease visual noise, (b) emphasize statistically relevant outliers, (c) deal with overlapping classes that result from the method and (d) provide visual tools to aid the understanding and interpretation of the classification.

Method

This use case applies the classification method to a polygon dataset that features the percentage of African American people in 755 counties of the southeast United States as reported by the US Census Bureau in 2000.

- The scatterplot features each county (map-polygon) as a point. The x-axis represents the attribute value of the county and the y-axis shows the mean of all its adjacent counties.
- The 5 blue areas (separated by thin black dotted lines) correspond to the class breaks resulting from a Jenk’s natural breaks classification which is only based on the county value (x-axis).
- The thick red dotted line is the regression through the point cloud and it’s slope equals the Moran’s I statistic for spatial autocorrelation.

The orthogonal projection of each point on the regression line and a redclassification based on their resulting distribution along the regression line produces 5 new diagonal areas separated by the thin red dotted lines. These are now used to reassign the class membership of each county. E.g. Point A was previously in class 1 and is now upgraded to class 2 while point B shifted from class 5 to class 4.

Conclusion

- This method results in a quantitatively measurable reduction of the visual complexity. (e.g. McEachran’s complexity index for choropleth maps).
- The outlier detection ensures to exclude statistically relevant features from recategorization and thereby emphasizes them in contrast to the overall smoother appearance of the map.
- ArcMap does not satisfyingly facilitate the implementation of overlapping classes. Tedious workarounds allow a rudimentary visualization of the class legend.
- The visual aids (scatterplot, histogram, reference classification) have proven to be useful (and necessary) tools to support the understanding and result interpretation of this classification method.

Outlook / temporal neighbors

Distinct time steps of a dataset can be additionally considered to extend the concept of “neighborhood” with the temporal dimension. This allows to apply the methods benefits to time-enabled visualizations in ArcMap. The weight of the spatial neighbors (adjacent polygons) and temporal neighbors (previous and later time steps of the same polygon) can be individually adjusted to suit the datasets characteristics.

References

Based on the online version at http://www.unsalzburg.at/.

Christoph Mayrhofer, University of Salzburg, Austria