Spatial and geostatistical analysis for regional dialectology

JACK GRIEVE
Aston University

Although methods developed for spatial and geostatistical analysis are commonly used in a variety of fields that work with spatial data, these methods have rarely been used in regional dialectology, despite their obvious application. This presentation therefore introduces a range of common statistical methods for the analysis of spatial data and demonstrates how these methods can be used to analyze regional linguistic variation. The specific methods described in this presentation are variogram analysis, ordinary kriging, point pattern analysis, global spatial autocorrelation analysis, and local spatial autocorrelation analysis, all of which can be conducted using common statistical software such as R. Taken together, these methods constitute a powerful, standardized, and statistically grounded set of techniques for the analysis, visualization, and modeling of regional linguistic variation. In particular, these methods can be used to model spatial patterns exhibited by linguistic variables, to predict the values of linguistic variables at unobserved locations, to analyze the distributions of individual linguistic forms, to test for overall patterns of spatial clustering in the values of quantitative linguistic variables, and to identify the location of spatial clusters in the distribution of quantitative linguistic variables. These methods also provide principled solutions to many of the most persistent methodological problems in regional dialectology, including the identification of linguistic variables exhibiting significant patterns of regional variation and the plotting of isoglosses. In addition to introducing and demonstrating the application of these methods through the analysis of dialect data from American English, this presentation will also provide the necessary R code so that these methods can be used by other researchers to analyze their own regional linguistic datasets.