

Spatial Data Analysis

Prepared by Rebecca E. Curtis

Definition: Spatial data analysis explicitly considers the spatial relationship present in an examined phenomenon. The initial procedures of analysis include the set of generic methods of exploratory analysis and the visualization of data, in general, through maps.

These techniques permit the description of the distribution of the variables of study, the identification of observations that are outliers not only in relation to the type of distribution but also in relation to its neighbors, and the exploration of the existence of patterns in the spatial distribution. The spatial inferential models are usually presented in three groups: continuous variation, discrete variation, and the point processes. The resolution of a spatial problem may involve the utilization of one of them or the interaction of some or even all of them.

There are three main types of data used in spatial data analysis:

- Events or point patterns – phenomena expressed through occurrences identified as points in space, denominated point processes. Some examples are: crime spots, disease occurrences, and the localization of vegetal species.
- Continuous surfaces – estimated from a set of field samples that can be regularly or irregularly distributed. Usually, this type of data results from natural resources survey, which includes geological, topographical, and ecological maps.
- Areas with Counts and Aggregated Rates – means data associated to population surveys, like census and health statistics, and that are originally referred to individuals situated in specific points in space. For confidentiality reasons these data are aggregated in analysis units, usually delimited by closed polygons (Câmara, et. al.).

Application:

In 1854, one of the many cholera epidemics was taking place in London and John Snow used a spot map to illustrate how cases of cholera were centered around a water pump. The map on the right presents the location of the deaths due to cholera and the water pumps that supplied the city, allowing the clear identification of one of the locations, in Broad Street, as the epicenter of the epidemics. This was one of the first examples of spatial data analysis, where the spatial relationship of the data significantly contributed to the advancement in the comprehension of a phenomenon (Crosier).



References/Sources:

Bailey, T. and A. Gattrel (1995). *Spatial Data Analysis by Example*. London, Longman.

Câmara, G., A. Monteiro, S. Fucks, and M. Carvalho. *Spatial Analysis and GIS: A Primer*. http://www.dpi.inpe.br/gilberto/tutorials/spatial_analysis/spatial_analysis_primer.pdf

Crosier, Scott. *John Snow: The London Cholera Epidemic of 1854*. <http://www.csiss.org/classics/content/8>

Fotheringham, A.S., C. Brunson and M.E. Charlton. 2000. *Quantitative Geography*. London: Sage.

Haining, R.P. 1990. *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge: Cambridge University Press.

Martin, D. 1995. *Geographic Information Systems: Socioeconomic Applications*. London: Rutledge.

O'Sullivan, D. and D. Unwin. 2002. *Geographical Information Analysis*. New York: Wiley and Sons.