

Objects as subjects: modeling the uncertainty in spatial big data.

Alfred Stein

University of Twente, Enschede, the Netherlands



Abstract

Current research interest on big data is increasing. With this research, clear statistical questions emerge. The increase is particularly relevant in spatial and spatio-temporal studies, where for example satellite images are of an increasingly fine resolution in space and time and the number of bands can be high. In this way, real 'big data' are produced. Still, uncertainty remains and spatial statistical tools are exceptionally suited to handle such uncertainty. It offers opportunities to summarize the data, and express measures of variation and uncertainty as well as their modeling in the spatial temporal domain. Much of the subjects of study is so far condensed as point observations, or administrative units. This presentation will show that treating spatial data as objects has various advantages. It starts with selecting objects directly from spatial data sources (in particular images). The objects are then characterized and possibly classified. Next they are used as a big data set for a particular modeling purpose. These steps may include uncertainties. For modeling of those uncertainties, random sets as a generalization of random variables are proposed. The presentation will be illustrated with examples on bushfires and slums. It is shown how the way from ontology to uncertainty analysis is relevant to address most of the problems.