Evaluation of Hotspot Detection Method based on Echelon Structure

Fumio Ishioka*
School of Law, Okayama University, Okayama, Japan
fishioka@law.okayama-u.ac.jp

Koji Kurihara
Graduate School of Environmental and Life Science Okayama University,
Okayama, Japan kurihara@ems.okayama-u.ac.jp

The importance of statistical analyses for spatial data has grown in various scientific fields. A statistical technique for the spatial data have been established by the radical advances of the computing power, for example geographical information system (GIS). Especially, the study for finding a hotspot, such as disease clustering or hazard map is one of important issue. There are several approaches to detecting hotspots from different kinds of spatial data. Among them, spatial scan statistics has been widely used as a hotspot detection method. The spatial scan statistic is based on the likelihood ratio associated with the number of events inside and outside a circular scanning window. However, it is noted that a non-circular shaped hotspot, such as the shape made by a river or a road cannot be detected. To solve this problem, several non-circular scanning techniques have been proposed. However, in these previously reported methods, a detected hotspot sometimes has an unlikely and uncommon shape which requires long computation times in cases where there is a large amount of regional data. In addition to these methods, we have proposed a technique using an echelon analysis as a non-circular shaped hotspot detection. The echelon analysis is a useful technique for systematically and objectively investigating the phase-structure of spatial regional data. The echelons are derived from changes in topological connectivity. To use the spatial structure based on echelon analysis as scanning methods, we can effectively detect hotspots. In this paper, we detect hotspots by using an echelon scan method for spatial data, and compare them with those detected by a previous study's method. In addition, we propose an all possible scan method which can absolutely detect a hotspot with the highest likelihood. We demonstrate the further validity of echelon scan method by comparison with all possible scan method for simulated data.

Key Words: Spatial scan statistic, cluster detection, spatial data