Spatial Statistics 11 (2015) 113–131

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Specialized agglomerations with Lattice data: Model and detection

Christian Haedo [a], Michel Mouchart [b,c,*]

[a] Centro Interuniversitario de Investigaciones sobre Desarrollo Económico, Territorio e Instituciones (CIDETI), Alma Mater Studiorum-Università di Bologna in the Argentine Republic, Argentina

[b] Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), UCLouvain, Belgium

[c] Center for Operations Research and Econometrics (CORE), UCLouvain, Belgium

ARTICLE INFO

Article history:
Received 26 November 2013
Accepted 24 November 2014
Available online 10 December 2014

Keywords:
Relative specialization
Specialized agglomeration
Lattice data
Spatial clustering
Spatial cluster detection
Permutation bootstrap

ABSTRACT

This paper develops new statistical and computational methods for the automatic detection of spatial clusters displaying an over- or under- relative specialization spatial pattern. A probability model is used to provide a basis for a space partition into clusters representing homogeneous portions of space as far as the probability of locating a primary unit is concerned. A cluster made of contiguous regions is called an agglomeration. A greedy algorithm detects specialized agglomerations through a model selection criteria. A random permutation test evaluates whether the contiguity property is significant. Finally this algorithm is run on Argentinean data. Evaluating the proposed methodology concludes the paper.

Contents

* Correspondence to: ISBA, voie du Roman Pays 20 - L1.04.01, B-1348 Louvain-la-Neuve, Belgium. Tel.: +32 10474318; Mob.: +32 472335927.

E-mail address: michel.mouchart@uclouvain.be (M. Mouchart).

http://dx.doi.org/10.1016/j.spasta.2014.11.003
2211-6753/

1. Introduction

1.1. Spatial clustering framework

Spatial cluster detection is an important issue in many fields, ranging for instance from astrophysics to forest ecology, including epidemiology and many others. There is accordingly quite a vast and rapidly expanding literature. In particular, Openshaw et al. (1988), Anselin (1995), Ester et al. (1996, 1998), Chaudhuri and Marron (1999, 2000), Wakefield et al. (2000), Glaz et al. (2001, 2009), Lawson and Denison (2002), Diggle (2003), Kulldorff et al. (2003), Tango (2010), Saini and Kaur (2014) illustrate the relevance of this topic in all these fields as well as the great variety of approaches for detecting spatial clusters. This paper is basically motivated by spatial economics.

A significant group of contributions uses the spatial scan statistic first proposed by Naus (1965) and further developed by Cressie (1977), Openshaw et al. (1988), Besag and Newell (1991), Kulldorff and Nagarwalla (1995), Kulldorff (1997), Glaz et al. (2001, 2009), Tango and Takahashi (2005), Kulldorff et al. (2006), Neill et al. (2006), Loh (2011), among many others; these contributions have been used to detect disease cluster in epidemiology. Our approach is essentially an elaboration of the basic ideas of the scan method proposed by Besag and Newell (1991) in which, given a complete set of regions, individual regions are extracted and progressively aggregated to contiguous regions in order to find the most significant cluster, evaluating all possible cluster schemes that can be formed from these regions. We work in a strictly discrete space, with lattice data on a fixed set of polygons. A model selection procedure, based on the BIC criterion, is used in the algorithm for building optimal clusters. In most works, significance testing is concerned with a null hypothesis of spatial uniformity whereas in this paper we are concerned with the significance of the contiguity condition, hence the reliance on permutation bootstrapping.

1.2. Economic background of industrial agglomeration

Economic activity is spatially concentrated. Spatial concentration generates agglomeration economies, notably upstream–downstream linkages, which help firms become more productive. These positive effects involve a critical mass of workers and infrastructure, and dense networks of suppliers and collaborators. The central role of agglomeration economies on the spatial structure of the economy has inspired a large literature focused on trying to understand their causes or origins and the dynamic of these economies, as well as most adequate industrial policies for their consolidation and promotion wherever the industrial agglomerations are weak or nonexistent. It has been recognized that the industrial policy objectives can be better fulfilled if they are more sensitive to places and sectors in design and delivery (Donato, 2002; Nathan and Overman, 2013).

The explanations about the existence and determinants of the industrial agglomerations started with the pioneer works of Marshall (1890), Scitovsky (1954), Arrow (1962), Becattini (1979), and Romer (1986), regarding the localization economies of firms located near other firms of the same industry; the literature often refers to it as forces that lead to the clustering of firms of similar activities and to regions that tend to specialize. Furthermore, Jacobs (1969), Henderson (1985), Lucas (1988) and Glaeser et al. (1992) are concerned with the aspect of the urban economies that refer to the advantages arising from the spatial concentration of firms of diverse activities (urbanization