Métodos de Investigación

Master en TIG, Universidad de Alcalá

Imparte: Francisco Escobar

SESIÓN 7

- Área de estudio y unidad de observación. El PUEM
- Presentación y discusión sobre la definición del área de estudio y la unidad de observación.

- **Área de estudio y unidad de observación. El PUEM (Eagleson, 2003)**

The MAUP is a form of ecological fallacy associated with the aggregation of individual data into areal units for geographical analysis. The MAUP is endemic to all spatially aggregated data. The MAUP consists of two interrelated parts: the scale effect, and the aggregation or zoning effect (Fotheringham & Wong 1991).

- The scale effect is the tendency, within a system of modifiable areal units, for different statistical results to be obtained from the same set of data when the information is grouped at different levels of spatial resolution. (See Figure 2.3.)
- The aggregation or zoning effect is the variability in statistical results obtained from a set of modifiable units as a function of the various ways these units can be grouped at a given scale, not as a result of the variation in the size of those areas. (See Figure 2.4.)

Census data is an example of the MAUP. It is collected from every household but released only at census boundaries every four years. When the values are averaged through the process of aggregation, variability in the dataset is lost, and the values of statistics computed at different boundary resolutions will be different. (This is the scale effect.) Additionally, the data analyst may receive different results depending on how the spatial aggregation occurs. (This is the zoning effect.) The MAUP is integral to the display of demographic data as the information relayed through mapping and statistics is a product of the size, shape and scale of the administrative boundaries used in the data aggregation process.

The MAUP is fundamental to the display of demographic data as the information that people perceive can be altered by the size, shape and scale that is used for display (Fotheringham & Wong 1991; Goodchild et al. 1993). Many researchers have researched the magnitude and effect of the MAUP. In particular, Openshaw and Taylor (1979; 1981); Monmonier (1991) and Openshaw and Alvanides (2001) have highlighted that the results of statistical analysis can be varied by altering spatial boundaries. (Research efforts into the MAUP are discussed further in chapter four.) This problem of boundary design is very important for ensuring that users of administrative boundary data are correctly informed about the data and the applications for which it is useful.

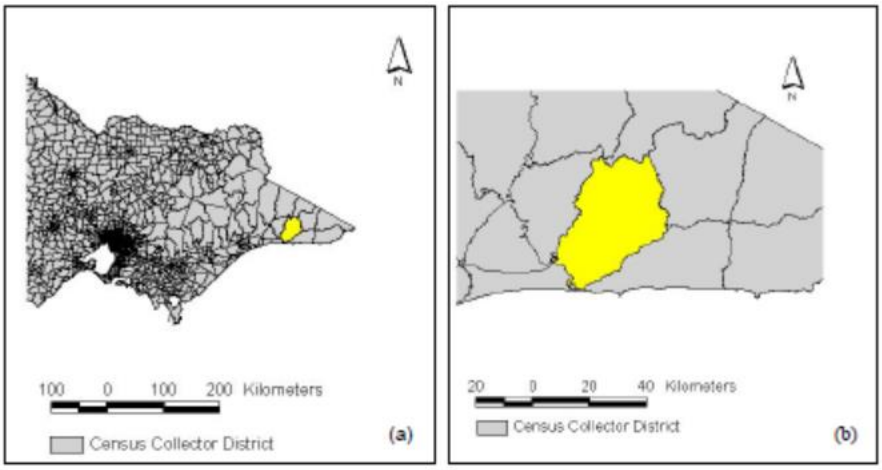

Figure 2.3: An illustration of the impact of scale. Depending on the scale the selected region has a greater/lesser impact on the analyst.

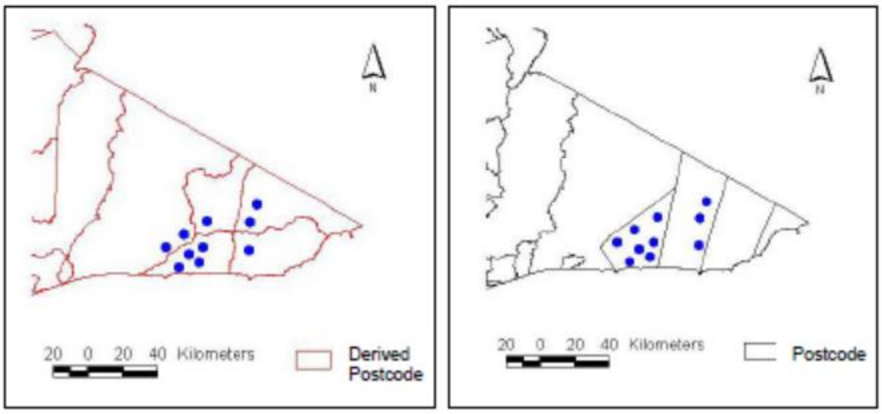

Figure 2.4: An illustration of the potentially different results when data is aggregated into two distinctly different boundary systems.

Postcode	SLA		Proportion of the population in the postcode common to the SLA
2339	5640	Otway (S)	100.0
2737	3760	Kernang (S)	100.0
3000	4601	Melbourne (C) - Inner	089.6
3000	4602	Melbourne (C) – Remainder	010.4
3002	4602	Melbourne (C) – Remainder	100.0
3003	4602	Melbourne (C) – Remainder	100.0
3004	4602	Melbourne (C) – Remainder	21.2
3004	6480	St Kilda (C)	19.6
3004	6880	South Melbourne (C)	59.2
3005	4602	Melbourne (C) – Remainder	100.0
3011	2840	Footscray (C)	99.9
3011	4602	Melbourne (C) – Remainder	0.1
3012	2840	Footscray (C)	64.9
3012	7080	Sunshine (C)	35.1
3013	2840	Footscray (C)	100.0
3015	2840	Footscray (C)	21.2
3015	8080	Williamstown (C)	78.8
3016	8080	Williamstown (C)	100.0

Table 3.1: A concordance between postcode and SLA. Source: Department of the Premier and Cabinet Department of Treasury Victoria (1992, p. 91)

Introducing possible solutions

1. A surface model approach to data integration

Bracken and Martin (1989) and Bracken (1994) have developed a method of surface model integration. This method of data integration transforms the zone-based census data and point-based address data into two continuous raster based surfaces. Due to the raster based structure of the data integration of the surfaces overcomes many of the limitations encountered with the original vector based data sets.

2. Areal interpolation

The problem of integrating data between non-coterminous boundary systems can be restated as the problem of transferring data from one set of boundary units (source zones) to another set of units (target zones). This process is commonly referred to as areal interpolation (Goodchild & Lam 1980; Goodchild et al. 1993; Flowerdew & Green 1994; Flowerdew & Green 1991; Trinidad & Crawford 1996; Xie 1995). Trinidad and Crawford (1996) detail data interpolation as a method of calculating corresponding attribute values in all regions of a target map, based on values collected in a source map.

Simplistically, the methodology without supplementary information involves calculating the area of overlap between the source zone, s, that lies within a target zone t. This area is denoted by a_{st}, and the known source-zone population by U_s. As a result, the target-zone population can be estimated by equation 1 (Goodchild et al. 1993):

$$V_t = \Sigma\, U_s(a_{st}/\Sigma a_{st})$$

The point-in-polygon method

A second example of data interpolation is the point-in-polygon method. This method of data interpolation, involves the transfer of attribute values from the source zone to a target zone (Preparata & Shamos 1985; Okabe & Sadahiro 1997; Burrough 1986). This process is completed using a point-location algorithm. This algorithm creates a grid over the initial map and then searches to find the centroids in the grid and the number of attributes contained within it. The number of attributes are added and interpolated onto a regular grid system (Okabe & Sadahiro 1997).

Supplementary information

In an attempt to improve the accuracy of interpolation between source and target zones, numerous methods of interpolation have been derived that utilise supplementary information.

- Land use data

- Street network

Bibliografía

- EAGLESON, Serryn, 2003, *GIS applied to administrative boundary design*, PhD thesis, University of Melbourne.
- Eagleson, S and Escobar, F, 2003, "Administrative boundary design in support of SDI objectives" in Williamson, Rajabifard and Fenney (eds) *Developing Spatial Data Infrastructures – From concept to reality*, London etc. Taylor & Francis, pp.249-263. (ISBN 0-415-30265-X)
- Jones, S, Eagleson, S, Escobar, F and Hunter, G, 2003, "Lost in the mail: the inherent errors of mapping Australia Post Postcodes to ABS derived postal areas", *Australian Geographical Studies*, 41(2), pp. 171-179.

Eagleson, S, Escobar, F, and Williamson, I, 2003, “Automating the administration boundary design process using Hierarchical Spatial Reasoning theory and Geographical Information Systems”, *International Journal of Geographic Information Sciences*, Volume 17, Number 2, pp 99 – 118

- Eagleson, S, Escobar, F, and Williamson, I, 2002, “Hierarchical Spatial Reasoning theory and GIS technology applied to the automated delineation of administrative boundaries”, Computers, Environment and Urban Systems, Vol. 26 (2002), pp 185-200.

- Tarea para las sesiones 10 y 11: Revisión bibliográfica. El alumno realizará una presentación resumiendo la revisión bibliográfica pertinente a su tema de investigación. Los criterios de evaluación tendrán en cuenta los métodos de búsqueda, la cantidad de referencias, la relación sintética realizada con ellas y la exposición. Este entregable supone el 20% de la nota final de la asignatura.