

**CONFERENCE OF EUROPEAN STATISTICIANS**

**UN/ECE Work Session on Methodological Issues Involving the Integration of Statistics and Geography**

(Neuchâtel, Switzerland, 10-12 April 2000)

Topic (iv): Spatial analysis in a statistical context and disclosure control procedures

**SPATIAL ANALYSIS – WHY BOTHER?**

Submitted by the Office for National Statistics, United Kingdom<sup>1</sup>

**Invited paper**

**ABSTRACT**

This paper is intended to promote discussion on the value of, and issues surrounding, the increased use of ‘spatial analysis’ in national statistical offices. It considers the benefits and, critically, some of the hurdles that lie in our way if we are to push GIS a little further into this area. The paper argues that there is real value in, and a responsibility on, statistical offices to develop and share experience in this area. The argument is based on a mix of personal experience and wild generalisation – so views are sought on the conclusions drawn here.

There is no attempt made here to formally define spatial analysis or to consider any of the techniques in any detail. This type of information will be provided by a publicly available Internet site which will be available to all delegates in advance of the WorkSession in Neuchatel in April 2000. The principles behind the website are explained at the end of the paper.

**I. SO WHAT’S THE ISSUE?**

1. For some time now ‘spatial analysis’ has been viewed as the next big thing in statistical GIS. It is increasingly identified as an area for further research – at some stage in the future when we have some spare time. A few of us (I discount myself here) have even known what we were talking about.

Brief aside:

For the sake of argument here let’s loosely define spatial analysis as ‘a set of procedures and techniques for analysing and modelling spatial data’.

All of us use our existing GIS systems for some operations that might be covered under this type of definition – data overlay, buffering, querying of data, area measurement – but these are only the start of the tools and techniques available. A few examples of the type of techniques included might include nearest neighbour analysis, network analysis, cluster analysis, automated zone design, spatial interaction models, and spatial modelling and interpolation.

---

<sup>1</sup> Prepared by Alistair Calder.

The actual definition doesn't matter much – what matters is that spatial analysis provides us with a set of clever analytical and modelling operations that can add value to our spatial data and that (most of us) don't use already.

2. Some offices are fairly far forward with the application of this type of analysis – but they are in the minority – and often the tools are only applied in a limited fashion. In many cases we use academics or researchers and we limit the use of these tools to individual projects. In some respects this is inevitable for now (see the discussion on problems below) but this approach severely limits the value we can get from these tools.

## II. SO WHY BOTHER?

3. Well, there are some good reasons and some bad reasons why we might think about developing spatial analysis and modelling in our organisations....

### II.1 Bad reasons

- Because it looks like an interesting new toy to play with. As we get on top of the technology and some of the issues start to fall away we are all getting a bit bored of conventional GIS.
- We will be able to create some nice 3D spatial models, fly-throughs and virtual worlds which will be fun and impress those people who we want to impress.
- Because people are now starting to understand and use GIS – so we might lose some of our power. Spatial analysis is good because it is something we can know about that no one else in the organisation will understand.
- Fortunately there are most powerful reasons than these why we should be adopting spatial analysis ...

### II.2 Good Reasons

- Because we have a responsibility towards our organisations to find out about potential uses for these tools. Statistical organisations care A LOT about every variable they collect and spend a great deal of time and money polishing their data and carefully analysing the relationships between variables. As our options and approaches to geo-referencing improve we now have access to a new variable for each data item – space. To ignore this fact is just plain irresponsible.
- Because geography is a special type of variable – if we are lucky it can be common to all of our data – whether they come from different topics or from external sources. GIS helps us pull together data in a way that isn't possible otherwise – and spatial analysis can make this integration of data easier still.
- Because the interactions between data items in space matter. For many of the phenomena we are trying to investigate the distance between items (or some aspect of their spatial relationship) has an influence on their characteristics – we should not ignore this.
- Because we are not always as good as spotting patterns as we think we are – maps are easy to produce and well understood but we do not always get the information from them we might. We are too easily fooled by poor map design and often have no chance of identifying patterns of complex interrelated variables. Spatial analysis can add some rigour to otherwise subjective analysis.
- Because modelling of complex interactions (with space as the glue) can help in analysis and estimation and save us money.

4. And finally if we must have a self centred reason here ...

- Because, with the improved referencing (and awareness) of many datasets in government, there may be a sudden political will to leverage more information and perform better analysis on the data we have. We would do well to be on top of the problems here before we are forced to by circumstances.

### **III. SO WHY HAVEN'T WE DONE MORE ABOUT THIS BEFORE?**

5. Again I suspect there are some good and some bad reasons ...

#### **III.1 Bad reasons**

- Because we don't even know what it is.
- This is certainly the case in some organisations. This is understandable but is soon going to be unacceptable - let's talk about ways of sharing experience to sort this out.
- Because we are a bit scared – what if spatial analysis turns out to be statistically out of our depth? – Let's ignore it and maybe it will go away.
- It won't. The statistics behind some (though not all) of these techniques are relatively sophisticated – and see the note below on the particular issues relating to spatial data - but that doesn't mean that we can afford to duck the issue here.
- Perhaps one of the issues here is that many Geography or GIS areas of statistical offices are made up of former mapmakers and survey takers. Where this is the case perhaps we need to bite the bullet and either develop the skills ourselves or form sensible allegiances with statisticians and methodologists.
- Because we have lots of proper work on – this is low priority - go away and leave us alone.
- A marginally better excuse - but again I suspect we won't get away with this approach for long.

#### **III.2 Good excuses**

- Spatial analysis is complicated

6. Yes, in some respects some of it is – but only in the way that multivariate analysis or writing computer programmes is complicated. We can handle it ... but we are going to have to learn new skills and new ways of thinking about our data if we are going to gain full value from the tools available.

7. Again this is a responsibility on Geography and GIS departments. Our organisations spend a massive amount of money collecting data – we need to ensure that we use all of the tools available to gain every ounce of value from the (relatively inexpensive) analysis.

- Geography causes problems

8. Nothing new here then – but yes geographic data does present some special issues here. Spatial analysis is already well developed and understood in some scientific disciplines. The types of spatial data we have to manage (and particularly the implications on propagation of error) do present us with different types of problem.

9. This doesn't mean we can ignore this – what it means is that we need to develop tools for spatial analysis which take proper account of the types of data and the degree of error which we have to deal with.

- GIS doesn't give us the tools we need

10. Well, that's true too. A few GIS systems (IDRISI, ESRI Spatial Analyst) go some of the way there.. but of course these tools are focussed on common requirements and are often better suited to environmental (environmental impact, terrain modelling) or specific commercial analysis (site location, trade area analysis).

11. Although we have discussed this approach before there would seem to be little value in statistical offices lobbying GIS companies to improve their implementation of tools particularly suited to our needs. Our requirements are not well enough defined yet and we are, let's face it, a limited market. There may be value in us make some statement of requirements – but we cannot afford to wait for implementation of spatial analysis tools better suited to our specific needs. We need to compare notes and use the tools available as best we can.

12. The literature on spatial analysis includes extended discussion on how spatial analysis tools should relate to GIS. Should new tools be embedded in our GIS systems as extra functions? – should we use sophisticated statistical modelling packages for the modelling work and just use GIS to visualise and present results? – or should we aim for something in-between? In truth the whole discussion is rather academic - for now we should look for pragmatic (though statistically sound) solutions using whatever tools offer us the best solutions.

#### **IV. CONCLUSION – AND THE WAY FORWARD**

13. It seems fairly clear that the future of GIS is going to depend on improved spatial analysis and our ability to leverage more information from the data we have already collected. In the light of this I believe that there is an overpowering case for us (or those of us who have not already done so) to look hard at the potential for increased use of spatial analysis and modelling within our organisations.

14. If the tools are not available – if the barriers really are too great or if the levels of error inherent in our data make the analysis invalid – then that is a good reason for holding back here. If, however, the key reason for not investigating these tools further is that we don't understand them, or are frightened to try, then we need to do some further work.

15. Equally I believe that this is a clear case where international co-operation between statistical offices can be of real value. I suspect that between us we actually have a considerable amount of experience in using these tools and that many of the technical issues are independent of national geographic differences and so transferable.

16. In the light of this I propose the establishment of a website where we can share information and experience on using spatial analysis tools. This site is proposed as a parallel development to the map-design website discussed at the Ottawa WorkSession (to be discussed and formalised at the Neuchatel meeting). A brief note on the proposed structure of the new site is attached below.

17. Both sites will be available to all delegates for review in advance of the WorkSession in Neuchatel in April 2000.

#### **Delegates are invited to :**

- Comment on the conclusions drawn in the above paper;

- Review the draft website in advance of the WorkSession;
- Provide comments on the value of this approach and suggestions for how we might take this work forward.

## V. NOTES ON THE PROPOSED SPATIAL ANALYSIS WEBSITE

18. At first it is intended that the proposed site will be little more than a framework to start discussion. In the longer term it should provide a simple introduction to the techniques and issues – and above all a framework to help us share our knowledge and experiences in using spatial analysis and modelling.

19. The basic structure will be as follows (though early comments are welcomed):

- a summary of what 'spatial analysis' means for us - a fairly wide definition I think;
- a discussion of the main types of spatial analysis tools which are available (and notes on statistical implications);
- a framework into which individual countries can place examples / case studies of work they have done in this area;
- links to important sites and resources.

20. An (electronic) template will allow individual countries to supply information about specific pieces of spatial analysis work which they think are of interest. The template might include :

- the 'problem' or purpose of the analysis;
- the technique used;
- the software used;
- the experience of doing it;
- the results & any recommendations;
- notes on any statistical issues of this type of analysis;
- any recommended contacts or links / sources of information.

21. I intend to obtain a few case studies before the session to act as examples – but it is hoped that other countries will provide further examples thereafter. I hope that over time the site will build to a databank of case studies and resources that will be of mutual value and will also act as a focus for discussion.

22. Delegates will be informed of the location of the draft website at the end of March (it currently exists only as single sheet of paper!). It has been already agreed that in the longer term this (and the map-design site) will be hosted by the UNECE.

23. It is hoped to formalise the management and plans for development of the website at the Neuchatel meeting. In the meantime comments on this approach, or offers of help (and particularly offers of contributions) are very welcome.