Map Line Ups: Using Graphical Inference to Study Spatial Structure

Roger Beecham* City University London

Jason Dykes City University London

Cagatay Turkay City University London

Aidan Slingsby City University London

Jo Wood City University London

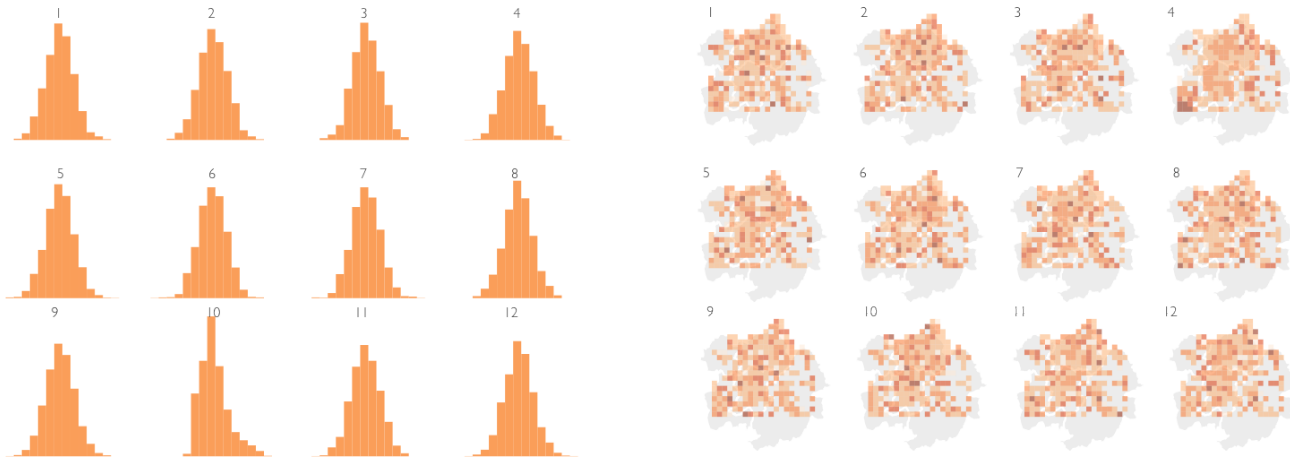

Fig. 1. Line-up method applied to two sample datasets.

Index Terms—graphical inference, confirmation bias, spatial data analysis.

1 Introduction

Wickham et al. [4] introduced the idea of *graphical inference* to Infovis research during the 2010 VisWeek conference. Here, the authors highlighted a distinction between data analysis undertaken by statisticians and that done by Infovis researchers. In Infovis, data analysis software tools are often built for insight discovery – for maximising the chances of finding new patterns and structures in datasets. By contrast, in statistics, researchers have at their disposal a set of techniques designed to help minimise the chances of finding false structures in datasets. An extreme in either direction is unhelpful, but a possible consequence of too little scepticism might be confirmation bias: a tendency to see mainly patterns that meet one's preconceptions. The techniques proposed by Wickham et al. [4] for graphical inference, specifically the line-up and Rorschach protocols, attempt to bridge these divides by "provid[ing] a tool for skepticism [sic.] that can be applied in a curiosity driven context".

In Wickham et al. [4], several datasets and numerous statistical graphics are used to demonstrate the Rorschach and line-up protocols. A particularly interesting area of application is in analysing *spatial* structure. A well-rehearsed concept in spatial analysis disciplines is that of spatial dependence, or *Tobler's first law of geography*. This states that: "Everything is related to everything else, but near things are more related than distant things" [3]. *Tobler's first law* thus describes the process of spatial autocorrelation: the co-variation of properties over geographic space. Clearly the extent to which *Tobler's first law* is true will vary with the data that are being observed. Testing for whether or not spatial autocorrelation exists, as well as how much it exists, is therefore a common question when analysing spatial data.

Whilst there are several quantitative measures of spatial autocorrelation, humans have difficulty discerning random from non-random spatial process when those patterns are depicted visually [1]. We believe that Wickham et al.'s [4] line-up protocol – where an analyst must select a 'real' dataset from a set of n decoy plots (as in Figure 1) – may therefore offer a practical means of supporting analysts in visually interpreting spatial structure.

The success or appropriateness of such a technique is nevertheless heavily contingent on its context of use: the specification of an appropriate *null hypothesis*, the sensitivity of the test and clearly the spatial process and geographic region being studied. We wish to evaluate the line-up method, considering these different constraints, and suggest guidelines for its use. Here, we discuss some challenges that we hope to address through experimental research.

2 Varying the 'null' datasets

The line-up protocol suggested by Wickham et al [4] works as follows:

- An analyst identifies a spatial pattern of interest: The spatial distribution of crime rates within a study region.
- A *null hypothesis* is defined: Crime rates within a study region are independent of location.
- n null datasets (or decoy plots) are created under this *null hypothesis*.
- The analyst is asked to correctly identify the 'real' dataset from the decoy plots created under the *null hypothesis*.

One method for creating decoy datasets under the *null hypothesis* of spatial independence as described above would be to create n copies of the observed data, maintaining the original crime attribute data, but permuting the location values (as in Figure 1). This is often how numeric significance tests of spatial independence are computed: a Monte Carlo simulation procedure is used to define pseudo significance values by shuffling the attribute values in a dataset among the

*e-mail: roger.beecham.1@city.ac.uk

locations [2]. Such a technique may be effective in certain analysis contexts. However, it may often be the case that, when represented visually, a *null hypothesis* of spatial independence leads to decoy plots that 'look' unrealistic: even when performing data transformations, in the case of crime rates normalising by local population size, it is reasonably easy to detect the real data from the decoys that appear in Figure 1. This may relate to a wider problem around using such statistical tests in spatial analysis: given *Tobler's first law*, it may not make sense to create null models that assume complete spatial randomness [2]. This criticism perhaps especially applies to the graphical inference example that appears in Figure 1. The outlier is the only true geographical distribution; since the decoys show no spatial dependence, they do not look geographical, and therefore the underlying *null hypothesis* is not plausible. In our experimental research, we hope to explore appropriate 'background' levels of spatial auto-correlation that might be used to create realistic or analytically-useful decoys: to vary the *null hypothesis* that is being tested.

3 Varying the geography

A separate consideration is that, when practically studying spatial process, the topology of a region being studied is likely to affect how spatial patterns in that region are interpreted. This is perhaps especially true when spatial processes are analysed in choropleth maps: polygons representing administrative areas in a region might be configured such that certain structures appear visually salient. Random spatial processes might then appear as clustered, purely due to the shape of the underlying spatial units.

There are other more inter-subjective, and therefore intractable, challenges. It is likely that individuals have certain preconceptions about how particular phenomena (for example, local crime rates) are spatially articulated, and these preconceptions may affect how individuals interpret spatial structure. Similarly, if 'real' geographic regions are used in our proposed experimental research, participants familiar with those regions may have particular ideas about how spatial processes are likely to manifest themselves in those regions.

Clearly, we will need to be consider these factors in our experimental design. We may wish to simulate a spatially regular geographic region and choose not to reveal to participants the phenomena that are described in the test dataset. However, since we are interested in applying the line-ups methods in real data analysis contexts, we may wish to explore the effect of at least varying the topology of study regions.

4 Varying the number of decoys

The primary objective of this research will be to test participants' performance when interpreting spatial line-up tests with differently defined null hypotheses. This will be achieved by simulating spatial structure under different definitions and levels of auto-correlation, and the aim will be to suggest spatial statistics, or thresholds in spatial statistics, that relate to visually perceived thresholds. A slightly more straight-forward factor for experimental research might also be to evaluate the effect of varying the number of decoy plots that are presented to research participants. Wickham et al. [4] suggest setting the number of decoys (n) at the traditional boundaries of statistical significance: where n=19, the probability of correctly identifying the 'real' dataset from the decoys by chance is therefore 5%. Such thresholds may not be so clearly defined when the line-up method is practically implemented, and we may wish to test for the difference between these numerically-defined and visually-perceived thresholds in our experimental research.

5 Conclusion

Graphical inference is a particularly promising technique for visually-inclined spatial data analysts: empirical research already demonstrates that there is not always a direct mapping between quantitative definitions of spatial structure and human visual interpretation of that structure [1]. This fact nevertheless emphasises that there are obvious practical implications for the ways in which tests for graphical inference are implemented. We wish to empirically evaluate the effectiveness one such technique – the line-ups method – for analysing spatial structure. In particular, we wish to explore the effect of altering the *null hypothesis* used to generate decoys in spatial line-up tests. This is early stage work and we hope the DECISIVe workshop will provide opportunity to discuss numerous controls and factors we may wish to explore in our research experiments.

Acknowledgments

This research has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) to Project VALCRI under the EC Grant Agreement No. FP7-IP-608142 awarded to Middlesex University and partners.

References

[1] A. Klippel, F. Hardisty, and R. Li. Interpreting spatial patterns: An inquiry into formal and cognitive aspects of Tobler's first law of geography. *Annals of the Association of American Geographers*, 101(5):1011–1031, 2011.

[2] D. O'Sullivan and D. Unwin. *Geographic Information Analysis*. John Wiley Sons, New Jersey, USA, 2002.

[3] W. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240, 1970.

[4] H. Wickham, D. Cook, H. Hofmann, and A. Buja. Graphical inference for infovis. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):973–979, 2010.