The Seventh Annual International Crime Mapping Research Conference
March 31 - April 3, 2004, Boston, Massachusetts

On Being More Robust About 'Hot Spots'

Allan J. Brimicombe
BA(Hons) M.Phil. Ph.D. C.Geog. FRGS FGS
Professor & Head, Centre for Geo-Information Studies,
University of East London, UK
a.j.brimicombe@uel.ac.uk

Abstract

Within GIS-based identification of hot spots, two new, complementary techniques are described and applied. First is a variable resolution approach to cluster detection which overcomes many problems of traditional point density estimation and provides a bridge between hot spot detection based on incident counts and hot spot detection based on risk. Second is a method of robust data normalisation which readily identifies outliers. When applied to traditional area-based choropleth mapping it identifies count- and/or risk-based hot spots consistent with point density approaches. It can also identify cool spots. When coupled with a geographical gains curve, these new techniques becomes a decision-making tool for problem-oriented policing and crime reduction.

Introduction

Underlying most approaches to problem-oriented policing and crime prevention is the understanding that crimes tend to form patterns (Brantingham & Brantingham, 1984). These patterns are the discernible manifestation that a process is at work which, once understood, can be modified or stopped through an appropriate set of interventions (legal, social or situational). The patterning of crime occurs in one or more key dimensions: spatially, temporally or in the attributes of the modus operandi (MO). The patterns of most interest would either form a clustering or exhibit degrees of regularity. Both have an element of predictability. Geographical Information Systems (GIS) have been used since the early 1990s to assist in identifying geographical clusters of crime, commonly known as *hot spots* (Harries, 1999).

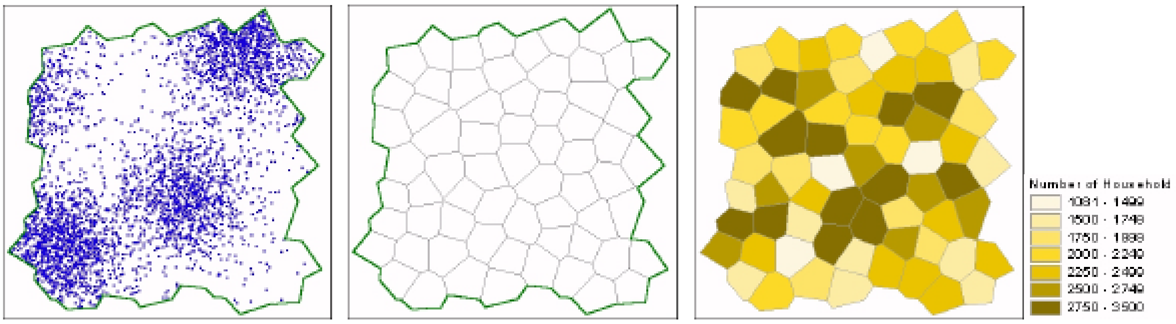
Records of crime are not just for the police. Though such data sets may be compiled and managed by the police and though they may be an imperfect enumeration of total crime, they also have social and political meaning. Crime statistics are widely used to inform political and legislative agenda, central and local government policy, strategic and operational planning in numerous agencies and in individual decision making around everyday activities. Crime data therefore need to be routinely shared, particularly with other agencies, not just as summary statistics but at much more detailed levels. Reiner (2000) and Garland (2000) have analysed the post World War II trend of rising crime in parallel with the rise of the consumer society, the loosening of social cohesion and growing levels of social exclusion, economic insecurity and inequality. The politicisation of crime first in the US (1960s) and then in the UK (1970s) was in response to high crime rates becoming a normal fact of society, the growing fear of crime and the perception that traditional penal-welfare solutions were failing. This has resulted in a two part strategy: an adaptive strategy which stresses partnership and prevention and a sovereign state strategy which stresses increased levels of control and punishment (Garland, 2000). Prevention through partnership have seen state and non-state agencies having to co-ordinate at strategic and operational levels so as to improve community safety by reducing criminal opportunities whilst raising crime-awareness. For example, in the UK, the 1998 Crime and Disorder Act legislated for the setting up of 376 local crime and disorder reduction partnerships (CDRP) throughout England and Wales. Their function is to audit local crime and disorder, identify causes, develop appropriate strategies based on an intelligence-led, problem-solving approach and to evaluate the effectiveness of responses. Strategies may include both social and situational responses. The CDRP were further instructed in Section 97 of the 2002 Police Reform Act to take due account “of the levels and patterns of crime and disorder in the area”, that is, to identify and respond to hot spots of crime and disorder. This responsibility for the integration and co-analysis of crime data with disorder data (e.g. graffiti, abandoned vehicles) and other social data (e.g. census, unemployment) means that all such data must be shared across partnerships. However, due to concerns over confidentiality, crime data are de-personalised usually by aggregation to area-based administrative units before release. Thus, whilst the police can detect hot spots using the highest resolution data, usually point events, partnerships using de-personalised aggregated data must find alternative area-based techniques. Most partnership analysts use popular, off-the-shelf tools (such as Microsoft® Excel and MapInfo®) and tend to be proficient mainly in the basic functionality.

This paper concerns itself not just with the detection of hot spots from point event crime data but goes on to consider how such analyses might be carried out in a partnership, crime reduction setting using de-personalised, aggregated data.

The Data Set

The release of high resolution crime data into the public domain can be problematic in terms of safeguarding individual privacy. Furthermore, such data can (and often do) contain errors and/or incomplete fields. To avoid these problems and have a data set which fulfils a number of research objectives, a data set has been carefully simulated to have many of the characteristics of real data for Greater London, UK. The point events shown in Figure 1(a) are for 3781 residential burglaries over a 12 month period. The bounded area (the jurisdiction) is approximately 80 square kilometres (31 square miles) and is subdivided into 65 districts to reflect administrative and census geography (Figure 1(b)). The number of households (Figure 1(c)) also reflects the reality of Greater London (2001 census) with 147,900 households in the bounded area giving an annual burglary rate of approximately 27 per thousand households.

For 90% of the burglaries, the distribution has been devised using a cluster simulation algorithm – after all, to test hot spot detection there must be some hot spots. The remaining 10% are randomly distributed. The burglaries visually form three distinct areas of higher point density diagonally from the south west to the north east of the bounded area with another possible higher point density area in the north west. These higher density areas of burglary do not necessarily correspond with the distribution of households at a district level.



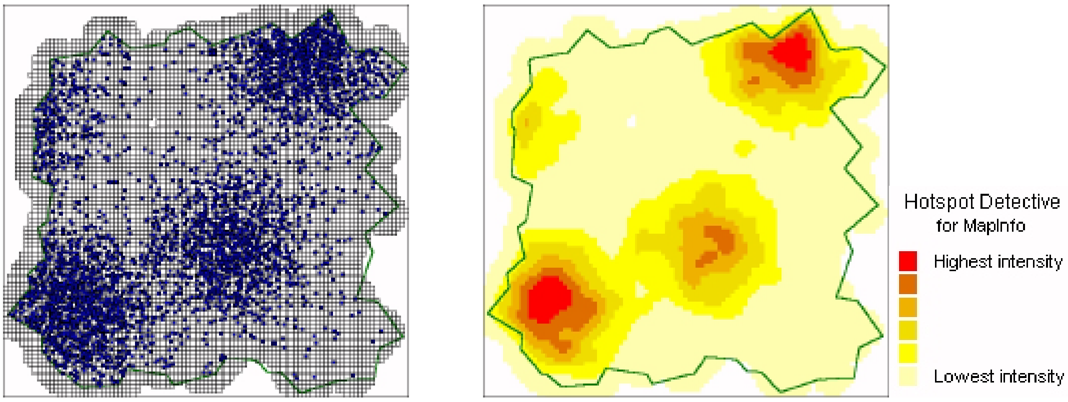
(a) (b) (c)

Figure 1: The data set: (a) point distribution of one year's residential burglaries; (b) districts; (c) number of households per district.

Hot Spot Detection

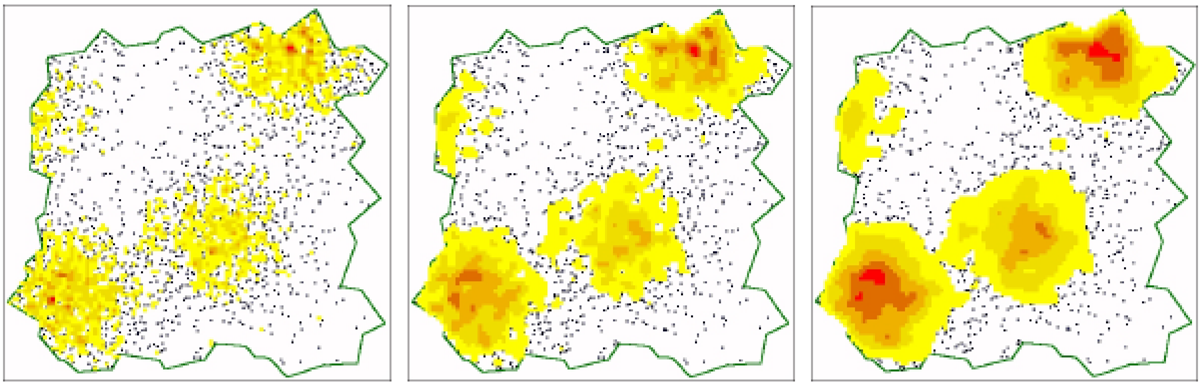
It has long been recognised that we absorb information from visualisations (images, graphics, maps) far quicker and in greater quantity than through any of our other senses. We are good at spotting patterns and inductively classifying things. But often these lack objectivity and consistency from one individual to another. Experiments have shown that considerable variation can arise in the visual perception of clusters in a distribution of point events (e.g. Sadahiro, 1997). Hence the need for more objective approaches in crime analysis and, in a digital age, computational techniques. GIS-based and other techniques for the detection of crime hot spots have been much discussed in the literature (e.g. Jefferis *et al*. 1998; Ratcliffe & McCullagh, 1999; McLafferty *et al*., 2000; Langworthy & Jefferis, 2000; Rogerson & Sun, 2001). Yet, there is no standard definition of a hot spot. The most commonly understood meaning focuses on crime *counts*, an elevated share of crime in a localised area. Hot spots based on counts inform the deployment of policing resources. Less commonly in crime analysis (but more commonly, for example, in epidemiology) are hotspots based on elevated *rates* (e.g. per thousand households). Such hot spots reflect the level of risk and thus inform deployment of crime reduction resources. For the same distribution of point events, hot spots based on counts will be different to those based on rates as the latter are not just a function of the distribution of point events but also the underlying population at risk. Ideally both should be used in problem-oriented policing.

One popular approach to hot spot detection based on counts is point density estimation with functionality available, for example, in CrimeStat®, Spatial Analyst extension to ArcView® and Hotspot Detective for MapInfo®. This is an interpolation that transforms the point events into a continuous surface which is then coloured up so that the peaks in the surface (the hot spots) are easily visualised. The algorithm of choice is kernel density estimation (Silverman, 1986; Bailey & Gatrell, 1995; Brunsdon, 1995) as this provides a smooth estimate of point density that is quick to calculate. An example is shown in Figure 2. As with any interpolation technique, parameters need to be set which are critical to the outcome. The first is the underlying grid size which will determine output resolution. The second is the radius of a circular window (or bandwidth) that steps across the grid cells. Reasonable values for these parameters can be difficult to estimate. There are a number of rules of thumb, but the author has not found them particularly helpful. Some of the software provide default values based on the size of area and number of points to be processed. Best practice would suggest a form of sensitivity analysis to identify optimum parameter values (Brimicombe, 2003a). Figure 3 shows such an approach for a fixed grid size (one hectare) and varying bandwidth. The maximum nearest neighbour distance between point events in Figure 1 is 574m or 12 times the median nearest neighbour distance of 47.5m; so the bandwidth was bracketed at three, six and nine times the median nearest neighbour distance. The effect with default legend is increased size and severity of hot spots! What then is the truth?



(a) (b)

Figure 2: Kernel density estimation using default settings in Hotspot Detective for MapInfo®: (a) grid pattern and points; (b) intensity surface (including a lowest intensity class).



(a) (b) (c)

Figure 3: Bracketing in search of an optimum – looking for objective hot spots or just an aesthetic graphic? (For legend see Fig 2(b); lowest intensity class not shown).

Theory seems to suggest that hot spots would be quite localised. High crime areas are primarily so because they are areas of high repeat offending and high repeat victimisation (Trickett *et al*., 1992; Townsley *et al*., 2003). In this regard Figure 3(a) with a few localised hot spots may be more correct than Figure 3(c) though the latter would undoubtedly be more popular with analysts due to its aesthetic qualities. Boundaries are also a problem for density estimation as densities inevitably decline towards the boundaries (Figure 2(b)) and hot spots thus defined cannot exist on the boundary. Software in the public domain by Atkinson & Unwin (2002) for MapInfo® does offer a buffer to avoid spurious values at boundaries but does not entirely overcome the problem of how to identify real hot spots that exist at boundaries. Finally, it is far from straightforward to equate density surfaces with the rate of occurrence (with reference to an at risk population) and hence an understanding of the risk to citizens. These criticisms of density estimation – sensitivity to parameters, boundary conditions, smoothing effect, choice of class intervals to identify hotspots, difficulty of identifying risk – lead us to look for alternative approaches.

A Variable Resolution Approach to Hot Spot Detection

The theory of adaptive recursive tessellations is given in Tsui and Brimicombe (1997a) with applications of their use for spatial analysis in Tsui and Brimicombe (1997b). Specific application to point pattern analysis can be found in Brimicombe and Tsui (2000), Brimicombe (2003b). At the heart of adaptive recursive tessellations is a variable resolution approach to space. No longer are scale and resolution treated as being uniform across an area but are allowed to vary spatially in response to the point pattern. This is achieved through a recursive decomposition of space, similar to quadtrees, but allowing variable decomposition ratios and rectangular cells. The algorithm makes no prior assumptions about the statistical or spatial distribution of points. Each point is treated as a binary occurrence of some phenomenon without further descriptive attributes. The decision to further decompose any one cell larger than the atomic cell size is based on the variance at the next level of decomposition and a heuristic on the number of empty cells that result. The atomic cell size is mediated between the average nearest neighbour distance and average cell size per point. Tests have shown the algorithm to be consistently effective in comparison with other approaches of point cluster detection (Brimicombe & Tsui 2000). The resulting clusters, or hot spots, are termed Geo-ProZones (GPZ) as they represent zones of geographical proximity in the point pattern. The important thing to note is that GPZ are not an interpolation, but a *segmentation* into polygons having internal consistency in the distribution and density of the point events within them. GPZ for Figure 1(a) are given in Figures 4(a) and 4(b).

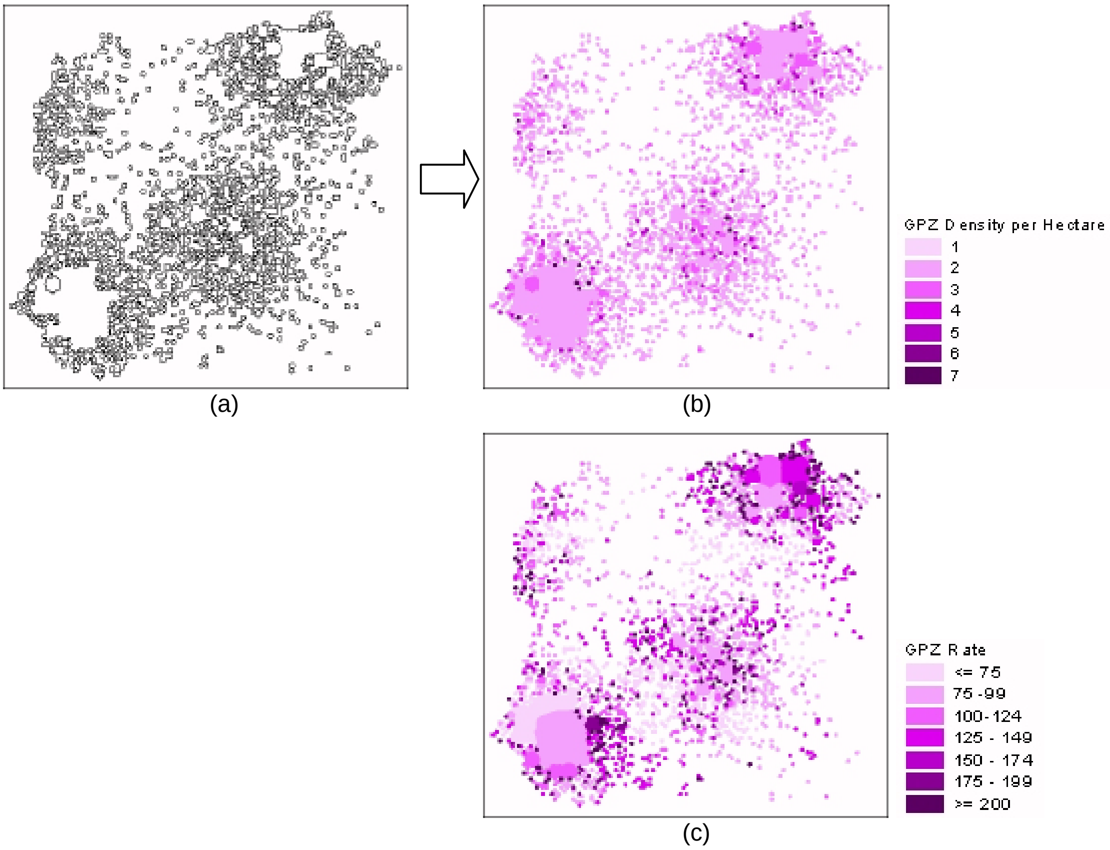
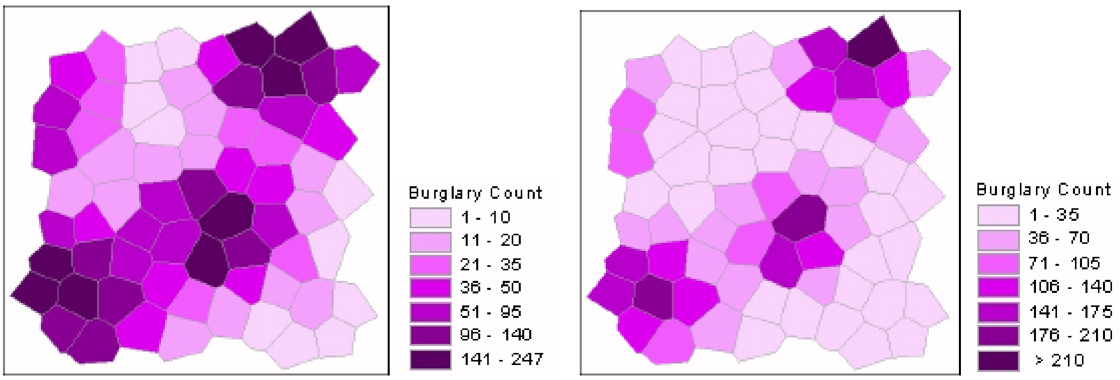


Figure 4: Variable resolution Geo-ProZone analysis: (a) polygon segmentation of point event density; (b) GPZ for density of burglaries per hectare; (c) GPZ re-segmentation for rate of burglaries per thousand households.

The pattern in Figure 4(b) reflects the pattern in Figure 3. The underlying speckle is because all events are mapped without smoothing. The highest densities, or what would be interpreted as hot spots, occur as more localised concentrations of repeat offending. The hotter areas of Figure 3(c) now come across as plateaux of uniform density. Because GPZ is a segmentation rather than an interpolation, it can be reconfigured for rates rather than point densities. Figure 4(c) shows GPZ as rates per thousand households from the underlying census data. The pattern of hot spots is quite different and identifies where citizens are at greatest risk. Some of these areas appear reasonably extensive, others quite localised.

Sharing and Analysing Aggregated Data

Approaches to hot spot detection using point event data (such as those described above) have come about due to the well-known limitations of area-based or choropleth mapping (see for example Monmonier, 1991). For each thematic map decisions must be made as to the number class intervals, the fixing of class boundaries and what colours to use. Aggregation of point events to area units suffers from the modifiable areal unit problem (Openshaw & Taylor, 1981). But some level of aggregation (usually to census units) is certainly necessary if crime counts are to be viewed as rates or correlated with possible causal variables. Of course, mapping rates further requires the correct identification of the appropriate 'at risk' base (households, population, school-age children and so on). Whilst the police may be able to avail themselves of point-based methods (though as we have seen, not without their own problems many of which, such as number of classes, class interval and application of colour, still persist), partnerships working on crime prevention issues receive de-personalised, aggregated data. The data in Figure 1(a) released by district as in Figure 1(b) would produce the well-known conundrum in Figure 5 when using defaults offered by mapping packages such as ArcView® or MapInfo® - where are the hot spots? But this is not dissimilar to the issue of malleable hot spots in Figure 3. Such problems are very much reduced, if not eliminated, by using some form of data normalisation.



(a) (b)

Figure 5: Problems of area-based mapping: (a) natural breaks; (b) equal interval.

Robust Normalisation

Normalisation is the application of a transformation to one or more data sets in order to make them fit the same distribution and thus exhibit some desirable characteristic(s). There are a number of such transformations, a very popular one being the Z transformation:

$$Z = (x - \bar{x})/\sigma$$

where $\bar{x}$ is the mean and σ is the standard deviation. Thus the mean will equal zero with all other data values being positive or negative in standard deviations. Standard deviations is also one of the default thematic mapping schemes in ArcView® and MapInfo®. However, Z transformation only works well with normally distributed data (bell-shaped), which crime and other social data rarely are. The burglary data in Figure 5 when presented as a histogram in Figure 6 shows that it is not normally distributed ($\bar{x}$ = 52; σ = 53) and that a Z transformation would result in a serious bias. A new alternative - *robust normalisation* – has recently been introduced (Brimicombe, 1999, 2000) and is proving effective for applications in crime mapping. The term 'robust' is used in naming the transformation because it uses the median and inter-quartile range from robust statistics (Hettmansperger & McKean, 1998). The outcome of robust normalisation is a distribution of median = 0, lower quartile = -1 and upper quartile = +1.

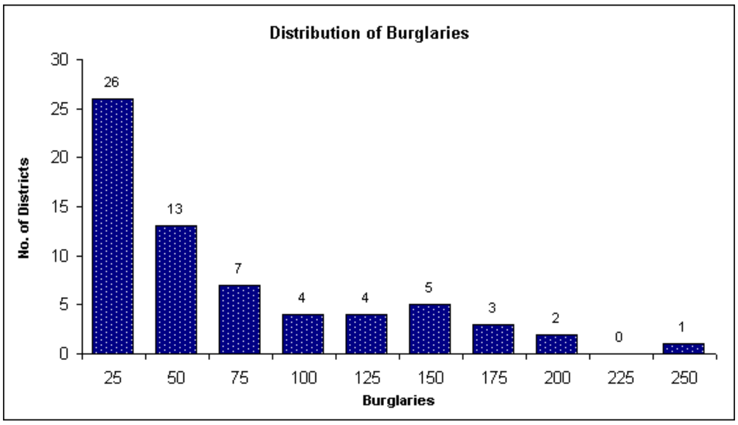


Figure 6: Histogram of burglaries by district in Figure 5.

Robust normalisation is achieved in the following manner:

$$RN = \frac{(x - median)}{(median - lower\_quartile)} \quad for\ x < median$$

$$RN = \frac{(x - median)}{(upper\_quartile - median)} \quad for\ x > median$$

$$RN = 0 \quad for\ x = median$$

The effect of the transformation is illustrated in Figure 7. This shows how the interquartile range (the middle 50% of values) from two dissimilar distributions are unified, whilst the extremes are allowed to 'float off' and still be identified. Within hot spot detection it is the extreme positive values (> 3) that are of most interest. The robust normalised distribution easily lends itself to five or seven class intervals with class boundaries at quartiles (in the seven class interval scheme the values immediately around the median are further separated). This can be used in a standardised way for all hot spot mapping and allows more objective comparisons between maps. The above algorithm is easily coded as a Microsoft® Excel macro.

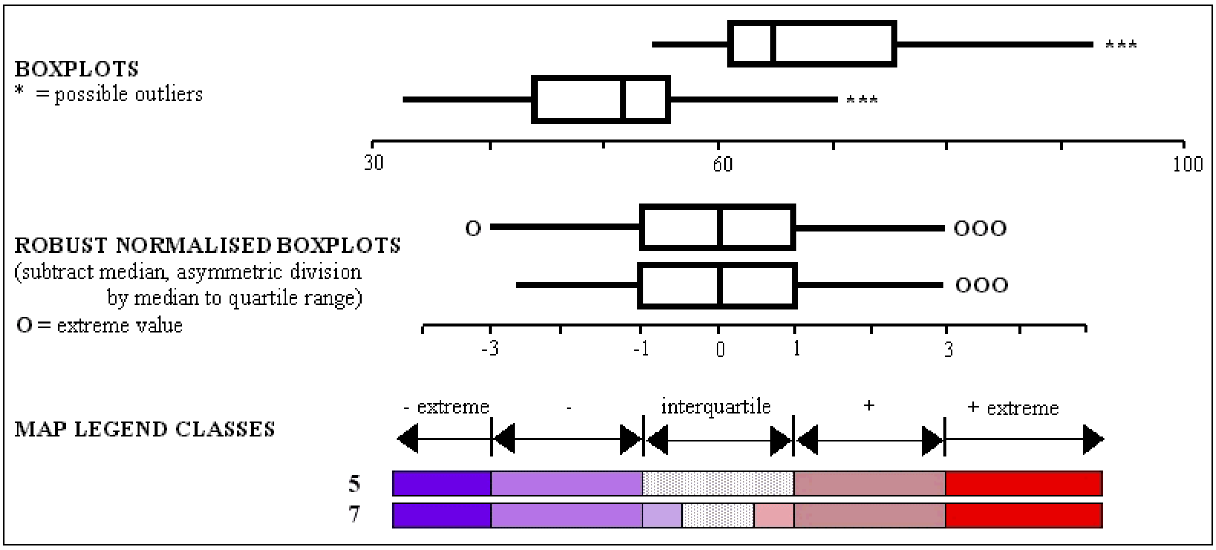


Figure 7: Illustration of robust normalisation and resulting map legends of 5 and 7 classes.

Robust normalisation can be applied to both to area-based data (Figure 8) and to density estimate interpolations. Figure 8(a) is a hot spot choropleth map consistent with the density estimate hot spot maps in Figure 3 and overcomes the problems of Figure 5. Figure 8(b) shows hot spots in the burglary rate and it is easy to spot where these conform and where they differ. Figures 8(c) and 8(d) show robust normalisation applied to GPZ densities and rates in Figure 4 and again overcome problems of arbitrary numbers of classes and class intervals. The two sets of maps in Figure 8 (district and GPZ) show reasonable conformance within the limits imposed by aggregation, and shows that partnerships working with aggregated data can also produce hot spot maps of counts and rates without being off target.

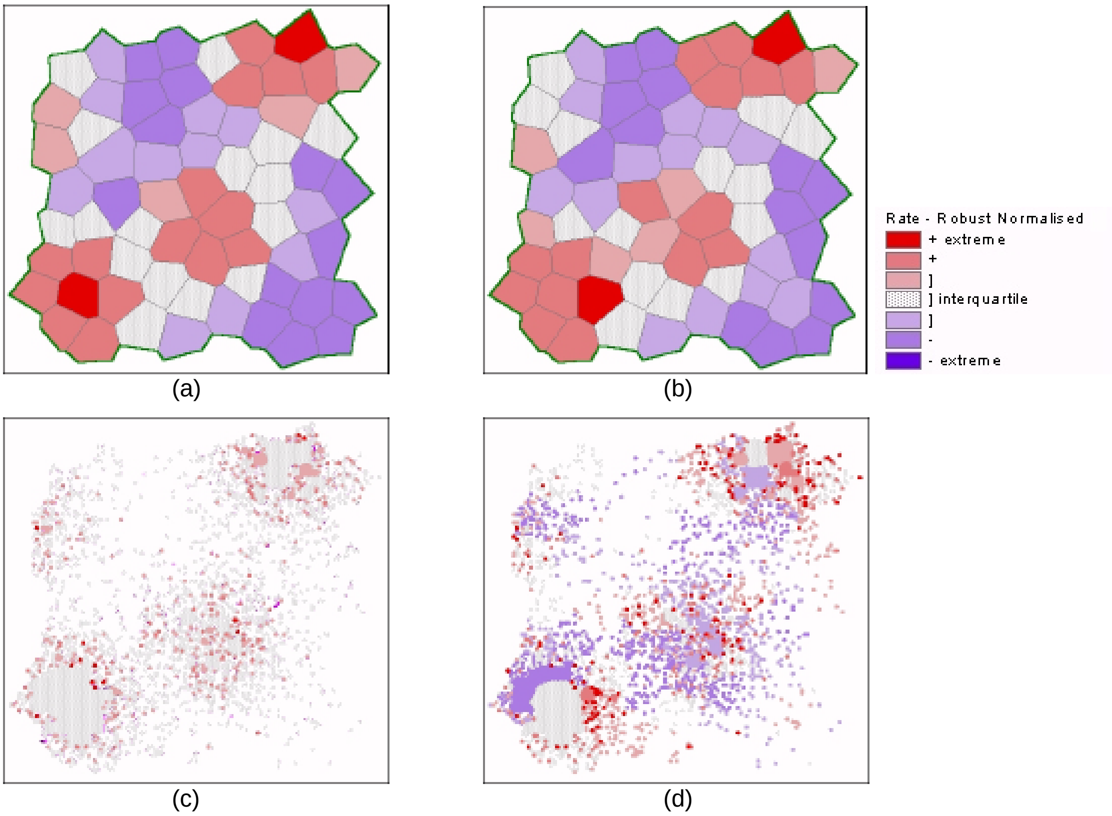
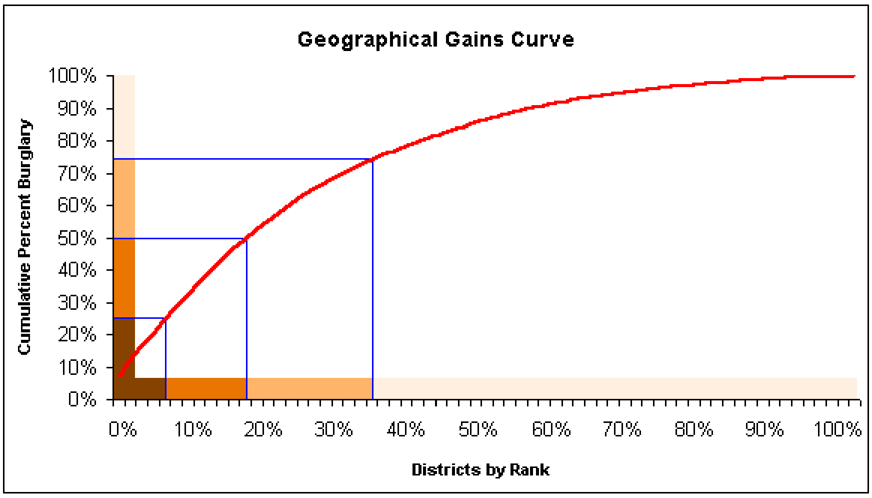


Figure 8: Applying robust normalisation for hot spot detection: (a) burglary counts by district; (b) burglary rates per thousand households; (c) GPZ density of burglaries; (d) GPZ burglary rates.

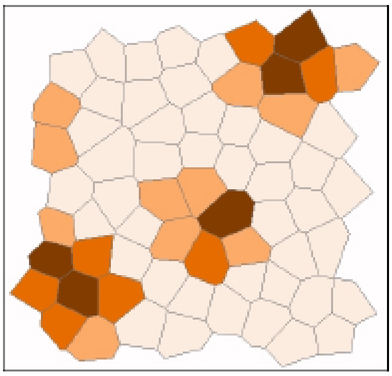
Taking Action: The Geographical Gains Curve

Having now the means to produce robust hot spot choropleth maps, partnerships working on crime reduction initiatives in a joined-up way (inter-agency) must reach considered decisions on targeting and commitment of resources to yield maximum gain. From a crime reduction perspective, initiatives should be targeted at those areas having the highest rates, that is, where citizens are most vulnerable and are at greatest risk of victimisation. As we have seen above, such areas may not geographically correspond with the highest incident counts simply because higher counts may only reflect larger numbers of households in an area and not any greater risk to the individual household owner. It would nevertheless be desirable for crime reduction initiatives to target areas with the highest rates and provide maximum gain in reducing counts (and hence resources in policing). A useful tool here is the geographical gains curve (Figure 9(a)). This plots cumulative percentage of burglary against ranked districts in descending order of burglary counts. This shows that 25% of the burglaries were committed in 8% of the districts and 50% of the burglaries in just under 20% of the districts.

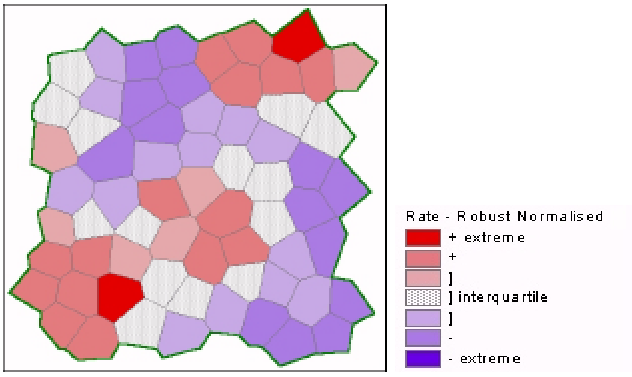
This shows very considerable gains when focusing on the top 8% of districts in order to maximise opportunities of reducing the overall number of incidents. Thus in Figure 9(b) the districts can be colour-coded by gain (in cumulative percent) and compared with the hot spot map for burglary rates. The first choice target tfor a crime reduction initiative is thus the northernmost district which couples extreme rates with highest quartile gain. This can then be called a *focus area* rather than just a hot spot. Similar rational choice can be extended to identifying other focus areas. Compare this outcome with the choice offered in Figure 3.



(a)



(b)



(c)

Figure 9: Applying the geographical gains curve: (a) geographical gains curve for burglary counts; (b) mapped geographical gain (for legend see colour coding in (a)); (c) robust normalised burglary rates (northernmost hot spot corresponds with highest gain in (b)).

Conclusions

This paper has followed a process from analysis of data through to decision-making. Despite the popularity of density estimate interpolation approaches to hot spot detection, many of the problems that detracted from traditional choropleth mapping still persist (numbers of classes, class intervals, application of colour) as well as new problems (smoothing, parameter estimation, boundaries). Geo-ProZones have been put forward here as an alternative approach to hot spot detection. Although GPZ is still being experimented upon, the important issue is that it provides a segmentation rather than an interpolation. Researchers should be focusing on segmentation techniques as they have the dual benefit of focusing on both hot spots of counts and hot spots of rates so that useful comparisons can be made. Robust normalisation, whether applied to interpolations, segmentations or aggregations to administrative units, provides a more objective means of hot spot detection as it allows a standardised approach in choosing the number of classes, the class intervals and the application of colour. Thus more objective comparisons can be made between maps of counts and rates. Finally, when coupled with the geographical gains curve to create a gains map, robust normalised choropleth maps of crime rates provide partnerships tasked with devising joined-up crime reduction initiatives the means to rationally identify focus areas that target high risk areas whilst maximising gains in reducing the numbers of incidents.

Acknowledgements

The research underscoring this paper was part-funded by the Metropolitan Police Service and the Home Office, UK. The author wishes to thank them for access to data, partnership analysts and experienced officers - all having significantly contributed to the author's understanding of policing and crime reduction.

References

Atkinson, P.J. and Unwin, D.J. (2002) Density and local attribute estimation of an infectious disease using MapInfo. *Computers and Geosciences* 28: 1095-1105

Bailey, T.C. and Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman, Harlow.

Brantingham, P. and Brantingham, P. (1984) *Patterns in Crime*. Macmillan, New York.

Brimicombe, A.J. (1999) Small may be beautiful - but is simple sufficient? *Geographical and Environmental Modelling* 3: 9-33

Brimicombe, A.J. (2000) Constructing and evaluating contextual indices using GIS: a case of primary school performance. *Environment & Planning A* 32: 1909-1933

Brimicombe, A.J. (2003a) *GIS, Environmental Modelling and Engineering*. Taylor & Francis, London.

Brimicombe, A.J. (2003b) A variable resolution approach to cluster discovery in spatial data mining. In Kumar *et al.* (eds) *Computational Science and its Applications*. Springer-Verlag, Berlin, vol. 3: 1-11

Brimicombe, A. J. and Tsui H. Y. (2000) A variable resolution, geocomputational approach to the analysis of point patterns. *Hydrological Processes* 14: 2143-2155

Brunsdon, C. (1995) Estimating probability surfaces for geographical point data: an adaptive kernel algorithm. *Computers and Geosciences* 21: 877-894

Garland, D. (2000) The culture of high crime societies. *British Journal of Criminology* 40 346-375

Harries, K. (1999) *Mapping Crime: Principle and Practice*. National Institute of Justice, Washington DC.

Hettmansperger, T.P. and McKean, J.W. (1998) *Robust Nonparametric Statistical Methods*. Arnold, London.

Jefferis, E. (and 16 others) (1998) *A Multi-Method Exploration of Crime Hot Spots*. National Institute of Justice, Washington DC.

Langworthy, R.H. and Jefferis, E.S. (2000) The utility of standard deviation ellipses for evaluating hotspots. In Goldsmith *et al*. (eds) *Analyzing Crime Patterns*, Sage, Thousand Oaks, CA: 87-104

McLafferty, S.; Williamson, D. and McGuire, P.G. (2000) Identifying crime hotspots using kernel smoothing. In Goldsmith *et al*. (eds) *Analyzing Crime Patterns*, Sage, Thousand Oaks, CA: 77-85

Monmonier, M. (1991) *How to Lie with Maps*. University of Chicago Press, Chicago.

Openshaw, S. and Taylor, P.J. (1981) The modifiable areal unit problem. In Wrigley, N. and Bennett, R.J. (eds) *Quantitative Geography*, Routledge, Henley-on-Thames: 60-70

Ratcliffe, J.H. and McCullagh, M.J. (1999) Hotbeds of crime and the search for spatial accuracy. *Journal of Geographical Systems* 1: 385-398

Reiner, R. (2000) Crime and control in Britain. *Sociology* 34 71-94

Rogerson, P. and Sun, Y. (2001) Spatial monitoring of geographical patterns: an application of crime analysis. *Computers, Environment and Urban Systems* 25: 539-556

Sadahiro, Y. (1997) Cluster perception in the distribution of point objects. *Cartographica* 34: 49-61

Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, New York.

Townsley, M.; Homel, R. and Chaseling, J. (2003) Infectious burglaries. A test of the near repeat hypothesis. *British Journal of Criminology* 43: 615-633

Trickett, A.; Osborne, D.K.; Seymour, J. and Pease, K. (1992) What is different about high crime areas? *British Journal of Criminology* 32: 81-90

Tsui, H. Y. and Brimicombe, A. J. (1997a) Adaptive recursive tessellations (ART) for Geographical Information Systems. *International Journal of Geographical Information Science* 11: 247-263

Tsui, H. Y. and Brimicombe, A. J. (1997b) Hierarchical tessellations model and its use in spatial analysis. *Transactions in GIS* 2: 267-279