Chapter 8
Predictive Mapping of Crime by ProMap: Accuracy, Units of Analysis, and the Environmental Backcloth

Shane D. Johnson, Kate J. Bowers, Dan J. Birks, and Ken Pease[1]

Abstract This chapter concerns the forecasting of crime locations using burglary as an example. An overview of research concerned with when *and* where burglaries occur is provided, with an initial focus on patterns of risk at the individual household level. Of central importance is evidence that as well as being geographically concentrated (at a range of geographic scales), burglary clusters in space *and* time more than would be expected if patterns of crime were simply the result of some places being more attractive to offenders than others. One theoretical framework regarding offender spatial decision making is discussed and consideration given to how features of the urban environment which affect the accessibility of places (e.g., road networks or social barriers) might shape patterns of offending. A simple mathematical model informed by the research discussed is then presented and tested as to its accuracy in the prediction of burglary locations. The model is tested against chance expectation and popular methods of crime hot-spotting extant and found to outperform both. Consideration of the importance of different units of analysis is a recurrent theme throughout the chapter, whether this concerns the intended policy purpose of crime forecasts made, the spatial resolution of different types of data analyzed, or the attention given to the dimension of time – a unit of analysis often overlooked in this type of work. The chapter concludes with a discussion of means of developing the approach described, combining it with others, and using it, inter alia, to optimize police patrol routes.

[1]The authors would like to thank Merseyside police and the Ordnance Survey for providing the data analyzed in this chapter, and three anonymous reviewers for their thoughtful comments. This research was supported by a British Academy International Collaborative Network grant, and additional funding from UCL Futures, NSCR and the Research Incentive Fund at Temple University.

S.D. Johnson
UCL Jill Dando Institute of Crime Science, London, United Kingdom
e-mail: shane.johnson@ucl.ac.uk

D. Weisburd et al. (eds.), *Putting Crime in its Place*,
DOI 10.1007/978-0-387-09688-9_8,

Introduction

Predictions about how crime will be distributed can be made for different levels of spatial aggregation, most being useful for some policy purpose. At the macro level, predictions may be made about towns or cities. These will be useful for national governments in the provision of resources and support to local policing. The meso-level employs smaller areal units such as police beats or census tracts which nonetheless may incorporate hundreds of households and businesses. These will be useful for resource decisions within police forces and for the understanding of offender spatial choices dependent on neighborhood characteristics (see e.g., Bernasco and Nieuwbeerta 2005). The micro level has the individual household or business as its unit of count and has been studied in the context of repeat victimization (e.g., see Farrell 2005) and household-specific risk factors (see Winchester and Jackson 1982; Coupe and Blake 2006). Prediction at this level is useful for the provision of advice and protective technology to individual locations at risk.

The writers contend that the unit of analysis selected should be informed by theory and/or an intended application. Data are often unavailable at the level of geographic resolution desirable, particularly where a disaggregated analysis would be appropriate. When analyses are performed using data from the less appropriate larger areal units, the temptation may be to assume that patterns observed across an area will apply equally to the mosaic of smaller areas (and individual locations) of which it is composed - the ecological fallacy. The assumption that crime experience is uniform within an area is certainly ill-founded (Bowers et al. 2005). The larger the areal unit used, the more probable it becomes that local decisions within the area will be sub-optimal. Of course, disaggregated data may be aggregated. Aggregated data cannot readily be disaggregated.

This preamble has two purposes. The first is to assert that the choice of areal unit in crime prediction should proceed in lockstep with the decisions to be informed. The second, more contentious, is to propose that there is an areal unit intermediate between micro and meso level which is the appropriate one for optimizing police patrol. This may (over-simply) be thought of as an area clearly visible within seconds by a police officer on foot, encapsulated as 'having a look around', having stopped a patrol vehicle in an area of potential interest. While vague, in the next section, we will suggest how a unit of analysis at roughly this scale emerges from research as useful in underpinning patrolling choices.

Before continuing, an overview of what follows will be provided. The general aim of this chapter is to present our most recent work and thinking concerning the forecasting of future locations of burglary, where the future relates to the next week, or so. The structure of the chapter is as follows. First, we briefly discuss research concerned with patterns of burglary at the most precise unit of analysis; repeat victimization of the individual household. This is followed by a discussion of more recent work which may be thought of an extension of the repeat victimization literature. In doing so, we discuss one theoretical framework regarding offender spatial decision making and how this might inform crime forecasting. Second, we discuss how other factors, which may be defined at different units of analysis (e.g., the street network, neighborhood characteristics), may influence offender decision making and how these might be integrated into a forecasting methodology. Third, we present an empirical test of the method discussed, comparing it to contending alternatives and chance expectation. Finally, the results are discussed, their implications articulated, and methodological issues regarding the importance of choosing the right spatial units of analysis revisited. The crime analyzed in this chapter is domestic burglary but we are actively conducting the analyses necessary for extension across crime types and across sequences of different crime types.

Risk, Contagion, and the Optimal Forager

The logic applied in the research refines and develops that focusing on repeat victimization (see Pease 1998). Briefly, for every type of crime studied (except homicide), the risk of victimization increases significantly following an initial event (e.g., Polvi et al. 1991); a small proportion of victims (e.g., 4%) account for a large proportion (e.g., 44%) of crime (Pease 1998); and, where repeat victimization occurs, it usually does so quickly, offering a narrow window of opportunity for intervention (e.g., Johnson et al. 1997). The consequence is spatio-temporal instability in crime risk at the micro-level unit of analysis. Importantly, interviews with offenders and analysis of police records suggest that repeat burglaries at the same location are overwhelmingly the work of the same offender or of groups having at least one offender in common (Ericsson 1995; Ashton et al. 1998: Everson and Pease 2001; Everson 2003; Bernasco 2008).

Research also suggests that offenders exhibit preferences, internalized as cognitive scripts (Cornish 1994; Wright and Decker 1994; Rengert and Wasilchick 2000) or templates (Brantingham and Brantingham 1993), for the types of property considered suitable targets. The prevalence of repeat victimization shows that despite the many and varied opportunities within a burglar's awareness space, he or she exhibits an inclination to return to homes already victimized. Interviews with convicted burglars (Ashton et al. 1998; Shaw and Pease 2000) suggest that where opportunities present themselves, offenders seek familiarity, which is to be found in returning to the same home or looking for near-replicas, favoring these over targets of which they know little. Prosaic reasons for return include this terse commentary of one Scottish burglar 'Big house, small van'!

Preferences (conceived as reward in relation to effort) seem to be maintained until they no longer offer an advantage over other opportunities (Brantingham and Brantingham 1981). Experience updates the templates. The discovery of a good opportunity for burglary should elevate the risk to similar households from that burglar. Thus, following a burglary at one home, those located nearby (which will share a range of physical and other characteristics) should be at an elevated risk for some time afterwards. This is termed a 'near repeat' (Morgan 2001; Pease 1998). Much recent research demonstrates this to be the case. Using techniques from epidemiology (Knox 1964), research has confirmed that burglary clusters in space and time (Townsley et al. 2003; Johnson and Bowers 2004a; Bowers and Johnson 2005). Communicable disease provides a useful simile from the policing and victim perspective, where the task at hand is to choose for attention places currently at high risk. When considering matters from the offender's viewpoint, the notion of the optimal forager (discussed below) is more apposite. In any event, we should not get too excited about similes beyond their heuristic value. For example, in disease contagion, each victim comes to carry the infective agent to his or her neighbors. In burglary contagion, the infective agent remains the same (the burglar), and it is the way in which their awareness of, or preferences for, particular opportunities may be temporarily shaped by their recent activity that is of interest.

Returning to the typical conclusions from the research concerned with near repeats, following a burglary, homes up to 400 m away have been shown to experience an elevated risk for up to two months. Importantly, data from five countries (UK, US, AUS, NDL, and NZ), demonstrated that the phenomenon is at least widespread, perhaps ubiquitous, in the developed world (Johnson et al. 2007). Arguing that the same underlying process is involved is premature but tempting. This so-far ubiquitous pattern suggests that, although some areas tend to experience enduring risks, hot spots tend to be 'slippery' (Johnson and Bowers 2004b). Offenders demonstrate by their 'spoor' (the trail of victimized homes or people) a search pattern which may be likened to foraging behavior (Johnson and Bowers 2004a). As a consequence of having exploited all favorable opportunities on one street segment, or because of a perceived elevation in the local risk of apprehension, an offender may move to other areas, typically those conveniently accessible, that is, nearby. Our understanding of the burglar's sequenced decision making is thus as follows. First, select an area/street segment, burgle the best presenting option, and then target the most similar available opportunities, returning to some before moving on when profit diminishes or physical change alerts to precautions being taken. As with foraging, the activity is intermittent.

Why is the image of the optimal forager appealing? Consider the foraging sheep. Sheep may take multiple bites of the most luscious grass, particularly since foraging involves a trade-off between, on the one hand, the energy value of food immediately available and, on the other, effort expended in reaching even more luscious grass elsewhere. Grass in the far corner of the field has to be more calorific to the extent that it offsets the energy (and possibly danger) involved in moving across the pasture. The food value of over-grazed grass diminishes until it re-grows, just as the take from repeatedly burgled homes declines until replacement goods are purchased. Even sheep get full sometimes, and so do the arms or vans of burglars. Burgled goods must be secreted or disposed of before foraging recommences. Once an area is perceived to have been grazed out (skimmed of the best opportunities) the forager moves on. This is in line with theory discussed elsewhere (e.g., Brantingham and Brantingham 1978; Cornish and Clarke 1986; Bernasco and Nieuwbeerta 2005) and is consistent with interviews conducted with offenders (e.g., Bennett and Wright 1984; Nee and Meenaghan 2006). The evident regularity in the space-time clustering of recorded burglary is, in short, a consequence of burglar as forager. It is emphatically *not* consistent with any theory predicated on a time-invariant distribution of risk factors. To put it bluntly, space *and* time are both crucial in the analysis of crime patterns. To ignore time is to diminish the value of spatial analysis and vice-versa. Ignoring time leads to a perceived paradox: that homes already victimized are liable to further victimization, but that crime risks move. The paradox is only apparent when the variable *time* is overlooked. The elevated risk to the recently burgled is transient. As time moves on, so does the foraging burglar.

We presume that crimes in a near repeat series are more often than not committed by the same offender(s). Interviews with offenders support this idea (Ashton et al. 1998), as do the findings of research which has examined patterns of offending for detected offences. For example, Everson (2003) found that having targeted one home on a street, burglars tended to target others nearby. Bernasco (2008) looked at pairs of detected burglaries at a range of spatio-temporal distances. He found that 98% of repeat burglary pairs occurring within 100 m and one week of each other were detected to the same burglar. The proportion of same-offender pairs declined as time and space between events increased. For example, events 100–200 m from an initial burglary 1–2 weeks later were the work of the same offender in only 55% of detected cases. That said, more research is required as analyses so far conducted may reflect only the targeting decisions made by the (potentially biased) samples of offenders for which data were available- those arrested or convicted.

A further test of the near repeat/same offender hypothesis (Bowers et al. 2004) involved the use of a simple mathematical model (hereafter, ProMap) to predict where burglary would next occur. The risk of crime at any location within a grid that represented the study area was derived by considering where and when burglaries had previously occurred. If burglaries clustering in space and time represent the foraging signatures of individual offenders (or co-offenders), grid locations where the greatest number of burglaries had occurred recently and nearby (rather than those locations that had just had the most crime at them, irrespective of when) were predicted to have the greatest imminent risk. Given the finding that such risk diminishes with time, the appropriate temporal interval for analysis of predictive accuracy should be short (e.g., one week into the future). The use of longer intervals (e.g., one month into the future) would be incongruent with the tempo at which offender and policing decisions are made. The areas for which the predictions were generated were also small (50 m grid cells) for the same reasons.

ProMap: Initial Tests and Introducing Accessibility

The results of the ProMap model were compared to the retrospective hot-spotting technique known as Kernel Density Estimation (KDE) and also to a thematic map generated using a police beat geography (Bowers et al. 2004). The KDE method was chosen as a comparator for two reasons. First, it is commonly used by police and researchers. Second, unlike thematic maps, the unit of analysis is flexible. 50 m grid cells were used for both the KDE and ProMap. The obvious justification for the use of thematic maps is that they correspond to areas of police responsibility (Poot et al. 2005) and that such analysis facilitates a comparison of forecast accuracy for maps derived using very different units of analysis.

To examine the relative accuracy of the approaches, we first compared the percentage of future burglaries that occurred in the 20% of cells (or beats when using thematic maps) with the highest predicted risks according to each method. ProMap proved more accurate than either KDE or thematic maps. The fact that only one prediction was generated and no comparison was made against chance expectation limits euphoria about ProMap's performance. Both of these considerations are addressed below.

A criticism of most crime mapping techniques is that crime data constitutes the sole input, (although this is not true of repeat victimization research (see, e.g., Tseloni et al. 2004; Tseloni and Kershaw 2005). The underlying opportunity structure or spatial field is essentially assumed to be uniform. This assumption is seldom tenable. Recent research concerned with spatial ecology (Matthiopoulos 2003) considered the impact on the foraging behavior of species under temporal and spatial constraints on the distribution of resources and the accessibility of locations. Where resources are unevenly distributed (the norm), the inclusion of an accessibility variable increases the predictive efficacy of mathematical models of foraging. People are likewise subject to both spatial and temporal constraints (Ratcliffe 2006), and location accessibility is hence a plausible influence on offender foraging.

In the urban environment, Beavon et al. (1994) examined how the risk of victimization at different street segments in Vancouver varied according to their ease of access. There was indeed a positive correlation between victimization and ease of access for a range of property crimes including burglary. Similar results have been found for research conducted in the UK (e.g., Armitage 2007; Hillier 2004).

The only study which considered such issues in the prediction of crime and of which the authors are aware was conducted by Groff and LaVigne (2001). They used a simple 'on-off' estimation of opportunity factors such as land use, housing tenure, and proximity to likely offenders, derived for a grid that represented the study area. These were aggregated to produce an overall risk score for every grid square (or cell). For each cell, the risk score incorporated the mean of the cells that comprised its Moore neighborhood (the surrounding cells). The accuracy of this weighted opportunity surface in predicting burglary was then tested. The results suggested that only 6% of burglaries occurred where the risk score was one to two standard deviations below the mean. At the other end of the scale, 20% of burglaries happened in cells with risk values one or more standard deviations above the mean. The model thus seemed better at predicting where crimes would not occur than where they would. Groff and LaVigne incorporated only time-invariant factors in their predictions. The model thus accurately identified areas where good opportunities for crime were always rare. Where opportunities did exist, the Groff and LaVigne procedure, neglecting variations over time, failed to help much in indicating which locations would be exploited.

How should one operationalize accessibility (or what might be thought of as pull factors)? The simplest and most obvious factor is the *number* of homes in a grid square and this is the first accessibility measure used in the study reported. The second concerns the street network and in the simplest sense the number and types of road in each grid cell. The rationale for using the latter was that the number of roads in a cell is likely to provide a crude index of how connected each cells is to those that surround it, whereas the type of road (small or large) provides an estimate of the likely volume of through traffic (pedestrian or vehicle).

A further factor that might affect offender movement is the existence of physical or perceived barriers. For example, using data for convicted offenders in The Hague, Bernasco and Nieuwbeerta (2005) examined the effect of area characteristics on offenders' target choices. After controlling for other relevant factors (such as proximity to the central business district), they found that an offender's decision to burgle was linked to ethnic heterogeneity, measured at the area level. More homogeneous neighborhoods appeared to generate impedance to offender movement, and this was particularly the case for areas hosting those not native to the Netherlands. In a further Dutch study, Poot et al. (2005) found that when offenders committed any place based crime (e.g., shoplifting, burglary, or violence) in neighborhoods other than their own, they were less likely to cross recognizable social barriers. In the same vein, LeBeau and Rengert (2006) showed that arrested drug dealers tend to deal drugs in areas with ethnic profiles similar to that of their own neighborhood. Reynald et al. (2006) also show the salience of social barriers in crime trips. In short, the evidence suggests that social barriers delineating areas shape offender choices. Observed consistencies in the target choices of offenders reflect their preferences (see Hakim et al. 2001).

Put simply, burglars have preferences for certain areas. We here hypothesize that where barriers may be perceived to separate areas of recent crime from different but contiguous areas, the established patterns whereby (offender foraging and hence) crime risk seems to spread will be impeded.

The Argument So Far and Measuring Predictive Success

Let us rehearse the arguments advanced to this point. We have argued

1. that a scale of spatial analysis between meso and micro is appropriate for patrolling choices.
2. that the range over which burglary risk might be thought of as being communicated (to use the simile) is consistent with this scale, being some 200–400 m
3. that the spatial communication of risk is temporally limited, making spatio-temporal analysis necessary.
4. that such patterns are overwhelmingly a reflection of the activity of individual offenders, with the analogy of the optimal forager being helpful in illuminating the likely search processes involved.
5. that factors of actual and perceived accessibility of places can usefully supplement the basic ProMap approach.

A further issue that is central to the theme of this book concerns the size of the areas to be identified as being most at risk. Police resources vary. A poorly resourced police area will be attracted by a predictive mapping model which identifies the most crime-prone 10% (say) of locations. The police commander in such an area will be indifferent to the model's performance for a higher proportion of the area to be policed, because he or she does not have the resources to cover more than 10% of the area. A more richly resourced area will be concerned about the model's performance up to (say) 20% of the area, if resources exist to police such a high proportion of the area. For this reason, a more sensitive means of establishing predictive accuracy than has been used elsewhere (e.g., Bowers et al. 2004; Groff and LaVigne 2001) is required. Instead of considering the fraction of events predicted by a particular proportion (such as the top 20%) of cells with the highest anticipated risks, the product which satisfies the information needs of police across a range of resourcing options is an accuracy concentration curve. This is simple to do, being generated by plotting the percentage of burglaries accurately predicted as a function of the incremental (risk ordered) percentage of cells considered. This provides a more complete understanding of how well the different models work. Crucially, it allows operational decisions to be optimized for *any* level of local resourcing and circumstance, and to be adjusted in response to changes in these factors. Consider the two hypothetical examples shown in Fig. 8.1. For the first four percent of cells considered, model 2 performs better than model 1, but thereafter the reverse is true. The functional form of model 1 is non-linear, which is desirable, whereas as that

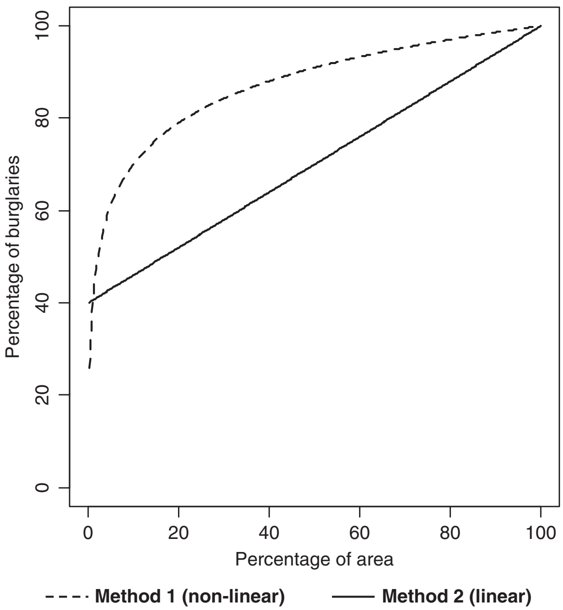

Fig. 8.1 An example accuracy concentration curve

for model 2 is linear, which is not. Patterns such as this are only detectable where a concentration curve is generated and will be missed where single cut points are used.

The aims of the analyses that follow were three-fold: (1) to provide a further test of the prospective mapping (ProMap) algorithm, by generating a series of predictions for comparison against chance expectation and the performance of retrospective methods generated using different (spatial) units of analysis (KDE and thematic mapping); (2) to examine the effect of using data concerning the distribution of opportunities and a crude index of their accessibility on the accuracy of the ProMap algorithm; and, (3) to develop the model to incorporate the potential impact of barriers. The last aim was particularly tentative as no relevant research had been conducted within the study area used in this chapter, or using the approach here proposed. It has yet to be established whether the existence of barriers of the kind discussed above affect offender targeting decisions within the UK, or what types of barrier might be chosen as most relevant. The attempt is worthwhile, if only as a marker for future research.

Data and Method

Residential burglary data were obtained for the county of Merseyside, UK. Reflecting processing limitations, data were analyzed for a (5 km $\times$ 5 km) 27,040,000 m^2 grid square in South Liverpool, a fragment of which is illustrated in Fig. 8.2. Each burglary record included information concerning date and location, the grid coordinates having a positional accuracy of one meter. The available data covered the period 1 September 1996 to 30 November 1997. Data for the first year (1 September 1996 to 31 August 1997) were used as a construction sample to generate a profile of space-time clustering, the remainder of the data were used (1 September 1997 to 30 November 1997) as a validation sample for the predictions generated[2]. Data covering a period of two months were used to generate all predictions, irrespective of the method used. To test the accuracy of predictions, data from the ensuing seven days were examined. In order to minimize spatial edge effects, data for a 500 m-buffer zone were included in the generation of the predictions.

Analysis of space-time clustering using historic data allowed calibration of the prospective mapping algorithm. Because this approach has been described elsewhere (see, Johnson et al. 2007), it will be merely outlined here. In the Knox approach (Knox 1964) and the Monte Carlo variant used here (Besag and Diggle 1977), the spatial and temporal interval between each crime and every other crime was computed. A contingency table was then populated to summarize the

[2] The validation sample was not used in the generation of the area-level profile of space-time clustering, but was of course used in the generation of the predictions. That is, to generate the predictions a rolling window of data were required. For the initial forecast only data from the construction sample were required, but for subsequent predictions additional data (which preceded the prediction) from the validation sample were necessary.

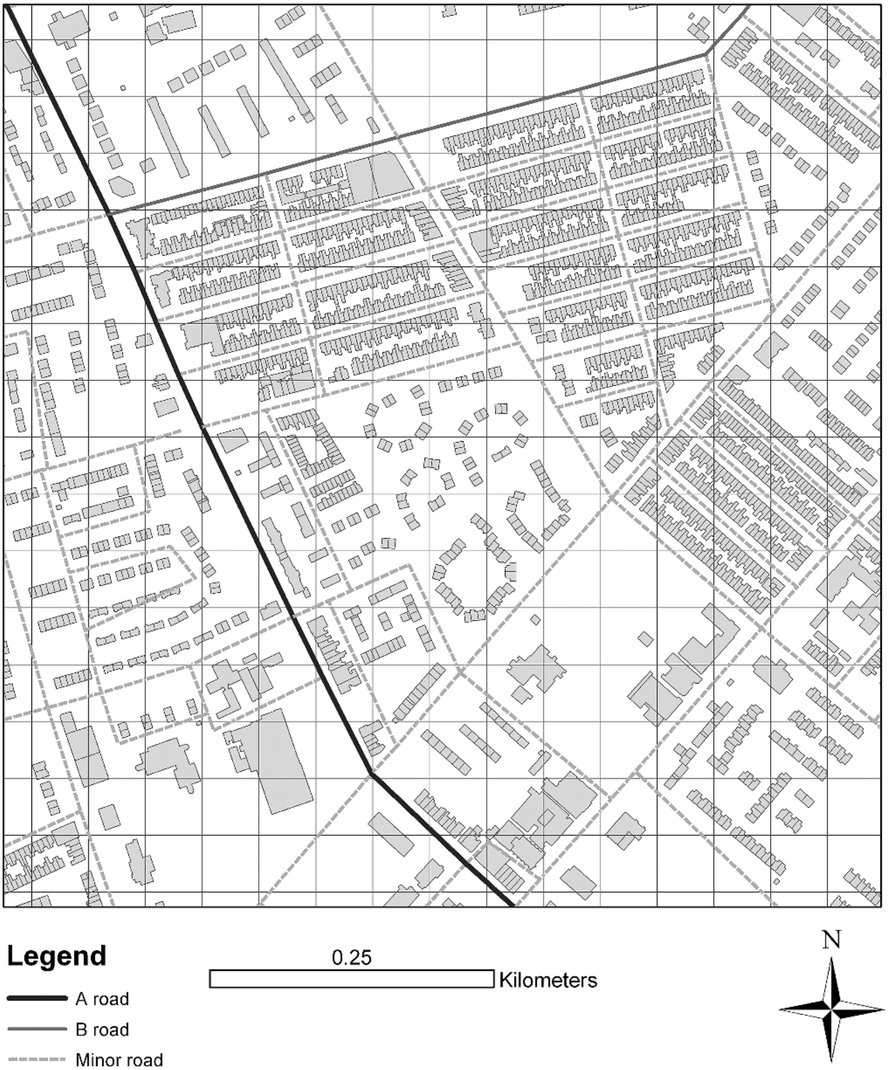

Fig. 8.2 An example of the study area grid and opportunity structure (Ordnance Survey © Crown Copyright. All Rights reserved)

resulting n*(n-1)/2 comparisons. To generate the frequencies expected (e.g., for burglaries that occurred within 100 m and 7 days of each other) under conditions where the timing and placement of crimes were independent (the null hypothesis), a Monte Carlo re-sampling approach (Besag and Diggle 1977) was used. This was completed 999 times to enable inference testing[3]. Twelve months' data from the

[3] The basic Monte Carlo re-sampling approach is discussed in a little detail in a later section. For the sake of brevity, readers are asked to accept that the approach was valid.

study area (N=1,662) as described above were used as the construction sample. The results showed that more burglaries occurred within 500 m and two months of another than would be expected on the basis of chance. If, as argued here, this pattern is the consequence of offender foraging behavior, the transient elevated risk for locations within close proximity of burgled homes reflects the behavior of the same offender.

As discussed, two factors contributing to the placement of burglary are the physical attributes of housing and the road infrastructure. The former defines the spatial distribution of opportunities for burglary, the latter a way of accessing homes and traveling between them. Two simple opportunity surfaces were derived to represent the two factors. Using a Geographical Information System (GIS), the study grid was divided into 50 m cells, 10,816 in all. Information on housing was generated using Ordnance Survey (OS) Land Line data. These data afforded a considerable advantage over census data, available only at the area level and consequently inviting the ecological fallacy. Processing OS data involved the conversion of building outlines to solid shapes using a GIS. The building shapes were then intersected with the 50 m grid squares which provided a target distribution 'layer'. This is illustrated in Fig. 8.2, which shows the relationship between the building outlines and the grid squares. It should be noted that some buildings were located on the boundary between two or more grid squares. Where this occurred, the building was allocated to the cell containing the midpoint of the building. It was then possible to quantify the number of buildings wholly or partially located within each grid square.

Information on road infrastructure was derived from OS Meridian data. This related to all roads in the area and provided details of the classification of the road (e.g., was it a major or minor road or a motorway?) as well as their location and length. This information was intersected with the grid square coverage and is also illustrated in Fig. 8.2. The width of the lines is proportional to the classification of the road (wide lines are wide roads). The figure shows that some grids contain no roads, others several. The following variables could then be quantified for each grid square:

- The number of road sections that were located in each grid square
- The length of each road encapsulated by each grid square
- The road classification. In the UK urban roads are categorized in terms of the volume of traffic use they are designed for, designated 'A', 'B', and 'minor', defining largest through smallest roads.

How these data inform predictions is described below. However, before proceeding, it is worth addressing a potential concern that might arise for readers most familiar with the Manhattan grid street configuration. Such readers might wonder about the orientation of the grid and why we did not align this with the street network. The simple answer is that in the UK, for most cities the street network is irregular and so in truth there was no (obvious) optimum orientation that could have been applied[4].

Crime Risk Surfaces

Thematic maps were generated in the conventional way. For a basic (concentration) map, the total number of burglaries occurring on each beat for the two months prior to each prediction was computed. To enable a direct comparison with other approaches (see below), some of the police beats were cropped so that only the area (of each beat) that was encapsulated by the 5 km grid was included. In addition to a basic police beat thematic map, others showed the historic concentration of burglary per meter squared and per 1000 households. These metrics standardize the risk experienced across beats using different denominators, with the rate per 1000 households being commonly used by both academic and operational analysts.

The ProMap and KDE event driven risk surfaces were derived as follows. Briefly, a moving window (see Bailey and Gatrell 1995) algorithm is used to generate a risk intensity value for every cell in the grid to reflect the risk experienced or anticipated for that location. For each cell, all crimes within a particular radius (bandwidth) from the midpoint of that cell are identified and a risk intensity value computed based on (for retrospective mapping) the number of crimes within the bandwidth and their spatial proximity to the center of the cell (crimes closest to the center of the cell are allocated a greater weighting). This produces a nicely smoothed map for display. A more detailed description of the approach is provided in Appendix 1.

The ProMap algorithm represents a refinement to the traditional KDE approach for three reasons. First, it provides a theoretical rationale for why risks should decay in space (distance decay is incorporated in KDE but not for theoretical reasons of which the authors are aware). Second, a temporal bandwidth is specified and an associated parameter used to model the effect of elapsed time on burglary risk. Third, the spatial bandwidth is selected for theoretical reasons and is empirically derived using the space-time clustering profile of the area concerned. Appendix 1 provides information on the equation used.

Event Driven Opportunity Surfaces

The opportunity surfaces described earlier were used to produce a variety of weighted ProMap risk surfaces, which were subsequently tested for predictive accuracy. Different layers (ProMap and the two opportunity surfaces) were generated independently and the values for each cell derived by taking the product of the relevant layers, stationary and event driven.

---

[4] Although not attempted here, it would of course be possible to repeat the analyses reported here a number of times using a different orientation of the grid each time.

The casual reader may wonder why only ProMap risk surfaces were subjected to the layering approach. The superficial answer is that the ProMap approach turned out to be the most predictive, and hence was used in the latter part of the research. The more satisfactory response is that, because the other approaches were time-invariant, other time-invariant factors like road structure (which are hypothesized to influence offender foraging) should already have their effects reflected in historic crime data. By contrast, for ProMap which particularly emphasizes short-run changes, time-stable factors like road structure provided a sort of scaffolding around which short-run changes could be modeled.

The values for each stationary surface could be generated in a number of ways, but in the current study the following simple rules were applied. A more detailed discussion of the rules used is provided in Appendix 1.

Road backcloth - this surface was produced by weighting grid cells on the basis of the presence of roads alone, their length and type.

Building backcloth – this surface was produced by weighting cells on the basis of the number of buildings hosted.

Combined backcloth – this combined the data from the road and building surfaces.

Barriers

The first consideration for this factor concerned the identification of the boundaries that would be used to model barrier effects. A number of approaches are possible, including the identification by local offenders or police officers of likely barriers (see Poot et al. 2005). Another uses existing administrative boundaries with barriers being identified using socioeconomic data. This is the approach used here, not least because a finding of the Poot et al. (2005) study was that these aligned with practitioner perceptions. UK census Output Area geography was used as proxy for barrier location, since socioeconomic information is thus readily available. For this initial analysis a geodemographic classification system was derived using census data. The super-profiles system (Brown 1991) is one of a number of target market classification systems and was selected because of its availability to the authors. Further research will explore other market segmentation systems and alternative approaches.

Essentially, the super-profiles classes (ranging from 1 to 10) provide an indication of area affluence. Low scores reflect affluence, high scores deprivation. Figure 8.3 shows the study area with the spatial distribution of super profile lifestyle areas. Where two areas vary in terms of affluence (measured at the area level), we presume social barriers. The greater the area affluence contrast, the higher the social barrier.

The imperfections of this approach are acknowledged, and in consequence the method used was simple. Following the rationale in the introduction, this was: when deriving estimates of risk intensity, rather than weighting each crime event equally, those occurring within the same type of area as the cell for which a prediction was being made were given a higher weighting than those occurring in a different type

Fig. 8.3 Census geography and Super-profile classification for the study area

of area. To do this, each crime had to be assigned an area-level characteristic using the spatial-join command in the GIS. Each cell also had to be assigned a value using the ‘intersect’ command. A discussion of the equation used to model the effects of barriers can be found in Appendix 1.

Measuring Predictive Accuracy

We now return to the basic question of the relative predictive accuracy of ProMap and alternatives. The approach used to assess this was, as noted earlier, to generate the maps using two months of historic data and to determine how many of the burglaries occurring over the ensuing seven days (on average there were 70 per week, SD=7) were correctly predicted as to location. Seven days was considered sensible as each prediction would include both weekdays and a weekend, with their different routine activities (see Rengert and Wasilchick 2000). Predictions for weekdays and weekends should arguably be considered separately in future research. Nine predictions were generated using each method. Predictions were generated for the weeks from 1st, 14th, and 24th of September, October, and November 1997. Thus, test periods did not overlap. The number of predictions generated was limited to nine for two reasons. First, it was necessary to build a space-time clustering profile (discussed above) for the area to calibrate the ProMap algorithm, and with limited resources we were reluctant to do this more than once. As it is likely that space-time clustering profiles change over time, (a topic of no little interest in itself) this was deemed prudent. The second reason was simply that the work took a long time.

Statistical inference is central to theory testing. The question of interest is whether the results observed could have occurred on the basis of chance. Various types of test and method have been derived for different types of data and research questions, but the authors are unaware of any that have been used to establish the accuracy of crime forecasting techniques in anything other than a descriptive way. In the current study we use a simple Monte Carlo (MC) simulation approach (for a general discussion, see North et al. 2002). For the current endeavor, chance performance would characterize a strategy whereby cells were assigned a risk level randomly. One approach would be randomly to select 10,816 balls, one to represent each cell of the grid, from a very large bag, recording their order of selection. The cells designated as most risky would be those selected first, those least at risk last. This would generate one random selection against which the performance of the different predictive approaches could be compared. However, making one selection is not reliable and hence a series of selections was deemed necessary. If we assume that the prediction under evaluation had been drawn from a larger population, then (say) 99 random sequences or samples would enable a reasonable distribution to be produced for the purposes of inferential testing[5]. If the volume of crime predicted

[5] For a one-tailed test, a minimum of 19 simulations are required if the 0.05 level of significance is used, but the standard error of the estimate is inversely related to N and so a larger sample is desirable.

by the model of interest were the most extreme then the null hypothesis maybe rejected. For consistency with convention, a particular model would be considered to deviate significantly from chance if the predictions made exceeded at least 95% of the (pseudo) random sequences.

Selecting balls in this way would take a long time and result in the de-motivation of research staff. Fortunately, this MC re-sampling technique can be easily implemented by those with a basic knowledge of computer programming. The method uses a recursive algorithm shown as Appendix 2.

One criticism is that not all cells within the grid contain houses (around 29% do not). Domestic burglars would not select locations without dwellings! Consequently, the MC approach was modified so that cells with houses were selected first, and only then were cells that did not contain housing considered. This removes potential bias in favor of judging ProMap better than chance.

ProMap Accuracy and Backcloth Influences

Figure 8.4 shows the accuracy concentration curves for the first prediction for four of the models (others are not shown due to limitations of space). Each graph shows what would be expected on the basis of chance, estimated using the MC simulation described above, alongside the particular model under consideration. This enables comparisons to be made with chance expectation and, by providing consistent reference lines, facilitates comparisons across models. Considering chance expectation first, as would be anticipated there is a linear relation between the number of cells searched and the concentration of burglary identified. The slope for the average of the simulations is greater than one, indicating the effect of first sampling cells containing housing. On average around 15% of burglaries are correctly identified by searching 10% of cells. The dotted line shows the area of the graph within which 95% of simulations fall. The maximum proportion of burglaries identified by searching 10% of the cells for 95% of the simulations was 22%. If the value for a particular model exceeds the 95th percentile of the simulations, we conclude that the model was significantly more predictive than chance (for a one-tailed test, and $p < 0.05$).

For this prediction, the ProMap model has a logarithmic functional form which deviates from chance expectation across most of the distribution. This pattern reflects known facts about near repeats. The pattern for the retrospective KDE method is quite different. The KDE method appears to exceed chance expectation until around 50% of burglaries have been identified, but the shape of the concentration curve is less impressive. For example, after around 67% of burglaries have been identified, the KDE model performs below chance expectation. The thematic mapping method (computed using the rate of burglary per 1000 households) performs the worst, failing to exceed chance across the entire distribution.

The ProMap models that include opportunity layers consistently exceed chance expectation, and the area of the graph between model performance and chance expectation is larger than for the basic ProMap model. Including a parameter to model the existence of barriers (graph not shown) has little effect on model accuracy.

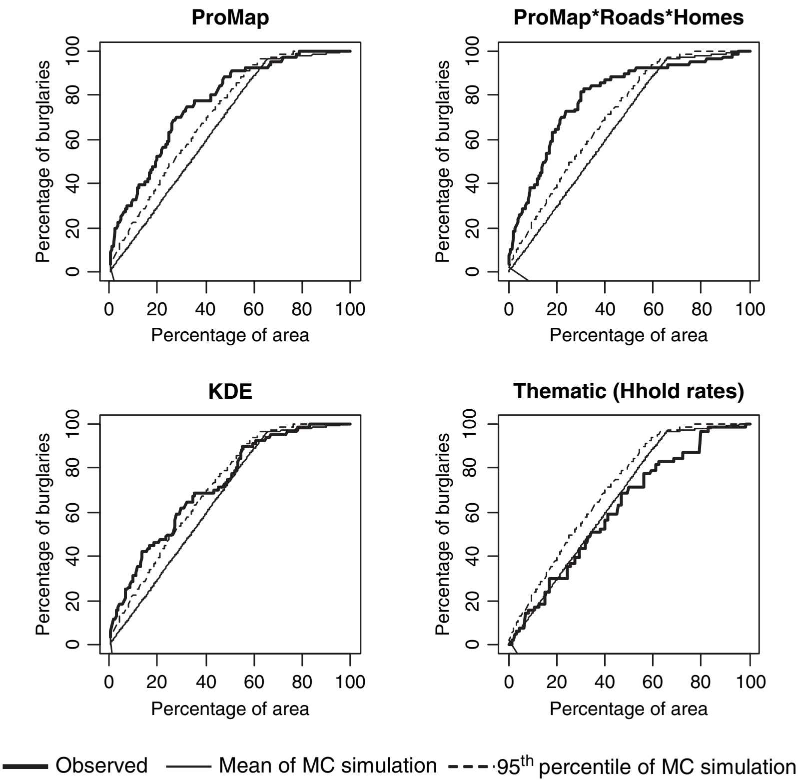

Fig. 8.4 Accuracy concentration curves for a sample of methods tested and chance expectation (for prediction 1)

The above results relate to only one prediction. Rather than reproducing each of the graphs for every prediction, a summary table was generated to provide an overview of the findings. Table 8.1 shows the median accuracy of each model across the nine predictions. It tabulates the fraction of cells that would have to be searched to identify x% of burglaries for five different cut points. Other thresholds could be used, but covering a range of values is instructive since they provide an accurate reflection of visual inspection of every finding[6].

For each cut point, the cell value for the best performing algorithm is highlighted in bold. Where ties occur, multiple cells are highlighted. Also shown in each cell as a subscript is the number of predictions for which that model exceeded chance expectation as defined by the value for the 95th percentile of the simulation results.

[6] An alternative approach to summarizing the data that we are yet to do is to compute a gini coefficient for each concentration curve. Frequently used to examine (for example) the inequality in the distribution of wealth across areas, the measure indicates the area between an observed (Lorenz) curve and the 45 degree line of equality. The advantage of so doing is that the coefficient summarizes the entire distribution rather than a subset of cut points.

Table 8.1 Median mapping algorithm accuracy for a subset of the results

	Percentage of burglaries identified					
	10	25	50	75	90	Percentage of cells searched
Prospective:						
ProMap	1.3_9	5.0_9	14.3_9	30.8_9	55.3_6	
ProMap*Houses	1.5_9	5.0_9	14.3_9	28.3_9	48.8_8	
ProMap*RDs	1.3_9	4.8_9	13.3_9	29.0_9	52.5_7	
ProMap*Houses*RDs	1.5_9	4.5_9	12.5_9	29.3_9	56.3_5	
ProMap*barriers	1.5_9	5.0_9	14.8_9	29.5_9	54.5_5	
Chance:						
Simulation 95th Percentile	3.8	11.5	27.3	44.8	56.8	
Simulation Mean	7.0	17.0	34.3	51.3	61.3	
Retrospective:						
KDE	2.0_9	6.5_9	16.8_9	34.8_7	59.0_4	
Thematic (concentration)	4.0_3	15.5_0	35.4_0	49.1_2	63.0_2	
Thematic (rate per area)	3.3_4	10.8_5	23.4_7	42.5_5	58.8_2	
Thematic (rate per 1000 Hholds)	6.4_1	17.6_2	31.7_0	50.0_2	69.1_1	

Note: subscripts indicate for how many of the nine predictions the model exceeded chance expectation

To take an extreme example, in predicting the locations of 90% of the burglaries, the retrospective KDE method exceeds chance on only four occasions, being worse than chance for five. For the thematic maps, two of the models (basic concentration and rate per 1000 households) failed to exceed chance for almost every trial and each cut point. It should be noted that in generating the graphs for the thematic maps, it was not possible to compute the percentage of area needed to identify a particular fraction of burglaries. This was because, unlike the regular cells used for the KDE and prospective maps, the units of analysis (police beats) for the thematic maps varied in size. Thus, when estimating how much area would need to be patrolled to identify a particular proportion of crime, one can either select the configuration of beats that identify slightly more than the threshold of interest or slightly less. The latter would overestimate the effectiveness of the thematic maps and is the approach used here. Thus the thematic approach appears to be egregious even when it is favored in the conventions applied.

We have thus established that the prospective models typically exceed chance expectation. A complementary perspective on model accuracy involves the comparison of the different models against each other. A variety of comparisons is possible. We report here comparisons between the prospective model most accurate for forecasts made over the smallest areas (that which included data on the homes and roads in each cell) and the KDE method. The results indicated that the ProMap model considered consistently performed better than the KDE method in identifying 10% ($z=1.70$, $p<0.05$, one-tailed)[7], 25% ($z=1.96$, $p<0.03$, one-tailed), 50% ($z=2.56$, $p<0.01$, one tailed), and 75% ($z=1.90$, $p<0.03$, one-tailed) of burglaries. In identi-

[7] A one-tailed test is here used as the direction of the hypothesis was specified a-priori.

fying 90% of burglaries, the difference was non-significant (z=0.29, p=ns). Thus, with one exception, the ProMap algorithm that incorporated data on the location and spatial concentration of homes and roads exceeded chance performance and consistently outperformed the KDE method. The inclusion in the ProMap algorithm of data concerning the concentration of homes *and* the road network seemed only to offer meager additional value when forecasts were generated for the largest areas (e.g., when forecasting 90% of burglary locations). Expressed in another way, the improvements in predictive accuracy offered by the ProMap model discussed were most evident for geographical areas that could realistically be targeted for police attention (i.e., smaller areas).

It is important to paint a picture of how to interpret differences in the values shown in the cells of the table. For example, what does it mean to say that to identify the locations of 50% of burglaries, using the ProMap*housing*RDs algorithm 12.5% of cells would need to be searched whereas for the KDE algorithm the equivalent fraction of cells would be 16.8? A potentially useful guide to the interpretation of these figures is that a one-percent difference in the search area required to identify a particular fraction of burglaries equates to a patrol area of 270,400 m^2 (or 0.27 km^2). If the purpose of a crime reduction intervention informed by predictive mapping is the detection of offences in progress, then even a one-percent difference in accuracy so measured will have substantial implications. The importance of a four or five percent difference needs no advocacy as to operational importance.

What Is Established? What Next?

The incorporation of short-run changes in risk via the ProMap approach to predicting domestic burglary does seem to yield a dividend in predictive performance across the spectrum of risk levels. The distances across which risks change chime well with the practice of patrolling and it is in the optimization of patrolling routes that ProMap has its most obvious potential. The various backcloth information contributed modestly to the accuracy of ProMap and with enough promise to motivate the writers to develop alternative ways of characterizing areal influence on foraging behavior. Further research will use parameters specifically calibrated to maximize prediction accuracy. Similar approaches to least squares estimation will be used, and where enough data are available, Monte Carlo Markov Chain approaches to parameter fitting may also be explored. We also intend to examine accessibility, measured using the street network, in different ways. Accessibility at a higher level of spatial aggregation than analyzed here may be considered by examining (for example) the number of roads that connect each cell to members of its Moore neighborhood. This would provide an indication of how accessible the local area is, as well as the number of roads contained within a particular grid cell. Such analysis may also be conducted at other levels of spatial aggregation.

To model the space-time variation in risk more precisely, a slightly different mathematical approach is desirable. Here, each cell (and eventually each home) would be considered a vertex or node in a graph. Each vertex would be connected by roads represented by edges in the graph. An adjacency matrix would be used to summarize how each vertex was connected to every other vertex, and weights applied to indicate the distance between each node. For nodes that are not connected by a single edge (a road), a shortest path analysis algorithm would be required (e.g., Dijkstra's 1959) to make the necessary calculations. In addition to distance, weights may also be used to indicate other factors that might encourage or impede the flow of risk through the graph network. For example, instead of assuming that the change in risk would be isotropic (uniform in all directions), the role, that the configuration of the street network plays upon people's ability to navigate their environment (and any directional biases that arise), could be modeled. Risk would be hypothesized to flow with a higher probability between nodes located near to each other, between those connected by multiple edges and along arcs which generate the least friction.

Unlike the other factors considered, the inclusion in the model of data concerned with social barriers had no observable impact on predictive accuracy. There are at least three explanations for this, some of which concern construct validity while others are more theoretical. First, the size of the effect that barriers have on offenders' target choices may be truly minimal. Second, the geography used to define the barriers considered may have been inappropriate and failed to reflect the kinds of barrier that offenders take into consideration. Given that the census geography was used, this is entirely plausible, but it should be noted that the same kinds of geographic boundary were used in the studies reviewed in the introduction for which barrier effects were revealed. Further research should use alternative geographies and, in line with the approach used by Poot et al. (2005) draw upon local knowledge when determining which boundaries to use and how to classify them.

Third, the variable used to differentiate between areas and hence to identify barriers may have been sub optimal. We here used a sociodemographic classification system developed for target marketing. Other methods of classification, such as the use of univariate or multivariate analyses of census data are possible and could be usefully explored in future research. Measures of social cohesion may be a useful next step (see Bernasco and Luykx 2003). Considering the current findings, our intuition is that the latter explanations are most likely. Whatever the answer, the current approach offers one empirical approach, easily programmed, for testing theories of this kind. A further caveat worthy of discussion, of course, is that in this study predictions were generated using the spatio-temporal distribution of historic crime events. Whilst these reflect the emergent patterns of the activity of offenders, they do not provide a direct test of the behavior of any distinct offender(s). It may (or may not) be the case that on some or all occasions there is too much 'noise' in the data to model this by proxy.

Areal Units and ProMap: Closing Thoughts

The writers believe that there is much scope for further work. Patrolling patterns are not dictated by one offence type, and establishing regularities in space-time patterns of other offence types, and integrating ProMap across offence types represents the research strategy. Research tactics should obviously incorporate salient other data as was attempted for domestic burglary in the later analyses reported above and discussed below. The point of central importance however is that the ProMap algorithm and (particularly) variants of it significantly outperformed methods extant. They did so in a way that is important not just for theory testing but also for operational policing. All variants of the ProMap algorithm produced maps that could identify the same volume of crime as other methods within much smaller areas.

With respect to units of analysis, the general theme of this book, unsurprising but important results emerged. As expected, the Thematic mapping approach was worst at predicting what happens next, but what is perhaps surprising is that the allocation of resources based on such maps may be no better than chance. This was particularly true where prioritization was informed by area-level crime rates expressed as the risk per 1000 households. It is worth briefly discussing why this might be so. The definition of the boundaries used is critical. For the task at hand, the ideal situation would be to have boundaries that circumscribe areas within which risks *tend* towards homogeneity. Unfortunately, the boundaries used by the police and others are dictated by diverse factors. Some of these are likely to be administrative and not pertinent for the anticipation of crime patterns. The units are also typically large, meaning that within them some degree of heterogeneity is inevitable. And, the boundaries (e.g., census collection areas) are often defined in advance of the data collected, rather than being derived on the basis of emergent phenomena. Rather than fully articulating these issues, which are covered in the chapter in this volume by Brantingham and Brantingham, one approach to boundary determination should be briefly discussed. This is outlined here because it may influence how the current authors attempt to delineate barriers in future work. The approach is inspired, in part, by 'lossy' image compression techniques from the field of software engineering, which attempt to reduce image file sizes by generalizing about similar elements of an image. Potential problems with any such approach are acknowledged. The idea is rehearsed here so that others might point out any weaknesses in the approach or any suitable refinements which overcome them.

To illustrate the approach, the first panel of Fig. 8.5 shows an area with three distinct geographies over which a relatively small 'meso-level' grid has been applied (one cell might encompass 10–20 properties as previously described). All cells within the area are ascribed one or a set of values relating to a multivariate description of their internal characteristics relevant to the investigator (see panel 2). For example, the values considered may include the type of housing in each cell, concentration of crime and so on. Each cell is then examined with respect to its neighbors and evaluated for similarity. Contiguous cells considered 'similar' (i.e., with values within some bandwidth of one another) are then grouped into single areas/geographies/unit of aggregation.

The degree to which cells are considered 'similar' is dictated by the level of 'compression' applied; the higher the level of compression, the less similar things need to be for them to be considered as such, as illustrated in panels 3 and 4. By applying this technique, new area boundaries can be defined which are wholly dictated by similarities important to the research question under consideration rather

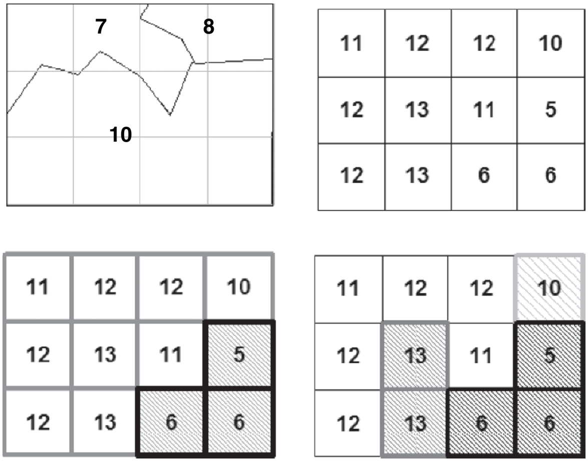

Fig. 8.5 An example of 'Lossy' Boundaries

than (say) the identification of a series of areas with similar population counts. Obviously this does not provide a complete solution to the problems associated with boundary definition as its viability relies purely upon the correct selection of those factors considered in the calculation of 'similarity'. However, it may offer an incremental step towards a more hypothesis driven approach to area boundary definition.

To reiterate in conclusion, the research demonstrates that an algorithm based on the findings from the near repeat literature predicts the future locations of crime at a level that exceeds chance expectation, and also outperforms other hot-spotting methods. Considering Thematic mapping in particular, for which the unit of analysis was police beats, this approach was not only worse than ProMap but typically failed to exceed chance expectation. Consequently, the findings provide support for the hypothesis on which the model was based. The inclusion of data concerned with the opportunity structure for the area considered improved the performance of the model, but an initial attempt at modeling barriers did not. The results have implications for both theory and practice. Maintaining a balance between theory refinement and applicability enhancement is difficult. It is contended that their implementation in practice should not wait upon further research. In the light of the research already reported, it seems difficult to justify basic KDE crime mapping in preference to ProMap as enhanced by road and building composition. The research agenda which takes enhanced ProMap as its starting point should contain at least the following elements:

1. Recalculating for periods which are operationally meaningful (e.g., police shifts and periods for which tasking and coordinating meetings plan);
2. Ensuring by calculation of patrolling routes that high risk areas are not so dispersed that transit times between high risk areas offset the advantages which ProMap offers;

3. Considering thresholds of the proportion of predicted burglaries which it is realistic to police, including consideration of policing styles which involve thorough scrutiny of the highest risk areas with more superficial search of a greater number of areas;
4. Considering other crime types, to produce a pan-crime ProMap, devising search patterns reflecting the relative seriousness of individual crime types.

Much of this drive towards applicability would appropriately be carried out by crime analysts in local areas, in a federated research programme. Winston Churchill, in World War II visited Washington to plead for assistance. He opined 'Give us the tools and we will finish the job'. A cynic's version is that for social scientists, the mantra would be 'Give us the tools, and we will spend the rest of our lives sharpening them'. This would be a fate worse than death for crime mapping. The techniques are already fit for purpose and need to be sympathetically but urgently integrated into the craft of policing.

Appendix 1 Equations Used in the Derivation of the KDE and ProMap Algorithms

KDE Equation and Callibration

To generate the KDE surfaces used, a quartic function (Bailey and Gatrell, 1995) was used. The formula is shown as equation (8.1). There exists no obvious consensus of opinion regarding the appropriate spatial bandwidth to use in KDE hotspot generation. Some suggest using a bandwidth of one-tenth of the smallest dimension of the study area. This yields a nicely smoothed map for visual display but is determined by the dimensions of the study area rather than criminological understanding or the requirements of operational policing. A variety of alternative bandwidths or methods for calibration have been proposed, but no theoretical explanations are offered to underpin the selection, or empirical studies provided to demonstrate any superior heuristic value. In the absence of a rationale for bandwidth selection, and on the basis of a suggestion by Ratcliffe (2000), we used a bandwidth of 200 m. This is relevant not least because it is used by many UK crime analysts with whom the authors have had contact.

$$\lambda_\tau(s) = \sum_{di \leq \tau} \frac{3}{\pi \tau^2} \left(1 - \frac{d_i^2}{\tau^2}\right)^2 \tag{8.1}$$

Where, $\lambda_\tau(s)$ = risk intensity value for cell s

τ = bandwidth
d_i = distance of each point (i) within the bandwidth from the centroid of the cell

ProMap Equation and Callibration

Equation (8.2) shows the formula used in the computation of the ProMap risk surface.

$$\lambda_\tau(s) = \sum_{c_i \leq \tau \cap e_i \leq \upsilon} \left(\frac{1}{(1 + c_i)}\right) \frac{1}{(1 + e_i)} \tag{8.2}$$

Where, $\lambda_\tau(s)$ = risk intensity value for cell s

τ = spatial bandwidth υ = temporal bandwidth
c_i = number of cells between each point (i) within the bandwidth and the cell
e_i = weeks elapsed for each point (i) within the temporal bandwidth

Similar in basic terms to that used to derive the KDE maps, the equation includes an additional parameter (e_i) to model the effect of elapsed time on the risk of crime at each location. The spatial bandwidth used is determined by the results of the Knox analysis. For example, if the Knox analysis suggests that events cluster in space and time more than would be expected on a chance basis for a distance of (say) 500 m and two months, then the spatial bandwidth used would be 500 m.

Modeling Accessibility

The basic ProMap approach was modified to try to reflect the influence of the urban backcloth on offender spatial target choices. Two factors were considered, the type and number of roads in each cell, and the number of homes:

Roads

This was used to provide a crude estimate of accessibility. The weighting considered the length, classification and number of road sections within each cell. Cells with more and larger roads were assigned a higher accessibility weighting to reflect the fact that they were likely to be more connected to surrounding cells than other cells. The following equations show the construction of the road weighting:

Road type = weighting	Length of A rds in cell +	0.8 * length B rds in cell +	0.6 * length of ‘minor’ rds
Road number = weighting	Number of A rds in cell +	Number of B rds in cell +	Number of ‘minor’ rds

Where the road type weighting was above 20 (the average length of road across cells) and the road number weighting was greater than zero, a weighting of 1.01 was applied. Where the road type weighting was less than 20 a weighting of 1.0 was used. Where there were no roads, a weighting of zero was assigned. The weights were arbitrary, but the results are robust across weighting options.

Housing

Where a cell had more than the average amount of houses (7 homes) a weighting of 1.01 was applied. Where a cell had no homes within it, a weighting of zero was used. In all other cases, a weighting of 1.0 was applied.

Roads and Homes

A multiplicative function was used to weight the cells, so that the combined weighting was the product of the above variables. These weightings were then used to generate three new event driven opportunity surfaces by taking the product of these layers and the ProMap surface.

Barriers

To model the effect of barriers, a modification of equation (8.2) was required to reflect whether contributing events had occurred in an area similar (or not) to the cell for which a prediction was being made. If we consider the ProMap formula as a form of regression equation, we can alter the slope or intercept (or both) of the model using a parameter to model the presence or absence of a barrier. Considering the effect of changing the intercept, an event will acquire a change in weight on the basis of the area within which it is located, irrespective of its proximity to a barrier. Changing the slope takes account both of the existence of a barrier and a cell's proximity to it and is thus the preferred approach here.

To model the effects of barriers, equation (8.2) is modified only slightly. In this case, the denominator in the first term of the equation $(1 + c_i)$ is replaced with the following:

$$(1 + c_i)(1 - \alpha)$$

where,

(1) $\alpha = 0$ when the lifestyle value of the cell for which the prediction is being derived and the event considered is located are the same, and
(2) $\alpha > 0$ when the lifestyle values differ.

A series of calibration trials were conducted to identify the optimum value of α. To do this, predictions were generated using values ranging from 0.5 to 0.99, with increments of 0.01 being tested across trials. The value ultimately selected of 0.98 was then fixed. That is, α was always 0.98 when condition (2) above was met. That the slope found to be optimal was virtually at the top of the range tested may be taken as a preliminary indication of the salience of social barriers as here measured.

Appendix 2 Monte Carlo Simulation Algorithm for Estimating Chance Expectation

The simulation worked in the following way:

1. Using a uniform random number generator (RNG) select the first cell from those available (1–10,816) and assign this the highest risk value
2. Select another number using the RNG

   a. If this has not previously been selected, assign this the next highest value. If it has already been selected, choose another value until a previously unselected cell is identified.
   b. Repeat step 2 until all cells have been allocated a risk value from 1 to 10,816

3. Rank order the data using the risk values assigned to each cell
4. Generate an accuracy concentration curve
5. Repeat the above steps 99 times, storing the results of each iteration

References

Armitage, R. (2007). Sustainability versus safety: Confusion, conflict and contradiction in designing out crime. In: G. Farrell, K. J. Bowers, S. D. Johnson & M. Townsley (Eds.), *Imagination for Crime Prevention: Essays in Honour of Ken Pease*. Monsey, NY: Criminal Justice Press.

Ashton, J., Brown, I., Senior, B., & Pease, K. (1998). Repeat victimisation: Offender accounts. *International Journal of Risk, Security and Crime Prevention*, 3(4), 269–279.

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis*. Harlow: Longman.

Beavon, D. J. K., Brantingham, P. L., & Brantingham, P. J. (1994). The influence of street networks on the patterning of property offenses. In: R. V. G. Clarke (Ed.), *Crime Prevention Studies*. Vol. 2, Monsey NY: Criminal Justice Press.

Brantingham, P. L., & Brantingham, P. J. (1993). Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology*, 13, 3–28.

Brantingham, P. J., & Brantingham, P. L. (1981) *Environmental criminology*. Rev. ed. 1991 Prospect Heights, IL: Waveland, 1991.

Brantingham P. J. (1978). A Theoretical model of crime site selection. In: M. D. Krohn & R. L. Ackers (Eds.), *Crime, Law and Sanctions: Theoretical Perspectives*. Beverly Hills: Sage.

Bernasco W. (2008). Them Again? Same-offender Involvement in Repeat and Near Repeat Burglaries. *European Journal of Criminology*, 5(4), forthcoming.

Bernasco, W., & Luykx, F. (2003). Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighbourhoods. *Criminology*, 41, 981–1001.

Bernasco, W., & Nieuwbeerta, P. (2005). How do residential burglars select target areas? *British Journal of Criminology*, 45, 295–315.

Bennett, T., & Wright, R. (1984). *Burglars on Burglary*. Aldershot: Gower.

Besag, J., & Diggle, P. J. (1977). Simple monte carlo tests for spatial pattern. *Applied Statistics*, 26, 327–333.

Bowers, K. J., & Johnson, S. D. (2005). Domestic burglary repeats and space-time clusters: The dimensions of risk. *European Journal of Criminology*, 2(1), 67–92.

Bowers, K. J., Johnson, S., & Pease, K. (2004). Prospective hotspotting: The future of crime mapping? *British Journal of Criminology*, 44(5), 641–658.

Bowers, K. J., Johnson, S., & Pease, K. (2005). Victimisation and re-victimisation risk, housing type and area: A study of interactions. *Crime Prevention and Community Safety*, 7, 7–17.

Brown, P. J. B. (1991). Exploring Geodemographics.' In: I. Masser & M. Blakemore (Eds.), *Geographic Information Management: Methodology and Applications*, London: Longman.

Coupe, T., & Blake, L. (2006). Daylight and darkness targeting strategies and the risks of being seen at residential burglaries, *Criminology* 44(2), 431–464.

Cornish, D. (1994). The procedural analysis of offending and its relevance for situational crime prevention. In: R. Clarke (Ed.), *Crime Prevention Studies*, (Vol. 3, pp. 151–196). Monsey, New York: Criminal Justice Press.

Cornish, D., & Clarke, R. (1986). *The Reasoning Criminal*. New York: Springer-Verlag.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1, 269–271.

Ericsson, U. (1995). Straight from the Horse's Mouth. *Forensic Update*, 43, 23–25.

Everson S. P., & Pease K. (2001). Crime against the same person and place: detection opportunity and offender targeting. In: G. Farrell & K.Pease (Eds.), *Repeat Victimisation*. Monsey NY: Criminal Justice Press.

Everson, S. (2003). Repeat victimisation and prolific offending: Chance or choice? *International Journal of Police Science and Management*, 5, 180–194.

Farrell, G. (2005). Progress and prospects in the prevention of repeat victimization. In: N. Tilley (Ed.), *Handbook of crime Prevention and Community Safety*. Cullompton, UK: Willan.

Groff, E. R., & LaVigne, N. G. (2001). Mapping an opportunity surface of residential burglary. *Journal of Research in Crime and Delinquency*, 38(3), 257–278.

Hillier, B. (2004). Can streets be made safe? *Urban Design International*, 9, 31–45.

Hakim, S., Rengert, G. F., & Schachmurove, Y. (2001). Target search of burglars: A revised economic model. *Papers in Regional Science*, 80, 121–137.

Johnson, S., Bowers, K., & Hirschfield, A. (1997). New insights into the spatial and temporal distribution of repeat victimisation. *British Journal of Criminology*, 37(2), 224–241

Johnson, S. D., Bernasco, W., Bowers, K. J., Elffers, H., Ratcliffe, J., Rengert, G., & Townsley, M. T. (2007). Near repeats: A cross national assessment of residential burglary. *Journal of Quantitative Criminology*, 23(3), 201–219.

Johnson, S. D., & Bowers, K. J. (2004a). The burglary as a clue to the future: the beginnings of prospective hot-spotting. *European Journal of Criminology*, 1(2), 237–255.

Johnson, S. D., & Bowers, K. J. (2004b). The stability of space-time clusters of burglary. *British Journal of Criminology*, 44(1), 55–65.

Knox, G. (1964). Epidemiology of childhood leukaemia in northumberland and durham. *British Journal of Preventative and Social Medicine*, 18, 17–24.

Matthiopoulos, J. (2003). The use of space by animals as a function of accessibility and preference. *Ecological Modelling*, 159, 239–268.

Morgan, F. (2001). Repeat burglary in a Perth Suburb: Indicator of Short-Term or Long-Term risk? In: G. Farrell, & K. Pease (Eds.), *Repeat Victimisation: Crime Prevention Studies*. Vol. 12, Monsey NY: Criminal Justice Press.

Nee, C., & Meenaghan, A. (2006). Expert decision making in burglars. *British Journal of Criminology*, 46(5), 935–949.

North, B. V., Curtis, D., & Sham, P. C. (2002). A note on the Calculation of Empirical P Values from Monte Carlo Procedures. *American. Journal. of Human. Genetics*, 71, 439–441.

Pease, K. (1998). *Repeat Victimization: Taking Stock*. The Home Office: Police Research Group: Crime Detection and Prevention Series Paper 90.

Polvi, N., Looman, T., Humphries, C., & Pease, K. (1991). The time course of repeat burglary victimisation. *British Journal of Criminology*, 31(4), 411–414.

Poot, C., Luykx, F., Elffers, H., & Dudnik, C. (2005). Hier wonen en daar plegen? Sociale grenzen en locatiekeuze. *Tijdschrift voor Criminologie*, 47(3), 255–268.

Ratcliffe, J. (2000). Aoristic analysis: The spatial interpretation of unspecific temporal events. *International Journal of Geographical Information Science*, 14(7), 669–79.

Ratcliffe, J. (2006). A temporal constraint theory to explain opportunity-based spatial offending patterns. *Journal of Research in Crime and Delinquency*, 43(3), 261–291.

Rengert, G., & Wasilchick, J. (2000). *Suburban Burglary: A Tale of Two Suburbs*. Springfield, IL: Charles Thomas.

Lebeau, J. L., & Rengert, G. (2006). The role of Racial/Ethnic barriers in the movement of drug offenders. Presentation at the *Environmental Criminology and Crime Analysis Seminar*, Vancouver, June, 2006.

Reynald, R., Averdijk, M., Elffers, H., & Bernasco, W. (2006). *Do social barriers affect urban crime trips? The effects of ethnic and socio-economic neighborhood compositions on the flow of crime in The Hague, The Netherlands*. (NSCR 2006-4) Leiden.

Shaw M., & Pease K. (2000). *Research on Repeat Victimisation in Scotland'*. Crime and Criminal Justice Research Findings 44. Edinburgh: Scottish Executive.

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

Tseloni A., Wittebrood, K., Farrell, G., & Pease K. (2004). Burglary Victimisation in the US, England and Wales and the Netherlands: Cross-national comparison of routine activity patterns. *British Journal of Criminology*, 44, 66–91.

Tseloni A., & Kershaw, C. (2005). Predicting crime rates: fear and disorder based on area information. Evidence from the British Crime Survey. *International Review of Victimology*, 12, 295–313.

Winchester, S. W. C., & Jackson, H. (1982). *Residential Burglary: The Limits of Prevention*. Home Office Research Study 74. London: HMSO.

Wright, R.T., & Decker, S. (1994). *Burglars on the Job: Streetlife and Residential Break-Ins.* Boston: Northeastern University Press.