Using circular statistics to analyse time patterns in crime incidence

Chris Brunsdon [a],[*], Jon Corcoran [b]

[a] Department of Geography, University of Leicester, University Road, Leicester LE1 7RH, UK
[b] Centre for Research into Sustainable Urban and Regional Futures, School of Geography, Planning and Architecture, University of Queensland, Australia

Abstract

A set of techniques referred to as *circular statistics* has been developed for the analysis of directional and orientational data. The unit of measure for such data is angular (usually in either degrees or radians), and the statistical distributions underlying the techniques are characterised by their cyclic nature—for example, angles of 359.9° are considered close to angles of 0°. In this paper, we assert that such approaches can be easily adapted to analyse time-of-day and time-of-week data, and in particular daily cycles in the numbers of incidents reported to the police. We begin the paper by describing circular statistics. We then discuss how these may be modified, and demonstrate the approach with some examples for reported incidents in the Cardiff area of Wales.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Circular statistics; Public disorder; Kernel density

1. Introduction

In a number of areas of police resource management, analysis of the relative frequency of events occurring at different times of day plays an important role. Identifying times of day when certain kinds of crime or disorder are more likely to occur offers a first step towards prioritising resources and improved targeting. Because of this, there has been

[*] Corresponding author. Tel.: +44 0116 252 3843.
 E-mail address: cb179@le.ac.uk (C. Brunsdon).

considerable interest in analytical and visual techniques which help to identify patterns in the time of day of crimes—see Ratcliffe and McCullagh (1998a) as an example. However, although the use of GIS to explore the spatial patterning of crime incidence has received growing attention over recent years (see for example Murrey, McGuffog, Western, & Mullins, 2001; Ratcliffe & McCullagh, 1998b), the analysis of temporal patterns of crime incidence has, in comparison, been somewhat neglected in the wake of such spatial studies. One reason for this might be the difficulty of representing and analysing temporal data in a GIS environment. The work presented in this paper attempts to address this imbalance through the promotion of inferential circular statistics as a novel method by which crime patterns can be investigated.

Time of day is also important from a theoretical viewpoint. Cohen and Felson (1979, p. 589) present their routine activities theory (RAT) in which they argue that the occurrence of a crime is a result of a convergence in time and space of three fundamental elements, namely; motivated offenders, suitable targets and absence of capable guardians against violation. The theory argues that an understanding of crime patterning rests upon knowledge of these three fundamental elements, and where they converge in time and space a crime is more likely to occur. The theory can then be used to design appropriate preventative measures that attempt to modify one of the three elements and thus avert a crime from taking place. Van Koppen and Jansen (1999) provide empirical support for this theory in their investigation of commercial robberies. Their study analysed commercial robberies over time to identify daily, weekly and monthly patterns concluding that the observed temporal variations were consistent with patterns implied by routine activity theory. Further examples that support the utility of this theory in the explanation of crime include Roncek and Maier (1991) in their study of bar-related crime, and Sherman, Gartin, and Buerger (1989) who investigated hot spots of predatory crime.

RAT was developed further into pattern theory (Brantingham & Brantingham, 1993) where interactions of people and places are modelled using the concept of nodes and pathways in space and time. People travel along known activity pathways to nodes around which the offender searches for potential targets. Similarly the potential victims routine activities are of importance and also have the power to affect the patterning of crime. The theory identifies the importance of triggers, described as cues to the committal of a crime. For example, an accidental physical contact that leads to a spillage of drink in a pub may trigger an aggravated response—see Brantingham and Brantingham (1981, p. 268).

Thus, pattern theory reinforces the idea that an understanding of crime patterning rests upon knowledge of activity and awareness spaces. Convergence of the three fundamental elements of RAT coupled with one or more triggering mechanisms create a scene with enhanced potential for crime to occur. Clearly, the circumstances contributing to these conditions will vary throughout the day. To investigate such ideas empirically, it is necessary not only to be able to model frequency distributions for a given crime type, but also to compare distributions between different crime types, or the same crime type under different circumstances.

The term *circular statistics* describes a set of techniques used to analyse and model distributions of random variables that are cyclic in nature (Mardia, 1975). For example angles or directions differing by an integer multiple of 360° or 2π radians are considered to be equivalent. These techniques have enjoyed popularity in a number of areas where exploration, modelling and testing hypotheses of directional or orientational information has played a rôle. Surprisingly, most work involving circular statistics has concentrated on directional data as described above, although time-of-day data is also cyclic and there is no

reason why circular statistics cannot be applied. In this paper, we will demonstrate their use in a number of tasks:

- Modelling daily cycles of frequencies of events.
- Comparing daily cycles for two data sets.
- Visualising geographical patterns in daily cycles.

It is also worth noting that although discussion up to now has considered *daily* cycles of crime frequency, these techniques can also be applied to *weekly* or even longer cycles. An alternative approach to circular statistics may be to use frequency domain time time series analysis (see for example, Chatfield, 2003, chaps. 6 & 7) to identify and model these cyclic effects. Although this may be the subject of future research, it should be noted that there are a number of complicating factors. Firstly, conventional time series analysis requires regularly spaced data. Although current data is rounded to the nearest hour and easily convertible to such a format, it is possible that in the future incident times could be recorded to a greater degree of accuracy, and that exploiting the extra detail would require a technique capable of analysing irregular data. Secondly, there is an assumption that time series data follows a Gaussian distribution—whereas hourly counts would clearly not follow such a distribution. The theory would need to be suitably adapted to allow for this. For these reasons, it seems that the circular statistical approach is most appropriate at least for an initial examination of cyclic crime patterns.

The paper will begin with an overview of basic circular statistics aimed at readers with a more general statistical knowledge. Following this, two examples will be given using police incident log from Cardiff, Wales. Consideration will then be given to the analysis of geographical patterns, again followed by an example from Cardiff.

2. An overview of circular statistics

This section provides a brief introduction to circular statistics. More complete texts on this subject are Fisher (1993), Jammalamadaka and SenGupta (2001) and Mardia and Jupp (1999). An important initial question is "why do you need circular statistics?". This is perhaps best answered by considering the following example. Suppose public disorder incidents were recorded at the times 2330, 0015, 0030 and 0045 during a given night. We could convert these times to decimal hours and obtain a mean value, which would be computed as $(23.5 + 0.25 + 0.5 + 0.75)/4 = 6.25$ or 0615 when converted back to a time of day. An average time of just after six in the morning makes little sense for a set of events clustered at around midnight—and illustrates the folly of applying a standard (non-circular) mean to cyclic measurements.

In situations such as this, a more appropriate *circular mean*, denoted by h, is defined: if

$$A = \sum_{i=1}^{n} \sin(x \, h_i) \quad \text{and} \quad B = \sum_{i=1}^{n} \cos(x \, h_i) \tag{1}$$

then

$$h = \begin{cases} x^{-1} \arctan(A/B) & \text{if } B \ge 0 \\ x^{-1}(\arctan(A/B) + \pi) & \text{if } B < 0 \end{cases} \tag{2}$$

where $\{h_i : i = 1, \ldots, n\}$ are the n observed cyclic data values and x is a constant that scales the cyclic values to the range $[0, 2\pi)$. For time of day data x $= \pi/12$, and for data on a weekly cycle x $= 2\pi/7$. For the example given, with the times converted to decimal hours, it may be verified that h ¼ 0:2508 or just after 0015.

There is a relatively simple geometrical interpretation of the circular mean.

If the h_i's are represented as points on the circumference of a circle, as in Fig. 1, then the point labelled **P** is their centroid in two-dimensional space. If it is assumed that the circle has a radius of one unit, then it may be shown that **P** has coordinates $(A/n, B/n)$. A line extended from the centre of the circle through **P** will eventually meet the circumference of the circle at point **Q**. The position of **Q** then corresponds to the value of h.

Another useful circular statistic is the index of dispersion, which measures the degree of spread for circular data. Again, this can be derived geometrically. In Fig. 2, there are two panels both showing circular data sets. That on the left is less spread out than that on the right. The centroids of these two data sets as considered earlier are shown as the points **P1** and **P2** on the two panels. **P1** is much further away from the centre of the left-hand circle than **P2** is from the centre of the right-hand circle. This is generally the case—data sets with a greater degree of spread have centroids closer to the centre of the circle.
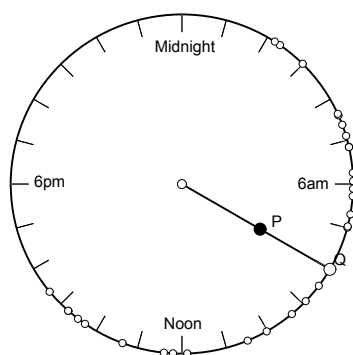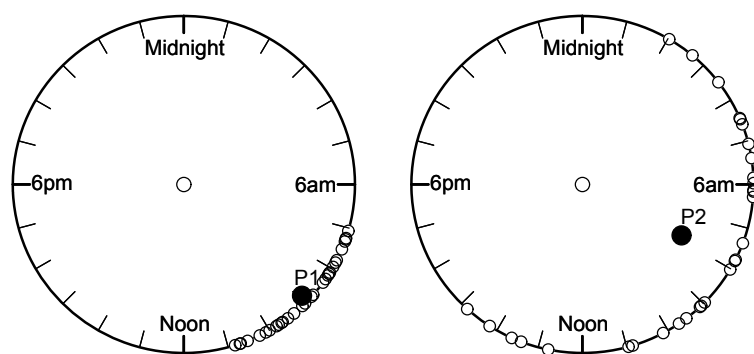

Fig. 1. Geometrical derivation of a circular mean.

Fig. 2. Geometrical derivation of a circular deviance.

This is the basis for an index of circular spread, $\overline{R}$:

$$\overline{R} = \frac{1}{n}(A^2 + B^2)^{\frac{1}{2}} \qquad (3)$$

where *A* and *B* are as defined in (1). Note that with $\overline{R}$, higher values are associated with *less* spread in the data. Another measure, the *Circular Dispersion*, $\hat{d}$, is defined by

$$\hat{d} = 1 - \frac{q_2}{2\overline{R}^2} \qquad (4)$$

where

$$q_2 = n^{-1} \sum_{i=1}^{n} \cos(2(x - h_i)) \qquad (5)$$
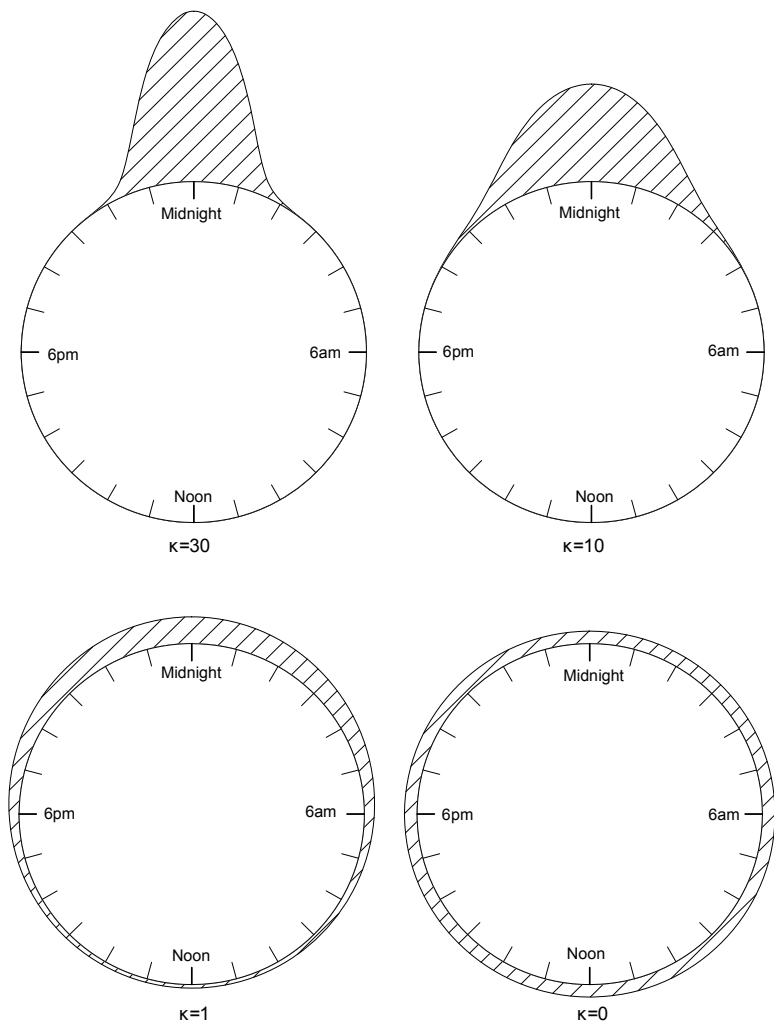

Fig. 3. The von Mises distribution.

Table 1
Some hypothesis tests for cyclic data

Purpose of test	Hypothesis	Name of test
Test whether h has a uniform distribution	$j = 0$	Rayleigh (1905)
Test whether h has the same mean for m distributions	$l_1 = l_2 = \ldots = l_m$	Watson (1983)

This is in some respects more akin to the non-circular measure of standard deviation, as it has no upper bound, and larger values correspond to greater degrees of spread.

As well as basic descriptive statistics for circular models, we may go on to consider probability density functions. The *von Mises distribution* is often regarded as a 'natural' distribution for cyclic data, and shares some properties with the normal distribution for non-cyclic data. The probability density of h is given by

$$vm(h|l;j) = \frac{1}{2\pi x\, I_0(j)} \exp(j\,\cos(x(h-l))) \qquad (6)$$

where x is a scaling constant as before, l is the mean value of h, j is the concentration parameter related to the circular spread, and $I_0(j)$ is a normalisation constant to ensure that the probability density integrates to one.

Fig. 3 illustrates the von Mises distribution for $j = \{30, 10, 1, 0\}$. From this it can be seen that the distribution tends towards the shape of the normal distribution when j gets large, and is a uniform distribution on the circle when j is zero. Given a set of observations $\{h_i : i = 1, \ldots, n\}$ the maximum likelihood estimators (MLE) of l and j are given by

$$\hat{l} = \bar{h}$$
$$\hat{j} = A_l^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} \cos(x(h_i - \bar{h}))\right) \qquad (7)$$

where A_l^{-1} is a function defined, for example, in Fisher (1993) or Jammalamadaka and SenGupta (2001).

The above gives a brief view of how one can summarise cyclic data, and also fit one kind of model, the von Mises distribution. In addition to this, it is also possible to carry out hypothesis tests. Two common tests used with cyclic data are set out in Table 1.

Other tests exist and will be introduced later in the text.

3. Example data: criminal damage in Cardiff

The data set described in this section relates to reports of criminal damage in the city of Cardiff, Wales. Cardiff, officially the capital of Wales since 1955 covers an area of 524 acres (46 square miles) and forms the selected study locality for this research. According to the 2001 UK Census of Population, the city houses a population of 305,353 inhabitants and consists of 127,476 households. The city also houses a large student population (in excess of 25,000 in 1991).

Police incident data and specifically reports of criminal damage and disorder/disturbance are used in this research. Each incident in Cardiff that is reported to the police is logged in this data set and includes information pertaining to the type of the incident, date, time and its location as an (x, y) co-ordinate pair. Despite the various weaknesses associated with this source of data such as both under and over-reporting (for a fuller discussion

see for example, Sherman et al., 1989) and the time lag between the occurence of an incident and the time of its eventual reporting (discussed in more detail later in the paper), the analysis of incident data still provides a valuable indicator of the demand on police resources. As such it can be used as a foundation from which scheduling and tactical deployment decisions can be established and their success subsequently evaluated. These span a period of two years from 1st July 1999 to 30th June 2001.

The remainder of this section commences with an inspection of the data using a circular model for time of day assuming a von Mises distribution to assess differences between days of the week. A more flexible, non-parametric method using kernel density estimation is then presented to model the circular distribution. The section concludes by presenting a method by which several kernel density estimated distributions can be compared using a Monte-Carlo randomisation approach.

3.1. Initial data inspection

The data contain information on 10,359 incidents of criminal damage. An initial inspection of the data (Fig. 4) reveals a possible bimodality, where frequency of reporting times peaks at around 2300, falls off at around 0500 and peaks again at around 1000, before falling again slightly, and then gradually rising throughout the day. Note that data here are rounded to the nearest hour. This is not seen as a major problem: Fisher (1993, p. 35) notes that rounding errors do not become problematic when estimating sample statistics until circular data are assigned to around 12 groups (that is, rounding to the nearest 2 h)—furthermore he specifically cites rounding to the nearest hour as not being problematic.
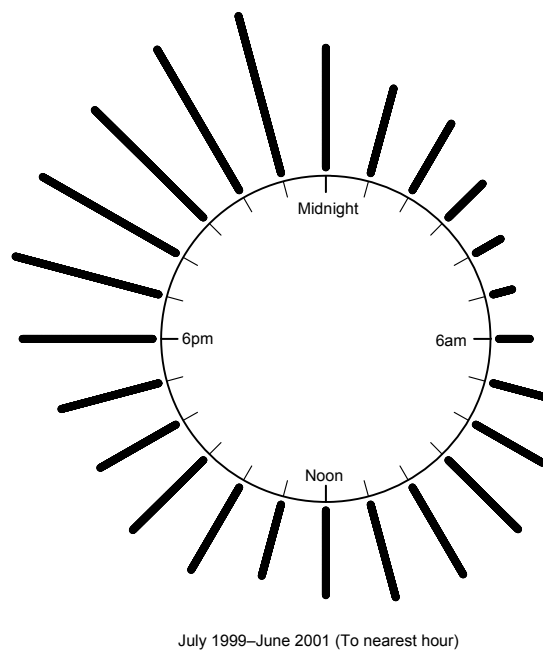

July 1999–June 2001 (To nearest hour)

Fig. 4. Temporal distribution of criminal damage incidents.

Assuming a von Mises distribution, the maximum likelihood estimate of l is 1912 (rounding to nearest minute), and that of j is 0.482. However, it is worth noting that there is some doubt whether a von Mises distribution is an appropriate model, since Fig. 4 suggests that the distribution is multimodal. We will discuss methods avoiding this distributional assumption later in the paper.

3.2. Differences by day of week

A useful initial analysis is to compare the distribution of incidents on different days of the week. This helps to address a practical policing issue of human resource deployment. In Table 2, estimated values of l and j are given for data grouped by day of week. Again, the von Mises distribution model is used in each case.

The dispersion parameters (j) are similar (although the ones toward the end of the week are slightly higher than the others) but the mean parameters for earlier days in the week are notably earlier (particularly Monday) and for Saturday the parameter is particularly high. The latter is possibly due to a greater frequency of incidents later on Saturday night, when some bars close at 2300, and drinkers move towards the later opening bars and clubs. We may test whether the differences between mean times for incidents are significant using a Watson test (Watson, 1983). This test assumes that the incident distributions for each day are all von Mises, with common j's. Under this assumption it tests the hypothesis that the means are also identical for each day. The test statistic for the Watson test is defined by

$$Y_r = 2(N - R_p)/\hat{d}_0 \tag{8}$$

where

$$R_p = \left[\left(\sum_{i=1}^{r} n_i \cos(\hat{l}_i) \right)^2 + \left(\sum_{i=1}^{r} n_i \sin(\hat{l}_i) \right)^2 \right]^{\frac{1}{2}} \tag{9}$$

and

$$\hat{d}_0 = \sum_{i=1}^{r} n_i \hat{d}_i / N \tag{10}$$

where N is the total number of incidents, n_i is the number of incidents on day i, r is the number of subgroups of data (here $r = 7$), $\hat{l}_i$ is the sample circular mean for group i and $\hat{d}_i$ is the sample circular dispersion for group i. The statistic Y_r is compared against

Table 2
von Mises distribution parameters for data grouped by day of week (criminal damage)

Day of week	$\hat{l}$	$\hat{j}$
Monday	1721	0.430
Tuesday	1816	0.499
Wednesday	1824	0.525
Thursday	1931	0.492
Friday	1931	0.521
Saturday	2040	0.523
Sunday	1949	0.517

a v^2 distribution with $r - 1$ degrees of freedom. In this example, $Y_r = 1207.6$ with 6 degrees of freedom, which has a *p*-value less than 0.01, suggesting that there is very strong evidence *against* the hypothesis that the mean time is the same for all days.

3.3. More flexible tests

The previous evidence for difference in the distributions appears impressive, but it should be remembered that it is conditional on the assumption that the data follows a von Mises distribution. Fig. 4 suggests that there is a degree of bimodality in the data—and other figures (not shown here) for individual days of the week suggest the same thing. Could it be that the previous encouraging result is due to the violation of the von Mises distribution assumption, rather than a genuine difference in daily patterns? In order to answer this, a more general although much more computationally intensive test is proposed. However, before outlining the test, it is necessary to consider a non-parametric approach to circular distribution modelling.

3.3.1. Kernel density estimation for circular data

In previous sections, we have discussed modelling circular distributions using a prescribed functional form, the von Mises distribution. In this case, model calibration is simply a matter of estimating the parameters in this distribution given the data. However, a very different approach is that of kernel density estimation (KDE) (Silverman, 1986). Here the aim is to construct a probability density curve solely on the basis of the data. With KDE there is no simple algebraic expression for the distribution, but the payback is that the distribution estimates are no longer constrained to follow a distinct functional form. If the data suggests, they may take on any degree of skewness or have any number of modes.

A kernel density estimate $\hat{f}(h)$ is obtained using the formula below:

$$\hat{f}(h) = (nh)^{-1} \sum_{i=1}^{n} K\left(\frac{h - h_i}{h}\right) \qquad (11)$$

where $K(\cdot)$ is any valid probability density function, and h is a constant called the *bandwidth*. Typically $K(\cdot)$ is a symmetrical unimodal function. The general idea is that the estimate is the mean of n roughly bell-shaped distributions, each centred on an observed data point h_i. Although $K(\cdot)$ is symmetrical and unimodal, the mean of n differently centred $K(\cdot)$'s need not be. The rôle of h is to control the smoothness of the estimate. Very low values of h yield 'spikey' density estimates with one spike associated with each data point, whilst high values of h lead to flat density estimates, where genuine features are 'smoothed out'.

Experimentation shows that the choice of $K(\cdot)$ is not critical provided it generally fits the 'bell-shaped' criterion. However, choice of h does have a strong influence on the final density estimate. In fact, under certain regularity conditions it can be shown (Silverman, 1986) that if h decreases at an appropriate rate as n increases, then $\hat{f}(h)$ will tend to the true value of $f(h)$ for all values of h.

For circular data the von Mises distribution is used as $K(\cdot)$, although a minor modification is made so that the bandwidth is j. This results in the revised circular KDE formula:

$$\hat{f}(h) = (nx\, I_0(j))^{-1} \sum_{i=1}^{n} vm(h|h_i; j) \qquad (12)$$
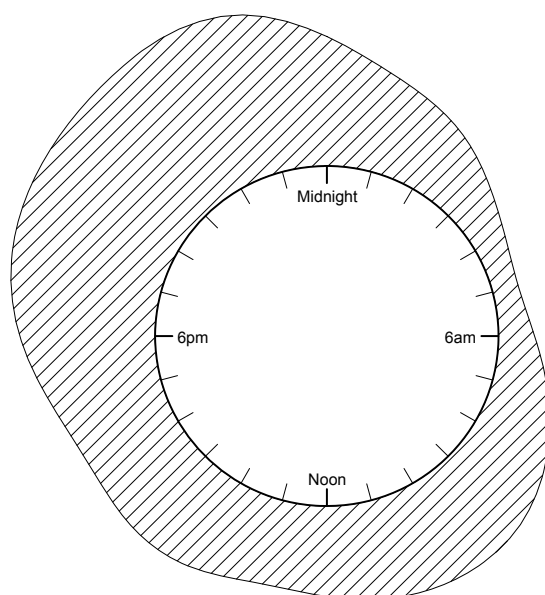
Fig. 5. Temporal distribution of criminal damage reports (KDE).

where vm(μ,κ) is as defined in Eq. (6). Unlike Eq. (11), j is not simply the standard deviation of the kernel distribution $K(\mu)$, and hence has a different normalising expression. However j has a monotone relationship with the standard deviation of the kernels and hence regulates the smoothness of the density estimate.

There are a number of ways of choosing an 'optimal' bandwidth for a given data set[1]. Perhaps the simplest is 'trial and error'—plotting a KDE at a number of different bandwidths often demonstrates levels at which the estimate has become too spikey or too flat. Choosing some level in between this is a matter of judgement. This approach is quite similar to the way in which many people choose the number of classes in a histogram. A number of more rigorous approaches exist—see for example Silverman (1986), Stone (1974), Bowman (1984), Hogg (1979), Sheather and Jones (1991) and Bowman and Azzalini (1997).

In this instance, a 'trial and error' approach was chosen, as experimentation with other methods tended to give spikey circular density estimates—this probably being due to the h_i's being rounded to the nearest hour. Using the 'trial and error approach' a value of j = 30 was chosen. This is shown in Fig. 5. This estimate captures the multimodality in the distribution, and even suggests the possibility of a third small mode at around 1430.

3.3.2. Comparing distributions without a parametric model

The discussion in the previous section addresses the problem of estimating a circular probability distribution when there is no prior information about the form of that distribution. In this section, we explain how this can be taken further, allowing one to compare

---

[1] Optimal is in quotes since there is some discussion about what this term means—and many practitioners of KDE's prefer to choose an optimal bandwidth graphically rather than on the basis of an algorithm.

several circular distribution without making any assumptions about their functional form. To do this, we make use of a general approach proposed in Kent (1983) and Joe (1989). They suggest, as a measure of correlation between two random quantities x and y the statistic

$$s = n^{-1} \sum_{i=1}^{n} \log\left(\frac{\hat{f}_{xy}(x, y)}{\hat{f}_x(x)\hat{f}_y(y)} \right) \qquad (13)$$

where $\hat{f}_{xy}$ is the estimated joint distribution of x and y and f_x and f_y are the respective estimated marginal distributions. If x and y were independent, then $f_{xy}(x, y) = f_x(x)f_y(y)$ and we would expect s to be close to zero. However, if they were not independent, $f_{xy}(x, y)$ would tend to be larger than $f_x(x)f_y(y)$, as it is a better model of the joint probability of x and y, and s would be expected to be larger. If the sampling distribution of s is known, it is possible test the null hypothesis of independance of x and y.

In cases where the sampling distribution is unknown, a Monte-Carlo randomisation approach can be used to simulate the sampling distribution of s (Bowman and Azzalini, 1997). Under the null hypothesis, an observed value of x in the data sample is equally likely to occur with any of the observed values of y. Thus, randomly permuting the x's against the y's and computing s will give a random value of s drawn from its sampling distribution under the null hypothesis. Drawing several[2] such random s's builds up a picture of this sampling distribution, and the value of s for the true data can be compared against this, providing a significance test for the null hypothesis.

For our data, x can be equated to the time of day of an incident, and y can be equated to the day of the week. Note here that y is a categorical variable. This poses no problem. The null hypothesis in this situation is equivalent to stating that the distribution of times of incidents is the same for all days, so that the probability of an incident on day j at time h is just $f(h)p_j$, where p_j is the probability that the event occurs on day j. However, if there is some dependency, then this probability is $f_j(h)p_j$—where f_j is the distribution of h on day j. Thus the test statistic may be written as

$$s = n^{-1} \sum_{i=1}^{n} \log\left(\frac{\hat{f}_j(h_i)}{\hat{f}(h_i)} \right). \qquad (14)$$

Since we make no prior assumptions about the distribution of the time of incidents, both f_j and f are unknown. However, we may use the KDE method outlined in the previous section to obtain estimates. An estimate of $f(h)$ is obtained by applying the KDE approach to all of the recorded h_i's, yielding the result shown in Fig. 5. By grouping the h_i's according to day of the week, and applying the method to each group in turn, we obtain estimates for each of the $f_j(h)$'s. We then compute an estimate of s by substituting these into Eq. (14). For the Monte-Carlo test, we compute s as above, but we use a random permutation of the day of the week against incidents. In this way, we obtain an estimate of the distribution of s under the null hypothesis that the distribution of h's is not dependent on the day of the week.

For the observed data, s is estimated as 0.0158. The results of 10,000 Monte-Carlo simulations are shown in Fig. 6. From this it is clear that the result of this test is still highly
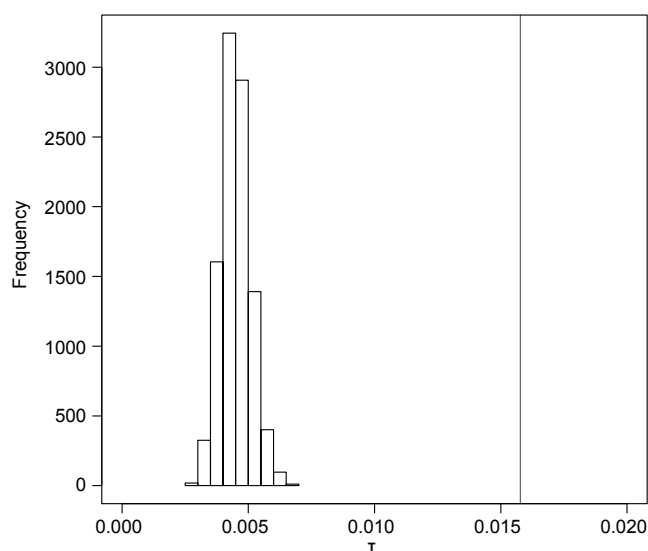
---

[2] Typically around 10,000.

Fig. 6. Results of Monte-Carlo randomisation test for s. The grey line indicates the value for the observed data set.

significant (the *p*-value is less than 0.0001), so that even allowing for the bimodality in the daily patterns of incidence, there is strong evidence that the time-of-day distributions do vary throughout the week.

It is also useful to consider variation for *parts* of the week. although results suggest that there is an obvious difference between the weekend and weekdays, it may be of interest to see if there are more subtle variations *within* weekdays. For example, we may wish to consider events occurring in the early part of the week, defined here as Monday, Tuesday and Wednesday. To do this, the data set is filtered, so that only the h_i's occurring on these days remain. After this, the s-based test is carried out on this subset of the data. When this method is applied to the criminal damage data, the result in Fig. 7 is obtained. Here we see that the observed s is not particularly unusual when compared to the distribution of s's generated under the randomisation hypothesis. In fact, the Monte-Carlo test gives $p = 0.3723$ (for 10,000 simulations) and so we do not have sufficient evidence to reject the null hypothesis that the distributions for Monday, Tuesday and Wednesday are identical. Thus, evidence so far suggests that although distributions are not the same for all days, those for the early part of the week are very similar.

Given this evidence, it is useful to consider the times of events on a weekly cycle instead of a 24-h cycle. In this case, h varies continuously from 0 to 7, with 0 (and 7) representing midnight on Sunday night[3]. Using this representation, a KDE of the time-in-week of incidents is shown in Fig. 8. Here, j was set to 30 once more.

This is a helpful visual aid, as it shows clearly how the daily patterns differ. Although the cycle for each day is bimodal, on the periods covering Friday night to Saturday morning, and Saturday night to Sunday morning, the first mode is noticeably larger than the second. However for the other days of the week, the modes are of roughly equal size.

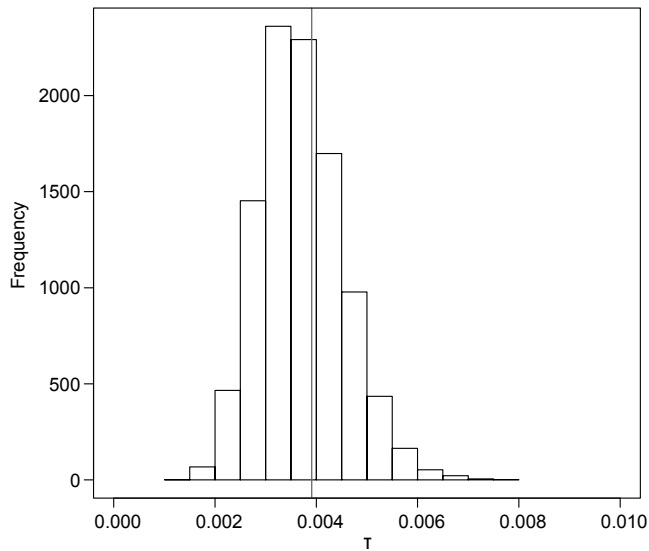[3] This location is arbitrary. There is no reason not to choose another datum point.

Fig. 7. Results of Monte-Carlo randomisation test for s (Monday, Tuesday, Wednesday). The grey line indicates the value for the observed data set.
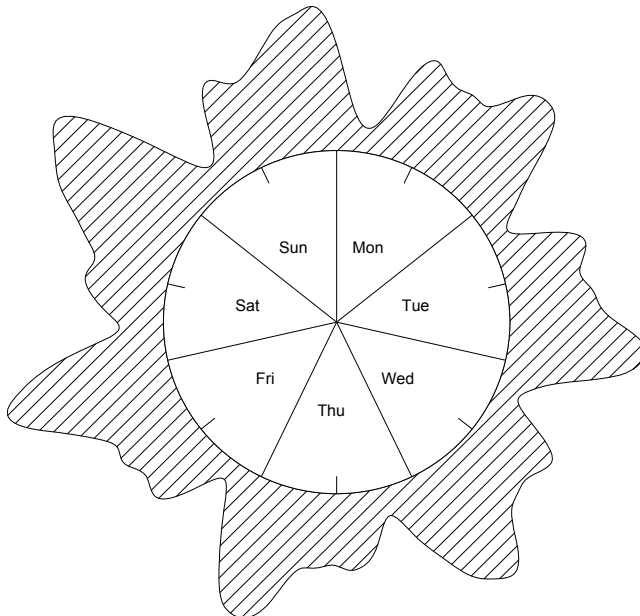

Fig. 8. Temporal distribution of criminal damage incidents (by time-in-week).

As suggested earlier, this accounts for the earlier mean times of crime on Saturday and Sunday seen in Table 2.

Having established a difference between daily patterns of incidence on different week-days, it is now of interest to see whether there are any seasonal differences in pattern.
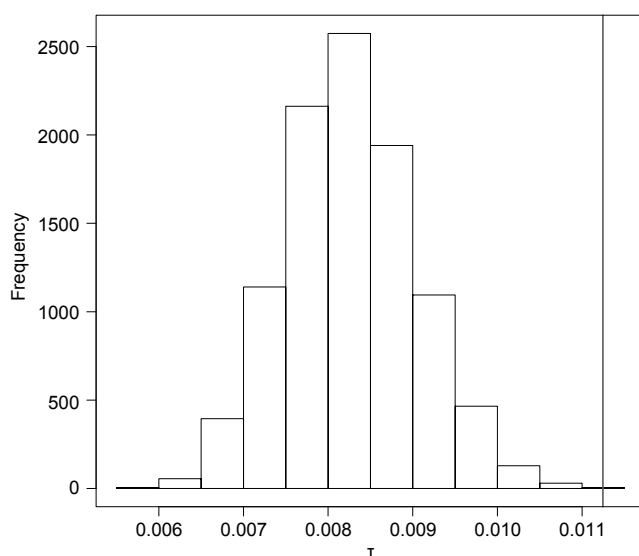
Fig. 9. Results of Monte-Carlo randomisation test for s (monthly effects). The grey line indicates the value for the observed data set.

To do this we now consider each event as a circular variable on a weekly cycle, and carry out a new s test, this time with *j* being the month that the event occurred in. Here the null hypothesis is that weekly cycles do not change between months. Using the same approach as before, we obtain a s of 0.0112. The result of the Monte-Carlo test is shown in Fig. 9.

Again it can be seen that s is highly significant ($p < 0.0001$), and there is strong evidence to reject the null hypothesis. Combining the result of this test and the previous one, we conclude that although there is a distinct weekly cycle in the reporting of criminal damage, this cycle also changes seasonally.

4. Two further examples

4.1. Another kind of incident

An important issue in terms of urban policing is that of *public disorder* which we consider in this section. Policing literature has suggested that the reduction of public disorder issues can be achieved through focused interventions (Braga et al., 1999). In their study of violent places in Jersey City, Braga et al. attributed the successful reduction of crime and disorder following the implementation of a problem-orientated approach to the alteration of the places, routine activities, and situations giving rise to violent behaviour. This problem-orientated approach to policing is, however, reliant upon a detailed analysis in order to identify key times and localities to which the necessary police resources can be deployed. Techniques such as those presented in this paper permit us to illicit key information concerning policing demands, and are central to our ability to design and implement suitable countermeasures. Alongside the criminal damage data set, we consider a data set of public disorder incidents, also from July 1st 1999 to June 30th 2001. As before, each of these is recorded to the nearest hour. The time distribution of these events is shown in Fig. 10, as a
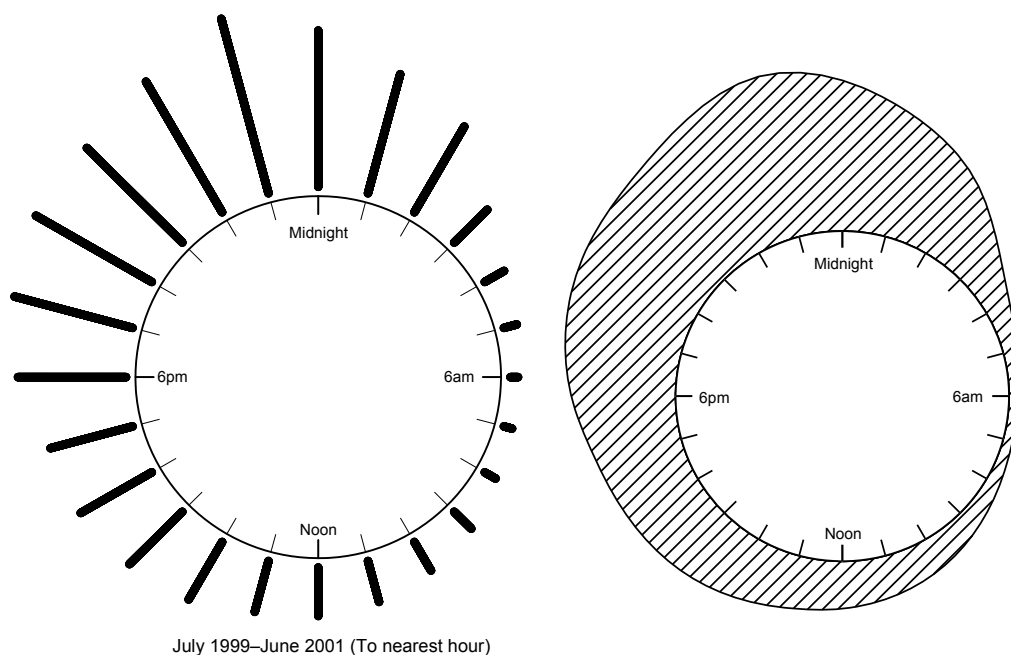
July 1999–June 2001 (To nearest hour)

Fig. 10. Temporal distribution of disorder reports (LHS = Dot Plot, RHS = KDE).

dot plot and as a kernel density estimate. Both of these diagrams suggest a different daily pattern to the criminal damage report, with just one modal time. This is a feature seen in the von Mises distribution. However, in the von Mises case, the minimum value of the distribution is always diametrically opposite the maximum value (i.e., 12 h for a 24-h cycle)— but here the maximum is at around 2300 and the minimum at around 0600.

A study of the weekly pattern again highlights differences from the patterns for criminal damage (Fig. 11). The daily cycle here is shown as quite clearly unimodal, although in common with criminal damage, larger peaks of incidence occur on Friday and Saturday nights.

Although the graphical evidence in this case provides compelling evidence for rejection of the null hypothesis that the distributions of criminal damage and disorder incidents are the same, it is possible to carry out a significance test . The test is the same s-based method outlined in Section 3.3.2. In this case, incidents for both criminal damage and disorder are pooled, and the distribution for the pooled times of day, f, is compared with the two distributions f_{cr} and f_{dam}, these being the distinct distribution functions for criminal damage and disorder, respectively. The histogram of results is omitted here, but the test gave the result that $p < 0.0001$, suggesting the expected recommendation that the null hypothesis be rejected.

It is also possible to apply to this data set all of the tests that were applied to the criminal damage data set. The results of these are given in Table 3. These suggest that, like criminal damage, the disorder patterns vary from month to month, and from day to day on a weekly cycle, although the patterns for the two types of incident are distinct.

Another difference is that for Monday, Tuesday and Wednesday there *is* sufficient evidence to reject the null hypothesis of identical distributions. This suggests that unlike
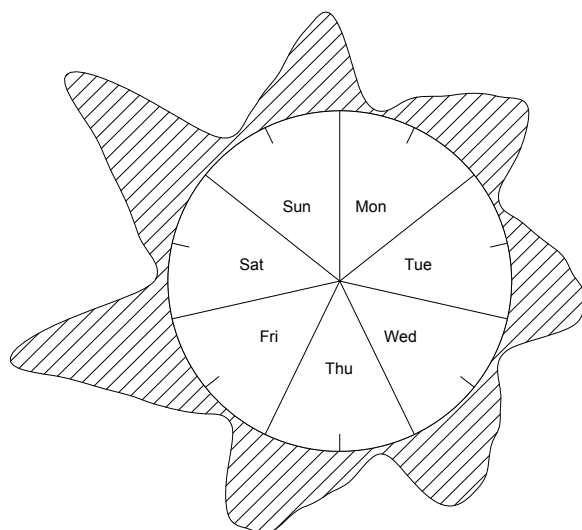
Fig. 11. Temporal distribution of disorder reports (by time-in-week).

Table 3
Results of tests on the disorder data

Test	p-Value
Variation between all days	<0.0001
Variation between months	<0.0001
Variation between Monday, Tuesday and Wednesday	0.0032

the criminal damage incidents, there is some variation in the daily patterns of disorder in the early part of the week.

4.2. Geographical differences

Another important issue in terms of police resource management, and of routine activity theory, is that of the geographical locations of incidents. In particular, it is hypothesised that a number of public disorder incidences may be linked to evening and night-time drinking activities in the city centre. This is well known, and is already considered in policing terms, but it is still helpful to obtain some quantitative insights into this phenomenon. In particular, it is helpful to obtain information about the relative differences in the distributions of events inside and outside of the city centre, in order to identify times at which the increased risk of crime incidence occurs, and also to estimate the relative increase in risk that happens during these times.

Here, we consider both of the Cardiff data sets (criminal damage and public disorder/disturbance). Each incident has a geographical coding, and here we use these to create an indicator variable (CITYCENTRE) taking the value 1 when an incident occurs within the city centre, and 0 when it does not. For the purposes of this study, the city centre is defined as three contiguous census enumeration districts containing the bars, clubs and commercial areas in the centre of Cardiff. These also include points of access to taxis and public

transport. A map of this area in relation to the rest of the Cardiff study area is given in Fig. 12.

For both criminal damage and disorder/disturbance, we group the data by the CITYCENTRE variable. The counts and percentages of each incidence category are given in Table 4.

From this it can be seen that the greater proportion of both incident types occurs outside of the city centre. Note that the totals for each incident type are a little less than those in the examples before, as some incidents were not properly spatially referenced. Applying s-tests to both crime types (grouping by CITYCENTRE in each case and using j = 30) gives *p*-values of less than 0.0001, so that for both crime types we have sufficient evidence to reject the null hypothesis that the temporal distribution of events is the same inside and outside of the city centre. However, although it is useful to have evidence that the 'city centre late night drinking' hypotheses is true, it is perhaps unsurprising. It is more informative in this example to estimate the relative differences in the distribution of incidents for the two values of CITYCENTRE.

One way of measuring this is to consider the conditional probability that an incident takes place inside the city centre, given that it occurred at time h. This is a function of h, say $/(h)$, defined by

$$/(h) = \frac{p_I f_I(h)}{p_I f_I(h) + p_O f_O(h)} \quad (15)$$

where p_I and p_O are the respective probabilities of an incident occurring inside or outside of the city centre regardless of time of day, and f_I and f_O are the respective probability

Fig. 12. Location of Cardiff city centre.

Table 4
Numbers of incidents broken down by type and location (percentage of incident type in brackets)

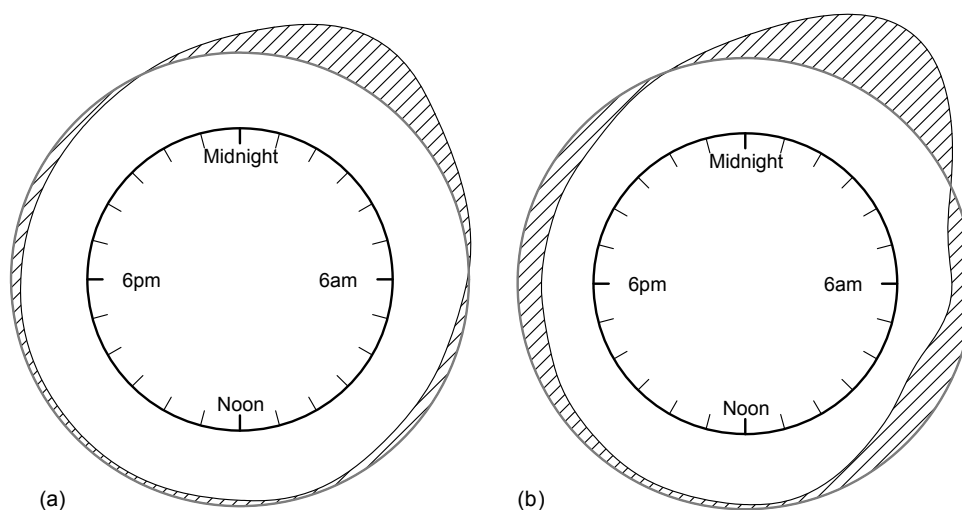| In city centre	Incident type	
	Criminal damage	Disorder/disturbance
Yes	585 (5.6)	2143 (16.2)
No	9774 (94.4)	11,098 (83.8)

Fig. 13. Estimates of / (h) for criminal damage (LHS) and disorder/disturbance (RHS).

density functions for h conditioned on the incident occurring inside or outside of the city. Clearly, / (h) can take values in the range $[0, 1]$ with values exceeding 0.5 meaning that an event at time h is more likely to occur inside the city centre zone.

We may estimate / (h) by using KDEs to estimate f_I and f_O, and obtaining estimates of p_O and p_I from Table 4. This is done for both criminal damage and disorder/disturbance data sets in Fig. 13.

In each diagram, the reference circle corresponds to p_I, the overall probability of an incident being within the city centre, so when the curve breaks outside of this circle, there is a greater than average probability that an event occurs within the city centre. Both curves tell a similar story, with a greater likelihood of incidents in the city centre between around 2300 and 0400. The start of this time period corresponds to the time when some bars close—the end of the period most likely corresponds to a time when the last few people remain in the city centre after the clubs have closed for the night. The effect for criminal damage is less marked than that for disorder/disturbance. One plausible reason for this is that disorder incidents tend to be reported very rapidly, there tends to be a greater time lag between the incident and time of reporting for criminal damage, which is sometimes not reported until the following morning, when workers or others return to the site of the incident and discover the damage. This has the effect of shifting some of the incident times from night-time to the next morning, and possibly reduces the day/night contrast of / (h) when it is estimated from the time that incidents are reported.

5. Computational issues

Since the statistical techniques used here are not readily available in 'off-the-shelf' statistical packages such as SAS or SPSS, much of the work here was done using the R statistical programming language (Ihaka and Gentleman, 1996). This is the case both for the numerical computation and the production of graphics. The Monte-Carlo tests

are fairly computationally intensive, mainly because for each of the random permutations of data, several kernel density estimates must be produced.

For this particular data set, we took advantage of the fact that data was grouped into one-hourly intervals. If there are m events occurring at say 0100, then it is only necessary to compute the value of $\log(f_j(h)/f(h))$ once, and multiply it by m, rather than computing m separate terms when calculating the value of s. Using this approach, a 2.4-GHz PC calculated 10,000 simulations of s in around 14 min for the time-of-day data. In the future it is planned to write part of this routine in FORTRAN, compile this as a dynamically linked library (DLL)[4] and link this to the R code. The current R is translated into machine instructions each time the program is run (*interpreted*), and it is hoped that pre-compiling the FORTRAN code will lead to a speed improvement.

6. Concluding discussion

The techniques outlined in this paper are useful in identifying temporal patterns in crime and policing incidents. The s-test and the other tests outlined in the paper provide a formal mechanism for testing hypotheses relating to patterns for different incident data sets, whilst the modelling and calibration of probability density functions also provides a visualisation tool to investigate daily and weekly patterns of incidence, and to compare these patterns for different subgroups of the data. These approaches could be used as a tool to aid police resource management. For example, there is some discussion as to whether changes in licensing hours of bars may reduce the incidence of disorder offences (see for example Royal College of Psychiatrists, 2003). If such a scheme underwent a trial period methods such as these could be used to identify any changes in the time patterns of disorder/disturbance incidents that might result.

There is perhaps some discrepancy between the true time of offence, and the time the offence was reported (particularly for the criminal damage data), but even in this case the analysis provides some useful information. Knowing the times of day certain times of incident are likely to be reported gives some idea about the nature and amount of demand for police work throughout the day.

There are a number of ways in which the ideas outlined here could be extended. Firstly, although the s-tests here were used to investigate association between a categorical variable and h, effectively by sub-grouping the observations by category, a similar methodology could be used to investigate the linkage between a continuous variable and h. For example if hourly rainfall data were available, one could investigate the relationship between this and incident rates.

Also, the geographical aspect could be considered in more detail. In a sense, the methods here complement existing work on identifying hotspots of crime on a geographical map—see for example Ratcliffe and McCullagh (1999). As well as a testing framework, the visual aspects of the method outlined here provide a way of identifying hotspots on a 24-h clock. We do also consider geographical patterns in a simple way when we identify a pattern difference between the city centre area and the remainder of the Cardiff study area, but at this stage we omit the geographical detail that map-based hotspot identification offers. However, we believe there is scope to combine the two kinds of hotspot analysis—further work is currently

---

[4] Equivalently a shared object file in Unix.

in progress to model and visualise space-time probability density functions for incidence, where location is treated as a point in two-dimensional space rather than a simple dichotomous variable.

References

Bowman, A. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika, 71*, 353–360.

Bowman, A., & Azzalini, A. (1997). *Applied smoothing techniques for data analysis: The kernel approach with S-Plus illustrations*. Oxford: Oxford University Press.

Braga, A. A., Weisburd, D., Waring, E., Mazerolle, L., Spelman, W., & Gajewski, F. (1999). Problem-oriented policing in violent crime places: a randomized controlled experiment. *Criminology, 37*, 541–580.

Brantingham, P., & Brantingham, P. (1981). *Environmental criminology*. Beverley Hills, CA: Sage Publications.

Brantingham, P., & Brantingham, P. (1993). Environment, routine and situation: Toward a pattern theory of crime. In R. Clarke & M. Felson (Eds.). *Routine activity and rational choice: Advantages in criminological theory* (Vol. 5, pp. 259–294). Transaction Publishers.

Chatfield, C. (2003). *The analysis of time series: An introduction* (6th ed.). Boca Raton, FL: CRC Press.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: a routine activity approach. *American Sociological Review, 44*, 588–608.

Fisher, N. I. (1993). *Statistical analysis of circular data*. New York: Cambridge University Press.

Hogg, R. V. (1979). Statistical robustness: one view of its use in applications today. *American Statistician, 33*, 108–116.

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics, 5*, 299–314.

Jammalamadaka, S. R., & SenGupta, A. (2001). *Topics in circular statistics*. Singapore: World Scientific Press.

Joe, H. (1989). Relative entropy measures of multivariate dependance. *Journal of the American Statistical Association, 84*, 157–164.

Kent, J. T. (1983). Information gain and a general measure of correlation. *Biometrika, 70*, 163–173.

Mardia, K. V. (1975). Statistics of directional data (with discussion). *Journal of the Royal Statistical Society (B), 37*, 349–393.

Mardia, K. V., & Jupp, P. E. (1999). *Directional statistics*. Germany: Wiley.

Murrey, A., McGuffog, J., Western, J., & Mullins, P. (2001). Exploratory data analysis techniques for examining urban crime. *British Journal of Criminology, 41*, 309–329.

Ratcliffe, J. H., & McCullagh, M. J. (1998a). Aoristic crime analysis. *International Journal of Geographical Information Science, 12*, 751–764.

Ratcliffe, J. H., & McCullagh, M. J. (1998b). Identifying repeat victimisation with a GIS. *British Journal of Criminology, 38*, 651–662.

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

Rayleigh, L. (1905). The problem of random walk. *Nature, 72*, 318.

Roncek, D., & Maier, P. (1991). Bars, blocks and crimes revisited: linking the theory of routine activities to the empiricism of "hot spots". *Criminology, 29*, 725–753.

Royal College of Psychiatrists (2003). National Alcohol Harm Reduction Strategy Consultation Document from the Strategy Unit/Department of Health. Available from http://www.rcpsych.ac.uk/college/parliament/responses/nahrs0103.htm

Sheather, S. J., & Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society (B), 53*, 683–690.

Sherman, L. W., Gartin, P., & Buerger, M. (1989). Hot spots of predatory crime: routine activities and the criminology of place. *Criminology, 27*, 27–55.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Stone, M. A. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society (Series B), 36*, 111–147.

Van Koppen, P., & Jansen, W. (1999). The time to rob: variations in time of number of commercial robberies. *Journal of Research in Crime and Delinquency, 36*, 7–29.

Watson, G. S. (1983). *Statistics on spheres*. New York: Wiley.