Article:

Birks, D. orcid.org/0000-0003-3055-7398, Townsley, M. and Hart, T. (2023) Predictive policing in an Australian context: assessing viability and utility. Trends and Issues in Crime and Criminal Justice (666). ISSN 0817-8542

https://doi.org/10.52922/ti78870

Australian Government
Australian Institute of Criminology

Trends & issues in crime and criminal justice

No. 666

Abstract | Studies in the United States and Europe have demonstrated that burglary and vehicle crime exhibit consistent patterns, supporting the application of crime prediction techniques to proactively deploy police resources to reduce incidents of crime. Research into whether these techniques are applicable in an Australian context is currently limited.

Using crime data from the Queensland Police Service, this study assessed the presence of spatio-temporal patterns in burglary, theft of motor vehicle and theft from motor vehicle offences in three distinct local government areas. After establishing the presence of spatio-temporal clustering, the forecasting performance of two predictive algorithms and a retrospective crime mapping technique was evaluated.

Forecasting performance varied across study regions; however, the prediction algorithms performed as well as or better than the retrospective method, while using less data. The next step in evaluating predictive policing within Australia is to consider and design effective tactical responses to prevent crime based on the forecast locations and identified patterns.

Predictive policing in an Australian context: Assessing viability and utility

Daniel Birks, Michael Townsley and Timothy Hart

Research consistently shows that crime is not uniformly or randomly distributed in space or time. In particular, patterns of several crime types, including burglary and vehicle crime, have consistently been shown to be spatio-temporally concentrated. Such phenomena have been observed in multiple western nations (Johnson et al. 2007). A direct consequence of these findings is that the spatial patterning of crime events can be used to prospectively forecast crime risk—identifying locations and times where there is an increased risk of future crime incidents occurring. Due to the finite nature of crime reduction resources, harnessing this observation may offer the means to increase the efficiency and effectiveness of resource deployment strategies. This may take many forms, but typically involves the deployment of short-term police resources (such as vehicle or foot patrols) to high-risk areas in the hope of preventing future victimisation (Weisburd & Majmundar 2018). It is this notion of short-term elevated crime risk that underpins a significant proportion of the rapidly expanding field of applied criminological research often referred to as predictive policing.

CRIMINOLOGY RESEARCH GRANT

This study assesses the effectiveness of short-term spatio-temporal crime prediction algorithms in three distinct Australian communities. In doing so, it presents, to the best of the authors' knowledge, the first large-scale academic study designed to evaluate the likely effectiveness of these techniques in an Australian context.

Background

Police and researchers have mapped crimes for centuries, consistently finding that crime clusters in particular areas, which are commonly referred to as crime 'hot spots'. Understanding where and when crime clusters provides vital information for designing and deploying effective crime reduction strategies. However, until recently, many researchers and police analysts have operated retrospectively—using crime incident location information to simply describe historical crime patterns through the use of data visualisation techniques or to design reactive policing strategies based on the assumption that previous crime patterns are reliable indicators of future problem areas. Unlike traditional crime hot spot analysis, predictive policing seeks to forecast the risk of future crime patterns, rather than simply summarising the past.

Defining predictive policing

In the current study, we define predictive policing as follows:

> The use of dynamic prediction models that apply spatio-temporal algorithms to core business data supplemented by secondary data sources, including internal corporate data and external environmental and socio-economic data, with the purpose of forecasting areas and times of increased crime risk, which could be targeted by law enforcement agencies with associated prevention strategies designed to mitigate those risks.

Similar to many other definitions of predictive policing, our definition distinguishes it from traditional crime hot spot mapping techniques because it emphasises the importance of: (1) dynamic prediction models that rely on an explicitly defined algorithm; (2) near-term forecasts of elevated crime risk; and (3) the use of this information to inform crime reduction strategies.

Predicting the locations and times of elevated crime risk is of great interest to law enforcement agencies around the world. Although predictive policing is a relatively new way to analyse crime data, several methods and techniques designed to forecast crime risk exist. These include crime hot spot mapping, regression methods, near-repeat analysis, spatio-temporal analysis, cluster analysis and risk terrain modelling (see Perry et al. 2013 for a comprehensive review). Regardless of method, the ultimate goal of predictive policing is to support a proactive approach to crime prevention. Traditional hot spot methods and the ways in which they differ from modern predictive policing techniques, including those applied in the current study, are described in the next section.

Retrospective hot spot analysis

Analysts and academics from around the world study crime event data to identify crime hot spots and to convey these patterns visually, in the form of hot spot maps. Crime hot spots are areas of concentrated incident locations that demonstrate a non-random pattern in space or time. Depending on the size of the geographic area being studied (eg address locations, streets, blocks, suburbs, jurisdictions etc) and the particular research question at hand, different types of hot spot methods can be employed. To this end, a variety of techniques have emerged from the academic literature (Bowers, Johnson & Pease 2004; Chainey & Ratcliffe 2005; Eck et al. 2005) and can be grouped into two general categories: methods based on aggregated incident locations and analysis of crime event point patterns.

Crime hot spot techniques that rely on aggregated crime counts include grid-based thematic mapping and local tests of spatial association (Anselin 1995; Getis & Ord 1992; Ord & Getis 1995). Aggregate tests of statistical association are used to identify patterns in crime incident data based on historical information and to determine whether clusters of incidents are arranged in significant, non-random patterns across a study area.

An alternative approach to crime hot spot mapping involves point-pattern analysis (Hartigan 1975; McBratney & deBruijter 1992; Spring & Block 1989). These techniques are also referred to as adaptive scanning methods. Like aggregate methods for determining spatial associations, point-pattern techniques share a common approach: identifying spatially clustered discrete crime event locations, based on certain input parameters used to conceptualise a crime hot spot. Resulting maps produced from these methods visually depict concentrations of previous crime—hot spots—as geometrically defined areas such as ellipses and convex hulls.

Prospective hot spot methods

Many traditional crime hot spot techniques use crime incident location information to describe retrospective crime patterns. However, other methods use historical data to produce prospective hot spot maps, based on the assumption that past crime hot spots are reliable indicators of future problem areas.

This assumption is underpinned by the observation that crime is often spatio-temporally concentrated. This phenomenon, commonly referred to as near-repeat victimisation, was first observed by Townsley, Homel and Chaseling (2003), who demonstrated how the chance of a residential burglary can more than double after an initial burglary and that this elevated risk extends to nearby locations, with this increase in risk decaying over time. Because of this patterning some scholars have likened victimisation risk to that of a contagion, whereby future victimisation risk has a spatial and temporal component arising from the reference victimisation.

Drawing on traditional repeat victimisation literature, several theoretical explanations for the presence of such patterns exist. While there are distinctions between explanations, they draw heavily on the notion that things that are spatially proximate are often functionally so; thus an advantage in offending at a particular location or property is proposed to be indicative of similar advantages nearby.

Although these underlying mechanisms may differ, the fundamental implication of the near-repeat phenomenon remains the same: that an understanding of the spatio-temporal patterns of previous crime can be harnessed to inform short-term forecasts of crime risk in the future, which in turn permit the implementation of policing strategies to prospectively address problem areas or times. Findings such as these have been consistently observed across different crime types (eg robbery, aggravated assaults, vehicle theft) and in several different countries (eg in Australia, the United States and Europe).

The first algorithm we use to test predictive policing is promap ('prospective mapping') which is based on the near-repeat phenomenon (Johnson, Birks et al. 2007). The prospective mapping method generates risk intensity values based on the spatial distribution of historical crime events, weighted by recency. A detailed description of this method is provided in later sections. The second algorithm we use was developed by YongJie Lee (Lee, O & Eck 2020) and combines two criminological theoretical approaches (population heterogeneity and state dependence) to determine short-term heightened crime risks in historically predictable hot spot locations. A more detailed description is also provided in a later section.

Predictive policing

The use of predictive policing algorithms represents a shift away from the retrospective crime mapping techniques described previously. A number of research studies have evaluated the crime forecasting accuracy of different algorithms and techniques in Europe and the United States (Gorr & Lee 2015; Milic et al. 2019; Mohler & Porter 2018; Mohler et al. 2020). Experimental evaluations of the crime reduction effects of predictive policing have been less frequent. A robust evaluation by Mohler and colleagues (2015) demonstrated an average 7.4 percent reduction in multiple crime types (burglary, vehicle theft, theft from vehicle) as a result of directing patrols to areas identified as being at an increased risk based on a predictive algorithm. In more targeted applications, reductions as large as 26 percent (in residential burglary) have been observed as a result of adopting predictive policing practices (Fielding & Jones 2012). Recently, a randomised experiment found that directed patrols by marked police vehicles in prediction locations reduced expected robbery crime counts by 31 percent, with a persistent diffusion effect for the eight hours subsequent to this also reducing expected offences by 40 percent (Ratcliffe et al. 2021).

Despite these promising results, a joint report on proactive policing by the US National Academies of Sciences, Engineering and Medicine recently concluded that 'there are insufficient robust empirical studies to draw any firm conclusion about either the efficacy of crime-prediction software or the effectiveness of any associated police operational tactics. Furthermore, it is as yet unclear whether predictive policing is substantively different from hot spots policing' (Weisburd & Majmundar 2018: 132).

It should be noted that the use of predictive algorithms in policing has also raised ethical concerns, with some commentators suggesting the implementation can have unintended negative consequences (Prins & Reich 2018; Ugwudike 2020; Završnik 2021). The use of algorithms in predictive policing falls under two main categories: individual-based risk assessments and location-based crime predictions. When using algorithms for location-based crime forecasting (as in the current study) researchers and practitioners need to be mindful of the potential for feedback loops that can result in the over-policing of particular areas (O'Donnell 2019; Richardson et al. 2019). This can arise if the location predictions are used as a blanket direction to target an area, rather than tailoring tactical responses to individual high-risk locations.

To date, there has been limited work in this area in Australia. In part, this is a result of the difficulties typically associated with researchers gaining access to recorded crime data at a sufficiently precise resolution to permit the application of predictive analytics. As a result, current knowledge concerning the effectiveness of predictive policing in Australia relies on the transferability of previous international findings, which may not be generalisable to other regions.

This study set out to determine the feasibility of the forecasting stage of predictive policing in Australia. Using nine years of recorded crime data, we used the first three years as a training dataset to develop parameters for a prospective algorithm. The final six years of data were used to test our prospective algorithm and a population heterogeneity and state dependence algorithm, along with a retrospective approach that mimicked conventional operational practice. Our results suggest that the two predictive algorithms were able to forecast future crime at higher rates (or equivalent rates with less data) than the retrospective technique across the offence types analysed. Consequently, we argue that predictive crime mapping shows promise for adoption in Australia.

Data and methods

To test the feasibility of crime forecasting in predictive policing, we took a two-step approach. First, we identified whether spatio-temporal clustering was present in three major offence categories: residential burglary, theft of motor vehicle (TOMV), and theft from motor vehicle (TFMV). The second step involved using the parameters of the spatio-temporal clustering to forecast the locations of future crime in an out-of-sample time window. A brief overview of the first step (which is covered in numerous other studies) will be given here, and the second step—applying prospective algorithms in an Australian context—will be described in detail.

This study only examines the forecasting stage of the predictive policing model; as such, it was important to minimise threats to external validity when we moved from desktop analysis to real-world resource deployment (the second stage). To do this we privileged pragmatic and generalisable approaches likely to mirror the decision-making of operational analysts in real-world applications.

Study regions

We used three study regions for this evaluation, all in Queensland: Logan, Brisbane and Townsville local government areas (LGAs). By using a range of urban contexts, we hope our results provide more generalisable insights than those based on a single study region. This choice of region provided a range of contrasting communities with which to explore the main research questions.

Logan is a city to the south of the state's capital, Brisbane, and is the smallest (by area) of the three study areas. Logan has a large number of environmental and recreational parks, including bushland reserves and wetlands. The majority of residential and commercial land use is located in the north of the study region.

Brisbane is a large metropolitan city, the capital of the state of Queensland. It covers a larger area than Logan but less than half that of Townsville, and it has by far the highest population density of the three areas.

Townsville is a regional community and covers the largest area of the three study regions while having the smallest population. Similar to Logan, the residential and commercial land use is spatially concentrated in a small part of the study region.

The median ages in all regions are similar and just a few years younger than the state-wide median age of 37 years. Queensland is a large, sparsely populated state, with the population concentrated in the major cities. Population density differs considerably between the three study areas, with Brisbane having almost three times the density of Logan, and nearly 18 times the density of Townsville. The unemployment rate of Logan and Townsville was, at the time of writing, just below 9 percent, while Brisbane had the lowest unemployment rate at almost 7 percent. Median income was roughly equivalent between Brisbane and Townsville, with Townsville's a fraction higher, while Logan had the lowest median income of the three (Australian Bureau of Statistics nd). Comparing the study areas by SEIFA (Socio-Economic Indexes for Areas), an index produced by the ABS to compare the socio-economic advantage and disadvantage for areas/communities, a little over a third of Logan's population was in the lowest SEIFA quintile, while less than 5 percent of Brisbane's population was in this range. Townsville is about midway between these two other areas on this measure (Queensland Government Statistician's Office 2018). Townsville and Logan had fairly similar crime rates, at 6,412 and 5,997 offences per 100,000 people respectively. Both these study areas had a higher crime rate than the overall Queensland rate of 4,706 (per 100,000 people), while Brisbane had the lowest rate at 4,417 per 100,000 people (Queensland Government Statistician's Office nd; Queensland Police Service 2017). Based on these statistics, Logan has the lowest socio-economic status and Brisbane the highest.

Crime data

Crime offences recorded by the Queensland Police Service for the years 2008 to 2016 (inclusive) were used in this study. Australian and New Zealand Standard Offence Classification (ANZSOC) codes were used to select the sample, which comprised all offences coded 07 (unlawful entry with intent/break and enter, burglary), 0811 and 0812 (theft of motor vehicle), and 0836 (theft from motor vehicle). The fields in the dataset include offence start and end times and dates, and the spatial coordinates, street address (in the case of apartment buildings) and clearance status/detection of the offences.

Data entry errors in the spatial coordinates were found for a small number of incidents. For instance, some incidents had extremely large latitude values (corresponding to no known location on Earth), some had positive values for latitude (positioning them in the Northern Hemisphere) and some appeared to have truncated coordinates ('rounded' to an even number). Another group of incidents had no spatial coordinates, but this was expected, as problematic or incomplete addresses often frustrate geocoding. A final check was made by confirming that all remaining points were located within one of the study region boundaries. The remaining incidents were split into three main offence categories (see Table 1).

Table 1: Crime counts for each crime classification across each of the study regions (2008–2016)

	Logan	Brisbane	Townsville
Burglary	20,230	64,483	14,821
Theft of motor vehicle	11,367	24,518	6,108
Theft from motor vehicle	20,982	50,437	12,988

As shown in Table 1, Brisbane had the highest crime counts for all three of the classifications, more than double those of Logan, while Townsville had the lowest counts. These differences largely reflect the population differences between the three regions. Annual crime trends (results not shown) were generally stable across each of the study regions, especially for TOMV and TFMV offences. Burglary offences in Brisbane demonstrated a noticeable decline from 2013 to 2016, but aside from this the crime counts in 2016 were similar to the counts in 2008.

As noted earlier, the dataset was split into two groups, a training and a test set. All incidents in the 2008–10 calendar years formed the training set, with data from 2011 to 2016 comprising the test set. Once the space-time parameters were established using the training set, all forecasts were made with the test set. This 'hold-out' sample approach is commonplace in forecasting research and ensures that evaluations of predictive accuracy best reflect how forecasts could be used in real-world applications.

Presence of spatio-temporal clustering

The first step in our analysis was to determine if spatio-temporal clustering was present in the crime data. To do this, a series of Knox tests were performed. These are described extensively in other studies (Johnson, Bernasco et al. 2007; Townsley, Homel & Chaseling 2003; Townsley & Oliveira 2015). Knox tests were run on the training set for each crime type and study area.

A full description of the results cannot be presented due to space restrictions; however, we provide a summary of the results in Table 2. Our analysis suggests that for all offence type and study region combinations, crime events do cluster spatio-temporally. The spatial parameters demonstrated a high level of stability; however, the temporal parameter varied across the training period. As such, we updated the forecasting parameters each year using the Knox results for the previous year, a decision that reflects police analyst practice.

Table 2: Summary of crime risk parameters for each crime/study region combination

Region	Crime type	Year	Time (weeks)	Space (metres)
Logan	Burglary	2008	3	200
		2009	1	200
		2010	1	200
	TFMV	2008[a]	2	200
		2009	1	200
		2010[a]	1	200
	TOMV	2008	3	200
		2009[a]	3–4	200
		2010[a]	7	1,000
Brisbane	Burglary	2008	2	200
		2009	1	200
		2010	2	200
	TFMV	2008	1	200
		2009	2	200
		2010	2	200
	TOMV	2008[a]	2–3	200
		2009	2	200
		2010[a]	1–2	200
Townsville	Burglary	2008	2	200
		2009	2	200
		2010	1	200
	TFMV	2008	3	200
		2009	2	200
		2010	4	200
	TOMV	2008	3	600
		2009	1	200
		2010	2	600

a: Indicates years in which parameters are based on lower concentrations of pairs

The next step was to determine how accurate the forecasts of future crime risk areas were, using these parameters. Here, the test dataset (calendar years 2011 to 2016) was used to evaluate the predictive power of the parameters established using the training dataset.

Forecasting future crime events

To assess how predictable crime risk is, we implemented three different methods of forecasting: a conventional retrospective kernel density estimation (KDE) surface, a prospective KDE surface, and a population heterogeneity and state dependence algorithm (Lee, O & Eck 2020). KDE is a widely used technique that provides an intuitive summary of a spatial distribution. It is a type of spatial smoothing technique used to estimate the probability surface of a population of events. In practical terms, the KDE comprises a grid of cells overlaid onto a study region. Intensity values for each cell are computed based on the amount of point data (crime events, in our case) proximate to each cell. Thus, areas with high counts of crime will have larger intensity values.

To make comparisons, we implemented two different KDE surfaces: a retrospective KDE surface and a prospective KDE surface. The retrospective KDE is merely the conventional KDE computation. It summarises a set of historical crime data into intensity values of cells according to the spatial distribution of the data. The prospective KDE surface is very similar, but data are also weighted by recency: events that occur in the recent past are weighted more than those taking place earlier in the time window. In contrast, the conventional KDE gives each point equal weight in generating the surface.

The logic is that if spatio-temporal patterns exist, then crimes occurring more recently contain more predictive information than crimes in the distant past. Our prospective KDE surface exploits this temporal patterning of crime events; the retrospective KDE surface ignores it.

The final forecasting method used was a population heterogeneity and state dependence algorithm developed by Lee, O and Eck (2020), which we will refer to as the YJL algorithm after its key inventor, YongJie Lee. This algorithm also uses a grid of cells overlaid onto a study region and operates as two models working in concert: a population heterogeneity model, which identifies the consistently high-risk crime locations, and a state dependence model, which identifies the short-term fluctuations in crime occurrence at these consistently high-risk locations (Lee, O & Eck 2020). In the population heterogeneity component, the algorithm ranks grid cells by the average true-positive rate across the previous 52 weeks of crime data. A true-positive is when the calculated probability of a crime occurring in that cell is greater than zero, and a crime is recorded in that period. The state dependence component then performs a sorting process as a second step based on the number of recorded crimes in the most recent week for each grid cell. In combination, these two processes identify the most recent crime locations among the most consistently predictable high-risk locations to forecast the most probable crime hot spots (Lee, O & Eck 2020).

Forecasting future crime

As mentioned earlier, we partitioned our time window into training and test sets. The former was used to determine the parameters of the prospective KDE surface. The test set (2011–2016) was used to evaluate the forecasting performance of the retrospective and prospective KDE techniques, and the YJL algorithm.

To assess forecasting performance, we created a regular grid of square cells 140 metres by 140 metres overlaid onto the study regions to represent the forecast risk surface. The cell size was retained across all regions to allow clearer comparisons of forecasting performance. We then chose spatial and temporal bandwidths for the algorithms, determining the amount of information used to estimate the probability surfaces. The prospective algorithm simply used the critical values determined by the previous Knox analyses. For the retrospective algorithm, we followed the advice of Williamson et al. (1999), who suggest that the average nearest neighbour distance should be used to inform the spatial bandwidth selection. The temporal bandwidth was set at three months of data, which in our experience is representative of typical analyst practice. The KDE methods both required decay parameters for weighting the spatial proximity to the cell midpoint, and the prospective method also required a temporal decay parameter. We chose a linear inverse distance weighting method for both, with events occurring in close proximity (spatially and temporally) to the reference cell midpoint or time point weighted at one, and events occurring at or beyond the respective bandwidth values receiving a weighting of zero.

To assess forecasting accuracy we created 313 forecasting periods, one for each week of the test dataset. We then applied the KDE algorithms to produce a retrospective risk surface and a prospective risk surface for each period, while the YJL algorithm was applied to generate an average true-positive value for each grid cell, indicating the predictability of the cell's value. For the retrospective and prospective forecasts, we ordered the cells according to the forecast risk, with the highest intensity estimate having a rank of 1 and lower intensity scores having higher ranks. The YJL forecast uses a two-step ranking system, which we applied. First, we ranked by how predictable the cells were, and then within these ranks they were ordered by the number of crimes in the week prior to forecast. Cells with a larger number of crimes were therefore ranked higher than cells with the same level of predictability but fewer crimes in the week prior. We then counted the number of 'future' crimes located in each cell, repeating this process for all three crime types for each of the 313 forecasting periods and the three study regions.

Metrics

Three metrics were used to assess the predictive performance of the three methods: hit rate (HR), predictive accuracy index (PAI) and recapture rate index (RRI), described briefly below.

The HR is the most basic measure of predictive accuracy and is defined as the percentage of crimes that fall within zones predicted to be at high risk of events occurring. The main drawback of this metric is that it does not account for the size of the prediction area used.

The PAI was developed by Chainey and colleagues in 2008 as an attempt to address the issue of the size of predicted crime hot spots. It takes into account the HR, the size of the study area and the size of predicted hot spot areas.

The RRI attempts to incorporate historical data into the assessment of prediction accuracy (Levine 2008). This metric compares the predicted hot spot density at a particular time to the hot spot density at a previous time. It is a simple ratio, with values closer to one reflecting greater consistency or reliability.

Results

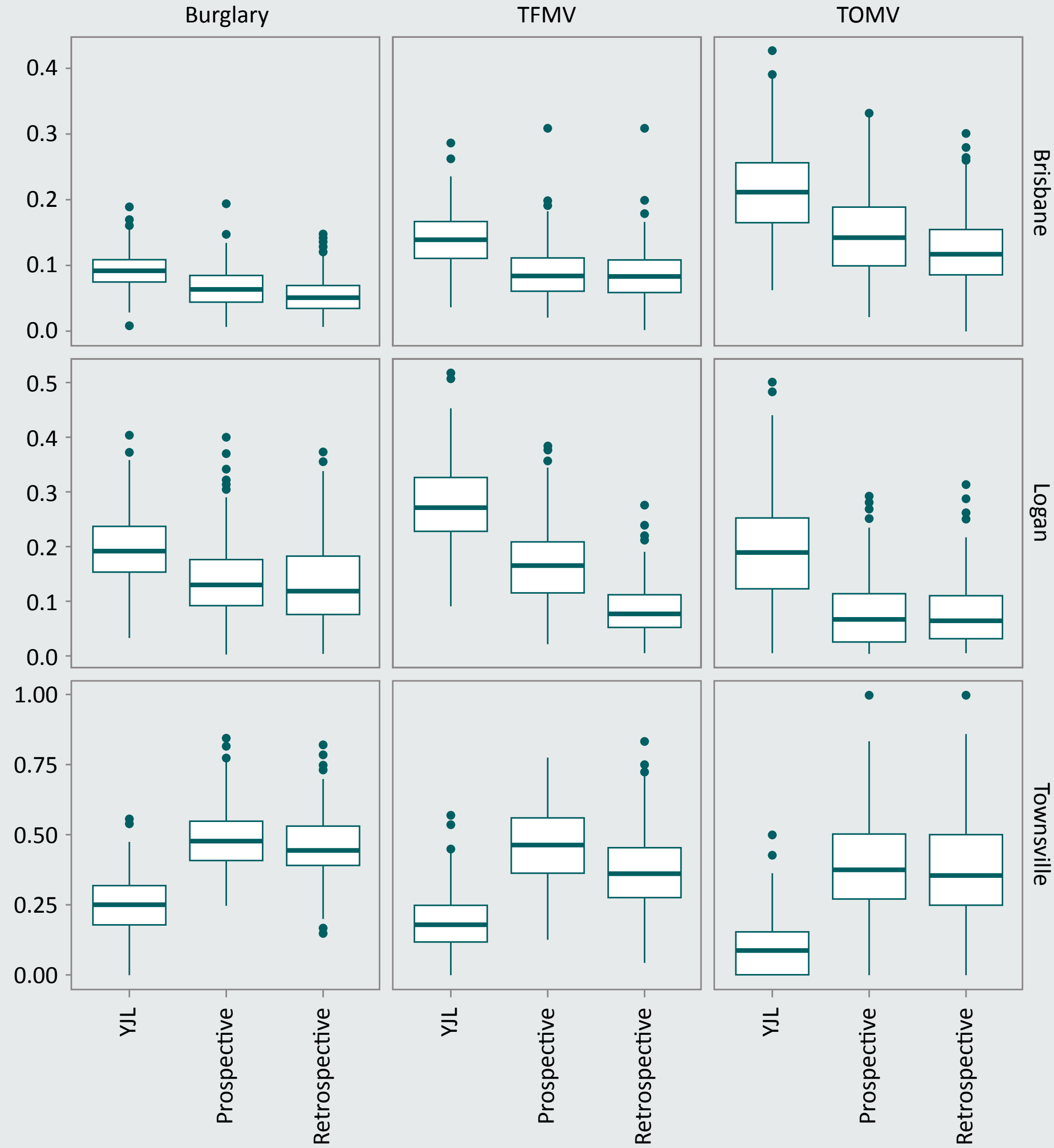
We have provided aggregate results below, with results displayed across the test period for each crime type, study region and algorithm combination. We can confirm we observed no instances of a varying or seasonal performance during the test period.

The HR, PAI and RRI metrics across the test period indicate that the forecasting methods differed in their performance depending on the region. The boxplots shown in Figures 1 and 2 summarise the weekly performance for all three metrics described above. The results are based on selecting the top one percent of grid cells by risk value for the three forecasting algorithms. As a result of this selection, and due to the way the PAI is computed, the HR and PAI metrics are the same if the 'high-risk' areas are the same. Thus we only display the results of the HR but discuss both.

The YJL model outperformed the prospective and retrospective models across all crime types in Brisbane and Logan. In contrast, the YJL model had the lowest forecasting performance in Townsville. This may be due to the lower count of each crime type in Townsville, the level of spatial concentration of offences, how the YJL algorithm is calculated, or a combination of these. This algorithm is based on 'hits': that is, crimes occurring within a grid cell that returned a non-zero probability of experiencing a crime. If there is a low count of recorded crimes, or these are highly concentrated in a small number of grids, then the number of potential cells the model can predict is limited. The prospective and retrospective models had similar performance in the majority of comparisons; however, when there was a difference this was in favour of the prospective method. It is worth noting that the prospective method used a shorter lookback period than the retrospective model, producing equal (or better) forecasting performance from a smaller pool of data.

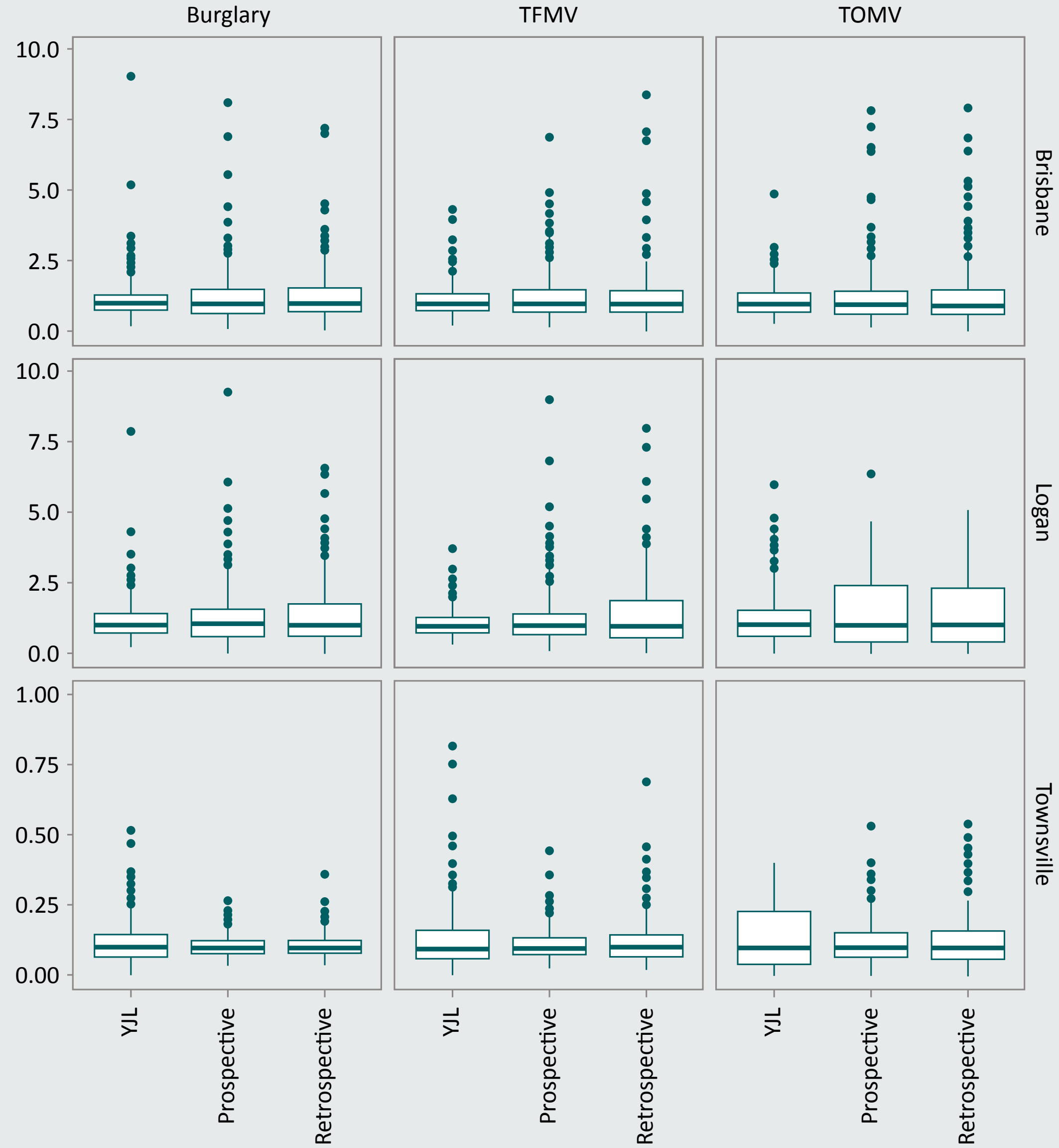
Townsville demonstrated the highest median HR and PAI for the retrospective and prospective methods across each of the crime types, despite having the lowest crime counts (see Figure 1). In contrast, Brisbane, which had the highest crime counts, demonstrated the lowest forecasting performance for burglary offences across all forecast models, and for TFMV offences using the YJL and prospective methods. This contrast between the crime volume and forecasting performance may relate to the spatial concentration of offences. Brisbane demonstrated a high concentration of crimes in Brisbane city, but had a broad diffusion of lower concentrations of offending across a large area around this. Townsville had a highly concentrated pattern, with a small area of high-intensity offending and a limited 'buffer' of lower-intensity offending surrounding this. The crime concentration patterns in Logan were in between these two, with clear high-intensity hot spots of offending and a diffusion of lower-intensity offending around these areas. This suggests the importance of tailoring both the forecast parameters and forecasting methods to the location of interest.

Figure 1: Hit rates by forecasting method: Aggregate boxplots for crime type and study region (2011–2016)



An examination of aggregate RRI results (Figure 2) shows that the three forecasting methods demonstrated similar median values, suggesting that overall they are generally effective at consistently encompassing future crime locations in the forecasts. The spread of the boxplots indicates the consistency of the RRI scores across the test period. In Brisbane and Logan the YJL model had the smallest interquartile range (IQR) across the different crime types (with the exception of TOMV in Brisbane), indicating this forecast model had the greatest week-to-week consistency in forecast accuracy. In general, the prospective and retrospective methods had a similar IQR; however, the retrospective method had a greater number of outliers (indicated by the dots) at a greater range in most comparisons (there were outliers that substantially exceeded the scale provided in Figure 2). TOMV offences demonstrated the greatest spread of scores out of the three offence types, suggesting it was the least consistently predictable crime type, regardless of the region.

Figure 2: Recapture Rate Index by forecasting method: Aggregate boxplots for crime type and study region (2011–2016)



Note: Scale has been limited to assist readability

These results show all three crime types are able to be forecast with varying degrees of accuracy in Australia. However, the differing performance across locations suggests that additional factors should be taken into account based on local patterns, such as the volume and diffusion/concentration of crime.

Preventing future crimes

A field trial was outside the scope of this project, so we are unable to provide empirical evidence of the actual impact of predictive policing on crime levels if implemented in the Australian context. In this next section we attempt to sketch out the likely impact on crime by looking at the current evidence on the efficacy of different policing tactics as well as highlighting important contextual factors that are necessary for predictive policing to be effective. To do this, we draw on both the extant literature on crime prevention and published randomised controlled trials (RCTs).

Insights from meta-analysis on effective police tactics

The US Committee on Proactive Policing conducted a review of the evidence for a range of policing practices aimed at preventing or reducing crime and disorder, rather than reacting and responding to crimes once they have occurred (Weisburd & Majmundar 2018). Within the practices reviewed, three are of relevance to the current discussion of predictive policing efficacy: problem-oriented policing, hot spots policing, and predictive policing.

Problem-oriented policing (POP) uses a systematic approach to identify and address the underlying causes of crime problems (Bullock et al. 2022). This process commonly follows the SARA (scanning, analysis, response, assessment) model, guiding stakeholders through identifying a persistent crime problem, analysing and identifying the likely causes, tailoring a response to these causes, and finally assessing the effectiveness of the response (Eck & Spelman 1987). Despite the strong support of practitioners, and many published successful problem-solving case studies, the approach has not been widely adopted (Borrion et al. 2020; Weisburd & Majmundar 2018). In terms of the quality of evidence, POP studies are considered to have relatively weak experimental design but reasonably strong effect sizes (Weisburd & Majmundar 2018).

Among the techniques reviewed, Weisburd and Majmundar (2018) concluded that hot spots policing had the strongest evidence of positive crime reduction effects. This approach involves the identification of micro-geographic hot spots and the allocation of police resources to these areas. This often takes the form of directed police patrols but can incorporate a range of tactics, including POP approaches. In a meta-analysis of hot spots policing, both directed patrols and POP were found to reduce crime occurrence; however, on average, interventions that used POP practices were twice as effective as other practices (Braga, Papachristos & Hureau 2014). In addition, studies that considered displacement effects often found a diffusion of benefits—that is, crime reduction was observed in areas outside the targeted area (Weisburd & Majmundar 2018).

Compared to hot spots policing and POP, predictive policing is a more recently developed strategy with few strong empirical evaluations. So it is perhaps not surprising that evaluations of predictive policing have demonstrated mixed results, providing a weak evidence base at this time. This is interesting, as Weisburd and Majmundar (2018) noted that predictive policing interventions are often indistinguishable from those used in hot spots policing, which, as noted above, has strong evidence of crime reduction effects. This suggests the problem either resides in the forecasting process or in the implementation of tactical responses.

Insights from experimental trials

We are aware of three RCTs evaluating predictive policing published in the academic literature. The first, by Hunt, Saunders and Hollywood (2014), evaluated a predictive algorithm to identify hot spots for property crime intervention compared to a conventional crime mapping technique. The predictive algorithms did not result in greater crime reductions than the conventional approach; however, a number of implementation issues were identified. Monthly strategic planning meetings were intended to provide uniform intervention strategies for the predicted areas, but these meetings did not occur, and strategies varied across locations (Hunt, Saunders & Hollywood 2014). There was also notable slippage, with monthly resources decreasing by 60 percent across the evaluation. In addition to the implementation issues, low monthly crime counts across the study areas impacted the ability to detect positive effects (Hunt, Saunders & Hollywood 2014).

The second RCT, by Mohler and colleagues (2015), evaluated the performance of a predictive algorithm relative to best practice hot spot maps generated by crime analysts. In this evaluation, each method generated patrol mission maps for the allocation of police resources on a day-to-day basis. Once both maps were generated, one was randomly selected to be deployed for the 24-hour period. Police were not mandated to use particular tactics in their patrols. Instead, officers were simply directed to patrol within the mission areas and could adopt tactics they deemed appropriate. Results showed a significant reduction in crime on days that used the maps generated using predictive policing techniques compared to those generated by crime analysts (Mohler et al. 2015). As the tactics used while patrolling the mission areas were not specified, this suggests the difference observed was due to the improved accuracy and efficiency of the predictive algorithm over the hot spot maps generated by analysts. This further supports Weisburd and Majmundar's (2018) assertion that the tactics of predictive policing do not differ from those in hot spots policing.

In the third RCT, Ratcliffe and colleagues (2021) compared expected and actual crime counts across three experimental conditions and one control condition. Either officers were made aware of the predicted areas at the start of shift, or a marked patrol car was assigned to the predicted areas, or an unmarked patrol car was assigned to the predicted areas, or a business-as-usual approach was adopted. Marked patrol cars were associated with a substantial (but not statistically significant) reduction in expected crime counts, but no other intervention demonstrated notable reductions in crimes (Ratcliffe et al. 2021). Like the RCT conducted by Hunt, Saunders and Hollywood (2014), this RCT also suffered from implementation problems. While patrol cars were notionally dedicated to the predicted location, competing demands (such as responding to incidents outside treatment areas) meant that, overall, officers were spending at least 30 minutes of the hour within the treatment area for less than half of the eight-hour shift (Ratcliffe et al. 2021). The largest challenge noted was the infrequency of crime occurrence when forecasting across eight-hour shifts within 500 by 500 foot square grids (about 150 by 150 metres), which severely limited the ability to detect significant positive effects if these existed. This becomes a balancing act between forecasting locations at a sufficient resolution to be operationally useful, and large enough to have measurable impacts.

Contextual factors that will facilitate or hinder success

Based on the published literature, we suggest there are four conditions required if predictive policing is to be effective. The first requirement, and the focus of the current study, is that crime needs to be predictable. Our results indicate that spatio-temporal patterns that enable accurate prediction are present in the Australian context. The second requirement is a replicable process for diagnosing the problem—that is, using analysis to identify the cause(s) of the increased crime risk in the forecast area. The early stages of analytic techniques such as problem-oriented policing or intelligence-led policing (Ratcliffe 2016) are well suited to this. The third requirement is the ability to identify appropriate treatment of the underlying cause(s) and subsequent high-quality implementation. For this to be effective, it should be tailored to the specific underlying causes identified in the location, rather than adopting a generalised approach. Finally, the outcome of the treatment should be assessed to determine if it was successful, with appropriate action taken depending on direction and level of impact.

Throughout all the stages of the predictive policing enterprise, the operational time horizon should be considered. Longer time horizons and the existence of consistent high-risk locations allow for a more detailed analysis and the design of an appropriate tactical response to effectively reduce crime at those locations. In contrast, transient high-risk locations prevent such a time investment, and lower investment responses, such as short-term directed patrols, may be sufficient to reduce the occurrence of crime.

The four predictive policing stages we have described mirror the SARA model of POP. In this way, predictive policing would serve the function of automating the scanning process to identify potential crime problems for further investigation. Nonetheless, if predictive policing is implemented in this way, it is likely to face the same barriers that have hindered the widespread adoption of POP. These barriers relate to five key areas: skill requirements, available resources, establishment of partnerships, organisational structure, and perceptions of 'real' police work.

The skills required (both technical and analytical) for in-depth analysis and evaluation of response strategies are quite specialised, and often exceed those typically available within police agencies (Goldstein 2003; Scott 2003). While this can be addressed through partnerships with universities and research institutions, this is less effective than having the skillsets in-house and may reduce police staff buy-in when the allocation of police resources is seen to be directed by 'outsiders'. Problem-oriented policing is also resource intensive, requiring a long-term investment of analysts and personnel for results to be achieved. Bullock and colleagues (2022) noted this was one of the primary barriers to (or facilitators of) POP, and that ideally there should be dedicated teams with time devoted to the analysis and evaluation. In addition to the time costs, there are the financial costs associated with tactical responses, which may take away from resources normally allocated to more traditional police roles (Goldstein 2003).

Effective POP approaches often require partnerships with agencies beyond the police. For example, police data may be insufficient to fully analyse and understand the problem, requiring collaboration with other data custodians (Goldstein 2003). Similarly, partner organisations may have additional powers or resources to enact changes (Bullock et al. 2022). While this can facilitate the POP process, there is also the risk in any partnership of competing perspectives and priorities hampering effective collaboration. The final two barriers—organisational structure and perceptions of 'real' police work—are interlinked. The rank structure in police agencies is at odds with the bottom-up approach of POP, in which problems are often identified by patrol officers (Borrion et al. 2020). The use of predictive policing could offset this, with algorithms somewhat automating the identification of crime problems suitable for POP. Finally, POP can be seen by police officers as secondary to 'real' police work, which involves responding to crimes and calls for service, conducting investigations, and making arrests (Bullock et al. 2022). Any task that detracts from this central role can be seen as mission drift. Addressing these issues would require top-down implementation and buy-in from senior officers, and a culture shift in what is perceived as 'real' police work. Smaller scale implementation is likely more feasible than attempting a substantial change to police organisations, perhaps with small units addressing three or four problems a year (Knuttson 2003).

Conclusions

This study identified space-time patterns in historical crime data. From these, two methods (the prospective method and the YJL method) were used to forecast likely crime locations in the immediate future for three crime types in three distinct study regions. An alternative prediction method emulating conventional police analyst practice was used as a baseline (the retrospective method). The key finding of the study was that the three crime types were able to be predicted in the Australian context. However, our results show the importance of tailoring parameters and methods to the location of interest.

As a preliminary study, our results offer cautious optimism for the predictive component of the predictive policing approach. Disaggregating the results revealed other observations that are worth considering alongside the headline findings:

- Across the study regions, Brisbane demonstrated the most stable performance, particularly when forecasting burglary offences. Logan and Townsville had more variable performance, occasionally exhibiting volatile week-to-week accuracy, especially for TOMV. This can be seen by the range of the box and whisker plots in the earlier figures. We suggest two candidate explanations for this. First, Brisbane had the largest sample size of the study regions, providing more data points to draw from, while Logan and Townsville had smaller sample sizes and less data with which to refine predictions. Second, the forecasting horizon was reasonably short: only a week at a time. This duration was selected due to operational constraints—we believe it is better to generate forecasts for the application for which they are intended (operational deployment) than for maximising predictive fit.

- While Brisbane demonstrated the most consistent performance, it also generally demonstrated the lowest HR. The first explanation we offer for this relates to the interplay between area size, crime levels, crime opportunities and crime dispersion. Brantingham (2016) supported the idea that larger areas typically provide opportunities for more types of crime as a result of hosting more diverse environmental settings. Similarly, one would expect that a larger area within a certain environmental setting would provide more opportunities for particular types of crimes. More dispersed crime can impact the performance of distance weighted forecasting methods, and Brisbane had the highest recorded crime counts and the greatest level of spatial dispersion. A possible explanation for this may relate to the type of land use and distribution of crime opportunities. Brisbane has the largest proportion of its land area that is developed, followed by Logan, and then Townsville. This has implications for the opportunity structure for the offence types forecast, with larger urban spread creating more dispersed opportunities suitable for these crime types, while highly concentrated population centres (as in Townsville and, to a lesser extent, Logan) limit suitable crime opportunities to a smaller area.

Like any study, the analytical strategy and data used here may have created potential weaknesses in validity and interpretation. We summarise these here in the interests of objectivity. First, we necessarily relied on recorded crime data, which are subject to a host of well-known organisational filters. For the crime types considered here, incidents need to be reported for them to be recorded, and police officers rarely observe property crimes being committed. According to the most recent national victimisation survey (Australian Bureau of Statistics 2019), burglary and theft of vehicles have high reporting rates (72% and 95% respectively), a function of insurance companies requiring a police incident number to process a claim. Theft from vehicles, by contrast, has a lower reporting rate (54%). Thus, the tendency to report crime is likely skewed towards more affluent suburbs, which has implications for the distribution of police resources. This weakness might also be recast as an advantage with respect to some of the more contentious issues surrounding the application of predictive policing. Crimes reported to the police are less likely a function of police activity than those typically discovered by police (eg drug-related offences), which simulation studies have shown may lead to problematic feedback loops in over-policing particular (often disadvantaged) communities.

Second, the veracity of the reported location of the two types of vehicle crime is difficult to determine. By their nature, vehicles can be associated with many different locations, which creates challenges for recording. Vehicles may be subject to crime at residential addresses, anywhere on the street network, in commercial car parks or at recreational parks. With the exception of residential addresses, each of these reports is likely to rely on a narrative description of location, which is subject to interpretation.

Our performance metric (amount of crime occurring in predicted cells) does not incorporate the influence of police actions. It may be that the observed crime distribution was actually in part a result of police operations. Without information about police locations and operations during the test period, we are not able to definitively claim the forecasting performance is solely due to the algorithm. However, given the efficacy of most police operations (short-term, modest impact; Farrell et al. 1998) we are confident these patterns are a product of offender location choice. An RCT, such as those described by Hunt, Saunders and Hollywood (2014), Mohler et al. (2015) and Ratcliffe et al. (2021), would be able to rule out this competing explanation.

In summary, predicting the most likely location for crime occurrence in the short term seems possible, and the two prediction algorithms were able to forecast at higher rates (or equivalent rates with less data) than our idea of current approaches. Moreover, these results were achieved for multiple crime types in different study regions. The second half of the predictive policing enterprise—crime reduction—has mixed evidence of effectiveness. We only evaluate the prediction component of the predictive policing enterprise here. The next step is to consider and design effective tactical responses to preventing crimes based on the identified patterns. There is considerable commentary in the academic literature on the organisational factors that inhibit effective crime reduction and problem-solving. Police leaders should be mindful of these when considering implementing such approaches.

Acknowledgements

The authors gratefully acknowledge the use of the services and facilities of the Griffith Criminology Institute's Social Analytics Lab at Griffith University.

References

Anselin L 1995. Local indicators of spatial association—LISA. *Geographical Analysis* 27(2): 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Australian Bureau of Statistics nd. Data by region. https://dbr.abs.gov.au/

Australian Bureau of Statistics 2019. *Crime victimisation, Australia, 2017–18*. ABS cat. no. 4530.0. Canberra: ABS

Borrion H et al. 2020. The problem with crime problem-solving: Towards a second generation POP? *The British Journal of Criminology* 60(1): 219–240. https://doi.org/10.1093/bjc/azz029

Bowers KJ, Johnson SD & Pease K 2004. Prospective hot-spotting: The future of crime mapping? *The British Journal of Criminology* 44: 641–658. https://doi.org/10.1093/bjc/azh036

Braga AA, Papachristos AV & Hureau DM 2014. The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly* 31(4): 633–663. https://doi.org/10.1080/07418825.2012.673632

Brantingham PJ 2016. Crime diversity. *Criminology* 54(4): 553–586. https://doi.org/10.1111/1745-9125.12116

Bullock K et al. 2022. Problem-oriented policing in England and Wales: Barriers and facilitators. *Policing and Society* 32(9): 1087–1102. https://doi.org/10.1080/10439463.2021.2003361

Chainey S & Ratcliffe JH 2005. *GIS and crime mapping*. San Francisco, CA: Wiley. https://doi.org/10.1002/9781118685181

Chainey S, Tompson L & Uhlig S 2008. The utility of hotspot mapping for predicting patterns of crime. *Security Journal* 21: 4–28. https://doi.org/10.1057/palgrave.sj.8350066

Eck JE, Chainey S, Cameron JG, Leitner M & Wilson RE 2005. Mapping crime: Understanding hotspots. Washington DC: National Institute of Justice Special Report

Eck JE & Spelman W 1987. *Problem-solving: Problem-oriented policing in Newport News.* Washington DC: Police Executive Research Forum

Farrell G, Chenery S & Pease K 1998. *Consolidating police crackdowns: Findings from an anti-burglary project.* London: Home Office, Policing and Reducing Crime Unit, Research, Development and Statistics Directorate

Fielding M & Jones V 2012. Disrupting the optimal forager: Predictive risk mapping and domestic burglary reduction in Trafford, Greater Manchester. *International Journal of Police Science & Management* 14(1): 30–41. https://doi.org/10.1350/ijps.2012.14.1.260

Getis A & Ord JK 1992. The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24(3): 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Goldstein H 2003. On further developing problem-oriented policing. In J Knutsson (ed), *Problem-oriented policing: From innovation to mainstream*. Crime prevention studies vol 15. Devon, UK: Willan Publishing: 13–48

Gorr WL & Lee Y 2015. Early warning system for temporary crime hot spots. *Journal of Quantitative Criminology* 31: 25–47. https://doi.org/10.1007/s10940-014-9223-8

Hartigan JA 1975. *Clustering algorithms.* New York: John Wiley & Sons

Hunt P, Saunders J & Hollywood JS 2014. *Evaluation of the Shreveport Predictive Policing Experiment.* RAND Corporation. https://www.rand.org/pubs/research_reports/RR531.html

Johnson SD, Bernasco W et al. 2007. Space-time patterns of risk: A cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology* 23(3): 201–219. https://doi.org/10.1007/s10940-007-9025-3

Johnson SD, Birks DJ, McLaughlin L, Bowers KJ & Pease K 2007. *Prospective crime mapping in operational context*. Home Office online report 19/07. London, UK: Home Office

Knutsson J 2003. Introduction. In J Knutsson (ed), *Problem-oriented policing: From innovation to mainstream*. Crime prevention studies vol 15. Devon, UK: Willan Publishing: 1–12

Lee YJ, O SH & Eck JE 2020. A theory-driven algorithm for real-time crime hot spot forecasting. *Police Quarterly* 23(2): 174–201. https://doi.org/10.1177/1098611119887809

Levine N 2008. The 'hottest' part of a hotspot: Comments on 'the utility of hotspot mapping for predicting spatial patters of crime'. *Security Journal* 21: 295–302. https://doi.org/10.1057/sj.2008.5

McBratney AB & deBruijter JJ 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. *Journal of Soil Science* 43(1): 159–178. https://doi.org/10.1111/j.1365-2389.1992.tb00127.x

Milic N, Popovic B, Mijalkaovic S & Marinkovic D 2019. The influence of data classification methods on predictive accuracy of kernel density estimation hotspot maps. *International Arab Journal of Information Technology* 16(6): 1053–1062

Mohler G & Porter MD 2018. Rotational grid, PAI-maximizing crime forecasts. *Statistical Analysis and Data Mining: The ASA Data Science Journal* 11(5): 227–236. https://doi.org/10.1002/sam.11389

Mohler G, Porter M, Carter J & LaFree G 2020. Learning to rank spatio-temporal event hotspots. *Crime Science* 9(3): 1–12. https://doi.org/10.1186/s40163-020-00112-x

Mohler GO et al. (2015). Randomized control field trials of predictive policing. *Journal of the American Statistical Association* 110(512): 1399–1411. https://doi.org/10.1080/01621459.2015.1077710

O'Donnell RM 2019. Challenging racist predictive policing algorithms under the equal protection clause. *New York University Law Review* 94(3): 544–580

Ord JK & Getis A 1995. Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis* 27(4): 286–306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x

Perry WL, McInnis B, Price CC, Smith SC & Hollywood JS 2013. *Predictive policing: The role of crime forecasting in law enforcement operations*. Washington DC: Rand Corporation. https://doi.org/10.7249/RR233

Prins SJ & Reich A 2018. Can we avoid reductionism in risk reduction? *Theoretical Criminology* 22(2): 258–278. https://doi.org/10.1177/1362480617707948

Queensland Government Statistician's Office nd. Queensland Thematic Maps – Crime and Justice – General. https://statistics.qgso.qld.gov.au/qld-thematic-maps?sub=80&ser=183180

Queensland Government Statistician's Office 2018. SEIFA Index of Advantage/Disadvantage, Census 2016 – LGA. Queensland Government Statistician's Office. https://www.qgso.qld.gov.au/statistics/theme/economy/prices-indexes/seifa-socio-economic-indexes-areas

Queensland Police Service 2017. Annual Statistical Review 2016–2017. Queensland Government. https://www.police.qld.gov.au/maps-and-statistics/annual-statistics

Ratlciffe JH 2016. *Intelligence-led policing,* 2nd ed. London: Routledge. https://doi.org/10.4324/9781315717579

Ratcliffe JH et al. 2021. The Philadelphia predictive policing experiment. *Journal of Experimental Criminology* 17(1):15–41. https://doi.org/10.1007/s11292-019-09400-2

Richardson R, Schultz JM & Crawford K 2019. Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *New York University Law Review* 94: 15–55

Scott MS 2003. Getting the police to take problem-oriented policing seriously. In J Knutsson (ed), *Problem-oriented policing: From innovation to mainstream*. Crime prevention studies vol 15. Devon, UK: Willan Publishing: 49–78

Spring JV & Block CR 1989. *STAC user's manual.* Chicago, IL: Criminal Justice Information Authority

Townsley M, Homel R & Chaseling J 2003. Infectious burglaries: A test of the near repeat hypothesis. *The British Journal of Criminology* 43(1): 615–633. https://doi.org/10.1093/bjc/43.3.615

Townsley M & Oliveira A 2015. Space-time dynamics of maritime piracy. *Security Journal* 28: 217–229. https://doi.org/10.1057/sj.2012.45

Ugwudike P 2020. Digital predictive technologies in the justice system: The implications of a 'race-neutral' agenda. *Theoretical Criminology* 24(3): 482–501. https://doi.org/10.1177/1362480619896006

Weisburd D & Majmundar MK 2018. *Proactive policing: Effects on crime and communities*. Washington, DC: National Academies Press. https://doi.org/10.17226/24928

Williamson D, McLafferty S, Goldsmith V, Mollenkopf J & McGuire P 1999. A better method to smooth crime incident data. *ESRI ArcUser Magazine*. https://www.esri.com/news/arcuser/0199/crimedata.html

Završnik A 2021. Algorithmic justice: Algorithms and big data in criminal justice settings. *European Journal of Criminology* 18(5): 623–642. https://doi.org/10.1177/1477370819876762

Daniel Birks (PhD) is Associate Professor of Quantitative Policing and Crime Data Analytics at the University of Leeds.

Michael Townsley (PhD) is Professor at the School of Criminology and Criminal Justice at Griffith University.

Timothy Hart (PhD) is an Associate Professor in the Department of Criminology at the University of Tampa.

General editor, *Trends & issues in crime and criminal justice* series: Dr Rick Brown, Deputy Director, Australian Institute of Criminology. Note: *Trends & issues in crime and criminal justice* papers are peer reviewed. For a complete list and the full text of the papers in the *Trends & issues in crime and criminal justice* series, visit the AIC website: www.aic.gov.au

ISSN 1836-2206 (Online) ISBN 978 1 922478 87 0 (Online)

https://doi.org/10.52922/ti78870

GPO Box 1936
Canberra ACT 2601, Australia

Tel: 02 6268 7166