CRIMINOLOGY
RESEARCH GRANT

Assessing the viability and utility of predictive policing in Australia

Daniel Birks
Michael Townsley
Timothy Hart

Report to the Criminology Research Advisory Council
Grant: CRG 32/16–17

April 2023

This project was supported by a Criminology Research Grant.
The views expressed are the responsibility of the author and are not necessarily those of the Council.

Published by the Australian Institute of Criminology
GPO Box 1936 Canberra ACT 2601
Tel: (02) 6268 7166
Email: front.desk@aic.gov.au
Website: www.aic.gov.au/crg

ISBN 978 1 922478 98 6 (Online)

This research was supported by a Criminology Research Grant. The views expressed are those of the author and do not necessarily reflect the position of the Criminology Research Advisory Council or the Australian Government.

This report was reviewed through a double-blind peer review process.

Edited and typeset by the Australian Institute of Criminology.

Contents

Figures

Tables

Acknowledgements

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

We would also like to thank the anonymous reviewers for thoughtful and constructive feedback on an earlier version of this report.

Acronyms and abbreviations

ETAS	epidemic-type aftershock sequence (model)
HR	hit rate
IQR	interquartile range
KDE	kernel density estimation
PAI	predictive accuracy index
RRI	recapture rate index
RTM	risk terrain modelling
RTM Dx	Risk Terrain Modeling Diagnostics Utility
SEIFA	Socio-Economic Indexes for Areas
TFMV	theft from motor vehicle
TOMV	theft of motor vehicle
YJL	YongJie Lee (inventor of algorithm)

Abstract

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.

Executive summary

Understanding the space-time signature of crime and its utility for crime forecasting

Police resources are finite. Consequently, understanding where and when to best deploy those available to maximise their benefit represents an important analytical problem with significant potential benefits.

Previous research has demonstrated that several key volume crimes, including burglary and vehicle crime, exhibit consistent patterns which may be used to inform proactive deployment of crime reduction resources.

The majority of this existing research has been undertaken in the United States and Europe. While a number of public and private sector organisations have developed software for use in applied policing environments, to date little research has sought to assess the generalisability of these techniques to Australian crime problems.

In this study we seek to address this significant research gap, analysing data covering nine years of crime occurring in Brisbane, Logan and Townsville local government areas and evaluating the effectiveness of short-term crime prediction algorithms for both burglary and vehicle-related offences.

Research objectives

In order to assess the potential effectiveness of predictive policing analytics in an Australian context, the research undertaken sought to address the following primary objectives:

- Assess if spatial and temporal patterns of crime risk observed elsewhere can be seen in Australian communities.
- Establish the degree to which these patterns differ across burglary, theft of and theft from motor vehicles, and within the regions of Brisbane, Logan and Townsville in Queensland, Australia.
- Identify methods capable of harnessing these regularities in the spatio-temporal structure of crime to prospectively identify locations at a heightened risk of crime in the short-term future, and thus able to usefully inform operational policing efforts.
- Evaluate these methods using historical data, comparing their predictive accuracy to baseline/business as usual measures of resource allocation, in order to estimate the potential benefits to be gained from allocating resources using prospective rather than retrospective analytics.

Methodology

The project used Queensland Police Service recorded crime data. The data encompassed nine years of recorded crime data pertaining to all burglary (domestic and other), theft of motor vehicle and theft from motor vehicle offences occurring within Brisbane, Logan and Townsville local government areas from 2008 to 2016. This produced nine unique datasets for study.

First, basic spatial and temporal patterns were explored, including hot spots and annual, monthly and daily prevalence of crime. This analysis was replicated across each of the nine datasets (three crime types in each of the three study regions).

Second, each dataset was divided into two disparate training and test datasets. The training set (spanning 2008–10) was used to identify the spatio-temporal structure of each crime occurring in each region. The results of these analyses were then used to identify significant parameters for use in the subsequent predictive algorithm development and evaluation phase.

Third, analysing the test dataset (spanning 2011–16), we systematically evaluated the predictive accuracy of two crime prediction algorithms (prospective mapping and a population heterogeneity and state dependence algorithm (YJL model)) and a retrospective mapping technique on a weekly basis over a six-year period. Three metrics were used to evaluate the forecasts: hit rate, the predictive accuracy index and the recapture rate index. The hit rate evaluates the proportion of future crimes that were captured within the forecast risk areas, while the predictive accuracy index takes into account the hit rate and the proportion of the study area covered by the predicted hot spot area. The recapture rate index is a measure of how consistent the predictive accuracy is across successive forecast periods.

Results

The primary findings of our research are twofold. First, patterns of burglary and vehicle crime in Brisbane, Logan and Townsville demonstrated spatio-temporal clustering consistent with patterns observed in other western nations. While there was some variability between crime type and locality, all crime types in all regions displayed characteristic spatio-temporal clustering at sufficient levels to warrant the investigation of the effectiveness of predictive policing analytics.

Second, while future crime locations were able to be forecast using historic data, the methods used varied in their performance across the three study regions and crime types. When aggregating performance, the YJL model outperformed the prospective and retrospective methods in forecasting each crime type in Brisbane and Logan. When forecasting the three crime types in Townsville, the YJL model had the poorest performance of the three methods. The prospective and retrospective methods generally had similar levels of performance, although when there were differences in performance these were in favour of the prospective method, despite it having a shorter lookback period with which to generate forecasts. The explanation we offer for the differing performance outcomes across the study regions relates to the different patterns of crime concentration, with the YJL algorithm performing well with more dispersed hot spots (Brisbane and, to a lesser extent, Logan), while the prospective and retrospective methods performed well with highly clustered hot spots (Townsville).

Research and policy implications

Policing is a complex task that encompasses varied dimensions of individuals and their interactions with neighbours. While the mission of police services is collectively to prevent, deter or detect crime, a critical aspect of this is, among other things, to allocate finite resources where they are most appropriate or effective. This research attempted to contribute to a greater understanding of this facet of the policing endeavour.

In terms of prediction, we broadly demonstrated that is it possible to predict locations of future crime over a short time horizon (a week). This varied by crime type and location, with the results demonstrating the importance of tailoring parameters and methods to the location of interest.

It is possible to generate more accurate forecasts than those presented in this study. Only two predictive algorithms were used here, but several candidate alternatives exist. It is possible that other approaches might perform better on smaller sample sizes and when forecasting across different patterns of crime dispersion. Critics may argue that we are running the risk of underplaying the effectiveness of these approaches by not exploring more algorithms and/or searching for greater accuracy or model fit. This will always be an inevitable tension between specificity (having the best fitting model for a dataset) and generalisability (being able to fit any dataset). While we performed some localising of the models, there is additional scope for this, and our results suggest it is an important component to explore.

The other important aspect to consider is the practical implications of this research. The methodological approach described in this document is known as secondary data analysis. We used data already collected, ran analyses on it, and used our findings to suggest the operational impact we *might* have had. It is extremely easy, of course, to take this thinking too far. Our study is useful for the implications it suggests, but they should remain as indications of operational utility. Stronger research designs, such as experiments and quasi-experiments, provide much better insights into the efficacy of predictive policing.

We selected three crime types for this study: burglary, theft of motor vehicles and theft from motor vehicles. Burglary has been used in many studies and is a fairly predictable type of crime. Vehicle crime is less common in predictive policing, but there is considerable literature on the space-time dynamics of vehicle crime. However, both burglary and vehicle crime make up a minority of the work of the police. It seems useful to consider other crime types (or calls for service types) and assess their predictability as well as their 'police-ability'.

Four key messages for policy and practice arise from this study. First, predicting the most likely location for crime occurrence in the short term seems possible in Australia. Second, forecast parameters should be tailored to local conditions. Third, the crime reduction component of predictive policing currently has mixed evidence of effectiveness. And fourth, predictive policing as an approach has the potential to be effective, as long as there is long-term commitment to applying the techniques in a thoughtful way.

To understand how predictive policing analytics might support operational policing in Australia, further empirical research is required. However, the findings contained herein suggest that a field trial of such techniques would not be premature.

Project background

Introduction

Research consistently shows that crime is not uniformly or randomly distributed in space or time. In particular, patterns of several key offences, including burglary and vehicle crime, have consistently been shown to be spatio-temporally concentrated. This means that when an initial crime occurs the risk of a subsequent crime occurring nearby and within a relatively short time period is significantly elevated (Townsley, Homel & Chaseling 2003). This spatio-temporal structure in volume crime has been observed in multiple western nations (Johnson, Bernasco et al. 2007). Importantly, a direct consequence of this finding is that analyses of previous crime events can be used to prospectively forecast crime—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. It is this notion that underpins a significant proportion of the rapidly expanding field of applied criminological research often referred to as predictive policing.

This report summarises research that seeks to assess the effectiveness of predictive policing, and particularly of such short-term spatio-temporal crime forecasting, in three Australian communities. Using data provided by the Queensland Police Service, the research is separated into three distinct phases. The first phase analyses burglary and vehicle-related recorded crime data to establish major spatio-temporal patterns and, in turn, assess the suitability of such data for the application of predictive policing analytics. The second phase examines patterns of residential burglary and vehicle crime in three distinct Queensland regions—Brisbane, Logan and Townsville. This phase involves quantifying the spatio-temporal structure of crimes occurring in these localities and identifying the potential predictability of crimes in them, with the ultimate aim of identifying key parameters to be used in the development of crime prediction algorithms developed and tested in the next phase. In the third phase the predictive accuracy of crime prediction algorithms in each of the three regions across three key crime types is evaluated using historic data. This involves a process of ‘hindcasting’, where historic data are used to forecast crime patterns based on operationally relevant time windows (eg weeks) and evaluate their accuracy over a six-year evaluation period.

Objectives of the current research

The primary objectives of the project were as follows:

- Assess if spatial-temporal patterns of clustering observed in other countries can be seen in Australian communities.
- Establish the degree to which these patterns differ across burglary, theft of and theft from motor vehicles, and within the regions of Brisbane, Logan and Townsville in Queensland, Australia.
- Identify methods capable of harnessing these regularities in the spatio-temporal structure of crime to prospectively identify locations at a heightened risk of crime in the short-term future, and thus able to inform operational policing efforts.
- Evaluate these methods using historical data, comparing their predictive accuracy to baseline/business as usual measures of resource allocation, in order to estimate the potential benefits to be gained from allocating resources using prospective rather than retrospective analytics.

Structure of the report

The remainder of the report is structured as follows. The second section outlines key research findings that underpin the enterprise of short-term crime prediction—namely, regularities in the spatial and temporal patterning of crime and the notion of crime risk communicability. Subsequently a summary of current prospective crime mapping analytics is provided, and these techniques are contrasted with retrospective crime mapping methods. The next section outlines the key data and methods used throughout the project. This section describes the original data source provided to the researchers by the Queensland Police Service, the study regions selected, and methods devised to address the first three primary analytic objectives discussed above. This is followed by a description of the data used in this study and an outline of the exploratory analysis of major space and time patterns in the data, followed by spatio-temporal analysis. The next section begins with an explanation of the metrics of forecasting accuracy and then presents the results obtained using these metrics across the three study regions and three crime types. In the final section we summarise our findings, discuss their implications for operational policing in Australian communities, discuss several limitations of the study and outline potential avenues for further research.

Predictive policing: Background

Introduction

On 18 November 2009, the National Institute of Justice, Bureau of Justice Assistance, and the Los Angeles Police Department convened the first federal government-sponsored Predictive Policing Symposium. The event brought together academics, crime analysts, data scientists, and law enforcement executives to discuss the future of predictive policing in the United States. Participants highlighted early applications of predictive policing strategies used at the time and areas in law enforcement where it could be applied to have meaningful and measurable impact in the future. Agencies were moving towards more empirically based proactive policing strategies and predictive policing was seen by many as a promising solution to improve resource allocation and to reduce crime.

Two years later, *Time* magazine released its list of the 50 best inventions (Grossman et al. 2011). In 2011, pre-emptive policing was number 47 on the list, and was described as computer software that enabled law enforcement agencies to predict areas within a city at highest risk of future crime. According to the article, the Santa Cruz Police Department in California was testing the software and suggested the agency could use it to figure out where crime was likely to occur before it happened so that it could 'have a member of the force at the ready' (Grossman et al. 2011: np). The software described in the article was released commercially the following year as PredPol and Santa Cruz was the first city in the United States to adopt and use its predictive policing algorithm.

In the years following PredPol's release, a growing body of scholarship offered evidence supporting predictive policing techniques. Studies showed predictive policing could be an effective proactive policing strategy that could improve resource deployment and reduce crime. For example, New York Police Department's predictive policing software built around their Domain Awareness System was being used throughout the agency's jurisdiction. Levine et al. (2017) evaluated the Domain Awareness System and found that it was responsible not only for generating approximately $50 million per year in savings to the city but also for helping reduce index crimes by six percent.

Mohler et al. (2015) also assessed the benefits of their predictive policing algorithm in two law enforcement jurisdictions. They compared results of their predictive policing algorithm, based on an epidemic-type aftershock sequence (ETAS) model, to crime analysts' predictions, using randomised control field trials. Results showed that their predictive ETAS model significantly outperformed traditional methods used by analysts. Their model predicted 1.40 to 2.20 times as much crime as analysts' predictions and crime decreased more than seven percent in areas where patrolling was informed by ETAS forecasts, compared to no significant reduction in crime in areas patrolled based on analysts' predictions.

Despite the growing evidence that predictive policing could be a beneficial and effective tool, the murder of George Floyd in 2020 brought greater scrutiny by many communities of the police and the various methods and techniques they use to fight crime, including predictive policing. Commercial software applications like PredPol, which many agencies were using at the time, were built on proprietary algorithms, and as such they lacked transparency. This caused many to question whether their shortcomings outweighed their benefits. Brannon (2017), for instance, raised concerns over the racial and social inequities arising from the use of the Kansas City No Violence Alliance predictive policing software and called on agencies using predictive policing methods to rethink their approach. As a result of growing community concern over predictive policing, especially among communities of colour, in 2020 the city of Santa Cruz, California—the city first in the country to adopt the PredPol algorithm—became the first city in the United States to ban predictive policing (Sturgill 2020).

Over the past decade, predictive policing has seen a meteoric rise and fall in popularity in the United States. Despite currently being out-of-favour with many in the public and a growing number of agencies moving away from commercial software applications like PredPol, elements of predictive policing still provide the foundation for many types of effective proactive policing strategies, including prospective crime hot spot mapping, repeat/near repeat pattern analysis, and crime risk estimation. Despite a large and growing body of empirical evidence supporting these methods, however, relatively little is known about their effectiveness in an Australian context. To better understand its potential, predictive policing must first be defined.

Defining predictive policing

The definition of predictive policing has evolved over the past several years. For instance, the definition that emerged from the first Predictive Policing Symposium defined it as follows:

> Predictive policing refers to any policing strategy or tactic that develops and uses information and advanced analysis to inform forward-thinking crime prevention. (Uchida 2009: 1)

In the years that followed the symposium, alternative definitions were proposed, including those that stressed the important role that data analytics played in many predictive policing processes. For instance, the Rand Corporation defined predictive policing as:

> The application of analytic techniques—particularly quantitative techniques—to identify likely targets for police intervention and prevent crime or solve past crimes by making statistical predictions. (Perry et al. 2013: 1–2)

More recent definitions of predictive policing mention the application of predictive algorithms that leverage big data in the process of identifying crime patterns. For instance, in a recent National Academies of Sciences report, predictive policing was defined as:

> …a strategy for proactive policing that uses predictive algorithms based on combining different types of data to anticipate where and when crime might occur and to identify patterns among past criminal incidents. (National Academies of Sciences, Engineering, and Medicine 2018: 49)

Others have defined it as part of recursive data processing cycle. For instance, Mohler et al. (2018: 2454) define predictive policing as a:

> …data cycle … where input may be generated from victim reports or police initiated arrests (for example). This data enters the police database and is then used by a predictive algorithm to inform police activity and patrol. That activity may then influence future suspects and victims in the areas the algorithm selects for patrol, as well as in those areas that do not receive police attention.

Elements of these and other existing definitions of predictive policing inform the definition of predictive policing used in the current study:

> The use of dynamic prediction models that apply spatio-temporal algorithms to administrative data for 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 and/or reduce those risks.

Collectively, three elements of this definition distinguish it from others. Firstly, it emphasises the dynamic nature of crime prediction models, which rely on specialised algorithms. Secondly, it stresses the importance of time and elevated crime risk informed by the repeat/near repeat literature. And finally, it underscores the importance of predictive policing being a tool to support other crime reduction strategies instead of a tool to replace them. Despite a general lack of consistency in how predictive policing is defined, there is growing body of scholarship on how predictive policing has been used as part of a broader approach to data-driven efforts to reduce crime, including more traditional prospective crime hot spotting methods.

Models, algorithms, and mechanisms

Unlike early retrospective crime hot spotting methods, which simply used historical crime location data to map where crime occurred as an estimate of where it would likely occur in the future, prospective hot spot mapping integrated a temporal component to spatial analysis. This assumption was underpinned by the observation that crime is often spatio-temporally concentrated, a phenomenon commonly referred to as near-repeat victimisation. 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. Because of this patterning, some scholars have likened victimisation risk to that of a contagion, whereby risk of victimisation can be communicated in space and time from an initial victimisation to a future crime.

Risk communicability

Following the repeat victimisation literature that underpins the near-repeat hypothesis, two non-mutually exclusive explanations have been proposed for this communicability of risk. The first, referred to as the boost hypothesis, offers a state dependence-based mechanism for near repeats. Under the boost account, the occurrence of an initial victimisation conveys an advantage to a returning offender (or their peers) for subsequent offending. This advantage may manifest in an increased awareness of the target in question—how it might be accessed, the potential risks involved in subsequent offending and how they might be overcome—and of the rewards likely to be obtained by doing so. Following this model of offender decision-making, when choosing where to commit a subsequent crime an offender is likely to select targets on which some prior information is available. Importantly, the spatial diffusion of this advantage is supported by the notion that objects that are spatially proximate are often also functionally similar. Consequently, this prior information regarding the risks, reward and effort associated with a target extends from the initial victim to other similar targets within some short distance of the original victim. To illustrate, if your home is burgled, your neighbours' homes, for a number of systematic reasons, are likely to offer similar methods of entry and means of escape, and have similar goods within them.

A second explanation, often referred to as the flag hypothesis, offers a population heterogeneity-based mechanism for near repeats. Under the flag hypothesis, the occurrence of an initial victimisation highlights the suitability of a given target for offending not just by the same offender or their peer group, but also by other similarly motivated offenders. Thus, a target is flagged as previously deemed suitable by an initial offender and is likely to be similarly assessed by other offenders. Again, the notion of spatial similarity dictates that this process applies not only to the original target but also to others within some finite distance of it.

Though the nature and form these underlying mechanisms may manifest differently, the fundamental implication of the near-repeat phenomenon is 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 can support proactive policing strategies. For example, Hu et al. (2018) demonstrated the utility of a spatio-temporal kernel density estimation algorithm used to prospectively identify crime hotspot in Baton Rouge, Louisiana. This algorithm outperformed both traditional retrospective hot spotting methods and earlier prospective methods in forecasting the locations of residential burglary. Findings demonstrating the effectiveness of considering space *and* time in crime forecasting can be found in existing scholarship for different crime types (eg robbery, aggravated assaults, auto theft) and in different jurisdictions, including jurisdictions in the United States, Europe and Australia.

In addition to prospective hot spotting approaches, the use of predictive policing algorithms to forecast crime has emerged in recent years as an alternative approach to crime forecasting. These techniques involve identifying and forecasting crime patterns, using various data and algorithms, such as point process algorithm previously described in Mohler's (2015) study. Studies demonstrating the effectiveness of predictive algorithms can be found in the existing literature, such as Liesenfeld, Richard and Vogler's (2017) study, which used a maximum likelihood efficient importance sampling model to predict Uniform Crime Reporting index crimes across census tracts in Pittsburgh, Pennsylvania. Similarly, using data from Portland, Oregon, Mohler et al. (2011) demonstrated how their predictive point process algorithm, based on a Hawkes (1971) process model, could maximise the predictive accuracy of crime forecasts. Spatio-temporal patterns identified with algorithms such as these can be exploited directly to prevent crime by more efficiently and effectively allocating resources.

Other methods for estimating crime risk, which can be used as part of a broader predictive policing strategy, have grown in popularity in recent years. Chainey (2021) suggests that predictive policing should be reconceptualised as crime risk estimation because 'prediction' implies an unrealistic expectation of precision. There is little doubt that his recommendation is also in some way related to the general backlash against predictive policing methods that became the focus of public ire following the George Floyd murder. Regardless, one of the most popular crime risk estimation applications used to support proactive policing strategies like predictive policing is risk terrain modelling (RTM; Caplan & Kennedy 2010; Caplan, Kennedy & Miller 2011). The process of RTM is relatively straightforward. It begins by identifying the spatial influence of multiple risk factors within a specific geographic area that could be associated with the locations where crime concentrates. Then this information is used to generate separate map layers for each risk factor, created with kernel density estimation (KDE). Finally, the separate crime risk map layers are combined to produce a composite risk terrain map. Once complete, RTM maps show the presence, absence or intensity of all risk factors included in a model at every location throughout a study area. RTM has been used in several jurisdictions to reduce crime (see, for example, Buccine-Schraeder & Kennedy 2021 and Caplan et al. 2020).

Despite promising results being associated with predictive policing, the United States' National Academies of Sciences' Committee on Proactive Policing recently concluded:

> …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)

To date, relatively little work in this area has been undertaken in Australia. In part, this is a result of the difficulties typically associated with gaining access to recorded crime data at a sufficiently precise resolution to permit the application of predictive analytics. Consequently, current knowledge concerning the effectiveness of predictive policing in Australia relies on the transferability of these international findings, which are dependent on three key factors:

- the quality of Australian data, analysis of which is required both to quantify the spatio-temporal patterning of events, and subsequently to produce timely forecasts of risk;
- the presence of non-random and consistent spatio-temporal signatures that permit risk to reliably estimated; and
- the degree to which these patterns can facilitate predictions of crime risk that are more effective than existing retrospective resource allocation models.

Algorithmic justice

In recent years several fields beyond policing have seen substantial growth in the use of predictive algorithms for decision-making and risk assessment; these applications range from welfare systems and credit score rankings to criminal justice estimations of the risk of reoffending, to name a few. Parallel to this has been the development of algorithmic justice as an area of research. This field is concerned with the application of algorithms to make decisions that can have substantial impact on individuals' wellbeing. Critics argue that algorithmic decision-making processes build in biases existing in the data sources, resulting in unfair and discriminatory harmful effects, sometimes despite specific efforts to prevent these occurring (Brayne 2017; Browning & Arrigo 2021; Jefferson 2018; Marjanovic, Cecez-Kecmanovic & Vidgen 2022; Prins & Reich 2018; Ugwudike 2020). In contrast, proponents of algorithmic justice argue that these algorithms are more objective than human decision-makers, as well as more efficient and more accurate in the estimation of risk and prediction (Brayne 2017; Marjanovic, Cecez-Kecmanovic & Vidgen 2022; Prins & Reich 2018). While much of this research has been focused on welfare systems, disability services and healthcare, there are concerns relating to the use of algorithms in predictive policing as well.

The use of algorithms in predictive policing falls into two main categories: individual-based risk assessments and location-based crime predictions/forecasts. Individual-level predictive policing typically relates to estimating the risk of reoffending for individuals coming into contact with the criminal justice system. Researchers have highlighted the risk of harm in these situations due to a lack of consideration for the bias inherent in the input data used by algorithms. This, combined with the assumed objectivity of algorithmic decision-making, raises concerns about entrenching discrimination and reducing the ability to contest or appeal decision-making (Prins & Reich 2018; Ugwudike 2020; Završnik 2019).

Of most relevance to the present study are concerns regarding location-based predictive policing algorithms. These generally highlight the risks of over-policing certain areas (eg areas with larger minority populations or low-income areas), increasing racialised policing practices such as stop-and-frisk/search with decreased transparency and accountability, and a general increase in negative police interactions with civilians. These criticisms largely focus on the application component of predictive policing—that is, the tactical or operational resource allocation. If location predictions are purely used as a blanket direction to target an area, this can be problematic, for feedback loops may result in the over-policing of minority populations and, some argue, can violate basic criminal justice tenets such as the presumption of innocence (O’Donnell 2019; Richardson, Schultz & Crawford 2019).

Brantingham, Valasik and Mohler (2018) conducted a randomised controlled trial assessing whether predictive policing practices led to biased arrests. The authors found that algorithmically predicted locations had a greater number of arrests than control locations, but that there were no significant differences between the two areas in the proportion of arrests across racial/ethnic groups (Brantingham, Valasik & Mohler 2018). Meijer and Wessels (2019) conducted a review of the benefits and drawbacks of predictive policing and found there was no empirical evidence to support the claimed drawbacks.

Predictive policing is a two-stage process. First, the prediction of high-risk locations provides useful information for police and other services about where demand for their services is likely to be located in the near future. The second stage comprises the tactical response and is likely to be heavily context dependent. Covert surveillance, high-visibility policing, situational crime prevention, crime prevention through environmental design, and community prevention are some of the approaches that might be appropriate, depending on the composition of the problem. Because of this variety in the response to future risk, evaluating both stages of predictive policing is not common, and thus generalisable findings are even more difficult to establish.

The current study

There is a growing body of international research concerning the effectiveness and application of predictive policing, an approach that seeks to generate short-term crime forecasts based on spatio-temporal analyses of previous crime patterns. Predictive policing, which draws on mathematics, epidemiology and computer science, has been productively applied to optimally allocate both operational policing and crime reduction resources in the United States, the United Kingdom and mainland Europe. However, to date, very little research has sought to assess the suitability of Australian crime problems and recording systems for these techniques. As such, the transferability of previous findings to an Australian context is, as yet, largely unknown.

The current study aims to identify spatio-temporal patterning of volume crime incidents using Australian recorded crime data, thereby contributing to the literature by considering the potential for predictive policing in an Australian context. The scope of this project is limited to the first stage of predictive policing: forecasting high crime risk locations across a short-term time horizon. Resource constraints ruled out conducting an operational trial into preventing crime using predictive policing forecasts. As this study is only examining one part of the predictive policing model—the forecasting stage—we needed to be careful to design the analysis appropriately to minimise threats to external validity once we move from pure desktop analysis to real-world resource deployment (the second stage).

There is an enormous difference between performing analysis on local crime trends and making operational decisions based on identified trends. How an analyst may conduct their forecast is likely to yield tremendous differences compared to an academic researcher using the same approach or algorithm. Unless both use the same code or software, there is a good chance there will be a divergence in risk profiling and thus in prioritisation. This will obviously have knock-on impacts once tactical options are chosen. To use an example unrelated to criminology, the academic finance literature is replete with peer-reviewed, published trading strategies that are successful on a risk-adjusted basis yet fail to perform in the real world. The main reasons for the failure of academic models are overfitting and unrealistic trading conditions in academic models (transaction costs, slippage, taxes not included). In other words, there is a danger that our analysis, devoid from real-world implementation, may provide unrealistic evidence in favour of or against predictive policing through premature optimisation, overfitting, and otherwise unrealistic operating conditions.

Our solution to minimise the threat to external validity is to make our forecasting regime as authentic and operationally representative of real-world conditions as possible. Thus, we privilege pragmatic and generalisable approaches over statistically complex and optimised choices. Where there are a wide range of choices in selecting a model's parameters, we favour choices most likely to be selected by an operational analyst. If there is a choice between an algorithm requiring intense computation requirements and one that is relatively easy to implement and 'run', we choose the latter.

We are aware there are different perspectives on this, and fully acknowledge the satisfaction of developing an optimised model with superior metrics. But we are concerned about the false comfort this might instil prior to real-world implementation. We are confident our approach will approximate real-world effectiveness because our findings will be robust to a host of practicalities that will arise once operational decisions need to be made.

As noted above, predictive policing is a two-stage process, involving first the forecasting of future crimes, and then deploying resources to prevent crime. When evaluating the approach in its entirety, the overall aim is preventing crime; as such, the performance is assessed by any observed crime reduction. If conducting an evaluation on the first stage of predictive policing, the aim is to forecast future crimes, and performance is assessed by the accuracy of these forecasts. This study only evaluates the first stage of the predictive policing process and our metrics reflect this, measuring only the accuracy of prediction. We make no claim about the number of crimes that would be prevented, as success in accurately forecasting future crime locations does not guarantee success in the prevention of crimes at these locations.

Data and methods

This section describes the data and methodological approaches used to address the objectives of this study. First, we define the study regions and time window used for the analyses. Next, the recorded crime data used to forecast future crime events are described, as is the data processing required to clean the data. We use three major offence categories to assess the predictability of crime in each study region: burglary, theft of motor vehicle (TOMV) and theft from motor vehicle (TFMV). Finally, the methods and results of basic exploratory data analysis conducted to explore the dataset are described.

In terms of forecasting of crime, we use two forecasting algorithms and a baseline model. The baseline model (retrospective mapping) is our attempt to depict the conventional practice used by crime analysts in day-to-day operational policing. The first algorithm we use to test predictive policing is promap (‘prospective mapping’; Johnson, Birks et al. 2007), which is based on the near-repeat phenomenon. A detailed description of our implementation of promap is provided in later sections. The second algorithm we use was developed by YongJie Lee (Lee, O & Eck 2020). This approach combines two major criminological theoretical approaches (population heterogeneity and state dependence) and performed extremely well in the Real-Time Crime Forecasting Challenge (Lee O & Eck 2017; National Institute of Justice 2016). In addition, it is extremely straightforward to implement and thus more likely to be used in practice. Again, a full description is outlined in a later section.

Both algorithms require historical data to calibrate parameter values to be used in the forecasting. To enable this, we used the first three years (2008–10) of the time window as a ‘training’ dataset to explore the space-time patterns for each study region and crime type combination. Assessment of the forecasting performance was conducted using the remaining five years (2011–16) as a ‘testing’ dataset.

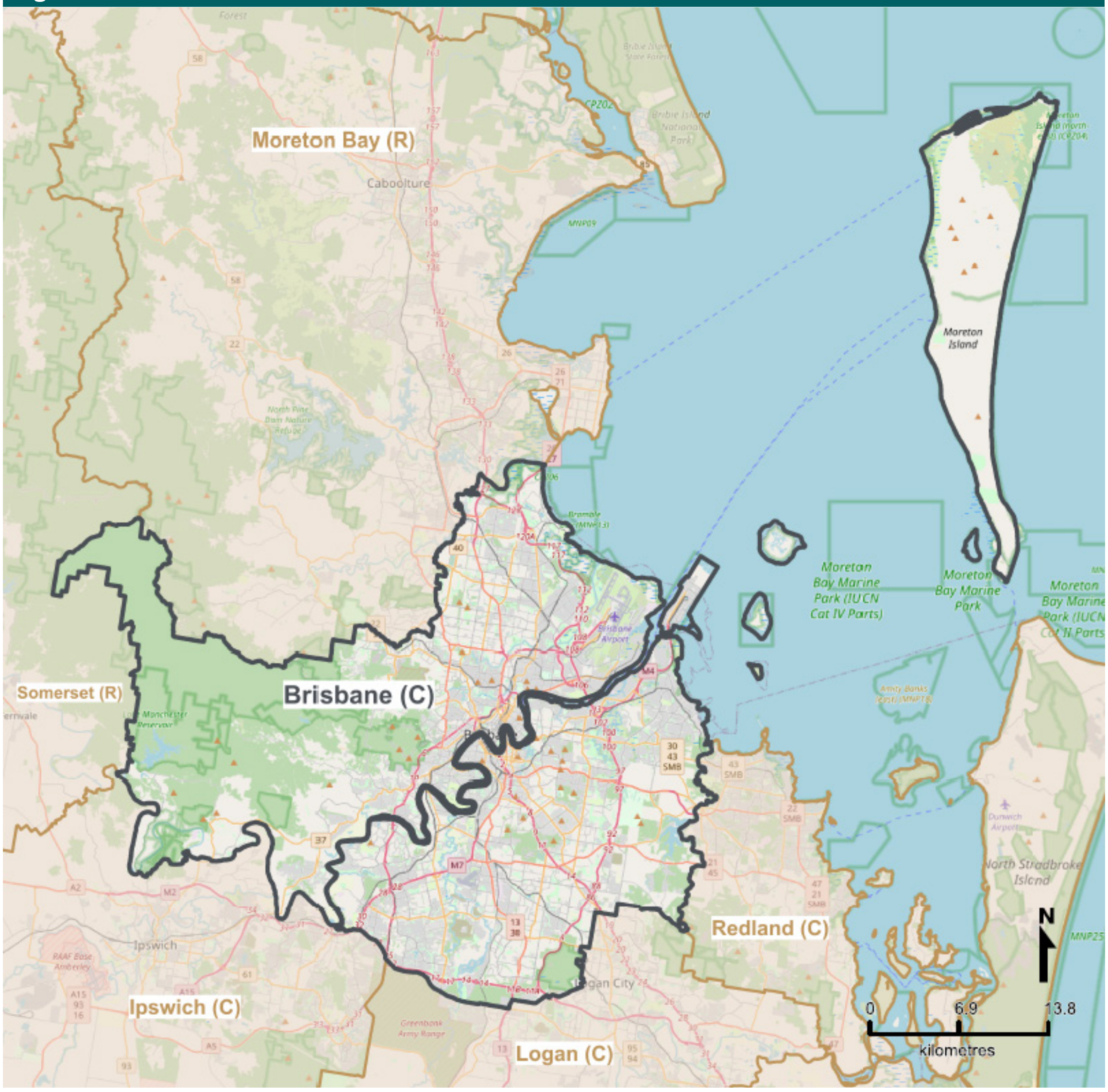
Study regions

The Brisbane, Logan and Townsville Queensland local government areas (LGAs) were used as the study areas for this project. This choice of regions provided a range of contrasting communities with which to explore the objectives of this research.

Brisbane

Brisbane is a large metropolitan city, the capital of the state of Queensland (see Figure 1). 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. Of the three LGAs, Brisbane had the lowest crime rate, unemployment rate, and percentage of population in the lowest SEIFA quintile (see Table 1).

Figure 1: Brisbane LGA

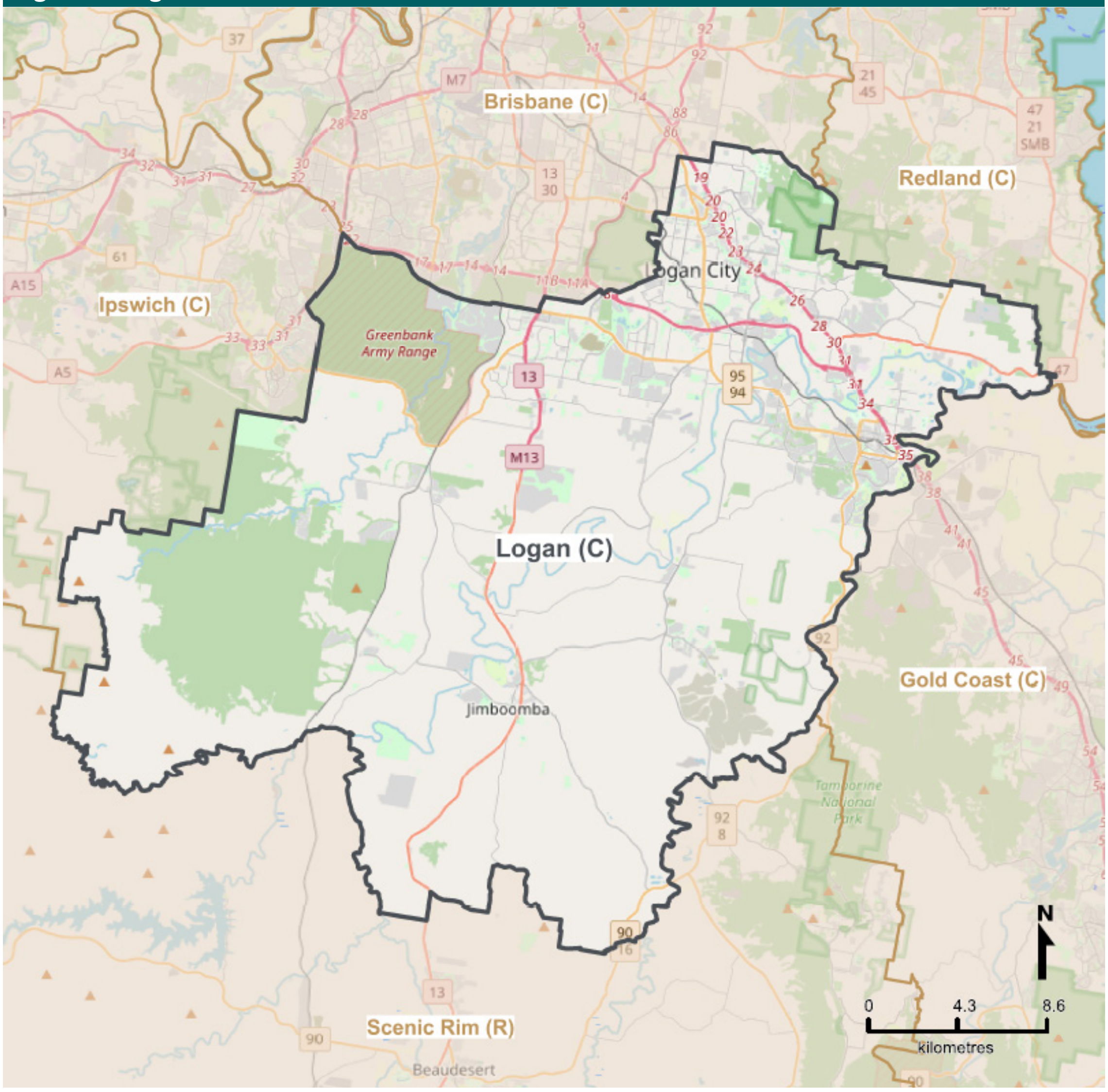


Note: Brisbane LGA has the Australian Standard Geographical Classification code 31000
Source: Queensland Government Statistician's Office 2021

Logan

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 LGA (see Figure 2) 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.

Figure 2: Logan LGA



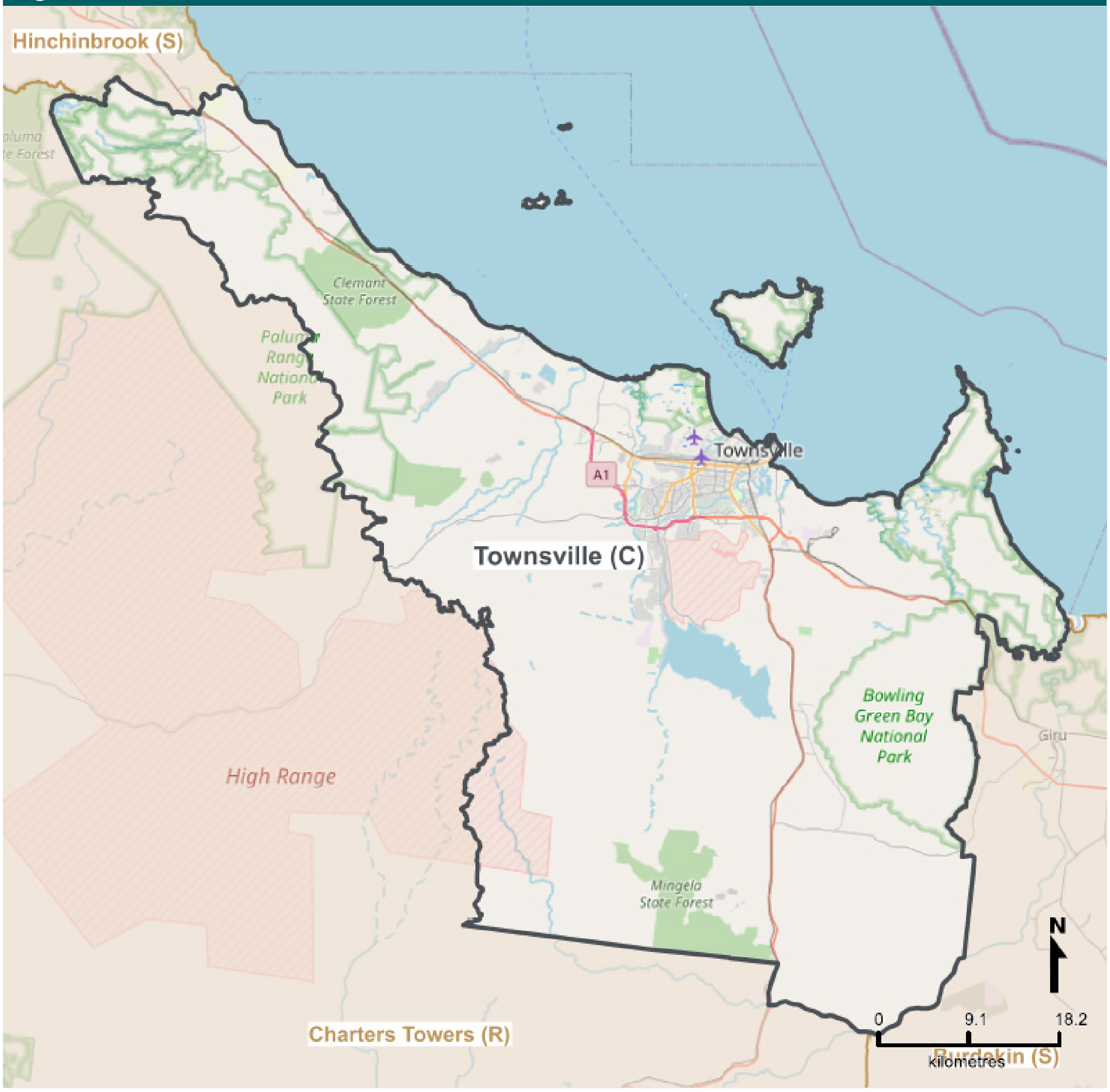
Note: Logan LGA has the Australian Standard Geographical Classification code 34590
Source: Queensland Government Statistician's Office 2021

Compared with the other study regions, Logan had the lowest median income, equal highest unemployment rate, and the highest percentage of the population in the most disadvantaged Socio-Economic Indexes for Areas (SEIFA) quintile (see Table 1).

Townsville

Townsville is a regional community and covers the largest area of the three study regions while having the smallest population (see Figure 3). Like Logan, the residential and commercial land use is spatially concentrated into a small part of the study region. Townsville's median income was just higher than that of Brisbane and the percentage of the population in the lowest SEIFA quintile was in between that of Brisbane and Logan. It had the same unemployment rate as Logan and a slightly higher crime rate, giving it the highest crime rate of the three sites used in this study (see Table 1).

Figure 3: Townsville LGA



Note: Townsville LGA has the Australian Standard Geographical Classification code 37010
Source: Queensland Government Statistician's Office 2021

Table 1: Summary statistics for the study regions

	Brisbane	Logan	Townsville	All Queensland
Area (km^2)	1,342.719	958.132	3,730.82	1,730,172.083
Population	1,184,752	314,511	191,348	4,845,152
Population density (people/km^2)	882.4	328.3	51.3	2.8
Crime rate (offences against property per 100,000 population)	4,416.9	5,996.8	6,412.2	4,706
Unemployment rate	6.8%	8.9%	8.9%	7.6%
SEIFA quintile 1 (%)	4.1%	35%	20.8%	21.4%
Median age (years)	34.5	33.8	33.8	37
Median income (per year)	$51,708	$46,810	$52,134	$46,869

Source: Crime rate statistics for study regions were obtained from Queensland Government Statistician's Office (nd) for financial year 2016–17. The crime rate for all Queensland was obtained from Queensland Police Service (2017). SEIFA quintile percentages were obtained from Queensland Government Statistician's Office (2018). The remaining statistics were obtained from the relevant LGA summary for 2016 from Australian Bureau of Statistics (nd)

As shown in Table 1, the median age in all regions is similar and just a few years younger than the median age across the entire state. The regions differ in area and population, however. 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 population density of Logan, and nearly 18 times the density of Townsville. The unemployment rate of Logan and Townsville is the same, a little higher than the state rate, while Brisbane has the lowest unemployment rate. Median income is very similar between Brisbane and Townsville, with Townsville's a fraction higher, while Logan has the lowest median income of the three areas. A little over a third of Logan's population is in the lowest SEIFA quintile, while less than five percent of Brisbane's population is in this range. Townsville falls about halfway between the two areas on this characteristic. Townsville and Logan have similar crime rates, with Townsville experiencing slightly more crime per 100,000 people. Both these study areas have a higher crime rate than the overall Queensland rate, while Brisbane has the lowest rate, a little below the state average. Based on these statistics, Logan has the lowest socio-economic status and Brisbane the highest.

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 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 the 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). These errors are difficult to explain given that the geocoding of crimes is automated by the information systems used by the Queensland Police Service. 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 each of the district boundaries. The remaining incidents were split into three main offence categories (see Table 2).

Table 2: Crime counts for each crime classification across each of the study regions

	Brisbane	Logan	Townsville
Burglary (2008–2016)	64,483	20,230	14,821
Burglary (2011–2016)	46,756	16,315	12,120
Theft of motor vehicle (2008–2016)	24,518	11,367	6,108
Theft of motor vehicle (2011–2016)	16,313	8,116	3,224
Theft from motor vehicle (2008–2016)	50,437	20,982	12,988
Theft from motor vehicle (2011–2016)	36,828	16,424	8,810

As shown in Table 2, Brisbane had the highest crime counts for all three of the classifications, more than double the counts of Logan, while Townsville had the lowest counts. These differences largely reflect the population differences between the three regions. The crime rates presented in Table 1 refer to all property crime, whereas the frequencies in Table 2 pertain to the subset of crime events.

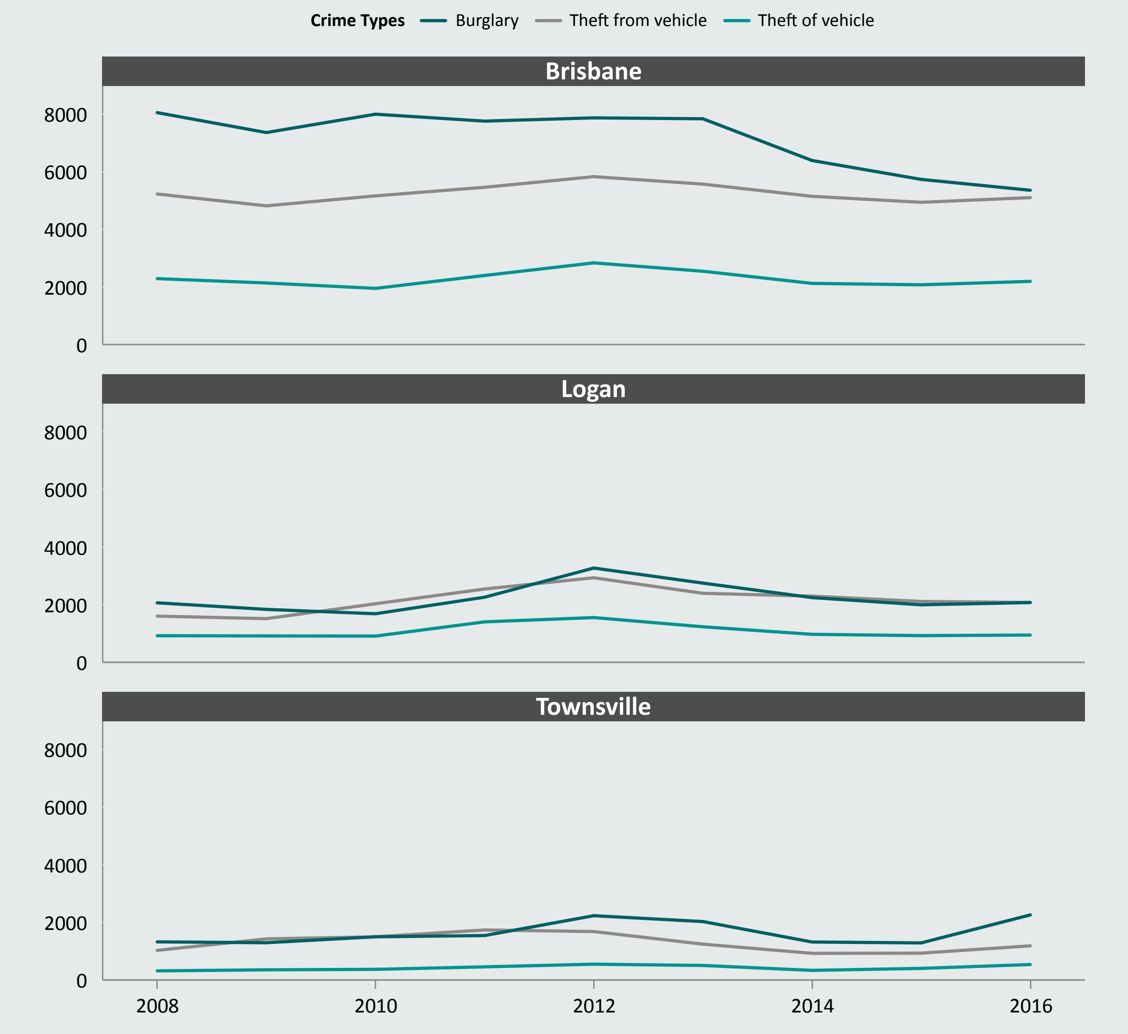
Exploratory data analysis

Patterns in time

We began by examining the patterns in the three major crime classifications on broad time scales before exploring them using finer units of time. Examining the rates of crime across the entire time window helps to show whether one-off 'spikes' in historical data might confound forecasts.

We first investigated how crime incidents changed on an annual basis (see Figure 4). This is important to consider in the event that major changes to levels of crime observed might weaken the predictability of crime events (such as the crime drop, as discussed in Bannister, Bates & Kearns 2018).

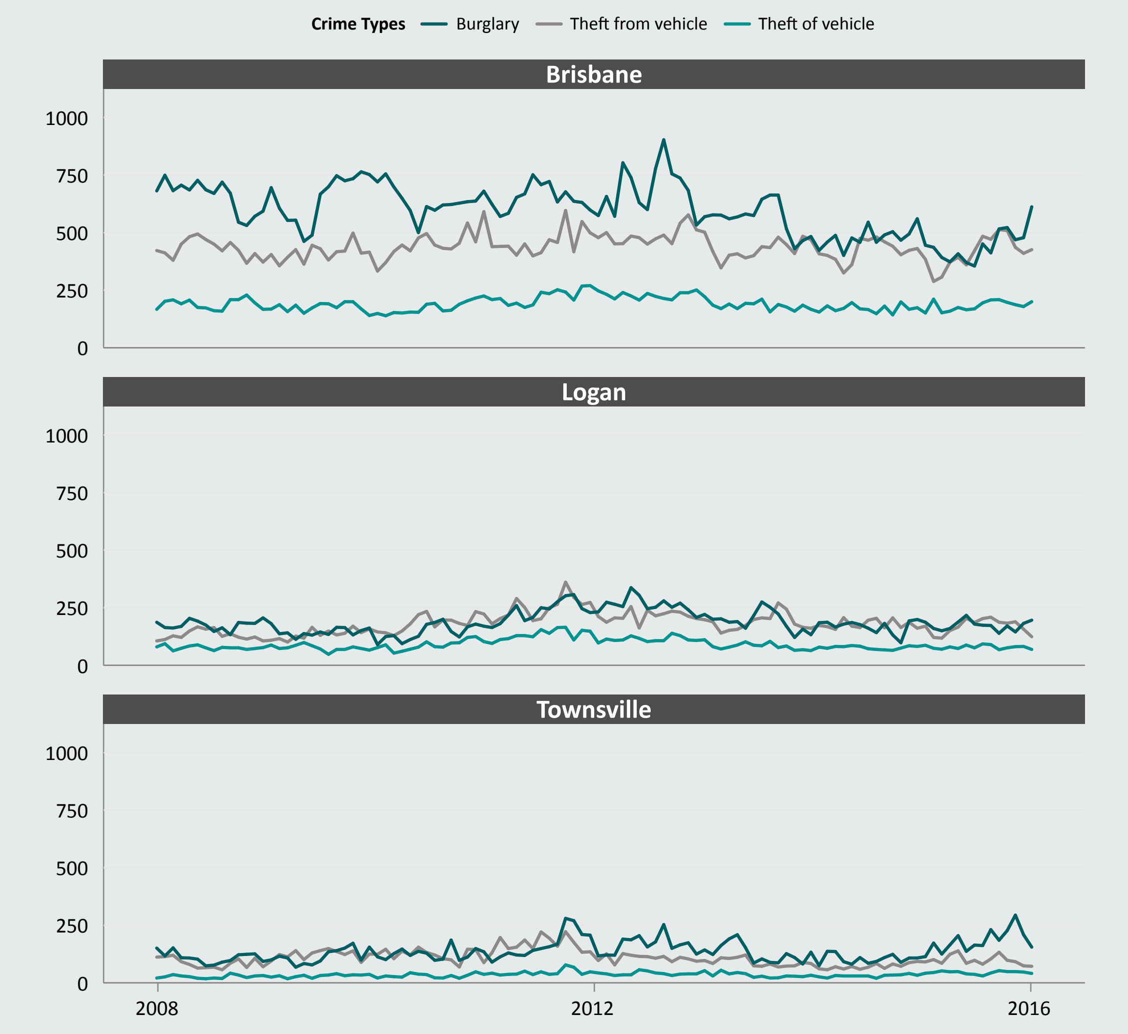
Figure 4: Annual crime counts for each classification across the three study areas, 2008–2016



The three crime types showed fairly similar trends in each study region, but with varying quantities. All three regions appeared to have had a peak around 2012 but overall remained fairly stable. Burglary in Brisbane had a noticeable decline from 2013 to 2016. Burglary in Townsville showed a small increase towards the end of the dataset period. However, with the exception of these two categories, and despite fluctuations, the other crime classification counts in 2016 were very similar to the counts in 2008. It is worth noting that Townsville had a very low rate of TOMV, which could impact on the ability to forecast offending. The trends for Logan are consistent for the three crime types, increasing in the first half of the time window and then decreasing in the second half.

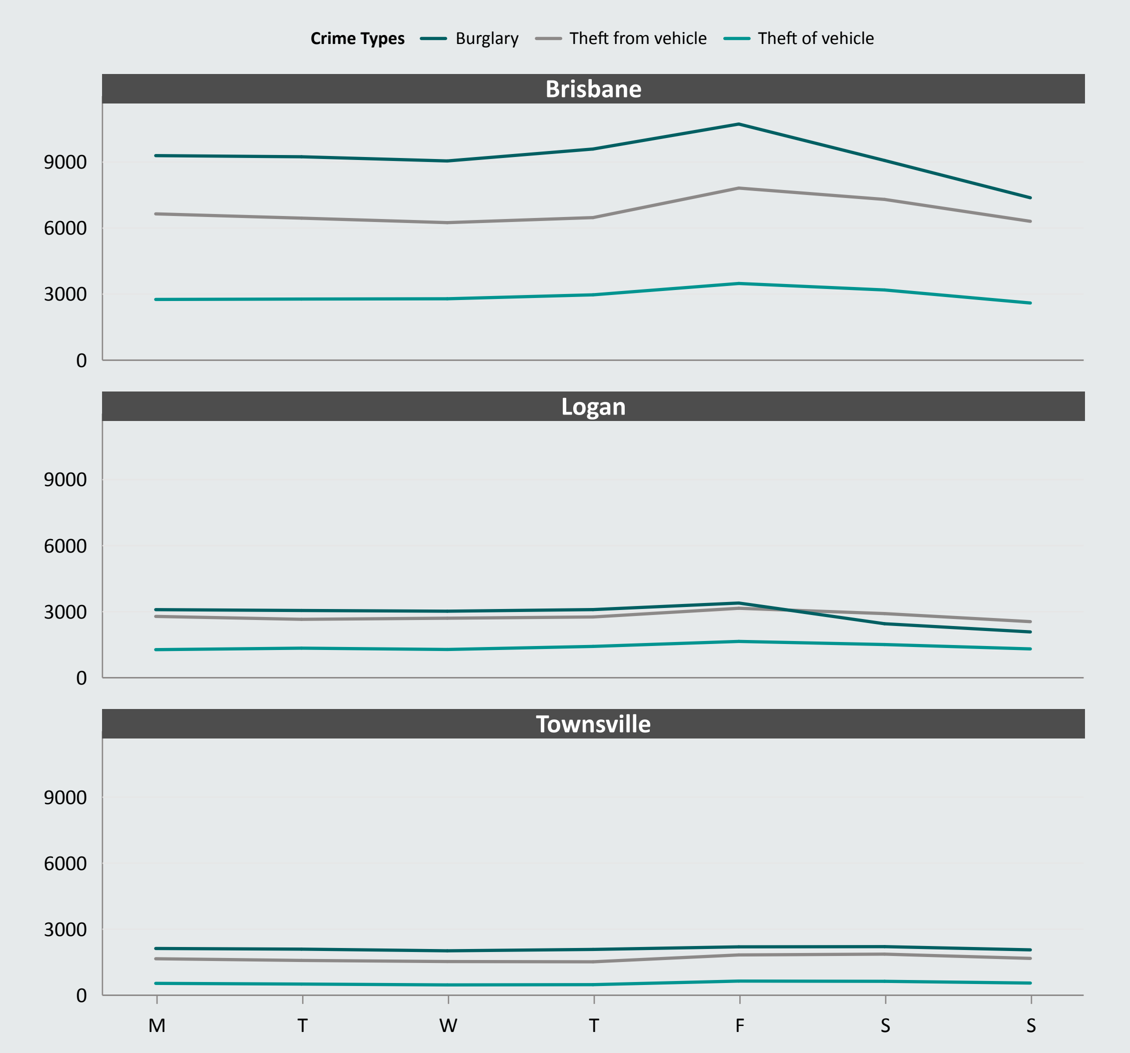
Next we looked at whether there were seasonal patterns for each of the offence types. Figure 5 shows the monthly counts of each crime type during the time window. Burglary displayed considerable fluctuation month to month, particularly in Brisbane with its much larger population. TFMV also showed month-to-month fluctuations but on a smaller scale than burglary. In comparison, TOMV appeared quite stable, especially in Townsville where there was very little variation throughout the study period. Despite the more frequent fluctuations, the seasonal patterns still reflect the overall, smoother patterns shown in the annual trends.

Figure 5: Monthly crime counts for each classification across the three study areas, 2008–2016



The next level of resolution is day of the week. If major intra-week differences were located, this might justify very short-term forecasting horizons (such as days rather than weeks). These results are shown in Figure 6. Townsville showed very little change based on the day of the week across all three crime classifications. TFMV had a small increase on a Friday and Saturday but other than that counts for these offences remained stable across the week. In contrast, Brisbane reached its maximum level on a Friday for burglary and TFMV, although this was less pronounced for TOMV. Burglary decreased to its lowest level on Sunday before returning to a stable level during the first few days of the working week. Logan mirrored this pattern to a degree as well, with a shallow increase to Friday and a less pronounced decrease to Sunday. TFMV and TOMV remained largely stable across the week.

Figure 6: Daily crime counts for each crime classification across the three study areas



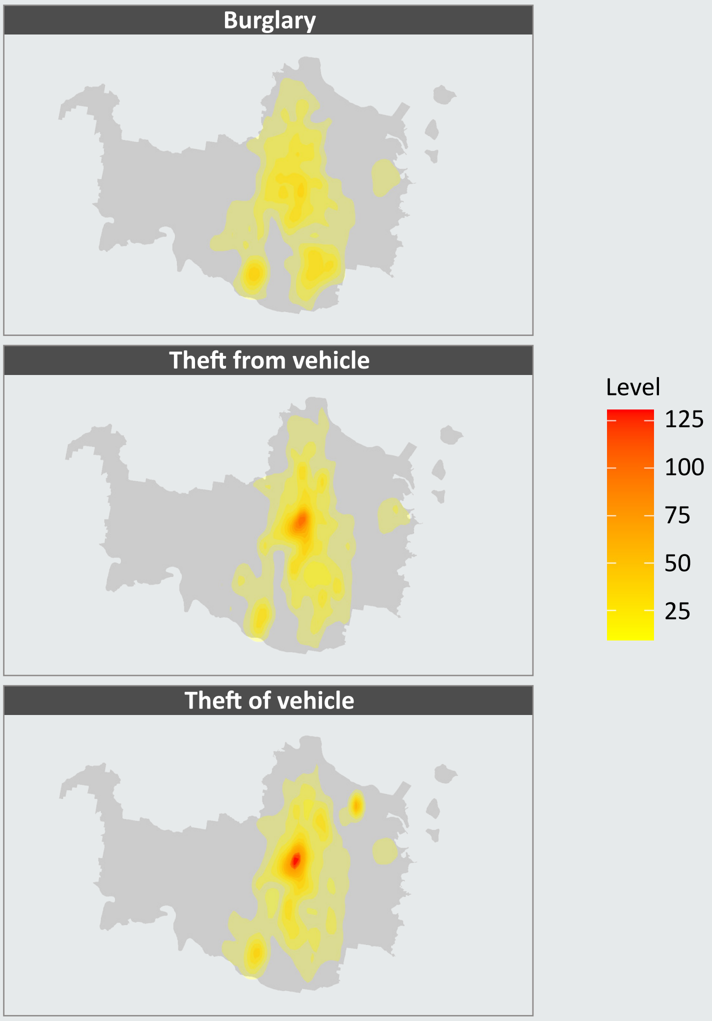
Patterns in space

It is important to get a general picture of the spatial concentration of offences before drilling down to more focused spatial patterns. In the figures below, spatial patterns are depicted through a kernel density estimation, an estimate of the population density. We include a legend for completeness but note that the unit ('level') does not correspond to crime counts or rates. It is a relative measure of concentration and does not have an inituitive analogue to real-world values. Figure 7 shows the aggregate concentration of crimes in Brisbane throughout the time window of the study.

Brisbane

Generally the three crime classifications showed similar overall shapes in Brisbane. Burglary demonstrated a diffuse pattern with no observable hot spots. This is consistent with the residential land use of Brisbane. There was a general increased intensity around Brisbane City itself, in the centre of the map, as well as two locations in the south corresponding to the Forest Lake/Inala area on the left and the Sunnybank/Sunnybank Hills area on the right. TFMV and TOMV were not as widely spread as burglary. TFMV showed a clear hot spot in the centre at Brisbane City and appears to have a 'tail' pointing southward. TOMV also had a clear hot spot in the centre around Brisbane City, as well as a small hot spot in the north-east near Brisbane Airport.

Figure 7: Brisbane hot spots map: Distribution of crime types, 2008–2016

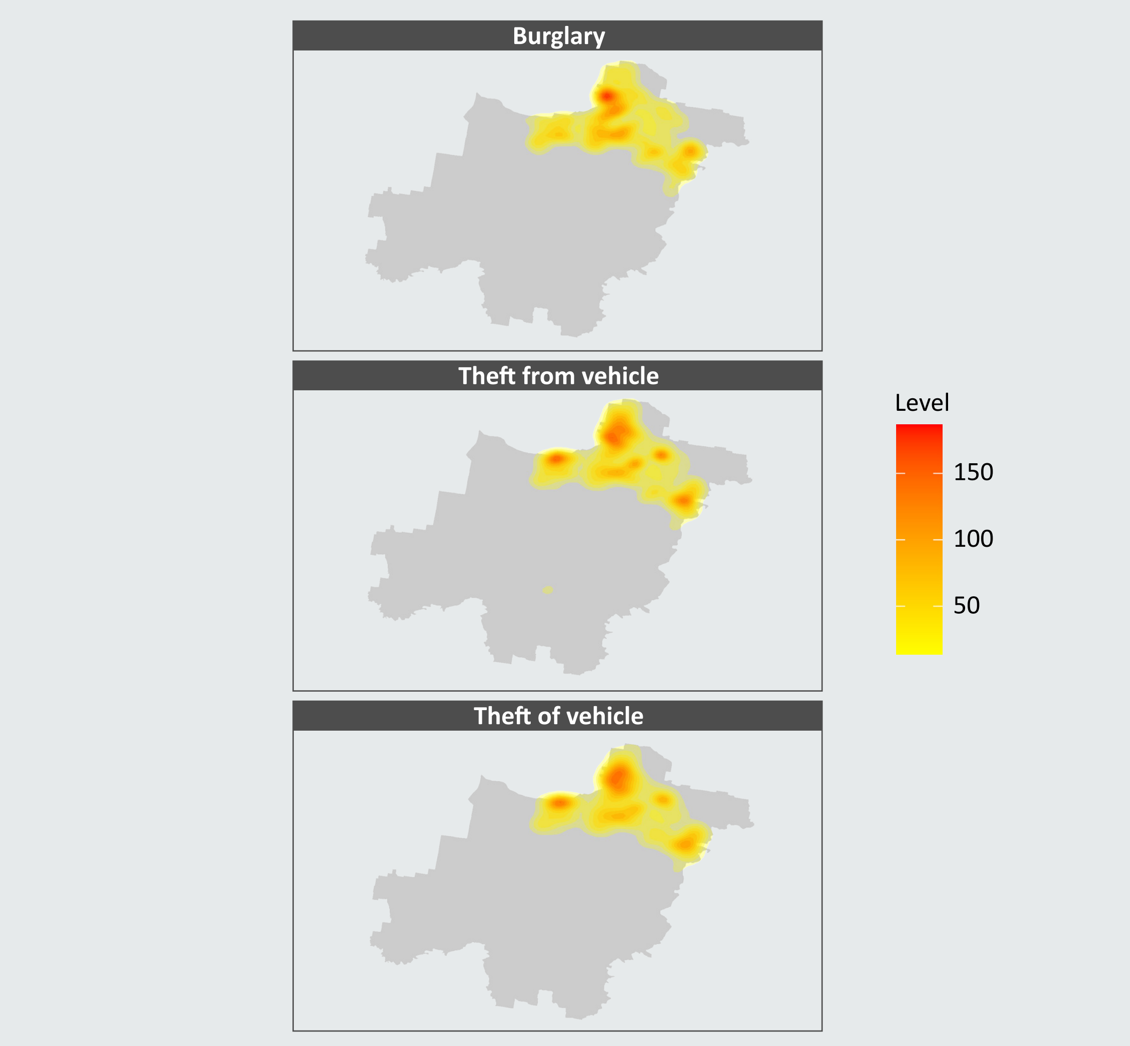


When looking at the three crime classifications by year, it is clear that the higher crime areas were quite stable across the study period (see Figure A1 in the *Appendix*). The overall pattern for burglary remained the same across the study period, with some fluctuations in the intensity of the hot spots. There appears to have been a general decrease in the density of burglary offences from 2011 to 2013, and then a small increase (in density) throughout the remainder of the study period. TFMV demonstrated more fluctuation than burglary across the eight years. While the hot spot around Brisbane City remained, it becomes less clustered, extending to a larger area and, as a result, fading in intensity, particularly through to 2014. In the last two years the intensity increased, but the area it covered remained the same. The TOMV hot spot in Brisbane City remained fairly stable from 2008 to 2016; however, a steady increase in intensity is observable year by year as a hot spot formed at Brisbane Airport in the north-east. The small pocket to the east of Brisbane, around Wynnum, fluctuated throughout the study period, but always remained an area of moderate density.

Looking at patterns across the days of the week, there was very little change across any of the three crime types (see Figure A2 in the *Appendix*). TFMV and TOMV both shared a similar pattern, appearing to cluster around Brisbane City on a Friday and Saturday, with a more intense hot spot, while for the rest of the week seeming to have a more general spread throughout the surrounding areas, particularly on a Monday. Burglary demonstrated no discernible weekday patterns.

Logan

Figure 8: Logan hot spots map: Distribution of crime types, 2008–2016



All three offences were heavily concentrated in the northern quarter of the study region, around Logan City and Beenleigh, which are the more populated areas. TFMV had a faint hot spot located towards the centre of the region around Jimboomba, while TOMV had a small pocket around Logan Village.

Next, we examined how consistent this pattern was by year. As above, if the spatial patterns differ year by year the feasibility of forecasting might be questioned. Year-to-year variation was not very high for any of the offences (see Figure A3 in the *Appendix*). The overall pattern of where burglaries occurred in Logan was largely unchanged from 2008 to 2016. The only variation was the intensity of the hot spots, but the locations themselves did not show much change. In 2008 there was a hot spot in the centre of the north around Browns Plains; however, this hot spot faded in and out every few years but did not reach the intensity of 2008 at any other point in the study period. TFMV had a similar pattern, with little variation in the locations of the hot spots across the study period. Between 2009 and 2013 a small cluster of offences formed in the centre of the region in Jimboomba, but this then disappeared after 2013 and did not reappear for the remainder of the study period. TOMV also had little movement in where the crimes were occurring, just small fluctuations in their volume from year to year. In 2016 a small hot spot began to form in the Logan Village area.

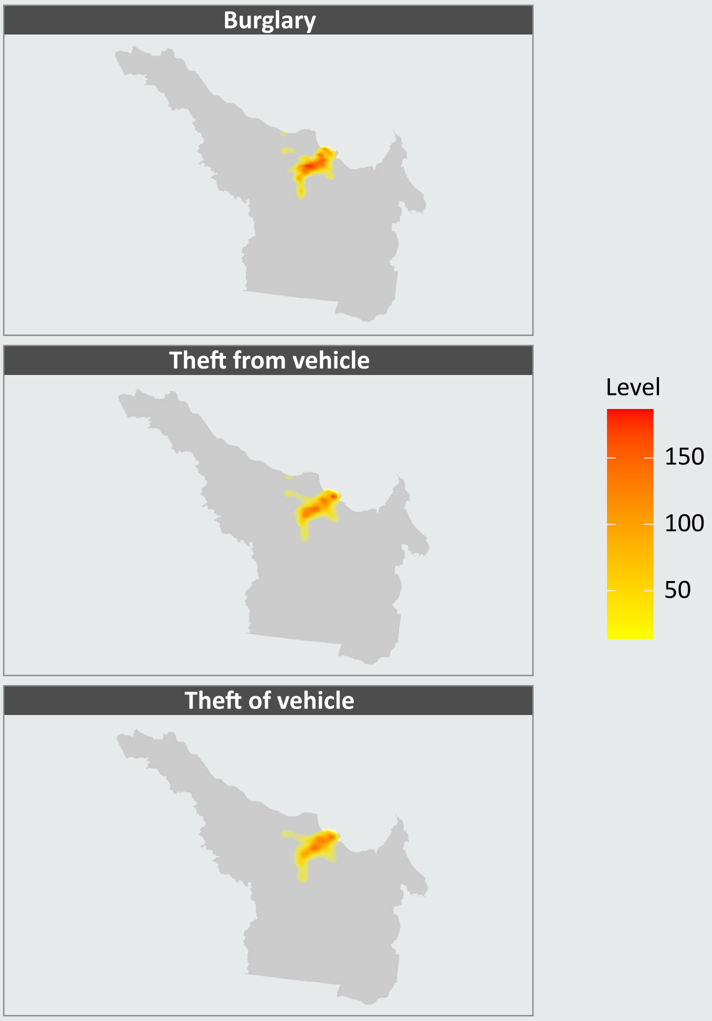
When looking at offences across the days of the week, again there was very little change between the overall pattern of where the offences occurred (see Figure A4 in the *Appendix*). TFMV and TOMV both demonstrated very similar patterns, and the frequency of offences seems to be highest during the working week. Burglary also showed very little change in the location of the hot spots; however, unlike the other two crime classifications, the intensity of the burglary hot spot was greatest on the weekend.

Townsville

As mentioned above, residential and commercial land use is concentrated in a small area of Townsville. In the maps that follow we display a subset of the Townsville study region focused on these areas.

The Townsville hot spots, shown in Figure 9, were very similar across crime classifications, clustering around Townsville City and the surrounding suburban areas. The burglary hot spot was most intense throughout the more densely populated suburban areas outside Townsville City and then there were two faint hot spots located around Bushland Beach in the north and Deeragun just south of there. TFMV and TOMV showed a slightly wider spread than burglary but covered roughly the same areas. Both of these crime classifications showed a more intense hot spot in Townsville City.

Figure 9: Townsville hot spots map: Distribution of crime types, 2008–2016



There was little fluctuation across the three crime classifications from year to year (as shown in Figure A5 in the *Appendix*). The hot spots covered largely the same areas across the study period, with a slight tendency for the intensity to fade and spread.

All three of the crime types demonstrated almost no change across the days of the week (see Figure A6 in the *Appendix*). The hot spots showed only very minor fluctuations in intensity and spread, with no real patterns emerging for any of the crime classifications.

Patterns in space-time (near repeats)

To explore space-time patterns, a series of Knox tests were used. These are described extensively in other studies (Johnson, Bernasco et al. 2007; Johnson, Birks et al. 2007; Townsley, Homel & Chaseling 2003; Townsley & Oliveira 2015), so only a brief description will be provided here.

One way of establishing space-time patterns in phenomena is to look at pairs of observations. If there are more 'close' pairs in space and time than would be expected on the basis of chance, then we can conclude that there is some space-time dependency for contagion or infectiousness.

The Knox procedure is a straightforward way to estimate this effect. N observations will generate $(N(N-1))/2$ unique pairs. For example, four cases will generate six unique pairs: {1,2}, {1,3}, {1,4}, {2,3}, {2,4}, {3,4}. Each observation has a location in space and time. For all pairs, we can calculate the difference in space (in metres) and time (in days).

Next, we look at the combination of distances in space and time. It is possible to generate an expectation if there were no space-time mechanism generating the phenomenon. We can compare the observed number of close pairs with the expectation. If there are more close pairs than are expected on the basis of space-time independence, it can be concluded that there are near repeats in the sample.

In this study, Knox tests were run for each crime type and study area. Because it is hard to define 'close', we examined the number of pairs at different combinations of the space-time distances. We conducted Knox analyses using space intervals of 200 metres, 500 metres and 1,000 metres and time intervals of one week and two weeks. We were constrained from selecting space intervals smaller than 200 metres due to the relatively low housing density in Australia compared to other studies. For instance, a spatial bandwidth of 50 metres covers approximately four houses, while a 200-metre bandwidth contains about 66 houses per cell. After trialling the various intervals, the combination of a space interval of 200 metres and a time interval of one week was selected, as this provided a high resolution capable of identifying space-time patterns. The results of this combination are presented below.

Burglary

The Knox analyses for the three study regions for burglary based on incidents across the training period (2008, 2009 and 2010) are summarised in the figures below. Each cell in the grid represents a space-time distance combination. Cells are shaded if there are more pairs observed in that combination than would be expected by chance. The depth of shading indicates the relative magnitude of the departure. Cells shaded as 'greater than 2' indicate that there are at least twice as many pairs of crimes in those cells as would be expected if there were no crime risk contagion. If there are space-time patterns, we expect a cluster of shaded cells in the bottom left corner of the figures, representing crime events occurring in close spatial and temporal proximity.

Figure 10 displays the results of the burglary Knox analyses for 2008 using one-week time intervals. Each location displays an excess of pairs close in space and time, indicated by the clusters of shaded cells in the bottom left corner of each plot. In Brisbane, the greatest number of pairs occur within two weeks and 200 metres, with the concentration decreasing as time and space increase. Logan demonstrates a very similar pattern, with more close pairs than would be expected occurring within three weeks and 200 metres of the initial offence. Townsville is very similar to the other locations, with the greatest number of close pairs occurring within two weeks and 200 metres.

Figure 10: Knox analysis of burglary for the three study regions with a one-week time unit using 2008 data

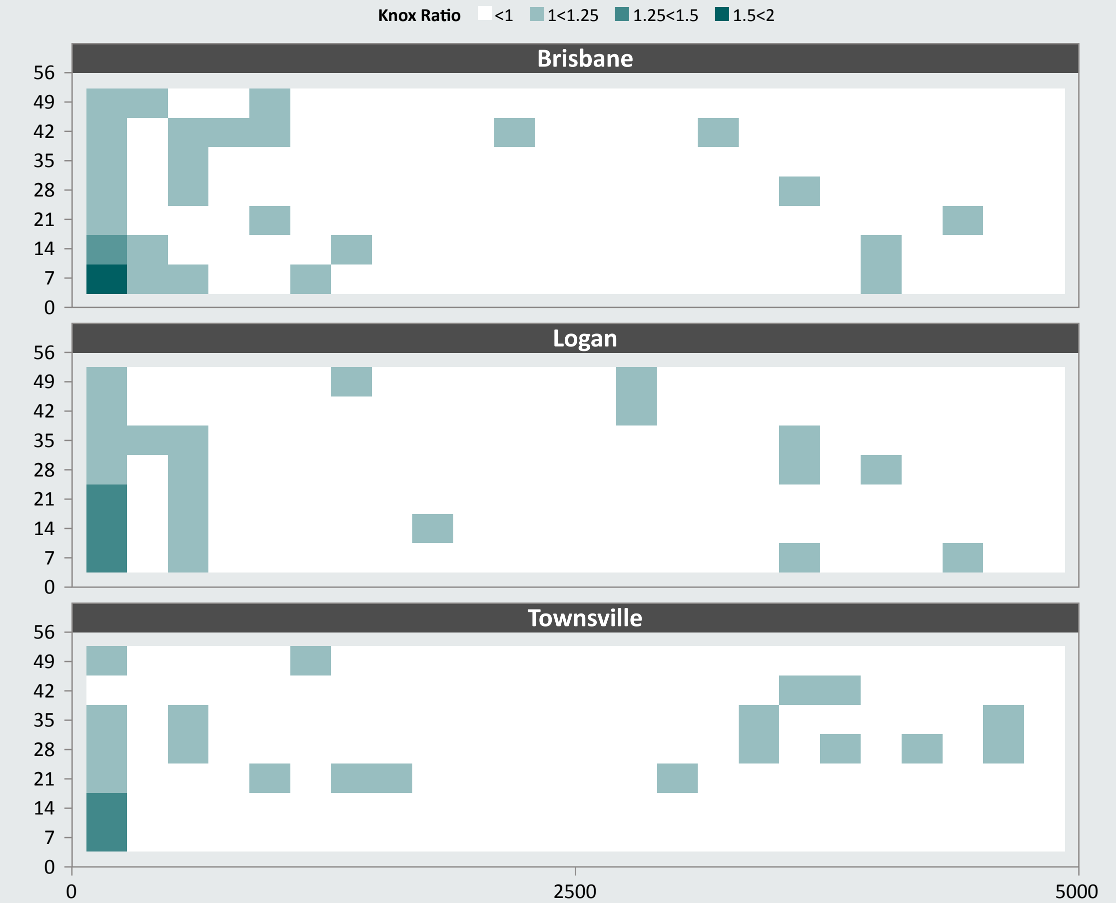
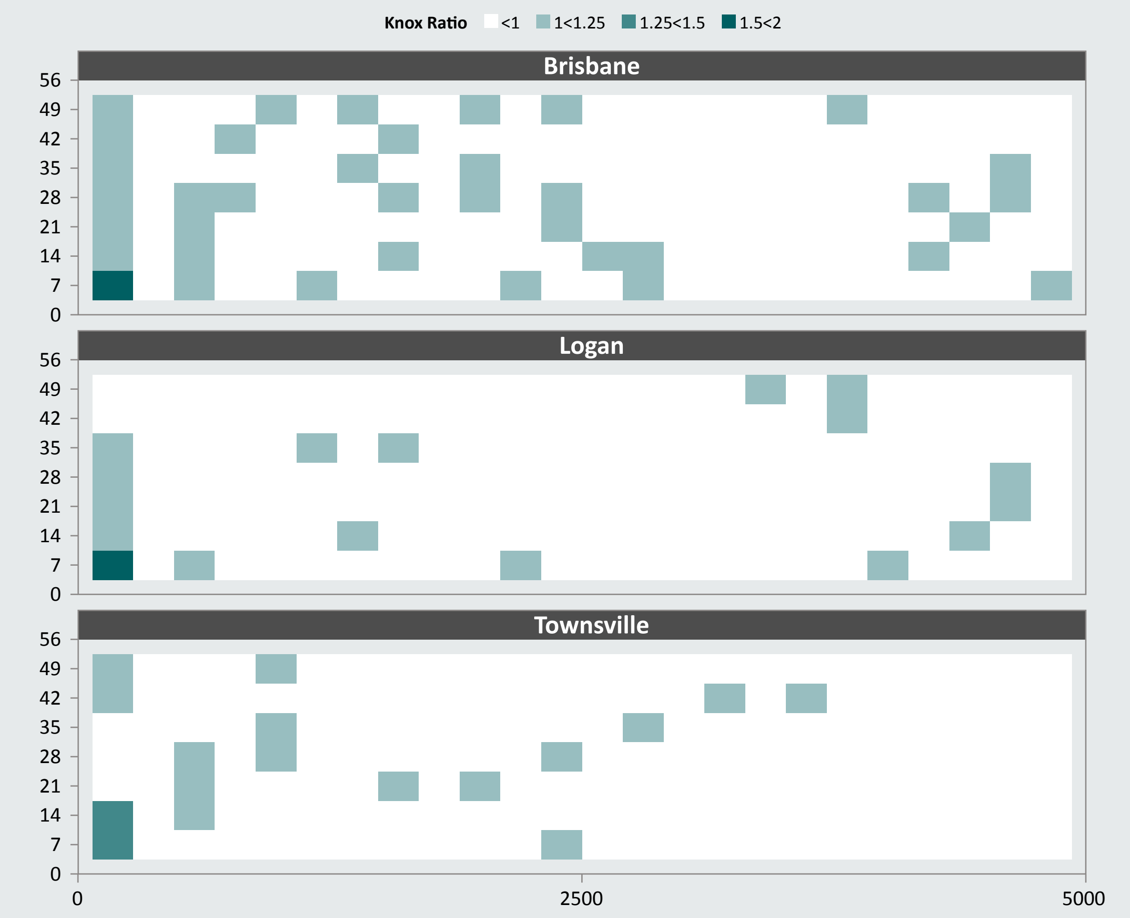


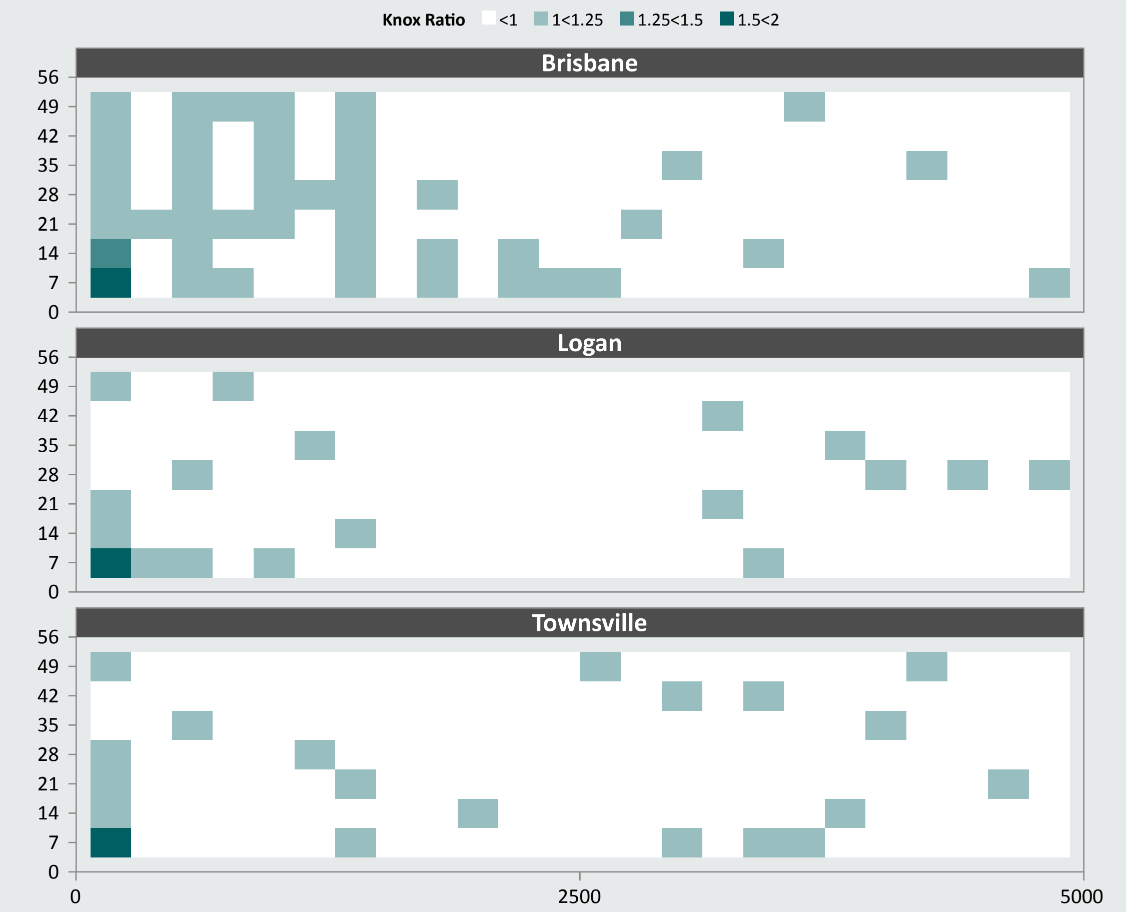
Figure 11 displays the Knox results for burglary offences in 2009, with one-week time intervals and spatial intervals of 200 metres. In Brisbane, the highest concentration of pairs still occurs within 200 metres of the initial offence, but this persists for only one week, with lower concentrations persisting out to seven weeks. Low concentrations of pairs are scattered throughout the seven-week window and within 2,800 metres. Logan demonstrates a clear cluster in the lower left, with the highest concentration within one week and 200 metres of the original offence, and low concentration extending out to five weeks. In Townsville, burglary offences have a higher concentration within 200 metres but out to two weeks after the initial offence. The high concentration spatial patterns for each site are similar to those of 2008, but the temporal patterns demonstrate some fluctuation, particularly in Logan, where the time period decreased from three weeks to one.

Figure 11: Knox analysis of burglary for the three study regions with a one-week time unit using 2009 data



When looking at the spatial-temporal patterns for burglary in 2010, we again see the highest concentration of pairs for all sites within 200 metres of the original offence (Figure 12). In Brisbane the temporal clustering is within the first two weeks of the original offence, as was the case in 2008. The lower concentration cells are again scattered throughout the left side of the figure. In Logan the highest concentration occurs within one week and 200 metres of the original offence, with lower concentrations of pairs within 600 metres and three weeks. Townsville also demonstrates high clustering within one week and 200 metres, with lower concentrations out to four weeks.

Figure 12: Knox analysis of burglary for the three study regions with a one-week time unit using 2010 data



These three figures demonstrate that the higher concentrations of pairs are quite stable spatially but the temporal parameter fluctuated between one and three weeks, depending on the location. The spatial and temporal patterns of the lower concentration cells varied across the years, particularly in Brisbane. These results suggest that rather than using set parameters for the forecasts, it may be more appropriate to dynamically update the spatial and temporal bandwidths based on the patterns from the previous year. This approach would also be more consistent with real-world applications, in which parameters are likely to be regularly updated based on more recent data.

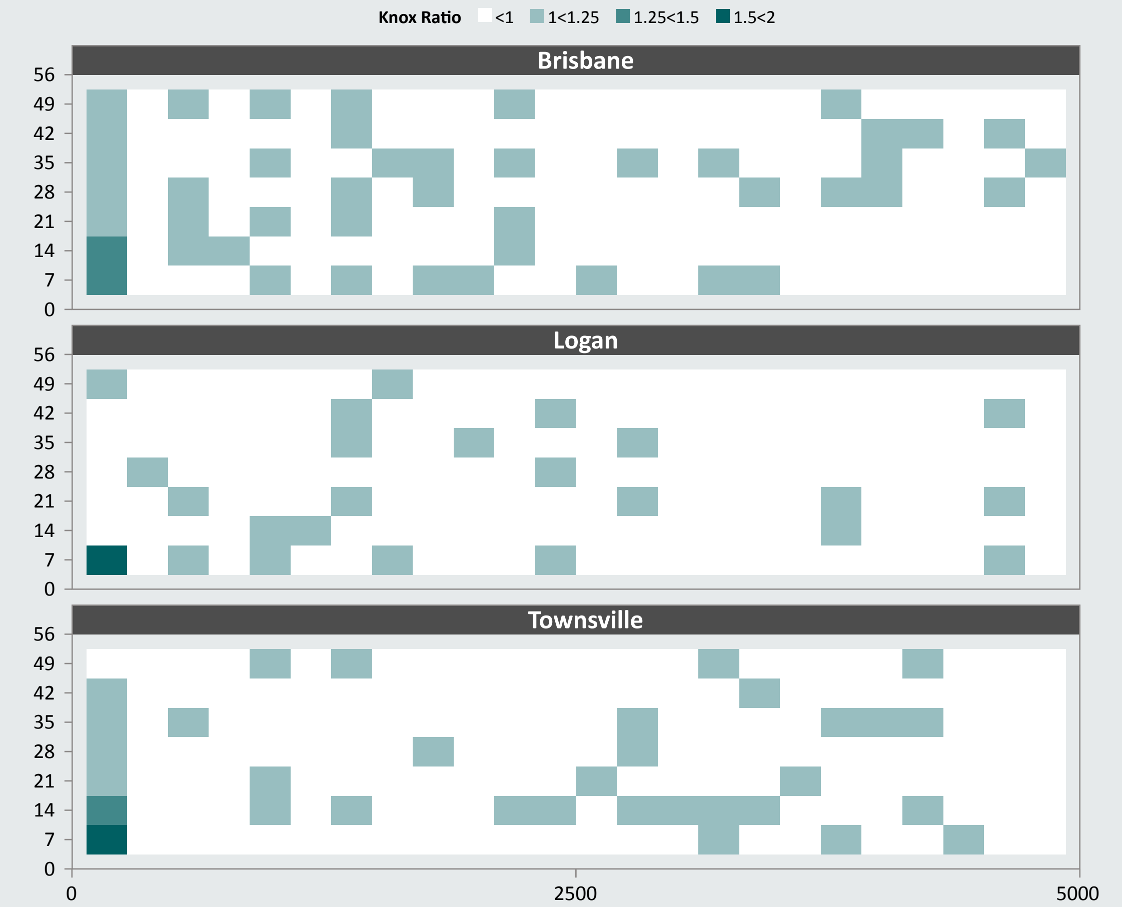
Theft from motor vehicles

Figure 13 displays the TFMV results of the Knox analyses for 2008 with a one-week temporal interval. Brisbane and Townsville demonstrate pairs clustering in the lower left corner of the plot, but Logan does not demonstrate a strong pattern. Brisbane has a clear concentration within one week and 200 metres, with lighter shading indicating some risk contagion within seven weeks and 600 metres. Logan has some risk contagion within two weeks and 200 metres and scattered light-shaded cells around 1,000 metres within four weeks. Townsville has a strong concentration of pairs within three weeks and 200 metres of the initial offence.

Figure 13: Knox analysis of TFMV for the three study regions with a one-week time unit using 2008 data

The 2009 TFMV Knox results show the high concentration cells clearly cluster in the lower left sections of Figure 14. In Brisbane two cells have a ratio between 1.25 and 1.5, indicating a higher concentration of pairs within 200 metres and two weeks of the initial offence. Lower concentrations of pairs appear within 200 metres out to seven weeks, and for several weeks after the initial offence at distances of 600 metres, 800 metres and 1,000 metres. Logan demonstrates clustering in the lower left quadrant, with high concentrations within 200 metres and one week, and lower concentrations of pairs regularly occurring within three weeks and 1,600 metres. TFMV offences in Townsville have a clear 200-metre spatial risk, with high concentrations of pairs occurring within this distance and two weeks of the original offence, and lower concentrations out to six weeks.

Figure 14: Knox analysis of TFMV for the three study regions with a one-week time unit using 2009 data



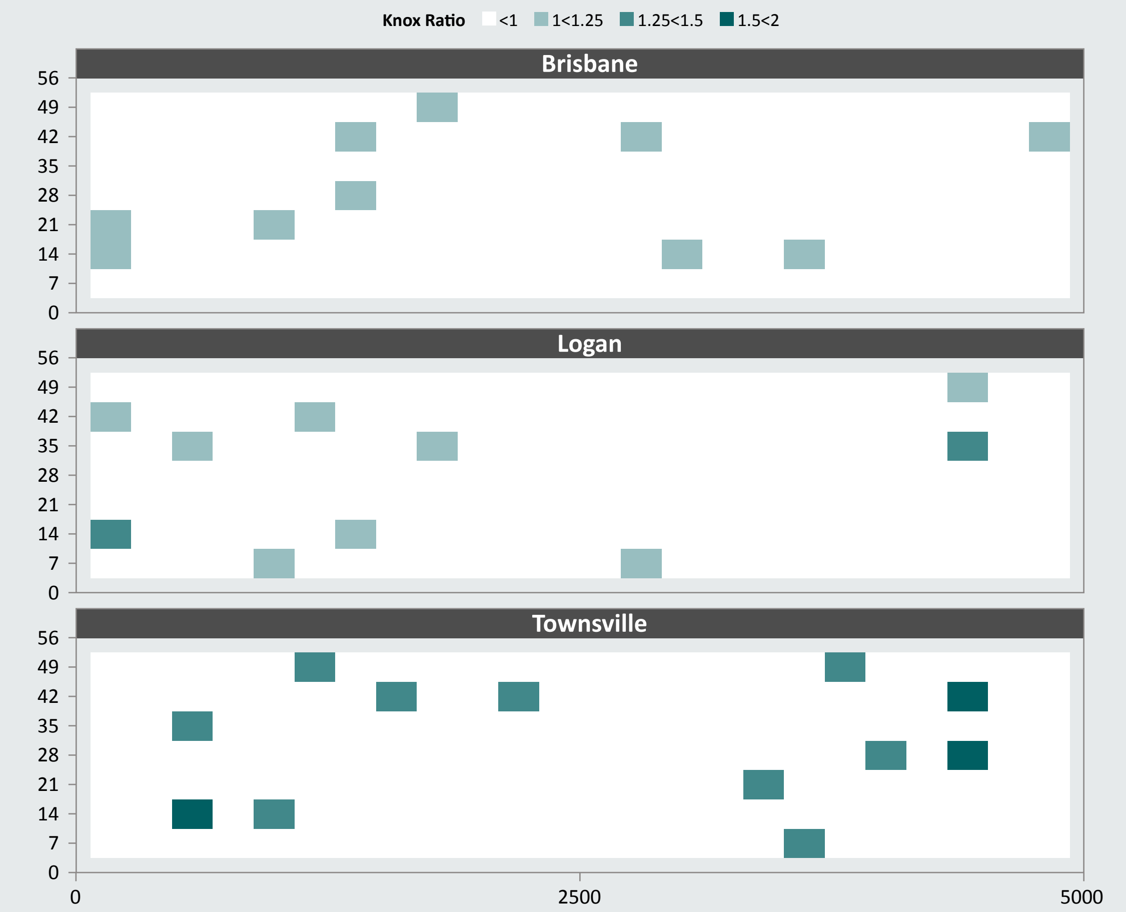
The results for the 2010 TFMV Knox analyses (Figure 15) show that Brisbane has less dispersion of low concentrations of pairs across the window analysed. The high concentration cells again occur within 200 metres and two weeks, with the low concentration cells extending this time window out to five weeks. This offence type does not demonstrate high concentration of any pairs in Logan, but the lower concentration cells do cluster towards the lower left section, broadly within 600 metres and three weeks. In Townsville there is a clear spatial clustering of pairs over 200 metres within four weeks, with some increased presence of pairs persisting for seven weeks after the initial offence. These results suggest that TFMV offences in Brisbane and Townsville demonstrate space-time patterns, though the lower concentration (lightly shaded cells) spatial and temporal patterns demonstrate variance across the training dataset window. In Logan these offences may not be as predictable, with 2008 and 2010 each not demonstrating cells with a high concentration of pairs, indicating that there are no strong spatio-temporal trends for TFMV offences in this location, although there were weaker concentrations present.

Figure 15: Knox analysis of TFMV for the three study regions with a one-week time unit using 2010 data

Theft of motor vehicles

Figure 16 displays the results of the Knox analyses for TOMV incidents in 2008 with a weekly time interval. In general, the TOMV incidents demonstrate weaker space-time patterns than the other crime types. This is likely due to the lower crime counts, with 2,290 in Brisbane, 923 in Logan, and 317 in Townsville in 2008. Brisbane has some light-shaded cells between two and three weeks and within 200 metres of the initial offence. It is of note that, in contrast to the previous Knox figures, there is not a greater number of pairs than expected within the first week after the offence. Logan has an increased concentration of pairs occurring two weeks after the initial offence within 200 metres, and some lower concentrations at 1,000 metres and 1,400 metres within one week, and in the second week after the offence. Townsville has some higher concentrations of pairs two weeks after the initial offence at distances of 600 metres and 1,000 metres, but there are higher concentration cells scattered throughout the assessed range.

Figure 16: Knox analysis of TOMV for the three study regions with a one-week time unit using 2008 data



The Knox analyses of TOMV offences in 2009 also show weak space-time patterns across the three sites (see Figure 17). Brisbane had a higher concentration of pairs clustering within 200 metres and a two-week temporal range. Lighter-shaded cells persist out to a five-week range, and there is also a cluster around 600 to 800 metres during the second and third weeks after the initial offence. Logan has lightly shaded cells scattered across a wide spatial range, with most of these occurring between five and seven weeks after the initial offence, providing limited evidence of space-time trends. Townsville has a high concentration of pairs occurring within 200 metres and one week; however, there were several high concentration cells occurring between 3,000 metres and 5,000 metres from the initial location. TOMV is the least prevalent crime type of the three crime types considered in this study, which may play a part in the modest patterns observed.

Figure 17: Knox analysis of TOMV for the three study regions with a one-week time unit using 2009 data

In Brisbane and Logan, the recorded TOMV offences from 2010 do not demonstrate space-time patterns (see Figure 18). Brisbane has a lightly shaded cell two weeks after the initial offence and within 200 metres, but there are other lightly shaded cells scattered throughout the plot. In Logan there are only four lightly shaded cells, indicating pairs occurring slightly more than would be expected, but these are not concentrated in the lower left quadrant. Townsville has a high concentration of pairs within one week and 200 metres; however, the largest grouping of shaded cells occurred between 2,400 and 2,800 metres and two weeks to seven weeks after the initial occurrence.

Figure 18: Knox analysis of TOMV for the three study regions with a one-week time unit using 2010 data

Conclusions

Based on the Knox analyses presented, it appears there is a communication of crime risk, with the parameters provided in Table 3 below. Burglary offences demonstrated the most consistent evidence of space-time patterns, with the spatial parameter in particular being stable across 2008–10 in all three sites. The temporal parameter varied a small amount across the different sites and years, ranging from one week to three weeks. TFMV offences showed greater than expected numbers of pairs in the lower left quadrant of the Knox analyses in Brisbane and Townsville.

Once again, the spatial parameter was consistent at 200 metres, while the temporal parameter fluctuated by region and year. The Knox analyses of TFMV offences in Logan did not consistently demonstrate strong support for space-time dependencies, suggesting that the prospective forecasting method may not perform as well for that offence type and location. TOMV offences demonstrated inconsistent evidence of space-time patterns, either having some cells with slightly more pairs than expected in the lower left, or demonstrating no clear patterns, with shaded cells scattered sporadically throughout the window. This may be a result of the small number of recorded motor vehicle thefts at the testing sites, or it may relate to an aspect of the offence itself being less conducive to forecasting.

Table 3: Summary of communication of crime risk parameters for each crime classification in each of the three study regions

Region	Crime type	Year	Time (weeks)	Space (metres)
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
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
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 lightly shaded cells (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 testing dataset (calendar years 2011 to 2016) was used to evaluate the predictive power of the parameters established using the training dataset. While the spatial parameters demonstrated a high level of stability, the temporal parameter varied across the training period. As a result of this finding, the forecast parameters for the prospective method will update every year, using the Knox results of the previous year. This is not only consistent with the predictive policing approach but also representative of operational practice.

Forecasting future crime events

The results of the near-repeat analysis were used to inform the implementation of the prospective algorithm (Bowers, Johnson & Pease 2004; Johnson, Birks et al. 2007; Johnson et al. 2009) for three offence categories: burglary, TFMV and TOMV.

Analytical set-up

Our project involved predicting the probable locations of future crime events and analysing the accuracy of these predictions. We implemented three different methods of forecasting: a conventional retrospective KDE surface, a prospective KDE surface, and a population heterogeneity and state dependence algorithm (Lee, O & Eck 2020). As noted earlier, 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. The principle underlying this approach is that if the surface depicts the population sufficiently well, it should be useful for predicting locations of interest in the future.

KDE works in the following way. A grid is placed over the study area. For each cell we identify the number of events located within a distance *d*, known as the spatial bandwidth, from the midpoint of each cell. For each cell, events are weighted according to their distance from the cell midpoint, so points close to a cell midpoint will contribute more to the cell's score than points located further away.

The prospective KDE algorithm is similar, but points are also weighted by recency: events that occur in the recent past contribute more to the surface than those taking place earlier in the time window. In contrast, conventional KDE gives each point equal weight in generating the surface.

At a technical level, each cell in the risk surface is represented by a number—the intensity value—and, when visualised on a map, the colouring schema simply reflects relative differences in these values. Because the cells have a numerical score, they can be ranked from highest to lowest. Assessing the predictive accuracy of different surfaces involves determining whether the locations of events in the future correspond to the highest-ranked cells.

In practical terms this means that locations can be prioritised in order of expected crime risk. If the prioritisation is predictive, we would expect a higher number of crimes to occur in cells with high intensity values. This ranking, in turn, provides preventive opportunities.

The final forecasting method used was a population heterogeneity and state dependence algorithm described by Lee, O and Eck (2020). In the interests of brevity, we will refer to this as the YJL algorithm to reflect its key inventor, YongJie Lee. The YJL algorithm is simple to implement (having been designed to run in Microsoft Excel) and provides transparency throughout the process. It 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 and risk increases at these consistent high-risk locations (Lee, O & Eck 2020). The population heterogeneity component is based on the premise that the features of a place signal the presence of crime opportunities to offenders, with some locations providing better opportunities than others. The high-risk locations based on these features tend to be more consistent and stable over time. Once these are identified, the state dependent component is based on the premise that once an offender has learned about an area and its suitability (after successfully committing an offence there) there is a period of increased risk of subsequent offences which decays over time. The state dependent model assumes that the locations that had more recent crime occurrences have a greater chance of crimes occurring due to this risk. In combination, these two processes identify the most recent crime locations among the most consistent high-risk locations to forecast the most probable crime hot spots (Lee, O & Eck 2020).

As with the two KDE processes described above, for the YJL algorithm, a grid is first placed over the study area. We used the same grid parameters across all three forecasting methods, with regular 140 metre by 140 metre grid cells. In the population heterogeneity component, the Poisson probability of a crime occurring in each week is calculated based on the previous 52 weeks of crime data for each grid cell. If the calculated probability is greater than 0, and a crime was recorded in that week, then the cell is assigned a value of 1 to indicate a successful forecast. If no crime is recorded or the calculated probability was 0, it is assigned a value of 0. By repeating this process over successive weeks for each grid cell, this provides an indication of the predictability of a given location based on the distribution of successful forecasts. The grid cells with a larger number of 1 values throughout this trial period indicate consistent high risk and are more predictable. This is ranked by averaging the monthly values to get an average true-positive value for the grid cell, with cells having values closer to 1 being more predictable.

The state dependence component measures the number of recorded crimes for the most recent week for each grid cell to indicate the short-term risk for the forecast week. By using a two-step sorting process of the grid cells—first from highest to lowest average true-positive value to select the most predictable cells, then from highest to lowest recorded crimes in the current week—the model forecasts the most predictable locations with the highest short-term risk of a crime occurring (Lee, O & Eck 2020).

The assessment of forecasting performance was conducted as follows:

1. Create a grid over the study region. This grid is used to represent the risk surface. The cell sizes represent the level of resolution, with smaller cell sizes providing a finer level of resolution, while larger cells provide less detail, as forecasts cover a larger area by default. We used a regular grid of square cells of 140 metres by 140 metres, providing a high level of resolution for forecast areas. Hart and Zandbergen (2014) suggest grid size has very little impact on predictive accuracy. This cell size was retained across all three regions to allow clearer comparisons of forecasting performance.

2. Choose spatial and temporal bandwidths. Bandwidth corresponds to the amount of information used to estimate the probability surface. There are two decisions that influence surface estimation: how much historical data to include and over what spatial distance will points confer risk.

The choice of these values is straightforward for the prospective algorithm: it is simply the critical values of space-time dependency, as determined the by the previous Knox analyses. As mentioned, these values demonstrated some variance across the training period; therefore, the two bandwidth values are dynamically updated for each forecast year based on the results for the previous year.

Choosing appropriate values for the retrospective KDE surface is less straightforward. The ideal parameter values would be those used by practitioners. It is possible to obtain these values through interviews, but data gathered this way are likely to be highly individualistic. We are aware there are many ways of computing a bandwidth for KDEs, so much so that practitioners may be overwhelmed by the volume and inconsistency of advice. In the absence of practitioner insight and given the myriad options, we used the advice of Williamson et al. (1999), who suggest that the average nearest neighbour distance should be used to inform bandwidth selection.

For the temporal bandwidth, our experience suggests that analysts typically do not use 12 months of data, but generally work with three or four months of data. Aligned with our experience, we used three months of data for the temporal bandwidth.

3. Choose decay parameters. For the prospective and retrospective KDE, every point falling within a given bandwidth is weighted by proximity to the corresponding cell midpoint. We chose a linear inverse distance weighting, as such an event occurring at or very near the cell midpoint would be counted as one event, while an event occurring 190 metres from the midpoint (with a bandwidth of 200 metres) would be counted as 0.05 of an event.

The prospective KDE also weights events temporally, placing greater emphasis on events that occurred more recently. Here we chose a linear inverse distance weighting, with events occurring immediately prior to the forecast period weighted one, with the weighting decreasing linearly out to the temporal bandwidth, at which point the weighting becomes zero.

We investigated a range of weighting functions and decay values but observed no significant changes across any combination.

4. Create forecasting periods. The testing period was 2011–16. To assess forecasting accuracy we created 313 forecasting periods, one for each week of the testing window. For each of these periods, we stored four sets of events for each grid cell: (i) the events taking place 12 months prior to the forecasting date (for the population heterogeneity component of the YJL algorithm); (ii) the events taking place one week prior to the forecasting date (for the state dependence component of the YJL algorithm); (iii) the events taking place within the temporal bandwidth (for the retrospective and prospective KDEs); and (iv) the events taking place one week after the forecasting date—these are the events we are trying to predict.

5. Create risk surfaces for each period. Here we apply the KDE algorithm to produce a retrospective surface and a prospective surface. The calculations are identical except the events are time weighted for the prospective version. The YJL algorithm is also applied to generate an average true-positive value for each grid cell, indicating how predictable it is.

6. For the retrospective and prospective forecasts we order cells according to the forecast risk. Our ranking works such that the cell with the highest intensity estimate has a rank of 1 and lower-intensity scores have higher ranks. We also use a cumulative distribution ranking, such that cells with tied intensity values are rounded up to the maximum rank value (eg the vector {10,9,8,7,7,7} would have rank values of {1,2,3,6,6,6}). The YJL forecast has a two-step ranking system in which the cells are first ranked by the average true-positive values, indicating the most predictable to least predictable cells. The cells are then ordered within these rankings by the number of crimes that occurred in the week prior to forecast. Cells that have a high level of predictability and a high number of crimes within the week prior to forecast are therefore ranked higher than crimes with the same level of predictability, but fewer crimes in the week prior to forecast.

7. Count the number of ‘future’ crimes located in each cell.

We repeated the above procedure for all three crime types for each of the 313 forecasting periods and three study regions.

Metrics

Three metrics were used to assess the performance of the prediction algorithms: hit rate (HR), predictive accuracy index (PAI) and recapture rate index (RRI). These are described in detail below.

Hit rate

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. HR equals the number of crimes in areas of interest divided by the total number of crimes, where ‘areas of interest’ are the hot spots or identified forecast areas.

This metric has the advantage of being easy to understand: the higher the HR, the more accurate the predicted risk surface. However, it does not account for the size of the prediction area used. Defining large proportions of an area as ‘hot’ will capture the majority of crime but will be infeasible from an operational perspective.

Predictive accuracy index

The PAI was developed by Chainey and colleagues in 2008 (Chainey, Tompson & Uhlig 2008a, 2008b, 2008c) as an attempt to address the issue of the size of predicted crime hot spots. It takes into account the hit rate, the size of the study area and the size of the predicted hot spot area, using the formula $PAI_t = HR_t/(HSA_t/TA_t)$, where HR_t is the hit rate at time *t*, as described above; HSA_t is the total area of the study region identified as crime hot spots at time *t*; and TA_t is the total area of the study region at time *t*.

Recapture rate index

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 ($time_{2)}$ to the hot spot density at a previous time ($time_{1)}$. It is a simple ratio, with higher values reflecting greater consistency or reliability. $RRI_p = PAI_p/PAI_h$, where PAI_p is the PAI at the present time and PAI_h is the PAI at a previous time. In this study we considered RRI from week to week, comparing PAI at week *n* to PAI at week *n*−1.

Results

The results presented in this section are based on selecting the top one percent of grid cells by risk value for the three forecasting algorithms. We also performed the analyses using other values for internal robustness, and these are briefly described at the end of the *Results* section.

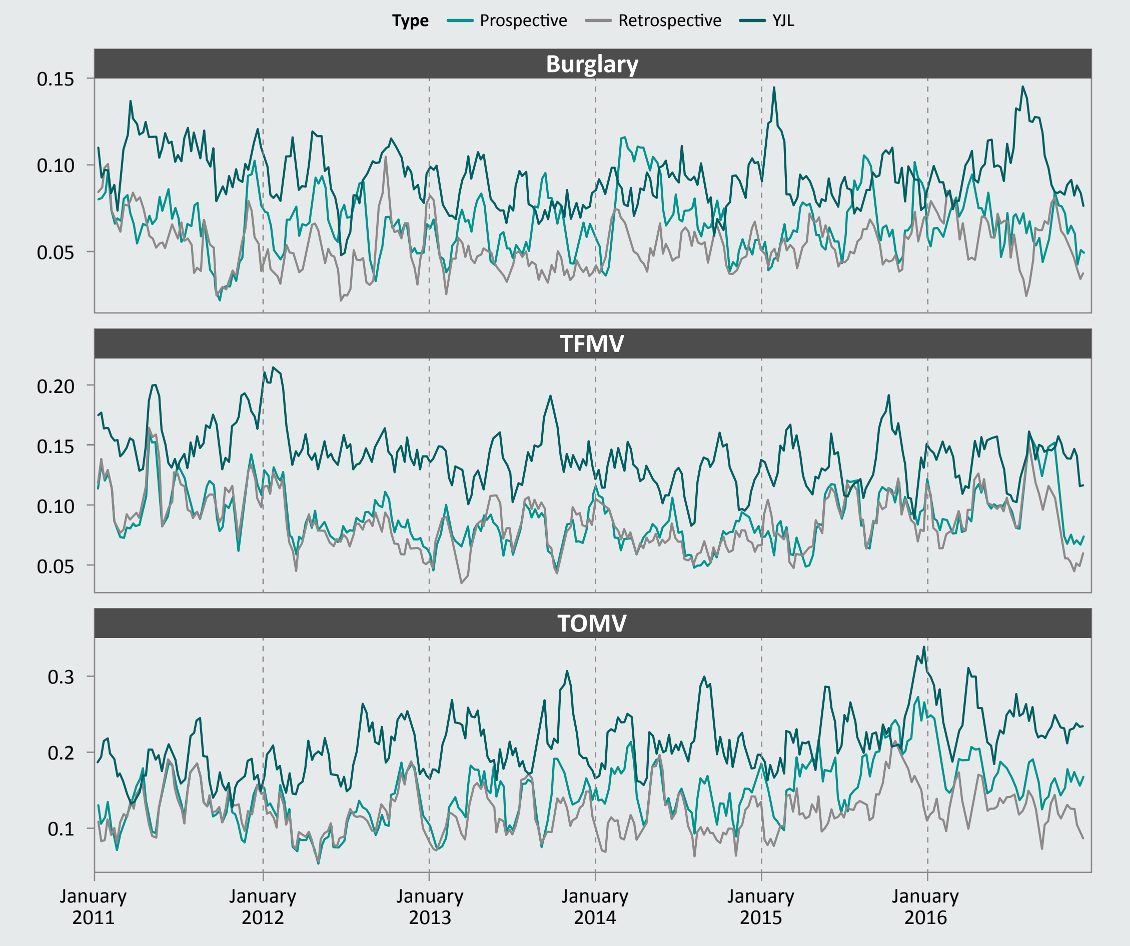
Assessing overall performance

It is cumbersome to interpret the individual results for all 313 forecast periods in the test set across the three crime types for Brisbane, Logan and Townsville. To get an overview of forecasting performance throughout the test period we can assess performance through the HR, PAI and RRI metrics.

Brisbane

The HR for each crime type in Brisbane is displayed in Figure 19. The YJL model routinely had the highest forecast HR for each crime type across the testing period. For burglary, the prospective method frequently had a slightly higher HR than the retrospective method; however, there were several peaks where the difference between these two methods became quite pronounced. The most notable of these occurred in the third quarter of 2013 and then in the second quarter of 2014, with the prospective method also outperforming the YJL model for these periods. Both KDE measures had very similar performance forecasting for TFMV offences throughout the forecast window. When looking at TOMV offences in the first three years of data, the two KDE methods demonstrated very similar performance. In the last three years, however, the prospective method outperformed the retrospective surface in almost every forecast period.

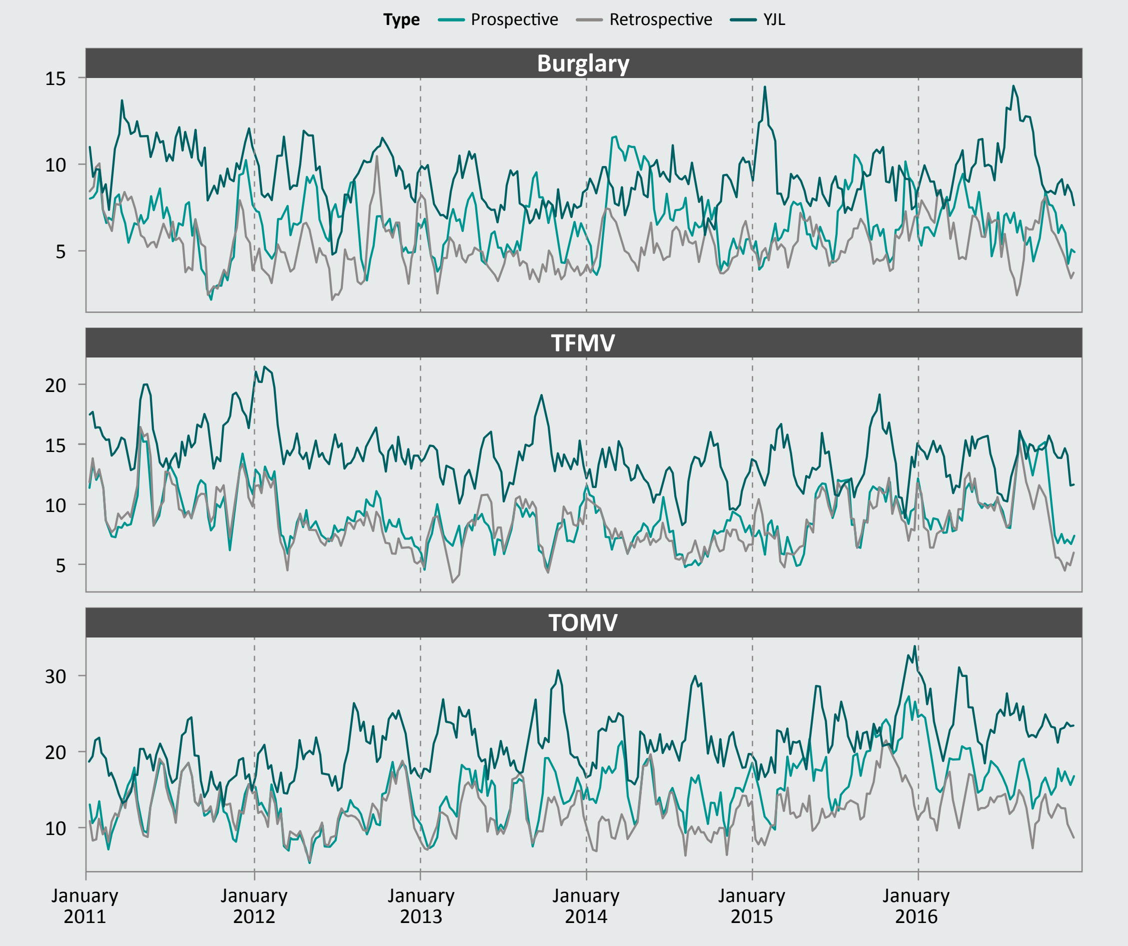
Figure 19: Forecasting hit rate for each crime in Brisbane across the entire test period



Note: The dashed lines indicate each year of the testing period

In Brisbane the YJL model made more accurate predictions than either of the other methods for all three crime types (see Figure 20). It regularly captured a greater number of crime events with the same size prediction area, regardless of the offence type. Overall, the prospective algorithm was more accurate in predicting burglary and TOMV offences than the retrospective algorithm, though both had very similar accuracy when forecasting TFMV offences.

Figure 20: Predictive accuracy index for each crime in Brisbane across the entire test period



Note: The dashed lines indicate each year of the testing period

Turning to the consistency of predictive accuracy, Figure 21 shows the RRI metric for Brisbane. Compared to the other study areas, Brisbane has the most stable predictive accuracy across the different crime types and forecasting methods. The retrospective method was the only forecasting algorithm that exhibited evidence of large changes in PAI from one forecast period to the next, and this was only observed for TFMV and TOMV offences. The consistency of PAI is likely a reflection of the higher recorded crime counts in Brisbane compared to the other regions.

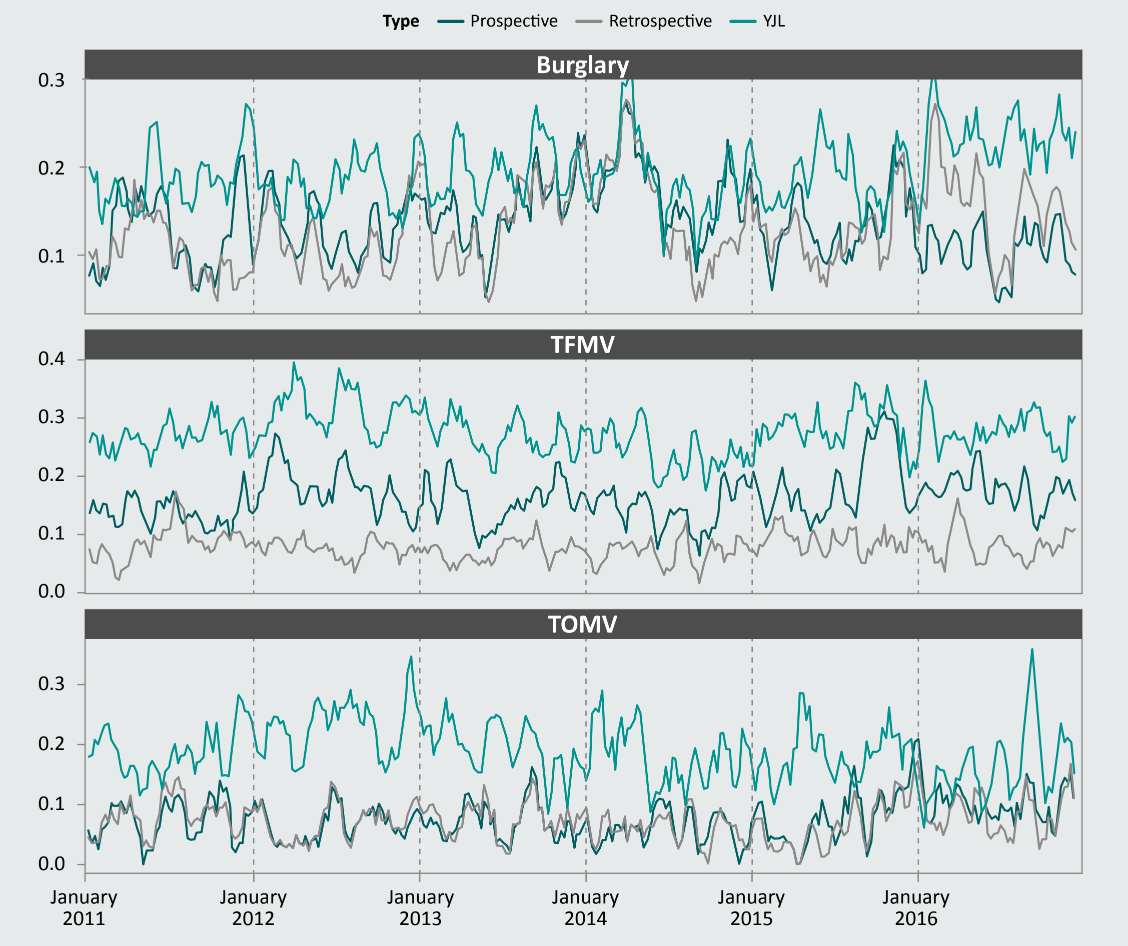
Figure 21: Recapture rate index for each crime in Brisbane across the entire test period

Note: Scales cropped to better display weekly fluctuation around one. The dashed vertical lines indicate each year of the testing period. The dashed horizontal line provides a reference for the RRI relative to a score of one

Logan

Figure 22 shows the HR for each crime in Logan across the entire test period. The HR has been smoothed using a four-week rolling average to assist with interpreting the results.

Figure 22: Forecasting hit rate for each crime in Logan across the entire test period



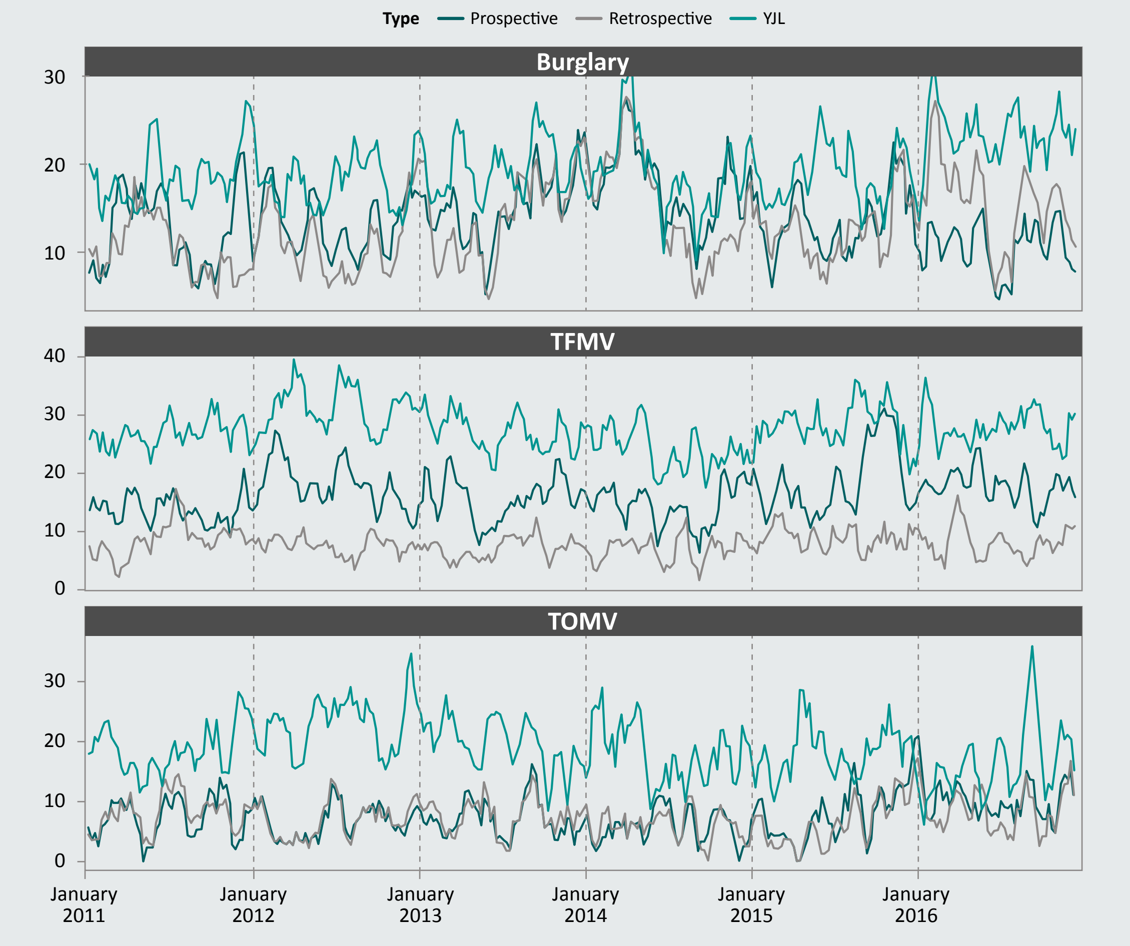
Note: Dashed lines indicate the start of each year

The YJL surface consistently captured more crime events in the forecast periods than either the prospective or retrospective surfaces. This is particularly evident for TFMV and TOMV where, apart from a small number of forecast periods, there is a notable separation between the HR lines. Burglary offences do not have the same level of separation between the different risk surfaces, although the YJL surface does outperform the two KDE-based surfaces overall. When looking at the two KDE surfaces, the prospective surface reliably captured more TFMV offences than the retrospective surface. For the burglary offences the prospective and retrospective surfaces captured a similar amount of crime events in most forecast periods but demonstrate substantial divergence at several points during the testing period. This is most noticeable towards the end of 2011, towards the end of 2014, and in the first half of 2016.

In the first two instances the prospective surface captured more crime events but in the first half of 2016 the retrospective surface outperformed the prospective surface. The prospective and retrospective methods were largely equivalent in their forecasting of TOMV offences, which was the lowest-volume offence.

Next, we turn to the PAI. This metric takes into account the size of the area used to define a hot spot. Figure 23 shows the PAI (smoothed with a four-week rolling average) for each crime in Logan across the entire test period. As these results are based on selecting the top one percent of grid cells as hot spots (the hot spot area), and the denominator for calculating the PAI is the hot spot area divided by the total area, in this situation the PAI amounts to dividing the HR by 0.01. As a result, this figure mirrors the patterns demonstrated in the HR figure above.

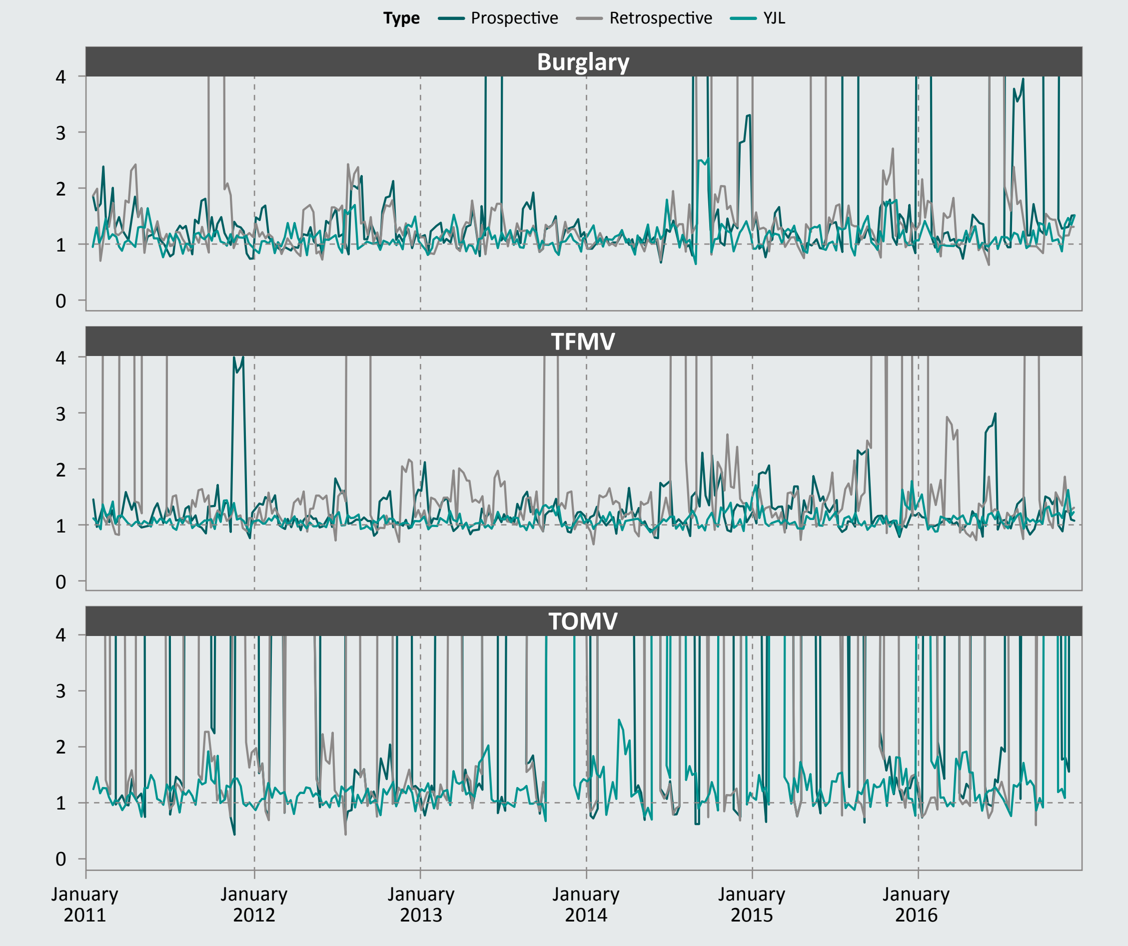
Figure 23: Predictive accuracy index for each crime in Logan across the entire test period



Note: Dashed lines indicate the start of each year

Figure 24 shows the four-week rolling average RRI for each crime in Logan across the entire test period. An RRI value of more than one indicates an increase in predictive accuracy between forecast periods, while a value below one indicates a decrease. To assist with viewing this fluctuation around one in the figure, the RRI scale has been zoomed in. Burglary and TFMV offences both demonstrate a high level of stability across the three algorithms, as shown by the RRI value fluctuating closely around one, while forecasting performance for TOMV offences was less consistent. As seen in the figure, the retrospective and prospective methods had a greater level of volatility for each offence type, with RRI values spiking up (at times beyond the scale shown). The RRI experiences these large spikes following a forecast period with a low PAI score. The YJL model demonstrated the most stable performance, with an overall average RRI of 1.2 for burglary, 1.1 for TFMV and 7.2 for TOMV. The overall average RRI for the prospective method was 3.3 for burglary, 1.3 for TFMV and 19 for TOMV. The retrospective method had averages of 3.5 for burglary, 5 for TFMV and 16 for TOMV.

Figure 24: Recapture rate index for each crime in Logan across the entire test period

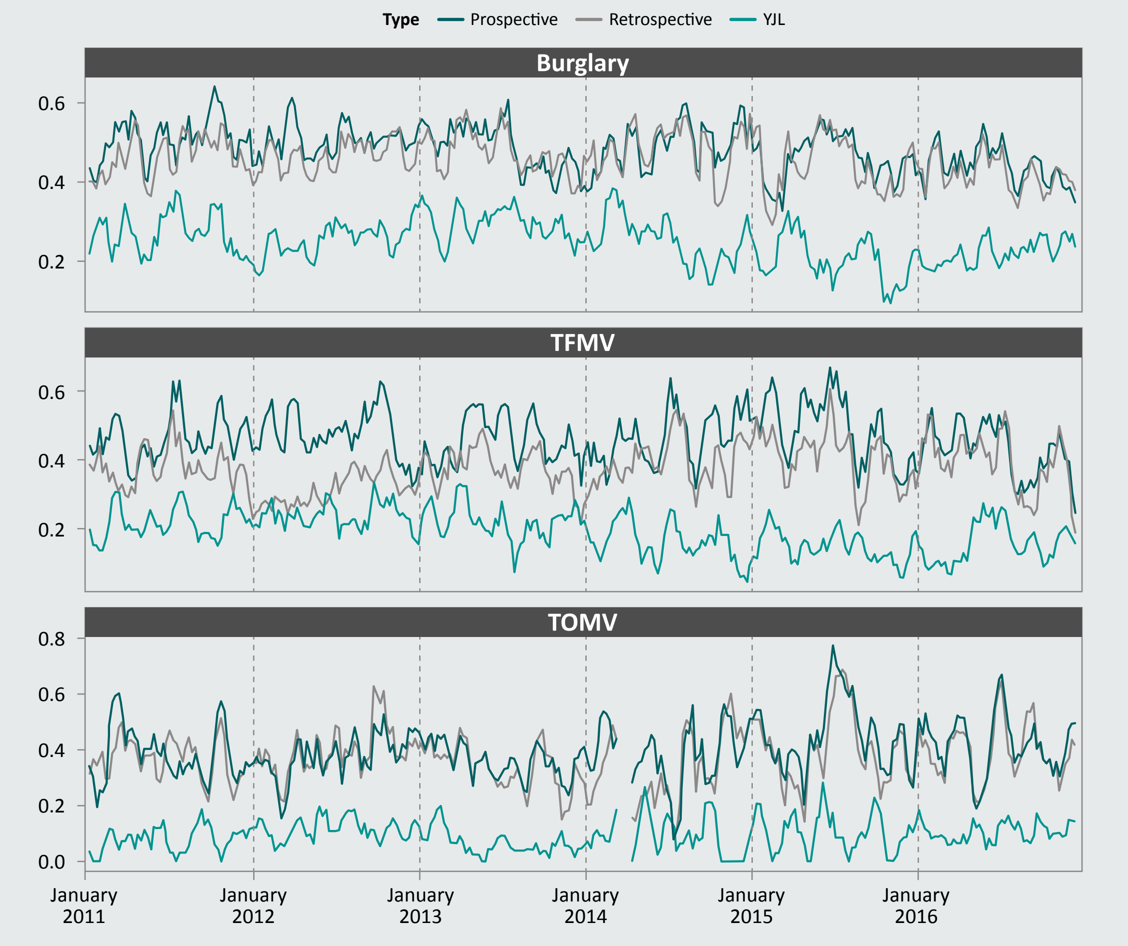


Note: Scales cropped to better display weekly fluctuation around one. The dashed vertical lines indicate the start of each year. The dashed horizontal line provides a reference for the RRI relative to a score of one

Townsville

Figure 25 shows the HR for each of the crime types in Townsville across the full test period. As in the previous figures, a four-week rolling average has been applied to smooth the graph. In contrast to the HR performance in Logan, the YJL model was consistently outperformed by the two KDE models in forecasting all three offence types in Townsville. The TFMV offence type was the only one that had a clear difference in HR performance between the two KDE surfaces, with the prospective method outperforming the retrospective method in most forecast periods. This is particularly evident from the second half of 2011 until the end of 2012. The two methods display very similar performance in forecasting burglary and TOMV offences, with the prospective method demonstrating slightly higher HRs on average. There was one week in 2014 in which there were no TOMV offences recorded in Townsville.

Figure 25: Forecasting hit rate for each crime in Townsville across the entire test period



Note: The dashed lines indicate each year of the testing period

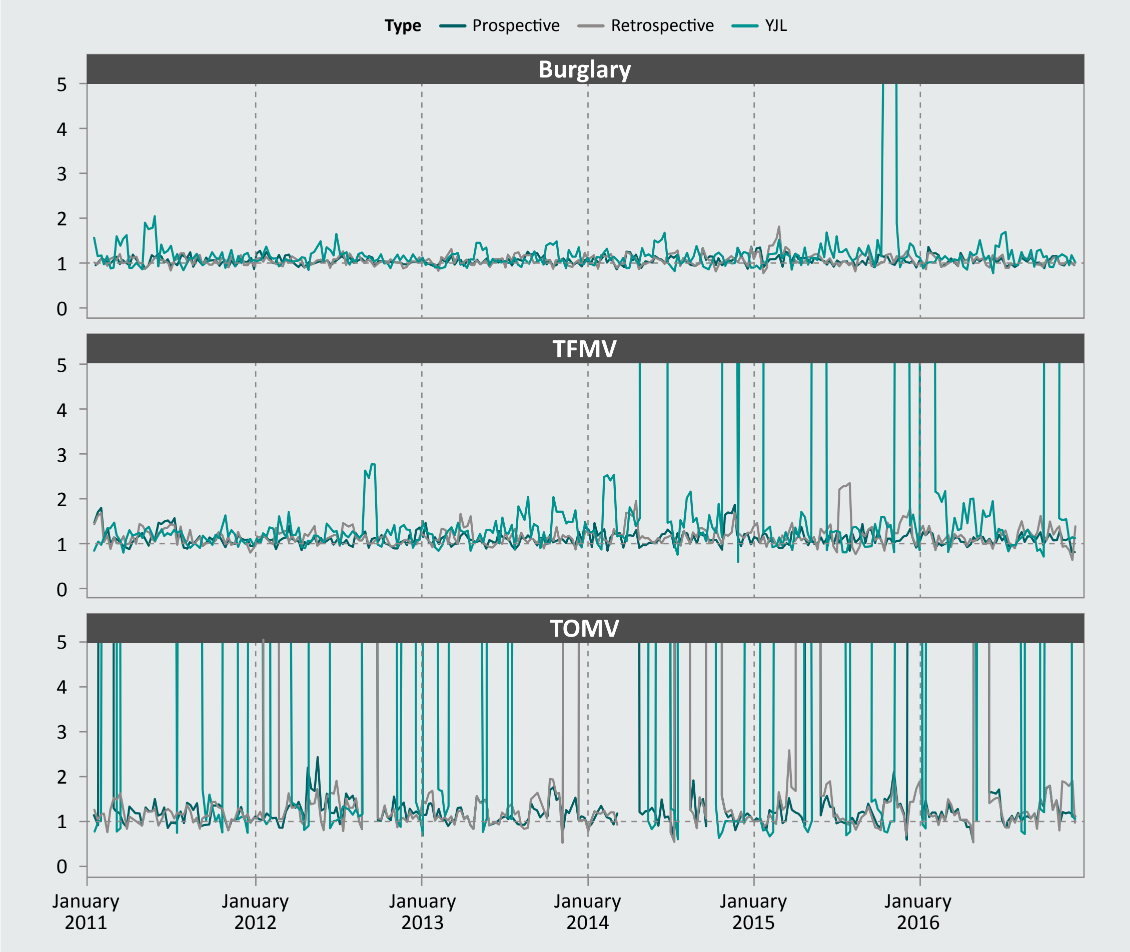
The PAI scores in Townsville for each forecasting method across the crime types is shown in Figure 26. As in the previous figures, a four-week rolling average has been used to smooth the results and aid interpretation. The YJL model consistently had the lowest PAI for each of the crime types. The prospective model outperformed the retrospective model in most forecast weeks for TFMV offences, although the two KDE models had similar performance in forecasting burglary and TOMV offences. The average PAI was slightly higher for the prospective model than the retrospective in both instances, with burglary averages of 48.2 (prospective) and 45.9 (retrospective), and TOMV averages of 39.4 and 37.1, respectively.

Figure 26: Predictive accuracy index for each crime in Townsville across the entire test period

Note: The dashed lines indicate each year of the testing period

Figure 27 displays the RRI for the different crime types in Townsville. The two KDE surfaces demonstrated stable, consistent performance in capturing crime events in their forecast areas for burglary and TFMV offences, as shown by the absence of large spikes in the RRI. The YJL method was slightly less reliable in its performance on burglary offences, with several small increases in RRI throughout the testing window and a large spike near the end of 2015. In the latter half of the testing window (from 2013 onwards) the YJL model was less consistent in its ability to accurately forecast TFMV offences, with six large spikes in the RRI suggesting the occurrence of weeks with very low PAI performance. TOMV offences proved difficult for all three algorithms to consistently forecast with accuracy, as evidenced by the large number of RRI spikes for all three methods. This offence type was the least frequent in each region, and Townsville recorded the lowest number of this offence type out of all the regions. The small sample size of crime events to draw upon for forecasts, and the low incidence rate in the forecast windows, may account for this poorer reliability.

Figure 27: Recapture rate index for each crime in Townsville across the entire test period



Note: Scales cropped to better display weekly fluctuation around one. The dashed vertical lines indicate each year of the testing period. The dashed horizontal line provides a reference for the RRI relative to a score of one

Summarising metrics across all regions and crime types

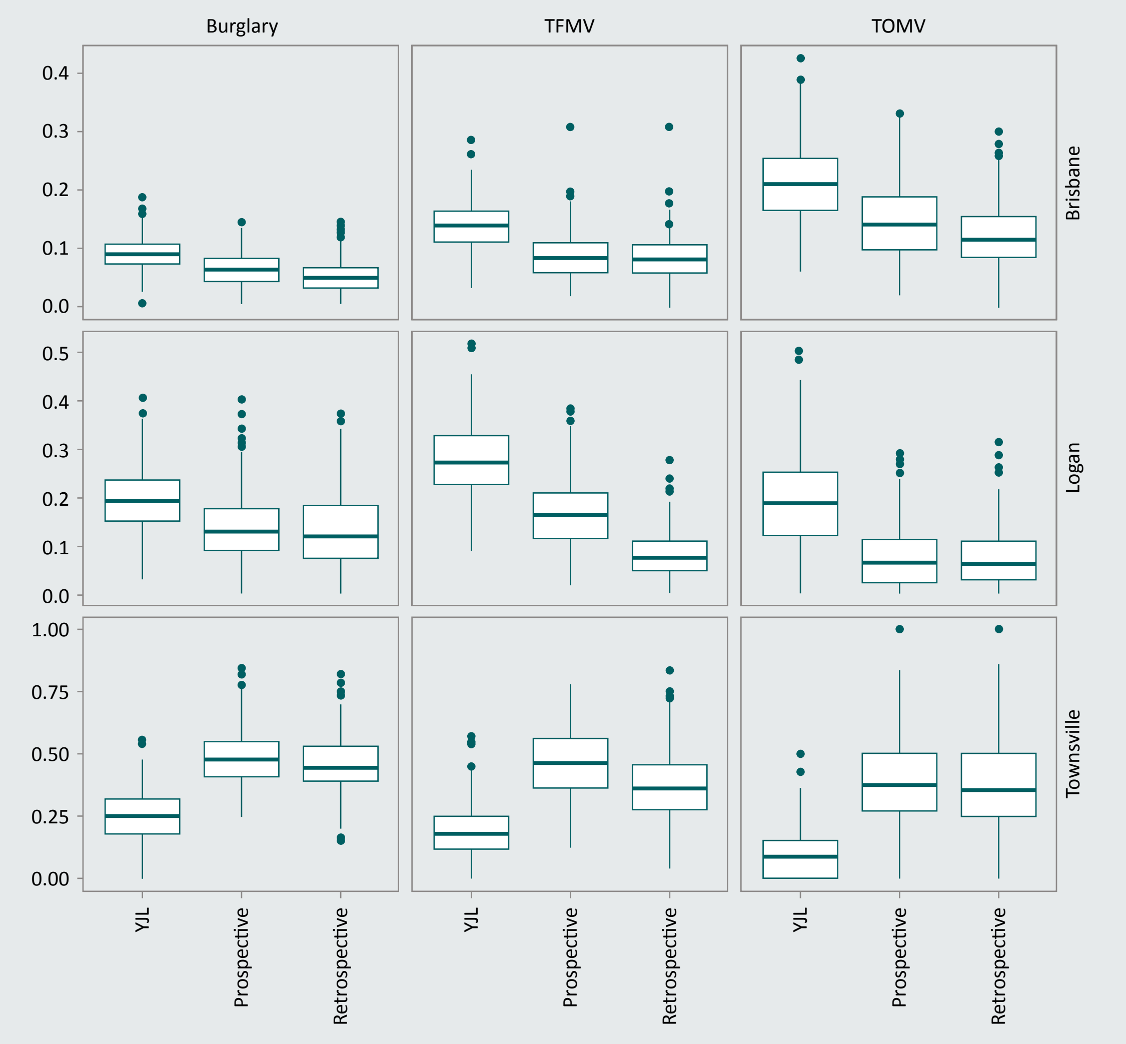
Computing aggregate values for HR, PAI and RRI across the test period indicates that the performance of the three methods varies across region and crime type. The boxplots below summarise the weekly performance of the three methods using each of the three metrics.

The aggregate HR boxplots (Figure 28) show that the YJL model outperformed the prospective and retrospective surfaces in every crime type in Brisbane and Logan. In contrast, the YJL model had the lowest performance of the three algorithms for all three crime types in Townsville. This may be due to a combination of the lower count of each crime type in Townsville, the level of spatial concentration of offences, and how the YJL model is calculated. This algorithm is based on crime 'hits': 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, then the number of grid cells with recorded crime hits is limited, which in turn limits the number of cells the model can predict, affecting the accuracy of the HR.

Brisbane burglary offences had the tightest spread of scores for each forecasting method, as evidenced by the small range of the box and whiskers (representing the interquartile range (IQR) and 1.5 times the IQR), indicating the HR scores were consistent across the forecast periods. Townsville TOMV offences had the largest spread of scores, indicating that despite these having some of the highest median HRs, they also had the greatest level of fluctuation across the forecast weeks.

For each crime type, Townsville demonstrated the highest median HR of the three regions for the retrospective and prospective methods, despite having the lowest counts for each crime type. In contrast, Brisbane, which had the highest counts for each crime type, demonstrated the lowest median HRs for burglary offences, and the lowest median HRs for the YJL and prospective methods for TFMV offences. This contrast between the volume of offences and HR may relate to the spatial concentration of offences. The hot spot maps provided earlier for each region (Figures 7, 8 and 9) demonstrated very different concentrations of offences. The hot spot map for Brisbane highlighted a high concentration around Brisbane City (in the centre of the map), however there was a broad diffusion of lower concentrations of offending throughout a large area. In contrast to this, Townsville had a highly concentrated pattern of offending, with a small area of high-intensity offending and a small 'buffer' around this area consisting of a lower intensity of offending. Logan was somewhat in between these two, with clear higher intensity hot spots, and a diffusion of lower-intensity offending spread around these. This highlights the importance of tailoring the parameters (and the forecasting method) for the location of interest.

Figure 28: Aggregate hit rate for each crime and study region and forecasting method

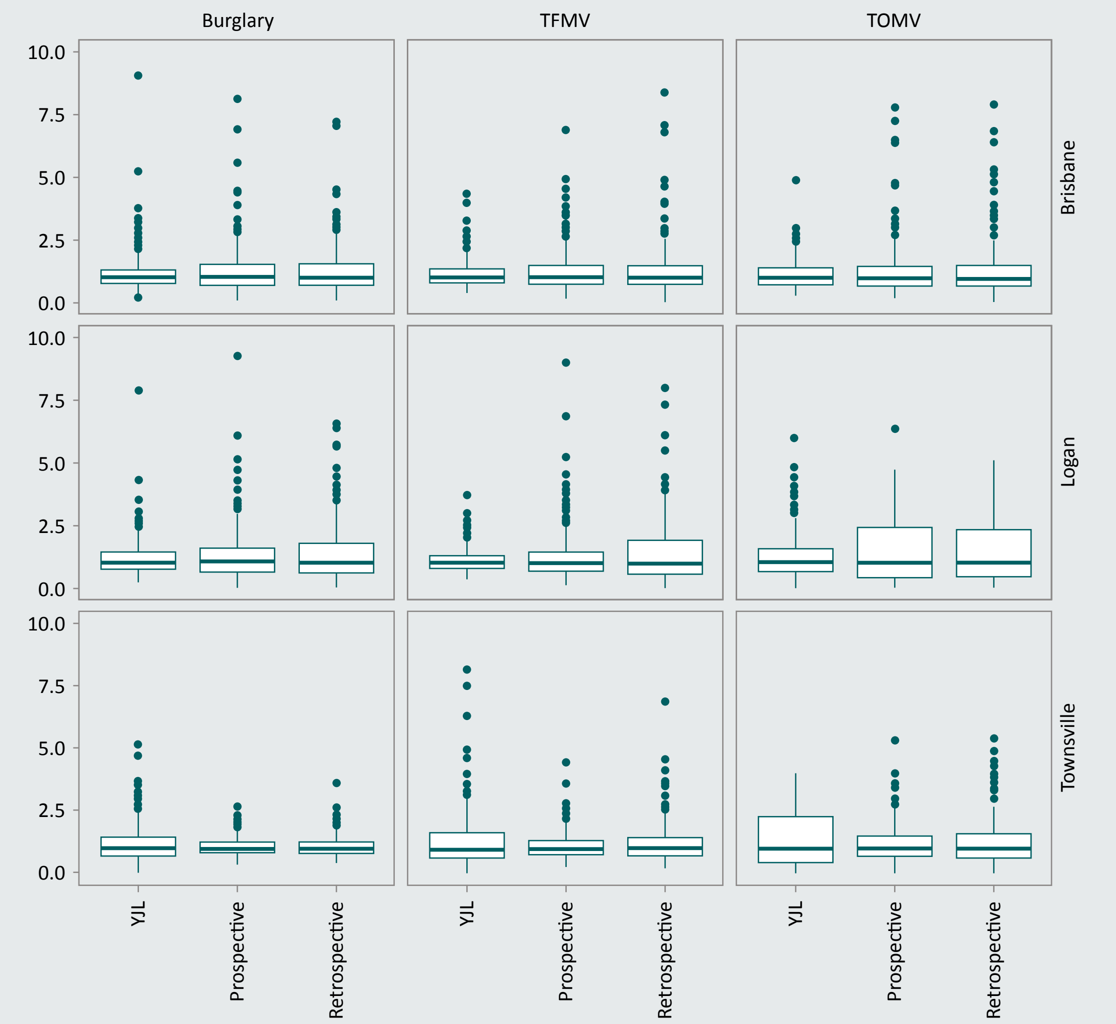


As mentioned previously, the PAI patterns mirror those displayed in the HR figures. Given this, rather than displaying the aggregate results for PAI we instead provide the PAI figures in the next section when discussing the differences in results when selecting the top one percent, five percent and 10 percent of grid cells as hot spots.

As Figure 29 shows, in each of the crime type and region combinations all three forecast algorithms demonstrated similar median RRI values. This suggests that, overall, they are equally effective at consistently encompassing future crime locations in the forecasts. The main finding of interest in Figure 29 relates to the consistency of RRI scores, indicated by the relative spread of the boxplots throughout the different dimensions of comparison.

In Brisbane and Logan, the YJL model had the smallest IQR across the different crime types (with the exception of TOMV in Brisbane), indicating this forecasting model had the greatest week-to-week consistency in its forecast accuracy. The two KDE methods had a similar IQR in all comparisons except TFMV in Logan, in which the retrospective method had a larger IQR, indicating it was less consistent. Despite these two methods having a similar IQR, the retrospective method had a greater number of outliers (indicated by the dots) at a greater range in most comparisons (many of the outliers vastly exceeded the scale provided in Figure 29). The exceptions to this are burglary offences in Brisbane and TOMV offences in Logan, for which the prospective method had outliers at a greater range. This result is consistent with the RRI figures for each region displayed previously, in which the retrospective method demonstrated more frequent spikes to high RRI values than the prospective method. TOMV offences demonstrated the greatest spread of scores among the three crime types, suggesting it was the least consistently predictable crime type, regardless of the region.

Figure 29: Aggregate boxplots of recapture rate index for each crime, study region and forecasting method



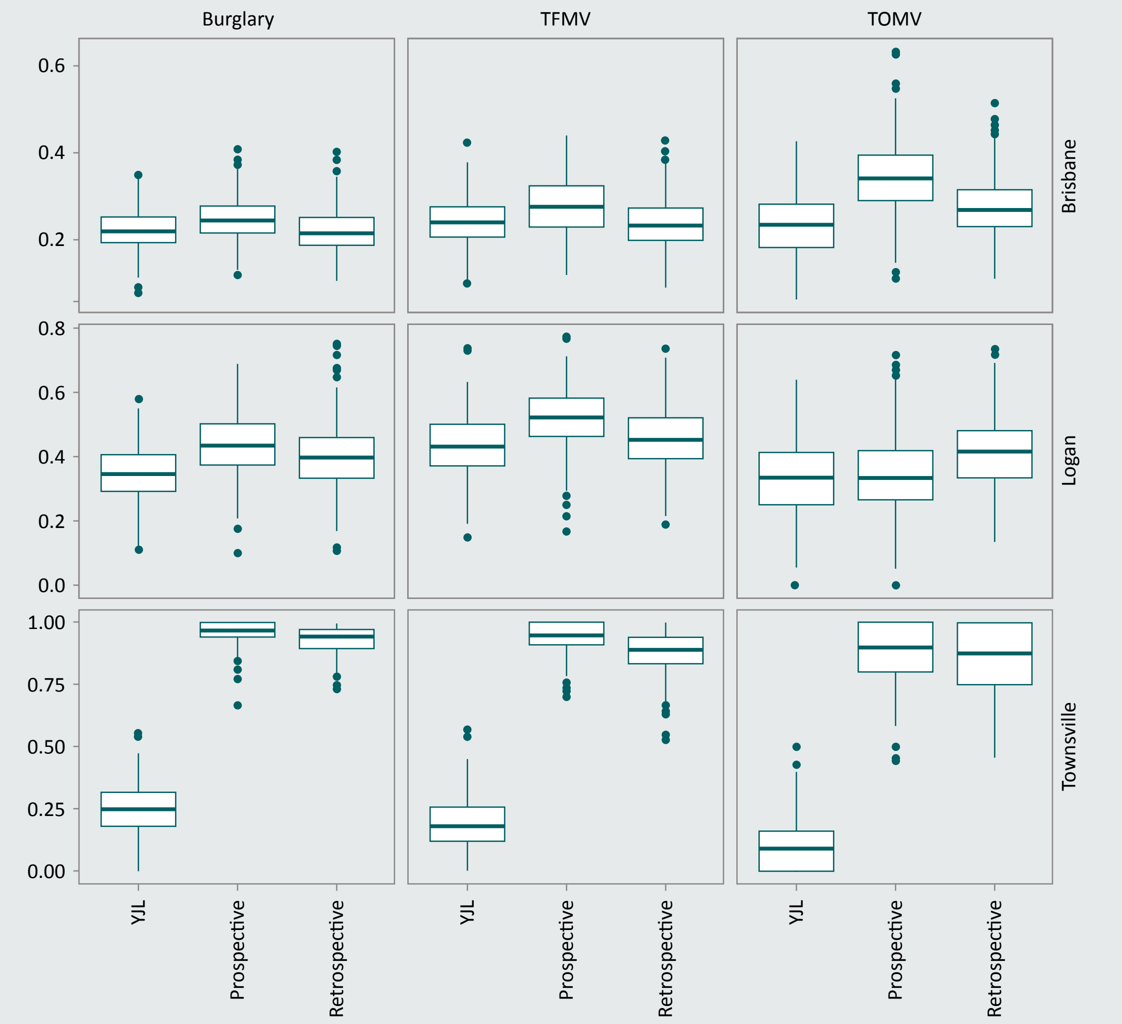
Additional considerations

For the purpose of internal robustness some additional analyses were performed, although not all of these warrant inclusion in this report. For instance, serial correlation analyses were performed on the different forecasting methods and the YJL model was run using both weekly and monthly forecast periods, as the original research proposing this algorithm (Lee, O & Eck 2020) provided forecasts on a monthly basis.

As mentioned at the beginning of the *Results* section, the results presented above represent the selection of the top one percent of forecast grid cells based on risk. The analyses were also performed using the top five percent and top 10 percent of forecast cells, even though it would likely be infeasible to target such a fraction of a study region. Selecting a larger number of the highest-ranked cells typically increases the HR observed for the forecasting method, but this is often at the cost of a reduction in the PAI as the size of the hot spot area increases. For the allocation of crime prevention resources, this becomes a balancing act of optimising both the HR and the PAI for the target area. This optimisation process goes beyond the scope of the current research, which simply aims to determine the feasibility of predictive forecasting methods in an Australian context, and as such only tests the three hot spot area selections.

When increasing the selected hot spots from the top one percent of cells to the top five percent of cells several changes to performance metrics were observed. The HR results for the top five percent of cells are shown in Figure 30 below. With the larger number of grid cells being included as forecast hot spots there was a notable increase in the observed HR for most forecasting methods across the different crime types and regions. The prospective method had the highest HR across most of the comparisons. In Townsville both the prospective and retrospective forecasting methods surpassed a median HR of 80 percent for all three crime types. The YJL method, however, did not receive an uplift in performance, suggesting the number of cells assigned a forecast risk value did not exceed one percent of the total number of grid cells.

Figure 30: Aggregate boxplots displaying the hit rate for each crime and study region when selecting the top 5% of forecast grid cells as hot spots



While increasing the percentage of grid cells classified as hot spots increased the HR, this also decreased the PAI in most instances, as shown in Figures 31, 32 and 33. To illustrate, in Townsville the prospective method had a burglary HR increase from just under 50 percent to near 100 percent, while the median PAI decreased from approximately 50 to 20 (as shown in Figure 33 below). There were a few exceptions to this, in which increasing the grid selection from the top one percent to the top five percent resulted in an increase in both the HR and the PAI, all of which occurred in Logan (Figure 32). In forecasting TOMV, both the prospective and retrospective methods had a small increase in the median PAI, and for TFMV offences the retrospective method also had a small improvement. This indicates that the improvement in the HR exceeded the percentage increase in hot spot area.

Similar patterns were evident when selecting the top 10 percent of grid cells as hot spots, with the retrospective and prospective methods having a further uplift in HR performance (results not shown), while the YJL model demonstrated very little improvement. As expected, the PAI suffered a decrease in performance, with most combinations' values reducing by approximately half.

Figure 31: Aggregate boxplots displaying the PAI for each crime type and forecasting method in Brisbane across 1%, 5% and 10% of grid cells selected as hot spots

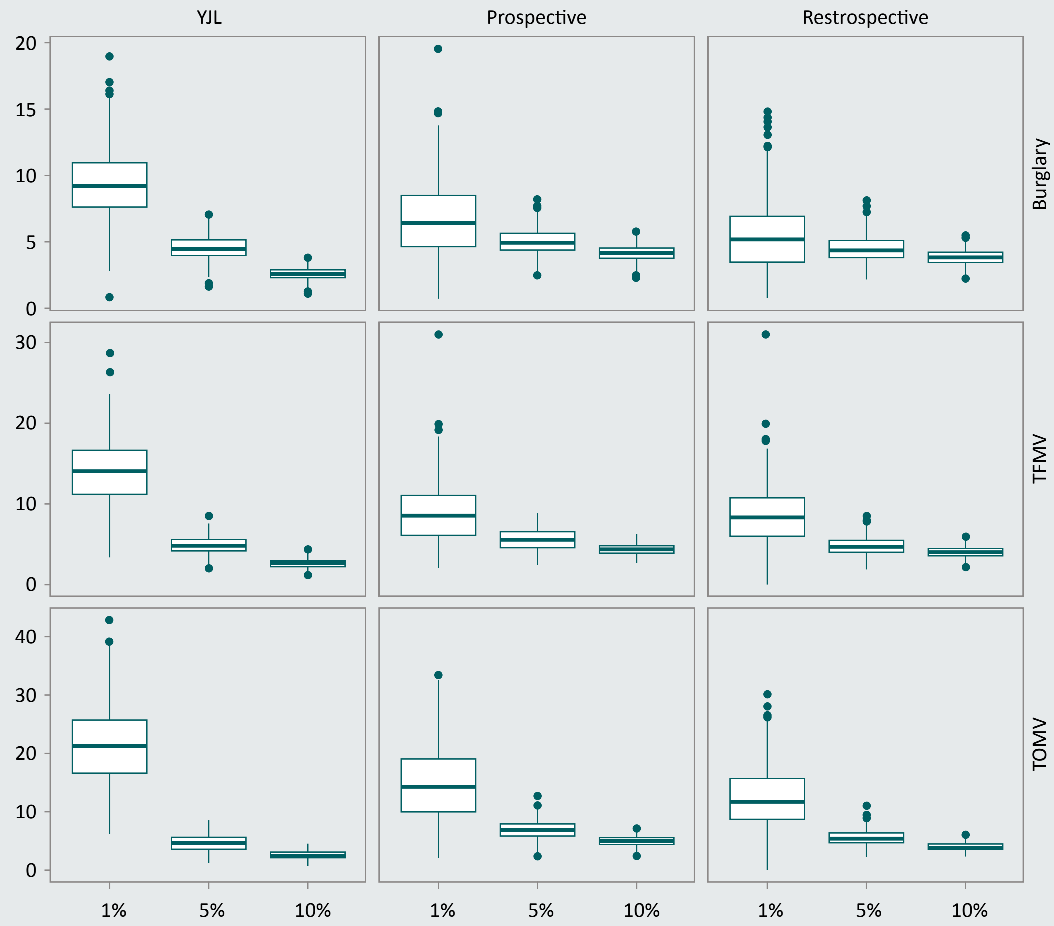


Figure 32: Aggregate boxplots displaying the PAI for each crime type and forecasting method in Logan across 1%, 5% and 10% of grid cells selected as hot spots

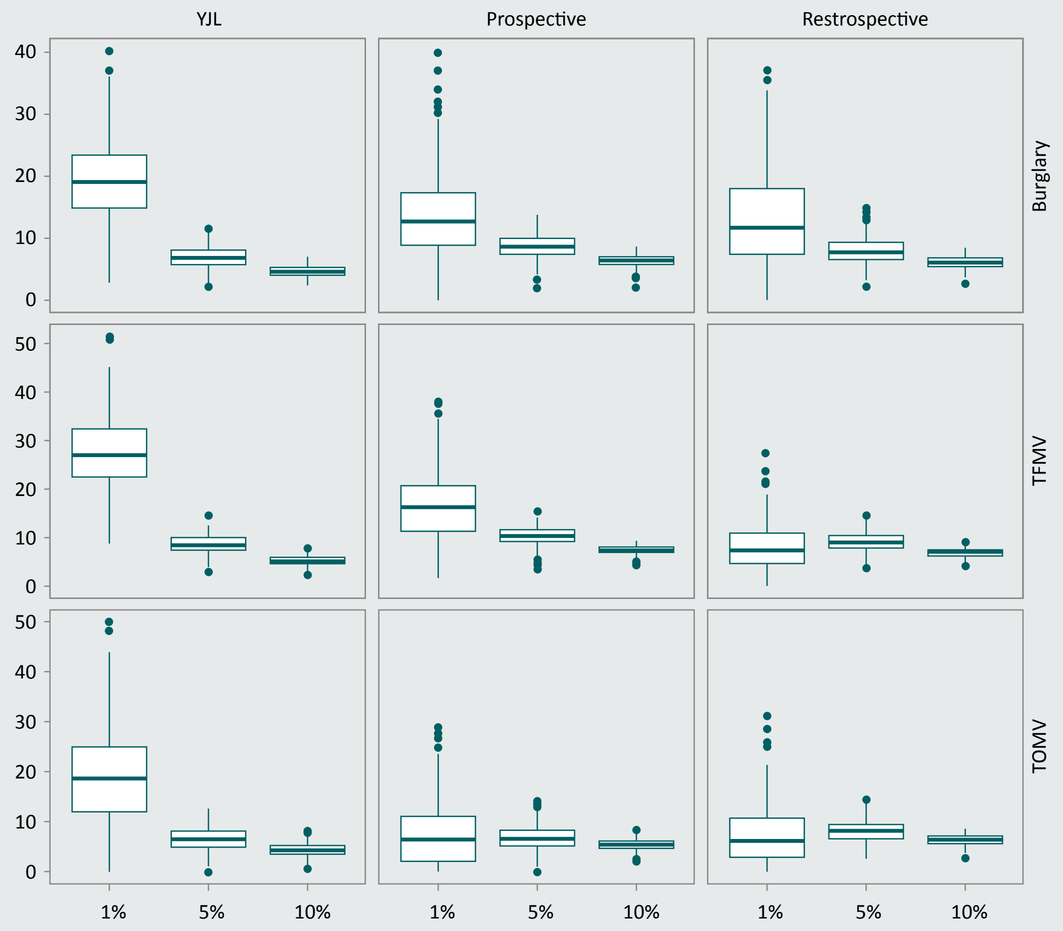
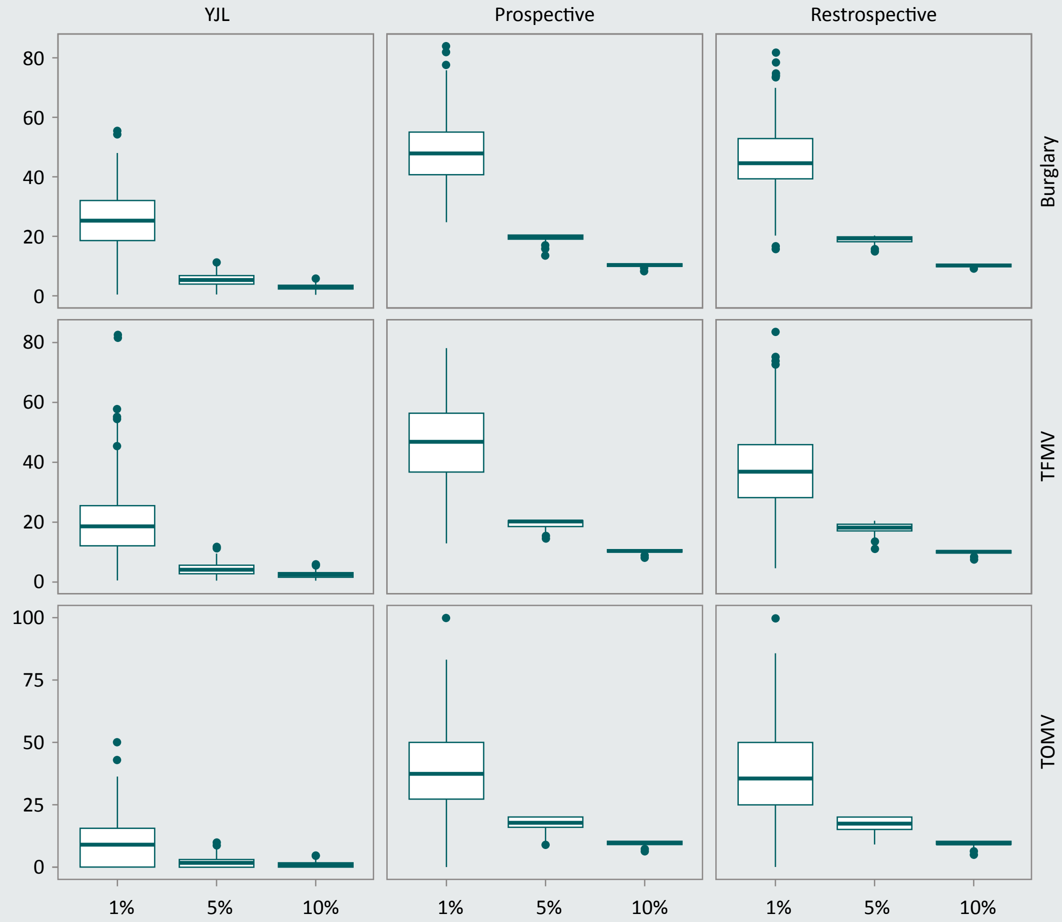


Figure 33: Aggregate boxplots displaying the PAI for each crime type and forecasting method in Townsville across 1%, 5% and 10% of grid cells selected as hot spots



Discussion and conclusions

This study identified space-time patterns in historic crime data and, from these, used two methods (the prospective method and the YJL method) to predict likely crime locations in the immediate future for three crime types in three distinct study regions. An alternative prediction method emulating conventional 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.

There were a few key differences between the methods used here: (i) the YJL method had reasonably long lookback periods compared to the two KDE methods; (ii) the prospective method weighted events by distance *and* recency—the retrospective method weighted only by distance; and (iii) the YJL method instead weighted grid cells first by predictability and then ranked these by the recent volume of crime occurrence. It seems intuitive that a method that relies on more data should be able to generate better forecasts. This is true if the phenomenon is stationary. However, our results suggest that, at least for the regions and crime types in this study, this is not the case for all combinations of crime types and regions.

There also appear to be several factors that influence how well the different methods can forecast crimes. The type of spatial concentration appears to play a non-trivial role in how effective the different methods were in forecasting crime, as seen by the differences in performance across the three regions. The YJL model performed well in locations in which the hot spots could be classified as either dispersed or clustered (such as Brisbane and Logan) according to Ratcliffe's (2004) hot spot matrix. When the spatial concentration was very high, as in Townsville, the YJL method did not perform as well. In contrast, the prospective and retrospective models both demonstrated strong HRs in Townsville, but weaker performance in Brisbane and Logan, which had more diffused spatial concentration levels. This likely reflects the key differences described above. Greater levels of spatial concentration are beneficial for the prospective and retrospective methods, as they weight events by distance. The YJL model in its first step weights grid cells by the consistency with which crimes occur in them, and cells can only be selected as a hot spot if crimes have been recorded there recently. If crime is highly spatially concentrated, then the possible number of grid cells that can receive a risk weighting is limited, reducing the possibility of capturing crimes that occur nearby, but not in a previously recorded cell.

In looking at predictive accuracy, we compared several metrics. As noted earlier, the HR is the percentage of crimes that fall within zones predicted to be at high risk of events occurring and does not take into account of the size of the prediction area used. The PAI addresses the issue of the size of predicted crime hot spots, taking into account the HR, the size of the study area, and the size of the predicted hot spot area.

The prospective method had a larger median HR than the retrospective, but the range of performance still overlapped considerably. The largest difference in performance occurred when forecasting TFMV offences in Logan, in which the 25th percentile performance of the prospective method was greater than the 75th percentile of the retrospective surface. The PAI metric mirrored this performance, again by virtue of selecting the top one percent of grid cells for classification as hot spots.

As a preliminary study, our results appear to offer findings that quite strongly support adopting a predictive policing approach, or at least the predictive component of this. Disaggregating the results reveals 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 performance, especially for TOMV. This can also be seen when looking at the range of the box and whisker plots in Figure 28, with the Brisbane metrics evidencing a tighter spread of values. A number of candidate explanations suggest themselves. First, Brisbane had the highest sample size of the study regions, providing more data points to draw from (regardless of the lookback period used for forecasting), while Logan and Townsville both had smaller sample sizes, and less data with which to refine the predictions. Supporting this assertion, at all three locations TOMV was the least frequently occurring offence and displayed similar (albeit more modest) predictive performance. Second, the forecasting horizon is reasonably short, using 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) rather than maximising predictive fit.
- While Brisbane demonstrated the most consistent performance, it also generally demonstrated the lowest HRs. There are a couple of possible explanations for this. The first one we offer relates to the interplay between area size crime levels, crime opportunities and crime dispersion. Brantingham (2016) supported the idea that larger areas can host a greater diversity of environmental settings and, as such, provide opportunities for more types of crime. Similar to this argument, one would expect that a larger area of a certain type of environmental setting would also provide more opportunities suitable for *particular* types of crimes. As mentioned above, more dispersed crime can impact the predictive performance of distance weighted forecasting methods, and Brisbane had the highest recorded crime counts and the greatest level of spatial dispersion. A second possible explanation that links to this is the type of land use and the distribution of crime opportunities. Brisbane has the largest proportion of its land area that is built up or developed, followed by Logan, and then Townsville. This has implications for the

opportunity structure for these offences, particularly burglary, and this is reflected in the hot spot maps presented earlier in the report (Figures 7, 8 and 9). The hot spots are clustered around the population centres with varying intensities. In Brisbane the urban spread covers much of the centre of the map. In Logan the majority of the population is concentrated in the northern section of the map, while Townsville has a highly concentrated population centre in one small area. Burglary offences in Brisbane are generally dispersed somewhat evenly across a large, medium-intensity hot spot throughout the centre of the map. The hot spots for the two vehicle crime types cover a similar area; however, both have a high-intensity cluster centred around the Brisbane CBD, which may explain why these both had slightly higher median HRs across all three forecasting methods.

Like any preliminary study, we were not able to comprehensively explore all factors in our analysis. We briefly list these here and explain what next steps should be explored in order to increase the accuracy of crime forecasting.

First, time of day information should be incorporated in our predictions. Previous studies have generated forecasts based around shift patterns of the policing agency, as crime types often display distinct time-of-day signatures. Our analysis does not consider time of day, and this seems a straightforward enhancement to the approach described here.

Second, optimising by crime type should be considered. We took a measured approach to tailoring the algorithms to local conditions, such as allowing parameters to be updated annually. However, we did constrain the analytic set-up in a bid to compare forecasting performance across the crime type and study region combinations. More optimal models are certainly possible and there is a place for that type of study. This study provides 'good enough' justification for further exploration without risking overfitting. The potential number of combinations to be explored is considerable, and beyond the scope of preliminary work here. However, it seems there is much to be gained by customising our method and tailoring the forecasts to each offence type and study region.

Third, forecasting could be customised for locations within each study region. Again, we have used the same parameters for the entire region, but it is conceivable that these dynamics may vary across space, especially considering two of the study regions had extreme land use concentration. One area to explore could be partitioning a district into subregions and performing forecasts distinct to those locations, based on their own historic data.

Fourth, bandwidths could be determined dynamically. While we updated bandwidths annually, future research could investigate forecasting performance based on more frequently updated parameters. While our analyses did not identify seasonal trends, these may emerge in the future or be present in other regions.

Fifth, demographics of the study regions could be incorporated into the modelling. While we describe the differences and consistencies of the land use and demographic composition of the study regions, at no stage was this information used to inform the forecasts. As others have demonstrated (Taylor, Ratcliffe & Perenzin 2015), the distribution and magnitude of crime is a product of factors that operate at long- and short-term time horizons. We contend that the influence of long-term factors is included in our analysis by virtue of using historical crime data, which are themselves the product of long- and short-term factors. Nonetheless, directly exploring and testing this contention may reveal opportunities for further refinement and increased accuracy. It is particularly important to exercise caution given the well-acknowledged unfair and harmful discriminatory effects that can occur when incorporating variables that might reinforce over-policing through feedback loops (Jefferson 2018; Marjanovic, Cecez-Kecmanovic & Vidgen 2022).

Sixth, other algorithms could be used in future research. This study used two algorithms—the near-repeat phenomenon and the YJL algorithm—but there are other approaches that have been used to forecast crime (Kulldorff et al. 2005; Mohler et al. 2011). Our results suggest that different algorithms have varying performance depending on particular contexts (land use, regional vs metropolitan), and these should be explored. There has been little comparison of algorithm performance in the literature to date.

The preceding list of future directions notwithstanding, 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 policies requiring an 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 toward more affluent suburbs, which has implications for the distribution of police resources.

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 sorts of 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 relies on a narrative description of location, which is subject to interpretation. This is an obvious area of concern in light of the criminological literature's move towards smaller units of analysis.

Third, 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, in part, a result of police operations. Without information about police locations and operations during the outofsample 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, Chenery & Pease 1998) we are confident these patterns are a product of offender location choice. A randomised controlled trial, such as those described by Hunt, Saunders and Hollywood (2014), Mohler et al. (2015) and Ratcliffe et al. (2020), would be able to rule out this competing explanation.

In summary, this study supports cautious optimism for the prospects of predictive policing in Australia. We make the following four observations based on the findings and the empirical literature.

Predicting the most likely location for crime occurrence in the short term seems possible in Australia. The location of future crime was able to be predicted at higher rates (or equivalent rates with much less data) than current approaches, using simple methods or methods that could be implemented relatively easily. Moreover, this was achieved for multiple crime types and in different study regions.

Forecast parameters should be localised. Our results suggest that forecasts do require some degree of parameter localisation to improve predictive performance. Even though we observed local differences and dynamically updated the forecasting models, the algorithms showed variable performance across different combinations of crime types and study regions.

The second half of the predictive policing enterprise—crime reduction—has mixed evidence of effectiveness. This study only considered half of the predictive policing enterprise: prediction. The next step is to consider and design effective tactical responses to preventing these types of crimes based on the identified patterns. As stated earlier, the accurate forecasting of future crime locations does not guarantee crimes occurring in these locations will be prevented. These represent separate stages of predictive policing that have distinct metrics for assessing performance. Few studies have evaluated predictive policing in its entirety, from prediction to implementation of tactical responses to prevent future crimes; collectively, however, the studies that have been undertaken have demonstrated mixed findings and weak evidence of effectiveness (Weisburd & Majmundar 2018). Previous predictive policing studies had a tendency to avoid longer-term sustained crime prevention approaches in favour of deploying police cars on a per shift basis, predominantly using either covert or overt policing. These types of time-sensitive tactics are difficult to implement as planned, however, due to a range of issues—for example, available resources, equipment failures and emergency response (Famega, Hinkle & Weisburd 2017; Ratcliffe et al. 2020). For instance, the final report for the Philadelphia predictive policing experiment noted that marked patrol cars allocated to spend time in a predicted hot spot spent only 64 percent of the time in the assigned areas and unmarked patrol cars were present only 68 percent of the time (Ratcliffe et al. 2018).

Despite the mixed evidence of effectiveness, predictive policing tactical responses generally overlap with those used in hot spots policing (eg problem-oriented policing, community policing), which does have a strong evidence base for crime reduction effects (Weisburd & Majmundar 2018; Weisburd et al. 2019). This mixed evidence could be a function of the shift-based implementation often seen, other implementation issues, or simply the lower number of rigorous evaluations of predictive policing response tactics.

Predictive policing as an approach has the potential to be effective, as long as there is long-term commitment to applying the techniques in a thoughtful way. For predictive policing to work—that is, to reduce crime—not only do the forecasts need to be valid, but appropriate prevention tactics need to be identified and effectively implemented. This means that once areas are identified as a high risk for future crime, further analysis should be conducted to inform the response tactics that are appropriate. For example, areas that have long-standing crime and disorder issues are most likely to benefit from a problem-solving approach that analyses and seeks to understand the underlying causes of the crime problem in order to design a solution to address them. Areas with heightened but transient risk are likely to benefit from 'traditional' hot spots policing tactics, such as directed patrols. Much like our recommendation that algorithm parameters be tailored to local areas, response tactics should be tailored to the individual areas predicted as having a higher risk of crime for predictive policing to be an effective crime reduction approach.

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

Bailey T & Gatrell A 1995. *Interactive spatial data analysis*. Essex, England: Longman Scientific and Technical

Ball GH & Hall DJ 1970. A clustering technique for summarizing multivariate data. *Behavioral Science* 12: 153–155

Bannister J, Bates E & Kearns A 2018. Local variance in the crime drop: A longitudinal study of neighbourhoods in Greater Glasgow, Scotland. *British Journal of Criminology* 58(1): 177–199. https://doi.org/10.1093/bjc/azx022

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

Brantingham PJ 2018. The logic of data bias and its impact on place-based predictive policing. *Ohio State Journal of Criminal Law* 15: 473–486

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

Brantingham PJ, Valasik M & Mohler GO 2018. Does predictive policing lead to biased arrests? Results from a randomized controlled trial. *Statistics and Public Policy 5*(1): 1–6. https://doi.org/10.1080/2330443X.2018.1438940

Brannon MM 2017. Datafied and divided: Techno-dimensions of inequality in American cities. *City & Community* 16(1): 20–24. https://doi.org/10.1111/cico.12220

Brayne S 2017. Big data surveillance: The case of policing. *American Sociological Review* 82(5): 977–1008. https://doi.org/10.1177/0003122417725865

Browning M & Arrigo B 2021. Stop and risk: Policing, data, and the digital age of discrimination. *American Journal of Criminal Justice* 46: 298–316

Buccine-Schraeder H & Kennedy LW 2021. Risk terrain modeling (RTM). In JC Barnes and David R Forde (eds), *The encyclopedia of research methods in criminology and criminal justice*. Wiley: New York, NY

Caplan JM & Kennedy LW 2010. *Risk terrain modeling manual: Theoretical framework and technical steps of spatial risk assessment*. Newark, NJ: Rutgers Center on Public Security

Caplan JM, Kennedy LW & Miller J 2011. Risk terrain modeling: Brokering criminological theory and GIS methods for crime forecasting. *Justice Quarterly* 28(2): 360–381. https://doi.org/10.1080/07418825.2010.486037

Caplan JM, Kennedy LW, Piza EL & Barnum JD 2020. Using vulnerability and exposure to improve robbery prediction and target selection. *Applied Spatial Analysis and Policy* 13: 113–138

Chainey S 2021. A comparison of burglary near repeat victimization between rural and urban areas using a target-based assessment of criminal opportunity. *International Criminal Justice Review* 31(4): 405–419. https://doi.org/10.1177/10575677211041920

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 2008a. The utility of hotspot mapping for predicting patterns of crime. *Security Journal* 21: 4–28

Chainey S, Tompson L & Uhlig S 2008b. Response to Pezzuchi. *Security Journal* 21: 293–294

Chainey S, Tompson L & Uhlig S 2008c. Response to Levine. *Security Journal* 21: 303–306

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

Famega C, Hinkle JC & Weisburd D 2017. Why getting inside the 'black box' is important: Examining treatment implementation and outputs in policing experiments. *Police Quarterly* 20(1): 106–132. https://doi.org/10.1177/1098611116664336

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

Flaxman S, Chirico M, Pereira P & Loeffler C 2019. Scalable high-resolution forecasting of sparse spatiotemporal events with kernel methods: A winning solution to the NIJ 'real-time crime forecasting challenge'. *The Annals of Applied Statistics* 13(4): 2564–2585

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

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

Grossman L et al. 2011. The 50 best inventions: The year's most inspired ideas, innovations and revolutions, from the microscopic to the stratospheric. *Time*: New York, NY

Hardyn W & Rummens A 2018. Predictive policing as a new tool for law enforcement? Recent developments and challenges. *European Journal on Criminal Policy and Research* 24(3): 201–218. https://doi.org/10.1007/s10610-017-9361-2

Hart TC & Zandbergen PA 2014. Kernel density estimation and hotspot mapping: Examining the influence of interpolation method, grid cell size, and bandwidth on crime forecasting. *Policing: An International Journal of Police Strategies & Management* 37(2): 305–323. https://doi.org/10.1108/PIJPSM-04-2013-0039

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

Hawkes AG 1971. Spectra of some self-exciting and mutually Exciting point processes. *Biometrika* 58(1): 83–90.

Hu Y, Wang F, Guin C & Zhu H 2018. A spatio-temporal Kernel density estimation framework for predictive crime hotspot mapping and evaluation. *Applied Geography* 99: 89–97

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

Jefferson BJ 2018. Predictable policing: Predictive crime mapping and geographies of policing and race. *Annals of the American Association of Geographers* 108(1): 1–16

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

Johnson SD, Bowers K, Birks D & Pease K 2009. Predictive mapping of crime by Promap: Accuracy, units of analysis, and the environmental backcloth. In DL Weisburd, W Bernasco & GJN Bruinsma (eds), *Putting crime in its place: Units of analysis in geographic* criminology. New York: Springer: 171–198

Kounadi O, Ristea A, Araujo Jr, A & Leitner M 2020. A systematic review on spatial crime forecasting. *Crime Science* 9: 1–22

Kulldorff M, Heffernan R, Hartman J, Assuncao R & Mostashari F 2005. A space–time permutation scan statistic for disease outbreak detection. *PLoS Medicine* 2(3):e59. https://doi.org/10.1371/journal.pmed.0020059

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

Lee YJ, O SH & Eck JE 2017. *A theory-driven algorithm for real-time crime hot spot forecasting*. United States: National Criminal Justice Reference Service, Department of Justice

Levine E, Tisch J, Tasso A & Joy M 2017. The New York City Police Department's Domain Awareness System. *Interfaces* 47(1): 70–84. https://doi.org/10.1287/inte.2016.0860

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

Liesenfeld R, Richard JF & Vogler J 2017. Likelihood-based inference and prediction in spatio-temporal panel count models for urban crimes. *Journal of Applied Econometrics 32*(3): 600–620

Marjanovic O, Cecez-Kecmanovic D & Vidgen R 2022. Theorising algorithmic justice. *European Journal of Information Systems* 31(3): 269–287. https://doi.org/10.1080/0960085X.2021.1934130

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

Meijer A & Wessels M 2019. Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration* 42(12): 1031–1039. https://doi.org/10.1080/01900692.2019.1575664

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 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

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 G, Raje R, Carter J, Valasik M & Brantingham J 2018. A penalized likelihood method for balancing accuracy and fairness in predictive policing. *2018 IEEE International Conference on Systems, Man, and Cybernetics*, Miyazaki, Japan, pp 2454–2459. https://doi.org/10.1109/SMC.2018.00421

Mohler G, Short M, Brantingham P, Schoenberg F & Tita G 2011. Self-exciting point process modeling of crime. *Journal of the American Statistical Association* 106(493): 100–108. https://doi.org/10.1198/jasa.2011.ap09546

National Academies of Sciences, Engineering, and Medicine 2018. *Proactive policing: Effects on crime and communities*. The National Academies Press

National Institute of Justice 2016. *Real-time crime forecasting challenge*. https://nij.ojp.gov/funding/real-time-crime-forecasting-challenge

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

Pezzuchi G 2008. A brief commentary on 'The utility of hotspot mapping for predicting spatial patterns of crime'. *Security Journal* 21: 291–292

PredPol 2020. *Predictive Policing: Guidance on where and when to patrol*. https://www.predpol.com/how-predictive-policing-works/

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 2021. Queensland local government areas (LGA), 2021. https://www.qgso.qld.gov.au/visualisations-geographies/geographical-boundaries/queensland-local-government-areas-lga-2021

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 2016–17*. Brisbane: Queensland Police Service. https://www.police.qld.gov.au/maps-and-statistics/annual-statistics

Ratcliffe J 2014. What is the future...of predictive policing? *Translational Criminology* 6: 4–5

Ratcliffe JH 2004. The hotspot matrix: A framework for the spatio-temporal targeting of crime reduction. *Police Practice and Research* 5(1): 5–23. https://doi.org/10.1080/1561426042000191305

Ratcliffe JH et al. 2020. The Philadelphia predictive policing experiment. *Journal of Experimental Criminology* 17(1): 15–41

Ratcliffe JH et al. 2018. *The Philadelphia predictive policing experiment: Final report*. Washington, DC: National Institute of Justice

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

Rosser G, Davies T, Bowers KJ, Johnson SD & Cheng T 2017. Predictive crime mapping: Arbitrary grids or street networks? *Journal of Quantitative Criminology* 33: 569–594

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

Sturgill K 2020. Santa Cruz become the first U.S. city to ban predictive policing. *Los Angeles Times*, 26 June. https://www.latimes.com/california/story/2020-06-26/santa-cruz-becomes-first-u-s-city-to-ban-predictive-policing

Taylor RB, Ratcliffe JH & Perenzin A 2015. Can we predict long-term community crime problems? The estimation of ecological continuity to model risk heterogeneity. *Journal of Research in Crime and Delinquency* 52(5): 635–657. https://doi.org/10.1177/0022427815586062

Thompson HR 1956. Distribution of distance to n^{th} neighbour in a population of randomly distributed individuals. *Ecology* 37: 391–394

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

Uchida CD 2009. *A national discussion on predictive policing: Defining our terms and mapping successful implementation strategies*. USGPO: Washington, DC

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

Ward JH 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58: 236–244

Weisburd D et al. 2019. Proactive policing: A summary of the report of the National Academies of Sciences, Engineering, and Medicine. *Asian Journal of Criminology* 14(2): 145–177

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 2019. Algorithmic justice: Algorithms and big data in criminal justice settings. *European Journal of Criminology* 18(5): 623–642. https://doi.org/10.1177/1477370819876762

Appendix

Figure A1: Brisbane hot spots map displaying the distribution of each crime type annually across the study period

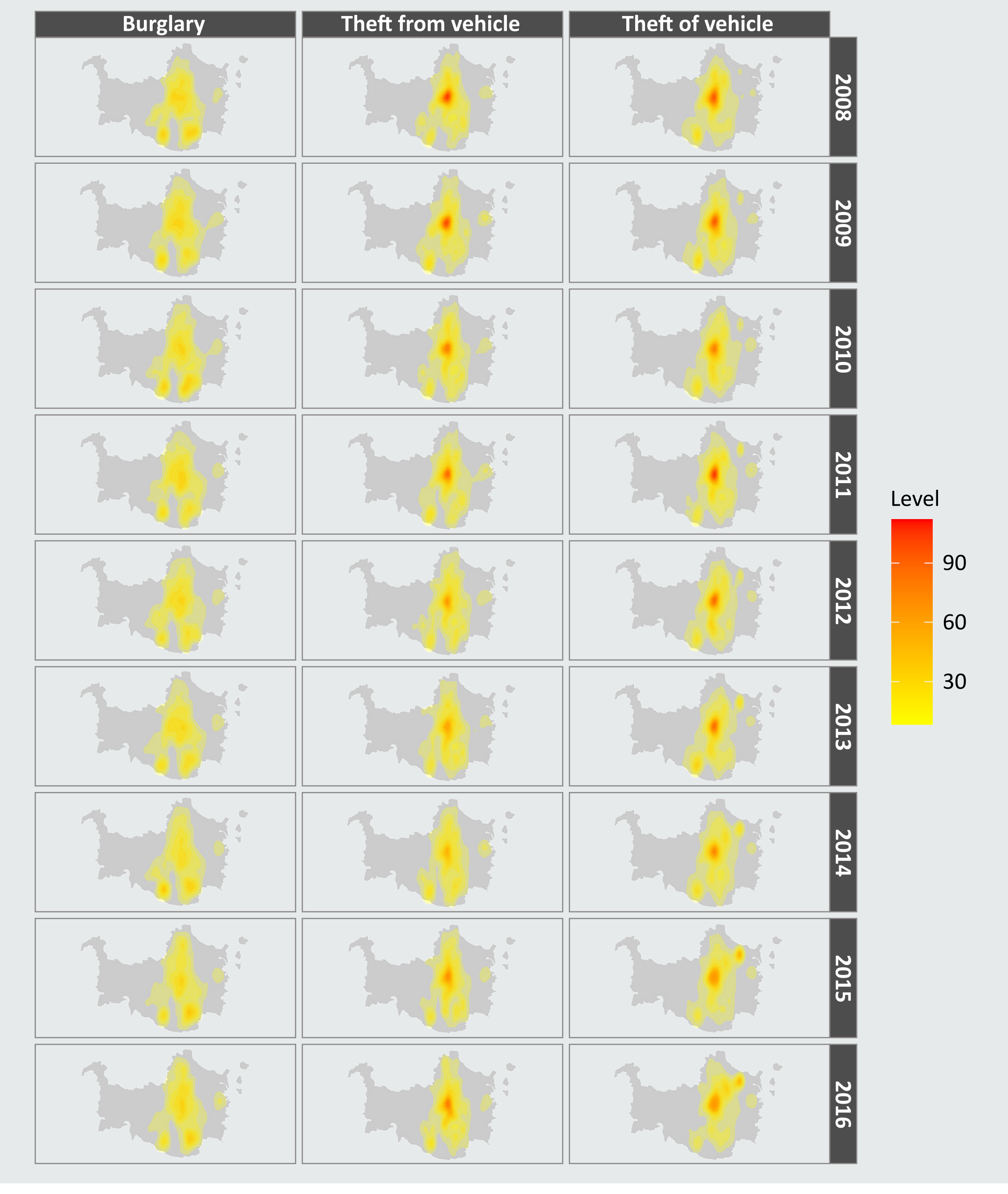


Figure A2: Brisbane hot spots map displaying the distribution of each crime type by day of the week across the study period

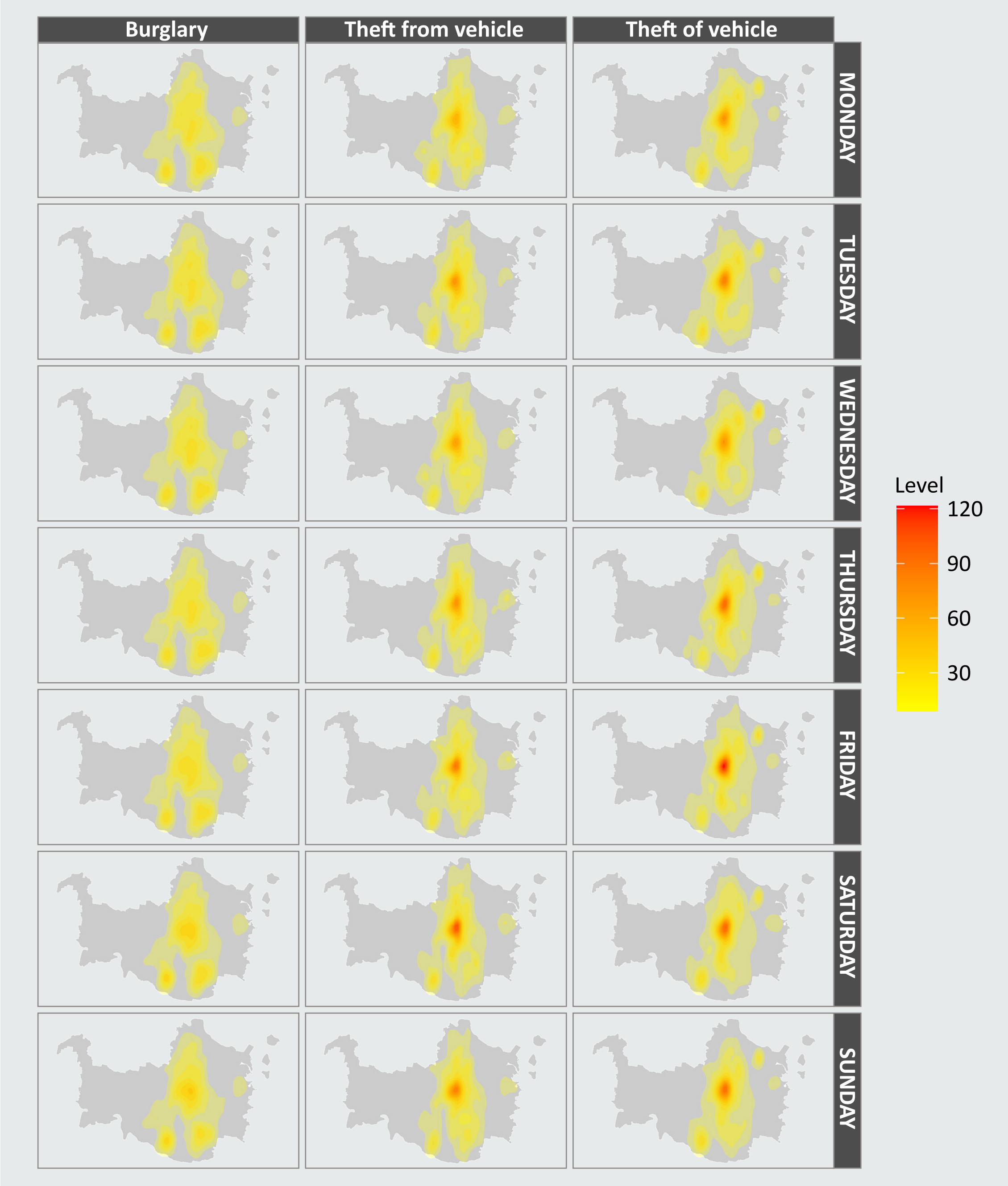


Figure A3: Logan hot spots map displaying the distribution of each crime type annually across the study period

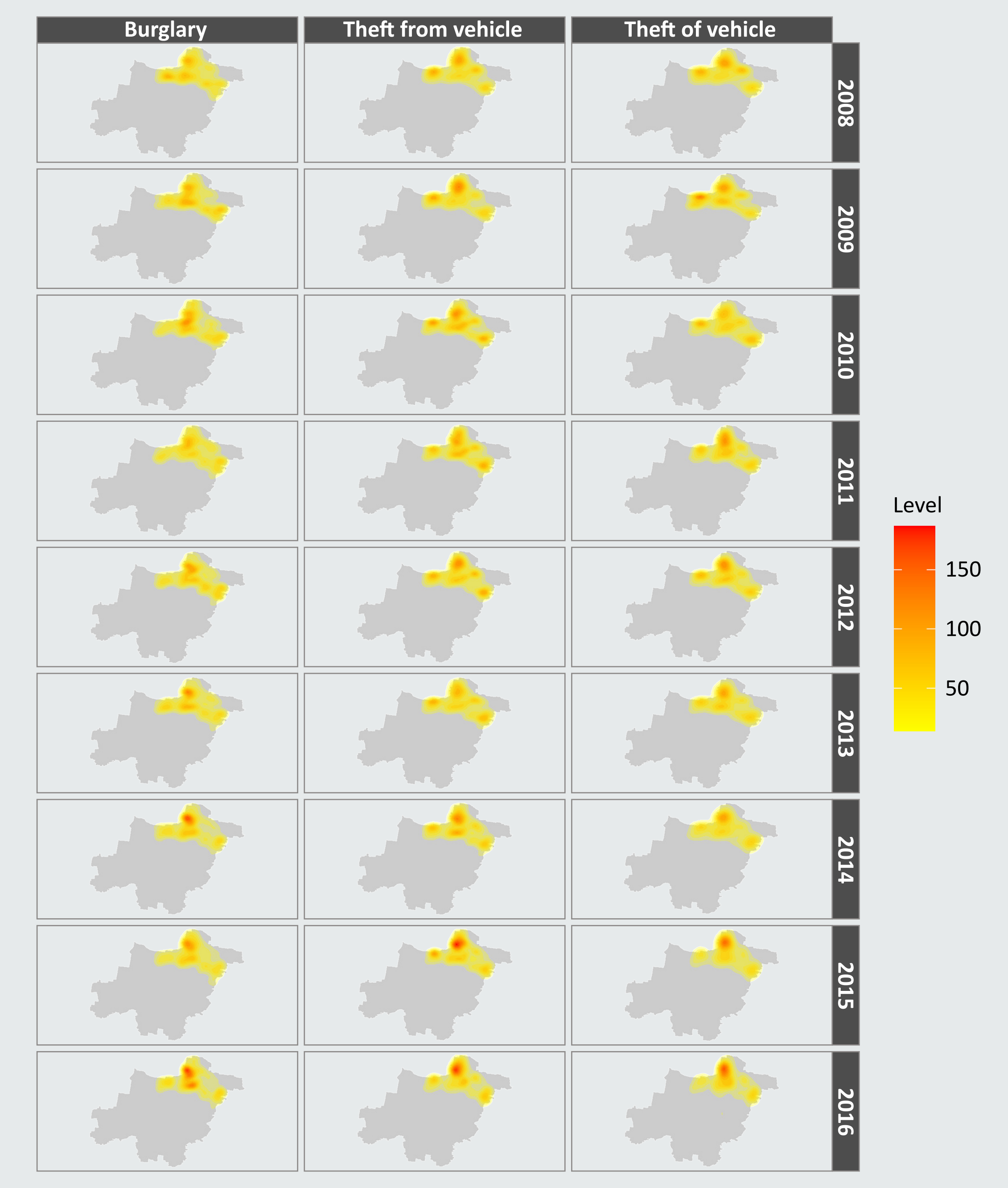


Figure A4: Logan hot spots map displaying the distribution of each crime type by day of the week across the study period

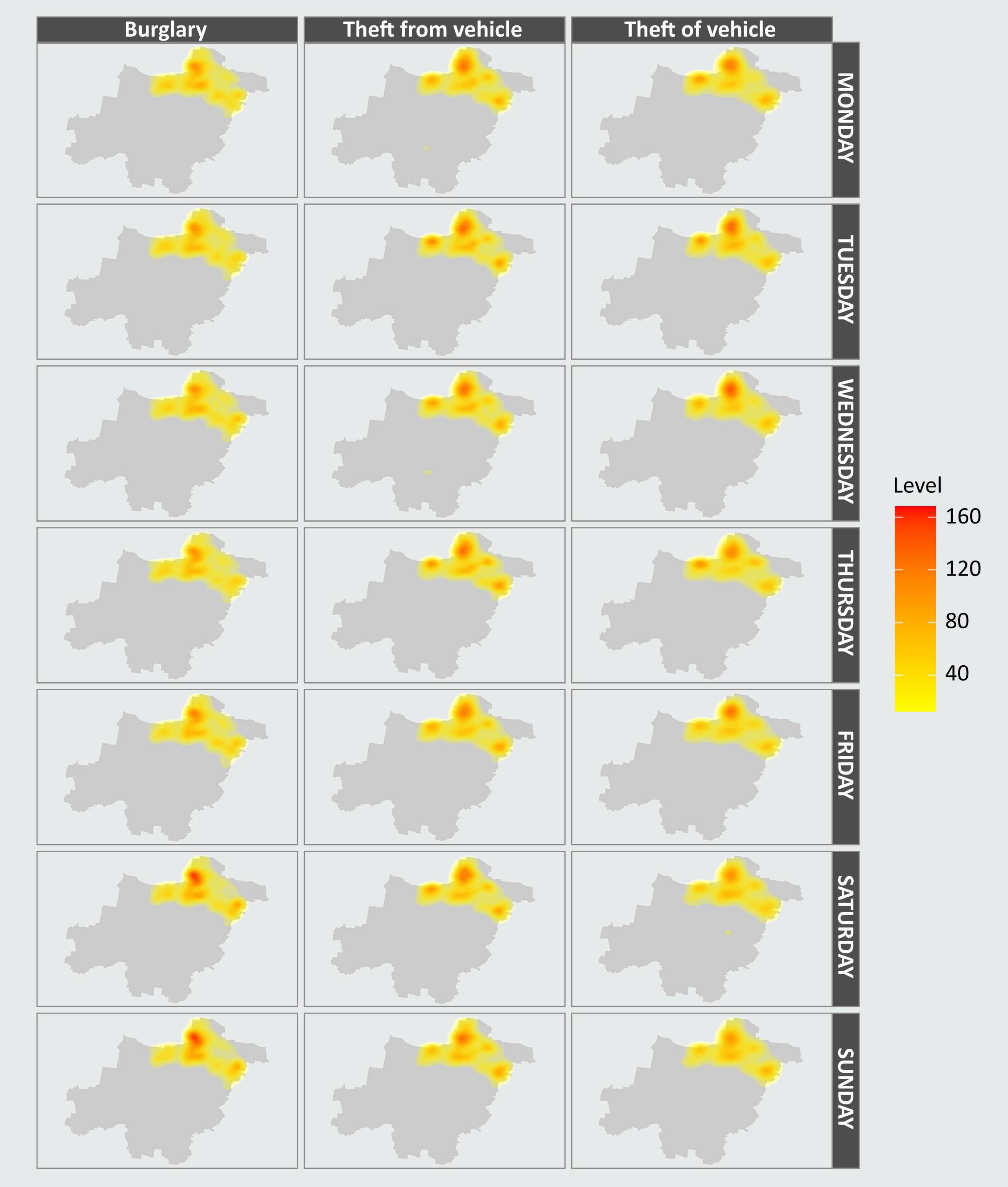


Figure A5: Townsville hot spots map displaying the distribution of each crime type annually across the study period

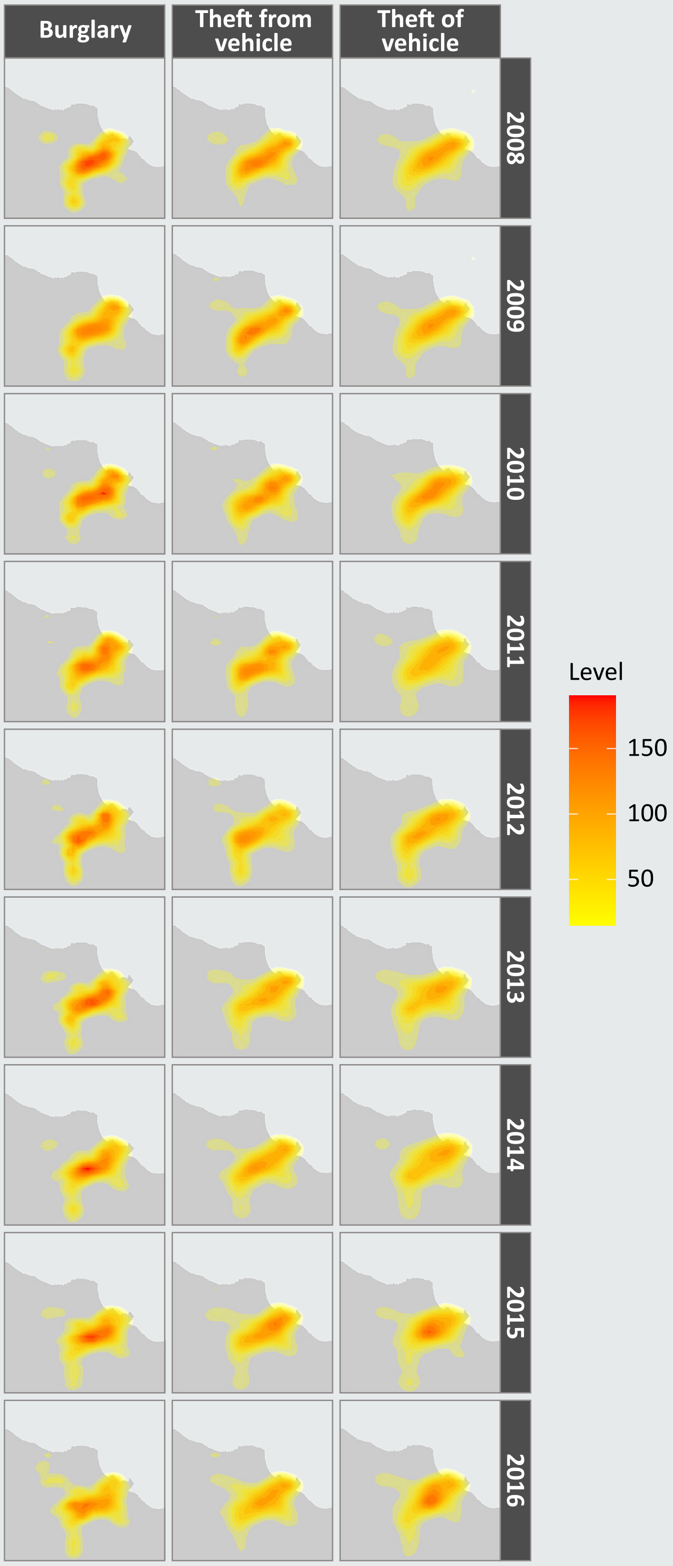
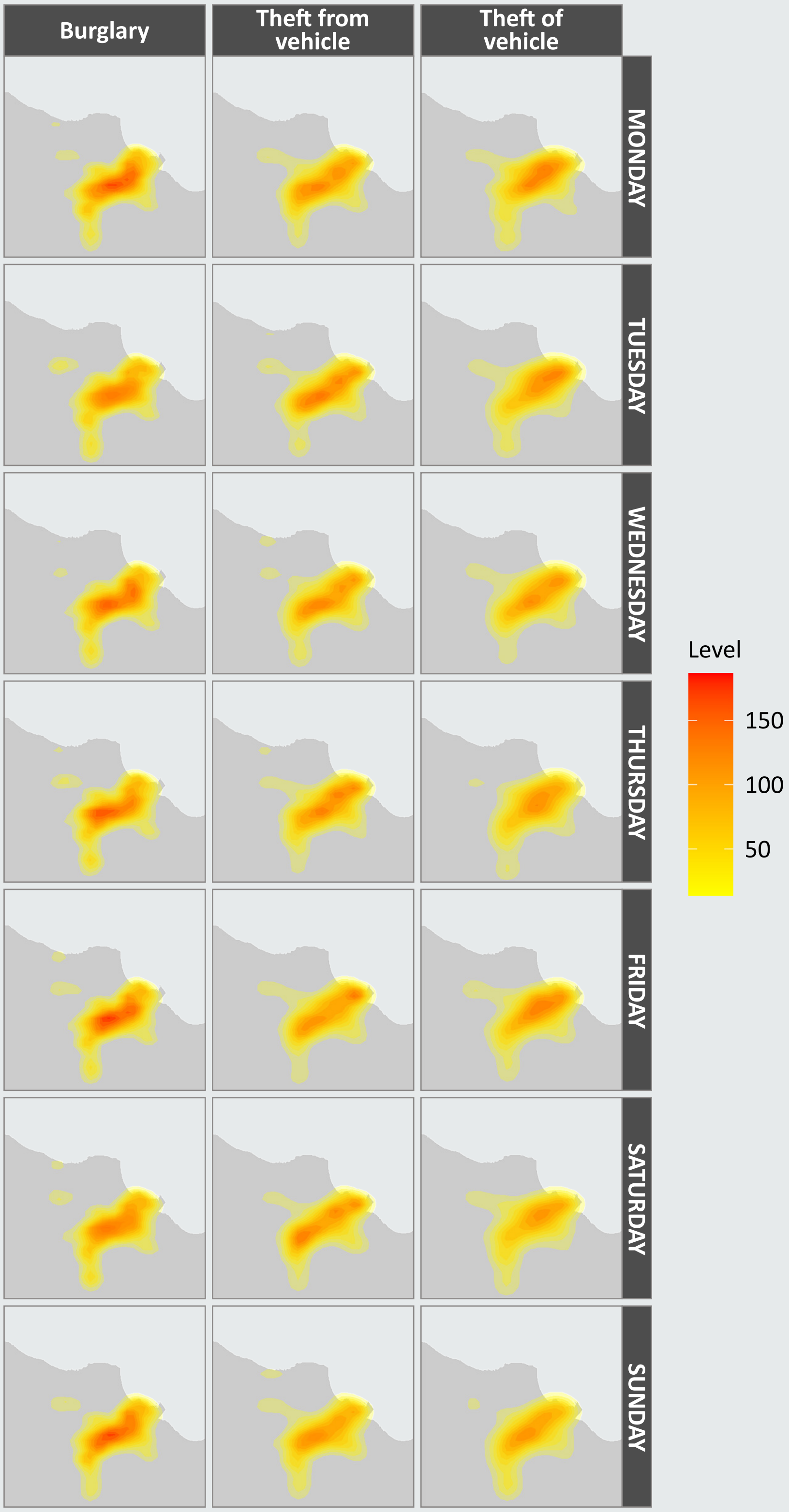


Figure A6: Townsville hot spots map displaying the distribution of each crime type by day of the week across the study period



CRG reports

CRG 32/16–17

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

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

Timothy C Hart, PhD, is an Associate Professor in the Department of Criminology and Criminal Justice at the University of Tampa.

www.aic.gov.au/crg