Nordic Journal of Studies in Policing

UNIVERSITETSFORLAGET

RESEARCH PUBLICATION

Volume 2, No. 1-2021, p. 1–21
ISSN online: 2703-7045
DOI: https://doi.org/10.18261/issn.2703-7045-2021-01-02

Exploring Violent and Property Crime Geographically

A Comparison of the Accuracy and Precision of Kernel Density Estimation and Simple Count

Maria Camacho Doyle
PhD student, School of Law, Psychology and Social Work, Örebro University, Sweden
https://orcid.org/0000-0002-1576-5079
maria.doyle@oru.se

Manne Gerell
Associate Professor, Department of Criminology, Malmö University, Sweden
https://orcid.org/0000-0002-2145-113X

Henrik Andershed
Professor, School of Law, Psychology and Social Work, Örebro University, Sweden
https://orcid.org/0000-0002-8163-6558

Abstract
There are multiple geographical crime prediction techniques to use and comparing different prediction techniques therefore becomes important. In the current study we compared the accuracy (Predictive Accuracy Index) and precision (Recapture Rate Index) of simply counting crimes: Simple Count with Kernel Density Estimation in the prediction of where people are reported to commit violent crimes (assault and robbery) and property crimes (residential burglary, property damage, theft, vehicle theft and arson), geographically. These predictions were done using a different number of years into the future and based on a different number of years combined to do the crime prediction, in a large Swedish municipality. The Simple Count technique performed quite well in comparison to simple Kernel Density Estimation no matter what crime was being predicted, making us conclude that it may not be necessary to use the more complex method of Kernel Density Estimation to predict where people are reported to commit crime geographically.

Keywords
Hotspot Mapping, Predictive Accuracy Index, Recapture Rate Index, Simple Count, Kernel Density Estimation

Introduction
Is it science fiction or reality to predict where people will commit crime with such an accuracy that we can prevent further crime from happening? Perfecting its reality is something the police and researchers have been working on for the past decades, with continuously evolving techniques (Drawve, 2016; Eck et al., 2005; Levine, 2008). There are several different techniques today aimed at predicting where people will commit crime geographically. Some techniques are simple, like putting pins on a map to portray where people have committed crimes and based on this history try to predict where people will commit crime

in the future (Eck et al., 2005). Other techniques are more advanced, using mathematical algorithms to calculate a prediction of future crimes (Mohler et al., 2015). This study aims to compare the accuracy and precision of two different geographical crime prediction techniques, more specifically in terms of where people will commit violent and property crime in a large Swedish municipality.

It is important to know with some accuracy where the next crime incident might happen because it can provide practitioners a possibility of preventing the potential incident altogether. Thus, subsequently to be able to prevent crime in a cost-effective way by concentrating preventive efforts on high-risk places, we first need to predict where crimes are likely to take place (Ratcliffe & McCullagh, 1999, 2001; Sherman, 1995). In doing this we need to know what prediction technique is most useful/accurate. While important strides have been taken in the research on prediction of crime, a lot remains to be learnt, in terms of both prediction but also concerning how to use such predictions to actually prevent crime.

The current paper aims to add to the literature by testing two different prediction techniques, Kernel Density Estimation (KDE) and simply counting crimes: Simple Count (SC), for several different crime types to assess how good they are at predicting people's criminal behavior geographically. An advantage of SC is its simplicity. It is somewhat surprising that so few papers have tested its viability as a crime prediction technique (e.g., Groff & La Vigne 2002), given that more complicated techniques may raise the bar for police departments and other practitioners to implement techniques of crime prediction in their work. The current study will also be one of the first to test the viability of crime prediction techniques in a Nordic setting, and as such adds to the literature by adding a new geographic context.

Geography of crime and hotspots

Mapping and trying to understand and predict where people will commit crime geographically is by no means new; it dates back to trying to understand crime geographically in France (Guerry, 1833), in Belgium (Quetelet, 1842), and in Chicago (Shaw & McKay, 1942), to mention a few. The early studies on where crime occurs were however mostly focused on large areas, comparing regions, cities or neighborhoods. In the 1980s and 90s, large strides in understanding the geography of crime were taken as smaller geographical locations came into the spotlight (Weisburd, Bruinsma & Bernasco, 2009). Multiple studies show that some locations tend to persistently have more crime (Braga et al., 2014; Eck et al., 2005; Weisburd et al.2004; Weisburd, Morris & Groff 2009). These locations are geographically small (Caplan et al., 2011; Eck et al., 2005; Kennedy, Caplan and Piza, 2011; Sherman, Gartin and Buerger, 1989; Weisburd et al. 2004; Weisburd et al. 2009) and are often referred to as crime hotspots (Caplan et al., 2013; Drawve 2016). A crime hotspot can be defined as a small geographical location with a concentration of crime incidents over time (Sherman & Weisburd 1995). The concept of hotspots has become very influential in both research and practice since its inception, not least in relation to hotspot policing. Hotspot policing is a crime prevention strategy that builds on identifying hotspots of crime and directing police resources to such locations to prevent crime (Braga & Weisburd 2010). It has consistently been shown to result in crime reduction (Braga et al., 2014), and has been described as one of the policing strategies with the strongest evidence base as of today (Abt & Winship, 2016). Hotspot policing is not without problems though and we refer the reader to Rosenbaum (2006). For example, identifying hotspots is not the same as understanding them; a more holistic perspective of these crime hotspots would be preferred. Nevertheless, hotspot policing is dependent on hotspot analysis, the process of identifying locations with a high concentration of crime appropriate for preventive interventions. In the current paper the focus is on the aforementioned hotspot analysis, and how such analysis can be done.

In hotspot analysis the aim is to identify as well as predict hotspots (Drawve, 2016). In traditional hotspot analysis, the analyst uses the particular location's crime history to predict future crime. Like all analysis depending on prior incidents, this is a retrospective technique, where historical events are used to predict the future (Chainey et al., 2008; Drawve, 2016; Groff & La Vigne, 2002). The criminal history of a location is a well-known risk factor for future crime (see e.g., Braga et al., 2014; Chainey et al., 2008). Crime history can therefore plausibly aid in identifying future hotspots of crime (Kennedy et al., 2016). The criminal history of a location can be analyzed in many different ways. The most straightforward way of doing it is to simply count the number of crimes at the location within a defined time frame, something that we in this paper refer to as Simple Count. Other techniques to a varying degree involve more complicated methods, such as calculating the density of crime, or weighting crimes depending on recency (see e.g., Bowers et al., 2004; Chainey et al., 2008; Hu et al., 2018).

Theoretical perspectives on the geography of crime

Retrospective crime mapping is somewhat atheoretical (Groff & La Vigne 2002). There are however different theoretical explanations that can be used to describe and understand why crimes cluster in one place. Mutual in the theories is the focus on "the underlying dynamics, situations, and attributes of the place" (Braga et al., 2014, p. 635), including the routine activity theory (Cohen & Felson, 1979), crime pattern theory (Brantingham & Brantingham, 1995), and flag and boost (Pease, 1998), to mention a few. There are for sure geographical crime hotspots. That is, crime can sometimes concentrate in and be repeated at small geographical areas (Weisburd, 2015). According to the flag hypotheses, certain weaknesses of the place itself "flag" the place (hotspot) as an open game for crime. The weaknesses could for example be low surveillance, easy access or high value targets. Potential offenders commit crime by taking advantage of these existing "flagged" weaknesses (e.g., see Bowers & Johnson, 2005). According to the boost hypothesis, previous victimization increases the risk of future victimization. An initial crime "boosts" the risk of another crime in the near vicinity (Farrell & Pease, 1993; Short et al., 2009). A burglary can increase the risk of a future burglary in the area within a certain time-period. The offender can make a rational choice and choose to come back to the area to commit another crime because he/she knows the opportunities or weaknesses of the place (Bowers & Johnson, 2004). Another example are shootings that increase, "boost", the risk of future shootings. The risk of future shootings based on the first shooting can possible be due to retaliation and/or escalation of the crime (Ratcliffe & Rengert, 2008). In short, crime begets crime. It might be that some crimes are more dependent on place while some are less so.

Crime prediction techniques

Kernel Density Estimation (KDE) is a retrospective technique that can be used to predict where people will commit crime geographically. In KDE a grid is put over the study area. In every grid cell, with a pre-specified grid cell size, every crime incident is calculated. The closer an incident is to the center of the grid cell; the higher density value is given to the incident. The grid cell is then given a density estimate. This will result in a map with a variation of crime density and aid in identifying crime hotspots in a larger area, kind of like a wheatear map (Eck et al., 2005; Levine 2013). KDE has been studied and compared in earlier research of crime hotspots whilst predicting for example burglary, thefts from vehicles and thefts of vehicles (Chainey et al., 2008; Levine, 2008), assault (Levine, 2008), robberies (Chainey et al., 2008; Drawve, 2016; Dugato, 2013; Levine, 2008; Van Patten et al., 2009), to mention a few. There are also more elaborated KDE methods where both place and time of the crime are considered simultaneously such as "prospective hot-spotting" (see Bowers et al., 2004).

Sophisticated crime prediction techniques with mathematical algorithms and special crime prediction computer programs are used in practice (see e.g., Kennedy et al., 2016; Mohler et al., 2015). These crime prediction techniques are however not always used or not always easily attainable in practice (see e.g., Caplan et al., 2011; Groff & La Vigne, 2002; Spelman, 1995). The alternative is that the crime analysts simply count the amounts of crime that have occurred in an area, and the area with the most crime incidents will be considered the hotspot (Groff & La Vigne, 2002; Johnson et al., 2007; Spelman, 1995). A Simple Count (SC) of crime incidents in places reflects the notion that "hotspots of today are hotspots of tomorrow"(Groff & La Vigne, 2002 p. 34).

Because multiple techniques exist, comparing different prediction techniques in different locations becomes important. Previous research that compares different hotspot techniques comes to inconsistent conclusions concerning what prediction technique might be preferable (Chainey et al., 2008; Dugato, 2013; Hart & Zandbergen, 2012; Levine, 2008; Van Patten et al., 2009). To date, there is no standard technique to identify specific locations with high crime that are promoted over others (Drawve, 2016). It has been suggested that using different prediction techniques together might be preferable, no matter what crime is being predicted (Caplan et al., 2013; Drawve, 2016; Kennedy et al., 2011; Van Patten et al., 2009). Many of the previously mentioned studies have analyzed techniques using data from the US (Caplan et al., 2015; Caplan et al., 2011; Drawve, 2016; Hart & Zandbergen, 2012; Kennedy et al., 2011; Kennedy et al., 2016) and larger European cities (Chainey et al., 2008; Dugato, 2013).

Recent studies recommend using a shared reference value when comparing different hotspot techniques (see recommendation from e.g., Chainey et al., 2008; Drawve, 2016). One reference value recommended is the predictive accuracy index (PAI) value for accuracy jointly with the recapture rate index (RRI) value for precision. With the PAI value you examine how accurate the technique is in finding the hotspot, by comparing the hit rate (how much crime you are able to predict in the hotspot compared to the total amount of crime in the study area) to the area of the hotspot and of the whole study area. The PAI formula was originally proposed by Chainey et al., (2008) and extended by Van Patten et al. (2009). With the RRI value you examine how precise the technique is in finding the hotspot over time (Levine, 2008). For a more in-depth description of what PAI and RRI are and how they are calculated, see Chainey et al. (2008), Drawve (2016), Levine (2008) and Van Patten et al. (2009).

When looking at crime history, using one year of crime history or even longer time periods might be good as some public places seem to be chronic hotspots. Using shorter timespans such as month to month only, might be misleading as there can be some fluctuations in the hotspot status in such short timespans. Even though an area might be a chronic hotspot, looking at shorter timespans might make it look as if it is not, due to the fluctuating data (see e.g., Adams-Fuller, 2001; Spelman, 1995). For some crimes though, such as burglary, using one month of data might render more accurate predictions, rather than one year would, due to time-sensitive repeat victimization (see e.g., Anderson et al., 1995; Farrell & Pease, 1993; Johnson et al., 2007; Polvi et al., 1991).

The use of KDE to predict crime is fairly well established (Chainey et al., 2008; Chainey, 2013; Drawve, 2016; Hu et al., 2018). It is however less common to find studies that simply count the number of crimes in an area and use that to predict future crime (see e.g. Groff & La Vigne, 2002). However, SC may be a technique more easily used by practitioners, as it is easy to conduct and interpret. As surprisingly little research appears to have been done on the accuracy of SC to predict future crimes, we aim to explore this issue and evaluate the difference between KDE and SC in crime prediction. This will shed some light on the merit of the law of parsimony that suggests simple explanations, and of the statement that “the more complicated methods are not always better predictors” (Groff & La Vigne 2002 p. 50). In addition, to find out what the “information and statistical analysis threshold” for accurate and precise predictions is, might be beneficial for the practice of crime prediction and in the end crime prevention. Furthermore, using KDE to compare with SC is a starting point and can provide a baseline for which more sophisticated methods (such as the KDE’s by Bowers et al., 2004 and Gorr and Lee’s 2015, 2017 or PredPol by Mohler et al., 2015) can be compared to in future research. Therefore, comparing a Simple Count with a retrospective hotspot technique such as KDE, using a predictive accuracy index (PAI) value for accuracy and a recapture rate index (RRI) value for precision, when predicting different crime types in a large municipality in northern Europe might therefore add to the research field.

The aim of the current study is to compare the accuracy and precision of Simple Count (SC) and Kernel Density Estimation (KDE) when it comes to predicting where people will commit violent crimes (assault and robbery) and property crimes (residential burglary, property damage, theft, vehicle theft and arson) geographically, in a large Swedish municipality. The current study adds to the literature by looking at whether Simple Count (SC) or Kernel Density Estimation (KDE) will render more accurate and precise predictions of violent and property crime, when compared to each other. In the analysis, we will therefore study differences between KDE and SC, in relation to property crimes and violence as well as for specific crime types. This will be done using a different number of years into the future and based on the different number of years combined to do the crime prediction. Overall then we will explore multiple aspects of crime prediction using KDE and SC to study the value of each technique under different circumstances.

In this study, we have chosen to define hotspots in a relative manner to create comparability across years and crime types. The share of locations in the municipality that together contain 30 percent of the crime in the municipality are considered hotspots. We then test how well we can predict what locations fall into such a definition of hotspot based on KDE and SC, using multiple different specifications for the calculations. This will be discussed further below under headline Simple Count.

Method

Study area

In the current study we compared the predictability accuracy and precision of SC and KDE using reported crime from Malmö, Sweden. Malmö is a municipality with an official population of 331 201 in June 2017 (SCB1) and is approximately 157 km^2 in size (SCB2). It has more foreign-born residents, more unemployment, a younger population, and more crime than Sweden in general does (Malmö stad, 2014; Statistics Sweden, 2013; Ekström et al., 2012). For the SC and KDE analysis, Malmö was divided into 100 meter by 100 meter grid cells, a total of 16737 grid cells.

Data

The study was approved by an ethics board in 2017 (Dnr 2017/479). Geocoded crime point data was obtained from the Malmö police department from January 2012 through December 2017. The crime points were police reported crimes (reported offences).

Table 1. Number of police reported crime incidents, by type of crime, for each calendar year used for analyses.

Year	Robbery	Assault	Threat to public servant	Sexual exhibition	Residential burglary	Arson	Property damage	Theft	Vehicle Theft	Crime involving public danger	Total
2012	134	872	177	49	1158	118	4962	3576	5417	112	16581
2013	167	864	155	43	1001	86	4216	3433	5217	79	15262
2014	123	982	135	37	942	78	4066	3084	4573	84	14104
2015	174	983	150	44	909	80	4414	3252	4319	84	14409
2016	212	1053	171	42	859	90	3733	3443	4654	115	14374
2017	196	882	157	35	736	109	3623	3537	4458	82	13938
Total	1006	5636	945	250	5605	561	25014	20325	28638	556	88601

The crime data either came with geocodes to the specific address of the crime incident or was subsequently geocoded for spatial analysis. Geocoding and an interactive geocoding correction procedure was performed using ArcGis online. All of the geocoded data were geocoded above the recommended 85 percent level recommended by Ratcliffe (2004). See appendix for a more detailed account of the geocoding process.

After analysis with CrimeStat IV all KDE-layers were clipped in ArcGIS 10.3 to the Malmö municipality area, to minimize the number of grid cells. A total of 16737 grid cells 100 meters by 100 meters were used for analysis. Some crime points ended up outside the Malmö municipality area. If the crime points were 50 meters or less from the Malmö municipality border, they were included in the analysis by a 50-meter extension of the Malmö border at that particular point, to account for potential mistakes in the geocoding process.

To compare KDE and SC the predictive accuracy index (PAI) value for accuracy and the recapture rate index (RRI) value for precision were calculated following the primary analysis based on recommendation (see e.g., Chainey et al., 2008; Drawve, 2016; Drawve et al., 2016; Dugato, 2013; Hart & Zandbergen, 2014; Van Patten et al., 2009).

Simple Count

Simple Count is counting the crime incidents in a certain area in a defined time period, often one year, and letting that information predict the subsequent year's hotspots. Simply counting crimes in an area is the most basic form of analysis that can be done for a geographical area, and as such, holds an advantage over more complicated techniques. To analyze this, a grid net with grid cell sizes of 100 meters by 100 meters was laid over the Malmö study area. This grid cell size was based on a recommendation for KDE cell sizes (Chainey 2013). See appendix for a more detailed account of the grid cell size selection process. The crime data were then spatially joined in ArcGIS 10.3 to these grid cells, so that a number of crime incidents could be attributed to all the grid cells. Many grid cells were empty; hence no crime incidents were counted there. Fewer grid cells had more crime incidents in them. The more crime incidents the hotter the hotspot.

The grid cells containing 30 percent of all the crime incidents were considered top hotspots. Hence, the grid cells that captured 30 percent of all crime in the years (for example the year 2016 or years 2014–2015) were used to predict crime in the year 2017. The cut-offs are arbitrary. To be able to capture 30 percent of all crime a SC method of finding top hotspots was used for violent crime, property crime, assault, property damage, theft and vehicle theft. For robbery, residential burglary and arson a KDE standard deviation method was used. The differences produced by these two methods of calculating a top hotspot was that the KDE standard deviation method generally produced lower PAI values for comparison, due to the greater area rate used in this method. In other words, more 100- by 100-meter grid cells were generally needed to reach 30 percent of all crimes. The hotspot patterns were however similar for both the SC and the KDE standard deviation cut-off methods no matter the crime type. Hence, it is unlikely that the method of locating the cut-offs substantially alter the main findings. See appendix for a more detailed account of the top hotspot selection process and the choices made for the different crime types. In all hotspot cut-offs we allowed for a 10 percent fluctuation, meaning the top hotspots used for analysis could include 25 percent to 35 percent of all crimes.

Kernel Density Estimation

KDE was calculated using the CrimeStat IV single-kernel density interpolation technique (Levine 2013). A direct (Euclidean) type of distance measurement was used. For some analyses the indirect (Manhattan) type of distance measurement was also used and compared. One type of measurement (direct or indirect) did not produce consistently better results, which is why the direct type of distance measurement was chosen and used. Within KDE certain parameters need to be set: method of interpolation, grid cell size and bandwidth. Because the interpolation method used and bandwidth chosen can affect the outcome result (Hart & Zandbergen 2014), a search for the "best KDE model" was completed. See appendix for a more detailed account of the search for the "best KDE model"

Once 18 analyses were run, (three interpolation methods times' six bandwidths) the best model was found by looking at the PAI and RRI averages, see Table 2.

Table 2. Average predictive accuracy index (PAI) and recapture rate index (RRI) by kernel density interpolation method.

Interpolation Method	Measured PAI	Predictive PAI	RRI
Normal	51.29	46.28	0.92
Quartic	67.07	55.07	0.83
Triangular	64.31	53.26	0.84

Note: The figures represent the average score of violent crime hotspots predicted in 2017 using 2016 data. The averages contain all the different bandwidths tested.

The interpolation method quartic rendered the most accurate result and the interpolation method normal the most precise over time result. Quartic was however chosen due to its higher PAI value. In comparison to SC, KDE still had better RRI values using the quartic interpolation. Furthermore, the results showed that as bandwidth decreased the PAI increased and the RRI decreased. Due to us trying to find the best model possible for KDE to compare against SC, quartic interpolation with a 100-meter cell size and a 100-meter bandwidth with a predictive PAI of 62.83 was chosen. While having a bandwidth that is as large as the cell size is somewhat counterintuitive when using KDE, as it means that the score will approximate a Simple Count in ranking grid cells, we nevertheless chose to use it due to its high PAI value.

Next, top hotspot grid cells were calculated within ArcGIS 10.3. As in SC the grid cells with about 30 percent of all the crime incidents were considered hotspots. In SC we identified the about 30 percent cut-off by selecting the grid cells with 30 percent of all crimes. We then looked at the same number of top grid cells for the KDE for comparison. Except for the crimes of robbery, residential burglary and arson, the KDE standard deviation was used, as mentioned before. Hence, for both SC and KDE, fixed grid cells (100 by 100 meters) were used for comparison, making the shape and size of the top hotspots identical for both techniques. SC and KDE were compared by examining the number of predicted crimes in 2017 in these 100- by 100-meter grid cells. The choice of fixed grid cells for the top hotspots (rather than the KDE buffers) was made for easy comparison of the two methods and for ease of practical implementation (see also Lee & Gorr, 2015), even though working with buffers is feasible in practice. Whether fixed grid cells or dynamic boundaries provide additional prediction benefits is an empirical question not answered in the current paper but is left for future research.

We viewed differences of five percent between the PAI and RRI values as substantial differences. We used a percentage of difference because PAI values vary when the area percentage changes; this might affect different crime types differently. The five percent is calculated by multiplying the PAI value by 0.05. For example, PAI 40.52 x 0.05 = 2. PAI 40.52 – 2 = 38.52. If the second PAI value is lower than 38.52 then there is more than five percent difference between the two PAI values, which is a substantial difference.

In sum, in the comparison of one SC technique and one KDE technique (the best one) was used. Both SC and KDE had fixed grid cells (100 by 100 meters) making the shape and size of the top hotspots identical in both methods. The geographic location of the fixed cells however could differ and were therefore compared. For violent crime, property crime, assault, property damage, theft and vehicle theft the SC method of finding hotspots was used. For robbery, residential burglary and arson the KDE standard deviation method was used.

Results

Only results with the highest prediction accuracy and best precision over time will be presented in the text. The year to be predicted in every analysis is 2017. As can be seen in Table 3, when viewing differences of five percent between the PAI and RRI values as substantial differences, there is generally no difference between SC and KDE in accuracy predicting the total amount of crime. In general, as seen in Table 3, PAI values are slightly higher for SC, though it varies a bit, and as mentioned above, the differences are not substantial. KDE generally renders more precise predictions over time (KDE RRI = 0.65 and SC RRI =0.61). Generally, the results show (see Table 3) that neither technique is more suitable for one crime type than the other. The prediction accuracy (PAI values) is quite similar across crime types for both prediction techniques.

Using two to five years of reported violent and property crime rather than one year alters the accuracy and precision of the predictions made (see Table 4). For KDE the combined two years of reported crime (2015–2016) prior to the year of prediction render the highest accuracy rate (PAI 47.91). For SC the year prior (2016) to the year of prediction renders the highest accuracy rate (PAI 50.60). Predicting the total crime rate using reported crime in Malmö, the year prior to prediction or the combined two years before prediction render the highest prediction accuracy. Using three to five years of reported crime renders the best precision over time (KDE RRI 0. 77 and SC RRI 0.74) predicting the total crime rate.

Predicting violent crime (KDE PAI 64.04, SC PAI 66.94) renders more accurate predictions than predicting property crime (KDE PAI 20.16, SC PAI 21.35). However, predicting property crimes renders more precise predictions over time (KDE RRI 0.85, SC RRI 0.84) compared to violent crime (KDE RRI 0.73, SC RRI 0.69).

As can be seen in Table 3, accuracy is increased when predicting assault (KDE PAI 80.94) rather than violent crime (KDE PAI 64.04), while precision over time does not change. Conversely, predicting robbery renders less accuracy (KDE PAI 37.09) and precision over time

(KDE RRI 0.35) compared to the violent crime umbrella. Accuracy is also increased when predicting different property crime types (for example theft KDE PAI 41.25) rather than the total property crime umbrella (PAI 20.16). One exception is residential burglary (KDE PAI 11.13). Precision over time is however less when property crime (KDE RRI 0.85) is broken down to actual crime types. For more specific examples see Table 3.

With both techniques and the year of reported crime prior to prediction, about 24 percent of property crime in 1 percent of the municipality was accurately predicted (KDE PAI = 23.80, SC PAI = 24.72). About 23 percent of vehicle thefts in 0.7 percent of the municipality were accurately predicted (KDE PAI= 34.51, SC PAI = 34.34). Adding years of reported crime to the analysis generally reduces the accuracy. Using another year of reported crime than 2016 renders less accuracy. Using two to five years of reported crime renders the best precision over time with KDE (Property crime RRI=0.94, Vehicle theft RRI=1). Using three to five years of reported crime renders the best precision over time with SC (Property crime RRI=0.94, Vehicle theft RRI=0.95). Using only one year of reported crime, precision over time generally worsens the further you get from the year of prediction and generally renders less precision over time compared to the precision with several years of reported crime combined.

With both techniques using reported violent crime from 2015–2016, about 23 percent of violent crime in 0.3 percent of the municipality was accurately predicted (KDE PAI = 69.49, SC PAI = 70.69). With both techniques and reported assault from 2015 of about 19 percent of assaults in 0.2 percent of the municipality were accurately predicted (KDE PAI = 97.25, SC PAI = 93.10). Adding years of reported crime to the analysis generally reduces the prediction accuracy for both violent crime and assault. For violent crime, using any year of reported crime renders similar results. For assault the prediction accuracy fluctuates when using different years of reported assault separately (see Tables 5–6 in the appendix). The precision over time is generally increased by adding years of reported crime to the analysis for both violent crime and assault. For example, four years of reported violent crime (KDE RRI=0.85, SC RRI=0.78) and five years of reported assault (KDE RRI= .86). Using separate years of reported crime, the year prior renders the best precision over time, then precision over time generally worsens the further you get from the year 2017.

With SC using the year of reported crime prior to prediction, about 13 percent of robberies in 0.2 percent of the municipality were accurately predicted (PAI = 60.01). About 24 percent of theft in 0.3 percent of the municipality was accurately predicted (PAI = 76.99). About 24 percent of property damage in 0.6 percent of the municipality was accurately predicted (PAI = 41.82). For robberies and theft, adding years of reported crime to the analysis reduces the prediction accuracy. For property damage adding two or four years of reported crime renders similar prediction accuracy. Using another single year of reported crime than 2016 renders less accuracy for robberies, theft and property damage alike. For robbery, using three years of reported robbery renders the best precision over time (KDE and SC RRI= 0.61). For theft, using three to five years of reported theft renders the best precision over time (SC RRI=0.92). For property damage, using three to four years of reported property damage renders the best precision over time (SC 2013-2016 RRI=0.95). Using only one year of reported crime renders less precision over time compared to the precision with several years of reported crime combined for robbery, theft and property damage alike.

With KDE and the year 2013 about eight percent of residential burglaries in 0.6 percent of the municipality were accurately predicted (PAI = 12.61). Adding more years of reported residential burglary to the analysis generally reduces the accuracy. Using only one year of reported residential burglary, the accuracy fluctuates depending on year used (see Table 5 in the appendix). Using five years of reported residential burglary render the best precision over time (KDE RRI=0.67). The fewer years of reported crime used for prediction, as well as using only one year of reported crime, renders less precision over time (see Table 7 in the appendix).

With KDE and the years 2015–2016 about 12 percent of arson in 0.1 percent of the municipality was accurately predicted (PAI = 83.17). Adding years of reported arson to the analysis generally reduces the accuracy. Using one year of reported arson in the analysis, the accuracy worsens the further you get from the year of prediction. Using one or two years of reported arson renders the best precision over time (KDE 2016 and 2015–2016 RRI=0.42). Using only one year of reported arson renders less precision over time.

KDE is sensitive to the parameters selected. Using a quartic interpolation rather than normal interpolation increased the predictive accuracy by 8.79 PAI, but reduced RRI by 0.09. As bandwidth decreased, PAI increased and RRI decreased. Comparing a quartic interpolation with 500-meter bandwidth to a 100-meter bandwidth, the difference in PAI is 12.47 and in RRI 0.19. SC should be less sensitive to parameters selected but using Malmö crime data the PAI value is affected by the cell size used. A cell size of 50 meters renders a 237 PAI value when predicting violent crime in 2017 using crime data from 2016. This compares with a PAI of 66.84 using a 100-meter cell size.

Table 3. Average PAI and RRI values for different crimes predicting where people will commit crime geographically in 2017

Crime type	KDE PAI	SC PAI	KDE RRI	SC RRI
Violent Crime	64.04	66.94	0.73	0.69
• Assault	80.94	81.86	0.77*	0.70*
• Robbery‡	37.09	35.93	0.35*	0.32*
Property Crime	20.16*	21.35*	0.85	0.84
• Residential burglary‡	11.13*	9.79*	0.41*	0.34*
• Theft	41.25*	43.64*	0.80	0.78
• Vehicle theft	29.69	29.58	0.87*	0.82*
• Property damage	34.39*	38.78*	0.78	0.79
• Arson‡	37.18	36.78	0.25*	0.19*
Total	39.54	40.52	0.65*	0.61*

Note: ‡ denotes the use of standard deviation in KDE rather than count in SC; * denotes a >5 percent difference between the PAI- and the RRI values of KDE and SC.

Table 4. Average PAI and RRI values for using different years predicting where people will commit crime geographically in 2017

Year	KDE PAI	SC PAI	KDE RRI	SC RRI
2012	33.21^	33.09^	0.47*^	0.44*^
2013	35.10^	35.79^	0.53*^	0.49*^
2014	35.97^	36.71^	0.58^	0.56^
2015	38.47*^	42.89*^	0.57^	0.55^
2016	45.37*^	50.60*	0.63*^	0.59*^
2012–2016	39.08^	39.62^	0.77	0.74
2013–2016	39.42^	40.19^	0.76	0.73
2014–2016	41.35^	41.64^	0.77	0.73
2015–2016	47.91*	44.13*^	0.71*^	0.65*^
Total	39.54	40.52	0.65*	0.61*

Note: * denotes a >5 percent difference between the PAI- and the RRI values of KDE and SC. ^ denotes a >5 percent difference between the PAI- and the RRI values of the different years within each technique. Base years are KDE PAI: 2015–2016, SC PAI: 2016, KDE and SC RRI: 2012–2016.

Specific places in Malmö had more than a few missing geocodes for the reported crime, hence robustness testing was performed. If the places had more than 10 points of missing reported crime per year, then polygons were drawn in ArcGIS 10.3 and the amount of missing crime points were simulated by creating random points to represent the missing points (434 points in total). The information (such as date, time and crime type) from the missing crime points was added to the randomly created points. Only property crime was affected by the missing points, with bike theft (part of vehicle theft) being the main crime affected. The PAI values did not increase much by creating random points for places with more than 10 crimes missing per year, which is why the decision was made not to create and include points for places with more than five crime points missing per year. Another argument for not creating polygons and adding random points to account for missing points is that it will most likely not be done in practice.

Discussion

The purpose of the current study was to compare KDE and SC when it came to predicting where people will commit crimes of violence (assault and robbery) and property crimes (residential burglary, property damage, theft, vehicle theft and arson) geographically, in a large Swedish municipality. The results indicated that KDE and SC were quite similar in their accuracy in predicting where people will commit different kinds of crime. The key take-away message and conclusion from this study however is that across all model specifications, the most basic technique possible, counting the number of crimes in the prior year within an area and using that Simple Count to predict crime, performed well. Using the year 2016, 22 percent of violent crimes in 2017 were accurately predicted. As a comparison, by randomly selecting locations 0.3 percent of violent crimes in 2017 were accurately predicted. Or when randomly selecting locations, from locations with at least one past violent crime, about eight percent of violent crimes in 2017 were accurately predicted. This means that, based on the current study, practitioners who want to predict where people will commit a lot of crime in the coming year, can use information about where people committed many crimes in the prior year.

There generally did not appear to be any added benefit of using more than one year, nor of using simple KDE. While the finding that more years added to the analysis and more complicated techniques generally do not give better predictions, may appear to be an underwhelming finding from an academic viewpoint, we believe it is of fair importance from a more practical viewpoint. A technique that is easy to use, understand, and apply, will inherently have better chances of implementation in practice than a more complicated technique, all else being equal. This means that hotspot analysis and hotspot policing is within easy reach for any authority that can map the locations of crime within their area.

Another important finding is that the possibility of accurately predicting where people will commit crime in a small number of locations differs substantially across crime types. Some crime types were easier to predict on a yearly basis. For public environment assault that registered the highest prediction rate, it meant that just by looking at the prior year of assaults, we identified 32 locations comprising 0.2 percent of the municipality, where 19 percent of the assaults took place. While 19 percent of assaults may sound like a low number, this is 156 crimes that take place in a limited number of locations where the police can have a reasonably high presence to prevent crimes, in particular if the analysis is complemented by the days and times when assaults are the most common to direct prevention efforts. For residential burglary, the analysis was less useful, as the analysis fluctuated and identified eight percent of residential burglaries in 0.6 percent of the municipality. To work on preventing burglary it may be necessary to strengthen the analysis by considering temporary hotspots, and prospective hotspotting (Johnson & Bowers 2004). The risk for burglary is significantly higher in the weeks, or months, following a burglary incident in the area; this is well established internationally (Bernasco, 2008; Johnson, 2013; Short et al., 2009; Townsley et al., 2003), as well as for the municipality of Malmö directly (Hoppe and Gerell 2019). Efforts might benefit from being more focused in time. It has recently been shown that efforts at preventing burglary based on such analysis can indeed have an impact on crime (Johnson et al., 2017; Stokes & Clare 2019).

Theft is another crime that fared well by simply counting crimes in the prior year as a basis for the prediction for the following year. Predicting 24 percent, or 846 thefts, in 52 locations (an area as small as 0.3 percent of the municipality), could mean that a focused effort of any practitioner with a crime preventive agenda in these few areas potentially could prevent more than one fifth of all thefts. Similarly to one other study, theft from vehicles has been shown to be easier to predict than both residential burglaries and vehicle theft (Chainey et al., 2008). In the current study, vehicle theft was most likely harder to predict than theft because it was more spread out geographically, vehicle theft covered 0.7 percent of the municipality's surface instead of 0.3 percent. Of interest though is that no matter what crime was predicted, with these data, the top hotspots to work with are on less than one percent of the municipality's surface. For arson, the geographical area where crimes occurred was stably quite small (about 0.1 percent of the municipality). However, using two or more years to inform the prediction with KDE only predicted about 12 percent of all accounts of arson. In common for arson, robbery and residential burglary was that it was hard to reach precision over time predicting these crimes. Maybe using shorter timeframes for prediction than a year, or letting environmental and contextual factors influence the predictions alongside crime history would increase the prediction accuracy and precision for these particular crimes. Finally, predicting 24 percent of all property damage on 0.6 percent of the municipality's surface could be useful information even for other stakeholders than the police. The municipality or property owners could use this information not only to potentially decrease future crime incidents but also to decrease future fear of crime, as property damage and fear of crime has shown to be related (see e.g. Doran and Burgess 2012; Farrall, Jackson and Gray 2009).

Using crime data from Malmö a hotspot map with geospatial data continue to offer value for one year up to five years depending on crime type (see Tables 5 through 8 in the appendix for specifics), however often with a diminishing value over time. Assault for example showed a relatively steady pattern over time and hence a hotspot map from five years ago still offers some value. It is better with a more recent hotspot map, but the old one still offers some information. Property damage also held a reasonably stable pattern, with a slight diminishing value over time. Hotspot maps of arson, robbery and theft on the other hand need to be more recent, as older hotspot maps of these crimes rapidly lose their value over the years. Maps of vehicle theft also lose value over the years, however not as rapidly as the other mentioned crimes. As for residential burglary, as previously mentioned, maybe one year of crime data is not suitable to predict residential burglaries as both accuracy and precision over time is quite low as well as fluctuating.

The current study has some weaknesses, as the analyses are atemporal. Adding a temporal aspect could have provided additional information and maybe helped in the prediction of for example residential burglary. Another interesting analysis that could have been beneficial is a weighted KDE (as recommended by Bowers et al. 2004; Chainey et al. 2008; Hu et al. 2018) where more years are added to the analysis but the year 2016 is given a higher weight. This might have made KDE a better prediction technique of choice compared to SC. Furthermore, maybe the nearest neighbor hierarchical prediction technique (Nnh) with convex hulls would have fared better in comparison with SC than KDE, as earlier research shows that Nnh is better than KDE at short term predictions, using one year to predict the following year (Van Patten et al., 2009).

One aspect not considered by SC is that every crime has a possible spillover effect, KDE hotspots reflect this near-repeat victimization pattern while SC hotspots do not (see for example Bernasco et al., 2015; Hoppe & Gerell, 2018 on near-repeat victimization for burglary). Crime displacement and diffusion effects due to targeted crime-reduction interventions in certain hotspots (see e.g., Braga et al., 2019) as well as weather (Cohn, 1990) are potential confounders that could affect the prediction of different crime types differently. These confounders could obscure the results for any prediction attempt. We have not measured these confounders in the current study. However, the same crime-data was used to compare KDE and SC, so possible changes in crime patterns due to interventions should not affect the comparison of the two techniques. Furthermore, it is important to emphasize that since no significance test was used; there is a chance of predicting false hotspots, for both KDE and SC.

It is possible that the results would differ slightly with a different top hotspot cut-off method than the ones used. However, as described in the appendix, since the hotspot patterns were similar for the 30, 50 and 70–80 percent cut-offs and the SC and KDE standard deviation method of locating top hotspot cut-offs, the cut-off methods are unlikely to substantially alter the main findings. Furthermore, the 10 percent allowance for the hotspot to fluctuate (top hotspots containing 25–35 percent of all crimes instead of a strict 30 percent) could potentially have affected the results. Using PAI as the measurement for comparison, even though recommended, could be critiqued. Differences in PAI values between techniques can be great, though it only concerns a difference of two or three crimes being predicted. On the other hand, the PAI value differences between techniques can be small but concern a greater difference in the amount of crime being predicted. The PAI value varies depending on a changing area percentage. This is why we chose to view a five percent difference between the PAI values as a substantial difference, rather than just looking at the PAI values in themselves. Another way of improving the PAI-measure could be to use a PAI curve instead of a PAI value (see e.g., Hu et al., 2018). Lastly, crime statistics are among other things influenced by public willingness to report the crime and can vary depending on crime type. This could influence the predictions made and comparisons between crime types.

The results of the current study may give some support to theories that emphasize that crime in one place bring about more crime. For example, the flag hypothesis (see e.g., Bowers & Johnson, 2005), could be used to explain the more stable predictions of property damage and assault over time. These areas could have certain characteristics conducing to more crimes. For burglary, which had a fluctuating accuracy, and for arson, robbery and theft, that were not as stable over time, maybe the boost hypothesis could be more explanatory. In the boost hypothesis (see e.g. Bowers & Johnson, 2005), the risk of crime over time is not constant, but increases due to previous (and more recent) crime. This needs further investigation though, perhaps by using smaller time-windows for crime predictions, or by adding characteristics of the place to the analysis. In the current study, only crime history was of interest and used and the shortest time period was one year.

In future research, it would be interesting to see a comparison of other prediction techniques such as the Nnh, or a weighted KDE with SC. Maybe there are other better-suited techniques for the current crime data. Furthermore, an added aspect of shorter prediction windows such as months, weeks or even days, while still keeping the history in mind, might provide additional prediction accuracy and precision and would make the more advanced techniques fare better in comparison with a Simple Count. Additionally, analyzing the adding of other information like environmental and contextual risk factors to the prediction techniques might further benefit the predictions and increase the number of crimes being predicted. It would also be interesting to see if these crime predictions such as, for example arson, were dependent on some environmental variable or some other contextual variable. Adding environmental and contextual factors could potentially bring about a more holistic approach to the hotspot study as described in Rosenbaum (2006). Many interesting analyses remain to be done.

Policy Implications

The results of the current study give some direction on how to prevent future crime and potentially ease the workload of practitioners preventing crime, through the means of hotspot mapping. First, any practitioner with a crime preventive agenda could use past crime in an area as a predictor of future crime in that particular area and guide their preventive efforts accordingly. The geographical stability of certain crimes, such as public environment assault and property damage could be used as a guide to curtail future crimes. Secondly, based on the current study, practitioners who want to predict crime in the coming year, could count crimes committed in the prior year. Results will however most likely be improved by adding information about the days and times when these crimes are most common. If crime data from the prior year is not available, it is possible to use earlier years to make predictions and still get good results predicting public environment assault, property damage and, more carefully, with vehicle theft. Not for theft, however. For residential burglary, robbery and arson the mapping should be more focused in time (not based on yearly data) as hotspots seem temporary. Third, based on the results of the current study it is possible to count the crime events in any given area of the city and let that guide the following year's preventive efforts. All you need is the ability to map the locations of crime within an area.

For practitioners wanting to use SC to predict future crime the procedure is quite simple. First, decide on a size of the location, for example 100 by 100 meters. The top hotspots to work with are small no matter what crime is being predicted. Then count how many crimes have been committed in each location in the past year. The locations with the most crime incidents are where attention should be focused the following year.

References

Adams-Fuller, T. (2001). *Historical Homicide Hot Spots: The Case of Three Cities* (Doctoral Dissertation), Howard University, Washington, DC.

Anderson, D., Chenery, S., & Pease, K. (1995). *Biting Back: Tackling Repeat Burglary and Car Crime.* Home Office Police Research Group, London.

Abt, T., & Winship, C. (2016). What Works in Reducing Community Violence: A Meta-Review and Field Study for the Northern Triangle, (Washington, DC: USAID) https://www. usaid. gov/sites/ default/files. USAID-2016-What-Works-in-Reducing-Community-Violence-Final-Report.pdf

Barnum, J. D., Caplan, J. M., Kennedy, L. W., & Piza, E. L. (2017). The crime kaleidoscope: A cross-jurisdictional analysis of place features and crime in three urban environments. *Applied Geography, 79,* 203–211. https://doi.org/10.1016/j.apgeog.2016.12.011

Bernasco, W. (2008). Them again? Same-offender involvement in repeat and near repeat burglaries, *European Journal of Criminology, 5*(4), 411–431. https://doi.org/10.1177/1477370808095124

Bernasco, W., Johnson, S. D., & Ruiter, S. (2015). Learning where to offend: Effects of past on future burglary locations. *Applied Geography, 60*, 120–129. https://doi.org/10.1016/j.apgeog.2015.03.014

Bowers, K.J. & Johnson, S.D. (2005). Domestic burglary repeats and space-time clusters: The dimensions of risk. *European Journal of Criminology, 2*(1), 67–92. https://doi.org/10.1177/1477370805048631

Bowers, K. J., Johnson, S. D., & Pease, K. (2004). Prospective hot-spotting: The future of crime mapping? *British Journal of Criminology, 44*(5), 641–658. https://doi.org/10.1093/bjc/azh036

Bowers, K. J. & Johnson, S. D. (2004). 'Who commits near repeats? A test of the boost explanation.' *Western Criminology Review 5*(3): 12–24.

Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: an update of an ongoing systematic review and meta-analysis. *Journal of Experimental Criminology, 15*(3), 289–311. https://doi.org/10.1007/s11292-019-09372-3

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

Braga, A. A., & Weisburd, D. L. (2010). *Policing Problem Places.* Oxford, England: Oxford University Press.

Brantingham, P., & Brantingham, P. (1995). Criminality of place. *European Journal on Criminal Policy and Research, 3*(3), 5–26. https://doi.org/10.1007/BF02242925

Brantingham, P. L., & Brantingham, P. J. (1993). Nodes, paths and edges: Considerations on the complexity of crime and the physical environment. *Journal of Environmental Psychology, 13*(1), 3–28. https://doi.org/10.1016/S0272-4944(05)80212-9

Caplan, J. M., Kennedy, L. W., Barnum, J. D., & Piza, E. L. (2015). Risk terrain modeling for spatial risk assessment. *Cityscape, 17*(1), 7. https://doi.org/10.1525/9780520958807

Caplan, J. M., Kennedy, L. W., & 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, J. M., Kennedy, L. W., & Piza, E. L. (2013). Joint utility of event-dependent and environmental crime analysis techniques for violent crime forecasting. *Crime and Delinquency,* 59(2), 243–270. https://doi.org/10.1177/0011128712461901

Chainey, S., Tompson, L., & Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal, 21*(1–2), 4–28. https://doi.org/10.1057/palgrave.sj.8350066

Chainey, S. P. (2013). Examining the influence of cell size and bandwidth size on kernel density estimation crime hotspot maps for predicting spatial patterns of crime. *Bulletin of the Geographical Society of Liege,* 60, 7–19. URL: https://discovery.ucl.ac.uk/id/eprint/1388639

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review, 44*(4), 588–608. https://doi.org/10.2307/2094589

Cohn, E. G. (1990). Weather and crime. *British Journal of Criminology, 30*(1), 51–64. https://doi.org/10.1093/oxfordjournals.bjc.a047980

Doran, B. J., & Burgess, M. B. (2012). *Putting Fear of Crime on the Map, Investigating Perceptions of Crime Using Geographic Information Systems.* London: Springer.

Drawve, G. (2016). A metric comparison of predictive hot spot techniques and RTM. *Justice Quarterly, 33*(3), 369–397. https://doi.org/10.1080/07418825.2014.904393

Drawve, G., Moak, S., C., & Berthelot, E., R (2016). Predictability of gun crimes: A comparison of hot spot and risk terrain modelling techniques. *Policing and Society, 26*(3), 312–331. https://doi.org/10.1080/10439463.2014.942851

Dugato, M. (2013). Assessing the validity of risk terrain modeling in a European city: preventing robberies in Milan. *Crime Mapping, 5*(1), 63–89.

Eck, J. E., Chainey, S., Cameron, J. G., Leitner, M., & Wilson, R. E. (2005). *Mapping Crime: Understanding Hot Spots.* Washington, DC: National Institute of Justice.

Ekström, E., Eriksson, A., Korsell, L., & Vesterhav, D. (2012). *Brottslighet och Trygghet i Malmö, Stockholm och Göteborg: En kartläggning*, Stockholm: Brottsförebyggande rådet.

Farrall, S. D., Jackson, J., & Gray, E. (2009). *Social Order and the Fear of Crime in Contemporary Times.* Oxford: Oxford University Press.

Farrell, G., & Pease, K. (1993). *Once Bitten, Twice Bitten: Repeat Victimisation and its Implications for Crime Preventio*n. Retrieved from https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/2149/3/CoverforPDF.pdf

Gerell, M. (2018). Quantifying the geographical (un)reliability of police data. *Nordisk Politiforskning, 4*(2). https://doi.org/10.18261/issn.1894-8693-2018-02-05

Gorr, W. L., & Lee, Y. (2015). Early warning system for temporary crime hot spots. *Journal of Quantitative Criminology, 31*(1), 25–47. https://doi.org/10.1007/s10940-014-9223-8

Groff, E. R., & La Vigne, N. G. (2002). Forecasting the future of predictive crime mapping. *Crime Prevention Studies*, *13*, 29–58.

Guerry, A.-M. (1833). Essai Sur La Statistique Morale De La France [Essay on the Moral Statistics of France]. Paris: Crochard.

Hart, T., & Zandbergen, P. (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 and Management, 37*(2), 305–323. https://doi.org/10.1108/pijpsm-04-2013-0039

Hart, T. C., & Zandbergen, P. A. (2012). *Effects of Data Quality on Predictive Hotspot Mapping.* Washington, DC: National Institute of Justice.

Hoppe, L., & Gerell, M. (2019). Near-repeat burglary patterns in Malmö: Stability and change over time. *European Journal of Criminology*, 16(1). https://doi.org/10.1177/1477370817751382

Hu, Y., Wang, F., Guin, C., and Zhu, H. J. A. G. (2018). A spatio-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation. *Applied Geography,* 99, 89–97. https://doi.org/10.1016/j.apgeog.2018.08.001

Johnson, D. (2013). The space/time behavior of dwelling burglars: Finding near repeat patterns in serial offender data. *Applied Geography,* 41, 139–146. https://doi.org/10.1016/j.apgeog.2013.04.001

Johnson, S. D., Birks, D. J., McLaughlin, L., Bowers, K. J., & Pease, K. (2007). *Prospective Crime Mapping in Operational Context: Final Report.* UCL, Jill Dando Institute of Crime Science: London, UK.

Johnson, S. D., & Bowers, K. J. (2004). The burglary as clue to the future: The beginnings of prospective hot-spotting. *European Journal of Criminology, 1*(2), 237–255. https://doi.org/10.1177/1477370804041252

Johnson, S. D., Davies, T., Murray, A., Ditta, P., Belur, J., & Bowers, K. (2017). Evaluation of Operation Swordfish: A near-repeat target-hardening strategy. *Journal of Experimental Criminology*, *13*(4), 505–525. https://doi.org/10.1007/s11292-017-9301-7

Kennedy, L. W., Caplan, J. M., & Piza, E. (2011). Risk clusters, hotspots, and spatial intelligence: Risk terrain modeling as an algorithm for police resource allocation strategies. *Journal of Quantitative Criminology, 27*(3), 339–362. https://doi.org/10.1007/s10940-010-9126-2

Kennedy, L. W., Caplan, J. M., Piza, E. L., & Buccine-Schraeder, H. (2016). Vulnerability and exposure to crime: Applying risk terrain modeling to the study of assault in Chicago. *Applied Spatial Analysis and Policy, 9*(4), 529–548. https://doi.org/10.1007/s12061-015-9165-z

Levine, N. (2008). The "hottest" part of a hotspot: Comments on "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime." *Security Journal, 21*(4). https://doi.org/10.1057/sj.2008.5

Levine, N. (2013). Kernel density interpolation, In N. Levine (Ed.). *CrimeStat IV: A Spatial Statistics Program for the Analysis of Crime Incident Locations.* Houston, TX, Washington, DC: Ned Levine and Associates, National Institute of Justice.

Malmö stad (2014). *Malmö i Korta Drag.* Retrieved from https://malmo.se/download/18.578ac132166b2bd8d60711e/1541506858205/Malm%C3%B6%2Bi%2Bkorta%2Bdrag%2B181106.pdf

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (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

Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., & Brantingham, P. J. (2015). Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association, 110*(512), 1399–1411. https://doi.org/10.1080/01621459.2015.1077710

Pease, K. (1998). *Repeat Victimisation: Taking Stock* (Vol. 90). London: Home Office Police Research Group.

Pezzuchi, G. (2008). A Brief Commentary on "The Utility of Hotspot Mapping For Predicting Spatial Patterns of Crime." *Security Journal, 21*, 291–292. https://doi.org/10.1057/sj.2008.6

Polvi, N., Humphries, C., Looman, T., & Pease, K. (1991). The time course of repeat burglary victimization. *British Journal of Criminology, 31*(4), 411–414. https://doi.org/10.1093/oxfordjournals.bjc.a048138

Quetelet, M. A. (1842). *A Treatise on Man.* Edinburgh: William and Robert Chambers.

Ratcliffe, J. H. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science, 18*(1), 61–72. https://doi.org/10.1080/13658810310001596076

Ratcliffe, J. H., & McCullagh, M. J. (2001). Chasing ghosts? Police perception of high crime areas. *British Journal of Criminology, 41*(2), 330–341. https://doi.org/10.1093/bjc/41.2.330

Ratcliffe, J., H., & McCullagh, M., J. (1999). Hotbeds of crime and the search for spatial accuracy. *Journal of Geographical Systems, 1*(4), 385–398. https://doi.org/10.1007/s101090050020

Ratcliffe, J.H., & Rengert, G.F. (2008). Near repeat patterns in Philadelphia shootings. *Security Journal 21*(1–2), 58–76.

Rosenbaum, D. P. (2006). The limits of hot spots policing. In Weisburd, D. & Braga A. (Eds), *Police Innovation: Contrasting Perspectives* (pp. 245–263). Cambridge University Press

SCB1, Folkmängd i Riket, Län och Kommuner 30 Juni 2017 och Befolkningsförändringar 1 April–30 Juni 2017 [cited 2018 24/09]; Available from: https://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/befolkningens-sammansattning/befolkningsstatistik/pong/tabell-och-diagram/kvartals–och-halvarsstatistik–kommun-lan-och-riket/kvartal-2-2017/

SCB2, Befolkning, Landareal och Andel Fritidshusområden Efter Region och Typ av Område, Vart femte år 2015 2016 [cited 2018 24/09]; Available from: http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__MI__MI0811/MI0811Tab01/?rxid=a0f29188-e0c0-4392-ab0f-50ec992dbd3a

Shaw, C. R., & McKay, H. D. (1942). *Juvenile Delinquency and Urban Areas.* Chicago, IL: University of Chicago Press.

Sherman, L. W. (1995). Hot spots of crime and criminal careers of places. *Crime and Place*, 4, 35–52.

Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, *27*(1), 27–56. https://doi.org/10.1111/j.1745-9125.1989.tb00862.x

Sherman, L. W., & Weisburd, D. (1995). General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial. *Justice Quarterly, 12*(4), 625–648. https://doi.org/10.1080/07418829500096221

Short, M. B., D'Orsogna, M. R., Brantingham, P. J., & Tita, G. E. (2009). Measuring and modeling repeat and near-repeat burglary effects. *Journal of Quantitative Criminology*, *25*(3), 325–339. https://doi.org/10.1007/s10940-009-9068-8

Spelman, W. (1995). Criminal careers of public places. *Crime and Place, 4*, 115–144.

Statistics Sweden. (2013). Fortsatt Ökning av Utrikes Födda i Sverige. Available online at http://www.scb.se/sv_/Hitta-statistik/Artiklar/Fortsatt-okning-av-utrikes-fodda-i-Sverige/

Stokes, N., & Clare, J. (2019). Preventing near-repeat residential burglary through cocooning: Post hoc evaluation of a targeted police-led pilot intervention. *Security Journal*, *32*(1). https://doi.org/10.1057/s41284-018-0144-3

Townsley, M., Homel, R., & Chaseling, J. (2003). Infectious burglaries, a test of the near repeat hypothesis. *British Journal of Criminology, 43*(3), 615–633. https://doi.org/10.1093/bjc/43.3.615

Van Patten, I. T., McKeldin-Coner, J., & Cox, D. (2009). A microspatial analysis of robbery: Prospective hot spotting in a small city. *Crime Mapping, A Journal of Research and Practice*, 1(1), 7–32.

Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology 53*(2), 133–157. https://doi.org/10.1111/1745-9125.12070

Weisburd, D., Bruinsma, G. J., & Bernasco, W. (2009). Units of Analysis in Geographic Criminology: Historical Development, Critical Issues, and Open Questions. In B. W. Weisburd D., Bruinsma G.J. (Ed.). *Putting Crime in its Place* (pp. 3–31), New York, NY: Springer

Weisburd, D., Bushway, S., Lum, C., & Yang, S. M. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology, 42*(2), 283–322. https://doi.org/10.1111/j.1745-9125.2004.tb00521.x

Weisburd, D., Morris, N. A., & Groff, E. R. (2009). Hot spots of juvenile crime: A longitudinal study of arrest incidents at street segments in Seattle, Washington. *Journal of Quantitative Criminology, 25*(4), 443. https://doi.org/10.1007/s10940-009-9075-9

Appendix

Geocoding

All of the geocoded data were geocoded above the recommended 85 percent level recommended by Ratcliffe (2004), the lowest geocode rate was at 90 percent (years 2013–2015). The data were geocoded to the specific address of the crime incidents. All years combined, a few crimes (1545, about 1.7 percent) had no known address and were subsequently removed. A slush-coordinate point was used for several crimes (1459, about 1.7 percent) these were either geocoded online if possible (85 about 0.001 percent) or discarded (1374 about 1.6 percent). Furthermore, some crimes were not committed in Malmö and hence discarded (36 about 0.0004 percent).

There were some discrepancies in the geographical reliability of the police geocoded crime data. The data tended to be slightly off from the locations it was supposed to capture. It is in line with research on the geographical reliability of police data for torched cars finding that it exhibits a median error of 83 meters (Gerell, 2018). In the current study where data is aggregated to grids it should have a minor aggregate effect on the results, although specific grids may be recorded for inaccurate values. Re-projecting the data gave better results in terms of precision of the data with most points being at the recorded crime location address, whilst a few points were at the street center line instead. After re-projection, the accuracy of the data points was assessed by drawing a random number of geocoded crimes for every year (2012–2017) and checking if ArcGIS 10.3 put them in the correct position. No major projection problem seemed to remain after re-projection.

Choice of grid cell size

The grid cell size in the current study was based on a recommendation for KDE cell sizes (Chainey, 2013). The extent of the shorter side of the study area was divided by 150. So, approximately 15400 meters divided by 150 = 103m^{2,} rounded down to 100 meter by 100 meter grid cells, resulting in a total of 16737 grid cells (after being cut to the Malmö municipality border). Using a grid cell-size of one third of an average block face of the study area (Caplan et al., 2011; Hart & Zandbergen, 2014; Kennedy et al., 2011) does not work in Malmö municipality, as the inner-city and the suburban areas of Malmö municipality differ quite a bit.

Top hotspot selection process

In SC we identified the 30 percent cut-off by locating and selecting the grid cells with 30 percent of all crimes. For robbery, residential burglary and arson it was not possible to locate the top hotspots (grid cells with about 30 percent of all crime) using this technique. For example, using arson 2016 to predict arson in 2017, grid cells with two or more crime incidents amounted to 19 percent of all crimes and grid cells with one or more crime incidents were 100 percent of all crime incidents. Hence, either this prediction would contain 19 or 100 percent of all crimes. This would make comparison to other crime types hard. For these crime types, we located the top hotspots by selecting 30 percent of all crimes using standard deviation in the KDE analyses.

The KDE standard deviation method of selecting top hotspots was also calculated and compared to the SC method of selecting top hotspots for all crime types. The differences produced by these two different methods of calculating top hotspots was that the KDE standard deviation method generally produced lower PAI values for comparison, due to the greater area rate used in this method. In other words, more 100- by 100-meter grid cells were generally needed to reach 30 percent of all crimes. The patterns were similar for both the SC and the KDE standard deviation cut-offs no matter crime type. Hence, it is unlikely that the method of locating the cut-offs substantially alter the main findings. We furthermore examined less hot hotspots, where 50 percent of all violent and property crime occurred and where 70–80 percent of all violent and property crime occurred. These 50 percent and 70–80 percent cut-offs also produced similar patterns when comparing KDE and SC predicting violent and property crime.

The search for the “best KDE model”

To find the best model of KDE, violent crime in 2016 was used to predict violent crime in 2017. A normal interpolation was used, due to the recommendation from Levine (2013), that with samples smaller than 10.000 points a normal interpolation should be used, due to the risk of finding “false hotspots” otherwise. The quartic and triangular interpolation methods were also examined due to the potential better prediction accuracy, as seen in previous studies (Drawve, 2016; Hart & Zandbergen 2014). Choice of cell size potentially does not affect the outcome of the KDE, as the other parameters do (Chainey, 2013; Hart & Zandbergen, 2014). The same cell size and the same method for choosing this cell size was used as in SC, 100 by 100 meters.

A starting point for the choice of bandwidth was to take the cell size times five (Chainey, 2013) that rendered a bandwidth of 500 meters (100 meters times five). Earlier research however shows that smaller bandwidth values may increase the predictability accuracy of crime patterns (Chainey, 2013; Hart & Zandbergen, 2014). Because of this, bandwidths of 400, 300, 200, 100 meters were also used. The KDE was also employed with the adaptive bandwidth as suggested by Pezzuchi (2008) and used in other research (Mohler et al., 2011). With an adaptive bandwidth, the bandwidth will be smaller in high crime areas and larger in areas with less crime. A minimum number of 25 points within the bandwidth radii was chosen to get a finer grained estimate rather than a smoother one. For each bandwidth experimentation, a cell size of 100 by 100 meters was used.

Table 5. KDE PAI values for different crime types predicting where people will commit crime geographically in 2017

	2012	2013	2014	2015	2016	2012-2016	2013-2016	2014-2016	2015-2016	Average value
Violent Crime	70.16	57.29*^	62.87*^	62.41^	62.83^	61.36^	63.85^	66.12^	69.49	64.04
Assault	83.08^	87.29^	71.65^	97.25	90.86^	77.45^	73.76^	71.25*^	75.91^	80.94
Robbery‡	29.45*^	32.02^	27.17*^	24.40*^	53.08*	40.99^	39.94^	40.76^	45.98*^	37.09
Property Crime	15.70*^	19.17*^	19.78^	19.37*^	23.80	19.80*^	20.70*^	21.28*^	21.87*^	20.16*
Residential burglary‡	9.40*^	12.61*	12.01	10.96*^	9.84^	11.41*^	12.20*	11.66*^	10.08*^	11.13*
Theft	24.92*^	33.47^	31.43*^	44.36^	67.98*	34.50*^	37.06*^	43.81^	53.74*^	41.25*
Vehicle theft	23.04^	27.06^	27.86^	29.72^	34.51	28.96^	31.13^	31.86^	33.03	29.69
Property damage	33.56*^	34.16^	32.53*^	33.48*^	36.14*	33.42*^	32.77*^	35.55*^	37.94*	34.39*
Arson‡	9.60^	12.80^	38.39^	24.24*^	29.25*^	43.87*^	43.40*^	49.90*^	83.17*	37.18
Average value	33.21^	35.10^	35.97^	38.47^	45.37^	39.08^	39.42^	41.35^	47.91	39.54

Note: ‡ denotes the use of standard deviation in KDE rather than count in SC. * denotes a >5 percent difference between the PAI values of KDE and SC. ^ denotes a >5 percent difference between the PAI values of the different years. Base year is the top PAI value for each crime type. Base year for average value ^ is 2015–2016.

Table 6. Simple Count PAI values for different crime types predicting where people will commit crime geographically in 2017

	2012	2013	2014	2015	2016	2012-2016	2013-2016	2014-2016	2015-2016	Average value
Violent Crime	72.48	62.40*^	66.72*^	64.65*^	66.83^	63.73^	66.35^	68.62^	70.69	66.94
Assault	81.03^	90.45	74.36^	93.10	89.71	80.94^	75.19^	75.19*^	77.00^	81.86
Robbery‡	23.56*^	32.02^	19.41*^	21.96*^	60.01*	42.13^	39.94^	42.70^	41.60*^	35.93
Property Crime	16.75*^	20.67*^	20.39^	20.40*^	24.72	21.24*^	21.98*^	22.67*^	23.30*^	21.35*
Residential burglary‡	8.19*^	8.11*^	11.64	9.98*^	9.84^	10.25*^	11.00*^	10.45*^	8.64*^	9.79*
Theft	26.91*^	33.47^	33.59*^	42.80^	76.99*	36.85*^	39.05*^	46.10^	56.98*^	43.64*
Vehicle theft	23.27^	26.55^	27.37^	29.41^	34.34	29.80^	30.86^	31.73^	32.89	29.58
Property damage	36.00*^	35.65^	38.48*^	39.02*	41.82*	37.19*^	40.63*	38.90*^	41.31*	38.78*
Arson‡	9.60^	12.80^	38.39^	64.65*	51.18*^	34.47*^	36.72*^	38.39*^	44.79*^	36.78
Average value	33.09^	35.79^	36.71^	42.86^	50.60	39.62^	40.19^	41.64^	44.13^	40.52

Note: ‡ denotes the use of standard deviation in KDE rather than count in SC. * denotes a >5 percent difference between the PAI values of KDE and SC. ^ denotes a >5 percent difference between the PAI values. Base year is the top PAI value for each crime type. Base year for average value ^ is 2016.

Table 7. KDE RRI values for different crime types predicting where people will commit crime geographically in 2017

	2012	2013	2014	2015	2016	2012-2016	2013-2016	2014-2016	2015-2016	Average value
Violent Crime	0.59*^	0.59^	0.72*^	0.65^	0.72^	0.80*^	0.82	0.85*	0.79^	0.73
Assault	0.67*^	0.69*^	0.75*^	0.73*^	0.77*^	0.86*	0.84*	0.81^	0.81*^	0.77*
Robbery‡	0.17*^	0.17*^	0.13*^	0.16*^	0.35*^	0.56*^	0.54*^	0.61	0.43*^	0.35*
Property Crime	0.67^	0.78^	0.88^	0.81^	0.83^	0.94	0.94	0.94	0.89	0.85
Residential burglary‡	0.29*^	0.29*^	0.30*^	0.29*^	0.27*^	0.67*	0.62*	0.57*^	0.43*^	0.41*
Theft	0.54^	0.75*^	0.80^	0.84*^	0.72^	0.92	0.90	0.88	0.83^	0.80
Vehicle theft	0.60*^	0.75*^	0.83*^	0.86*^	0.85*^	0.96	0.98*	1.00*	0.98*	0.87*
Property damage	0.66^	0.72*^	0.70*^	0.71^	0.74^	0.87	0.84*^	0.91	0.84^	0.78
Arson‡	0.03^	0.03^	0.14^	0.08^*	0.42*	0.39*	0.39*	0.39*	0.42*	0.25*
Average value	0.47^	0.53^	0.58^	0.57^	0.63^	0.77	0.76	0.77	0.71^	0.65*

Note: ‡ denotes the use of standard deviation in KDE rather than count in SC. * denotes a >5 percent difference between the RRI values of KDE and SC. Base year is the top RRI value for each crime type. Base year for average value ^ is 2015–2016.

Table 8. Simple Count RRI values for different crime types predicting where people will commit crime geographically in 2017

	2012	2013	2014	2015	2016	2012-2016	2013-2016	2014-2016	2015-2016	Average value
Violent Crime	0.55*^	0.58^	0.68*^	0.62^	0.69^	0.76*	0.78	0.79*	0.76	0.69
Assault	0.57*^	0.61*^	0.70*^	0.63*^	0.66*^	0.81*	0.78*	0.77	0.74*^	0.70*
Robbery‡	0.14*^	0.18*^	0.09*^	0.14*^	0.38*^	0.53*^	0.47*^	0.61	0.37*^	0.32*
Property Crime	0.66^	0.77^	0.86^	0.79^	0.80^	0.93	0.94	0.93	0.87^	0.84
Residential burglary‡	0.24*^	0.17*^	0.27*^	0.24*^	0.25*^	0.57*	0.53*^	0.47*^	0.35*^	0.34
Theft	0.54^	0.70^*	0.78^	0.74^*	0.77^	0.92	0.90	0.88	0.83^	0.78
Vehicle theft	0.57*^	0.69*^	0.77*^	0.81*^	0.80*^	0.95	0.93*	0.94*	0.93*	0.82*
Property damage	0.64^	0.68*^	0.77*^	0.73^	0.75^	0.89^	0.95*	0.91	0.82^	0.79
Arson‡	0.03^	0.03^	0.14^	0.21*^	0.19*^	0.31*	0.29*^	0.28*^	0.22*^	0.19*
Average value	0.44^	0.49^	0.56^	0.55^	0.59^	0.74	0.73	0.73	0.65^	0.61

Note: ‡ denotes the use of standard deviation in KDE rather than count in SC * denotes a >5 percent difference between the RRI values of KDE and SC. Base year is the top RRI value for each crime type. Base year for average value ^ is 2015–2016.