Harinam *et al. Crime Science* (2022) 11:14
https://doi.org/10.1186/s40163-022-00176-x

Crime Science

RESEARCH Open Access

Spatial distribution and developmental trajectories of crime versus crime severity: do not abandon the count-based model just yet

Vincent Harinam[1*], Zeljko Bavcevic[2] and Barak Ariel[3,4*]

Abstract

Purpose/background: A new body of research that focuses on crime harm scores rather than counts of crime incidents has emerged. Specifically in the context of spatial analysis of crime, focusing on crime harm suggests that harm is more concentrated than counts, at the level of crime hot spots. It remains presently unclear what drives the concentration distributions, and whether the count-based model should be abandoned.

Methods: Cross-sectional and longitudinal analysis of 6 year of spatiotemporal crime data in Toronto, Canada, to compare patterns and concentration of crime harm (measured in terms of the Crime Severity Index (CSI) against crime counts. Conditional probabilities, trajectory analyses, power few concentrations, and spatial Global Moran's I are used to infer generalised trends from the data.

Findings: Overall CSI and crime counts tend to exhibit similar concentrations at the spatial micro levels, except against-the-body crimes such as violence which seems to drive nearly all the variations between the two measurement types. Violence harm spots tend to be more dispersed citywide and often do not remain constant year-to-year, whereas overall crime hotspots are more stable over time. Nevertheless, variations in disproportionally high crime hot spots are associated with total variations in crime, with as little as 1% increase in crime levels in these hot spots translating into substantial overall gains in recorded crime citywide.

Conclusions: Abandoning count-based models in spatial analysis of crime can lead to an incomplete picture of crime concentrations. Both models are needed not just for understanding spatial crime distributions but also for cost-effective allocation of policing resources.

Keywords: Hot spots, Harm spots, Crime counts, Crime harm index, Spatial analysis

Introduction

Crimes are non-random events, distributed unevenly in space and time. Dubbed "the law of crime concentration" by Weisburd (2015); see also Weisburd et al., (2012), a small percentage of geospatial units account for an outsized proportion of counted crimes in both urban and nonurban settings (e.g., 5% of street segments account for 50% of crimes). Moreover, this power law distribution exhibits marked temporal stability, with high crime locales present year-on-year (Andresen et al., 2017a, 2017b; Groff et al., 2010; Weisburd et al., 2004; but *cf.* Weinborn et al., 2017). Importantly, while traditional place-based analyses of crime have focused on macro and meso-geographic units of analysis, the "criminology of place" (Sherman et Al., 1989; Weisburd et al., 2012) prioritizes micro-places such as street segments between

*Correspondence: vh315@cam.ac.uk; ba285@cam.ac.uk

[1] Institute of Criminology, University of Cambridge, Sidgwick Ave, Cambridge CB3 9DA, UK
[3] Experimental Criminology, Institute of Criminology, University of Cambridge, Sidgwick Ave, Cambridge CB3 9DA, UK
Full list of author information is available at the end of the article

intersections, specific addresses, or specific "risky" facilities.

While the criminology of place has altered discussion parameters within criminological circles, a growing and largely unexamined line of research has argued that focusing solely on counts is an imprecise and ineffective practice (Sherman, 2013:422): "all crimes are not created equal; some crimes cause horrible injuries and deaths. Others cause scant harm to anyone." With ever-shrinking police budgets, not all crime can (or should) be policed equally. Indeed, a triage approach which prioritizes high severity crimes over low severity crimes is required. Moreover, given the robustness of the criminology of place, it may be time to shift from a count-based hot spots model to a model which incorporates harm spots into the existing framework. It is important to note that is not universally held as while crime harm is established in the UK and US, this is not necessarily the case in parts of Europe.

However, this call to shift our focus from a count-based model—in which all crime types are given the same weight—to a policy that is sensitive to the harm caused by a crime, has largely remained unanswered. In this paper, we answer the call for more research that focuses on crime harm over crime counts, specifically in the context of place-based criminology. We assess the spatial concentration of crime counts versus crime harm in Toronto, Canada.

The purpose of this study is threefold. First, we seek to replicate study comparing the spatial concentration of crime counts and crime harm outside the United Kingdom to test the generalizability of these findings. Our study adds to this emerging literature by observing crime trends within hexagonal tessellations as its spatial unit of analysis no longer than 124 m long—as opposed to much larger spatial units (e.g., Massey et al., 2019). To estimate crime harm, we utilize the Canadian Crime Severity Index (CCSI; see Hiltz et al., 2020) to assign weights to the crime categories in our data. For the purposes of study, it is important to note that crime harm and crime severity are the same concept. Second, we seek to compare, via trajectory analysis, the spatial concentration of crime hot spots and harm spots over time at the aggregate level and across crime types. The literature on spatial concentrations of crime demonstrates the importance of examining individual crime types (see Andresen, 2007; Weisburd & Mazerolle, 2000). This, however, has not been applied to crime harm. As a natural extension of both the crime and place and crime harm literatures, we conduct our analyses in a number of dimensions of crime: violent crime, auto-theft, burglary, robbery, and theft. Finally, as the temporal patterns of crime are understudied, specifically in the context of hot spots, we highlight their significance in the study using a novel measurement of temporal crime harm trends.

Literature review

Concentration of crime in micro-places

In one of the first examinations of crime patterns at the micro-spatial level, Sherman et al. (1989) found that 1.2%, 2.7% and 2.2% of addresses and intersections in Minneapolis accounted for 50% of rapes, motor vehicle thefts and robberies, respectively. Following this, Weisburd et al. (2004) reported that 50% of crime incidents occurred in 4.5% of street segments in Seattle. These results appear to be ubiquitous (Eck et al., 2017; Lee et al., 2017). In Canada, Curman et al. (2014) attempted to replicate Weisburd et al. (2004) and found that 7.8% of street segments and intersections in Vancouver accounted for 60% of the city's crime. More recently, Andresen et al. (2017a) and Andresen et al. (2017b) have demonstrated that disaggregated crime types (assault, burglary, robbery, etc.) are highly concentrated in space with stable trajectories across time, and that property crimes are highly concentrated in a small percentage of street segments but demonstrate spatially stable trajectories in the most recent years.

Given the consistency of these findings, Weisburd (2015) has proposed an empirical principle which explains the spatial distribution of crime at micro-levels: the law of crime concentration. This law stipulates that "for a defined measure of crime at a specific micro-geographic unit, the concentration of crime will fall within a narrow bandwidth of percentages for a defined cumulative proportion of crime" (Weisburd, 2015:138). We note that an Auerbach or Zipf distribution is expected with these datasets, but the contribution of this law of concentration is the tight bandwidth in which crime concentrates in small areas of land called hot spots.

Concentration of harm in micro-places

According to Greenfield and Paoli (2013:864), "neither criminology nor the adjacent social sciences have made a serious effort to systematically identify, evaluate or compare the harms associated with different crimes." This is particularly the case in place-based analyses of crime and disorder (Macbeth & Ariel, 2019).

While many were never put to use for practical and theoretical purposes, there are several methods of measuring crime harm. These include measuring the moral culpability of an offender (Hall, 1960); calculating the "true" cost of crime (Brand & Price, 2000; Cohen & Bowles, 2010; see more broadly; weighting each offence based on the average sentence meted out to offenders (Wallace et al., 2009); and applying gravity score guidelines (Ratcliffe, 2015). A UK-based model, Sherman et al.'s

(2016) Cambridge Crime Harm Index (CHI), which is based on the sentencing guidelines of England and Wales. The CHI reflects the recommended number of days of imprisonment for first time offenders without extraordinary circumstances accused of a crime. CHI systems that rely on sentencing guidelines have been used in several jurisdictions (e.g., Andersen & Mueller-Johnson, 2018; House & Neyroud, 2018; Mitchell, 2019; Ransley et al., 2018; see also the Home Office 2019 for comparison).

Despite the growing popularity of crime harm indices (e.g., Andersen & Mueller-Johnson, 2018; Bland & Ariel, 2020; Carter et al., 2021; Frydensberg et al., 2019; Link & Losel, 2022; Simon & Kichova, 2020; see also review in Ashby, 2018, but *cf.* critique in Sarnecki, 2021), the list of studies examining crime harm remains short. Weinborn et al. (2017) compared the spatial concentration of crime counts and crime harm in Birmingham, UK. The authors generated "harm spots" using Sherman et al.'s (2016) CHI to weight crime incidents. The results of this analysis indicated that harm-weighted crime is more concentrated than raw crime counts. The authors found that 50% of crime events were concentrated in 3% of all street segments, whereas 50% of harm was concentrated in only 1% of street segments. Based on these findings, crime harm is more likely than crime counts to adhere to Weisburd's (2015) "law of crime concentration". However, Fenimore (2019), examining crime counts and harm in Washington, D.C., found that counts and harm were equally concentrated in space, which may be indicative of the types of crimes that are more prevalent in the study locations: harm is driven by *either* the infrequent but more serious crime categories (e.g., against-the-person crime categories in the night-time economy) *or* low-level harm but voluminous categories (e.g., theft from person in the rail environment). Nevertheless, the benefits from using a recording system that is sensitive to the harm caused to society, rather than a binary measure of crime, are well noted (Sherman, 2020).

Present study: comparing hot spots and harm spots using the Toronto spatiotemporal crime harm index

While the scholarly literature on crime at place and crime harm is robust, there are several gaps which this study seeks to fill. In particular, we know very little about the developmental trajectory of crime harm at micro-places. To this extent, it is unclear whether high harm micro-places remain as such year-on-year. Moreover, we compare the spatial concentration of crime counts and crime harm. Crucially, the scholarly literature is bereft of studies comparing the spatial concentration of crime hot spots and harm spots over time at the aggregate level and across crime types. As such, we seek to understand the developmental trajectory of crime types by crime count and harm. Finally, we attempt to demonstrate the utility of bespoke hexagonal tessellations as a spatial unit of analysis comparable to street segments (Weisburd & Amram, 2014).

Methods

Settings

The city of study is Toronto, the provincial capital of Ontario. The city covers 630 km^2 (243 square miles) and has a shoreline of 46 km on Lake Ontario. Toronto is comprised of 140 neighbourhoods within six boroughs: Downtown Toronto, East York, Etobicoke, North York, Scarborough and York, home to approximately 6.3 M residents in the greater Toronto region who earn approximately $104,000 p.a., but with 20.2% low-income families in the City of Toronto (www.toronto.ca, 2022). To contextualize the crime levels in the city of study, according to official statistics in Canada (Statistics Canada, 2020), the census metropolitan areas with the lowest crime severity indices (CSIs) were Québec (42.1, out of a benchmark of 100), Saguenay (43.9) and Barrie (44.9), followed by Toronto (46.2), Trois-Rivières (47.7), Sherbrooke (47.8) and Ottawa (48.3).

Data and procedures

Crime counts

This study uses data from the Toronto Police Service's Open Data Portal. Two datasets are used and juxtaposed for the purpose of the analyses: the Major Crime Indicator (MCI) dataset and the Toronto Police Homicide dataset. The MCI dataset contains data from crimes occurring between 2014 and 2019 and covers five crime categories: assault, auto-theft, burglary, robbery and theft. These crime categories are aggregates, grouping 49 different offences (see Additional file 1: Materials S1 for a full list). From a GIS perspective, the crime location occurrences are deliberately offset to the nearest road intersection to protect the privacy of the individuals involved in each event. The second dataset is the "Homicide ASR RC TBL 002" dataset, which included all homicides from 2004 to 2019.

As the MCI dataset is recorded at the "offence and/or victim" level, there may be multiple rows in the dataset for each victim present at each crime event. These duplicate rows inflate crime counts within each MCI by measuring raw counts of occurrences as opposed to distinct counts of occurrences. All were reduced to a single incident by deleting rows with a duplicate crime event ID. Furthermore, as the MCI data were comprised of crime events occurring between 2014 and 2019, the homicide data, which ranged from 2004 to 2019, was filtered to remove all pre-2014 homicides. These datasets were merged into a single file, with homicides being added to the assault MCI. This aggregate crime category was renamed "violent crimes". In total, we counted 180,867 offences in this dataset, or roughly 30,000 recorded crimes per year.

Crime severity index

Next, the study utilizes the Canadian Crime Severity Index (CSI) to assign severity weights. Whereas Sherman et al.'s (2016) Cambridge Crime Harm Index (CHI) calculates a severity score from baseline or "starting point" sentences established by the English-Welsh sentencing guidelines, CSI weights are based on the violation's incarceration rate, as well as the average length of prison sentence handed down by criminal courts in all Canadian provinces and territories. To create a severity weight for a specific offence, the incarceration rate for the offence is multiplied by the mean sentence length. Weights are based on the five most recent years of available sentencing data, and are updated every five years in order to account for changes in sentencing patterns. The most recent update was carried out in 2018 and is utilized by this study.

Overall, there were 49 offence types across the five MCIs for which a CSI score was available. The weakest sanction was "Assault—Level 1", which possessed a severity weight of 26.39 prison days. "Murder 1st Degree", the most serious offence, yielded an average of 7656.16 days in prison.

Spatial unit of analysis and geospatial procedure

Studies of developmental trajectories (Andresen et al., 2017a, 2017b; Groff et al., 2010) have used street segments as spatial units of analysis. While adequate as geospatial units, street segments have two glaring limitations: variable size and data loss. First, the length of street segments within a centerline map is inconstant and based on features of the physical geography. As a result, larger street segments may capture more crimes than smaller street segments. Second, street segments omit crime events that take place at intersections. The resultant level of non-random data loss can be considerable, such that analyses capture an incomplete picture of crime at a place.

To overcome the shortcomings of a street segment-based approach, the present analysis employs custom-generated "hexagonal tessellations" of the City of Toronto. The concept of a hexagon is an elementary pillar of Euclidian geometry: a shape on a flat surface with six equal straight sides and six equal angles. Our reasons for preferring six-sided areas are two-fold: (1) mitigating the "border issue", and (2) reducing variance in surface area. According to Zhang et al. (2012), the border issue is a methodological flaw in which specific occurrences cluster around geospatial borders due to the use of administrative boundaries. A custom hexagonal layer overcomes this issue by creating sub-areas to capture spatial variations in offences while minimizing classification problems. Second, the use of custom-generated hexagons reduces differences in surface area that are present when using census-based administrative zones.

Developments in digital analysis now allow anyone to design their own borders of urban areas, with different numbers of hexagons of widely varying sizes. Rather than accepting the limitations of areas drawn for historic reasons of administrative data collection, analysts can deploy computer programs to break down a city into equally sized hexagons. These allow crime analysts to be more sensitive to the needs of operational decision-making, especially the allocation of scarce resources in different micro-areas.

We initially subdivided the City of Toronto into 65,170 hexagons by using the method of "tessellation" or tiling, in which an arrangement of shapes is closely fitted together in a repeated pattern without gaps or overlapping. These hexagonal zones were 124.08 m (side-to-side), with a combined surface area of 10,000 square metres. This specific size was randomly selected. However, because the layer was clipped to the City of Toronto municipal boundary, not all hexagons were uniform. Our data were subsequently joined to this layer using the Spatial Join feature in ArcMap software.

We note that a risk of using this aggregate layer is that a number of hexagonal tessellations would cover areas where crime is unlikely to occur due to a lack of human activity (e.g., lakes or industrial zones). This would inflate the zero term, as there would be an artificially higher number of hexagons with no reported crimes during the study period. In order to reduce zero inflation, all hexagonal zones that did not intersect with a linear road feature were removed. This was done via a spatial selection in ArcMap and produced a reported data loss of less than 1% for each MCI. Following this procedure, a total of 51,673 hexagons remained (or 79.3% of all hexagons). These hexagons were each 124 m long, balancing granularity and sufficient sample sizes per unit.

Analytic approach

Aside from standard descriptive statistics which summarize spatial and temporal crime trends, we employ k-means longitudinal clustering to identify the trajectory of hexagonal subgroups based on their crime count and crime severity. Originally conceived by Calinski and Harabasz (1974), k-means is a non-parametric statistical technique used to analyze longitudinal data with the goal of identifying clusters of cases that

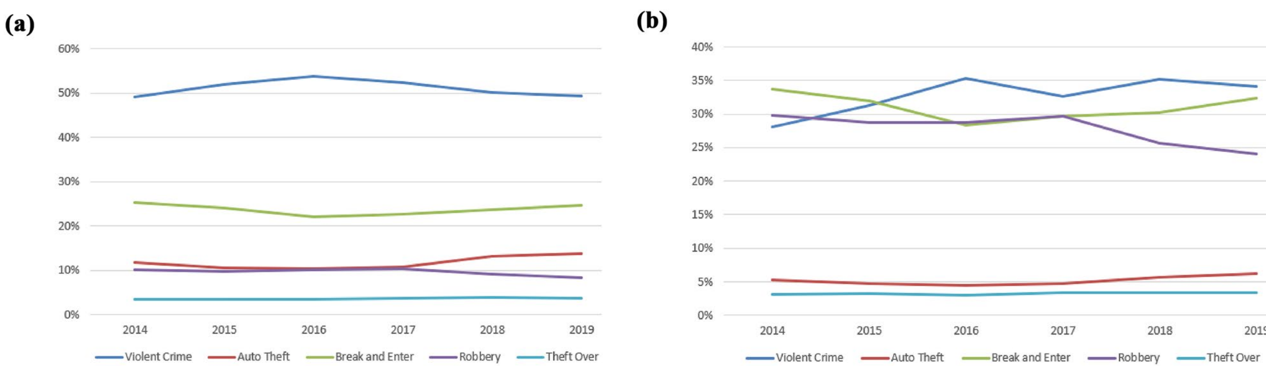

Fig. 1 **a** Year-on-Year MCI Count Distribution. **b** Year-on-Year MCI Severity Distribution

share similar traits (Genolini & Falissard, 2010). Unlike group-based trajectory modelling (Groff et al., 2010; Nagin & Land, 1993; Weisburd et al., 2004; Wheeler et al., 2015), k-means longitudinal does not require data to fit a specific distribution and is better able to accommodate larger counts.

More specifically, k-means is a hill-climbing algorithm belonging to the Expectation–Maximization (EM) class of algorithms. EM algorithms initially assign each observation to a cluster, then progress towards optimal clustering by recomputing each cluster and moving each observation to its "nearest cluster" (Genolini & Falissard, 2010). This is repeated until no further changes occur in the clusters.

All trajectory models were constructed in *R* using the KmL package to estimate k-means longitudinal clusters. An unsettled problem with k-means is the need to know, a priori, the number of clusters. We mitigate this limitation by using the Calinski criterion to evaluate the various trajectory solutions and identify the optimal number of trajectory groups for each crime type. As such, the models in this study possess four to six trajectories befitting their optimality. The Calinski criterion is a relative metric that compares the different group solutions. Importantly, the criterion used in our trajectory models was constructed for both the total crime count and total crime severity, for each of the five MCIs. There were twelve models in total, with four to six trajectories in each model.

To summarize, our first analysis consists of spatial crime concentrations measured for crime count and severity, and considering the full extent of hexagonal tessellations with and without a recorded crime and with only a recorded crime. Second, we conduct trajectory analyses of crime count and severity at hexagons and visualize the results to find differences at the aggregate level and across crime types.

Findings

Descriptive results

Figure 1A and b present the year-on-year percentage distributions of each MCI for crime count and crime severity, respectively (see Additional file 1: Materials S2 for the full breakdown). Violent offences are the most frequently occurring counted crime, comprising 49–54% of all crimes in a given year. Burglaries are the second most frequently occurring counted crime, comprising 22–25% of all crimes across the time series. In general, there is a marked consistency in the distribution of counted MCIs from 2014 to 2019. It is important to stipulate that this data does not include theft under which generally constitutes an outsized proportion of crime reported to Canadian police forces.

Nevertheless, crime severity exhibits a very different pattern. Violent crime and burglary are interchangeable, competing for the highest severity MCI across the time series. Violent crime accounted for 28% to 35% of crime severity in a given year, while burglaries ranged between 28 and 34%. Curiously, the crime count and severity of violent crime are not proportional as the weighted severity exceeds the number of counted crimes relative to other MCIs. This can be attributed to murders as they represented 794 counted crimes over the time series but produced a total severity of 3,299,805. Put differently, murders made up 0.86% of all counted crimes in the violent crime MCI but accounted for 34.4% of the total severity.

Spatial distribution of crime count and crime severity

Figure 2A and b present the year-on-year spatial concentration of crime count and crime severity at hexagons in Toronto by quartiles. Examined across the time series, Toronto demonstrates a tight bandwidth of percentages for hexagons accounting for 25%, 50%, 75%, and 100% of both crime count and crime severity. These

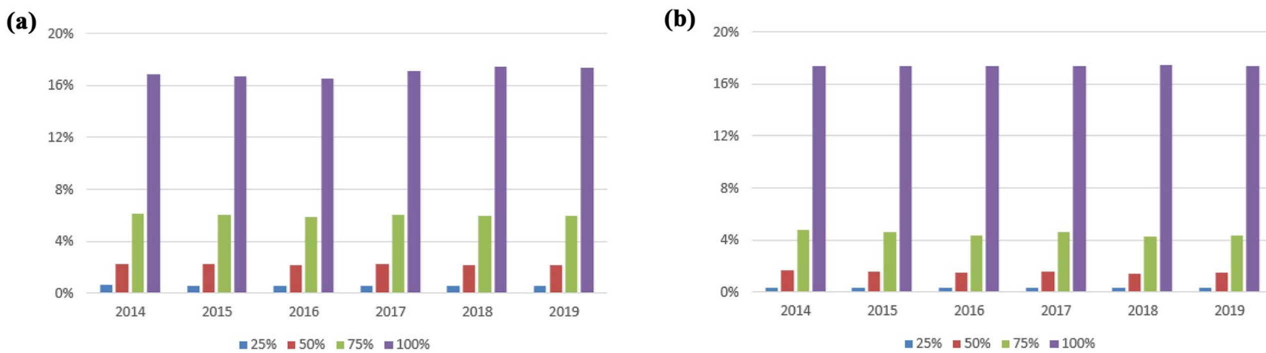

Fig. 2 **a** Distribution of hexagons that account for crime counts by quartiles (year). **b** Distribution of hexagons that account for crime severity by quartiles (year)

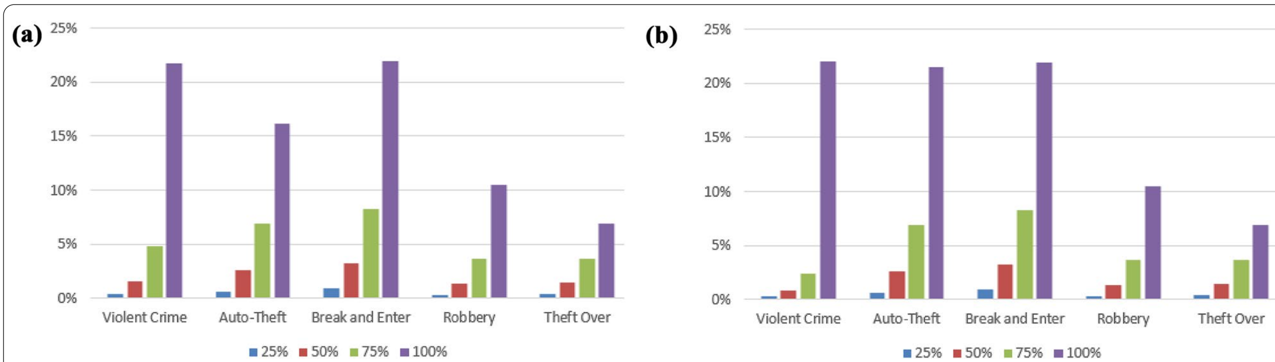

Fig. 3 **a** Distribution of hexagons that account for crime counts by quartiles (MCI). **b** Distribution of hexagons that account for crime severity by quartiles (MCI)

percentages do not vary substantially between 2014 and 2019. For example, 50% of counted crimes occurred in 0.56% to 0.58% of hexagons in a given year, whereas 50% of crime severity occurred in 0.32% to 0.36% of hexagons. It is clear from these findings that crime severity is more spatially concentrated than counted crime as computed with MCI data, with an approximately 32–37% change between count and harm model. Within the first, second and third quartiles, crime severity is more spatially concentrated relative to crime count.

Figure 3a and b present the spatial concentration of crime count and severity at hexagons for each MCI in quartiles. Crime count and crime severity are equally or near equally spatially concentrated at all quartiles across auto-theft, burglary, robbery, and theft. However, the crime severity of violent crime is more concentrated in space relative to crime count.

It may be assumed that a lower overall percentage of street segments with crime equates to greater micro-level clustering. However, percentage distributions, while useful for validating the law of crime concentration, should be cautiously considered when the total number of crimes in given year is exceeded by the total number of spatial units within the study. By definition, such crimes **must** have a high degree of "concentration" because there are fewer criminal events than there are spatial units. This caveat is applicable here as, with the exception of the total number of violent crimes (n = 92,526), the total number of spatial units in this study (n = 51,672) exceeds both aggregate counts and yearly totals. This is particularly problematic when there are more hexagons without a recorded crime event than there are crimes in a given year.

As Fig. 4a and b indicate, over 80% of hexagons in any given year had no counted crime or recorded crime severity. Moreover, less than 1% of hexagons in any given year contained 16 or more counted crimes, whereas less than 5% of hexagons in a given year contained a weighted severity greater than 500. As such, the aforementioned percentages do not necessarily offer a clear rendering of the spatial concentration of crime count and crime severity in Toronto over time (See Fig. 5a and b). Nevertheless, we can still determine the level of spatial concentration

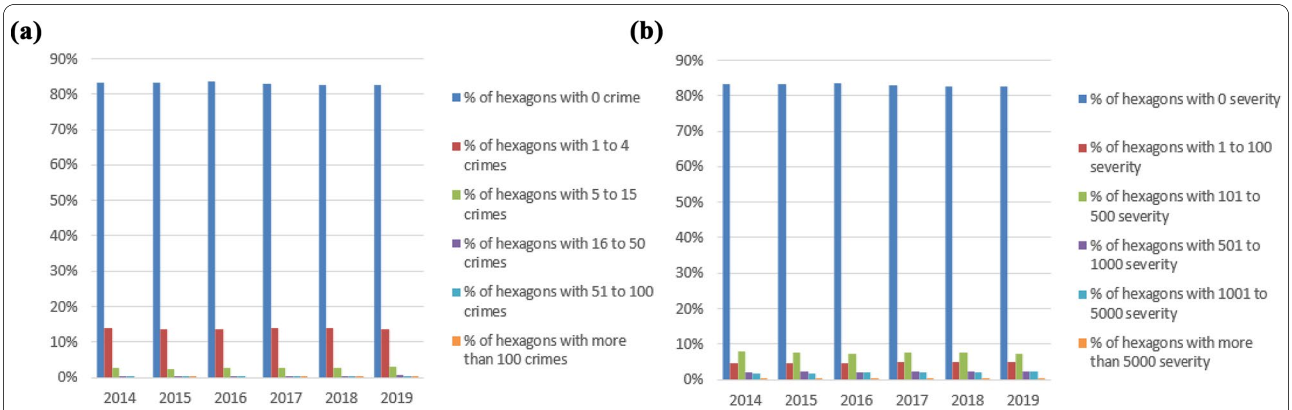

Fig. 4 **a** Distribution of hexagons by crime count. **b** Distribution of hexagons by crime severity

for crime count and severity by only examining hexagons with a recorded crime event.

Figure 6a and b present power law distributions of crime count and severity over the 6-year time series, respectively. Importantly, these figures exclude all hexagons without a recorded crime. From these results, crime severity is more spatially concentrated than crime count across the times series and in individual years—but only slightly so. Whereas 50% of counted crimes were found in 8.4% of hexagons between 2014 and 2019, 7.3% of hexagons accounted for 50% of crime severity during this period. A similar pattern is observed across all individual years. While 13.4%, 13.4%, 13.3%, 13.1%, 12.4% and 12.4% of hexagons accounted for 50% of the crime count in 2014, 2015, 2016, 2017, 2018 and 2019, respectively, 9.9%, 9.5%, 9%, 9.4%, 8,2% and 8.7% of hexagons accounted for 50% of crime severity in the same years. Furthermore, these percentages reflect stable levels of spatial concentration across time for both crime count and crime severity. These findings parallel observations made by where 50% of counted crime and crime severity were concentrated in 3% and 1% of street segments, respectively. To deal with zero inflation, we have opted omit hexagons without a crime event. An alternative solution would have involved the use of the generalized Gini coefficient (Bernasco and Steenbeek, 2017).

Figure 7a and b present power distributions of crime count and severity, respectively, for each MCI. Crime count and crime severity demonstrate equal levels of spatial concentration across auto-theft, burglary, robbery, and theft. To this extent, 16.4%, 14.9%, 12.8% and 20.8% of hexagons accounted for 50% of both crime count and severity for auto-theft, burglary, robbery, and theft-over, respectively. Crime severity is a weighted multiple of the crime count and, as such, the cumulative distribution function used to calculate these distributions produced similar results for crime count and crime severity.

However, this is not the case for violent crime, as crime severity is more spatially concentrated than crime count with 3.7% of hexagons accounting for 50% of crime severity, while 7.4% of hexagons accounted for 50% of counted crimes. This difference is largely attributed to variances in the weighted value of offence types within the violent crime MCI. Crime severity in this category ranged from 26.39 (Assault—Level 1) to 7656.1 (first degree murder). This variance is not observed in the four other MCIs as there are a lower number of unique offence types with similar severity weights.

Table 1 presents the Global Moran's I of counted crimes and crime severity in each year. The Moran's I measures the autocorrelation of a spatial feature class (Li et al., 2007). We calculate the Moran's I using queen's continuity. It measures how similar a feature class is to those surrounding it—and in the case of hot spots, two hexagons bordering each other. If objects are attracted (or repelled) by each other, it means that the observations are not independent. As such, the Global Moran's I is a metric which compares the standardized magnitude of difference between the spatial concentration of counted crimes relative to crime severity. Based on the results presented in Table 1, it is evident that crime severity is more spatially variable than counted crimes. Indeed, in each year, counted crimes possess a higher Global Moran's I relative to crime severity. Crime severity demonstrates greater "street-by-street variability" (Weisburd et al., 2012) than counted crime. Moreover, while crime severity is more spatially concentrated than counted crimes, these hotspots are not clustered altogether. Rather, high crime severity hexagons are mostly diffused throughout Toronto than are hexagons with high crime counts.

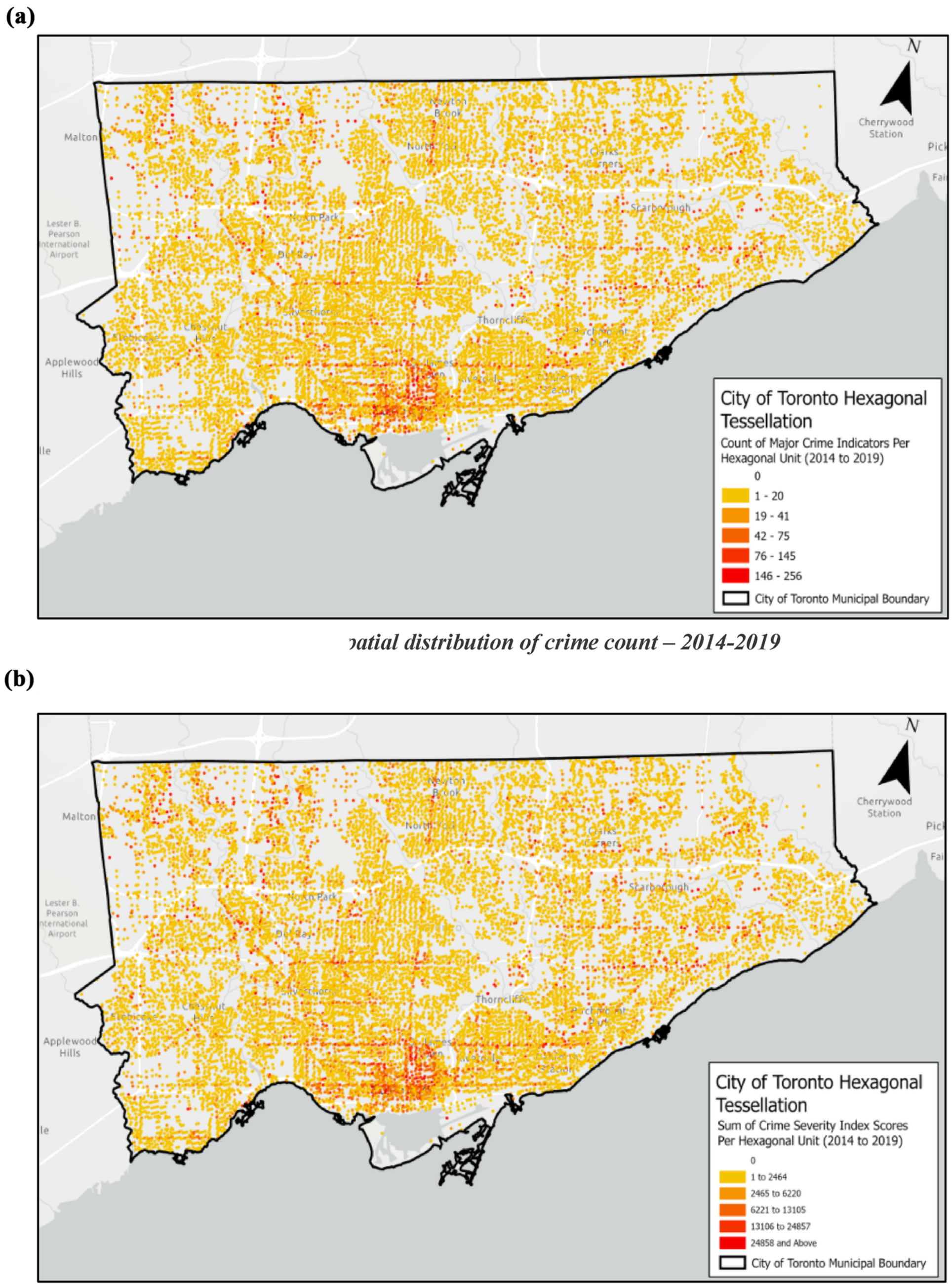

Fig. 5 **a** Spatial distribution of crime count—2014–2019. **b** Spatial distribution of crime severity—2014–2019

Trajectory results

The results from the k-means trajectory analyses are summarized in Table 2. This table shows the various MCIs by measurement, the number of trajectories, the level of crime for each trajectory (relative to the specific MCI), its base crime count or severity in 2014, the trend, and the percentage of hexagons within each trajectory group. As with previous research (Curman et al., 2014; Weisburd et al., 2004), these trends are defined by regression analyses of the hexagons over time within each trajectory group. If the slope parameter on the time axis is close to zero (e.g. between −0.2 and +0.2), the trajectory is considered stable. The trajecory is considered decreasing if the slope parameter is less than −0.2 and increasing if it is greater than +0.2. For all stable trajectories, we also include the sign of the slope parameter for time.

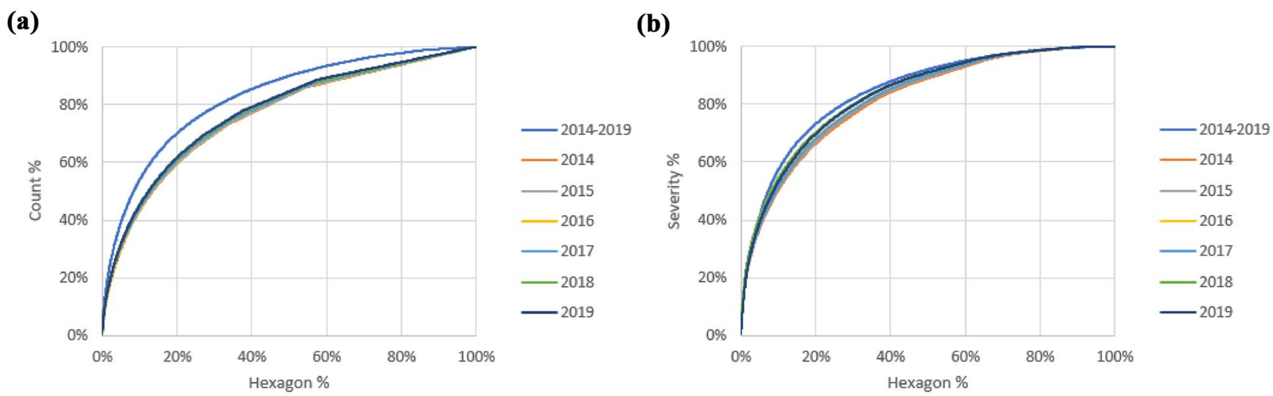

Fig. 6 **a** Power law distribution of hexagons by crime count (year). **b** Power law distribution of hexagons by crime severity (year)

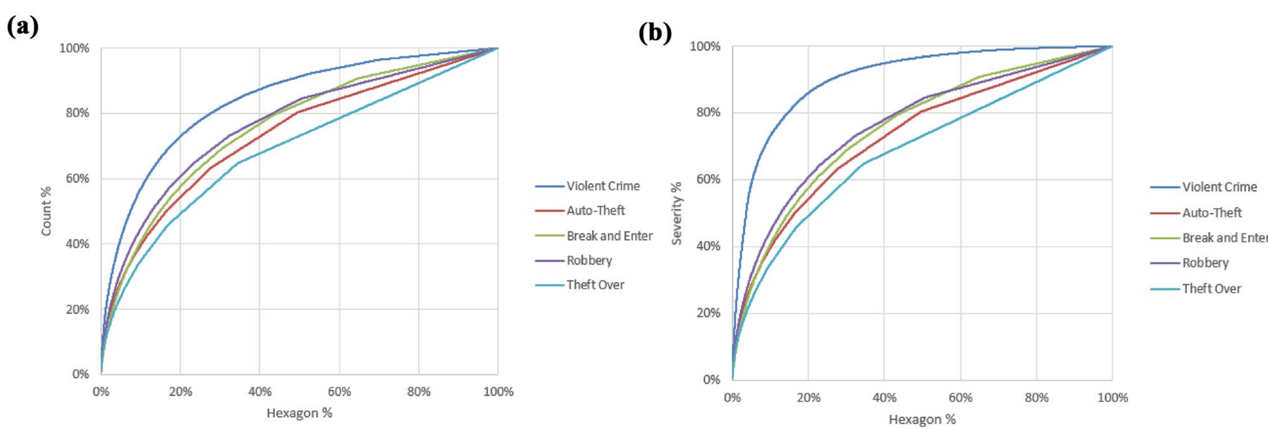

Fig. 7 **a** Power law distribution of hexagons by crime count (MCI). **b** Power law distribution of hexagons by crime severity (MCI)

Table 1 Comparison of forecasting value by model type t

Year	Counted crimes global moran's I (Z-score)	Crime severity global moran's I (Z-score)
2014	21.18	14.94
2015	20.75	13.76
2016	24.77	17.44
2017	24.27	19.51
2018	22.65	17.65
2019	29.69	25.00

Based on the Calinski criterion score, we identified an optimal k-means partition of five groups for all crime count models. In each of the six models, the trajectory contained 90.5–98.8% of all hexagons within the subgroup. More importantly, these trajectories were stable, indicating that the majority of hexagons covering Toronto consistently had little to no crime over time.

The MCI-specific crime count models are not entirely dissimilar from the aggregate count model. In each model, the first two to three trajectories accounted for 99% of all hexagons and were stable. As such, hexagons with little to no violent crime, auto-thefts, burglaries, robberies, or thefts in 2014 remained this way in the following five years.

Curiously, trajectories 2, 3, 4 and 5 in the total count model exhibited increasing trends but accounted for a small percentage of the total number of hexagons. These trajectories encapsulated moderate to high crime count hexagons. A similar pattern was observed by Weisburd et al. (2004) who noted that only 14% of street segments demonstrated a decrease overtime. This reflects a localized crime drop. This pattern is also present among the MCI-specific crime count models: hexagons with high numbers of violent crime, auto-thefts, burglaries, robberies, and/or thefts are likely to become more harmful over time (see Fig. 8). This seemingly reflects a law of

Table 2 Summary of k-means trajectories

Crime type	Trajectory	Level	Base, 2014	Trend	% of Hexagons	Crime Type	Trajectory	Level	Base, 2014	Trend	% of Hexagons
Total count	1	Low	0.17	Stable (+)	93.60	Total Severity	1	Low	33.37	Stable (−)	95.60
	2	Low	3.60	Increasing	5.02		2	Moderate	797.03	Increasing	3.50
	3	Moderate	11.00	Increasing	1.17		3	High	2137.10	Increasing	0.60
	4	High	27.30	Increasing	0.21		4	High	938.32	Increasing	0.12
	5	High	52.90	Increasing	0.03		5	High	9331.33	Increasing	0.10
							6	High	8652.35	Increasing	0.04
Violent crime count	1	Low	0.09	Stable (+)	96.50	Violent Crime Severity	1	Low	11.40	Increasing	99.10
	2	Low	3.50	Stable (−)	2.98		2	High	1709.28	Increasing	0.61
	3	Moderate	11.36	Increasing	0.48		3	High	508.25	Increasing	0.15
	4	High	26.1	Increasing	0.07		4	High	176.38	Increasing	0.11
	5	High	50.00	Increasing	0.00		5	High	403.97	Decreasing	0.02
							6	High	16,595.73	Decreasing	0.01
Auto-theft count	1	Low	0.03	Stable (−)	95.10	Auto-Theft Severity	1	Low	1.82	Stable (+)	95.10
	2	Low	0.57	Stable (+)	4.60		2	Moderate	40.78	Increasing	4.60
	3	Moderate	3.46	Stable (+)	0.28		3	Moderate/High	245.67	Increasing	0.28
	4	Moderate/High	20.89	Increasing	0.02		4	High	1484.16	Decreasing	0.02
	5	High	11.50	Increasing	0.00		5	High	817.08	Increasing	0.00
Break and enter count	1	Low	0.036	Stable (−)	90.50	Break and Enter Severity	1	Low	7.61	Stable (−)	90.50
	2	Low	0.85	Stable (−)	7.43		2	Moderate	178.25	Decreasing	7.43
	3	Low	1.30	Increasing	1.42		3	Moderate/High	274.96	Increasing	1.42
	4	Moderate	3.97	Stable (−)	0.55		4	High	836.60	Increasing	0.55
	5	Moderate	3.43	Increasing	0.08		5	High	722.74	Increasing	0.08
Robbery count	1	Low	0.02	Stable (−)	97.60	Robbery severity	1	Low	9.77	Stable (−)	97.60
	2	Low	1.09	Stable (+)	2.09		2	High	510.87	Decreasing	2.09
	3	Moderate	3.26	Stable (+)	0.29		3	High	1516.52	Increasing	0.29
	4	High	11.5	Increasing	0.02		4	High	5355.32	Increasing	0.16
	5	High	25.00	Increasing	0.00		5	High	11,642.00	Increasing	0.00
Theft count	1	Low	0.01	Stable (−)	98.80	Theft severity	1	Low	1.52	Stable (+)	98.80
	2	Low	0.69	Stable (+)	1.11		2	Moderate	98.82	Increasing	1.13
	3	Low/moderate	1.44	Increasing	0.03		3	High	417.18	Increasing	0.02
	4	High	10.50	Increasing	0.00		4	High	1516.31	Increasing	0.00
	5	High	6.00	Increasing	0.00						

accumulated advantage: high crime micro-places continue to accrue counted crimes over time. Similar to findings from Andresen et al., (2017a, 2017b), this indicates that changes in the level of counted crimes in these specific hexagons will most likely impact the overall crime situation in Toronto, as these hexagons accounted for an outsized proportion of counted crime in the city (see also Weisburd et al., 2004). In short, changes in under 1% of hexagons drive crime trends in Toronto.

We identified an optimal k-means partition ranging from four to six groups for all crime severity models. This is dissimilar from the count models where five trajectories were found to be the optimal number based on the Calinski criteria. As it pertains to the total severity model, the first trajectory accounted for 95.6% of hexagons and was stable, indicating that low crime severity hexagons remained this way across time. Similar to the total count model, each of the other trajectories within this partition, all reflecting moderate to high crime severity hexagons, were increasing. This pattern can be observed among the models measuring the severity of auto-thefts, burglaries, robberies and thefts. To summarize, while low severity hexagons remain low across time, moderate and high severity hexagons, constituting a small percentage of all hexagons, increased over time. As with crime counts, a law of accumulated advantage is present: moderate and

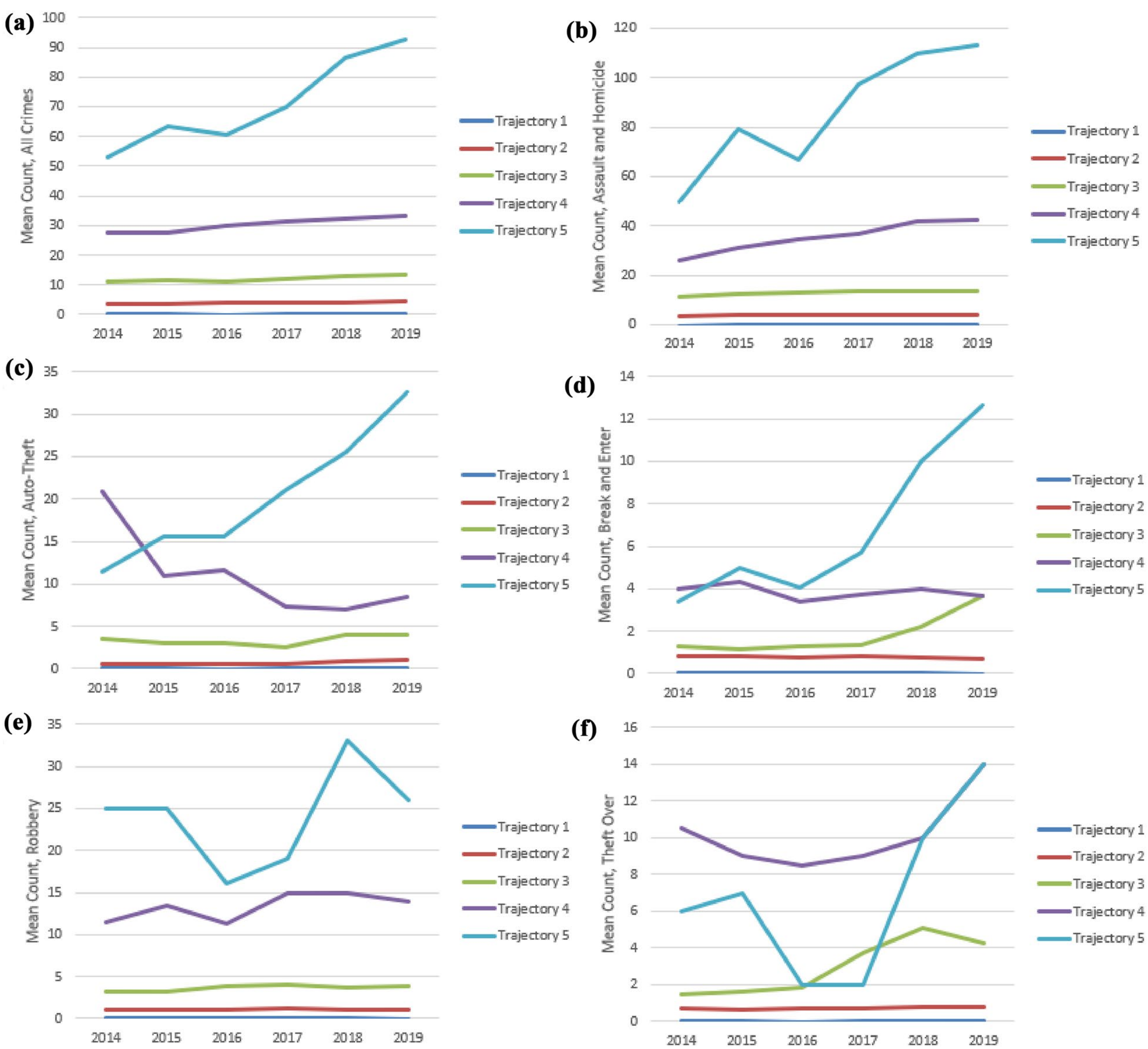

Fig. 8 K-means trajectories by crime count, 2014–2019: **a** total crime, **b** violent crime, **c** auto-theft, **d** break and enter, **e** robbery, and **f** theft over

high severity hexagons constitute an infinitesimally small percentage of all hexagons, but these hexagons translate into large overall severity in the city (see Weisburd & Amram, 2014). This pattern is not present within the violent crime severity model. The first four trajectories in this model demonstrate an increasing trend whereas the final two were decreasing. Furthermore, with the exception of the first trajectory, all others were high severity hexagons. This is likely a function of the high severity scores within the violent crime MCI. We explain further in the following paragraph.

Figures 8 and 9 plot the trajectories over time, with each line representing the result of the statistical model, showing average values. Most models exhibit a smooth trajectory with subtle year-on-year increases or decreases in the average value. However, several trajectories in both the total severity and violent crime severity models are sporadic, increasing and decreasing by large amounts—which is expected given the smaller number of events acting as outliers that pull the means to extreme scores. For example, trajectory 4 in the total severity model began with an average severity score of 938.3 in 2014, subtly increasing to 1032.7 in 2015, but ballooning to 9782.7 in 2016. Similar irregularities can be observed in trajectories 5 and 6 of the total severity model and trajectories 3, 4, 5 and 6 of the violent crime severity models.

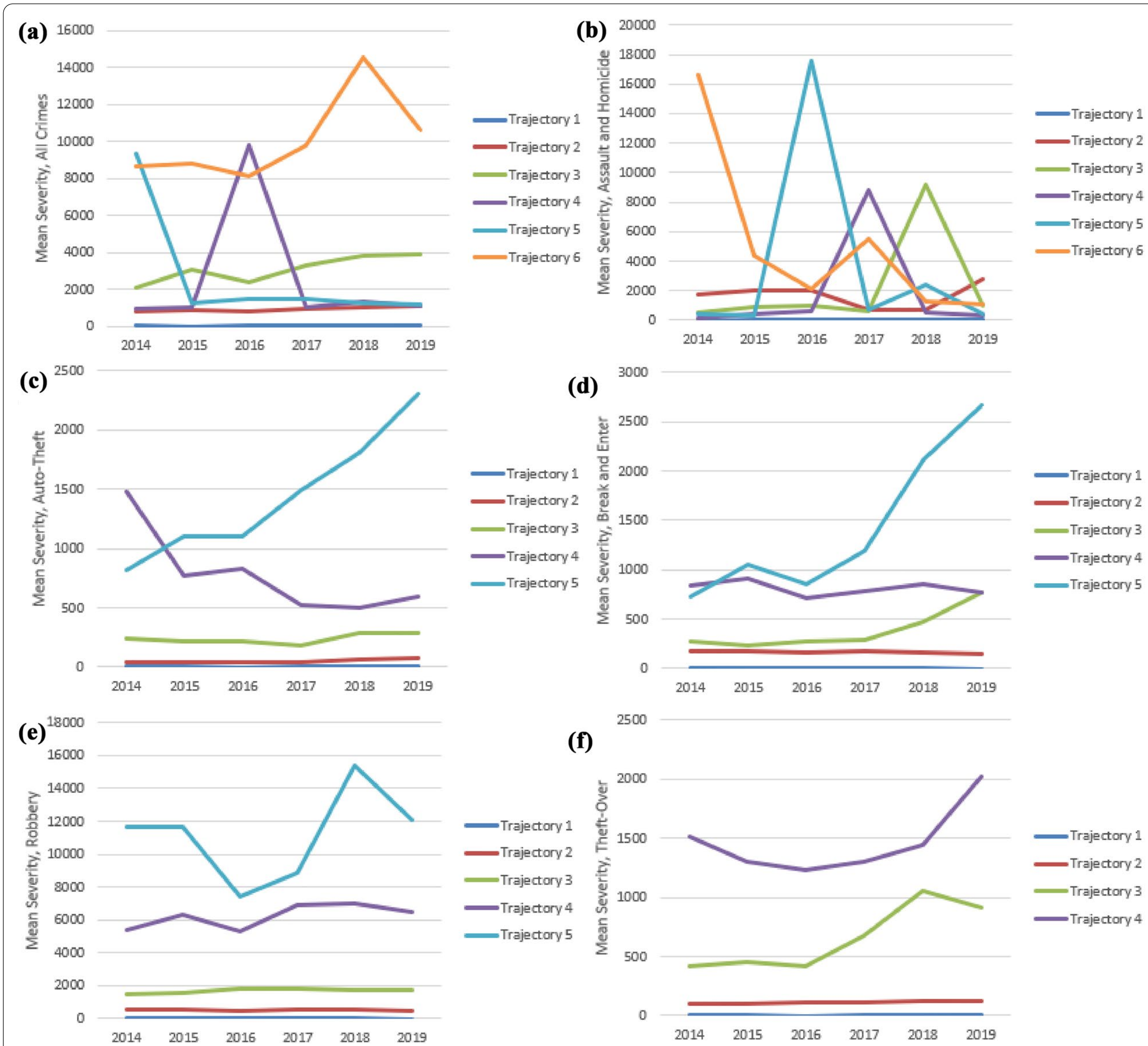

Fig. 9 K-means trajectories by crime severity, 2014–2019: **a** total severity, **b** violent crime, **c** auto-theft, **d** break and enter, **e** robbery, and **f** theft over

This spontaneity can be attributed to the lopsided nature of homicide severity scores. Indeed, first and second-degree murders have the potential to skew a trajectory, turning a hexagon with one or two counted crimes into a high severity locale. It is probable that these trajectories would have been smoother had murders been removed from the analysis.

Discussion

In this paper, we investigate the spatial concentration of crime counts and severity at the micro-level though analyses of trajectories for aggregated and disaggregated crime types in Toronto, Canada. Our analyses considered violent crime, auto-theft, burglary, robbery and theft. The implications of our results can be organized as follows: the comparability of the spatial concentration of crime count and crime severity, the longitudinal stability of crime severity at the micro-level, and the functional utility of bespoke geospatial units, like the hexagon, in crime analysis.

Based on the descriptive statistics, both crime count and severity abide by the law of crime concentration (Weisburd, 2015). However, crime severity is more spatially concentrated than counted crimes, but only slightly: 7.3% of hexagons accounted for 50% of crime severity between 2014 and 2019, and 8.4% of hexagons accounted for 50% of counted crime across the time series. Nevertheless, this difference is more pronounced when examining individual years. Whereas 13.4%, 13.4%, 13.3%, 13.1%, 12.4% and 12.4% of hexagons accounted for 50% of crime counts in 2014, 2015, 2016, 2017, 2018 and 2019, respectively, 9.9%, 9.5%, 9%, 9.4%, 8.2% and 8.7% of hexagons accounted for 50% of crime severity in the same years. These results reflect stable levels of spatial concentration across time for both crime count and crime severity, paralleling observations made by and Fenimore (2019).

As it pertains to disaggregated crime types, crime count and crime severity demonstrate equal levels of spatial concentration across auto-theft, burglary, robbery and theft. This is not altogether surprising as crime severity is a weighted multiple of the crime count. As such, the cumulative distribution function used to calculate these distributions produced similar results for crime count and crime severity when there are a large number of incidents, even when each incident weighs relatively little. In short, the overall distribution of a harm-based model and the overall distribution of a count-based model follow "power few" functions.

However, this was not the case for violent crime, as this crime category contained a myriad of crime types with varying yet high crime severity weights. First, the crime severity of violent crimes is highly spatially concentrated, with 3.7% of hexagons accounting for 50% of crime severity. In comparison, 7.4% of hexagons accounted for 50% of counted violent crimes. This suggests that the spatial dispersion of violence in society is larger than other crime categories, with a more diffused nature of against-body crimes than property crimes. This conclusion is also apparent from the Global Moran's I results, which demonstrate that crime severity, while more spatially concentrated than counted crime at the micro-level, was more dispersed across the city. This is similar to findings made by Fenimore (2019) and lays bare significant practical implications as it suggests that a strict focus on crime counts will draw resources to specific meso-geographic locales within a city. In contrast, a focus on crime severity will draw resources to multiple and discrete micro-locales where crime is most serious. To this extent, counted crimes obscure micro-level crime trends that portend to the cost-effective allocation of resources. Crime severity measures represent a more accurate means by which micro-level violent hot spots can be identified.

Based on the results of the trajectory analysis, counted crime and crime severity demonstrate for the most part similar patterns of crime concentrations at the micro-level over time. Indeed, hexagons with little counted crime or crime severity remained as such across time, whereas the small number of hexagons with moderate to high amounts of crime became increasingly more harmful. The pattern remains the same across disaggregated crime types: this reflects a law of accumulated advantage where high crime count and severity micro-places continue to accrue crimes over time. Similar to findings from Andresen et al., (2017a, 2017b), this indicates that changes in the level of counted crimes in these specific hexagons will most likely impact the overall crime situation in Toronto, as these hexagons accounted for an outsized proportion of counted crime in the city. In short, changes in under 1% of hexagons will drive crime trends in Toronto.

Finally, whereas most models exhibited a smooth trajectory with subtle year-on-year changes in the average value, several trajectories in both the total severity and violent crime severity models were sporadic. This was attributed to homicides, as a hexagon with one or two counted homicides will be designated a high severity locale. This is vitally important for resource allocation: allocation of resources cannot be based solely on a harm index, but must take into account crime counts as well. Indeed, while crime harm measures allow us to identify the weight of crime in a geographic locale, counts provide information on the frequency of crime therein, and in particularly violent, high-harm hot spots, this becomes more pertinent. Harm measures naturally lend themselves to the inclusion of statistical outliers which skew perceptions of where crime harm is located. Crime counts allow us to account for these outliers and therefore, while all crimes are not created equal, discounting the count-based model altogether is inefficient, at least in the context of place-based criminology.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40163-022-00176-x.

Additional file 1. Counted Crime and Crime Severity Scores by Crime Type in Toronto, Canada (2014-2019).

Acknowledgements
None.

Author contributions
Dr VH—report drafting; statistical analyses. Mr ZB—data curation; statistical analyses; report drafting. Dr BA—report drafting; review of analysis. All authors read and approved the final manuscript.

Funding
None.

Data availability
Data are available upon request from the corresponding author.

Declarations

Ethics approval and consent to participate
No ethics approval was required for this research, which is based solely on publicly available, non-personal data routinely collected by law enforcement agencies.

Competing interests
We declare no potential competing interests.

Author details
[1]Institute of Criminology, University of Cambridge, Sidgwick Ave, Cambridge CB3 9DA, UK. [2]Calgary Police Service, Calgary, Canada. [3]Experimental Criminology, Institute of Criminology, University of Cambridge, Sidgwick Ave, Cambridge CB3 9DA, UK. [4]Criminology, Institute of Criminology, The Hebrew University of Jerusalem Mt. Scopus, 9190501 Jerusalem, Israel.

Received: 12 March 2022 Accepted: 7 November 2022
Published online: 29 November 2022

References

Andersen, H. A., & Mueller-Johnson, K. (2018). The Danish Crime Harm Index: How it works and why it matters. *Cambridge Journal of Evidence-Based Policing, 2*(1), 52–69.

Andresen, M. A. (2007). Location quotients, ambient populations, and the spatial analysis of crime in Vancouver, Canada. *Environment and Planning A, 39*(10), 2423–2444.

Andresen, M. A., Curman, A. S., & Linning, S. J. (2017a). The trajectories of crime at places: Understanding the patterns of disaggregated crime types. *Journal of Quantitative Criminology, 33*, 427–449.

Andresen, M. A., Linning, S. J., & Malleson, N. J. (2017b). Crime at places and spatial concentrations: exploring the spatial stability of property crime in Vancouver BC, 2003–2013. *Journal of Quantitative Criminology, 33*, 255–275.

Ashby, M. P. (2018). Comparing methods for measuring crime harm/severity. *Policing: A Journal of Policy and Practice, 12*(4), 439–454.

Bernasco, W., & Steenbeek, W. (2017). More places than crimes: Implications for evaluating the law of crime concentration at place. *Journal of Quantitative Criminology, 33*, 451–467.

Bland, M. P., & Ariel, B. (2020). Measuring harm. *Targeting domestic abuse with police data* (pp. 63–82). Springer.

Brand, S. & Price, R. (2000). *Home office research study 217: The economic and social costs of crime*, London

Calinski, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics, 3*(1), 1–27.

Carter, J. G., Mohler, G., Raje, R., Chowdhury, N., & Pandey, S. (2021). The Indianapolis harmspot policing experiment. *Journal of Criminal Justice, 74*, 101814.

Cohen, M., & Bowles, R. (2010). Estimating costs of crime. In A. Piquero & D. Weisburd (Eds.), *Handbook of quantitative criminology* (pp. 143–162). Springer.

Curman, A. S. N., Andresen, M. A., & Brantingham, P. J. (2014). Crime and place: A longitudinal examination of street segment patterns in Vancouver, BC. *Journal of Quantitative Criminology, 31*, 127–147.

Eck, J. E., Lee, Y., O, S.-H., & Martinez, N. (2017). Compared to what? Estimating the relative concentration of crime at places using systematic and other reviews. *Crime Science, 6*(1), 1–17.

Fenimore, D. (2019). Mapping harmspots: An exploration of the spatial distribution of crime harm. *Applied Geography, 109*, 1–10.

Frydensberg, C., Ariel, B., & Bland, M. (2019). Targeting the most harmful co-offenders in Denmark: A social network analysis approach. *Cambridge Journal of Evidence-Based Policing, 3*(1), 21–36.

Genolini, C., & Falissard, B. (2010). KmL: k-means for longitudinal data. *Computational Statistics, 25*(2), 317–328.

Greenfield, V., & Paoli, L. (2013). A framework to assess the harms of crimes. *British Journal of Criminology, 53*(5), 864–885.

Groff, E., Weisburd, D., & Yang, S. (2010). Is it important to examine crime trends at a local "micro" level? A longitudinal analysis of street to street variability in crime trajectories. *J Quant Criminology, 26*(1), 7–32.

Hall, J. (1960). *General principles of criminal law (second)*. The Lawbook Exchange.

Hiltz, N., Bland, M., & Barnes, G. C. (2020). Victim-offender overlap in violent crime: Targeting crime harm in a Canadian suburb. *Cambridge Journal of Evidence-Based Policing, 4*(3), 114–124.

House, P. D., & Neyroud, P. W. (2018). Developing a crime harm index for Western Australia: The WACHI. *Cambridge Journal of Evidence-Based Policing, 2*(1), 70–94.

Lee, Y., Eck, J. E., O, S.-H., & Martinez, N. (2017). How concentrated is crime at places? A systematic review from 1970 to 2015. *Crime Science, 6*(1), 1–16.

Li, H., Calder, C., & Cressie, N. (2007). Beyond Moran's I: Testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis, 39*(4), 357–375.

Link, E., & Lösel, F. (2022). Development and application of an offense severity index in the evaluation of treatment of individuals convicted of sexual crimes. *Psychology, Crime & Law,*. https://doi.org/10.1080/1068316X.2022.2032054

Macbeth, E., & Ariel, B. (2019). Place-based statistical versus clinical predictions of crime hot spots and harm locations in Northern Ireland. *Justice Quarterly, 36*(1), 93–126.

Massey, J., Sherman, L. W., & Coupe, T. (2019). Forecasting knife homicide risk from prior knife assaults in 4835 local areas of London, 2016–2018. *Cambridge Journal of Evidence-Based Policing, 3*(1), 1–20.

Mitchell, R. J. (2019). The usefulness of a crime harm index: Analyzing the Sacramento Hot Spot Experiment using the California Crime Harm Index (CA-CHI). *Journal of Experimental Criminology, 15*(1), 103–113.

Nagin, D., & Land, K. (1993). Age, criminal careers and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology, 31*(3), 327–362.

Ransley, J., Murphy, K., Karstedt, S., Bartlett, D., Forrester, L., & Carless, M. (2018). Developing and applying a Queensland crime harm index—Implications for policing serious and organised crime. *Organised Crime Research in Australia, 1*, 105–114.

Ratcliffe, J. (2015). Towards an index for harm-focused policing. *Policing, 9*(2), 164–182.

Sarnecki, J. (2021). Comment on Kärrholm et al. "Designing the Swedish Crime Harm Index: an Evidence-Based Strategy." *Cambridge Journal of Evidence-Based Policing, 5*(1), 76–90.

Sherman, L. (2013). The rise of evidence-based Policing: Targeting, testing, and tracking. *Crime and Justice, 42*(1), 377–451.
Sherman, L. W., Gartin, P., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology, 27*(1), 27–55.
Sherman, L., Neyroud, P., & Neyroud, E. (2016). The Cambridge Crime Harm Index: Measuring total harm from crime based on sentencing guidelines. *Policing*. https://doi.org/10.1093/police/paw003
Šimon, M., & Jíchová, J. (2020). Crime count and crime harm in a post-socialist city: How does the law of crime concentration at places apply? *European Journal of Criminology*. https://doi.org/10.1177/1477370820973993
Statistics Canada (2020). Police-reported crime statistics in Canada, 2020. Retrieved 14/02/2022 from https://www150.statcan.gc.ca/n1/pub/85-002-x/2021001/article/00013-eng.htm
Wallace, M., Turner, J., Matarazzo, A., & Babyak, C. (2009). *Measuring crime in Canada: Introducing the crime severity index and improvements to the uniform crime reporting survey*. Canadian Centre for Justice Studies.
Weinborn, C., Ariel, B., Sherman, L., & O'Dwyer, E. (2017). Hot spots vs. harmspots: Shifting the focus from counts to harm in the criminology of place. *Applied Geography, 86*, 226–244.
Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology, 53*(2), 133–157.
Weisburd, D., & Amram, S. (2014). The law of concentrations of crime at place: The case of Tel Aviv-Jaffa. *Police Practice & Research, 15*, 101–114.
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*, 283–321.
Weisburd, D., Groff, E. R., & Yang, S. M. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Oxford University Press.
Weisburd, D., & Mazerolle, L. G. (2000). Crime and disorder in drug hot spots: Implications for theory and practice in policing. *Police Quarterly, 3*(3), 331–349.
Wheeler, A., Worden, R., & McLean, S. (2015). Replicating group-based trajectory models of crime at micro-places in Albany, NY. *Journal of Quantitative Criminology, 32*, 589–612.
Zhang, Y., Li, A., & Fung, T. (2012). Using GIS and multi-criteria decision analysis for conflict resolution in land use planning. *Procedia Environmental Sciences, 13*, 2264–2273.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.