ISS INSTITUTE FOR SECURITY STUDIES

Point-level analysis of crime in Khayelitsha

A critical review

Ian Edelstein and Robert Arnott

This report analyses point-level data on murders and robberies for a high-crime community in South Africa to discern spatial trends. It explores the data at a granular level to understand the nature of several hotspots. Issues with the quality of police data are discussed, both as limitations for analysis, and in terms of the need for more accurate and readily available point-level data for ongoing crime analysis and response.

Key findings

- From 2006–2016, crimes across Greater Khayelitsha are dispersed over place and time. They are however concentrated near intersections, shopping centres, medical facilities, train stations and police stations. This may be the result of police geolocating crimes at landmarks, rather than at their actual locations (when such geolocations are not provided).
- In the 2010–2015 period, many crimes are geolocated in perfect grid patterns. This gridded data affects as much as 10% of all crime data in the 10-year period, interfering with accurate spatial-temporal analysis that could improve intelligent policing.
- Such imprecision may result in areas being misidentified as crime hotspots, while actual crime hotspots may not be identified at all. This precludes deeper analysis that could inform more strategic crime interventions and policing.
- Without effective data creation and collection methods, little can be done to reduce violence through data-driven, evidence-based methods.
- South African Police Service (SAPS) data may be too unreliable to be effectively used for hotspot analysis or the development of an effective violence prevention strategy.

Recommendations

- SAPS should publish details of how incident-level crime data is collected, described and geocoded (including devices and methods used), clearly indicating the limitations of each record. Partial, incomplete data is far better than incorrect data which prevents adequate analyses. This information will allow analysts to determine the precision with which each piece of data can be analysed. In addition, if victim, perpetrator and anonymised incident details are attached to each record, it can substantially increase the research potential.
- Crime data should indicate when a single crime may fall into multiple categories. For example, a house robbery with a firearm may fall into two categories (house robbery and robbery with a firearm). Both categories should be reported and it should be clearly indicated that only a single crime occurred.
- Adequate financial, technical and human resources are required at station level to provide the ongoing capacity to collect, clean, verify and provide timeous crime data to inform policing and other crime prevention efforts.
- This data should be shared, timeously and responsibly, with the National Safety, Crime and Violence Prevention Centre (as proposed in the 2016 White Paper on Safety and Security), as well as external researchers. This will allow for much deeper crime data analysis than current practice allows and will help to improve analysis and response.

Introduction

The spatial analysis of crime and its correlates dates back to 19th century Europe. In France, property crimes and violent crimes were mapped to the district level, demonstrating both spatial variation between the two types of crime, as well as seasonal fluctuation.[1] In England, Mayhew showed a high concentration of criminal offenders in the band separating the City of London (one square mile with a separate police service) from Greater London.[2] The jurisdictional boundaries of the two police services allegedly provided offenders with more opportunities to avoid apprehension.

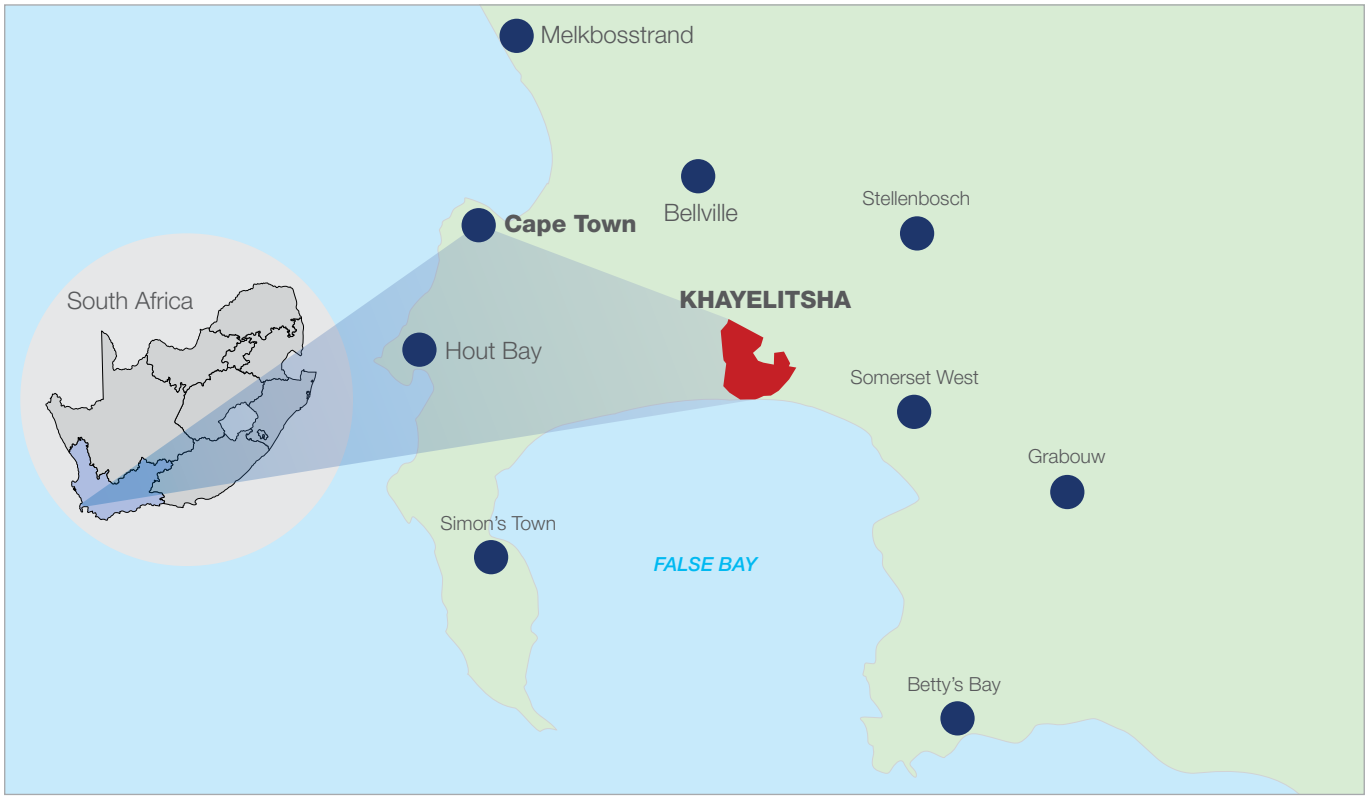
The field of spatial crime analysis was advanced in the 20th century with studies exploring the nature of urban growth in American cities (largely outward growth in concentric bands), identifying the businesses, industries and residents drawn to each zone, and the nature of crime that would follow.[3] American cities grew more rapidly during industrialisation than did the older cities of Europe.

Geographic information systems (GIS),[4] dating back to the 1980s, enabled an explosion of spatial and temporal crime research in the ensuing decades, often demonstrating the existence of crime patterns (e.g. house robberies by day when residents are at work, business robberies at night when shops are closed). Several key theories seek to understand such crime patterns.

The routine activity theory[5] suggests that a would-be offender, an appropriate target and the absence of a guardian or protector of the target are the three key components that enable opportunities for crime. Crime prevention efforts may seek to do the following: alter the behaviours of would-be offenders (providing alternative opportunities and increasing the penalties if caught); make the targets less vulnerable or obvious (e.g. self-defence training, high walls, electric fences); and increase both active and passive guardianship (e.g. CCTV cameras, visible policing, community watches, private security guards, police raids, random searches).

Rational choice theory[6] holds that many offenders make conscious, rational decisions to commit crimes based on the potential rewards and the likelihood of getting caught. In contexts of high unemployment and limited policing, such as South Africa, higher crime rates are hardly a surprise. However, not all offenders and offences are

Figure 1: Location of Khayelitsha in South Africa



rational. Both strain theory[7] and bounded rationality[8] suggest, respectively, that the pressures to accumulate wealth in a highly unequal society and the effects of alcohol, drugs, inadequate (and unequal) education and pressure from peers or family may induce crimes of a less rational nature.

General, or collective strain theory,[9] where a concentration of individuals, strained by relative deprivation, live in close proximity, alongside poorly protected targets (e.g. people carrying cellphones and salary packets, shacks that are easily forced open, informal areas that are poorly monitored) would seem to explain why overcrowded South African townships and adjacent informal settlements, with many lower-income residents, are hotbeds of crime.

Crime pattern theory[10] suggests that all people move between places (e.g. home, school, work, shopping, entertainment) on a regular basis, becoming familiar with these places and the routes they normally travel. Offenders look for opportunities along their own routes, as well as in select target areas, that they can become familiar with (e.g. shopping centres, homes in affluent neighbourhoods). Thus, within Khayelitsha, you may expect to find crime patterns that follow regular movements and opportunities (e.g. transportation hubs during crowded periods, shopping centres after payday, points of alcohol consumption in the evenings and during weekends, main intersections, walking routes to schools and taxi ranks, businesses during evenings and periods of closure).

Spatial crime analysis has a strong visual component that can both tell a story and, potentially, mislead. Readers are always drawn to darker (or, if in colour, redder) areas on a map, assuming they depict higher levels of crime. In a map of exclusively low-crime areas, even a single crime may give the impression of a concentrated risk. Likewise, the areal unit selected for aggregation can give the impression that the entire area reflects equally high-crime levels, when in fact, it may be much smaller hotspots within the aggregated units that contain most of the crime. This is known as the modifiable areal unit problem.[11] It is especially the case for low-income neighbourhoods leading to the (false) perception that poor areas and poor people are uniformly dangerous.

> Spatial analysis and mapping can show crime concentration and trends but only if the data is largely accurate

SPATIAL ANALYSIS IS ACCESSIBLE AND VISUAL

Generally speaking, the real value of spatial analysis is found in its accessibility to lay audiences and visual storytelling. Academic and statistically rigorous approaches to spatial analysis have corrected some of the problems with coarse generalisations and uncertain data but have also been criticised for providing little actionable evidence to address the very problems under study.[12]

This brief follows an approach intended to be straightforward, visually intuitive, as well as empirically grounded. Thus, spatial analysis and mapping have the potential to reveal truths in terms of crime concentration and trends but only if the underlying data is largely accurate. This in turn enables more intelligent

responses to crime. However spatial analysis and mapping can also mislead if the data is inaccurate and the levels of analysis are too coarse.

Methodology

Data

Point-level crime data for the three police stations comprising Greater Khayelitsha (Khayelitsha police station, Lingelethu West, and Harare) were obtained through a public information request to the SAPS Western Cape Division. The datasets list all reported contact crimes by police precinct, crime category, date and time stamp, and geolocation from 1 April 2006 until 31 March 2016, a 10-year period. The analyses focus on murders and the five most common robbery categories (business robbery, common robbery,[13] house robbery, robbery with a weapon or instrument other than a firearm, and robbery with firearm).

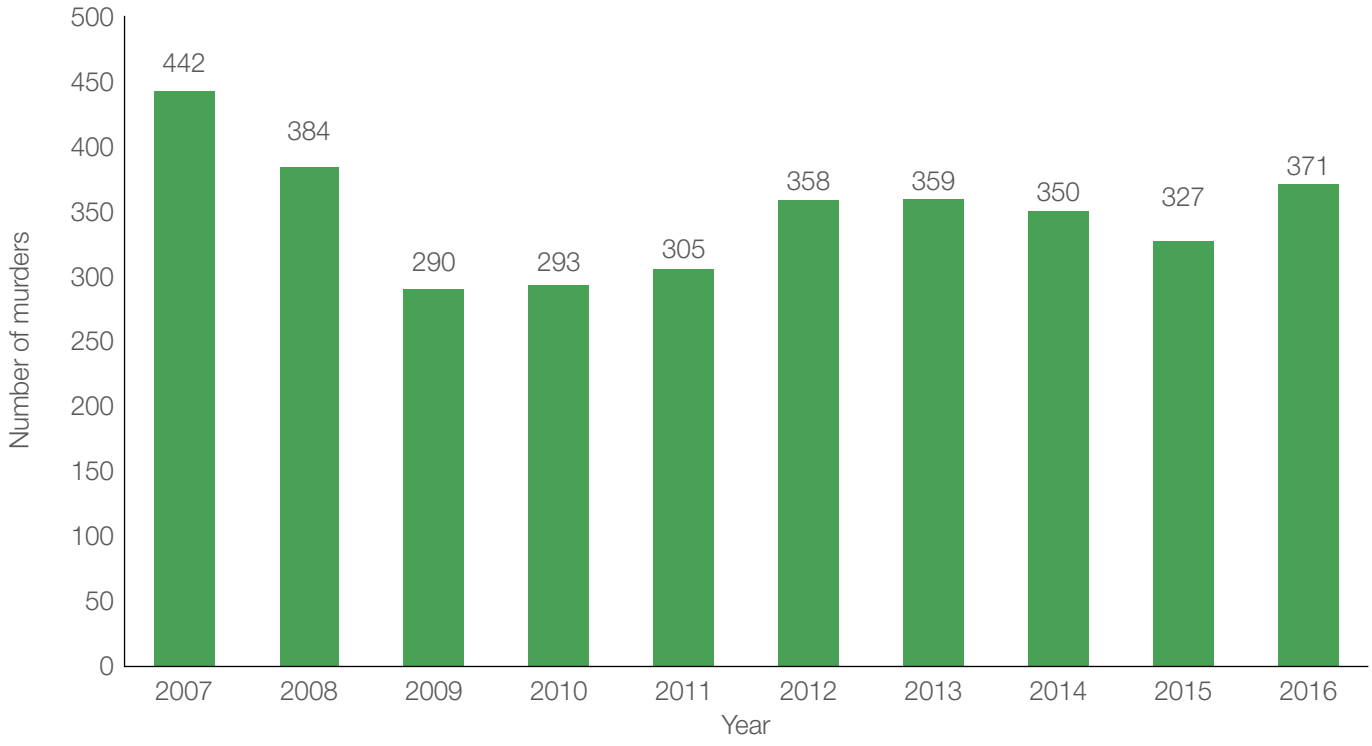
Murders are assumed to be a more complete dataset – in most cases a body is proof of the incident and police investigative efforts may be more thorough in specifying time and location. Robberies are also included in the analyses as a category of crimes that may be more preventable through predictive policing[14] and social and infrastructural interventions. These interventions will increase passive surveillance (e.g. eyes on the streets) and improve the social fabric (e.g. reduce the broken windows theory that neglected and unsurveilled spaces encourage anti-social and criminal behaviour).

In total, over the 10-year period (1 April 2006 to 31 March 2016), there were 3 479 murder cases in Greater Khayelitsha: 1 531 murders in the Khayelitsha police station precinct, 462 in Lingelethu West and 1 486 in Harare.

For all combined robberies in the 10-year period (18 550 in total), there were 7 719 in the Khayelitsha police station precinct, 3 466 in Lingelethu West and 7 365 in Harare. The overall breakdown is as follows: 9.8% (1 823 cases) business robbery; 24.2% (4 497 cases) common robbery; 11.2% (2 072 cases) house robbery; 13.7% (2 541 cases) robbery with a weapon or instrument other than a firearm; and 41.1% (7 617 cases) robbery with a firearm.

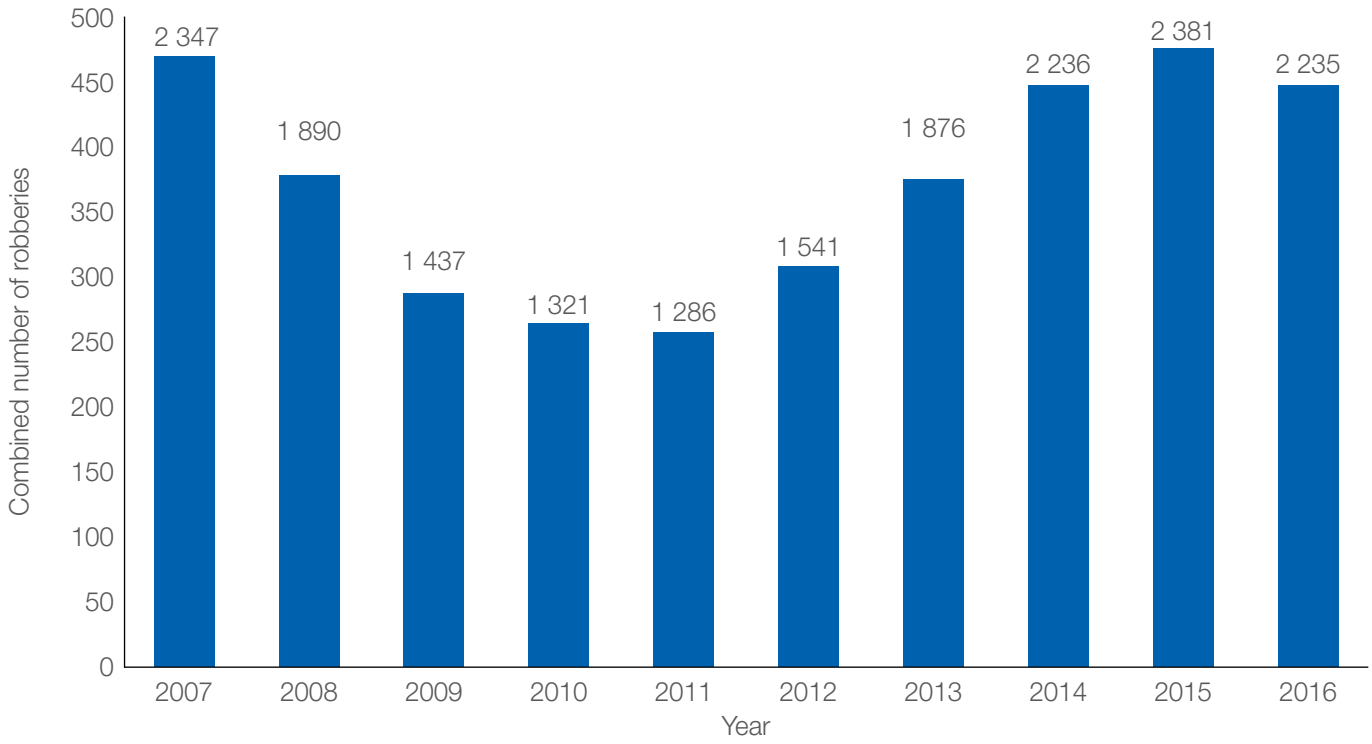
The aggregated year-by-year levels are shown in Figure 2 and 3. It is notable that Khayelitsha police station does not report any cases in the 10-year period under the category robbery with a weapon or instrument other than a firearm.

Figure 2: Murders by year (1 April–31 March of the following year) for Greater Khayelitsha



Data source: South African Police Service

Figure 3: Combined robberies by year (1 April–31 March of the following year) for Greater Khayelitsha



Data source: South African Police Service

Analysis methods

This report focuses primarily on the spatial and visual depiction of crime concentrations and it attempts to understand the underlying factors that may account for these concentrations over the 10-year period. As such, detailed temporal analysis is beyond the scope of this report but could be undertaken in future research. After identifying hotspots using various methods, an inspection of several of these hotspots is conducted to understand what is happening, both in the data and, through satellite imagery, on the ground in terms of the built environment. Issues with the way crime data has been geocoded become evident and these are visualised and discussed.

The data is visualised at various stages using dots to represent the crime incidents. Colour coding is used to delineate the crime category, police station or other variables of interest. This is done using open-source GIS mapping software, QGIS,[15] Crime Stat IV,[16] as well as custom-built spreadsheets.

Three methods are used to identify and examine hotspots. The first method breaks the entire Khayelitsha area down into 267 m square grids[17] and all crimes are binned within each grid cell by category (murders and combined robberies). The 267 m dimension is internationally accepted as the length of a city block. Conveniently 3 x 3 grids also equate to roughly 800 m x 800 m squares which can be thought of as larger hot regions. The grid boxes contain the total number of crimes occurring in each one over the 10-year period.

This visualisation is then laid over hybrid satellite imagery street maps to provide a sense of the built environment beneath. Lastly, points of interest (or possible crime association) are geolocated and marked on the maps: the three police stations, all public schools, transportation hubs (taxi ranks and train stations) and all public health facilities (clinics and hospitals).

Although aggregating crimes into Stats SA small area layers (the smallest unit of spatial-demographic aggregation in the SA Census) was explored, it was found that the varying sizes and shapes were problematic for hotspot analysis.[18] Such a technique would however allow for some degree of population-based standardisation of crime rates.

The second method for identifying hotspots uses the robbery locations themselves as analysis centre points, using the fuzzy mode (F-mode) analysis method.[19] In this method, the number of robberies within a 50 m radius of each analysis point (robbery location) is added up. Any analysis point with fewer than 10 robberies within this radius is removed. Plotting the remaining points gives a good indication of areas with very high concentrations of robberies (for example, busy street corners or train stations).

This procedure is followed for each year of analysis data, with varying parameters. This method (as well as the third method described below) is conducted using only robbery data as murders are not frequent enough to generate meaningful hotspots on their own, nor to influence the robbery data concentrations.

The third method employs nearest neighbour hierarchical spatial clustering (Nnh).[20] This is a technique that identifies groups of points, based on the distances between points, as well as the total number of points, within that group. It is useful for identifying crime hotspots that are perhaps not as dense as those identified using the F-mode method. Examples would be hotspots that span a large shopping complex, a portion of a suburb, or a stretch of road.

In this study, the clustering was performed with varying parameters – one example was a fixed maximum distance of 200 m and a fixed minimum of 30 points per group. Ostensibly, the higher the threshold number of crimes set, the fewer hotspots will emerge.

In the first step of the Nnh analysis, groups of 30 points (robberies) that are all within 200 m of each other are identified and formed into groups, representing robbery hotspots. These are called the first order hotspots.

The method is hierarchical so the next step determines the centres of each of these first order hotspots, and groups any hotspots that have centres that are within 200 m of each other, forming larger hotspots. These are called the second order hotspots.

The process is then repeated with these larger hotspots and with the next group of larger hotspots, until no further hotspots can be grouped together. Finally, convex hulls are drawn around all the points that form each of the hotspots. A convex hull represents the area that encloses all the points that form a hotspot, as if an elastic band were stretched around all of the outermost points. This Nnh analysis is conducted for the 10-year period to explore enduring hotspots.

Finally, after establishing the hotspots with the aforementioned techniques, several hotspots are explored in greater detail to examine their spatial aspects. Applied across a police precinct, and in conjunction with temporal analysis, this form of deep hotspot-analysis would potentially allow for predictive policing[21] and other localised crime prevention strategies to be targeted.

Analysis and visualisations

Global visualisation: all crimes from 2006–2016

In the first spatial analysis image (Figure 4), all murder and robbery crime points are shown and colour coded for the respective police stations: Khayelitsha police station in green, Lingelethu West in red, Harare in blue.

At this level of aggregation, there appears to be both a saturation of crime events across the Khayelitsha police station precinct (in green), as well as some spatial gaps in areas that include built environment but have no crimes reported – surrounded by areas with crime events (Lingelethu West in red, Harare in blue). This is especially the case for the southern regions of Harare, despite the reporting of Harare crimes further south, near the Monwabisi Resort on the False Bay coast. It also appears that a small section in the middle of the two (green) triangular regions that comprise the Khayelitsha police station precinct is occupied by Lingelethu West reported crimes. This may speak to a jurisdictional challenge.

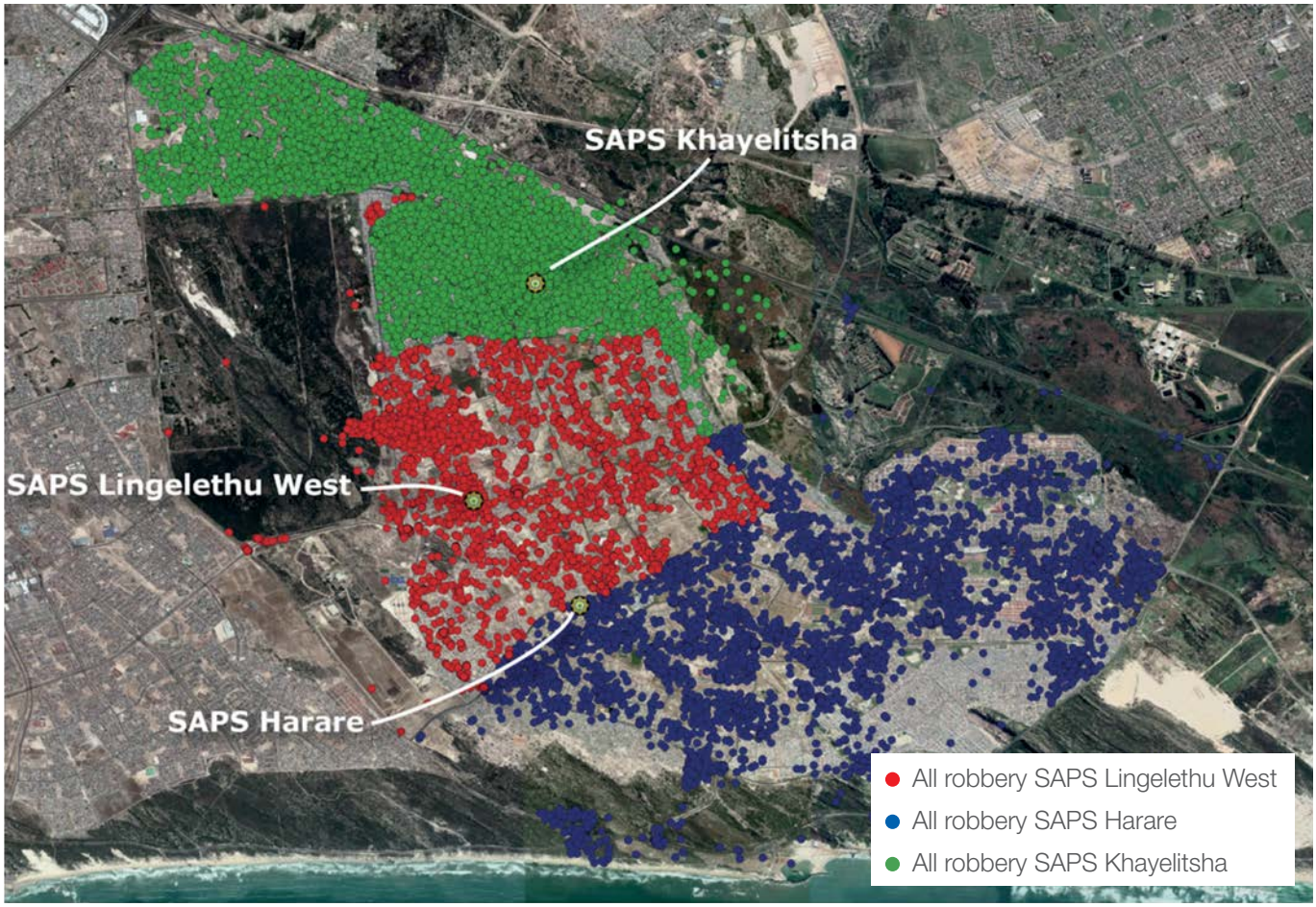
Trends by station

Looking at murders and combined robberies by year and by station (Figure 5), a general reduction trend is evident from the peaks in 2007 to the lower 2009–2011 period. Increases in reported robberies are evident in all stations from 2011 to 2015, but tapered in 2016. Murder levels appear to be fairly consistent in all stations from 2011 to 2016, but rose across all stations from 2015 to 2016 while the number of robberies mostly dropped.

Grid quadrant analyses

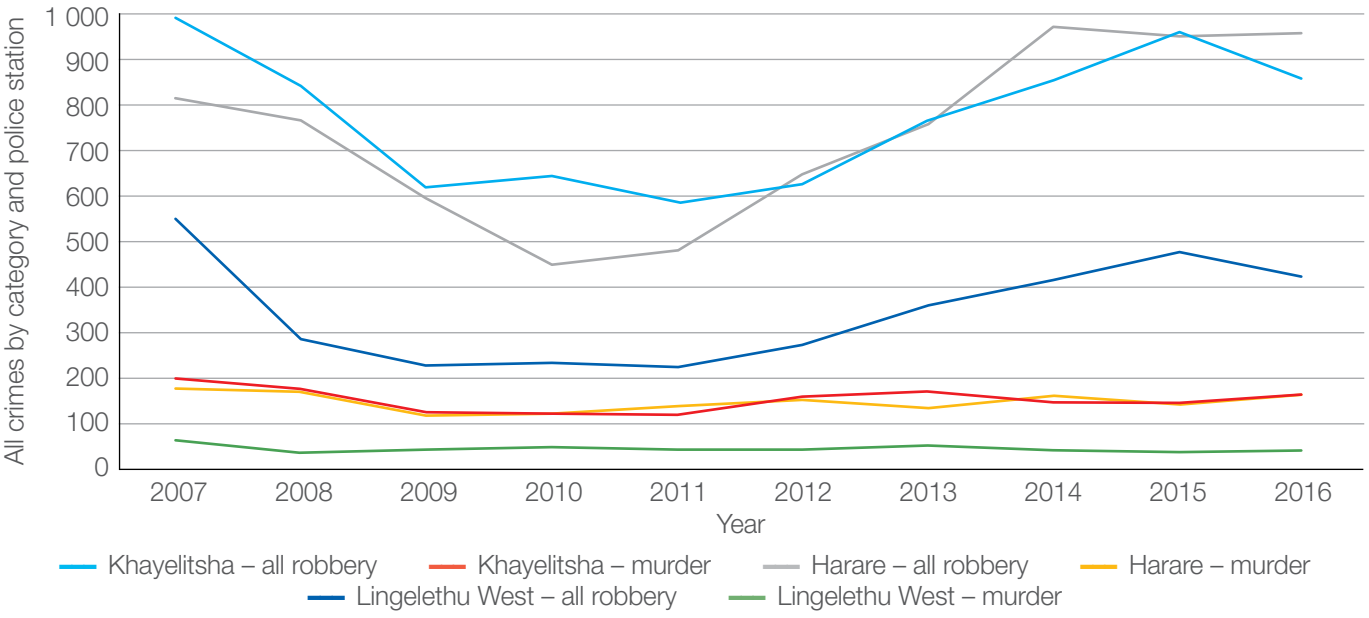
In the 267 m grids in Figure 6 murders are widely dispersed over the 10-year period. There are also concentrations of murders in individual blocks and many adjacent or contiguous areas reporting 20 or more

Figure 4: All Khayelitsha geolocated robberies over the 10-year period, colour coded by police station



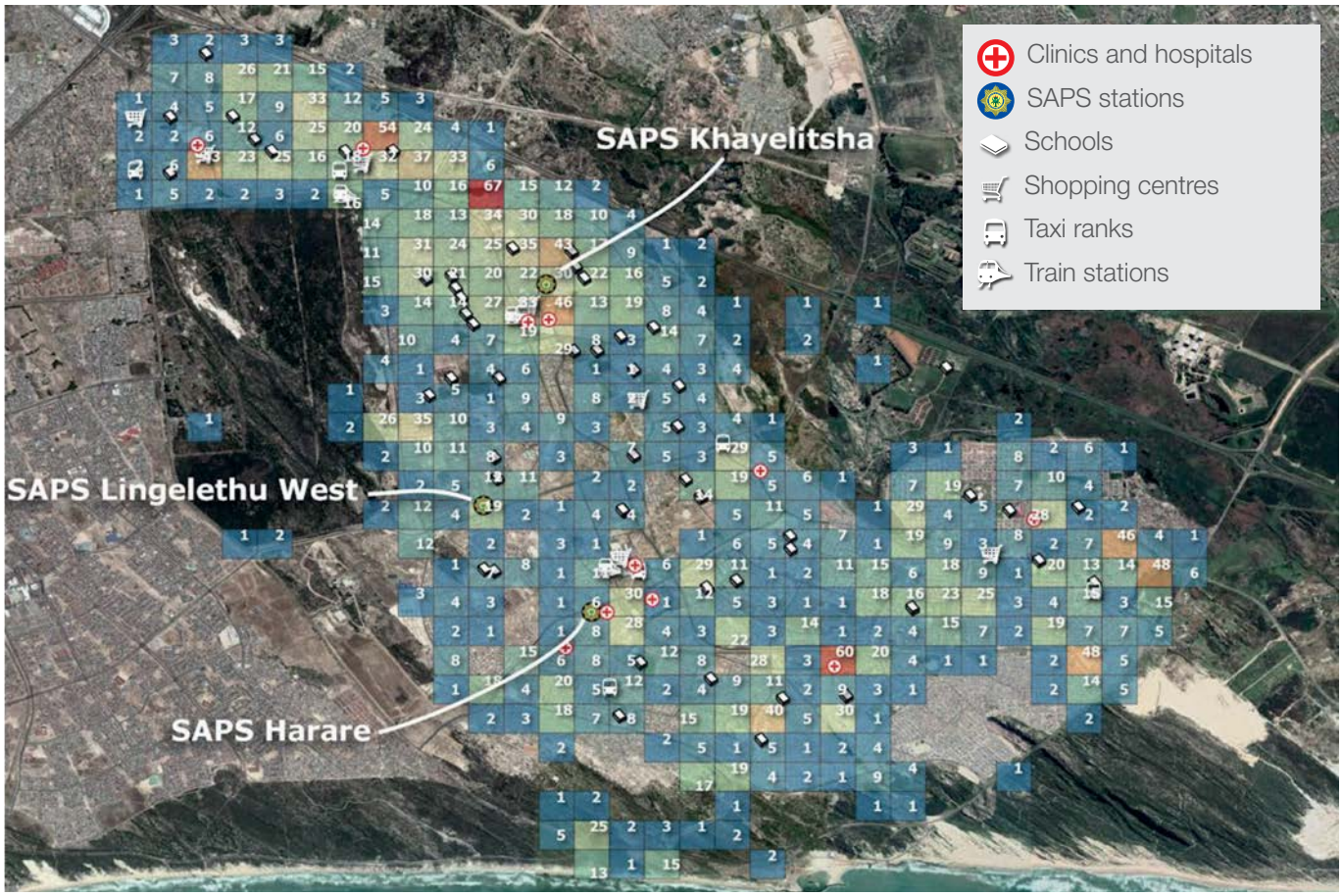
Data source: South African Police Service

Figure 5: All crimes by category and police station



Data source: South African Police Service

Figure 6: All Khayelitsha 10-year murder incidents binned into 267 m square block grids



Data source: South African Police Service

murders. Contiguous high-murder blocks are especially apparent in large parts of the Khayelitsha police station precinct.

Overall, many of the high-murder blocks appear to be centred over medical facilities, such as the Khayelitsha Day Hospital in Harare (with 60 murders), with other blocks containing or bordering medical facilities reporting 54, 43, 30, 29 and 28 murders, respectively. Furthermore, the Khayelitsha police station is itself immediately adjacent to a residential block containing 43 murders.

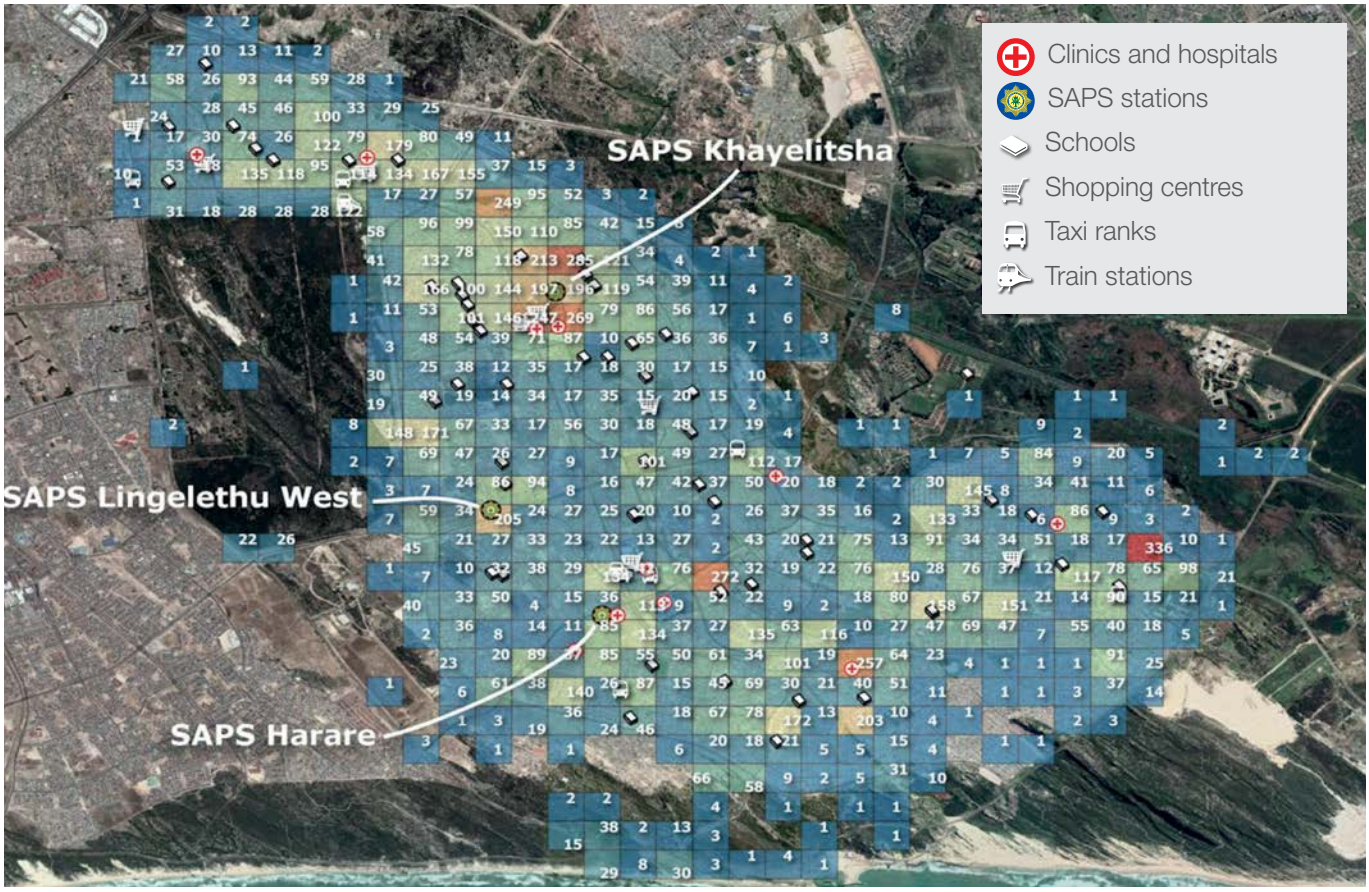
In the 10-year combined robbery incidents grids (Figure 7), several very high-robbery hotspots are apparent that are isolated and surrounded by low-robbery level blocks. This may be the result of a tendency for the SAPS to georeference crimes on specific intersections when they may have occurred nearby in a sizable informal settlement area without identifiable roads. As with the murder data, high numbers of robberies are seen in blocks containing larger medical facilities and adjacent to two of the police stations.

Just as murders with unknown locations (with bodies arriving at medical facilities) may be tagged at the medical facility or police station, aggravated robbery victims may also have their incidents referenced to the medical facility or police station when they are unable to provide more information at the time or if follow-up by the crime data analysts does not occur.

F-mode analysis

In this F-mode analysis, robberies within a 50 m radius of each analysis point (each robbery location) are totalled and any centre points with totals less than 10 robberies are removed. This technique identifies small areas with very high concentrations of robberies. Figure 8 displays visualised results of this analysis, focused on Khayelitsha police precinct and Lingelethu West.

Figure 7: All Khayelitsha 10-year robbery incidents binned into 267 m square block grids



Data source: South African Police Service

The more prominent F-mode clusters are magnified in the bubbles to examine the patterns of concentration. Small clusters have not been magnified. These analyses are conducted on two separate, yet contiguous time frames, 2007–2009 and 2010–2015. The same analysis was conducted on the Harare data, although those findings are not presented here.

As seen in the previous grid visualisation of all robberies (Figure 7), it is evident in the 2007–2009 F-mode visualisation (Figure 8) that the Khayelitsha police station data shows more concentrated crime radii and these radii appear more clustered. The majority of the F-mode hotspots in this time period tend to be located at major intersections or adjacent to a shopping centre, medical facility or school. Several appear to be clusters located in specific neighbourhoods (in both cases, adjacent to informal settlement areas).

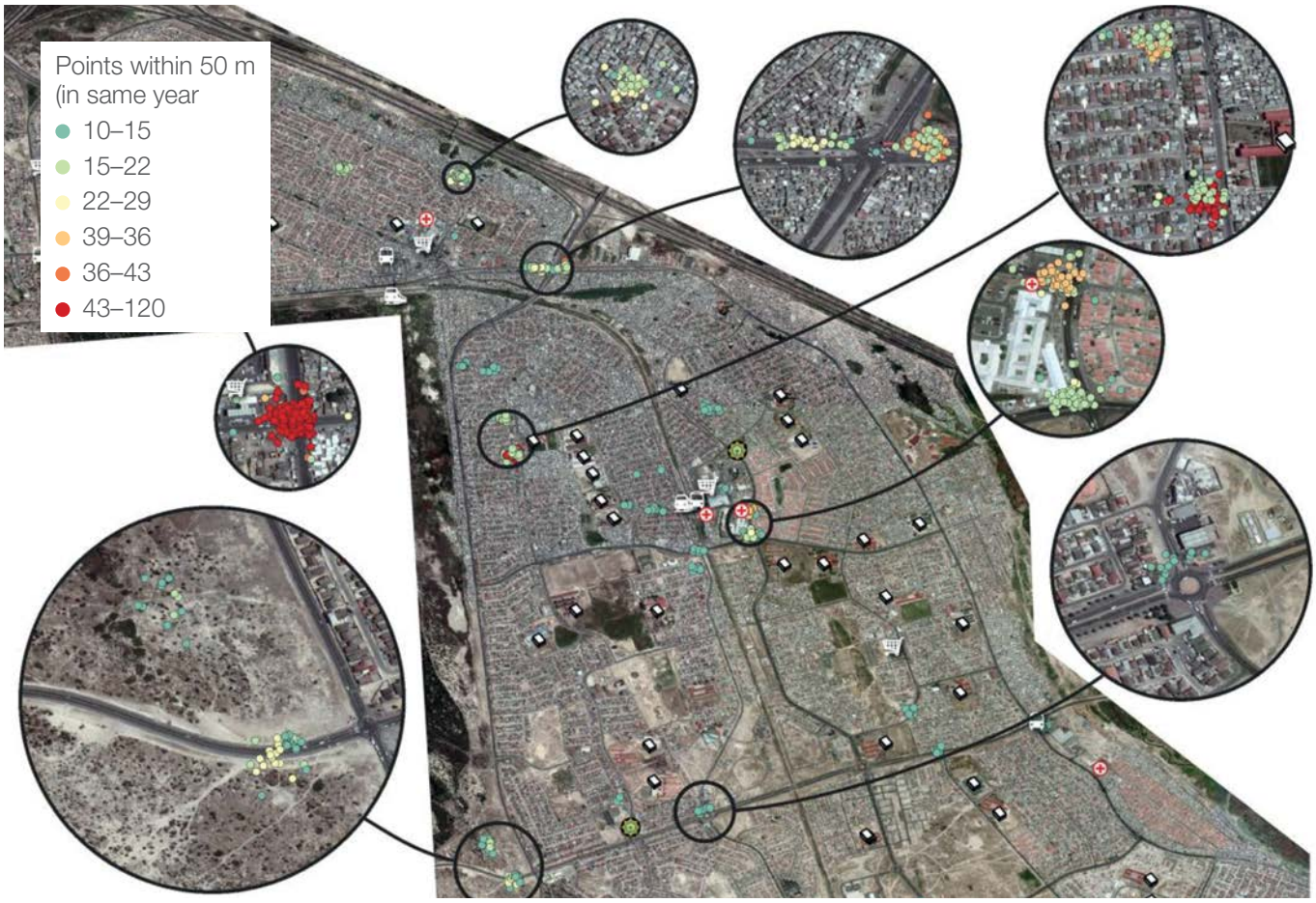
In the 2010–2015 F-mode analysis (Figure 9), the picture changes dramatically. Only eight significant F-mode clusters emerge during this period across the same area (as Figure 8). In the zoom windows, grid patterning is apparent in six of the eight significant hotspots and is further analysed in the visual data anomalies section below.

Nearest neighbour hierarchical spatial clustering

Viewed more closely, the nearest neighbour convex hulls[22] attempt to refine hotspot shapes based on the concentration of crimes within a specified distance of each other. In this case, 200 crimes within 200 m of each other were used as the threshold levels. In Figure 10, two major hotspots are evident on the northern and southern fringes of the Khayelitsha police station.

The hotspot to the north, the hottest spot overall with 555 robberies, is the location of suspect (gridded) data (explored in the next section), potentially an artefact of the geotagging of data at and next to the police station when no other address or coordinates were available. The hotspot to the south, however, includes

Figure 8: F-mode analysis showing all points with 10 or more incidents in the same year within a 50 m radius for 2007–2009



Data source: South African Police Service

both the Nonqubela Mall, a large shopping centre in Khayelitsha and clearly a potential attractor or generator of robberies, as well as the Site B Khayelitsha Community Health Centre and the Ubuntu Clinic. Many murder and aggravated robbery victims may be identified at these medical facilities and these crimes then mistakenly geolocated here when it is not possible to obtain more accurate incident details.

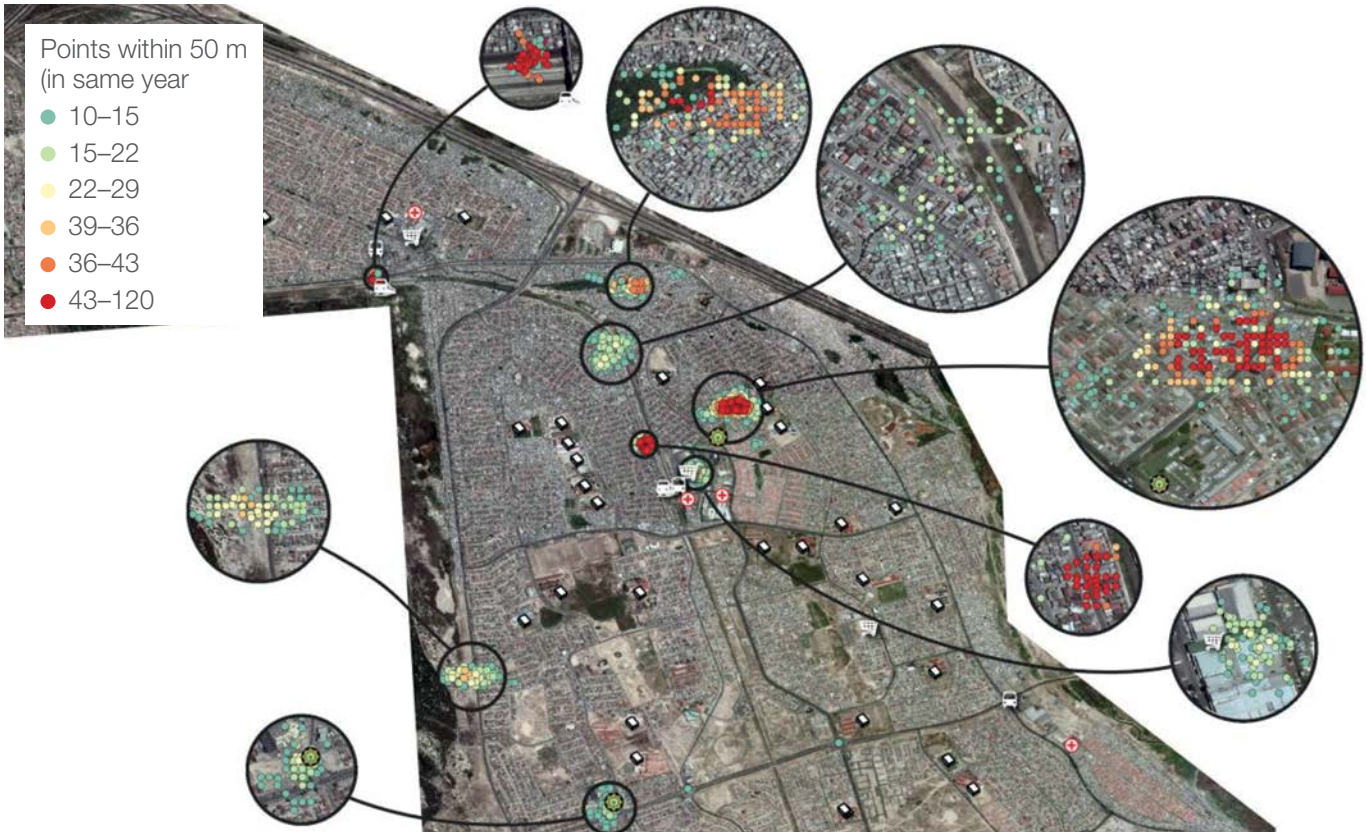
It also appears that the Khayelitsha police station area hotspots roughly follow the Metrorail corridor that splits Khayelitsha east to west, but the hotspots are predominantly located on the northern and eastern side of the railway. This may again be the result of several major intersections and railroad crossings near large informal settlements serving as the go-to snap points for approximating locations. The two Lingelethu West hotspots are centred on the police station and a point along Mew Way where grid-patterned data is apparent. This seems to compromise the reliability of the Lingelethu West hotspot data. In Harare, the hotspots appear to be more distributed and were not detected in the area in the F-mode or Nnh analyses.

Visual data anomalies

Within the hotspots identified above, there is visual evidence of unrealistic data patterns. At least four prominent groupings emerge in visualisations of the data bearing grid-patterned characteristics.

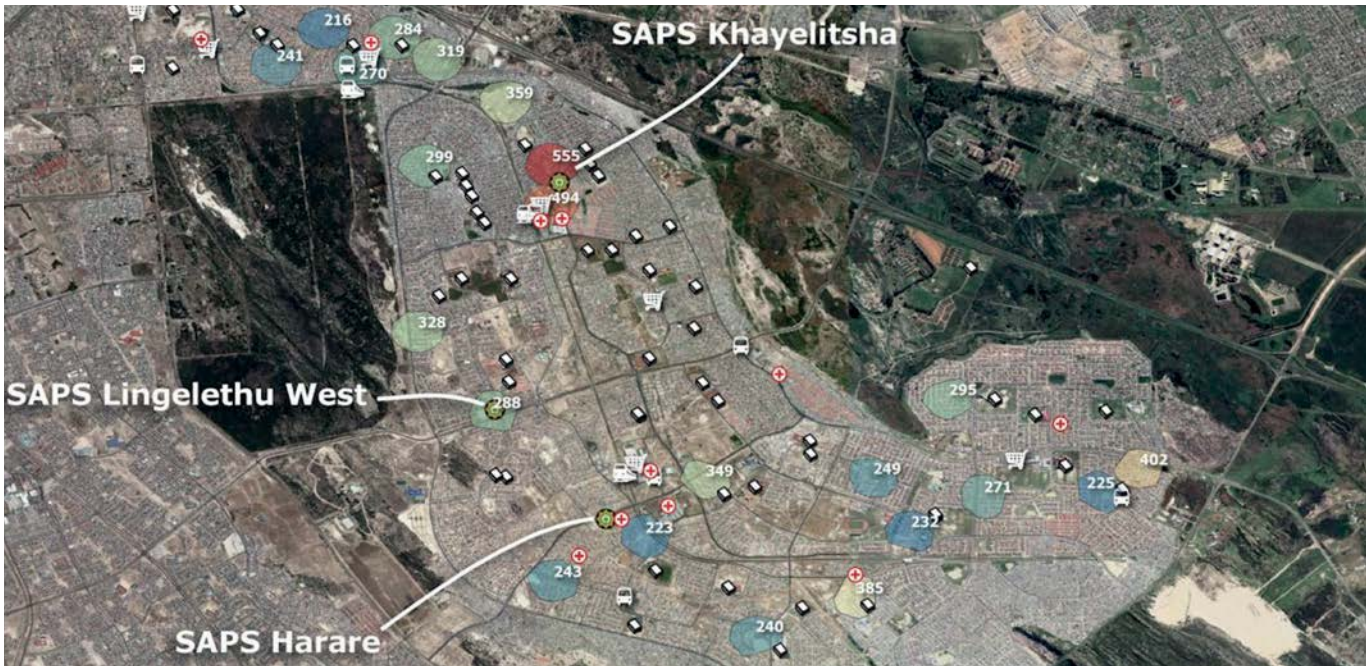
Looking closely at the data points immediately to the northeast of Khayelitsha police station in 2010-2015 (Figure 11), a prominent rectangular grid pattern is evident that extends across intersections and neighbourhoods and into, what appears to be, a school complex. Individual inspection of the data points did not reveal any clear temporal relationships (some were

Figure 9: F-mode analysis showing all points with 10 or more incidents in the same year within a 50 m radius for 2010–2015



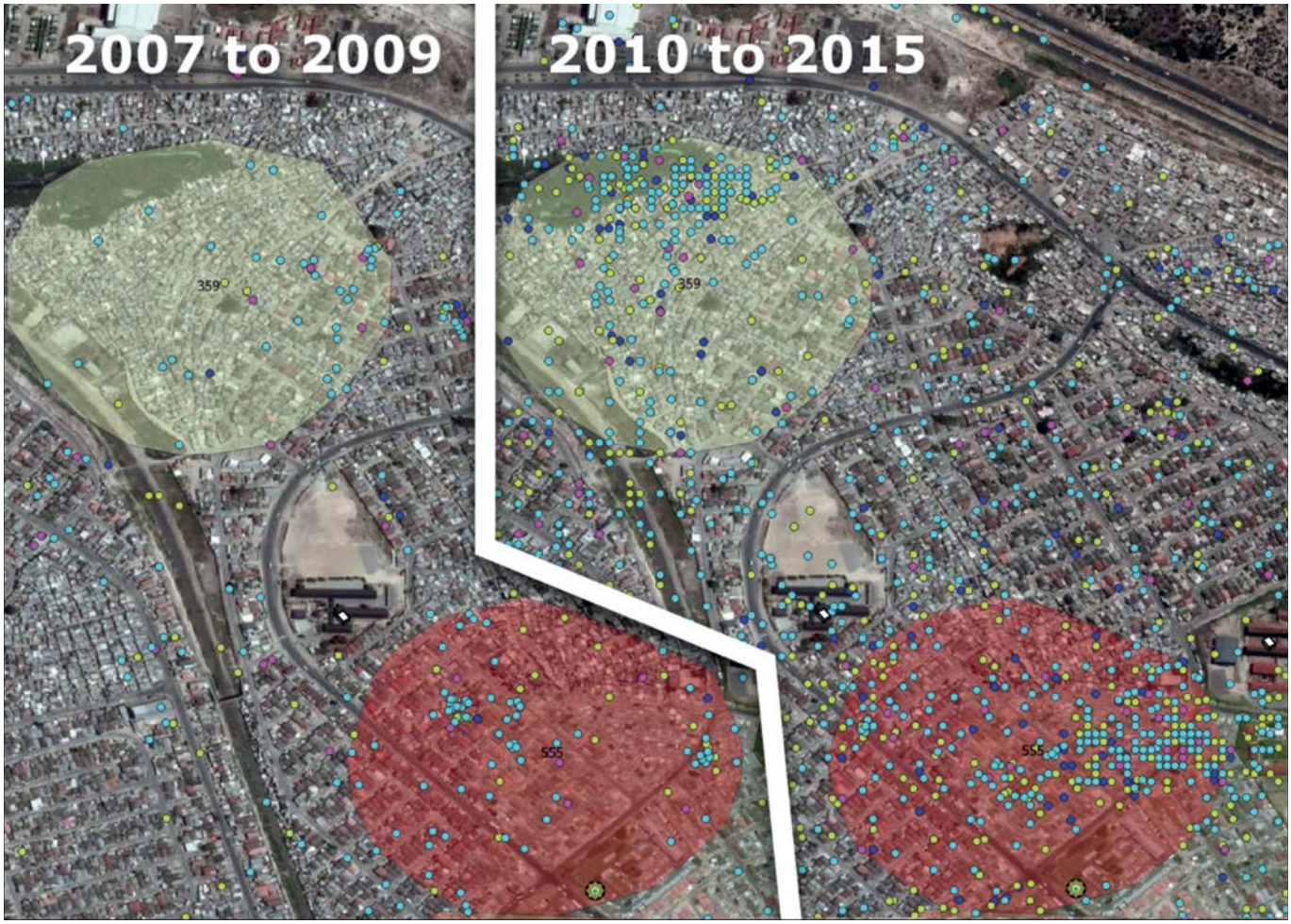
Data source: South African Police Service

Figure 10: Nearest neighbour hierarchical spatial clustering for 10-year robbery data only – set at minimum 200 incidents within 200 m of each other



Data source: South African Police Service

Figure 11: Grid-patterned data in Khayelitsha police station precinct: 2007–2009 vs. 2010–2015



Data source: South African Police Service

close in time, others were separated by months and, in some cases, years). Most of this grid-patterning adjacent to the police station seems to appear in 2013 and increases through 2015, before it tapers off in 2016. More recent data (from 1 April 2016 onwards) has yet to be accessed and analysed.

Inspection of the GPS coordinates of the points to the north of Khayelitsha police station reveals that the points are precisely 0.0000863 decimal degrees (0.3107 seconds) apart, in both the longitudinal and latitudinal directions. Similar patterns (with identical spacing) are seen near the northern edge of Khayelitsha (Figure 11), extending from an informal settlement directly into a green belt, in the 2010–2015 period.

The gridded points to the northwest of Lingelethu West police station and those surrounding the police station (Figure 12) are exactly 0.00019664 decimal degrees (0.7079 seconds) apart. The larger grouping of these points extends from formal neighbourhoods east of Mew Way straight across the road and into an undeveloped area.

These grid-patterned points do not lie precisely on lines of latitude or longitude, nor do they lie on rounded values (whether using decimal degrees, minutes or seconds). The spacing between the points is also not a round number (whether using decimal degrees, minutes or seconds). The two preceding conditions indicate that the gridded nature of the points is not due to rounding of the coordinates during the data capturing process, unless an input system other than decimal degrees or DMS (degrees minutes seconds) was used.

This patterning of crime data is possibly the result of a computer programme that takes a starting location (near the Khayelitsha police station, for instance) and assigns crimes to GPS coordinates across a grid, in such a way that they are interspersed through a designated area.

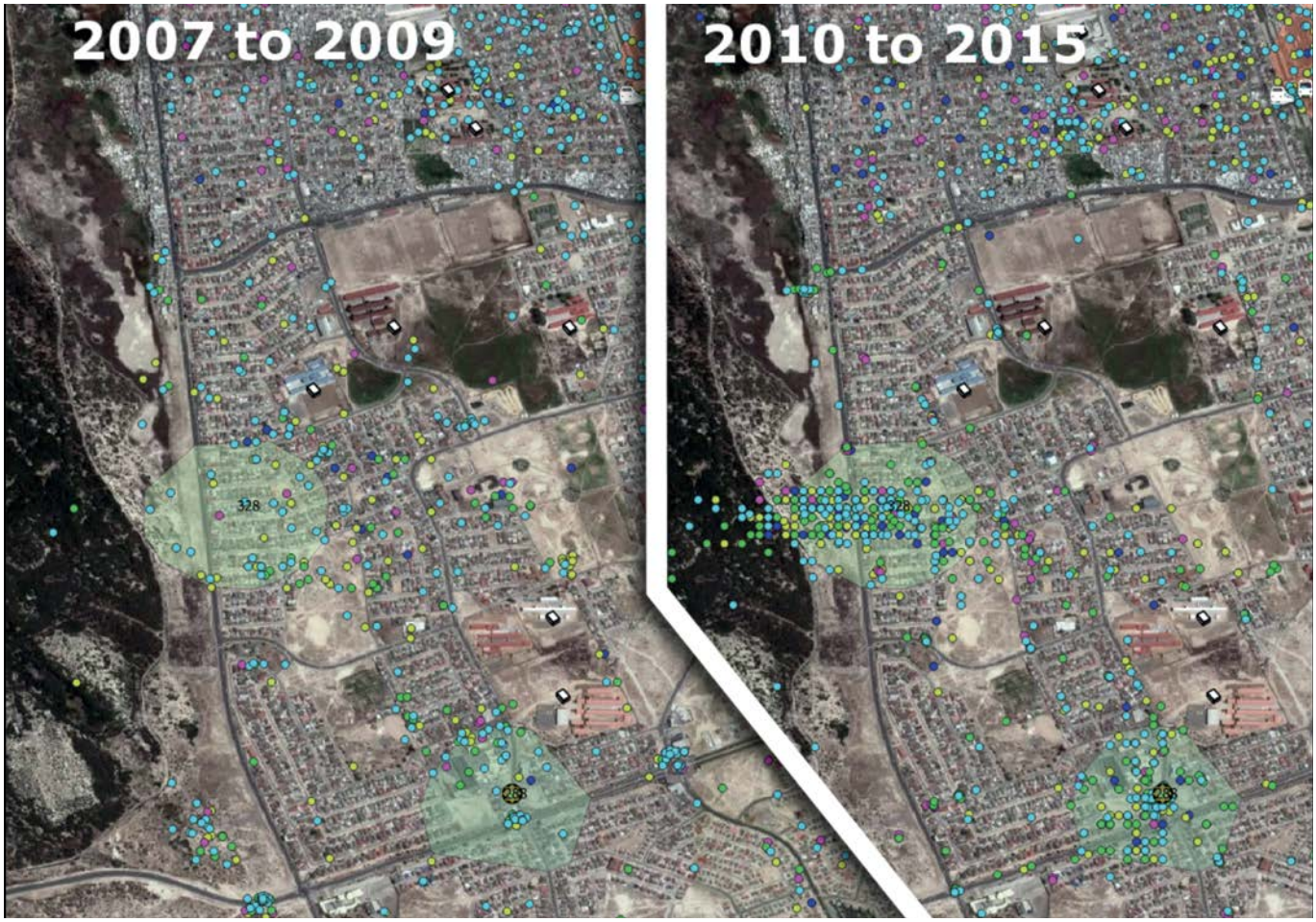
Another possible explanation for the grid patterning could be if an automatic geolocation device was used, and the device introduced course rounding errors before converting the data to decimal degrees (although this would not explain the sudden decline in hotspots that were present in 2007–2009, and the sudden increase in crime in the gridded areas). By doing a cursory count, it is estimated that this grid-patterning (and potential faulty geolocation of crime incidents) may be affecting approximately 2 000 of the crime points in this dataset, or roughly 10% of the data.

Closer inspection of hotspots

In this section a closer look is taken at several prominent hotspots identified in the Nnh analysis that appear to have more random crime distributions. In the first convex hull hotspot in Site B Khayelitsha (Figure 13), there appear to be two clusters accounting for a significant proportion of the crimes, particularly robberies with a firearm (making this a second order F-mode hotspot).

The cluster to the bottom of the convex hull appears to be at an intersection (Tandazo Drive and Bathembu Crescent) adjacent to both a school (Ikhusi Primary School) and an informal settlement (known locally as UT Section). Relatively few crimes have been geolocated inside UT Section over the 10-year period despite its local reputation as a high-crime area.[23]

Figure 12: Grid-patterned data in Lingelethu West precinct: 2007–2009 vs. 2010–2015



Data source: South African Police Service

This suggests the possibility that many crimes occurring in this informal settlement are being geotagged at this nearby intersection. The second cluster within Figure 13 is somewhat less concentrated and appears to be centred on a formal residential block. However, it also is in close proximity to informal settlement areas to the north of Ikhusi Primary School, which again show relatively few crimes over the 10-year period. It is further apparent in this visualisation that some house robberies are geolocated in the middle of streets, within non-residential school facilities (to the east of the convex hull in Figure 13), and along the railway line (far right of Figure 13).

The second convex hull hotspot is located in the Makhaza area of the Harare police precinct, immediately to the west of Sinako High School (Figure 14). This area, aside from the school, comprises formal housing with a dense distribution of backyard shacks. Here, two clusters are apparent in the northeast and southwest regions of the convex hull, both near intersections (also a second order F-mode hotspot). The surrounding area shows very low-density crime.

Figure 13: Site B Khayelitsha convex hull hotspot adjacent to informal settlements and Ikhusi Primary School

Data source: South African Police Service

Figure 14: Harare convex hull hotspot adjacent to Sinako High School

Data source: South African Police Service

- Business robbery
- Common robbery
- House robbery
- Robbery with a weapon or instrument (other)
- Robbery with firearm

It is clear in this Harare crime data that robbery with a weapon or instrument other than a firearm is a common occurrence. However, as stated previously, there are no such incidences recorded over the 10-year period in the Khayelitsha police station data. From the nature of the geolocation of this data, it is difficult to discern if Sinako High School is acting as a crime generator and attractor, if these particular cross-streets are truly high-crime locations surrounded by low-crime areas, or if these intersections are serving as georeference points for most crime occuring in the broader community.

Mapping crime to guide interventions

The 2016 White Paper on Safety and Security[24] clearly states the need for both a holistic understanding of the drivers of crime and violence (this need goes beyond the exclusive use of this knowledge for policing) and the urgent need for complete, accurate, readily available and disaggregated crime data to facilitate the intended National Safety, Crime and Violence Prevention Centre.

To date, relatively little point-level crime data in South Africa has been made available to independent researchers. Analyses of this point-level data have

been limited to several years of data (now more than 10 years old) in the City of Tshwane[25] and a cursory test of the theory of law of crime concentration at places[26] which only used Khayelitsha police station data over an 8-year period.[27] This current study has analysed far more data points over a greater time frame but is restricted to spatial analysis.

The findings and visualisations of the current study provide evidence that crimes across Greater Khayelitsha are both dispersed over place and time, yet have tended to concentrate near certain landmarks (intersections, shopping centres, medical facilities, train stations and police stations). Some of these features of the built environment may be genuine crime generators or attractors. For instance, Nonqubela Mall is likely to be the location of many robberies. However, the mall is adjacent to a large medical facility and it appears that many crimes are geotagged at an intersection equidistant from the mall and the medical facility.

> The potential for spatial-temporal analysis of high crime locations in South Africa is enormous

The apparent concentration of murder and robbery at this intersection may be the result of victims arriving at the medical facility without further data being available on the actual location of the crime. This underlines the uncertainty and potential imprecision that appear in the data and precludes deeper analysis that could inform more strategic crime interventions and intelligent, data-informed and predictive policing efforts.

Further uncertainty relates to the apparently artificial, repeated grid-spacing of data from 2010–2015 (in Figures 7 and 8). Given that these problematic grid-patterned data points may represent roughly 10% of all the crime points in the dataset, any spatial and temporal analyses and conclusions drawn must be circumspect.

Chainey and Ratcliffe state, 'We also assume that to identify hotspots in crime data these data need to be fit for the purpose. For example, if crime hotspots at the street level need to be identified, crime data that are accurate, precise, complete, consistent and reliable need to be available.'[28] This must hold true for all crime data and analyses – the quality of the analyses and conclusions will always be limited by the veracity of the input data.

Conclusion

The potential for spatial-temporal analysis of high-crime locations in South Africa is enormous and the technology is (probably) in place to make this work accessible, affordable and productive. Sufficient financial and human resources are required at the places where crime is most prevalent, alongside the technical capacities and management systems that will induce fidelity in geocrime data collection, verification, compilation and coding practices. This includes explaining to all data users exactly how geocoded crime data is captured and the uncertainties (in space and time) that accompany all records with incomplete information.

As long as SAPS crime data does not meet Stats SA criteria for official statistics,[29] it cannot be assumed that the data will enable genuine diagnosis and treatment of crime and violence. In addition, the crime data should indicate when a single crime falls into multiple categories. For example, a house robbery with a firearm may fall into two categories: house robbery and robbery with a firearm. Both categories should be r eported for the same incident record, but it should be made clear that only a single crime occurred.

South Africans concerned with crime, violence and insecurity, particularly regarding the poorest and most marginalised citizens, must take heed of the recommendations with regard to crime data in the cabinet-approved 2016 White Paper on Safety and Security:[30]

The ability to effectively plan and monitor implementation of the White Paper and assess delivery is predicated on reliable data.

- Reliable and up-to-date data must be collected across the range of departments and sectors to:
  - Identify and define the incidence and prevalence of crime and violence reported and unreported;

  - Identify the scale, scope and location of safety problems;
  - Identify specific risk and protective factors;
  - Identify availability and gaps in services;
  - Assess effectiveness of allocation of resources;
  - Identify, develop and test interventions, which can then be implemented; and
  - Evaluate what works and develop a depository of evidence based knowledge for future use.
- Data must be disaggregated to facilitate analysis and identification of drivers and risks factors.

These recommendations must be adequately resourced and absolutely central to the notional national development plan of South Africa.

Further areas of research

Temporal analysis was largely beyond the scope of this initial report which was focused on spatial analysis and crime-pattern visualisation. Further research with the existing data could explore temporal aspects and variations such as time of day, day of the week, day of the month, months of the year and seasonal variations (for instance, between winter and summer months when there is the greatest variance in hours of darkness).

It would be both possible and prudent to compare spatial and temporal aspects between the three Khayelitsha precincts themselves and with other high-crime communities and stations. This would allow researchers to discern if the data anomalies seen in parts of Greater Khayelitsha are unique and if other high-crime areas show different crime patterns, both spatially and temporally.

Crucially, point-level crime data analysis offers the possibility to assess the localised effects of crime prevention and reduction initiatives, both through the security and social sectors. Many crime and violence interventions in South Africa could be analysed for crime reduction impact by comparing spatial-temporal crime trends before and after interventions and with suitable comparisons and controls. A growing body of western research has shown that data-informed policing (when accompanied by the appropriate analysis and interpretation) can reduce crime levels.[31]

Endnotes

1 A Guerry, *Essai sur la statistique morale de la France: precede d'un rapport a l'Academie de sciences*, Paris: Chez Crochard, 1833.

2 H Mayhew, *London labour and the London poor*, London: Griffin Bohn, 1862.

3 E Burgess, The growth of the city: an introduction to a research project, in R Park, E Burgess and R McKenzie (eds.), *The city*, Chicago: University of Chicago Press, 1925, 47–62. C Shaw and H McKay, *Juvenile delinquency and urban areas*, Chicago: Chicago University Press, 1942.

4 Geographic information systems and software allow geographic data to be captured, manipulated, analysed, managed, stored and presented.

5 L Cohen and M Felson, Social change and crime rate trends: a routine activity approach, *American Sociological Review*, 44, 1979, 588–608. M Felson, *Crime and everyday life: impact and implications for society*, Thousand Oaks, California: Pine Forge Press, 1998.

6 R Clarke and M Felson, Introduction: criminology, routine activity, and rational choice, in RV Clarke and M Felson (eds.), *Routine activity and rational choice*, New Brunswick: Transaction Publishers, 1993, 259–294.

7 R Merton, Social structure and anomie, *American Sociological Review*, 3, 1938, 672–682.

8 G Newman, Introduction: towards a theory of situational crime prevention, in G Newman, R Clarke and SG Shoham (eds.), *Rational choice and situational crime prevention: theoretical foundations*, Dartmouth, UK: Ashgate, 1997, 1–23.

9 R Agnew, A general strain theory of community differences in crime rates, *Journal of Research in Crime and Delinquency*, 36(123), 1999, 123–155.

10 PJ Brantingham and PL Brantingham, *Patterns in crime*, New York: Macmillan, 1984.

11 S Chainey and J Ratcliffe, *GIS and crime mapping*, Chichester: John Wiley, 2005.

12 J Ratcliffe, The hotspot matrix: a framework for the spatio-temporal targeting of crime reduction, *Police Practice and Research*, 5(1), 2004, 5–23.

13 SAPS publicly reported crime statistics refer to *street robbery* and not *common robbery*. However, in this dataset supplied by SAPS, the *common robbery category is used.*

14 S Decker and C Kohfeld, Crimes, crime rates, arrests, and arrest ratios: implications for deterrence theory, *Criminology*, 23(3), 1985, 437–450.

15 QGIS, Free and open source geographic information system, *Open Source Geospatial Foundation Project*, 2015.

16 N Levine, Introduction to *CrimeStat IV: a spatial statistics program for the analysis of crime incident locations*, Version 4.0, Washington, DC: National Institute of Justice, 2013.

17 Ratcliffe, 2004. J Eck et al, *Mapping crime: understanding hotspots*, Washington: US Dept. of Justice, Office of Justice Programs, 2005.

18 However, it should also be acknowledged that the 2011 census data is now out of date. Khayelitsha is dominated (more than 50% of housing) by informal dwellings that can frequently move and expand, and many robberies may have more to do with patterns of movement and opportunities for crime than solely the number of residents. It is further apparent in the data that informal settlement areas, anecdotally 'notorious' for high levels of crime, frequently feature little to no crime recorded over certain years, likely due to the geolocation practices of the police.

19 N Levine, Crime mapping and the Crimestat program, *Geographical Analysis*, 38(1), 2006, 41–56.

20 Ibid.

21 W Perry, *Predictive policing: the role of crime forecasting in law enforcement operations*, Santa Monica: Rand Corporation, 2013.

22 Levine, 2006, 2013.

23 I Edelstein, Measuring pathways to youth violence and the possible effects of a sports-based development intervention on youth in South Africa, PhD thesis, University of Cape Town, Cape Town, 2016.

24 Civilian Secretariat of Police, *White Paper on Safety and Security*, Pretoria: Government Printer, 2016.

25 For example: G Breetzke, Modeling violent crime rates: a test of social disorganization in the city of Tshwane, South Africa, *Journal of Criminal Justice*, 38(4), 2010, 446–452.

26 D Weisburd and S Amram, The law of concentrations of crime at place: the case of Tel Aviv-Jaffa, *Police Practice and Research*, 15(2), 2014, 101–114.

27 G Breetzke and I Edelstein, The spatial concentration and stability of crime in a South African township, *Security Journal*, 2018, 1–16.

28 Chainey and Ratcliffe, 2005, 148.

29 Parliament of South Africa, Crime statistics 2017/18 analysis: Stats SA, Academic Civil Society, CSPS and SAPS, 24 October 2018.

30 Civilian Secretariat of Police, 2016, 14–15.

31 GO Mohler et al, Randomized controlled field trials of predictive policing, *Journal of the American Statistical Association*, 110(512), 2015, 1399–1411.

INSTITUTE FOR SECURITY STUDIES

About the authors

Dr Ian Edelstein is a research specialist at the Human Sciences Research Council. His research interests include interpersonal violence, youth development, and poverty and inequality studies. He specialises in quantitative analysis of programme effects and developmental trajectories.

Robert Arnott is a mechanical engineer specialising in energy efficient design through computer modelling and analysis – a skill central to GIS analysis of crime data.

About ISS Southern Africa Reports

Southern Africa Reports provide the results of in-depth research on the latest human security challenges in the region. Some reports analyse broad conflict trends and threats to peace and security in specific countries. Others focus on challenges in the region such as electoral reform, corruption or intra-state conflict.

About the ISS

The Institute for Security Studies (ISS) partners to build knowledge and skills that secure Africa's future. The ISS is an African non-profit organisation with offices in South Africa, Kenya, Ethiopia and Senegal. Using its networks and influence, the ISS provides timely and credible policy research, practical training and technical assistance to governments and civil society.

Acknowledgements

Hanns Seidel Foundation

Bayerische Staatskanzlei

This report is funded by the Hanns Seidel Foundation and the Bavarian State Chancellery. The ISS is grateful for support from the members of the ISS Partnership Forum: the Hanns Seidel Foundation, the European Union and the governments of Canada, Denmark, Finland, Ireland, the Netherlands, Norway, Sweden and the USA.

Photo credit: Masixole Feni/GroundUp

ISSN 2617-8206 Print
ISSN 2617-8214 Digital

9 772617 821002