

---

# Spatiotemporal Interaction of Urban Crime in Nairobi, Kenya

Lucy MBURU

Heidelberg University / Germany · lucy.waruguru@geog.uni-heidelberg.de

This contribution was double-blind reviewed as full paper.

## Abstract

Crime is regarded as one of the major threats to socio-economic development. The variable offending behaviour of criminals is nonetheless little understood. Clusters of crime have direct implications on crime prevention strategies. This study examines variability in offending patterns, based on the spatial and temporal patterns of historical crime. Point patterns of criminal events are mapped over a study area in eastern Nairobi, and the distribution of various types of crimes is examined throughout a 24-months study period. Although various other factors influential to crime are identified, the association between rainy seasons and high rates of crime is distinctly observed.

## 1 Introduction

Concern over the effects of criminal activities is growing throughout the world and especially in major cities. In the city of Nairobi, Kenya, the frequency and fatality of crime have rapidly increased over the past few years. Most of the prevalent information regarding the existence and changes of crime patterns, as well as on the impact of observed patterns on individuals and communities, has been gathered in the developed world. Nevertheless, even here the understanding of factors which drive different types of crime in the spatial and temporal spaces is still limited. In this regard, many fundamental questions remain unsolved.

PEASE (1998) and JOHNSON et al. (2007) contribute a solution to this puzzle with the observation that clusters of crime emanate from the presence of crime generators, as well as the encouragement that successfully committed crime offers criminals the possibility to continue offending within their surroundings. This offending behaviour is further demystified by the criminological principle of the offender's mental map (PUTWAIN & SAMMONS 2013), which purports offenders to gradually build a "safety zone" around the areas they frequent and within which they feel comfortable. The safety zone expands with increased offending and elapsed time, and largely accounts for repeat victimization and serial offending (MBURU & HELBICH 2014). SHAW & PEASE (2000), for example, observed 68 percent of burglaries to occur on the same side of a city block. Such consistent offending within similar neighbourhoods generates hotspots of criminal activity, as has been identified in various studies (e.g., JOHNSON et al. 2007, RATCLIFFE 2010, UITTENBOGAARD & CECATTO 2011, HELBICH & LEITNER 2012). Multiple social and economic consequences

of this clustering exist (e.g., increased fear, depreciating housing prices), and their discovery necessitates prompt and focussed police attention. From the perspective of crime prevention, an ideal solution is to anticipate elevated risk of crime by observation of possible stimulants.

## 2 Modelling the Seasonality of Crimes

Concentrations of criminal activity are triggered by particular attractors such as drug trading, specific land uses (e.g. thefts around bus stops), and weather patterns (CECCATO & UITTENBOGAARD 2011, 2014). Voluminous literature exists on the seasonal variability of crime (e.g., JOHNSON ET AL. 2007, NAKAYA & YANO 2010, ANDRESEN & MALLESON 2013, CECCATO & UITTENBOGAARD 2014), and reviews of such literature (e.g. BAUMER & WRIGHT 1996), but elementary questions still exist. A common consensus across findings is that offenses, perhaps with the exception of homicide (AJDACIC-GROSS 2010), vary across the yearly seasons. The degree and direction of this variability is nevertheless reported differently among different studies. FIELD (1992), for example, observed the cold season to precipitate violent crime and the summer to be characterized by peaks of non-violent crime, while FARRELL and PEASE (1994) observed opposite cycles for violent and non-violent offenses. Later studies (e.g., ROTTON & COHN 2000, HIPPEL et al. 2004) observed peaks in both non-violent and violent crimes during summer, but noted other complex associations between weather and crime.

With regard to co-variability in the location and time of offending, CECCATO and UITTENBOGAARD (2014) discovered peaks of crime in the evening, at night, on holidays, and on weekends at underground stations. They also observed higher crime in winter, as well as summer clusters of crime in the close vicinity of alcohol-selling establishments. Further substantial differences exist when the observation shifts from the broader classification of crime to individual offenses. HIRD and RUPAREL (2007), for example, observed increased sexual offenses during the hot summer period, while ROTTON and COHN (2000), and CECCATO and UITTENBOGAARD (2011) observed robbery peaks in the same season.

In contrast to the above areas of study, Nairobi experiences no extreme weather changes during winter and summer. Temperatures vary between 10°C and 23°C in winter, and between 25°C and 33°C in summer. Fog, snow and ice are not characteristic of the study area. The yearly seasons are however characterized by two rainy periods: long rains, often between February and June, and short rains, between September and December. Rains are often characterized with thunderstorms, which frequently occur at night, and cause interruptions of electricity. Darkness, noise from rain and thunder, and slow police response during this time enticing criminals, are all major crime generators. Cries for help and the commotion caused by burglars remain obscured during thunderstorms. TOMPSON and BOWERS (2013) observed darkness, rather than a fall in temperature, to generate robbery during winter in the UK. The focus of this study is to establish the degree to which these patterns generate crime. Other possibly influential factors, such as environmental characteristics and vicinity of police stations are, however, also considered.

### 3 Methods

#### 3.1 Study Site and Data

The site chosen for observation is located in eastern Nairobi and is part of city's metropolitan area. This site lies only 10 km east of Nairobi's central business district (CBD), an extremely busy commercial area, traversed daily by over 500,000 people in addition to its normal population (UNEP 2008). Extending 43 square km, the study area is spatially subdivided into 38 housing estates and sub-estates (termed as "phases"). Two informal settlements contribute to the dense population of this area. The city's rubbish dumping site is also located in this area, and provides various hideouts to criminals. Almost half of crime incidents in the city within the months of November and December, 2010 (47 percent) were reported in eastern Nairobi, as opposed to 13% of crime within the CBD (KENYA POLICE 2010). This makes the area an ideal observation ground for studying the spatial and temporal distribution of crime, as well as its dynamics within a natural community and on a relatively small geographic scale.



**Fig. 1:** Residential housing estates and sub-estates in eastern Nairobi

After extensive geo-coding, 2,641 offences reported between January 1st, 2009 and December 31st, 2010, together with their associated spatial information, were gathered for this study. Crime events were obtained from the police station in Buruburu, Nairobi as individual dated cases. A 24-months study was necessary in order to obtain sufficiently conclusive patterns that allow cluster detection within ordinary offending reports. Associated weather data for the entire period of study was also obtained from TuTiempo, a web service which provides historical and current information about weather in major world cities. Only data from the weather station 637400 (long 36.917, lat -1.317) which lies four km from the study area was used.

## 3.2 Analyzing Spatial Clusters of Crime

### 3.2.1 Cluster Identification

RIPLEY'S (1976)  $K$  index quantified the clustering of crime in the study area in terms of the degree and spatial scale.  $\hat{K}$ , whose circumflex indicates the derived estimate to stem from a single realization of an underlying process, was computed as;

$$\hat{K}(d) = \frac{1}{\lambda^2 A} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \frac{I_d(d_{ij})}{w_{ij}}$$

$A$  represents the entire study area.  $\lambda$  is the intensity, i.e. number of crimes divided by the study area,  $n$  is the number of crime locations,  $d_{ij}$  is the distance between any crime locations,  $i$  and  $j$ , and the Boolean variable,  $I_b(d_{ij})$  indicates whether (1) or not (0) a crime location exists within a distance,  $d$  from the first-observed crime,  $i$ . To account for border effects, a weighting factor  $w_{ij}$  was used to define the proportion of distance band,  $d$  situated within the study area (DIGGLE 1983).

With input from the  $K$ -function, a randomized test was formulated to establish whether the  $n$  crime locations found in the study area were significantly aggregated in space as compared to aggregations that would be expected by random chance.  $K$  index was transformed into  $L$  function ( $\hat{L}(d)$ , BESAG 1977), resulting in a linear expected value:

$$\hat{L}(d) = \sqrt{\frac{K(d)}{\pi}} - d$$

Clusters were evaluated for the entire dataset, as well as for each of the two temporal windows: high- and low-crime periods. Additionally for the high-crime period, 999 simulation tests were used to compute a critical envelope,  $L(d)$  with ranges between  $L_{lo}$  and  $L_{hi}$ , and the observed value,  $L_{obs}$ . If  $L_{obs}$  fell below the average expected value, the crime distribution was termed as random, or at least as not statistically significant from any random observation.  $L_{obs}$  appearing above the expected line indicated a significantly non-random distribution  $\alpha=5\%$ . Values falling outside the envelope were considered to be outliers. This test was used for both non-violent and violent categories of crime.

### 3.2.2 Density Estimation

Kernel density estimation is a particularly popular approach owing to its high ability for visualizing concentrations of crime. Estimating density involved placing a symmetrical surface, called the kernel function,  $k(\cdot)$ , over each event,  $S_i$ , and summing-up the values of all surfaces onto a regular reference grid which covered the whole observation area,  $A$ . Such a kernel function falls off with distance from each event at a rate which depends on the shape of this function and its bandwidth,  $b$ . Density estimation is nevertheless limited by its provision of purely spatial evidence of changes. SOM was therefore used to allow simultaneous modelling of spatial and temporal characteristics of events.

### 3.2.3 The Self-organizing Map

SOM (KOHONEN 2001) is a form of artificial neural network (ANN), which is recursively trained using unsupervised learning to produce a low-dimensional (typically 2D), representation of higher dimensional entities. Geometric relationships between the image points indicate their relative similarity. A neighbourhood function preserves the topological properties of the input space, akin to multidimensional scaling, and makes SOMs useful for visualizing high-dimensional data (HAGENAUER & HELBICH 2011, HELBICH & HAGENAUER 2013). SOM is also easy to interpret and can handle large complex datasets.

Like most artificial neural networks, major phases of the SOM consist of training and mapping. The training phase uses input samples to build a map in a competitive process called "vector quantization", while mapping classifies a new input vector. The complete process underlying the SOM involves six steps:

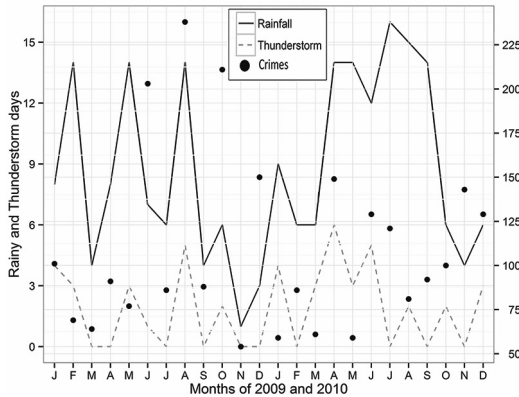
- Randomly initializing all weight vectors  $w_j$  for all nodes  $j$
- Selecting an input vector  $x = [x_1, x_2, x_3, \dots, x_n]$  from the training set
- Comparing  $x$  with weights  $w_j$  for each neuron  $j$  to determine winner
- Updating the winner to become akin to  $x$ , together with the winner's neighbours
- Adjusting the parameters: learning rate & 'neighbourhood function'
- Repeating the second step until the map is converged (i.e., there exist no observable changes in the weights) or pre-defined no. of training cycles have passed.

For this study, SOMs were produced using the visualization system for space-time and multivariate patterns (VIS-STAMP, GUO et al., 2006). As input for generating SOMs, crime events were aggregated according to multi-dimensional variables: spatial variables comprised the housing estate phases depicted in Fig. 1 above, while temporal variables were the individual months which constituted the observation period. Computational methods were combined with visual- and cartographic methods to train and map matrices. The resultant U- distance matrix was linked to a parallel coordinate plot and a central matrix of maps, each map representing a timestamp of the study period. SOMs were based on a network of 7-by-7 nodes for all crimes in each of the study area's sub-estates for the entire 24-month study period.

## 4 Results and Discussion

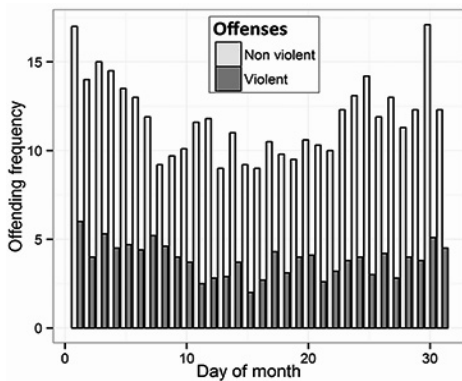
### 4.1 Descriptive Statistics

Basic analysis examined each month for the presence of seasonal offending. Also observed for each month were the numbers of rainy days, as well as the number of thunderstorms to assess the influence of weather on offending patterns (Fig. 2). In Kenya, many thunderstorms occur at night, or commence during the evening hours and extend into the night. It was therefore assumed that a majority of the observed rains and thunderstorms were nocturnal and that they caused the absence of safety guardians to a certain extent.



**Fig. 2:** Monthly offending patterns from Jan, 2009 to Dec, 2010 plotted against the number of rainy days (bold line) and thunderstorm days (dashed line)

Weather, and particularly thunderstorms, significantly affected the rate of offending. All months with more than three days of thunderstorms experienced more than 80 crimes. Months characterized by many rainy days (i.e., July, August, September, October and December of 2009, and April, June, July, November and December of 2010) also showed high offending peaks. November, 2009, which had the least rain and no thunderstorms, had the lowest number of crimes. This pattern of routine activity is akin to that observed RANSON (2012), who reports rain to extensively generate crime. But despite the appearance of some clear patterns of offending, non-uniform offending was observed across different months. Hence crime events for the entire observation period were examined for each day of the month. Further investigation of different offending behaviour also necessitated observation of events as two separate groups: violent (involving armed robbery and rape) and non-violent (involving non-armed robbery, theft, burglary and drug dealing). These two categories were observed separately (Fig. 3).



**Fig. 3:** Monthly offending patterns observed between Jan 1, 2009 and December 31, 2010

Non-uniform offending across different days of the month accounted for the variability observed in Fig. 2. Fewer crimes occurred during the middle of the month than at the start and the end. The highest peak of non-violent crime appeared on the 30<sup>th</sup>. This is a typical last day of the month by which time many employees have received their pay, thereby

posing attraction to criminals. Violent crimes had less fluctuation than non-violent crimes, but the highest peak for this category was observed at the beginning of each month.

## 4.2 Testing Offense Coverage of Space

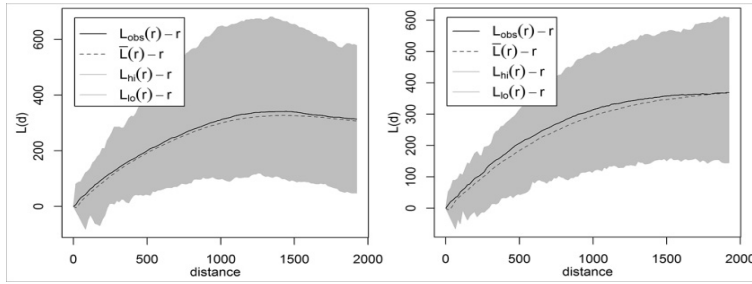
To identify the extent with which crime clusters deviated from complete spatial randomness (CSR), crime events were overlapped with predictors using the L-function described in Section 3.2.1. The clustering of non-violent crimes was tested separately from that of violent crimes. Events occurring during the observed high offending period (i.e. July, August, September, October and December in 2009, and April, June, July, November and December in 2010) were further extracted from the rest of the dataset and comparatively evaluated for spatial clustering.

**Tab. 1:** Spatial description of violent and non-violent crime events throughout the observation period and for two periods: low crime, and high crime periods

Category	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Non-violent crimes	0.0	647.1	1221.8	1180.3	1731.8	2194.6
Violent crimes	0.0	695.5	1295.9	1253.4	1839.0	2321.3
<b>High-Crime period</b>						
Non-violent Crimes	0.0	700.7	1295.2	1251.0	1828.5	2293.1
Violent Crimes	0.0	659.4	1244.7	1212.2	1781.5	2280.5
<b>Low-Crime period</b>						
Non-violent Crimes	0.0	773.1	1424.5	1369.6	1994.9	2508.4
Violent Crimes	0.0	853.5	1534.6	1476.7	2142.4	2667.4

Although the first and third quartiles described the internal structure of the event data (Table 1), the median which is more resistant to outliers was more informative. Observing events from the entire time period, for example, the smaller inter-quartile range for non-violent crimes (1,084.7) than violent crime (1,143.5) depicted denser clustering of these non-violent offences around the median. For both violent and non-violent crimes, events observed in the high-crime period were more densely clustered around the median than events observed in the low-offense period. This is despite the fact that the high-crime period comprised fewer months (10 months) than the low period (14 months).

The observation above called for the need to further test the spatial clustering observed during the high-offense period. Separating violent from non-violent crimes, critical envelopes of event clustering from this period were generated using the “spatstat” package of R (Fig. 5).

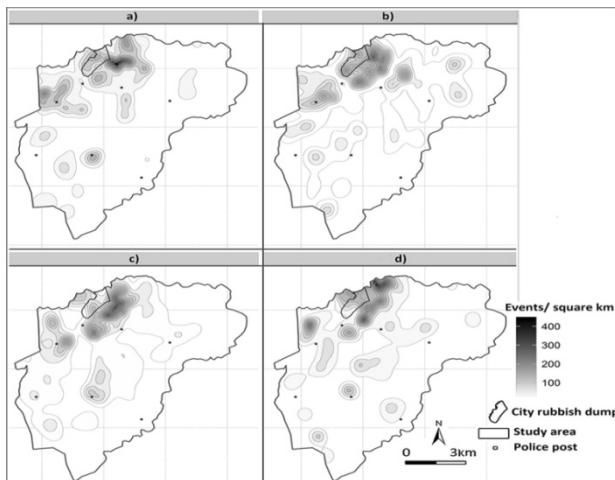


**Fig. 4:** Point-wise critical envelopes for non-violent crimes (left) and violent crimes (right) during the high crime season, at  $p=0.05$ . The dashed line represents the expected (theoretical) L-value for a CSR process. The solid line represents the observed value.

Crime density higher than CSR (Fig. 4) was observed at almost all distances. Violent crimes were especially clustered at very short distances (i.e., between 400m and 1,200m). Nevertheless, observations of both violent and non-violent crimes highlighted a satisfactory behaviour of the L-function in a spectacular manner.

### 4.3 Spatial Hotspot Identification

Hotspot mapping with kernel density estimation confirmed the locations of the crime clusters identified. Density estimation used Diggle's edge correction, which is slower to compute, but yields a more accurate output. Clusters from each year were estimated separately, with each sub-dataset further split into violent and non-violent categories. Four groups constituted the resultant dataset for this analysis.



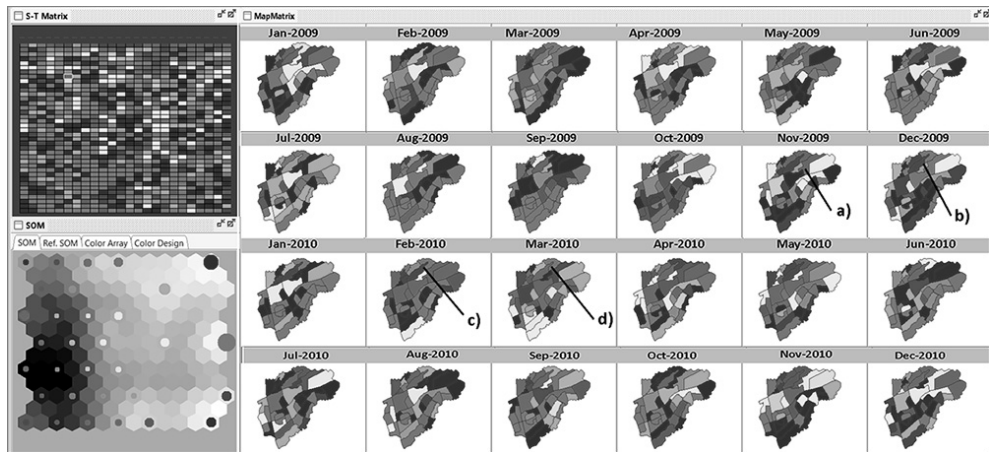
**Fig. 5:** Peak-period clusters in 2009 (left) and 2010 (right) of violent (a and b) and non-violent (c and d) crimes



Numerous clusters of non-violent crime were observed throughout the study area, especially in 2010. Four police posts are located at the heart of these clusters. Violent clusters of crimes also significantly appeared in this area, confirming the critical measuring of clustering in Fig.4. The most dense clusters appeared near the city's dumping site which is situated north of the study area. Clusters east and south of the area were less prominent.

#### 4.4 Space-time Statistics

Interlinked analysis and display of re-orderable and colour-coded matrices with VIS-STAMP assessed clusters in the spatio-temporal space. The entire crime dataset was observed in monthly windows throughout the entire period (Fig. 6).



**Fig. 6:** SOM clusters of crime events in each sub-estate and their relative distances (bottom right) with a linked spatio-temporal matrix (top right) and a map matrix (center)

24 temporal rows and 38 spatial columns made up the spatio-temporal matrix (top right). Rows represented a fixed order of each month. Columns showed the different sub-sections of the study area, with their ordering computationally derived from similarity in offending patterns over time. The third phase of Dandora appearing on the first row, for example, experienced the highest cumulative crime. The Donholm estate, which lies far south of the study area, experienced the lowest crime and is last on the spatio-temporal queue. The coupled SOM matrix (bottom right) provides a key to the colour coding. It also displays the spatio-temporal clusters (circles) and their respective distances (nodes) from each other. Dark shades depict smaller Euclidian distances while lighter shades show larger distances between the clusters. The blue circles show closely located clusters, most of which contain few events. On the spatio-temporal matrix (top right) these clusters were more prominent in the months of February to June, 2010. With the exception of the month of February, each of these months experienced more than 12 days of rainfall. This period also experienced the most thunderstorms. The map matrix (central panel) shows the blue-shaded clusters to surround the housing estates which lie in close vicinity of the city's dumping site. The

longest outbreak of crimes (between November, 2009 and March, 2010, marked “a” till “d”) appeared in the third phase of Dandora which lies south of this dump. Brown and red circles represent outlier clusters which contain relatively many data items. These represent the cold spots on the spatio-temporal matrix, appearing mainly from August to November, 2009. Little rainfall occurred during this period. The cold spots are located south of the study area where fewer housing estates exist and the road network is less busy.

## 5 Conclusion

This paper has presented various methods which support tactical and strategic decision making processes by retrospectively surveying criminal activities. Descriptive statistics to identify a temporal window of interest, as well as prior testing for clusters using Ripley’s statistics were helpful in narrowing down the focus of interest. Progressive step-by-step analysis with spatial and temporal cluster detection methods allowed the discovery of offending patterns. Results of this study contribute to the existing body of knowledge by providing evidence of how weather influences the rise and fall of crime. An interesting observation from the kernel density estimation (KDE) was the existence of police administrative sites within or near offence hotspots. Although some of the posts were set up after the observation period, following increased insecurity, the Dandora police post which is situated near the city’s dumping site where many crimes occur was set up in 2008. This finding, like that of STAVROU (2002), suggests the impunity of offenders in Nairobi, and highlights the need by law enforcement to develop more rigorous crime prediction and prevention counteractive measures. A WORLD BANK (2006) study reported a majority of offences in Nairobi to occur in areas exhibiting certain physical characteristics e.g., narrow and uneven paths, open drainages, unlit areas, constant electrical-power cuts and dumping areas, which provide hideouts and operational areas to criminals. Indeed the areas of the study area in which sporadic outbreaks were observed both with KDE and the SOM exhibit such characteristics. Observations of this study can aid the strategic decision-making of the police and have substantial implication on safety policymaking.

Techniques used in this study can be extended to survey clusters in other areas to identify factors influential to crime and to pinpoint problematic areas. Observations reveal the potential usefulness of clustering techniques in advising effective ways of allocating and distributing policing resources, and increasing emergency responsiveness.

## References

- AJDACIC-GROSS, V., BOPP, M., RING, M., GUTZWILLER, F. & ROSSLER, W. (2010), Seasonality in suicide – A Review and Search of New Concepts for Explaining the Heterogeneous Phenomena. *Social Science & Medicine*, 71 (4), 657-666.
- ANDRESEN, M. A. & MALLESON, N. (2013), Crime Seasonality and its Variations Across Space. *Applied Geography*, 43, 25-35.
- BAUMER, E. & WRIGHT, R. (1996), Crime Seasonality and Serious Scholarship: a Comment on Farrell and Pease. *Brit. J. Criminology*, 36, 579.

- BESAG, J. (1977), Efficiency of Pseudolikelihood Estimation for Simple Gaussian Fields. *Biometrika*, 616-618.
- CECCATO, V. & UITTENBOGAARD, A. C. (2011), Space-time Clusters of Crime in Stockholm, Sweden. In *Procedia Environmental Sciences*, 1-8.
- CECCATO, V. & UITTENBOGAARD, A. C. (2014), Space-Time Dynamics of Crime in Transport Nodes. *Annals of the Association of American Geographers*, 104 (1), 131-150.
- DIGGLE, P. J. (1983), *Statistical Analysis of Spatial Point Processes*.
- FARRELL, G. & PEASE, P. (1994), Crime Seasonality: Domestic Disputes and Residential Burglary in Merseyside 1988-90. *British Journal of Criminology*, 34 (4), 487-498.
- FIELD, S. (1992), The effect of Temperature on Crime. *British Journal of Criminology*. 32 (3), 340-351.
- GUO, D., CHEN, J., MACEACHREN, A. M. & LIAO, K. (2006), A Visualization System for Space-time and Multivariate Patterns (VIS-STAMP). *Visualization and Computer Graphics, IEEE Transactions on*, 12 (6), 1461-1474.
- HAGENAUER, J., HELBICH, M. & LEITNER, M. (2011), Visualization of Crime Trajectories with Self-organizing Maps: a Case Study on Evaluating the Impact of Hurricanes on Spatio-temporal Crime Hotspots. 25th conference of the International Cartographic Association, Paris, 2007.
- HELBICH, M. & LEITNER, M. (2012), Evaluation of Spatial Cluster Detection Algorithms for Crime Locations. In *Challenges at the Interface of Data Analysis, Computer Science, and Optimization*, Springer, 193-201.
- HELBICH, M., HAGENAUER, J., LEITNER, M. & EDWARDS, R. (2013), Exploration of unstructured narrative crime reports: an unsupervised neural network and point pattern analysis approach. *Cartography and Geographic Information Science*, 40 (4), 326-336.
- HIPP, J. R., CURRAN, P. J., BOLLEN, K. A. & BAUER, D. J. (2004), Crimes of Opportunity or Crimes of Emotion? Testing Two Explanations of Seasonal Change in Crime. *Social Forces*, 82 (4), 1333-1372.
- HIRD, C. & RUPAREL, C. (2007), *Seasonality in Recorded Crime: Preliminary Findings*. Home Office, Research, Development and Statistics Directorate.
- JOHNSON, S. D., BERNASCO, W., BOWERS, K. J., ELFFERS, H., RATCLIFFE, J., RENGERT, G., & TOWNSLEY, M. (2007), Space-time Patterns of Risk: a Cross National Assessment of Residential Burglary Victimization. *Journal of Quantitative Criminology*, 23 (3), 201-219.
- KENYA POLICE (2010), *Annual Crime Report for the Year 2010*.  
<http://www.kenyapolice.go.ke/resources> (last accessed 02 January, 2014).
- KOHONEN, T. (2001), *Self-organizing Maps*, Springer, 30.
- MBURU, L. & HELBICH, M. (2014), Evaluating the Accuracy and Effectiveness of Criminal Geographic Profiling Methods: The Case of Dandora, Kenya. *The Professional Geographer*, (ahead-of-print).
- STAVROU, A. (2002), *Crime in Nairobi: Results of a Citywide Victim Survey*. ISBN: 92-1-131614-6.
- TOMPSON, L. & BOWERS, K. (2013), A Stab in the Dark? A Research Note on Temporal Patterns of Street Robbery. *Journal of Research in Crime and Delinquency*, 50 (4), 616-631.
- WORLD BANK (2006), *Inside Informality: Poverty, Jobs, Housing and Services in Nairobi's Slums*. Report No. 36347-KE, Kenya, May 31, 2006.

