Evaluating the Accuracy and Effectiveness of Criminal Geographic Profiling Methods: The Case of Dandora, Kenya. The Professional Geographer

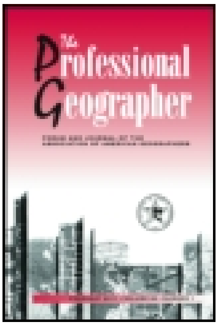
Article *in* The Professional Geographer · December 2014

DOI: 10.1080/00330124.2014.886921

2 authors:

Lucy Mburu
KCA University

Marco Helbich
Utrecht University



The Professional Geographer
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rtpg20

Evaluating the Accuracy and Effectiveness of Criminal Geographic Profiling Methods: The Case of Dandora, Kenya

Lucy Mburu [a] & Marco Helbich [a]
[a] University of Heidelberg
Published online: 19 Mar 2014.

Evaluating the Accuracy and Effectiveness of Criminal Geographic Profiling Methods: The Case of Dandora, Kenya

Lucy Mburu and Marco Helbich
University of Heidelberg

Criminal geographic profiling (CGP) prioritizes offender search, extensively reducing the resources expended in criminal investigations. The utility of CGP has, however, remained unclear when variations in environmental characteristics and offense type are introduced. This study evaluates several CGP strategies with data from Dandora, a small but densely populated suburb of Nairobi, Kenya. The research employs error distance and search-cost measures to determine CGP accuracy. Characterized by much shorter journeys to crime than those observed in Western cities, this study discovers significant applicability of CGP strategies in prioritizing offender searches. The negative exponential CGP strategy is identified to generate the most accurate geo-profiles. **Key Words: crime analysis, criminal geographic profiling, Nairobi (Kenya), offender travel, serial offense.**

犯罪地裡剖绘 (CGP) 将寻找犯罪者列为首要目标, 并因而大幅减少耗费在犯罪侦查的资源。当引进环境特徵与犯罪型态的变异时, CGP 的效用却不甚明确。本研究藉由肯尼亚奈洛比中一个小型却人口稠密的郊区丹多拉 (Dandora) 的数据, 评估部分的 CGP 策略。本研究运用误差距离与搜寻成本方法来决定 CGP 的准确性。相较于在欧洲城市所观察到的犯罪而言, 该地的犯罪移动距离短得多, 本研究因此发觉, CGP 在以搜寻犯罪者为首要的目标中具有显着的可运用性。本文亦指认出负指数 CGP 策略, 以生产最为准确的地理剖绘。 **关键词: 犯罪分析, 犯罪地理剖绘, 奈洛比 (肯尼亚), 犯罪旅程, 连续犯罪。**

La construcción de perfiles criminales en plano geográfico (CGP, según el acrónimo inglés) da prioridad a la búsqueda del delincuente, disminuyendo así notablemente los recursos que se gastan en investigaciones criminales. Sin embargo, la utilidad de los CGP sigue siendo poco clara cuando se le incorporan las variaciones en las características ambientales y el tipo de delito. Este estudio evalúa varias estrategias de CGP con datos obtenidos en Dandora, un pequeño aunque densamente poblado suburbio de Nairobi, Kenia. La investigación utiliza medidas de error de distancia y costo de la búsqueda para determinar la exactitud del CGP. Caracterizado por viajes al lugar del crimen mucho más cortos de lo que se observa en las ciudades occidentales, en este estudio se descubre una aplicabilidad significativa de estrategias del CGP para priorizar la búsqueda de delincuentes. La estrategia CGP exponencial negativa se identifica como la que genera los geo-perfiles más exactos. **Palabras clave: análisis del crimen, perfil geográfico criminal, Nairobi (Kenia), recorrido delincuencial, delito seriado.**

Investigation of serial offenses generates a myriad of suspects. It is, therefore, resource-intensive, and focusing investigations using manual methods (e.g., suspect interrogation, crime-scene analysis) becomes complex. Supported by spatial analysis and geographic information technology (see Chainey and Ratcliffe 2005; Hagenauer, Helbich, and Leitner 2011; Leitner and Helbich 2011; Wang 2012; Helbich, Hagenauer, et al. 2013; Leitner 2013), criminal geographic profiling (CGP; Rossmo 2000) has greatly aided such investigations. This investigative tool allows researchers to analyze a series of known crime locations and predict the most likely residence of a serial offender (Canter 2009). CGP methods are, however, limited. Although they are meant to prioritize offender searches, attempts to evaluate their efficacy have used data from already-solved crimes that involve known offender residences (Paulsen 2006). Additionally, the algorithms used by CGP are designed to locate only the offender homes that lie within the designated hunting area and are therefore not practical for location of commuting offenders (Rossmo 2000).

Bennell and Corey (2007) define the behavior of a serial offender as highly structured. This enables CGP systems to produce operationally useful locational profiles of the offenders' anchor points. Despite this structured nature, variations in behavior exist among offenders in different spatial settings. Studies from urbanized Western countries (e.g., Kent and Leitner 2007; Hammond and Youngs 2011; Iwanski et al. 2011; Levine and Lee 2013) and a study from semiurban India (Sarangi and Youngs 2006) have identified such variations. These studies, modeling offender travel, commonly discover that variations in the urban fabric (e.g., city vs. suburb) can result in CGP performance differences and that limitations in analyzed data (e.g., data size, crime types) affect adequate assessment of CGP.

Until now, CGP has been a popular method for modeling offender travel. Nevertheless, few studies (e.g., Snook, Taylor, and Bennell 2005; Kent, Leitner, and Curtis 2006; Canter et al. 2013) have evaluated CGP accuracy, and all have analyzed data from Western cities. Because of the lack of reported data from less developed non-Western nations, there has emerged an imbalance in our understanding of global CGP performance. It remains unclear whether areas with completely dissimilar socioeconomic and

The Professional Geographer, XX(X) XXXX, pages 1–11
Initial submission, May 2013; revised submissions, July and August 2013; final acceptance, August 2013.
Published by Taylor & Francis Group, LLC.

demographic landscapes (e.g., African countries) can benefit from the studies conducted on CGP. Previous studies also employ limited offense types and series (e.g., Rossmo 2000; Snook, Taylor, and Bennell 2004), simulated or biased sampling (e.g., Snook, Canter, and Craig 2002), or inconsistently use identical study areas (e.g., Levine 2002; Paulsen 2006). A clear establishment of associations between environmental variations and offender travel patterns, through sufficient analysis, can be instrumental in selecting a suitable CGP strategy.

This article seeks to redress the identified research gaps and aims to identify alternative travel characteristics of serial offenders in a non-Western suburban setting. Using the case study of Dandora, Kenya, the study comparatively analyzes several raster-based CGP strategies, alongside basic centrographic strategies. The research addresses the following questions:

1. To what extent can CGP efficiently and accurately predict offender residences through analysis of crime distributions?
2. What possibilities exist for the empirical application of CGP methods to focus on criminal investigations in differing environments?

Theoretical Basis and Current Progress in Geoprofiling Studies

CGP methods are developed around theories that model offender travel characteristics and the factors influencing target choices. The routine activity theory of environmental criminology (Cohen and Felson 1979), for example, claims that offender travels do not only begin at home but also from other locations (e.g., work, leisure) where motivated offenders and unsupervised targets meet in space and time (Franklin et al. 2012; Mohler and Short 2012). Nevertheless, with the exception of crimes of passion (e.g., assault, murder), the behavior of most serial offenders is seldom impulsive and has been ascribed to the rational choice theory (Chainey and Ratcliffe 2005; Brantingham and Brantingham 2008). This theory recognizes the offender's ability to weigh the costs and benefits of a crime through careful premeditation, resulting in travel that often begins from home and is characterized by a no-offending buffer zone.

Another theory regards crime locations as emergent from an offender's opportunity space. Brantingham and Brantingham (2008) applied the cognitive mapping principle to demonstrate cognitive processes behind patterns identified in crime distributions. Cognitive maps represent specific areas within which an individual feels comfortable and secure. Bernasco and Block (2011) defined an offender's crime distributions as comprising awareness spaces that an offender frequents during routine activities and that expand with increasing knowledge of the area.

A common implication of these theories is that criminals do not unjustifiably expend time and resources in traveling to offend. The "hunting area" for a serial offender's residence or other anchor point, therefore, constitutes a cluster of previously identified crime locations and is defined by a diameter composed of the two furthermost points (Warren, Reboussin, and Hazelwood 1995; Rossmo 2000). Nevertheless, the understanding of specific offender travel patterns and their influencing characteristics (e.g., the source of attractions, the street-network layout) increases the precision of investigative predictions.

Studies from the United Kingdom and North America have consistently shown that serial offenders travel usually not farther than 3.2 km from their homes to their crimes (Warren, Reboussin, and Hazelwood 1995; Snook, Taylor, and Bennell 2004). Hewitt, Beauregard, and Davies (2012) notably discovered that rapists seek out their victims in specific areas and often travel much shorter distances than offenders targeting material attractors. Similarly, Godwin and Canter (1997), in studying the behavior of U.S. serial murderers, determined that offenders travel relatively short distances, between 0.32 km and 0.8 km, to abduct their victims and up to 39 km to dump their remains. Meaney (2004) noted that serial offenders travel longer journeys of up to 8 km in less urbanized environmental settings. This counters the findings of Sarangi and Youngs (2006), who showed that serial burglars in semiurban India make short journeys (approximately 1.6 km) to offend.

The average distance that a criminal travels to offend is influenced by factors such as the offense type, the offending strategy, the environmental design, and the target value (Kent, Leitner, and Curtis 2006; Sarangi and Youngs 2006; Law and Quick 2012). Additionally, commuters have the propensity to travel far to seek their targets, whereas marauders seek attractions near their homes or other anchor points (Meaney 2004). Evaluation of the applicability of CGPs should consider such factors. Although many studies have applied CGP to model offender travel patterns, few studies have attempted to evaluate the quality of the CGP strategies. Table 1 lists several of these studies.

Snook, Canter, and Craig (2002) and Paulsen (2006) concluded that CGP models are in many respects not more accurate than human predictors. Levine (2002) and Kent, Leitner, and Curtis (2006), however, discovered the negative exponential (NE) function built within the Crimestat model (Levine 2010) to be the most accurate, particularly for modeling abodes of property criminals. In contrast, Rossmo (2000) rated the RIGEL model highest for its ability to determine anchor points, observing a 6 percent average cost of searching murderers' abodes.

The studies listed in Table 1, although they provide a crucial overview of the relative usefulness of CGP, also demonstrate limitations in methodology and data selection. Many of these studies have focused on single or limited crime types, involved too few crimes, or used a single data set for several studies. Rossmo (2000) and Kent, Leitner, and Curtis (2006), for example, used series from a Baton Rouge, Louisiana, serial

Table 1 *Existing studies on criminal geographic profiling evaluation*

Study	Series/avg. per series	Crime	Sample area	Applied model	Evaluation measure
Rossmo (2000)	15/11	Murder	Canada	RIGEL	Search cost
Canter et al. (2000)	79	Murder	United States	Dragnet	Search cost
Levine (2002)	50	Four crimes[a]	Baltimore (U.S.)	Crimestat,[b] centrographic statistics	ED
Snook et al. (2002)	10/5	Murder	United States	Dragnet, HP	ED
Snook et al. (2004)	10/3	Murder	Germany	Crimestat[c]	ED
Snook et al. (2005)	16/11	Burglary	UK	Crimestat,[a] centrographic statistics	ED
Paulsen (2006)	25/7	Six crimes[d]	Baltimore (U.S.)	Crimestat,[b] Dragnet, HP	ED, search reduction
Kent et al. (2006)	1/301	Murder	United States	Crimestat[e]	ED
Canter et al. (2013)	63	Burglary	UK	Dragnet	Density

Notes: ED = error distance; HP = human predictors.
[a]This study uses data for street burglary, residential burglary, robbery, and vehicle theft.
[b]Five distance decay functions (the negative exponential (NE), normal, lognormal, linear, and truncated negative exponential [TNE] are used).
[c]Only the NE function is explored.
[d]The offenses used for the study are larceny, arson, auto theft, and commercial, street, and residential burglaries.
[e]The NE and TNE functions are computed.

killer, and Snook, Canter, and Craig (2002), Levine (2002), Paulsen (2006), Leitner and Kent (2009), and Levine and Block (2011) analyzed serial offenses committed in Baltimore County, Maryland. Although the persistent use of similar data might pinpoint certain variations that result from applying different methodologies, a lack of diversity in input data greatly narrows the empirical application of results. Regarding offense types, several studies (see Table 1) have focused on murder. This is a grave issue because CGP systems are principally developed to profile "routine-based" serial offenses, such as burglaries and robberies.

Only a few studies specify how data sampling was conducted. Rossmo (2000) is reported to have selected samples that "were appropriate for the assumptions of the strategy being tested" (Snook, Taylor, and Bennell 2005, 4) and that could therefore not be fully representative of real-life case scenarios. Levine (2002) employed a selection that balanced the number of incidents committed by each offender and estimated the offense distribution. Snook, Canter, and Craig (2002, 118) stated that their "samples could be biased in favor of Dragnet," because they included geometries complex to the human subjects employed but manageable with the computer algorithm. Such inappropriate sampling designs have often resulted in biases of CGP output.

Several studies have employed limited crime samples. For instance, Sarangi and Youngs (2006) analyzed some series with two crimes. Paulsen (2006) employed a larger variety of crime types and used various evaluation strategies, but limitations in the participation allowance of his human subjects necessitated the use of twenty-five offense series. There exist no standards for determining the minimum crime sample size, but small samples could hinder the reliability of offense-distribution representation. Previous studies (e.g., Rossmo 2000) have also demonstrated the sample size to be directly proportional to geoprofile accuracy. Observations from analysis using insufficient sample size might therefore indicate not the accuracy of applied methods but the limitations within geoprofiles.

Finally, all of the reviewed studies have analyzed data from the urbanized areas of developed countries. These areas share common characteristics (e.g., street-network development, ownership of motor vehicles) that heavily influence offender travel behavior. In contrast, the limited infrastructure of the study area might create the need to seek good crime targets, resulting in far distances, or causing travel constraints, limiting the offenses' span. Levine (2002), Kent, Leitner, and Curtis (2006), and Helbich and Arsanjani (forthcoming) cited environmental and sociodemographic characteristics as principal determinants of the appropriateness of a crime-modeling application. Given the specific conditions exhibited in different urban environments, a variation of analyzed data is necessary in assessing the usefulness of CGP approaches.

To summarize, many variations exist in offender travel behaviors in different environments that can affect the performance of CGP. Building on the body of research hitherto examined, this study models offender travel to determine suitable models. It aims to resolve the discussed limitations through an empirical evaluation of CGP strategies using an African case study.

Study Site and Data

Dandora (1.25° S, 36.9° E) is a small eastern suburb of Nairobi, Kenya. It borders two informal settlements and the city's dump site, and its serial criminality is fueled by overpopulation, high unemployment, and drug consumption among the youth (Blacksmith Institute 2007; Kummsa and Mwangi 2011). The study area of Dandora and its surroundings, stretching 12.5 km east to west, has an estimated population of 507,000 inhabitants (Kenya Open Data 2011). The area has an underdeveloped and constantly changing road network; its infrastructure, therefore, contrasts with the static infrastructure of urbanized areas cited in existing offender travel literature (e.g., Kent, Leitner, and Curtis 2006; Levine 2006; Van Daele and Beken 2010; Canter et al. 2013). Figure 1 shows its street layout,

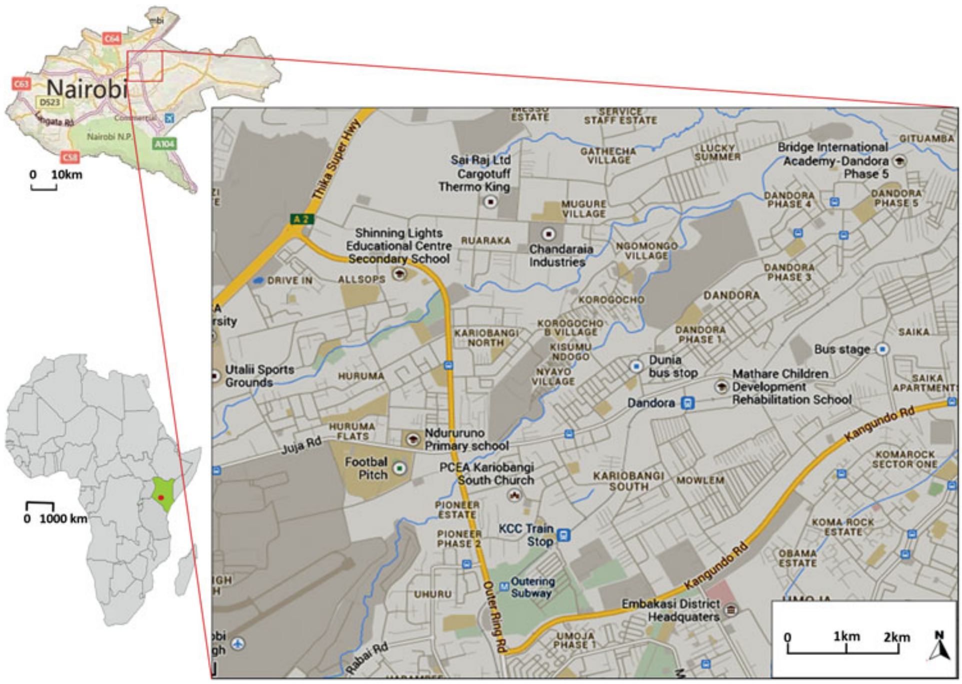


Figure 1 *Dandora area of Nairobi, Kenya (Courtesy of Google Maps™). (Color figure available online.)*

which is primarily composed of footpaths. The built environment constitutes 80 percent of the total area and includes substandard residential structures. There is no clear-cut distinction between commercial and residential areas, a factor directly influencing offender travel. With few personal automobiles, residents rely on walking and short-distance public transportation (Kenya Open Data 2011). Decreased mobility reduces the possibilities of traveling long-range distances yet introduces freedom in travel direction, as an automobile is confined to the structure of the road network.

An initial data set of 1,493 solved offenses committed between 2008 and 2011 was collected from the Dandora Police station. All crimes committed within Dandora and its surrounding areas are reported at this station and copied to the Buruburu main station, with summary reports periodically updating the central police headquarters. Although manual bookkeeping is practiced, crime reports are reliable and detailed.

The data included specific offender and crime descriptions, which included physical addresses where offenses were executed and fingerprint information linking a particular offender to his or her previous crimes. Personal information was removed after address geocoding. In geocoding, the authors collected waypoints from crime locations and offender residences using handheld Global Positioning System (GPS) receivers and processed the output with a geographic information system (GIS). Because many postal codes in Kenya are independent of the geographic ordering of buildings, location identification relied on further descriptions of a place and the help of law enforcers and residents. The total area covered by crimes and offender residences is 107.6 km^2. It must be noted that crime series of offenders who lived outside the study area (commuters) were not incorporated into the analysis. After cleaning, the final data set contained 1,422 solved offenses and included ten offense types that ranged from nonviolent (e.g., burglary) to violent (e.g., murder) crimes. With 346 crime series, each with four crimes on average, the study obtains a balance between data and series size.

Methods

The CGP evaluation procedure applies spatial distribution analysis and raster-based CGP analysis.

Descriptive and Centrographic Statistics

The study first computes the mean distances to crime from offenders' homes and comparatively analyzes this information according to offense types. The results test the broad hypothesis that offender anchor points predominantly determine offense locations. Subsequently, observation is made of whether differences in offender travel behavior within the semiurban environment can influence the performance of CGP, the utility in the urbanized setting of which is confirmed. For each crime distribution, the study identifies the distance between the two furthest crime locations and uses this output to compute the expanse covered by crime locations. This analysis models the extent of criminal travel and also obtains the search area to be used for computing the search cost, a measure of CGP accuracy.

The mean center, median center, and center of minimum distance (CMD) are used in generating spatial estimates of offender abodes and to provide a criterion for comparing raster-based CGP strategies. These centrographic statistics have been used by Levine (2002), Snook, Taylor, and Bennell (2005), and Paulsen (2006) to predict offender residences with low average error. Their technical description is provided by Levine (2010). Nevertheless, distributional statistics are limited. They only provide a central tendency to a crime distribution, without considering other influential factors, such as offender travel propensity. The result is a generalization of a point distribution to a single location, rather than the estimation of an area or probability surface (for alternatives, see also Helbich and Leitner [2012]). A solution to this limitation is raster-based journey-to-crime approaches.

Journey to Crime Modeling

Raster-based journey-to-crime models output surface entities estimating offenders' residences. This research uses the Crimestat CGP model (Levine 2010) due to its flexible analysis options and its availability. Crimestat primarily used five functions to characterize distance decay: (1) the normal function estimates the peak likelihood of offending at some optimal distance from the offender's home; (2) the log-normal function performs like the normal but is skewed to the left or right and can potentially model a rapid offense increment near the criminal's home; (3) the NE models the highest likelihood of offending to be very near the offender's home (almost with no buffer zone) and to constantly decline with distance from home to zero likelihood; (4) the linear function provides a control, modeling the highest offending likelihood near the offender's residence, with a constant decline to zero likelihood; and (5) the truncated negative exponential (TNE) fuses the linear and NE functions. A positive linear function is applied, starting at the offender's home and quickly increasing to its peak. Before leveling off, the function adopts an NE characteristic that quickly declines with distance from home. As an alternative to using prespecified parameters for each function, a study area–specific calibration data set, defined with data of known offender homes and offense locations, can be used. Recently, probabilistic journey-to-crime strategies have also been extended toward Bayesian modeling (Canter 2009; Leitner and Kent 2009; Levine and Block 2011). Due to the data requirements associated with calibration and Bayesian modeling, which are unavailable for the case study, analysis uses the five functions just listed. These functions assume different hypotheses and suppositions regarding the factors influencing offending distance and offer unique behavioral descriptions for the decay in travel distance. Using crime locations, each of the functions is used to estimate offender residences and generate probability surfaces. Every surface is made up of a rectangular grid that covers the offender search area. The cell size is based on specific urban morphology, with each grid cell depicting a specific likelihood of offender residence.

Model Accuracy Assessment

In comparing the relative usefulness of CGP strategies for predicting offender residence, the study employs two basic measures: error distance (ED) and search cost. ED measures the distance between the criminal's actual home and his predicted home (the grid cell with the highest likelihood estimate), thereby evaluating a geoprofile's accuracy and utility. Existing studies (e.g., Levine 2002; Snook, Canter, and Craig 2002; Snook, Taylor, and Bennell 2004, 2005; Paulsen 2006) show this measure to be fairly effective. Different offense types are also separately analyzed to model variations in accuracy. Additionally, the individual EDs are summarized with the root mean squared error (RMSE), a widely used model quality assessment measure (e.g., Helbich, Brunauer et al. 2013). A smaller RMSE value denotes less prediction error and hence a more accurate geoprofile. For each CGP strategy, the study also analyzes offenders' homes that lie within equal distance intervals of the cell with the highest probability.

A search-cost measure computes the percentage of the offense domain to be searched until the offender's residence is found (Paulsen 2006; Hammond and Youngs 2011). Because a geoprofile should reduce the offender search area, a smaller ratio denotes a more accurate geoprofile and an increased ability to prioritize the search area. Rossmo (2000) obtained the search cost as an offense domain where investigation should be focused. Canter et al. (2013) and Snook, Taylor, and Bennell (2005) used a search-cost variation that computes the percentage of overlaid grid cells to be searched to find the offender's home. For geoprofiles created by raster-based journey-to-crime functions, this study obtains the search cost by computing the percentage of the highest profile area automatically generated from each series, over the total hunting area. Centrographic distribution strategies cannot automatically produce a high-profile area. Therefore, a circular area, with a radius that is the distance from the point of peak probability to the offender residence, is used.

Results

Descriptive Statistics

The study analyzes the crime data (n = 1,422) to determine the average number of crimes and the distance traveled to offend. Specific information regarding the different offense types is also retrieved. Table 2 shows that the average distance to offend is 2 km, or approximately one quarter of the study area. This distance is almost akin to the 1.3-km distance traveled by burglars in semiurban India (Sarangi and Youngs 2006) but much shorter than travels observed in Western cities (e.g., Kent, Leitner, and Curtis 2006; Paulsen 2006). Short distances traveled to offend are due to the lack of substantial travel by offenders into the more urbanized Nairobi central business district. Such distinct travel patterns can aid in the selection of CGP models.

Classifying travel groups by offense type reveals that vehicle thieves travel the longest distances (51 percent farther than the average distance) to offend. This could

Table 2 Description of the data (standard deviations in parentheses)

Offenses	Number of series	Number of crimes	Average per series	Average travel	Average Longest travel	Average maximum distance between crimes	Hunting area
Total sample	346	1,422	4.15	2.08 km (1.53)	3.23 km (2.00)	5.18 km (2.50)	28.46 km^2 (27.11)
Narcotics	94	352	3.77	1.70 km (1.25)	2.45 km (1.53)	4.41 km (2.19)	19.05 km^2 (17.48)
Stealing	52	187	3.63	2.02 km (1.39)	3.19 km (1.38)	5.21 km (2.80)	27.39 km^2 (36.71)
Burglary	54	238	4.41	2.25 km (1.66)	3.44 km (2.33)	5.52 km (2.39)	28.36 km^2 (29.41)
Robbery	19	92	4.84	2.16 km (1.81)	3.96 km (2.25)	5.24 km (2.72)	27.01 km^2 (27.57)
RWV	54	214	3.96	2.86 km (1.52)	3.71 km (1.84)	5.57 km (2.04)	27.58 km^2 (18.20)
Vehicle theft	22	74	3.36	3.16 km (2.27)	4.61 km (3.03)	6.84 km (2.52)	43.76 km^2 (47.37)
Rape	16	79	4.94	1.89 km (1.56)	2.85 km (1.87)	4.58 km (3.09)	23.56 km^2 (33.04)
Defilement	20	80	4.00	1.90 km (1.56)	3.19 km (1.92)	4.61 km (2.43)	21.06 km^2 (18.76)
Abduction	7	29	4.14	2.38 km (1.50)	3.40 km (1.71)	6.34 km (2.52)	35.82 km^2 (23.32)
Murder	17	77	4.53	1.92 km (1.55)	3.60 km (2.27)	5.77 km (2.50)	30.76 km^2 (22.56)

Note: RVW = robbery with violence.

be due to assurance of increased mobility or an effort to avoid arrest, as a motor vehicle is a valuable property. Standard deviation values depict a lot of variation in general travel distances by offenders, particularly among rapists and vehicle thieves. The greater distances traveled by rapists might be due to the lack of red-light districts near the study site, whereas for the vehicle thieves this might be caused by the lack of specialized commercial or residential areas. Armed robbers (with an offense classified as robbery with violence, or RWV) travel 14.3 percent farther on average than nonarmed robbers. Generally, armed robbers seek higher returns, and this could explain their willingness to travel greater distances. Drug dealers travel the least distance, with an average of 1.7 km and a minimum of 70 m (i.e., within the same building). This is a significant finding, given the large amount of data for this offense type. The need to deal with people and establish regular customers for their trade could explain this short distance.

Figure 2 depicts the offending frequency with increasing distance from the offender's home. The study showcases specific groups based on 2-km distance intervals.

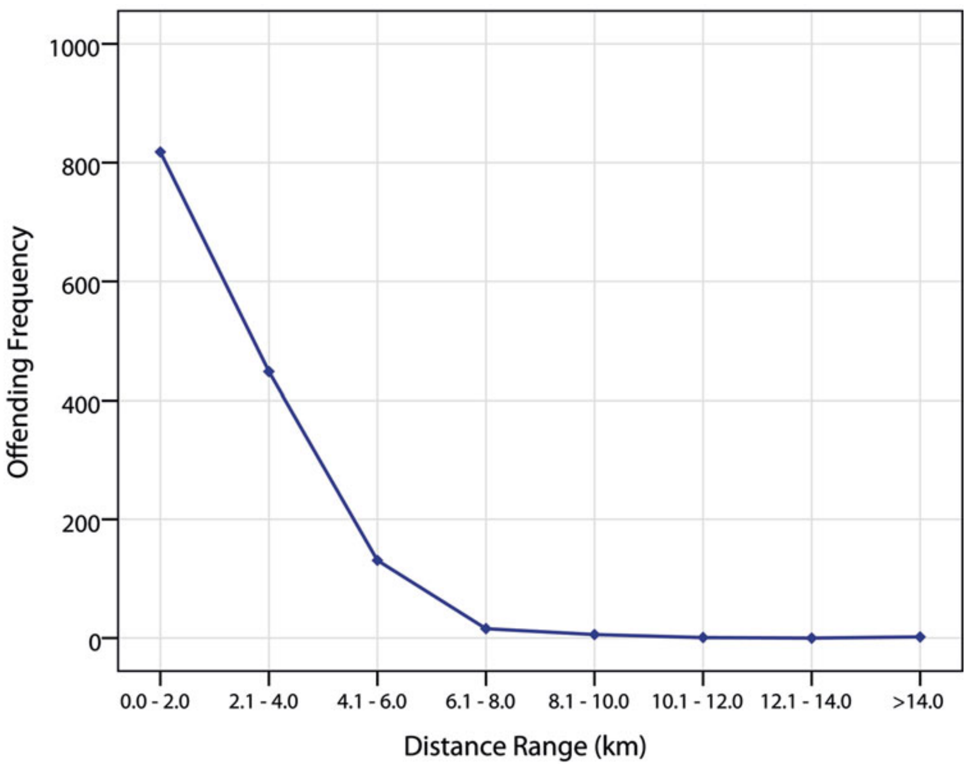


Figure 2 *Offending frequency by increasing distance from home (n = 1,422). (Color figure available online.)*

Table 3 *Results of profiling strategies*

	Spatial centrographic strategies			Probability distance strategies				
Category	***M***	**CMD**	**Median**	**NE**	**Normal**	**Log-normal**	**Linear**	**TNE**
Narcotics	0.76	0.63	0.74	**0.60**	0.96	1.29	0.61	0.77
Stealing	0.98	**0.90**	0.94	0.91	1.23	1.84	1.04	1.00
Burglary	1.25	1.22	1.29	**1.22**	1.36	2.07	1.21	1.35
Robbery	1.51	1.05	1.27	**0.92**	1.98	2.34	**0.92**	1.22
RWV	1.16	0.89	1.06	**0.80**	1.38	1.78	0.84	1.06
Vehicle theft	1.58	1.30	1.66	**1.09**	1.55	2.70	1.09	1.28
Rape	0.87	0.83	0.91	**0.77**	1.13	1.60	0.89	0.82
Defilement	**1.02**	1.15	1.21	1.26	1.30	1.49	1.20	1.27
Abduction	1.64	1.72	1.77	1.72	**1.54**	2.45	1.72	1.65
Murder	1.10	**1.01**	1.05	1.04	1.50	1.09	1.04	1.12
Offenses *M*	1.05	0.94	1.04	**0.90**	1.25	1.80	0.92	1.05
RMSE	1.38	1.27	1.37	**1.23**	1.49	2.41	1.27	1.33

Note: Values depicting the most accurate methods are shown in bold. CMD = center of minimum distance; NE = negative exponential; TNE = truncated negative exponential; RWV = robbery with violence; RMSE = root mean squared error.

Most of the criminal activities are concentrated within a distance shorter than 1 km from home, with this frequency dropping off quickly to slightly above 9 percent at 4 km from home and to almost no activity beyond 8 km. Offender travel behavior in the study area conforms to distance decay, where the area nearest to the offender's residence is characterized by a lot of criminal activities that gradually decrease with offending distance.

Evaluation of CGP Models

Table 3 contains the results of the ED evaluation performed for eight CGP strategies. The grid-cell size of 350 m reflects the study area's urban morphology, where cells roughly represent a housing estate. The Friedman nonparametric analysis of variance test compares the ED values to analyze significant overall differences (Hsu 1996). From this analysis, substantial performance differences among the strategies are evident ($p < 0.001$).

The NE function, in particular, performs with the least amount of error, both on average and for 50 percent of offense types. The RMSE, however, depicts this function to have a cumulative error of 1.23 km, or 37 percent less accurate than the average ED. This illustrates a high variation in distance predictions among the geoprofiles. The CMD performs only slightly worse than the NE, with about 40 m difference in error. This finding is similar to that of Levine (2002), Snook, Taylor, and Bennell (2005), and Paulsen (2006) and can be attributed to the high capability by centrographic statistics for measuring central tendency. The NE function performs well for many offense types but is outperformed by simple centrographic strategies for defilement, abduction, and murder, offenses involving human targets.

On $p < 0.001$, a paired t test shows no significant mean differences between the best strategy (NE) and second best strategy (linear CGP), both of which are raster-based strategies. There is, however, a highly significant difference between the NE and the worst-performing strategy (the log-normal raster-based CGP) on $p < 0.001$. To identify specific changes in performance levels, the study employs a post hoc pair-wise multiple comparisons test using Nemenyi's procedure with Bonferroni correction (Hollander and Wolfe 1997). This test assigns ranking values to all ED measurements and then merges the data and ranks each measurement from the lowest to the highest (see Table 4). All sequences of ties are assigned an average rank. The precision is measured from A (most precise) to D (least precise).

To evaluate CGP efficiency, Table 5 displays offender residences that lie within 1-km intervals from the highest probability point. As can be observed, 59 percent of offender residences have been predicted within a 1-km search radius, and only 2 percent have their residences predicted outside a 4-km radius. The CMD and two raster-based estimation strategies—the NE and linear functions—perform with the same level of efficiency, constituting almost 50 percent of overall method performance. With the exception of the normal and log-normal functions, however, the other strategies do not perform substantially worse. All of the methods, except the normal and log-normal functions, perform with the same efficiency at a 2-km to 3-km range.

Table 4 *Multiple two-tailed error distance comparison with Nemenyi's test of ranked data*

Profiling strategy	**Sum of ranks**	**Mean of ranks (Sum of ranks/*n*)**	**Rank groups**
NE	1,253.5	3.6	**A**
Linear	1,281.5	3.7	**A**
CMD	1,330.0	3.8	**A**
Mean center	1,545.5	4.5	B
Median center	1,545.5	4.5	B
TNE	1,594.0	4.6	B
Normal	1,799.5	5.2	C
Log-normal	2,106.5	6.1	D

Note: Values depicting the most accurate methods are shown in bold. NE = negative exponential; CMD = center of minimal distance; TNE = truncated negative exponential.

Table 5 *Method performance by strategy*

	% of strategy performance by distance intervals (km)					
Profiling strategy	**0–1[a]**	**> 1–2[b]**	**> 2–3[c]**	**> 3–4[d]**	**> 4–5[e]**	**> 5[f]**
Mean center	13%	14%	11%	10%	9%	0%
CMD	14%	13%	10%	9%	9%	5%
Median center	13%	13%	10%	13%	9%	5%
NE	**15%**	14%	11%	8%	6%	0%
Normal	9%	16%	15%	8%	9%	5%
Log-normal	9%	9%	21%	30%	47%	81%
Linear	14%	9%	11%	10%	6%	5%
TNE	13%	12%	12%	13%	3%	0%

Note: Values depicting the most accurate methods are shown in bold. CMD = center of minimal distance; NE = negative exponential; TNE = truncated negative exponential.
[a]$n = 204$.
[b]$n = 90$.
[c]$n = 34$.
[d]$n = 12$.
[e]$n = 4$.
[f]$n = 3$.

Table 6 demonstrates evaluation of the CGP strategies using search cost. A generalized t test identifies significant differences in accuracy among the geoprofiles. Based on these differences, a Dunn's post hoc test (Hsu 1996) classifies the strategies into distinct groups.

Characterized by the lowest values, results from the search-cost measure demonstrate the NE CGP function to be the most accurate in offender residence prediction. A standard deviation measure depicts the NE as most consistent in generating the high-profile area and producing accurate geoprofiles. The normal and the TNE CGPs have the least consistent search costs. A Dunn's test ranks the NE function in the top-most A group. The linear function and the CMD are, however, on average only 1 percent less accurate. Although the normal CGP function computes the largest areas of offender-home probability, it is the least accurate in prioritizing offender search. Both observable and statistical results show the log-normal and normal functions to perform much worse than all other strategies.

Figure 3 displays exemplary performance of the NE journey-to-crime estimation function using a probability surface generated from a ten-crime distribution. The output is displayed alongside the results of centrographic statistics. The highlighted cell portrays the highest likelihood of offender residence. Using an ED measure, the NE function predicts the offender's residence at 0.34 km, or just two blocks from the actual residence. This performance is four to five times more accurate than the centrographic statistics.

Discussion and Conclusion

Although CGP has performed with success, the independent empirical evaluation of CGP strategies with data from developing-world areas has until now been nonexistent. Establishing differences between travel patterns of the offenders in the study area and those in Western cities was important in assessing the applicability of the data for CGP. Although the study shows differing travel behaviors, these differences do not affect the ability of CGP to model offender travel. Overall, it has been discovered that about 75 percent of criminal activities occur within 2 km of the offender's home. These short distances partly explain the good performance of the NE CGP, as this strategy performs best where the no-offending buffer zone is small.

The ineffective performance of the log-normal and the normal CGP functions for many of the crime series raises concerns about their suitability for modeling offender travel. A possible explanation is their increased performance with increasing distance from home, a condition that is untrue for most trips observed. The inability of a single CGP strategy to consistently perform with excellence also necessitates either the development of an optimal CGP strategy through testing with widely varied data or the inclusion of several probability-estimating options into a single CGP package, to cater for changing offense variables.

The assessment of results elicits further vital information with respect to the applicability of CGP methods in estimating offender residences. We

Table 6 *Mean search costs (standard deviations in brackets)*

				Dunn's post hoc procedure[a]		
Profiling strategy	**Highest probability area**	**Search cost**	**One-sample t test (observed value)***	**Ranks sum**	**Ranks mean**	**Group rank**
NE	2.78 km^2 (3.89)	**10.3%** (09.8%)	19.67	1,273	3.68	A
CMD	3.08 km^2 (4.67)	11.2% (10.4%)	19.94	1,320	3.82	A & B
Linear	3.06 km^2 (5.25)	11.5% (14.3%)	14.96	1,346	3.89	A & B
Median center	3.74 km^2 (5.77)	13.9% (12.7%)	20.32	1,498	4.33	B & C
TNE	3.51 km^2 (5.74)	14.1% (15.0%)	17.43	1,587	4.59	C
Mean center	4.21 km^2 (6.18)	14.3% (10.4%)	25.61	1,597	4.62	C
Log-normal	3.60 km^2 (5.15)	14.5% (13.9%)	19.47	1,730	5.00	D
Normal	6.94 km^2 (9.01)	32.0% (28.6%)	20.82	2,150	6.08	D

Note: Values depicting the most accurate methods are shown in bold. NE = negative exponential; CMD = center of minimal distance; TNE = truncated negative exponential.
[a]Group ranks: A (most accurate) to D (least accurate).
*Significant at $p < 0.001$.

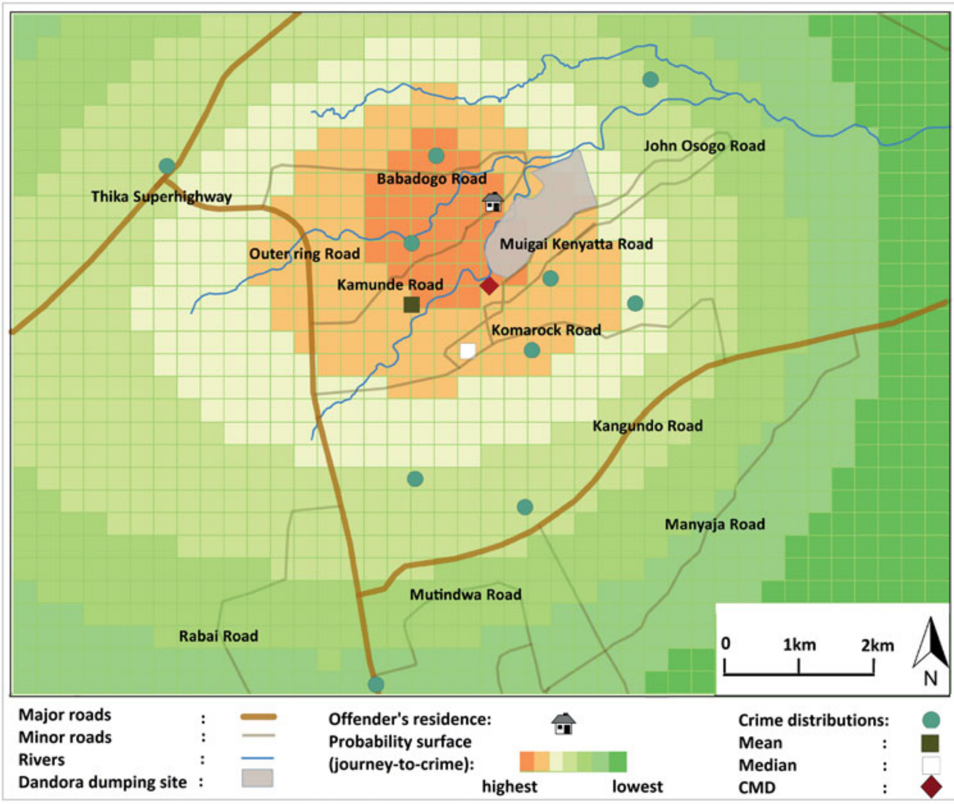


Figure 3 *Journey to crime analysis for a sample crime series. CMD = center of minimum distance. (Color figure available online.)*

discover that the performance of CGP in the developing-world semiurban environment is similar to that of developed cities. Due to similarities in infrastructural characteristics between the study area and other areas in developing countries, such as many African and Latin American towns, investigators in these areas could benefit from the results of this assessment when selecting appropriate CGP strategies. Another observation is that offender travel probability estimation using the raster-based NE function leads to accurate offender-residence determination. The high performance of the NE through assessment with both ED and search-cost measures, as consistent with findings by Levine (2002), Kent, Leitner, and Curtis (2006), and Paulsen (2006), makes this strategy ideal to investigators. Additionally, and corresponding with the findings of Snook, Canter, and Craig (2002), Snook, Taylor, and Bennell (2005), and Paulsen (2006), this study has found that simple spatial statistics, and specifically the CMD, are useful in predicting an offender's residence with almost the same accuracy and efficiency as raster-based CGP strategies. Considering the simplicity of this spatial distribution approach, the CMD's good performance sheds doubt on the benefits of the more complex raster-based CGP to crime investigators, especially if implementing raster-based systems requires purchase and training.

One could argue that small differences in linear computation measures become profound when these measures are converted into search areas, as offender searches seldom follow a linear manner. The inability of spatial centrographic statistics to automatically generate the high-profile area also limits their ability to efficiently compute the search cost, a highly effective measure of geoprofile accuracy. Raster-based CGPs have increased utility because geoprofiles determine search areas, and the linear nature of the ED measure, with which spatial statistics perform best, can underestimate error at an ever-increasing rate and with increasing distance from offender residence. Finally, the study area, with its high population density and subsequent pool of suspects, has complex search prioritization. Even slight improvements in performance measures therefore substantially determine the effectiveness of a strategy, warranting its use.

This research has demonstrated the empirical applicability of CGP. While providing great insight, the research has limitations. First, the study data are exclusively extracted from solved crimes. Given that many crimes in the area remain unsolved, the results therefore lack essential predictive factors available in crime data from ongoing investigations. Another possible consequence might be a bias of the analysis toward the travel behavior of apprehended offenders and against that of offenders from unsolved crimes. Second, the study data is geographically restricted to a case study, and differences in spatial and environmental designs of other developing world areas, which might also affect offender travel, have not been explored. Third, the study analysis assumes all offenders begin their travels at home. These travels could have begun from their work, leisure, or other anchor point. Finally, offenders select targets based on a variety of reasons, with associated travel distance being only one factor. Continued research using data from the developing world, and consideration of all offender anchor points during analysis, can point out variations existent in offender travel behavior and provide further assessment of CGP utility. Furthermore, to improve the geoprofile quality, analysis should also consider other crime-influencing factors, such as terrain variations, occurrence time and season, elapsed duration from encounter to disposal, interval between offenses, the offender's gender and age, and the offender's social and economic status before and during crimes, as well as the importance of these factors to crime-site selection. Finally, future research can benefit from using crime data from ongoing investigations to cater to changes through time. ■

Funding

The first author would like to express her gratitude to the law enforcement in Dandora and to Alexander Zipf (University of Heidelberg, Germany) for their support, as well as to the German Academic Exchange Service (DAAD) for funding this research. The authors would also like to acknowledge the constructive comments and feedback from the reviewers.

Literature Cited

Bennell, C., and S. Corey. 2007. Geographic profiling of terrorist attacks. In *Criminal profiling: international perspectives in theory, research, and practice*, ed. R. N. Kocsis, 189–203. New York: Springer.

Bernasco, W., and R. Block. 2011. Robberies in Chicago: A block-level analysis of the influence of crime generators, crime attractors and offender anchor points. *Journal of Research in Crime and Delinquency* 48 (1): 33–57.

Blacksmith Institute. 2007. The world's worst polluted places: The top ten of the Dirty Thirty, New York. http://www.blacksmithinstitute.org (last accessed 19 March 2013).

Brantingham, P. J., and P. L. Brantingham. 2008. Crime pattern theory. In *Environmental criminology and crime analysis*, ed. R. Wortley and L. Mazerolle, 78–93. Devon, UK: Willan.

Canter, D. 2009. Developments in geographic offender profiling: Commentary on Bayesian journey to crime modeling. *Journal of Investigative Psychology and Offender Profiling* 6 (3): 161–66.

Canter, D., T. Coffey, M. Huntley, and C. Missen. 2000. Predicting serial killers' home using a decision support system. *Journal of Quantitative Criminology* 16 (4): 457–78.

Canter, D., L. Hammond, D. Youngs, and P. Juszczak. 2013. The efficacy of ideographic models for geographic offender profiling. *Journal of Quantitative Criminology* 29 (3): 423–46.

Chainey, S., and J. H. Ratcliffe. 2005. *GIS and crime mapping*. London: Wiley.

Cohen, L. E., and M. Felson. 1979. Social change and crime rate trends: A routine activity approach. *American Sociological Review* 44:588–608.

Franklin, C., T. Franklin, M. Nobles, and G. Kercher. 2012. Assessing the effect of routine activity theory and self control on property, personal and sexual assault victimization. *Journal of Criminal Justice and Behavior* 39 (10): 1296–1315.

Godwin, M., and D. Canter. 1997. Encounter and death: The spatial behavior of U.S. serial killers. *International Journal of Police Strategies and Management* 20 (1): 24–38.

Hagenauer, J., M. Helbich, and M. Leitner. 2011. Visualization of crime trajectories with self-organizing maps: A case study on evaluating the impact of hurricanes on spatio-temporal crime hotspots. Paper presented at the International Cartographic Conference 2011, Paris, France.

Hammond, L., and D. Youngs. 2011. Decay functions and criminal spatial processes: Geographical offender profiling of volume crime. *Journal of Investigative Psychology and Offender Profiling* 8 (1): 90–102.

Helbich, M., and J. Arsanji. Forthcoming. Spatial eigenvector filtering for spatiotemporal crime mapping and spatial crime analysis. *Cartography and Geographic Information Science*.

Helbich, M., W. Brunauer, J. Hagenauer, and M. Leitner. 2013. Data-driven regionalization of housing markets. *Annals of the Association of American Geographers* 103 (4): 871–89.

Helbich, M., J. Hagenauer, M. Leitner, and R. Edwards. 2013. Exploration of unstructured narrative crime reports: An unsupervised neural network and point pattern analysis approach. *Cartography and Geographic Information Science* 40 (4): 326–36.

Helbich, M., and M. Leitner. 2012. Evaluation of spatial cluster detection algorithms for crime locations. In *Challenges at the interface of data analysis, computer science, and optimization. Studies in classification, data analysis, and knowledge organization*, ed. W. Gaul, A. Geyer-Schulz, L. Schmidt-Thieme, and J. Kunze, 193–201. Berlin: Springer.

Hewitt, A., E. Beauregard, and G. Davies. 2012. Catch and release: Predicting encounter and victim release location choice in serial rape events. *International Journal of Police Strategies and Management* 35 (4): 835–56.

Hollander, M., and D. A. Wolfe. 1997. *Nonparametric statistical methods*. New York: Wiley.

Hsu, J. C. 1996. *Multiple comparisons: Theory and methods*. New York: Chapman and Hall.

Iwanski, N., R. Frank, V. Dabbaghian, A. Reid, and P. J. Brantingham. 2011. *Analyzing an offender's journey to crime: A criminal movement model (CriMM)*. Washington, DC: Institute of Electrical and Electronics Engineers Computer Society.

Kent, J., and M. Leitner. 2007. Efficacy of standard deviation ellipses in the application of criminal geographic profiling. *Journal of Investigative Psychology and Offender Profiling* 4 (3): 147–65.

Kent, J., M. Leitner, and A. Curtis. 2006. Evaluating the usefulness of functional distance measures when calibrating journey-to-crime distance decay functions. *Journal of Computers, Environment and Urban Systems* 30 (2): 181–200.

Kenya Open Data. 2011. Kenya county fact sheets, Dec 2011. https://opendata.go.ke/Counties/Kenya-County-Fact-Sh eets-Dec-2011/zn6m-25cf (last accessed 26 February 2013).

Kummsa, A., and I. K. Mwangi. 2011. Challenges of sustainable urban development: The case of Umoja 1. In *Eco-city planning: Policies, practice and design*, ed. T. Wong and B. Yuen, 181–97. New York: Springer.

Law, J., and M. Quick. 2012. Exploring links between juvenile offenders and social disorganization at a large map scale: A Bayesian spatial modeling approach. *Journal of Geographic Systems* 15 (1): 89–113.

Leitner, M. 2013. *Crime modeling using geospatial techniques*. Dordrecht, The Netherlands: Springer.

Leitner, M., and M. Helbich. 2011. The impact of hurricanes on crime: A spatio-temporal analysis in the city of Houston, Texas. *Cartography and Geographic Information Science* 38 (2): 214–22.

Leitner, M., and J. Kent. 2009. Bayesian journey-to-crime modeling of single and multiple crime-type series in Baltimore County. *Journal of Investigative Psychology and Offender Profiling* 6 (3): 213–36.

Levine, N. 2002. *Crimestat: A spatial statistics program for the analysis of crime incident locations*. Houston, TX: Ned Levine and Associates and the National Institute of Justice.

———. 2006. Crimestat and the crime mapping program. *Journal of Geographic Analysis* 38 (1): 41–56.

———. 2010. Crimestat3: A spatial statistics program for the analysis of crime incident locations, version 3.3. http://www.icpsr.umich.edu/CrimeStat/ (last accessed 4 April 2013).

Levine, N., and R. Block. 2011. Bayesian journey-to-crime estimation, an improvement in geographic profiling methodology. *The Professional Geographer* 63 (2): 213–29.

Levine, N., and P. Lee. 2013. Journey-to-crime by gender and age group in Manchester, England. In *Crime modeling and mapping using geospatial technologies*, ed. M. Leitner, 145–78. New York: Springer.

Meaney, R. 2004. Commuters and marauders: An examination of the spatial behavior of serial criminals. *Journal of Investigative Psychology and Offender Profiling* 1 (2): 121–37.

Mohler, G. O., and M. B. Short. 2012. Geographic profiling from kinetic models of criminal behavior. *SIAM Journal of Applied Mathematics* 72 (1): 163–80.

Paulsen, D. 2006. Human versus machine: A comparison of the accuracy of geographic profiling methods. *Journal of Investigative Psychology and Offender Profiling* 3 (2): 77–89.

Rossmo, K. 2000. *Geographic profiling*. Boca Raton, FL: CRC Press.

Sarangi, S., and D. Youngs. 2006. Spatial patterns of Indian serial burglars with relevance to geographic profiling. *Journal of Investigative Psychology and Offender Profiling* 3 (2): 105–15.

Snook, B., D. Canter, and B. Craig. 2002. Predicting the home location of serial offenders: A preliminary comparison of the accuracy of human judges with a geographic profiling system. *Journal of Behavioral Sciences and the Law* 20 (1–2): 109–18.

Snook, B., P. J. Taylor, and C. Bennell. 2004. Geographic profiling: The fast, frugal and accurate way. *Applied Cognitive Psychology* 18 (1): 105–21.

———. 2005. Shortcuts to geographic profiling: A reply to Rossmo. *Applied Cognitive Psychology* 19 (5): 655–61.

Van Daele, S., and T. V. Beken. 2010. Journey to crime of itinerant crime groups. *International Journal of Police Strategies and Management* 33 (2): 339–53.

Wang, F. 2012. Why police and policing need GIS: An overview. *Annals of GIS* 18 (3): 159–71.

Warren, J. I., R. Reboussin, and R. R. Hazelwood. 1995. *The geographic and temporal sequencing of serial rape*. Washington, DC: U.S. Government Printing Office.

LUCY MBURU is a doctoral researcher with the GIScience research group, University of Heidelberg, Berliner Straße 48, 69120 Heidelberg, Germany. E-mail: lucy.waruguru@geog.uni-heidelberg.de. She is also a lecturer at the KCA University, 56808–00200, Nairobi, Kenya. Her interests are in geographic information analysis and modeling.

MARCO HELBICH was a postdoctoral researcher at the Institute of Geography, University of Heidelberg, Germany. E-mail: m.helbich@uu.nl. He is an Assistant Professor at the Department of Human Geography and Planning at the Utrecht University in The Netherlands. His interests lie in the areas of urban geography and GIScience. In particular, he combines computational models and GIS-based approaches to investigate research challenges in real estate, crime, and health, among others.