Applied Geography 86 (2017) 262–273

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

What does Gunshot Detection Technology tell us about gun violence?

Yasemin Irvin-Erickson [a, *], Nancy La Vigne [a], Ned Levine [b], Emily Tiry [a], Samuel Bieler [c]

[a] Urban Institute, 2100 M St. NW, Washington, DC, 20037, United States
[b] Ned Levine & Associates, United States
[c] New York University, United States

ARTICLE INFO

Article history:
Received 29 October 2016
Received in revised form
8 May 2017
Accepted 6 June 2017
Available online 21 July 2017

Keywords:
Gunshot Detection Technology
Gun violence
Calls for service
Spatial
Temporal

ABSTRACT

Purpose: Evaluate the sensitivity of Gunshot Detection Technology (GDT) relative to Calls for Service. Existing crime data sources have biases that do not present a complete picture of gun-related crime. GDT may offer a new metric for firearm crimes. However, few studies have assessed the accuracy of GDT and its relationship to other measures of violence.
Methods: GDT and gun crime-related Calls for Service in Washington, DC during 2010 were studied. Using temporal comparisons for month, day of year, weekday, and hour of the day, spatial comparisons on a quarter-mile square grid, and a Poisson-Lognormal-CAR spatial regression model on a combined grid by time period dataset, we examined the sensitivity of GDT activations relative to gunshot-related calls for service.
Results: The results showed that relative GDT sensitivity changed by time and by space. In particular, the relative sensitivity of GDT was much stronger in the evening and at nighttime than in the daytime. In terms of spatial variation, we found that GDT sensitivity decreased with distance from the nearest zone centroid. In addition, there were two small geographic areas in the study area in which the relative GDT sensitivity was lower than expected.
Conclusions: GDT systems identify the frequency and locational accuracy of gunshot incidents, particularly at nighttime. This technology has the potential to improve data collection on gun use and violence and produce more accurate representations if the temporal and distance limitations of the technology are understood. GDT may improve gun detection and, thereby, improve police operations and public support for police.

1. Introduction

This study examines the relationship between Gunshot Detection Technology (GDT) and calls for service relating to firearms discharges. Researchers have long known that information in available statistical measures of crime reveal only a portion of total offenses, with another portion, sometimes referred to as “the dark figure of crime,” going unrepresented (Coleman & Moynihan, 1996; Penney, 2014). GDT offers the possibility of identifying some of that unrepresented crime related to firearms use. However, few studies have assessed the accuracy of GDT and its relationship to other measures of violence.

1.1. Under-reporting of gun violence

Reporting bias, either due to individuals' reluctance to disclose information or failure to recall past crime incidents poses a particularly significant threat to research on gun violence. Public reports (Calls for Service) have been the primary means by which police become aware of gunshots (Mazerolle, Watkins, Rogan, & Frank, 1999). However, gun violence in communities tends to be concentrated within small geographic areas and affect a limited network of people (mostly young males), many of whom have been both victims and perpetrators of illegal gun activity (Braga, 2003; 2007) and who may be socially unconnected and disinclined to report incidents (Fox, 1996; Jones- Brown, 2007; Solis, Portillos, & Brunson, 2009). Lack of reporting may also occur where gunshots are commonplace and perceptions of police response and effectiveness are low (Kidd & Chayet, 1984; Langton, Berzofsky, Krebs, & McDonald, 2012; Mazerolle, Frank, Rogan, & Watkins, 1999) or when there are non-fatal firearm-related crimes such as gun

* Corresponding author.
E-mail address: yasemin.irvin@gmail.com (Y. Irvin-Erickson).

http://dx.doi.org/10.1016/j.apgeog.2017.06.013
0143-6228/

discharges with no victim or gun assaults with minor wounds (Mazerolle, Frank et al., 1999).

To overcome reporting bias researchers generally focus on firearm-related homicides. However, most gunshot-related crime incidents do not result in a homicide (Alavarado & Massey, 2010; Archer & Gartner, 1984; La Free, 2005). These limitations of traditional data on gun violence underscore the need for additional information on gun-related crimes that can provide a more accurate identification.

1.2. Gunshot Detection Technology

Data from GDT offer a new source of information on firearm-related crimes. GDT uses a network of acoustic sensors to identify the sound of a gunshot and relay this information to emergency services personnel (Eng, 2004; Showen, 1997; Siuru, 2007). Beginning in the 1990s increasing numbers of law enforcement agencies began adopting GDT to facilitate improved response to gun violence and by 2014, a total of fifty cities and 267 square miles were covered by just one GDT vendor's system (SST Inc., 2014).

The technology uses the sensors to identify a gunshot sound and triangulate its position (see Fig. 1). Gunshots have a distinctive acoustic signature composed of the sound of the explosion, the muzzle blast, the sound of the firearm's projectiles and, to lesser degree, the mechanical sound of the firearm and vibration from any solid surfaces near the discharge of the weapon (Maher, 2007). A byproduct of the detection process is standardized data recording of the location and time of gunshot with considerable precision. The most recent versions of this technology have also been found to accurately record gunshots under most conditions (Goode, 2012).

1.3. Evaluation of the accuracy of GDT systems

Findings on the accuracy of GDT systems for detecting and triangulating gunshots are mixed but do suggest that it has improved over time. Early deployments detected approximately 81 percent of gunshots fired with the technology being able to triangulate the location of the gunshot in 84 percent of those cases (Mazerolle, Watkins et al., 1999; Watkins, Mazerolle, Rogan, & Frank, 2002). However, a two-city assessment yielded sharply divergent results with a 97 percent success rate identifying gunshot at one of the cities and a failure to identify shots of more than 70 percent at the other (Litch & Orrison, 2011). Another study found that the correlation between activations and reported calls for service differed substantially within Oakland, CA and Washington, DC (Carr & Doleac, 2015a, 2015b). That study found only 12% of the gunshots detected by the GDT system resulted in 911 calls to report gunshots.

2. Methodology

Our study focuses on concentrations of detected gunshots and Calls for Service in Washington, DC. In 2010, the city had 17.3 square miles of GDT coverage which represented about 25% of the total area of the city. The GDT installations were placed in areas with high amounts of citizen-reported gunshots and gun-related crimes. The city also has ample crime data of sufficient quality, collected by the Washington DC Metropolitan Police Department (MPD).

In 2010, Washington, DC had approximately 1241 violent crimes per 100,000 persons, putting its violent crime rate (819 violent crimes per 100,000 persons) substantially above the national average for cities with populations between 500,000 and 999,999. According to official records, in 2010 in the city, 76 percent of homicides, 19 percent of aggravated assaults, and 40 percent of robberies were committed with firearms (FBI, 2011).

2.1. Data sources

The biggest methodological problem in evaluating GDT is that there is not a complete database of gunshot events. Based on the data we have, actual gun-related crimes appear to be fewer than 10% of all gunshot events. Consequently, it was necessary to approximate the number of gunshot events by comparing the GDT detection with Calls for Service for gun-related events.

Three different sources of data were used for this study: 1) Activations of the Metropolitan Police Department (MPD)'s GDT system; 2) Gun-related calls for police service reported to the MPD in 2010 (hereinafter referred to as Calls for Service), and 3) Gun-related crimes reported to the MPD in 2010. These data sources were only collected for the GDT coverage areas. The GDT data was obtained from ShotSpotter, the largest GDT manufacturer (SST, 2016), which divided its coverage within the city into four zones. To ensure better accuracy, a buffer zone of 0.25 miles beyond the boundaries was added for the identification of gunshot incidents.

The data were obtained by accessing a data file made publically available online by the MPD in response to a Freedom of Information request. The data provide the geographic coordinates of the detected gunshot event (within 25 m), the date and time, file of the gunshot, and an indicator of whether the incident involved single or multiple shots. These data provided an initial total of 5745 detected incidents in 2010. Eliminating GDT counts outside the coverage area (the coverage area plus a quarter mile buffer zone beyond) reduced the number to 5437. Calls for Service data were obtained from the MPD. In 2010, there were 6855 Calls for Service related to gunshot incidents within Washington DC. Of these, 6072 (or 89%) occurred within the study area. However, for many gunshot events, there were multiple calls received.

To identify unique calls, that is unique gun events for which one or more calls were received, we selected three time windows of 10 (N = 4251), 20 (N = 3592), and 30 min (N = 1708) and only identified the first call within each window. The aim was to identify a unique set of gunshot events based on one or more persons calling the police for a gunshot sound. When we made comparisons by month, day of year, day of week, hour of day, and spatial variability, the three sub-sets correlated highly with each other.

We chose the 20 min window (N = 3592) as representing a reliable estimate of the number of unique gun events identified by the public.[1] We chose the 20 min window for three reasons. First, considering the number of unique events in 10, 20, and 30 min windows, we decided that a 20 min window was a good break point. The number of GDT events identified was 5,437, which is more frequent than the calls received in any of those windows. The longer the time window, the greater the ratio since the number of unique events identified decreases. Clearly, without having an independent dataset of actual gun shots, we cannot easily test which window is the most accurate.

Second, we did not want to bias the results by either overestimating or underestimating. Having a higher overall ratio makes GDT appear more accurate than it is; the converse would be true for a 10 min window where the ratio was lower (i.e., we would get more hours where the ratio fell below 1.0).

Third, there is a potential 'false negative' problem in taking a longer window. If two separate gun events came within 30 min of each other, then the 30 min window would categorize them as a single event, rather than two. On the other hand, using a 10 min window might create 'false positive', identifying two 'events' which actually was only one. Thus, the 20 min window is a good balance

[1] All three subsets produced virtually identical results on all temporal and spatial tests.

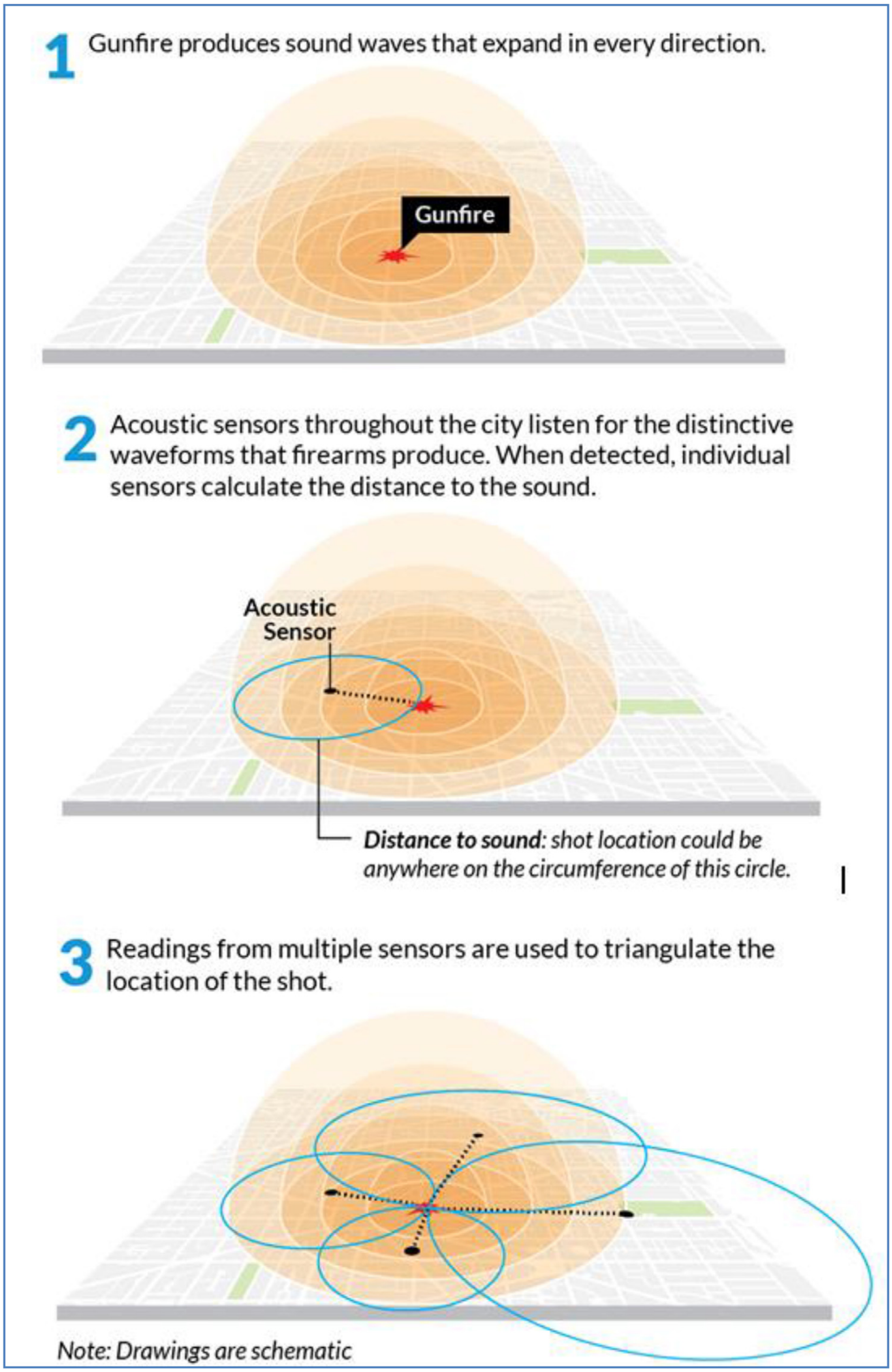

Fig. 1. How the Gunshot Detection Technology works.
Source: Urban Institute.

between the two 'false' conditions.

Crime incident data was limited to gun-involved assaults (N = 396) and gun homicides (N = 60) for the types of crimes most likely to be captured by the GDT system. While robberies are another category of crimes for which a firearm might be used, these were excluded due a lower likelihood of firearm discharge (around 40%; see Cook, Ludwig, Venkatesh, & Braga, 2007; Hales, Lewis, & Silverstone, 2006). Unfortunately, a time stamp for this data set was not available to us. Consequently, we could not use it to match GDT and calls for service but, instead, used it as an indicator of the overall spatial distribution of gun events in the space-time model discussed below.

2.2. Dimensions of evaluation

2.2.1. GDT-to-Calls Ratio

In the public health and medical fields, and increasingly in other fields, the concepts of sensitivity and specificity of a method are used to evaluate it (Emory University, 2016). The *sensitivity* of a test is the proportion of 'true positives' that are correctly identified while the *specificity* of a test is the proportion of 'true negatives' that are correctly identified. The problem with using these concepts is that an independent database of events is needed to evaluate each measure which, unfortunately, we do not have.

Consequently, we used a comparative index of sensitivity, namely the ratio of GDT-to-Calls. This is an indicator of ***Relative GDT sensitivity,*** that is, the ratio of GDT detection relative to gun-related Calls for Service. If GDT is an accurate technology, then it should identify all gun events that humans identify and call into the police. In addition, it should identify gun events for which there are no calls to the police. Therefore, the ratio of GDT-to-Calls should be equal to or greater than 1.0. Overall, the ratio of total GDT events to Calls varied by the time window. For the 10-min window, the ratio was 1.28. For the 20-min window (our default choice), it was 1.52 and for the 30-min window it was 3.18.[2]

2.2.2. Temporal and spatial comparisons

A temporal comparison of the GDT-to-Calls ratio was conducted by month, day of the year, weekday, and hour of the day. In addition, six 4-h time periods were used to group both the GDT and Calls for Service data.

For spatial analysis, both the GDT and Calls for Service data were assigned to 628 grid cells, one quarter mile on a side that covered the coverage area. An additional space-time data base was created by assigning the GDT and calls to grid cells subdivided by the six time periods. That is, for each of six 4-h time periods, the data were assigned to the 628 grid cells yielding a space-time dataset of 3768 grid cells (i.e., 628 grid cells x 6 time periods). Fig. 2 below shows the Washington DC boundary, the four Shotspotter coverage areas, the 1320 foot buffer around those coverage areas, the 628 grid cells, and the centroids of the grid cell.

2.2.3. Distance from the nearest zone

Because the location of the GDT sensors is proprietary to the vendor, we do not know precisely where they are located. As a rough approximation, we took the centroid of each of the four coverage zones on the assumption that the sensors were placed symmetrically within the zone. We then calculated the distance from each of the 628 grid cells to the nearest zone centroid as a proxy for nearness to the acoustical sensors.

2.3. Statistical modeling

To test a simple relationship between GDT and Calls for Service by the time categories, two statistical tests were used: 1) a Pearson 'r' correlation coefficient between the two data series (Kanji, 2006); and 2) a Kolmogorov-Smirnov two-sample test to indicate whether the two samples came from the same distribution (Kanji, 2006). The Pearson 'r' indicates the overall degree of similarity in the pattern of GDT and Calls for Service; it is a parametric test based on the assumption that the two series are approximately normally distributed. The Kolmogorov-Smirnov test is a non-parametric test that looks for the maximum difference in the cumulative proportions of each data series and compares this to a theoretical distribution about likely variation in that maximum difference if the two series were from the same underlying distribution.

For space-time modeling, we tested whether the GDT-to-Calls ratio changed according to time of day and distance from the centroid of the nearest coverage zone. For this test, a Poisson-Lognormal-CAR risk model was used (Levine et al., 2013). Because each of the two data series is a count that is highly skewed in space, the number of GDT events detected relative to the number of Calls for Service is assumed to be Poisson distributed and independent over all segments, and has the form:

$$y_i|\lambda_i \sim Poisson\ (\lambda_i) \tag{1}$$

In turn, the mean of the Poisson, λ_i, is modeled as:

$$\mu_i = \nu_i \lambda_i \tag{2}$$

where ν_i is an *exposure* measure and λ_i is the *rate* (or risk). The exposure variable is the baseline variable to which the number of events is related (Besag, Green, Higdon, & Mengersen, 1995), in our case the number of Calls for Service for each grid cell-time period. The rate is further structured as a Poisson- Lognormal model:

$$\mu_i = \nu_i \lambda_i = \nu_i \exp\left(x_i^K \beta + \varphi_i + \varepsilon_i\right) \tag{3}$$

where *exp* is the base of the natural logarithm (an exponential function), β is a vector of unknown coefficients for the *K* covariates plus an intercept, ϕ_i is a spatial random effect which is estimated using a Conditional Autoregressive (CAR) function (Besag, 1974), and ε_i is the model (residual) error independent of all covariates. The error term, ε_i, is assumed to follow a lognormal distribution with a mean equal to 0 and a variance equal to $\sigma_\varepsilon^{-2} = \tau_\varepsilon \sim Gamma\ (a_\varepsilon,\ b_\varepsilon)$. Note that there is no coefficient for the risk (exposure) variable, ν_i (i.e., it is 1.0). Because of this, it is sometimes called an *offset* variable (or exposure offset).

Normally with a count dependent variable, a negative binomial regression model is used. In that model, the mean is assumed to be Poisson-distributed while the dispersion is assumed to be Gamma distributed (Cameron & Trivedi, 1998). However, when there is a low sample mean, as with these data, the negative binomial function is unreliable (Ma, Kockelman, & Damien, 2008; Park & Lord, 2007). With these data, we have 5437 GDT detections and 3592 calls for service which are then assigned to 3768 grid cells, an average of 1.44 and 0.95 per cell respectively. The Poisson-Gamma model typically becomes reliable when the sample mean exceeds 3 or 4 (Park & Lord, 2007).

The models were estimated with a Markov Chain Monte Carlo (MCMC) algorithm using the *CrimeStat IV* statistical program (Lord, Levine, Park et al., 2013). The MCMC method (sometimes called Bayesian Hierarchical Modeling; Gelman, Carlin, Stern, & Rubin, 2004) is used with complex functions where maximum likelihood estimation does not work. To produce reliable estimates of the parameters, each model was run with 100,000 samples with the first 50,000 samples being discarded (called 'Burn in'). More information on the method can be found in Lynch (2007) and Gelman et al. (2004). The coefficients were tested with approximate t-values and 95% credible intervals.

The overall fit of a model was measured through several statistics: 1) The log-likelihood (smaller negative value is better); 2) The Aikaike Information Criteiron (AIC) and the Bayesian Information Criterion (BIC) which control for degrees of freedom (smaller positive value is better for both statistics); 3) The Mean Absolute Deviation and Mean Squared Predicted Error which test for goodness of fit (smaller is better for both statistics); and 4) The dispersion multiplier which indicates how much the mixed function model differs from a pure Poisson model since skewness in spatial distributions of crime events are more skewed than

[2] A second dimension for evaluation would be the locational accuracy of GDT relative to Calls for Service. However, it is well known that calls are not particularly locationally accurate for gunshots unless the caller is very close to the incident location. Consequently, we did not attempt to evaluate the relative spatial accuracy of GDT.

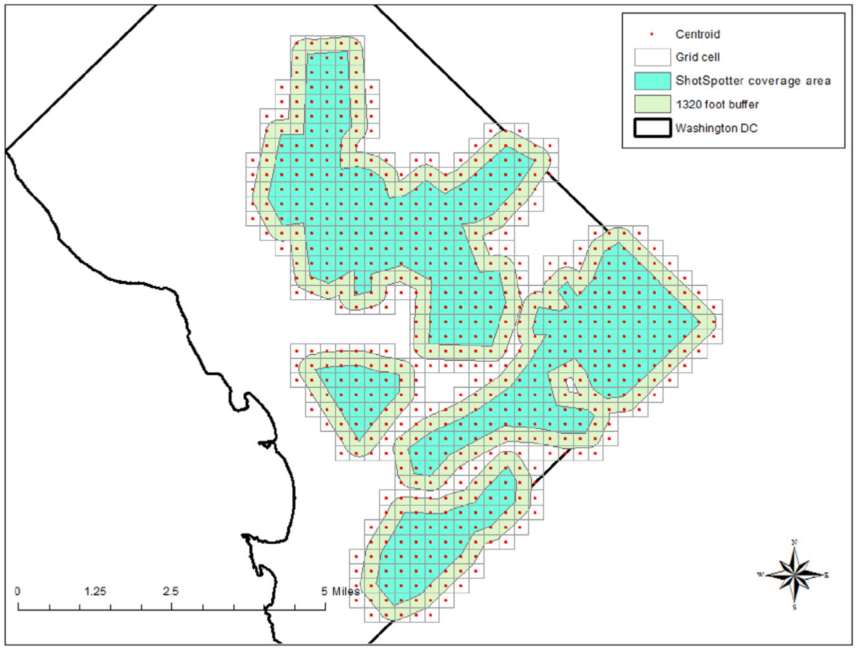

Fig. 2. Shotspotter coverage and grid cells used in analysis.

expected by a Poisson distribution. See Lord, Levine, & Park (2013) for a discussion of these measures.

The effects of the independent variables were measured through coefficients and their statistical significance was tested with an approximate *t*-test as well as by 95% credible intervals (Lord, Levine, Park et al., 2013). The models are presented in linear form with the log of the dependent variable (GDT) being a linear function of the log of the exposure variable (Calls) plus the independent variables with coefficients.

We also ran Poisson-Gamma-CAR models for comparison but found the Poisson-Lognormal-CAR to be better both in terms of model fit as well as producing more realistic predictions.[3] In addition, separate models were run for each time period using the above Poisson-Lognormal-Car formulation.

3. Results

3.1. Temporal comparison

Temporal comparisons were made by month, day of year, weekday, and hour of the day. For month, both the GDT identification and Calls for Service tended to be higher in the summer than at other times with the exception of January where both indices peaked. However, the GDT-to-Calls ratio varied by month from a high of 2.41 in January to a low of 1.12 in August. The Pearson 'r' correlation between the two series was moderately high, 0.69, and was statistically significant ($t = 3.87$; $p \leq 0.01$). However, the Komolgorov-Smirnov two-sample test was also statistically significant at the $p \leq 0.001$ level ($D_{max} = 0.050$; Critical $D_{.001} = 0.042$) indicating that the two series cannot be considered as coming from the same distribution. That is, the GDT-to-Calls ratio is not constant and varies by month.

For day of the year comparisons, the GDT-to-Calls ratio averaged 1.41 per day. However, it varied from a low of 0 (2 days; March 14th and December 14th) to a high of 15.27 (January 1st). In fact, the number of gun events on January 1st was almost three times more frequent than the next highest day (June 26th). The top 20 days of 2010 in terms of frequency of GDT activations were examined and the ratio averaged 4.29 with a median of 2.62. Even excluding January 1st, the average for the next highest 19 days was 3.52. Thus, on days with many GDT detections, the ratio of GDT detected events to Calls for Service was substantially higher than the average for the year.

Over all 365 days, the Pearson 'r' correlation between GDT and Calls for Service was 0.48 and statistically significant ($t = 12.71$; $p \leq 0.001$). The Komolgorov-Smirnov two-sample test was also significant at the $p \leq 0.001$ level ($D_{max} = 0.075$; Critical $D_{.001} = 0.042$). Therefore, even though there was a moderate correlation between the two series, they have to be considered as coming from different underlying distributions (i.e., the GDT-to-Calls ratio is not constant throughout the year).

For weekday comparisons, both the GDT detections and the calls were higher on the weekends than on the weekdays (see Fig. 3 below). The GDT-to-Calls ratio varied from a low of 1.18 (Wednesdays) to a high of 2.16 (Fridays). The weekend days (Friday, Saturday, Sunday) had a higher average (1.78) than weekdays (1.25). The Pearson 'r' correlation between GDT and Calls over the seven days was 0.83 and statistically significant ($t = 4.54$; $p \leq 0.01$). The Komolgorov-Smirnov two-sample test was also statistically significant at the $p \leq 0.001$ level ($D_{max} = 0.089$; Critical $D_{.001} = 0.042$) and, thereby, indicates that the two series were from different underlying distributions.

For hour of the day, both GDT and the calls for service were higher in the evening and the early morning hours than at other hours (see Fig. 4 below). However, the ratio of GDT-to-Calls also followed this pattern, varying from a high of 1.85 between Midnight and 1:00 a.m. to a low of 0.58 between 6:00 a.m. and 6:59 a.m. (see Fig. 5). For the nighttime hours, the ratio indicates that the GDT technology is particularly sensitive, having a ratio that is at least 1.5 times that of the calls for service. But, during the daytime (from around 5:00 a.m. through 5:59 p.m.), the ratio was not different than parity (1.0) and for six individual hours actually fell below 1.0.

[3] The Poisson-Gamma-CAR models produced much too large predictions.

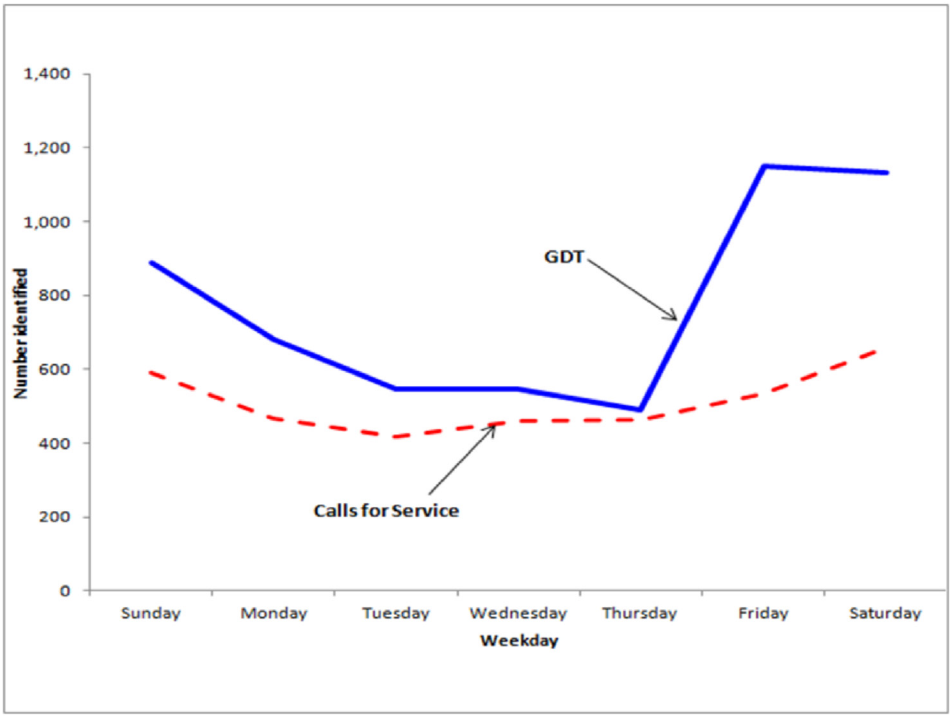

Fig. 3. Weekday comparison of GDT detection and calls for service.

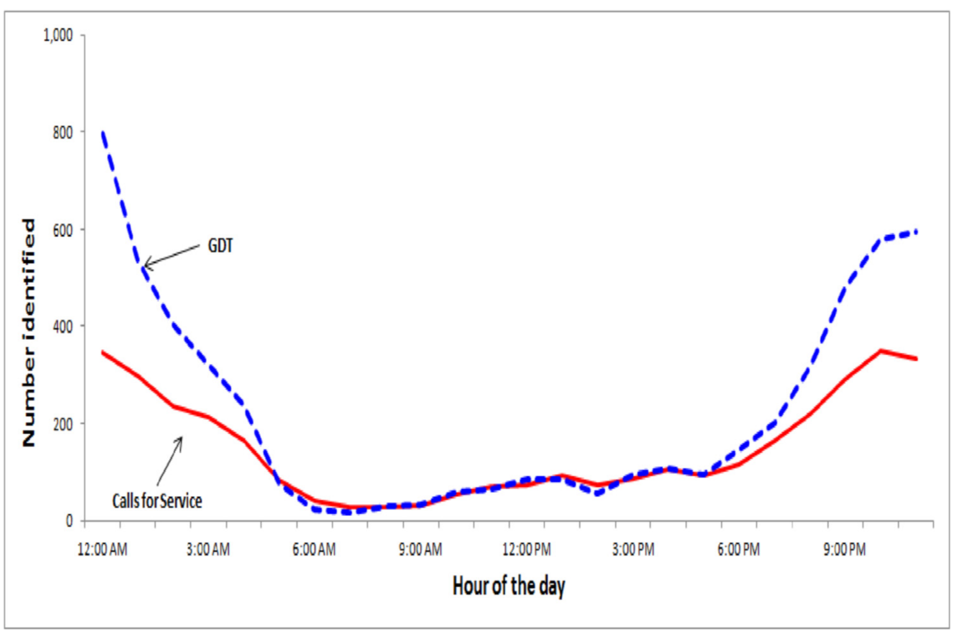

Fig. 4. GDT detection and calls for service by hour of the day.

The Pearson 'r' correlation between the two series for hour of the day was 0.98, highly significant ($t = 30.81$, $p \le 0.001$). However, again, the Komolgorov-Smirnov two-sample test was also significant at the $p \le 0.001$ level ($D_{max} = 0.075$; Critical $D_{.001} = 0.042$) and, thereby, indicated that the two series were from different underlying distributions.

In summary, GDT detection and the Calls for Service using the 20-min window showed parallel patterns. But, in each case, the Komolgorov-Smirnov test indicated that they were not from the same underlying distributions with the subsequent GDT-to-Calls ratio varying by time.

3.2. Spatial comparison

Both GDT detections and Calls for Service (20-min window) were assigned to 628 grid cells, one quarter mile on a side. The overall pattern of the two distributions was similar with more events being identified in the center of the coverage zones for both

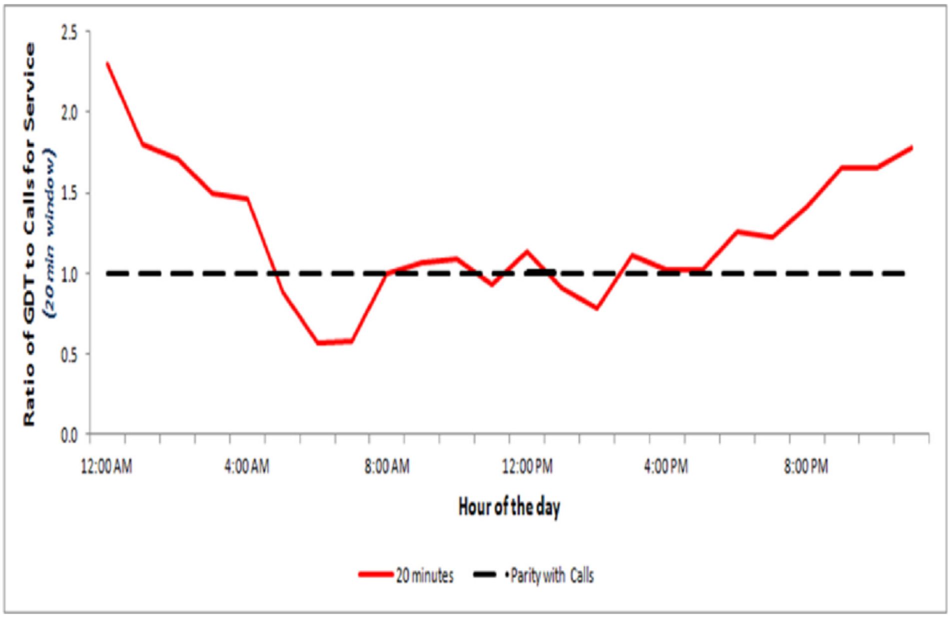

Fig. 5. Ratio of GDT detection to calls for service.

indices (see Fig. 6 for GDT events and Fig. 7 for Calls for Service). The Pearson 'r' correlation was 0.84 between the two distributions and was highly significant ($t = 51.85$; $p \leq 0.001$). Further, the Komolgorov-Smirnov two-sample test was not significantly different for the two distributions over all 628 grid cells ($D_{max} = 0.026$; Critical $D_{.05} = 0.029$). Therefore, the overall spatial pattern was similar between GDT and Calls for Service and can be considered as having the same underlying distributions, unlike the temporal patterns.

However, the GDT-to-Calls ratio varied considerably within the study area, also being higher where there was a concentration of GDT and Calls (see Fig. 8). We suspect that this was due to more acoustical sensors being located near to these concentrations. But, since we do not know the sensor locations, we cannot test that hypothesis.

3.3. Space-time comparison

As was shown above, the temporal patterns in the GDT-to-Calls ratio varied and this affected the spatial patterning, too. When we

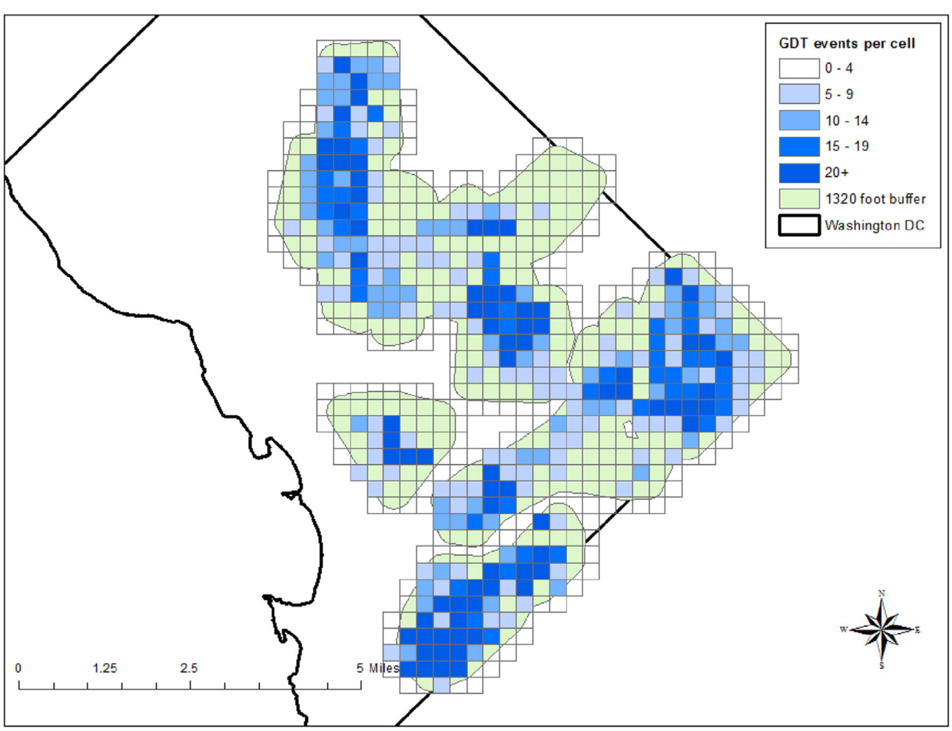

Fig. 6. GDT Gunshot Identification by 0.25 mile Grid Cells.

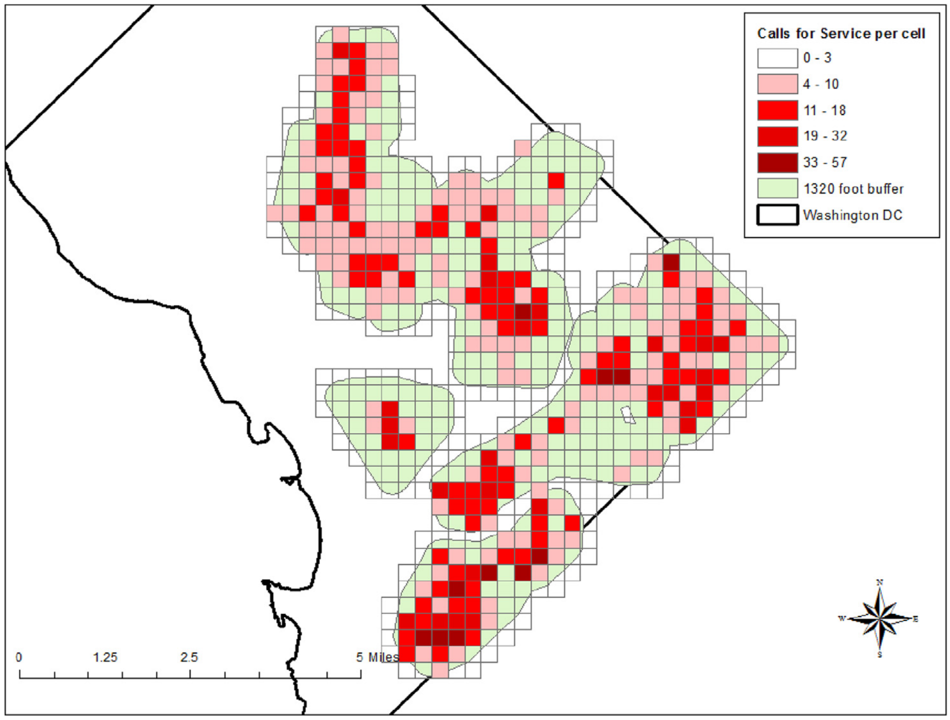

Fig. 7. Calls for Service Identification by 0.25 mile Grid Cells.

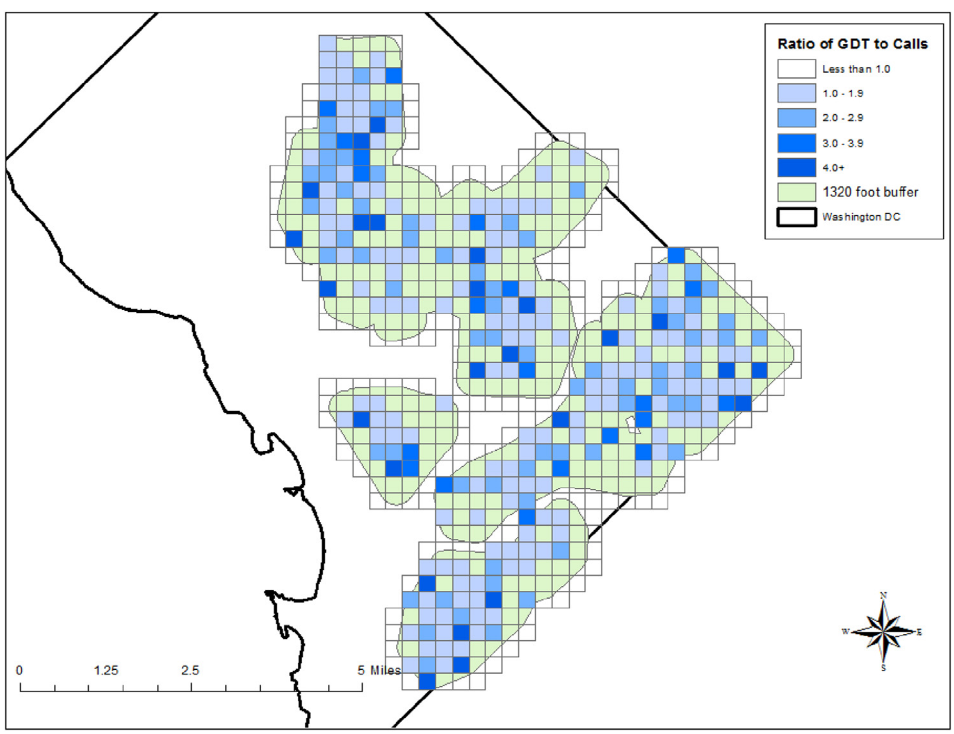

Fig. 8. Ratio of GDT gunshot identification to calls for service.

examined the space-time data set that combined GDT and Calls for Service by six 4-h time periods, we found substantial variations that indicated the GDT-to-Calls ratio was not constant in space and time. The Pearson ‘r’ correlation coefficient over the 3728 cells was 0.72 and was statistically significant ($t = 93.97$; $p \leq 0.001$). But, the Pearson ‘r’ correlation was lower for the space-time data set than for the grid cells for all time periods. Further, unlike the grid cells for all time periods, the Komolgorov-Smirnov two-sample test was statistically significant for the two distributions over all 3768 grid cells ($D_{max} = 0.084$; Critical $D_{.05} = 0.044$). In other words, the spatial pattern of GDT and Calls for Service were not from the same distribution but varied by time period.

Table 1
Variables used in space-time model.

Dependent variable:	**GDT-to-Calls Ratio by Time Period by Grid Cell**
Independent variables:	**Time Period** (dummy variables)
	Midnight – 3:59 a.m.
	8:00 a.m.–11:59 a.m.
	Noon – 3:59 p.m.
	4:00 p.m.–7:59 p.m.
	8:00 p.m.–11:59 p.m.
	Distance From the Nearest Zone Centroid (miles)
	Number of Gun-related Homicides and Assaults
	Border Zone (dummy variable)

Table 2
Space-time Models of GDT Sensitivity. Poisson-Gamma-CAR Models for GDT-to-Calls Ratio as a Function of Time and Spatial Location (N = 3768 space-time grid cells).

Dep. Var: Log of GDT			
	Model 1:	**Model 2:**	**Model 3:**
N:	3768	3768	3768
Df:	3759	3761	3759
No. of samples:	100,000	100,000	100,000
'Burn in' samples:	50,000	50,000	50,000
Log likelihood:	−7568.85	−7578.09	−7569.82
AIC:	15,155.70	15,170.19	15,157.64
BIC:	15,211.81	15,213.83	15,213.75
Mean Absolute Deviation:	1.56	1.55	1.54
Mean-squared Error:	16.75	16.62	16.50
Dispersion multiplier:	10.20***.	10.26***	10.26***
Modeled Parameters	**Coefficient**	**Coefficient**	**Coefficient**
Exposure variable			
calls for service:	1.0000	1.0000	1.0000
Independent variables			
Intercept:	0.4788 n.s	**0.8012***	**0.9440***
Midnight – 3:59 a.m.:	**1.1740***	**0.8601***	**0.8443***
8:00 a.m. – 11:59 a.m.:	**0.6434***	–	–
Noon – 3:59 a.m.:	−0.1702 n.s.	**−0.5122***	**−0.4989***
4:00 p.m. – 7:59 p.m.:	0.3754 n.s.	–	–
8:00 p.m. – 11:59 p.m.:	**1.1077***	**0.7903***	**0.7783***
Distance from nearest zone centroid:	**−0.3764****	**−0.3844****	**−0.3752****
Number of gun-related homicides/assaults:	–	–	−0.0900 n.s.
Border zone:	–	–	−0.2419 n.s.
Spatial Autocorrelation			
CAR (Average φ)	0.0000 n.s.	0.0000 n.s.	−0.0000 n.s

n.s. Not significant.
*p ≤ 0.05.
**p ≤ 0.01.
***p ≤ 0.001.

3.3.1. Space-time modeling

To understand variation in spatial pattern by time, we modeled the GDT-to-Calls ratio over the space-time grid cells using a Poisson-Lognormal-CAR model. To measure time, five dummy variables were created corresponding to five of the six 4-h time periods. To avoid over-fitting the model, the time period of 4:00 a.m.–7:59 a.m. was excluded from the model. To measure spatial location, the distance of each grid cell from the nearest zone centroid was used.

Space and time are the primary independent variables. However, to control for possible distorting influences, two other variables were examined: 1) The number of gun-related crimes (homicides and assaults) that were reported in 2010; and 2) whether the grid cell was on the border of the four coverage zones. The gun-related homicides and assaults might be expected to correlate with both the number of GDT detections and Calls for Service. However, it would not necessarily correlate with the GDT-to-Calls ratio. Therefore, it can be considered as a potential 'adjustment factor' in improving the space-time model. The border grid cell might also have an effect on the GDT-to-Calls Ratio since we added a quarter mile buffer zone to the coverage areas. Table 1 summarizes the variables used in the model.

Three models were tested. In the first, we included the five time periods and the distance variable. In the second model, we included only variables that were statistically significant. In the third model, we added the ancillary variables of number of gun-related homicides and assaults and the border zone variable. Table 2 presents the results for the three models. The estimated coefficients are shown along with the model log-likelihood, AIC, BIC, mean absolute deviation, mean squared predictive error, and dispersion multiplier statistics. The value of the CAR (spatial autocorrelation) component shown, φ, was an average over all 3768 grid cells. The actual CAR value that was output is observation-specific, φ_i, and was a local adjustment to the predicted value (not shown). In all three models, the coefficient for Calls for Service was 1.0 since this was a risk/exposure type model within a Poisson mixed model.

Model 1 shows that there were variations in the GDT-to-Calls ratio by time of day (evening, nighttime, and late morning had higher ratios while the afternoon period had a lower ratio). Similarly, the GDT-to-Calls ratio decreased with distance from the centroid of the coverage zones. Model 2 produced a reduced form of this model where all coefficients were statistically significant. In particular, the GDT-to-Calls ratio was significantly higher in the evening and nighttime periods (essentially between 8:00 p.m. and 3:59 a.m.) and significantly lower in the afternoon period. Again, the distance from the coverage zone centroid was significantly negative. Model 3 added the gun-related homicide and assault counts and the border zone identifier. However, neither of these variables was statistically significant above-and-beyond the variables found for the reduced form model (model 2).

An alternative model was run in which the spatial autocorrelation component was tested only for a distance of one mile. Supplementary material presents this alternative. However, the results were essentially the same. In summary, in 2010 the GDT-to-Calls ratio varied by time of day (evening and nighttime were higher; the daytime and especially the afternoon were lower) and by distance from the zone centroids.

3.3.2. Distance decay by time period

We ran the Poisson-Lognormal-CAR exposure model for each of the six 4-h time periods to estimate the decline in the GDT-to-Calls ratio by distance from the zone centroid. We examined the coefficient of the distance variable for each of the six time periods as a partial prediction of the GDT-to-Calls ratio (see Fig. 9). That is, this figure shows the partial prediction from the distance component only (i.e., neither the number of Calls for Service nor the CAR component are included in the prediction on the Y-axis). Initially, we had expected that the distance decay would be stronger during the daytime hours than at nighttime. However, inconsistent results were found that only partially supported that hypothesis.

The sharpest decay occurred during the morning period (8:00 a.m.–11:59), as expected. Similarly, the decay for the late night period (Midnight – 3:59 a.m.) and late afternoon (4:00 p.m.–7:59 p.m.) showed a less steep decay. The flattest decay was for the afternoon period (Noon – 3:59 p.m.), while the evening period (8:00 p.m.–11:59 p.m.) actually showed a positive slope (i.e., the sensitivity of GDT relative to Calls increased with distance). In short, the results are only partially consistent with the hypothesis that the distance decay in relative GDT sensitivity would be less sharp at nighttime due to less ambient noise.

3.3.3. Spatial autocorrelation and residual errors

The CAR function is an observation-specific spatial adjustment

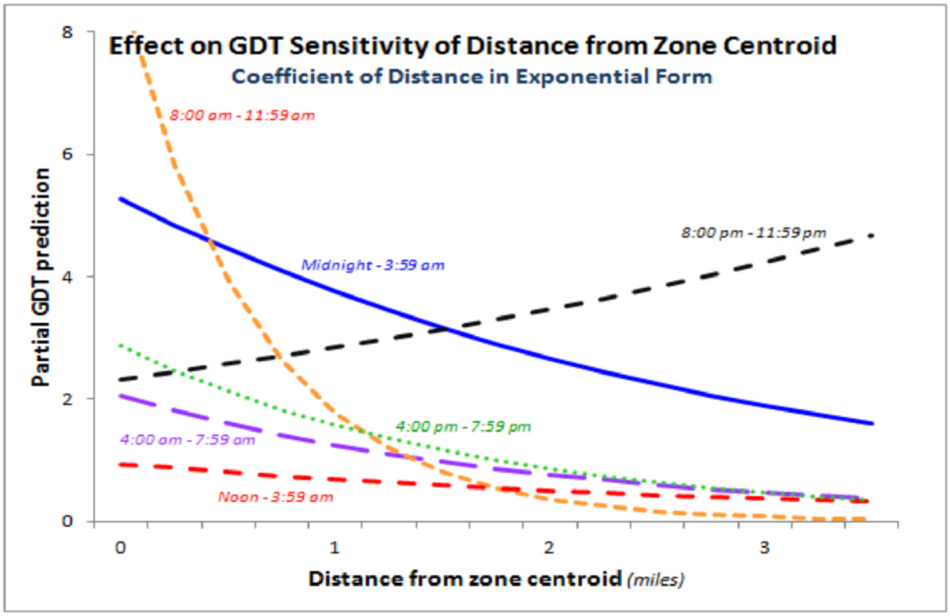

Fig. 9. Effect on GDT sensitivity of distance from zone centroid.

for local spatial autocorrelation, ϕ_i (or Phi). We examined the individual Phi values for the full model (model 1), but they involved less than a 1% adjustment for all grid cells. Therefore, they were very minor adjustments to the predicted values for each grid cell-time period combination.

We also examined the residual errors to see if there were areas where the model under-estimated (i.e., the relative GDT sensitivity was higher than expected by the model) or over-estimated (i.e., the relative GDT sensitivity was lower than expected by the model). Because the Poisson-Lognormal-CAR model overestimated GDT detections relative to Calls for Service by 67%, it was necessary to re-scale the predicted values to equal the actual number of GDT detections.

Fig. 10 shows these residuals. There were two areas where the model underestimated, one in the northern part of the coverage area (around Columbia Heights in Ward 1) and a second in the central part of the coverage area (around Gallaudet, Ivy City and Carter Langston in Ward 5). For these areas, GDT detections were greater than that predicted by the model.

On the other hand, there were two areas where the relative GDT

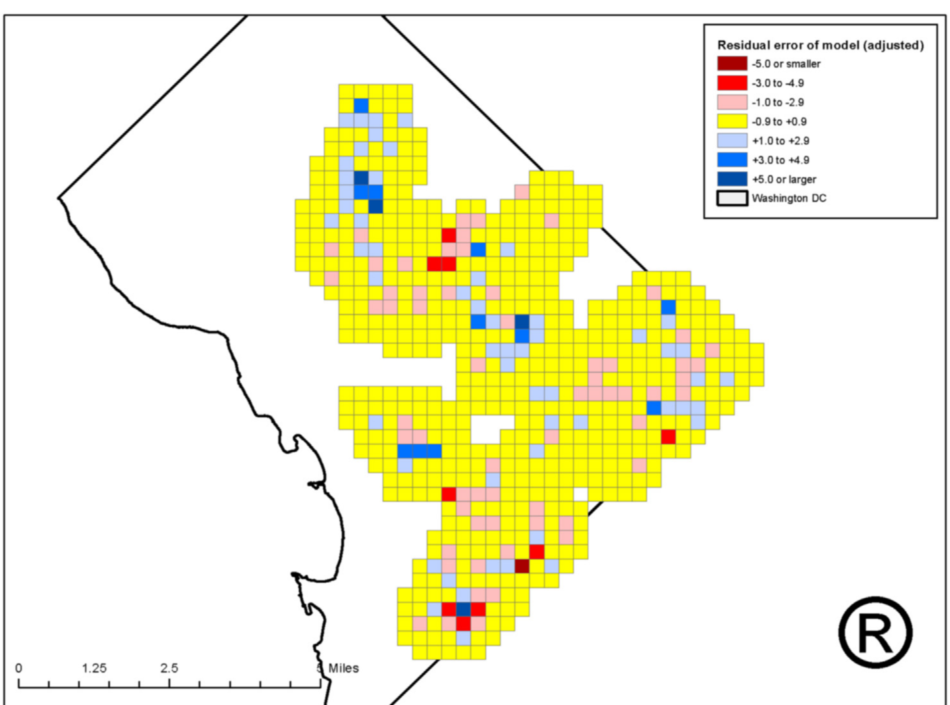

Fig. 10. Overestimation and Underestimation of Space-Time Models (overestimation in red; underestimation in blue).

sensitivity was less than that predicted by the model, one in the southwest corner of the coverage (around Washington Highlands in Ward 8) and the other north of the coverage area center (around Edgewood and Bloomingdale in Ward 5). It is possible that more acoustical sensors are needed in those areas to improve GDT sensitivity.

4. Discussion

The results of these analyses show that relative GDT sensitivity (the GDT-to-Calls ratio) varies both by time and by space. In particularly, the relative sensitivity of GDT was much stronger in the evening and at nighttime than in the daytime, varying between 1.5 and 2.3 times as much as the number of Calls for Service. On the other hand, the GDT-to-Calls ratio was close to parity during the daytime hours and actually fell below 1.0 for 6 h during the daytime. The reason for this lack of daytime sensitivity is undoubtedly the level of ambient noise, presumably from traffic and construction noise.

In terms of spatial variation, we found that GDT sensitivity decreased with distance from the nearest zone centroid with the exception of the evening period. We suspect this has to do with where the acoustical sensors are located, with distance being a limiting factor. But since we do not know those locations, we could not estimate an optimal distance for locating the GDT sensors. However, the data certainly point to a general distance decay pattern and also point to certain areas of the city where the relative sensitivity is weaker. Adding more sensors in those locations would certainly help improve the sensitivity of the technology.

We do not know why the GDT-to-Calls ratio increased by distance during the evening period. Both GDT and Calls for Service separately decreased with distance from the zone centroid during this period. However, the sensitivity of GDT relative to the calls increased in the evening, but decreased for all other time periods. More research on this point is necessary.

These results corroborate early research on the differences between GDT and reported crimes. Carr and Doleac (2015a) found strong correlations between gunshot and both reported crimes and Calls for Service but noted different trends in gunshots and reported crimes across a city based on local land usage. Our results suggest a similar relationship.

The limited relationship between firearm-related Calls for Service and GDT activations is to be expected. Gunshots that do not result in death or serious injury are frequently underreported (Mazerolle, Watkins et al., 1999) so it is not surprising that Calls for Service for gun-related crimes are only partially related to activations of the system. Also, in some neighborhoods, people may be more sensitive to hearing gunshots and, therefore, more likely to call in, whereas in other neighborhoods the sound of a gunshot may be so routine as to not result in calls, perhaps because residents perceive that the calls do not result in any reductions in violence. In other words, GDT and Calls for Service may be capturing different phenomena.

Another consideration is the spatial accuracy of GDT compared to Calls for Service. Unless a caller is close to where the gunshots were fired, that person cannot accurately identify the spatial location, only the general direction. On the other hand, GDT is very accurate with respect to the spatial location of a gunshot. Even indoors, the sound of a gunshot can be reliably identified from at least 1000 feet away (Bieler & La Vigne, 2014). Thus, GDT offers better accuracy for those events that it detected. But, as we saw, its sensitivity is not better than human response during the daytime.

These results also demonstrate the challenge in correctly identifying multiple calls for the same gunshot event. Individuals who call the police may reside in different directions from the gunshot so that their directional accuracy is relative. It is possible, though not practical, to triangulate calls from multiple callers to approximate the location of a gunshot. However, few police departments are going to assign staff to do that or fund the development of software for triangulating human response to a gunshot. GDT, on the other hand, can do this quickly with very good spatial accuracy (25 m).

4.1. Limitations

There are several limitations to our study. First, we did not have an independent data base of gunshot events by which both GDT and Calls for Service could be evaluated. The gun-related crime data set that we used covered fewer than 10% of the detections and also did not have a time of day identifier. It would be possible to get a more accurate data set of gun crimes that did have a time stamp in order to evaluate the accuracy of GDT, but this would only cover those gun shots that led to a crime, not all gun shots.

Second, in modeling the GDT-to-Calls ratio by space and time, we only examined the hourly variation in time, grouped into 4 h periods. But, as mentioned, there were also seasonal variations that could alter the sensitivity, particularly on January 1st where there is an excessive number of gunshots fired, mostly in celebration.

Third, the Poisson-Lognormal-CAR model that was used to examine the space-time dataset has its own problems. In our use of it, the model over-estimated the number of GDT detections relative to Calls for Service. In other studies, however, this type of model has underestimated the dependent variable (Levine, 2017). Better statistical models may produce more accurate representations in the future.

4.2. Conclusion

Future research could greatly improve the use of GDT data by better understanding how different ranges and conditions affect GDT's ability to detect gunshots and about the extent to which police should depend on GDT as opposed to public Calls for Service as well as how the two data sources can be used in combination. Nevertheless, based on these findings as well as prior analyses, this study concludes that GDT may offer a valuable new source of data on gun violence, but with some limitations in terms of time of day and distance from acoustical sensors.

Acknowledgements

This research was supported by Award No. 2012-R2-CX-0001 awarded by the National Institute of Justice, U.S. Department of Justice. We would like to thank our NIJ program officer Eric D. Martin for his guidance and to David Chipman, James Bledrock, Mark Jones, and Joe Hawkins of ShotSpotter SST. for their insight into the operation and dynamics of DC's GDT system and its history.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.apgeog.2017.06.013.

References

Alavarado, S. E., & Massey, D. S. (2010). In search of Peace: Structural adjustment, violence, and international migration. *Annuals of the American Academy of Political Science, 630*(1), 137–161.

Archer, D., & Gartner, R. (1984). *Violence and crime in cross-national perspective*. New Haven: Yale University Press.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of the Royal Statistical Society: Series B, 36*, 192–236.

Besag, J., Green, P., Higdon, D., & Mengersen, K. (1995). Bayesian computation and stochastic systems (with discussion). *Statistical Science, 10*, 3–66.
Bieler, S., & La Vigne, N. (2014). *Close-range gunshot around DC schools*. Washington, DC: Urban Institute.
Braga, A. A. (2003). Serious youth gun offenders and the epidemic of youth violence in boston. *Journal of Quantitative Criminology, 19*(1), 33–54.
Braga, A. A. (2007). *Gun violence among serious young offenders*. Washington, DC: Department of Justice.
Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data*. Cambridge, U.K: Cambridge University Press.
Carr, J. B., & Doleac, J. L. (2015a). *The geography, incidence, and underreporting of gun Violence: New evidence using ShotSpotter data*. Charlottesville, VA: University of Virginia.
Carr, J. B., & Doleac, J. L. (2015b). *Keep the kids Inside: Juvenile curfews, bad weather, and urban gun violence*. Charlottesville, VA: University of Virginia.
Coleman, C., & Moynihan, J. (1996). *Understanding crime Data: Haunted by the dark figure*. Buckingham UK: Open University Press.
Cook, P., Ludwig, J., Venkatesh, S., & Braga, A. A. (2007). Underground gun markets. *The Economics Journal, 117*(524), F588–F618.
Emory University. (2016). *Sensitivity, specificity, predictive value for a positive result, predictive value for a negative result*. Emergency Medical-Ambulatory Care, Emory University Accessed October 15, 2016 https://www.med.emory.edu/EMAC/curriculum/diagnosis/sensand.htm.
Eng, T. R. (2004). Population health Technologies: Emerging innovations for the health of the public. *American Journal of Preventative Medicine, 26*(3), 237–242.
Federal Bureau of Investigation (FBI). (2011). Crime in the United States, 2010. Available at: https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/tables/10tbl08.xls/view (accessed 10 August 2015).
Fox, J. A. (1996). *Trends in juvenile violence: A report to the United States attorney general on current and future rates of juvenile offending*. Washington, DC: Bureau of Justice Statistics.
Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
Goode, E. (2012). *Shots fired, pinpointed and argued over*. New York Times. Available at: http://www.nytimes.com/2012/05/29/us/shots-heard-pinpointed-and-argued-over.html?pagewanted=all&_r=0 (accessed 16 December 2015).
Hales, G., Lewis, C., & Silverstone, D. (2006). Gun crime: The market in and use of illegal firearms. In *Home office research studies 298*. London, United Kingdom: Home Office.
Jones- Brown, D. (2007). Forever the Symoblic Assailant: The More Things Change, the More They Remain the Same. *Criminology and Public Policy, 6*(10), 103–121.
Kanji, G. K. (2006). *100 statistical tests* (3rd ed.). Los Angeles: Sage Publications.
Kidd, R. F., & Chayet, E. F. (1984). Why do victims fail to Report? The psychology of criminal victimization. *40*(1), 39–50.
La Free, G. (2005). Evidence for elite convergence in cross-national homicide victimization trends, 1956 to 2000. *The Sociological Quarterly, 46*(1), 191–211.
Langton, L., Berzofsky, M., Krebs, C., & McDonald, H. (2012). *Victimizations not reported to the police, 2006-2010*. Washington, DC: Bureau of Justice Statistics.
Levine, N. (2017). The location of late night bars and alcohol-related crashes in houston, Texas. *Accident Analysis & Prevention* (in press).
Levine, N., Lord, D., Park, B.-J., Geedipally, S., Teng, H., & Sheng, L. (2013). Chapter 19: Spatial regression modeling. In N. Levine (Ed.), *CrimeStat: A spatial statistics program for the analysis of crime incident locations (v. 4.0)*. Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. June. http://www.nij.gov/CrimeStat.
Litch, M., & Orrison, G. A. (2011). *Draft technical report for secures demonstration in hampton and newport news, Virginia* (US DOJ report).
Lord, D., Levine, N., & Park, B.-J. (2013). Chapter 16: Poisson regresson modeling. In N. Levine (Ed.), *CrimeStat: A spatial statistics program for the analysis of crime incident locations (v. 4.0)*. Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. June. http://www.nij.gov/CrimeStat.
Lord, D., Levine, N., Park, B.-J., Geedipally, S., Teng, H., & Sheng, L. (2013). Chapter 17: Estimating complex models with Markov Chain Monte Carlo simulation. In N. Levine (Ed.), *CrimeStat: A spatial statistics program for the analysis of crime incident locations (v. 4.0)*. Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. June. http://www.nij.gov/CrimeStat.
Lynch, S. M. (2007). *Introduction to applied bayesian statistics and estimation for social scientists*. New York: Springer.
Maher, R. C. (2007). Acoustical characterization of gunshots. In *Proc. IEEE SAFE 2007: Workshop on signal processing applications for public security and forensics, Washington, DC* (pp. 109–113).
Ma, J., Kockelman, K. M., & Damien, P. (2008). A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis & Prevention, 40*, 964–975.
Mazerolle, L. G., Frank, J., Rogan, D., & Watkins, C. (1999). *Field evaluation of the ShotSpotter gunshot location System: Final report on the redwood city field trial*. Washington, DC: National Institute of Justice.
Mazerolle, L. G., Watkins, C., Rogan, D., & Frank, J. (1999). *Random gunshot problems and gunshot detection systems*. Washington, DC: National Institute of Justice.
Park, E. S., & Lord, D. (2007). Multivariate poisson-lognormal models for jointly modeling crash frequency by severity. In *Transportation research record 2019: Journal of the transportation research board, TRB* (pp. 1–6). Washington, D.C: National Research Council. https://ceprofs.civil.tamu.edu/dlord/Papers/Park_%26_Lord_Joint_Modeling.pdf.
Penney, T. L. (2014). Dark figure of crime (problems of estimation). In G. Bruinsma, & D. Weisburd (Eds.), *The encyclopedia of criminology and criminal justice*. Hoboken, NJ: Wiley-Blackwell.
Showen, R. (1997). Operational gunshot location system. In *Proc. SPIE 2935, surveillance and assessment technologies for law enforcement*. Available at: http://proceedings.spiedigitallibrary (accessed 17 December 2015).
Siuru, B. (2007). Gunshot location systems. *Law and Order, 55*(10). pp. 10, 12, 15, 17.
Solis, C., Portillos, E. L., & Brunson, R. K. (2009). Latino youths' experiences with and perceptions of involuntary police encounters. *The Annals of the American Academy of Political and Social Science, 629*, 39–51.
SST Inc. (2014). *Additional cities sign-on to use ShotSpotter to combat illegal urban gunshot*. SST. Available at: http://www.shotspotter.com/press-releases/article/additional-cities-sign-on-to-use-shotspotter-to-combat-illegal-urban-gunshot.
SST, Inc. (2016). *ShotSpotter*. Newark, CA: SST, Inc. http://shotspotter.com/.
Watkins, C., Mazerolle, L. G., Rogan, D., & Frank, J. (2002). Technological approaches to controlling random Gunshot: Results of a gunshot detection system field test. *Policing: An International Journal of Police Strategies and Management, 25*(2), 345–370.