Testing the Spatial Accuracy of Address Based Geocoding for Gun Shot Locations

Andrew P. Wheeler^{1,**}, Manne Gerell², & Youngmin Yoo¹

*Data and code to replicate the findings can be obtained from
<https://www.dropbox.com/sh/bceyldwgj84ztlw/AABdQBnjKGdO3GUxWM0ZMd3Ya?dl=0>

**Corresponding author, email: apwheele@gmail.com

1. School of Economic, Political, and Policy Sciences – Program in Criminology & Criminal Justice

The University of Texas at Dallas

800 West Campbell Road, Mail Station GR 31

Richardson, Texas 75080-3021

2. Department of Criminology, Malmö University, Sweden.

Title: Testing the Spatial Accuracy of Address Based Geocoding for Gun Shot Locations

Abstract: We assess the positional accuracy of address based geocoding of shooting incidents relative to the location recorded via acoustic gun-shot detection technology. This provides a test of the accuracy of typical address based geocoding methods used in crime analysis, as well as provides evidence for how much accuracy one gains when using sensors. Examining over 1,000 shooting incidents in Wilmington, North Carolina, we find that the majority of address-based incidents are quite accurate, on average within 60 feet of the actual location (using a street centerline geocoder), or within 90 feet (using Google rooftop geocoding). However, based on the incident narrative we identify a subset of transcription errors in over 10% of the cases that increases the distance between the true shooting location and that geocoded using address data. This suggests mechanisms to prevent human errors may be more frugal than those relying on sensors in geocoding shooting incidents.

Keywords: Gunshot-Detection, Geocoding, Spatial-Accuracy, Shootings

Data to replicate the analysis can be downloaded from

<https://www.dropbox.com/sh/bceyldwgi84ztlw/AABdQBnJKGdO3GUxWM0ZMd3Ya?dl=0>.

Introduction

In the present analysis we test the accuracy of address-based geocoding for shooting incidents relative to the location recorded by acoustic gun-shot detection technology (GDT), ShotSpotter, in Wilmington, North Carolina. Prior analyses have focused on the ability of GDT to record shooting incidents that are not reported to the police, and thus potentially provide more accurate estimates of the incidence of shootings (Carr & Doleac, 2016; Carr & Doleac 2018; Irvin-Erickson et al., 2017; Mazeika & Uriarte 2018). Although it is controversial whether those non-reported incidents represent actual shootings, or whether they are the result of false positives detected by the technology (Drange, 2016; Ratcliffe et al., 2019).

An additional factor that is often used to motivate GDT is more spatially accurate locations of the shootings (WECT News 6, 2014). Here we take the GDT recorded locations as true, and evaluate the error associated with address based geocoding of the same incidents. This provides quantitative estimates of how much spatial accuracy one gains with GDT, an important consideration given the costs of the technology (Pizzi, 2015).

Place-based analysis of crime has become more prevalent in recent decades (Weisburd et al., 2011). This is to a large extent driven by technological innovation coupled with research findings on the concentration of crime at hot spots (Sherman et al., 1989; Weisburd, 2015) and the overall success of hot spots policing to reduce crime (Braga et al., 2012). Identifying hot spots and deploying police resources to those locations to reduce crime is a wide spread policing tactic, with one study finding that 70% of police departments in the U.S. with at least 100 officers using crime mapping to identify hot spots (Weisburd et al., 2003). Using crime mapping to identify high crime locations, and deploying resources based on such mapping works to reduce crime, and has become main-stream practice in policing.

GDT, and the use of sensors in general, can be seen as a further technological extension to traditional GIS based technologies. Much remains to be learnt on the extent to which such technology can improve on data quality, crime analysis and police responses to crime. Such technology should improve the accuracy of spatial analyses, and increases in accuracy should theoretically translate into improved policing responses to crime (Lawrence et al., 2018).

With modern geographic information systems it is simple to plot crimes on a map to see what appears to be the exact locations of crime incidents and to perform advanced statistical analysis based on those locations. However, it is well known that there are reliability issues with crime data (Tompson et al., 2015; Ratcliffe, 2004; Carr & Doleac, 2016), and the crime maps may portray an image of crime as more exact than what police recorded data can actually capture (Field, 2011; Singleton & Brunsdon, 2014). Crimes can take place anywhere humans can travel to, but are typically recorded to an address or an intersection (Levine & Kim, 1998; Wheeler, Worden, & McLean, 2016), and thus only represent an approximation to where the crime actually took place.

The majority of prior analyses assessing geocoding accuracy of address level crime data have lacked a gold standard for where the crime actual occurred in evaluating positional error (Bichler & Balchak, 2007; Hart & Zandbergen, 2013; Ratcliffe, 2001), and thus the extent of error which such address based approximations generate is for the most part unknown. For a recent exception, Gerell (2018) used GPS coordinates obtained by the fire department at the exact scene of 114 car arsons in Sweden, and he finds that address based incidents had an average error of 83 meters. That study was for just one crime type, among a small sample of incidents, and in the country of Sweden. Subsequently this analysis can be seen as an additional evidence estimating the extent of error generated via address based geocoding of crime incidents with a gold standard reference.

A final contribution of the current article is assessing not only address based geocoding error, but also *human* based errors that contribute to inaccuracy of geocoding. Upon undertaking the analysis, we identified that many of the most egregious errors appeared to be the result of human transcription errors. Using reported discrepancies in the listed address versus those described in the incident narratives we estimate of the proportion of such errors, which is over 10% in this sample. While address based geocoding resulted in little spatial error (under 100 feet on average), the high proportion of human errors suggests other mechanisms to audit and correct address based data are potentially a much more frugal way to obtain spatially accurate shooting data.

Literature Review

Shotspotter and Gunshot Detection Technology

GDT has quickly gained popularity among US police departments, and in 2016 it was reported that more than 90 police departments had adopted ShotSpotter technology (Carr & Doleac, 2016). GDT is based on acoustic sensors that triangulate the location of gun shots based on multiple acoustic sensors (Irvin-Erickson et al., 2017). One independent study found that the geographical error of a ShotSpotter system had a median error of 25 feet (Mazerolle et al., 2000),

while another paper suggests 10-25 meters (30 to 75 feet) error depending on wind and other conditions (Aguilar, 2015).

To date one of the main focuses of research on GDT has been on its ability to identify underreporting of gun violence through detecting incidents that are not reported to the police. Carr & Doleac (2016) found that only 12.3% of GDT recorded gun fire incidents resulted in a call for service for shots fired, and 2.3% (Washington) and 6.4% (Oakland) of gun fire incidents resulted in a police report for assault with a deadly weapon. They also noted large variations within cities, with some areas of a city having almost twice the rate of calls for service relative to GDT registered incidents.

Some of those findings are further elaborated on in a paper that considered how the relationship of GDT data with calls for service might vary both in space and time (Irvin-Erickson et al., 2017). The study found that the relationship was significantly stronger for evenings and nights than daytime gunfire incidents for GDT data in Washington, D.C. They also noted that GDT appears to function less well with distance from sensors, so that some areas would have a smaller chance of detection of shots fired. While they note that GDT provides a much better spatial accuracy than calls for service, they also noted that this advantage disappears during daytime, likely due to increased noise from traffic and construction (Irvin-Erickson et al., 2017).

While the overall picture gained from Carr & Doleac (2016) and Irvin-Erickson et al., (2017) is one of GDT improving data-quality, some concerns are raised on this topic by Ratcliffe et al., (2019). Employing GDT coupled with closed captioned television (CCTV) they find that GDT technology indeed result in far more (+259%) gunfire incidents recorded, but this fails to materialize as confirmed incidents upon investigation. This is suggestive that at least parts of the GDT detected gun shots are false positives, and the GDT technology increased police work load while not resulting in significant increases in confirmed gun violence incidents. Ratcliffe et al., (2019) conclude by suggesting that GDT may be of more value if coupled with CCTV that can be automatically directed towards gunfire to collect video footage of the scene around an incident. Work by Drange (2016) shows similar patterns in a lack of evidence collected after GDT detections in multiple jurisdictions, so the findings by Ratcliffe et al. (2019) do not appear to be idiosyncratic.

While the above-mentioned studies have focused on studying gun detection sensors from a methodological point of view, some studies have also employed such data in relation to empirical questions. Carr & Doleac (2018) for instance used GDT data to examine the effect of juvenile curfews, and found that the curfew was associated with more shootings in Washington, D.C. They do not find the same relationship when using reported calls for service, and thus suggest that juvenile curfews themselves likely make reporting shootings less likely to happen.

For other examples of analyses using GDT data, Loeffler & Flaxman (2018) and Wheeler (2017) evaluate spatial predictive models using the same D.C. GDT data as others. Mazeika & Uriarte (2018) used GDT data in Trenton, New Jersey, to examine near repeat patterns for gun violence, and found that the use of GDT data led to more near repeat patterns identified relative to calls for service.

Overall the research on GDT technology suggest that it can reduce false negatives and yield better results for analysts, but at the same time the risk for false positives may mean that

analytical gains are offset by increased workloads for police. Its cost effectiveness can thus be questioned based on current evidence.

Geocoding accuracy

As mentioned in the introduction, a key element for geographical analysis is how to convert the location of an event (such as a crime) into a location on a map. Sometimes this is easier, for instance when a GPS device is used from the exact location of an event, an example of which is how the rescue services record the location of a torched car in Sweden (Gerell, 2018). In relation to crime however, the norm is that the event gets coded to an address, which then needs to be geocoded to a location. This introduces several potential sources of bias which could impact data quality (Zandbergen, 2009).

One type of bias occurs when the exact location of the event is unknown, or at least very difficult to determine. This could for instance be the case when a call for service for shots fired come in from a caller that has heard gunshots, but no victim is found. Gun shots are unique compared to most crimes in the sense that they can be detected at a greater distance via sound; Bieler & La Vigne (2014) suggest they can be reliably detected by sound up to 1,000 feet away. Whereas most crimes need to be visibly seen before an individual can be expected to report the incident to police. In such instances it would appear plausible that GDT can be of great help in providing an accurate location, rather than a general area based on the caller's judgement.

Once the (approximate) location of the crime has been established, it will be assigned to an address. With crimes taking place outdoors this will typically mean the crime gets assigned to a nearby building, resulting in yet another potential source of error. In some cases incidents that occur outdoors are assigned to the nearest intersection (Levine & Kim, 1998; Wheeler, Worden, & McLean, 2016). This error will usually be fairly small, but can be larger if the location has no nearby buildings, for instance when it is by a park (Gerell, 2018). Block & Block (2011) describe examples of geocoding homicide locations (based on where the body was found) can result in locations that are too ambiguous to assign to an address (such as in a river), and as such cannot be reliably located using a standard address or street centerline file. It is plausible that GDT could be helpful in circumstances, such as a shooting in a park, as it is not reliant on addresses to determine locations.

A third type of error then finally comes when the address needs to be assigned a set of spatial coordinates. Some addresses will be incomplete or incorrect, or otherwise fail to be recorded (Bichler & Balchak, 2007). In addition, even when geocoding is successful the reliability of it may vary. There is also a potential tradeoff between successful geocoding and positional accuracy of the results. Higher geocoding rates can be achieved by lowering the threshold for success, but this may mean some of the geocoding is inaccurate (Bichler & Balchak, 2007; Hart & Zandbergen 2013; Mazerika & Summerton, 2017). These two types of errors within the geocoding process will be discussed below.

There are a few different methods for geocoding currently employed. It has traditionally been done through the use of street geocoding, where a database matches street segments to locations and then interpolates the exact location based on the range of addresses on the street segment (Mazeika & Summerton, 2017). There are multiple databases that can be

used for street geocoding, in the US the TIGER database covers the whole country and is the most common (Hart & Zandbergen 2013). A potentially better alternative, when available, is address point geocoding, where each address has an exact location coded, thus not reliant on interpolation of addresses along a street segment (Ratcliffe, 2001; Zandbergen & Hart, 2009). Hart and Zandbergen (2013) mention that while this method is not used often in the US, it has been used in Australia, Canada and the UK, and it is also used by Swedish police, as with the data in Gerell (2018). Hart and Zandbergen (2013) used address point data compared with street geocoding to determine the positional accuracy of different street geocoding and noted that local street line reference data outperformed the national data such as TIGER. With the fast paced technology development there are now multiple commercial services that can provide geocoding, and although Mazeika and Summerton (2017) note that this likely was the least common method within the criminal justice system this is subject to rapid change, and currently the extent of its usage is unknown.

Several studies have attempted to quantify positional error due to geocoding, which is the most relevant prior work to this study. Ratcliffe (2001) examine the spatial distance of parcel based address geocoding compared to street centerline database, find that 49% of the crimes are within 15 meters of the parcel centroid, and that 96% of crime incidents are within 100 meters. Using a similar research design, Hart and Zandbergen (2013) finds that the median error of the street centerline geocoded data tend to be about 60 to 80 meters. While the use of address point data to measure error is less exact than using a gold standard of where the crime actually took place, these estimates never the less serve to give an indication of a minimum bar for errors based on such geocoding methods. Research appear to be fairly consistent in identifying median errors tend to be smaller than 100 meters, or roughly 200-300 feet.

Mazeika and Summerton (2017) used exact locations determined through google street view to establish mean errors for geocoding of burglaries. They found that the mean error was between 40 and 100 meters when using street geocoding. This number was improved when instead using google to geocode, dropping the error to 10-25 meters. While their study had a more exact measurement of crime locations and was not reliant on address point reference data, it should be noted that the crime studied, burglary, should always be associated with a specific building. This will reduce many of the errors associated with the geocoding process when assigning crimes that take place outdoors (e.g. Thompson et al., 2015).

To sum up then, research appears to fairly consistently show that geocoding produces errors that are around 100 meters. This is however based on either inexact approximations of the true location for a crime (address points), or in the case of Mazeika and Summerton (2017) on a crime that is much easier to geocode than most other crimes. The true error associated with crimes that are more difficult to geocode, such as shootings, remain to be showed. The finding that the positional accuracy of geocoding was increased through the use of new technology, google (Mazeika & Summerton 2017), also points to a need for further study on how technological innovation, such as GDT, may be of use in accurate geocoding.

The second part of geocoding errors deals with the geocoding hit rate. What share of the addresses could be successfully converted to geographic coordinates? For geocoding not to have a global impact on analysis it has been shown that at least 85% need to be coded correctly (Ratcliffe, 2004). While this threshold appears to be quite reachable, a study by Cohen (2006)

noted large yearly differences, with a low 72% for one year and a high 91% for another year. The latter number was echoed when Mazeika and Summerton (2017) successfully geocoded about 93% of burglaries in their study. Using google instead of the street geocoding they achieved a much higher match rate, over 99%. However, even if 85% geocoding can be deemed as acceptable, it should still be noted that anything below 100% will mean some loss in data quality due to geocoding attrition.

It has also been shown that the match rates may differ quite dramatically between crime types. Hart and Zandbergen (2013) found that while overall match rates across different reference data, crime types, and jurisdictions tended to be around 80%, this number was substantially lower when the data was referencing a street block rather than an exact location. When exact address data was available the match rate was between 83% and 86% depending on crime type. But, when the address was for an intersection or a street block the matching rate achieved dropped to between 15% and 39% depending on crime type. This means that when no exact address is recorded for a crime there is an additional error in geocoding on top of the uncertainty and imprecision already inherent with the inexact address. This again suggests a potential benefit for GDT technology, since GDT technology uses triangulation to determine locations it is not likely to suffer from such errors.

Overall then there are multiple sources of geographical error in police crime data. Hypothetically all these sources of error could be alleviated through the use of GDT technology to help determine a more exact location of crimes. The subsequent analysis focuses on the question of positional accuracy of GDT incidents relative to the more typical status-quo of address based crime incident geocoding.

Data and Methods

The gun shot location data used in the analysis have been publicly disseminated by the Wilmington, North Carolina Police Department. We downloaded the data for analysis from <http://justicetechlab.org/shotspotter-data/>, which includes shootings spanning 1/1/2013 through 9/28/2015 (Carr & Doleac, 2016). Wilmington is a city on the eastern coast of the United States, with a population of over 100,000. It has slightly higher crime rates compared to cities of the same size, with a homicide rate of over 12 per 100,000 in 2015 (Wheeler & Kovandzic, 2018).

The dataset includes set of latitude and longitude coordinates as recorded by the GDT, as well as an incident address and that address geocoded in a local coordinate system (the North Carolina State Plane System, EPSG 2264) using a local street centerline geocoding reference system. The original shooting dataset contained 1,268 incidents. For analysis we eliminated cases in which no latitude and longitude were recorded by the GDT ($n = 82$), and then we eliminated duplicated incidents in the records (those recorded as “DUP” in the *closecode* field), an additional $n = 181$ cases. This resulted in a total of 1,005 incidents being used in the final analysis. Thus for this analysis we only focus on distance errors in geocoding addresses, not on errors associated with geocoding hit rates and subsequent analysis.

Besides the street centerline geocoded address provided by the Wilmington Police Department, we generated an additional set of address based coordinates using Google’s online geocoding API, which is a hybrid approach geocoding to the rooftop where available and using

street interpolation when not (Roongpiboonsopit & Karimi, 2010a; 2010b). This resulted in 10 cases in which results were either geocoded to the generic center of the city, or to only the midpoint of the street, or to a location not in North Carolina. These 10 cases are not included in the analysis of the accuracy of the Google geocoding results.

Because the GDT data are in spherical coordinates, whereas the centerline geocoded data are in projected coordinates, we needed to transform one set to the other to calculate the distance between the two sets. The analysis we report is based on transforming the spherical coordinates (both GDT and the Google geocoding results) to the projected coordinates. We additionally conduct the analysis by projecting the local coordinate system to spherical coordinates, and then use the Vincenty formula to calculate distances. The differences are immaterial for the results, with the distances based on the projected data being smaller by only around one third of a foot (or less).

The final analysis step was the manual coding of cases into different transcription error categories. This was done by examining the incident narrative, which often had a plain text description of the address where the incident took place (along with other descriptive characteristics of the incident). We subsequently coded instances in which the narrative location agreed with the recorded incident address field (*Same Address*), cases where the narrative and the address field had the same street but a different address number listed (*Same Street, Different Address*, e.g. 53 Shipyard Walk being listed in the narrative, but 1 Shipyard Walk being the address in the address field), cases where the incident narrative had a totally different street listed (*Different Street*, e.g. 59 S Liberty Ct in the narrative, and 1012 S 12th St in the address field), and finally cases where no address was listed in the narrative (*No Address in Narrative*).

Data and code to replicate the analysis can be downloaded from the supplementary material to the article.

Results

Table 1 displays the distribution of the error distances, along with the proportions of cases in each of the coded categories. Like prior analyses of geocoding errors, they appear to be heavy tailed (Zandbergen et al., 2012), and as such we use quantiles to describe the overall distribution and use the logarithmic transform when estimating regression models and prediction bands (Zandbergen, 2008).

The majority of the cases were classified into the same address category, 86%, where the address listed in the narrative and in the address field were the same. These cases appear to have subsequently the smallest amount of error, with a median error of 50 feet, and 25th and 75th percentile errors of 31 feet and 82 feet respectively. The other three categories make up approximately 5% of the sample each. Same street, different address cases have a slightly higher median error of 65 feet, and those on a different street have a median error of 94 feet. The undeterminate category of no address in the narrative had the highest median error distance of 123 feet, but the 75th percentile appears to be larger for the different street category, and one can see the outlying values in the maximum column.

Comparing the error distributions between the street centerline geocoding and the google rooftop, the street centerline distances are generally smaller, but only by a small amount. For the

same address category, the median error distance is 29 feet larger for the google rooftop versus the street centerline. While the error distances for the same street, different address category have very similar distributions for both street centerline and google rooftop errors, the other two categories (No Address in Narrative and Different Street) have slightly larger error distances in the median error. One potential explanation of this is that google geocoder places the incident at the rooftop location (where available), which might place it further from the street location compared to the centerline locator (Zandbergen, 2009). Given that the centerline geocoded incidents were provided in the data, we cannot be certain as to the exact parameters used to the geocode the data. A smaller offset though would result in a smaller distance to the street.

Table 1: Distribution of Error Distances (in Feet) by Percentiles, Broken Down by Each Incident Type

	<i>Min.</i>	<i>25th</i>	<i>50th</i>	<i>75th</i>	<i>Max.</i>	<i>Sample Size</i>	<i>%</i>
<i>Street Centerline Differences (n = 1,004)</i>							
Same Address	1	31	50	82	3,152	862	86
Same Street, Different Address	19	47	65	138	1,281	57	6
No Address in Narrative	1	67	123	257	2,679	50	5
Different Street	1	41	94	660	12,201	35	4
All Cases	1	33	52	87	12,201	1,004	100
<i>Google Rooftop Differences (n = 994)*</i>							
Same Address	3	46	79	159	10,863	854	86
Same Street, Different Address	7	42	65	147	1,248	56	6
No Address in Narrative	6	62	145	292	2,685	49	5
Different Street	9	47	141	597	12,183	35	4
All Cases	3	46	81	165	12,183	994	100

*There are 10 fewer cases in the google rooftop results, as 10 cases were not geocoded within Wilmington.

Figure 1 shows using a dot plot that displays the entire distribution of those errors, according to each category, on a log base ten scale. Based on this graph one can see that the errors are approximately normally distributed on the log scale for cases in which there was no transcription error in the address (Zandbergen et al., 2012). The other categories appear to be heavy tailed, and so the log transformation appears to be appropriate, although are too sparse to determine whether they are normally distributed.

Even ignoring the No Address in Narrative category, these results show that over 10% of the sample have noticeable human transcription errors in the listed address relative to the narrative. While in most cases it appears to only add a slight amount of geocoding error to the results, in outlying cases can result in extreme errors in locating the incident. Even in the case with the same address the maximum error distance is not trivial, with an outlying case of over 3,000 feet, as well as several cases of error distances of over 1,000 feet.

Figure 1: Dot plot of distance in feet between shooting location (based on sensor data) versus location based on street centerline address geocoding. Different types of address errors were based on the incident narrative.

Table 2 displays regression estimates for each of the four categories based on the natural logarithm of the error distances, for both the street centerline and the Google geocoding results. While the regression is estimated on the logged error distances, the confidence intervals of the mean error are displayed in feet for easier interpretation. Because the transformation is monotonic, the confidence intervals of the mean logged error can simply be exponentiated and still have the correct coverage on the feet scale (Duan, 1983).

This shows that the logged median error slightly increases for each of the different categories for the street centerline data, going from same address, same street different address, no address in narrative, and different street. It also shows that the error distances for google rooftop tend to be slightly higher than the street centerline for each associated category, with the exception of the same street different address is smaller for google rooftop. But, given the size of the standard errors, a contrast hypothesis test comparing errors for centerline vs google would not be statistically significant for any of the categories.

Table 2: Average distances in feet between sensor data locations, and types of geocoding error. Provided for both street centerline and Google Rooftop based address geocoding.

	<i>Mean*</i>	<i>Std. Error</i>	<i>95% CI Low**</i>	<i>95% CI High**</i>
<i>Street Centerline Differences (n = 1,004)</i>				
Same Address	3.91	0.04	47	54
Same Street, Different Address	4.34	0.14	59	100
No Address in Narrative	4.88	0.15	98	175
Different Street	5.09	0.17	116	229
<i>Google Rooftop Differences (n = 994)</i>				
Same Address	4.42	0.04	77	89
Same Street, Different Address	4.28	0.14	55	95
No Address in Narrative	4.96	0.15	106	192
Different Street	5.39	0.18	154	313

*Mean and Standard Error are on the natural log of the error in feet scale. **CI Low and High are the 95% confidence intervals of the mean on the original feet scale.

Note the confidence intervals provide an error estimate of the *mean* error distance. It does not give the probability that any individual case will be that far apart, which requires a prediction interval (Hyndman & Athanasopoulos, 2018). Given that the more extreme quantiles do not appear to perform as well in descriptive statistics for log-normal data (Zandbergen, 2008), we might expect this to apply to the prediction intervals based on the prior regression. As such, we use a non-parametric resampling approach to formulate the prediction intervals (Lei et al., 2018). Table 3 provides the 95% upper prediction interval for events overall, as well as for each category and for both street centerline and rooftop geocoders based on the prior regression model estimates. What this number signifies is that for future shooting events, one would expect 95% of the cases to have an error distance of under that amount.

This can be interpreted in two ways. One is that under the current status quo of recording shooting incidents, one does not know whether an event will be correctly inputted, and so the upper prediction interval for *all cases* is the best guess of the upper bound of the error for the majority of cases. Another though is that it is clearly within the police departments control for whether a case is assigned the correct address. So if the Wilmington Police Department takes steps to ensure the correct address is inputted into the address field, they can limit the potential error in locating the shooting incident to typically under 261 feet using a street centerline geocoder versus 312 feet for the general error distance.

Table 3: Upper 95% Prediction Intervals for the Error Distance (in Feet)

	<i>Street Centerline</i>	<i>Google Rooftop</i>
Same Address	261	449
Same Street, Different Address	410	397
No Address in Narrative	710	802
Different Street	903	1,251
<i>All Cases</i>	<i>312</i>	<i>476</i>

Figure 2 displays a map of the spatial coordinates of the shooting incidents, placed at the location of the GDT detection. Points are displayed as graduated circles according to the log of the error distance to the street centerline address, to visualize potential spatial patterns in the geocoding errors. Like other cities, Wilmington only has GDT coverage in a small part of the city (Irvin-Erickson et al., 2017).

While spatial autocorrelation is not easily visible in the map, examining the Moran's I statistic (using an inverse distance weighted, row normalized spatial weights matrix), is positive and statistically significant with a value of 0.14. This value is the same for both the street centerline and the Google geocoding results. Examining the distribution, areas with less dense streets towards the north area of the map appear to have further errors on average, which is likely a limitation of using a street centerline geocoder. While all of the study area is an urban area, it is the same fundamental problem of geocoding using street centerlines in rural areas – longer street segments result in larger interpolation errors (Zandbergen, 2011).

Shooting Locations Based on Sensor Data

Log Error Distance (Feet)

- -0.4 to 2.7
- 2.8 to 3.7
- ◐ 3.8 to 4.6
- ◑ 4.7 to 6.0
- 6.1 to 9.4

Figure 2: Map of the shooting locations in Wilmington, North Carolina. Points are located according to sensor data, and graduated circles are based off of error distances to the street centerline address geocoded data.

Discussion

The results of this paper show that while geographical inaccuracy of common geocoding practices for police data exist, the problem is mostly fairly minor as long as the address was recorded correctly to begin with. While this means that there indeed can be issues with the reliability of police data, it suggests that the solution to the problem lies with better recording practices in the police, rather than with the technical solutions used to transform an address to a location on a map.

These findings replicate some of the findings of Gerell (2018), the most similar prior work assessing the positional accuracy using data on the exact locations of criminal events, although the errors here are a bit smaller. The median error found here is 52 feet (16 meters) for comparison to street centerline geocoded addresses, and 81 feet (25 meters), as compared to the 250 feet (83 meters) found from the Swedish study. While Gerell (2018) speculated on different types of error that could explain why the difference between actual and geocoded locations, he offered no data for the explanation. In the current study the use of incident narratives helps shed more light on this issue. When the police have an accurate address recorded in the police report the error is quite low. When the police record the correct street, but wrong address on that street the median error (for street centerline geocoding) jumps to 65 feet (20 meters). When recording on a totally different street the median error jumps to 94 feet (29 meters). Given that estimates of GDT technology can only pinpoint events to within 30 to 75 feet (Aguilar, 2015), to expect any lower error for when the correct address is noted on average is unrealistic, as we cannot estimate the error due to the initial GDT provided location in this analysis.

Two additional findings are worth noting. First, the analysis showed that the Google geocoding results were on average farther away from the true locations than the street centerline results. This is not surprising, given that Google often returns rooftop based geocoding results. This result seems likely to extend to other crimes that occur outdoors. Even if they are associated with a particular address for simplicity, thus counter to many arguments street centerline geocoding may be more accurate in some circumstances, as opposed to parcel based geocoding (Ratcliffe, 2001; Zandbergen, 2011).

The second is that positional error distances do appear to follow a log-normal distribution (Zandbergen, 2008). Thus we provide analysis of the logged errors, as well as estimate non-parametric prediction intervals for error distances. These results show that the upper 95% prediction intervals are much larger than median errors typically reported, with the upper 95% interval at 261 feet (80 Meters) for street centerline geocoding and 449 feet (137 Meters) for the Google geocoding results. These are the results for when the address and the incident narrative report the same location, so are under the best case error scenario. If using these more robust error metrics, the use of GDT appears more reasonable, although we cannot say what is an acceptable amount of error for the typical types of crime mapping applications to take place within a police department. While we did not conduct analysis of how such error resulted in loss in efficiency in typical analyst products (Mazeika & Uriarte, 2018; Zandbergen et al., 2012), it seems not an unreasonable amount of error to live with for crime analysts. Given it is popular

currently to aggregate crime incidents to street segments to conduct analysis (Weisburd et al., 2004; Wheeler et al., 2016), such aggregation will likely make such error immaterial. The exception is if the coding would result in the wrong street segment to be registered, as discussed regarding polygon aggregation by (Ratcliffe, 2001), but as long as the input address is correct this is likely to be a fairly uncommon problem. Focusing on address level or exact repeat analysis though such error will become more problematic (Block and Block, 1995; Eck, Clarke, & Guerette, 2007). Also given that much crime data is disseminated at larger aggregations to preserve privacy (Field, 2011; Singleton & Brunsdon, 2014; Thompson et al., 2015), it suggests this aggregation may be reasonable, given that lower address level dissemination is likely to contain some error.

There are several limitations that should be kept in mind when interpreting the results. One is that the nature of the GDT may change police behavior in recording incident addresses, and this could impact the findings. One scenario would be that given the GDT location is already provided, the incident address is more easily assigned, and as such GDT provides an *underestimate* of the positional error due to street centerline or address based geocoding. If this is the case however, it may still be reasonable to generalize the positional accuracy estimates to other crimes that occur outdoors, such as interpersonal robberies, which need to have a visible witness to report the event.

A second aspect in which the use of GDT may bias the results is that given the incident location is available, those inputting the address locations for the crime incident are not as careful. If this is the case, this could result in a larger proportion of cases that were coded as having some transcription errors, and could result in larger distance errors. We cannot say whether the resulting 10% of cases that had transcription errors is the norm, but prior analysis of geocoding accuracy would not have uncovered this type of error (with perhaps the exception of Mazeika & Summerton (2017) if they were unable to find a building at a listed address).

One way future research on crime events may tackle this problem is via forensic evidence collected in the field, such as taking photographs that can be geo-stamped to the specific location, or using GPS traces on officers (Wain & Ariel, 2014). Either of those data sources provides a potential gold standard of where the event took place, but would be less likely to bias data entry in a way that knowing the GDT is already recorded may for this analysis.

A simpler strategy to prevent such errors in the police department itself though is to put in place a mechanism to audit and correct such transcription errors when they are found. While it is the case an automatic system can be put into place when a true location is known, this is not feasible for the majority of crime incidents. It is within any police departments capabilities to ensure that correct address coding occurs when digitizing crime incidents though. While frequent audits may not be realistic for less serious crimes of high volume, it could be argued to be well worth the effort for serious crimes with smaller volumes such as the gun violence studied in the present study.

Given this small of an error, which is overall consistent with prior work on assessing positional error in geocoded crime incidents (Gerell, 2018; Mazeika & Summerton 2017; Ratcliffe, 2001), this research does not provide strong evidence that GDT is necessary to effectively map gun shot incidents relative to the more typical address based methods. Given that additional work has questioned that the number of false positives with the technology are

substantial (Ratcliffe et al., 2019), the empirical results of either the necessity or cost-effectiveness for GDT are not conclusive. The findings here suggest that simpler audits of address based systems are likely an easier (and cheaper) alternative for police departments to increase the accuracy of address based geocoding incidents.

References

- Aguilar, J.R. (2015). Gunshot detection systems in civilian law enforcement. *Journal of the Audio Engineering Society* 63, 280-291.
- Bichler, G., & Balchak, S. (2007). Address matching bias: Ignorance is not bliss. *Policing: An International Journal of Police Strategies & Management* 30, 32-60.
- Bieler, S., & La Vigne, N. (2014). Close-Range Gunfire around DC Schools. *Urban Institute*.
- Block, C.R., & Block, R.L. (2011). Case Study 7.1: Forty years of codifying and mapping homicides in Chicago: Impacts on police, research, and community well-being. In *Public Sociology: Research, Action, and Change*. Ed. Nyden, P., Hossfeld, L., & Nyden, G. pgs. 245-251. Sage.
- Block, R.L., & Block, C.R. (1995). Space, place and crime: Hot spot areas and hot places of liquor-related crime. *Crime Prevention Studies* 4, 145-184.
- Cohen, J. (2006). Development of crime forecasting and mapping systems for use by police. *NCJRS*, 2001-IJ-CX-0018. <http://www.justicestudies.com/pubs/forecasting.pdf>
- Carr, J., & Doleac, J. L. (2016). The geography, incidence, and underreporting of gun violence: new evidence using ShotSpotter data. *Incidence, and Underreporting of Gun Violence: New Evidence Using Shotspotter Data (April 26, 2016)*.
- Carr, J. B., & Doleac, J. L. (2018). Keep the Kids Inside? Juvenile Curfews and Urban Gun Violence. *Review of Economics and Statistics*, 100(4), 609-618.
- Drange, M. (2016). ShotSpotter alerts police to lots of gunfire, but produces few tangible results. *Forbes*. <https://www.forbes.com/sites/mattdrange/2016/11/17/shotspotter-alerts-police-to-lots-of-gunfire-but-produces-few-tangible-results>
- Duan, N. (1983). Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association* 78, 605-610.
- Eck, J.E., Clarke, R.V., & Guerette, R.T. (2007). Risk facilities: Crime concentration in homogeneous sets of establishments and facilities. *Crime Prevention Studies* 21, 225-264.
- Field, K. (2011). Editorial: Reflections on a Criminal Crime Map. *The Cartographic Journal* 48, 1-3.
- Gerell, M. (2018). Quantifying the Geographical (Un)reliability of Police Data. *Nordisk Politiforskning*, 5(2), 157-171, [10.18261/issn.1894-8693-2018-02-05](https://doi.org/10.18261/issn.1894-8693-2018-02-05)

- Hart, T.C., & Zandbergen, P.A. (2013). Reference data and geocoding quality: Examining completeness and positional accuracy of street geocoded crime incidents. *Policing: An International Journal of Police Strategies & Management* 36, 263-294.
- Hyndman, R.J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. OTexts: Melbourne, Australia.
- Irvin-Erickson, Y., La Vigne, N., Levine, N., Tiry, E., & Bieler, S. (2017). What does Gunshot Detection Technology tell us about gun violence? *Applied geography*, 86, 262-273.
- Lawrence, D.S., La Vigne, N.G., Goff, M., & Thompson, P.S. (2018). Lessons learned implementing gunshot detection technology: Results of a process evaluation in three major cities. *Justice Evaluation Journal* 1, 109-129.
- Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R.J., & Wasserman, L. (2018). Distribution-free predictive inference for regression. *Journal of the American Statistical Association* 113, 1094-1111.
- Levine, N., & Kim, K.E. (1998). The location of motor vehicle crashes in Honolulu: A methodology for geocoding intersections. *Computers, Environment and Urban Systems*, 22, 557-576.
- Loeffler, C., & Flaxman, S. (2018). Is gun violence contagious? A spatiotemporal test. *Journal of Quantitative Criminology* 34, 999-1017.
- Mazeika, D. & Summerton, D. (2017). The impact of geocoding method on the positional accuracy of residential burglaries reported to police. *Policing: An International Journal of Police Strategies and Management*, 40(2), 459-470
- Mazeika, D. M., & Uriarte, L. (2018). The near repeats of gun violence using acoustic triangulation data. *Security Journal*, 1-21.
- Mazerolle, L. G., Frank, J., Rogan, D., & Watkins, C. (2000). A field evaluation of the ShotSpotter gunshot location system: Final report on the Redwood City field trial. *US DOJ report*.
- Pizzi, J. (2015) ShotSpotter: Wilmington gunshots decrease 42 percent. *Delaware Online*. <https://www.delawareonline.com/story/news/crime/2015/07/06/wilm-officials-fewer-gunshots-detected-downtown/29793477/>.
- Ratcliffe, J. (2001). On the accuracy of TIGER-type geocoded address data in relation to cadastral and census areal units. *International Journal of Geographical Information Science* 15:5, 473–486, <https://doi.org/10.1080/13658810110047221>
- Ratcliffe, J. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science*, 18(1), 61–72. <https://doi.org/10.1080/13658810310001596076>
- Ratcliffe, J. H., Lattanzio, M., Kikuchi, G., & Thomas, K. (2019). A partially randomized field experiment on the effect of an acoustic gunshot detection system on police incident reports. *Journal of Experimental Criminology*, 15(1), 67-76.

- Roongpiboonsopit, D. and Karimi, H. A. (2010a). Comparative evaluation and analysis of online geocoding services. *International Journal of Geographical Information Science*, 24, 1081-1100.
- Roongpiboonsopit, D. and Karimi, H. A. (2010b). Quality assessment of online street and rooftop geocoding services. *Cartography and Geographic Information Science*, 4, 301-318.
- Singleton, A., & Brunson, C. (2014). Escaping the pushpin paradigm in geographic information science: (re)presenting national crime data. *Area* 46, 294-304.
- Tompson, L., Johnson, S., Ashby, M., Perkins, C., & Edwards, P. (2015). UK open source crime data: Accuracy and possibilities for research. *Cartography and Geographic Information Science*, 42(2), 97–111. <https://doi.org/10.1080/15230406.2014.972456>
- Sherman, L., Gartin, P., & Buerger, M. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology* 27, 27-56.
- Wain, N., & Ariel, B. (2014). Tracking of police patrol. *Policing: A Journal of Policy and Practice* 8, 274-283.
- WECT News 6. (2014). WPD asks for expansion of ShotSpotter technology, more coverage. <http://www.wect.com/story/24855532/wpd-asks-for-expansion-of-shotspotter-technology-more-coverage/>.
- Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology* 53, 133-157.
- Weisburd, D., Bushway, S., Lum, C., & Yang, S.M. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology* 42, 283-322.
- Weisburd, D., Mastrofski, S., McNally, A.M., Greenspan, R., & Willis, J. (2003). Reforming to preserve: Compstat and strategic problem solving in American policing. *Criminology and Public Policy* 2, 421 – 456.
- Wheeler, A.P. (2017). Geospatial analytics. In *SPSS® Statistics for Data Analysis and Visualization*. Eds. McCormick, K., & Salcedo, J., with contributions by Peck, J. and Wheeler, A.P. Wiley: Indianapolis, IN.
- Wheeler, A.P., & Kovandzic, T.V. (2018). Monitoring volatile homicide trends across U.S. cities. *Homicide Studies* 22, 119-144.
- Wheeler, A.P., Worden, R.E., & McLean, S.J. (2016). Replicating group-based trajectory models of crime at micro-places in Albany, NY. *Journal of Quantitative Criminology* 32, 589-612.
- Zandbergen, P.A. (2008). Positional accuracy of spatial data: Non-normal distributions and a critique of the national standard for spatial data accuracy. *Transactions in GIS* 12, 103-130.
- Zandbergen, P.A. (2009). Geocoding quality and implications for spatial analysis. *Geography Compass* 3, 647-680.

- Zandbergen, P.A. (2011). Influence of street reference data on geocoding quality. *Geocarto International* 1, 35-47.
- Zandbergen, P.A., & Hart, T.C. (2009). Geocoding accuracy considerations in determining residency restrictions for sex offenders. *Criminal Justice Policy Review* 20, 62-90.
- Zandbergen, P.A., Hart, T.C., Lenzer, K.E., & Camponovo, M.E. (2012). Error propagation models to examine the effects of geocoding quality on spatial analysis of individual-level datasets. *Spatial and Spatio-Temporal Epidemiology* 3, 69-82.