Criminal Justice
Policy Review
Volume 20 Number 1
March 2009 62-90

10.1177/0887403408323690
http://cjp.sagepub.com
hosted at
http://online.sagepub.com

Geocoding Accuracy Considerations in Determining Residency Restrictions for Sex Offenders

Paul A. Zandbergen
University of New Mexico, Albuquerque
Timothy C. Hart
University of Nevada, Las Vegas

Geocoding is commonly employed to determine the location of addresses for use in spatial analysis, including the establishment of residency restriction zones for sex offenders. Street geocoding has known limitations in terms of positional accuracy, which may adversely affect spatial analytic methods. A case study on sex offenders in Orange County, Florida, is used to characterize the positional accuracy of street geocoding and its impact on spatial analysis. Positional accuracy of street geocoded locations of sex offenders' residences, schools, and day care facilities was determined using Geographic Information Systems (GIS) by measuring the distance to the correct property boundaries. Results show that positional errors in street geocoding are substantial and may bias conclusions drawn from proximity analysis. Findings strongly suggest that street geocoding is not appropriate for assessing residency restriction violations for sex offenders. These findings have important implications for criminal justice policies related to residency restrictions for sex offenders.

Keywords: *street geocoding; parcel geocoding; positional accuracy; Geographic Information Systems (GIS)*

Law enforcement agencies throughout the United States rely on Geographic Information Systems (GIS) as a valuable tool in combating crime. Results from a government survey show that nearly two thirds of the nation's largest law enforcement agencies use GIS to map reported crimes, more than one half use GIS to map calls for service, and more than two-in-five use GIS to map arrest data (Reaves & Hart, 2000). GIS is also used by law enforcement for the monitoring and tracking of registered sex offenders, especially within jurisdictions that have enacted laws that prohibit registered sex offenders from living within a specified distance of places where children congregate (i.e., residency restrictions laws). Many GIS applications

Authors' Note: Please direct all correspondence to Paul A. Zandbergen, Department of Geography, Bandelier West Room 111, MSC01 1110, 1 University of New Mexico, Albuquerque, NM 87131; e-mail: zandberg@unm.edu.

by law enforcement agencies include the use of address geocoding, a technique that assigns coordinates to a street address. The use of geocoding has become widespread, and it is arguably one of the most successful applications of GIS. As with other techniques used to create spatial data, there is some uncertainty associated with the results of geocoding, which may affect its usefulness. Uncertainty in geospatial information is a fundamental issue that continues to attract attention as a research challenge (see Couclelis, 2003; Foody & Atkinson, 2002; MacEachren et al., 2005; Zhang & Goodchild, 2002). Although uncertainty in geospatial information has many components (Chrisman, 1991), completeness and positional error are of particular interest in the case of geocoding. This article explores the positional error of geocoding and its impact on the analysis of residency restrictions for registered sex offenders. The following sections review the literature on residency restrictions as well as positional errors in geocoding.

Registration and Residency Restrictions for Sex Offenders

When Jesse Timmendequas was convicted of the rape and murder of 7-year-old Megan Kanga more than a decade ago, it prompted nationwide demand for tighter control of known sex offenders and predators. The culmination of these efforts helped create Megan's law, which set forth requirements for states to manage and track sex offenders and predators. To date, all 50 states have enacted variations of Megan's law. Legislative responses to dealing with sex offenders not only require registration and community notification but some also place restrictions on where they can live (Welchans, 2005).

To date, 29 states have enacted legislation that establishes some form of residency restrictions that prohibit sex offenders from living within close proximity to places where children gather (Council of State Governments, 2007). Although some states leave it up to parole boards to determine the specific nature of these restrictions, most statutes include specific distances and specific locations to which these distances apply. Distances range from 500 ft to 2,000 ft, with 1,000 ft being the most common. The most common locations are schools and day care facilities, but some states add locations such as parks, playgrounds, and school bus stops.

Florida statutes prohibit certain offenders[1] from living within a 1,000-ft radius of sensitive facilities, such as schools, day care centers, parks, and any place where children congregate. When a sex offender or predator moves within Florida or into Florida from another state, the Florida Department of Corrections (FDOC) or the county sheriff's office must be notified.[2] Upon notification, the local law enforcement agency must verify whether the offender's or predator's registered address is outside any 1,000-ft restricted buffer. Within 48 hr of notification, agencies are also required to inform public schools and licensed day care facilities that a sexual predator resides within a 1-mile radius.[3] Finally, a probation officer from the Department of Corrections is required to periodically make contact with the offender or predator to verify whether they reside at their registered address and are not in violation of any other court order.

Although current Florida statutes prescribe that sex offenders and predators must live more than 1,000 ft from places where children gather, many local jurisdictions have gone further and have imposed limits of 1,500 ft or 2,500 ft. Currently, there are approximately 40 local ordinances in effect throughout the state of Florida with more restrictive residency requirements for sexual offenders and predators. The American Civil Liberties Union of Florida is concerned these laws are too restrictive, preventing sex offenders from residing anywhere within some of these jurisdictions. Indeed, a recent study shows that Florida's residency restriction law has significantly reduced the housing options for registered sex offenders (Zandbergen & Hart, 2006).

How Is Distance Determined?

Chapter 948.30(1b) of Florida's Criminal Procedure and Corrections code states that if the victim of certain crimes[4] is under the age of 18, the court must impose on the convicted offender "a prohibition on living within 1,000 ft of a school, day care center, park, playground, or other place where children regularly congregate." In the original statute, only schools, day care facilities, parks, and playgrounds were specifically mentioned, but in 2004 public school bus stops were added (F.S.S. 947.1405 [7a][2]).

Chapter 948.30(1b) of Florida's Criminal Procedure and Corrections code also describes how the distance is to be determined:

> The 1,000-feet distance shall be measured in a straight line from the offender's place of residence to the nearest boundary line of the school, day care center, park, playground, or other place where children congregate. The distance may not be measured by a pedestrian route or automobile route.

Local law enforcement officials have employed a number of techniques to determine whether the 1,000-ft (or other) distance requirement is violated. Typical field-based techniques include the use of a distance wheel, a laser device, global positioning system (GPS) unit, or traditional surveying methods. Alternatively, other jurisdictions have turned to GIS-based methods. For these agencies, sex offender data as well as data representing where children regularly congregate are address geocoded. Agencies not already using GIS to determine residency restriction may adopt GIS for this purpose in the future as they may already be using GIS for other purposes while other methods are very costly and time consuming.

If accurate and up-to-date parcel boundaries were readily available in GIS format for all areas where children regularly congregate and for all residences of sex offenders, the determination of the residency restrictions would be relatively straightforward using a simple buffer function in GIS. However, this type of data is often not readily available, and the use of street geocoding of addresses is much more common (Goldberg, Wilson, & Knoblock, 2007; Rushton et al., 2006). Given the potential inaccuracies of the various types of geocoding, a closer examination of these techniques is required.

Address Geocoding

Address geocoding is the process of creating an XY location of an address (i.e., associating an address record with a point on a map). In the most common approach to address geocoding, a street network is represented as street line segments that hold street names and the range of house numbers on each side of the street. Geocoding is accomplished by first matching the street name, then the segment that contains the house numbers, and finally placing a point along the segment based on linear interpolation within the range of house numbers. An optional offset can be employed to show on which side of the street line segment the address is located. Although this is not the only method of address geocoding, it remains the most widely employed technique among law enforcement agencies (Bichler & Balchak, 2007; Harries, 1999).[5] This technique is referred to as street geocoding.

There are many potential problems with street geocoding, which have been well described in the literature (Goldberg et al., 2007; Harries, 1999; Krieger, Waterman, Lemieux, Zierler, & Hogan, 2001; Ratcliffe, 2001; Rushton et al., 2006). For the results of street geocoding to be meaningful, the geocoding process needs to meet certain quality expectations. Recent research on the quality of geocoding has emphasized a consideration of completeness, positional accuracy, and repeatability (Whitsel et al., 2004). These three components can be used to characterize the overall quality of any geocoding result. Completeness is the percentage of records that can be reliably geocoded, also referred to as the match rate or hit rate. Positional accuracy indicates how close each geocoded point is to the true location of the address. Repeatability indicates how sensitive the geocoding results are to variations in the street network input, the matching algorithms of the geocoding software, and the skill/interpretation of the analyst. Of these three components, the positional accuracy will be discussed in more detail and is the focus of the current study.

Several studies have determined quantitative estimates of the positional accuracy of geocoding. Estimates of typical positional errors for residential addresses range from 25 to 168 meters (Bonner et al., 2003; Cayo & Talbot, 2003; Dearwent, Jacobs, & Halbert, 2001; Karimi & Durcik, 2004; Ratcliffe, 2001; Schootman et al., 2007; Strickland, Siffel, Gardner, Berzen, & Correa, 2007; Ward et al., 2005; Whitsel et al., 2006; Zandbergen, 2007; Zhan, Brender, De Lima, Suarez, & Langlois., 2006; Zimmerman, Fang, Mazumdar, & Rushton, 2007) based on median values of the error distribution. Results in urban areas are generally more accurate than in rural areas (Bonner et al., 2003; Cayo & Talbot, 2003; Ward et al., 2005). It should also be noted that the occurrence of major positional errors is relatively common. For example, in one of the more thorough studies by Cayo and Talbot (2003), 10% of a sample of urban addresses was geocoded with errors larger than approximately 96 m and 5% was geocoded with errors larger than 152 meters. For rural addresses these distances were 1.5 and 2.9 km, respectively.

The research on the positional accuracy of street geocoding reveals a number of important points. First, a number of different techniques are employed to determine the positional accuracy, including the use of field measurements using GPS, the centroid of the parcel, and the location of the residence using aerial photography. Each of these methods makes different assumptions about what the correct location is, and each method has its own inherent error. Second, different statistical estimates are used to characterize the positional accuracy, including the range, mean, median, standard deviation, root mean square error (RMSE), and percentiles (75th, 90th, 95th, etc.). There is no agreement in the published research on the most meaningful statistic; however, several studies that report the actual distribution of error estimates (Cayo & Talbot, 2003; Karimi & Durcik, 2004; Zandbergen, 2008) suggest that this distribution is log-normally distributed. This indicates that traditional statistics, such as range, mean, standard deviation, and RMSE, employed by several studies are not very meaningful. Third, positional accuracy is partly dependent on population density: street geocoding results in urban areas are more accurate than in rural areas due to the nature of rural residential addresses (larger parcels as well as a more uneven size distribution of parcels). Collectively, the current body of knowledge suggests that the positional errors in street geocoding can be very substantial and needs to be characterized in a meaningful manner relevant to the use of the street geocoding results.

The positional error in geocoded addresses may adversely affect spatial analytic methods. Specific effects includes inflation of standard errors of parameters estimates and a reduction in power to detect such spatial features as clusters and trends (Jacquez & Waller, 2000; Waller, 1996; Zimmerman, 2008). Even relatively small positional errors can have an impact on local statistics for detecting clusters (Burra, Jerrett, Burnett, & Anderson, 2002). Research on this topic has been mostly confined to the health field. For example, typical street geocoding is not sufficiently accurate for the analysis of exposure to traffic-related air pollution of children at short distances of 250 m to 500 m (Zandbergen, 2007; Zandbergen & Green, 2007). Similar errors in misclassification of exposure potential have been identified by Whitsel et al. (2006). However, the effect of the positional error in geocoding has received very limited attention in the criminal justice literature.

Geocoding the locations of registered sex offenders is commonly accomplished using street geocoding. Almost all state agencies make the list of registered sex offenders available online, and these state registries have been integrated into the online National Sex Offender Public Web site (United States Department of Justice, 2007). Several of the these registries provide a direct link to the mapped location of

Figure 1
Screenshot of the California Sex Offender Locator Map

the residential address using an online geocoding service, such as MapQuest, Yahoo Maps, Google Maps, etc. Some registries have also developed interactive online GIS mapping systems to show not only the residential locations of registered sex offenders but also points of interest such as schools. Figure 1 shows an example screenshot of the California Sex Offender Locator Map (State of California Department of Justice, 2007), which displays sex offenders and public schools in a Web–GIS environment. Several online tools developed by third parties produce similar maps by obtaining the sex offenders registries from public agencies and creating a map mashup to display the locations. Figure 2 shows an example screenshot of a sex offender mapping tool developed by a private company (Vision 20/20, 2007), which displays the locations of sex offenders and public schools within Microsoft Virtual Earth. Displays can be generated in the form of roadmaps, using traditional satellite image views or in the form of oblique high resolution aerial photography. Although no tools are provided to determine the residency restrictions zones, the nature of the geographic information presented does suggest that determining residency restrictions is straightforward and reliable. The widespread availability of this geocoded

Figure 2
Screenshot of a Third-Party Online Sex Offender Mapping Tool

information raises the questions as to how accurate the geographic representations are and how reliable they are in the monitoring of residency restriction zones.

Published research studies that have investigated residency restriction zones have used both street geocoding and more accurate alternatives. Zandbergen and Hart (2006) used parcel geocoding exclusively to examine the effect of residency restrictions on housing availability. Grubesic, Mack, and Murry (2007) in a similar study employed mostly parcel geocoding, but they relied on street geocoding for approximately 10% of the sex offenders where parcel geocoding proved unsuccessful. Maghelal and Olivares (2003) relied exclusively on parcel geocoding to determine child safety zones. A study on residential proximity and sex offender recidivism in Minnesota (Minnesota Department of Corrections, 2007), however, used street geocoding in Google Earth to determine the locations of residential addresses. Finally, Clontz and Mericle (2004) employed street geocoding to determine clustering of sex offenders by neighborhood. This raises the question as to whether the studies that have employed street geocoding are sufficiently accurate to meet the requirements of the particular analysis.

Research Questions

The current study investigates how the positional error in street geocoding affects the determination of residency restrictions of registered sex offenders. More specifically, the following two research objectives are undertaken: (a) Determine the positional error in street geocoding of sex offenders' residences and the locations of schools and day care facilities and (b) determine the degree to which geocoding inaccuracies introduce error and bias in the results of the analysis of sex offender residency restrictions. Results of this research have very direct implications for how residency restrictions are determined and enforced as well as for assessing the utility of street geocoding for these types of applications.

Data and Method

The current study design relies on a comparison between two geocoding techniques: street geocoding and parcel geocoding.[6] For the purpose of this analysis, parcel geocoding is assumed to be very accurate and is therefore used as a control. The Euclidean distance between the street geocoded location and the parcel boundary is used as an error estimate. The parcel boundary is considered the most meaningful comparison in this case as the Florida statutes make specific reference to the use of property boundaries in determining residency restrictions. The effect of this positional error is determined by carrying out a 1,000-ft buffer analysis using various scenarios. The following sections describe in detail the sex offender, school, and day care facility data sets used in the current study, along with the analytic approach employed.

Sex Offenders

This data set contains the complete list of registered sex offenders for the state of Florida for 2003. The data contain selected fields denoting the physical address, physical description, date of birth, as well as information about the victim and offender. The data were originally released in 2003 by the Florida Department of Law Enforcement (FDLE) and street geocoded by Geoplan. Data were obtained in DBF file format, geocoded against ARC Logistics Route GDT Roads, and converted into a shapefile with the FGDL Albers HPGN coordinate system. Interactive matching was employed to catch any unmatched cases due to small errors in street naming, but only very reliable matches were included in this process. Unreliable matches and ties can result in large positional errors. Given the focus on the effect of positional error on proximity analysis, including only reliable matches reduces the potential confounding effects from large positional errors encountered in unreliable matches and ties.

From these geocoded locations for all of Florida, a subset of data was created for those offenders located within Orange County. Offenders registered at locations within a 3,000-ft buffer around the Orange County boundary were also included because preliminary analysis showed that inaccuracies in some street geocoded locations resulted in sex offenders residing in Orange County but placed just outside of the county boundary. The use of a small buffer captures these locations. Modifications to the overall data described above resulted in a preliminary set of 1,114 offenders. Offender status information was also used to select only those cases where the offender's reported address was expected to be a reliable indicator of his/her actual residence. Only those offenders whose status is (a) on administrative probation, (b) under community control, (c) under federal supervision, (d) on parole, (e) released from custody, (f) had their release revoked, or (g) under court-ordered supervision were included.[7] This excluded any case where the offender whose status was deceased, incarcerated, or fled from supervision. Only those cases where the victim was a minor were included as the residency restrictions only apply to these cases. Duplicates were also removed by identifying individual offenders by their tracking number, full name, and address. Many of these duplicates occurred because the offender was recorded as a predator after previously being registered as an offender. Other cases were simply duplicates. These cleaning steps resulted in a total of 744 sexual offender records.[8]

Addresses for the 744 records were exported to a DBF file and parcel geocoded using the Orange County 2004 1:2,000 parcel database. A total of 623 records were accurately matched. To calculate a realistic match rate, the 26 offenders who were not matched to a parcel but who were within the 3,000-ft buffer area immediately around the county border were removed from the comparison. This resulted in a match rate of 86.8% (623 out of 718).

Schools

The location of public and private schools in Florida in 2003 was obtained from Geoplan. The original data were provided by the Florida Department of Education and street geocoded by Geoplan. Data were obtained in DBF file format, geocoded against ARC Logistics Route GDT Roads, and converted into a shapefile with the FGDL Albers HPGN coordinate system. A subset of data was created that consisted of those schools located in Orange County. Schools located within a 3,000-ft buffer around the county border were also included. Schools identified as adult education, such as vocational/technical schools, community colleges, and universities were excluded, as well as any schools identified as inactive. In addition, several private schools were found to operate under multiple names at the same physical address. These duplicates were removed as well. This resulted in a total of 309 schools. Addresses for the 309 records were exported to a DBF file and parcel geocoded using the Orange County 2004 parcel database. A total of 261 records were accurately matched, including all public schools. To calculate a realistic match rate, the 12 schools that were not matched but that were within the 3,000-ft buffer area immediately around the county border were removed from the comparison. This resulted in a match rate of 87.8% (261 out of 297).

Day Care Facilities

The location of Florida day care facilities in 2003 was obtained from Geoplan. The original data were provided by the Florida Department of Children and Families and street geocoded by Geoplan. Data were obtained in DBF file format, geocoded against ARC Logistics Route GDT Roads, and converted into a shapefile with the FGDL Albers HPGN coordinate system. Several day care facilities were found to operate under multiple names at the same physical address; these duplicates were removed. This resulted in a total of 612 day care facilities. Addresses for the 612 records were exported to a DBF file and parcel geocoded using the Orange County 2004 parcel database. A total of 498 records were accurately matched. To calculate a realistic match rate, the 13 day care facilities that were not matched but that were within the 3,000-ft buffer area immediately around the county border were removed from the comparison. This resulted in a match rate of 82.9% (498 out of 601).

Comparing Street and Parcel Geocoding Results

The original street geocoded locations for sex offenders, schools, and day care facilities consisted of point locations placed along the street network without an offset. Parcel geocoding places a point in the centroid of the matched parcel. New data layers were created of the actual parcel boundaries that each address was matched with, resulting in new polygon data layers for sex offenders, schools, and day care facilities. In the end, three versions of geocoded locations for sex offenders, schools, and parcels were produced: (a) points of street geocoded locations, (b) centroids of parcel geocoded locations, and (c) polygons of parcel geocoded locations.

A total of 623 sex offenders resided in 571 unique parcels. Many of the offenders live at the exact same address, or at the same property, including shelters, apartment complexes, mobile home parks, hotels/motels, etc. The 261 schools are located on 254 unique parcels as several of the school sites include multiple schools, for example, an elementary and a middle school on the same property. Finally, the 498 day care facilities are located on 498 unique parcels.

For each of the three categories (sex offenders, schools, and day care facilities), the street and parcel geocoding results were compared. For each unique observation, the Euclidean distance between the street geocoded location and the nearest boundary of the corresponding parcel polygon was determined. In this analysis, the distance to the parcel boundary is used an estimate of the positional error of the street geocoded locations. Any street geocoded location that fell inside the correct parcel would have generated a distance of zero and was considered perfectly accurate. Summary statistics and cumulative distribution functions were used to characterize the positional error estimates.

Analysis of the 1,000-Ft Buffer Zones

One thousand feet buffers were created around each of the three locations for schools and day care facilities (street geocoded points, parcel centroids, and parcel polygons). Overlay analysis was conducted to determine whether the location of each sex offender's residence fell inside any of the 1,000-ft buffer zones; all three locations for each sex offender were considered: street geocoded points, parcel geocoded centroids, and parcel geocoded polygons. Only the results of the most meaningful combinations were used in the final analysis. In the comparison of results from different techniques, a kappa index of agreement (KIA) was determined.[9] The strength of agreement has been determined by Landis and Koch (1977): <0 (poor), 0-.20 (slight), .21-.40 (fair), .41-.60 (moderate), .61-.80 (substantial), and .81-1.00 (almost perfect).

The use of parcel centroids instead of parcel boundaries was also considered. Although Florida statutes specify the use of property boundaries, for reasons of computational simplicity or data availability, the use of parcel centroids has been suggested as an alternative. In some jurisdictions, property boundaries with address information are not available in GIS compatible format that allows for geocoding, and point representations may represent the only viable alternative. Many states, for example, have created databases of schools with XY locations representing some arbitrary location within the school grounds. In the absence of the actual boundaries, law enforcement may have no other option than to use the available data instead. In our analysis, for each of the three locations considered (sex offenders, schools, and day care facilities), the number of sex offenders residing within the 1,000-ft zone is determined for each possible polygon/centroid combination.

Results and Discussion

Figures 3 through 9 show example comparisons between the street and parcel geocoded locations of sex offenders, schools, and day care facilities. The first common scenario for sex offenders is shown in Figures 3 and 4 for single-family residential neighborhoods, where a single residential address is associated with a single residential parcel. Parcels are relatively small and fairly uniform in size. As can be seen in these examples, the street geocoded location is placed on the correct street segment, but the location along this segment relative to the parcel varies. Sometimes the location is placed almost directly in front of the parcel, but sometimes it is placed at the opposite end of the street. This incorrect placement reflects errors in the street number ranges in the street network data and inconsistent parcel numbering along a street segment. The effect of parcel size in these areas is relatively limited.

The second common scenario for sex offenders is shown in Figures 5 and 6 for multiunit neighborhoods, where multiple residential addresses are associated with a single parcel, such as apartment complexes and mobile home parks. Figure 5 shows

Figure 3
Comparison of Street and Parcel Geocoded Locations of a Sex Offender Residing in a Single-Family Residential Area

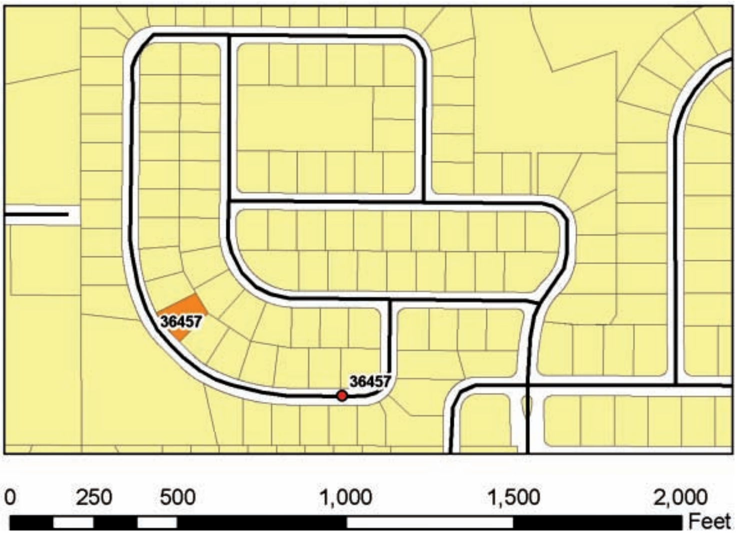


Figure 4
Comparison of Street and Parcel Geocoded Locations of Three Sex Offenders Residing in a Single-Family Residential Area

the example of a single sex offender residing in a mobile home park, whereas Figure 6 shows the example for 6 sex offenders living in the same mobile home park. Such multiunit residential areas in most cases do have a street network that allows for individual street geocoding of the residential addresses, but the properties are mostly rented and the legal boundary of the property is a single parcel. Multiple residential addresses can therefore be associated with a single parcel. Although it can be argued that the street geocoded locations may in fact be a more accurate representation of the physical location of the sex offender's residence, F.S.S. 948.30(1b) clearly indicates that the legal property boundary is to be used for assessing violations of residency restriction law. A total of 46 multiunits properties with at least 1 sex offender were identified, consisting of apartment complexes and mobile home parks. A total of 77 sex offenders were found to reside on these properties or 12.4% of all sex offenders in Orange County. In most of these 77 cases, the street geocoded locations fell inside the parcel, resulting in a positional error estimate of zero.

Example results for schools are shown in Figures 7 and 8. Schools are located mostly on moderate- to very large-size parcels. If the street segment is fairly short and the street geocoded location is placed on the correct street segment, this often results in the street geocoded location being placed very close to the school parcel and produces a very small positional error. In other cases, however, the street segment is much longer, or the street geocoded location is placed along the incorrect segment. In this scenario, a much larger positional error is produced.

Finally, example results for day care facilities are shown in Figure 9. Day care facilities are located mostly on small- to medium-size parcels in single-family residential areas. Many day care facilities are home based, and the results are very similar to the first scenario described above for sex offenders.

Positional Error of Street Geocoding

The positional error of street geocoded locations is determined as the Euclidean distance from the street geocoded location to the nearest boundary of the corresponding parcel. If the street geocoded location falls inside the parcel, the positional error is zero. Results of the positional error associated with sex offenders, schools, and day care facilities are shown in Figure 10 and summarized in Table 1. Figure 10 shows the cumulative distribution function for the positional error for the three categories. The distribution is created by ordering the results by the values for positional error and determining how many observations fall at or below each value. Within each category (sex offenders, schools, and day care facilities), all observations were considered equal. Results are plotted as a percentage of the total.

The cumulative distribution functions of the three categories reveal that the positional error estimates are very similar, up to the 65th percentile, after which the positional error values for schools are much larger than for sex offenders or day care facilities. At very large distances of several thousand feet, the percentiles for offenders

Figure 5
Comparison of Street and Parcel Geocoded Locations of a Sex Offender Residing in a Mobile Home Park

Figure 6
Comparison of Street and Parcel Geocoded Locations of Six Sex Offenders Residing in a Mobile Home Park

Figure 7
Comparison of Street and Parcel Geocoded Locations of Two Schools

Figure 8
Comparison of Street and Parcel Geocoded Locations of Three Schools

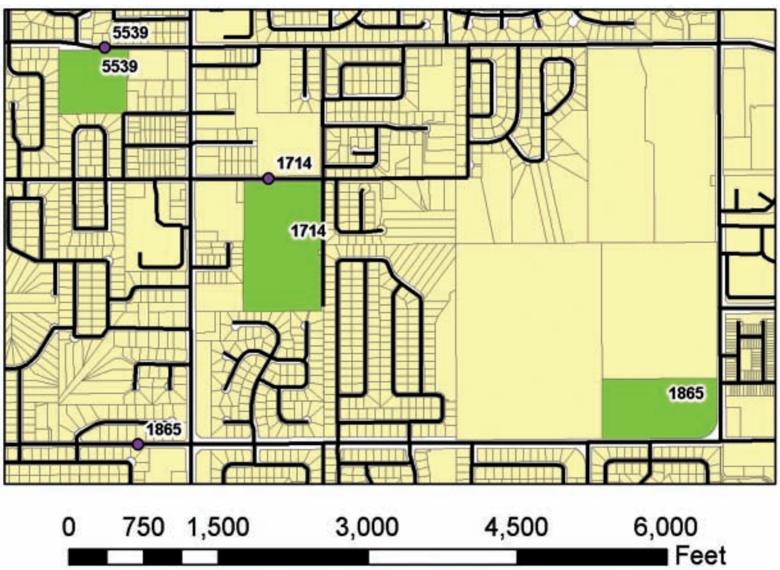


Figure 9
Comparison of Street and Parcel Geocoded Locations of Four Day Care Facilities

and day care facilities approaches 100%, whereas it is still around 97% for schools. Figure 10 also indicates that the positional error estimates are not normally distributed (a normal distribution would be represented by a symmetrical sigmoid curve) but are more like a log-normal distribution. This suggests that care should be taken in the interpretation of conventional statistics, such as the mean and standard deviation.

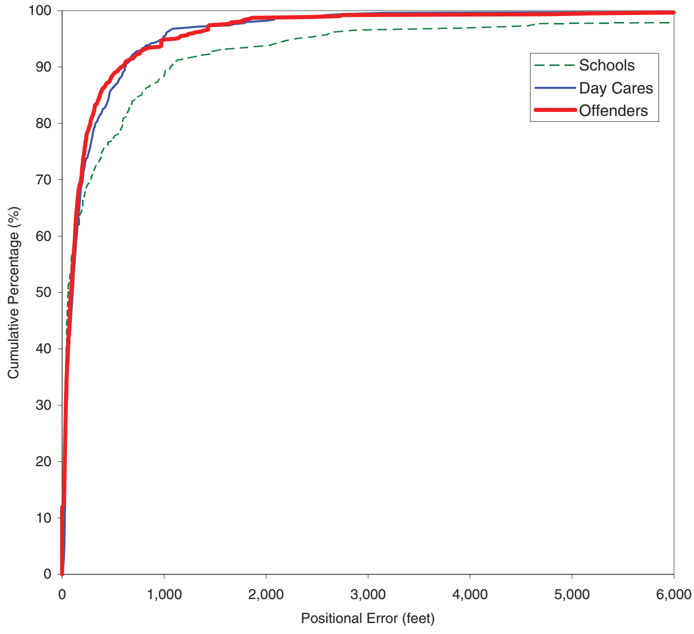
Descriptive statistics of the positional error estimates are provided in Table 1. The first observation is that the number of perfectly placed street geocoded locations (i.e., within the parcel boundaries) is 4 for schools, 4 for day care facilities and 74 for offenders. The high value for the offenders reflects the occurrence of multiunit properties (i.e., apartment complexes and mobile home parks) discussed previously. When excluding these properties, there were no perfectly placed street geocoded locations for offenders. The use of an offset from the street centerline in street geocoding results in a greater number of geocoded locations placed inside the correct parcel. A trial using an offset of 60 ft resulted in a total of 118 locations placed correctly. However, the overall effect of this offset on the distribution of positional errors was minimal, as for many locations the offset resulted in a very similar error and for some locations the error in fact increased. This is in agreement with previous studies that have determined the relationship between varying offsets and positional errors in geocoding (Cayo & Talbot, 2003; Zandbergen, 2007). The second observation is that the maximum value for all three categories is very large: approximately

Table 1
Positional Error Estimates of Street Geocoded Locations

	School Locations	Day Care Locations	Sexual Offenders' Residences
N	261	498	623
N with an error of 0	4	4	74
Positional error estimates in feet			
Min	0	0	0
Max	27,430	38,066	93,649
Mean	774	336	509
RMSE	3,260	1,816	4,826
Median	60	103	87
75th percentile	396	266	218
90th percentile	1,065	617	571
95th percentile	2,292	985	1,113

Note: RMSE = Root mean square error.

Figure 10
Cumulative Distribution Functions of Positional Error of Street Geocoded Schools, Day Care Facilities, and Sex Offenders



5 miles for schools, 7 miles for day care facilities, and 18 miles for offenders. These represent a very low number of cases where the street geocoding was completely incorrect. Review of these cases revealed that these errors can be attributed to incorrect street names in the street network database. This is common in street geocoding but usually limited to handful of cases. In our sample of 623 sex offenders, we identified that only 18 street geocoded locations were not placed along the same street segment as the correct parcel. All these locations have very large positional errors. Their small number, however, suggests that much of the positional error can be attributed to incorrect placement along the correct street segment. This interpolation error is inherent in the street geocoding algorithm and cannot be addressed by improved address cleaning.

Due to the nonnormal distribution of the positional error values, the mean and the RMSE, which are commonly used in error analysis, are not very meaningful as they are strongly influenced by the small number of very large errors. Instead, the use of percentiles is adopted to allow for a meaningful comparison between the results of the three types of locations. Similar to their use in the national standard for spatial data accuracy (NSSDA; Federal Geographic Data Committee [FGDC], 1998) and the national map accuracy standard (NMAS; United States Geologic Survey [USGS], 1999), the 90th and 95th percentile are chosen as the most meaningful characterization of the positional error estimates.

When considering the mean (50th percentile), the schools appear to be most accurate (60 ft), followed by offenders (87 ft) and day care facilities (103 ft). When considering the 75th percentile, schools become the least accurate (396 ft vs. 218 ft for offenders and 266 ft for day care facilities). This reflects the pattern already observed in Figure 10. Above the 65th percentile, the positional error of the schools is noticeably higher than for the other two. At the 90th percentile, the positional error estimate for schools is 1,065 ft or nearly double that of day care facilities (617 ft) and offenders (671 ft). A similar pattern occurs at the 95th percentile: 2,292 ft for schools, 985 for day care facilities, and 1,113 for offenders. Based on the results for the 90th and 95th percentiles, the accuracy of the geocoded locations for schools is much lower than for the other two, with the results for day care facilities and offenders so close that they can be considered equivalent in terms of accuracy.

Another look at Figures 3 through 9 helps explain why the positional error for the schools is so much higher. In general, parcels for schools are much larger than for day care facilities and offenders; multiunits parcels for offenders are an exception to this, but in these cases the positional error of the street geocoding is in fact very low due to the presence of a street network within the parcel. Such large parcels present difficulties in the street geocoding process as the algorithm to place a location along a street segment applies a linear interpolation within the street number range for the segment. Although a large parcel along a very short street segment results in low positional error, as the street segments get longer the error increases. This is reflected in the fact that the mean positional error for schools is lower than for the other two, but the values for the 90th and 95th percentiles are much higher. Parcels for offenders (other than multiunit) and day care facilities are mostly small and located in an area where parcel size along the same street segment is very uniform. In these cases, the linear interpolation provides more accurate and consistent results.

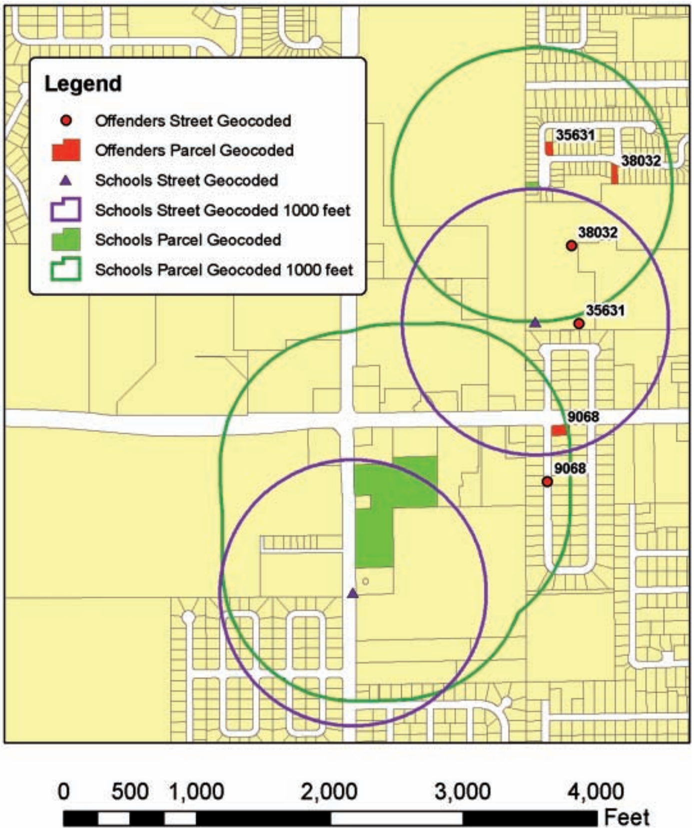
To recognize more clearly what the error estimates mean, they can be used to provide a formal description of positional accuracy. For example, the street geocoded location of a school is accurate to within 1,065 ft of the school boundary 90% of the time. The values for day care facilities and offenders are 617 and 571 ft, respectively. To put this in another perspective, the NMAS (USGS, 1999) can be applied to these results. NMAS states that 90% of test locations must fall within 1/50th of an inch on a printed paper copy of the map. This implies that the results for school geocoding are equivalent to the quality of a 1:639,000 printed map; for day care facilities and offenders, the results are 1:370,200 and 1:342,600, respectively. Again, the control in this case is the property boundary: anywhere on or within the property boundary would be considered perfectly accurate.

Analysis of the 1,000-Ft Residency Restrictions

The positional error estimates documented so far provide a strong indication that the analysis of the 1,000-ft residency restrictions using street geocoding is not very reliable. In the case of schools, the 90th percentile of the positional error is close to this 1,000-ft value, and the values for day care facilities and offenders are a bit more than half this value. Although it makes intuitive sense that such large errors do not allow for a distance-based analysis of this magnitude, the impact of the positional error in the street geocoded locations on the residency restrictions was quantified. In the first part of this analysis, the street geocoded point locations and parcel polygons were used (not the parcel centroid).

Figure 11 shows a close-up example result and will be used to describe how the analysis results were obtained. The street and parcel geocoded locations of two schools and three sex offenders' residences are shown in Figure 11. One thousand feet buffers were created around each school location. For offender 9,068, both the street and parcel geocoded locations fall within the 1,000-ft buffer of the parcel geocoded school location. The positional error in the street geocoded location of the offender's residence in this case does not affect the analytic result. However, both locations fall outside the buffer of the street geocoded school location due to its positional error. For offender 35,631, the street geocoded location falls outside the 1,000-ft buffer of the parcel geocoded school location, but the parcel geocoded location falls within it. It should be pointed out here that the parcel geocoded locations are considered the correct ones, allowing for an identification of Type I and Type II errors resulting from street geocoding. Type I errors include those offenders who reside within a 1,000 buffer zone but are not identified as such using street geocoding. Type II errors include those offenders who do not reside within a 1,000 buffer zone

Figure 11
Example Analysis of 1,000-Ft Residency Restrictions Using Parcel and Street Geocoding



but are identified as such using street geocoding. Both errors are obviously of concern, but Type I errors are expected to be much more common because a 1,000-ft buffer around a school parcel is often much larger than a 1,000-ft buffer around a point. The occurrence of Type I errors, therefore, can only be attributed partially to the positional errors in street geocoding, whereas the occurrence of Type II errors are strictly due to these positional errors.

As the discussion of Figure 11 makes clear, many different scenarios are possible, but only the most meaningful ones are quantified for the entire data set and described below. A summary of the results is presented in Table 2 for schools and Table 3 for day care facilities. In each case, a determination was made of how many

Table 2
Agreement Between 1,000-Ft Buffer Analysis Results for Schools Using Street and Parcel Geocoding Techniques

		Street Geocoding of Schools and Offenders		
		In	Out	Total
Parcel geocoding of schools and offenders	In	55	103	158
	Out	5	460	465
	Total	60	563	623
Kappa Index of Agreement: .424				

		Parcel Geocoding of Schools and Street Geocoding of Offenders		
		In	Out	Total
Parcel geocoding of schools and offenders	In	125	33	158
	Out	7	458	465
	Total	132	491	623
Kappa Index of Agreement: .820				

of the 623 sex offenders resided inside the 1,000-ft zone, and the agreement between the two methods was determined using the KIA. A KIA value of 1.0 would indicate perfect agreement between results.

Schools/sex offenders: parcel–parcel versus street–street. In this scenario, parcel geocoding for both schools and sex offenders is compared with street geocoding for both schools and sex offenders. Results indicate that parcel geocoding produces an outcome where many more sex offenders fall inside the 1,000-ft zone (158 vs. 60). This substantial increase is largely due to the fact that the 1,000-ft buffer zone around the parcel polygon is much larger than around the street geocoded points, in particular for schools. Of the 158 offenders living within 1,000 ft of a school, analysis using street geocoding only identifies 60 of them and 55 correctly: 5 sex offenders were found to live within the buffer zone using street geocoding that were outside the buffer using parcel geocoding (Type II error). The overall KIA value is 0.42, indicating only a moderate level of agreement.

Schools/sex offenders: parcel–parcel versus parcel–street. In this scenario, parcel geocoding for both schools and sex offenders is compared with parcel geocoding for schools and street geocoding for sex offenders. This represents a somewhat more realistic scenario: A correctly placed geocoded point location does not represent a school boundary very well; so even if a reliable parcel database is not available for an entire study area, the boundaries of schools might be available through other means, in particular for public schools. Results indicate that parcel geocoding

Table 3
Agreement Between 1,000-Ft Buffer Analysis Results for Day Care Facilities Using Street and Parcel Geocoding Techniques

		Street Geocoding of Day Care Facilities and Offenders		
		In	Out	Total
Parcel geocoding of day care facilities and offenders	In	161	85	246
	Out	10	367	377
	Total	171	452	623
Kappa Index of Agreement: .663				

		Parcel Geocoding of Day Care Facilities and Street Geocoding of Offenders		
		In	Out	Total
Parcel geocoding of day care facilities and offenders	In	184	62	246
	Out	12	363	377
	Total	196	427	623
Kappa Index of Agreement: .742				

(of offenders) also produces an outcome where many more sex offenders fall inside the 1,000-ft zone, although the result is not as biased as the previous case (158 vs. 132). Seven sex offenders were found to live within the buffer zone using street geocoding that were outside the buffer using parcel geocoding (Type II error). The overall KIA value is 0.82, indicating almost perfect agreement.

Day care facilities/sex offenders: parcel–parcel versus street–street. In this scenario, parcel geocoding for both day care facilities and sex offenders is compared with street geocoding for both day care facilities and sex offenders. Results indicate that parcel geocoding produces an outcome where many more sex offenders fall inside the 1,000-ft zone (246 vs. 171), but the relative difference is less compared to the same scenario for schools. This can be attributed to the fact that the parcels for day care facilities are much smaller than for school, resulting in a lower bias when using street geocoded point locations. The street geocoding results, however, continue to represent a major underestimation of the actual number of offenders living within the 1,000-ft zone. As with the schools, a number of offenders are incorrectly identified as living within the 1,000-ft zone; in this case, 10 were identified (Type II error). The overall KIA value is 0.66, indicating substantial agreement.

Day care facilities/sex offenders: parcel–parcel versus parcel–street. In this scenario, parcel geocoding for both day care facilities and sex offenders is compared with parcel geocoding for day care facilities and street geocoding for sex offenders. Results indicate that that parcel geocoding (of offenders) also produces an outcome where many more sex offenders fall inside the 1,000-ft zone (246 versus 196), but

Table 4
Comparison Between the Use of Parcel Boundaries and Centroids in 1,000-Ft Buffer Analysis

	Combination of School/Sex Offender Parcel Geometry			
	Polygon/Polygon	Polygon/Centroid	Centroid/Polygon	Centroid/Centroid
In	158	128	81	57
Out	465	495	542	566
	Combination of Day Care/Sex Offender Parcel Geometry			
	Polygon/Polygon	Polygon/Centroid	Centroid/Polygon	Centroid/Centroid
In	246	202	202	177
Out	377	421	421	446

the improvement is not as dramatic as for the schools. Twelve sex offenders were found to live within the buffer zone using street geocoding that were outside the buffer using parcel geocoding (Type II error). The overall KIA value is 0.74, indicating substantial agreement.

These results provide strong evidence of the error and bias introduced by the use of street geocoding for determining residency restriction violations for registered sex offenders. First, the use of street geocoding produces substantial underestimates of the actual number of sex offenders residing in the 1,000-ft zone, under all scenarios considered. Second, despite this underestimate, a small number of offenders are identified as living within the 1,000-ft whereas, in fact, they are not (Type II errors).

Use of Parcel Boundaries Versus Centroids

The final analysis presented is the use of parcel centroids instead of parcel boundaries. For each of the three locations considered (offenders, schools, and day care facilities), the number of offenders residing within the 1,000-ft zone is determined for each possible polygon/centroid combination. Table 4 presents a summary of the results. For both schools and day care facilities, the true number is represented by the polygon/polygon combination: 158 and 246 offenders, respectively. Logically, all other combinations result in lower numbers, and the highest underestimate is represented by the centroid/centroid combination. As expected, the effect of the use of centroids is strongest for schools (from 158 to 81), reflecting the large size of school parcels. The effect of the use of centroids for day care facilities and offenders is identical (a decrease from 246 to 202). These results indicate that using parcel centroids produces substantial underestimates of the number of offenders within the 1,000-ft zone, but no false positives are identified.

Conclusions

Street geocoding of the locations of schools, day care facilities, and sex offenders' residences was found to result in substantial positional errors. Under the assumption that the true location is best represented by the parcel boundaries, the 90th percentile of the positional errors is 1,065 ft for schools, 617 ft for day care facilities, and 571 ft for sex offenders' residences. These positional error estimates fall within the ranges reported by other research on the positional accuracy of street geocoding. The values for all three categories are considered very high relative to the need to reliably determine 1,000-ft residency restriction zones. The distribution of the positional errors precludes the use of conventional error estimates, such as the mean or the RMSE, which have been commonly used to report geocoding accuracy. The use of the 90th or 95th percentiles is proposed here as a more meaningful alternative.

The use of street geocoded locations results in substantial underestimates of the true number of offenders living within the 1,000-ft restricted zone around schools and day care facilities: 60 versus 132 for schools and 171 versus 196 for day care facilities. The use of street geocoded locations also results in a small number of Type II errors. As a result, the determination of residency restrictions using street geocoded locations should be considered very unreliable. The use of parcel centroids instead of property boundaries also results in substantial underestimates of the number of offenders residing within a 1,000-ft zone but no Type II errors. Given this major bias, the use of centroids is to be avoided not only for the locations represented by large parcels (such as the schools) but also those represented by small parcels (such as most offenders and day care facilities).

Findings from this study should be considered when residency restriction policies are created and applied. Residency restriction laws should explicitly state how geographic data should be geocoded for the purposes of assessing residency restriction violations. Such detailed language is not completely absent in current policies. Recall Chapter 948.30(1b) of Florida's Criminal Procedure and Corrections code:

> The 1,000-feet distance shall be measured in a straight line from the offender's place of residence *to the nearest boundary line* [italics added] of the school, day care center, park, playground, or other place where children congregate.

In crafting their residency restriction laws, Florida's policy makers considered the various ways the 1,000-ft distance could be measured. They attempted to reduce the ambiguity of the statute by spelling out how the measurement used to assess violations must be made.[10] Similar consideration should be given to how data is to be geocoded, when a GIS methodology is employed to assess residency restriction violations. The use of property boundaries is the preferred technique and produces reliable results. Therefore, the current study suggests that statues aimed at restricting sex offenders' residence contain language that requires a parcel geocoding technique, when a GIS application is used to assess such restrictions. Failure to do so opens the door to the use of techniques such as street geocoding that will produce unreliable assessments of violations.

Limitations of this study include the fact that the sample of sex offenders used in the analysis is not a complete and perfect sample of all sex offenders residing in Orange County. A number of records could not be geocoded with certainty due to incomplete or ambiguous address information. Although this presents a potential source of error, the objective of the study was to determine the accuracy of the street geocoding and its effect on applying the 1,000-ft rule, which does not require a complete and perfect sample of sex offenders. A second limitation is that not all sex offenders' residencies are restricted. Offenders who are released from custody or control of the FDOC, for example, are no longer on probation and therefore no longer have to abide by a residency restriction that was issued by the court as a part of their probation condition. Of all the sex offenders being considered in the study, only a small percentage must comply with the 1,000-ft rule. As this information is not easy to obtain and as the analysis was aided by a larger sample, all released offenders to whom the residency restrictions could apply were used. The results therefore do not represent a reliable estimate of how many sex offenders are currently in violation of residency restrictions placed on them as a condition of their probation. An additional limitation is that the parcel database also represents some positional error. Given the source of the data (Orange County Property Appraisers Office) and the reported scale (1:2,000), the positional error of the property boundaries considered is not expected to exceed approximately 10 ft, which has very little effect on the analysis presented here. Finally, this study was limited to a single county, and the positional error in geocoding may vary in other geographic areas.

Despite the aforementioned limitations, the results of this study strongly indicate commonly used street geocoding techniques are insufficient to establish reliable residency restriction zones around schools and day care facilities. The implications of this are several fold. First, it casts doubt on the findings of previous studies that have employed street geocoding (e.g., Minnesota Department of Corrections, 2007) as a substantial portion of the sample may have been incorrectly identified as residing within a certain distance of a restricted location. Any of the online sex offender registries linked to online mapping tools should also not be interpreted as being reliable for determining residency restrictions. Second, any future study on residency restrictions (for example, determining violation rates or determining the effectiveness of residency restrictions) should not employ street geocoding but should use a more accurate technique such as parcel geocoding. Third, it suggests that statutes should be very clear in terms of how residency restrictions are defined, with clear reference to spatial data and analysis techniques that can be implemented in a GIS environment. In particular, statutes should include reference to the use of legal property boundaries in determining proximity, if that is in fact what is intended. Without such clear and unambiguous language, statutes are likely to be challenged and may prove very difficult to enforce in a practical sense. Fourth, it suggests that the enforcement of residency restrictions would greatly benefit from utilizing digital parcel data that allows for geocoding using GIS.

The findings from this study confirm those in other fields, in particular public health, that typical street geocoding is of insufficient positional accuracy to allow for spatial analysis at relatively short distances (Whitsel et al., 2006; Zandbergen, 2007; Zandbergen & Green, 2007).

Although digital parcel data is becoming much more available, current estimates suggest that only 60% of all approximately 140 million parcels in the United States are available in a format that can be utilized in a GIS environment (Stage & Von Meyer, 2003). Many smaller communities in particular have limited spatial data in digital format and limited GIS resources. This suggests that establishing accurate residency restrictions zones using digital parcel data in many jurisdictions is not yet possible. Enforcement in these cases may have to rely on ad hoc field determinations using either GPS or traditional surveying methods. Such methods are time consuming, have their own data accuracy limitations, and require specialized skills that may not be available. Although such alternatives may be possible for a limited number of cases, determining reliable residency restrictions for hundreds of offenders is only feasible in a consistent and cost-effective manner using reliable spatial data and analysis techniques.

Notes

1. Those convicted of F.S.S. 800.04 (lewd or lascivious offenses committed on or in the presence of persons less than 16 years of age), F.S.S. 827.071 (sexual performance by a child), or F.S.S. 847.0145 (selling or buying of minors).

2. If an offender/predator is under the control or supervision of the Florida Department of Corrections (FDOC), they must register with FDOC, who then provides the registrant's information to the Florida Department of Law Enforcement (FDLE). If the offender or predator is not under the control or supervision of the FDOC, then they must initially register in person with the sheriff of the county in which they are temporarily or permanently residing.

3. Although not required by Florida law, some agencies also notify residences living within a 1-mile radius of a registered sexual predator.

4. See Note 1 for a list of specific crimes.

5. Representatives working for 26 of the country's largest law enforcement agencies that analyze data spatially were polled informally about how their agency geocodes nongeographic data. The question was not specific to sex offender locations. Results indicate that two thirds geocode their data using the street geocoding method.

6. Parcel geocoding is the process of identifying the property boundaries associated with an address. Instead of a point location placed along the street segment, like in street geocoding, parcel geocoding results in a polygon for each address, representing the legal boundaries of the property associated with the address.

7. The FDLE defines administrative probation as a "form of noncontact supervision in which an offender who is determined to represent a low risk of harm to the community and, upon satisfactory completion of half the term of regular probation, is placed on nonreporting status until expiration of the term

of supervision." Community control is defined as a "form of intensive supervision, housed in the community, including surveillance on weekends and holidays, administered by officers with limited caseloads." Federal supervision includes the "care custody or control of federal criminal justice authorities." Parole includes any "post-prison supervision program" were "eligible inmates have terms and conditions of parole set by the Florida Parole Commission." Release is defined as "no longer under FDOC confinement, supervision or any other court imposed sanction." Revoked includes situations where the "Florida Department of Corrections Supervision/Community Control/Administrative Probation has been cancelled/terminated." Supervision is defined as "probation under a court-ordered term of community monitoring to include specified conditions and a specific period of time that cannot exceed the maximum sentence for the offense. Any form which requires directions and or monitoring by the Department of Corrections staff and/or the Florida Parole Commission" (see http://offender.fdle.state.fl.us/offender/LegalStatusList.jsp, viewed on 11/11/07).

8. Information related to all released offenders to whom the residency restrictions could apply were used, including those offenders who have been released from correctional supervision. Results therefore do not represent a reliable estimate of how many sex offenders are currently in violation of residency restrictions placed on them as a condition of their probation.

9. Kappa is generally considered a better measure than the proportion of agreement as it corrects the proportion of agreement due to chance.

10. Although the statute is explicit about where the measurement should be made to, the policy is less clear with respect to where the measurement should be made from. The statute would be clearer if it stated that the measurement should be taken from the nearest parcel boundary of the sex offender's residence to the nearest parcel boundary of the restricted area.

References

Bichler, G., & Balchak, S. (2007). Address matching bias: Ignorance is not bliss. *Policing: An International Journal of Police Strategies & Management, 30*, 32-60.

Bonner, M. R., Han, D., Nie, J., Rogerson, P., Vena, J. E., & Freudenheim, J. L. (2003). Positional accuracy of geocoded addresses in epidemiologic research. *Epidemiology, 14*, 408-412.

Burra, T., Jerrett, M., Burnett, R. T., & Anderson, M. (2002). Conceptual and practical issues in the detection of local disease clusters: A study of mortality in Hamilton, Ontario. *Canadian Geographer, 46*, 160-171.

Cayo, M. R., & Talbot, T. O. (2003). Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics, 3*, 5.

Chrisman, N. R. (1991). The error component in spatial data. In D. J. Maguire, M. F. Goodchild, & D. W Rhind (Eds.), *Geographical information systems* (pp. 165-174). London: Longman.

Clontz, K. A., & Mericle, J. G. (2004). Testing of apparent child sex offender clustering. Retrieved November 11, 2007, from http://gis.esri.com/library/userconf/proc04/docs/pap1261.pdf

Couclelils, H. (2003). The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS, 7*, 165-175.

Council of State Governments. (2007). *Sex offender residency restriction laws*. Lexington, KY: Author. Retrieved November 15, 2007, from http://www.csg.org/policy/pubsafety/documents/ResidencyRestrictionLaws.pdf

Dearwent, S. M., Jacobs, R. R., & Halbert, J. B. (2001). Locational uncertainty in georeferencing public health datasets. *Journal of Exposure Analysis and Environmental Epidemiology, 11*, 329-334.

Federal Geographic Data Committee. (1998). *Geospatial positional accuracy standards Part 3: National standard for spatial data accuracy.* Reston, VA: Author.

Foody, G. M., & Atkinson, P. M. (Eds.). (2002). *Uncertainty in remote sensing and GIS.* Hoboken, NJ: Jossey-Bass.

Goldberg, D. W., Wilson, J. P., & Knoblock, C. A. (2007). From text to geographic coordinates: The current state of geocoding. *URISA Journal, 19*, 33-46.

Grubesic, T. H., Mack, E., & Murry, A. T. (2007). Geographic exclusion: Spatial analysis for evaluating the implications of Megan's Law. *Social Science Computer Review, 25*, 143-162.

Harries, K. (1999). *Mapping crime: Principle and practice*. Washington, DC: U.S. Department of Justice, National Institute of Justice.

Jacquez, G. M., & Waller, L. A. (2000). The effect of uncertain locations on disease cluster statistics. In H. T. Mowrer & R. G. Congalton (Eds.), *Quantifying spatial uncertainty in natural resources: Theory and applications for GIS and remote sensing* (pp. 53-64). Chelsea, MA: Arbor Press.

Karimi, H. A., & Durcik, M. (2004). Evaluation of uncertainties associated with geocoding techniques. *Computer-Aided Civil and Infrastructure Engineering, 19*, 170-185.

Krieger, N., Waterman, P., Lemieux, K., Zierler, S., & Hogan, J. W. (2001). On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. *American Journal of Public Health, 91*, 1114-1116.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., et al. (2005). Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science, 32*, 139-160.

Maghelal, P., & Olivares, M. (2003). *Critical risk zones: Violators of Megan law*. Retrieved November 11, 2007, from http://gis.esri.com/library/userconf/proc05/papers/pap2035.pdf

Minnesota Department of Corrections. (2007). *Residential proximity and sex offense recidivism in Minnesota.* St. Paul, MN: Author.

Ratcliffe J. H. (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*, 473-485.

Reaves, B. A., & Hart, T. C. (2000). Law enforcement management and administrative statistics, 1999: Data for individual state and local agencies with 100 or more officers (NCJ 184481). Washington, DC: United States Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Rushton, G., Armstrong, M., Gittler, J., Greene, B., Pavlik, C., West, M., et al. (2006). Geocoding in cancer research: A review. *American Journal of Preventive Medicine, 30*, S16-S24.

Schootman, M., Sterling, D. A., Struthers, J., Yan, Y., LaBoube, T., Emo, B., et al. (2007). Positional accuracy and geographic bias of four methods of geocoding in epidemiologic research. *Annals of Epidemiology, 17*, 464-470.

Stage, S., & Von Meyer, N. (2003). *An assessment of parcel data in the United States*. Reston, VA: United States Geological Survey, Federal Geographic Data Committee's Subcommittee on Cadastral Data.

State of California Department of Justice. (2007). *California sex offender locator map*. Retrieved November 11, 2007, from http://www.meganslaw.ca.gov

Strickland, M. J., Siffel, C., Gardner, B. R., Berzen, A. K., & Correa, A. (2007). Quantifying geocode location error using GIS methods. *Environmental Health, 6*, 10.

United States Department of Justice. (2007). *National sex offender public Web site*. Retrieved November 11, 2007, from http://www.nsopr.gov/

United States Geological Survey. (USGS). (1999). *Map accuracy standards* (United States Geological Survey Fact Sheet 171-99). Reston, VA: Author.

Vision 20/20. (2007). *Vision 20/20 POM offender locator*. Available from www. thevision2020.com

Waller, L. A. (1996). Statistical power and design of focused clustering studies. *Statistics in Medicine, 15*, 765-782.

Ward, M. H., Nuckols, J. R., Giglierano, J., Bonner, M. R., Wolter, C., Airola, M., et al. (2005). Positional accuracy of two methods of geocoding. *Epidemiology, 16*, 542-547.

Welchans, S. (2005). Megan's law: Evaluations of sexual offender registries. *Criminal Justice Policy Review, 16*, 123-140.

Whitsel, E. A., Quibrera, P. M., Smith, R. L., Catellier, D. J., Liao, D., Henly, A. C., et al. (2006). Accuracy of commercial geocoding: Assessment and implications. *Epidemiologic Perspectives & Innovations, 3*, 8.

Whitsel, E. A., Rose, K. M., Wood, J. L., Henley, A. C., Liao, D. & Heiss, G. (2004). Accuracy and repeatability of commercial geocoding. *American Journal of Epidemiology, 160*, 1023-1029.

Zandbergen, P. A. (2007). Influence of geocoding quality on environmental exposure assessment of children living near high traffic roads. *BMC Public Health, 7*, 37.

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., & Green, J. W. (2007). Error and bias in determining exposure potential of children at school locations using proximity-based GIS techniques. *Environmental Health Perspectives, 115*, 1363-1370.

Zandbergen, P. A., & Hart, T. C. (2006). Reducing housing options for convicted sex offenders: Investigating the impact of residency restriction laws using GIS. *Justice Research and Policy, 8*, 1-24.

Zhan, F. B., Brender, J. D., De Lima, I., Suarez, L., & Langlois, P. H. (2006). Match rate and positional accuracy of two geocoding methods for epidemiologic research. *Annals of Epidemiology, 16*, 842-849.

Zhang, J., & Goodchild, M. F. (Eds.). (2002). *Uncertainty in geographical information.* New York: Taylor and Francis.

Zimmerman, D. L. (2008). Estimating the intensity of a spatial point process from locations coarsened by incomplete geocoding. *Biometrics, 64*, 262-270.

Zimmerman, D. L., Fang, X., Mazumdar, S., & Rushton, G. (2007). Modeling the probability distribution of positional errors incurred by residential address geocoding. *International Journal of Health Geographics, 6*, 1.

Paul A. Zandbergen, PhD, is an associate professor in the department of geography at the University of New Mexico. He obtained his PhD in resource management and environmental studies at the University of British Columbia in Vancouver, Canada. He considers himself a geographic information scientist with interests in both the fundamentals of GI Science as well as the applications of geospatial technologies to several fields, including water resources, spatial ecology, environmental health, and criminal justice. His current research has focused on issues of scale, error, and uncertainty in spatial analysis as well as on the robustness of spatial analytical techniques.

Timothy C. Hart, PhD, is an assistant professor in the department of criminal justice at the University of Nevada, Las Vegas (UNLV). He also serves as the director of the Center for the Analysis of Crime Statistics. His areas of interest include survey research, applied statistics, Geographic Information Systems (GIS), and victimization. Prior to joining the faculty at UNLV, he worked as a statistician for the Bureau of Justice Statistics (BJS), a program analyst for the Drug Enforcement Administration (DEA) and a research analyst for the Hillsborough County Sheriff's Office (HCSO).