Geographic disparity of positional errors and matching rate of residential addresses among geocoding solutions

Article *in* Annals of GIS · October 2015

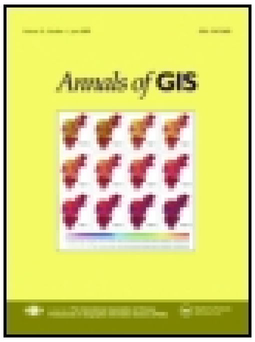
DOI: 10.1080/19475683.2015.1085437

3 authors:

Edwin Chow
Texas State University

Nathaniel Dede-Bamfo
Texas State University

Khila Raj Dahal
Boise State University



Annals of GIS

ISSN: 1947-5683 (Print) 1947-5691 (Online) Journal homepage: http://www.tandfonline.com/loi/tagi20

Geographic disparity of positional errors and matching rate of residential addresses among geocoding solutions

T. Edwin Chow, Nathaniel Dede-Bamfo & Khila R. Dahal

To cite this article: T. Edwin Chow, Nathaniel Dede-Bamfo & Khila R. Dahal (2015): Geographic disparity of positional errors and matching rate of residential addresses among geocoding solutions, Annals of GIS, DOI: 10.1080/19475683.2015.1085437

To link to this article: http://dx.doi.org/10.1080/19475683.2015.1085437

Published online: 05 Oct 2015.

http://dx.doi.org/10.1080/19475683.2015.1085437

Taylor & Francis
Taylor & Francis Group

Geographic disparity of positional errors and matching rate of residential addresses among geocoding solutions

T. Edwin Chow[a], Nathaniel Dede-Bamfo[a] and Khila R. Dahal[b]

[a]Department of Geography, Texas State University, San Marcos, TX 78666, USA; [b]Public Policy Research Center, Boise State University, Boise, ID, USA

ABSTRACT

The purpose of this study was to systematically examine the geographic disparity of geocoding error among different geocoding solutions. The research questions include: (1) What are the positional accuracies and matching rates of various geocoding techniques? (2) Are there any significant differences of geocoding quality in rural and urban areas? In this study, 1100 residential addresses scattered across Texas, USA, were address-matched using eight different geocoding platforms, including the ESRI ArcGIS Address Locator, CoreLogic PxPoint, Google Maps API, Yahoo! PlaceFinder, Microsoft Bing, Geocoder.us, Texas A&M University Geocoder, and OpenStreetMap (OSM). The geocoded locations for each method were validated against the GPS data and manual digitization. Using GPS data as reference data, the desktop geocoding using parcel data achieved the highest positional accuracy with a mean error of 24.8 m, whereas the Google Maps API was the best among the six Internet solutions with a mean error of 31.7 m. All geocoding solutions, except Geocoder.us and OSM, achieved a matching rate >95%. It is important to note, however, that the OSM geocoding revealed decent positional accuracy in terms of median and 5% trimmed mean errors, indicating the potential of crowdsourcing approach to produce an accurate geospatial data set. The geocoding errors between urban and rural areas were significantly different in most geocoding solutions but there was no consistent and monotonic trend. Excluding the errors of outliers, defined as the lowest 5 percentile, indeed changed the direction of urban versus rural accuracy in OSM, PxPoint, and ArcGIS geocoding.

ARTICLE HISTORY
Received 18 March 2015
Accepted 11 June 2015

KEYWORDS
Desktop and Internet geocoding; address-matching; urban/rural disparity; quality assessment; crowdsourcing

1. Introduction

All social and environmental processes take place in geographic locations that are often referenced by place names and street addresses. In order to capture and analyse the spatial processes, these locational attributes are processed in a Geographic Information System (GIS) through address-matching – a process converting textual location data (e.g. postal addresses) into spatial coordinates (e.g. latitude and longitude). Address-matching, which is also commonly known as geocoding, is the digital assignment of 'geographic identifiers or geocodes' to the aspatial records of locations (Rushton et al. 2006). The *x*- and *y*-coordinates for the place names or addresses are determined by matching the corresponding features in a reference data set (Goldberg, Wilson, and Knoblock 2007). The process of geocoding includes parsing the input data, querying the reference data by attributes, determining the point location by assigning the same coordinates, or interpolating the appropriate location based on the feature's geometry and topology (Longley et al. 2010; Zandbergen 2009). Common reference data sets used for geocoding include street networks, areal units such as parcels and census enumerations, and discrete address or place location (Zandbergen 2008).

Geocoding has been widely used in social, environmental, and health-related studies. It has been employed as an integral method of spatial referencing for subsequent investigation in public health (e.g. Krieger et al. 2001; Kravets and Hadden 2007; Mazumdar et al. 2008), population studies (e.g. Chow et al. 2011), criminal justice (e.g. Hay et al. 2009; Zandbergen and Hart 2009), epidemiology (e.g. Whitsel et al. 2006; Zimmerman and Li 2010), and air pollution (e.g. Gilboa et al. 2006). Generally, geocoding is carried out by a 'geocoder', which can be either part of a GIS package installed in the computer (i.e. desktop geocoding) or a tool that can be provoked by sending request as an Internet-based service (i.e. Internet geocoding).

CONTACT T. Edwin Chow chow@txstate.edu

Desktop geocoding has been the conventional practice for quite some time. It requires a trained GIS professional to preprocess the reference data set, create an address locator, customize matching parameters, perform quality control (QC), and manually correct any mismatches. Desktop geocoding, especially the manual inspection and correction, is often a laborious and time-consuming process. On the other hand, Internet geocoding involves the same process but the reference data sets are stored and maintained by a remote hosting server. The main responsibility of the users (i.e. geocoding clients) is to input the address or place name into the web interface of the geocoder, send a geocoding request to the server, and interpret the results (Chow 2008). Compared to desktop geocoding, the web alternative is [semi-] automatic and hence is relatively easy to use without much users' intervention. Therefore, Internet geocoding is often faster, less labour-intensive, requires no client-side software, and keeps the users free from the burden of acquiring and preparing the reference data set(s) (Roongpiboonsopit and Karimi 2010). Complemented by the proliferation of Web 2.0 technologies, such as mash-ups and Map Application Programming Interface (API), Internet geocoding is gaining popularity (Chow 2008; Batty et al. 2010). Various Internet geocoders, such as Google and Yahoo! Maps, have been widely used by GIS professionals and ordinary users alike.

Moreover, there has been a new paradigm in geocoding practices due to recent emergence of the citizen science. The advancement of Web 2.0 technologies provides the web infrastructure that empowers mapping citizens as voluntary sensors, producing enormous data sets known as Volunteered Geographic Information (VGI) (Goodchild 2007). As local geographic knowledge is valuable to the production of high-quality geographic databases (De Leeuw et al. 2011), the reference database of many Internet geocoding services can be updated and revised through crowd-sourcing. This bottom-up approach transforms millions of conventional consumers to producers who share the responsibilities to yield accurate and comprehensive geographic databases. In fact, OpenStreetMap (OSM) is a wiki-project that attempts to map the whole world mostly through citizen mapping. Some Internet geocoding services, such as Google, adopt a hybrid approach that integrates both licensed proprietary commercial databases and 'peer-reviewed' VGI (Dobson 2013).

However, Internet geocoding is not without its limitations. It is criticized primarily for the black box operation of its results (Swift, Goldberg, and Wilson 2008; Karimi, Sharker, and Roongpiboonsopit 2011). The quality of Internet geocoders has not been improved markedly although they are increasingly being favoured by the public and have been more easily available (Swift, Goldberg, and Wilson 2008). However, the essential components of Internet geocoding, including both the matching algorithm and reference data, remained unknown to most users, and hence the quality of geocoding results is subjected to unknown accuracy. Depending on the quality as well as the updating policy of reference data sets maintained by the geocoding service, there may be geographic disparity of geocoded results in terms of positional accuracy, temporal currency, and matching completeness. The errors in the geocoded results could introduce potential geographic bias in subsequent spatial analysis, conclusion, and decision (Zimmerman et al. 2007; Jacquemin et al. 2013). Researchers should be careful in choosing a particular Internet geocoding service over other alternatives. As different reference data and matching algorithms are used by various Internet geocoding services, accuracy and precision of the geocoded results vary (Karimi, Durcik, and Rasdorf 2004). Owing to the trade-offs in some aspects of the geocoding performance (e.g. positional accuracy, matching rate), the selection of appropriate geocoding service with optimum outcomes can become difficult. Moreover, information about the geocoding service, such as reference database(s), matching techniques, associated errors, and other metadata, are typically kept proprietary (Karimi, Sharker, and Roongpiboonsopit 2011). Therefore, it is important to understand and acknowledge the quality of various geocoding options to enhance scientific interpretation and decision-making.

The purpose of this study was to examine the geographic disparity of geocoding error among different GIS solutions. The specific research questions include: (1) What are the positional accuracies and matching rates of various geocoding techniques? (2) Are there any significant differences of geocoding quality in rural and urban areas? In this study, 1100 residential addresses scattered across Texas, USA, were address-matched using eight geocoding platforms, including the ESRI ArcGIS Address Locator, CoreLogic PxPoint, Google Maps API, Yahoo! PlaceFinder, Microsoft Bing, Geocoder.us, Texas A&M University WebGIS Geocoder, and OSM. The geocoded locations for each geocoding method were validated against the GPS point data collected from a field survey and manual digitization from in-house QC.

2. Geocoding quality

The quality of geocoding outcomes depends on the degree of urbanization, land use, the address model of the reference data set(s) (e.g. address points, parcels, or street networks), as well as the mode of geocoding platform that would affect the tolerance of input textual data, complexity of matching techniques, and the associated parameters (Zandbergen 2008). Geocoding errors can be considerable in amount and non-random in distribution (Zandbergen and Green 2007).

An important factor affecting geocoding quality is the degree of urbanization. Previous research suggested that the median error distance ranges from about 30 m to about 170 m in urban areas (Whitsel et al. 2006; Zimmerman et al. 2007; Schootman et al. 2007). Positional accuracy is lower in rural areas compared to their urban counterparts (Bonner et al. 2003; Cayo and Talbot 2003; Ward et al. 2005; Kravets and Hadden 2007). In addition to positional accuracy, rural areas are also more likely to produce higher unmatched or ambiguous records (Hay et al. 2009; Roongpiboonsopit and Karimi 2010). Wey et al. (2009) also reported a lower matching rate of geocoding in less-populated areas and Indian reservations in South Dakota. The lower positional accuracy and matching rate in the rural areas could be attributed to the use of rural routes and post office box, which would be different from the physical location of a rural address (Vieira et al. 2010). It is also possible that there is a prolonged temporal interval to maintain and update the map database in the rural areas for address-matching.

In addition, geocoding quality is also dependent on land use. Roongpiboonsopit and Karimi (2010) reported that agricultural and industrial addresses were less accurate than residential and commercial ones. Specifically, matching rates are generally higher for single-family residential addresses compared to that of multi-unit residential and commercial addresses (Zandbergen 2008). Interestingly, Roongpiboonsopit and Karimi (2010) reported no significant differences in positional accuracy across five Internet geocoders for rural addresses but for urban and suburban addresses.

Moreover, the performance (quality of geocoding results) of a specific geocoding algorithm is dependent on the address model of its associated reference data. Common address models include street network, address points, and parcels (Zandbergen 2008). The former method generally requires linear interpolation to derive the precise *x*- and *y*-coordinates in a one-to-many relationship, whereas the latter two methods use a one-to-one specific point location (e.g. the centroid of a parcel or an assigned location from an aerial photo interpretation) for geocoding. Zandbergen (2008) reported that the matching rate of address points is slightly lower but comparable to street address, followed by the parcel geocoding at a much lower matching rate.

Comparative studies examining commercial geocoders have reported varying geocoding qualities (Krieger et al. 2001; Whitsel et al. 2004; Zhan et al. 2006; Schootman et al. 2007). Therefore, it is recommended that analysts take caution when choosing the ideal algorithm, reference database, and geocoding parameters such as spelling sensitivity. Despite the proprietary nature of many commercial geocoders, conventional geocoding practice relies on desktop platforms, limited sets of reference data, and relatively well-documented geocoding quality. On the other hand, Internet geocoding as a new platform opens up a new line of research, as the reference data used for geocoding now is more open to various modes of operation (i.e. free, licensed, or VGI) and multi-source reference data. Many Internet geocoding services licensed proprietary commercial geographic databases that are now value-added versions of the free public data (e.g. TIGER files).

Swift, Goldberg, and Wilson (2008) reviewed five Internet geocoders (Google Maps API, Yahoo! API, Geocoder.us, Google Earth, and the hierarchy-based University of Southern California Geocoder) along with three common desktop alternatives (Centrus US, Geolytics, and ESRI Address Locator) in terms of cost, licensing, match rate, and concordance at the block-group level. They concluded that none of the geocoders outperformed the others in all criteria evaluated. Roongpiboonsopit and Karimi (2010) compared the matching rate, positional accuracy, and similarity of five Internet geocoding platforms, and reported superior performance from MapPoint, Google Map, and Yahoo! Map over Geocoder.us and MapQuest. They corroborated the findings of previous research that substantial variations occur among different metrics but none stands out in all specified criteria. Among six Internet geocoders, Karimi, Sharker, and Roongpiboonsopit (2011) concluded that Google Map and Bing Map-Streets performed better in terms of matching rate, whereas the results of Bing Map-Rooftops and Google Map were more accurate in terms of positional accuracy. When comparing the performance of Google Map API with a frequently updated local address-point database for a Brazilian city, Davis and De Alencar (2011) found that the accuracy of Google Map geocoding was comparatively more irregular with the address points clustered in certain parts of the streets. Goldberg (2011) examined the positional accuracy of Google Maps, Microsoft Bing, Yahoo! Maps, USC Geocoder, and a new approach, and found that all geocoders performed comparably.

The above-noted errors produce significant bias in geographic analysis and interpretation of geocoding results and subsequent use in spatial analysis (Kravets and Hadden 2007). For example, positional errors lead to inaccurate results of spatial pattern modelling (Oliver et al. 2005) and misclassification (Hay et al. 2009). The description and detection of spatial features such as clusters and trends are also adversely impacted (Zimmerman et al. 2007). Using manual geocoding that matches residential addresses to the building level as reference data, Jacquemin et al. (2013) reported that the three interpolation-based geocoding methods, including NavTEQ, Google, and Bing, estimated higher concentrations of urban air pollution but a decreasing statistical association with lung function parameters. In another instance, by overestimating the spatial accessibility in high-income suburban areas while underestimating those in urban core, geographical disparity of geocoding error could lead to social implications as well (McLafferty et al. 2012). Since the quality of geocoded results has a direct implication on spatial decisions, the errors and uncertainties associated with them should be communicated in measurable terms so that the end user can better interpret and utilize them. It is therefore recommended that studies involving geocoding evaluate and report the level of accuracy of their methods (Krieger et al. 2001; Shi 2007; Hay et al. 2009).

Drawing upon the mentioned scientific literature, the present study has made a comparative assessment of the quality of eight commonly used desktop GIS and Internet geocoding platforms. A unique feature in this study is that the comparative analysis includes a representative collection of various geocoding approaches, including desktop geocoding as well as Internet geocoding services with reference databases acquired by various modes of operations. It is also the first study that evaluates the performance of a VGI-empowered geocoding service such as OSM. The research sheds useful insights into the effectiveness of VGI towards improving geocoding quality. Moreover, this study employs a multi-scale validation so that the geocoded results of each platform are validated with (1) field-collected GPS points within a region, and (2) ortho-imagery-derived point data across a state. The advantages of field validating against GPS data are high accuracy (i.e. decimeters) and known positional errors. The evaluation metrics used include matching rate and positional accuracy.

3. Methodology

3.1. Data

Geocoding has been widely used in health and medical geography, which often highlights the geographic disparity of spatial accessibility among ethnic minority to medical facilities and their health outcomes (Krieger et al. 2001; Kravets and Hadden 2007; Mazumdar et al. 2008). This study investigated geocoding quality and examined its possible implications on ethnic minorities by using a subset of residential addresses drawn from a database of Vietnamese Americans in Texas, United States. Using stratified random sampling based on population distribution in both rural and urban areas across Texas, this study drew a representative sample of residential addresses to assess positional accuracy and matching rate among desktop and Internet GIS solutions.

The residential addresses were obtained from WhitePages (www.whitepages.com), a common people-finder site that supports many other similar websites, such as SwitchBoard (www.switchboard.com), 411 (www.411.com), and Address (www.address.com). The data were compiled into a database, which was then preprocessed and filtered to remove fictitious, duplicate, and incomplete entries using record linkage algorithms (Chow, Lin, and Chan 2011). The residential addresses contain street-level address information, with street number and name, city, state, and five-digit zip code. This process ensured that the textual quality of postal addresses to be geocoded was maintained.

3.2. Research design

3.2.1. Research questions and hypotheses

The purpose of this study was to systematically examine the geographic disparity of geocoding quality among different desktop GIS and Internet GIS solutions. In this study, a total of 1100 residential addresses scattered across Texas were address-matched using eight geocoding platforms, including the ESRI ArcGIS Address Locator, CoreLogic PxPoint, Google Maps API, Yahoo! PlaceFinder, Microsoft Bing, Geocoder.us, Texas A&M University WebGIS Geocoder, and OSM. This study also examined three geocoding data supported by the PxPoint platform, including parcels, NAVTEQ data sets, and USGS address files. Table 1 highlights the characteristics (e.g. platform, reference data, license information, etc.) of the eight geocoding solutions examined in this research.

In this study, the specific research questions include:

(1) Are there any significant differences of geocoding quality, in terms of positional error and matching rates, among various geocoding solutions?
(2) Are there any significant differences of geocoding quality between rural and urban areas?

Table 1. Characteristics of the eight geocoding solutions examined in this research.

Source	Classification*	Reference data	Data type	API limitation**
ESRI ArcGIS Address Locator	Desktop	TIGER	Public	-
CoreLogic PxPoint	Desktop	a) NAVTEQ b) national parcel database c) USPS address file	Licensed	
Google	Internet	TeleAtlas, proprietary data sets	Licensed & limited VGI	15,000/day
Yahoo!	Internet	NAVTEQ	Licensed	5000/day
Microsoft Bing	Internet	NAVTEQ, proprietary	Licensed	50,000/day
Geocoder.us	Internet	TIGER	Public	50,000/day
Texas A&M Geocoder	Internet	TIGER, parcel, ZCTA, City, County	Public	50,000/day
OpenStreetMap	Internet	TIGER, VGI	Public & VGI	1/sec

Note: *Both ESRI and CoreLogic have Internet solutions as well. Only the desktop solutions were examined in this study.
** This study only examined the free version.

The corresponding null hypotheses associated with the above-mentioned two research questions would state no significant differences in positional error (ε) and matching rate (δ) (1) across the eight geocoding solutions (i.e. $\varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_n$; $\delta_1 = \delta_2 = \ldots = \delta_n$), and (2) between urban and rural areas for a given geocoding solution *i* (i.e. $\varepsilon_{\text{urban},\ i} = \varepsilon_{\text{rural},\ i}$; $\delta_{\text{urban},\ i} = \delta_{\text{rural},\ i}$).

3.2.2. Geocoding and validation procedure

The 1100 residential addresses were address-matched using the eight desktop and Internet geocoding solutions with either the batch mode or the programming approach (through their corresponding API) to automate the geocoding process. The geographic coordinates of latitude/longitude generated from the geocoding solutions were defined in World Geodetic System 1984 (WGS 1984). These points were then transformed and re-projected into the Universal Transverse Mercator Zone 14 North (UTM Zone 14N) coordinate system based on the North American Datum 1983 (NAD 83). This state plane coordinate system provides the most accurate projection for mapping across Texas.

The geocoded locations for each geocoding method were validated against reference data collected by (1) manual digitization of valid addresses across the state of Texas and (2) a survey-grade Global Positioning System (GPS) within Central Texas. In light of the stated research questions, this study examined the geocoding quality using the best available reference data at both local and regional scales.

All 1100 residential addresses were manually digitized based on National Agricultural Imagery Program (NAIP) and Texas Orthoimagery Program (TOP) 2008–2009 Digital Ortho Quadrangle Quads (DOQQ) available from Texas Natural Resources Information System (TNRIS). The 0.75-m colour-infrared ortho-imagery has a horizontal accuracy of ±6 m at 95% confidence interval across the state (TNRIS 2009). In addition, the authors also cross-referenced with high-resolution ortho-imagery available from Google Map, Bing Map, Yahoo Map, and Zillow (an online real estate database) to ascertain the location of the residential address. After determining the location of an address, a point was placed based on the centroid of the property. Out of the 1100 sample addresses, only 940 valid residential addresses were identified. It was noted that some addresses were non-existent, business, or inaccessible locations, and hence were disregarded from further analysis.

Owing to practical constraints (e.g. time, cost), the field collection of survey-grade GPS coordinates was limited to the residential addresses within Central Texas. Specifically, this confined study area included residential addresses in selected counties within the Austin-San Antonio corridor (Figure 1). In this focus study, the subset included 78 residential addresses out of the 1100 samples noted above. For each address, the researchers visited the physical address in person and used a Trimble GeoXH GPS unit to record the average geographic coordinate over 200 readings. Whenever possible, the unit was placed either on the driveway or on the front door of the property. These locations were transformed and re-projected into UTM Zone 14N NAD83. The survey-grade GPS positioning has an average error of 0.61 m using differential correction based on the two Continuously Operating Reference Stations (CORS) nearby (Table 2). Figure 2 presents the data flow diagram that highlights the stated procedure.

The reference geocoded addresses were then overlaid on urban areas to determine whether their locations fall within or beyond urban areas defined by the Census Bureau 2010 data. The reference data sets were labelled as either 'urban' or 'rural' for comparison and statistical analyses to answer the second research question.

3.2.3. Statistical assessment

To evaluate the positional error of a location (*x, y*), the signed mean errors in both *x*- and *y*-coordinates were applied as a measure for detecting directional bias

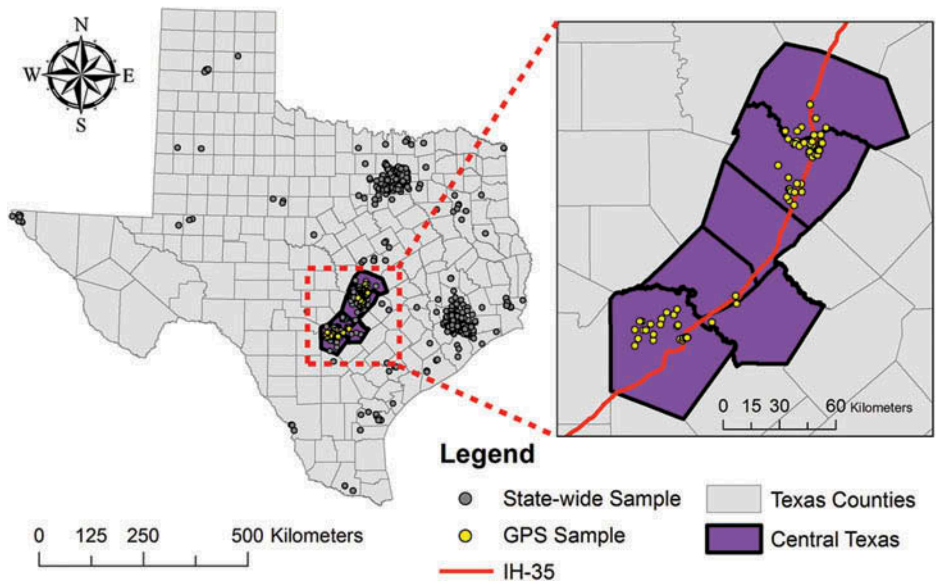


Figure 1. The residential address database of Vietnamese-Americans (VA) in Texas.

Table 2. Estimated accuracies for 12,373 corrected positions of GPS points.

Range (m)	Percentage (%)
0.00–0.15	16.0
0.15–0.30	52.1
0.30–0.50	16.7
0.50–1.00	10.5
1.00–2.00	4.5
2.00–5.00	0.2
>5.00	0.0

while the unsigned mean error quantifies the distance between observed and geocoded locations. The signed mean error (ε_x, ε_y) is simply the difference between the reference (x_{ref}, y_{ref}) and geocoded coordinates (x_{geoc}, y_{geoc}) as

$$\varepsilon_x = x_{\text{ref}} - x_{\text{geoc}}$$
$$\varepsilon_y = y_{\text{ref}} - y_{\text{geoc}}$$

The unsigned mean error (ε_{xy}) is defined as the Euclidean distance, which was calculated as

$$\varepsilon_{xy} = \sqrt{\varepsilon_x{}^2 + \varepsilon_y{}^2}$$

By using either GPS or manually digitized location as the reference data, the mean positional errors were calculated for each geocoding solution in the statewide (n = 940) and Central Texas (n = 78) experiments, respectively. Descriptive statistics, including mean, 5% trimmed mean, median, and standard deviation, of positional errors were derived to quantify the error distribution and its variability. These were compared

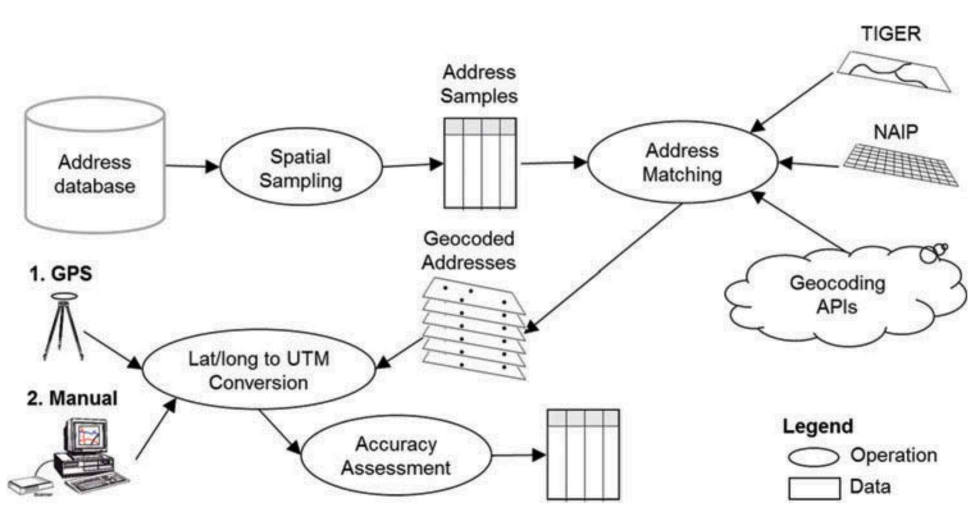


Figure 2. The data flow diagram of the geocoding and validation procedure.

among all the different geocoding engines. The matching rate is simply a ratio between successful geocoding over all attempts expressed in percentage.

The Shapiro–Wilk test revealed significant difference of positional error from normal distribution. To answer the first question, the non-parametric Kruskal–Wallis test was used to examine any significant differences among the geocoding solutions with a post hoc test to further refine any pair-wise comparison. Similarly, the Mann–Whitney U test was used to examine any significant differences of geocoding quality between urban and rural areas. To mitigate possible impact from any extreme cases (i.e. outliers) affecting the statistical analyses, the above-mentioned procedure was conducted in all points available as well as a 95 percentile subset.

4. Results

4.1. Geocoding quality across solutions

This study addressed the first research question by examining the geocoding quality across eight geocoding platforms, including the Google Maps API (Google), Yahoo! PlaceFinder (Yahoo!), Microsoft Bing (Bing), Texas A&M University WebGIS Geocoder (TAMU), Geocoder.us, OSM, CoreLogic PxPoint (PxP), and ESRI ArcGIS Address Locator (ArcGIS). The CoreLogic PxPoint geocoder uses the 'best' result by comparing three data sets, including parcels (PxP-P), NAVTEQ data sets (PxP-N), and USGS address files (PxP-U), against proprietary data source. For quality assurance and quality control (QA/QC) purposes, manual digitization (Manual) in this study was also compared against the GPS reference data set.

Validating against GPS data collected in Central Texas, the commercial geocoding using parcel data (PxP-P) achieved the highest positional accuracy with a mean error of 24.8 m, and ranked the best in terms of standard deviation and maximum error (Table 3). Other comparable geocoding solutions that achieved mean positional errors less than 50 m included PxPoint, Google, manual digitization, PxP-N, and Yahoo! in a descending order. In the intermediate range, the mean positional errors of PxP-U, Bing, and Geocoder.us were around 100 m. The TAMU and ArcGIS geocoders had mean positional errors around 10 km. Among the PxP geocoding products, it was interesting to note that the best solution (PxPoint) had a larger mean positional error compared to the parcel data set (PxP-N). This was attributed to the difference in reference data between GPS used in this study and the proprietary source adopted by CoreLogic. Notwithstanding, all geocoding solutions, except Geocoder.us and OSM, achieved a high matching rate >95%.

Furthermore, all the geocoding solutions had a lower median than mean positional error, indicating that their error distributions were positively skewed with a small proportion of outliers with high errors (Table 3). Manual digitization had the best 5% trimmed mean, median, and minimum positional error. It was noted that TAMU and ArcGIS had comparable median positional errors but their 5% trimmed mean remained large, revealing a bimodal distribution of low and high errors. The signed mean error within Central Texas revealed a moderate random directional distribution of geocoded locations (Figure 3). With the exception of Geocoder.us (GeoC), the signed mean errors of most geocoded solutions were in the north – and/or western direction from the GPS-observed location.

Using manual digitization as reference data, the commercial solutions PxP-P, PxPoint, and PxP-N achieved the closest matching with mean positional deviations of 21.5 m, 37.1 m, and 51.6 m across the state of Texas, respectively (Table 4). Around the intermediate range, Yahoo! and Bing had a mean deviation slightly below the 100-m threshold while Google, PxP-U, OSM, and Geocoder.us were above it. The mean deviation of TAMU geocoding was around 3 km whereas the ArcGIS batch geocoding was around 36 km. It is important to note that some geocoding platforms suffered from excessive positional errors in some outliers. The 5% trimmed mean revealed that Google, Bing, TAMU, and Geocoder.us reduced their mean positional errors to more than one-fifth after removing the lowest and highest 5 percentile from

Table 3. Positional error of geocoding solutions validated against GPS reference data in Central Texas (n = 78).

Error (m)	Manual	Google	Yahoo!	Bing	TAMU	Geocoder.us	OSM	PxPoint	PxP-N	PxP-P	PxP-U	ArcGIS
Mean	38.0	31.7	49.8	173.3	10,072.8	713.0	229.3	27.0	39.5	**24.8**	95.3	9764.9
5% Trimmed mean	**13.6**	24.8	44.1	23.4	7436.1	155.7	161.9	22.3	33.4	20.3	84.0	3468.4
Median	**10.8**	20.6	36.1	17.1	41.1	106.2	31.4	18.1	22.9	17.0	74.9	122.6
Min	**1.0**	5.2	2.7	5.0	5.9	13.1	4.9	5.0	2.5	5.0	6.9	7.5
Max	1384.6	370.7	391.9	10,316.8	70,021.7	21,434.1	1666.2	**194.1**	303.3	**194.1**	458.0	135,155.3
Std. dev.	159.9	46.2	52.9	1183.8	19,947.0	3369.9	489.8	30.9	46.6	**28.5**	86.8	32,460.7
Matching rate (%)	**100**	**100**	98.8	97.4	**100**	51.3	14.1	**100**	**100**	96.2	**100**	**100**

Note: Bold number indicates the best geocoding solution with minimal positional error in that category or maximum matching rate.

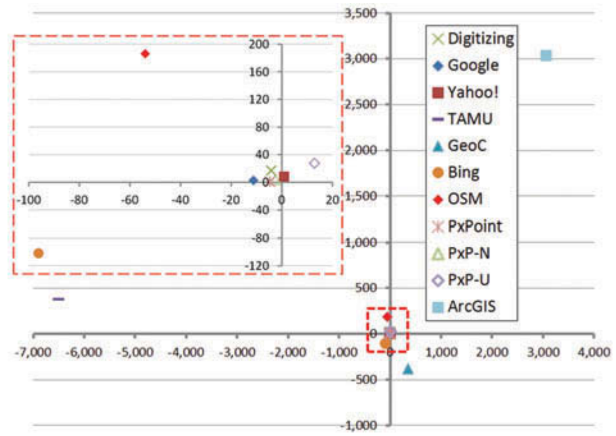


Figure 3. The signed mean error of geocoding solutions against GPS data within Central Texas.

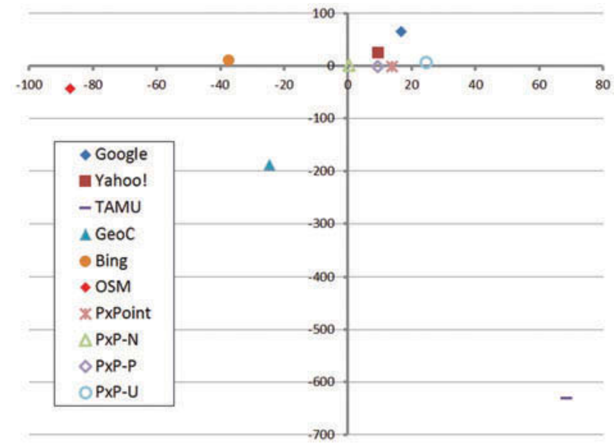


Figure 4. The signed mean error of geocoding solutions against manual digitization across Texas (excluding ArcGIS).

Table 4. Positional deviation (m) of geocoding solutions of all samples (n = 940) verified against manual digitization across Texas.

Error (m)	Google	Yahoo!	Bing	TAMU	Geocoder.us	OSM	PxPoint	PxP-N	PxP-P	PxP-U	ArcGIS
Mean	169.6	88.6	90.8	2181.5	777.0	360.0	37.1	51.6	**21.5**	176.8	36,350.9
5% Trimmed mean	16.2	51.9	12.8	197.7	114.9	136.7	12.5	39.7	**7.5**	103.3	15,946.5
Median	5.8	43.3	5.2	23.2	82.5	64.4	5.2	25.7	**4.7**	84.0	89.7
Min	0.1	1.7	**0.1**	0.2	3.6	5.6	0.1	1.7	0.1	13.3	2.5
Max	66,690.2	23,495.8	32,547.2	94,767.1	60,614.7	10,583.4	7968.9	**2159.2**	7620.0	22,106.6	920,521.2
Std. dev.	2357.2	771.5	1194.0	10,006.6	4582.9	1199.6	326.6	**100.0**	264.0	843.5	111,866.4
Matching rate (%)	**100**	99.9	97.8	**100**	84.5	26.6	99.9	99.8	89.4	99.9	**100**

Note: Bold number indicates the best geocoding solution with minimal positional error in that category or maximum matching rate.

the sample (i.e. the 'middle-90' percentile) (Table 4). Indeed, the 5% trimmed mean positional error improved in all samples of the geocoding platforms, indicating the strong influence of the few outliers. In fact, the median errors of PxP-N, PxPoint, Bing, and Google improved to around 5 m, implying 50% of all the geocoded observations were within that distance from the reference data.

Comparing the statewide samples with the Central Texas experiment, it was noted that most of the geocoding solutions, except PxP-P and TAMU, had higher positional deviation across the state (Tables 2 and 3). The PxP-N sample had the lowest root mean square error (RMSE) and circular map accuracy standard at 100 m and 151.8 m, respectively. The matching rates for most of the geocoding solutions were high and comparable between the two experiments, whereas Geocoder.us and OSM improved from 51.3% to 84.5% and from 14.1% to 26.6% in Central TX and across the whole state, respectively. Relative to the manual digitization, the signed mean error across Texas revealed a random directional distribution of geocoded locations among all geocoding solutions (Figure 4). It must be noted that the mean signed errors of ArcGIS, with ε_x and ε_y of 11,132 m and 15,175.5 m, respectively, were not plotted on the figure to maintain legibility.

4.2. Geographic disparity of geocoding quality

In this study, the research question regarding the geographic disparity of geocoding quality was investigated by exploring the statistical differences between urban and rural areas. Using the GPS data within Central Texas as the reference, results from the Mann–Whitney U test revealed significant differences of geocoding error between urban and rural areas in manual digitization, Google, Yahoo!, Bing, PxPoint, PxP-P, and PxP-U (Table 5). In particular, most of the geocoding solutions, except Bing, had less positional error (i.e. more accurate) in urban areas than in rural areas. It was noted, however, that the rural sample in Central Texas was relatively small (around five) in general. The matching rate of OSM was poor and had no geocoded addresses in the rural areas for further analysis.

Using all samples of manually digitized addresses as reference data (n = 940), the geographic disparity of geocoding quality across Texas was statistically significant among Google, TAMU, Geocoder.us, Bing, PxPoint, PxP-P, and ArcGIS (Table 6). Based on the results from the Mann–Whitney U test, however, there was no consistent and monotonic trend where urban areas always had higher or lower positional errors than rural areas across all geocoding solutions examined. Specifically, the mean positional errors in urban areas were

Table 5. Statistical differences of positional errors between all urban and rural samples (n = 78) using GPS reference data in Central Texas.

	Mann–Whitney U	Urban sample	Mean error	Std. error	Urban-rural	Rural sample	Mean error	Std. error
Manual	**0.02**	73	37.9	165.0	<	5	39.58	47.5
Google	**0.02**	73	30.4	46.6	<	5	50.1	38.7
Yahoo!	**0.05**	72	43.8	33.9	<	5	136.3	151.4
TAMU	0.19	73	9644.1	19,854.4	<	5	16,331.6	22,613.6
Geocoder.us	0.72	36	772.6	3551.8	>	4	176.2	160.1
Bing	**0.04**	71	182.2	1224.8	>	5	47.0	43.8
OSM	-	11	229.3	489.8		0	-	-
PxPoint	**0.04**	73	25.7	29.8	<	5	47.4	43.5
PxP-N	0.48	73	37.9	45.1	<	5	64.1	66.3
PxP-P	**0.02**	70	23.1	26.8	<	5	47.4	43.5
PxP-U	**0.05**	73	87.1	73.3	<	5	216.3	169.2
ArcGIS	0.15	73	8894.5	31,275.5	<	5	22,473.0	49,610.1

Note: Bold numbers are statistically significant.

Table 6. Statistical differences of positional errors between all urban and rural samples (n = 940) using manually digitized reference data across Texas.

	Mann–Whitney U	Urban sample	Mean error	Std. error	Urban-rural	Rural sample	Mean error	Std. error
Google	**<0.01**	780	190.6	2573.5	>	160	67.6	594.3
Yahoo!	0.62	779	93.2	845.3	>	160	66.1	120.5
TAMU	**<0.01**	780	1757.2	8940.4	<	160	4249.7	13,948.3
Geocoder.us	**0.01**	661	732.4	4564.0	<	133	998.6	4686.8
Bing	**<0.01**	761	61.4	573.5	<	158	232.4	2592.2
OSM	0.21	231	376.6	1245.6	>	19	158.4	189.3
PxPoint	**<0.01**	779	31.1	217.1	<	160	66.5	630.7
PxP-N	0.48	778	50.3	94.6	<	160	58.0	123.2
PxP-P	**<0.01**	779	23.9	292.0	>	160	10.3	26.7
PxP-U	0.42	779	179.0	881.2	>	160	165.9	630.4
ArcGIS	**0.01**	780	37,654.5	114,769.2	>	160	29,996.0	96,543.2

Note: Bold numbers are statistically significant.

significantly lower (i.e. more accurate) among TAMU, Geocoder.us, Bing, and PxPoint; however, Google, PxP-P, and ArcGIS revealed the opposite trend where urban areas were worse than rural areas.

The skewed distribution of positional error, however, illustrates the varying degree of influence from outliers across all geocoding platforms affecting geographic disparity. Using 95% samples of the manually digitized addresses as reference data (n = 894), the same set of geocoding platforms, including Google, TAMU, Geocoder.us, Bing, PxPoint, PxP-P, and ArcGIS, had statistically significant geographic disparity of geocoding quality across Texas (Table 7). Among them, Bing and PxPoint had a polarity change of geographic disparity between the 95% sample and the full data set (Tables 5 and 6). It was also noted that OSM, PxP-N, and PxP-U had an opposing pattern of geographic disparity between urban and rural areas although they were not statistically significant.

Figure 5(a)–(k) depicts the spatial distribution of positional errors across all platforms in Texas. For example, Google has a high density of low positional errors in major metropolitan areas, (e.g. Dallas-Fort Worth, Austin-San Antonio corridor, and Houston), but occasional high positional errors in the suburbs of those metropolitan areas and a few rural areas (e.g. Texarkana, Laredo, College

Table 7. Statistical differences of positional errors between 95% sample of urban and rural observations (n = 894) using manually digitized reference data across Texas.

	Mann–Whitney U	Urban sample	Mean error	Std. error	Urban-rural	Rural sample	Mean error	Std. error
Google	**<0.01**	745	21.6	36.8	>	149	12.6	29.3
Yahoo!	0.25	742	55.8	43.9	>	152	49.8	38.5
TAMU	**<0.01**	748	586.2	3075.7	<	146	1837.0	7069.0
Geocoder.us	**<0.01**	643	141.5	301.3	<	128	187.9	328.6
Bing	**<0.01**	741	16.5	28.4	>	153	11.1	24.8
OSM	0.07	209	77.3	71.5	<	18	123.7	117.0
PxPoint	**<0.01**	743	28.6	16.6	>	151	9.8	21.1
PxP-N	0.87	742	43.2	43.0	>	152	41.6	43.5
PxP-P	**<0.01**	666	9.0	13.2	>	151	6.7	11.5
PxP-U	0.25	740	107.6	78.5	<	154	110.8	97.8
ArcGIS	**<0.01**	739	25,585.8	83,136.8	>	155	21,926.3	76,212.8

Note: Bold numbers are statistically significant.

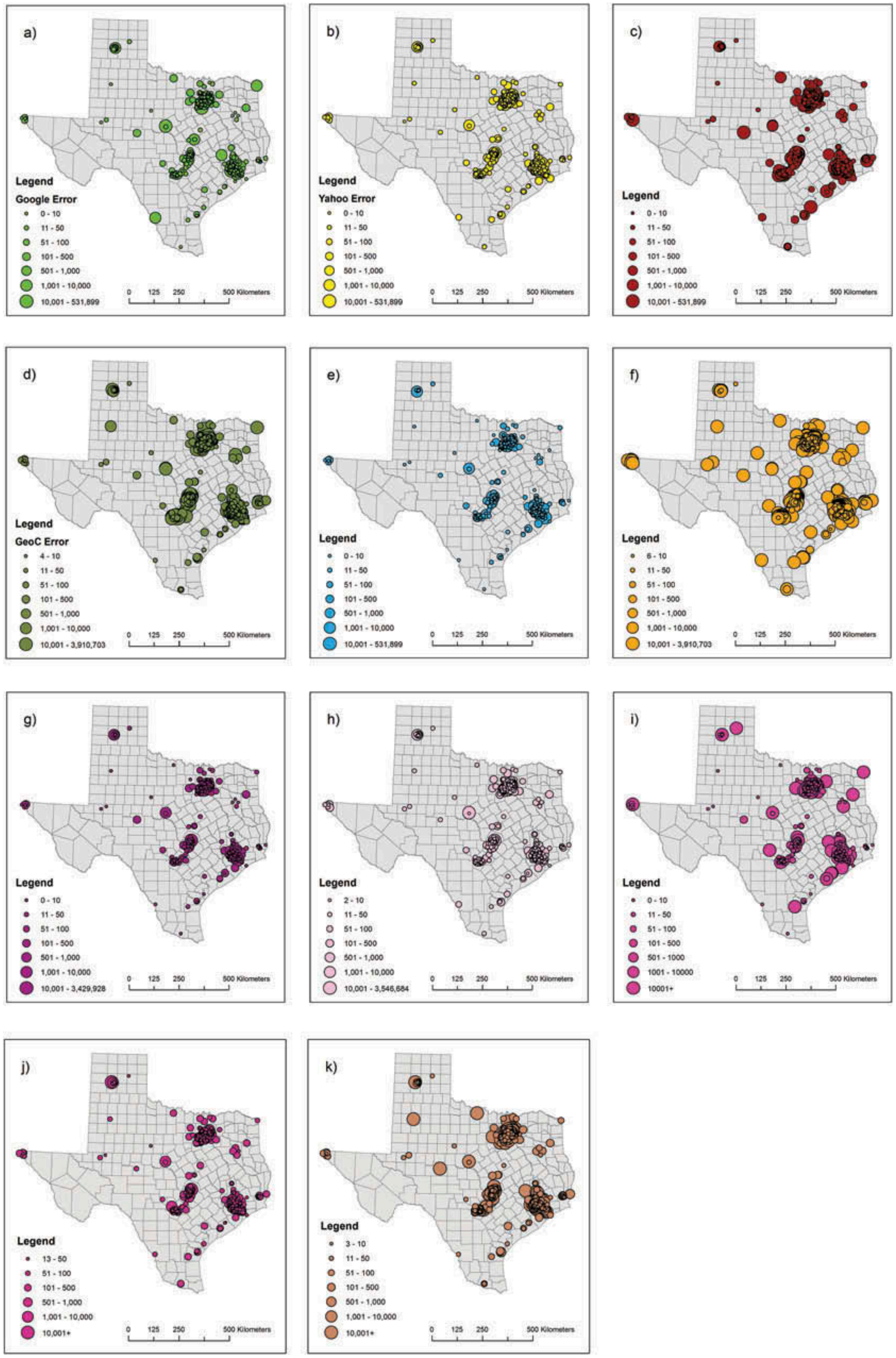


Figure 5. The spatial distribution of positional error against manual digitization across Texas by various geocoding platforms, including (a) Google, (b) Yahoo!, (c) TAMU geocoding, (d) Geocoder.us, (e) Bing, (f) OSM, (g) PxPoint, (h) PxP-N, (i) PxP-P, (j) PxP-U, and (k) ArcGIS.

Station) (Figure 5(a)). On the other hand, OSM has some accurate geocoded locations in major metropolitan areas but also poorly geocoded locations in both rural and urban areas (Figure 5(f)).

5. Discussion and conclusion

The results illustrate the importance of using high-quality reference data for geocoding purposes. Besides the manual digitization, the commercial solution, PxPoint, and its associated products, including PxP-P, PxP-N, and PxP-U, had superior positional accuracy than other solutions within Central Texas and across the state. This finding is perhaps not surprising because CoreLogic specializes in real estate industry and has devoted tremendous effort and resources into cadastral data (e.g. parcel and property), which is crucial in providing accurate base maps for geocoding. Dominant Internet geocoding services, such as Google, Yahoo!, and Bing, provided good positional accuracy similar in range as previously reported (Swift, Goldberg, and Wilson 2008) by licensing map data from other geospatial data vendors. The free geocoding services, OSM and Geocoder.us, demonstrated decent positional accuracy by leveraging on crowdsourcing or open sources. Although the issue of QC remains a challenge regarding the maintenance of large geographic data sets over a vast region, this study presents empirical evidence for the potential of VGI-driven base maps in OSM to deliver accurate geocoding (Haklay 2010). Compared to the study of Swift, Goldberg, and Wilson (2008), there were substantial differences of positional accuracy in Geocoder.us, TAMU geocoding, and ArcGIS geocoding. This finding was likely a result of the different reference data used between the two studies. For example, TAMU geocoding was originally developed in southern California (formerly USC geocoding) and its national parcel database was not available for use in this study (Goldberg, personal communication, 1 November 2013). The better performance of ArcGIS geocoding in other studies (Zhan et al. 2006; Schootman et al. 2007; Swift, Goldberg, and Wilson 2008) could be attributed to the use of enhanced StreetMap USA data instead of TIGER/Line data in this study. It is noteworthy, however, that this statewide study is different from most previous studies conducted in urbanized metropolitan areas (Zhan et al. 2006; Schootman et al. 2007; Swift, Goldberg, and Wilson 2008).

The median as well as the 5% trimmed mean was lower than the mean positional errors in all the geocoding solutions, indicating a positively skewed distribution with a small but important proportion of outliers with very high errors (Tables 2 and 3). In general, the bimodal error distribution from this study seemed to be consistent with that of Swift, Goldberg, and Wilson (2008), where many observations generally had low positional error around 10 m, but a fewer number of wrong matches with high positional error up to 1 km or more. The maximum positional errors across all platforms reported in this study were much larger than the maximum error of around several hundred metres reported in other studies (Zhan et al. 2006; Schootman et al. 2007). Nevertheless, as revealed from Tables 2 and 3, the median positional errors across Texas ranged from 5 to around 100 m and are similar to median errors in similar studies (Whitsel et al. 2006; Zimmerman et al. 2007; Schootman et al. 2007). High positional errors from a small but important number of outliers also illustrate the practical challenge of maintaining geocoding quality in a vast region like Texas.

Comparing the two experiments conducted within Central Texas and across the state, the positional error of most geocoding solutions, except PxP-P and TAMU, was higher in the experiment using manual digitization as reference data than the one using GPS (Tables 2 and 3). While there would be differences between the use of GPS and manual digitization as reference data, the signed mean error between manual digitization and GPS was negligible and did not reveal any directional bias compared to the other geocoding solutions (Figures 3 and 4). Ruling out the trivial impact due to difference in reference data, the general increase of mean positional error might be due to the geographic areas investigated in the two experiments. It was likely that the incorporation of vast rural areas across Texas in the study might have resulted in higher positional errors in general. Interestingly, this study also found that the ArcGIS geocoding tends to cluster in the east side of the reference data set (Figure 3) – an observation similar to the bearing bias reported in Schootman et al. (2007).

While most geocoders have very high matching rates better than 95%, Geocoder.us and OSM have relatively low matching rates especially within Central Texas. While it is understandable that the matching rate of OSM depends on the completeness of VGI from crowdsourcing, Geocoder.us suffers from the lack of resources to update its map databases as a free and open source public service. It is also interesting to note that there were slight increases in the matching rate across Texas than in Central Texas for all geocoding solutions except for a small drop across the CoreLogic platform (i.e. PxPoint, PxP-N, PxP-P, and PxP-U). One possible explanation is that the rapid development in Central Texas, one of the fastest growing regions in the nation, has resulted in a longer lag time in updating the base maps than other parts of Texas. For example, the metropolitan statistical area of Austin-Round Rock-San Marcos, Texas, had a population increase of 37.3% between 2000 and 2010 (U.S. Census Bureau 2010).

To maintain high matching rates among competing geocoding services, it is expected that geocoding providers will continue to license high-quality reference data from geospatial data vendors and explore ways to integrate VGI to improve their geocoding products (Dobson 2013). Despite the low matching rates of 14% within Central Texas and 27% across Texas, OSM geocoding revealed decent accuracy in terms of trimmed 5% mean error and median error, indicating the potential of the crowdsourcing approach to produce geospatial data set for geocoding with comparable positional accuracy along those measures insensitive to outliers (e.g. median and trimmed 5% mean error). As mentioned, the QC in VGI would need to be enforced more rigorously in order to mitigate the huge error induced by outliers.

Within Central Texas (Table 5), all geocoders, except Bing and Geocoder.us, had lower positional error in urban areas than in rural areas, as suggested in previous studies (Bonner et al. 2003; Cayo and Talbot 2003; Ward et al. 2005; Kravets and Hadden 2007). Across the whole state, however, six of the geocoders had higher positional errors in urban areas than in rural areas, among which Google, PxP-P, and ArcGIS revealed significant geographic disparity (Table 6). It was also noted that five out of eleven geocoding platforms, including Bing, OSM, PxPoint, PxP-N, and PxP-U, revealed a changing polarity of geographic disparity between the 95% and full samples (Table 7). More importantly, the geographic disparities of Bing and PxPoint were both statistically significant ($p < 0.001$), but revealed an opposite trend using 100% and 95% samples, respectively (Tables 5 and 6). Moreover, the changing polarity had no definite pattern, which could be either because the positional error in urban areas was larger than the rural areas or vice versa. While any attempt to discard outliers from any analysis must be performed cautiously and backed by scientific justification, this comparative analysis revealed the importance of spatial sampling in light of the positively skewed nature of error distribution. This finding suggests the need for a well-designed spatial sampling strategy to further investigate geographic disparity, if there is any evidence to accept the research hypothesis one way or the other. There were at least seven geocoding platforms that reported a significant difference between urban and rural areas from the two experiments at varying scales; this seemed to suggest the existence of geographic disparity of geocoding quality across Texas (Tables 4–6). While the Mann–Whitney U values within Central Texas (Table 5) tend to be higher than the state-wide experiment (Tables 5 and 6), the few samples in rural areas within Central Texas were inconclusive to suggest that there is less geographic disparity in the local region.

This statewide study provides a systematic evaluation of geocoding quality across a vast region and the basis to facilitate a spatially representative sampling strategy to further explore geographic disparity. The empirical results pinpointed and synthesized the spatial distribution of positional errors for each geocoding platform across Texas (Figure 5(a)–(k)). For geocoding providers, these maps identify areas with varying degrees of geocoding quality and assist the decision process to prioritize places susceptible to high positional error and allocate resources accordingly.

A limitation of this study, as well as any similar research, is that the accuracy assessment only provides a snapshot of the geocoding quality of common platforms, which are subjected to dynamic business partnership to fulfil the continual need of maintenance of reference data for geocoding. Moreover, the data licensing agreement is often proprietary and therefore this study was inconclusive about making inferences to the relationship between the quality of the reference data used by individual geocoding platforms and the positional accuracy. For example, after the completion of this project, Yahoo! began shutting down its Map Web Services and moved to the HERE platform owned by Nokia. Microsoft Bing has also licensed spatial data from Nokia as one of their data suppliers. TAMU geocoding has proprietary parcel-level data for address-matching, but it was not available for free geocoding due to licensing agreement.

As shown in this study, existing evidence calls for further research to investigate the geographic disparity of positional errors among various geocoding platforms. Since many epidemiological and health studies rely on geocoded location of patients to explore their ecological relationship and environmental pathways, future research should examine the propagation of geocoding error and its geographic disparity, if there is any and in which polarity would affect subsequent analysis and policy implications (Zinszer et al. 2010; Goldberg and Cockburn 2012). While Zhan et al. (2006) reported no significant difference in classifying geocoded locations within 1500 m from toxic release inventory, the varying degree and higher range of positional errors reported in this study calls for further investigation across emerging geocoding platforms. Depending on the biased bearing and positional error relative to the census boundary, this study echoed a previous study (Schootman et al. 2007) that reported small but important wrong allocation of census geography.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Batty, M., A. Hudson-Smith, R. Milton, and A. Crooks. 2010. “Map Mashups, Web 2.0 and the GIS Revolution.” *Annals of GIS* 16 (1): 1–13. doi:10.1080/19475681003700831.

Bonner, M. R., D. Han, J. Nie, P. Rogerson, J. E. Vena, and J. L. Freudenheim. 2003. “Positional Accuracy of Geocoded Addresses in Epidemiologic Research.” *Epidemiology* 14 (4): 408–412. doi:10.1097/01.EDE.0000073121.63254.c5.

Cayo, M. R., and T. O. Talbot. 2003. “Positional Error in Automated Geocoding of Residential Addresses.” *International Journal of Health Geographics* 2 (1): 10. doi:10.1186/1476-072X-2-10.

Chow, T. E. 2008. “The Potential of Maps APIs for Internet GIS Applications.” *Transactions in GIS* 12 (2): 179–191. doi:10.1111/j.1467-9671.2008.01094.x.

Chow, T. E., Y. Lin, and W. D. Chan. 2011. “The Development of a Web-Based Demographic Data Extraction Tool for Population Monitoring.” *Transactions in GIS* 15 (4): 479–494. doi:10.1111/tgis.2011.15.issue-4.

Chow, T. K. E., Y. Lin, N. T. Huynh, and J. M. Davis. 2012. “Using Web Demographics to Model Population Change of Vietnamese-Americans in Texas between 2000 and 2009.” *GeoJournal* 77: 119–134. doi:10.1007/s10708-010-9390-6.

Davis, J. C. A., and R. De Alencar. 2011. “Evaluation of the Quality of an Online Geocoding Resource in the Context of a Large Brazilian City.” *Transactions in GIS* 15 (6): 851–868. doi:10.1111/tgis.2011.15.issue-6.

De Leeuw, J., M. Said, L. Ortegah, S. Nagda, Y. Georgiadou, and M. DeBlois. 2011. “An Assessment of the Accuracy of Volunteered Road Map Production in Western Kenya.” *Remote Sensing* 3: 247–256. doi:10.3390/rs3020247.

Dobson, M. W. 2013. “VGI as a Compilation Tool for Navigation Map Databases.” In *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, edited by D. Sui, S. Elmwood, and M. F. Goodchild, 307–328. New York: Springer.

Gilboa, S. M., P. Mendola, A. F. Olshan, C. Harness, D. Loomis, P. H. Langlois, D. A. Savitz, and A. H. Herring. 2006. “Comparison of Residential Geocoding Methods in Population-Based Study of Air Quality and Birth Defects.” *Environmental Research* 101 (2): 256–262. doi:10.1016/j.envres.2006.01.004.

Goldberg, D. W. 2011. “Improving Geocoding Match Rates with Spatially-Varying Block Metrics.” *Transactions in GIS* 15 (6): 829–850. doi:10.1111/tgis.2011.15.issue-6.

Goldberg, D. W., and M. G. Cockburn. 2012. “The Effect of Administrative Boundaries and Geocoding Error on Cancer Rates in California.” *Spatial and Spatio-temporal Epidemiology* 3: 39–54. doi:10.1016/j.sste.2012.02.005.

Goldberg, D. W., J. P. Wilson, and C. A. Knoblock. 2007. “From Text to Geographic Coordinates: The Current State of Geocoding.” *URISA Journal* 19 (1): 33-47.

Goodchild, M. F. 2007. “Citizens as Sensors: The World of Volunteered Geography.” *GeoJournal* 69 (4): 211–221. doi:10.1007/s10708-007-9111-y.

Haklay, M. 2010. “How Good Is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets.” *Environment and Planning B: Planning and Design* 37 (4): 682–703. doi:10.1068/b35097.

Hay, G., K. Kypri, P. Whigham, and J. Langley. 2009. “Potential Biases Due to Geocoding Error in Spatial Analyses of Official Data.” *Health & Place* 15 (2): 562–567. doi:10.1016/j.healthplace.2008.09.002.

Jacquemin, B., J. Lepeule, A. Boudier, C. Arnould, M. Benmerad, C. Chappaz, J. Ferran, et al. 2013. “Impact of Geocoding Methods on Associations between Long-Term Exposure to Urban Air Pollution and Lung Function.” *Environmental Health Perspectives* 121 (9): 1054–1060. doi:10.1289/ehp.1206016.

Karimi, H. A., M. Durcik, and W. Rasdorf. 2004. “Evaluation of Uncertainties Associated with Geocoding Techniques.” *Journal of Computer-Aided Civil and Infrastructure Engineering* 19 (3): 170–185. doi:10.1111/mice.2004.19.issue-3.

Karimi, H. A., M. Sharker, and D. Roongpiboonsopit. 2011. “Geocoding Recommender: An Algorithm to Recommend Optimal Online Geocoding Services for Applications.” *Transactions in GIS* 15 (6): 869–886. doi:10.1111/tgis.2011.15.issue-6.

Kravets, N., and W. C. Hadden. 2007. “The Accuracy of Address Coding and the Effects of Coding Errors.” *Health & Place* 13 (1): 293–298. doi:10.1016/j.healthplace.2005.08.006.

Krieger, N., P. Waterman, K. Lemieux, S. Zierler, and J. Hogan. 2001. “On the Wrong Side of the Tracts? Evaluating the Accuracy of Geocoding in Public Health Research.” *American Journal of Public Health* 91 (7): 1114–1116. doi:10.2105/AJPH.91.7.1114.

Longley, P. A., M. F. Goodchild, D. J. Maguire, and D. W. Rhind. 2010. *Geographic Information Systems and Science*, 560p. 3rd ed. John Wiley and Sons.

Malone, N., K. F. Baluja, J. M. Costanzo, and C. J. Davis. 2003. The Foreign-Born Population: 2000. Census 2000 Brief. C2KBR-34. Washington, DC: U.S. Census Bureau. Accessed April 6, 2013. http://www.census.gov/prod/2003pubs/c2kbr-34.pdf

Mazumdar, S., G. Rushton, B. J. Smith, D. L. Zimmerman, and K. J. Donham. 2008. “Geocoding Accuracy and the Recovery of Relationships between Environmental Exposures and Health.” *International Journal of Health Geographics* 7 (1): 13–18. doi:10.1186/1476-072X-7-13.

McLafferty, S., V. L. Freeman, R. E. Barrett, L. Luo, and A. Shockley. 2012. “Spatial Error in Geocoding Physician Location Data from the AMA Physician Masterfile: Implications for Spatial Accessibility Analysis.” *Spatial and Spatio-temporal Epidemiology* 3: 31–38. doi:10.1016/j.sste.2012.02.004.

Oliver, M. N., K. A. Matthews, M. Siadaty, F. R. Hauck, and L. W. Pickle. 2005. “Geographic Bias Related to Geocoding in Epidemiologic Studies.” *International Journal of Health Geographics* 4 (1): 29. doi:10.1186/1476-072X-4-29.

Roongpiboonsopit, D., and H. A. Karimi. 2010. "Comparative Evaluation and Analysis of Online Geocoding Services." *International Journal of Geographical Information Science* 24 (7): 1081–1100. doi:10.1080/13658810903289478.

Rushton, G., M. P. Armstrong, J. Gittler, B. R. Greene, C. E. Pavlik, M. M. West, and D. L. Zimmerman. 2006. "Geocoding in Cancer Research: A Review." *American Journal of Preventive Medicine* 30 (2): SS16–24. doi:10.1016/j.amepre.2005.09.011.

Schootman, M., D. A. Sterling, J. Struthers, Y. Yan, T. Laboube, B. Emo, and G. Higgs. 2007. "Positional Accuracy and Geographic Bias of Four Methods of Geocoding in Epidemiologic Research." *Annals of Epidemiology* 17 (6): 464–470. doi:10.1016/j.annepidem.2006.10.015.

Shi, X. 2007. "Evaluating the Uncertainty Caused by Post Office Box Addresses in Environmental Health Studies: A Restricted Monte Carlo Approach." *International Journal of Geographical Information Science* 21 (3): 325–340. doi:10.1080/13658810600924211.

Swift, J. N., D. W. Goldberg, and J. P. Wilson. 2008. Geocoding Best Practices: Review of Eight Commonly Used Geocoding Systems. University of Southern California GIS research laboratory, Technical report No. 10. Accessed January 17, 2015. http://spatial.usc.edu/wp-content/uploads/2014/03/gislabtr10.pdf

TNRIS. 2009. TOP – 2008–2009 0.75 Meter Level-2 Metadata. Unpublished documents. Accessed June 9, 2015.

U.S. Census Bureau. 2010. American fact finder. Washington, DC. Accessed April 6, 2013. http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml

Vieira, V. M., G. J. Howard, L. G. Gallagher, and T. Fletcher. 2010. "Geocoding Rural Addresses in a Community Contaminated by PFOA: A Comparison of Methods." *Environmental Health* 9: 18. doi:10.1186/1476-069X-9-18.

Ward, M. H., J. R. Nuckols, J. Giglierano, M. R. Bonner, C. Wolter, M. Airola, W. Mix, J. S. Colt, and P. Hartge. 2005. "Positional Accuracy of Two Methods of Geocoding." *Epidemiology* 16 (4): 542–547. doi:10.1097/01.ede.0000165364.54925.f3.

Wey, C. L., J. Griesse, L. Kightlinger, and M. C. Wimberly. 2009. "Geographic Variability in Geocoding Success for West Nile Virus Cases in South Dakota." *Health & Place* 15 (4): 1108–1114. doi:10.1016/j.healthplace.2009.06.001.

Whitsel, E. A., P. M. Quibrera, R. L. Smith, D. J. Catellier, D. Liao, A. C. Henley, and G. Heiss. 2006. "Accuracy of Commercial Geocoding: Assessment and Implications." *Epidemiologic Perspectives & Innovations* 3: 8. doi:10.1186/1742-5573-3-8.

Whitsel, E. A., K. M. Rose, J. L. Wood, A. C. Henley, D. Liao, and G. Heiss. 2004. "Accuracy and Repeatability of Commercial Geocoding." *American Journal of Epidemiology* 160: 1023–1029. doi:10.1093/aje/kwh310.

Zandbergen, P. A. 2008. "A Comparison of Address Point, Parcel and Street Geocoding Techniques." *Computers, Environment and Urban Systems* 32 (3): 214–232. doi:10.1016/j.compenvurbsys.2007.11.006.

Zandbergen, P. A. 2009. "Geocoding Quality and Implications for Spatial Analysis." *Geography Compass* 3 (2): 647–680. doi:10.1111/j.1749-8198.2008.00205.x.

Zandbergen, P. A., and J. W. Green. 2007. "Error and Bias in Determining Exposure Potential of Children at School Locations Using Proximity-Based GIS Techniques." *Environmental Health Perspectives* 115 (9): 1363–1370. doi:10.1289/ehp.9668.

Zandbergen, P. A., and T. Hart. 2009. "Geocoding Accuracy Considerations in Determining Residency Restrictions for Sex Offenders' Criminal Justice." *Policy Review* 20 (1): 62–90.

Zhan, F. B., J. D. Brender, I. DE Lima, L. Suarez, and P. H. Langlois. 2006. "Match Rate and Positional Accuracy of Two Geocoding Methods for Epidemiologic Research." *Annals of Epidemiology* 16 (11): 842–849. doi:10.1016/j.annepidem.2006.08.001.

Zimmerman, D. L., X. Fang, S. Mazumdar, and G. Rushton. 2007. "Modeling the Probability Distribution of Positional Errors Incurred by Residential Address Geocoding." *International Journal of Health Geographics* 6 (1): 1. doi:10.1186/1476-072X-6-1.

Zimmerman, D. L., and J. Li. 2010. "Spatial Autocorrelation among Automated Geocoding Errors and Its Effects on Testing for Disease Clustering." *Statistics in Medicine* 29 (9): 1025–1036.

Zinszer, K., C. Jauvin, A. Verma, L. Bedard, R. Allard, K. Schwartzman, L. De Montigny, K. Charland, and D. L. Buckeridge. 2010. "Residential Address Errors in Public Health Surveillance Data: A Description and Analysis of the Impact on Geocoding." *Spatial and Spatio-temporal Epidemiology* 1: 163–168. doi:10.1016/j.sste.2010.03.002.