Impacts of Catchments Derived from Fine-Grained Mobility Data on Spatial Accessibility

Alexander Michels
CyberGIS Center for Advanced Digital and Spatial Studies, University of Illinois Urbana-Champaign, IL, USA

Jinwoo Park
CyberGIS Center for Advanced Digital and Spatial Studies, University of Illinois Urbana-Champaign, IL, USA

Bo Li
Department of Statistics, University of Illinois Urbana-Champaign, IL, USA

Jeon-Young Kang
Department of Geography, Kyung Hee University, Dongdaemun-gu, Seoul, South Korea

Shaowen Wang
CyberGIS Center for Advanced Digital and Spatial Studies, University of Illinois Urbana-Champaign, IL, USA

Abstract

Spatial accessibility is a powerful tool for understanding how access to important services and resources varies across space. While spatial accessibility methods traditionally rely on origin-destination matrices between centroids of administrative zones, recent work has examined creating polygonal catchments – areas within a travel-time threshold – from point-based fine-grained mobility data. In this paper, we investigate the difference between the convex hull and alpha shape algorithms for determining catchment areas and how this affects the results of spatial accessibility analyses. Our analysis shows that the choice of how we define a catchment produces differences in the measured accessibility which correlate with social vulnerability. These findings highlight the importance of evaluating and communicating minor methodological choices in spatial accessibility analyses.

2012 ACM Subject Classification Applied computing → Earth and atmospheric sciences; Applied computing → Health informatics; Applied computing → Transportation

Keywords and phrases Spatial accessibility, alpha shape, convex hull, cyberGIS, social vulnerability

Digital Object Identifier 10.4230/LIPIcs.GIScience.2023.52

Category Short Paper

Funding This paper is based upon work supported in part by the Institute for Geospatial Understanding through an Integrative Discovery Environment (I-GUIDE) that is funded by the National Science Foundation (NSF) under award No. 2118329. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF. The work also received support from the Taylor Geospatial Institute. Our computational experiments used Virtual ROGER that is a geospatial supercomputer supported by the CyberGIS Center for Advanced Digital and Spatial Studies and the School of Earth, Society and Environment at the University of Illinois Urbana-Champaign.

1 Introduction

Spatial accessibility is an important field of research that examines access across space to vital resources and services like healthcare [11, 12, 14]. This makes spatial accessibility a powerful tool for identifying and analyzing disparities in access across space. Access is especially crucial for socially vulnerable populations – those who are socio-economically disadvantaged, disabled, with limited transportation, etc. – who may be less likely to overcome the barriers between them and the services they need. This means that spatial accessibility work must always be cognizant of how various methodological choices impact measures of accessibility and how these different patterns of access correlate with social vulnerability.

While spatial accessibility traditionally relies on origin-destination matrices between centroids of administrative zones, recent work in spatial accessibility has created polygonal catchments from fine-grained travel data [10, 11, 14]. These works have used fine-grained point data, such as travel-time on OpenStreetMap road networks [3] and Floating Car Data (FCD) [10], to more accurately determine catchments and service areas. To calculate these catchments from point data, researchers used convex hulls in Kang et. al. [11, 12] and alpha shapes in Jiao et. al. [10]. However, convex hulls have the potential to exaggerate the catchment area as they oversimplify the shape of accessible locations.

In this paper, we examine the implications of using the convex hull and alpha shape algorithms for defining catchments in spatial accessibility analysis with a case study in Cook County, Illinois, USA. In Section 2 we discuss our methods for determining spatial accessibility and catchments and Section 3 details our data. Section 4 gives our findings for our two research questions: (1) what are the differences in the accessibility measures when we compare the two approaches and (2) how do these differences correlate with social vulnerability? Section 5 concludes with a discussion of our findings and their implications.

2 Methods

2.1 Measuring Spatial Accessibility

Spatial accessibility analyzes the distribution of supply and/or demand across space. The Enhanced Two-Step Floating Catchment Area (E2SFCA) method is a common tool for calculating spatial accessibility [13]. The first step of E2SFCA determines the weighted ratio of supply and demand (R_j) for each supply location j using Equation 1:

$$R_j = \frac{S_j}{\sum_{k \in \{d_{kj} \in D_r\}} P_k W_r} \tag{1}$$

where S_j is the degree of supply at location j, P_k is the degree of the demand or population at location k and W_r is the weight for travel-time zone r [13]. The travel-time between the supply location k and demand location j is given by d_{kj} and each step of the summation only considers supply/demand pairs k, j if the travel-time is within that step's travel-time threshold D_r ($d_{kj} \in D_r$). In the second step, each demand location sums the weighted supply-to-demand ratios of supply locations within the travel-time zones. The equation for Step 2 of the E2SFCA method is:

$$A_i = \sum_{j \in \{d_{ij} \in D_r\}} R_j W_r \tag{2}$$

where A_i is the access at demand location i and R_j, W_r are ratios and weights from step one. This yields a measure which can be interpreted as supply-to-demand ratios across space.

2.2 Calculating Catchments

An explosion in high-quality geospatial data and the development of cyberGIS for high-performance geospatial analysis [17] in recent years has led to a greater diversity in how travel-time catchments are defined. Mobility information is often given in the form of points

– nodes on road networks [12], mobile phone data [16], social media data [8], etc. – but there is some uncertainty in how we determine a service area from a set of points.

Our study examines two well-known approaches: convex hull [5] and alpha shapes [6]. The convex hull $\mathcal{CH}(S)$ of a set of points S is convex – meaning that the line between any two points in $\mathcal{CH}(S)$ is completely contained in $\mathcal{CH}(S)$ – and is the smallest convex set containing S [5]. Kang et. al. [12, 11] created driving-time polygons by calculating the ego-centric graph – the network around a node up to some distance threshold – on the road network around each supply location and used the convex hull to produce polygons. A similar approach was employed by Park & Goldberg using travel speed data in addition to the street network data [14]. The convex hull on a road network is given on the left of Figure 1.

Alpha shapes instead use the Delaunay triangulation of the points [6]. Using the triangulation, the alpha shape algorithm filters out triangles based on their circumradius using an alpha parameter [1]. We follow the convention used by the Python `alphashape` package [2], by filtering out triangles in the Delaunay triangulation which have a circumradius greater than $1/\alpha$. The convex hull and alpha shape are related in that the convex hull can be thought of as an alpha shape with $\alpha = 0$; the Delaunay triangulation with all triangles [6]. Whereas the convex hull is like a rubber band around the points, the alpha shape is like shrink wrap being fitted to the points, with α telling us how long to apply the heat. Jiao et. al. (2020) [9] calculated hospital service areas (HSAs) using alpha shapes and isolated forest algorithm on taxi trajectory data and Jiao et. al. (2022) [10] calculated accessibility using these service areas. An alpha shape on a road network is given on the right of Figure 1.

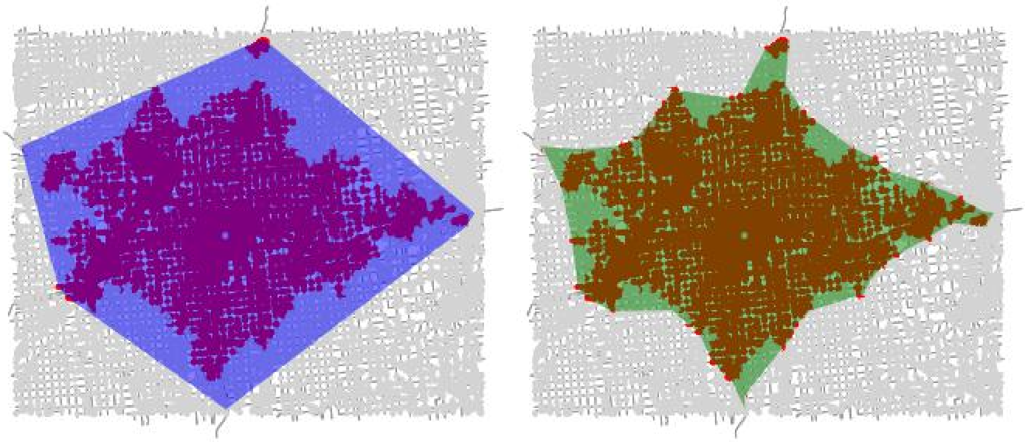


Figure 1 An example of the difference between convex hulls and alpha shapes on road network data. The street network nodes in red are within 30 minutes of the Carle Hospital in Urbana, IL while the grey nodes are not. The convex hull around the red nodes is on the left and the alpha shape (with $\alpha = 2^{-13}$ using Albers Equal Area Conic projection) is given on the right.

To determine catchments in our experiments, we calculated travel-time with the `osmnx` package [3]. First, we cleaned the road networks to remove all but the largest weakly and strongly connected components of the network, ensuring each hospital and census tract were reachable, and determined free-flow travel-times for each edge. Distance between nodes on the graph were calculated with the Python `networkx` package using Dijkstra's Algorithm. We collected the coordinates of each node within the travel-time threshold and created collections of points using the Python `geopandas` package. Using the collection of points, we were able to calculate polygonal catchments using the convex hull and alpha shape algorithms.

3 Study Area and Data

This study examined spatial accessibility for the general population to Intensive Care Unit (ICU) beds and its relationship with social vulnerability in Cook County, Illinois. This analysis required several different sets of data: (1) population and social vulnerability, (2) hospitals and ICU beds, and (3) road network data. Our population and social vulnerability data comes from the Centers for Disease Control (CDC) Social Vulnerability Index (SVI) which includes population estimates at the census tract level from the Amercian Community Survey 5-Year (2014-2018) [4]. The hospitals and ICU beds per hospital were obtained from the Homeland Infrastructure Foundation-Level Data Geoplatform[1]. Our road network dataset was obtained from OpenStreetMap using the Python `osmnx` package [3].

4 Results

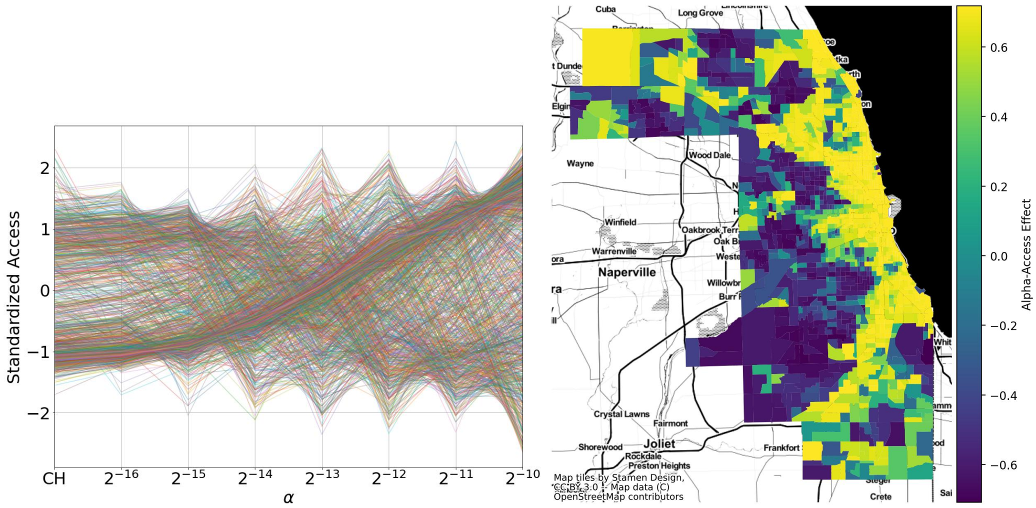
Our experiments answer two key research questions. First, what is the relationship between the alpha parameter and measured access across space? Second, does the relationship between alpha and access correlate with the CDC SVI? To accomplish this, we calculated alpha shapes using a range of alphas $(2^{-16}, 2^{-15}, \cdots, 2^{-6})$ and convex hulls around the 10, 20, and 30 minute travel zones for each hospital. Then, we calculated spatial accessibility using the E2SFCA method with the catchments produced.

To understand the relationship between alpha and access, we plotted for each census tracts standardized accessibility score based on each tested α value as shown in Figure 2a. For each census tract, we compiled the distribution of spatial accessibility measures for each value of alpha and standardized the data such that the mean was zero and standard deviation is one. It is hard to determine a clear relationship here: access in some census tracts increase while others decrease as alpha rises. However, it is clear from Figure 2a that our choice of convex hull versus alpha shape makes a significant impact on the measured spatial accessibility.

To quantify the relationship between alpha and access, we computed an "alpha-access effect" metric for each census tract, mapped in Figure 2b. The metric is the linear regression coefficient between log base two of the alpha values and standardized mean accessibility for each census tract. A positive alpha-access value indicates a positive relationship between the alpha value and the measured accessibility, whereas a negative value means the measured accessibility tends to decrease as the alpha value increases. The clustering in the map prompted us to check for spatial autocorrelation and we found a Moran's I of 0.465 and p-value of zero given by `pysal` using a Gaussian weight matrix. The map in Figure 2b shows positive values in downtown Chicago, against Lake Michigan, and generally declining values as we move west with some exceptions like O'Hare International Airport (north-western corner), the I-90 corridor (the yellow strip running from downtown Chicago to O'Hare), and the I-57 corridor (the yellow strip running south-west from the lake).

Lastly, we found a weak negative correlation (Kendall's τ: -5.45e-02, p-value: 3.06e-03) between the alpha-access effect and SVI. This is a statistically significant result at the 0.01 significance level and indicates that census tracts with high social vulnerability tend to also be the ones where measured spatial accessibility decreases as a function of alpha. Practically, this means convex hulls and low α alpha shapes tend to over-report access in socially vulnerable communities relative to higher α alpha shapes.

[1] https://hifld-geoplatform.opendata.arcgis.com/datasets/geoplatform::hospitals-1



(a) Standardized access as a function of α for census tracts in Cook County, IL

(b) Map of the "alpha-access effect" for census tracts in Cook County, IL

Figure 2 (Left) Plots of standardized accessibility by census tracts as a function of alpha for Cook County, IL. CH stands for Convex Hull and is an alpha shape with alpha equal to zero. (Right) A map of Cook County, IL giving the alpha-access effect of each census tract.

5 Concluding Discussion

In this paper, we explored how different catchment constructions (i.e., convex hull and alpha shape) affect spatial accessibility metrics that employ granular travel-time data. Our work shows that the differences are spatially autocorrelated and vary greatly depending on the alpha value used. In addition, we demonstrated that these differences in spatial accessibility – arising from the choice between convex hulls and alpha shapes – correlate with social vulnerability. This suggests that using convex hulls and low α alpha shapes for spatial accessibility may overestimate access for socially vulnerable populations which could have unintended policy implications.

While we cannot claim that either convex hulls or alpha shapes provide a ground truth for mobility, alpha shapes with appropriate values of alpha more accurately represent the data we have, as seen in Figure 1. Therefore, we can conclude that using convex hulls and inappropriately low α alpha shapes for constructing catchments tend to over-report access to ICU beds for those who are socially vulnerable in Cook County, IL. This may lead to policy-makers providing less support to socially vulnerable populations.

There is future work to do in this vein of research as more diverse spatial datasets and tools become available. It would be illuminating to apply this methodology to a variety of cities to see how much of our findings hold in cities generally. Additionally, our work used OpenStreetMap data, but there is a variety of mobility and transportation data which have potential for use in spatial accessibility studies including the Floating Car Data [10] and temporally dynamic mobility data [15]. Further work could also help to identify the circumstances in which convex hulls and alpha shapes more accurately describe real-world mobility which varies heavily based on individual-level characteristics [7].

References

1 Nataraj Akkiraju, Herbert Edelsbrunner, Michael Facello, Ping Fu, EP Mucke, and Carlos Varela. Alpha shapes: definition and software. In *Proceedings of the 1st international computational geometry software workshop*, volume 63, 1995.

2 Ken Bellock, Neil Godber, and Philip Kahn. bellockk/alphashape: v1.3.1 Release, April 2021.

3 Geoff Boeing. OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139, 2017.

4 Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research, Analysis, and Services Program. CDC/ATSDR Social Vulnerability Index 2018 Database US, 2018.

5 Mark de Berg, Otfried Cheong, Marc van Kreveld, and Mark Overmars, editors. *Computational Geometry: Algorithms and Applications.* Springer, Berlin, Heidelberg, 2008.

6 H. Edelsbrunner, D. Kirkpatrick, and R. Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29(4):551–559, July 1983.

7 Amin Gharebaghi and Mir Abolfazl Mostafavi. Space-Time Representation of Accessible Areas for Wheelchair Users in Urban Areas (Short Paper). In Stephan Winter, Amy Griffin, and Monika Sester, editors, *10th International Conference on Geographic Information Science (GIScience 2018)*, volume 114 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 28:1–28:6, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

8 Yingjie Hu and Jimin Wang. How Do People Describe Locations During a Natural Disaster: An Analysis of Tweets from Hurricane Harvey. In *11th International Conference on Geographic Information Science (GIScience 2021) - Part I*, volume 177 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 6:1–6:16, Dagstuhl, Germany, 2020.

9 W. Jiao, H. Fan, and Y. Wang. Analyzing the Spatiotemporal Patterns of Emergency Medical Travels from FCD Data. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume V-3-2020, pages 751–756. Copernicus GmbH, August 2020. `doi:10.5194/isprs-annals-V-3-2020-751-2020`.

10 Wei Jiao, Wei Huang, and Hongchao Fan. Evaluating spatial accessibility to healthcare services from the lens of emergency hospital visits based on floating car data. *International Journal of Digital Earth*, 15(1):108–133, December 2022. `doi:10.1080/17538947.2021.2014578`.

11 Jeon-Young Kang, Bita Fayaz Farkhad, Man-pui Sally Chan, Alexander Michels, Dolores Albarracin, and Shaowen Wang. Spatial accessibility to HIV testing, treatment, and prevention services in Illinois and Chicago, USA. *PLOS ONE*, 17(7):e0270404, July 2022.

12 Jeon-Young Kang, Alexander C Michels, Fangzheng Lyu, Shaohua Wang, Nelson Agbodo, Vincent L Freeman, and Shaowen Wang. Rapidly Measuring Spatial Accessibility of COVID-19 Healthcare Resources: A Case Study of Illinois, USA. *International Journal of Health Geographics*, 2020. `doi:10.1186/s12942-020-00229-x`.

13 Wei Luo and Yi Qi. An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place*, 15(4):1100–1107, December 2009. `doi:10.1016/j.healthplace.2009.06.002`.

14 Jinwoo Park and Daniel W. Goldberg. An Examination of the Stochastic Distribution of Spatial Accessibility to Intensive Care Unit Beds during the COVID-19 Pandemic: A Case Study of the Greater Houston Area of Texas. *Geographical Analysis*, July 2022. `doi:10.1111/gean.12340`.

15 Jinwoo Park, Alexander Michels, Fangzheng Lyu, Su Yeon Han, and Shaowen Wang. Daily changes in spatial accessibility to ICU beds and their relationship with the case-fatality ratio of COVID-19 in the state of Texas, USA. *Applied Geography*, page 102929, March 2023.

16 Michael Sinclair, Qunshan Zhao, Nick Bailey, Saeed Maadi, and Jinhyun Hong. Understanding the use of greenspace before and during the COVID-19 pandemic by using mobile phone app data. In *GIScience 2021*, September 2021. `doi:10.25436/E2D59P`.

17 Shaowen Wang. A CyberGIS Framework for the Synthesis of Cyberinfrastructure, GIS, and Spatial Analysis. *Annals of the Association of American Geographers*, 100(3):535–557, June 2010. `doi:10.1080/00045601003791243`.