Using Extended Two-Step Floating Catchment Area to Map Children’s Level of Access to Physical Books in West Virginia

Heather A. Maxey
hamaxey@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Geographic Information Sciences Commons, Human Geography Commons, Library and Information Science Commons, and the Spatial Science Commons

Recommended Citation

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Using Extended Two-Step Floating Catchment Area to Map Children’s Level of Access to Physical Books in West Virginia

Heather Maxey

Thesis submitted to the Eberly College of Arts and Sciences at West Virginia University in partial fulfillment of the requirements for the degree of

Master of Arts in Geography

Dr. María Pérez, PhD. (Chair)
Dr. Jamison Conley, PhD.
Dr. Kathleen Ryan, PhD.

Department of Geology and Geography
Morgantown, West Virginia
2023

Keywords: Book Access; Book Deserts; Library Access; West Virginia; E2SFCA; Mapping Access

© 2023 Heather Maxey
Abstract

Using Extended Two-Step Floating Catchment Area to Map Children’s Level of Access to Physical Books in West Virginia

Heather Maxey

The state of West Virginia lacked a comprehensive, in-depth analysis of the availability of physical books across the state. Access to reading material is vital to a child’s educational development as well as lifelong learning beyond childhood. To guide policy decisions and planning for outreach which addresses reading levels and beyond, I mapped levels of access stemming from socioeconomic barriers, spatial barriers, and an integrated index of both types of barriers. To map spatial access, I utilized a modified Extended Two-Step Floating Catchment Area model, deriving techniques from Hong et al’s 2023 article addressing access for health care in WV. The final product was an integration of the spatial and socioeconomic indices to model overall levels of access in the state. The integrated index demonstrates the areas with the lowest access levels in West Virginia are the most rural, with a particular emphasis on the Southeastern part of the state. Clay County, WV is noted as the low access zone with the highest estimated child population.
Table of Contents

I. Introduction

II. Background
   2.1 Assessing Spatial Accessibility
   2.2 Assessing Socioeconomic Accessibility
   2.3 Prior Works Examining Access to Public Libraries and Books
   2.4 Mapping Public Library Access: Major Works and Implications

III. Methods
   3.1 Study Area and Data
   3.2 Socioeconomic Access
   3.3 Spatial Access
   3.4 Integrated Access Index

IV. Results

V. Discussion and Concluding Remarks

VI. References

Acknowledgements:
The WVLC was an invaluable source of data and general information, without their resources this thesis would have been impossible. Thank you to Dr. Pérez, Dr. Conley, and Dr. Ryan for their kindness and assistance throughout the entire process. A special thank you to Koby Tazwell and Erik Nevicky for sharing their knowledge and time. Last and certainly not least, thank you to my family and friends for their support.
I. Introduction

Educators and activists have raised a number of concerns regarding the impacts of lack of funding, curriculum cuts in favor of teaching for standardized testing, and lack of access to technology on student learning (Helfenbein, 2021). It is widely accepted in educational research that access to reading material is an important factor in childhood education and development, as demonstrated by over half of poorly performing students lacking easily accessible reading material (Helfenbein, 2021). In the United States, over 60% of kindergarteners with poor academic performance have zero books available in their homes (Neuman, 1999). Households below the poverty level are less likely to own books, due to issues of space, ability to travel, affordability, and educational attainment of the adults in the home, among other factors. This implies a state struggling with poverty is also a state struggling with access to books. West Virginia has one of the five highest poverty rates in the country as of 2019 U.S. Census data. Despite the known critical importance of access to physical books, however, conversations on issues of digital access often overshadow those of access to physical books.

Despite multiple studies reporting a higher level of reading comprehension, retention of information, and general enjoyment compared to digital formats, scholars have paid less attention to the manifestations and impacts of a lack of physical book access on children (Kretzschmar et al., 2013; Mangen et al., 2013). Current literature shows how, especially during long breaks from school, children’s literacy rates decline when they experience this lack of physical books—a trend that disproportionately affects minorities and those in poverty (Alexander et al., 2016; Allington et al., 2010). Specifically, there is a critical need to document and assess accessibility to physical books prior to moving towards potential solutions, such as where to provide outreach and services. To promote and assure that quality access exists for all we first need to identify where it does not.
This project combines conceptual and methodological geographic information science (GIS) tools to examine and map children’s level of access to physical books in West Virginia.

Scholarship examining physical book access for children in the U.S. is mostly outside of geographic scholarship even though this phenomenon has uniquely spatial aspects. Of this scholarship, all were either published in education research journals or do not focus on geographic elements, like travel time to access a particular good or service. Neuman (1999) hints at the spatiality of access to literacy by discussing how the presence of books alters our classrooms while championing the importance of literature in a child’s life and education. Neuman and Moland (2016) discuss the concept of “book deserts”, an idea they compare to food deserts, and the consequences of income segregation on a child’s access to print. In the field of Geographic scholarship\textsuperscript{1}, the concept of book deserts and access to books remains unexplored. In the context of West Virginia, this is even more so.

**This thesis aims to examine the levels of access to literature for children in West Virginia through a quantitative multivariate analysis of socioeconomic and spatial accessibility factors.** Here, I define children as minors under the age of eighteen and above the age of three. Also, in the context of this study, I define literature as physical books in libraries open to children. I define the concept of access as the ease of ability for a child to physically travel and obtain this literature (assuming the presence of a caretaker to facilitate such travel). In line with the scope of a master’s thesis, my analysis focuses on accessibility to public libraries as a proxy for access to physical books. Using Extended Two-Step Floating Catchment Area (E2SFCA), a model which assesses accessibility by producing a supply-to-population ratio, I measure spatial

\textsuperscript{1} The literature review was limited to English-language journals, predominantly authored by scholars in US and European institutions. However, see Baridon (2016), Krashen et al. (2021), Guo et al. (2017), Mushtaq and Arshad (2022), and Witten et al. (2011) for examples outside of English (Baridon) and US and European institutions.
access to physical books, as well as socioeconomic indicators, with a categorical mapping index. I pilot the usability of this analysis in collaboration with WVU’s Energy Express summer reading and food outreach program with the goal to provide insight to further improve their program.

**Research Questions:**

1. What levels of access do children have to physical literature in West Virginia?
   a. What percentage of the state is considered low access?
   b. What percentage of minors in West Virginia live in low access areas?

2. Which of the chosen socioeconomic factors have the highest impact on a child’s level of accessibility?
   a. Do spatial factors or socioeconomic factors have a higher level of access?
   b. Which individual factors have the highest levels?

**II. Background**

This section begins with a discussion on the independent variables I have selected and how they are measured. I start with spatial accessibility, then move to socioeconomic accessibility, and why I have chosen the five variables I did for the purpose of this thesis. Afterwards, I examine how other studies have approached comparable topics concerning these variables, comparing, and contrasting methodological approaches and their results. To conclude, I summarize major themes and important research, as well as general implications of the works referenced in this section.

2.1 Assessing Spatial Accessibility

Concerning spatial accessibility, two commonly used methods are container approaches and travel impedance approaches. The container approach measures the number of facilities or
opportunities in a geographical unit and has been used to measure the accessibility of public facilities and opportunities (Cheng et al., 2021; Guo et al., 2017; Hong et al., 2023; Sin, 2011; Witten et al., 2011). The ‘container’ is the geographic unit which can be defined by a multitude of parameters from political boundaries to physical boundaries. Examples of this include Sin’s (2011) approach of identifying locations of public libraries within Census tracts, or the statistical subdivisions of a county, using the Census tract as their container. A major limitation of this approach is the averaging of the accessibility for all households within the container, which is less suitable for rural areas because the container scales depending on the population density, making rural Census tracts larger. Secondly, in relation to public libraries, limiting the analysis to sites only within a certain container is less applicable as services are often offered based on state residency. A Census tract, for example, could exclude many potential sites due to the scaled container size mentioned above. The container itself could be large but with over 50% its shape covering uninhabited areas. If one were to assign a centroid to this large container, that centroid could land in the middle of a forest or fields nowhere near any population.

The other main approach, travel impedance, focuses on the effort required to travel to a site or opportunity (Cheng et al., 2021). Travel time, travel distance, and travel cost are the three commonly measured variables for this approach. The traditional method is a straight distance or Euclidean distance measurement, which does not consider road networks or topography. Road network-based analysis is a more advanced technique which requires road network data and is more commonly used in recent methodology with the availability of data on the rise in recent years (Efiong, 2019; Guo et al., 2017; Hong et al., 2023; Huhndorf & Dzialek, 2017; Rosichan, 2020). Straight-line distance does not accurately account for travel distance, especially in rural,
mountainous regions like West Virginia, although Apparicio et al. (2008) found a strong correlation between straight-line distance and road network distance in metropolitan areas.

The main limitation of road network distance is that it usually utilizes an individual’s location, such as a home address, which is not feasible on a large scale due to the difficulty of obtaining said data and computing the data. Park (2012) utilized this method, taking registered members’ addresses from libraries in a county in Florida and measuring the distance between their address sites and the library sites. Other factors that road network analysis accounts for are temporal factors and travel mode. Delafontaine et al. (2011) and Allen (2019) both examined these aspects and found them significant to accessibility. Traditionally used in assessing the accessibility of health care (facility size or number of physicians present), size of the facility and its supply (in the case of libraries, book collection size) are also found to be significant factors (Hong et al., 2023; Huhndorf & Działek, 2017; Luo & Qi, 2009; Rosichan, 2020). The distance to the facility can be considered worthwhile if the service provided is of higher quality or availability. Assuming that the further a facility is, the less likely it is to be accessed, accessibility can be represented as the number of opportunities divided by the travel impedance (Monzón et al., 2013; Witten et al., 2011), or the facility’s size or attractiveness divided by the travel impedance (Cheng et al., 2021; Huhndorf & Działek, 2017; Luo & Qi, 2009). Distance friction reflecting distance decay usually is associated with travel impedance (Cheng et al., 2021; Efiong, 2019; Hong et al., 2023; Huhndorf & Działek, 2017; Luo & Qi, 2009; Zakrajšek & Vodeb, 2020).

Overall, the literature illustrates an increase in accuracy of assessing spatial accessibility, especially at the local, urban scale. However, there are still challenges which limit our ability to accurately measure accessibility of rural, large-scale areas with complex topography such as West Virginia. Hong et al. (2023) seeks to improve on these methods specifically in the context of WV
and form the basis of what I seek to build upon. Without extremely powerful processing capabilities, complex, small-scale data, and extended periods of time—these improvements are difficult to manifest.

2.2 Assessing Socioeconomic Accessibility

Aside from physical barriers such as distance, there are a multitude of other factors that affect accessibility. Even if a facility is within walking distance, the time to go there, the ability to afford the travel, the knowledge that the option even exists, and that it is a resource worth visiting and utilizing, are all barriers of entry. As many factors affect accessibility, for feasibility of analysis, researchers must carefully choose the variables that are most relevant. As there is no major consensus in the literature for which factors are necessary, this is a challenging task. A second limitation is how few studies have been performed which address more than the spatial barriers affecting library accessibility. The two factors that remain consistent are poverty rate and educational attainment of the householder (Cheng et al., 2021; Guo et al., 2017; Hertel & Sprague, 2007; Rosichan, 2020; Zakrajšek & Vodeb, 2020). Based on literature which examines health care accessibility, two more factors are often examined: householders without a partner or spouse and access to a vehicle (Asanin & Wilson, 2008; Daly et al., 2019; Hong et al., 2023; Paez et al., 2010; Shah et al., 2015; Wang & Luo, 2005). Few studies have examined the integration of spatial accessibility with aspatial accessibility and there is no consensus on methods to best do so (Bissonnette et al., 2012; Hong et al., 2023; Wang & Luo, 2005).

\[2\] In the case of children, there is the key factor of their dependence on caretakers and caretaker networks such as school systems as well. I acknowledge but do not pursue this in this study.
2.3 Prior Work Examining Public Libraries and Access to Books

Libraries themselves are a unique space with a multitude of uses and purposes. Sociologist Ray Oldenburg’s concept of ‘third place’ describes this uniqueness and the importance of the public library as a place that is separated from home (first place) and work (second place); however, a library can blur the lines between all three (Aabø & Audunson, 2012; Kuno, 2023; Mushtaq & Arshad, 2022). A patron’s role can adapt and shift across many spheres, from family and neighbor to colleague and citizen (Aabø & Audunson, 2012). More recent literature has a focus on the digital and technological age and how it affects the place and space of a library and the experience of their patrons (Baridon, 2016; Bradley, 2016; DeMaagd et al., 2013; Mann, 2001; Mushtaq & Arshad, 2022). All reviewed studies concerning the use of public libraries highlight the fluidity of their use and their importance as community building, lifelong learning centers (Aabø & Audunson, 2012; Audunson et al., 2019; Baridon, 2016; DeMaagd et al., 2013; Krashen et al., 2021; Kuno, 2023; Mushtaq & Arshad, 2022; Schloffel-Armstrong et al., 2021).

Most of the previous literature examines access to and use of libraries (or health care sites) and not the specific issue of access to physical books. Studies that do examine access to physical books generally do so from an angle of the importance of the ability to access and do not address the level of accessibility (Neuman, 1999; Neuman & Celano, 2001, 2012; Neuman & Moland, 2016). An important example of this scholarship is the work of Susan Neuman, who has performed multiple, seminal studies on the importance of access to physical books. Her 1999 paper was critical to understanding the importance of this thesis, as she examines the impact of a prominent level of access to books on economically disadvantaged children in the United States. Other critical research from 2016, Neuman and Moland introduced the concept of ‘book deserts’ as areas lacking books, but they have yet to be mapped. The studies in the previous sections are concerned with
access to public libraries, which due to the limitations of a thesis is how I am choosing to approach the question of access to physical books. Overall, the literature on the concept of book deserts and a lack of access to literature needs to be expanded upon. We know there are gaps but finding the gaps and addressing them is the next step.

2.4 Mapping Public Library Access: Major Works and Implications

The most common approach to mapping public library (or similar facility) access is through the Two-Step Floating Catchment Area method (2SFCA) or an extended variation of it (Allen, 2019; Bissonnette et al., 2012; Guo et al., 2017; Huhndorf & Dzialek, 2017; Luo & Qi, 2009). Other common methods include basic network analysis and population demand, which is also a vital part of the E2SFCA model (Cheng et al., 2021; Efiong, 2019; Hertel & Sprague, 2007; Park, 2012; Witten et al., 2011). Hong et al. (2023) utilize a modified E2SFCA model alongside multiple GIS methods to map both spatial and aspatial access to public health care in the state of West Virginia. Their methodology addresses many of the concerns other studies have when attempting to apply them to a statewide scale for a state such as West Virginia. I have developed a variation of their methodology suitable for my area and theme of interest.

Aside from Hong et al. (2023), the studies that most closely inform this research are Neuman and Moland (2019), Guo et al. (2017), and Cheng et al. (2021). Each outline important considerations and methodologies for examining and mapping physical book access despite not dealing specifically with this topic. Major themes throughout the literature discuss facility supplies (book count), facility site numbers, population demand, travel distance, and distance decay. As there have not been studies specifically mapping children’s accessibility to physical books despite the multiple studies highlighting the importance of physical books, the need for research is present.
Aside from poverty rates and reading levels, West Virginia needs information on spatially informed accessibility levels before it can truly address the problem. The rural status of the state is a secondary factor of importance as most studies examine urban mapping methodology. This study aims to address this gap by mapping these levels of access.

**III. Methods**

This section first examines the study area and data used, then describes the three main expected indices that I calculated and their analytical power. First, I discuss the socioeconomic accessibility index and the five variables chosen for this analysis. Next, I examine the spatial accessibility index and the two steps of the model used to create it. Finally, I discuss the integrated accessibility index.

3.1 Study Area and Data

Data on public libraries open to children, their service capacity in book collection size, and socioeconomic variables were obtained for the state of West Virginia. Figure 1 shows the updated 2022 WV library site data and 2022 Census block groups. Block groups are the statistical subdivision of Census blocks, which are the subdivisions of the Census tracts as discussed previously. They are the smallest unit available for the data I utilized. The combination of prevalent poverty rate, lack of publicly available and mapped data, and lower than average reading levels in the state are all factors highlighting the significance of WV as a case study. There is an understanding that people who live on the border of the state can travel over state lines to use resources, but this factor was not accounted for due to data availability and project feasibility of the research.
West Virginia addresses literature outreach to children in a variety of ways, mostly at schools both in classes and through extracurricular outreach programs—such as Energy Express and Read Aloud WV—and through public libraries. For the scope of this thesis, I focus on public library locations exclusively. Spatial and socioeconomic factors affecting access are also exacerbated in WV by poor state infrastructure and a lack of public transportation. Children across the state cannot drive until at least 16 years of age and most will rely on being provided transportation to reach places which provide books. Rural areas in particular experience a higher rate of this deficit in infrastructure.

---

3 Little libraries, bookmobiles, and outreach programs are other methods of book distribution which are often locally provided and difficult to map, but their existence is acknowledged as important to the provision and distribution of physical literature.
Figure 1: Study area with updated 2022 public library sites and 2020 Census block groups.
There are three main pieces of data I utilized, outlined in Table 1. I derived the library site points from data courtesy of the West Virginia Library Commission (WVLC) and the West Virginia GIS Technical Center (WVGISTC). The WVGISTC currently has a publicly accessible geocoded shapefile of library locations provided by the WVLC in 2002. As the current data set is from 2002, I removed and added sites depending on their availability to children and if they are currently open as of 2022 for a total of 169 libraries. The main limitation was a physical address as sites sometimes used a P.O. box which required me to use a combination of the WVLC’s library directory and Google Maps to ascertain. Using this physical address, Google Maps and Streetview, and imagery of the state, I checked the placement of each point was accurate and moved any that were not. The other piece of information I needed for each site was a book count, which I obtained via the WVLC’s November 2022 statistical report available on its website.

The state block group population data was pulled from the 2020 U.S. Census and the block group geometry was derived from the Census TIGER/Line data. Each block group has both the total population and specifically the population of people in grade school as separate values, as well as the totals for the separate survey answers. I calculated the analysis with the total population to better represent the demand on services. This serves as a proxy for children’s access. The block groups also have socioeconomic factors attached to each polygon from the Census’ 2022 American Community Survey (ACS) estimates. The ACS is a continuous survey which collects and disseminates vital information on a yearly basis from U.S. residents via randomly selected addresses. I aggregated the individual variable spreadsheets and joined them to the corresponding block groups, as well as calculated and combined data for certain variables such as households with no spouse or partner present as it is separated by men and women. Removing unnecessary

---

4 The partner data focused on family households where the head of house had no partner and did not account for cohabitation with other adults such as grandparents or other relatives.
data for factors I was not including was also a key step prior to analysis. A large limitation of this data was less so a fault of the data and more so a complication of a multitude of factors contributing to levels of accessibility. Only so many factors could be chosen for analysis due to project constraints and software limitations.

The final set of main data I needed for this project is built into ArcGIS Pro and accessible with WVU’s institutional ESRI license. I needed both prior data sets to use it. The library sites only required location correction and the addition of the book collection count. The Census data needed to have each factor aggregated from the individual survey results. After combining each of the surveys, I joined the table to the Census block groups and calculated the percentage of each factor. With the ACS data joined, the largest limitation of the data was ArcGIS Pro’s OD Cost Matrix tool. I could only run up to 1,000 points at a time, and there are 1,592 Census block groups in WV. To address this, I grouped the block group centroids based on their proximity to each other by utilizing the Point Aggregation tool in ArcGIS Pro with a five-kilometer distance buffer. Once I had the resulting polygons, I created a centroid point for each and joined the new data with the original layer for a total of 818 points. These new centroid points represented the location of the aggregated points. Statistics were updated to represent the combined totals of each.

I set the maximum travel time threshold for the tool to 20 minutes’ drive time during the weekdays in the afternoon, accounting for school and work end times. This considers the topography of WV and its road networks. As a result, I extracted 1,054 pairs of facilities and block group centroids. A limitation of this data is the rural aspect of the state of West Virginia. Not every road is perfectly mapped, especially in more rural areas of the state. Secondary data consists of state boundary data, also attainable from the WVGISTC, and a base map of the state from ESRI’s ArcGIS Pro.
### Table 1: Data used in study.

<table>
<thead>
<tr>
<th>DATA</th>
<th>SOURCE</th>
<th>NEEDED</th>
<th>INCLUDES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Library Site Points</td>
<td>WVLC and WVGISTC</td>
<td>Updated, normalized, checked for positional</td>
<td>Physical address and book collection count</td>
</tr>
<tr>
<td></td>
<td></td>
<td>accuracy</td>
<td></td>
</tr>
<tr>
<td>Census Block Groups</td>
<td>U.S. Census</td>
<td>Normalized, aggregated with ACS survey results</td>
<td>Block group geometry and socioeconomic factors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and population, aggregated points</td>
<td></td>
</tr>
<tr>
<td>Road Networks</td>
<td>ArcGIS Pro</td>
<td>Ran in OD Cost Matrix tool using library sites</td>
<td>Travel time along road networks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>and Census block group centroids</td>
<td></td>
</tr>
</tbody>
</table>

#### 3.2 Socioeconomic Accessibility Index

To model socioeconomic access, a complex issue which affects many variables, and many variables affect, five statistics from the 2020 American Community Survey five-year estimates were chosen for their impact and relevance to approximating children’s ability to access libraries (Table 2). Using only five variables reduces the accuracy of the assessment, but due to feasibility and complexity there had to be a limit. There is also no real consensus of which variables are the most relevant, so each variable was narrowed down from a list of potential factors from the literature to ones that best fit (Hong et al., 2023). Following Hong et al. (2023) I assigned each of the variables to Census block groups as percentages. The resulting combined percentage of each variable represents the ranking of the individual Census block group, ranked into five separate classes from the top 20% to the bottom 20%.

The first statistic is the percentage of households whose income in the past 12 months is below the Federal Poverty Level. This is one of two variables most commonly utilized in previous works (Cheng et al., 2021; Guo et al., 2017; Hertel & Sprague, 2007; Rosichan, 2020; Zakrajšek
Poverty is an economic impediment to affording travel (such as cost of gas, car maintenance, and public transportation fees) or library services (such as costs of late fees, damaged books, and falsely assumed cost of entry). Second is the percentage of households with no vehicles available, following recommendations from Hong et al. (2023). This is an indirect indicator of economic status. It also suggests transportation barriers, as public transportation is minimal in much of West Virginia and walking is not always an option. Most minors rely on vehicle transportation as the driving age in West Virginia is sixteen. The third statistic is the percentage of households without internet, which was included from multiple discussions with reading outreach program providers. This is another indirect indicator of economic status like households in poverty but also correlates with information access and state infrastructure. For example, so many things are digital, now more than ever. Having access to library webpages highly facilitates knowing about library locations, their services, or events.

Fourth, I included the percentage of heads of household without a high school diploma. Educational attainment is the second of two persistent variables used in library accessibility assessments (Cheng et al., 2021; Rosichan, 2020). This is an indicator of influence of the education level on access. This links back to Neuman’s 1999 analysis of low ownership of books in the home correlating with low academic achievement. Another example of its relevance is how higher education often requires the ownership of books by default. Fifth, I have selected the percentage of householders with no spouse or partner present once again following the recommendations of Hong et al. (2023). This is another indirect indication of economic hardship in households. Guardians often have difficulty with juggling work, their needs, and the needs of their children.

---

5 It is important to note that households below the poverty level or heads of households with less educational attainment may highly value access to books and own many despite their decreased levels of access to them. This is a complex issue beyond the scope of this project.
even when there is a second guardian present. A single parent with limited time to make trips to a library will struggle with this as well. Finally, population counts for individuals 3 years old and over enrolled in school (kindergarten to 12th grade) was also pulled as a reference statistic. This population count was not used in the calculation of the index.

<table>
<thead>
<tr>
<th>Variable 1: Households in Poverty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable 2: Householders Without High School Diploma</td>
</tr>
<tr>
<td>Variable 3: Households Without Vehicle</td>
</tr>
<tr>
<td>Variable 4: Households Without Internet</td>
</tr>
<tr>
<td>Variable 5: Householders Without Spouse or Partner</td>
</tr>
</tbody>
</table>

*Table 2: List of variables for the socioeconomic index.*

### 3.3 Spatial Accessibility Index

To map spatial accessibility I scripted and utilized a modified Extended Two Step Floating Catchment Area tool based on a combination of library access studies and specifically Hong et al.’s 2023 article (Huhndorf & Działek, 2017; Luo & Qi, 2009). The code which assisted the creation of my own tool was created by Lina Zhang (Zhang, 2021). To utilize the tool, I needed the three data sets discussed in section 3.1. The first was the geocoded locations of library sites and their supply capacities (number of books in their collections); second, the population of each block group applied to a centroid of the block group; and third, the calculation of the distance from these block group centroids to the library sites. Figure 2 shows the first step of E2SFCA which calculates $R_j$, a population-to-supply ratio of the site of interest $j$ using the above pieces of data courtesy of Hong et al.’s 2023 approach.
\[ R_j = \frac{S_j}{\sum_{j \in \{d_{ij} \in \tau\}} P_i W} \]

*Figure 2: Equation for step one of the E2SFCA tool.*

\( S_j \) is the supply at the library location \( j \), \( P_i \) is the population of Census block group centroid location \( i \). The travel time between the locations is \( d \) and \( W \) represents the distance weight for said travel. Figure 3 shows the second step also used by Hong et al. (2023), the accessibility index \( A_i^F \) which is the summation of the results from the previous equation. Each equation repeats the function for each library site within each catchment area for each Census block group.

\[ A_i^F = \sum_{j \in \{d_{ij} \in \tau\}} R_j W \]

*Figure 3: Equation for step two of the E2SFCA tool.*

Due to my use of the OD Cost Matrix tool to calculate the travel time instead of the Google Maps API as Hong et al. (2023), the way I approached distance weight differed here based on what I was capable of reproducing through ArcPy code. Figure 4 shows the equation I utilized to calculate the distance weights, \( W \). The distance decay coefficient is represented by \( \beta \).

\[ W = e^{-(d_{ij} - 2)^2 / \beta} \]

*Figure 4: Equation for the distance weight model.*

### 4.4 Integrated Accessibility Index

For a comprehensive spatial-socioeconomic index, I combined the two previous indices once again modeling Hong et al.’s approach (2023). As previously mentioned, there is no consensus on how to best integrate spatial and aspatial accessibility into a single assessment. I
gave equal weight to both indices as I have no reason to weigh one higher than the other, then added them together to obtain their average (Hong et al., 2023). I normalized the data to account for any differences. The resulting index was classified in the same percentile classification as the previous two with the five classes of low to high accessibility levels.

IV. Results

I created three indices to assess locations in West Virginia with low accessibility to physical books. All three indices were calculated for Census block group units. The first index looks at socioeconomic barriers to access, the second at spatial barriers, and the third a combination of both. I symbolized all indices in five classes to compare the lowest 20% through the highest 20%. An outline of WV counties is included to assist with location recognition. Low access values are in shades of red, high access values are in shades of blue, and moderate values are in purple. Figure 5, the socioeconomic index, reveals a scattered pattern of access trending towards lower values in more rural areas and southern West Virginia. Table 3 examines the statistics for each of the variable percentages.

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>MAXIMUM*</th>
<th>AVERAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Households in Poverty</td>
<td>62.5%</td>
<td>17.7%</td>
</tr>
<tr>
<td>House Owners Without Diploma</td>
<td>40.7%</td>
<td>12.7%</td>
</tr>
<tr>
<td>Households Without Vehicle</td>
<td>62.3%</td>
<td>8.8%</td>
</tr>
<tr>
<td>Households Without Internet</td>
<td>64.2%</td>
<td>18.3%</td>
</tr>
<tr>
<td>Householders Without Partner</td>
<td>49.9%</td>
<td>15.9%</td>
</tr>
</tbody>
</table>

Table 3: Maximum and average of percentages of each variable, rounded to the nearest tenth. Highest values highlighted for clarity. *Maximum values do not count outliers.
The variable with the highest overall average was households without internet. The largest cluster of no internet households is in Wayne County between Sydney and Wilsondale, WV. The highest percentage, including outliers, is in Martinsburg, WV at 85% of the Census block group having no internet. Households below poverty are second highest overall, with the highest percentage of respondents below poverty being 100% in a block group in Morgantown, WV. The next highest percentage is also in Morgantown at 87%.

Figure 6 shows the spatial index. Clustered patterns of higher access levels emerge around major urban areas while more mountainous and rural areas trend towards lower access levels. Figure 7 is the combination of the spatial and socioeconomic indices, highlighting the combined average access level from the previous two indices. A total of 314,212 individuals, or 17.4% of the total population of the state live in low access areas. Table 4 shows the total population distribution for each level of access for each index.

<table>
<thead>
<tr>
<th></th>
<th>SOCIAL INDEX</th>
<th>SPATIAL INDEX</th>
<th>INTEGRATED INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>VERY HIGH ACCESS</td>
<td>24.6%</td>
<td>39.6%</td>
<td>24.1%</td>
</tr>
<tr>
<td>HIGH ACCESS</td>
<td>34.1%</td>
<td>29.7%</td>
<td>32.4%</td>
</tr>
<tr>
<td>MODERATE ACCESS</td>
<td>26.5%</td>
<td>17.2%</td>
<td>26.6%</td>
</tr>
<tr>
<td>LOW ACCESS</td>
<td>12.1%</td>
<td>1.3%</td>
<td>4.7%</td>
</tr>
<tr>
<td>VERY LOW ACCESS</td>
<td>2.7%</td>
<td>12.2%</td>
<td>12.2%</td>
</tr>
</tbody>
</table>

Table 4: Population distribution for each of the three indices. Percentages rounded to the nearest tenth. Highest values highlighted for clarity.

Referring to the Integrated Index, the total population in high access zones is 1,082,577 individuals, accounting for 56% of the total population of WV. The high access zone school population is approximately 170,047 children, or 60.5% of the total minor population. In low
access zones, there are 314,212 individuals, or 16.9% of the total population with approximately 48,793 minors or 17.4% of the total minor population in the state. Approximately 26.6% of the total population lives in moderate access zones.

Figure 8 isolates the Census block groups that are within the low access categories and shows the population of grade school students in each of these sections. Approximately 17.4% of students live in low access zones. The town of Clay in Clay County and the immediate surrounding areas are highlighted as the most vulnerable with the highest population. A secondary area of concern is Dunn, WV in northern Mercer County. Figure 9 is an analysis of the zones which have zero to one library within a 20-minute drive of their center. Approximately 13% of the state lacks an accessible connection to a library. Many of these zones with no connections are low in population and rural, but that does not mean there is not a need for access.
Figure 5: Socioeconomic Access Index map.
Figure 6: Spatial Access Index map.
Figure 7: Integrated Access Index map.
Figure 8: Map of estimated grade school student population in low access areas. Highlighted in the inset map is Clay County and its surrounding area. Two of three zones with the highest population are in the center of the state.
Figure 9: Map of zones with no libraries or a single library within 20 minutes’ drive time.
V. Discussion and Concluding Remarks

Part of the inspiration behind this thesis comes from my personal experience as a child with a school system that lacked the same reading resources as the rest of the elementary schools I attended. In the northeast corner of the inset map highlighting Clay County as a vulnerable zone, is the town of Sutton, West Virginia. I attended Sutton Elementary School for part of the 4th grade, and even then, at the age of 8, was I able to discern the difference in quality between my previous schools and this one. Occasionally we would read a story aloud in class, but I only recall a total of two books and a single teacher who ran the reading program for the entire school. There were no art or music teachers. The experiences I had in this elementary school led to a slow decline of my once intense love of reading, which is something I wish no child experiences.

To begin aiding in the work of preventing similar experiences, this thesis created multiple maps and tables indicating the levels of spatial, socioeconomic, and socio-spatial access children have to physical books across West Virginia through public libraries. Additionally, an updated layer of libraries and Census block groups with more current population counts across WV will be provided to the WVGISTC. The map and data will be publicly accessible and available for reproducibility and future research projects. In addition, this project contributes to the scholarship on accessibility mapping methodology by creating data on the use and variation of methods specifically on a rural state scale. West Virginia is a state that has complex topography and is sometimes difficult to perform geographic research for. It is my hope that this analysis will aid not only to my own future research but to others combating insufficient childhood access to books. The analysis scale of Census block groups utilizing E2SFCA gives us a more accurate representation of library accessibility than we previously had access to.
Scale is one of the major challenges of this thesis. Ideally the scale with the most accuracy would be the smallest possible, in this case I would utilize home addresses. A common scale of analysis, often out of necessity due to data, is the county. However, detail and accuracy are lost in the summation and averaging at this level. The next level down from this is the Census tract, then the Census block, and finally the block group. Of course, a local, neighborhood analysis would be even more optimal, but obtaining data for this level of specificity is difficult. There are pros and cons to looking at an area such as a state for spatial analysis. It is important for the state to have access to state level data, as local data often leaves out surrounding complexities and statewide context. Most of it comes down to an issue of political versus actual boundaries as well. In the case of libraries, most offer membership based on state residency, blurring the boundaries of counties.

Something that also surpasses county boundaries is topography. West Virginia is heavily forested and covered in mountainous terrain. This is one of the reasons straight line distances are not optimal, as road networks often require navigation around difficult terrain and cannot travel along a straight path, increasing drive times. A factor that can alleviate the difficulty in accessing a resource would be public transportation, something made difficult by both the terrain and complex road networks as well as rural communities. Public transportation is expensive and requires quality infrastructure. Rural road networks are less likely to be complete and the quality of roadways themselves is also called into question.

All these factors contribute to the resulting levels of access in West Virginia, and all of them affect each other, whether in positive or negative manners. To understand the full picture, one must understand the individual parts which surmise the issue. It is in this spirit that I examine both spatial and socioeconomic factors separately as well as in conjunction. It is not only one part which affects accessibility. Overall, one fifth of the state had little to no access to a library within
20 minutes of drive time, much of this a result of rural areas with low populations, but not one of these block groups had no population. Almost one fifth of the state’s population of minors live in these areas. While most children have decent access, any amount that does not is an issue worth addressing. A lack of internet data of any kind and prevalent poverty levels are factors that highly contributed to low levels of access. When considering population distribution, spatial access was better overall with 68% living in high access zones. This still leaves 32% of our state with moderate access or worse.

The initial limitations of the analysis producing these results was data availability and data age. There was no current spatial data for libraries in the state of West Virginia, the previous data being from 2002. In the past 20 years we have gone from 178 to 169 libraries, including openings as well, but there is a question of whether the new libraries are an improvement upon the old or not. There are a multitude of key factors which needed to be included but due to the scope of a thesis I had to cut content. Without solid consensus on which variables and aspects are the most vital to this analysis it is hard to narrow down the list. A third limitation dealt once again with the geography of a rural leaning state, as most models are not developed with these areas in mind and the quality of data is questionable. After the data was collected and updated, most issues were data manipulating, and data analysis related. The first issue was the limitation on points used in the OD Cost Matrix tool being 1,000 when there are over 1,600 Census block groups, leading to the aggregation of about eight hundred centroid points. In retrospect using a smaller grid for point aggregation to be closer to that 1,000-point limit would be a better representation of the data. Much was learned during the process and will impact my future research.

The main purpose of this thesis was to build a foundation for this future research, as it is hard to know where to look at lived experience without knowing where that experience is being
lived. Future research to improve upon the methodologies used here often deal with the issues of scale and rural analysis. As Hong et al. 2023 outline, the distance decay weights need to be supplemented with real travel times. They also outline weighing the socioeconomic variables, which requires in depth study on what factors most affect a person's ability to access a public library. I would like to see this analysis done on a smaller scale within major population centers in the state, as the results hint towards more realized variability. More research into how WV residents utilize their libraries and when would also benefit knowing what time of day to base the drive time analysis on and how much travel time is too much. A minor’s view specifically is important and complex due to the dependence on caregivers. Aside from the promotion of equitable access to valuable reading material, libraries also offer other community services. The benefit they can provide, especially in a rural setting, is something I aimed to garner support for.
VI. References


Rosichan, S. (2020). Application of GIS in Mapping of Adult Literacy Outreach Programs and
https://doi.org/10.1080/01616846.2019.1657789

https://doi.org/10.1111/gec3.12592


https://doi.org/10.1068/a43219

https://doi.org/10.1016/j.lisr.2020.101013