RECALCULATING THE WALKABILITY INDEX FOR POCATELLO AND CHUBBUCK, IDAHO

Danielle Huerta, GIS Technician, Idaho State University, GIS Training and Research Center
Keith Weber, GISP, GIS Director, Idaho State University, GIS Training and Research Center

ABSTRACT

Walkability for the Cities of Pocatello and Chubbuck, Idaho were determined following the same process used by the US Environmental Protection Agency (EPA) and the 2019 National Walkability Index (NWI). The recalculation was done to determine if changes that occurred since 2019 (e.g., the 2020 Census) had an effect on walkability scores. The results of this study show that while overall trends are very similar, several areas in these cities resulted in higher walkability scores. This was not due to changes in population demographics but rather to the inclusion of transit stop data which was omitted by the EPA in their 2019 assessment. It is recommended that transit stop data be provided to the EPA when a revision to the NWI is announced. In addition, it was learned that the current NWI is calculated using road data as a surrogate for sidewalks that actually enable pedestrian traffic. An incorrect assumption is made in the NWI that roads always have sidewalks adjacent to them. This is not a correct assumption and could be corrected by acquiring sidewalk data in addition to current road data.

KEYWORDS: *National Walkability Index, Sidewalks, Streets, GIS, Transit stops, bus stops, Census blocks, Census block groups*

INTRODUCTION

A 2015 study conducted by the National Association of Realtors and Portland State University, found roughly half of the people surveyed, responded they would prefer smaller living quarters if it meant being closer to amenities over living in larger quarters with less accessibility to goods and services. People are more likely to use nearby amenities and very likely will walk to these amenities if it can be done in a safe manner (i.e., well-lit and maintained sidewalks). This is an example of an area that has high walkability. Thus, walkability encourages the community to be more active by travelling to work, school, or community events by walking or biking. This encourages more socialization among people and in turn can aid in overall physical and mental health[1]. One can also see a decrease in pollution in areas where there is less vehicle traffic[2]. In addition, walkable areas require less infrastructure, because fewer streets and land are needed.

With a growing concern of a community's walkability, the EPA created the National Walkability Index (NWI) scoring system. This is used to calculate a walkability index for a given area. The NWI uses census block groups (CBG) as the area of analysis in order to make the assessment more meaningful. As a result, NWI CBG scores can be compared in relation to one another. The NWI score is created using data in the Smart Location Database (SLD). The SLD summarizes demographic, employment, and built environment factors. This includes residential and employment *density*, land use *diversity*, *design* of the built environment, access to *destinations*, and *distance* to transit. The EPA also allows access for individuals to view an interactive map that shows SLD scores across the country[3]. A SLD technical document and user guide are also available[4] (Ramsey and Bell 2014).

---

[1] Glazier, R.H., et al. "Density, Destinations or Both? A Comparison of Measures of Walkability in Relation to Transportation Behaviors, Obesity and Diabetes in Toronto, Canada." *PLoS ONE* 9.1 (2014). https://doi.org/10.1371/journal.pone.0085295
[2] Younger, M., et al. "The Built Environment, Climate Change, and Health: Opportunities for Co-Benefits." *Journal of Preventive Medicine* 35.1 (2008): 517-526. https://doi.org/10.1016/j.amepre.2008.08.017
[3] https://epa.maps.arcgis.com/home/webmap/viewer.html?webmap=137d4e512249480c980e00807562da10
[4] Ramsey, Kevin. Bell, Alexander. "Smart Location Database." Environmental Protection Agency (EPA), 2014. https://www.epa.gov/sites/default/files/2014-03/documents/sld_userguide.pdf

For this study, the NWI and SLD were used to reassess walkability for the cities of Pocatello and Chubbuck, in Bannock County, Idaho using current data. This was completed using two spatial scales, the first was CBG --with recalculated walkability scores-- and the second used more spatially resolved census blocks (CB) to determine walkability in a more granular approach. Results from this study were compared to results from the EPA and their SLD walkability scores which were last updated in February, 2022.

METHODS

There are three factors used to calculate the walkability index: (1) diversity of land uses, (2) road intersection density, and (3) proximity to nearest transit stop. The diversity of land use factor used two sub-factors; employment mix and employment household information. Intersection density was based upon the density of roads at each intersection. Finally, the nearest transit stop factor used the distance to bus stops within service areas. Service areas are those areas adjacent to a transit stop and its connectivity to areas within 0.5 miles of the transit stop. Areas outside each service area received a null response in the resulting geodatabase. Each of these factors help to better understand walkability. Once each factor was calculated following the same methodology described in the SLD technical document, these factors were assigned a weighted value which was then used to calculate the final walkability score. All processing and calculations followed the same techniques used by the EPA and SLD (**Equation 1**). To ensure this was completed correctly, each factor and sub-factor result was carefully scrutinized and compared to SLD data for the same area before the methodology was applied to CB areas.

Equation 1. The formula used to calculate the Walkability Index score

$$\text{Walkability Index score} = \left({}^{w}/_{3}\right) + \left({}^{x}/_{3}\right) + \left({}^{y}/_{6}\right) + \left({}^{z}/_{6}\right)$$

Where w = CBG ranked score for intersection density
x = CBG ranked score for proximity to transit stops
y = CBG ranked score for employment mix
z = CBG ranked score for employment and household mix

Each factor required preprocessing before the layers/attributes were ready for use. These preprocessing steps are described in the following sections. Once preprocessing was complete, ArcGIS Pro model builder tools were used to replicate the EPA's process of finding the *w, x, y, and z* factor scores for use in equation one and to derive the final walkability index score.

Intersection Density (w)

To find the intersection density factor (*w)*, a comprehensive and accurate street layer was needed. This layer includes streets, pedestrian pathways, and biking trails. This was accomplished by appending the pathway and trail layers to a comprehensive streets layer. Each street, path, and trail were described using attributes (fields) that identified street type (e.g., ramp, street, loop), speed limit, one-way direction and street use type (auto, pedestrian, both). Intersections were found by creating points where the lines features met using the Generate Points Along Line tool in ArcGIS Pro. This tool did not produce perfect result however. The results were carefully reviewed using aerial imagery and other supporting basemaps and intersection points were placed manually by systematically reviewing the study area using a fishnet layer polygon grid to ensure no area was omitted from review. Intersection points were deleted if an automated point not a true intersection. In cases where two streets with different names met each other and formed a corner, a point was placed and was considered an intersection. To further validate the intersections layer, a comparison was made with an intersection point layer provided by Bannock Transportation Planning Organization (BTPO).

Once the intersection layer was created and validated, each point was buffered using the Pairwise Buffer tool in ArcGIS Pro and a buffer distance of three meters was used. The number of streets entering an intersection was determined by intersecting the streets layer with the intersection buffer layer followed by using the Aggregate tool in ArcGIS Pro.

Speed limits were determined and added to each street/line where that data was not already present. Speed limit is taken into consideration in the SLD, because pedestrians are more likely to use a sidewalk where vehicles are driving at a slower speed. The speed limit on bikeways was that of the street it was adjacent to, and the speed limit for pedestrian pathways were set to a value of null. In addition, one-way directions for street were identified and classified as such. This was used to help determine street type use. Once all these attributes were populated, these data were ready to be run through the `Intersection_Cat` model (this and all models are included in the project deliverables zip file, available at https://giscenter.isu.edu/research/Techpg/BannockGIS/zip/walkabilitydeliverables.zip). This model builder tool used the following factors to determine the intersection density score (*w*); speed limit, street type, presence of one-way direction roads, and the count of streets at each intersection.

Proximity to Transit Stops (x)

For the proximity to transit stops factor (x), all bus stops were first identified and mapped. The inside service area (ISA) polygon layer was also used. The ISA encompasses all areas that the City of Chubbuck and City of Pocatello provide bus services for. These areas are within 0.5 mile from a bus stop and have complete sidewalk connectivity to the bus stop" Once bus stops and the ISA layer were added to the map, these data were run through the `D4a_CBG_TransitProx` and `D4a_CB_TransitProx` models. These models calculated the score for the *x* factor.

Employment Mix (y) and Employment/Household Mix (z)

For the employment and household factors (*y and z*), the location of all structures in the study area was needed. Creating these data followed a process using a fishnet polygon grid overlaid onto the study area. Points were digitized for each structure seen on current aerial imagery (2018 Pictometry and 2021 NAIP imagery). The population count for each home was determined using the occupancy rate and population per household data from the 2020 Census in each census block, respectively. Once the structures layer and population attribute were complete, a population-weighted centroid was determined for both the CBG and CB areas using the Mean Center tool in ArcGIS Pro.

To determine the score for the *y* factor, the EPA used an 8-tier entropy value based on employment classifications to calculate the employment mix. This was completed in the `D2b_E8MixA` model. Another entropy calculation was used to find the *z* factor, but this included both employment and occupied housing values as well. This calculation was completed using 5-tier employment categories. The `D2a_EmpHHm_CBG` and `D2a_EmpHHm_CB` models were used for this process.

Spatial Interpolation of Walkability Scores

The Inverse Distance Weighted (IDW) interpolation of walkability scores was done using CB data to help gain a better understanding of where walkability should be highest and lowest if walkability can be considered a linear or continuous variable within a city. To complete this analysis, CB centroid points with walkability scores were used in the IDW interpolation tool in ArcGIS Pro. This resulted in a raster layer for the study area predicting where walkability should be high or low, assuming walkability is a continuous variable.

Maps

Maps were created for intersection density as well as proximity to transit stops to aid in the visualization of each data set and layer. Maps were also created for both CBG and CB using the walkability scores calculated in earlier steps. In addition, an interpolation was done where CB centroids were used to for hot spot analysis. The results from the hot spot analysis were used in the Inverse Distance Weighted (IDW) tool to create a raster of interpolated high and low walkability values. This was raster was clipped to include only the cities of Pocatello and Chubbuck.

RESULTS AND DISCUSSION

One can see from **Figures 1 and 2** that the areas with high walkability are located in the core of Pocatello and Chubbuck. Though it seems the highest scoring areas are located on the outside of the main high-scoring grouping. The areas with low walkability scores are areas that do not include much housing, businesses, intersections, or easy access to transit stops. This outcome reinforces the idea that calculations represent reality, since low and high scoring areas are where one would expect them to be.

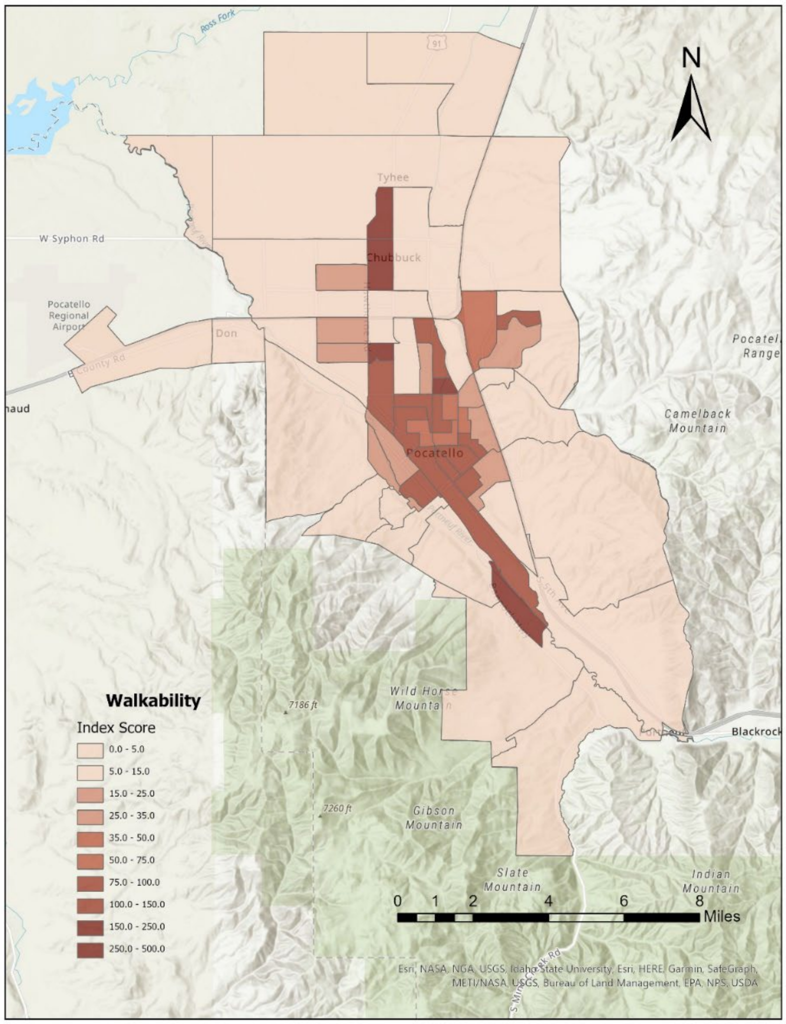

Figure 1. Walkability Index scores of Bannock County, Idaho base on census block groups (CBG).

Errors may be present due to misrepresentations of the factors used to calculate the walkability score. Some new streets may not have been considered if they were not already included in pedestrian pathway, street, or bike trail layers provided for this project. In addition, there was a chance for human error to occur when intersections were being identified, even with the aid of the fishnet polygon and systematic approach to identifying intersections. Speed limits of streets may also not be representative of reality. A posted speed limit does not always mean vehicles will travel this speed. It is also important to note the score of the CBG may not be representative of the entire area within the sometimes very large census block group polygons. There are likely areas within the CBG that have very high or very low walkability score. One can see this to be true when looking at just CB (**Figure 2**).

The CB map (**Figure 3**) shows a similar pattern as seen in the CBG analysis, but with higher walkability scores being more centralized and granular. This gives a more complete breakdown of blocks, making it easier for one to find the true area where walkability is high or low. One can see that the centralized area with high scores becomes even larger when looking at CB, because of the way CBG are grouped. When grouped, high scoring CB areas may not have as large an impact on the CBG score because of the other low scoring CB grouped with it.

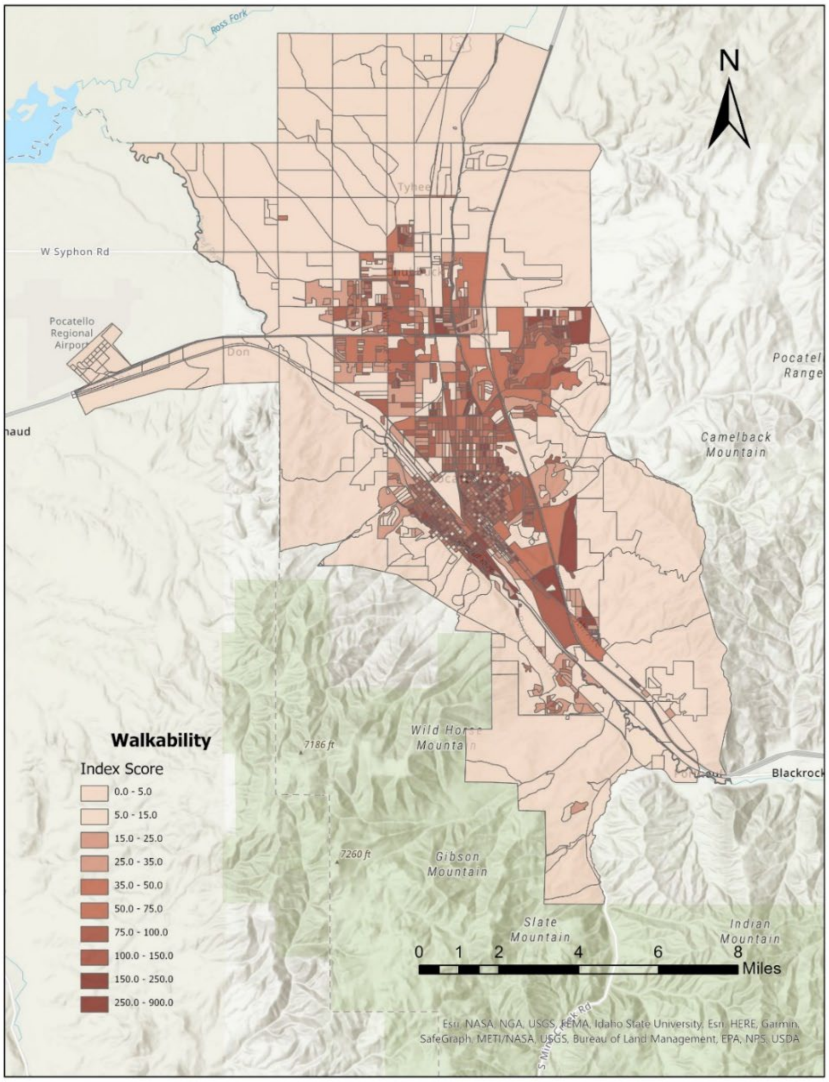

Figure 2 Walkability index scores for Bannock County, Idaho based on census blocks (CB).

Statistics were calculated for both CBG and CB. The polygons used only included those where all data was available. In other words, polygons were included when their score was not equal to zero or null. There were 53 CBG polygons that contained a full suite of data (*n* = 53). While the walkability index can range up to a value of 900, the mean walkability score for the study area was 55.6 and the median score was 24.4 (s.d. = 72.9). This difference between mean and median indicates the average walkability score was impacted by several high scoring areas since the median is quite a bit lower than the mean. In addition, overall walkability is quite low outside the cities of Pocatello and Chubbuck. This is unsurprising however as Bannock County is largely rural. Furthermore, the range of values and differences observed between mean and median demonstrate these data are not normally distributed. This is not surprising in a built environment however. The high standard deviation (exceeding the mean) indicates high variability exists throughout the study area. Applying one standard deviation to the mean, allows the median to be encompassed at the 66% CI.

The CB polygons were treated in the same way as CBGs, resulting in a total of 1,201 polygons that had scores greater than zero (n = 1,201). The mean walkability score for CBs was 94.9 and the median score was 73.2 (s.d. = 95.6). The average score for CB was noticeably higher than CBG with a difference of 39.3. The median score also showed an increase. This suggests the finer granularity provided by the CB is advantageous to understanding walkability.

Based on the IDW interpolation (**Figure 3)** the concentration of higher walkability scores is located in the heart of Pocatello (i.e., historic old town). One could have likely inferred this on their own, because old town Pocatello has a high concentration of businesses and households, smaller street blocks, and access to transit stops. These results further reinforce the results described above.

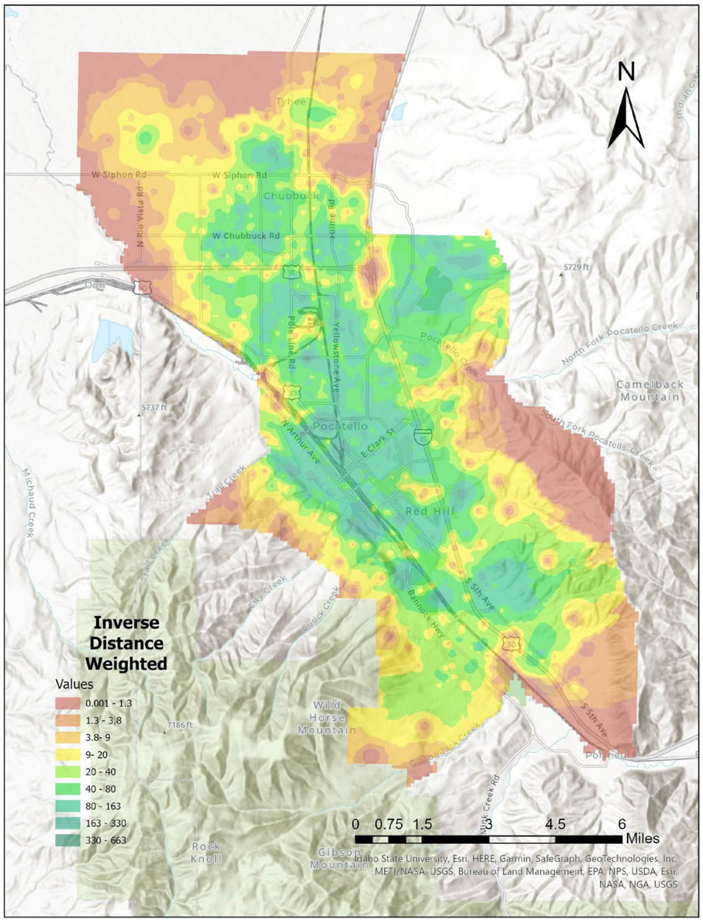

Figure 3. An Inverse Distance Weighted interpolation of walkability in Bannock County, Idaho. Since walkability is not a continuous variable the fit between mapped.

There are differences between the results produced by this project and the EPA's SLD (**Figure 7**). The biggest source of this difference was due to transit stop information. The EPA did not use proximity to transit stop information when calculating their walkability scores for Pocatello and Chubbuck as evidenced by null values in the smart database tables. This omission may be because the EPA did not have a GIS layer of transit stops for Pocatello and Chubbuck, Idaho. However, the calculations completed in this study did include transit stop locations and resulting scores. This directly impacted the final score with the EPA's scoring being lower (as expected) than the walkability scores calculated for this study. Since the transit stop factor was considered in our recalculation, the scores found in this study should be considered more complete and better representative of walkability.

An error in the logic of the walkability index scoring system is the assumption that all roads have sidewalks on either side of the road. This assumption does not reflect reality and the use of actual sidewalk locations (like that done in this study) is a better approach. However, a comprehensive GIS layer of sidewalks for the entire United States is not readily available and this was likely the reasoning for the use of roads as an approximate/surrogate for sidewalks.

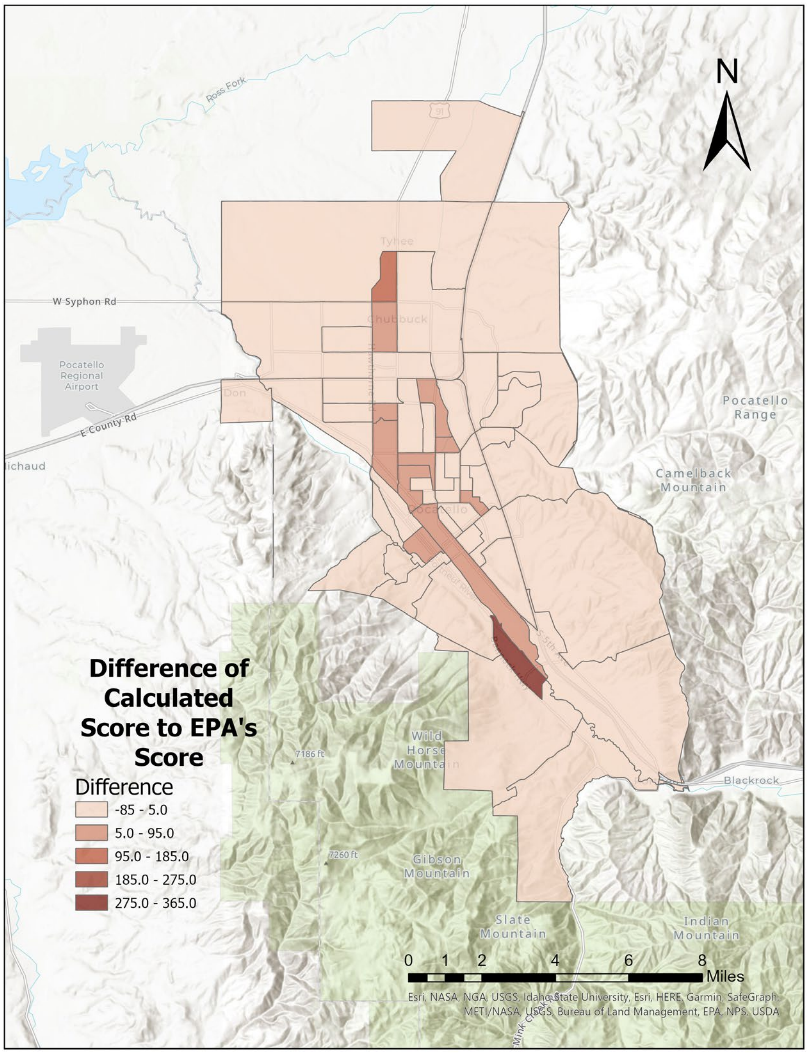

Figure 4. Difference between walkability index scores calculated in this study and the EPA's reported walkability index scores based on census block groups (CBG). The majority of differences occurred where transit stops actually exist in the cities of Pocatello and Chubbuck, Idaho.

CONCLUSIONS

The SLD user guide allowed the authors to follow the EPA's process of calculating walkability scores. Following this process, the EPA's results and the resulting walkability scores calculated for this study were found to follow similar trends with high and low walkability scores occurring in similar areas. While overall trends were similar, quite a few CBG scores increased in this study's 2023 analysis. This difference was attributed to including transit stops which were omitted from the EPA's 2019 SLD. Proximity to transit stop is one of four primary factors used to determine walkability.

An error in the logic of the walkability index scoring system is the assumption that all roads have sidewalks on either side of the road. This assumption does not reflect reality and the use of actual sidewalk locations is a better approach.

APPENDIX A: Map of the streets layer and intersection points.

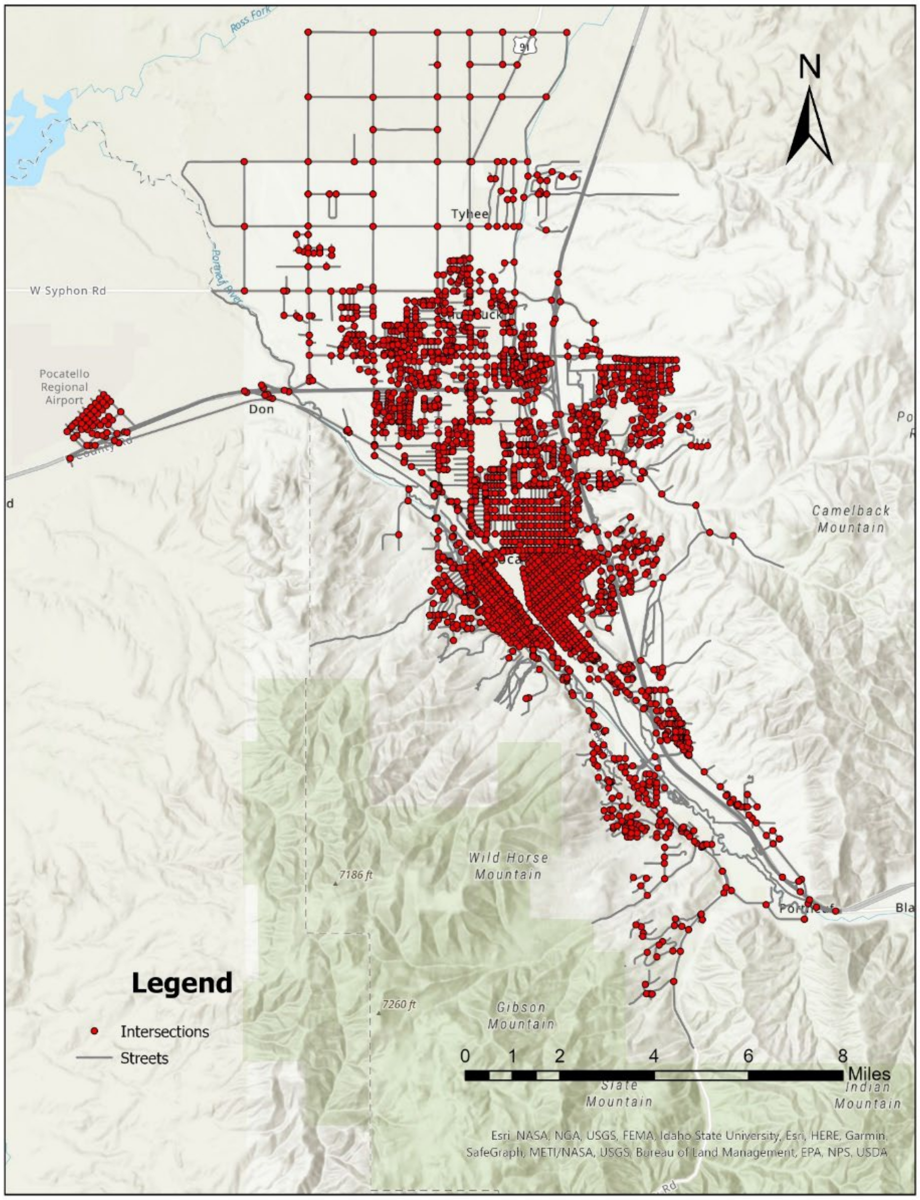

Figure 5 "Streets", shown by gray lines, and intersections, shown by red dots, found in the cities of Pocatello and Chubbuck

APPENDIX B: Map of bus stops and Inside Service Areas polygons

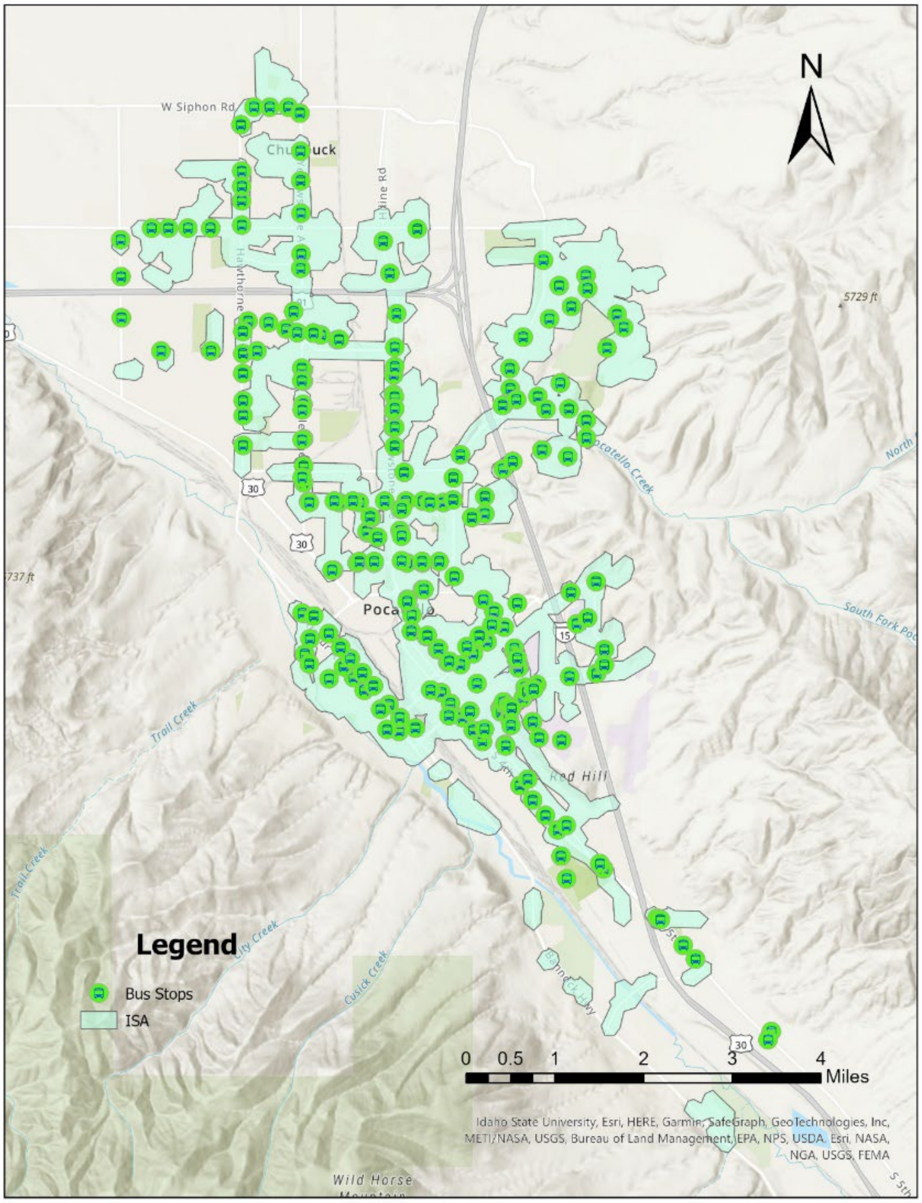

Figure 6 Pocatello's and Chubbuck's bus stops as points. The polygons shown are areas are the "Inside Service Areas."