OAK RIDGE NATIONAL LABORATORY

MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

ORNL/TM-2014/54

Final Report for Phase II Study: Prototyping the Sketch Planning Visualization Tool for Non-Motorized Travel

February 2014

Prepared by
Ho-Ling Hwang, Ph.D.
Daniel Wilson
Timothy Reuscher
Shih-Miao Chin, Ph.D.
Rob Taylor

UT-BATTELLE

ORNL-27 (4-00)

ORNL/TM-2014/54

Developing a Visualization-Based Sketch Planning Tool for Non-Motorized Travel

Final Report for Phase II Study
Prototyping the Sketch Planning Visualization Tool for Non-Motorized Travel

February 2014

Prepared for
Federal Highway Administration
U.S. Department of Transportation

Prepared by
Ho-Ling Hwang, Ph.D.
Daniel Wilson
Timothy Reuscher
Shih-Miao Chin, Ph.D.
Rob Taylor

Center for Transportation Analysis
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831-6283
Managed by
UT-BATTELLE, LLC
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACS	American Community Survey
ADA	American Disabilities Act
BTS	Bureau of Transportation Statistics
D	Destination
EPA	Environmental Protection Agency
ESRI	Environmental Systems Research Institute
FARS	Fatality Analysis Reporting System
FEMA	Federal Emergency Management Agency
FHWA	Federal Highway Administration
GIS	Geographical Information System
GTFS	General Transit Feed Specification
HH	Households
HTS	Household Travel Survey
ID	Identification
JTW	Journey-to-Work
LUM	Land Use Mix
MD	Maryland
MPO	Metropolitan Planning Commission
MWCOG	Metropolitan Washington Council of Governments
NCHRP	National Cooperative Highway Research Program
NHTS	National Household Travel Survey
NMT	Non-Motorized Travel
O	Origin
O-D	Origin and Destination
ORNL	Oak Ridge National Laboratory
SLD	Smart Location Database
SPT	Sketch Planning Tool
TIGER	Topologically Integrated Geographic Encoding & Referencing
TPB	Transportation Planning Board
VA	Virginia
WMATA	Washington Metropolitan Area Transit Authority

ACKNOWLEDGEMENT

This project has benefited from the contributions of many individuals and organizations. The authors would like to specifically thank Robert Griffiths of the National Capital Regional Transportation Planning Board at the Metropolitan Washington Council of Governments for providing access to their 2007-2008 Regional Household Travel Survey data, as well as their 2011 Geographically-Focused Area Survey data.

The authors would also like to thank Mr. Martin Howes of Prince George’s County, Maryland and Mr. Christopher McGovern of Montgomery County, Maryland for providing transportation infrastructure data for their given jurisdictions.

And finally, the authors would like to thank Brian Gardner from the Office of Planning, Environment & Realty (HEP) in the Federal Highway Administration (FHWA) for funding this research. In addition we acknowledge Mike Barry (HEP-FHWA) for providing Washington D.C. sidewalk inventory data and Mark Sarmiento (HEP-FHWA) for providing assistance with ArcGIS Online. And most importantly, the authors are extremely grateful to Elaine Murakami (HEP-FHWA) for her help and guidance throughout this project.

1. INTRODUCTION

1.1 BACKGROUND

The idea of livable communities suggests that people should have the option to utilize non-motorized travel (NMT), specifically walking and bicycling, to conduct their daily tasks. Forecasting personal travel by walk and bike is necessary as part of regional transportation planning, and requires fine detail not only about individual travel, but also on transportation and neighborhood infrastructure.

This project was funded by the Office of Planning, Federal Highway Administration (FHWA), and was carried out in two phases. Phase I of this project was completed in 2012, which focused on using data collected under the 2009 National Household Travel Survey (NHTS) and supplemented with data from the American Community Survey (ACS), as well as several other public and private databases, to characterize the "market" potential for NMT under a national perspective. The final report documenting the effort and findings of Phase I research can be found at this website: http://info.ornl.gov/sites/publications/Files/Pub36798.pdf.

The Phase I NMT research exercise revealed that developing estimates on the propensity to travel non-motorized was, unfortunately, too large a task on a national scale. Although the reduction of geography to nine add-on areas with better NHTS data coverage did produce a relatively reasonable overall model performance under Phase I, the NHTS sampling limitation indeed impacted the model's ability to accurately estimate characteristics in certain areas, specifically in the City of Alexandria, Virginia.

To further examine how factors such as those identified from the Phase I NMT study, and the modeling framework developed under that effort could be applied to local/regional level planning activities, FHWA decided to pursue a Phase II study. It was determined that a small geographic area with more detailed local data would be necessary. Although Washington D.C. was not one of the 2009 NHTS add-ons, it did conduct a household travel survey of 11,000 households in 2007-2008. The National Capital Region Transportation Planning Board at the Metropolitan Washington Council of Governments (MWCOG) conducted the household travel survey. The data coverage under the MWCOG survey is much higher than that of the NHTS. As a part of the Phase II study, a prototype of a Geographic Information System (GIS)-based sketch planning visualization tool was also to be developed. The intent was to use a neighborhood in the Washington D.C. region as a case study for this prototype application.

1.2 PHASE II OBJECTIVES

As stated, the objectives for Phase II were:

- To utilize data and models developed under the Phase I study to conduct a case study using a selected small neighborhood, and

- To develop a prototype GIS-based sketch planning visualization tool to demonstrate how local planners can examine potential impacts of regional characteristics on walking behaviors within their community.

1.3 ORGANIZATION OF REPORT

This report documents findings from the Phase II study. This includes discussions of data sources, data limitations, descriptions of modeling approaches, and discussions on the process of developing the sketch planning tool under this NMT study. Note that the literature review associated with this NMT subject was conducted under Phase I which was detailed in the Phase I final report, thus not being repeated in this Phase II report.

The geographic regions defined for the Phase II study are presented in Section 2 of this report. In Section 3 of this report data sources considered in the Phase II study are discussed. Data sources that were new to the Phase II study are described in more detail than those used previously in Phase I. That section is followed by the descriptions of modeling approaches in Section 4. A brief discussion of results and major factors identified are also included in Section 4. Section 5 provides a description of the Non-Motorized Travel Sketch Planning Tool (NMT-SPT) prototype system along with its system requirements. Finally, a general summary and conclusions are provided in Section 6 of this report.

Several appendices are included at the end of this report. Appendix A includes lists of data elements specifically from the MWCOG Travel Survey data and the Environmental Protection Agency's (EPA) Smart Location Database (SLD). Outputs from the Discriminant Analysis run for the Walkability Index model is presented in Appendix B. The system requirements for the prototype NMT-SPT is provided in Appendix C. Finally, the list of block groups within the study area of the prototype system is shown in Appendix D, along with a few of their associated walkability measures, followed by Appendix E describing briefly how to navigate the ArcGIS online map display for the NMT Phase II project.

2. GEOGRAPHIC SCOPE

2.1 SELECTION OF GEOGRAPHIC REGION

Based on findings from the Phase I NMT project, the Phase II effort was to develop a prototype sketch planning tool using a Washington D.C. metro area neighborhood as a case study. The selection of this case-study neighborhood was initially anticipated to be from the Virginia portion of the Washington D.C. metro area because Virginia was a 2009 NHTS Add-On region included in the Phase I model development work. After the Phase II NMT study started, this initial plan had to be altered to avoid duplication with the NCHRP-08-78 study which includes Arlington, Virginia as one of its case study areas. The other option to use Alexandria, Virginia, which is widely considered as a well-developed walking community, did not appear to fit the original intention of developing a prototype sketch planning tool for a "typical" neighborhood.

Based on these concerns, and the anticipation of the newly available MWCOG Geographically-Focused Survey data (specifically the Fall-2011 Survey), the search for a candidate geographic region for this study was then shifted to Maryland. Because this region was completely outside the geographic boundary of Phase I scope, new data collection and modelling efforts became necessary. Specifically, this move of location changed the original project scope from "conducting a case study using the data/model already collected/developed from Phase I study" to performing a nearly standalone project.

2.2 STUDY AREA

After examining the regions covered under the MWCOG Fall-2011 Focused Survey (discussed in Section 3.2), the research team determined that the "Purple Line International Corridor" in Montgomery and Prince George's Counties, Maryland, should be used as the case Study Area for developing a prototype sketch planning tool under this study. Figure 2-1 shows the proposed Washington Metropolitan Area Transit Authority (WMATA) Purple Line Metrorail route [1] as well as the proposed stations.

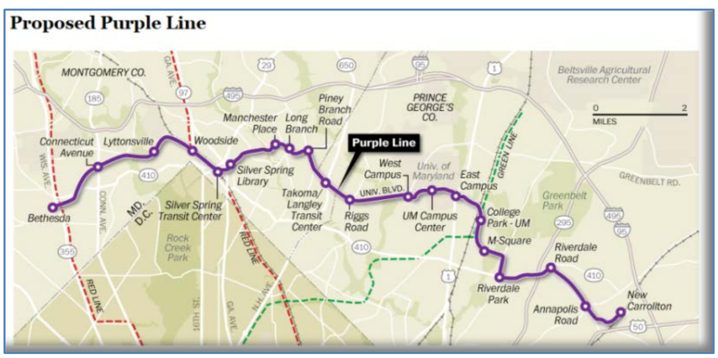


Figure 2-1. Proposed Purple Line.

The Purple Line International Corridor case study "neighborhood" includes a total of thirty-five Census block groups along the proposed WMATA Purple Line Metrorail route; east of Silver Spring, Maryland (MD) and the west of College Park, MD. These thirty-five Census block groups formed the "Study Area" under this Phase II project and are displayed under the purple-shaded region in Figure 2-2. This Study Area was then used for the development of a prototype sketch planning tool and its associated visualization applications. Note that many detailed local data (e.g., parks, crime information, fatal accident locations, etc.), which could provide valuable insight to support local planning activities, were also collected for this project.

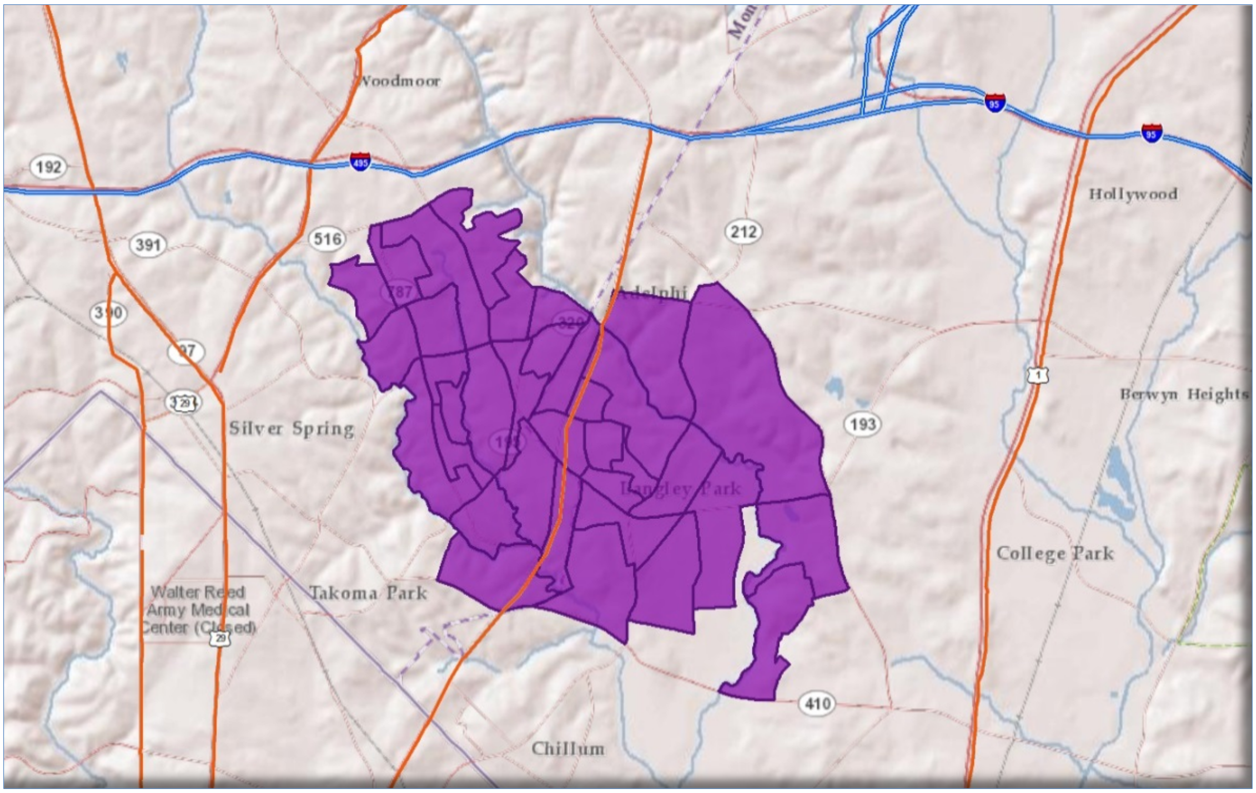


Figure 2-2. Study Area with 35 Census Block Groups.

2.3 MODELING REGION

As stated previously, Phase I developed models were not implemented directly to this Phase II Study Area mainly because of the changes in geographic coverage and considerations of local data elements (e.g., sidewalk, slope, transit accessibility, etc.). Specifically, the Phase II Study Area is completely outside the scope of Phase I geography and many local-level data elements were not possible or feasible to collect under a national-level study such as the Phase I NMT study. In order to have a model formulation that could better fit local planning needs and consequently be built into the prototype sketch planning tool, an overhaul of the Phase I model was necessary.

Given that the Study Area contains only 35 block groups (the unit of measure for this study), only 35 "records" were available for developing models, less if there were any missing data involved. To allow sufficient degrees of freedom in the resulting models, a larger region which includes 722 block groups in Washington D.C., and portions of Montgomery County, MD, and Prince-George's County, MD (see green-shaded region in Figure 2-3) was defined to be the

"Modeling Region" for this project. For reference, the "Study Area" of the 35 block groups (shaded in purple) is overlaid on the "Modeling Region" in this map.

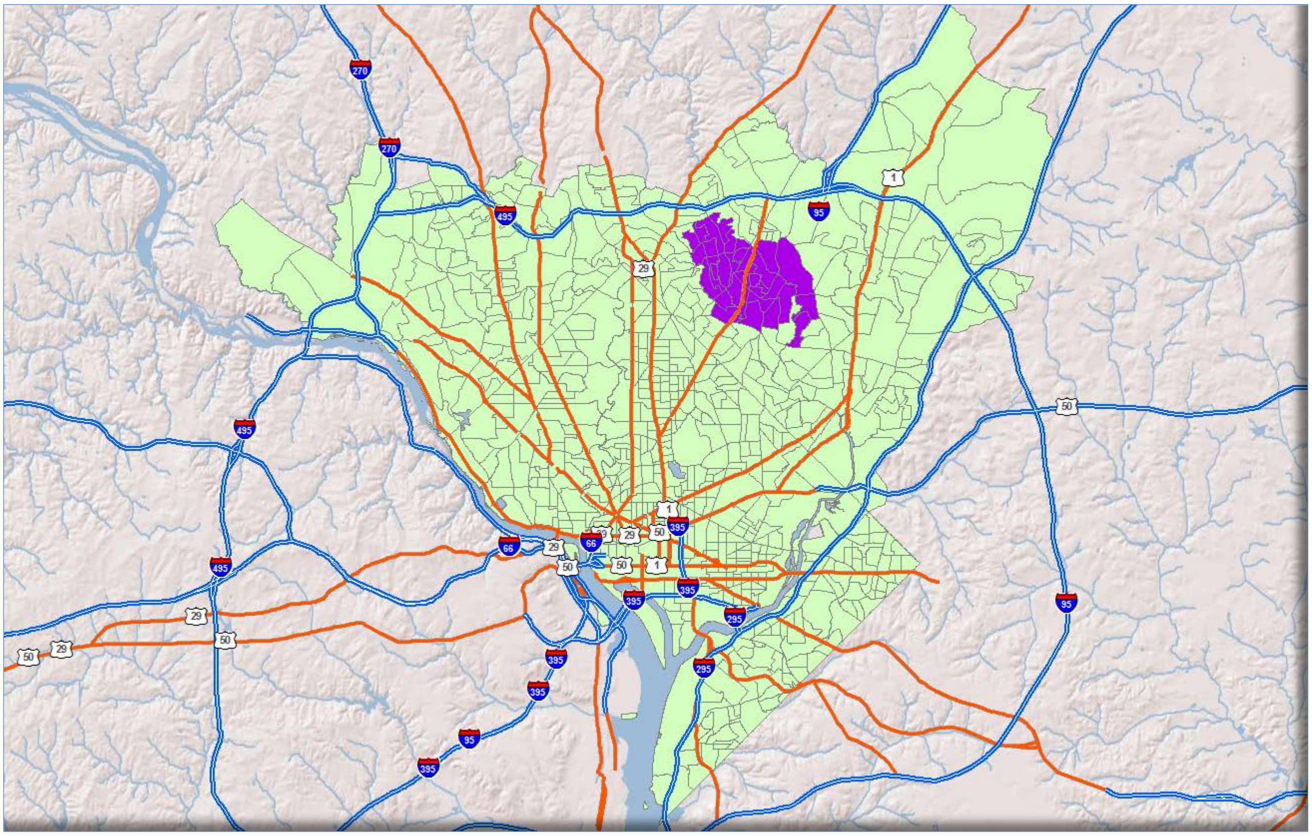


Figure 2-3. The Modeling Region with 722 Census Block Groups.

3. DATA SOURCES AND MEASURES

Major travel survey data used for this NMT Phase II study are the MWCOG 2007-2008 Transportation Planning Board Household Travel Survey (TRB HTS) data and the MWCOG 2011 Geo-Focused Area Survey data. In addition to these travel data, the EPA SLD [2], the 2006-2010 ACS data from the U.S. Census Bureau, and information collected from other data sources, as well as geospatial data variables generated by the project team, were used in this study. More detailed discussions on the data sources, and their associated data elements selected for the models, are presented in this section.

3.1 2007-2008 TPB REGIONAL HOUSEHOLD TRAVEL SURVEY DATA

In fulfilling its role as the Metropolitan Planning Organization (MPO) for the Washington D.C. region, the MWCOG periodically conducts regional household travel surveys to monitor changes in daily travel and to gather information on the demographic, socioeconomic, and trip making characteristics of metropolitan Washington D.C. area residents. From February 2007 through April 2008, a TPB HTS was conducted to collect information on the demographic and travel behavior characteristics of persons living in households in the Washington D.C. metro area and adjoining jurisdictions. This TPB survey, conducted by NuStats, captured a total of 11,000 sample households residing in this region.

Figures 3-1 and 3-2 display trip origin and destination locations (using block centroids), respectively, based on data collected under the TPB HTS. These two maps closely resemble each other in pattern, mainly because most daily trips made by a household typically would have reversed return trips made at a later time.

Under this study, the project team signed a data access agreement with MWCOG and obtained the 2007-2008 travel survey data files. Trip data provided to the team were linked-trips, which are different from trip-segment based data, as in the NHTS dataset. Geographic boundaries used in assigning Census regions were based on the 2000 Census definition blocks (2000 TIGER (Topologically Integrated Geographic Encoding & Referencing)). This data was aggregated to the 2000 Census block group level (block group being the measuring unit for this NMT project) by the research team. Several data elements were then derived from this MWCOG dataset, most are measures that associated with travel characteristics (e.g., mode of transportation), household demographic information (e.g., driving status, vehicle ownership, etc.). A list of data elements examined during the modeling effort is provided in Appendix A at the end of this report.

Based on the TPB HTS data, about 65% of walk trips (with “walk” as the main mode of transportation) traveled within 0.5 mile distance; another 21% traveled between 0.5 and 1 mile. This confirmed the general consensus that walking is typically conducted as the main mode of transportation for trips shorter than 1 mile in distance. This distance, thus, is implemented under the “Transit Accessibility” function in the prototype NMT application discussed in Section 5 of this report, as the upper bound distance for access to egress from transit stations.

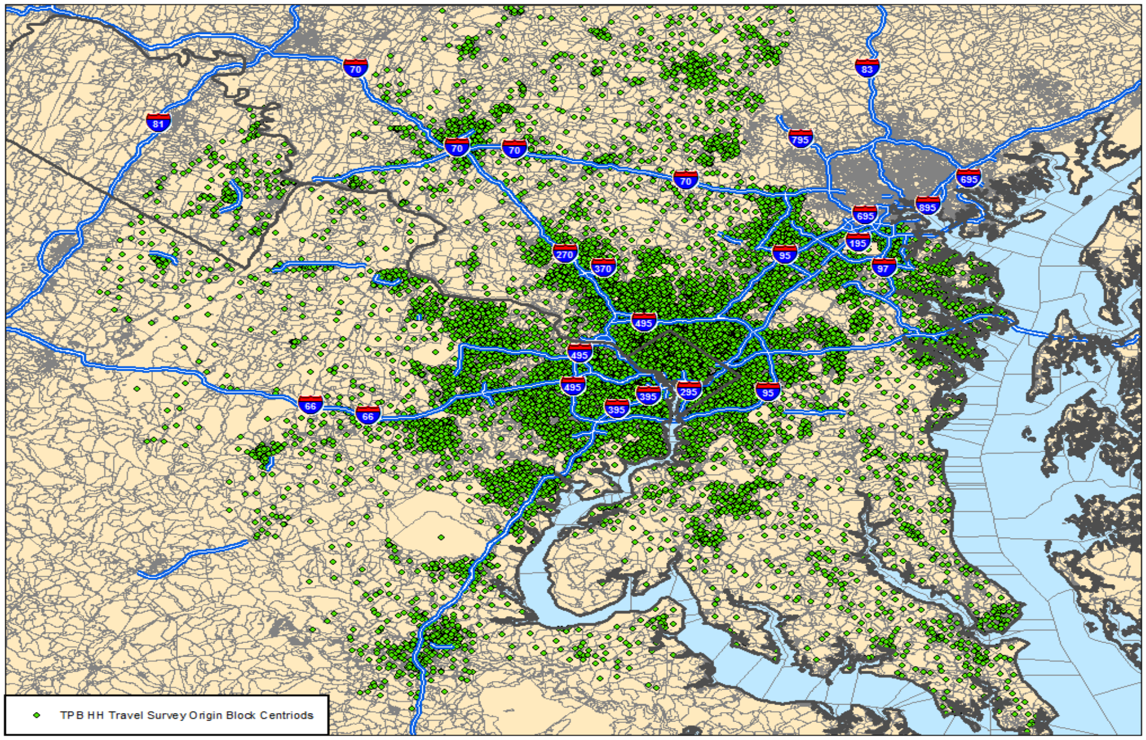


Figure 3-1. TPB Household Travel Survey Trip Origins, Using 2000 Census Block Centroids.

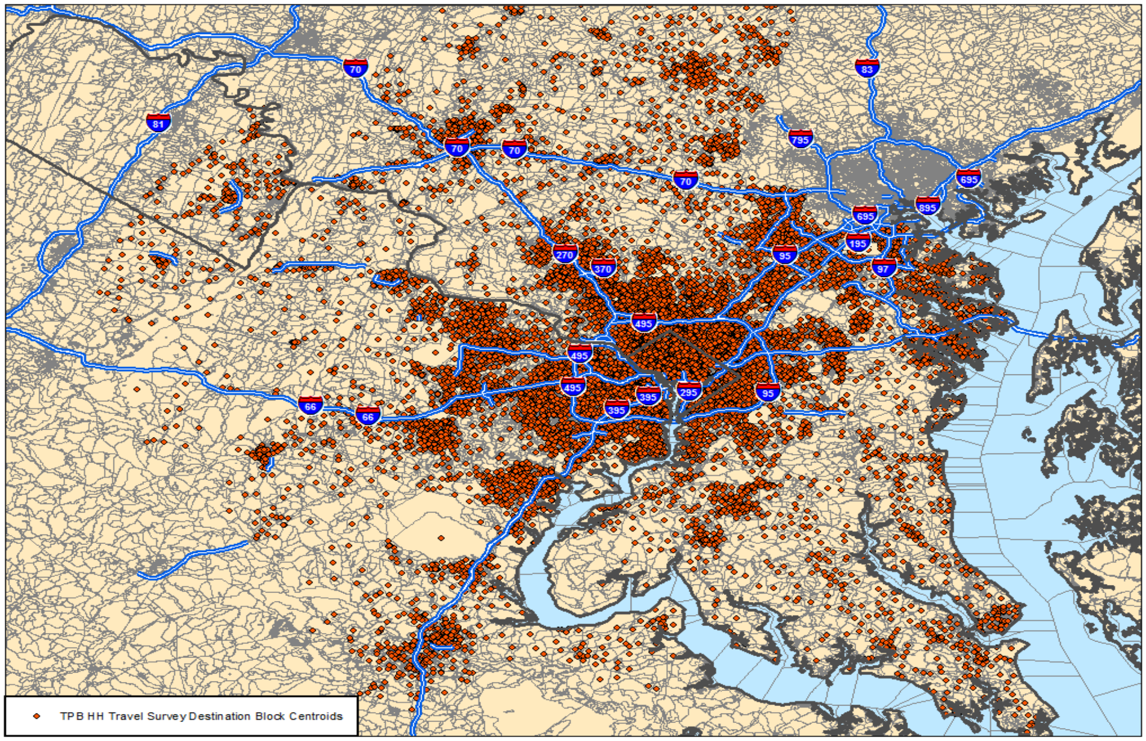


Figure 3-2. TPB Household Travel Survey Trips by Destination 2000 Census Block Centroids.

3.2 2011 MWCOG GEOGRAPHICALLY-FOCUSED AREA SURVEY DATA

Although 2007-2008 TPB HTS data provided sufficient samples to support models for large counties and cities within the MWCOG region, it did not have enough household samples to permit analysis of daily travel behavior in smaller areas within its jurisdiction. To enable specific smaller geographic areas to examine their travel patterns, a HTS in Focused Geographic Subareas of the region was added to the TPB work program in 2011. The first wave of the follow-on geographically-focused area survey was conducted in the fall of 2011, between mid-September and mid-November. Under this wave, travel information from 2,200 households in seven geographically-focused subareas of the MWCOG region was collected. These seven subareas include:

- The Logan Circle/14th St North West in the District of Columbia (from Massachusetts Avenue North West to north of Florida Avenue North West),
- The White Flint area in Montgomery County, MD,
- The Purple Line International Corridor in Montgomery and Prince George's Counties, MD (University Boulevard from south of I-495 to Adelphi Road),
- The Largo area in Prince George's County, MD,
- The City of Frederick, MD,
- The Reston area in Fairfax County, Virginia (VA), and
- The Woodbridge area in Prince William County, VA.

For the Phase II NMT study, trip data from the Fall-2011 focused area survey was requested and obtained from the MWCOG. The data files received contained similar variables and followed a similar format as those in the 2007-2008 TPB HTS dataset. However, geographic boundaries (e.g., block) provided in this dataset were coded based on the 2010 Census boundary definition. This data was aggregated to the 2010 Census block group level by the research team. Data elements considered for the Phase II NMT study were the same as those listed under Appendix A.

3.3 SMART LOCATION DATABASE

The SLD is produced by the EPA's Office of Sustainable Communities and contains a set of variables that are related to factors known as the "five Ds" [3, 4]which include: density, diversity (i.e., land-use mix), design, destination, and distance (to transit, in particularly). These "build environment" factors are known to have a significant effect on transportation and travel behaviors. The SLD is a nationwide geographic data resource that contains more than 90 attributes summarizing characteristics such as demographics, neighborhood design, housing density, land-use diversity, transit accessibility, etc. Most of these attributes are available at the Census block group level which is the geographic unit used in this NMT project. The EPA utilized data from Census as well as other data sources (e.g., Navteq) to create the SLD. Specific data elements obtained or derived from SLD, and are considered for this project, are provided in Appendix A of this report. Many of these data elements were used as indirect measures to reflect employment opportunities available to households in the block groups, substituting the role played by the Nielsen Employment Data used in the Phase I study.

Examples of this SLD data include "working-age population within travel distance of 30 miles," "cumulative number of jobs within travel distance of 30 miles," etc.

3.4 AMERICAN COMMUNITY SURVEY (ACS)

The ACS [5] is a continuous survey conducted by the U.S. Census Bureau of about 3 million households each year (250,000 per month), subject to the constraint that households should not be surveyed more than one time in any five-year period. Thus, the ACS is very intensive and is repeated every year. Each year's ACS sample includes, on average, almost 15 households per Census block group. Data on demographic, social, and economic characteristics is collected in the ACS. The ACS also collects data on commuting, i.e., Journey to Work (JTW), including mode of transportation and travel time to work. The 5-year ACS for 2006-2010 was used for this Phase II NMT project.

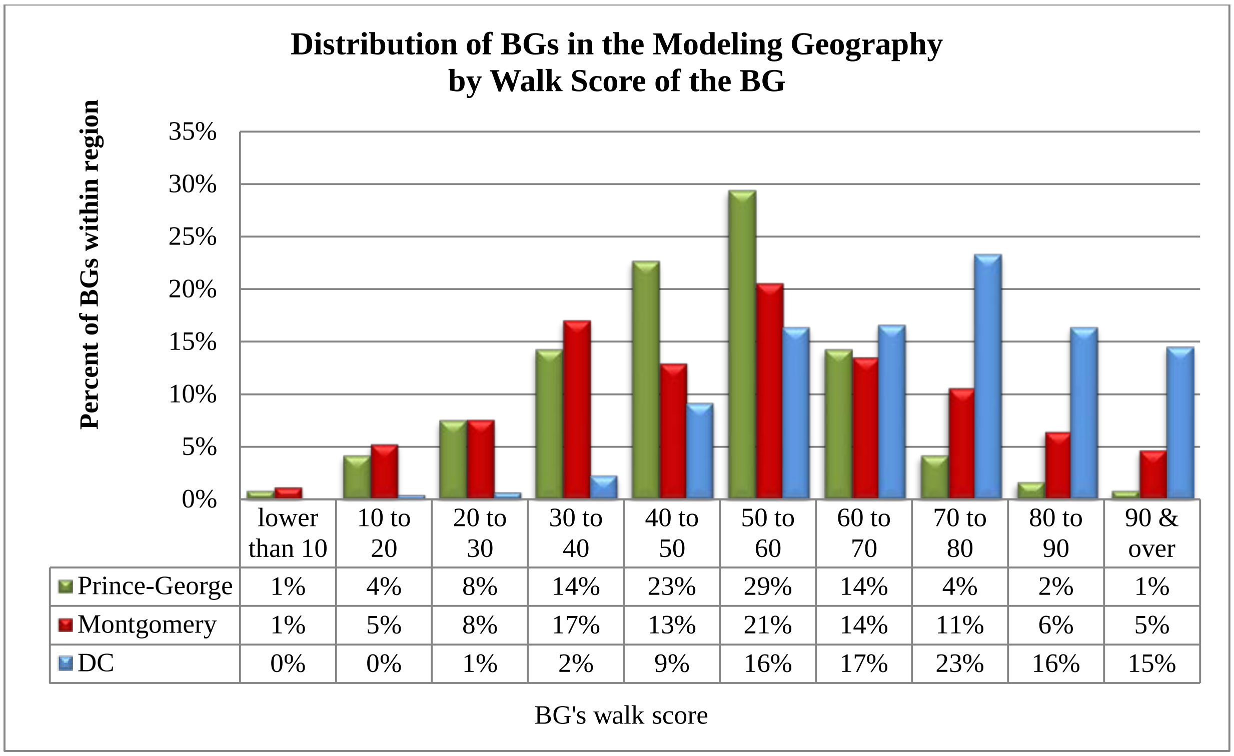
Note that, in order to increase transferability of the resulting models and the associated prototype tool developed under this project, the team opted to utilize publically available data as much as possible. For example, if a similar measure (e.g., "Percent White") is available through the public ACS data and also from a private data source (e.g., MWCOG), the ACS data element would be selected for the model formulations under this study. The rationale of taking this approach was to permit more regions to benefit from the study results and adapt the tool developed from this project without requiring extensive data collection efforts.
A large number of demographic and JTW variables were considered during the modeling development effort of this study. The most significant ones included:

- Percent of White Persons,
- Percent of Households with Incomes Above $99,000,
- Percent of Owner-Occupied Households,
- Percent of Persons with a Graduate Degree,
- Percent of Workers that Go to Work 7 – 9 AM,
- Percent of Spanish Speaking Households,
- Average Household Size,
- Average Worker Count of Households,
- Percent of Workers,
- Total Person Density, and
- Average Vehicles per Household.

3.5 WALK SCORE

Most of the walk scores [6] (from www.walkscore.com) for block groups within the Washington D.C. area were obtained during Phase I of this NMT study. Only a small number of block groups within the Modeling Region were not covered under Phase I and required manual processing to obtain their walk scores via the WalkScore.com web site. Ultimately, these walk scores were calculated at the block group centroid, based on distance to a diverse set of nearby amenities, e.g., grocery stores, restaurants, parks, schools, etc., and certain categories are weighted more heavily than others to reflect destinations associated with more walking trips. For instance, amenities within .25 miles receive maximum points and no points are awarded for amenities further than one mile. The “scoring” of a single location at the block group centroid might not necessarily have captured general neighborhood walkability as accurately in every block group, however.

Figure 3-3 shows the distribution of Walk Scores for the 722 block groups included in the Modeling Region of this project (see Section 2). Clearly, most Washington D.C. block groups (87%) have scores of 50 or higher reflecting the region’s high walkability. On the other hand, only about half of the block groups in each portion of Prince George’s and Montgomery Counties have the same level of walkability.



	lower than 10	10 to 20	20 to 30	30 to 40	40 to 50	50 to 60	60 to 70	70 to 80	80 to 90	90 & over
Prince-George	1%	4%	8%	14%	23%	29%	14%	4%	2%	1%
Montgomery	1%	5%	8%	17%	13%	21%	14%	11%	6%	5%
DC	0%	0%	1%	2%	9%	16%	17%	23%	16%	15%

Figure 3-3. Distribution of Walk Scores (at Block Group Centroids) in the Modeling Region.

Note that Walk Score defines several walkability categories:

0 – 24	very car-dependent,
25 – 49	car-dependent,
50 – 69	somewhat walkable,
70 – 89	very walkable, and
90 & over	walker’s paradise.

Although WalkScore.com also provides similar measures for transit accessibility (Transit Score), the research team did not use this information due to concern of its data quality, especially outside of Washington D.C. Walk Score and Transit Score are discussed in more detail in the Phase I Final Report of this project [7].

3.6 GEOSPATIAL DATA

Walkable places are streets and districts with physical attributes that encourage walking/biking for functional and recreational purposes. Based on this, knowledge on land-use characteristics of a neighborhood's physical environment is also an important component needed for studying factors that impact NMT activities. The geospatial nature of a neighborhood (or community) is best addressed by the use of GIS software. Factors such as roadway density, intersection density, and population density are commonly used as attributes in studying neighborhood land-use mix. Major geospatial data sources and variables considered under this study are discussed below.

3.6.1 Building Occupancy Square Footage Data for Land-use Proxy Measure

It is known that activity-friendly neighborhoods promote healthy living for their residents. Mixed land use in a neighborhood is important in this regard, as it reflects the availability of destinations to which residents can walk. Unfortunately, there is no direct measure of land-use mix available for the region covered by this study.

As in the Phase I study, the square footage occupancy database available from the Federal Emergency Management Agency (FEMA) was used to generate the land-use proxy measure for this project. The data is available at the Census block level and collected for use in emergency management purposes [8]. FEMA categorizes the data into seven classes: residential, commercial, industrial, agricultural, religious, government, and education. With the exception of religious and agricultural, further sub-classifications of land use are also provided. For example, the residential class can be further categorized into number of units per residence, temporary lodging, nursing home, etc. For this study, block level data was processed and aggregated to block group for areas identified under this Phase II study. Several ratios based on various combinations of the major building occupancy categories were analyzed. The "commercial" category was further separated into "retail" and "non-retail commercial" groups: where "retail" consists of FEMA's subcategories of retail store, restaurant, and theater; and the "non-retail commercial" consists of the remaining subcategories within the original "commercial" category.

This project utilizes the concept of entropy to combine information, in this case, using the square footage data on major building occupancy types to generate a quantitative measure of the land-use mix. Instead of using the Shannon diversity index [9] (Shannon entropy) as in Phase I NMT study for land-use mix measure, a slightly different entropy formulation as published in the 2004 *American Journal of Preventive Medicine* by Frank, et al., [10] was applied in this Phase II study. Specifically, the land-use mix (LUM) measure is expressed as:

$$LUM = -\sum_{i=1}^{n} p_i \ln p_i \,/\, \ln n$$

Where p_i is the proportion of estimated square-footage attributed to a given building occupancy category i, and n is the number of "land use" categories. The resulting LUM has a value that ranges from zero to one. When there is only one type (i.e., category) of building use in the area (e.g., purely residential neighborhood), this LUM value equals zero, while LUM of 1 represents a perfectly even distribution of square-footage across all land use categories. The hypothesis here was that more diverse land-use mix would encourage more daily activities being conducted by walking in a community.

Two slightly varied measures of land-use mix entropy were estimated under this study and applied in the models. The two variables are:

- ENTROPY: Considered square-footage for building occupancy types of Residential, Retail, Non-Retail Commercial, Industrial, Agricultural, Religious, Government, and Educational.
- ENTROPY_R: Considered all square-footage for building occupancy types as specified above except Residential.

Figure 3-4 shows the estimated block group land-use mix measures using variable “ENTROPY.”

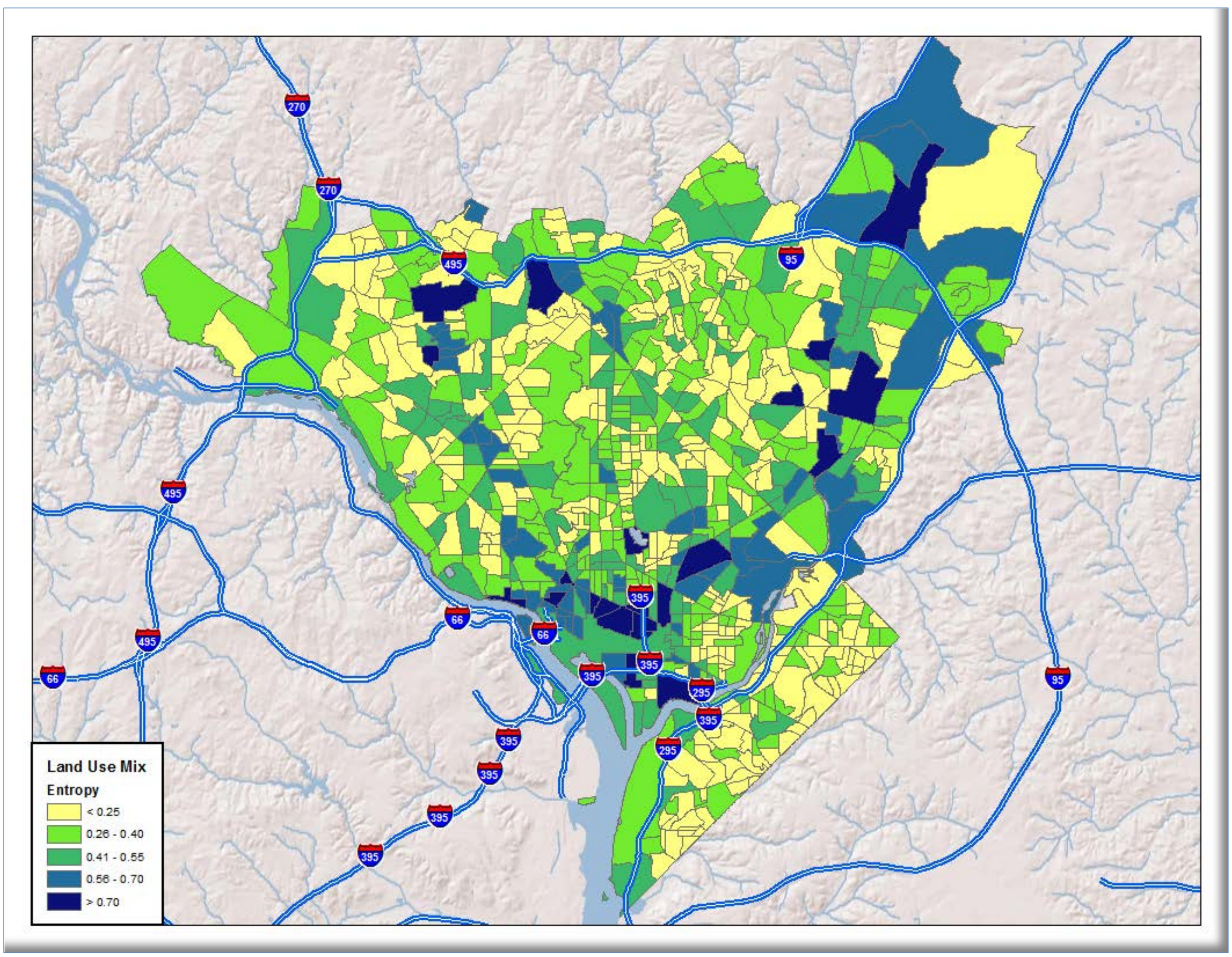


Figure 3-4. Measure of Land-Use Mix (Variable Entropy) for block Groups in the Modeling Region.

3.6.2 Sidewalk Inventory

Availability of sidewalks within a community is commonly considered an important factor for promoting walking. Sidewalks (particularly paved sidewalks) provide people with space to travel separated from motor vehicles and on-road bicycles, thus could have a less likelihood of being a crash [11]. The Safe Routes to School program specifically emphasizes that roads around schools not equipped with sidewalks can be unsafe for walking. Aside from the safety point of view, according to a Bureau of Transportation Statistics (BTS) study report release in December 2004, the *BTS Issue Brief,* "the presence of sidewalks has a slight positive effect on the tendency for adults to take walks," although BTS also stated that the presence of sidewalks "has no effect on the frequency of those walks." [12]

The lack of effect on the frequency of walking might be due to decision of walking are made with considerations of many other characteristics such as sidewalk surface, width, street lighting, neighborhood security, etc. However, data on sidewalk networks and associated detailed attributes are generally difficult to find. Under this study, sidewalk inventory data was provided from three jurisdictions which intersect the Modeling Region (described in Section 2). Note that no link-level attributes were available from any of these sidewalk inventory data files.

The level of detail, and data quality, of these three sidewalk inventory data files are quite different. Specifically,

- Washington D.C.: obtained the ***Polygon Shape-files*** from the FHWA [13]. The data appears to be very accurate and high quality when examined with Satellite imagery of the region. It includes fine detail and provides a very accurate measure of sidewalks (see the example displayed in Figure 3-5).

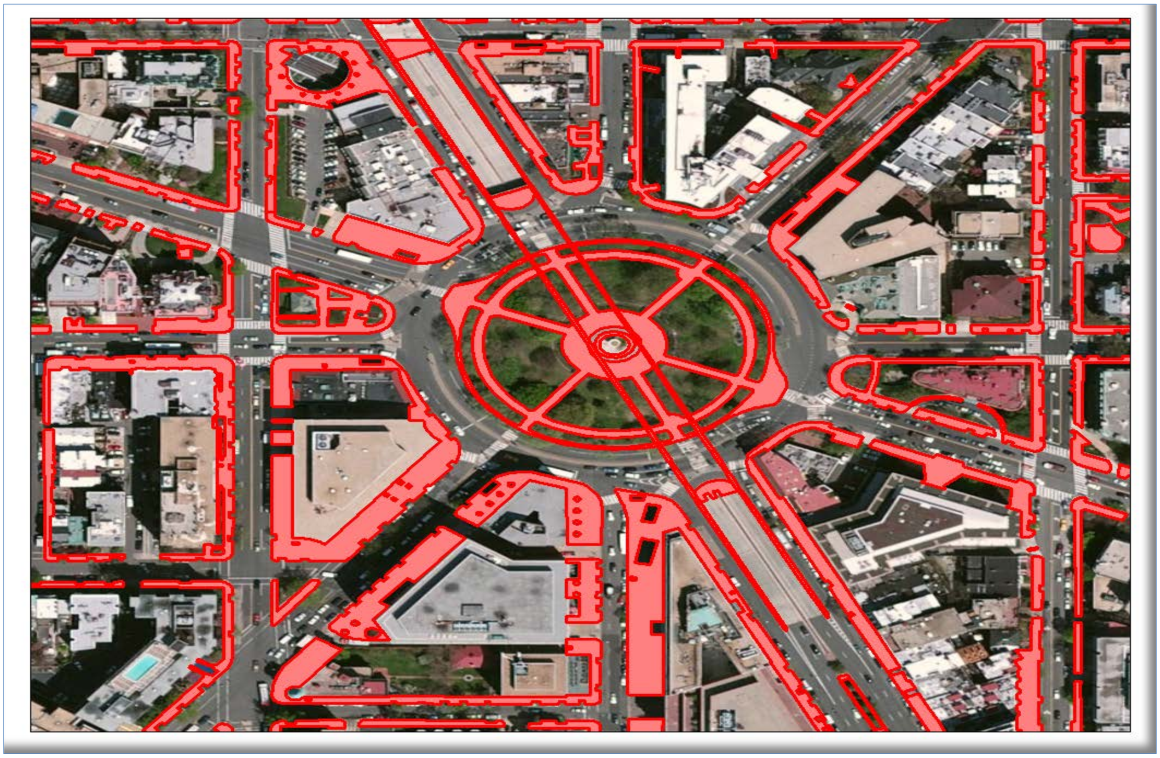


Figure 3-5. Washington D.C. Polygon Sidewalk Inventory Data in DuPont Circle Area.

- Prince-George's County MD: provided ***Polygon Shape-files*** which covers about 90% of all sidewalks in the county [14]. It also includes walkups to houses (from sidewalk to front door), which could inflate the measure of sidewalk within a block group. An example is shown in Figure 3-6 for this dataset.

- Montgomery County MD: provided ***Poly-line Shape-files*** which covers only sidewalk segments that are over 200 ft. in length [15]. Therefore reducing the quality of the dataset, specifically in more downtown areas (e.g., Silver Springs, MD example is shown in Figure 3-7) where streets and sidewalk segments are denser and shorter in length (also where sidewalks are more prevalent to be). Also due to the fact that it is a line file, an assumption on the width is needed in order to create a measure of sidewalk within a block group, i.e., a measure consistent to other sets.

Figure 3-6. Prince George's County MD Polygon Sidewalk Inventory Data in College Park Area.

Figure 3-7. Montgomery County MD Polygon Sidewalk Inventory Data in Silver Springs Area.

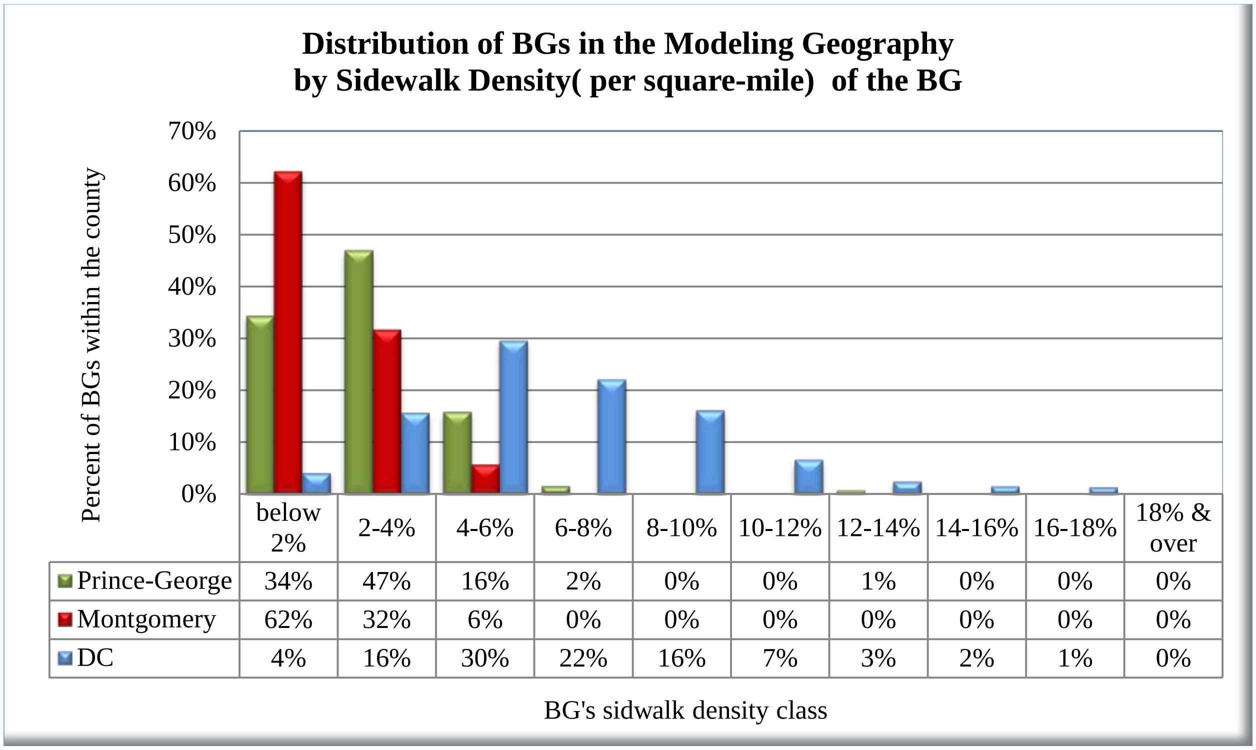
The above described differences among the three sidewalk inventory datasets posed a great deal of complication in fusing these sets into one consistent measure for this study. A “Sidewalk Density” measure (i.e., variable) was created using the square mileage of sidewalks within a block group divided by the total square mileage of the given block group. The process of calculating this measure varied depending on the data sources.

For Washington D.C., sidewalk density was created by calculating the square mileage of the provided sidewalk polygon dataset, after breaking it at the block group boundaries, and dividing by the block group’s square mileage to obtain the density.

Prince-George’s County sidewalk density was created by calculating the square mileage of the provided sidewalk polygon dataset, after breaking it at the block group boundaries, and then deducting 20% from it to account for the large amount of “walkups” (20% reduction was determined after a significant amount of sampling and examining effort from various block groups within the County). This adjusted value was then divided by its corresponding block group’s square mileage to create the density.

Montgomery County sidewalk density was created by calculating the linear mileage of the segments in the provided line dataset, after breaking it at the block group boundaries, and multiplied by .00136 miles (or 7.2 feet). The estimated 7.2 feet was estimated by taking the sampled 6-feet average width of sidewalks in Montgomery County from aerial imagery and adding 20% to it to account for the missing sidewalks in Montgomery County. (Recall Montgomery County data only included sidewalks over 200 feet in length). This adjusted value was then divided by its corresponding block group’s square mileage to create the density.

Using per square-mile density as a measure, Figure 3-8 shows that Washington D.C. has a significantly larger percent of block groups with higher sidewalk densities than in the other two regions. As an alternative to the Sidewalk Density variable, another variable measuring the Ratio of Sidewalk Square Mileage to the total linear Road Miles for each block group was also calculated and considered in the modeling process.



	below 2%	2-4%	4-6%	6-8%	8-10%	10-12%	12-14%	14-16%	16-18%	18% & over
Prince-George	34%	47%	16%	2%	0%	0%	1%	0%	0%	0%
Montgomery	62%	32%	6%	0%	0%	0%	0%	0%	0%	0%
DC	4%	16%	30%	22%	16%	7%	3%	2%	1%	0%

Figure 3-8. Distribution of Sidewalk Density by County.

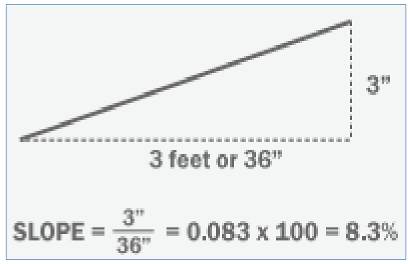
3.6.3 Elevation Data for Slope Measure

Because the slope of a roadway, or a sidewalk, also influences walkability and no such attribute information was provided from the sidewalk dataset described above, an alternate method was used to estimate slope within the Study Area. Using the National Map Viewer [16] created and hosted by the United States Geological Survey, a digital elevation model of the Washington D.C. metro area and the neighboring region was downloaded. The downloaded data contained raw raster data at 1/3 arc-second resolution (i.e., 3 meter by 3 meter, or 3-m x 3-m).

The study team processed the raw elevation data into a dataset containing slope information. This slope-estimation process involved several steps. First, to adhere strictly to the geographic region selected for this project (see Section 2), all fragments that fell outside of the specific area geographies were discarded. Generally speaking the slope-estimation tool processes the data cell by cell (each 3 meter by 3 meter), beginning with marking a target cell and then evaluating the maximum changes between the surrounding cell and the target cell using an algorithm. The specific processes used within the "Slope Tool" are described in more details under Environmental Systems Research Institute's (ESRI) ArcGIS Resources website [17].

The output of the "Slope Tool" is a raster file containing the slope values for each cell. These slope values can be reported in either degrees of angle or percent of rise. The percent of rise measure of slope is typically used to facilitate a better understanding of the steepness of the slopes involved. Figure 3-9 shows general diagrams on how percent rise measure of the slope and degrees angle measure of slope are calculated.

Percent rise measure of the slope



Degree angle measure of the slope

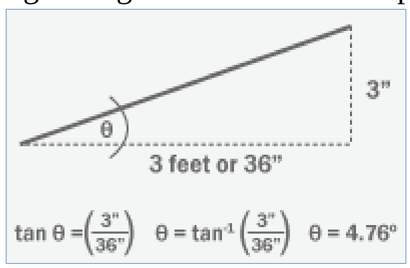


Figure 3-9. Measures of the Slope.

Once slope measurements for the selected Modeling Region were calculated, the raster file was converted into a vector point file for use in this project. According to the Americans with Disabilities Act (ADA), a walking space with a slope greater than 5% (or 2.86 degrees) is considered a ramp [18]. Using this information as a baseline, the slope data was divided into three categories:

Category One: 0% to 5% (2.86 degrees),
Category Two: 5% to 10% (5.72 degrees), and
Category Three: greater than 10%.

A count of slope points was created where each point represents one 3-m x 3-m cell that falls within each of the Census block groups in the project region. Three slope-related variables were created based on the percent-rise slope measures and used in the modeling effort described in Section 4 of this report. These variables provide estimates for the percentage of areas within a block group (based on counts of 3-m x 3-m cells) that are within: <u>easily walkable slope</u> (i.e., 5% or less), <u>moderately walkable slope</u> (5.1 – 10%), and <u>difficult walking slope</u> (over 10%). For example, a block group with a 77% value in the "easily walkable slope" variable means that 77% of its block group region has a slope of "5% or less."

Figures 3-10 and 3-11 together provide a supplementary view of "walkable" areas within the Modeling Region of this project. Block groups with a darker green (Figure 3-10) shading indicate that they have a higher share of flatter land (measured as 5% or less in slope) within their areas; while block groups with a darker red (Figure 3-11) shading represent those with a much lesser percent of walk-friendly land (i.e., more area in a block group with over 10% slope).

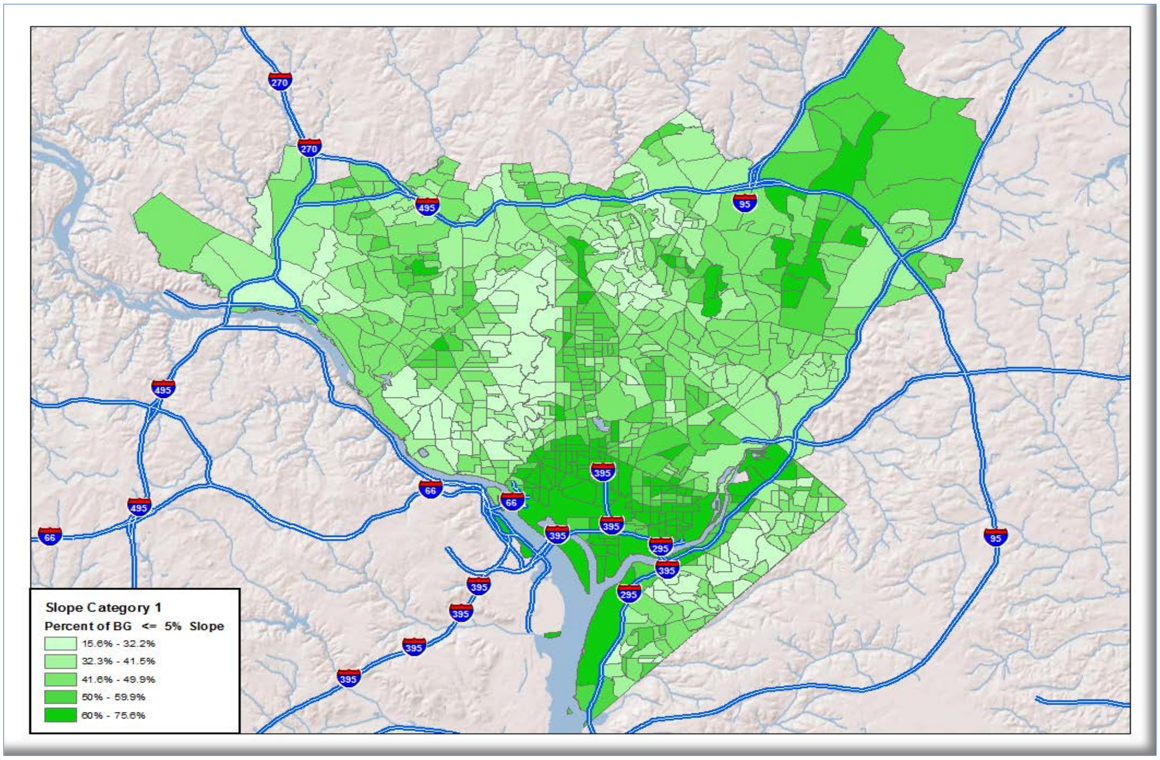


Figure 3-10. Percentage of a Block Group with Easily Walkable Slope (5% or Less).

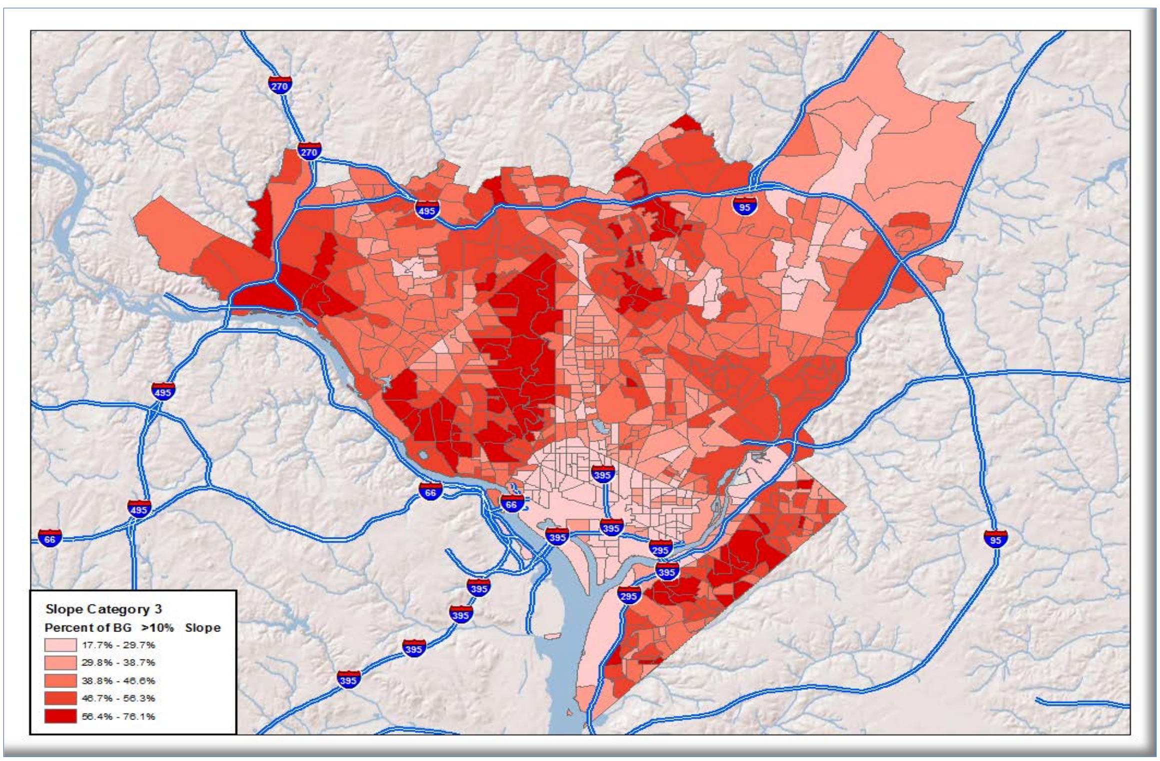
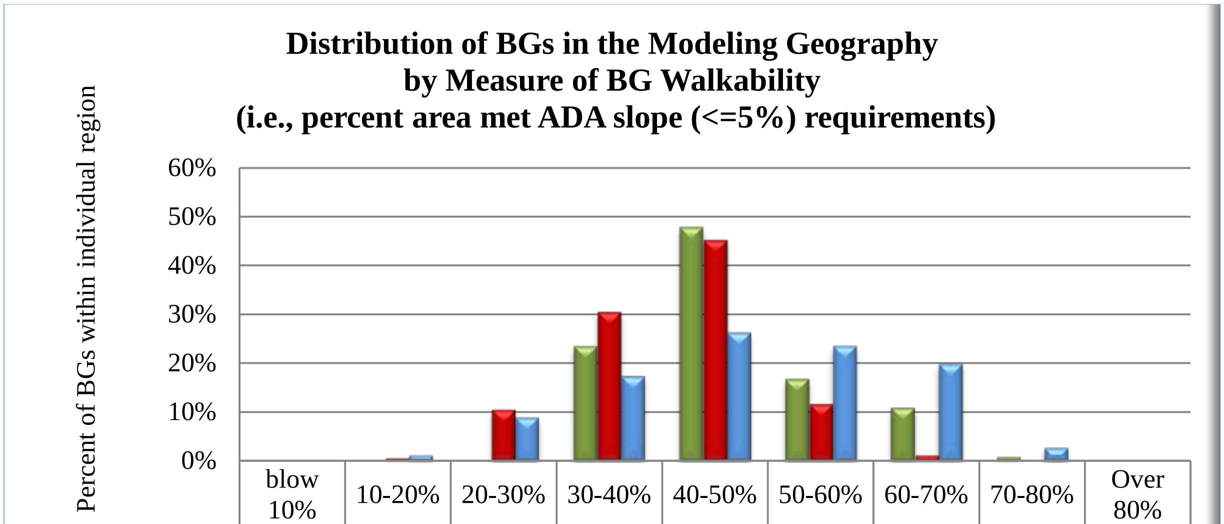


Figure 3-11. Percentage of a Block Group with Difficult Walking Slope (Over 10%).

Overall, distributions of block groups with a "walkable" slope (5% of less) in the three jurisdictions studied show similar patterns (Figure 3-12), with distribution for Washington D.C. shifted to the right slightly more indicating a higher number of "flatter" block groups in its region.



	blow 10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	Over 80%
Prince-George	0%	0%	0%	24%	48%	17%	11%	1%	0%
Montgomery	0%	1%	11%	31%	45%	12%	1%	0%	0%
DC	0%	1%	9%	17%	26%	24%	20%	3%	0%

Figure 3-12. Distribution of Block Groups with Walkable Slope (5% or Less).

3.6.4 Transit Information – General Transit Feed Specification and Beyond

Clearly, access to transit stations is one of the reasons for people making walk trips. The presence of a transit station, by itself however, does not necessarily reflect accessibility. The frequency of transit services at a given location, as well as routes being served from the given location, are all important factors in determining accessibilities for that neighborhood. Under this study, several transit-related variables were created in an attempt to measure transit accessibility. Data elements considered under the modeling processes include:

- total number of Metrorail transit stations within a block group;
- total number of all transit stations/stops (bus and rail) within a block group;
- total number of daily transit arrivals at stations/stops within a block group;
- average hourly transit arrival rate based on an assumed 18-hour operation per day within a block group; and
- average number of transit arrivals per station/stop within a block group.

Figure 3-13 displays the distribution of transit station counts for the Modeling Region, based on the total number of transit stations (or stops) within each block group. Block groups around downtown Washington D.C. appear to be in a much darker color, reflecting a higher concentration of transit stations/stops located within that region. The number of transit stations/stops within an area does not always truthfully reflect its transit availability, however. This is especially true for smaller size block groups, as in the areas north of downtown Washington D.C. (can be observed in Figure 3-13). Frequency of transit arrivals at stations, for example, might deliver a better measure for the level of service within certain areas. As presented in Figure 3-14, which displays the distribution of daily transit arrivals at stations within the Modeling Region block groups, the pattern is different from what was seen in Figure 3-13. For example, the areas north of the downtown region discussed above have a higher transit arrival rate (darker shade) than the far left block-group in the map (Figure 3-14); it was the other way around for the transit station count (Figure 3-13), however.

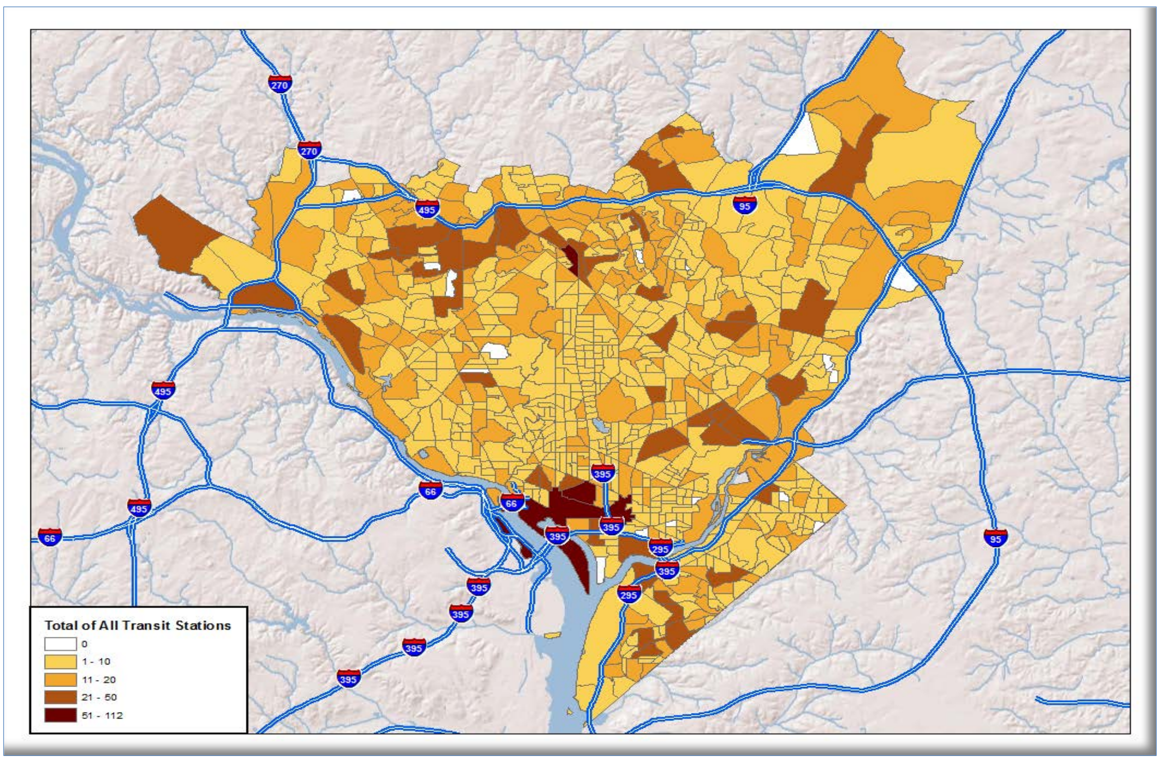


Figure 3-13. Number of Transit Stations/Stops within Modeling Region Block Groups.

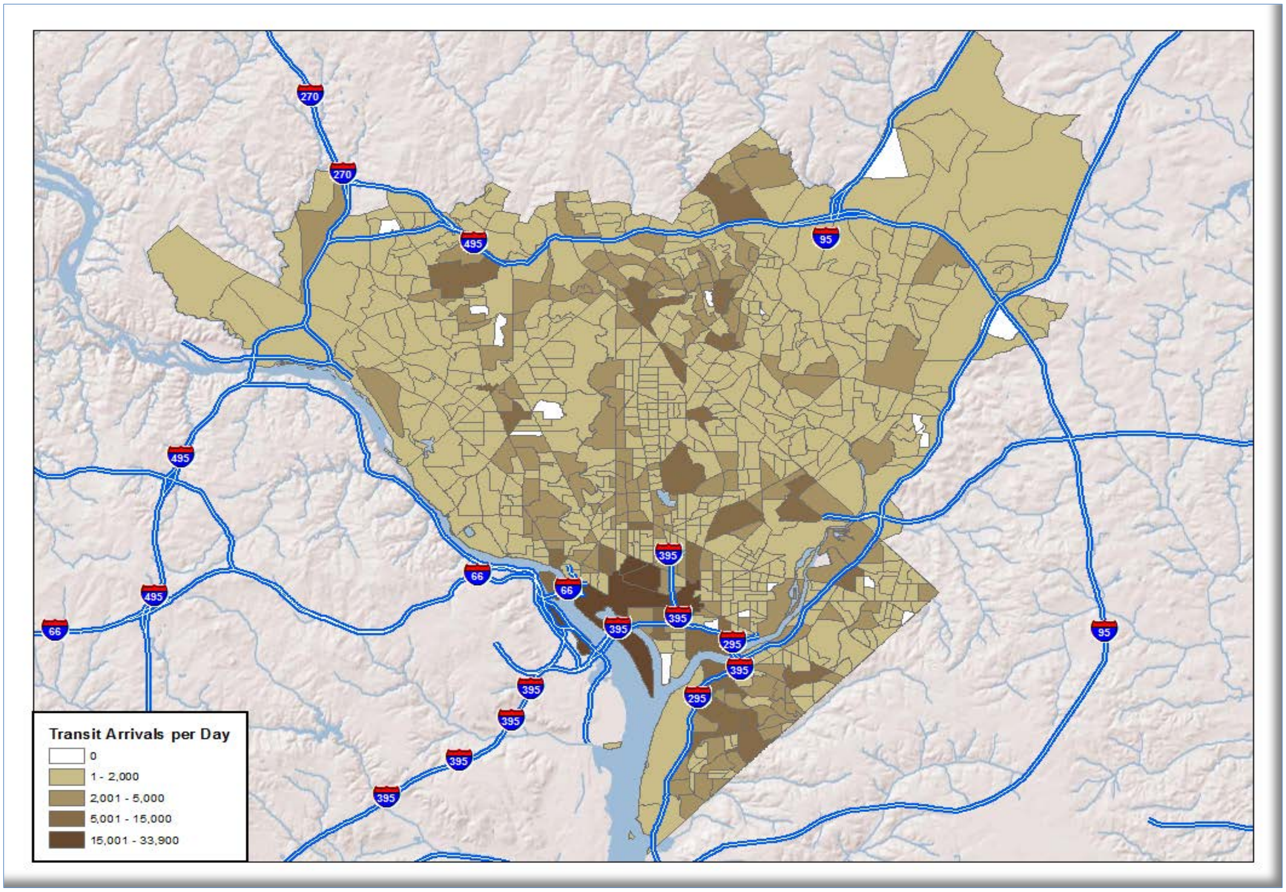


Figure 3-14. Number of Daily Transit Arrivals at Stations within Modeling Region Block Groups.

Figure 3-15 displays the average transit arrivals per hour assuming an 18-hour per day operating time period. This measure is basically the same, except for their scales (one uses the total count, the other uses the average). When considering the average transit arrivals per transit station as displayed in Figure 3-16, it is not surprising to see that most of the high arrival frequency stations/stops are concentrated in the Washington D.C. area block groups, including Metrorail stations such as Gallery Place/Chinatown and Metro Center.

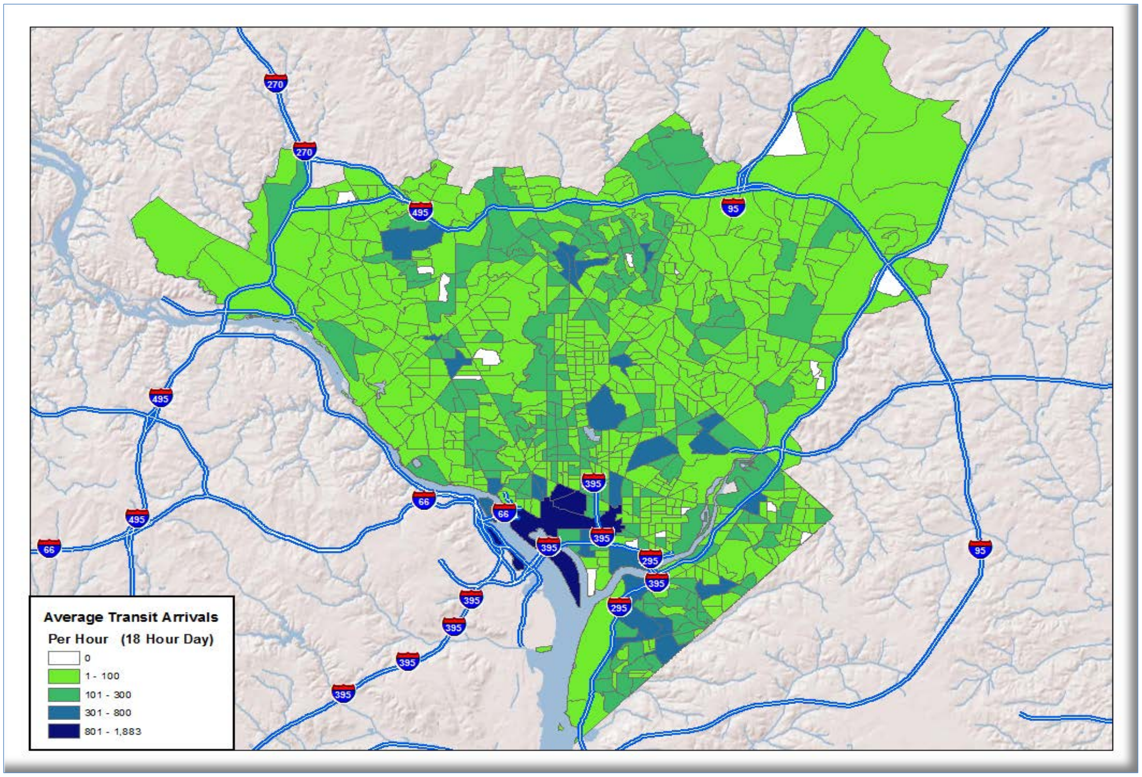


Figure 3-15. Average Transit Arrival Rate per Hour (Assumed 18-hour Operation Day).

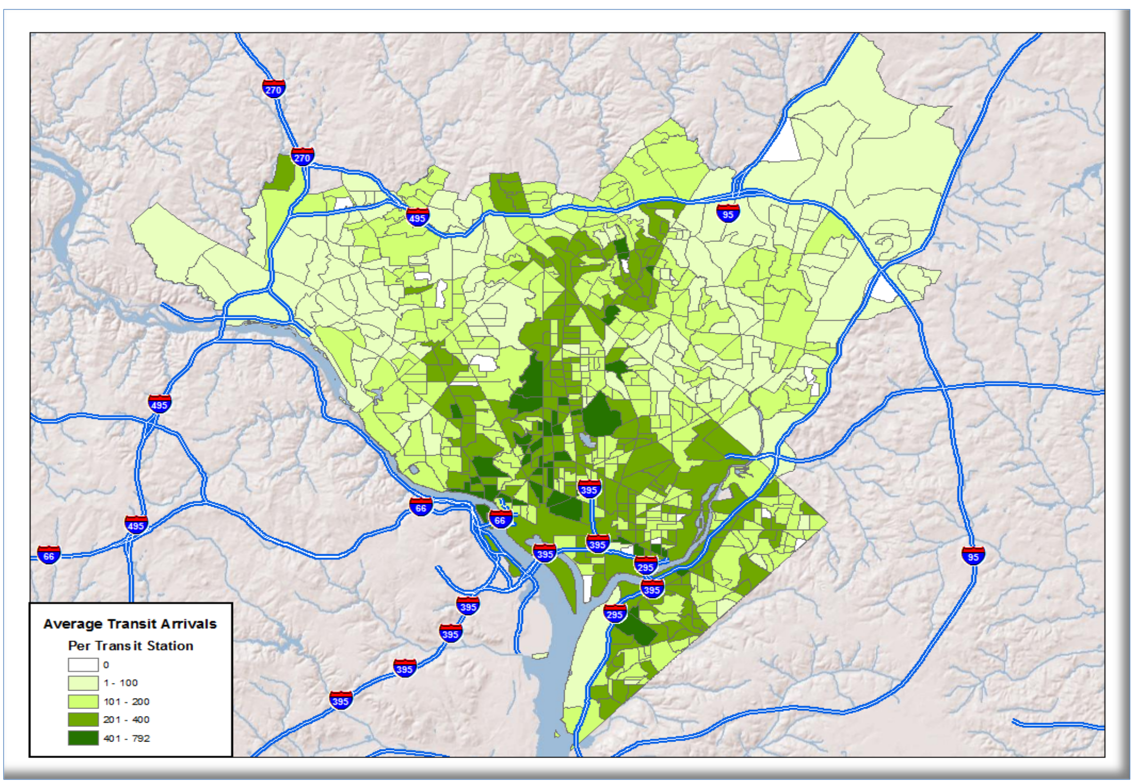


Figure 3-16. Average Transit Arrivals per Station.

The main source of information for transit routes and associated stations was based on data obtained from the General Transit Feed Specification (GTFS) public feeds for the area covered under the WMATA (both bus and rail) system. Additional transit data for Montgomery County's "Ride On" local bus system was also obtained from GTFS. Additional local bus transit data for Prince George's County's "The Bus" had to be manually created by the research team. The dataset was developed following the GTFS format for data consistency.

3.6.5 Other Geospatial Data

As in Phase I of the NMT project, Oak Ridge National Laboratory (ORNL) derived several geospatial data elements for density measures, such as roadway density, intersection density, daytime population density, and nighttime population density. These derived geospatial data elements were also considered in the model formulation process for this project. Definitions and calculation methods of these data elements are not repeated in this report, however. Readers interested in more details are referred to the Phase I report at http://info.ornl.gov/sites/publications/Files/Pub36798.pdf.

Although not used in the models, several new datasets were also added to enrich the prototype GIS-tool. These include:

- Transportation Infrastructure,
- Parks,
- Fatal accident locations from the 2011 Fatality Analysis Reporting System (FARS),
- Crime Data (Crime Reports.com), and
- Schools (K-12 and Colleges).

An example of the GIS data for parks in the Prince George's and Montgomery Counties is shown in Figure 3-17. Locations where fatal accidents occurred with pedestrian involvement status, around the Proposed Purple Line region in 2011 are displayed in Figure 3-18.

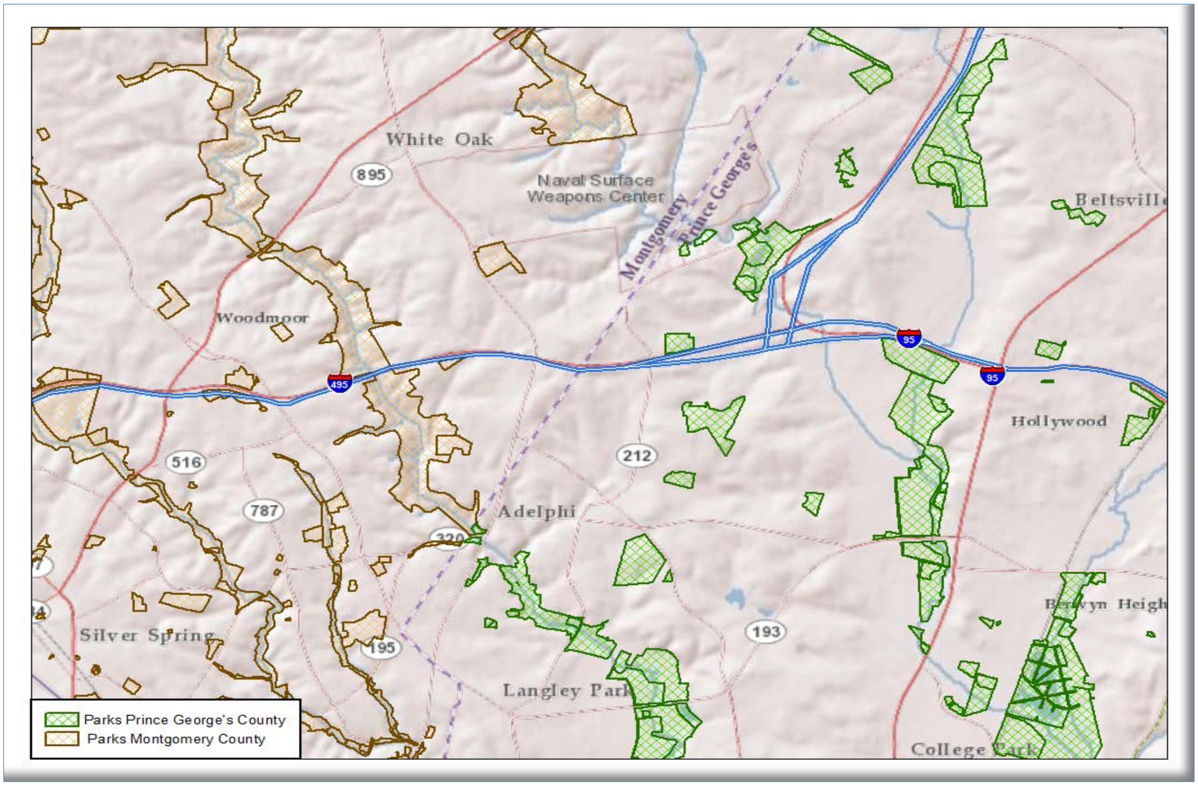


Figure 3-17. Parks in Prince George's and Montgomery Counties, MD.

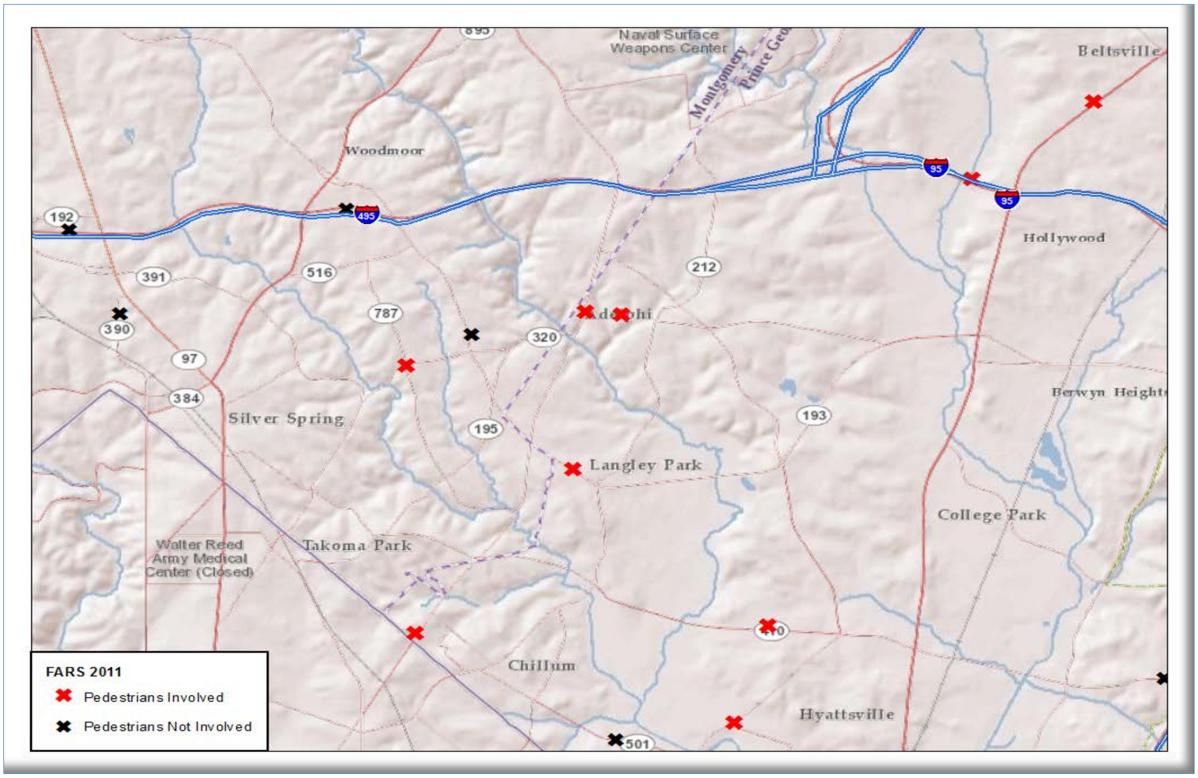


Figure 3-18. Location of Fatal Accidents with Pedestrian-Involvement Status in 2011 (FARS Data).

3.7 DATA LIMITATION AND CHALLENGES

Clearly, travel data obtained from the MWCOG regional travel survey has much better coverage for the geographic region selected for this project than what the national level NHTS data could have provided. Of the 722 block groups in the Modeling Region (explained in Section 2), a total of 124 block groups did not have any walk trips, including walking as the egress or access mode. Among those 124 block groups, 15 block groups had no sampled households in the 2007-2008 survey. This lack of walk trips in some block groups might have resulted because only one household in the given block group was sampled and no walking trips were made by that household on its given travel day. This is basically a sample survey data limitation.

3.7.1 Linked Trip vs. Unlinked Trip

Another likely reason for the lack of walking trip records could have resulted from the way trip data was coded. Note that the MWCOG data as provided to the research team was linked-trip records (from an origin to a destination for a primary purpose), not detailed trip-leg (one-way travel to any place by a single mode of transportation) records. The following hypothetical example can be used to illustrate the difference between unlinked trip legs and a single linked-trip record. Assuming a person (Person-1) of a given household travels from home to work, which involved:

1. Drive to Metrorail station at location X,
2. Take a Metrorail trip from X to Metrorail station at Y,
3. Walk from Metrorail station at Y to Bus stop at L1,
4. Take Bus ride from L1 to Bus stop at L2, and
5. Walk from Bus stop at L2 to work place.

Using this hypothetical travel example, Table 3-1 shows the 5 unlinked trip legs for this person's commute trip. The corresponding linked trip for this specific example is presented in Table 3-2. The linked trip record shows a commute trip from home (origin) to work (destination) with primary mode of "Metrorail", access mode of "Auto" and egress mode of "Bus." Note that, none of the "walk" segments made into the linked trip records.

The processes MWCOG used in converting unlinked trip legs into single-purpose linked trips utilized a hierarchy system to assign the primary mode and the secondary modes of access and egress for linked transit trips. This hierarchy of modes is listed in Table 3-3 with "Metrorail" on the top of the mode hierarchy and "Walk" being the lowest. Because of this hierarchical sequence, "walk" could be a primary mode only if the travel involved walking exclusively. Similarly, "walk" could be the access or egress mode only if no other mode of transportation was used as the secondary mode (e.g., walk to a station to access the primary mode, or walk from the ending point of the primary mode to a destination). To "capture" as much walking activities as possible from this dataset for the modeling used in this study, all walking activities including walking as the primary mode and those with access/egress mode of walk or any type of buses were considered in the counts as "trips involving walking."

Table 3-1. Example of the Unlinked Trip Records

Person Number	Trip Number	Trip Origin	Origin Purpose	Travel Mode	Trip Destination	Destination purpose
1	1	home in location A	Home	Auto driver	Metrorail station X	Change of mode
1	2	Metrorail station X	Change of mode	Metrorail	Metrorail station Y	Change of mode
1	3	Metrorail station Y	Change of mode	Walk	Bus stop @ L1	Change of mode
1	4	Bus stop @ L1	Change of mode	Local bus	Bus stop @ L2	Change of mode
1	5	Bus stop @ L2	Change of mode	Walk	Workplace @ location B	Work

Table 3-2. A Linked Trip Record Created from the Unlinked Trip Legs Shown in Table 3-1

Person Number	Trip Number	Trip Origin	Origin Purpose	Travel Mode	Trip Destination	Destination purpose
1	1	home in location A	Home	Transit /Metrorail	Workplace @ location B	Work

Person Number	Trip Number	Mode of Access	Mode of Egress
1	1	Auto Driver	Local Bus

Table 3-3. Trip Linking Hierarchy

Trip Linking Hierarchy Mode
1. Metrorail
2. Commuter Rail
3. Light Rail
4. Commuter Bus
5. Local Bus
6. School Bus
7. Metro Access/DAR
8. Shuttle Bus
9. Taxi
10. Auto Passenger
11. Auto Driver
12. Heavy Truck
13. Motorcycle
14. Bike
15. Other
16. Walk

3.7.2 Data Consistency among Jurisdictions

More often than not, data provided from different jurisdictions would have varying degrees of accuracy, coverage, or sometimes, even inconsistent definitions. This is always a challenge when one needs to fuse data from different sources together and make it as consistent as possible, so as to provide a smooth and complete picture of whatever the targeted subject is. The sidewalk inventory data mentioned under Section 3.6.2 is an example of this data issue/challenge. Most of the time, assumptions and adjustments are needed in order to allow datasets to be integrated as seamlessly as possible so comparisons could be made (as fairly as possible). These are only estimates; no ground-truth information is available so accuracy of the measure itself (e.g., total area of sidewalk) is difficult to judge. The focus of this study is to use available data for model formulation and integrate that formulation into a prototype visualization tool that allows local planners to perform sketch planning regarding walking in their community. It is assumed that local planners could have better data to substitute than what is currently used in this model.

3.7.3 Geographic Boundary Changes

Block group boundary changes between Census years 2000 and 2010 also posed challenges for the research team during the data processing and development of the sketch planning tool. Specifically, the 2007-2008 MWCOG data was coded using the 2000 Census geographic definition and the Fall-2011 MWCOG Geographically-focused area data was coded according to 2010 Census geographic definition. These differences involved: change of block group ID (no boundary changes), change of block group boundary (kept the same ID), or combinations of ID and boundaries in various fashions.

Of the thirty-five 2010 Census block groups in the Study Area, twenty-one remained intact when moving from their 2000 Census block group definitions. Some block groups split from one block in 2000 into two block groups in 2010. An example of this is Census 2000 block group "240317021011" that was split into two block groups "240317021011" and "240317021014" under the Census 2010 definition (See Figures 3-19 and 3-20). In other cases, some block groups maintained their boundary but changed their block group ID (See Figures 3-21 and 3-22).

Figure 3-19. Block Group "240317021011" Boundary with 2000 Census Definition.

Figure 3-20. 2010 Definition of Census Block Groups (240317021011 & 240317021014).

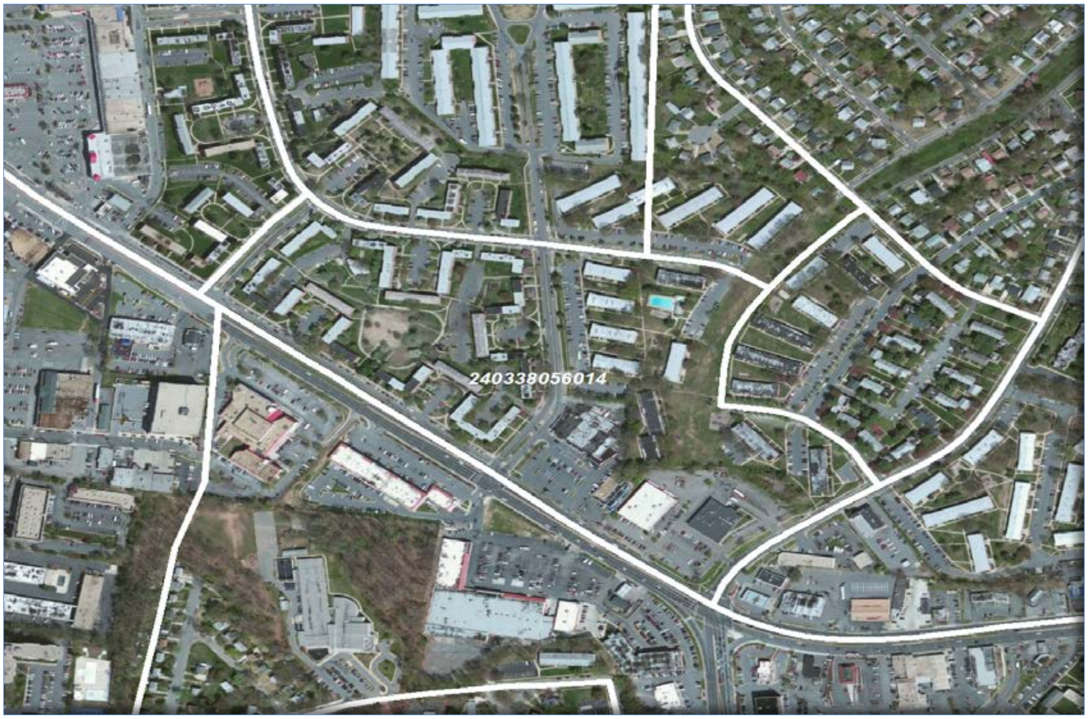


Figure 3-21. 2000 Census Block Group 240338056014.

Figure 3-22. Change of Block Group ID: 2010 Census Block Group 240338056012.

4. MODELING APPROACHES AND FACTORS

As stated previously, the Phase II Study Area is completely outside the scope of Phase I geography and Phase II also included many local-level data elements that were not used under the Phase I NMT study. Thus updated models were needed to estimate walkability in communities over the Phase II region. Obviously, major factors impacting a community's walking activities identified from Phase I (e.g., walk score, demographic characteristics, land-use mix) are still relevant to the Modeling Region of this project. No major rescanning of data sources as conducted in Phase I was performed for Phase II.

Furthermore, in addition to a model that produces a walkability index of a community i.e., block group (similar to the Phase I model), a new model that generates a "per-unit" measure, i.e.., average number of walk trips per day, was also established under this study. Both models were implemented into the prototype sketch planning tool developed under this Phase II study. Discussions on the modeling approaches and their resulting formulations are presented in this Section.

To increase transferability of the models and the resulting prototype tool developed under this project, publically available data elements (variables) took preference in the models being developed, against variables based on private data sources. For example, "Percent White" information derived from the public ACS data was used in the models, instead of using a similar measure calculated from the MWCOG data. The rationale of taking this approach was to allow more regions to benefit from the study results and adapt the tool developed from this project without requiring extensive data collection efforts.

Note that because dependent variables of the models (i.e., percent of walk trips in a block group, or number of walk trips per person per day in a block group) were based on MWCOG data, data elements taken from the same MWCOG source would generally be more consistent (e.g., collected via same method under same survey weighting scheme) and likely be associated with each other better (in terms of correlations) when used as independent variables in the model. Although most travel surveys use Census information (i.e., ACS) in their sample weighting calculations, most do not weight at disaggregated levels, such as a block group. Thus certain inconsistencies among MWCOG and ACS measurements at the more detailed geographic levels might exist, which in turn influences the goodness-of-fit in the developed models. This might be considered a drawback from the decision to substitute private variables with public data sources in the models.

Nevertheless for this project, the main focus was to develop a prototype sketch planning tool that demonstrates how such a system could benefit local planners in examining potential impacts from land-use or policy changes on walking behaviors in their communities. In other words, local governments who do not have the budget to conduct their own data collection efforts could use public data to get some rough estimates for their planning purposes. Therefore, the goodness-of-fit measure of a model was not the top priority here and thus was slightly relaxed. For example, under the model development process, 0.2 was used instead of the typical 0.05 significance level for allowing independent variables to enter into the models. This, however, did not result in much reduction in the goodness-of-fit of the models. All resulting models considered during the Phase II modeling development process have R-squared values of around 0.50, with most of their variables having a p-value of less than 0.05[1].

4.1 MEASURE OF WALKABILITY INDEX

4.1.1 Logistic Regression and Discriminant Analysis Based Models

A similar model formulation as applied in the Phase I NMT study was also employed in this Phase II project to produce a model that measured walkability index using data for the 722 block groups in the Modeling Region. Specifically, linear logistic stepwise regression procedures were used to select a set of variables (major factors) from the large number of potential data elements (discussed in Section 3). This selected set of variables was then applied with discriminant analysis to produce the final formulation of models for integration into the prototype system.

A few important changes from the original Phase I modeling approaches were made to (1) simplify the equation (model formulation), (2) properly account for walking trips in block groups, and (3) improve performance of the model (in terms of R-squared values). These changes are briefly described below.

Simplify the Model Formulation: Transformation (e.g., squared-value or taken the log of a variable) or interaction variables (e.g., multiply two variables to form a new variable) for certain factors were considered under Phase I models for performance improvement purposes. Under Phase II, it was determined that user-friendliness of the intended sketch planning tool should be a priority over the use of slightly better-fit but more complicated formulations, resulting from the inclusion of transformed or interaction variables. That is, for the targeted users of the prototype sketch planning tool, when they make changes to a transformed-factor, or to a factor that is involved in an interaction term, the resulting "impacts" might be confusing or rather difficult to comprehend. Thus, the model formulation was kept to a simple linear form.

Properly Account for Walk Trips in Block Groups: Under the Phase I model, the counts of walk trips per block group was household (HH) based. That is, each HH walk trip "credits" the block group in which the HH belongs, regardless where the trip took place (i.e., a HH-based approach). In Phase II, several alternatives were considered during the model development process, including counting walk trips by the block group of trip origin (O-based), counting by the block group of trip destination (D-based), as well as counting a trip that has ended within the given block group (OD-based). Alternatives that include walk trip as the access/egress mode to/from transit were also considered for each of the HH, O, D, and OD-based models (i.e., a total of 8 different alternatives were considered). Obviously,

[1] The P value is used to determine whether a variable is having a "significant" effect on the dependent variable in a regression model, assuming the model is specified correctly. A P value of 5% or less (i.e., 0.05 or less) is commonly used, which implies a 95% or more probability that the variable is having some effect on the dependent variable of the model. Although the P value is a measure of significance, its size does not necessarily reflect the size of the effect that variable has on the dependent variable. For instance, a highly significant variable (with a very small P-value) might have a miniscule effect.

including access/egress mode provided a higher number of walking trips for analyses than those concerning only the primary mode of trips.

Improve Model Performance: Typically a value of 0.05 is used in stepwise regression models to determine a factor to enter into or stay within a model. To allow for more factors to enter the regression equation, this value was increased to 0.1 (also experimented with 0.2) during the modeling process, which also resulted in a slightly increased R-squared value.

Note that the dependent variable in a logistic linear regression model is a binary term, in this case the "walk" or "no-walk" status of a block group, based on the 2007-2008 MWCOG travel survey data. Given that the Modeling Region consists of 722 block groups, there are potentially 722 records that can be used to fit the regression equations. A total of 15 block groups (of the 722) have missing values on the dependent variables (i.e., determination of walk or no-walk could not be made) because there were no sampled-HHs in the dataset for those block groups. Clearly, models including the access/egress mode in their trip counts would have more records for their regression estimation procedures than those without. A great deal of effort was devoted to investigating various formulations of the logistic regression models (under each of the eight alternatives) and to examining major factors resulting from those models. Generally speaking, models generated from origin-based and destination-based closely resembled each other. This is likely due to the "symmetric" nature of trips collected under a daily travel survey, where an out-going trip typically would have a reversed return trip.

The HH-based logistic regression models behaved less desirably than those considering trip-involved locations (i.e., origin (O), destination (D), and OD-based models), and thus were excluded from further consideration. Models that considered only the primary mode of trips were also dropped from later examinations. Furthermore, for simplicity, the team decided to restrain the total number of major factors in the final model implemented into the prototype system to under 20 variables.

4.1.2 Major Factors

Reviews of the findings from the O/D/OD-based models that considered both "walking as the primary mode" and "walking as the access/egress mode" resulted in 18 major factors. These factors were significant in having effects on the walk/no-walk status of a block group, and are used in the discriminant analysis to formulate the discriminant function for measuring the walkability index. These significant factors are listed in Table 4-11.

Note that these factors were listed in two columns based on their "changeable or "controllable" capability implemented under the prototype sketch planning tool. Specifically, those listed on the left of Table 4-1 are variables that allow for changes to be made by users of the prototype sketch planning tool, while those on the right were kept in their original measures so no changes can be made by the users.

Table 4-1. Major Factors with Significant Impacts on Measures of Walkability Index of a Block Group

Walk Score	Percent drivers
Number of transit arrivals per day	Average worker count per household
Percent workers	Percent of population with graduate degree
Land-use entropy	Percent of owner-occupied households
Cumulative number of jobs within 30 miles	Percent white
Working age population within 30 minutes transit plus walking commute	Percent of Spanish speaking households
Percent of workers who go to work at 7-9 am	Average household size
Population density	Working age population within 30 miles
Residential density	
Intersection density	

4.1.3 Formulation for Sketch Planning Tool

Discriminant analysis performed using the 722 block group dataset and the above listed major factors produced the estimated parameters for the associated discriminant functions shown in Table 4-2. The fundamental model development processes for Phase II followed similar approaches as developed under the Phase I study. The main purpose was not to develop new models, but rather to apply additional information from measures of regional characteristics (e.g., sidewalk and slope) to refine the list of discriminating factors (in a similar fashion as in Phase I) and to recalibrate the models for better use by the prototype system.

Table 4-2. Linear Discriminant Function Estimates

Variable	Label	Model_1	Model_0
Constant		-322.37822	-316.78131
Res_den	Residential Density = RES_SQ_FT / square-mile	1.31E-06	1.13E-06
walk_score	Walk score at block group centroid	-0.30525	-0.33384
TRANST_APD	Number of Transit Arrivals at All Transit Stations per day in a block group	-5.45E-06	-0.0000366
ENTROPY	Land use mix, All types	51.18981	49.77161
D5ae	Working-age population within 30 miles adjusted by travel time using gravity model, (in 1,000)	1.67917	1.65851
D5ar	Cumulative number of jobs within 30 miles adjusted by travel time using gravity model (in 1,000), EPA SLD	-0.86811	-0.85098
D5be	Working-age population within 30-minute transit + walking commute (in 1,000), EPA SLD	-0.08169	-0.08506

Table 4-2. Linear Discriminant function Estimates (Continued)

Variable	Label	Model_1	Model_0
AvgHHSize	Average household Size of households in block group, ACS Data	15.23461	14.81587
AvgWorkers	Average Worker Count of households in block group, ACS Data	-36.08337	-35.11394
PctWorkers	Percent Workers in block group, ACS Data	95.36266	93.81284
PctDrivers	Percent Drivers in block group, MWCOG Data	3.43252	4.83748
PctGrad	Percent grad degree, ACS Data	-40.04743	-41.02666
PCT_7_9_AM	Percent workers who go to work, 7-9 am, ACS Data	38.98053	37.44454
PCT_ESP_HH	Percent Spanish speaking households, ACS Data	-78.19101	-76.21577
PERDEN	Density of (ACS) Total persons divided by square-mile	0.0004363	0.000525
Inters_den	Intersection Density = INTERSECTI / square-mile	0.03281	0.03438

The walkability index which is calculated based on the probability estimated from the discriminant function can be expressed mathematically as: Walkability Index = 100 * Prob., where

Prob. = Probability of the given block group being a "walk-block group"
= EXP(model_1) / [EXP(model_1) + EXP(model_0)]

Here, model_1 and model_0 can be expressed in the simple linear regression equation formats using variables and estimated parameters listed in Table 4-2. That is, both model_1 and model_0 can be constructed into the simple linear regression form of

$$Y = \sum_0^{18} \beta_i V_i$$

In this equation,

Y represents the dependent variable;
V_i is the i-th independent variable listed in Table 4.2 for $i = 0, 1, \ldots, 18$, and V_0 is the constant term (i.e., intercept), V_1 is the variable "Res_den", … etc.; and
β_i is the corresponding parameter estimated for the i-th variable for $i = 0, 1, \ldots, 18$.

Results generated from this discriminant analysis are included in Appendix B of this report for further reference.

These equations and their associated dataset were then integrated under the "Sketch Planning Tool" feature of the prototype visualization system, which is discussed in Section 5 of this report. Walkability indices for the 35 block groups within the Study Area are calculated by applying their data into abovementioned equations.

4.2 PER-UNIT MEASURE

4.2.1 Log-linear Regression Based Model

In addition to estimating the walkability index of a community as discussed above, a second model using a "per-unit" measure (number of walk trips per capita per day) as the dependent variable was also created and integrated within the sketch planning tool under the prototype system. The motivation for developing such a model was to provide a simple and easily understood measure on the amount of walking activities with a community (block group), in addition to a more abstract measure of the walkability index.

Note that trips specify the primary mode being walking and those with walking as access/egress modes were combined in producing the walk-trip counts for this model estimation process. Because the distribution of these walk-trip counts showed a long tail on the right (i.e., skew to the left), natural-log scale of the block group walk-trip counts was used as the dependent variable with a simple linear regression approach. As in Phase I and in the process of developing the walkability index model, a stepwise regression procedure was used to select major factors influencing the "per-capita walk-trip counts" dependent variable, from the large set of variables collected from data sources discussed in Section 3. Again, for model simplicity, no transformation or interaction variables were considered in the model.

4.2.2 Major Factors

The log-linear regression model identified 13 variables with significant effects on the dependent variable of "number of walk trips per capita per day" using data for block groups within the Modeling Region. These factors along with their parameters estimated are listed in Table 4-3.

Table 4-3. Significant Factors for Number of Walk Trips per Capita per Day

	Variable	Description	Parameter Estimated
0	Intercept	Constant term	3.69102
1	walk_score	Walk score	0.00851
2	TRANSAPH18	Transit average arrival per hour in 18-hour operation, ORNL	0.00162
3	ENTROPY_R	Land-use mix, excluding residential, HazUS	1.22095
4	Inters_den	Intersection density, HazUS	0.00245
5	PctWhite	Percent white, ACS	1.38931
6	HU2010_DEN	Housing unit density in 2010, EPA SLD	-0.00005359
7	nrtl_den	Non-retail commercial density, HazUS	6.26E-08
8	PCT_MORE99K	Percent income 100k+, ACS	-1.22896
9	PctGrad	Percent with grad degree, ACS	1.19406

10	AvgVehs	Average number of vehicles, ACS	-0.2247
11	D5be	Working age population within 30 min. transit + walking commute (in 1,000), EPA SLD	0.0014
12	PctDrivers	Percent drivers, MWCOG	-0.57897
13	ls_day_den	Day time population density, ORNL	0.00000969

As mentioned earlier, interaction terms were not included in the models for simplicity reason. Several factors considered during the modeling process, however, were somewhat correlated. Individual effects from these factors, thus, might be lessened due to the lack of independencies among them. For instance, variable "PCT_MORE99K" was negatively associated with the dependent variable of "Number of Walk Trips per Capita per Day," while the "PctGrad" variable has a positive relationship with the dependent variable.

4.2.3 Formulation for Sketch Planning Tool

The log-linear regression model in this case can be expressed in the form of:

$$\ln Y = \sum_0^{13} \beta_i V_i$$

Where

> Y is the dependent variable, i.e., "number of walk trips per capita per day";
> V_i is the i-th independent variable listed in Table x, for i = 0, 1, …, 13, and V_0 is the constant term (i.e., intercept); and
> β_i is the corresponding parameter estimated for the i-th variable, for i = 0, 1, …, 13.

Alternatively, this model can be expressed in the alternative form of:

$$Y = e^{\sum \beta i V i}.$$

The later expression retains the original measure of the dependent variable (Y), i.e., number of walk trips, rather than its natural-log form ($\ln Y$), making it easier to communicate the results to users of the developed sketch planning tool. Thus this equation was selected and integrated into the prototype NMT Sketch Planning Tool (NMT-SPT) application discussed in the next Section.

5. PROTOTYPE SKETCH PLANNING TOOL

5.1 INTRODUCTION

Note that both models, the walkability index and the per-capita measure of walk trips as discussed in Section 4, were developed using data collected for the 722 block groups defined as the Modeling Region, which includes the 35 block groups of the Study Area. The NMT-SPT is designed for use within the targeted Study Area only, to illustrate the feasibility of building such a sketch planning tool for “local” planning purposes. Data from the Fall-2011 Geographically-Focused Survey was only used to obtain updated MWCOG-specific information, i.e., trip counts and percent of drivers by block group.

As pointed out in Section 3, Census geographic boundaries for the 35 block groups changed between 2000 and 2010. Such changes created challenges when applying data from Fall-2011 survey, defined using the 2010 boundary, to the models developed based on data elements collected under the 2000 geographic definitions. Therefore, all 2000 based block group significant data elements had to be converted to 2010 block group definitions for inclusion into the NMT-SPT application. This conversion was conducted using simple approaches as described in the following steps:

- If there was no change to a block group between 2000 and 2010 no corrections were needed;
- If the change was only to a block group ID number between 2000 and 2010, the 2000 ID number was converted to the 2010 ID number;
- If a 2000 block group was split into two 2010 block groups, the 2000 block group variables that were counts were halved and applied to each new 2010 block group. 2000 block group variables that were densities or percentages were left alone and applied to each new 2010 block group.
- If two 2000 block groups were merged into one 2010 block group, the 2000 block group variables that were counts were added together and applied to the new 2010 block group. The averages of 2000 block group variables that were densities or percentages were then applied to the new 2010 block group.

5.2 SYSTEM DEVELOPMENT PLATFORM

5.2.1 ArcGIS Server versus ArcGIS Online

The concept of building a prototype GIS-based visualization system as a case-study for the NMT project (i.e., as a Phase II task) was formed when scoping the initial Phase I effort of the project. Geospatial data collected/generated during this NMT project (both phases) were processed using ERSI ArcGIS software. The application was originally designed and developed under the ArcGIS JavaScript Application Program Interface, utilizing data and geo-processing services within the ArcGIS Server environment.

However, during the Phase II project development, FHWA requested that the NMT SPT application be developed under the FHWA ArcGIS Online platform, as an effort to increase data accessibility and consistency, as well as to follow the FHWA internal GIS data sharing concept. Although all project-related data layers were successfully migrated to ArcGIS Online, investigation by the research team found that development of the application was not achievable within the ArcGIS Online environment. ArcGIS Online provides a cloud based platform that allows for visualization of data through maps and specific applications. However, the current iteration of ArcGIS Online lacks the full suite of geo-processing tools needed for the required functionality of the NMT-SPT application. While the project-related data layers are available for visualization in ArcGIS Online, the NMT-SPT had to be developed using the ArcGIS Server environment mentioned previously.

5.2.2 System Configurations

The prototype NMT-SPT system was developed using ESRI's JavaScript API in Aptana Studio 3.0. The application currently draws on web-based services that are operating on ArcGIS Online as well as on an internal ORNL server through ArcGIS Server. The server specifications are:

- Windows Server 2008 R2 Operating System,
- 32 GB RAM,
- 2 Quad-core AMD Opteron processors running at 2.4 GHZ, and
- ESRI License: ArcServer 10.0, no installed service packs.

More details on the system-related information are included under Appendix C of this report.

5.3 NMT-SPT FEATURES

The NMT-SPT prototype system offers local planners a tool that can query and visualize the data in a GIS-based environment, as well as examine major contributing factors determining the walkability measure of a Census block group. The application can aid planning processes by highlighting factors that promote walking and healthier communities. Several functions available under the NMT-SPT prototype system are shown in a flowchart presented in Figure 5-1. This prototype NMT-SPT system is currently hosted at: http://faf.ornl.gov/NMT_Sketch/index.html.

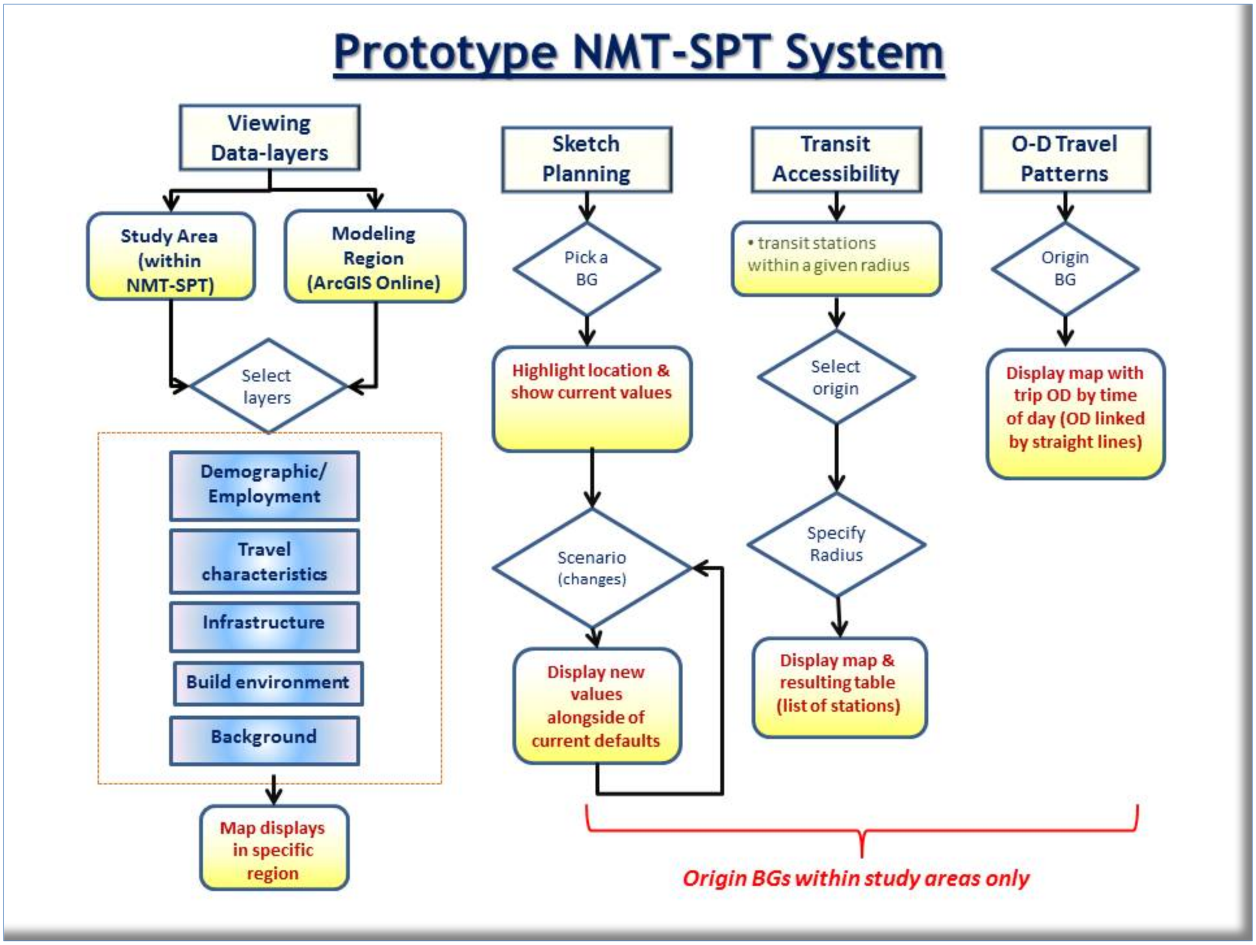


Figure 5-1. Flowchart Showing Components of the Prototype NMT-SPT System.

When the NMT-SPT is executed, the opening screen shows the Study Area (35 block groups) and provides four tabs (each representing a specific function) at the top of the display screen for users to select from (see Figure 5-2). These four tabs each correspond to one of the major functions displayed in the above flowchart, where ***Layers Overview*** allows a user to select one or more data layers for displaying on the map; ***Scenario Planning*** provides a user the ability to examine potential impacts from changes made on specific factors to the community's walkability measures; ***Transit Accessibility*** enables the user to find transit stations located within a user-specified walking-distance radius; and ***Travel Patterns*** shows the Origin-Destination travel patterns involving walking as captured by the Fall-2011 Geographically-Focused MWCOG Survey data. A brief discussion on each of the features is provided below.

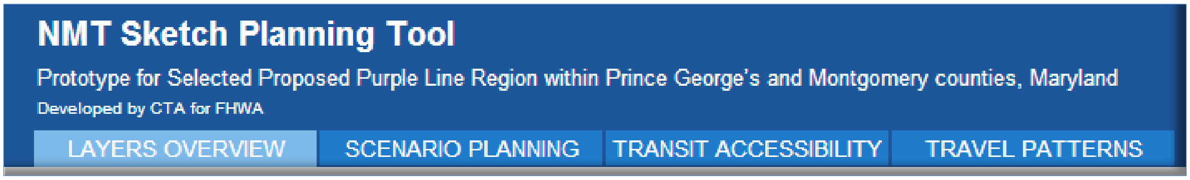
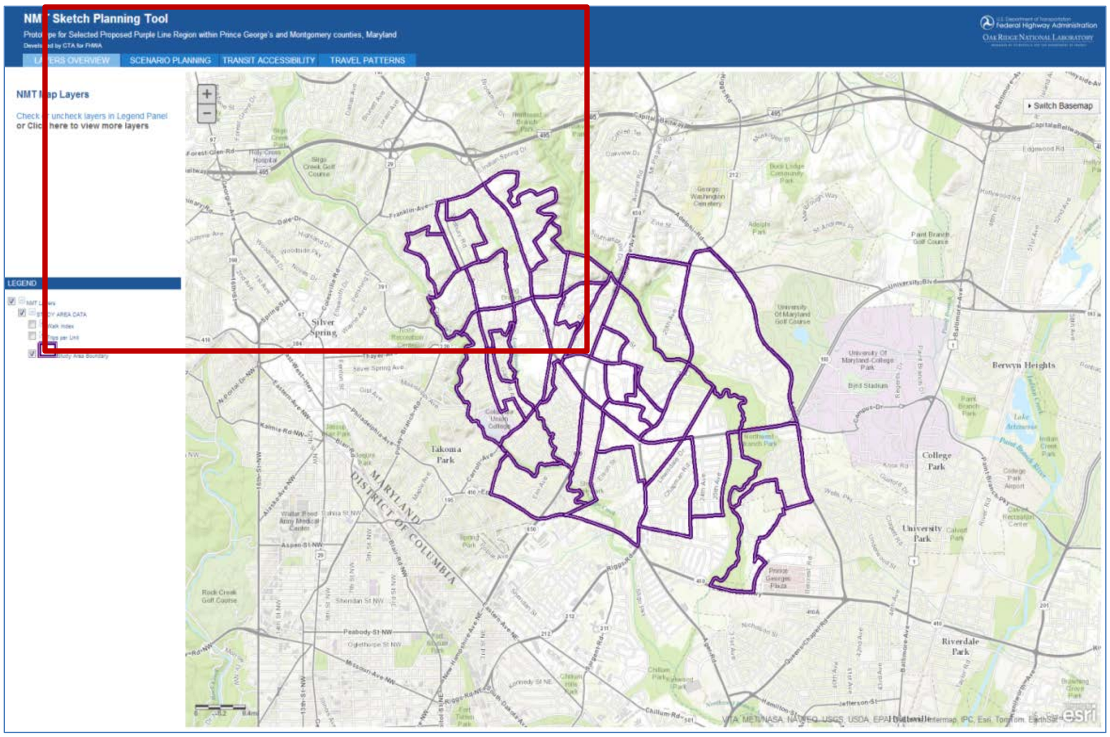


Figure 5-2. Menu of Functions Available in the NMT-SPT Application.

5.3.1 Viewing Data-Layers

The ***Layer Overview*** Tab is the default starting place for users of this application. This function allows the user to visually explore the different variables (data elements) used by the NMT-SPT application. The user can turn on/off a data layer, by a mouse click, to view information on the selected variable across block groups. Two sets of the data visualization platforms were implemented in the NMT-SPT system with ArcGIS Online for displaying data associated with the entire Modeling Region, as well as data layers related to the focused Proposed Purple-Line Study Area only.

From the top lefthand side window of the NMT-SPT openning screen (shown in Figure 5-3), users can click on the "Click here to view more layers" option and be connected to the ArcGIS Online map display website to view a large collection of geospatial data layers produced under the NMT Phase II project. Note that, currently users need to be authorized by the FHWA (i.e., having an account set up) in order to access these data layers under the FHWA ArcGIS Online platform (see Appendix E for more information on navigating the ArcGIS Online map display for the NMT Phase II project) . This top window typically contains contextual instruction/informaiton about the tool under other functions (e.g., ***Scenario Planning*** etc.). No user credentials are required to access the NMT-SPT application or its associated data layers (via the bottom part of the lefthand side window under the heading LEGEND). Data layers included in the NMT-SPT application are pertinent to the 35 block groups in the selected Purple Line Study Area, and are limited to variables used in the associated models.



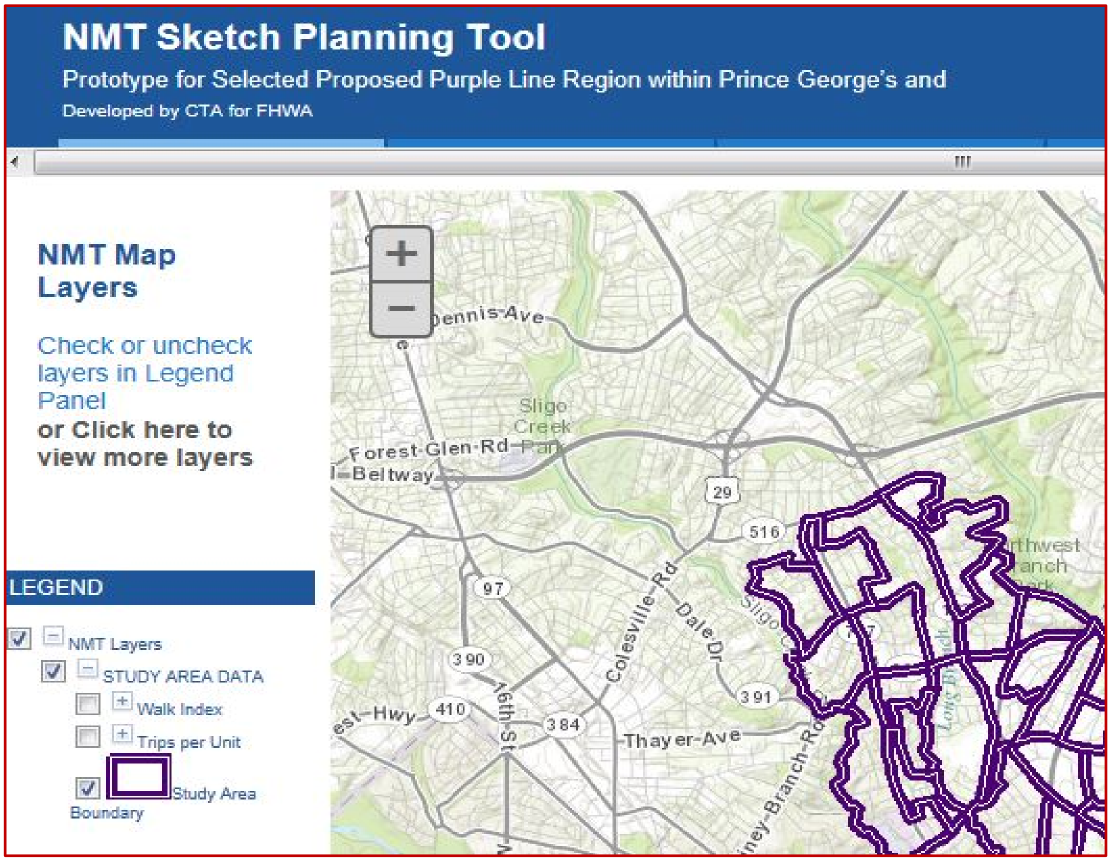


Figure 5-3. Opening Screen of the NMT-SPT (top) with a Zoomed-In View of the Map Layers (bottom).

On the upper right corner of the map display screen, a ***Switch Basemap*** option is provided to allow users to switch between various background base-map choices. This option is available in every function tab. Nine background base-maps are seen in Figure 5-4, including satellite imagery, topographic maps (the default setting), a street map, a plain gray canvas, and others, and can be selected at any time under any NMT-SPT function. The users can choose any background they wish to view the application with a simple click.

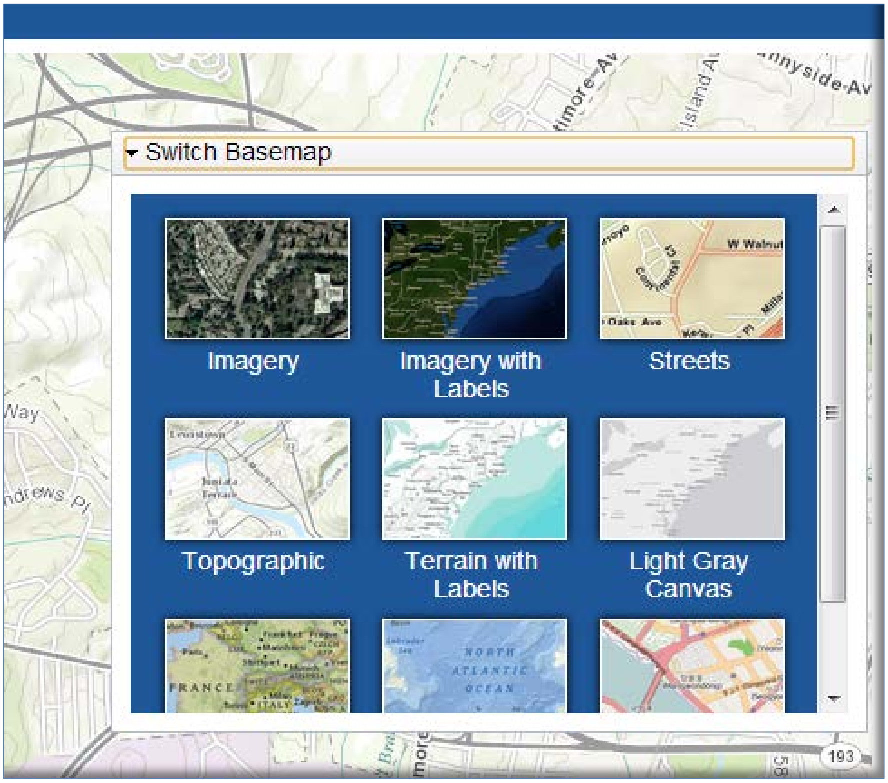


Figure 5-4. Base-Map Options for Background Display in NMT-SPT.

5.3.2 Scenario Planning Tool for Sketch Planning

This is the main function developed under this project, with the intention to allow local planners to conduct "what-if analysis" on block groups within the Study Area**.** This ***Scenario Planning*** tool was built around two models developed specifically under the Phase II project, one estimates the walkability index measure and the other produces per-unit walk trip counts (i.e., per capita per day). Users of this system can make changes on certain measures of community characteristics for a user-selected block group within the Study Area and see the impacts of those changes, in terms of changes in walkability index (category) or in the number of walk trips per-capita per-day. Figure 5-5 shows an example screen of the Scenario Planning Tool.

Specifically, the users can pick any one of the 35 block groups within the proposed Purple Line Study Area to examine potential effects due to changes in factors that influence walkability. Default values of measures (based on data used in the model developments) are shown on the thematic map when this ***Scenario Planning*** function is initiated, depending on whether "Walk Index" tab or the "Trips per Unit" tab was selected by the user.

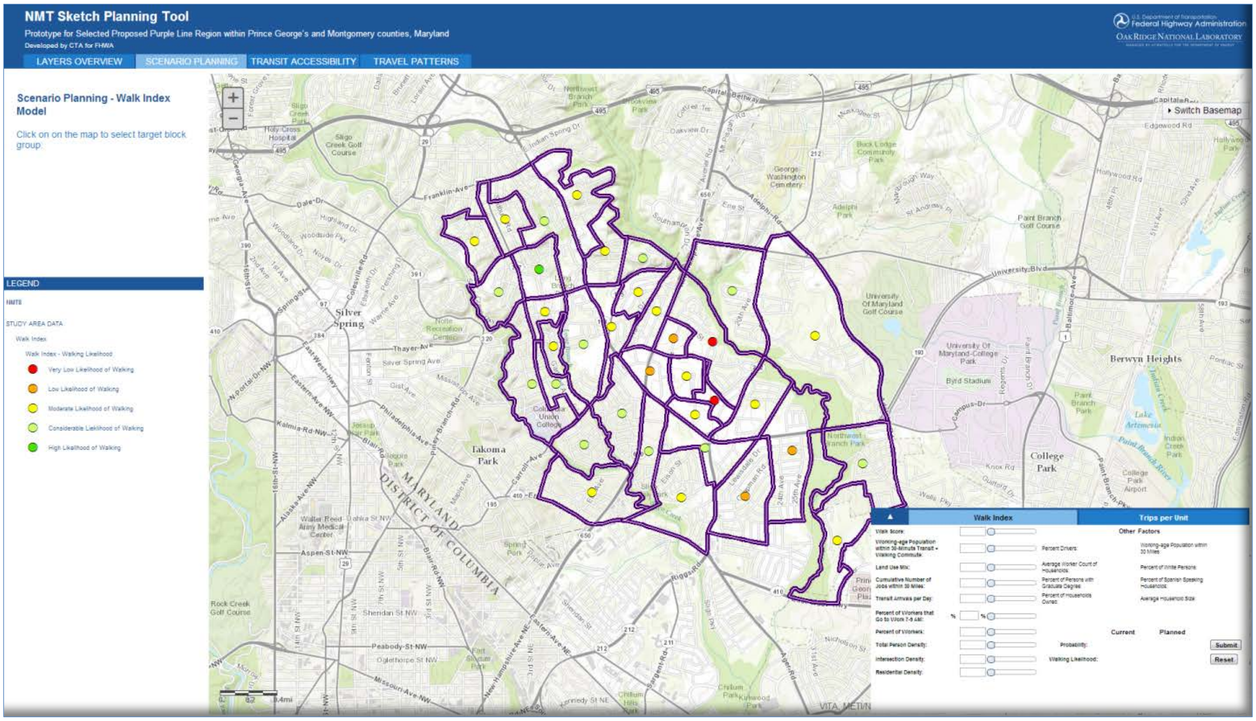


Figure 5-5. Scenario Planning Tool Screen for "Walkability Index."

The Walk Index Tab is used to explore the factors that determine a block group's Walking Likelihood (walkability). The user would select the block group of interest from the map display. Current values for all factors/variables used in the model, as well as a current estimate of the Walk Index (ranges from 0 to 100) and Walking Likelihood category are displayed. To make changes to any of the variables, one can simply type in a new value in the provided text box or adjust the value manually with the slider bar. Note that variables impacting walkability (i.e., included in the model) but are not adjustable are also shown in the window for reference.

Figure 5-6 gives an example of a closer view of the Walk Index window, where variables, both adjustable and non-adjustable, are shown with their current values and corresponding model-generated result for the walkability measures. The user can make changes to a selected variable, or on multiple variables, and click on the "Submit" button to update the model results. The new measure of the Walk Index (probability) and its corresponding Walking Likelihood classification are then displayed under the "planned" column. To return to the initial setting (i.e., original values as used for model development), the user can click on the "Reset" button.

The "Trips per Unit" Tab option utilizes the log-linear regression model discussed in Section 4.2 which measures the walkability of a block group in terms of number of trips per capita per day. As in the "Walk Index" option, some variables of this model allow user adjustments and others do not. The operations of both estimation options are identical: where the user can enter a new value for a selected variable or adjust the value by using the slider bar. When adjustments are made and the user clicks the "Submit" button, a new "Walk Trips per Person per Week" value is then displayed under the "Planned" column. An example is shown in Figure 5-7. Note that the selected block group is highlighted with a red boundary in the map.

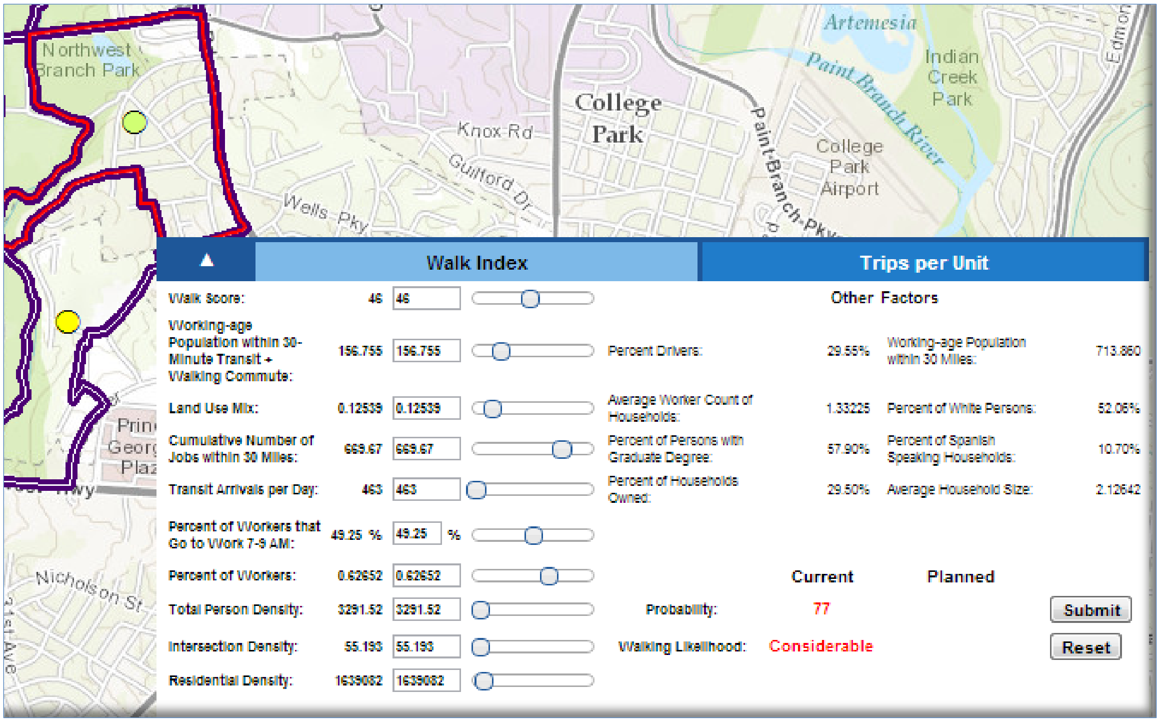


Figure 5-6. Walk Index Option of the NMT-SPT, with Variables and Slider-bars.

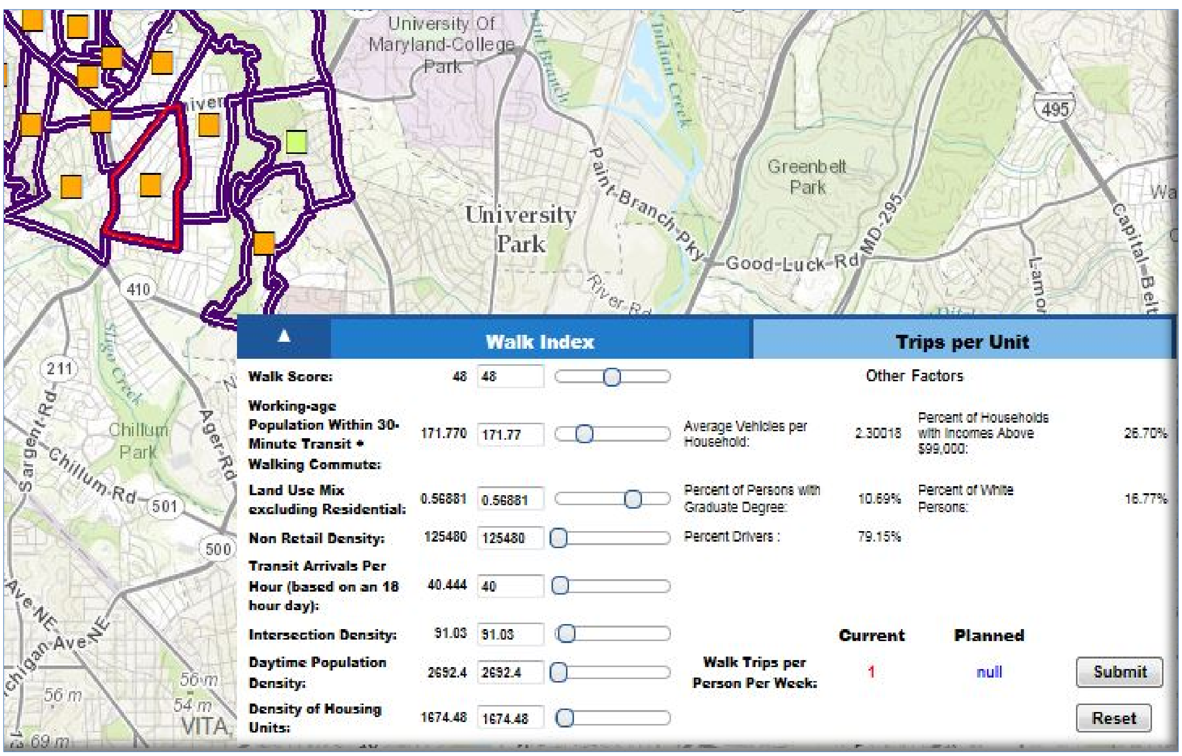


Figure 5-7. Trips per Unit Option of the NMT-SPT.

To increase user-friendliness of this NMT-SPT application, contextual information for all variables used in the estimation models was also built into the system. A user can click on any desired variable name to bring up a dialog box containing a brief description of the variable and related URL, if available, for further reference. See Figure 5-8 for an example on the variable "Walk Score."

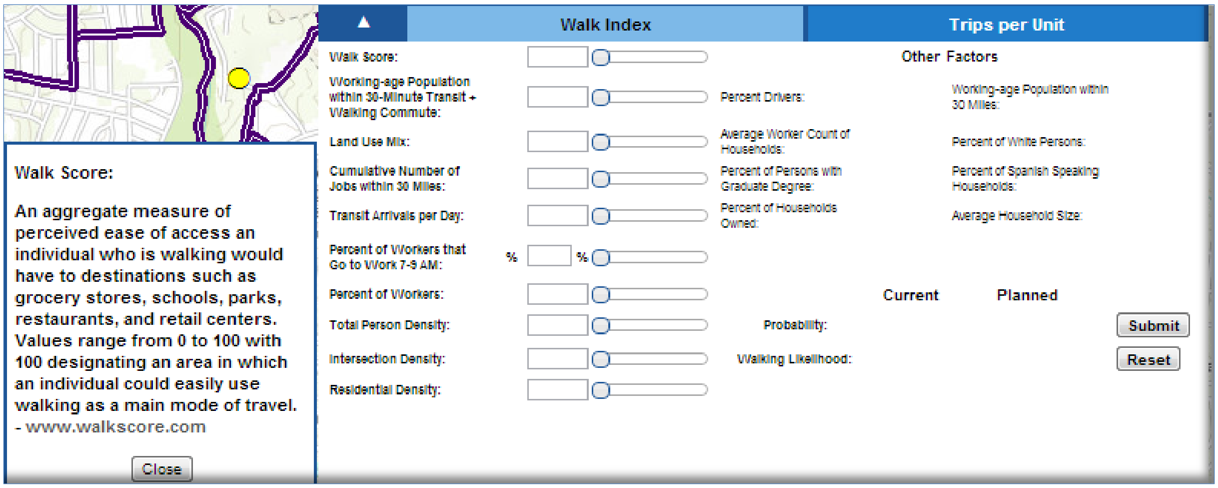


Figure 5-8. Example of the Variable Description Box.

5.3.3 Transit Accessibility Function

The ***Transit Accessibility*** function is a tool that allows users to view available transit stations within a specific radius of a user-specified location within the Study Area. Because the focus is on walking trips, an upper bound of one mile was implemented to ensure a reasonable walking distance. Figure 5-9 shows all transit station locations within the entire Study Area (35 block groups), along with the proposed Metrorail Purple Line and its associated stations (shown in purple).

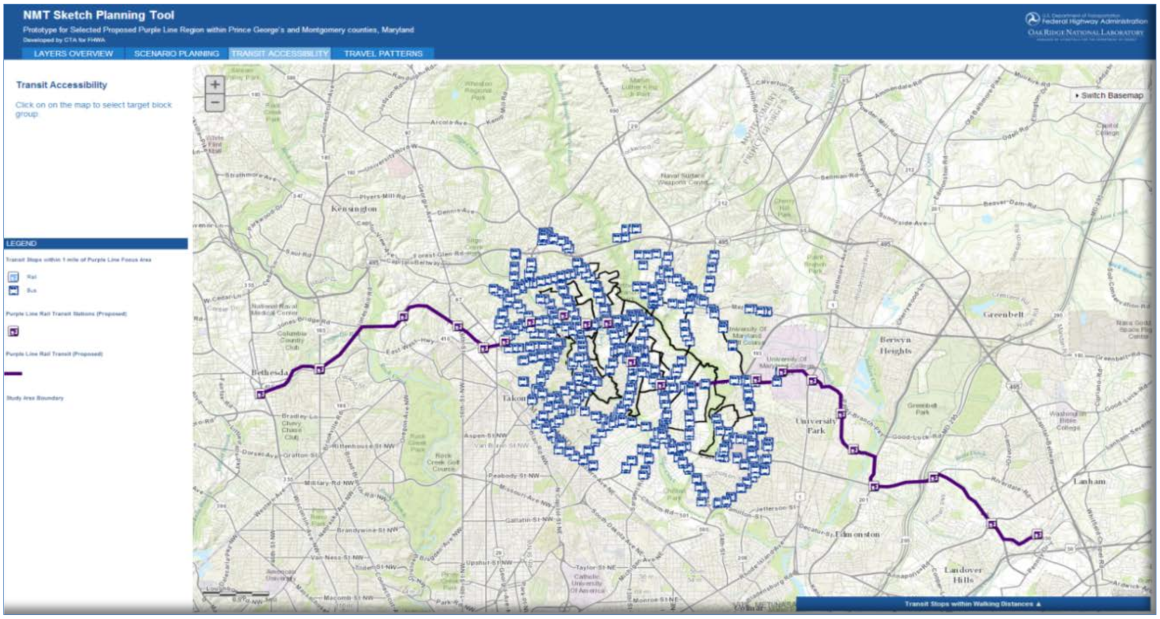


Figure 5-9. Transit Accessibility Function Display Screen.

To initiate this ***Transit Accessibility*** function, a user first selects a desired location by clicking on a point anywhere within the 35 block groups, which brings up the "Walking Distance" dialog box as shown in Figure 5-10.

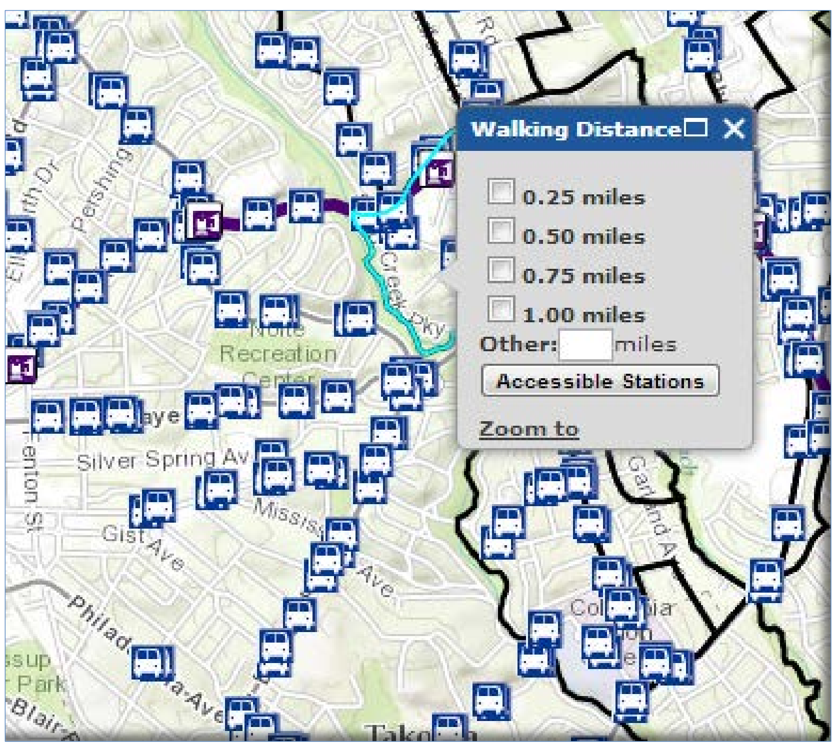


Figure 5-10. Dialog Box for Walking Distance Options under the *Transit Accessibility* Function.

The user can select one or more of the pre-determined radii, or enter another desired distance of up to 1 mile great-circle distance, and then click on the “Accessible Stations” button. Once the process is completed, the results (i.e., all stations fall within the user-specified radius) will be listed under the “Transit Stops within Walking Distance” window on the screen.

Moreover, a map showing the shaded area around the user-selected point (with user-desired radius) will be displayed. The “Zoom To” button from the “Walking Distance” window can be used to quickly move the map display to the user-specified point. Furthermore, if any of the Proposed Purple Line Rail Transit Stations are within the radius identified by the user, these proposed stations will also be listed in the results window. Figure 5-11 provides an example of results displayed in the “Transit Stops within Walking Distances” box. Note that existing transit stops are listed on the left and the proposed Purple Line rail transit stations are shown on the right of the results window.

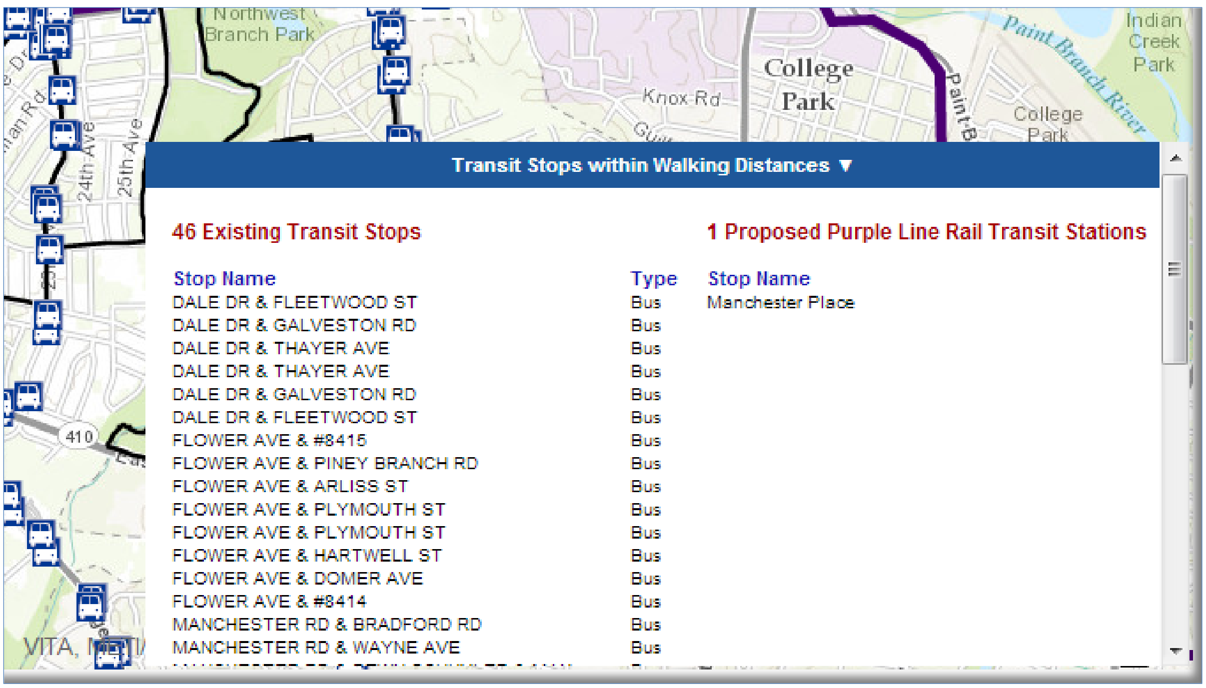


Figure 5-11. Example of Results Displayed in the *Transit Accessibility* Function

5.3.4 Travel Patterns Function

The purpose of the ***Travel Patterns*** function provided in the NMT-SPT application is to illustrate the use of animation in visually displaying travel patterns using O-D travel data for trips involving walking, and originating from the Study Area (see Figure 5-12). Data from the Fall-2011 MWCOG Geographically-focused Survey was used. At the current time, this “function” is more of a visualization tool (an animation) of the trip data, than as an actual tool. The O-D travel data is currently shown by time of day over a period of 18 hours. Future developments could enhance this function to allow, for instance, selections of trips by purpose, trips by block group of origin/destination, etc.

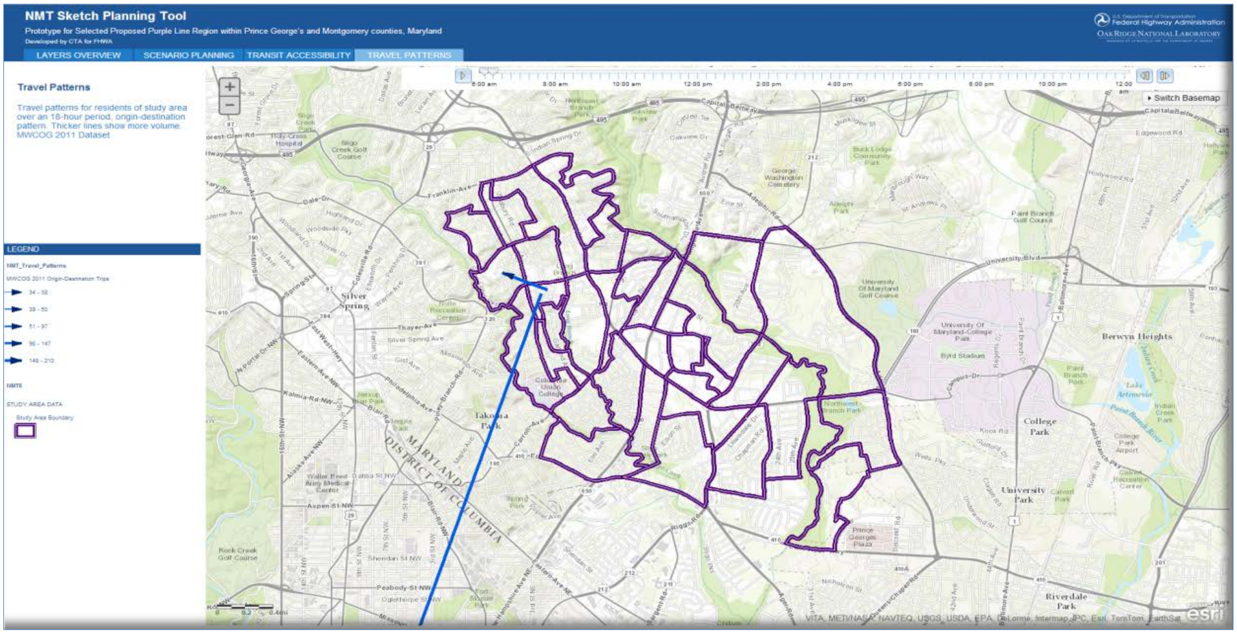


Figure 5-12. Screen Display of the Travel Patterns Tab within the NMT Sketch Planning Tool.

Note that the origin and destination points shown in this visualization are on the associated block group centroids, which do not represent any specific physical locations within those involved block groups.

6. SUMMARY AND CONCLUSIONS

The twofold goal of this project was to:

1. Apply the fundamental modeling approach as developed from the NMT Phase I study to estimate the measure of walkability for a selected small community in the Washington D.C. area based on the set of major discriminant factors identified under Phase I; and
2. Develop a GIS-based visualization prototype sketch planning tool that implements the model and data from above.

Because of changes made during the scoping of the Phase II study, additional efforts for conducting data collection as well as major recalibration of models were required. Thus, the Phase II study was in fact more "standalone" than was originally designed.

There were improvements made to the models in the Phase II research, which expanded the model formulation from one that estimates a probability-based walkability index to also include another model that provides a more user-friendly per-capita measure. Information needed for the models was shifted, as much as possible, from private data to publically available data sources. This was done to permit more regions to benefit from the study results, and allow them to adapt the tool developed from this project without requiring extensive data collection effort.

This study accomplished its goal of integrating travel survey data with other data sources in models that measure walkability and, with those models and data, developed a GIS-based prototype system that allows local planners to make "what-if" scenario analysis regarding walking in their communities. Although this prototype system was designed using a small community around the Proposed Purple Line area as the case study, it is expected that transferability of the system to other communities can be established with relatively minimal effort.

The prototype NMT-SPT system provides four functions, including data visualization, scenario planning, transit accessibility, and an animated view of travel patterns as captured by the Fall-2011 MWCOG survey data. Other tools or new functions can be added as individual functions at a later time if desired. For example, allowing users to select a block group to view travel patterns for out-bound trips (trips originating from the selected block group) or in-bound trips (travel ending in a selected block group). Additional functionality, such as options for printing tables or maps as shown on the screen, viewing of attribute data for a specific transit station, etc. can also be included to enhance the application.

As mentioned in the report, aside from considering only some limited attributes associated with walking behavior, Walk Scores from the WalkScore.com website are point-estimates, measured at block group centroids in this project. These scores do not likely account for overall activities regarding "walkability" within the block groups where their centroid points are located. On the other hand, the estimation of the walkability index measure produced from this study took into account various factors at the block group level, i.e., not at any specific point location. This difference made the two fundamentally not comparable to each other. Table D.1 in Appendix D is provided only for a side-by-side view of what these two walkability measures look like for block groups in the Proposed Purple Line Study Area. The new per-capita measures generated from this project are also included in the Table D.1 for reference.

Table 6-1 provides a cross tabulation of the counts of block groups within the Proposed Purple Line Study Area by walkability categories, between walkability categories assigned by using Walk Score and those based on Walkability Index estimated from the model developed under this study. Although, as pointed out above, these two measures are not directly comparable, Table 6-1 shows in a broad sense that these two methods "agree" on walkability measures for about 80% of the block groups in this Study Area. Figures 6-1 and 6-2 provide visual displays of the categories for block groups in the Study area, by their measures of Walk Score and Walkability Index, respectively.

Table 6-1. Walkability Categories for Block Groups in Study Area

Walk Score (at block group centroids)	**Walkability Index** (of block groups)					
	very low likelihood	low likelihood	moderate likelihood	considerable likelihood	high likelihood	Total
very car-dependent	**0**	**0**	1	0	0	1
car-dependent	**1**	**2**	1	2	0	6
somewhat walkable	1	2	**12**	**9**	**0**	24
very walkable	0	0	**1**	**2**	**1**	4
All	2	4	15	13	1	35

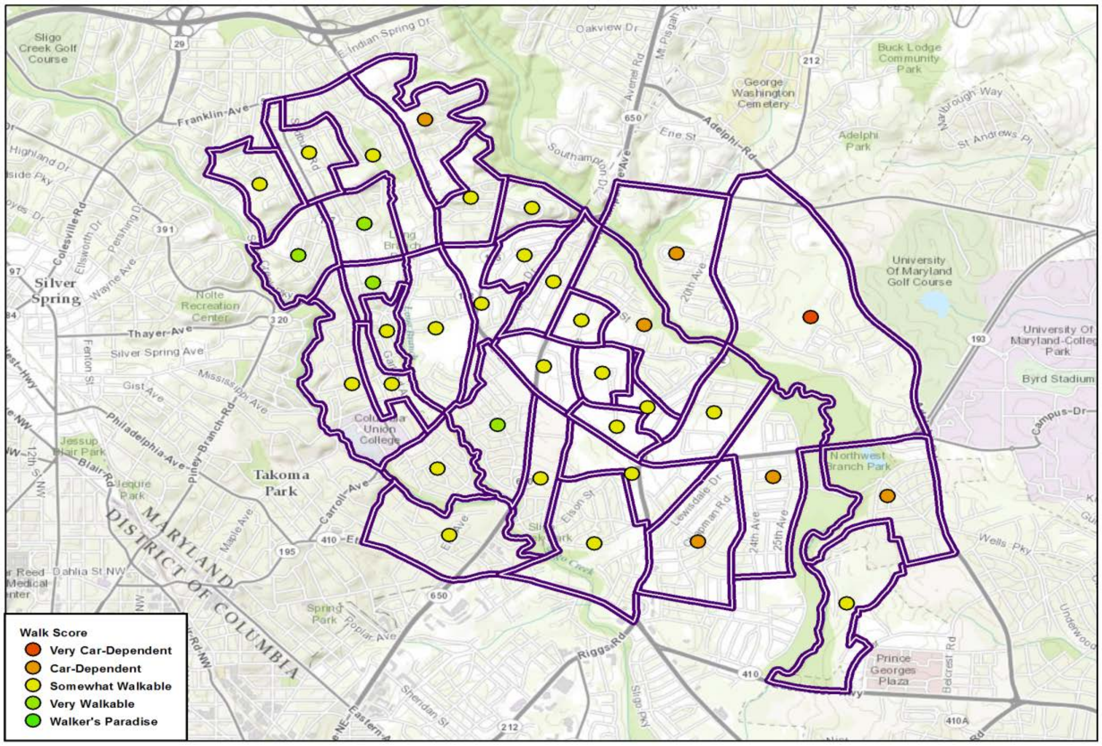


Figure 6-1. Walk Score from walkscore.com for the Study Area Block Groups.

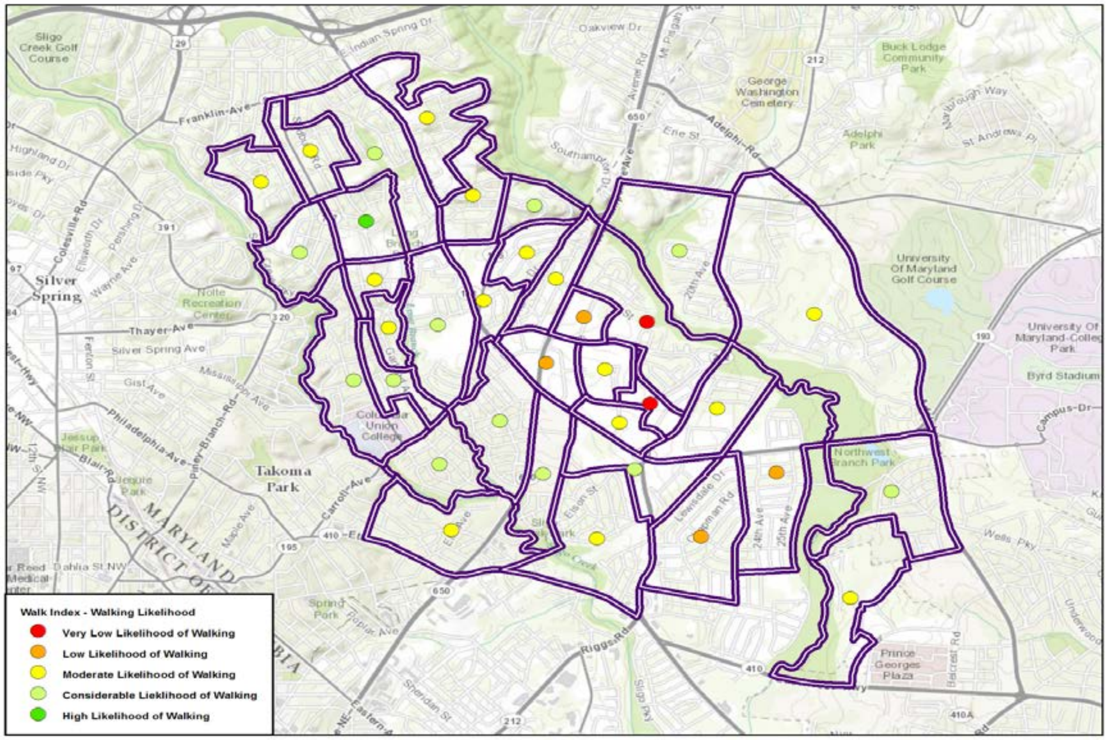


Figure 6-2. Walkability Index (Likelihood) for the Study Area Block Groups.

As previously mentioned, a study of Arlington County, VA was conducted as a part of the NCHRP-08-78 research [19]. In the Arlington Study, the walk-accessibility score was network-based and created at each trip end. The Phase II NMT project is very different in concept from the Arlington Study. Under the Phase II NMT project, the unit of measure is at the block group level. Thus the focus of this NMT study is on the community as a whole, rather than individual households or trips. The NMT-SPT application produced from this project is intended for community sketch planning, and not for studying mode choice or other trip demand modeling purposes.

As stated in Section 5.2.1, the current iteration of ESRI's ArcGIS Online is limited in its capability to provide processing functionality for a customized application like NMT-SPT. However, ESRI continues to expand the abilities of ArcGIS Online and quite possibly will allow for the implementation of the tools necessary for a customized application to exist in the ArcGIS Online framework in the future.

REFERENCES

1. Maryland Transit Administration, Gene Thorp/The Washington Post, published on January 11, 2012, 9:46 p.m. http://www.washingtonpost.com/local/more-communities-may-be-affected-by-purple-line/2012/01/11/gIQAAefEsP_graphic.html.

2. Smart Location Database, U.S. Environmental Protection Agency. http://www.epa.gov/dced/smartlocationdatabase.htm.

3. Cervero, R. and K. Kockelman "Travel Demand and the 3Ds: Density, Diversity, and Design." *Transportation Research Part D*, Vol. 2, 1997, pp. 199–219.

4. Ewing, R., K. Bartholomew, S. Winkelman, J. Walters, and D. Chen, *Growing Cooler: The Evidence on Urban Development and Climate Change*. Urban Land Institute, 2007.

5. American Community Survey, U. S. Census, website: http://www.census.gov/acs/www/.

6. Walk Score Methodology, www.walkscore.com/methodology.shtml.

7. Hwang, H. L., T. Reuscher, D. Wilson, R. Schmoyer, *Final Report for Phase I Study to Characterize the Market Potential for Non-Motorized Travel*, ORNL/TM-2012/223, Oak Ridge National Laboratory, June, 2012. http://info.ornl.gov/sites/publications/Files/Pub36798.pdf.

8. HAZUS, The Federal Emergency Management Agency's (FEMA's) Methodology for Estimating Potential Losses from Disasters. http://www.fema.gov/hazus.

9. Diversity Indices: Shannon's H and E, http://www.tiem.utk.edu/~gross/bioed/bealsmodules/shannonDI.html.

10. Frank, L. D., M. A. Andresen, and T. L. Schmid, "Obesity Relationships with Community Design, Physical Activity, and Time Spent in Cars." *American Journal of Preventive Medicine*, Elsevier, page 87-96, Vol. 27(2), 2004. http://download.journals.elsevierhealth.com/pdfs/journals/0749-3797/PIIS074937970400087X.pdf.

11. Federal Highway Administration, *An Analysis of Factors Contributing to "Walking Along Roadway" Crashes: Research Study And Guidelines for Sidewalks And Walkways*, Report No. FHWA-RD-01-101, February 2002.

12. *Sidewalks Promote Walking, BTS Issue Brief*, Number 12, December 2004, Bureau of Transportation Statistics, U.S. Department of Transportation, http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/special_reports_and_issue_briefs/issue_briefs/number_12/pdf/entire.pdf.

13. ArcGIS Sidewalk Inventory shape file provided by Mr. Mike Barry of the FHWA.

14. Personal conversation with Mr. Martin Howes of the Prince-George’s County, MD.

15. Personal conversation with Mr. Christopher McGovern of Montgomery County, MD.

16. *The National Map Viewer*, U.S. Geological Survey, http://viewer.nationalmap.gov/viewer/.

17. *How Slope Works*, *ArcGIS Help 10.1,* ArcGIS Resources, ESRI, http://resources.arcgis.com/en/help/main/10.1/index.html#//00q90000001r000000.

18. *ADA Standards for Accessible Design,* 28 CFR Part 36, http://www.ada.gov/adastd94.pdf.

19. *Estimating Bicycling and Walking for Planning and Project Development,* Practitioner Guidebook, NCHRP Project 08-78a, October 2013.

APPENDIX A: DATA ELEMENTS FROM MWCOG TRAVEL SURVEY DATA AND EPA SMART LOCATION DATABAS

Variables Derived from MWCOG 2007-2008 Household Travel Survey Data

Data elements considered as **Independent Variables** during this study include:

- AvgIncome — Average Income for HHs in BG
- AvgHHSize — Average HH Size of HHs in BG
- AvgVehCount — Average Vehicle Count of HHs in BG
- AvgWorkers — Average Worker Count of HHs in BG
- AvgDrivers — Average Driver Count of HHs in BG
- PctOwnHH — Weighted Percent of Owner-occupied HHs in BG
- PctWhitePer — Weighted Percent of White Person in BG
- PctAfAmPer — Weighted Percent of African American Persons in a BG
- PctAsAmPer — Weighted Percent of Asian American Persons in a BG
- PctHispPer — Weighted Percent of Hispanic Persons in a BG
- PctMales — Percent of Males in a BG
- PctWorkers — Percent of Workers in a BG
- PctDrivers — Percent of Drivers in a BG
- PctBikeJTW — Percent of Workers in a BG Who Biked to Work in the Last Week
- PctWalkJTW — Percent of Workers in a BG Who Walked to Work in the Last Week
- Pct5_15 — Percent of Persons in a BG age 5-15
- Pct16_24 — Percent of Persons in a BG age 16-24
- Pct25_34 — Percent of Persons in a BG age 25-34
- Pct35_44 — Percent of Persons in a BG age 35-44

Data elements considered as **Dependent Variables** include:

- WtdWalkTrpsMain — Weighted number of Walk Trips as Main Mode, HH Location
- WtdWalkTrpsAll — Weighted number of Walk Trips as Main or Access-Egress Mode, HH Location
- WtdWalkTrpsMainO — Weighted number of Walk Trips as Main Mode, Origin
- WtdWalkTrpsAllO — Weighted number of Walk Trips as Main or Access-Egress Mode, Origin
- WtdWalkTrpsMainOD — Weighted number of Walk Trips as Main Mode, Origin or Destination
- WtdWalkTrpsAllOD — Weighted number of Walk Trips as Main or Access-Egress Mode, Origin or Destination

- WtdWalkTrpsMainD Weighted number of Walk Trips as Main Mode, Destination
- WtdWalkTrpsAllD Weighted number of Walk Trips as Main or Access-Egress Mode, Destination

EPA Smart Location Database Variables

- P_wrk2000 Proportion of Population of Working Age in 2000
- Prop025 Proportion of BG within ¼ Mile of a Fixed-Guideway Transit Station
- Prop05 Proportion of BG within ½ Mile of a Fixed-Guideway Transit Station
- D1a Housing Density (Units per Unprotected Acre in 2010)
- D1b Population Density (People per Unprotected acre in 2010)
- D1c Employment Density (Employees per Unprotected Acre in 2008)
- D2 Entropy Index of Commercial, Industrial, Institutional, Retail, Recreational, and Residential within a BG
- D3 Weighted Measure of Intersection Density (per Square Mile)
- D4a Number of Fixed-Guideway Transit Stations within ¼ Mile
- D4b Number of Fixed-Guideway Transit Stations within ½ Mile
- D5ae Working –Age Population within 30 Miles (Adjusted by Travel Time using Gravity Model)
- D5ar Cumulative Number of Jobs Within 30 Miles (Adjusted by Travel Time using Gravity Model)
- D5ce Working-Age Population within ½ Mile of Transit Stations along a Transit System that is Accessible within ¼ Mile of Home BG
- D5cr Number of Jobs within ¼ Mile of Transit Stations along a Transit System that is Accessible within ½ Mile of Home BG
- Ac_Unpro Land area that is Privately Owned and Unprotected from Development (e.g. not Parks, Lakes, or Conservation Areas)
- HU2010 Number of Housing Units in 2010
- Emp2008 Total Number of Jobs in 2008
- EmpInd08 Total Number of Industrial Jobs in 2008 (NAICS Sectors 11, 21-23, 31-33, 42, 48-49)
- EmpCom08 Total Number of Commercial Jobs in 2008 (NAICS Sectors 11, 21-23, 31-33, 42, 48-49)
- Pop2010 Population 2010
- Ac_Land Total Land Area of Block Group in Acres
- D5br Number of Jobs within 30 Minute Transit + Walking Commute
- D5be Working-Age Population within 30 Minute Transit + Walking Commute

APPENDIX B: DISCRIMINANT ANALYSIS RESULTS

(WALKABILITY INDEX MODEL)

The FREQ Procedure

Annual Walk Trips (incl ACC/EG) Per Cap (Interval mid-point)				
WTPC_AE_OD	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
-4	134	18.69	134	18.69
1	15	2.09	149	20.78
2	318	44.35	467	65.13
3	234	32.64	701	97.77
4	15	2.09	716	99.86
5	1	0.14	717	100.00

Frequency Missing = 5

Annual Walk Trips per Person, incl Acc/Egr, 50-50 Priors

The DISCRIM Procedure

Total Sample Size	646	**DF Total**	645
Variables	18	**DF Within Classes**	644
Classes	2	**DF Between Classes**	1

Number of Observations Read	722
Number of Observations Used	646

Class Level Information					
WTPC_AE_OD_I	**Variable Name**	**Frequency**	**Weight**	**Proportion**	**Prior Probability**
ONE	ONE	550	550.000	0.851393	0.500000
ZERO	ZERO	96	96.000	0.148607	0.500000

Pooled Covariance Matrix Information	
Covariance Matrix Rank	**Natural Log of the Determinant of the Covariance Matrix**
18	69.27301

Annual Walk Trips per Person, incl Acc/Egr, 50-50 Priors

The DISCRIM Procedure

Generalized Squared Distance to WTPC_AE_OD_I		
From WTPC_AE_OD_I	**ONE**	**ZERO**
ONE	0	1.32741
ZERO	1.32741	0

Linear Discriminant Function for Wt_WTPC_AE_OD_I			
Variable	**Label**	**ONE**	**ZERO**
Constant		-322.37822	-316.78131
res_den	Residential Density = RES_SQ_FT / SQMI	1.31009E-6	1.1295E-6
walk_score	Walks core at block group centroid from Phase 1	-0.30525	-0.33384
TRANST_APD	Number of Transit Arrivals at All Transit Stations per day in a Block Group	-5.4487E-6	-0.0000366
ENTROPY	Land Use Mix, All HAZUS Types	51.18981	49.77161
D5ae	Working-age population within 30 miles adjusted by travel time using gravity model	1.67917	1.65851
D5ar	Cumulative number of jobs within 30 miles adjusted by travel time using gravity model	-0.86811	-0.85098
D5be	Working-age population within 30-minute transit + walking commute	-0.08169	-0.08506
AvgHHSize	Average HH Size of HHs in BG, ACS Data	15.23461	14.81587
AvgWorkers	Average Worker Count of HHs in BG, ACS Data	-36.08337	-35.11394
PctOwnHH	Percent of owner-occupied HHs in Block Group, ACS	3.33436	4.28907
PctWhitePer	Wtd Pct White Pers in BG, ACS Data	30.66145	29.47652
PctWorkers	Percent Workers in BG, ACS Data	95.36266	93.81284
PctDrivers	Percent Drivers in BG, MWCOG Data	3.43252	4.83748
ACSPctGrad		-40.04743	-41.02666
PCT_7_9_AM	% worker go to work, 7-9 am	38.98053	37.44454
PCT_ESP_HH	% spanish speaking HH	-78.19101	-76.21577
TOTPERS_ACSDEN	Density of (ACS) TotPers Divided by SQMI (from GIS Variables)	0.0004363	0.0005250
inters_den	Intersection Density = INTERSECTI / SQMI	0.03281	0.03438

Annual Walk Trips per Person, incl Acc/Egr, 50-50 Priors

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.DATATMP
Re-substitution Summary using Linear Discriminant Function

Number of Observations and Percent Classified into WTPC_AE_OD_I			
From WTPC_AE_OD_I	**ONE**	**ZERO**	**Total**
ONE	385 70.00	165 30.00	550 100.00
ZERO	22 22.92	74 77.08	96 100.00
Total	407 63.00	239 37.00	646 100.00
Priors	0.5	0.5	

Error Count Estimates for WTPC_AE_OD_I			
	ONE	**ZERO**	**Total**
Rate	0.3000	0.2292	0.2646
Priors	0.5000	0.5000	

APPENDIX C: NMT-SPT APPLICATION CONFIGURATION AND SYSTEM REQUIREMENTS

Application Link: http://faf.ornl.gov/NMT_Sketch/index.html

Application Description:

- Developed using ESRI's JavaScript API in Aptana Studio 3.0.
- The JavaScript program is currently running on the FAF.ORNL.GOV server, and was published under the server's www root folder. The code for the application can be accessed in the application folder.
- The application draws on web-based data services that are currently operating on ArcGIS Online as well as the FAF Server (through ArcGIS Server)

Server Specifications:

- Operating System: Windows Server 2008 R2
- RAM: 32GB
- Processors: 2 Quad-Core AMD Opteron running at 2.4 GHz
- ESRI ArcGIS License: ArcServer 10.0, no installed service packs

System Operations:

- The data and tools used by the application are hosted server-side, and are distributed through separate services running under ArcServer and ArcGIS Online.

  http://faf.ornl.gov/FAF/rest/services/NMT_STUDY_AREA_DATA/MapServer

  This service provides the primary source of data for the application, including the layers seen on the Layers Overview tab as well as the data processed by the models and geo-processing tools in the other tabs.

  The data used by this service is located within the _Map Data folder and is named NMT_STUDY_AREA_DATA.mxd

- The application requires several support services running on the server to undertake various processing tasks:

  http://faf.ornl.gov/FAF/rest/services/Geometry/GeometryServer

  This is a generic ESRI Geometry service required by the JavaScript API to undertake geo-processing tasks. The function of the service is to respond to requests from the application for any geo-processing actions or calculations, such as acquiring the geometry of a targeted block group and performing buffer operations and spatial calculations.

- The Walk Index - Walking Likelihood and the Walk Trips per Person per Week layers as well as the topographic map reference and associated symbology are retrieved from the following ArcGIS Online web maps:

  - http://hep.maps.arcgis.com/home/item.html?id=5af3210b658d4428a3cc655d6b84552f
  - http://hep.maps.arcgis.com/home/item.html?id=c5263b457cc54a8498ac044d888ab651

Steps Needed For Setting Up New Server:

To deploy the application on a new server the following steps must be taken:

- Copy the application data folder (_Map Data) to the target server.
- Deploy the MXD file (NMT_STUDY_AREA_DATA) as a service running under ArcServer 10.0.
- Deploy the Geometry service if one is not currently running.
- The Application code will need to be modified to access the new URLs of the services, this can be done using a programming IDE by simply replacing all references to the old URL addresses with the new addresses.
- Once the application code is updated, it will need to be recompiled and deployed under the www root directory of the server.
- Testing will need to be done in ensure that the services are being called on correctly and adjustments may need to be made in the application code and then redeployed to the server.

APPENDIX D: MEASURES OF WALK SCORE WITH WALKABILITY INDEX AND PER-CAPITA WALK TRIPS

Table D.1 Walk Score[2] Compared with Walkability Index and Walk Trips per Person per Week

BG ID	*Walk Score*	*Walk Score Category*	*Walkability Index*	*Walkability Index Category*	*Walk Trips per Person per Week*
240317017032	62	somewhat walkable	65	considerable likelihood	3.5
240317020002	63	somewhat walkable	40	moderate likelihood	5.1
240317021011	52	somewhat walkable	61	considerable likelihood	3.5
240317019001	78	very walkable	49	moderate likelihood	2.9
240317022001	63	somewhat walkable	69	considerable likelihood	5.2
240317017041	54	somewhat walkable	60	considerable likelihood	6.1
240317021014	52	somewhat walkable	55	moderate likelihood	3.1
240317023012	82	very walkable	84	high likelihood	5.0
240317017033	55	somewhat walkable	65	considerable likelihood	2.7
240317017021	67	somewhat walkable	62	considerable likelihood	6.1
240317020001	54	somewhat walkable	48	moderate likelihood	2.6
240317020003	55	somewhat walkable	68	considerable likelihood	6.4
240317023011	75	very walkable	70	considerable likelihood	4.4
240317019002	68	somewhat walkable	48	moderate likelihood	3.5
240317022004	60	somewhat walkable	40	moderate likelihood	2.5
240317021012	42	car-dependent	40	moderate likelihood	2.1
240317019003	63	somewhat walkable	73	considerable likelihood	6.5

[2] A point estimate at the given block centroid.

240317017031	71	very walkable	62	considerable likelihood	4.7
240317023021	57	somewhat walkable	43	moderate likelihood	2.0

Table D.1 Walk Score Compared with Walkability Index and Walk Trips per Person per Week (Continued)

BG ID	*Walk Score*	*Walk Score Category*	*Walkability Index*	*Walkability Index Category*	*Walk Trips per Person per Week*
240338055001	51	somewhat walkable	46	moderate likelihood	2.4
240338055002	57	somewhat walkable	65	considerable likelihood	2.9
240338058011	46	car-dependent	23	low likelihood	1.5
240338059092	66	somewhat walkable	51	moderate likelihood	2.3
240338059041	23	very car-dependent	33	moderate likelihood	1.7
240338059042	32	car-dependent	66	considerable likelihood	1.8
240338056013	68	somewhat walkable	22	low likelihood	4.5
240338056022	56	somewhat walkable	2	very low likelihood	1.4
240338058021	48	car-dependent	26	low likelihood	1.4
240338059091	46	car-dependent	77	considerable likelihood	7.3
240338057002	42	car-dependent	8	very low likelihood	1.7
240338057001	55	somewhat walkable	45	moderate likelihood	2.1
240338057003	51	somewhat walkable	49	moderate likelihood	2.5
240338056021	60	somewhat walkable	40	moderate likelihood	2.4
240338056012	52	somewhat walkable	46	moderate likelihood	2.1
240338056011	65	somewhat walkable	26	low likelihood	2.6

APPENDIX E: NAVIGATING THE ARCGIS ONLINE MAP DISPLAY FOR THE NMT PHASE II PROJECT

As stated in section 5.3.1 of this report, additional data layers that were produced for the entire Modeling Region used in this study, as well as data layers related to the focused Proposed Purple-Line Study Area only, are also provided within the ArcGIS Online platform. To view these additional data layers the user must first select “Click here to view more layers” on the openning screen of the NMT-SPT. The user will then be directed to the ArcGIS Online website where they will be prompted to log in. Note that, currently users are requiared to be authorized by the FHWA (i.e., having an account set up) in order to access these data layers under the FHWA ArcGIS Online platform.

Once in the ArcGIS Online Non-Motorized Phase II map display website (shown in Figure E.1), the user can then navigate through the data layers in a similar fashon to the NMT-SPT.

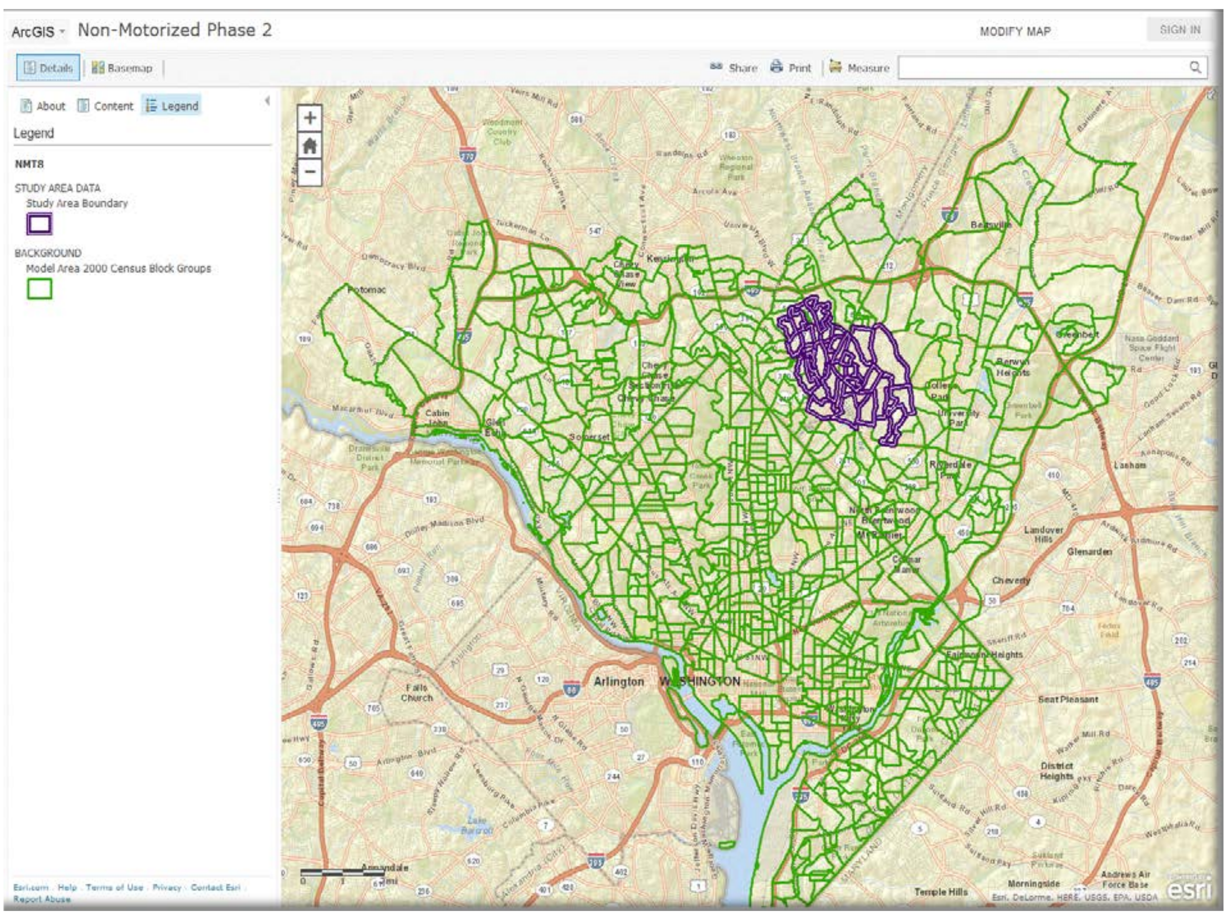


Figure E.1. ArcGIS Online Non-Motorized Phase II Website.

To see the displays of available data layers, the user must select the "***Content***" tab in the left upper corner of the screen. The user can expand the list of data layers by clicking on heading and layer names in the Contents (Note there are numerous headings that must be expanded to view all available data layers), see example screen shown in Figure E.2. The user can turn on/off a data layer, by a mouse click, to view the selected layer. However, the user must ensure that a layer's headings are also turned on when wanting to view a layer.

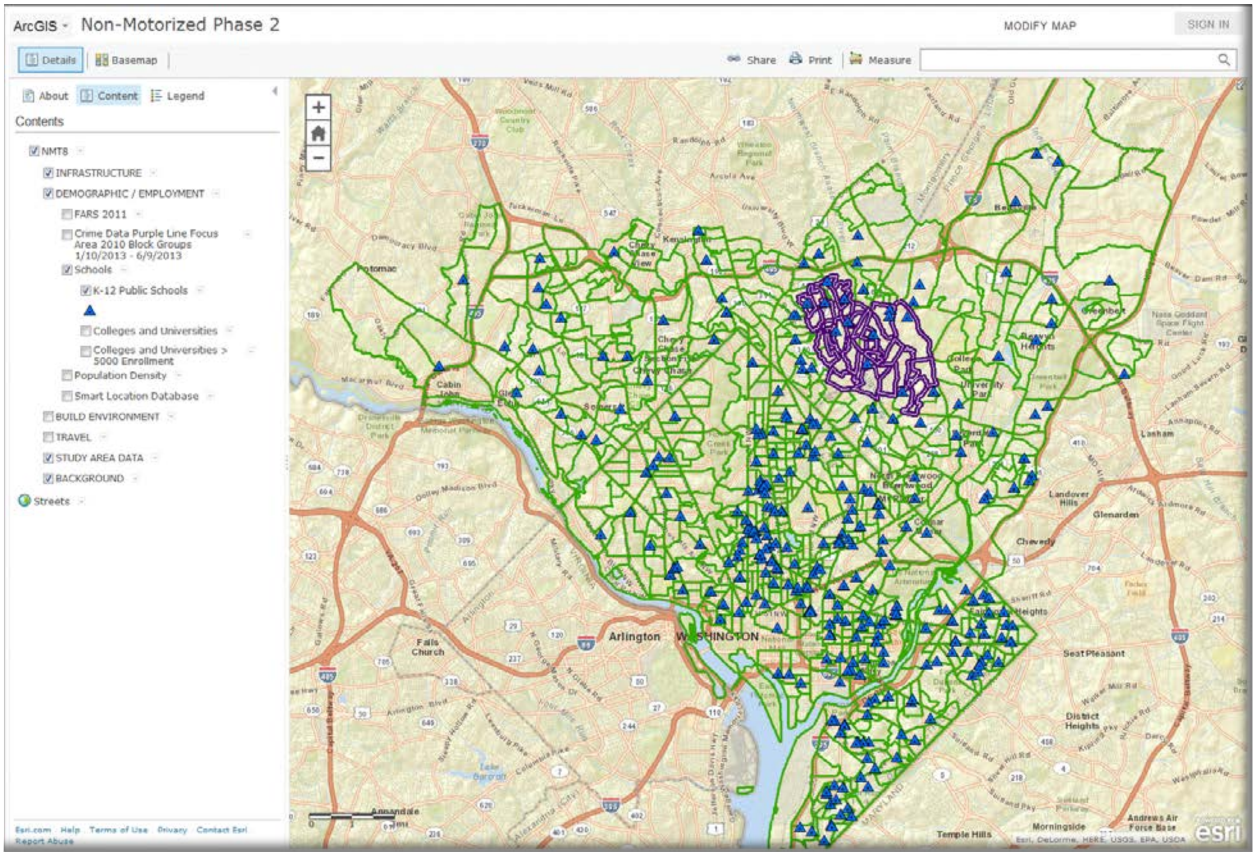


Figure E.2. Data Layer Navigation within ArcGIS Online.

Much like the NMT-SPT system, a ***Basemap*** option is provided to allow users to switch between various background base-map choices. The ***Basemap*** tab is located in the top left corner of the screen.

Navigation tools are also provided for the user (top left corner of the map), allowing the user to zoom in/out by clicking the "+" or "-"symbol respectively. The user can also zoom in/out by using their mouse wheel while hovering over the map; and can pan the map by clicking in a location on the map and moving their mouse. ArcGIS Online also provides tools to "Share" the map with others, as well as to "Print," "Measure," and "Search." These tools are all provided along the top left portion of the screen.