Draft Technical Report

EPA's Smart Location Database: A National Dataset for Characterizing Location Sustainability and Urban Form

By Kevin Ramsey[1], Dave Theobald[2], and John Thomas[1]

[1] U.S. Environmental Protection Agency Office of Sustainable Communities
[2] Colorado State University

Last updated: October 26, 2012

Abstract:
There is a growing demand for data products and tools that enable users to consistently compare multiple places based on their location sustainability and urban form. EPA's Smart Location Database (SLD) is intended to help address this crucial need. It characterizes every Census 2000 block group in the U.S. using several variables which are demonstrated in the transportation research literature to have an effect on residential travel behavior. These variables are all related to factors known as the "five Ds" (Cervero and Kockelman 1997; Ewing et al. 2007): *density* (of population, housing, or jobs), land use *diversity*, urban *design*, *destination* accessibility, and *distance* to transit. While this database is still undergoing refinement and quality assurance testing, it is freely available for public use. This report documents each of the variables included in the SLD, discusses limitations that must be considered before use, and describes ongoing efforts to improve or expand the SLD.

Acknowledgements:
The original version of this report and the SLD was created by Dave Theobald and his colleagues at CSU. Thanks to Aaron Poresky, Alex Bedig, and their colleagues at Geosyntec Consultants for their assistance developing data used to further enhance the SLD. Appendix A was authored by Aaron and Alex. Finally, thanks to Jeremy Borrego who helped bring in the 2010 Census data and assess protected area datasets.

Contents

1. Introduction

Studies indicate that the location of new homes and businesses can have an enormous effect on the travel behavior of people who will live and work there. For example, buildings in walkable communities that are well-served by transit can support a far greater variety of transportation choices than those in more automobile-dependent locations. This is particularly important when considering energy use. The energy consumed in transportation by building residents, employees and/or visitors can potentially dwarf the impact of energy saving technologies adopted inside a building. In 2008, the transportation sector was responsible for 28% of US energy consumption, compared to just 11% for residential and commercial buildings (USEIA 2009).

A growing body of research has demonstrated that "location sustainability" and urban form can have a measurable effect on travel behavior. However there has been no consistent national database that measures factors associated with location sustainability and urban form at the local (sub-Census tract) scale. As a result, it can be quite challenging to consistently compare alternative locations based on their location sustainability. EPA sought to address this crucial need through the development of the Smart Location Database (SLD). This spatial database characterizes location sustainability and urban form at the Census 2000 block group level. This database is both national in coverage and allows for regional contextualization. It is hoped that this public database can be of use to the research community and also be integrated into simple tools that enable the proliferation of location sustainability analysis.[1]

The framework used to develop this database is grounded in the latest developments in land use and transportation research. Ewing and Cervero's (2001; 2010) meta-studies of empirical research on land use and transportation behavior have resulted in quantitative estimates of the relationship between several "*D"* variables and travel behavior. The "D" variables describe several aspects of urban form, most notably: *density* (people, housing, jobs), *diversity* (mix of land uses), *design* (street network), *destination* (accessibility) and *distance* (to transit) (Cervero and Kockelman 1997; Ewing et al. 2007). Destination accessibility captures regional-scale patterns of movement and has been found to be the single most important aspect of urban form in shaping overall vehicle travel (Ewing et al. 2007). We used the "five Ds" framework to develop a national database to characterize location sustainability through metrics of urban form. Table 1 provides a summary of the "five Ds" framework and factors.

This database includes variables that are meant to characterize location sustainability from both the workplace location (employer) and residential perspectives. This distinction is most notable in the destination accessibility metrics (*D5*). Whereby residential travel behavior studies focus on accessibility of residential locations to jobs, our database also characterizes the accessibility of a location to working-age population. This later variable may be of particular interest to those who wish to understand trade-offs in siting new facilities or employment locations.

[1] Note that although we do not explore them in this report, there are likely additional uses of these datasets to assist in examining non-transportation residential energy use as identified by Ewing and Rong (2008): electric transmission and distribution losses, energy requirements of different housing stocks, and space heating/cooling associated with urban heat islands.

Table 1. Typical factors and metrics used in the "five Ds" framework to characterize urban form and location sustainability based on Cervero and Kockelman (1997) and Ewing et al. (2007).

Context	Factor	Typical Metrics	Metrics included in the SLD
Local/ neighbor-hood	*Density*	• Household/population density • Job density	• Housing units per acre • Population per acre • Jobs per acre
	Land use *diversity*	• Entropy index of land use mix • Jobs-housing balance	• Entropy index of land use mix
	Pedestrian-friendly urban *design*	• Density of street intersections • % 4-way intersections	• Intersection density
	Distance to transit	• Distance to transit stop • Transit station/stop density	• Number of fixed-guideway transit stops within ¼ mile • Number of fixed-guideway transit stops within ½ mile
Regional	*Destination* accessibility	• Job accessibility by auto • Job accessibility by transit • Distance to downtown	• Accessibility to jobs by auto • Accessibility to working-age population by auto • Accessibility to jobs by transit • Accessibility to working-age population by transit

It is important to note here that the 5D conceptual framework was derived through studies of *residential* travel behavior. In developing variables that reflect an employer perspective, we are extending this framework beyond its initially intended scope. To date, there is limited research that examines the relationship between employment location characteristics and employees' travel behavior. Ongoing analysis and literature review is being conducted to refine our understanding of the relative importance of the D factors in explaining employee travel behavior and/or VMT.

Finally, work has begun to develop a Smart Location Index (SLI) for both residential locations and employment locations. This index would score Census 2000 block groups based on their overall location sustainability relative to a regional mean. SLI values are not yet available in this release of the database.

2. Methods

2.1 Developing measures of urban form and location sustainability

Density (D1)

The SLD includes three different density measures: housing density, population density, and employment density. Housing density is distinguished from population density in that population estimates (based on primary residence) often belie the magnitude of activity, particularly in areas of temporary, second-home, or tourist/vacation resorts. We measured housing density directly using housing unit counts from the 2010 Census at the block group level[2].

$$D1_a = u/\text{ac}$$

where $D1_a$ is housing density in units per acre (*ac*), and *u* is housing units in 2010. Additionally, we also computed population density ($D1_b$) and employment density ($D1_c$).

When calculating gross housing and employment density, efforts were made to refine the metric to reflect density in areas where development can occur. For instance, a block group that includes a large city park would appear to be much lower in density than it actually is if the area of the park is included in the analysis. Calculating the total area of the block group *excluding* the park provides a more realistic measure of density. Therefore two national datasets representing public and protected lands were analyzed to estimate the total land area of each block group that is privately-owned and unprotected from development:

- **Navteq** land use (Navteq, 2011) – Local, state, and regional parks; animal parks (zoos); cemeteries; and beaches;
- **Navteq** water features (Navteq, 2011)
- **Protected Areas Dataset – US** (PADUS V1.2, April 2011): Public lands (primarily federal and state, as well as some local government) and voluntarily provided private conservation lands (e.g., Nature Conservancy Preserves or land trust easements) from authoritative data sources.[3]

There are some important limitations to the use of the PAD-US dataset to identify areas that are protected from development. For instance, some of the public lands in PAD-US are not necessarily protected from development. Many tribal lands and military bases fall into this category.[4] Likewise some of the public lands (including tribal and others) include housing and jobs. Therefore removing that land area from calculations of gross housing or employment density results in inaccuracies (sometimes quite significant inaccuracies).

Steps were taken to address the clearest examples of inaccuracies that resulted from the above limitations. After calculating housing density on privately owned unprotected land, quality checks were conducted to identify block groups with density metrics that appeared to be far outside the range of what reasonably could be expected.

---

[2] Note that the SLD uses Census 2000 block groups. All variables that use 2010 Census data are summarized by 2000 block group geographies. There are 207,507 Census 2000 block groups in the conterminous US.

[3] In the use of PAD-US for this analysis, an initial level of screening was conducted based on the "Status" field. Features that were identified as "Designated – Legally or administratively decreed" were included, while features designated as "Not Known – Current site status unknown" or "Proposed – local government level approval" were not included.

[4] In retrospect, all tribal lands and military bases should *not* have been included as protected land type, due to the fact that people live and work on these lands. We hope to address this problem in a future release of this dataset.

- Little or no privately-owned unprotected land AND the presence of homes

This check was conducted to identify block groups that are more or less completely contained by public/conservation land but none-the-less had housing development on it. In these cases it was assumed that the land is not protected from development (e.g., tribal land) and adjusted the density measures to consider the entire land area of the block group. Two separate selection rules were used to select block groups in this category:

  - <0.2 acres privately-owned unprotected land AND >1 housing unit
  - <1 acre privately-owned unprotected AND >10 housing units

- Identify block groups with high residential density AND relatively low destination accessibility

Residential density above 20 units per acre is almost exclusively found in centrally located areas of large or medium-sized cities. Block groups that are *not* in such areas can be identified using a destination accessibility metric (called "D5ae") that counts the total number or working age adults living within 30 miles of a block group (gravity weighted so that population further away are counted less). In cases where high residential density was found in areas with low destination accessibility, it was assumed that at least some of the public land is not protected from development. Therefore the density measures were adjusted to consider the entire land area of the block group.[5] Two separate selection rules were used to identify block groups in this category:

  - >20 housing units per acre AND D5ae<75,000
  - >30 housing units per acre AND D5ae<200,000

- Identify block groups in Hawaii with high residential density (>20) AND quite large in total land area (>300 acres)

As noted above, destination accessibility metrics were not available for Hawaii. 65 total block groups on the island of Oahu were identified as high in residential density. Upon inspection, many of these were in the downtown area were such densities would be expected. Isolating only those block groups that were over 300 acres in area identified places that, upon spot checking, clearly did not have any residential density. Generally these block groups were on military bases where much of the land area is publicly owned. For all of these block groups, the density measures were adjusted to consider the entire land area of the block group.

- Identify block groups with very high residential density (>50) AND large in total land area (>200 acres)

Finally, a review of the density measures in a few very large metropolitan regions outside of HI revealed isolated examples of very high residential density in large block groups far outside of city centers. Spot checking a few of these revealed that this unexpected result was due to a large portion of the total acreage being publically owned (e.g., a military base), yet homes are located in the publically owned area. This selection rule appeared to catch most, if not all, of these problems. For all block groups selected, the density measures were adjusted to consider the entire land area of the block group.

[5] Note that destination accessibility metrics were not available for Alaska and Hawaii. The few block groups in AK were spot checked to confirm that they too were in low accessibility areas. Block groups in Hawaii were analyzed separately.

These criteria resulted in identification of approximately 420 CBGs (0.2% of total) for which the unprotected area was deemed to be unreliable. In these cases, the total land area of the CBG (LAND_AC) was used instead of the unprotected area for the purpose of computing density metrics.

Land use diversity (D2)

The *Land Use Diversity* factor characterizes the mixture of land use types within a local (neighborhood) area, particularly residential and commercial land uses. Ideally this factor would incorporate some measure of "balance" between uses and operate on parcel-based data to better incorporate the richness of land uses. However, parcel-level data are not available for many regions of the country (let alone a consistent national parcel-level database). Therefore, we used employment and population data as a proxy for land use. Specifically, we assembled data about four different land use categories at the block group level: commercial/industrial/institutional jobs, retail jobs, recreational jobs (parks, recreation, stadiums)[6], and residential (number of housing units). We computed *D2* (*E*) using an entropy index as:

$$E = -\left\{\left[\sum_{i=1}^{n}\left(P_i * \ln P_i\right)\right] / \ln(n)\right\}$$

where P_i is the proportion of the employees in land use type i found in a block group (denominator is total employees + number of housing units), and n is the number of types ($n=4$). The value of E ranges from 0.0 (heterogenous) to 1.0 (even distribution across all types). We normalized the entropy index value by simply dividing E by the area of each block group (acres) to account for wide variability in the size of block groups.

One notable limitation of this metric is that it does not consider the range of services available within a given block group. Instead it only recognizes the balance of different broad categories (commercial, residential, industrial, and recreation).

Urban design (D3)

Note: This metric is not available for block groups in Alaska and Hawaii. Additionally, there are known reliability issues associated with this measure – resulting in unusually high values in some block groups. Use with caution—particularly in regards to high outlying values. EPA has plans to update this metric and release more reliable data in 2012.

The *Urban Design* factor characterizes the connectivity of the street network, where a well-connected network has many short links, numerous intersections, and minimal dead-ends (cul-du-sacs) – so that travel distances are minimal and route options are maximized. A traditional "gridded" road network typically has high connectivity, while a curvilinear design (typical in some suburban areas) has many cul-du-sacs and is organized in a hierarchy requiring travel on arterial or major roads for most trips (Kulash et al. 1990). This factor serves as a surrogate for the degree to which the street network facilitates (or restricts) pedestrian movement.

The most common metric to measure urban design is the density of road intersections in a local area, which is typically measured as the number of intersections per square mile (Frank et al. 2005). This analysis takes a slightly different approach. It is a measure of intersection density that is weighted by the number of road segments that form the intersection. D3 values were calculated by simply summing the end of each road segment, for each square mile. So, a 4-way intersection would count 4, a 3-way intersection 3, a two-way intersection 2, and a one-way intersection (a cul-de-sac) would count as 1.

---

[6] Job types were identified by NAICS code: retail = 44, 62; recreation = 71, 72; commercial/industrial = all other NAICS;

Note that there are **known reliability issues** associated with the values for this metric. A number of block groups have extremely high values, far in excess of nearby block groups with similar urban design characteristics. The cause of this problem is not entirely clear. Spot checking block groups with values far outside of the expected range indicates that one problem may be the inclusion of intersections between freeways, on/off ramps, and surrounding streets. So the intersection density value is unusually high in some areas with freeway on/off ramps despite the fact that walkability is often quite compromised in those locations. (This appears to be a common problem. WalkScore's new Street Smart calculator also overestimates intersection density near a particular freeway offramp in north Seattle). These findings indicate that the metric might significantly benefit from being recalculated with better attention to the type of intersections being counted. Clearly any intersections associated with freeways or off/on ramps should be removed. But also overpasses, roadways inside of parking lots, etc. It might be useful to also consider alternatives street network data that offer more detailed information regarding the presence and attributes of intersections (e.g., NavTeq).

Another limitation to this metric is that often either adjacent land uses (e.g., a park) or transportation corridor (e.g., railway line or interstate highway) can severely limit pedestrian mobility, yet are not fully accounted for in the metric. Additionally, this measure could obviously be improved through attention to issues such as sidewalk completeness, directness of pedestrian routes, bicycle pathways, etc. Unfortunately there is a lack of nationwide consistent spatial data regarding these additional themes.

Distance to fixed-guideway transit (D4)

The *Distance to Transit* factor characterizes the availability of transit within the nearby vicinity. Currently it only considers fixed-guideway[7] transit stations. Two variations on this measure are offered in the dataset. $D4_a$ measures the number of stations with ¼ mile. $D4_b$ measures the number of stations with ½ mile.

Note that transit is only a viable option if it reaches desired destinations within a reasonable period of time. Therefore the *destination accessibility via transit* ($D5_b$ and $D5_c$) metrics described below may provide a stronger measure of transit service for some purposes.

Destination accessibility (D5)

The *Destination Accessibility* factor characterizes the cumulative accessibility to jobs or workers from a given location as a function of distance, travel time, or mode. Three separate metrics are offered to get at this issue from three different perspectives. For each of these measures, we distinguish two types of accessibility: access from a location to *working-age population* [8] (employer perspective) or *jobs* (residential perspective).

Accessibility to jobs/workers by automobile ($D5_a$)

Note: $D5_a$ metrics are not currently available for Hawaii and Alaska.

This metric summarizes working-age population ($D5_{ae}$) or jobs ($D5_{ar}$) in nearby Census tracts as well as the relative accessibility of each of those tracts to the origin block group via automobile (in terms of linear distance). We think of this as a "workers-shed" ($D5_{ae}$) or "jobs-shed" ($D5_{ar}$).

[7] "Fixed-guideway" commonly refers to transit with a dedicated right-of-way that separates transit vehicles from traffic, such as rail, subways, and ferries. The dataset used for this measure also includes streetcars, trolleys, and light rail routes that may not be consistently separated from traffic or grade separated to avoid traffic signals.

[8] A person of working-age is defined as someone aged 18-64. For $D5_a$ and $D5_c$, working-age population was calculated by multiplying 2009 CBG population by the fraction of population that was working-age in 2000. For $D5_b$, working-age population was calculated by multiplying 2010 CBG population by the fraction of population that was working-age in 2000.

Accessibility A was calculated for the $D5_a$ metrics using a gravity model based on the work of Handy and Niemeier (Handy 1993; Handy and Niemeier 1997):

$$A_i = \sum_j \left[a_j / \exp\left(t_{ij} * \beta \right) \right]$$

where A_i is the cumulative accessibility in tract i, a_j is the activity measured by number of workers (employer perspective) or the number of jobs (residential perspective) in tract j, and t_{ij} is the average travel time in minutes between tracts i and j. *ß* is an impedance value for regional accessibility (Handy 1993; ß=0.1302). Separate A metrics were generated for the two perspectives: employer (A_e) and residential (A_r). This calculation only considers tracts within 30 miles of the origin tract.

Note that in calculating t_{ij}, we made two simplifying assumptions. First, to reduce the computational demand of calculating travel time on over 60,000 tracts to all neighbors within 30 miles (~39 million pairs), we simply computed distance as straight-line from centroid-to-centroid. Tract centroids were computed by weighting block centroids by the number of housing units in 2000. Future work is needed to determine workers- and jobs-sheds that consider likely travel times based on available transportation infrastructure between tracts and/or block groups as well as time lost due to traffic congestion. More work is also necessary to determine if the impedance value for regional accessibility requires refinement. Preliminary analysis of data indicates that impedance may be too low (particularly when used without attention to transportation infrastructure), resulting in index values that do not capture sufficient geographic variation in accessibility within metropolitan regions.

Accessibility to jobs/workers within a 30-minute transit/walking commute ($D5_b$)

Note: This variable is only available in 34 metropolitan regions, detailed below. Additionally, there are a handful of block groups within these regions for which data was not returned. The source of that problem is still being investigated. A more detailed description of the methodology used to calculate is provided in Appendix A.

This metric summarizes working-age population ($D5_{be}$) or jobs ($D5_{br}$) in nearby block groups that are accessible within a 30 minute transit/walking commute. Both local bus and rapid transit are considered where data is available. Values for this metric are only available in metropolitan regions served by the Mapnificent API[9]. Mapnificent is a website that generates a unique travel-shed (via transit and walking) for any geographic point of interest. Travel times are estimated based on available transit service data shared by transit agencies in the General Transit Feed Specification (GTFS)[10]. A custom script was designed to repeatedly query the Mapnificent API to generate a database of all feasible trips from origin block groups to destination block groups[11] within each metro region served. The assumed start and end point of feasible commutes is the population-weighted block group centroid.

This analysis assumes that the commuter will spend no more than 15 minutes walking as part of the 30 minute transit/walking commute. Furthermore, a transit trip is not a required element of the commute. Therefore block groups that are not served by transit will have a value for this metric that includes at least the total working-age population ($D5_{be}$) or jobs ($D5_{br}$) within the immediate block group and any nearby

[9] See http://www.mapnificent.net/docs/

[10] For details see: http://code.google.com/transit/spec/transit_feed_specification.html and http://www.citygoround.org/agencies/?public=public. Additionally, here is a listing of transit agencies whose data is not yet public: http://www.citygoround.org/agencies/?public=no_public

[11] Origin and destination points were set at the population weighted centroid based on Census 2000 data.

block groups that can be reached within a 15 minute walk. However values are *not* offered for block groups that cannot access any other block group via transit and walking.

This metric has some important limitations to keep in mind. It does not consider available transportation or pedestrian infrastructure in the calculation of walking travel time. Therefore major obstacles to pedestrian mobility will be ignored by this metric. Furthermore, transit travel times are only roughly estimated. A detailed discussion of methodology and accuracy is offered on the Mapnificent author's website.[12] Finally, in some regions served by Mapnificent, not all transit service providers have shared their data in GTFS format. Therefore transit travel time metrics in those regions only consider travel via those transit providers who do share data.

Finally, in some regions values were not returned for a small number of block groups that are clearly within a 15 minute walk of neighboring block groups. The source of this problem is still being investigated and an updated dataset will be circulated if the problem is resolved.

Table 2. Metropolitan regions for which $D5_b$ data is available (see Appendix A for detailed description of area covered)

Albany, NY	Eugene, OR	Portland, OR
Ann Arbor, MI	Fayetteville, AK	Rochester, NY
Austin, TX	Hampton Roads, VA	Sacramento, CA
Blacksburg, VA	Houston, TX	Salt Lake City, UT
Boston, MA	Kansas City, MO	San Francisco Bay Area, CA
Chicago, IL	Las Vegas, NV	San Diego, CA
Cleveland, OH	Los Angeles, CA	Sarasota, FL
Columbus, OH	Madison, WI	St. Louis, MO
Colorado Springs, CO	Milwaukee, WI	Spokane, WA
Dallas, TX	Minneapolis, MN	Tampa, FL
Delaware Region, DE	Nashville, TN	Washington, D.C.
Detroit, MI		

Accessibility to jobs/workers by fixed-guideway transit ($D5_c$)

While this metric also measures accessibility via transit, it does so in a fundamentally different way than $D5_b$ and it is based on an entirely different set of data. Therefore the values for this metric should not be considered as directly comparable to those in $D5_b$, despite the fact that the unit of measurement (working-age population or jobs) is the same as $D5_b$.

This metric takes advantage of a complete national dataset of fixed-guideway transit station locations that was developed in 2010 as part of the TOD Database. This database does *not* include any local bus service. Rather it focuses only on existing commuter station locations for heavy rail, light rail, commuter rail, streetcars, bus rapid transit (with dedicated right of way), cable cars, and ferries. There is some overlap with the transit service covered in $D5_b$, but this database include stations locations in 13 regions not covered by Mapnificent, while Mapnificent includes transit service information in several regions not covered by TOD Database.

Rather than considering travel time via transit, this measure summarizes the total number of jobs (or working-age population) that are accessible via transit from the origin census block group. It assumes that all fixed-guideway transit stations within the same metropolitan region (as defined in the TOD database) are functionally integrated and does not consider the distance between stations.

[12] See http://blog.stefanwehrmeyer.com/post/1448498820/a-mapnificent-world

Two separate metrics were calculated. $D5_{ce}$ reflects the employer perspective by measuring accessibility to working age population. It is calculated only for block groups that have a fixed guideway transit station within ¼ mile of the origin block group population weighted centroid. The value for $D5_{ce}$ is the total working age population (residents age 18-64) within ½ mile of all stations in the entire regional transit system.[13]

$D5_{cr}$ reflects the residential perspective by measuring accessibility to jobs. It is calculated only for block groups that have a fixed guideway transit station within ½ mile of the origin block group population weighted centroid. The value for $D5_{cr}$ is the total number of jobs within ¼ mile of all stations in the entire regional transit system.

We differentiated the threshold walking distance (¼ vs. ½ mile) to reflect findings in the research literature which indicate people are willing to walk further from their home to a transit stop than they are to walk from a transit stop to their work location. Because the block group units are relatively large for computing relatively small distances, we computed the distance from transit stations in a raster GIS (at 90 meter or 0.05 mile resolution) and then calculated the number of accessible jobs (or working-age population) of each block group using the proportion of the block group within the ¼ (or ½) mile buffer computed using the raster calculator.

As noted above, one major drawback of this measure is that it lacks any consideration of travel time or convenience. A block group near a single station downtown is scored the same as a block group near one station in a commuter suburb. However in most cases, the centrally located station is more accessible (in terms of travel time and convenience) to more people throughout the region than would be the peripheral one. $D5_b$ addresses this problem as well as the problem of leaving out local bus service. Therefore we may consider phasing out the $D5_c$ measure and focus instead on continuing to refine and update $D5_b$ as more and more transit service providers share their data in GTFS format.

2.2 Notes for interpreting flags in the SLD geodatabase

There are three different flags in the SLD geodatabase for drawing attention to instances where variables are calculated in a way that differs from the standard methodologies described above. The field named "Flag1" refers to acreage and density (D1) metrics. "Flag2" refers to employment counts used in calculating $D1_c$. Each flag is described in more detail here.

Flag1: Notes on density calculations

Two different nuances regarding acreage and density calculations are captured in this flag. First, the note: "**Ac_Land based on Census 2000 total area**" refers to the source of data for determining the total land area of the block group. For the majority of block groups the variable Ac_Land (total land area in acres) is derived from the ALAND00 variable released with the Census TIGER 2010 boundaries (which includes an update to Census 2000 block group boundaries). For some reason, some block group ids included in the original Census 2000 block groups were not included in the 2010 update to the Census 2000 block groups. For these few block groups, land area values from the original 2000 release are included. The note "**D1 metrics based on Ac_Land**" refers to the denominator in density calculations. As noted in the discussion of the D1 metrics above, for most block groups the denominator is Ac_Unpro (land area that is privately-owned and unprotected from development). However, as noted above, there are some instances where use of Ac_Unpro results in inaccurate or misleading density values. Ac_Land is used as a backup in these instances. This note is used to flag all such instances where density metrics reflect the use of Ac_Land as the denominator.

[13] The "region" field in the TOD Database was used to define transit systems. Note that this assumes inter-connection of regional systems, which may or may not be the case.

Flag2: Notes on employment counts

Emp2008 (total employment in 2008), EmpInd08 (industrial employment in 2008), EmpCom08 (commercial employment in 2008) and $D1_c$ (employment density) are all based on Census LED data. However Census LED data is not yet available in all states. Therefore, in areas where LED data is not available, alternative employment counts are used in the calculation of $D1_c$. The note "**Alternative employment data source used for D1c**" is used to flag these block groups. See Table 2 or the SLD Data Dictionary for information about employment count data sources used for other SLD variables.

Table 3. "D" Metrics included in the Smart Location Database version 0.2

Con-text	Factor	Metric	Primary datasets
Local/neighborhood	*Density*	$D1_a$: Housing density (units per unprotected acre) in 2010	*Housing units*: Census 2010
		$D1_b$: Population density (people per unprotected acre) in 2010	*Population*: Census 2010
		$D1_c$: Job density (jobs per unprotected acre)	*Jobs:* Census LED 2008
	Land Use *Diversity*	D2: Entropy index of commercial/industrial/institutional, retail, recreational, and residential within a block group	*Jobs and housing units:* ESRI Business Demographics 2009
	Urban *Design*	D3: Intersections per sq mile (weighted by intersection type)	US Census TIGER/Line Shapefile 2009
	Distance to Transit	$D4_a$: # of fixed-guideway stations w/in ¼ mile of block group $D4_b$: # of fixed-guideway stations w/in ½ mile of block group	*Stations:* Center for Transit Oriented Development. 2009. TOD Database [online] URL: http://toddata.cnt.org/
Regional	*Destination* Accessibility	$D5_{ae}$: Total working-age population[14] within 30 miles adjusted by travel time using gravity model $D5_{ar}$: Cumulative number of jobs within 30 miles adjusted by travel time using gravity model	*Jobs/Population:* ESRI Business Demographics 2009 (by tract)
		$D5_{be}$: Total working-age population within a 30 minute transit/walking commute $D5_{br}$: Cumulative number of jobs within a 30 minute transit/walking commute	*Jobs:* Census LED 2008 *Travel time:* Derived from Mapnificent API, which is based on public GTFS data shared by individual transit service providers.
		$D5_{ce}$: Working-age population w/in ½ mile of transit stations along a transit system that is accessible within ¼ mile of home CBG $D5_{cr}$: Number of jobs w/in ¼ mile of transit stations along a transit system that is accessible within ½ mile of home CBG	*Jobs/Population*: ESRI Business Demographics 2009 *Stations*: Center for Transit Oriented Development. 2009. TOD Database [online] URL: http://toddata.cnt.org/

[14] Calculated using proportion of population in 2000 that were of working age (18-64) * population in 2010.

3. Discussion

It is important to clarify that we *do not* anticipate that the SLD variables alone can be used to accurately *predict* household or employee VMT associated with a given location. For instance there are a number of other variables known to have a significant impact on both residential and employee travel behavior. These include (but are not limited to) policies regulating parking supply and pricing, incentives (such as transit subsidies), socio-economic characteristics, the availability of services within easy walking distance, the quality (not just quantity) of transit service, quality of pedestrian and cycling infrastructure, etc. Similarly, regional-scale differences in land use and transportation infrastructure among different metropolitan areas have a significant impact on the transportation options of residents, and consequently travel behavior. Data about these factors are either unavailable in a nationally consistent format or difficult to quantify in a meaningful manner (or both). The SLD provides a consistent set of variables known to shape travel behavior, *all else being equal*. As such, SLD variables may be most useful for evaluating the *potential* of a location to support reduced VMT per resident or employee if policies and other factors are favorable to promoting alternatives to SOV travel.

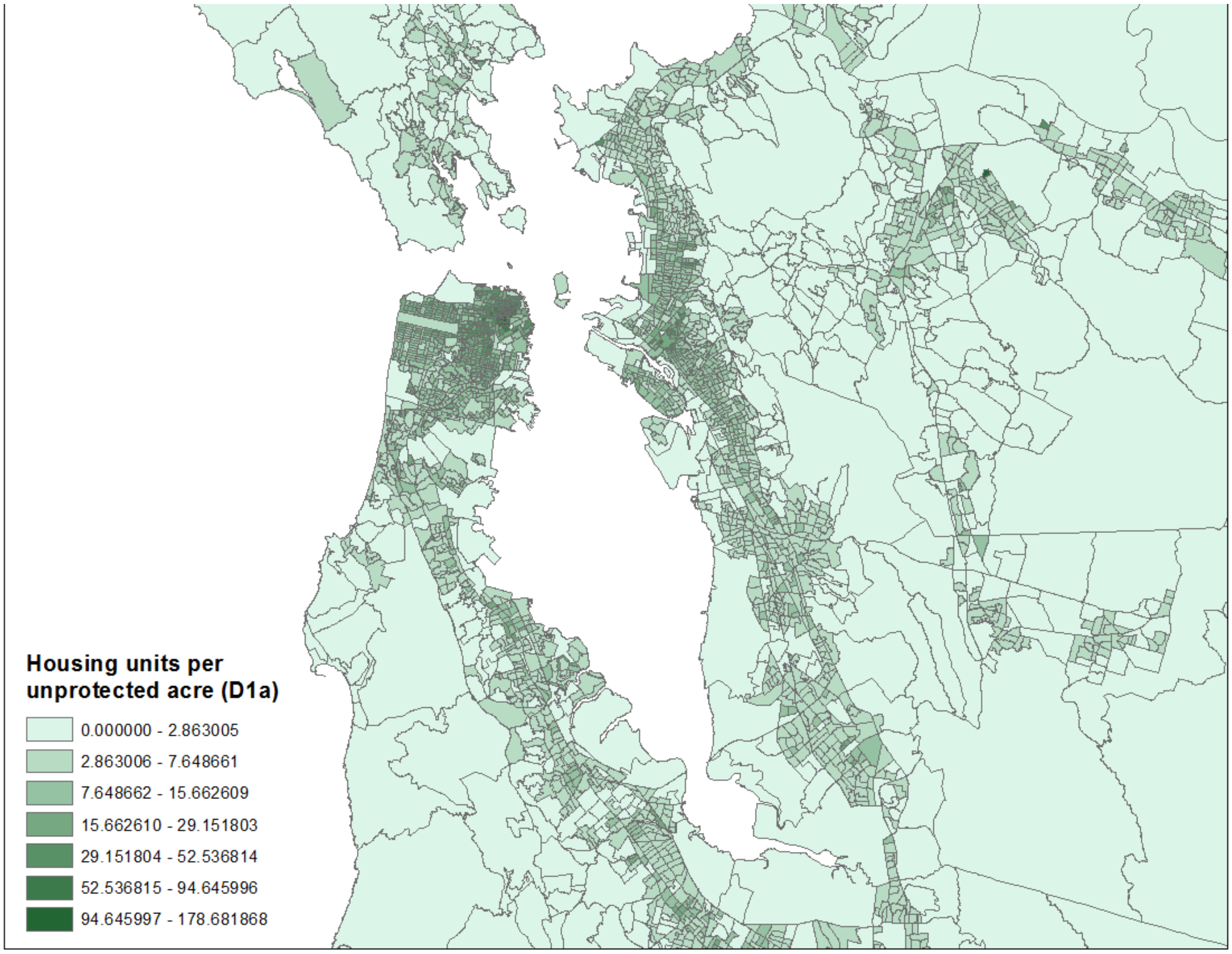
The SLD variables may be of use to inform decisions regarding sustainable employment facility location. For instance, businesses or government agencies could identify locations that would be expected to offer employees and visitors viable alternatives to automobile travel and would be more likely to reduce overall VMT compared to alternative locations. However, these measures only assess the characteristics of a Census block group *before* a new employment facility is developed. If the new facility is associated with a major redevelopment initiative, those built and environment and urban form characteristics could change significantly. Most notable of these measures is employment density, residential density, and land use diversity. Furthermore, cities often make infrastructure and transit improvements to better serve redeveloped property—improvements that could potentially affect urban design, distance to transit, and accessibility via transit. All such changes would improve a location's SLD metrics and therefore its expected sustainability benefits. Such considerations should be taken into account when using this dataset for decision support.

Finally, as noted above, our ability to develop meaningful measure the five D factors is limited by availability of datasets with nationwide coverage. For example, the land use diversity measure does not consider the range of services available within a given block group. Instead it only recognizes the balance of different broad land use categories (commercial, residential, industrial, and recreation). Many of these limitations are detailed in the previous section. Users should consider supplementing this national data with relevant local data when resources are available to do so.

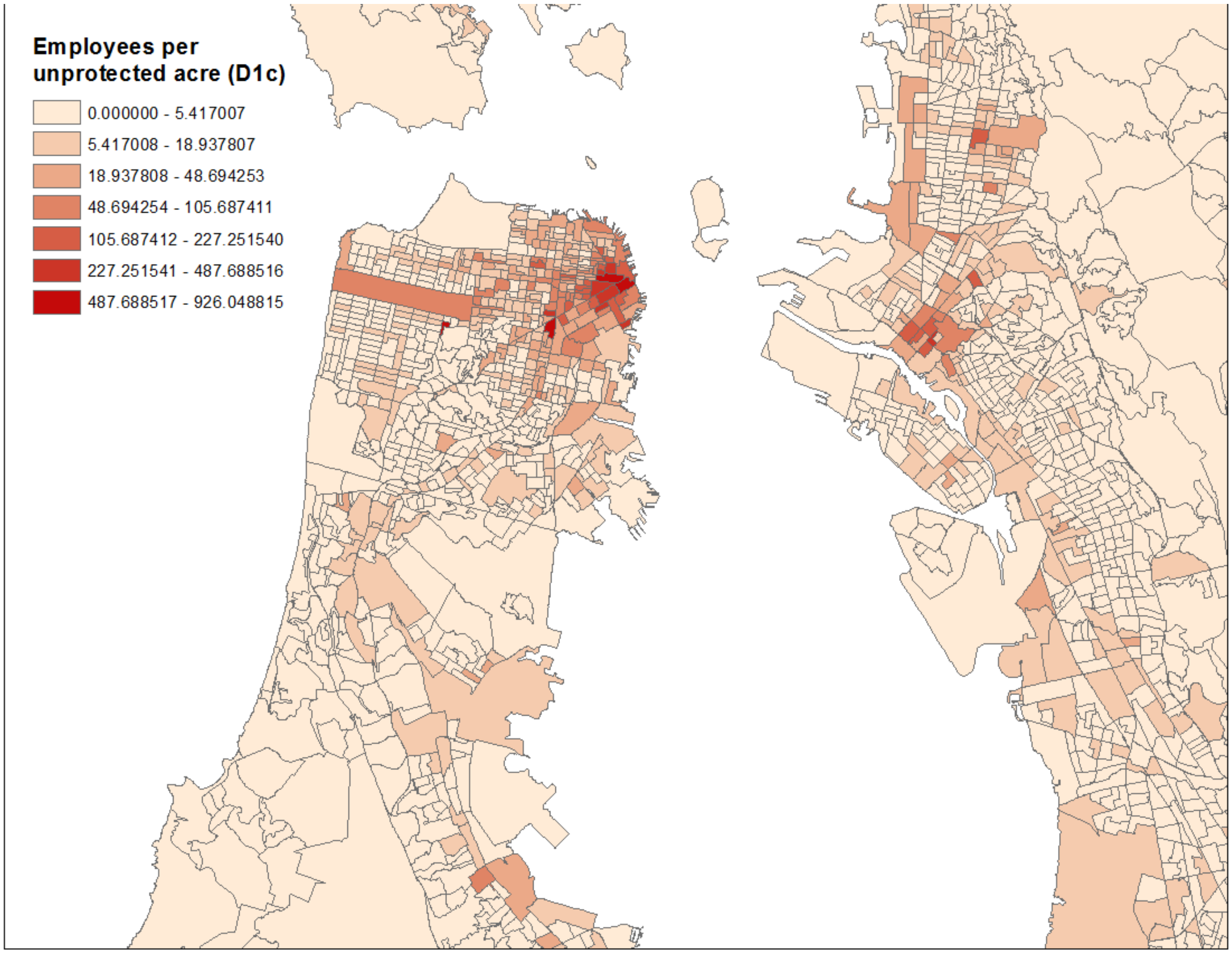
4. Map Illustrations of D Variables

We developed a single Geodatabase that includes D measures and associated attributes for every Census 2000 block group in the U.S. An accompanying data dictionary defines each variable in more detail. To illustrate the resulting measures, we include a number of example maps from the San Francisco Bay Area and Portland, OR.

Residential Density (D1a)

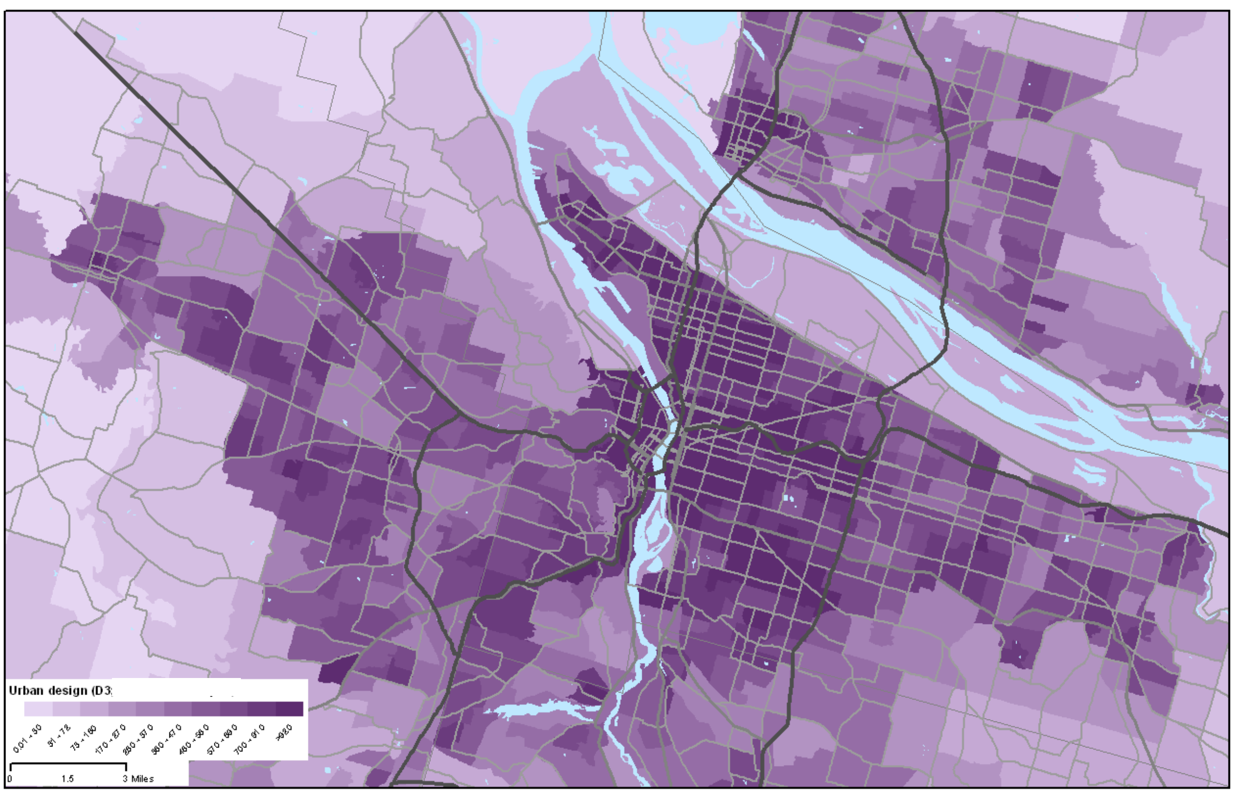


Employment Density (D1c)



Land use diversity (D2)

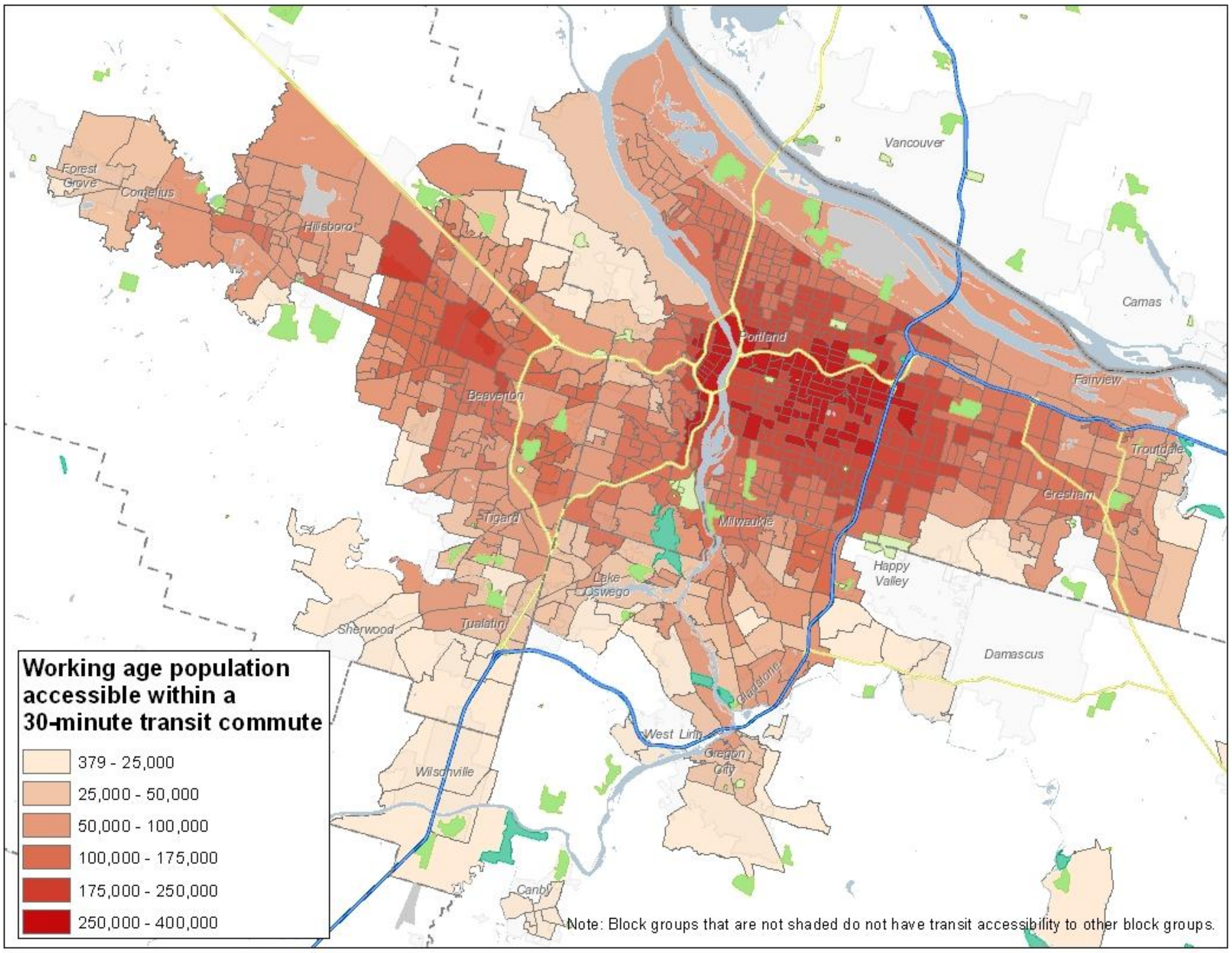
Urban design, Intersection Density (D3)



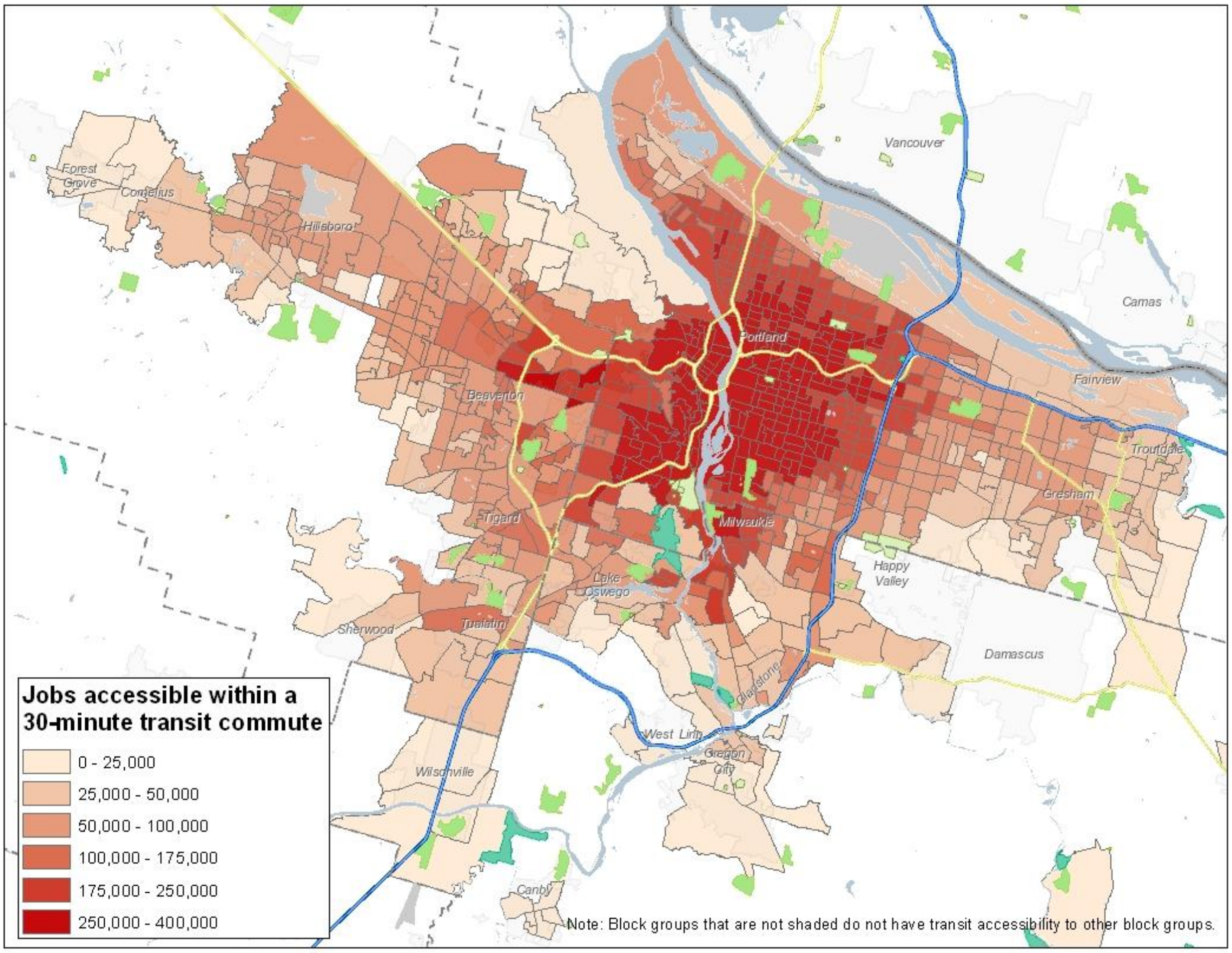
Destination accessibility, working-age population by car ($D5_{ae}$)

Destination accessibility, jobs by car ($D5_{ar}$)

Destination Accessibility, working-age population by transit ($D5_{be}$)



Destination Accessibility, jobs by transit ($D5_{br}$)



Literature Cited

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

Ewing, R. 1996. *Best Development Practices; Doing the Right Thing and Making Money at the Same Time*, Planners Press (www.planning.org), 1996.

Ewing, R., and R. Cervero. 2001. Travel and the Built Environment: A Synthesis. *Transportation Research Record* 1780: 87-114.

Ewing, R., and R. Cervero. 2010. Travel and the built environment. *Journal of the American Planning Association* 76(3): 265-294.

Ewing, R., K. Bartholomew, S. Winkelman, J. Walters, and D. Chen. 2007. *Growing cooler: The evidence on urban development and climate change*. Urban Land Institute.

Ewing, R. and F. Rong. 2008. The impact of urban form on US Residential energy use. *Housing Policy Debate* 19(1): 1-30.

Frank, L.D., T.L. Schmid, J.F. Sallis, J. Chapman, and B.E. Saelens. 2005. Linking objectively measured physical activity with objectively measured urban form. *American Journal of Preventative Medicine* 28: 117-125.

Galster, G., R. Hanson, H. Wolman, and S. Coleman. 2000. *Wrestling sprawl to the ground: defining and measuring an elusive concept.* Report for Fannie Mae Foundation, Washington, D.C., USA.

Glaeser, E.L. and M. Kahn. 2008. The greenness of cities. John F. Kennedy School of Government, Harvard University Policy Brief

Handy, S. 1993. Regional versus local accessibility: implications for nonwork travel. *Transportation Research Record* 1400: 58-66.

Handy, S.L. and D.A. Niemeier. 1997. Measuring accessibility: an exploration of issues and alternatives. *Environment and Planning A* 29: 1175-1194.

Hu, P.S. and T.R. Reuscher. 2004. National Household Travel Survey. US Department of Transportation, Federal Highway Administration. http://nhts.ornl.gov/2001/pub/STT.pdf.

Kulash, W., J. Anglin, and D. Marks. 1990. Traditional neighborhood development: will the traffic work? Development 21(July/August): 21-24.

Lewis, T.G. 2009. *Network Science: Theory and applications*. John Wiley & Sons. 512 pgs.

Song, Y., and G. J. Knaap. 2004. Measuring urban form: is Portland winning the war on sprawl? *Journal of the American Planning Association* 70 (2): 210–225.

Theobald, D.M. 2005. Landscape patterns of exurban growth in the USA from 1980 to 2020. *Ecology and Society* 10(1): 32. [online] URL: http://www.ecologyandsociety.org/vol10/iss1/art32/.

Theobald, D.M. In preparation. A national land use dataset for the US.

US Energy Information Administration (USEIA). 2009. Annual energy review. Report No. DOE/EIA-0384(2008).

Appendix A: Methodology for calculating accessibility via transit

*This appendix is adapted from a report authored by **Aaron Poresky** and **Alex Bedig**, Geosyntec Consultants.*

This appendix describes the methodology used to develop estimates of transit accessibility: $D5_{be}$ (accessibility to working age population via transit) and $D5_{br}$ (accessibility to jobs via transit).

General Approach

The Mapnificent Application Programming Interface (API) was used to develop a database of feasible trips between the population weighted centroids of census block groups (CBG). This inventory of feasible trips was then coupled with records of employment and working age population data available CBG to compute transit accessibility metrics for each CBG.

The approach makes use of the Mapnificent API to access transit data feeds supplied by transit agencies. Feeds are supplied to Mapnificent in a standardized format known as the General Transit Feed Specification (GTFS). The Mapnificent API provides a interface for accessing these data. The Mapnificent API is implemented entirely in Javascript and is designed to run in the local browser environment.

The approach involves development of a set of questions (i.e., can I reach this point from this point) followed by query of the Mapnificent API to answer each question. For each API query, inputs include: (a) two Lat/Long coordinate pairs, each defining the centroid[15] of the origin and destination CBG, (b) a transit system (defined as a City in Mapnificent), and (c) a set of other configuration parameters such as walking speed, time of week of trip, maximum walking time to a stop and so on. All routing pairs (to/from lat/lon) were run in both directions to account for potential asymmetry of transit trips.

Input/output and data management was accomplished using a SQL Server database. This database contains the routing pairs and all configuration parameters. Every time a point is accessible from another point given a set of parameters, this is considered to be a “feasible trip” and a record is created in the database including the start point, end point, and the set of configuration parameters that were used. After completing the API queries, the database included all feasible trips within transit systems supported by the API. Specific assumptions related to the Mapnificent API queries are documented in the following subsection.

The final output from the Mapnificent API query consists of a database containing all feasible trips between CBG centroids. These feasible trips were then coupled with records of employment and working age population data by CBG to compute transit accessibility metrics.

[15] In this document, centroid always refers to the population-weighted Census 2000 Block Group centroid, as obtained from the US Census Bureau, via USEPA.

Mapnificent API Input Parameters

Transit Systems

In the Mapnificent API, transit systems are generally defined by the primary city of metropolitan regions. The Mapnificent API bounding boxes for each transit system are defined in Table 1. The API does not return valid results outside of these bounding boxes.

Table 1: Mapnificent Cities and Bounding Box Coordinates

CityName	State Abbrev.	Mapnificent Bounding Box Coordinates				API Functional?
		CityNLat	CityWLong	CitySLat	CityELong	
Boston	MA	42.80029	-71.793881	41.7266	-70.625616	Y
Kansas City	MO	39.30503	-94.829681	38.89727	-94.272018	Y
San Francisco Bay Area	CA	38.10015	-122.53867	36.97492	-121.54903	Y
Sacramento	CA	38.72253	-121.55301	38.42318	-121.18047	Y
Los Angeles	CA	34.32634	-118.86083	33.70685	-117.80931	Y
San Diego	CA	33.25689	-117.2784	32.54336	-116.18448	Y
Eugene	OR	44.22097	-123.35293	43.7868	-122.11527	Y
Salem	OR	45.31135	-123.57633	44.86203	-122.77525	Y
Portland	OR	45.63835	-123.11545	44.93884	-122.27309	Y
Seattle	WA	47.97933	-122.50673	47.18911	-121.78244	N
Spokane	WA	47.76914	-117.70576	47.47667	-117.0895	Y
Las Vegas	NV	36.30094	-115.33082	35.9676	-114.8259	Y
Salt Lake City	UT	41.52802	-112.46961	40.03124	-111.57766	Y
Denver	CO	40.22544	-105.58415	39.45688	-104.67243	N
Colorado Springs	CO	39.75292	-105.15169	38.66438	-104.68163	Y
Austin	TX	30.58679	-97.9911	30.15121	-97.554732	Y
Dallas	TX	33.08427	-97.328195	32.55905	-96.541924	Y
Houston	TX	30.05429	-95.772764	29.54129	-94.986966	Y
Fayetteville	NC	36.12753	-94.206026	36.04762	-94.143164	Y
Minneapolis	MN	45.56013	-94.157372	44.70715	-92.80558	Y
St. Louis	MO	38.83771	-90.679491	38.3337	-89.792918	Y
Madison	WI	43.15825	-89.564477	42.98829	-89.276537	Y
Milwaukee	WI	43.41104	-88.116404	42.87121	-87.849137	Y
Chicago	IL	42.58584	-88.6175	41.41833	-87.525699	Y
Nashville	TN	36.22898	-86.773584	36.16185	-86.297334	Y
Ann Arbor	MI	42.31784	-83.799816	42.21352	-83.543361	Y
Detroit	MI	42.45887	-83.333282	42.25542	-82.911599	Y
Columbus	OH	40.14755	-83.161046	39.83017	-82.768662	Y
Sarasota	FL	27.52978	-82.69469	26.94897	-82.132418	Y
Tampa	FL	28.24814	-82.73403	27.70568	-82.157672	Y
Cleveland	OH	41.63667	-81.967606	41.2279	-81.436691	Y
Blacksburg	VA	37.25608	-80.45329	37.16157	-80.39342	Y
Miami	FL	26.1225	-81.075723	24.71614	-80.050118	N

CityName	State Abbrev.	Mapnificent Bounding Box Coordinates				API Functional?
		CityNLat	CityWLong	CitySLat	CityELong	
Pittsburgh	PA	40.66895	-80.255573	40.27302	-79.707664	N
Buffalo	NY	43.17912	-79.1178	42.45582	-78.48495	N
Rochester	NY	43.29738	-77.966888	42.90382	-76.993382	Y
Washington D.C.	DC	39.19176	-77.449779	38.59877	-76.669386	Y
Baltimore	MD	39.60381	-77.961655	38.31753	-76.07318	N
Hampton Roads	VA	37.27637	-76.70829	36.70991	-75.97241	Y
Delaware Region	DE	39.83581	-75.831878	38.4494	-75.05148	Y
Philadelphia	PA	40.33375	-75.763621	39.67056	-74.754376	N
New York	NY	43.29514	-80.255573	39.36357	-71.95388	N
Albany	NY	43.29514	-74.20323	42.43669	-73.39606	Y

Y = API operational and quality reviewed
N = API not operational at time of analysis; does not return data or does not return valid data

The API for several cities was not functional at the time of publication or did not return reliable results. Results for these cities could not be provided. The issue with the majority of these cities appears to in some way relate to the fact that the Mapnificent API defines some cities as having bounding boxes that overlap with other cities. The bounding box coordinates were obtained directly from the Mapnificent.net website, and overlaps were confirmed via inspection of the Mapnificent.net user interface. Despite the API convention of loading individual transit systems as completely different entities, in cases of overlaps it was observed that only one of the overlapping transit systems returned reliable results. The other city(ies) only returned results where it (they) overlapped with the reliable city.

For example, the Baltimore and Washington D.C. bounding boxes overlap. When the API was queried for all points in the Washington D.C. system, the results that were returned were reliable and complete. However, when the API was queried for all points in the Baltimore transit system, the only results returned were those areas that overlapped with the Washington D.C bounding box. Additionally, these results did not appear reliable or complete for the region that did return results. These two systems were run completely independently with a separate transit system and an independent set of input locations.

Several other investigations were conducted to attempt to rectify this issue, but none returned reliable results. Based on behavior such as this, we can only conclude at this time that there is a glitch in the Mapnificent API that is preventing reliable results.

Cities without bounding box overlap issues, but which did not return any results include Seattle and Miami. The cause of the issue in these cities could not be identified or corrected.

Trips

Trips are defined by a given start and end point. Each start and end point represents a population-weighted centroid of a CBG. To generate potential trips to use for Mapnificent API queries, CBGs were first filtered to remove CBGs with centroids located outside of any transit system bounding box. A stored procedure, sp_CreateLocationsFromInputLocationsForCity in the database executes this transaction with the recorded bounding box of the City specified as an input parameter. From the remaining CBG centroids, a database of all potentially feasible trips was developed based on all potential permutations, both forward and backward. Trips longer than 40 miles were not considered to be potentially feasible. This is based on an 80 mph max average transit speed and 30 min trip time.

Configuration Parameters

Table 2 describes the configuration parameters assumed for the Mapnificent API queries.

Table 2: Mapnificent API Configuration Parameters

Parameter	Unit	Description	Allowable Range	Selected for Analysis
Walking speed	Seconds/km	The speed it is assumed commuters walk to and from transit stations.	400-4000	720
Bicycles allowed?	Boolean	Whether or not commuters use bicycles to move to and from transit stations	True/False	False
Bicycling speed	Seconds/km	The speed it is assumed commuters bicycle to and from transit stations.	200-4000	NA
Maximum walking/biking time	Seconds	The maximum allowed non-transit commute to a transit stop	0 - n	900[16]
Maximum travel time	Minutes	The maximum allowed total commute time for this trip.	0 - n	30
Transit Period	Mapnificent API Code[17,18]	The temporal period in the week during which this commute takes place.	m(0-4), a(0-4), u(0-4)	m(2) Monday (6 am to noon)

In addition to the configuration parameters, the Mapnificent API includes several default assumptions[19]:

1. Commuters time their journeys so that they do not have to wait for their first transit option.
2. When transit connections must be made (e.g. train to bus), the commuter is assumed to wait for headway/3 minutes, where headway is the time between vehicles of the upcoming transit route (in the above example, of the bus). This is a hybrid assumption that implies some level of coordination of transit vehicles; possible distance traveled per time for commutes in cities that coordinate transit vehicles to minimize wait time will be underestimated.
3. Successful trips are possible without involving a transit trip in cases where CBG centroids are separated by less than 15 minutes walking radius (approximately 0.78 miles).

[16] While 900 seconds (15 minutes) is somewhat longer than typically allowed for walking time in accessibility calculations (≈10 minutes), this value is increased from typical values to reduce artificial discontinuities in results. Interim observations in this analysis noted that artificial discontinuities may arise when the distance between a CBG centroid to the nearest transit stop is slightly higher than the allowable walk time. If a rider were to walk slightly further to reach the stop, much greater accessibility would be possible. This assumption effectively provides a 5 minute walking buffer to account for errors associated with the use of centroid points instead of actual population distribution and rider discretion. The overall 30 minute maximum trip time is not changed.

[17] Mapnificent divides the transit week into days, and each day into four 6-hour periods. The documented codes specify Monday (m), Saturday (a), and Sunday (u), implying that Monday is a characteristic weekday.

[18] Note, the sensitivity of this parameter is significantly reduced by the Mapnificent default assumptions of wait times and headway discussed herein. It is believed that a single assumed weekday daytime travel period yields a reliable estimate for the purpose of this analysis.

[19] http://blog.stefanwehrmeyer.com/post/1448498820/a-mapnificent-world

Accessibility Metric Summations

Based on the database of feasible trips compiled from queries of the Mapnificent API, transit accessibility metrics were generated for each CBG from both the employer/workplace perspective and the home/residential perspective. The metrics and tabulations are described below.

Transit Accessibility - Workplace Perspective. "How many working age people can reach a given work CBG within a 30 minute transit trip?"

1) Calculate working age population of each CBG by multiplying 2010 CBG population [Pop2010] by the fraction of population that is working age [p_work2000]

2) For each work CBG, sum the working age population for all home CBGs that can reach the work CBG within a 30 minute travel time. Include the working age population that resides in the work CBG.

Transit Accessibility – Home Perspective. "How many jobs can be reached from a given home CBG within a 30 minute transit trip?"

1) Tabulate jobs in each CBG [Emp2008]

2) For each home CBG, sum the number of jobs in all workplace CBGs that can be reached from the home CBG within a 30 minute travel time. Include the jobs that are located in the home CBG.

Note: No data is provided for block groups that cannot access any other block groups via transit and walking, given the assumptions and parameters described above.

Limitations

The following limitations exist with the use of the Mapnificent API as part of this methodology:

Bounding box edge effects. In some cases, the limits of bounding boxes of transit systems cut through the middle of a populated area. Generally this is believed to occur where a certain part of the populated area is served by a transit system supported by Mapnificent but the other part of the populated place is served by an unsupported transit system. The accessibility results for CBGs located close to the edges of the bounding box would tend to be biased low in these cases.

Areas not supported by API but within bounding boxes. In some cases, unsupported transit systems may exist within the bounding box of a supported transit system. In cases where the API is queried for a location that is served by the unsupported transit system, results do not reflect trips that may be possible on the unsupported transit system. For example, Providence, RI, is located within the bounding box of the Boston transit system; however the Providence transit system is not supported by the Mapnificent API. Therefore the presence of feasible trips between CBGs in the Providence area is likely significantly underrepresented in the API.

Lead time and headway assumptions. The embedded Mapnificent assumptions regarding lead time and headway may over-represent actual transit accessibility in cases where transit exists but is less frequent. This is considered reliable for a general estimate of transit accessibility when start times are not known.

System Asymmetry. The API was queried assuming Monday morning travel schedule. For systems with significant asymmetry or directionality (i.e., more frequent trips in one way during the morning period and more frequent trips in the opposite direction during the afternoon/evening), results would be somewhat different if an evening or weekend travel time was assumed instead. However, generally the most important direction and time for transit accessibility calculations is the trip from a home CBG to a work CBG in the morning (at the start of the most prevalent shift).

Working age population. Working age population is derived from the product of the 2010 population and the 2000 estimate of the percentage of the population that is working age. While the percent working age is not highly variable between CBGs, it is expected that in some cases the 2000 percentage working age would not accurately represent this percentage in 2010.

CBG centroids. The analysis uses point locations of population-weighted CBG centroids to represent the entire working age population and number of jobs in a CBG. This inherently introduces error into the analysis, particularly in CBGs that have multiple population centers where the centroid may be located in a relatively unpopulated area (and therefore not well served by public transit).

Static data. The estimates returned from the API represent a snapshot in time of transit system characteristics and systems supported as of October/November 2011. Systems characteristics will continue to change and additional systems will likely be added as time passes.

Quality of Mapnificent API. The Mapnificent API provides no warrantee for data quality. In the course of this analysis, several issues have been identified in the Mapnificent API that have resulted in removal of cities from the study area. This indicates that there are errors present in the API, and may be others that were not detected. While significant effort has been investing in reality-checking results against other transit information and knowledge of transit systems, it is not possible to rule out the possibility that quality issues exist with API function that have not been identified.

Walking radius anomalies. If adjacent CBG centroids are within a 15 minute walking radius of each other (as the crow flies), a successful trip should be returned. However in some cases such as this, no successful trips were returned. This is hypothesized to be a function of numerical simplifications made in the API to describe walking radius, and may also influence the ability of these CBGs to reach a transit stop. This may bias accessibility downward for some CBGs where the walking distance from a CBG centroid to the nearest CBG centroid or nearest transit stop is relatively far. However, an offsetting bias may be in effect if the numerical simplifications also tend to over-predict walk radius in some cases. Also, because the analysis was run with a maximum walking radius of 15 minutes as the crow flies, which is somewhat longer than typically used in accessibility calculations, the net effect may not be substantial. While this may influence results for a significant subset of CBGs, results in general are still believed to provide valid approximations of transit accessibility.

CBGs without population. In rare cases, CBGs are defined by the Census Bureau that have no population, but do contain area and employment (approximately 0.3% of all CBGs). Point data for these CBGs were not present in the population-weighted CBG centroid dataset and were not included in this analysis. This has the tendency to bias accessibility low in some cases; however these CBGs occur very infrequently and are not believed to have a significant overall effect on the resulting accessibility dataset.

Quality Management

Geosyntec's quality management system (QMS) requires peer review and senior review of all deliverables, as well as development of workflows to minimize potential for quality issues. The following quality control practices were employed for this analysis:

Code review. Prior to executing queries, code was thoroughly reviewed and results manually checked against the visual display at www.mapnificent.net. Likewise, prior to executing accessibility summation queries, code and calculations were reviewed and checked manually for several locations.

Real-time plotting. While API queries were being executed, realtime visualization of results was enabled through the use of Google FusionTables. A periodic instances, API results were evaluated for reasonableness and checked against the visual display at www.mapnificent.net. A stored procedure, sp_GetAccessibleCBGsFromCBG, will return a table of the locations accessible from the specified starting CBG in all existing scenarios that can be exported as a csv from SQL Server Management Studio and uploaded directly into Google FusionTables for spatial visualization. Note this should be updated to filter to a specific CalculationHistory.CalculationHistoryID if multiple scenarios are ever present in the same database.

Post-hoc visualization. Following compilation of accessibility metrics, CBG results were plotted in a geographic information system (GIS) and checked for reasonableness.

Manual results validation and code step-through. Based on the data plotting in GIS with manual selection of accessible CBGs, manual summations were executed to spot check the end results. Additionally, summation procedures were "watched" at each step for several case studies to ensure that expected results were obtained.