Twin Cities Walking Study

Environment and Physical Activity:
GIS Protocols
Version 4.1, June 2007

Editor: Ann Forsyth

Contributors (alphabetically):

Edward D'Sousa
GIS Research Assistant

Ann Forsyth, PhD
PI, RWJF ALR Grant

Joel Koepp
GIS Research Fellow

J. Michael Oakes, PhD
Co-PI, RWJF ALR Grant

Kathryn H. Schmitz, PhD
Co-PI, RWJF ALR Grant

David Van Riper
GIS Research Fellowr

Jason Zimmerman
GIS Research Fellow

Extra material from:

Daniel Rodriguez and Yan Song
University of North Carolina

T h i s   i s   a   w o r k   i n   p r o g r e s s

TABLE OF CONTENTS

1. CONCEPTUAL ISSUES

By Ann Forsyth (major draft), Michael Oakes (statistics and sampling), Kathryn Schmitz (measurement); fact-checking by Joel Koepp and Jason Zimmerman (data); additional data descriptions by Ed D'Sousa.

Note: This chapter has had only minor revisions compared with versions 2.0 and 2.1. Additional information about data storage and sharing, however, were included in chapter 8 of version 2.2. Version 3.1 included a new section on work environments. Version 4.0 updated the chapter to include measures beyond those used in the Twin Cities Walking Study.

1.1. Protocol Purposes and Audiences

This manual provides protocols for measuring environmental variables associated with walking. It reflects work in progress. This is version 4.0 of the manual.

The manual was initially developed for a grant for the Robert Wood Johnson Foundation, and has four purposes:

- To record the methods for environmental measurement used in the Twin Cities Walking Study so that the research team can replicate its own findings at a later date and different measurement staff share common definitions and processes. Geographic Information Systems (GIS) are so multifaceted that seemingly simple decisions made at one time have ramifications for future work. Without a formal communications mechanism that makes such decisions explicit and organizes them, errors can be left unrecognized, thereby compounding measurement problems.

- To document methods for replication. The Robert Wood Johnson Foundation (RWJF), which funded the Twin Cities Walking Study, expressed interest in documenting methods well enough to have at least some measurements replicated across other RWJF study teams in order to allow later meta-analysis.

- To provide a preliminary prototype for other such manuals that may be produced by different teams. It is also possible for other teams to merely modify this manual, however, with attribution.

- To make GIS research methods and data sources less opaque, particularly to public health researchers who have not used such data.

These four purposes correspond to the audiences for this manual: (1) the initial research team, (2) other RWJF grantees, (3) other research groups doing this kind of environmental measurement, and (4) researchers not particularly familiar with GIS, but interested in understanding the strengths and limitations of GIS-based measures of environmental features potentially correlated with physical activity.

Versions of the manual after version 4.0 will refer to a companion protocols manual focusing on variables not used in the initial Twin Cities Walking Study, including variables related to the food environment and to the recreational environment of youth.

The manual responds to a general problem in the literature on measuring environmental features thought to be associated with physical activity, namely, a lack of standards for how to define key measurement variables. While there are notable exceptions that we cite in particular protocols, researchers rarely provide detailed formulae for the calculations of specific variables or even their detailed approaches to conceptualization (e.g., was water included in the land area calculations?). Further, few publication venues are available for explaining the details of measurement that are highly dependent on the capacities of specific GIS software.

1.2 Organization of the Protocols Manual

The manual is organized into eight sections.

- Chapter 1, on conceptual issues, deals with the organization of the manual, how it was produced, data sources, and measurement geographies.

- Chapter 2 includes protocols for measuring items that are not, in themselves, possible measures of walkability, but that are components of such measures. An example is land area.

- Chapters 3 though 6 look, in turn, at measurements of density, pedestrian infrastructure, mixed use, and street pattern.

- Chapter 7, outlines some measures of variables related to the built environment, e.g., traffic and safety.

- Chapter 8 examines a number of crucial issues related to data storage and sharing.

- Appendices and references are at the end of the manual.

Protocols themselves take a specific format. Each one contains six main parts. At the beginning of each protocol there may also be a note stating how the protocol needs to be tested to refine it:

1. Basic Concept. A rudimentary idea, often backed up by references to the literature from which it is drawn. While limited, in journal articles it is quite common to find it as the only type of description of a variable.

> Example: gross population density, without water, is the overall residential population divided by the land area.

> In this protocol manual, this section also cites sources that discuss why this variable is important as a measure of the environment.

2. Basic Formula, Basic Definition, or Basic Procedure. A more specific formula, definition, or procedure.

> Example: Population per Unit Land Area (without water) = Persons in housing units per unit gross area, excluding water.

3. Detailed Formula or Detailed Definition. A more specific formula that includes data sources or a more detailed definition.

> Example: Population per Unit Land Area (without water) = Persons in housing units as

measured in the US Census at the block level per unit land area, excluding area of water features as measured in the Ramsey County layer or the Metropolitan Council water layer for areas outside Ramsey (with the nature of these layers explained, e.g., Ramsey County includes all visible water).

4. Comments and Explanations. Includes answers to more detailed questions that will likely arise as measurements are made. This is also the location for warnings about potential sources of error.

> Example: where census blocks are cut by the end of the measurement geography, the population is apportioned according to the percentage of land area falling inside and outside the line.

> The Ramsey County water layer includes all water visible in very high-resolution digital orthophotos, though not such features as swimming pools (see more detail in chapter 2). For the one area outside Ramsey County, we used the Metropolitan Council's 200 Water Features Layer (again, more details are in chapter 2).

5. GIS Approach. Outlines the big steps involved in making the measurement. An expert GIS user could work from this list to perform the measures and would need to do so if using software other than the ESRI ArcGIS suite of software.

6. GIS Steps. Involves detailed GIS instructions, in smaller print, using ArcGIS 8 or ArcGIS 9 and is designed to be comprehensible by someone who does not use GIS every day (e.g., a principal investigator). The study team completed them because they

- Meant the team didn't have to reinvent the wheel when going back to do measurements a second time;

- Allowed investigators and other GIS fellows to discuss how well the GIS functions actually represented the concepts being measured;

- Allowed some measurements to be conducted by lower-level users, such as research assistants (and principal investigators); and

- Allowed us to avoid some inconsistencies. When locating the centroid of crescent-shaped polygons, for example, one function in ArcGIS 8 allows it to be outside the polygon and one does not. Because we had recorded the exact function we used, we were able to check that we had been consistent in this measurement.

It should be noted that these steps often indicate one of several ways to do measurements.

In addition, after version 1.0 of this manual was completed, we were required by University policies to upgrade our GIS to ArcGIS 9. To do this we had to uninstall ArcGIS 8. Version 4.0 has been largely converted to ArcGIS 9.1, particularly where that saved significant time. Version 9.1 introduced a number of new tools including a new form of network analyst. However, at times these new tools, while saving processing time, deliver different results compared with earlier versions of ArcGIS. Where this is the case we have discussed these differences and provided recommendations about how to respond.

In some cases we have not yet updated the protocols fully and the need for updating is noted at the beginning of those protocols.

Additional Information on this Version of the Manual

This is version 4.0 of the protocols. There are a number of refinements that we will undertake in upcoming versions.

- *Links to the Literature:* These measures are based on a systematic review of the literature. The sources, uses, and limits of variables and measurement approaches have been tracked in our internal documentation. This version does not, however, reflect all this work. Later versions will be increasingly complete in this area or will refer to other recently published papers.

- *GIS Steps*: We are continually experimenting with automating more steps. Because this is a work in progress, some of the GIS steps have been tested only once by the person who wrote them. Feedback is welcomed. In addition, some were developed for ArcGIS 8 and have not yet been converted to ArcGIS 9.

1.3 The Context of the Manual: The Twin Cities Walking Study

The manual was written in the context of a research project measuring the walkability of residential environments in the environmentally diverse northern sector of the Minneapolis-St. Paul metropolitan area.

Partly in order to undertake additional environmental fieldwork in areas known to be environmentally diverse, and partly due to recruitment logistics for the people in the study, the research focused on 36 discrete *focus areas*, 0.5*0.5 miles (805*805 meters) in size, selected for variation along two dimensions hypothesized to be important for physical activity: (A) gross population density and (B) street pattern (measured as median block size). [1]The study sites were selected to represent a subset of the most extreme areas—ranking high or low on each of these two dimensions—nine sites with high density and small blocks, nine with high density and big blocks, nine with low density and small blocks, and nine with low density and big blocks. Low population density was set at below 5 persons per gross acre (12.4 per hectare) for a 0.5*0.5-mile (805*805 meter) area; high density we defined as 10 per acre (24.7 per hectare) and above. All density calculations excluded the land area covered by water, an important issue in

---

[1] Gross density was measured over a land area, i.e., excluding water, using census blocks. Where blocks were cut by the edge of a focus area, population was apportioned in proportion to their area (e.g., if 40 percent of a census block was in the focus area, it received 40 percent of its population). Median block size as calculated for census blocks that had a centroid within the focus area.

"The study sites were selected to represent a subset of the most extreme areas—ranking high or low on each of these two dimensions—nine sites with high density and small blocks, nine with high density and big blocks, nine with low density and small blocks, and nine with low density and big blocks. Of course, given North American development patterns densities are not particularly extreme in an international sense. Low population density was set at below 5 persons per gross acre for a 0.5*0.5-mile area, high density we defined as 10 per acre and above. All density calculations excluded the land area in water. Originally we had wanted to be closer to 13 persons per acre, but there were not enough candidates from which to sample, particularly for large blocks. Small median block size was defined as below 5 acres, which was related to standard block sizes in the area. Large blocks were larger than 8 acres. Areas also had to have at least 20 percent of their land area in residential land use" (Forsyth et al. 2004a).

Minnesota. Small median block size was defined as below 5 acres (2.0 hectares), which was related to standard block sizes in the area. Large blocks were larger than 8 acres (3.2 hectares). Areas also had to have at least 18.75 percent of their land area in residential land use and a population of at least 200 (paragraph adapted from Forsyth et al. 2004a).[2]

In each focus area, approximately twenty residents were recruited to wear a motion detector (an accelerometer) for seven days, complete a seven-day travel diary, have their height and weight measured, and answer a number of surveys dealing with demographic, attitudinal, and socioeconomic issues (Forsyth et al. 2004a). The final total was 717 and 20 percent had repeated the measures although there were only 715 good addresses. This is one of the first large studies to have very detailed objective physical activity measurement and very detailed environmental measurement.

Several aspects of this study are important to understand when using this manual:

- *Results yet to come:* At the time of preparing this version (3.1) of these protocols, all data collection is complete and much measurement of environmental variables but analysis is not yet complete. However, preliminary analysis indicates that many environmental features affect the reasons for physical activity (leisure, travel) rather than the overall amount of physical activity of walking.

- *Utilitarian walking*: The study focused on utilitarian walking—earlier research indicated that such walking is most affected by environment (compared with recreational walking, for example). While the study is measuring all kinds of walking and physical activity though accelerometry (motion detection), surveys, and a travel and walking diary, the environmental measures were selected through the lens of those features most likely to be associated with utilitarian walking (walking for transportation to work, errands, etc.). The main focus of the research is on the residential environment, although as data were collected on work environments, where applicable, many measures were also applied to those environments.

- *Four categories of measures:* Based on a review of earlier research, the study proposed that four kinds of variables might be associated with how much people walk: population density, pedestrian infrastructure and amenities, mixed use or destinations, and street pattern or connectivity. We are cast our net widely, however, in terms of actual environmental measurements and will add more transportation-related measures in a later version. One of the results of the study will be better knowledge about which environmental variables are most correlated with measured walking and other physical activity, and at what scales. The study is also measuring a variety of other variables correlated with walking, however, including those with a spatial component (traffic) and those that are more purely social or psychological.

- *GIS data:* The study area was selected because it had excellent GIS data. We still faced some difficulties, however, with less consistent data on the edge of the study sector, and with specific topics such as street trees and sidewalks that were not collected consistently. Specific data issues are discussed below.

- *Urban design inventory*: As RWJF funding recipients, we promised to build on an earlier round of RWJF funding to develop measurement instruments. Of the measurement grants involving original data collection, one draft inventory was available at the time of

---

[2] One additional block was cut out of the sample because its population was only 60 people.

our data collection—the inventory by a team led by Marlon Boarnet and Kristen Day of the University of California, Irvine. The Twin Cities Walking Study team modified this inventory and the two groups are currently combining inventories into what will be the Irvine-Minnesota inventory (Forsyth et al. 2004a; Day et al. 2005). This inventory is based on a sample of road segments, for focus areas and a 0.25 mile (402 meter) buffer around them, and some protocols deal with these sampled segments.

However, from version 4.0, the manual will refer to a companion protocols manual from a new study examining additional aspects of the environment related to food.

1.4 Steps in Developing the Manual

The Twin Cities Walking Study essentially consists of three overlapping teams: environmental measurement, physical activity measurement, and statistics. This manual was developed primarily by and for the environmental measurement team, though with input from and in coordination with the other two teams. In addition, it enabled the three teams to communicate clearly about environmental measurements.

The manual was developed in several steps:

- *Listing and revising the list of variables:* A list of variables was developed based on variables used in research to date. Additional variables were added related to those that should be important according to the literature, modified by what data were available. This was revised a number of times to reflect recent developments in the literature, the availability of new data, developments in our own research, and to coordinate with other RWJF grantees.

- *Creating the basic formulae*: Where variables had been used by others, these sources were reviewed and are cited. Drawing either on this literature, or working from first principles or at a more conceptual level, the equivalent of the "basic concept" and "basic formula" were developed (see section above). These two steps were mostly done by Ann Forsyth, with useful input from Kathryn Schmitz and Michael Oakes. It should be noted that this version does not contain citations for every instance of use of particular measures. We will be incorporating them incrementally in subsequent versions but cannot hope to find all uses. If readers have used some of these measures and want to be cited, please email the editor of this manual.

- *Troubleshooting the details*: At this stage the GIS technical team was brought on board to develop more detailed methods for measuring the variables, trying them out on focus and context areas as the final set of participant addresses was not yet known and many of the buffer measures could not be developed (see below for an explanation of measurement geographies). Operationalizing many of the measurements was not at all straightforward. The kinds of issues dealt with included those we outline in the "detailed formula" and "comments and explanations" sections of the protocols. The issues were discussed in weekly GIS team meetings that were partially minuted, emails, and via research logs. These modes of recording are discussed in more detail in section 1.7. While Ann Forsyth had the final say on measurement, Jason Zimmerman, Joel Koepp, and Ed D'Sousa did important work identifying problems and discussing how to operationalize measurements in GIS.

- *GIS Work:* GIS strategies were then written by Jason Zimmerman and Joel Koepp, and later refined and simplified by Ed D'Sousa, and are in two forms. The *GIS Approach* section gives an overview of tasks, and the *GIS Steps* are intended to be useable by moderately trained GIS users. Some GIS steps also come from the UNC team of Daniel Rodriguez and Yan Song.

- *Assembly and refinement*: At this stage the protocols could be assembled, edited, and refined. This version of the protocols is part of that development stage, allowing protocols to be tested operationally and reviewed to ensure that the detailed measurement steps fully reflect the concepts. The protocols can also be tested by others who may offer critiques, as well.

Note: authorship is indicated in each chapter as a way of enhancing internal coordination in the study team.

1.5 Data

Problems

There are a number of problems with available GIS data, apart from common errors (see Forsyth et al. 2006).

Geographies: Different environmental variables are measured using different geographies, e.g., parcels, blocks, street segments. Not all data are available for very small geographical areas—for example, much economic data is only released in aggregate form. It is possible, however, to have very small aggregations constructed—akin to census blocks.

Consistency: GIS data are not consistently collected across the United States or even within municipalities. Where national data are available these protocols can be used directly. Where this is not the case, care needs to be taken in transforming the data that do exist, collecting new data, or limiting inappropriate comparisons where data cannot be matched.

Purpose: Most GIS data are not collected to answer questions about physical activity and the built environment, but are, rather, collected in order to manage infrastructure investments, collect taxes, keep government records, and manage transit systems.

Accuracy: Different kinds of data are available at different levels of accuracy, many used for broad environmental monitoring and not accurate enough for research into physical activity, e.g., tree coverage data at the 1-2ha level.

Completeness: Many forms of data are only available for part of the environment—from urban design inventories using a sample to traffic counts only available for major roads.

Time: Data are complicated to collect and are, therefore, only available for certain times, so it is common for parcel data, roads, and aerial photos to come from slightly different times and thus be slightly different to each other.

Finally, where data are collected, they are not always freely available. Conducting a study such as the Twin Cities Walking Study involved a number of separate data licensing agreements and thousands of dollars in direct purchasing costs.

Specific data sources that we used included:

US Census

www.census.gov

Description: SF1 files at the block and block-group level and SF3 at the block-group level.

Metadata: Metadata for census variables are available in a number of parts of the census Web site. For details on downloading large SF# files see
http://www.census.gov/support/SF3ASCII.html

MetroGIS

www.metrogis.org
www.datafinder.org

Description: This is the GIS group supported by the Metropolitan Council of the Twin Cities, a state agency that does planning and analysis for seven counties at the core of the Twin Cities metropolitan area. Its data include a number of useful layers complete with metadata and accessible from http://www.datafinder.org/catalog.asp. Not all data are free and many require specific agreements to be made between the user and the data owner:

General data (useful but not as detailed as Ramsey/I-35W data)

2000 Digital Orthophotos (2 feet per pixel resolution)
Generalized Land Use, 2000
Counties, Cities and Townships
MetroGIS Regional Parcel Dataset – Current (updated quarterly) (unavailable in 2004 due to a legal dispute but made available in 2005)
Water Features from 2000 Land-Use Data (supplements Ramsey County water layer)

Parcels beyond Ramsey
Anoka County Parcel Layer
Dakota County Parcel Boundaries
Hennepin County Parcel Layer

The Lawrence Group (TLG) Data—Streets, Points of Interest, and Census
Census Geography 2000 – TLG Aligned; Blocks, Block Groups & Tracts This is the MetroGIS-endorsed version of the census data and the one we used. It has the census geographies better aligned with the TLG roads layer
TLG Landmarks – Areas of Interest
TLG Landmarks – Lines of Interest
TLG Landmarks – Points of Interest
TLG Street Centerline Edits Table
TLG Street Centerlines and Address Ranges

Buses
Bus Route Segments
Bus Routes
Bus Service
Bus Shelters

Bus Stops
Transit Centers

Other
Regional Trails – Existing and Proposed
Metadata: Most are available linked to the data finder Web site at
http://www.datafinder.org/catalog.asp

Ramsey County, Cities in Ramsey County, and I-35W Corridor Coalition

http://www.ramseygis.com/Webpages/
http://www.i35w.org/gis.html

City of New Brighton
Boulevard Trees
Sidewalks

City of Roseville
Sidewalks

City of St. Paul
Lighting (incomplete)
Park Property
Recreation Centers
Sidewalks (incomplete)
Street Trees (incomplete)

Description: We used local data from two overlapping sources. One was the I-35W Corridor Coalition, a group of seven suburban municipalities that had cooperated to produce consistent GIS data. Six of the cities were in Ramsey County. Ramsey County had created a data guide with much of the metadata. The I-35W Corridor Coalition had a GIS consultant who provided other Corridor metadata via email. In general, Ramsey County had a great deal of fairly well documented data, including recent high-resolution digital orthophotos (resolution of one foot per pixel). The I-35W Corridor Coalition, however, had some already existing Excensus data (see below) and while corridor land uses were in broader categories than Ramsey County, they were updated more regularly.

The following list gives a sense of the sources available in Ramsey County, which has produced a data guide (Roman 2001). The Coalition member in Anoka County sometimes lacked some of these data layers (although in our case, only one focus area fell in Anoka). Data include:

▪ Parcel boundaries from county assessor and surveyor records, updated monthly

▪ Physical features, such as structures and water, from aerial photography, updated early 2002

▪ Digital orthophotos in one-foot-per-pixel photos from the county surveyor's office, 2003

▪ Land-use data in six major land-use classes (agricultural; industrial; commercial; residential; exempt, which includes many public buildings; and other). Each major

classification is further subdivided into between 6 and 50 detailed subclasses. The commercial land uses, for example, include about 50 subclasses such as "office building 1-3 stories," "gas station," "automotive service station," "community shopping center," "restaurant, cafeteria, and/or bar" and "flex industrial center." Some data subclasses are further subdivided and available at even more detailed levels, for example "Dairy Queen" or "TGI Fridays."

Metadata: Some metadata are available via the *Ramsey County Data User's Handbook*, online at http://www.ramseygis.com/Webpages/Hand%20Book.html. The I-35W Corridor Coalition has separate land-use metadata (see Appendix 1A).

Excensus

http://excensus.gritechnologies.com/about.pl

Description: The employment data that we used evolved over time. At first we proposed using data from a private Twin Cities consulting firm, Excensus, that had created very small area data sets, often updated annually, compiled from state, county, and local government sources, such as Minnesota postal files and data from water utilities, drivers' licenses and vehicle registrations, the school district census, and property assessors. State employment data are also available but had to be aggregated and we selected a 100*100m (1 hectare) grid cell for aggregating data.

For some complex reasons, Excensus generally used Dun and Bradstreet data for our study, checked against other sources. We then used the original point files and did not aggregate into one-hectare cells.

Metadata: See Appendix 1B.

We also collected or mapped new data.

Digitized Street Trees, Street Lamps, and Sidewalks

Description: There were no consistent street tree, street lamp, and sidewalk data. Working from the Ramsey County digital orthophotos, we first digitized these for all study areas and then field checked them. In a few cases we had data from cities, but these also needed to be checked and updated. Digitizing was done by research assistants, but checking of all data was done by a GIS Research Fellow.

Metadata: See Appendix 1C.

Irvine-Minnesota/Modified Boarnet-Day Inventory

Description: The research team used a 160+ question inventory to collect information about urban design features and pedestrian infrastructure in 20 percent of street segments. The Twin Cities Walking Study Team modified an inventory prepared by another team of researchers from UC Irvine. More information is available in (Forsyth et al. 2004a). The two versions are currently being combined and will be available on the Active Living Research site in 2005. They are presently available in draft form at http://webfiles.uci.edu/kday/public/index.html.

Metadata: See Forsyth et al. 2004a and forthcoming publications.

infoUSA Business Listings (Bundled with ESRI ArcGIS Business Analyst)

Description: The infoUSA business listings bundled with ESRI's Business Analyst is a Geocoded, proprietary database compiled by infoUSA from a variety of sources, including Yellow and White page directories, state and county records as well as phone follow-up by infoUSA staff. The database contains both businesses and institutions such as schools, libraries and government offices. The quality and completeness of the database is difficult to assess, but in general, for categories with large numbers of business, the listing appears to have reasonable completeness, classification accuracy and geocoded position. Categories with relatively few features (such as libraries) should be reviewed and validated prior to using in any analysis. In all cases, it can be expected that 5-10% of the features may be in error. For a complete discussion of how the infoUSA data was used in the Twin Cities Walking Study and subsequent research, please see Appendix 5D.

1.6 Measurement Units

In the US there is a decision about Metric or Imperial/English units. Many people in GIS use both as we did in the study, although we migrated from Imperial to Metric. We used a 0.5*0.5-mile-square focus area largely because in the study area it matched the arterial grid. We used feet in discussing many features that had been created in feet—from parcels to sidewalks. Miles were used in many of the standard surveys used in the rest of the project, and in the travel diary adapted from the National Household Travel Survey measures were in blocks and miles.

As we increasingly used measurements involving buffering of individuals, however, the measurement geographies became so idiosyncratic that metric made as much sense to use. Version 2.0 of these protocols is the first fully metric version (we hope!)

We were also saved by some easy correspondences that indicated converting to metric meant that measurements were close to those that had been developed in Imperial/English units:

> 1 mile = 1609.3 meters (rounded to 1600 for measurements)
> 0.5 miles (half) = 804.7 meters (rounded to 800)
> 0.25 miles (quarter) = 402.3 meters (rounded to 400)
> 0.125 miles (eighth) = 201.2 meters (rounded to 200)
> 0.0625 miles (sixteenth) = 100.6 meters (rounded to 100)

1.7 Measurement Logistics

Another note concerns the organization and practices of measurement teams using these protocols. The protocols were in part designed for the Twin Cities Walking Study team's internal use. They make explicit the myriad decisions that go into even fairly simple environmental measurement. In particular,

- They allow the investigators and GIS technical staff to communicate with each other so that it is clear that measures conform to the concepts being tested.

- They also help communication among the investigators.

We hope the protocols will both allow measurements to be replicated more accurately and streamline the technical process of measurement.

Given the quite varied data availability and formats, however, it is vital that the investigators and GIS technical staff communicate about variations, as decisions that seem small and innocuous to GIS staff can alter measures in significant ways and concepts of interest to investigators may not be simple to operationalize with different GIS data. These protocols were one way of communicating about these issues; however, we also used (1) decision logs, (2) emails, and (3) minutes. An excerpt from a decision log is on the next page. Emails were the basis for some of the protocols. Minutes of weekly meetings focused on action items and major decisions and were also emailed. Emailed minutes of weekly project meetings are a standard procedure in the office.

RWJF Decision Log – Jason Zimmerman
[Note, Jason worked one day per week during this period]

Week ending:	GIS decisions made:

02.06.04   Will not try to convert New Brighton street light layer to GIS due to complexity of conversion and time lost to process. Instead, will use printout as a reference when digitizing layers from the orthophoto (data is from 1986 and would need to be carefully checked anyway).

02.13.04   [None – prepared questions/issues related to GIS variables]

02.20.04   Census blocks used for calculations if center of block (centroid) falls within the site area (both focus areas and context areas).

average census block size – selected all census blocks that had their centers in the site areas; converted area (in meters) to acres by multiplying by 0.0002471; opened table in Excel and calculated average size

median census block size – selected all census blocks that had their centers in the site areas; converted area (in meters) to acres by multiplying by 0.0002471; opened table in Excel and calculated median size

02.27.04   We decided on the following set of rules regarding **access points**:

1. The first set of access points will be created from the intersections of all roads from the TLG data layer and the boundaries of the focus and context areas.

2. The second set of access points will remove points that result from intersections with Interstates as well as on- and off-ramps. It will also remove one access point from each pair of points created by parkways and other divided at-grade roads.

3. The third set of access points will remove the points in Set Two and also remove marginal points that are just inside the edge of the site boundary, do not extend into the site, and have no other roads with access into the site within 330 feet.

Set One was created in ArcMap and the points counted automatically for both the focus and context areas. Set Two was created by manually removing the Interstate, ramp, and divided road points. For Set Three, a buffer of 330 feet was created around each of the access point. Marginal points which did not connect to a road with access to the site within this distance were removed manually.

03.05.04   Revised decision about **access points** to remove Set Three, consideration of marginal points, from the analysis. Because the site boundaries were created based on a fixed grid system, a number of access points were left just inside or just outside the site boundaries. This opened the door to many subjective decisions about inclusion. It was decided that overall the number of these access points actually counted would balance out and we would not worry about trying to classify or analyze them further.

03.12.04   Method for determining **road lengths**:

1. Select 1st pass of non-contiguous polygons

2. Clip road layer to polygons (use selected features only)

3. Recalculate road lengths with XTools

4. Perform spatial join to assign IDs to road segments (if there are road segments that are common to more than one polygon – as with the

1.8 General Measurement Geographies

There are many possible measurement geographies for analyzing the walkability of an environment. Which geographies are used depends on:

- Data availability (e.g., some data are available only in aggregated format)
- The nature of the variable (for example, if walking to a corner store is thought to be a key motivation for walking for transportation, the distance to the nearest store may be more important than the percentage of a buffer area taken up by corner stores)

Typical geographies include the following, and these protocols focus on those that are underlined. "Objective" geographies independent of experience:

- Natural neighborhood, variously defined
- Existing measurement unit, e.g., census tract, TAZ
- Large grid cell (e.g., 400-1600 meters)
- Small grid cell (e.g., 100-200 meters)
- Straight line or airline buffer (at different radii, e.g., 100, 200, 400 meters)
- Network or street distance buffer (at different radii, e.g., 100, 200, 400 meters)
- Straight line or airline distance to nearest feature
- Network or street distance to nearest feature
- Parcel of residence

Other geographies

- Perceived environment (from surveys/interviews)
- Experienced environment (from diaries, GPS)

The measurement geographies are described in more detail in chapter 2. Appendix 1D contains a matrix of measures by measurement geographies in the study.

1.9 Creating the Initial Focus/Context Area Study Geography for the Twin Cities Walking Study

This section explains the key decisions we made in creating the study geography for the Twin Cities Walking Study. This is important to include for two reasons. First, some people may wish to replicate our approach. Second, even if study teams choose a quite different strategy for selecting areas to study, understanding the scheme used in the Twin Cities Walking Study provides useful background to this list of protocols. More detail about how we did the actual work to select the focus and context areas is described in a protocol in chapter 2.

Major Decisions

Several major decisions shaped the overall study and in particular the selection of 36 0.5*0.5-mile (805*805-meter) square *focus areas* embedded in 1.0*1.0-mile (1609*1609-meter) *context areas*:

- Deciding to focus on residential environments rather than, say, work environments

- Deciding to concentrate on specific places rather than sampling people from across the metropolitan areas

- Using a grid to demarcate neighborhoods rather than using "natural" neighborhoods, individually defined neighborhoods, census geographies, etc.

- Stratifying by density and street pattern, rather than other factors, such as mixed use or pedestrian infrastructure

- Choosing which density and street pattern measures to use

The section concludes with some reflections on things we would do differently if doing it again.

Deciding to Look at the Residential Environment

Some people in the US conduct most daily activities in their neighborhoods, but more people live their lives across a metropolitan area and daily walking activities are conducted in spatially diverse locations linked by motorized transportation. This poses dilemmas for analyzing the effects of environment. People from the same residential area may have vastly different levels of physical activity, partly depending on the pedestrian orientation of the distant locations where they work, shop, socialize, and recreate (Cervero and Gorham 1995). Residential environments are important locations, however, as they are used by everyone not just paid workers.

In order to capture larger activity patterns, we used a seven-day travel diary (with format based on the National Household Travel Survey) in association with seven-days of readings from an accelerometer. Individual-level surveys also asked for the address of the work environment, allowing additional analysis of those places.

However, as we also collected work address data in the surveys, we were able to measure work environments for many participants—either work at home or away from it.

Deciding to Recruit from and Measure Specific Places Rather than Sampling People from Across the Metropolitan Area.

We decided to focus on specific areas, rather than a random sample from across a larger area, for three main reasons:

- *Finding extreme areas*: As outlined above, we looked at areas that were extreme in terms of gross population density and median block size.

- *Urban Design and GIS fieldwork logistics*: Because detailed measurements of environmental characteristics involved collecting and checking data, having 36 consolidated focus and context areas rather than a couple of hundred smaller cells made this data collection viable.

- *Recruitment and measurement*: By focusing on 36 areas we could target recruitment better and find a limited number of "neutral" places, such as community centers and libraries, in which to conduct face-to-face interviews.

Using a Grid to Demarcate Neighborhoods Rather than using "Natural" Neighborhoods, Census Geographies, etc.

Many earlier studies had examined neighborhoods defined as census tracts, or defined by such features as arterial roads. We instead created a 0.5*0.5-mile (805*805-meter) grid across the entire area. The grid helped us get around some of the measurement and analysis problems associated with the differences in size and shape of these other geographies. We also used a number of individualized geographies, however, such as buffers around individuals for actual measurements so that we were not stuck with the relatively coarse grid for all measurements. We called the 36 0.5*0.5-mile (805*805 meter) grid cells that we ultimately selected for study our "focus areas." Areas of 1.0*1.0 miles (1609*1609 meters) surrounding the focus areas were called "context areas."

There were over 500 populated cells in the study area—of these we selected 36 (see more detail in chapter 2).

We chose 805*805-meter squares as a workable size. The 0.25-mile (402-meter) and 0.5-mile (405-meter) radius from a nodal point is frequently used as a baseline for walkability. The focus and context areas reflect this scale and the specific size was chosen for three reasons related to the reasons for focusing on specific places mentioned above. This size of space bounds an area that is small enough to be the focus for a person's pedestrian activity; large enough to have a reasonable population to sample; and, for much of the study area, arterial roads are laid out on a 0.5- or 1.0-mile (805- or 1609-meter) grid. This means that census blocks and block groups will have boundaries that match these grids making census analysis easier. A central consideration is that the amount of walking recommended for health benefits is 30 minutes on most days (USDHHS 1996), which translates to about two miles (or approximately 3200 meters) per day. Since we are interested in walking for transportation in addition to walking for walking sake, however, we chose not to evaluate elements of the physical environment that would allow for a 30-minute, 3200-meter walking loop for the purpose of leisure-time exercise. We assumed that there would be multiple transportation trips per day and that the addition of some transportation trips would contribute to, but may not fulfill completely, the accumulation of 30 minutes of brisk walking on most days.

Of course as Krizek (2003a; 2003b) points out, many commercial and civic uses are located at intersections of arterials, and thus in measuring mixed use one commercial area could be divided between four focus areas. These commercial and civic areas will, however, still be within the context measurement area, and within buffers around individuals, and so the effects of these uses on walking will be considered explicitly in that analysis.

Stratifying by Density and Street Pattern, Rather than Other Factors such as Mixed Use or Pedestrian Infrastructure

We knew that there were many environmental variables that may matter in utilitarian walking. The study design selects two to stratify the sample: density (gross population density) and street pattern (median census block area). This was to maximize variation. The logic is described below.

Statistically, sampling is critical to both internal and external validity in observational designs. Consider that in order to best estimate effect of socio-environmental contexts on walking, we should randomly assign (randomly selected) people to a large number diverse contexts (or visa versa) and observe walking behavior with perfect measures. Although such an experiment is obviously not possible, the exercise reveals the importance of measurement error, variability across neighborhood contexts, and the requirement that similar people reside in each (i.e., assignment mechanisms). Without control over assignment mechanisms, researchers must appreciate that "like people" do not reside in diverse neighborhoods: the rich reside in one type of neighborhood, the poor another; whites in one, Blacks in another; urbanites in one, farmers in another. It is such social stratification that, in fact, defines neighborhoods, for the most part. This means that the background characteristics of residents (composition) are typically confounded with the contexts. It follows that our sampling goal was to maximize variability in contexts while minimizing variability in resident's background characteristics—a difficult (sampling) problem, indeed.

Conceptually, we were interested in the effects of four dimensions of context: density, street-pattern, mixed use, and pedestrian amenities or infrastructure. Units in each dimension may be classified into high, medium, and low levels but to maximize variability our interest here rests with only the high and low. We would have liked to sample neighborhoods in this space, but that would have meant that we needed data in all of the $2^4 = 16$ cells of the sampling-matrix cells. Given known levels of social stratification, there is a high probability that many of the cells will have no, or very few, observable neighborhoods (Oakes and Rossi 2003; Oakes *In Press-a*). The upshot of this confounding is that inferences are "off-support" of the data and rely on strong assumptions. To avoid this, we sampled 2-space defined on density and pattern (4 sampling-matrix cells), and measured mixed use and amenities as any other covariate, such as percentage of white, percentage of poor, etc. This plan was to strike the proper balance between internal and external validity. We may make inference in the 4-space, but only with appropriate care and caution and an eye on actual observed data. By analyzing a number of measures of mixed use and pedestrian infrastructure as variables (co-variates), rather than using them in the sampling frame, this study could still examine the association of these characteristics with walkability. Details about how the final 36 areas were chosen are described in chapter 2.

Our sampling plan may thus be characterized as a stratified cluster design. But unlike much work in survey sampling, we had direct interest in cluster (i.e., PSU or primary sampling unit)

effects. We staggered participant recruitment in each neighborhood to minimize any seasonality effects.

From the environmental perspective, the stratification ensured that there would be variation on many measures because some are correlated with the sampling frame variables (e.g., sidewalks go with small blocks), and others are randomly distributed (e.g., street trees are in most places).

Choosing Which Density and Street Pattern Measures to Use

Once density and street pattern were selected for stratification, we had to select one of the many measures of each of these. We chose gross density over, for example, the density on residential parcels, because people walking through a neighborhood experience the whole environment, not just the residential part of the environment.

We chose block size as an easy measurement of connectivity. Our major innovation here was to use median block size—which is not totally straightforward in ArcGIS, where average sizes are far easier to calculate. This was because statistically medians are a preferred measure of central tendency. In doing both measures we found significantly different distributions of means and medians.
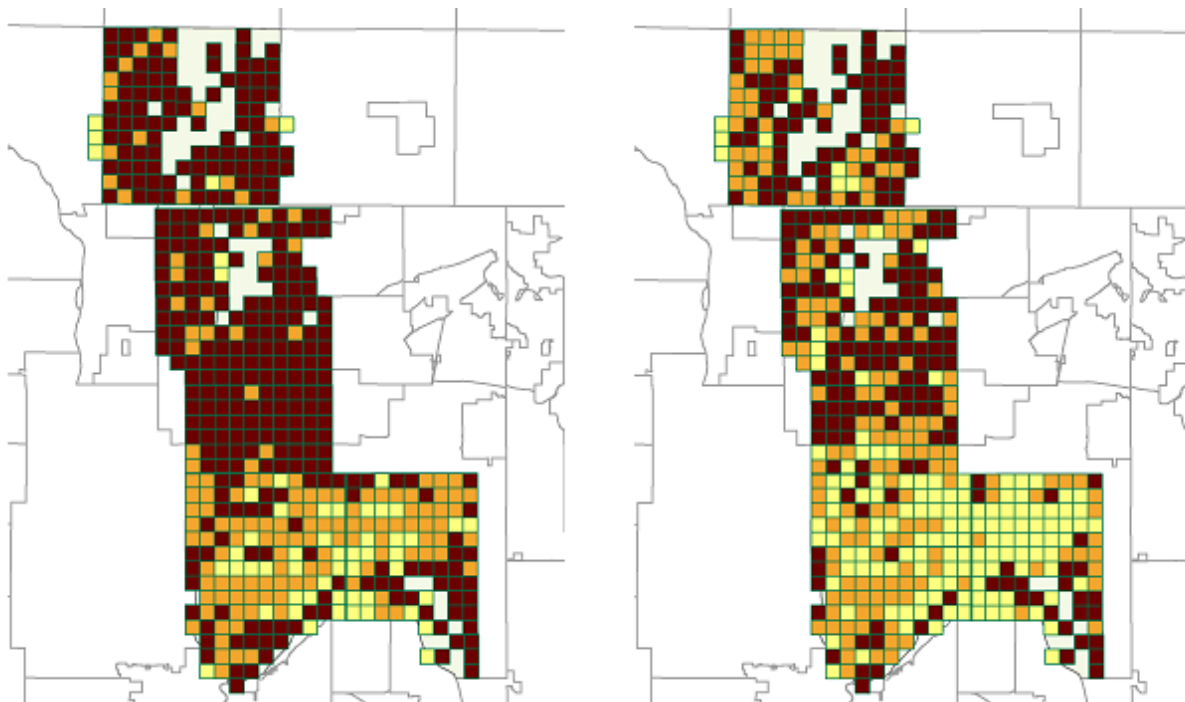
Fig. 1.1 and 1.2. Map on left shows average census block size for study area and on the right the median.
Maroon = 8+ acres (3.2+ hectares), Orange = 5-8 acres (2.0-3.2 hectares)

In particular we picked up a number of areas with mostly small blocks but one large land use (e.g., a golf course or large industrial area) that gave them a medium *average* block size but a low *median* block size. They were, thus, included in our study.

Things We Would Do Differently

There are a number of decisions that we would make differently next time.

Most importantly, we would not sample close to the edge of a county or other important data collection unit. We sampled from all 500 eligible areas, but that meant some of our sampled focus areas were right on the edge of a county, and the context area went into the next county. This added to our data collection efforts, as we had to buy parcel data and match land-use codes. While not overly onerous, it would have been simpler to avoid this.

1.10 Definitions of Terms

Context Area—1609*1609-meter area around a focus area, used for data collection.

Focus Area—805*805-meter area selected for study.

Study Area—the entire I-35W Corridor Coalition area plus St. Paul, from which 36 focus areas were selected.

Study Site—a smaller area, generally a focus or context area.

Several terms have a specific meaning in ArcGIS and are frequently used:

Summarize the data: involves creating a variety of "summary" statistics including the total sum of values but the same function can create an average, maximum, minimum, standard deviation, and variance but not a median.

Spatial join: joins map layers when their underlying attribute tables don't share a common field—such as a parcel ID or census tract number—so they are joined based on location.

Intersect or spatial intersect: "sandwiches" two distinct data layers together, creating an output file combining the extent and features common to both layers. It preserves all boundaries, shapes, and attributes from both layers, so in the output a feature from one layer will typically be "cut" into multiple pieces by the boundaries of features from the other layer.

Appendix 1E contains some other relevant information such as coordinate systems used.

2. FUNDAMENTAL PROTOCOLS AND PROCEDURES

By Ann Forsyth (drafting, conceptual issues and formulae, some comments and GIS approach text); Joel Koepp and Jason Zimmerman (GIS Steps and the rest of the comments and GIS approach text for versions up to 3.1) and Ed D'Sousa (substantial revisions for version 4.0 including notes on the need to test protocols where needed). David Van Riper updated some measures for ArcGIS versions 9.1 and 9.2.

Note: the main revision to this chapter following version 2.0 was to protocol 2.8, where steps were added; in version 2.2 additional information about sampling was added to protocol 2.3. In 3.0, the GIS steps in protocol 2.6 were substantially expanded. In versions 4.0 and 4.1, the measures were substantially updated to reflect ArcGIS version 9.1 and to streamline measurements.

Background

This chapter deals with issues fundamental either to other general measurements or to the Twin Cities Walking Study. These generally involve measurements of areas.

2.1 Net Land Area (without water)
Note: As of ArcGIS ArcInfo 9.1, the operations performed by ET GeoWizards and XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest ET GeoWizards and XTool versions are used.

1. Basic Concept
While it is easy just to measure total or gross area in a study geography (see next protocol), people generally do not walk on water. In locations with significant numbers of water bodies, a more useful version of the land area calculation excludes significant water bodies. This is, of course, particularly important in Minnesota, which is just such a place, i.e., with many water bodies, even in urban areas. In other locations, total area would be appropriate.

2. Basic Formula
Net Land Area (without water) = Area within defined geography minus water bodies.

3. Detailed Formula
Net Land Area (without water) = Area within defined geography minus water bodies as interpreted by Ramsey County and the Metropolitan Council (for the one study site we had in Anoka County).

4. Comments and Explanations
The water area was defined differently by Ramsey and Anoka counties. Fortunately, we only had one focus (study) area in Anoka, and slivers of context areas in Anoka, Hennepin, and Dakota, all without a great deal of water.

The definitions included:

Ramsey County: "This data set primarily contains open water areas including creeks, lakes and rivers. Open water areas are derived from aerial photography observation and interpretation utilizing stereo processing techniques. Vector line work was

---

originally stereo plotted from 2003 photography. All graphical lines contain a character based layer name that identifies the different mapping features. The flight height for the aerial photography utilized in this project is approximately 5,000 feet [1,524 meters]. The mapping scale for aerial photography utilized in this project is 1 inch equals 833 feet [1 meter equals 10,000 meters]. All features of this type photogrammetrically identifiable within the Ramsey County coverage are included."
http://www.datafinder.org/metadata/rams_water.htm

The Metropolitan Council's 2000 Water Features layer: "It generally includes lakes larger than 3 acres [1.2 hectares] and rivers wider than 200 feet [61.0 meters]. It may also include some smaller open water features."
http://www.datafinder.org/metadata/hydro_luse2000.htm

5. GIS Approach
The GIS approach involves using the water layer to clip the study site polygons to end up with net land area.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Water Layer** (polygon)
3. **ET GeoWizards** tool (available for free download at http://www.ian-ko.com)
4. **Xtools** extension (available from the Downloads section of the ESRI Web site)

Determine Net Land Area (without water) for each Site:
1. Add the **Study-Site Layer** and **Water Layer** to your **Map Document**.
2. Using the **ET GeoWizard** tool, remove any areas of water from the study sites (follow instructions included with the software for installation).
   a. Open **ET GeoWizard**.
   b. Under the **Geoprocessing** tab, select **Erase** and click on **GO**.
   c. Select your study site polygon layer as the layer to be erased.
   d. Give your output layer a name.
   e. Click on **Next**.
   f. Select your water polygon layer as the layer to erase with.
   g. Select either **Dense** or **Normal** data type (Dense works better for polygons with many vertices).
   h. Click on **Finish**.
   i. The new layer will be added to the **Map Document**.
3. From the **Xtools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length, Acres, and Hectares** (follow instructions included with the extension for **Xtools** installation).
   a. Under **Select layer to measure**, select your new study site polygon layer.
   b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
   c. If you do not wish to generate a parameter, such as **Perimeter**, **Length** or **Hectares**, deselect it now (leave **Area** and **Hectares** selected).
   d. Click **OK**.
4. New fields called **Area** and **Hectares** will be created in the **Attribute Table** of your site polygon layer with the areas of each polygon calculated in the units you chose (Area) and in hectares (Hectares).

2.2 Gross Area (including water)

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
Land area is one of the most basic measurements. Area, including water, is a gross measure that is appropriate for places without coastlines, lakes, or similar water features. Because the Twin Cities has considerable water, however, this is not used in these protocols.

2. Basic Formula
Gross Area (including water) = Area within defined geography.

3. Detailed Formula
Same as above.

4. Comments and Explanations
Includes water, areas not in parcels, basically everything.

5. GIS Approach
This is a plain-vanilla land area calculation.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Xtools** extension (available from the Downloads section of the ESRI Web site)

Calculate Land Area for Each Site:
1. Add the **Study-Site Layer** to your **Map Document**.
2. From the **Xtools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **Xtools** installation).
   a. Under **Select layer to measure**, select your new study site polygon layer.
   b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
   c. If you do not wish to generate a parameter, such as **Perimeter**, **Length**, or **Hectares**, deselect it now (leave **Area** and **Hectares** selected).
   d. Click **OK**.
3. New fields called **Area** and **Hectares** will be created in the **Attribute Table** of your site polygon layer with the areas of each polygon calculated in the units you chose (Area) and in hectares (Hectares).

2.3 Creating a Focus Area Sampling Grid and Sampling Extreme Areas (followed by random sampling addresses)

Note: Note: The Create Fishnet tool has not been tested with ArcGIS 9.1.

1. Basic Concept
This procedure identifies potential focus areas for the study and obtains a random sample of the most extreme areas, a strategy we used to guarantee environmental diversity.

This protocol is slightly different from others as it focuses on a procedure rather than a formula or a definition.

2. Basic Procedure
This procedure:

- Places an 805*805-meter grid over the entire study area.

  o There were a total of 780 of these. In the future, we would eliminate all grid cells on the edge of the study geography unless identical data are available outside it.

- Finds those grid cells with more than 18.75 percent (30 acres, if there is no water) of their parcel land area in residential uses.

  o There were a total of 388 of these.

- Calculates the population per hectare (without water) for each cell, with population in blocks cut by the edge of a cell apportioned according to the land area on either side of the boundary (e.g., if 40 percent of the land area is in the cell, it gets 40 percent of the block's population).

  o See chapter 3 for instructions on calculating population per hectare (without water).

- Calculates the Median Block Size for blocks with a centroid within the cell (using the modified census layer from TLG and Metro GIS. For a description, see chapter 1). Block centroids are used so that each block is counted as a whole and each block is allocated to only one grid cell.

  o See chapter 6 for instructions on calculating Median Block Size.

- Arranges grid cells into a 3*3 matrix of low (up to 5 persons per acre/12.4 per hectare), medium, and high (over 10 persons per acre/24.7 per hectare) population density and small (under 5 acres/2.0 hectares), medium, and large (over 8 acres/3.2 hectares) blocks. These cut-off points reflect the character of the study area and some of the literature on population density where 10 persons per acre (the cut-off for "high" density, was the lowest cutoff we could justify in terms of walkable areas (see Frank and Pivo 1994, 46, Hess et al. 1999, 1).

- Eliminates grids cells with fewer than 200 people.

- Randomly selects nine grid cells to study from the four extreme cells of this matrix.

- Checks that none of the cells will be undergoing large-scale redevelopment during the study period.

  o In the case of the Twin Cities Walking Study, we checked with all the relevant planning departments and two cells were each going to have hundreds of units constructed during the period of the study (causing disruption to current residents and changing the residential density and street pattern over the course of the year). We excluded these areas and re-sampled.

3. Detailed Procedure
As above.

4. Comments and Explanations
Some steps require additional explanation.

In placing a 805*805-meter grid over the study area, we adjusted it to fit the arterial grid so that for many of the grid cells the boundaries are roads. The logic behind this is explained in chapter 1, but was to do with simplifying some measurements. The actual arterial grid was offset, however, so our sampling grid needed to be broken and manipulated a little. The diagram below shows these offsets. There were some minor overlaps and gaps, but quite minor.
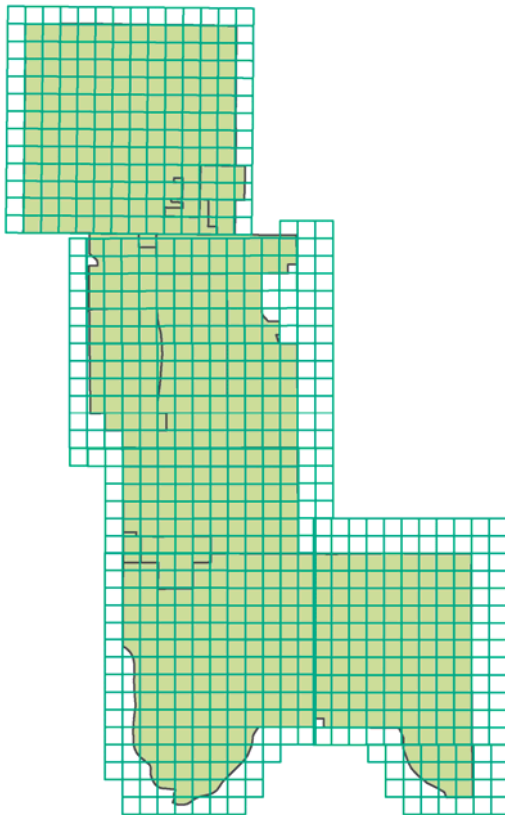
Fig. 2.1. Study area showing 805*805-meter grid with offsets.

These grid cells were used for sampling the participants in the study. The basic procedure includes:

- Creating an address list of all residential parcels. As there is no land-use code in the Ramsey County Assessor's data for mixed use, we simply included parcels matching the land-use codes in the types R (residential) and M (multifamily).[3] The specific parcel database did not include separate addresses for properties with multiple units under one ownership—typically rental units. The parcels could be identified through use type M (multifamily) and associated land-use codes (401,402,403,404) In this case we purchased more accurate address lists from a private vendor.

---

[3] In the Design Center's own computer filing system these are found in LUC.DBF on \\Environment\GISdata\Ramsey County\Ramsey\CAMA\ParcelDataLookup).

- The revised address list for each focus area was then placed in a column in an Excel spreadsheet. The random number generating function in Excel was used to generate random numbers in the adjacent column. The two columns were then sorted and the first 20 addresses (or whatever number was needed at that time) were selected.

- Because we deliberately spread the individual level measures over eight months, so that we were not measuring in winter and people from each area were being measured in spring through fall, we did multiple random selections for each phase without replacing those who had been participants (or in some cases we had promised not to contact). It would be equally appropriate, however, to do one random sample at the beginning, and just move down the list as more people were needed.

5. GIS Approach
For density and block size measures, see other chapters referenced above. An ArcGIS tool, Fishnet, was used to create a shapefile of grid squares in an appropriate projection.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. A layer containing the maximum extent for analysis
2. **Create Fishnet** tool (available from ESRI's ArcScripts Web site)

Create the **grid cell shapefile**:
1. Ensure that the data view is using the desired units for measurement by checking the lower-right corner of the status bar at the bottom of **ArcMap**.
2. Move the mouse to the lower-left corner of the area for which grid cells will be created. Note the coordinates.
3. Measure the approximate horizontal and vertical distances that need to be covered by the grid and divide each distance by the intended size of each grid cell. This will yield the number of rows and columns needed in the grid.
4. Click the **Create Fishnet** button that was added to the toolbar during installation.
5. Enter the X-Y coordinates of the lower-left corner recorded above.
6. Enter the number of rows and columns estimated above.
7. Enter the height and width of each grid cell.
8. Choose a location and name for the output shapefile.
9. Click the **Edit** button to select the coordinate system of the output shapefile.
10. Make sure it is set to product **Polygons** and click **OK**.

2.4 Measurement Geography Type 1: Grid Cell
Note: The Create Fishnet tool has not been tested with ArcGIS 9.1 and the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used. The steps required to "convert to graphics" can be performed in an ArcMap Edit Session, simplifying the workflow.

1. Basic Concept
This measurement geography divides an area into equal square cells for sampling and measurement. These protocols use three scales of grid cell: a focus area, a context area, and a 100-meter grid cell. The logic behind the first two is described in chapter 1 and they are primarily used to create the sample frame for the study and to guide data collection.

The small, 100*100-meter grid cells are used in measurement where data are only available in aggregated form or where measures require grid cells for other reasons described in the protocols for those measures (e.g., the land-use dissimilarity index).

2. Basic Definition
The focus area is an area 0.5*0.5 miles/805*805 meters used as basic measure for sampling.

The context area is an area 1.0*1.0 miles/1609*1609 meters that buffers each focus area by 0.25 miles/402 meters on all sides.

3. Detailed Formula
Focus area = Area 0.5*0.5 miles/805*805 meters used as basic measure for sampling. The basic focus area grid ignores water bodies, but these are generally excluded from land area calculations.

Context area = Area 1.0*1.0 miles/1609*1609 meters that buffers each focus area by 0.25 miles/402 meters on all sides. The basic context area grid ignores water bodies, but these are generally excluded from land area calculations. Used to collect data so that there can be a 0.25 mile straight-line buffer (at least) measured around all participants.

4. Comments and Explanations
None.

5. GIS Approach
The GIS approach is to use the Create Fishnet add-on for ArcGIS to create grid cells for one study site. Convert the grid cell shapefile to graphics and copy the graphics to each study site, aligning them as well as possible with each study site. Convert all the graphics into a new shapefile and assign the correct study site ID to each group of grid cells.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. A layer containing the maximum extent for analysis
2. **Xtools** extension (available from the Downloads section of the ESRI Web site)
3. **Create Fishnet** tool (available from ESRI's ArcScripts Web site)

Create the grid cell shapefile (this procedure assumes you are creating grid cells for a 1600m-square study site):
1. Ensure that the data view is using meters by checking the lower-right corner of the status bar at the bottom of **ArcMap**.
2. Move the mouse to the lower-left corner of the study site for which grid cells will be created. Note the coordinates. Click the **Create Fishnet** button that was added to the toolbar during installation.
3. Enter the **X-Y coordinates** of the lower-left corner recorded above.
4. Enter **16** for the number of rows and columns.
5. Enter **100** for the height and width of each grid cell.
6. Choose a location and name for the output shapefile.
7. Click the **Edit** button to select the coordinate system of the output shapefile.
8. Make sure it is set to product **Polygons** and click **OK**.

Save grid cells to graphics and duplicate:
1. Right-click on the new grid cell layer in the **Table of Contents** and choose **Convert Features to Graphics**.
2. Select the graphic grid cells (all 256 of them) so they are highlighted.
3. Right-click and choose **Group** to combine the cells into one selectable graphic.

---

4. Right-click to **Copy** and **Paste** the graphic to other study sites.
5. Align the graphics to each study site.
6. Select all the graphics, right-click and choose **Ungroup**.

Create **grid cell shapefile**:
1. Open the **Xtools** menu in the toolbar and choose **Feature Conversions→Convert Graphics to Shapes**.
2. Select the **Polygon** features and click **OK**.
3. The new shapefile will be added to the **ArcMap Document**.

2.5 Measurement Geography Type 2: Straight Line or Airline Buffer

Note: The Buffer Wizard has been replaced with the Buffer Tool in the ArcMap Toolbox. The Buffer Wizard tool is still available through the Menu Tools->Customize->Commands->Tools.

1. Basic Concept

This measurement places a buffer around a point or polygon and the radius of the buffer is equal at all points. This is useful in measuring the general context of a person. While the street network buffer (see next protocol) focuses on street distances, pedestrians do not always walk on streets and in some locations, the straight line or airline buffer may well capture parts of the environment accessible to pedestrians through parks, parking lots, trails, and informal paths.

It should be noted that we also use very small buffers in defining intersections, largely to eliminate classification errors from slightly offset intersections. This is complex to describe briefly, so see chapter 6 for more detail.
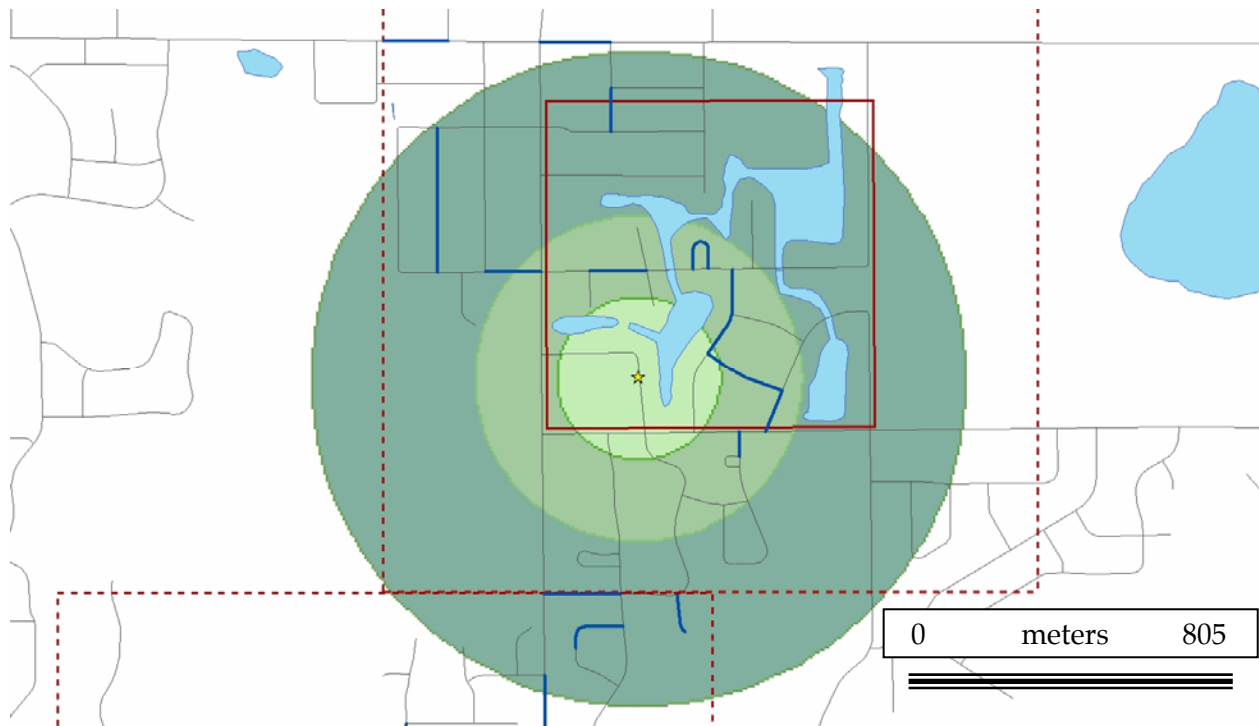
Fig. 2.2. Straight line or airline buffers of three different radii (in this case 0.125, 0.25, and 0.5 miles or 201, 402, and 805 meters) around a hypothetical individual.

Buffers can also be created around polygons such as parks. The following diagram shows straight-line buffers of 0.25 miles/402 meters around parks in a suburban area of the Twin Cities. In this case buffers around several parks have been merged to show areas within 0.25 miles/402 meters of any park (shaded) and beyond that distance (white). These maps are from separate work on park access for the Trust for Public Land.
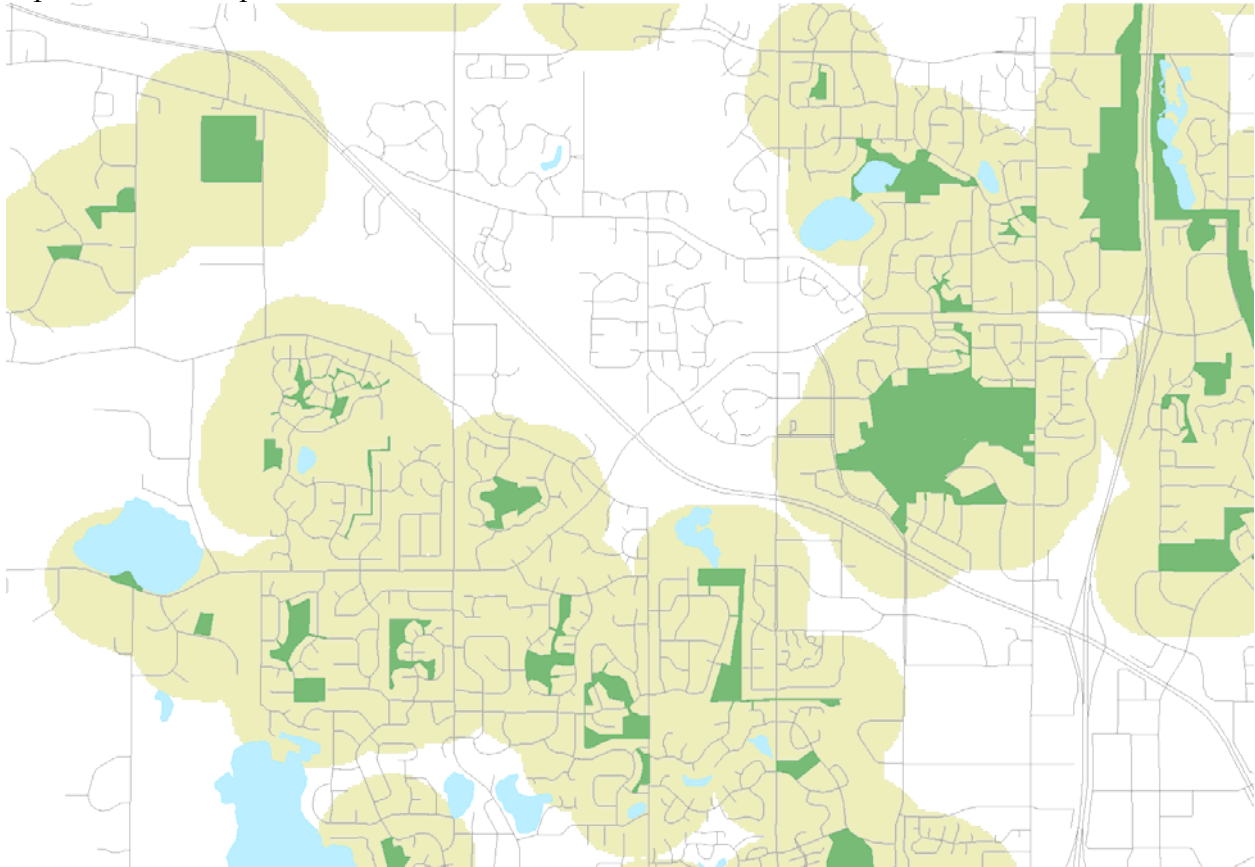
Fig. 2.3. Straight-line or airline buffers of 0.25 miles/402 meters shown in beige around parks shown in green. Areas beyond that distance are in white.

2. *Basic Definition*
A straight-line buffer is an area within "X" meters straight-line distance of a point (or a polygon).

3. *Detailed Definition*
A straight-line buffer is an area within "X" meters straight-line distance of a point (or a polygon). In the case of this study, with a sample based on residential environments, the straight-line distance is from a participant's home. More specifically, it is the straight-line distance from the residence on the participant's parcel, located visually from aerial photos. However, when using work addresses, the formula is similar.

4. *Comments and Explanations*

This is a straightforward process in ArcGIS 8.3, using the Buffer Wizard. The software allows you to choose whether to keep the buffers distinct (one per feature) or merge them into one big feature. For the parks analysis shown in the image above, the buffers could be merged together, as it was not necessary to retain a relationship between an individual park and its buffer. For the Twin Cities Walking Study, each buffer was tied to a participant's home location, so the buffers had to remain distinct.

Picking the starting point was a bit more time consuming in this process as we only partially automated it, checking, rather, the geocoding from the parcel database. The geocoding process works with the length of the street segment and the address range associated with it, and places points along the length of the line proportionally using the address and the known range. In order to have the points in the correct place we used the parcel data (to match exact addresses) and the orthophotos (to find the building locations) and manually shifted points to reflect the actual location of participants. In a larger study, we would automate this step to use the centroid of the parcel of residence. The manual placement of points makes the most difference in larger or longer lots where the centroid may be some distance from the actual house, affecting small buffers in particular.

Buffering can be a large task. For census-based measures, for instance, it is important to intersect straight-line buffers and census blocks; that is, to cut census blocks into sections that overlap with each individual buffer. This is a large processing task that requires two stages. First is the normal intersect between a set of buffers and the census blocks. Then, because the buffers of participants in a limited geographical area tend to overlap, the resulting file has to be run through the Dissolve procedure in the GeoProcessing Wizard. After intersecting census blocks with many overlapping buffers, each buffer gets split into many pieces. By dissolving this layer by both the unique ID of the buffer and the unique census block ID, the file is put back together into all the unique combinations of the two attributes. The result is the equivalent of simply doing a spatial intersect with non-overlapping buffers. The first step results in extremely large files. Intersecting 715 1600-meter buffers and the approximately 6500 census blocks they touch, for example, took nearly a half hour and generated a file containing 2.65 million records. In order to avoid a computer crash, this had to be broken up into six chunks of approximately half a million records each because the dissolve tool would crash trying to operate on more records than that. When all six pieces were done, the append tool was used to merge them back together—the final file contains 156,000+ records (i.e., census blocks and parts of census blocks).

Fig. 2.4. In the picture to the left you can see the parcel centroid marked with a green X, while the house sits at the back of the parcel with a long

driveway out to the road. In this picture the distance from the point on the house out to the road is just over 400 meters.

5. GIS Approach
Use the Buffer Wizard in ArcGIS to create one or more buffers a specified distance or distances from a selected set of features.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Participant Locations Layer** (point) with unique ID field
2. **Census Block Layer** (polygon) with field containing population

Create Participant buffers:
1. Under the **Tools** dropdown menu, select **Buffer Wizard**.
2. In the **Buffer Wizard** window, under **features of a layer**, select your **Participant Locations Layer** and click on **Next**.
3. Click **At a specified distance** and enter **XX** (your desired distance).
4. After **Distance units are**, select **meters**.
5. Click **Next**.
6. After **Dissolve barriers between**, select **No**.
7. Give the layer a name and save as a new layer.
8. Click **Finish**.

2.6 Measurement Geography Type 3: Network or Street Distance Buffer

Note: Network Analyst is available as of ArcGIS 9.1, but produces network buffer results that are significantly different than the ArcView 3.3 version. With versions 9.1 and 9.2of ArcGIS there is a radical change in the way that network buffers are made. This is dealt with extensively in the companion volume: Environment, Food, and Youth: GIS Protocols. In this

1. Basic Concept
This kind of buffer uses the street network as the organizing geography. A buffer of 400 meters would start by finding all points on the street network that are 400 meters from the starting point and then joining the dots. Its basic idea is that people use the street network to move about and something that is 400 meters as the straight line may be several times that far in an area with a winding or discontinuous street pattern. The illustration below shows this for a 0.5-mile/805-meter buffer around the same point as in the previous diagram. The network buffer is significantly smaller than the straight-line buffer.

It is important to note that the network buffer is not always better than a straight-line buffer as a measure of accessible areas. In some locations there may be many pedestrian paths and cuts, although they are not streets. This means that it may be useful to test both kinds of buffers.

2. Basic Definition
A network buffer is an area within X-meters street distance of a point (or a polygon).

3. Detailed Definition
A network buffer is an area within X-meters street distance of a point (or a polygon). In the case of this study, with a sample based on residential environments, the street network distance is from a participant's home. More specifically, it is the network distance from the house or building on the participant's parcel, where the building is located visually from aerial photos.

Where these buffers are zero or very small, due to the lack of cross streets, we created a work-around that created a minimum buffer size of 12,000 square meters. When using work addresses the procedure is similar.
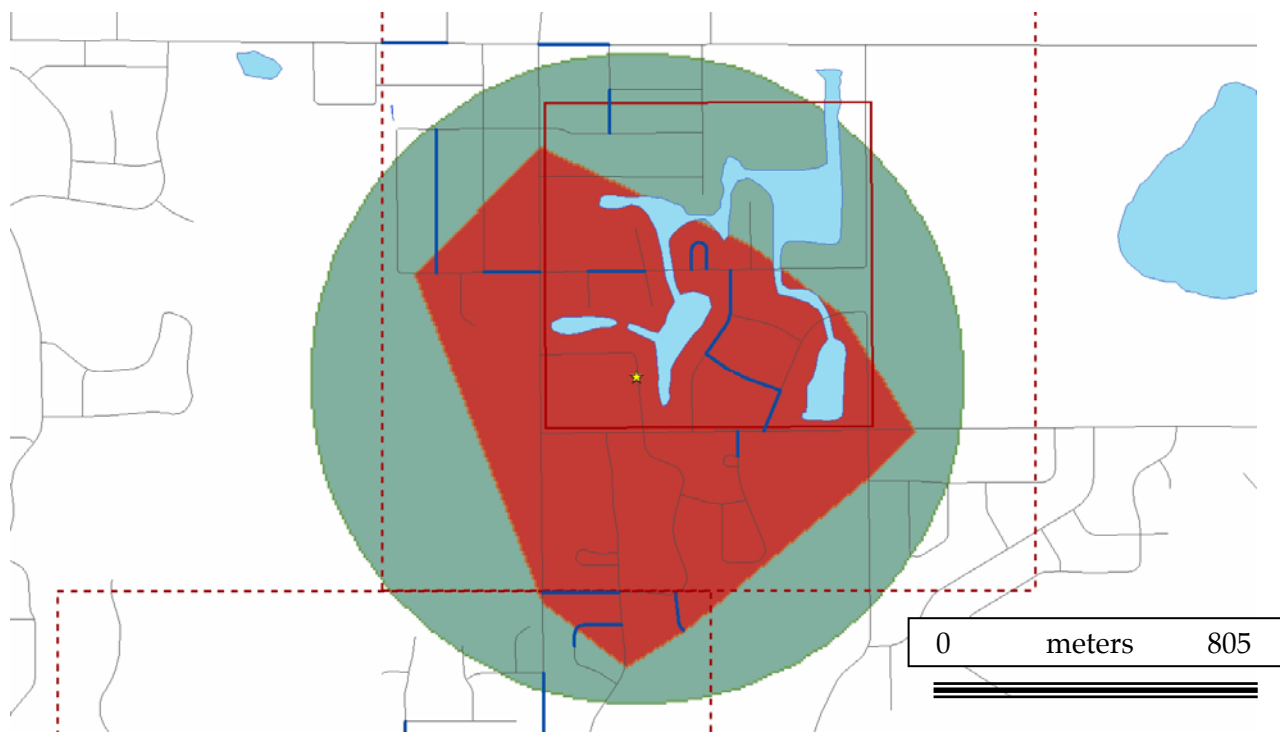
Fig. 2.5. Straight-line versus network buffers for a radius of 0.5 miles/805 meters around a hypothetical individual.

4. Comments and Explanations
This is a little trickier to do in GIS than the straight-line buffer. Currently Network Analyst, the tool that creates network buffers, has not been released for ArcGIS, so this must be done in ArcView 3.3 and then re-imported into ArcGIS.

In addition, there are some added complications, particularly for smaller buffers in street patterns dominated by cul-de-sacs. Network Analyst works by finding points on the network—in this case the street—a specified distance along the network from the starting point. It then joins these points into some kind of polygon. For 200-meter buffers in areas without many cross streets, 200 meters might just result in a long, skinny line up and down a single street. This results in a buffer that does not intersect with land use, street trees, or other items outside of the road right of way. We could not find any direction for this in the literature, so created a minimum buffer for those of 15 meters on each side of the road center line for 185 (200 minus 15) meters each way or approximately 12,000 meters squared or 1.2 hectares. Overall, approximately 60 buffers, or roughly 8 percent of buffers, received this minimum size. We in fact came to this in two steps. First we ran minimum buffers on zero or almost-zero network buffers, but these minimums ended up being bigger than some of the standard network buffers up to 12,000 square meters, given slight curves in the road and the odd intersection. So then we expanded the number of minimum buffers to all those under 12,000 square meters. This also involved moving starting points out to the right of way, where they were close (in two cases they were far and were not moved).

We ended up with approximately 656 unmodified features; 57 that were edited by moving the points out to the road and running them through Network Analyst, taking the resulting road features, and buffering them by 15 meters; and 2 features set to have 200-meter network buffers of 0 square meters.

*5. GIS Approach (**Original version; updated approach and steps for ArcGIS 9.1 are below and it is important to read the companion manual to understand the differences.**)*
Use the Network Analyst extension for ArcView 3.3 to determine a service area around a set of point features. This requires lines features such as roads or trails that will become the network used to calculate all possible routes of a specified distance from the selected points.

6. GIS Steps
The following procedures were completed using ArcView 3.3.

To complete this protocol you will need:
1. **Roads Layer** (polyline)
2. A set of features from which to buffer—in this example, participant locations (point)
3. The **Network Analyst** extension for **ArcView**

Prepare the view and data:
1. Add the roads theme and the participant theme to a new view.
2. Go to the **View** menu and choose **Properties**.
3. Set the map units and distance units appropriate for the coordinate system of the layers.
4. In the view, select the participant theme, open the **Theme** menu, and select **Table…**
5. Go to the **Table** menu and select **Start Editing**.
6. Go to the **Edit** menu and select **Add Field**.
7. In the **Field Definition** window, name the field **REPORTCOST**.
8. Choose the field type **Number**, then click **OK**.
9. In the theme's table, click on the name of the newly added field (it should turn dark gray).
10. Go to the **Field** menu and choose **Calculate…**
11. Enter the buffer distance and click **OK**.

Calculate network buffers:
1. In the view, highlight the roads theme.
2. Go to the **File** menu and choose **Extensions**.
3. Scroll to **Network Analyst**, check the box, and click **OK**. There should now be a **Network** menu at the top of the **ArcView** window.
4. Go to the **Network** menu and choose **Find Service Area**.
5. Check the radio button for **Travel** from site.
6. Click the **Load Sites** button.
7. When asked to choose a theme, select the participant locations theme and click **OK**.
8. The **Find Service Area** window will now be titled **Sarea1** and **Snet1**. Click the button in the upper right with a picture of a grid and a black arrow on it.
9. When the processing finishes, the polygon theme **Sarea1** will contain the network buffers.
10. To export the features to a fresh shapefile, click on the **Sarea1** theme, go to the **Theme** menu, and choose **Convert to shapefile**.
11. Choose a location and name for the output file and click **OK**.

If, once the buffers have been created, there are buffers with no – or very little – area due to their long skinny shape and their location on a road without cross streets within the network buffer distance, use the following GIS Steps to modify these buffers.

Move the participant points (using **ArcMap 9.0**):
1. Open the attribute table for the network buffers
2. Create a new field called **MODIFY**
3. Select the buffers that will need to be modified

4. In the **MODIFY** field, give these buffers a value of 1 (the rest of the buffers should get a value of 0)
5. Join the network buffer layer to the participant point layer
6. Open the attribute table for the participant point layer
7. Create a new field called **MODIFY**
8. Set this new field equal to the **MODIFY** field from the network buffers
9. Remove the join
10. Select only the points to be modified (**MODIFY** = 1) and export the data to a new shapefile
11. In the **Editor Toolbar**, choose **Start Editing**
12. Select the correct source for the subset of participant points and click on OK
13. In the **Editor Toolbar**, choose **Snapping**
14. For the roads layer, check the box under **Edge**
15. For each point to be moved, click and drag the point perpendicular to the nearest road until it snaps to the road
16. When complete, in the **Editor Toolbar** choose **Stop Editing** and save changes

Prepare the view and data (using ArcView 3.3):
1. Add the roads theme and the modified participant theme to a new view
2. Go to the **View** menu and choose **Properties**
3. Set the map units and distance units appropriate for the coordinate system of the layers
4. In the view, select the modified participant theme, open the **Theme** menu, and select **Table…**
5. Go to the **Table** menu and select **Start Editing**
6. Go to the **Edit** menu and select **Add Field**
7. In the Field Definition window, name the field **REPORTCOST**
8. Choose the field type **Number**, then click OK
9. In the theme's table, click on the name of the newly added field (it should turn dark gray)
10. Go to the **Field** menu and choose **Calculate…**
11. Enter the buffer distance (we entered our network distance minus 15m, which is what we used as a buffer distance around the selected street segment) and click OK

Calculate network street segments (using ArcView 3.3):
1. In the view, highlight the roads theme
2. Go to the **Network** menu and choose **Find Service Area**
3. Check the radio button for **Travel from site**
4. Click the **Load Sites** button
5. When asked to choose a theme, select the modified participant locations theme and click OK
6. The Find Service Area window will now be titled Sarea1 and Snet1. Click the button in the upper right with a picture of a grid and a black arrow on it.
7. When the processing finishes, the polyline theme Snet1 will contain the network street segments
8. To export the features to a fresh shapefile, click on the Snet1 theme, go to the **Theme** menu, and choose **Convert to shapefile**
9. Choose a location and name for the output file and click **OK.**

Create modified network buffers (using **ArcMap 9.0**):
1. Add the new network street segments theme to the Map Document
2. Open **ArcToolbox** and select **Analysis Tools**, **Proximity**, **Buffer**
3. Use your network street segment layer as the **Input Feature**
4. Give your **Output Feature** a name and location
5. Enter a linear distance of X and select the units (we used 15 meters)
6. Keep **Side Type** as **Full**, **End Type** as **Round**, and change **Dissolve Type** to **List**
7. From the **Dissolve Field(s)** options check the box for your participant ID
8. Click OK

Append original and modified network buffer polygons
1. From the original network buffer layer, select only those buffers with MODIFY = 0 and export the data to a new shapefile

---

2. Open **ArcCatalog** and create a new polygon shapefile that will contain your complete modified network buffer layer
3. In **ArcToolbox**, select **Data Management Tools, General, Append**
4. For **Input Features**, select your two network buffer layers (the newly modified network buffers and the subset of original "unmodified" buffers)
5. For **Output Features** select your newly created polygon shapefile
6. Under **Schema Type** select **No_Test**
7. Click **OK**
8. The two buffer layers will now be combined into one modified network buffer layer

5. GIS Approach *(updated version)*

Use Network Analyst extension for ArcGIS 9.1to determine the new service area around the set of point features, in this case participants or schools. Line features such as roads or trails are used to create the network which is used to calculate all possible routes of a specified distance from the seleceted points.

6. GIS Steps

The following procedure was completed using ArcInfo 9.1

Requirements:
1. **Road Layer** polyline
2. **Point Features** for buffer
3. **Network Analyst Extension** for ArcGIS 9.1

I. Create a New Network Dataset
   1. Start **ArcCatalog**.
   2. On the Main Menu click **Tools** and Click on **Extensions.**
      a. Check **Network Analyst** Extension.
      b. Click Close.

   3. In the catalog tree, navigate to the feature dataset if you are creating a geodatabase based network dataset. If you are creating a shapefile based network dataset navigate to the shapefile location in the catalog tree.

   4. If you are creating a geodatabase based network dataset right click on the feature dataset and point to **New** and then click **Network Dataset**. If you are creating a shapefile based network dataset right click on the shapefile and click **New Network Dataset**.

   5. In the **New Network Dataset Wizard**
      a. Type a name for the network dataset and Click **Next**.
      b. Check the **Road feature classes** that will be used as network source in the network and Click **Next**.
      c. Accept the **default connectivity** settings and Click **Next**.
      d. Accept the **default** option **NO** to modify the connectivity with **elevation** field data or elevation field settings and Click **Next**.
      e. Accept the **default YES** to model **Turn** in the Network and Click **Next**.
      f. Click **Add**
         i. Type **Length** in the Name field.
         ii. Select **Cost** as **Usage Type**, if it is not already selected.
         iii. Select **Meters** as the **Units**.
         iv. Select **Double** as the **Data type**, if it is not already selected.
         v. Click **OK**.
         vi. Click **Next**.
      g. Accept the default **Yes** to use the **default directions** settings to establish driving directions for the network dataset and Click **Next**.

---

h. Review the network dataset summary to check the options are correctly selected and Click **Finish**.

6. Click **Yes** to build the Network Dataset.

II. Create the Network or Street Distance Buffer
1. Start **ArcMap**
2. On the Main Menu click **Tools** and Click on **Extensions.**
   a. Check **Network Analyst** Extension.
   b. Click Close

3. Click the **Add Data** button on the **Standard Toolbar**
   a. Navigate to the **Network Dataset** location, created earlier in Step I
   b. Click on the **Network Dataset** and click **Add** to add the Network Dataset to ArcMap.
   c. Click **Yes** to add all the feature classes, that participate in the network, to the map.

4. Click the **Add Data** button on the **Standard Toolbar**
   a. Navigate to the **Point Layer** to be used for analysis, in this case **Participant location/School Locations**
   b. Click on the **Point Layer** and click **Add** to add it to ArcMap.

5. In the Table of Contents Right Click the **Point Layer** and click on **Open Attribute Table**
   a. In the **Attribute Table**
      i. Click **Options**
      ii. Click **Add Field**
      iii. Add the **School_Id/Participant_Id** as **Text**
         ▪ Type **School_Id/Participant_Id** in the **Name** field
         ▪ Select **Text** (50) as **Type**
         ▪ Accept **default** for **Length**
         ▪ Click **OK**
      iv. Add the **School_Id/Participant_Id** as **Integer**
         ▪ Type **School_Id/Participant_Id** in the **Name** field
         ▪ Select **Short Integer** as **Type**
         ▪ Accept **default** (0) for **Precision**
         ▪ Click **OK**
      v. Enter the assigned Id for the participants and schools

6. On the main menu click **View**
   a. Click on **Toolbars**
   b. In the side menu click **Network Analyst** to add the Network Analyst toolbar to ArcMap

7. Click **Network Analysis Toolbar**

   a. Click **Network Analyst Window** button to add the Network Analyst Window to the map

8. On the **Network Analyst Toolbar** Click **Network Analyst Dropdown Menu**
   a. Click **New Service Area**
   b. The **Table of contents** and the **Network Analyst** Window contains empty list of **Facilities**, **Barriers** and **Polygon** Categories.

9. On the **Network Analyst Window** right Click **Facilities** and Click **Load Locations**
   a. In **Load Locations Window**
      i. Click **Browse** and Navigate to the **point feature layer**

---

    ii. Click on the **Point feature layer** and click **Add**
    iii. Choose the **Participant ID/School Id** as the **Name** field. Accept the **default** for other **Location Analysis Properties**
    iv. Accept the **default Location Position**
    v. Click **OK**

10. To set the properties to be used in the Analysis, In the Table of Contents Right Click the **Service Area Layer** and Click **Properties**
    a. Click on **Analysis Settings Tab**
        i. Choose **Length** (Meters) for **Impedance**.
        ii. Type the **Buffer Distance** for **Default Breaks**
        iii. Choose **Away From Facility** as **Direction**
        iv. Choose **Everywhere** for **Allow U-Turns**
    b. Click on **Polygon Generation Tab**
        i. Check **Polygon generation**
        ii. Choose **Generalized** as the **Polygon Type**
        iii. Choose **Separate Polygon per facility** for **Multiple Facility Option**
        iv. Choose **Rings** as the **Overlap Type**
    c. Click **Ok**

11. On the Network Analyst Toolbar, Click **Solve** Button
12. Joining attributes with Source Point Layer to carry over the data.
    a. Right Click the **Routes Layer**
    b. Click **Joins and Relates** and then **Join**
        i. Choose **Facility Id** as the Field to base the join on
        ii. Choose the **Restaurant Layer** as the Table to Join to this layer
        iii. Choose the **Restaurant ID** as the field in the table to base the join on
    d. Right Click the **Routes Layer**
    e. Click on **Data** and then **Export Data**
        i. Choose **this layer's source data** for the coordinate system
        ii. Specify the **Output location**

2.7 Measurement Geography Type 4: Straight Line or Airline Distance to Nearest Feature
Note: The protocol has been modified and tested in ArcMap 9.1

1. Basic Concept
If people walk at least partly in order to reach an important destination, such as a park or a supermarket, then the distance to the closest example of a specific kind of destination is an important issue.

2. Basic Definition
The straight-line distance to the nearest feature is the distance to the closest X.

3. Detailed Definition
The straight-line distance to the nearest feature is the distance to the closest X where X is a type of destination.

4. Comments and Explanations

As of version 3,1, this measurement used an ArcView extension called Nearest Features. It was developed by Jeff Jenness, a wildlife biologist, and can be downloaded here: http://www.jennessent.com/arcview/nearest_features.htm

However, for version 4.0, protocol 2.7 has been revised to use ArcGIS 9.1 and the toolbox tool "Near".

For determining the shortest distance from a participant to a business, the distance is between the two points. When determining the nearest polygon feature, for example a park, the distance is calculated between the participant and the closest vertex. Vertices are more useful representations of access than centroids, particularly if the polygon feature (e.g., a park) is very large or oddly shaped. In this case, the centroid may be located far from the nearest road.

5. GIS Approach

This calculation simply involves determining the straight-line distance between an origin and the closest of the possible destinations of the desired type (businesses, parks, etc.). For non-point destinations, this measurement requires a decision about what part of the feature to call the endpoint of the shortest distance line—the centroid, vertex, or closest edge.

6. GIS Steps

The following procedures were completed using **ArcGIS ArcInfo 9.1.**
Note: as of ArcGIS version 9.1, the **Near** tool is only available in **ArcInfo**. For **ArcView** 9.1 users, equivalent results can be obtained using "Hawth's Analysis Tools for ArcGIS" **Distance Between Points (Between Layers)** found at http://www.spatialecology.com/htools/overview.php.

To complete this protocol you will need:
1. A set of origin features (point) (Participants in this example)
2. A set of destination features (point) (Bus Stops in this example)
3. If using ArcGIS ArcView - Hawth's Analysis Tools for ArcGIS

Prepare the view and data:
1. Add the **Origin** and **Destination** themes to a new **Data Frame**.
2. Go to the **View** menu and choose **Data Frame Properties**..
3. Set the map units and distance units appropriate for the coordinate system of the layers
4. Click **OK**.

Calculate distance to nearest feature (if using ArcGIS ArcView use Hawth's **Distance Between Points (Between Layers)** tool)
1. Open **ArcToolbox**.
2. Choose the **Near** tool under **Analysis Tools -> Proximity**.
3. For the **Input Features**, select the theme of **origin features** Participant.
4. For the **Near Features**, select the **destination features** Bus Stops.
5. Click **OK** to begin the processing.
The distance to the closest destination feature will be added to the origin features

2.8 Measurement Geography Type 5: Network or Street Distance to Nearest Feature

Note: The protocol has been modified and tested in ArcMap 9.1

1. Basic Concept

As with the straight-line distance, network or street distance is important as a measure of accessibility of destinations. As stated in the previous protocol, if people walk at least partly in order to reach an important destination, such as a park or a supermarket, then the distance to the closest example of a specific kind of destination is an important issue.

The straight-line distance, however, may not be a realistic measure of the actual distance to the nearest feature as people typically walk along streets and, therefore, the street distance is a more accurate measure. It is important to note that it is not always better, as in some areas there may be many pedestrian paths and cuts, although they are not streets. This means that it may be useful to test both kinds of distance measures.

2. Basic Definition

The street distance to the nearest feature is the distance to the closest X.

3. Detailed Formula

The network distance to the nearest feature is the distance to the closest X where X is a type of destination and the closest X is the closest X in terms of street distance. There may well be closer destinations of this type in straight-line terms, but barriers or street patterns make them further along the network.

4. Comments and Explanations

This is also done using Network Analyst in ArcGIS 9.1 using a function called Find Closest Facility. You can specify a point layer to use as your "facilities" and then also a point to serve as your origin. It will then take the road network and calculate the shortest path to a "facility" (and it can do this using travel costs, as well, if, for instance, certain types of roads were considered more or less easy to use).

For points that do not fall on the network, there is a dialog box when using this extension that allows the user to decide how far away from the network to search for points to use in the processing. Points outside that distance will not be used.

5. GIS Approach

The approach for this measure is to load a set of origins and destinations. Then use the Closest Facility feature in Network Analyst to calculate the network distance from each origin to the nearest destination.

6. GIS Steps

The following procedures were completed using **ArcGIS ArcInfo 9.1**. This protocol will also work with ArcGIS ArcView 9.1 without modification.

To complete this protocol you will need:
1. **Roads Network Layer** (Network Dataset) (See Appendix 5E for general information regarding the creation of network datasets.)
2. A set of origin features—in this example, **Participant Locations** (point)
3. A set of destination features—in this example, **Bus Stop Locations** (point)
4. **Network Analyst** extension for **ArcGIS 9.1**

Prepare the view and data:
1. Start a new **ArcView** project.
    a. Add the **Roads**, **Origins**, and **Destinations** layers to the **View**.

---

b. Open the **Network Analyst Toolbar** and choose **New Closest Facility** From the Network Analyst Menu (on the Toolbar). A **Closest Facility Group Layer** will be added to the active data frame.
c. Open the **View** menu and choose **Data Frame**
d. Make sure the map units (meters, feet, etc.) match the units of the data layers.

e. On the **Network Analyst Toolbar**, click the **Network Analyst Window** button   .

2. On the **Network Analyst Window**, right-click **Incidents** and click **Load Locations**.
   a. Choose the **Participant Locations** for your **Load From** feature set.
   b. Choose a unique ID field for the **Name** property under Location Analysis Properties. In this case the **Name** will be the participants unique ID number.
   c. Choose a suitable **Search Tolerance** for the **Location Position.** If features cannot be located on the network dataset, this tolerance may need to be increased.
   d. To verify all features were located on the network dataset do one of the following. If any feature has a status of **Not Located,** the search tolerance set in 2.c may need to be increased or the feature reviewed for correct location.
      i. For small datasets, Scroll through the resulting lists of Facilities in the **Network Analyst Window,** looking for features with RED Question Marks
      ii. For large datasets, it may be easier to Right Click on the **Facilities** in the **Network Analyst Window, Open Attribute Table** and sort descending on the **STATUS** column, the "Not Located" features should sort to the top.
3. Repeat steps 2.a through 2.d for the **Facilities** features in the **Network Analyst Window.** The features to be loaded are the **Bus Stop Locations.** In this case, the unique ID field for the **Name** property will be a unique identifier for the bus stops.

Calculate network distance to nearest facility
4. Right-click on the **Closest Facility Group Layer** created in step 1b and set the following **Properties** under the **Analysis Settings** tab:
   a. Impedance: Length (Meters or whatever measurement units your data is in)
   b. Default Cutoff Value: <None>
   c. Facilities To Find: 1
   d. Travel From: Incident to Facility
   e. Allow U-Turns: Everywhere (this may be different for particular datasets)
   f. Output Shape Type: True Shape
   g. Check any Restrictions particular to your **Roads Network Layer** that you would like to enforce (such as one-way streets – for walking, one-way probably is not a true restriction)
5. Right-click on the **Closest Facility Group Layer** and click on **Solve.** This will perform the nearest facility analysis and create the save the shortest network path in the **Routes** feature class**.** The network distance is recorded in the **Total_Length** field.

Network Analyst creates its own unique ID for incidents and facilities. To be able to relate the network distance back to the Participants ID the following steps need to be performed
6. Right-click on the **Routes** feature class in the **Closest Facility Group Layer** and choose **Joins and Relates -> Joins…**
   a. What do you want to join to this layer? **Join attributes from a table**
   b. The field in this layer the join will be base on: **IncidentID**
   c. Table to join: **Incidents**
   d. Field in table to base the join on: **ObjectID**
7. The resulting routes are stored in the **Roads Network Dataset.** To save the routes to a shapefile, right-click **Routes,** choose **Data…Export Data.**

The **Closest Facility Group Layer** can be reused for several different facilities types and the same **Participants Locations** by reloading new facilities. To delete the existing facilities, right-click **Facilities** in the **Network Analyst Window** and choose **Delete All.** Repeat steps 3, 4, 5 and 7 to perform the analysis with a different set of facilities. Step 6 only needs to be performed once, as the table join will persist with the **Closest Facility Group Layer.**

2.9 Measurement Geography Type 6: Parcel of Residence

1. Basic Concept
Recent work by Anne Vernez Moudon and her research team has found more of a correlation between features of the parcel of residence and walking than some larger geographies—in her case 1-kilometer and 3-kilometer buffers. This may be due to features of the parcel of residence being a proxy for non-environmental features such as levels of car ownership or due to the lack of intermediate scales in her analysis. We originally intended using parcel of residence to check some of these variables however eventually decided against it.

2. Basic Definition
The parcel in which a participant resides.

3. Detailed Formula
As above.

4. Comments and Explanations
None.

5. GIS Approach
This merely involves selecting the parcel of residence.

6. GIS Steps
Steps are in other protocols.

3. DENSITY

By Ann Forsyth (drafting, conceptual issues and formulae, some comments and GIS approach text) and Joel Koepp (GIS Steps and the rest of the comments and GIS approach text) and Ed D'Sousa and David Van Riper GIS steps from version 4.0 on.

The only change between version 2.1 and version 2.0 was some additional detail in protocol 3.3. For version 2.3 an error was corrected in protocol 3.4 (it does exclude water). In version 3.1, protocol 3.5 was edited to simplify steps.

Background

Density is the number of units (people, trees, housing units, etc.) in a given area. Some researchers propose that density is primarily a proxy for other variables that actually cause walking and others propose it is more directly causal. Those on the proxy side argue that it is highly correlated with older neighborhoods with gridded streets, mixed use, low-income people, etc. For those seeing a more direct link, density provides a critical mass of people and places, a physical sense of community, auto congestion, and makes transit viable so people walk to transit (for reviews, see Ewing 1994, Steiner 1994, Cervero and Kockelman 1997, Handy 2003, Audirac 1999).

There are many measures of density depending on what the density is of (residents, employees, buildings) and the base land area (a parcel, a whole neighborhood) (Forsyth 2003, see fig. 3.1 below). A single home may be part of different measurement areas that give it different overall densities depending on the scale of the measures (see fig. 3.2 below). *Note: the figures use a different terminology from the protocols, largely because the protocols are independent of size of base area.*

Not all density measures are reported in this section. Chapter 4 on pedestrian infrastructure, for example, contains a number of density measures, such as street tree density.

Fig. 3.1. Selected Residential Density Definitions

Note: this uses a different terminology from the protocols, largely because the protocols are independent of size of base area.

Source: Forsyth 2003; reprinted from Forsyth 2004a

Parcel or Site Density: dwelling units (DUs) or residential population (RP) divided by total site/parcel area (all uses). This is often used by developers. It is easy to calculate with GIS but also fairly simple by hand if there is only one parcel. However, since parcel boundaries are not always visible on the ground this form of density can be hard to calculate from physical observations.

Block Density: DUs or RP divided by block area measured to the curb. This is relatively easy to measure from aerial photos and census data, and reflects a unit that is part of the experience of place, the block. However, if the block is not surrounded by roads, for example where it abuts open space, the boundaries can be less clear.

Net Neighborhood Residential Dwelling/Population Density: DU or RP divided by total land area devoted to residential facilities. This is a calculation that involves defining both a neighborhood and residential land within that neighborhood. Care must be taken in assigning land to residential uses rather than, say, recreation—the key is to find equivalent elements in different residential designs.

Net Neighborhood Density (NND): DU or RP divided by the neighborhood area with the base land area calculated to exclude citywide uses in the neighborhood. Included in the neighborhood land area are residential land, streets, and neighborhood type uses—schools, parks, churches/synagogues/temples etc. and neighborhood shopping. Excluded are citywide businesses, public uses, high schools and colleges, major arterials, major regional parks, and vacant and unusable land. These exclusions can be difficult to calculate (adapted from Alexander 1993).

Gross Neighborhood Density (GND)/Gross Census Tract Density (GCTD): DU or RP divided by the total neighborhood area. This is easy to calculate although it may be skewed by regional uses such as regional parks. The Gross Census Tract Density is particularly useful as it is available across the United States from census information and does not rely on local data.

City Density (CD): DU or RP divided by the entire developed area of the city or town. In built out local government areas this is in practical terms the entire city. On the urban edge, it includes only developed land, a more complex calculation (adapted from Alexander 1993). This is also a gross density.

Metropolitan Density (MD): DU or RP for US Census Metropolitan Statistical Area divided by total land area. This calculation includes undeveloped areas, which will lower the overall figures, but is nationally comparable. The US Census prepares such density figures. This is also a gross density.

Fig. 3.2. Comparison of Residential Density Measures for the Same Location

Source: Forsyth 2003

Site density	10 DUs per acre/25 per hectare
Block density	8 DUs per acre/20 per hectare
Net neighborhood residential density	10 DUs per acre/25 per hectare
Net neighborhood density	6 DUs per acre/15 per hectare
Gross neighborhood density	5 DUs per acre/12 per hectare
City density	4 DUs per acre/10 per hectare
Metropolitan density	3 DUs per acre/7 per hectare

3.1 Raw Population per Unit Land Area

Note: The Geoprocessing Wizard has been replaced with the Intersect Overlay Tool in the ArcMap Toolbox and the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

Note: Previous versions of this protocol used the Geoprocessing Wizard and Xtools. ArcGIS 9.2 has replaced the Geoprocessing Wizard with the Intersect Tool, and it offers a Calculate Geometry dialog box, negating a need for Xtools. The protocol detailed below has been tested in ArcGIS 9.2.

1. Basic Concept

This is the most basic gross measure of population density, defined as the population of those who reside in housing in the area. It should be noted that this measure is independent of scale and could be used to measure the population per hectare of a small grid cell or an entire metropolis (and therefore would have different names, e.g., raw grid cell density, raw metropolitan density).

Sources that have used this measure, typically at a neighborhood or census tract level, and that will be more fully discussed in the next draft of this manual include Frank and Pivo (1994, 46), Hess et al. (1999, 1), Schimek (1996, 121), and Dunphy and Fisher (1996, 91).

Because the Twin Cities has a great many lakes, we generally used the next protocol, which excludes water.

2. Basic Formula

Raw Population per Unit Land Area = Persons in housing units per unit area with no exclusions.

3. Detailed Formula

Raw Population per Unit Land Area = Persons in housing units as measured in the US Census per unit area with no exclusions.

4. Comments and Explanations

Where census blocks are cut by the end of the measurement geography, the population is apportioned according to the percentage of land area falling inside and outside the line.

5. GIS Approach

Use ArcGIS to intersect study sites with census blocks. Retain the original area of each census block and then use Calculate Geometry to update the areas after performing the intersect procedure. Find the proportion of new area to original area and multiply the census population figure by that same proportion. Finally, divide this apportioned population figure by the updated area.

6. GIS Steps

The following procedures were completed using **ArcMap (ArcInfo) 9.2**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field
2. **Census Block Layer** (polygon) with field containing population

Determine the apportioned census population figure for each site:
1. Add the **Study-Site Layer** and **Census Block Layer** to your **Map Document**.
2. Add the area fields to the **Census Block Layer**
   a. Right-click on the **Census Block Layer** and select **Open Attribute Table**.
   b. Click on the **Option** button and select **Add Field…**
   c. Name the field **Area** and change the Type to **Double**.
      i. If a field named **Area** already exists, change the name in the **Add Field…** dialog box to something else
   d. Click **OK**.
   e. Click on the **Option** button again and select **Add Field…**
   f. Name the field **Acres** and change the Type to **Double**.

     i. If a field named **Acres** already exists, change the name in the **Add Field…** dialog box to something else

  g. Click **OK**.
  h. Click on the **Option** button again and select **Add Field…**
  i. Name the field **Hectares** and change the Type to **Double**

     i. If a field named **Hectares** already exists change the name in the **Add Field…** dialog box to something else

  j. Click **OK**.

3. Calculate the values for Area, Acres, and Hectares
  a. For the **Area** field,
     i. Right-click the **Area** field in the **Census Block Layer** attribute table and select **Calculate Geometry…**
     ii. Set the **Property:** box to **Area**.
     iii. Set the **Coordinate System** to **"Use coordinate system of the data source:"**
     iv. Set the **Units** to the desired type. For the **Area** attribute, square meters, miles or feet is probably appropriate.
     v. Click **OK**.
  b. For the Acres field,
     i. Right-click the **Acres** field in the **Census Block Layer** attribute table and select **Calculate Geometry…**
     ii. Set the **Property:** box to **Area**.
     iii. Set the **Coordinate System** to **"Use coordinate system of the data source:"**.
     iv. Set the Units to **Acres US [ac]**.
     v. Click **OK**.
  c. For the Hectares field,
     i. Right-click the **Hectares** field in the **Census Block Layer** attribute table and select **Calculate Geometry…**
     ii. Set the **Property:** box to **Area**.
     iii. Set the **Coordinate System** to **"Use coordinate system of the data source:"**.
     iv. Set the Units to **Hectares [ha]**.
     v. Click **OK**.

4. Open **ArcToolbox** by clicking on the icon and navigate to the **Intersect tool** (Analysis Tools -> Overlay -> Intersect). Double-click the **Intersect tool** to open it.
  a. For the **Input Features**, click on the down-arrow and select the **Census Block Layer** from the list of layers.
  b. Repeat 4.a and select the **Study –Site Layer**.
  c. Specify a name and location for the output layer in the **Output Feature Class** box.
  d. Set the **JoinAttributes** drop-down box to **ALL**.
     i. The ALL option will transfer the all attributes from the input features to the output feature class.
  e. Leave the **XY Tolerance** box empty.
  f. Set the **Output Type** to **INPUT**.
     i. The INPUT option will return intersections with the same geometry type as the Input Feature with the lowest geometry dimension.
  g. Click **OK** to execute the Intersect. The output feature class will be added to the **Map Document**.

5. Open the **Attribute Table** of the clipped **Census Block Layer** created in the previous step.
  a. Click on **Options** and choose **Add Field**.
  b. Name the field **OrigHect** and change the Type to **Double** and click OK.
  c. Right-click on the **OrigHect** field.
  d. Choose **Field Calculator**.
  e. In the **Field Calculator**, click **Hectares** and click **OK**.
  f. You have now copied the values from **Hectares** to **OrigHect**.

6. Update the **Hectares** attribute in the clipped **Census Block Layer**
  a. Right-click on the **Hectares** field and select **Calculate Geometry…**
  b. Set the **Property:** box to **Area**.

---

      c. Set the **Coordinate System** to **"Use coordinate system of the data source:"**.
      d. Set the Units to **Hectares [ha]**.
      e. Click **OK**. This will overwrite the values in the Hectares field with the new correct values (since the shapes have been clipped to the extent of the Study-Site Layer), yet the original number of hectares is preserved in the OrigHect field.

7. To perform the population apportionment, open the **Attribute Table** of the clipped **Census Block Layer** again.
      a. Click on **Options** and choose **Add Field**.
      b. Name the field **PopSplit** and make it of type **Double**.
      c. Back in the **Attribute Table**, right-click on the name of the **PopSplit** field.
      d. Choose **Field Calculator...**
      e. In the **Field Calculator**, click **Hectares**, then click the / button, then click **OrigHect**, click the * button, and finally click the name of your **Population field**
      f. Click **OK**.

8. The result is a field called **PopSplit** that contains the same proportion of the original census population figure as the proportion of the clipped census block area to the original area.

Perform final calculation
1. In the **Attribute Table** of the clipped **Census Block Layer**, right-click on the name of the site ID field and choose **Summarize**.
      a. In the box beneath **Choose one or more summary statistics**, click the + next to **PopSplit**.
      b. Check the box next to **Sum**.
      c. Choose a name and location for the **Output Summary Table**.
      d. Click **OK**.
2. Add PopPerHect attribute to **Study-Site Layer Table**.
      a. Open the **Study-Site Layer** attribute table.
      b. Click the **Option** button and select **Add Field…**
      c. Add a new field called **PopPerHect** and select type **Double**.
3. Join this **Summary Table** into the **Study-Site Layer Table** using the **ID** field.
      a. Right-click the **Study-Site Layer** and select **Joins and Relates -> Join…**
      b. From the box entitled "**1. Choose the field in this layer that the join will be based on:**", choose the unique **ID** field for the Study-Site Layer.
      c. From the box entitled "**2. Choose the data to join to this layer, or load the table from disk**:", choose the **Summary Table** created in the previouw step.
      d. From the box entitled "**3. Choose the field in the table to base the join on:**", choose the unique **ID** field for the Study-Site Layer.
      e. Click **OK**.
4. Calculate the population per hectare for the **Study-Site Layer**.
      a. Open the **Study-Site Layer** attribute table.
      b. Right-click on this new field and choose **Field Calculator**.
      c. In the **Field Calculator**, click on the **PopSplit** field, then click the / button, and finally click on the **Hectares** field from the **Study-Site Layer**.
      d. Click **OK**.

The **PopPerHect** field now contains the total apportioned population per hectare for each study site.

3.2 Population per Unit Land Area (without water)

Note:Previous versions of this protocol used the Geoprocessing Wizard and Xtools. ArcGIS 9.2 has replaced the Geoprocessing Wizard with the Intersect Tool, and it offers a Calculate Geometry dialog box, negating a need for Xtools. The protocol detailed below has been tested in ArcGIS 9.2.

1. Basic Concept

In areas with many lakes, water provides an important barrier to walking. As the Twin Cities is one such area, many of our calculations used this population density measure, which is identical to the one above except that it excludes water.

2. Basic Formula
Population per Unit Land Area (without water) = Persons in housing units per unit gross area excluding water.

3. Detailed Formula
Population per Unit Land Area (without water) = Persons in housing units as measured in the US Census at the block level per unit land area excluding area of water features as measured in the Ramsey County layer.

4. Comments and Explanations
Where census blocks are cut by the end of the measurement geography, the population is apportioned according to the percentage of land area falling inside and outside the line.

5. GIS Approach
The approach to this measurement is to intersect census blocks with study sites, preserving the original area of each block and then recalculating it after the spatial intersect procedure. Calculate the apportioned population and divide it by the total area that has been recalculated after removing the area of all water features.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.2**.

1. To follow this protocol you will need:
2. **Study-Site Layer with Water Features removed** (polygon) with unique ID field
3. **Census Block Layer** with field containing population

Create the Census Block Layer with water features removed:
1. Add the Census Block Layer and Water Features layer to your Map Document
2. Open ArcToolbox and navigate to the Erase tool (Analysis Tools -> Overlay). Open the Erase tool by double-clicking it.
3. In the Input Features drop-down box, select the Census Block layer.
4. In the Erase Features drop-down box, select the Water Features layer.
5. Specify the location and name of the output feature class.
6. Click OK.
7. The output feature class will be generated and added to the map document.

Determine the apportioned census population figure for each site:
1. Add the **Study-Site Layer with Water Features removed** and **Census Block Layer with Water Features removed** to your **Map Document**.
   a. If the **Study-Site Layer with Water Features removed** does not exist, please complete TREC Protocol 2.1 before continuing.
2. Add the area fields to the **Census Block Layer with Water Features removed**
   a. Right-click on the **Census Block Layer** and select **Open Attribute Table**.
   b. Click on the **Option** button and select **Add Field…**
   c. Name the field **Area** and change the Type to **Double**.
      i. If a field named **Area** already exists, change the name in the **Add Field…** dialog box to something else.
   d. Click **OK**.
   e. Click on the **Option** button again and select **Add Field…**
   f. Name the field **Acres** and change the Type to **Double**.

---

i. If a field named **Acres** already exists, change the name in the **Add Field…** dialog box to something else.
g. Click **OK**.
h. Click on the **Option** button again and select **Add Field…**
i. Name the field **Hectares** and change the Type to **Double**.
    i. If a field named **Hectares** already exists, change the name in the **Add Field…** dialog box to something else.
j. Click **OK**.
3. Calculate the values for Area, Acres, and Hectares
   a. For the **Area** field,
      i. Right-click the **Area** field in the **Census Block Layer with Water Features removed** attribute table and select **Calculate Geometry…**
      ii. Set the **Property:** box to **Area**.
      iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**"
      iv. Set the **Units** to the desired type. For the **Area** attribute, square meters, miles or feet is probably appropriate.
      v. Click **OK**.
   b. For the Acres field,
      i. Right-click the **Acres** field in the **Census Block Layer with Water Features removed** attribute table and select **Calculate Geometry…**
      ii. Set the **Property:** box to **Area**.
      iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**".
      iv. Set the **Units** to **Acres US [ac]**.
      v. Click **OK**.
   c. For the Hectares field,
      i. Right-click the **Hectares** field in the **Census Block Layer with Water Features removed** attribute table and select **Calculate Geometry…**
      ii. Set the **Property:** box to **Area**.
      iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**".
      iv. Set the **Units** to **Hectares [ha]**.
      v. Click **OK**.
4. Open **ArcToolbox** by clicking on the icon and navigate to the **Intersect tool** (Analysis Tools -> Overlay -> Intersect). Double-click the **Intersect tool** to open it.
   a. For the **Input Features**, click on the down-arrow and select the **Census Block Layer with Water Features removed** from the list of layers.
   b. Repeat 4.a and select the **Study-Site Layer with Water Features removed**.
   c. Specify a name and location for the output layer in the **Output Feature Class** box.
   d. Set the **JoinAttributes** drop-down box to **ALL**
      i. The ALL option will transfer all the attributes from the input features to the output feature class.
   e. Leave the **XY Tolerance** box empty.
   f. Set the **Output Type** to **INPUT**.
      i. The INPUT option will return intersections with the same geometry type as the Input Feature with the lowest geometry dimension.
   g. Click **OK** to execute the Intersect. The output feature class will be added to the **Map Document**.
5. Open the **Attribute Table** of the clipped **Census Block Layer with Water Features removed** created in the previous step.
   a. Click on **Options** and choose **Add Field…**
   b. Name the field **OrigHect** and change the Type to **Double** and click **OK**.
   c. Right-click on the **OrigHect** field.
   d. Choose **Field Calculator…**
   e. In the Field Calculator, click **Hectares** and click **OK**.
   f. You have now copied the values from **Hectares** to **OrigHect**.
6. Update the **Hectares** attribute in the clipped **Census Block Layer with Water Features removed**
   a. Right-click on the **Hectares** field and select **Calculate Geometry…**

---

      b. Set the **Property:** box to **Area**.
      c. Set the **Coordinate System** to "**Use coordinate system of the data source:**".
      d. Set the **Units** to **Hectares [ha]**.
      e. Click **OK**. This will overwrite the values in the Hectares field with the new correct values (since the shapes have been clipped to the extent of the Study-Site Layer with Water Features removed), yet the original number of hectares is preserved in the OrigHect field.
7. To perform the population apportionment, open the **Attribute Table** of the clipped **Census Block Layer with Water Features removed** again.
      a. Click on **Options** and choose **Add Field**.
      b. Name the field **PopSplit** and make it of type **Double**.
      c. Right-click on the name of the **PopSplit** field.
      d. Choose **Field Calculator…**
      e. In the **Field Calculator**, click **Hectares**, then click the / button, then click **OrigHect**, then click the * button, and finally click the name of the **Population field**.
      f. Click **OK**.
8. The result is a field called **PopSplit** that contains the same proportion of the original census population figure as the proportion of the clipped census block area to the original area.

Perform final calculation
1. In the **Attribute Table** of the clipped **Census Block Layer with Water Features removed**, right-click on the name of the site ID field and choose **Summarize**.
      e. In the box beneath **Choose one or more summary statistics**, click the + next to **PopSplit**.
      f. Check the box next to **Sum**.
      g. Choose a name and location for the **Output Summary Table**.
      h. Click **OK**.
2. Add PopPerHect attribute to **Study-Site Layer with Water Features removed Table**.
      i. Open the **Study-Site Layer with Water Features removed** attribute table.
      j. Click the **Option** button and select **Add Field…**
      k. Add a new field called **PopPerHect** and select type **Double**.
3. Join this **Summary Table** into the **Study-Site Layer Table with Water Features removed** using the **ID** field.
      l. Right-click the **Study-Site Layer with Water Features removed** and select **Joins and Relates -> Join…**
      m. From the box entitled "**1. Choose the field in this layer that the join will be based on:**", choose the unique **ID** field for the Study-Site Layer with Water Features removed.
      n. From the box entitled "**2. Choose the data to join to this layer, or load the table from disk**:", choose the **Summary Table** created in the previouw step.
      o. From the box entitled "**3. Choose the field in the table to base the join on:**", choose the unique **ID** field for the Study-Site Layer with Water Features removed.
      p. Click **OK**.
4. Calculate the population per hectare for the **Study-Site Layer with Water Features removed**.
      a. Open the **Study-Site Layer with Water Features removed** attribute table.
      b. Right-click on this new field and choose **Field Calculator**.
      c. In the **Field Calculator**, click on the **PopSplit** field, then click on the / button, and finally click on the **Hectares** field from the **Study-Site Layer with Water Features removed**.
      d. Click **OK**.

The PopPerHect field now contains the total apportioned population per hectare for each study site.

3.4 Residential Population in Residential Parcels (Residential Density) (without water)

Note:Previous versions of this protocol used the Geoprocessing Wizard and Xtools. ArcGIS 9.2 has replaced the Geoprocessing Wizard with the Intersect Tool, and it offers a Calculate Geometry dialog box, negating a need for Xtools. The protocol detailed below has been tested in ArcGIS 9.2.

---

1. Basic Concept
In areas with many lakes, water provides an important barrier to walking. As the Twin Cities is one such area, many of our calculations used this population density measure, which is identical to the one above except that it excludes water.

2. Basic Formula
Population per Unit Land Area (without water) = Persons in housing units per unit gross area excluding water.

3. Detailed Formula
Population per Unit Land Area (without water) = Persons in housing units as measured in the US Census at the block level per unit land area excluding area of water features as measured in the Ramsey County layer.

4. Comments and Explanations
Where census blocks are cut by the end of the measurement geography, the population is apportioned according to the percentage of land area falling inside and outside the line.

5. GIS Approach
The approach to this measurement is to intersect census blocks with study sites, preserving the original area of each block and then recalculating it after the spatial intersect procedure. Calculate the apportioned population and divide it by the total area that has been recalculated after removing the area of all water features.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.2**.

4. To follow this protocol you will need:
5. **Study-Site Layer with Water Features removed** (polygon) with unique ID field
6. **Census Block Layer** with field containing population

Create the Census Block Layer with water features removed:
8. Add the Census Block Layer and Water Features layer to your Map Document
9. Open ArcToolbox and navigate to the Erase tool (Analysis Tools -> Overlay). Open the Erase tool by double-clicking it.
10. In the Input Features drop-down box, select the Census Block layer.
11. In the Erase Features drop-down box, select the Water Features layer.
12. Specify the location and name of the output feature class.
13. Click OK.
14. The output feature class will be generated and added to the map document.

Determine the apportioned census population figure for each site:
9. Add the **Study-Site Layer with Water Features removed** and **Census Block Layer with Water Features removed** to your **Map Document**.
    a. If the **Study-Site Layer with Water Features removed** does not exist, please complete TREC Protocol 2.1 before continuing.
10. Add the area fields to the **Census Block Layer with Water Features removed**
    a. Right-click on the **Census Block Layer** and select **Open Attribute Table**.
    b. Click on the **Option** button and select **Add Field…**
    c. Name the field **Area** and change the Type to **Double**.
        i. If a field named **Area** already exists, change the name in the **Add Field…** dialog box to something else.

---

d. Click **OK**.
e. Click on the **Option** button again and select **Add Field…**
f. Name the field **Acres** and change the Type to **Double**.
    i. If a field named **Acres** already exists, change the name in the **Add Field…** dialog box to something else.
g. Click **OK**.
h. Click on the **Option** button again and select **Add Field…**
i. Name the field **Hectares** and change the Type to **Double**.
    i. If a field named **Hectares** already exists, change the name in the **Add Field…** dialog box to something else.
j. Click **OK**.

11. Calculate the values for Area, Acres, and Hectares
    a. For the **Area** field,
        i. Right-click the **Area** field in the **Census Block Layer with Water Features removed** attribute table and select **Calculate Geometry…**
        ii. Set the **Property:** box to **Area**.
        iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**"
        iv. Set the **Units** to the desired type. For the **Area** attribute, square meters, miles or feet is probably appropriate.
        v. Click **OK**.
    b. For the Acres field,
        i. Right-click the **Acres** field in the **Census Block Layer with Water Features removed** attribute table and select **Calculate Geometry…**
        ii. Set the **Property:** box to **Area**.
        iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**".
        iv. Set the **Units** to **Acres US [ac]**.
        v. Click **OK**.
    c. For the Hectares field,
        i. Right-click the **Hectares** field in the **Census Block Layer with Water Features removed** attribute table and select **Calculate Geometry…**
        ii. Set the **Property:** box to **Area**.
        iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**".
        iv. Set the **Units** to **Hectares [ha]**.
        v. Click **OK**.

12. Open **ArcToolbox** by clicking on the icon and navigate to the **Intersect tool** (Analysis Tools -> Overlay -> Intersect). Double-click the **Intersect tool** to open it.
    a. For the **Input Features**, click on the down-arrow and select the **Census Block Layer with Water Features removed** from the list of layers.
    b. Repeat 4.a and select the **Study-Site Layer with Water Features removed**.
    c. Specify a name and location for the output layer in the **Output Feature Class** box.
    d. Set the **JoinAttributes** drop-down box to **ALL**
        i. The ALL option will transfer all the attributes from the input features to the output feature class.
    e. Leave the **XY Tolerance** box empty.
    f. Set the **Output Type** to **INPUT**.
        i. The INPUT option will return intersections with the same geometry type as the Input Feature with the lowest geometry dimension.
    g. Click **OK** to execute the Intersect. The output feature class will be added to the **Map Document**.

13. Open the **Attribute Table** of the clipped **Census Block Layer with Water Features removed** created in the previous step.
    a. Click on **Options** and choose **Add Field…**
    b. Name the field **OrigHect** and change the Type to **Double** and click **OK**.
    c. Right-click on the **OrigHect** field.
    d. Choose **Field Calculator…**
    e. In the Field Calculator, click **Hectares** and click **OK**.

---

f. You have now copied the values from **Hectares** to **OrigHect**.
14. Update the **Hectares** attribute in the clipped **Census Block Layer with Water Features removed**
    a. Right-click on the **Hectares** field and select **Calculate Geometry…**
    b. Set the **Property:** box to **Area**.
    c. Set the **Coordinate System** to "**Use coordinate system of the data source:**".
    d. Set the **Units** to **Hectares [ha]**.
    e. Click **OK**. This will overwrite the values in the Hectares field with the new correct values (since the shapes have been clipped to the extent of the Study-Site Layer with Water Features removed), yet the original number of hectares is preserved in the OrigHect field.
15. To perform the population apportionment, open the **Attribute Table** of the clipped **Census Block Layer with Water Features removed** again.
    a. Click on **Options** and choose **Add Field**.
    b. Name the field **PopSplit** and make it of type **Double**.
    c. Right-click on the name of the **PopSplit** field.
    d. Choose **Field Calculator…**
    e. In the **Field Calculator**, click **Hectares**, then click the / button, then click **OrigHect**, then click the * button, and finally click the name of the **Population field**.
    f. Click **OK**.
16. The result is a field called **PopSplit** that contains the same proportion of the original census population figure as the proportion of the clipped census block area to the original area.

Perform final calculation
4. In the **Attribute Table** of the clipped **Census Block Layer with Water Features removed**, right-click on the name of the site ID field and choose **Summarize**.
    q. In the box beneath **Choose one or more summary statistics**, click the + next to **PopSplit**.
    r. Check the box next to **Sum**.
    s. Choose a name and location for the **Output Summary Table**.
    t. Click **OK**.
5. Add PopPerHect attribute to **Study-Site Layer with Water Features removed Table**.
    u. Open the **Study-Site Layer with Water Features removed** attribute table.
    v. Click the **Option** button and select **Add Field…**
    w. Add a new field called **PopPerHect** and select type **Double**.
6. Join this **Summary Table** into the **Study-Site Layer Table with Water Features removed** using the **ID** field.
    x. Right-click the **Study-Site Layer with Water Features removed** and select **Joins and Relates -> Join…**
    y. From the box entitled "**1. Choose the field in this layer that the join will be based on:**", choose the unique **ID** field for the Study-Site Layer with Water Features removed.
    z. From the box entitled "**2. Choose the data to join to this layer, or load the table from disk**:", choose the **Summary Table** created in the previouw step.
    aa. From the box entitled "**3. Choose the field in the table to base the join on:**", choose the unique **ID** field for the Study-Site Layer with Water Features removed.
    bb. Click **OK**.
5. Calculate the population per hectare for the **Study-Site Layer with Water Features removed**.
    a. Open the **Study-Site Layer with Water Features removed** attribute table.
    b. Right-click on this new field and choose **Field Calculator**.
    c. In the **Field Calculator**, click on the **PopSplit** field, then click on the / button, and finally click on the **Hectares** field from the **Study-Site Layer with Water Features removed**.
    d. Click **OK**.

The PopPerHect field now contains the total apportioned population per hectare for each study site.

3.5 Population plus Employment per Unit Land Area (excluding water)
Note: The protocol needs to be revised to reflect the replacement of Geoprocessing Wizard with ArcToolbox

1. Basic Concept
Traditionally, population densities have been calculated in terms of people residing in an area. However, if density has some direct relation to walking (e.g., though there being a critical mass of people walking, via congestion, or via transit viability) then some kind of overall activity density is a better measure than merely residential density. Of course there are many kinds of activities that bring people to a place—e.g., employment, school, tourism, leisure. Of these, employment is a major generator of movement and is also measured at a fine scale—although these very localized data are difficult to access.

A later version of this manual will discuss sources that have used this measure, including Messenger and Ewing (1996, 149-50).

2. Basic Formula
Population plus Employment per Unit Land Area = Persons in housing units plus total employees per unit land area (excluding water).

3. Detailed Formula
Population plus Employment per Unit Land Area = Persons in housing units as measured by the US Census at the block level plus total employees per unit land area (excluding water).

4. Comments and Explanations
Where census blocks are cut by the end of the measurement geography, the residential population is apportioned according to the percentage of land area falling inside and outside the line.

Employment data are quite varied. We changed the data we planned to use during the course of our study.

5. GIS Approach
[Note: this uses employment data in grid cells, the result of a strategy to avoid confidentiality problems with state employment data. We ended up using Dun and Bradstreet point data, however, so will redo these steps in the next version of the protocols.]

Add together the apportioned population and grid-cell-based employment for an area. Divide this figure by the total land area after adjusting for the area of the removed water features.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field
2. **Employment Data** (point)
3. **Census Block Layer** with population data (polygon)
4. **Water Features Layer** (polygon)
5. **XTools** extension (available from the Downloads section of the ESRI Web site)

Refer to the fundamental protocol for calculating Net Land Area (excluding water). That protocol will create a **Study-Site Layer** that excludes water features and has updated area. In the rest of this protocol, any references to the **Study-Site Layer** assume that water features have been removed.

Perform a spatial join to summarize the employment data:
1. Add the **Study-Site Layer** and the **Employment Data Layer** to your **Map Document**.
2. Right-click the **Study-Site Layer** and choose **Joins and Relates→Join**.
3. In the top dropdown box, make sure it is set to **Join data from another layer based on spatial location**.
4. In the next item down, select the **Employment Data Layer** to be the layer to join to the study sites.
5. Select **Each polygon will be given a summary of the numeric attributes** and check the box below next to **Sum**.
6. At the bottom, choose a name and location for the output file.
7. Click on **OK**.

The resulting file will have a field in its **Attribute Table** called **Sum_EMPL** or similar (depending on the name of the field containing the number of employees in the original employment layer). The new layer should have the same number of features as the **Study-Site Layer**.

Determine the apportioned census population figure for each site:
1. Add the **Census Block Layer** to your **Map Document**.
2. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres and Hectares** (follow instructions included with the extension for **XTools** installation).
   a. Under **Select layer to measure**, select the **Census Block Layer**.
   b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
   c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected).
   d. Click **OK**.
3. New fields called **Area** and **Hectares** will be created in the **Attribute Table** of your new census block polygon layer with the areas of each polygon calculated in the units you chose (Area) and in hectares (Hectares).
4. Open the **GeoProcessing Wizard**.
   a. Choose **Intersect two layers** and click **Next**.
   b. Select your census blocks as the input layer.
   c. Select your **Study-Site Layer** as the overlay layer .
   d. Choose a name and location for the output layer.
   e. Click **Finish**.
   f. The new layer will be added to the **Map Document**.
5. Open the **Attribute Table** of the **Census Block Layer** created in the previous step.
   a. Click on **Options** and choose **Add Field**.
   b. Name the field **OrigHect** and make it of type **Double**
   c. Back in the **Attribute Table**, right-click on the name of the **OrigHect** field.
   d. Choose **Calculate Values**.
   e. In the **Field Calculator**, double-click **Hectares** and click **OK**.
   f. You have now copied the values from **Hectares** to **OrigHect**.
6. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres and Hectares** (follow instructions included with the extension for **XTools** installation).
   a. Under **Select layer to measure**, select the clipped **Census Block Layer** created in step 4.
   b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters** or **Feet**, for example)
   c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected)
   d. Click **OK**.

---

7. This will overwrite the values in the **Hectares** field with the new correct values (since the shapes have been clipped to the extent of the **Study-Site Layer**), yet the original number of hectares is preserved in the **OrigHect** field.
8. To perform the population apportionment, open the **Attribute Table** of the clipped **Census Block Layer** again.
   a. Click on **Options** and choose **Add Field**.
   b. Name the field **PopSplit** and make it of type **Double**.
   c. Back in the **Attribute Table**, right-click on the name of the **PopSplit** field.
   d. Choose **Calculate Values**.
   e. In the **Field Calculator**, double-click **Hectares**, then click the / button, then double-click **OrigHect**, click the * button, and finally double-click the name of your **Population** field.
   f. Click **OK**.
9. The result is a field called **PopSplit** that contains the same proportion of the original census population figure as the proportion of the clipped census block area to the original area.

Perform final calculation
1. In the **Attribute Table** of the clipped **Census Block Layer**, right-click on the name of the site ID field and choose **Summarize**.
   a. In the box beneath **Choose one or more summary statistics**, click the + next to **PopSplit**.
   b. Check the box next to **Sum**.
   c. Choose a name and location for the **Output Summary Table**.
   d. Click **OK**.
2. Join the **Summary** Table of **Apportioned Population** and the **Summary Table of Employment** into the **Study-Site Layer Table** using the **ID** field.
   a. Add a new field of type **Double** that will contain the result of the population plus employment per hectare calculation.
   b. Right-click on this new field and choose **Calculate Values**.
   c. In the **Field Calculator**, click the **)** button, double-click on the **PopSplit** field, click the + button, double-click the field containing the **sum of employees**, click the **)** button, click the / button, and finally double-click on the **Hectares** field from the **Study-Site Layer**.
   d. Click **OK**.

The field you just calculated now contains the sum of apportioned population and total employment divided by the study site's area in hectares (excluding water).

3.6 Employment per Unit Land Area

1. Basic Concept

While residential population is the most often used density measure, and an overall activity density is likely important, a number of researchers have propose that employment density may be associated with walking. In downtown areas this is obvious as congestion makes automobiles less viable for transportation. In terms of the residential populations in the Twin Cities Walking Study, however, employment density in these neighborhoods is more likely a proxy of mixed use.

A later version of this manual will address prior studies using various employment density measures, although many were only of land in parcels, e.g., Cervero and Kockelman (1996, 8), Ewing (1995, 55), and Frank and Pivo (1994, 46-48).

2. Basic Formula

Employment per Unit Land Area = Total employees per unit land area (excluding water).

3. Detailed Formula

Employment Per Unit Area = Total employees per unit land area (excluding water).

4. Comments and Explanations
Area of water features is excluded in the calculation of area. The methods explained below are
narrowly tailored to the nature of the employment data we purchased.

Employment is from Dun and Bradstreet with some checking of data.

5. GIS Approach
The approach to this measurement is to perform a spatial join between the employment data
and the study sites for which employment density is needed. When doing the spatial join,
calculate the sum of all employment. Finally, divide the total employment for each study site by
its area.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.2**.

To follow this protocol you will need:
    1. **Study-Site Layer** (polygon) with unique ID field
    2. **Employment Data Layer** (point)
    3. **Water Features Layer** (polygon)

Determine Net Land Area (without water) for each Site:
    1.   Add the Study-Site Layer and Water Layer to your Map Document.
    2.   Open ArcToolbox by clicking on the icon and navigate to the Erase tool (Analysis Tools ->
         Overlay -> Erase). Double-click the Erase tool to open it.
           a.   In the Input Features drop-down box, select the Study-Site layer.
           b.   In the Erase Features drop-down box, select the Water Features layer.
           c.   Specify a name and location for the output layer.
           d.   Click OK.
           e.   The output feature class will be generated and added to the map document.
                                                    2.   Calculate the Net Land Area for the new
                                                         Study-Site layer (without water).
           a.   Right-click on the new Study-Site layer (without water) and select Open Attribute Table.
           b.   Click on the Option button and select Add Field…
           c.   Name the field Area and change to type to Double.
                  i.   If a field named Area already exists, change the name in the Add Field… dialog
                       box to something else.
           d.   Click OK.
           e.   Click on the Option button again and select Add Field…
           f.   Name the field Hectares and change the Type to Double.
                  i.   If a field name Hectares already exists, change the name in the Add Field…
                       dialog box to something else.
                                                    3.   Calculate the values for Area and
                                                         Hectares.
           a.   For the Area field,
                  i.   Right-click the Area field in the new Study-Site layer (without water) attribute
                        table and select Calculate Geometry…
                  ii.   Set the Property: box to Area.
                  iii. Set the Coordinate System to "Use coordinate system of the data source:"
                  iv.   Set the Units to the desired type. For the Area attribute, square meters, miles or
                        feet are probably appropriate.
                  v.   Click OK.
           b.   For the Hectares field,
                  i.   Right-click the Hectares field in the new Study-Site layer (without water) and
                        select Calculate Geometry…
                  ii.   Set the Property: box to Area.

iii. Set the Coordinate System to "Use coordinate system of the data source:".
iv. Set the Units to Hectares [ha].
v. Click OK.

Find the employment data points within the study-sites (without layer)
1. Add the Study-Site layer (without water) and the Employment Data layer to an ArcMap Document.
2. Right-click on the Study-Site layer (without water) and choose Joins and Relates -> Join.
3. Under the drop-down box labeled What do you want to join to this layer? choose Join data from another layer based on spatial location.
4. In Step 1 of the dialog box, choose the Employment Data layer.
5. In Step 2, click the button to choose Each polygon will be given a summary of the numeric attributes… and also click the check-box next to Sum.
6. In Step 3, choose a name and location for the output layer.
7. Click OK.
8. The output layer now contains a Count field with the number of Businesses found within each Study Site. The Sum_Employment attribute contains the sum of employment for those businesses within each Study Site.

Calculate the employment per unit land area
1. Add an attribute to hold the Employment per Unit Land Area value to the Study-Site Layer (without water).
    a. Right-click on the Study-Site Layer (without water) and select Open Attribute Table.
    b. Click the Options button and select Add Field…
    c. Name the field EmpPerHect and change the Type to Double.
    d. Click OK.
2. Join the output layer from the previous procedure to the Study-Site Layer (without water).
    a. Right-click on the Study-Site Layer (without water) and choose Joins and Relates -> Join.
    b. Under the dropdown box labeled What do you want to join to this layer? choose Join attributes from a table.
    c. Choose the unique ID for each study site as the field to base the join on.
    d. Choose the output layer from the previous procedure as the Table to Join to this layer.
    e. Choose the unique ID for each study site as the field to base the join on.
    f. Click OK.
3. Perform the final calculation
    a. Right-click the EmpPerHect field in the Study-Site Layer (without water) and select Field Calculator…
    b. In the Field Calculator, click the field Sum_Employment, click the / button, and then click the Hectares field.
    c. Click OK.

The EmpPerHect field now contains the number of employees per hectare (excluding the water of water features).

3.7 Housing Units per Unit Land Area (census data) (without water)
Note: The Geoprocessing Wizard has been replaced with the Intersect Overlay Tool in the ArcMap Toolbox and the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
Housing units per hectare are a proxy for population density. Housing units rather than population, however, are what are regulated by such mechanisms as zoning.

This measure uses census housing units. It would also be possible to use housing unit numbers from parcel data, and this might well be more accurate toward the end of the intercensal period and in areas with a great deal of construction.

In the next version of the manual we will discuss earlier studies that used this measure, e.g., Krizek (2003; APA 269); and McNally and Ryan (1993, 70).

2. Basic Formula
Housing Units per Unit Land Area = Total housing units as counted by the census per unit land area (excluding water).

3. Detailed Formula
Housing Units per Unit Land Area = Total housing units as counted by the census, including both occupied and unoccupied units, per unit land area (excluding water).

4. Comments and Explanations
Where census blocks are cut by the end of the measurement geography, the number of housing units is apportioned according to the percentage of land area falling inside and outside the line.

5. GIS Approach
The approach to this measure is to intersect census blocks with the measurement geography and apportion the housing units to the measurement geographies. Divide the housing units by the total area.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field (results of protocol 2.1 – Net Land Area)
2. **Census Block Layer** (polygon) with field containing housing units
3. **XTools** extension (available from the Downloads section of the ESRI Web site)

Determine the apportioned census population figure for each site:
1. Add the **Study-Site Layer** and **Census Block Layer** to your **Map Document**.
2. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
   a. Under **Select layer to measure**, select the **Census Block Layer**.
   b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
   c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected).
   d. Click **OK**.
3. New fields called **Area** and **Hectares** will be created in the **Attribute Table** of your new census block polygon layer with the areas of each polygon calculated in the units you chose (Area) and in hectares (Hectares).
4. Open the **GeoProcessing Wizard**.
   a. Choose **Intersect two layers** and click **Next**.
   b. Select your census blocks as the input layer.
   c. Select your **Study-Site Layer** as the overlay layer.
   d. Choose a name and location for the output layer.
   e. Click **Finish**.
   f. The new layer will be added to the **Map Document**.
5. Open the **Attribute Table** of the **Census Block Layer** created in the previous step.

---

a. Click on **Options** and choose **Add Field**.
b. Name the field **OrigHect** and make it of type **Double**.
c. Back in the **Attribute Table**, right-click on the name of the **OrigHect** field.
d. Choose **Calculate Values**.
e. In the **Field Calculator**, double-click **Hectares** and click **OK**.
f. You have now copied the values from **Hectares** to **OrigHect**.

6. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
a. Under **Select layer to measure**, select the clipped **Census Block Layer** created in step 4.
b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected).
d. Click **OK**.

7. This will overwrite the values in the **Hectares** field with the new correct values (since the shapes have been clipped to the extent of the **Study-Site Layer**), yet the original number of hectares is preserved in the **OrigHect** field.

8. To perform the housing unit apportionment, open the **Attribute Table** of the clipped **Census Block Layer** again.
a. Click on **Options** and choose **Add Field**.
b. Name the field **HUSplit** and make it of type **Double**.
c. Back in the **Attribute Table**, right-click on the name of the **HUSplit** field.
d. Choose **Calculate Values** in the **Field Calculator**, double-click **Hectares**, then click the / button, then double-click **OrigHect**, click the * button, and finally double-click the name of your **Housing Units** field.
e. Click **OK**.

9. The result is a field called **HUSplit** that contains the same proportion of the original census population figure as the proportion of the clipped census block area to the original area.

3.8 Lot Coverage (without water)

Note: The Geoprocessing Wizard has been replaced with the Intersect Overlay Tool in the ArcMap Toolbox and the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
This is a measure of perceived density and is literally how much ground is covered with buildings.

2. Basic Formula
Lot Coverage = Building footprint area divided by area in parcels, excluding vacant or agricultural land uses.

3. Detailed Formula
Same as above.

4. Comments and Explanations
Where buildings and/or parcels are cut by the end of the measurement geography, just slice them. This is an overall coverage measure for an entire measurement geography.

5. GIS Approach
Intersect building footprints with study sites and calculate area for the resulting layer. Perform a definition query on the parcels layer to exclude vacant and agricultural areas, then intersect it

with the measurement geographies and calculate area for the resulting layer. Summarize the tables from the two layers resulting from the previous intersect procedures by measurement geography ID and generate the sum of the areas. Divide the sum of building footprint hectares by the sum of parcel areas for each measurement geography.
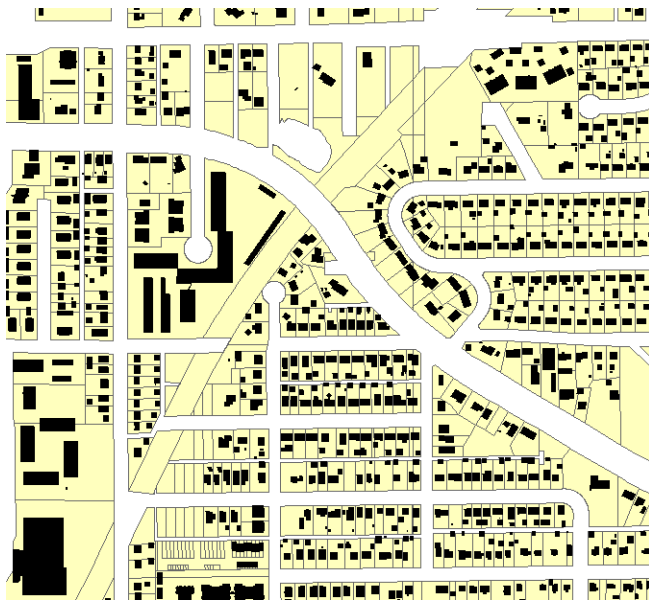
Fig. 3.1: Typical building footprint map.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field
2. **Parcels Layer** (polygon)
3. **Building Footprints Layer** (polygon)
4. **XTools** extension (available from the Downloads section of the ESRI Web site)

Determine total area of building footprints for each site:
1.   1.   Add the **Study-Site Layer** and **Building Footprints Layer** to your **Map Document**.
2.   Open the **GeoProcessing Wizard**.
    a.   Choose **Intersect two layers** and click **Next**.
    b.   Select your building footprints as the input layer.
    c.   Select your **Study-Site Layer** as the overlay layer .
    d.   Choose a name and location for the output layer.
    e.   Click **Finish**.
    f.   The new layer will be added to the **Map Document**.
3.   From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
    a.   Under **Select layer to measure**, select the layer created in the previous step.
    b.   Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
    c.   If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected).
    d.   Click **OK**.
4.   New fields called **Area** and **Hectares** will be created in the **Attribute Table** of your new building footprints polygon layer with the areas of each polygon calculated in the units you chose (Area) and in hectares (Hectares).

Determine total area of parcels for each site:
1. Add the **Study-Site Layer** and **Parcels Layer** to your **Map Document**.
2. Open the **GeoProcessing Wizard**.
    a. Choose **Intersect two layers** and click **Next**.
    b. Select your parcels as the input layer.
    c. Select your **Study-Site Layer** as the overlay layer.
    d. Choose a name and location for the output layer.
    e. Click **Finish**.
    f. The new layer will be added to the **Map Document**.
3. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
    a. Under **Select layer to measure**, select the layer created in the previous step.
    b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
    c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected).
    d. Click **OK**.
4. New fields called **Area** and **Hectares** will be created in the **Attribute Table** of your new parcel polygon layer with the areas of each polygon calculated in the units you chose (Area) and in hectares (Hectares).

At this point you have a new **Parcels Layer** and a new **Building Footprints Layer**. Both have been clipped to the extent of your study sites, and each feature in the two new layers should have a field containing the unique ID of the study site in which it is located.

For each of the two new layers, do the following:
1. Open the layer's **Attribute Table**.
2. Right-click on the name of the field containing the unique ID of the study site.
3. Choose **Summarize**.
4. In the box beneath **Choose one or more summary statistics**, click the + next to **Hectares**.
5. Check the box next to **Sum**.
6. Choose a name and location for the **Output Summary Table**.
7. Click **OK**.

Final calculation
1. Join the **Summary Table of Building Footprint Area** to the **Summary Table of Parcel Area** based on the study site **ID** field.
2. Add a field of type **Double** named **LotCover** or similar (this will contain the final desired result).
3. Right-click on the name of this new field and select **Calculate Values**.
4. Click **Yes** to bypass the question about not being in an edit session.
5. In the **Field Calculator**:
    a. Using the list of available fields, double-click on the **Sum_Hectares** field from the **Building Footprint Summary Table**.
    b. Click on the / button to the right.
    c. Double-click on the **Sum_Hectares** field from the **Parcel Area Summary Table**.
    d. Click **OK** to perform the calculation.

The result is a table with one record for each study site that also contains the lot coverage (building footprints area divided by parcel area). This table can be joined into the **Study-Site Layer** and the values copied over into a new field if desired.

The following variables are not proposed to be used in the study, so only a detailed formula is included.

3.9 Housing Units per Unit Land Area (parcel data)—Not Used in Study

Detailed Formula
Housing Units per Unit Land Area (parcel data) = Total housing units in parcels in parcel data, including occupied and unoccupied units of all sorts, per unit land area, including land inside and outside parcels *but excluding water). Where buildings and/or parcels are cut by the end of the measurement geography, just slice them. This is an overall density measure for an entire measurement geography. Care needs to be taken with group quarters.

3.10 Housing Units per Unit Residential Land Area—Not Used in Study

Detailed Formula
Housing Units per Unit Residential Land Area = Total housing units as counted by the census per area of land in residential parcels (excluding water). Where census blocks are cut by the end of the measurement geography, the number of housing units is apportioned according to the percentage of land area falling inside and outside the line.

Note: this is the housing unit equivalent of Residential Population in Residential Parcels. Later drafts of this manual will relate this variable to dwellings per net area in Ewing (1995, 55).

3.11 Households per Unit Land Area (excluding water)—Not Used in Study

Detailed Formula
Households per Unit Land Area (excluding water) = Total households as counted by the census in blocks per unit land area (excluding water). Where census blocks are cut by the end of the measurement geography, the population is apportioned according to the percentage of land area falling inside and outside the line. See, e.g., Rutherford et al. (2000, 15).

4. PEDESTRIAN INFRASTRUCTURE

By Ann Forsyth (drafting, conceptual issues and formulae, some comments and GIS approach text), Joel Koepp (GIS Steps and the rest of the GIS approach/comments text), and David Van Riper (updates to version 4.1).

Between versions 2.0 and 2.2, there were significant updates to protocols 4.3, 4.6, 4.7, 4.8, 4.10, and 4.11. In version 3.0 we corrected an error in protocol 4.1. In 3.1 we added the description of approach in 4.9. In 4.1 we updated transit stop protocols using ArcGIS 9.2.

Background

Pedestrian infrastructure includes the built and planted features that provide pedestrian amenities or that affect pedestrian mobility, safety, and comfort. From marked pedestrian crossings to street trees, these are more open to modification than the basic street pattern. This dimension matters because:

- Pedestrians are exposed to the outdoor elements—amenities can make walking much more enjoyable.

- Pedestrians have unique safety concerns, some of which are affected by design features, e.g., streetlights, pedestrian crossings.

- The street pattern and pedestrian infrastructure, such as sidewalks, do not always coincide—some areas with gridded street patterns do not have sidewalks; some areas with superblocks and culs-de-sac have pedestrian path and trail networks (Forsyth 2002). Examining pedestrian infrastructure can uncover some of this complexity.

Measurements of this dimension, however, are the least developed because data are incomplete or unavailable, so measurement includes data collection as well as analysis. Some studies assume that more "traditional" neighborhoods with gridded streets and high connectivity possess good pedestrian infrastructure. More detailed measures, however, have been developed based on field observations, interviews, and GIS based indicators, such as completeness of the sidewalk system (Hess et al.. 1999).

4.1 Sidewalk Length per Unit Area

Note: The protocol needs to be revised to reflect the replacement of Geoprocessing Wizard with ArcToolbox and add steps to compute the Length per Unit Area.
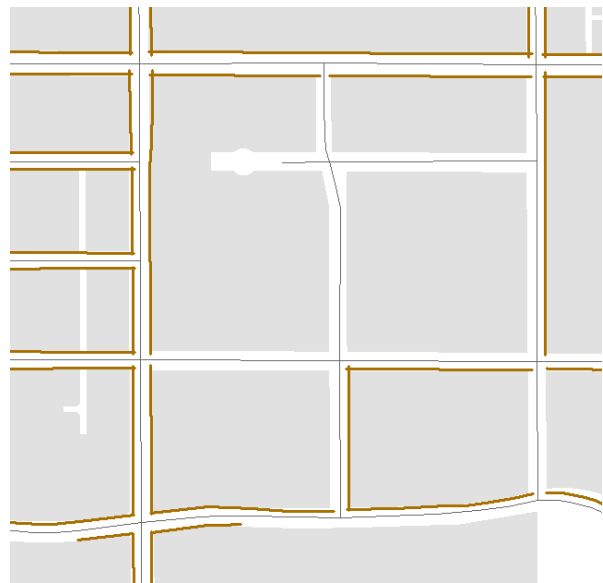
Fig.4.1. Sidewalks in a neighborhood along some streets and not others

1. Basic Concept
Sidewalks are thought to support walking. Intuitively, this is likely to be the case on busy roads, although the relationship is less clear in the case of slower roads. Sidewalk length in an area also reflects the density of streets compared with the ratio of sidewalks to roads. One could imagine a moderate value on this measure with both a very dense street network with sidewalks on some streets or a street pattern with larger blocks but a complete sidewalk system. Both may promote walking.

2. Basic Formula
Sidewalk Length in an Area = Sum together the length of all sidewalk segments in a particular area and divide by the area.

3. Detailed Formula
Sidewalk Length in an Area = Total sidewalk length of all sidewalks in area where they are located along public rights of way divided by the area of the measurement geography. This excludes sidewalks in parks, informal paths, and cuts-through.

4. Comments and Explanations
Sidewalk data are inconsistently collected. As explained in chapter 1, we digitized all sidewalks from high-resolution orthophotos of the area and field checked the data.

5. GIS Approach
The GIS approach is to perform a spatial intersect between the sidewalks and the study sites. Open the attributes of the resulting layer and summarize the data using the unique study site ID field. When doing the table summary, calculate the Sum of the Length field. The new table created will be the sum of the length of all sidewalk segments for each study site.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study -Site Layer** (polygon) with unique ID field
2. **Sidewalks Layer** (polyline)
3. **XTools** extension (available from the Downloads section of the ESRI Web site)

Add the **Study -Site Layer** and **Sidewalks Layer** to your **Map Document**.
1. Open the **GeoProcessing Wizard**.
    a. Choose **Intersect two layers** and click **Next**.
    b. Select your sidewalks as the input layer.
    c. Select your **Study-Site Layer** as the overlay layer.
    d. Choose a name and location for the output layer.
    e. Click **Finish**.
    f. The new layer will be added to the **Map Document**.
2. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
    a. Under **Select layer to measure**, select the clipped **Sidewalks Layer** created in the previous step.
    b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
    c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now.
    d. Click **OK**.
3. Open the **Attribute Table** of the new **Sidewalks Layer**.

4.2 Sidewalk Length Divided by Road Length

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
More complete and continuous sidewalks are thought to support walking. This measure focuses on the proportion of the road network that has sidewalks.

The next version of these protocols will discuss the previous uses of this measure, e.g., Hess et al. 1999, 10.

2. Basic Formula
Sidewalk Length Divided by Road Length = Sum together the length of all sidewalk segments along public rights of way and divide by the length of all road centerlines.

3. Detailed Formula
Sidewalk Length Divided by Road Length = Sum together the length of all sidewalk segments along public rights of way and divide by the length of all road centerlines for roads where pedestrians have access, i.e., excluding interstates and other limited access roads. Divided roads are counted as one.

4. Comments and Explanations
Where sidewalks and/or roads are cut by the end of the measurement geography, just slice them. This is an overall ratio measure for an entire measurement geography.

The results of this calculation can vary based on which definition of "roads" is used, i.e., whether or not freeway entrance/exit ramps are included, length of divided roads is averaged, etc. See the protocol for Road Length per Hectare for more details.

The area that we studied did not have private roads.

5. GIS Approach
Use the GeoProcessing Wizard to intersect sidewalks with measurement geographies. Use XTools to recalculate the lengths of all sidewalk segments. Finally, summarize the intersected sidewalks by measurement geography ID, calculating the sum of the Length field in the process.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field

---

2. **Road Features Layer** (polyline)
3. **Sidewalk Layer** (polyline)
4. **XTools** extension (available from the Downloads section of the ESRI Web site)

Refer to the earlier protocol 4.1 on calculating sidewalk length. Once a **Summary Table of Sidewalk Length** by study site exists, perform calculation of road length.

Add the **Roads Layer** and **Study-Site Layer** to a new **ArcMap Document**:
1. In **Arc Toolbox**, go to **Analysis Tools→Overlay** and select **Intersect**.
2. Under **Input Features**, select both the **Roads Layer** and **Study-Site Layer**.
3. Under **Output Feature Class**, choose a name and location for the output layer.
4. Click **OK**.
5. The new layer will be added to the **Map Document**.
6. In **Arc Toolbox**, go to **Data Management Tools→Generalization** and select **Dissolve**.
7. For **Input Features**, select the layer created in step 3.
8. Under **Output Feature Class**, choose a name and location for the output file.
9. Under **Dissolve_Field(s)**, check the box for the field containing the unique ID of your study sites.
10. Click **OK**.
11. The resulting file should consist of road segments and have the same number of features as you have study sites.

To complete the calculation, divide the length of all sidewalks in a study site (from the **Summary Table** created earlier) by the length of all roads in that study site.

4.3 Sidewalk Length per Length of Major Road

Note: The Geoprocessing Wizard has been replaced with the Intersect Overlay Tool in the ArcMap Toolbox and the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
The basic idea behind this measure is that sidewalks are more important on major roads than on side streets. Ann Vernez Moudon (2004) and her team have used the measure.

2. Basic Formula
Sidewalk Length Divided by Length of Major Road = Length of sidewalks along major roads only divided by the length of those major roads in an area.

Fig. 4.2.. Major roads (collectors and arterials) differentiated from other roads. These roads often have attractive destinations, but also problems of traffic safety.

3. Detailed Formula
Sidewalk Length Divided by Length of Major Road = Using major roads only, sum together the length of all sidewalk segments along public rights of way on those roads and divide by the length of all road centerlines after removing roads with no pedestrian access and averaging the length of divided roads.

4. Comments and Explanations

As Vernez Moudon had a different road file, we defined major roads as roads larger than neighborhood streets but smaller than interstates and US highways (four-lane divided, limited access roads not available to pedestrians). This is possible using the F_CLASS field in the roads data. Selecting roads of F_CLASS 'A30' and 'A25' will identify the collectors and arterials to be used as major roads. We manually removed certain roads within this group known to be four-lane divided with no pedestrian access.

5. GIS Approach
Select features with F_CLASS = 'A25' OR 'A30' from the roads file, and average the lengths of divided roads to create a major roads layer. Intersect the major roads with measurement geographies. Summarize the Attribute Table of the resulting layer and calculate the length of all major roads in each study site. Use Select by Location to identify features in the sidewalks layer that are within a certain distance (e.g., 10 meters) of the major roads and save them as a new layer. Use the GeoProcessing Wizard to intersect the sidewalks along major roads with measurement geographies. Use XTools to recalculate the lengths of all sidewalk segments. Summarize the attribute table of intersected sidewalks by measurement geography ID, calculating the sum of the Length field in the process. Divide each site's sidewalk length by the length of its major roads.

6. GIS Steps
The following procedures were completed using ArcMap (ArcInfo) 8.3.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field
2. **Sidewalk Layer** (polyline)
3. **Roads Layer** (polyline)
4. **XTools** extension (available from the Downloads section of the ESRI Web site)

Calculate length of major roads:
1. Right-click on the **Roads Layer** and choose **Data→Export Data** to make a copy of the **Roads Layer** for the purpose of editing.
2. Choose a location for the output file and name it **Major_Roads** or similar.
3. Click **Yes** to add the new layer to the view.
4. Go to the **Editor** toolbar and select **Editor→Start Editing**.
   a. Select the directory that contains the **Major Roads Layer** and click **OK**.
   b. Open the **Selection** menu and choose **Select by Attributes**.
   c. Select the major **Roads Layer** in the top dropdown box.
   d. Make sure the method is **Create** a new selection.
   e. Type out an expression that will identify non-walkable roads (freeways and on/off ramps), such as **F_CLASS <= 'A25'** or **F_CLASS > 'A30'**.
   f. Double check that all the roads intended to be removed are selected, then click the **Delete** key to remove them.
   g. For every divided road, use the **Edit** tool from the **Editor Toolbar** to select one of the two segments and delete it.
   h. When all edits are complete, go to the **Editor Toolbar** and choose **Editor→Stop Editing**, clicking **Yes** when prompted to save edits.
5. Open the **GeoProcessing Wizard**.
   a. Choose **Intersect two layers** and click **Next**.
   b. Select your major roads as the input layer.
   c. Select your **Study-Site Layer** as the overlay layer.
   d. Choose a name and location for the output layer.
   e. Click **Finish**.
   f. The new layer will be added to the **Map Document**.

---

6. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
   a. Under **Select layer to measure**, select the **Intersected Major Roads Layer** created in the previous step.
   b. Under **Desired output units**, choose **Meters**.
   c. Click **OK**.
7. Open the **Attribute Table** of the **Intersected Major Roads Layer**.
   a. Right-click on the name of the unique study site ID field and choose **Summarize**.
   b. Click the + to expand the **Length** field and check the box next to **Sum**.
   c. Select a name and location for the **Output Summary Table**.
   d. Click **OK**.

This **Summary Table** contains records for all the study sites, along with a **Sum_Length** field that contains the total length of all major roads in the study site.

Add the **Study-Site Layer** and **Sidewalks Layer** to your **Map Document**:
1. Open the **Selection** menu and choose **Select by Location**.
   a. In the top dropdown box, **choose select features from**.
   b. Check the box next to the **Sidewalks Layer**.
   c. In the second dropdown box, choose **are within distance of**.
   d. In the last dropdown box, choose the **Major Roads Layer**.
   e. Check the box next to **Apply a buffer to the features** and enter a distance and the units (e.g., **10 meters**).
   f. Click **Apply**.
   g. In the **Table of Contents**, right-click on the **Sidewalks Layer** and choose **Data→Export Data**
   h. Save the selected features to a new layer named **Major_Sidewalks** or similar.
2. Open the **GeoProcessing Wizard**.
   a. Choose **Intersect two layers** and click **Next**.
   b. Select your major sidewalks as the input layer.
   c. Select your **Study-Site Layer** as the overlay layer.
   d. Choose a name and location for the output layer.
   e. Click **Finish**.
   f. The new layer will be added to the **Map Document**.
3. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
   a. Under **Select layer to measure**, select the **Sidewalk Layer** created in the previous step.
   b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example).
   c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now.
   d. Click **OK**.
4. Open the **Attribute Table** of the new major **Sidewalk Layer**.
   a. Right-click on the name of the unique study site ID field and choose **Summarize**.
   b. Click the + to expand the Length field and check the box next to **Sum**.
   c. Select a name and location for the **Output Summary Table**.
   d. Click **OK**.

This **Summary Table** contains records for all the study sites, along with a **Sum_Length** field that contains the total length of all sidewalks along major roads in the study site.

To complete the calculation, use the two summary tables created using these instructions. For each study site, divide the length of sidewalks along the major roads by the length of all major roads.

4.4 Streetlights per Length of Road
Note: The Geoprocessing Wizard has been replaced with the Intersect Overlay Tool in the ArcMap Toolbox and the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol

---

detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
Safety from accidents and crime is an important issue for pedestrians. Lighting improves both forms of safety at night.

The next version of these protocols will review previous uses of related variables, e.g., Cervero and Kockelman (1996, 8,16), which examined the average distance between lights.

2. Basic Formula
Streetlights per Length of Road = Number of streetlights divided by total length of road centerlines.

3. Detailed Formula
Streetlights per Length of Road = Number of streetlights (multiple headed fixtures classed as one) divided by total length of road centerlines for roads that are accessible to pedestrians (i.e., not limited access highways).

4. Comments and Explanations
Where roads are cut by the end of the measurement geography, just slice them. This is an overall ratio measure for an entire measurement geography.

5. GIS Approach
Perform a spatial join between the streetlights and the measurement geographies to determine the number of streetlights in each measurement geography. Intersect the

Fig. 4.3. Streetlights in a neighbourhood indicated by dots at the location of the base of the post

roads layer with the measurement geographies. With the resulting layer, open the attributes and summarize the table using the field containing the unique measurement area ID. When summarizing the table, calculate the sum of all fields containing numbers, which will result in an output table with a field containing the length of all road segments. Divide the number of streetlights in each measurement geography by the total road length in that measurement geography.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field
2. **Streetlights Layer** (point)
3. **Roads Layer** (polyline)
4. **XTools** extension (available from the Downloads section of the ESRI Web site)

Add the **Study-Site Layer** and **Streetlights Layer** to your **Map Document**.
1. Perform a **Spatial Join** on your new buffer layer.
   a. Right-click on your new layer and select **Joins and Relates** and then **Join**.
   b. Under **What do you want to join to this layer**, select **Join data from another layer based on spatial location**.

---

c. Under **1. Choose the layer to join to this layer**, select your clipped **Intersection Layer** (point)
d. Under **2. You are joining**, select the first option **Each polygon will be given a summary**.
e. Under **How do you want the attributes to be summarized?** check the box for *Sum*.
f. Under **3. The result of the join will be saved**, give your new layer a name.
g. Click on **OK**.

Add the roads layer to your **Map Document**.
1. Open the **GeoProcessing Wizard**.
    a. Choose **Intersect two layers** and click **Next**.
    b. Select your roads as the input layer.
    c. Select your **Study-Site Layer** as the overlay layer.
    d. Choose a name and location for the output layer.
    e. Click **Finish**.
    f. The new layer will be added to the **Map Document**.
2. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
    a. Under **Select layer to measure**, select the clipped **Roads Layer** created in the previous step.
    b. Under **Desired output units**, choose **Meters**.
    c. If you do not wish to generate a parameter, such as **Perimeter** or **Length** (do not delete **Area** or **Hectares**), deselect it now.
    d. Click **OK**.
3. Open the **Attribute Table** of the new **Roads Layer**.
    a. Right-click on the name of the unique study site ID field and choose **Summarize**.
    b. Click the + to expand the **Length** field and check the box next to **Sum**.
    c. Select a name and location for the **Output Summary Table**.
    d. Click **OK**.

You now have a **Summary Table of Road Length** in meters by study site. Join this table into the table of the layer that resulted from the spatial join of lights to study sites above (based on the site ID field).
1. Add a field of type **Double** to hold the result of your streetlights per meter calculation.
2. Right-click on this new field in the **Attribute Table** and choose **Calculate Values**.
3. In the field selector box, double-click the field containing the **number of streetlights**, then click the / button, and finally double-click the field containing the **length of roads**.
4. Click **OK**.

4.5 Street Trees (trees within an X distance buffer) per Length of Road
Note: The protocol already uses ArcGIS version 9.0 tools, but has not been fully tested in ArcGIS 9.1.

1. Basic Concept
Street trees are thought to improve pedestrian comfort (shade, wind protection) and the attractiveness of the street. While this is a measure per length of road, it is in many ways a density measure weighted by road length, as all trees within a certain area are being counted.

A later version of these protocols will examine the use of related concepts by earlier researchers, e.g., Cervero and Kockelman (1996, 8,), who looked at the proportion of blocks with street trees.

2. Basic Formula
Street Trees (trees within a 15/20 meter buffer) per Length of Road = Number of street trees within a buffer distance of 15/20 meters of the road centerline per length of road.

Fig. 4.4. Only trees found within the road buffer are counted

3. Detailed Formula
The size of the road buffer depends on the
character of the road. For city streets, the buffer is 15 meters from the street centerline. For arterials and larger roads, the buffer is 20 meters from the centerline. These roads were identified by querying the F_CLASS (functional class) field for values less than or equal to A30 (A30 is typically a major arterial, A20 a state or US highway, and A10 an interstate).

4. Comments and Explanations
We digitized trees from high-resolution digital orthophotos. Where trees had a continuous canopy so that marking individual trees was difficult, tree stands were photographed and then sorted into high, medium, and low density. We measured densities of trees in three areas and used these figures, and some research on tree density ranges, to allocate trees into general density ranges with a clear upper limit. The upper limit was so that a very dense stand in one area, which may be perceived similarly to a stand that is much less dense but still has a lot of trees, did not bias the tree density too much. The working density figures are 60 trees per acre for low density, 200 for medium density, and 400 for high density. These figures are subject to review, however, as we refine these measurements. Trees can reach very high densities, well over 1,000 per acre.

5. GIS Approach
The GIS approach to this calculation requires two stages. First, a spatial join between the road buffer and the tree points will yield the count of trees per study site. Next, the road buffer must be intersected with the tree polygons (which represent dense stands of trees) and the count of trees must be determined by calculating the area of the resulting features and multiplying it by the assigned tree density. Finally, the count of tree points and count of trees in polygons must be added together and divided by the area of the road buffer.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon) with unique ID field
2. **Roads Layer** (polyline)
3. **Trees Layer** (point)
4. **Tree-Stands Layer** (polygon) with a field containing the density
5. **XTools** extension (available from the Downloads section of the ESRI Web site)

Prepare roads layer for buffering:
1. Open the **Attribute Table** of the **Roads Layer**.
2. Click on **Options** and choose **Add Field**.
3. Name the field **BuffDist** and make it of type **Short Integer**, then click **OK**.
4. Click on **Options** and choose **Select by Attributes**.
5. Double-click the field **F_CLASS**, then click the **<=** button, and then select **A30** from the unique values.
6. Click **Apply** to select major arterials from the **Roads Layer**.
7. In the **Attribute Table**, right-click the name of the **BuffDist** field and choose **Calculate V**alues.
8. Type in **20** and click **OK** to assign a value of 20 to the selected records.
9. In the **Attribute Table**, click **Options** and choose **Switch Selection**.
10. Right-click the name of the **BuffDist** field and choose **Calculate Values**.

11. Type in **15** and click **OK** to assign a value of 15 to the selected records.

Buffer road features:
1. In **ArcToolbox**, open **Analysis Tools→Proximity**, and select **Buffer**.
2. In the **Buffer** tool window, under **Input Features** select your **Roads Layer**.
3. Under **Output Feature Class**, choose a name and location for the output shapefile.
4. For **Distance**, click **Field** and select the **BuffDist** field from the list.
5. Click **OK**.

Create road buffers for each study site:
1. In **ArcToolbox**, go to **Analysis Tools→Overlay** and select **Intersect**.
2. Under **Input Features**, select both the buffered **Roads Layer** from the previous step and the study sites.
3. Under **Output Feature Class**, choose a name and location for the output layer.
4. Click **OK**.
5. The new layer will be added to the **Map Document**.
6. In **ArcToolbox**, go to **Data Management Tools→Generalization** and select **Dissolve**.
7. For **Input Features**, select the layer created in step 3.
8. Under **Output Feature Class**, choose a name and location for the output file.
9. Under **Dissolve_Field(s)**, check the box for the field containing the unique ID of your study sites
10. Click **OK**.
11. The resulting file should consist of road buffers and have the same number of features as you have study sites.

Count tree points:
1. In the **Table of Contents** pane, right-click on the **road buffers layer** created in the last step and go to **Join and Relates→Join**.
2. Choose to **Join data from another layer based on spatial location**.
3. Below, click the radio button that provides a summary of numeric attributes.
4. Choose a name and location for the output file and click **OK**.

Count trees in tree stands:
1. In **Arc Toolbox**, go to **Analysis Tools→Overlay** and select **Intersect**.
2. Under **Input Features**, select both the buffered **Roads Layer** from the previous step and **Tree Stands Layer**.
3. Under **Output Feature Class**, choose a name and location for the output layer.
4. Click **OK**.
5. The new layer will be added to the **Map Document**.
6. In **Arc Toolbox**, go to **Data Management Tools→Generalization** and select **Dissolve**.
7. For **Input Features**, select the layer created in step 3.
8. Under **Output Feature Class**, choose a name and location for the output file.
9. Under **Dissolve_Field(s)**, check the boxes for the field containing the unique ID of your study sites and the unique ID of the tree stands.
10. Under **Statistics Field(s)**, choose **Density**.
11. Set the **Statistic Type** to be **First**.
12. Click **OK**.
13. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
    a. Under **Select layer to measure**, select the output layer from the dissolve above
    b. Under **Desired output units**, choose your units (depending on the study site size, try **Miles**, **Meters**, or **Feet**, for example)
    c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected)
    d. Click **OK**.

e. New fields called **Area** and **Hectares** will be created in the **Attribute Table** of your new census block polygon layer with the areas of each polygon calculated in the units you chose (Area) and in Hectares (Hectares).
14. Open the **Attribute Table** of the layer of created in the dissolve procedure above.
15. Click on **Options** and choose **Add Field**.
16. Name the field **Trees** and set it to type **Double**, then click **OK**.
17. Right-click on the name of this field and do **Calculate Values**.
18. Double-click the name of the **Density** field (**First_Dens**), then click the * button, finally double-click the **Hectares** field and click **OK**.
19. This will calculate the number of trees in each tree stand based on the density and area.
20. In the **Attribute Table**, right-click on the field containing the study site unique ID and choose **Summarize**.
21. For the **Trees** field, check the box the calculate the summary statistic.
22. Choose a name and location for the output table.
23. This **Summary Table** should contain one record for each study site and the total number of trees in tree stands for each.

Add together the count of tree points and count of trees in tree stands for each study site. Refer to the protocol for calculating road length to calculate it, then divide the total trees for each study site by its total road length.

Note: the next variables use data from an urban design inventory that uses sample data and may need to be modified for 100 percent data.

4.6 Percent of Street Segments with Marked Pedestrian Crossings at One or Both Ends
Note: T*he protocol already uses ArcGIS version 9.0 tools, but has not been fully tested in ArcGIS 9.1.*

1. Basic Concept
Marked pedestrian crossings may help pedestrian movement, though they may also indicate busy streets. This measure is very broad and looks for at least some marked crossings and does not require all potential crossing areas to be marked.

This variable is based on the following inventory question that was filled in for both ends of the segment.

2a. Consider the places on the segment that are **intended** for pedestrians to cross the street. Are these places marked for pedestrian crossing?
All = 3; Some = 2; none = 0; NA = 8 (lines 7 and 211 of inventory)

2. Basic Formula
Percent of Street Segments with Marked Pedestrian Crossings at One or Both Ends = Total street segments in area that was sampled, divided by the number that score 2 or 3 on this variable at one or both ends.

3. Detailed Formula
As above.

4. Comments and Explanations
None.

5. GIS Approach
The GIS approach to creating this variable begins with identifying which sampled road segments have marked pedestrian crossings. Perform a spatial intersect between the sampled road segments and the measurement geography. The output needs to be dissolved using both the unique measurement geography ID and the unique road segment ID (in case the features of the measurement geography overlap each other). The attribute table of the dissolved output must then be summarized using the measurement geography ID field. This will yield a new table containing both a count of the number of road segments with marked crossings and the total number of road segments. Divide the number of segments with marked crossings by the total to determine the percentage.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Roads Layer** (polyline) with unique ID for each road segment
3. **Urban Design Inventory** (table) with road segment IDs

Identify sampled roads with marked pedestrian crossings:
1. Add the sampled **Roads Layer** and **Urban Design Inventory Table** to an **ArcMap Document**.
   a. Right-click on the **Inventory Table** and choose **Joins and Relates→Relate**.
   b. In item **1** of the **Relate** dialog box, choose the field in the **Inventory Table** that contains the unique road segment ID (in our case, Line1).
   c. In item **2**, choose the sampled **Roads Layer**.
   d. In item **3**, select the field from the sampled **Roads Layer** that contains the unique ID (in our case, TLGID).
   e. Click **OK**.
2. Right-click the **Inventory Table** and choose **Open**.
   a. A road segment with a marked crossing at either end will be counted as having marked crossings.
   b. Click **Options** and choose **Select by Attributes**.
   c. Enter the expression **Line7<> 0 OR Line211<> 0** and click **Apply**, then **Close**.
   d. This will select the road segments that have one or more ends with marked crossings.
3. Click **Options** and choose **Related Tables→Relate1** to open the sampled **Roads Layer** and select the matching records.
   a. In the sampled roads **Attribute Table**, select **Options→Add Field**.
   b. Name the field **Crossings** and make it of type **Short Integer**, then click **OK**.
   c. Right-click on the new **Crossings** field and choose **Calculate Values**.
   d. Enter **1** and click **OK** to assign a value of 1 to the highlighted records (sampled roads that have crossings at one or both ends).

Combine sampled roads and measurement geographies.
1. Intersect the roads and measurement geographies.
   a. In **ArcToolbox**, open the **Analysis Tools→Overlay→Intersect tool**.
   b. Select the sampled roads and the measurement geography as the **Input Features**.
   c. Choose a name and location for the **Output Feature Class**.
   d. Click **OK**.
2. Open the **Data Management Tools→Generalization→Dissolve tool**.
   a. Select the output layer from the **Intersect** tool above as the **Input Feature**.
   b. Choose a name and location for the **Output Feature Class**.
   c. Select both the unique measurement geography ID and the unique sampled road ID as the **Dissolve Fields**.

d. Choose **Crossings** as the **Statistics** field and set **the Statistic Type** to **MAX** or **MIN** (it doesn't matter which one, this is just a way of preserving the field and carrying it over to the output file).
e. Click **OK**.

Count roads and roads with crossings
1. Open the **Attribute Table** of the output layer from the **Dissolve** tool process.
   a. Right-click on the field containing the unique ID of the measurement geographies and choose **Summarize**.
   b. In the dialog box that appears, click the + symbol next to the **MAX_Crossings** field and check the box for **Sum** to calculate the sum of that field as a summary statistic.
   c. Choose a name and location for the **Output Summary Table**.
   d. Click **OK**.
   e. Click **Yes** to add the table to the **ArcMap Document**.
2. Open the **Summary Table**.
   a. Click **Options** and choose **Add Field**.
   b. Name it **PctCross** or similar and set it to type **Double**.
   c. Right-click the added field in the table and choose **Calculate Values**.
   d. In the dialog box, enter the expression **[SUM_MAX_Cr] / [Count]** (the field names may be slightly different depending on the names you chose earlier).
   e. Click **OK**.

This divides the number of segments with marked crossings by the total number of sampled segments for each feature in the measurement geographies.

4.7 Percent of Street Segments with Visible Litter, Graffiti, or Dumpsters
Note: The protocol already uses ArcGIS version 9.0 tools, but has not been fully tested in ArcGIS 9.1.

1. Basic Concept
General maintenance and cleanliness is thought to affect attractiveness and a sense of safety for pedestrians.

This variable is based on the following inventory questions.

> 44. How much **graffiti** is apparent on this segment?
> **Some/a lot** = 3; little = 2; none = 0 (line 166)
> 45. How much **litter** is apparent on this segment?
> **Some/a lot** = 3; little = 2; none = 0 (line 167)
> 46. Are there **dumpsters** visible on this segment?
> **Some/a lot** = 3; few = 2; none = 0 (Line 168)

2. Basic Formula
Percent of Street Segments with Visible Litter, Graffiti, or Dumpsters = The total number of inventory segments in an area with at least one instance of "some/a lot" (i.e., score 3) of graffiti, litter, or dumpsters visible, divided by the total number of segments sampled in an area.

3. Detailed Formula
As above.

4. Comments and Explanations
None.

5. GIS Approach

The GIS approach to creating this variable begins with identifying which sampled road segments have visible litter, graffiti, or dumpsters. Perform a spatial intersect between the sampled road segments and the measurement geography. The output needs to be dissolved using both the unique measurement geography ID and the unique road segment ID (in case the features of the measurement geography overlap each other). The attribute table of the dissolved output must then be summarized using the measurement geography ID field. This will yield a new table containing both a count of the number of road segments with visible litter, etc., and the total number of road segments. Divide the number of segments with visible litter, etc., by the total to determine the percentage.

6. GIS Steps

The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Sampled Roads Layer** (polyline) with unique ID for each road segment
3. **Urban Design Inventory** (table) with road segment IDs

Identify sampled roads with visible litter, graffiti, or dumpsters:
1. Add the **Sampled Roads Layer** and **Urban Design Inventory Table** to an **ArcMap Document**.
   a. Right-click on the **Inventory Table** and choose **Joins and Relates→Relate**.
   b. In **Item 1** of the **Relate** dialog box, choose the field in the **Inventory Table** that contains the unique road segment ID (in our case, Line1).
   c. In **Item 2**, choose the sampled **Roads Layer**.
   d. In **Item 3**, select the field from the sampled **Roads Layer** that contains the unique ID (in our case, TLGID).
   e. Click **OK**.
2. Right-click the **Inventory Table** and choose **Open**.
   a. Click **Options** and choose **Select by Attributes**.
   b. Enter the expression **Line166= 3 OR Line167= 3 OR Line168= 3** and click **Apply**, then **Close**.
   c. This will select the road segments that are defined as having visible litter, graffiti, or dumpsters.
3. Click **Options** and choose **Related Tables→Relate1** to open the sampled **Roads Layer** and select the matching records.
   a. In the sampled roads **Attribute Table**, select **Options→Add Field**.
   b. Name the field **Litter** and make it of type **Short Integer**, then click **OK**.
   c. Right-click on the new **Litter** field and choose **Calculate Values**.
   d. Enter **1** and click **OK** to assign a value of 1 to the highlighted records (sampled roads that have visible litter, etc.).

Combine sampled roads and measurement geographies:
1. Intersect the roads and measurement geographies.
   a. In **ArcToolbox**, open the **Analysis Tools→Overlay→Intersect tool**.
   b. Select the sampled roads and the measurement geography as the **Input Features**.
   c. Choose a name and location for the **Output Feature Class**.
   d. Click **OK**.
2. Open the **Data Management Tools→Generalization→Dissolve tool**.
   a. Select the output layer from the **Intersect** tool above as the **Input Feature**.
   b. Choose a name and location for the **Output Feature Class**.
   c. Select both the unique measurement geography ID and the unique sampled road ID as the **Dissolve Fields**.

d. Choose **Litter** as the **Statistics** field and set the **Statistic Type** to **MAX** or **MIN** (it doesn't matter which one, this is just a way of preserving the field and carrying it over to the output file).
e. Click **OK**.

Count roads and roads with visible litter, graffiti, or dumpsters.
1. Open the **Attribute Table** of the output layer from the **Dissolve** tool process.
    a. Right-click on the field containing the unique ID of the measurement geographies and choose **Summarize**.
    b. In the dialog box that appears, click the + symbol next to the **MAX_Litter** field and check the box for **Sum** to calculate the sum of that field as a summary statistic.
    c. Choose a name and location for the output **Summary Table**.
    d. Click **OK**.
    e. Click **Yes** to add the table to the **ArcMap Document**.
2. Open the **Summary Table**.
    a. Click **Options** and choose **Add Field**.
    b. Name it **PctLitt** or similar and set it to type **Double**.
    c. Right-click the added field in the table and choose **Calculate Values**.
    d. In the dialog box, enter the expression **[SUM_MAX_Li] / [Count]** (the field names may be slightly different depending on the names you chose earlier).
    e. Click **OK**.

This divides the number of segments with visible litter, graffiti, or dumpsters by the total number of sampled segments for each feature in the measurement geographies.

4.8 Percent of Street Segments with Traffic Calming, Broadly Defined
Note: The protocol already uses ArcGIS version 9.0 tools, but has not been fully tested in ArcGIS 9.1.

1. Basic Concept
This variable is based on the following inventory questions.

51. Are there measures on this segment that could **slow down traffic**? Mark all that apply. Speed bump/speed hump/raised crosswalk; or dips (that are intended to slow down traffic)
yes = 1; no = 0 (line 177)
Rumble strips or bumps (includes dots, reflectors, raised concrete strips, etc.)
yes = 1; no = 0 (line 178)
Curb bulb out/curb extension
yes = 1; no = 0 (line 179)
Traffic circle/roundabout
yes = 1; no = 0 (line 180)
Median
yes = 1; no = 0 (line 181)
Angled/ On-street parking (that runs along most or the entire segment - does not have to be on both sides of segment)
yes = 1; no = 0 (line 183)
Other
yes = 1; no = 0 (line 184)

2. Basic Formula
Percent of Street Segments with Traffic Calming, Broadly Defined = The total number of inventory segments in an area with at least one instance of one of the features above, divided by the total number of segments sampled in an area.

3. Detailed Formula
As above.

4. Comments and Explanations
None.

5. GIS Approach
The GIS approach to creating this variable begins with identifying which sampled road segments have traffic calming. Perform a spatial intersect between the sampled road segments and the measurement geography. The output needs to be dissolved using both the unique measurement geography ID and the unique road segment ID (in case the features of the measurement geography overlap each other). The attribute table of the dissolved output must then be summarized using the measurement geography ID field. This will yield a new table containing both a count of the number of road segments with traffic calming and the total number of road segments. Divide the number of segments with traffic calming by the total to determine the percentage.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Sampled Roads Layer** (polyline) with unique ID for each road segment
3. **Urban Design Inventory** (table) with road segment IDs

Identify sampled roads with visible litter, graffiti, or dumpsters:
1.   Add the **Sampled Roads Layer** and **Urban Design Inventory table** to an **ArcMap Document**.
     a.   Right-click on the **Inventory Table** and choose **Joins and Relates→Relate**.
     b.   In **Item 1** of the **Relate** dialog box, choose the field in the **Inventory Table** that contains the unique road segment ID (in our case, Line1).
     c.   In **Item 2**, choose the sampled **Roads Layer**.
     d.   In **Item 3**, select the field from the sampled **Roads Layer** that contains the unique ID (in our case, TLGID).
     e.   Click **OK**.
2. Right-click the **Inventory Table** and choose **Open**.
     a.   A road segment with a **Value of 1** for at least one of the seven questions will be counted as having traffic calming.
     b.   Click **Options** and choose **Select by Attributes**.
     c.   Enter the expression **Line177 <> 0 OR Line178<> 0 OR Line179 <> 0 OR Line180 <> 0 OR Line181 <> 0 OR Line183 <> 0 OR Line184 <> 0** and click **Apply**, then **Close**.
     d.   This will select the road segments that have at least one element of traffic calming.
3. Click **Options** and choose **Related Tables→Relate1** to open the sampled **Roads Layer** and select the matching records.
     a.   In the sampled roads **Attribute Table**, select **Options→Add Field**.
     b.   Name the field **Calming** and make it of type **Short Integer**, then click **OK**.
     c.   Right-click on the new **Calming** field and choose **Calculate Values**.
     d.   **Enter 1** and click **OK** to assign a **Value of 1** to the highlighted records (sampled roads that have traffic calming).

Combine sampled roads and measurement geographies:
1. Intersect the roads and measurement geographies.
   a. In **ArcToolbox**, open the **Analysis Tools→Overlay→Intersect tool**.
   b. Select the sampled roads and the measurement geography as the **Input Features**.
   c. Choose a name and location for the **Output Feature Class**.
   d. Click **OK**.
2. Open the **Data Management Tools→Generalization→Dissolve tool**.
   a. Select the output layer from the **Intersect** tool above as the **Input Feature**.
   b. Choose a name and location for the **Output Feature Class**.
   c. Select both the unique measurement geography ID and the unique sampled road ID as the **Dissolve Fields**.
   d. Choose **Calming** as the **Statistics** field and set the **Statistic Type** to **MAX** or **MIN** (it doesn't matter which one, this is just a way of preserving the field and carrying it over to the output file).
   e. Click **OK**.

Count roads and roads with visible traffic calming:
1. Open the **Attribute Table** of the output layer from the **Dissolve** tool process.
   a. Right-click on the field containing the unique ID of the measurement geographies and choose **Summarize**.
   b. In the dialog box that appears, click the + symbol next to the **MAX_Calming** field and check the box for **Sum** to calculate the sum of that field as a summary statistic.
   c. Choose a name and location for the output **Summary Table**.
   d. Click **OK**.
   e. Click **Yes** to add the table to the **ArcMap Document**.
2. Open the **Summary Table**.
   a. Click **Options** and choose **Add Field**.
   b. Name it **PctCalm** or similar and set it to type **Double**.
   c. Right-click the added field in the table and choose **Calculate Values**.
   d. In the dialog box, enter the expression **[SUM_MAX_Ca] / [Count]** (the field names may be slightly different depending on the names you chose earlier).
   e. Click **OK**.

This divides the number of segments with traffic calming by the total number of sampled segments for each feature in the measurement geographies.

4.9 Other fieldwork-based measurements of neighborhood identification, street crossing, street characteristics, views, land uses, barriers, sidewalks, bicycle lanes, mid-block crossings, steepness, sidewalk amenities, street trees, buildings, windows, other features of buildings, garages, parking, driveways, maintenance, lighting, freeways, traffic features, and architecture/design.

These will be added in subsequent editions, however, for the purposes of testing the inventory measures we divided the Irvine Minnesota Inventory into dichotomous questions (yes/no), ordinal questions (scale), and categorical questions (where there was an incomplete scale or an NA). At the focus area and buffer level, we calculated the following:
    a. Dichotomous questions: we calculated the average score in an area.
    b. Ordinal questions: we calculated the average score.
    c. Categorical questions: we calculated the percentage of each answer for focus and context areas only.

4.10 Distance to Nearest Transit Stop

Note: Network Analyst is available as of ArcGIS 9.1. The network distance calculation workflow has also changed as of ArcGIS 9.1. The workflow has been revised.

1. Basic Concept

People walk to transit. The distance to the nearest transit stop may be associated with physical activity.

This relates to the TREC-IDEA survey question:
- SS178 About how long would it take to get from your home to the nearest bus stop if you walked? (Adapted from NEWS B20 (14))

2. Basic Definition
Distance to Nearest Facility = Straight line or network distance to the centroid of the nearest store or hardware store.

3. Detailed Definition
As above.

4. Comments and Explanations
This is done using Network Analyst in ArcMap (ArcInfo) 9.2 using a function called Find Closest Facility. You can specify a point layer to use as your "facilities" and another point layer to serve as your origins or "incidents". It will then take the road network and calculate the shortest path to a facility (and it can do this using travel costs as well, if certain type of roads were considered more or less easy to use, for instance).

For points that do not fall on the network, there is an option in the Load Locations dialog box that allows the user to decide how far away from the network to search for points to use in the processing.

5. GIS Approach
The approach for this measure is to load a set of origins or "incidents" (in this case, participant locations) and destinations or "facilities" (transit stops). Then use the Find Closest Facility function in Network Analyst to calculate the network distance from each origin to the nearest transit stop.

6. GIS Steps
The following procedures were completed in ArcMap (ArcInfo) 9.2.

To follow this protocol you will need:
1. Road Network dataset
2. Transit Stop Locations Point Feature Layer for facilities
3. Participants' Locations Point Feature Layer for incidents
4. **Network Analyst Extension** for ArcGIS 9.2

   I. Calculate the Network Distance to the nearest transit stop
1. Start **ArcMap**
2. On the Main Menu click **Tools** and Click on **Extensions.**
   a. Check **Network Analyst** Extension.

---

      b. Click **Close**

3. Click the **Add Data** button ⬇ on the **Standard Toolbar**
      a. Navigate to the **Network Dataset** location.
      b. Click on the **Network Dataset** and click **Add** to add the Network Dataset to ArcMap.
      c. Click **No** to add all the feature classes, which participate in the network, to the map.

4. Click the **Add Data** button ⬇ on the **Standard Toolbar**
      a. Navigate to the **Point Layers** to be used for analysis, in this case **Participant Home location and Transit Stop Locations**
      b. Click on the **Point Layers** and click **Add** to add it to ArcMap.

5. If the **Transit Stop ID's** and **Participant ID's** are already assigned go to next step else, In the Table of Contents Right Click the **Point Layer** and click on **Open Attribute Table**
      a. In the **Attribute Table**
            i. Click **Options**
            ii. Click **Add Field**
            iii. Add the **Tran_ID_T** as Text
                ▪ Type **Tran_ID_T** in the **Name** field
                ▪ Select **Text** as **Type**
                ▪ Set the **Length** to 20
                ▪ Click **OK**
             v. Click Options and select Add Field
            iv. Add the **Tran_ID** as **Integer**
                ▪ Type **Tran_ID** in the **Name** field
                ▪ Select **Short Integer** as **Type**
                ▪ Accept default (0) for **Precision**
                ▪ Click **OK**
            vi. Right-click on the **Tran_ID** field and select **Field Calculator…**
                ▪ Set the **Tran_ID** equal to **[FID] + 1**
                ▪ Click **OK**
            vii. Right-click on the **Tran_ID_T** field and select **Field Calculator…**
                ▪ Set the **Tran_ID_T** equal to **Tran_ID**
                ▪ Click **OK**
      e. Right-click on the **Participant locations** point layer and select **Open Attribute Table**
            iv. In the **Attribute Table**
                ▪ Click the **Options** button and select **Add Field**
                ▪ Add **Par_ID_T** as **Text**
                    a. Type **Par_ID_T** in the **Name** field
                    b. Select **Text** as **Type**
                    c. Set the **Length** to **20**
                    d. Click **OK**
                ▪ Click the **Options** button and select **Add Field**
                ▪ Add **Par_ID** as **Integer**
                    a. Type **Par_ID** in the **Name** field
                    b. Select **Short Integer** as **Type**
                    c. Accept **default (0)** for **Precision**
                    d. Click **OK**
                ▪ Right-click on the **Par_ID** field and select **Field Calculator…**
                    a. Set the **Par_ID** equal to **[FID] + 1**
                    b. Click **OK**
                ▪ Right-click on the **Par_ID_T** field and select **Field Calculator…**
                    a. Set the Par_ID_T field equal to **Par_ID**
                    b. Click **OK**

---

6. Load the Network Analyst Toolbar
   a. On the Main Menu, click on **View**
   b. Click on **Toolbars**
   c. In the side menu click **Network Analyst** to add the Network Analyst toolbar to ArcMap

7. Load the Network Analyst Window

   a. Click on the **Show/Hide Network Analyst Window** button to add the Network Analyst Window to the map

8. Create a **New Closest Facility**
   a. Click on the **Network Analyst** button on the Network Analyst toolbar
   b. Select **New Closest Facility** from the dropdown menu
   c. The Table of Contents and the Network Analyst Window will now contain a **Closest Facility** layer, which contains Facilities, Incidents, Barriers, and Routes categories

9. Load the Transit Stop locations in the Facilities category in Network Analyst
   a. In the Network Analyst window, right-click **Facilities** and click **Load Locations**. This will open the Load Locations Window
   b. In the Load Locations Window
      i. Select the point layer containing the **Transit Stop locations** from the dropdown box next to **Load From**. If the required point layer is not in the dropdown box, click on the **Browse for Point Features to Load** icon and navigate to the point layer.
      ii. Choose the **Tran_Stop_ID** as the Sort field
      iii. Choose the **Tran_Stop_ID_T** for the **Name** property under **Location Analysis** properties
      iv. Accept the defaults for the other **Locational Analysis** properties.
      v. Choose a suitable **Search Tolerance** for the **Location Position**. If features cannot be located on the network dataset, this tolerance may need to be increased.
      vi. Click **Ok**
   c. To verify all Transit Stop locations were located on the network dataset, do one of the following. If any feature has a status of **Not Located**, the search tolerance in 9.b.v may need to be increased or the feature reviewed for correct location.
      i. For small datasets, scroll through the resulting lists of Facilities in the **Network Analyst Window**, looking for features with RED questions marks.
      ii. For large datasets, it may be easier to right-click on the **Facilities** in the **Network Analyst Window**, **Open Attribute Table** and sort decending on the **STATUS** column. The **Not Located** features should sort to the top.

10. Load the Participant locations into **Incidents** category in Network Analyst
   a. In the Network Analyst window, right-click **Incidents** and click **Load Locations**. This will open the Load Locations Window.
   b. In the Load Locations Window
      i. Select the point layer containing the **Participant locations** from the dropdown box next to **Load From**. If the required point layer is not in the dropdown box, click on the **Browse for Point Features to Load** icon and navigate to the point layer.
      ii. Choose the **Par_ID** as the Sort field
      iii. Choose the **Par_ID_T** for the **Name** property under **Locational Analysis** properties
      iv. Accept the defaults for the other **Locational Analysis** properties.
      v. Choose a suitable Search Tolerance for the Location Position. If features cannot be located on the network dataset, this tolerance may need to be increased.
      vi. Click **OK**

b. To verify all Participant locations were located on the network dataset, do one of the following. If any feature has a status of **Not Located**, the search tolerance in 10.b.v may need to be increased or the feature reviewed for correct location.

    i. For small datasets, scroll through the resulting lists of Facilities in the **Network Analyst Window**, looking for features with RED questions marks.

    ii. For large datasets, it may be easier to right-click on the **Facilities** in the **Network Analyst Window**, **Open Attribute Table** and sort decending on the **STATUS** column. The **Not Located** features should sort to the top.

11. Set the properties for the analysis

    a. In the Table of Contents, right-click **Closest Facility** and click **Properties**. This will open the Closest Facility properties dialog box.

    b. Click on the **Analysis Settings Tab**

        i. Choose **Length (meters)** for **Impedence**

        ii. Choose **None** for **Default Cutoff Values**

        iii. Set **Facilities to Find** to **1**

        iv. Set the **Travel From** option to **Incident to Facility**

        v. Set **Allow U-Turns** to **Everywhere**

        vi. Set **Output Shape Type** to **True Shape**

        vii. Check the box to **Ignore Invalid Locations**

    c. Click **OK**

12. To calculate the closest Transit Stop to a Participant location, click the **Solve Button** 🔲 on the Network Analyst Toolbar.

    a. When Network Analyst has solved the calculations, the **Routes** layer will be populated with line segments joining each participant location to its closest transit stop. The Routes attribute table contains the unique IDs for the participant and the transit stop along with the network length separating them and the route name (a concatenation of the participant ID and the transit stop ID).

13. Join the Transit Stops and the Participants to the results of Find Closest Facility analysis

    a. Right click the Routes layer

    b. Select **Joins and Relates** and then click **Join**

        i. Choose Facility_ID as the field to base the join on

        ii. Choose the **Transit Stop** layer as the **Table to Join** to this layer

        iii. Choose the **Tran_ID** as the field in the table to base the join on

        iv. Click **OK**

    c. Right-click the **Routes** layer

    d. Select **Joins and Relates** and then click **Join**

        i. Choose **Incident_ID** as the field to base the join on

        ii. Choose the **Participant** layer as the **Table to Join** to this layer

        iii. Choose the **Par_ID** as the field in the table to base the join on

        iv. Click **OK**

    e. Right-click the **Routes** layer

    f. Select **Data** and then click **Export Data**

        i. Choose this **layer's source data for the coordinate system**

        ii. Specify the **Output Location**

        iii. A new shapefile will be created in the output location. This shapefile will contain the routes from each participant location to the nearest transit stop. Each route will contain the attributes of the participant and the nearest transit stop.

Calculate straight-line distance to nearest transit stop

To follow this protocol, you will need:

    1. **Participant Locations Point Feature Layer** for incidents (origins)

2. **Transit Stop Point Feature Layer** for facilities (destinations)

Calculate the straight-line distance to the nearest transit stop
1. Start **ArcMap**.
2. Click the **Add Data** button on the **Standard Toolbar**.
   a. Navigate to the Point Layers to be used in the analysis, in this case the **Participant locations** and the **Transit Stops**.
   b. Select the Point Layers and click **Add** to add them to ArcMap.
3. Open **ArcToolbox** by clicking on the icon.
4. Navigate to the **Near** tool (Analysis Tools -> Proximity) and double-click to open it.
5. Click the down-arrow in the Input Features box and select the **Participant Locations** feature layer.
6. Click the down-arrow in the Near Features box and select the **Transit Stops** feature layer.
7. Leave the **Search Radius (optional)** box empty. This will guarantee that every Participant Location will have its nearest transit stop found.
8. Leave the **Location (optional)** and **Angle (optional)** check boxes unchecked.
9. Click **OK** to execute the tool.
10. The tool will find the nearest transit stop to each Participant Location and measure the straight-line distance between those points. The tool will add two attributes to the participant location attribute table, NEAR_DIST and NEAR_FID. NEAR_DIST will contain the distance, in the units of the Layers (e.g., meters, feet). NEAR_FID will contain the feature ID of the closest transit stop. Use the NEAR_FID attribute to join the Transit Stop Layer to the Participant Location layer to determine the characteristics of the nearest transit stop.

Join the Participant Location Layer to the Transit Stop Layer using the NEAR_FID attribute.
1. **Right-click** on the Participant Location layer and select **Joins and Relates -> Join…**
2. Choose **NEAR_FID** as the field to base the join on.
3. Choose the **Transit Stop** layer as the Table to Join to this layer.
4. Choose the **FID** attribute as the field in the table to base the join on.
5. Click **OK**. The Transit Stop attributes will be joined to the Participant locations.
6. **Right-click** on the Participant Locations Layer and select **Data -> Export Data**.
   a. Choose this layer's source data for the coordinate system.
   b. Specify the **Output Location**.
   c. Click **OK**.

A new shapefile will be created in the output location. This shapefile will contain the participant locations and the attributes of the nearest transit stop to each participant location.

4.11 Transit Stop Density
Note: The protocol has been updated for ArcGIS 9.2

1. Basic Concept
Transit stops may also be a proxy for a walkable neighborhood. Transit stop density may correlate with overall activity.

2. Basic Formula
Transit Stop Density = Number of transit stops per unit area.

3. Detailed Formula
As above.

4. Comments and Explanations
None.

5. GIS Approach
The GIS approach is to perform a spatial join between the measurement geographies and the transit stops. This will count the number of stops in each measurement geography. Then divide this number by the net land area of each measurement geography.

6. GIS Steps
The following procedures were completed using ArcMap (ArcInfo) 9.2.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Transit Stops Layer** (point)

Refer to protocol 2.1 for the procedures to calculate net land area (land area with water features removed).

Add the **Transit Stops** and **Study-Site** layers to an **ArcMap Document**.
1. Right-click on the **Study-Site Layer** and choose **Joins and Relates→Join**.
2. Under the dropdown box labeled **What do you want to join to this layer?** choose **Join data from another layer based on spatial location**.
3. In **Step 1** of the dialog box, choose the **Transit Stops Layer**.
4. In **Step 2**, click the button to choose **Each polygon will be given a summary of the numeric attributes**…..
5. In **Step 3**, choose a name and location for the output layer.
6. Click **OK**.
7. The output layer now contains a **Count** field with the number of transit stops found in each feature of the measurement geography.

Calculate transit stops per net land area:
1. Right-click on the **output layer** of the **spatial join** and select **Open Attribute Table**.
2. Click **Options** and choose **Add Field**.
3. Name the field **StopDens** and change the type to **Double**.
4. Click **OK** to add the field.
5. In the **Table of Contents**, right-click on the **output layer** from the **spatial join** performed above and choose **Joins and Relates→Join**.
6. In the dropdown box, choose to **Join attributes** from a table.
7. In **Step 1**, choose to base the join on the field containing the unique ID of the measurement geographies.
8. In **Step 2**, load the table from the result of following **protocol 2.1, Net Land Area**.
9. In **Step 3**, choose the field in the **Net Land Area Layer's** table containing the unique ID of the measurement geographies.
10. Click **OK**.
11. Open the **Attribute Table** of the output layer from the spatial join, which now has the net land area table joined into it.
12. Right-click on the **StopDens** field and choose **Field Calculator…**.
13. In the **Field Calculator** dialog box, click the **Count** field, then click the / button, and finally click the name of the field containing the net land area.
14. Click **OK**.

This field now contains the density of transit stops (number of transit stops divided by the net land area in whichever units you used).

The following variables are not proposed to be used in the study, so only a detailed formula is included.

4.12 Percentage of Blocks with Full Length Sidewalks

Detailed Formula
Percentage of Blocks with Full Length Sidewalks = Percentage of block faces that have a sidewalk along the entire length of at least one side without gaps (see related measure in Cervero and Kockelman 1996, 8).

4.13 Percentage of Blocks with Streetlights

Detailed Formula
Percentage of Blocks with Streetlights = Percentage of block faces that have at least one streetlight. This will involve creating some kind of logical formula to capture this (see related measure in Cervero and Kockelman 1996, 8, 16).

5. LAND-USE MIX

By Ann Forsyth (drafting, conceptual issues and formulae, some comments and GIS approach text) and Jason Zimmerman (GIS Steps, many comments/GIS approach text), Ed D'Sousa (refinement of steps and comments), and David Van Riper (updates to ArcGIS 9.2), with reference to additional materials by the team of Active Living Research grantees at the University of North Carolina, led by Daniel Rodriguez.

In version 2.1 compared with version 2.0, the basic outlines of several new protocols were added. These include 5.8, 5.9, 5.10, and 5.15. Significant changes were made to 5.13 and 5.14. An error was corrected in 5.6 (now similar land uses = 0 in both 5.6 and 5.7). In version 2.2, the dissimilarity variables were significantly elaborated, and the "North Carolina" variables were also expanded. In version 2.3 the "North Carolina" variables were clarified. In version 3.1, protocol 5.1 had a number of steps substantially revised and 5.14 had substantial revisions to the approach and steps for measuring distance to parks and to work. Version4.0 updated and simplified several protocols.

Background

Mixed use is a measure of different activities. As Grant (2002) outlines, however, "mixed-use" is often used to indicate a mix of housing types, e.g., detached houses and apartments. Transportation planning is, in part, based on the insight that people move to go between different activities or different destinations.

It is proposed that certain kinds of mixed land uses are particularly important, particularly ones that can provide:

- A greater variety of attractive destinations within walking distance. This raises a number of issues, however: is variety or attractiveness key? Once a critical mass of land-use variety has been reached, will more mix make a difference (Krizek 2003 a, b)? In measurement there has been a focus on commercial destinations, but other destinations are also important: schools, faith-based facilities, or parks (Audirac 1999).

- More visual variety and interest for pedestrians. Varied land uses are seen as promoting architectural and landscape variety.

- Greater street safety due to informal policing—one is less likely to be alone with an attacker. This assumes that uses have a mixture of opening hours, particularly in the evening, and generate pedestrian traffic. Some uses may be perceived, however, to undermine safety—e.g., rowdy bars.

Measures of mixed use have typically been either quite broad—looking at any mix—or have focused on retail destinations. Other destinations are potentially important, however, such as parks and civic uses. As Krizek (2003a, b) outlines, common measures have included:

- Field observations or self report—simple reports on whether mix is present.

- Density and distance measures, including of total employment or businesses; density of specific kinds of land uses, businesses and jobs; or distances to particular kinds of businesses (e.g., Cervero and Gorham 1995, Handy 1992).

- Entropy and dissimilarity indices that measure overall mix (e.g., Frank and Pivo 1994, Cervero and Kockelman 1996, 1997).

The now fairly wide availability of parcel-level land-use data allows other measurements based on particular kinds of land uses (see McNally and Kulkarni 1997, 109; Crane and Crepeau 1998).

Several different sorts of data are used in mixed-use measures. Land-use data show how land is actually used at the parcel level. This is distinct from the zoning for a parcel, which is what is allowed to be built in the future. Employment data can indicate establishments and is useful in determining the scale of economic activity. It is less useful, however, for ascertaining the presence of some destinations, such as parks.

5.1 Percentage of Total Parcel Area in Major Land Uses (seven land uses in this case)

Note: The Geoprocessing Wizard has been replaced with the Intersect Overlay Tool in the ArcMap Toolbox and the operations performed by XTools can be performed from the ArcGIS Toolbox. We are deciding whether to update this protocol fully to use the toolbox and modify the water removal calculation—a different but incomplete protocol that does this is available in the companion Environment, Food, and Youth Protocols as of version 1.1. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
A simple measure of land-use percentage in major uses, this is less a measure of mix per se than the relative amount of each land use: residential, commercial, office, industrial, institutional, park and recreation, vacant.

The next draft of these protocols will review earlier work that has used similar measures, e.g., McNally and Kulkarni (1997, 109) examined a ratio of commercial area to total area.

2. Basic Formula
Percentage of Total Parcel Area in Major Land Uses = Total area in each of seven major land-use categories (residential, commercial, office, industrial, institutional, park and recreation, vacant) divided by total land area in parcels.

3. Detailed Formula
Percentage of Total Parcel Area in Major Land Uses = Total area in parcels in each of the seven major land-use categories (residential, commercial, office, industrial, institutional, park and recreation, vacant) divided by total land area in parcels, with water area removed. Note that some land uses (agricultural, airport, non-roadbed right of way) are not included in our seven major land-use categories, but are included in the total land area figures.
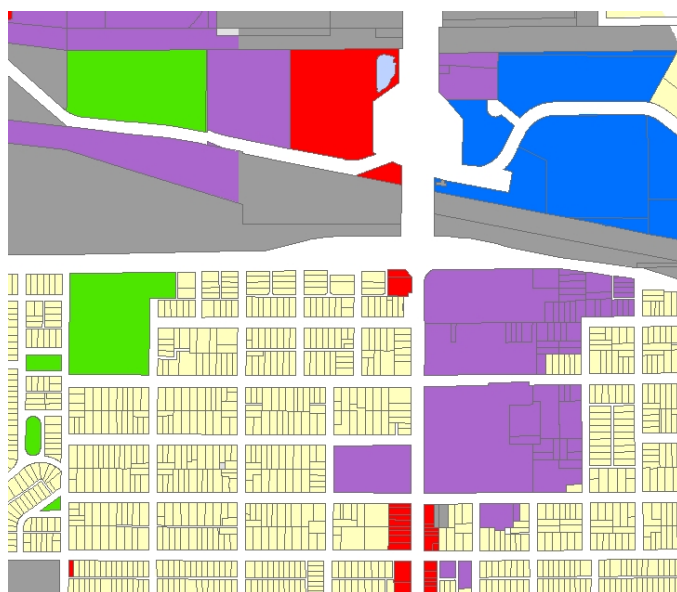
Fig. 5.1. Parcels coded with the seven major land-use categories:

Residential – yellow
Commercial – red
Institutional – purple
Office – blue
Park and recreation – green
Industrial – dark grey
Vacant - not marked

4. *Comments and Explanations*

Land uses were assigned at the parcel level and reflected the most detailed data available. These data were then generalized to our seven major land-use categories.

Complicating factors may, however, be unique to our data:

    a. Two different water layers were used for exclusions (see earlier discussion in chapter 1).

    b. We used two methods for assigning land uses, depending on whether the parcels were in our original Context Areas or fell within our extended buffer areas (see appendix 5C for details on how this was done).

    c. In general, we field-checked in the context areas, and where large buffers fell outside the context areas, we felt that the errors from using Metropolitan Council data would not be so large.

5. *GIS Approach*

The parcel layer (with areas of water removed) was intersected with the measurement geographies and the areas of the parcels recalculated to account for those that were split by the study site boundaries. The land uses associated with each parcel were generalized into seven major land-use categories. For each study site, the total area of each land-use category was summarized. The percentages of the total parcel area for each study site were then calculated.

6. *GIS Steps*

Future versions will update for buffer-related steps and refer to appendix 5C for details on creating Context Area land uses]
The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To complete this protocol you will need:
    1. **Study-Site Layer** (polygon)
    2. **Land-Use Parcels Layer** (polygon)
    3. **XTools** extension (available from the Downloads section of the ESRI Web site)

4. **Microsoft Excel** or other spreadsheet program

Determine Net Land Area (without water) for each Site: Follow the steps outlined in the protocol above.

Convert land uses into seven categories:
1. Open the **Land-Use Parcel Attribute Table**.
2. Click on the **Options** button and select **Add Field.**
3. Create a field called **7LUC** that is type **Text**.
4. Reclassify your parcels so that the land uses are generalized into the seven major land-use categories (residential, commercial, office, industrial, institutional, park and recreation, vacant). How this is done depends on the level of detail of your existing land-use categories – for example, if you have Small-scale Commercial, Strip Commercial, and Shopping Center land uses, these would all be generalized to a new land-use code of Commercial.
5. Enter the new land-use codes in the **7LUC** field.

Intersect parcels with study-site layer:
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using the **GeoProcessing Wizard**, intersect the **Land-Use Parcels Layer** with the new (net area) site polygon layer
   a. Under the **Tools** menu, select **GeoProcessing Wizard**.
   b. Select **Intersect two layers** and click on **Next**.
   c. Under **Select the input layer to intersect**, choose your **Land-Use Parcels Layer**.
   d. Under **Select a polygon clip layer**, choose your **Study-Site Layer**.
   e. Give your file a name and click **Finish**.
   f. The parcels will now be clipped to the study-site boundaries, have areas of water removed, and be assigned the IDs associated with each study site. Note that parcels that exist in two or more overlapping study sites will now be represented by two or more identical (except for ID number) rows in the **Attribute Table**).

Dissolve intersect results to unique combinations of land use and site ID and add area:
1. Under Toolbox, Data management Tools, Generalization, open the Dissolve tool
   a. Input Features are the results of the parcel, study-site intersect.
   b. Giver you output file a name.
   c. Under Dissolve_Field(s), select (check) the 7LUC and the Study-Site unique ID field.
   d. Ensure the Multipart option (at the bottom of the dialog) is checked.
2. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
   a. Under **Select layer to measure**, select your new **dissolved layer**.
   b. Under **Desired output units**, choose your units (depending on the parcel size, try **Meters** or **Feet**, for example).
   c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Area** and **Hectares** selected).
   d. Click **OK**.
3. The result will have a multipart polygon, with it's area, for each land use present at each study-site.
4. Under the **XTools** dropdown menu, select **Table Operations** and **Export to Excel**.
5. Select your **Land-Use Parcels Layer** and the following three fields: **ID**, **Hectares**, and **7LUC**.
6. Click **OK**. Depending on the number of parcels you have selected, it may take some time to export the data to **Excel**.

Calculate land-use percentages for each site:
1. In **Excel** open the Excel file containing the dissolve results.
2. Create a PivotTable using the PivotTable Wizard.
   a. The data to analyze will be the ID, Hectares and 7LUC columns
   b. The PivotTable columns are the 7LUC
   c. The PivotTable rows are the study-site IDs

d. The PivotTable data is the Hectares (area)
3. Use Excel to calculate percentages land use for each site. The right most column of the PivotTable will have the total area for each site.

5.2 Percentage of Land Area in Nighttime Uses

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
Part of the benefit of mixed use is thought to be the increased activity that it provides at various times of the day and night. This increases safety and the time period in which destinations are available to give a reason for walking.

2. Basic Formula
Percentage of Land Area in Nighttime Uses = Total area in parcels with Nighttime Uses divided by total land area in parcels.

3. Detailed Formula
Percentage of Land Area in Nighttime Uses = Total area in parcels with Nighttime Uses divided by total land area in parcels with water area removed. The specific uses classed as nighttime are the following Ramsey County uses:

a. HOTELS
b. SUPERMARKETS
c. RESTAURANTAFETERIA, AND/OR BAR
d. CONVENIENCE STORE
e. THEATERS
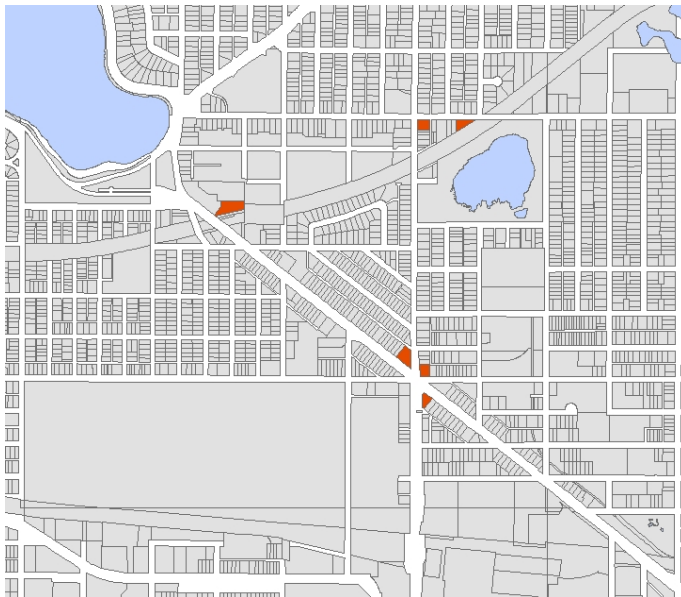
5.2, Parcels with Nighttime Uses highlighted.

4. Comments and Explanations
Ramsey County parcels included detailed land-use codes. Parcels outside of Ramsey County were field-checked to populate a land-use data field with the same level of detail.

5. GIS Approach
Using the parcel layer that was intersected with the study sites (from above), the parcels with Nighttime Uses were selected. For each study site, the total hectares of the parcels with Nighttime Uses was summarized. The percentages of the land area in Nighttime Uses for each study site were then calculated.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To complete this protocol you will need:
1. **Net Area Land-Use Parcels Layer** (polygon) created above
2. **XTools** extension (available from the Downloads section of the ESRI Web site)
3. **Microsoft Excel** or other spreadsheet program

Select parcels with nighttime uses:
1. Open the **Land-Use Parcel Attribute Table**.
2. Click on the **Options** button and select **Add Field**.
3. Create a field called **NIGHT** that is type **Text**.
4. Select any parcels that qualify as Nighttime Uses. The following are the uses we selected:
   a. HOTELS
   b. SUPERMARKETS
   c. RESTURANT, CAFETERIA, AND/OR BAR
   d. CONVENIENCE STORE
   e. THEATERS
5. Enter a new land-use code in the **NIGHT** field.

Summarize land-use hectares for each site:
1. Right-click on the **Land-Use Parcels Layer** and select **Properties**.
2. Click the **Definition Query** tab and click the **Query Builder** button.
3. In the **SELECT * FROM . . . WHERE** box select the parcels that are Nighttime Uses.
4. Click **OK**, then click **Apply**. Now, only parcels that are Nighttime Uses will be shown in the **Attribute Table**.
5. Under the **XTools** dropdown menu, select **Table Operations** and **Export to Excel**.
6. Select your **Land-Use Parcels Layer** and the **ID** and **Hectares** fields.
7. Click **OK**. Depending on the number of parcels you have selected, it may take some time to export the data to **Excel**.
8. The new **Excel** file will contain the hectares of the Nighttime Uses in each study site; use Excel's **Subtotals** command (under the **Data** dropdown menu) to summarize and group this information.
9. Re-query the **Land-Use Parcels Layer** to once again include all parcels (do not include areas of water or roads). Follow steps 1-4, but modify the criteria for the parcels to be selected.
10. Return to the **Land-use Parcel Attribute Table**.
11. Right-click on the field that contains the IDs and select **Summarize**.
12. Under **2. Choose one or more summary statistics**, expand the box labeled **Hectares** and check the box marked **Sum**.
13. Select or create an output file name and click **OK**.
14. The output .dbf file will contain the total hectares of the land use in each study site.

Calculate land-use percentages for each site:
1. In **Excel** or other spreadsheet program, combine the land-use hectares and your total parcel hectares based on the site ID numbers.
2. For each land-use category in each site, divide the land-use hectares by total parcel hectares to get the percentages of total parcel area in the Nighttime Uses.

5.3 Percentage of Land Area in Social Uses

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
Land uses that perform a social or civic function may be important destinations for walking. These are often missed in measures of mixed use that focus on retail and other commercial uses.

2. Basic Formula
Percentage of Land Area in Social Uses = Total area in parcels with Social Uses divided by total land area in parcels.

3. Detailed Formula
Percentage of Land Area in Social Uses = Total area in parcels with Social Uses divided by total land area in parcels with water area removed. Social Uses include the following Ramsey County uses:

a. DAY CARE CENTERS
b. MEDICAL CLINICS AND OFFICES
c. THEATERS
d. BOWLING ALLEYS
e. LODGE HALLS AND AMUSEMENT PARKS
f. SPORT/PUBLIC ASSEMBLY FACILITY
g. EXEMPT COMMUNITY RECREATIONAL FAC.
h. LIBRARY
i. EXEMPT PROPERTY OWNED BY BD. OF EDUC.
j. EXEMPT PROP OWNED BY PRIVATE SCHOOLS
k. CHURCHES ETC PUBLIC WORSHIP PRIVE
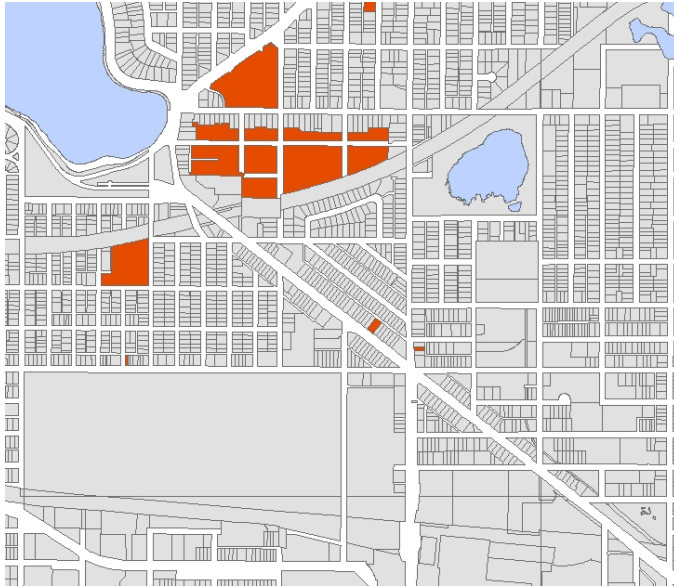
Fig. 4.3. Parcels with Social Uses highlighted.

4. Comments and Explanations

Ramsey County parcels included detailed land-use codes. Parcels outside of Ramsey County were field checked to populate a land-use data field with the same level of detail.

5. GIS Approach

Using the parcels layer that was intersected with the study sites (from above), the parcels with Social Uses were selected. For each study site, the total hectares of the parcels with Social Uses were summarized. The percentages of the land area in Social Uses for each study site were then calculated.

6. GIS Steps

The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To complete this protocol you will need:
1. **Net Area Land-Use Parcels Layer** (polygon) created above
2. **XTools** extension (available from the Downloads section of the ESRI Web site)

Select Parcels with Social Uses:
1. Open the **Land-Use Parcel Attribute Table**.
2. Click on the **Options** button and select **Add Field**.
3. Create a field called **SOCIAL** that is type **Text**.
4. Select any parcels that qualify as Social Uses. The following are the uses we selected:
   a. DAY CARE CENTERS
   b. MEDICAL CLINICS AND OFFICES
   c. THEATERS
   d. BOWLING ALLEYS
   e. LODGE HALLS AND AMUSEMENT PARKS
   f. SPORT/PUBLIC ASSESMBLY FACILITY
   g. EXEMPT COMMUNITY RECREATIONAL FAC.
   h. LIBRARY
   i. EXEMPT PROPERTY OWNED BY BD OF EDUC.
   j. EXEMPT PROP OWNED BY PRIVATE SCHOOLS
   k. CHURCHES ETC PUBLIC WORSHIP PRIVE

5. Enter a new land-use code in the **SOCIAL** field.

Summarize Land-use Hectares for each Site:
1. Right-click on the **Land-Use Parcels Layer** and select **Properties**.
2. Click the **Definition Query** tab and click the **Query Builder** button.
3. In the **SELECT * FROM . . . WHERE** box, select the parcels that are Social Uses.
4. Click **OK**, then click **Apply**. Now, only parcels that are Social Uses will be shown in the **Attribute Table**.
5. Under the **XTools** dropdown menu, select **Table Operations** and **Export to Excel**.
6. Select your **Land-Use Parcels Layer** and the **ID** and **Hectares** fields.
7. Click **OK**. Depending on the number of parcels you have selected, it may take some time to export the data to **Excel**.
8. The new **Excel** file will contain the hectares of the social uses in each study site; use Excel's **Subtotals** command (under the **Data** dropdown menu) to summarize and group this information.
9. Re-query the **Land-use Parcels Layer** to once again include all parcels (do not include areas of water or roads). Follow steps 1-4, but modify the criteria for the parcels to be selected.
10. Return to the **Land-Use Parcel Attribute Table**.
11. Right-click on the field that contains the **IDs** and select **Summarize**.
12. Under **2. Choose one or more summary statistics**, expand the box labeled **Hectares** and check the box marked **Sum**.
13. Select or create an output file name and click **OK**.
14. The output .dbf file will contain the total hectares of the land use in each study site.

Calculate Land-use Percentages for each Site:
1. In **Excel** or other spreadsheet program, combine the land-use hectares and your total parcel hectares based on the site ID numbers.
2. For each land-use category in each site, divide the land-use hectares by total parcel hectares to get the percentages of total parcel area in the Social Uses.

5.4 Percentage of Land Area in Retail Uses
Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below is updated for Version 9.2.

1. Basic Concept
Retail uses are an important destination for trips. These protocols use a number of different retail measures, including land use (this protocol) and employment data (see below).

A later version of these protocols will deal with how retail measures have been used in previous studies, e.g., Crane and Crepeau (1998, 26), which examined the share of each census tract in commercial use and Cervero and Kockelman (1996, 8), which looked at retail uses per developed hectare.

2. Basic Formula
Percentage of Land Area in Retail Uses = Total area in parcels with Retail Uses divided by total land area in parcels.

3. Detailed Formula
Percentage of Land Area in Retail Uses = Total area in parcels with Retail Uses divided by total land area in parcels with water area removed. Retail Uses include the following Ramsey County uses:

   a. SMALL (UNDER 10,000SF) DETACH RETAI
   b. SUPERMARKETS

---

c. DISCOUNT STORES & JR. DEPT STORES
d. MEDIUM DETACHED RETAIL
e. FULL LINE DEPARTMENT STORES
f. NEIGHBORHOOD SHOPPING CENTER
g. COMMUNITY SHOPPING CENTER
h. RESTURANT, CAFETERIA, AND/OR BAR
i. SMALL STRIP CENTER
j. CONVENIENCE STORE
k. FULL SERVICE BANKS
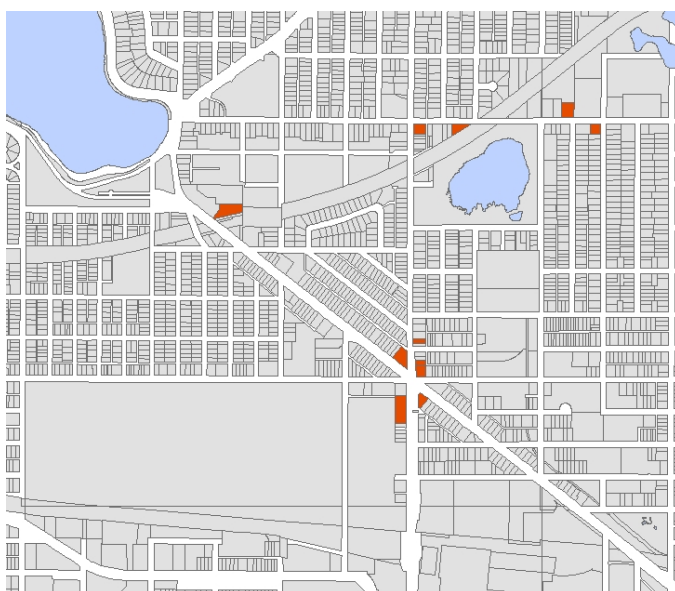
Fig. 4.4. Parcels with retail uses highlighted.

4. Comments and Explanations
Ramsey County parcels included detailed land-use codes. Parcels outside of Ramsey County were field-checked to populate a land-use data field with the same level of detail.

5. GIS Approach
Using the parcels layer that was intersected with the study sites (from above), the parcels with Retail Uses were selected. For each study site, the total hectares of the parcels with Retail Uses were summarized. The percentages of the land area in Retail Uses for each study site were then calculated.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.2**.

To complete this protocol you will need:
1. Study-Site Layer (polygon)
2. Land-Use Parcels Layer (polygon))

If you have already intersected the Land-Use Parcels Layer with the Study-Site Layer (as described in Protocol 5.7), you may skip to the step entitled "Select Parcels with Retail Uses". Otherwise, please complete the steps entitled "Use a Definition Query to exclude water area from the Land-Use Parcels Layer", "Intersect parcels with study-site layer" and "Calculate the area of each parcel within a study-site".

Use a Definition Query to exclude water area from the Land-Use Parcels Layer
1. In the Table of Contents, right-click on the Land-Use Parcels layer and select Properties.
2. Go to the Definition Query tab and click Query Builder.
3. Construct an SQL expression that applies to whichever field contains the land-use code and exclude values the represent water.
4. Click OK to close the Query Builder and then OK to close the Layer Properties window.
5. Ensure that water parcels in the Land-Use Parcels Layer are excluded.

Intersect parcels with study-site layer:
3. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
   a. Right-click on the Study-Site Layer and select Open Attribute Table.
   b. Click on the Options button and select Add Field.
   c. Name the field Study_ID.
   d. Select Short Integer as Type and accept default (0) for Precision.
   e. Click OK.
   f. Right-click on the Study_ID field and select Field Calculator…
   g. Set the Study_ID equal to [FID] +1.
   h. Click OK.
4. Open ArcToolbox by clicking on the icon and navigate to the Intersect tool (Analysis Tools -> Overlay -> Intersect). Double-click the Intersect tool to open it.
   a. For the Input Features, click on the down-arrow and select the Land-Use Parcels Layer from the list of Layers.
   b. Repeat 2.a and select the Study-Site Layer.
   c. Specify a name and location for the output layer in the Output Feature Class box.
   d. Set the JoinAttributes drop-down box to ALL.
      i. The ALL option will transfer all the attributes from the input features to the output feature class.
   e. Leave the XY Tolerance box empty.
   f. Set the Output Type to INPUT.
      i. The INPUT option will return intersections with the same geometry type as the Input Feature with the lowest geometry dimension.
   g. Click OK to execute the Intersect. The output feature class will be added to the Map Document.
   h. The parcels will now be clipped to the study-site boundaries and be assigned the ID associated with the study-site in which it is located. Note that parcels which exist in two or more overlapping study sites will be represented by two or more identical (except for Study-Site ID number) rows in the Attribute Table.

Calculate the area of each parcel within a study-site:
1. Right-click on the layer created by the Dissolve tool and select Open Attribute Table.
2. Click the Options button and select Add Field…
3. Type Hectares in the Name field and set the Type to Double.
4. Click OK.
5. Right-click on the Hectaresfield and select Calculate Geometry…
6. Set the Property: box to Area.
7. Set the Coordinate System to "Use coordinate system of the data source:".
8. Set the Units to Hectares [ha].
9. Click OK. This will calculate the area, in hectares, of each land-use type within a given study-site.

Select Parcels with Retail Uses:
1. Open the **Intersected Study-Site/Land-Use Parcel Attribute Table**.
2. Click on the **Options** button and select **Add Field**.
3. Name the field Retail, set the Type to Text, and set the Length to 10.
4. Click on the Options button and select "Select By Attributes…"
5. Construct an SQL querion that chooses parcels that qualify as Retail Uses. The following are the uses we selected:

    a.   SMALL (UNDER 10,000SF) DETACH RETAI
    b.   SUPERMARKETS
    c.   DISCOUNT STORES & JR. DEPT STORES
    d.   MEDIUM DETACHED RETAIL
    e.   FULL LINE DEPARTMENT STORES
    f.   NEIGHBORHOOD SHOPPING CENTER
    g.   COMMUNITY SHOPPING CENTER
    h.   RESTURANT, CAFETERIA, AND/OR BAR
    i.   SMALL STRIP CENTER
    j.   CONVENIENCE STORE
    k.   FULL SERVICE BANKS

6. After choosing Retail parcels, right-click on the Retail field and select Field Calculator…
7. Type "Yes" in the Field Calculator box and click OK. This will set the Retail attribute to Yes for each retail parcel.

Summarize land-use hectares for each site:
1. To find the total hectares of land parcels within a study site:
    a.   Open the Intersected Study-Site/Land-Use Parcel Attribute Table.
    b.   Right-click on the field containing the unique ID for each Study-Site and select Summarize…
    c.   Select the field containing the unique ID for each Study-Site in "1. Select a field to summarize:"
    d.   Click on the box next to Hectares in "2. Choose one or more summary statistics to be included in the output table:".
    e.   Specify the name and location of the output table.
    f.   Click OK. This will create a table containing one row for each study-site and a column holding the total hectares in land parcels in the study site.
2. To find the hectares in retail uses within a study site:
    a.   Open the Intersected Study-Site/Land-Use Parcel Attribute Table.
    b.   Click the Options button and select "Select By Attributes…"
    c.   Click on the Retail, then click =, and then type "Yes" (including the quotation marks).
    d.   Click OK.
    e.   Right-click on the field containing the unique ID for each Study-Site and select Summarize…
    f.   Select the field containing the unique ID for each Study-Site in "1. Select a field to summarize:"
    g.   Click on the box next to Hectares in "2. Choose one or more summary statistics to be included in the output table:"
    h.   Specify the name and location of the output table.
    i.   Make sure the box next to "Summarize on the selected records only" is checked.
    j.   Click OK. This will create a table containing one row foe each study-site and a column holding the total hectares in retail use in the study site.

Calculate the percentage of land in retail uses for each study site:
3. Open the table containing the total hectares of land parcels within a study site.
4. Add a field that will hold the percentage of land in retail uses.
    a.   Click the Options button and select "Add Field…"
    b.   Type Perc_Ret in the Name field and set the Type to Double.
    c.   Click OK
5. Join the table containing the total hectares in retail use in each study site to the table containing the total hectares of land parcels within a study site.
    a.   Right-click on the table containing the total hectares of land parcels within a study site and select Joins and Relates -> Join.
    b.   Choose the unique ID for each study-site as the field to base the join on.
    c.   Choose the table containing the total hectares in retail use in each study site as the Table to Join to this layer.

d. Choose the unique ID for each study-side as the field in the table to base the join on.
e. Click OK.
6. Calculate the percentage of land parcel area in retail uses.
    a. Right-click on the table containing the total hectares of land parcels within a study site and select Open Attribute Table.
    b. Right-click on the Perc_Ret field and select Field Calculator…
    c. Click on the field holding the hectares of retail use, then click on the / button, then click on the field holding total hectares, then click on the * button, and finally type 100.
    d. Click OK. This will calculate the percentage of land parcel area in retail uses.

5.5 Percentage of Land Area in Industrial and Auto-Oriented Uses

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
Industrial and auto-oriented land uses are thought to deter walking.

2. Basic Formula
Percentage of Land Area in Industrial and Auto-Oriented Uses = Total area in parcels with Industrial and Auto-Oriented Uses divided by total land area in parcels.

3. Detailed Formula
Percentage of Land Area in Industrial and Auto-Oriented Uses = Total area in parcels with Industrial and Auto-Oriented Uses divided by total land area in parcels, with water area removed. Industrial uses include the following Ramsey County uses:

a. INDUSTRIAL, VACANT LAND
b. FOOD&DRINK PROCESS PLANTS AND STORAE
c. FOUNDERIES&HEAVY MANUFACT PLANTS
d. MANUFACTURING & ASSEMBLY LIGHT
e. GRAIN ELEVATORS
f. OTHER INDUSTRIAL STRUCTURES
g. CAR WASHES
h. AUTO CAR SALES AND SERVICES
i. PARKING GARAGE STRUCTURE & LOTS
j. FLEX INDUSTRAIL CENTER
k. COMMERCIAL WAREHOUSES
l. MINI WAREHOUSE
m. COMMERCIAL TRUCK TERMINALS
n. RAILROAD REAL PROP USED IN OPERATION

4. Comments and Explanations
Ramsey County parcels included detailed land-use codes. Parcels outside of Ramsey County were field checked to populate a land-use data field with the same level of detail.

5. GIS Approach
Using the parcels layer that was intersected with the study sites (from above), the parcels with Industrial and Auto-Oriented Uses were selected. For each study site, the total hectares of the

parcels with Industrial and Auto-Oriented Uses were summarized. The percentages of the land area in Industrial and Auto-Oriented Uses for each study site were then calculated.
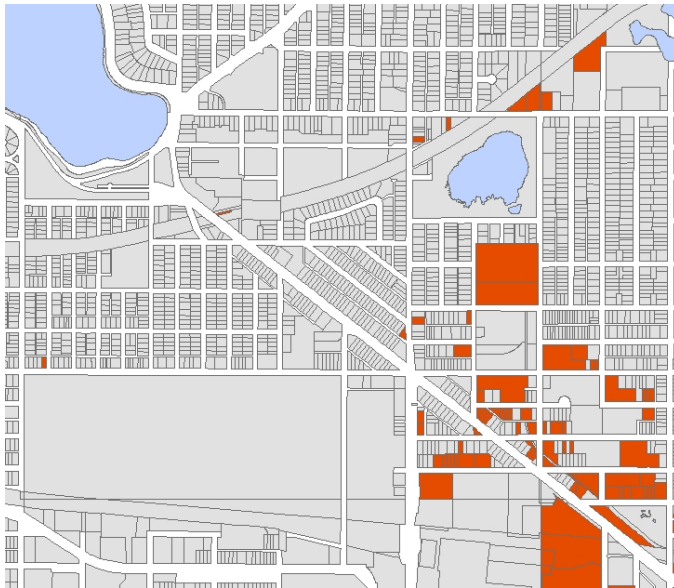
Fig. 4.5. Parcels with industrial and auto-oriented uses highlighted

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To complete this protocol you will need:
1. **Net Area Land-Use Parcels Layer** (polygon) created above
2. **XTools** extension (available from the Downloads section of the ESRI Web site)

Select Parcels with Industrial and Auto-Oriented Uses:
1. Open the **Land-Use Parcel Attribute Table**.
2. Click on the **Options** button and select **Add Field**.
3. Create a field called **AUTO** that is type Text.
4. Select any parcels that qualify as Industrial and Auto-Oriented Uses. The following are the uses we selected:
    a. INDUSTRIAL, VACANT LAND
    b. FOOD&DRINK PROCESS PLANTS AND STORAE
    c. FOUNDERIES & HEAVY MANUFACT PLANTS
    d. MANUFACTURING & ASSEMBLY LIGHT
    e. GRAIN ELEVATORS
    f. OTHER INDUSTRIAL STRUCTURES
    g. CAR WASHES
    h. AUTO CAR SALES AND SERVICES
    i. PARKING GARAGE STRUCTURE & LOTS
    j. FLEX INDUSTRAIL CENTER
    k. COMMERCIAL WAREHOUSES
    l. MINI WAREHOUSE
    m. COMMERCIAL TRUCK TERMINALS
    n. RAILROAD REAL PROP USED IN OPERATION

5. Enter a new land-use code in the **AUTO** field.

Summarize Land-Use Hectares for each Site:
1. Right-click on the **Land-Use Parcels Layer** and select **Properties**.
2. Click the **Definition Query** tab and click the **Query Builder** button.
3. In the **SELECT * FROM . . . WHERE** box select the parcels that are Industrial and Auto-Oriented Uses.
4. Click **OK**, then click **Apply**. Now, only parcels that are Industrial and Auto-Oriented Uses will be shown in the **Attribute Table**.
5. Under the **XTools** dropdown menu, select **Table Operations** and **Export to Excel**.
6. Select your **Land-Use Parcels Layer** and the **ID** and **hectares** fields.
7. Click **OK**. Depending on the number of parcels you have selected, it may take some time to export the data to **Excel**.
8. The new **Excel** file will contain the hectares of the are Industrial and Auto-Oriented Uses in each study site; use Excel's **Subtotals** command (under the **Data** dropdown menu) to summarize and group this information.
9. Re-query the **Land-Use Parcels Layer** to once again include all parcels (do not include areas of water or roads). Follow steps 1-4, but modify the criteria for the parcels to be selected.
10. Return to the **Land-Use Parcel Attribute Table**.
11. Right-click on the field that contains the **IDs** and select **Summarize**.
12. Under **2. Choose one or more summary statistics**, expand the box labeled **Hectares** and check the box marked **Sum**.
13. Select or create an output file name and click **OK**.
14. The output .dbf file will contain the total hectares of the land use in each study site.

Calculate Land-Use Percentages for each Site:
1. In **Excel** or other spreadsheet program, combine the land-use hectares and your total parcel hectares based on the site ID numbers.
2. For each land-use category in each site, divide the land-use hectares by total parcel hectares to get the percentages of total parcel area in the Industrial and Auto-Oriented Uses.

5.6 Proportion of Dissimilar Land Uses Among Grid Cells by Cell, Formula 1

Note: The Create Fishnet tool has not been tested with ArcGIS 9.1. The workflow can be modified to replace the MSExcel portion of the workflow to use native Spatial Analyst Map Algebra. This revised workflow is less error prone and does not rely on complex MSExcel formulas. The protocol is under consideration for revision or may be provided as an alternative protocol.

1. Basic Concept
This measure, adapted from work by Cervero and Kockelman (1996, 8), places a 100*100-meter grid over the study site, assigns a major land use to each grid cell, and then calculates the proportion of dissimilar land uses adjacent to or kitty-corner to the central cell. It is a way of standardizing mixed-use calculations although initially it seems to have been used because data were already available in the 100-meter cell format. Cervero and Kockelman go on to create an index basically of average dissimilarity across an area, such as a tract, similar to the method described below. Values for individual grid cells, however, are also important.

2. Basic Formula
Proportion of Dissimilar Land Uses Among Grid Cells by Cell = Working from a grid, pick one cell and calculate the dominant land use for that cell, calculate the dominant land uses in the adjacent grid cells, then calculate the percentage of cells with different land uses to the central cell.

3. Detailed Formula
Proportion of Dissimilar Land Uses Among Grid Cells by Cell = Working from a grid, such as the one in the diagram below, select one cell and calculate the dominant land use for that cell, calculate the dominant land uses in the adjacent grid cells, then calculate the percentage of cells with land uses different to the central cell. Land uses include the seven major land uses above, as well as water and an "other" category for land uses that are not among our seven major land uses (e.g., agriculture, airport). Majority land uses are assigned based on the area within parcels, as well as water bodies (road right of ways are not included in the calculations).

4. Comments and Explanations
We found two ways to do this.

> A. Using a grid to lay over the study areas, measuring land uses, and then calculating the dissimilarity index in Excel.

> B. We later found a way to automate the grid measurement process to create a dominant land use, based on Steiner (2004). This just allocates a dominant land use, however, and Excel is still needed to perform the dissimilarity.

For both, the land use for a grid cell (see center cell A in the diagram below) is compared to the land uses in each of its eight neighboring cells (B, C, D, E, F, G, H, and I), scoring 0 for an identical use and 1 for a different use. These eight values are then averaged to generate the index for that single cell. Sample grid cell pattern:

D	F	I
C	**A**	H
B	E	G

The steps for transferring the land-use data for the grid cells to Excel and performing the necessary calculations to generate the dissimilarity indices are complicated. We have attempted to break out the steps to show both a generalization and an actual example from our data.

As part of the GIS steps used to create the grid (outlined below), each grid cell is assigned a unique ID number. These numbers are assigned consecutively, from the bottom left-hand corner of the grid to the top right-hand corner.

The dissimilarity formula was applied, for example, in our study to a grid of 207 columns and 287 rows (enough to cover completely our largest buffered areas—see image). Using the steps outlined below, the grid cells were numbered consecutively from bottom to top (rows 1-287) beginning on the left and working to the right (columns 1-207). Since the steps for assigning unique ID numbers produce a beginning number of 0, the leftmost bottom cell was number 0,

the leftmost top cell was number 286, the rightmost bottom cell was number 59122, and the rightmost top cell was number 59408 (which is [207*287]-1).
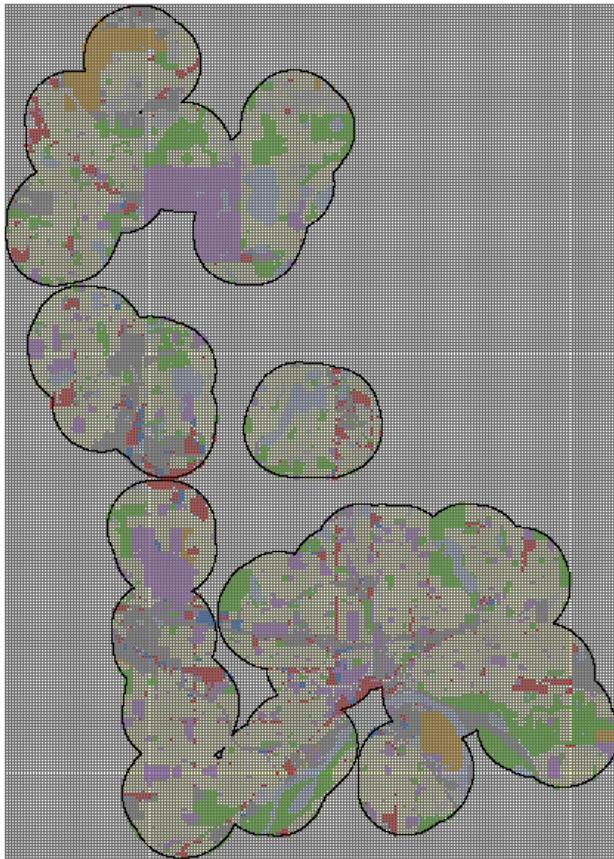
Fig. 4.6. Study area with 100*100-meter grid overlay

Looking more closely at a handful of grid cells from our study area, the pattern of unique ID numbers becomes clearer (compare to the sample grid-cell pattern above).

Fig. 4.7. 100*100-meter grid cells with land-use coding and unique grid cell ID numbers

When the data for each cell is exported to Excel (in order to carry out the dissimilarity calculations), each grid cell also assumes an Excel cell reference number. Notice how the cell reference numbers from the image below and the unique ID numbers from the image above are also found in the clipped section of the Excel spreadsheet.

Fig. 4.8. 100*100-meter grid cells with land-use coding and associated Excel cell reference numbers

	A	B	C
1	Unique ID	Land-Use Code	Dissimilarity Index
. . .			
6946	14711	I	0.5
6947	14712	I	0.375
6948	14713	I	0.5
6949	14714	R	0.875
6950	14715	WAT	0.5
. . .			
7231	14996	R	0
7232	14997	R	0.125
7233	14998	R	0.25
7234	14999	R	0.375
7235	15000	R	0.375
7236	15001	PR	1
7237	15002	WAT	0.625
7238	15003	WAT	0.625
7239	15004	I	0.25
. . .			
7520	15285	R	0
7521	15286	R	0
7522	15287	R	0.125
7523	15288	R	0.5
7524	15289	I	0.625
. . .			

The spreadsheet that calculates the dissimilarity index for each cell in the 100*100-meter grid uses the following methodology:

1. Column A contains the unique ID number for each grid cell.

2. Column B contains the land-use code associated with each grid cell (calculated to be the majority land use in each cell).

3. Column C contains the formula to calculate the dissimilarity index for each cell, utilizing the land-use codes from the adjoining 8 grid cells. The formula does the following:

    a. Checks the unique ID number for that cell

    b. Calculates the unique ID numbers for the adjoining 8 cells (since the cells are numbered consecutively, these offsets are always the same—in our example, based on the overall size of the grid, to calculate the number of the grid cell to the lower left subtract 288)

Fig 4.9. Cells

    c. Looks up or down the unique ID column to find the matching ID numbers

    d. In the row of each matching ID number, compares the land-use code to the land-use code of the center cell

    e. If the land-use codes match, a value of 0 is returned; if the land-use codes do not match, a value of 1 is returned

    f. The 8 values (0s or 1s) are summed and divided by 8 to get an index number between 0 and 1

In Microsoft Excel, here is the detailed <u>formula for calculating the Cell-by-Cell Dissimilarity Index (with values based on our data and being calculated for the cell with unique ID 15000):</u>

=SUM(IF(B7235=(VLOOKUP(A7235-288,$A$2:$B$26640,2,FALSE)),0,1),
IF(B7235=(VLOOKUP(A7235-287,$A$2:$B$26640,2,FALSE)),0,1),
IF(B7235=(VLOOKUP(A7235-286,$A$2:$B$26640,2,FALSE)),0,1),
IF(B7235=(VLOOKUP(A7235-1,$A$2:$B$26640,2,FALSE)),0,1),
IF(B7235=(VLOOKUP(A7235+1,$A$2:$B$26640,2,FALSE)),0,1),
IF(B7235=(VLOOKUP(A7235+286,$A$2:$B$26640,2,FALSE)),0,1),
IF(B7235=(VLOOKUP(A7235+287,$A$2:$B$26640,2,FALSE)),0,1),

IF(B7235=(VLOOKUP(A7235+288,$A$2:$B$26640,2,FALSE)),0,1))/8

So, generally, for each center cell A,

• **SUM(B-I)**/8 sums the values returned for B-I and divides the sum of the returned values by 8 to get the dissimilarity index for cell A

> **where B-I are the statements:**
> • **IF(W=(Y),0,1)** which check to see if the land-use values of the eight adjoining cells are the same as that of cell A
>
> > **where W is:**
> > • the land-use value of cell A
> >
> > **where Y is the statement:**
> > • **VLOOKUP(Z,range,column,FALSE)** which looks up data in a specified cell according to set of rules (the V in VLOOKUP indicates that Excel will be searching vertically in the spreadsheet in contrast to an HLOOKUP which searches horizontally)
> >
> > > **where Z is the series of statements:**
> > > • **V-offset1; V-offset2; V-offset3; V-offset4; V+offset4; V+offset3; V+offset2; V+offset1**
> > >
> > > > **where V is:**
> > > > • the unique numerical identifier of cell A, adjusted in each of the 8 entries to jump up or down the spreadsheet to access the land-use data associated with each of the 8 adjoining cells of cell A (here, based on our grid size, the offsets are 288, 287, 286, and 1)
> > >
> > > **and range is:**
> > > • the extent of the data in the worksheet (here, from cell A2 to cell B26640—the area that contains the data on both the unique ID numbers and the land-use codes)
> > >
> > > **and column is:**
> > > • the column from which the comparison value is to be selected (here, column 2, which contains the land-use codes)
> > >
> > > **and FALSE is:**
> > > • the command that forces VLOOKUP to search for an exact match.
> >
> > **and 0,1:**
> > • returns a value of 0 if W matches Y and returns 1 if it does not.

So from our example above we see that cell A with unique ID number 15000 has a land-use code R (for Residential). It is surrounded by cells with land uses I, I, R, R, PR, R, R, and R. The dissimilarity formula will return values of 1, 1, 0, 0, 1, 0, 0, and 0. Summing these eight values and dividing by eight yields a dissimilarity index for cell A of 0.375.

We did find one "glitch" in the way we used the MAJORITY function to calculate the dominant land use in each cell. Since our land-use layer was parcel based and did not include polygons for road right of ways, when the MAJORITY function was used it only counted area that was within our data layer. Therefore, a 100*100-meter grid cell that was 95 percent road right of way and 5 percent residential land use (such as a cell located in the middle of a large highway) would be coded as residential. The result is dissimilarity data that indicates less variety (or more similarity) of land uses than actually exists along these wide road right of ways.

A fix to this problem would be to "fill" the road right of ways with an additional polygon layer and take these areas into account (much like water bodies are handled). After discussing the issue we decided not to take this step and continued using the data as it was originally calculated. By taking these road right of ways into account, we would be adding dissimilarity into our results, but it would be based on wide road right of ways that are not pedestrian friendly instead of a mix of land uses, which are suggested to support pedestrian activity. Adding a road right of way layer would, therefore, actually be working against our study goals.

Also, it should be noted that the Convert to Raster function used to change the land-use parcels layer into a 10*10-meter grid cell raster does NOT use a MAJORITY function to calculate the underlying land use. In fact, it simply takes the land use beneath the grid cell center point and assigns it to the grid cell. Although this step results in some cells that misrepresent the true nature of the underlying land uses, we devised a method that broke the 100*100-meter grid cells into 10*10-meter grid cells, took the land use under the center point, and then used that as the basis for the majority land use. We estimated that this should remove most, if not all, of this variation.

The following images demonstrate the cell-by-cell dissimilarity index in one study site.

The first shows the land uses as defined by the parcel boundaries (with the 100*100-meter grid cell overlay).
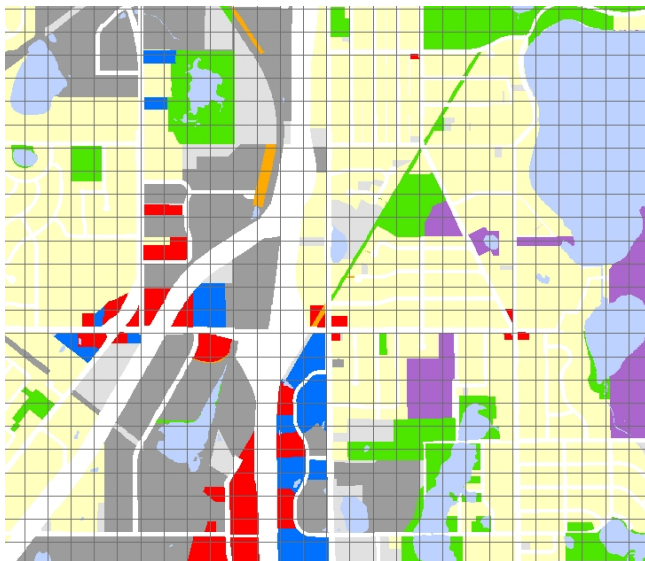
The next set of three images shows a close-up of the transition of land-use designations from parcels to 10*10-meter grid cells to 100*100-meter grid cells.
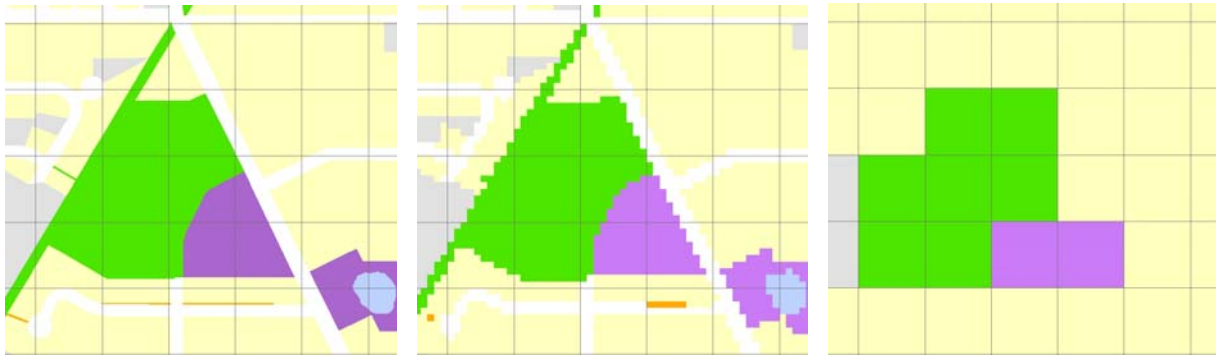
The next image shows the land uses in the broader context as defined by the 100*100-meter grid cell boundaries with the dominant land uses applied.
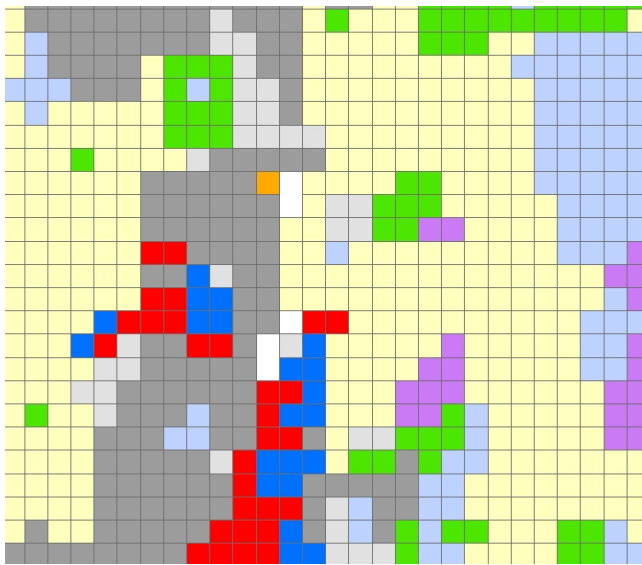
The last shows the cell-by-cell dissimilarity index. The index ranges from 0 – low dissimilarity (the lightest shade of blue, where the dominant land use of the cell is the same as all of the dominant land uses of the eight adjacent cells) to 1 – high dissimilarity (the darkest shade of blue, where the dominant land use of the cell is different from any of the dominant land uses of the eight adjacent cells).
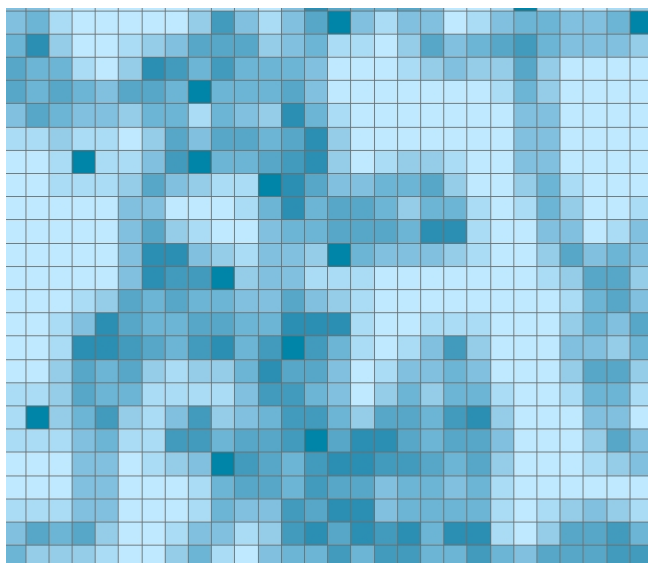
5. GIS Approach

The basic steps to calculate the level of difference between each cell and its eight neighbors are:

1. Use FISHNET tool to create a 100*100-meter raster grid of the land-use data that aligns with the left and bottom edges of the data.
2. Create a raster of the land-use data in a 10*10-meter grid (this should align with the 100*100-meter grid).
3. Create unique ID numbers for the 100*100-meter grid (copy the existing the Object_ID into a new field).
4. Use the Zonal Statistics tool to calculate the MAJORITY of land uses for each 100*100-meter grid cell of the 100 10*10-meter raster cells contained in each (this creates a raster).
5. Convert the raster back into polygons (don't simplify the polygons!).
6. Create centroids for the original 100*100-meter grid.
7. Assign the underlying (majority) land uses to the centroids using a spatial join.
8. Assign the centroid land uses back to the 100*100-meter grid.
9. Export (using **XTools**) the unique ID numbers and the land-use codes to Excel.
10. Plug in the dissimilarity formula to produce the dissimilarity index for each cell.
11. Save the unique ID number and the index columns in dbf format.
12. Join the newly indexed dbf file to the 100*100-meter grid cell file.
13. Create a new field for the dissimilarity numbers and copy the indexed column before removing the join.
14. Join this data to the Participant Point File (using a spatial join) to be able to produce dissimilarity indices for each of the participants (there may be some index numbers that will be repeated since some 100*100-meter grid cells may contain more than one participant).

6. GIS Steps

The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To complete this protocol you will need:
1. **Land-Use Parcels Layer** (polygon)
2. **Water Layer** (polygon) if water not included in **Land-Use Parcels Layer**
3. **Participants Layer** (point)

4. **Create Fishnet** tool (available from ESRI's ArcScripts Web site)
5. **XTools** extension (available from the Downloads section of the ESRI Web site)
6. **Excel** or other spreadsheet program

Create Grids:
1. Ensure that the data view is using the desired units for measurement by checking the lower-right corner of the status bar at the bottom of **ArcMap**.
2. Move the mouse to the left-most extent of the area for which grid cells will be created. Zoom in and note the **X** coordinate.
3. Move the mouse to the bottom-most extent of the area for which grid cells will be created. Zoom in and note the **Y** coordinate.
4. Measure the approximate horizontal and vertical distances that need to be covered by the grid and divide each distance by the intended size of each grid cell. This will yield the number of rows and columns needed in the grid.
5. Click the **Create Fishnet** button that was added to the toolbar during the installation of the tool.
6. Enter the X-Y coordinates recorded above (this will identify the lower left corner of the area).
7. Enter the number of rows and columns estimated above.
8. Enter the height and width of each grid cell (100*100 meters for our purposes).
9. Choose a location and name for the output shapefile.
10. Click the **Edit** button to select the coordinate system of the output shapefile.
11. Make sure it is set to product **Polygons** and click **OK**.

Create Unique ID Numbers for Grid Cells:
1. Open the **Attribute Table** for the 100*100-meter grid cells.
2. Click on **Options** and then **Add Field**.
3. Name the field and click on **OK**.
4. Right-click on the field name and select **Calculate values**.
5. In the formula box, set the new field to be equal to the **FID**.
6. Click **OK**. Each cell will now have a unique ID, numbered consecutively from bottom to top beginning with the left-most column.

Calculate Majority Land Use for each Grid Cell
1. Open the **ArcToolbox**.
2. Click on **Conversion Tools, to Raster, Feature to Raster**.
3. For **Input features**, select your **Land-Use Layer**.
4. For **Field**, select the field containing your land-use codes.
5. Name your output file.
6. For **Output cell size**, enter **10** (meters).
7. Click on **OK**. A raster with 10*10-meter grid cells containing land-use data will have been created (the first cell will be located in the lower-left corner of the Data Layer's extent).
8. In the **ArcToolbox**, click on **Spatial Analyst Tools, Zonal, Zonal Statistics**.
9. For **Input raster or feature zone data**, select your 100*100-meter grid cell file.
10. For **Zone** field, select your unique ID field.
11. For **Input value raster**, select your 10*10-meter land-use raster file.
12. Name your output file.
13. Under **Statistics type**, select **MAJORITY**.
14. Check the **Ignore NoData** option.
15. Click on **OK**. The majority land use in each 100*100-meter grid cell (calculated based on the 100 10*10-meter grid cells contained within) will be assigned in the new 100*100-meter grid cell layer (raster).
16. In the **ArcToolbox**, click on **Conversion Tools, From Raster, Raster to Polygon**.
17. For **Input raster**, select your 100*100-meter grid cell raster (just created).
18. For **Field**, select the field containing the land-use codes.
19. Name your output file.
20. Deselect the **Simplify polygons** option

21. Click **OK**. The raster file will have been converted to a polygon file (although not yet divided back into 100*100-meter grid cells.
22. Using **XTools**, click on **Feature Conversions, Shapes to Centroids**.
23. For **Input Features Layer**, select your **ORIGINAL** 100*100-meter grid-cell layer.
24. Name your output file.
25. For **Centroid location**, select **Centerpoint**.
26. Click **OK**. Your 100*100-meter grid cells will now have centroids.
27. Right-click on your new centroid layer and select **Joins and Relates, Join**.
28. For **What do you want to join to this layer? select Join data from another layer based on spatial location**.
29. For **Item 1**, select your new 100*100-meter grid-cell layer (now with land uses).
30. For **Item 2**, select the first option, **Each point will be given all the attributes of the polygon that . . . it falls inside**.
31. For **Item 3**, name your output file.
32. Click **OK**. The centroids of the 100*100-meter grid cells will now contain the land-use data.
33. Right-click on your **ORIGINAL** 100*100-meter grid-cell layer and select **Joins and Relates, Join**.
34. For **What do you want to join to this layer?** select **Join data from another layer based on spatial location**.
35. For **Item 1**, select your centroid layer (now with land uses).
36. For **Item 2**, select the second option, **Each polygon will be given all of the attributes of the point that is closets to its boundary**.
37. For **Item 3**, name your output file.
38. Click **OK**. You now have a 100*100-meter grid-cell layer that contains data on the dominant underlying land use.

Calculate Dissimilarity Indices in **Excel**:
1. Using **XTools**, click on **Table Operations, Export Data to MS Excel**.
2. For **Feature layer**, select your new 100*100-meter grid-cell layer with land uses.
3. For **Select fields to export**, select the Unique ID number and the land-use code.
4. Click **OK**.
5. In **Excel**, create a new column to contain the dissimilarity indices.
6. For each cell (Unique ID number), build a formula that calculates the dissimilarity index for that cell (see *Comments and Explanations* above for a detailed description of our Excel formula).
7. For cells that contain no data (especially along the edges of irregular study areas), you may wish to assign an easily identifiable code in the dissimilarity index column (we used 999). This will make it easy to later identify and remove (or not display) these cells.
8. When finished, save the worksheet in DBF 4 (dBase IV) format.

Join Dissimilarity Data to Cells of Interest (Participants):
1. Right-click on your new 100*100-meter grid-cell layer and select **Joins and Relates, Join**.
2. For **What do you want to join to this layer?** select **Join attributes from a table**.
3. For **Item 1**, select your Unique ID number field.
4. For **Item 2**, select your dbf file that contains the dissimilarity numbers.
5. For **Item 3**, select your Unique ID number field.
6. Click **OK**. The dissimilarity number will now be joined to your 100*100-meter grid-cell layer.
7. Right-click on your grid-cell layer and select **Open Attribute Table**.
8. Under **Options**, select **Add Field**.
9. Name your field. Under **Type**, select **Double**. For **Precision and Scale** use numbers that fit your data (we used 8 for Precision and 3 for Scale).
10. Click **OK**.
11. Right-click on the name of your new field and select **Calculate Values**.
12. Set your new field equal to the dissimilarity index field and click on **OK**.
13. Right-click on your new 100*100-meter grid-cell layer and select **Joins and Relates, Remove Join(s), Remove All Joins**.
14. Your 100*100-meter grid-cell layer will now contain dissimilarity data for each grid cell. If desired, you can pull out the data for specific cells of interest (such as those that contain the

location of individual study participants) by performing a **Join** of the grid cell data to the participant points file (note that some grid cells may be repeated since there may be more than one participant in a given grid cell).

5.7 Proportion of Dissimilar Land Uses Among Grid Cells in an Area, Formula 1

Note: The protocol already uses ArcGIS version 9.0 tools, but has not been fully tested in ArcGIS 9.1.

1. Basic Concept

The basic idea here is to use the dissimilarity calculation in the previous protocol to come up with an average dissimilarity over a larger area. Cervero and Kockleman (1996) examined a census tract scale but a number of different geographies could be used, including ones of more uniform size than tracts. For example, the index could be calculated for whole neighborhoods or for cells inside or intersecting with a small buffer distance around an individual.

2. Basic Formula

Following Cervero and Kockelman (1996, 8), the formula is: Proportion of Dissimilar Land Uses Among Grid Cells in an Area $=\{\Sigma_J^k \Sigma_1^8 (X/8)]/K\}$, where K is the number of developed grid cells in the tract and "$X_1 = 1$ if land-use category of neighboring…. Hectare grid cell differs from hectare grid cell j (0 otherwise)."

3. Detailed Formula

Following Cervero and Kockelman (1996, 8), the formula is: Proportion of Dissimilar Land Uses Among Grid Cells in an Area $=\{\Sigma_J^k \Sigma_1^8 (X/8)]/K\}$, where K is the number of developed grid cells in the study geography such as a buffer, and "$X_1 = 1$ if land-use category of neighboring …. Hectare grid-cell differs from hectare grid cell j (0 otherwise)."

4. Comments and Explanations

We considered changing the formula to a median dissimilarity, but decided to use the average of grid cell indices instead of the median because the spread is low and we are trying to pick up differences (median will hide differences).

In order to be consistent with other measures in this manual, we decided to use only the grid cells that had their centroids within each buffer. This required an intermediate step in our process that selects only those cells whose original (not intersected) centroid falls within the buffer boundary.

Very large layers of grid cells intersected with buffers may have to be split into smaller pieces before they can be dissolved (then merged back together). The intersection of grid cells with our largest buffer (1600m) resulted in 4.1 million records!

5. GIS Approach

The basic steps to calculate the average dissimilarity index across an area are:

1. Begin by creating cell by cell dissimilarity as outlined above.
2. Intersect the grid cells with buffers.
3. Dissolve the layer based on buffer ID and unique grid cell ID.

4. Select only the cells that contain centroids from the original grid cells. Export as a new data layer.
5. Join to original grid-cell layer (via unique ID) and copy over the dissimilarity indices.
6. Use the Summarize function to average the indices for each buffer.

6. GIS Steps
The following procedures were completed using ArcMap (ArcInfo) 9.0.

To complete this protocol you will need:
1. **Cell-by-Cell Dissimilarity Layer** (polygon)
2. **Grid-Cell Centroid Layer** (point)
3. **Buffer Layer** (polygon)

Generate cell-by-cell dissimilarity indices: follow steps outlined above in protocol 5.6.

Roll up grid cells into areas
1. From **ArcToolbox**, select **Analysis Tools, Overlay, Intersect**.
2. For **Input Layers**, select your **Grid-Cell Layer** and your **Buffer Layer**.
3. Name your output layer.
4. Click **OK**.
5. Once your **Intersected Layer** has been completed, from **ArcToolbox,** select **Data Management Tools, Generalization, Dissolve**.
6. For **Input Layer**, select your **Intersected Layer** (grid cells and buffer).
7. Name your output layer.
8. For **Dissolve Fields**, select your buffer ID and the unique ID of the grid cells.
9. Click **OK**.

Assign dissimilarity indices
1. Under **Selection**, choose **Select by Location**.
   a. For **I want to**, choose **select features from**.
   b. For **the following layer(s)**, choose your newly dissolved layer.
   c. For **that** choose **contain**.
   d. For **the features in this layer**, choose your **Grid Cell Centroid Layer**.
   e. Click **Apply**.
2. When this is complete, right-click on the dissolved layer and select **Data, Export Data**.
3. Choose **Export: Selected features** and check **Use the same Coordinate System as this layer's source data**.
4. Name your output layer and click **OK**.
5. Right-click on your new layer and select **Joins and Relates, Join**.
6. Select **Join** attributes from a table.
   a. For **Choose the field in this layer that the join will be base on**, choose your unique ID.
   b. For **Choose the table to join to this layer**, choose your original **Grid -Cell Layer**.
   c. For **Choose the field in the table to base the join on**, choose your unique ID.
   d. Click **OK**.
7. Once the join is complete, right-click on your layer and select **Open Attribute Table**.
8. Under **Options**, select **Add Field**.
9. Name your field. Under **Type**, select **Double**. For **Precision and Scale** use numbers that fit your data (we used 8 for Precision and 3 for Scale)
10. Click **OK**.
11. Right-click on the name of your new field and select **Calculate Values**.
12. Set your new field equal to the dissimilarity index field and click on **OK**.
13. Once the values have been copied over, right-click on your layer and select **Joins and Relates, Joins, Remove all Joins**.
14. The layer will now contain the dissimilarity indices for all selected grid cells

Calculate Average Dissimilarity Indices
1. Right-click on your layer and select **Open Attribute Table**.
2. Right-click on the buffer ID field and select **Summarize**.
3. Expand the options for your dissimilarity index field and check **Average**.
4. Click **OK**.
5. A new dbf file will be created that contains the average of the dissimilarity indices of all the associated grid cells for each buffer ID.

5.8 Dissimilarity Index, Formula 2 (new in Version 2.1)

1. Basic Concept
This is a measurement based on work by Daniel Rodriguez and colleagues at the University of North Carolina (Song and Rodriguez 2005a). **Please contact us before using this measure as we are still checking it and making sure that we are using similar scales and other characteristics to the UNC team.** This second version of a Dissimilarity Index is familiar to those working in the area of racial segregation. As they outline:

> Definition: The degree to which different land uses exist within the smaller unit of analysis (e.g., grids or neighborhood) and this distribution pattern is typical throughout the larger unit of analysis (e.g., neighborhood or the city). The value of the index ranges between 0 and 1. A value of 0 means perfect similarity; a value of 1 means perfect dissimilarity. This measure is borrowed from Duncan and Duncan (1955) who employed it to reflect the level of residential segregation and has received strong endorsement from recent findings that it is effective to capture the evenness dimension of residential segregation (Massey and Denton 1988). The measure is implemented in the Appendix for two land uses (e.g., residential and non-residential land uses) and for various land-use types simultaneously.

> Unit of analysis: Different levels of neighborhoods (e.g., grids, census boundaries, or metropolitan areas). (Song and Rodriguez 2005a, 9)

The UNC team's Web site is
http://planningandactivity.unc.edu/Activities%20and%20News.htm

2. Basic Formula
The UNC team chose to use a formula focused on two basic land-use types, residential and non-residential. As they explain:

"Measurement 1: The following formula can be used to measure the evenness between residential use and non-residential land uses.

$$D = 0.5 \sum_{i=1}^{n} |X_i - Y_i|$$

The first implementation of the index, D_N, is to examine if the distribution of land uses at sub-neighborhood level (i.e., a grid cell) is similar to the distribution pattern at the neighborhood level. We denote n as the number of grids in the neighborhood, X_i as the ratio of the residential area in grid i to the residential area in the neighborhood, and Y_i as the ratio of the non-residential area in grid i to the non-residential area in the neighborhood" (Song and Rodriguez 2005a, appendix 4).

In their format, the index is measured for grid cells (we chose 100m) and then compared with the larger study geography. In their case this is a neighborhood. **We are still determining what geography we will use for comparison.**

 "Then the comparison of the distribution of land uses within each of the grids to the neighborhood as a whole can be made. The index of dissimilarity equals to 0 if distribution of land uses within each grid is similar to the distribution at the neighborhood level, i.e., the proportions of the residential and non-residential land uses within each of the grids are identical to the proportions at the neighborhood level respectively. The index equals to 1 if there are either only residential land uses, or there are only non-residential land uses within each of the grids" (Song and Rodriguez 2005a, appendix 4-5).

2. Basic Formula

$$D = 0.5 \sum_{i=1}^{n} |X_i - Y_i|$$

In other words, Dissimilarity = the ratio of the residential area in grid cell number 1 to the ratio in the larger comparison measurement geography (the "neighborhood") minus the ratio of the non-residential area in grid cell number 1 to the non-residential area in the neighborhood + the same calculation for grid cells 2, 3, 4, etc., all divided by 2.

 3. Detailed Formula
This formula uses the simplified land uses from earlier protocols.

4. Comments and Explanations
We are still coordinating the exact operation of this measurement with the University of North Carolina team. The key issue is defining the "neighborhood" or higher-level measurement geography.

5. GIS Approach
This is taken directly from draft work by Song and Rodriguez and should be attributed to them. The citation is: Song, Yan and Daniel A. Rodríguez (2005b). *Steps for Computing the Dissimilarity, Entropy, Herfindahl-Hirschman and Accessibility (Gravity with Competition) Indices.* Draft manuscript. Please, however, check their Web site for updated versions of these steps: http://planningandactivity.unc.edu/Activities%20and%20News.htm

1. Decide to which neighborhood each taxlot belongs:
    a. Convert the taxlot polygons into centroids by using **Xtool**.
    b. Spatial join the neighborhood polygons and the taxlot centroids so that each taxlot would be assigned with a neighborhood ID and all taxlot centroids that are inside a neighborhood will have the same neighborhood ID.
2. Construct grids which would intersect with the study area: Create grids with user-defined size by using "Coordinate Grid Maker 2.290" ArcView script, which can be downloaded from ESRI's Web site.
3. Assign each taxlot by grid: Spatial join the grid (polygon) shape file and the taxlot centroids (point) shape file so that all taxlot centroids that are inside a grid cell will have the same grid cell ID.
4. Calculate total area of residential uses for each grid cell by summarizing the table of taxlot centroids shape file based on the grid cell ID. You would obtain a table with the

grid cell ID as the key. Each grid cell also has a neighborhood ID indicating the neighborhood to which the grid cell belongs.

5. Calculate total area of non-residential uses for each grid cell by summarizing the table of taxlot centroids shape file based on the grid cell ID. You would obtain a table with the grid cell ID as the key. Each grid cell also has a neighborhood ID indicating the neighborhood to which the grid cell belongs.

6. Based on the common grid cell ID, combine the table created from step 4 and the table created from step 5 into a single table. The resulted new table includes the following fields: grid cell ID (key), residential area by each grid cell, non-residential area by each grid cell, and neighborhood ID.

7. Using the attribute table of taxlot centroids, calculate total residential area by each neighborhood by summarizing residential area based on neighborhood ID. This would result a table with neighborhood ID as the key.

8. Using the attribute table of taxlot centroids, calculate total non-residential area by each neighborhood by summarizing non-residential area based on neighborhood ID. This would result a table with neighborhood ID as the key.

9. Based on the common grid cell ID, combine the table created from step 7 and the table created from step 8 into a single table. The resulted new table contains the following fields: neighborhood ID (key), residential area by each neighborhood, and non-residential area by each neighborhood.

10. Based on the common neighborhood ID, join the table created from step 9 to the table created from step 6. In addition to the existing fields in the table created from step 6, the new table will also have information on total area of neighborhood residential and non-residential uses by each grid cell.

11. In the table created from step 10, divide total residential area in grid cell by total residential area in neighborhood. A new field containing the information on the ratio is added to the table.

12. In the table created from step 11, divide total non-residential area in grid cell by total non-residential area in the neighborhood. A new field containing the information on the ratio is added to the table.

13. Calculate the difference between the two ratios obtained from step 11 and 12 and add the new field—ratio difference—to the table.

14. Calculate Dissimilarity Index:
    a. Use the table created from step 13 to summarize the grid cell ratio differences by neighborhood ID.
    b. Divide the summarized total differences by 2 for each neighborhood. This is the value of dissimilarity index for each neighborhood."(Song and Rodriguez 2005b).

6. GIS Steps
[To add]

5.9 Entropy Index (new in Version 2.1)

1. Basic Concept
This is a measurement based on work by Daniel Rodriguez and colleagues at the University of North Carolina (Song and Rodriguez 2005). *Please contact us before using this measure as we are still checking it and making sure that we are using similar scales and other characteristics to the UNC team.* As they outline:

Definition: The entropy index is a measure of variation, dispersion or diversity (Turner et al. 2001). It measures the degree to which land uses are heterogeneously distributed

within a neighborhood. A value of 0 indicates homogeneity, wherein all land uses are of one single type; a value of 1 means heterogeneity, wherein area is evenly distributed among all land-use categories.

Unit of analysis: Neighborhoods. (Song and Rodriguez 2005a, 10).

As they explain:

"The entropy index (Shannon Index) is commonly calculated through the following formula: $Entropy = \{-\sum_k [(p_i)(\ln p_i)]\}/(\ln k)$" (Song and Rodriguez 2005a, appendix 6)

There are several ways to operationalize this and the UNC team has chosen the following:

Measurement: To examine the distribution pattern of different land uses within a

neighborhood, the index spells out that p_i = Proportions of each of the six land-use types, such as single-family residential, multifamily residential, commercial, industrial, public institutional and park uses, and k = The number of land uses. In this case, k=6 (Song and Knaap 2004, Song and Rodriguez 2005b).

The UNC team's Web site is
http://planningandactivity.unc.edu/Activities%20and%20News.htm

2. Basic Formula
Entropy = negative of proportion of land use (p) of type 1 times the log of that proportion, plus proportion of land use (p) of type 2 times the log of that proportion, and so on for each of "k" land-use categories all divided by the log of "k."

3. Detailed Formula
If the proportion of a land-use type in a given neighborhood is 0, taking the log of that land use would result in an undefined term, preventing the completion of the calculation. To work around this problem, Rodriguez suggests the following fix:

For each land use i, the term is P_i*ln(P_i). To calculate, we take the following shortcut: If P_i is zero, then the whole term would be zero, even if ln(P_i) is undefined (in other words, that land use does not contribute to the entropy—the key is to view entropy as the sum of proportions of each land use, weighed by the natural log of those proportions) (from personal e-mail correspondence, Rodriguez 2005).

4. Comments and Explanations
The University of North Carolina team defined "neighborhoods" as buffers of ⅛, ¼, and ½ mile around an individual's home. To be consistent with our other measures, we used similarly sized buffers of 200m, 400m, and 800m (both straight-line and network).

Some of our 200m network buffers had no land-use area associated with them (see explanation in section 2.6). For these buffers, we recorded "No Data," rather than using values of 0, to distinguish them from buffers with calculated indices of 0 (which indicate perfect homogeneity).

5. GIS Approach

This is taken directly from draft work by Song and Rodriguez and should be cited as being by them. The citation is: Song, Yan and Daniel A. Rodríguez. 2005b. *Steps for Computing the Dissimilarity, Entropy, Herfindahl-Hirschman and Accessibility (Gravity with Competition) Indices.* Draft manuscript. Please, however, check their Web site for updated versions of these steps: http://planningandactivity.unc.edu/Activities%20and%20News.htm

1. "Decide to which neighborhood each taxlot belongs:
   a. Convert the taxlot polygons into centroids by using **Xtools**.
   b. Spatial join the neighborhood polygons and the taxlot centroids so that each taxlot would be assigned with a neighborhood ID and all taxlot centroids that are inside a neighborhood will have the same neighborhood ID.
2. Calculate the total area of each of the six land-use types in each neighborhood: Using the taxlot centroids attribute table, summarize the area of taxlots by each land-use type and by neighborhood ID. The resultant table includes the following fields: neighborhood ID, land-use type, and the total area by land-use type.
3. Calculate total area of all land uses in each neighborhood: Using the taxlot centroids attribute table, summarize the area of taxlot by neighborhood ID.
4. Based on the common neighborhood ID, join the table created from step 3 to the table created from step 2. The new table includes the following fields: Neighborhood ID, land-use type, area by land-use type, and total neighborhood area.
5. Using the table created from step 4, divide the area of each single land-use type by total

   neighborhood area. New fields – proportions of each land-use type (p_i) – are added to the table.
6. Calculate $(p_i)(\ln p_i)$ and create the new fields with the values.
7. Calculate $\{\sum_k [(p_i)(\ln p_i)]\}$: Summarize the field $(p_i)(\ln p_i)$ for all 6 land-use types by neighborhood ID and create new fields.
8. Finally calculate entropy index for each neighborhood using the formula:

   $Entropy = \{-\sum_k [(p_i)(\ln p_i)]\} / (\ln k)$." (Song and Rodriguez 2005b. Note, this has been slightly edited for consistency with usage in the rest of the manual.)

6. GIS Steps
No additional GIS Steps were taken for this measurement. All calculations were done in Excel based on previous GIS work for Percentage of Total Land Area in Major Uses (section 5.1).

5.10 Herfindahl-Hirschman Index, HHI (new in Version 2.1)

1. Basic Concept
This is a measurement based on work by Daniel Rodriguez and colleagues at the University of North Carolina (Song and Rodriguez 2005). *Please contact us before using this measure as we are still checking it and making sure that we are using similar scales and other characteristics to the UNC team.*

As they outline:

"Definition: The Herfindahl-Hirschman Index (HHI index), a commonly accepted measure of market concentration used to detect market monopoly, can be used to assess the level of land-use mixture. The HHI index is the sum of squares of the percentages of each type of land uses in the user-defined neighborhoods. If there is only one land-use

type in the neighborhood, HHI index will equal 10,000. The higher the value of HHI Index, the lower the level of land-use mixture.

Unit of analysis: Neighborhoods. (Song and Rodriguez 2005, 9)"

As Song and Rodriguez (2005b) elaborate: "Following the same notation as the entropy index, the formula of calculating HHI among *k* uses of land is:

$$HHI(k) = \sum_{i=1}^{K} (P_i * 100)^2$$

where p_i is the percentage of each type of land use in the neighborhood, and *K* is the number of land-use types."

The UNC team's Web site is
http://planningandactivity.unc.edu/Activities%20and%20News.htm

2. Basic Formula
The Herfindahl-Hirschman Index = the percentage of land-use type 1 times 100 and then squared, plus the percentage of land-use type 2 times 100 and then squared (and so on for the seven major types (residential, commercial, office, industrial, institutional, park and recreation, vacant).

3. Detailed Formula
As above.

4. Comments and Explanations
Similar to the calculations of the Entropy Index, we used buffers (neighborhoods) of 200m, 400m, and 800m around each individual.

In addition, we based our area calculations on the range of seven major land uses defined in section 5.1.

5. GIS Approach
This is taken directly from draft work by Song and Rodriguez and should be attributed to them. The citation is: Song, Yan and Daniel A. Rodríguez. 2005b. *Steps for Computing the Dissimilarity, Entropy, Herfindahl-Hirschman and Accessibility (Gravity with Competition) Indices.* Draft manuscript. Please, however, check their Web site for updated versions of these steps:
http://planningandactivity.unc.edu/Activities%20and%20News.htm

1. "Using the table created from step 4 in the entropy index, divide the area of each single land-use type by total neighborhood area. New fields—proportions of each land-use

    type (p_i)—are added to the table.
2. For each newly created field in step 1, calculate $(p_i \cdot 100)^2$ and create *i* new fields with these new values.

---

3. Calculate $\sum\limits_{i=1}^{K}(P_i*100)^2$: Summarize the field created in step 2, $(p_i*100)^2$, for all k land-use types by neighborhood ID and create new fields" (Song and Rodriguez 2005b).

Employment Per Unit Land Area (see 3.6 in density section)

5.11 Retail Employment per Unit Area
Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below hasbeen updated to version 9.2.

1. Basic Concept
Retail uses are thought to promote walking.

2. Basic Formula
Retail Employment per Unit Area = Total retail employees per unit land area (excluding water).

3. Detailed Formula
Retail Employment per Unit Area = Total retail (SIC codes 5200-5999) employees per unit land area (excluding water).

4. Comments and Explanations
We used SIC codes to correspond with earlier research. In later versions of these protocols we will provide the NAICS equivalent.

Retail data came from Dun and Bradstreet but were checked against other employment data.

In an earlier version of the protocols we proposed that the measure be made in increments of 100*100-meter grid cells. We gained access to point data, however, and could then use standard buffers.

5. GIS Approach
Divide the total retail employment in each buffer by that buffer's area (excluding water).

6. Click **OK** to close the **Query Builder** and **OK** again.
7. The **Employment Data Layer** should be updated to show only retail businesses in the data view.

Create a study-site layer with the water feature removed and the area field updated. Skip this step if you have already finished protocol 2.1:
1. Add the **Study-Site Layer** and the **Water Features Layer** to the **Map Document**.
2. Open **ArcToolbox** and navigate to the **Erase tool** (Analysis Tools -> Overlay). Double-click the **Erase** tool to open it.
3. In the **Input Features** drop-down box, select the **Study-Site layer**.
4. In the **Erase Features** drop-down box, select the **Water Features layer**.
5. Specify the location and name of the output feature class.
6. Click OK.
7. The output feature class will be generated and added to the map document.

Update the areas of the study sites after removing the water:
1. Add area fields to the output feature class created in the previous step.
   a. Right-click on the **Output Feature Class** and select **Open Attribute Table…**
   b. Click on the **Option** button and select **Add Field…**
   c. Name the field **Area** and change the **Type** to **Double**. If an **Area** field already exists, you can **Cancel** this operation and overwrite the area values.
   d. Click **OK**.
   e. Click on the **Option** button again and select **Add Field…**
   f. Name the field **Hectares** and the change the type to **Double**. If a **Hectares** field already exists, you can **Cancel** this operation and overwrite the hectares values.
2. Calculate the values for **Area** and **Hectares**
   a. For the **Area** field,
      i. Right-click the **Area** field in the **Output Feature Class** attribute table and select **Calculate Geometry…**
      ii. Set the **Property:** box to **Area**.
      iii. Set the **Coordinate System** to **"Use coordinate system of the data source:"**
      iv. Set the **Units** to the desired type. For the **Area** attribute, square meters, miles, or feet is probably appropriate.
      v. Click **OK**.
   b. For the **Hectares** field,
      i. Right-click the **Hectares** field in the **Output Feature Clas**s attribute table and select **Calculate Geometry…**
      ii. Set the **Property:** box to **Area**.
      iii. Set the **Coordinate System** to **"Use coordinate system of the data source:"**
      iv. Set the **Units** to **Hectares [ha]**.
      v. Click **OK**.

Perform a spatial join between the study-site layer and employment layer:
8. Right-click the **Study-Site Layer** and choose **Joins and Relates→Join**.
9. In the top dropdown box, make sure it is set to **Join data from another layer based on spatial location**.
10. In the next item down, select the **Employment Data Layer** to be the layer to join to the buffers.
11. Select **Each polygon will be given a summary of the numeric attributes** and check the box below next to **Sum**.
12. At the bottom, choose a name and location for the output file.
13. Click on **OK**.

The resulting file will have a field in its **Attribute Table** called **Sum_EMPL** or similar (depending on the name of the field containing the number of employees in the original employment layer). The new layer should have the same number of features as the **Study-Site Layer**.

Perform the final calculation of the number of retail employees per hectare (excluding the area of water features).

1. Open the attribute table of the output file created in the previous step.
2. Click on the **Option** button and select **Add Field…**
3. Name the field **RetPerHect** and set the type to **Double**.
4. Right-click on the **RetPerHect** field and select **Field Calculator…**
5. In the **Field Calculator**, click on the field containing the sum of retail employees, click the / button, and then click the **Hectares** field.
6. Click **OK**.

The **RetPerHect** field now contains the number of employees per hectare (excluding the area of water features).

5.12 Density of Employees in Major Retail Subcategories (Separate Measures): General Merchandise (old SIC 53xx); Food Stores (54xx); Eating and Drinking Places (58xx); Miscellaneous Retail (59xx).
Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
While retail in general is thought to be an important predictor of walking, there is considerable interest as to whether a smaller subset of retail is the most important (Krizek 2003a, 2003b). This protocol enables these retail subcategories to be measured individually and as a combination, rather as the major land uses protocol above did.

2. Basic Formula
Density of Employees in Major Retail Subcategories = Total retail employees in each of the four categories above (or all four together) per unit land area (excluding water).

3. Detailed Formula
Density of Employees in Major Retail Subcategories = Total retail employees in each of the four categories above (or all four together) per unit land area (excluding water).

4. Comments and Explanations
We used SIC codes to correspond with earlier research. In later versions of these protocols we will provide the NAICS equivalent.

These data will come from both state employment figures and Dunn and Bradstreet—however, so far we have only received the Dunn and Bradstreet data. Data are for businesses in that SIC code with fewer than 200 employees (following Krizek 2003a, 2003b).

In an earlier version of the protocols we proposed that the measure be made in increments of 100*100-meter grid cells. We gained access to point data, however, and could then use standard buffers.

5. GIS Approach

Divide the general merchandise employment in each grid cell by that cell's area (excluding water). To determine general merchandise employment density for an entire focus or context area, summarize the attributes of the grid cells by focus or context area ID and include the sum of the general merchandise employment field and the sum of the grid cell hectares field.

The process is similar for other geographies, such as buffers, although a decision needs to be made whether grid cells that are cut by a measurement geography have their value included, excluded, or apportioned.

Repeat for food stores, eating and drinking places, and miscellaneous retail.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.0**.

To complete this protocol you will need:
 Participant Buffer Layer (polygon)
 Employment Data Layer (point)
 Water Features Layer (polygon)
 XTools extension (available from the Downloads section of the ESRI Web site)

Start a new **ArcMap Document** and add the **Employment Data Layer** and **Participant Buffer Layer** to the **View** (see protocols 2.5 or 2.6 for steps to create buffers).

Query the **Employment Data Layer** to show only general merchandise businesses:
1. Add the **Employment Data Layer** to your **Map Document**.
2. Right-click on the **Employment Data Layer** in the **Table of Contents** and choose **Properties**.
3. Click the **Definition Query** tab.
4. Click on the **Query Builder** button.
5. Build a query that will identify all general merchandise businesses (in our case, "SIC_GM" = 1).
6. Click **OK** to close the **Query Builder** and **OK** again.
7. The **Employment Data Layer** should be updated to show only general merchandise businesses in the data view.

Perform a spatial join between the participant buffers and **Employment Data Layer**:
1. Right-click the **Participant Buffer Layer** and choose **Joins and Relates→Join**.
2. In the top dropdown box, make sure it is set to **Join data from another layer based on spatial location**.
3. In the next item down, select the **Employment Data Layer** to be the layer to join to the buffers.
4. Select **Each polygon will be given a summary of the numeric attributes** and check the box below next to **Sum**.
5. At the bottom, choose a name and location for the output file.
6. Click on **OK**.

The resulting file will have a field in its **Attribute Table** called **Sum_Empl** or similar (depending on the name of the field containing the number of general merchandise employees in the original **Employment Data Layer**). The new layer should have the same number of features as the buffer layer.

Refer to the protocol 2.1 for calculating Net Land Area (excluding water). Create a **Study-Site Layer** (polygon) that has had water features erased from it and area data updated. Join the **Attribute Table** of the layer created in the spatial join above to the **Attribute Table** of the **Study-Site Layer** with water removed based on the ID field.
1. Add a new field to the **Attribute Table** of type **Double** called **GMPerHect** or similar.
2. Right-click on this field and choose **Calculate Values**.
3. In the **Field Calculator**, double-click the field containing the sum of general merchandise employees, click the / button, and then double-click the **Hectares** field from the **Study-Site Table**.
4. Click **OK**.

The **GMPerHect** field now contains the number of general merchandise employees per hectare (excluding the area of water features).

5.13 Distance to Nearest Facility with Employees, and with Employees in Major Retail Subcategories (Separate Measures): General Merchandise (old SIC 53xx); Food Stores (54xx); Eating and Drinking Places (58xx); Miscellaneous Retail (59xx).

Note: The network and straightline distance calculations are similar to protocol 2.7 and 2.8, which have been modified and tested in ArcMap 9.1. This protocol has not been modified with the updated ArcGIS Network Analyst workflow. Please refer to 2.7, 2.8 or 5.14 for a revised workflow.

1. Basic Concept
More important than the density of retail may be the distance to the nearest retail establishment. The same may function for employment more generally.

2. Basic Formula
Distance to Nearest Facility = Straight-line or network distance to the centroid of the nearest facility of a certain type.

3. Detailed Formula
Same as above.

4. Comments and Explanations
The straight-line distance part of this measurement uses an ArcView extension called Nearest Features. It can be downloaded here:
http://www.jennessent.com/arcview/nearest_features.htm

The network distance to nearest part of this measurement uses an ArcView extension called Multiple Closest Facilities. It can be downloaded from the ESRI Web site at
http://arcscripts.esri.com/details.asp?dbid=13125

5. GIS Approach
The GIS approach involves determining the straight line and network distance between an origin (in this case participant locations) and the closest of the possible destinations (in this case, locations of various subcategories of retail, , or employment more generally).

6. GIS Steps

I. Straight-line distance to nearest facility

The following procedures were completed using **ArcView 3.3**.

To complete this protocol you will need:
   1. **Roads Layer** (polyline)
   2. A set of origin features—in this example, participant locations (point)
   3. A set of retail subcategory locations (point)
   4. The **Nearest Features** extension for **ArcView**

Prepare the view and data:
   1. Add the **Roads** theme and the **Participant Locations/Retail Locations** themes to a new **View**.

---

2. Go to the **View** menu and choose **Properties**..
3. Set the map units and distance units appropriate for the coordinate system of the layers
4. Go to the **File** menu and choose **Extensions**.
5. Find the **Nearest Features** extension and check the box to turn it **On**.
6. Click **OK**.

Calculate distance to nearest feature
1. Click on the **Nearest Features** button in the toolbar.
2. In the first dialog box, select the theme of **Participant Locations** and the field in its table that contains a unique ID.
3. Click **OK**.
4. In the next dialog box, select one or more **Destination Themes** (if you want to calculate the distance to nearest feature for more than one kind of feature at a time), and for each of them choose a field containing a unique ID.
5. Click **Add** for each combination of feature and unique ID.
6. Click **OK** when finished.
7. In the next dialog box, enter the number of nearest features you want to find (1 = nearest feature, 2 = nearest and second-nearest features, etc) and check the boxes for **Comparison Feature ID** and **Distance to Centroid**.
8. Click **OK**.
9. In the final dialog box, you may choose to draw lines connecting each participant location and its nearest retail location (this will increase processing time).
10. Click **OK** to begin the processing.

The output will be a table containing the unique ID of each participant location, the unique ID of the nearest retail location, and the distance between them in the chosen units of the view.

II. Network distance to nearest facility

The following procedures were completed using **ArcView 3.3**.

To complete this protocol you will need:
1. **Roads Layer** (polyline)
2. A set of origin features—in this example, **Participant Locations** (point)
3. A set of destination features—in this example, **Retail Subcategory Employment Locations** (point)
4. **Network Analyst** extension for **ArcView**
5. **Multiple Closest Facilities** extension for **ArcView**.

Calculate network distance to nearest facility
1. Start a new **ArcView** project and create a blank **View**.
   a. Open the **View** menu and choose **Add Theme**.
   b. Add the **Roads**, **Origins**, and **Destinations** layers to the **View**.
   c. Open the **View** menu and choose **Properties**.
   d. Make sure the map units (meters, feet, etc.) match the units of the data layers.
   e. Open the **File** menu and choose **Extensions**.
   f. Check the box next to the **Multiple Closest Facilities** extension and click **OK**.
2. Open the **Network Tools** menu that has been added and choose **Multiple closest facilities**.
3. In the dialog box that appears, select the **Participant Locations Layer** in the **From** dropdown box.
4. Select the unique ID of the **Participant Locations Layer** in the **From ID** field dropdown box.
5. In the **To** dropdown box, select the **Retail Locations Layer**.
6. Choose the unique ID of the **Retail Locations Layer** in the **To ID** field dropdown box.
7. In the **Network** dropdown box, select the **Roads Layer**.
8. For **Facilities to find**, enter **1**.
9. Set the **Cutoff Distance** to **0**.

---

10. The **Search Tolerance** is a distance in the current map units that **Network Analyst** will search away from the road features to find points. Make this number large enough so the process will not miss any of the participant locations.
11. Click **Travel to facility**.
12. Check the box for **Create Join Table**.
13. Click **Process**.

When the process ends, the extension will add a new polyline feature to the **View** consisting of the network segments that trace the shortest path from each participant location to its nearest retail location. The table of the new layer contains an **Origin** field with the unique ID of the participant locations layer, a **Destin** field with the unique ID of the closest retail location, and a **Length** field with the distance in the current map units between the participant location and the closest retail location.

5.14 Distance to Nearest Facility for Comparison with NQLS Survey

Note: This protocol's GIS steps were substantially revised in version 4.0.

1. Basic Concept
This set of measures has two purposes. It starts with a set of facilities that have been identified in the Neighborhood Quality of Life Survey (NQLS) developed by James Sallis and others. Using GIS we measure the physical distance between homes of subjects and these facilities. We can then assess (a) whether walking and physical activity is associated with measured distance to these facilities and (b) the correlations between the distances to the facilities measured using GIS and the survey responses (perceived distances).

2. Basic Formula
Distance to Nearest Facility = Straight line or network distance to the centroid of the nearest facility of a certain type.

3. Detailed Formula
Same as above.

4. Comments and Explanations
The NQLS survey includes a listing of "stores, facilities, and other things in your neighborhood." Respondents can say they are within the following walking distances: 1-5 minutes, 6-10, 11-20, 20-30, or 30+. The following land uses are asked about in the NQLS survey, but only those highlighted in bold initially had readily available data sources. The remaining measures were performed using data supplied with ESRI Business Analyst 9.1 InfoUSA dataset. For a complete discussion of the quality, contents and application of the InfoUSA dataset provided with Business Analyst, please see Appendix 5D.

1. convenience/small grocery store
2. supermarket
3. hardware store
4. fruit/vegetable market
5. laundry/dry cleaners
6. clothing store
7. post office
8. library
9. elementary school
10. other schools
11. book store

---

12. fast food restaurant
13. coffee place
14. bank/credit union
15. non-fast food restaurant
16. video store
17. pharmacy/drug store
18. salon/barber shop
19. your job (available from survey)
20. bus or trolley stop
21. park
22. recreation center
23. gym or fitness facility

To do this, one has to convert walking times to distances. We used 3 miles per hour as the comparison and took the outer limit of the time (i.e., 5 minutes, 10, 20, 30, and further).

5. GIS Approach
The GIS approach involves determining the straight line and network distance between an origin (in this case participant locations) and the closest of the possible destinations (in this case, locations of bus stops, participant work addresses, and parks).

In our study, the steps for calculating distances for bus stops were straightforward and the same as the steps for calculating distances in Section 5.13. We found, however, that a slight modification needed to be made in order to match each working participant (of which there were 345 – see Chapter 7 under Work Environments) with his or her specific work address location. For both straight-line and network distances, this was accomplished by calculating the distances between each participant and ALL work address locations, then selecting out only those routes for which the participant ID and the work address ID matched. This increased the processing time for both, and even required splitting the data into three pieces for the network distances and running each separately before merging them back together, but was the only way we could determine – without running each unique pair of workers and jobs separately – to accomplish this measure for a large data set.

In early versions of this protocols we created a work around do do this measure, given that ArcGIS did not have a simple method for calculating distance toa complex polygon. The tricky component of this measure is that a park is an area and so the GIS program has difficulty finding the closest point as it slowed down trying to find the exact nearest boundary point of the park. Thus, the distance to the nearest park calculation in fact starts with the park, creates a series of buffers at, say, 200m intervals, and measures whether homes or schools are within those buffers. As we described in earlier versions>=:

> We also found that calculating distances to the centroids of the polygons in our parks layer would not paint a very accurate picture of the true distance to the park, since many of our parks were large and/or had intricate shapes, placing the centroid far from the nearest edges of the park. A better way to complete this measure, we decided, was to measure the distance to the nearest park vertex. Unfortunately, the extremely large number of park vertices (due, again, to the intricate shapes of the parks) crashed the computer repeatedly and we were unable to calculate the measure in this manner.

> Finally, we decided revise our distance to parks measures and instead use a technique developed for a different Center project that best finds access points to parks. It creates a

15 meter buffer around each park polygon and then finds the points of intersection between the edge of the buffer and any road centerlines that cross this boundary. These points are then used as access points to the park and distances can be calculated to these points. (For more information, see the Metropolitan Design Center's Technical Paper Number 2, Mapping Park Buffers, at http://www.designcenter.umn.edu/projects/direct_design_asst/2004/pdfs/ParkBuffering_Tech112305.pdf.)

In addition, we decided to generalize our distances to parks to show if, for a given participant, there was a park within 200m, 400m, 800m, or 1600m (both straight-line and network distances) instead of focusing on specific distances to access points.

In version 4.1. we were able to directly create the parks measure. Please see the companion *Environment, Food, and Youth: GIS Protocols* 5.16.

6. GIS Steps
I. Straight-line distance to nearest facility

The following procedures were completed using **ArcGIS ArcInfo 9.1.**
Note: as of ArcGIS version 9.1, the **Near** tools is only available in **ArcInfo**. For **ArcView** 9.1 users, equivalent results can be obtained using "Hawth's Analysis Tools for ArcGIS" **Distance Between Points (Between Layers)** found at http://www.spatialecology.com/htools/overview.php.

To complete this protocol you will need:
    4.   A set of origin features—in this example, **Participant Locations** (point)
    5.   A set of destination features – in this example, **Bus Stop Locations** (point)
    6.   If using ArcGIS ArcView - Hawth's Analysis Tools for ArcGIS

Prepare the view and data:
    1.   Add the **Roads** theme and the **Participant Locations/Retail Locations** themes to a new **Data Frame**.
    2.   Go to the **View** menu and choose **Data Frame Properties**..
    3.   Set the map units and distance units appropriate for the coordinate system of the layers
    4.   Click **OK**.

Calculate distance to nearest feature
    6.   Open **ArcToolbox**.
    7.   Choose the **Near** tool under **Analysis Tools -> Proximity**.
    8.   For the **Input Features**, select the theme of origin features **Participant**.
    9.   For the **Near Features**, select the destination features Bus Stops.
    10. Click **OK** to begin the processing.
    11. The distance to the closest destination feature will be added to the origin features.

II. Network distance to nearest facility

The following procedures were completed using **ArcGIS ArcInfo 9.1**. This procedure will also work with ArcGIS ArcView 9.1 without modification.

To complete this protocol you will need:
    1.   **Roads Network Layer** (Network Dataset) (See Appendix 5E for general information regarding the creation of network datasets.)
    2.   A set of origin features—in this example, **Participant Locations** (point)
    3.   A set of destination features—in this example, **Bus Stop Locations** (point)

4. **Network Analyst** extension for **ArcGIS 9.1**

Prepare the view and data:

1. Start a new ArcMap project
    a. Add the **Roads**, **Origins**, and **Destinations** layers to the **View**.
    b. Open the **Network Analyst Toolbar** and choose **New Closest Facility** From the Network Analyst Menu (on the Toolbar). A **Closest Facility Group Layer** will be added to the active data frame.
    c. Open the **View** menu and choose **Data Frame**
    d. Make sure the map units (meters, feet, etc.) match the units of the data layers.

    e. On the **Network Analyst Toolbar**, click the **Network Analyst Window** button   .
1. On the **Network Analyst Window**, right-click **Incidents** and click **Load Locations**.
    a. Choose the **Participant Locations** for your **Load From** feature set.
    b. Choose a unique ID field for the **Name** property under Location Analysis Properties. In this case the **Name** will be the participants unique ID number.
    c. Choose a suitable **Search Tolerance** for the **Location Position.** If features cannot be located on the network dataset, this tolerance may need to be increased.
    d. To verify all features were located on the network dataset do one of the following. If any feature has a status of **Not Located,** the search tolerance set in 2.c may need to be increased or the feature reviewed for correct location.
           i. For small datasets, Scroll through the resulting lists of Facilities in the **Network Analyst Window,** looking for features with RED Question Marks
          ii. For large datasets, it may be easier to Right Click on the **Facilities** in the **Network Analyst Window, Open Attribute Table** and sort descending on the **STATUS** column, the "Not Located" features should sort to the top.
2. Repeat steps 2.a through 2.d for the **Facilities** features in the **Network Analyst Window.** The features to be loaded are the **Bus Stop Locations.** In this case, the unique ID field for the **Name** property will be a unique identifier for the bus stops.

Calculate network distance to nearest facility

3. Right-click on the **Closest Facility Group Layer** created in step 1b and set the following **Properties** under the **Analysis Settings** tab:
    a. Impedance: Length (Meters or whatever measurement units your data is in)
    b. Default Cutoff Value: <None>
    c. Facilities To Find: 1
    d. Travel From: Incident to Facility
    e. Allow U-Turns: Everywhere (this may be different for particular datasets)
    f. Output Shape Type: True Shape
    g. Check any Restrictions particular to your **Roads Network Layer** that you would like to enforce (such as one-way streets – for walking, one-way probably is not a true restriction)
4. Right-click on the **Closest Facility Group Layer** and click on **Solve.** This will perform the nearest facility analysis and create the save the shortest network path in the **Routes** feature class**.** The network distance is recorded in the **Total_Length** field.

Network Analyst creates it's own unique ID for incidents and facilities. To be able to relate the network distance back to the Participants ID the following steps need to be performed

5. Right-click on the **Routes** feature class in the **Closest Facility Group Layer** and choose **Joins and Relates -> Joins…**
    a. What do you want to join to this layer? **Join attributes from a table**
    b. The field in this layer the join will be base on: **IncidentID**
    c. Table to join: **Incidents**
    d. Field in table to base the join on: **ObjectID**
6. The resulting routes are stored in the **Roads Network Dataset.** To save the routes to a shapefile, right-click **Routes,** choose **Data…Export Data.**

The **Closest Facility Group Layer** can be reused for several different facilities types and the same **Participants Locations** by reloading new facilities. To delete the existing facilities, right-click **Facilities** in the **Network Analyst Window** and choose **Delete All.** Repeat steps 3, 4, 5 and 7 to perform the analysis with a different set of facilities. Step 6 only needs to be performed once, as the table join will persist with the **Closest Facility Group Layer.**

5.15 Gravity Measure with Competition (new in Version 2.1)

1. Basic Concept
This is a measurement based on work by Daniel Rodriguez and colleagues at the University of North Carolina (Song and Rodriguez 2005). *Please contact us before using this measure as we are still checking it and making sure that we are using similar scales and other characteristics to the UNC team.* As they describe:

> Definition: The simplest gravity-based measure of land-use mix can be defined as the sum of accessibility of residential land use to all other given types of non-residential land uses, discounted by the distance decay function between these two points.
> Unit of analysis: Individual land parcels or neighborhoods. (Song and Rodriguez 2005, 6).

Mathematically, Song and Rodriguez (2005b) outline the following:

$$AG_i = \sum_{j=1}^{n} \frac{A_j d_{ij}^{-\beta}}{\sum_{k=1}^{m} d_{kj}^{-\beta}}$$

where

AG_i = accessibility of residential land use (e.g., house units or neighborhoods) i to non-residential land use (e.g., non-residential land parcels or activity centers) j

A_j = attractiveness of non-residential land use j, in this case, it's measured by floor space of retailing stores

d_{ij} or d_{kj} = distance from residential use i, or housing unit k, to non-residential use j

$d_{ij}^{-\beta}$ or $d_{kj}^{-\beta}$ = impedance function based on the inverse power function

β = distance decay parameter
n = number of one type of non-residential land uses
m = number of residential land uses.

The UNC team's Web site is
http://planningandactivity.unc.edu/Activities%20and%20News.htm

2. Basic Formula
See above.

3. Detailed Formula
See above.

4. Comments and Explanations

We are still coordinating the exact operation of this measurement with the University of North Carolina team. The specific geography still needs to be defined and some of the variables need to be more clearly defined.

5. GIS Approach
This is taken directly from draft work by Song and Rodriguez and should be attributed to them. The citation is: Song, Yan and Daniel A. Rodríguez. 2005b. *Steps for Computing the Dissimilarity, Entropy, Herfindahl-Hirschman and Accessibility (Gravity with Competition) Indices.* Draft manuscript. Please, however, check their Web site for updated versions of these steps:
http://planningandactivity.unc.edu/Activities%20and%20News.htm

1. Import the attribute table of single-family taxlots to Microsoft Access database named as **db1.mdb** and name the table as **SFR**.
2. Import the attribute table of commercial stores to **db1.mdb** database and name the table as **STORE**.
3. Create a new table named as "RES" to store the information on distances and store floor space. Modify the table by applying the developed C++ program.
4. Create a new table named as **NEWSTORE** to store the information on the summation of distances among single-family housing units to commercial stores. Apply the developed C++ program to obtain the values.
5. Create a new table named as **ACCESS** to store the final values of the index. Apply the developed C++ program to obtain the final values.

Note: you need Microsoft Visual C++ along with ADO to run the developed C++ program. The code snippet is available upon request.

6. GIS Steps
[To Add]

6. STREET PATTERN

By Ann Forsyth (drafting, conceptual issues and formulae, some comment text) and Jason Zimmerman (all the terrific illustrations, most comments, GIS Steps and the GIS approach text); comment text represents discussions between Ann Forsyth, Joel Koepp, and Jason Zimmerman. Additional work, in version 4.1, is from David Van Riper.

Compared with version 2.0 of these protocols, there are few changes, except for moving the Link-Node ratio to the end and to add an explanation to 6.2. There were basically no changes between versions 2.1 and 3.1. In version 4.1 some steps were updated to ArcGIS 9.2.

Background

Streets and paths are the routes that people generally use to move about the environment, although this is not always the case as people can take shortcuts and cut-throughs in such environments as parks and parking lots.

In general, street pattern measures get at issues such as the number and directness of routes for pedestrians. Pedestrians appreciate the option of direct routes because they are exposed to the weather and they appreciate options for reasons of variety and safety.
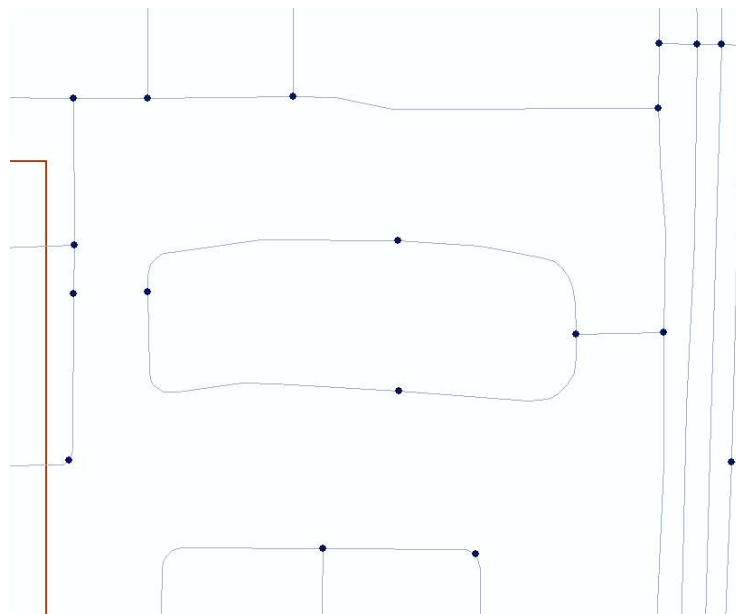
Fig. 6.1. The vast majority of our roads have nodes only at the beginning and end of the arc. This image shows both kinds—the nodes in the middle of the cul-de-sac segment get a valence of 2 and can be identified and ignored with a query when counting intersections.

In the first round of Active Living Research grants, Jennifer Dill (2003) was funded to examine the issue of measures of street pattern and we drew on her work in creating this chapter. Subsequent versions will also draw on more recent research by Steiner et al. (2004).

The next version of this chapter will contain more detail about these conceptual issues, including more of a review of how earlier studies have dealt with the issue of street pattern.

One last issue is the construction of GIS databases. In some, every street segment is made of one or more lines (in ArcGIS, called "arcs" that begin and end with "nodes"). In others, multiple segments are contained in one line. This has implications for how street intersections are counted.

6.1 Average Census Block Area
Note: This protocol's GIS steps are being considered for a revised, streamlined workflow in future versions.

1. Basic Concept
The average census block area is an easy-to-calculate approximation for street pattern. Simply put, bigger blocks have fewer streets and/or more culs-de sac and unless they also have many additional pedestrian routes they will be harder to get around. As Dill (2003, 3) explains, block size is only an approximation as the shape of the blocks also affects the street pattern. This has been often used in previous research because it uses some form of census block, which is easily available and average block size is easy to calculate in the ArcGIS programs (Hess et al. 1999, 1-2,10; Krizek 2003 (APA), 269; Cervero and Duncan 2003, 1480).

In general, however, we used median block size rather than average. This is because medians are, statistically, a preferred measure. As was explained in chapter 1, this meant that we picked up a number of areas with mostly small blocks but one large land use (e.g., a golf course or large industrial area) that gave them a medium average block size but a low median.

[In a future version of these protocols, we will discuss other related measures, such as Ewing et al. (2003, 50-51).]

2. Basic Formula
Average Census Block Area = Average area for census blocks with a centroid falling within the measurement geography.

3. Detailed Formula
Same as above.

4. Comments and Explanations
The Centroid Rule
Because the whole block size is what matters, it is not possible to clip the blocks to the measurement geography. This leaves the question of which census blocks to include, however, in the measure. We examined a number of options including:

- All census blocks that intersect the measurement geography
- All census blocks with a centroid within the measurement geography
- All census blocks completely contained within the measurement geography
-

The middle approach was selected because with this approach and a measurement geography that divided up an entire study area, for example into uniform grid cells, each block would fall within one and only one cell but also, every block would be counted as part of some cell.

Centroid Calculations—What Can Go Wrong

When using version 8.3 of Arc GIS, we calculated centroids using XTools, which creates a separate point file. XTools uses a centroid calculation method that does not require the centroid to be within the boundaries of the shape. That is, in the coordinate system, it places the centroid at the average of the shape's minimum/maximum X and Y values. The Select by Location tool in ArcMap, however, from all the evidence, requires the "center" of a shape to be inside that shape's boundaries. In Arc 9, the toolbox for calculating centroids of polygons allows you to explicitly choose whether or not the centroid must be located within the bounds of the polygon, however the choice is still important in particular circumstances, such as odd-shaped census blocks (e.g., that bend around lakes). In this case, results are different.

For consistency we used XTools, because these protocols do not require the centroid to be within the polygon. For processes requiring selecting features based on whether or not their centroids are inside a focus/context area, that selection process needs to use an actual centroid point file in the Select by Location dialog, instead of using the polygons from which those centroids were derived and letting ArcMap use its different "have their centers in" method. [This will be explained more clearly in future versions of the manual.]

Change in Protocol

The earlier TCWS version of this protocol calculated the average area of census blocks within a study-site differently from this version. TCWS first selected census block centroids that fell within a study-site. Then it used those centroids to select the actual census block polygons. Finally, it used two methodologies, one for non-overlapping study sites and one for overlapping study sites, to calculate the average area. The methodology for overlapping sites involved selecting each site individually and calculating the average area. Individual attention was required because of ArcGIS' Spatial Join function. When a spatial join is performed on two polygon layers, the software intersects the polygons and transfers the attributes based on the intersection. When one of the layers contains overlapping features, the spatial join only transfers the attributes of the first study-site that intersects with the census blocks. Thus, the second study-site will be missing the attributes of some census blocks with which it intersects. While the methodology for overlapping sites works, the individual attention required is not feasible for hundreds of study-sites.

Instead of performing the spatial join between two polygon layers, this protocol spatially joins the census block centroids (a point layer) to the study-sites. For overlapping study-sites, the spatial join will assign points that fall withing the overlap area to all study-sites that share the overlap. Thus, individual attention is not required for overlapping study-sites when joining points to polygons.

5. GIS Approach
Identify all census blocks with centroids in the measurement geography. Use a spatial join or an intersect to associate each census block with the ID of the measurement geography in which it is located. Summarize the attributes by the measurement geography ID, calculating the average of the Hectares field in the process.

After testing this protocol with different-sized buffers around schools, we have eliminated the 200m buffer from this protocol. In a number of cases, the 200m buffer failed to intersect with any census block centroids (Figure X). Instead of including buffers with an average census block

area of 0, we have eliminated this small buffer and use only the 400m, 800m, and 1600m buffers in this protocol.
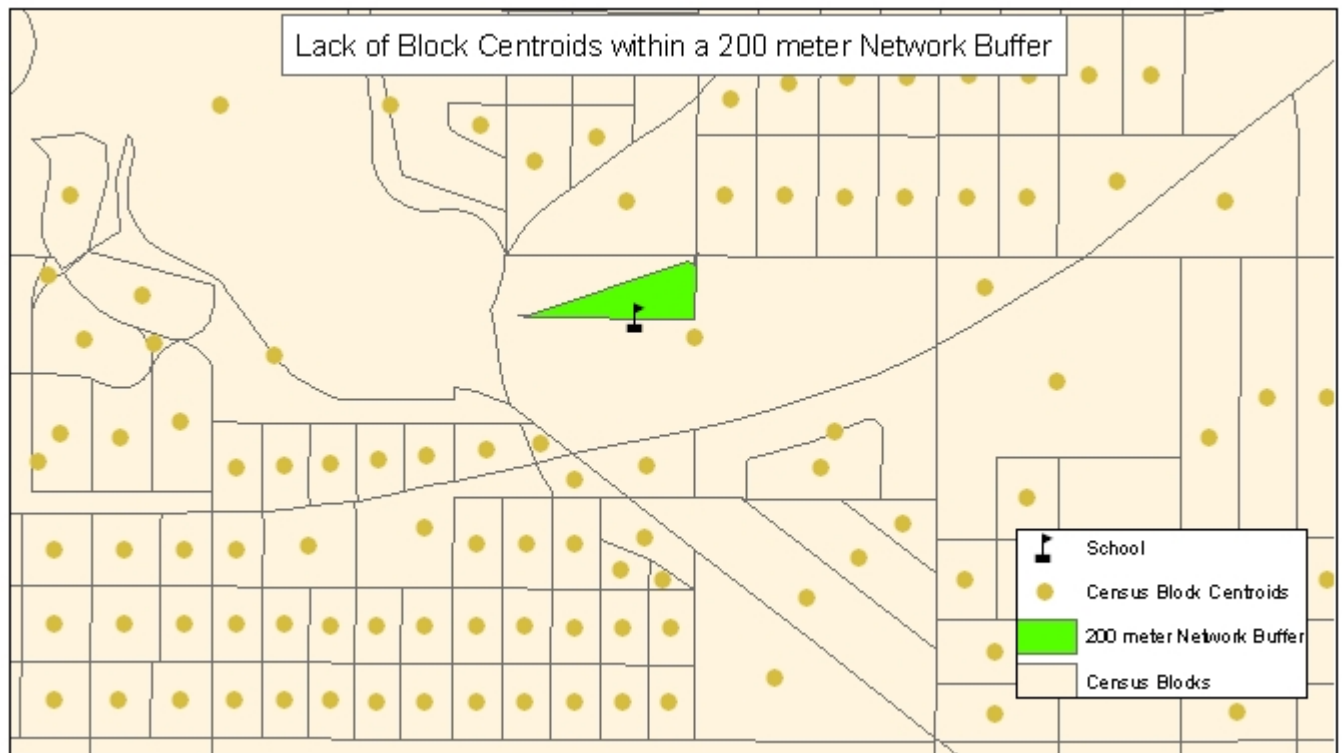
6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.2**.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon) - Buffers
2. **Census Block Layer** (polygon)

Calculate Census Blocks Areas:
1. Add the area fields to the Census Block Layer
    a. Right-click on the **Census Block Layer** and select **Open Attribute Table**.
    b. Click on the **Option** button and select **Add Field…**
    c. Name the field **Area** and change the Type to **Double**.
        i. If a field named **Area** already exists, change the name in the **Add Field…** dialog box to something else.
    d. Click **OK**.
    e. Click on the **Option** button again and select **Add Field…**
    f. Name the field **Hectares** and change the Type to **Double**.
        i. If a field name **Hectares** already exists, change the name in the **Add Field…** dialog box to something else.
    g. Click **OK**.
2. Calculate the values for Area and Hectares
    a. For the **Area** field,
        i. Right-click the **Area** field in the Census Block Layer attribute table and select **Calculate Geometry…**
        ii. Set the **Property:** box to **Area**.

ENVIRONMENT AND PHYSICAL ACTIVITY: GIS PROTOCOLS VERSION 4.1, JUNE 2007, WORK IN PROGRESS   139
University of Minnesota and Cornell

iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**"
iv. Set the **Units** to the desired type. For the Area attribute, **square meters**, **miles** or **feet** is probably appropriate.
v. Click **OK**.
b. For the **Hectares** field,
i. Right-click the **Hectares** field in the Census Block Layer attribute table and select **Calculate Geometry…**
ii. Set the **Property:** box to **Area**.
iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**"
iv. Set the Units to **Hectares [ha]**.
v. Click **OK**.

Create Census Block centroid layer:
1. Open **ArcToolbox** by clicking on the icon.
2. Navigate to the **Feature to Point** tool (**Data Management -> Features**). Double-click on the tool to open it.
3. Click on the down arrow in the **Input Features** box and select the **Census Block Layer** from the list of layers.
4. Specify the name and location of the output feature class.
5. Do NOT click the check box next to Inside. By not click the check box, ArcGIS will create the actual geometric centroid of each census block. The centroid may not fall within its census block, but it is not forced inside the block.
6. Click **OK**. The centroid layer will be created and added to the Map Document.

Calculate the Average Census Block Area for the Study-Site Layer:
1. If the Study-Site Layer does not have unique IDs or names for each site, create them now.
2. Use the **Spatial Join** to calculate the average census block area for each study-site
a. Right-click on the Study-Site Layer and select **Joins and Relates -> Join…**
b. Under "**What to you want to join to this layer?**", choose **Join data from another layer based on spatial location**.
c. Under **1. Choose the layer to join to this layer, or load spatial data from disk:**, click the down-arrow and select the Census Block centroid layer from the list.
d. Click the radio button for "**Each polygon will be given a summary of the numeric attributes of the points that fall inside it, and a count field showing how many points fall inside it**".
e. Under **How do you want the attributes to be summarized?**, click the check-box next to **Average**.
f. Specify the location and name of the output layer.
g. Click **OK**.
3. The output layer will contain one polygon for each study-site. One of the attributes of this layer will be the average hectares for all centroids falling inside a study-site.

6.2 Median Census Block Area

This protocol combines both versions available in the TCWS Protocol 6.2 in versions up to 4.0..

1. Basic Concept
As with the average census block area, the median block area is an approximation for street pattern where bigger blocks have fewer streets and/or more culs-de sac and, unless they also have many additional pedestrian routes, they will be harder to get around.

This measure is not as easy to calculate as an average, because the calculation needs to be done partly in Excel.

---

2. Basic Formula
Median Census Block Area = Median area for census blocks with a centroid falling within the measurement geography.

3. Detailed Formula
Same as above.

4. Comments and Explanations
The centroid rule and the inconsistent centroid calculations issues in the previous protocol are the same for this one.

5. GIS Approach
The GIS approach is to perform a spatial identity between the census block centroids and study sites. The attriubte table of the resulting layer is imported into Microsoft Excel and summarized using the PivotTable feature. This creates a list of all the census blocks whose centroids are in a given study site. The median area can then be calculated automatically.

6. GIS Steps

The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Census Block Layer** (polygon)
3. **Microsoft Excel**

Calculate Census Block Areas:
1. Add the area fields to the **Census Block Layer**.
   a. Right-click on the **Census Block Layer** and select **Open Attribute Table**.
   b. Click on the **Option** button and select **Add Field…**
   c. Name the field **Area** and change the Type to **Double**.
      i. If a field named **Area** already exists, change the name in the **Add Field…** dialog box to something else.
   d. Click **OK**.
   e. Click on the **Option** button again and select **Add Field…**
   f. Name the field **Hectares** and change the Type to **Double**.
      i. If a field named **Hectares** already exists, change the name in the **Add Field…** dialog box to something else.
   g. Click **OK**.
2. Calculate the values for Area and Hectares
   a. For the **Area** field,
      i. Right-click the **Area** field in the Census Block Layer attribute table and select **Calculate Geometry…**
      ii. Set the **Property:** box to **Area**.
      iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**"
      iv. Set the units to the desired type. For the **Area** attribute, **square meters**, **miles**, or **feet** is probably appropriate.
      v. Click **OK**.
   b. For the **Hectares** field,
      i. Right-click the **Hectares** field in the Census Block Layer attribute table and select **Calculate Geometry…**
      ii. Set the **Property:** box to **Area**.
      iii. Set the **Coordinate System** to "**Use coordinate system of the data source:**"

    iv. Set the **Units** to **Hectares [ha]**.
    v. Click **OK**.

Create Census Block centroid layer:
1. Open **ArcToolbox** by clicking on the icon.
2. Navigate to the **Feature to Point** tool (**Data Management -> Features**). Double-click on the tool to opent it.
3. Click on the down arrow in the **Input Features** box and select **Census Block Layer** from the list of layers.
4. Specify the name and location of the output feature class.
5. Do NOT click the check box next to Inside. By not clicking on the check box, ArcGIS will create the actual geometric centroid of each census block. The centroid may not fall within its census block, but it it not forced inside the block.
6. Click **OK**. The centroid layer will be created and added to the Map Document.

Intersect the Census Block centroids with the Study-Site Layer:
1. Add the **Study-Site Layer** to the **Map Document** if it is not there already.
2. Open **ArcToolbox** by clicking on the icon.
3. Navigate to the **Intersect tool** (**Analysis Tools -> Overlay**). Double-click on the tool to open it.
4. Click on the down arrow in the **Input Features** box and select the **Census Block centroid layer** from the list of layers.
5. Click on the down arrow in the **Identity Features** box and select the **Study-Site layer** from the list of layers.
6. Specify the name and location of the output feature class.
7. Set the **JoinAttributes** option to **ALL**.
8. Set the **XY Tolerance** to **0** and leave the **Keep Relationships check box** unchecked.
9. Click **OK**.
10. A feature class containing the census block centroids with the appropriate study-site attributes attached will be created and added to the Map Document. If a census block centroid intersects with two or more overlapping study-sites, the output feature class will contain one record for each study-site with which the centroid intersected.

Import data into Microsoft Excel:
1. Right-click on the layer created in the previous **Identiy** step and select **Open Attribute Table**.
2. Click on the **Option** button and select **Export**.
3. Set the **Export:** box to **All Records**. Then, specify a name and location for the output DBF file.
4. Click **OK**.
5. Open **Microsoft Excel** and then open the DBF file created in the previous step.
6. Open a blank Microsoft Excel workbook and copy and paste all the data from the DBF file into the workbook.
7. From the **Data** menu, select **PivotTable and PivotChart Report…**
8. Select all the data in the workbook for the **PivotTable** and click **Next**. Then click **Finish** to create the Pivot Table.
9. From the **PivotTable Field List**, drag and drop the field containing the **unique study-site ID** onto the large window where it says **Drop Column Fields Here**.
10. From the **PivotTable Field List**, drag and drop the field containing the **unique Census block centroid ID** onto the large window where it says **Drop Row Fields Here**.
11. From the **PivotTable Field List**, drag and drop the field containing the **Census Block Hectares values** onto the large window where it says **Drop Data Items Here**.
12. Find the row beneath the last row containing any values (each column will have a different number of values depending on how many census block centroids were found in each measurement geography).
13. Calculate the median of all cells in each column. This will provide the median census block size, in hectares, for the census blocks whose centroids fall within a study-site.

6.3 Ratio of Area within X Street Distance to Area within X Distance Radius

Note: The network and straight line distance calculations are similar to protocol 2.7 and 2.8, which have been modified and tested in ArcMap 9.1. This protocol has not been modified with the updated ArcGIS Network Analyst workflow. Please refer to 2.7, 2.8

1. Basic Concept
Also called the effective walking area (Dill 2003), this is a measure of how connected the street pattern is. As can be seen in the diagram below, the area that can be reached along the street network (the central, red area) is smaller in the loop and cul-de-sac road arrangement than in the more gridded street pattern to the right.
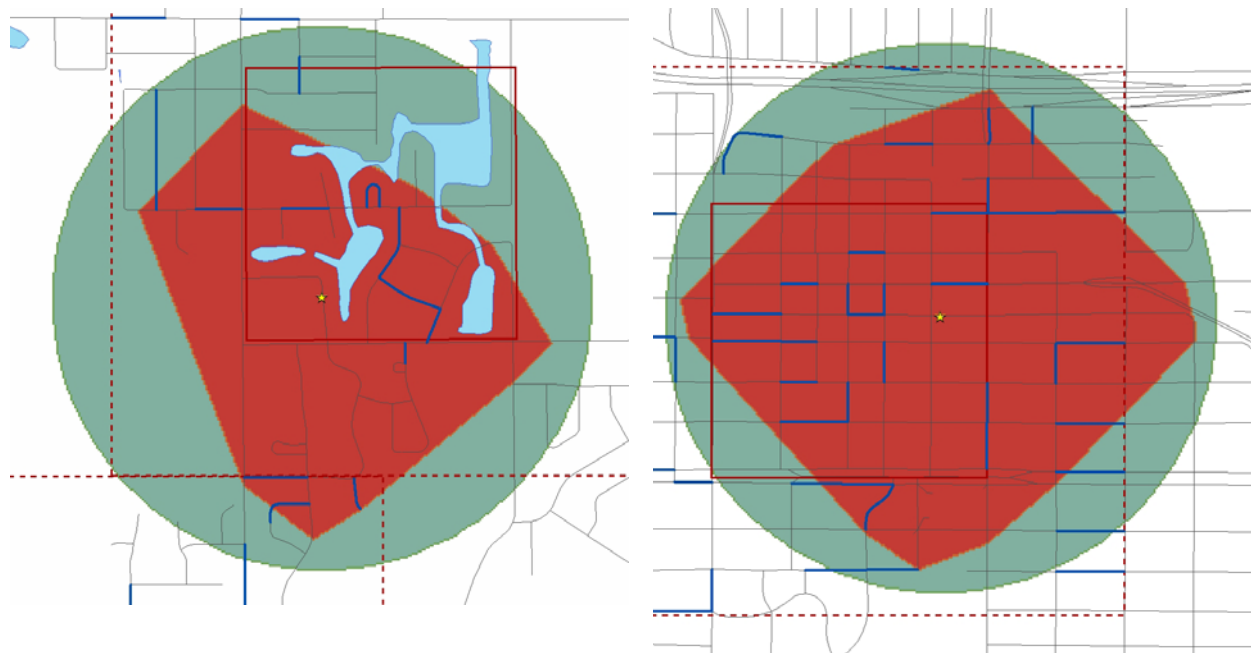
Fig. 6.1. Both diagrams show two 0.5 mile-/402-meter buffers—a straight line or crow-flies buffer and 0.5 mile-/402-meter street distance buffer. The left diagram was generated in an area with many loops and culs-de-sac and the right in a more gridded area with smaller median block size. As can be seen visually, the ratio of the area within a network buffer to area within the crow-flies buffer is much smaller in the loop and cul-de-sac area than in the gridded area, indicating that it is harder to get around.

Previous work in this area had looked at the ratio of number of parcels within the network distance to parcels within the crow-flies distance (Dill 2003, 5). We, however, decided to look at the areas.

It should be noted that the network buffer diagram shapes generate a set of points along the road network at x distance from the starting point, then join the points. This means that this is an approximation of the area people can walk to as it cuts through large parcels. It also uses the street network rather than the pedestrian network as the base and some researchers think this may be a less accurate measure of pedestrian accessibility in the suburbs where cuts-through parks and parking lots are typical (Moudon 200X). As an approximation, however, it seems to be relevant.

2. Basic Formula
Ratio of Area within X Street Distance to Area within X Distance Radius = Ratio of area within a x distance network buffer distance divided by area within a x distance crow-flies buffer.

3. Detailed Formula
Ratio of Area within X Street Distance to Area within X Distance Radius = Ratio of area within a x distance network buffer distance divided by area within a x distance crow-flies buffer with water excluded from both measures.

4. Comments and Explanations
As was explained in chapter 2, Network Analyst is currently only available in ArcView so those using ArcGIS will need to use this second program.

5. GIS Approach
Use the Buffer tool to create 400 meter buffers, then use Network Analyst to create the 400 meter network buffers. Use XTools to calculate the area of each buffer and then divide the area of the network buffer by the area of the crow flies buffer.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.0 and ArcView 3.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Participant Locations Layer** (point)
3. **Road Centerline Layer** (polyline)
4. **XTools** extension (available from the Downloads section of the ESRI Web site)

Calculate area of Straight-Line Buffers:
1. Start an **ArcMap 9.0 Document** and add the Participant Locations to the **View**.
2. In **Arc Toolbox**, open **Analysis Tools➜Proximity**, and select **Buffer**.
3. In the **Buffer** tool window, under **Input Features** select your **Participant Locations Layer** .
4. Under **Output Feature Class**, choose a name and location for the output shapefile.
5. For **Distance**, click **Linear Distance**, enter **400**, and select **Meters** from the dropdown list.
6. Click **OK**.
7. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
   a. Under **Select layer to measure**, select the new buffer layer.
   b. Under **Desired output units**, choose your units (depending on the parcel size, try **Meters** or **Feet**, for example).
   c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Hectares** selected).
   d. Click **OK**.
8. A field called **Hectares** will be added to the **Attribute Table** of the **Straight-Line Buffer Layer**.

Calculate area of Network Buffers:
1. Start an **ArcView 3.3** project and add the **Roads Layer** and **Participant Locations Layer** to a new **View**.
   a. Go to the **View** menu and select **Properties** to set the map and distance units.
   b. Go to the **File** menu and select **Extensions**.
   c. Check the box to turn on the **Network Analyst** extension.
2. Open the **Participant Locations Table**.
3. Go to the **Table** menu and choose **Start Editing**.
4. Go to the **Edit** menu and choose **Add Field**.

---

5. Name the field **REPORTCOST** and set it to type **Number**, then click **OK**.
6. Click the heading of the new **REPORTCOST** field, open the **Field** menu, and choose **Calculate**.
7. Type **400** and click **OK**.
8. Go to the **Table** menu and choose **Stop Editing**, then click **Yes** when prompted to save edits.
9. Select the **Roads** theme in the **View**.
10. Go to the **Network** menu and choose **Find Service Area**.
11. Click the **Load Sites** button and choose the **Participant Locations Layer**.
12. Make sure **Travel from site** is selected.
13. Click the button in the upper-right of the window with the grid and black arrow to begin the processing.
14. When finished, click on the **Sarea1** theme in the **View**.
15. Go to the **Theme** menu and choose **Convert to shapefile**.
16. Select a name and location for the new shapefile.
17. Switch to **ArcMap** and add this new shapefile to the **View**.
18. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length, Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
    a. Under **Select layer to measure**, select the new network buffer layer.
    b. Under **Desired output units**, choose your units (depending on the parcel size, try **Meters** or **Feet**, for example).
    c. If you do not wish to generate a parameter, such as **Perimeter** or **Length**, deselect it now (leave **Hectares** selected).
    d. Click **OK**.
19. A field called **Hectares** will be added to the **Attribute Table** of the **Network Buffer Layer**.

For each study site, divide the area of its network buffer by the area of its straight-line buffer to calculate the ratio.

6.4 Number of Access Points
Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used. The Add Points at Crossing script has not been tested in ArcGIS 9.1.

1. Basic Concept
Access points are a measure of the amount of connection an area has to its surroundings. We are not convinced it is a good measure of walkability, but we are testing it.

At its most extreme, one can imagine a gated neighborhood, inside which the neighborhood had small blocks and a highly connected street pattern, but had only one way in and out. In this case, walking to destinations outside the development would be made far more complex. Access point measures create a measurement geography, generally a grid cell, and measure how porous its edges are (Southworth 1997, McNally and Kulkarni 1997, 109).

2. Basic Formula
Number of Access Points = Number of times a road crosses the edge of the measurement geography.

3. Detailed Formula
Number of Access Points = Number of times a road crosses the edge of the measurement geography with divided at-grade roads counted once and limited access highways excluded.

Fig. 6.2. Access points are points of intersection between road centerlines and site boundaries excluding points resulting from the intersections of site boundaries with Interstates, on- and off-ramps, and other limited access highways.

In the diagram to the left, the four green points are access points created by the intersection of the site boundary and an Interstate and its on- and off-ramps, and therefore are excluded from the analysis. The red point is a valid access point.

Fig. 6.3. Site boundary intersections with divided at-grade roads, such as parkways, are also counted once (as one access point) instead of twice. The green point in the diagram to the left is half of a set of paired access points created by the intersection of the site boundary and a divided parkway. Only the red access points are included in the analysis.
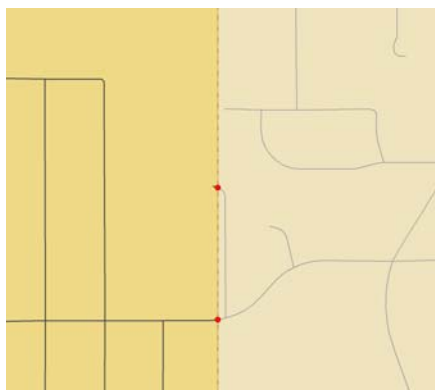
Fig. 6.4. Many access points, as measured, are somewhat marginal in their quality. That is, these points may exist but not provide significant access to the site, or may fall just inside the site boundary and therefore not be counted (see examples below). The red point in the middle of the image is technically an access point because the road centerline just crosses the site boundary, even though it provides no significant access to the site. However, it is extremely complex to develop rules to exclude such points—see next image for more examples.

Many access points, as measured, are somewhat marginal in their quality. That is, these points may exist but not provide significant access to the site, or may fall just inside the site boundary and therefore not be counted (see examples below). The red point in the middle of the image is technically an access point because the road centerline just crosses
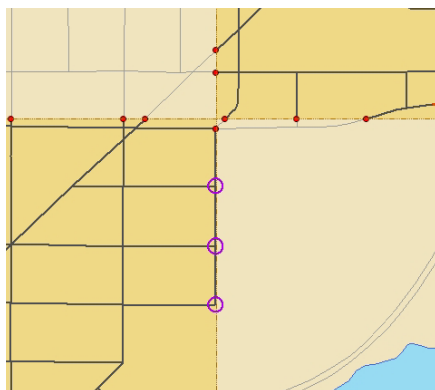
Fig. 6.5. The three purple circles identify points that should be included in the analysis, since they provide significant access to the site, are not because the road centerlines lie just inside the site boundary and do not intersect it.

We experimented with methods of buffering these points and derived various rules for exclusion or inclusion but in the end, determined that this opened the door to many subjective decisions. It seemed that the number of access points included that probably should be excluded would roughly offset the number of access points excluded that probably should be included and, therefore, we decided to set aside the issue of marginal access points in our analysis.

4. Comments and Explanations
None.

5. GIS Approach
A point file was created that contained all points were road centerlines intersected site
boundaries. From this set, Interstate points were deselected. For any divided roads (those with
two distinct centerlines, such as parkways), one half of the access points were deselected. The
total number of access points for each study site was summed.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Site Boundary Layer** (polyline)
   If you only have a polygon **Study-Site Layer**, convert the polygons to polylines using the **XTools**
   extension (available from the Downloads section of the ESRI Web site).
3. **Road Centerline Layer** (polyline)
4. **AddPointsAtCrossings ArcScript** available from the Downloads section of the ESRI Web site

Create the Access Point Layer:
1. Using **ArcMap**, install the **AddPointsAtCrossings** script following the instructions included with
   the file.
2. Add to your **Map Document** the **Road Centerline** and **Site Boundary** layers.
3. In **ArcCatalog**, create a point file (layer) that will eventually contain the output of the access point
   procedure.
4. Begin editing the point layer.
5. Highlight the two polyline layers.
6. Run the **AddPointsAtCrossings macro**.
7. Your point layer will now contain points created from the intersection of the two polyline layers.

Correct for Interstate Access Points:
1. Open the **Attribute Table** of the **Access Point Layer**.
2. Click on the **Options** button at the bottom of the window and select **Add Field**.
3. Enter the field name **Corrected**.
4. Click **OK**.
5. In your **View**, manually select the access points that result from the intersection of a site
   boundary and an Interstate or its on- or off-ramps.
6. Return to the **Attribute Table** of your **Access Point Layer**.
7. Click on the **Show: Selected** button at the bottom of the screen.
8. Right-click on the field name you just created (**Corrected**).
9. Select **Calculate Values**.
10. Assign these records a value of **1**.
11. Click on the **Show: All** button at the bottom of the screen.
12. Click on the **Options** button at the bottom of the screen and select **Clear Selection**. Close the
    Attribute Table.

Correct for Parkway Access Points:
1. In your **View**, manually select one of each of the pair of access points resulting from divided
   parkways.
2. Open the **Attribute Table** of the **Access Point Layer**.
3. Click on the **Show: Selected** button at the bottom of the screen.
4. Right-click on the field name you just created (**Corrected**).
5. Select **Calculate Values**.
6. Assign these records a value of **2**.

7. Click on the **Show: All** button at the bottom of the screen.
8. Click on the **Options** button at the bottom of the screen and select **Clear Selection**. Close the **Attribute Table**.

Remove Unwanted Access Points from View/Table:
1. Right-click on your **Access Point Layer** and select **Properties**.
2. Click the tab **Definition Query**.
3. Click the button **Query Builder**.
4. Under **Fields** double-click on the field you just created (**Corrected**).
5. Click or type =. Under **Unique Values**, double-click on **0**.
6. Click **OK**. The **Query Builder** window will close.
7. Click **OK**. The **Layer Properties** window will close.
8. Now, access points resulting from the intersection of a site boundary and an Interstate or its on- and off-ramps and one of the paired divided parkway points will not be displayed. To display all access points again, repeat steps 1-7, but remove the existing equation in the **Query Builder**.

Summarize Access Point Counts for *Multiple Non-Overlapping Sites without Shared Edges:*
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using a **Spatial Join**, assign the IDs from the sites to each access point.
3. Open the **Attribute Table** of the **Access Point Layer** and select the field that contains the **IDs**.
4. Right-click on the field name and select **Summarize**.
5. Select or create an output file name and click **OK**.
6. The output .dbf file will contain the counts of access points in each site.

Summarize Access Point Counts for *Multiple Non-Overlapping Sites with Shared Edges:*
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using a **Spatial Join**, assign the IDs from the site to each access point (use **closest distance** features).
3. Access points that fall on shared edges will each have two records in the **Attribute Table**. When joined, both points will be given the same ID. Each shared access point must have one of the two IDs changed manually in the table.
4. Once step 3 is complete, open the **Attribute Table** of the **Access Point Layer** and select the field that contains the **IDs**.
5. Right-click on the field name and select **Summarize**.
6. Select or create an output file name and click **OK**.
7. The output .dbf file will contain the counts of access points in each site.
8.

Summarize Access Point Counts for *Multiple Overlapping Sites:*
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using your site polyline layer, select (highlight) the first study site.
3. Perform a **Select by Location**.
4. Under **I want to**, choose **select features from**.
5. Under **the following layers**, choose your **Access Point Layer**.
6. Under **that**, choose are **within a distance of**.
7. Under **the features in this layer**, choose your **Site Polyline Layer**.
8. Make sure the box for **Use selected features** is checked.
9. Check the box for **Apply a buffer…of** and fill in a very small distance (e.g., 1 meter).
10. Click **Apply**.
11. Open the **Attribute Table** of the **Access Point Layer**.
12. Text at the bottom of the window should show you how many access points are selected. This is the number of access point for the study site you currently have selected.
13. Record this number elsewhere (in a spreadsheet or Word file).
14. Repeat steps 2-6 for each study site.

6.5 Road Length per Unit Area

Revision Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used.

1. Basic Concept
Road length per hectare is related to both block size and connectivity.

The next version of these protocols will contain more detail about earlier uses of this measurement.

2. Basic Formula
Road Length per Unit Area = Length of road with both interstates and ramps removed and divided roads averaged, per measurement area.

3. Detailed Formula
Road Length per Unit Area = Length of road with both interstates and ramps removed and divided roads averaged, per measurement area with water removed from land area calculation.

4. Comments and Explanations
Road lengths are simply the sums of the lengths of all road centerlines bounded by the study sites. Similar to the process used for access points, we ignored non-pedestrian lengths by removing Interstates and on- and off-ramps from our calculations. Also, we removed one of each pair of centerlines from all parkways and other divided at-grade roads in order to count their lengths once instead of twice.

Once this was done we used the area of each site to calculate the road length per hectare figures for each site.

5. GIS Approach
Delete all Interstates and on- and off-ramps. Delete one of each pair of centerlines for parkways and at-grade divided roads. Intersect the roads layer with the measurement geographies. Recalculate the length of the road segments. Summarize the intersected layer by the measurement geography ID, calculating the sum of the Length field in the process. Divide the resulting sum of road lengths by the area in hectares of the associated measurement geography.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.
To follow this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Road Centerline Layer** (polyline)
3. **XTools** extension (available from the Downloads section of the ESRI Web site)
4. The protocol for determining **Net Land Area (without water)**

Calculate Road Lengths:
1. Using the **GeoProcessing Wizard**, intersect the **Road Centerline Layer** with the **Study-Site Layer**.
   a. Under the **Tools** menu, select **GeoProcessing Wizard**.
   b. Select **Intersect Two Layers** and click on **Next**.
   c. Under **Select the input layer to intersect**, choose your **Roads Layer**.
   d. Under **Select a polygon overlay layer**, choose your **Study-Site Layer**.

---

e. Give your file a name and click **Finish**.
2. Using **XTools**, recalculate the length of the road segments in the newly intersected layer.
    a. Under the **XTools** menu, select **Table Operations**.
    b. Choose **Calculate Area, Perimeter, Length, Acres, and Hectares**.
    c. Under **Select layer to measure**, choose your new **Roads Layer**.
    d. Choose your desired output units.
    e. Click **OK**. The lengths of the road segments in the **Length** field will be recalculated.

Correct for Interstate Road Segments:
1. Open the **Attribute Table** of your new **Roads Layer** (right-click on the layer).
2. Click on the **Options** button at the bottom of the window and select **Add Field**.
3. Enter a name for a field that will indicate if a road segment is an Interstate or an on- or off-ramp.
4. Click **OK**.
5. In your **View**, manually select the road segments that are Interstates or on- or off-ramps.
6. Return to the **Attribute Table** of your **Roads Layer**.
7. Click on the **Show: Selected** button at the bottom of the screen.
8. Right-click on the field name you just created.
9. Select **Calculate Values**.
10. Assign these records a value of **1**.
11. Click on the **Show: All** button at the bottom of the screen.
12. Click on the **Options** button at the bottom of the screen and select **Clear Selection**. Close the **Attribute Table**.

Correct for Divided At-grade Road Segments:
1. Open the **Attribute Table** of your new **Roads Layer**.
2. Click on the **Options** button at the bottom of the window and select **Add Field**.
3. Enter a name for a field that will indicate if a road segment is a divided at-grade road segment.
4. Click **OK**.
5. Click on the **Options** button at the bottom of the window and select **Add Field**.
6. Enter a name for a field that will indicate it contains an average length of a divided at-grade road segment.
7. Click **OK**.
8. For each site that contains divided at-grade road segments, select the pair of roadways that make up the divided at-grade road.
9. Return to the **Attribute Table** of your **Roads Layer**.
10. Click on the **Show: Selected** button at the bottom of the screen.
11. Right-click on the field name you just created.
12. Select **Calculate Values**.
13. Assign these records a value of **1**.
14. Calculated the total length of the two roadways.
15. Divide by **2** to find the average length.
16. Enter this length into the new **AVERAGE** field of any one record of those currently selected (leave the remained selected records as **0**).
17. Repeat steps 8-16 for each set of divided at-grade roads in the site.
18. Repeat steps 8-17 for each site.

Remove Unwanted Road Segments from View/Table:
1. Right-click on your **Access Point Layer** and select **Properties**.
2. Click the tab **Definition Query**.
3. Click the button **Query Builder**.
4. Under **Fields**, double-click on the field you just created for Interstates and on- and off-ramps.
5. Click or type =. Under **Unique Values**, double-click on **0**.
6. Click **OK**. The **Query Builder** window will close.
7. Click **OK**. The **Layer Properties** window will close.

8. Now, only those road segments that are NOT Interstates or on- and off-ramps will be displayed. To display all road segments again, repeat steps 1-7, but remove the existing equation in the **Query Builder**.

Summarize Road Lengths:
1. Open the **Attribute Table** of the new **Roads Layer**.
2. Click on the **Options** button at the bottom of the window and select **Select by Attributes**.
3. Under **Fields** double-click on the field you created for divided at-grade roadways.
4. Click or type =. Under **Unique Values**, double-click on **0**.
5. Click **Apply**. (You may now close the **Select by Attributes** window at any time.)
6. Click on the **Show: Selected** button at the bottom of the screen. Only the non-divided road segments will now be displayed.
7. Right-click on the **ID** field name and select **Summarize**.
8. In the window that opens, under step 2 select the **Length** field and check the box marked **sum**.
9. Select or create an output file name and click **OK**.
10. At the bottom of the window check the box marked **Summarize on the selected records only**.
11. The output .dbf file will contain the sums of the lengths of non-divided roads in each site.
12. Click on the **Options** button at the bottom of the window and select **Switch selection**. The divided road segments will now be selected.
13. Click on the **Show: Selected** button at the bottom of the screen. Only the divided road segments will now be displayed.
14. Right-click on the ID field name and select **Summarize**.
15. In the window that opens, under step 2 select the **Average Length** field and check the box marked **sum**.
16. Select or create an output file name and click **OK**.
17. At the bottom of the window check the box marked **Summarize on the selected records only**.
18. The output .dbf file will contain the sums of the average lengths of divided roads in each site.
19. In **Excel** or other spreadsheet program, open the two .dbf files and sum the road lengths to arrive at the total road lengths for each site.

Determine Net Land Area (without water) for each Site: Follow the steps outlined in the protocol above.
Calculate Road Length per Hectare:
1. Open the **Attribute Table** of the **Study-Site Layer** and right-click on the **ID** field name. Select **Summarize**.
2. In the window that opens, under step 2 select the **Hectares** field and check the box marked **sum**.
3. Select or create an output file name and click **OK**.
4. The output .dbf file will contain the total hectares in each site.
5. In **Excel** or other spreadsheet program, combine your total road length figures and your total hectares figures based on the site ID numbers.
6. For each site, divide total road length by total hectares to arrive at road length per hectare figures.

6.6 Intersections per Unit Area

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used. The FNode TNode script has not been tested in ArcGIS 9.1.

1. Basic Concept
The number of intersections per hectare is a measure of the connectivity of the street network (Cervero and Kockelman 1996, 8; McNally and Kulkarni 1997, 109). As can be seen from this and later protocols, however, judging an intersection can be complex, particularly in older areas, where intersections are not regularly spaced, road rights of way have varying sizes, and centerlines may be slightly offset within the right of way, creating false positives in terms of intersection counts (see below for more detail).

This is obviously an area where more research would help, to discover how much of a jog in an intersection is needed before it is perceived as two intersections. In the absence of such research, we spent some time observing difficult cases in the field.

2. Basic Formula
Intersections per Unit Area = Number of intersections per unit land area.

3. Detailed Formula
Intersections per Unit Area = Number of intersections per unit land area, excluding freeways with land area excluding water. Intersections are measured raw, and buffered by 10-meter and 15-meter buffers to deal with slightly offset centerlines and divided roads.
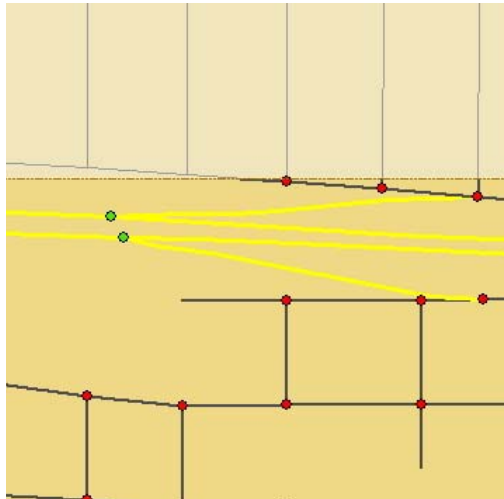
4. Comments and Explanations
As was explained in the introduction to this chapter, a key issue in measuring intersections is how the streets are entered into the GIS database—as a series of segments that end at each intersection or as long lines that overlap. These instructions are for the segment-by-segment type of database—other types will need to use slightly different ArcGIS operations although the concepts will be basically the same.

Fig. 6.5. All of our intersection counts excluded intersections formed by the convergence of Interstates and on- or off-ramps (see diagram to left), but included intersections between on- and off-ramps and local streets.

We also tried to correct for intersections where roads converged on an area, but did not perfectly align and, therefore, failed to provide our GIS operations with a single intersection point. To do this, we created buffers around each intersection point (both 10- and 15-meter buffers), merged them where they overlapped, and created new centroids to act as intersection points for our new intersection polygons (see diagrams below).
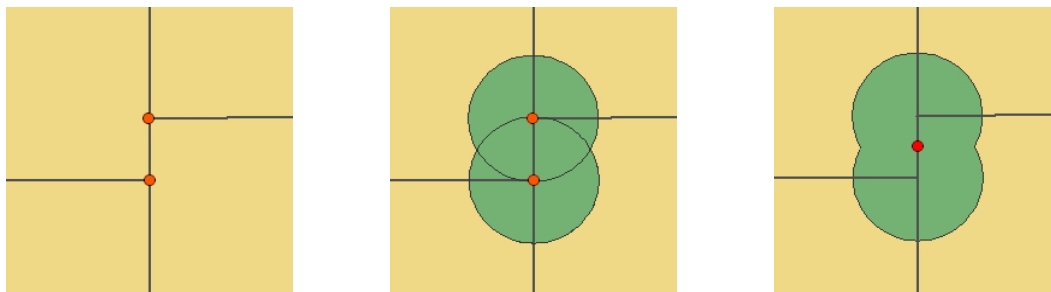
Fig. 6.6. This diagram shows the buffering process where two road centerlines are slightly offset, creating two T intersections in the GIS measurement but in the field, because the road has a width on either side of the center line, they would be perceived as one four-way intersection. In these cases we used buffering to merge the two intersections into one.

A buffer distance of 20 meters was also tried, but it was determined though fieldwork by the environmental measurement team, and then independently by the Principal Investigator, that the distance between roads was too great to consider the area as a singular functional

intersection. In order to make these judgments we took two field trips. First, the environmental team walked key intersections at the University and measured them to get a sense of road dimensions. Second, the Principal Investigator located part of the study area with nearby intersections that were 20 (10-meter buffers), 30, and 40 meters apart and went out on a sunny afternoon and observed how walkers were using the intersections. The first two functioned as one intersection, the last did not, although the specific case could vary with the width of the paved areas, the presence or absence of sidewalks, and the location of street trees.)

5. GIS Approach
Remove Interstates from the road centerline file. Use the Calculate Fnode Tnode script to create the intersection point file. Remove points with a valence of 1 or 2. Buffer intersections with 10- and 15-meter buffers and create new intersection centroids. Intersect point file with measurement geography and summarize intersection counts per unit area.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Road Centerline Layer** (polyline)
3. **Calculate Fnode Tnode ArcScript** available from the Downloads section of the ESRI Web site
4. **XTools** extension (also available from the Downloads section of the ESRI Web site)
5. The protocol for determining **Net Land Area (without water)**

Create the intersections layer:
1. Using **ArcMap**, install the **Calculate Fnode Tnode** script following the instructions included with the file.
2. Add to your **Map Document** the **Road Centerline Layer**.
3. Create a point file (layer) that will eventually contain the output of the intersection node procedure.
4. Run the **Calculate Fnode Tnode** tool.
5. When prompted, select your **Node Centerline Layer** as your polyline layer and the new point layer as your point layer.
6. Select **0** as the beginning ID number for the nodes.
7. After the procedure concludes, your **Point Layer** will now contain points created from the nodes of the **Road Centerline Layer**. The numbers in the **Valence** field represent the number of roads that converge to create the node (e.g., a valence of 1 means the node is a dead end or cul-de-sac; a Valence of **3** indicates a three-way intersection, etc.).

Clip intersections to study sites:
1. Using **GeoProcessing Wizard**, clip the **Intersections Layer** to your study site(s).
   a. In the **GeoProcessing Wizard**, select **Clip one layer based on another**.
   b. Click on **Next**.
   c. Under **Select the input layer to clip**, select your **Intersections Layer**.
   d. Under **Select a polygon clip layer**, select your **Study-Site Layer**.
   e. Give your clipped layer a name and click on **Finish**.
2. The new clipped layer will open separately.

Correct for Interstate/ramp intersections:
1. Open the **Attribute Table** of the (clipped) **Intersections Layer**.
2. Click on the **Options** button at the bottom of the window and select **Add Field**.
3. Enter the field name **Corrected**.
4. Click **OK**.
5. In your **View**, manually select the intersections formed by the convergence of Interstates and on- or off-ramps.

---

6. Return to the **Attribute Table** of your **Intersections Layer**.
7. Click on the **Show: Selected** button at the bottom of the screen.
8. Right-click on the field name you just created (**Corrected**).
9. Select **Calculate Values**.
10. Assign these records a value of **1**.
11. Click on the **Show: All** button at the bottom of the screen.
12. Click on the **Options** button at the bottom of the screen and select **Clear Selection**. Close the **Attribute Table**.

Remove Unwanted Intersections from View/Table:
1. Right-click on your **Intersections Layer** and select **Properties**.
2. Click the tab **Definition Query**.
3. Click the button **Query Builder**.
4. Under **Fields**, double-click on the field you just created (**Corrected**).
5. Click or type **=**. Under **Unique Values**, double-click on **0**.
6. Click or type **AND**.
7. Under **Fields** double-click on **valence**.
8. Click or type **>**. Under **Unique Values**, double-click on **2**.
9. Click **OK**. The **Query Builder** window will close.
10. Click **OK**. The **Layer Properties** window will close.
11. Now, only those intersections NOT formed by the convergence of Interstates and on- or off-ramps AND created by the intersection of three or more road segments (no culs-de-sac) will be displayed. To display all intersections again, repeat steps 1-10 but remove the existing equation in the **Query Builder**.

Create Buffered Intersections:
1. Under the **Tools** dropdown menu, select **Buffer Wizard**.
   a. In the **Buffer Wizard** window, under **Features of a layer**, select your (clipped) **Intersections Layer** and click on **Next**.
   b. Click **At a specified distance** and enter **XX** (your desired distance)
   c. After **Distance units are**, select **meters**.
   d. Click **Next**.
   e. After **Dissolve barriers between**, select **Yes**.
   f. Give the layer a name and save as a new layer
   g. Click **Finish**.
2. Using **XTools**, select **Feature Conversions** and then **Shapes to Centroids**.
   a. Select the new buffer layer (polygon)
   b. Give the point layer you are about to create a name and click **OK**.
   c. The new layer will contain points that represent merged intersections—those resulting from two intersections that are within 2(XX) meters of each other (a XX meter radius around each intersection) and essentially functioning as one intersection for the purposes of this analysis.

Summarize Intersection Counts for Multiple Non-overlapping Sites:
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using a **Spatial Join**, assign the IDs from the sites to each intersection.
3. Open the **Attribute Table** of **the Intersections Layer** and select the field that contains the IDs.
4. Right-click on the field name and select **Summarize**.
5. Select or create an output file name and click **OK**.
6. The output .dbf file will contain the counts of intersection nodes in each site.

Summarize Intersection Counts for Multiple Overlapping Sites:
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using your site polygon layer, select (highlight) the first study site.
3. Perform a **Select by Location**.
   a. Under **I want to**, choose **select features from**.
   b. Under **the following layers**, choose **your intersection node layer**.

      c.   Under **that**, choose **have their center in**.
      d.   Under the **features** in this layer, choose **your site polygon layer**.
      e.   Make sure the box for **Use selected features** is checked.
      f.   Click **Apply**.
4.   Open the **Attribute Table** of the **Intersection Node Layer**.
5.   Text at the bottom of the window should show you how many intersections are selected. This is the number of intersections in the study site you currently have selected.
6.   Record this number elsewhere (in a spreadsheet or Word file).
7.   Repeat steps 2-6 for each study site.

Determine Net Land Area (without water) for each Site: Follow the steps outlined in the protocol above. Calculate intersections per hectare for each site:
1.   Open the **Attribute Table** of the **Study-Site Layer** and right-click on the **ID** field name. Select **Summarize**.
2.   In the window that opens, select the **Hectares** field under step 2 and check the box marked **sum**.
3.   Select or create an output file name and click **OK**.
4.   The output .dbf file will contain the total hectares in each site.
5.   In **Microsoft Excel** or other spreadsheet program, combine intersection node counts and your total hectares figures based on the site ID numbers.
6.   For each site, divide the number of intersection nodes by total hectares to get the number of intersections per hectare.

6.7 Ratio of Four-Way Intersections to All Intersections

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used. The FNode TNode script has not been tested in ArcGIS 9.1.

1. Basic Concept
Four-way intersections are an indicator of grid street patterns, believed by some to be more connected. One can imagine a high proportion of four-way intersections in an area, however, with enormous blocks, so the density of intersections should also be considered.

2. Basic Formula
Ratio of Four-Way Intersections to All Intersections = Raw percentage of four-way intersections versus other intersections (valence 3+).

3. Detailed Formula
Ratio of Four-Way Intersections to All Intersections = Raw percentage of four-way intersections versus other intersections (valence 3+), with the added complication that some higher valence intersections are actually three-way and four-way intersections.

4. Comments and Explanations
As outlined in Intersections per Unit Area, intersection counts in our study sites (1) did not include intersections formed by the convergence of Interstates and on- or off-ramps and (2) corrected for misaligned intersections by creating buffers around each intersection node, merging them where they overlapped, and creating a centroid for each new intersection polygon.

In order to be able to classify these new nodes, we created "rules" to assign types (three-way, four-way) to the intersections. These assignments were based on the sum of the valences of the original intersections (the valence of an intersection, based on the script that creates the

intersection nodes, is the count of the number of road segments that converge to create the node—a valence of 1 is a dead end or cul-de-sac, a valence of 3 is a three-way intersection, a valence of 4 is a four-way intersection, etc.).

We realized, in particular, that intersections with valences of 6 were typically the result of two intersections that were not quite aligned but that functioned as four-way intersections. Intersections with valences of 8 were almost always the result of intersections of divided and non-divided roads that functioned as four-way intersections. Intersections with valences of 16 were the result of intersections of two divided roads that functioned as four-way intersections.

Fig. 6.7. Four-way intersections:

Valence of 4   Valence of 6   Valence of 8

Valence of 16
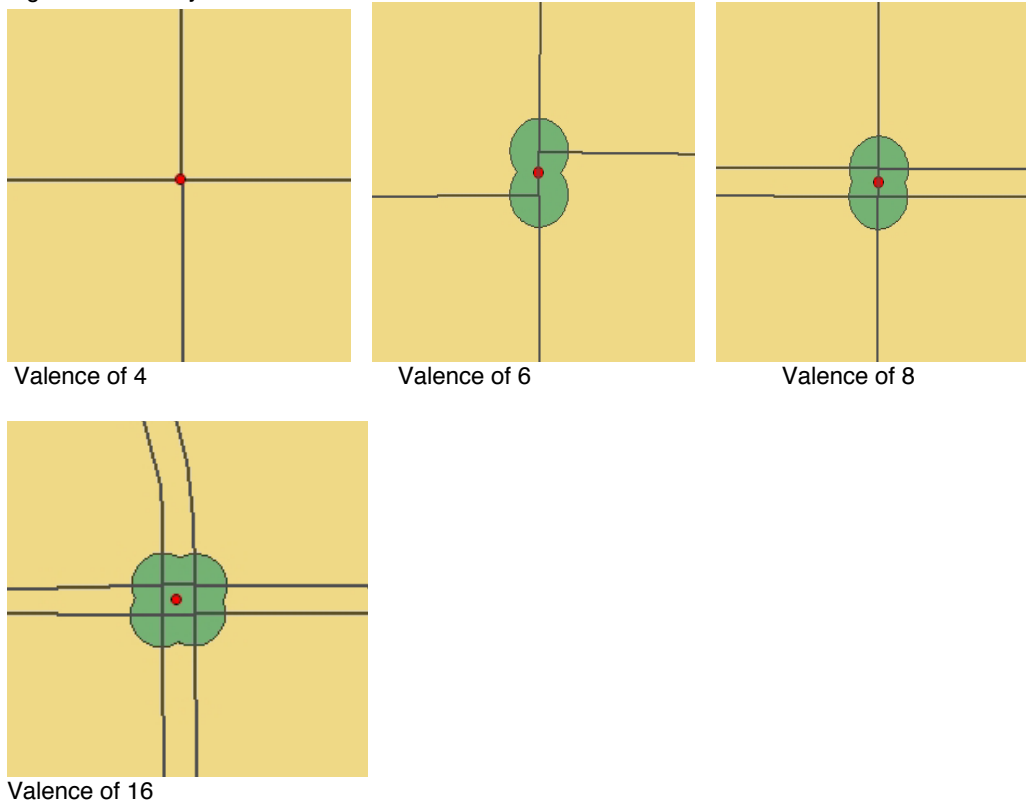
We, likewise, found that most newly formed intersections with valences of 7 were the result of three-way intersections of divided and non-divided roads. Similarly, about half of the intersections with valences of 12 were the result of three-way intersections of two divided roads. These seemed to function as three-way intersections.

Fig. 6.8. Three-way intersections:

Valence of 3   Valence of 7   Valence of 12
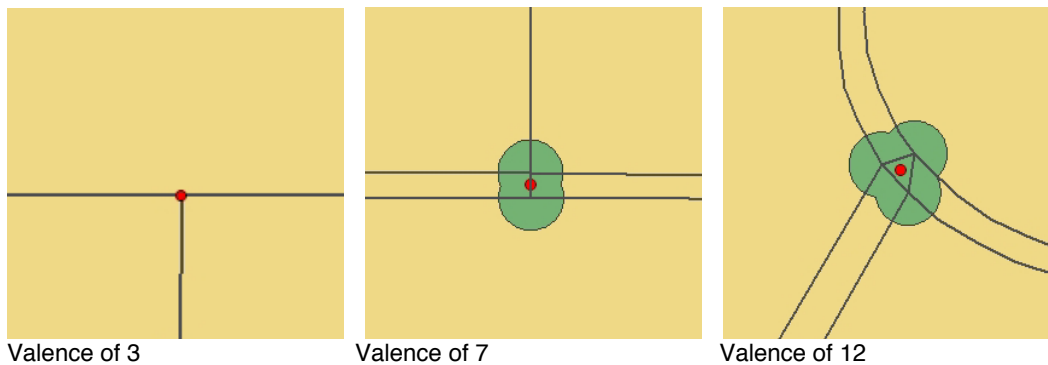
There were some intersections with valences of 7 and 12 that did not conform. Specifically, there were a handful of merged intersections with valences of 7 and 12 that were the formed by the convergence of the gridded street pattern with a diagonal. In addition, there was one intersection with a valence of 12 that seemed to function more like a four-way intersection. Because there were only three intersections of valence 12 using 10-meter buffers (two three-way and one four-way) and five using 15-meter buffers (two three-way and three four-way), we chose to apply our three-way and four-way intersection rules by hand instead of using a blanket rule.

Fig. 6.9. The images below were classed as four-way.

Valence of 7(b)   Valence of 12(b)   Valence of 12(c)
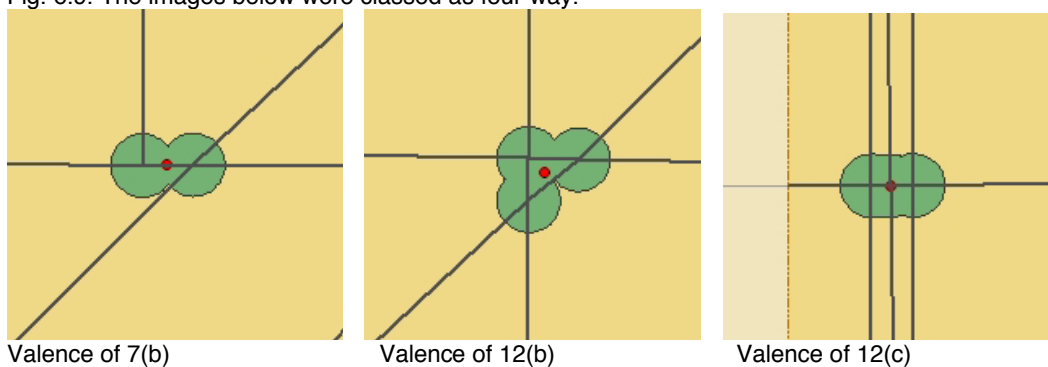
Finally, there were a number of other new intersection types that were more complex than either three-way or four-way intersections. These were partly artifacts of our trying to deal with classifying functional three- and four-way intersections through buffering. These are intersections that show a high level of connectivity, which is the point of the four-way intersection count, so they were counted as four-way intersections. *Note: the intersection referenced is outlined with a dotted oval.*

Fig. 6.10. Other variations:
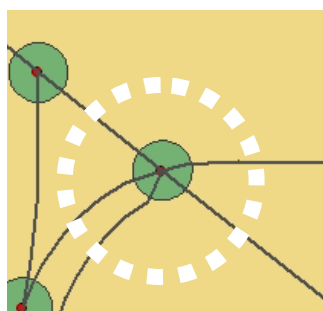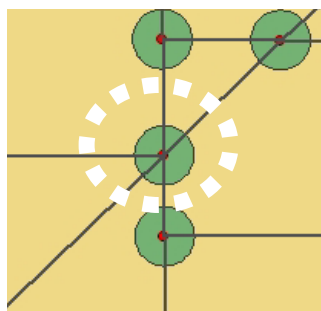
Valence of 5(a) Valence of 5(b) Valence of 9 (should really be 13)
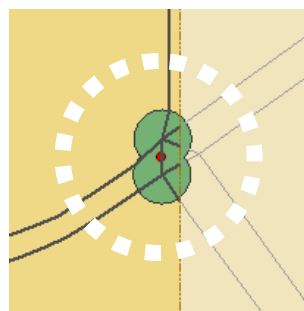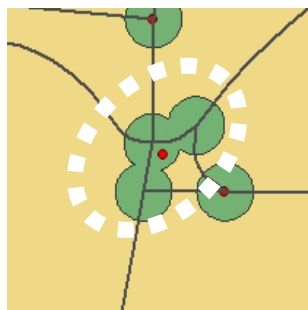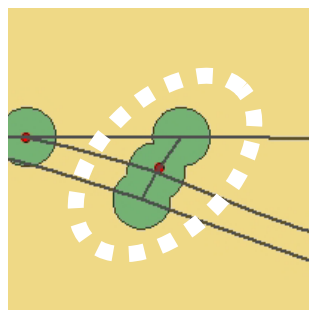
Valence of 10(a) Valence of 10(b) Valence of 10(c)
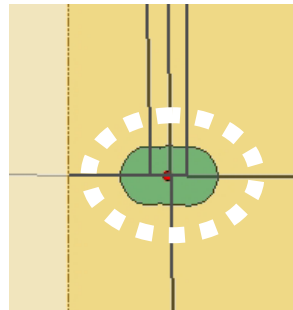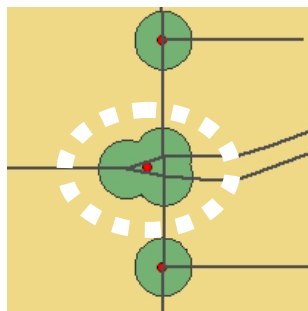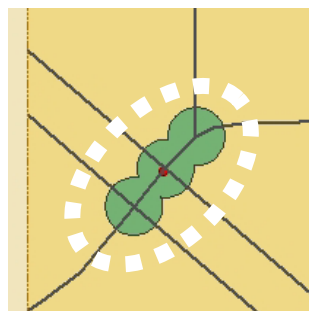
Valence of 11(a) Valence of 11(b) Valence of 13
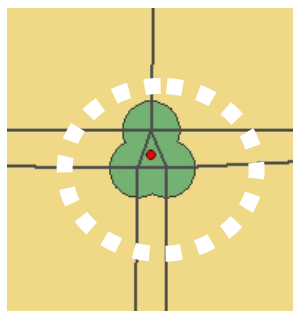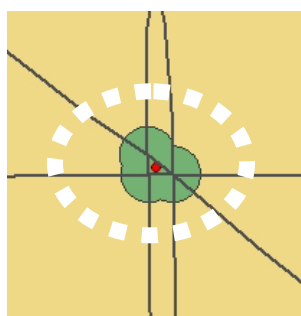
Valence of 14

For all our intersection counts, we ran our numbers three times. During the first run, we counted intersections without any buffers. During the second run, we counted intersections that were buffered 10 meters and adjusted for valences of 6, 7, 8, 12, and 16 and other adjustments as detailed above. Finally, in our third run, we counted intersections that were buffered 15 meters and adjusted for valences of 6, 7, 8, 12, and 16 and other adjustments as detailed above.

5. GIS Approach
Remove Interstates from the road centerline file. Use the Calculate Fnode Tnode script to create the intersection point file. Remove points with a valence of 1 or 2. Buffer intersections with 10- and 15-meter buffers and create new intersection centroids with reassigned valences. Intersect pointfile with measurement geography and calculate ratios of four-way intersection counts to total intersection counts.

6. GIS Steps [Note: these need to be adjusted for valences over 4.]
The following procedures were completed using ArcMap (ArcInfo) 8.3.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Road Centerline Layer** (polyline)
3. **Intersection Node Layer** (point)
4. **Calculate Fnode Tnode** ArcScript available from the Downloads section of the ESRI Web site
5. **XTools** extension (also available from the Downloads section of the ESRI Web site)

Total intersection counts from protocol for calculating **Intersections per Unit Area**

Create the intersections file:
1. Using **ArcMap**, install the **Calculate Fnode Tnode** script following the instructions included with the file.
2. Add to your **Map Document** the **Road Centerline Layer**.
3. Create a point file (layer) that will eventually contain the output of the intersection node procedure.
4. Run the **Calculate Fnode Tnode** tool.
5. When prompted, select your **Node Centerline Layer** as your polyline layer and the new point layer as your point layer.
6. Select **0** as the beginning ID number for the nodes.
7. After the procedure concludes, your point layer will now contain points created from the nodes of the **Road Centerline Layer**. The numbers in the **valence** field represent the number of roads that converge to create the node (e.g., a valence of 1 means the node is a dead end or cul-de-sac; a Valence of **3** indicates a three-way intersection; etc.).

Clip intersections to study sites:
1. Using **GeoProcessing Wizard**, clip the **Intersection Layer** to your **Study-Site Layer**.
2. In the **GeoProcessing Wizard**, select **Clip one layer based on another**.
3. Click on **Next**.
4. Under **Select the input layer to clip**, select your **Intersection Layer**.
5. Under **Select a polygon clip layer**, select your **Study-Site Layer**.
6. Give your clipped layer a name and click on **Finish**.
7. The new clipped layer will open separately.

Correct for Interstate/ramp intersections:
1. Open the **Attribute Table** of the (clipped) **Intersection Layer**.
2. Click on the **Options** button at the bottom of the window and select **Add Field**.
3. Enter the field name **Corrected**.
4. Click **OK**.
5. In your view, manually select the intersections formed by the convergence of Interstates and on- or off-ramps.
6. Return to the **Attribute Table** of your **Intersection Layer**.
7. Click on the **Show: Selected** button at the bottom of the screen.
8. Right-click on the field name you just created (**Corrected**).

9. Select **Calculate Values**.
10. Assign these records a value of **1**.
11. Click on the **Show: All** button at the bottom of the screen.
12. Click on the **Options** button at the bottom of the screen and select **Clear Selection**. Close the **Attribute Table**.

Remove unwanted intersections from view/table:
1. Add the **Intersection Layer** to your **Map Document**.
2. Right-click on the **intersection layer** and select **Properties**.
3. Click the tab that says **Definition Query**.
4. Click the button that says **Query Builder**.
5. Under **Fields**, double-click on the field **Corrected**.
6. Click or type =. Under **Unique Values**, double-click on **0**.
7. Click or type **AND**.
8. Under **Fields**, double-click on **Valence**.
9. Click or type >. Under **Unique Values**, double-click on **2**.
10. Click **OK**. The **Query Builder** window will close.
11. Click **OK**. The **Layer Properties** window will close.
12. Now, only those intersections NOT formed by the convergence of Interstates and on- or off-ramps AND created by the intersection of three or more road segments (no culs-de-sac) will be displayed. To display all intersections again, repeat steps 1-10, but remove the existing equation in the **Query Builder**.

Create buffered intersections:
1. Under the **Tools** dropdown menu, select **Buffer Wizard**.
2. In the **Buffer Wizard** window, under **features of a layer**, select your clipped **Intersection Layer** and click on **Next**.
3. Click **At a specified distance** and enter **XX** (your desired distance).
4. After **Distance units** are, select **meters**.
5. Click **Next**.
6. After **Dissolve barriers between**, select **Yes**.
7. Give the layer a name and save as a new layer.
8. Click **Finish**.
9. Using **XTools**, select **Feature Conversions** and then **Shapes to Centroids**.
10. Select the new buffer layer (polygon).
11. Give the point layer you are about to create a name and click **OK**.
12. The new layer will contain points that represent merged intersections – those resulting from two intersections that are within 2(XX) meters of each other (a XX meter radius around each intersection) and essentially functioning as one intersection for the purposes of this analysis.

Assign new valences to buffered intersections:
1. Perform a Spatial Join on your new buffer layer.
   a. Right-click on your new layer and select **Joins and Relates** and then **Join…**
   b. Under **What do you want to join to this layer?** select **Join data from another layer based on spatial location.**
   c. Under **1. Choose the layer to join to this layer**, select your clipped **Intersection Layer** (point).
   d. Under **2. You are joining**, select the first option **Each polygon will be given a summary**.
   e. Under **How do you want the attributes to be summarized?** check the box for **Sum**.
   f. Under **3. The result of the join will be saved**, give your new layer a name.
   g. Click **OK**.
2. Open the **Attribute Table** of your new layer and remove all fields except **Count_ and Sum_valenc** (TIP: use **XTools** for multiple-field deletes).
3. Perform a **Spatial Join** on your centroid layer (created in the previous step).
   a. Right-click on your centroid layer and select **Joins and Relates** and then **Join**…

---

b. Under **What do you want to join to this layer?** select **Join data from another layer based on spatial location**.
c. Under **1. Choose the layer to join to this layer**, select your newly joined buffer layer.
d. Under **2. You are joining**, select the first option **it falls inside**.
e. Under **3. The result of the join will be saved**, give your new layer a name
f. Click **OK**.
4. Open the **Attribute Table** of your new layer and remove all fields except **Count_ and Sum_valenc**.
5. Click on the **Options** button at the bottom of the screen and select **Add Field…**.
6. Give your field the name **New_valenc** and click **OK**.
7. Right-click on **New_valenc** and select **Calculate Values…**
8. Double-click **Sum_valenc** and click **OK**. The values from the **Sum_valenc** field will have been copied over to the **New_valenc** field.
9. Click on the **Options** button and select **Select by Attributes**.
10. Set **Count** to **> 1**.
11. Click on **Apply**.
12. Click on the **Show: Selected** button at the bottom of the screen.
13. Click on the **Options** button and select **Select by Attributes** (this window may still be open).
14. After **Method**, select **Select from current selection**.
15. Set **Sum_valenc** to **= 6**.
16. Right-click **New_valenc** and select **Calculate Values**….
17. Enter **4** in the box below **New_valenc =** and click **OK**.
18. Repeat steps 9 through 17 twice using **Sum_valenc = 8** and **= 16** in step 15.
19. Click on the **Show: All** button at the bottom of the screen.
20. Click on the **Options** button at the bottom of the screen and select **Clear Selection**.
21. Right-click the layer and select **Properties**…
22. Click the tab **Definition Query**.
23. Click the button **Query Builder**.
24. Under **Fields** double-click the field **New_valenc**.
25. Click or type =. Under **Unique Values**, double-click on **4**.
26. Click **OK**. The **Query Builder** window will close.
27. **Click "OK."** The **Layer Properties** window will close.
28. Now, only those intersections with a valence of 4 (four-way intersections) will be displayed. To display all intersections again, repeat steps 21-27, but remove the existing equation in the **Query Builder**.

Summarize four-way intersection counts for multiple non-overlapping sites:
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using a **Spatial Join**, assign the IDs from the sites to each intersection.
3. Open the **Attribute Table** of the **Intersection Layer** and select the field that contains the IDs.
4. Right-click the field name and select **Summarize**.
5. Select or create an output file name and click **OK**.
6. The output .dbf file will contain the counts of intersection nodes in each site.

Summarize four-way intersection counts for multiple overlapping sites:
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using your **site polygon layer**, select (highlight) the first study site.
3. Perform a **Select by Location**.
4. Under **I want to**, choose **select features from**.
5. Under **the following layers**, choose your **Intersection Node Layer**.
6. Under **that**, choose **have their center in**.
7. Under **the features in this layer**, choose your site polygon layer.
8. Make sure the box for **Use selected features** is checked.
9. Click **Apply**.
10. Open the **Attribute Table** of the **Intersection Node Layer**.

11. Text at the bottom of the window should show you how many intersections are selected. This is the number of intersections in the study site you currently have selected.
12. Record this number elsewhere (in a spreadsheet or Word file).
13. Repeat steps 2-6 for each study site.

Calculate four-way intersection ratios for each site
1. In a spreadsheet program such as **Microsoft Excel**, open the file containing the summary of total intersection counts for each site created in protocol for **Intersections per Unit Area**.
2. Open the file containing the summary of four-way intersection counts for each site.
3. Divide the numbers of four-way intersections by the total numbers of intersections to get the four-way intersection ratios for each site.

6.8 Four-Way Intersections per Unit Land Area

1. Basic Concept
Four-way intersections are an indicator of grid street patterns, thought to be supportive for walking. The measure of four-way intersections per hectare adjusts this measure to reflect different block sizes.

2. Basic Formula
Four-Way Intersections per Unit Land Area = Number of four-way intersections per unit land area.

3. Detailed Formula
Four-Way Intersections per Unit Land Area = Number of four-way intersections per unit land area excluding water, where four-way intersections are defined as defined in the comments and explanations.

4. Comments and Explanations
For a discussion of what we consider to be four-way intersections, see the previous protocol on the Ratio of Four-Way Intersections to all Intersections.

5. GIS Approach
For each study site, the four-way intersection counts were matched to the net land area through the site ID numbers. The intersection counts were divided by the net hectares to derive the number of four-way intersections per hectare for each study site.

6. GIS Steps
The following procedures were completed using ArcMap (ArcInfo) 8.3.

To follow this protocol you will need:
1. **Four-way Intersection Counts** from protocol for calculating **Ratio of Four-Way Intersections to All Intersections**
2. The protocol for determining **Net Land Area (without water)**

Determine Net Land Area (without water) for each Site:
1. Follow the steps outlined in the protocol above.

Calculate Four-Way Intersections per Hectare for each Site:
1. In **Microsoft Excel** or other spreadsheet program, combine the four-way intersection counts and your net land area figures based on the site ID numbers.

---

2. For each site, divide the number of four-way intersections by net hectares to get the number of four-way intersections per hectares.

6.9 Ratio of Three-Way Intersections to All Intersections

Note: As of ArcGIS ArcInfo 9.1, the operations performed by XTools can be performed from the ArcGIS Toolbox. The protocol detailed below has not been tested in ArcGIS 9.1, but should work provided the latest XTools version is used. The FNode TNode script has not been tested in ArcGIS 9.1.

1. Basic Concept
This is another take on the issue of street connections.

2. Basic Formula
Ratio of Three-Way Intersections to All Intersections = Raw percent of three-way intersections versus other intersections (valence 3+).

3. Detailed Formula
Same as above except to note that some higher valence intersections were reclassified as three-way if that is how they operated, i.e., some valences of 7 and of 12. Buffering was at both 10 and 15 meters.

4. Comments and Explanations
As outlined in Intersections per Unit Area, intersection counts in our study sites (1) did not include intersections formed by the convergence of Interstates and on- or off-ramps and (2) corrected for misaligned intersections by creating buffers around each intersection node, merging them where they overlapped, and creating a centroid for each new intersection polygon.

In order to be able to classify these new nodes, we created "rules" to assign types (three-way, four-way) to the intersections. These are described above in the protocol on the Ratio of Four-Way Intersections to all Intersections.

For our intersection ratios, we ran our numbers three times. During the first run, we counted intersections without any buffers. During the second run, we counted buffered intersections, but did not adjust for new valence figures (therefore, a valence of 7 resulting from the intersection of roads with valences of 3 and 4 would remain a valence of 7). Finally, in our third run, we included some valences of both 7 and of 12 in our count of three-way intersections.

5. GIS Approach
Remove Interstates from the road centerline file. Use the Calculate Fnode Tnode script to create the intersection point file. Remove points with a valence of 1 or 2. Buffer intersections with 10- and 15-meter buffers and create new intersection centroids with reassigned valences. Intersect pointfile with measurement geography and calculate ratios of three-way intersection counts to total intersection counts.

6. GIS Steps [Also check this for manual corrections to valence=7 and =12]
The following procedures were completed using ArcMap (ArcInfo) 8.3.

To follow this protocol you will need:
   1. **Study-Site Layer** (polygon)
   2. **Road Centerline Layer** (polyline)
   3. **Intersection Node Layer** (point)

---

4. **Calculate Fnode Tnode ArcScript** available from the Downloads section of the ESRI Web site
5. **XTools** extension (also available from the Downloads section of the ESRI Web site)
6. Total intersection counts from protocol for calculating **Intersections per Unit Area**

Create the Intersections file:
1. Using **ArcMap**, install the **Calculate Fnode Tnode** script following the instructions included with the file.
2. Add to your **Map Document** the **Road Centerline Layer**.
3. Create a point file (layer) that will eventually contain the output of the intersection node procedure.
4. Run the **Calculate Fnode Tnode** tool.
5. When prompted, select your **Node Centerline Layer** as your polyline layer and the new point layer as your point layer.
6. Select **0** as the beginning ID number for the nodes.
7. After the procedure concludes, your point layer will now contain points created from the nodes of the **Road Centerline Layer**. The numbers in the "valence" field represent the number of roads that converge to create the node (e.g., a valence of 1 means the node is a dead end or cul-de-sac; a valence of 3 indicates a three-way intersection, etc.).

Clip Intersections to Study Sites:
1. Using **GeoProcessing Wizard**, clip the **Intersection Layer** to your study site(s).
2. In the **GeoProcessing Wizard**, select **Clip one layer based on another**.
3. Click on **Next**.
4. Under **Select the input layer to clip**, select your **Intersection Layer**.
5. Under **Select a polygon clip layer**, select your **Study-Site Layer**.
6. Give your clipped layer a name and click **Finish**.
7. The new clipped layer will open separately.

Correct for Interstate/ramp intersections:
1. Open the **Attribute Table** of the clipped **Intersection Layer**.
2. Click on the **Options** button at the bottom of the window and select **Add Field**.
3. Enter the field name **Corrected**.
4. Click **OK**.
5. In your view, manually select the intersections formed by the convergence of Interstates and on- or off-ramps.
6. Return to the **Attribute Table** of your **Intersection Layer.**
7. Click **Show: Selected** button at the bottom of the screen.
8. Right-click on the field name you just created (**Corrected**).
9. Select **Calculate Values**.
10. Assign these records a value of **1**.
11. Click on the **Show: All** button at the bottom of the screen.
12. Click on the **Options** button at the bottom of the screen and select **Clear Selection**. Close the **Attribute Table**.

Remove unwanted intersections from view/table:
1. Add the intersection layer to your Map Document.
2. Right-click on the **Intersection Layer** and select **Properties**.
3. Click the tab **Definition Query**.
4. Click the button **Query Builder**.
5. Under **Fields** double-click the field **Corrected**.
6. Click or type =. Under **Unique Values**, double-click on **0**.
7. Click or type **AND**.
8. Under **Fields** double-click on **valence**.
9. Click or type **>**. Under **Unique Values**, double-click on **2**.
10. Click **OK**. The **Query Builder** window will close.
11. Click **OK**. The **Layer Properties** window will close.
12. Now, only those intersections NOT formed by the convergence of Interstates and on- or off-ramps AND created by the intersection of three or more road segments (no culs-de-sac) will be displayed. To display all intersections again, repeat steps 1-10 but remove the existing equation in the **Query Builder**.

Create buffered intersections:
1. Under the **Tools** dropdown menu, select **Buffer Wizard**.
2. In the **Buffer Wizard** window, under **features of a layer**, select your clipped **Intersection Layer** and click on **Next**.
3. Click **At a specified distance** and enter **XX** (your desired distance).
4. After **Distance units are**, select **meters**.
5. Click **Next**.
6. After **Dissolve barriers between**, select **Yes**.
7. Give the layer a name and save as a new layer.
8. Click **Finish**.
9. Using **XTools**, select **Feature Conversions** and then **Shapes to Centroids**.
10. Select the new buffer layer (polygon).
11. Give the point layer you are about to create a name and click **OK**.
12. The new layer will contain points that represent merged intersections – those resulting from two intersections that are within 2(XX) meters of each other (a XX meter radius around each intersection) and essentially functioning as one intersection for the purposes of this analysis.

Assign new valences to buffered intersections:
1. Perform a Spatial Join on your new buffer layer
   a. Right-click on your new layer and select **Joins and Relates** and then **Join**.
   b. Under **What do you want to join to this layer?** select **Join data from another layer based on spatial location**.
   c. Under **1. Choose the layer to join to this layer**, select your clipped **Intersection Layer** (point).

d. Under **2. You are joining**, select the first option **Each polygon will be given a summary**.
e. Under **How do you want the attributes to be summarized?** check the box for **Sum**.
f. Under **3. The result of the join will be saved**, give your new layer a name.
g. Click on **OK**.
2. Open the **Attribute Table** of your new layer and remove all fields except **Count_ and Sum_valenc** (TIP: use **XTools** for multiple-field deletes).
3. Perform a Spatial Join on your centroid layer (created in the previous step)
a. Right-click on your centroid layer and select **Joins and Relates** and then **Join**.
b. Under **What do you want to join to this layer?** select **Join data from another layer based on spatial location**.
c. Under **1. Choose the layer to join to this layer**, select your newly joined buffer layer.
d. Under **2. You are joining**, select the first option **it falls inside**.
e. Under **3. The result of the join will be saved**, give your new layer a name.
f. Click on **OK**.
4. Open the **Attribute Table** of your new layer and remove all fields except **Count_ and Sum_valenc**.
5. Click on the **Options** button at the bottom of the screen and select **Add Field**.
6. Give your field the name **New_valenc** and click on **OK**.
7. Right-click on **New_valenc** and select **Calculate Values**.
8. Double-click on **Sum_valenc** and click on **OK**. The values from the Sum_valenc field will have been copied over to the New_valenc field.
9. Click on the **Options** button and select **Select by Attributes**.
10. Set **Count** to **> 1**.
11. Click on **Apply**.
12. Click on the **Show: Selected** button at the bottom of the screen.
13. Click on the **Options** button and select **Select by Attributes** (this window may still be open).
14. After **Method**, select **Select from current selection**.
15. Set **Sum_valenc** to **= 7**.
16. Right-click on **New_valenc** and select **Calculate Values**.
17. Enter **3** in the box below **New_valenc =** and click on **OK**.
18. Repeat steps 9 through 17 using **Sum_valenc** = **12** in step 15.
19. Click on the **Show: All** button at the bottom of the screen.
20. Click on the **Options** button at the bottom of the screen and select **Clear Selection**.
21. Right-click on the layer and select **Properties**.
22. Click the tab **Definition Query**.
23. Click the button **Query Builder**.
24. Under **Fields**, double-click on the field **New_valenc**.
25. Click or type =. Under **Unique Values**, double-click on **3**.
26. Click **OK**. The **Query Builder** window will close.
27. Click **OK**. The **Layer Properties** window will close.

Now, only those intersections with a valence of 3 (three-way intersections) will be displayed. To display all intersections again, repeat steps 21-27, but remove the existing equation in the **Query Builder**.

Summarize three-way intersection counts for multiple non-overlapping sites:
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using a **Spatial Join**, assign the IDs from the sites to each intersection.
3. Open the **Attribute Table** of the **Intersection Layer** and select the field that contains the IDs.
4. Right-click on the field name and select **Summarize**.
5. Select or create an output file name and click **OK**.
6. The output .dbf file will contain the counts of intersection nodes in each site.

Summarize three-way intersection counts for multiple overlapping sites:
1. If the **Study-Site Layer** does not have unique IDs or names for each site, create them now.
2. Using your site polygon layer, select (highlight) the first study site.
3. Perform a **Select by Location**.
4. Under **I want to**, choose **select features from**.
5. Under **the following layers**, choose your **Intersection Node Layer**.

6. Under **that**, choose **have their center in**.
7. Under **the features in this layer**, choose your site polygon layer.
8. Make sure the box for **Use selected features** is checked.
9. Click **Apply**.
10. Open the **Attribute Table** of the intersection node layer.
11. Text at the bottom of the window should show you how many intersections are selected. This is the number of intersections in the study site you currently have selected.
12. Record this number elsewhere (in a spreadsheet or Word file).
13. Repeat steps 2-6 for each study site.

Calculate three-way intersection ratios for each site
1. In a spreadsheet program, such as **Microsoft Excel**, open the file containing the summary of total intersection counts for each site created in protocol for **Intersections per Unit Area**.
2. Open the file containing the summary of three-way intersection counts for each site.
3. Divide the numbers of three-way intersections by the total numbers of intersections to get the three-way intersection ratios for each site.

6.10 Connected Node Ratio

1. Basic Concept
This is a measure of the street pattern and a ratio of the number of intersections (presumably valence 3 or more) divided by the number of intersections plus culs-de sac (with culs de sac valence 1) (Song and Knaap 2004, 214; Dill 2003, 3). A value close to 1 means that there are few culs-de-sac and, therefore, the street pattern is more highly connected.

Of course there could be a perfect grid with enormous blocks that were hard to get around and such a pattern would fare well on this measure, so it is not a perfect indicator

2. Basic Formula
Connected Node Ratio = Number of street intersections divided by number of intersections plus culs-de-sac (Dill 2003, 3)

3. Detailed Formula
Connected Node Ratio = Number of street intersections (valence 3 or more) divided by the number of intersections plus culs-de-sac (valence 1) (Dill 2003, 3)

4. Comments and Explanations
A connected node ratio is the ratio of total intersections to intersections plus culs-de-sac in a given study site. The maximum value of a connected node ratio is 1.00 indicated that there were no culs-de-sac within the study site. Dill (2003, 3) reports that ratios of 0.70 are preferred by those promoting highly-connected street patterns.

For our analysis, we calculated connected node ratios with the intersections created between Interstates and on- and off-ramps removed. Given the coarse nature of the measure—and the fact that the main distinction is between valence 1, valence 2 (omitted), and all other valences—there is not much difference between using buffered and unbuffered intersections. We calculated all three types, however—non-buffered, 10m-buffer, and 15m-buffer—to be consistent with our other intersection counts.

5. GIS Approach
From the intersection node layer, all culs-de-sac were selected (valence of 1) and a new cul-de-sac layer created. For each study site, the number of culs-de-sac and the number of total

intersections were summarized and the ratio of intersections to intersections plus culs-de-sac was calculated.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 8.3**.

To follow this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Intersection Node Layer** (point)
3. **Total Intersection Counts** from protocol for calculating **Intersections per Unit Area** (these have been buffered)

Create culs-de-sac layer:
1. Right-click on your **Intersection Layer** and select **Properties**.
   a. Click the tab **Definition Query**.
   b. Click the button **Query Builder**.
   c. Under **Fields**, double-click **valence**.
   d. Click or type **=**. Under **Unique Values**, double-click **1**.
   e. Click **OK**. The **Query Builder** window will close.
   f. Click **OK**. The **Layer Properties** window will close.
   g. Now, only those intersections created by one road segment (culs-de-sac) will be displayed.
2. Right-click on your **Intersection Layer** and select **Data**.
   a. Select **Export Data**.
   b. In the **Export Data** window under **Export**, select **All features**.
   c. Check **Use the same Coordinate System as this layer's source data**.
   d. Create a name for your new culs-de-sac layer and click **OK**.
   e. The new **Culs-de-sac Layer** will open separately.
3. Right-click on your **Intersection Layer** and select **Properties**.
   a. Click on the tab for **Definition Query**.
   b. Click on the button for **Query Builder**.
   c. Your previous entry of **valence = 1** should still be shown. Change **= 1** to **> 2**.
   d. Click **OK**. The **Query Builder** window will close.
   e. Click **OK**. The **Layer Properties** window will close.
   f. Now, the **Intersection Layer** will once again only show intersections that are created by three or more road segments.

Summarize cul-de-sac counts for multiple non-overlapping sites:
1. Using a **Spatial Join**, assign the IDs from the sites to each cul-de-sac point.
2. Open the **Attribute Table** of the cul-de-sac layer and select the field that contains the IDs.
3. Right-click on the field name and select **Summarize**.
4. Select or create and output file name and click **OK**.
5. The output .dbf file will contain the total number of culs-de-sac in each site.

Summarize cul-de-sac counts for multiple overlapping sites:
1. Using your site polygon layer, select (highlight) the first study site.
2. Perform a **Select by Location**.
   a. Under **I want to**, choose **select features from**.
   b. Under **the following layers**, choose your cul-de-sac layer
   c. Under **that**, choose **have their center in**.
   d. Under **the features in this layer**, choose your site polygon layer
   e. Make sure the box for **Use selected features** is checked
   f. Click **Apply**.
3. Open the **Attribute Table** of the **Cul-de-sac Layer**.
4. Text at the bottom of the window should show you the number of culs-de-sac in the study site you currently have selected.
5. Record this number in the spreadsheet.

6. Repeat steps 1-5 for each site.

Calculate Connected Note Ratios for each Site:
1. In a spreadsheet program such as **Microsoft Excel**, open the file containing the summary of **total intersection counts** for each site created in protocol for **Intersections per Unit Area**.
2. Open the file containing the counts of culs-de-sac in each site.
3. Add the total intersection counts to the counts of culs-de-sac in each site.
4. Divide the [number of intersections] by the [number of intersections plus culs-de-sac] (from step 3) to get the connected node ratio for each site.

6.11 Median Perimeter of Block

1. Basic Concept
This is a measure of the road length per block, and captures the issue of curvilinear street patterns that might have smaller block areas, but circuitous routes.

Future drafts of these protocols will examine how other researchers have used this measure (Song 2003).

2. Basic Formula
Median Census Block Perimeter = Median perimeter for census blocks with a centroid falling within the measurement geography.

3. Detailed Formula
Same as above.

4. Comments and Explanations
None.

5. GIS Approach
The GIS approach is to select all census blocks that have their centers in the measurement geography. Next, use XTools to calculate the perimeter of the identified census blocks. Perform a spatial intersect between the identified census blocks and the measurement geography. Import the Attribute Table of the resulting layer into Microsoft Access and analyze it using the PivotTable feature. Finally, copy the desired data out to an Excel spreadsheet and calculate the median perimeter.

6. GIS Steps
The following procedures were completed using ArcMap (ArcInfo) 8.3.

To complete this protocol you will need:
1. **Study-Site Layer** (polygon)
2. **Census Block Layer** (polygon)
3. **XTools** extension (available from the Downloads section of the ESRI Web site)
4. **Microsoft Access**
5. **Microsoft Excel**

Calculate Census Block Areas:
1. From the **XTools** menu, select **Table Operations** and then **Calculate Area, Perimeter, Length, Acres, and Hectares** (follow instructions included with the extension for **XTools** installation).
    c. Under **Select layer to measure**, select your new **Census Block Layer**.
    d. Under **Desired output units**, choose your units (depending on the census block size, try **Miles**, **Meters** or **Feet**, for example).

---

     e.   Uncheck all boxes except **Perimeter**.
     f.   Click **OK**.
2. A new field called **Perimeter** will be created in the **Attribute Table** of your new **Census Block Layer** with the perimeter of each polygon calculated in the units you chose.

Create centroid layer:
1. From the **XTools** dropdown menu, select **Feature Conversions** and choose **Shapes to Centroids**.
2. Under **Select the layer you wish to convert**, select your **Census Block Layer**.
3. Give the layer a name and click on **OK**.
4. The new layer will be added to your **Map Document**.

Select census blocks:
1. Add the **Census Block Centroid Layer** and the **Study-Site Layer** to your **Map Document**.
2. From the **Selection** dropdown menu, choose **Select by Location**..
3. Under **I want to**, choose **select features from**.
4. Under **the following layers**, select your centroid layer.
5. Under **that**, choose **have their center in**.
6. Under the **features in this layer**, select your **Study-Site Layer**.
7. Click **Apply**. Now the centroids that fall inside your study sites will be selected.
8. From the **Selection** dropdown menu, choose **Select by Location**.
9. Under **I want to**, choose **Select features from**.
10. Under **the following layers**, select your **Census Block Centroid Layer**.
11. Under **that**, choose **are completely within**.
12. Under **the features in this layer**, select your **Study-Site Layer**.
13. Click **Apply**. Now the census blocks centroids that are inside the study sites will be selected.
14. Right-click on your **Census Block Centroid Layer** and select **Data** and then **Export Data**.
15. After **Export**, choose **selected features**.
16. Choose **Use the same Coordinate System as this layer's source data**.
17. Give the new layer a name and click **OK**.
18. When prompted, add the new layer to the **Map Document**.

Open the GeoProcessing Wizard:
1. Choose **Intersect two layers** and click **Next**.
2. Select the exported subset of census block centroids from the previous step as the input layer.
3. Select your **Study-Site Layer** as the overlay layer.
4. Choose a name and location for the output layer.
5. Click **Finish**.
6. The new layer will be added to the **Map Document**.

Import data into Microsoft Access:
1. Open the **Attribute Table** of the layer created in the previous intersect step
2. Choose **Export** and select a location and name for the output DBF. *(Note: make the filename 8 characters or less so it can be imported into an Access database.)*
3. Open **Microsoft Access** and create a new blank database.
4. Go to **File….Get External Data….Import**.
5. Find the DBF table you exported from **ArcMap** above and import it.
6. Open the newly imported table.
7. From the **View**, menu, select **PivotTable View**.
8. In the **Field List** window, select the name of the unique study site ID and drag it onto the large window **Drop Column Fields Here**.
9. From the **Field List**, drag the field containing the **Perimeter** onto the large window **Drop Totals or Detail Fields Here**.

Copy data into Microsoft Excel and calculate median:
1. Open a new blank **Microsoft Excel** spreadsheet.

---

2. In **Microsoft Access**, click and drag the column headings to select them, and then CTRL-C to copy them
3. Switch to Excel and do CTRL-V to paste them. *(Note: if you have more than 255 study site IDs, you will need to paste into multiple worksheet tabs, as Excel worksheets are limited to 255 columns.)*
4. Find the row beneath the last row to contain any values (each column will have a different number of values depending on how many census block centroids were found in each measurement geography)
5. Calculate the median of all the cells above containing perimeter values

The following variables are not proposed to be used in the study, so only a brief description and detailed formula are included.

6.12 Link-Node Ratio

This is the ratio of segments to nodes (intersections or the ends of culs-de-sac, i.e., valences apart from 2). A perfect grid has a ratio of 2.5 (Dill 2003, 4). Given the character of the data in the Twin Cities Walking Study, this was very complicated to calculate.

Link-Node Ratio = Number of segments divided by the number of nodes in a measurement geography where a node is an intersection or an end of a cul-de-sac (i.e., valence 1 or 3 and above).

CHAPTER 7. OTHER BUILT-ENVIRONMENT RELATED/SPATIAL VARIABLES

Initial outline by Ann Forsyth. Steps by Joel Koepp, Jason Zimmerman, and David Van Riper. In version 3.0. In 3.1 we added significant work address information. In version 4.1 we updated some protocols to ArcGIS 9.2.

Overview

The initial part of this manual focused on measuring the built environment, with a focus on residential areas. This chapter expands those measurements to the work environment (although most of the measures are virtually identical) and to other related topics such as crime, climate and air quality, and traffic.

This section will be expanded in subsequent versions.

Work Environments

Work environments have tended to be less studied than residential ones. This has been for theoretical reasons—everyone has a home environment but not all have a work one; the residential environment is an important environment in the parts of one's day with discretionary time. It is also for practical reasons—it is easy to get census data about where people reside but not where they work.

In the Twin Cities Walking Study we collected survey data including work addresses. It should be noted, however, that not everyone goes to (paid) work. Many people are retired, homemakers, or unemployed. Others are at colleges or universities that they do not class as work. Still others work part-time or in multiple locations. Overall, studying the work environment is a complex matter.

Of the 715 participants in our study, we were able to establish valid work addresses for 345 individuals. Excluded from this set were those working from home, unemployed, retired, etc., as well as those for whom a work address was incomplete or was not able to be accurately mapped. For a small number of participants, for whom two interviews were required to obtain all the necessary data, work status changed between the first and subsequent interviews. In order to simplify the process and be consistent, we only used the results of the first interview and ignored subsequent ones. In some cases, individuals were unemployed in the first interview and employed in the second; they were not included in our work address analysis. Others changed jobs between the first and second interviews; only the data for the first job was used.

For this study, we examined the following measurements with respect to the work address of those participants working outside of their homes:

- Population per Unit Land Area (without water) – Section 3.2
- Population per Developed Land Area – Section 3.3
- Housing Units per Unit Land Area (gross, census data) – Section 3.7

- Distance to Nearest Transit Stop – Section 4.10
- Transit Stop Density – Section 4.11
- Percentage of Total Parcel Area in Major Land Uses – Section 5.1
- Entropy Index – Section 5.9
- Herfindahl-Hirschman Index, HHI – Section 5.10
- Median Census Block Area – Section 6.2
- Ration of Area within X Street Distance to Area within X Distance Radius – Section 6.3
- Number of Access Points – Section 6.4
- Road Length per Unit Area – Section 6.5
- Intersections per Unit Area – Section 6.6
- Ratio of 4-way Intersections to All Intersections – Section 6.7
- 4-way Intersections Per Unit Land Area – Section 6.8
- Ratio of 3-way Intersections to All Intersections – Section 6.9
- Connected Node Ratio – Section 6.10
- Median Perimeter of Block – Section 6.11

Our GIS steps were the same as those performed for home addresses with a few minor exceptions. Because the extent of the work addresses for these 345 participants expanded far beyond our original study area, in a few cases we were forced to use more generalized data in place of the detailed data we had obtained for Ramsey and Anoka County. Specifically, instead of using the Ramsey County water layer in conjunction with the Met Council's water layer, we used only the Met Council's water layer (as it covered a larger area). Similarly, in place of the detailed land use information we had developed for Ramsey and Anoka County, we were forced to rely on more generalized Met Council land use data for our work address analysis.

Crime

Crime affects where, when, and how much people walk. There are good theoretical reasons to consider that perceptions of crime are more important than the reality of crime, however. The Twin Cities Walking Study asked participants to respond to a number of statements about crime, taken from the Neighborhood Quality of Life Survey (NQLS):

1. My neighborhood streets are well lit at night.
2. Walkers and bikers on the streets in my neighborhood can be easily seen by people in their homes.
3. I see and speak to other people when I am walking in my neighborhood.
4. There is a high crime rate in my neighborhood.
5. The crime rate in my neighborhood makes it unsafe to go on walks during the day.
6. The crime rate in my neighborhood makes it unsafe to go on walks at night.

As local crime statistics were very unevenly available in the study area, we used these perception variables.

Climate and Air Quality

Climate and air quality both vary with a region and have some differences within a region. The Twin Cities Walking Study only measured individuals in the warmer months of April to November.

Climate data are available from central sources and we will use these data in our analyses.

Air quality is very complex to measure and the variation is not very different in the Twin Cities.

Traffic

Traffic affects where, when, and how much people walk. The Twin Cities Walking Study asked a number of questions about perceptions of traffic:

1. There is so much traffic along the street I live on that it makes it difficult or unpleasant to walk in my neighborhood.
2. There is so much traffic along nearby streets that it makes it difficult or unpleasant to walk in my neighborhood.
3. The speed of traffic on the street I live on is usually slow (30 mph or less).
4. The speed of traffic on most nearby streets is usually slow (30 mph or less).
5. Most drivers exceed the posted speed limits while driving in my neighborhood.
6. There are crosswalks and pedestrian signals to help walkers cross busy streets in my neighborhood.
7. The crosswalks in my neighborhood help walkers feel safe crossing busy streets.
8. When walking in my neighborhood there are a lot of exhaust fumes (such as from cars, buses).

While significant data are available for traffic, each has a number of problems of coverage or relevance. We investigated the following potential measures:

- Traffic counts. These were only available for major roads that were spaced quite widely. Many of the smaller measurement geographies did not overlap with any of these roads.

- Posted traffic speed. This was available for all roads in the study area. There was a huge range of roads with a posted speed of 30 miles per hour, however, including some of the most and least busy. As a measurement it did not seem useful.

- Functional classification. This was available for all roads in the study area. It more accurately reflects the "busy-ness" of major streets.

7.1 Proportion of High Functional Classifications
Note: this has been updated in version 4.1.

1. Basic Concept
Functional classifications of roads divide them into a hierarchy. In the Metropolitan Council's dataset prepared by The Lawrence Group the classifications are modified versions of the US Census Bureau's classification in TIGER Line Files including many subclasses of the following road types (US Census 2005, 3-28):

- Primary Highway With Limited Access—mostly interstates
- Primary Road Without Limited Access—mostly US and state highways
- Secondary and Connecting Road—mostly state and some county highways
- Local, Neighborhood, and Rural Road
- Vehicular Trails

This measure assesses the ratio of higher road classes to local and neighborhood roads. It would also be possible to examine the density of higher classifications.

2. Basic Formula
Proportion of High Functional Classifications = Length of primary, secondary, and connecting roads to length of all roads.

3. Detailed Formula
Proportion of High Functional Classifications = Length of primary, secondary, and connecting roads to length of all roads excluding vehicular trails, and gated roads.

4. Comments and Explanations
We converted the following local road classifications to classifications of high functional classes:

A10   213   Primary highway with limited access
A20   122   Primary road without limited access
A25   388   Primary road without limited access
        Primary road without limited access, US highways, separated
A30   2144   Secondary and connecting road
A63   348   Road with special characteristics
        Access ramp, the portion of a road that forms a cloverleaf or limited access interchange
A64   34   Road with special characteristics
        Service drive, the road or portion of a road that provides access to businesses, facilities, and rest areas along a limited-access highway

Other roads:
A40   13651   Local, neighborhood, and rural road
A71   47   Road as other thoroughfare
        Walkway or trail for pedestrians, usually unnamed
A73   3   Road as other thoroughfare
        Alley, road for service vehicles, usually unnamed, located at the rear of buildings and property

Excluded:
A66   114   Road with special characteristics
        Gated barrier to travel
A98   5   * not listed *
        This is the University's bus transitway, which has bike but not pedestrian access.

5. GIS Approach
Using the F_Class field from the road layers that had already been intersected with our buffers in previous protocols, each road segment was reclassified based on the groupings listed above (high, other, and excluded). For each buffer, the lengths of the high functional class roads were summed and divided by the sum of the lengths of all roads (not including those in the "excluded" category) to obtain the proportion of high functional class roads to all roads.

6. GIS Steps
The following procedures were completed using **ArcMap (ArcInfo) 9.2**.

To complete this protocol you will need:
4. **Road Centerline Layer** (polyline)

---

5. **Study-Site Layer** (polygon)

The basic steps are:
- Reclassify Road Segments
- Intersect Road Segments with Study-Site Layer
- Summarize Road Segment Lengths
- Calculate Proportion of High Functional Classifications

Reclassify Road Segments:
2. Right-click the **Road Centerline Layer** and select **Open Attribute Table…**
3. Click the **Option** button and select **Add Field…**
4. Name the field **CLASS** and change the type to **Short Integer**.
5. Click the **Option** button and select **Select by Attributes**.
6. Construct an SQL query to select all road segments that can be classified as "**High Functional**".
   a. Click **OK** to execute the query.
7. Right-click the **CLASS** attribute and select **Field Calculator…**
   a. Set the selected records equal to **1**.
8. Click the **Option** button and **Select by Attributes**.
9. Construct an SQL query to select all road segments that can be classified as "**Other Roads**".
   a. Set the selected records equal to **2**.
10. Click the **Option** button and **Select by Attributes**.
11. Construct an SQL query to select all road segments that can be classifies as "**Excluded**".
   a. Set the selected records equal to **3**.

Intersect Road Centerlines and Study-Sites and calculate road lengths:
1. Click on the **Add Data** icon and load the **Road Centerline Layer** and the **Study-Site Layer**.
2. Open **ArcToolbox** by clicking on the icon.
3. Navigate to the **Intersect** tool (**Analysis Tools** -> **Overlay**) and open the tool by double-clicking on it.
4. Click on the down-arrow by the **Input Features** box and select the **Study-Site Layer**.
5. Again, click on the down-arrow by the **Input Features** box and select the **Road Centerline Layer**.
6. Specify a name and location for the output layer.
7. Set the **JoinAttributes** option to **ALL**.
8. Accept the **XY Tolerance** default.
9. Set the **Output Type** to **LINE**. This will generate a line layer containing the intersections of the study-sites with the road centerlines.
10. Click **OK**. A new output layer containing only road segments intersecting with study-sites will be added to the Map Document. If a road segment falls in overlapping study-sites, the output layer will contain two records, one for each study-site, for the segment.
11. Recalculate the length of the road segments in the newly intersected layer.
   a. Right-click on the new layer and select **Open Attribute Table**.
   b. Click on the **Option** button and select **Add Field…**
   c. Name the field **Length** and change the Type to **Double**.
       i. If a field named **Length** already exists, change the name in the **Add Field…** dialog box to something else.
   d. Click **OK**.
   e. Right-click the **Length** field in the attribute layer and select **Calculate Geometry…**
   f. Set the **Property:** to **Length**.
   g. Set the **Coordinate System** to "**Use coordinate system of the data source**."
   h. Set the units to the desired type.
   i. Click **OK**. The lengths of the road segments will be calculated.

Summarize Intersected Road Segment Lengths

---

1. Right-click on the layer containing the intersected road segments and select **Open Attribute Table...**
2. Click the **Option** button and select **Select by Attributes**.
3. Construct an SQL query to select the "High Functional" roads.
   a. Click the **Class** field, then click **=**, and then type **1**. Click **OK**.
4. Right-click on the study-site ID field and select **Summarize**.
5. For **Select a Field to Summarize** select your ID field.
6. Under **Choose One or More Summary Statistic...** expand the **Length** field and check the box for **Sum.**
7. Specify a name and location for the output table.
8. Check the box that is marked **Summarize on the Selected Records Only**.
9. Click **OK**.
10. Click the **Option** button and choose **Select by Attributes**.
11. Construct an SQL query to select all roads except those classified as "Excluded".
    a. Click the **Class** field, then click **<>**, and then type **3**. Click **OK**.
12. Right-click on the study-site ID field and select **Summarize**.
13. For **Select a Field to Summarize** select your ID field.
14. Under **Choose One or More Summary Statistic...**' expand the **Length** field and check the box for **Sum**.
15. Specify a name and location for the output table.
16. Check the box that is marked **Summarize on the Selected Records Only**.
17. Click **OK**.

Calculate Proportion of High Functional Classifications
1. For each study-site, you now have two tables of road length – one with the sums of the lengths of High Functional roads and one with the sums of the lengths of all roads (except those that may be excluded from the analysis).
2. Right-click on the study-site layer and select **Open Attribute Table...**
3. Click the **Options** button and select **Add Field...**
4. Name the field **Pro_High** and set the type to **Double**. Click **OK**.
5. Join the two tables of road length to the study-site layer.
   a. Right-click on the study-site layer and select **Joins and Relates** -> **Join...**
   b. Under **What do you want to join to this layer?**, click on the down-arrow and select **Join attributes from a table**.
   c. Select the study-site ID for **1. Choose the field in this layer that the join will be based on**:
   d. Under **2. Choose the table to join to this layer, or load the table from disk**, click the down-arrow and select the table containing the road length summaries for the **High Functional** class.
   e. Under **3. Choose the field in the table to base the join on:**, click the down-arrow and select the study-site ID.
   f. Click **OK**.
   g. Right-click on the study-site layer and select **Joins and Relates** -> **Join...**
   h. Under **What do you want to join to this layer?**, click on the down-arrow and select **Join attributes from a table**.
   i. Select the study-site ID for **1. Choose the field in this layer that the join will be based on**:
   j. Under **2. Choose the table to join to this layer, or load the table from disk**, click the down-arrow and select the table containing the road length summaries for the **All Roads except Excluded** class.
   k. Under **3. Choose the field in the table to base the join on:**, click the down-arrow and select the study-site ID.
   l. Click **OK**.
6. Right-click on the **Pro_High** attribute and select **Field Calculator...**

7. Click the attribute containing the sum of High Functional road lengths, then click the / button, and then click the attribute containing the sum of all roads except "Excluded". Click **OK** to execute the calculation.
8. This calculation will yield the proportion of High Functional roads to all roads (except those excluded) in a study-site.

CHAPTER 8. DATA STORAGE AND SHARING

By Jason Zimmerman, Joel Koepp, and Ann Forsyth

Note: This was new in version 2.2.

Steps/problems with obtaining our data

MetroGIS data server
In the Twin Cities, MetroGIS—an organization that coordinates data sharing among a number of local governments—maintains a data server that allows access to a number of metro-wide data sets free of charge (municipal boundaries, water bodies, and land-use data sets are among those provided). These data have been aligned and cleaned as much as possible to provide consistency across municipalities. Much of our base data was obtained from the DataFinder Web site (http://www.datafinder.org/).

User license agreements (Regional parcels)
In order to access the regional parcel data set (an item on the MetroGIS DataFinder), it was necessary to enter into a separate license agreement. At the time the data was requested, this involved two license agreements: one that covered six of the seven metro counties and another that covered the seventh. This second agreement proved to be a much harder agreement to reach, and involved exchanges of agreement language between University attorneys and attorneys for the county. By the time this hurdle was crossed, however, a new obstacle emerged. At the end of the approaching calendar year the contract for data sharing between MetroGIS and the seven counties expired and was in the process of being renegotiated. This ended up taking many months to finalize, greatly extending the time before we were able to finally obtain access to the regional parcel data set.

Luckily, while all of this was going on, we were able to rely on our previous relationships with two of the counties to use their parcel data free of charge. We did, however, need to purchase small areas of parcel data from two additional counties where the largest buffers of our study area spilled over across county boundaries.

I-35W Coalition agreement/Access to Ramsey County/St. Paul data
A memorandum of understanding with the I-35W Coalition (a coalition of seven cities within two counties) allowed us to access a large amount of high-quality data for free (a lump sum was paid to the Coalition to cover time for the GIS Coordinator to assist with data problems, but there was no charge to access the data itself).

This arrangement had additional benefits when we began to look for data outside the boundaries of the Coalition. The Coalition's membership in the Ramsey County Users Group allowed us to access data for the whole of Ramsey County, not just for the Ramsey portion of the coalition cities. It would have been possible just to join the Ramsey County Users Group, however, and in the end this would likely have been less expensive. This highlights some of the challenges in deciding which of many complex arrangements one must decide on in accessing data.

Data for the City of St. Paul (which lies within Ramsey County but is not a member of the Coalition) was also easier to obtain because of this relationship, although we spent some time in negotiations.

Problems that arose in obtaining data from the City of St. Paul were mostly due to the dispersed nature of the data. Data for street trees needed to be obtained directly from the Parks Department. Data for sidewalks and streetlights needed to be obtained from Public Works. Difficulty in coordinating messages, confirming approvals, and exchanging data were the biggest obstacles in obtaining information from St. Paul.

Storing/Sharing data

Central server
GIS data, documentation, and tracking spreadsheets are all stored on a single file server. With a small project team, issues of file locking are infrequent and easily handled. All data are backed up to a separate dedicated server on an incremental basis nightly, with a fresh full backup occurring once a week.

Working files kept locally
Many measurements require several large processing operations on data before it is ready for final calculations. To aid the efficiency of computationally intensive GIS tasks, working data files are kept on local machines rather than on a network. In the process of calculating median block size for 1600m straight-line buffers of the 715 participants with good addresses, for example, one of the files generated contained over 2.5 million features and consumed nearly a gigabyte of disk space. This intermediate file then had to be broken into smaller pieces so ArcGIS would be able to process them. Even working off local hard drives these processes are intensive, and it would take much longer even on the fastest of local networks if the files were saved on a server due to network latency.

"Final" files on server
Important base data and final versions of modified files are stored on a central file server for access by multiple users. Rather than storing shapefiles and dealing with the resulting clutter of having dozens of layers in separate folders, we used the geodatabase storage format, introduced to ArcGIS in version 8. The geodatabase can store multiple layers and tables together in one large file (which is actually a modified Microsoft Access database). We created several geodatabases named by topic, e.g., Street Features, Density, Buffers, etc. Among the benefits of using the geodatabase format is that line and polygon features automatically have their length and area fields recalculated, when modified. It is important to note that the file size of a geodatabase cannot exceed 2 GB, so when working with very large datasets, multiple geodatabases may be needed.

Issues with multiple users
While the geodatabase is convenient for data storage, one limitation is that the basic ArcGIS or ArcINFO install does not support some of the advanced features of the format. With an installation of the ESRI product ArcSDE (designed for file servers that share and manage GIS data), it is possible for multiple users to edit a geodatabase layer simultaneously. The other capability afforded by ArcSDE is versioning, which would help keep track of changes made to a layer throughout the stages of geoprocessing performed on base data.

Our college did not have a license for ArcSDE.

Tracking changes/updates

Work was tracked in a number of ways:

Shapefile tracking sheet
This Excel spreadsheet is to keep track of which layers get put in the geodatabases on the central server. There are fields for the name of the layer, the geodatabase in which it is found, the date created, the date last modified, the features in the layer, and miscellaneous notes. It is updated each time a layer is added or replaced. This is a way of avoiding having to purchase ArcSDE.

Decision logs
Logs were kept by both GIS Research Fellows throughout the study. These documents, updated at least weekly, record key decisions that were made regarding GIS, conversations that were held, and tasks that were performed. These logs carry out two functions: they allow the Principal Investigator to be kept in the loop and they act as a record of GIS-related information over time that can be reviewed when questions arise at a later date.

Updating/tracking measurement results
A master measurement spreadsheet with tabs for each variable stores variables.

Checklist
A checklist of variables was maintained, with numbering from protocols, required geographies in columns, and filled in with name and date when completed. A version of this checklist is in Appendix 1D.

Other issues

Projections

The standard coordinate system for MetroGIS datasets is Universal Transverse Mercator (UTM) North American Datum 1983 (NAD83), Zone 15. Any working or final files created from MetroGIS data was kept in the same coordinates. Some data layers that came from individual cities or from Ramsey County—such as streetlights, street trees, and sidewalks—came in the Ramsey-Hennepin coordinate system. Because ArcGIS has the capability to correctly display data from multiple projections in the same view, these layers remained in Ramsey-Hennepin coordinates.

Much of the street feature data was hand-digitized using high-resolution color aerial photos of Ramsey County. Imagery cannot have coordinate system information attached to it like shapefiles or geodatabase feature classes, so the data view must be set to the known coordinates of the imagery. As a result, when the hand-digitized features were saved from graphics to a shapefile, they were saved in Ramsey-Hennepin coordinates.

Converting Measures: Blocks and Minutes to Meters

A number of survey measures asked participants to estimate distances in blocks. We investigated how long to make the block. Susan Liss from the National Household Transportation Survey indicated 1/9 of a mile as appropriate. Other experts, however,

including Kevin Krizek, used 1/10 of a mile as a rule of thumb. We estimated the median block length in the area as closer to 1/10 of a mile or 160 meters and used that measure.

APPENDICES

Note: if there is a specific author for each section he, she, or it is named at the beginning of the section. Appendices 1D and 1F are new in version 2.3.

Appendix 1A: I-35W Corridor Coalition Land-Use Metadata

North Metro I-35W Corridor Coalition
Land-Use Designations
(Definitions Adopted as of 1/15/98)
(Updated 1/19/99)

Following are the land-use designations, definitions, notes, and samples developed and adopted by the North Metro I-35W Corridor Coalition.*

*1. Land-use designations, unless otherwise specified, refer to the predominant identifiable land use on a parcel or portion thereof.

2. The sample uses listed after each category indicate the types of land uses that fit into each land-use category assuming they meet the criteria outlined in the definition. Some land uses may fit into more than one category (for example, gas stations may be categorized as Neighborhood Commercial, Community Commercial, or Highway/Convenience Commercial depending on there size, location, etc.). Sample land uses are simply intended to give the user an idea of the <u>typical</u> land uses within each category. They are not intended to be the only criteria for categorizing land use.

3. The land-use categories, definitions, and samples outlined below were developed to serve areas ranging from first-ring suburbs to outer-ring suburbs. These categories and definitions may not be sufficient for rural or core urban areas.

AGRICULTURAL (AG)

Definition: Land used for agricultural purposes, including farming, dairying, pasturage, horticulture, floriculture, viticulture, and animal and poultry husbandry and accessory uses; provided that such accessory uses shall be incidental to the agricultural activities.*

*1. Agricultural properties with a residential density of less than one unit per twenty acres may be considered Residential if a majority of the land is not used for agricultural production (Hobby Farms).

2. Some traditional Agricultural uses may not be considered Agricultural if they conform more closely with definitions and samples for commercial or industrial land uses.

Samples: Agricultural land uses would typically include any of the following that meet the previously defined criteria:

> Cash Grain or General Farm
> Other Agricultural Use
> Sod Farms

RESIDENTIAL (RES)

This section contains a general definition of residential land use followed by several subset definitions. It is intended that land uses falling within a subset definition meet the criteria outlined in the general definitions plus the additional criteria outlined in the subset definition.*

General Definition: Land containing a building or portion thereof used exclusively for residential purposes, including one-family, two-family, and multiple-family dwellings, but not including hotels, motels, and boarding and lodging houses, nursing homes, or elderly care facilities; or open space within or related to a residential development, not in individually owned lots or dedicated for public use, but which is designed and intended for the common use or enjoyment of the residents of the development.

* Residential properties with a density of less than one unit per twenty acres may be designated as Agricultural if a majority of the land is used for agricultural production (Hobby Farms).

i.) Rural (R)

Definition: Land meeting the general definition of residential and that is outside the MUSA boundary containing a residential building with not more than one dwelling unit entirely surrounded by open space on the same lot.

ii.) Single-Family, Detached (SFD)

Definition: Land meeting the general definition of residential and containing a residential building with not more than one dwelling unit entirely surrounded by open space on the same lot.

> Single-Family Dwelling, Platted Lot
> Mobile Home on Real Estate

iii.) Single-Family, Attached (SFA)

Definition: Land meeting the general definition of residential and containing one or more buildings containing two or more dwelling units, each of which has primary ground floor access to the outside, and which are attached to each other by party walls without openings.

> Two-Family Dwelling, Platted Lot
> Conversion
> Three-Family Dwelling, Platted Lot
> Double Dwelling
> Townhome-Inner
> Townhome-Outer

Townhome Non-Tax Outlotiv.) Multifamily (MF)

Definition: Land meeting the general definition of residential and containing one or more buildings containing three or more dwelling units, one or more of which does not have primary ground floor access to the outside and which are attached to each other by party walls without openings.

> Apartment, 1-9 Rental Units
> Apartment, 10-19 Rental Units
> Apartments, 20-49 Rental Units
> Apartment, 50 or More Unites
> Apartment Misc. Improvements
> Condominium
> Condominium Non-Tax Lot

v.) Manufactured Housing Park (MHP)

Definition: Land meeting the general definition of residential and under single ownership that has been planned and improved for the placement of manufactured housing for dwelling purposes.

Trailer or Mobile Home Park

COMMERCIAL/RETAIL

i.) Neighborhood (NC)

Definition: Land use primarily engaged in the provision of goods or services with a primary service radius of 1 to 1.5 miles. Neighborhood Commercial/Retail Land Uses would typically be located in a predominantly residential area and have a Gross Leasable Area (GLA) of approximately 25,000 square feet or less.

Samples: Neighborhood Commercial/Retail land uses would typically include any of the following that meet the previously defined criteria.

Small (Under 10,000 sq. ft.) Detached Retail
Neighborhood Shopping Centers
Restaurant, Cafeteria, and/or Bar
Dry cleaning Plants and Laundries
Automotive Service Station
Barber/Beauty Shops

ii.) Community (CC)

Definition: Land use principally engaged in the provision of goods or services with a primary service radius of 3 to 5 miles. Community Commercial/Retail Land Uses would typically be located in a small commercial/retail center with a Gross Leasable Area (GLA) of approximately 50,000 to 150,000 square feet.

Samples: Community Commercial/Retail land uses would typically include any of the following that meet the previously defined criteria.

>Supermarkets
>Discount Stores & Jr. Dept. Stores
>Community Shopping Center
>Other Retail Structures
>Funeral Homes
>Full Service Banks
>Savings & Loans
>Auto Car Sales and Services
>Bowling Alleys
>Restaurant, Cafeteria, and/or Bar
>Automotive Service Station
>Other Commercial Structures
>Car Wash
>Day Care Centers

iii.) Regional (RC)

Definition: Land use principally engaged in the provision of goods or services with a primary service radius of 10 or more miles. Regional Commercial/Retail Land Uses would typically be located in or near a large commercial/retail center with a Gross Leasable Area (GLA) of approximately 350,000 square feet or more.

Samples: Regional Commercial/Retail land uses would typically include any of the following that meet the previously defined criteria.

>Nurseries
>Full Line Department Stores
>Regional Shopping Center
>Parking Garage Structure & Lots
>Parking Ramps
>Theaters
>Drive-in Theaters
>Lodge Halls and Amusement Parks
>Marine Service Facilities
>Sport/Public Assembly Facility

iv.) Highway/Convenience (HC)

Definition: Land use principally engaged in the provision of goods or services that primarily services customers attracted from a nearby major transportation arterial.

Samples: Highway/Convenience Commercial/Retail land uses would typically include any of the following that meet the previously defined criteria.

>Motels and Tourist Cabins
>Hotels
>Drive-in Rest/Food Service Facility
>Gas Station
>Automotive Service Station

OFFICE (OFC)

Definition: Land use predominantly involved in administrative, professional, or clerical services.

Samples: Office land uses would typically include any of the following that meet the previously defined criteria.

> Office Building, 1-2 Stories
> Office Building, 3 or More Stories Walkup
> Office Building, 3 or More Stories Elevator
> Veterinary Clinic
> Medical Clinics and Offices

MIXED USE (MU)

Definition: Land containing a building with significant amounts of residential uses in combination with commercial and/or office uses.[*]

*Land is identified as Mixed Use only if it contains buildings with stacked or multiple uses in which there is a significant amount of residential use. Properties with only a small amount of residential use (such as caretakers quarters, etc.) are assigned the predominant land use.

Samples: Mixed-use land would typically include any of the following that meet the previously defined criteria.

> Residential apartments over commercial or office uses

INDUSTRIAL

i.) Light (LI)

Definition: Land used primarily in the manufacture, processing, fabrication, assembly, packaging, incidental storage, sales, and distribution of predominantly previously prepared materials, finished products or parts. Light Industrial land uses would typically have all processing within buildings, require limited exterior storage, generate limited amounts of truck or rail traffic, and be free of hazardous or objectionable elements such as noise, odor, dust smoke, glare, or other pollutants.

Samples: Light Industrial land uses would typically include any of the following that meet the previously defined criteria.

> Greenhouses, Vegetables & Floriculture
> Food & Drink Process Plants and Storage
> Manufacturing & Assembly, Light
> Small Shops (Machine, Tool & Die, etc.)
> Other Industrial Structures
> Flex Industrial Center

Mini Warehouse
Research and Development Facility

ii.) Heavy (HI)

Definition: Land used primarily in the manufacture and/or processing of products from large bulky, predominantly raw, extracted, or hazardous materials; or use engaged in the storage of flammable, explosive, or other materials that may pose a threat to public health or safety. Heavy Industrial land uses may require exterior storage of large equipment or material, be engaged in outside processing or assembly, generate significant amounts of truck and/or rail traffic, or emit limited amounts of objectionable elements such as noise, odor, dust, smoke, glare, or other pollutants.

Samples: Heavy Industrial land uses would typically include any of the following that meet the previously defined criteria.

Foundries & Heavy Manufacturing Plants
Grain Elevators
Commercial Warehouses
Commercial Truck Terminals

INSTITUTIONAL (IN)

Definition: Land used for primarily religious, governmental, educational, social or health care facilities excluding clinics.*

* Symbols may be used to differentiate between different types of institutional uses.
Samples: Institutional land uses would typically include any of the following that meet the previously defined criteria.

Nursing Homes & Private Hospitals
Exempt Property Owned by USA
Exempt Property Owned by State of MN
Exempt Property Owned by Counties
Fire Stations
Exempt Property Owned by Board of Education
Exempt Property Owned by Private Schools
Charitable Exempt Hospital/Nursing Home
Welfare/Charitable Facility
Churches, etc., Public Worship – Private
Graveyards, Monuments, Cemeteries

SPORT/RECREATION

i.) Indoor (SRI)

Definition: Land containing a building(s) developed, used, and maintained primarily for recreational activities.

Samples: Indoor Sport/Recreation land uses would typically include any of the following that meet the previously defined criteria.

Ice/Hockey Arenas
Indoor Soccer Fields
Dome Recreational Facilities
Large-Scale Health/fitness Clubs
Indoor Swimming Pools
Sport/Public Assembly Facility

ii.) Outdoor (SRO)

Definition: Land used primarily for public recreation activities improved with playing fields, playground or exercise equipment and associated structures.

Samples: Outdoor Sport/Recreation land uses would typically include any of the following that meet the previously defined criteria.

Baseball/Softball Fields/Stadiums
Soccer/Football Fields/Stadiums
Golf courses
Golf Driving Ranges
Mini-Gold Courses
Playground Equipment

OPEN SPACE

i.) Passive (OSP)

Definition: Land used and maintained for resource protection, amenity, or buffer, which supports unorganized public recreational activities and which may contain trails, picnic areas, public fishing/boating docks, etc.

Samples: Passive Open Space uses would typically include any of the following that meet the previously defined criteria.

Exempt Cultural & Nature Exhibits

ii.) Natural (OSN)

Definition: Predominantly undeveloped or unaltered land preserved in its natural state for environmental or aesthetic purposes.

Samples: Natural Open Space uses would typically include any of the following that meet the previously defined criteria.

Wetlands
Timber or Forest Lands
Nature Preserves

UTILITY (UTL)

Definition: An area or strip of land, either public or private, occupied by a power plant or substation, electric transmission line, oil or gas pipeline, water tower, municipal well, reservoir, pumping station, water treatment facility, communications tower, or similar use.

Samples: Utility uses would typically include any of the following that meet the previously defined criteria.

> Comm Ld & Impro Own by Pub Uti Th Rail

TRANSPORTATION

i.) Vehicular/Pedestrian Right of Way (ROW)

Definition: An area or strip of land, either public or private, on which an irrevocable right-of-passage has been recorded for the use of vehicles or pedestrians or both.

Samples: Transportation Right of Way uses would typically include any of the following that meet the previously defined criteria.

> Skyways
> Street Right of Way
> Sidewalks/Paths/Trails

ii.) Railway (RLW)

Definition: An area or strip of land occupied or intended to be occupied by a railroad or other similar use.

Samples: Railway uses would typically include any of the following that meet the previously defined criteria.

> Railroad Real Property Used in Operation
> Railroad Real Property Not Used in Operation

iii.) Airport (AP)

Definition: Land used for the operation of aircraft and associated accessory uses; provided accessory uses are incidental to the airport activity.

Samples: Airport land uses would typically include any of the following that meet the previously defined criteria.

> Landing strips/runways
> Airplane Hangers
> Helipads
> Air-Traffic Control Facilities
> Boarding Terminals

VACANT (VAC)

Definition: Land not currently used for any defined purpose that may or may not contain buildings or other structures.

Samples: Vacant land uses would typically include any of the following that meet the previously defined criteria.

Agricultural, Vacant Land
Industrial, Vacant Land
Commercial, Vacant land
Apartment, Vacant Land
Residential, Vacant land, Lot
Vacant Land
Exempt HRA Vacant Land

OPEN WATER (WAT)

Definition: Permanently flooded open water not including wetlands, or periodically flooded areas.

Samples: Open Water would typically include any of the following that meet the previously defined criteria.

Lakes
Rivers
Streams (not international)
Ponds (permanently flooded)

Appendix 1B: Excensus Metadata

Twin Cities Walking Study - Employment Summary
Prepared by Excensus LLC

Overview of Methodology for St. Paul Area

I. Obtained database of current employers within study area from Dun & Bradstreet (DNB)

II. Verified coverage of DNB dataset against requested geographies

III. Cleaned and standardized DNB employer records

IV. Geocoded DNB employer records and resolved unmatched records

V. Manually reviewed points (with 5 or more employees) placed just outside of context area boundaries to determine whether they should be included

VI. Merged DNB data with previous employment data from I35W Corridor project

VII. At this stage there were two products. One assigned employer points to 100-meter grid cells, taking into account overlapping context areas. The other gave the "raw" data which was in fact what we used.

VIII. Summarized employment counts in each cell, including the requested retail breakouts

Structure of Employment Summary Shapefile (Employment_Summary_100m.shp)

Field Name	Description
Shape	Type of shapefile
ID	100-meter grid identifier
Context	Area identifier
Group	Allows context areas to be grouped into non-overlapping groups
Type	Indicates whether grid cell is in a context or a focus area
Num_Empl	Total number of employees within grid cell
Num_Retail	Number of employees in retail businesses (SIC 52xx – 59xx)
Num_GenMer	Number of employees in general merchandise businesses (SIC 53xx)
Num_Food	Number of employees in food stores (SIC 54xx)
Num_EatDri	Number of employees in eating & drinking establishments (SIC 58xx)
Num_MisRe	Number of employees in miscellaneous retail businesses (SIC 59xx)

Appendix 1C: GIS Digitizing and Fieldwork for Street Trees, Street Lamps, and Sidewalks

By Joel Koepp

This is an account of how the research team created street tree, street lamp, and parcel data layers for the 1.0*1.0-mile/1609*1609-meter context areas. Digitizing was done by two undergraduate research assistants—Jorge Salcedo and Ian Kaminski-Couglin. Data conversion was done by research fellows Joel Koepp and Jason Zimmerman. Joel Koepp did all field-checking of data. Ann Forsyth did a small amount of work to figure out a method for dense stands of trees that had continuous canopies in the digital orthophotos.

Digitizing
The color high-resolution digital orthophotos served as the basis for the digitizing of street features. The study areas were divided into groups so the undergraduate research assistants could be working in parallel to complete the digitizing.

For places where no sidewalk data could be obtained from the cities, sidewalks identifiable in the orthophotos were digitized using the basic drawing tools in ArcMap. A polyline was drawn down the middle of the sidewalk for each sidewalk segment, starting and ending where the sidewalk met the road. In the cases where a sidewalk meets a driveway and continues on the other side, the line was drawn so that it continued across the driveway rather then being broken into two separate lines. When all the sidewalks were digitized and checked, they were converted into a GIS shapefile using the XTools extension, which can convert graphics into spatial data with attributes.

Streetlights were digitized by inspecting the orthophotos and using the drawing tools to place a dot at the base of each light pole. The majority of streetlights were identifiable due to the unique shadows created on the ground. Along major arterials, particularly in the suburbs, most lights consist of a pole and an attached arm that hangs out over the road. The shadow it created was very distinctive, so it was not difficult to locate the base of the pole using this shadow. In the City of Saint Paul, residential streets use one or two styles of light fixtures in a consistent fashion throughout entire neighborhoods. They are located at regular intervals and so were relatively easy to identify and digitize. In places where a streetlight could not be distinguished from a utility pole or tree, the point was digitized and left to be resolved in field checking.

Trees were digitized in much the same way as streetlights. The drawing tools were used to place a point on each tree for all trees up to the front face of a house, and for a similar distance away from the street in non-residential areas. In places where the density of trees was too high to distinguish individual trees with enough accuracy, the cluster of trees was drawn using a polygon. These polygons were to be field checked later to assign a tree density and arrive at a count.

Field Checking
The laptop was loaded up with the GIS software (both ArcMap and ArcView) and all the relevant data: focus and context areas, streets, parcels, and the digitized sidewalks, street trees, and streetlights. The GPS receiver was also taken.

In general terms, the goal of field checking was to verify the existence of the features we had digitized in the office. Where necessary, features not found or not identifiable on the orthophotos would be added, and features mistakenly digitized (e.g., telephone pole mistaken for light pole) would be removed. In addition, fieldwork also included taking photos of tree stands where the tree density on the orthophotos was too high to distinguish individual trees.

The first context area was done in conjunction with two research assistants doing the separate urban design inventory. For this, area checking used paper maps and any changes were marked directly on them. Back at the office changes were made in GIS, based on the map notations. This checking also involved walking the streets, doing a small area on foot, then driving the car a short distance to a new area, getting out and walking again.

After one day it was clear that walking was not a viable long-term option for completing fieldwork in the time frame the research team had laid out for it. It also turned out that the paper maps strategy was cumbersome, because this involved making edits on the paper and then having to repeat them in the GIS to change the actual data. From then on, edits were made directly to the data in GIS, adding and deleting features as necessary.

To check data on a given street, the field checker drove slowly down the street and parked along the side two or three times to check what he was looking at against the data on the screen. In a few places there were streets that were either too busy to park on or where parking was not allowed. For those situations, the checker found a nearby place to park and walked with the laptop.

The first feature checked was sidewalks, as it was the feature that required the least amount of editing. Typically a checker could tell immediately by looking down the street if there were sidewalks on one or both sides, or none at all. If a sidewalk was there for which there was no data, the checker drew it in for the whole length of street by default. One or two times that sidewalk would be edited if it turned out not to extend to the end of the street. For streetlights, the normal procedure was to count the number of lights on each side of the street and check against the GIS. In suburban areas, lights were usually found only on one side of a street. On most streets in Saint Paul, the pattern is to have evenly spaced streetlights on alternating sides of the street, and then usually two lights on opposite corners at an intersection. In places where the GIS data clearly indicated that a streetlight was missing, the checker would locate it on the street and make a note of how many houses away from the corner or end of the street it was. That information allowed me to use the parcel data in the GIS as a reference for adding in the point. For trees, the number of them in a front yard was checked against what was on the screen, and in places where the GIS data was missing a significant number of trees they were digitized. Obviously without using the GPS the tree locations were not precise, but for the purposes of the eventual variable the number of trees is going to be crucial.

For keeping track of progress, the field checker checked off street segments on a paper map, as this was more efficient than doing the equivalent within GIS.

After every day of fieldwork, the checker compressed the GIS data into a zipfile and transferred it to a computer at work via FTP for backup purposes.

The field checker took all the tree photos with a digital camera. The process began by erasing all data on the camera memory card and all waypoints on the GPS. When the checker came to a stand of trees that needed to be photographed (as indicated by the "dense tree area" polygons digitized), the car was stopped, a picture taken, and then immediately the position was

recorded with a GPS waypoint. This way, each photo had a corresponding waypoint. At the end, the checker had the same number of photos and waypoints, and both were in the order in which they had been recorded so they could be easily associated with each other. The Garmin software and an ArcView extension called AVGarmin were used by the checker to retrieve the waypoints from the GPS and convert them into a shapefile. Finally, an ArcMap project was created with all the waypoints and linked in the tree images.

Appendix 1D: Matrix of Measures by Measurement Geography Used in Twin Cities Walking Study (for residential environments, valid as of version 3.0)

Measurement	STRAIGHT				NETWORK					grid	Dist. to Near	Focus/ Context
	200	400	800	1600	200 a	200 b	400	800	1600			
3.1 Raw Population per Unit Land Area										▓	▓	
3.15 Net Land Area										▓	▓	
3.2 Population per Unit Land Area (without water)										▓	▓	
3.3. Population per Developed Land Area										▓	▓	
3.4 Residential Population in Residential Parcels (Residential Density)										▓	▓	
3.5 Population plus Employment per Unit Land Area (excluding water)			▓	▓				▓	▓	▓	▓	
3.6 Employment per Unit Land Area			▓	▓				▓	▓	▓	▓	
3.7 Housing Units per Unit Land Area (gross, census data)										▓	▓	
3.8 Lot Coverage			▓	▓				▓	▓	▓	▓	
4.1 Sidewalk Length per Unit Area			▓	▓				▓	▓	▓	▓	
4.2 Sidewalk Length Divided by Road Length			▓	▓				▓	▓	▓	▓	
4.3 Sidewalks Length divided by Major road Length			▓	▓				▓	▓	▓	▓	
4.4 Streetlights per Length of Road			▓	▓				▓	▓	▓	▓	
4.5 Street Trees (trees within an X foot/meter buffer) per Length of Road			▓	▓				▓	▓	▓	▓	
4.6 Percent of Street Segments with Marked Pedestrian Crossings at One or Both Ends			▓	▓				▓	▓	▓	▓	
4.7 Percent of Street Segments with Visible Litter, Graffiti, or Dumpsters			▓	▓				▓	▓	▓	▓	

4.8 Percent of Street Segments with Traffic Calming, Broadly Defined			■	■				■	■	■			
4.9 Other fieldwork based measurements of neighborhood identification, street crossing, street characteristics, views, land uses, barriers, sidewalks, bicycle lanes, mid block crossing, steepness, sidewalk amenities, street trees, buildings, windows, other features of buildings, garages, parking, driveways, maintenance, lighting, freeways, traffic features, architecture/design.													
4.10 Distance to nearest Bus Stop	■	■	■	■	■	■	■	■	■				■
4.11 Bus Stop Density										■	■		
5.1 Percentage of Total Parcel Area in Major Land Uses (seven land uses in this case)										■	■		
5.2 Percentage of Land Area in Nighttime Uses			■	■				■	■	■			
5.3 Percentage of Land Area in Social Uses			■	■				■	■	■			
5.4 Percentage of Land Area in Retail Uses			■	■				■	■	■			
5.5. Percentage of Land Area in Industrial and Auto-Oriented Uses			■	■				■	■	■			
5.6 Dissimilarity Index, Formula 1										■	■		
5.7 Proportion of Dissimilar Land Uses Among Grid Cells in an Area (Formula 1)										■	■		
5.8 Dissimilarity Index, Formula 2										■	■		
5.9 Entropy Index				■					■	■	■		
5.10 HH Index				■					■	■	■		
5.11 Retail Employment per Acre			■	■				■	■	■			
5.12 Density of Employees in Major Retail Subcategories: General Merchandise (old SIC 53xx); Food Stores (54xx); Eating and Drinking Places (58xx); Miscellaneous Retail (59xx).	■	■	■	■	■	■	■	■	■	■			
5.13 Distance to nearest gen merchansise			■	■				■	■	■			
Food stores			■	■				■	■	■			

Eating and drinking			■	■				■	■	■		
Misc retail			■	■				■	■	■		
5.14 Distance to nearest facility (for comparison with NQLS)	■	■	■	■	■	■	■	■	■		■
5.15 Gravity measure with competition										■	
6.1 Average Census Block Area									■	■	
6.2 Median Census Block Area									■	■	
6.3 Ratio of Area within X Street Distance to Area within X Distance Radius	■	■	■	■					■	■	■
6.4 Number of Access Points									■	■	
6.5 Road Length per Unit Area									■	■	
6.6 Intersections per Unit Area									■	■	
6.7 Ratio of 4-Way Intersections to All Intersections									■	■	
6.8 Four-Way Intersections per Unit Land Area									■	■	
6.9 Ratio of Three-Way Intersections to All Intersections									■	■	
6.10 Connected Node Ratio									■	■	
6.11 Median Perimeter of Block									■	■	

Appendix 1E: Other Data and Measurement Issues

By Joel Koepp and Jason Zimmerman, with Ann Forsyth

1. Projections

For these measurements we used the UTM NAD83 Zone 15 coordinate system, which uses the meter as its fundamental unit. This is the coordinate system used by the Metropolitan Council for all of its endorsed regional datasets. Some datasets received from counties were used in their native coordinate system, such as the Ramsey County parcels, which come in the Ramsey-Hennepin coordinate system (this uses the foot as its fundamental unit). Datasets digitized from the Ramsey County orthophotos (sidewalks, street trees, streetlights) were created in Ramsey-Hennepin coordinates originally but then converted to UTM coordinates later.

2. Transformation of Raw Data to Final Form

Water

Sources: Ramsey County water layer, Met Council hydro layer (from 2000 land-use data)

Steps:

1. Merge Ramsey County and Met Council layers to get most complete coverage
2. Along Rice Creek in Anoka County, make manual corrections due to poor digitizing and poor alignment of Ramsey and Anoka parcels

Parcels

Sources: Ramsey County, Anoka County, Hennepin County, Dakota County (or Met Council Regional Parcel Dataset)

Steps:

1. Clip and merge parcels from various counties (or clip from Met Council data)
   a. Remove Dakota right of way (included as a polygon)
2. Union water layer
3. Correct manually some parcels due to updates in parcel attribute information (see below)
4. Create manually some new parcels along Mississippi (see below)
5. Correct old plat of right of way in Dakota County (along river)
6. Add manually a parcel in Mounds View that was left out
7. Manually adjust alignment of some parcels in St. Paul
8. Create road right of way layer from space left between parcels

Land Use

Sources: I-35W land-use data (polygon matched to parcel lines), Ramsey County land-use data (parcel), Ramsey County online property information search, Ramsey County online maps (parks), Metropolitan Council land-use data (polygons), City of St. Paul park data (parcel data), field-checked land-use data in Anoka, Dakota, and Hennepin Counties (parcel data)

Steps:

1. Begin with complete parcel data set.
2. Add I-35W land-use codes to parcels
   a. Convert from polygons to parcels (check centroid locations)
3. Convert Ramsey County land-use codes to I-35W codes; add to parcels
   a. Correct missing parcels (due to differing dates of parcels and attribute data; required some manual adjustments to parcel lines)
   b. Correct some parcel data without codes (common area around attached townhomes)
4. Make corrections to Ramsey County data (too much vacant and institutional)
   a. Add parks from St. Paul data
   b. Add parks from Ramsey County Web site (park layer)
5. Merge water layer (from two sources) and code as water
6. Correct vacant land use around water bodies (to open space)
7. Add land-use data from field-checking in other counties
   a. Use Met Council data to confirm use
   b. Remove Dakota County parcel within highway right of way; code others to vacant
8. Correct parcel slivers (usually changes to adjacent land uses)
9. Decide to leave government land uses (around Capitol) as institutional (not open space)
10. Carve out new open space parcels along Mississippi River
11. Recode parcels crossed by new Shepard Road
12. Check orthophotos or field-check for remaining mystery parcels
13. Convert to our seven General Land-Use Categories, as well as "Water" and "Other"
14. Based on Ramsey County land-use codes, mark nighttime, social, retail, anti-pedestrian uses); field-check commercial parcels outside Ramsey County or use online property information search

In addition: Decided, for land use, to use our detailed land-use data for inside the context areas and use general Met Council data for areas outside the context areas (See appendix 5C for more information on determining land uses in context areas).

Met Council residential uses were considered to be the following:

- Farmstead
- Seasonal/vacation
- Single-family detached
- Manufactured housing park
- Single-family attached
- Multifamily
- Mixed-use residential

Roads

Sources: The Lawrence Group centerline data

Steps: Remove random Macalaster College road segments (for intersection analysis, adjust these valences to reflect new intersection types).

Appendix 1F: Detailed Instructions for Downloading and Cleaning Census Data

By Joel Koepp

Introduction
The US Census Bureau's American FactFinder Web site (http://factfinder.census.gov) is a quick and easy way to obtain data for normal geographies—cities, counties, census tracts, etc.—but is very cumbersome for situations requiring non-standard geographies.

Documentation
http://www.census.gov/prod/cen2000/doc/sf1.pdf
http://www.census.gov/prod/cen2000/doc/sf3.pdf

The links above are the Technical Manuals for using Census Summary Files 1 and 3. They explain the geographies and contain the complete documentation for all variables, along with metadata.

Importing SF1 into Access
Begin by downloading the Access database template provided by the Census Bureau at
http://www.census.gov/support/2000/SF1/Access97.zip

Open this database and update it to the most recent version of the Access file format. This is done by opening the Tools menu, then the Database Utilities submenu, and finally the Convert Database option. You will be prompted to save the file as a new database with a new name. At this point the database contains only empty tables that have the correct field names and types for the data they will eventually contain.

The raw data is downloaded from the Census FTP site (http://www2.census.gov/census_2000/datasets/) in compressed form. Individual files uncompress into very large text files with .uf1 file extensions. Each one must be renamed to have a .txt extension before importing into Access (e.g., rename mn00009.uf1 to mn00009.txt). The benefit of downloading the Census Access template is to get the import specifications built in. Import specifications ensure that when importing one of the text data files, all the fields are assigned the correct data type and receive the proper names.

Summary File 1 is split into 39 segments and 1 geographic header file. Chapter 7 of the Technical Manual contains a table matrix that shows the segment in which to find each table and individual variable. The most commonly used segments will be 1-11 and 37-39, which contain raw data at the block level. Segments 12-36 contain percentage tables.

With a segment file downloaded, uncompressed, and properly renamed, you can import it into the database.

1. From the **Access** menu, choose **File→Get External Data→Import**.
2. Change the dropdown box **Files of Type** to look for text files.
3. Select the file to import (e.g., mn00001.txt for Segment 1) and click **Import**.
4. In the **Import Text Wizard** that appears, click the Advanced button at the lower left.
5. Click **Specs** to choose an import specification.
6. Find the import specification that matches the segment being imported (in this case, SF10001 Import Specification) and click **Open**.
7. Click **OK**.

8. The wizard will correctly determine that the data is in a delimited format, so click **Finish**.
9. When prompted to choose a table in which to store the data, select the existing empty SF10001 table, click **Next**, and then **Finish**.
10. Right-click on the table and choose **Design View**.
11. Right-click on the **LOGRECNO** field and select **Primary Key**.
12. Click **Save** to preserve the changes in the table.

The geographic header file (called SF1GEO in the database, it gets imported just like the segment files and has its own import specification) and one segment table are the minimum requirements to start making custom queries.

Open the **Tools** menu and choose **Relationships**. Add the SF1GEO table and one of the segment tables to the view. Drag the **LOGRECNO** field (which is the primary key in all tables) from **SF1GEO** to **SF10001**. This establishes the one-to-one relationship between the two tables based on that field. Create this relationship between SF1GEO and any other segment tables (SF10002, SF10003, etc.) you wish to use.

Basic Queries
The first piece of information necessary in constructing a query of Census variables is the geography. It can be a county, a group of cities, a tract, a customized list of specific census blocks, or a combination of multiple geographies. To get the right records out of the database, a query requires the use of the proper summary level codes for the geographies. These are listed in a table on page 83 (chapter 4) of the Technical Manual. The more common codes are 160 for a city, 080 for a census tract, and 101 for a census block.

The SF1GEO table contains one record for every census unit in a state. There are a number of fields that can be used in a query to identify a desired geography. The most helpful are SUMLEV, COUNTY, PLACE, TRACT, BLKGRP, BLOCK, and NAME. For example, to find all the census blocks in Ramsey County, you would build a query that looks for records where COUNTY = 123 and SUMLEV = 101 (this will be explained in more detail later).

Making Useful Extra Tables

After doing a few queries it will probably become apparent that searching for a particular county or city using codes like "123" and "40382" is not very intuitive. It may help save time in the long run to quickly create some tables that contain a correspondence between the county/place codes and the actual names.

In the main **Microsoft Access** window, click on **Queries**. Double-click the icon on the right labeled **Create query in Design View**. In the blank query that appears, add the SF1GEO table and click **Close**. Click and drag the following fields into the columns at the bottom of the query: **SUMLEV**, **COUNTY**, and **NAME**. Under the **Criteria for SUMLEV**, enter **050** (this is the summary level code for a county, as found on page 83 in chapter 4 of the Technical Manual). Click the **red exclamation point icon** in the **Toolbar** to run the query, which should display all the counties in the state. Save the query with a useful name.
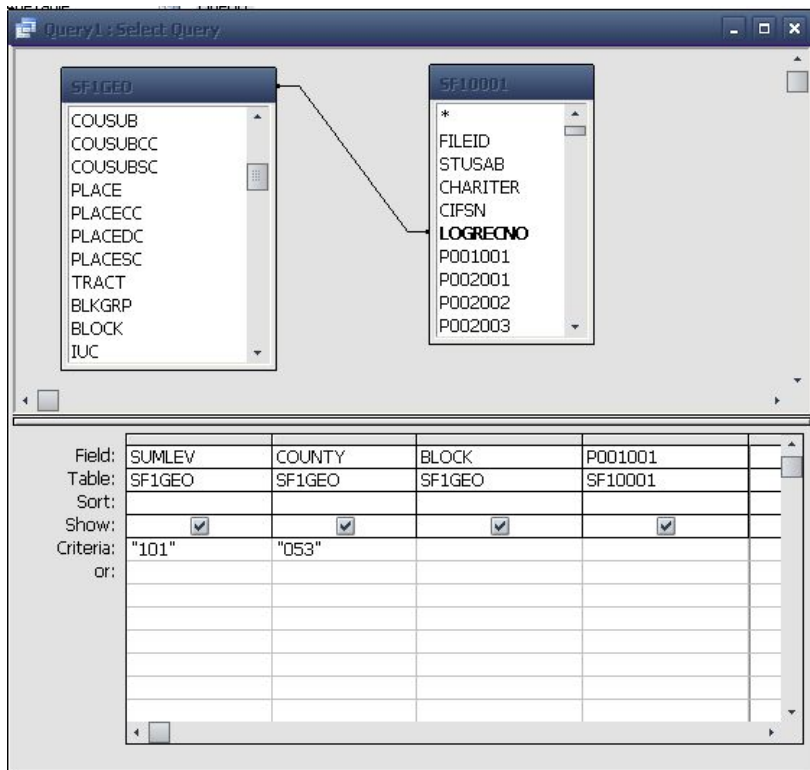
Another useful lookup query to have is one that matches place codes to city names. In the main Access window, click on **Queries**. Double-click the icon on the right labeled **Create query in Design View**. In the blank query that appears, add the SF1GEO table and click **Close**. Click and drag the following fields into the columns at the bottom of the query: **SUMLEV**, **PLACE**, and **NAME**. Under the **Criteria for SUMLEV**, enter **160** (this is the summary level code for a place, as found on page 83 in chapter 4 of the Technical Manual). Uncheck the box under **SUMLEV**, because it doesn't need to be displayed in the query results. Under the **NAME** column, set the **Sort** to be **Ascending** to display the city names alphabetically. Click the **red exclamation point icon** in the **Toolbar** to run the query, which should display all the counties in the state. Save the query with a useful name.

Example Query: Population in all Hennepin County census blocks
This section will describe in detail how to use a query to identify a geography and link in a segment table to pull out a specific census variable.

In the main Access window, click on **Queries**. Double-click the icon on the right labeled **Create query in Design View**. In the blank query that appears, add the **SF1GEO** and **SF10001** tables and click **Close**. The two tables should automatically be joined on the **LOGRECNO** field, assuming a relationship was established between them after they were imported. Click and drag the following fields from **SF1GEO** into the columns at the bottom of the query: **SUMLEV**, **COUNTY**, and **BLOCK**. Then click and drag the **P001001** field from the **SF10001** table into a column at the bottom. Under the **Criteria for SUMLEV**, enter **101** (this is the summary level code for a census block, as found on page 83 in chapter 4 of the Technical Manual). Uncheck the box under **SUMLEV**, because it doesn't need to be displayed in the query results. Under the **Criteria for COUNTY**, enter **053** (found using the county names query created earlier), which will match Hennepin County. Click the **red exclamation point icon** in the **Toolbar** to run the query. **Complex Queries and Calculating Fields**

Age data is a very frequently used part of Census Summary File 1. In addition to simply querying variables, Access allows a user to perform calculations in a query and store the results in a new field. Since the age ranges are stored in separate variables for male and female, it is convenient to use an Access query that sums together a corresponding pair or pairs of variables (e.g., adding together P012003 and P012027 from table SF10002 would yield the total population of children under 5 years old). Similarly, several variables could be added together to identify a particular group of ages, like the key 18- to 34-year-old male demographic.

As an example, the following steps make a query that counts the number of children between the ages of 5 and 14 at the block group level in Ramsey and Dakota Counties.

In the main Access window, click on **Queries**. Double-click the icon on the right labeled **Create query in Design View**. In the blank query that appears, add the **SF1GEO** and **SF10002** tables and click **Close**. The two tables should automatically be joined on the **LOGRECNO** field, assuming a relationship was established between them after they were imported. Click and drag the following fields from **SF1GEO** into the columns at the bottom of the query: **SUMLEV**, **COUNTY**, and **BLKGRP**. Under the **Criteria for COUNTY**, enter **037** and **123** on separate lines (these are the county codes for Dakota and Ramsey Counties, respectively, as found from the county code query created earlier). Under the **Criteria for SUMLEV**, enter **091** on the top two lines (this is the summary level code for a census block group, as found on page 83 in chapter 4 of the Technical Manual). In an empty column at the end of the query, enter the following in the top row marked **Field**:

Age5to14:
P012004+P012005+P012028+P012029

This will create a new field in the output of the query named Age5to14, and the values in that column will be the sum of the four variables that contain the male and female children of ages 5 to 9 and 10 to 14 (see chapter 7, table P12 on page 243 of the SF1 Technical Manual).

New fields can be created using mathematical operations of all kinds—for example, a user could sum together counts of various groups of the population and then calculate it as a percentage of the total population in one line, rather than having to do the work in a spreadsheet. Access also has many text manipulation functions, for concatenating multiple strings of text or stripping off characters from the left or right ends of a string.

Combining GIS and a Census Database
The ESRI Web site provides access to free Census GIS data for the entire United States. The Census Bureau's 2000 TIGER/Line data can be downloaded for individual counties after
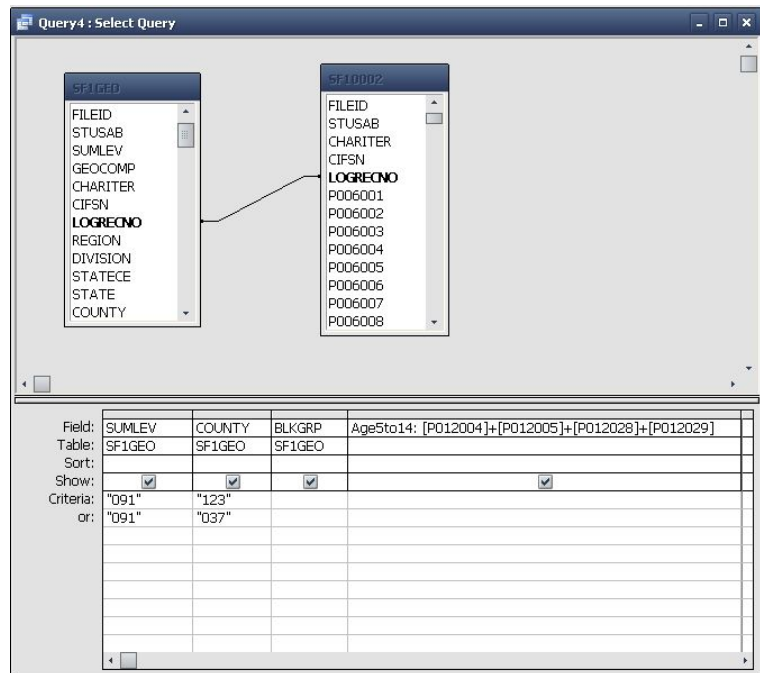
selecting a state. The TIGER/Line data contains line features (roads, railroads, etc.), landmarks, and boundary files (census geographies, city and county boundaries, traffic analysis zones, etc.).

Access Census data for download here:
http://arcdata.esri.com/data/tiger2000/tiger_download.cfm

Once a shapefile is downloaded, uncompressed, and added to an **ArcGIS Document**, it is ready for analysis or processing. Each census geography shapefile will have in its **Attribute Table** a field called **STFID**, which serves as the unique identifier for each feature. For example, a layer of census tracts will have an 11-character value in the **STFID** field, with each value unique to a particular census tract. A layer of block groups will have **STFID** values with 12 characters, and census blocks will have 15-character **STFID** values.

The **STFID** is created by concatenating the values of the **STATE**, **COUNTY**, **TRACT**, **BLKGRP**, and **BLOCK** fields (as found in the **SF1GEO** table. The **STATE** portion is two characters, **COUNTY** is three characters, **TRACT** is six characters, **BLKGRP** is one character, and **BLOCK** is four characters. Note that the first character of the **BLOCK** field is identical to the **BLKGRP** value.

To construct a value in the Census database that will match the **STFID** values in the GIS shapefiles, use the following:

Tract:	**[STATE] & [COUNTY] & [TRACT]**
Block Group:	**[STATE] & [COUNTY] & [TRACT] & [BLKGRP]**
Block:	**[STATE] & [COUNTY] & [TRACT] & [BLOCK]**

Appendix 1G: Table of Data Sources Required by Protocols

By Ed D'Sousa

Introduction
The following table documents the data requirements for each protocol. This table is meant to be a quick reference when planning GIS measures in a study area. For a detailed description of the data sources used in the Twin Cities Walking Study, please refer to section 1.5 of this document.

Additionally, the table indicates the source for each data set used in the Twin Cities Walking Study. The Twin Cities area has a strong GIS tradition and subsequently has a good selection of high quality data sets available. This may not be the case in other areas and each data sets availability and quality should be assessed prior to starting any study involving the GIS measures.

Definitions
- Study Specific/Collected: The data for this data set was collected specifically for the current study by the researchers and their staff. The Participant Locations are integral to most protocols in some form, either used directly as point data or in the form of a buffer. The Urban Design Inventory requires on site surveys. Both data sets represent a significant investment in time to collect and verify.

- Derived Data or Data Modified from Original Source (with source indicated): These data sets are derived from another data set used in the study. Generally, the modification involves a change in the data type necessary to complete a specific protocol, for example the buffers used in many protocols are derived from the participant locations, and do not represent a significant investment in time, but are used by many of the protocols and need only be created once and saved. Some of the derived data sets are modified to remove or enhance data provided by third parties. For example, the roads require significant modification to eliminate unwanted features. These modified data sets can represent a significant investment in time.

- Collected for this study but may be available from Public/Private Source: These data sets are often available from locale government agencies, but may be incomplete, out of date or require verification. Additionally, if the study area covers a large area that involves several jurisdictions, there many only be data available a portion of the area. To complete the data for a entire study area may represent a significant investment in time. Often, several sources, of varying quality and accuracy, will need to be reconciled and combined to create a single data set to cover the area. Again, this process can represent a significant investment in time.

- Public/Private Data: These data sets are generally available through various data providers for most areas in the United States and were used "as is" without modification or enhancement. The scale, quality and accuracy of the data was suitable for the study and was well documented the originating sources.

Data Feature Types:
P – Point, L – Line, A – Polygon/Area

		Key:	Study Specific/Collected			Derived Data or Data Modified from Original Source (with source indicated)			Collected for this study but may be available from Public/Private Source or is a Combiation of Collected and Public/Private Data			Public/Private Source - Original Data			
Data Feature Type:		**P**	**A**	**P**	**A**	**P/A**	**A**	**A**	**L**	**A**	**A**	**A**	**A**	**A**	**P**	**P**
Protocol		Participant Locations	Urban Design Inventory	Study Sites	Points of Interest	Sidewalks	Trees	Water	Study Area (Analysis extent)	Roads	Parcels	Census	Employment Data	Building Footprints	Street Lights	Transit Stops	Retail Locations
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2.1	Study Sites w/o Water			(1,8)				■									
2.2	Study Sites w/Water			(1,8)													
2.3	Focus Area Sampling Grid								■								
2.4	Grid Cell Measurement Geography								■								
2.5	Straight Line Buffer Measurement Geography	■															
2.6	Network Buffer Measurement Geography									(9)							
2.7	Straight Line Distance to Nearest Feature	■			■												
2.8	Network Distance to Nearest Feature	■			■					(9)							
2.9	Parcel of Residence	■									■						
3.1	Raw Population per Unit Land Area			(1,8)								■					

Protocol

Protocol		Participant Locations	Urban Design Inventory	Study Sites	Points of Interest	Sidewalks	Trees	Water	Study Area (Analysis extent)	Roads	Parcels	Census	Employment Data	Building Footprints	Street Lights	Transit Stops	Retail Locations
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3.2	Population per Unit Land Area (w/o Water)			(1,8)				■				■					
3.3	Population per Developed Land Area (w/o Water)			(1,8)				■			■	■					
3.4	Residential Population in Residential Parcels (Residential Density) (w/o Water)			(1,8)				■			■	■					
3.5	Population plus Employment per Unit Land Area (w/o Water)			(1,8)				■				■	■				
3.6	Employment per Unit Land Area (w/o Water)			(1,8)				■					■				
3.7	Housing Units per Unit Land Area (Census Data) (w/o Water)			(1,8)								■					
3.8	Lot Coverage (w/o Water)			(1,8)							■			■			
3.9	Housing Units per Unit Land Area (Parcel Data) (w/o Water)																
3.10	Housing Units per Unit Residential Land Area (Parcel Data) (w/o Water)																

Protocol

	Participant Locations	Urban Design Inventory	Study Sites	Points of Interest	Sidewalks	Trees	Water	Study Area (Analysis extent)	Roads	Parcels	Census	Employment Data	Building Footprints	Street Lights	Transit Stops	Retail Locations
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3.11 Households per Unit Land Area (Parcel Data) (w/o Water)																
4.1 Sidewalk Length per Unit Area			(1,8)		■											
4.2 Sidewalk Length Divided by Road Length (Sidewalk Length PER? Road Length)			(1,8)		■				■							
4.3 Sidewalk Length per Length of Major Road			(1,8)		■				(9)							
4.4 Streetlights per Length of Road			(1,8)											■		
4.5 Street Trees per Length of Road			(1,8)			■			(9)							
4.6 Percent of Street Segments with Marked Pedestrian Crossings at One or Both Ends		■	(1,8)						■							
4.7 Percent of Street Segments with Visible Litter, Graffiti or Dumpsters		■	(1,8)						■							
4.8 Percent of Street Segments with Traffic Calming		■	(1,8)						■							

Protocol

	Participant Locations (1)	Urban Design Inventory (2)	Study Sites (3)	Points of Interest (4)	Sidewalks (5)	Trees (6)	Water (7)	Study Area (Analysis extent) (8)	Roads (9)	Parcels (10)	Census (11)	Employment Data (12)	Building Footprints (13)	Street Lights (14)	Transit Stops (15)	Retail Locations (16)
4.9 Other Fieldwork-based measurements of neighborhood characteristics																
4.10 Distance to Nearest Transit Stop	■								(9)						■	
4.11 Transit Stop Density			(1,8)												■	
4.12 Percentage of Blocks with Full Length Sidewalks																
4.13 Percentage of Blocks with Streetlights																
5.1 Percentage of Total Parcel Area in Major Land Uses			(1,8)							■						
5.2 Percentage of Land Area in Nighttime Uses										(10)						
5.3 Percentage of Land Area in Social Uses										(10)						
5.4 Percentage of Land Area in Retail Uses										(10)						
5.5 Percentage of Land Area in Industrial and Auto-Oriented Uses										(10)						
5.6 Percentage of Land Area in Industrial and Auto-Oriented Uses	■						???			■						

Protocol

		Participant Locations	Urban Design Inventory	Study Sites	Points of Interest	Sidewalks	Trees	Water	Study Area (Analysis extent)	Roads	Parcels	Census	Employment Data	Building Footprints	Street Lights	Transit Stops	Retail Locations
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
5.7	Proportion of Dissimilar Land Uses Among Grid Cells in an Area, Formula 1																
5.8	Dissimilarity Index (Formula 2)			(1,8)							■						
5.9	Entropy Index										(10)						
5.10	HHI - Herfindahl-Hirschman Index										(10)						
5.11	Retail Employment per Unit Area	???		(1,8)				???					■				
5.12	Density of Employees in Major Retail Subcategories (…..)			(1,8)				???					■				
5.13	Distance to Nearest Facility with Employees and … in … Subcategories	■			■												■
5.14	Distance to Nearest Facility for Comparison with NQLS Survey	■			■												■
5.15	Gravity Measure with Competition	■									■						
6.1	Average Census Block Area			(1,8)								■					
6.2	Median Census Block Area			(1,8)								■					

Protocol

		Participant Locations	Urban Design Inventory	Study Sites	Points of Interest	Sidewalks	Trees	Water	Study Area (Analysis extent)	Roads	Parcels	Census	Employment Data	Building Footprints	Street Lights	Transit Stops	Retail Locations
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6.3	Ratio of Area within X Street Distance to Area within X Distance Radius			(1,8)													
6.4	Number of Access Points			(1,8)						▉							
6.5	Road Length per Unit Area			(1,8)						▉							
6.6	Intersections per Unit Area			(1,8)						▉							
6.7	Ratio of Four-Way Intersection to All Intersections			(1,8)						▉							
6.8	Four-Way Intersection per Unit Land Area			(1,8)						▉							
6.9	Ratio of Three-Way Intersections to All Intersections			(1,8)						▉							
6.10	Connnected Node Ratio			(1,8)						▉							
6.10	Connnected Node Ratio			(1,8)						▉							
6.11	Median Perimeter of Block			(1,8)						▉							
6.12	Link-Node Ratio			(1,8)						▉							

Appendix 4A: Modified Boarnet-Day Inventory, Now the Irvine-Minnesota Inventory

By Ann Forsyth drawing on collaborative work by Kristen Day, Marlon Boarnet, and Mariela Alfonzo

Note, this is the inventory as used in the Twin Cities Walking Study. Developed by a team led by Marlon Boarnet and Kristin Day, the inventory was in draft form in April 2004.

The inventory as used is an abridged version of the April 2004 draft of the inventory. As such it kept the original numbering of the inventory, meaning that there are some "missing" numbers.

A revised version has been created, however, and is available on the Active Living Research Web site. Called the Irvine-Minnesota Urban Design Inventory, it includes reliability information, http://www.activelivingresearch.org/index.php/Irvine_Minnesota_Inventory/334. It has also been published (Day et al 2006a, 2006b).

In the inventory as used, a number of coding protocols were used to clarify skipped items in particular and a revised version has been prepared that removes the "skip question and code as 9" type items, but instead provides a "skip/not applicable" option in each question. A separate paper is available that describes how this inventory was used (Forsyth et al. 2004a).

It should be noted that the inventory was conducted on a 20 percent sample of roads with pedestrian access, with divided roads treated as one road (if both sides were selected, one was removed and a new random segment drawn). Because the road data contained a number of very small segments, those less than 50 feet long were excluded.

Use 8 not slash to indicate NA; in excel code a skipped answer as 9 unless otherwise instructed below

early line numbering—to track data	Question / Item	Coding	Date	1	2	3	4	5	6	7	8	9	10
1	Observer / SEGMENT >												
	Answer questions 1-6 based on this end of the segment												
3	INTERSECTION--FILL IN!!!!												
4	**Neighborhood Identification**												
5	1. Are there monuments or markers including neighborhood entry signs that indicate that one is entering a special district or area?	yes = 1; no = 0											
6	**Street Crossing**												
7	2a. Consider the places on the segment that are intended for pedestrians to cross the street. Are these places marked for pedestrian crossing?	All = 3; Some = 2; none = 0; N A = 8											
8	If 2a = "None" or "N/A", skip to question 3 [after coding 2b as 0]												
9	2b. What type of marking do the crosswalks have? Mark all that apply.												
10	White painted lines	yes = 1; no = 0											
11	Colored painted lines	yes = 1; no = 0											
12	Zebra striping	yes = 1; no = 0											
13	Different road surface or paving (e.g. tiles, colored concrete, marble, etc)	yes = 1; no = 0											
14	Other	yes = 1; no = 0											
15	3. Are there curb cuts at all places where crossing is expected to occur?	All = 3; Some = 2; none = 0; N A = 8											
	4. What type of traffic/pedestrian signal(s)/system(s) is/are provided? Mark all that apply.												
17	Traffic signal	yes = 1; no = 0											
18	Stop sign	yes = 1; no = 0											
19	Yield sign	yes = 1; no = 0											
20	Pedestrian activated signal	yes = 1; no = 0											
21	Pedestrian crossing sign	yes = 1; no = 0											
22	Pedestrian overpass/underpass/bridge	yes = 1; no = 0											
23	None	yes = 1; no = 0											
24	5. For an individual who is on this segment, how safe (traffic wise) do you think it is to cross the street from this segment?	Pretty/very safe = 1; not very safe/ unsafe = 0 Cul de sac = 8											
25	6. For an individual who is on this segment, how convenient (traffic wise) do you think it is to cross the street from this segment?	pretty/very convenient =1; not very/inconvenient= 0; cul de sac = 8											
	Answer questions 7-11 while standing at the beginning of the segment												
27	**Neighborhood Identification**												
28	7. Does the segment have banners that identify the neighborhood?	Some/a lot = 3; few = 2; none = 0											
29	**Street Characteristics**												
30	8a. Is this a pedestrianized street?	yes = 1; no = 0											
31	8b. Is the street a …	one way = 1; two way = 2											
32	9. Is this segment an alley?	yes = 1; no = 0											

Use 8 not slash to indicate NA; in excel code a skipped answer as 9 unless otherwise instructed below

	line # (to track data)	Date / code	1	2	3	4	5	6	7	8	9	10
10. How many vehicle lanes are there for cars? (include turning lanes) [changed scale]	33	count										
Views	34											
11a. Is this segment characterized by having a significant open view of an object or scene that is not on the segment? The view must be a prominent one.		yes = 1; no = 0										
	35											
If 11a = No, skip to question 12.	36											
11b. How attractive is the open view?	37	3 =Attractive; 2 =neutral; 1=Unattractive.										
Begin walking along segment to answer questions 12-68												
12c. How many of the buildings in this segment contain vertical-mixed use, that is, the building has different land uses on different floors of the building?	39	Some/a lot = 3; few = 2; none = 0; NA no buildings>1 story = 8										
12d. Mark if any of these building types are present (focusing on the form of the building).	40											
Big box shops (includes super stores or warehouse stores)	41	yes = 1; no = 0										
Shopping mall	42	yes = 1; no = 0										
Strip mall/row of shops	43	yes = 1; no = 0										
Drive-thru	44	yes = 1; no = 0										
None of the above	45	yes = 1; no = 0										
13a. Mark off all types of public space(s) on this area and how attractive it is	47	3 = attractive; 2 = neutral; 1 = unattractive; 0 = no space;										
Park/playground	48	above										
Playing or sport field	49	above										
Plaza /square /courtyard	50	above										
Public garden	51	above										
Beach	52	above										
Other	53	above										
13d. Is it possible for the general public to use the public space(s)?	54	unclear = 2; yes = 1; no = 0										
Other Land Uses	55											
14. How many of these land uses are present on this segment?	56											
Bars/night clubs	57	**Some/a lot = 3; few = 2; none = 0**										
Adult uses	58	**Some/a lot = 3; few = 2; none = 0**										
Check cashing stores/pawn shops/bail bond stores	59	**Some/a lot = 3; few = 2; none = 0**										
Liquor stores	60	**Some/a lot = 3; few = 2; none = 0**										
15. How many of the following gathering places are on this segment?	61											
Restaurants	62	**Some/a lot = 3; few = 2; none = 0**										
Coffee shops	63	**Some/a lot = 3; few = 2; none = 0**										
Libraries/bookstores	64	**Some/a lot = 3; few = 2; none = 0**										
"Corner" store	65	**Some/a lot = 3; few = 2; none = 0**										

	early line numbering— to track data	Date	1	2	3	4	5	6	7	8	9	10
Art or craft galleries	66	Some/a lot = 3; few = 2; none = 0										
Farmers market	67	yes = 1; no = 0										
16. Are these nature features present on this segment?												
Open field/golf course	68	yes = 1; no = 0										
Lake/pond	69	yes = 1; no = 0										
Fountain/reflecting pool	70	yes = 1; no = 0										
Stream/River/Canal/Creek	71	yes = 1; no = 0										
Forest or woods	72	yes = 1; no = 0										
Mountain or hills	73	yes = 1; no = 0										
Barriers	74											
17a. Are the following **barriers** present on this segment. Check all that apply. **Mark response in column A.** and whether barrier can be overcome e.g. there's a pedestrian bridge.		0 = no barrier, 1 = can be overcome; 2 can be somewhat overcome; 3 can not be overcome										
Highway (elevated or below ground)	75	above										
Railroad track	76	above										
Impassable land use (e.g., gated community, major industrial complex, etc.)	77	above										
River	78	above										
Drainage ditches	79	above										
Fences/walls	80	above										
Road with 6 or more lanes	81	above										
Other	82	above										
	83	above										
Sidewalks	84											
18a. How many sides of the street have sidewalks?	84a	Count 1 or 2										
18b. Is the sidewalk **complete** on one or both sides?		yes = 1; no = 0; 8 = NA										
18c. What is the **predominant sidewalk width**? (If no predominant width, mark all that apply).	85	0 = under 4 ft; 1 = 4-7 ft; 2 = 7+; 8 = NA (no sdiewalk)										
18d. What is the **condition** or maintenance of the sidewalk?	86	0 = Poor; 1 = moderate or good; 2 = under repair										
18e. Is there a decorative or unique paving on that **covers most or all of the sidewalk** on the segment? (e.g., bricks, tile, etc.)	87	yes = 1; no = 0										
18f. Determine how much of the sidewalk is covered by these **features that provide protection** from sun, rain, and/or snow.		1 = some/ much of s'walk covered; 0 = No/little covered;										
Arcades	88	above										
Awnings	89	above										
Other	90	above										
	91											
18g. Is there is a **buffer** (for example, parked cars, landscaped "buffer" strip, etc.) between sidewalk or street.	92	2 = both sides; 1 = 1 side; 0 = no buffer										
If 18g = "No buffer", skip to question 19 [after coding 18h as 0].	93											
18h. Mark the **forms of buffers** that are present. Mark all that apply. **must be present along most or the entire segment on 1+ sides***	94											

Page 3

Use 8 not slash to indicate NA; in excel code a skipped answer as 9 unless otherwise instructed below

	early line numbering—to track data	Date	1	2	3	4	5	6	7	8	9	10
Angled parking of cars	95	yes = 1; no = 0										
Parallel parking of cars	96	yes = 1; no = 0										
Street trees that run along the most of or the entire segment, between sidewalk and the street	97	yes = 1; no = 0										
Landscape "buffer" strip or bollards (posts) between sidewalk and street	98	yes = 1; no = 0										
Other	99	yes = 1; no = 0										
19. Are there sidewalks/greenbelts/trails/paths other than sidewalks along street?	100	yes = 1; no = 0										
Bicycle Lanes	101											
20a. Are there bicycle lanes on the segment?	102	yes = 1; no = 0										
If 20a = "No", skip to question 21.	103											
20b. How are the bicycle lanes demarcated?	104	On road, painted line/ reflectors =3; on road physical separation = 2; off road = 1										
Mid Block Crossing	105											
21a. Is there a marked mid-block crosswalk for pedestrians?	106	yes = 1; no = 0										
If 21a = No, skip to question 22 [After coding 21b as 0].	107											
21b. What type of marking does the crosswalk have? Mark all that apply												
White painted lines	109	yes = 1; no = 0										
Colored painted lines	110	yes = 1; no = 0										
Zebra striping	111	yes = 1; no = 0										
Different road surface or paving (e.g. tiles, colored concrete, marble, etc)	112	yes = 1; no = 0										
Other	113	yes = 1; no = 0										
Steepness	114											
22. How steep or hilly is this segment? Mark all that apply.	115	steep = 1; flat or gentle = 0										
Sidewalk Amenities	116											
23. Are there outdoor dining areas (e.g., cafes, outdoor tables at coffee shops or plazas, etc) located on the segment?	117	Some/a lot = 3; few = 2; none = 0										
24a. Indicate how many of each of the following street furniture/sidewalk amenities is/are present on the segment.	118											
Benches (not a bus stop), chairs and/or ledges for sitting	119	Some/a lot = 3; few = 2; none = 0										
Bus stops with seating	120	Some/a lot = 3; few = 2; none = 0										
Heat lamps	121	Some/a lot = 3; few = 2; none = 0										
Bike racks	122	Some/a lot = 3; few = 2; none = 0										
Trash cans	123	Some/a lot = 3; few = 2; none = 0										
If 24a = "None" for all types of street furniture, skip to question 25.												
24b. Looking at the street furniture on this segment overall, would you describe it as generally attractive?	125	3 = Attractive; 2 = neutral; 1 = Unattractive										

Page 4

Use 8 not not slash to indicate NA; in excel code a skipped answer as 9 **unless** otherwise instructed below

Question	early line numbering— to track data	Date (response codes)	1	2	3	4	5	6	7	8	9	10
25. Are there obvious public restrooms on this segment that are **clearly open to the public**?	126	yes = 1; no = 0										
Buildings	127												
28. How many buildings on this segment can you enter **directly from the street**?	128	**Some/a lot** = 3; few = 2; none = 0										
31. Are there **abandoned buildings or lots** on this segment?	129	**Some/a lot** = 3; few = 2; none = 0										
32a. Does at least **50% of the segment have buildings?**	130	yes = 1; no = 0										
If 32a = "No", skip to question 33.	131												
Windows	132												
32b. From how many buildings on the segment can pedestrians easily be seen?	133	0 = none, 1 = < 50%, 2 = > 50%										
33. How many buildings on this segment have **windows with bars**? (proportion)	134	**Some/a lot** = 3; few = 2; none = 0; NA = 8										
Other features of buildings	135												
34. How many buildings on this segment have **front porches**? (proportion–porches you can sit on)	136	**Some/a lot** = 3; few = 2; none = 0; NA = 8										
35. How many buildings on this segment have **balconies** facing the segment? (proportion)	137	**Some/a lot** = 3; few = 2; none = 0; NA = 8										
36. How much of the segment has **blank walls or buildings with blank walls**?	138	All = 3; Some = 2; none = 0; NA = 8										
37a. On this segment, what percentage of the properties contain **fences/walls/hedges/elevated ground** that separate the building(s) or property(ies) from the street? (Hedges must be acting as fences/walls to be counted)	139	0 = none, 1 = < 50%, 2 = > 50%										
If 37a = "none", skip to question 38a.	140												
37b. Describe the **attractiveness** overall of the walls and/or fences on this segment.	141	3 = Attractive; 2 = neutral; 1 = Unattractive										
37c. What is the approximate **height** of the fences/walls on the segment? Mark all that apply	142												
Less than 4 feet	143	yes = 1; no = 0										
4 feet and over	144	yes = 1; no = 0										
Garages	145												
38a. How many buildings have **garage doors** facing the street?	146	**Some/a lot** = 3; few = 2; none = 0; NA = 8										
If 38a = "none or one" or "N/A", skip to question 39a.	147												
38b. How **prominent** are most garage doors when looking at the front of the buildings?	148	Very = 3; somewhat = 2; not very/not visible = 0										
Parking	149												
39a. Is there a **parking structure** visible on this segment (do not include parking structures that are completely underground)?	150	yes = 1; no = 0										
If 39a = "No" skip to question 40a.	151												

ENVIRONMENT AND PHYSICAL ACTIVITY: GIS PROTOCOLS VERSION 4.1, JUNE 2007, WORK IN PROGRESS 218
University of Minnesota and Cornell

Use 8 not slash to indicate NA; in excel code a skipped answer as 9 unless otherwise instructed below

Question	line numbering—to track data	Date	1	2	3	4	5	6	7	8	9	10
39b. Looking at the front of the parking structure on the street level floor, what is the **predominant use** that is visible to you?	152	Parking = 2; varied = 1; not parking/other uses = 0										
39c. What percentage of the segment is covered by the parking structure?	153	all or almost all = 3; about half = 2; less than half = 1										
40a. Is there a **parking lot** that fronts/faces the street on either side of the segment?	154	2 = both sides; 1 = 1 side; 0 = no lots										
If 40a = "neither side has a parking lot", skip to question 41.												
40c. How much of the side(s) that has the parking lot fronting the street is(are) covered by the parking lot? (Consider the largest parking lot if more than one)	157	all or almost all = 3; about half = 2; less than half = 1										
40d. How many **trees** are there in the parking lot, compared to the number of parking spaces?	158	**Some/a lot = 3**; few = 2; none = 0										
Driveways	159											
41. How many **driveways** are visible on the segment?	160	**Some/a lot = 3**; few = 2; none = 0										
Maintenance	161											
42a. Describe the general **maintenance of the buildings** on this segment.	162	3 = Attractive; 2 = neutral; 1 = Unattractive; 8= NA										
42b. Describe the general **maintenance of the landscaping** on this segment.	163	3 = Attractive; 2 = neutral; 1 = Unattractive; 8= NA										
If 42b = unattractive or NA, skip to question 44.												
43. **How much** of the segment contains well-maintained (fair or good) landscaping?	165	1 = < 50%, 2 = > 50%										
44. How much **graffiti** is apparent on this segment?	166	**Some/a lot = 3**;little = 2;none = 0										
45. How much **litter** is apparent on this segment?	167	**Some/a lot = 3**;little = 2;none = 0										
46. Are there **dumpsters** visible on this segment?	168	**Some/a lot = 3**; few = 2;none = 0										
47. Is there visible **electrical wiring** overhead on the	169	**Some/a lot = 3**; few = 2; none = 0										
Lighting	170											
48b. Are the outdoor light fixtures **mostly less than or greater than 13 feet tall** (i.e. one story)?	171	1 = more than 13; 0 = less										
Freeways	172											
49. Is there a **freeway overpass/underpass** connected to this segment?	173	under a freeway overpass =3; next to freeway = 2; IS a freeway overpass = 1; none of the above = 0										
Traffic Features	174											
50. What is the **posted speed limit** on this segment? Only include those on the segment itself.	175	Use number; 8 = not posted										
51. Are there measures on this segment that couldslow down traffic? Mark all that apply.	176											
Speed bump/speed hump/raised crosswalk, or dips (that are intended to slow down traffic)	177	yes = 1; no = 0										
Rumble strips or bumps (includes dots, reflectors, raised concrete strips, etc.)	178	yes = 1; no = 0										
Curb bulb out/curb extension	179	yes = 1; no = 0										
Traffic circle/roundabout	180	yes = 1; no = 0										

Page 6

Use 8 not slash to indicate NA; in excel code a skipped answer as 9 unless otherwise instructed below

Item	Line #	Coding / Date	1	2	3	4	5	6	7	8	9	10
		Date										
Median	181	yes = 1; no = 0										
Street Trees (that run along most or the entire segment – do not have to be on both sides; can be in yards)	182	yes = 1; no = 0										
Angled/ On-street parking (that runs along most or the entire segment - does not have to be on both sides of segment)	183	yes = 1; no = 0										
Other	184	yes = 1; no = 0										
52a. Is there a cul-de-sac or permanent street closing on this segment?	186	yes = 1; no = 0										
If 52a = "No" skip to question 53.	187											
52b. Is there a pedestrian access point or cut through point that allows pedestrians to go from one segment to another even though vehicular traffic may not be able to)?	188	yes = 1; no = 0; Don't know = 7										
Architecture/Design	189											
53. Rate the attractiveness of the segment. (design + maintenance)	190	3 = Attractive; 2 = neutral; 1 = Unattractive										
54. Does this segment have buildings that appear to be historic? (old + detailed)	191	yes = 1; no = 0; NA = 8										
55. How interesting is the architecture/urban design of this segment?	192	3 = interesting; 2 = somewhat int.; 1 = Uninteresting										
56. How alike are the designs of the buildings on the segment?	193	2= all similar design; 1 = range of different designs; 8 = NA (0-1 buildings)										
If 56 = "All of similar designs" or "N/A" skip to question 58.												
57. How much do the houses or buildings appear to fit or go together (how compatible they are), in terms of architectural appearance? (Buildings do not have to be exactly the same to go together)	195	Some/a lot = 3; a little = 2; none = 0										
Other features of the segment	196											
62. How many street vendors or stalls are on this segment? (do not count newspaper racks; there must be a person vending)	197	Some/a lot = 3; few = 2; none = 0										
63. Is there public art that is visible on this segment?	198	yes = 1; no = 0										
64. Are there billboards present on this segment?	199	Some/a lot = 3; few = 2; none = 0										
People	200											
65. How many other people were present outdoors while you were observing this segment?	201	Some/a lot = 3; few = 2; none = 0										
If 65 = "None" skip to question 67.	202											
66. Thinking of the people on this segment, how safe do you feel walking on this segment?	203	Pretty/very safe = 1 not very safe/ unsafe = 0										
Dogs	204											
67. Are there any loose/unsupervised/barking dogs on this segment that seem menacing?	205	yes = 1; no = 0										
Olfactory Character	206											
68. Is the dominant smell unpleasant?	207	yes = 1; no = 0										

Page 7

Use 8 not slash to indicate NA; in excel code a skipped answer as 9 unless otherwise instructed below

	line numbering—to track data early	Date		1	2	3	4	5	6	7	8	9	10
Neighborhood Identification (OTHER END OF SEGMENT)													
1. Are there **monuments or markers** including neighborhood entry signs that indicate that one is entering a special district or area?	209	yes = 1; no = 0											
Street Crossing	210												
2a. Consider the places on the segment that are **intended** for pedestrians to cross the street. Are these places marked for pedestrian crossing?	211	All = 3; Some = 2; none = 0; NA = 8											
If 2a = "None" or "N/A", skip to question 3 [after coding 2b as 0]	212												
2b. What type of **marking** do the crosswalks have? Mark all that apply.	213												
White painted lines	214	yes = 1; no = 0											
Colored painted lines	215	yes = 1; no = 0											
Zebra striping	216	yes = 1; no = 0											
Different road surface or paving (e.g. tiles, colored concrete, marble, etc)	217	yes = 1; no = 0											
Other	218	yes = 1; no = 0											
3. Are there **curb cuts** at all places where crossing is expected to occur?	219	All = 3; Some = 2; none = 0; NA = 8											
4. What type of **traffic/pedestrian signal(s)/system(s)** is/are provided? Mark all that apply.	220												
Traffic signal	221	yes = 1; no = 0											
Stop sign	222	yes = 1; no = 0											
Yield sign	223	yes = 1; no = 0											
Pedestrian activated signal	224	yes = 1; no = 0											
Pedestrian crossing sign	225	yes = 1; no = 0											
Pedestrian overpass/underpass/bridge	226	yes = 1; no = 0											
None	227	yes = 1; no = 0											
5. For an individual who is on this segment, how **safe** (**traffic wise**) do you think it is to cross the street from this segment?	228	Pretty/very safe = 1; not very safe/ unsafe = 0 Cul de sac = 8											
6. For an individual who is on this segment, how **convenient** (**traffic wise**) do you think it is to cross the street from this segment?	229	pretty/very convenient =1; not very/inconvenient= 0; cul de sac = 8											

Appendix 5A: Ramsey County Land-Use Codes

Land Use Code	Type	Land-Use Description
109	A	GREENHOUSES, VEGETABLES & FLORACULT.
300	I	INDUSTRIAL, VACANT LAND
310	I	FOOD&DRINK PROCESS PLANTS AND STORAE
320	I	FOUNDERIES&HEAVY MANUFACT PLANTS
340	I	MANUFACTURING & ASSEMBLY LIGHT
390	I	GRAIN ELEVATORS
399	I	OTHER INDUSTRIAL STRUCTURES
400	C	COMMERCIAL VACANT LAND
401	M	APARTMENTS 1-9 RENTAL UNITS
402	M	APARTMENTS 10-19 RENTAL UNITS
403	M	APARTMENTS 20-49 RENTAL UNITS
404	M	APARTMENTS 50 OR MORE UNITS
405	M	APARTMENT VACANT LAND
406	M	APARTMENT MISC IMPROVEMENTS
410	C	MOTELS AND TOURIST CABINS
411	C	HOTELS
412	C	NURSING HOMES & PRIVATE HOSPITALS
415	C	TRAILER OR MOBILE HOME PARK
420	C	SMALL (UNDER 10,000SF) DETACH RETAI
421	C	SUPERMARKETS
422	C	DISCOUNT STORES & JR. DEPT STORES
423	C	MEDIUM DETACHED RETAIL
424	C	FULL LINE DEPARTMENT STORES
425	C	NEIGHBORHOOD SHOPPING CENTER
426	C	COMMUNITY SHOPPING CENTER
428	C	VETERINARY CLINIC
429	C	OTHER RETAIL STRUCTURES
430	C	RESTURANT, CAFETERIA, AND/OR BAR
431	C	SMALL STRIP CENTER
432	C	CONVENIENCE STORE
435	C	DRIVE-IN REST/FOOD SERVICE FACILITY
437	C	DAY CARE CENTERS
441	C	FUNERAL HOMES
442	C	MEDICAL CLINICS AND OFFICES
444	C	FULL SERVICE BANKS
447	C	OFFICE BUILDING 1-2 STORIES
448	C	OFFICE BULD 3 OR MORE STORIES WALKUP
449	C	OFFICE BULD 3 OR MORE STORIES ELEVAT
450	C	CONDOMINIUM OFFICE UNITS
451	C	GAS STATION
452	C	AUTOMOTIVE SERVICE STATION
453	C	CAR WASHES
454	C	AUTO CAR SALES AND SERVICES
456	C	PARKING GARAGE STRUCTURE & LOTS
458	C	CI/CONDO/CO-OP NON-TAX OUTLOT
460	C	THEATERS

463	C	GOLF COURSES
464	C	BOWLING ALLEYS
465	C	LODGE HALLS AND AMUSEMENT PARKS
479	C	FLEX INDUSTRAIL CENTER
480	C	COMMERCIAL WAREHOUSES
481	C	MINI WAREHOUSE
482	C	COMMERCIAL TRUCK TERMINALS
499	C	OTHER COMMERCIAL STRUCTURES
500	R	RESIDENTIAL, VACANT LAND, LOT
501	E	RESID/EXEMPT VACANT LAND
510	R	SINGLE-FAMILY DWELLING, PLATTED LOT
511	R	RES/EXEMPT SINGLE-FAMILY DWELLING
515	R	TWO RESIDENCES ON ONE PARCEL
520	R	TWO FAMILY DWELLING, PLATTED LOT
521	E	RES/EXEMPT MULTI-FAMILY DWELLING
525	R	CONVERSION
530	R	THREE-FAMILY DWELLING, PLATTED LOT
540	R	DOUBLE DWELLING
545	R	1/2 DOUBLE DWELLING
546	E	RES/EXEMPT 1/2 DOUBLE DWELLING
550	R	CONDO/CO-OP
552	R	CONDO/CO-OP GARAGE OR STORAGE
555	R	CONDO/CO-OP NON-TAX OUTLOT
570	R	TOWNHOME-INNER
573	R	TOWNHOME-DETACHED
575	R	TOWNHOME-OUTER
577	R	TOWNHOME NON-TAX OUTLOT
579	R	RESIDENTIAL, TOWNHOME VACANT
600	E	EXEMPT PROPERTY OWNED BY USA
610	E	EXEMPT PROPERTY OWNED BY STATE OF MN
615	E	EXEMPT PROPERTY OWNED BY MET COUNCIL
620	E	EXEMPT PROPERTY OWNED BY COUNTIES
635	E	FIRESTATIONS
640	E	EXEMPT PROPERTY OWNED BY MUNICIPALS
641	E	EXEMPT LAND
642	E	EXEMPT CLINICS
644	E	SPORT/PUBLIC ASSESMBLY FACILITY
645	E	EXEMPT HRA VACANT LAND
646	E	EXEMPT CULTURAL & NATURE EXHIBITS
647	E	EXEMPT COMMUNITY RECREATIONAL FAC.
649	E	LIBRARY
650	E	EXEMPT PROPERTY OWNED BY BD OF EDUC.
660	E	EXEMPT PROPERTY OWNED PARK DIST.PUB
670	E	EXEMPT PROP OWNED BY PRIVATE SCHOOLS
680	E	CHARITABLE EXEMPT HOSP./NURSG HOMES
682	E	WELFARE/CHARITABLE FACILITY
685	E	CHURCHES ETC PUBLIC WORSHIP PRIVE
687	E	EXEMPT OFFICE BUILDINGS
690	E	GRAVEYARDS MONUMENTS, CEMETERIES
830	U	COMM LD&IMPRO OWN BY PUB UTI TH RAIL
840	U	RAILROAD REAL PROP USED IN OPERATION

850	U	RAILROAD REAL PROP NOT USED OPERATOS
880	U	PUB UTIL PER PROP OTHER THAN RRS

OTHERS NOT PRESENT IN STUDY SITES

101	A	CASH GRAIN OR GENERAL FARM
108	A	NURSERIES
120	A	TIMBER OR FOREST LANDS
190	A	OTHER AGRICULTURAL USE
370	I	SMALL SHOPS (MACHINE,TOOL & DIE ETC)
407	M	FRATERNITY/SORORITY HOUSE
427	C	REGIONAL SHOPPING CENTER
440	C	DRY CLEANING PLANTS & LAUNDRIES
445	C	SAVINGS AND LOAN
457	C	PARKING RAMP
461	C	DRIVE-IN THEATERS
462	C	GOLF DRIV RANGE & MINI GOLF COURSES
485	C	RESEARCH AND DEVELOPMENT FACILITY
490	C	MARINE SERVICE FACILITIES
496	C	MARINA (SMALL BOAT)
509	R	RESID, NEW PLATS INITIAL YEAR
519	R	RESID, NEW PLATS INITIAL + 1/3 DIFF
529	R	RESID, NEW PLATS INITIAL + 2/3 DIFF
539	R	RESID, NEW PLATS THIRD YEAR
549	R	RESID, PROBLEM VACANT PINS
559	R	RESIDENTIAL, CONDO/CO-OP VACANT
560	R	MOBILE HOME ON REAL ESTATE
598	E	EXEMPT OTHER RESIDENTIAL
599	R	OTHER RESIDENTIAL
625	E	ST PAUL AIRPORT & MAC PROPERTY
695	E	SKYWAYS
821	U	VACANT LAND

Appendix 5B: I-35W Corridor Coalition Land-Use Codes

AG	Agriculture
SFD	Single-Family Detached
SFA	Single-Family Attached
MF	Multiple-Family
MH	Mobile Home Park
R	Rural Residential
NC	Neighborhood Commercial
CC	Community Commercial
RC	Regional Commercial
HC	Highway Convenience
OFC	Office
MU	Mixed Use
LI	Light Industrial
HI	Heavy Industrial
IN	Institutional
SRO	Outdoor Sport/Recreation
SRI	Indoor Sport/Recreation
OSN	Natural Open Space
OSP	Passive Open Space
ROW	Right of Way
RR	Railway
UTL	Utility
VAC	Vacant
WAT	Water
AP	Airport

Appendix 5C: Determining Land Uses for Context Areas and Buffers

By Jason Zimmerman

Our land-use analysis was the result of the merging of two separate measurement areas. Outside of our Context Areas, land-use data was based on the categories created as part of the Metropolitan Council's 2000 Land Use data set. Within the Context Areas, a more detailed set of steps was carried out to achieve the most accurate data possible.

Process for determining Land Uses for parcels in our Context Areas developed as follows:
1. Began with I-35W land-use codes for parcels in coalition cities
2. Added parcel data from other Ramsey County cities
3. Translated Ramsey County land-use codes to I-35W land-use codes (used I-35W land-use definitions and best estimates as to comparable categories)
4. Some Vacant and Institutional land uses were really Park and Recreation land uses. Using a combination of a St. Paul parks layer, an online Ramsey County park map, and county orthophotos, large vacant areas were checked to see if they should be recoded as Park and Recreation. They were recoded if other sources indicated they were parks—vacant parcels that were not clearly identified elsewhere as parks were left vacant.
5. Manually corrected some Ramsey County parcels with no codes (e.g., land surrounding townhomes, typically common areas, re-coded from Vacant to Single-Family Attached)
6. Manually corrected small slivers created through merging and clipping of parcel data sets (coded to match surrounding land uses)
7. Added Hennepin, Dakota, and other Anoka county parcels and land uses field-checked by staff
8. Railroad parcels inconsistently classified across cities and counties. In some places they were coded Exempt Land, others Operational RR, and still others Non-Operational RR. We did not change this. As a result, some railroad parcels are coded Tax Exempt and others Industrial.
9. When in doubt, used Metropolitan Council 2000 land-use data, online property tax information, and orthophotos to check questionable parcels
10. Overlaid Ramsey County water layer (most detailed available)
11. Outside Ramsey County, overlaid Met Council water layer (more general)
12. Parcel remnants around ponds were reassigned to open space categories—some had been vacant (see below, image 1 has the parcel surrounding the pond coded Vacant; image 2 is the orthophoto that shows the area is really open space; image 3 is the corrected data with the parcel coded as open space)

Process for determining Land Uses for parcels outside our Context Areas (Buffers) developed as follows:

1. Had to assign Met Council land-use data from polygons to Regional parcel data using centroids and a spatial join—this meant we had to correct by hand when parcel centroids did not fall within land-use polygons

   The image below demonstrates this point. The colored areas are land-use polygons; the thin grey lines show parcel boundaries; and the points are parcel centroids. Odd-shaped parcels, such as the one outlined in purple, have centroids that may not fall within the correct land-use polygon. Here, the parcel centroid (highlighted in purple) falls inside the lake. Assigning the underlying land use to the centroid, and thus to the parcel, would result in the large open space parcel being coded as Water. Instead, the centroid was manually coded as Park and Recreation.

2. Converted general Met Council land-use categories to our seven major land-use categories

Met Council Code	Met Council Land Use	RWJF Land Use
100	Agriculture	OTHER
111	Farmsteads	RESIDENTIAL
112,113,116	Single-Family Residential	RESIDENTIAL
114,115	Multifamily Residential	RESIDENTIAL
120	Commercial - Retail	COMMERCIAL
130	Commercial - Office	OFFICE
141	Mixed Use -	OFFICE

142	Residential Mixed Use - Industrial	INDUSTRIAL
143	Mixed Use - Commercial	COMMERCIAL
151	Industrial and Utility	INDUSTRIAL
153	Extractive	INDUSTRIAL
160	Institutional	TAX EXEMPT
170	Park, Rec, and Preserve	PARK AND RECREATION
173	Golf Course	PARK AND RECREATION
201	Major Highway	OTHER
202	Railway	INDUSTRIAL
203	Airport	OTHER
210	Undeveloped	VACANT
220	Water	WATER

Appendix 5D: SIC/NAICS Coding and the ESRI Business Analyst infoUSA Database

By Edward D'Sousa

This appendix documents the SIC/NAICS codes that were used for the Twin Cities Walking Study and highlights some of the quality and usability issues of ESRI's Business Analyst infoUSA database.

The SIC/NAICS codes used in our analysis were based on a best match of the NQLS Survey categories to the 4 digit Industry SIC Description and refined using the 6 digit extended codes where necessary. While the SIC 4 digit Industry code is standardized, the 6 digit extended codes appear to vary between data providers and are not defined as part of the national standard. To complicate matters, the transition from SIC to NAICS, and how the SIC codes are mapped to the new NAICS codes is not always clear, and again, may vary between the various data providers. It is important to note that private data providers (such as infoUSA and Dun & Bradstreet) databases are not based on the classifications assigned to businesses by the US Census Bureau.

Our use of the infoUSA database required we have a definitive listing of the 6 digit extended codes used by infoUSA. We also desired to know the mapping between the 6 digit SIC code and 8 digit NAICS codes in the database. The codes and descriptions were not part of the delivered documentation with Business Analyst, but were provided by the ESRI Business Analyst support group. Currently the Codes and Descriptions can be found at
http://forums.esri.com/Thread.asp?c=154&f=1667&t=185984#550618

Assignment of businesses/institutions to SIC/NAICS Categories

As noted above, both the SIC and NAICS are standardized (4 digits for SIC and 6 digits for NAICS). The North American Industry Classification System (NAICS) is the current classification system used by the US Office of Budget and Management (OMB) for statistical reporting and is the official business classification standard for all US Federal agencies. NAICS replaced the Standard Industrial Classification (SIC) system in 1997, partially as a response to North American Free Trade Agreement and provides common classification system for businesses in Canada, Mexico and the United States.

Both the SIC and NAICS appear to have a large set of 4 and 6 digit codes (respectively) for industry classifications and the assignment of businesses to the 4 digit SIC/6 digit NAICS within the infoUSA database appears to be fairly reasonable. infoUSA (along with other data providers) has extended both SIC and NAICS with 2 (or more) additional digits which allows for greater resolution of different business types within a category. Unfortunately, these extended codes can differ significantly between the various private data providers and do not translate well when mapping SIC to NAICS. In many cases, these extended codes are not consistently applied, often do not reflect a businesses activity, and add little to the ability to isolate focused business activities. Understanding how businesses are assigned their SIC/NAISC codes is a key issue. For higher quality data requirements, the list may need to be cross referenced against other sources, a possibly expensive and time consuming task. The cross referencing may only provide a marginal increase in data quality, but for certain categories, this may be desirable.

As an example, in part of the Twin Cities we examined, infoUSA data for Food Services (SIC 5812 - NAICS 7221/7222/7223) has nearly 85% of the Food Services restaurants (SIC 581208) as Limited Service Restaurants (NAICS 72221105). This appears to be a catch all that includes businesses that would appear to fit better in other categories. SIC 581208 includes pizza places, cafés, full service restaurants, McDonald's/Burger King, etc. While it is a reasonable to extract all restaurants in a study area from the infoUSA database using SIC 5812, it is not possible to isolate fast-food from the full service, sit down restaurants. Additionally, some of the mappings between SIC and NAICS seem unreasonable (Table 5D-1), such as Cafes (SIC 581214) mapped to Cafeterias (NAICS 72221202) or Box Lunches (SIC 581215) to Full Service Restaurants (NAICS 72211002).

Table 5D-1: Restaurant and Food Services SIC/NAICS Classification mappings and feature counts for the Study Area

SIC Code	SIC Description	NAICS Code	NAICS Description	Count
581203	ICE CREAM PARLORS	72221313	SNACK & NONALCOHOLIC BEVERAGE BARS	142
581206	FOODS-CARRY OUT	72221104	LIMITED-SERVICE RESTAURANTS	21
581208	RESTAURANTS	72221105	LIMITED-SERVICE RESTAURANTS	3254
581209	DELICATESSENS	72221103	LIMITED-SERVICE RESTAURANTS	57
581212	CATERERS	72232001	CATERERS	158
581214	CAFES	72221202	CAFETERIAS	20
581215	BOX LUNCHES	72211002	FULL-SERVICE RESTAURANTS	5
581219	SANDWICHES	72211020	FULL-SERVICE RESTAURANTS	7
581220	RESTAURANT MANAGEMENT	72231009	FOOD SVC CONTRS	10
581222	PIZZA	72211016	FULL-SERVICE RESTAURANTS	114
581223	BANQUET ROOMS	72231001	FOOD SVC CONTRS	46
581224	BARBECUE	72221101	LIMITED-SERVICE RESTAURANTS	3
581225	BEVERAGES-NON-ALCOHOLIC-RETAIL	72221320	SNACK & NONALCOHOLIC BEVERAGE BARS	3
581226	THEATRES-DINNER	71111006	THEATER COMPANIES & DINNER THEATERS	1
581227	ITALIAN FOOD PRODUCTS	72211011	FULL-SERVICE RESTAURANTS	1
581229	DELI-BAKERY	72221102	LIMITED-SERVICE RESTAURANTS	4
581230	RESTAURANTS-FOODS DELIVERY	49221004	LOCAL MESSENGERS & LOCAL DELIVERY	1
581231	FOOD SERVICE-INDUSTRIAL	72231004	FOOD SVC CONTRS	1

[Source: Mapping of SIC to NAICS - ESRI]

Data Quality

For the seven counties in the Twin Cities Metro area, 93,695 records exist in the InfoUSA database. The following database description is from the Business Analyst InfoUSA documentation, and demonstrates that only 90 percent of businesses are coded at the street address level:

> **Description:** InfoUSA collects information on approximately 11 million private and public US companies. Individual businesses are located by address geocoding—not all will have an exact location. Businesses can be retrieved by their Standard Industrial Classification Code (SIC) as well as by North American Industry Classification System (NAICS) Code and Location. The infoUSA Business File can be used for locating both competitors and marketing opportunities. This data is current as of January 2005.
> **Enhancements for 2005:**
> With the 2005 infoUSA business listings data, a collection of geocoding processes were incorporated into the process to help improve the overall geocoding match rate. This included several passes through different geocoders. Overall, 90 percent of the businesses are coded at the address level; 99.8 percent are assigned to a census block group.

Geo-coding and Positional Accuracy

As stated in the documentation, the business listings are processed through several geocoders, with 90 percent of the data geocoded at the address level with the majority of the remaining businesses located at the census block group. Each feature is coded with its geocoding status, so those features located at the address level can be separated from the other, "less accurate"

locations. While the geocoding status can help indicate the level of positional accuracy the feature was processed to, it does not completely describe how well individual, address level, features are placed. In general, business located in strip malls and other complexes that cover large areas may be off by several hundred meters. Businesses in individual buildings along public roads are generally located within 10 meters or so.

Completeness

For most categories, the data was assumed to be accurate and complete enough for the analysis purpose. For categories, such as libraries, where very few locations existed in the database (less than 100), the analysis could be significantly affected if one or two key locations were missing. For categories where a relatively large number of locations exist (greater than 100) the list was reviewed to see if it was reasonable. For categories with greater than 500 locations, the list was assumed to be sufficient without any review.

Several examples can help illustrate the types of errors and the approach used to verify the data produced reasonable analysis results.

Ninety libraries were extracted from the infoUSA database for the seven counties in the Twin Cities Metro area. The library locations were verified against a listing from MELSA (an alliance of public libraries in the Twin Cities). Overall, 13 percent of the libraries were missing from the infoUSA database and 26 percent of the infoUSA listings were deleted for a variety of reasons (duplicates, incorrect listings, incorrect address). In the City of St. Paul, which covered the heart of the study area, 5 of the 12 library locations (>40 percent) were not in the infoUSA database. The missing libraries were geocoded and combined with the other, verified infoUSA libraries.

While libraries represent a critical neighborhood asset that people may walk to, post office locations were deemed to be less critical, as many of the functions that can be performed at a post office can also be performed elsewhere. For example, stamps can be purchased from ATMs and supermarkets, mail delivery people collect letters for delivery from residences, and private businesses also specialize in mailing and shipping services. In all, 142 post office locations were extracted for the metro area. Three post offices were removed from the Business Analyst dataset, as they were deemed not to be actual post offices.
- HITCHING POST (2 occurrences)
- POST FACTO INC

The post offices locations were not checked against an authoritative source (i.e. the USPS list of locations).

Additionally, a very small set of businesses are listed more than once in the database, often within the same category. Usually the business name is spelled differently (Di Napoli Café <u>vs</u> Dinapoli Café), and/or the address is slightly different (814 Main St. <u>vs.</u> 816 Main St.). While this is a relatively rare occurrence, it does represent a source of noise in the database.

ESRI Business Analyst

It is important to note the infoUSA dataset shipped with Business Analyst is stored in a non-standard, proprietary data file that can only be accessed if a Business Analyst license is available. The point locations for business can be exported to other formats (shapefile, Geodatabase) but the business address will not be accessible. Only the business's name, city, state/zip, SIC/NAICS, sales volume, number of employees and geocoding status are stored in

tables accessible without a Business Analyst license. Address information is stored in an encrypted data file that can only be accessed using Business Analyst tools. Additionally, only one address can be displayed at a time. For our study, the location and business classification were the primary concern. If we had required the businesses addresses, a different data source would probably have been chosen.

NQLS Based SIC Code Listing for the Twin Cities Walking Study

The following SIC codes were used for the analysis performed in Protocol "5.14 Distance to Nearest Facility for Comparison with NQLS Survey". This list represents 1) the SIC categories that best matched the NQLS category and 2) SIC codes that are represented within the study area. This is not an exhaustive listing of all SIC codes that could possibly be matched to the NQLS categories, but rather represents only the features found in the Twin Cities Metro. Other geographic areas may have additional features that are relevant, but were not present in our region.

Category Type	NQLS Category	SIC Code	SIC Description
Institution	Post office	431101	POST OFFICES
	Library	823106	LIBRARIES-PUBLIC
	Elementary school WON'T WORK	821103	SCHOOLS
		821105	MONTESSORI SCHOOLS
		821107	SCHOOLS WITH SPECIAL ACADEMIC EDUCATION
		821114	HOME SCHOOLING
	Other schools	821108	SCHOOLS-ADULT EDUCATION OMIT--NOTE SOMEWHERE THAT IT IS OMITTED
		822101	SCHOOLS-UNIVERSITIES & COLLEGES ACADEMIC
		822108	SCHOOLS-MEDICAL
		822109	DENTAL SCHOOLS
		822298	JUNIOR COLLEGES & TECHNICAL INSTITUTES

Category Type	NQLS Category	SIC Code	SIC Description
Food	Convenience/small grocery store	541103	CONVENIENCE STORES
	Supermarket	541101	FOOD MARKETS
		541104	FOOD PRODUCTS-RETAIL
		541105	GROCERS-RETAIL
	Fruit/vegetable market	543101	FRUITS & VEGETABLES & PRODUCE-RETAIL
	Bakery	546101	BAGELS
		546102	BAKERS-RETAIL
		546105	DOUGHNUTS
	Coffee place	549915	COFFEE & TEA
		549938	ESPRESSO & ESPRESSO BARS
		581228	COFFEE SHOPS
	Restaurant	581203	ICE CREAM PARLORS

Note: there does not appear to be any reliable or reasonable way to distinguish between full service, Limited service or fast food restaurants in the infoUSA dataset

	SIC Code	SIC Description
	581206	FOODS-CARRY OUT
	581208	RESTAURANTS
	581209	DELICATESSENS
	581214	CAFES
	581215	BOX LUNCHES
	581219	SANDWICHES
	581222	PIZZA
	581224	BARBECUE
	581226	THEATRES-DINNER
	581227	ITALIAN FOOD PRODUCTS
	581229	DELI-BAKERY

Category Type	NQLS Category	SIC Code	SIC Description
Service	Laundry/dry cleaners	721101	LAUNDRIES
		721201	CLEANERS
		721202	VALET SERVICE
		721501	LAUNDRIES-SELF SERVICE
	Bank	602101	BANKS
		602201	STATE COMMERCIAL BANKS
		603501	SAVINGS & LOAN ASSOCIATIONS
		603698	SAVINGS INSTITUTIONS-EXCEPT FEDERAL
		606101	CREDIT UNIONS
		606102	FEDERALLY CHARTERED CREDIT UNIONS
	Salon/barber shop	723101	SKIN TREATMENTS
		723102	MANICURING
		723119	SPAS-BEAUTY & DAY
		724101	BARBERS

Category Type	NQLS Category	SIC Code	SIC Description
Retail	Clothing store	561101	MEN'S CLOTHING & FURNISHINGS-RETAIL
		561106	JEANS
		561107	LEATHER CLOTHING-RETAIL
		561108	NECKWEAR-RETAIL
		561109	WORK CLOTHING-RETAIL
		562101	WOMEN'S APPAREL-RETAIL
		562103	MATERNITY APPAREL
		562104	BRIDAL SHOPS
		562105	BOUTIQUE ITEMS-RETAIL
		563202	FUR BUSINESS-RETAIL
		563205	BRASSIERES
		563206	HOSIERY-RETAIL
		563207	HANDBAGS
		563208	KNIT GOODS-RETAIL
		563210	LINGERIE

		563211	WOMEN'S SPECIALTY SHOPS
		563215	ACCESSORIES-FASHION
		564101	BABY ACCESSORIES
		564103	CHILDRENS & INFANTS WEAR-RETAIL
		565101	CLOTHING-RETAIL
		569901	FASHION DESIGNERS
		569904	DANCING SUPPLIES
		569905	COSTUMES-MASQUERADE & THEATRICAL
		569906	DRESSMAKERS
		569907	WESTERN APPAREL
		569908	RAINWEAR
		569911	RIDING APPAREL & EQUIPMENT
		569913	SPORTSWEAR-RETAIL
		569915	SWIMWEAR & ACCESSORIES-RETAIL
		569917	T-SHIRTS-RETAIL
		569919	TAILORS
		569921	UNIFORM SUPPLY SERVICE
		569922	UNIFORMS
		569934	CLOTHING-CUSTOM MADE
		569947	APPAREL & GARMENTS-RETAIL
	Book store	593201	BOOK DEALERS-USED & RARE
		594201	BOOK DEALERS-RETAIL
		594202	BIBLES
		594203	LAW BOOKS
		594205	COMIC BOOKS
	Video store	573501	RECORDS TAPES & COMPACT DISCS-RETAIL
		573502	VIDEO TAPES DISCS & CASSETTES
		784102	VIDEO TAPES & DISCS-RENTING & LEASING
	Hardware store	525101	SAWS
		525102	TOOLS-HAND
		525104	HARDWARE-RETAIL
		525115	TOOLS-NEW & USED
	Pharmacy/drug store	591202	HEALTH CARE PRODUCTS
		591203	FIRST AID SUPPLIES
		591205	PHARMACIES

Category Type	NQLS Category	SIC Code	SIC Description
Recreation/Fitness	Recreation center	799701	RECREATION CENTERS
	Gym or fitness facility	799101	HEALTH CLUBS STUDIOS & GYMNASIUMS
		799102	GYMNASIUMS
		799105	SPAS (HEALTH)
		799106	PERSONAL TRAINERS-FITNESS
		799107	PILATES

REFERENCES

Alexander, Ernest. 1993. Density measures: A review and analysis. *Journal of Architectural and Planning Research* 10, 3: 181-202.

Andersen, R. E., T. A. Wadden, S. J. Bartlett, B. Zemel, T. J. Verde, and S. C. Franckowiak. 1999. Effects of lifestyle activity vs. structured aerobic exercise in obese women: a randomized trial. *Journal of the American Medical Association.* 281, 4:335-40, Jan. 27.

Audirac, Ivonne. 1999. Stated preference for pedestrian proximity. *Journal of Planning Education and Research.* 19: 53-66.

Ball, Kylie, Adrian Bauman, Eva Leslie, and Neville Owen. 2001. Perceived environmental aesthetics and convenience and company are associated with walking for exercise among Australian adults. *Preventive Medicine* 33: 434-440.

Boarnet, M. G., and S. Sarmiento. 1998. Can land-use policy really affect travel behavior? A study of the link between non-work travel and land-use characteristics. *Urban Studies* 35(7):1155-1169.

Boarnet, Marlon, and Kristen Day. 2004a. Assessing built environment features linked to physical activity. Presentation at Active Living Research Annual Conference, San Diego, January.

Cervero, R., and R. Gorham. 1995. Commuting in transit versus automobile neighborhoods. J*ournal of the American Planning Association*; 61:210-225.

Cervero, Robert, and Michael Duncan. 2003. Walking, bicycling, and urban landscapes: Evidence from the San Francisco Bay area. *American Journal of Public Health* 93(9): 1478-1483.

Cervero, Robert, and Kara Kockelman. 1996. Travel demand and the 3 Ds: Density, diversity, and design. Working Paper 674. Berkeley: UC Institute of Urban and Regional Development.

_____. 1997. Travel demand and the 3 Ds: Density, diversity, and design. *Transportation Research*; D, 2, 3: 199-219.

Churchman, Arza. 1999. Disentangling the concept of density. *J Plan Lit* 13, 4: 389-411.

Crane, Randall. 2000. The influence of urban form on travel: An interpretive review. *Journal of Planning Literature* 15, 1: 3-23.

Crane, Randall, and Richard Crepeau. 1998. Does neighborhood design influence travel? Behavioral analysis of travel diary and GIS data. UCTC Working Paper. Berkeley: University of California Transportation Center.

Day, Kristen, and Mariela Alfonzo. 2004. Measuring urban design features related to walking and bicycling. Presentation at the Assoc. of Collegiate Schools of Planning Conf., Portland, October.

Day, Kristen, Marlon Boarnet, Mariela Alfonzo, Ann Forsyth. 2005. *The Irvine-Minnesota Urban Design Inventory* (inventory, training package and instructions). Irvine: University of California Irvine.

Day, Kristen, Marlon Boarnet, Mariela Alfonzo, Ann Forsyth, 2006. The Irvine Minnesota Inventory to Measure Built Environments: Development. *American Journal of Preventive Medicine* 30, 2: 144-152.

Day, Kristen,, Marlon Boarnet, Mariela Alfonzo, Ann Forsyth, J. Michael Oakes, 2006. The Irvine Minnesota Inventory to Measure Built Environments: Reliability Tests. *American Journal of Preventive Medicine* 30, 2: 153-259.

Dill, Jennifer. 2003. Measuring network connectivity for bicycling and walking. July 17 draft. Presented at the Joint Cong. of ACSP-AESOP, Leuven, Belgium.

Duncan, O. D., and B. Duncan. 1955. A methodological analysis of segregation indexes. *American Sociological Review* 20: 210-217.

Dunphy and Fisher. 1996. [Citation to add.]

Ewing, Reid, Tom Schmid, Richard Killingsworth, Amy Zlot, and Stephen Raudenbush. 2003. Relationship between urban sprawl and physical activity, obesity, and morbidity. *American Journal of Health Promotion* 18, 1: 47-57.

Forsyth, Ann. 2000. Analyzing public space at a metropolitan scale: Notes on the potential for using GIS. *Urban Geography* 21, 2: 121-147.

_____. 2004a. *Community Design: A Toolkit for Building Physical Activity into Daily Life.* Minneapolis: Design Center.

_____. Forthcoming. Density. *Encyclopedia of the City.* Edited by Roger Caves. New York: Routledge.

_____. 2003. *Measuring Density: Working Definitions for Residential Density and Building Intensity.* Design Brief 9. Minneapolis: Design Center.

Forsyth, Ann, Michael Oakes, and Kathryn H. Schmitz. 2004a. Observational measures of the physical environment: Testing and modifying the Boarnet-Day Inventory. Revised after presentation at the Assoc. of Collegiate Schools of Planning Conf., November.

Forsyth, Ann, Kathryn H. Schmitz, Michael Oakes, Jason Zimmerman, and Joel Koepp. 2006. Standards for Environmental Measurement using GIS: Toward a Protocol for Protocols. *Journal of Physical Activity and Health..*

Frank, Lawrence D., and Gary Pivo. 1994. Impacts of mixed use and density utilization of three modes of travel: Single-occupant vehicle, transit, and walking. *Transportation Research Record* 1466:44-52.

Frank, Lawrence D., and P. O. Engelke. 2001. The built environment and human activity patterns: Exploring the impacts of urban form on public health. *Journal of Planning Literature*;16, 2: 202-218.

Giles-Corti, Billie, and Robert J. Donovan. 2003. Relative influences of individual, social environmental, and physical environmental correlates of walking. *American Journal of Public Health* 93, 9: 1583-1589.

Grant, Jill. 2002. Mixed use in theory and practice. *Journal of the American Planning Association* 68, 1: 71-84.

Handy, Susan. 1996a. Methodologies for exploring the link between urban form and travel behavior. *Transportation Research Record* 1, 2:151-165.

_____. 1992. Regional versus local accessibility: Implications for nonwork travel. *Transportation Research Record* 1400:58-66.

_____. 1996c. Understanding the link between urban form and nonwork travel behavior. *Journal of Planning Education and Research* 15, 3: 183-198.

_____. 1996b. Urban form and pedestrian choices: Study of Austin neighborhoods. *Transportation Research Record* 1552:135-144.

_____. 2003. Critical assessment of the literature on the relationships among transportation, land use, and physical activity. Paper prepared for the TRB and the Institute of Medicine Committee on Physical Activity, Health, Transportation, and Land Use; Nov. 24. Draft.

Handy, Susan, Marlon G. Boarnet, Reid Ewing, and Richard E. Killingsworth. 2002. How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine* 23, 2S: 64-79.

Hess, Paul M., Anne Vernez Moudon, Mary Catherine Snyder, and Kiril Stanilov. 1999. Neighborhood site design and pedestrian travel. *Transportation Research Record,* 1674:9-19.

Hoehner, Christine M., Laura K. Brennan, Ross C. Brownson, Susan L. Handy, and Richard Killingsworth. 2003. Opportunities for integrating public health and urban planning approaches to promote active community environments. *American Journal of Health Promotion* 18(1): 14-20.

Humpel, Nancy, Neville Owen, and Eva Leslie. 2002. Environmental factors associated with adults' participation in physical activity. *American Journal of Preventive Medicine* 22, 3: 188-199.

Krizek, Kevin. 2003a. Residential relocation and changes in urban travel: Does neighborhood-scale urban form matter? *Journal of the American Planning Association* 69 ,3: 265-281.

_____. 2000. Pretest-posttest strategy for researching neighborhood scale urban form and travel behavior. *Transportation Research Record*, 1722:48-55.

_____. 2003b. Operationalizing neighborhood accessibility for land use-travel behavior research and regional modeling. *Journal of Planning Education and Research* 22, 3: 270-287.

Lund, Hollie. 2003. Testing the claims of new urbanism: Local access, pedestrian travel, and neighboring behaviors. *Journal of the American Planning Association* 69, 4: 414-429.

Massey, D. S., and N. A. Denton. 1988. The dimension of residential segregation. *Soc Forces* 67: 281-315.

McCann, Barbara A., and Reid Ewing. 2003. Measuring the health effects of sprawl: A national analysis of physical activity, obesity and chronic disease. Smart Growth America Surface Transportation Policy Project.

McNally, Michael G. and Anup Kulkarni. 1997. Assessment of influence of land-use transportation system on travel behavior. *Transportation Research Record* 1607: 105-115.

McNally and Ryan. 1993. [Reference to add.]

Messenger and Ewing. 1996. [Reference to add.]

Moudon, Anne Vernez, and Chanam Lee. 2003. Walking and bicycling: An evaluation of environmental audit instruments. *American Journal of Health Promotion* 18(1): 21-37.

Moudon, Anne Vernez, and P. Hess. 2000. Suburban clusters: The nucleation of multifamily housing in suburban areas of the central Puget Sound. *Journal of the American Planning Association* 66, 3: 243-264.

Moudon, Anne Vernez, P. Hess, M. Snyder, and K. Stanilov. 1997. Effects of site design on pedestrian travel in mixed use, medium density environments. Report WA-RD 431.1. Olympia: Washington State Department of Transportation.

Moudon, Anne Vernez, Chanham Lee, Allen Cheadle, Cheza Collier, Donna Johnson, Robert Weathers, Jean-Ives Courbois, and Phil Hurvitz. 2004. Walkable and bikable communities project: A draft report on findings from the walking models. Seattle: University of Washington Urban Form Lab.

Oakes, J. M. 2004. The (mis)estimation of neighborhood effects: Causal inference for a practicable social epidemiology. *Social Science and Medicine* 58:1929-52 (with discussion and rejoinder).

Oakes, J. M. and P. H. Rossi. 2003. The measurement of SES in health research: current practice and steps toward a new approach. *Social Science and Medicine* 56:769-84.

Roman, M. 2001. Users' handbook version 1.0. Ramsey County GIS Users' Group.

Saelens, Brian, James Sallis and Lawrence Frank. 2003. Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine* 25, 2: 80-91.

Sallis, J. F., A. Bauman, and M. Pratt. 1998. Environmental and policy interventions to promote physical activity. *American Journal of Preventive Medicine,*15 4:379-397.

Sallis, J. F., M. Hovell, C. Hofstetter, J. Elder, M. Hackley, C. Caspersen, and K. Powell. 1990. Distance between homes and exercise facilities related to frequency of exercise among San Diego residents. *Public Health Report* 105:179-185.

Sallis, J. F., M.F. Johnson, K. J. Calfas, S. Caparosa, J. Nichols. 1997. Assessing perceived physical environment variables that may influence physical activity. *Research Quarterly Exercise Sport* 68:345-351.

Schimek, Paul. 1996. Household motor vehicle ownership and use: How much does residential density matter? *Transportation Research Record* 1552: 120-125.

Song, Yan, and Gerrit-Jan Knaap. 2004. Measuring urban form. *Journal of the American Planning Association* 70, 2: 210-225.

Song, Yan, and Daniel A. Rodríguez. 2005b. Steps for computing the dissimilarity, entropy, Herfindahl-Hirschman and accessibility (gravity with competition) indices. Draft manuscript.

_____. 2005a. The measurement of the level of mixed land uses: A synthetic approach. Draft manuscript.

Southworth, M. 1997. Walkable communities? An evaluation of neotraditional communities on the urban edge. *Journal of the American Planning Association*, 63, 1: 28-44.

Steiner, R. L. 1994. Residential density and travel patterns: Review of the literature. *Transportation Research Record*, 1466:37-43.

Steiner, R., A. Bond, D. Miller, and P. Sand. *Future Directions for Multimodal Areawide Level of Service Handbook: Research and Development*. Florida Department of Transportation, Office of Systems Planning, Contract BC-345-78, 2004.

Takano, T., K. Nakamura, and M. Watanabe. 2002. Urban residential environments and senior citizens' longevity in megacity areas: The importance of walkable green spaces. *Journal of Epidemiology and Community Health* 56: 913-918.

Turner, M. G., R. H. Gardner, and R.V. O'Neill. 2001. *Landscape Ecology in Theory and Practice: Pattern and Process.* New York: Springer Verlag.

US Census Bureau. 2005. 2004 First Edition TIGER/Line® Technical Documentation. Washington, DC: US Census.

US Department of Transportation, Federal Highway Administration. 1999a. *Guidebook on Methods to Estimate Non-Motorized Travel: Overview of Methods*. FHWA-RD-98-165. McLean, VA: FHWA.

_____. 1999b. *Guidebook on Methods to Estimate Non-Motorized Travel: Supporting Documentation*. FHWA-RD-98-166. McLean, VA: FHWA.