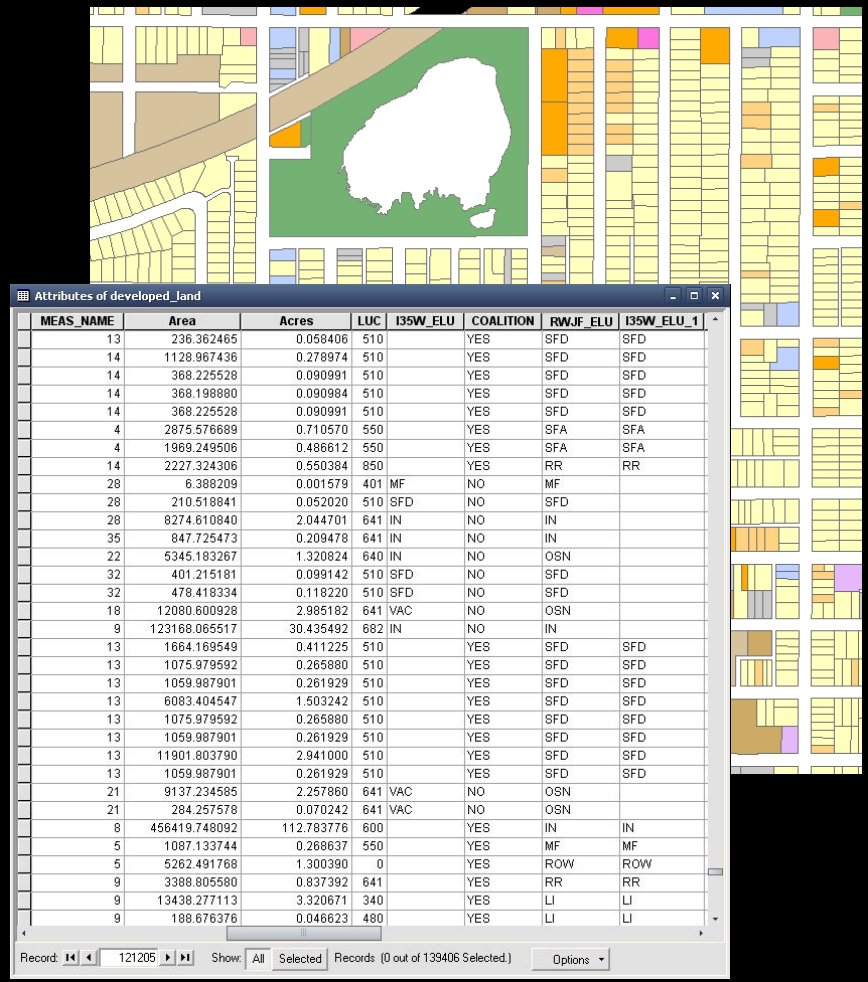
GIS for the Non-Expert

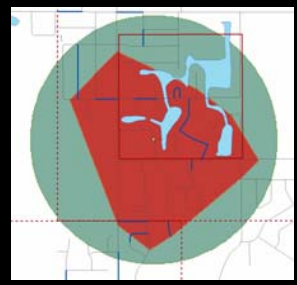
Ann Forsyth

University of Minnesota

February 2006



GIS for the Non-Expert



1. Definitions and problems
2. Measures being tested in Twin Cities Walking Study
   - Basic approach, data, variables and geographies
3. Protocols
   - Concept, formulae, approach, and steps
4. Typical measurement dilemmas
   - Intersections, small network buffers, counting trees
5. Analysis issues

GIS for the Non-Expert

- Who has done a GIS class?
- Who has used GIS in a project?
- Who has created new data in GIS?

1. Definitions and Problems
Definitions

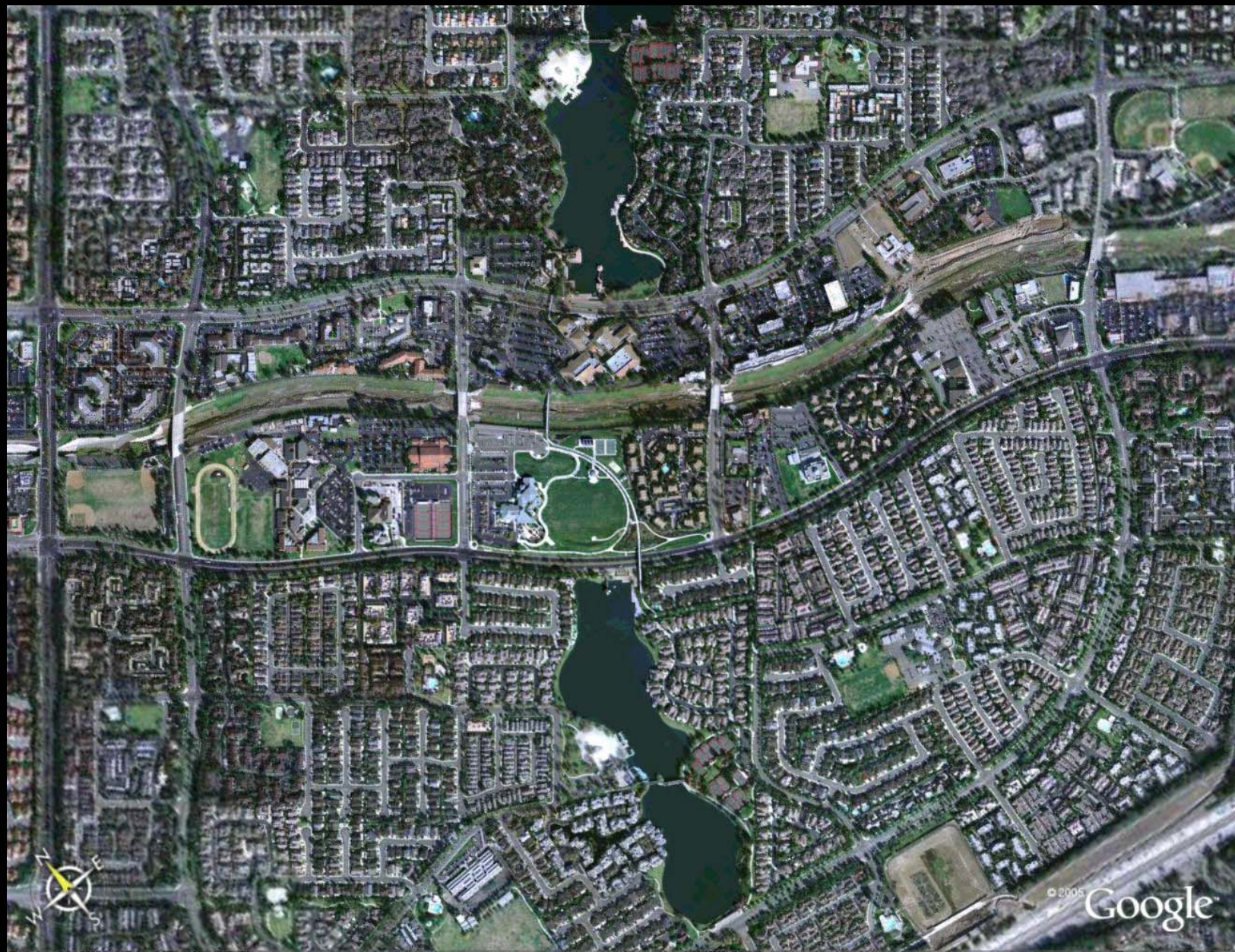
- A Geographic Information System (GIS) is a system of software and data that allows mapping and analysis
- ESRI is the main vendor in the US (ArcGIS suite including ArcMap, ArcCatalog, Spatial Analyst, ArcInfo, Network Analyst)
- Rapid increase in available software and data over the past 10 years

1. Definitions and Problems
Problems

- Selecting variables
- Data—existing vs. new
- Specific definitions of variables
- Documentation of data, variables, analysis
- Several GIS cultures—no perfect fit
  - Environmental/remote sensing (automation, large scale analyses)
  - Social analysis (inequality, access)
  - Transportation (facilities, modeling)
  - Land use/urban design (use, experience, aesthetics)

Irvine 7,000 feet (approx 1.5 miles across)

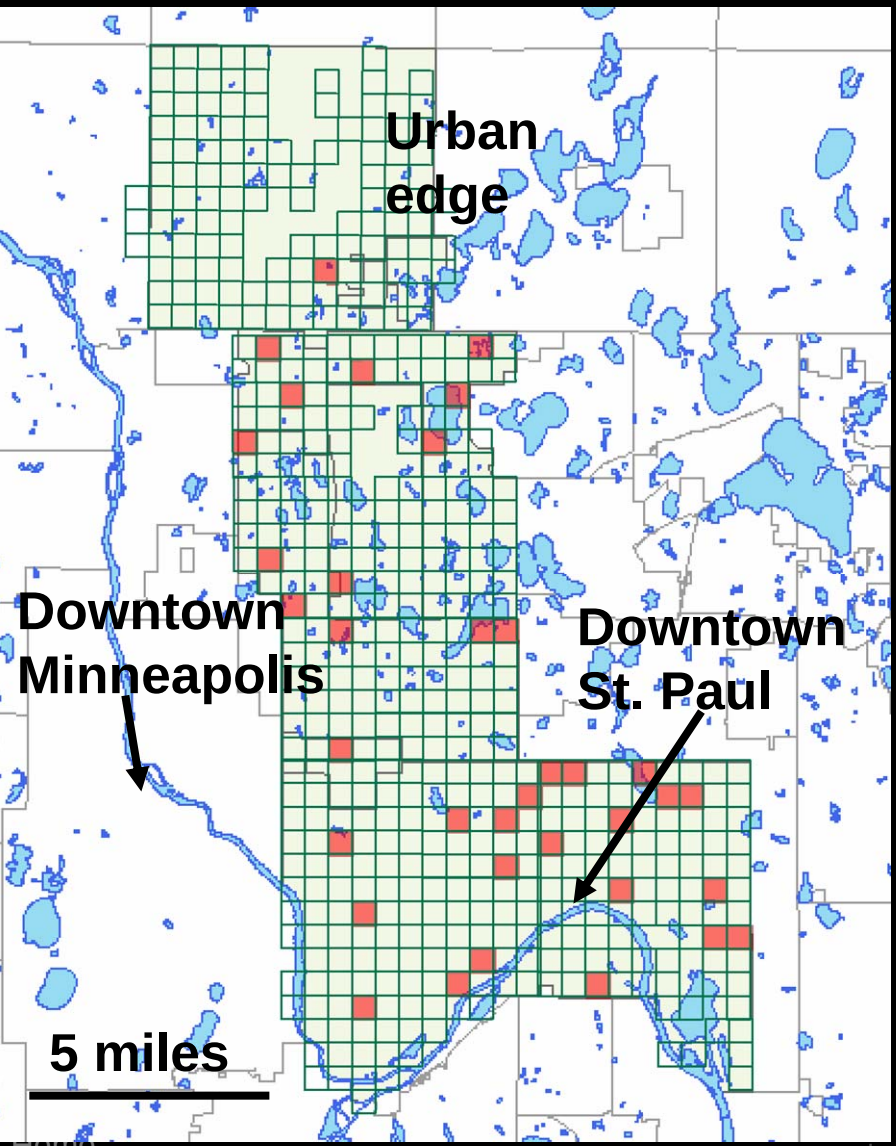
Environmental/remote sensing; Social analyses; Transportation; Land use/urban design



555 Hudson St, 7000 ft

Environmental/remote sensing; Social analyses; Transportation; Land use/urban design

2. Measures
Twin Cities Walking Study



- Ann Forsyth, Kathryn Schmitz, Michael Oakes
- Measuring how density, street pattern, mixed use, and pedestrian infrastructure affect walking vs. social and economic factors (many variables)
- 36 0.5*0.5 mile areas in corridor from St. Paul to Blaine—selected to be varied in gross density and median block size
- Environmental measures (GIS, observations); surveys, travel diaries, measured height and weight, and accelerometer data for 718 mostly randomly selected people during 2004—work in progress

2. Measures
Variables We Are Examining

- Everything plausible and not too expensive, based on review of literature
- Do we know which ones are best? After examining correlations between different categories of those plausible environmental variables and accelerometry, measured BMI, travel diaries, and self reported PA—it depends on reason for PA
- Environmental variables include:
  - Density measures
  - Street pattern/connectivity
  - Pedestrian infrastructure
  - Mixed use/destinations
  - Traffic
- Also comparing survey vs. measured environmental features at the individual level (e.g. distance to grocery store, park)

2. Measures

How to Think About Measures

Measurement of variables associated with the built environment occurs in several levels

- Dimension or topic e.g. density
- Measurement of variable associated with that dimension e.g. gross population density, employment density
- Geography of that measurement e.g. 200 meter street network distance/buffer, 100 meter grid cell
- There can be easily 50-100 variable/geography combinations for one dimension or topic e.g. eight different densities measured at 4-11 different geographies

- Tip: for the built environment to significantly increase physical activity it needs to be significantly different to current North American environments + supported culturally and economically

2. Measures
Data Sources

Data Sources

- Census e.g. block—consistent, national data
- GIS data —streets, parcels, building footprints—consistent at municipal and sometimes regional level
- GIS digital orthophotos—very high resolution for some areas only (all our study area)
- Fieldwork—only for study areas; at segment level

Issues

- Consistency from place to place
- Purpose of collection
- Geographies of collection and measurement
- Data resolution
- Accuracy
- Completeness
- Time
- Errors

More about data

Twin Cities Walking Study

Environment and Physical Activity: GIS Protocols

Version 3.1, February, 2006

Editor: Ann Forsyth

Contributors (alphabetically):

Edward D'Sousa
GIS Research Assistant, Metropolitan Design Center

Ann Forsyth, PhD
PI, RWJF ALR Grant, Metropolitan Design Center

Joel Koepp
GIS Research Fellow, Metropolitan Design Center

J. Michael Oakes, PhD
Co-PI, RWJF ALR Grant, School of Public Health

Kathryn H. Schmitz, PhD
Co-PI, RWJF ALR Grant, School of Public Health

Jason Zimmerman
GIS Research Fellow, Metropolitan Design Center

Extra material from:

Daniel Rodriguez and Yan Song
University of North Carolina

Twin Cities Walking Study, Metropolitan Design Center
1 Rapson Hall, 89 Church Street SE, Minneapolis, MN 55455-0109USA
http://www.designcenter.umn.edu Forsyth@umn.edu

Preparation of this manual was assisted by grants from the Robert Wood Johnson Foundation Active Living Research Program and The University of Minnesota Center for Transportation Studies

This is a work in progress

On Design Center web site at www.designcenter.umn.edu

2. Measures

Fundamental Measures

Area

2.1 Net Land Area (without water)

2.2 Gross Area (including water)

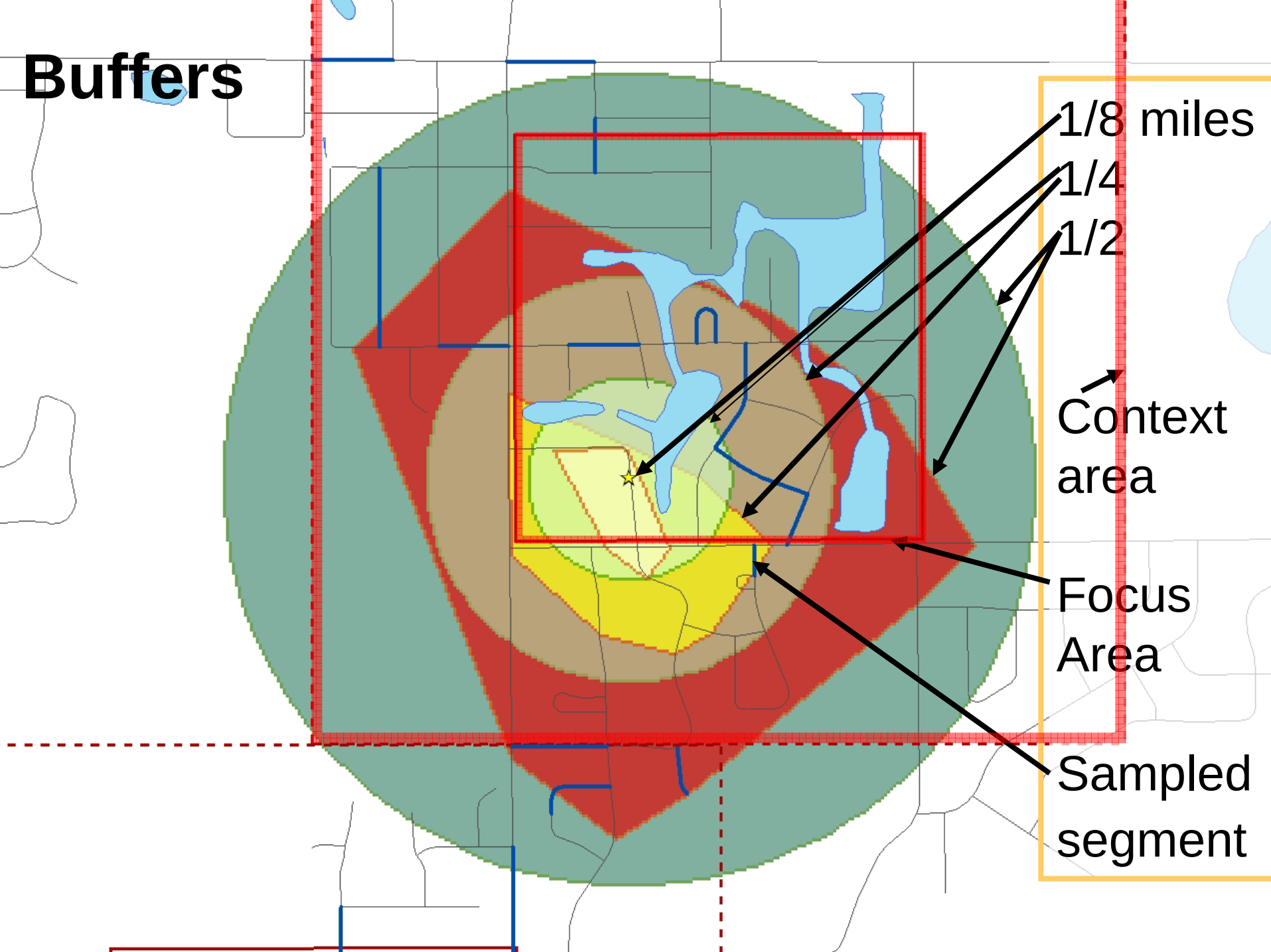
Geographies

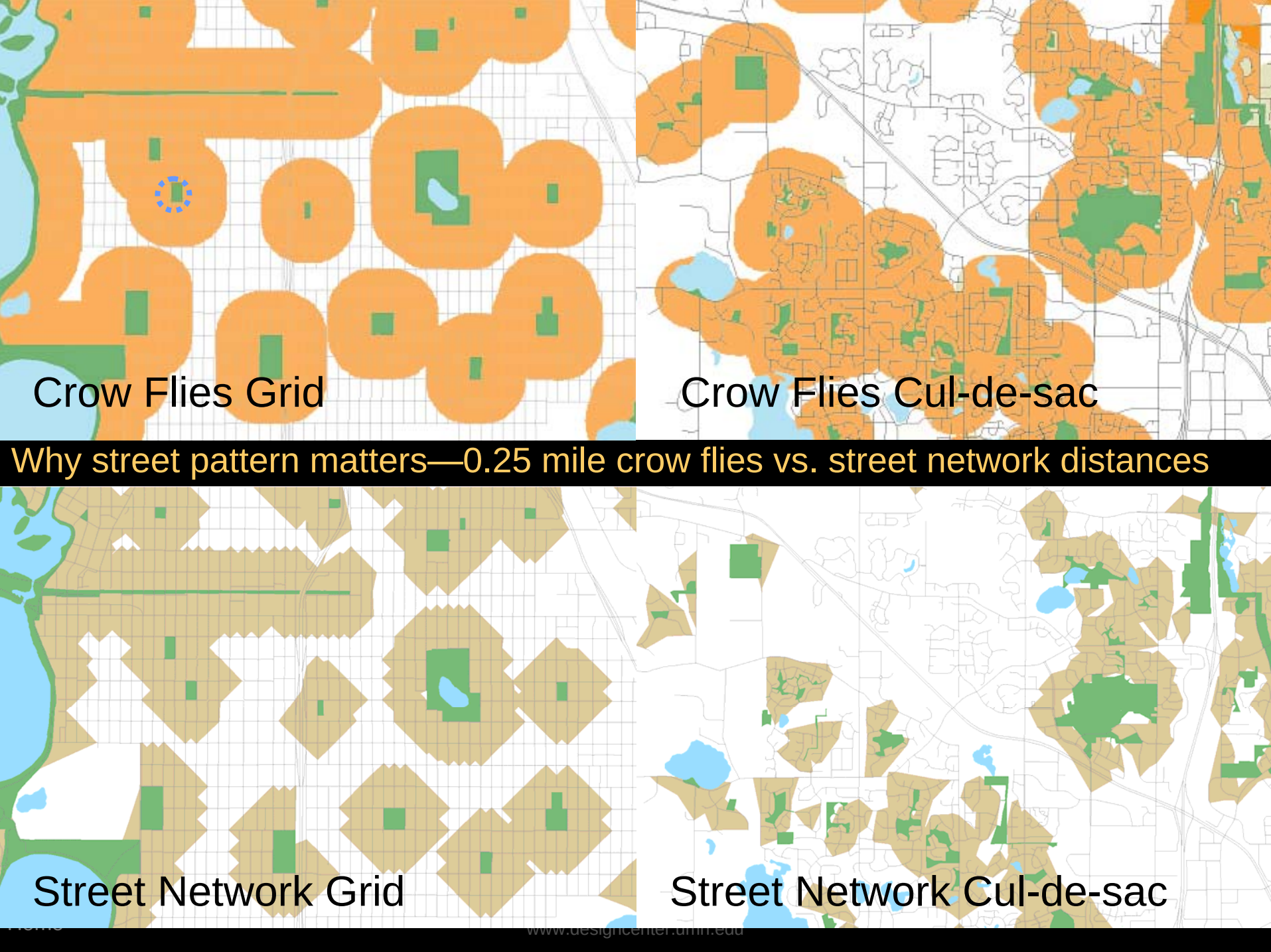
2.3 Creating a Focus Area Sampling Grid and Sampling Extreme Areas

2.4-2.9 Measurement Geographies

- 1: Grid Cell
- 2: Straight Line or Airline Buffer
- 3: Network or Street Distance Buffer
- 4: Straight Line or Airline Distance to Nearest Feature
- 5: Network or Street Distance to Nearest Feature
- 6: Parcel of Residence

Buffers




Crow Flies Grid
Crow Flies Cul-de-sac
Why street pattern matters—0.25 mile crow flies vs. street network distances
Street Network Grid
Street Network Cul-de-sac

2. Measures
Density

Why it matters: critical mass; physical sense of community; transit viability; auto congestion; proxy variable

Measures

Residential

3.1/3.2 Population per Unit Land Area: raw; without water

3.3 Population per Developed Land Area

3.4 Residential Population in Residential Parcels (Residential Density)

Activity

3.5 Population plus Employment per Unit Land Area

3.6 Employment per Unit Land Area

Buildings

3.7 Housing Units per Unit Land Area (gross, census data)

3.8 Lot Coverage

2. Measures

Pedestrian Infrastructure

Why it matters: Comfort, safety, interest

Sidewalks, street lights, street trees

4.2-4.4 Sidewalk Length: Divided by Road Length; per Unit Area; per Length of Major Road

4.5 Street Lights: per Length of Road

4.6 Street Trees within an X meter buffer per Length of Road + other measures

Inventory (Irvine-Minnesota, sample)

4.7-4.9 Percent of Street Segments with: Marked Pedestrian Crossings at One or Both Ends; Visible Litter, Graffiti, or Dumpsters; Traffic Calming, Broadly Defined

4.10+ Other fieldwork based measurements of neighborhood identification from crosswalks to architecture. 160+ question inventory


Street lights

Where are the street trees?

Street tree zone up to 90 m wide– individual trees digitized.

Where continuous cover, this was photographed and allocated to 3 categories of minimum densities (field based)

2. Measures

Mix

Why it matters: Movement between destinations; safety/activity

Land Use

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

5.2-5.5 Percentage of Land Area in Night Time Uses; Social; Retail; Industrial and Auto-Oriented

5.6-5.7 Proportion of Dissimilar Land Uses Among Grid Cells

Employees

5.8 Total Employment per Unit Area

5.9 Density of Employees in each of Major Retail Subcategories (Separate Measures)

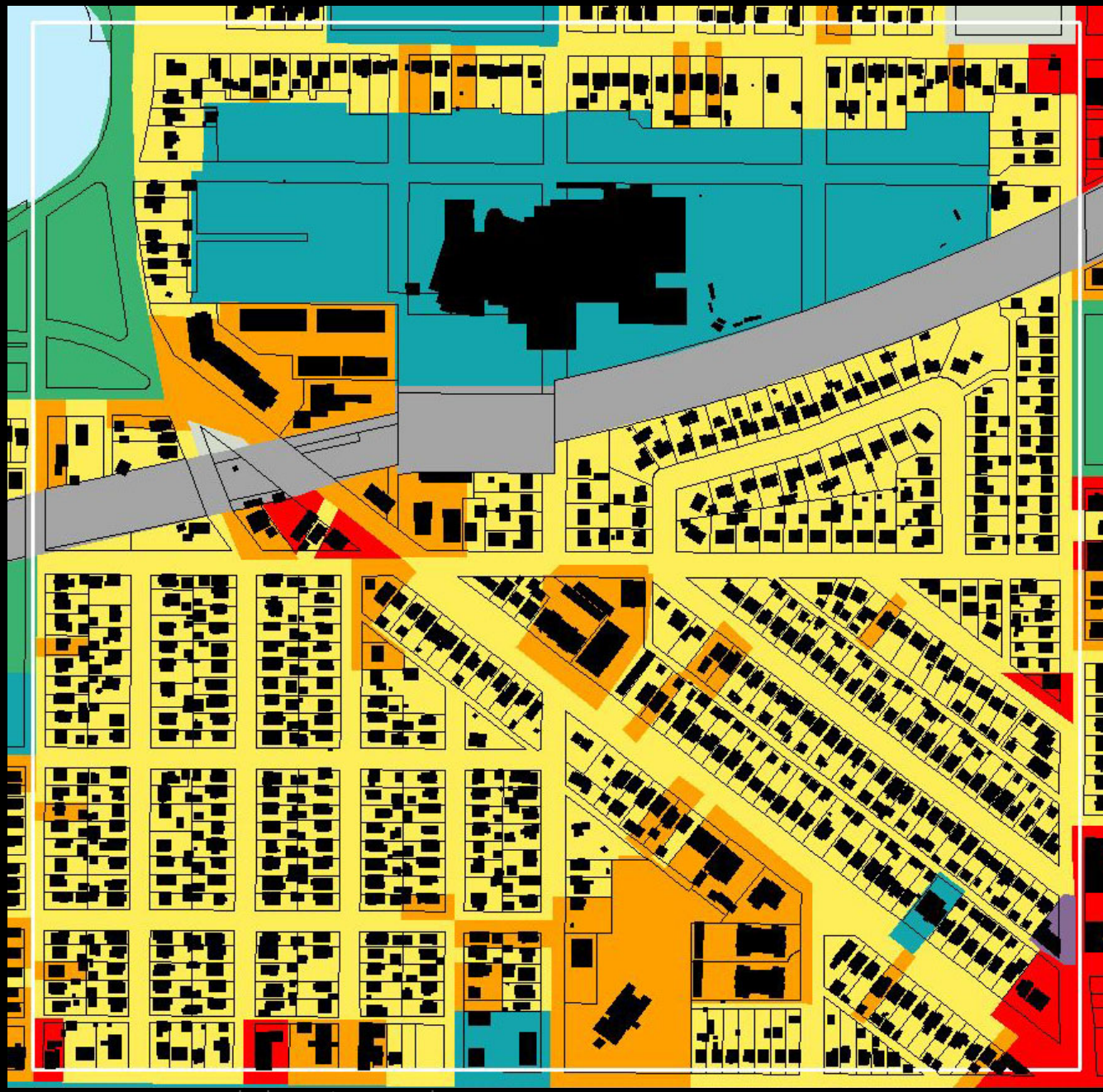
Facilities

5.10 Distance to Nearest Facility (same facilities as 5.9)

+ UNC measures e.g. Gini coefficient, HHI index, and Entropy index.

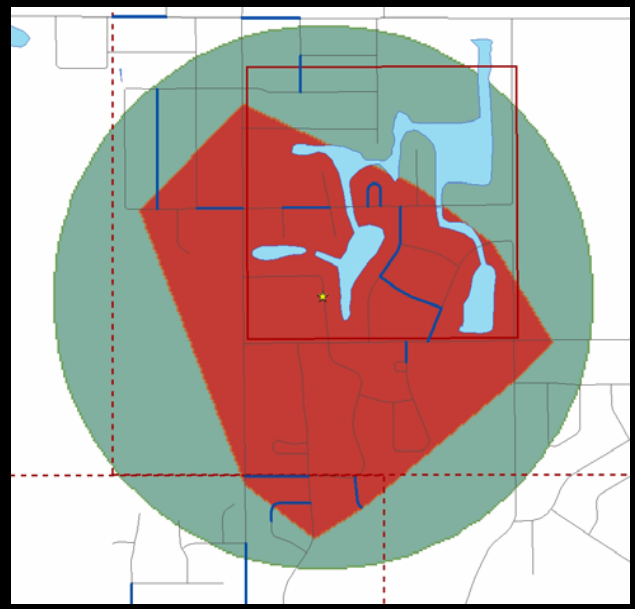
Mix

•use e.g. commercial—bowling alley



2. Measures

Street Pattern



Why it matters: Directness; alternate routes

Blocks

6.1/6.2 Average/Median Census Block Area

6.3 Ratio of Area within X Street Distance to Area within X Distance Radius

6.12 Median Perimeter of Block

Roads

6.4 Number of Access Points

6.5 Road Length per Unit Area

Intersections

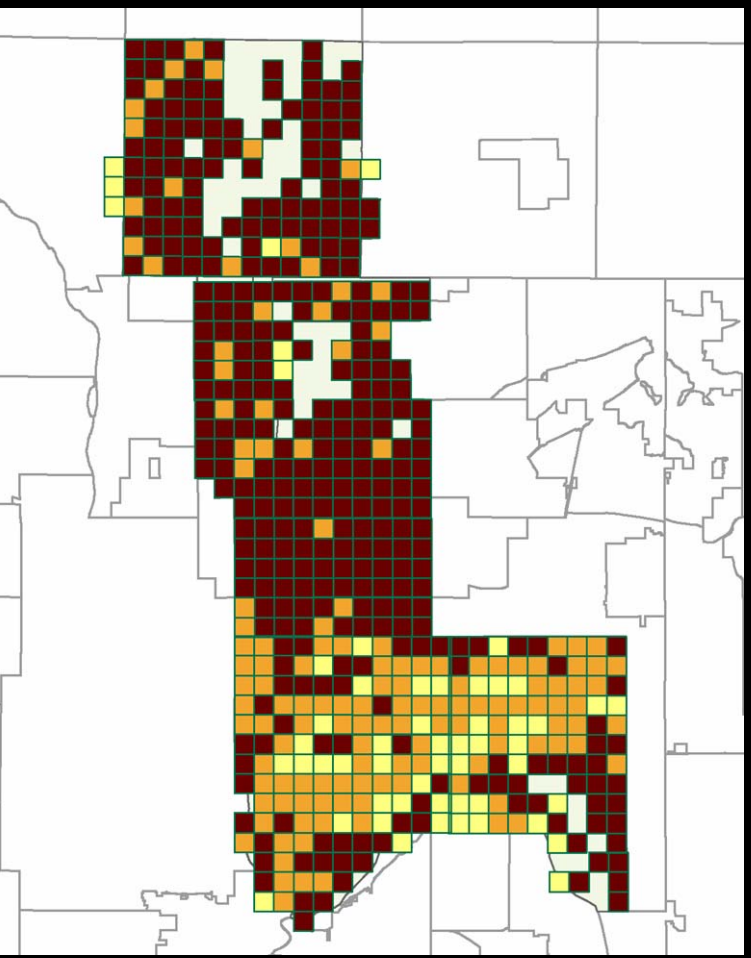
6.6 Intersections per Unit Area

6.7-6.9 Ratio of 3 or 4-Way Intersections to All Intersections; per unit area

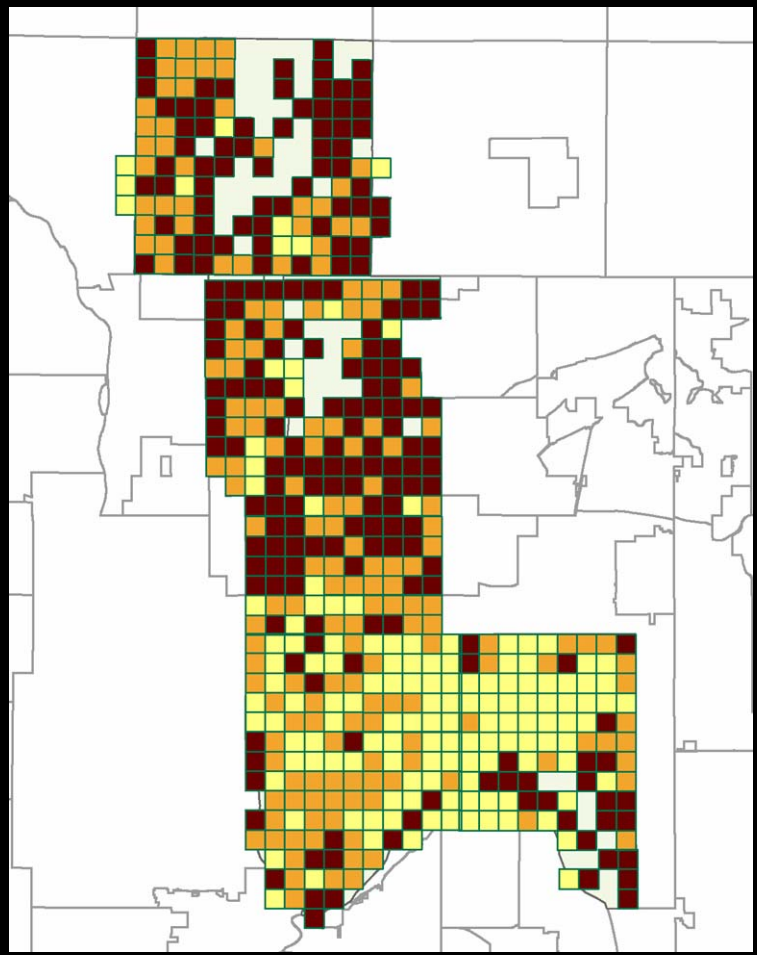
6.10-6.11 Node Ratios

2. Measures

Average vs. Median Block Size



Average



Median

Maroon: 8+ ac, 3.2+ha
Orange: 2-3.2 ha
Yellow: 0 – 5 ac, 0-2 ha

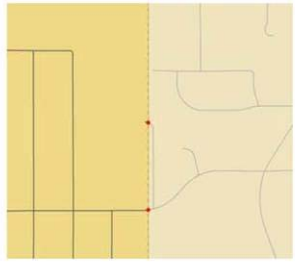
3. Protocols
Structure

1. Basic Concept
2. Basic Formula, or Basic Definition, Basic Procedure

   Population per Unit Land Area (without water) = Persons in housing units per unit gross area excluding water.
3. Detailed Formula or Detailed Definition

   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.
4. Comments and Explanations
5. GIS Approach
6. GIS Steps

3. Protocols
What a Protocol Looks Like



Many access points 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 [dark] 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.

The three purple [dark] circles identify points that should be included in the analysis, since they provide significant access to the site, but 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.

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 were summed.

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

To complete this protocol you will need:

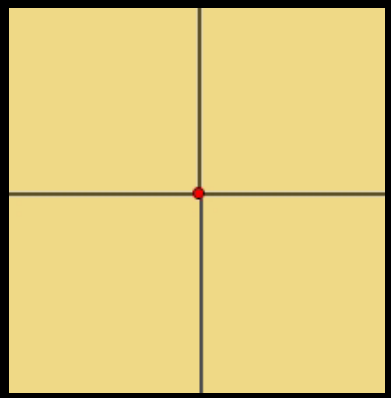
1. Site layer (polygon)
2. Site boundary layer (polyline)
   If you only have a polygon site layer, convert the polygons to polylines using the XTools extension (available from the Downloads section of the ESRI website).
3. Road centerline layer (polyline)
4. AddPointsAtCrossings ArcScript available from the Downloads section of the ESRI website

The basic steps are:

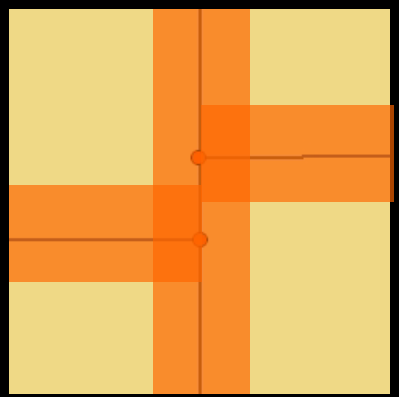
- *Create the Access Point file*
- *Correct for Interstate Access Points*
- *Correct for Parkway Access Points*
- *Remove Unwanted Access Points from View/Table*

4. Typical Problems

A. What is an Intersection?



A

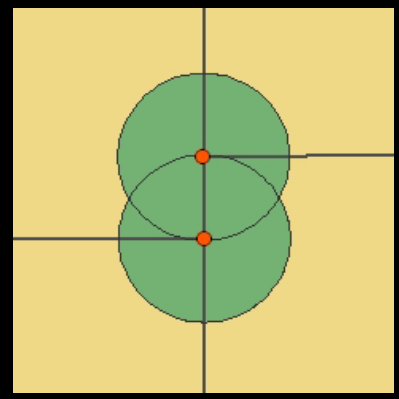


B

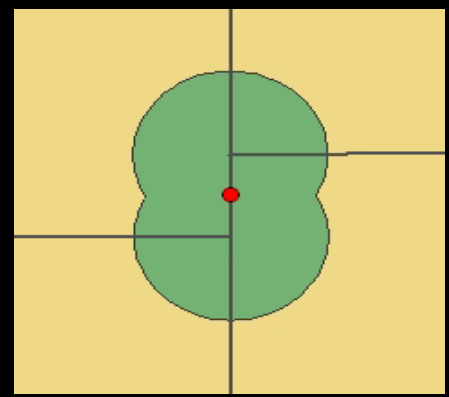
4-Way or X intersections indicate higher connectivity than 3-way or T intersections

Issue simple 4-way intersections vs. functional 4-way

Solution, is buffering: 10m (20m offset); 15m



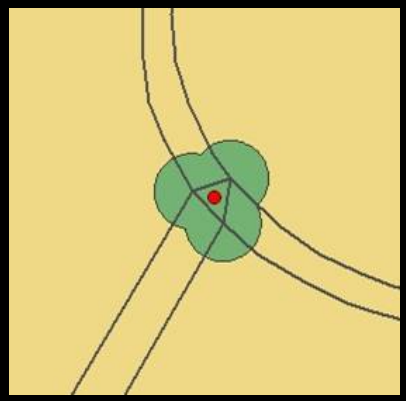
C



D

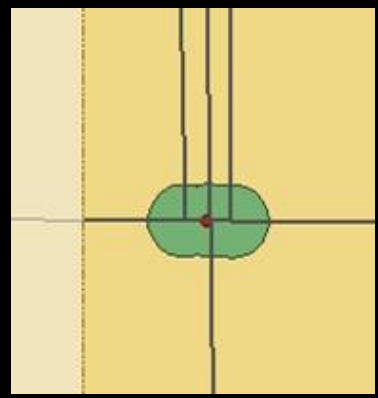
4. Typical Problems

A. What is an Intersection?

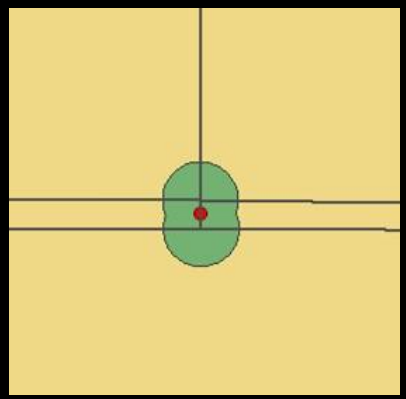


Valence 12

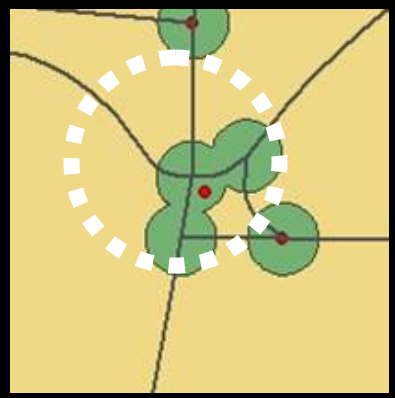
Complex intersections—high valences but not all function with high connectivity (valence 7 and 12 are often 3 way)



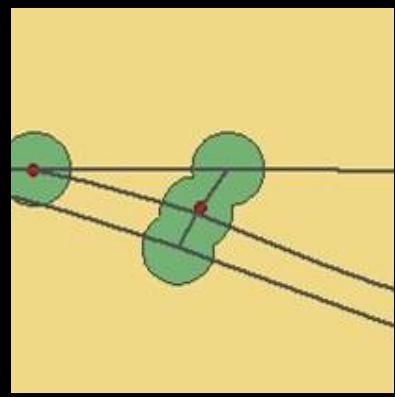
Valence 10



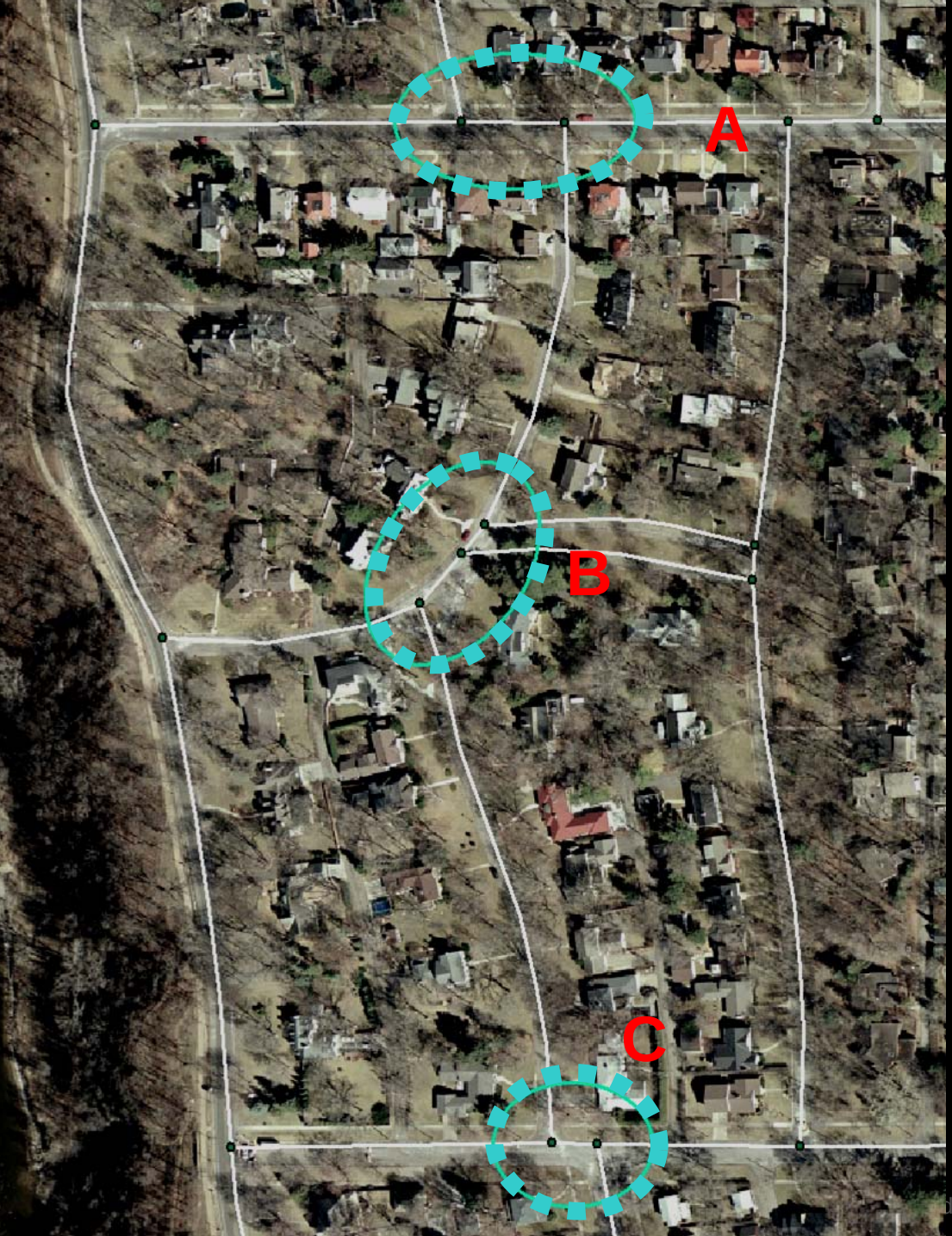
Valence 7



Valence 10



Valence 10



Are the T-intersections in this neighborhood distinct or experienced as functional 4-way intersections?

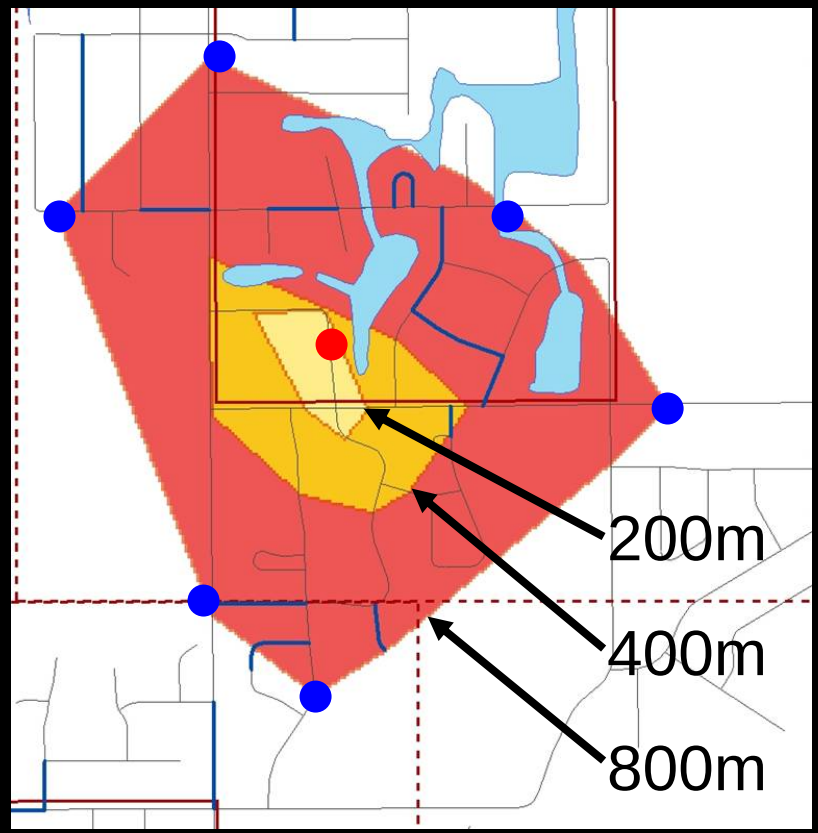
A. 130ft/40m apart
B. 100ft/30m apart
C. 65ft/20m apart

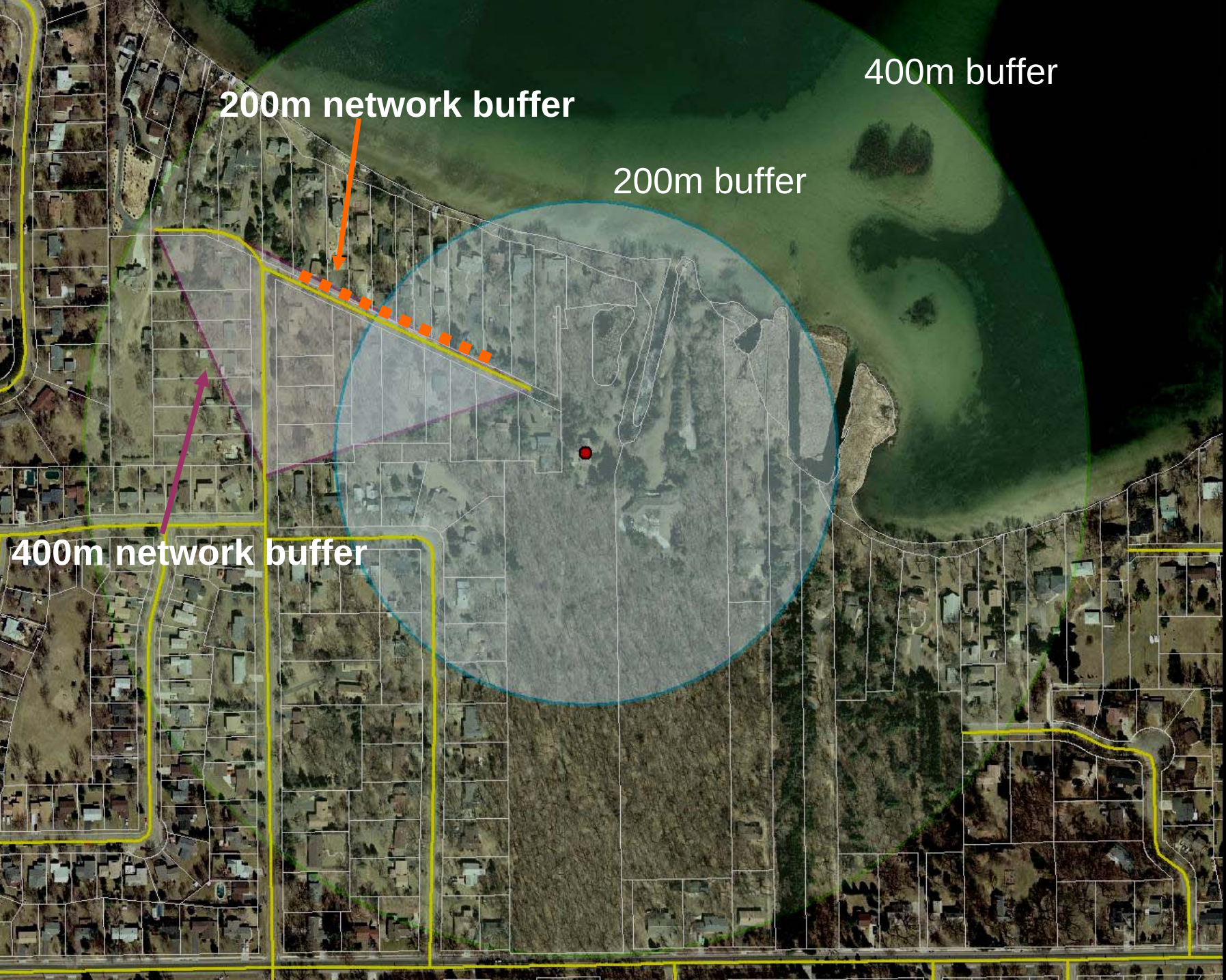

C. Buffers at:
10m/30ft (60 ft between centerlines)
15m
20m

4. Typical Problems

Making Network Buffers

- Network buffers indicate the area reachable within a certain street distance
- They are created by joining the points X meters (blue dots) from the starting point (red dot)
- However, some street patterns end up with network buffers with no area, which is an artifact of how buffers are constructed and does not reflect the experience of the place (see next slide)




400m buffer
200m network buffer
200m buffer
400m network buffer

15m buffer/30m width for network buffers under minimum 12,000+m2 (30m*200m*2)—still refining

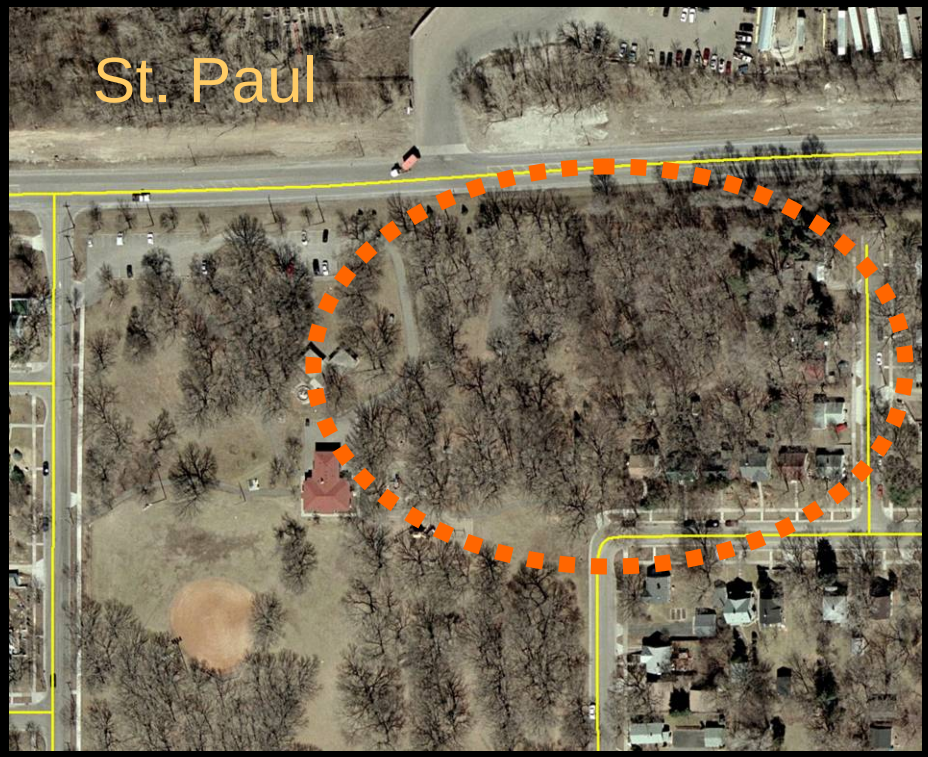
4. Typical Problems

C. How Many Trees in a Forest?

- Continuous canopy in aerial photos
- Perception of tree density--after a certain level there are just a lot of trees so it probably makes no difference in perception but saying there are 2700/ha vs. 500 or 150 would skew tree densities
- Need more research on perception but in interim…

4. Typical Problems

C. How Many Trees in a Forest?



Trees are too dense to count from the photo

4. Typical Problems

C. How Many Trees in a Forest?

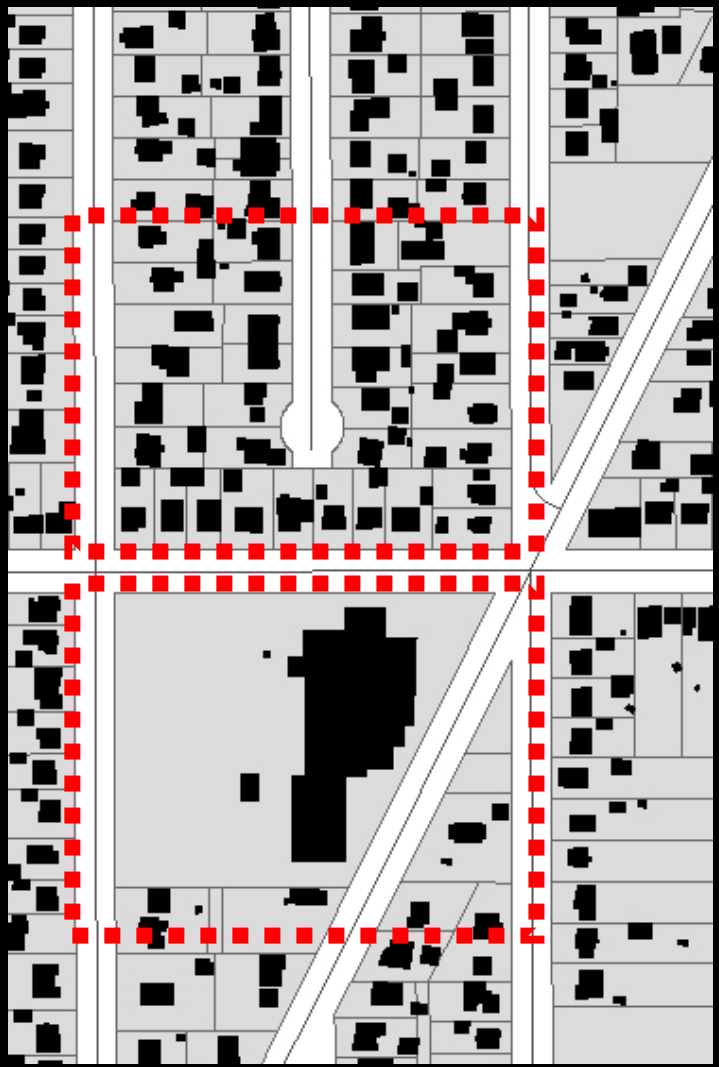
St. Paul, approx 60/acre
150/ha

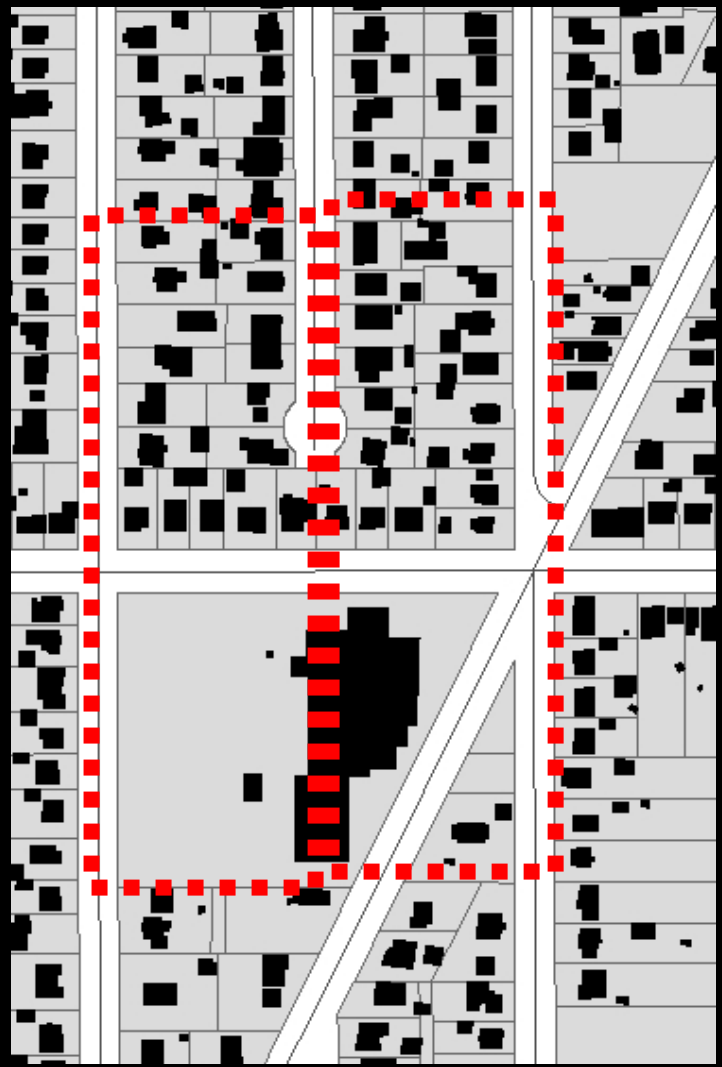
St. Paul, approx 1100/acre
2700/ha above about 4 inch diameter

5. Analysis Issues

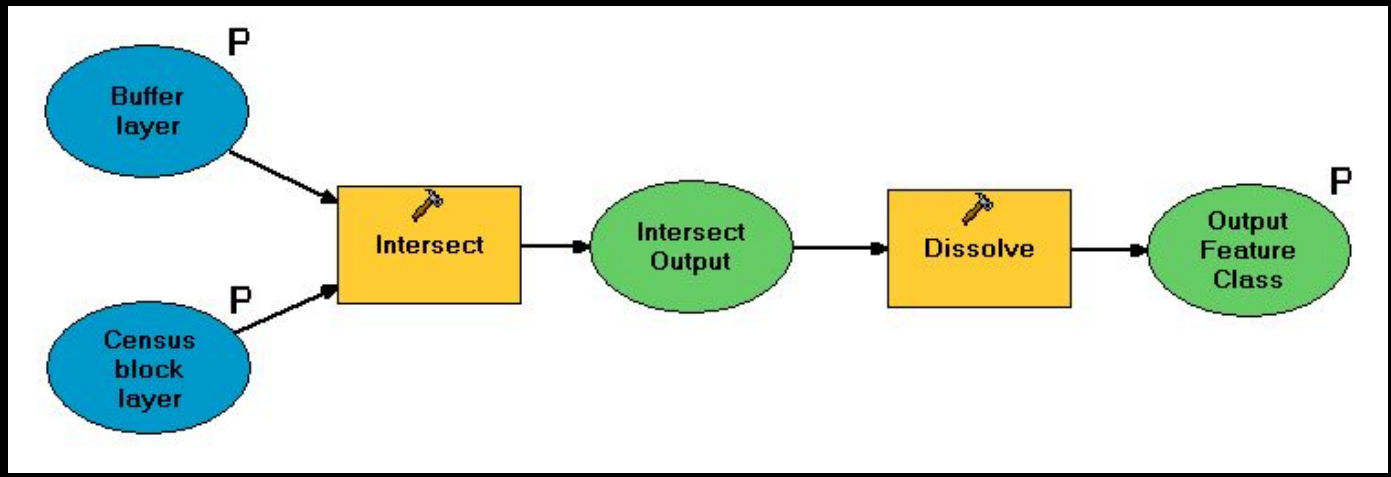
- Software—scripting to approximate measures
- Modifiable Areal Unit Problem (http://www.geog.ubc.ca/courses/geog516/notes/maup.htm)
  "The scale effect is the tendency…. for different **statistical r**esults to be obtained from the same set of data when the information is grouped at different levels of spatial resolution (e.g., enumeration areas, census tracts, cities, regions)."

  "The aggregation or zoning effect is the variability in statistical results obtained within a set of modifiable units as a function of the various ways these units can be grouped at a given scale, and not as a result of the variation in the size of those areas." (**geographical**)
- Spatial autocorrelation—level, nature, and strength of interdependence between variables; close by things are similar (new tools Arc 9.0)



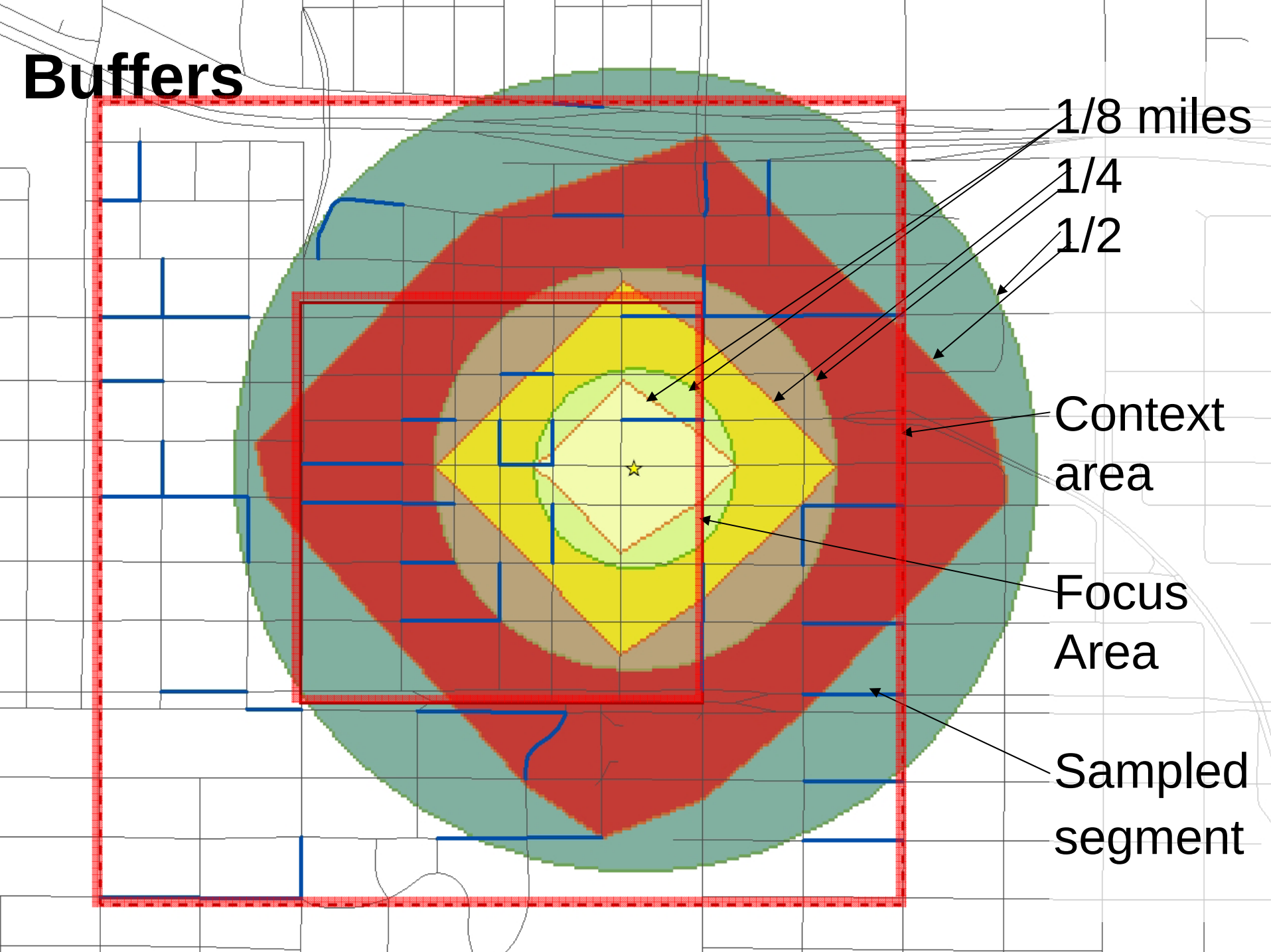


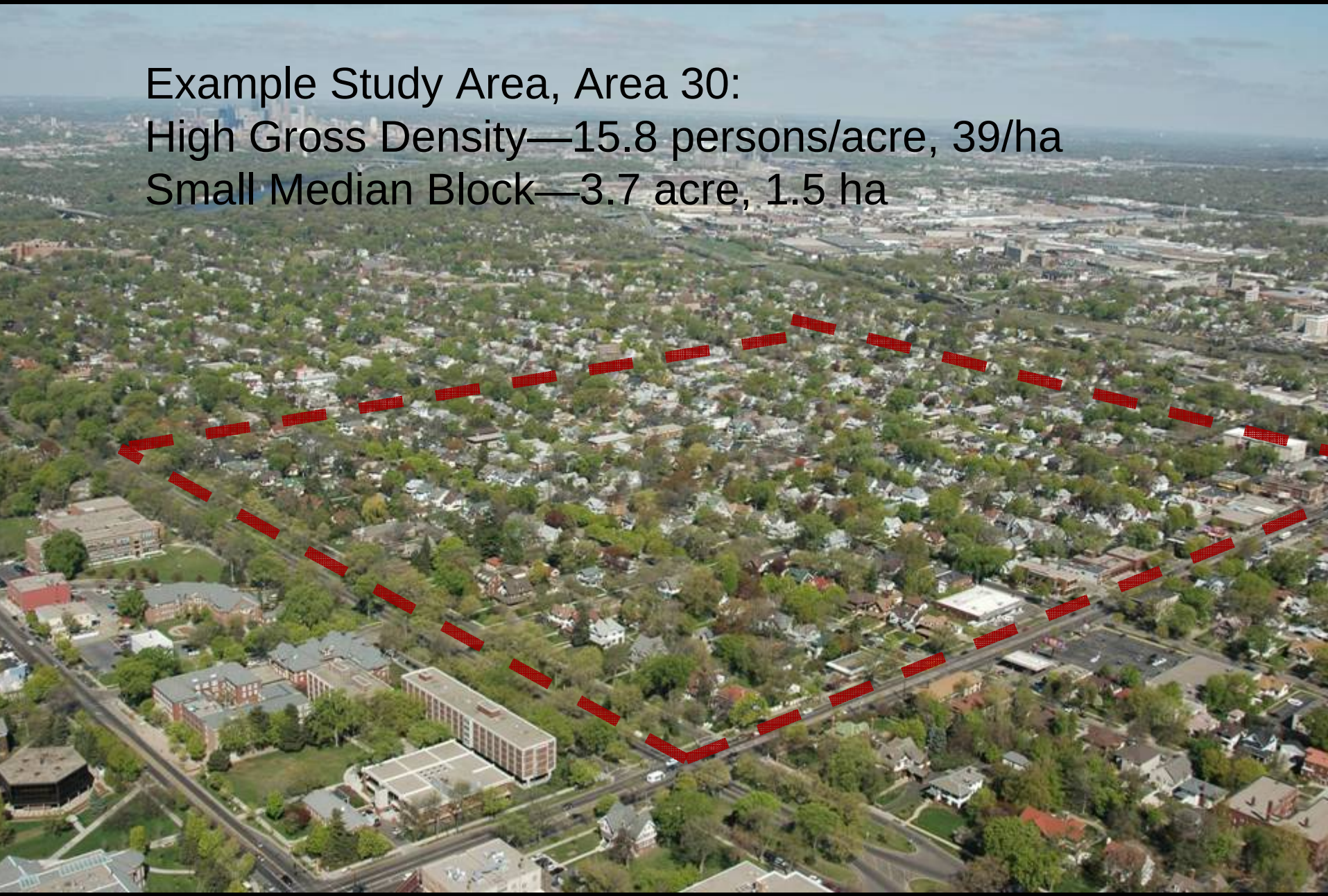
6. Final Tips for Researchers



- Read the GIS documentation—you'll get new ideas
- Try things out yourself rather than leaving it all to the GIS specialist—you'll at least get to look at the data in detail
- Look at the ESRI discussion forums—that's how the GIS technical folks figure out how to do things
- Communicate among the technical team in writing—that way you'll identify places where you are not communicating well

Buffers




Example Study Area, Area 30:
High Gross Density—15.8 persons/acre, 39/ha
Small Median Block—3.7 acre, 1.5 ha

GIS and Physical Activity

PowerPoint Suite

Created by the Metropolitan Design Center, University of Minnesota, for Blue Cross Blue Shield of Minnesota

Draft June 2005

Main Menu

- Main Shows

1 Community Design and Active Living
2 Environmental Features and Physical Activity
3 Twin Cities Walking Study
4 GIS-Based Measures
   a. Geographies
   b. Data layers
   c. Techniques
   d. Density
   e. Pedestrian infrastructure
   f. Mixed use
   g. Street pattern

- 5 Protocols

Supporting Shows

Buffers

Trees/Forests

Intersections

GIS for the Non-Expert

1. Definitions and problems
2. Measures being tested in Twin Cities Walking Study
   - Basic approach, data, variables and geographies
3. Protocols
   - Concept, formulae, approach, and steps
4. Typical measurement dilemmas
   - Intersections, small network buffers, counting trees
5. Analysis issues