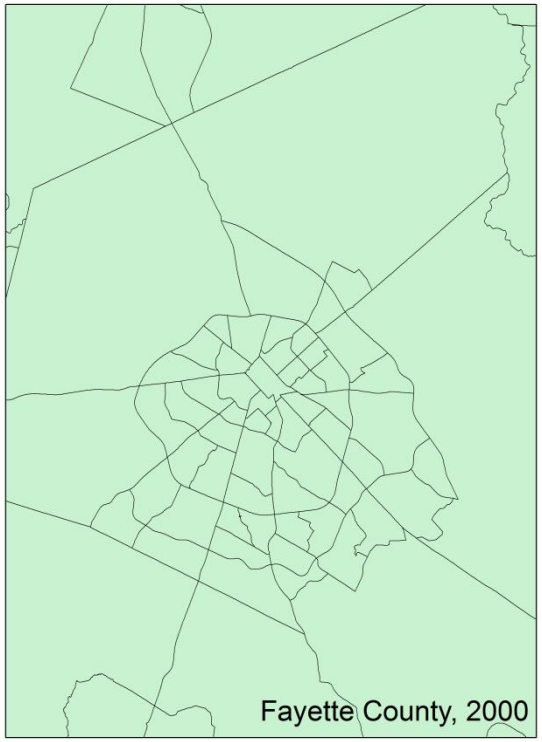
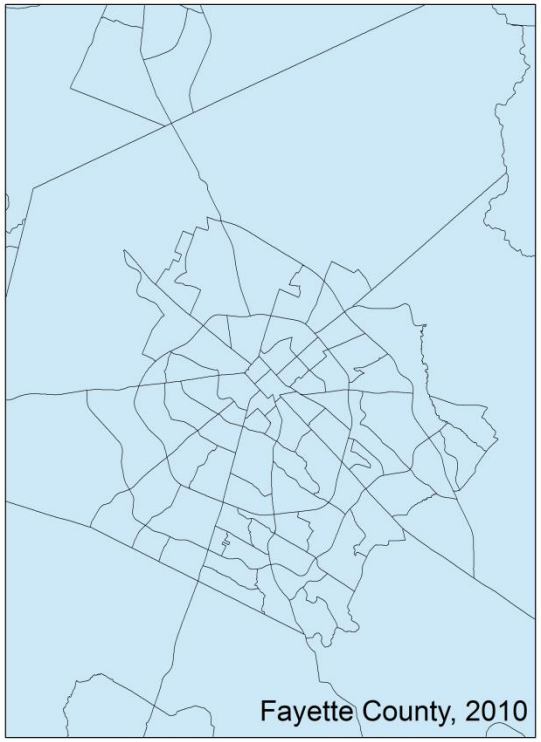




Integrating Areal Interpolation and Dasymetric Refinement to Resolve Temporal Incompatibilities in Zoning Systems

Sarah Ehresman and Matt Ruther, Ph.D.
Kentucky State Data Center, University of Louisville

KSDC
KENTUCKY STATE DATA CENTER

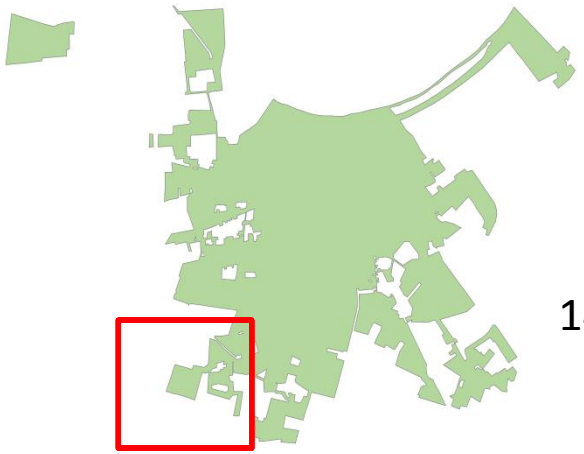
UNIVERSITY OF LOUISVILLE

Temporal incompatibility in zoning systems

- Temporal analysis of Census (and other) data is limited by changing enumeration boundaries over time
- Difference between social/physical science data
- What do we want to do?
  - Estimate small area “populations” over time
  - Not just population….any data that is enumerated (school districts, poverty data, unemployment data)
- Accessible, generalizable, and accurate

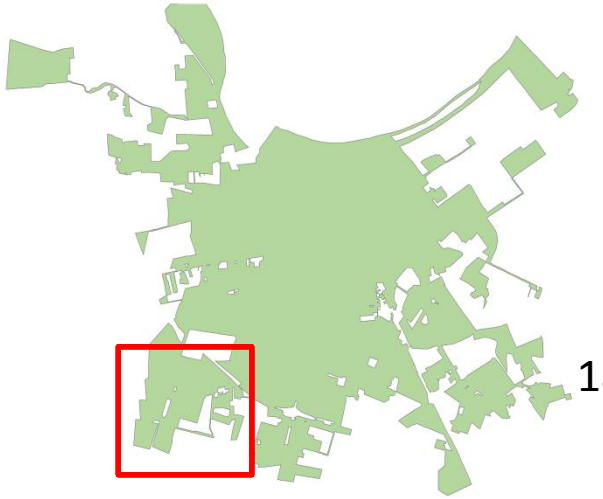
Incompatible temporal boundaries

City of Owensboro



14.2 sq. mi.
1980

16.2 sq. mi.
1990



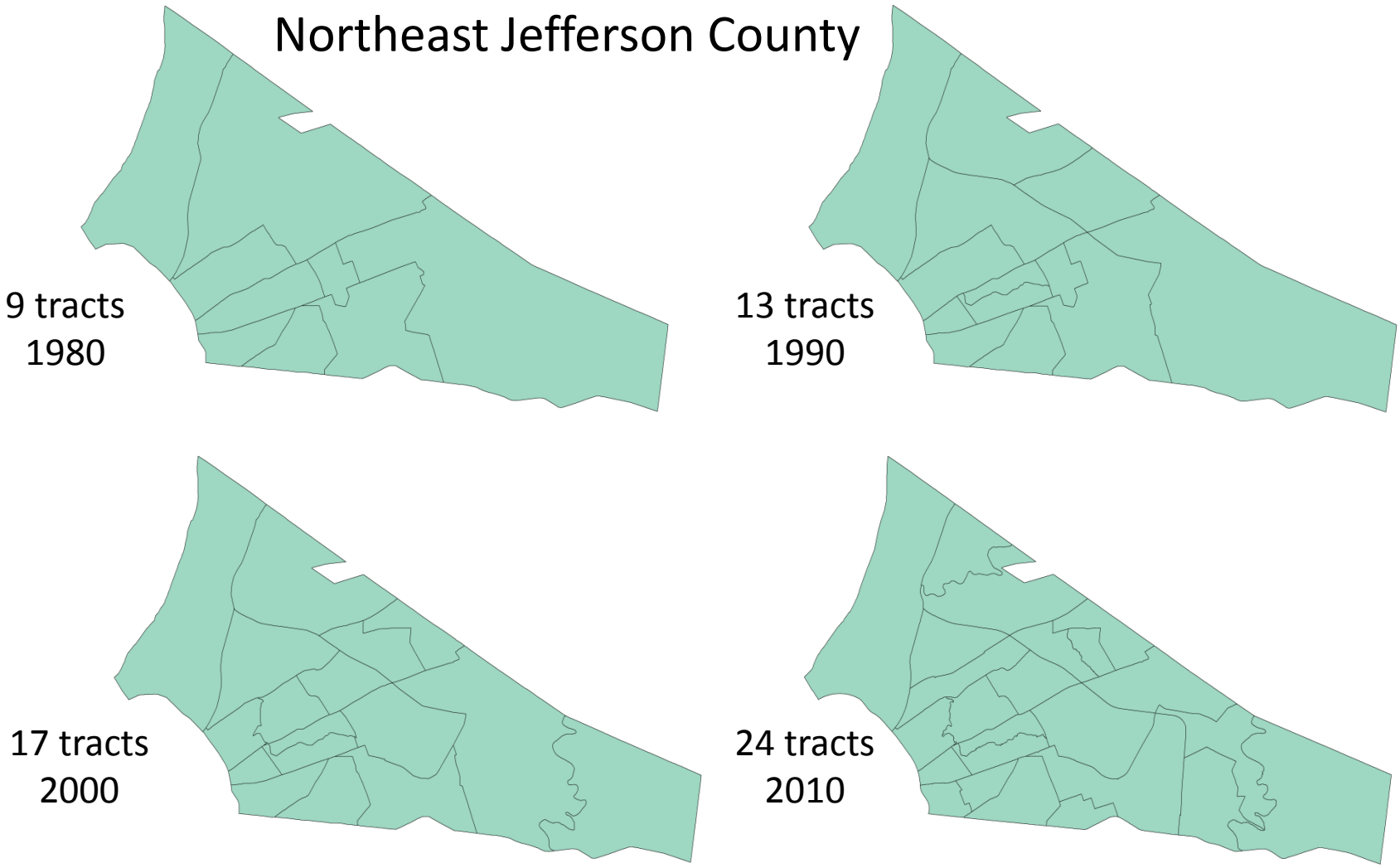
18.5 sq. mi.
2000

20.4 sq. mi.
2010

KSDC
Kentucky State Data Center

University of Louisville

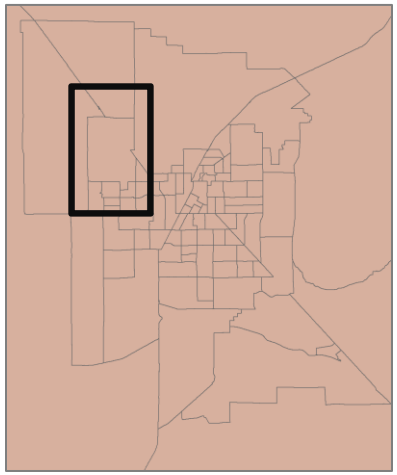
Incompatible temporal boundaries – Census tracts



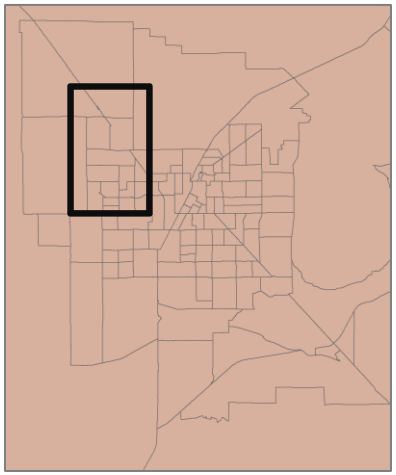
How to deal with changing Census tract boundaries?

- Commercial databases, such as the Neighborhood Change Database
  - $$$
  - Inflexible
  - Rely on areal interpolation methods (DIY!)
- Generate minimum comparable areas
- Areal interpolation

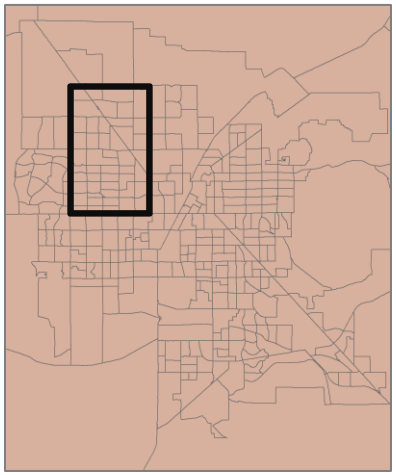
Minimum comparable areas



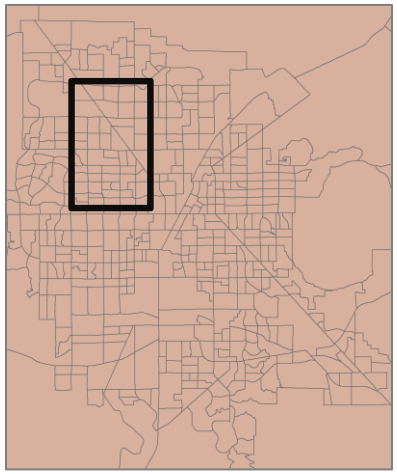
1980



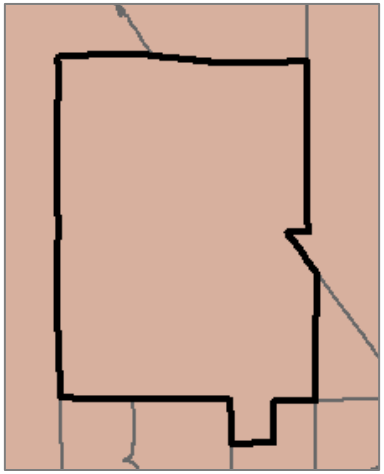
1990

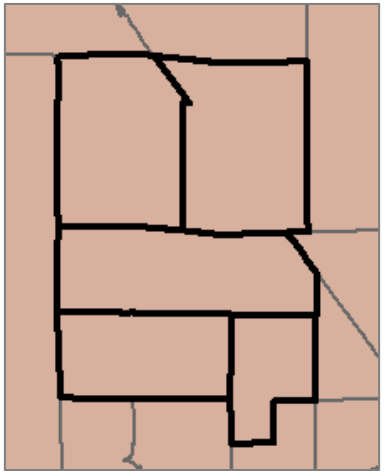


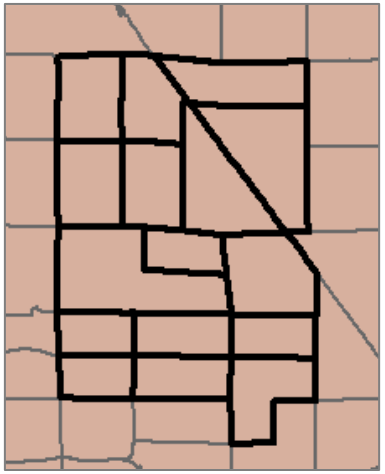
2000

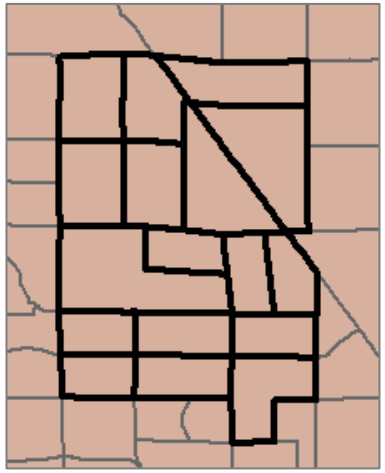


2010





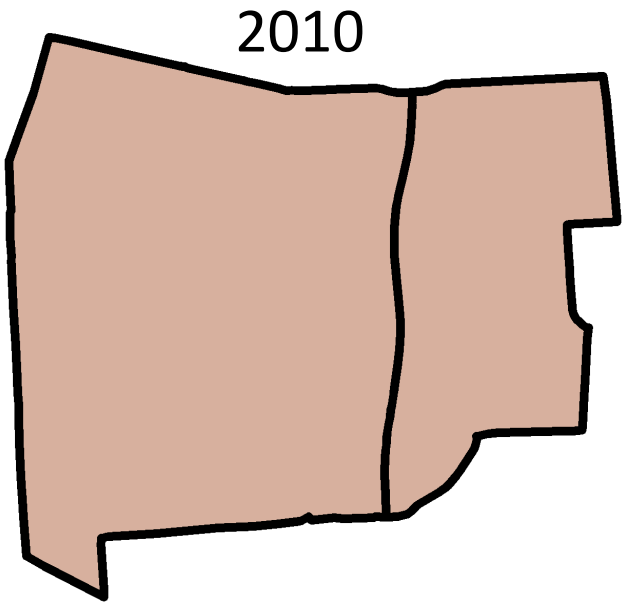


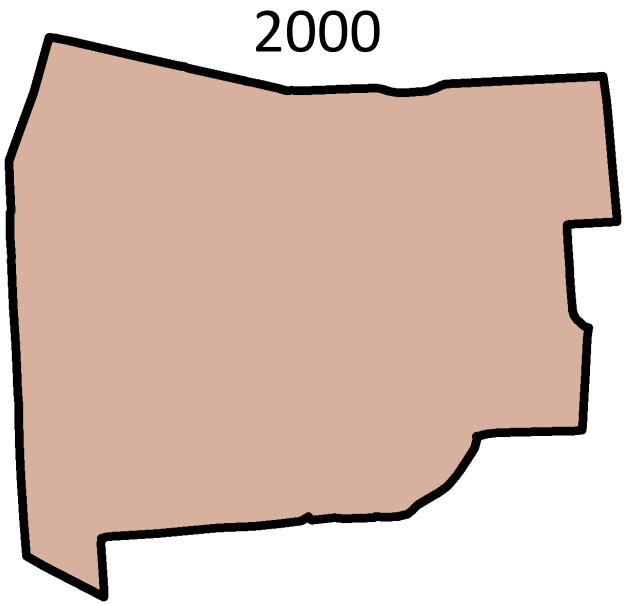


KSDC
Kentucky State Data Center

University of Louisville

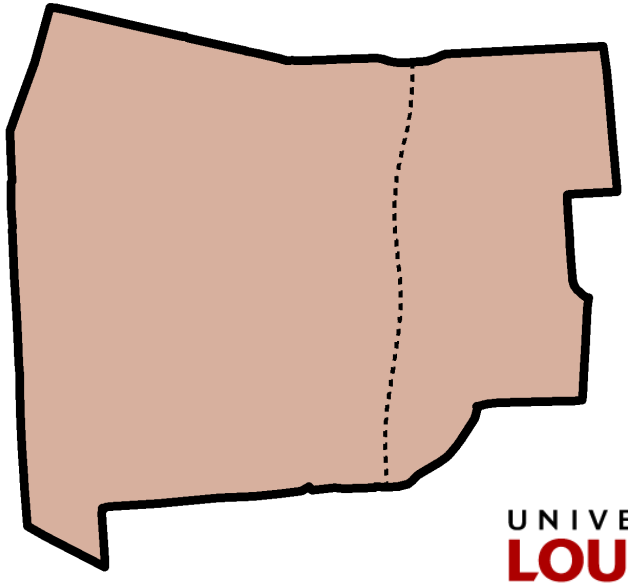
Areal interpolation methods





2000 data (e.g. population) within 2010 boundaries

2010 = Target census
2000 = Source census



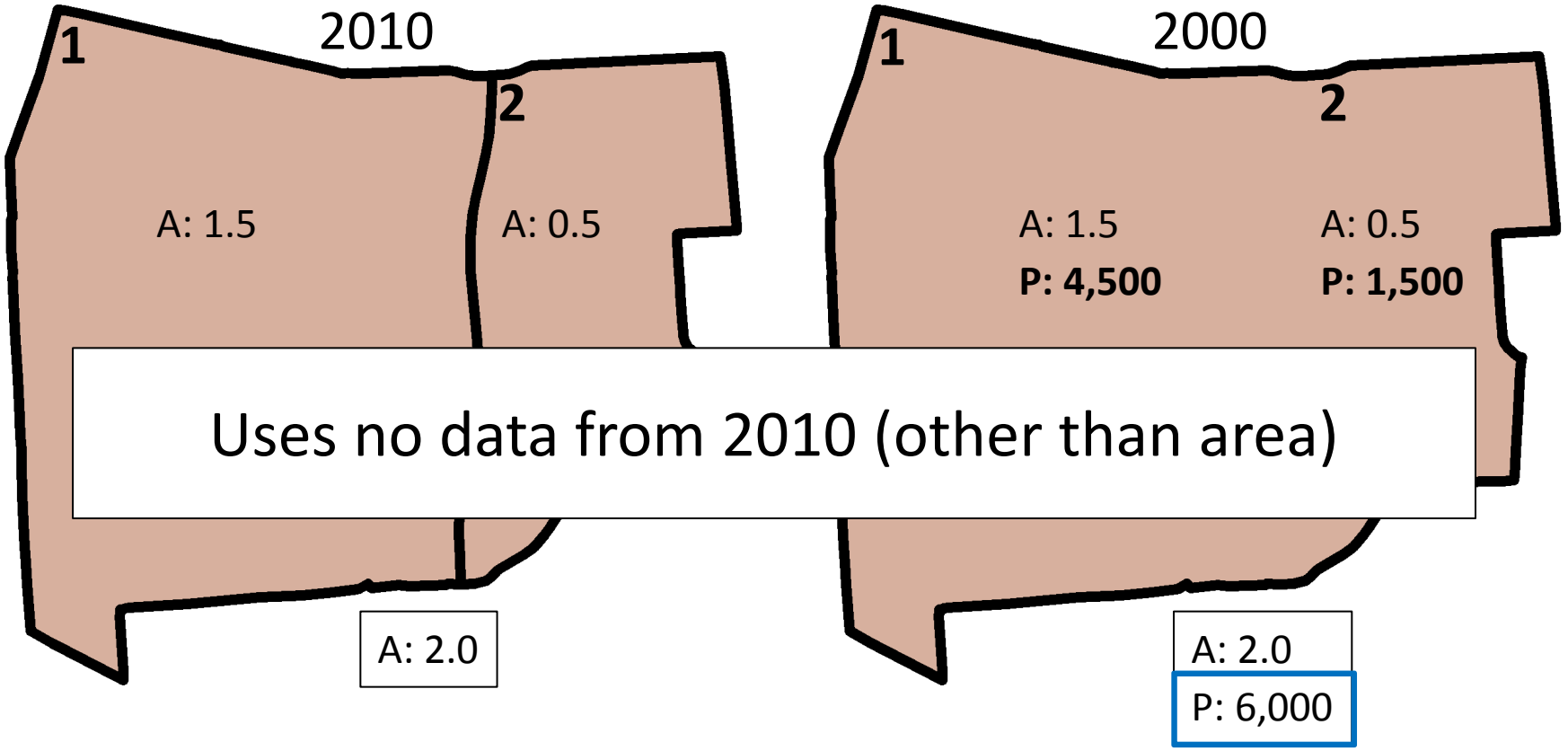
Method 1: Areal weighting

Estimation of source populations within target zones is based on the proportion of areal overlap between target zones and source zones…

Based only on geometry!

$$\hat{y}_t = \sum_s \frac{A_{st}}{A_s} y_s$$

Method 1: Areal weighting



Uses no data from 2010 (other than area)

Sector 1: 1.5/2.0 = 0.75 0.75 × 6000 = **4500**

Sector 2: 0.5/2.0 = 0.25 0.25 × 6000 = **1500**

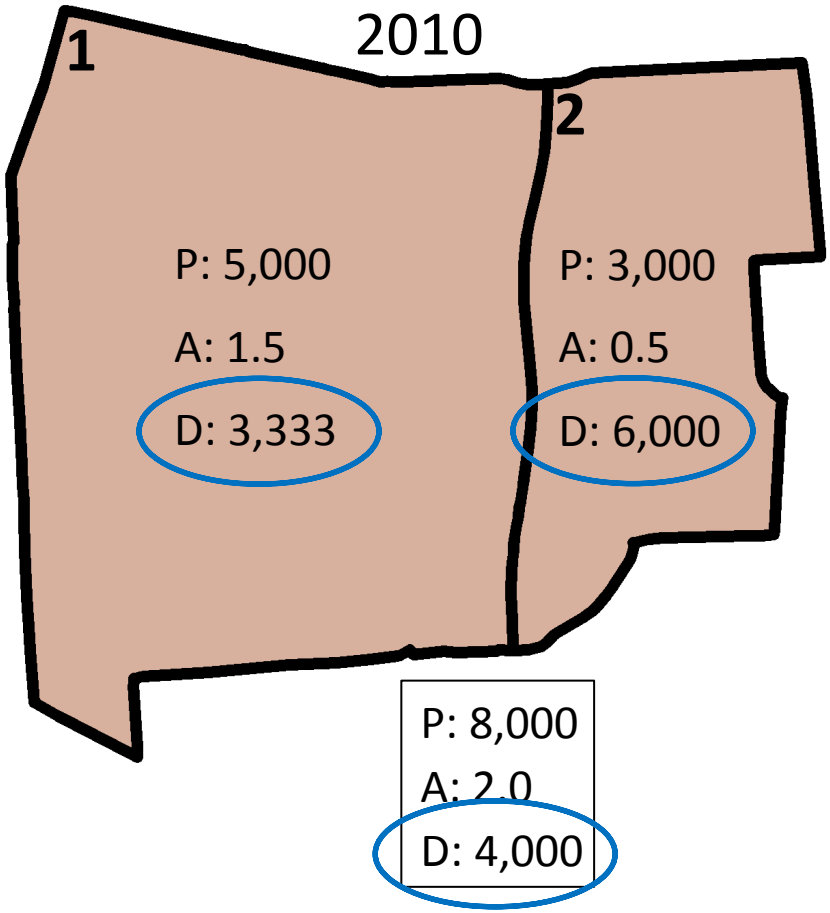
Method 2: Target density weighting

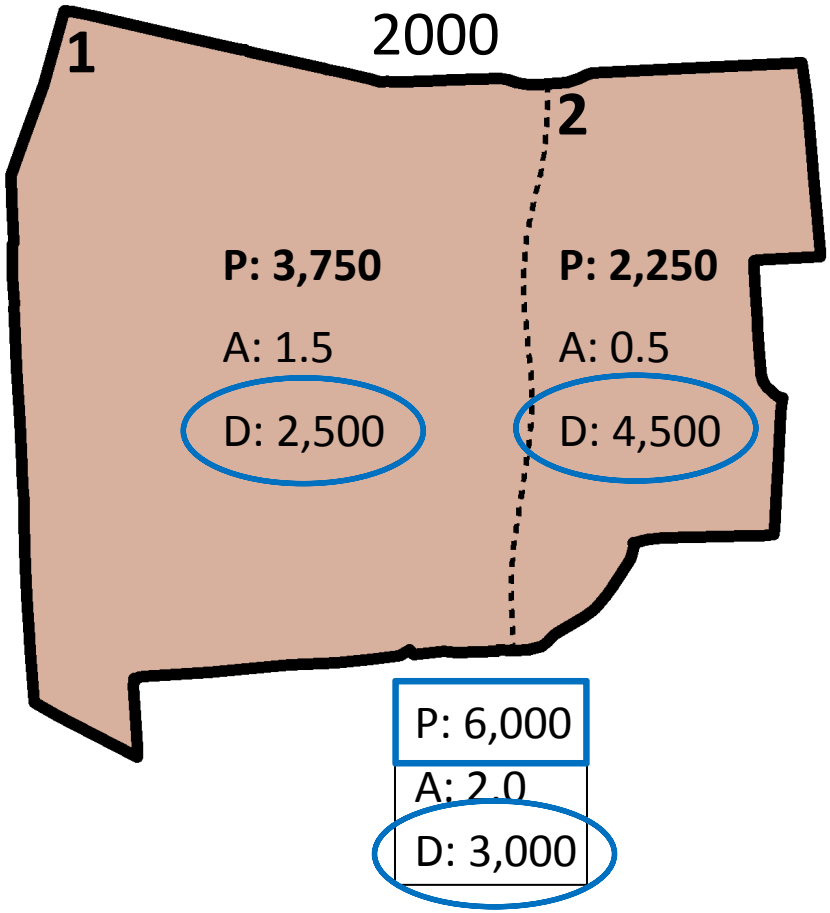
Estimation of source populations within target zones is based on the ratio of densities of target populations (or some other ancillary variable) within target zones

Ratio of densities in source year is set equal to ratio of densities in target year

$$\hat{y}_t = \sum_s \frac{\left(\frac{A_{st}}{A_t}\right) z_t}{\sum_\tau \left(\frac{A_{s\tau}}{A_\tau}\right) z_\tau} y_s$$

Method 2: Target density weighting





Sector 1: $\frac{3333}{4000} = \frac{D}{3000}$ $D = 2500$ $2500 \times 1.5 = \mathbf{3750}$

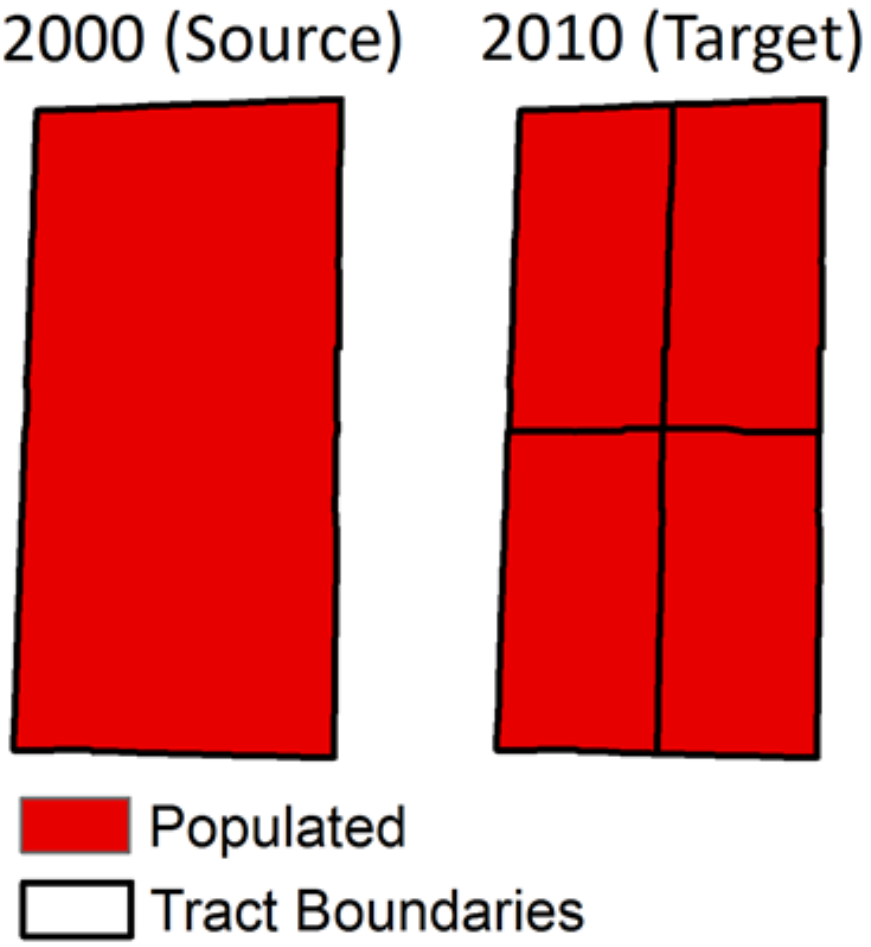
Sector 2: $\frac{6000}{4000} = \frac{D}{3000}$ $D = 4500$ $4500 \times 0.5 = \mathbf{2250}$

KSDC Kentucky State Data Center

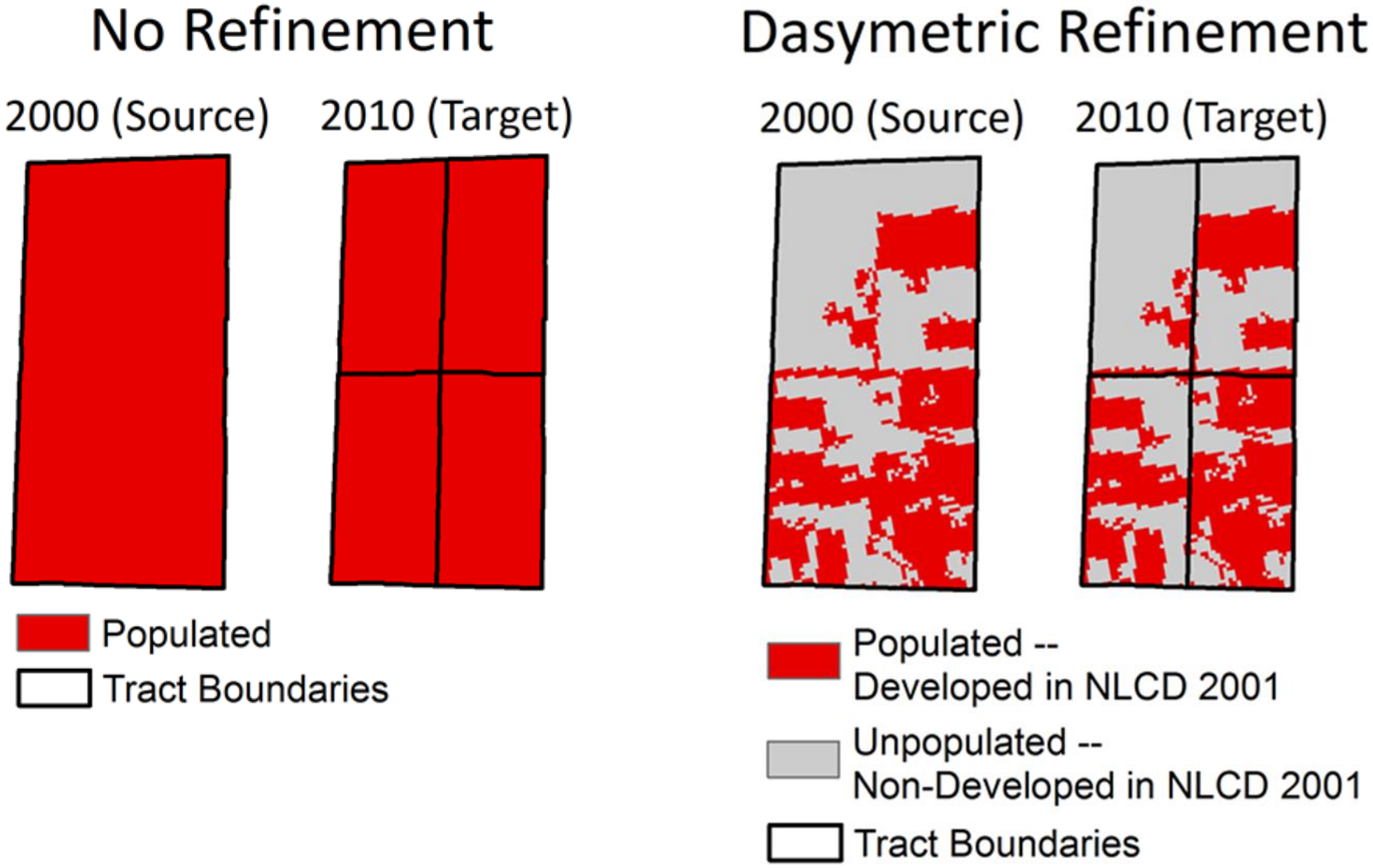
UNIVERSITY OF LOUISVILLE

How can these methods be improved?

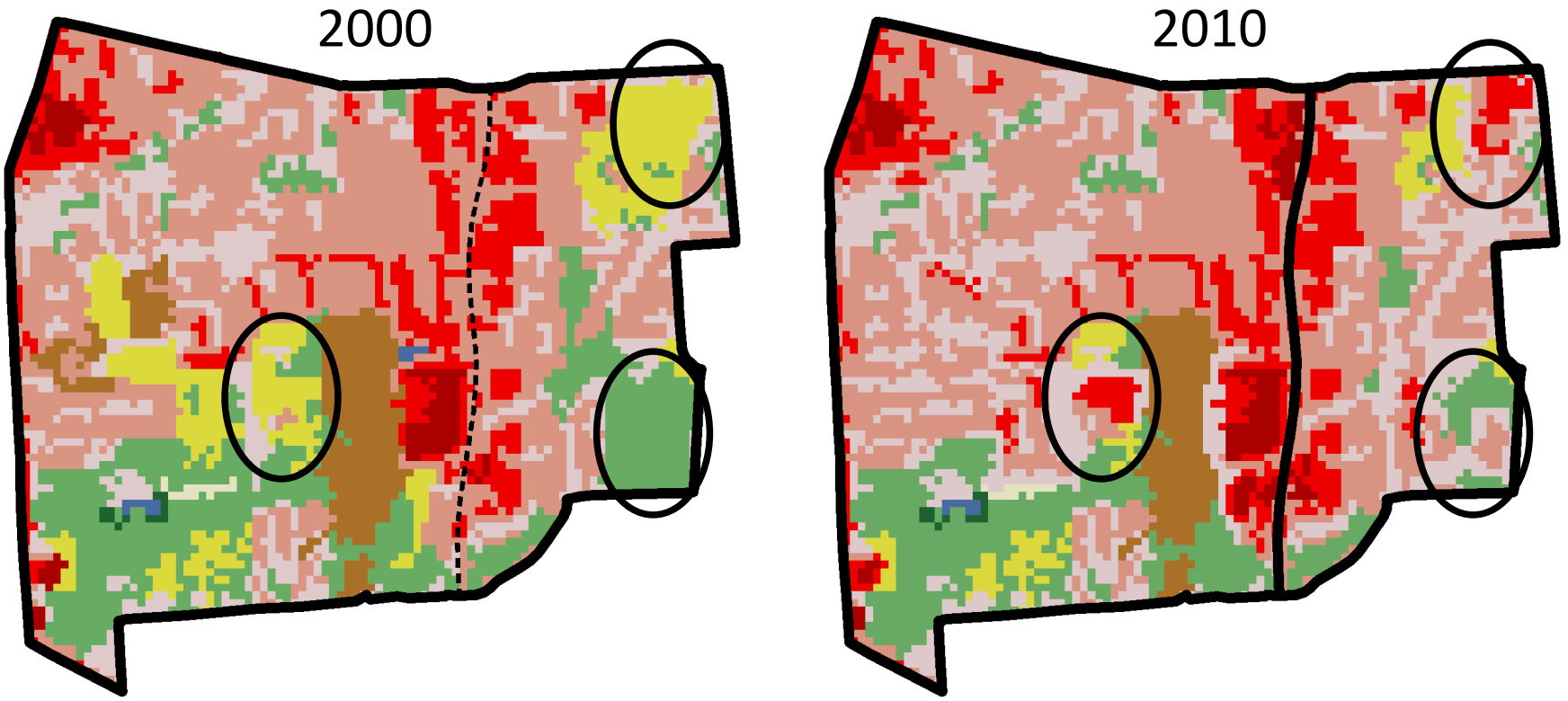
Both AW and TDW assume that the population is evenly distributed – this assumption is likely inaccurate!



Improvement through dasymetric refinement



Identifying development with NLCD



Method 3: Expectation-maximization algorithm

A third method, the expectation maximization algorithm, uses an iterative likelihood process to determine the different population densities for each NLCD class within a county.

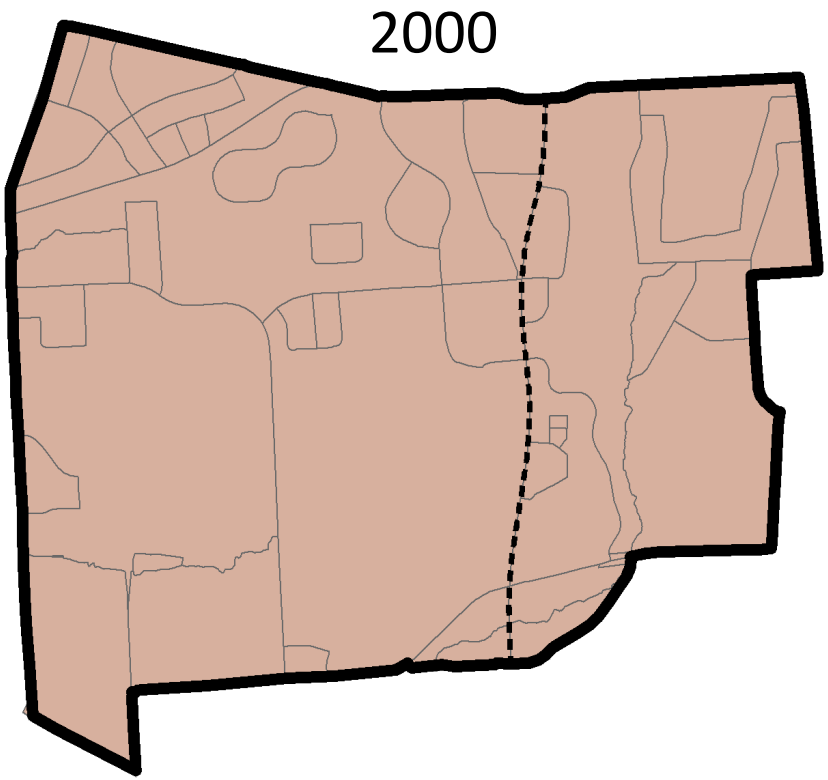
Once the appropriate population density in each NLCD class is determined, the population in the source year is distributed to the target zones based on the land cover within each target zone.

Goals of the project

- Areal interpolation is not a new idea but integrating it with spatial refinement over time has not been studied thoroughly
  - Which method (AW, TDW, EM) performs best?
  - How does the spatial refinement affect accuracy?
  - Does the time between Censuses matter?
- Create state-wide dataset of temporally consistent population and housing estimates at the neighborhood level

Results

Validating the results for Census tracts

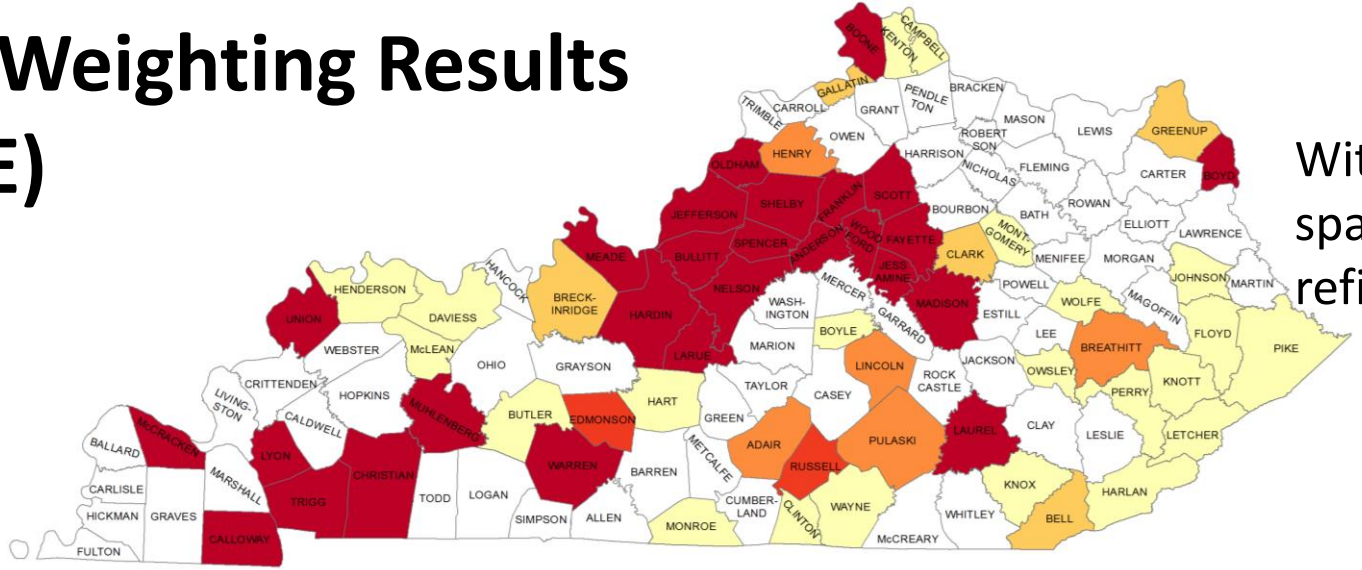


Blocks sometimes do not align *exactly* with tract boundaries

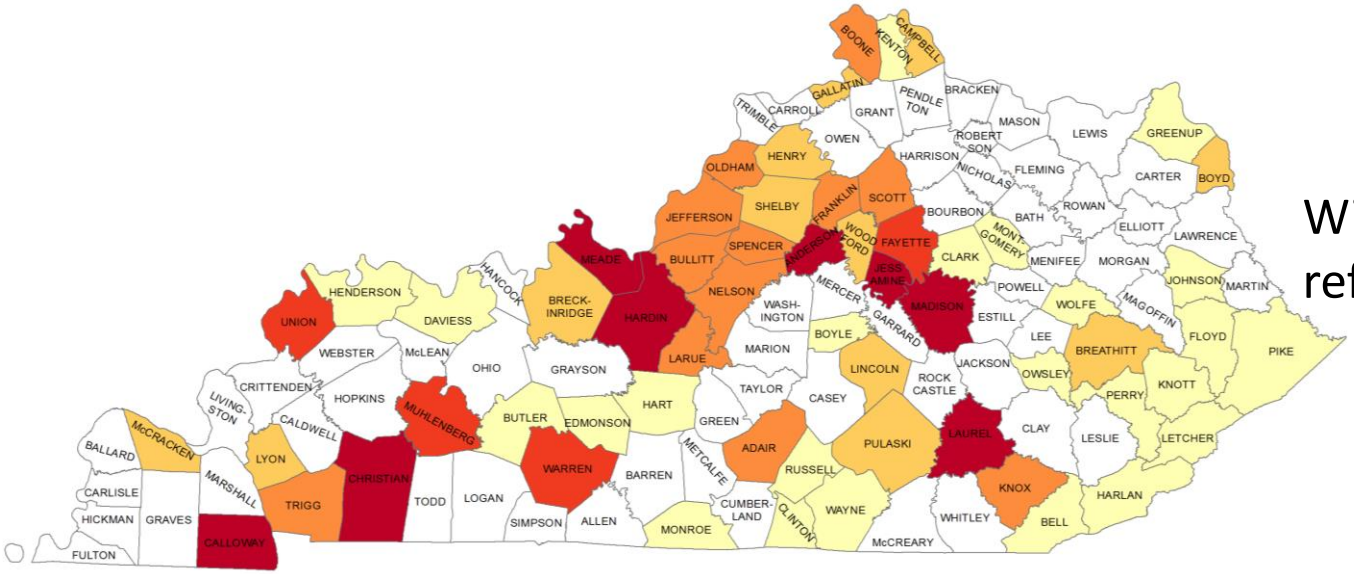
Want a method that is generalizable to any geography

Block data is only available for a few variables

Areal Weighting Results (RMSE)



Without spatial refinement



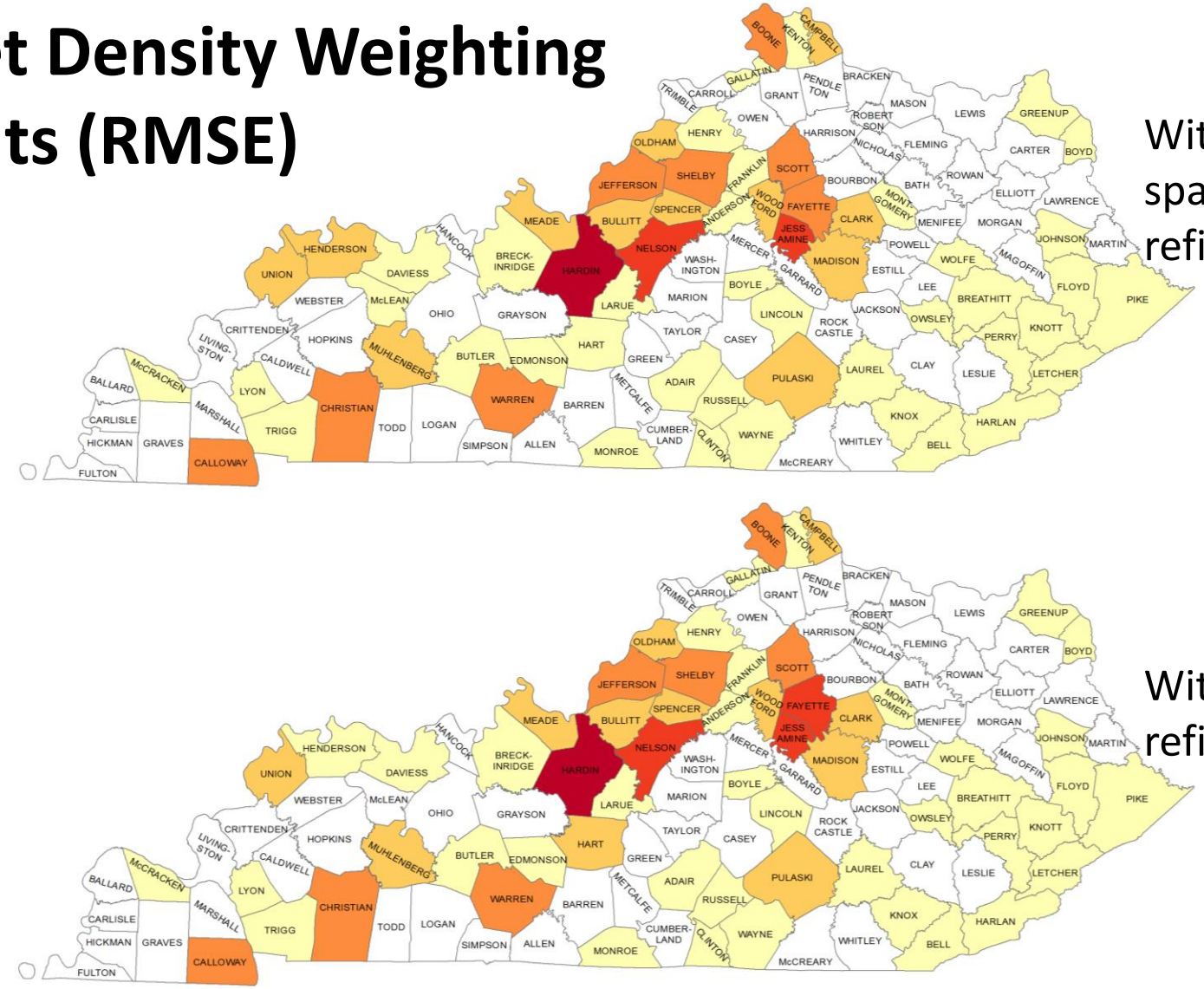
With spatial refinement

1 - 249 | 250 - 499 | 500 - 749 | 750 - 999 | 1000 and Over | No Changed Tracts

KSDC
Kentucky State Data Center

UNIVERSITY OF LOUISVILLE

Target Density Weighting Results (RMSE)



Without spatial refinement

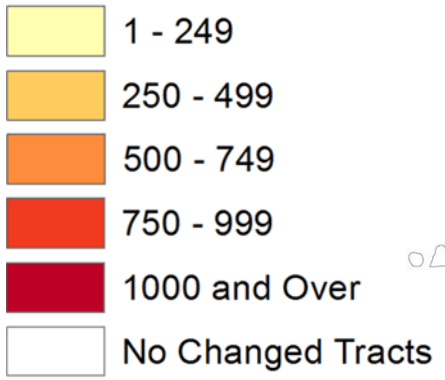
With spatial refinement

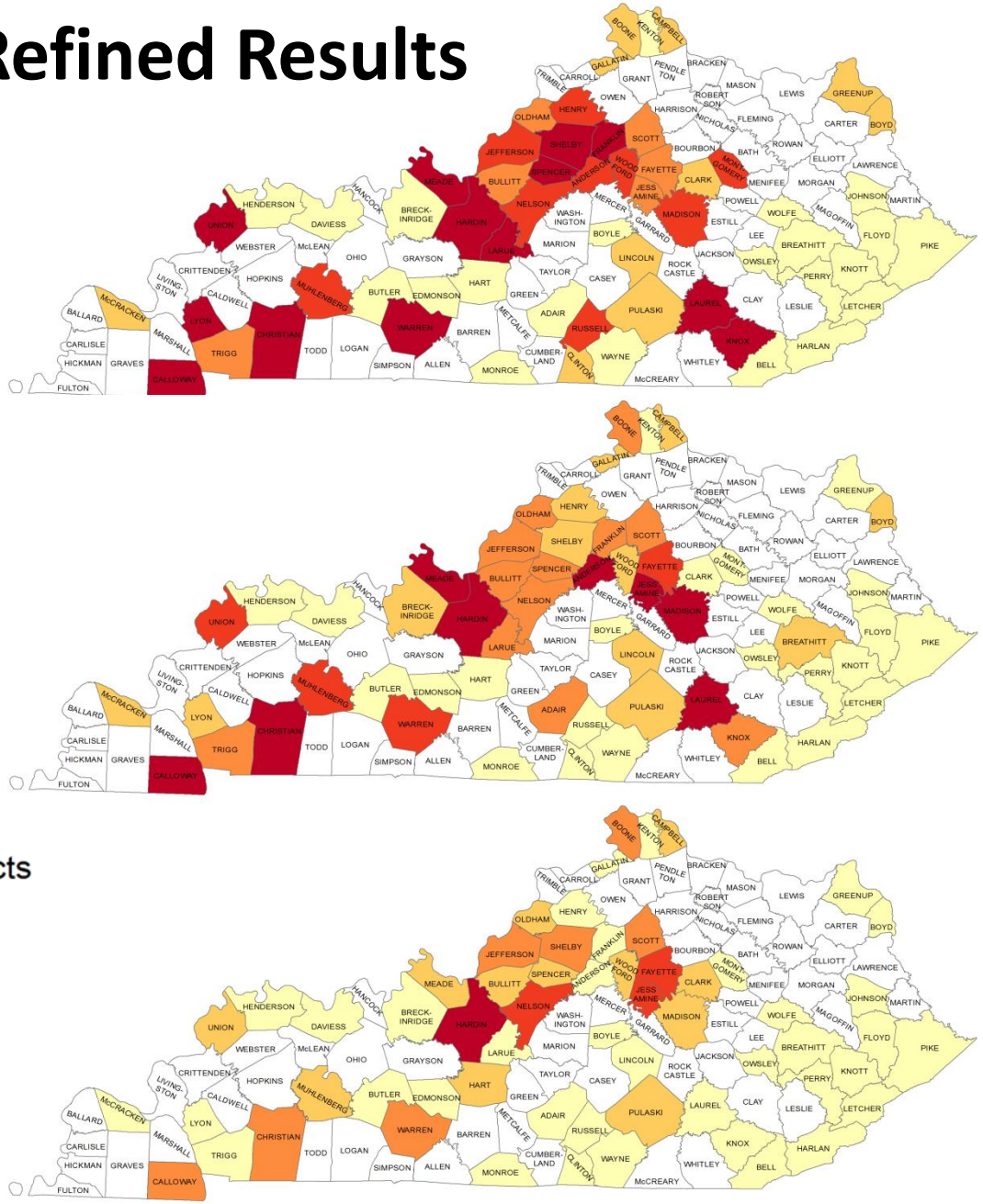
1 - 249 | 250 - 499 | 500 - 749 | 750 - 999 | 1000 and Over | No Changed Tracts

KSDC
KENTUCKY STATE DATA CENTER

UNIVERSITY OF LOUISVILLE

Spatially Refined Results (RMSE)





Expectation maximization

Areal weighting

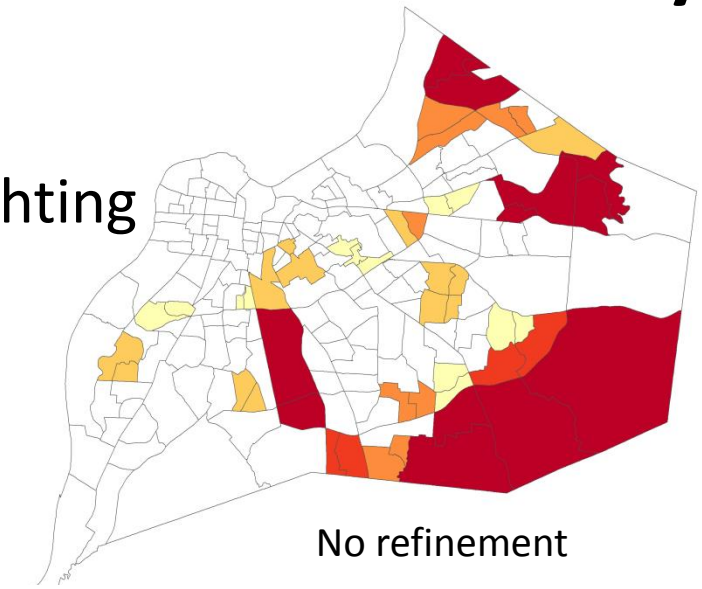
Target density weighting

KSDC
Kentucky State Data Center

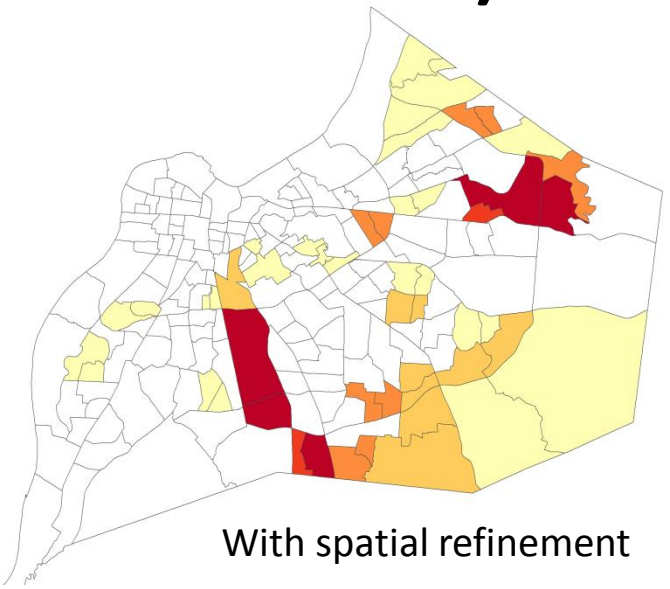
University of Louisville

Results for Jefferson County (Absolute Error)

Areal weighting



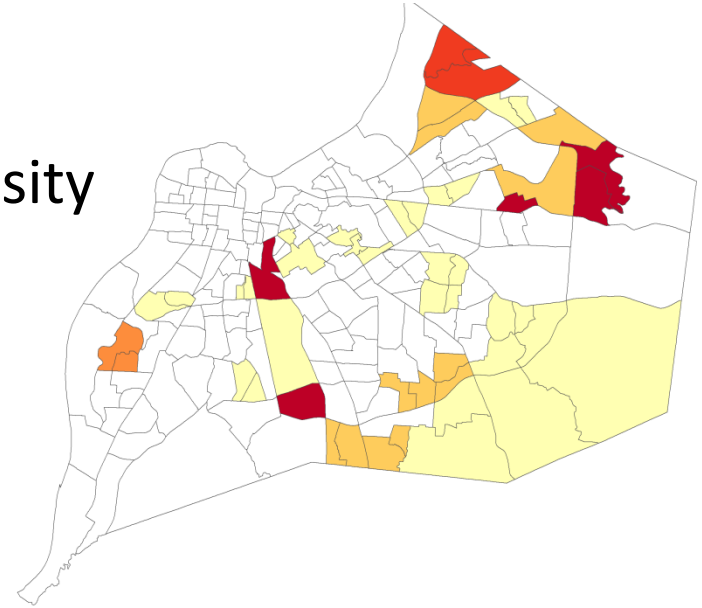
No refinement

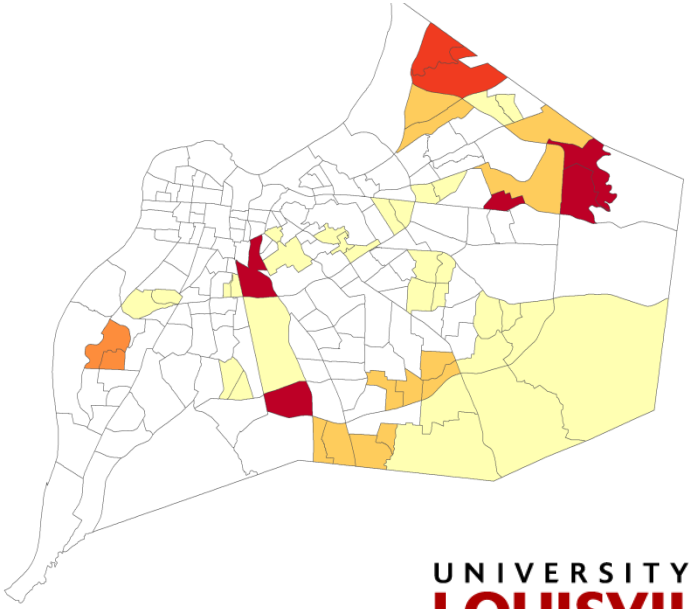


With spatial refinement

1 - 249 | 250 - 499 | 500 - 749 | 750 - 999 | 1000 and Over | No Change

Target density weighting





KSDC
KENTUCKY STATE DATA CENTER

UNIVERSITY OF LOUISVILLE

Results for Kentucky

2000

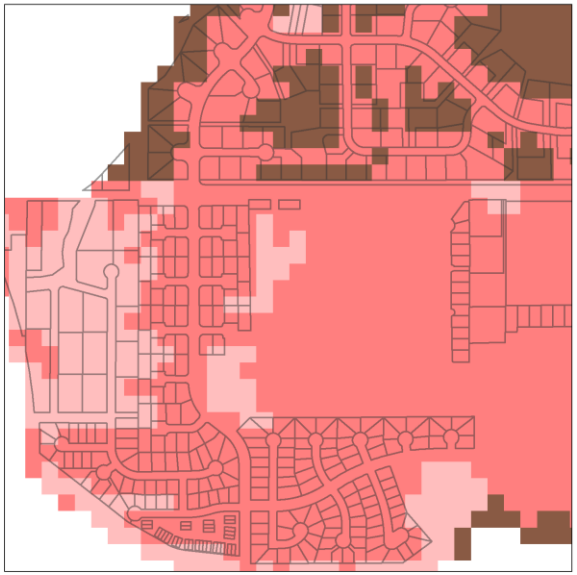
	AW	AW-Refined	TDW	TDW-Refined	EM
Mean Error	1046	537	346	337	552
Median Error	645	330	181	171	333
RMSE	1663	933	546	542	865
n=319					

1990

	AW	AW-Refined	TDW	TDW-Refined	EM
Mean Error	783	353	426	393	355
Median Error	364	224	219	198	198
RMSE	1390	569	730	713	559
n=435					

Limitations and Future Directions

- Results depend on the accuracy of ancillary information
  - Developed land in rural areas may be under-estimated
  - Temporal instability
  - Questionable using NLCD at small resolution
- Other ancillary information could be integrated
  - Parcel data(?)



NLCD 2011

- Open Space Developed (21)
- Low/Medium Intensity Developed (22/23)
- All Other Classes

KSDC
KENTUCKY STATE DATA CENTER

UNIVERSITY OF LOUISVILLE

Integrating Areal Interpolation and Dasymetric Refinement to Resolve Temporal Incompatibilities in Zoning Systems

Sarah Ehresman
sarah.ehresman@louisville.edu

Matt Ruther, Ph.D.
matthew.ruther@louisville.edu

Kentucky State Data Center
http://ksdc.louisville.edu
502.852.7990

KSDC
KENTUCKY STATE DATA CENTER

UNIVERSITY OF LOUISVILLE