Identification of Regional Subcenters Using Spatial Data Analysis for Estimating Traffic Volume

Fang Zhao and Nokil Park
Lehman Center for Transportation Research
Department of Civil & Env. Engineering
Florida International University
Miami, FL

Introduction

- To predict traffic volume, a roadway's function, local land use pattern, and spatial patterns of land use should be considered
- Significant activity centers such as CBD or population centers influence traffic volume
- It is difficult to determine urban population and employment centers
- Previous methods for identifying subcenters require detailed local knowledge

Objectives

- Develop a systematic method to determine activity centers or subcenters
- Develop regional accessibility measures based on the subcenters identified

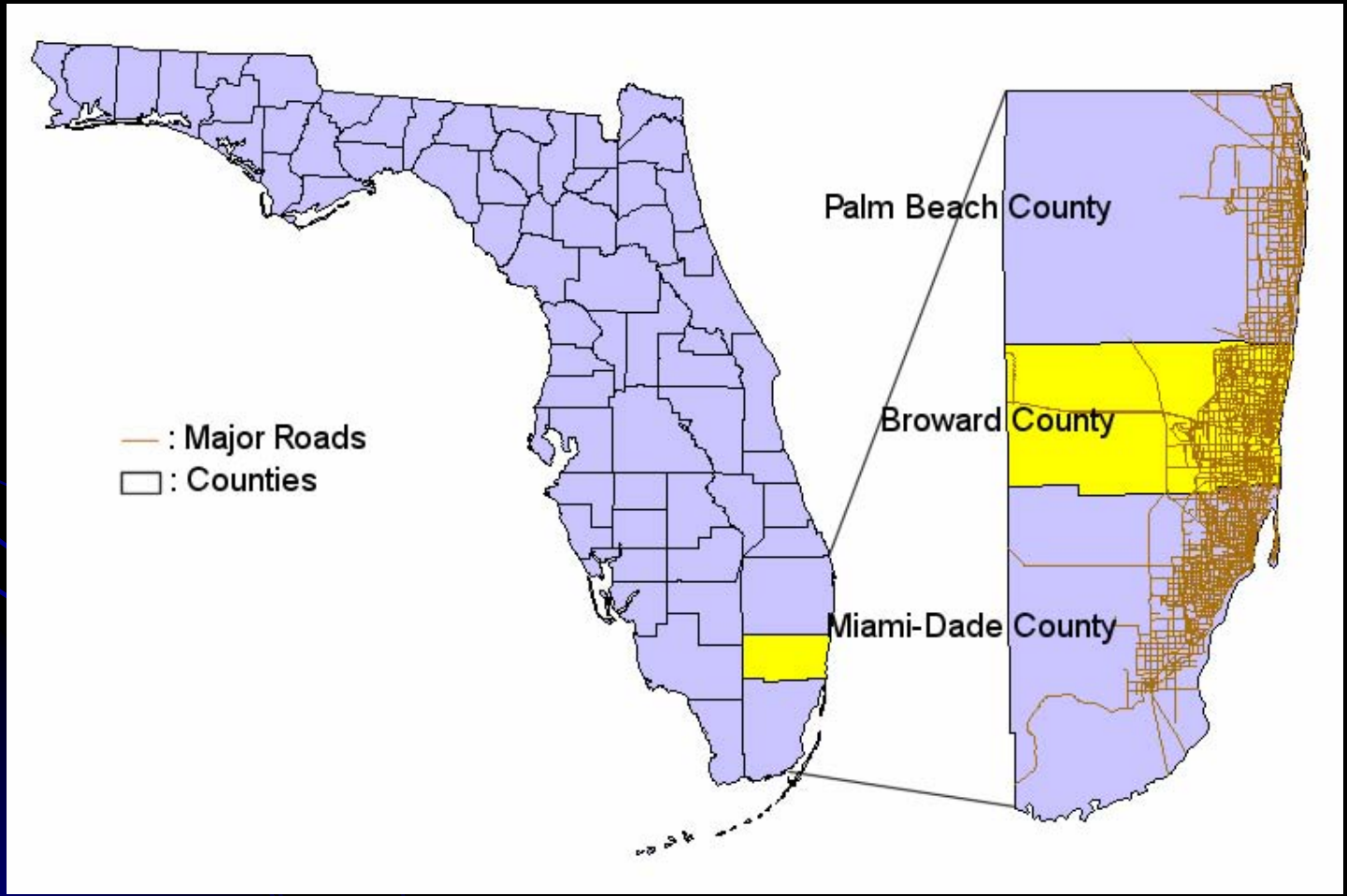
Background

- Dividing an area into zones and defining a subcenter as a zone whose measure of employment concentration (gross employment density or employment-population ratio, with employment) is higher than adjacent zones (McDonald 1987).
- A contiguous set of zones, each with an employment > 10 employees per acre, and together with a minimum of 10,000 employees (Giuliano and Small 1991)
- Population and employment regional centers estimated based on distribution and density maps and three-dimensional surface maps of population and employment at TAZ level, and contour maps generated from the surface map (Zhao and Chung 2001).

About This Study

- Define subcenters based on
  - Moran's I and LISA
  - GIS tool
- Replace original accessibility measure with the new one
- Calibrate regression model using both accessibility measures
- Compare regression models with different regional subcenters
- Broward County, FL as the study area - urbanized with about 1.5 million population

Study Area


Palm Beach County
Broward County
Miami-Dade County
: Major Roads
: Counties

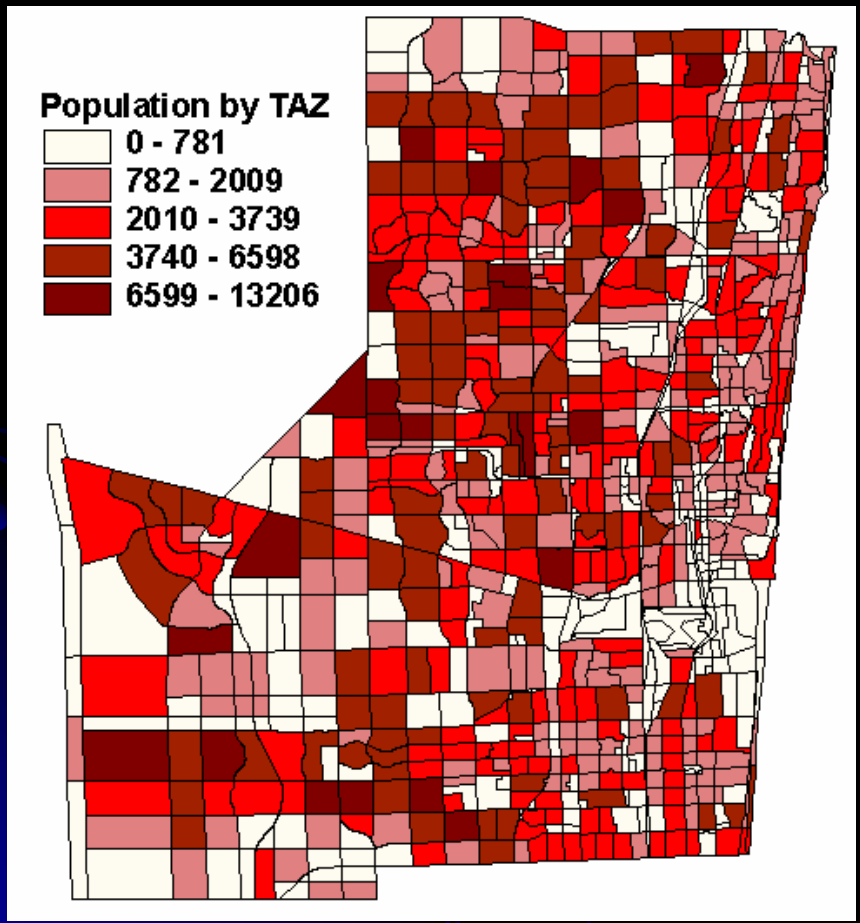
Moran Scatter Plot

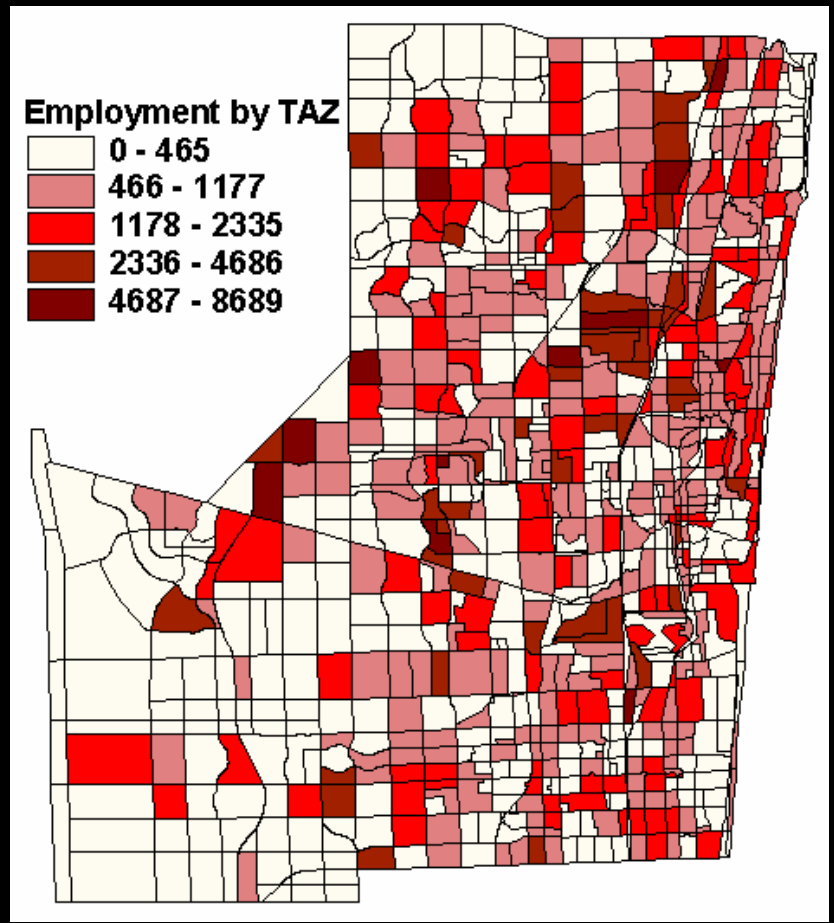
- Moran's *I* reflects the level of the spatial autocorrelation
- Moran scatterplot helpful to visualize four types of spatial autocorrelation in a quadrant of the scatterplot.
- Quadrant HH (LL) refers to an observation with high (low) value surrounded by observations with high (low) values
- Quadrant LH (HL) refers to an observation with low (high) value surrounded by observation with high (low) values

Local Indicators of Spatial Association (LISA)

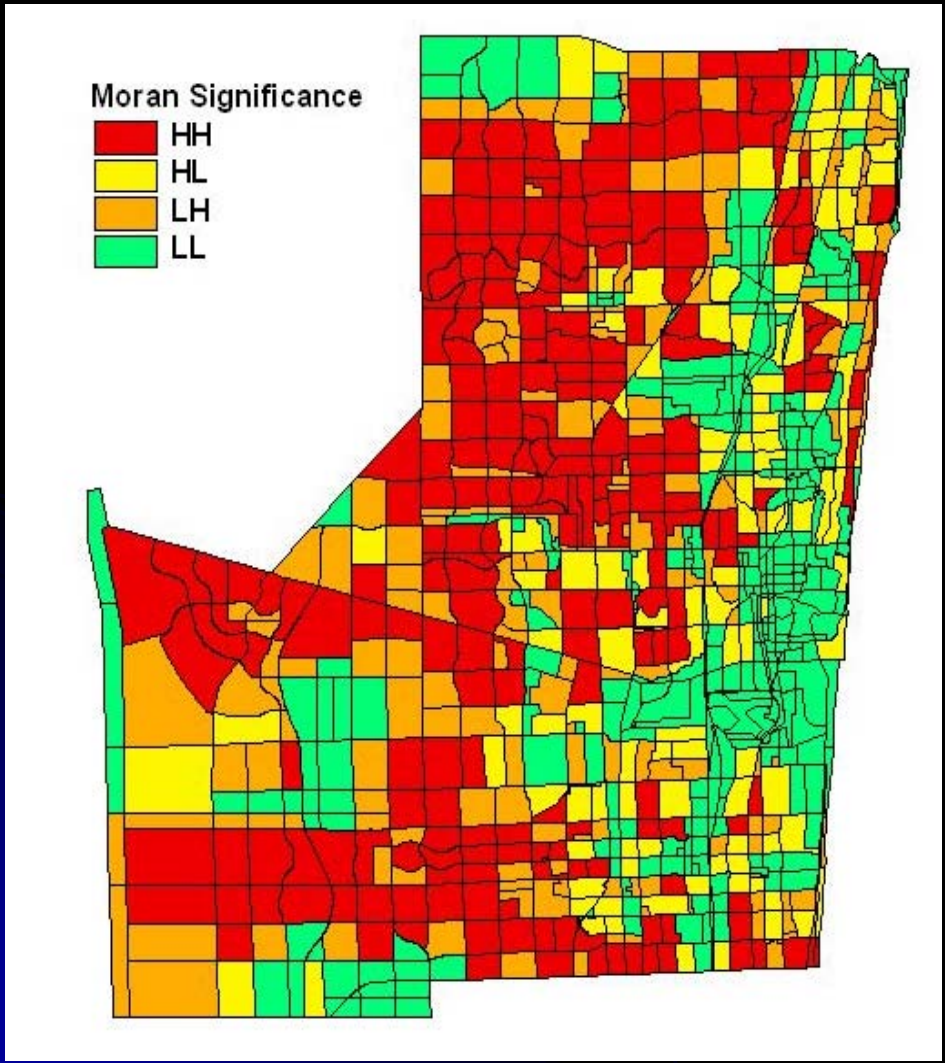
- LISA statistics (Anselin 1995) was introduced to assess the significance of spatial autocorrelation
- LISA indicates the level of spatial autocorrelation at the local scale
- LISA refers to the local version of Moran's *I*
- A positive value for LISA indicates spatial clustering of similar values (can be high or low)
- A negative value indicates spatial clustering of dissimilar values between a observation and those in its neighborhood.

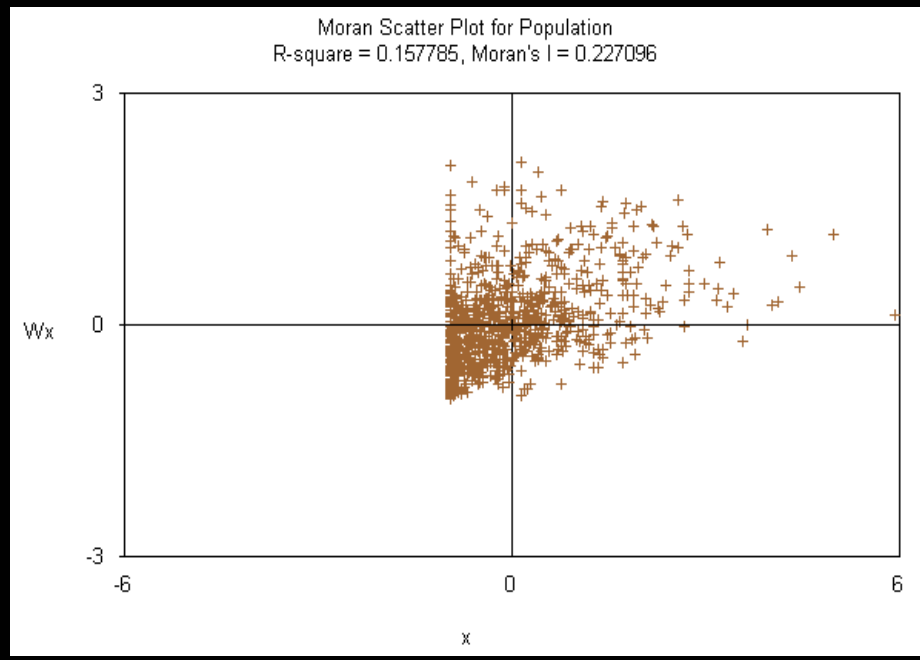
Population and Employment Distribution



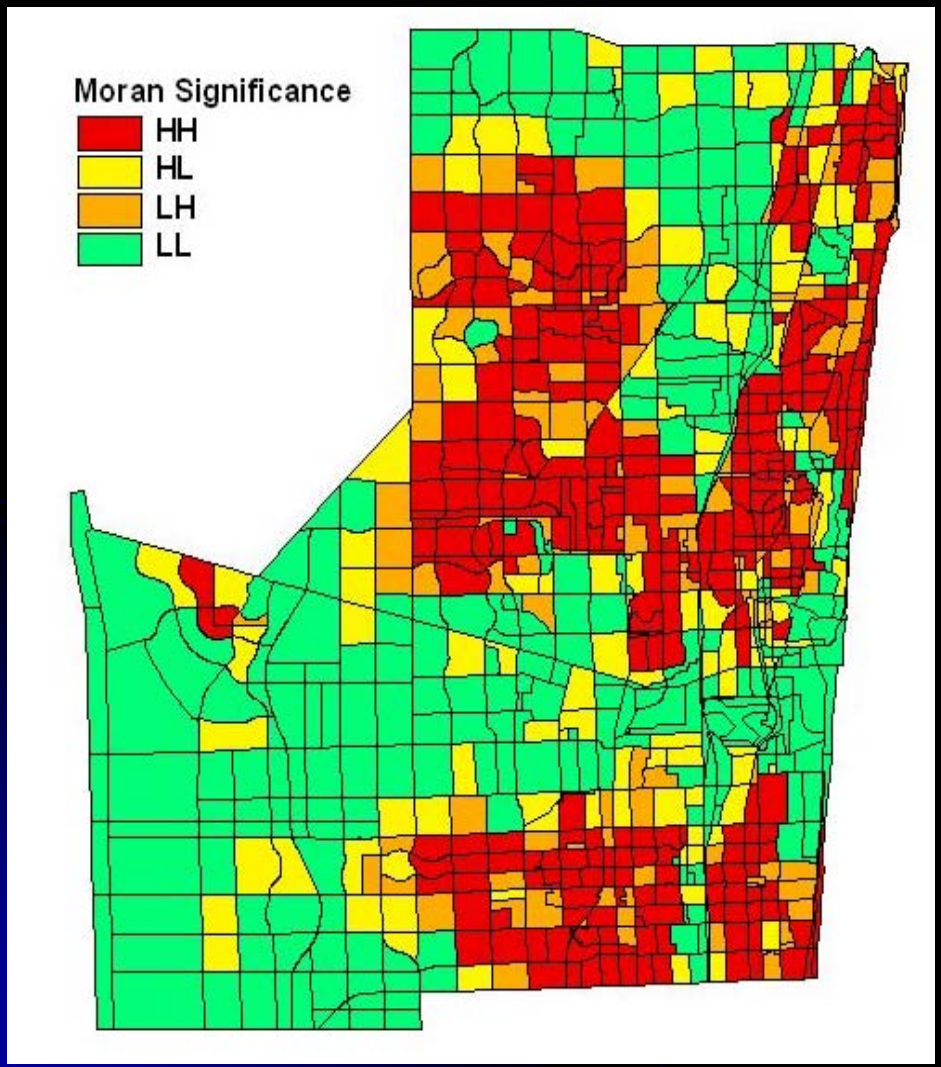


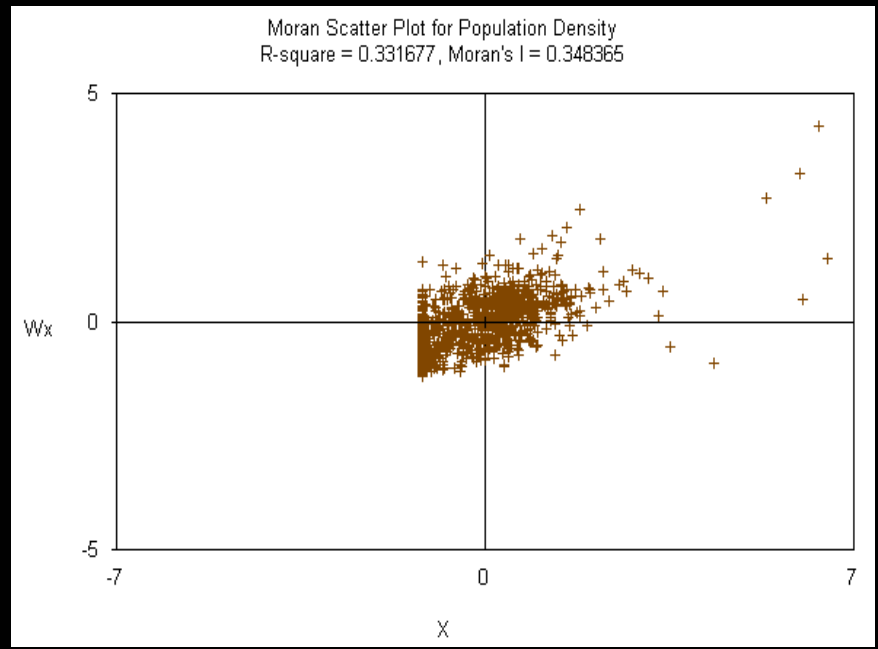
Moran Significance Map and Moran Scatter Plot for Population



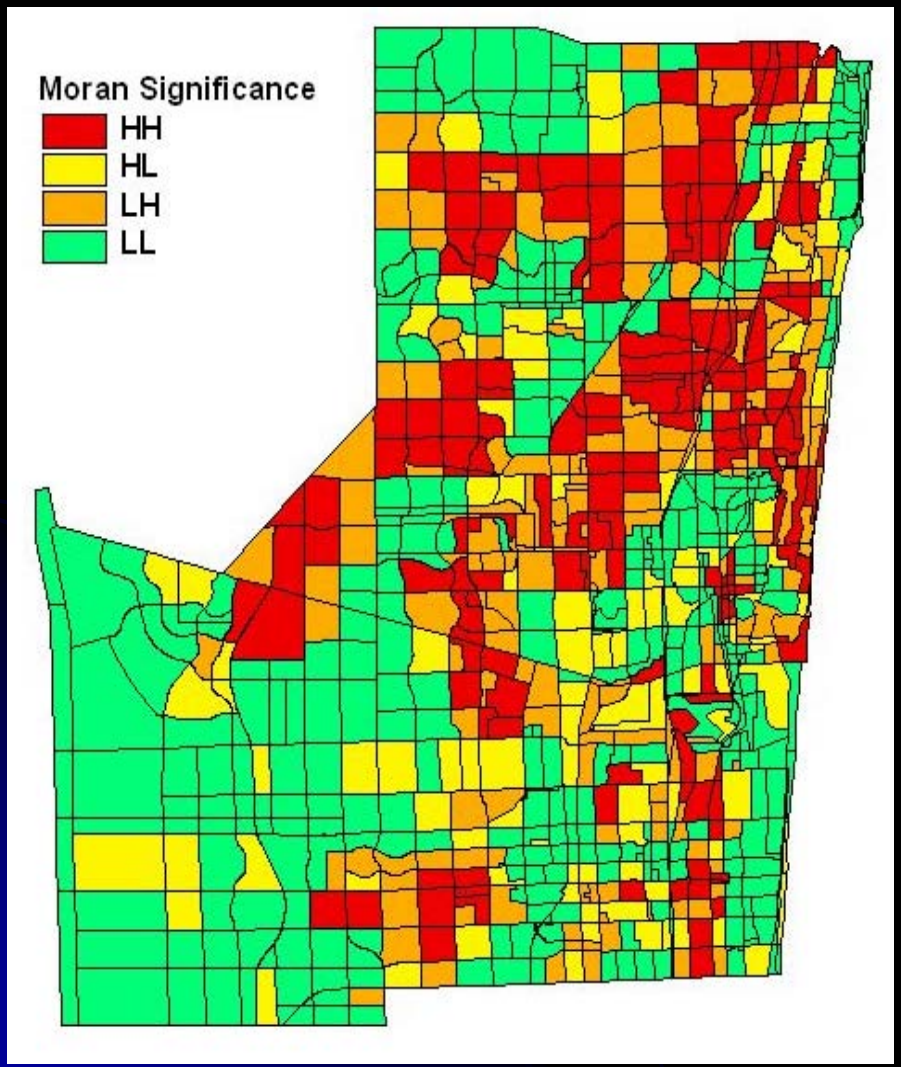


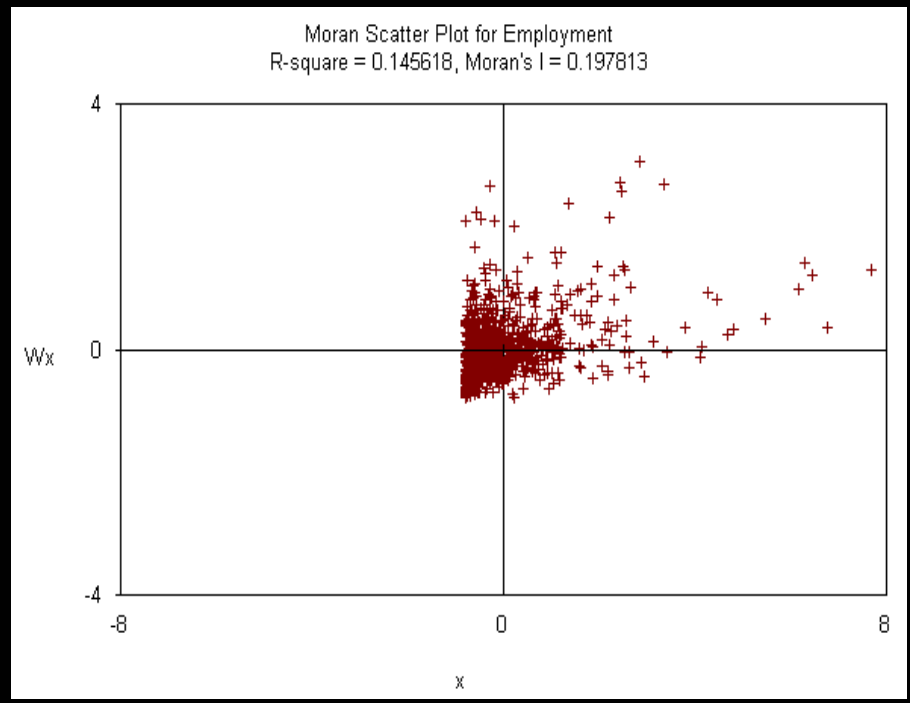
Moran Significance Map and Moran Scatter Plot for Population Density



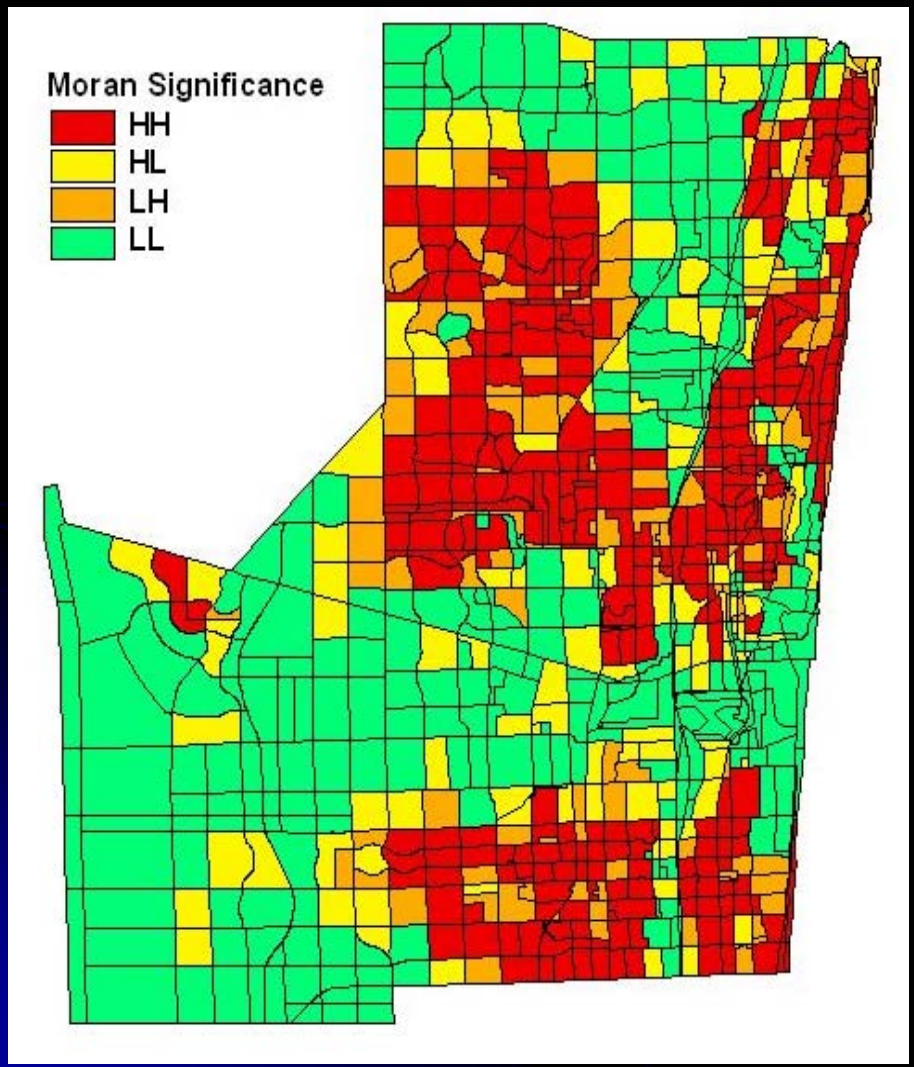


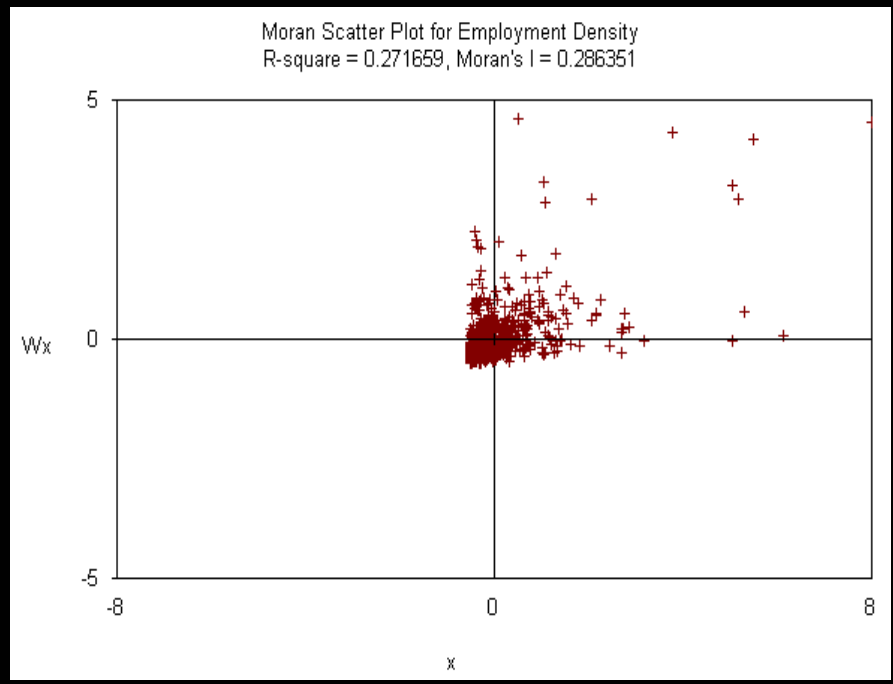
Moran Significance Map and Moran Scatter Plot for Employment





Moran Significance Map and Moran Scatter Plot for Employment Density

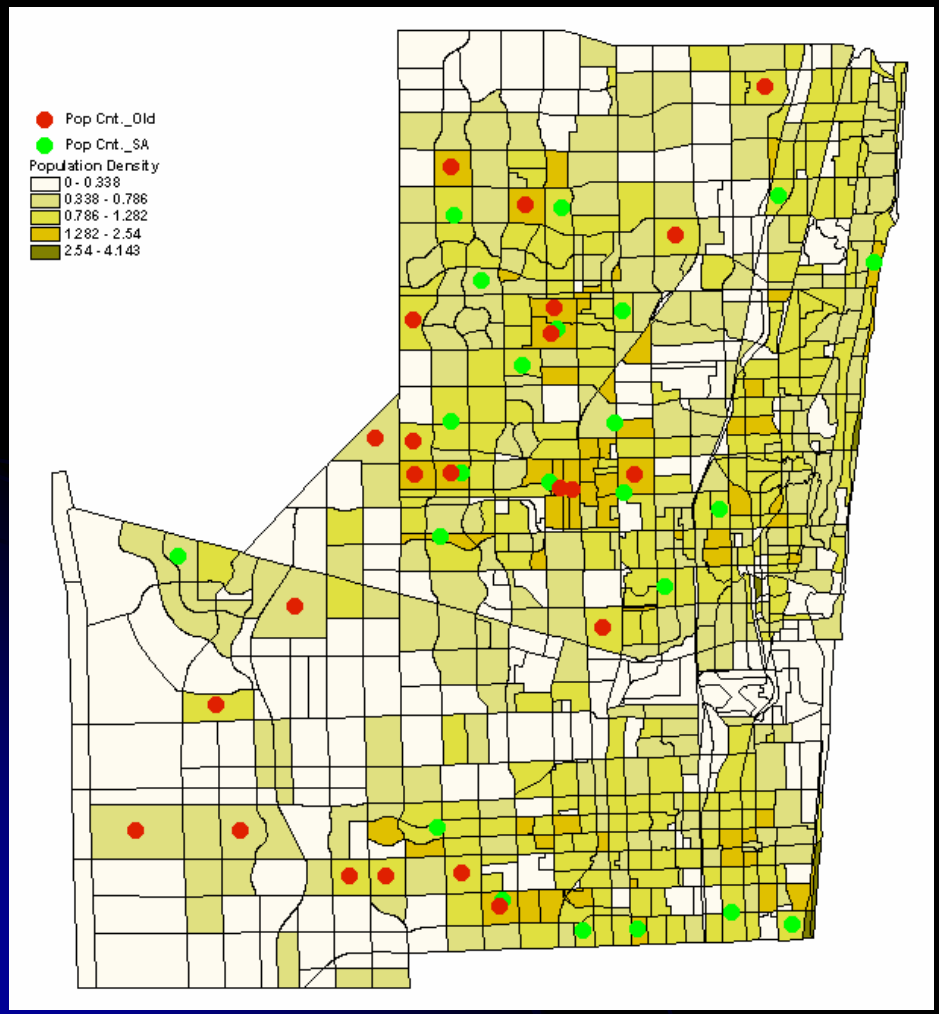


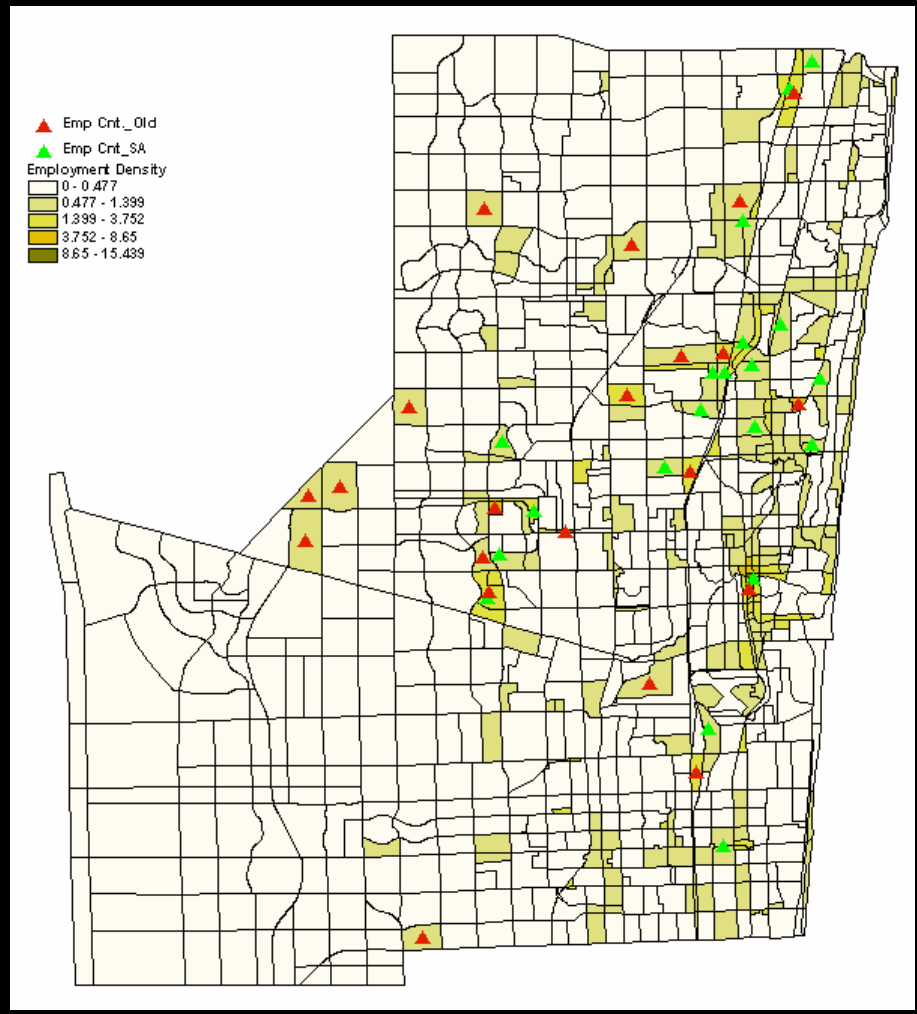


Processes for Identification of Subcenters

1. Selecte TAZs with significant HH spatial association for both population/employment and population/employment density variables
2. Determined the cut-off value : population/employment of TAZ to include more than 50% of the total population/employment of County
3. Merge the contiguous TAZs that have at least one common side
4. Create geographical centers of the polygon

Population and Employment Subcenters





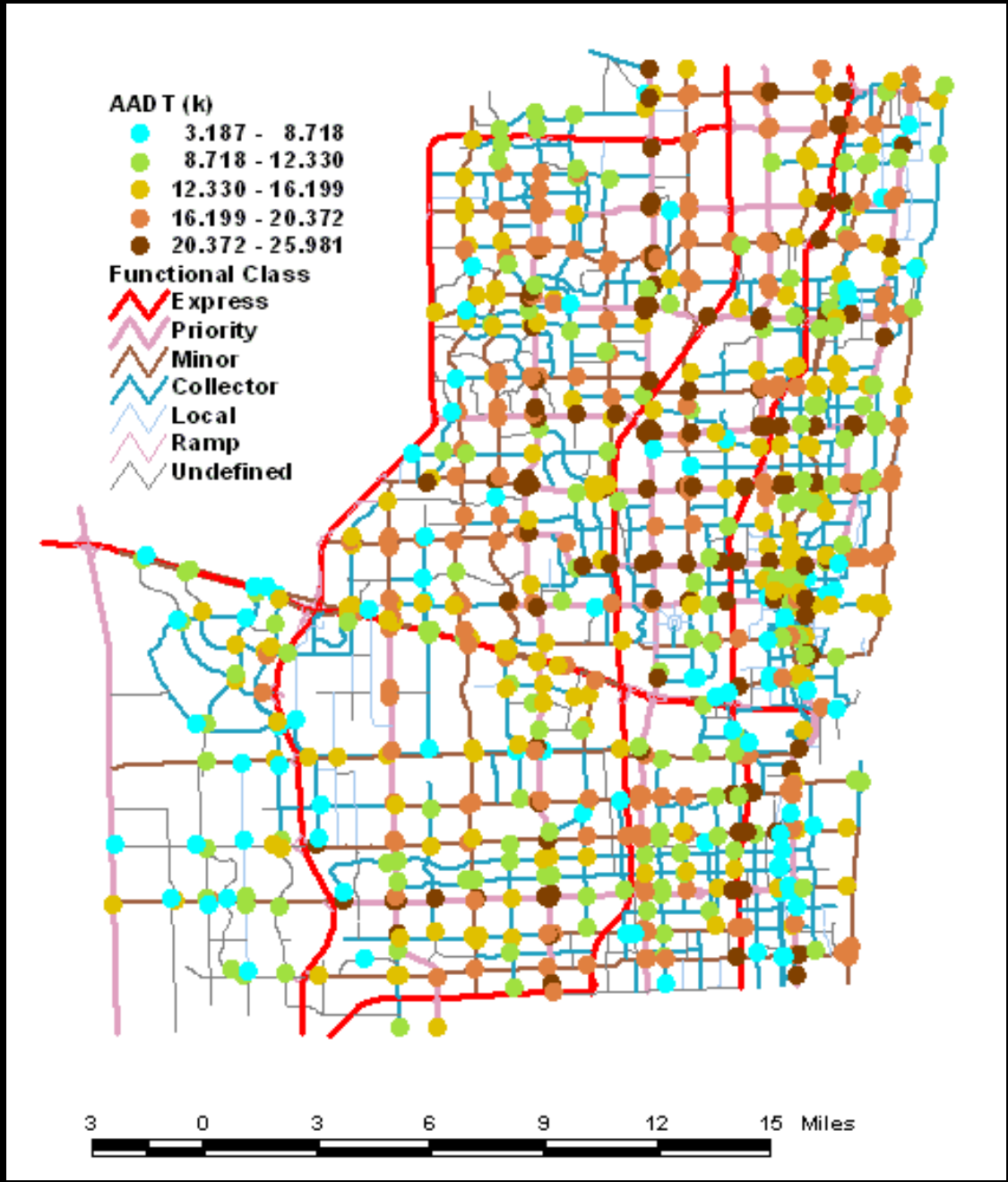
Regional Accessibility

- Regional accessibility to population and employment centers (*RPEAccess*)

$$RPEAccess_k = \left(\sum_{i=1}^{N_P} P_i e^{-0.0954 t_{ki}} \right) \left(\sum_{j=1}^{N_E} E_j e^{-0.0954 t_{kj}} \right)$$

*P_i, E_j = the population/employment of *i*, *j*-th population/employment center
*t_{ki} , t_{kj} = the network travel time from count station *k* to the *i*, *j*-th population/ employment center
*N_P , N_E = the number of population/ employment centers

AADT and Functional Classification



Data Compilation and Processing

- Dependent Variable - 2000 AADT
  - 723 for model development
  - 145 on principal arterials, 235 on minor arterials, 289 on collectors, 54 on local roads
- Predictors
  - roadway characteristics (number of lanes)
  - land uses (population, employment)
  - Accessibility to subcenters

Land Uses

- Population (*POPBUF*) and employment (*EMPBUF*) size around a count station

Location	Buffer Size (miles)			
	Principal Arterial	Minor Arterial	Collector	Local / Undefined
East	1.0	0.5	0.25	0.12
Central	1.5	0.5	0.25	0.12
Central West	2.0	1.0	0.50	0.25
West	3.0	2.0	0.50	0.25

Expressway Accessibility

- Direct access to expressway (*DIRECTAC*) as a binary variable
  - 1 – for a count station on a road that connects to an expressway,
  - 0 – otherwise

Estimation Models

- Model_Old

  AADT = -2.775 + 5.074 LANE + 6.460 DIRECTAC
  + 3.019 EMPBUF + 1.437 POPBUF
  + 2.219 RPEAccess_Old

- Model_SA

  AADT = -5.044 + 4.959 LANE + 7.111 DIRECTAC
  + 2.389 EMPBUF + 2.228 POPBUF
  + 2.571 RPEAccess_SA

Model Performance

Model	Model_Old	Model_SA
R-square	0.757	0.764
Adj R-square	0.755	0.763
MSE	59.333	57.507
Mean Error(%)	41.49	40.65
Total Error(%)	-0.0054	0.00082

Conclusions

- New method is systematic and automatic
- Spatial Autocorrelation is a useful tool to identify regional subcenters
- Model_SA is able to better explain the variation in the data than Model_Old