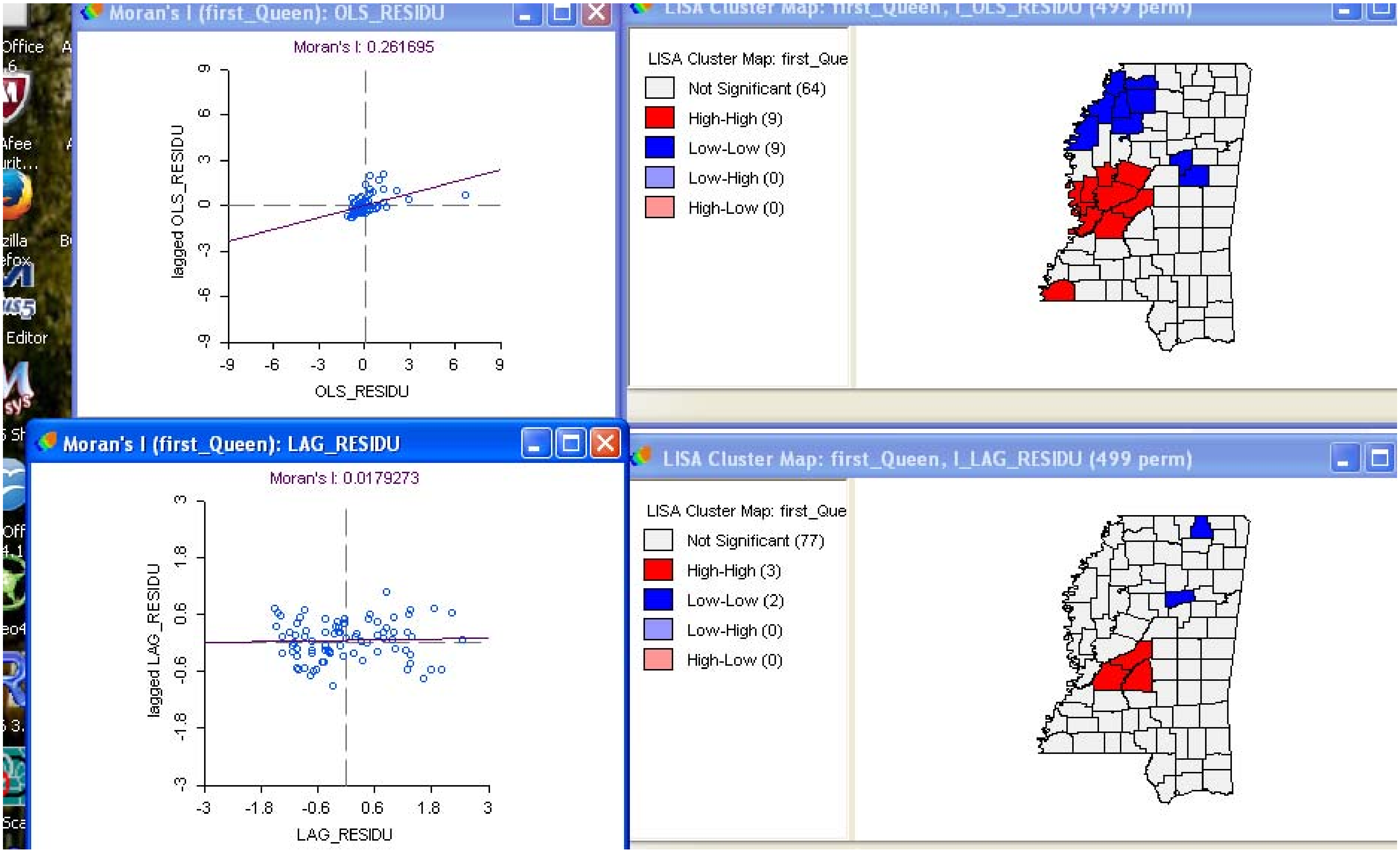
Geographically Weighted Regression (GWR): An Introduction

Spatial Regression

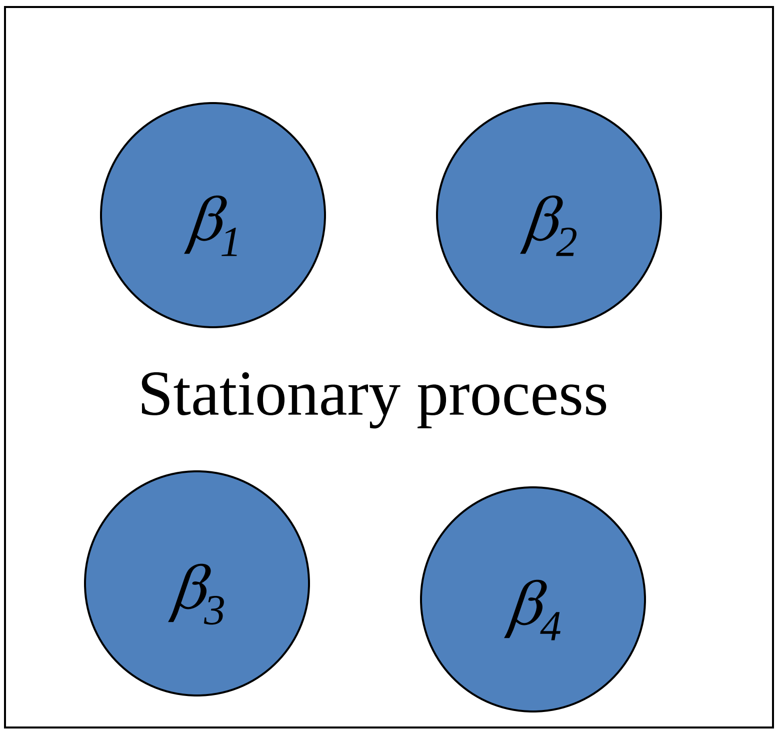
- *Appropriate when OLS assumptions are violated due to spatial distribution of values….*
  - Independence
  - *Linearity*
  - *Normality*
  - Constant Variance

- *Lag:* $y = \rho W y + X\beta + \varepsilon$
- *Error:* $y = X\beta + \varepsilon$; *where:* $\varepsilon = \lambda W \varepsilon + \xi$


Moran's I (first_Queen): OLS_RESIDU
Moran's I: 0.261695
lagged OLS_RESIDU
OLS_RESIDU
LISA Cluster Map: first_Queen, I_OLS_RESIDU (499 perm)
LISA Cluster Map: first_Que
Not Significant (64)
High-High (9)
Low-Low (9)
Low-High (0)
High-Low (0)
Moran's I (first_Queen): LAG_RESIDU
Moran's I: 0.0179273
lagged LAG_RESIDU
LAG_RESIDU
LISA Cluster Map: first_Queen, I_LAG_RESIDU (499 perm)
LISA Cluster Map: first_Que
Not Significant (77)
High-High (3)
Low-Low (2)
Low-High (0)
High-Low (0)

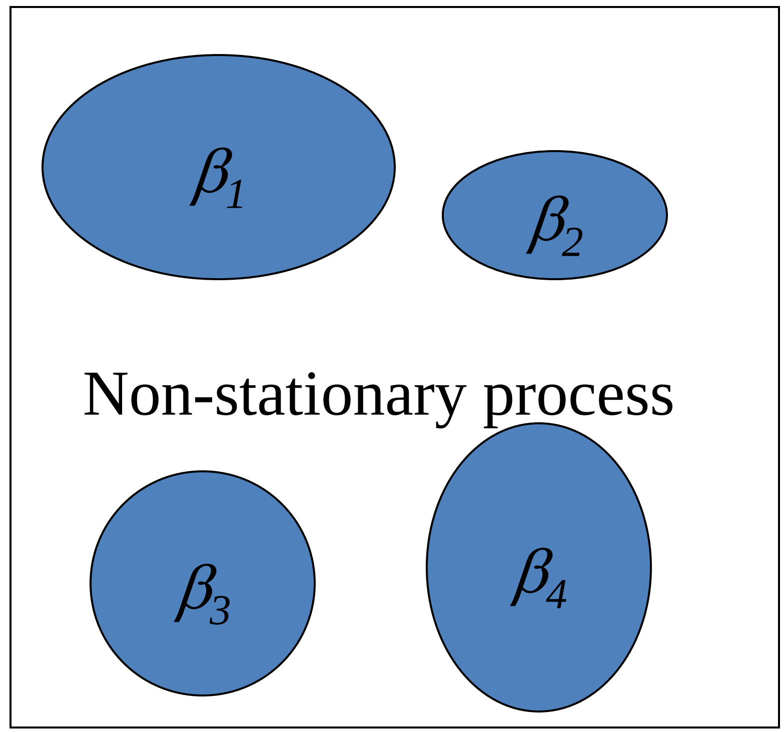
Stationary v.s non-stationary

$y_i = \beta_0 + \beta_1 x_{1i}$



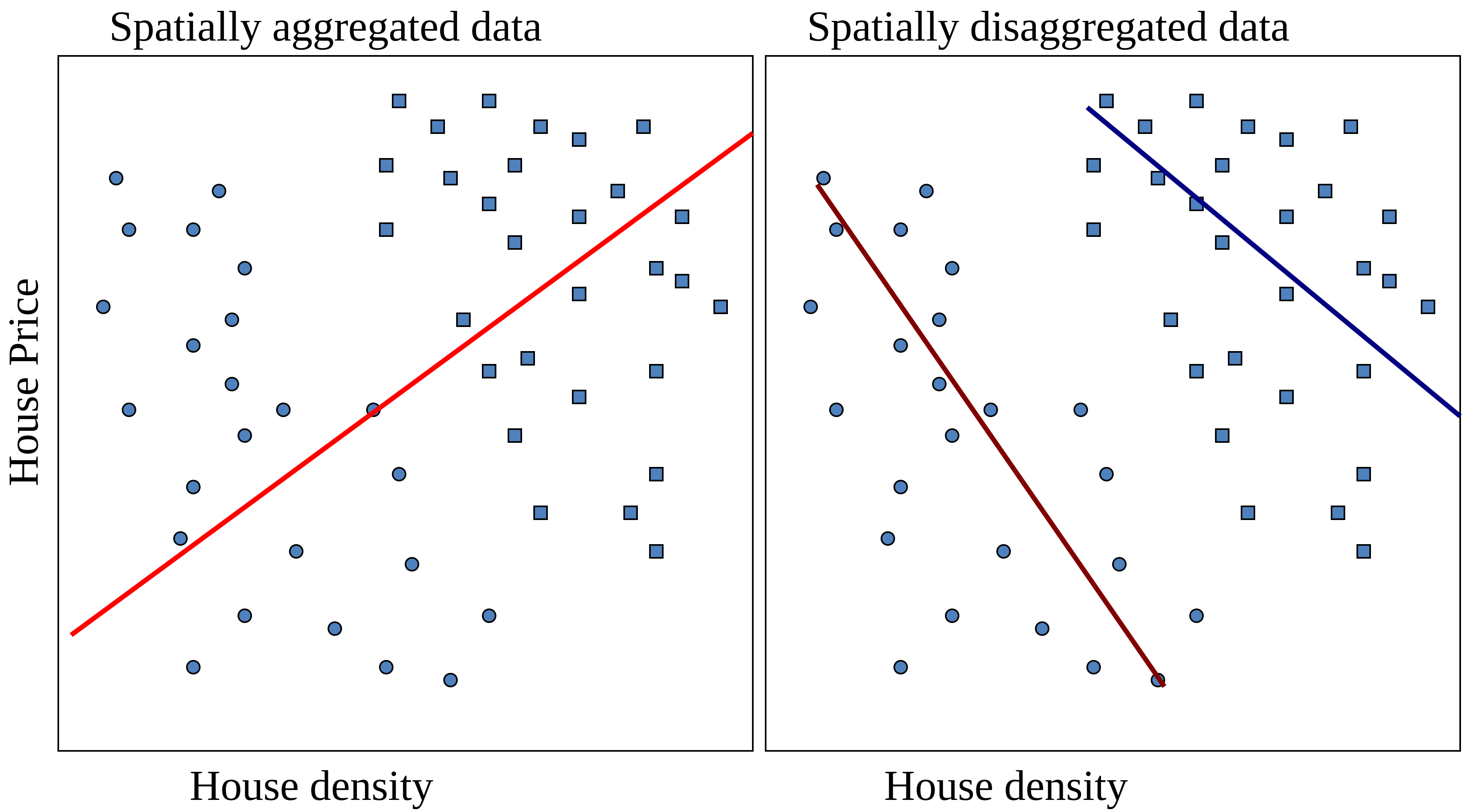
Assumed

$y_i = \beta_{i0} + \beta_{i1} x_{1i}$



More realistic

Simpson’s paradox



Some definitions

- Spatial non-stationarity: the same stimulus provokes a different response in different parts of the study region

Some definitions

- Global models: are statements about processes which are assumed to be stationary and as such are *location independent*
  - *General approach to social science modeling*
    - *Exceptions include families of fixed effect models (most notably HLM/multilevel models)*

- Local models: spatial decompositions of global models, the results of local models are *location dependent* – a characteristic we usually anticipate from geographic (spatial) data and, more generally, social data

Regression

- Regression establishes relationship among a dependent variable and a set of independent variable(s)
- A typical linear regression model looks like:
- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots + \beta_n x_n + \varepsilon$
  - This is an example of a Global

GWR

- Addresses the non-stationarity directly
  - Allows the relationships to vary over space, i.e., βs do not need to be everywhere the same
  - This is the essence of GWR, in the linear form:
  - $Y_i = \beta_{i0} + \beta_{i1}x_{1i} + \beta_{i2}x_{2i} + \ldots\ldots + \beta_{in}x_{ni} + \varepsilon_i$
  - Instead of remaining the same everywhere, βs now vary in terms of locations (*i*)

Global v.s. local statistics

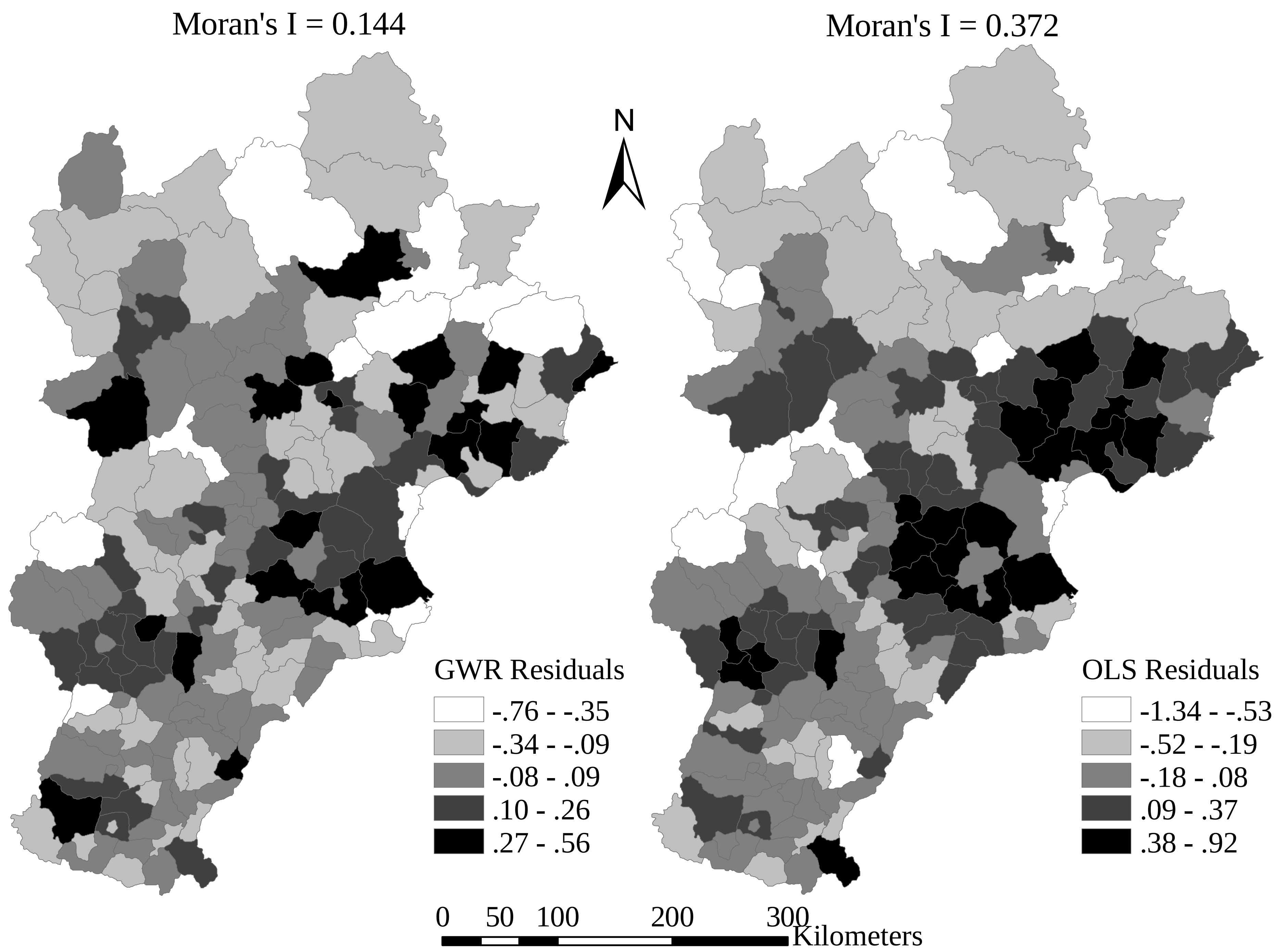
- Global statistics
  - Similarity across space
  - Single-valued statistics
  - Not mappable
  - GIS “unfriendly”
  - Search for regularities
  - aspatial
- Local statistics
  - Difference across space
  - Multi-valued statistics
  - Mappable
  - GIS “friendly”
  - Search for exceptions
  - spatial

Where does GWR fit?

Residuals from GWR are generally much lower and usually much less spatially dependent

- GWR models give much better fits to data, EVEN accounting for added model complexity and number of parameters (decrease in degrees of freedom)

- GWR residuals are usually much less spatially dependent


Moran's I = 0.144
Moran's I = 0.372
N
GWR Residuals
-.76 - -.35
-.34 - -.09
-.08 - .09
.10 - .26
.27 - .56
OLS Residuals
-1.34 - -.53
-.52 - -.19
-.18 - .08
.09 - .37
.38 - .92
0
50
100
200
300
Kilometers

Where does GWR fit?

GWR as a "spatial microscope"

- A series of bandwidths are examined before the selection of an optimal bandwidth. This statistically identifies the optimal "catchment" areas in which the local regression models will be estimated.

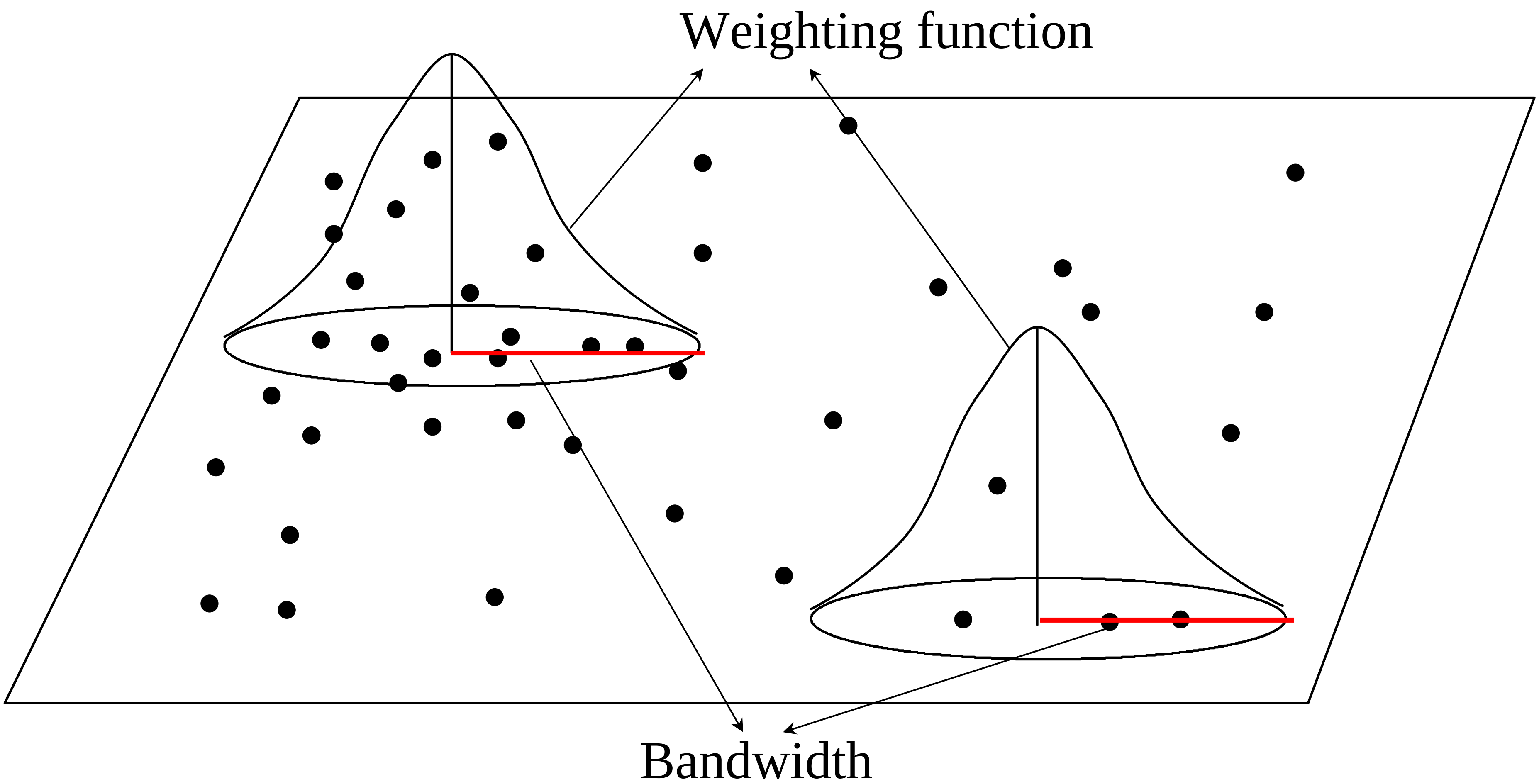
  - In comparison to Geoda… where the neighborhood must by manually developed once at a time to compare different neighborhood definitions.

Calibration of GWR

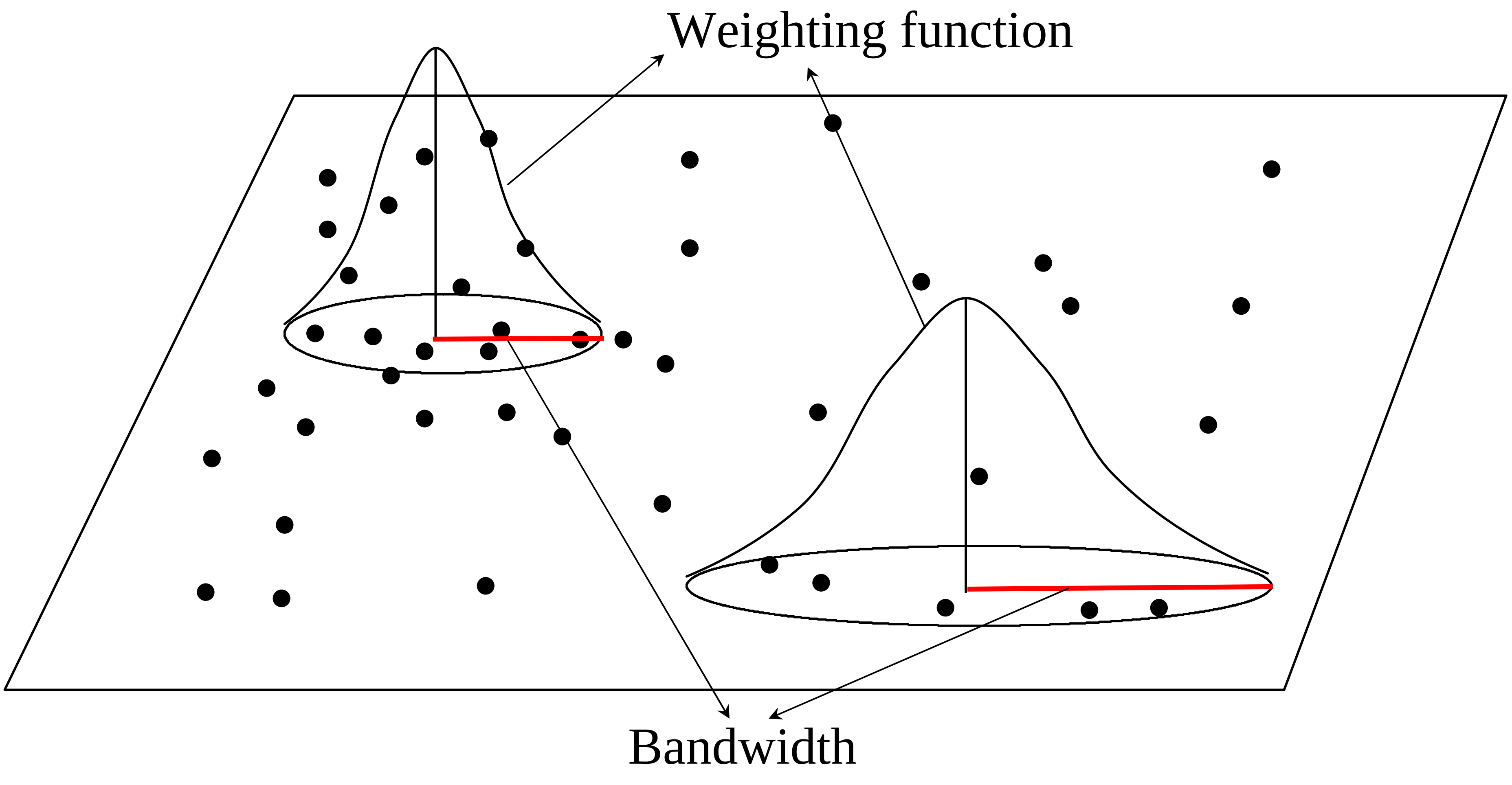
– Weights are attached with locations
  - Can be fixed by user (not recommended)
  - Fixed by system
  - Adaptive by system

– Most schemes tend to be Gaussian or Gaussian-like reflecting the type of dependency found in most spatial processes

Fixed weighting scheme



Adaptive weighting schemes



Calibration

- Surprisingly, the results of GWR appear to be relatively insensitive (robust) to the choice of weighting functions as long as it is a continuous distance-based function (Gaussian or Gaussian-like functions)
  - Additionally, in most cases the results from Fixed vs. Adaptive weighting schemes also do not vary significantly

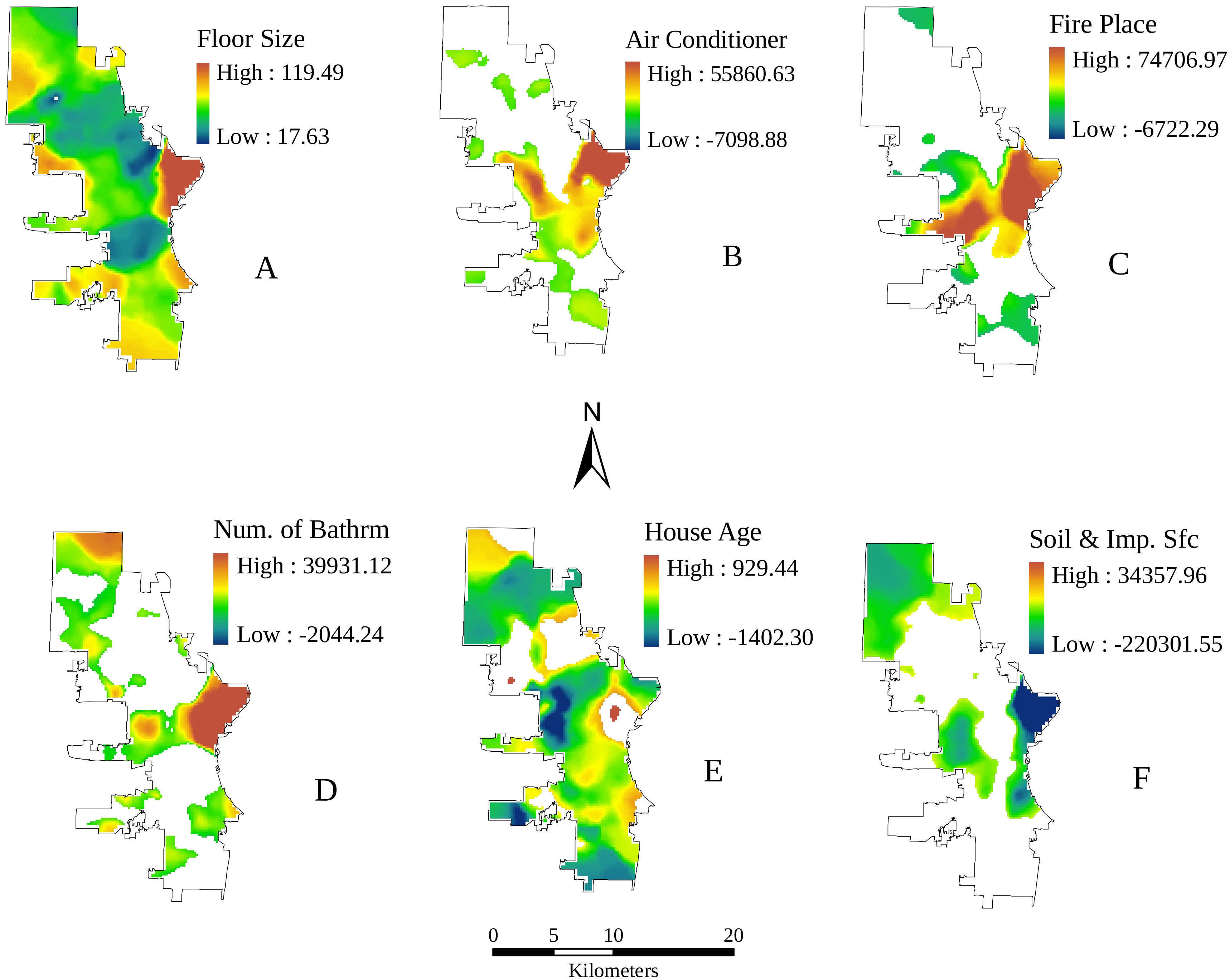
An example

- Housing demand model in Milwaukee
  - Data: MPROP 2004 – 3430+ samples used
    - Master Property Record
  - Dependent variable: the assessed value (price)
  - Independent variables: air conditioner, floor size, fire place, house age, number of bathrooms, soil and Impervious surface (remote sensing acquired)

The global model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18944.05	4112.79	4.61	4.25e-06
Floor Size	78.88	2.00	39.42	<2e-16
House Age	-508.56	33.45	-15.20	<2e-16
Fireplace	14688.13	1609.53	9.13	<2e-16
Air Conditioner	13412.99	1296.51	10.35	<2e-16
Number of Bathrooms	19697.65	1725.64	11.42	<2e-16
Soil&Imp. Surface	-27926.77	5179.42	-5.39	7.44e-08

Residual standard error: 35230 on 3430 degrees of freedom
Multiple R-Squared: 0.6252, Adjusted R-squared: 0.6246
F-statistic: 953.7 on 6 and 3430 DF, p-value: < 2.2e-16
Akaike Information Criterion: 81731.63


Floor Size
High : 119.49
Low : 17.63
A
Air Conditioner
High : 55860.63
Low : -7098.88
B
Fire Place
High : 74706.97
Low : -6722.29
C
N
Num. of Bathrm
High : 39931.12
Low : -2044.24
D
House Age
High : 929.44
Low : -1402.30
E
Soil & Imp. Sfc
High : 34357.96
Low : -220301.55
F
0
5
10
20
Kilometers

Spatially Varying GWR (Local) Regression Results:
Model Fit and the Effect of TRI Exposure on Quality of Life