Geographically Weighted Regression (GWR)

Revealing Spatial Non-Stationarity with Geographically Weighted Regression

Yifan Yang, Ph.D. Student

Advisor: Dr. Lei Zou (lzou@tamu.edu)

Department of Geography, Texas A&M University

GEAR Lab Website: https://www.geoearlab.com

GEAR Lab
Explore & Resolve

Self Introduction

University of Southern California
Master of Science, Spatial Data Science
August 2022- May 2024

Hainan University
Bachelor of Engineering, Software Engineering
September 2018 - June 2022

Research Interests

1. Spatial Data Science & GeoAI Integration

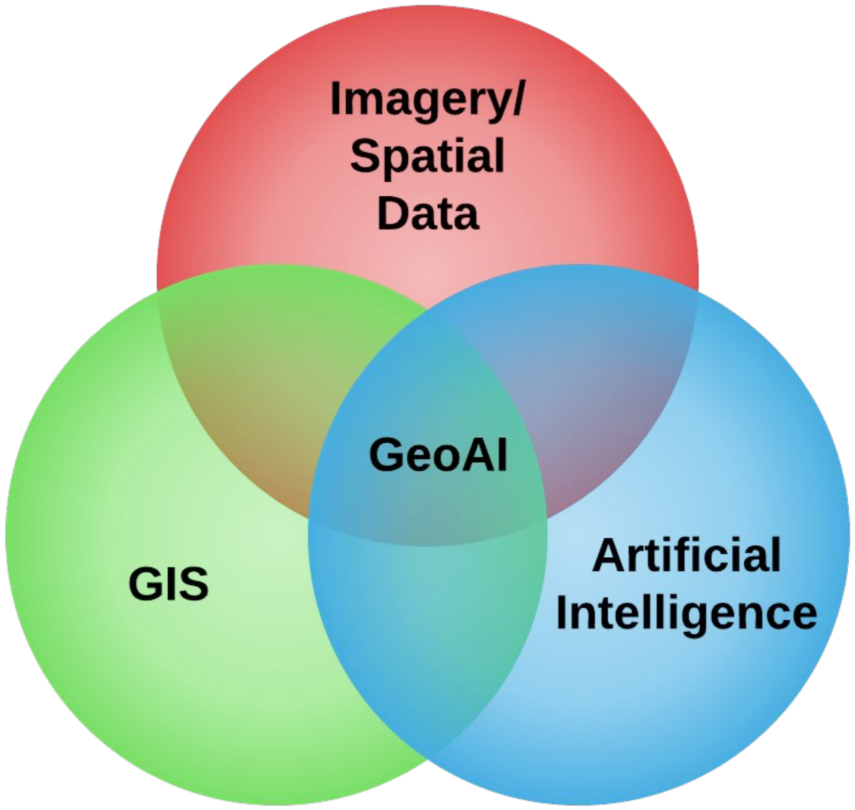
To integrate spatial data science with artificial intelligence for understanding and modeling geographic patterns.

Disaster Resilience: Predicting, assessing, and mitigating natural disasters using multi-source spatial data such as remote sensing, GIS, and street-view imagery.

2. Responsible and Autonomous GeoAI

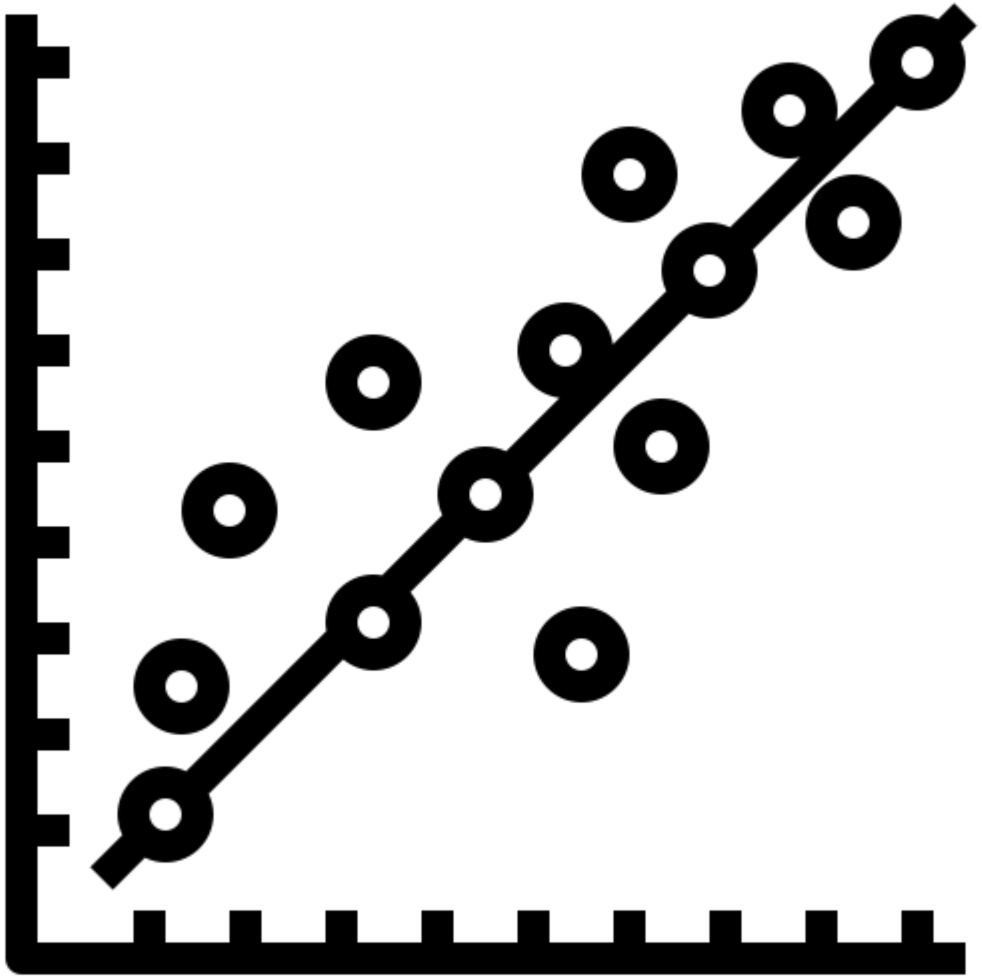
To develop responsible and autonomous geographical AI systems that can perceive, reason, and act in complex real-world environments.

https://rayford295.github.io/



GeoAI = GIS + Artificial Intelligence

EAR Lab
Explore & Resolve



The key to all of this is statistics.

If you cannot understand statistics, you cannot understand data, and naturally you will not understand information science, or GIScience.

Learning Objectives

By the end of this session, students will be able to:

1. **Explain the conceptual motivation of Geographically Weighted Regression (GWR).**
2. **Distinguish between global regression and local regression models.**
3. **Identify the core concepts required to apply GWR correctly.**
4. **Interpret spatially varying regression coefficients and diagnostics.**
5. **Implement a basic GWR workflow using Python in a Jupyter/Colab environment**

Minimum learning objectives

Before you leave, there are at least three questions you need to understand.

1. **The Purpose of Regression.**
2. **Why is Geographic Data Special?**
3. **Does it satisfy the Law of Large Numbers?**

Conceptual Motivation: Why Do We Need GWR?

Key Question

- Why might a single global regression model be insufficient for geographic data?
- What are the unique characteristics of geographic data?

https://news.mit.edu/2010/explained-reg-analysis-0316

Conceptual Motivation: Why Do We Need GWR?

Why might a single global regression model be insufficient for geographic data?

In classical regression analysis (e.g., Ordinary Least Squares, OLS), we assume that the relationship between variables is **spatially stationary**—that is, the same regression coefficients apply everywhere in the study area.

However, many geographic processes violate this assumption.

Examples:

- Vegetation may cool urban temperatures more effectively in dense downtown areas than in suburban neighborhoods.

- Building density may increase land surface temperature differently across industrial, residential, or coastal zones.

This phenomenon is known as **spatial non-stationarity**, where relationships between variables vary across space.

Geographically Weighted Regression (GWR) was developed to explicitly model this spatial variation by allowing regression coefficients to change from one location to another.

Conceptual Motivation: Why Do We Need GWR?

What are the unique characteristics of geographic data?

Global models assume that all data points are independent and that the relationships between variables are constant throughout the entire study area.

However, geographical data often violate these assumptions:

1. **Spatial Heterogeneity (Non-stationarity)**: This is the main reason. The relationships between variables change with location.

The assumption of a global model is that the effect of the independent variable X on the dependent variable Y (i.e., the regression coefficient) is the same regardless of location.

Real-world situation:

Example: Suppose you want to study the impact of "house size" on "house price."

In the city center, each additional square meter might increase the house price by **100,000** (a large coefficient). In a remote suburb, each additional square meter might only increase the house price by **10,000** (a small coefficient). If you use a global model, you will only get an "average value" (e.g., **50,000**), which cannot accurately predict prices in either the city center or the suburbs.

Conceptual Motivation: Why Do We Need GWR?

2. **Spatial Dependence (Autocorrelation)**: The first law of geography states: "Everything is related to everything else, but near things are more related than distant things."

The assumption of a global model is that all observations (sample points) are independent of each other, and the error terms (residuals) are randomly distributed.

The reality is: Geographic data often exhibits clustering. If housing prices are high in one area, the prices in surrounding neighborhoods are usually also high.

If global regression is used, the model's residuals (the difference between predicted and actual values) often show a clear clustering pattern on the map (for example, the model's predicted values are consistently too low in a certain area). This means the model is missing some crucial spatial information, leading to invalid statistical inferences.

Conceptual Motivation: Why Do We Need GWR?

3. **Simpson's Paradox**:

In geographical data, if we ignore local grouping structures (such as different climate zones or different administrative regions) and only look at the overall data, we may arrive at conclusions that are completely opposite to the local facts. Global models often mask these true local trends.

Feature	Global Model (e.g., OLS)	Geographic Reality
Relationship	Fixed everywhere (Stationary)	Varies by location (Non-stationary)
Data Points	Independent	Spatially Dependent (Clustered)
Coefficients	Single average value	Local values map out spatial patterns
Best For	General, non-spatial summary	Analyzing local dynamics

Global Regression vs. Local Regression

Global Regression vs. Local Regression

Global Regression (OLS)

A global regression model estimates **one single set of coefficients** for the entire study area:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \epsilon$$

Key characteristics:

- One global relationship
- Easy to interpret
- Assumes spatial stationarity
- Often produces spatially clustered residuals when the assumption is violated

OLS answers the question:

What is the average relationship across the entire study area?

Global Regression vs. Local Regression

Local Regression (GWR)

GWR relaxes the assumption of spatial stationarity by estimating **location-specific coefficients**:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \epsilon_i$$

where (u_i, v_i) are the coordinates of location i.

Key characteristics:

- Coefficients vary across space
- Captures local relationships
- Produces maps of regression coefficients
- More complex to interpret than OLS

GWR answers the question:

How does the relationship change from place to place?

Core Concepts Required to Apply GWR Correctly

Core Concepts Required to Apply GWR Correctly

To use GWR appropriately, several foundational concepts must be understood.

1. **Spatial Non-Stationarity**

This is the fundamental reason for using GWR (Geographically Weighted Regression).

Meaning: The relationship between variables is not the same everywhere, but varies with geographical location.

Simple explanation: For example, consider the relationship between "rainfall" and "crop yield." In arid regions, more rainfall leads to higher yields (positive correlation); but in low-lying, flood-prone areas, too much rain can drown the crops (negative correlation). This phenomenon of the relationship changing with location is called spatial non-stationarity.

The function of GWR: GWR doesn't simply "assume" this phenomenon exists; it's a detection tool. After running the model with GWR, if significant differences in coefficients are found across different locations, it proves the existence of spatial non-stationarity.

Core Concepts Required to Apply GWR Correctly

To use GWR appropriately, several foundational concepts must be understood.

2. Distance Decay and Spatial Weighting

This is the core mechanism of GWR (Geographically Weighted Regression).

Meaning: When calculating the regression equation for a specific point on the map (the target point), data samples closer to this point have greater influence (weight), while those further away have less influence.

Simple explanation: It's like if you want to understand the housing prices in a particular neighborhood in Beijing, you would focus on the prices of neighboring neighborhoods (high weight), and wouldn't pay much attention to the prices of a neighborhood in Shanghai (low or zero weight).

Implementation: This is achieved by using a "distance decay function" to assign a score (weight) to each sample.

Core Concepts Required to Apply GWR Correctly

To use GWR appropriately, several foundational concepts must be understood.

3. Kernel Functions

This is the mathematical formula that defines the specific form of "distance decay."

Meaning: The kernel function specifies how the weight decreases as the distance increases.

Two common types of kernels:

Gaussian kernel: The decrease is relatively smooth. Regardless of the distance, the weight never completely becomes 0, only becomes very, very small (it has a global influence).

Bisquare kernel: There is a hard cutoff distance. Beyond this distance, the weight directly becomes zero (it only has a local influence).

Note: While the choice of kernel function does have an impact, it is usually not as important as choosing the correct "bandwidth."

Core Concepts Required to Apply GWR Correctly

4. Bandwidth Selection

This is the most crucial step in GWR (Geographically Weighted Regression).

Meaning: Bandwidth determines the size of the "local" area (the size of the circle) considered when calculating local regression.

Too small bandwidth: The circle is too small, only considering the immediate vicinity. The model becomes very fragmented and highly volatile, prone to overfitting.

Too large bandwidth: The circle is too large, encompassing the entire province or even the whole country. The model degenerates into a regular global regression model, losing the significance of GWR.

How to choose: We usually don't guess, but let the computer automatically select an optimal value.

AICc (Corrected Akaike Information Criterion): An indicator for measuring the quality of a model, aiming for a model that is both accurate and simple.

CV (Cross-validation): Another method for automatic optimization.

Core Concepts Required to Apply GWR Correctly

To use GWR appropriately, several foundational concepts must be understood.
5. Local Regression Coefficients

Meaning: The influence of each independent variable (predictor, such as vegetation cover NDVI) on the dependent variable (such as temperature) varies at every point on the map.

Visualization: We usually don't look at tables, but instead visualize these coefficients as maps (Coefficient Maps).

Red areas may indicate a positive correlation (higher NDVI means higher temperature).

Blue areas may indicate a negative correlation (higher NDVI means lower temperature).

How to interpret: Your interpretation must be "localized."

Example: "In this specific area (e.g., the city center), a one-unit increase in vegetation leads to a decrease of X degrees in temperature; but in that area (e.g., the suburbs), the same increase in vegetation only leads to a decrease of Y degrees in temperature."

Core Concepts Required to Apply GWR Correctly

To use GWR appropriately, several foundational concepts must be understood.
6. Local R^2

Local R^2: This is a tool used to check "where the model is accurate and where it is not."

Meaning: R^2 represents the model's ability to explain the data (goodness of fit). Ordinary regression only has one global R^2 (e.g., 0.6), which tells you that the model is generally acceptable. However, GWR calculates an R^2 for each point.

High local R^2: This indicates that in this area, your model is perfect and can explain what is happening very well.
Low local R^2: This indicates that the model fails in this area. Implication: This usually means you have missed some key variables.

For example, if the R^2 is very low in a certain area, it might be because there is a special chemical plant there causing abnormal temperatures, but your model does not include the "distance to the factory" variable. This is very useful for disaster research, as it can help you discover unknown risk driving factors.

From Concepts to Practice: The GWR Workflow

Link: 👇
https://colab.research.google.com/drive/1o2G9PTpvzrkJToewi_5NKpqK1GuT34Re?usp=sharing

0206_GWR.ipynb

文件 修改 视图 插入 代码执行程序 工具 帮助

命令 + 代码 + 文本 全部运行

Geographically Weighted Regression (GWR): A Hands-on Urban Heat Island Example

Context (GEOG 475: Advanced GIS)
In many urban environmental problems, relationships vary across space. A single global regression model may hide local patterns.

What we will do today (workflow)

1. Simulate a city-like dataset (points + urban variables)
2. Fit a global regression model (OLS)
3. Diagnose spatial patterns in residuals (Moran's I)
4. Fit a local regression model (GWR) with bandwidth selection
5. Visualize local coefficients and compare OLS vs GWR

Key concept: spatial non-stationarity (spatially varying relationships)

```
# Install packages (Colab)
!pip -q install numpy pandas matplotlib scipy statsmodels
!pip -q install libpysal esda
!pip -q install mgwr

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

from scipy.spatial.distance import cdist
import statsmodels.api as sm

from libpysal.weights import KNN
from esda.moran import Moran

from mgwr.gwr import GWR
from mgwr.sel_bw import Sel_BW

np.random.seed(42)
```

Step 1 — Create sampling locations (a simplified "city")

We simulate point locations in a 2D plane:

- A denser cluster near the city center (CBD)
- A sparser ring of points in the outer area

This is more realistic than a uniform grid because real sampling / urban features are not evenly distributed.

OLS Regression Results

Dep. Variable:	y	R-squared:	0.555
Model:	OLS	Adj. R-squared:	0.551
Method:	Least Squares	F-statistic:	139.9
Date:	Thu, 05 Feb 2026	Prob (F-statistic):	7.75e-59
Time:	17:41:43	Log-Likelihood:	-569.11
No. Observations:	340	AIC:	1146.
Df Residuals:	336	BIC:	1162.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	34.2414	0.640	53.530	0.000	32.983	35.500
x1	-6.1677	0.931	-6.627	0.000	-7.999	-4.337
x2	3.4042	0.984	3.461	0.001	1.469	5.339
x3	0.4199	0.295	1.424	0.155	-0.160	1.000

Omnibus:	3.161	Durbin-Watson:	1.967
Prob(Omnibus):	0.206	Jarque-Bera (JB):	2.367
Skew:	-0.020	Prob(JB):	0.306
Kurtosis:	2.593	Cond. No.	19.8

Geographically Weighted Regression (GWR) Results

Spatial kernel:	Adaptive bisquare
Bandwidth used:	50.000

Diagnostic information

Residual sum of squares:	206.600
Effective number of parameters (trace(S)):	64.861
Degree of freedom (n - trace(S)):	275.139
Sigma estimate:	0.867
Log-likelihood:	-397.752
AIC:	927.227
AICc:	959.471
BIC:	1179.406
R2:	0.838
Adjusted R2:	0.799
Adj. alpha (95%):	0.003
Adj. critical t value (95%):	2.981

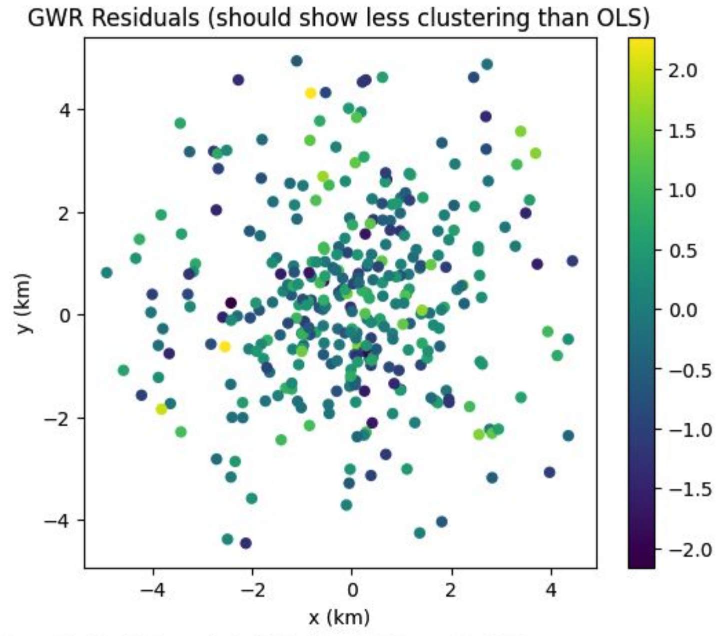
Summary Statistics For GWR Parameter Estimates

Variable	Mean	STD	Min	Median	Max
X0	33.290	2.489	27.856	33.438	39.570
X1	-6.286	2.478	-12.608	-5.982	2.233
X2	3.923	3.655	-3.168	3.477	13.286
X3	0.055	1.721	-3.773	-0.006	4.917

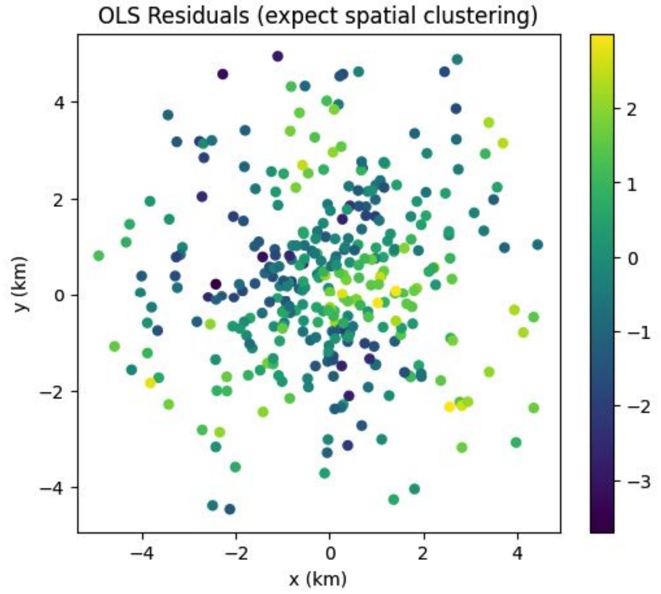
None

Goodness of Fit (R^2):
OLS (Global): 0.555. This means the global model can only explain 55.5% of the data variation. GWR (Local): 0.838. The explanatory power of GWR soars to 83.8%!

Conclusion: This indicates that a large amount of local patterns are hidden in the data, which the global model failed to capture, while GWR successfully captured them. Information Criterion (AIC/AICc): OLS: 1146 GWR: 959.471 (AICc) Conclusion: A lower AIC value is better. The significantly lower value for GWR indicates that even though GWR is more complex (using more parameters), the improvement in accuracy is worthwhile. This is a more "superior" model.



Moran's I (GWR residuals): I=-0.002, p=0.4870



Moran's I (OLS residuals): I=0.477, p=0.0010

Residual Diagnostics

This is a check of the model's residuals.
Visual inspection: The colored dots on the map appear to be "randomly scattered" (salt and pepper pattern).
There are no large areas of all yellow (overestimation) or all purple (underestimation).

Statistical check (Moran's I):
The text below the graph shows:
Moran's I = -0.002, p = 0.4870.

I \approx 0: This indicates that the residuals are randomly distributed, and there is no spatial autocorrelation.

p > 0.05: This indicates that this randomness is statistically significant (not a coincidence).

Conclusion: This shows that the GWR model has "extracted" all the spatial patterns from the data, leaving only pure random noise. If I were large here (e.g., 0.5), it would indicate that the model is poor and has missed crucial spatial information. But the current structure is perfect.

Is GWR truly perfect? What are its drawbacks?

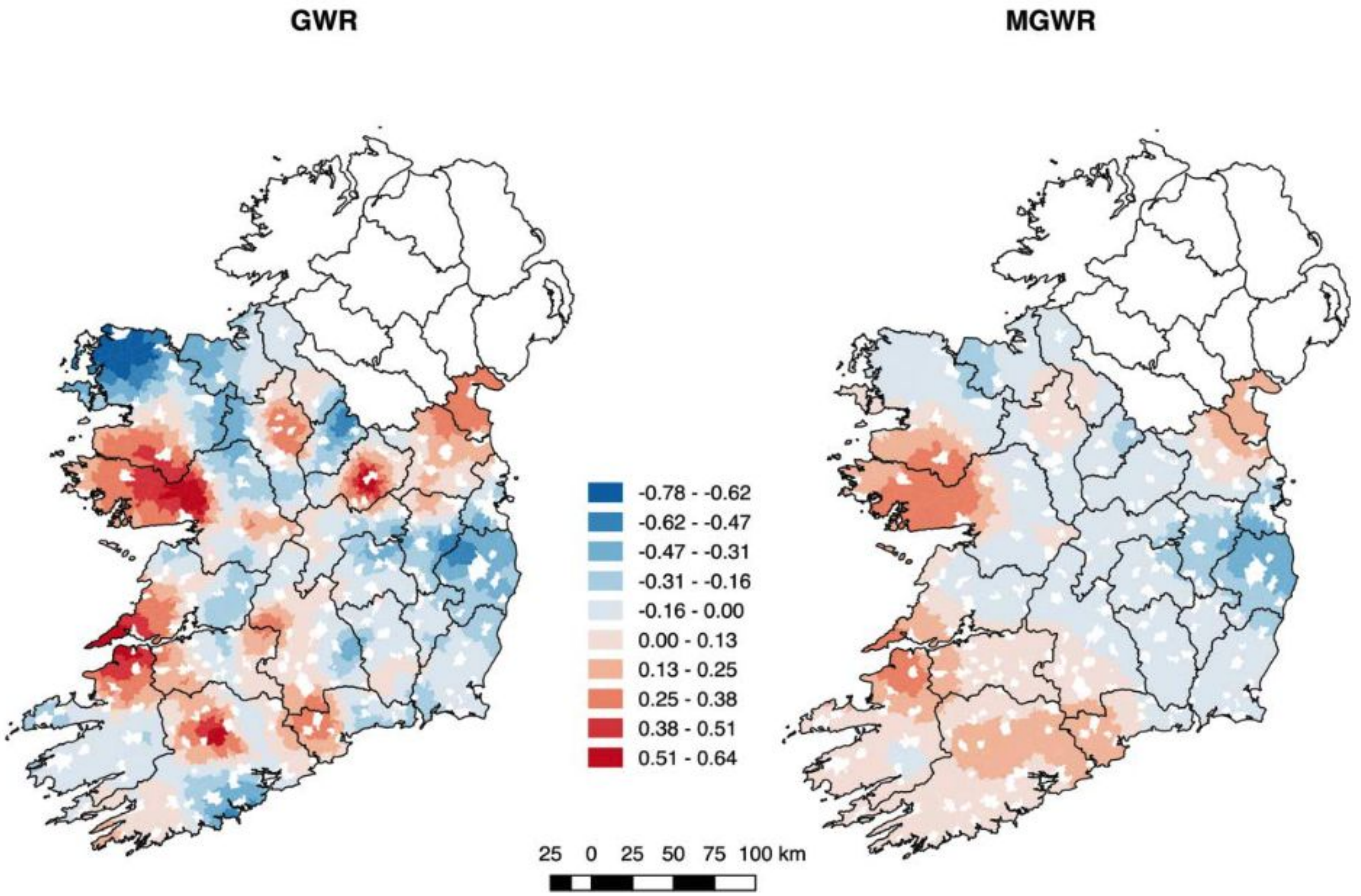
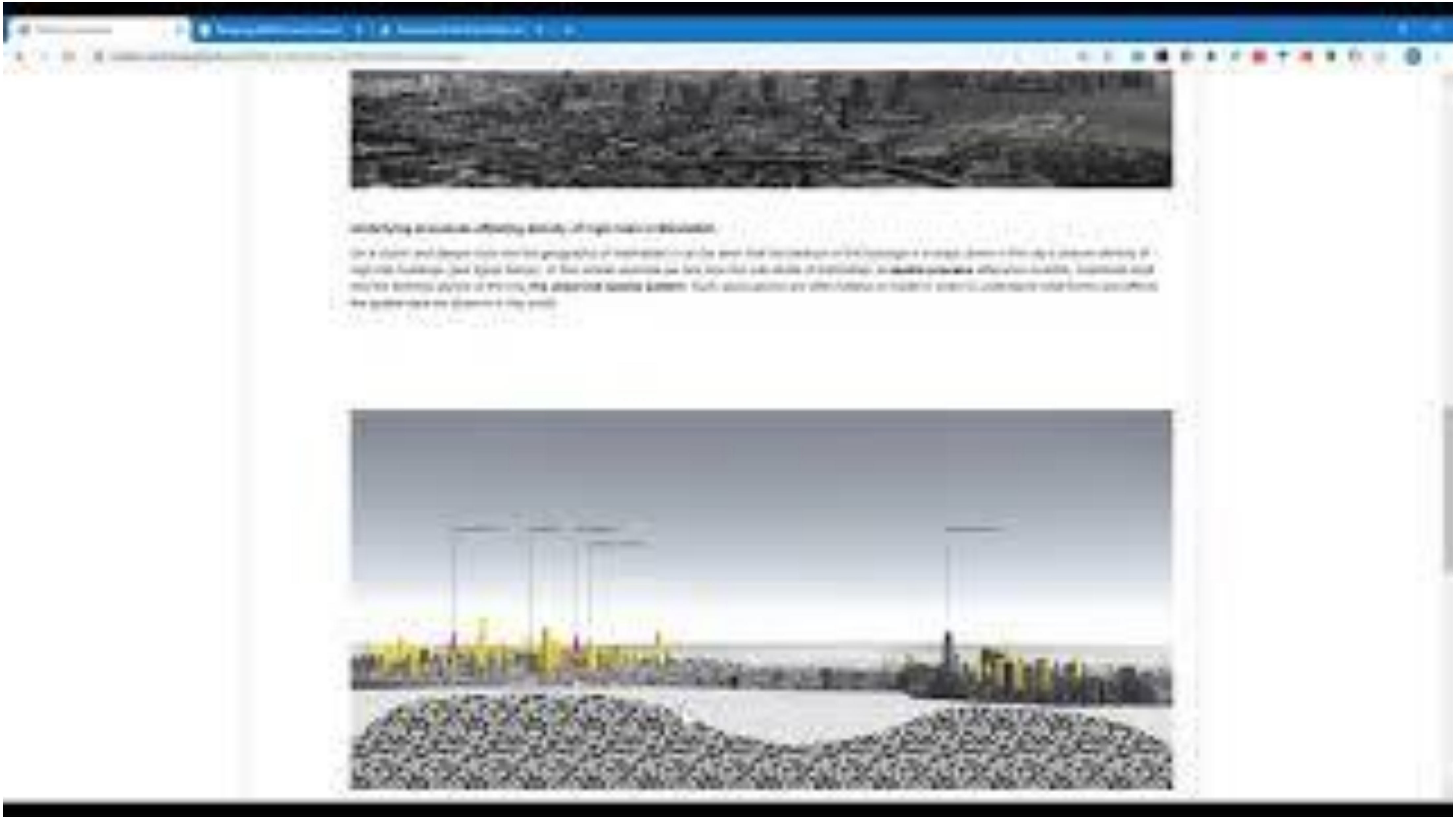


Figure 17. Geographically weighted regression and multiscale geographically weighted regression local estimates for workhouse proximity. GWR = geographically weighted regression; MGWR = multiscale geographically weighted regression. (Color figure available online.)

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (MGWR). ***Annals of the American Association of Geographers*****,** ***107*****(6), 1247-1265.**

https://youtu.be/o8lDAJaFZfU?si=_L0300AGLxdKjZYr



Summary

Before you leave, there are at least three questions you need to understand.

1. **The Purpose of Regression.**

   Description & Explanation, Prediction, Causal Inference & Control

2. **Why is Geographic Data Special?**

   Spatial Autocorrelation, Spatial Heterogeneity, MAUP - Modifiable Areal Unit Problem

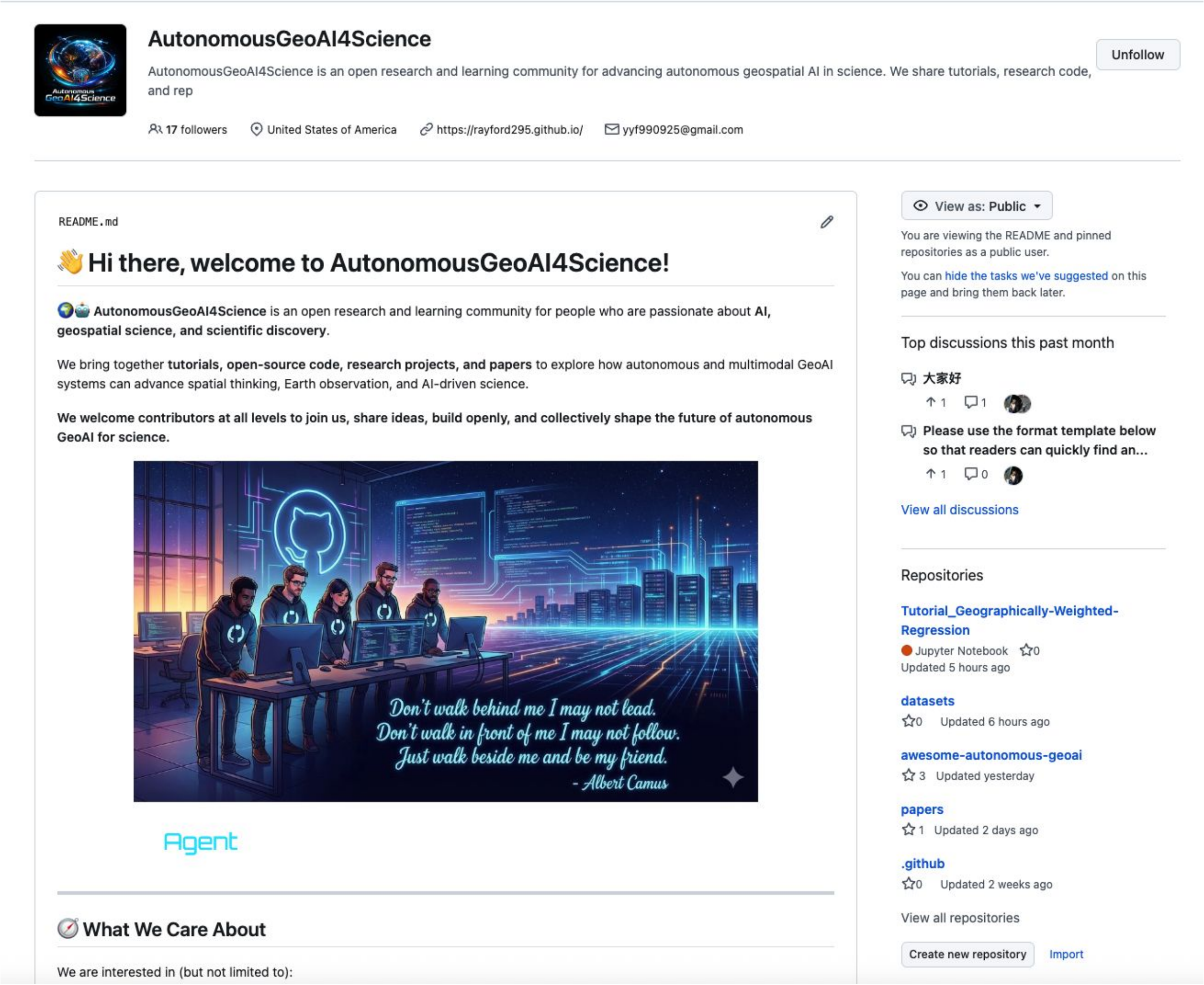
3. **Does it satisfy the Law of Large Numbers?**

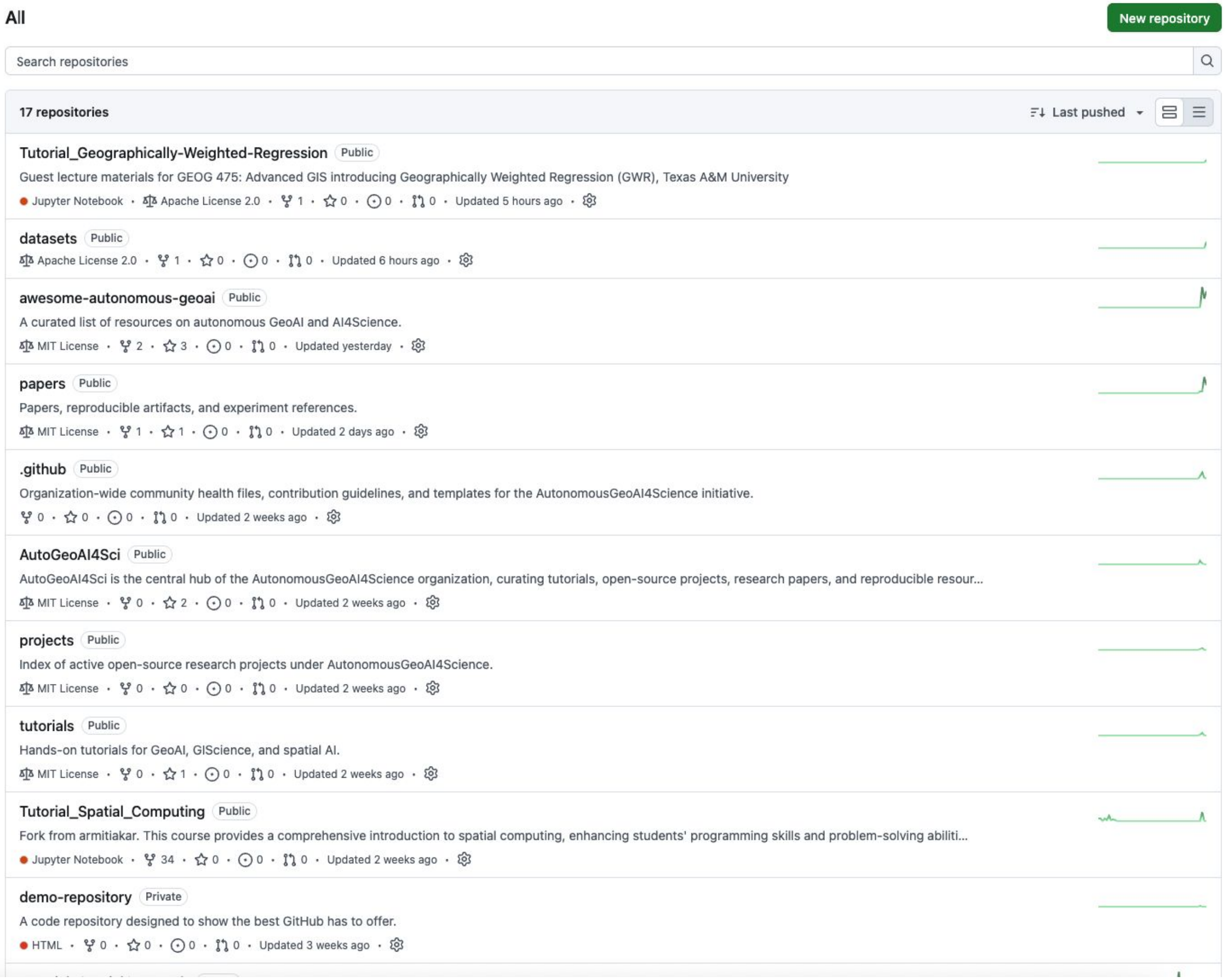
   The premise for this is usually that the samples are independent. In geographical data, due to the existence of spatial autocorrelation, increasing the sample size n does not necessarily bring you closer to the truth.

Course Resources

https://github.com/rayford295/Tutorial_Geographically-Weighted-Regression

https://github.com/AutoGeoAI4Sci/Tutorial_Geographically-Weighted-Regression





Q&A

yyang295@tamu.edu

EAR Lab
Explore & Resolve