



# 地理加权回归模型回顾与新进展

## 摘要

地理加权回归模型 (geographically weighted regression, GWR) 是一种局部加权回归模型，将空间不平稳性即空间异质性纳入模型，从而量化变量空间分布及其对因变量解释程度的变化情况。虽然面临运算量大和多重共线性强等挑战，GWR 模型已在众多领域内被广泛应用并衍生出众多改良模型，成为重要的空间建模工具。本文回顾了 GWR 模型的发展，系统介绍了 GWR 模型的基本形式及求解与推断方法，并根据实际问题介绍了三个新发展的 GWR 衍生模型。并以武汉市城市建筑景观与地表温度相关性分析为例，进一步介绍了 GWR 技术软件工具的使用。

**关键字：**地理加权回归 广义加性模型 空间异质性 空间统计

## 1 | 引言

在空间统计中，一般线性回归 (ordered linear regression, OLR) 模型中确定变量见关系的重要统计方法，然而由于变量在空间上经常呈现不平稳的分布，因此 OLR 模型基于全局平稳的假设进行建模的方法难以对这样的空间不平稳性进行解释。因此人们提出了一些将变量空间结构不平稳性纳入模型的方法，以 (Cleveland 1979; Anselin 1995, 1988) 为代表的局部地理回归模型得以发展。

区别于基于全局平稳假设的空间统计技术，地理加权回归模型 (geographic weighted regression model, GWR) 是一种基于局部统计的空间统计方法，其核心思想是借用估计点周围的信息用于预测该点的值。GWR 一方面提供了相对于全局回归方法更好的统计效果，另一方面为解释空间的不平稳性即异质性提供了新的技术 (Brunsdon et al. 1996, 1998; Fotheringham and M.E. Charlton & C. Brunsdon 1997)。并迅速从早期的一种探索性空间统计方法 (Brunsdon et al. 1998) 发展出一套完整的假设检验 (Brunsdon et al. 1999) 从而被广泛使用。在地质学 (Atkinson et al. 2003)，环境科学 (Mennis and Jordan 2005)，城市景观生态学 (Buyantuyev and Wu 2010)，生命科学 (Oshan et al. 2020) 和社会经济学 (Lu et al. 2011) 等领域均有众多应用。

与此同时，众多 GWR 衍生模型被提出。如 (Fotheringham et al. 2002) 中提出的混合地理加权回归 (semiparametric geographically weighted regression, SGWR) 模型，(Huang et al. 2010) 中提出引入时间不平稳性并被 (Wu et al. 2014) 和 (Fotheringham et al. 2015) 完善为地理时空加权 (geographically weighted temporally regression, GTWR) 模型，(Lu et al. 2015) 中提出的使用非欧距离加权的回归模型等都拓宽了 GWR 的应用场景。

然而，GWR 模型的运用在很长一段时间受到不少批评，主要集中在变量的异质性程度评价和多重共线性风险。由于 GWR 模型比相比需要大得多的运算资源，因此对变量的异质性特征的评价以确认使用 GWR 模型的必要性是 GWR 模型运用首先要回答的问题。与其他非参数方法相同，早期的空间异质性检验多基于蒙特卡洛模拟法 (Brunsdon et al. 1996) 和空间自适应滤波法 (Foster and Gorr 1986)，随后 brunsdon 等总结了前人使用的检验方法并引入 F 检验 (Brunsdon et al. 1999)，并被 leung 等完善 (Leung et al. 2000b,a)。而 (Griffith 2008) 等提出伴随着变量异质性出现的变量间较强的相关性会导致模型求解结果失真。同时 GWR 模型将所有变量统一到同一地理尺度很可能引入额外的多重共线性 (Comber et al. 2018)。为了回答这个问题，Fotheringham 等构建了一系列模拟实验验证了在一般情况下 GWR 技术对共线性风险的稳健性 (Fotheringham and Oshan 2016)。Wheeler 等通过构建地理加权 Lasso 回归 (geographically weighted lasso, GWL) 回归的方法探讨了缓解共线性风险的可能 (Wheeler 2009)。在实际应用中，对于变量间的共线性风险的评价仍然是 GWR 模型运用中的一个重要问题。

引文网络分析是一种基于文献之间引用关系的分析方法，可以用来识别研究领域内的核心文献和知识结构。本文以地理加权回归 (GWR) 为主题抓取 web of science 上共 1400 篇期刊文章，使用 VOSviewer 基于引文网络分析，确定核心文献结果如图1所示。

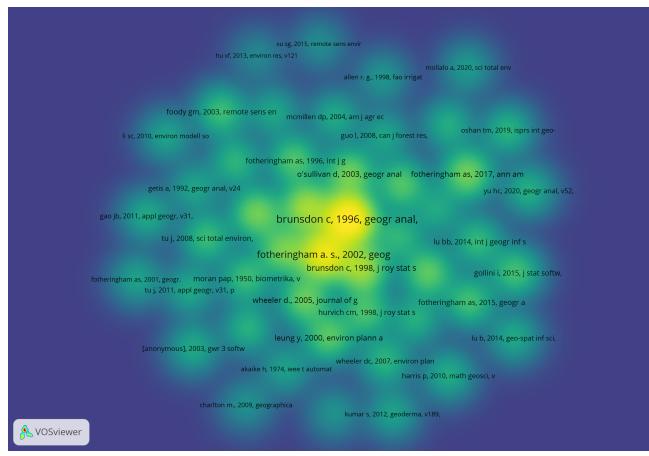


图 1: GWR 核心文献热点图

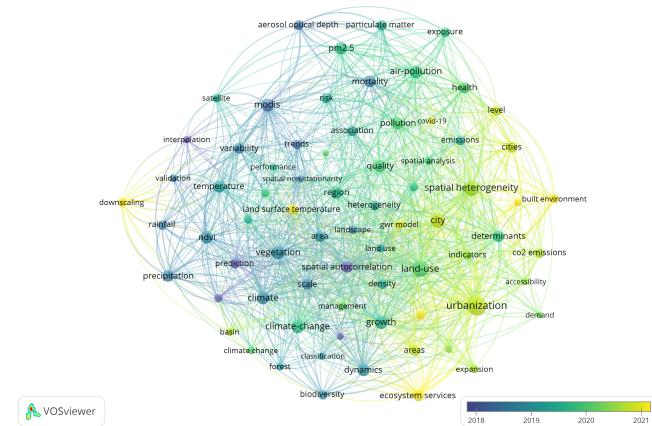


图 2: GWR 关键词变化趋势

从图2关键词网络中可以发现，近年来 GWR 相关方向关注的问题从回归模型的“本职工作”拟合预测本身 (trend, autocorrelation, prediction, etc.) 更多转向揭示时空不平稳性 (heterogeneity, downscaling, etc.)，也从更多关心自然地理现象 (ndvi, temperature, etc.) 转向城市研究 (urbanization, city, etc.) 和人居环境 (ecosystem services, building environment, etc.)。这体现了学界对地理加权回归这一方法认识的深化，也反映了 GWR 模型在以如城市等具有高度异质性的研究对象研究时具有的良好的解释性被学界越来越多地重视。

本文在此基础上结合高被引论文及其引文网络确定综述文献阅读整理范围，以下部分将按照此顺序组织：第二节介绍地理加权回归模型的基本原理，第三节从不同用途介绍各种地理加权回归衍生模型，第四节介绍近些年发展的从广义加性模型视角转化地理加权回归模型及其意义，最后以研究武汉市建筑景观与城市地表温度解释异质性为例介绍地理加权回归模型工具箱的使用并总结。

## 2 | 地理加权回归模型基本原理

地理加权回归模型 (geographically weighted regression, GWR) 是一种局部回归模型，其基本形式为：

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p X_{ik}\beta_k(u_i, v_i) + \varepsilon_i, i = 1, 2, \dots, n \quad (1)$$

其中  $y_i$  为因变量,  $x_{ij}$  表示第  $i$  个样本的第  $j$  个变量,  $(u_i, v_i)$  表示第  $i$  个样本的经纬度坐标,  $\beta_i = [\beta_1(u_i, v_i), \beta_2(u_i, v_i), \dots, \beta_p(u_i, v_i)]^T$  表示第  $i$  个样本第  $j$  个变量的回归系数。

其中  $W_i$  为距离加权矩阵, 为对角矩阵:

$$W_i = \begin{pmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{pmatrix} \quad (2)$$

其中对角线值  $w_{ij}$  表示第  $j$  个数据点到回归分析点  $i$  的权重值。加权函数称为核函数 (kernel function), 根据值域分布特折可分为连续型和截断型。(卢宾宾 et al. 2020) 总结了常用的核函数如下:

$$\begin{aligned} \text{Gaussian: } w_{ij} &= \exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right] & \text{Exponential: } w_{ij} &= \exp\left(\frac{|d_{ij}|^2}{b}\right) \\ \text{Box-car: } w_{ij} &= \begin{cases} 1 & d_{ij} \leq b \\ 0 & d_{ij} > b \end{cases} & \text{Bi-square: } w_{ij} &= \begin{cases} (1 - (\frac{d_{ij}}{b})^2)^2 & d_{ij} \leq b \\ 0 & d_{ij} > b \end{cases} \end{aligned} \quad (3)$$

其中  $w_{ij}$  为第  $i$  个样本到第  $j$  个样本的距离权重,  $d_{ij}$  为两点间距离,  $b$  为带宽。其中 Gaussian 核函数和 Bi-square 核函数是最常用的核函数, 但截断型核函数对 GWR 的“局部”这一语义有更好的解释性, 但容易造成投影矩阵奇异, 在实际运用中需要权衡考量。

其最小二乘估计方程正规化形式为:

$$\begin{aligned} L(\beta_i) &= \sum_{j=1}^n w_{ij} \cdot (y_j - \beta_{i0}(u_i, v_i) - \sum_{k=1}^p X_{jk}\beta_{ik}(u_i, v_i))^2 \\ \hat{\beta}_i &= \operatorname{argmin}_{\beta_i} L(\beta_i) \end{aligned} \quad (4)$$

结果可表示为:

$$\hat{\beta}_i = (X^T W_i X)^{-1} X^T W_i y, i = 1, 2, \dots, n \quad (5)$$

从而可根据  $\hat{\beta}$  构建投影矩阵  $S$ :

$$S = \begin{pmatrix} x_1(x^T W_1 x)^{-1} x^T W_1 \\ \vdots \\ x_n(x^T W_n x)^{-1} x^T W_1 \end{pmatrix}_{n \times n} \quad (6)$$

则可以发现其  $S$  在观测值  $x_{ij}$  确定时由加权矩阵  $W_i$  确定, 进一步的在函数形式和距离  $d_{ij}$  计算方式确定时  $W_i$  由带宽  $b$  唯一确定, 因此选择合适带宽是求解 GWR 的关键步骤。若带宽过小会导致局部模型过拟合, 带宽过大则可能导致 GWR 模型趋于全局化。从此视角出发可将 GWR 模型分为固定型带宽和可变型带宽两种形式。固定型带宽需要人为试验确定合适的带宽, 可变型带宽则通过依据  $S$  构建损失函数确定最佳带宽  $b$ 。目前基于修正赤池信息量准则 (Akaike information criterion,  $AIC_c$ ) 和交叉验证 (cross validation,  $CV$ ) 方法被广泛使用, 其表达式为:

$$AIC_c = 2n \ln(\hat{\sigma}) + n \ln(2\pi) + n \frac{n + tr(S)}{n - 2 - tr(S)} \quad (7)$$

$$CV(b) = \sum_{i=1}^n [y_i - \hat{y}_{\neq i}(b)]^2$$

其中  $\hat{\sigma}$  为最小二乘估计结果的标准差,  $tr(S)$  为投影矩阵  $S$  的迹, 最佳带宽的选择原则是使得选定指标最小。Lu 等近年来还提出了使用 Minkowski 方法实现寻优带宽 (Lu et al. 2016) 等方法, 为 GWR 带宽寻优选取提供了更多选择。同时回归的残差平方和 (residual sum of squares,  $RSS$ ) 和校正决定系数 (adjusted  $R^2$ ) 可以反映模型的预测精度和拟合优度, 将其与一般线性回归模型 (Original Linear Regression Model, OLS) 的相关诊断统计量比较可以体现 GWR 模型相对于 OLS 方法的改进程度。一般地, 当  $AIC_c$  的变化值大于 3 时可以认为模型结果之间具有显著不同。

### 3 | 地理加权回归衍生模型及其应用

#### 3.1 解释异质性的内涵

方创琳院士等在 (Sun et al. 2023) 对新城市化区域对城乡差异的异质性影响研究中提出了解释异质性的两种内涵:

1. 自变量的变化对因变量的影响不一致。
2. 自变量对因变量的影响大小和方向随因变量的变化而变化。

其中第一种异质性被广泛考虑, 表现为局部回归模型中各地点上自变量对因变量的解释程度不同, 而第二种异质性并没有广泛揭示, 而 GWR 模型已经具有根据局部回归量统计性描述的不同提供异质性的直观解释。但基本的 GWR 模型受限于要将全部变量统一到统一尺度, 且较难考虑变量在时间上的分布特征, 使得其对地理现象的时空异质性的描述不够充分。因此在近十几年来一些新的 GWR 衍生模型得到构建和运用, 以下简单介绍两类 GWR 衍生模型。

#### 3.2 使用地理加权回归模型解释时空二维异质性

GTWR 模型的基本形式是在 GWR 模型的基础上将二维的空间位置增加时间维度为三维的时空位置:

$$Y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p X_{ik} \beta_k(u_i, v_i, t_i) + \varepsilon_i, i = 1, 2, \dots, n \quad (8)$$

然而由于时间和空间具有不同的度量单位，且时间距离与空间距离的“接近度”度量并不相同，因此需要定义时空接近度。记  $d_{ij}^{ST}$  为  $i$  和  $j$  两点间的距离，则：

$$d^{ST} = d^S \otimes d_T = \lambda d^S + \mu d^T = \sqrt{\lambda[(u_i - u_j)^2 + (v_i - v_j)^2] + \mu(t_i - t_j)^2} \quad (9)$$

其中  $\lambda$  和  $\mu$  为空间和时空距离的权重参数。则对于 Gaussian 核函数和 bi-square 等使用二次项的核函数，以 Gaussian 核函数为例其回归权重为：

$$\begin{aligned} w_{ij} &= \exp\left\{-\frac{\lambda[(u_i - u_j)^2 + (v_i - v_j)^2] + \mu(t_i - t_j)^2}{h_{ST}^2}\right\} \\ &= \exp\left\{-\left[\frac{(u_i - u_j)^2 + (v_i - v_j)^2}{h_S^2} + \frac{(t_i - t_j)^2}{h_T^2}\right]\right\} \\ &= \exp\left\{-\frac{d_{ij}^S}{h_S^2}\right\} \times \exp\left\{\frac{d_{ij}^T}{h_T^2}\right\} = w_{ij}^S \times w_{ij}^T \end{aligned} \quad (10)$$

更一般地可以表示为：

$$w_{i_{(S,T)}}^{(t)} = k_S(d_{ij}^S, b_S) \times k_T(d_{ij}^T, b_T) \quad (11)$$

其中  $k_S$  和  $k_T$  是空间和时空距离的权重函数， $b_S$  和  $b_T$  是空间和时空距离的带宽参数， $d_{ij}^S$  和  $d_{ij}^T$  是空间和时空距离。相应地，将式2中的空间距离加权矩阵  $W$  拓展为时空距离加权矩阵：

$$W_i = \text{diag}(W_{i_{(S,T)}}^{(t)}, W_{i_{(S,T)}}^{(t-1)}, \dots, W_{i_{(S,T)}}^{(t-q)}) \quad (12)$$

$W_i$  为分块矩阵，表示在空间上取  $n$  个数据点和在时间上去  $q$  取个时刻，其中：

$$W_{i_{(S,T)}}^{(t)} = \text{diag}(w_{i1_{(S,T)}}^t, w_{i2_{(S,T)}}^t, \dots, w_{in_{(S,T)}}^t) \quad (13)$$

其中  $w_{ij_{(S,T)}}^t$  是数据点在  $t$  时刻根据两点的时空距离得到。

进一步地可使用交叉验证的方式来确定最佳带宽参数  $b_S$  和  $b_T$ ：

$$CV(b_S, b_T) = \sqrt{\sum_{i=1}^n \frac{(y_i - \hat{y}^{(-i)}(b_S, b_T))^2}{n}} \quad (14)$$

将时间异质性纳入模型，其具体为：

$$CV(b_{St}, b_{S(t-1)}, \dots, b_{S(t-q)}, b_T) = \sqrt{\sum_{i=1}^n \frac{(y_i - \hat{y}^{(-i)}(b_{St}, b_{S(t-1)}, \dots, b_{S(t-q)}, b_T))^2}{n}} \quad (15)$$

以上我们将二维回归模型推广到三维回归模型。此外 (Wu et al. 2014) 还提出了时空地理加权自回归 (geographically and temporally weighted autoregressive regression, GWTAR) 模型以及 (Liu et al. 2018) 还提出了时空地理加权逻辑回归 (geographically and temporally weighted logistic regression, GWTALR) 模型，此处不再展开。

### 3.2.1 使用 MGWR 提取自变量在空间分布解释异质性

因此可以发现，传统的 GWR 模型需要将各自变量数据统一到同一尺度下即使用同一带宽，而不同的自变量可能具有不同的最佳带宽。( Fotheringham et al. 2017) 在 SGWR 和距离-变量对应的地理加权回归 (geographically weighted regression with parameter-specific distance metrics, PSDM GWR) 模型 ( Lu et al. 2015) 的基础上提出了多尺度地理加权回归 (multiscale geographically weighted regression, MGWR) 模型，突破了所有自变量需要使用同一带宽的限制，使更强大的多尺度回归模型的构建成为可能。MGWR 的基本形式可写作：

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p X_{ik} \beta_{bk}(u_i, v_i) + \varepsilon_i, i = 1, 2, \dots, n \quad (16)$$

即将1式的  $\beta_k$  从依赖于自变量  $k$  的值修正为同时依赖自变量  $k$  与该自变量选取带宽  $bw$  的值。由于回归方程的建立在此时依赖于带宽的选取，从而原先基于最小化  $AIC_c$  的带宽选择方法不再适用。因此卢宾宾教授和 Fotheringham 教授在 2017 年分别提出了使用 back-fitting 的方式确定各自变量的最佳带宽的方法，其基本思想是通过改变参数估计和带宽选取的顺序采用先确定回归系数  $\beta$ ，再依据  $\beta$  确定带宽  $\tau$  并不断迭代的方法，依次确定各自变量的最佳带宽，从而得到  $\hat{\beta}$ 。(Lu et al. 2017) 中提出的算法可描述如下：

---

#### Algorithm 1 MGWR back-fitting 法确定最佳带宽

---

- 1: Initialize all coefficients by:  $\hat{\beta}^{(0)} = \langle \hat{\beta}_0^{(0)}, \hat{\beta}_1^{(0)}, \dots, \hat{\beta}_p^{(0)} \rangle$  using specific bandwidth.
  - 2: Initialize all estimated terms by:  $\hat{Y}^{(0)} = \hat{\beta}^{(0)} \cdot X$ .
  - 3: Calculate the first residual sum of squares  $RSS^0$ .
  - 4: Set iteration termination maximum number to  $N$  and threshold value  $\tau$ .
  - 5: **for**  $k \leftarrow 1$  to  $N$  **do**
  - 6:     Calculate the residual by:  $\sum_{k=1}^p (\varepsilon_i^{(k)}) = y - \sum_{j=1}^{i-1} \hat{y}_j^{(k)} - \sum_{j=i+1}^p \hat{y}_j^{(k-1)}$
  - 7:     Do the weighted regression for each variable and get the new set of coefficients  $\hat{\beta}^{(k)}$
  - 8:     Update the estimated terms by:  $\hat{Y}^{(k)} = \hat{\beta}^{(k)} \cdot X$
  - 9:     Calculate the  $k^{th}$  residual sum of squares  $RSS^k$ .
  - 10:    Calculate the change value of RSS:  $CVR^{(k)} = \frac{RSS^{(k)} - RSS^{(k-1)}}{RSS^{(k-1)}}$
  - 11:    **if**  $CVR^{(k)} > \tau$  **then**
  - 12:       **break**
  - 13:    **end if**
  - 14: **end for**
- 

从地理学角度出发，该变量的最佳带宽即为该自变量对因变量解释程度最大的最佳地理尺度。而一个自变量的最佳地理尺度可以反映在该样本集中自变量的空间异质性 ( Gollini et al. 2015): 若该变量带宽较大，即说明其在各个局部对因变量的解释程度接近，即其解释异质性较低；相反若该变量带宽较小，即说明其在各个局部对因变量的解释程度差距较大，即其解释异质性较高。

但是，由于在 MGWR 中不同变量使用了不同带宽，使得 GWR 模型的推断难以进行。该问题将在后文进一步讨论。

### 3.2.2 使用 GWQR 提取自变量在因变量值分布解释异质性

地理加权回归模型可以识别解释变量对因变量的影响是否存在空间异质性，但是传统的 GWR 模型使用最小二乘法对局部模型进行拟合，即主要反映了数据的均值特征，而对于数据的分布特征解释程度较弱。

Koenker 与 Bassett 在 1978 年在中位数回归方法上进一步发展了分位数回归 (quantile regression) 基本形式为：

$$Y_i = X_i^T \beta^\tau + \varepsilon_i \quad (17)$$

其中  $\beta^\tau$  表示对  $y$  的  $\tau^{th}$  分位数回归的系数，其含义为  $q_\tau(x_i) \equiv \inf\{y : F(y|X_i) \geq \tau\} = X_i^T \beta^\tau$ ，有  $P(Y \leq Q_Y(\tau)) = \tau$ ，并有  $P(\varepsilon \leq 0) = \varepsilon$ 。

其正规化形式为：

$$\begin{aligned} L(\beta) &= \sum_{i=1}^n \rho_\tau(y_i - x_i \beta) \\ \hat{\beta}(\tau) &= \operatorname{argmin}_\beta L(\beta) \end{aligned} \quad (18)$$

最小，其中  $\rho_\tau$  为修正函数，有  $\rho_\tau(u) = u \cdot [\tau - I(u < 0)]$ ， $I(\cdot) = \begin{cases} 1 & u < 0 \\ 0 & u \geq 0 \end{cases}$  为判定函数。

Chen 等人将其引入 GWR 模型并提出了地理加权分位数回归模型 (geographically weighted quantile regression, GWQR)( Chen et al. 2012)，并使用该技术在不同人口占比分位数下各自变量的显著性指数。该模型能更好地反映以因变量不同分位数的为目标下各解释变量的影响，并且通过 QR 解决了 GWR 中不同地区残差项独立但不同分布的问题，从而提高了模型的解释能力。GWQR 的基本形式为：

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p X_{ik} \beta_k^\tau(u_i, v_i) + \varepsilon_i, i = 1, 2, \dots, n \quad (19)$$

其中  $\tau^{th}$  分位数的含义为：

$$q_\tau(X_i, u_i, v_i) = X_i^T \beta^\tau(u_i, v_i) = \beta_0(u_i, v_i) + \sum_{k=1}^p X_{ik} \beta_k^\tau(u_i, v_i) \quad (20)$$

(Chen et al. 2012) 提出的估计方式为：

$$\beta_k^\tau(u_i, v_i) \approx \beta_k^\tau(u_0, v_0) + \beta_k^{\tau(u)}(u_0, v_0)(u - u_0) + \beta_k^{\tau(v)}(u_0, v_0)(v - v_0) \quad (21)$$

对于  $(u_0, v_0)$  的一个邻域  $(u_0, v_0)$ ，其中

$$\beta_l^{\tau(u)}(u_0, v_0) = \frac{\partial}{\partial u} \beta_k^\tau(u, v) |_{(u,v)=(u_0,v_0)}, \beta_l^{\tau(v)}(u_0, v_0) = \frac{\partial}{\partial v} \beta_k^\tau(u, v) |_{(u,v)=(u_0,v_0)} \quad (22)$$

从而将原先的  $\beta$  估计方法 4 中的损失函数修改修改为：

$$L_i(h) = \sum_{i=1}^n \rho_t(Y_i - \tilde{X}_i^T \theta^\tau(u_0, v_0)) \cdot K\left(\frac{d_{ij}}{h}\right) \quad (23)$$

其中  $\tilde{X}_j$  和  $\theta_i^\tau(u_i, v_i)$  分别为对  $X$  和  $\beta$  及其邻域的线性估计, 即  $\tilde{X}_j = [1, X_{j1}, X_{j2}, \dots, X_{jp}, (u_j - u_i), X_{j1}(u_j - u_i), X_{j2}(u_j - u_i), \dots, X_{jp}(u_j - u_i), (v_j - v_i), X_{j1}(v_j - v_i), X_{j2}(v_j - v_i), \dots, X_{jp}(v_j - v_i)]^T$ ,  $\theta^\tau(u_i, v_i) = [\beta_0^\tau(u_i, v_i), \dots, \beta_p^\tau(u_i, v_i), \beta_0^{\tau(u)}(u_i, v_i), \dots, \beta_p^{\tau(u)}(u_i, v_i), \beta_0^{\tau(v)}(u_i, v_i), \dots, \beta_p^{\tau(v)}(u_i, v_i)]^T$ 。

从而在 GWQR 模型中设立不同的分位数, 可以进一步地解释各自变量对因变量的不同分位值的贡献情况。例如分位数  $\tau$  接近 1 时, 其回归系数大的自变量对因变量的高值影响较大, 而分位数  $\tau$  接近 0 时, 其回归系数大的自变量对因变量的低值影响较大。通过对自变量对不同分位数的因变量的解释程度可以衡量其在因变量上的解释异质性。

其余部分与 GWR 处理方法基本保持一致, 但是使用  $AIC_c$  作为最佳带宽确定准则的方式不再适用, 为此Chen et al. 2012对其构建交叉验证得分求解的方法, 表达式为:

$$CV(h) = \sum_{i=1}^n \rho_\tau(Y_i - \hat{q}_\tau^{(-i)}(X_i, u_i, v_i)) \quad (24)$$

使得  $CV(h)$  最小即为最佳带宽的位置。

实际上使用  $CV$  值寻找最佳带宽本身也是在基本的 GWR 模型中寻找最佳带宽被广泛使用的方法, 而由于其可以很好地求解广义加性模型, 使得  $CV$  值在 GWR 模型求解中有了更多应用, 以下将进一步阐述。

## 4 | 从广义加性模型视角理解地理加权回归模型

使用Jia et al. 2019中提出的方法构建投影矩阵  $S$  的构造方式, 若将 GWR 模型写成广义加性模型 (Generalised Additive Model, GAM) 的形式, 则有:

$$\hat{\varepsilon} = y - \sum_{k=1}^p \hat{f}_j \quad (25)$$

则等价的 GAM 的投影矩阵族  $R_{1,\dots,p}$  满足:

$$\begin{aligned} \hat{f}_j &= R_j y \\ \hat{y}_j &= S y = \sum_{k=1}^p R_j y \Rightarrow \sum_{k=1}^p R_j = S \end{aligned} \quad (26)$$

记  $A_j$  为最优化前  $j-1$  个回归参数的投影矩阵, 则用于最优化第  $j$  个回归参数的 GAM 基函数簇为:

$$\hat{f}_j^* = A_j(\hat{f}_j + \hat{\varepsilon}) \quad (27)$$

将25式和26式带入, 得到:

$$\hat{f}_j + \hat{\varepsilon} = \hat{f}_j + y - \sum_{k=1}^p \hat{f}_k = R_j y + y - S y \quad (28)$$

与27联立并带入26, 得到:

$$R_j^* = A_j R_j + A_j - A_j S \quad (29)$$

对应的得到更新后的投影矩阵  $S^*$ :

$$S^* = S - R_j + R_j^* \quad (30)$$

从而在通用的视角将 GWR 每一个局部看作一个非参数估计过程，用于解决各类 GWR 衍生模型的解释问题。

### 讨论: 将 GWQR 推广到多尺度形式

总结上文讨论，MGWR 模型的核心是改变参数估计和带宽选取的顺序并不断迭代从而得到各自变量的最优带宽，而 GWQR 模型则是改变回归方法，使用分位数回归代替最小二乘回归进行计算。从而可以发现这两种方法是对 GWR 在不同环节上的改变，这给统一 MGWR 和 GWQR 模型提供了契机。统一这两个模型的原理是将分位数回归方法得到的  $\hat{\beta}$  作为投影矩阵  $S$  的构造依据，可以根据不同的分位数得到不同自变量的最佳带宽。从而同时揭示两种解释异质性。然而使用 GWQR 时可能出现某一自变量在某一分位数下显著而在另一分位数下不显著的情况，这给基于变量投影矩阵迭代的 MGWR 模型在寻优上造成麻烦，使得回归结果不够稳健。进一步的结论有待探索。

## 参考文献

- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, volume 4 of *Studies in Operational Regional Science*. Springer Netherlands, Dordrecht.
- Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2):93–115. Publisher: Wiley.
- Atkinson, P. M., German, S. E., Sear, D. A., and Clark, M. J. (2003). Exploring the Relations Between Riverbank Erosion and Geomorphological Controls Using Geographically Weighted Logistic Regression. *Geographical Analysis*, 35(1):58–82.
- Brunsdon, C., Fotheringham, A. S., and Charlton, M. (1999). Some Notes on Parametric Significance Tests for Geographically Weighted Regression. *Journal of Regional Science*, 39(3):497–524. Publisher: Wiley.
- Brunsdon, C., Fotheringham, A. S., and Charlton, M. E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28(4):281–298. Publisher: Wiley.
- Brunsdon, C., Fotheringham, S., and Charlton, M. (1998). Geographically Weighted Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3):431–443.
- Brunsdon, C., Fotheringham, S., and Charlton, M. (1998). Geographically weighted regression - Modelling spatial non-stationarity. *Journal of the Royal Statistical Society Series D: The Statistician*, 47(3):431–443. Publisher: Wiley-Blackwell.
- Buyantuyev, A. and Wu, J. (2010). Urban heat islands and landscape heterogeneity: linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. *Landscape Ecology*, 25(1):17–33.
- Chen, V., Deng, W.-S., Yang, T.-C., and Matthews, S. (2012). Geographically Weighted Quantile Regression (GWQR): An Application to U.S. Mortality Data. *Geographical Analysis*, 44.
- Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*.
- Comber, A., Chi, K., Quang Huy, M., Nguyen, Q., Lu, B., Phe, H., and Harris, P. (2018). Distance metric choice can both reduce and induce collinearity in geographically weighted regression. *Environment and Planning B: Urban Analytics and City Science*, 47:239980831878401.
- Foster, S. A. and Gorr, W. L. (1986). An Adaptive Filter for Estimating Spatially-Varying Parameters: Application to Modeling Police Hours Spent in Response to Calls for Service. *Management Science*, 32(7):878–889. Publisher: INFORMS.
- Fotheringham, A., Brunsdon, C., and Charlton, M. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. *John Wiley & Sons*, 13.

- Fotheringham, A. S., Crespo, R., and Yao, J. (2015). Geographical and Temporal Weighted Regression (GTWR): Geographical and Temporal Weighted Regression. *Geographical Analysis*, 47(4):431–452. Publisher: Wiley.
- Fotheringham, A. S. and Oshan, T. M. (2016). Geographically weighted regression and multicollinearity: dispelling the myth. *Journal of Geographical Systems*, 18(4):303–329. Publisher: Springer Science and Business Media LLC.
- Fotheringham, A. S., Yang, W., and Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *ANNALS OF THE AMERICAN ASSOCIATION OF GEOGRAPHERS*, 107(6):1247–1265. Place: 2-4 PARK SQUARE, MILTON PARK, ABINGDON OX14 4RN, OXON, ENGLAND Publisher: ROUTLEDGE JOURNALS, TAYLOR & FRANCIS LTD Type: Article.
- Fotheringham, S. and M.E. Charlton & C. Brunsdon (1997). Two Techniques for Exploring Non-stationarity in Geographical Data. *Journal of Geographical Systems*, 4:59–82. Publisher: Springer-Verlag.
- Gollini, I., Lu, B., Charlton, M., Brunsdon, C., and Harris, P. (2015). **GWmodel** : An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software*, 63(17).
- Griffith, D. (2008). Spatial-Filtering-Based Contributions to A Critique of Geographically Weighted Regression (GWR). *Environment and Planning A*, 40:2751–2769.
- Huang, B., Wu, B., and Barry, M. (2010). Geographically and Temporally Weighted Regression for Modeling Spatio-Temporal Variation in House Prices. *International Journal of Geographical Information Science*, 24:383–401.
- Jia, Y., Tang, L., Xu, M., and Yang, X. (2019). Landscape pattern indices for evaluating urban spatial morphology -A case study of Chinese cities. *Ecological Indicators*, 99:27–37.
- Leung, Y., Mei, C., and Zhang, W.-X. (2000a). Statistical Tests for Spatial Nonstationary Based on the Geographically Weighted Regression Model. *Environment and Planning A*, 32:9–32.
- Leung, Y., Mei, C., and Zhang, W.-X. (2000b). Testing for Spatial Autocorrelation among the Residuals of the Geographically Weighted Regression. *Environment and Planning A*, 32:871–890.
- Liu, Y., Lam, K.-F., Wu, J. T., and Lam, T. T.-Y. (2018). Geographically weighted temporally correlated logistic regression model. *Scientific Reports*, 8(1):1417. Publisher: Nature Publishing Group.
- Lu, B., Brunsdon, C., Charlton, M., and Harris, P. (2017). Geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*, 31(5):982–998.
- Lu, B., Charlton, M., Brunsdon, C., and Harris, P. (2016). The Minkowski approach for choosing the distance metric in geographically weighted regression. *International Journal of Geographical Information Science*, 30(2):351–368.
- Lu, B., Charlton, M., and Fotheringham, A. S. (2011). Geographically Weighted Regression Using a Non-Euclidean Distance Metric with a Study on London House Price Data. *Procedia Environmental Sciences*, 7:92–97.

Lu, B., Harris, P., Charlton, M., and Brunsdon, C. (2015). Calibrating a Geographically Weighted Regression Model with Parameter-specific Distance Metrics. Procedia Environmental Sciences, 26:109–114.

Mennis, J. L. and Jordan, L. (2005). The Distribution of Environmental Equity: Exploring Spatial Nonstationarity in Multivariate Models of Air Toxic Releases. Annals of the Association of American Geographers, 95(2):249–268.

Oshan, T. M., Smith, J. P., and Fotheringham, A. S. (2020). Targeting the spatial context of obesity determinants via multiscale geographically weighted regression. INTERNATIONAL JOURNAL OF HEALTH GEOGRAPHICS, 19(1). Place: CAMPUS, 4 CRINAN ST, LONDON N1 9XW, ENGLAND Publisher: BMC Type: Article.

Sun, B., Fang, C., Liao, X., Liu, M., Liu, Z., and Guo, X. (2023). Revealing the heterogeneous effects of new urbanization on urban-rural inequality using geographically weighted quantile regression. Applied Geography, 159:103082.

Wheeler, D. (2009). Simultaneous Coefficient Penalization and Model Selection in Geographically Weighted Regression: The Geographically Weighted Lasso. Environment and Planning A, 41:722–742.

Wu, B., Li, R., and Huang, B. (2014). A geographically and temporally weighted autoregressive model with application to housing prices. International Journal of Geographical Information Science, 28(5):1186–1204. Publisher: Taylor & Francis.

卢宾宾, 葛咏, 秦昆, and 郑江华 (2020). 地理加权回归分析技术综述. 武汉大学学报 (信息科学版), 45(9):1356–1366. 152 citations(CNKI)[6-29-2024].