第30卷第1期 安徽建筑大学学报 Vol.30 No.1
2022年1月 Journal of Anhui Jianzhu University Feb.2022

DOI:10.11921/j.issn.2095-8382.20220110

安徽省县域粮食产量及生产投入要素的空间差异分析

刘 斌，张华琼

（安徽建筑大学 数理学院，安徽 合肥 230601）

摘 要：基于2017年安徽省县域粮食产量和生产投入要素的数据，在OLS分析基础上，运用空间自相关分析和GWR模型，探讨了研究区域内粮食产量及其影响因素的空间分布特征。结果表明：（1）粮食产量存在显著的空间正自相关性；（2）农业机械总动力、有效灌溉面积和粮食作物面积对粮食产量的影响均呈现显著的空间非平稳性，表现为同一要素在不同县域对粮食产量的影响有明显差异；（3）GWR模型能有效地消除OLS模型中残差的空间自相关，其各项评价指标均优于OLS模型。

关键词：空间自相关；地理加权回归；粮食产量

中图分类号：F307.11 **文献标识码：**A **文章编号：**2095-8382(2022)01-058-06

Spatial Difference Analysis of Grain Yield and the Factors of Production at the County Level in Anhui Province

LIU Bin，ZHANG Huaqiong

（School of Mathematics and Physics，Anhui Jianzhu University，Hefei 230601，China）

Abstract：Based on the grain yield and the factors of production of counties in Anhui province in 2017，the spatial distribution characteristics of grain yield and its influencing factors are studied with OLS analysis，the spatial auto-correlation and GWR model. The results show that：(i) there is a positive spatial autocorrelation in the grain yield；(ii) the effects of agricultural machinery power，the effective irrigation area and the grain planting area on the grain yield is significant spatially non-stationary，and the effect of the same factor on grain yield varied significantly across counties；(iii) the GWR model can effectively eliminate the spatial autocorrelation of the residual errors in the OLS model，and its evaluation indexes are better than those of the OLS model.

Keywords：spatial auto-correlation；geographical weighted regression；grain yield

民以食为天，保护和提高粮食产量是我国农业发展过程中面临的长期艰巨任务。粮食问题涉及国计民生，关系到国家粮食安全和经济社会的健康发展[1-2]。影响粮食生产的因素多且复杂，由于不同地区的地理条件和发展水平存在差异，粮食产量和其影响因素具有显著的空间变异性[3-5]。了解粮食生产区域格局及影响因素对于制定合理的粮食生产政策和贸易政策有重要的意义[6]。

粮食生产和布局问题一直是国内学术界关注的热点领域。周立青等人采用空间自相关和多元回归模型等方法，探讨了黑龙江省粮食生产的时空动态及其主要影响因素[7]。杨宗辉等人采用空间自相关分析和空间杜宾模型分析了我国省域粮食生产格局变动的影响因素，得出有效灌溉面积、化

收稿日期：2021-08-30

基金项目：国家自然科学基金资助项目（71571002）；安徽建筑大学科研项目（2016QD118）

作者简介：刘斌（1965-），男，教授，博士，主要研究方向为随机运筹学和应用概率；
张华琼（1996-），女，硕士研究生，主要研究方向为统计方法在随机模型中应用。

肥施用量、粮食播种面积及人均耕地面积均对我国粮食产量有显著正效应的结论[8]。胡慧芝等人采用探索性空间数据分析、重心转移模型和空间误差模型研究了长江流域粮食生产的时空格局演变特征及影响因素,结果表明粮食播种面积和农业化肥施用量对粮食生产均具有显著的正向效益[9]。这些研究揭示了区域粮食生产格局的演变特征及影响因素,为优化粮食生产布局及制定粮食生产政策提供了科学依据。然而,现有的研究往往忽视了地理空间要素对粮食生产的影响,鲜有针对粮食产量及生产投入要素的空间差异分析。

县域是我国最基本的行政单元,是粮食生产发展的重点[10]。本文以 2017 年安徽省 58 个县和县级市的粮食产量及影响因素为研究对象,在全局回归分析的基础上,采用空间自相关和地理加权回归分析,对安徽省县域粮食产量及其生产投入要素的空间分布差异进行探索,以期对安徽省的粮食生产的政策制定提供参考。

1　研究区域、数据来源和研究方法

1.1　研究区域

安徽省位于中国华东长江三角洲地区,常年农作物种植面积超过 8.67×10^4 km^2,其中粮食作物面积占 75% 以上。地势有平原、丘陵、山地等类型,全省有淮北平原、江淮波状平原、皖西山地、沿江丘陵平原和皖南山地 5 个地貌区(图 1),地跨淮河、长江、新安江三大水系[11]。安徽省处于中纬度地区,四季分明,全省年平均气温在 14~17 ℃之间,平均日照 1 800~2 500 h,平均无霜期 200~250 天,平均降水量 800~1 800 mm,农业气候条件适宜,农业资源丰富,是典型的农业大省[12]。

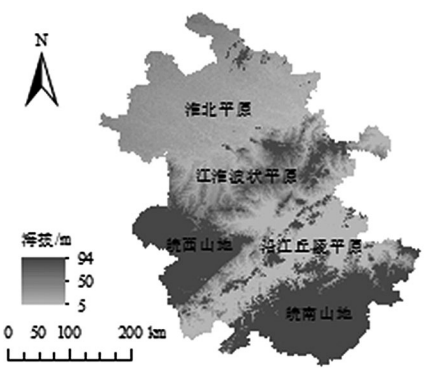

图 1　安徽省地貌分区图

1.2　数据来源及处理

考虑到数据的可得性和完整性,本文研究的样本空间单元为安徽省下辖的 58 个县和县级市(歙县、和县和宁国市由于数据不全未包括在内),自变量为农业机械总动力(kW)、有效灌溉面积(khm^2)和粮食作物面积(khm^2),因变量为粮食产量(10^4t)。数据资料来源于 2018 年《安徽省统计年鉴》以及各市县 2017 年的《国民经济和社会发展统计公报》。

1.3　研究方法

1.3.1　空间自相关分析

空间自相关分析是用空间距离的函数来度量给定变量的样本之间的相似性、描述样本在区域内分布的空间特征的方法[13]。Moran's I 指数是空间自相关性分析中最常用的指数,其计算公式如下:

$$I=\frac{N\sum_i\sum_j w_{ij}\left(x_i-\bar{x}\right)\left(x_j-\bar{x}\right)}{\left(\sum_i\sum_j w_{ij}\right)\sum_i\left(x_i-\bar{x}\right)^2} \tag{1}$$

其中 N 是空间单元样本的数量,x_i 和 x_j 是空间单元 i 和 j 的属性值,w_{ij} 是 i 和 j 之间的空间权重,$\bar{x}$ 是 N 个空间单元的属性值的平均值。Moran's I 的取值范围为 [−1,1],I 值为正且越接近 1,表明空间单元的正相关性越强;I 值为负且越接近 −1,表明空间单元的负相关性越强;I 值越接近于 0,表明样本单元的空间自相关性越低。

1.3.2　GWR 模型

通常的全局回归模型由于不考虑地理位置的信息,只表示在平均意义下因变量和自变量之间的相关关系,因此不能反映出回归参数的空间特征。设 Y 为因变量,$X_1,X_1,\cdots,X_P$ 为自变量,(Y_i, $X_{i1},\cdots,X_{ip}$),i=1,2,$\cdots$,n 是 n 组观测数据,则一般线性回归模型的样本形式可表示为:

$$Y_i=\beta_0+\beta_1X_{i1}+\cdots+\beta_pX_{ip}+\varepsilon_i \tag{2}$$

其中 $\beta_0,\beta_1,\cdots,\beta_p$ 为未知的常值参数,ε_i 为随机误差项,满足 $E(\varepsilon_i)=0$,$Var(\varepsilon_i)=\sigma^2>0$,$Cov(\varepsilon_i,\varepsilon_j)=0$,$i\neq j$。对于式(2)中未知参数的估计一般选用普通最小二乘(简称 OLS)估计。

地理加权回归(简称 GWR)模型作为一种局部回归的技术,将数据的空间位置属性纳入回归模型中,利用回归系数函数在各空间位置处的估计值,分析回归关系随空间位置变化的特征[14]。

GWR 模型将一般线性回归模型改写为如下形式：

$$Y_i=\beta_1(u_i,v_i)+\beta_1(u_i,v_i)X_{i1}+\cdots+\beta_p(u_i,v_i)X_{ip}+\varepsilon_i \quad (3)$$

其中 $Var(\varepsilon_i)=\sigma^2$，$\beta_j(u,v)$，$j=0,1,2,\cdots,p$ 为空间地理位置 (u,v) 的未知函数；$\varepsilon_i(j=1,2,\cdots,n)$ 为独立同分布的误差项且 $E(\varepsilon_i)=0$，对于式（3）中的参数可以通过求解下列矩阵方程来估计：

$$\begin{aligned}\hat{\beta}(u_i,v_i)&=\left(\hat{\beta}_0(u_i,v_i),\hat{\beta}_1(u_i,v_i),\ldots,\hat{\beta}_p(u_i,v_i)\right)^T\\&=\left(X^TW(u_i,v_i)X\right)^{-1}X^TW(u_i,v_i)Y\end{aligned} \quad (4)$$

式中 $\hat{\beta}(u_i,v_i)$ 是 β 的无偏估计，$W(u_i,v_i)$ 是加权矩阵，其作用是确保靠近特定位置的观测点具有更大的权重。本文的研究中使用了高斯加权核函数形式：$W_{ij}=\exp\left(-\frac{d_{ij}^2}{h^2}\right)$，其中 d_{ij} 是回归点 i 和相邻观测点 j 之间的欧氏距离，h 表示核函数的光滑参数，称为带宽，本文中的 h 由修正后的 AICC 准则确定[15]。

2 结果

2.1 OLS 模型结果及分析

以粮食产量为因变量，农业机械总动力、有效灌溉面积和粮食作物面积为自变量，运用 OLS 方法进行回归分析，其估计结果在表 1 中列出。表中 VIF 为方差膨胀因子，是用来度量自变量之间相关程度的指标，VIF 越大，表示共线性越严重。R^2a 为自由度调整的复决定系数；AIC_C 是模型性能的一种度量，如果两个模型的 AIC_C 值相差大于 3，具有较低的 AIC_C 值的模型被视为更佳模型[14]。由表 1 可知，模型的拟合优度达到 98.4%，各个自变量以及模型整体均通过了 5% 水平的显著性检验。每个自变量的 VIF 均不大于 7，模型中不存在严重的多重共线性。通过比较表 1 中各自变量的系数可知，粮食作物面积对粮食产量的影响最大，其次为有效灌溉面积，而农业机械总动力对粮食产量的作用较弱。

通过 GeoDa 软件对粮食产量数据进行全局空间自相关分析可得，Moran’s I=0.488，p 值为 0.001，表明粮食产量的 Moran’s I 指数在 1% 水平上显著为正，安徽省县域粮食产量的空间分布具有显著的正相关性。

为了全面分析安徽省县域粮食产量的空间差异特征，以粮食产量为横轴，粮食产量的空间滞后向量为纵轴绘制的 Moran 散点图，如图 2 所示。Moran 散点图由四个象限构成：分布在第一象限和第三象限的点为空间正自相关的点数据，分别表示某地区与其相邻地区皆有较高（低）的粮食产量，说明区域具有集聚性；分布在第二象限和第四象限的点为空间负自相关的点数据，分别表示粮食产量较低（高）的地区，其相邻地区的粮食产量却较高

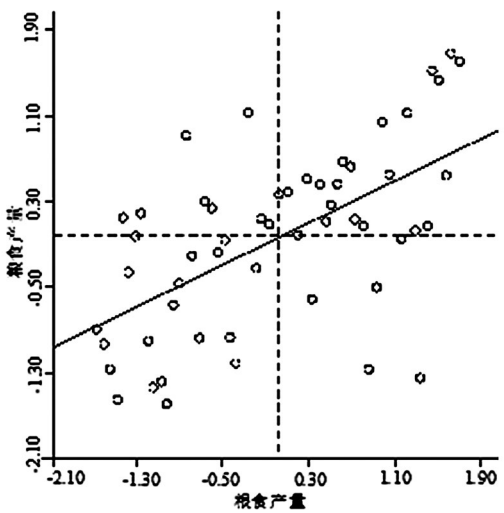

图 2 粮食产量的 Moran’s I 散点图

表 1 OLS 模型参数估计及检验结果

	估计参数	Std.Error	t 值	p 值	VIF	评价指标	指标值
常数项	0.271 992	1.556 251	0.174 774	0.861 912		R^2	0.984
农用机械总动力	−0.000 009	0.000 004	−2.434 998	0.018 224	6.963 568	R^2_a	0.983
有效灌溉面积	0.088 134	0.038 548	2.286 349	0.026 181	3.392 512	AIC_C	392
粮食作物面积	5.924 675	0.215 058	27.549 238	0.000 000	4.729 904		

(低),说明区域具有异质性。如果分布在四个象限的点的个数是均匀的,则说明地区之间不存在空间自相关性。

根据图 2 显示,分布在第一、三象限的点明显多于分布在第二、四象限的点,说明多数县呈现出明显的空间聚集特征。安徽省县域粮食产量的空间分布特征为:粮食产量高的县趋于与粮食产量高的县相邻,粮食产量低的县趋于与粮食产量低的县相邻。进一步对残差进行空间自相关检验,得到 Moran' s I=0.45,p 值为 0.001,说明该 OLS 模型的回归残差并不是随机分布,在空间上存在自相关性。

根据以上分析,安徽省县域粮食产量及 OLS 模型的估计残差皆存在空间相关性。但是,由于全局回归不考虑地理位置的信息,只表征了在平均意义下因变量和自变量的相关关系,不能反映安徽省县域粮食产量在空间上的非平稳性,为此接下来采用 GWR 模型引入空间位置信息以提高模型精度。

2.2　GWR 模型结果及分析

2.2.1　模型构建

为了分析回归关系随空间位置变化的特征,本文采用 GWR 方法,以安徽省县域粮食产量为因变量,农业机械总动力、有效灌溉面积和粮食作物面积为自变量,将数据的空间位置属性纳入回归模型中,有效挖掘回归关系的局部特征。表 2 列出了 GWR 模型的参数估计及检验结果。

根据表 2 显示的结果,GWR 模型的拟合优度达到 99.4%,AICC 为 375.1748,其模型性能较 OLS 模型有显著提高。表 3 给出了 GWR 模型的描述性统计结果,表中包括了各自变量回归系数的最小值、上四分位数、中位数、下四分位数和最大值。进一步对模型中的每个参数进行显著性检验,均有 p 值 <0.001,这表明粮食产量与各变量的关系在空间上是非平稳的。

表 2　GWR 模型参数估计及检验结果

模型参数	Bandwidth	Effective Number	AIC_C	R^2	R^2_a
数值	121 017.864 0	21.994 7	375.174 8	0.994 0	0.990 6

由于农业机械总动力、有效灌溉面积和粮食作物面积对粮食产量的影响存在空间差异性,因此,根据两种模型的估计结果得出的结论不完全一致。具体来看,根据表 1 中 OLS 模型的估计结果,农业机械总动力(–0.000 009)对粮食产量有负向影响,有效灌溉面积(0.088 134)对粮食产量有正向影响,而表 3 中农业机械总动力和有效灌溉面积的 GWR 估计值均有正有负,且绝对值有大有小。以上说明这两个变量对粮食产量的影响方向和程度在不同地区表现不一样,而平均意义上的正向或者负向影响无法体现地区间的差异性。无论是 OLS 模型还是 GWR 模型,在三个自变量中,粮食作物面积对粮食产量的影响皆为最大,且回归系数均为正值,说明其在各地区对粮食产量的影响方向一致。对比粮食作物面积的 GWR 估计值和 OLS 估计值可知,其 GWR 估计值的四分位数范围大于 OLS 估计值的 ±1 倍标准差范围(5.709 617~6.139 733),说明粮食作物面积对粮食产量的影响程度存在较大的空间差异性。

从 GWR 模型的标准化残差分布图(图 3)可以看出,仅 1 个县的回归标准化残差落在了 2.5 倍标准差外。进一步对残差进行空间自相关检验,得到 Moran' s I=0.085,p 值为 0.368,未能通过 1% 的显著性检验,即残差在空间上是随机分布的,这说明 GWR 是一种能有效处理回归分析中空间非平稳性问题的技术。

表 3　GWR 模型回归系数的描述性统计

变量	最小值	上四分位数	中位数	下四分位数	最大值
常数项	–10.583 354	–2.331 338	–0.418 381	0.900 295	5.799 196
农业机械总动力	–0.000 007	–0.000 005	–0.000 003	0.000 002	0.000 008
有效灌溉面积	–0.156 524	–0.014 258	0.026 711	0.072 777	0.161 812
粮食作物面积	5.098 990	5.556 148	5.919 144	6.216 007	6.650 971

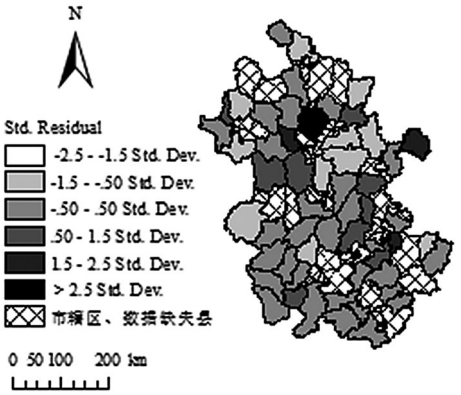

图 3 GWR 模型标准化残差的空间分布

2.2.2 结果分析

由于 GWR 模型的估计参数因地区而异，本文通过每个变量系数的空间分布图进一步解释单个参数的空间非平稳性特征。

（1）农业机械总动力对粮食产量影响的空间变异特征

从图 4 的空间分布来看，农业机械总动力对粮食产量的效应有正有负，回归系数总体上从西南向东北递减。农业机械总动力对粮食产量为正效应的地区主要分布在安徽省的山地和丘陵地域，充分显示了这些区域的农业机械化还有一定的发展空间。这是由于山地和丘陵地区的耕地较分散，耕地形状大小不一且大多根据地势而建，其农业机械化水平的发展受限。由于自然条件的限制，山地和丘陵区域的农业机械化发展水平要落后于平原区域。因此，在这些区域加大农机化的扶持政策，扩大农机化的投入，提升农机化水平，是提高粮食产量的重要手段。农业机械总动力对粮食产量呈负效应的地区主要分布在安徽省的平原区，说明平原区的农业机械化水平较高，处于相对饱和状态。

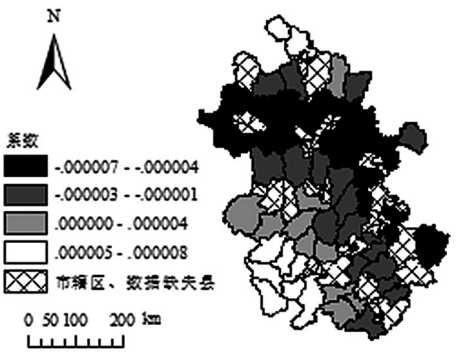

图 4 农业机械总动力回归系数的空间分布

（2）有效灌溉面积对粮食产量影响的空间变异特征

从图 5 可以看出，有效灌溉面积的回归系数由西部向东北和东南两个方向呈逐渐递减趋势，正值主要见于安徽省的江淮波状平原、皖西山地、淮北平原中南部和沿江丘陵平原；而负值则主要分布在淮北平原的东部及北部和皖南山地。从地理位置上看，淮北平原中南部和沿江丘陵平原除降水直接利用外，粮食作物还得到地下水的补给，因此这些区域的粮食生产对水利灌溉的依赖程度比较小；江淮波状平原和皖西山地水资源较为紧张，自然降水一般难以满足作物的需求，因此粮食生产对补充灌溉的依赖比较高，加之这些区域的水利工程还不完善，有效灌溉面积对粮食产量有显著的正影响。江淮波状平原和皖西山地对农业灌溉的依赖程度比较高，有效灌溉面积的提高对粮食产量的影响最为明显，这些地区应当加大农田水利工程投入，完善灌溉工程的建设。

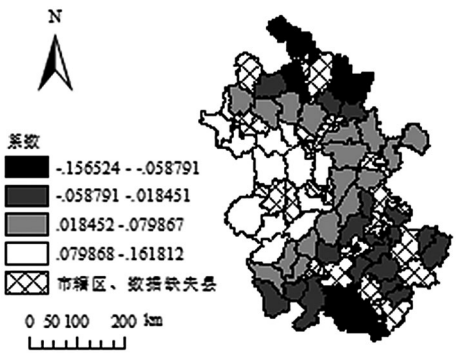

图 5 有效灌溉面积回归系数的空间分布

（3）粮食作物面积对粮食产量影响的空间变异特征

根据图 6 可知，研究区域内的粮食作物面积对粮食产量的影响皆为正向，总体空间效应呈现出由安徽省西部向东北和东南两个方向递增的态势。在农业机械总动力、有效灌溉面积和粮食作物面积中，仅粮食作物面积的回归系数在所有区域中皆为正值，且回归系数的值普遍较高，这说明粮食作物面积是影响粮食产量的最主要因素，扩大粮食作物的种植面积能显著提高粮食产量。从回归系数的空间分布来看，粮食作物面积对粮食产量影响较大的区域主要位于沿江丘陵平原和皖南山地，粮食作物面积对粮食产量影响较小的区域主要分布在安

徽省西部地区。

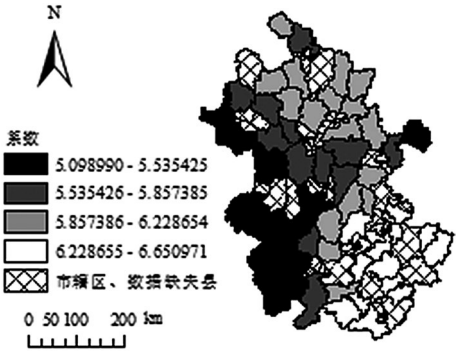

图 6　粮食作物面积回归系数的空间分布

3　结论和讨论

安徽省县域粮食产量存在显著的空间正相关性,在空间上表现出明显的聚集特征:粮食产量高的县趋于与粮食产量高的县相邻,粮食产量低的县趋于与粮食产量低的县相邻。在县域粮食产量与生产投入要素的回归分析中,GWR 模型较 OLS 模型在性能上有明显提升,GWR 可以有效地处理回归关系的空间非平稳性问题。在安徽省各区域,粮食作物面积始终是影响粮食产量的最主要因素,保障粮食的种植面积是促进安徽省粮食生产持续发展的重要途径。山地和丘陵区域主要受农业机械总动力的正向影响,在这些区域加大农机化投入,提升农机化水平,可进一步提高粮食产量。江淮波状平原和皖西山地受有效灌溉面积的正向影响,这些地区应当重视农业水利的建设,提高有效灌溉面积。

本文的回归分析仅考虑了空间因素,而粮食产量与其影响因素同样也受到时间因素的影响,因此,用 GWR 方法得出的分析结果还有局限性。另外,虽然 GWR 模型性能优于 OLS 模型,但仍存在一些不足之处,比如它假定所有的变量都具有空间非平稳性,而实际中可能会出现部分变量不受空间位置影响的情况。未来的研究可以把空间和时间因素都纳入回归模型中,改变变量或增加变量的多样性,以期得到更加客观科学的分析结果。

参考文献:

[1] 尹成杰 . 关于提高粮食综合生产能力的思考 [J]. 农业经济问题,2005,26(1):5-10,79.

[2] 郭燕枝,郭静利,王秀东 . 我国粮食综合生产能力影响因素分析 [J]. 农业经济问题,2007,28(S1):22-25.

[3] 李月,孔祥斌,张安录,等 . 基于 LMDI 模型的我国省域粮食生产变化影响因素分析 [J]. 中国农业大学学报,2016,21(1):129-140.

[4] 张利国,鲍丙飞 . 我国粮食主产区粮食全要素生产率时空演变及驱动因素 [J]. 经济地理,2016,36(3):147-152.

[5] 李蓬勃,闫晓冉,徐东瑞 .BP 神经网络和多元线性回归在粮食产量空间分布预测中的比较 [J]. 干旱区资源与环境,2014,28(9):74-79.

[6] 罗万纯,陈永福 . 中国粮食生产区域格局及影响因素研究 [J]. 农业技术经济,2005(6):156.

[7] 周立青,程叶青 . 黑龙江省粮食生产的时空格局及动因分析 [J]. 自然资源学报,2015,30(3):491-501.

[8] 杨宗辉,蔡鸿毅,覃诚,等 . 我国粮食生产的时空格局及其影响因素分析 [J]. 中国农业科技导报,2018,20(9):1-11.

[9] 胡慧芝,王建力,王勇,等 .1990—2015 年长江流域县域粮食生产与粮食安全时空格局演变及影响因素分析 [J]. 长江流域资源与环境,2019,28(2):359-367.

[10] 郭成,储金龙,李咏 . 公共产品视角下安徽省县域城镇化水平时空差异分析 [J]. 安徽建筑大学学报,2015,23(06):87-94.

[11] 程言新,张福生,王婉茹,等 . 安徽省地貌分区和分类 [J]. 安徽地质,1996,6(1):63-69.

[12] 李想 . 粮食主产区农户技术采用及其效应研究——以安徽省水稻可持续生产技术为例 [D]. 北京:中国农业大学,2014.

[13] Moran P A P.Notes on continuous stochastic phenomena[J]. Biometrika,1950,37(1/2):17-23.

[14] McMillen D P.Geographically weighted regression:the analysis of spatially varying relationships[J].American Journal of Agricultural Economics,2004,86(2):554-556.

[15] Brunsdon C,Fotheringham A S,Charlton M.Geographically weighted summary statistics -- a framework for localised exploratory data analysis.Computers,Environment and Urban Systems,2002,26 (6):501-524.