

环渤海沿海区域耕地格局及影响因素分析

吴 莉^{1,2}, 侯西勇¹, 徐新良³

(1. 中国科学院烟台海岸带研究所, 烟台 264003; 2. 中国科学院大学, 北京 100049;

3. 中国科学院地理科学与资源研究所, 北京 100101)

摘 要: 为分析环渤海省市沿海区域耕地格局与影响因素的关系, 以耕地在 5 km×5 km 网格单元所占比例为因变量, 选用地形、距离、气候及人口等 10 个影响因素为自变量, 分别建立普通最小二乘法线性回归模型、空间滞后模型、空间误差模型、地理加权回归模型。结果表明: 耕地格局及各影响因素均呈现较强的空间正相关, 并随距离增大而减少; 针对该研究, 空间滞后模型、空间误差模型和地理加权回归模型模拟效果均优于普通最小二乘法线性回归模型, 空间误差模型优于空间滞后模型; 从全局上来讲, 高程、坡度、到最近公路距离与耕地格局呈负相关影响, 距最近海岸线、铁路、居民点距离、多年平均气温和多年平均降水与耕地格局呈正相关。从局部上来讲, 除了多年平均降水对各网格单元内耕地面积均呈正向影响外, 其余影响因素随网格单元变化正负向影响均存在。多年平均气温和多年平均降水是主要的、最敏感的正向影响因素, 高程、坡度和距最近水系距离为主要的、最敏感的负向影响因素。

关键词: 土地利用; 回归分析; 农村地区; 环渤海沿海区域; 耕地; 线性回归; 空间自回归; 地理加权回归

doi: 10.3969/j.issn.1002-6819.2014.09.001

中图分类号: F301.21

文献标志码: A

文章编号: 1002-6819(2014)-09-0001-10

吴 莉, 侯西勇, 徐新良. 环渤海沿海区域耕地格局及影响因素分析[J]. 农业工程学报, 2014, 30(9): 1 - 10.
Wu Li, Hou Xiyong, Xu Xinliang. Analysis of spatial pattern of farmland and its impacting factors in coastal zone of Circum Bohai[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(9): 1 - 10. (in Chinese with English abstract)

0 引 言

耕地是全球最普遍的土地利用类型之一, 其变化对粮食安全、社会稳定, 甚至一个国家或地区的社会经济可持续发展均有重要影响^[1]。中国人多地少, 人口持续增长、快速的城市化和工业化导致人地矛盾日益加剧, 城镇扩展, 退耕还林、还草, 国土资源开发等宏观政策均会对耕地产生影响^[2-3]。中国总的耕地面积由改革开放初期的 $1.32 \times 10^6 \text{ km}^2$ 减少到 2011 年的 $1.21 \times 10^6 \text{ km}^2$, 2011 年末人均耕地面积仅为 0.00091 km^2 ^[4], 因此耕地乃至后备耕地资源得到越来越多的关注^[5]。目前已有不少学者从全国或区域尺度对耕地的总量变化、时空变化、总量变化影响因素以及耕地保护价值等问题进行了

研究^[6-10], 然而专门针对耕地格局及影响因素的定量化研究较少。已有的研究多是在土地利用格局模拟过程中建立各种土地利用类型空间分布的适宜性图像, 耕地仅是其中的一种用地类型, 而并未对耕地格局与影响因素之间的关系进行深入地分析和讨论^[11-12]。

耕地格局作为一种空间数据具有空间依赖性的特点, 即存在空间自相关性。传统回归模型的理论假设前提是数据本身在统计上是独立且均匀分布的^[13], 忽略了空间自相关的影响, 如果用于空间数据分析会产生偏差。Anselin^[14]提出了适合用来处理空间数据的空间自回归模型, 并在土地利用领域中广泛应用^[15-18]。由于空间自回归模型在本质上属于全局模型, 在耕地格局影响因素研究时, 忽略了影响因素的空间分异性。地理加权回归模型为局部模型, 在各地理空间位置上所估计的参数随研究的地理空间位置变化^[19-26], 可用于反映各影响因素对耕地格局作用的空间非平稳性。

环渤海沿海区域位于中国沿太平洋西岸的北部, 经济增长和城镇化进程较快, 耕地损失及质量下降等问题较为突出^[27-28], 因此对该区域耕地格局研究具有重要意义。本文以该区域耕地格局为研究对象, 针对地形、距离、气候及人口等 10 个影响

收稿日期: 2013-11-24 修订日期: 2014-03-26

基金项目: 中国科学院战略性先导科技专项 (XDA05130703); 中国科学院重点部署项目 (KZZD-EW-14); 中国科学院知识创新工程重要方向项目 (kzcx2-yw-224); 国家自然科学基金项目 (40801016)。

作者简介: 吴 莉 (1987 -), 女 (汉族), 山东菏泽人, 博士生, 主要从事海岸带土地利用/覆盖变化研究。烟台 中国科学院烟台海岸带研究所, 264003。Email: lwu@yic.ac.cn

通信作者: 侯西勇 (1975 -), 男, 山东泰安人, 研究员, 博士生导师, 主要从事海岸带土地利用/覆盖变化、海岸带综合管理方面的研究。烟台 中国科学院烟台海岸带研究所, 264003。Email: xyhou@yic.ac.cn

因子,分别应用经典线性回归模型、空间自回归模型(空间滞后模型和空间误差模型)、地理加权回归模型对耕地的空间格局进行建模,并依据相关评价指标对模型的拟合精度进行评价。从全局和局部的2个水平上探讨耕地格局与影响因子的回归关系及空间变异性。

1 研究区概况及数据来源

1.1 研究区概况

环渤海沿海区域跨16个地级市和1个直辖市,由北向南依次为丹东、大连、营口、盘锦、锦州市、葫芦岛、秦皇岛、唐山、天津、沧州市、滨州市、东营、潍坊市、烟台市、威海、青岛、日照。研究区总面积约17万 km^2 ,

位于 $115^{\circ}41' \sim 125^{\circ}42'E$ 与 $35^{\circ}4' \sim 42^{\circ}8'N$ 之间。

1.2 数据来源与处理

本文选取2009/2010年云量小于5%的Landsat-5 TM影像为数据源,共涉及20景影像,轨道号为p120r31、p118r32、p119r32、p120r32、p121r32、p122r32、p19r33、p120r33、p121r33、p122r33、p123r33、P119r34、p120r34、p121r34、p122r34、p123r34、p119r35、p120r35、p121r35和p120r36。选择4—3—2波段假彩色合成,在ArcGIS软件中参照DEM、坡度等信息,目视解译得到2010年耕地分布矢量数据。将研究区划分为 $5\text{ km} \times 5\text{ km}$ 规则网格,与耕地矢量数据进行叠加,统计耕地在每个网格单元中所占的百分比,作为回归分析的因变量。

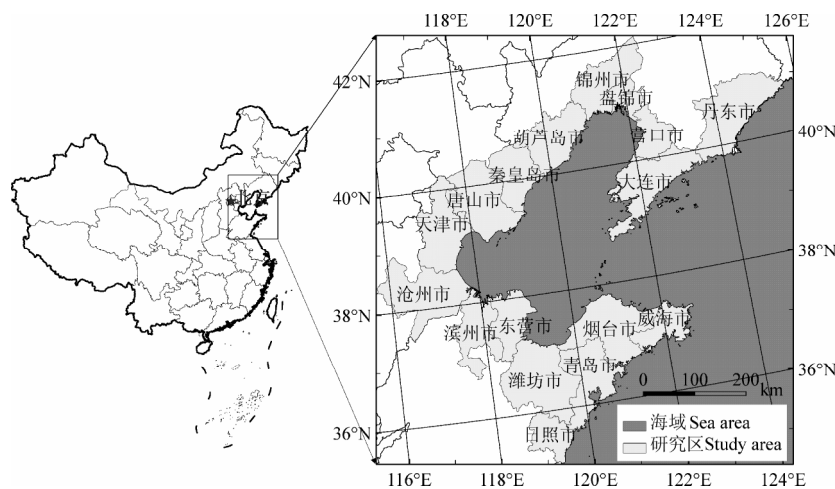


图1 研究区位置

Fig.1 Location of study area

耕地格局受自然和社会经济多种因素交互影响。本研究在综合考虑各种自然环境和社会经济因子的基础上,兼顾数据的可获取性以及满足定量化、空间化的要求,避免因子选择过于繁琐而影响信息的采集以及造成信息冗余,或因因子太少而影响评价的准确性。最终选取10个耕地格局影响因子,具体如下:1)地形因子,包括高程(x_1)和坡度(x_2);2)可达性因子,包括到最近海岸线距离(x_3)、到最近铁路距离(x_4)、到最近公路距离(x_5)以及到最近居民点距离(x_6);3)水资源因子,即到最近水系距离(x_7);4)气候因子,包括多年平均降水(x_8)和多年平均气温(x_9);5)社会经济因子,即人口密度(x_{10})。各影响因子信息的基本统计单元统一为 $5\text{ km} \times 5\text{ km}$ 。其中,降水和气温是通过气象站数据空间插值得到;DEM数据来自于USGS网站(<http://glovis.usgs.gov/>),并据其生成坡度数据;距离因子通过ArcGIS的空间分析模块计算各点到最近城镇、公路等的距离而得到;公路数据来源于中国科学院资源环境科学数据库,包括高速公

路、国道、省道和一般公路等类型。查阅中国区域经济统计年鉴(2011)获得环渤海省市沿海区域各县的行政区划面积和年底总人口数据。年底总人口除以行政区划面积得到县级行政单元的平均人口密度。每个 $5\text{ km} \times 5\text{ km}$ 网格中人口密度的赋值原则是:1)若网格完全落入某县级行政单元内,则将该县级行政单元的人口密度赋予网格。2)若网格落在县级行政单元的交界线处,则将网格内面积最大的县级行政单元的人口密度赋予该网格。将各影响因子图层的值标准化到0~1之间。对10个影响因子进行Pearson相关分析、容差和方差膨胀因子检验,表明10个因子之间不存在多重共线性或共线性现象较弱。

2 研究方法

本研究以耕地格局为因变量,以高程、坡度、到最近海岸线距离、到最近铁路距离、到最近公路距离、到最近居民点距离、到最近水系距离、多年平均降水、多年平均气温以及人口密度为自变量,

判断因变量、自变量是否存在空间自相关，然后分别采用普通最小二乘法线性回归模型、空间自回归模型（空间滞后模型和空间误差模型）和地理加权回归模型对耕地格局及其影响因子进行建模。

2.1 空间自相关

Moran's I 指数是一种应用非常广泛的全局空间自相关统计量，公式如下^[13]：

$$I = \frac{m}{S_0} \frac{\sum_{k=1}^m \sum_{l \neq i}^m w_{kl} (x_k - \bar{x})(x_l - \bar{x})}{\sum_{k=1}^m (x_k - \bar{x})^2} \times 100\% \quad (1)$$

式中： m 为研究区空间单元的总数； x_k 和 x_l 分别代表第 k 个和第 l 个空间单元上的观测值； \bar{x} 是观测值的平均值； w_{kl} 是空间权重； S_0 是空间权重矩阵 w_{kl} 中所有元素之和。Moran's I 指数取值在 $-1 \sim 1$ 之间。在给定显著性水平下，当 Moran's I 指数显著且为正时，表明观测值之间显著正相关，相似属性聚集在一起；当 Moran's I 指数显著且为负时，表示观测值之间存在显著负相关，呈空间分散格局；当 Moran's I 指数接近于 0，观测值在空间上随机分布。

2.2 空间自回归模型

Anselin 空间自回归方程的通用形式为^[14]：

$$\begin{aligned} y &= \rho W_1 y + X\beta + \varepsilon \\ \varepsilon &= \lambda W_2 + \mu \\ \mu &\sim N(0, \sigma^2 I_n) \end{aligned} \quad (2)$$

式中： y 为因变量； X 为自变量； W_1 、 W_2 均为 $n \times n$ 阶权重矩阵，分别反映因变量自身、残差的空间趋势； ρ 为空间滞后变量 $W_1 y$ 的系数； β 是与自变量 X 相关的 $k \times 1$ 参数向量； ε 为随机误差项向量； λ 为空间自回归结构 W_2 的系数， μ 为正态分布的随机误差向量； σ^2 为方差； I_n 为残差的空间自相关系数。

当式 (2) 参数向量的不同子向量设置为 0 时，可以产生 4 种常见的空间模型结构^[29-30]，本文只讨论其中的 2 种：当 $\rho = 0$ ， $\beta = 0$ ， $\lambda = 0$ 时，为空间滞后模型 (spatial lag model, SLM)，它考虑了因变量的空间相关性，即某一空间对象上的因变量不仅与同一对象自变量相关，还与相邻对象的因变量有关；当 $\rho = 0$ ， $\beta = 0$ ， $\lambda \neq 0$ 时，为空间误差模型 (spatial lag model, SEM)，即某一空间对象上的因变量不仅与同一对象自变量相关，还与相邻对象的自变量和因变量有关。

2.3 地理加权回归模型

普通最小二乘法 (ordinary least squares linear regression, OLS) 线性回归模型为全局线性回归模型，用全部自变量值估计因变量的值，用最小二乘

法估计多元线性回归方程的未知参数。其表达式为：

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i \quad (3)$$

式中： $i=1, \dots, n$ 是观测值的个数； y_i 为因变量，它是自变量 x_{ij} ($j=1, \dots, k$) 的一个线性组合； β_0 为常数项， β_j 为第 j 个回归参数， ε_i 为符合正态分布的随机项误差。

地理加权回归 (geographically weighted regression, GWR) 模型是对 OLS 的扩展，实质是局部加权最小二乘法，通过在每个观测点使用加权最小二乘法对参数向量进行估计^[31]。模型表达式如下：

$$y_i = \beta_0(\mu_i, \nu_i) + \sum_{j=1}^k \beta_j(\mu_i, \nu_i) x_{ij} + \varepsilon_i \quad (4)$$

式中： (μ_i, ν_i) 第 i 个观测点的坐标 (如经纬度)，即位置 i 的坐标； $\beta_j(\mu_i, \nu_i)$ 是第 i 个观测点上的第 j 个回归参数；

在 GWR 中，观察值的权重在回归过程中不再保持不变，而是通过与位置 i 的邻接程度来进行加权，即

$$\hat{\beta}(\mu_i, \nu_i) = [X^T W(\mu_i, \nu_i) X]^{-1} X^T W(\mu_i, \nu_i) X Y \quad (5)$$

式中： $\hat{\beta}$ 是 β 的估计值； X 为由自变量观测值构成的矩阵； Y 是由因变量观测值构成的矩阵； W 是对位置 (μ_i, ν_i) 刻画模型时赋予数据点的空间权重矩阵。在空间权重矩阵函数选择上，本研究采用高斯函数。

2.4 模型拟合度的测量

经典决定系数 R^2 和类决定系数 R^2 分别用于评价传统回归模型和空间自回归模型，经典决定系数 R^2 和类决定系数 R^2 两者之间不可比较，但可以比较不同空间自回归模型的类决定系数 R^2 。最大似然对数值 (maximum likelihood logarithm, LIK)、Akaike 信息指标 (Akaike information criterion, AIC) 和 Schwartz 指标 (Schwartz criterion, SC) 也可用来评价模型的拟合精度。如果模型中 LIK 越大 (AIC、SC 越小)，则模型的拟合优度越高，但 LIK 不像经典决定系数 R^2 ，不能作为评价模型拟合优度的绝对指标。残差平方和可用来评价回归模型的拟合效果，数据残差平方和越小，其拟合优度越高。

3 结果与分析

3.1 耕地格局及空间自相关性分析

3.1.1 耕地格局

2010 年环渤海省市沿海区域耕地面积为 92 558 km²，约占区域总面积的 53%。耕地在覆盖研究区的 5 km × 5 km 网格单元中所占比例的分布

如图 2 所示。由图 2 可以看出,耕地主要分布在锦州市东部、盘锦、沧州、潍坊、威海、青岛,在丹东市、葫芦岛市南部、秦皇岛市北部分布比例较低。

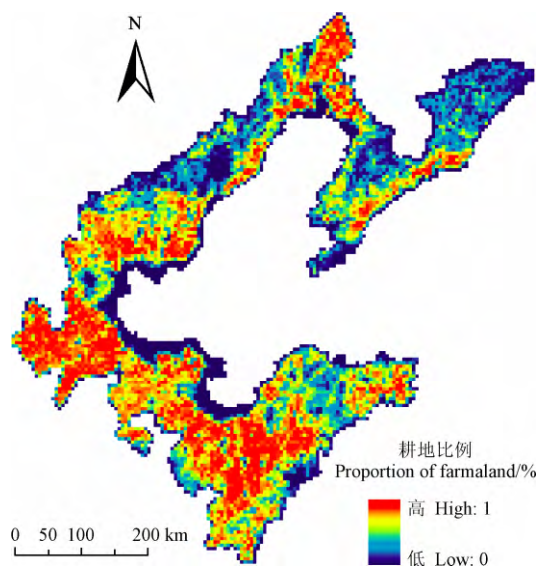
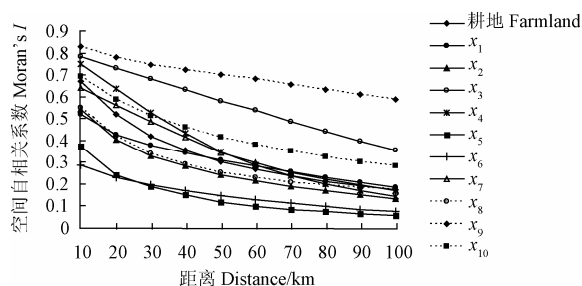


图 2 2010 年环渤海省市沿海区域耕地
5 km×5 km 网格分布图

Fig.2 Distribution of farmland at grids of 5 km×5 km in coastal zone of Bohai Rim region in 2010

3.1.2 空间自相关性分析

耕地格局与各影响因子的空间自相关分析(图 3)表明:耕地格局及各影响因子均呈现出较强的空间正相关,并随着距离的增大而减少,10 种影响因子中,多年平均降水在各距离权重下均表现出最强的空间相关性。



注: X_1 、 X_2 、 X_3 、 X_4 、 X_5 、 X_6 、 X_7 、 X_8 、 X_9 、 X_{10} 分别表示高程、坡度、到最近海岸线距离、到最近铁路距离、到最近公路距离、到最近居民点距离、到最近水系距离、多年平均降水、多年平均气温、人口密度,下同。
Note: X_1 、 X_2 、 X_3 、 X_4 、 X_5 、 X_6 、 X_7 、 X_8 、 X_9 、 X_{10} represent elevation, slope, distance to nearest coastline, distance to nearest railway, distance to nearest road, distance to nearest residential area, distance to nearest river, average yearly precipitation, average yearly temperature, population density, the same below.

图 3 环渤海沿海区域耕地格局及影响因子空间自相关图

Fig.3 Spatial correlograms of cultivated land and influencing factors in coastal zone of Bohai Rim region

3.2 回归模型建立

OLS 模型和空间自回归模型可以从整体上研

究各影响因子对耕地格局的影响。OLS 模型基于最小二乘法估计回归参数,空间自回归模型基于最大似然估计法估计回归参数。从传统线性回归模型的 4 个拉格朗日乘数统计量检验结果(表 1)可以看出,LM-lag、Robust LM-lag、LM-error、Robust LM-error 均显著,不能判断出空间滞后模型和空间误差模型在刻画耕地格局中的优劣。因此,分别建立空间滞后模型和空间误差模型来确定模拟效果最好的空间自回归模型。输入全部自变量,构建最小二乘法线性回归模型-1 (ordinary least squares linear regression model-1, OLS-1)、空间误差模型-1 (spatial error model-1, SEM-1)、空间滞后模型-1 (spatial lag model-1, SLM-1),然后剔除不显著的影响因子建立最小二乘法线性回归模型-2 (ordinary least squares linear regression model-2, OLS-2)、空间误差模型-2 (spatial error model-2, SEM-2)、空间滞后模型-2 (spatial lag model-2, SLM-2) (表 2)。

表 1 传统线性回归模型拉格朗日乘数统计量检验

Table 1 Lagrange multiplier test statistics of standard linear regression

Lagrange 乘数统计量 Lagrange multiplier test statistics	t 值 t value	p 值 p value
LM-lag	7 453.0787	<0.001
Robust LM-lag	430.3178	<0.001
LM-error	8 986.2542	<0.001
Robust LM-error	1 963.4934	<0.001

OLS 模型和空间自相关模型属于全局模型,计算出的同一个影响因子在不同地理单元中的贡献是相同的,不能反映变量的空间分异特征。因此,同时建立 GWR 模型,以探测空间数据的空间非平稳性。

3.3 模型比较

将输入全部影响因子与剔除掉不显著影响因子的 OLS、SEM、SLM 分别比较(表 3)可知: OLS-2 与 OLS-1 相比, OLS-2 和 OLS-1 的 R^2 值相同,残差平方和、最大似然对数值则表明 OLS-1 模型的模拟优度略高于 OLS-2,而 AIC 值和 SC 值则表明 OLS-2 模型的模拟优点略微优于 OLS-1。由于 OLS-1 和 OLS-2 的各评价指标值差异很小,即模型的拟合度变化不大,而 OLS-2 中所有的自变量都呈显著性水平(表 2),因此选用 OLS-2 模型代表最小二乘法线性回归模型与其他类型模型进行比较;针对于空间误差模型,在类 R^2 值相同的情况下,残差平方和、AIC 和 SC 值均表明 SEM-2 优于 SEM-1;空间滞后模型中 SLM-2 的类 R^2 高于 SLM-1 且残差平方和较 SLM-1 大幅度下降,显著优于 SLM-1;综合评价,认为 OLS-2、SEM-2 与 SLM-2 的拟合

效果分别优于 OLS-1、SEM-1 与 SLM-1。

全局模型 SEM-2、SLM-2 与 OLS-2 相比较，由残差平方和、LIK、AIC 及 SC 值可知，SEM-2 和 SLM-2 模型均优于 OLS-2 模型。比较 SEM-2 和 SLM-2，发现 SEM-2 模型类 R^2 和 LIK 值均大于 SLM-2，且其残差平方和、AIC 和 SC 值低于 SLM-2，

均表明 SEM-2 的模拟优度要优于 SLM-2 模型。通过其以上分析可知，在全局水平上，针对于本区域耕地格局影响因子分析 SEM-2 的拟合效果最好。

GWR 模型能在局部水平上对耕地格局进行的解释，其解释程度比 OLS-2 提高了 15% 以上，但其残差平方和与 AIC 值均高于 SEM-2 和 SLM-2 模型。

表 2 回归模型参数

Table 2 Model parameters of different regression models

回归模型 Regression model	标量 Variables	回归系数 Coefficient	标准 误差 S.E.	t 值 t value	p 值 p value	回归模型 Regression model	标量 Variables	回归系数 Coefficient	标准 误差 S.E.	t 值 t value	p 值 p value
OLS-1	常数	0.1336	0.0062	21.44	<0.001	OLS-2	常数	0.1333	0.0062	21.44	<0.001
	x_1	-0.5233	0.0281	-18.63	<0.001		x_1	-0.5218	0.0280	-18.63	<0.001
	x_2	-0.3389	0.0286	-11.85	<0.001		x_2	-0.3393	0.0286	-11.87	<0.001
	x_3	0.3952	0.0126	31.38	<0.001		x_3	0.3942	0.0125	31.51	<0.001
	x_4	0.1070	0.0124	8.66	<0.001		x_4	0.1079	0.0123	8.78	<0.001
	x_5	-0.1973	0.0365	-5.41	<0.001		x_5	-0.1961	0.0364	-5.38	<0.001
	x_6	0.0601	0.0294	2.04	0.041		x_6	0.0579	0.0293	1.98	0.048
	x_7	-0.4916	0.0456	-10.79	<0.001		x_7	-0.5020	0.0433	-11.60	<0.001
	x_8	0.4708	0.0094	50.32	<0.001		x_8	0.4684	0.0088	53.39	<0.001
	x_9	0.0942	0.0115	8.17	<0.001		x_9	0.0946	0.0115	8.22	<0.001
	x_{10}	-0.0154	0.0208	-0.74	0.461		x_{10}	—	—	—	—
SEM-1	常数	0.2202	0.0144	15.30	<0.001	SEM-2	常数	0.2210	0.0144	15.32	<0.001
	x_1	-0.3370	0.0203	-16.61	<0.001		x_1	-0.3383	0.0202	-16.72	<0.001
	x_2	-0.0933	0.0226	-4.14	<0.001		x_2	-0.0943	0.0225	-4.19	<0.001
	x_3	0.3466	0.0181	19.16	<0.001		x_3	0.3498	0.0178	19.63	<0.001
	x_4	0.0801	0.0175	4.58	<0.001		x_4	0.0789	0.0174	4.52	<0.001
	x_5	-0.1487	0.0277	-5.36	<0.001		x_5	-0.1519	0.0276	-5.51	<0.001
	x_6	0.0425	0.0196	2.16	0.030		x_6	0.0405	0.0195	2.08	0.030
	x_7	-0.0512	0.0450	-1.14	0.255		x_7	—	—	—	—
	x_8	0.2276	0.0099	23.02	<0.001		x_8	0.2271	0.0096	23.77	<0.001
	x_9	0.1648	0.0195	8.45	<0.001		x_9	0.1646	0.0195	8.45	<0.001
	x_{10}	0.0095	0.0248	0.38	0.702		x_{10}	—	—	—	—
SLM-1	λ	0.8878	0.0063	141.62	<0.001	SLM-2	λ	0.8883	0.0063	142.09	<0.001
	ρ	0.7871	0.0068	115.39	<0.001		ρ	0.7894	0.0067	118.10	<0.001
	常数	-0.0718	0.0041	-17.55	<0.001		常数	-0.0729	0.0041	-17.95	<0.001
	x_1	-0.2236	0.0183	-12.21	<0.001		x_1	-0.2352	0.0162	-14.49	<0.001
	x_2	-0.0246	0.0184	-1.34	0.182		x_2	—	—	—	—
	x_3	0.1346	0.0087	15.46	<0.001		x_3	0.1344	0.0086	15.57	<0.001
	x_4	0.0688	0.0079	8.67	<0.001		x_4	0.0685	0.0079	8.64	<0.001
	x_5	-0.0851	0.0233	-3.65	<0.001		x_5	-0.0856	0.0233	-3.68	<0.001
	x_6	0.0603	0.0188	3.21	0.001		x_6	0.0584	0.0187	3.12	0.002
	x_7	-0.0214	0.0293	-0.73	0.465		x_7	—	—	—	—
	x_8	0.2036	0.0066	30.91	<0.001		x_8	0.2034	0.0066	30.95	<0.001
	x_9	0.0894	0.0074	12.04	<0.001		x_9	0.0884	0.0074	12.01	<0.001
	x_{10}	-0.0284	0.0133	-2.13	0.033		x_{10}	-0.0319	0.0126	-2.53	0.012

注：OLS-1、OLS-2、SEM-1、SEM-2、SLM-1、SLM-2、 λ 、 ρ 分别代表普通最小二乘法模型-1、普通最小二乘法模型-2、空间误差模型-1、空间误差模型-2、空间误差模型-1、空间误差模型-2，空间误差项回归系数，空间滞后项回归系数。

Note: OLS-1, OLS-2, SEM-1, SEM-2, SLM-1, SLM-2, λ , ρ represent ordinary least squares model-1, ordinary least squares model-2, spatial error model-1, spatial error model-2, spatial lag model-1, spatial lag model-2, coefficient to the spatial error, coefficient to the spatial lag.

表 3 回归模型参数比较
Table 3 Comparison of the parameters of regression results

	R^2	类 R^2 Preudo R^2	残差平方和 residual sum of squares	最大似然对数值 Maximum likelihood logarithm	Akaike 信息指标 Akaike information criterion	Schwartz 指标 Schwartz criterion
OLS-1	0.5027	—	377.48	707.29	-1392.58	-1316.09
OLS-2	0.5027	—	377.51	707.02	-1394.04	-1324.50
SEM-1	—	0.8127	142.18	3797.53	-7573.05	-7496.56
SEM-2	—	0.8127	142.16	3796.88	-7575.76	-7513.18
SLM-1	—	0.7968	377.50	3693.63	-7363.27	-7279.82
SLM-2	—	0.7970	154.11	3692.46	-7364.91	-7295.37
GWR	0.6635	—	255.43	—	-4132.72	—

3.4 耕地格局影响因子分析

3.4.1 全局水平影响因子分析

由 SEM-2 模型中各回归系数值 (表 2) 可以看出, 从区域整体上来讲, 耕地格局与距水系距离 x_7 和人口密度 x_{10} 的回归关系不具有统计学意义, 与高程 x_1 、坡度 x_2 和到最近公路距离 x_5 负相关, 而其他因子呈正相关关系。也就是说, 耕地表现出倾向于向降水、温度高, 地势低、坡度小的区域以及靠近公路分布的特征。从回归系数的绝对值大小可以看出, 耕地受到最近海岸线距离和高程影响较

大, 其次为多年平均气温和多年平均降水, 受到最近居民点距离影响最小。

3.4.2 局部水平影响因子分析

表 4 列出了 GWR 模型参数估计值的描述性统计分析结果。可见, 多年平均降水 x_8 对各网格内耕地比例均呈正相关影响, 而其余影响因子随网格单元变化正负向影响都存在, 但正负比例有差异。高程 x_1 、坡度 x_2 、到最近公路距离 x_5 、到最近水系距离 x_7 和人口密度 x_{10} 对各网格内耕地比例产生负向影响的比例超过正向影响, 其余因素与之相反。

表 4 GWR 模型参数估计值的统计描述
Table 4 Statistical description of GWR's parameter estimation

描述指标 Descriptive index	最小值 Minimum	中位数 Median	最大值 Maximum	平均值 Mean	正值比率 Percentage of positive value/%	负值比率 Percentage of negative value/%
常数项	0.0139	0.0866	0.2333	0.1189	100.00	0
高程 x_1	-2.6295	-0.6420	22.8288	0.9584	15.13	84.87
坡度 x_2	-8.7645	-0.3453	4.7009	-0.7210	6.45	93.55
到最近海岸线距离 x_3	-0.4450	0.3903	1.1973	0.3599	86.07	13.93
到最近铁路距离 x_4	-0.1982	0.0614	0.7320	0.0997	73.53	26.47
到最近公路距离 x_5	-0.8908	-0.1638	0.2650	-0.1822	32.44	67.56
到最近居民点距离 x_6	-0.3142	-0.0012	0.7227	0.0097	50.22	49.78
到最近水系距离 x_7	-2.2952	-0.6931	0.2362	-0.8373	1.09	98.91
多年平均降水 x_8	0.1181	0.3724	0.8624	0.4294	100.00	0
多年平均气温 x_9	-0.6755	0.3995	1.8584	0.4576	79.26	20.74
人口密度 x_{10}	-7.4913	-0.1806	4.6912	0.0028	26.51	73.49

在每个 $5 \text{ km} \times 5 \text{ km}$ 网格上, 耕地格局对各影响因子的敏感程度可以从该网格的 GWR 模型回归系数来反映。分别针对每个 $5 \text{ km} \times 5 \text{ km}$ 网格计算其回归系数中最大正值和最小负值, 对应的影响因子作为耕地格局最敏感的正向和负向影响因子, 其格局如图 4。表 5 则统计了研究区中各影响因子作为敏感的正向、负向影响因子的栅格数及百分比。

最敏感的正向影响因子主要是多年平均气温 x_9 和多年平均降水 x_8 , 两者所占面积之和超过 60%, 耕地趋向于在气温和降水高的区域分布。此外, 高程 x_1 在沧州市、滨州市、东营市北部、天津市南部为最敏感的正向影响因子, 约占区域总面积的 14.61%。这个区域高程在 $-3 \sim 20 \text{ m}$ 之间, 地势平坦, 经济发达, 耕地更倾向于分布在地势高区域。

较敏感的负向影响因子由高程 x_1 、坡度 x_2 、到最近公路距离 x_5 、到最近水系距离 x_7 和人口密度 x_{10} 组成。其中, 高程 x_1 、坡度 x_2 和到最近水系距离 x_7 , 三者所占比例之和接近 90%。以高程 x_1 为最敏感的负向影响因子区域占研究区总面积的 42.59%, 主要分布在锦州市、唐山市南部、潍坊东部以及青岛、烟台、威海和日照。该区域以低山丘陵为主, 高程值在 $-3 \sim 800 \text{ m}$ 之间, 平均高程为 70m, 耕地表现出随高程的升高而降低。以到最近水系距离 x_7 为最敏感的负向影响因子区域占研究区总面积的 32.06%, 分布范围较广, 除青岛、烟台、威海和日照, 在其余地市均有分布。研究区约 13.85% 的区域以坡度 x_2 为最敏感负向区域, 集中在天津、滨州和东营。

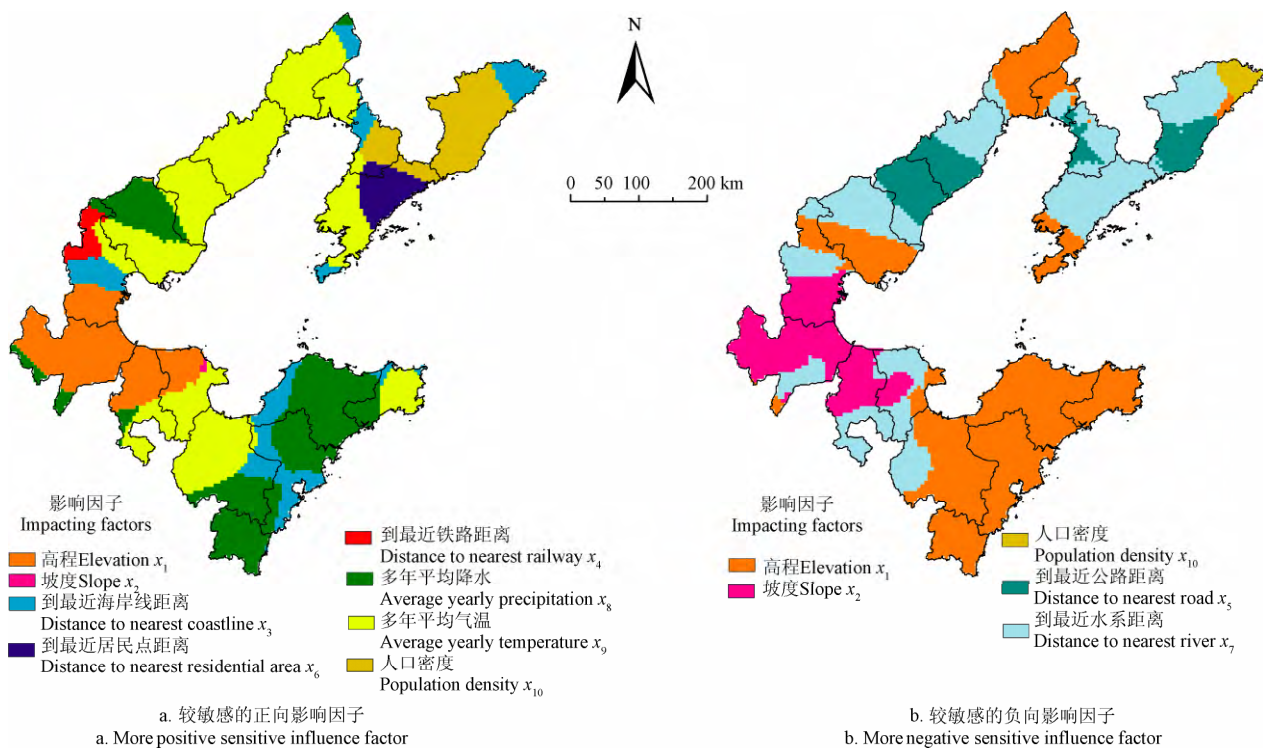


图 4 环渤海沿海区域耕地较敏感正向因子和较敏感负向因子分布

Fig.4 Distribution of the more positive or negative sensitive influence factors of cultivated land in coastal zone of Bohai Rim region

表 5 环渤海沿海区域耕地格局敏感影响因子统计

Table 5 Statistics of more sensitive influence factors of distribution of cultivated land in coastal zone of Bohai Rim region

影响因子 Influence factors	正向影响网格数 Number of grids of positive influence	网格百分比 Percentage of grids/%	负向影响网格数 Number of grids of negative influence	网格百分比 Percentage of grids/%
高程 x_1	1 013	14.61	2 953	42.59
坡度 x_2	6	0.09	960	13.85
到最近海岸线距离 x_3	722	10.41	0	0
到最近铁路距离 x_4	101	1.46	0	0
到最近公路距离 x_5	0	0	699	10.08
到最近居民点距离 x_6	227	3.27	0	0
到最近水系距离 x_7	0	0	2 223	32.06
多年平均降水 x_8	2670	22.30	0	0
多年平均气温 x_9	648	38.51	0	0
人口密度 x_{10}	1 546	9.35	98	1.41
合计	6 933	100.00	6 933	100.00

4 结论与讨论

从全局和局部的两个尺度上探讨耕地格局与影响因子的回归关系及空间变异性。研究结果表明：

1) SEM、SLM 和 GWR 的拟合效果均显著优于 OLS 模型。其中，SEM 在该研究中优于 SLM。虽 GWR 拟合度弱于 SEM，但可以获得不同地理位置影响因子对耕地格局的影响程度大小。

2) 从区域整体来讲，根据 SEM 模型的拟合结果，距水系距离和人口密度与耕地格局的关系不具有统计学意义，高程、坡度、到最近公路距离与耕地格局呈负相关影响，而其余因素与耕地格局呈正相关。

3) 从局域来看，除了多年平均降水对各网格单元内耕地面积均呈正相关影响，其余因子随网格单元变化正负向影响都存在，但正负所占比例有差异。多年平均气温和多年平均降水是主要的最敏感的正向影响因子，高程、距水系距离和坡度为主要的、最敏感的负向影响因子。

环渤海省市沿海区域空间范围广阔（南北跨纬度近 7° ，东西跨经度 10° 以上），且跨 16 地级市和 1 个直辖市，区域内部地形、交通网路、河流、气候以及人口等差异较大。空间自回归模型在经典回归分析的基础上引入了处理空间依赖性的技术，但对回归系数是全局或平均意义上的估计值，不随空

间位置变化而变化,本质上属于全局模型。地理加权回归模型是局部模型,对每个数据点都有一组局部的参数估计,回归系数随空间位置而变化。在研究过程中可采用空间自回归模型作为基础,对整个区域的信息有整体、直观地把握,明确了各影响因子的整体影响力。再结合地理加权回归模型分析各影响因子的空间变异性。两者相结合可能提供更多的有用信息,为研究区土地利用规划和宏观调控提供依据。

[参 考 文 献]

- [1] 李永乐, 吴群. 中国经济增长与耕地资源数量变化阶段性特征研究——协整分析及 Granger 因果检验[J]. 长江流域资源与环境, 2011, 20(1): 33 - 39.
Li Yongle, Wu Qun. Economic growth and quantity change of cultivated land: Co-integration analysis and Granger causality test[J]. Resources and Environment in the Yangtze Basin, 2011, 20(1): 33 - 39. (in Chinese with English abstract)
- [2] Jiang Li, Deng Xiangzheng, Seto K C. The impact of urban expansion on agricultural land use intensity in China[J]. Land Use Policy, 2013, 35: 33 - 39.
- [3] Zhong Taiyang, Huang Xianjin, Zhang Xiuying, et al. Temporal and spatial variability of agricultural land loss in relation to policy and accessibility in a low hilly region of southeast China[J]. Land Use Policy, 2011, 28(4): 762 - 769.
- [4] 肖思思, 吴春笃, 储金字. 近 15 年中国耕地驱动因素的空间差异分析[J]. 中国土地科学, 2013, 27(7): 55 - 61.
Xiao Sisi, Wu Chundu, Chu Jinyu. Spatial variation of driving factors on cultivated land change in China for the past 15 years[J]. China Land Sciences, 2013, 27(7): 55 - 61. (in Chinese with English abstract)
- [5] 易玲, 张增祥, 汪潇, 等. 近 30 年中国主要耕地后备资源的时空变化[J]. 农业工程学报, 2013, 29(6): 1 - 12.
Yi Ling, Zhang Zengxiang, Wang Xiao, et al. Spatial-temporal change of major reserve resources of cultivated land in China in recent 30 years[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2013, 29(6): 1 - 12. (in Chinese with English abstract)
- [6] 张国平, 刘纪远, 张增祥. 近 10 年来中国耕地资源的时空变化分析[J]. 地理学报, 2003, 58(3): 323 - 332.
Zhang Guoping, Liu Jiuyan, Zhang Zengxiang. Spatial-temporal changes of cropland in china for the past 10 years based on remote sensing[J]. Acta Geographica Sinica, 2003, 58(3): 323 - 332. (in Chinese with English abstract)
- [7] 封志明, 刘宝勤, 杨艳昭. 中国耕地资源数量变化的趋势分析与数据重建: 1949 - 2003[J]. 自然资源学报, 2005, 20(1): 35 - 43.
Feng Zhiming, Liu Baoqin, Yang Yanzhao. A study of the changing trend of Chinese cultivated land amount and data reconstructing: 1949-2003[J]. Journal of Natural Resources, 2005, 20(1): 35 - 43. (in Chinese with English abstract)
- [8] 邵晓梅, 谢俊奇. 中国耕地资源区域变化态势分析[J]. 资源科学, 2007, 29(1): 36 - 42.
Shao Xiaomei, Xie Junqi. Analyzing regional changes of the cultivated land in China[J]. Resources Science, 2007, 29(1): 36 - 42. (in Chinese with English abstract)
- [9] 田光进, 庄大方, 刘明亮. 近 10 年中国耕地资源时空变化特征[J]. 地球科学进展, 2003, 18(1): 30 - 36.
Tian Guangjin, Zhuang Dafang, Liu Mingliang. The spatial-temporal dynamic change of cultivated land in China in 1990s[J]. Advance in Earth Sciences, 2003, 18(1): 30 - 36. (in Chinese with English abstract)
- [10] Jin Jianjun, Jiang Chong, Li Lun. The economic valuation of cultivated land protection: A contingent valuation study in Wenling City, China[J]. Landscape and Urban Planning, 2013, 119: 158 - 164.
- [11] 戴声佩, 张勃. 基于 CLUE-S 模型的黑河中游土地利用情景模拟研究——以张掖市甘州区为例[J]. 自然资源学报, 2013, 28(2): 336 - 348.
Dai Shengpei, Zhang Bo. Land use change scenarios simulation in the middle reaches of the Heihe river basin based on CLUE-S model: A case of Ganzhou district of Zhangye city[J]. Journal of Natural Resources, 2013, 28(2): 336 - 348. (in Chinese with English abstract)
- [12] 吴莉, 侯西勇, 徐新良, 等. 山东沿海地区土地利用和景观格局变化[J]. 农业工程学报, 2013, 29(5): 207 - 216.
Wu Li, Hou Xiyong, Xu Xinliang, et al. Land use and landscape pattern changes in coastal areas of Shandong province, China[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2013, 29(5): 207 - 216. (in Chinese with English abstract)
- [13] Cliff A D, Ord J K. Spatial Processes: Models and Applications[M]. London: Pion, 1981.
- [14] Anselin L. Spatial Econometrics: Methods and Models[M]. Dordrecht: Kluwer Academic Publishers, 1988.
- [15] Overmars K P, De Koning G H J, Veldkamp A. Spatial autocorrelation in multi-scale land use models[J]. Ecological Modelling, 2003, 164(2/3): 257 - 270.
- [16] 谢花林, 刘黎明, 李波, 等. 土地利用变化的多尺度空间自相关分析——以内蒙古翁牛特旗为例[J]. 地理学报, 2006, 61(4): 389 - 400.
Xie Hualin, Liu Liming, Li Bo, et al. Spatial autocorrelation analysis of multi-scale land-use changes: A case study in Ongniud Banner, Inner Mongolia[J]. Acta Geographica Sinica, 2006, 61(4): 389 - 400. (in Chinese with English abstract)
- [17] 邱炳文, 王钦敏, 陈崇成, 等. 福建省土地利用多尺度空间自相关分析[J]. 自然资源学报, 2007, 22(2): 311 - 320.
Qiu Bingwen, Wang Qinmin, Chen Chongcheng, et al. Spatial autocorrelation analysis of multi-scale land use in Fujian province[J]. Journal of Natural Resources, 2007,

- 22(2): 311 - 320. (in Chinese with English abstract)
- [18] 刘敏, 赵翠薇, 施明辉. 贵州山区土地利用变化多尺度空间自相关分析[J]. 农业工程学报, 2012, 28(20): 239 - 246.
- Liu Min, Zhao Cuiwei, Shi Minghui. Spatial autocorrelation analysis of multi-scale land use change at mountainous areas in Guizhou province[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2012, 28(20): 239 - 246. (in Chinese with English abstract)
- [19] Fortheringham A S, Brunsdon C, Charlton M. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships[M]. New York: Wiley, 2002.
- [20] 董冠鹏, 张文忠, 武文杰, 等. 北京城市住宅土地市场空间异质性模拟与预测[J]. 地理学报, 2011, 66(6): 750 - 760.
- Dong Guanpeng, Zhang Wenzhong, Wu Wenjie, et al. Spatial heterogeneity in determinants of residential land price: Simulation and prediction[J]. Acta Geographica Sinica, 2011, 66(6): 750 - 760. (in Chinese with English abstract)
- [21] 张金萍, 秦耀辰. 县域粮食单产及其生产投入因素的空间异质性——以河南省为例[J]. 自然资源学报, 2011, 26(3): 373 - 381.
- Zhang Jinping, Qin Yaochen. Spatial heterogeneity of grain yield per hectare and factors of production inputs in counties: A case study of Henan province[J]. Journal of Natural Resources, 2011, 26(3): 373 - 381. (in Chinese with English abstract)
- [22] 杨扬, 杨建宇, 李绍明, 等. 玉米倒伏胁迫影响因子的空间回归分析[J]. 农业工程学报, 2011, 27(6): 244 - 249.
- Yang Yang, Yang Jianyu, Li Shaoming, et al. Spatial regression analysis on influence factors of maize lodging stress[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2011, 27(6): 244 - 249. (in Chinese with English abstract)
- [23] 刘琼峰, 李明德, 段建南, 等. 农田土壤铅、镉含量影响因素地理加权回归模型分析[J]. 农业工程学报, 2013, 29(3): 225 - 234.
- Liu Qiongfeng, Li Mingde, Duan Jiannan, et al. Analysis on influence factors of soil Pb and Cd in agricultural soil of Changsha suburb based on geographically weighted regression model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2013, 29(3): 225 - 234. (in Chinese with English abstract)
- [24] 宁秀红, 郭龙, 张海涛. 基于空间自回归和地理加权回归模型的不同尺度下土地利用程度研究[J]. 华中农业大学学报, 2013, 32(4): 48 - 54.
- Ning Xiuhong, Guo Long, Zhang Haitao. Comprehensive degree of land-use at different scales based on spatial autocorrelation regression and geographically weighted regression models[J]. Journal of Huazhong Agricultural University, 2013, 32(4): 48 - 54. (in Chinese with English abstract)
- [25] 邵一希, 李满春, 陈振杰, 等. 地理加权回归在区域土地利用格局模拟中的应用——以常州市孟河镇为例[J]. 地理科学, 2010, 30(1): 92 - 97.
- Shao Yixi, Li Manchun, Chen Zhenjie, et al. Simulation on regional spatial land use patterns using Geographically Weighted Regression: A case study of Menghe town, Changzhou[J]. Scientia Geographica Sinica, 2010, 30(1): 92 - 97. (in Chinese with English abstract)
- [26] 马宗文, 许学工, 卢亚灵. 环渤海地区 NDVI 拟合方法比较及其影响因素[J]. 生态学杂志, 2011, 30(7): 1558 - 1564.
- Ma Zongwen, Xu Xuegong, Lu Yaling. Comparison of NDVI simulation models for Bohai Rim region and the factors affecting NDVI[J]. Chinese Journal of Ecology, 2011, 30(7): 1558 - 1564. (in Chinese with English abstract)
- [27] 郭丽英, 王道龙, 王介勇. 中国沿海地区耕地变化及其成因分析[J]. 中国农业资源与区划, 2012, 33(1): 6 - 10.
- Guo Liyong, Wang Daolong, Wang Jieyong. Analysis of the arable land change and its causes in the eastern coastal China over the last decades[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2012, 33(1): 6 - 10. (in Chinese with English abstract)
- [28] 郭丽英, 李刚. 环渤海地区城镇化进程中耕地面积及其重心转移格局分析[J]. 中国农业资源与区划, 2013, 34(4): 32 - 36.
- Guo Liying, Li Gang. Analysis of spatial pattern on farmland area change and the gravity center transfer under urbanization of Bohai Rim[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2013, 34(4): 32 - 36. (in Chinese with English abstract)
- [29] 霍霄妮, 李红, 孙丹峰, 等. 北京耕地土壤重金属空间自回归模型及影响因素[J]. 农业工程学报, 2010, 26(5): 78 - 82.
- Huo Xiaoni, Li Hong, Sun Danfeng, et al. Spatial autogression model for heavy metals in cultivated soils of Beijing[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2010, 26(5): 78 - 82. (in Chinese with English abstract)
- [30] Chen Xiwei, Dai Erfu. Comparison of spatial autoregressive models on multi-scale land use[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2011, 27(6): 324 - 331.
- 陈曦炜, 戴尔阜. 土地利用的多尺度空间自相关模型比较[J]. 农业工程学报, 2011, 27(6): 324 - 331. (in English with Chinese abstract)
- [31] 王库. 基于地理权重回归模型的土壤有机质空间预测[J]. 土壤通报, 2013, 44(1): 21 - 28.
- Wang Ku. Spatial estimation of soil organic matter by using Geographically Weighted Regression Model[J]. Chinese Journal of Soil Science, 2013, 44(1): 21 - 28. (in Chinese with English abstract)

Analysis of spatial pattern of farmland and its impacting factors in coastal zone of Circum Bohai

Wu Li^{1,2}, Hou Xiyong¹, Xu Xinliang³

(1. Yantai Institute of Coastal Zone Research, Chinese Academy of Science, Yantai 264003, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Science, Beijing 100101, China)

Abstract: In this paper, coastal zone of Circum Bohai Sea Region which covers an area of approximately 170,000 km² was selected as the study area. The spatial distribution characteristics of farmland of this study area were analyzed and the relationship between farmland distribution and natural, social or economic impacting factors was explored. Based on Landsat TM images acquired in 2009/2010, farmland distribution map was created through visual interpretation with auxiliary data in ArcGIS 9.3. Then farmland distribution map was overlaid with a lattice map to statistic area of farmland in each 5 km × 5 km lattice. Impacting factors of farmland consisted of elevation, slope, distance to nearest coastline, distance to nearest railway, distance to nearest road, distance to nearest residential area, distance to nearest river, average yearly precipitation, average yearly temperature and population density, which were compiled into raster format data with a spatial resolution of 5 km × 5 km and normalized between 0 and 1 in ArcGIS 9.3. As conventional statistical methods assumed that the data to be analyzed was statistically independent, it was inappropriate to use traditional statistical method to analyze spatial land use data which had a tendency to be dependent. In this study, ordinary least square linear regression model (OLS), spatial error model (SEM), spatial lag model (SLM) and geographically weighted regression model (GWR) were established from global and local perspectives. Several evaluation indexes were selected to assess the performance of those models. The results showed that: 1) Farmland was the main land use type, which occupied 53% of the whole study area. Positive spatial autocorrelation that decreased gradually with distance was detected in both farmland distribution and impacting factors; 2) Spatial autoregressive models and GWR had a better goodness-of-fit than conventional linear regression model. As to spatial autoregressive models, SEM performed better than SLM in this study, as was indicated by higher pseudo R^2 value and maximum likelihood logarithm (LIK) value, and lower Akaike information criterion (AIC) value, Schwartz criterion (SC) value and residuals for the former model; 3) GWR could be used to explore spatial variation in the relations between cultivated land distribution and different impacts factors, providing more detailed information, while SEM could only explore the relations from a global view; 4) The SEM showed a positive correlation between farmland and elevation, slope, distance to the nearest roads, as well as a negative correlation between farmland and distance to nearest shoreline, distance to nearest railroad, distance to nearest settlements, average yearly temperature, average yearly precipitation from a global perspective; and 5) The GWR revealed both positive and negative correlations between farmland and impacting factors (except for average yearly precipitation). Among the most sensitive factors affecting farmland distribution, average yearly temperature and average yearly precipitation were the main positive factors, while elevation, slope and distance to nearest residential area were the main negative factors.

Key words: land use; regression analysis; rural areas; coastal zone of Circum Bohai Sea Region; farmland; linear regression model; spatial autoregressive model; geographically weighted regression model