

基于栅格数据的土地利用格局解释模型及应用

邓祥征¹ 林英志¹ 战金艳² 何书金¹

(1. 中国科学院地理科学与资源研究所, 北京 100101;

2. 北京师范大学环境学院水环境模拟国家重点实验室, 北京 100875)

摘 要:在一定空间分辨率的栅格上用二值型变量表征土地利用格局,是一种常见的土地利用格局表征形式。解释基于这一表征形式的土地利用格局与自然环境条件、社会经济因素之间的关系多需引入非线性模型。本文基于非线性统计方法构建了基于栅格数据的土地利用格局解释模型,不仅刻画了土地利用格局与影响因子之间的非线性统计关系,解决了分类变量引入统计模型的难题,还消除了土地利用格局解释模型中概率估计值常常超过概率值域的问题。太仆寺旗的应用研究表明,栅格尺度土地利用格局解释模型能准确表达土地利用分布格局与自然环境条件、社会经济因素间的定量关系,所得结论为太仆寺旗土地利用规划、土地资源优化利用提供了决策参考信息。

关 键 词:土地利用;土地利用格局;RBEMLUP;栅格;太仆寺旗

土地利用格局是自然环境条件和社会经济因素相互作用、共同约束的结果。土地利用格局影响因素分析旨在判定影响因子与土地利用格局间的统计关系^[1-3]。基于栅格数据开展统计分析是解释自然环境条件和社会经济因素影响土地利用格局的主要途径之一^[4-8]。采用线性回归模型或相关分析有助于从复杂的土地利用系统中提取出土地利用格局的影响因子^[9-10],但在处理一定空间分辨率栅格上用二值型变量表征(基于土地利用栅格数据)的土地利用格局问题时,则显得颇为乏力^[11-15]。一方面,土地利用格局与影响因子之间的关系错综复杂,并非简单线性相关关系;另一方面,土地利用栅格数据往往含有非连续型变量,不满足线性回归对模型变量类型的限制性要求^[16-17]。

本文基于非线性统计学方法,构建了适用于土地利用格局影响因素分析的栅格尺度土地利用格局解释模型(RBEMLUP),并选择太仆寺旗作为案例区,阐明了当地土地利用格局与自然环境条件、社会经济因素间的定量关系,揭示了当地土地利用格局的影响因素与变化规律,所得结论具有重要决策参考价值。

收稿日期:2009-07; 修订日期:2009-09.

基金项目:国家自然科学基金项目(70873118、70821140353)、国家科技支撑项目(2006BAC08B03、2006BAC08B06、2008BAK47B02)、科技部国际科技合作项目(2006DFB919201)。

作者简介:邓祥征(1971-),男,博士,副研究员,主要从事区域环境变化、土地系统变化及效应研究。

E-mail: dengxz.ccap@igsnrr.ac.cn

1 RBEMLUP

1.1 模型假设

RBEMLUP 适用于解释在一定空间分辨率栅格上用二值型变量表征的土地利用分布格局。该模型假定编号 i ($i=1, 2, \dots, n$) 的栅格(以下简称栅格 i) 上的第 k ($k=1, 2, \dots, M$) 种土地利用面积为 $P^k (\geq 0)$, 其与二值型变量 z^k 之间的对应关系定义为:

$$\begin{cases} z^k = 1 & \text{当 } P^k > 0 \text{ 时} \\ z^k = 0 & \text{当 } P^k = 0 \text{ 时} \end{cases} \quad (1)$$

式中: z^k 为土地利用类型。模型自变量 X 是一系列自然环境和社会经济因素组成影响因子协变向量。

$$X = (x_1, x_2, \dots, x_l, \dots, x_L)^T \quad (2)$$

1.2 模型推导

栅格出现第 k 种土地利用类型的概率为 $p^k = P(z^k = 1 | X)$, RBEMLUP 假定这个条件概率可以用 logistic 函数的形式表达,即:

$$p^k = \frac{1}{1 + \exp[-(a_0^k + X^T a^k)]} \quad (3)$$

式中: p^k 表示栅格出现第 k 种土地利用的概率; $a^k =$

$(a_1^k, a_2^k, \dots, a_l^k, \dots, a_L^k)^T$ 为各自然环境条件和社会经济因素对应系数组成的向量。

由 p^k 可以求得栅格“出现第 k 种土地利用”的发生比:

$$\frac{p^k}{1-p^k} = \exp(a_0^k + X^T a^k) \quad (4)$$

这个比值通常也被称为似然比(odds)^[18]。对发生比进行自然对数变换,可以得到:

$$\text{logit}(y^k) = \ln(p^k / (1-p^k)) = a_0^k + X^T a^k \quad (5)$$

此即以 logit 形式表达的 RBEMLUP。此模型具有 S 型分布(图 1)。

RBEMLUP 消除了土地利用格局影响因子分析研究中概率估计值超过概率值域的问题。随着 logit(y^k)取值的增大,栅格“出现第 k 种土地利用”的发生比 $p^k / (1-p^k)$ 和发生概率 p^k 不断增加。一个自变量如果对 logit(y^k)有正向作用,则说明该变量对第 k 种土地利用的出现有促进作用;反之,有抑制作用。RBEMLUP 的估计结果可以用 ROC 方法检验^[19]。

1.3 模型估算

本文介绍如何利用最大似然估计法来估算 RBEMLUP 的系数。假设有由 N 个栅格构成的总体,分别为 $Z_1^k, Z_2^k, \dots, Z_N^k$ 。从中随机抽取 n 个作为样本,观测值分别为 $z_1^k, z_2^k, \dots, z_N^k$ 。而

$$\begin{cases} P(z_i^k = 1 | X_i) = p_i^k \\ P(z_i^k = 0 | X_i) = 1 - p_i^k \end{cases} \quad (6)$$

因此,得到观测值 z_i^k 的概率为:

$$P(z_i^k) = (p_i^k)^{z_i^k} (1-p_i^k)^{1-z_i^k} \quad (7)$$

由于各个观测之间相互独立,因此可得这 n 个观测值的似然函数:

$$L(\theta) = \prod_i^n (p_i^k)^{z_i^k} (1-p_i^k)^{1-z_i^k} \quad (8)$$

根据最大似然估计的基本原理,求解的参数估计值要使得似然函数的值最大,即取得给定的样本观测值的可能性最大的参数估计。然而要使 $L(\theta)$ 最大化,其运算过程比较复杂,但可以证明不等式

$$d \ln[L(\theta)] / dL(\theta) = \frac{1}{L(\theta)} > 0 \quad (9)$$

恒成立,即 $\ln[L(\theta)]$ 为 $L(\theta)$ 的单调递增函数,当 $L(\theta)$ 取得最大值时 $\ln[L(\theta)]$ 达到最大,反之亦然。所以,通常做法都是在最大化的基础上计算参数的估计值。

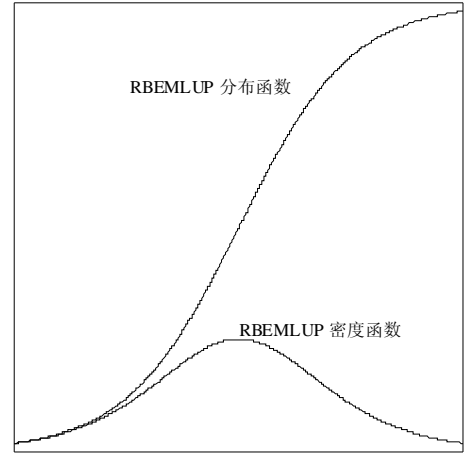


图 1 RBEMLUP 分布

Fig.1 Distribution of RBEMLUP

$$\begin{aligned} \ln[L(\theta)] &= \ln \left[\prod_i^n (p_i^k)^{z_i^k} (1-p_i^k)^{1-z_i^k} \right] \\ &= \sum_{i=1}^n [z_i^k \ln(p_i^k) + (1-z_i^k) \ln(1-p_i^k)] \\ &= \sum_{i=1}^n [z_i^k (a_0^k + a^k X_i) - \ln(1 + e^{a_0^k + a^k X_i})] \end{aligned} \quad (10)$$

(10)式称为对数似然函数,将它分别对 a_0^k 和 a^k 求偏导,并令得到的偏导数等于零:

$$\begin{aligned} \frac{\partial \ln[L(\theta)]}{\partial a_0^k} &= \sum_{i=1}^n \left(z_i^k - \frac{e^{\frac{a_0^k + a^k X_i}{a_0^k + a^k X_i}}}{1 + e^{\frac{a_0^k + a^k X_i}{a_0^k + a^k X_i}}} \right) = 0 \\ \frac{\partial \ln[L(\theta)]}{\partial a_1^k} &= \sum_{i=1}^n \left(z_i^k - \frac{e^{\frac{a_0^k + a^k X_i}{a_0^k + a^k X_i}}}{1 + e^{\frac{a_0^k + a^k X_i}{a_0^k + a^k X_i}}} \right) x_{i1} = 0 \\ \frac{\partial \ln[L(\theta)]}{\partial a_L^k} &= \sum_{i=1}^n \left(z_i^k - \frac{e^{\frac{a_0^k + a^k X_i}{a_0^k + a^k X_i}}}{1 + e^{\frac{a_0^k + a^k X_i}{a_0^k + a^k X_i}}} \right) x_{iL} = 0 \end{aligned} \quad (11)$$

显然,此方程组的求解仍需借助计算机迭代运算实现。解得的 $\hat{a}_0^k, \hat{a}_1^k, \dots, \hat{a}_L^k$, 分别为 $a_0^k, a_1^k, \dots, a_L^k$ 的估计值。设 $\hat{a} = (\hat{a}_1^k, \dots, \hat{a}_L^k)$, 则称 \hat{a} 为 a 的估计, RBEMLUP 亦可写为:

$$\text{logit}(y^k) = \ln(p^k / (1-p^k)) = \hat{a}_0^k + \hat{a}^k X_i \quad (12)$$

虽然 RBEMLUP 的估算过程与一般线性回归模型相类似,模型本身也具有很多线性回归模型的性质,但其与一般线性回归模型在土地利用格局解释的原理方面存在显著差异。首先,线性回归模型通常都会假设因变量的观测值服从正态分布,而在 RBEMLUP 中,因变量的观测值服从二项分布;其次,栅格尺度土地利用格局解释模型中不存在类似

线性回归模型的残差项。

总之, RBEMLUP 在刻画影响因子与土地利用格局之间的关系时, 较线性回归模型具有诸多优点。第一, 相对于线性概率模型, RBEMLUP 能够在引入离散变量的同时保证参数估计是无偏的; 第二, RBEMLUP 具有良好的函数形式, 在一定程度上克服了线性回归模型难以精确拟合非线性关系的缺点。

2 太仆寺旗案例研究

本文以太仆寺旗为案例区, 详细介绍利用 RBEMLUP 开展土地利用格局影响因子分析的过程。太仆寺旗土地利用数据来自中国科学院资源环境科学数据中心 2000 年土地利用数据集。土地利用格局影响因子按照来源可以划分为自然环境条件和社会经济因素。其中, 自然环境条件自内蒙古自治区 1:25 万地形图与 DEM 数据提取; 社会经济数据源于内蒙古自治区统计调查资料。

2.1 土地利用数据

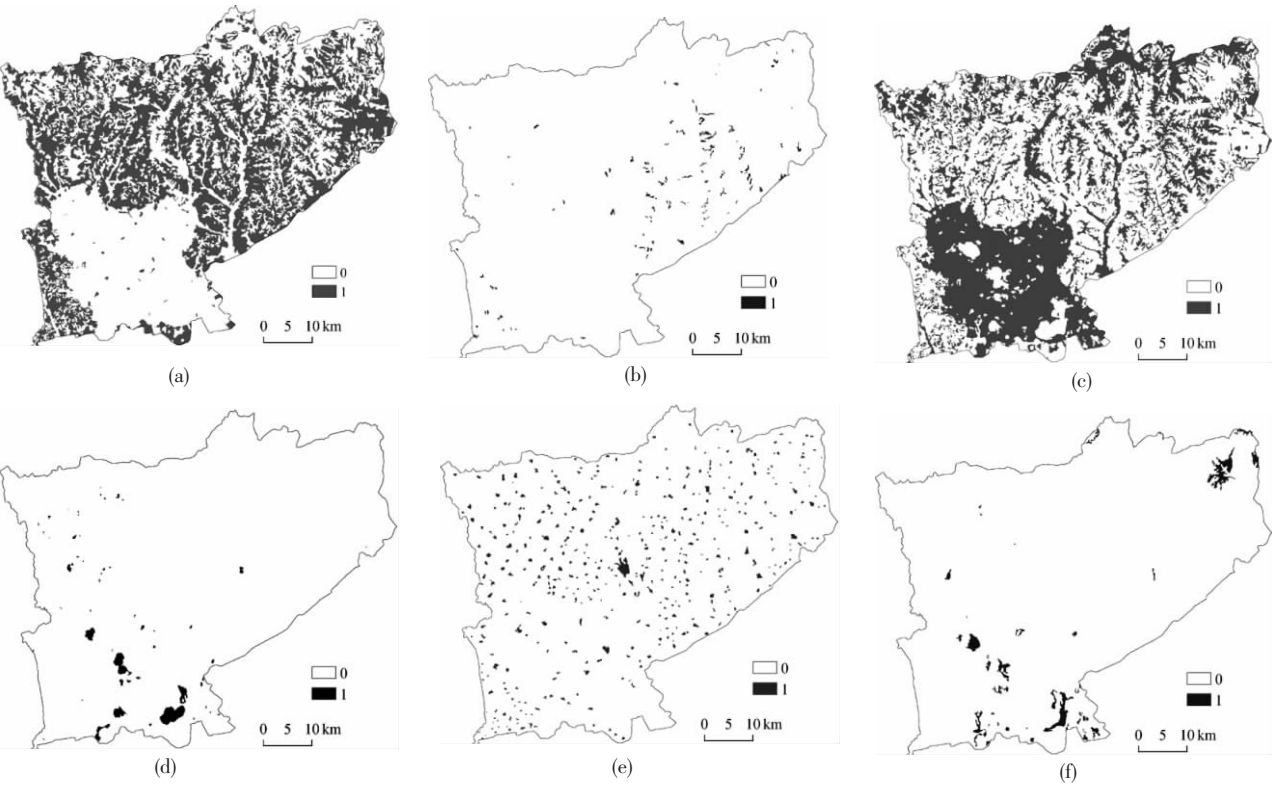
太仆寺旗基于栅格尺度土地利用格局解释模

型自变量的构造需要将 2000 年土地利用数据转换成二值型表示的形式。耕地、林地、草地、水域、建设用地与未利用地六种土地利用类型的 250 m×250 m 栅格上的二值型变量精确地记录了每个栅格上是否存在该土地利用类型(图 2)。

2.2 自然环境条件数据

自然环境条件对区域土地利用格局的影响主要体现为长期控制作用。本文将一系列的自然环境条件与太仆寺旗土地利用格局进行空间叠加, 依据不同海拔高度、坡度坡向、地形地貌、气候条件、土壤质地条件下的土地利用类型空间分异规律, 决定选取地貌类型、高程、坡度、坡向、土壤侵蚀程度、大于等于 10℃积温、年均气温、年均降水量、日照时数和土壤钾含量作为解释太仆寺旗土地利用格局的自然环境条件变量。

用以解释太仆寺旗土地利用格局的自然环境条件变量均被离散化或重采样至 250 m×250 m 栅格尺度。本文通过设计地貌类型重映射表, 实现了地貌类型的重分类, 将其转化为具有空间信息的二值型变量^[16]。利用太仆寺旗的 1:25 万等高线图, 本文首先生成了 50 m×50 m 栅格尺度的高程数据, 进



注: (a)、(b)、(c)、(d)、(e)和(f)分别代表耕地、林地、草地、水域、建设用地与未利用地。

图 2 太仆寺旗 2000 年土地利用格局的二值型表达

Fig.2 Land use pattern identified by binary variables of Taips League in 2000

而生成太仆寺旗坡度与坡向栅格数据,并将三者重采样至 250 m×250 m 栅格尺度。坡向数据进而经由坡向重分类表将连续型变量转化成具有实际意义的序列变量;坡度数据为连续型变量,直接参与土地利用格局影响因子分析。本文将遥感解译获得的 1:10 万土壤侵蚀数据按照土壤侵蚀程度进行了重分类,实现了土壤侵蚀程度的序列化表达。运用距离权重法,基于全国 669 个气象台站的日均气温与降水量观测数据,本文对全国日均气温与降水量数据进行了空间插值^[16]。在此基础上,利用太仆寺旗的行政边界,切出当地日均气温与降水量信息,进而计算出每个 250 m×250 m 栅格上的大于等于 10℃积温、年均气温与年均降水量。基于全国气象台站的实测数据和全国土壤普查数据,采用类似处理方法,本文生成了太仆寺旗 250 m×250 m 栅格尺度日照时数和土壤钾含量数据。

2.3 社会经济数据

社会经济因素在较短时间上内决定着区域土地利用格局的变化。根据太仆寺旗技术相对落后、交通不发达、生活封闭、灾害频繁、人类活动影响强度大的特点,本文主要选择能够反映区位条件的到省会城市距离、到港口城市距离、到水系的距离和到主要公路距离,反映人们生活条件的第二产业比例,反映灾害风险的灾损率,以及反映人类活动强度的人口密度、国民生产总值和复种指数,作为解释太仆寺旗土地利用格局的社会经济因素。

本文分层计算了每个 250 m×250 m 栅格距省会城市和港口城市的平均距离。利用 2000 年覆盖太仆寺旗的 Landsat TM 几何精纠正影像,勾画出旗内主要水系与交通线,计算出每个 250 m×250 m 栅格距离水系与主要公路的距离。人口密度、国民生产总值、复种指数、第二产业比例、灾损率的离散化过程较为复杂,此处不再赘述,可参见[16]。

3 结果与分析

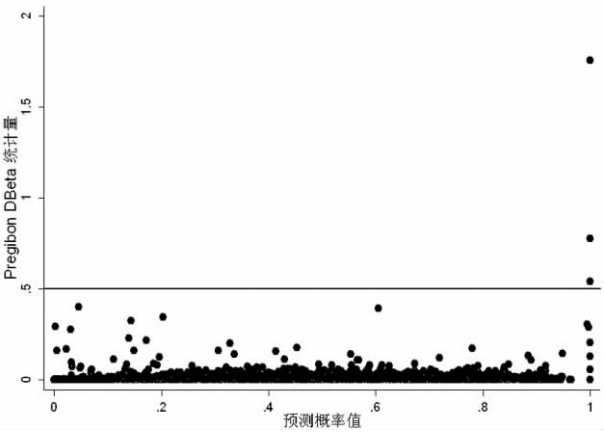
自变量之间的多重共线性是导致 RBEMLUP 估计结果偏误甚至符号错误的主要原因之一。因此,在对土地利用格局与影响因子展开分析之前,本文选用方差膨胀因子 VIF 测度了自变量的多重共线性情况(表 1)。测试结果显示,所有变量的方差膨胀因子都在 4.0 以下,说明包含太仆寺旗 RBEMLUP 内的自变量之间的共线性并不严重。

在此基础上,本文进一步判断了奇异值及其对估计结果的影响。诊断结果显示,用于耕地分布格局影响因子分析的样本中存在 3 个奇异值(图 3)。

表 1 太仆寺旗 RBEMLUP 主要解释变量的共线性诊断*

Tab.1 Diagnosis of the multicollinearity between independent variables included in the RBEMLUP of Taips League		
变量	VIF	1/VIF
到省会城市的距离	3.91	0.256
日照时数	3.87	0.258
到水系的距离	3.85	0.260
到港口城市的距离	3.84	0.260
年均降水量	3.80	0.263
地貌:平原	3.78	0.265
地貌:丘陵	3.78	0.265
年均气温	3.75	0.267
人口密度	3.37	0.297
国民生产总值	3.22	0.311
土壤侵蚀强度:中度	3.09	0.324
土壤侵蚀强度:轻度	3.04	0.329
≥10℃积温	2.95	0.339
坡向:西南坡	2.79	0.358
坡向:东坡	2.78	0.360
坡向:东南坡	2.61	0.383
高程	2.58	0.387
坡向:东北坡	2.52	0.396
坡向:西坡	2.46	0.407
坡向:南坡	2.38	0.420
坡向:西北坡	2.32	0.432
坡向:北坡	2.31	0.434
灾损率	1.75	0.572
到主要公路距离	1.56	0.641
土壤侵蚀强度:重度	1.48	0.677
地貌:台地	1.46	0.684
土壤钾含量	1.44	0.695
复种指数	1.30	0.767
坡度	1.20	0.832
第二产业比例	1.01	0.985
VIF 均值	2.67	

* 1/VIF 是容忍度,其值为 VIF 的倒数。



* 奇异值判定标准: db>0.500

图 3 太仆寺旗 2000 年耕地分布格局解释模型样本奇异值判定
Fig.3 Identification of the singular value of the samples used in RBEMLUP for the pattern of cultivated land distribution Taips League in 2000

本文用 2×2 维交叉表计算 RBEMLUP 的预测精度。耕地分布格局影响因子分析得到的预测精度交叉表表明,剔除奇异值能提高模型的预测精度。可以看出,剔除奇异值前,RBEMLUP 对耕地为 0 的栅格的预测正确率为 73.5%,对耕地为 1 的栅格的预测正确率为 77.9%,总体正确率 75.7%;剔除奇异值后,RBEMLUP 对耕地为 0 的栅格的预测正确率为 73.5%,对耕地为 1 的栅格的预测正确率为 78.2%,总体正确率为 75.9%(表 2)。

利用绘制的 ROC 曲线可以判定模型的拟合效果^[20]。ROC 曲线图中曲线下方面积值越大,表示模型的拟合效果越好。从图 4 可以看出,剔除奇异值之前,基于 RBEMLUP 开展的耕地分布格局影响因子分析的 ROC 曲线下方面积为 0.839,剔除奇异值之后的 ROC 曲线下方面积为 0.840。渐进显著性概率值为 0.000,拒绝零假设。因此,可以认为基于 RBEMLUP 估计出的耕地分布格局与影响因子之间的关系具备较高的拟合优度。

剔除奇异值前后的估计结果的差异表明奇异值的剔除提高了模型的精度。采用同样的方法,本文剔除了太仆寺旗其他五种栅格尺度土地利用格局解释模型数据中的奇异值并判定了模型的拟合优度。结果表明,RBEMLUP 对土地利用类型空间分布与影响因子关系的估计结果是可信的。

表 3 汇总了基于 RBEMLUP 估计的各影响因子与各种土地利用发生比间的定量关系。可以看出,台地相对于山地而言,对耕地和草地的影响较显著,其中耕地的发生比率为 0.726,草地的发生比率为 1.119。丘陵只对草地的影响较显著,其相对于山地的草地发生比率为 1.319。

平原之相对于山地,对建设用地影响最大,其次为林地,发生比率分别为 1.791 和 0.521。随着高程的增加,林地、草地和未利用地的发生比均增加,而耕地、水域和建设用地的发生比则减少。高程每增加 1 个单位,草地的发生比平均增加 0.28%,林地的发生比平均增加 0.56%,未利用地的发生比平均增加 1.58%。地形坡度较大的地区多为草地、未利用地覆盖,其中坡度每增加 1 个单位,草地的发生比平均增加 2.74%,未利用地的发生比平均增加 7.36%。

坡向对耕地和未利用地的影响均较显著,相对于平地而言,其它坡向上耕地发生比率均小于 1,未利用地发生比率均大于 1。由此可见,相对而言,平地更有利于耕地的扩张,而未利用地则易出现在坡地上。土壤侵蚀强度是阻碍耕地扩张的另一个重要因素。相对于微度侵蚀区来讲,轻度侵蚀区、中度侵蚀区和重度侵蚀区的耕地发生比率依次减少,分别为 0.967、0.941、0.904。

气候对土地利用格局产生重要影响。年均气温每增加 0.1℃,耕地发生比平均增加 19.87%,林地发生比平均增加 0.46%。降水量每增加 1mm,草地发生比平均增加 2.75%,未利用地发生比则降为原来的 77.10%。

表 2 奇异值处理前后太仆寺旗耕地分布格局预测精度的交叉检验

Tab.2 Comparison of the prediction accuracy of cultivated land pattern before and after singular value checking for the RBEMLUP of Taips league

		奇异值处理前的预测值		奇异值处理后的预测值	
		耕地		耕地	
观测值		0	1	判对率	
耕地	0	4917	1776	73.5	4937
	1	1526	5391	77.9	1506
总体判对率		75.7		75.9	

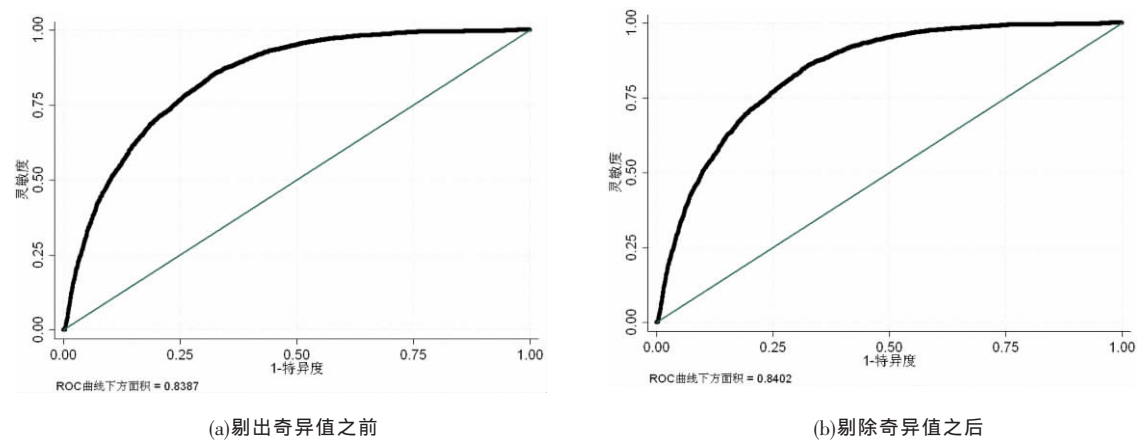


图 4 表征基于 RBEMLUP 模拟耕地分布格局拟合优度的 ROC 曲线

Fig.4 ROC curves identifying the goodness of fit for the RBEMLUP for the prediction of cultivated land pattern: (a) before singular value elimination, and (b) after singular value elimination

省会城市往往位于地势平坦地带,周围大部分为耕地和建设用地,所以一般距离省会城市越远,耕地、建设用地发生比越低,而未利用地、草地发生比越高。随着人口密度的增加,耕地、林地、水域发生比也相应增加,而草地和未利用的发生比则减少。自然灾害一般是针对农田而言,耕地发生比大的地区,灾损率的一般较大。国民生产总值对草地和建设用地发生比有正效应,国民生产总值密度每增加 1 万元/km²,建设用地发生比平均增加 2.56%。

4 结论与讨论

本文介绍了适用于区域土地利用格局影响因子分析的基于栅格数据的土地利用分布格局解释模型(RBEMLUP),并以太仆寺旗为案例区,详细说明了应用该模型解释区域土地利用分布格局的步骤。研究表明,RBEMLUP 能够准确刻画区域土地利用格局与自然环境条件和社会经济因素间的定量关系。太仆寺旗自然环境条件和社会经济因素影响土地利用格局的分析结论对当地土地利用规划、土地资源开发利用的调控及其优化管理具有重要的决策参考价值。

需要说明的是,基于非线性统计方法建立的 RBEMLUP 所描述的是影响因子与土地利用格局之间的统计关系,而并非确定的因果关系。不过,基于 Logistic 分布的 RBEMLUP 突破了变量类型对模型形式的限制,为基于栅格数据考察影响因子与土地利用格局的定量关系提供了方法,并在一定程度上揭示了在影响因子作用下各种土地利用在栅格尺度上的竞争与互补关系。

参考文献

[1] 李秀彬. 全球环境变化研究的核心领域: 土地利用/覆盖变化的国际研究动向. 地理学报, 1996, 51(6): 553-557.

[2] 陈百明, 刘新卫, 杨红. LUCC 研究的最新进展评述. 地理科学进展, 2003, 22(1): 22-29.

[3] Verburg P H, Schulp C J E, Witte N, et al. Downscaling of land use change scenarios to assess the dynamics of European landscapes. Agriculture, Ecosystems & Environment, 2006, 114(1): 39-56.

[4] 陈述彭, 陈星. 地球信息科学的理解与实践. 地球信息科学, 2004, 6(1): 4-10.

[5] 摆万奇, 赵士洞. 土地利用变化驱动力系统分析. 资源科学, 2001, 23(3): 39-41.

表 3 太仆寺旗 2000 年各种土地利用发生比与影响因子之间的定量关系

Tab.3 Relationship between the odds of land use types and driving factors estimated from the RBEMLUP of Taips League in 2000

	耕地	林地	草地	水域	建设用地	未利用地
地貌: 丘陵	—	—	1.319	—	—	—
地貌: 台地	0.726	—	1.119	—	—	—
地貌: 平原	—	0.521	—	—	1.791	—
高程	0.996	1.006	1.003	0.953	0.998	1.016
坡度	0.996	—	1.027	—	0.834	1.074
坡向: 北坡	—	—	—	—	—	0.286
坡向: 东北坡	0.835	—	0.750	—	—	—
坡向: 东坡	0.842	—	0.840	—	—	1.006
坡向: 东南坡	0.921	—	—	—	1.546	1.003
坡向: 南坡	—	—	1.159	0.403	—	1.001
坡向: 西南坡	0.925	—	—	—	—	1.005
坡向: 西坡	0.881	—	—	0.724	—	1.003
坡向: 西北坡	0.876	4.625	0.832	—	—	1.162
土壤侵蚀强度: 轻度	0.967	—	0.924	0.871	—	1.500
土壤侵蚀强度: 中度	0.941	—	0.853	0.928	—	1.581
土壤侵蚀强度: 重度	0.904	—	0.832	—	—	—
土壤钾含量	—	—	—	—	—	1.135
年均温度	1.199	1.228	1.322	—	0.426	1.180
?10℃积温	1.003	1.005	—	—	—	0.981
日照时数	0.901	0.869	1.061	—	0.954	1.074
年降水量	—	—	1.028	0.872	—	0.771
到省会城市的距离	0.944	0.932	1.019	—	0.973	1.138
到港口城市距离	0.983	0.854	—	—	—	0.939
到道路距离	1.018	1.116	0.968	1.069	—	1.117
到水域距离	1.056	1.108	—	0.834	—	0.890
人口密度	1.028	1.002	0.962	1.010	—	0.967
国民生产总值	0.936	—	1.086	0.952	1.026	0.744
第二产业比例	—	—	—	—	—	—
复种指数	—	—	—	0.627	—	—
灾损率	1.031	—	0.978	0.891	—	—

[6] 刘纪远, 庄大方, 张增祥, 等. 中国土地利用时空数据平台建设及其支持下的相关研究. 地球信息科学, 2002, 4(3): 3-7.

[7] 何书金, 李秀彬. 环渤海地区耕地变化及其动因分析. 自然资源学报, 2002, 17(3): 345-352.

[8] 战金艳, 邓祥征, 岳天祥, 等. 内蒙古农牧交错带土地利用变化及其环境效应. 资源科学, 2004, 26(5): 80-88.

[9] 庄大方, 刘纪远. 中国土地利用程度的区域分异模型研究. 自然资源学报, 1997, 12(2): 105-111.

[10] 刘彦随. 土地利用优化配置中系列模型的应用: 以乐清

- 市为例. 地理科学进展, 1999, 18(1): 26–31.
- [11] 王秀兰, 包玉海. 土地利用动态变化研究方法探讨. 地理科学进展, 1999, 18(1): 81–87.
- [12] 徐建华. 现代地理学中的数学方法. 北京: 高等教育出版社, 1994.
- [13] 于兴修, 杨桂山. 中国土地利用/覆被变化的研究现状与问题. 地理科学进展, 2002, 21(1): 51–57.
- [14] 史培军, 宫鹏, 李晓冰, 等. 土地利用/覆被变化研究的方法与实践. 北京: 科学出版社, 2000.
- [15] Turner II B L, Meyer W B, Skole D L. Global land–use/land–cover change: Towards an integrated program of study. *Ambio*, 1994, 23(1): 91–95.
- [16] 邓祥征. 土地系统动态模拟. 北京: 中国大地出版社, 2008.
- [17] 邓祥征. 土地用途转换分析. 北京: 中国大地出版社, 2008.
- [18] 王济川, 郭志刚. Logistic 回归模型: 方法与应用. 北京: 高等教育出版社, 2001.
- [19] Pontius R G, Schneider L C. Land–cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment*, 2001, 85(1–3): 239–248.
- [20] Hanley J A, McNeil B J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 1982, 143(1): 29–36.

Robust Estimation of the Raster–Based Explanatory Model for Land Use Pattern

DENG Xiangzheng¹, LIN Yingzhi¹, ZHAN Jinyan², HE Shujin¹

(1. Institute of Geographic Sciences and National Resources Research, CAS, Beijing 100101, China;

2. State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing 100875, China)

Abstract: The non–linear models with binary variables to identify the existence of some land use types at the pixel–specific dimension are widely used to analyze the relationship between land use pattern and its influential factors. In this paper, we develop and use a Raster–Based Explanatory Model for Land Use Pattern (RBEMLUP) to solve the problems we are always confronted with when we build and estimate the models to figure out the drive mechanism of land use pattern after we introduce the binary variables in the statistical estimation model and when there exist non–linear relationship between the independent variables and dependent variables. The case study in Taips League shows that RBEMLUP is capable of accurately estimating the relationship between the influential factors from the geophysical and social–economic dimensions and the land use pattern at regional extent, through a series of treatments including mainly the multicollinearity diagnoses and singular value elimination. The research results for the case study of Taips League are of significance to supply the decision making information for land use planning and the optional utilization of land resource.

Key words: land use; land use pattern; RBEMLUP; pixel; Taips League