

基于地理回归的农作物播种面积统计数据空间化方法

夏天^{1,2}, 吴文斌^{2*}, 周清波², 周勇¹, 罗静¹, 杨鹏², 李正国²

(1. 地理过程分析与模拟湖北省重点实验室/华中师范大学城市与环境科学学院, 武汉 430079;

2. 农业部农业信息技术重点实验室/中国农业科学院农业资源与农业区划研究所, 北京 100081)

摘要: 农作物空间格局反映了农作物种植结构和特征, 是了解生产资源利用状况及进行农业结构调整的重要依据。研究旨在探索农作物播种面积统计数据空间化的方法, 实现对历史农业统计数据的空间化表达。将传统的农业统计调查与先进的遥感技术、空间地理信息技术相结合, 通过多元 Logistic 回归分析农作物格局与自然地理因素和社会经济因素之间的关系, 构建农作物空间适宜性分布概率, 在此基础上将农作物播种面积统计数据利用空间迭代分配方法, 实现统计数据空间化的研究。论文选取中国东北三省作为方法研究区, 实现了对该地区 2000—2010 年农作物播种面积空间化表达, 经检验, 该方法对东北三省水稻空间化精度为 0.76, 能够较好地完成农作物时空播种面积统计数据空间化工作。该方法是农作物调查和遥感时空格局解译研究的有效补充, 为丰富农作物空间数据提供了技术手段。

关键词: 农作物; 播种面积; 统计数据; 空间化; 地理回归

中图分类号: F302.5

文献标志码: A

文章编号: 1000-3037(2016)10-1773-10

DOI: 10.11849/zrzyxb.20151259

农作物空间格局反映了空间范围内的农业生产资源利用状况, 是了解农作物种植结构和进行农业结构调整的重要依据^[1]。同时, 农作物空间格局特征对研究农业生态系统影响碳足迹变化, 以及评价气候变化对农业生产的影响具有重要作用。近年来, 随着全球环境变化和经济一体化的加快, 我国农作物种植格局发生了显著变化。研究显示, 我国传统的“南粮北运”格局逐渐转换成“北粮南运”格局, 粮食生产的重心不断北移和西扩^[2]。这种宏观格局变化对国家粮食安全、区域水土资源可持续利用、生态安全与社会经济发展等具有重要影响^[3-4]。因此, 及时准确地掌握区域农作物空间格局及其动态变化具有重要的科学和实际意义。

纵观国内外已有研究, 遥感技术因其大范围、高效率、快速等特点, 成为区域农作物空间格局及其动态变化研究的主流方法^[5]。受卫星数据覆盖年代短、历史遥感数据获取难度大、图像解译误差多等限制, 遥感技术无法应用于长时段的农作物空间格局及其动态变化研究。农业统计数据的优势在于, 它基于行政单元进行调查, 时间维可以覆盖几十年或百年, 在分析农作物播种面积的数量、幅度和类型变化中得到了广泛应用, 但往

收稿日期: 2015-11-16; 修订日期: 2016-03-01。

基金项目: 国家自然科学基金项目 (41201089, 41271112); 中央高校科研基本业务费 (CCNU15A05058)。

[Foundation items: National Natural Science Foundation of China, No. 41201089 and 41271112; Fundamental Research Funds for the Central Universities, No. CCNU15A05058.]

第一作者简介: 夏天 (1981-), 男, 湖北省武汉市人, 副教授, 主要从事农业定量遥感、土地资源与环境遥感和空间模型等方面的研究工作。E-mail: xiatian@mail.ccnu.edu.cn

*通信作者简介: 吴文斌 (1977-), 男, 湖北省潜江市人, 研究员, 主要从事农业遥感基础与应用基础、空间建模理论与技术和农业生态系统对全球变化响应与反馈评估等方面的研究工作。E-mail: wuwenbin@caas.cn

往缺乏空间表达。随着计算机技术和空间信息技术的发展,已有学者开展了属性或统计数据空间化研究。以往研究侧重社会经济的统计数据空间化方法研究,如研究人口数据^[6-7]、GDP数据^[8-9]等,部分研究自然地理数据空间化,多集中于气象数据^[10-11]、林地数据^[12]或耕地数据^[13-14]等方面空间化的研究。然而,针对耕地内部的农作物播种面积统计数据的空间化研究较少,缺乏对历史长时间序列农作物数据空间化方法研究。事实上,农作物空间格局是人类-环境相互作用而形成,具有高度的复杂性,使得播种面积统计数据的科学空间化难度更高。基于此,本研究尝试综合利用遥感数据和空间模型,建立农作物空间分布模型,实现农作物播种面积统计数据的空间化。

1 农作物播种面积统计数据空间化方法

1.1 理论基础

农作物播种面积统计数据是空间地块农作物种植的数量汇总表达。影响农作物空间种植的因素很多,包括地形、热量、水分、土壤等自然因素,以及人口状况、经济发展水平、农业政策等社会因素。这些因素共同作用和相互反馈,驱动农作物种植的类型、数量和空间发生变化,进而影响宏观层次的农作物播种面积统计数据发生变化。基于以上关系,本研究引入空间动力学的思想来构建农作物播种面积统计数据的空间化方法。空间动力学能够描述区域系统空间相互作用、相互反馈机制及区域系统空间格局的演变^[15],已经在土地利用领域广泛使用^[16-17]。该统计数据空间化的总体思路如图1所示,通过分析农作物空间分布与影响因素(自然地理和社会经济等)关系,利用地理回归(Logistic回归)方法分析农作物空间分布适宜性,构建农作物空间分布概率;在此基础上进行统计数据空间迭代分配,将农作物播种面积统计数据空间分配到网格单元中,实现统计数据的空间化表达。

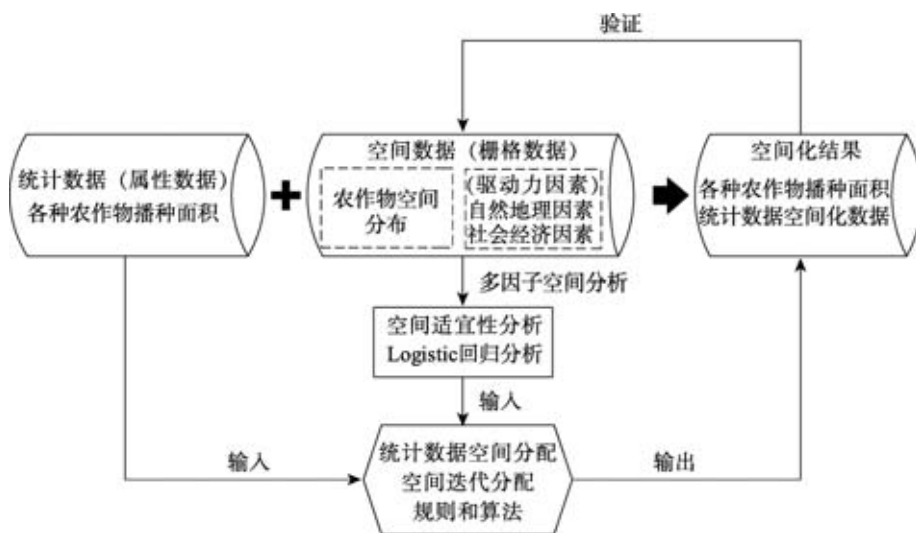


图1 统计数据空间化方法总体框架

Fig. 1 The flowchart of spatialization of statistic data

1.2 方法描述

该技术方法主要包括空间适宜性分析和空间分配两个环节。空间适宜性分析是统计

数据空间分配的基础, 本研究采用二元Logistic回归方法来进行农作物空间分布适宜性分析。具体地, 将农作物空间分布设为因变量, 将影响农作物空间分布的自然地理、社会经济等驱动因素设为自变量, 构建农作物空间分布概率模型, 描述多种驱动因素共同作用下的农作物空间分布概率, 即农作物空间分布适宜性。概率计算公式如下:

$$\text{Log}\left(\frac{1-P_i}{P_i}\right) = \beta_0 + \beta_1 \times X_{1i} + \beta_2 \times X_{2i} + \dots + \beta_n \times X_{ni} \quad (1)$$

式中: P_i 表示农作物*i*空间分布的概率; X 表示各驱动因素; β 是各驱动因子相应的回归系数。回归系数 β 运用统计软件进行二元Logistic回归计算时求得, 方程需要置信度一般大于95% (即 $\alpha \leq 0.05$), 对每一种农作物空间分布相关性较高的驱动因子则入选回归方程, 相关性较低的驱动因子不选入回归方程。

空间分配是统计数据空间化方法的关键, 本研究采用空间迭代方法将统计数据分配到空间网格单元中。首先, 基于农作物空间分布概率、农作物转换规则和空间迭代变量计算总农作物空间分布的概率总和 [式 (2)]。农作物空间分布概率通过上述Logistic回归方程计算得到。农作物转换规则主要包括转移次序和转移弹性系数设定。各农作物类型之间的转移次序利用转移矩阵来表征 (表1), 矩阵中1表示可以转变, 0表示不能转变, 如水稻不能直接转为玉米或大豆, 即转入和转出都设置为0; 玉米可以转为种植大豆, 即设置为1, 该参数决定了模拟结果中的作物变化类型。转移弹性系数 ($ELAS_u$) 用0~1的数值表示, 0代表某一作物在栅格单位上分布极易发生转变, 1代表不易发生转换。

表1 农作物转移矩阵

Table 1 Conversion matrix of different crops

	水稻	玉米	大豆
水稻	1	0	0
玉米	0	1	1
大豆	0	1	1

注: 此矩阵表可能根据实际情况进行调整设置。

$$TPROP_{i,u} = P_{i,u} + ELAS_u + ITER_u \quad (2)$$

式中: $TPROP_{i,u}$ 是栅格单元*i*上农作物类型*u*分布的总概率; $P_{i,u}$ 是通过Logistic回归方程求得的农作物空间分布概率; $ELAS_u$ 是农作物类型*u*的转移弹性; $ITER_u$ 是农作物类型*u*的迭代变量。

其次, 空间化分配控制区用于控制属性数据空间化变化过程和范围, 如耕地空间分布对农作物空间分布起控制作用, 农作物只能分布在耕地范围之内。最后, 依据概率总和, 在控制因素的约束下进行多次空间迭代, 具体的迭代步骤为: 对各农作物类型赋相同的迭代变量值 $ITER$, 并且按照每一个栅格单元中各种农作物类型分布的总概率 $TPROP$ 从大到小对每一个栅格中农作物变化进行初次分配, 将每一个栅格设置为总概率最高的农作物。对各农作物类型初次分配面积和农作物统计面积进行对比, 如农作物初次分配面积小于农作物统计面积, 则增加迭代变量 $ITER$ 值, 反之则减小迭代变量 $ITER$ 值, 随后进行第二次农作物面积空间分配。直到农作物空间分配面积与统计面积相等, 实现统计数据的空间化分配 (图2)。

2 方法验证

2.1 研究区域

东北三省位于中国大陆的东北部 (115°05'~135°02'E, 38°40'~53°34'N), 土地总面积

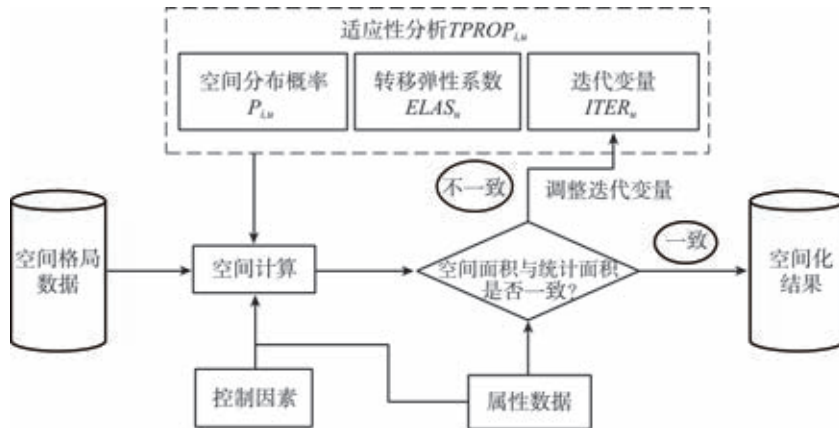


图2 空间迭代分配流程

Fig. 2 Iterative allocation procedure of statistic data

约 $79.18 \times 10^4 \text{ km}^2$ 。区域北部土壤主要为黑钙土、草甸土和白浆土，南部主要为暗棕壤，地貌以平原为主^[18-19]。东北三省耕地为 $26.44 \times 10^4 \text{ km}^2$ ，约占全国耕地总面积的 16.5%，是我国 21 世纪粮食增产和粮食供给潜力最大的地区^[20]。该地区主要种植玉米、大豆、水稻，分别占到粮食作物播种面积的 43.1%、33.0% 和 14%，是我国重要的商品粮生产基地。“十二五”期间，国家计划到 2015 年将东北地区粮食综合生产能力稳定在 $1 \times 10^8 \text{ t}$ 以上，东北三省的粮食综合生产能力对保障我国粮食安全起着至关重要的作用^[21]。因此，本研究选取中国东北三省作为统计数据空间化方法的试验区域。

2.2 数据收集及预处理

本研究所需数据共分为 4 个类型：统计数据、土地利用数据、自然地理数据和社会经济数据（表 2）。首先，将黑龙江省、吉林省和辽宁省农业统计数据按照年份进行数字化处理。然后，对农作物空间分布数据、自然地理数据和社会经济数据进行标准化处理。将农作物空间分布数据 MIRCA 2000 进行标准化处理，提取出重点研究的水稻、玉米、小麦和大豆 4 种农作物空间分布。其次对自然地理和社会经济因素数据进行标准化处理。将 DEM、气温（年平均）、降水量、 $\geq 0 \text{ }^\circ\text{C}$ 积温、 $\geq 10 \text{ }^\circ\text{C}$ 积温、人口和人均 GDP 数据均处理为 1 km 的栅格数据。土壤图按照研究区内土壤类型拆分成单独的土类分布图，共划分为 20 类土壤类型，按照类型进行编号。研究区域的坡度和坡向数据均由 DEM 数据生成，并对生成的坡向数据根据向南和向北进行区分处理，这样的处理方式有助于判断不同方位对农作物空间分布的影响。利用交通路网分布、河流水域分布和中心集镇分布数据生成研究区域内每一个栅格中心点到研究对象最近的距离。

2.3 结果与分析

将各类农作物空间数据分别与选取的自然地理和社会经济因子进行二元 Logistic 回归分析，置信度设置大于 95%（即 $\alpha \leq 0.05$ ），低于该值的驱动力因子不选入回归方程，农作物与驱动因子之间的回归分析结果如表 3 所示，然后对于每种农作物回归方程的拟合度采用 ROC 曲线进行检验^[22]。ROC 取值 0.5 ~ 1 之间，值越大，农作物概率分布与真实的农作物分布之间的一致性越好。同时，通过回归方程计算出每一种农作物在研究区域中的空间分布概率，如果空间上越接近于真实的农作物分布情况，说明构建的回归方程与真

表2 数据列表及数据来源
Table 2 Data list and data sources

数据类型	数据名称	数据格式	说明
统计数据	2000—2010年各农作物面积	属性表	黑龙江省、吉林省和辽宁省农业统计数据
土地利用数据	耕地空间分布数据	栅格	中国科学院地理科学与资源研究所数据, 1 km 网格
	农作物空间分布数据	栅格	MIRCA 2000 数据, 全球农作物分布数据
自然地理数据	DEM	栅格	中国科学院地理科学与资源研究所数据, 空间分辨率 1 km
	多年均气温分布图	栅格	国家气象资料汇编, 空间分辨率 500 m
	多年平均降雨分布图	栅格	国家气象资料汇编, 空间分辨率 1 km
	多年平均 ≥ 0 °C 积温分布图	栅格	国家气象资料汇编, 空间分辨率 500 m
	多年平均 ≥ 10 °C 积温分布图	栅格	国家气象资料汇编, 空间分辨率 500 m
	土壤图	栅格	FAO 土壤图
	1~3级交通路网分布图	栅格	国家基础地理信息数据
	1~3级河流水域分布图	栅格	国家基础地理信息数据
	中心集镇分布图	栅格	国家基础地理信息数据
社会经济数据	人口数据分布图	栅格	中国科学院地理科学与资源研究所数据, 1 km 网格, 单位: 人/km ²
	GDP 数据分布图	栅格	中国科学院地理科学与资源研究所数据, 1 km 网格, 单位: 10 ⁴ 元/km ²

实的农作物分布具有较高的一致性。表3列出了水稻、小麦、玉米、大豆和其他作物回归方程的ROC, 取值为0.85~0.95之间, 这表明, 模型得到的农作物分布概率与真实农作物空间分布具有较高的一致性, 回归方程能够较好地解释农作物的空间分布。玉米作物的ROC最高, 达到0.95, 这表明选取的影响因素能够较好地描述玉米的空间分布, 这也和东北玉米集中连片大面积种植有关, 回归样本量较大, 一定程度上提高了回归方程的解释性。水稻和大豆回归方程的ROC为0.91, 小麦为0.88, 回归方程总体解释能力都较好。其他作物的回归方程ROC相对较低, 为0.85, 主要原因是其他作物包含作物类型多, 选取的因子难以综合解释这些多作物空间分布。综上所述, 本研究的Logistic回归分析结果较好, 回归方程对农作物空间分布解释度高, 能够支撑统计数据空间化。

为进一步评价空间适宜性分析的精度, 将Logistic模型计算得到的农作物空间分布概率和农作物真实空间分布进行比较(图3)。结果显示, 农作物空间分布概率与2000年农作物真实空间分布具有较好的一致性, 其中水稻、玉米和大豆的空间分布一致性较好, 小麦的空间分布与真实值一致性相对较差, 其主要原因是小麦的样本量较少。该结果与Logistic的分析结果一致, 也表明构建的农作物分布概率方程能够较好地解释东北三省的农作物真实空间分布。

利用该方法可以对2000—2010年逐年的东北三省农作物播种面积统计数据进行空间化, 图4显示了2000、2005和2010年的农作物空间分布情况。由于缺乏2000年以后农作物真实空间分布数据, 本研究中每一种农作物逐年的空间化方法是一致的, 因此, 利用中国科学院地理科学与资源研究所2005年土地利用数据中的水田数据层视为水稻空间分布的真值, 基于Costanza提出的多尺度不同分辨率分析对比法^[23]对2005年水稻播种面积空间化结果进行了精度验证。结果表明, 东北三省水稻播种面积空间化的总体精度为0.76, 由于两者数据本身的差异性(分辨率和分类对象差异等), 一定程度上降低了精度

表3 农作物与驱动因素回归分析结果

Table 3 Regression analysis of crops and driving factors

驱动因子	水稻	小麦	玉米	大豆	其他作物
常数	-19.599 0	45.132 0	-802.484 0	268.845 0	50.763 0
坡向	0.002 0	0.002 2			0.001 1
DEM	-0.006 0	-0.000 7	-0.004 4	-0.000 8	-0.003 0
人均GDP	0.001 0	-0.000 4	-0.001 7	0.003 8	
人口密度	0.004 0	0.005 1	0.008 1	0.003 1	0.005 6
降雨	0.001 0	-0.000 3	0.000 4	-0.000 6	-0.000 3
距中心集镇距离	-0.007 0	-0.003 7	0.003 1	-0.001 0	
距河流域距离			0.000 4		
距交通路网距离		-0.000 4	-0.091 0		
坡度	-0.093 0	-0.022 6	1.383 3		-0.015 9
≥0 °C积温	0.081 0			-0.678 5	-0.049 3
≥10 °C积温	-0.062 0	-0.043 2	-0.612 7	0.422 4	
年积温	-0.026 0	0.054 5	-0.313 3	0.059 1	0.031 3
土壤类型_1		0.845 3			-0.702 0
土壤类型_2		-2.434 8			-0.699 0
土壤类型_3		-1.182 2			0.857 6
土壤类型_4		-2.221 6			-0.841 9
土壤类型_5		-1.753 0			-0.860 0
土壤类型_6					
土壤类型_7		-1.773 8			
土壤类型_8		-1.388 7		-4.502 4	
土壤类型_9					
土壤类型_10					-2.360 7
土壤类型_11					
土壤类型_12		0.546 0			
土壤类型_13		-0.679 4			
土壤类型_14					2.005 9
土壤类型_15		2.579 7			-2.535 6
土壤类型_16					-0.188 5
土壤类型_17		0.776 6			-0.357 4
土壤类型_18					-2.356 2
土壤类型_19		-3.585 1			
土壤类型_20					
土壤质地_1		0.735 2		1.329 4	0.467 0
土壤质地_2		-2.848 9		-0.451 4	
土壤质地_3		3.101 2	2.318 0		
土壤质地_4		1.530 3	1.600 0		
土壤质地_5		1.803 1		1.180 3	0.414 1
土壤质地_6			1.999 2	-2.130 6	-0.343 6
土壤质地_7		1.941 6	1.268 9	1.372 4	
土壤质地_8			1.812 9		
土壤质地_9		1.028 7		-3.131 9	
土壤质地_10		0.790 9		1.456 3	1.215 3
ROC	0.91	0.88	0.95	0.91	0.85

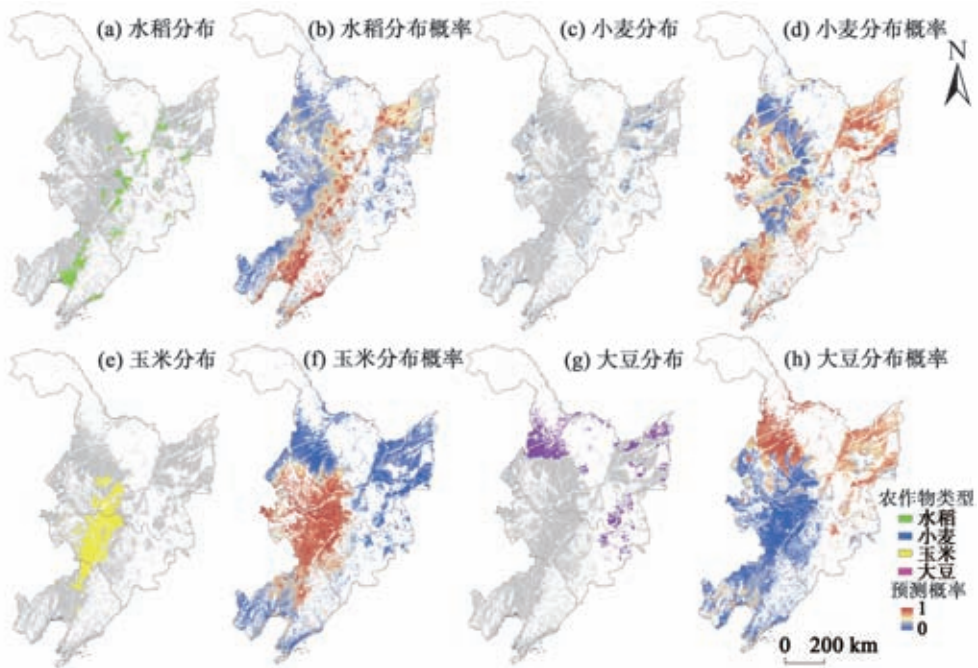


图3 农作物真实空间分布与空间分布概率对比

Fig. 3 Comparison of real distribution and distribution probability of four crops

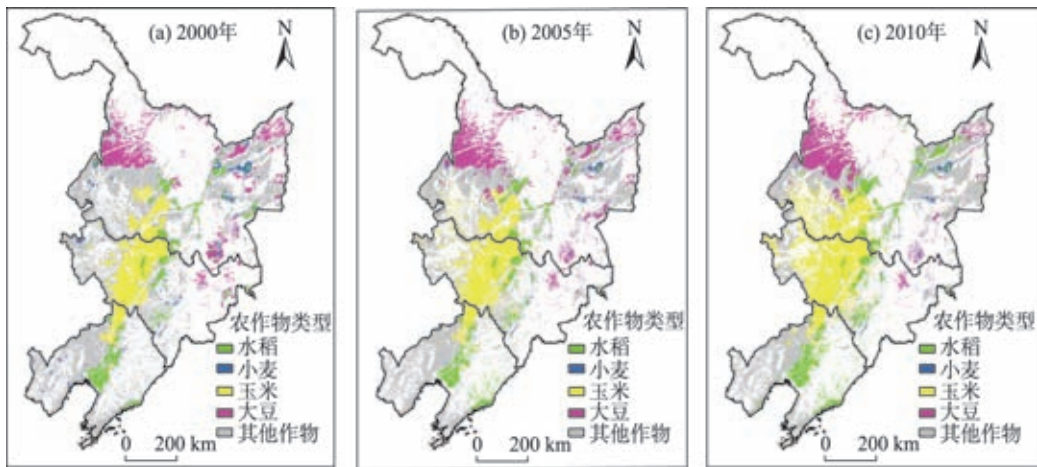


图4 东北三省2000、2005和2010年农作物空间分布

Fig. 4 Spatialized crop patterns in the Northeast China in 2000, 2005 and 2010

评价结果。可见，利用统计数据空间化的方法，能够较精确地对该地区的水稻空间格局进行表达，这表明本研究中的技术方法总体上科学可行。

3 结论和讨论

针对长时间序列历史农作物空间分布数据缺乏的现状，本研究提出了一种基于地理回归的农作物播种面积统计数据空间化方法。该方法综合考虑了影响农作物空间分布的

自然地理和社会经济等因子,利用Logistic回归方法建立农作物空间分布概率模型,结合其他转换规则和空间限制进行空间迭代,实现了农作物播种面积的空间化表达。模型在我国东北三省进行了应用,取得了较好的结果。本研究中的空间适宜性分析采用了Logistic回归分析,其优点是能够从复杂系统中剥离出主要影响因子,确定其系统变化与驱动因素的定量关系,该因果关系对统计数据空间化结果具有直接影响。然而,利用该因果关系进行长时间序列的统计数据空间化时,可能会带来一定误差,因为这种因果关系可能会随着时间发生变化,某一时段的统计关系不一定适用于其他时段。因此,利用该方法构建长时间序列数据集时,也需要对空间适宜性关系进行更新,这对基础数据提供了新的要求。

基础数据质量对空间化结果精度具有重要影响。由于目前高精度农作物空间分布数据缺乏,本研究在东北应用时采用了2000年10 km的全球农作物空间分布数据(MIRCA 2000)。虽然借助其他辅助数据对MIRCA数据进行了预处理,但基础数据质量还是会一定程度上影响模拟精度。同时,受数据限制,仅采用土地利用数据中的水田数据层对水稻空间化结果进行了检验,总体精度达到0.76,这表明该方法能够较好地实现区域尺度农作物播种面积统计数据空间化。如果进一步提高基础数据的质量和精度,统计数据空间化结果质量也会提升。

农作物播种面积统计数据空间化方法基于遥感影像解译结果,可以较好地实现历史农作物空间分布数据的重建,将时间连续的统计数据重建为时间和空间连续的农作物分布数据,有效地弥补了遥感数据在大面积、长时间序列数据生产效率不足的问题。该数据集可为农作物空间分布变化过程和规律研究提供基础数据,为农作物遥感解译提供参考补充,丰富农业土地利用研究的技术手段。

参考文献(References):

- [1] 唐华俊,吴文斌,杨鹏,等.农作物空间格局遥感监测研究进展[J].中国农业科学,2010,43(14):2879-2888. [TANG H J, WU W B, YANG P, et al. Recent progresses in monitoring crop spatial patterns by using remote sensing technologies. *Scientia Agricultura Sinica*, 2010, 43(14): 2879-2888.]
- [2] 刘玉杰,杨艳昭,封志明.中国粮食生产的区域格局变化及其可能影响[J].资源科学,2007,29(2):8-14. [LIU Y J, YANG Y Z, FENG Z M. The change of the main regions for China's foodgrain production and its implications. *Resources Science*, 2007, 29(2): 8-14.]
- [3] 刘彦随,陈百明.中国可持续发展问题与土地利用/覆被变化研究[J].地理研究,2002,21(3):324-330. [LIU Y S, CHEN B M. The study framework of land use/cover change based on sustainable development in China. *Geographical Research*, 2002, 21(3): 324-330.]
- [4] KRISHNAMURTHY P K, LEWIS K, CHOULARTON R J. A methodological framework for rapidly assessing the impacts of climate risk on national-level food security through a vulnerability index [J]. *Global Environmental Change*, 2014, 25: 121-132.
- [5] LEI Z, BINGFANG W, LIANG Z, et al. Patterns and driving forces of cropland changes in the Three Gorges Area, China [J]. *Regional Environmental Change*, 2012, 12(4): 765-776.
- [6] 戚伟,李颖,刘盛和,等.城市昼夜人口空间分布的估算及其特征——以北京市海淀区为例[J].地理学报,2013,68(10):1344-1356. [QI W, LI Y, LIU S H, et al. Estimation of urban population at daytime and nighttime and analyses of their spatial pattern: A case study of Haidian District, Beijing. *Acta Geographica Sinica*, 2013, 68(10): 1344-1356.]
- [7] 廖顺宝,李泽辉.基于人口分布与土地利用关系的人口数据空间化研究——以西藏自治区为例[J].自然资源学报,2003,18(6):659-665. [LIAO S B, LI Z H. Study on spatialization of population census data based on relationship between population distribution and land use—Taking Tibet as an example. *Journal of Natural Resources*, 2003, 18(6):

- 659-665.]
- [8] 韩向娣, 周艺, 王世新, 等. 夜间灯光遥感数据的GDP空间化处理[J]. 地球信息科学学报, 2012, 14(1): 128-136. [HAN X T, ZHOU Y, WANG S X, et al. GDP spatIALIZATION IN CHINA BASED ON NIGHTTIME IMAGERY. *Journal of Geo-information Science*, 2012, 14(1): 128-136.]
- [9] 任斐鹏, 江源, 董满宇, 等. 基于遥感和GIS的流域社会经济数据空间化方法研究[J]. 长江科学院院报, 2015, 32(3): 112-116. [REN P F, JIANG Y, DONG M Y, et al. Approach of spatialising socioeconomic data in watershed scale by remote sensing and GIS. *Journal of Yangtze River Scientific Research Institute*, 2015, 32(3): 112-116.]
- [10] 宁晓菊, 秦耀辰, 崔耀平, 等. 60年来中国农业水热气候条件的时空变化[J]. 地理学报, 2015, 70(3): 364-379. [NING X J, QIN Y C, CUI Y P, et al. The spatio-temporal change of agricultural hydrothermal conditions in China from 1951 to 2010. *Acta Geographica Sinica*, 2015, 70(3): 364-379.]
- [11] 方修琦, 侯光良. 中国全新世气温序列的集成重建[J]. 地理科学, 2011, 31(4): 385-393. [FANG X Q, HOU G L. Synthetically reconstructed holocene temperature change in China. *Scientia Geographica Sinica*, 2011, 31(4): 385-393.]
- [12] 李士成, 何凡能, 张学珍. 中国历史时期森林空间格局网格化重建方法研究——以东北地区为例[J]. 地理学报, 2014, 69(3): 312-322. [LI S C, HE F N, ZHANG X Z. An approach of spatially-explicit reconstruction of historical forest in China: A case study in Northeast China. *Acta Geographica Sinica*, 2014, 69(3): 312-322.]
- [13] 曹雪, 金晓斌, 王金朔, 等. 近300年中国耕地数据集重建与耕地变化分析[J]. 地理学报, 2014, 69(7): 896-906. [CAO X, JIN X B, WANG J S, et al. Reconstruction and change analysis of cropland data of China in recent 300 years. *Acta Geographica Sinica*, 2014, 69(7): 896-906.]
- [14] 李士成, 张德镗, 何凡能. 过去百年青海和西藏耕地空间格局重建及其时空变化[J]. 地理科学进展, 2015, 34(2): 197-206. [LI S C, ZHANG Y L, HE F N. Reconstruction of cropland distribution in Qinghai and Tibet for the past one hundred years and its spatiotemporal changes. *Progress in Geography*, 2015, 34(2): 197-206.]
- [15] 陈述彭. 城市化与城市地理信息系统[M]. 北京: 科学出版社, 1999. [CHEN S P. *Urbanization and Urban Geographic Information System*. Beijing: Science Press, 1999.]
- [16] VERBURG P H, SOEPBOER W, VELDKAMP A, et al. Modeling the spatial dynamics of regional land use: The CLUE-S Model[J]. *Environmental Management*, 2002, 30(3): 391-405.
- [17] 王晓琴. 基于空间动力学的城市用地扩张策略研究[D]. 武汉: 武汉大学, 2005. [WANG X Q. *Study on Strategy about Urban Land Expanding Based on Space Dynamics*. Wuhan: Wuhan University, 2005.]
- [18] 李奇峰, 张海林, 陈阜. 东北农作区粮食作物种植格局变化的特征分析[J]. 中国农业大学学报, 2008, 13(3): 74-79. [LI Q F, ZHANG H L, CHEN F. Changes in spatial distribution and planting structure of major crops in Northeast China. *Journal of China Agricultural University*, 2008, 13(3): 74-79.]
- [19] 方修琦, 盛静芬. 从黑龙江省水稻种植面积的时空变化看人类对气候变化影响的适应[J]. 自然资源学报, 2000, 15(3): 213-217. [FANG X Q, SHENG J F. Human adaptation to climate change: A case study of changes in paddy planting area in Heilongjiang Province. *Journal of Natural Resources*, 2000, 15(3): 213-217.]
- [20] LI Z, TANG H, YANG P, et al. Spatio-temporal responses of cropland phenophases to climate change in Northeast China[J]. *Journal of Geographical Sciences*, 2012, 22(1): 29-45.
- [21] 赵俊芳, 杨晓光, 刘志娟. 气候变暖对东北三省春玉米严重低温冷害及种植布局的影响[J]. 生态学报, 2009, 29(12): 6544-6551. [ZHAO J F, YANG X G, LIU Z J, et al. Influence of climate warming on serious low temperature and cold damage and cultivation pattern of spring maize in Northeast China. *Acta Ecologica Sinica*, 2009, 29(12): 6544-6551.]
- [22] PONTIUS J R G, SCHNEIDER L C. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA[J]. *Agriculture, Ecosystems & Environment*, 2001, 85(1/3): 239-248.
- [23] COSTANZA R. Model goodness of fit: A multiple resolution procedure[J]. *Ecological Modelling*, 1989, 47(3/4): 199-215.

Spatialization of Statistical Crop Planting Area Based on Geographical Regression

XIA Tian^{1,2}, WU Wen-bin², ZHOU Qing-bo², ZHOU Yong¹, LUO Jing¹,
YANG Peng², LI Zheng-guo²

(1. Key Laboratory for Geographical Process Analysis & Simulation, Hubei Province/ College of Urban & Environmental Science, Central China Normal University, Wuhan 430079, China; 2. Key Laboratory of Agri-Informatics, Ministry of Agriculture / Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China)

Abstract: The spatial pattern of crops reflects the planting structure and characteristics of crops, which is an important basis for understanding agricultural resource utilization and adjusting crop planting structure. This study aims to explore the method for specializing statistical data of crop planting area, and thus spatially express historical agricultural statistics data. This study used the traditional agricultural statistical survey data and remote sensing imagery data with geographic information technologies. The spatial probability distributions of suitabilities of crops are estimated using the Binary Logistic regression analysis that characterizes the relationships between the crop planting structure and the geographical factors as well as social-economic factors. Based on the spatial probability distribution, the statistical data of crop planting area were spatially distributed by using spatial iterative allocation. Northeast China was taken as the study area and the spatial expression of sown area in this area during 2000–2010 was completed. The spatial accuracy of 0.76 was achieved by using this multi-scale and multi-resolution analysis method, which demonstrated it is superior in spatially expressing statistical data of crop planting. The method can be taken as an effective complement for crop field survey and remote sensing-based crop interpretation, and thus provides novel technical means for enriching crop spatial data.

Key words: crop; planting area; statistical data; spatialization; geographical regression