基于加权马尔可夫模型的土地利用预测
——以徐州市为例

汪 婧[1]， 吴绍华[1*]， 牟守国[2]， 张 燕[1]

（1 南京大学地理与海洋科学学院，南京 210093；2 中国矿业大学环境与测绘学院，江苏徐州 221116）

摘 要： 以正在进行功能转型的徐州市区为研究对象，通过遥感影像解译将该区域内土地分为水体、林地、建设用地、农田、未利用地 5 种用地类型，并计算了 2001 年、2005 年和 2008 年的转移矩阵概率，利用加权马尔可夫方法预测徐州市未来 9 年内的土地利用变化。研究结果表明：徐州市在 2009—2017 年内，建设用地将增加 7.363 km²，水体和林地略有上升，耕地面积持续减少 9.61 km²；采用加权马尔可夫链进行修正，能提高预测结果的准确性。

关键词： 土地利用变化；预测；马尔可夫模型；徐州

中图分类号： F301.24

徐州市位于江苏省西北部，总面积 11 258 km²，占江苏省总面积的 11%，全市 2008 年总人口 940.95 万。徐州市是苏北最大城市，全国重要的交通枢纽、能源和工业基地，中国工程机械生产基地，江苏省重点规划建设的三大都市圈核心城市和四个特大城市之一，新亚欧大陆桥中国段五大中心城市之一和淮海经济区的中心城市[1]。因此，准确预测徐州市近十几年各类型土地利用的需求量，对今后土地利用管理和社会经济发展具有重要的现实意义。

国内外已有大量的学者利用马尔可夫模型对区域土地利用状况进行预测[2-6]。大部分此类预测都是以两年的土地利用转移矩阵概率为变化依据进行计算。但这种方法无法消除由于政策变化、经济发展等原因所带来的误差，预测的精确性有待提高。因此，本次研究以徐州市行政区（包括九里区、鼓楼区、云龙区、泉山区等 4 区，总面积 327 km²）为研究对象，以 2001 年、2005 年及 2008 年 3 年的土地利用变化数据，用两组步长的土地利用转移矩阵概率预测土地利用面积，将加权的思想融入到土地面积的预测中，以期提高土地利用预测的精度。

1 资料来源与研究方法

1.1 资料来源

采用 2001 年 3 月 31 日、2005 年 3 月 18 日和 2008 年 4 月 29 日 3 期不同时相的 LANDSAT-5 卫星影像作为主要的基础资料，分析研究区 2001—2008 年的土地利用变化。

1.2 研究思路

采用 ENVI 软件处理遥感影像，结合对影像的目视判读，进行土地利用分类，统计各类土地的面积，得到各时期内土地利用类型变化的转移矩阵，最后利用加权的马尔可夫模型对土地利用进行预测。

1.2.1 遥感解译 采用ENVI软件，通过几何校正、拼接和裁剪、数据融合及波段选择等技术对遥感数据进行预处理[7]。在此基础上，采用google earth，使用人工目视解译结合计算机自动识别的方法对遥感影像中的土地进行分类[8]。在分类的过程中先选择训练区，采用监督分类方法中的支持向量机进行分类。再利用ENVI导出 2001—2005 年及 2005—2008 年徐州市土地利用变化的转移矩阵，计算各期的转移概率。经过实地路线抽查检验，解译精度高达 79%，满足研究需要。

1.2.2 土地利用分类 依据国家土地资源遥感宏观调查采用的土地利用分类系统，结合研究区实际情况，以土地的用途、经营特点、利用方式和覆盖特征等因素作为分类依据，将土地利用分为建设用地、农田、林地、水体和未利用地等 5 大类。

1.2.3 马尔可夫模型原理 马尔可夫过程是随机过程的一个分支，最基本特征是“无后效性”[9]，适用于研究土地利用的动态变化[10]。在对于土地利用的动态预测中，其关键在于确定土地利用类型之间相互

* 通讯作者（shaohuawu@126.com）

作者简介：汪婧（1987—），女，安徽安庆人，硕士研究生，主要从事土地利用与土地规划等方面的研究。E-mail: wangjing198711@163.com

转化的初始转移概率矩阵 p，其描述如下：

$$P = P_{ij} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix} \quad (1)$$

$$P_{(n)} = P_{(n-1)} P_{ij}$$

式中，n为土地利用类型数目，P_{ij}为i类土地利用类型转化为j类的概率，P_{ij}满足 2 个条件：①$0 \leqslant P_{ij} \leqslant 1$；②$\sum_{i}^{n} P_{ij} = 1$（$i$，$j=1$，2，3，…，$n$）[1]。

根据马尔可夫模型，系统所研究的对象在任何时刻的状态概率向量$P_{(n)}$，可由它的初始状态概率向量$P_{(n-1)}$和转移概率矩阵P_{ij}来确定：$P_{(n)} = P_{(n-1)}\ P_{ij}$[1]。

1.2.4 加权马尔可夫模型原理 冯耀龙和韩文秀[11]将“加权”的思想引进马尔可夫链分析中，其后这种方法得到一定的研究与运用[12]。其基本思想是：一列相依的随机变数序列，各阶自相关系数刻画了各种步长序列的相依关系及其强弱。因而，可考虑分别依其前面若干时段的数据，根据相依关系，利用加权求和思想，充分利用已知信息，对未来序列的状态作出合理预测。

确定权重是加权马尔可夫法的关键，常用自相关系数作为其定量的测度。设 X_t（$t=1,2,\cdots n$）是一序列，时间间隔（$t_1 - t_2$）记为k年，计算序列的各阶自相关系数的公式为[12]：

$$r_k = \sum_{t=1}^{n-k}(X_t - \overline{X})(X_{t+k} - \overline{X}) / \sum_{t=1}^{n}(X_t - \overline{X})^2 \quad (2)$$

式中，r_k 表示第 k 阶（滞时为 k 年）自相关系数，n 为序列长度。

归一化各阶自相关系数，即：

$$\overline{w_k} = |r_k| / \sum_{k=1}^{m} |r_k| \quad (3)$$

将它们作为各阶步长的马尔可夫链的权重（m 为按预测需要计算的最大阶数）。预测未来土地利用面积的计算公式为：

$$S_t = \sum_{i=1}^{m} \overline{w_i} \times S_{ti} \quad (4)$$

式中，S_t 表示预测年份土地面积，$\overline{w_i}$ 表示第 i 年的权重，S_{ti} 代表第 i 年的土地面积。

2 结果与讨论

2.1 徐州市土地利用类型状况

根据遥感影像分类结果，得到 2001 年、2005 年及 2008 年 3 期徐州市区土地利用状况，结果见表 1 及图 1。

表 1 徐州市 2001，2005 和 2008 年土地利用面积（km^2）

Table 1 Areas of land use types in Xuzhou city in 2001, 2005 and 2008

年份	建设用地	林地	农田	水体	未利用地
2001 年	174.544	15.404	115.503	17.788	4.087
2005 年	189.190	16.400	99.870	18.110	3.741
2008 年	196.163	16.740	93.778	18.135	2.500

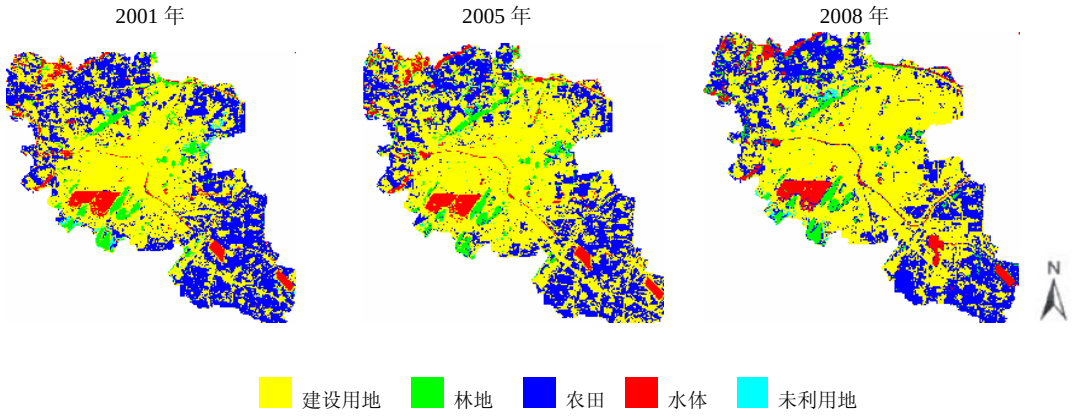

图 1 徐州市 2001、2005 和 2008 年土地利用变化

Fig. 1 Land use types of Xuzhou in 2001, 2005 and 2008

从图 1 可以看出，目前，徐州市建设用地面积持续增加，耕地逐渐减少，林地和水体在一定程度上都有所增加。建设用地主要向东南方向扩张，其次是东北方向。这主要是因为徐州市自 2004 年起开始在徐州市东南部启动建设新城区，作为一个总规划占地 50 km^2，建设用地 38 km^2 的行政区，其必然带动周围经济的发展，从而提高对建设用地的需求。水体增加主要是由于近年来徐州市塌陷区的积水，仅仅是九里区的矿区塌陷地面积就高达 28.55 km^2[13]。林地增加是由于近年来徐州市加大了生态建设，城市绿化投入资金逐年增多，另外，对一些城内原有林地比如泉山、云龙山保护力度加大。因此，新增绿化面积总量逐年增大。

利用 ENVI 软件，得出 2001—2005 年及 2005—2008 年土地利用类型之间相互转化的转移矩阵（表 2、3）。并算出土地利用类型之间相互转化的初始转移概率矩阵（表 4、5）。

表 2　徐州市 2001—2005 年土地利用类型转化的转移矩阵（km^2）

Table 2　Transition matrixes of land use types in Xuzhou from 2001 to 2005

	建设用地	林地	农田	水体	未利用地
建设用地	151.677	2.424	29.756	2.472	2.868
林地	2.807	10.925	1.335	1.209	0.127
农田	15.102	0.853	82.937	0.300	0.683
水体	3.032	0.907	0.426	13.723	0.023
未利用地	1.929	0.295	1.048	0.083	0.386

表 3　徐州市 2005—2008 年土地利用类型转化的转移矩阵（km^2）

Table 3　Transition matrixes of land use types in Xuzhou from 2005 to 2008

	建设用地	林地	农田	水体	未利用地
建设用地	154.103	2.194	36.265	1.936	1.666
林地	3.324	9.436	2.228	1.349	0.401
农田	27.018	3.639	59.630	2.186	1.305
水体	3.343	1.061	0.913	12.637	0.182
未利用地	1.402	0.070	0.834	0.003	0.187

表 4　徐州市 2001—2005 年土地利用类型转化的转移概率矩阵（km^2）

Table 4　Transition probability matrixes of land use types in Xuzhou from 2001 to 2005

	建设用地	林地	农田	水体	未利用地
建设用地	86.899	15.738	25.762	13.899	70.164
林地	1.608	70.920	1.156	6.798	3.111
农田	8.652	5.536	71.805	1.687	16.714
水体	1.737	5.890	0.369	77.149	0.558
未利用地	1.105	1.916	0.907	0.468	9.454

表 5　徐州市 2005—2008 年土地利用类型转化的转移概率矩阵（km^2）

Table 5　Transition probability matrixes of land use types in Xuzhou from 2005 to 2008

	建设用地	林地	农田	水体	未利用地
建设用地	81.454	13.380	36.312	10.689	44.524
林地	1.757	57.539	2.231	7.447	10.715
农田	14.281	22.189	59.708	12.068	34.890
水体	1.767	6.468	0.914	69.781	4.872
未利用地	0.741	0.424	0.835	0.015	5.000

由表 1 ~ 表 5 可见：徐州市建设用地面积有较大的增长；水体和林地因面积较小且少有转化；而农田和未利用的面积逐年减少；2001—2008 年期间耕地主要转化为建设用地（2001—2005 年转换的面积为 29.756 km²，2005—2008 年为 36.265 km²）主要是因为在此期间，徐州市发展加快，对建设用地的需求加大；而未利用地主要转化为农田和建设用地。

2.2 徐州市土地利用变化预测

根据徐州市土地利用面积 2001—2005 年的转移概率矩阵，以 4 年为步长，2005 年数据为初始值，进行 3 次马尔可夫过程预测 2009 年、2013 年和 2017 年的徐州市土地利用面积；通过对 2005—2008 年徐州市土地利用类型转化的转移概率矩阵，结合公式（1）和公式（2），以 2008 年的现状值为初始值，进行 3 次马尔可夫过程可预测 2011 年、2014 年和 2017 年的土地利用面积，预测结果见表 6。

表 6 加权后各期土地利用面积预测值（km²）

Table 6 Predicted areas of land use types after weighted

年份		建设用地	林地	农田	水体	未利用地
步长为 4 年	2009 年	197.856	17.175	89.919	18.613	3.749
	2013 年	203.020	17.783	83.576	19.161	3.772
	2017 年	206.062	18.263	79.516	19.686	3.788
步长为 3 年	2011 年	199.127	16.789	90.782	18.183	2.435
	2014 年	200.436	16.799	89.411	18.241	2.429
	2017 年	201.009	16.801	88.786	18.293	2.427
修正值	2017 年	203.526	17.529	84.168	18.987	3.105

根据预测，无论步长为 3 年的预测还是步长为 4 年的预测，土地利用变化的总趋势都是建设用地上升，农田面积减少，林地和水体的面积基本不变或略有上升，未利用地的面积下降。

2.3 加权马尔可夫法计算预测值

为增加预测结果的准确性，可利用加权马尔可夫模型预测 2017 年的徐州市土地利用状况。通常，步长越长，预测的准确性越低；相反，步长越短，准确性应越高。因此，预测的结果和各年面积资料的相关性有很大的关系，资料相关性越大，用其转移概率预测的结果就越准确。用 Excel 计算了 2001 年和 2005 年及 2005 年和 2008 年各土地类型面积的相关性，得出相关系数，再进行归一化处理，从而求出权重。未来土地利用面积为：

$$S_{2017} = \overline{w_1} \times S_{2013} + \overline{w_2} \times S_{2014} \quad (5)$$

加权马尔可夫法进行预测的关键是确定权重。由于本文根据 2001 年、2005 年和 2008 年的遥感数据，以 2005 年或 2008 年为基期，分别以 4 年或 3 年为步长均都可得到 2017 年的预测值，为准确预测 2017 年的土地利用状况，需要确定以 4 年和以 3 年为步长的预测值的权重。

用公式（1）和公式（2），计算 2001—2005 年及 2005—2008 年数据的相关性，得$r_1 = 0.991\,2$，$r_2 = 0.998\,7$；再用公式（3），得权重 $\overline{w_1} = 0.498$，$w_2 = 0.502$；最后用公式（5），计算修正后的 2017 年土地面积预测值，结果见表 6。

由于步长不同，故其转移概率不同。且我国的土地政策受政府导向的影响较大，具有不稳定性和不连续性：其中，在 2001—2005 年间，建设用地面积增长幅度较大，农田减少较多；而在 2005—2008 年间，由于国家加大了基本农田保护的力度，广泛推行耕地占补平衡制度，对各省市建设占用耕地即农转非的审批更为严格，使农田转为建设用地的概率减小。同时，徐州市矿区开展了采煤（矿）塌陷地复垦工作。另外，近几年徐州市加强土地整理的工作，对城市“空心村”的一些低效建设用地进行整合，提高了土地集约程度，并结合河道整治开发河滩地、整理沙碱地，提高了土地利用率和产出率，建设用地剧增及农田面积亟减的势头得到有效控制。因此，如果仅凭一组步长，不仅其数据值间相关性不高，而且也没有考虑 2005—2008 年内其他因素的变化对土地利用的影响，可能会降低预测的准确性；此外，由于政策因素的影响，采用较大步长的数据无法体现政策变动的推动作用。所以本文依据土地利用的自相关性分别对步长为 4 和 3，这两组短步长数据的预测结果进行加权，以各年数据的

自相关性为权重，以减小各年数据的误差，提高对 2017 年土地利用面积预测的准确性。

3　结论与建议

（1）2001—2005 年间，徐州市建设用地面积急速增加，耕地面积减少；而 2005—2008 年间由于基本农田保护的进一步加强，虽然建设用地面积持续上升，耕地继续减少，但耕地减少的速度有所降低。

（2）预测在 2008—2017 年间，徐州市建设用地将继续增加，农田减少，林地和水体略有增加。虽然在 2008—2017 年间徐州市几类土地面积都有较大的变化，但是变化幅度减小，这与国家严格实施耕地保护政策有关。为了控制 18 亿亩耕地红线，国家近年对农转非的审批标准更为严格，除了紧抓每年占补平衡中补充农用地的数量、质量之外，对非法占用农用地的打击力度也加强，农田减少的速率相对前几年较缓。

（3）在预测过程中，仅以一组土地利用面积转移矩阵去计算未来各类土地面积必然存在误差。为提高预测精度，可以以不同步长数据的相关性为权重，利用加权马尔可夫法预测，且为了减少因政策因素带来的误差，宜用短步长预测。

（4）近年来，江苏省提出了支持苏北发展计划，极大地推动了徐州市经济发展，但也提高了建设用地的需求量。因此，徐州市在今后的土地管理中应该更加注重控制建设用地的增长势头，以内涵挖潜代替外延扩大。同时，提高水体质量和绿地面积，对于采矿遗留下来的塌陷地，应因地制宜进行合理开发利用，使徐州市尽快步入生态型城市的行列。

参考文献：

[1] 胡召玲. 基于马尔可夫模型对徐州市土地利用动态变化的研究. 安徽农业科学, 2007, 35(8): 2344−2345

[2] 夏乐天, 朱元生, 沈永梅. 加权马尔可夫链在降雨预测中的应用. 水利水电科技进展, 2006(12): 20−27

[3] 毛凌云. 基于马尔可夫模型预测长汀县土地利用动态变化. 福建地质, 2009(1): 52−57

[4] 肖翔, 黄海涛, 吴海均, 濮励杰, 朱明, 涂小松. 基于马尔可夫模型的苏州市区土地利用变化预测. 江西农业学报, 2008, 20(11): 134−139

[5] 周兴东, 于胜文, 赵长胜. 徐州市土地利用/覆盖动态变化分析. 测绘通报, 2008(1): 33−39

[6] Bell EJ, Hinojosa RC. Markov analysis of land use change: Continuous time and stationary processes. Socio-Econ Plan Sci, 1977, 11(1): 13−17

[7] 赵文吉, 段福州, 刘晓萌, 徐智勇. ENVI 遥感影像处理专题与实践. 北京: 中国环境科学出版社, 2008

[8] 周宽久, 张世荣. 支持向量机分类算法研究. 计算机工程与应用, 2009, 45(1): 159−162

[9] 张保法. 经济预测与经济决策. 北京: 经济科学出版社, 2004: 3

[10] 张君, 刘丽. 基于马尔可夫模型的绿洲土地利用变化预测研究. 统计与信息论坛, 2007, 21(4): 73−76

[11] 冯耀龙, 韩文秀. 加权马尔可夫链在河流丰枯状况预测中的应用. 系统工程理论与实践, 1999(10): 89−98

[12] 王艳, 毛明志, 赵东方. 最优分割法确定的加权马尔可夫链预测. 中国科技论文在线, http://211.68.23.123/was40/search

[13] 王巧妮. 采煤塌陷地复垦模式综合效益评价与对策研究——以徐州九里为例 (硕士学位论文). 南京: 南京林业大学, 2008

Forecast of Land Use Based on Weighted Markov Model: A Case Study of Xuzhou

WANG Jing[1], WU Shao-hua[1], MU Shou-guo[2], ZHANG Yan[1]

(1 *Geographical and Ocean College, Nanjing University, Nanjing* 210093, *China*;

2 *School of Enviroment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, Jiangsu* 221116, *China*)

Abstract: This paper took Xuzhou city, which is ongoing function transformation, as the study region, divided land resources by interpreting remote sensing images into 5 types, i.e., water, forest, construction land, farmland and unused land, calculated the probability of transfer matrix in 2001, 2005 and 2008, and then predicted the changes of land use in the coming 9 years with weighted Markov model which could improve the accuracy of prediction. The results showed that from 2009 to 2017 the area of construction land will increase 7.363 km², the areas of water and forest lands will rise slightly, but the area of farmland will continuously decrease 9.61 km².

Key words: Land use change, Predict, Markov model, Xuzhou