

## 基于灰色预测法的芜湖市耕地变化研究

李 祥<sup>1</sup>, 王心源<sup>1,2</sup>, 顾继光<sup>3</sup>, 张广胜<sup>1</sup>

(1. 安徽师范大学国土资源与旅游学院遥感实验室, 芜湖 241000;

2. 中国科技大学地球与空间科学系, 合肥 230026; 3. 暨南大学水生物研究所, 广州 510632)

**摘 要:**应用灰色系统预测理论, 建立芜湖市耕地预测模型 GM(1, 1), 利用 1996~2003 年耕地面积建立时间数列, 对其求解、检验、预测, 可得到今后数年的芜湖市耕地面积数量, 对合理利用耕地资源、实现耕地的可持续发展有重要意义。最后并对灰色系统模型应用于耕地面积预测作了可行性分析, 研究表明: 灰色模型适合于任何一个无规律到有规律的系统。耕地数据在主导因素——政府政策没有太大的变化的情况下, 它本身就是一个理想的灰色数列。应用灰色模型作耕地面积预测为政府做耕地长期规划提供了一个较精确的预算方案。

**关键词:**灰色模型; 耕地; 预测; 芜湖市

中图分类号: F301.24

文献标识码: A

文章编号: 1005-3409(2006)05-0014-03

## Research on the Cultivated Land Change of Wuhu City Based on Grey Theory

LI Xiang<sup>1</sup>, WANG Xin-yuan<sup>1,2</sup>, GU Ji-guang<sup>3</sup>, ZHANG Guang-sheng<sup>1</sup>

(1. Laboratory of RS Application of College of National Territorial  
Resource and Tourism, Anhui Normal University, Wuhu 241000;

2. School of Earth and Space Sciences, University of Science and Technology of China, Hefei 230026, China;

3. Institute of Hydrobiology of Jinan University, Guangzhou 510632, China)

**Abstract:** Applying the grey prediction theory to establish model GM(1, 1) of the cultivated land of Wuhu City, and set up time sequence using 1996~2003 cultivated land area, and then calculate it, check out, forecast, at last the cultivated land area in the future in Wuhu City is gained, which makes it of significance of using land resources reasonably and realizing the sustainable development of cultivated land. Lastly the calculation result proved that this method is feasible and effective, the result shows: gray model is propitious to any system from rule to no rule. Cultivated land data is on the condition of steady leading factors – the government policy has made great change, and it itself is also an ideal gray sequence. It offers a better budget plan that the appliance of gray model forecasting makes lasting cultivated land project.

**Key words:** gray model; cultivated land; forecast; Wuhu City

土地是最基本的自然资源, 耕地作为土地中的精华, 为人类提供了绝大部分的农产品, 还是轻工业原料的主要来源地, 可以说耕地资源及其利用状况直接影响着农业乃至决定着国民经济的发展过程, 其动态变化和可持续利用无疑是影响区域可持续发展的关键问题<sup>[1]</sup>。特别是对于我国这样一个人口多耕地少的国家, 耕地数量的变化已成为直接影响我国农业的可持续发展、制约我国国民经济的发展速度。因此, 加强区域耕地动态变化研究, 进一步开展耕地数量变化研究, 对合理利用耕地资源、实现耕地的可持续利用具有重要意义<sup>[2]</sup>。但由于研究时间短, 加之受科学水平总体发展水平的制约, 耕地数量变化研究中存在着一定的问题。首先, 目前, 耕地数量变化的研究以宏观性大的研究区域为主, 主要集中于全国与一些省耕地变化的研究, 对中、小区域耕地数量变化的“微观”研究极少<sup>[3]</sup>。其次, 研究方法也不尽完善, 对耕地数量变化的研究总体上以定性研究为主, 很少应用定量研究, 即使运用也是诸如相关分析等较为简单的数学方法, 定量研究程度和研究精度不高<sup>[4]</sup>。因而, 本文为弥补

在耕地研究中以上两方面不足, 特从小区域, 精确定量的角度分析耕地数量的变化。

### 1 研究区概况

芜湖市位于皖东南长江沿岸, 历史上著名的四大米市之一, 是皖江的渔米之乡, 是国家批准的对外开放港口和沿江对外开放城市, 行政上辖三县四区。芜湖市位于安徽省东南部, 距安徽省省会合肥市 119 km, 地处东经 117°58′~118°43′, 北纬 30°38′~31°31′, 为北亚热带季风气候。至 2001 年底, 全市总人口 215.10 万人。耕地是芜湖市的主要土地利用类型之一, 根据芜湖市国土局 2004 年土地变更详查数据, 芜湖市现有耕地总面积 149 554.77 hm<sup>2</sup>, 占芜湖市土地总面积的 45.62%, 主要为水田和旱地, 也有少量其他类型耕地分布(表 1)。耕地在芜湖市的分布有较大差异。总体而言, 芜湖市的耕地主要集中在东部和南部地区, 所辖各个县区的耕地比例也不尽相同, 芜湖县耕地面积 3.42 万 hm<sup>2</sup>, 占总面积的 35.42%, 人均耕地 0.07 hm<sup>2</sup>; 繁昌县拥有耕地 3.53 万

\* 收稿日期: 2005-12-20

基金项目: 国家自然科学基金项目(40571162); 安徽省自然科学基金项目(050450401)

作者简介: 李祥(1977-), 男, 安徽和县人, 硕士研究生, 主要从事数学建模在地理学中的应用研究和教学; 通讯作者: 王心源(1964-), 男, 安徽六安人, 博士, 教授, 主要从事数学建模在地理学中的应用, 遥感地学应用, 遥感考古等研究。

hm<sup>2</sup>, 占该县土地总面积的 38.5%, 耕地资源较丰富; 南陵县人均耕地面积 0.26 hm<sup>2</sup>。

表 1 芜湖市耕地及其构成

耕地类型	面积/hm <sup>2</sup>	耕地构成比例/%	占总土地面积比例/%
灌溉水田	106385.42	69.7	31.63
水浇地	1042.24	0.68	0.31
旱地	38662.77	25.3	11.49
菜地	7477.29	4.32	2.22
耕地合计	152524.48	100	45.65

\* 资料来源: 芜湖市国土局 2004 土地详查数据整理。

## 2 模型的建立

### 2.1 灰色系统理论简述

灰色系统理论是由我国学者邓聚龙教授于 20 世纪 80 年代首创的一种系统科学理论。灰色系统是指既含有已知信息, 又含有未知和非确定信息的系统<sup>[5]</sup>。地理系统是一类典型的灰色系统, 因此, 自灰色系统理论产生以来, 它就被广泛应用于地理研究之中。灰色数列预测就是对某一指标的发展变化情况所作的预测, 其预测的结果是该指标在未来各个时刻的具体数值, 其预测模型相对于其他预测模型, 有其独特的优点——要求数据少, 计算简单, 将其无规律的数据, 整合成有规律。

### 2.2 数据的预处理

以给定年份内芜湖市耕地面积为一时间数据序列, 已知该序列  $\{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(N)\}$  为一离散的数据集, 由于随机因素影响, 该数据序列可能出现较大的波动。

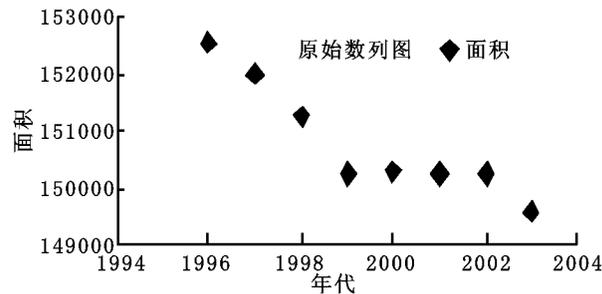


图 1 耕地面积数列图

### 2.3 建模的机制

在采用灰色模型进行预测时, 最常用的是 GM(1, 1) 模型, 它是由一个只包含单变量的一阶微分方程构成的模型。由上面的数据预处理得到一个新数列:

$$X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(N)\}$$

对于这样的新数列, 其变化趋势可以近似地用如下微分方程描述:

$$\frac{dx}{dt} + ax^{(1)} = u \quad (1)^{[7]}$$

$dx/dt$  为导数,  $ax$  项中的  $x$  为  $dx/dt$  的背景值, 记为  $\Phi$ ,  $a, u$  为参数。按导数定义, 有

$$\frac{dx}{dt} = \lim_{\Delta t \rightarrow 0} \frac{x(x + \Delta t) - x(x)}{\Delta t}$$

写成离散形式为:

$$\frac{x}{t} = \frac{x(k+1) - x(k)}{k+1 - k} = x(k+1) - x(k) = (1)(x(k+1))^{[8]}$$

这表示  $\frac{x}{t}$  是  $x(k+1)$  的一次累减生成, 我们认为, 当

$\Delta t = 1$  的很短时间内, 变量  $x(t)$  与  $x(t + \Delta t)$  之间不会出现突变, 那么, 在  $\Delta t$  很短时间内,  $dx/dt$  的背景值可取其平均

根据灰色系统预测的建模理论, 可采用多种方法进行滤波处理, 通常采用较为简便的一次累加生成, 将数据序列离散程度限制在一定的范围之内, 生成非随机数据序列, 以达到建立灰色预测模型 GM(1, 1) 的目的<sup>[6]</sup>。如果经一次累加生成后随机性仍偏大, 则可进行两次乃至多次累加生成, 以达到建模的要求。一般情况下, 对于非负的数据序列, 累加次数越多, 其随机性弱化也越明显, 规律性越强, 这样就比较易用指数曲线逼近。但累加的次数越多, 还原计算的比例就增加, 这样就增大了误差, 因而, 考虑到本文数据的特点和预测的精度, 就采用了一次累加。

表 2 芜湖市历年耕地面积 hm<sup>2</sup>

年代	1996	1997	1998	1999	2000	2001	2002	2003
面积	152524.48	151998.32	15270.07	150236.54	150247.29	150214.88	150221.51	149554.77

\* 资料来源: 芜湖市国土局 2004 土地详查数据整理。

设  $X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(8)$  是所要预测的原始数据, 对其作一次累加生成处理, 即

$$X^{(1)}(1) = X^{(0)}(1) = 152524.48$$

$$X^{(1)}(2) = X^{(0)}(1) + X^{(0)}(2) = 304522.80$$

$$X^{(1)}(3) = X^{(0)}(1) + X^{(0)}(2) + X^{(0)}(3) = 455792.87$$

⋮

$$X^{(1)}(8) = X^{(0)}(1) + X^{(0)}(2) + X^{(0)}(3) + X^{(0)}(4) + X^{(0)}(5) +$$

$$X^{(0)}(6) + X^{(0)}(7) + X^{(0)}(8) = 1206267.86$$

则得到一个新的数列 [152524.48 304522.80

455792.87 606029.41 756276.70 906491.58

1056713.09 1206267.86]。这个新数列与原数列相比, 其

随机性程度大大弱化, 平稳程度大大增加, 如图 1 所示:

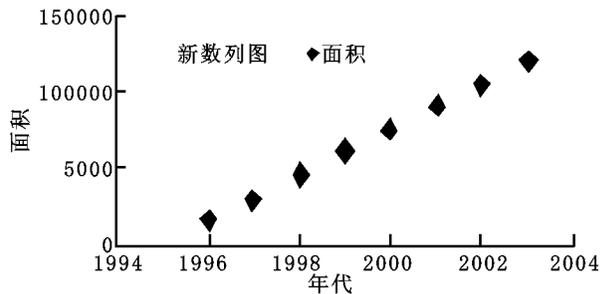


图 1 耕地面积数列图

$$\Phi(t) = 1/2[x(k) + x(k+1)] \quad k = 1, 2, \dots, n \quad (2)$$

经过微分替代运算, 即可得到以下方程:

$$X^{(0)}(k+1) = a[-1/2[x^{(1)}(k) + x^{(1)}(k+1)]] + u \quad (3)$$

上述方程中,  $X^{(0)}(k+1)$  和  $x^{(1)}(k+1)$  为已知量, 因而只有两个变量  $a, u$ , 而方程数又大于 2 个, 所以方程组在通常意义下无解。但可以通过如下最小二乘法拟合得到

$$\begin{Bmatrix} a \\ u \end{Bmatrix} = (B^T B)^{-1} B^T Y_M \quad (4)$$

根据最小二乘法的运算, 可构造数据矩阵:

$$Y_M = [X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(M)]^T$$

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \dots & \dots \\ -\frac{1}{2}[x^{(1)}(M-1) + x^{(1)}(M)] & 1 \end{bmatrix}$$

微分方程(1)所对应的时间响应函数为:

$$X^{(1)}(t+1) = [X^{(0)}(1) - u/a]e^{-a(t-1)} + u/a \quad (5)$$

上式即为数列预测的基础公式。

### 3 模型求解与检验

由公式 (2)、(3)、(4) 和原始数据可联解得到:  $a = -0.00368$ ,  $u = -7.312 \times 10^7$ 。所以芜湖市耕地预测模型为:

$$X^{(1)}(t+1) = [X^{(0)}(1) - u/a]e^{-a(t-1)} + u/a$$

$$= 7.727 \times 10^7 e^{0.00368(t-1)} - 7.312 \times 10^7$$

任何一个模型它的合理性, 只有通过检验以后才能获得, 灰色系统模型也不例外。因而, 我们可按下述方法进行精度检验:

首先, 求的原始数的还原值:

$$\hat{X}_c^{(0)}(t) = x_c^{(1)}(t) - x_c^{(1)}(t-1) \quad (6)$$

然后, 计算出原始数据的还原值与其实际值之间的残差值  $\varepsilon(0)(t)$  和相对误差值  $q(t)$  如下:

$$\varepsilon^{(0)}(t) = X^{(0)}(t) - \hat{X}_c^{(0)}(t) \quad (7)$$

$$q(t) = \varepsilon^{(0)}(t) / X^{(0)}(t) \times 100\% \quad (8)$$

最后, 计算方差比  $C = S_2/S_1$  以及小误差概率  $P\{|\varepsilon^{(0)}(t) - \varepsilon(0)| < 0.6745S_1\}^{[9,10]}$ 。由于受文章篇幅所限, 这里的  $S_1$ 、 $S_2$  以及  $\varepsilon^{(0)}$  的具体算法未做一一列出。

由公式(5)、(6)、(7)、(8) 和原始数据可联解得到以下数据(表 3)

表 3 GM(1, 1) 模型预测值与实际值的比较  $\text{hm}^2$

时间	预测数据	还原数据	原始数据	残差/q	误差率/%
1	152524.48	152524.48	152524.48	0	0
2	305547.02	153548.70	151998.32	-1550.38	-1.02
3	461672.85	156125.83	151270.07	-4855.76	-3.21
4	612330.01	150657.16	150236.54	-420.62	-0.28
5	756176.77	143846.76	150247.29	6400.53	4.26
6	908104.09	151927.32	150214.88	1712.44	1.14
7	1054945.62	146841.53	150221.51	3379.98	2.25
8	1202496.36	147550.74	149554.77	2004.03	-1.34

通过上表数据我们可以看到, 平均预测精度 98.18%, 最高的达到 99.72%, 最低的也有 95.74%, 平均误差率为 1.92%。进一步分析误差、残差都得到理想的效果。再者利用公式计算得:  $C = S_2/S_1 = 0.0452$ ,  $P\{|\varepsilon^{(0)}(t) - \varepsilon^{(0)}| < 0.6745S_1\} = 1$ , 从灰色预测精度检验等级标准(表 3)可知  $C = 0.0452$ (合格),  $P = 1$ (好)<sup>[11]</sup>。

因此, 根据上述预测公式计算, 得到 2004~2010 年芜湖市耕地面积的预测值(表 5)。用模型回测 1993 年的耕地面积为 142 913.23  $\text{hm}^2$ , 而 1993 年的耕地面积统计值是参考文献:

[1] 严长清, 袁林旺, 李满春. 江苏省耕地后备资源的空间分异及开发时序模型研究[J]. 地理与地理信息科学, 2005, 21(2): 58-61.

[2] 党安荣等. 地理信息系统支持下的中国耕地面积变化研究[J]. 中国人口、资源与环境, 1998, (4): 45-50.

[3] 蔡玉梅, 任国柱. 中国耕地数量的区域变化及调控研究[J]. 地理学与国土研究, 1998, 14(3): 13-18.

[4] 李秀彬. 中国近 20 年来耕地面积的变化及其政策启示[J]. 自然资源学报, 2004, 19(4): 329-333.

[5] 邓聚龙. 灰色系统基本理论和方法[M]. 武汉: 华中工学院出版社, 1987. 12-16.

[6] Eiumuon A. Erosion estimation of gramtic derived soils by remote sensing techniques[C]. Proceedings of the Fifth International soil Conservation Cofference, Bangkok, Thailand, 1988.

[7] 付强. 现代建模与优化技术研究-在三江平原井灌水稻节水灌溉系统中的应用[J]. 中国工程科学, 2002, 4(6): 44-45.

[8] 罗党, 刘思峰, 党耀国. 灰色模型 GM(1, 1) 优化[J]. 中国工程科学, 2003, 5(5): 50-54.

[9] 邓聚龙. 农业系统灰色理论与方法[M]. 济南: 山东科技出版社, 1998. 24-58.

[10] 徐建华. 地理数学方法[M]. 北京: 高等教育出版社, 2000. 138-195.

[11] 孙德祥, 等. 宁夏盐池半荒漠区沙漠化土地综合治理生态工程效益评价[J]. 水土保持学报, 2003, 17(1): 80-82.

141 783.53  $\text{hm}^2$ , 再看 2005 年的预测值为 143 317.85  $\text{hm}^2$  与 2005 年的统计值只相差 290.32  $\text{hm}^2$ , 进一步预测到 2010 年芜湖市耕地面积得到 101 232.65  $\text{hm}^2$ , 与政府的规划目标 99 868.92  $\text{hm}^2$  相差 1 363.73  $\text{hm}^2$ , 误差率仅为 1.36%。

表 4 灰色预测精度检验等级标准

检验指标	P	C
好	> 0.95	< 0.35
合格	> 0.80	< 0.50
勉强	> 0.70	< 0.65
不合格	≤ 0.70	≥ 0.65

表 5 芜湖市耕地预测面积(2004~2010)及回测值(1990~1995)  $\text{hm}^2$

	1990	1993	1995	2004	2005	2008	2010
预测值	143651.53	142913.23	145617.12	150229.44	143317.85	121926.80	101232.65
真实值	143232.21	141783.53	143269.58	148868.76	143164.28	115893.33	99868.92

\* 本表只列出部分预测值与真实值的比较。

### 4 耕地预测可行性分析

(1) 灰色系统预测模型本身的可行性。它适合任何一个无规律到有规律的系统, 既包含已知信息, 又含有未知和非确定信息的系统。

(2) 耕地数据的适用性。在主导因素——政府政策没有太大的变化的情况下, 耕地数据就是一个理想的灰色数列。

(3) 预测结果的正确性。从预测结果值与真实值的比较来看, 误差较小, 平均误差为 1.17%, 基本能达到一致性的要求, 并且, 做中长期预测时, 结果与耕地规划的目标相协调, 也充分说明我们政府规划的理性化。

### 5 结语与讨论

灰色预测模型相对于其他预测模型, 有其独特的优点, 那就是要求数据少, 计算简单, 将其无规律的数据, 整合成有规律。但任何一种模型有其优点, 也有其缺点, 灰色系统也不例外, 灰色系统是针对一种指数性的预测, 而且需要再还原, 因而就增大了误差的可能性, 鉴于本文数据的特点——数据加权后符合指数变化规律, 且原始数据也没有突变的过程, 其结果与检验都较为理想, 但笔者深思, 如何将灰色系统预测模型更加具有普遍适用性, 对其模型进行修正是必然的选择。这也是笔者进一步想要探究的内容, 能否将其预测模型与其他预测模型相结合使用, “取其之长, 补自之短”, 怎样修正更加完美化, 数据本身怎样处理更能与实际相统一。