

灰色动态模型 GM(1. 1)在大豆产量 预测中的应用*

胡喜平 王泽奇* 齐 宁 郭 泰 刘忠堂 张荣昌

(黑龙江省农科院合江农科所)

摘要 本文利用灰色系统理论建立高产、稳产大豆品种合丰 25产量动态模型。其模型达到了一级水平。经检验该模型平均相对误差小于 4.3%,最高精度达 98%以上。可见利用 GM(1. 1)模型预测大豆产量具有现实可信性。

关键词 灰色动态模型 大豆产量 预测

中图分类号 S565. 1

GM模型是灰色系统理论核心的一部分。GM(1. 1)模型是灰色系统预测中的基本模型,是GM(1. n)模型的特例,即n常为1;是通过数据到数据的“映射”,时间序列到时间序列的“映射”处理随机变量建立微分方程的;是一种不同于统计学方法实现新的数学逼近途径和建模思想。其特点是着眼点在现在,在未来,根据现在和过去已有的信息建立一个从过去延伸到未来的模型,从而为生产服务。

1 材料与方 法

1. 1 数据

作物产量是许多因素综合作用的结果。可以把产量的形成过程看做是既含已知信息又含未知信息的灰色系统,把作物产量看作是这个复杂系统的输出结果,可以应用GM(1. 1)模型进行模拟和预测。因此,本文建立模型数据引用合江农科所1993~1997年比试验CK产量结果(合丰25为CK),按小区实收面积折算成亩产量。即:

$$X^{(0)} = \{ 221.4 \quad 279.6 \quad 333.1 \quad 310.4 \quad 345.3 \}$$

1. 2 数据生成

作物灰色育种学认为,一切随机量都是在一定时间段上,一定范围中变化的灰色量及灰色过程。对于灰色量的处理不象数理统计方法那样去寻找概率分布,寻求统计规律,而是用数据处理的方法来找数据间的规律。某种数据处理方式被称为一种生成方式。这是一种就数找数的规律的途径。

一般对给定的原始数据列

$$X^{(0)} = \{ X_{(1)}^{(0)}, X_{(2)}^{(0)}, \dots, X_{(n)}^{(0)} \}$$

进行一次累加生成(1-AGD),可生成数据列为 $X^{(1)}$:

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

* 收稿日期 1998-05-24

* 工作单位黑龙江省农垦总局

并且, $X^{(0)}$ 与 $X^{(1)}$ 满足下列关系:

$$X^{(k)} = \sum_{i=1}^k X^{(i)} \quad (k=1, 2, \dots, n)$$

新生成的数据列弱化了原始数据列的随机性,使原来无规律的数据明显地接近指数关系规律的数据。即累加生成灰指数律,这是光滑离散函数的一种性质,数学上已有证明

1.3 GM(1,1)模型的建立原理

用数据 $X^{(1)}$ 可以建立下列的微分方程: $\frac{dx^{(1)}}{dk} + ax^{(1)} = U$, 其涵义为一阶一个变量的微分方程模型

$$\begin{array}{l} \text{系数向量} \quad \hat{a} = [a, u]^T \\ \\ \text{累加矩阵} \quad B = \begin{bmatrix} -0.5 & [X^{(1)}_1 & X^{(2)}_1] & 1 \\ -0.5 & [X^{(1)}_2 & X^{(3)}_1] & 1 \\ \\ -0.5 & [X^{(1)}_{n-1} & X^{(n)}_1] & 1 \end{bmatrix} \end{array}$$

$$\text{常数向量} \quad Y_n = [X^{(0)}_2, X^{(0)}_3, \dots, X^{(0)}_n]$$

用最小二乘法求解系数 \hat{a}

$$\hat{a} = [a, u]^T = (B^T B)^{-1} B^T Y_n$$

将其代入微分方程,得到白化微分方程的解为: $\hat{X}^{(1)}_{(k-1)} = \left[X^{(0)}_1 - U/a \right] e^{-ak} + U/a \dots (1)$

$$\hat{X}^{(1)}_{(k-1)} = -a \left[X^{(0)}_1 - U/a \right] e^{-ak} \dots (2)$$

(1)、(2)两个 GM(1,1)模型是进行灰色预测的基本计算公式。

1.4 检验

1.4.1 回代检验(残差检验)

$$E^{(1)}_k = X^{(1)}_k - \hat{X}^{(1)}_k$$

$$R^{(1)}_k = |E^{(1)}_k| / X^{(1)}_k$$

$R^{(1)}_k$ 为相对误差, $X^{(1)}_k$ 为实际生成值, $\hat{X}^{(1)}_k$ 为模型生成值, $E^{(1)}_k$ 为残差。

$$E^{(0)}_k = X^{(0)}_k - \hat{X}^{(0)}_k$$

$$R^{(0)}_k = |E^{(0)}_k| / X^{(0)}_k$$

$R^{(0)}_k$ 为相对误差, $X^{(0)}_k$ 为实际值, $\hat{X}^{(0)}_k$ 为还原值, $E^{(0)}_k$ 为残差。

1.4.2 后验差检验 利用公式如下六个:

$$\bar{X}^{(0)} = \frac{1}{n} \sum_{k=1}^n X^{(0)}_k$$

$$S = \frac{1}{n-1} \sum_{k=1}^n (X^{(0)}_k - \bar{X}^{(0)})^2 \quad \bar{E}^{(0)} = \frac{1}{n-1} \sum_{k=2}^n |E^{(0)}_k|$$

$$S_e = \frac{1}{n-2} \sum_{k=2}^n (|E^{(0)}_k| - \bar{E}^{(0)})^2$$

$$C = \frac{S_e}{S}$$

$$P = P\{11|E^{(0)}_k| - \bar{E}^{(0)} < 0.6745S\}$$

其中 $\bar{X}^{(0)}$ 为原始数据均值, S 为原始数据的离差, $\bar{E}^{(0)}$ 为残差绝对值均值, S_e 为残差绝对值离差, C 为残差绝对值离差与原始数据的离差之比, P 小误差概率。

2 结果与分析

2.1 数据来源

本文数据借助计算机利用 Qbasic程序计算得到

$$a = -0.05434601 \quad U = 271.7864$$

于是 GM(1,1)模型基本计算公式为:

$$\hat{X}_{(k)}^{(1)} = 5222.438e^{0.0543k} - 5001.037 \dots\dots\dots (3)$$

$$\hat{X}_{(k)}^{(0)} = 283.5783e^{0.0543k} \dots\dots\dots (4)$$

2.2 检验

GM(1,1)模型是否符合要求,需要对其精度进行检验。一般检验方法有回代检验和后验差检验两种

2.2.1 回代检验(残差检验) 相对误差 $R_k^{(1)}$ (或 $R_k^{(0)}$)是衡量 GM(1,1)模型的重要参数之一。一般 $R_k^{(1)}$ (或 $R_k^{(0)}) \leq 0.01$ 时,模型为一级精度; $0.01 < R_k^{(1)}$ (或 $R_k^{(0)}) \leq 0.05$ 时,模型为二级精度; $0.05 < R_k^{(1)}$ (或 $R_k^{(0)}) \leq 0.1$ 时为三级精度; $R_k^{(1)}$ (或 $R_k^{(0)}) > 0.1$ 时,模型为四级精度。对公式(3)进行检如下表 1

表 1 模型生成值相对误差检验表

| 序号 | 实际生成值 $X_{(k)}$ | 模型值 $\hat{X}_{(k)}$ | 残差 $E_{(k)}$ | 相对误差 $R_{(k)}$ | 序号 | 实际生成值 $X_{(k)}$ | 模型值 $\hat{X}_{(k)}$ | 残差 $E_{(k)}$ | 相对误差 $R_{(k)}$ |
|-----|--------------------|------------------------|-----------------|-------------------|----|--------------------|------------------------|-----------------|-------------------|
| 2 | 501.0 | 512.82 | - 11.82 | 0.024 | 4 | 1144.5 | 1145.35 | - 0.85 | 0.001 |
| 3 | 834.1 | 820.50 | 13.60 | 0.016 | 5 | 1489.8 | 1488.33 | 1.47 | 0.001 |
| 平均值 | | | | | | | | | 0.010 |

注: 相对误差平均值 $0.01 < \bar{R}_{(k)} = 0.0101 < 0.05$ 模型为二级精度

对公式(3)进行检验,可见此模型拟合的较好,相对误差为 1%,其精度达 98%以上。虽属二级精度,但非常接近一级精度水平,可信度很高

对公式(4)进行检验如表 2

表 2 模型还原值相对误差检验表

| 序号 | 实际值 $X_{(k)}$ | 模型值 $\hat{X}_{(k)}$ | 残差 $E_{(k)}$ | 相对误差 $R_{(k)}$ | 序号 | 实际值 $X_{(k)}$ | 模型值 $\hat{X}_{(k)}$ | 残差 $E_{(k)}$ | 相对误差 $R_{(k)}$ |
|-----|------------------|------------------------|-----------------|-------------------|----|------------------|------------------------|-----------------|-------------------|
| 2 | 279.6 | 291.42 | - 11.82 | 0.0423 | 4 | 310.4 | 324.85 | - 14.45 | 0.0466 |
| 3 | 333.1 | 307.68 | 25.42 | 0.0763 | 5 | 345.3 | 342.98 | 2.32 | 0.0067 |
| 平均值 | | | | | | | | | 0.0430 |

注: 相对误差平均值 $0.01 < \bar{R}_{(k)} = 0.043 < 0.05$ 模型为二级精度

相对误差为 4.3%,其精度达 95%以上,模型为二级精度,其拟合的也较好。但较(3)误差略大,不如前者模型方程拟合的更好。同时也说明,数据累加生成后,原始数据的随机性被弱化,增加了规律性,模型值更精确,较好地反映了呈指数规律变化的规律性

通过对两个基本公式精度的检验,可以看出随着时间变量 K值的增大,模型的预测值精度增高,一般呈不太规则的递增趋势。也就是说信息越新建立的模型的精度越高。

表 3 模型后验差检验精度等级

| C | P | 精度等级 | 评语 | C | P | 精度等级 | 评语 |
|----------|----------|------|----|------------|-----------|------|-----|
| < 0.35 | > 0.95 | 1 | 好 | < 0.65 | $> = 0.7$ | 3 | 勉强 |
| < 0.5 | > 0.8 | 2 | 合格 | $> = 0.65$ | < 0.7 | 4 | 不合格 |

注: 本模型的 $C = 0.1917$ $P = 0.99$ (查正态分布累积 2函数表)

2.2.2 后验差检验 残差绝对值离值与原始数据离差之比 $C=0.1917 < 0.35$,小误差出现的概率 $P=0.99 > 0.95$,可见 GM(1.1)模型达到了后验差检验精度等级的一级,好的水平。证明,本文建立的模型具有很高的可信度

当 $K=5$ 时,由公式(3)算出预计合丰 251998年产量为 181.1kg,较前 5年有所增产。

3 结 论

3.1 通过三种方法检验,证明运用灰色预测理论建立的大豆品种合丰 25年单产量灰色动态预测模型 GM(1.1),精度在 98%以上,可以作为灰色预测模型在生产上应用

3.2 大豆品种合丰 25自 1984年推广至今 14年,仍为我省主栽第一大品种,其稳产性非常好,回归系数 $b_i=0.99988$ 线性回归离差 $S^2_d=0.05725^{[3]}$ 。因此本文预测 1998年合丰 25产量与现实也是相符合的。如在保持好品种纯度提高耕作水平适当施肥的情况下,合丰 25大豆产量再提高是可以实现的,同时说明合丰 25大豆确实是一个高产稳产的优良大豆品种。

3.3 由于灰色预测是根据过去和现在的信息,寻求现在和未来的发生或运行状况,所以对大豆品种建立 GM(1.1)模型,可为该品种综合评价及生产上应用前景预测等提供重要信息。另外,建立灰色动态模型数据不易太多,应当不断地采用等维新陈代谢模型来代替旧模型,掌握最新动态,提高预测效果,利用于作物生产预测预报。农业生态值预测原始数据要具有地区的代表性,建立的模型才具有地区产量的真实性

3.4 建立不合格的 GM(1.1)模型,需要建立 GM(1.1)残差模型进行修正。原理与 GM(1.1)模型相似,本文不予赘述

参 考 文 献

- 1 郭瑞林.作物灰色育种学.中国农业科技出版社,1995,12
- 2 刘仕英,赵星.灰色动态模型在油料作物产量预测中的应用.中国油料,1990,2:53~57
- 3 王玖,杨庆凯.大豆品种稳定性分析及主要农艺性状相关性的初据.黑龙江农业科学,1985,5:6~10
- 4 王建国,陈洪涛.对提高灰色预测模型精度的探讨.农业系统科学与综合研究,1997,13(1):14~16

The Application of Gray Development Model in Yield-forecasting of Soybean

Hu Xiping Wang Zeqi Qi Ning Guo Tai Liu Zhengtang Zhang Rengchang

(Hejiang Agricultural Institute of Heil Hongjiang Academy of Agricultural Sciences)

Abstract The gray development model of soybean variety Hefeng 25 was built in the paper. The model has reached grade A. The average relative error of the development model is less than 4.3%. The highest accurate level is more than 98%. So the development model CM(1.1) may be used in forecasting soybean yield.

Key words Gray development model; Soybean yield; Forecast