安徽农业科学,Journal of Anhui Agri. Sci. 2008,36(17):7053,7056 责任编辑 理雪莲 责任校对 卢瑶

农药残留降解模型的参数估计

朱成莲 (淮阴师范学院数学系,江苏淮安 223300)

摘要 参数拟合是数学建模的重要方面。在农药残留分析中,改进的 Rayleigh 模型是农药降解规律的数学模型。介绍了改进的 Rayleigh 模型一种参数估计的方法,为研究农药残留规律提供理论依据。

关键词 农药残留;农药降解;数学模型

中图分类号 S481⁺.8 **文献标识码** A **文章编号** 0517-6611(2008)17-07053-01

Parameter Estimation of the Degradation Model of Pesticide Residue

ZHU Cheng-lian (Department of Mathematics, Huaiyin Teacher's College, Huaian, Jiangsu 223300)

Abstract Parameter fitting was an important aspect in mathematics model. A modified Rayleigh model was a mathematics model for the degradation law of pesticide. Parameter estimation method of modified Rayleigh model was introduced, which offered theoretical basis for the study on pesticide residue law.

Key words Pesticide residue; Pesticide degradation; Mathematics model

由于病、虫、草害,全世界每年损失的粮食约占总产量的 50%,使用农药可以挽回总产量的 15%左右[1]。但是由于农药的大量使用,已经严重污染了植物、空气、水、土壤,破坏了生态系统,引起了人和动物急慢性中毒。随着人们环保意识的增强,农药残留问题越来越引起人们的关注。研究农药的降解规律,选择适当的数学模型来描述农药残留的动态过程,对分析和预测农药残留量有着重要意义。

1 农药残留降解模型

目前国内外有关农药残留问题研究中,许多学者认为农药浓度的降解规律可用一级动力学方程来描述[2],即在不考虑其他因素的条件下,农药的降解速度与其浓度成正比,即:

$$\frac{dy}{dx}=-ky(k>0,y(0)=a) \tag{1}$$

式中,y 为农药在 t 时刻的浓度,t 为施药后时间,k 为农药降解速度常数,a 为农药在 $t=0$ 时的浓度(初始浓度)。解微分方程(1)可得:

$$y=ae^{-kt} \tag{2}$$

由实测数据估计参数 a、k。

农药的降解规律是随时间的增加,其浓度逐步递减,据此,方一平等提出了改进的 Rayleigh 模型[3]:

$$c=ax^{\alpha}e^{bx^2} \tag{3}$$

式中,$c=c(x)$,为 x 时刻农药的浓度,a、α、b 为待估参数。改进后的 Rayleigh 模型是 1 种较好的新方法,与常用的指数模型相比,具有误差小的特点[3]。用实测数据估计参数是建立数学模型的一个重要方面,许多学者求解参数方法是将方程(2)、(3)线性化,作代换,然后用最小二乘估计参数。而笔者用麦夸(Marquardt)算法估计模型(3)中的参数,优化农药降解模型参数拟合的算法,为更好地利用模型(3)描述农药在土壤、作物中的降解规律提供了 1 种方法。

2 模型的参数估计

对于农药残留量函数 $c=ax^{\alpha}e^{bx^2}$($c=c(x)$,为 x 时刻农药的浓度,a、α、b 为待估参数),设有 n 对试验数据$(x_i,y_i)(i=1,2,\cdots,n)$,先给出 a、α、b 的一个初始值 $a^{(0)}$、$\alpha^{(0)}$、$b^{(0)}$,记初值与真值之差(未知)为 h_1、h_2、h_3。即 $a=a^{(0)}+h_1,\alpha=\alpha^{(0)}+h_2,b=b^{(0)}+h_3$。这样,确定 a、α、b 的问题就转化为求修正值 h_1、h_2、h_3 的问题。为了方便起见,记:$c=f(x;a,\alpha,b)=ax^{\alpha}e^{bx^2}$;

$$A=\frac{\partial f(x;a^{(0)},\alpha^{(0)},b^{(0)})}{\partial a}=x^{\alpha}e^{bx^2},A_i=\frac{\partial f(x_i;a^{(0)},\alpha^{(0)},b^{(0)})}{\partial a};B=\frac{\partial f(x;a^{(0)},\alpha^{(0)},b^{(0)})}{\partial \alpha}=a\alpha x^{\alpha-1}e^{bx^2},B_i=\frac{\partial f(x_i;a^{(0)},\alpha^{(0)},b^{(0)})}{\partial \alpha};D=\frac{\partial f(x;a^{(0)},\alpha^{(0)},b^{(0)})}{\partial b}=ax^{\alpha+2}e^{bx^2},D_i=\frac{\partial f(x_i;a^{(0)},\alpha^{(0)},b^{(0)})}{\partial b}。$$

将关于 a、α、b 的三元函数 $f(x;a,\alpha,b)$在$(a^{(0)},\alpha^{(0)},b^{(0)})$处展开为一级 Taylor 展式,即:

$$f(x;a,\alpha,b)\approx f(x;a^{(0)},\alpha^{(0)},b^{(0)})+Ah_1+Bh_2+Dh_3 \tag{4}$$

要使得:

$$I=\sum_{i=1}^{n}(y_i-f(x_i;a,\alpha,b))^2\approx\sum_{i=1}^{n}(y_i-(f(x_i;a^{(0)},\alpha^{(0)},b^{(0)})+A_ih_1+B_ih_2+D_ih_3))^2=\min \tag{5}$$

应满足$\frac{\partial I}{\partial h_i}=0,i=1,2,3$。

可得方程:

$$\begin{cases}a_{11}h_1+a_{12}h_2+a_{13}h_3=e_1\\a_{21}h_1+a_{22}h_2+a_{23}h_3=e_2\\a_{31}h_1+a_{32}h_2+a_{33}h_3=e_3\end{cases} \tag{6}$$

其中,$a_{11}=\sum_{i=1}^{n}A_i^2,a_{12}=\sum_{i=1}^{n}A_iB_i,a_{13}=\sum_{i=1}^{n}A_iD_i,a_{21}=\sum_{i=1}^{n}A_iB_i,a_{22}=\sum_{i=1}^{n}B_i^2,a_{23}=\sum_{i=1}^{n}D_iB_i,a_{31}=\sum_{i=1}^{n}A_iD_i,a_{32}=\sum_{i=1}^{n}D_iB_i,a_{33}=\sum_{i=1}^{n}D_i^2,e_1=\sum_{i=1}^{n}(y_i-f(x_i;a^{(0)},\alpha^{(0)},b^{(0)}))A_i,e_2=\sum_{i=1}^{n}(y_i-f(x_i;a^{(0)},\alpha^{(0)},b^{(0)}))B_i,e_3=\sum_{i=1}^{n}(y_i-f(x_i;a^{(0)},\alpha^{(0)},b^{(0)}))D_i$ (7)

麦夸(Marquardt)算法是将式(6)改为:

$$\begin{cases}(a_{11}+d)h_1+a_{12}h_2+a_{13}h_3=e_1\\a_{21}h_1+(a_{22}+d)h_2+a_{23}h_3=e_2\\a_{31}h_1+a_{32}h_2+(a_{33}+d)h_3=e_3\end{cases} \tag{8}$$

式中,实数 d 为“阻尼因子”。

给定允许误差 ε 和阻尼因子 d,由试验数据$(x_i,y_i)(i=1,2,\cdots,n)$和初始值 $a^{(0)},\alpha^{(0)},b^{(0)}$。按式(7)计算 $a_{ij}(i=1,2,3;j=1,2,3)$及 $e_i(i=1,2,3)$,将其代入式(8)解出 h_1、h_2、h_3。再利用(8),用 $a^{(0)}+h_1$ 代换 $a^{(0)}$,$\alpha^{(0)}+h_2$ 代换 $\alpha^{(0)}$,$b^{(0)}+h_3$ 代换 $b^{(0)}$。再

(下转第 7056 页)

基金项目 江苏省教育厅自然科学基金资助项目(05KJD110034)。
作者简介 朱成莲(1966-),女,江苏涟水人,硕士,副教授,从事概率论与数理统计研究。
收稿日期 2008-04-14

有所降低(r_{1y}=0.232 0),表明密度过高,将会导致株铃数下降,对产量不利。衣分(x_4)的直接通径系数($p_{3\to y}$=0.307 0)居第 3 位,因受到通过密度间接正效应的影响,其最后效应值 r_{4y}=0.386 0 比其直接作用要大。铃重(x_3)的直接效应值 $p_{3\to y}$=0.118 0,由于它通过另外 3 个因素对皮棉的间接作用皆取正值,其净效应值 r_{3y}=0.319 0 远超过其直接效应。考虑到单铃重和衣分 2 个性状比较稳定(变异系数分别为 9.90%和 3.81%),人为促变的作用不大,而单株成铃数的伸缩性较大,其变异系数高达 25.82%。根据式(5)计算,得到各因素对产量的决定系数为 $R^2(1)$=0.037 2,$R^2(2)$=0.255 6,$R^2(3)$=0.061 4,$R^2(4)$=0.142 8。因此,对皮棉产量影响的综合作用排序为:每株铃数>衣分>单铃重>密度。根据这一结果,更加明确该品种高产栽培应在合理密植的基础上,以力争株铃、兼顾铃重、稳定衣分为主攻目标。

2.2 产量结构模拟

为探求 4 个产量因素间相互影响的数量关系,分别以每个因素变量为依变量,以皮棉产量和另外 3 个产量因子为自变量,建立回归方程$\hat{x}_i=b_i+b_y y+\sum_{j=1(j\neq i)}^{4} b_j x_j$,计算结果如下:

$$\hat{x}_1=36.782+0.011\,4y-0.983x_2+1.87x_3-0.004x_4 (F=2.558^{\Delta})$$

$$\hat{x}_2=41.124-0.318x_1+0.01y+0.295x_3-0.597x_4 (F=4.815^{**})$$

$$\hat{x}_3=4.617+0.017\,8x_1+0.008\,7x_2+0.000\,4y-0.009x_4 (F=0.626)$$

$$\hat{x}_4=37.078-0.000\,2x_1-0.104x_2-0.053\,5x_3+0.002\,6y (F=1.166)$$

模拟产量结构方程表明,密度每增加 1 000 株/hm²,株铃减少 0.318 个,铃重增加 0.017 8 g,衣分减少 0.20%;株铃每增加 1 个,铃重增加 0.008 7 g,衣分减少 0.10%;铃重每增加 1.000 0 g,衣分减少 0.05%。这表明了皮棉产量构成因素间制约与促进的数量关系。依据这个结构方程,导出不同产量水平的产量结构指标值,结果见表 3。

表 3 不同产量水平各产量构成因素的模拟值

Table 3 Simulation value of yield components at different yield levels

产量 Yield kg/hm²	密度 Density 万株/hm²	每株铃数 Number of bolls per plant//个	株铃重 Boll weight per plant//g	衣分 Lint percentage//%
750	4.040	15.20	5.43	37.15
975	4.154	16.89	5.56	37.55
1 200	4.267	18.57	5.68	37.95
1 425	4.381	20.25	5.80	38.36
1 650	4.494	21.94	5.92	38.76
1 875	4.607	23.62	6.04	39.16

2.3 产量预测 利用回归方程(7),可进一步对 SGK791棉花品种的产量进行预测。根据该品种在各试点的试验数据及生产试验资料,取密度为 48 000 株/hm²,株铃数为 21 个,铃重为 6 g,衣分为 40%,代入回归方程(7),得到预测皮棉产量为$\hat{y}_1$=1 484.2 kg/hm²,此时的预测误差为:

$$s^2(\hat{y}_1)=MSE[1+x_0'(X'X)^{-1}x_0]=226.276\,4(1+0.189\,5)=229.115\,8,\ s(\hat{y}_1)=16.406\,1$$

式中,均方误差 MSE=226.276 4,X 为回归方程(7)的设计矩阵。于是得皮棉产量 y_1 的 95%的预测区间为:

$$(\hat{y}_1-t_{0.025}(16)s(\hat{y}_1),\hat{y}_1+t_{0.025}(16)s(\hat{y}_1))=(1\,449.4,\ 1\,519)。$$

3 模型评价

一般采用多元线性回归方程$\hat{y}=b_0+b_1x_1+b_2x_2+b_3x_3+b_4x_4$来描述产量及其构成因子之间的相互关系,但因诸产量因素的量纲不同以及它们之间相互制约关系的存在,不能简单从偏回归系数的大小确定其对皮棉产量 y_1 影响的重要程度,而应从标准化回归方程的结果确定重要因素。该试验中,回归方程(7)诸偏回归系数:$b_1=16.585^{\#}$ 符号“#”表示在 α=0.1 水平上显著,$b_2=45.089^{**}$,b_3=66.033,$b_4=67.792^{\#}$。这表明单铃重比每公顷株数对产量的影响更重要,但偏相关分析的结果和通径分析的结果都表明密度比单铃重对产量更重要。这说明在分析产量构成因子的关系时,不能仅依赖于偏回归系数的结果,而要进行深入的分析,以排除因子间相互关系的影响。

通径分析和偏相关分析的结果均显示,每株铃数对产量的正向作用最大,其次是密度和衣分,作用最小的是单铃重。研究表明,单铃重和衣分 2 个性状比较稳定(变异系数分别为 9.25%和 3.81%),每株铃数是个不稳定因子,其变异系数高达 25.18%,说明它有较大的伸缩性。因此,SGK791 高产栽培应在合理密植(45 000 株/hm² 左右)的基础上,以力争株铃数、兼顾铃重、稳定衣分为主攻目标。

参考文献

[1] 袁志发,周静芋.多元统计分析[M].北京:科学出版社,2003:100-145.
[2] 卢纹岱.SPSS for Windows 统计分析[M].2 版.北京:电子工业出版社,2002:218-240.
[3] 南策雄,李尉.鄂棉 18 产量构成因子分析及高产栽培方向探讨[J].中国棉花,1997,24(7):7-9.
[4] 承泓良,何旭平,潘光照,等.棉花产量育种的数量性状分析[J].棉花学报,1998,10(6):285-291.
[5] 乔国庆,肖春鸣.海岛棉品种间铃部性状的通径分析[J].中国棉花,2005,32(4):10-11.
[6] 张取仁,施六林.棉花产量构成因素的相关和通径分析及其优质高产栽培途径[J].安徽农业科学,1997,25(3):225-226.

(上接第 7053 页)

按式(7)计算 a_{ij}(i=1,2,3;j=1,2,3)及 e_i(i=1,2,3),求出 h_1、h_2、h_3,反复迭代,直到 $|h_i|<\varepsilon$,(i=1,2,3)。文献[4]已经证明该迭代过程一定收敛。

3 结语

参数拟合是数学建模的重要方面。利用计算机,按麦夸(Marquardt)算法编程迭代运算求解已不是困难的事。参数初值的选取是迭代算法的关键。如果初值选取不恰当,则迭代可能得不到理想的结果。通常初值选取要参考问题的实际意义,一般可用随机搜索法试探性选取初值[5]。若输出结果不理想,则换初值,再重新计算。

参考文献

[1] 赵红杰,叶非.农药降解与残留分析中数学模型的应用[J].东北农业大学学报,2007,38(1):68-72.
[2] 梅进俊,施昌亚.试用数学方法预报土壤中农药积累残留量[J].环境科学,1980,1(1):44-48.
[3] 方一平,张国庆.Rayleigh 模型在农药残留上的应用[J].安徽农业大学学报,1995,22(4):462-463.
[4] 中科院计算中心.概率统计计算[M].北京:科学出版社,1979.
[5] 王莽莽,李典谟.用麦考方法最优拟合逻辑斯谛曲线[J].生态学报,1986,6(2):142-147.