

文章编号: 2095-2163(2022)11-0041-03

中图分类号: TP391.1

文献标志码: A

# 基于 Logistic 曲线的草莓生长模型研究

倪子凡, 张云华

(浙江理工大学 信息学院, 杭州 310018)

**摘要:** 草莓作为一种深受广大消费者喜爱的水果,在国内设施栽培及观光采摘领域发展迅速。但目前少有关于其生长模型的研究,多数关于草莓的研究均是基于对田间试验的统计分析来确定种植策略。Logistic 曲线模型自提出以来便一直被广泛应用于各类统计数据的拟合,本文以草莓的生长数据为研究对象,使用 Logistic 曲线模型来定量分析草莓生长周期中累积重量的增长过程,并提出一种改进的生长模型。实验结果显示,本文提出的改进生长模型相较于 Logistic 曲线模型,与实测曲线相关性更好,在不同种植策略的草莓数据上的拟合效果可决系数( $R^2$ )均优于 Logistic 曲线模型,应用前景良好。

**关键词:** Logistic 曲线模型; 生长模型; 种植策略

## Research on strawberry growth model based on Logistic curve model

NI Zifan, ZHANG Yunhua

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

**[Abstract]** Strawberry, as a kind of fruit favored by many consumers, has developed rapidly in the field of domestic cultivation and sightseeing picking. However, there are few studies on its growth model, and most studies on the strawberry are based on statistical analysis of field trials to determine planting strategies. The Logistic curve model has been widely used in the fitting of various statistical data since it was proposed. In this paper, the growth data of the strawberry is taken as the research object, and the Logistic curve model is used to quantitatively analyze the growth process of the cumulative weight in the strawberry growth cycle, and an improved growth model is proposed. The experimental results show that compared with the Logistic curve model, the improved growth model proposed in this paper has better correlation with the measured curve, and the fitting effect on strawberry data of different planting strategies ( $R^2$ ) is better than the Logistic curve model. The model is fine to be used in other fields.

**[Key words]** Logistic curve model; growth model; planting strategies

## 0 引言

在中国,草莓作为一种重要的经济作物,其年产量约 200 万吨,产值约 300 亿元<sup>[1]</sup>。随着草莓种植在国内的普及,使用科学的方法对草莓的种植策略进行指导也渐渐成为了一个热点问题。

生长模型、即研究生物生长情况的模型,一般是根据生物实际累积生长曲线进行拟合的理论生长曲线,可由累积重量和时间的函数表示。生物的生长一般情况下符合先快速增长,后平缓接近于某一临界值的规律,实际生产实践与应用中常用 S 型曲线表示,一般具有非线性特点,故可以使用非线性的数学模型来模拟生物的生长规律<sup>[2]</sup>。研究作物累积重量随时间变化的规律,建立可靠的统计模型,可以更加全面地了解作物的生长情况以及不同生长阶段的生长潜力,有助于对作物的生长进行预测,分析

种植效果,从而针对不同的情况制定不同的种植策略,以期达到最理想的种植效果。

本文使用 Logistic 曲线模型对草莓的生长过程进行研究,并提出了一种改进的生长模型,该改进模型相较于 Logistic 曲线模型更好地拟合和重现了草莓在生长周期中累积重量的变化。

## 1 相关知识

Logistic 曲线模型又被称为皮尔(Pearl)曲线模型,是一种应用非常广泛的非线性增长曲线模型。该曲线模型可用于描述事物发生、发展、成熟和极限四个阶段的生长变化过程,最早是应用于生物繁殖、人口发展统计和产品生命周期分析等方面<sup>[3]</sup>,后逐渐推广到人工神经网络、生态学、经济学等诸多学科<sup>[4]</sup>。该曲线增长率大于零,有唯一的拐点和一条水平渐近线<sup>[5]</sup>,是一种简单的饱和增长模型<sup>[6]</sup>。国

**作者简介:** 倪子凡(1997-),男,硕士研究生,主要研究方向:软件工程技术;张云华(1965-),男,博士,研究员,主要研究方向:软件工程、系统仿真、智能信息处理。

**通讯作者:** 张云华 Email:605498519@qq.com

**收稿日期:** 2022-03-10

内外学者数年来的理论和实验研究都证明该模型具有可信的识别、预测和推广能力<sup>[7]</sup>。

假设作物累积重量的自然增长率,即净增长率在一定时期内为常数  $k$ 。由该假设可知,作物的累积重量增长率,即单位时间内累积重量的增长量与当时作物的重量成正比,其比例系数为  $k$ ,且  $k > 0$ 。由于增长率本质上反映的是变化速度,故可以利用微分方程来建立模型<sup>[8]</sup>,列出方程见式(1):

$$P'(t) = kP(t) \quad (1)$$

其中,  $t$  表示时间,  $P(t)$  表示  $t$  时刻作物的累积重量,这样便得到了一个一阶常微分方程。将式(1)分离变量后积分可得公式(2):

$$P(t) = C e^{kt} \quad (2)$$

其中,当时间  $t$  取 0 时,  $P(t)$  的值为  $C$ ,即  $C$  表示初始情况下的作物累积重量<sup>[9]</sup>。该式即为马尔萨斯于 1798 年提出的指数增长模型<sup>[10]</sup>,在该模型下作物累积重量的增长是非常快的,后期将形成爆炸式增长,这显然是不符合实际情况的。回到基本假设中可以发现,实际上作物在生长到一定程度后,其自然增长率在受到资源、环境等因素的制约后并不会维持不变<sup>[11]</sup>,而是随着时间逐渐减少,即基本假设中的自然增长率  $k$  应该是  $P(t)$  的减函数。为表示出增长率受资源限制的关系,设当前环境下作物所能达到的最大累计重量为  $M$ ,则  $M - P(t)$  表示当前时间  $t$  时作物累计重量的剩余增长空间。 $k$  与  $M - P(t)$  为正相关关系,令  $k = \lambda(M - P(t))$ ,代入式(1),可得式(3):

$$P'(t) = \lambda(M - P(t))P(t) \quad (3)$$

对式(3)进行变形,得到式(4):

$$\frac{P'(t)}{M - P(t)} + \frac{P'(t)}{P(t)} = M\lambda \quad (4)$$

由式(4)可得式(5):

$$(\ln P(t) - \ln(M - P(t)))' = (M\lambda t)' \quad (5)$$

式(5)两边导数相等,则求导前的函数相差一个任意实数,设为  $C$ ,得到式(6):

$$\ln \frac{P(t)}{M - P(t)} = M\lambda t + C \quad (6)$$

两边取  $e$  的对数,得到式(7):

$$\frac{P(t)}{M - P(t)} = e^{M\lambda t + C} \quad (7)$$

令  $e^C = 1/a$ ,  $\lambda M = b$ ,解得式(8):

$$P(t) = \frac{M}{1 + a e^{-bt}} \quad (8)$$

式(8)即为 Logistic 曲线模型的积分形式。

## 2 模型改进

本文在 Logistic 曲线模型的基础上对作物累积重量的自然增长率进行了优化,使其更符合实际生产中的情况。

由前文可知,作物累积重量的自然增长率  $k$  与剩余增长空间  $M - P(t)$  之间成正相关,但实际生产中二者之间的关系并不一定是线性的<sup>[12]</sup>,而更可能是时间  $t$  的函数,且随着时间的增长持续下降<sup>[13]</sup>,可设其为  $r(t)$ ,即推得式(9):

$$k = r(t)(M - P(t)) \quad (9)$$

本文讨论  $r(t)$  为一次函数的情况,这里令  $r(t) = at + b$ , $a$  和  $b$  为参数,得到公式(10):

$$k = (at + b)(M - P(t)) \quad (10)$$

同式(8)的推导过程,可得公式(11):

$$P(t) = \frac{M}{1 + e^{-\frac{1}{2}atM^2 - Mbt - c}} \quad (11)$$

其中,  $c$  为任意实数。

## 3 实验与分析

### 3.1 实验数据及工具

为了对比 2 种模型的拟合效果,本文采用了 Universidade Federal de Santa Maria 的草莓实验数据集,该数据集使用了 16 种不同的处理方法对草莓进行种植,在完全成熟阶段,每周 2 次、共进行 37 次收获并称量。

拟合的工具使用的是 SPSS26,这是一款图形操作界面友好的统计软件,可以对目标数据集进行回归分析,并给出参数的估计值和拟合优度<sup>[14]</sup>。

### 3.2 实验数据的预处理

由于 SPSS26 在拟合数据较大时会出现拟合失败的情况,故需对数据进行预处理。本文采用计算频率的方式处理数据,即将每一时刻的累积重量除以总累积重量,时间同理,以此将数据大小限制在 0 到 1 之间,方便进行拟合<sup>[15]</sup>。

### 3.3 实验结果

通过使用 SPSS26 软件对 2 种模型各自在 5 组数据上进行拟合,对比 Logistic 曲线模型和本文提出的改进模型的拟合效果,采用可决系数 ( $R^2$ ) 作为指标来评价拟合优度,计算方法见式(12):

$$R^2 = 1 - \frac{\sum (P(t) - \hat{P}(t))^2}{\sum (P(t) - \bar{P}(t))^2} \quad (12)$$

将一组数据预处理后输入 SPSS26 后使用改进

生长模型曲线进行拟合, 拟合结果如图 1 所示。图 1 中, 横坐标为迭代次数, 纵坐标为残差平方和。图 1 中, 从图 1 可以看出, 数据拟合效果较好, 虽然前几次迭代中残差平方和有所上升, 但在接下来的迭代中残差平方和迅速降低, 并趋于稳定。

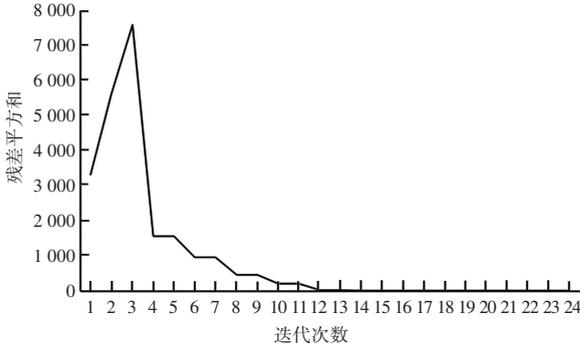


图 1 拟合结果残差平方和

Fig. 1 RSS of fitting results

将 2 种模型在相同数据集上进行拟合, 拟合的评价结果见表 1。

表 1 2 种模型的实验结果

Tab. 1 Results of two models

| dataset | $R^2$ (Logistic) | $R^2$ (本论文) |
|---------|------------------|-------------|
| 1       | 0.996            | 0.998       |
| 2       | 0.994            | 0.996       |
| 3       | 0.988            | 0.991       |
| 4       | 0.994            | 0.995       |
| 5       | 0.994            | 0.996       |

由表 1 可以看出, 本文提出的改进生长模型对比 Logistic 曲线模型在可决系数的指标上都有所提高, 可以更准确地重现草莓在生长过程中的累积重量变化, 具有良好的指导意义。

## 4 结束语

本文提出了一种基于 Logistic 曲线模型的生长模型, 对 Logistic 曲线模型中的作物累积重量自然增长率进行了优化, 使其更符合实际情况。本文对多

种不同处理下的草莓生长数据分别使用 Logistic 曲线模型和改进生长模型进行拟合, 实验结果表明改进后的生长模型的拟合效果更好。考虑到生长模型曲线泛用性强的特点, 今后的研究工作中将会尝试将其推广到更多的领域, 并结合实际情况加以改进, 以便为更多领域的数据统计提供一个可靠的统计模型。

## 参考文献

- [1] 李竞涵, 张艳玲, 杨彤, 等. 品种创新技术改良—中国草莓育种开始发力[J]. 果农之友, 2014(04): 44-46, 50.
- [2] 王志跃, 陈伟亮, 白群安, 等. 新扬州鸡生长模型比较研究[J]. 黑龙江畜牧兽医, 2002(12): 4-5.
- [3] PEARL R, REED L J. On the rate of growth of the population of the United States since 1790 and its mathematical representation [M]//Mathematical Demography. Biomathematics. Berlin/Heidelberg: Springer, 1977, 6: 341-347.
- [4] YIN Xi, ZELENAY P. Kinetic models for the degradation mechanisms of PGM-free ORR catalysts [J]. ECS Transactions, 2018, 85(13): 1239-1250.
- [5] 朱珉仁. Gompertz 模型和 Logistic 模型的拟合[J]. 数学的实践与认识, 2002(05): 705-709.
- [6] 杨义群, 吴良欢, 张火法. Gompertz 曲线与 logistic 增长曲线之比较[J]. 生物数学学报, 1996, 11(03): 166-168.
- [7] 朱华锋. Logistic 模型的参数估计及其实证研究分析[J]. 科技信息, 2011(01): 169-170.
- [8] 段成林, 董胜, 王智峰. 基于 Logistic 模型的北极海冰生长曲线研究[J]. 海洋湖沼通报, 2021, 43(04): 1-6.
- [9] 俞启泰. 论 Usher、Logistic 和 Gompertz 三种增长曲线的使用价值[J]. 新疆石油地质, 2001(02): 136-141, 87.
- [10] 万昌秀, 梁中宇. 逻辑斯谛曲线的一种拟合方法[J]. 生态学报, 1983(03): 288-296.
- [11] 崔启武, LAWSON G. 一个新的种群增长数学模型——对经典的 logistic 方程和指数方程的扩充[J]. 生态学报, 1982(04): 403-415.
- [12] 乔松珊, 孙成金. 优化的 Logistic 模型及其在城镇化水平预测中的应用[J]. 许昌学院学报, 2016, 35(05): 10-15.
- [13] 李文灿. 对 Logistic 方程的再认识[J]. 北京林业大学学报, 1990, 12(02): 121-128.
- [14] 王玉梅. 用 SPSS 拟合人口增长模型实例[J]. 电子技术, 2013, 42(10): 10-11.
- [15] 赵晓进, 梁芝栋, 邵立杰, 等. SPSS 软件非线性回归功能的分析与评价[J]. 统计与决策, 2021, 37(23): 20-22.