

# 两类生长曲线模型的参数辨识及对疫情数据的拟合

钟铭兴\*, 孙丹, 戴兰, 张潇腾, 李景

长沙理工大学数学与统计学院, 湖南 长沙

收稿日期: 2021年9月21日; 录用日期: 2021年10月14日; 发布日期: 2021年10月25日

---

## 摘要

本文研究了Logistic和Richards两类种群生长模型的差异。Logistic方程是Richards方程异速增长参数取定值2的特例。从微分方程出发, 我们求出了两类生长方程的显式通解, 通过取对数、求导、换元和差分等操作, 利用线性最小二乘法对参数进行估计, 最后结合湖南省的新冠疫情实例对比说明Logistic和Richards增长曲线方程的适用性及参数估计的有效性。

## 关键词

Logistic方程, Richards方程, COVID-19, 曲线拟合

---

# Parameter Identification and Epidemic Data Fitting of Two Types of Growth Curve Models

Mingxing Zhong\*, Dan Sun, Lan Dai, Xiaoteng Zhang, Jing Li

School of Mathematics and Statistics, Changsha University of Science and Technology, Changsha Hunan

Received: Sep. 21<sup>st</sup>, 2021; accepted: Oct. 14<sup>th</sup>, 2021; published: Oct. 25<sup>th</sup>, 2021

---

## Abstract

The differences between Logistic and Richards population growth models were studied. The Logistic equation is a special case of the constant value 2 of allometric growth parameter of Richards equation. Starting from the differential equation, we calculate the growth equations of the two kinds

---

\*通讯作者。

of explicit conditions and by the exponential, derivation, changing and difference operations such as, using the linear least square method to estimate the parameters, and finally combined with the outbreak of the new champions league of Hunan Province to explain the applicability of Logistic and Richards growth curve equation and the validity of parameter estimation.

## Keywords

Logistic Equation, Richards Equation, COVID-19, Curve Fitting

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

种群生态数学模型可以很好地刻画种群或种群间的变化规律，如预测人口变化趋势或传染病的传播和爆发。其中较为知名的有 Logistic 模型、Gompertz 模型和 Richards 模型等。对于前两种模型，它们的生长曲线具有特定的拐点，只能准确描绘一种特定形状的 S 型曲线。因此对给定生长过程建模时，在生长方程的比较和筛选方面需要做大量工作。Richards 在 1931 年研究流体[1]通过多孔介质中毛细管传导作用时建立了 Richards 方程，情况才得以改观。Richards 认为 Bertalanffy 方程中的异速生长常数限定为 2/3 太严格，这一限制使得该模型不能应用到大部分生物生长上，将其设为一般化参数  $m$ ，得到一个含有 4 个参数的非线性 S 型增长模型。这种模型对生长过程描述的准确性达到了前所未有的水平[2]。

邢黎峰[3]在研究 Richards 方程时发现，当异速生长参数  $m$  变化时，不仅包括了 Brody、Logistic 等生长方程，而且包含了它们的中间过渡形式和更为广义的形式，并得出 Richards 模型比众多生长模型更为准确的结论。对种群生态数学模型的参数估计是模型得以应用的前提，因而对该类模型的参数估计受到许多学者的关注和研究。吴新元[4]指出常用于 Logistic 方程拟合的方法有三点法、麦考特法和枚举优选法，但这些方法在应用上都有其局限性；魏冠军[5]等将参数视为随机变量，利用 Bayes 理论进行参数估计，同遗传算法和最小二乘法进行对比，得出前者的估计效果优于后两者；徐文科[6]提出双向差分拟合方法和双向差分广义加权最小二乘法[7]，并对初始预报值进行了改进，通过与优选法、麦考特法等方法比较，结果表明精度有所提高；刘洋[8]利用双向差分广义加权最小二乘法对 Gompertz 增长曲线方程的参数估计进行研究，并在 SARS 的疫情实例中得到了较好的拟合效果；乔钰[9]利用线性最小二乘法和双向差分加权最小二乘法对 Richards 增长曲线的参数进行估计，并对具有相关关系的 Richards 方程组建立半相依回归模型[10]。

综上所述，对于一系列生物生长方程，已有许多的数值方法和统计模型对参数进行估计，在实际应用中，选择合适的模型和参数估计方法就显得尤为重要。本文研究了 Logistic 和 Richards 两类种群生长模型，利用线性最小二乘法对参数进行估计，最后结合湖南省的新冠疫情实例说明增长曲线方程的适用性及参数估计的有效性。

## 2. Logistic 生长模型模拟新冠疫情

### 2.1. Logistic 方程的产生

种群动态研究一直是生态学研究的核心问题之一。Logistic 方程指出了有限空间种群增长的基本规律：

当一个物种迁入到一个新生态系统中后,其数量会发生变化。当物种的起始数量小于环境最大容纳量时,数量会增长;若此生态系统中食物、空间等资源有限(非理想环境),则种群数量呈S型增长。Logistic方程是研究种群动态的重要工具,其微分形式为:

$$\begin{cases} \frac{dN}{dt} = r \left( 1 - \frac{N}{N_m} \right) N \\ N(t_0) = N_0 \end{cases} \quad (1)$$

其中  $N$  为种群个体总数,  $t$  为时间,  $r$  为种群增长潜力指数,  $N_m$  表示环境最大容纳量,  $N_0$  为  $t_0$  时刻的种群数量。利用分离变量法并代入初值条件  $N(t_0) = N_0$ , 解得:

$$N = \frac{1}{\frac{1}{N_m} + \left( \frac{1}{N_0} - \frac{1}{N_m} \right) e^{-r(t-t_0)}} \quad (2)$$

Logistic 方程的性质:

- 1) 由式(1)知,  $N(t)$  是关于  $t$  的递增函数;
- 2) 对式(1)求导并结合式(2), 得到拐点坐标

$$t = \frac{1}{r} \ln \left( \frac{N_m}{N_0} - 1 \right) + t_0, \quad N_t = \frac{N_m}{2} \quad (3)$$

- 3) 存在两条渐近线, 当  $t \rightarrow t_0$  时,  $N \rightarrow N_0$ ; 当  $t \rightarrow \infty$  时,  $N \rightarrow N_m$ 。

## 2.2. 生物生长曲线的一般特征

生物体在一个完整的生长过程中其生长速度通常具有慢-快-慢的变化特征。它的累积生长量最初较小, 随着时间的延长逐渐增大, 最终稳定在一个饱和值附近。这一过程若用曲线来表示, 将是一条拉长的S型曲线(见图1)。

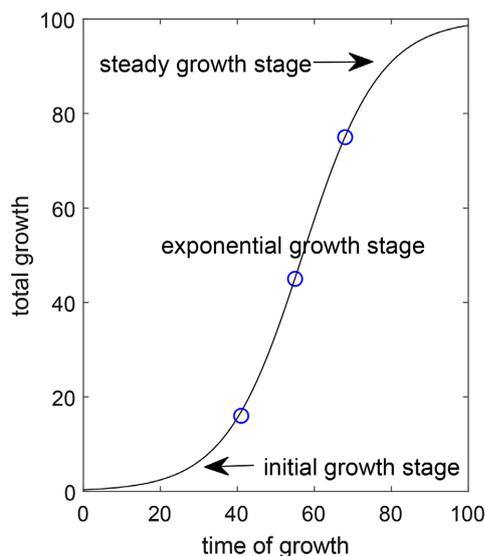


Figure 1. Growth curve of "S" type

图1. "S"型生长曲线

COVID-19 累计病人数具有慢 - 快 - 慢的变化特征, 且初值较小, 随着时间的延长逐渐增大, 最终稳定在一个饱和值上。Logistic 和 Richards 方程能够很好地描述其初始生长阶段、快速增长阶段和稳定生长阶段的累计病人数的变化情况。而任意两个生长过程都不会完全相同, 它们的差异受生物物种、生态环境、观测指标和时间序列等诸多因素的影响[11]。

### 2.3. Logistic 增长曲线的参数估计

由式(2), Logistic 方程的一般形式为  $y(t) = \frac{1}{A + Be^{-Kt}}$ , 取对数得:

$$\ln y(t) = -\ln(A + Be^{-Kt}). \quad (4)$$

对式(4)求导得:

$$\frac{1}{y} \frac{dy(t)}{dt} = \frac{BKe^{-Kt}}{A + Be^{-Kt}}. \quad (5)$$

由式(4)、(5)可得:

$$\frac{dy(t)}{dt} = BKe^{-Kt} y^2(t). \quad (6)$$

对式(6)取对数得:

$$\ln \frac{dy(t)}{dt} = \ln BK - Kt + 2 \ln y(t). \quad (7)$$

进一步令  $\beta_0 = BK$ ,  $\beta_1 = -K$ , 上式可表为

$$\ln \frac{dy(t)}{dt} - 2 \ln y(t) = \beta_0 + \beta_1 t. \quad (8)$$

再用向后差分  $\Delta y(t) = y(t) - y(t-1)$  近似微分, 设误差为  $e(t)$ , 得差分方程

$$\ln \Delta y(t) - 2 \ln y(t) = \beta_0 + \beta_1 t + e(t) \quad (9)$$

$$\text{令 } X = \begin{pmatrix} 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \\ 1 & n \end{pmatrix}, \quad Y = \begin{pmatrix} \ln \Delta y(2) - 2 \ln y(2) \\ \ln \Delta y(3) - 2 \ln y(3) \\ \vdots \\ \ln \Delta y(n) - 2 \ln y(n) \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \quad e = \begin{pmatrix} e(2) \\ e(3) \\ \vdots \\ e(n) \end{pmatrix},$$

将式(9)写成矩阵形式即  $Y = X\beta + e$ , 进而得到  $\beta$  的最小二乘估计

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = (X^T X)^{-1} X^T Y. \quad (10)$$

对于式(10), 已求得  $B$ 、 $K$  的估计值, 只需将所有观测值代入求得  $A$  的一系列观测值, 并剔除异常点, 取均值即得  $A$  的估计值。

至此, Logistic 方程的参数估计值全部得到。

对 2020 年湖南省的疫情历史数据进行复现拟合, 利用上述方法得到一组初值  $[0.00129, 0.01956, 0.2981]^T$ , 进而利用 LM 非线性拟合方法得到参数, 最后得到的 Logistic 方程为:

$$y = \frac{1}{0.001 + 0.0156 \times e^{-0.2797x}}$$

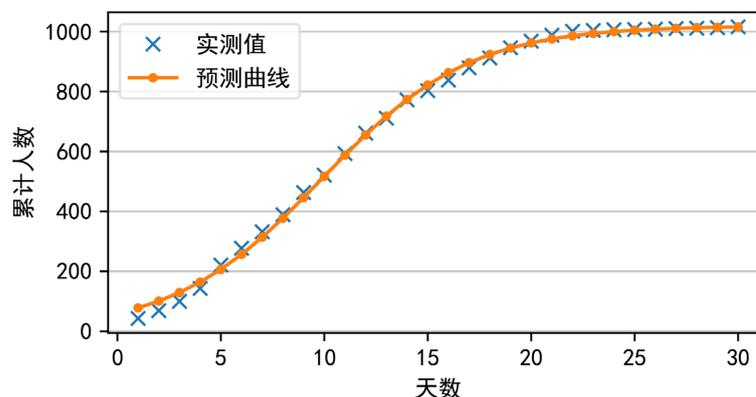


Figure 2. Fitting results of logistic equation of “S” growth curve

图2. “S”型生长曲线 logistic 方程拟合结果

Logistic 方程的拟合结果与实测值的均方根误差(RMSE)为 15.3300, 由图 2 可以看出, Logistic 方程对前期数据的拟合效果较差, 通过查找文献, 我们发现了一类适应度更加广阔的生长方程, 即 Richards 方程。

### 3. Richards 方程模型

#### 3.1. Richards 方程的产生

Richards 认为 Bertalanffy 方程中的异速生长常数限定为  $2/3$  太严格, 应将其设为一般化参数  $m > 0$ , 它的取值与生物物种和环境条件有关, 并提出假设:

$$\frac{dV}{dt} = \eta V^m - \gamma V \quad (11)$$

这是  $m$  次的 Bernoulli 微分方程,  $V$  是种群个体数,  $\eta$ 、 $\gamma$  和  $m$  为异速增长参数。令  $Z = V^{1-m}$ , 化简得非齐次微分方程:

$$\frac{dZ}{dt} = \gamma(m-1)z + (1-m)\eta. \quad (12)$$

利用通解公式解得:

$$Z = \exp\left\{\int \gamma(1-m)dt\right\} \left( \int (1-m)\eta \exp\left\{-\int \gamma(1-m)dt\right\} dt + C \right) \quad (13)$$

代入初值条件  $V(t_0) = V_0$  得到特解[10]:

$$V(t) = P \left( 1 - Qe^{-Ht} \right)^{1/(1-m)} \quad (14)$$

式中,  $P = \left( \frac{\eta}{\gamma} \right)^{1/(1-m)}$ ,  $Q = \left( 1 - \frac{V_0^{1-m}}{\eta/\gamma} \right) e^{Ht_0}$ ,  $H = \gamma(1-m)$ 。通常把式(14)叫做 Richards 增长曲线方程,

这里  $P$ 、 $Q$ 、 $H$ 、 $m$  为四个待估参数, 分别含有不同的生物学意义。 $P$  是研究对象的生物学上限,  $Q$  是生长初始值参数,  $H$  是生长速率参数,  $m$  是异速生长参数, 其决定着曲线图像的走势以及拐点的位置。

#### 3.2. Richards 生长方程的性质与可塑性

在文献[12]中, Fengri Li 研究了 Richards 方程各参数不同取值的情况, 并依据解结构和生物学解释

划分为 8 种类型。因为 COVID-19 疫情的变化特点, 在本文我们只考虑  $P$ 、 $H > 0$ ,  $Q > 0$  且  $m > 1$  的情形, 此时 Logistic 方程是  $m = 2$  时的特例。

- 1) 对式(14)求导知,  $V(t)$  是关于  $t$  的递增函数;
- 2) 对式(14)求导并令二阶导数为 0, 得到拐点, 其坐标为

$$t = \frac{\ln(|Q|/(1-m))}{H}, V_t = Pm^{1/(1-m)} \quad (15)$$

在该点生物增长率达到峰值  $\frac{dV}{dt} = PHm^{m/(1-m)}$ , 称之为拐点;

- 3) 存在两条渐近线,  $t \rightarrow t_0$  时,  $V \rightarrow V_0$ ;  $t \rightarrow \infty$  时,  $V \rightarrow P$ 。

Richards 方程具有很强的可塑性, 当方程中的参数  $m$  连续取不同的值时, 曲线形状会不断发生变化(见图 3)。由(15)式可知, 参数  $m$  的取值影响拐点的横、纵坐标大小, 也即影响拐点的相对位置。当  $m$  取值较小时, “S” 曲线的拐点位置偏低, 该曲线的下半臂较短, 上半臂更长; 反之, 当  $m$  取值较大值时, “S” 曲线的拐点位置偏高, 该曲线的下半臂较长, 上半臂更短。特别地, 通过变换 Richards 增长曲线方程中参数  $m$  的值, 可以分别得到许多著名且应用广泛的方程。例如令参数  $m$  为 0、2/3、2, 分别得到 Mitscherlich 方程、Bertalanffy 方程和 Logistic 方程。

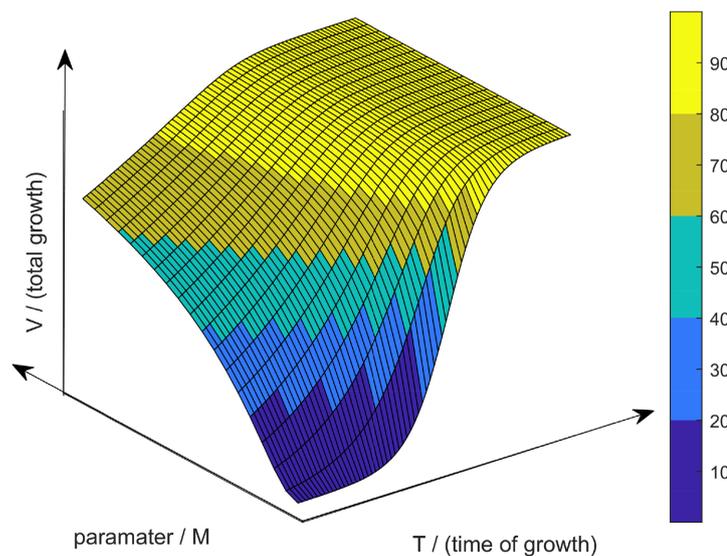


Figure 3. Plasticity of Richards equation  
图 3. Richards 方程的可塑性

### 3.3. Richards 增长曲线的参数估计

同上, 对式(14)取对数求导得:

$$\ln \frac{dy(t)}{dt} = \ln \left( -\frac{HQ}{\delta} P^{-\delta} \right) + (\delta + 1) \ln y(t) - Ht. \quad (16)$$

进一步令  $-QP^{-\delta} = C$ ,  $\beta_0 = \ln \frac{H}{\delta} C$ ,  $\beta_1 = \delta + 1$ ,  $\beta_2 = -H$  上式可表为:

$$\ln \frac{dy(t)}{dt} = \beta_0 + \beta_1 \ln y(t) + \beta_2 t. \quad (17)$$

再用向后差分  $\Delta y(t) = y(t) - y(t-1)$  近似微分, 设误差为  $e(t)$ , 得差分方程

$$\ln \Delta y(t) = \beta_0 + \beta_1 \ln y(t) + \beta_2 t + e(t) \tag{18}$$

$$\text{令 } X = \begin{pmatrix} 1 & \ln y(2) & 2 \\ 1 & \ln y(3) & 3 \\ \vdots & \vdots & \vdots \\ 1 & \ln y(n) & n \end{pmatrix}, \quad Y = \begin{pmatrix} \ln \Delta y(2) \\ \ln \Delta y(3) \\ \vdots \\ \ln \Delta y(n) \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad e = \begin{pmatrix} e(2) \\ e(3) \\ \vdots \\ e(n) \end{pmatrix},$$

将式(18)写成矩阵形式即  $Y = X\beta + e$ , 假设  $E(e) = 0$ ,  $Cov(e) = \sigma^2 I$ , 进而得到  $\beta$  的最小二乘估计

$$\beta = (\beta_0, \beta_1, \beta_2)^T = (X^T X)^{-1} X^T Y. \tag{19}$$

对于式(19), 已求得  $\delta$ 、 $H$ 、 $C$  的估计值, 将所有观测值代入求得  $P$  的一系列观测值, 剔除异常点, 取均值得至此, Richards 方程的参数估计值全部得到。

利用上述方法得到一组初值  $[1031.0, -9.7, 0.2774, 1.726]^T$ , 进而利用 LM 非线性拟合方法得到参数, 最后得到的 Richards 方程  $y = 1037.3 \times (1 + 2.4 \times e^{-0.2x})^{\frac{1}{1-1.4}}$ 。

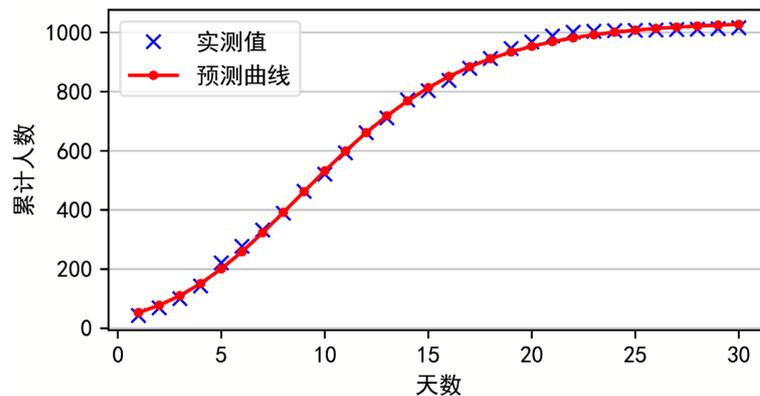


Figure 4. Fitting results of Richards equation  
图 4. Richards 方程拟合结果

Richards 方程的拟合结果与实测值的 RMSE 为 10.6444, 由图 4 可以看出, Richards 方程较 Logistic 的结果大大改进了, 但后期的值存在波动, 所以我们利用残差倒数对两者的结果结合进行预测。

### 4. 组合预测方法

采用下式所示的误差倒数法对两个模型的结果进行加权组合, 其中  $\omega_i$  为权值系数,  $f_{1t}$  和  $f_{2t}$  分别为 Logistic 和 Richards 的预测值。权值系数中  $\varepsilon_1$ 、 $\varepsilon_2$  分别为两者的误差。

$$\begin{cases} f_t = \omega_1 f_{1t} + \omega_2 f_{2t}, \quad t = 1, 2, \dots, n, \\ \omega_1 = \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2}, \quad \omega_2 = \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2}. \end{cases} \tag{20}$$

组合模型的 RMSE 为 9.8681, 从下图 5 可以看出对两类曲线方程的拟合效果都较好, 使用残差倒数法对结果进行校正, 该方法对误差小的模型会赋予较大的权值系数, 从而使整个组合模型误差减小, 达到提升整体预测精度的效果。

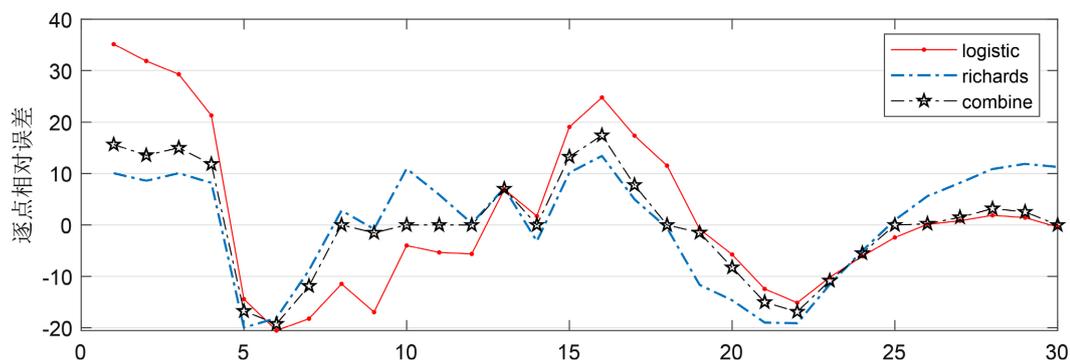


Figure 5. Point by point relative error

图 5. 逐点相对误差

## 5. 结论

对于具有微分方程形式的生物数学模型, 本文利用线性最小二乘法, 从定量的角度对两类增长曲线方程的参数进行估计, 并将其应用到湖南省的新冠疫情实例中。结果表明利用增长曲线模型预测传染病的传染增长规律具有一定的可行性, Richards 方程具有更为广泛的适应性、良好的解析性和预测性。这种方法易于编程实现, 不仅适合于 Logistic 和 Richards 曲线方程的拟合, 而且适用于其他非线性曲线的拟合, 因此在生物学中有广泛的应用。

## 基金项目

湖南省大学生创新创业训练项目(NO.202010536019)。

## 参考文献

- [1] 李凤日, 吴俊民, 鲁胜利. Richards 函数与 Schnute 生长模型比较[J]. 东北林业大学学报, 1993(4): 15-24.
- [2] 陈华豪, 徐文科. 生物种群动态数学模型的参数辨识方法[J]. 农业系统科学与综合研究, 1993(1): 1-5+50.
- [3] 邢黎峰, 孙明高, 王元军. 生物生长的 Richards 模型[J]. 生物数学学报, 1998, 13(3): 348-353.
- [4] 吴新元. 逻辑斯谛曲线拟合的一种数值方法[J]. 生物数学学报, 1990(1): 26-32.
- [5] 魏冠军, 党亚民. 基于 Bayes 理论的 Logistic 增长曲线模型参数估计方法研究[J]. 工程勘察, 2012, 40(5): 64-67.
- [6] 徐文科, 张石生. 生物种群 Logistic 模型双向差分拟合方法及对初始预报值的改进[J]. 农业系统科学与综合研究, 1997(1): 1-3+9.
- [7] 徐文科. 基于微分方程的生态数学模型统计分析[D]: [博士学位论文]. 哈尔滨: 东北林业大学, 2009.
- [8] 刘洋, 徐文科. Gompertz 增长曲线方程的参数估计及对初始预报值的优化[J]. 黑龙江大学自然科学学报, 2015, 32(5): 600-606.
- [9] 乔钰, 徐文科. Richards 增长曲线的参数估计[J]. 哈尔滨师范大学自然科学学报, 2015, 31(5): 23-26.
- [10] 乔钰. Richards 增长曲线方程及其方程组的参数估计[D]: [硕士学位论文]. 哈尔滨: 东北林业大学, 2017.
- [11] 焦莹莹. 基于 Richards 增长曲线的全球新冠疫情分析[D]: [硕士学位论文]. 哈尔滨: 黑龙江大学, 2021.
- [12] Liu, Z.G. and Li, F.R. (2003) The Generalized Chapman-Richards Function and Applications to Tree and Stand Growth. *Journal of Forestry Research*, **14**, 19-26. <https://doi.org/10.1007/BF02856757>