

孙 鹏, 孙 源. 3 种方法对中肋骨条藻 Logistic 生长模型拟合的比较研究[J]. 江苏农业科学, 2019, 47(2): 229–232.
doi:10.15889/j.issn.1002-1302.2019.02.057

3 种方法对中肋骨条藻 Logistic 生长模型拟合的比较研究

孙 鹏¹, 孙 源²

(1. 曲阜师范大学生命科学学院, 山东曲阜 273165; 2. 青岛袁策生物科技有限公司, 山东青岛 266000)

摘要:为探索高效准确的 Logistic 生长模型拟合方法, 对中肋骨条藻在不同浓度二价镉[$\text{Cd}(\text{II})$]胁迫下的 Logistic 生长曲线, 分别用 Excel、Origin 和 Matlab 3 种软件进行拟合。结果表明, 拟合后的 R^2 大小顺序为 Matlab > Origin > Excel; 采用 Excel 非线性拟合中的四点法拟合效果优于三点法。提示 Matlab(用 Levenberg–Marquardt 算法进行非线性化多次迭代)是 3 种软件中对 Logistic 生长曲线拟合的最优方法。

关键词:海洋硅藻; 中肋骨条藻; Logistic 生长曲线; 相关系数 R^2 ; Matlab; 模型拟合; 最优方法筛选

中图分类号: Q141; S184 **文献标志码:** A **文章编号:** 1002-1302(2019)02-0229-04

Logistic 曲线(或 Logistic 函数)模型由比利时数学家 Verhulst 首次提出^[1], 最初主要是描述和研究人口增长的一种 S 形曲线。其方程为 $y = \frac{a}{1 + b \cdot e^{-kx}}$, 随后 Logistic 曲线在各领域得到了广泛应用, 如 Pearl 等首次将 Logistic 曲线应用于生物领域^[2]; Fisk 在经济学领域引入 Logistic 曲线^[3]; Anderson 将 Logistic 曲线应用到医疗诊断中^[4]。如今, Logistic 曲线几乎替代了一切的 S 形曲线, 成为描述种群生态、动植物及微藻生长变化的一条最为普遍的曲线^[5]。

收稿日期: 2017-08-25

基金项目: 山东省自然科学基金(编号: ZR2015DL009); 曲阜师范大学校级课题(编号: xkj201512)。

作者简介: 孙 鹏(1980—), 男, 山东济宁人, 硕士, 讲师, 主要从事微藻生物技术及生物统计学研究。E-mail: hisunpeng@163.com。

支管间距为 a 、支管间距为 b , 则不发生漏喷的临界状态是喷头 p_1 、 p_2 、 p_3 、 p_4 以射程 r 为半径的圆相交于典型计算区域中的零点, 因此, 不发生漏喷的条件是 $r \geq \frac{\sqrt{a^2 + b^2}}{2}$ 。就本例而言, $30 \text{ m} \times 30 \text{ m}$ 布置要求的喷头喷洒半径 r 应大于 21 m , 而喷头实测最大喷洒半径为 21 m , 有效半径约为 20 m , 因此 $30 \text{ m} \times 30 \text{ m}$ 布置将会发生漏喷现象或喷洒水量极低的情况。可见, 尽管计算的均匀度满足规范要求, 但实际上是不可行的。(2)从控制面积上看, 按照 $26 \text{ m} \times 26 \text{ m}$ 布置时, 平均单喷头控制面积为 676 m^2 , 而 $20 \text{ m} \times 20 \text{ m}$ 布置时平均单喷头控制面积为 400 m^2 , 相差 69%, 这意味着同一项目, $20 \text{ m} \times 20 \text{ m}$ 布置比 $26 \text{ m} \times 26 \text{ m}$ 布置的投入资金要多得多。(3)从平均喷灌强度来看, 按照 $26 \text{ m} \times 26 \text{ m}$ 布置时, 平均喷灌强度为 6.4 mm , 而按照 $20 \text{ m} \times 20 \text{ m}$ 布置时平均喷灌强度 11.21 mm , 相差 75%, 这意味着在不考虑土壤入渗能力的情况下, $26 \text{ m} \times 26 \text{ m}$ 布置比 $20 \text{ m} \times 20 \text{ m}$ 布置喷洒效率要低得多, 同一地块作业时间也要长得多。

从以上计算分析可以看出, 在具体喷灌项目设计时, 利用

Logistic 曲线的拟合就是确定方程中的 3 个参数值 a 、 b 、 k , 经典方法包括线性和非线性方法。线性方法是 Logistic 曲线方程转为线性方程, 然后采用最小二乘法让残差平方和达到极小值以对参数进行无偏估计。非线性方法可直接用相关算法得到 Logistic 模型参数的最小二乘无偏估计^[6]; 现采用的算法有莱文贝格–马夸特方法(Levenberg–Marquardt)方法、数值方法、三次样条插值函数法等^[7]。上述参数估计的过程较复杂, 须借助各种软件及编程语言来进行拟合分析, 如统计学软件 SPSS、SAS, 程序设计语言 R、Python。目前热门的机器学习算法也涉足非线性拟合领域(如支持向量机回归^[8]), 而掌握这些软件和语言对生物从业者较为困难。

中肋骨条藻(*Skeletonema costatum*)是常见的海洋硅藻, 分布较广泛。近年来, 研究者多对海洋微藻重金属胁迫下的生长过程用 Logistic 方程进行描述, 进而反映出藻生物量与生

上述方法可快速计算出喷洒均匀度、平均喷灌强度等技术参数, 通过土壤入渗强度和作物轮灌周期以及项目资金投入等的综合对比分析, 进而确定合理的喷头布置间距。

参考文献:

- [1] 喷灌工程技术规范: GB/T 50085—2007[S]. 2007.
- [2] 杨路华, 刘玉春, 柴春玲, 等. 应用 Surfer 软件进行喷(微)灌均匀度分析[J]. 节水灌溉. 2004, (5): 14–16.
- [3] 张志宇. 喷头水量分布的智能仿真与组合间距的优化[D]. 保定: 河北农业大学, 2006.
- [4] 劳冬青, 韩文霆. 喷头水量分布仿真及组合优化软件系统研究[J]. 节水灌溉. 2010, (1): 42–45.
- [5] 张 洋. 基于 C# 与 OpenGL 喷头水量分布动态模拟及组合优化[D]. 杨凌: 西北农林科技大学, 2012.
- [6] 刘晓扬, 杨路华, 柴春玲, 等. 微喷头水量分布仿真及组合优化研究[J]. 节水灌溉. 2016, (3): 24–26.
- [7] 张玉龙. 农田水利学[M]. 3 版. 中国农业出版社, 2013.
- [8] 李小平. 喷灌系统水量分布均匀度研究[D]. 武汉: 武汉大学, 2005.

长时间、重金属浓度的关系^[9]。本研究以中肋骨条藻为模式生物,分别采用 Excel、Origin、Matlab 3 种常用的作图和计算软件,对其在二价镉[$\text{Cd}(\text{II})$]胁迫下的生长用 Logistic 模型进行拟合,对拟合方法和效果进行介绍和比较,为生命科学从业者进行 Logistic 曲线的拟合、参数估计和数据处理提供借鉴和参考。

1 材料与方法

1.1 微藻培养及毒性试验

中肋骨条藻接种自中国海洋大学海洋生命学院,在曲阜师范大学生命科学学院环境生态学实验室采用 Guillard F/2 (f/2) 培养基培养,培养条件为温度 $(20 \pm 1)^\circ\text{C}$,光源为白色日光灯,光照强度为 3 000 lx,12 h 光照 12 h 黑暗培养;各步骤均进行灭菌处理。

取 20 mL 藻液于 250 mL f/2 培养基中,微藻起始细胞密度均为 10^4 个/mL,加入 4 种不同浓度梯度 (0、0.1、1.0、5.0 mg/L) 的 $\text{Cd}(\text{II})$ 试验,每个浓度下均设平行样 3 个,采用与微藻培养相同的培养条件,每隔 24 h 取 10 mL 藻液,用 Coulter 计数仪测定藻液细胞密度,重复测定 3 次^[10],连续测定 7 d。

1.2 数据处理软件及系统

软件:Microsoft Excel 2003、Origin 8.0、Matlab R2010b;系统:Windows 7 × 86 with Service Pack 1。

1.3 微藻生长拟合

“1.1”节中的微藻生长曲线用“1.2”节中的 3 种软件进行 Logistic 回归拟合。

1.3.1 Excel 对 Logistic 曲线的线性化拟合

1.3.1.1 三点法和四点法确定 a 值^[11] 三点法计算公式:

$$a = \frac{2N_1N_2N_3 - N_2^2(N_1 + N_3)}{N_1N_3 - N_2^2}, 2x_2 = x_1 + x_3。 \quad (1)$$

式中: (x_1, N_1) 、 (x_2, N_2) 、 (x_3, N_3) 分别表示实测数据的起点、中点、终点。

四点法计算公式:

$$a = \frac{N_1N_4(N_2 + N_3) - N_2N_3(N_1 + N_4)}{N_1N_4 - N_2N_3}, x_2 + x_3 = x_1 + x_4。 \quad (2)$$

式中: (x_1, N_1) 、 (x_4, N_4) 分别为实测数据的起点、终点, (x_2, N_2) 、 (x_3, N_3) 为中间 2 点。

1.3.1.2 转化为直线回归方程,求得 b 、 k 值 a 已知时可将方程由非线性模型转为线性模型^[12],在本模型中 x 为藻类培

养时间 $t(\text{h})$,转化如下:

$$y = \frac{a}{1 + b \cdot e^{-k \cdot x}}; \quad (3)$$

$$\ln\left(\frac{a}{y} - 1\right) = \ln b - kx; \quad (4)$$

$$\text{设 } y' = \ln\left(\frac{a}{y} - 1\right); \quad (5)$$

$$\text{则有 } y' = \ln b - kx。 \quad (6)$$

利用公式(5)和 a 值可以先求出 y' ,利用 Excel 对 y' 和 x 进行线性回归分析,拟合后可得到 b 和 k 值。在 Excel 2003 中的操作如下:

(1)[工具]→[加载宏]→勾选“分析工具库”。(2)[工具]→[数据分析]→[回归]→弹出回归对话框→单击“Y 值输入区域(Y)”输入框右边的折叠按钮,选取 y' ,单击“X 值输入区域(X)”输入框右边的折叠按钮,选取 x 的数据;→在“输出选项”框中选择“输出区域”单选框,点击右边的折叠按钮,任选一输出区域;→勾选“残差”框中的“线性拟合”;→单击[回归]对话框中的[确定],显示回归结果,并绘得散点图。(3)利用散点图,添加趋势线可得到线性回归公式。

1.3.2 Origin 进行 Logistic 曲线的非线性回归拟合 (1) 打开 Origin 8.0,点选 [column]→[add new columns]输入须要增加的数据列数(本试验 3 列)。(2)按列输入数据后,选中所有数据,菜单栏中选择 [Plot]→[Symbol]→[Scatter] 绘制散点图(图 1-a 图)。(3)关闭坐标轴设置对话框,点选 1 - 选择黑色方框数据组 b 进行拟合→菜单栏中选择 [Analysis]→[Fitting]→[Nonlinear Curve Fit]→[Open Dialogue] 打开对话框;按步骤点选 2 - Category - Growth/Sigmonial; 3 - Function - Slogistic3; 4 - Fit,即完成拟合和作图(图 1-b),同时给出参数和统计处理数据。

1.3.3 用 Matlab 进行 Logistic 曲线的非线性回归拟合 打开 Matlab R2010b,用简单编程语言和自带的控件工具箱对 Logistic 曲线快速拟合:(1)在 Command Window 对话框中输入原始数据:横坐标 x 以及 y_1 、 y_2 、 y_3 、 y_4 这 4 条生长曲线的数据。(2)依次点击软件左下角: [Start]→[Toolboxes]→[Curve Fitting]→[Curve Fitting Tool (cftool)],弹出对话框“Curve Fitting Tool”。(3)在 Curve Fitting Tool 对话框中点击 Data→弹出 Data 对话框→选择: X Data: x ; Y Data: y_1 ;→点击 Create data set 按钮可绘制散点图;(4)在 Curve Fitting Tool 对话框中点击 Fitting,弹出 Fitting 对话框→在 Fitting 对话框中点选 New fit, Type of fit: 选择 Custom Equations, 点击 New 按

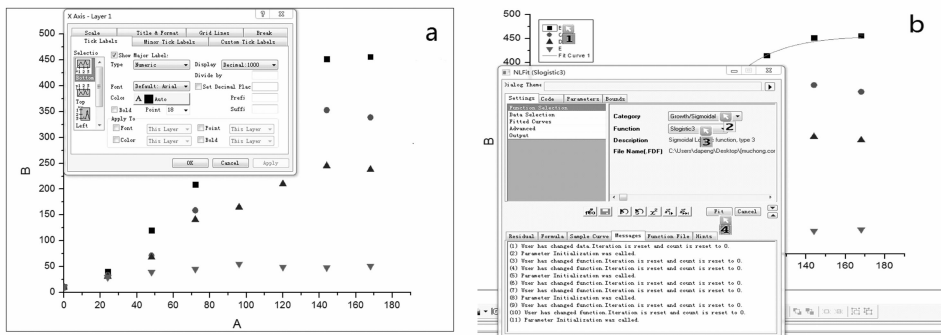
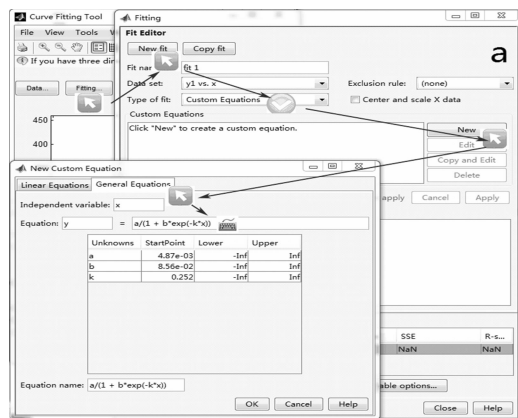


图1 Origin 8.0 对生长曲线拟合操作截图

钮→弹出 New Custom Equation 对话框,单击选择 General Equations 选项→在 General Equations 选项中输入 Equation: $y = a / [1 + b \cdot \exp(-k \cdot x)]$ →点击 OK(图 2-a)。(5)回到



Fitting 选项卡,点击 Fit options → Algorithms: Levenberg - Marquardt 算法→Close→Apply,即可得拟合后的 Logistic 曲线图(图 2-b)。

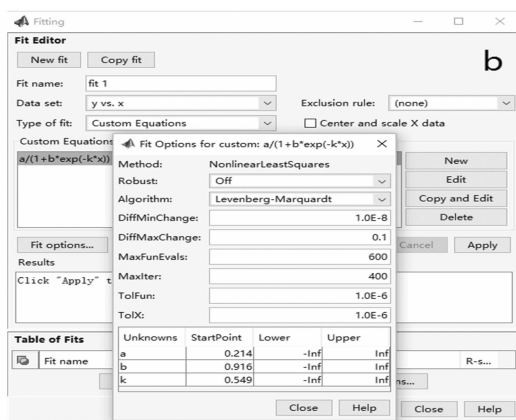


图2 Matlab R2010b 对生长曲线拟合操作截图

2 结果与分析

2.1 3 种方法绘制的不同浓度 Cd(Ⅱ) 下中肋骨条藻的 S 形生长曲线

由图 3 至图 5 可得,(1)图 3 中用 Excel 作出的曲线只能是连接各试验数据点的平滑曲线,无法拟合作图;(2)图 5 中 Matlab 所得曲线与各数据点的吻合程度明显高于图 4,说明 Matlab 对生长曲线的拟合较优。

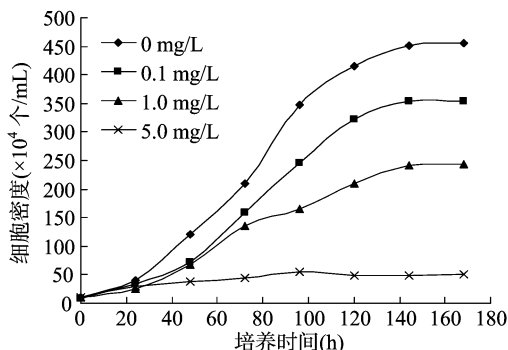


图3 Excel 对中肋骨条藻 Logistic 生长曲线的拟合

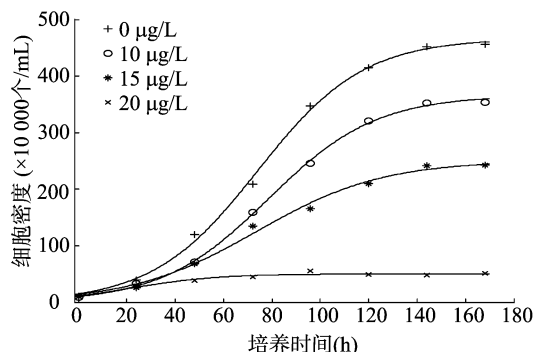


图5 Matlab 对中肋骨条藻 Logistic 生长曲线的拟合

于 Origin 的单个迭代。

3 结论与讨论

Logistic 曲线作为一种广义的线性回归模型,它的拟合方法可分为线性和非线性法。本研究 3 种软件中,Excel 对应的是 Logistic 曲线的线性方法拟合;Origin 和 Matlab 对应的是曲线的非线性回归拟合,其中 Origin 采用单次迭代,而 Matlab 是多次迭代。曲线拟合精度的高低都可用 R^2 的大小来判定。由试验结果可知,3 种软件对 Logistic 的拟合精度均较高, R^2 值均接近于 1[表 4 显示 Cd(Ⅱ) 浓度较大时,对微藻的生长会产生较大影响, R^2 值减小],这表明 3 种软件所获得的 Logistic 回归方程均与实测数据拟合很好。同时结果提示,(1)Excel 可以对 Logistic 曲线模型的 3 个参数进行估计,但不能作出拟合曲线图,只能得到平滑曲线图。(2)Excel 进行拟合时,四点法各参数估计量与非线性回归最为接近,因此本研究中四点法比三点法拟合效果优。这是由于本研究数据点数为 8 个,通常四点法较适合于试验数据序列数为偶数的情况,三点法适用于奇数的情况,这与以往报道^[13]一致,但三点法具有简单快速的优势。(3)Excel 拟合中所采用的线性拟合方法虽然简便,但对 Logistic 函数进行线性变换后会对估计参数的性质产生影响,如不再具有无偏性;所以其拟合精度不如 Origin 和 Matlab 所采用的非线性拟合。(4)在非线形拟合中,多次迭代会提高 R^2 ,因此 Matlab 的拟合效果优于 Origin。(5)综合数据可得,不同拟合方法拟合后的 R^2 大小顺序为非

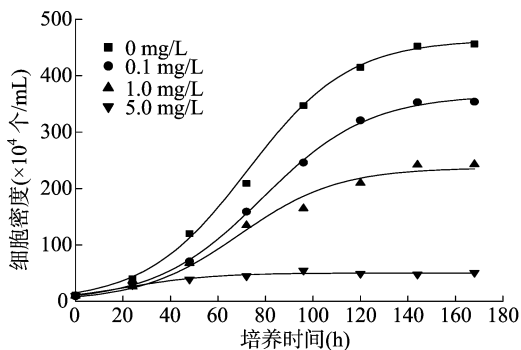


图4 Origin 对中肋骨条藻 Logistic 生长曲线的拟合

2.2 3 种拟合方法对微藻生长曲线的拟合效果比较

由表 1 至表 4 可得,3 种软件拟合后的 R^2 值整体大小顺序为 Matlab > Origin > Excel,其中 Excel 方法中四点法 R^2 值均大于三点法。数据同时显示,四点法的参数值与非线性方法更为接近。在非线形拟合中,Matlab 多次迭代的 R^2 值均高

表 1 3 种拟合方法对微藻生长曲线一[Cd(Ⅱ) 浓度为 0 mg/L] 的拟合效果比较

拟合软件	回归方法	<i>a</i>	<i>b</i>	<i>k</i>	<i>R</i> ²
Excel	三点线性化	468.78 ± 4.10	30.88 ± 0.61	0.044 7 ± 0.005 4	0.989 3
	四点线性化	460.38 ± 3.76	37.88 ± 0.54	0.050 2 ± 0.002 6	0.995 3
Origin	非线性化单次迭代	464.09 ± 5.01	29.35 ± 0.34	0.046 7 ± 0.004 1	0.996 7
Matlab	非线性化多次迭代	466.70 ± 4.12	29.89 ± 0.35	0.045 8 ± 0.001 9	0.998 4

表 2 3 种拟合方法对微藻生长曲线二[Cd(Ⅱ) 浓度为 0.1 mg/L] 的拟合效果比较

拟合软件	回归方法	<i>a</i>	<i>b</i>	<i>k</i>	<i>R</i> ²
Excel	三点线性化	368.32 ± 4.17	31.52 ± 0.94	0.042 9 ± 0.012 9	0.984 7
	四点线性化	359.28 ± 3.91	38.49 ± 0.91	0.048 7 ± 0.002 4	0.988 5
Origin	非线性化单次迭代	366.01 ± 3.85	33.91 ± 0.88	0.044 8 ± 0.003 9	0.998 7
Matlab	非线性化多次迭代	366.10 ± 3.92	34.59 ± 0.81	0.045 1 ± 0.002 1	0.999 1

表 3 3 种拟合方法对微藻生长曲线三[Cd(Ⅱ) 浓度为 1.0 mg/L] 的拟合效果比较

拟合软件	回归方法	<i>a</i>	<i>b</i>	<i>k</i>	<i>R</i> ²
Excel	三点线性化	249.24 ± 3.01	22.28 ± 0.67	0.042 0 ± 0.003 7	0.983 3
	四点线性化	247.24 ± 2.99	24.06 ± 0.71	0.044 1 ± 0.003 2	0.990 5
Origin	非线性化单次迭代	236.89 ± 2.01	28.53 ± 0.77	0.048 6 ± 0.003 3	0.981 3
Matlab	非线性化多次迭代	250.50 ± 2.49	16.84 ± 0.53	0.034 0 ± 0.021 0	0.992 9

表 4 3 种拟合方法对微藻生长曲线四[Cd(Ⅱ) 浓度为 5.0 mg/L] 的拟合效果比较

拟合软件	回归方法	<i>a</i>	<i>b</i>	<i>k</i>	<i>R</i> ²
Excel	三点线性化	50.88 ± 1.72	1.86 ± 0.10	0.029 2 ± 0.003 3	0.875 7
	四点线性化	50.30 ± 1.67	2.18 ± 0.14	0.031 7 ± 0.010 4	0.908 1
Origin	非线性化单次迭代	50.29 ± 1.70	3.34 ± 0.13	0.055 23 ± 0.011 5	0.951 1
Matlab	非线性化多次迭代	50.35 ± 1.88	3.34 ± 0.17	0.055 02 ± 0.002 9	0.965 0

线性 > 四点线性 > 三点线性, 3 种软件拟合后的 *R*² 值为 Matlab > Origin > Excel, 其中, Matlab 采用 Levenberg - Marquardt 算法进行非线性化多次迭代拟合, 提示 Matlab 是 3 种软件中对 Logistic 生长曲线拟合的最佳方法。

Logistic 曲线拟合的优度检验除了 *R*², 还有如卡方值、标准偏差等其他统计指标^[14], 除了本研究的 3 种软件, 对 Logistic 的拟合还可借助其他各种计算机语言。因此各种不同软件、方法对 Logistic 拟合的比较还有更多的探讨空间。

参考文献:

[1] Verhulst P J. Notice sur la loi que la population suit dans sons accroissement[J]. Corr Math et Physique,1838,10:113 - 121.
[2] Pearl R, Reed L J. The Logistic Curve and the Census Count of 1930 [J]. Science,1930,72(1868):399 - 401.
[3] Fisk P R. The Graduation of income distributions[J]. Econometrica, 1961,29(2):171 - 185.
[4] Anderson J A. Sepatate sample Logistic discrimination [J]. Biometrika,1972,59:19 - 35.
[5] 王福林,王吉权. 生长曲线参数估计的一种新方法——优化回归组合法[J]. 生物数学学报,2007,22(3):533 - 538.
[6] 陈冬,程维虎. 利用样本分位数的 Logistic 总体分布参数的近似最佳线性无偏估计[J]. 应用数学学报,2005,28(2):325 - 332.

[7] 许小勇,钟太勇. 三次样条插值函数的构造与 Matlab 实现[J]. 兵工自动化,2006,25(11):76 - 78.
[8] Zhang L J, Song Z J, Pan X D, et al. Comparison of v - support vector regression and Logistic equation for descriptive modeling of *Lactobacillus plantarum* growth [J]. African Journal of Biotechnology,2011,10(32):6162 - 6171.
[9] 栾红艳,赵卫红,苗辉. Cd²⁺胁迫下中肋骨条藻细胞内多胺的生理响应[J]. 中国环境科学,2015,35(5):1487 - 1494.
[10] Torres E, Mera R, Herrero C, et al. Isotherm studies for the determination of Cd(Ⅱ) ions removal capacity in living biomass of a microalga with high tolerance to cadmium toxicity [J]. Environmental Science & Pollution Research, 2014, 21 (22): 12616 - 12628.
[11] 殷祚云. Logistic 曲线拟合方法研究[J]. 数理统计与管理, 2002,1(21):41 - 46.
[12] 胡文冉,范玲,田晓莉,等. Excel 在 Logistic 曲线拟合中的应用[J]. 农业网络信息,2013(3):14 - 16.
[13] 范国兵. 一种估计 Logistic 模型参数的方法及应用实例[J]. 经济数学,2010,27(1):105 - 110.
[14] 梁玉麟,赵建东. 基于 logistic 方程的供应商 - 采购商合作关系的经济模型研究[J]. 鲁东大学学报(自然科学版),2015,31(3):200 - 204,213.