

预测微生物模型的评价方法

王 军^{1,2}, 董庆利³, 丁 甜^{2,*}

(1. 临沂大学生命科学学院, 山东 临沂 276005; 2. 国立江原大学食品科学与生物技术学院, 韩国 春川 200701;
3. 上海理工大学食品科学与工程研究所, 上海 200093)

摘 要: 在预测微生物模型构建过程中, 模型验证是必不可少的步骤, 其可靠性和适用性得到合理评价后才具有应用价值。本文综述预测微生物模型的常用评价方法及对应的参数, 着重归纳总结相关评价参数的基本概念及意义, 并结合前期工作的论文数据阐述演示各常用评价参数的应用方法及评价依据。

关键词: 预测微生物; 预测模型; 评价; 验证

An Overview of Evaluation Methods for Predictive Microbial Growth Models

WANG Jun^{1,2}, DONG Qing-li³, DING Tian^{2,*}

(1. College of Life Science, Linyi University, Linyi 276005, China;
2. Department of Food Science and Biotechnology, Kangwon National University, Chuncheon 200701, Korea;
3. Institute of Food Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: In predictive microbial growth modeling, it is necessary to validate the reliability and applicability of developed models before applications. The major objective of this review is to outline common evaluations methods for predictive microbial growth models with emphasis on the concept and role of 10 related evaluation parameters. Additionally, the square root and Ratkowsky models which has been proposed in our previous work were validated by the internal validation and external validation methods based on the evaluation parameters described in this review.

Key words: predictive microbiology; predictive model; evaluation; validation

中图分类号: TS201.3

文献标识码: A

文章编号: 1002-6630(2011)21-0268-05

预测微生物学(predictive microbiology)是一门结合微生物学、化学、数学、统计学和应用计算机技术的交叉性学科, 它采用数学的方法描述不同环境条件下, 细菌数变化和外部环境因素之间的响应关系, 并对微生物的生长动力学做出预测。微生物预测模型按照 Whiting 和 Buchanan 的分类方法分为一级模型、二级模型和三级模型, 一级模型主要表达一定生长条件下微生物生长、失活与时间之间的函数关系; 二级模型主要表达通过一级模型得到的参数与环境因子变量之间的函数关系; 三级模型是主要指建立在一级和二级模型基础上的计算机软件程序, 可以预测相同条件或不同环境条件下同一种微生物的生长或失活情况^[1-3]。预测微生物模型的建立主要有以下几个步骤, 收集实验数据、一级和二级模型的建立及模型验证。一般来说, 预测模型构建以后需要采用一定的方法进行拟合度的评估, 并且在应

用之前需要经过验证来对该模型的可靠性作出评价。本文结合前人所做的工作, 介绍已发表论文中常用的预测微生物模型评价方法及其应用, 并结合相关实验数据使用不同的参数对所建立的微生物预测模型进行评价, 更直观的阐述和理解各种不同的预测模型的评价方法。

丁甜等^[4]研究了不同温度下单增李斯特菌在营养肉汤中的生长动力学预测模型的建立和验证, 结果表明 Ratkowsky 模型对实验数据的拟合度最高。本文结合该文献中已发表的实验数据及预测模型, 来介绍常用预测微生物模型评价方法的应用, 以期对预测微生物模型的验证方法进行详细的介绍和说明, 具有一定的指导意义。

1 预测微生物模型的评价方法

完成预测微生物建模工作以后, 需要对已经建立好

收稿日期: 2010-12-21

作者简介: 王军(1980—), 男, 博士研究生, 研究方向为食品微生物风险评估。E-mail: wangjun@kangwon.ac.kr

* 通信作者: 丁甜(1985—), 男, 博士研究生, 研究方向为食品微生物风险评估。E-mail: tding@kangwon.ac.kr

的微生物预测模型进行检验和验证,评价所构建模型准确无偏差描述实验观测数据能力,衡量预测模型的适用性和可靠性。预测微生物模型的验证是将实际的实验数据或者他人文献中的数据带入到已经建立的模型中,将得到的预测值和实际数值进行比较,判断所得模型的可靠性和适用性。模型的验证有两种形式,内部验证和外部验证。内部验证(internal validation)或自我验证(self validation)是指采用建立模型时的实验数据和模型预测值进行比较来验证模型的适用性和可靠性。外部验证(external validation)是有两种情况,一种是指额外做一些类似建立模型时环境条件的实验,采用新的实验数据进行模型验证;另外一种情况是引用参考文献中具有相同或相似实验环境的实验数据,并将其带入到所建预测模型中,根据相关评价参数来评估预测模型的适用性和可靠性^[5]。一般情况下,预测模型内部验证的结果要好于外部验证的结果,但是外部验证的结果更具有说服力,更能显示出所建预测模型的适用性和可靠性。预测模型适用性和可靠性的评价可以通过图表或者基于数学或统计学的参数完成。

1.1 图形分析比较(graphical comparison)

图形分析比较是模型验证中常用的一个方法,它可以通过图表很直观的反映预测值是否在合理的范围之内^[6],并且通过对代表预测值的散点进行线性回归分析,判断模型预测值可以解释实验观测值的程度。

1.2 数学或统计学的比较

除了常用的图形分析比较法以外,通常还采用基于数学或统计学参数的方法进行模型评价。评价时通常采用以下参数。

1.2.1 相关系数

相关系数 R (式(1))衡量两个变量线性相关密切程度的量。用于模型验证时,可以衡量模型预测值和实验观测值之间线性密切程度。

$$R = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (1)$$

相关系数 R 的值在-1和+1之间(负号代表两个变量呈负相关),当 $0.8 \leq R \leq 1$ 或 $-1 \leq R \leq -0.8$ 时,说明实际值和预测值具有很强的相关性;当 R 值介于 $0.5 \sim 0.8$ 或 $-0.8 \sim -0.5$ 之间时,说明实际值和预测值具有一般的相关性;当 $-0.5 \leq R \leq 0.5$ 时,说明实际值和预测值之间相关性较差^[7]。

1.2.2 决定系数

决定系数 R^2 (式(2))表示的在因变量 Y 的总平方和中,由自变量 X 引起的平方和所占的比例。决定系数一般用来对预测模型拟合程度做一个总的评价^[8]。决定

系数的大小决定了自变量与因变量相关的密切程度,即决定了模型预测值和实验数据的相关程度。 R^2 的值在 $0 \sim 1$ 之间,当 R^2 越接近1时,表示相关的预测模型参考价值越高;相反,越接近0时,表示该模型参考价值越低^[9]。

$$R^2 = 1 - \frac{SSE}{SST} \quad (2)$$

式中: SSE 代表误差平方和; SST 代表总离差平方和。

1.2.3 修正决定系数

修正决定系数 R^2_{Adj} 是在决定系数 R^2 的基础上考虑了样本大小和变量参数数目对模型的影响,避免了因变量参数数目增加引起的 R^2 值增大,而模型拟合度并没有得到提高的情况所带来的负面影响^[10]。 R^2_{Adj} 可以用来评价预测模型的拟合程度, R^2_{Adj} 的取值范围小于或等于1, R^2_{Adj} 愈接近1,预测模型对实验数据的拟合度愈高^[11]。当模型包含的变量参数对模型预测没有帮助的时候,有可能出现负值。

$$R^2_{Adj} = 1 - \frac{(1 - R^2)(n - 1)}{n - N - 1} \quad (3)$$

式中: n 代表观测值的个数; N 代表预测模型中变量参数的个数。

1.2.4 均方根误差(root mean square error, RMSE)

$$RMSE = \sqrt{\frac{\sum (\mu_{observed} - \mu_{predicted})^2}{n}} \quad (4)$$

式中: $\mu_{observed}$ 代表实验观测值; $\mu_{predicted}$ 代表预测值; n 代表观测值个数。

均方根误差可作为衡量预测准确度的一种数值指标,可以说明模型预测值的离散程度^[12]。

1.2.5 均方误差(mean square error, MSE)

对于同一组实验数据,可以建立多个预测模型,得到不同的模型预测值,而平均误差的大小通常用来评价预测模型的优劣。 MSE 是衡量“平均误差”的一种较方便的方法, MSE 可以评价数据的变化程度, MSE 的值越小,说明预测模型描述实验数据具有更好的精确度^[13]。

$$MSE = \frac{\sum (\mu_{observed} - \mu_{predicted})^2}{n} \quad (5)$$

1.2.6 准确因子(accuracy factor, A_i)和偏差因子(bias factor, B_i)

精确因子 A_i (式(6))一般用来验证预测模型的准确度。 A_i 代表了每一个预测值的点与等值线之间的平均距

离, 可以衡量预测值和观测值之间的接近程度。如果是一个完美的预测模型, 则 $A_f = 1$, 表明所有的预测值和观测值均相等^[14]; A_f 值越大, 表明该模型预测的平均精确度越低^[15]。

$$A_f = 10^{\left(\frac{\sum \lg(\mu_{\text{predicted}} / \mu_{\text{observed}})}{n}\right)} \quad (6)$$

偏差因子 B_f (式(7))一般作为判断预测模型偏差度的参数。 B_f 用来判断预测值在等值线的上方还是下方以及评价预测值偏离等值线的程度。它表明了所建预测模型的结构性偏差。对于微生物代时(generation time, GT), 如果 $B_f < 1$, 说明该模型是失效保护模型, 即 GT 预测值平均比观测值要小, 这样预测模型可以给出一个安全的预测界限^[5,15]。

$$B_f = 10^{\left(\frac{\sum \lg(\mu_{\text{predicted}} / \mu_{\text{observed}})}{n}\right)} \quad (7)$$

式中: A_f 和 B_f 是预测模型验证中比较常用的两个参数, 通常用于预测模型的外部验证。

1.2.7 中位数相对误差 (median relative error, MRE)

MRE 是指各个预测值和实验观测值相对误差的中位数, 相对误差(relative error, RE)的计算公式如下所示:

$$RE = 10 \frac{\mu_{\text{predicted}} - \mu_{\text{observed}}}{\mu_{\text{observed}}} \quad (8)$$

相对误差则是绝对预测误差与实验观测值的比值, 一般来说, 相对误差更能反映预测值的可信程度, 而 MRE 在一定程度上可以反映预测模型的准确度。MRE 越接近于 0, 说明预测模型的可信程度越高^[16]。MRE 的不足之处就是对没有典型性的异类预测值或极端预测值不敏感。

1.2.8 平均相对误差绝对值(mean absolute relative error, MARE)

平均相对误差绝对值 MARE(式(9))一般用来验证预测模型的偏差度^[17], MARE 值越小, 说明预测模型的预测偏差越小^[18], 同时 MARE 避免了类似 MRE 对极端预测值不敏感的不足。

$$MARE = \frac{1}{n} \sum_{i=1}^n |RE_i| \quad (9)$$

1.2.9 预测标准误差 (standard error of prediction, SEP)

预测标准误差 SEP(式(10))是指预测模型预测值和实验观测值差异的标准偏差, 它可以衡量和验证预测模型的准确度, SEP 越小, 说明预测模型能够更好的描述实验数据^[19]。

$$SEP/\% = \frac{100}{\bar{\mu}_{\text{observed}}} \sqrt{\frac{\sum (\mu_{\text{observed}} - \mu_{\text{predicted}})^2}{n}} \quad (10)$$

2 评价方法的应用

在评价所构建的预测模型时, 需要最大限度的衡量预测模型的适用性和可靠性, 而不是采用尽可能多的评价参数对模型进行一一验证。丁甜等^[4]在研究中是以内部验证的方法采用图形分析比较法及 R^2 、 R^2_{Adj} 、RMSE 等参数对模型进行验证的。

ComBase(Combined Database, www.combase.cc)是英国食品标准局(Food Standards Agency)、英国食品研究所(Institute of Food Research)、美国农业部农业研究服务中心(USDA Agricultural Research Service)及其下属东部地区研究中心(Eastern Regional Research Center)和澳大利亚食品安全中心(Food Safety Centre)共建和开发的预测微生物数据库。ComBase 系统包含了微生物如何响应不同的环境条件, 它不仅能预测同一种微生物在恒定环境条件下的生长情况, 还能预测一种微生物在不同环境中的生长情况, 并且可以对这些不同情况下的生长情况进行比较和分析。Growth Predictor (version 1.0)是 ComBase 网站提供的一款免费的微生物生长预测软件, 它提供了不同生长条件下各种常见食源性病原菌在营养肉汤中的生长情况与时间之间的函数关系。为了介绍常用的预测微生物模型评价方法的应用, 说明内部验证和外部验证的关系, 本实验使用 Growth Predictor 选择单增李斯特菌在 4、10、15、20、25、30、35 °C 条件下生长的实验数据, 对文献[4]中所建立的 Square root 模型和 Ratkowsky 模型进行验证, 并将本文提到的评价参数进行比较, 以更好的探讨各种验证参数的应用情况和其意义。验证结果如图 1、2 和表 1、2 所示。图 1、2 展示了内部验证和外部验证时, 实验观测值与 Square root 和 Ratkowsky 模型预测值的比较情况, 结果表明不论是内部验证还是外部验证, Ratkowsky 模型要比 Square root 模型具有更好的预测能力。

通过表 1、2 的对比可以看到, 不论是内部验证还是外部验证, Ratkowsky 模型的 R 、 R^2 和 R^2_{Adj} 值均大于 Square root 模型; MRE、MARE、MSE、RMSE 和 SEP 均小于 Square root 模型。以上结果说明了 Ratkowsky 预测模型具有很好的适用性和可靠性, 要强于 Square root 模型。在内部验证结果中还可以看到 Ratkowsky 模型的 $A_f = 1.0556$, A_f 值可接受的范围与环境参数的个数有关, 每增加一个参数变量, A_f 值可接受的范围增加 0.10~0.15 个单位^[20]。Ratkowsky 模型和 Square root 模型中的参数变量只有一个,

表 1 单增李斯特菌生长动力学预测模型内部验证的参数

Table 1 Internal validation parameters for Square root model and Ratkowsky model describing the growth of *L. monocytogenes* in broth medium

参数	R	R^2	R^2_{Adj}	MRE	MARE	MSE	RMSE	SEP/%	A_f	B_f
Square root 模型	0.9886	0.9774	0.9729	-0.0143	0.0207	0.000542	0.0233	10.1587	1.2849	0.8283
Ratkowsky 模型	0.9998	0.9997	0.9996	0.0045	0.0026	0.000008	0.0028	1.2121	1.0556	0.9677

表 2 单增李斯特菌生长动力学预测模型外部验证的参数

Table 2 External validation parameters for Square root model and Ratkowsky model describing the growth of *L. monocytogenes* in broth medium

参数	R	R^2	R^2_{Adj}	MRE	MARE	MSE	RMSE	SEP/%	A_f	B_f
Square root 模型	0.9887	0.9775	0.973	0.1746	0.078	0.0094	0.0968	33.7123	1.5074	0.9256
Ratkowsky 模型	0.9951	0.9902	0.9882	0.1111	0.0665	0.0085	0.092	32.0455	1.2354	1.0814

也就是说 A_f 值在 1.0~1.15 的范围内是可接受的。表 1 中内部验证的数据结果显示, Ratkowsky 模型比 Square root 模型具有更高的预测精确度, 平均只有 5.56% 的预测值与观测值不同。同样, 表 2 中外部验证的数据结果显示, Ratkowsky 模型比 Square root 模型具有更高的预测精确度, 平均有 23.54% 的预测值与观测值不同。前面提到外部验证的结果要逊色于内部验证, 但其结果更具说服力。虽然 $1.2354 > 1.15$, 但是结合其他评价参数, 认为 Ratkowsky 模型在预测能力和准确度方面要强于 Square root 模型。 B_f 值在 0.9~1.05 的范围内较好, 说明预测模型在预测时具有较小的偏差; B_f 值在 0.7~0.9 或 1.06~1.15 的范围内是可接受的; 而如果 $B_f < 0.7$ 或 $B_f > 1.15$ 是不可接受的, 这说明预测模型预测时偏差较大, 是不可信的^[15]。与 A_f 的情况类似, 表 1 和表 2 的结果同样显示 Ratkowsky 模型比 Square root 模型具有更高的预测能力、适用性和可靠性。

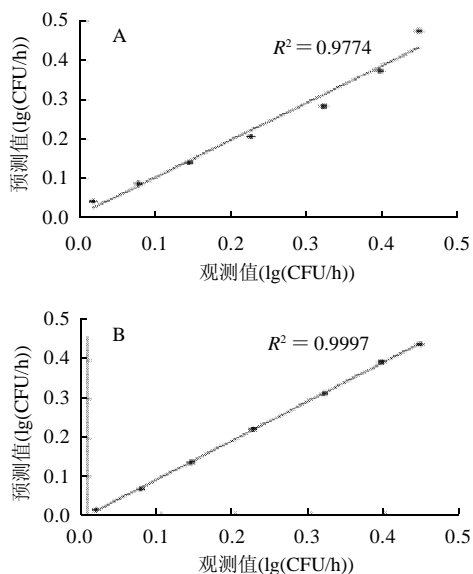


图 1 内部验证时, 实验观测值与 Square root (A) 和 Ratkowsky (B) 模型预测值的比较

Fig.1 Observed values versus predicted values obtained from Square root model (A) and Ratkowsky model (B) for internal validation

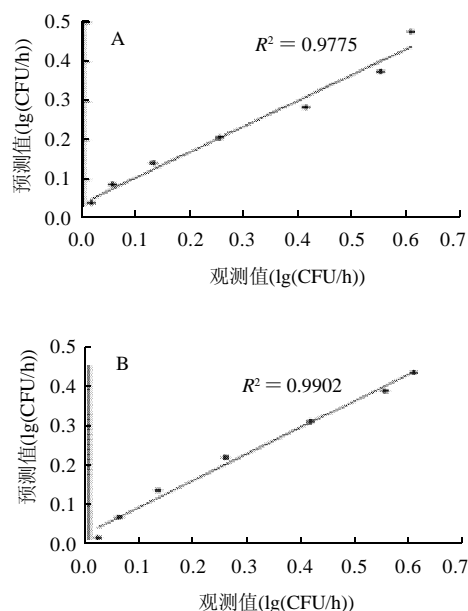


图 2 外部验证时, 实验观测值与 Square root (A) 和 Ratkowsky (B) 模型预测值的比较

Fig.2 Observed values versus predicted values obtained from Square root model (A) and Ratkowsky model (B) for external validation

3 结 语

综上所述, 由于环境因素的多变, 使得环境因素与微生物生长动力学之间的关系更复杂。所以, 加强对建立预测微生物模型后的评价工作的研究显得尤为重要。对于微生物预测模型的建立, 模型评价是在模型应用前十分必要且关键的一步。模型评价方法和参数有很多种, 我们应该选择恰当的方法和参数来评价模型的拟合度及验证模型, 以最大限度的衡量所建立模型的适用性和可靠性, 使其具有更高的可信度。

参考文献:

- [1] DING Tian, SHIM Y H, CHOI N J, et al. Mathematical modeling on the

- growth of *Staphylococcus aureus* in sandwich[J]. Food Science Biotechnology, 2010, 19(3): 763-768.
- [2] SWINNEN I A M, BERNAERTS K, DENS E J J, et al. Predictive modelling of the microbial lag phase: a review[J]. International Journal of Food Microbiology, 2004, 94(2): 137-159.
- [3] FERRER J, PRATS C, LOPEZ D, et al. Mathematical modelling methodologies in predictive food microbiology: a SWOT analysis[J]. International Journal of Food Microbiology, 2009, 134(1/2): 2-8.
- [4] 丁甜, 董庆利, 王璐, 等. 单增李斯特菌在营养肉汤中最大比生长速率预测模型的建立与验证[J]. 华中农业大学学报, 2010, 29(4): 71-75.
- [5] TE GIFFEL M C, ZWIETERING M H. Validation of predictive models describing the growth of *Listeria monocytogenes*[J]. International Journal of Food Microbiology, 1999, 46(2): 135-149.
- [6] HEO S K, LEE H S, HA S D. A predictive model for the growth rate of *Bacillus cereus* in broth by response surface methodology[J]. Biotechnology and Bioengineering, 2009, 14(2): 202-206.
- [7] KARUNAKARAN K, THAMILARASU P, SHARMILA R. Statistical study on physicochemical characteristics of groundwater in and around Namakkal, Tamilnadu, India[J]. E Journal of Chemistry, 2009, 6(3): 909-914.
- [8] DING Tian, RAHMAN S M E, PUREV U, et al. Modelling of *Escherichia coli* O157:H7 growth at various storage temperatures on beef treated with electrolyzed oxidizing water[J]. Journal of Food Engineering, 2010, 97(4): 497-503.
- [9] STEELE R. Understanding and measuring the shelf-life of food[M]. Cambridge: Woodhead Publishing, 2004: 189-206.
- [10] SNEDECOR G W, COCHRAN W G. Statistical methods[M]. 8th ed. Iowa: Iowa University Press, 1989: 360-362.
- [11] DAVEY K R, DAUGHTRY B J. Validation of a model for predicting the combined effect of three environmental factors on both exponential and lag phases of bacterial growth: temperature, salt concentration and pH[J]. Food Research International, 1995, 28(3): 233-237.
- [12] HUANG Lihan. Thermal inactivation of *Listeria monocytogenes* in ground beef under isothermal and dynamic temperature conditions[J]. Journal of Food Engineering, 2009, 90(3): 380-387.
- [13] CAYRE M E, VIGNOLO G, GARRO O. Modeling lactic acid bacteria growth in vacuum-packaged cooked meat emulsions stored at three temperatures[J]. Food Microbiology, 2003, 20(5): 561-566.
- [14] LEBERT I, ROBLES-OLVERA V, LEBERT A. Application of polynomial models to predict growth of mixed cultures of *Pseudomonas* spp. and *Listeria* in meat[J]. International Journal of Food Microbiology, 2000, 61(1): 27-39.
- [15] ROSS T. Indices for performance evaluation of predictive models in food microbiology[J]. Journal of Applied Bacteriology, 1996, 81(5): 501-508.
- [16] OSCAR T P. Response surface models for effects of temperature, pH, and previous growth pH on growth kinetics of *Salmonella typhimurium* in brain heart infusion broth[J]. Journal of Food Protection, 1999, 62(2): 106-111.
- [17] BANG W S, CHUNG H J, JIN S S, et al. Prediction of *Listeria monocytogenes* growth kinetics in sausages formulated with antimicrobials as a function of temperature and concentrations[J]. Food Science and Biotechnology, 2008, 17(6): 1316-1321.
- [18] OSCAR T P. Response surface models for effects of temperature and previous temperature on lag time and specific growth rate of *Salmonella typhimurium* on cooked ground chicken breast[J]. Journal of Food Protection, 1999, 62(10): 1111-1114.
- [19] DONG Qingli, TU Kang, GUO Liyang, et al. Response surface model for prediction of growth parameters from spores of *Clostridium sporogenes* under different experimental conditions[J]. Food Microbiology, 2007, 24(6): 624-632.
- [20] ROSS T, DALGAARD P, TIENUNGOON S. Predictive modelling of the growth and survival of *Listeria* in fishery products[J]. International Journal of Food Microbiology, 2000, 62(3): 231-245.