

试验设计和优化及其在发酵培养基优化中的应用

代志凯^{1,2} 张翠^{1,2} 阮征^{1,2*}

(1. 南昌大学 食品科学与技术国家重点实验室 江西 南昌 330047)

(2. 南昌大学 生命科学与食品工程学院 江西 南昌 330031)

摘要: 在发酵工业中, 发酵培养基的优化对发酵水平的提高起着举足轻重的作用。而在寻求最佳发酵培养基的过程中, 试验设计及统计优化发挥着重要作用。对近年来常用的试验设计及优化方法进行了综述, 内容包括单因素轮换法、析因设计、均匀设计、响应面设计、人工神经网络和遗传算法等多种试验设计和优化方法, 并对其进行分析和比较。

关键词: 培养基, 优化, 试验设计, 神经网络, 遗传算法

The Application of Experimental Design and Optimization Techniques in Optimization of Microbial Medium

DAI Zhi-Kai^{1,2} ZHANG Cui^{1,2} RUAN Zheng^{1,2*}

(1. State Key Laboratory of Food Science and Technology, Nanchang, Jiangxi 330047, China)

(2. College of Life Science and Food Engineering, Nanchang University, Nanchang, Jiangxi 330031, China)

Abstract: The optimization of fermentation medium is important for improving microbial production or quality in fermentation industry. Experimental design and optimization techniques play an important role for optimizing medium. In this paper, we reviewed the methodologies and techniques used in optimization of microbial medium, including one-factor-at-a-time design, factorial design, uniform design, response surface optimization (RSM), artificial neural network (ANN) and genetic algorithms (GA). Meanwhile, the methodologies and techniques were analyzed and evaluated comprehensively.

Keywords: Fermentation medium, Optimization, Experimental design, Artificial neural network, Genetic algorithm

现代发酵工业已发展成为一个重要的产业体系, 涉及食品、医药、化工等领域, 在国民经济中占有重要的地位, 而发酵培养基的优化在微生物发酵产业化生产中举足轻重, 是从实验室到工业生产的必要环节。由于微生物发酵过程是一个高度非线性、非结构化的复杂系统, 要建立一个准确适合

的发酵模型存在许多困难。因此选择适当试验设计和优化方法, 设计合理的培养基优化方案就显得尤为重要。

目前国内外关于培养基优化的文献仅仅局限于一种或少数几种试验设计和优化方法^[1-5], 鉴于培养基优化的特殊性, 仅仅依靠单一方法很难取得满

基金项目: 南昌大学“赣江学者奖励计划”项目; 中国博士后科学基金资助项目(No. 20080440166); 食品科学与技术国家重点实验室探索项目(No. SKLF-TS-200817)

* 通讯作者: Tel: 86-791-8304447-8325; ✉ ezruan@yahoo.com

收稿日期: 2009-11-11; 接受日期: 2010-03-08

© 中国科学院微生物研究所期刊联合编辑部 <http://journals.im.ac.cn>

意的效果。而目前介绍培养基优化策略的文献综述往往局限于传统的优化方法,对最近流行的神经网络和遗传算法在培养基优化中的应用介绍不多^[6-8],只是粗略的涉及。因此为了方便广大科技工作者了解各种培养基优化方法的优缺点和适用范围,更加方便、快捷的进行相关研究,本文对培养基优化中常见的试验设计和统计优化方法进行了分析和比较。

1 单次单因子设计

实验室最常用的优化方法是单次单因子法(One-factor-at-a-time design,简称“单因素法”),这种方法是在假设因素间不存在交互作用的前提下,通过一次改变一个因素的水平而其它因素保持不变,然后逐个因素进行考察。这种方法的优点是简单、容易,结果很明了,不需要专门的统计分析,因此一直都是发酵培养基组分优化的常用方法^[9-21]。

但是由于微生物培养基成分之间经常存在交互作用,使得该方法并非总能获得最佳的优化配比。另外,当考察的因素较多时,需要太多的试验次数和较长的试验周期^[22]。因此单因素法在培养基优化过程中常常用于初步的筛选试验,且主要用于筛选最佳碳源、氮源、无机盐和金属离子等营养成分,筛选出主要因素后,再利用正交、均匀或者响应面设计对主要因素着重考察。陶申傲等^[9]利用单因素从众多碳氮源中筛选到蔗糖、玉米淀粉、麸皮和脱脂蛋白粉 4 种成分是榆耳发酵培养基的最佳碳氮源,然后利用均匀设计对这 4 种成分的含量进一步优化。左爱连等^[10]筛选出葡萄糖, $(\text{NH}_4)_2\text{SO}_4$ 和玉米浆为主要影响因素后利用的是响应面设计, Xu 等^[11]利用的是正交设计。

2 析因设计

析因设计(Factorial design)是一种多因素多水平交叉分组的试验设计,它可以研究两个或两个以上因素的主效应以及各因素的交互作用,而且通过比较各种组合,可以找到最佳组合。因此在微生物发酵培养基的优化中发挥着重要的作用。析因设计分全因子和部分因子设计两种。

2.1 全因子设计

全因子设计(Full factorial design)中各因素的不

同水平间的各种组合都将被试验。其优点是所获得的信息量很多,可以准确地估计各试验因素的主效应的大小,还可估计因素之间各级交互作用效应的大小。缺点是所需要的试验次数较多,因此当所考察的试验因素和水平较多时,研究者很难承受。所以一般利用全因子设计对培养基进行优化试验都为 2 水平,是能反映因素间交互作用的最小设计,其效率比单次单因子试验高。

2.2 部分因子设计

当全因子设计因所需试验次数多而不可行时,从全部水平组合中抽取部分有代表性的水平组合来作试验,这种方法称为部分因子设计(Fractional factorial design, FFD)也称全因子设计的部分实施。在微生物培养基的优化中 2 水平的部分因子设计用的最多,主要用来从众多因素中筛选主要因素着重考察。与全因子试验设计相比,部分因子设计在保留主要信息的前提下,大大地减少了试验次数,并且可以根据试验数据拟合出一次多项式,并确定主要影响因素。此外在 2 水平部分因子设计中还可以适当的添加中心点以估计误差和对弯曲性的保护(利用 2 水平析因设计时,假定因素效应是线性的)。除了 2 水平部分因子设计外,我们常见的正交设计、均匀设计、Plackett-Burman 设计(简称 PB 设计)也都属于部分因子设计,这些设计在培养基优化过程中使用得较为广泛。

2.2.1 正交设计:正交试验设计(Orthogonal experimental design)是从“均匀分散、整齐可比”的角度出发,是以拉丁方理论和群论为基础,从全面试验中挑选出部分有代表性的点进行试验。利用正交表可以来安排多因素试验,相对于全面试验而言,它只是部分试验,但是却可用较少的试验获得能基本上反映全面情况的试验资料,克服了单次单因子试验不能全面考察因素的影响与因素之间交互效应的缺点^[23],是一种高效率、快速、经济的试验设计方法。

2.2.2 均匀设计:均匀设计(Uniform design)是 1978 年由我国科学家方开泰和王元共同提出的,也是一种用设计好的表格安排试验的方法。相比正交设计均匀分散和整齐可比的特性,均匀设计只是考虑试验点在试验范围内的均匀分散性,而忽略掉整齐可比性的一种试验设计方法。它的优点是当因素数目

较多时所需要的试验次数也不多。但由于均匀设计舍弃了整齐可比性,故不能像正交设计那样用方差分析,而只能用回归分析,因此在数据处理上比较麻烦,还好在计算机高度普及的今天已不是件很难的事情。DPS (Data processing system)、SAS (Statistical analysis system)、SPSS (Statistical product and service solutions)等软件均能做回归分析。由于均匀设计每个因素每个水平只做 1 次试验,故当试验条件不易控制时,不宜使用均匀设计,对波动相对较大的微生物培养试验,每一实验组最好重复 2-3 次,以便估计误差。

2.2.3 Plackett-Burman 设计: 该试验设计由 Plackett 和 Burman 提出,是一种以不完全平衡块 (Balanced incomplete blocks)为原理的部分因子设计法,它试图用最少的试验次数达到使因素的主效果得到尽可能精确估计的目的,适用于从众多的考察因素中快速有效地筛选出最为重要的几个因素,供进一步研究用。对于 N 次实验至多可研究 N-1 个因素,但实际因素应该要少于 N-1 个,至少要有 1 个虚构变量用以估计误差。每个因素取 2 个水平,低水平为原始培养条件,高水平约取低水平的 1.25 倍。但对某些因素高低水平的差值不能过大,以防掩盖了其它因素的重要性,应依试验条件而定^[24]。该设计不能区分主效应与交互作用的影响,但对显著影响的因子可以确定出来,从而达到筛选的目的,避免在后期的优化试验中由于因子数太多或部分因子不显著而浪费试验资源。

上述几种部分因子设计方法,各有特点和适用场合,应根据实际情况加以选用。例如进行培养基初步筛选时,除了单因素法外,还可以利用 PB 设计和 2 水平部分因子设计。这两种设计均适合从众多因素中筛选出主要的影响因素,以便以后深入研究各主要因素的影响,很适合培养基的初步筛选,且试验次数比单因素法少,尤其是 PB 设计,次数最少,效率最高。Wu^[25]等利用 PB 设计,仅通过 20 次试验,就从 15 种培养基成分中筛选出葡萄糖、NaNO₃、K₂HPO₄、MnCl₂ 和 ZnSO₄对重组枯草芽孢杆菌发酵产核黄素具有显著影响。陈宁等^[26]也利用 PB 设计,通过 12 次试验,从 L-异亮氨酸初始发酵培养基的 10 种成分中筛选出葡萄糖、(NH₄)₂SO₄ 和 KH₂PO₄对黄色短杆菌(*Brevibacterium flavum*)发酵

产酸有显著影响。

当初步筛选出较重要的因素后可以采用正交、均匀或者响应面设计。如果试验者不希望遗漏被考察因素,且使考察的因素水平范围广,可采用均匀设计安排试验,它可以在一个方案里同时考察多因素多水平效应,能以较少的实验工作量确定大致合适的试验条件。苏国成等^[27]通过单因素法从初始培养基中筛选出葡萄糖、(NH₄)₂SO₄为石蜡节杆菌(*Arthrobacter paraffineus*)发酵产 L-鸟氨酸的最适碳源和氮源,同时又希望综合考察 K₂HPO₄、酵母膏、大豆蛋白胨以及尿素等营养强化因素对菌种产酸的影响,所以作者选择了 6 因素 10 水平 10 组试验方案的均匀设计。

如果试验者通过前期预实验或者参考文献初步确定相对重要的因素之后且因素又不是很多(一般小于 5),因素之间需进一步估计主要与次要因素的影响程度,则采用正交试验设计比较合适,虽然试验次数比均匀设计多一些,但又不是多很多,采用方差分析法或极差分析可以很方便地了解各因素影响的相对大小,确定主、次要因素,这是单因素法所无法比拟的。而且针对主要因素的重要性不同,还可以采用混合正交表安排试验,主要影响因素多安排几个水平,次要因素少安排几个水平,可以有针对性地对因素进行试验研究。

3 最速上升法

最速上升法(Steepest ascent design, SAD)又称最陡爬坡法,是一种能够快速寻找各变量变化最优区域的方法。在发酵培养基的优化过程中常常需要通过回归方程来寻找最优培养基配比,如果建立的回归方程不在最优区域,就不适合用来寻求最优配比,而通过最速上升法就能很好的解决这类问题。

最速上升法是沿着最速上升的路径,即响应面有最大增量的方向逐步移动的方法。当然,如果是最小值,则称最速下降法。拟合的一阶模型是

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^k \hat{\beta}_i x_i$$

与一阶响应曲面相应的 \hat{y} 的等高线,是一系列平行的直线,如图 1 所示,最速上升的方向就是 \hat{y} 增加最快的方向,这一方向平行于拟合响应曲面等高线的法线方向。

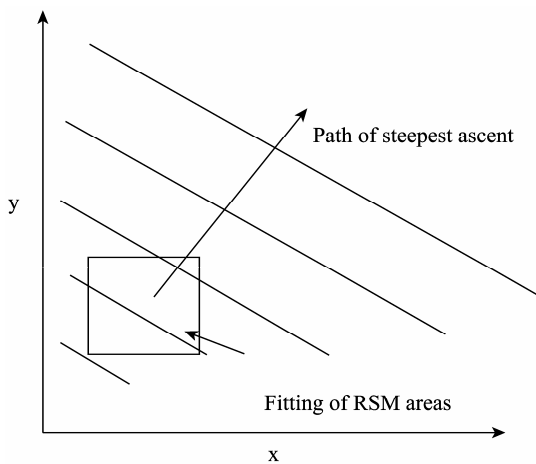


图1 等高线和最速上升路径

Fig. 1 Contour plot of a response and the path of the steepest ascent

在培养基的优化过程中, 可以根据前期试验拟合出的方程来确定各变量的最速上升路径和变化步长, 若系数为正, 则该因素水平为递增, 反之递减。以系数最大的变量为基准确定基本步长, 以其它变量与基准变量系数的比值来确定其它变量的步长, 确定了上升或下降方向及变化步长后便可以选取中心点进行试验设计。

Xiao 等^[28]前期通过 PB 试验拟合出 3-羟基丁酮 (Acetoin) 的产量和 7 种培养基成分的一阶模型:

$$\hat{Y} = 24.5417 + 2.1917X_1 + 2.2917X_2 - 0.0083X_3 - 0.6417X_4 - 0.3917X_5 + 0.2417X_6 + 0.2250X_7$$

然后通过方差分析发现 X_1 (糖蜜) 和 X_2 (豆粕水解液) 对 3-羟基丁酮的产量有显著影响, 一阶模型中 X_1 和 X_2 的系数均为正, 即在试验中水平应为递增, 两者之比接近于 1 ($2.2917/2.1917 \approx 1$), 也就是说 X_2 增加一个步长, X_1 也相应增加一个步长, 即爬坡直线的斜率为 1。然后选定中心点进行爬坡试验, 逼近到最佳区域后即可利用响应面进一步优化(优化过程详见下节)。

4 响应面设计

响应面设计法(Response surface methodology, RSM)是一种寻找多因素系统中最佳条件的数学统计方法, 是数学方法和统计方法结合的产物, 它可以用来对受多个变量影响的响应问题进行建模与分析, 并将该响应进行优化。该法不但具备试验

次数少, 周期短、精度高等优点, 而且可以建立连续变量曲面模型; 同时对影响试验指标的各因子水平及其交互作用进行优化和评价, 可快速有效的确定多因子系统的最佳条件。在发酵培养基的优化中较常用的响应面设计有 Box-behnken 设计 (Box-behnken design, BBD) 和中心组合设计 (Central composite design, CCD) 又称 Box-wilson 设计。

4.1 Box-behnken 设计

该设计是由 Box-behnken 于 1960 年提出拟合响应曲面的 3 水平设计^[29], 该设计是由 2 水平因子设计与不完全区组设计组合而成。每个因素取 3 个水平, 分别以 (-1, 0, 1) 编码, 然后根据实验表进行设计, 运用响应面法对试验后的数据进行分析。

4.2 中心组合设计

由 Box 和 Wilson 开发的中心组合设计是国际上一种较为常用的响应面法。图 2 是 3 因素的 CCD 设计, 该设计由 2^k 析因或部分析因设计 (立方体顶点) 添加 $2k$ 个坐标轴点 ($\pm \alpha, 0, 0 \dots$), ($0, \pm \alpha, 0 \dots$), ($0, 0, \pm \alpha \dots$)... ($0, 0, \dots, \pm \alpha$) 和 N 个中心点 ($0, 0, 0 \dots$) 所组成, 对于 3 因素来说 $k = 3$, 每个因素有 5 个水平。同时 α 值可以使 CCD 试验点具有旋转性, 其意义在于使设计在最优值为主的情况下各个方向上提供等精度的估计, 而且通过调整中心点的次数 N 还可以使 CCD 具有正交性和通用性。CCD 适用于 2-5 因素, 5 水平的优化试验。

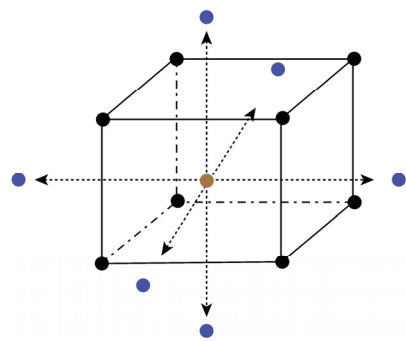


图2 3因素中心组合设计示意图

Fig. 2 Three factors CCD

在发酵培养基的优化过程中, 当因素少于 4 个时, BBD 比 CCD 更具优势, 首先 BBD 每个因素取 3 水平, 故在相同因素下试验次数较 CCD 少, 并且他们是可旋转或近似可旋转的; 其次该设计是球面设计, 并不包含各个变量的极值点, 因此当由于试验

限制而不可能对顶点所代表的因子水平组合进行试验时,此设计就显出其特有的优势^[30]。

目前响应面设计已广泛应用于微生物发酵培养基的优化中(表 1),该试验设计一般步骤为^[31]:(1) 影响因子的确认;Xiao 等^[28]通过文献确定糖蜜、豆粕水解液、磷酸盐等 7 种培养基成分是枯草芽孢杆菌(*Bacillus subtilis*)发酵产 3-羟基丁酮(Acetoin)的影响因素。(2) 因子的筛选,以确定各个因子的影响程度,一般使用 PB 或者 2 水平部分因子设计;文献^[28]利用 7 因素 12 次的 PB 设计试验,从 7 种培养基成分中筛选出糖蜜和豆粕水解液对 3-羟基丁酮的产量有显著影响。(3) 最陡爬坡法确定主要影响因素的合适水平;Xiao 等通过最陡爬坡试验(SAD),发

现糖蜜为 21% (V/V)、豆粕水解液为 25% (V/V)时,3-羟基丁酮的产量(37.5 g/L)接近了响应区间的稳定点。(4) BBD 或 CCD 设计进行试验分析: BBD 或 CCD 方法确定试验点,通过统计分析软件,寻求最佳的试验条件以及因素间的交互作用;Xiao 等利用 SAD 逼近稳定点后,紧接着以 21%糖蜜和 25%豆粕水解液作为响应面设计的中心点,选用 2 因素 13 次试验的中心组合设计(CCD),拟合二次多项式模型,通过 SAS 软件求解出当糖蜜为 22.0%,豆粕水解液为 27.8%时,3-羟基丁酮产量的理论最大值为 37.8 g/L。(5) 最佳条件的验证。最后作者经过验证试验,得出 3-羟基丁酮产量的实际值为 37.9 g/L 与理论值基本一致,优化过程非常成功。

表 1 响应面设计在培养基优化中的应用
Table 1 The application of RSM in optimization of microbial medium

作者 Author	因素筛选 Screening factor	爬坡试验 SAD	响应面设计 RSM	分析软件 Analysis software	文献 Reference
Xiao et al.	7 因素 N = 12 PB 设计	有	2 因素 CCD	SAS	[28]
Chen et al.	6 因素 N = 20 FFD 设计	有	2 因素 CCD	SAS	[32]
Yuan et al.	8 因素 N = 12 PB 设计	有	2 因素 CCD	Minitab	[33]
Cai et al.	10 因素 N = 12 PB 设计	无	3 因素 CCD	Design-expert	[34]
Souza et al.	4 因素 N = 10 PB 设计	无	2 因素 CCD	Statgraphics	[35]
Pan et al.	8 因素 N = 12 PB 设计	有	3 因素 BBD	Minitab	[36]
Li et al.	12 因素 N = 20 PB 设计	无	4 因素 CCD	Design-expert	[37]
Li et al.	6 因素 N = 20 FFD 设计	有	2 因素 CCD	SAS	[38]
李艳华等 Li Yanhua et al.	7 因素 N = 12 PB 设计	有	3 因素 BBD	Minitab	[39]
刘小杰等 Liu Xiaojie et al.	5 因素 N = 20 FFD 设计	有	2 因素 CCD	SAS	[40]
徐广宇等 Xu Guangyu et al.	6 因素 N = 8 PB 设计	有	3 因素 BBD	SAS	[41]
潘春梅等 Pan Chunmei et al.	8 因素 N = 12 PB 设计	无	3 因素 BBD	Minitab	[42]
王雪莲等 Wang Xuelian et al.	10 因素 N = 12 PB 设计	有	3 因素 BBD	Statistica	[43]
赵丽坤等 Zhao Likun et al.	7 因素 N = 12 PB 设计	无	3 因素 BBD	Minitab	[44]
徐子钧等 Xu Zijun et al.	10 因素 N = 24 PB 设计	无	3 因素 BBD	SAS	[45]
张义萍等 Zhang Yiping et al.	10 因素 N = 24 PB 设计	无	3 因素 BBD	SAS	[46]
曹小红等 Cao Xiaohong et al.	6 因素 N = 12 PB 设计	有	3 因素 BBD	SAS	[47]
钟环宇等 Zhong Huanyu et al.	7 因素 N = 8 FFD 设计	无	3 因素 BBD	SAS	[48]
周海鸥等 Zhou Haiou et al.	6 因素 N = 12 PB 设计	有	2 因素 CCD	SAS	[49]
张伟等 Zhang Wei et al.	单因素设计	无	5 因素 CCD	Design-expert	[50]

响应面设计有许多方面的优点,但它仍有一定的局限性。首先,如果因素水平选得太宽,或选得关键因素不全,将会导致响应面出现吊兜和鞍点,因此事先必须进行调研、查询和充分的论证;其次,通过回归分析得到的结果只能对该类实验做估计;第三,当回归数据用于预测时,只能在因素所限的范围内进行预测。即使如此,响应面设计仍不失为一种研究多因素问题的强有力工具^[51]。

5 人工神经网络和遗传算法

人工神经网络(Artificial neural network, ANN)简称神经网络,是20世纪80年代重新兴起的一种模拟人脑及其活动的理论化的数学模型,由多个非常简单的处理单元按某种分时相互联接而形成的计算系统,具有自组织、自适应、自学习等特点,对解决非线性问题特别有效,还有很强的输入输出非线性映射能力,易于学习和训练等优点。它基本上类似黑箱理论,只根据输入数据和输出数据来建立模型,网络的统计信息储存在数量巨大的加权矩阵内,可以反映十分复杂的关系^[52]。目前对于不同目的的神经网络有60多种,其中在发酵培养基优化中应用最多的是误差反传网络即BP(Back-propagation)神经网络。典型的BP神经网络如图3所示。

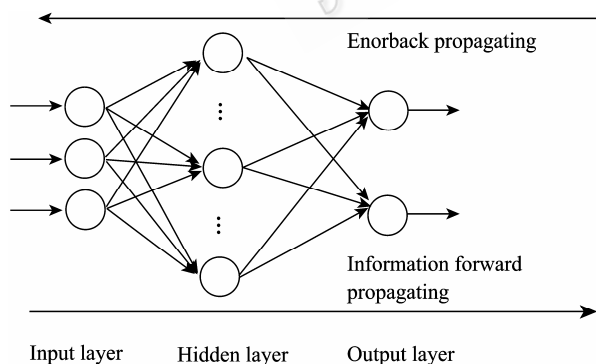


图3 BP神经网络模型

Fig. 3 BP neural network structure for the modeled system

遗传算法(Genetic algorithm, GA)是模拟生物遗传和进化并借用达尔文的“适者生存”学说而发展起来的一种组合优化算法。它摒弃了传统的搜索方式,模拟自然界生物进化过程,采用人工进化的方式对目标空间进行随机化搜索。它将问题域中的可能解

看作是群体的一个个体或染色体,并将变量进行二进制或十进制编码,确定个体性能的评价依据,即适应度函数,然后模拟生物进化过程,对群体反复进行类似于遗传的操作(选择、交叉和变异),根据预定的目标适应度函数对每个个体进行评价,依据适者生存,优胜劣汰的进化规则,不断得到更优的群体,同时以全局并行方式来搜索群体中的最优个体,最终获得适应度最好的个体^[53]。

神经网络有很强的输入输出非线性映射能力,特别适用于微生物发酵这种高度非线性、非结构化的复杂模型中。而遗传算法又是一种有导向的全局随机搜索方法,它对于目标函数和搜索空间没有任何限制,因此非常适合神经网络模型等无明确分析函数形式的优化问题。BP神经网络与遗传算法结合的方式如图4所示。

在图4中,实验培养基配比的组合被分成训练组和预测组,训练组用来训练BP神经网络,然后预测组用来对训练好的网络进行测试,由此构建神经网络模型。并以该模型的输出为GA的目标函数,通过遗传算法的全局寻优,找到最优培养基组合。

目前在发酵培养基优化过程中,除了响应面设计外,神经网络与遗传算法结合使用较为流行。表2列举了一些近几年来国内外学者利用神经网络与遗传算法结合优化培养基的文献。例如Nagata^[54]等利用中心组合设计的试验数据,将30组数据分为训练(27组)和预测(3组)两部分,建立拓扑结果为(4-6-1)的神经网络,然后利用遗传算法全局寻优。结果神经网络和遗传算法建立的模型拟合和预测能力均优于响应面设计得出的二次多项式模型。利用神经网络和遗传算法优化培养基时,并不需要额外单独的设计试验,目前最普遍的法是利用响应面设计的实验数据,建立网络,然后利用遗传算法直接寻优,为响应面设计又提供一个强有力的优化手段。

除上文提到的试验设计和统计优化方法外,还有调优操作法(Evolutionary operation, EVOP),各种回归设计(回归正交设计,回归旋转设计,回归最优设计),改进单纯形设计(Modified simplex method)等,这些设计和优化方法在发酵培养基的优化中使用不多,在此就不再累述。

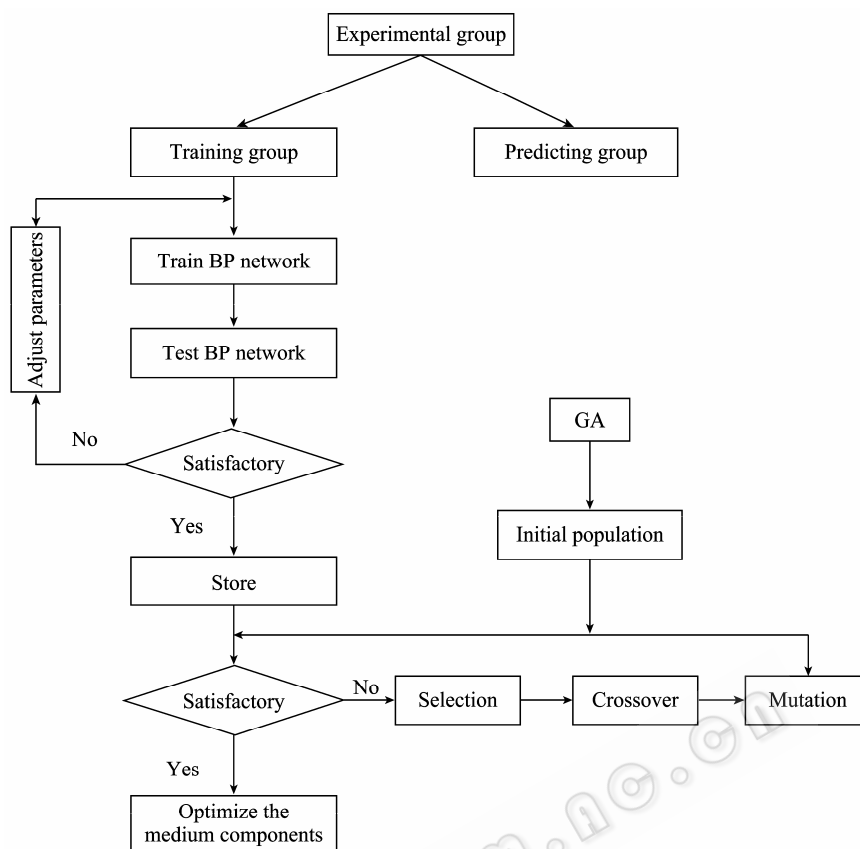


图 4 ANN 与 GA 结合的框图

Fig. 4 Schema of the working principle for ANN coupled GA

表 2 神经网络和遗传算法在培养基优化中的应用
Table 2 The application of ANN-GA in optimization of microbial medium

作者 Author	考察因素 Study factor	优化结果 Optimization results	文献 Reference
Nagata et al.	Molasses、NH ₄ NO ₃ 、NaH ₂ PO ₄ 和 MnCl ₂	ANN-GA 模型的拟合能力优于 RSM 模型	[54]
罗剑飞等 Luo Jianfei et al.	NaNO ₂ 、KH ₂ PO ₄ 、MgSO ₄ 、NaCl 和 NaHCO ₃	ANN-GA 优化结果比原始培养基配方的“性价比”提高了 27.36%	[55]
彭锋等 Peng Feng et al.	Sucrose、(NH ₄) ₂ SO ₄ 、MgSO ₄ ·7H ₂ O、KH ₂ PO ₄ 、Yeast extract 和 CaCl ₂	ANN-GA 优化的类胡萝卜素产量比原始培养基提高了 95.90%	[56]
方柏山等 Fang Baishan et al.	YEP、EP、(NH ₄) ₂ SO ₄ 、(NH ₄) ₂ HPO ₄ 、KH ₂ PO ₄ 和 MgSO ₄	ANN-GA 优化的木糖醇产量比单独使用 GA 提高了 15%	[57]
张峰等 Zhang Feng et al.	胰蛋白胨、牛肉膏、葡萄糖、酵母粉和起始 pH 值	ANN-GA 优化的 A 群链球菌生物产量比用正交试验优化所得产量提高 18%	[58]
蔡宇杰等 Cai Yujie et al.	玉米粉、豆饼粉、麸皮、NH ₄ NO ₃ 、NaH ₂ PO ₄ 和组份 A	ANN-GA 优化法使酶活提高了 15%，大大缩短了实验周期及成本	[59]
赵亮等 Zhao Liang et al.	豆饼粉、玉米浆、葡萄糖	豆饼粉 26.88 g/L，玉米浆 11.32 g/L，葡萄糖 8.97 g/L，氨基丁酸的最高产量 3.84 g/L	[60]
Desai et al.	Sucrose、Yeast extract、K ₂ HPO ₄ 和 MgSO ₄	ANN-GA 模型预测精度和泛化能力均优于 RSM	[61]
朱新术等 Zhu Xinshu et al.	组分 A、蛋白胨、无机盐、酵母膏和组分 E	ANN-GA 结果优于均匀设计，比均匀设计最优组合提取多糖的得率高出 6.85%	[62]
Singh et al.	MgSO ₄ 、NaCl、Glucose、Soybean 和 CaCO ₃	ANN-GA 模型的最优值比 RSM 的最优值提高了 36.7%	[63]
Babu et al.	Glutamic acid、Citric acid、Glycerol 和 NH ₄ Cl	ANN-GA 模型的拟合能力优于 RSM 模型	[64]
Imandi et al.	Sucrose、Ammonium chloride、Ferrous sulphate 和 Zinc sulphate	ANN-GA 模型的拟合能力优于 RSM 模型	[65]
He et al.	Glucose、Soybean、Peptone 和 NaCl	ANN-GA 优化的 PCA 产量高于 RSM 优化的产量	[66]

6 微生物培养基优化的一般步骤及其软件实现

微生物发酵培养基的优化步骤主要由试验设计、数学建模和优化设计 3 部分组成。其中试验设计部分是整个优化过程的关键所在。合理的试验设计可以用更少的实验数据获得更优的建模效果, 合理的模型能够准确预测培养基配比对所考察目标的影响规律, 优化方法则是在已建立数学模型的基础上运用最优化技术, 搜索出因子空间中的最优解。

随着统计软件的不开发, 上述步骤(包括试验设计、模型拟合、方差分析和最优化条件的求解)都可用一些商业软件来完成。这些常用软件有 Design Expert、DPS、SAS、SPSS、Matlab 和 Minitab 等。

例如国际著名的 SAS 统计软件有专门的试验设计版块, 能非常方便地进行以上提到的绝大多数试验设计。以 SAS 8.0 版本为例, 首先进入 SAS 操作界面依次点击 Solutions→Analysis→Design of Experiment 界面中, 然后在试验设计的版块下点击“Create Two-Level Design”小图标进入 Two-Level Design 模块, 点击右边的“Select Design”选项进入, 里面几乎包含了各种情况下的 2 水平设计方案(包括部分析因和 PB 设计), 可根据实际需要选择。当利用筛选试验找到主要的影响因素后, 可以进行常见的响应面设计, 还是在“Design of Experiment”界面中选择“Create Response Surface Design”小图标进入响应面设计模块, 选择需要的响应面试验设计方案。然后依次点击 Fit、Optimize、Report 命令就可以求出最优培养基配方。

Matlab 软件则更擅长数学运算和建模, 例如神经网络和遗传算法, Matlab 中有专门的神经网络和遗传算法工具箱, 而且有一套固定的程序, 试验者只需改动其中参数的取值就能非常方便的建立神经网络和遗传算法的模型并求取最优值。

总之, 虽然发酵培养基的优化是一项复杂的过程, 但在发酵工业产业化过程中又是一项极其重要的工作。我们相信随着统计优化技术的发展, 借助一些统计软件, 能迅速、可靠、方便地安排优化实验和数据处理, 把机理分析、统计优化和计算机软件三者灵活地结合起来, 可以更科学地安排试验, 更方便寻找培养基的最优配比, 提高科研效率。

参 考 文 献

- [1] 郝学财, 余晓斌, 刘志钰, 等. 响应面方法在优化微生物培养基中的应用. 食品研究与开发, 2006, 27(1): 38-41.
- [2] 李季, 蒲彪. 人工神经网络在食品微生物发酵中的应用. 食品研究与开发, 2009, 30(4): 171-174.
- [3] 袁哲明, 左斌, 谭泗桥, 等. 基于均匀设计与支持向量回归的发酵配方优化. 过程工程学报, 2009, 9(1): 148-152.
- [4] Marteiijn RC, Jurrius O, Dhont J, *et al.* Optimization of a feed medium for fed-batch culture of insect cells using a genetic algorithm. *Biotechnol Bioeng*, 2003, 81(3): 269-278.
- [5] Lara EF, Link H, Botz DW. Evaluation of artificial neural networks for modelling and optimization of medium composition with a genetic algorithm. *Proc Biochem*, 2006(41): 2200-2206.
- [6] Botz DW. Experimental design for fermentation media development: statistical design or global random search? *J Biosci Bioeng*, 2000, 90(5): 473-483.
- [7] Wang JL, Wan W. Experimental design methods for fermentative hydrogen production: a review. *International Journal of Hydrogen Energy*, 2009, 34(1): 235-244.
- [8] Kennedy M, Krouse D. Strategies for improving fermentation medium performance: a review. *J Ind Microbiol Biotechnol*, 1999, 23(6): 456-475.
- [9] 陶申傲, 蔡永峰, 岳国海, 等. 采用均匀设计法优化榆耳液态发酵培养基的研究. 食品与发酵工业, 2008, 34(1): 73-76.
- [10] 左爱连, 张伟国. 利用 Design-Expert 软件优化丝氨酸羟甲基转移酶产酶培养基. 生物技术, 2008, 18(3): 45-49.
- [11] Xu CP, Kima SW, Hwang HJ, *et al.* Optimization of submerged culture conditions for mycelial growth and exo-biopolymer production by *Paecilomyces tenuipes* C240. *Proc Biochem*, 2003(38): 1025-1030.
- [12] 周虓, 郑毅, 叶海梅, 等. 响应面分析法优化耐高温蛋白酶发酵培养基. 生物数学学报, 2007, 22(1): 113-118.
- [13] 范洪臣, 李艳华, 梁金钟, 等. 响应面法优化枯草芽孢杆菌产 γ -PGA 的条件. 生物加工过程, 2008, 6(3): 17-23.
- [14] 郑辉杰, 张洪起, 邸进申, 等. 响应面法优化高产海藻糖菌株的培养条件. 食品研究与开发, 2008, 29(8): 53-57.
- [15] 郑晓婷, 赵新淮. 毛霉的产蛋白酶发酵条件优化. 微生物学通报, 2009, 36(2): 193-197.

- [16] 潘艳, 刘海龙, 田野, 等. 统计学分析方法应用于富硒米曲霉发酵培养基的优化. 中国调味品, 2009, **34**(4): 50-55.
- [17] Lu HY, Deng X, Chen Z. Enhanced production of hybrid extracellular β -glucanase by recombinant *Escherichia coli* using experimental design method. *Chin J Chem Eng*, 2007, **15**(2): 172-177.
- [18] 张萃荟, 杨瑞金, 张文斌, 等. 适合乳果糖生产的 β -半乳糖苷酶高产株的筛选及培养基优化. 食品工业科技, 2008, **29**(4): 68-70.
- [19] Achary A, Hariharan KA, Bandhyopadhyaya S, et al. Application of numerical modeling for the development of optimized complex medium for D-Hydantoinase production from *Agrobacterium radiobacter* NRRL B 11291. *Biotechnol Bioeng*, 1997, **55**(1): 148-154.
- [20] Liu CH, Hwang CF, Liao CC. Medium optimization for glutathione production by *Saccharomyces cerevisiae*. *Proc Biochem*, 1999, **34**(1): 17-23.
- [21] Chen HW, Fang BS, Hu ZD. Optimization of process parameters for key enzymes accumulation of 1,3-propanediol production from *Klebsiella pneumoniae*. *Biochem Eng J*, 2005, **25**(1): 47-53.
- [22] 褚以文. 微生物培养基优化方法及其 OPTI 优化软件. 国外医药抗生素分册, 1999, **20**(2): 58-66.
- [23] 李波, 徐泽民, 李方, 等. 试验设计与优化. 中国皮革, 2003, **32**(1): 26-28.
- [24] 欧宏宇, 贾士儒. SAS 软件在微生物培养条件优化中的应用. 天津轻工业学院学报, 2001, **36**(1): 14-17.
- [25] Wu QL, Chen T, Gan Y, et al. Optimization of riboflavin production by recombinant *Bacillus subtilis* RH44 using statistical designs. *Appl Microbiol Biotechnol*, 2007, **76**(4): 783-794.
- [26] 陈宁, 常高峰, 张克旭. L-异亮氨酸发酵培养基的响应面法优化. 食品与发酵工业, 2004, **30**(2): 33-37.
- [27] 苏国成, 黄爱清, 蔡慧农, 等. DPS 数据处理系统在 L-鸟氨酸发酵培养基优化中的应用. 食品科学, 2006, **27**(10): 402-406.
- [28] Xiao ZJ, Liu PH, Qin JY. Statistical optimization of medium components for enhanced acetoin production from molasses and soybean meal hydrolysate. *Appl Microbiol Biotechnol*, 2007, **74**(1): 61-68.
- [29] Montgomery DC. Design and analysis of experiments. 5th Edition. New York: John Wiley and Sons, 2001: 72-93.
- [30] 张楠, 夏尚远, 刘训理, 等. 统计优化技术在微生物发酵中的应用. 山东农业大学学报, 2009, **40**(3): 465-468.
- [31] 刘志祥, 曾超珍. 响应面法在发酵培养基优化中的应用. 北方园艺, 2009(2): 127-129.
- [32] Chen QH, He GQ, Ali MAM. Optimization of medium composition for the production of elastase by *Bacillus* sp. EL31410 with response surface methodology. *Enzyme and Microbial Technology*, 2002, **30**(5): 667-672.
- [33] Yuan LL, Li YQ, Wang Y, et al. Optimization of critical medium components using response surface methodology for phenazine-1-carboxylic acid production by *Pseudomonas* sp. M-18Q. *J Biosci Bioeng*, 2008, **105**(3): 232-237.
- [34] Cai MH, Zhou XS, Sun XQ, et al. Statistical optimization of medium composition for aspergillide a production by marine-derived *Fungus Aspergillus glaucus*. *J Ind Microbiol Biotechnol*, 2009, **36**(3): 381-389.
- [35] Souza MC, Roberto IC, Milagres AMF. Solid-state fermentation for xylanase production by *Thermoascus aurantiacus* using response surface methodology. *Appl Microbiol Biotechnol*, 1999, **52**(6): 768-772.
- [36] Pan CM, Fan YT, Xing Y, et al. Statistical optimization of process parameters on biohydrogen production from glucose by *Clostridium* sp. Fanp2. *Bioresource Technology*, 2008, **99**(8): 3146-3154.
- [37] Li YQ, Jiang HX, Xu YQ, et al. Optimization of nutrient components for enhanced phenazine-1-carboxylic acid production by *gacA*-inactivated *Pseudomonas* sp. M18G using response surface method. *Appl Microbiol Biotechnol*, 2008, **77**(6): 1207-1217.
- [38] Li C, Bai JH, Cai ZL. Optimization of a cultural medium for bacteriocin production by *Lactococcus lactis* using response surface methodology. *Journal of Biotechnology*, 2002, **93**(1): 27-34.
- [39] 李艳华, 梁金钟, 范洪臣. 响应面法优化 γ -聚谷氨酸发酵培养基的研究. 食品科技, 2008, **33**(3): 45-48.
- [40] 刘小杰, 何国庆, 陈启和, 等. 黑曲霉 β -葡萄糖苷酶发酵培养基的优化. 中国食品学报, 2006, **6**(2): 6-10.
- [41] 徐广宇, 骆健美, 杨德山, 等. 纳他霉素发酵培养基及发酵条件的优化. 微生物学杂志, 2007, **27**(4): 73-78.
- [42] 潘春梅, 樊耀亭, 赵攀. 发酵法产氢培养基的响应面分析优化. 中国农学通报, 2008, **24**(1): 38-44.
- [43] 王雪莲, 王敏, 骆健美, 等. 枯草芽孢杆菌生防菌 B579 最佳培养基响应面法优化. 江苏农业学报, 2009, **25**(1): 212-215.
- [44] 赵丽坤, 张利平, 张秀敏, 等. 利用 Minitab 软件优化阿维菌素产生菌发酵培养基. 河南农业科学, 2009(6): 90-93.
- [45] 徐子钧, 李剑, 梁凤来, 等. 利用 SAS 软件优化 L-乳酸发酵培养基. 微生物学通报, 2004, **31**(3): 85-87.
- [46] 张义萍, 张伟国. 利用 SAS 软件优化 L-精氨酸发酵培养基. 化学与生物工程, 2006, **23**(2): 44-46.
- [47] 曹小红, 蔡萍, 李凡, 等. 利用响应面法优化 *Bacillus*

- natto TK-1 产脂肽发酵培养基. 中国生物工程杂志, 2007, 27(4): 59-65.
- [48] 钟环宇, 许建军, 江波. 利用响应面分析法优化 γ -氨基丁酸发酵培养基. 无锡轻工大学学报, 2004, 23(3): 19-22.
- [49] 周海鸥, 汪传高, 张益波, 等. 统计学分析方法应用于桑黄菌发酵培养基的优化. 食品研究与开发, 2009, 30(5): 44-48.
- [50] 张伟, 黄耀坚, 郑忠辉, 等. 响应面法优化南强菌素发酵培养基. 厦门大学学报: 自然科学版, 2009, 48(4): 585-589.
- [51] 林敏, 宁喜斌. 优化方法在木糖醇发酵培养中的应用. 食品研究与开发, 2006, 27(5): 26-27.
- [52] 张蓓, 熊明勇, 张克旭, 等. 人工神经网络在发酵工业中的应用. 生物技术通讯, 2003, 14(1): 74-76.
- [53] 秦华明, 赵颖怡, 梁世中. 遗传算法在微生物发酵培养中的应用. 食品与发酵工业, 2001, 28(3): 51-55.
- [54] Nagata Y, Chu KH. Optimization of a fermentation medium using neural networks and genetic algorithms. *Bio-technol Lett*, 2003, 25(1): 1837-1842.
- [55] 罗剑飞, 林炜铁. 基于神经网络和遗传算法培养基优化的发酵经济学. 食品与生物技术学报, 2009, 28(3): 424-428.
- [56] 彭锋, 李安明. BP 神经网络和遗传算法在红发夫酵母培养基优化中的应用. 应用与环境生物学报, 2008, 14(6): 834-837.
- [57] 方柏山, 陈宏文, 谢晓兰, 等. 基于神经网络和遗传算法的木糖醇发酵培养基优化研究. 生物工程学报, 2000, 16(5): 648-650.
- [58] 张峰, 王志兵, 郎维, 等. 神经网络与遗传算法对 A 群链球菌发酵培养基的优化. 长春工业大学学报, 2008, 29(3): 263-266.
- [59] 蔡宇杰, 诸葛斌, 张锡红, 等. 遗传算法与神经网络耦联法优化生淀粉酶发酵培养基. 无锡轻工大学学报, 2001, 20(4): 421-423.
- [60] 赵亮, 刘俏, 宋莉. 用 BP 神经网络与遗传算法优化 γ -氨基丁酸的发酵培养基. 计算机与应用化学, 2008, 25(10): 1273-1276.
- [61] Desai KM, Survase SA, Saudagar PS, *et al.* Comparison of artificial neural network and response surface methodology in fermentation media optimization: Case study of fermentative production of scleroglucan. *Biochemical Engineering Journal*, 2008, 41(3): 266-273.
- [62] 朱新术, 杨志强, 李建喜, 等. 基于神经网络和遗传算法的黄芪发酵培养基优化. 中兽医医药杂志, 2008, 27(3): 18-20.
- [63] Singh V, Khan M, Khan S, *et al.* Optimization of actinomycin V production by *Streptomyces triostinicus* using artificial neural network and genetic algorithm. *Appl Microbiol Biotechnol*, 2009, 82(2): 379-385.
- [64] Babu IS, Rao GH. Optimization of culture medium for the production of the poly- γ -glutamic acid using artificial neural networks and genetic algorithm. *Research Journal of Microbiology*, 2006, 1(6): 520-526.
- [65] Imandil SB, Karanam SK, Garapati HR. Optimization of fermentation medium for the production of lipopeptide using artificial neural networks and genetic algorithms. *International Journal of Natural and Engineering Sciences*, 2008, 2(2): 105-109.
- [66] He L, Xu YQ, Zhang XH. Medium factors optimization and fermentation kinetics for phenazine-1-carboxylic acid production by *Pseudomonas* sp. M18G. *Biotechnol Bioeng*, 2008, 100(2): 250-259.

稿件书写规范

论文中统计学符号书写规则

统计学符号一般用斜体。本刊常用统计学符号介绍如下, 希望作者参照执行。

样本的算术平均数用英文小写 x , 不用大写 X , 也不用 $Mean$ 。标准差用英文小写 s , 不用 SD 。标准误差用英文小写 $s_{\bar{x}}$, 不用 SE 。 t 检验用英文小写 t 。 F 检验用英文大写 F 。卡方检验用希文小写 χ^2 。相关系数用英文小写 r 。样本数用英文小写 n 。概率用英文大写 P 。