

文章编号:1009-038X(2004)03-0051-06

人工神经网络结合遗传算法在焙炒大米过程中的应用

朱益波, 张建华, 史仲平*, 毛忠贵

(江南大学 工业生物技术教育部重点实验室, 江苏 无锡 214036)

摘 要: 焙炒是一种全新的使原料淀粉糊化的方法. 它用热风替代水蒸气, 在高温, 短时间的条件下处理原料米, 具有无废水污染, 容易保存等优点, 为食用酒酿制过程中的一种新型技术. 通过对焙炒大米的 3 个指标: 糊化率、脂肪含量和氨基氮进行测定, 利用人工神经网络技术 ANN 对上述性能指标和操作参数的数据进行训练学习, 得到可以描述焙炒过程操作条件和性能指标之间关系的模型. 在所得模型的基础上, 利用遗传算法 GA 对大米的焙炒条件实施优化, 对未参与 ANN 建模的数据进行评价和比较, 结果发现, 结合使用 ANN 和 GA, 能够比较准确地预测对应于期望指标的操作条件, 预测结果与实验数据吻合.

关键词: 人工神经网络; 遗传算法; 焙炒; 糊化率

中图分类号: TP 389.1

文献标识码: A

Application of Genetic Algorithm Associated with Artificial Neural Network Techniques for Optimization of Rice Roasting Process

ZHU Yi-bo, ZHANG Jian-hua, SHI Zhong-ping, MAO Zhong-gui

(The Key Laboratory of Industrial Biotechnology, Ministry of Education, Southern Yangtze University, Wuxi 214036, China)

Abstract: Rice roasting process is a novel technique for raw starch gelatinization. It uses heated air to replace water steam for processing rice under the conditions of high temperature and short time, and is a new technique for making rice wine with the characteristics of easy storage and zero water pollution. This study measured three major process performance index, starch α -ratio, total fat and amino nitrogen contents of the roasted rice under various operating conditions, then the relationship between the performance index and the corresponding operation variables were modeled by artificial neural network (ANN) by learning and training the corresponding data sets. Based on the models obtained, genetic algorithm was used to optimize the rice roasting process, which is to search the optimal operation variables corresponding to the expected performance index. The results showed that the combination of artificial neural network and genetic algorithms could accurately estimate the operating variables for the expected performance index, by evaluating several data sets unused for the ANN models' learning and training.

Key words: artificial neural network; genetic algorithm; roasting; starch α -ratio

收稿日期: 2003-08-12; 修回日期: 2003-12-30.

作者简介: 朱益波(1980-), 男, 江苏靖江人, 环境工程硕士研究生; * 通讯作者.

焙炒是一种全新的使原料淀粉糊化的方法,即用热空气替代水蒸气,高温短时间处理原料米。由于在原料处理全过程中“干法”作业,和传统的蒸饭法相比,处理后的原料不会发生淀粉老化现象,具有易保存、无废水污染等优点,被国内众多的黄酒生产单位和研究单位所关注^[1]。在利用焙炒大米进行黄酒与清酒的清洁生产中,焙炒过程的操作参数对焙炒米的指标具有决定作用,而这些指标又对焙炒大米的储藏和酿造有着重要影响,直接关系到成酒的品质以及生产成本。焙炒过程的性能指标包括:糊化率、脂肪、氨基氮和水分质量分数等,而主要影响性能指标的操作参数则包括:焙炒温度、焙炒时间和米的浸泡时间等。由于焙炒过程是一个复杂的物理和化学反应过程,很难建立一个能够准确描述过程特性的数学模型,并利用这个模型进行过程优化。

人工神经网络(ANN)计算仿真技术是近年来迅速发展起来的一门应用研究学科。随着计算机技术的不断更新完善,ANN技术在包括发酵过程控制和优化等许多研究领域得到了广泛应用^[2~4]。由于人工神经网络不需要传统的反映透明机制的生物或物理化学模型,且其处理复杂、非线性系统(数据)的能力很强,因此与传统的建模方式相比,具有很大的优越性。BP-ANN网络是当前应用较为广泛的一种多层前向传播神经网络模型,训练学习采用误差反向传播算法。BP网络结构简单,状态稳定,计算条件易于满足,可有效地用于非线性函数逼近和基于不规则数据结构的复杂系统仿真^[5]。

遗传算法(Genetic Algorithm)简称GA,是模拟生物的遗传和长期进化过程建立起来的一种搜索和优化算法,它模拟生物界“生存竞争,优胜劣汰,适者生存”的机制,用逐次迭代法搜索寻优,所得到的最优解为全局最优解。遗传算法模拟了自然选择和自然遗传中发生的繁殖,交配和突变现象,将问题中每个可能的解看作是种群中的一个个体,并将每个个体编码成字符串的形式,根据预定的目标函数对每个个体进行评价,给出一个适合度值,进行逐次择优传代。遗传算法的3种基本算子是:选择,交叉,变异^[6]。近年来,GA在发酵过程中的应用大多集中在确定动力学模型参数^[7,8],决定最佳底物流加策略^[9,10]以及优化培养基成分^[11]等方面。

由于食用酒酿造过程是一个非常复杂的过程,过程优化控制往往依靠操作者的经验。作者尝试使用ANN和GA的方法,并使用MATLAB 6.5编写的程序对大米焙炒过程进行建模和过程优化。

1 材料与方法

1.1 大米焙炒试验

1.1.1 实验材料 市售糯米;糖化酶:取2 mL糖化酶(酶活 10×10^4 u/mL)稀释到100 mL;中性蛋白酶GC710,酸性蛋白酶GC106:无锡杰能科公司产品;SBA-40B生物传感分析仪,山东科学院生物研究所产品;索氏抽提器;Sartorius MA40水分天平。

1.1.2 大米的焙炒 在确保焙炒大米完全流化的情况下,影响焙炒的3个主要因素分别是大米的含水量、焙炒温度以及焙炒时间,根据实际操作条件按照正交表 $L_9(3^4)$ 安排多批试验,根据各批对应的试验条件进行焙炒试验,各批焙炒大米分装待测定其指标。为便于工业上的操作,大米的含水量以30℃水的浸湿时间来代替。

1.1.3 焙炒性能指标测定 糊化率的测定方法采用焙炒米的糊化率快速测定法^[12]。脂肪质量分数的测定使用索氏抽提法^[13]。焙炒米中氨基氮的质量分数在模拟发酵条件下,对蛋白酶水解焙炒米中的蛋白质产生的 α -氨基氮进行测定。测定方法采用中性甲醛滴定法。

1.2 人工神经网络的结构以及模型的建立

作者设计了一个3层BP神经网络结构,见图1。输入层神经元数为3,每个输入单元分别代表相应的输入变量,即大米浸泡时间、焙炒温度和焙炒时间。中间隐层的层数为10。输出层层数为1,输出变量分别代表焙炒性能指标,即糊化率、脂肪含量和 α -氨基氮含量。因此,完整的神经网络模型由3个 $3 \times 10 \times 1$ 的网络所构成。其中, x_i 代表输入层和中间隐层中的第*i*个输入变量; $O_j^{(2)}$ 和 $O_j^{(3)}$ 分别代表中间隐层和输出层中的第*j*个输出变量,其值由对应的 x_i 和各层各单元间的权值 W_{ij} 的乘积之和,以及Sigmoid形式的传递函数所决定。输入层和中间隐层各有一个偏差单元。这里 $N1 = 3 + 1 = 4$, $N2 = 10 + 1 = 11$ 。各层各单元间的权值 W_{ij} 在网络学习和训练过程中经过优化调整,通过误差,即“期望”输出值与神经网络输出层的计算值之差达到最小而求出。采用MATLAB中的Neural Network Toolbox系统仿真环境提供的网络学习,训练和仿真模块进行计算,计算出各层各单元间的权值 W_{ij} ,进而得到大米焙炒过程的操作参数与对应的性能指标间的模型。

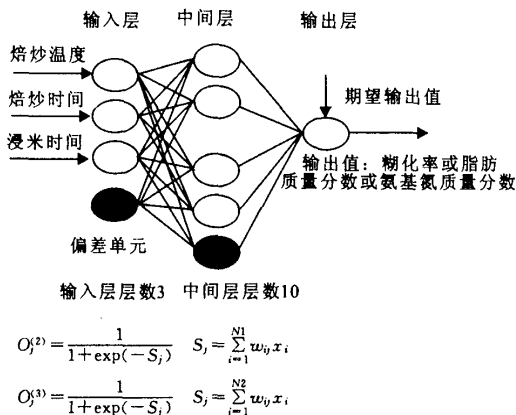


图1 3层BP神经网络模型结构

Fig. 1 Three-layer BP neural network model

1.3 遗传算法的建立

在神经网络建模后,用遗传算法求取对应于目标性能指标(糊化率或脂肪质量分数或 α -氨基氮质量分数,或复合指标)的优化操作条件(即焙炒温度、焙炒时间和大米浸泡时间),具体步骤如下:

1)将操作参数遗传基因-染色体化,即将3个操作变量统合并编码成一个二进制串结构数据,由计算机随机地产生由2进(0或1)文字序列构成的初代染色体(基因)种群N个。

2)遗传基因序列的解码化,将解码后种群的二进制串结构数据转换成原有的操作参数。

3)将解码化的参数代入ANN模型求解,算出各染色体的适合度(Fitness)。其中适合度(Fitness)一般可以定义为 $1 - \sum \text{ABS}(y_i - y_i^*)$, y_i 是由操作参数和ANN模型求得的性能指标, y_i^* 为目标性能指标。适合度可以是单一适合指标($i=1$;如仅以糊化率为指标),也可以是复合适合指标($i=1, N$;如同时以糊化率和脂肪质量分数为指标 $N=2$)。

4)根据各种群适合度的大小,从种群中选择适合度较大的一些个体组成交配池。

5)通过交叉和变异两种遗传算子对交配池中的个体进行遗传操作,并形成新一代的种群。

6)反复执行步骤2)到5),直至达到规定的遗传代数。

作者在种群选择算子中采用了最优保存策略和比例选择法相结合的思路,即首先找出当前群体中适合度最高和最低的个体,将最佳个体保留并用其替换掉最差个体。为保证当前最佳个体不被交叉和变异操作所破坏,允许其不参与交叉变异而直接进入下一代。这种做法的优点是能保证最佳个体的适合度不减小,并使最高适合度逐渐增加并得到最

佳的个体^[14]。本研究遗传算法的计算参数为:种群大小51,搜索步长0.01,交叉概率为0.6,变异概率为0.1,传代数数为40~100。

2 结果与讨论

2.1 大米焙炒中各指标随操作变量变化的基本趋势

将大米浸泡5 min后,用炒米机高温流化处理45 s测定糊化率、脂肪质量分数和 α -氨基氮质量分数随焙炒温度变化的曲线,结果见图2。在同一焙炒时间下糊化率随着温度的升高而升高,糊化率在温度达到240℃时,增长趋势减缓,并开始达到稳定状态,糊化率接近90%。同时,同一温度下焙炒时间长的大米糊化率高(数据未列出),较低温度下适当延长焙炒时间提高大米的糊化率。焙炒米的脂肪质量分数随着焙炒温度的升高逐渐降低,到250℃后开始达到稳定。同样,在相同的焙炒温度下,焙炒时间长,脂肪质量分数也较小。大米中易于利用的氨基氮质量分数开始随着焙炒温度的升高而上升,到了150℃左右,随着焙炒温度的升高则反而开始下降。一般而言,所得到的焙炒米糊化率较高,脂肪质量分数较低,酿制的酒具有特殊的香味,属于清爽型的黄酒。而用于酿造黄酒的大米氨基氮的质量分数控制,应根据酿造的具体要求而定。

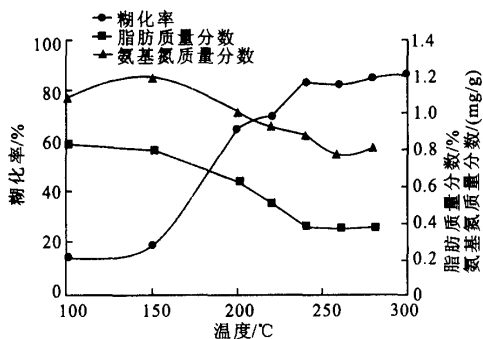


图2 大米焙炒中各指标随操作变量变化的基本趋势

Fig. 2 Basic change pattern of various performance index versus operating conditions

2.2 ANN人工神经网络模型的预测性能

为建立得到ANN人工神经网络模型,将33套糊化率与操作参数(浸泡时间、焙炒温度和焙炒时间)的数据,23套脂肪质量分数与操作参数的数据,23套氨基氮质量分数与操作参数的数据,分别用来训练和学习从而得到各自神经网络模型中各层各单元间的权值 W_{ij} 。另外,各6套未用于训练和学习的数据,用来评价所得到的ANN人工神经网络模型的准确性和通用能力。糊化率和脂肪质量分数的

神经网络模型分别经过约 500 次和 250 次的训练和优化调整,相应的网络模型的误差平方和分别达到 5.52×10^{-6} 和 8.38×10^{-5} 。图 3a、图 3b 分别表示归一化(实际值/最大值)的糊化率和脂肪质量分数的模型预测值与实际值之间的相互关系。如果模型足够准确,所有的数据点都应该处在通过原点的直线上。从图中可以看出,用于训练学习的数据和用于评价的数据均处在该直线的附近,这表明所得到的人工神经网络模型能够比较准确的预测焙炒的操作参数与相应性能指标之间的关系,从而为后续的遗传算法的运行实施提供了比较准确的模型。有关氨基酸质量分数的神经网络模型的预测性能,因篇幅所限未在本文中给出。

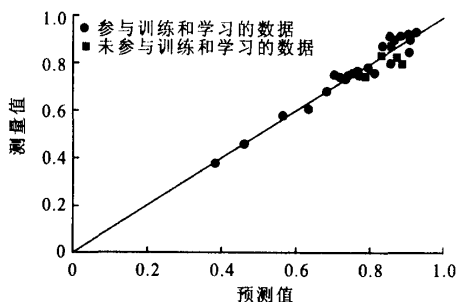


图 3a ANN 对糊化率的预测性能

Fig. 3a Estimation performance for starch α -ratio by ANN

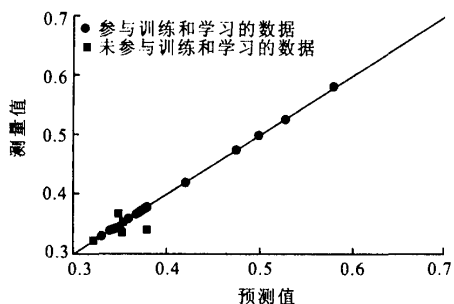


图 3b ANN 对脂肪质量分数的预测性能

Fig. 3b Estimation performance for fat content by ANN

2.3 利用遗传算法求解对应于目标性能指标的最优操作参数

2.3.1 以糊化率为惟一性能指标时的最优操作参数以及遗传算法的性能 以大米焙炒的糊化率作为惟一性能指标,利用遗传算法进行优化计算时的糊化率以及对操作参数的变化趋势见图 4a 和图 4b。这时,目标糊化分数被设定在 0.921 (糊化率 92.1%)。

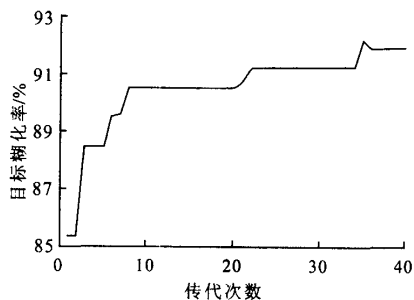


图 4a 糊化率随世代次数的变化趋势

Fig. 4a Starch α -ratio versus genetic generation

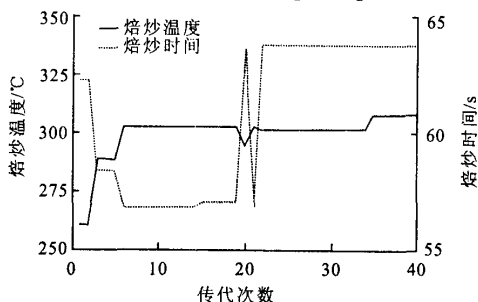


图 4b 操作参数随世代次数的变化趋势

Fig. 4b Operating variables versus genetic generation

由图 4 可以看到遗传操作中每一代的糊化率通过改变操作参数而逐渐增大,经过 40 代的遗传操作,糊化率趋近于其目标值(92.1%),而对应于目标值的最优焙炒温度和焙炒时间分别为 308 °C 和 64 s。此时,遗传算法的性能用 6 套未参与神经网络训练和学习的数据来加以评价,即比较对应于实际目标值的计算操作参数与实际操作参数,其结果见表 1。由表 1 可以看到焙炒时间和焙炒温度的计算值与实际值比较接近,但浸米时间的相对误差及其变动则较大,其原因是与焙炒时间和焙炒温度相比,浸米时间对目标性能影响较小的缘故。另外,焙炒时间和焙炒温度的增大均会造成糊化率的升高,这会出现同一个糊化率可以对应多个操作参数的情况。

2.3.2 以糊化率和脂肪质量分数为复合性能指标时的最优操作参数以及遗传算法的性能 以大米焙炒的糊化率和脂肪质量分数为复合性能指标,利用遗传算法进行优化计算时的糊化率、脂肪质量分数以及相对应操作参数的变化趋势见图 5a 和图 5b。这时,目标糊化率和脂肪质量分数分别被设定在 0.921(糊化率 92.1%)和 0.314%。

由图 5 可以看到与以糊化率为惟一性能指标时相比,为使性能指标趋近于其目标值,需要较多的世代次数。本研究需要约 70 代左右的遗传操作,

糊化率和脂肪质量分数才能趋近于其目标值,但却不能完全达到所设定的目标值.这时,对应于复合目标值的最优焙炒温度和焙炒时间分别为 298 ℃ 和 62 s.以复合性能指标为基础的遗传算法的性能用另外 2 套未参与神经网络训练和学习的数据来加以评价,其结果见表 2.

表 1 以糊化率为惟一性能指标时的实际操作参数与预测参数对照

Tab. 1 Comparison results of actual and estimated operating variables with starch α -ratio as the sole performance index								
项 目	批次	糊化率/%	浸米时间/min	相对误差/%	温度/℃	相对误差/%	焙炒时间/s	相对误差/%
实际操作参数	1	55.1	5		150		55	
	2	74.8	5		220		55	
	3	83.1	5		240		55	
	4	87.7	5		260		55	
	5	81.7	5		280		55	
	6	92.3	5		310		55	
遗传算法预测参数	1	55.1	7.6	52.0	173.9	15.9	46.5	15.6
	2	74.8	3.5	30.0	218.6	0.6	55.3	9.6
	3	83.1	5.0	0.0	254.0	5.8	57.4	4.3
	4	87.7	5.1	2.0	289.2	11.2	50.7	7.8
	5	81.7	1.5	70.0	256.8	8.3	53.8	2.2
	6	92.3	5.5	10.0	313.0	1.0	61.6	12.0

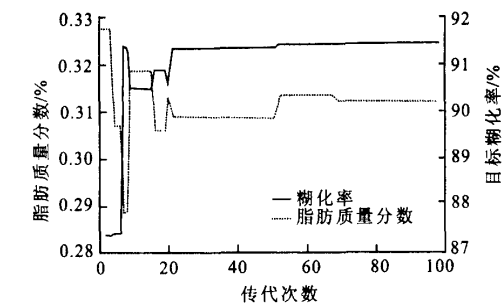


图 5a 糊化率和脂肪质量分数随传代次数的变化趋势

Fig. 5a Starch α -ratio & fat content versus genetic generation

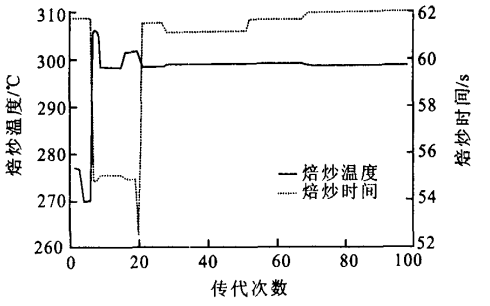


图 5b 操作参数随传代次数的变化趋势

Fig. 5b Operating variables versus genetic generation

表 2 以糊化率和脂肪含量为复合性能指标时的实际操作参数与预测参数对照

Tab. 2 Comparison results of actual and estimated operating variables with both starch α -ratio and fat content as the complex performance index

项 目	批次	糊化率/%	脂肪质量分数/%	浸米时间/min	温度/℃	焙炒时间/s
实际操作参数	1	91.55	0.328	5.0	300	55
	2	91.63	0.321	7.0	300	60
遗传算法预测参数	设定	91.55	0.325	6.5	303	62

由表 2 可以看到焙炒时间和焙炒温度的计算值与实际值非常接近,说明以复合性能指标为基础的遗传算法有很高的预测精度.

3 结 论

在大米焙炒过程中结合使用 ANN 和 GA 的方法,能够比较准确地预测对应于期望指标的操作条

件,预测结果与实验数据吻合。以上建模和过程参数优化的方法,对于大米焙炒过程的实际应用和优化具有一定的指导意义。建立准确的过程模型需要大量的实验数据,而生产和试验过程中积累的数据则可以用来不断地完善既存的模型,从而使得最优

操作条件的预测值更加准确。另外,对于影响焙炒大米其它品质的参数也可以通过类似的方法来建模和优化操作条件。作者提出的方法简单可行,具有一定应用价值。

参考文献:

- [1] 贝沼禎介. 焙炒技術の清酒製造への応用[A]. '97 国际酒文化研讨论文集[C]. 上海: 学林出版社, 1998, 39—43.
- [2] 苗志奇, 赵凌云, 元英进. 谷氨酸流加发酵过程的神经网络优化研究[J]. 生物工程学报, 1998, 14: 198—202.
- [3] 未作君, 苗志奇, 元英进. 酵母流加发酵过程的神经网络优化研究[J]. 无锡轻工大学学报, 2000, 19: 547—551.
- [4] Shi Z, Shimizu K. Neural-fuzzy control of bioreactor systems with pattern recognition[J]. *J Biosci Bioeng*, 1992, 74: 39—45.
- [5] 薛定宇. 控制系统计算机辅助设计, MATLAB 语言及应用[M]. 北京: 清华大学出版社, 1996.
- [6] 席裕庚, 柴天佑. 遗传算法综述[J]. 控制理论与应用, 1996, 13(6): 697—708.
- [7] Ranganath M, Renganathan S, Gokulnath C. Identification of bioprocess using genetic algorithm[J]. *Biopro Eng*, 1999, 21: 123—127.
- [8] 陈宏文, 方柏山, 胡宗定. 遗传算法应用分批发酵动力学模型参数估算[J]. 华侨大学学报, 2000, 21(1): 71—75.
- [9] Na J G, Chang Y K, Chung B H, *et al*. Adaptive optimization of fed-batch culture of yeast by using genetic algorithms [J]. *Biopro Eng*, 2002, 24: 299—308.
- [10] 王丹, 高玲, 林建强, 等. 用遗传算法优化流加培养的底物流加轨迹[J]. 工业微生物, 2002, 32(3): 40—43.
- [11] 方柏山, 陈宏文, 胡宗定. 基于遗传算法的木糖醇发酵培养基优化研究[J]. 化学反应工程与工艺, 2000, 16: 12—16.
- [12] 张建华, 彭昌亚, 毛忠贵, 等. 焙炒米的糊化率快速测定方法[J]. 酿酒, 2002, 29(1): 80—81.
- [13] 无锡轻工大学. 工业发酵分析[M]. 北京: 中国轻工业出版社, 1999.
- [14] 闵惜琳. 人工神经网络结合遗传算法在建模和优化中的应用[J]. 计算机应用研究, 2002, (1): 79—80.

(责任编辑: 朱 明)