

# 基于 GRNN 的米糠蛋白提取条件优化研究

刘 静<sup>1</sup>, 管 骁<sup>2,\*</sup>, 李景军<sup>3</sup>

(1.上海海事大学信息工程学院,上海 201306;

2.上海理工大学医疗器械与食品学院,上海 200093;

3.江苏雨润食品产业集团有限公司,江苏南京 210041)

**摘 要:**针对米糠蛋白碱法生产中存在的得率低、色泽深等普遍问题,本文在单因素实验确定的几个蛋白提取关键影响因素基础上,采用中心组合实验的方法探讨了因素之间交互作用的问题,最后利用广义回归神经网络技术(general regression neural network, GRNN)对蛋白提取率和提取液色泽进行了多目标优化问题研究。GRNN 模型结果显示:在温度 36.5℃,液料比 11.5:1 (v/w), pH10.9, 辅助剂用量 0.56% 的条件下,米糠蛋白提取率理论值为 61.0%, 色度值为 53.3, 与实际值的误差分别为 3.6% 和 4.7%, 且相比较目前工业生产普遍使用的生产条件,蛋白提取率与色度值分别提高了 31.4% 与 43.3%, 采用 GRNN 方法优化的米糠蛋白提取条件具有很好的实用价值。

**关键词:**广义回归神经网络,米糠,蛋白,提取,优化

## Study on the extraction optimization of rice bran protein based on GRNN

LIU Jing<sup>1</sup>, GUAN Xiao<sup>2,\*</sup>, LI Jing-jun<sup>3</sup>

(1.College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China;

2.School of Medical Instruments and Food Engineering, University of Shanghai for Science and Technology,

Shanghai 200093, China; 3.Jiangsu Yurun Food Industry Group Co., Ltd., Nanjing 210041, China)

**Abstract:** According to the general problems including the low yield and the deep colour of product during the alkali-extraction process of rice bran protein, the cross-action of different factors was discussed through central composite experiments on the basis of the key factors determined by single factor experiments. Furthermore, the optimization of interaction-effect including protein yield and color were performed by generalized regression neural network (GRNN) method. The GRNN model results showed that 61.0% protein yield and 53.3 color value were obtained under the conditions of temperature 36.5℃, ratio of solve to material 11.5:1 (v/w), pH value 10.9, and auxiliary agent amount 0.56%, which existed a deviation of 3.6% and 4.7% with actual values, respectively. Compared with the conditions for industrial production, the protein extraction yield and colour value raised 31.4% and 43.3%, respectively, under the optimized conditions. It was proved that the optimized extraction conditions of rice bran protein by GRNN had great pragmatic value for industrial production.

**Key words:** GRNN; rice bran; protein; extraction; optimization

中图分类号: TS210.9

文献标识码: B

文章编号: 1002-0306(2012)15-0187-04

米糠是稻谷加工过程中的副产物,我国年产量在 1000 万 t 以上。米糠中蛋白含量较高,达 14%~20%,且米糠蛋白营养价值高,功能性质较好,特别是还具有低过敏性的优点,因此开发米糠蛋白产品具有广阔的市场前景<sup>[1-2]</sup>。米糠蛋白提取通常有碱法、酶法和物理法,不同方法各有优缺点,但一般认为碱法操作简便、成本低廉、适合于工业化生产<sup>[3]</sup>。碱提蛋白法是根据蛋白质在碱性环境中具有较高溶

解度的原理,影响因素较多,包括提取液 pH、液料比、提取温度等,如一般提取液 pH 越高,蛋白提取率越高,然而也带来负面影响,如容易发生脱赖反应生成有毒物质、促进美拉德反应导致产品颜色过深等,且不同因素之间还存在交互作用<sup>[4-5]</sup>。因此碱提蛋白法存在的主要问题是:如何优化提取条件以平衡蛋白提取率与产品色泽的矛盾问题。广义回归神经网络 (Generalized Regression Neural Network, GRNN) 是径向基神经网络的一种,具有很强的非线性映射能力和柔性网络结构以及高度的容错性和鲁棒性,适用于解决非线性问题;特别是在逼近能力和学习速度上较径向基神经网络 (Radical basis function, RBF)、BP 神经网络具有更强的优势,网络最后收敛于样本量积聚较多的优化回归面,并且在样本数据较少时,预测效果也较好<sup>[6-7]</sup>。此外,网络还可以处

收稿日期:2011-11-14 \* 通讯联系人

作者简介:刘静(1979-),女,博士,副教授,研究方向:智能算法与机器学习。

基金项目:国家自然科学基金(31101348,31000780);上海市晨光计划项目(2008CG55,09CG50)。

理不稳定的数据<sup>[6]</sup>。因此,GRNN 在能源化工、控制决策系统、食品工程、结构分析等诸多领域都得到了广泛的应用<sup>[7-9]</sup>。本研究针对米糠蛋白的碱法提取问题,采用中心组合实验设计探讨了各提取条件的综合影响,在此基础上利用 GRNN 建立米糠蛋白提取模型,以蛋白提取率和提取液色泽为目标进行实验参数的优化,为开发出理想的蛋白产品提供指导。

## 1 材料与方法

### 1.1 材料与设备

米糠粕(低温) 实验室自制;氢氧化钠、盐酸、氧化镁、氯化钙等 均为分析纯。

磁力搅拌器 上海亚荣生化仪器厂;CS501 型超级恒温器 重庆实验设备厂;玻璃夹套反应器 玻璃厂定制;离心机 上海医疗器械集团;721 型分光光度计 上海光学仪器厂;WB-2000IXA 型全自动测色色差仪 北京新恒能分析仪器有限公司。

### 1.2 实验方法

1.2.1 低温米糠粕蛋白的提取 低温米糠粕原料中的蛋白含量为 17.1%。称取 10g 米糠粕原料置于玻璃反应器中,同时设定水浴温度,并按比例加入一定体积的提取辅助剂溶液,搅拌均匀后调节 pH 至设定值,提取结束后,离心分离,收集上清液,测定上清液的体积与蛋白质浓度以及提取液的色度。

#### 1.2.2 蛋白质提取率的计算

$$\text{蛋白提取率}(\%) = \frac{\text{提取液中蛋白的质量}}{\text{原料中总蛋白质量}} \times 100$$

1.2.3 蛋白质含量测定 固体采用微量凯氏定氮法(GB5511-85)(氮对蛋白转换系数取 5.95);液体采用福林-酚法。

1.2.4 蛋白提取液色度测定<sup>[10]</sup> 采用 L, a, b 均匀色空间表色系统来评价蛋白提取液的颜色深浅。L 表示亮度, a、b 表示色彩, L 值越小表明提取液色泽越深。

1.2.5 实验设计 本研究采用四因素、五水平的中心组合实验设计(Central Composite Rotatable Design),共计 31 次实验来考察提取条件对蛋白提取率及提取液色泽的影响。选取考察的提取条件及其水平是基于单因素实验(结果另文发表)。各自变量的水平按式(1)转换为无因次代码表示:

$$x_i = (X_i - X_0) / \Delta X_i \quad \text{式(1)}$$

式中: $x_i$ -自变量的代码值; $X_i$ -自变量的真实值; $X_0$ -自变量中心水平处真实值; $\Delta X_i$ -步长。各自变量及其水平如表 1 所示。31 个条件下的提取实验随机进行以减小系统误差,每个条件实验重复三次,结果取平均值。实验结果如表 2 所示。

1.2.6 数据处理 实验设计采用 SAS 8.0 软件进行。由于涉及到多目标优化问题,实验数据分析及结果处理采用 Matlab 2009a 软件进行。

1.2.7 模型验证 GRNN 由四层构成,分别为输入层、模式层、求和层和输出层<sup>[8,10]</sup>(如图 1 所示),对应网络输入为温度、液料比、pH 和辅助剂用量,其输出分别为蛋白提取率和提取液色泽。GRNN 网络仅需要调整一个光滑因子参数,因此可更快地预测网络,

具有更强的计算优势。本研究采用 GRNN 方法获得最优的米糠蛋白提取条件,需同时满足较高的蛋白提取率与较浅的蛋白色泽要求。为验证模型预测结果的准确性,根据预测的优化条件进行蛋白提取实验,重复三次,结果取平均值。将实验值与预测值对比,以验证模型的准确性。

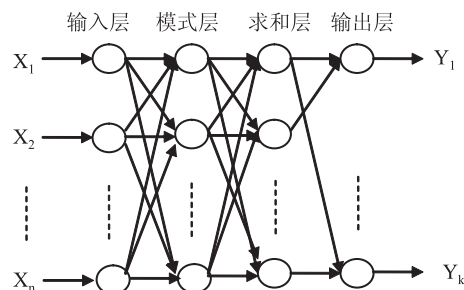


图 1 GRNN 网络结构示意图

Fig.1 Diagram of GRNN network structure

## 2 结果与讨论

### 2.1 中心组合实验设计结果与分析

根据单因素实验结果,确定了温度、液料比、pH、辅助剂用量四个因素为影响蛋白提取率和产品色泽的重要因素,对上述 4 个因素进行 5 水平的中心组合实验设计<sup>[11-12]</sup>,同时考察蛋白提取率与提取液色度值的变化情况。全部实验包括 31 个实验点,因素及水平的选取见表 1,实验结果如表 2 所示。

表 1 实验影响因素与水平

Table 1 The experiment factors and levels

水平	因素			
	X <sub>1</sub> 温度 (°C)	X <sub>2</sub> 液料比 (v/w)	X <sub>3</sub> pH	X <sub>4</sub> 辅助剂 用量(%)
-2	30	6:1	9	0.2
-1	35	8:1	9.5	0.3
0	40	10:1	10	0.4
1	45	12:1	10.5	0.5
2	50	14:1	11	0.6

### 2.2 基于 GRNN 神经网络的数据处理与分析

根据中心组合实验的结果,利用 MATLAB 2009a 软件以温度、液料比、pH、辅助剂用量为自变量,蛋白提取率、提取液色泽两项为因变量构建 GRNN 模型。对米糠蛋白提取而言,理想情况是在某一提取条件下,得到高蛋白提取率的同时,提取液的颜色较浅为宜(即 L 值较大),即属于多目标优化问题,GRNN 方法是解决这一问题的良好途径。基于此,先将两项因变量作归一化处理,分别建立不同的模型,以便寻找最优值。

2.2.1 样本的选取与归一化处理 本文选取实验号 1~30 作为样本的训练集,31 号作为样本的测试集。将温度、液料比、pH、辅助剂用量 4 个因素的实际值作为自变量,将两项因变量按式(2)进行归一化处理 [0, 1] 后作为输入层,归一化处理结果如表 3 所示。在输出层两项因变量根据式(3)将数值换回。

$$x_i = (X_i - X_{\min}) / (X_{\max} - X_{\min}) \quad \text{式(2)}$$

$$Y_i = x_i (x_{\max} - x_{\min}) + x_{\min} \quad \text{式(3)}$$

式中:  $x_i$  表示归一化后的数据;  $X_i$  为归一化前的真实值;  $X_{\min}$  为数据最小值;  $X_{\max}$  为数据最大值。

表2 中心组合实验设计与结果

Table 2 The central composite rotatable design and results

实验号	$X_1$	$X_2$	$X_3$	$X_4$	蛋白提取率 ( $Y_1, \%$ )	L 值 ( $Y_2$ )
1	-2	0	0	0	48.14	51.92
2	-1	-1	1	-1	56.40	38.63
3	-1	-1	1	1	54.22	49.66
4	-1	-1	-1	-1	46.78	51.20
5	-1	1	1	-1	59.75	54.31
6	-1	-1	-1	1	46.83	52.76
7	-1	1	1	1	61.28	53.10
8	-1	1	-1	1	56.66	57.42
9	-1	1	-1	-1	51.53	58.86
10	0	-2	0	0	50.36	42.30
11	0	0	2	0	60.37	46.31
12	0	0	0	-2	57.41	50.76
13	0	0	0	0	57.06	53.50
14	0	0	0	0	57.16	53.55
15	0	0	0	0	57.13	53.69
16	0	0	-2	0	51.99	59.82
17	0	0	0	2	55.74	60.76
18	0	2	0	0	54.08	71.45
19	0	0	0	0	55.84	53.51
20	0	0	0	0	56.85	52.95
21	0	0	0	0	55.60	53.71
22	0	0	0	0	55.69	53.74
23	1	-1	1	-1	58.33	39.07
24	1	-1	1	1	53.38	50.65
25	1	1	1	-1	58.88	51.19
26	1	-1	-1	-1	48.89	52.50
27	1	1	1	1	60.73	53.05
28	1	-1	-1	1	47.28	54.70
29	1	1	-1	1	57.67	59.40
30	1	1	-1	-1	55.24	57.73
31	2	0	0	0	57.35	56.88

2.2.2 样本的训练与测试 为寻找蛋白提取率与 L 值的最优解, 本文通过 MATLAB 软件分别建立不同的 GRNN 模型, 包括提取因素与蛋白提取率, 提取因素与 L 值, 以及提取因素与蛋白提取率和 L 值的加权模型。

2.2.2.1 不同提取条件对蛋白提取率的影响 根据建立的提取条件与蛋白提取率的 GRNN 模型, 蛋白提取率的实际值与网络输出的比较如图 2 所示。通过最优化分析, 光滑因子确定为 0.55, 此时对 31 号实验网络的测试输出最大值为 0.5703, 与实际值 0.7290 相比, 误差达 21.8%。由于训练样本容量较小, 难以得到更精确的预测结果。

2.2.2.2 不同提取条件对 L 值的影响 根据建立的提取条件对 L 值的 GRNN 模型, L 值的实测值与网络输出的比较如图 3 所示。通过最优化分析, 光滑因子为 0.55, 对 31 号实验网络的测试输出最大值为 0.4161, 误差高达 25.2%, 同样受限于训练样本容量较小的原因, 预测误差相对较大。

表3 归一化处理结果

Table 3 Results of normalization

实验号	归一化结果[0,1]	
	蛋白提取率	L 值
1	0.0938	0.4049
2	0.6634	0
3	0.5131	0.3361
4	0	0.3830
5	0.8945	0.4778
6	0.0034	0.4305
7	1.0000	0.4409
8	0.6814	0.5725
9	0.3276	0.6164
10	0.2469	0.1118
11	0.9372	0.2340
12	0.7331	0.3696
13	0.7090	0.4531
14	0.7159	0.4546
15	0.7138	0.4589
16	0.3593	0.6456
17	0.6179	0.6743
18	0.5034	1.0000
19	0.6248	0.4534
20	0.6945	0.4363
21	0.6083	0.4595
22	0.6145	0.4604
23	0.7966	0.0134
24	0.4552	0.3662
25	0.8345	0.3827
26	0.1455	0.4226
27	0.9621	0.4394
28	0.0345	0.4896
29	0.7510	0.6328
30	0.5834	0.5820
31	0.7290	0.5561

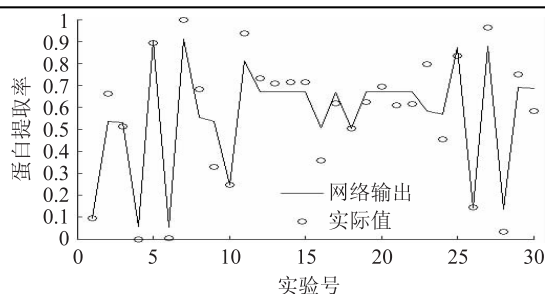


图2 GRNN 模型中蛋白提取率的实际值与网络输出的比较

Fig.2 Comparison of actual value and network output of protein extraction in GRNN model

2.2.2.3 不同提取条件对蛋白提取率与 L 值加权值的影响 蛋白提取率与产品色泽是蛋白提取问题中主要考虑的两个方面, 为使提取率与 L 值同时达到最优, 本研究将它们归一化后的数值各按权重 50% 进行加和处理, 建立模型。此时建立的 GRNN 模型中, 实际值与网络输出的比较如图 4 所示。通过最优化分析, 光滑因子为 0.17。将测试集代入, 蛋白提取率与 L 值的归一化测试结果分别为 0.7575 和



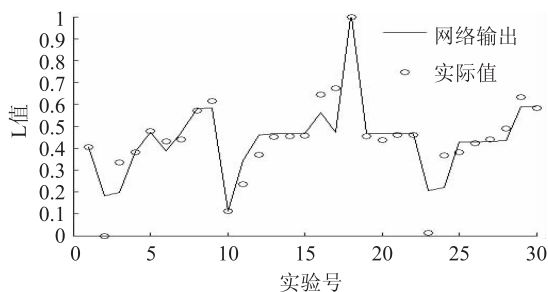


图3 GRNN模型中L值的实测值与网络输出的比较

Fig.3 Comparison of actual value and network output of L value in GRNN model

0.5526, 经过反归一化处理, 结果为 57.76 和 56.53, 误差分别为 0.7% 和 0.6%。说明该模型具有很好的预测能力, 可以用来预测最优提取条件。

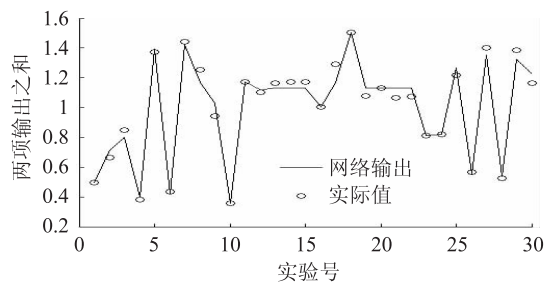


图4 蛋白提出率与L值加权之和的实际值与网络输出的比较

Fig.4 Comparison of actual value and network output of weighted sum of protein extraction and L value

2.2.3 优化提取条件验证实验 利用 MATLAB 软件在以上加权模型中进行寻优, 得到米糠蛋白的最优提取工艺条件为温度 36.5℃, 液料比 11.5:1, pH10.9, 辅助剂用量 0.56%, 在此条件下蛋白提取率预测值为 61.0%, 色度 L 值为 53.3, 结果均比较理想。较之于目前工业生产普遍采用的米糠蛋白提取条件<sup>[13]</sup> (温度 45~55℃, 液料比 10:1 (v/w), pH11.5, 无辅助剂添加), 本研究可将蛋白提取率提高 31.4%, 色度值提高 43.3%, 说明本研究对米糠蛋白提取条件的优化是有效的。另外, 将优化前后的米糠蛋白提取条件进行对比, 不难看出优化后的条件尽管对液料比的要求稍有提高, 并增加了辅助剂的使用, 但大大降低了提取过程液料的温度, 同时对碱性环境的要求也有所降低, 即降低了蛋白提取成本的同时, 产品的得率与外观方面有显著改善, 因此在工业生产中的应用必将有很好的前景。

为进一步检验 GRNN 模型预测值的可靠性, 进行了验证实验。按模型给出的最优提取条件进行米糠蛋白提取实验, 得到蛋白实际提取率为 58.9% ± 2.1%, 色度值 L 为 55.9 ± 1.9, 与预测值的误差分别为 3.6% 和 4.7%, 在可接受的范围内, 说明采用 GRNN 模型预测得到的米糠蛋白最优提取条件是真实可靠的, 具有实用价值。

### 3 结论

本文通过中心组合实验考察了温度、液料比、pH 以及辅助剂添加量对蛋白提取率和提取液色泽的综合影响, 并采用 GRNN 方法建立了提取条件预测模型, 确定了米糠蛋白的最佳提取工艺条件为: 温度 36.5℃, 液料比 11.5:1, pH10.9, 辅助剂用量 0.56%, 在此条件下提取率预测值为 61.0%, 色度 L 预测值为 53.3。对以上预测结果进行了验证实验, 结果表明蛋白提取率和提取液色泽预测偏差分别为 3.6% 和 4.7%, 在可接受范围内。与目前工业生产普遍采用的提取条件相比, 本研究优化的条件可将米糠蛋白提取率提高 31.4%, 色度 L 值提高 43.3%。

### 参考文献

- [1] 李新华, 富艳鑫, 郑煜焱. 米糠蛋白提取工艺条件的优化[J]. 食品科学, 2010, 31(22): 251-254.
- [2] 陈季旺, 孙庆杰, 夏文水, 等. 碱酶两步法制备大米蛋白的研究[J]. 农业工程学报, 2006, 22(5): 169-172.
- [3] 陈季旺, 姚慧源, 张小勇, 等. 米糠可溶性蛋白的提取工艺和特性研究[J]. 中国油脂, 2003, 28(2): 46-50.
- [4] Hamada J S. Use of proteases to enhance solubilization of rice bran proteins [J]. Journal of Food Biochemistry, 1999, 23: 307-321.
- [5] Kaewka K, Therakulkait C, Cadwallader K R. Effect of preparation conditions on composition and sensory aroma characteristics of acid hydrolyzed rice bran protein concentrate [J]. Journal of Cereal Science, 2009, 50: 56-60.
- [6] 刘秀清, 字仁德, 范东凯. 基于广义回归神经网络的交通事故预测[J]. 山东理工大学学报: 自然科学版, 2007, 21(2): 28-31.
- [7] Kulkarni S G, Chaudhary A K, Nandi S, et al. Modeling and monitoring of batch processes using principal component analysis (PCA) assisted generalized regression neural networks (GRNN) [J]. Biochemical Engineering Journal, 2004, 18: 193-210.
- [8] 张德丰. MATLAB 神经网络应用设计[M]. 北京: 机械工业出版社, 2008.
- [9] 章德宾, 徐家鹏, 许建军, 等. 基于监测数据和 BP 神经网络的食物安全预警模型[J]. 农业工程学报, 2010, 26(1): 221-226.
- [10] 陶菲. 真空预冷处理延长白蘑菇贮藏期的研究[D]. 无锡: 江南大学, 2006.
- [11] 魏艳强. 基于 RBF 神经网络的货运量预测模型研究[D]. 天津: 天津理工大学, 2007.
- [12] Cochran W G, Cox G M. Some methods for the study of response surfaces. In *Experimental designs* [M]. New York: Wiley, 1992: 335-375.
- [13] 温焕斌, 曹晓虹, 李翠娟, 等. 米糠蛋白提取工艺优化及其特性研究[J]. 扬州大学学报: 农业与生命科学版, 2010, 31(2): 72-77.