

基于近红外光谱技术与 BP-ANN 算法的豆粕品质快速检测

周新奇¹, 杨伟伟¹, 房兆华², 桑强¹, 叶华俊¹, 陈智锋¹

(1. 聚光科技(杭州)股份有限公司, 浙江 杭州 310052; 2.)

摘要:应用近红外漫反射光谱技术结合误差反向传递人工神经网络(BP-ANN)算法,建立豆粕品质(包括水分、粗蛋白、残油)的定量分析模型。将豆粕漫反射吸收光谱数据进行 SNV、DT、SG 求导、SG 平滑和均值中心化处理,然后采用偏最小二乘方法(PLS)降维获取主成分,并优化选择合适的隐含层节点数、隐含层和输出层转化函数,建立校正模型,并用验证样品对校正模型进行验证。结果显示,BP-ANN 法建立的水分、粗蛋白和残油的预测相关系数(R)分别为 0.981、0.988、0.982,预测标准偏差(SEP)分别为 0.120、0.216、0.036,均优于 PLS 建模方法结果,且满足传统分析方法的重复性要求,表明 BP-ANN 方法可用于生产过程豆粕品质的快速监控。

关键词:神经网络; 近红外光谱; 豆粕

中图分类号: TS 207.3

文献标识码: A

文章编号:

Rapid determination of soybean meal quality based on Near Infrared Spectroscopy coupled with BP-ANN

ZHOU Xin-qi¹, YANG Wei-wei¹, FANG Zhao-hua², SANG Qiang¹, YE Hua-jun¹, CHEN Zhi-feng¹

(1. Focused Photonics (Hangzhou), Inc., Hangzhou 310052, China; 2.)

Abstract:The models of quantitative analysis of moisture, protein and residual oil in soybean meal were established by back propagation- artificial neural network method (BP-ANN) based on near infrared spectroscopy (NIR). Firstly, the original absorbance spectra of soybean meal samples were pretreated by SNV, DT, Savitzky-Golay derivative, Savitzky-Golay smoothing and mean-centering. Secondly, the principal components were obtained by PLS dimension-reducing, and the number of hidden node, transfer functions of hidden layer and output layer were optimized; Finally, all the parameters were inputted into BP-ANN to establish the calibration model. Then the models were validated by prediction set. The results showed that the correlation coefficients (R) of prediction for moisture, crude protein and residual oil were 0.989, 0.982 and 0.994 respectively; and the standard errors of prediction (SEP) were 0.120, 0.216, 0.036, respectively. It shows that BP-ANN was more accurate compared with the partial least square method (PLS). Furthermore, the results meet the repeatability of traditional analysis method, it can be applied to rapid monitoring of soybean meal quality.

Keywords: ANN; NIR; soybean meal

豆粕是大豆油脂企业重要的产品,也是饲料企业重要的蛋白原料。水分、粗蛋白及残油含量是衡量豆粕品质的关键指标。检测上述品质指标一般采用化学分析方法,但这些传统的分析方法分析周期长、消耗大,很大程度上限制了豆粕品质管理水平。

近红外光谱(NIRS)分析是一种新兴的快速分析技术。与传统分析方法相比,NIRS技术具有样品前处理简便、无需化学试剂、操作简单、检测速度快等优点^[1-5]。该技术利用物质组成与其近红外吸收光谱的对应关系,通过化学计量学方法建立模型,进行定性和定量分析。NIRS技术采用的建模方式有很多,如主成分分析^[6]、偏最小二乘回归分析^[7]、人工神经网络^[8]等。

豆粕的主要成分如水分、粗蛋白、残油在近红外区域具有特征吸收,因此可用近红外光谱分析技术对其成分和性质进行分析,目前国内外已在此方面开展了研究^[9-11],并成功应用于生产企业。这些研究主要采用线性的偏最小二乘方法进行分析。

由于温度、背景变化等因素影响,样品的近红外光谱与其待测组分含量或质量指标之间的关系往往具有非线性、多因子的复杂关系。采用非线性校正方法在理论上有望建立更好的预测模型。人工神经网络(ANN)作为一种新兴的化学计量学方法,可以逼近任何连续的非线性曲线,有效反应非线性数据依赖关系,是一种用于非线性校正分析的首选方法。

本文探讨将近红外漫反射光谱结合误差反传人工神经网络方法(BP-ANN)应用于豆粕品质分析,并将分析

结果与 PLS 方法所得结果进行比较。

1 材料与方法

1.1 实验材料和方法

豆粕样品由山东三维粮油公司提供，共 560 个样品。采用聚光科技（杭州）股份有限公司生产的 SupNIR-2700 近红外光谱分析仪对样品进行光谱分析。光谱分析的波长范围 1000-1800nm，分辨率 10nm，波长间隔 1nm，波长准确性小于 0.2nm，吸收噪声小于 5E-5。将每份豆粕样品经粉碎机粉碎后，置于分析仪旋转样品台上进行光谱分析。为保证样品分析结果更好的一致性，将样品受光表面刮平。样品的吸收光谱如图 1 所示。本文采用光谱分析专用软件 RIMP 中的 BP-ANN 方法建立分析模型。

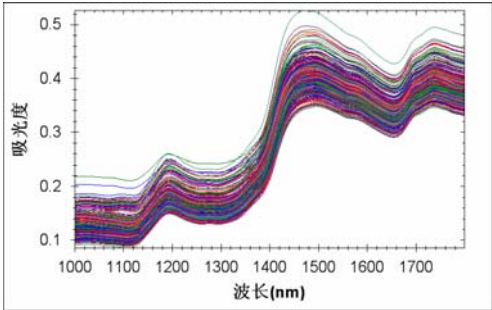


图 1 豆粕的近红外吸收光谱

1.2 样品参考值的测定

豆粕水分测定参照《GB/T 10358-2008 油料饼粕 水分及挥发物含量的测定》进行检测；豆粕粗蛋白测定参照《GB/T 6432-94 饲料中粗蛋白测定方法》进行检测，检测结果用湿基表示；豆粕残油测定参照《GB/T 10359-2008 油料饼粕 含油量的测定 第 1 部分：己烷（或石油醚）提取法》进行检测，检测结果用干基表示。样品的水分、粗蛋白和残油参考值统计结果如表 1 所示。

表 1 豆粕水分、粗蛋白和残油的参考值统计结果

成分	最小值	最大值	极差	均值	标准偏差
水分	10.82	15.19	4.37	12.63	0.63
粗蛋白	41.44	47.55	6.11	44.55	1.50
残油	0.36	1.99	1.63	0.85	0.36

2 结果与讨论

2.1 近红外光谱预处理

为了减少高频随机噪声、基线漂移、样本不均匀、光散射等影响，需要进行光谱预处理。所测的样品近红外光谱数据依次应用 SNV、DT、SG 求导（窗口参数 7，拟合次数 2，求导次数 1）、SG 平滑（窗口参数 7，拟合次

数 2）和均值中心化运算。

2.2 神经网络参数选择

影响 BP-ANN 法建立 NIR 模型的主要参数有主因子数、隐含层的节点数、隐含层传递函数、输出层传递函数、学习速率、训练次数等^[12]。

在本实验的 BP 神经网络的建立过程中，选择合适的 PLS 主成分作为神经网络的输入变量，通过交互验证，利用预测残差平方和（PRESS 值）比因子选择数法^[13]确定 PLS 降维的最佳主成分数。根据 Kolmogorov 经验定理^[12]和多次试验确定隐含层节点数。

因此，水分、粗蛋白、残油三个 ANN 模型的参数选择如下：水分 ANN 模型参数：输入层节点数为 12，隐含层节点数为 4，隐含层传递函数为线性函数 purelin，输出层传递函数为线性函数 purelin；粗蛋白 ANN 模型参数：输入层节点数为 16，隐含层节点数为 4，隐含层传递函数为非线性函数 logsigmoid，输出层传递函数为线性函数 purelin；残油 ANN 模型参数：输入层节点数为 16，隐含层节点数为 4，隐含层传递函数为线性函数 purelin，输出层传递函数为线性函数 purelin。

其他参数：输入层到隐含层及隐含层到输出层的初始权重均为 -0.5~0.5 之间的随机数，动量项为 0.9，学习速率取 0.1，训练最大步数 2500，网络训练 100 步左右即可达到预先设定的学习误差 0.1。

2.3 BP-ANN 模型

从所有豆粕样品中（其中水分 325 份，粗蛋白 452 份，残油 295 份），随机取 80% 作为校正集，剩余 20% 的样品作为验证集。分别采用 BP-ANN 与 PLS 方法建立的豆粕水分、粗蛋白和残油模型统计指标如表 2 所示。图 2 表示使用 BP-ANN 方法获得的豆粕样品粗蛋白化学值（参考值）与相应预测值的相关曲线，横轴为化学值，纵轴为预测值。

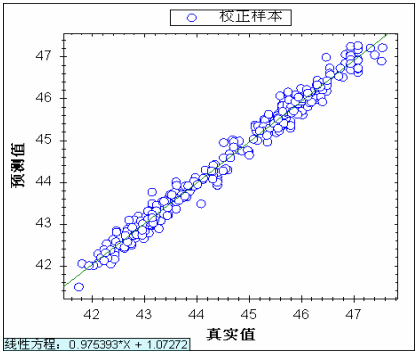


图 2 验证样品的粗蛋白化学值-预测值相关性

表 2 豆粕水分、粗蛋白和残油模型的统计指标

性质 名称	建模 方法	SEC	SEP	RC	RP	RPD
水分	BP-ANN	0.123	0.130	0.982	0.981	4.815
	PLS	0.139	0.166	0.975	0.968	3.771
粗蛋白	BP-ANN	0.219	0.246	0.989	0.988	6.081
	PLS	0.235	0.274	0.987	0.984	5.460
残油	BP-ANN	0.037	0.039	0.994	0.982	9.179
	PLS	0.044	0.049	0.992	0.981	7.306

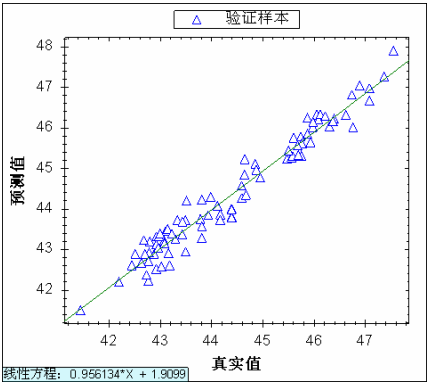


图 3 验证样品的粗蛋白化学值-预测值相关性

2.5 稳定性验证

为验证 BP-ANN 豆粕模型预测的稳定性,对豆粕样品进行 10 次扫描,结果如表 4 所示,10 次预测结果的水分、粗蛋白及残油的标准差分别为 0.031, 0.035, 0.035, 满足 1.2 节所示的国标检测重复性要求。

表 4 BP-ANN 豆粕模型预测 10 次扫描的结果

次数	水分 (%)	粗蛋白 (%)	残油 (%)
1	14.135	44.841	0.799
2	14.105	44.822	0.890
3	14.171	44.825	0.900
4	14.151	44.858	0.891
5	14.166	44.877	0.907
6	14.138	44.820	0.897
7	14.210	44.863	0.869
8	14.200	44.904	0.928
9	14.151	44.870	0.889
10	14.143	44.782	0.907
极差	0.105	0.122	0.129
标准差	0.031	0.035	0.035

3 结论

本文采用近红外光谱技术结合三层 BP-ANN 算法,建立豆粕水分、粗蛋白和残油的定量分析模型。对验证样品的预测结果及稳定性试验结果表明,该方法分析结果准确性和稳定性满足传统分析方法的国标要求。该方法操作简便,分析结果准确,适用于豆粕产品的水分、粗蛋白及残油品质指标的快速分析。

2.4 准确性验证

从豆粕的水分、粗蛋白、残油的所有样品集中,随机取 20%作为验证集。豆粕验证集的粗蛋白真实值与预测值的相关性如图 3 所示。表 3 为部分验证集水分的预测误差。从表 3 可看出,水分的平均绝对偏差分别为 0.103,满足 1.2 节所示的国标检测准确性要求(限于篇幅,未列出蛋白及残油预测偏差),表明模型的预测能力很好,可满足快速检测豆粕水分、粗蛋白及残油的分析要求。

表 3 部分验证样品水分参考值与 NIR 预测值误差

序号	水分参考 值	水分预 测值	绝对 偏差	相对 偏差
1	14.36	14.46	0.1	-0.70
2	13.03	12.87	0.16	1.21
3	12.90	12.93	0.03	-0.26
4	12.18	12.10	0.08	0.67
5	12.03	12.20	0.17	-1.44
6	11.98	11.94	0.04	0.36
7	11.64	11.66	0.02	-0.14
8	11.51	11.35	0.16	1.42
9	11.29	11.55	0.26	-2.28
10	10.97	10.98	0.01	-0.09
平均值	12.189	12.204	0.103	-0.125

参考文献:

- [1] 陆婉珍.现代近红外光谱分析技术(第二版)[M]. 北京:中国石化出版社, 2007:1
- [2] 严衍禄.近红外光谱分析基础及应用[M].北京:中国轻工业出版社,2005:24
- [3] NORRIS K H, Roman J D.Qualitative Applecaiton of Near-Infrared Reflectance Spectroscopy. Agriculture engineering. 1962,43:154
- [4] WILLIAMS P C,NORRIS K H. Near-infrared Technology in the agricultural and food industries (2nd edition)[M].USA:American Association of Cereal Chemists Inc.,St.Paul,MN,2001
- [5] 徐广通,袁洪福,陆婉珍. 现代近红外光谱技术及应用进展. 光谱学与光谱分析.2000,20(2): 134 -142
- [6] 吉海彦,严衍禄.主成分-人工神经网络在近红外光谱定量分析中的应用.分析测试学报.1999,18(3):12
- [7] 张明祥,闵顺耕,李宁.乙醇混合燃料近红外定量分析.分析测试学报.2003,22(6):15
- [8] 任瑞雪,汤真,刘福强.傅里叶变换近红外全谱回归分析的应用研究.光谱学与光谱分析.2001,21(4):521
- [9] 卢利军,庄树华,李爱军,吴剑,张少杰. 应用近红外技术测定黄豆粕中水分蛋白质和粗脂肪. 分子科学学报.2001,17(2): 115 -120
- [10] 杨海锋,吕小文,秦玉昌,等. 近红外光谱分析技术在豆粕质量监控中的应用研究.饲料工业.2006,27(19): 31-34
- [11] 孟兆芳,张玺,陈峥.近红外光谱分析技术在豆粕豆质分析中的应用.天津农业科学.2002,8(4): 30-32
- [12] 杨建.人工神经网络实用教程[M].杭州:浙江大学出版社, 2001:41
- [13] 许禄.化学计量学一些重要方法的原理及应用[M].北京:科学出版社,2004:23