

文章编号:1671-1513(2012)05-0077-04

# 基于BP神经网络检测面粉中滑石粉含量的研究

刘翠玲, 董秀丽, 孙晓荣, 吴静珠, 吴胜男

(北京工商大学 计算机与信息工程学院, 北京 100048)

**摘要:** 利用近红外光谱技术对掺杂滑石粉的小麦面粉进行了检测,采用多元散射校正对谱图进行预处理,利用BP神经网络中的SCG反向传播算法训练函数建立了面粉中滑石粉的定量分析模型,并对校正集和预测集进行了定量分析,分析结果为 $R^2 = 0.9973$ , $RMSEC = 0.4367$ , $RMSEP = 1.7088$ 。结果表明,BP神经网络结合近红外光谱技术检测面粉中滑石粉含量具有快速、精度高、泛华能力强的优点,可用于面粉中滑石粉含量的快速准确检测。

**关键词:** BP神经网络; 近红外光谱; 滑石粉; 小麦面粉

**中图分类号:** TS205

**文献标志码:** A

面粉中掺杂滑石粉是食品安全问题中亟待解决的重要环节,研究者们针对此问题做了大量工作,目前有多种方法用于检测面粉中滑石粉的含量,但是这些方法有的不够精确,有的过程复杂耗时较长。根据实地调查,常用的检测方法仍以传统的物理-化学方法为主,尚无统一的更加快捷检测面粉中滑石粉含量的方法和标准。针对此研究课题,本文使用BP神经网络结合近红外光谱技术,旨在研究面粉中滑石粉快速定量检测的可行性<sup>[1]</sup>。

## 1 面粉中掺杂滑石粉检测现状和BP神经网络

### 1.1 面粉中滑石粉检测现状

小麦面粉及面制品是人们尤其是北方居民的主食,其质量的好坏对人们的身体健康有重要的影响,国标法对小麦粉面粉的检测做了详细的规定<sup>[2]</sup>。国家对小麦面粉的添加剂种类做了明确的规定,然而,在暴利的驱动下,部分不法厂家向小麦面粉中添加大量的滑石粉,有些厂家甚至在25 kg面粉中添加多达3~4 kg的滑石粉。因此,对面粉中滑石粉的检测急需找到一种快速简便的方法,以保证广大人民的身体健康<sup>[3]</sup>。

滑石粉的主要成分是天然的水合硅酸镁( $Mg_3[Si_4O_{10}](OH)_2$ ),为白色或类白色、微细、无砂性的粉末,手摸有滑腻感,无色无味,不溶于水、稀硝酸或稀氢氧化碱溶液,可作药用。滑石粉中常含有对人体有害的重金属,长期服用含有滑石粉的小麦面粉会增加肝脏的负担,甚至会危及人的生命<sup>[4]</sup>。

目前,国标法GB/T21913—2008<sup>[5]</sup>对食品中滑石粉的检测主要采用化学法测定镁和二氧化硅的含量,间接换算成滑石粉的量,操作繁琐且耗时较长,而且检测成本较高。其他方法<sup>[6]</sup>等存在对样品前处理试剂消耗量大的缺点,X射线衍射分析法<sup>[7,8-12]</sup>需用有机试剂分离面粉或将面粉灰化后再进行检验,处理时间长且X射线对人体有辐射,不利于长时间工作。

与化学计量学方法结合使近红外光谱技术得到广泛的关注。近红外光谱技术使用化学计量学方法建立可靠的校正模型,具有分析速度快、效率高、成本低等优点,可直接实现对样品进行无损检测,近年来在农业、食品、药品、烟草等领域有着广泛应用。

本文应用近红外漫反射光谱法结合BP神经网络建立面粉中滑石粉含量的定量分析模型,并用所建模型对预测集样品进行预测,得到较好的结果,该方法有望成为一种代替现行面粉中滑石粉含量测定

收稿日期: 2012-08-29

作者简介: 刘翠玲,女,教授,博士,主要从事食品安全检测方面的研究;

董秀丽,女,硕士研究生,研究方向为食品安全检测技术。通讯作者。

的快速绿色分析方法.

## 1.2 BP神经网络

BP(back propagation)神经网络是多层前馈网络,采用误差反向传播的学习过程,输入信号由输入层接收并传递给隐含层,经隐含层信息变换处理后,传递给输出层,并将此输出与期望输出进行计算得出误差,若误差不在可接受范围内,则将误差反向传播给神经网络,周而复始并以此计算各层的阈值和权值,直至得到可接受的结果,训练结束<sup>[13]</sup>.

BP神经网络的权值训练使用非线性可微函数,并且,只含有一个隐含层的BP神经网络即可对任意连续函数进行任意精度的映射,因此本文亦采用3层BP神经网络进行建模.使用BP神经网络建模不必预知模型的各项参数即可对函数进行拟合,因此,BP神经网络的建模能力和非线性映射能力很强,可以很容易进行复杂模型的建立<sup>[14]</sup>.

由于没有明确的隐含层神经元选择公式和规范,BP神经网络建模的难点之一为隐含层神经元数目的选取,隐含层神经元数目过多会导致模型过于复杂、过拟合和泛化能力差等问题,导致模型的预测效果极差.因此,通过多次调试选择隐含层神经元数目,得到最优模型.

另外,由于训练函数根据自身特点,需多次试验选择出最优训练函数对神经网络进行训练.根据多次实验结果,选择量化共轭梯度(scaled conjugate gradient, SCG)反向传播算法训练函数.

## 2 材料与方法

### 2.1 仪器

VERTEX70 红外光谱分析仪,漫反射积分球附件,德国 Bruker 公司;光谱采集及分析软件使用 OPUS6.5.

### 2.2 试验样品

试验用的小麦面粉是来自北京古船食品有限公司不同批次的面粉,用电子分析天平准确称量,在面粉中分别掺入浓度为 0.4% ~ 20% 的滑石粉,共制备了 1 ~ 50 号共 50 个掺杂样本,并以掺杂后的面粉中滑石粉的浓度作为样本的真值.

### 2.3 光谱采集

仪器工作参数设定为:谱区范围 12 500 ~ 4 000  $\text{cm}^{-1}$ ,扫描次数 32 次;分辨率 8  $\text{cm}^{-1}$ .将上述所配面粉样本放置在旋转台的样品杯中,开启旋转台,室

内扫描温度 25  $^{\circ}\text{C}$ .每次扫描前仪器预热 30 min,以仪器内置参比做背景校正.

## 2.4 建立近红外预测模型的方法

定量分析采用偏最小二乘法(PLS)建立近红外光谱与小麦面粉中滑石粉含量的回归模型.其评价标识有决定系数( $R^2$ )、交叉验证标准差平方根(root mean square error of cross validation, RMSECV)和预测标准差平方根(root mean square error of prediction, RMSEP),其计算公式如下.

$$R^2 = \left[ 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_n)^2}{\sum_{i=1}^n (y_i - y_n)^2} \right] \times 100\%, \quad (1)$$

其中  $n$  为样本数; $\hat{y}_i$  为第  $i$  个样本的预测值; $y_m$  为样本真值的平均值; $y_i$  为第  $i$  个样本的参考值(真值).

RMSECV 和 RMSECP 表示近红外测定值与真值间的相似程度,其计算公式如下.

$$RMSECV = \sqrt{\frac{1}{N-P-1} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (2)$$

其中, $P$  为回归因子数目;

$$RMSECP = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}. \quad (3)$$

## 3 结果与讨论

### 3.1 采集样本的近红外光谱

图 1 为用近红外光谱仪扫描含滑石粉的小麦面粉所得原始近红外光谱图.由图 1 可见,不同配比的掺杂滑石粉的小麦面粉样本的近红外光谱趋势一致性较好,没有光谱异常现象,且因化学值和光谱异常的判断较为复杂,故在此不做异常样品的剔除.

### 3.2 数据与处理方法和模型建立

近红外光谱往往包含一些与待测样品性质无关的因素带来的干扰,导致了近红外光谱的基线漂移和光谱的重复,为了减弱或消除基线漂移、散射等各种非目标因素对光谱的影响,对近红外光谱仪采集的掺杂滑石粉的小麦面粉光谱进行光谱预处理是必要的.本文分别以多元散射校正对光谱进行预处理后,建立定标模型;由于要限制频率范围到 9 000  $\text{cm}^{-1}$ ,故选用波段 9 000 ~ 4 000  $\text{cm}^{-1}$ ,可以达到满意的效果.采用 PLS 对数据进行分析,可以有效滤除噪声和背景对建模的干扰,同时也减少了变量个数和模型的随机性,提高了模型的运算速度,改善了模

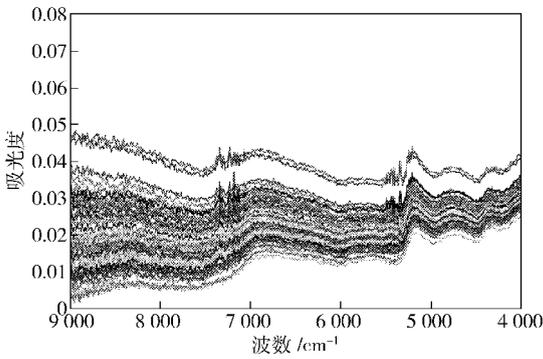


图1 含滑石粉的小麦面粉原始近红外光谱图  
Fig.1 Raw near-infrared spectra of talc-containing wheat flour

型的预测精度和线性相关性. 图2为预处理后的面粉样本光谱图,其中横坐标是波数( $\text{cm}^{-1}$ ),纵坐标是吸光度.

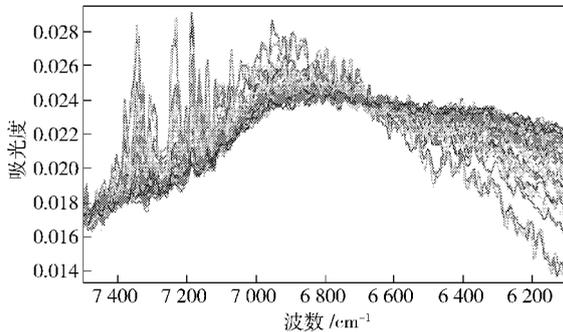


图2 预处理后的面粉样本光谱图  
Fig.2 Flour sample spectra after pretreatment

根据光谱预处理结果,选择  $7\,502.1 \sim 6\,098.1 \text{ cm}^{-1}$  波段的特征吸收峰对面粉中滑石粉建立定量分析模型. 使用 MATLAB R2009a 神经网络工具箱中量化共轭梯度(scaled conjugate gradient, SCG)反向传播算法训练函数(trainscg 训练函数),同样采用以校正集为训练样本对网络进行训练和以平均绝对误差(mean squared error, MSE)为性能函数,隐含层选为11个神经元,建模时对BP神经网络训练500次. 所建立模型的各个评价参数为:  $R^2 = 0.997\,3$ ,  $RMSEC = 0.436\,7$ ,  $RMSEP = 1.708\,8$ . 图3为内部交叉验证拟合结果,横座标代表目标值即真实值  $T$ ,纵坐标代表BP网络输出值即预测值  $Y$ ,线性拟合函数为  $Y = 0.99T + 0.095$ ,圆圈代表真实值,红线代表拟合值,黑线代表  $Y = T$ ;图4为BP神经网络对预测集的预测结果,横座标代表样本、滑石粉掺杂量,纵坐标蓝线代表真实值、绿线代表预测值.

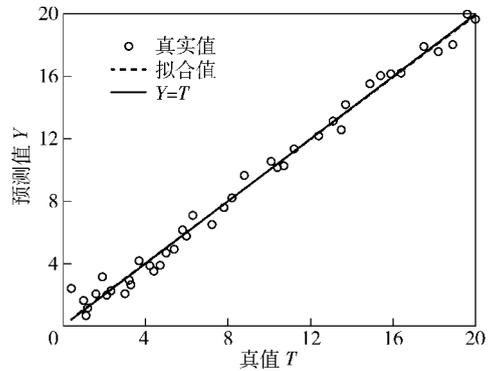


图3 BP神经网络内部交叉验证拟合结果  
Fig.3 BP neural network internal cross-validation fitting results

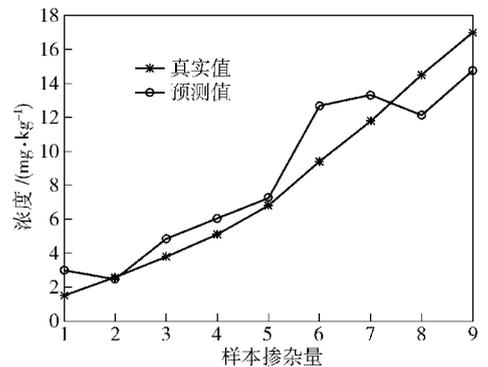


图4 预测集预测结果  
Fig.4 Prediction results

## 4 结论

面粉中滑石粉的定量检测,由于受到仪器精度及稳定性、实验方法设计和实验过程中的操作方法的影响,检测结果往往和预想的结果有较大差距,化学计量学方法在面粉中滑石粉检测技术中的应用是本课题的重要内容. 本文通过对近红外技术光谱数据的预处理,不但放大了特征信息,而且大大减小了后期的建模过程计算量.

结合近红外光谱技术,使用BP神经网络建立面粉中滑石粉的定量分析模型有效利用了近红外光谱技术的快速、无损检测特点和BP神经网络非线性映射能力强、模型概括性和推广型强的优点.

在使用BP神经网络进行建模的过程中,最优训练函数的选取和隐含层神经元数目的选择是本文的难点之一,根据采用不同训练函数和隐含层神经元数的BP神经网络对面粉中滑石粉特点所建立的模型的建模结果为:模型的相关系数在0.99以

上,交叉验证标准差和预测标准差分别为  $R^2 = 0.9973$ ,  $RMSEC = 0.4367$ ,  $RMSEP = 1.7088$ , 得到了较好的评价参数. 试验结果表明:BP神经网络应用于面粉中滑石粉的检测技术是可行的,但是,样本集所包含样本数对BP神经网络建模结果的影响仍需要进一步研究.

#### 参考文献:

[1] Osborne B G. Near-infrared spectroscopy in food analysis [J]. *Cereal Foods World*, 2000, 45(1): 11.

[2] 国家质量监督检验检疫总局. GB/T5055—2008. 粮油检验灰分测定法[S]. 北京:中国标准出版社,2008.

[3] 董福祥. 面粉掺假方法介绍[J]. *宁夏农学院学报*, 2001, 22(4): 75-76.

[4] 薛笑莉,周激. 面粉及面制品中滑石粉-石膏-碳酸钙含量的测定[J]. *山西化工*, 2003, 23(2): 33-34, 61.

[5] 国家质量监督检验检疫总局. GB/T21913—2008. 食品中滑石粉的测定[S]. 北京:中国标准出版社,2008.

[6] 郭光美,王云清. 面粉中掺有滑石粉的快速检测方法[J]. *食品研究与开发*, 2000, 21(1): 40-42.

[7] 王承明,潘峰. 面粉中滑石粉的X射线衍射分析[J]. *分析科学学报*, 2006, 22(6): 651-654.

[8] 金凤,王冬. 面粉中滑石粉的现场快速检测方法及其装置研究[J]. *食品与发酵工业*, 2008, 34(3): 160-163.

[9] 魏峰,霍军生,徐宇洋. 面粉中滑石粉含量的检测[J]. *黄山学院学报*, 2008, 10(5): 48-49.

[10] 由健,刘成雁. 面粉中滑石粉的X射线检测法[J]. *科协论坛*, 2009(1): 85-86.

[11] 杨彦丽,林立. 微波消解-电感耦合等离子发射光谱法测定面粉中滑石粉的含量[J]. *化学分析计量*, 2011, 20(3): 62-64.

[12] 赵延华,韩旭. 傅里叶变换-红外光谱法快速测定面粉中滑石粉[J]. *理化检验化学分册*, 2011, 47(2): 208-210.

[13] MATLAB中文论坛. MATLAB神经网络30个案例分析[M]. 北京:北京航空航天大学出版社,2010: 36-44.

[14] 刘翠玲,索少增. BP神经网络在ATR-FTIR技术微量农药溶液检测中的应用[J]. *北京工商大学学报:自然科学版*, 2011, 29(4): 64-67.

## Rapid Determination of Talc-containing Flour Based on BP Neural Network

LIU Cui-ling, DONG Xiu-li, SUN Xiao-rong, WU Jing-zhu, WU Sheng-nan

(School of Computer Science and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

**Abstract:** Near infrared spectral technology (NIR) was used to test talc-containing wheat flour. The spectrum was preprocessed with multiplicative scatter correction. The quantitative analysis model of talc-containing flour was built using SCG back propagation algorithm training function of BP neural network, and the calibration set and prediction set were quantitatively analyzed.  $R^2$  was 0.9973, the root mean square error of calibration (RMSEC) was 0.4367, and the root mean square error of prediction (RMSEP) was 1.7088. The results showed that BP neural network with NIR for the determination of talc-containing flour has the advantages of fast, high precision, and the ability of Fanhua, and can be used for talc-containing flour.

**Key words:** BP neural network; NIR; talcum powder; wheat flour