第27卷, 第5期
2007年5月

光谱学与光谱分析
Spectroscopy and Spectral Analysis

Vol. 27, No. 5, pp911-915
May, 2007

应用多种近红外建模方法分析梨的坚实度

傅霞萍, 应义斌*, 陆辉山, 于海燕, 徐惠荣

浙江大学生物系统工程与食品科学学院, 浙江 杭州 310029

摘 要 近红外光谱(near infrared spectroscopy, NIRS)技术是一种快速、无损的仪器分析方法, 在农产品品质检测方面引起了广泛的关注, 在近红外光谱信息和品质指标之间建立一个稳健的模型是近红外光谱分析中十分重要且有一定难度的过程, 常见的多元校正方法有偏最小二乘回归(PLSR)、主成分回归(PCR)和逐步多元线性回归(SMLR)等, 该研究中除了常用的线性方法外, 还采用了一种结合非线性方法的组合算法[结合了SMLR和径向基神经网络(RBFN)]用于梨坚实度的近红外光谱检测。比较常用的线性建模方法, 原始光谱的PLSR模型的得到了较好的结果: 校正集相关系数 r= 0.87, 校正均方根误差RMSEC= 3.88 N, 预测集 r= 0.84, 预测均方根误差RMSEP= 4.26 N; 组合算法的建模结果比SMLR和PCR的结果好, 但比PLSR的结果稍差: 校正集 r= 0.85, RMSEC= 4.15 N, 预测集 r= 0.82, RMSEP= 4.67 N。结果表明: NIRS可用于梨的坚实度检测, 但是建模方法的选择值得进一步研究以提高预测的精度。

关键词 NIRS; 多元建模方法; 坚实度; 梨

中图分类号: S37; TP216 文献标识码: A 文章编号: 1000-0593(2007)05-0911-05

引 言

近期曾有过多种方法(例如文献[1, 2])检测水果的坚实度, 但是仍存在了不少缺点。本文提出用多种近红外建模方法分析梨的坚实度, 取得很好的结果。

NIRS[3]分析中最常用的多元线性统计方法主要有逐步多元线性回归(stepwise multiple linear regression, SMLR)、主成分回归(principal component regression, PCR)和偏最小二乘回归(partial least squares regression, PLSR)等, 这些方法在样品的被测参数和光谱吸光度之间建立起一个线性关系, 属于线性建模方法。与此相对的, 人工神经网络(artificial neural networks, ANN)是最常用的非线性建模方法[4,5], 是在样品的被测参数与光谱吸光度之间建立起一个非线性关系。神经网络有很多种不同的训练方法, 径向基函数网络(radial basis function networks, RBFN)是其中一种, 它是由Swierenga, Windig, Davies等建立并发展起来的[6-8], 至今已有不少文献报道了RBFN在化学、光谱数据分析方面的应用[9, 10]。

到目前为止, 也有不少学者在应用近红外光谱技术检测水果坚实度方面作了许多研究。Cho等[11]研究了用NIRS技术检测苹果坚实度, 并认为由于果胶物质的变化及水分在1 900 nm附近的吸收使结果受到一定的影响。Cho等的研究中只用到了1 100~2 400 nm光谱范围, 但Elridge[12]通过研究苹果和桔子果胶的近红外光谱, 确证在980 nm左右有一个果胶的吸收峰。McGlone等[1]应用800~1 100 nm的反射光谱建立了预测猕猴桃坚实度的多元模型, 他们将较差的预测结果归因于水果成分变化与坚实度没有直接的关系。Lu等[13, 14]分别在800~1 700 nm和900~1 500 nm波长范围检测甜樱桃和苹果坚实度, 研究结果表明没有一个单一波长与坚实度有明显的相关性。Park等[15]报道了应用PCR和马氏距离法(mhalanobis distance, MD)分析两个不同品种的苹果的坚实度, 在进行MD分析方法分类苹果坚实度时他们研究了800~1 100 nm和400~1 800 nm两个不同的波长范围, 结果表明进行坚实度分类时较宽的波长范围比较窄的波长范围更精确。

本研究的主要目的是:

(1) 应用不同的建模方法建立近红外漫反射光谱与梨坚实度之间的校正模型;

(2) 比较原始光谱、一阶微分光谱和二阶微分光谱建立的梨坚实度预测模型的性能;

(3) 比较线性和非线性多元统计方法建立的梨坚实度预测模型的性能。

收稿日期: 2006-05-08, 修订日期: 2006-09-23

基金项目: 国家自然科学基金项目(30671197), 教育部"新世纪优秀人才支持计划"基金项目(NCET-04-0524)资助

作者简介: 傅霞萍, 女, 1980年生, 浙江大学生物系统工程与食品科学学院博士生 * 通讯联系人 e-mail: ybying@zju.edu.cn

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

1 材料和方法

1.1 材料

实验所用的雪青梨是在 2004 年 7 月 20 日至 8 月 10 日之间从浙江大学实验果园分 3 次采摘的。前 2 个采收时间分别在正常采收期的前 20 天和前 10 天，最后一次是在梨的正常采收期。采收的梨先经清理和标记，所有实验均在 20 ℃左右室温下进行。本实验共有 144 个样品，其中 101 个用于建立校正模型，剩余 43 个用于模型的验证。

1.2 近红外光谱采集

完整水果的近红外漫反射光谱是用 NEXUS 智能型 FT-IR 光谱仪(尼高力仪器公司，美国) 及其所带的光纤漫反射附件采集的，光谱仪的光源是一个 50 W 的石英卤素灯，检测器选用了 InGaAs(其光谱检测范围为 800~ 2 630 nm)，扫描次数为 64 次，动镜速度为 0.949 4 cm·s^{-1}。由于高于 2 500 nm 的光谱信噪比较低，所以数据分析时只用了 800~ 2 500 nm 的光谱信息。光纤探头是一个 16 mm 直径的不锈钢圆柱形管，入射光纤束和接收光纤束随机地分布其中。

光纤固定在水果支架上，测量时将水果平稳地放置在支架之上，果柄-果蒂轴线呈水平。在水果和光谱探头之间有一个橡皮垫圈，既作为密封圈防止表面反射，同时又作为一个柔性支架以适应不同大小的水果。每个水果在赤道位置间隔 120°的 3 个位置采集 3 条漫反射光谱，并在相应的位置做好标记。在水果光谱采集之前，先要测量标准白板在相同参数设置下的背景光谱作为参比。

1.3 坚实度测量

每个水果的坚实度在 NIR 漫反射光谱采集的相应位置进行测量，用固定在生物材料万能试验机(Instron 5543, USA) 上的一个直径为 6 mm 的探头以 20 mm·min^{-1}的加载速度，通过联接的计算机获取连续的应力变化曲线，曲线反应了每次检测的刺入力和深度的变化，探头在果皮破裂时会自动卸载并回到起始位置，取应变曲线上的应力最大值作为水果的坚实度。

1.4 光谱数据处理和分析

光谱数据的采集和保存通过计算机运行光谱仪自带的 OMNIC6.01(尼高力仪器公司，美国) 软件实现，这个软件同时还能实现其他许多操作如波数与波长之间的转换、光谱平滑处理、光谱微分处理等等。本实验获取的光谱以每一波长下的吸光度值 log 1/R 表示，R 是指样品的反射光强与参比白板的反射光强之比。应用 TQ Analyst(尼高力仪器公司，美国) 软件进行数据的分析及 SMLR，PCR 和 PLS 等线性模型的建立，用 Matlab 软件建立 ANN 非线性模型。

人工神经网络是一个从输入到输出的非线性映射，每个结点(神经元) 通过连接权接收来自输入层结点的信息，通过输入输出转移函数给出输出信息，并根据输出信息的正误程度，通过一定的训练规则对连接权值进行修正[16]，使网络不断朝误差减小的方向进行，直至输出值与目标值的偏差满足工作所需的精度，达到训练的目的，RBFN 是众多训练方法中的一种。RBFN 的结构如图 1 所示，n 和 m 分别表示输入层和隐含层的神经元个数。在本研究中，光谱范围为 800~ 2 500 nm，如果将每个样品在每一个波长下的吸光度视作一个向量，那么所有样品在每一波长的吸光度就构成一个矩阵，这一矩阵含 1 000 多个向量，即 n> 1 000，这么大的输入层神经元数是很不实际的，因而在本研究中 n 由 SMLR 建模结果确定，将 TQ 软件推荐的 2 个波长向量作为网络的输入，建模过程应用 Matlab 软件的 GUI 工具箱实现。

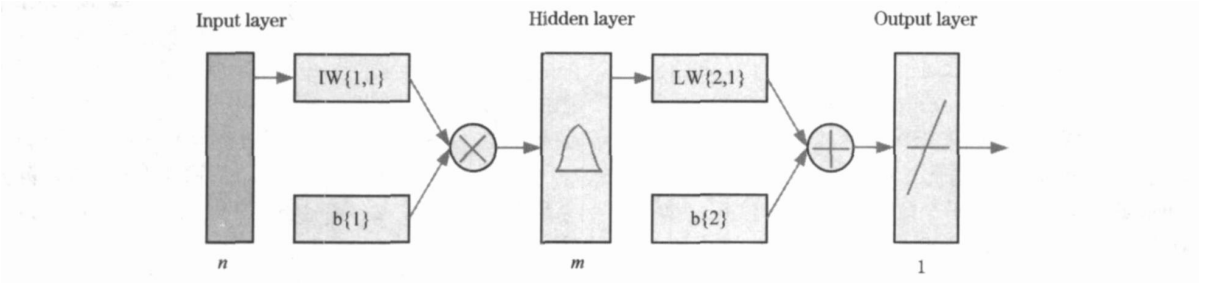

Fig 1 RBFN architecture

2 结果与讨论

2.1 坚实度测量结果

样品坚实度检测结果为：最大值 57.58 N，最小值 22.53 N，平均值 39.44 N，标准偏差 7.85N。表 1 所列为校正集和验证集样品的坚实度检测结果。

2.2 SMLR，PCR 和 PLSR 分析结果

图 2 分别是梨的原始光谱、一阶微分光谱和二阶微分光谱。原始光谱中在 970，1 190，1 450，1 780 和 1 940 nm 五处的宽峰是水的典型吸收峰，代表了水果中水分的吸收特征。

Table 1 Pear firmness measurement results

项目	校正集	验证集
样品数量	101	43
最大值/N	57.58	57.37
最小值/N	22.53	26.25
平均值/N	39.38	39.60
标准偏差/N	7.87	7.89

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

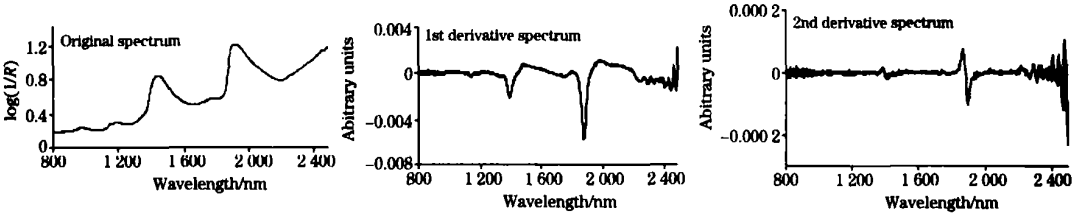

Fig 2　Typical curves of original spectrum, 1st derivative and 2nd derivative spectra of pears

用 SMLR, PCR 和 PLSR 三种方法, 在 800~ 2 500 nm 波长范围分别建立原始光谱、一阶微分光谱和二阶微分光谱的模型, 建模结果如表 2 所示。从表中可以看出原始光谱的建模结果比一阶微分光谱和二阶微分光谱都要好; 三种方法中, PLSR 模型的结果最好, 校正集的相关系数 r= 0. 87, 校正均方根误差 RMSEC= 3 88, 预测集 r= 0. 84, 预测均方根误差 RMSEP= 4 26, 实验结果见图 3。

Table 2　Prediction results of original, first derivative and second derivative spectra using SMLR, PCR and PLSR method

建模方法	光谱类型	校正集(101 个样品)		验证集(43 个样品)	
		r	RMSEC/ N	r	RMSEP/ N
SMLR	原始光谱	0 83	4 40	0 81	4 50
	一阶微分光谱	0 81	4 56	0 75	5 25
	二阶微分光谱	0 77	4 97	0 79	4 54
PCR	原始光谱	0 81	4 59	0 80	4 68
	一阶微分光谱	0 73	5 37	0 77	5 00
	二阶微分光谱	0 72	5 40	0 69	5 69
PLSR	原始光谱	0 87	3 88	0 84	4 26
	一阶微分光谱	0 85	4 15	0 75	5 41
	二阶微分光谱	0 86	4 06	0 72	5 64

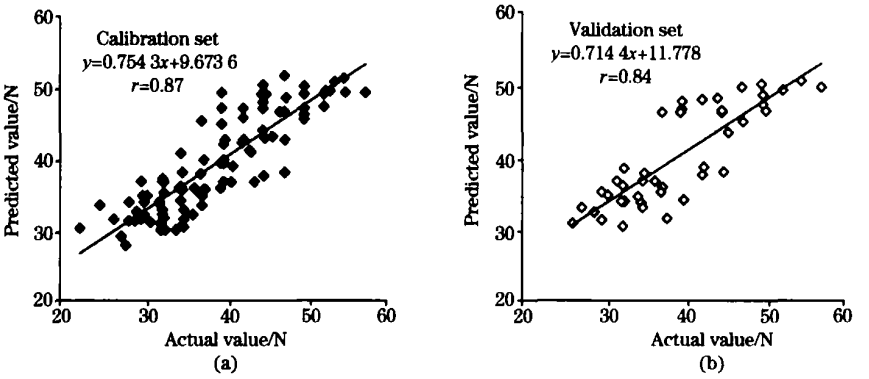

Fig 3　Calibration and validation results of PLSR model for pear firmness prediction

2. 3　ANN 分析结果

本研究还选用了 RBFN 网络进行非线性建模, 这种网络的训练时间较短, 权值也不需要有所限制, 远离判别边界的样品在处理过程中对结果几乎没有影响。RBFN 的输入向量由 SMLR 建模结果决定。用 TQ 软件建立 SMLR 模型时, 软件推荐了 2 个主要的波长向量(1 643 和 1 733 nm) , 将这 2 个向量作为 RBFN 的输入向量。并用 Matlab 的 GUI 工具箱建立 RBFN 模型, 其结果为: 校正集 r= 0 85, RMSEC= 4. 15 N; 验证集 r= 0 82, RMSEP= 4 67 N(见图 4)。

不少报道提出由于 NIRS 技术的一些特点和属性, 非线性方法比线性方法更适用于 NIRS 建模。比较本文中的分析结果, RBFN 的结果比 PLSR 的结果要稍差一些, 这与 RBFN 的输入向量少有极大关系, 在 RBFN 分析中只用了 SMLR 模型推荐的两个波长向量作为网络的输入, 若对网络的输入向量选择方法进行一定的改进, 相信建模结果也会有一定的提高。

3　结　论

应用不同建模方法建立近红外漫反射光谱模型预测雪青梨的坚实度, 根据结果可以认为 NIR 漫反射光谱技术可用于梨的坚实度检测, 但预测精度有待提高。原始光谱的 PLSR

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

模型得到了较好的预测结果：校正集 r= 0.87，RMSEC= 3.88 N；验证集 r= 0.84，RMSEP = 4.26 N。SMLR 和 RBFN 组合算法的建模结果比 SMLR 和 PCR 模型的结果较好，但比 PLSR 模型的结果稍差：校正集 r= 0.85，RMSEC = 4.15 N；r= 0.82，RMSEP= 4.67 N。从结果可知，用多种方法组合进行近红外光谱建模检测梨的坚实度是可行的，但是在神经网络训练方法的选择上以及网络输入向量的选择上可以进行更多尝试，相信会有助于提高模型的预测精度。

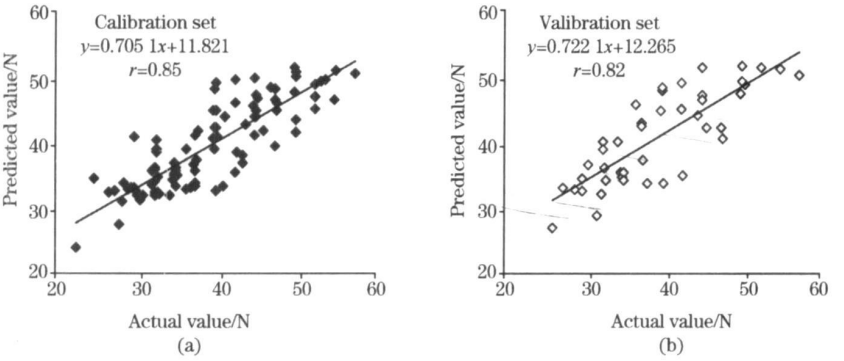

Fig 4 Calibration and validation results of RBFN model for pear firmness prediction

参 考 文 献

[1] McGlone V A, Kawano S. Postharvest Biology and Technology, 1998, 13: 131.
[2] Fidêncio P H, Poppi R J, Andrade J C. Anal. Chim. Acta, 2002, 453: 125.
[3] WANG Duo-jia, ZHOU Xiang-yang, JIN Tong-ming, et al(王多加, 周向阳, 金同铭, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2004, 24(4): 447.
[4] LUO Yi-fan, GUO Zhen-fei, ZHU Zhen-yu, et al(罗一帆, 郭振飞, 朱振宇, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(8): 1230.
[5] BAI Yin-kui, MENG Xian-jiang, DING Dong, et al(白英奎, 孟宪江, 丁 东, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(3): 381.
[6] Swierenga H, Groot P J, Weijer A P, et al. Chemometrics and Intelligent Lab System, 1998, 41: 237.
[7] Windig W, Stephenson D A. Anal. Chem., 1992, 64: 2735.
[8] Davies A M C. NIR News, 1993, 4(4): 10.
[9] Fischbacher C, Jagemann K U, Danzer K, et al. Fresenius' J. Anal. Chem., 1997, 359: 78.
[10] Carlin M, Kavli T, Lillekjendlie B. Chemometrics and Intelligent Lab System, 1994, 23: 163.
[11] Cho R K, Lee K H, Iwamoto M. Spectr. Conf. R. A. Taylor, ed. West Sussex, U.K.: NIR Publications, 1992, 71.
[12] Elridge C D. Int. J. Remote Sensing, 1995, 11(6): 1175.
[13] Lu R. Trans. American Society Agriculture Engineers, 2001, 44(5): 1265.
[14] Lu R, Ariana D. Applied Engineering in Agriculture, 2002, 18(5): 585.
[15] Park B, Abbott J A, Lee K J, et al. Trans. American Society Agriculture Engineers, 2003, 46(6): 1721.
[16] Ramadan Z, Hopke P K, Johnson M J, et al. Chemometrics and Intellignet Lab Systems, 2005., 75(1): 23.

Application of Some Different Modeling Algorithms to Pear MT-Firmness Detection Using NIR Spectra

FU Xia-ping, YING Yi-bin*, LU Hui-shan, YU Hai-yan, XU Hui-rong

College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310029, China

Abstract Near infrared (NIR) spectroscopy is an instrumental method, which was widely studied and used for rapid and nondestructive detection of internal qualities of agricultural products. Statistical modeling is a very important and difficult process in NIR detection to establish the relationship between nondestructive NIR spectral data and interested quality index of the products. Classical multivariate calibration methods such as partial least square regression (PLSR), principle component regression

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

(PCR), stepwise multilinear regression (SMLR) were often used for modeling. In the present study, besides these algorithms, another mixed algorithm was adopted for establishing a nonlinear model of NIR spectra and Magness Taylor(MT) firmness of "Xueqing" pears. The mixed algorithm was combined with SMLR and artificial neural network (ANN). NIR diffuse reflectance spectra of intact pears were measured in the spectral range of 800-2 630 nm using InGaAs detector. However, only spectral information between 800 and 2 500 nm was used for modeling because of the low signal to noise ratio beyond 2 500 nm. Comparing the classical multivariate calibration methods of PLSR, PCR and SMLR, the modeling results using PLSR method were much better than the other two methods. Moreover, models based on original spectra turned out better results than models based on derivative spectra for all the three methods. The best results were r= 0 87, RMSEC= 3. 88 N of calibration and r= 0 84, and RMSEP= 4. 26 N of validation by using PLSR method based on original spectra. The mixed algorithm also performed better than SMLR and PCR, but was a bit worse than PLSR: r= 0 85, RMSEC= 4 15 N of calibration and r= 0 82, and RMSEP= 4 67 N of validation. The results indicated that fruit NIR spectra could be used for MT-firmness prediction when a proper algorithm was chosen, however, further study on statistic modeling is still necessary to improve the predicting performance.

Keywords NIRS; Multivariate calibration methods; Firmness, Pears

(Received May 8, 2006; accepted Sep. 23, 2006)

* Corresponding author

《光谱学与光谱分析》2007 年征订启事

欢迎投稿　　欢迎订阅

《光谱学与光谱分析》1981 年创刊，国内统一刊号：CN 11-2200/O4，国际标准刊号：ISSN 1000-0593，CODEN 码：GYGFED，国内外公开发行，大 16 开本，208 页，月刊；是中国科协主管，中国光学学会主办，钢铁研究总院、中国科学院物理研究所、北京大学、清华大学共同承办的学术性刊物。北京大学出版社出版，每期售价 30. 00 元，全年 360 元；国内邮发代码 82-68，国外发行代码 M905。刊登主要内容：激光光谱测量、红外、拉曼、紫外、可见光谱、发射光谱、吸收光谱、X 射线荧光光谱、激光显微光谱、光谱化学分析、国内外光谱化学分析领域内的最新研究成果、开创性研究论文、学科发展前沿和最新进展、综合评述、研究简报、问题讨论、书刊评述。

《光谱学与光谱分析》适用于冶金、地质、机械、环境保护、国防、天文、医药、农林、化学化工、商检等各领域的科学研究单位、高等院校、制造厂家、从事光谱学与光谱分析的研究人员、高校有关专业的师生、管理干部。

《光谱学与光谱分析》为我国首批自然科学核心期刊，中国科协优秀科技期刊，中国科协择优支持基础性、高科技学术期刊，中国科技论文统计源刊，"中国科学引文数据库"，"中国物理文摘"，"中国学术期刊文摘"，同时被国内外的 CSCI，SCI，AA，CA，EI，PX，MEDLINE 等文献机构收录。根据国家科技部信息研究所发布信息，中国科技期刊物理类影响因子及引文量《光谱学与光谱分析》都居前几位。欢迎国内外厂商在《光谱学与光谱分析》发布广告（广告经营许可证：京海工商广字第 8094 号）。

《光谱学与光谱分析》的主编为黄本立院士。

欢迎新老客户到全国各地邮局订阅，若有漏订者可直接与光谱学与光谱分析期刊社联系。

联系地址：北京市海淀区学院南路 76 号，光谱学与光谱分析期刊社

邮政编码：100081

联系电话：010-62181070，62182998

电子信箱：chngpxygpfx@vip.sina.com；修改稿专用邮箱：mengzh1018@vip.sina.com

网　　址：http://www.gpxygpfx.com

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net