

近红外光谱法对茶叶中游离氨基酸含量的快速测定

董二慧^{1,2}, 谭红², 杨鸿波², 赵红波^{1,2}, 马凯^{1,2}, 袁鑫^{1,2}

(1. 贵州大学 化学与化工学院, 贵州 贵阳 550002; 2. 贵州省理化测试分析研究中心, 贵州 贵阳 550002)

摘要: 为建立一种快速检测茶叶中游离氨基酸含量的方法, 应用近红外光谱分析技术对贵州省不同产地的 162 个绿茶样品中的游离氨基酸含量进行检测, 并结合偏最小二乘法(PLS), 建立茶叶中氨基酸含量的预测模型。结果表明: 通过标准归一化(SNV)光谱预处理, 光谱范围选择 $6\ 657.74 \sim 6\ 178.16, 5\ 793.11 \sim 5\ 688.98, 4\ 493.33 \sim 4\ 350.62\ \text{cm}^{-1}$, 主因子数为 11, 得到模型的内部交互验证相关系数(R)为 0.981 56, 交互验证均方差($RMSECV$)为 0.501; 模型的预测值与实测值的相关系数为 0.998, 预测均方差($RMSEP$)为 0.312, 稳定性试验得到的相对标准偏差(RSD)小于 0.4%。综合分析, 傅里叶变换红外光谱法与化学计量学方法相结合可以实现茶叶游离氨基酸的快速检测, 满足检测要求。

关键词: 近红外光谱法; 游离氨基酸; 偏最小二乘法; 模型

中图分类号: S571.1 O657.33 **文献标志码:** A **文章编号:** 1004-3268(2012)10-0044-05

Rapid Determination of Free Amino Acids in Tea by NIR Spectroscopy

DONG Er-hui^{1,2}, TAN Hong², YANG Hong-bo², ZHAO Hong-bo^{1,2},
MA Kai^{1,2}, YUAN Xin^{1,2}

(1. College of Chemical Engineering and Chemistry, Guizhou University, Guiyang 550002, China;

2. Guizhou Research Center of Physical Test and Chemical Analysis, Guiyang 550002, China)

Abstract: In order to establish a method for rapidly detecting free amino acids of tea, 162 green tea samples of different origins in Guizhou province were detected by near infrared spectroscopy (NIR). Combining with PLS, a mathematical model of amino acids in green tea was set up. According to the standard normal variate processing method, the selected spectra ranged between $6\ 657.74 \sim 6\ 178.16, 5\ 793.11 \sim 5\ 688.98, 4\ 493.33 \sim 4\ 350.62\ \text{cm}^{-1}$. The number of principal factor was 11. The correlation coefficient of the model and the root mean square error of cross validation ($RMSECV$) were 0.981 56 and 0.501, respectively. Moreover the correlation coefficient and the root mean square error of prediction ($RMSEP$) between predicted value of the model and chemical determination value were 0.998 and 0.312, respectively, and the RSD was below 0.4% in the stability test. The result showed that the combination of FTIR spectroscopy and chemometrics was proved to be quick and reliable for free amino acid detection, which would satisfy the requirements of detection.

Key words: near infrared spectroscopy; free amino acid; partial least squares; model

中国是茶的故乡, 人类发现茶叶至今已经有 5 000 多年历史, 研究茶叶成分并利用茶使之造福于人类是现代科技发展的目的。随着人们生活水平的

提高, 饮茶、品茶已经成为很多人生活的一部分, 茶叶已被公认为是世界上最具有生命力的纯天然饮料^[1]。茶叶中有 400 多种成分, 其中氨基酸、咖啡碱

收稿日期: 2012-05-25

基金项目: 贵州省科技支撑计划项目(黔科合 NY 字[2009]3021); 贵州省国际合作计划(黔科合外 G 字[2009]700131)

作者简介: 董二慧(1986-), 男, 山东滕州人, 在读硕士研究生, 研究方向: 化学分析与检测。E-mail: dong0613goole@163.com

和茶多酚是主要成分,与茶叶品质有很大的关系,影响着茶色、茶味和茶香的形成^[2]。其中,游离氨基酸是茶叶中一类重要的含氮物质^[3],它不仅是合成茶叶蛋白的基本单位,还是合成许多与代谢有关的生理活性物质的基质,而且与茶叶品质密切相关^[4]。因此,快速、可靠的氨基酸检测技术逐渐受到大家的重视。传统的氨基酸总量检测方法有酮量法和茚三酮比色法^[5],而在实际应用中发现,采用这 2 种方法得到的数据重复性较低,结果差异较大^[6]。近年来,随着近红外光谱(NIR)分析技术的发展,以及其在食品安全检测中表现出的快速、准确、高效、不损伤样品以及重现性好等优点,已使得近红外光谱分析技术广泛应用于各行业的检测中^[7]。鉴于此,利用近红外光谱分析技术,对茶叶中游离氨基酸含量进行快速检测,比较不同预处理方法对模型效果的影响,建立茶叶游离氨基酸的快速预测模型,以期建立一种快速检测茶叶品质相关指标的有效方法。

1 材料和方法

1.1 仪器

近红外光谱仪(ANTARIS II 型,美国 THERMO FISHER 公司)以及配套的 Result 操作软件和 TQ Analyst 光谱分析软件。仪器工作参数:扫描范围 12 000~4 000 cm^{-1} ,分辨率 8 cm^{-1} ,扫描次数 64 次(所需时间约为 30 s),测量方式为漫反射式。

1.2 样品收集与制备

试验中选取的绿茶包括贡茶、一级、二级和三级 4 个档次,绿茶样品产自贵州省茶叶主产区都匀、贵定、湄潭、凤岗等,总共 162 份,收集的茶叶品种包括贵州所有绿茶品种,具有很好的代表性。分别取适量样品,按照农业部标准(GB/T8303-2002)通过孔径为 600~1 000 μm 的筛,过筛样品置冰箱中密封保存。采用“取五留一法”^[8]将样品每 5 个分为 1 组,取其中间含量样品作为验证集,其余为定标集,共选出 27 个验证集样品,135 个定标集样品(剔除 6 个异常样品)。

1.3 游离氨基酸的化学测定

茶叶中游离氨基酸含量采用茚三酮比色法^[9]测定。为了保证所测数据的准确性,所有样品的测定均重复 6 次,结果取其平均值。

1.4 光谱采集

扫描前,仪器需预热 2 h,自检和仪器校正通过后方可正常使用。将茶叶样品从冰箱中取出,密封下使其恢复至室温。扫描过程中室温保持 25 $^{\circ}\text{C}$ 左右,并严格控制室内湿度,保持环境的一致性。绿茶

近红外光谱的采集采用漫反射方式。样品光谱分 3 次不同时间测量,时间间隔为 2 d,每次光谱采集 3 次,最后样品光谱取 9 次的平均值,以在一定程度上消除因时间变化和样品不均匀造成的影响。

1.5 数据处理方法

光谱数据经标准归一化(standard normal variate, SNV)、Savitzky-Golay 卷积平滑和微分预处理后采用偏最小二乘法(partial least square, PLS)建立定量校正模型。在建模过程中,以定标集样品的交互验证均方差(RMSECV)及相关系数(R)为指标来优化模型,以验证集样品的预测均方差(RMSEP)和相对标准偏差(RSD)^[10]为指标考察模型的预测准确度。其中,

$$R = \sqrt{1 - \frac{\sum(\bar{C}_i - C_i)^2}{\sum(C_i - \bar{C})^2}}$$

$$RMSECV = \sqrt{\frac{\sum(\bar{C}_i - C_i)^2}{n-1}}$$

$$RMSEP = \sqrt{\frac{\sum(\bar{C}_i - C_i)^2}{m}}$$

$$RSD = RMSEP / \bar{C}_m \times 100\%$$

式中, \bar{C}_i 是样品实测值, C_i 是样品预测值, \bar{C} 是全部样品的实测平均值, \bar{C}_m 是验证集样品实测平均值, n 是定标集样品数, m 是验证集样品数。

2 结果与讨论

2.1 茶叶样品中游离氨基酸的测定结果

茶叶中的游离氨基酸在 pH 值 8.0 条件下与茚三酮共热,形成紫色络合物,用分光光度法在特定的波长下测定其含量。从图 1 可以看出,氨基酸与茚三酮形成的紫色络合物在可见光谱区有最大吸光度,其最大吸收波长为 570 nm。从表 1 可以看出,定标集中样品的游离氨基酸含量范围包含任一样品含量的变化范围,且在整个变化范围内是均匀分布的,因此可以满足定量分析的要求。

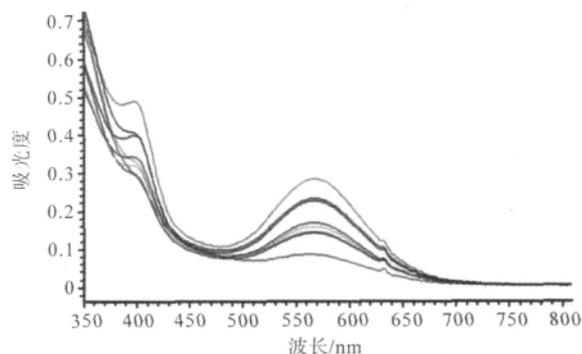


图 1 绿茶样品中游离氨基酸的紫外光谱

表 1 绿茶样品中游离氨基酸含量

类别	样品数/份	范围/%	平均值/%	标准差/%
全部样	162	1.82~3.91	2.51	0.28
定标集	129	1.82~3.91	2.51	0.30
验证集	27	1.96~3.80	2.51	0.25

注:定标集原为 135 个样品,此处为剔除 6 个异常样品后的样品数。

2.2 茶叶样品的近红外漫反射光谱

从图 2 可以看出,光谱曲线大致呈 3 个台阶形状,将其划分为 3 个波段^[11],分别对应于 3 个台阶,即 I:10 000~7 143 cm^{-1} ; II:7 143~5 376 cm^{-1} ; III:

5 376~4 000 cm^{-1} 。其中 I 波段是近红外光谱的二级倍频谱区,其内的 8 400 cm^{-1} 是甲基、亚甲基 C-H 伸缩振动的二级倍频峰; II 波段是近红外光谱的一级倍频谱区,区间的 6 860 cm^{-1} 和 5 780 cm^{-1} 分别是 N-H 键和 C-H 键伸缩振动的一级倍频峰,它主要反映了葡萄糖、纤维素等的分子结构信息; III 波段是近红外光谱的合频谱区,信息比前 2 个区要强,它主要反映了茶叶中氨基酸分子的结构信息,其内的 4 690 cm^{-1} 是氨基酸分子 N-H 键和 C=O 键的合频谱峰,4 330 cm^{-1} 是亚甲基 C-H 伸缩振动和弯曲振动的合频谱^[12]。

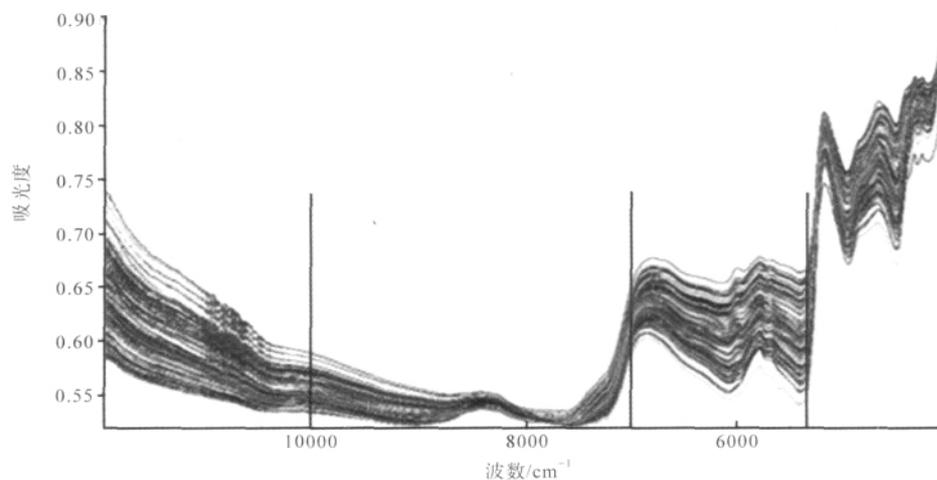


图 2 绿茶样品的近红外光谱

2.3 光谱预处理方法的选择

近红外光谱分析往往不可避免地受到样品状态、光的散射、杂射光及仪器响应等与待测样品性质无关因素的干扰,导致近红外光谱的基线漂移和光谱的不重复,通过光谱数据预处理,可以消除偏移或基线的变化,从而保证光谱数据和组成(性质)之间很好的相关性^[13]。光谱的标准归一化是光谱数据预处理方法之一,可消除光程长短或样品浓度

对光谱响应值的影响^[14]。通过对原始光谱数据标准正态化处理,可扣除样品间因散射而引起的光谱误差;同时,通过对光谱数据进行平滑和求微分处理,来提高分析信号的信噪比,消除基线漂移,强化谱带特征,克服谱带重叠^[13]。从表 2 可以看出,原始光谱经二阶导数和 Savitzky-Golay 卷积平滑预处理后, RMSECV 值最小, R 值最大,模型效果最好。

表 2 不同光谱预处理对氨基酸模型效果的影响

预处理方法	R	RMSECV	主因子数
SNV	0.851 45	0.826	3
SNV+1 st Der+Savitzky-Golay	0.760 42	0.915	16
SNV+2 nd Der+Savitzky-Golay	0.981 56	0.551	11

注:1st Der 表示一阶导数,2nd Der 表示二阶导数。

2.4 近红外光谱波长的选择

游离氨基酸是既含氨基(-NH₂)又含羧基(-COOH)的有机化合物,区别于结合氨基酸,特征基团是 -COOH 和 -NH₂,一级倍频分别在 6 661.67~6 172.84 cm^{-1} 和 5 885.82~5 455.54 cm^{-1} ,在手动优选谱区时可加以考虑。

试验采用了全谱区、软件自动优化谱区及波谱解析加手动优化谱区,通过三者比较确定最优波长范围。从表 3 可以看出,通过对谱区进行自动和手动优化之后,模型效果有了明显提高。全谱区和自动优化谱区的建模效果都不佳,因此,建模的谱区范围为 6 657.74~6 178.16, 5 793.11~5 688.98, 4 493.33~4 350.62 cm^{-1} 。

表 3 不同谱区对氨基酸模型效果的影响

选取方式	谱区范围/ cm^{-1}	R	RMSECV	主因子数
全谱区	11 000.00~3 800.00	0.903 75	0.921	5
自动优化	5 793.11~5 688.98,4 493.33~4 350.62	0.851 45	0.826	3
自动和手动优化	6 657.74~6 178.16,5 793.11~5 688.98,4 493.33~4 350.62	0.981 56	0.501	11

2.5 主因子数的选择

偏最小二乘法的基本思想是对光谱矩阵和茶叶中的游离氨基酸含量矩阵同时进行分解,并在分解时考虑两者之间的关系,加强对应计算关系,从而保证获得最佳的校正模型^[13],因此 PLS 在近红外光谱分析中得到了广泛的运用。但是,在建立定量模型时,为了避免欠拟合或过拟合现象的发生,需要对 PLS 所用的主因子数进行合理选择^[15]。试验中采用 PLS 模型统计方法和 SNV+2nd Der+Savitzky-Golay 光谱预处理方法,通过改变主因子数分析其对模型的影响。从图 3 可以看出,当主因子为 11 时,模型的交互验证均方差最小,效果最好。

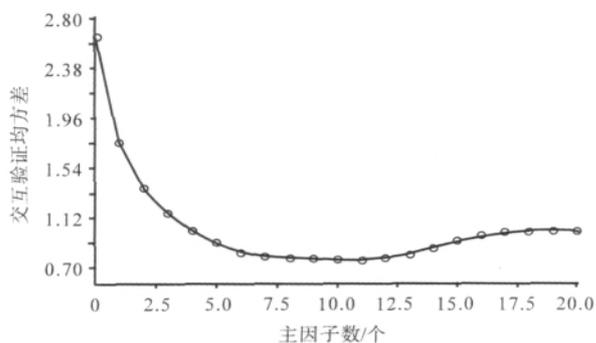


图 3 不同主因子数对氨基酸模型效果的影响

2.6 建立数学模型

综上所述,确定的绿茶中游离氨基酸数学模型的各参数为:

- ① 化学统计方法:偏最小二乘法(PLS);
- ② 光谱预处理方法:SNV+2nd Der+Savitzky-Golay;
- ③ 谱区范围:6 657.74~6 178.16,5 793.11~5 688.98,4 493.33~4 350.62 cm^{-1} ;
- ④ 主因子数为 11。

由以上各参数所作的数学模型如图 4 所示。从图 4 可以看出,模型的内部交互验证相关系数(R)为 0.981 56, RMSECV=0.501,表明氨基酸模型较好。

2.7 数学模型的可靠性评价

近红外光谱分析方法通过数学模型预测样品的待测量,因此,该项技术的分析精确度等特征决定于

数学模型的预测可靠性,试验中可通过样品预测值与实测值的偏差来评价所建立的数学模型的可靠性。选择预测验证集中的 10 个样品,其实测值与预测值如表 4 所示,由表 4 可以看出,样品的预测值与实测值相差不大,通过计算预测值与实测值的相关系数(R)为 0.998,预测均方差(RMSEP)为 0.312。由此表明,所建立的数学模型具有一定的可靠性,能够应用于茶叶中游离氨基酸含量的快速测定。

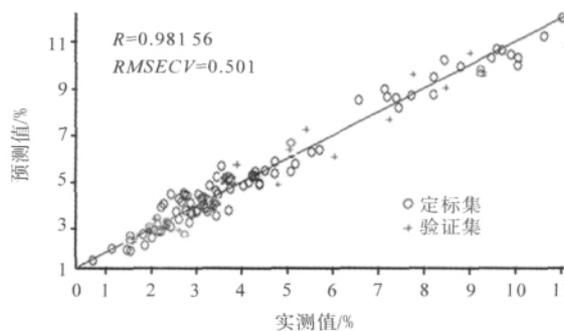


图 4 游离氨基酸模型拟合

表 4 茶叶验证集游离氨基酸的预测值与化学测定值比较

序号	实测值/%	预测值/%	偏差/个百分点
1	1.98	2.01	0.03
2	2.01	2.04	0.03
3	1.99	2.11	0.12
4	2.81	2.88	0.07
5	3.01	3.02	0.01
6	2.51	2.51	0.00
7	2.60	2.70	0.10
8	2.72	2.73	0.01
9	1.96	2.03	0.07
10	2.66	2.71	0.05

2.8 稳定性试验

模型的稳定性指数学模型对样品的类型或其背景信息的适应性,取决于建模样品集所覆盖范围大小,这就需要从大量样品中选择具有代表性的样品才能实现。在稳定性试验中,选择具有代表性的样品 5 个,重复测量 10 次,得到各样品的相对标准偏差(RSD)均小于 0.4%。说明所建数学模型的稳定

性较好,该模型能够适用于一定范围内的样品测量。

3 结论

本试验应用近红外光谱分析技术,对采自贵州省不同地区的 162 个样品进行红外扫描,实现了茶叶中游离氨基酸含量的快速测定。采用 PLS 法建立了茶叶中游离氨基酸含量的分析模型,模型相关系数(R)为 0.981 56, $RMSECV$ 为 0.501,主因子数为 11,此时的校正结果和预测结果都很好;可靠性试验中,10 个样品的预测值与实测值的 R 为 0.998, $RMSEP$ 为 0.312,说明模型的预测效果较好,可以满足茶叶游离氨基酸含量检测的要求;稳定性试验得到的 RSD 均小于 0.4%。因此,红外光谱与化学计量学方法相结合,建立茶叶游离氨基酸含量预测模型在实际应用中是可行的,并且结果稳定可靠,为茶叶的客观评审提供了一种快捷准确的新途径。

参考文献:

- [1] Wang L, Xu R J, Hu B, *et al.* Analysis of free amino acids in Chinese teas and flower of tea plant by high performance liquid chromatography combined with solid-phase extraction[J]. *Food Chem*, 2010, 123: 1259-1266.
- [2] Thomas J, Mandal A K A, Kumar R, *et al.* Role of biologically active amino acid formulations on quality and crop productivity of tea (*Camellia sp.*) [J]. *Int J Agric Res*, 2009, 4(7): 228-236.
- [3] Hung Y T, Chen P C, Richie L C, *et al.* Sequential determination of tannin and total amino acid contents in tea for taste assessment by a fluorescent flow-injection analytical system[J]. *Food Chem*, 2008, 118: 876-881.
- [4] Zhang G. Amino acid species and leaching characters in tea (*Camellia sinensis*) using ion chromatography with integrated amperometric detection[J]. *Int J Food Eng*, 2011, 7(5): 1-12.
- [5] 倪君, 须海荣. 茶叶氨基酸检测方法的研究进展[J]. *茶叶*, 2007, 33(2): 63-66.
- [6] Williams P, Norris K. Near-infrared technology in the agricultural and food industries [M]. Minnesota (USA): American Association of Cereal Chemists, 2001.
- [7] 王志岚, 陈华才, 尹军峰. 近红外光谱分析技术在茶叶中的应用[J]. *现代农业科技*, 2008(15): 110-111.
- [8] 徐玮, 谭红, 包娜, 等. 近红外光谱法快速测定白酒中的酒精度[J]. *中国农学通报*, 2010, 26(2): 1-3.
- [9] 中华全国供销合作总社杭州茶叶研究所. GB/T 8314—2002 茶游离氨基酸总量测定[S]. 北京: 中国标准出版社, 2002.
- [10] 孟兆芳, 赵龙莲, 程奕, 等. 近红外光谱法测定玉米品质指标的研究[J]. *华北农学报*, 2008, 23(2): 147-150.
- [11] 郑咏梅, 张军, 李荣富, 等. 小麦近红外特征波长提取及蛋白质含量测定[J]. *激光与红外*, 2003, 33(2): 125-127.
- [12] 徐立恒, 吕进, 林敏, 等. 茶叶中 3 类主要组分的近红外光谱分析作为茶叶质量的快速评定方法[J]. *理化检测: 化学分册*, 2006, 42(5): 334-336.
- [13] 陆婉珍. 现代近红外光谱分析技术[M]. 北京: 中国石化出版社, 2007.
- [14] 吴瑾光. 近代傅里叶变换红外光谱技术及应用[M]. 北京: 科学技术文献出版社, 1994.
- [15] 严衍禄, 赵龙莲. 近红外光谱分析基础与应用[M]. 北京: 中国轻工业出版社, 2005.