

## 基于近红外光谱的安吉白茶溯源模型的建立与验证

廖金燕<sup>1</sup>, 洪雪珍<sup>2</sup>, 付贤树<sup>3</sup> (1. 浙江育英职业技术学院商务贸易分院, 浙江杭州 310018; 2. 中国计量大学质量与安全工程学院, 浙江杭州 310018; 3. 中国计量大学生命科学学院, 浙江杭州 310018)

**摘要** [目的]研究安吉白茶源产地追溯方法,建立高识别率追溯方法。[方法]对266个来自7大源产地安吉白茶和118个来自六大茶叶主产区的非安吉白茶进行近红外光谱检测,预处理方法分别采用SNV和MSC,通过LDA模型识别率比较二者优劣,并通过PCA、K-means聚类法和HCA聚类方法进行模型验证。采用SPSS分析聚类结果与样品地理位置的关系。[结果]MSC预处理效果优于SNV,LDA分析模型识别率可达100%。模型验证显示,经MSC预处理后,PCA方法可在前3个PC下清晰聚类,K-means和HCA聚类方法都可将266个产地内和118个产地外样品成功预测出。SPSS聚类分析显示,7个安吉白茶的聚类结果与样品采集地的地理位置远近存在对应关系。[结论]安吉白茶近红外光谱经MSC预处理,LDA建模可作为安吉白茶源产地追溯方法。

**关键词** 安吉白茶;近红外光谱;源产地追溯;数据划分;模型验证

**中图分类号** S126 **文献标识码** A

**文章编号** 0517-6611(2020)13-0224-05

**doi**: 10.3969/j.issn.0517-6611.2020.13.060



开放科学(资源服务)标识码(OSID):

### Establishment and Verification of Traceability Model of Anji White Tea Based on Near Infrared Spectroscopy

LIAO Jin-yan<sup>1</sup>, HONG Xue-zhen<sup>2</sup>, FU Xian-shu<sup>3</sup> (1. Business and Trade Branch, Zhejiang Yuying College of Vocational Technology, Hangzhou, Zhejiang 310018; 2. College of Quality and Safety Engineering, China Jiliang University, Hangzhou, Zhejiang 310018; 3. College of Life Sciences, China Jiliang University, Hangzhou, Zhejiang 310018)

**Abstract** [Objective] To study traceability model of Anji white tea and to establish a high origin tracing method. [Methods] 266 real and representative Anji white tea from 7 tea plantations in their original producing areas, as well as 118 non-Anji white tea samples of similar appearance from 6 important tea producing areas, were collected and detected by NIR. For data pretreatment, SNV and MSC were respectively applied to improve the raw spectra. The advantage and disadvantage of the two were compared by the recognition rate of LDA model. Model validation was subsequently verified by PCA, K-means and HCA clustering. The SPSS clustering analysis was used to analyze the relationship between clustering results and geographical location of samples. [Results] The effect of MSC pretreatment was significantly better than that of SNV. After MSC pretreatment, the recognition rate of LDA analysis model could be up to 100%. Model verification demonstrated that PCA method with MSC spectra could clear cluster under the first three PCs. For the clustering models, K-means and HCA clustering methods could successfully predict 266 Anji white tea and 118 non-Anji white tea from 384 samples. According to the SPSS clustering analysis, there was positive correlation between the clustering results of Anji white teas and the geographical locations of the samples. [Conclusions] The proposed method, which combined NIR and LDA with MSC spectra, can be a feasible identification method for Anji white tea.

**Key words** Anji white tea; Near infrared (NIR); Origin tracing; Data splitting; Model validation

安吉白茶产自浙北安吉县<sup>[1]</sup>,是典型地理标志产品,虽冠以“白茶”之名,但安吉白茶却是绿茶。茶叶的品质和特征主要受到茶叶种植地理位置的影响<sup>[2]</sup>,由于经纬度、光照、气候、温度、湿度等影响,不同种植地的茶叶口感、香味等有很大差异。安吉白茶由于其独特的口感而广受消费者的喜爱,2014年后,安吉县将1.13万hm<sup>2</sup>安吉白茶茶园进行分级保护<sup>[3]</sup>,这对保障安吉白茶原产地的真实性具有重要作用,但随着安吉白茶销售量的剧增,掺杂、掺假及以次充好等不良现象屡屡出现,使得安吉白茶的声誉受到严重影响。

目前,常用的茶叶源产地追溯方法主要有感官评价、滋味因子分析(电子舌、电子鼻等)、气相色谱法、毛细管电泳法等<sup>[4]</sup>。但是这些方法有操作复杂、仪器昂贵、检测时间长、预处理麻烦、灵活度低等缺点。近红外光谱技术是近年发展起来的一种新型特征指纹检测技术,具有快速、无损、低成本等优点。其原理是C—H、O—H、N—H等氢原子团的振动包含了样品分子结构和成分信息,通过倍频、频合、频差叠加等方法检测氢原子团的振动,从而确定样品的物理性质和化学

成分<sup>[5]</sup>。

应用化学计量学对近红外光谱进行建模分析,已广泛应用于产品化学成分含量测定<sup>[6-7]</sup>、质量级别判定<sup>[8]</sup>、品种判定<sup>[9]</sup>及产地鉴别<sup>[10-20]</sup>等方面。现有文献报道显示,近红外光谱技术在茶叶源产地追溯上的研究对象主要是绿茶、黑茶和乌龙茶,2018年绿茶的全国产量约172万t,占茶叶总产量的64.9%,而目前绿茶研究主要针对的是西湖龙井、碧螺春、信阳毛尖等家喻户晓的地标产品,对于安吉白茶研究甚少,安吉作为习总书记“绿水青山就是金山银山”科学论断的发祥地,对作为排头兵的安吉白茶进行源产地追溯研究尤为显得重要。

笔者以产地内、外安吉白茶为研究对象,基于不同预处理方法和模式识别算法对安吉白茶进行源产地追溯分析,茶样近红外光谱采用标准正态变换(Standard Normal Variation, SNV)和多元散射校正(Multiplicative Scatter Correction, MSC)进行预处理,并采用线性判别分析(Linear Discriminant Analysis, LDA)进行建模分析判断SNV和MSC的优劣;然后,采用主成分分析(Principal Component Analysis, PCA)、K-means聚类分析和层次聚类(Hierarchical Cluster Analysis, HCA)分析等不同分类算法进行模型验证;最后,采用SPSS分析聚类结果与样品地理位置的关系。

**基金项目** 浙江省公益研究计划项目(LGN19C160004);杭州市农业与社会发展科研项目(20191203B23)。

**作者简介** 廖金燕(1977—),女,湖南岳阳人,讲师,硕士,从事企业质量管理研究。

**收稿日期** 2020-04-10

## 1 材料与方 法

**1.1 材料与试剂** 以安吉白茶为对象,共采摘 384 个安吉白茶样品,其中正宗安吉白茶样品 266 个,采集于安吉县 7 个地区,分别为高禹镇、递铺镇、上墅乡、昆铜乡、良朋镇、溪龙乡、梅溪镇;安吉县外非安吉白茶样品 118 个,采集于 6 个地区,分别为安徽广德、浙江长兴、湖南永州、浙江湖州、江西瑞金、浙江文成。样品具体情况如表 1 所示。

表 1 第 1 批样品采集汇总

Table 1 Summary of the samples of the first batch

产地 Production area	标记 Mark	地区 Area	组别 Group	样品数 Sample number
产地内 Inside production area	IN/0	递铺镇	1	74
		高禹镇	2	15
		昆铜乡	3	72
		良朋镇	4	15
		梅溪镇	5	15
		上墅乡	6	15
		溪龙乡	7	60
		(小计)	—	266
产地外 Out- side production area	OUT/1	安徽广德	1	13
		浙江长兴	2	45
		湖南永州	3	15
		浙江湖州	4	15
		江西瑞金	5	15
		浙江文成	6	15
		(小计)	—	118

**1.2 仪器与设备** Tensor37 FT-NIR 光谱仪(配备 InGaAs 检测器)购自德国 Bruker 公司。

### 1.3 方法

**1.3.1 样品原始光谱采集方法。**白茶样品无需处理,直接装于石英管中进行检测。光谱波数为 12 000~4 000  $\text{cm}^{-1}$ ,分辨率为 8  $\text{cm}^{-1}$ ,扫描间隔设置为 1.928  $\text{cm}^{-1}$ ,扫描 64 次,每条白茶样品原始光谱中共有 4 148 个数据点,因此原始数据集的大小为 384 个样品 $\times$ 4 148 个变量。所有测量均在(25 $\pm$ 1)  $^{\circ}\text{C}$ 室温下进行。

**1.3.2 train-test-split 数据划分方式。**采用随机划分方式为 train-test-split 数据划分方式,随机划分方式即随机选取产地内和产地外样品的 70%为训练集,剩余 30%为预测集,用于验证模型的准确性(表 2)。

表 2 建模样品与预测样品数量比较

Table 2 Comparison of the quantity of model samples and forecast samples

地区 Area	训练集 Training set	测试集 Test set	总计 Total
产地内 Inside production area	186	80	266
产地外 Outside production area	82	36	118

**1.4 原始光谱数据预处理方法** 近红外光谱仪扫描得到的光谱数据往往含有噪音,这主要由 2 个原因造成:一方面受到仪器自身和测量环境的影响;另一方面可能来自于光源自身和样品基质的差异<sup>[21]</sup>。为了降低上述 2 种因素对分析模

型的影响、提高分析模型的准确性和稳定性,必须对近红外光谱数据进行预处理。

**1.4.1 SNV 标准正态变换。**SNV 是一种转换光谱的预处理方法,通常用于消除散射、噪声、粒径和光程变化的多元综合干扰<sup>[22]</sup>。在 SNV 中,每个频谱都被标准化为零均值和单位方差。

**1.4.2 MSC 多元散射校正。**MSC 是一种常用的光谱预处理技术,用于在建模前去除光谱数据矩阵中不希望出现的散射效应<sup>[23]</sup>,还用于补偿光谱数据中由于不同粒径的不均匀分布所引起的不均匀散射的影响。其工作原理是将每个光谱线性化为与校准装置的平均光谱相对应的“理想”光谱,而平均光谱主要通过最小二乘法确定<sup>[24]</sup>。

### 1.5 数据分析方法

**1.5.1 LDA 线性判别分析。**LDA 通过标记数据的线性组合构造判别函数,是一种监督方法,主要包括 2 个阶段:分离和分配。前一阶段是寻找能够很好地将群体分离的判别函数,后一阶段是利用判别函数将未知对象赋给其中一个群体。

**1.5.2 PCA 主成分分析。**PCA 是对原始变量进行线性组合,并用较少的新变量尽可能多地表征原始数据的结构特征,这几个相互正交的新变量即为主成分<sup>[25]</sup>。PCA 已用于观察多个传感器数据集中可能的分组。性能最好的 PC 通常显示最重要的信息。因此,相似的样品彼此之间的分组更接近,反之亦然。

**1.5.3 K-means 聚类分析。**K-means 算法的中心思想是给定一个数据库和聚类数  $K$ ,根据样品到聚类中心的距离将样品划分到离它最近的类别中,在空间聚类算法中有重要的运用。

**1.5.4 HCA 层次聚类分析。**HCA 是聚类算法的一种,其主要原理是通过计算 2 类数据点间的相似性,将数据中最相似的 2 类点进行组合,并反复迭代,创建 1 棵有层次结构的嵌套聚类树。

## 2 结果与分析

**2.1 近红外光谱分析** 图 1 显示了产地内、外安吉白茶 9 000~4 000  $\text{cm}^{-1}$  波段的原始光谱图(a)、平均光谱图(b)以及在不同预处理方法下(SNV、MSC)的平均光谱图(c和d)。在 9 000~4 000  $\text{cm}^{-1}$  波段内有一些密集的光谱峰,这些峰主要是由 C—H、N—H、O—H 和 C—O 键的拉伸或变形振动产生的,C—H、N—H、O—H 和 C—O 键是有机分子的主要结构成分,而这些振动主要是由茶叶中的儿茶素、氨基酸、多酚、生物碱以及一些芳香化合物等成分引起的<sup>[26]</sup>。

图 1b 是产地内、外 2 个茶叶群组的原始平均光谱图。图中显示,产地内、外茶叶样品的平均吸光度趋于平行,且产地内的略高于产地外。在 12 000~9 000  $\text{cm}^{-1}$  区域内,产地内、外茶叶样品的平均吸光度基本保持不变,产地内保持在 0.55 左右,产地外保持在 0.50 左右,可认为在该波段不包含待测对象的有用信息,其灵敏度和信噪比较低<sup>[27]</sup>,因此该试验将 9 000~12 000  $\text{cm}^{-1}$  的波段排除在外。

由图 1a 可知,原始光谱图较杂乱,需进行预处理。比较

图 1c 的 SNV、图 1d 的 MSC 预处理后的平均光谱图与原始平均光谱图(图 1b)可知,SNV 预处理后产地内、外茶叶样品的平均吸光度趋近,肉眼难以区分;而经 MSC 预处理后的平均

光谱图更加接近于原始平均光谱图,且肉眼可分。对 SNV 和 MSC 预处理后的数据分别进行 LDA 建模分析,进一步比较和探讨不同预处理方法的效果。

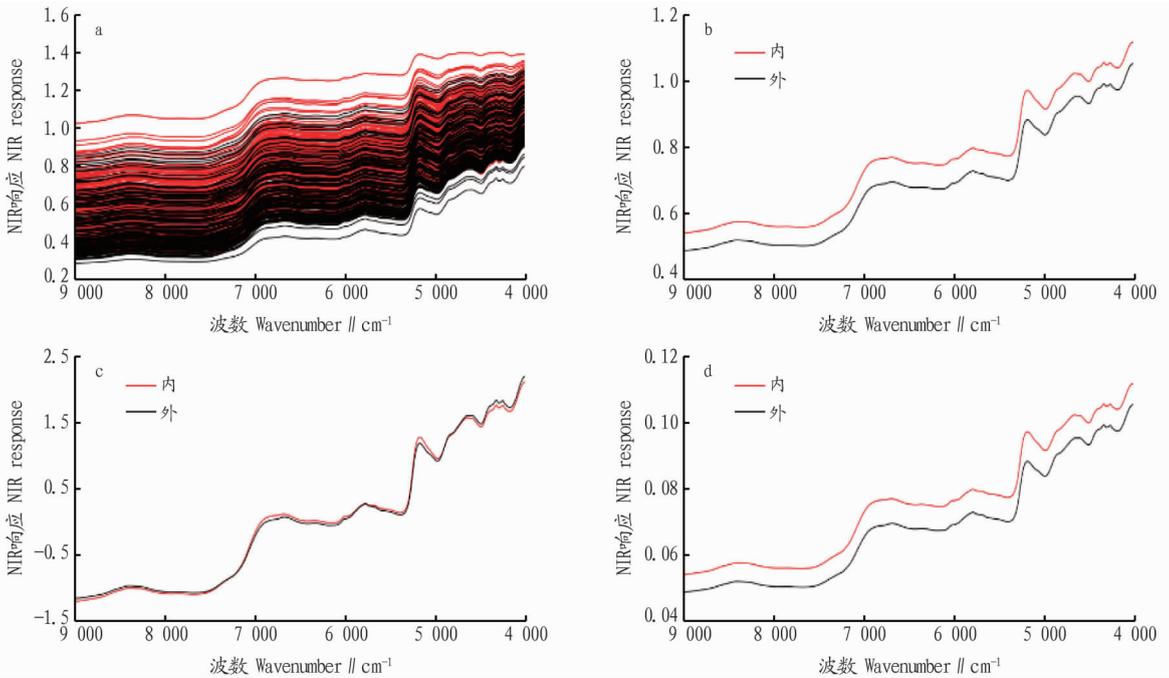


图 1 不同预处理下 9 000~4 000  $\text{cm}^{-1}$  波段的产地内、外光谱图

Fig. 1 NIR spectra of Anji white tea from 4 000 to 9 000  $\text{cm}^{-1}$  with raw (a), mean (b), SNV mean (c) and MSC mean (d) spectra

**2.2 LDA 区分产地内、外茶叶的结果** 用于分类鉴定的模式识别算法有很多,该试验采用 LDA 来建立分类模型,主要是因为 LDA 在初始样品降维中就表现出较好的效果。图 2 为 LDA 将初始样品降至二维的结果图。从图 2 可知,产地内、外的茶叶样品数据可以清楚地区分开来,且 LD1 的方差贡献率高达 100%。

通过 LDA 模型进行准确地区分。为进一步验证 MSC 和 SNV 预处理效果优劣差异,通过 PCA、K-means 聚类和 HCA 聚类等方法进行模型验证。

表 3 安吉白茶 LDA 分类

Table 3 LDA classification of Anji white tea

数据 Data	PCA 维	正确率 Accuracy/%	
		训练集 Training set	测试集 Test set
NIR9000	50	98.99	96.87
NIR9000 + SNV	50	98.97	96.34
NIR9000 + MSC	50	100	100

注: NIR9000 表示 9 000~4 000  $\text{cm}^{-1}$  波段的原始数据

Note: NIR9000 indicated original data at 9 000~4 000  $\text{cm}^{-1}$  wave band

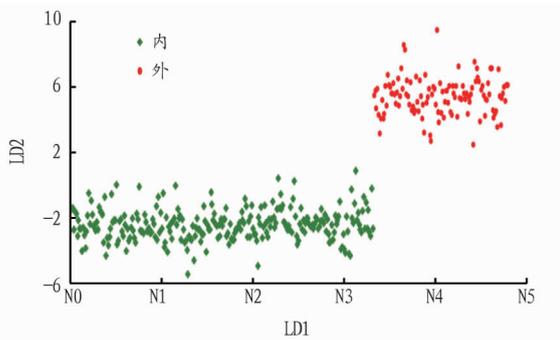


图 2 产地内、外样品 LDA 降维图

Fig. 2 LDA dimension reduction figures of white tea in and out of original production area

采用 LDA 建立分类模型,随机选取样品的 70% 作为训练集,用于模型的建立;30% 作为预测集,用于模型的验证。同时采用随机数的方法,做 50 次分析,并取所得准确度的平均值作为最终的数据结果和评判标准,具体结果如下表 3 所示。由表 3 可知,SNV 预处理后对分类准确率并没有提升,即其预处理效果并不明显,而 MSC 预处理后的分类准确率均达 100%,说明产地内、外样品数据经 MSC 预处理后可以

**2.3 SNV 和 MSC 在 PCA 中的效果比较** 图 3 显示样品数据经 SNV 和 MSC 预处理后在 PCA 降至三维中的效果图。由图 3a 可知,SNV 预处理后降至三维时,PC1 的方差占比达到 87.88%,PC2 的方差占比为 8.04%,前 2 个主成分的方差占比和高达 95.92%。由图 3b 可知,MSC 预处理后降至三维时,PC1 的方差占比达到 89.86%,PC2 的方差占比为 8.93%,前 2 个主成分的方差占比和高达 98.79%,明显高于 SNV。

从图 3 可以看出,经过 PCA 降至三维可视化后,MSC 的结果非常好,产地内、外的样品能清晰地区分开,而 SNV 的结果较模糊。这一现象可以解释为 MSC 在该试验中效果更好,具有较强的校正散射效应的能力。

**2.4 SNV 和 MSC 在 K-means 和 HCA 聚类分析比较** 由图 4 可知,SNV 预处理后的聚类结果没有 MSC 预处理后的

结果好。在 SNV 预处理后的聚类结果中,K-means 算法预测出 155 个产地内样品,229 个产地外样品,分数为 470.14,识别率为 42.2%;HCA 算法预测出 221 个产地内样品,163 个产地外样品,分数为 393.60,识别率为 78.6%。而在 MSC 预处理

后的聚类结果中,K-means 算法和 HCA 算法都成功地将 266 个产地内样品和 118 个产地外样品预测出来,分数均为 3492.59,识别率亦为 100%,明显高于 SNV。这同样验证 MSC 预处理在该试验中表现效果较优。

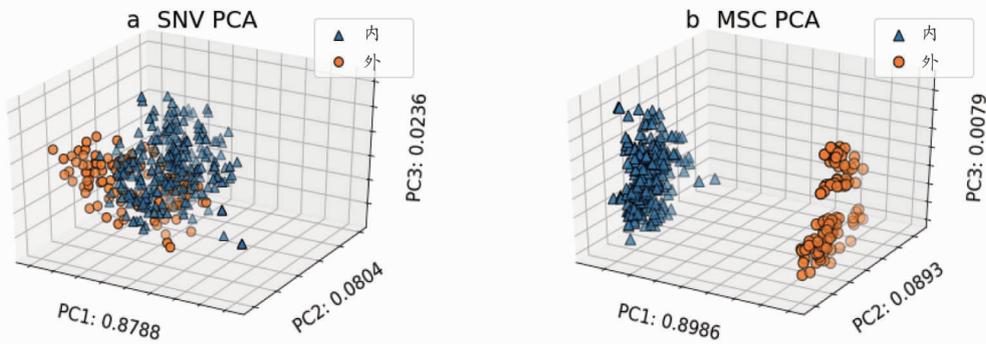
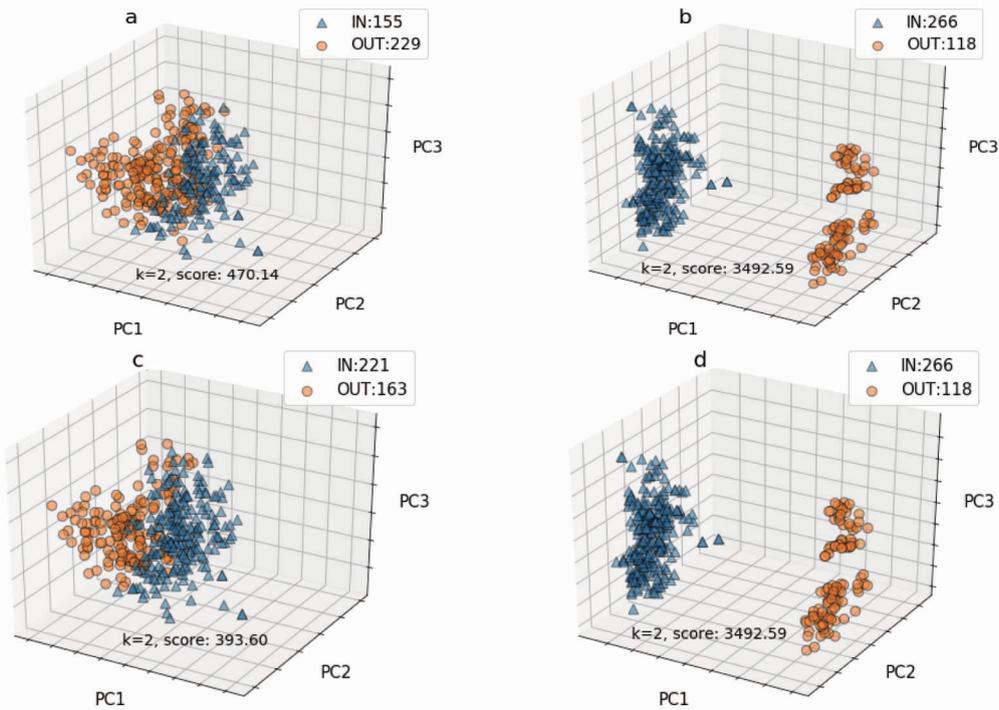


图 3 SNV (a) 和 MSC (b) 预处理后的 PCA 三维效果图

Fig. 3 Three-dimensional results of PCA of Anji white tea and non-Anji white tea with SNV (a) and MSC (b) pretreatments



注:a 和 b 分别是 SNV、MSC 的 K-means 聚类结果;c 和 d 分别是 SNV、MSC 的 HCA 聚类结果

Note:a and b were K-means cluster results of SNV and MSC;c and d were HCA cluster results of SNV and MSC

图 4 SNV、MSC 预处理的 K-means 和 HCA 聚类结果

Fig. 4 K-means cluster results with SNV (a) and MSC (b) pretreatment and HCA cluster results with SNV (c) and MSC (d)

**2.5 产地内样品 SPSS 聚类结果与地理位置的关系** 已知产地内茶叶样品主要采集于安吉县内 7 个不同的镇(乡),包括高禹镇(15 个样品)、递铺镇(74 个样品)、上墅乡(15 个样品)、昆铜乡(72 个样品)、良朋镇(15 个样品)、溪龙乡(60 个样品)、梅溪镇(15 个样品)。其镇(乡)的具体地理位置见图 5。

采用 SPSS 软件对经 MSC 预处理后的产地内茶叶样品数据进行聚类,聚类结果见图 6。由图 6 可知,递铺镇和昆铜乡明显聚为一类,高禹镇和良朋镇明显聚为一类,梅溪镇和溪龙乡可大致聚为一类,上墅乡单独成一类,这与地图上的镇(乡)地理位置的远近基本吻合。

根据上述树状图可知,可将产地内茶叶的 7 个采集地聚为 3 类,其中递铺镇、昆铜乡、高禹镇、良朋镇为一类;梅溪镇和溪龙乡为一类;上墅乡单独为一类,具体情况见表 4。

表 4 产地内茶叶采集地聚类

Table 4 SPSS cluster results of Anji white tea inside production area

序号 Code	类别 Type	采集地 Collection site
1	第 1 类	上墅乡
2	第 2 类	溪龙乡、梅溪镇
3	第 3 类	高禹镇、昆铜乡、递铺镇、良朋镇



图5 产地内样品采集区域

Fig. 5 Collection sites of Anji white tea samples

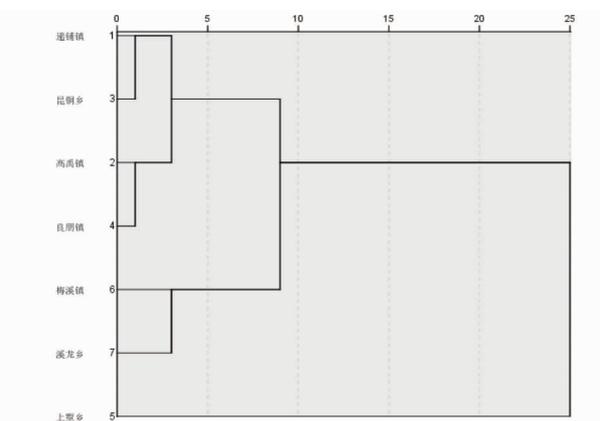


图6 安吉白茶7个原产地SPSS聚类树状图

Fig. 6 SPSS cluster results with seven different original areas of Anji white tea

### 3 结论

本研究对产地内、外共384个茶叶样品建立定性模型,进行源产地追溯研究,其中产地内为正宗安吉白茶,产地外为非安吉白茶茶样。结果表明,MSC的预处理效果明显优于SNV。利用MSC预处理方法,建立的LDA分类模型的准确率高达100%,同时通过PCA、K-means聚类和HCA聚类等方法进行模型验证,结果显示MSC优于SNV。经MSC预处理后,PCA方法可在前3个PC下清晰聚类,K-means和HCA聚类方法都可将266个产地内和118个产地外样品成功预测出。对正宗安吉白茶的7个茶样采集地进行SPSS聚类分析,其聚类结果与采集地的地理位置远近有对应关系。该研究表明,近红外光谱技术结合MSC预处理和LDA建模能够很好地运用于安吉白茶的源产地追溯,对保护安吉白茶的产地真实性具有重要的作用,同样可借鉴用于其他地标产品的地理位置溯源与保护。

### 参考文献

[1] 韩树根. 实施标准化战略 推进茶产业发展——安吉县建设“全国茶叶标准化示范县”历程[J]. 中国标准化, 2019(9): 32-35.  
 [2] 胡茶根, 赵红霞, 边文亮. 近红外技术在茶叶快速无损检测方面的研究与应用[J]. 食品科学, 2007, 28(10): 638-641.  
 [3] 白艳, 赖建红, 汤丹. 浅谈安吉白茶一二三产融合发展之路[J]. 中国茶叶, 2018, 40(12): 40-42.

[4] 袁玉伟, 胡桂仙, 邵圣枝, 等. 茶叶产地溯源与鉴别检测技术研究进展[J]. 核农学报, 2013, 27(4): 452-457.  
 [5] 范方媛, 杨梦璇, 龚淑英, 等. 基于近红外光谱技术的白茶3种典型感官滋味特征属性定量评价模型研究[J]. 茶叶科学, 2018, 38(3): 296-304.  
 [6] 陈美丽, 张俊, 龚淑英, 等. 茉莉花茶主要品质成分定量近红外光谱分析模型的建立[J]. 茶叶科学, 2013, 33(1): 21-26.  
 [7] LI L Q, WEI L D, NING J M, et al. Detection and quantification of sugar and glucose syrup in roasted green tea using near infrared spectroscopy[J]. Journal of near infrared spectroscopy, 2015, 23(5): 317-325.  
 [8] 王曼, 张正竹, 宁井铭, 等. 基于近红外光谱的黄山毛峰茶鲜叶品质分析及等级快速评价[J]. 食品工业科技, 2014, 35(22): 57-60, 64.  
 [9] 周健, 成浩, 叶阳, 等. 基于近红外的 Fisher 分类法识别茶叶原料品种的研究[J]. 光学学报, 2009, 29(4): 1117-1121.  
 [10] CHEN Q S, ZHAO J W, LIN H. Study on discrimination of Roast green tea (*Camellia sinensis* L.) according to geographical origin by FT-NIR spectroscopy an supervised pattern recognition[J]. Spectrochimica acta part A: Molecular & biomolecular spectroscopy, 2009, 72(4): 845-850.  
 [11] 曾智朋, 尹春玲, 胡乐乾, 等. 近红外光谱结合 SVM 识别4种茶叶[J]. 河南工业大学学报(自然科学版), 2013, 34(5): 53-57.  
 [12] 顾玉琦, 刘瑞婷, 寿国忠, 等. 应用近红外光谱技术快速鉴别铁皮石斛的产地[J]. 江苏农业科学, 2016, 44(5): 365-368.  
 [13] 刘明地, 宋萍, 王博. 基于大叶三七红外光谱系统聚类分析的产地鉴别[J]. 华中师范大学学报(自然科学版), 2019, 53(2): 222-228.  
 [14] FU X S, YU X P, YE Z H, et al. Analysis of antioxidant activity of Chinese brown rice by fourier-transformed near-infrared (NIR) spectroscopy and chemometrics[J]. Journal of chemistry, 2015, 2015(4): 1-5.  
 [15] XU L, FU X S, CAI C B, et al. The feasibility of using near infrared spectroscopy for rapid discrimination of aged shiitake mushroom (*Lentinula edodes*) after long-term storage[J]. Journal of chemistry, 2015, 2015: 1-7.  
 [16] FU X S, XU L, YU X P, et al. Robust and automated internal quality grading of a Chinese green tea (Longjing) by near-infrared spectroscopy and chemometrics[J]. Journal of spectroscopy, 2013(1): 367-383.  
 [17] HONG X Z, FU X S, WANG Z L, et al. Tracing geographical origins of teas based on FT-NIR spectroscopy: Introduction of model updating and imbalanced data handling approaches[J]. Journal of analytical methods in chemistry, 2019, 2019: 1-8.  
 [18] 宋雪健, 钱丽丽, 张东杰, 等. 基于漫反射傅里叶变换近红外光谱技术对不同年份的大米产地溯源检测[J]. 食品科学, 2017, 38(18): 286-291.  
 [19] TEYE E, AMUAH C L Y, MCGRATH T F, et al. Innovative and rapid analysis for rice authenticity using hand-held NIR spectrometry and chemometrics[J]. Spectrochimica acta part A: Molecular and biomolecular spectroscopy, 2019, 217: 147-154.  
 [20] EISENSTECKEN D, STÜRZ B, ROBATSCHER P, et al. The potential of near infrared spectroscopy (NIRS) to trace apple origin: Study on different cultivars and orchard elevations[J]. Postharvest biology and technology, 2019, 147: 123-131.  
 [21] 刘广昊, 祝诗平, 袁嘉佑, 等. 基于近红外光谱的胡椒产地鉴别方法研究[J]. 中国调味品, 2019, 44(5): 58-62, 66.  
 [22] BARNES R, DHANOA M S, LISTER S J. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra[J]. Applied spectroscopy, 1989, 43(5): 772-777.  
 [23] CORONEL-REYES J, RAMIREZ-MORALES I, FERNANDEZ-BLANCO E, et al. Determination of egg storage time at room temperature using a low-cost NIR spectrometer and machine learning techniques[J]. Computers and electronics in agriculture, 2018, 145: 1-10.  
 [24] WANG H L, PENG J Y, XIE C Q, et al. Fruit quality evaluation using spectroscopy technology: A review[J]. Sensors, 2015, 15(5): 11889-11927.  
 [25] 吴习宇, 祝诗平, 黄华, 等. 近红外光谱技术鉴别花椒产地[J]. 光谱学与光谱分析, 2018, 38(1): 68-72.  
 [26] 刘丽霞. 茶叶中6种主要儿茶素的高效液相色谱方法建立及应用[D]. 南京: 南京理工大学, 2013.  
 [27] MENG W J, XU X N, CHENG K K, et al. Geographical origin discrimination of oolong tea (*TieGuanYin*, *Camellia sinensis* (L.) O. Kuntze) using proton nuclear magnetic resonance spectroscopy and near-infrared spectroscopy[J]. Food analytical methods, 2017, 10(11): 3508-3522.