茶皂素提取过程的原位实时监测技术

张雪莉[1]， 丁艳华[1]， 马海乐[1*]， 洪 晨[1]， 张 涛[2]， 唐艳萍[2]
([1]江苏大学食品与生物工程学院 江苏镇江 212013
[2]湖南新金浩茶油股份有限公司 湖南永州 426100)

摘要 采用近红外光谱技术结合化学计量学方法，实时原位监测茶皂素提取过程中的茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度的变化。研究结果表明，一阶求导(1[st]Der)为最优的光谱预处理方法，3 个指标的校正集决定系数(R_C)和预测集决定系数(R_P)皆最高，交叉校验残差均方根误差(RMSECV)和预测残差均方根误差(RMSEP)最小。相较于偏最小二乘法(PLS)和区间偏最小二乘法(iPLS)算法对回归模型进行校正，以 R_P 和 RMSEP 为评价指标，联合区间偏最小二乘法(Si-PLS)算法下所建模型稳健性最佳。茶皂素浓度模型中的 R_P=0.9889，RMSEP=1.36；蛋白质浓度的 R_P=0.9859，RMSEP=0.354；多糖浓度的 R_P=0.9919，RMSEP=0.359。结论：近红外光谱技术联合 Si-PLS 模型可较好的实时监测茶皂素提取过程。

关键词 油茶籽饼粕；茶皂素；近红外光谱；原位监测；联合区间偏最小二乘法
文章编号 1009-7848(2022)11-0326-09 **DOI**：10.16429/j.1009-7848.2022.11.034

油茶籽粕中营养成分丰富，主要含有多糖 40%、茶皂素 10%~15%、粗蛋白 15%、粗脂肪 5% 以及多种氨基酸，是一种潜在价值很高的饲料原料[1]。茶皂素因溶血作用[2]，而可广泛应用于农作物害虫防治，且无污染、无毒害，耐于储存。茶皂素还是一种性能良好的天然非离子表面活性剂[3]，可广泛应用于日用化工、纺织和食品等行业[4-5]。此外，茶皂素具有杀菌[6]、消炎抗渗[7]以及抗氧化[8]等生物活性。从油茶籽饼粕中提取茶皂素对于提高油茶籽饼粕的经济价值，充分利用油茶籽饼粕资源，具有着重要的现实意义，应用前景十分广阔。

目前，茶皂素含量常用香草醛-浓硫酸显色法测定[9]，也有重量法、色谱法等[10-11]。蛋白质含量检测方法较多，主要有凯式定氮法、lowry 法和杜马斯燃烧法等[12-13]。检测多糖的方法多为苯酚硫酸法和蒽酮硫酸法[14-15]。这些传统检测方法过程繁琐、耗时较长、污染环境，且实际测量的化学值与提取过程之间存在延迟或误差，迫切需要一种连续、快速、准确、环保的方法监测提取过程中的茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度的变化，以确定提取终点。

收稿日期：2021-11-16
基金项目：国家自然科学基金面上项目(31471698)
第一作者：张雪莉，女，博士生
通信作者：马海乐 E-mail：mhl@ujs.edu.cn

相较于传统分析技术，近红外光谱技术具有快速、准确、成本低和绿色环保等特点[16-17]。Hua 等[18]对茯苓多糖的提取过程进行近红外光谱监测；Tian 等[19]利用近红外光谱和化学计量学技术联合的方法测定全麦粉中的总酚含量。此外，近红外光谱技术也被用于藏红花[20]和牛肉[21]的鉴别和掺假检测。可见，近红外光谱技术已广泛应用于在线和现场监测[22-23]。

本研究应用便携式近红外光谱技术，实时监测茶皂素提取过程中的主要质量参数变化，研究其适用性并选择一种有效的算法建立稳健的预测模型，实现提取过程中对各质量参数的预测，为终点的判断提供技术支持。

1 材料与方法

1.1 材料与试剂

油茶籽饼粕，湖南新金浩茶油股份有限公司；茶皂素标品，北京索莱宝科技有限公司；牛血清白蛋白，上海麦克林生化科技有限公司；香兰素、乙醇和浓硫酸等试剂均为分析纯级。

T-6 新世纪可见分光光度计，北京普析仪器有限公司；NIRQUEST256-2.5 近红外光谱仪、TP300 浸入式光纤探头、DH-2000-BAL UV-VIS-NIR 光源，美国海洋光学公司；台式高速冷冻离心机，盐城市凯特实验仪器有限公司；HH-4 数显恒温水浴锅，上海力辰邦西仪器科技有限公司。

1.2 试验方法

1.2.1 提取过程中原位实时光谱采集　近红外原位在线监测提取过程的系统装置图如图 1 所示。75%乙醇以料液比 1∶10 在 60 ℃下提取，每 2 min 取样一次，共计 88 个样品。将所有样品离心(10 000 r/min 离心 10 min)，收集上清液，放置于-20 ℃保存，待测化学值。光谱采集参数：光谱范围为 900~2 500 nm，分辨率为 6.4 nm，扫描 8 次，光程 4 mm。以 75%乙醇(60 ℃)作为背景，每个样品重复采集 3 次，取其平均值作原始光谱图。

1.2.2 茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度的离线测定　茶皂素质量浓度采用香草醛-浓硫酸显色法测定[24]。蛋白质质量浓度采用 lowry 法[12]测定，以牛血清白蛋白作为基准物质作标准曲线。采用苯酚硫酸法[14]测定多糖质量浓度，以干燥的葡萄糖标准品作为基准物质制作标准曲线。提取的茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度均按下式计算：

$$c=C\times N$$

式中，c——样品质量浓度，mg/mL；C——测定质量浓度，mg/mL；N——稀释倍数。

1.2.3 原位实时光谱预处理　采用标准正态变换(SNV)、多项式卷积平滑(SG)、一阶导数(1st Der)和二阶导数(2nd Der)4 种常见的预处理方法对提取过程中采集的原始光谱进行预处理。采用 PLS 模型进行预处理方法的对比研究，选取最优的预处理方法。

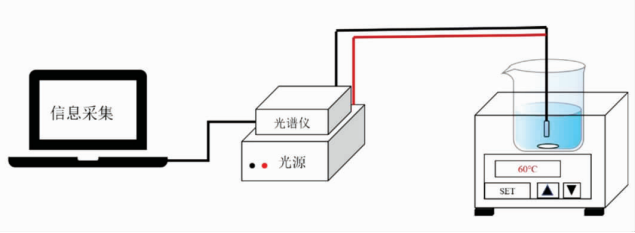


图 1　原位实时光谱在线监测茶皂素提取过程装置图

Fig.1　Device diagram for *in-situ* real-time spectroscopy on-line monitoring of tea saponin extraction process

1.2.4 提取过程原位实时监测定量模型的建立　选用 Set Partitioning Based On Joint X-Y Distance(SPXY)方法对近红外光谱仪采集的光谱信息进行校正集(66 个)和预测集(22 个)样本的划分，共计 88 个样本，如表 1 所示。

采用偏最小二乘法 (Partial least squares, PLS)、区间偏最小二乘法 (Interval partial least squares, ipLS)和联合区间偏最小二乘法(Synergy interval PLS, Si-PLS)进行模型构建，分别筛选出与茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度相关度高、预测及稳健性较好的模型。

表 1　校正集和预测集中茶皂素、蛋白质和多糖质量浓度的实测值

Table 1　Measured values of tea saponin, protein and polysaccharide mass concentrations in correction and prediction set

指标	分集	S.N	范围	平均值	S.D
茶皂素/mg·mL^{-1}	校正集	66	31.16~66.60	47.06	8.79
	预测集	22	29.70~62.15	47.82	9.17
蛋白质/mg·mL^{-1}	校正集	66	5.98~16.64	9.95	2.21
	预测集	22	6.54~14.36	9.93	2.13
多糖/mg·mL^{-1}	校正集	66	9.99~19.57	15.37	2.59
	预测集	22	10.29~20.12	15.23	3.00

注：S.N 表示样品数；S.D 表示标准偏差。

1.2.5 数据分析　所有数据处理和分析均在 Matlab 2009b 中进行。从 http://www.models.kvl.dk/免费下载 Si-PLS Matlab 代码。

2 结果与分析

2.1 提取过程中茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度的变化

提取过程中茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度变化如图 2 所示，随着提取时间延长，各样品质量浓度虽不断增加，但增长缓慢，呈现波动变化。可以发现，提取过程中，茶皂素提取量最高，多糖及蛋白质提取量较低，可见此提取方法能较大限度的提取较高纯度的茶皂素。同时监测茶皂素、蛋白质和多糖的质量浓度变化，可以更加全面的分析数据并建立模型，对提取中主要物质的变化进行更好的预测。

2.2 提取过程中光谱预处理

通过近红外光谱仪采集到茶皂素提取过程中的原始光谱图，选取标准正态变换(SNV)、多项式卷积平滑(SG)、一阶求导(1st Der)和二阶求导(2nd Der)4 种预处理方法对原始光谱进行预处理。4 种预处理方法结合最小二乘法(PLS)对提取过程采集的光谱分别与茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度进行建模，以校正集决定系数(R_C)、交叉校验残差均方根误差(RMSECV)和预测集决定系数 (R_p)、预测残差均方根误差(RMSEP)为模型评价指标，比较其建模效果从而筛选出最优的预处理方法，结果见表 2。由表 2 可知，在已构建的光谱信息和茶皂素质量浓度之间的模型中，1st Der 预处理方法建立的校正模型、预测模型的决定系数 R_C 值、R_p 值均最大，分别为 0.9819 和 0.9932，且 RMSECV 值、RMSEP 值比较低，分别为 1.67 和 1.08。在光谱信息和蛋白质质量浓度之间的模型中，1st Der 预处理方法建立的校正模型、预测模型的决定系数 R_C 值、R_p 值分别为 0.9718 和 0.9879，而 RMSECV 值、RMSEP 值也较低，分别为 0.518 和 0.323。已构建的光谱信息和多糖质量浓度之间的模型中，1st Der 预处理方法建立的校正模型、预测模型的决定系数 R_C 值、R_p 值也最大，分别为 0.9798 和 0.9947，RMSECV 值、RMSEP 值较低，分别为 0.531 和 0.309。因此，与其它 3 种预处理方法相比，使用 1st Der 预处理方法对原始光谱进行预处理后，得到的模型预测能力较强、稳健性好。因此，本研究后续将采用 1st Der 预处理后的光谱开展模型的构建。

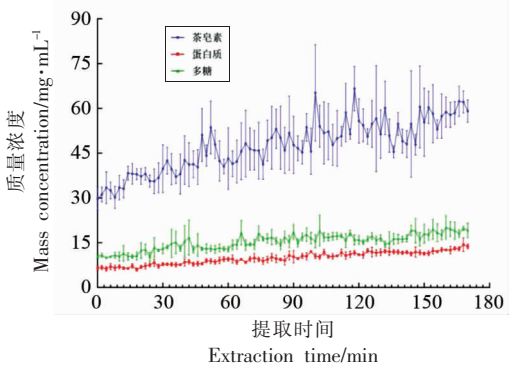


图 2 提取过程中茶皂素、蛋白质和多糖质量浓度变化

Fig.2 Off-line measured data of tea saponin, protein and polysaccharide mass concentration during the extraction process

表 2 不同方法预处理近红外光谱的结果

Table 2 Results of near-infrared spectroscopy by different pretreatment methods

指标	预处理方法	主成分因子数	R_C	RMSECV	R_P	RMSEP
茶皂素/mg·mL⁻¹	SNV	6	0.8779	4.240	0.9166	3.580
	SG	6	0.8648	4.490	0.8881	4.070
	1st Der	10	0.9819	1.670	0.9932	1.080
	2nd Der	4	0.7751	5.590	0.7849	5.390
蛋白质/mg·mL⁻¹	SNV	5	0.8356	1.210	0.9417	0.709
	SG	5	0.8256	1.260	0.8970	0.917
	1st Der	10	0.9718	0.518	0.9879	0.323
	2nd Der	4	0.7290	1.510	0.7546	1.370
多糖/mg·mL⁻¹	SNV	6	0.8798	1.270	0.9183	1.110
	SG	6	0.8698	1.320	0.8861	1.270
	1st Der	10	0.9798	0.531	0.9947	0.309
	2nd Der	10	0.8451	1.450	0.9038	1.160

2.3 近红外光谱与提取过程的建模研究

2.3.1 茶皂素质量浓度的 Si-PLS 模型 表 3 为提取过程中茶皂素的 Si-PLS 模型结果。从表中可看出，将光谱划分为 16 个区间，并联合其中的 4 个子区间 [2,3,7,14] 时可得到最优的 Si-PLS 模型，所对应的光谱区间分别为 999.07~1 096.61，1 103.10~1 200.36，1 516.35~1 612.58，2 229.86~2 324.62 nm。999.07~1 096.61 nm 区间包含 1 065 nm 和 1 029 nm 处的组合频吸收谱带[25]，主要吸收 C-H 和 O-H 的三级倍频和二级倍频[26]。

在最佳联合子区间下，对提取过程中的茶皂素质量浓度建立定量分析模型，其校正集及预测集样本实测值与预测值之间的散点图如图 3a 和图 3b 所示。由图可知，对提取过程中茶皂素质量浓度，校正模型的决定系数 R_C 值为 0.9858，RMSECV 为 1.48，预测模型的决定系数 R_P 值为 0.9889，RMSEP 为 1.36。

2.3.2　蛋白质质量浓度的 Si-PLS 模型　从表 3 可看出，将光谱划分为 16 个区间，并联合其中的 4 个子区间[2,3,7,14]时可得到蛋白质质量浓度的最优的 Si-PLS 模型，所对应的光谱区间分别为 999.07~1 096.61，1 103.10~1 200.36，1 516.35~1 612.58，2 229.86~2 324.62 nm。其中，包含的 1 530~1 600 nm 为 NH 一级倍频吸收谱带，对蛋

表 3　提取过程中茶皂素、蛋白质和多糖质量浓度的 Si-PLS 模型优选结果

Table 3　Optimization results of Si-PLS model for tea saponin, protein and polysaccharide mass concentration in the extraction process

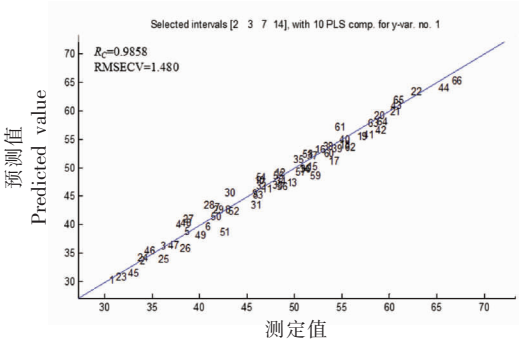
指标	区间数	区间	PCs	R_C	RMSECV	R_P	RMSEP
茶皂素	15	[1,3,5,10]	10	0.9838	1.580	0.9712	2.080
	16	[2,3,7,14]	10	0.9858	1.480	0.9889	1.360
	17	[3,7,8,10]	10	0.9852	1.510	0.9816	1.690
	18	[3,7,8,15]	10	0.9862	1.460	0.9788	1.890
	19	[1,3,7,14]	10	0.9865	1.450	0.9840	1.570
	20	[2,9,12,17]	10	0.9871	1.410	0.9784	1.800
	21	[2,10,13,18]	10	0.9864	1.460	0.9792	1.760
	22	[2,9,10,20]	9	0.9867	1.440	0.9834	1.580
	23	[2,4,10,20]	10	0.9871	1.420	0.9623	2.430
	24	[2,4,9,10]	10	0.9863	1.460	0.9664	2.280
	25	[2,4,8,10]	10	0.9891	1.310	0.9786	1.790
	26	[4,10,12,17]	10	0.9844	1.550	0.9835	1.670
	27	[4,10,12,16]	10	0.9847	1.540	0.9852	1.520
	28	[4,11,13,17]	10	0.9854	1.510	0.9847	1.560
	29	[3,11,17,20]	10	0.9803	1.750	0.9792	1.870
	30	[3,5,13,25]	10	0.9802	1.750	0.9544	2.760
蛋白质	15	[1,3,7,10]	10	0.9752	0.488	0.9691	0.520
	16	[2,3,7,14]	10	0.9766	0.472	0.9859	0.354
	17	[2,3,8,15]	10	0.9755	0.482	0.9855	0.353
	18	[2,3,7,11]	10	0.9765	0.474	0.9806	0.415
	19	[2,3,7,14]	10	0.9762	0.477	0.9811	0.416
	20	[2,9,12,17]	10	0.9762	0.478	0.9727	0.487
	21	[2,10,13,18]	10	0.9764	0.476	0.9729	0.482
	22	[2,3,4,9]	10	0.9768	0.471	0.9793	0.423
	23	[2,3,4,8]	10	0.9760	0.479	0.9749	0.464
	24	[4,9,10,14]	9	0.9739	0.498	0.9781	0.443
	25	[2,4,8,10]	10	0.9776	0.463	0.9780	0.436
	26	[3,4,11,22]	10	0.9752	0.486	0.9767	0.455
	27	[3,16,21,23]	10	0.9759	0.480	0.9825	0.387
	28	[3,17,22,24]	10	0.9759	0.480	0.9825	0.387
	29	[3,4,21,24]	10	0.9712	0.524	0.9714	0.516
	30	[3,5,13,25]	10	0.9711	0.524	0.9530	0.667

(续表 3)

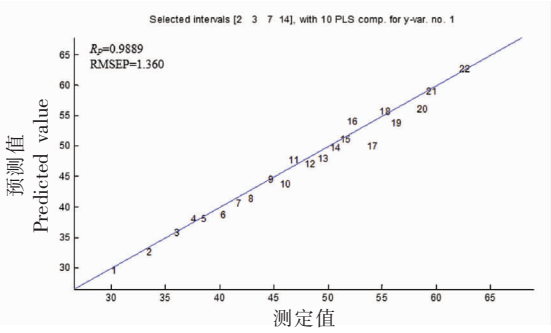
指标	区间数	区间	PCs	R_C	RMSECV	R_P	RMSEP
多糖	15	[5,6,7,11]	10	0.9841	0.471	0.9879	0.430
	16	[3,5,7,12]	10	0.9865	0.436	0.9894	0.417
	17	[3,7,8,14]	10	0.9840	0.473	0.9824	0.526
	18	[6,7,8,13]	10	0.9841	0.471	0.9919	0.359
	19	[3,7,14,15]	10	0.9870	0.427	0.9815	0.522
	20	[2,9,12,17]	10	0.9847	0.464	0.9842	0.480
	21	[2,10,13,18]	10	0.9855	0.452	0.9837	0.493
	22	[5,9,11,15]	10	0.9856	0.451	0.9839	0.495
	23	[2,8,9,17]	10	0.9830	0.488	0.9802	0.559
	24	[4,9,17,18]	10	0.9832	0.487	0.9819	0.525
	25	[2,4,8,10]	10	0.9879	0.412	0.9858	0.455
	26	[4,10,12,17]	10	0.9821	0.501	0.9832	0.507
	27	[4,9,10,11]	10	0.9819	0.503	0.9825	0.517
	28	[4,10,11,12]	10	0.9850	0.459	0.9828	0.515
	29	[4,5,12,20]	10	0.9788	0.547	0.9673	0.709
	30	[3,13,16,23]	10	0.9807	0.521	0.9829	0.506

白质分析非常重要[27]。

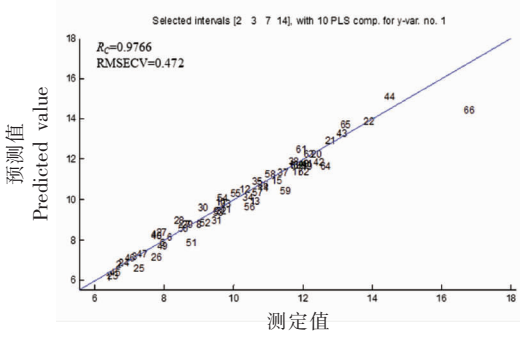
在最佳联合子区间下，对提取过程中的蛋白质浓度建立定量分析模型，其校正集及预测集样本实测值与预测值之间的散点图如图 3c 和图 3d 所示。由图可知，对提取过程中蛋白质质量浓度，校正模型的决定系数 R_C 值为 0.9766，RMSECV 为 0.472，预测模型的决定系数 R_P 值为 0.9859，RMSEP 为 0.354。



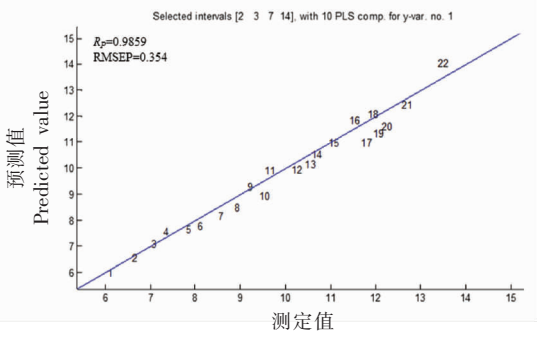
(a)茶皂素质量浓度校正集



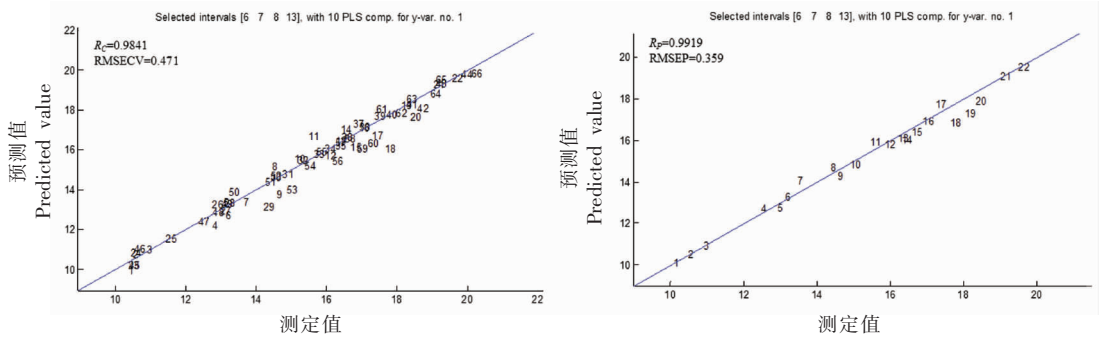
(b)茶皂素质量浓度预测集



(c)蛋白质质量浓度校正集



(d)蛋白质质量浓度预测集



(e)多糖质量浓度校正集　　(f)多糖质量浓度预测集

图 3　Si-PLS 模型中茶皂素、蛋白质和多糖质量浓度实测值与预测值之间的散点图

Fig.3　Calibration set and prediction set of Si-PLS joint optimal interval model

2.3.3　多糖质量浓度的 Si-PLS 模型　由表 3 可得，将光谱划分为 18 个区间，并联合其中的 4 个子区间[6,7,8,13]时可得到多糖质量浓度的最优的 Si-PLS 模型，所对应的光谱区间分别为 1 349.00~1 432.76，1 439.20~1 522.77，1 529.19~1 612.58，1 976.30~2 058.85。其中，1 440 nm 附近通常为游离 OH 伸缩振动吸收峰[25]。

在最佳联合子区间下，对提取过程中的多糖质量浓度建立定量分析模型，其校正集及预测集样本实测值与预测值之间的散点图如图 3e 和图 3f 所示。由图可知，对于提取过程中多糖质量浓度，校正模型的决定系数 R_C 值为 0.9841，RMSECV 为 0.471，预测模型的决定系数 R_P 值为 0.9919，RMSEP 为 0.359。

2.4　3 种模型的结果比较

为了确定适合提取过程的定量模型，提高模型的建模效果和预测性能，将 PLS、iPLS 和 Si-PLS 建立的模型进行了比较，结果如表 4 所示。从表可看出，Si-PLS 模型优于其它 2 种模型，对于这 3 个关键物质中的任一个，Si-PLS 模型都能得到较好的预测结果，且结果表明利用所开发的原位光谱学系统测量提取过程中的物质是可行的。

用经典 PLS 算法对 900~2 500 nm 的全光谱信息（包含 256 个变量）的回归模型进行校正时，提取工艺中存在的许多共线变量或无关变量（如温度波动）等冗余信息都会削弱 PLS 模型的稳健性。从表 4 中还可看出，对于提取中茶皂素、蛋白质和多糖质量浓度的模型建立，iPLS 模型甚至比 PLS 模型更差。原因可能是 iPLS 建模时只有一个子区间，许多数据并没有得到充分应用，从而使校正模型和预测模型都不理想[28]。相较于 PLS 和 i-PLS，Si-PLS 建模效果最优。因为 Si-PLS 模型是在同一次的区间划分下，联合了精度较高的几个光谱子区间建立的 PLS 模型[29]。与 PLS 和 iPLS 相比，Si-PLS 能更好的捕获利用信息，较好的监测提取过程中的茶皂素浓度、蛋白质浓度和多糖浓度。

表 4　茶皂素、蛋白质和多糖质量浓度最佳模型筛选结果

Table 4　Screening results of optimal model for tea saponin, protein and polysaccharide mass concentration

指标	模型	主成分因子数	R_C	RMSECV	R_P	RMSEP
茶皂素/mg·mL^{-1}	PLS	10	0.9819	1.670	0.9932	1.080
	iPLS	7	0.8860	4.140	0.9077	3.720
	Si-PLS	10	0.9858	1.480	0.9889	1.360
蛋白质/mg·mL^{-1}	PLS	10	0.9718	0.518	0.9879	0.323
	iPLS	7	0.8918	0.998	0.9077	0.897
	Si-PLS	10	0.9766	0.472	0.9859	0.354
多糖/mg·mL^{-1}	PLS	10	0.9798	0.531	0.9947	0.309
	iPLS	7	0.8806	1.270	0.8967	1.220
	Si-PLS	10	0.9841	0.471	0.9919	0.359

3 结论

本研究表明，近红外光谱技术在测定茶皂素提取过程中的 3 个重要的质量参数——茶皂素质量浓度、蛋白质质量浓度和多糖质量浓度方面具有很大潜力。在对样品无损的前提下，能快速准确的测定提取过程中的物质变化，以判断提取终点。在建立预测模型时，Si-PLS 与传统的 PLS 和 iPLS 校正方法相比，能剔除大量冗余信息，筛选出精度较高的光谱区间，具有很大的优越性。采用近红外光谱技术联合化学计量学方法对茶皂素提取过程进行实时、原位检测，可显著提高提取的质量控制效率和质量保证。

参 考 文 献

[1] 田璨熙. 茶粕微生物脱毒(皂素)研究[D]. 长沙：湖南农业大学，2012.
TIAN C X. Study on microbial detoxification of camellia oil waste[D]. Changsha：Hunan Agricultural University，2012.

[2] 贺国文. 茶籽饼中茶皂素的提取纯化研究[D]. 长沙：中南大学，2007.
HE G W. Study on extraction and purification of tea saponin from tea seed cake[D]. Changsha：Central South University，2007.

[3] FENG J，CHEN Y，LIU X，et al. Efficient improvement of surface activity of tea saponin through Gemini-like modification by straightforward esterification[J]. Food Chemistry，2015，171：272-279.

[4] JIAN H L，LIAO X X，ZHU L W，et al. Synergism and foaming properties in binary mixtures of a biosurfactant derived from *Camellia oleifera* Abel and synthetic surfactants[J]. Journal of Colloid and Interface Science，2011，359(2)：487-492.

[5] 李金洪. 茶皂素生物活性探索及其在饲料中的应用研究[D]. 杭州：浙江大学，2001.
LI J H. Study on the biological activity of tea saponin and its application in feed[D]. Hangzhou：Zhejiang University，2001.

[6] ZHAO Y，SU R Q，ZHANG W T，et al. Antibacterial activity of tea saponin from *Camellia oleifera* shell by novel extraction method[J]. Industrial Crops & Products，2020，153(10)：112604.

[7] 禹琪芳. 茶皂素对黄羽肉鸡抗炎活性及抗氧化能力的影响[D]. 长沙：湖南农业大学，2015.
YU Q F. Effects of tea saponin on anti-inflammatory activity and antioxidant capacity in yellow broiler chicken[D]. Changsha：Hunan Agricultural University，2015.

[8] 郭海阳，莫林兰，谭海生，等. 茶皂素的美白功效及其抑菌和抗氧化活性研究[J]. 中国粮油学报，2020，35(6)：83-89.
GUO H Y，MO L L，TAN H S，et al. Whitening effect，antibacterial and antioxidant activity of tea saponin[J]. Journal of the Chinese Cereals and Oils Association，2020，35(6)：83-89.

[9] WU H，LI C Z，LI Z M，et al. Simultaneous extraction of oil and tea saponin from *Camellia oleifera* Abel. seeds under subcritical water conditions[J]. Fuel Processing Technology，2018，174：88-94.

[10] 孙劲毅. 茶皂素的分离纯化及性能研究[D]. 无锡：江南大学，2017.
SUN J Y. Study on separation，purification and properties of tea saponin[D]. Wuxi：Jiangnan University，2017.

[11] WANG Q Y，DONG X，YANG J，et al. Vesicle based ultrasonic-assisted extraction of saponins in *Panax notoginseng*[J]. Food Chemistry，2020，303(1)：125394.

[12] 王丽，王吟，胡业勤，等. BCA 法检测组分百日咳疫苗中间品蛋白含量的可行性评价[J]. 中国生物制品学杂志，2020，33(9)：1048-1053.
WANG L，WANG Y，HU Y Q，et al. Feasibility evaluation of BCA method for determination of protein content in intermediate products of component pertussis vaccine[J]. Chinese Journal of Biological Products，2020，33(9)：1048-1053.

[13] 庄玉伟，褚艳红，郭辉，等. FTIR 光谱法和杜马斯燃烧法联用鉴定奶粉品质[J]. 光谱学与光谱分析，2018，38(S1)：31-32.
ZHUANG Y W，CHU Y H，GUO H，et al. FTIR spectroscopy and Dumas combustion method were used to identify the quality of milk powder[J]. Spectroscopy and Spectral Analysis，2018，38(S1)：31-32.

[14] ZHANG W H，WU J，WENG L Y，et al. An improved phenol-sulfuric acid method for the determination of carbohydrates in the presence of persulfate

[J]. Carbohydrate Polymers, 2020, 227(1): 115332.

[15] 王思琪，胡彦波，翟丽媛，等. 豆渣可溶酸性多糖的分离纯化及结构解析[J]. 食品科学，2021，42(10)：52-57.
WANG S Q, HU Y B, ZHAI L Y, et al. Purification and structure analysis of soluble acid polysaccharide from soybean dreg[J]. Food Science, 2021, 42(10): 52-57.

[16] SIMON D, CECILE L G, DIDIER K, et al. Using near-infrared spectroscopy to determine moisture content, gel strength, and viscosity of gelatin[J]. Food Hydrocolloids, 2021, 115(6): 106627.

[17] CUI P D, ZHAO J, LIU M, et al. Non-invasive detection of medicines and edible products by direct measurement through vials using near-infrared spectroscopy: A review[J]. Infrared Physics & Technology, 2021, 115(6): 103687.

[18] HUA H M, YI Y, LI H L, et al. Monitoring the extraction process of acidic polysaccharides in *Poria cocos* by near infrared spectroscopy combined with chemometrics[J]. Infrared Physics & Technology, 2020, 113(3): 103613.

[19] TIAN W F, CHEN G J, ZHANG G R, et al. Rapid determination of total phenolic content of whole wheat flour using near-infrared spectroscopy and chemometrics[J]. Food Chemistry, 2021, 344(5): 128633.

[20] ARIAN A, NASTARAN N, MARYAM A, et al. Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy based on chemometrics for saffron authentication and adulteration detection[J]. Food Chemistry, 2021, 344(5): 128647.

[21] 白京，李家鹏，邹昊，等. 近红外光谱定性定量检测牛肉汉堡饼中猪肉掺假[J]. 食品科学，2019，40(8)：287-292.
BAI J, LI J P, ZHOU H, et al. Qualitative and quantitative detection of pork in adulterated beef patties based on near infrared spectroscopy[J]. Food Science, 2019, 40(8): 287-292.

[22] PU Y Y, O'DONNELL C, TOBIN J T, et al. Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing[J]. International Dairy Journal, 2020, 103(4): 104623.

[23] HUANG H B, YU H Y, XU H R, et al. Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review[J]. Journal of Food Engineering, 2008, 87(3): 303-313.

[24] TANG Y Y, HE X M, SUN J, et al. Comprehensive evaluation on tailor-made deep eutectic solvents (DESs) in extracting tea saponins from seed pomace of *Camellia oleifera* Abel[J]. Food Chemistry, 2021, 342(4): 128243.

[25] 沃克曼，文依. 近红外光谱解析实用指南[M]. 北京：化学工业出版社，2009：49-67.
WORKMAN J, WEYER L. A practical guide to near-infrared spectral analysis[M]. Beiijing: Chemical Industry Press, 2009: 49-67.

[26] BERHOW M A, SINGH M, BOWMAN M J, et al. Quantitative NIR determination of isoflavone and saponin content of ground soybeans[J]. Food Chemistry, 2020, 317(7): 126373.

[27] MOHAMED H, NAGY P, AGBABA J, et al. Use of near and mid infra-red spectroscopy for analysis of protein, fat, lactose and total solids in raw cow and camel milk[J]. Food Chemistry, 2021, 334(1): 127436.

[28] CHEN Q S, JIANG P, ZHAO J W. Measurement of total flavone content in snow lotus (*Saussurea involucrate*) using near infrared spectroscopy combined with interval PLS and genetic algorithm[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2010, 76(1): 50-55.

[29] 杨心浩. 基于红外光谱分析蜂王浆品质及鉴别麦卢卡蜂蜜掺假的方法研究[D]. 广州：暨南大学，2020.
YANG X H. Determination of royal jelly quality and detection of manuka honey adulterated based on infrared spectroscopy[D]. Guangzhou: Jinan University, 2020.

In-situ and Real-time Monitoring Technology of Tea Saponin Extraction

Zhang Xueli[1], Ding Yanhua[1], Ma Haile[1*], Hong Chen[1], Zhang Tao[2], Tang Yanping[2]
([1]*School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, Jiangsu*
[2]*Hunan New Jinhao Tea Oil Co. Ltd., Yongzhou 426100, Hunan*)

Abstract In this study, near-infrared spectroscopy technology combined with chemometric methods was used to monitor the changes of tea saponin mass concentration, protein mass concentration and polysaccharide mass concentration in the process of tea saponin extraction in real time. The results showed that the first-order derivation (1stDer) was the optimal preprocessing method with the highest RC, R_P value and the least RMSECV, RMSEP in all pretreatments. The performance of a model was evaluated by the correlation coefficient (R_P) and the rootmean square error (RMSEP) in the prediction set. Compared with partial least squares (PLS) and interval partial least squares (iPLS) modeling, the model built under the synergy interval partial least squares (Si-PLS) algorithm had the best robustness. The tea saponin concentration R_P=0.9889, RMSEP=1.36; protein concentration R_P=0.9859, RMSEP=0.354; polysaccharide concentration R_P=0.9919, RMSEP=0.359. Conclusion: Near-infrared spectroscopy technology combined with Si-PLS model can better monitor the extraction process of tea saponin in real time.

Keywords camellia seed cake; tea saponin; near infrared spectroscopy; *in-situ* monitoring; synergy interval partial least squares