

文章编号: 1672-8785(2022)07-0041-08

两批次番茄炒蛋蛋白质近红外光谱检测模型的维护方法

邢淑娟^{1,2} 曹 凯^{1,2} 魏文松^{1,2} 艾 鑫^{1,2} 张春江^{1,2*}

(1. 中国农业科学院农产品加工研究所, 北京 100193;

2. 农业农村部农产品加工综合性重点实验室, 北京 100193)

摘 要: 为解决光谱数据差异导致模型不稳定的问题, 研究了不同批次中式菜肴营养素含量预测模型的传递方法。以间隔 3 个月制作的番茄炒蛋样本为例, 采集光谱数据并利用理化方法测定蛋白质含量(每批次 120 个样本); 选择预测效果较好的第二批次模型作为主模型, 将分段直接标准化(Piecewise Direct Standardization, PDS)算法、模型更新(Model Updating, MP)和斜率/截距(Slope/Bias, S/B)修正法联合(PDS-MP-S/B)用于菜肴类模型传递, 分析不同 PDS 窗口数和标准集数对预测结果的影响。当 PDS 窗口数为 11 且标准集数为 100 时, PDS-MP-S/B 算法对蛋白质含量的预测结果明显优于无模型传递和单独使用 3 种算法时, 预测模型的预测集决定系数 $R^2(\text{Pred})$ 为 0.9628, 相对预测偏差(Relative Prediction Deviation, RPD)为 5.6731, 预测均方根误差(Root Mean Square Error of Prediction, RMSEP)为 0.3157。从光谱、模型、结果三个方面实现了模型传递, 提高了模型的通用性, 减少了建模成本, 为中式菜肴的快检提供了理论支持。

关键词: 近红外光谱; 不同批次样本; 模型传递; PDS-MP-S/B

中图分类号: TP79 文献标志码: A DOI: 10.3969/j.issn.1672-8785.2022.07.007

Maintenance Method of Protein Detection Model of Two Batches of Scrambled Tomatoes and Eggs Based on Near Infrared Spectroscopy

XING Shu-juan^{1,2}, CAO Kai^{1,2}, WEI Wen-song^{1,2}, AI Xin^{1,2}, ZHANG Chun-jiang^{1,2*}

(1. Institute of Food Science and Technology, Chinese Academy of Agricultural Sciences, Beijing 100193, China;
2. Key Laboratory of Agricultural Product Processing, Ministry of Agriculture and Rural Affairs, Beijing 100193, China)

Abstract: In order to solve the problem of model instability caused by spectral data difference, the transfer method of nutrient content prediction model of different batches of Chinese dishes is studied in this paper. Taking the samples of scrambled tomatoes and eggs prepared at an interval of 3 months as an example, spectral

收稿日期: 2021-12-24

基金项目: 中国农业科学院农产品加工研究所“N 专项”项目(SN2020-13)

作者简介: 邢淑娟(1996-), 女, 内蒙古赤峰人, 硕士研究生, 主要研究方向为农产品光学无损检测技术与装备。

*通讯作者: E-mail: chjiang1976@126.com

data are collected and protein content is determined by physical and chemical methods (120 samples per batch). The model of the second batch with better prediction effect is selected as the main model. A combination of PDS, MP and S/B (PDS-MP-S/B) is applied in dish model transfer to analyze the influence of different PDS window numbers and standard set numbers on the predicted results. When the number of PDS Windows is 11 and the number of standard sets is 100, the prediction result of protein content by PDS-MP-S/B algorithm is significantly better than that by no model transfer and by using the three algorithms separately. The absolute coefficient of prediction set ($R^2(\text{Pred})$) of prediction model is 0.9628, the relative prediction deviation is 5.6731, and the root mean square error of prediction is 0.3157. The model transfer is realized from three aspects of spectrum, model and result, which improves the universality of the model, reduces the cost of modeling, and provides theoretical support for the fast inspection of Chinese dishes.

Key words: near infrared spectrum; samples from different batches; model transfer; PDS-MP-S/B

0 引言

光谱技术具备快速、无损、实时分析、操作简单等优势, 现已广泛用于多个领域的定量、定性分析^[1-2]。它利用计量学的方法对光谱数据和浓度数据建立相关模型, 从而实现对未知浓度样品的预测。但有关模型仅适用于与建模样品同源的样品。若改变测试条件或样品, 该模型便缺少适配性。因此, 研究人员针对测量条件、温度、光谱仪种类及型号、地区等因素的差异性提出了模型传递, 但需要根据实际问题采用不同的算法来解决^[3]。常规使用的方法主要有直接校正(Direct Standardized, DS)算法^[4]、PDS 算法^[5]、Shenk's 算法^[6]、S/B 修正算法^[7-8]等。

目前已有模型传递研究的相关报道。Salguero-Chaparro L 等人为了实现橄榄油游离脂肪酸含量与含水率预测模型的传递, 采用 S/B 算法、PDS 算法、正交投影算法进行了传递^[9]。Ji W J 等人使用 DS 算法对水稻氮含量、有机物含量、pH 值的模型传递有效提高了模型的适用性^[10]。Schimleck L R 等人^[11]利用添加新样本的方式对桉树制浆得率的近红外预测模型进行了修正。Candolfi A 等人^[12]通过添加新样品对药物辅料的近红外定性判别模型维护进行了研究。王靖、李倩倩等人采用 S/B 算法分别对猕猴桃可溶性固形物含量、不同产地羊肉蛋白质含量进行模型修正, 有效地提高了主机的预测性能^[13-14]。除此之外, 关于不同品种

的牛肉、猪肉的评价模型修正的研究也已存在。例如, 文东东等人^[15]采用 MP 和光谱吸光度修正对不同品种牛肉的 TVB-N 定量检测模型进行了维护; 李小昱等人^[16]采用 S/B 算法、MP 算法及光谱值校正传递算法修正了山黑猪模型; 刘娇等人^[17]采用光谱和预测值同步校正(CSPV)对不同品种冷鲜猪肉 pH 高光谱检测模型进行了修正, 有效提高了模型的适用性。

根据上述内容可知, 目前针对中式菜肴近红外光谱模型在不同批次之间的传递问题鲜有研究。因此, 本文以间隔 3 个月制作的番茄炒蛋样本为例, 采集两批次的光谱数据并利用理化方法测定蛋白质含量(每批次 120 个样本), 分别建立蛋白质含量预测模型。对比选择预测效果较好的第二批次模型作为主模型, 分别采用 PDS、MP、S/B 算法对第一批数据数据进行模型传递。同时提出 PDS-MP-S/B 结合的方式, 进一步提高模型准确性, 为中式菜肴营养素含量的快检提供支持。

1 实验部分

1.1 仪器与样品

1.1.1 实验仪器

在实验过程中, 利用厨师机(KENWOOD-AT320B)进行样品的粉碎; 用全自动凯氏定氮仪(FOSS-KJEJTEC2300)进行蛋白质含量的检测; 用海洋光学公司的光谱系统进行光谱数据的采集。其中, 硬件系统由 FLAME-S-VIS-NIR-ES 光谱仪(900~1700 nm)、HL-2000-

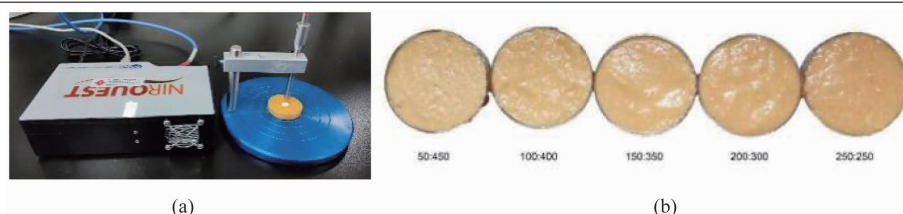


图 1 检测实物图

FHSA 光源、漫反射标准板、反射探头支架和 400 μm 实验室级反射探头组成(见图 1(a)), 光谱数据的采集软件为 OceanView1.6.7。

1.1.2 样品制备与处理

将分别采购于北京幸福荣耀超市和北京宏缘科技公司的番茄(精装)和鸡蛋用于菜肴的制备。首先,用 100 $^{\circ}\text{C}$ 的水进行漂烫 30 s,将西红柿去皮后切块(每块重 8 g 左右),并将鸡蛋搅至蛋清与蛋白均匀混合。然后按表 1 中的样品比例称重,番茄与鸡蛋的总重量为 500 g (见图 1(b))。依据中国居民膳食指南^[18],分别加入占总重 10%、1%、1%的油、盐、糖。油热后加入鸡蛋,待其被炒制成块状后加入西红柿,炒熟后盛出备用。用厨师机(功率设置为三档,时间设置为 30 s)将样品搅碎,然后采集光谱数据。

表 1 番茄炒蛋的构成比例

	鸡蛋/g	番茄/g	总重/g
1	50	450	500
2	100	400	500
3	150	350	500
4	200	300	500
5	250	250	500

1.2 实验方法

1.2.1 数据采集

样品的光谱采集时间为北京时间 10:00~14:00。此时间段内太阳高度角大于 45 $^{\circ}$,减小了太阳高度角变化对光谱测量结果的影响^[19]。采集前,将光源及光谱仪预热 30 min,采集波长范围为 900~1700 nm 的近红外光谱。测量参数选为反射率,光谱分辨率为 3.1 nm。设置光谱采集参数如下:多次扫描平均 10 次,滑动平均宽度为 5,积分时间为 196 ms。测量时,仪器探头保持垂直向下,探头与样本表面的垂直距离控制在 7 mm。每次采集目标光谱前进行白板校正,每个样本重复采集 10 次光谱。对 10 条光谱的反射率按波长求平均值,并将其作为该样本的原始光谱(见图 2)。

采用 GB 5009.5-2016《食品中蛋白质的测定》中的凯氏定氮法测定样品的蛋白质含量。

1.2.2 校正集、预测集和标准集分类方法

采用 K-S 算法划分两批次番茄炒蛋样品的校正集和预测集。其中,校正集 100 份,预测集 20 份。利用两批次光谱数据分别建立了蛋白质含量预测模型(评价结果见表 2)。

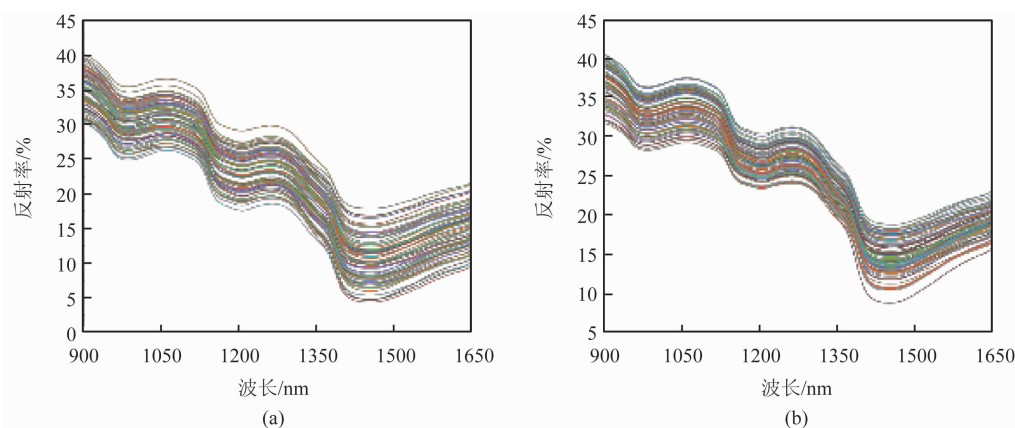


图 2 番茄炒蛋样品的反射光谱曲线图:(a)第一批次;(b)第二批次

表 2 番茄炒蛋蛋白质含量模型评价结果

样本	模型	预处理	RMSEP(g/100 g)	R ² (Pred)	RPD
第一批次	PLS	/	0.8536	0.7802	2.0982
		GLSW	0.4190	0.9622	4.2745
	PCR	/	1.1145	0.6440	1.6070
		GLSW	0.7664	0.8209	2.3369
	SVM	/	0.9109	0.7439	1.9662
		GLSW	0.5553	0.9132	3.2253
第二批次	PLS	/	0.3646	0.9631	4.9287
		GLSW	0.3879	0.9567	4.6326
	PCR	/	0.5772	0.9124	3.1133
		GLSW	0.3972	0.9548	4.5242
	SVM	/	0.6499	0.8869	2.7650
		GLSW	0.2583	0.9827	6.9570

由表 2 可知, 第二批次的模型评价结果均高于第一批次的模型。因此, 以第二批次的模型为标准, 从仪器通过模型传递算法计算后, 代入主仪器模型进行预测。采用 K-S 算法划分第二批次校正集中的标准集。标准集个数分别取 10、20、30、40、50、60、70、80、90、100。基于第一批次样品的预测集对用模型传递算法后的预测效果进行验证。

1.2.3 模型传递算法

模型传递采用多种方法联用的方式, 具体如下: 首先利用 PDS 进行光谱数据的传递, 进而利用 MP 重新建立校正模型; 最后采用 S/B 计算最终预测值, 实现 PDS^[20]、MP^[21] 和 S/B^[22] 算法联用 (PDS-MP-S/B), 对样品进行模型传递。其中, PDS 是 DS 的简化版, 主要通过调整窗口大小来实现相邻波长光谱的标准化, 从而校正受噪声影响的光谱。DS 受高变异水分的影响较弱, 因此具有一定的普适性^[20]。MP 算法主要用于重新建立校正模型。它主要是将 PDS 算法传递后第一批次标准集的光谱数据及其对应的蛋白质含量值添加到第二批次校正集中, 从而建立新的定量模型。分别将第一批次标准集与预测集传递后的光谱数据代入新模型, 进而预测蛋白质含量。S/B 算法用于计算最终预测值: 将用模型预测的第一批次标准集浓度值和标准集的实际浓度值进行一元线性拟合, 然后求出最小二乘解并将其作

为该线性模型的斜率 s 和截距 b , 最后利用 s 、 b 计算第一批次预测集的最终预测值。

1.2.4 建模方法

利用 MATLAB 软件选择全光谱范围^[23] 进行建模, 通过广义偏最小二乘 (GLSW) 预处理来消除光谱信号的基线漂移、随机噪声、颗粒度不均匀以及水分造成的影响^[24]。分别建立 GLSW^[25]、主成分回归^[26]、支持向量机模型^[27], 并根据最高 RPD 确定最佳定标模型。

1.2.5 预测效果评价

采用 R²(Pred)、RPD 和 RMSEP 对预测效果进行综合评价。其中, R²(Pred) 值描述了数据与模型 (回归线) 之间在 0~1 范围内的拟合程度, RMSEP 值测量预测的平均精度, RPD 比较不同模型的预测误差。RPD<1.5 表示该模型不能使用, RPD 为 1.5~2.0 表示该模型可以区分高值和低值, RPD 为 2.0~2.5 表示演化出近似模型预测能力, RPD 为 2.5~3.0 表示它是个良好的预测模型, RPD>3 要求该模型具有优秀的预测能力^[28-29]。

$$RMSEP = \sqrt{\frac{\sum_{i=1}^r (\hat{y}_i - y_i)^2}{r_{ul}}} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^r (y_i - \hat{y}_i)^2}{\sum_{i=1}^r (y_i - \bar{y})^2} \quad (2)$$

$$RPD = D_{pcal} / RMSEP \quad (3)$$

式中, y_i 是样本蛋白质含量的实际值, $\hat{y}_i$ 是蛋白质含量预测值, $\bar{y}$ 代表所有样本蛋白质含

表 3 不同 PDS 窗口个数下番茄炒蛋样品中蛋白质含量的预测效果($R^2(\text{Pred})$ 和 RMSEP 值)

窗口数	无预处理								
	SVM			PLS			PCR		
	RMSEP	$R^2(\text{Pred})$	RPD	RMSEP	$R^2(\text{Pred})$	RPD	RMSEP	$R^2(\text{Pred})$	RPD
3	2.2509	0.3935	0.7957	1.8979	0.4034	0.9437	1.9818	0.3894	0.9037
5	2.2533	0.3878	0.7948	1.8957	0.3991	0.9448	1.9867	0.3859	0.9015
7	2.2442	0.3850	0.7981	1.9088	0.3915	0.9383	2.0036	0.3800	0.8939
9	2.2149	0.3916	0.8086	1.8838	0.4019	0.9507	1.9794	0.3892	0.9048
11	1.4498	0.7359	1.2353	1.8776	0.2919	0.9539	1.9205	0.3923	0.9326
13	2.0946	0.3396	0.8551	1.8398	0.3878	0.9735	1.8933	0.3655	0.9460
15	2.1096	0.3277	0.8490	1.8361	0.3858	0.9754	1.8861	0.3605	0.9496
17	2.1272	0.3188	0.8420	1.8338	0.3843	0.9767	1.8829	0.3568	0.9512
19	2.1394	0.3172	0.8372	1.8292	0.3866	0.9791	1.8763	0.3579	0.9545
21	2.1366	0.3237	0.8382	1.8130	0.3978	0.9879	1.8545	0.3691	0.9658
23	2.1298	0.3256	0.8409	1.8048	0.4016	0.9924	1.8441	0.3738	0.9712
25	2.1182	0.3263	0.8455	1.7973	0.4041	0.9965	1.8331	0.3777	0.9770
27	1.6885	0.5572	1.0607	1.7892	0.4076	1.0010	1.8223	0.3821	0.9828
29	2.1010	0.3264	0.8525	1.7827	0.4109	1.0047	1.8152	0.3855	0.9867
31	2.0942	0.3248	0.8552	1.7772	0.4129	1.0078	1.8082	0.3872	0.9905

量的平均值, r 是校正集的样本数, r_{val} 是预测集的样本数, D_{pcal} 是标准差。

2 结果与讨论

2.1 不同 PDS 窗口的模型传递效果

PDS 算法中的窗口个数是可变的。不同的窗口数会影响模型传递后蛋白质含量的预测效果。在实验过程中, 将 PDS 窗口个数设置为 3~31(窗口间距为 2), 分别建立 PLS、PCR、SVM 模型(具体结果见表 3)。

不同 PDS 窗口个数对 $R^2(\text{Pred})$ 值影响较大。在无预处理前提下, 建立 PLS 预测模型。 $R^2(\text{Pred})$ 值在 0.3172~0.7359 范围内; RMSEP 的波动也相对较大, 在 1.4498~2.2533 g/100g 之间。在 PCR 预测模型中, $R^2(\text{Pred})$ 值在 0.2919~0.4109 范围内; RMSEP 的波动也相对较大, 在 1.7772~1.9088 g/100g 之间。在 SVM 预测模型中, $R^2(\text{Pred})$ 值在 0.3568~0.4129 范围内; RMSEP 的波动也相对较大,

在 1.8082~2.0036 g/100g 之间。采用 PDS 确定窗口宽度时, 若窗口宽度过小, 选定的数据则会缺失光谱仪器之间差异的信息, 传递的结果差; 若窗口宽度过大, 选定的数据则会有与光谱仪器不相关的信息, 计算量大^[30]。经过反复筛选确定, 传递窗口为 11 时模型的 $R^2(\text{Pred})$ 值、RMSEP 值综合效果最佳, 可得 PDS 方法的传递结果在此条件下达到最佳水平。因此, 在后续的蛋白质不同标准集个数的模型传递中, PDS 窗口个数定为 11 且传递前不经预处理。

2.2 不同标准集个数与 PDS 联合的模型传递效果

为探讨不同标准集个数对模型传递后蛋白质含量预测效果的影响, 设定标准集样品个数分别为 10、20、30、40、50、60、70、80、90、100, 研究不同标准集个数对蛋白质含量预测效果的影响(结果见表 4 和图 3)。

表 4 不同样本数下番茄炒蛋样品中蛋白质含量的预测效果($R^2(\text{Pred})$ 和 RMSEP 值)

样本数	无预处理						GLSW 预处理					
	PLS		PCR		SVM		PLS		PCR		SVM	
	$R^2(\text{Pred})$	RMSEP	$R^2(\text{Pred})$	RMSEP	$R^2(\text{Pred})$	RMSEP	$R^2(\text{Pred})$	RMSEP	$R^2(\text{Pred})$	RMSEP	$R^2(\text{Pred})$	RMSEP
10	0.4075	1.8036	0.4064	1.8098	0.4048	1.7217	0.6212	1.8506	0.595	1.9596	0.7113	1.2197
20	0.4065	1.7389	0.4096	1.7581	0.4135	1.6256	0.6156	1.4058	0.5994	1.4456	0.7571	0.9289
30	0.543	1.5372	0.4117	1.6949	0.683	1.0407	0.5891	1.3754	0.6672	1.2283	0.8465	0.7277
40	0.6729	1.1989	0.762	1.0764	0.7209	0.9696	0.7162	1.1374	0.6756	1.2215	0.9131	0.5402
50	0.7439	1.0707	0.7852	0.995	0.7178	0.9786	0.7932	0.9856	0.7795	1.0428	0.9487	0.4083
60	0.7322	1.071	0.7702	0.9974	0.6741	1.0627	0.7877	0.9601	0.5029	1.9329	0.9499	0.4124
70	0.7896	0.9824	0.7976	0.9432	0.7053	1.0021	0.8052	0.9182	0.7996	0.9581	0.9608	0.3547
80	0.797	0.9439	0.7941	0.9303	0.7086	0.9993	0.8151	0.8801	0.7904	0.956	0.9511	0.4029
90	0.7885	0.948	0.7851	0.9346	0.7661	0.8816	0.6946	1.0713	0.8474	0.8379	0.9721	0.3001
100	0.7516	0.9805	0.7811	0.9297	0.7694	0.8618	0.7750	0.9256	0.7192	1.0375	0.9727	0.2970

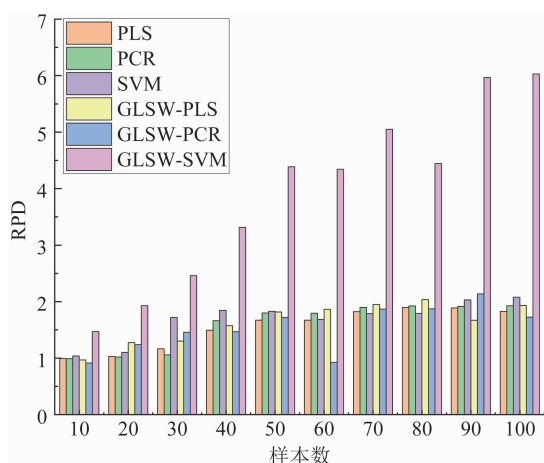


图 3 不同样本数下番茄炒蛋样品中蛋白质含量的预测效果(RPD 值)

由表 4 和图 3 可知,随着标准集样品个数的增加,模型传递后预测效果随标准集个数的增加而提升,具体如下:蛋白质含量的 $R^2(\text{Pred})$ 值和 RPD 值总体呈上升趋势,但略有波动, RMSEP 值总体呈下降趋势。对比预处理方法与 PLS、PCR、SVM 三种建模方式的结果表明:当标准集个数为 100 时,模型传递并经 GLSW 预处理后建立的 SVM 模型预测效果最好。 $R^2(\text{Pred})$ 值为 0.9727, RMSEP 值为 0.2970 g/100g 且 $\text{RPD} > 3$, 可实现第一批次光谱数据在第二批次蛋白质含量模型的预测。当

标准集个数为 90 时,模型传递并经 GLSW 预处理后所建 PCR 模型的 $R^2(\text{Pred})$ 值为 0.8474, RMSEP 值为 0.8379 g/100g。但其他模型的 RMSEP 值较高, $\text{RPD} < 3$ 不满足模型需求,其原因仍有待进一步研究。实验结果符合预测效果随标准集个数的增加而提升的规律,能够实现第一批次光谱数据在第二批次蛋白质含量模型的预测。

标准集样品个数的选择是 PDS-MP 算法达到模型传递预测效果的关键。与模型更新方法相同, PDS-MP 算法的预测效果也随着标准集个数的增加而提高。尽可能选择多个标准集样品,对提升 PDS-MP 算法的预测效果具有重要作用。后续研究将在少量标准集样品个数的情况下,实现更好的模型传递预测效果。

2.3 不同传递算法的修正效果比较

取两批次的光谱数据,将第一批次光谱数据分别采用 PDS、MP、S/B、PDS-MP-S/B 进行模型传递,代入第二批次蛋白质含量模型中,并比对未进行模型传递的第一批次光谱数据预测效果。表 5 列出了蛋白质含量预测的 $R^2(\text{Pred})$ 和 RMSEP 值等参数。其中, PDS 采用 11 个窗口,标准集个数采用 100 个样品。

表 5 采用不同传递算法时番茄炒蛋样品中蛋白质含量的预测效果($R^2(\text{Pred})$ 和 RMSEP 值)

模型传递方法	建模方法	预处理方法	$R^2(\text{Pred})$	RMSEP	RPD
PDS-MP-S/B	PLS	/	0.7476	1.0478	1.7093
		GLSW	0.7767	0.9847	1.8188
	PCR	/	0.7709	0.9961	1.7980
		GLSW	0.7215	1.1048	1.6211
	SVM	/	0.7801	0.8688	2.0615
		GLSW	0.9628	0.3157	5.6731

由表 5 可知, 分别采用 PDS、MP 和 S/B 三种传递算法对第一批光谱数据传递后所建的模型对蛋白质含量的预测结果并未显著提高。PDS 或 S/B 算法传递后, 经 GLSW 预处理建立的 PLS 模型的 $R^2(\text{Pred})$ 较低, 且 RPD 值仅满足区分高值和低值。显然, 单独使用 PDS 或 S/B 方法不能够实现两批次样本之间的模型传递。MP 算法传递后, 经 GLSW 预处理建立的 PLS 模型的 $R^2(\text{Pred})$ 较高, 且 RPD 值满足模型具有较好预测能力的要求, 但 $R^2(\text{Pred})$ 还需进一步提高。采用 PDS-MP-S/B 算法对蛋白质含量的预测结果明显优于无模型传递和单独使用 3 种算法的情况。所建模型的 $R^2(\text{Pred})$ 为 0.9628, RPD 为 5.6731, RMSEP 值为 0.3157 g/100g。因此, PDS-MP-S/B 可以作为不同批次样本间的模型传递方法。

3 结束语

由于不同批次的光谱数据采集时间、环境、原料等存在很大差异, 单纯使用常规模型传递方法对所有样本并非全部适用的。本文首先用 PDS 算法实现两批次光谱数据矩阵转换, 从而完成第一步模型传递; 然后利用 MP 的方式添加第一批的光谱数据和浓度值到第二批次的原模型中来重新建立模型, 从而实现第二步模型传递; 最后利用 S/B 法修正第一批浓度值数据来达到模型传递的目的。与单一的模型传递方法相比, PDS-MP-S/B 联用从光谱、模型、结果三个方面实现模型传递, 提高了模型的通用性, 并降低了建模成本。传递后模型的 $R^2(\text{Pred})$ 为 0.9628, RPD 为 5.6731, RMSEP 值为 0.3157 g/100g。本文可为近红外

光谱技术在中式菜肴中快检的应用提供支持。未来的研究可尝试寻找能够实现机器自动校准的模型传递算法, 进而解决现实应用时遇见的瓶颈问题。

参考文献

- [1] 郭志明, 郭闯, 王明明, 等. 果蔬品质安全近红外光谱无损检测研究进展 [J]. 食品安全质量检测学报, 2019, 10(24): 8280-8288.
- [2] 崔艳莉, 冀晓磊, 古丽菲娅, 等. 近红外光谱在果蔬品质无损检测中的研究进展 [J]. 农产品加工(学刊), 2007, 6(7): 84-86.
- [3] 谢丽娟, 应义斌, 应铁进, 等. 可见/近红外光谱分析技术鉴别转基因番茄叶 [J]. 光谱学与光谱分析, 2008, 28(5): 1062-1066.
- [4] 柳培忠, 张丽萍, 李卫军, 等. 基于 DS 算法的玉米近红外定性分析光谱校正方法研究 [J]. 光谱学与光谱分析, 2014, 34(6): 1533-1537.
- [5] 刘翠玲, 李天瑞, 位丽娜, 等. 直接标准化算法在食用油酸值和过氧化值上的近红外光谱模型传递的研究 [J]. 光谱学与光谱分析, 2017, 37(10): 3042-3050.
- [6] 高云, 高建芹, 王小天. 油菜籽品质指标近红外数学模型的仪器间传递优化 [J]. 江苏农业科学, 2012, 40(3): 292-295.
- [7] 钟雄斌. 基于高光谱技术的不同品种猪肉品质检测模型维护方法研究 [D]. 武汉: 华中农业大学, 2014.
- [8] 刘娇. 基于高光谱技术的不同品种猪肉品质检测模型传递方法研究 [D]. 武汉: 华中农业大学, 2015.
- [9] Salguero-Chaparro L, Palagos B, Peña-Rodríguez F, et al. Calibration Transfer of Intact Olive NIR Spectra Between a Pre-dispersive Instrument and a

- Portable Spectrometer [J]. *Computers and Electronics in Agriculture*, 2013, **96**: 202–208.
- [10] Ji W J, Li S, Chen S C, et al. Prediction of Soil Attributes Using the Chinese Soil Spectral Library and Standardized Spectra Recorded at Field Conditions [J]. *Soil & Tillage Research*, 2016, **155**: 492–500.
- [11] Schimleck L R, Kube P D, Raymond C A, et al. Extending Near Infrared Reflectance (NIR) Pulp Yield Calibrations to New Sites and Species [J]. *Journal of Wood Chemistry and Technology*, 2006, **26**(4): 299–311.
- [12] Candolfi A, Massart D L. Model Updating for the Identification of NIR Spectra from a Pharmaceutical Excipient [J]. *Applied Spectroscopy*, 2000, **54**(1): 48–53.
- [13] 王靖, 郭中华, 何凤杰, 等. 不同产地羊肉蛋白质高光谱检测模型的维护方法 [J]. *食品工业*, 2018, **39**(6): 118–121.
- [14] 李倩倩, 刘大洋, 杨彪, 等. 斜率/截距算法在猕猴桃可溶性固形物含量便携式检测中的应用 [J]. *食品科学*, 2018, **39**(14): 257–262.
- [15] 文东东, 李小昱, 赵政, 等. 不同品种牛肉新鲜度光谱检测模型的维护方法 [J]. *食品安全质量检测学报*, 2012, **3**(6): 621–626.
- [16] 李小昱, 钟雄斌, 刘善梅, 等. 不同品种猪肉 pH 值高光谱检测的模型传递修正算法 [J]. *农业机械学报*, 2014, **45**(9): 216–222.
- [17] 刘娇, 李小昱, 金瑞, 等. 不同品种冷鲜猪肉 pH 值高光谱检测模型的传递方法研究 [J]. *光谱学与光谱分析*, 2015, **35**(7): 1973–1979.
- [18] 《中国居民膳食指南》(2007)(中国营养学会制定) 第三部分 中国居民平衡膳食宝塔(25) [J]. *饮食科学*, 2016, **23**(4): 32–33.
- [19] 李鸿强, 孙红, 李民赞. 基于可见/短波近红外光谱检测结球甘蓝维生素 C 含量 [J]. *农业工程学报*, 2018, **34**(8): 269–275.
- [20] 林振兴, 邬蓓蕾, 王豪, 等. PDS 用于不同温度下的近红外光谱模型传递研究 [J]. *分析测试学报*, 2008, **27**(12): 1330–1333.
- [21] 褚小立, 袁洪福, 王艳斌, 等. 近红外稳健分析校正模型的建立(I)——样品温度的影响 [J]. *光谱学与光谱分析*, 2004, **24**(6): 666–671.
- [22] 吉纳玉, 李明, 吕文博, 等. 基于 Slope/Bias 算法的相近种类水果模型传递研究 [J]. *光谱学与光谱分析*, 2017, **37**(1): 227–231.
- [23] Balabin R M, Smirnov S V. Variable Selection in Near-infrared Spectroscopy: Benchmarking of Feature Selection Methods on Biodiesel Data [J]. *Analytica Chimica Acta*, 2011, **692**(1): 63–72.
- [24] 齐海军, Karnieli A, 李绍稳. Y-梯度广义最小二乘加权校正的土壤速效氮野外原位光谱预测 [J]. *光谱学与光谱分析*, 2018, **38**(1): 171–175.
- [25] 郝勇, 杜娇君, 张书敏, 等. 不同颜色冬枣可溶性固形物可见-近红外光谱分析模型构建 [J]. *光谱学与光谱分析*, 2021, **41**(11): 3385–3391.
- [26] Guy S, Eyal B D, Gil E. Quantitative Assessment of Hydrocarbon Contamination in Soil Using Reflectance Spectroscopy: A “Multipath” Approach [J]. *Applied Spectroscopy*, 2013, **67**(11): 1–14.
- [27] Moros M, Emeis K, Risebrobakken B, et al. Sea surface temperatures and ice rafting in the Holocene North Atlantic: climate influences on northern Europe and Greenland [J]. *Quaternary Science Reviews*, 2004, **23**(20): 2113–2126.
- [28] Saeys W, Mouazen A M, Ramon H. Potential for Onsite and Online Analysis of Pig Manure using Visible and Near Infrared Reflectance Spectroscopy [J]. *Biosystems Engineering*, 2005, **91**(4): 393–402.
- [29] Brasil Y L, Cruz-Tirado J P, Barbin D F. Fast online estimation of quail eggs freshness using portable NIR spectrometer and machine learning [J]. *Food Control*, 2022, **131**: 108418.
- [30] 韩君, 郝远, 方洪壮. 缙沙坦胶囊辅料含量近红外光谱模型传递研究 [J]. *云南民族大学学报(自然科学版)*, 2017, **26**(5): 365–368.