

• 药剂与工艺 •

在线近红外光谱监测桂枝茯苓胶囊流化床干燥过程水分的方法研究

王 晴^{1,2}, 徐芳芳^{2,3,4}, 张 欣^{2,3,4}, 徐 冰⁵, 李执栋^{2,3,4}, 杜 慧¹, 张永超¹, 相继芬^{2,3,4},
包乐伟^{2,3,4}, 王振中^{2,3,4}, 肖 伟^{1,2,3,4*}

1. 南京中医药大学, 江苏 南京 210000
2. 江苏康缘药业股份有限公司, 江苏 连云港 222001
3. 中药制药过程新技术国家重点实验室, 江苏 连云港 222001
4. 中成药智能制造国家地方联合工程研究中心, 江苏 连云港 222001
5. 北京中医药大学, 北京 100029

摘要: **目的** 采用在线近红外光谱(NIRS)技术,建立桂枝茯苓胶囊流化床干燥过程水分实时监测模型。**方法** 通过 NIRS 漫反射探头采集 16 个生产批次共 176 个样本进行建模,优选移动窗口平滑法进行光谱预处理,采用间隔偏最小二乘法(siPLS)结合移动窗口偏最小二乘法(mwPLS)筛选特征变量为 4 759.45~5 338.00 cm^{-1} 、5 503.84~6 101.67 cm^{-1} 、8 512.25~8 809.24 cm^{-1} ,采用偏最小二乘(PLS)法建立水分含量多变量校正模型。**结果** 水分预测的交叉验证均方根误差(RMSECV)为 0.243%,性能偏差比(RPD)值为 13.384,预测相对偏差(RSEP)为 0.270%。以 8 个生产批次对在线监控方法的可靠性进行持续验证,结果 40 个样本的相对预测误差均小于 4.7%。干燥过程水分实时监测趋势图显示可准确判断干燥终点,终点样本水分含量位于控制限内。**结论** 在线 NIRS 结合 PLS 建立的定量模型,可应用于生产规模桂枝茯苓胶囊流化床干燥过程水分含量在线监控且预测性能稳健、准确。

关键词: 桂枝茯苓胶囊;流化床干燥;在线近红外光谱技术;过程分析技术;实时监测;间隔偏最小二乘法;移动窗口偏最小二乘法;水分;多变量校正模型;交叉验证均方根误差;性能偏差比;预测相对偏差

中图分类号: R283.6 文献标志码: A 文章编号: 0253-2670(2019)22-5429-10

DOI: 10.7501/j.issn.0253-2670.2019.22.008

Study on on-line near-infrared spectroscopic for moisture monitoring in a fluidized bed drying process of Guzhi Fuling Capsules

WANG Qing^{1,2}, XU Fang-fang^{2,3,4}, ZHANG Xin^{2,3,4}, XU Bing⁵, LI Zhi-dong^{2,3,4}, DU Hui¹, ZHANG Yong-chao¹, XIANG Ji-fen^{2,3,4}, BAO Le-wei^{2,3,4}, WANG Zhen-zhong^{2,3,4}, XIAO Wei^{1,2,3,4}

1. Nanjing University of Chinese Medicine, Nanjing 210000, China
2. Jiangsu Kanion Pharmaceutical Co., Ltd., Lianyungang 222001, China
3. State Key Laboratory of New-tech for Chinese Medicine Pharmaceutical Process, Lianyungang 222001, China
4. National & Local Joint Engineering Research Center on Intelligent Manufacturing of TCM, Lianyungang 222001, China
5. Beijing University of Chinese Medicine, Beijing 100029, China

Abstract: Objective To establish a real-time moisture monitoring model for the fluidized bed drying process of Guizhi Fuling Capsules (GFC) by using online near-infrared spectroscopy (NIRS). **Methods** A total of 176 samples from 16 production batches were collected by NIRS diffuse reflection probe for modeling. The moving window average smoothing method was used for spectral

收稿日期: 2019-06-06

基金项目: 国家科技重大专项: 基于功效成分群的中药口服固体制剂先进制药与信息化技术融合示范应用(2018ZX09201010-004); 国家标准化专项: 桂枝茯苓胶囊标准化建设(ZYBZH-C-JS-28); 国际重大新药创制专项: 桂枝茯苓胶囊美国 III 期临床试验研究(2018ZX09737015); 国家工信部智能制造综合标准化与新模式应用项目(KYYY20170820)

作者简介: 王 晴(1994—),女,硕士研究生,研究方向为新型中药制剂的研究与开发。E-mail: 1477017198@qq.com

***通信作者** 肖 伟,研究员级高级工程师,博士,研究方向为中药新药的研究与开发。Tel: (025)87181851 E-mail: kanionlunwen@163.com

preprocessing. The characteristic variables were 4 759.45—5 338.00 cm^{-1} , 5 503.84—6 101.67 cm^{-1} , and 8 512.25—8 809.24 cm^{-1} , which were screened by the interval partial least squares method (siPLS) combined with the moving window partial least squares (mwPLS). The partial variable least squares (PLS) method was used to build a multivariate correction model for moisture. **Results** The root mean square error of cross-validation (RMSECV) of predicted moisture was 0.243%, the ratio of prediction to deviation (RPD) was 13.384, and the relative standard error of prediction (RSEP) was 0.270%. The reliability of the online monitoring method was continuously verified by eight production batches. The relative error of 40 samples was less than 4.7%, indicating that the PLS quantitative model prediction performance was robust and accurate. The real-time monitoring trend chart of the moisture in the drying process can accurately determine the drying end point, and the moisture content of the end sample was within the control limit. **Conclusion** The quantitative model established by online NIRS combined with PLS can be applied to the on-line monitoring of moisture content in the fluidized bed drying process of production scale GFC and the prediction performance was robust and accurate. **Key words:** Guizhi Fuling Capsules; fluidized bed drying; on-line near infrared spectroscopy; process analytical technology; real-time monitoring; interval partial least squares method (siPLS); moving window partial least squares (mwPLS); moisture; multivariate correction model; root mean square error of cross-validation (RMSECV); ratio of prediction to deviation (RPD); relative standard error of prediction (RSEP)

桂枝茯苓胶囊 (Guizhi Fuling Capsules, GFC) 成型工艺由混合、湿法制粒、干燥、整粒和总混等工序组成,其中湿颗粒干燥由流化床干燥设备完成。颗粒含水量影响颗粒物料的流动性、可压缩性和吸湿稳定性,进而影响胶囊装量和稳定性等关键质量属性^[1]。而目前生产采用的离线检测方法实时性差,影响干燥终点的判断,不利于质量控制。因此,流化床干燥过程水分在线控制是实现 GFC 质量批内和批间一致性的关键步骤。

近年来,基于在线近红外光谱法 (near-infrared spectroscopy, NIRS) 的颗粒流化床干燥过程水分实时监测的研究是工业生产中一个热点问题^[2]。从研究对象方面,已发表的文献侧重于化学药品^[3],中药相关报道较少^[4-7]。从 NIRS 在工业现场的实施方面,离线取样并建模分析的研究较多,而实时在线监控研究较少。Otsuka 等^[8]应用 NIRS 法实时监测实验室级别流化床制粒过程颗粒含水量及粒径;张广仁等^[9]离线采集了散结镇痛胶囊干燥过程中 67 个样本 NIRS,对水分含量检测的可行性进行了分析研究;易夏等^[10]对注射用五头孢唑林钠干燥过程样本水分含量在线监测方法进行探究,实现了样本生产规模取样,但光谱获取方式为离线采集,且未进行实际生产验证。因此,在中药生产过程复杂工况条件下通过在线 NIRS 法实现流化床干燥过程水分测定的可行性和精确度仍有待进一步研究。

本实验以 GFC 生产规模条件下的流化床干燥过程为载体,对流化床设备进行改造,搭建透光蓝宝石视窗和近红外无损检测平台,对生产过程代表性样品进行近红外分析,优选光谱预处理和变量筛选方法,建立颗粒干燥过程水分预测模型,并对该

模型在实际干燥过程中应用的准确性和稳健性进行验证。

1 仪器与试剂

1.1 仪器

Antaris MX 傅里叶变换近红外光谱仪,配备 Reflector 探头,美国 Thermo 公司;XY-105MW 型卤素快速水分测定仪,常熟幸运电子设备有限公司。

1.2 试剂

GFC 湿颗粒干燥过程中不同水分含量颗粒样品,由江苏康缘药业股份有限公司提供。28 批样品批次分别为 180808、180809、180810、180811、180922、180923、180924、180925、181010、181011、181012、181013、181118、181119、181120、181121、190407、190408、190409、190410、190411、190412、190501、190502、190503、190504、190505、190506。

2 方法

2.1 样品采集

近红外探头安装示意图见图 1-a。因为制粒和干燥是不连续过程,故将近红外探头置于可移动支架上,在工人上料、卸料时可移动至旁边。开始干燥时,将支架推至窗口处,探头平面与锅体窗口相切,防止扫描到空气引起运行异常。

流化床视窗和取样口位置示意图见图 1-b,圆形透光蓝宝石视窗开口于锅体中间位置,透过窗口可清晰地观察到物料运动状态,干燥过程中物料不黏附,取样口和视窗位于同一水平线上。探头采集的信号经光纤传递给控制系统。

采用积分球漫反射方式于室温下在线采集近红外光谱,光谱采集时间间隔为 3 min,扫描范围 10 000~4 000 cm^{-1} ,分辨率 8 cm^{-1} ,4 倍增益,扫

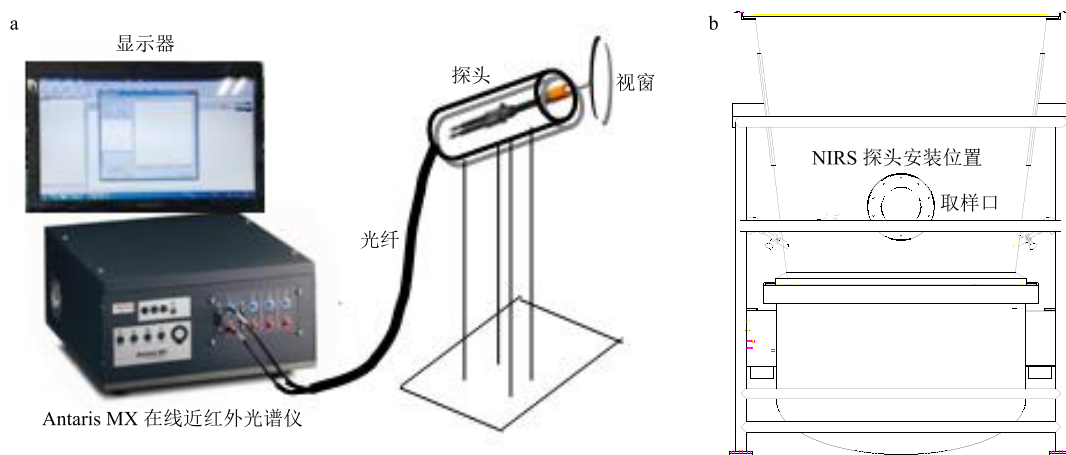


图 1 NIRS 探头及取样口位置示意图

Fig. 1 NIRS probe and sampling port position

描次数 32 次。扫描光谱的同时，从取样口中手动取样。收集 4 个月内 16 个生产批次共 176 个样本，每份样本约 10 g 颗粒。详细记录操作组、干燥时间、停止干燥时颗粒温度及干燥后颗粒的水分含量。

2.2 水分测定

精密称取约 5.0 g 待测样品，平铺在水分测定仪样品盘上，于 105 ℃ 下加热 10 min，根据仪器读数即得。

2.3 光谱预处理

环境温湿度变化、设备振动、现场噪音等因素会使 NIRS 产生噪声信号，流化床干燥过程中颗粒粒径、密度变化等亦会引起基线漂移。为了减少光谱中信号中的噪声以及无关信息，必须采用适当的预处理方法进行光谱预处理。移动窗口平均算法 (window moving average algorithm) [11] 是将光谱数据定义为矩阵 X ， X 中的每个元素为 x_{ij} (其中 i 表示第 i 条光谱， j 表示第 j 个波数变量)，校正后的元素定义为 x_{ij}^s ，则 k 个窗口平滑之后的校正值 x_{ij}^s 的计算公式如下。

$$x_{ij}^s = \left(\sum_{i=i-\frac{k-2}{2}}^{i+\frac{k-2}{2}} x_{ij} \right) / k, \text{ 其中 } k=2n+1 \text{ (} n \text{ 为正整数)}$$

其他常用的预处理方法还有卷积平滑 (Savitzky-Golay smooth)、矢量归一化 (normalization)、S-G 平滑加一阶导数 (S-G 1st)、S-G 平滑加二阶导数 (S-G 2nd)、基线校正 (baseline)、多元散射校正 (multiplicative scatter correction, MSC)、标准正态变换 (standard normal variate transformation, SNV)。S-G 平滑是通过多项式来对移动窗口内的数据进行多项式最小二乘拟合以消除噪声；矢量归一化可消除光程变化或样品稀

释变化等对光谱产生的影响；一阶导数可以消除背景的常数平移；二阶导数可以消除线性背景平移；基线校正主要是扣除仪器背景或漂移对光谱信号的影响；多元散射校正与标准正态变换效果近似，可降低样本中固体颗粒分布不均、大小不同和附加的散射对光谱的影响 [12]。

2.4 数据处理

采用 SIMCA 13.0 (Umetrics, Sweden) 软件进行主成分分析和光谱异常值辨别，采用 Unscramble X 10.4 (Camo software AS, Norway) 软件对光谱进行预处理，采用 Matlab 7.0 (Mathwork Inc., 美国) 软件进行样本划分、变量筛选及模型构建，采用 Origin 8.0 (OriginLab, 美国) 软件绘图。

3 结果和分析

3.1 光谱相似度评价

于 6 个批次光谱 (180808、180809、180810、180811、180922、180923) 中各随机选取 1 条光谱，利用夹角余弦法进行光谱相似度评价，结果如表 1 所示，6 条光谱间相似度均大于 0.999，表明保持

表 1 光谱相似度评价结果

Table 1 Evaluation of spectral similarity

光谱	相似度					
	S1	S2	S3	S4	S5	S6
S1	1.000 0	1.000 0	0.999 9	0.999 5	0.999 7	0.999 7
S2	1.000 0	1.000 0	0.999 9	0.999 6	0.999 8	0.999 8
S3	0.999 9	0.999 9	1.000 0	0.999 7	0.999 9	0.999 9
S4	0.999 5	0.999 6	0.999 7	1.000 0	0.999 5	0.999 7
S5	0.999 7	0.999 8	0.999 9	0.999 5	1.000 0	0.999 9
S6	0.999 7	0.999 8	0.999 9	0.999 7	0.999 9	1.000 0

NIRS 探头与锅体视窗相切时,得到的 NIRS 相似度高,重复性好,仪器安装方法和光谱采集方法稳定可靠。

3.2 水分参考值测定结果

GFC 工艺规程规定,干燥后颗粒水分含量应为 3%~5%,操作人员根据干燥时间和物料温度判断干燥终点。批次 180808~181121 内 176 个样本水分含量测定结果见表 2。终点样本水分含量(w)指结束干燥后物料温度降至室温时样本水分含量;干燥时间(t)指加热时间;停止温度(T)指停止加热时物料温度,由嵌在锅体上的温度计读数得到;操作组指不同的操作工人,分为 2 组,每组 3 人。当干燥过程持续时间长,停止温度高时,终点样本含水量较低。

表 2 流化床干燥过程样本水分含量测定结果

Table 2 Determination of moisture content of sample in fluidized bed drying process

批次	样本数	$w/\%$	t/min	$T/^\circ\text{C}$	操作组
180808	7	4.97	29	42	1
180809	10	3.98	38	47	1
180810	10	4.40	33	47	1
180811	8	3.95	35	52	1
180922	10	4.17	30	48	2
180923	11	4.67	30	47	2
180924	11	4.02	31	56	2
180925	12	4.16	30	55	2
181010	15	4.92	28	54	1
181011	13	4.24	30	54	1
181012	12	3.75	33	57	1
181013	11	4.55	33	54	1
181118	8	4.34	34	56	2
181119	13	4.52	31	55	2
181120	13	4.84	31	53	2
181121	12	4.23	30	55	2

3.3 离群样本辨识和样本划分

对 176 组光谱数据做主成分分析(PCA),绘制所有批次的主成分 1 和主成分 2 的得分图,即图 2-a,第 1 主成分可解释 93.9%的变量信息,第 2 主成分可解释 4.72%的变量信息。将 180809、180922 和 181118 批次中部分样本在第 1 主成分和第 2 主成分的得分绘制于图 2-b 中,可以看出,同一批次样本在得分图中分布在相近的区域,不同批次相距较远。表明生产过程中批间差异较为明显。离群样本会影

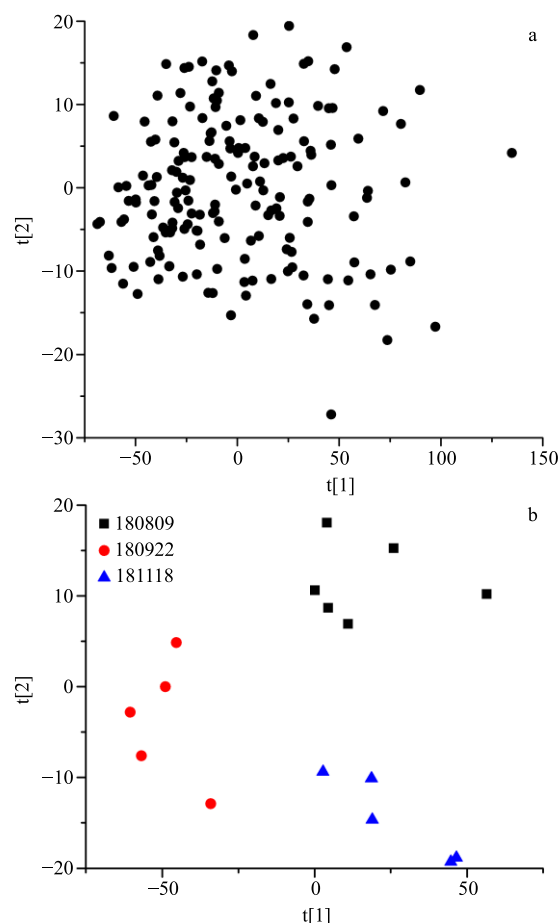


图 2 所有样本 (a) 以及 180809、180922、181118 批次部分样本 (b) 的主成分得分图

Fig. 2 Scores plot of all samples (a) and samples from 180809, 180922, and 181118 batches (b)

响模型稳健性和准确性,因此在建模之前应准确识别离群样本。Hotelling's T^2 统计量表征的是采样点在主成分的特征空间中与均值之间的距离,指示每个样本信息在幅值和变化趋势上偏离主成分模型的程度^[13]。不同样本 Hotelling's T^2 距离如图 3 所示,有 6 个样本超出 Hotelling's T^2 95%置信区间,分别是 26、109、121、128、130、131 号样本,建模之前剔除异常样本。针对 170 个样本对应的 NIRS,采用 Kennard-Stone 算法^[14]将所有样本划分为校正集和验证集,如表 3 所示,其中校正集样本占 80%,验证集样本占 20%。校正集含水量 3.75%~15.32%,验证集含水量 3.95%~13.14%,校正集和验证集的样本水分含量的平均值和标准偏差分别是 8.25%、3.05%和 7.03%、2.61%,可见校正集数据范围覆盖了验证集数据范围。

3.4 光谱预处理

3.4.1 光谱预处理方法的选择 使用移动窗口平滑

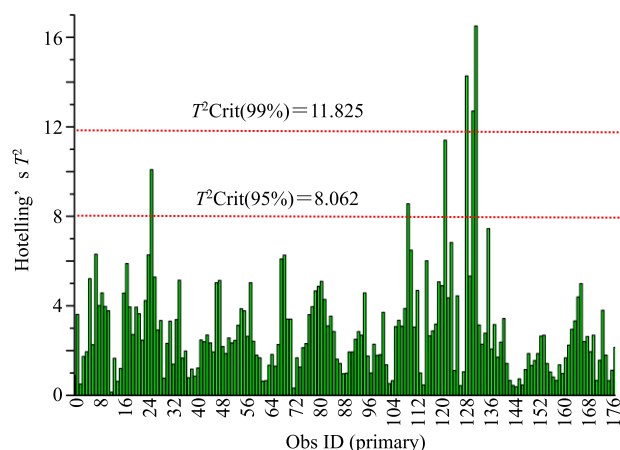
图 3 Hotelling's T^2 柱状图Fig. 3 Hotelling's T^2 column plot

表 3 校正集和验证集样本参考方法测定值的数据统计结果

Table 3 Reference measurement results of samples in calibration and prediction sets

样本集	样本数	参考值范围/%	平均值/%	SD/%
校正集	136	3.75~15.32	8.25	3.05
验证集	34	3.95~13.14	7.03	2.61

(9 点)、S-G 平滑 (9 点)、矢量归一化、S-G 1st (9 点)、S-G 2nd (9 点)、基线校正、MSC、SNV 对原始光谱进行预处理之后, 建立水分含量偏最小二乘 (PLS) 定量模型, 不同预处理方法对模型性能的影响见表 4。本研究以性能偏差比 (ratio of performance to deviation, RPD) 和预测相对偏差 (relative standard

表 4 不同预处理方法对水分含量定量模型性能的影响

Table 4 Effect of different pretreatment methods on performance of quantitative model of moisture content

预处理方法	潜变量	校正集				验证集				
		R_{cal}	RMSEC/%	RMSECV/%	$BIAS_{cal}$	R_{pre}	RMSEP/%	$BIAS_{pre}$	RPD	RSEP/%
无预处理	15	0.998	0.179	0.295	0.123	0.995	0.254	0.205	10.263	3.397
移动窗口平滑	8	0.997	0.223	0.262	0.167	0.997	0.229	0.184	11.377	3.036
S-G 平滑	8	0.997	0.218	0.330	0.187	0.996	0.249	0.188	10.499	3.217
归一化	8	0.996	0.282	0.381	0.211	0.996	0.234	0.185	11.149	3.128
S-G 1st	8	0.996	0.278	0.377	0.215	0.995	0.273	0.230	9.572	3.664
S-G 2nd	15	0.992	0.392	0.592	0.301	0.993	0.354	0.281	7.370	4.732
基线校正	9	0.997	0.232	0.336	0.174	0.996	0.266	0.217	9.796	3.560
MSC	10	0.998	0.208	0.318	0.162	0.997	0.230	0.190	11.367	3.069
SNV	10	0.998	0.209	0.319	0.163	0.997	0.229	0.190	11.367	3.065

error of prediction, RSEP) 为评价指标, RPD 越大表示模型预测能力越高, RSEP 越小, 表示预测偏差越小, 模型性能越好^[15]; RMSEC (root mean square errors of calibration)、RMSECV (root mean square errors of cross validation)、RMSEP (root mean square error of prediction) 分别表示校正均方根误差、交叉验证均方根误差和预测均方根误差, R_{cal} 和 R_{pre} 分别表示校正集和验证集的相关系数, $BIAS_{cal}$ 和 $BIAS_{pre}$ 分别表示校正的偏差和预测的偏差, RMSEC、RMSECV、RMSEP、 $BIAS_{cal}$ 和 $BIAS_{pre}$ 越小, R_{cal} 和 R_{pre} 越大, 表示模型校正和预测性能越好。移动窗口平滑优于其他预处理方法, RMSEC、RMSECV、RMSEP 分别为 0.223%、0.262%、0.229%, RPD 为 11.377, RSEP 为 3.036%。

过程湍流是干扰流化床干燥过程在线近红外光谱检测准确度的原因之一。探头周围物料密度的改

变以及颗粒粒径变化, 导致在线光谱的整体强度的强烈随机波动, 这通常被称为“散射效应”。3 种最常用的散射校正方法是 MSC、SNV 和求导法。在流化床干燥过程在线漫反射近红外监测的文献报道中, 数据分析以 MSC、SNV 或求导法为主^[16-17]。而本研究发现, 经过移动窗口平均算法对光谱数据进行预处理之后, 建立的水分含量 PLS 定量模型效能明显优于经 MSC 或 SNV 处理后的模型, 即传统的散射校正方法可能会消除一些与水分相关的有用变量信息。该结果与 Bogomolov 等^[18]提出的观点类似, 同时, Bogomolov 等^[18]为了提高流化床干燥过程中颗粒水分在线监测精确度, 提出一种基于时间维度的光谱平滑算法, 该算法保留了更多与颗粒含水量相关的变量。Chen 等^[19]也研究了散射校正数据包含的信息的性质, 认为对于包含可提取信息的散射校正参数而言, 散射校正不一定是最佳的预处理

方法。周宏^[20]也发现在应用 NIRS 系统获取斜生栅藻的透射光谱信息时, 基于经移动窗口平滑预处理之后的光谱油脂含量建立定量模型效果较好。

3.4.2 平滑窗口宽度的选择 窗口宽度(k)的选择是一个值得探讨的问题, k 太小, 平滑去噪效果不理想, k 太大, 会导致光谱失真。建立了不同 k (3~21 点) 下的 PLS 定量模型。 k 与 RPD、RSEP 的关系如图 4 所示, 图 4-a、b 中均出现 2 个明显的拐点, 此时均有较大的 RPD 和较小的 RSEP。当 k 为 5 时, RPD=12.129, RSEP=2.873%, 当 k 为 15 时, RPD=12.293, RSEP=2.833%, 即当采用 15 点平滑时, 模型预测性能最佳, 此时模型性能参数见表 5。180808 批次 7 个 GFC 流化床干燥过程样本的原

始光谱图和经移动窗口平滑预处理之后的光谱图见图 5-a、b。随着干燥的进行, $\lg(1/R)$ 逐渐减小, 与颗粒含水量减小趋势一致, 如图 5-a 中箭头 1 所示方向。同时, 随着干燥过程中物料温度升高, 整体光谱有向高波数方向移动的趋势, 尤其是 5 150 cm^{-1} 处的吸收峰位移极为明显, 如图 5-a 中箭头 2 所示方向。

3.5 特征变量筛选

筛选特征波段可以剔除无关变量和非线性变量, 简化模型, 提高模型预测精确度和稳健性。本研究考察了 3 种不同变量筛选方法及其组合应用时对模型的影响。

3.5.1 间隔偏最小二乘法 (interval PLS, iPLS)^[21]

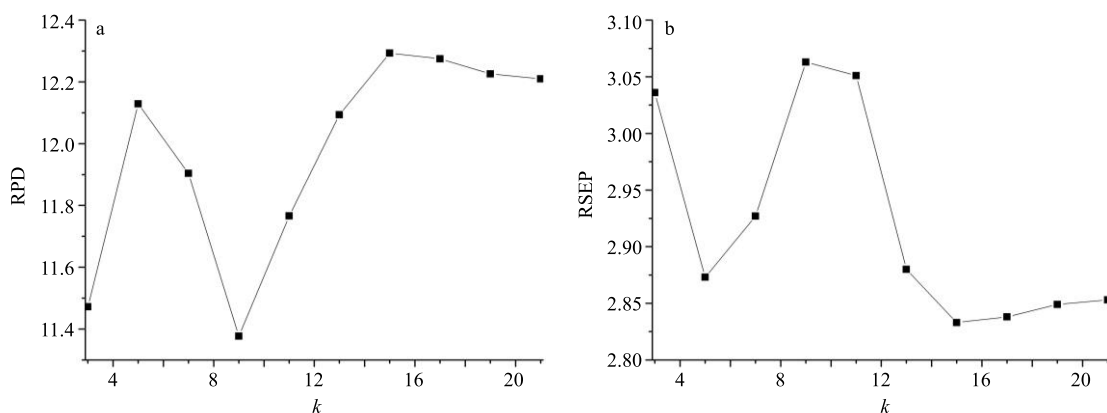


图 4 RPD (a) 和 RSEP (b) 对窗口宽度的依赖性关系

Fig. 4 RPD (a) and RSEP (b) dependence on window size

表 5 移动窗口平滑后水分含量定量模型

Table 5 PLS Model developed after window moving average something

预处理方法	校正集				预测集				
	R_{cal}	RMSEC/%	RMSECV/%	$BIAS_{cal}$	R_{pre}	RMSEP/%	$BIAS_{pre}$	RPD	RSEP/%
移动窗口平滑	0.997	0.224	0.260	0.171	0.997	0.212	0.171	12.293	2.833

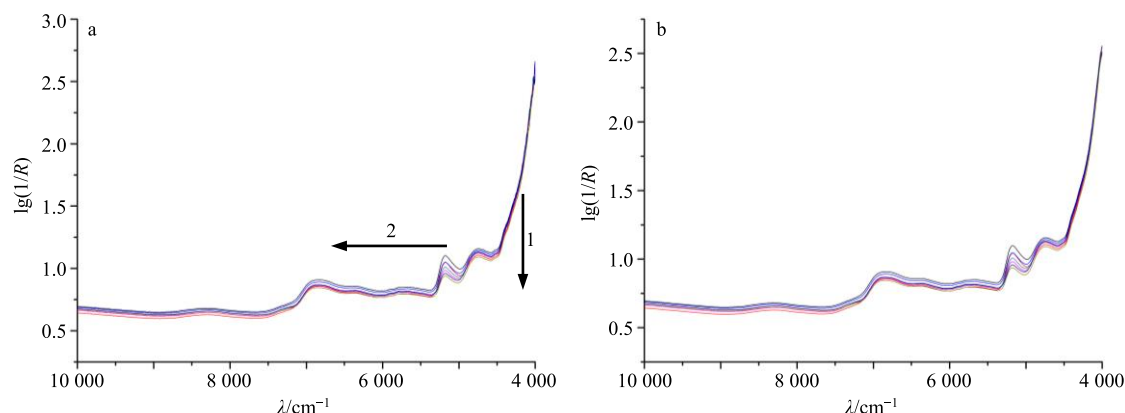


图 5 180808 批次样品干燥过程原始 NIRS 图 (a) 和经移动窗口平滑 (15 点) 之后的 NIRS 图 (b)

Fig. 5 Raw NIRS (a) and spectra preprocessed by window moving average algorithm with 15 points (b) of samples collected from 180808 batch

iPLS 是把光谱数据等分成多个子区间,并在每个子区间进行 PLS 回归,建立多个局部回归模型。本研究将全波段光谱划分为 20 个子区间,以 RMSECV 为评价指标,筛选最佳建模区间。

3.5.2 组合间隔偏最小二乘法(synergy interval PLS, siPLS) siPLS 是间隔偏最小二乘法的优化。siPLS 将数据集划分为多个等距间隔,通过不同区间数的任意组合(组合数一般为 2、3 或 4)建立 PLS 模型而筛选出最优的组合区间。本研究采用组合间隔偏最小二乘法筛选特征波段,将全波段光谱划分为 20 个子区间,以 RMSECV 为评价指标,筛选最佳建

模区间。

3.5.3 移动窗口偏最小二乘法(moving window PLS, mwPLS) [22] mwPLS 是根据移动窗口概念计算 PLS 模型,基本原理是将 1 个窗口沿着光谱轴连续移动,每移动 1 个波长点,采用交互验证方式确定最佳因子数并建立相应 PLS 最优模型。本研究考察窗口宽度了对 RMSECV 的影响(设计的窗口宽度分别 77、135、231、309、385、463 点),然后以 RMSECV 为评价指标,筛选最佳建模区间。

3.5.4 筛选特征变量区间 以不同区间变量为自变量,建立水分含量 PLS 模型,结果见表 6。仅使用

表 6 以不同变量建立 PLS 定量模型结果

Table 6 PLS regression statistics with different variables

变量筛选方法	变量区间/cm ⁻¹	RMSEC/%	RMSEP/%	R _{cal}	R _{pre}	RMSECV/%	RPD
无	3 999.64~10 001.03	0.224	0.212	0.9973	0.997	0.260	12.293
siPLS	4 902.16~5 199.15	0.226	0.198	0.997	0.997	0.245	13.198
	5 503.84~5 800.83						
	5 804.69~6 101.67						
	8 512.25~8 809.24						
iPLS	5 272.43~5 499.99	0.287	0.279	0.995	0.995	0.377	9.349
mwPLS	4 759.45~5 338.00	0.344	0.315	0.994	0.993	0.357	8.287
siPLS+iPLS	4 902.16~6 101.67	0.216	0.210	0.997	0.997	0.243	12.428
	8 512.25~8 809.24						
siPLS+mwPLS	4 759.45~5 338.00	0.226	0.195	0.997	0.997	0.243	13.384
	5 503.84~6 101.67						
	8 512.25~8 809.24						
iPLS+mwPLS	4 759.45~5 499.99	0.299	0.304	0.994	0.994	0.353	8.574

iPLS、mwPLS 或者联合 iPLS 和 mwPLS 时,模型性能低于全谱建模,提示 iPLS、mwPLS 在本研究中处于弱势地位。使用 siPLS 联合 mwPLS 筛选的组合变量 4 759.45~5 338.00 cm⁻¹、5 503.84~6 101.67 cm⁻¹ 和 8 512.25~8 809.24 cm⁻¹ 建模,结果显示优于其他方法单独使用。5 150 cm⁻¹ 处的强吸收峰是水分子反对称伸缩振动和弯曲震动的组合频谱带,包含于第 1 区间;水在 5 620 cm⁻¹ 和 8 310 cm⁻¹ 附近有 1 个宽的比较弱的组合频吸收,也大体包含于第 2、3 区间,第 3 区间变量波数高于 8 310 cm⁻¹,可能是由于温度升高导致了吸收峰发生位移,即该变量组合包含水分子吸收峰,同时消除了温度对校正模型的影响[23],故模型预测效果较好。

3.6 建立水分含量定量模型

以移动窗口平滑法对原始光谱进行预处理,采

用 siPLS 联合 mwPLS 筛选特征变量,建立 GFC 流化床干燥过程中间颗粒水分含量 PLS 模型。合适的潜变量数可以缩减建模子空间。潜变量过多时,会引入噪声和不相关部分,导致过拟合;潜变量过少又会导致欠拟合,由于模型缺少了部分相关信号,致使校正和预测阶段都无法保证模型准确度。最常用的检验技术是交互验证法。本实验采用留一交叉验证法进行优化,以预测误差平方和(predicted residual error sum square, PRESS)和 RMSECV 为指标,确定最优潜变量数。结果如图 6 所示,潜变量为 8 时, RMSECV 最小, PRSEE 趋于稳定。PLS 模型性能参数如表 3 所示,与无预处理条件下建立的 PLS 模型相比, RPD 提高至 13.384, RESP 降低至 2.602%,校正集和验证集参考值和预测值相关关系如图 7 所示,校正集回归曲线 $Y=0.9952X+$

0.038 9, $r^2=0.994\ 4$, 验证集回归曲线 $Y=0.969\ 2X+0.185\ 3$, $r^2=0.995\ 1$, 模型性能明显提高。

3.7 生产持续验证

取 190407、190408、190409、190410、190501、10502、190503、190504 批次流化床干燥过程颗粒作为外部验证样本, 考察模型稳定性, 参考值和预

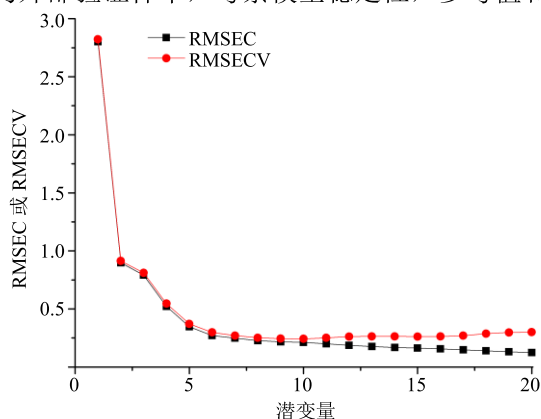


图 6 RMSEC、RMSECV 与潜变量数的关系

Fig. 6 Relationship between RMSEC, RMSECV, and number of latent variables

测值如表 7 所示, 相对预测误差 (相对预测误差 = $|\text{预测值} - \text{参考值}| / \text{参考值}$), 均小于 4.62%, 表明经过生产实际验证, 预测结果准确, 该模型较为准确和稳健。

3.8 实时在线监测

将建立好的 PLS 模型载入控制系统, 对新 1 批

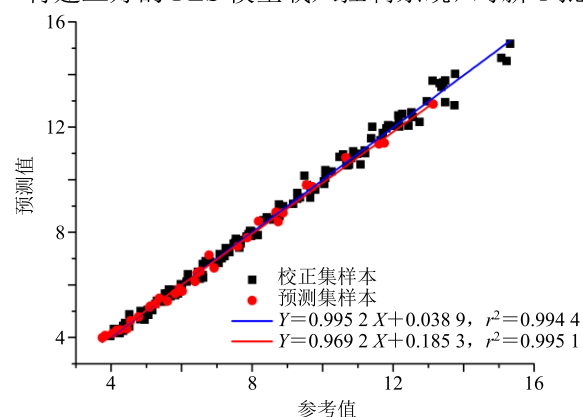


图 7 预测值与参考值相关关系

Fig. 7 Correlation between predicted values and experimental values

表 7 外部验证样本参考值和预测值结果

Table 7 Experimental values and predicted values for external validation samples

批次	参考值	预测值	相对预测 误差/%	批次	参考值	预测值	相对预测 误差/%	批次	参考值	预测值	相对预测 误差/%
190407	13.96	14.12	1.15	190411	12.24	12.26	0.16	190503	9.28	9.69	4.42
	10.94	10.97	0.27		8.81	9.12	3.52		7.26	7.34	1.10
	9.01	9.27	2.89		7.12	7.32	2.81		6.18	6.17	0.16
	6.35	6.36	0.16		6.08	6.21	2.14		4.89	4.84	1.02
	4.24	4.11	3.07		4.86	4.89	0.62		4.93	4.73	4.06
190408	11.83	12.01	1.52	190412	12.12	11.86	2.15	190504	10.84	11.14	2.77
	8.38	8.44	0.72		9.52	9.74	2.31		7.20	7.52	4.44
	7.02	7.33	4.42		7.72	7.78	0.78		4.86	5.09	4.73
	5.12	5.11	0.20		6.30	6.37	1.11		4.56	4.42	3.07
	4.04	3.87	4.21		4.49	4.40	2.00		4.07	3.94	3.19
190409	12.34	12.91	4.62	190501	8.98	8.92	0.67	190505	11.60	11.79	1.64
	8.47	8.62	1.77		7.75	8.09	4.39		9.78	9.89	1.12
	6.80	6.81	0.15		5.12	5.07	0.98		7.44	7.65	2.82
	5.51	5.32	3.45		4.76	4.81	1.05		6.23	6.16	1.12
	4.74	4.65	1.90		4.66	4.88	4.72		4.62	4.57	1.08
190410	8.52	8.66	1.64	190502	10.11	10.39	2.77	190506	11.40	11.54	1.23
	6.48	6.59	1.70		8.07	8.33	3.22		7.56	7.89	4.37
	5.10	5.12	0.39		6.35	6.56	3.31		5.59	5.83	4.29
	4.54	4.48	1.32		5.50	5.41	1.64		4.64	4.57	1.51
	4.27	4.29	0.47		4.66	4.52	3.00		4.26	4.19	1.64

次(批号 181122)GFC 干燥过程进行实时在线监测。探头会及时扫描到光谱信息并经光纤传递给控制系统,这时显示器就会迅速给出当前水分含量值,当显示值达到操作规程要求时,停止干燥。图 8 是流化床干燥过程中 PLS 模型在线 NIR 预测值与测定值的趋势对比图。预测值和参考值重叠性良好,相关系数 $r^2=0.9956$, 平均相对预测误差为 1.98%。终点样本颗粒水分含量为 4.27%,在控制限(3%~5%)范围内,符合生产规定。

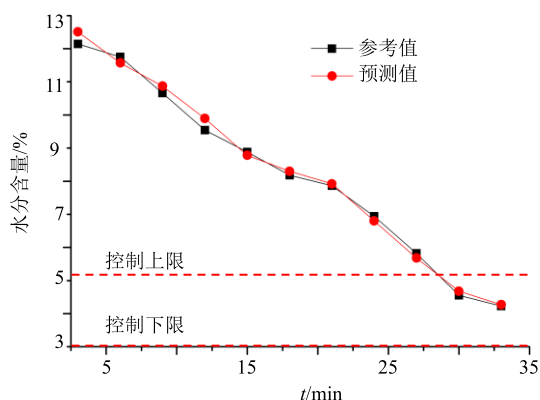


图 8 流化床干燥过程中间水分含量预测值和参考值趋势
Fig. 8 Comparison of predicted values with experimental values of process of fluidized bed drying process

4 讨论

中药原料不均一、生产过程缺乏在线控制是导致中药批间差异的重要因素,而过程分析技术在中药过程质量控制中的应用可显著改善批间差异。上述结果证明了应用 NIRS 技术实时监测流化床干燥过程颗粒水分含量的可行性。同时,与在线红外水分检测仪、在线微波水分检测仪等在线水分检测工具相比,近红外具有绿色无损、可快速同时检测多个指标等优点,可进一步应用于 GFC 流化床干燥过程中化学性质(如有效成分含量)和物理性质(如密度、粒径等)实时监测,从而提高批间一致性和产品稳定性。

根据上述结果,提出以下 3 点关于流化床干燥和类似工艺类型中水分含量的在线 NIRS 监测的建议:(1)准确的模型应建立在大量生产规模一级数据的基础上,要通过使用足够多的代表性样品和工艺条件来研究光谱数据与水分相关性。考虑到原料及工艺波动等造成的批间不均一性,模型校正集数据尽可能包含不同生产批次下的样本,如不同操作组、不同生产线、不同批次原料等。(2)建议在进行光谱预处理之前分析原始光谱数据与样本的联

系,不鼓励在多变量校准之前对在线过程光谱进行散射校正,因为它会导致原始光谱数据中与水分有关的信息丢失,使基于此建立的模型预测效果欠佳。(3)生产应用中,应定期进行模型外部验证,必要时进行模型更新。

参考文献

- [1] Faulhammer E, Llusa M, Radeke C, *et al.* The effects of material attributes on capsule fill weight and weight variability in dosator nozzle machines [J]. *Int J Pharm*, 2014, 471(1/2): 332-338.
- [2] Peinado A, Hammond J, Scott A. Development validation and transfer of a near infrared method to determine in-line the end point of a fluidised drying process for commercial production batches of an approved oral solid dose pharmaceutical product [J]. *J Pharm Biomed Anal*, 2011, 54(1): 13-20.
- [3] Wildfong P L, Samy A S, Corfa J, *et al.* Accelerated fluid bed drying using NIR monitoring and phenomenological modeling: Method assessment and formulation suitability [J]. *J Pharm Sci*, 2002, 91(3): 631-639.
- [4] 刘雪松,李梦茹,王致远,等.基于近红外光谱的驴胶补血颗粒浓缩过程研究 [J]. *中草药*, 2016, 47(22): 3997-4002.
- [5] 杜晨朝,赵安邦,吴志生,等.近红外光谱结合不同变量筛选方法用于金银花提取过程中绿原酸量的在线监测 [J]. *中草药*, 2017, 48(16): 3317-3321.
- [6] 吴莎,刘启安,吴建雄,等.统计过程控制结合近红外光谱在栀子中间体纯化工艺过程批放行中的应用研究 [J]. *中草药*, 2015, 46(14): 2062-2069.
- [7] 杨越,王磊,刘雪松,等.近红外光谱结合多变量统计过程控制(MSPC)技术在金银花提取过程在线实时监控中的应用研究 [J]. *中草药*, 2017, 48(17): 3497-3504.
- [8] Otsuka M, Koyama A, Hattori Y. Real-time release monitoring for water content and mean particle size of granules in lab-sized fluid-bed granulator by near-infrared spectroscopy [J]. *RSC Adv*, 2014, 4(34): 17461-17468.
- [9] 张广仁,吴云,孙仙玲,等.近红外光谱分析技术在散结镇痛胶囊干燥过程水分检测中的应用 [J]. *世界科学技术—中医药现代化*, 2016, 18(2): 313-317.
- [10] 易夏,冯艳春,胡昌勤,等.近红外光谱法探索在线监测注射用五水头孢唑林钠干燥过程中的水分含量 [J]. *沈阳药科大学学报*, 2017, 34(7): 558-564.
- [11] 梁逸曾,俞汝勤.化学计量学[M].北京:高等教育出版社,2003.
- [12] 倪力军,张立国.基础化学计量学及其应用[M].上

- 海: 华东理工大学出版社, 2013.
- [13] 徐冰, 史新元, 乔延江, 等. 金银花醇沉多阶段多变量统计过程控制研究 [J]. 中华中医药志, 2012, 27(4): 784-788.
- [14] 李华, 王菊香, 邢志娜, 等. 改进的 K/S 算法对近红外光谱模型传递影响的研究 [J]. 光谱学与光谱分析, 2011, 31(2): 362-365.
- [15] 王永香, 郑伟然, 米慧娟, 等. 热毒宁注射液青蒿金银花浓缩过程近红外快速定量检测方法的建立 [J]. 中草药, 2017, 48(1): 102-108.
- [16] Märk J, Karner M, Andre M, *et al.* Online process control of a pharmaceutical intermediate in a fluidized-bed drier environment using near-infrared spectroscopy [J]. *Anal Chem*, 2010, 82(10): 4209-4215.
- [17] Kona R, Qu H, Mattes R, *et al.* Application of in-line near infrared spectroscopy and multivariate batch modeling for process monitoring in fluid bed granulation [J]. *Int J Pharm*, 2013, 452(1/2): 63-72.
- [18] Bogomolov A, Mannhardt J, Heinzerling O. Accuracy improvement of in-line near-infrared spectroscopic moisture monitoring in a fluidized bed drying process [J]. *Front Chem*, 2017, 6(388): 1-10.
- [19] Chen Y C, Thennadil S N. Insights into information contained in multiplicative scatter correction parameters and the potential for estimating particle size from these parameters [J]. *Anal Chem Acta*, 2012, 746(2012): 37-46.
- [20] 周宏. 基于光谱技术的斜生栅藻油脂含量检测及其空间分布的研究 [D]. 杭州: 浙江大学, 2018.
- [21] Zou X B, Zhao J W, Povey M J W, *et al.* Variables selection methods in near-infrared spectroscopy [J]. *Anal Chem Acta*, 2010, 667(1/2): 14-32.
- [22] 史智佳, 李鹏飞, 吕玉, 等. 移动窗口偏最小二乘法优选猪油丙二醛近红外光谱波段 [J]. 中国食品学报, 2014, 14(11): 207-213.
- [23] Woekman J, Weyer J L. 近红外光谱解析实用指南 [M]. 北京: 化学工业出版社, 2009.