

# 制浆材木质素含量近红外分析模型传递研究

刘耀瑶<sup>1</sup> 杨浩<sup>1</sup> 熊智新<sup>1,\*</sup> 梁龙<sup>2</sup> 房桂干<sup>2</sup>

(1. 南京林业大学江苏省制浆造纸科学与技术重点实验室, 江苏南京, 210037;

2. 中国林业科学研究院林产化学工业研究所, 江苏南京, 210042)

**摘要:** 为研究制浆材中木质素含量近红外分析模型在两台便捷式近红外光谱仪间的传递, 对制浆材木质素样品近红外光谱数据集进行代表性样本的选取、光谱预处理和界外样本的剔除, 建立了源机的优化偏最小二乘(PLS)校正模型。分别采用斜率截距算法(S/B)、直接校正算法(DS)和典型相关分析算法(CCA)进行源机与目标机间的模型传递并比较了预测效果。结果表明, S/B算法模型传递效果较差, 而经DS算法和CCA算法模型传递后的预测效果均有大幅提升。DS算法模型传递后决定系数( $R^2$ )、预测标准差(RMSEP)和相对标准差(RPD)分别为0.9643、1.0370%和5.3513; CCA算法模型传递后 $R^2$ 为0.9540、RMSEP为1.1766%、RPD为4.7711。因此, DS算法和CCA算法均可实现制浆材木质素含量近红外分析模型在两台便携式近红外光谱仪之间的传递。

**关键词:** 近红外光谱; 模型传递; 木质素; 制浆材

**中图分类号:** TS7; O657.3

**文献标识码:** A

**DOI:** 10.11981/j.issn.1000-6842.2019.03.43

制浆材材性的差异会直接影响制浆过程的工艺参数及纸浆质量; 其中, 木质素的含量直接决定着漂白剂的用量<sup>[1]</sup>。因此, 在实际生产过程中, 实现制浆材中木质素含量的在线、快速检测对于及时调整制浆工艺和优化工艺参数具有重要意义<sup>[2]</sup>; 但传统的分析方法耗时且步骤繁琐。目前, 多数企业采用提高制浆过程中化学品用量的方法以尽可能地脱除木质素, 从而保证纸浆质量, 这在很大程度上导致了化学品的浪费和环境污染问题, 进而导致生产成本的增加<sup>[3]</sup>。近红外光谱(Near-infrared Spectroscopy, NIRS)分析技术作为一种快速分析手段, 具有分析简便、分析速度快、无损、操作技术要求低等优势, 已在农业、石油、化工、食品、制药等行业被广泛使用<sup>[4-5]</sup>。但在实际生产应用中, 近红外光谱受测量仪器或测量条件的影响较大<sup>[6]</sup>, 经常出现已建好的模型无法在另一台仪器或另一种条件下适用的情况, 而建立一个能满足实际应用要求的近红外校正模型需要花费大量的人力和物力; 因此, 利用合适的模型传递技术实现仪器间模型的共享和有效利用非常必要<sup>[7]</sup>。

模型传递的本质是克服样品在不同仪器上的测量信号(或光谱)间的不一致性<sup>[4]</sup>。按照是否需要在所有仪器上采集一一对应的标准光谱, 模型传递可以划分为有标样和无标样两种。大部分模型传递算法为有

标样算法<sup>[6]</sup>, 这类算法必须取一定数量的样品组成标样集, 并分别在源机(Master)和目标机(Slaves)上测得所取标样集的光谱, 从而找出两台仪器间的数学传递关系, 如斜率截距算法(Slope/Bias, S/B)<sup>[8]</sup>、直接校正算法(Direct Standardization, DS)<sup>[9]</sup>、典型相关分析算法(Canonical Correlation Analysis, CCA)<sup>[10]</sup>、分段直接校正算法(Piecewise Direct Standardization, PDS)<sup>[11]</sup>和Shenk's算法(Shenk's Algorithm)<sup>[12]</sup>等; 另一类是无标样方法, 如小波变换(Wavelet Transform, WT)<sup>[13]</sup>、有限脉冲响应(Finite Impulse Response, FIR)<sup>[4]</sup>方法等, 使用这类算法不需要任何标准样品。近年来, 国内外学者在模型传递方面做了很多相关研究。信晓伟等<sup>[14]</sup>在传统S/B算法的基础上, 提出引入变量高次幂、使用Lagrange插值法与Newton插值法求待定数和插值多项式的新方法, 实现了烟草中总糖、总氮、还原糖、总烟碱含量模型在两台仪器间的传递。Eskildsen等<sup>[15]</sup>通过改变模型传递评估标准, 实现了面粉中蛋白质含量模型在多台仪器间更准确地传递。目前, 多数模型传递研究主要集中在大型实验室近红外仪器方面, 此类仪器的性能指标(如光谱范围、分辨率、信噪比、自身稳定性和仪器间的一致性)相对较高<sup>[16]</sup>。然而, 近年来随着仪器的小型化发展, 近红外技术也从实验室逐步走向了各行各业的实际生产应用中, 各种性能适中但

收稿日期: 2018-02-05

基金项目: 国家林业局948项目“农林剩余物制机械浆节能和减量技术引进”(2014-4-31)。

作者简介: 刘耀瑶, 女, 1994年生; 在读硕士研究生; 主要研究方向: 制浆造纸过程控制与信息智能处理。

\* 通信联系人: 熊智新, 副教授; 主要研究方向: 制浆造纸过程控制与信息智能处理; E-mail: leo\_xzx@njfu.edu.cn。

价格相对较为便宜的便携式近红外光谱仪被广泛应用；因此，实现这些便携式近红外光谱仪之间的模型传递也成为了一个亟待解决的问题。

由于 PDS 算法和 Shenk's 算法在进行模型传递的过程中，需预先确定合适的窗口宽度、建立多个多元回归模型、计算量大且不利于便捷式近红外光谱仪间模型的传递。故本研究以 5 种常用制浆材（马尾松、杉木、相思木、桉木、杨木）为对象，分别采用 S/B、DS 以及 CCA 的模型传递算法，研究了制浆材中木质素含量近红外光谱分析模型在两台便捷式近红外光谱仪间的传递，比较了 3 种算法的模型传递效果。

## 1 模型传递原理及算法

模型传递是通过数学方法建立源机 (Master) 和目标机 (Slaves) 所测光谱之间的函数关系，由确定的函数关系对光谱进行转换来实现模型的通用性<sup>[16]</sup>。

### 1.1 S/B 算法

S/B 算法是对模型预测结果的标准化校正方法<sup>[4]</sup>。用源机上建立的校正模型分别预测标样集在源机、目标机上测得的光谱矩阵  $X_m$ 、 $X_s$ ，得到光谱的预测值  $y_m$ 、 $y_s$ 。假设  $y_m$  和  $y_s$  存在以下关系：

$$y_s = \text{Slope} \cdot y_m + \text{Bias} \quad (1)$$

其中，Slope 为线性方程斜率；Bias 为线性方程截距；可通过最小二乘法 (Least Square, LS) 计算求得。对于在目标机上测得的未知样品光谱矩阵  $X_{s,un}$ ；首先，用源机上建立的校正模型预测得到  $y_{s,un}$ ，然后，用下式便可求得传递后的预测值  $y_{s,tr}$ ：

$$y_{s,tr} = \text{Slope} \cdot y_{s,un} + \text{Bias} \quad (2)$$

其中， $y_{s,un}$  为未知样品的预测值； $y_{s,tr}$  为未知样品传递后的预测值。

### 1.2 DS 算法

DS 算法是通过转换矩阵  $F_{DS}$  将目标机测得的未知样品光谱矩阵  $X_{s,un}$  转换为未知样品传递后的光谱矩阵  $X_{s,tr}$ 。理论上， $X_{s,tr}$  与源机上测得的未知样品光谱矩阵  $X_{m,un}$  是相等的。首先，分别测得源机和目标机的标样集光谱矩阵  $X_m$  和  $X_s$ ，则转换矩阵  $F_{DS}$  可通过 LS 计算得到：

$$X_m = X_s \cdot F_{DS} \quad (3)$$

$$F_{DS} = X_s^+ \cdot X_m \quad (4)$$

式中， $X_m$  为源机标样集光谱矩阵； $X_s$  为目标机标样集光谱矩阵； $X_s^+$  为  $X_s$  的广义逆矩阵， $F_{DS}$  为  $m \cdot m$  维的转换矩阵 ( $m$  为波长点数)。

对于在目标机上测得的未知样品光谱矩阵  $X_{s,un}$ ，

根据下式：

$$X_{s,tr} = X_{s,un} \cdot F_{DS} \quad (5)$$

计算传递后的  $X_{s,tr}$ ，再将  $X_{s,tr}$  代入源机建立的校正模型中计算出光谱最终的预测值。

### 1.3 CCA 算法

CCA 算法是一种利用综合变量间的相关关系来反映两组指标间整体相关性的多元统计分析方法。在运用 CCA 算法进行模型传递时，首先对测得的源机和目标机的标样集光谱矩阵  $X_m$  和  $X_s$  进行典型相关分析：

$$L_m = X_m \cdot W_m \quad (6)$$

$$L_s = X_s \cdot W_s \quad (7)$$

式中， $W_m$  和  $W_s$  分别是  $X_m$  和  $X_s$  的典型向量， $L_m$  和  $L_s$  分别是  $X_m$  和  $X_s$  的典型变量。

则转换矩阵  $F_{CCA}$  可通过 LS 计算得到：

$$F_{CCA,1} = L_s^+ \cdot L_m \quad (8)$$

$$F_{CCA,2} = L_m^+ \cdot X_m \quad (9)$$

$$F_{CCA} = W_s \cdot F_{CCA,1} \cdot F_{CCA,2} \quad (10)$$

式中， $L_s^+$  和  $L_m^+$  分别为  $L_s$  和  $L_m$  的广义逆矩阵。

最后，采用式 (10) 得到的转换矩阵  $F_{CCA}$  对目标机上测得的未知光谱矩阵  $X_{s,un}$  进行转换：

$$X_{s,tr} = X_{s,un} \cdot F_{CCA} \quad (11)$$

式中， $X_{s,tr}$  是传递后的未知样品光谱矩阵；将计算得到的  $X_{s,tr}$  带入源机模型便可直接得到光谱最终的预测值，从而实现基于典型相关分析的近红外光谱传递。

### 1.4 评价指标

模型传递效果和模型预测能力是由样品预测值和标准值间的决定系数 ( $R^2$ )、预测标准差 (RMSEP) 和相对标准偏差 (RPD) 3 个指标进行综合评价<sup>[16]</sup>。其中， $R^2$  越接近 1、RMSEP 越小，则表明模型的传递效果越好；RPD 用来验证模型的稳定性和预测能力，当 RPD > 3 时，表明模型具有较高的稳定性和良好的预测能力<sup>[17]</sup>。

## 2 实验

### 2.1 实验仪器

实验采用无锡迅杰光远科技有限公司开发的 2 台同批次生产的光栅扫描型便携式近红外光谱仪 IAS-1000，分别标记为 1 号机 (源机) 和 2 号机 (目标机)。仪器扫描范围为 900 ~ 1700 nm，分辨率为 10 nm，扫描次数为 20 次。

### 2.2 制浆材样品制备与光谱采集

实验样品为中国林业科学研究院林产化学工业研

究所提供的5种常用制浆材,其中有2种针叶木制浆材(马尾松、杉木)和3种阔叶木制浆材(相思木、桉木、杨木),共计87个样品,各种制浆材样品数见表1。将原木去皮、切削、粉碎后,选取40~60目的木粉分别在源机和目标机上采集光谱。采集样品光谱时,将样品放入测量杯中,以50g砝码压平使其均匀分布,每个样品重复装样3次取平均光谱。每扫描完一个样品,用毛刷将样品杯中的残留木粉去除,以免影响后续样品光谱采集的准确性。采集得到源机和目标机的近红外光谱如图1所示。由图1可知,目标机测得的各木粉样品近红外光谱的吸光度明显大于源机测得的数值,即相同样品在同批次的2台近红外光谱仪上测得的近红外光谱有明显区别。

表1 5种制浆材样品木质素含量分布

制浆材品种	样品数/个	最大值/%	最小值/%	平均值/%
马尾松	10	28.95	28.47	28.63
杉木	23	34.20	32.55	33.43
相思木	12	27.15	24.62	25.69
桉木	26	27.56	21.49	23.64
杨木	16	20.50	14.81	18.14
总计	87	34.20	14.82	26.08

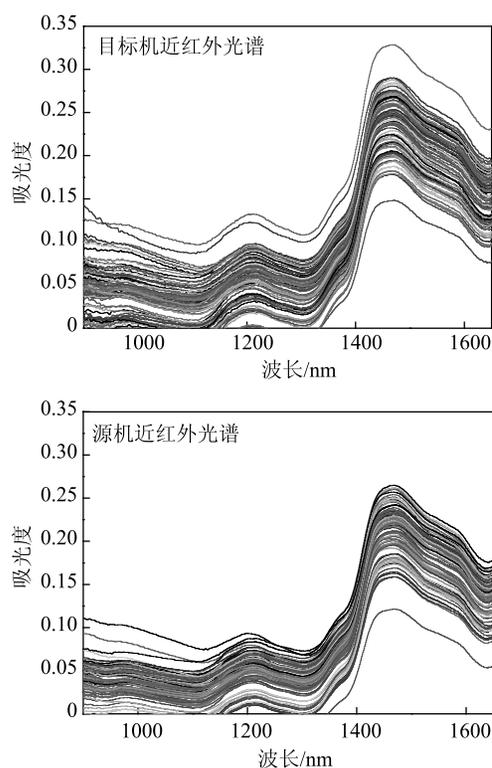


图1 木粉样品在源机和目标机采集的近红外光谱

## 2.3 木质素标准值测定

按照 GB/T 2677.8—1994, 采用硫酸法对制浆材

样品木质素进行测定。木粉样品经苯-醇混合液抽提后,用质量分数为(72±0.1)%的硫酸进行水解,然后依次测得所有制浆材样品水解残余物(即制浆材的木质素)的质量,测量结果如表1所示。由表1可知,87个制浆材样品木质素的含量范围为14.82%~34.20%,分布范围较广,表明选取的87个样品具有良好的代表性。

## 2.4 数据分析

本研究采用 NIRSA 4.5 系统以及 Matlab7.0 软件平台进行数据分析。NIRSA 4.5 系统是本实验室自主研发的专门用于近红外光谱数据处理的化学计量学软件(计算机软件著作权登记号为2007SR06801),主要用于近红外光谱预处理、制浆材木质素近红外光谱分析模型的建立以及转换集标样数的选取。Matlab7.0 平台则主要用于模型传递算法的程序编写和数据处理。

## 3 结果与分析

### 3.1 定量校正模型的建立

#### 3.1.1 建模集样本的选取与光谱预处理

用于建立模型的样品是否具有较强代表性对于所建模型的可靠性具有重要的影响。目前,常用的建模集样本选取方法有 Kennard-Stone (K-S)、Duplex 等。为确保建立的分析模型可以实现多种制浆材木质素含量的分析检测并具有较广的适应性;本研究在保证建模集中必须含有5种制浆材样品的基础上,采用 K-S 算法<sup>[17]</sup>选择了65个样品作为建模集,剩余22个样品作为预测集,样品集划分情况如表2所示。

表2 5种制浆材样品使用情况

制浆材品种	样品/个	
	建模集	预测集
杉木	17	6
马尾松	8	2
相思木	9	3
桉木	19	7
杨木	12	4
总计	65	22

近红外光谱仪所采集的光谱除样品的自身信息外,还包含了其他无关信息和噪声。因此,在用化学计量学方法建模时,消除光谱数据无关信息和噪声的预处理方法变得十分关键和必要<sup>[17]</sup>。本研究分别采用多元散射校正(Multiplication Scatter Correction, MSC)、一阶微分、二阶微分、Savitzky-Golay (S-G) 卷

积平滑和标准正态变量变换(Standard Normal Variate Correction, SNV)等不同的方法对标样集光谱进行预处理,经过多次处理对比发现,采用 MSC 进行预处理时,最终建模效果最好。由图 1 已知,同台仪器不同样品的光谱吸光度在纵轴上存在较大的平移,而 MSC 可以消除测量杯中因木粉样品颗粒形状各异、大小不同及分布不均匀等因素导致的散射系数差异对其漫反射光谱的影响<sup>[18]</sup>;因此,经 MSC 预处理后,标样集光谱的最终建模效果最佳。

### 3.1.2 建立制浆材木质素含量近红外光谱分析偏最小二乘(PLS)模型

利用 PLS 进行近红外光谱定量模型的建立,是目前近红外光谱分析中应用最广泛的一种方法<sup>[17,19]</sup>。因此,本研究利用 PLS 方法在源机上建立制浆材木质素近红外光谱定量校正模型。

在 PLS 建模中,随着主成分数增多,载荷向量对建模的重要程度逐渐减小,到一定程度后,载荷向量将变成模型的噪声<sup>[17]</sup>;因此,选择合适的主成分数尤为重要。本实验选定最大主成分数为 15,并采用留一法交叉验证以选取预测残差平方和(PRESS)最小的主成分数作为最终的主成分数。根据文献[16],在交互验证的过程中,如果预测值和参考方法测定值之间的偏差大于交互验证标准偏差(SECV)的 2 倍,则该样品被视为参考值异常样品,应予以剔除。本实验中,经计算检验,有 2 个桉木样品作为异常样品被剔除。因此,最终选定 63 个样本作为建模集(标样集),进行 PLS 建模。建模集 PRESS 随主成分数的变化如图 2 所示。由图 2 可知,当主成分数为 10 时, PRESS 达到最小值。故本实验中,选取 10 个主成分数以建立制浆材木质素近红外光谱定量校正模型。

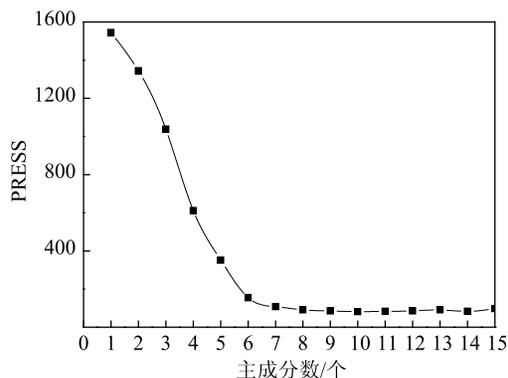


图 2 建模集 PRESS 随主成分数的变化

## 3.2 模型传递前预测结果及分析

利用源机所建制浆材木质素近红外光谱定量校正

模型分别对源机和目标机预测集的 22 个样品进行预测分析,结果见表 3。由表 3 可知,源机建立的制浆材木质素近红外光谱定量校正模型用于源机样品预测时,具有良好的预测精度, $R^2$ 、RMSEP 和 RPD 分别为 0.9660、1.0116% 和 5.5493,基本可以满足实际生产的要求。而利用该模型对目标机预测集样品进行预测时, $R^2$ 、RMSEP 和 RPD 等评价指标均变差,分别为 0.8614、2.0415% 和 2.7498,表明该模型不能直接用于目标机样品的预测。

表 3 源机校正模型对源机和目标机预测集的预测结果

仪器	RMSEP/%	$R^2$	RPD
源机	1.0116	0.9660	5.5493
目标机	2.0415	0.8614	2.7498

通常,导致源机校正模型在目标机上“失效”的原因主要有样品物理或化学性质变化、仪器差异及测量环境的影响<sup>[18]</sup>。在本实验中,所有木粉样品在实验室中两台相邻放置的仪器上进行平行测量;因此,样品化学性质和所处环境对光谱测量几乎无影响。光谱的不一致主要体现在同一样品在两台仪器上进行平行测量过程中,两次装填时样品表面的平整度、木粉颗粒分布的均匀性等很难保证一致,因此,可能会导致两者的漫反射系数不同;同时,两台仪器光源(卤钨灯)的波段能量差异、光栅狭缝宽度(约 3  $\mu\text{m}$ )不完全一致、光学器件安装工艺的细微偏差等,经精密的光路和电子器件的放大均会产生明显的仪器台间差。虽然,以更苛刻的工艺条件可以减少台间差,但这意味着制造成本的大幅增加,将极不利于小型化便捷式近红外光谱仪的市场化推广。因此,为实现目标机能共享源机模型,选择合适的模型传递算法非常必要,即通过软件,以算法修正弥补硬件制造上的不足,降低仪器成本的同时提高仪器性能。

## 3.3 木质素含量近红外光谱分析模型传递

### 3.3.1 转换集样品数的选择

S/B、DS、CCA 算法均为有标样算法,需要从源机和目标机的标样集分别选取转换集样品。转换集样品数的选取对传递效果有重要的影响,样品数目太少,则转换信息不充分;样品数目太多,则实际处理过程太复杂<sup>[20]</sup>。本研究采用 Kennard-Stone (K-S) 算法从源机和目标机的标样集中分别取 10、20、30、40、50、63 个样品作为模型传递的转换集。转换集样品数与 RMSEP 的关系如图 3 所示。由图 3 可知,S/B 算法中, RMSEP 随转换集样品数的增加逐渐减小,且变化趋势较为平缓;由此可知,转换集样品数

的选择对 S/B 算法的模型传递效果影响不大。DS 算法中,当转换集样品数为 50 时, RMSEP 最小,此时得到的是最佳转换集样品数;转换集样品数在 20~40 间, RMSEP 基本无变化。CCA 算法中,当转换集样品数为 50 时, RMSEP 最小;转换集样品数在 20~40 间, RMSEP 急剧降低。DS 算法和 CCA 算法中,当转换集样品数 > 50 时, RMSEP 均出现增大趋势。由于 S/B 算法模型传递效果较差,且转换集样品数的选取对其影响较小,因此,在选择转换集样品数时只需考虑 DS 算法与 CCA 算法即可。若考虑到模型传递的低复杂度<sup>[6]</sup>,可选择转换集样品数为 40,此时 CCA 算法的 RMSEP 为 1.4645%,模型传递的效果较好;若考虑达到最佳传递效果,则选择转换集样品数选为 50,此时 CCA 算法和 DS 算法的 RMSEP 都最小,即模型传递效果最佳。因此,本研究选用 50 个标样作为转换集。

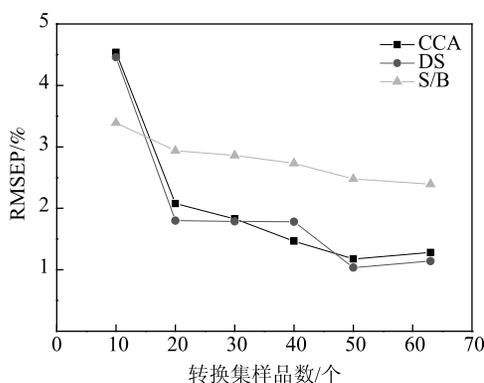


图3 RMSEP 随转换集样品数的变化

### 3.3.2 模型传递结果与分析

本研究分别采用了 S/B、DS 和 CCA 算法选取的最佳转换集样品数对目标机进行了模型传递,传递后的各项预测指标结果如表 4 所示。由表 4 可知,经 S/B 算法传递后的各项指标均较未传递前差,而经另外两种算法传递后,模型的预测效果较未传递前的预测效果均有大幅提升。这可能是由于 S/B 算法只能对预测结果进行线性校正,当两台仪器间差异性比较复杂时,此算法的模型传递效果不佳。而 DS 算法和 CCA 算法主要是利用信号处理方法消除或降低光谱间存在的差异性,因此,这两种算法的模型传递效果更好。经 DS 算法传递后,  $R^2$ 、RMSEP 和 RPD 分别达到 0.9643, 1.0370% 和 5.3513,且经此算法传递后,PLS 校正模型对目标机的预测效果与对源机的预测效果基本相同。CCA 算法在模型传递过程中需要提取标样光谱间的最大相关典型变量,由于本研究所使用的仪器为低分辨率、低信噪比的便携式近红外光

谱仪,在进行 CCA 算法时,若仅考虑提取的典型变量的最大相关,则可能引入与目标无关的冗余信息<sup>[6]</sup>;所以,经 CCA 算法传递后,模型的预测结果要略低于 DS 算法的预测结果。DS 算法计算转换矩阵时,能把仪器之间的差异以及标样在两仪器中测量状态的变化都校正到模型中去<sup>[18]</sup>,这也使得 DS 算法较 CCA 算法的传递效果略好,但 CCA 算法的 3 项指标  $R^2$ 、RMSEP 和 RPD 分别为 0.9540、1.1766% 和 4.7711,基本也可满足实际应用的需求。

表4 3种算法传递后的预测结果

算法	$R^2$	RMSEP/%	RPD
未传递前	0.8614	2.0415	2.7498
S/B	0.7959	2.4779	2.7155
DS	0.9643	1.0370	5.3513
CCA	0.9540	1.1766	4.7711

## 4 结论

本研究分别采用了斜率截距算法(S/B)、直接校正算法(DS)和典型相关分析算法(CCA)研究了制浆材中木质素含量近红外光谱模型的建立与传递问题。结果表明:

(1) S/B、DS 和 CCA 算法都是采用转换集进行模型传递的算法,研究中发现,转换集样品数对 S/B 算法的模型传递效果影响较小,但对 DS 和 CCA 算法则有不同程度的影响;当选取较少样品数时,CCA 算法的预测效果要优于 DS 算法的预测效果;而选择较多样品数时,DS 算法的预测效果优于 CCA 算法,但两者差异较小。

(2) 3 种算法中,基于线性校正的 S/B 算法不能得到满足精度要求的模型传递效果,而 DS 算法和 CCA 算法则利用信号处理方法消除或降低光谱间存在的差异性,较好地实现了多台同型号、低分辨率的便携式近红外光谱仪间的模型传递。

(3) 任何一种模型传递方法不可能适用于所有的实际情况;因此,本研究中得出 DS 算法和 CCA 算法进行模型传递的有效性是针对本研究所采用的仪器及其分析对象。实际中,每种模型传递方法各有优劣,在实际应用中应根据具体情况选择最适合的解决办法。研究中转换集样品数选取较多,而在实际生产应用中,较大的转换集样品数会使样品采集及数据处理工作量变大,不利于模型的传递。因此,采用少标样或无标样进行模型传递将是今后研究的重点。

## 参 考 文 献

- [1] He Wen-ming, Xue Chong-yun, Nie Yi, et al. Rapid Prediction of Wood Cellulose, Pentosan and Klason Lignin Contents Using Near Infrared Spectroscopy [J]. Transactions of China Pulp and Paper, 2010, 25(3): 9.  
贺文明, 薛崇韵, 聂怡, 等. 近红外光谱法快速测定木材纤维素、戊聚糖和木质素含量的研究[J]. 中国造纸学报, 2010, 25(3): 9.
- [2] Wu Ting, Fang Gui-gan, Liang Long, et al. Four Kinds of Algorithms Used for the Determination of Pulpwood Properties by Near Infrared Spectroscopy[J]. Chemistry and Industry of Forest Products, 2016, 36(6): 63.  
吴珽, 房桂干, 梁龙, 等. 四种算法用于近红外测定制浆材材性的对比研究[J]. 林产化学与工业, 2016, 36(6): 63.
- [3] Liang Long, Shi Ying-qiao, Cui Hong-hui, et al. Application of Near-infrared Spectroscopy in Pulp and Paper Industry[J]. Transactions of China Pulp and Paper, 2015, 30(4): 61.  
梁龙, 施英乔, 崔宏辉, 等. 近红外光谱分析技术在制浆过程中的应用[J]. 中国造纸学报, 2015, 30(4): 61.
- [4] Lu Wan-zhen. Modern Near Infrared Spectroscopy Analytical Technology[M] 2nd Ed. Beijing: China Petrochemical Press, 2000: 87.  
陆婉珍. 现代近红外光谱分析技术[M]. 2版. 北京: 中国石化出版社, 2000: 87.
- [5] Tsuchikawa S, Schwanninger M. A Review of Recent Near-Infrared Research for Wood and Paper (Part 2) [J]. Applied Spectroscopy Reviews, 2013, 48(7): 560.
- [6] Zhang Jin, Cai Wen-sheng, Shao Xue-guang. New Algorithms for Calibration Transfer in Near Infrared Spectroscopy[J]. Progress in Chemistry, 2017(8): 101.  
张进, 蔡文生, 邵学广. 近红外光谱模型转移新算法[J]. 化学进展, 2017(8): 101.
- [7] Liu Xian, Dong Su-xiao, Han Lu-jia, et al. Calibration Transfer of NIRS on Silage by Slope/Bias and Local Centering Technique[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(5): 153.  
刘贤, 董苏晓, 韩鲁佳, 等. 青贮饲料近红外光谱分析模型转移研究[J]. 农业机械学报, 2009, 40(5): 153.
- [8] Zhao Long-lian, Lao Cai-lian, Yan Yan-lu. Near-infrared Calibration Transfer between Different Diffuse-reflectancing Sample Cells [J]. Journal of China Agricultural University, 1998(S3): 91.  
赵龙莲, 劳彩莲, 严衍禄. 不同漫反射样品池间近红外数学模型的转移[J]. 中国农业大学学报, 1998(S3): 91.
- [9] Li Qing-bo, Zhang Guang-jun, Xu Ke-xin, et al. Application of DS Algorithm to the Calibration Transfer in Near-Infrared Spectroscopy [J]. Spectroscopy and Spectral Analysis, 2007, 27(5): 873.  
李庆波, 张广军, 徐可欣, 等. DS算法在近红外光谱多元校正模型传递中的应用[J]. 光谱学与光谱分析, 2007, 27(5): 873.
- [10] Wei F, Liang Y, Yuan D, et al. Calibration model transfer for near-infrared spectra based on canonical correlation analysis[J]. Analytica Chimica Acta, 2008, 623(1): 22.
- [11] Wang Yan-bin, Yuan Hong-fu, Lu Wan-zhen. A New Calibration Transfer Method Based on Target Factor Analysis [J]. Spectroscopy and Spectral Analysis, 2005, 25(3): 398.  
王艳斌, 袁洪福, 陆婉珍. 一种基于目标因子分析的模型传递方法[J]. 光谱学与光谱分析, 2005, 25(3): 398.
- [12] Chen Bin, Wang Hao. Calibration Transfer Between Near-infrared Spectrometric Instrument for the Determination of Wine Alcoholicity Using Shen's Algorithm [J]. Infrared Technology, 2006, 28(4): 245.  
陈斌, 王豪. 专利算法在白酒酒精度近红外光谱分析模型转移中的应用[J]. 红外技术, 2006, 28(4): 245.
- [13] Tian Gao-you, Chu Xiao-li, Yuan Hong-fu, et al. Application of Wavelet Transform-Piecewise Direct Standardization on the Near Infrared Analysis Model Transfer [J]. Analytical Chemistry, 2006, 34(7): 927.  
田高友, 褚小立, 袁洪福, 等. 小波变换-分段直接校正法用于近红外光谱模型传递研究[J]. 分析化学, 2006, 34(7): 927.
- [14] Xin Xiao-wei, Gong Hui-li, Ding Xiang-qian, et al. Study on Calibration Model Transfer for the Near Infrared Spectrum Based on Improved S/B Algorithm [J]. Spectroscopy and Spectral Analysis, 2017, 37(12): 3709.  
信晓伟, 宫会丽, 丁香乾, 等. 改进S/B算法的近红外光谱模型转移[J]. 光谱学与光谱分析, 2017, 37(12): 3709.
- [15] Eskildsen C, Hansen P, Skov T, et al. Evaluation of multivariate calibration models transferred between spectroscopic instruments: applied to near infrared measurements of flour samples [J]. Journal of Near Infrared Spectroscopy, 2016, 24(2): 151.
- [16] Chu Xiao-li. Molecular Spectroscopy Analytical Technology Combined with Chemometrics and its Applications [M]. Beijing: Chemical Industry Press, 2011: 89.  
褚小立. 化学计量学方法与分子光谱分析技术 [M]. 北京: 化学工业出版社, 2011: 89.
- [17] Zhang Xiao-chao, Wu Jing-zhu, Xu Yun. Near infrared spectroscopy and its application in modern agriculture [M]. Beijing: Publishing House of Electronics Industry, 2012: 49.  
张小超, 吴静珠, 徐云. 近红外光谱分析技术及其在现代农业中的应用 [M]. 北京: 电子工业出版社, 2012: 49.
- [18] Hu Chang-qin, Feng Yan-chun. Rapid analysis of drugs by near infrared spectroscopy [M]. Beijing: Chemical Industry Press, 2010: 68.  
胡昌勤, 冯艳春. 近红外光谱法快速分析药品 [M]. 北京: 化学工业出版社, 2010: 68.
- [19] Zhang Hong-guang. Fundamental research on novel modeling methods and applications of near infrared spectroscopy [D]. Hangzhou: Zhejiang University, 2015.  
张红光. 近红外光谱新型建模方法与应用基础研究 [D]. 杭州: 浙江大学, 2015.
- [20] Yang Hui-hua, Zhang Xiao-feng, Fan Yong-xian, et al. Near Infrared Spectroscopic Model Transfer Based on Simple Linear Regression [J]. Analytical Chemistry, 2014(9): 1229.  
杨辉华, 张晓凤, 樊永显, 等. 基于一元线性回归的近红外光谱模型传递研究[J]. 分析化学, 2014(9): 1229.

### Study on Near-infrared Calibration Model Transfer for Lignin Content in Pulpwood

LIU Yaoyao<sup>1</sup> YANG Hao<sup>1</sup> XIONG Zhixin<sup>1,\*</sup> LIANG Long<sup>2</sup> FANG Guigan<sup>2</sup>

(1. *Jiangsu Provincial Key Lab of Pulp and Paper Science and Technology, Nanjing Forestry University, Nanjing, Jiangsu Province, 210037*; 2. *Institution of Chemical Industry of Forestry Products, CAF, Nanjing, Jiangsu Province, 210042*)

(\* E-mail: leo\_xzx@njfu.edu.cn)

**Abstract:** The near-infrared calibration model transfer for lignin content in pulpwood was investigated between two portable near-infrared spectrometers. An optimal calibration model of master was established by partial least square (PLS) after the selection of representative infrared spectroscopy data net samples, preprocessing and eliminating outlier samples. The near-infrared spectroscopy calibration model was transferred between master and slave by the algorithms of slope/bias (S/B), direct standardization (DS) and canonical correlation analysis (CCA), respectively, and the prediction results were compared. The results indicated that the models transferred by DS and CCA improved the prediction accuracy significantly comparing to the algorithm of S/B with a poor performance. The coefficient of determination ( $R^2$ ), root mean square error of prediction (RMSEP) and ratio of performance to standard deviate (RPD) by DS were 0.9643, 1.0370%, 5.3513, and by CCA were 0.9540, 1.1766%, 4.7711, respectively. Therefore, both DS and CCA algorithms could achieve the calibration model transfer between the two portable near-infrared spectrometers.

**Keywords:** near-infrared spectroscopy; model transfer; lignin; pulpwood

(责任编辑:杨 艳)