第42卷 第3期　　华 中 农 业 大 学 学 报　　Vol.42 No.3
2023年 5月　　Journal of Huazhong Agricultural University　　May 2023,195～202

李旭,陈柏林,周保平,等.利用叶片高光谱反射率预测棉花叶绿素含量[J].华中农业大学学报,2023,42(3):195-202.
DOI:10.13300/j.cnki.hnlkxb.2023.03.023

利用叶片高光谱反射率预测棉花叶绿素含量

李旭[1],陈柏林[2],周保平[1],石子琰[1],洪国军[1]

1.塔里木大学信息工程学院,阿拉尔 843300; 2.塔里木大学化学化工学院,阿拉尔 843300

摘要　为提高棉花叶绿素含量预测的准确性,利用连续小波分析和传统光谱变换对棉花叶片原始光谱进行分解和变换,以特征小波系数和光谱特征波段为自变量,并利用单变量、逐步回归和偏最小二乘法建立反演棉花叶片叶绿素含量的数学模型。结果显示,不同的光谱处理方法使得棉花叶片叶绿素和光谱反射率的相关性都有不同程度的提升,对于传统光谱变换,倒数对数一阶微分$\lg(1/R')$对棉花叶片叶绿素相关性提高了0.41。结果表明,连续小波分析在信息降噪和挖掘特征信息方面优于传统光谱模型,建立的模型RPD>2,具有很好的稳定性,对样本数据都具很好的预测能力。

关键词　高光谱; 无损检测; 连续小波分析; 传统光谱变换; 叶绿素; 棉花

中图分类号　S562; S127　　**文献标识码**　A　　**文章编号**　1000-2421(2023)03-0195-08

2020年新疆棉花产量为520万t,占国内棉花总产量的87%左右,对新疆的经济和社会发展起着举足轻重的地位。叶绿素是影响作物光合作用的重要因素,与作物的生长和产量息息相关。研究表明,叶绿素与叶片氮浓度也具有相关性[1]。氮肥的过量使用会对土壤和地下水资源造成一定破坏和污染[2-3]。传统的叶绿素检测一般采用化学分析方法,虽然其具有较高的精度,但是化学分析过程复杂、耗时并且需要破坏作物,不能满足现代农业快速无损检测的要求。

光谱分析技术具有快速、无损等特点,随着光谱分析技术的不断发展,已被广泛应用于农作物生化参量的监测分析,进而为农业生产提供科学化管理。研究人员使用卫星、无人机和手持高光谱设备对玉米[4]、小麦[5-6]、水稻[7]和棉花[8]等农作物的营养元素和叶绿素建立了较高水平的预测模型。楚万林等[9]采用原始光谱、一阶导数和植被指数等手段建立了棉花叶绿素与高光谱参数之间的估算模型,其结果表明经一阶导数处理建立的多元逐步回归模型的预测效果最优。王登伟等[10]发现棉花冠层叶绿素在762 nm处具有最大的相关系数,对于一阶微分光谱,单叶的叶绿素和冠层叶绿素的敏感波段均在750 nm处。Chen等[11]将棉花黄萎病叶片光谱反射率和一阶微分相应特征参数与叶绿素进行相关分析,其研究结果对运用高光谱定量估算病害叶片叶绿素含量具有实用价值。另外,光谱分析技术在作物识别[12]和病虫害[13]信息检测方面也有较多的应用。虽然卫星和无人机已经广泛应用于高光谱监测中,并且监测精度也具有较高水平,但是卫星和无人机由于自身的特点,在时间分辨率和重复性监测方面存在缺陷,并且还受到成本高和技术操作等方面的限制[14]。近地面高光谱,例如手持式高光谱仪,可以避免云层和减少大气环境的影响,获得较高的空间分辨率,识别出细尺度的特征,从而大大提高作物生长信息的获取准确度。原始光谱反射率会受到大气、光照等背景噪声以及相邻波段的干扰,因此,对高光谱数据进行处理和变换可以挖掘出隐藏的特征信息、减少背景干扰和数据冗余。

利用数学分析方法对高光谱反射率进行变换,已应用于植被作物氮素和叶绿素预测研究[15-16]。小波分析具有多分辨率和方向性变化的多尺度分析功能,已应用于作物叶绿素含量预测以及重金属元素、生理信息、病虫害等监测[17-18]。本研究利用连续小波分析与传统光谱变换对棉花叶片原始光谱进行处

收稿日期: 2022-07-31
基金项目:国家自然科学基金项目(61563046);绿洲生态农业兵团重点实验室开放课题(202002)
李旭,E-mail:lixu2866@126.com
通信作者: 周保平,E-mail:502805150@qq.com

理，比较连续小波分析和传统光谱变换这2种方法所建立的模型对棉花叶片叶绿素含量的估测能力，并采用单变量分析、逐步回归分析和偏最小二乘法等方法比较建立的预测模型的精度和稳定性，旨在挖掘更利于反映棉花叶片叶绿素的特征波段，进一步提高棉花叶片叶绿素含量预测的准确性，在提高作物产量和品质同时可以避免氮肥的过量使用。

1 材料与方法

1.1 试验地点与仪器

试验于2021年6月和10月在塔里木大学园艺实验站开展，塔里木大学位于新疆维吾尔自治区阿拉尔市，阿拉尔市北起天山南麓山地，南至塔克拉玛干沙漠北缘，属于温带大陆性气候，年平均高温20 ℃、平均低温4 ℃，年均降水量为40.1～82.5 mm，全年棉花种植面积15.5万 hm^2。试验使用FieldSpec Hand Held 2便携式地物光谱仪采集棉花叶光谱特征，该仪器适用于植物光谱和土壤营养元素关系研究，波长范围为325～1 075 nm，光谱分辨率小于3 nm，视场角25°；使用SPAD-502叶绿素通过测量在650、940 nm波长下的透光系数，以确定当前叶片的相对叶绿素含量，SPAD测量范围为－9.9～199.9，精度在±1.0。

1.2 数据采集

1)光谱反射率测定。使用FieldSpec Hand Held 2便携式地物光谱仪测定，光谱采样间隔为1 nm，使用之前先在白板上进行校正测验，选择天气晴朗，无风或微风的条件下在10:00—14:00采集棉花叶片光谱数据，由于太阳位置的变化，每10 min将进行1次校正。测量时手持光谱仪，保持传感器垂直向下，探头视角25°，均匀地选择具有代表性的棉花叶片，每个叶片测量10次，取平均值。

2)叶片叶绿素测定。采用SPAD-502手持式叶绿素仪测定棉花叶片叶绿素。每次测量前利用仪器自带的检验卡进行校正，确保每次测量仪器正常，误差可控。每个试验小区均匀选取3片典型叶片，在每片叶片不同位置测量3次，取其平均值作为试验小区的叶片叶绿素含量。

1.3 研究方法

1)连续小波分析。连续小波分析方法由于具有多尺度分解特性，相比于传统的基于几个波段的植被指数，可以从高光谱遥感数据中获取生化参数的吸收特征和叶面积指数对反射率的振幅影响等信息。基本步骤是对原始反射光谱进行连续小波变换，即通过平移和缩放的母小波函数对反射光谱进行卷积运算，以此获得小波系数。

$$\varphi_{a,b}(\lambda)=\frac{1}{\sqrt{a}}\varphi\left(\frac{\lambda-b}{a}\right) \tag{1}$$

$$W_f(a,b)=\int_{-\infty}^{+\infty}f(\lambda)\varphi_{a,b}(\lambda)\mathrm{d}_\lambda \tag{2}$$

其中，$\varphi(\lambda)$为母小波函数，$\varphi_{a,b}(\lambda)$为平移和缩放后的母小波函数，a为缩放因子，也可以称为尺度，b为平移因子，即波段位置。$W_f(a,b)$为小波系数(特征小波)，可以看做小波母函数在尺度a和b波段处反射率数据的相似度。

本研究使用Mexican Hat小波作为连续小波变换的母函数，本质为高斯函数的二阶导函数，公式为：

$$\varphi(x)=\frac{2}{\sqrt{3}}\pi^{\frac{-1}{4}}(1-x^2)e^{\frac{-x^2}{2}} \tag{3}$$

$$\int_{-\infty}^{+\infty}\varphi(x)\mathrm{d}(x)=0 \tag{4}$$

高斯二阶导函数与反射光谱的反射峰和反射谷的形状相似，能够很好地进行局部优化和有效除去部分环境噪音的干扰，起到较好的平滑作用。

2)传统光谱变换。将已经平滑处理后的反射率进行倒数对数、一阶微分和一阶倒数对数微分的变换，具体计算公式如式(5)、(6)。

$$R'_\lambda=\frac{R_{\lambda+1}-R_{\lambda-1}}{2\Delta\lambda} \tag{5}$$

$$\lg\left(\frac{1}{R'_\lambda}\right)=\frac{\lg\left(\frac{1}{R_{\lambda+1}}\right)-\lg\left(\frac{1}{R_{\lambda-1}}\right)}{2\Delta\lambda} \tag{6}$$

其中，R'_λ是波长λ处的一阶微分值，$R_{\lambda+1}$和$R_{\lambda-1}$分别是波长(λ＋1)和(λ－1)处光谱反射率值；$\lg(\frac{1}{R'_\lambda})$是倒数对数在波长λ处的一阶微分值，$\lg(\frac{1}{R_{\lambda+1}})$和$\lg(\frac{1}{R_{\lambda-1}})$分别是在波长(λ＋1)和(λ－1)处倒数对数值。

3)模型评价指标。本研究选择决定系数(R^2)、均方根误差(RMSE)和相对分析误差(RPD)对叶绿素和全氮反演模型的拟合优度和预测精度进行评价。其中，R^2越趋近于1，RMSE越趋近于0，表明模型的拟合优度越好。RPD用于模型评价时，当RPD<1.4时，说明回归模型的精度最差，无法对样

本数据进行有效的预测；当1.4≤RPD<2时，说明回归模型的精度一般，可以对样本数据进行粗略的预测；当RPD≥2时，说明回归模型的精度较好，能对样本数据进行有效的预测。

1.4　数据处理

本研究所涉及的连续小波分析、相关系数计算、决定系数等高线图绘制以及模型回归分析均在 MATLAB R2019b、Origin 2021 和 SPSS 中完成。

2　结果与分析

2.1　叶片叶绿素样本数据统计分析

在6月蕾期和10月吐絮期共采集棉花叶片叶绿素样本200个，随机选取30%的样本为验证集，70%为建模集，叶绿素数据分布和统计如表1所示。由表1可知，3组叶绿素样本变化范围和变异系数变化不大，正态曲线数据分布均匀，3组样本经单因素分析无显著性差异，表明建模集合验证集划分合理，适合用于建模和验证。

表 1　叶绿素样本数据统计

Table 1　Chlorophyll sample statistics table

数据集 Data set	样本数 Sample size	最小值 Minimum value	最大值 Maximum value	Average 平均值	标准差 Standard deviation	方差 Variance	变异系数/% Coefficient of variation
总样本 Total sample	200	38.80	62.83	54.64	6.345 1	40.260 4	11.61
建模集 Modeling set	140	38.80	62.83	54.71	6.372 4	40.607 2	11.65
验证集 Validation set	60	40.41	62.50	54.49	6.331 8	40.091 9	11.62

注：叶绿素含量单位为SPAD。Note：The unit of chlorophyll content is SPAD.

2.2　连续小波分析的棉花叶片叶绿素相关性分析

通过将2021年采集的200组光谱数据进行连续小波分解后，得到同一波段对应不同分解尺度的小波系数，将得到的小波系数依次与实测值进行相关性分析，最终得到棉花叶片叶绿素的小波系数敏感度R^2的分布图(图1)。由图1可知，叶片叶绿素敏感的小波系数主要分布在480 nm蓝光波段附近的中低尺度区域和719 nm红光附近的中高尺度区域，这是因为绿色植被叶片的光谱特性主要由叶绿素和各种色素决定，700～780 nm波段是叶绿素a在红波段的强吸收区域，480 nm附近是叶绿素b的强吸收区域。值得注意的是，在500 nm附近有小波系数对叶片叶绿素和叶片含氮量共同敏感的区域，如不区分共同敏感区域，直接建立预测模型，势必会导致预测模型的反演精度偏低，模型稳定性较差。为了进一步提升数据模型的精确性，选取叶片叶绿素的小波特征系数需要排除叶片叶绿素与叶片含氮量共同敏感区域。以排除共同敏感区为原则，选取相关性最大的10个小波系数进行对比分析。

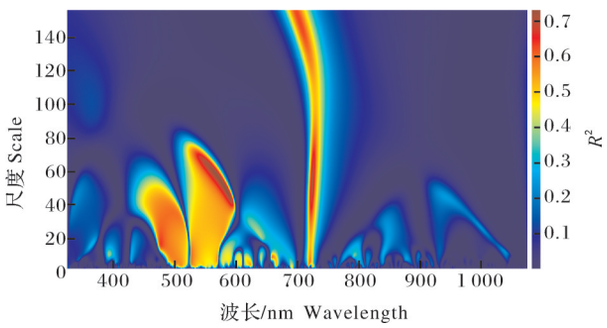


图 1　棉花叶片叶绿素小波系数敏感度 R^2 分布图

Fig.1　R^2 distribution of the sensitivity of chlorophyll wavelet coefficients in cotton leaves

2.3　传统光谱变换的棉花叶片叶绿素相关性分析

受测量环境的影响，反映在光谱曲线上表现为“毛刺”，需对原始光谱反射率进行平滑预处理，再进行微分、倒数对数变换。本研究在Origin中采用Savitzky-Golay卷积平滑消除“毛刺”，将获得的不同光谱变换后数据与叶片叶绿素进行相关性分析，得到棉花叶片叶绿素和叶片光谱变换的相关系数(图2)。图2反映了棉花叶片叶绿素与原始光谱反射率及光谱反射率变换后数据的相关性情况，从图2中可以发现，不同光谱变换之间的差异，其原始光谱反射率在整条波段范围内与叶片叶绿素几乎成负相关，在325～652 nm和689～837 nm，达到极显著性水平($P<0.01$)，相关系数为−0.442～−0.180，经过数学变换后，在原始光谱中的负相关峰逐渐出现部分正相关和负相关峰；lg(1/R)处理后的数据与叶片叶绿素在波长范围内几乎成正相关，在325～830 nm达到极显著性水平，相关系数0.211～0.468；R'和lg(1/R')处理后的数据相关性在波长范围内呈现正、负交替变化，曲线波动较大，相关系数为−0.8～0.7。从图2可以得到原始光谱提供的超过0.01显著水平的特征波长大概有477个，经过数学变

换后，通过0.01显著性检验的特征波长数有所减少，相关性更加突出，同时图2表明，棉花叶片叶绿素大部分敏感波段主要位于蓝边(490～530 nm)和红边(670～760 nm)边缘区域。

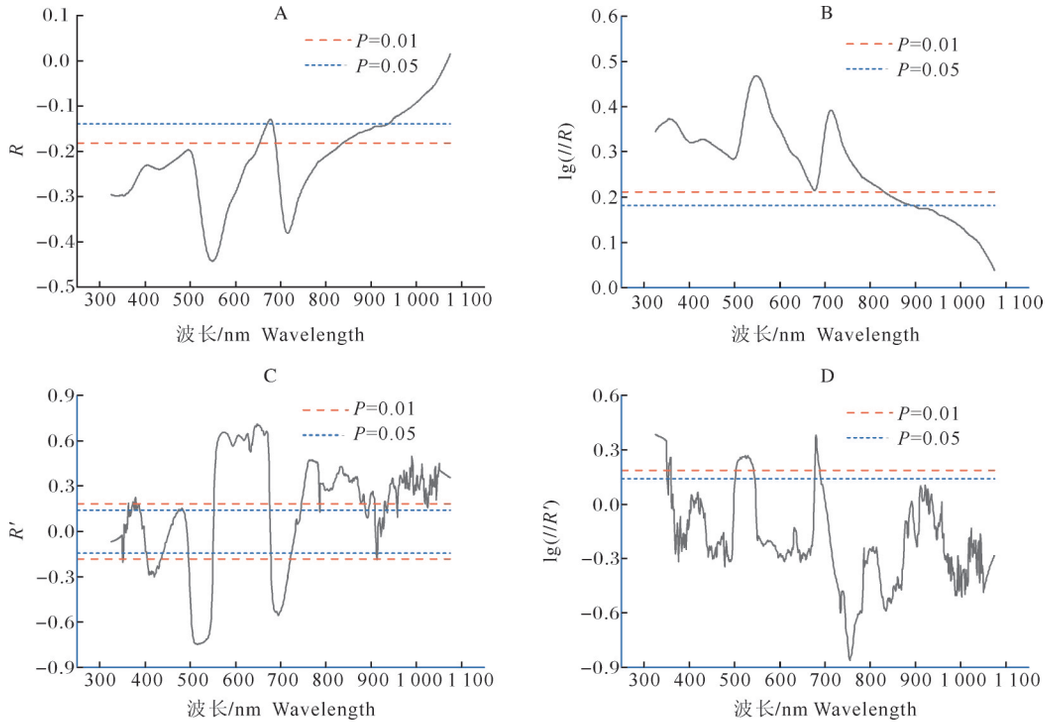


图2 传统光谱变换与叶片叶绿素的相关系数分布图

Fig.2 Distribution of correlation coefficients between spectral transformations and leaf chlorophyll

2.4 棉花叶片叶绿素反演模型的建立

1)单变量回归分析模型。分别以经连续小波分析获得特征小波系数以及原始光谱反射率和3种传统光谱变换的光谱参数为自变量，以棉花叶片叶绿素为因变量，通过单自变量回归方法构建棉花叶片叶绿素估算模型，其建模和验证结果如表2所示。

由表2可知，经过不同形式和方法处理后的光谱数据，在一定程度上提升了光谱反射率与棉花叶片叶绿素的相关性和反演模型的精度。以原始光谱反射率R和倒数对数$\lg(1/R)$处理的数据为自变量建立的单变量模型的决定系数R^2较小，RPD小于1.4，说明模型预测效果很差。而通过一阶微分变换后的数据建立的模型决定系数R^2有明显的提高，RMSE有所减小，即模型精度高于前者，但RPD小于2，模型对样本数据的预测效果一般。其中，经倒数对数一阶微分$\lg(1/R')$处理后的数据所建立的二次多项式模型，建模集和预测集决定系数R^2分别0.687 8和0.741 8，建模集和预测集RMSE分别为3.536 4和3.316 0，说明由倒数对数一阶微分$\lg(1/R')$处理后建立的模型具有较高的精度。以连续小波分析后得到特征小波系数为自变量的模型，在建模集中R^2均在0.719 1～0.726 7，RMSE整体差异变化不大且多项式模型明显小于线性模型，其中$W_f(48,585)$的二阶多项式模型R^2最大，RMSE为3.319 3。在验证集模型中，所有模型的R^2值均在0.186 1～0.817 5，其中R^2的最大值(0.817 5)是位于$W_f(48,585)$的线性模型。RMSE在2.859 4～2.907 1，整体差异不大，RPD在2.178～2.214，均大于2，表明由连续小波分析处理后建立的模型具有较好的精度，稳定性更好，能够对样本数据进行有效的预测，其中，$W_f(48,585)$线性模型的预测精度更高，稳定性更好。

不同方法的光谱变换和处理都在一定程度上增强光谱值与叶片叶绿素的相关性，使得具有高相关性的波段相对集中，在回归分析时，模型精度比原始光谱反射率模型精度有不同程度的提升，从而证明了不同光谱变换的有效性。

2)多变量回归分析模型。分别以经连续小波分析得到的特征小波系数以及原始光谱反射率和3种传统光谱变换的光谱参数为自变量，以棉花叶片叶绿素为因变量，通过逐步回归和偏最小二乘回归方法构建棉花叶片叶绿素估算模型，其建模和验证结果如表3所示。由表3可知，经过连续小波分析分解后建立的多变量模型具有更好的模型精度，预测偏

表 2　单自变量回归模型建模及验证结果

Table 2　Univariate regression model modelling and validation results

变量 Variable	模型 Model	回归方程 Regression equation	建模集 Modelling set		验证集 Validation set		
			R^2	RMSE	R^2	RMSE	RPD
$W_f(51,580)$	线性 Linear	$y=-73.169x+57.976$	0.719 9	3.360 6	0.817 1	2.8641	2.211
	多项 Multiple	$y=61.473x^2-83.254x+57.970$	0.724 0	3.335 7	0.814 7	2.893 1	2.189
$W_f(50,582)$	线性 Linear	$y=-78.186x+57.749$	0.721 5	3.350 9	0.817 1	2.867 3	2.208
	多项 Multiple	$y=62.759x^2-87.390x+57.714$	0.724 9	3.330 4	0.814 9	2.893 3	2.188
$W_f(49,583)$	线性 Linear	$y=-79.860x+58.432$	0.721 1	3.353 3	0.816 9	2.859 5	2.214
	多项 Multiple	$y=76.960x^2-92.340x+58.496$	0.725 8	3.325 0	0.815 1	2.883 0	2.196
$W_f(49,584)$	线性 Linear	$y=-84.540x+57.410$	0.721 5	3.350 7	0.815 1	2.884 5	2.195
	多项 Multiple	$y=58.773x^2-91.933x+57.347$	0.723 8	3.337 3	0.813 2	2.906 3	2.179
$W_f(48,584)$	线性 Linear	$y=-81.737x+59.074$	0.719 1	3.365 2	0.814 7	2.865 3	2.210
	多项 Multiple	$y=87.692x^2-97.058x+59.262$	0.724 6	3.331 9	0.813 3	2.884 8	2.195
$W_f(48,585)$	线性 Linear	$y=-87.168x+58.095$	0.723 2	3.340 6	0.817 5	2.860 2	2.214
	多项 Multiple	$y=78.251x^2-98.066x+58.101$	0.726 7	3.319 3	0.815 9	2.881 1	2.198
$W_f(47,586)$	线性 Linear	$y=-90.349x+58.696$	0.721 6	3.350 1	0.816 7	2.859 4	2.214
	多项 Multiple	$y=94.323x^2-104.349x+58.794$	0.726 1	3.323 4	0.815 2	2.878 9	2.199
$W_f(47,587)$	线性 Linear	$y=-96.715x+57.551$	0.721 9	3.348 5	0.814 3	2.890 8	2.190
	多项 Multiple	$y=68.690x^2-104.392x+57.503$	0.723 7	3.337 5	0.812 8	2.907 1	2.178
$W_f(46,588)$	线性 Linear	$y=-101.337x+58.156$	0.722 0	3.347 9	0.816 0	2.872 7	2.204
	多项 Multiple	$y=90.955x^2-112.195x+58.163$	0.724 6	3.331 8	0.814 6	2.888 7	2.192
$W_f(45,589)$	线性 Linear	$y=-106.782x+58.705$	0.719 5	3.362 6	0.815 5	2.869 1	2.207
	多项 Multiple	$y=111.273x^2-120.539x+58.781$	0.722 8	3.343 3	0.814 2	2.884 5	2.195
R_{715}	线性 Linear	$y=-17.706x+60.473$	0.145 5	5.850 6	0.186 1	5.666 2	1.117
	多项 Multiple	$y=46.140x^2-49.946x+65.229$	0.170 9	5.763 1	0.201 4	5.621 7	1.126
R'_{516}	线性 Linear	$y=-6077.566x+64.471$	0.556 4	4.215 3	0.609 7	3.923 8	1.614
	多项 Multiple	$y=-281413.7x^2-4992.8x-63.623$	0.557 4	4.210 7	0.613 4	3.906 3	1.621
$\lg(1/R_{548})$	线性 Linear	$y=12.703x+43.520$	0.219 3	5.592 4	0.264 7	5.387 5	1.175
	多项 Multiple	$y=-3.793x^2+19.898x+40.334$	0.221 3	5.585 0	0.262 8	5.391 2	1.174
$\lg(1/R'_{755})$	线性 Linear	$y=-14872.623x+37.120$	0.621 1	3.895 7	0.621 4	4.038 3	1.619
	多项 Multiple	$y=-1.462x^2-49787.670x+17.930$	0.687 8	3.536 4	0.741 8	3.316 0	1.972

表 3　多变量回归模型建模及验证结果

Table 3　Multivariate regression model modelling and validation results

回归模型 Regression model	建模集 Modelling set		验证集 Validation set		
	R^2	RMSE	R^2	RMSE	RPD
CWT-SR	0.723 0	3.340 6	0.817 5	2.860 2	2.214
R-SR	0.196 1	5.674 9	0.223 3	5.544 2	1.142
R'-SR	0.768 0	4.053 7	0.659 8	3.672 8	1.724
$\lg(1/R)$-SR	0.219 3	5.592 4	0.187 7	5.666 7	1.117
$\lg(1/R')$-SR	0.649 0	3.749 5	0.688 0	3.703 1	1.766
CWT-PLSR	0.723 1	3.341 3	0.817 8	2.858 9	2.215
R-PLSR	0.174 9	5.749 1	0.209 8	5.581 4	1.134
R'-PLSR	0.565 0	4.174 5	0.641 3	3.760 5	1.684
$\lg(1/R)$-PLSR	0.190 4	5.694 9	0.230 4	5.508 2	1.150
$\lg(1/R')$-PLSR	0.654 7	3.719 1	0.713 9	3.584 1	1.824

差更小，RPD 均大于 2，表明由连续小波分析建立的模型对样本数据具有有效的预测。

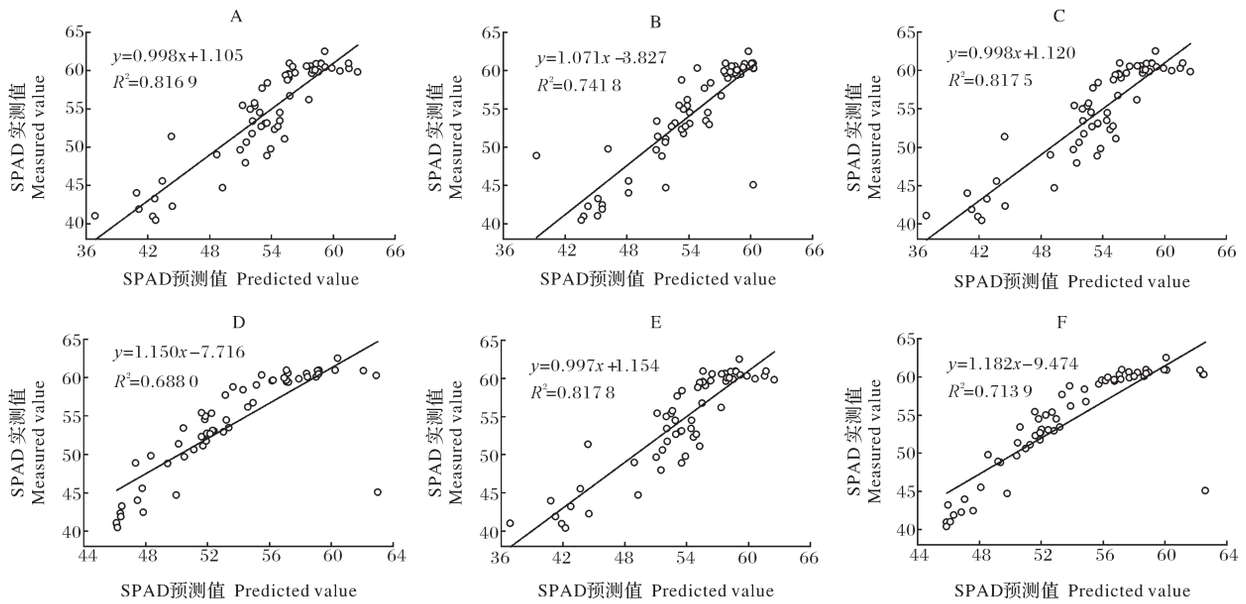
2.5　棉花叶片叶绿素回归模型对比分析

为分析不同建模方法间的差异，选择不同建模方法中的最优模型进行对比分析，结果如表 4 所示。从表 4 可知，经连续小波分析建立的回归模型中，验证集 CWT-PLSR 的决定系数 R^2 最大、RMSE 最小，即模型优度更好，预测精度更高；在倒数对数一阶微分 $\lg(1/R')$ 模型中，单自变量的模型优于多变量的模型，精度更高，与 CWT 模型不同的原因是，CWT 提取的特征小波系数处于的波长位置相对更集中，而 $\lg(1/R')$ 曲线波动较大，敏感波段位置相对分散，使得在引入多变量计算时受背景条件的影响更多，从而导致单变量模型优于多变量模型。为了更好地展现回归模型的结果，叶绿素反演模型线性结果如图 3 所示。

表4 回归模型建模及验证结果

Table 4 Modeling and validation results of the regression model

回归模型 Regression model	建模集 Modelling set		验证集 Validation set		
	R^2	RMSE	R^2	RMSE	RPD
$W_f(48,585)$	0.723 2	3.340 6	0.816 9	2.860 2	2.214
$\lg(1/R')$	0.687 8	3.536 4	0.741 8	3.316 0	1.972
CWT-SR	0.723 0	3.340 6	0.817 5	2.860 2	2.214
$\lg(1/R')$-SR	0.649 0	3.749 5	0.688 0	3.703 1	1.766
CWT-PLSR	0.723 1	3.341 3	0.817 8	2.858 9	2.215
$\lg(1/R')$-PLSR	0.654 7	3.719 1	0.713 9	3.584 1	1.824



A: $W_f(48,585)$; B: $\lg(1/R')$; C: CWT-SR; D: $\lg(1/R')$-SR; E: CWT-PLSR; F: $\lg(1/R')$-PLSR.

图3 叶绿素反演验证模型线性拟合结果

Fig.3 Linear fitting results of chlorophyll inversion validation model

3 讨 论

本研究利用连续小波分析与传统光谱变换对棉花叶片原始光谱中的特征信息进行提取，比较连续小波分析和传统光谱变换这2种方法所建立的模型对棉花叶片叶绿素含量的估测能力，并采用单变量分析、逐步回归分析和偏最小二乘法等方法比较建立的预测模型的精度和稳定性。结果表明，连续小波分析在信息降噪和挖掘特征信息方面优于传统光谱模型，建立的模型RPD>2，具有很好的稳定性，对样本数据都具有很好的预测能力。

光谱反射率曲线本质上属于电磁波信号的一种，在棉花生长过程中，棉花内部的各种生化参数对不同波段光谱的吸收和反射存在着差异，在反射光谱上表现为不同位置连续的“峰”和“谷”。小波变化是利用连续小波变化寻找出各种生化参数隐藏在光谱反射率中的特征吸收，连续小波分析方法已被大量应用于提取生化参数和叶面积指数，包括叶片高光谱和地面高光谱[19-20]。本研究采用连续小波分析方法中的Mexh小波分析，以保证数据分解过程的连续性，分解尺度a在已有的研究中为了简化计算过程的繁琐，一般在$a=2^n(n=0,1,2\cdots)$的尺度下进行分解，虽然此种处理方法可以缩短数据处理时间，但是由于数据的间断，可能会造成重要特征参数的缺失，因此，本研究采取尺度a从1到160的连续尺度分解。

传统光谱变换主要是采用数学处理方法，包括微分处理、倒数对数处理等常见数学处理方法，其作用可以在一定程度上减小大气、光照等背景噪声和

仪器因素所带来的干扰，同时也可以提高原始光谱反射率曲线的清晰度，与原始光谱相比，一些不明显的特征峰可能将被解析，以显示可被监测的信息[21]。研究结果表明，棉花叶片原始光谱经过一阶微分、倒数对数和倒数对数一阶微分的变换后，对棉花叶片叶绿素的相关性都有不同程度的提升，因此，经过传统数学处理后的光谱可在一定程度上解析出不太明显的特征信息，其中倒数对数一阶微分对棉花叶片叶绿素相关性提高了0.41，倒数对数对相关性的提升不太明显，倒数对数一阶微分可以有效地提取不同的敏感光谱参数，更好地反映植物自身的特性，而倒数对数主要是放大可见光区的光谱差异，将部分光谱数据由非线性关系转为线性关系。

本研究在采用连续小波和传统数学方法处理的基础上，还用单自变量回归分析、逐步回归分析和偏最小二乘法回归分析建立模型。单因变量回归分析是通过直接寻找y和x之间的相关性，建立一个y与x之间的函数关系。逐步回归分析（SR）是多元线性回归分析中的一种，通过选择自变量的方式建立最优回归方程，主要解决多变量共线性问题。偏最小二乘法回归分析（PLSR）是一种研究多因变量或者单因变量对多自变量的回归建模方法，包括了使用多元线性回归分析、主成分分析和典型相关性分析这3种基本分析方法，综合其优点，能够使在自变量存在严重多重相关性的情况下进行回归建模，有效地解决自模型共线问题，其回归模型具有很强的稳定性[22]。

棉花叶片叶绿素的小波系数敏感度R^2的分布图是经连续小波分析处理，而建立的叶绿素与特征小波系数之间的决定系数图，可以直观地看出敏感信息波段主要集中于蓝边波段和红边波段区域，Clevers[23]认为红边区域是个特殊区域，可以用于估算叶绿素和氮含量。与传统光谱变换与叶片叶绿素的相关系数分布图中的敏感波段相比，经连续小波分析所表现的敏感波段位置更加集中，原因在于连续小波可将原始光谱分解为多个不同尺度，使得更多的隐藏信息被挖掘，说明连续小波分析对原始光谱中隐藏信息的挖掘更为有效。由连续小波分析而建立的偏最小二乘法模型具有很好的预测能力，优于同样处理方式的其他模型，因为由棉花叶片叶绿素的小波系数敏感度R^2的分布图可知，连续小波分析使得敏感波段位置更加集中，而由传统数学方法处理得到的敏感波段位置相对分散，在引入多变量计算时受背景条件的影响更多。

参考文献 References

[1] YODER B J, PETTIGREW-CROSBY R E. Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400-2 500 nm) at leaf and canopy scales[J]. Remote sensing of environment, 1995, 53(3): 199-211.

[2] 杨德生，黄冠军，李勇，等. 水稻氮高效栽培技术、品种改良和生理机制研究进展[J]. 华中农业大学学报, 2022, 41(1): 62-75. YANG D S, HUANG G J, LI Y, et al. Progress on cultivation technologies, variety improvements and physiological mechanisms of rice with high nitrogen utilization efficiency[J]. Journal of Huazhong Agricultural Universtiy, 2022, 41(1): 62-75 (in Chinese with English abstract).

[3] CURRAN P J. Remote sensing of foliar chemistry[J]. Remote sensing of environment, 1989, 30(3): 271-278.

[4] CILIA C, PANIGADA C, ROSSINI M, et al. Nitrogen status assessment for variable rate fertilization in maize through hyperspectral imagery[J]. Remote sensing, 2014, 6(7): 6549-6565.

[5] DUAN D D, ZHAO C J, LI Z H, et al. Estimating total leaf nitrogen concentration in winter wheat by canopy hyperspectral data and nitrogen vertical distribution[J]. Journal of integrative agriculture, 2019, 18(7): 1562-1570.

[6] LI Z H, JIN X L, WANG J H, et al. Estimating winter wheat (*Triticum aestivum*) LAI and leaf chlorophyll content from canopy reflectance data by integrating agronomic prior knowledge with the PROSAIL model[J]. International journal of remote sensing, 2015, 36(10): 2634-2653.

[7] MOHARANA S, DUTTA S. Spatial variability of chlorophyll and nitrogen content of rice from hyperspectral imagery[J]. ISPRS journal of photogrammetry and remote sensing, 2016, 122: 17-29.

[8] MA L L, CHEN X Y, ZHANG Q, et al. Estimation of nitrogen content based on the hyperspectral vegetation indexes of interannual and multi-temporal in cotton[J/OL]. Agronomy, 2022, 12(6): 1319[2022-07-31]. https://doi.org/10.3390/agronomy12061319.

[9] 楚万林，齐雁冰，常庆瑞，等. 棉花冠层叶片叶绿素含量与高光谱参数的相关性[J]. 西北农林科技大学学报(自然科学版), 2016, 44(9): 65-73. CHU W L, QI Y B, CHANG Q R, et al. Relationship between chlorophyll content and hyperspectral parameters in canopy leaves of cotton[J]. Journal of Northwest A & F University (natural science edition), 2016, 44(9): 65-73 (in Chinese with English abstract).

[10] 王登伟，黄春燕，张伟，等. 高光谱数据与棉花叶绿素含量和叶绿素密度的相关分析[J]. 棉花学报, 2008, 20(5): 368-371. WANG D W, HUANG C Y, ZHANG W, et al. Relationships analysis between cotton chlorophyll content, chlorophyll density and hyperspectral data[J]. Cotton science, 2008, 20(5): 368-371 (in Chinese with English abstract).

[11] CHEN B, LI S K, WANG K R, et al. Study on hyperspectral estimation of pigment contents in leaves of cotton under disease stress[J]. Spectroscopy and spectral analysis, 2010, 30(2): 421-425.

[12] ZHANG J, HE Y, YUAN L, et al. Machine learning-based spectral library for crop classification and status monitoring[J/OL]. Agronomy, 2019, 9(9): 496[2022-07-31]. https://doi.org/10.3390/agronomy9090496.

[13] ZHANG N, YANG G J, PAN Y C, et al. A review of advanced

technologies and development for hyperspectral-based plant disease detection in the past three decades[J/OL].Remote sensing, 2020, 12 (19) : 3188 [2022-07-31]. https://doi. org/10.3390/rs12193188.

[14] LU B, DAO P, LIU J G, et al. Recent advances of hyperspectral imaging technology and applications in agriculture [J/OL]. Remote sensing, 2020, 12 (16) : 2659 [2022-07-31]. https://doi.org/10.3390/rs12162659.

[15] ABULAITI Y, SAWUT M, MAIMAITIAILI B, et al. A possible fractional order derivative and optimized spectral indices for assessing total nitrogen content in cotton[J/OL].Computers and electronics in agriculture, 2020, 171: 105275 [2022-07-31]. https://doi.org/10.1016/j.compag.2020.105275.

[16] LI C C, WANG Y L, MA C Y, et al. Hyperspectral estimation of winter wheat leaf area index based on continuous wavelet transform and fractional order differentiation[J/OL]. Sensors, 2021, 21 (24) : 8497 [2022-07-31]. https://doi. org/10.3390/s21248497.

[17] CAI Q, LI E, PAN J, et al. Retrieval of the canopy chlorophyll density of winter wheat from canopyspectra using continuous wavelet analysis[J].Nature environment & pollution technology, 2019, 18(4): 1211-1218.

[18] YANG Y K, NIE J, KAN Z, et al. Cotton stubble detection based on wavelet decomposition and texture features [J/OL]. Plant methods, 2021, 17 (1) : 113 [2022-07-31]. https://doi. org/10.1186/s13007-021-00809-3.

[19] HUANG J F, ALAN BLACKBURN G. Optimizing predictive models for leaf chlorophyll concentration based on continuous wavelet analysis of hyperspectral data[J].International journal of remote sensing, 2011, 32(24): 9375-9396.

[20] ZHANG J Y, SUN H, GAO D H, et al. Detection of canopy chlorophyll content of corn based on continuous wavelet transform analysis[J/OL].Remote sensing, 2020, 12(17): 2741[2022-07-31].https://doi.org/10.3390/rs12172741.

[21] WIGGINS K, PALMER R, HUTCHINSON W, et al. An investigation into the use of calculating the first derivative of absorbance spectra as a tool for forensic fibre analysis[J]. Science & justice, 2007, 47(1): 9-18.

[22] HANSEN P M, SCHJOERRING J K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression [J]. Remote sensing of environment, 2003, 86 (4) : 542-553.

[23] CLEVERS J G P W. The use of imaging spectrometry for agricultural applications [J]. ISPRS journal of photogrammetry and remote sensing, 1999, 54(5/6): 299-304.

Predicting the content of chlorophyll in cotton using hyperspectral reflectance of leaves

LI Xu[1], CHEN Bailin[2], ZHOU Baoping[1], SHI Ziyan[1], HONG Guojun[1]

1.*College of Information Engineering*, *Tarim University*, *Alar* 843300, *China*;
2.*School of Chemistry and Chemical Engineering*, *Tarim University*, *Alar* 843300, *China*

Abstract With the development of hyperspectral remote sensing technology, hyperspectral prediction of crop growth can provide scientific management for agricultural production, which can improve crop yields and quality while avoiding excessive use of nitrogenous fertilizers. A mathematical model to invert the content of chlorophyll in cotton leaves was developed using continuous wavelet analysis and conventional spectral transformation to decompose and transform the raw leaf spectra of cotton. The characteristic wavelet coefficients and spectral characteristic bands were used as independent variables. Methods including univariate, stepwise regression and partial least squares were used. The results showed that different spectral treatments improved the correlation between the content of chlorophyll and spectral reflectance of cotton leaves. For the conventional spectral transformation, the inverse logarithmic first order differential $\lg(1/R')$ improved the chlorophyll correlation of cotton leaves by 0.41. It is indicated that the continuous wavelet analysis is superior to traditional spectral models in terms of information noise reduction and mining of feature information. The model established has good stability with RPD>2 and good prediction ability for data sampled.

Keywords hyperspectral; non-destructive testing; continuous wavelet analysis; traditional spectral transformation; chlorophyll; cotto n

(责任编辑:陆文昌)