

植冠下土壤类型差异对遥感估算冬小麦叶面积指数的影响

高 林^{1,2*} 王晓菲^{1,2,3} 顾行发⁴ 田庆久^{1,2,3**} 焦俊男^{1,2} 王培燕^{1,2} 李 丹⁴

¹南京大学国际地球系统科学研究所, 南京 210023; ²南京大学江苏省地理信息技术重点实验室, 南京 210023; ³南京大学中国南海研究协同创新中心, 南京 210023; ⁴中国科学院遥感与数字地球研究所, 北京 100101

摘 要 遥感是从田块到区域乃至全球范围无损探测叶面积指数(LAI)的有效方法。土壤背景是LAI遥感研究的重要制约因素之一, 而土壤类型是组成土壤背景的主要部分, 对植被冠层-土壤的光学性质有重要影响, 但目前植冠下土壤类型背景对遥感LAI估算的影响尚不明确。该文通过分析归一化差异植被指数、修正型土壤调节植被指数、修正的叶绿素吸收比率指数、红边拐点、红边振幅、红边面积、红边对数指数和归一化差异光谱指数在不同土壤类型下对LAI的敏感性, 挖掘最不敏感的光谱参数; 通过比较两种回归模型(偏最小二乘回归和随机森林回归)在单一土壤类型和多种土壤类型区对LAI的预测精度, 探究将单一土壤类型下发展的LAI估算模型应用到复杂土壤类型地区时可能出现的问题。结果表明: (1)虽然8种光谱指数对LAI的敏感性因土壤类型不同而差异明显, 但红边拐点受植冠下土壤类型影响最小; “lambda-by-lambda”波段优选算法不仅可以提供对LAI最敏感的光谱区间, 而且可在一定程度上为抵抗植冠下土壤类型差异影响的光谱指数构建提供可行思路; (2)回归模型的LAI预测精度因是否考虑土壤类型而不同, 但在小区域尤其是田块尺度研究时, 对变量的解释能力是选择模型的第一考虑, 而偏最小二乘回归在此方面优于随机森林回归; 在未知地表先验知识的前提下, 随机森林回归对大区域LAI估算比偏最小二乘回归适合, 但地表先验知识的获取对LAI遥感估算仍然十分必要。

关键词 遥感; 土壤类型; 冬小麦叶面积指数; 光谱指数; 偏最小二乘回归; 随机森林回归

引用格式: 高林, 王晓菲, 顾行发, 田庆久, 焦俊男, 王培燕, 李丹 (2017). 植冠下土壤类型差异对遥感估算冬小麦叶面积指数的影响. 植物生态学报, 41, 1273–1288. doi: 10.17521/cjpe.2017.0231

Exploring the influence of soil types underneath the canopy in winter wheat leaf area index remote estimating

GAO Lin^{1,2*}, WANG Xiao-Fei^{1,2,3}, GU Xing-Fa⁴, TIAN Qing-Jiu^{1,2,3**}, JIAO Jun-Nan^{1,2}, WANG Pei-Yan^{1,2}, and LI Dan⁴

¹International Institute for Earth System Science, Nanjing University, Nanjing 210023, China; ²Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing 210023, China; ³Collaborative Innovation Center of South China Sea Studies, Nanjing University, Nanjing 210023, China; and ⁴Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

Abstract

Aims Remote sensing is an effective and nondestructive way to retrieve leaf area index (LAI) from plot, regional and global range. Soil background is one of the confounding factors limiting remotely estimating LAI. And soil type contains a large proportion of soil background information, which can influence the optical properties of vegetation canopy and soil. However, our knowledge on the effects stemmed from soil types underneath the canopy on LAI remote estimating have been in shortage. Thus, this study aims to explore the influences of soil types underneath the canopy on winter wheat LAI remote estimating.

Methods We analyzed the sensitivity variation of eight spectral indices, named normalized difference vegetation index (NDVI), modified soil-adjusted vegetation index (MSAVI), modified chlorophyll absorption ratio index 2 (MCARI₂), red edge inflection point (REIP), red edge amplitude (Dr), red edge area (SDr), red edge symmetry (RES), normalized difference spectral index (NDSI), to LAI in different soil types, and then we identified some spectral intervals or parameters that were insensitive to soil type variations underneath the canopy. We also compared the accuracy of two commonly used regression models, partial least squares regression (PLSR) and random forest regression (RFR), in estimating LAI for different soil types. We also explored the problems arising from

收稿日期Received: 2017-09-06 接受日期Accepted: 2017-11-22

* E-mail: gaol081115@126.com

** 通信作者Author for correspondence (E-mail: tianqj@nju.edu.cn)

applying the regression model developed in single soil type area to complex soil types area in retrieving *LAI*.

Important findings This paper demonstrates the effects of soil types underneath the canopy on *LAI* retrieving. 1) The sensitivity of spectral indices to *LAI* is significantly different due to the soil type variation, but *REIP* has the least effects from soil type variation among the eight spectral indices. Meanwhile, the band selection algorithm of lambda-by-lambda not only chooses the most sensitive spectral interval for *LAI*, but also provides a feasible way to construct the spectral index that exhibits strong resistances to the effects of soil types underneath the canopy. 2) The accuracy of *LAI* estimation by regression models differs under soil type considered or not. So we suggest that in small scale researches, especially in a field scale, the ability of regression models in explaining variables is the priority consideration, while the PLSR is superior to RFR in this respect. Under the premise of unknown priori knowledge of land surfaces, the RFR is more suitable for retrieving *LAI* than PLSR, but land surface priori knowledge is still necessary. These findings provide the theoretical basis and methods for developing remotely sensing estimating *LAI* models adapted to various land surfaces. Further analysis is needed in applying the findings in more crop types, cultivars and growth stages.

Key words remote sensing; soil type; winter wheat leaf area index; spectral indices; partial least squares regression; random forest regression

Citation: Gao L, Wang XF, Gu XF, Tian QJ, Jiao JN, Wang PY, Li D (2017). Exploring the influence of soil types underneath the canopy in winter wheat leaf area index remote estimating. *Chinese Journal of Plant Ecology*, 41, 1273–1288. doi: 10.17521/cjpe.2017.0231

叶面积指数(*LAI*)是表征植物几何结构的关键参数,被定义为单位地表面上绿叶表面积总和的一半(Chen & Black, 1992)。它与光合作用、蒸腾作用等有机物积累相关过程关系密切,也是分析作物长势状况和预测产量的重要依据(Fang *et al.*, 2011; Bolton & Friedl, 2013)。

20世纪70年代,遥感在揭示植物形态特征、物候状态与其光学性质之间的联系方面起到关键作用(Gausman *et al.*, 1969; Woolley, 1971; Allen *et al.*, 1973)。它为星-机-地遥感框架下农艺管理的精细化、科学化发展提供了技术支撑。高通量、非侵入式、广覆盖、实时观测等遥感优势不仅有效地弥补了传统直接测量*LAI*费时费力、具有一定破坏性、仅获得小区域值的缺陷;也使遥感成为快速、高效开展从田块到区域乃至全球范围的农作物*LAI*无损探测的有效方法。过去30年,遥感工作者们沿着以实际应用为主线,提高精度与其紧密咬合的框架不断探索*LAI*估算方法(Thenkabail *et al.*, 2012): (1)关注作物类型差异、种间*LAI*变化差异对模型外推的影响,寻找普适性更强的光谱指数(Viña *et al.*, 2011); (2)侧重于减小叶绿素浓度干扰,尝试发展对*LAI*变化更敏感的新植被指数(Broge & Leblanc, 2001); (3)着眼非参数算法在提高*LAI*预测精度中的作用(Wang *et al.*, 2011); (4)注重宽、窄波段信息以及多源遥感数据对估算*LAI*潜力的挖掘(Lee *et al.*, 2004; Li *et al.*, 2017)。这些研究在一定程度上降低了*LAI*估算模型对作物类型、种植品种、生长阶段以及地理环境的

特殊依赖性,拓宽了*LAI*估算模型的应用范围;同时,也提高了*LAI*的预测精度。但大多研究是在单一土壤类型或未知土壤类型先验知识的前提下开展的,忽略了植冠下土壤类型差异对估算作物*LAI*的影响。

土壤背景对植冠光谱的影响主要涉及两点:一是土壤的表面状况,包括朝向、粗糙度、阴影、结壳、秸秆残留物等;二是影响土壤光谱特征的内在物理和化学性质,包括矿物、有机质、水分、颗粒大小、土壤结构。许多学者分析了土壤湿度对遥感估算*LAI*的影响(Bausch, 1993; Todd & Hoffer, 1998; Yu *et al.*, 2014; Liu *et al.*, 2016);但土壤的光谱特性与土壤的组成有重要关系,当土壤类型发生变化时,常伴有颜色、矿物、有机质、颗粒大小等影响土壤表面反射率理化性质的相应变化(Nolet *et al.*, 2014; Sadeghi *et al.*, 2015);有学者(Rondeaux *et al.*, 1996)就指出相比土壤湿度或粗糙度,土壤类型能造成对植冠光谱更大数量级的影响。再者,不少学者(Rouse *et al.*, 1974; Huete *et al.*, 1985; Huete, 1988; Farrar *et al.*, 1994; Qi *et al.*, 1994; Haboudane *et al.*, 2004)证实近红外-红波段的光谱指数与土壤线关系密切;而土壤类型是造成土壤线变化的主要因素之一(Jasinske & Eagleson, 1989),但能够代表所有土壤类型的土壤线并不存在。因此,若发展一种适应于各种陆表条件的*LAI*遥感估算方法,则探究植冠下土壤类型背景差异对*LAI*遥感估算的影响显得极其必要,而目前还鲜有此方面的研究。

鉴于此,本文从以下方面研究植冠下土壤类型

背景对遥感估算作物 LAI 的影响, 以期为发展一种适应各种陆表环境的 LAI 遥感估算方法和宽覆盖卫星影像下基于土壤背景影响去除模型的 LAI 遥感估算研究提供理论依据和方法途径。

1) 分析现有的一些光谱指数在不同土壤类型区对 LAI 的敏感差异, 进而挖掘对植冠下土壤类型背景影响不敏感的光谱区间或光谱参数。

2) 比较国际上估算 LAI 比较主流的两种回归模型——偏最小二乘回归(PLSR)和随机森林回归(RFR), 在复杂土壤类型区估算 LAI 的适用性, 进而探究单土壤类型区发展的 LAI 估算模型应用到复杂土壤类型区时可能产生的问题。

1 材料和方法

1.1 研究区概况与数据采集

野外试验于2016年4月5日至12日在山东省济宁、曲阜和邹城地区(图1)开展。研究区地处鲁南泰

沂低山丘陵与鲁西南黄淮海平原的交接地带, 属暖温带季风气候, 四季分明, 夏季高温, 雨量充沛, 冬季低温少雨; 全区地势平坦, 土质肥沃, 土壤类型复杂多样(主要土壤类型有潮土、水稻土、棕壤、褐土、砂姜黑土以及粗骨土), 是我国重要的冬小麦(*Triticum aestivum*)种植区。

本研究参考2009年全国第二次土壤普查成果——中国1:100万数字化土壤图, 分别在潮土、水稻土、棕壤、砂姜黑土以及褐土5种土壤类型区选择地势平坦、土壤表面干燥的大田, 设置面积为105 m × 105 m的冬小麦样区, 并沿样区对角线以5 m为间隔设置30个样方(样方直径0.6 m)。5种土壤类型区种植的冬小麦品种均为‘济麦22号’, 播种日期在2015年10月1日前后, 播种间距0.3 m, 前茬作物均为玉米(*Zea mays*)。试验期间冬小麦正值拔节期, 所有样方的土壤表面干燥。

在北京时间11:00–14:00, 天气晴朗无云, 太阳

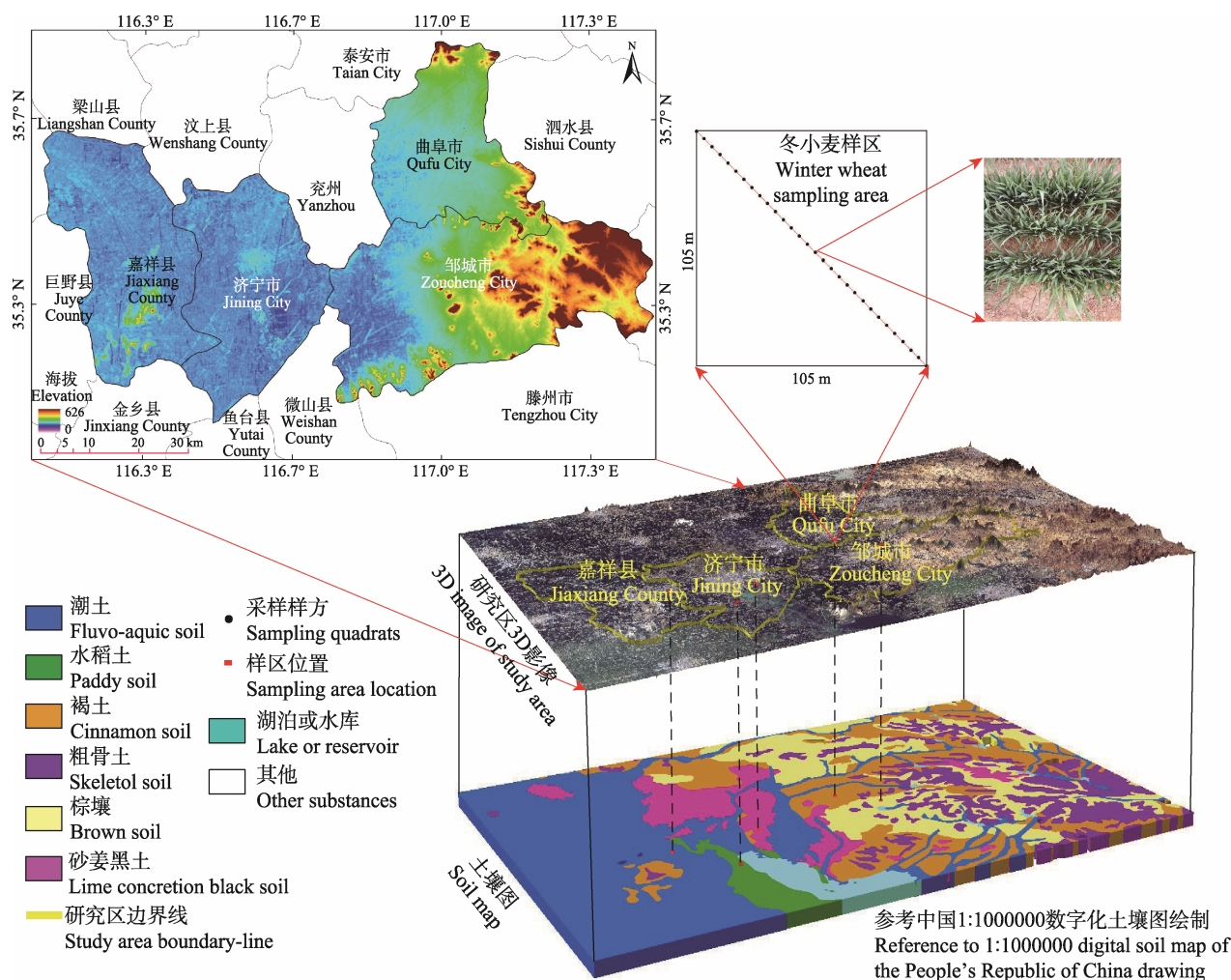


图1 研究区地理位置及土壤类型分布。研究区位于济宁、曲阜和邹城。

Fig. 1 Geographic location of the study area and soil types distribution. The study area is located in Jinan, Qufu and Zoucheng.

doi: 10.17521/cjpe.2017.0231

光强度稳定时,使用ASD FieldSpec Pro2500光谱辐射仪(Analytical Spectral Device, Boulder, USA)采集样方内冬小麦冠层光谱反射率,并且待每一个样方的光谱采集工作结束,随即插上标杆以便下一步进行LAI测量。ASD的光谱采集范围是350–2 500 nm,其中350–1 000 nm范围的光谱采样间隔为1.4 nm;1 000–2 500 nm范围的光谱采样间隔为2 nm (Zhao *et al.*, 2012)。光谱测量过程中,我们始终让ASD光纤探头与冬小麦冠层保持约1.0 m的垂直距离(地面视场范围约0.15 m²);并在每一个样方内随机测量10次,测量前后均通过标准白板进行校正,最后取10次测量的平均值作为该样方的冬小麦冠层光谱反射率(高林等, 2017)。使用相同方法对样区内无植被覆盖的纯裸土表面进行光谱测量,获取到5种土壤类型的光谱反射率。

在光谱测量当天的17:00–18:00,采用LAI-2000植物冠层分析仪(LI-COR, Nebraska, USA)无损测量样方内冬小麦的LAI。LAI-2000利用LAI-2050光学探头内的“鱼眼”镜头拍摄半球视场($\leq 74^\circ$)影像,通过光电二极管监测器探测由中心向外张开的5个观测方向(7° 、 23° 、 38° 、 53° 、 68°)的透射光线,进而利用间隙率模型计算LAI。考虑到LAI-2000的感应波长范围在320–490 nm,野外测量时应尽量避免直射光,在背向太阳一侧,将LAI-2000探头放置于冠层上方,先测定1个天空光,再将其放入样点地面,依次在冠层下不同位置重复测量5个目标值,测量前后始终保持探头水平,取其平均值作为样点冬小麦的LAI。每种土壤类型区均获取30组冬小麦的LAI,统计分析结果(表1)表明:5种土壤类型区的LAI均集中在2–4范围(个别样点> 5),且LAI的变异系数(CV)较低(< 0.5),说明5种土壤类型区冬小麦的LAI变化范围比较接近;再者,除褐土地区以外,潮土、水稻土、棕壤和砂姜黑土地区冬小麦的LAI样本数据均呈正态分布。下文中涉及5种土壤区冬小麦的LAI依次简称为LAI_{潮土}、LAI_{水稻土}、LAI_{棕壤}、LAI_{砂姜黑土}、LAI_{褐土}。

1.2 研究方法

本研究围绕冬小麦冠层光谱、光谱指数、LAI估算模型三方面系统分析植冠下土壤类型差异对遥感估算冬小麦LAI的影响,研究思路如图2所示。首先分析了植冠下土壤类型差异对冬小麦冠层光谱的影响;其次,比较了不同LAI区间光谱指数的变异系数,并分析了现有光谱指数在不同土壤类型区对

表1 冬小麦叶面积指数(LAI)统计分析(5种土壤类型)
Table 1 Descriptive statistics of winter wheat leaf area index (LAI) values (five soil types)

土壤类型 Soil type	LAI					P
	最小值 Min	最大值 Max	平均值 Mean	标准差 SD	变异系数 CV (%)	
潮土 Fluvo-aquic soil	1.970	5.870	3.650	1.031	28.24	0.105
水稻土 Paddy soil	2.240	5.880	4.114	0.962	23.39	0.593
棕壤 Brown soil	1.100	6.710	3.482	1.557	44.71	0.610
砂姜黑土 Lime concretion black soil	1.700	5.500	3.463	1.018	29.39	0.536
褐土 Cinnamon soil	1.320	5.860	4.483	1.235	27.54	< 0.005
5种土壤 Five type soils	1.100	6.710	3.838	1.232	32.09	0.242

p是基于Anderson-Darling的正态检验值。 $p > 0.05$ 说明数据服从正态分布。
SD, standard deviation; CV, coefficient of variation. p is normality test value using Anderson-Darling. $p > 0.05$ means that data obey a normal distribution.

LAI的敏感性;最后讨论不同土壤类型下两种回归模型对LAI的估算潜力。

1.2.1 光谱指数

光谱指数是利用代数运算(加、减、乘、除及其组合)将植被-土壤系统中不同波段反射率因子组合而成的(徐希孺, 2005),用以强调植物生理活性的一种参数。参考已有文献,本研究选取8种在LAI遥感估算方面较为常用且效果良好的光谱指数(表2)。

1.2.2 偏最小二乘回归(PLSR)

PLSR法是一种集多元线性回归分析、主成分分析和典型相关分析于一体的,基于自变量(X)的潜变量关于因变量(Y)的潜变量的线性回归模型,间接描述自变量与因变量关系的回归建模方法(王惠文等, 2006)。PLSR最初发展是为了提高光谱检测化学成分的精度的(Wold *et al.*, 2001),由于其有效地解决了光谱集合内部高度线性相关问题,从理论上降低了光谱变量之间的相关性对估测模型的影响,同时还构建了一个具有预测功能的“响应”矩阵,使得PLSR成为高光谱研究中处理自变量多重共线性问题以及变量个数大于样本量问题的重要工具(Abdi, 2003; Atzberger *et al.*, 2010)。本研究使用Minitab (<http://www.minitab.com>)进行PLSR分析。PLSR主要由3步构成: 1)所有自变量X和因变量Y的分解(公式1、2); 2) T和U的关系方程(公式3); 3)将公式3代入公式2得到关于因变量Y的新预测方程(公式4)

表2 本文估算冬小麦叶面积指数的光谱指数

Table 2 Spectral indices for winter wheat leaf area index estimating in this paper

光谱指数 Spectral indices	公式 Equation	来源 Reference
归一化差异植被指数 Normalized difference vegetation index (NDVI)	$NDVI = \frac{R_{800} - R_{680}}{R_{800} + R_{680}}$	Rouse <i>et al.</i> , 1974
修正型土壤调节植被指数 Modified soil-adjusted vegetation index (MSAVI)	$MSAVI = \frac{1}{2} \times \left[2 \times R_{800} + 1 - \sqrt{(2 \times R_{800} + 1)^2 - 8 \times (R_{800} - R_{670})} \right]$	Qi <i>et al.</i> , 1994
修正的叶绿素吸收比率指数 Modified chlorophyll absorption ratio index 2 (MCARI ₂)	$MCARI_2 = \frac{1.5 \times [2.5 \times (R_{800} - R_{670}) - 1.3 \times (R_{800} - R_{550})]}{\sqrt{(2 \times R_{800} + 1)^2 - (6 \times R_{800} - 5 \times \sqrt{R_{670}})} - 0.5}$	Haboudane <i>et al.</i> , 2004
红边拐点 Red edge inflection point (REIP)	$REIP = 700 + 40 \times \frac{(R_{670} + R_{780})/2 - R_{700}}{R_{740} - R_{700}}$	Danson & Plummer, 1995
红边振幅 Red edge amplitude (Dr)	$d\lambda_{red-edge}$	Feng <i>et al.</i> , 2009
红边面积 Red edge area (SDr)	$\sum d\lambda_{(680-750)}$	Filella & Penuelas, 1994
红边对称指数 Red edge symmetry (RES)	$RES = \frac{R_{718} - R_{675}}{R_{755} - R_{675}}$	Ju <i>et al.</i> , 2010
归一化差异光谱指数 Normalized difference spectral index (NDSI)	$NDSI_{(i,j)} = \frac{R_i - R_j}{R_i + R_j}$	Li <i>et al.</i> , 2013

R 是光谱反射率。NDSI采用“lambda-by-lambda”波段优选算法计算(Darmawan *et al.*, 2009; Elsayed *et al.*, 2015)。 i, j 是一定波段范围内任意2个波段($i < j$)。 λ 是波长。

R is spectral reflectance. NDSI calculated by band optimized algorithm of “lambda-by-lambda” (Darmawan *et al.*, 2009; Elsayed *et al.*, 2015). i and j are arbitrary bands in a certain wavelength range ($i < j$). λ is wavelength.

(Yeniay & Goktas, 2002):

$$X = TP^T + E \quad (1)$$

$$Y = UQ^T + F \quad (2)$$

$$U = bT + e \quad (3)$$

$$Y = bTQ + G \quad (4)$$

其中, T 、 U 分别是 X 和 Y 的主成分矩阵(score矩阵); P 、 Q 分别是 X 和 Y 的回归系数矩阵(loading矩阵); E 、 F 分别是 X 和 Y 的残差矩阵; P 、 Q 的上标T表示矩阵转置运算; b 是方程系数; e 是 T 和 U 的残差矩阵; G 是预测 Y 的新残差矩阵。目前, PLSR在高光谱遥感探测植物物理化学参数方面得到广泛使用(Hansen & Schjoerring, 2003; Cho *et al.*, 2007; Herrmann *et al.*, 2011; Li *et al.*, 2014)。

1.2.3 随机森林回归(RFR)

随机森林(Breiman, 2001)是基于决策树发展的一种将随机且有放回抽样生成的 n 棵树集成于一体进行分类和回归学习的算法; 森林中每棵CART树独立且不经修剪地生长至最大。

每次随机且有放回的抽样过程中, 大约有36.8%的样本未用于生成树, 这些样本称为“外袋”(OOB)数据。就分类问题而言, 预测“外袋”数据的响应值是通过多数投票方式完成的, 而回归问题则基于对所有树的平均得到。随机森林回归模型中, 3个参数需要根据“外袋”预测误差(the OOB estimate of

error rate)进行优化: 1)构成森林的树个数 n_{tree} (默认值是500); 2)树节点二叉分裂的参与变量个数 m_{try} (默认值是 $p/3$); 在随机森林回归中, 参与树节点的二叉最优分裂的子变量集是根据节点不纯度最小原则(the input variable with the lowest Gini index)从全部变量 p 中随机选取 m ($m \ll p$)个子变量构成的; 3)树末端节点的最小尺寸($nodesize$, 默认值是5), 树生长过程中, 节点拥有的尺寸若小于末端节点最小尺寸则不再进行二叉分裂。

随机森林回归在R语言(<https://www.r-project.org/>)环境下执行。

1.2.4 LAI反演精度评价

本文选择决定系数(R^2)、均方根误差(RMSE)和相对分析误差(RPD)(Rossel & Webster, 2012; Siegmann & Jarmer, 2015)作为评价LAI估测模型鲁棒性的指标。 R^2 用于评价模型拟合程度, 其取值范围在0–1之间, R^2 越接近1, 模型拟合精度越高; RMSE用于评价预测数据与实测数据之间的差异, RMSE越小, 两者差异越小, 模型的预测精度也相对越高; RPD用于评价模型鲁棒性, RPD越大, 模型鲁棒性越高。此外, PLSR采用留一法交叉验证(leave-one-out cross validation), 即每次用 $n - 1$ 个样本进行训练, 留下1个样本对训练模型进行检验, 获得该样本的估测值, 如此循环 n 次, 共建立 n 个训练模型, 得到

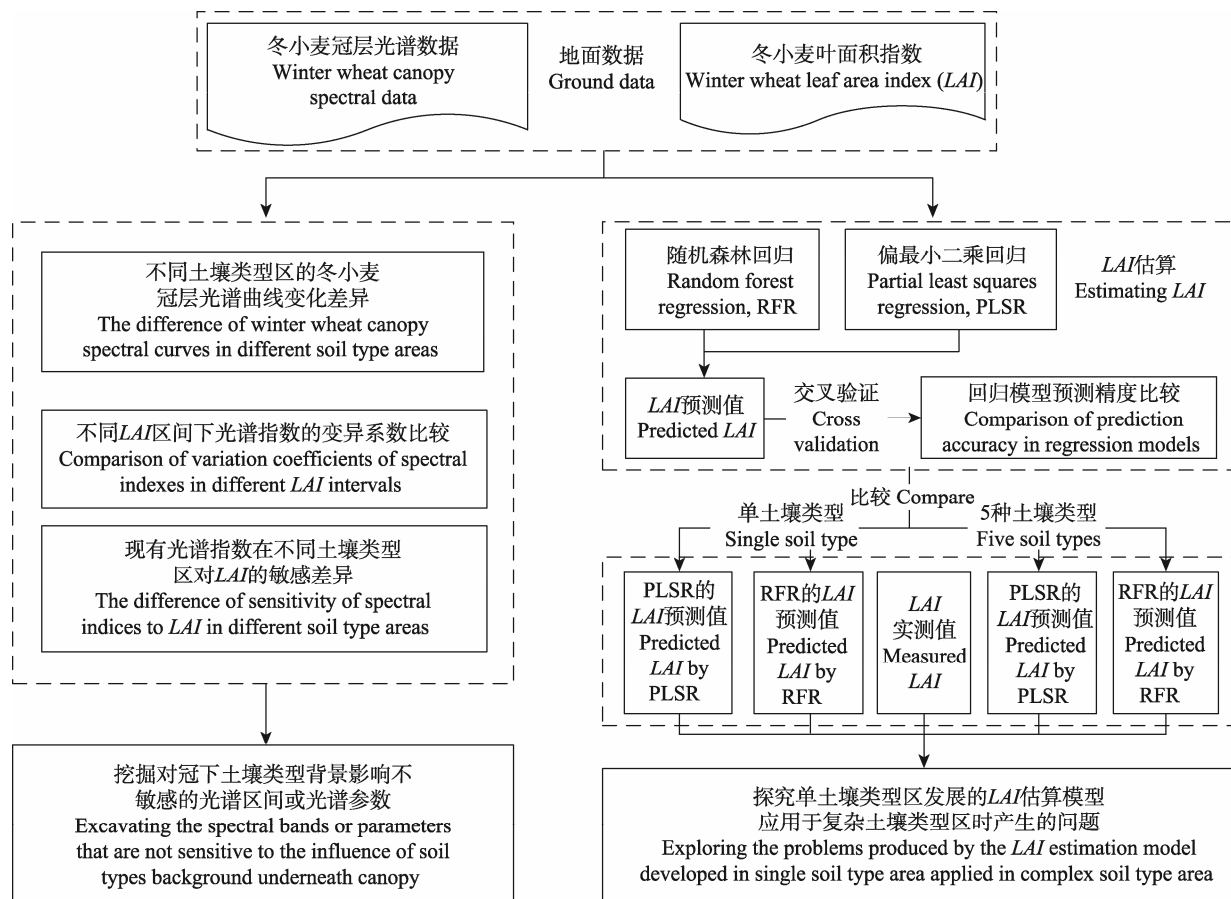


图2 研究路线图。

Fig. 2 Flowchart of research in this study.

每一个参与验证的样本的估测值(McLachlan, 2004)。

2 结果和讨论

2.1 光谱特征分析

图3是剔除存在较大反射率波动噪声(噪声存在于1 350–1 419 nm、1 796–1 950 nm和2 450–2 500 nm光谱范围)后的土壤光谱反射率曲线。从图中可以发现: 在350–2 500 nm范围, 褐土的反射率(ρ_C)最高, 砂姜黑土的反射率(ρ_L)最低; 在550–1 950 nm范围, 棕壤的反射率(ρ_B)高于潮土的反射率(ρ_F)和水稻土的反射率(ρ_P), 其中550–750 nm范围, ρ_F 和 ρ_P 比较接近, 1 950–2 450 nm范围, $\rho_F > \rho_B > \rho_P$ 。土壤线方面, 不同土壤类型的红波段反射率 $\rho(\lambda_{Red})$ 和近红外波段反射率 $\rho(\lambda_{NIR})$ 的线性差异显著: 土壤线斜率, 水稻土>潮土>砂姜黑土>棕壤>褐土; 褐土的土壤线截距最大, 棕壤次之, 潮土和砂姜黑土接近, 水稻土最小。

图4描述了5种土壤类型区的冬小麦冠层平均光

谱曲线变化。从图中可以发现: 在550 nm“绿峰”附近, 冬小麦“绿峰”的高度是潮土>棕壤>水稻土, 砂姜黑土和褐土的冬小麦“绿峰”的高度接近且低于水稻土; 800–900 nm“近红外反射平台”附近, 冬小麦冠层“近红外反射平台”的反射率高度为棕壤>褐土>砂姜黑土>水稻土>潮土; 680–750 nm红边范围, 5种土壤类型区的冬小麦红边有差异但不明显。

上述分析表明在样区内冬小麦没有封垄时, 由于冬小麦冠层光谱掺杂了植冠下土壤的某些光学特征, 导致不同土壤类型区的3个冬小麦冠层光谱特征(“绿峰”、“红边”和“近红外反射平台”)发生不同程度的变化, 其中土壤类型背景对近红外波段光谱的影响最突出; 但对红边光谱影响最不明显, 主要原因在于绿色健康的冬小麦红光吸收位置与近红外“肩部”连线的斜率随LAI增大会发生微小变化, 但这种微小变化在680–750 nm范围反射率由低值急剧变化到高值的趋势中不足以使不同土壤类型区的绿色健康的冬小麦红边光谱曲线发生较大位移。

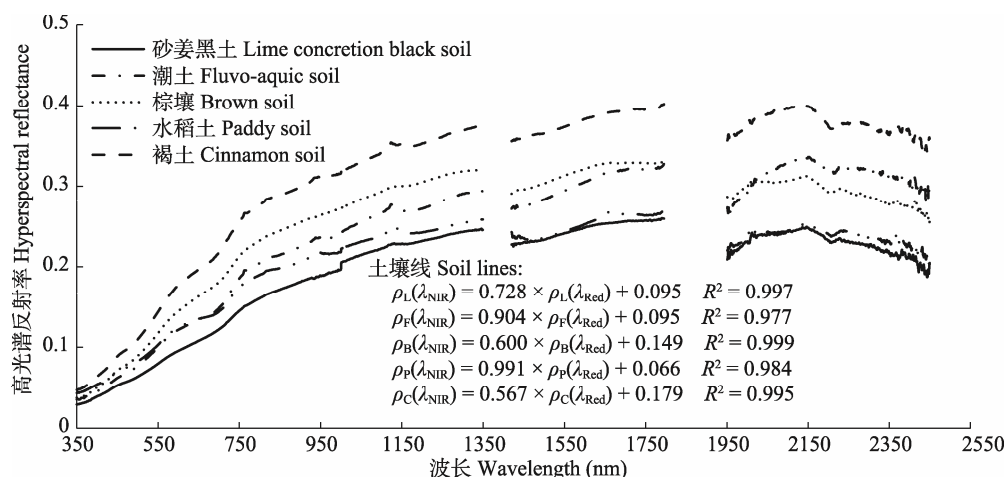


图3 不同类型土壤的光谱反射率曲线和土壤线。 $\rho_L(\lambda_{NIR})$ 和 $\rho_L(\lambda_{Red})$ 分别表示砂姜黑土的近红外波段反射率和红波段反射率; $\rho_F(\lambda_{NIR})$ 和 $\rho_F(\lambda_{Red})$ 分别表示潮土的近红外波段反射率和红波段反射率; $\rho_B(\lambda_{NIR})$ 和 $\rho_B(\lambda_{Red})$ 分别表示棕壤的近红外波段反射率和红波段反射率; $\rho_P(\lambda_{NIR})$ 和 $\rho_P(\lambda_{Red})$ 分别表示水稻土的近红外波段反射率和红波段反射率; $\rho_C(\lambda_{NIR})$ 和 $\rho_C(\lambda_{Red})$ 分别表示褐土的近红外波段反射率和红波段反射率。

Fig. 3 Different soil spectral reflectance curves and soil lines. $\rho_L(\lambda_{NIR})$ and $\rho_L(\lambda_{Red})$ are spectral reflectance of lime concretion back soil in near infrared band and red band, respectively; $\rho_F(\lambda_{NIR})$ and $\rho_F(\lambda_{Red})$ are spectral reflectance of fluvo-aquic soil in near infrared band and red band, respectively; $\rho_B(\lambda_{NIR})$ and $\rho_B(\lambda_{Red})$ are spectral reflectance of brown soil in near infrared band and red band, respectively; $\rho_P(\lambda_{NIR})$ and $\rho_P(\lambda_{Red})$ are spectral reflectance of paddy soil in near infrared band and red band, respectively; $\rho_C(\lambda_{NIR})$ and $\rho_C(\lambda_{Red})$ are spectral reflectance of cinnamon soil in near infrared band and red band, respectively.

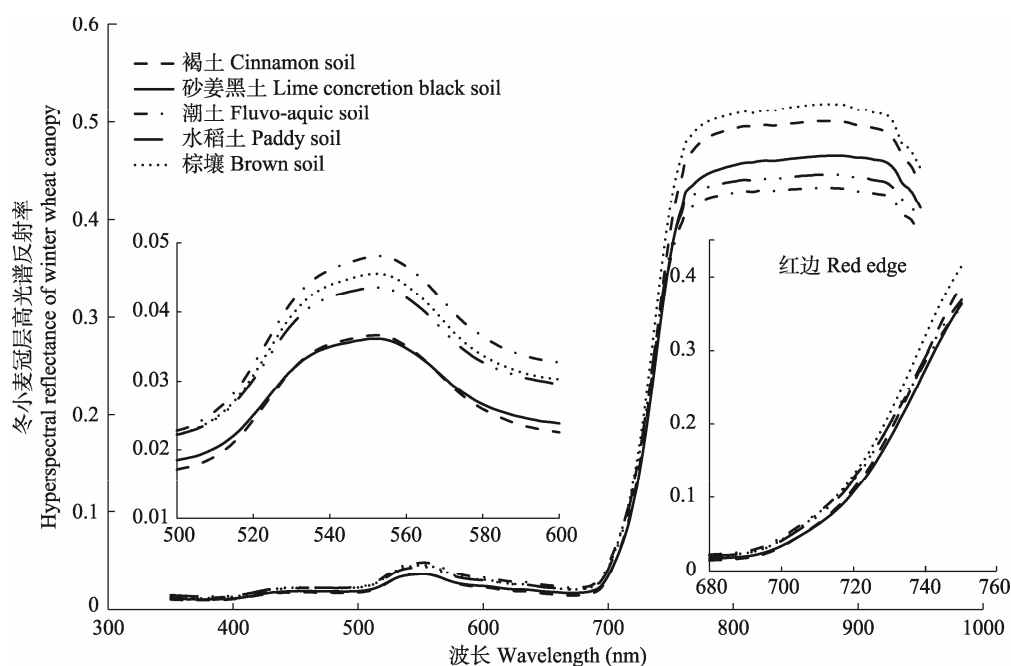


图4 5种土壤类型的冬小麦冠层光谱反射率比较。

Fig. 4 Comparison of winter wheat canopy spectral reflectance in five soil types.

2.2 光谱指数与冬小麦LAI的相关关系探讨

表3是将150个冬小麦LAI样本及其对应的光谱指数按LAI为1.0–1.5、1.5–2.0、2.0–2.5、2.5–3.0、3.0–3.5、3.5–4.0、4.0–4.5、4.5–5.0、>5.0依次划分9组,并分别计算每组光谱指数的变异系数;结果表明8个光谱指数中只有NDVI和REIP具有较低的变异系数(CV_{NDVI} 、 CV_{REIP});且当LAI < 3时, CV_{NDVI} 具有高

值,说明在低冠层覆盖下,植被冠层光谱容易混入土壤光谱,导致不同土壤类型区的NDVI变化特别强烈;当LAI > 3时, CV_{NDVI} 逐渐降低,且在LAI > 4以后, CV_{NDVI} 始终保持低值,说明高冠层覆盖下的组内NDVI值变化不大。相反, CV_{REIP} 随LAI取值区间的变化是8种光谱指数中最小的,始终在0.36上下浮动,且 CV_{REIP} 随LAI的递增出现先增大再减小最后不变

的变化趋势,说明相比其他7种光谱指数,植冠下土壤类型背景的改变对红边位置的影响较小。

表4描述了8种光谱指数与冬小麦LAI的相关性。如表所示,虽然NDVI是表征植被光合作用及其生长发育动态变化的最佳指示因子(Baret & Guyot, 1991; Myneni *et al.*, 1995),但NDVI-LAI的相关性因植冠下裸土成分差异而表现不同($r_{\text{水稻土}} > r_{\text{潮土}} > r_{\text{褐土}} > r_{\text{砂姜黑土}} > r_{\text{棕壤}}$)。尽管MSAVI和MCARI₂作为NDVI的改进形式,在不考虑土壤类型时提高了与LAI的相关性,但土壤类型背景影响仍然十分突出,不同土壤类型区MSAVI-LAI和MCARI₂-LAI的相关性均为 $r_{\text{水稻土}} > r_{\text{潮土}} > r_{\text{褐土}} > r_{\text{砂姜黑土}} > r_{\text{棕壤}}$;这说明基于一种土壤线假设改进的光谱指数在小区域研究中似乎合理,但在复杂土壤类型地区使用基于一种土壤线发展的光谱指数分析植被必然会产生较大误差。同样,不同植冠下土壤类型的变化也会引起红边反射特征指数与LAI的相关性差异:REIP与LAI的相关性为 $r_{\text{水稻土}} > r_{\text{潮土}} >$

$r_{\text{棕壤}} > r_{\text{褐土}} > r_{\text{砂姜黑土}}$; RES与5种土壤类型的冬小麦LAI均呈负相关关系,且 $r_{\text{水稻土}} > r_{\text{潮土}} > r_{\text{褐土}} > r_{\text{棕壤}} > r_{\text{砂姜黑土}}$; Dr与LAI的相关性为 $r_{\text{潮土}} > r_{\text{水稻土}} > r_{\text{砂姜黑土}} > r_{\text{褐土}} > r_{\text{棕壤}}$; SDr与LAI的相关性(无论正相关还是负相关,均以绝对值作比较)为 $r_{\text{潮土}} > r_{\text{水稻土}} > r_{\text{砂姜黑土}} > r_{\text{棕壤}} > r_{\text{褐土}}$ 。NDSI与LAI相关性是8种光谱指数中最高的,这是由于“lambda-by-lambda”波段优选算法构建NDSI时不需要如NDVI、MSAVI和MCARI₂那般依赖固定波段(680 nm、800 nm),也不需要如红边反射特征指数那般受限于特定的光谱区域(680–750 nm红边范围),它可以使用350–950 nm范围内的任意光谱波段进行两两组合,通过计算归一化差异光谱指数与LAI的相关性等值线图,挑选出与LAI有最高相关性的NDSI_(i,j)。然而,不同土壤类型区NDSI与LAI的相关性仍存在差异($r_{\text{水稻土}} > r_{\text{潮土}} > r_{\text{棕壤}} > r_{\text{砂姜黑土}} > r_{\text{褐土}}$),且不同土壤类型区优选的光谱波段组合(i,j)也不同;这说明即使采用这种波段优选算法也不能忽视土壤类

表3 不同叶面积指数(LAI)区间下光谱指数的变异系数(%)

Table 3 Coefficient of variation (%) of spectral indices in different leaf area index (LAI) intervals

光谱指数 Spectral indices	LAI								
	1–1.5	1.5–2	2–2.5	2.5–3	3–3.5	3.5–4	4–4.5	4.5–5	>5
NDVI	6.32	10.91	7.18	7.94	4.37	1.95	1.9	2.01	1.44
MSAVI	13.86	24.48	18.57	24.42	12.77	8.72	10.01	8.02	4.25
MCARI ₂	15.72	25.98	20.51	26.03	13.95	8.94	9.97	8.28	4.84
RES	20.54	25.8	24.06	18.15	16.26	8.36	8.49	11.69	7.8
REIP	0.44	0.45	0.53	0.39	0.41	0.23	0.25	0.29	0.24
NDSI	100.97	206.38	187.41	139.56	154.61	155.92	114.89	158.33	150.22
Dr	19.51	28.93	20.25	31.48	15.98	15.68	12.38	13.61	7.38
SDr	17.01	23.61	14.44	27.18	12.2	16.32	11.07	11.35	6.68

光谱指数同表2。

Spectral indices see Table 2.

表4 光谱指数与不同土壤类型背景的冬小麦叶面积指数(LAI)相关性

Table 4 Correlations between spectral indices and winter wheat leaf area index (LAI) in different soil types

光谱指数 Spectral indices	不同土壤类型区的冬小麦LAI相关系数 Winter wheat LAI in different soil types/Correlation coefficient (r)					
	水稻土 Paddy soil (n = 30)	潮土 Fluvo-aquic soil (n = 30)	棕壤 Brown soil (n = 30)	砂姜黑土 Lime concretion black soil (n = 30)	褐土 Cinnamon soil (n = 30)	5种土壤 Five type soils (n = 150)
NDVI	0.789	0.655	0.079	0.145	0.630	0.379
MSAVI	0.753	0.724	0.037	0.283	0.313	0.383
MCARI ₂	0.769	0.722	–0.008	0.258	0.401	0.386
NDSI _(i,j)	0.854/NDSI _(727, 932)	0.852/NDSI _(416, 429)	0.788/NDSI _(781, 783)	0.701/NDSI _(669, 676)	0.679/NDSI _(629, 636)	0.561/NDSI _(628, 636)
REIP	0.836	0.735	0.500	0.208	0.426	0.408
Dr	0.714	0.732	0.016	0.375	0.186	0.340
SDr	0.624	0.704	–0.086	0.371	0.008	0.265
RES	–0.834	–0.675	–0.368	–0.176	–0.477	–0.404

光谱指数同表2。

Spectral indices see Table 2.

型背景影响,但值得欣慰的是该算法可以在一定程度上缓解植冠下土壤类型差异对 $NDSI$ 与 LAI 相关性的影响。

基于上述分析发现:不同土壤类型区,同一种光谱指数与冬小麦 LAI 的相关性不同;同一种土壤类型区,不同光谱指数对冬小麦 LAI 的敏感性也迥异;且波段位置的选择对减弱植冠下土壤类型背景影响极其重要。

2.3 PLSR模型精度分析

37种变量组合($NDSI$ 和800 nm固定波段反射率计算的 $NDVI$ 、 $MSAVI$ 、 $MCARI_2$ 随机且不重复组合, $NDSI$ 和红边特征参数 RES 、 $REIP$ 、 Dr 、 SDr 随机且不重复组合)的PLSR模型交叉验证结果如图5所示。整体上,水稻土地区的PLSR模型精度(简称 $PLSR_{\text{水稻土}}$)较高,且 $PLSR_{\text{水稻土}}$ 不因输入变量组的不同而有较大差异, R^2 在0.65上下波动, $RMSE_{CV}$ 在0.56上下变化。潮土、棕壤、砂姜黑土和褐土地区的PLSR模型精度(简称 $PLSR_{\text{潮土}}$ 、 $PLSR_{\text{棕壤}}$ 、 $PLSR_{\text{砂姜黑土}}$ 、 $PLSR_{\text{褐土}}$)在 $NDSI$ 参与变量组合前后有较大变化。

基于对模型过度拟合和优化PLSR模型输入变量个数的考虑,选择具有最小 $RMSE_{CV}$ 和最大的 R^2 的PLSR模型作为估算冬小麦 LAI 的最优线性模型(表5)。当考虑土壤类型时,PLSR模型提高了单独使用红边反射特征指数估算冬小麦 LAI 的精度,由 $NDSI$ 和红边反射特征指数构成的4种变量组合,实现对水稻土、潮土、棕壤和砂姜黑土地冬小麦 LAI 的理想估算:其中 $PLSR_{RES+NDSI}$ 解释了水稻土地区68.4%的冬小麦 LAI 变量($RMSE_{CV} = 0.533$), $PLSR_{RES+REIP+NDSI}$ 解释了潮土地区74.3%的冬小麦 LAI 变量($RMSE_{CV} = 0.518$), $PLSR_{REIP+Dr+SDr+NDSI}$ 解释了棕壤地区77.1%的冬小麦 LAI 变量($RMSE_{CV} = 0.735$), $PLSR_{REIP+NDSI}$ 解释了砂姜黑土地地区49%的

冬小麦 LAI 变量($RMSE_{CV} = 0.721$);由 $NDVI$ 和 $NDSI$ 构建的PLSR仅解释了褐土地区38.5%的冬小麦 LAI 变量($RMSE_{CV} = 0.954$);同时,棕壤地区的 $PLSR_{REIP+Dr+SDr+NDSI} > \text{潮土地区的} PLSR_{RES+REIP+NDSI} > \text{水稻土地区的} PLSR_{RES+NDSI} > \text{砂姜黑土地地区的} PLSR_{REIP+NDSI} > \text{褐土地区的} PLSR_{NDVI+NDSI}$ 。当不考虑土壤类型因素时,5个地区150个样本的分析结果表明,由 RES 、 $REIP$ 、 Dr 和 $NDSI$ 组合的PLSR精度最高,解释了整个研究区34.1%的冬小麦 LAI 变量($RMSE_{CV} = 0.997$)。然而,相比任何单一土壤类型区的冬小麦 LAI 估算模型精度,不考虑土壤类型因素时的 $PLSR_{RES+REIP+Dr+NDSI}$ 精度较低。

值得注意的是,虽然红边信息对 LAI 的敏感性是一个复杂问题,长久以来都没有一致性的结论,多数学者也更倾向于指出红边反射特征指数在描述(叶片或冠层)叶绿素或氮素含量(Schlemmer *et al.*, 2013)、病虫害胁迫(Vogelmann *et al.*, 1993)等方面的研究价值,但从PLSR模型精度分析结果和不同 LAI 区间下光谱指数的变异系数比较(表3)中不难得出与已有研究(Horler *et al.*, 1983; Delegido *et al.*, 2013; 高林等, 2017)相似的结论,即红边信息依然比近红外波段更适宜预测 LAI ,以及红边特征光谱指数(如 $REIP$)可认为是近红外-红波段光谱指数预测 LAI 的理想替代。

2.4 RFR模型精度分析

RFR模型是根据 n_{tree} 与 m_{try} 的不同组合构建而成,其中 n_{tree} 从1 000–6 000范围以1 000为间隔选取, m_{try} 从1–8范围以1为间隔选取。研究比较了RFR估算的 LAI 与实测 LAI 的拟合精度(R^2 和 $RMSE$),确定出最优 n_{tree} 和 m_{try} (表6),并分析了8种光谱指数在最优RFR估算冬小麦 LAI 中的重要性(图6)。

由表6可知不同土壤类型区的优选 n_{tree} 和 m_{try} 有

表5 优选的偏最小二乘回归(PLSR)模型估测冬小麦叶面积指数(LAI)结果

Table 5 Results of estimating winter wheat leaf area index (LAI) by optimal partial least squares regression (PLSR) models

土壤类型 Soil type	PLSR模型 PLSR models	潜变量个数 Number of latent variables	建模 Calibration		交叉验证 Cross validation		
			R^2	$RMSE_C$	R^2	$RMSE_{CV}$	RPD_{CV}
水稻土 Paddy soil ($n = 30$)	$RES + NDSI$	2	0.733	0.489	0.684	0.533	1.805
潮土 Fluvo-aquic soil ($n = 30$)	$RES + REIP + NDSI$	3	0.803	0.450	0.743	0.518	1.990
棕壤 Brown soil ($n = 30$)	$REIP + Dr + SDr + NDSI$	4	0.837	0.617	0.771	0.735	2.118
砂姜黑土 Lime concretion black soil ($n = 30$)	$REIP + NDSI$	2	0.579	0.649	0.490	0.721	1.412
褐土 Cinnamon soil ($n = 30$)	$NDVI + NDSI$	1	0.466	0.887	0.385	0.954	1.295
5种土壤 Five type soils ($n = 150$)	$RES + REIP + Dr + NDSI$	4	0.378	0.968	0.341	0.997	1.236

$NDVI$ 、 $MSAVI$ 、 $MCARI_2$ 、 RES 、 $REIP$ 、 Dr 、 SDr 、 $NDSI$ 同表2。

$NDVI$ 、 $MSAVI$ 、 $MCARI_2$ 、 RES 、 $REIP$ 、 Dr 、 SDr 、 $NDSI$ see Table 2.

表6 优选的随机森林回归(RFR)模型估测冬小麦叶面积指数(LAI)结果
Table 6 Results of estimating winter wheat leaf area index (LAI) by optimal random forest regression (RFR) models

土壤类型 Soil type	随机森林回归模型的优化参数 The optimized parameters in RFR		模型精度 Model precision			
	n_{tree}	m_{try}	变量解释 Variable explained (%)	R^2	RMSE	RPD
水稻土 Paddy soil ($n = 30$)	1 000	2	60.08	0.604	0.598	1.609
潮土 Fluvo-aquic soil ($n = 30$)	1 000	8	72.86	0.733	0.528	1.953
棕壤 Brown soil ($n = 30$)	2 000	6	43.07	0.441	1.155	1.348
砂姜黑土 Lime concretion black soil ($n = 30$)	1 000	6	48.65	0.488	0.717	1.420
褐土 Cinnamon soil ($n = 30$)	1 000	3	24.58	0.258	1.054	1.172
5种土壤 Five type soils ($n = 150$)	2 000	4	42.61	0.427	0.930	1.325

RMSE, 均方根误差; RPD, 相对分析误差。
RMSE, root mean square error; RPD, relative percent deviation.

差异, 且最优RFR估算LAI的精度也不同: 水稻土地区, 由 $n_{tree} = 1\ 000$, $m_{try} = 2$ 组成的RFR精度最高, $R^2 = 0.604$, $RMSE = 0.598$; 潮土地区, 由 $n_{tree} = 1\ 000$, $m_{try} = 8$ 组成的RFR精度最高, $R^2 = 0.733$, $RMSE = 0.528$; 棕壤地区, 由 $n_{tree} = 2\ 000$, $m_{try} = 6$ 组成的RFR精度最高, $R^2 = 0.441$, $RMSE = 1.155$; 砂姜黑土地区, 由 $n_{tree} = 1\ 000$, $m_{try} = 6$ 组成的RFR精度最高, $R^2 = 0.488$, $RMSE = 0.717$; 褐土地区, 由 $n_{tree} = 1\ 000$, $m_{try} = 3$ 组成的RFR精度最高, $R^2 = 0.258$, $RMSE = 1.054$; 且 $RFR_{潮土} > RFR_{水稻土} > RFR_{砂姜黑土} > RFR_{棕壤} > RFR_{褐土}$ 。然而, 当不考虑土壤类型时, 由 $n_{tree} = 2\ 000$, $m_{try} = 4$ 组成的RFR估算冬小麦LAI精度最高, $R^2 = 0.427$, $RMSE = 0.930$; 但相比于考虑土壤类型(水稻土、潮土和砂姜黑土)时优选的RFR, 不考虑土壤类型时优选的RFR估算冬小麦LAI的精度明显较低。图6显示了8种光谱指数在最优RFR模型中估算冬小麦LAI的重要性: 整体上, NDSI对5种土壤类型区的冬小麦LAI估算最重要; NDVI、MSAVI、MCARI₂、RES、REIP、SDr在RFR估算LAI中发挥了不同程度的作用, 且它们之间的重要性差异不大(均小于10)。

2.5 PLSR与RFR模型精度比较

图7描述了PLSR和RFR两种算法最优模型的冬小麦LAI预测精度。当考虑土壤类型背景时, PLSR最优模型对冬小麦LAI的预测精度整体优于RFR, 特别是在棕壤地区, $PLSR_{REIP+Dr+SDr+NDSI}$ 的冬小麦LAI预测值与LAI实测值的拟合度远大于 $RFR_{(n_{tree}=2000, m_{try}=6)}$ (R^2 高0.33, $RMSE$ 低0.42, RPD 高0.77); 当不考虑土壤类型背景时, $RFR_{(n_{tree}=2000, m_{try}=4)}$ 对冬小麦LAI的解释能力略高于 $PLSR_{RES+REIP+Dr+NDSI}$ (R^2 高0.086, $RMSE$ 低0.067, RPD 高0.089)。

由于考虑土壤类型背景(30个样本)和不考虑土

壤类型背景(150个样本)直接反映样本数量的变化, 而随机森林采用Bagging思想, 即通过平均若干个弱分类器的回归结果, 选择出一个强分类器, 从而提高模型预测的准确性, 使得RFR对样本数量的依赖性极强, 这在一定程度上造成分析少量样本时(即考虑土壤类型背景), RFR优选模型对LAI的解释能力弱于PLSR; 分析大量样本时(即不考虑土壤类型背景), RFR算法估算LAI的能力得到极大改善。因此, Yuan等(2017)曾指出开展大样本数据集(如: 多或全生育期、作物品种或样本区域面积相当大)的LAI预测, 随机森林模型更理想; 但是, Bagging集成学习思想在提高预测精度同时, 也降低了RFR对变量的解释性, 虽然我们能够得知各变量在模型中的重要性(图6), 但对具体哪些参与树节点二叉最优分裂的变量对模型预测起到作用是未知的。相比RFR, 最小二乘法在回归迭代每个输入变量过程中, 不仅将自变量映射为主成分, 确定其系数, 也通过组合那些与因变量相关性大的主成分构建PLSR模型, 因此, PLSR对变量的解释可靠, 且整体性强; 但是, 如果没有红边反射特征指数组合的PLSR或红边反射特征指数与NDSI组合的PLSR对LAI具有理想的解释能力等类似先验知识, 那么从众多变量组合中筛选出最优PLSR所花费的时间和精力远多于优化RFR所花费的时间。

为了进一步探讨土壤类型对回归模型预测LAI的影响, 我们将不考虑土壤类型时预测的冬小麦LAI按已知的5种土壤类型区划分, 将划分的LAI预测值与考虑土壤类型时的LAI预测值以及LAI实测值进行对比(表7; 图8), 结果表明: 5个土壤类型区, 不考虑土壤类型背景时预测的LAI与实测LAI的拟合精度整体低于考虑土壤类型背景时; 不同土壤类型区,

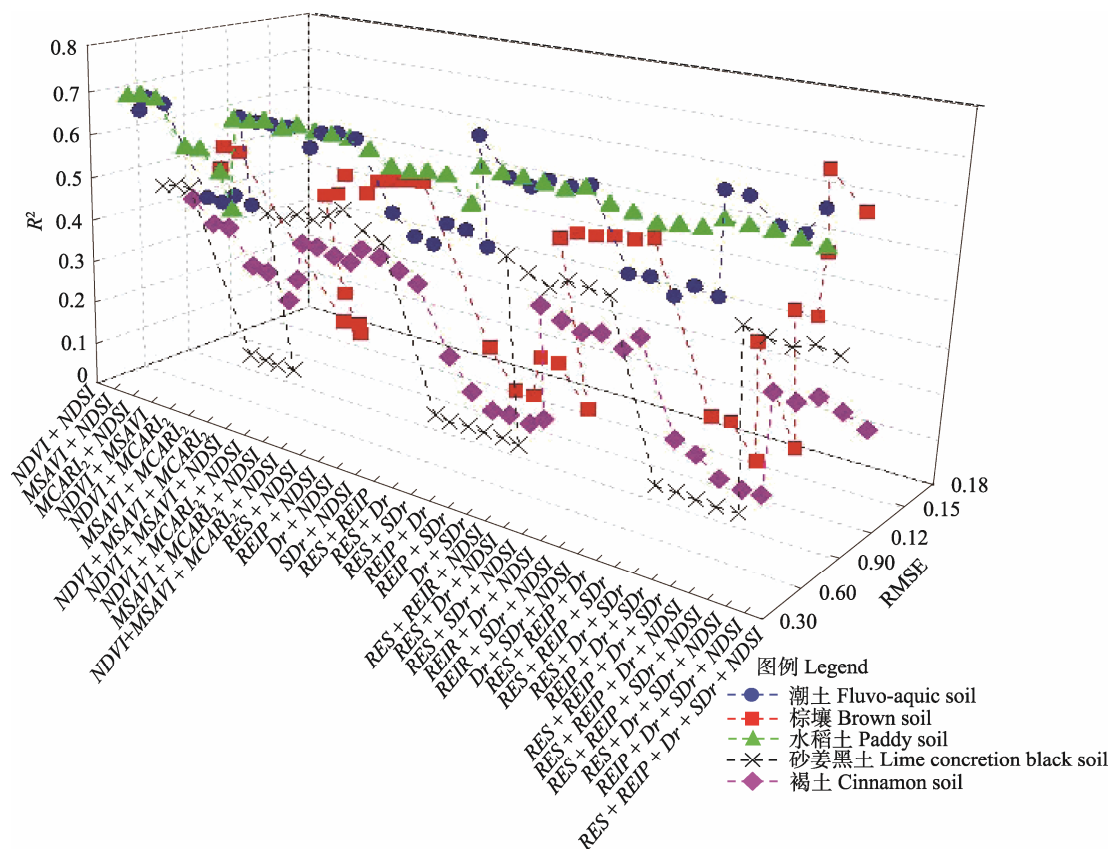


图5 不同组合的偏最小二乘回归模型估测叶面积指数(LAI)结果。NDVI、MSAVI、MCARI₂、RES、REIP、Dr、SDr、NDSI 同表2。RMSE, 均方根误差。

Fig. 5 Results of different groups of partial least squares regression model for estimating leaf area index (LAI). NDVI, MSAVI, MCARI₂, RES, REIP, Dr, SDr, NDSI see Table 2. RMSE, root mean square error.

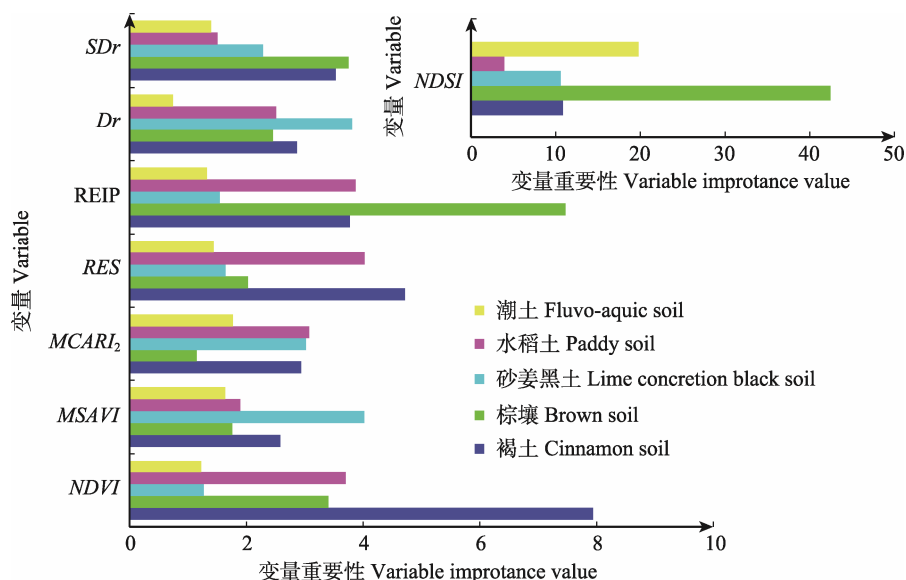


图6 优选的随机森林回归(RFR)模型中各变量的重要性。NDVI、MSAVI、MCARI₂、RES、REIP、Dr、SDr、NDSI同表2。

Fig. 6 Variable importance values in optimal random forest regression (RFR) models. NDVI, MSAVI, MCARI₂, RES, REIP, Dr, SDr, NDSI see Table 2.

模型预测的LAI趋势不同(潮土、棕壤、砂姜黑土地
区预测的LAI有高估趋势, 褐土地地区预测的LAI有低

估趋势, 水稻土地地区预测的LAI高、低估趋势均有);
是否考虑土壤类型因素对棕壤地区PLSR的LAI预测

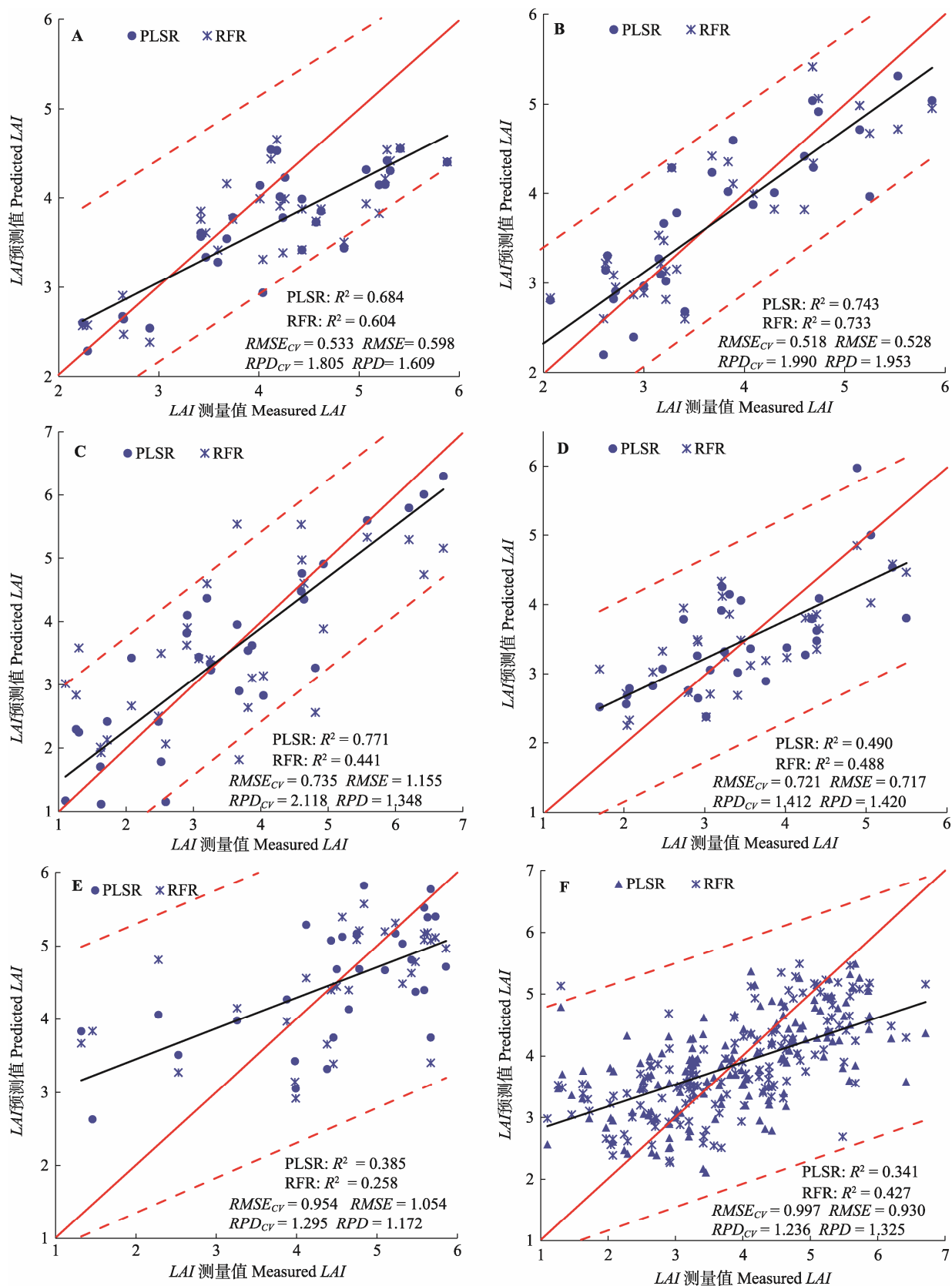


图7 冬小麦叶面积指数(LAI)测量值与预测值的拟合精度。A, 水稻土。B, 潮土。C, 棕壤。D, 砂姜黑土。E, 褐土。F, 5种土壤类型。虚线为预测区间线。PLSR, 偏最小二乘回归; RFR, 随机森林回归; $RMSE$, 均方根误差; RPD , 相对分析误差。

Fig. 7 Measured vs. predicted leaf area index (LAI). A, Paddy soil. B, Fluvo-aquic soil. C, Brown soil. D, Lime concretion black soil. E, Cinnamon soil. F, Five soil types. Dash lines indicate the confidence intervals of prediction. PLSR, partial least squares regression; RFR, random forest regression; $RMSE$, root mean square error; RPD , relative percent deviation.

表7 叶面积指数(LAI)估算精度

Table 7 Leaf area index (LAI) predicted accuracy

土壤类型 Soil type	考虑土壤类型因素 Considering the soil background				不考虑土壤类型因素 Neglecting the soil background			
	PLSR		RFR		PLSR		RFR	
	R^2	$RMSE_{CV}$	R^2	$RMSE$	R^2	$RMSE_{CV}$	R^2	$RMSE$
水稻土 Paddy soil	0.684	0.533	0.604	0.598	0.529	0.689	0.722	0.553
潮土 Fluvo-aquic soil	0.743	0.518	0.733	0.528	0.559	0.682	0.635	0.632
棕壤 Brown soil	0.771	0.735	0.441	1.155	0.098	1.467	0.244	1.387
砂姜黑土 Lime concretion black soil	0.490	0.721	0.488	0.717	0.173	0.953	0.413	0.768
褐土 Cinnamon soil	0.385	0.954	0.258	1.054	0.349	0.986	0.287	1.052

PLSR, 偏最小二乘回归; RFR, 随机森林回归。RMSE, 均方根误差。

PLSR, partial least squares regression; RFR, random forest regression. RMSE, root mean square error.

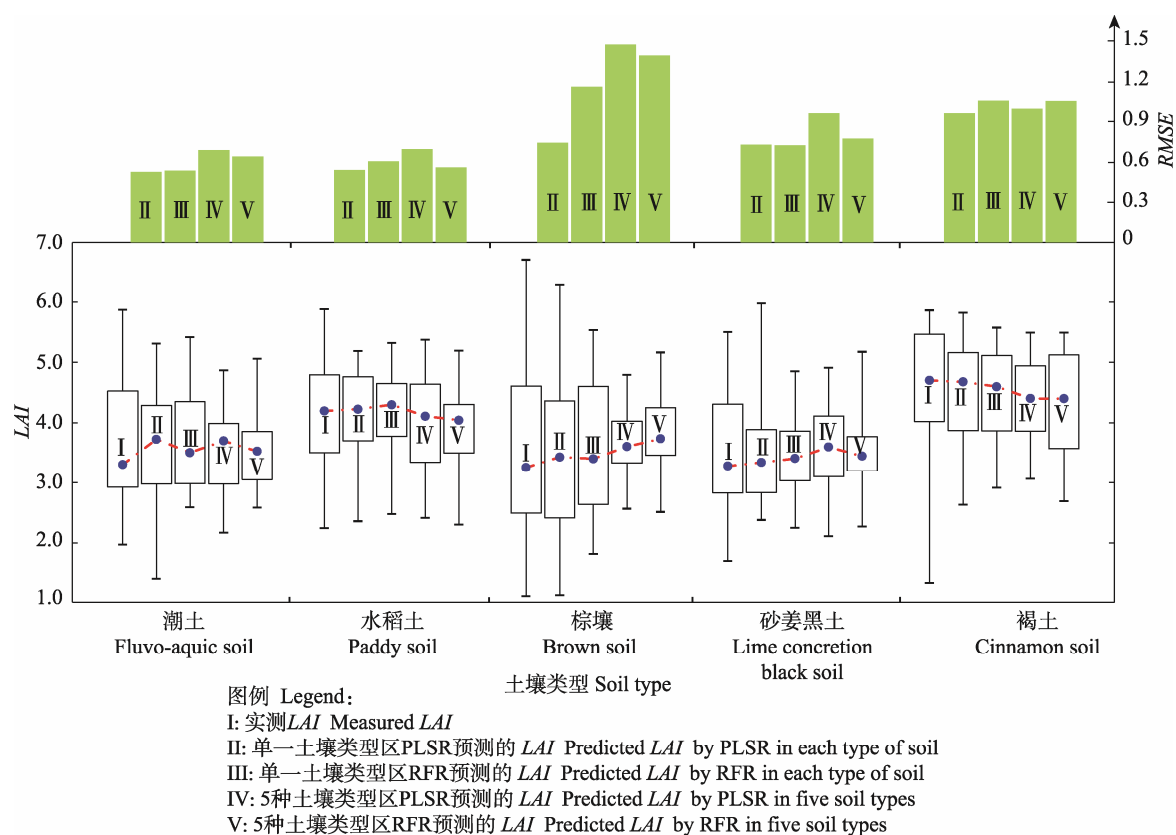


图8 估算叶面积指数(LAI)与实测LAI的箱线图比较。PLSR, 偏最小二乘回归; RFR, 随机森林回归。RMSE, 均方根误差。箱线图从上到下依次表示一组数据的最大值、上四分位数、中位数、下四分位数和最小值。

Fig. 8 Comparison of measured leaf area index (LAI) and predicted LAI in box plots. PLSR, partial least squares regression; RFR, random forest regression. RMSE, root mean square error. From top to bottom, the box plot represents the maximum, upper quartile, median, lower quartile, and minimum of a set data, respectively.

精度影响较大, 对褐土地区PLSR的LAI预测精度影响较小。

基于上述分析, 作者认为: 1)土壤类型不同很可能造成长势相同的作物具有不同的光谱指数信息, 因此不考虑土壤类型而直接使用光谱指数估算LAI, 会降低预测LAI的有效性, 容易使人们对作物长势产生错误认知; 2)在小区尺尤其是田块尺度的研究

中, 模型对变量的解释能力可作为模型选择的第一考虑, 因为参与构建最优模型的变量是农作物长势和墒情最佳描述的数学表达, 对它们的理解有助于改善对作物长势的认知, 便于农业管理的进行, 而在此方面PLSR优于RFR; 3)在未知地表先验知识前提下, RFR对大区域LAI估算比PLSR适合, 但先验知识的获取对LAI遥感估算仍然十分必要。

3 结论

面对当前宽幅、高空间(光谱)分辨率的星载传感器发展趋势,一景影像往往涵盖多种土壤类型,农田土壤背景愈加复杂,而植被反射率是各种因素物理耦合的结果,它反映了植被-土壤的光学性质以及它们在冠层内的光子交换过程;随着植被物候时序变化,土壤背景的影响随植被覆盖度的增大而明显减小,但仍然存在。因此,研究分析常用的8种光谱指数和估算方法在不同土壤类型区估算LAI的差异,得出以下结论:

1)光谱指数对LAI的敏感性因土壤类型背景不同而差异明显,因此,波段位置的合理选择是降低植被下土壤类型背景影响的首要保证。8个光谱指数中,REIP受植被下土壤类型背景影响最小;“lambda-by-lambda”波段优选算法不仅可以提供对LAI最敏感的光谱区间,也可在一定程度上为抵抗植被下土壤类型差异影响的光谱指数构建提供可行思路。因此,后期研究可借鉴“lambda-by-lambda”算法发展一种适用于各种陆表条件的可见光-红边-近红外波段的光谱指数。

2)回归模型的LAI预测精度因是否考虑土壤类型而差异明显;同时,土壤类型对回归模型估算LAI精度的影响程度也因所使用的回归模型不同而有差异。在小区域研究中,模型对变量的解释能力可作为模型选择的第一考虑,而在此方面PLSR优于RFR;在未知地表先验知识前提下,RFR对大区域LAI估算比PLSR适合,但先验知识的获取对LAI遥感估算仍然十分必要。

这些研究结论为发展一种适应各种陆表条件的LAI遥感估算方法和宽覆盖卫星影像下基于土壤背景影响去除模型的LAI遥感估算研究提供了理论依据和方法途径;但鉴于本研究涉及的作物品种、生长阶段以及样本数量有限,今后我们会在此基础上,对更复杂的陆表环境开展更深层次的探索。

基金项目 国家自然科学基金(41771370)和国家重点研发计划(2017YFD0600903)。

参考文献

- Abdi H (2003). Partial least square regression (PLS regression). In: Salkind N ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, Sage, USA.
- Allen WA, Gausman HW, Richardson AJ (1973). Willstätter - -

- Stoll theory of leaf reflectance evaluated by ray tracing. *Applied Optics*, 12, 2448–2453.
- Atzberger C, Guérif M, Baret F, Werner W (2010). Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Computers and Electronics in Agriculture*, 73, 165–173.
- Baret F, Guyot G (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, 35, 161–173.
- Bausch WC (1993). Soil background effects on reflectance-based crop coefficients for corn. *Remote Sensing of Environment*, 46, 213–222.
- Bolton DK, Friedl MA (2013). Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and Forest Meteorology*, 173, 74–84.
- Breiman L (2001). Random forests. *Machine Learning*, 45, 5–32.
- Broge NH, Leblanc E (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76, 156–172.
- Chen JM, Black TA (1992). Defining leaf area index for non-flat leaves. *Plant, Cell & Environment*, 15, 421–429.
- Cho MA, Skidmore A, Corsi F, van Wieren SE, Sobhan I (2007). Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation*, 9, 414–424.
- Danson FM, Plummer SE (1995). Red-edge response to forest leaf area index. *Remote Sensing*, 16, 183–188.
- Delegido J, Verrelst J, Meza CM, Rivera JP, Alonso L, Moreno J (2013). A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *European Journal of Agronomy*, 46, 42–52.
- Darmawan A, Nadirah, Wibowo A, Evri M, Mulyono S, Nugroho AS, Sadly M, Hendiarti N, Kashimura O, Kobayashi C, Uchida A, Uruguchi A, Sekine H (2009). Quantitative analysis from unifying field and airborne hyperspectral in prediction biophysical parameters by using partial least square (PLSR) and Normalized Difference Spectral Index (NDSI). <https://www.researchgate.net/publication/289258580>. Cited: 2017-9-1.
- Elsayed S, Elhoweity M, Schmidhalter U (2015). Normalized difference spectral indices and partial least squares regression to assess the yield and yield components of peanut. *Australian Journal of Crop Science*, 9, 976.
- Fang HL, Liang SL, Hoogenboom G (2011). Integration of MODIS LAI and vegetation index products with the CSM-CERES-Maize model for corn yield estimation.

- International Journal of Remote Sensing*, 32, 1039–1065.
- Farrar TJ, Nicholson SE, Lare AR (1994). The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. II. NDVI response to soil moisture. *Remote Sensing of Environment*, 50, 121–133.
- Feng W, Zhu Y, Yao X, Tian YC, Guo TC, Cao WX (2009). Monitoring nitrogen accumulation in wheat leaf with red edge characteristics parameters. *Transaction of the Chinese Society of Agricultural Engineering*, 25, 194–201. (in Chinese with English abstract) [冯伟, 朱艳, 姚霞, 田永超, 郭天财, 曹卫星 (2009). 利用红边特征参数监测小麦叶片氮素积累状况. 农业工程学报, 25, 194–201.]
- Filella I, Penuelas J (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing*, 15, 1459–1470.
- Gao L, Yang GJ, Li CC, Feng HK, Xu B, Wang L, Dong JH, Fu K (2017). Application of an improved method in retrieving leaf area index combined spectral index with PLSR in hyperspectral data generated by unmanned aerial vehicle snapshot camera. *Acta Agronomica Sinica*, 43, 549–557. (in Chinese with English abstract) [高林, 杨贵军, 李长春, 冯海宽, 徐波, 王磊, 董锦绘, 付奎 (2017). 基于光谱特征与PLSR结合的叶面积指数拟合方法的无人机画幅高光谱遥感应用. 作物学报, 43, 549–557.]
- Gausman HW, Allen WA, Myers V, Cardenas R (1969). Reflectance and internal structure of cotton leaves, *Gossypium hirsutum* L. *Agronomy Journal*, 61, 374–376.
- Haboudane D, Miller JR, Patten E, Zarco-Tejada PJ, Strachan IB (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90, 337–352.
- Hansen PM, Schjoerring JK (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment*, 86, 542–553.
- Herrmann I, Pimstein A, Karnieli A, Cohen Y, Alchanatis V, Bonfil DJ (2011). LAI assessment of wheat and potato crops by VEN μ S and Sentinel-2 bands. *Remote Sensing of Environment*, 115, 2141–2151.
- Horler DNH, Dockray M, Barber J (1983). The red edge of plant leaf reflectance. *International Journal of Remote Sensing*, 4, 273–288.
- Huete AR (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295–309.
- Huete AR, Jackson RD, Post DF (1985). Spectral response of a plant canopy with different soil backgrounds. *Remote Sensing of Environment*, 17, 37–53.
- Jasinski MF, Eagleson PS (1989). The structure of red-infrared scattergrams of semivegetated landscapes. *IEEE Transactions on Geoscience and Remote Sensing*, 27, 441–451.
- Ju CH, Tian YC, Yao X, Cao WX, Zhu Y, Hannaway D (2010). Estimating leaf chlorophyll content using red edge parameters. *Pedosphere*, 20, 633–644.
- Lee KS, Cohen WB, Kennedy RE, Maersperger TK, Gower ST (2004). Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. *Remote Sensing of Environment*, 91, 508–520.
- Li F, Misteel B, Hu Y, Chen X, Schmidhalter U (2013). Comparing hyperspectral index optimization algorithms to estimate aerial N uptake using multi-temporal winter wheat datasets from contrasting climatic and geographic zones in China and Germany. *Agricultural and Forest Meteorology*, 180, 44–57.
- Li H, Chen ZX, Jiang ZW, Wu WB, Ren JQ, Liu B, Tuya H (2017). Comparative analysis of GF-1, HJ-1, and Landsat-8 data for estimating the leaf area index of winter wheat. *Journal of Integrative Agriculture*, 16, 266–285.
- Li XC, Zhang YJ, Bao YS, Luo JH, Jin XL, Xu XG, Song XY, Yang GJ (2014). Exploring the best hyperspectral features for LAI estimation using partial least squares regression. *Remote Sensing*, 6, 6221–6241.
- Liu L, Zhang R, Zuo Z (2016). The relationship between soil moisture and LAI in different types of soil in central eastern China. *Journal of Hydrometeorology*, 17, 2733–2742.
- McLachlan G (2004). *Discriminant Analysis and Statistical Pattern Recognition*. John Wiley & Sons, Hoboken.
- Myneni RB, Hall FB, Sellers PJ, Marshak AL (1995). The interpretation of spectral vegetation indices. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 481–486.
- Nolet C, Poortinga A, Roosjen P, Bartholomeus H, Ruessink G (2014). Measuring and modeling the effect of surface moisture on the spectral reflectance of coastal beach sand. *PLOS ONE*, 9, e112151. doi: 10.1371/journal.pone.0112151.
- Qi J, Chehbouni A, Huete AR, Kerr YH, Sorooshian S (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48, 119–126.
- Rondeaux G, Steven M, Baret F (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55, 95–107.
- Rossel RA, Webster R (2012). Predicting soil properties from the Australian soil visible-near infrared spectroscopic database. *European Journal of Soil Science*, 63, 848–860.
- Rouse JW, Haas RH, Schell JA, Deering DW, Harlan JC (1974). Monitoring the vernal advancements and retrogradation of natural vegetation, NASA/GSFC, Type III, Final Report. <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022555.pdf>. Cited: 2017-9-1.
- Sadeghi M, Jones SB, Philpot WD (2015). A linear physically-based model for remote sensing of soil moisture using short wave infrared bands. *Remote Sensing of Environment*, 164, 66–76.

- Schlemmer M, Gitelson A, Schepers J, Ferguson R, Peng Y, Shanahan J, Rundquist D (2013). Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal of Applied Earth Observation and Geoinformation*, 25, 47–54.
- Siegmann B, Jarmer T (2015). Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data. *International Journal of Remote Sensing*, 36, 4519–4534.
- Thenkabail PS, Lyon JG, Huete A (2012). *Hyperspectral Remote Sensing of Vegetation*. CRC Press, New York.
- Todd SW, Hoffer RM (1998). Responses of spectral indices to variations in vegetation cover and soil background. *Photogrammetric Engineering and Remote Sensing*, 64, 915–922.
- Viña A, Gitelson AA, Nguy-Robertson AL, Peng Y (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment*, 115, 3468–3478.
- Vogelmann JE, Rock BN, Moss DM (1993). Red edge spectral measurements from sugar maple leaves. *International Journal of Remote Sensing*, 14, 1563–1575.
- Wang FM, Huang JF, Lou ZH (2011). A comparison of three methods for estimating leaf area index of paddy rice from optimal hyperspectral bands. *Precision Agriculture*, 12, 439–447.
- Wang HW, Wu ZB, Meng J (2006). *Partial Least Squares Regression-Linear and Nonlinear Methods*. National Defense Industry Press, Beijing. (in Chinese) [王惠文, 吴载斌, 孟洁 (2006). 偏最小二乘回归的线性与非线性方法. 国防工业出版社, 北京.]
- Wold S, Sjöström M, Eriksson L (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, 109–130.
- Woolley JT (1971). Reflectance and transmittance of light by leaves. *Plant Physiology*, 47, 656–662.
- Xu XR (2005). *Physics of Remote Sensing*. Peking University Press, Beijing. (in Chinese) [徐希孺 (2005). 遥感物理. 北京大学出版社, 北京.]
- Yeniay O, Goktas A (2002). A comparison of partial least squares regression with other prediction methods. *Hacettepe Journal of Mathematics and Statistics*, 31, 99–101.
- Yuan HH, Yang GJ, Li CC, Wang YJ, Liu JG, Yu HY, Feng HK, Xu B, Zhao XQ, Yang XD (2017). Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. *Remote Sensing*, 9, 309. doi: 10.3390/rs9040309.
- Yu K, Lenz-Wiedemann V, Chen X, Bareth G (2014). Estimating leaf chlorophyll of barley at different growth stages using spectral indices to reduce soil background and canopy structure effects. *ISPRS Journal of Photogrammetry and Remote Sensing*, 97, 58–77.
- Zhao CJ, Wang ZJ, Wang JH, Huang WJ (2012). Relationships of leaf nitrogen concentration and canopy nitrogen density with spectral features parameters and narrow-band spectral indices calculated from field winter wheat (*Triticum aestivum* L.) spectra. *International Journal of Remote Sensing*, 33, 3472–3491.

特邀编委: 郭庆华 责任编辑: 李 敏



扫码向作者提问