doi:10.6041/j.issn.1000-1298.2021.12.020

基于小波能量系数和叶面积指数的冬小麦生物量估算

李长春　李亚聪　王艺琳　马春艳　陈伟男　丁　凡
（河南理工大学测绘与国土信息工程学院，焦作 454000）

摘要：生物量是评价作物长势及产量估算的重要指标，科学、快速、准确地获取生物量信息，对于监测冬小麦生长状况以及产量预测等具有重要意义。以冬小麦为研究对象，通过相关性分析，选取相关性较好的小波能量系数，同时耦合叶面积指数，基于支持向量回归算法、随机森林算法、高斯过程回归 3 种算法构建冬小麦生物量估算模型。结果显示，基于小波能量系数，分别利用支持向量回归算法、随机森林算法、高斯过程回归进行生物量估算，4 个生育期的验证 R^2 分别是 0.55、0.40、0.39；0.75、0.70、0.83；0.84、0.92、0.93；0.84、0.89、0.85。表明高斯过程回归模型估算精度最优。叶面积指数耦合小波能量系数，利用支持向量回归算法、随机森林回归算法、高斯过程回归进行生物量估算，4 个生育期的验证 R^2 分别是 0.76、0.73、0.77；0.76、0.72、0.84；0.87、0.94、0.94；0.85、0.90、0.91。表明高斯过程回归算法估算精度最优，并且在一定程度上能够克服冠层光谱饱和现象，提高模型估算精度。以小波能量系数和叶面积指数为输入变量结合高斯过程回归算法建立冬小麦生物量估算模型，可以提高生物量估算精度，为基于遥感技术的作物参数快速估算提供参考。

关键词：冬小麦；生物量；小波能量系数；叶面积指数；高斯过程回归
中图分类号：S532　　**文献标识码**：A　　**文章编号**：1000-1298(2021)12-0191-10　　**OSID**：

Winter Wheat Biomass Estimation Based on Wavelet Energy Coefficient and Leaf Area Index

LI Changchun　LI Yacong　WANG Yilin　MA Chunyan　CHEN Weinan　DING Fan
(*School of Surveying and Land Information Engineering*, *Henan Polytechnic University*, *Jiaozuo* 454000, *China*)

Abstract: Biomass is an important indicator for evaluating crop growth and yield estimation. Obtaining biomass information scientifically, quickly and accurately is of great significance for monitoring the growth status of winter wheat and yield prediction. Taking winter wheat as the research object, through correlation analysis, the wavelet energy coefficient with good correlation was selected, and the leaf area index was coupled at the same time. Based on the support vector regression algorithm, random forest algorithm, and Gaussian process regression, three algorithms were used to construct a winter wheat biomass estimation model. The verification R^2 of the four growth periods were 0.55, 0.40 and 0.39; 0.75, 0.70 and 0.83; 0.84, 0.92 and 0.93; 0.84, 0.89 and 0.85, respectively. It was showed that the estimation accuracy of Gaussian process regression model was the best. Leaf area index coupled with wavelet energy coefficients, using the three algorithms to estimate biomass, the verification R^2 of the four growth periods were 0.76, 0.73 and 0.77; 0.76, 0.72 and 0.84; 0.87, 0.94 and 0.94; 0.85, 0.90 and 0.91, respectively, indicating that the Gaussian process regression algorithm had the best estimation accuracy, and to a certain extent, it can overcome the canopy spectrum saturation phenomenon and improve the estimation accuracy of the model. Using wavelet energy coefficient and leaf area index as input variables combined with Gaussian process regression algorithm to establish a winter wheat biomass estimation model, which can improve the accuracy of biomass estimation and provide a scientific reference for the rapid estimation of crop parameters based on remote sensing technology.

Key words: winter wheat; biomass; wavelet energy coefficient; leaf area index; Gaussian process regression

收稿日期：2021-07-31　修回日期：2021-09-13
基金项目：国家自然科学基金项目(41871333)和河南省科技攻关项目(212102110238)
作者简介：李长春(1976—)，男，副教授，主要从事农业定量遥感研究，E-mail：lichangchun610@126.com
通信作者：李亚聪(1996—)，女，助理研究员，主要从事农业定量遥感研究，E-mail：1632540199@qq.com

0 引言

生物量是指单位面积内作物在一定时间内累积的有机物[1]。生物量作为作物生长发育过程中的一个重要生理参数,其变化能够直接表征作物进行净光合作用积累有机物的能力,反映作物的生长情况[2]。因此,科学、快速、准确地获取生物量信息,对监测冬小麦长势以及产量估算具有重要的意义。

传统生物量获取方法主要是野外实测方法,该方法应用于小区域地区精度高,但对于大区域而言,工作过程繁琐、费时费力,时效性差,且大面积的田间破坏性取样会影响作物的生长[3]。遥感技术凭借其宏观、动态和实时等特点在作物监测方面得到广泛的关注,为估算作物生物量提供有效手段。高光谱遥感技术具有较高的光谱分辨率和较强的波段连续性,是对地观测领域的重大成就,是当今遥感科学的前沿技术[4]。在利用遥感技术进行作物生物量估算方面,构建植被指数基于经验统计的方法是目前生物量估算比较常用的方法,植被指数是由一个或多个光谱波段以一定的数学方式进行结合,能够反映植被生长状况。孙奇等[5]通过分析可见光-近红外植被指数、短波红外植被指数和 8 个融合植被指数与生物量的关系,结果显示融合植被指数对冬小麦生物量的估算能力更高,其中 MTVI2 结合 NDMI 估算精度最高。李岚涛等[6]利用波长 400 ~ 950 nm 处的冠层反射率,采用支持向量机和偏最小二乘回归方法构建不同生育期的冬小麦生物量估算模型,结果显示抽穗期效果最好,R^2 大于 0.85。郭超凡等[7]利用单变量植被指数,基于多元线性和随机森林方法对小麦地上部生物量进行估算,发现绿色叶绿素指数和归一化差异水体指数与生物量具有最优拟合关系,随机森林模型估算效果最优,R^2 达到 0.74。陶惠林等[8]基于植被指数、红边参数结合多元线性回归模型估算冬小麦生物量,结果表明植被指数耦合红边参数能够提高模型估算精度。使用植被指数建立生物量估算模型方便简单,但是光谱信息利用率较低,仅使用了其中的几个波段,不能充分利用有效光谱信息,并且使用植被指数估算多生育期的生物量容易存在饱和现象。近年来,为了进一步提高生物量估算精度,一些学者通过对光谱进行变换,如小波变换[9-11]、光谱微分[12-14]、连续统去除法[15-16]等多种方法挖掘潜在的有效光谱信息,提取隐藏信息,增强光谱信息的敏感度,有效解决了模型的饱和问题。

然而,这些研究仅仅利用光谱及其变换信息估算作物生物量,未综合考虑叶面积指数与生物量的关系。本文基于高光谱数据,通过连续小波变换得到的小波能量系数,耦合叶面积指数,分别利用支持向量回归算法、随机森林算法和高斯过程回归算法,构建冬小麦生物量估算模型,分析不同模型生物量估算效果,寻求最佳估算方法,以期为基于遥感技术的作物高通量表型参数快速估算提供参考。

1 材料与方法

1.1 研究区概况

研究区位于北京市昌平区小汤山国家精准农业研究示范基地。基地位于北纬 40°00′ ~ 40°21′,东经 116°34′ ~ 117°00′,平均海拔为 36 m,气候类型是典型的暖温带半湿润大陆季风气候,春季干旱多风,夏季炎热多雨,秋季凉爽,冬季寒冷干燥,四季分明。一般为每年 9 月末至 10 月初播种冬小麦,次年 3 月左右返青,经过拔节、孕穗、开花、灌浆 4 个生育期,6 月中上旬收获。研究区共设置 48 个小区,每个小区面积为 48 m^2(6 m × 8 m),每 16 个小区为 1 个重复,共 3 个重复。设置 4 个氮肥水平:N1:0 kg/hm^2,N2:15 kg/hm^2,N3:30 kg/hm^2,N4:45 kg/hm^2。冬小麦管理模式参考一般的大田模式,具有优越的肥力、灌溉和排水条件。其他按照田间实际管理进行操作,研究区地理位置如图 1 所示。

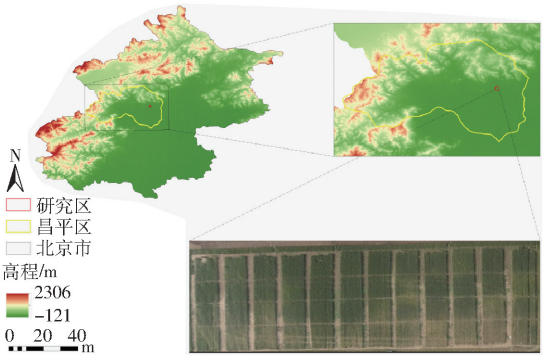

图 1 研究区地理位置

Fig. 1 Geographical location of study area

1.2 数据获取

从 4 月 13 日—5 月 27 日,每隔 13 ~ 15 d,在实验区里选取样本点进行小麦冠层高光谱数据、叶面积指数和生物量测量。

1.2.1 冠层高光谱数据获取

采用美国 ASD(Analytica Spectra Deviecs, Inc)生产的 FieldSpec 便携式地物光谱仪获取冬小麦冠层高光谱数据。该光谱仪可测量的波长为 350 ~ 2 500 nm,采样间隔为 1.4 nm(350 ~ 1 000 nm)和 2 nm(1 000 ~ 2 500 nm),仪器内部重采样后可达到 1 nm,测量时间选择在每天 10:30—14:00。为减少土壤背景的影响,测量过程中,传感器探头垂直向

下，距地表 1.3 m 处，视场角为 25°。每个实验小区均匀采集 20 条光谱数据，采集完成后利用与 ASD 光谱仪配套的光谱数据处理软件 ViewSpecPro 导出无量纲的反射率，以每个实验小区 20 条光谱数据的平均值作为各小区的冠层反射率。由于水分吸收带严重，所以选取波段 350 ~ 1 350 nm 内的光谱数据。在测量过程中，当每个样点测量完成后，应及时利用白板进行标准白板校准（标准白板反射率为 1，获得的目标光谱为相对反射率）。

1.2.2　叶面积指数获取

采集光谱数据的同时，在采样点上获取小麦植株，带回室内，将取样的冬小麦进行茎叶分离，利用 CI－203 型激光叶面积仪测定叶片叶面积，求其总和，进而计算出单茎叶面积，然后乘以单位土地面积单茎数得到冬小麦叶面积指数。4 个生育期共获取到 192 个冬小麦叶面积指数样本数据。

1.2.3　生物量获取

利用收获法获取生物量数据。每个小区内，随机选取 10 株长势均匀冬小麦植株作为样本，经茎叶分离、称量等操作后编号放入纸袋中，将干燥温度保持在 105℃，运行 30 min 杀青，随后将温度设置为 75℃干燥至质量恒定（48 h 左右），测定各器官干质量，最后根据群体密度及样本干质量，计算单位面积冬小麦地上生物量，最终获取生物量。

1.3　数据处理与模型构建方法

1.3.1　连续小波变换

小波变换是由傅里叶变换发展起来的一种数据处理方法，通过伸缩和平移等运算功能对函数或信号进行多尺度细化分析[17]。小波变换包括连续小波变换（Continuous wavelet transform，CWT）和离散小波变换（Discrete wavelet wransform，DWT），后者是前者的离散化。连续小波变换作为一种线性变换的方法，广泛应用于高光谱数据处理[18-20]。本文采用 CWT 对 ASD 高光谱数据进行分解，其原理就是将冠层高光谱数据分解成不同尺度和不同波长的小波能量系数。将一维空间的高光谱反射率转换为二维空间尺度和小波能量系数。利用 CWT 分别获取 4 个生育期前 10 尺度下的小波能量系数，计算式为

$$W_f(a,b) = \int_{-\infty}^{+\infty} f(\lambda)\Psi_{a,b}(\lambda)\mathrm{d}\lambda \tag{1}$$

其中

$$\Psi_{a,b}(\lambda) = \frac{1}{\sqrt{a}}\psi\left(\frac{\lambda - b}{a}\right) \tag{2}$$

式中　$f(\lambda)$——冠层高光谱反射率

λ——波段 350 ~ 1 350 nm 范围的光谱波段

$\Psi_{a,b}$——小波基函数

a——尺度因子　　b——平移因子

1.3.2　模型构建方法

（1）高斯过程回归

高斯过程回归（Gaussian process regression，GPR）由 RASMUSSEN 和 WILLIAMS 在 2005 年提出[21]，是将统计学习理论和贝叶斯理论相结合的一种机器学习方法，对于处理小样本、高维度和非线性复杂关系的回归和分类问题都有较好的效果，近年来成为作物参数估算最高效的机器学习算法之一[22-23]。

（2）支持向量回归

支持向量回归（Support vector regression，SVR）算法是由支持向量机算法演化而来，作为一种监督学习的机器学习算法，基本原理是通过建立一个最优决策超平面，使距样本之间距离达到最小化，从而拟合样本数据[24]，与其他算法相比，模型准确性较高，对高维度、小样本数据有较好的处理能力，具有良好的泛化能力和鲁棒性[25]。

（3）随机森林

随机森林（Random forest，RF）基本原理是利用多个决策树算法对相同现象做重复的预测[26]，作为一种组成式的有监督学习方法，是机器学习方法的分支之一，即集成学习方法，在处理大量输入变量问题中优势明显，预测性能好、鲁棒性强[27]。

本文以不同生育期小波能量系数、小波能量系数耦合 LAI 为自变量，生物量为因变量，在每个生育期，分别挑选 2/3 样本数据（32 个）作为建模集，1/3 样本数据（16 个）作为验证集，构建冬小麦生物量估算模型。

1.3.3　相关系数

采用皮尔逊相关系数 R 来表示两个随机变量之间的相关程度，相关系数取值范围为［－1，1］，相关系数绝对值越大，相关性越强，相关系数绝对值越接近于 0，相关度越弱[28]。相关系数计算式为

$$R(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{3}$$

式中　$R(X,Y)$——两随机变量相关系数

$\mathrm{cov}(X,Y)$——两随机变量协方差

σ——标准差

1.3.4　模型精度评价

采用决定系数（R^2）、均方根误差（RMSE）和标准均方根误差（nRMSE）3 个指标作为模型精度评价指标。R^2越大，表明该模型的拟合效果越高，并且越稳定。nRMSE 小于 10% 表示估算值和实测值一致性极好；nRMSE 为 10% ~ 20% 表示一致性较好；nRMSE 为 20% ~ 30% 表示一致性中等；nRMSE 大于等于 30% 表示一致性较差。

2 结果与分析

2.1 冬小麦生物量与叶面积指数变化特征

为了研究冬小麦生物量与叶面积指数之间的关系,对 4 个生育期生物量和叶面积指数的结果进行统计,结果如表 1 所示。

表 1 不同生育期实测生物量和叶面积指数统计分析

Tab. 1 Statistical analysis of measured biomass and leaf area index in different growth periods

参数	生育期	平均值	最大值	最小值	标准差	变异系数/%
生物量/(kg·hm^{-2})	拔节期	2 631.90	4 396.05	1 477.38	691.66	26.28
	孕穗期	5 113.59	8 266.23	2 194.37	1 467.54	28.70
	开花期	5 940.82	9 713.50	2 384.55	1 612.87	27.15
	灌浆期	5 321.95	9 009.55	2 398.05	1 696.16	31.87
叶面积指数	拔节期	3.71	5.49	2.16	0.94	25.35
	孕穗期	4.19	8.81	1.30	1.61	38.35
	开花期	3.29	5.89	1.24	1.19	35.98
	灌浆期	1.59	3.81	0.35	0.88	55.58

由表 1 可知,从拔节期到灌浆期,冬小麦生物量总体呈先升高后下降的变化趋势,拔节期生物量最小,平均值为 2631.90 kg/hm^2,开花期达到最大,平均值为 5940.82 kg/hm^2,开花期后生物量逐渐下降。变异系数在 26.28% ~31.87% 之间,离散程度较稳定,表明在完整生育期内,冬小麦生物量值域分布随着时间推移趋于集中。从拔节期到开花期,总生物量迅速上升,随着叶片数量增加和叶面积扩大,光合速率提高,使得各构件生物量迅速积累;开花期到灌浆期,即成熟初期,由于叶片和茎秆逐渐老化变黄,光合作用能力下降,生成的光合产物低于自身消耗,乃至脱落,小麦生物量呈很小的下降趋势,与侯学会等[29]的研究结果一致。

冬小麦叶面积指数在整个生育期内先增大后减小,呈抛物线型变化。在孕穗期之前,冬小麦营养生长和生殖生长较为活跃,叶片快速生长,叶面积指数迅速增加,且在拔节期与孕穗期之间达到峰值,孕穗期之后,随着营养向生殖器官的快速转移,叶片逐渐老化变黄、脱落,因而叶面积指数呈现下降趋势。拔节期离散程度较小,变异系数为 25.35% ,而灌浆期叶面积指数离散程度较大,变异系数为 55.58% 。

2.2 相关性分析

2.2.1 原始光谱与生物量相关性分析

对冬小麦拔节期、孕穗期、开花期和灌浆期去噪后的高光谱数据与相应的生物量数据进行相关性分析,计算相关系数 R,绘制不同生育期相关系数曲线,结果如图 2 所示。

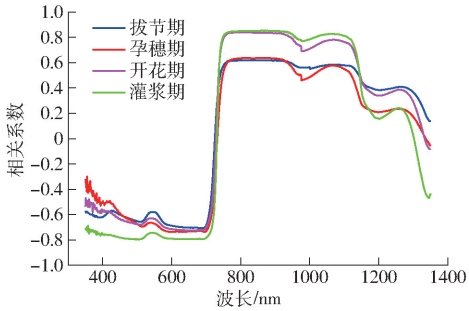

图 2 原始光谱与生物量相关性

Fig. 2 Correlation between original spectrum and biomass

由图 2 可知,各个生育期的冬小麦冠层光谱反射率与生物量的相关系数变化趋势表现出相似的特点。在波段 500 ~700 nm 附近,冬小麦冠层光谱反射率与生物量都有很高的负相关性,波段 800 ~900 nm 范围内相关系数基本保持不变,冬小麦冠层光谱反射率与生物量都有很高的正相关性。在拔节期,波长 691 nm 处负相关程度最大,相关系数 R 为 −0.71,波长 823 nm 处正相关程度最大,相关系数 R 为 0.62;在孕穗期,波长 695 nm 处负相关程度最大,相关系数 R 为 −0.74,波长 868 nm 处正相关性最大,相关系数 R 为 0.64;在开花期,波长 691 nm 处负相关性最大,相关系数 R 为 −0.73,波长 819 nm 处正相关性最大,相关系数 R 为 0.84;在灌浆期,波长 505 nm 处负相关性最大,相关系数 R 为 −0.80,波长 877 nm 处正相关性最大,相关系数 R 为 0.85。

2.2.2 小波能量系数与生物量相关性分析

对于冬小麦 4 个生育期,利用连续小波变换对冬小麦冠层高光谱数据进行 1 ~10 尺度分解,不同波段数与不同分解尺度对应的相关系数,表征小波能量系数与冬小麦生物量之间的相关性,通过连续小波变换间接对目标光谱信息进行优化。在每个分解尺度下,各波段生成特定的小波能量系数,然后分别与生物量进行相关性分析,生成相关性矩阵,结果如图 3 所示。

由图 3 看出,4 个生育期小波能量系数与生物量的相关性随着分解尺度的增加,呈现先升高后降低的趋势;通过 0.01 极显著水平检验的光谱波段数量不断增加,当分解尺度为 10,4 个生育期均达到 1 001 个波段。由图 3a 可知,在拔节期当分解尺度为 1 ~8 时,小波能量系数与生物量的最大相关系数绝对值均在 0.71 以上,分解尺度为 2 时,在波长 948 nm 处 $|R|$ 最大可达 0.77;当分解尺度为 9 和 10 时,$|R|$ 最大分别为 0.64、0.52,小波能量系数和生物量相关性较小,分解得到的小波能量系数已经无法表征冬小麦的生物量信息。由图 3b 可知,在孕穗

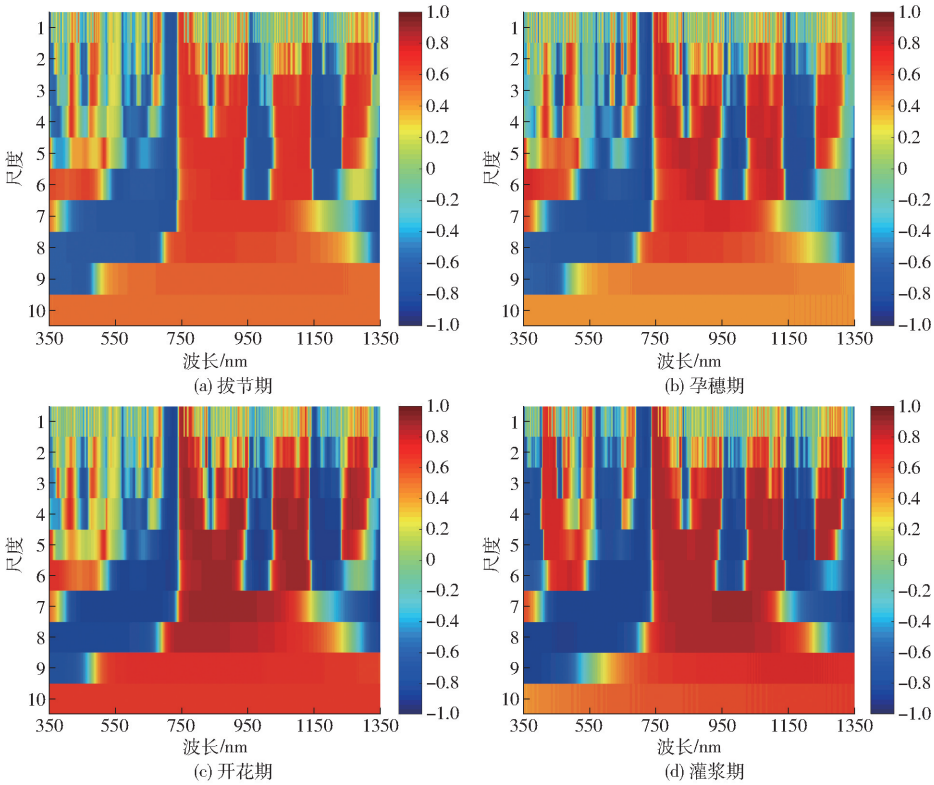

图 3　不同生育期小波能量系数与生物量相关性矩阵图

Fig. 3　Correlation matrix diagrams of wavelet energy coefficient and biomass in different growth periods

期当分解尺度为 1 ~ 8 时，小波能量系数与生物量的最大相关系数绝对值均在 0.80 以上，二者呈高度相关，分解尺度为 2 时，在波长 768 nm 处 $|R|$ 最大可达 0.84；当分解尺度为 9 和 10 时，$|R|$ 最大分别为 0.64、0.44，小波能量系数和生物量中低度相关，无法表征冬小麦的生物量信息。由图 3c 可知，在开花期当分解尺度为 1 ~ 8 时，小波能量系数与生物量的最大相关系数绝对值均在 0.89 以上，二者呈高度相关，当分解尺度为 3 时，在 912 nm 处 $|R|$ 最大可达 0.91；当分解尺度为 9 和 10 时，$|R|$ 最大分别为 0.74、0.64，小波能量系数和生物量中度相关，不能充分表征冬小麦的生物量信息。由图 3d 可知，在灌浆期当分解尺度为 1 ~ 9 时，小波能量系数与生物量的最大相关系数绝对值均在 0.86 以上，二者呈高度相关，分解尺度为 7 时，在波长 1 001 nm 处 $|R|$ 最大可达 0.88；当分解尺度为 10 时，$|R|$ 最大为 0.67，小波能量系数和生物量中度相关，不能充分表征冬小麦的生物量信息。

2.2.3　叶面积指数与生物量相关性分析

将实测得的叶面积指数与生物量数据利用式(3)计算相关系数。4 个生育期叶面积指数与生物量相关系数有所差异，4 个生育期 $|R|$ 分别为 0.91、0.82、0.93 和 0.87，均大于 0.8，且达到 0.01 显著性水平，说明叶面积指数与生物量具有很高的相关性。因此，在各生育期生物量预测模型中，可将叶面积指数与小波能量系数耦合共同作为模型输入变量以提高冬小麦生物量估算的精度。

2.3　敏感波段筛选

利用图 3 对每个生育期生物量与小波能量系数进行相关性分析，筛选出相关系数绝对值较高的前 10 个波段作为输入变量，估算冬小麦生物量，筛选结果如表 2 所示。

由表 2 可知，不同生育期不同尺度下最优波段有所不同，在拔节期，最优波段尺度在 1、2、3、4、6、7；在孕穗期，最优波段尺度在 1、2、3、4、5、6；在开花期，最优波段尺度在 3、4、5、6；在灌浆期，最优波段尺度在 4、5、6、7、8。综合分析来看，4 个生育期中，最优波段尺度主要集中在 3、4、5、6，并且每种尺度下筛选的敏感波段与生物量均达到 0.01 显著水平。

2.4　模型构建与精度评价

基于以上相关性分析结果，每个生育期以冬小麦生物量作为因变量，以小波能量系数、小波能量系能数耦合 LAI 为自变量，分别采用支持向量回归算法(SVR)、随机森林算法(RF)、高斯过程回归(GPR)算法构建冬小麦 4 个生育期的生物量高光谱估算模型，并且进行验证，结果见表 3、4。

由表 3、4 可知：

(1)利用相关性较好的小波能量系数为输入变量，分别利用支持向量回归算法、随机森林回归算法、高斯过程回归算法进行冬小麦生物量估算。拔

表 2 前 10 敏感波段及对应尺度

Tab. 2 Top ten sensitive bands and corresponding scales

序号	拔节期			孕穗期			开花期			灌浆期		
	尺度	波长/nm	|R|	尺度	波长/nm	|R|	尺度	波长/nm	|R|	尺度	波长/nm	|R|
1	C2	948	0.77	C2	768	0.84	C3	912	0.91	C7	1 001	0.88
2	C2	1 279	0.74	C1	755	0.84	C5	1 103	0.91	C7	967	0.88
3	C3	1 117	0.74	C6	727	0.84	C4	1 109	0.90	C4	1 171	0.88
4	C2	763	0.74	C5	900	0.83	C4	926	0.90	C7	929	0.88
5	C3	1 257	0.73	C5	725	0.83	C4	1 158	0.90	C5	1 192	0.88
6	C4	1 120	0.73	C3	728	0.83	C6	875	0.90	C8	475	0.88
7	C1	762	0.73	C5	839	0.83	C5	1 178	0.90	C6	1 108	0.87
8	C7	1 341	0.73	C4	726	0.83	C5	1 086	0.90	C8	502	0.88
9	C6	1 347	0.73	C3	775	0.83	C3	1 003	0.90	C7	493	0.87
10	C2	1 150	0.73	C2	728	0.83	C6	1 107	0.90	C4	1 258	0.87

表 3 不同生育期不同方法冬小麦生物量的建模精度

Tab. 3 Modeling results of winter wheat biomass in different growth periods and different methods

生育期	方法	小波能量系数			小波能量系数耦合叶面积指数		
		R^2	RMSE/(kg·hm^{-2})	nRMSE/%	R^2	RMSE/(kg·hm^{-2})	nRMSE/%
拔节期	SVR	0.65	451.18	17.74	0.86	292.95	11.44
	RF	0.44	543.52	20.86	0.64	430.66	16.47
	GPR	0.80	385.64	14.81	0.86	321.04	12.35
孕穗期	SVR	0.72	803.26	16.10	0.81	667.35	13.15
	RF	0.56	985.65	19.31	0.59	951.44	18.63
	GPR	0.79	679.71	13.32	0.89	502.69	9.85
开花期	SVR	0.79	715.36	13.70	0.92	517.43	8.69
	RF	0.77	788.74	13.29	0.81	709.35	11.91
	GPR	0.85	641.07	10.78	0.93	422.87	7.11
灌浆期	SVR	0.79	715.36	13.70	0.85	609.15	11.45
	RF	0.73	814.34	15.54	0.73	804.92	15.53
	GPR	0.82	657.74	12.56	0.91	459.08	8.77

表 4 不同生育期不同方法估算冬小麦生物量的验证精度

Tab. 4 Validation results of different growth periods and different methods to estimate winter wheat biomass

生育期	方法	小波能量系数			小波能量系数耦合叶面积指数		
		R^2	RMSE/(kg·hm^{-2})	nRMSE/%	R^2	RMSE/(kg·hm^{-2})	nRMSE/%
拔节期	SVR	0.55	506.01	20.10	0.76	348.06	13.47
	RF	0.40	505.83	19.74	0.73	360.12	13.75
	GPR	0.39	523.26	20.19	0.77	363.73	13.87
孕穗期	SVR	0.75	912.6	19.14	0.76	874.55	18.00
	RF	0.70	889.28	18.24	0.72	858.84	17.59
	GPR	0.83	974.78	21.69	0.84	669.61	13.86
开花期	SVR	0.84	629.47	10.76	0.87	565.00	9.47
	RF	0.92	500.96	8.43	0.94	457.02	7.59
	GPR	0.93	537.82	9.27	0.94	382.32	6.34
灌浆期	SVR	0.84	1 018.09	20.67	0.85	884.71	17.36
	RF	0.89	954.50	19.67	0.90	911.92	18.60
	GPR	0.85	943.25	18.63	0.91	613.97	11.61

节期和灌浆期利用随机森林算法进行生物量估算效果最好，拔节期模型验证 R^2、RMSE、nRMSE 分别为 0.40、505.83 kg/hm^2、19.74%，灌浆期模型验证 R^2、RMSE、nRMSE 分别为 0.89、954.50 kg/hm^2、19.67%。孕穗期和开花期利用高斯过程回归算法对生物量估算效果最佳，孕穗期模型验证 R^2、RMSE、nRMSE 分别为 0.83、974.78 kg/hm^2、21.69%，开花期模型验证 R^2、RMSE、nRMSE 分别为 0.93、537.82 kg/hm^2、9.27%。综合考虑模型建立和模型验证的评价指标结果以及模型的估算能力，

开花期利用高斯过程回归算法构建模型具有最优估算精度。

（2）将叶面积指数信息融合相关性较好的小波能量系数，分别利用支持向量回归算法、随机森林回归算法、高斯过程回归算法进行冬小麦生物量估算，模型建立和模型验证散点图如图 4 所示。4 个生育期均利用高斯过程回归算法进行生物量估算效果最好，拔节期模型验证 R^2、RMSE、nRMSE 为 0.77、363.73 kg/hm^2、13.87%，孕穗期模型验证 R^2、RMSE、nRMSE 为 0.84、669.61 kg/hm^2、13.86%，开花期模型验证 R^2、RMSE、nRMSE 为 0.94、382.32 kg/hm^2、6.34%，灌浆期模型验证 R^2、RMSE、nRMSE 分别为 0.91、613.97 kg/hm^2、11.61%。综合考虑模型建立和模型验证的评价指标结果以及模型的估算能力，表明在开花期利用高斯过程回归构建的模型具有最优估算精度。通过对比分析，叶面积指数耦合小波能量系数估算生物量，在一定程度上能够克服冠层光谱饱和现象，提高模型估算精度。

（3）与仅仅基于小波能量系数相比，耦合叶面积指数数据，利用支持向量回归、随机森林回归、高斯过程回归方法分析进行生物量估算时，拔节期验证的 R^2 分别提高 0.21、0.33、0.38，孕穗期验证的 R^2

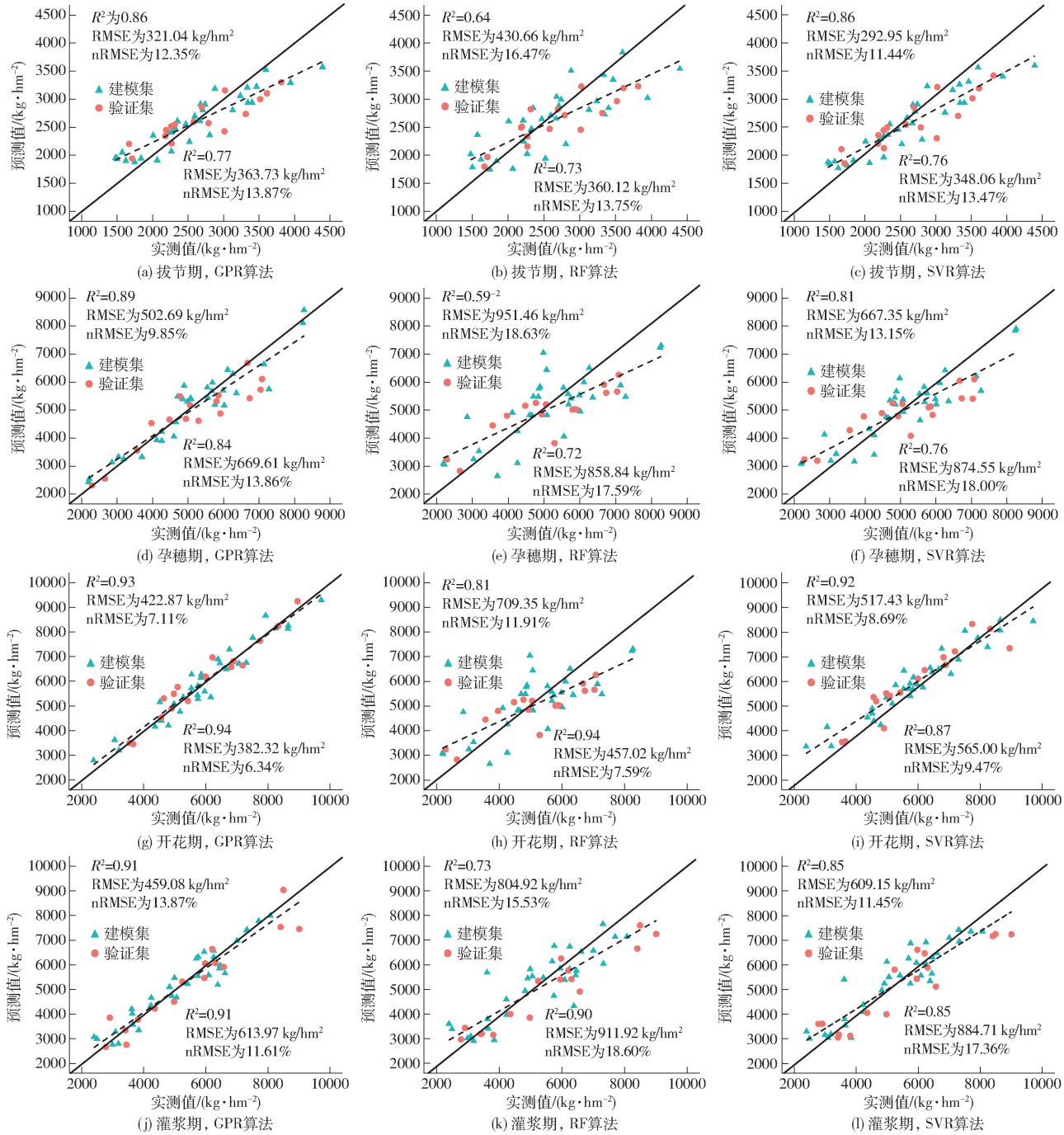

图 4　小波能量系数耦合叶面积指数估算模型和验证模型散点图

Fig. 4　Scatter plots of wavelet energy coefficient coupled leaf area index estimation model and verification model

分别提高 0.01、0.02、0.01,开花期验证的 R^2 分别提高 0.03、0.02、0.01,拔节期验证的 R^2 分别提高 0.01、0.01、0.06,表明小波能量系数耦合叶面积指数数据,可以提高生物量估算精度。

3 讨论

将 4 个生育期原始光谱、小波能量系数分别与生物量进行相关性分析,经对比可知,原始光谱和小波能量系数与生物量的相关性绝对值均呈先升高后降低的变化趋势,开花期的相关系数绝对值最大。可能的原因是从拔节期到开花期,叶片数量增加和叶面积扩大,光合速率大大提高,使得各构件生物量迅速积累和增加,冬小麦生物量在开花期达到最大,使得原始光谱和小波能量系数与生物量的关系由弱变强。开花期到灌浆期,即成熟初期,叶片和茎秆逐渐老化变黄,植被覆盖度明显降低,光合作用能力下降,生成的光合产物低于自身消耗,乃至脱落,小麦生物量下降,使得上述 2 个变量与生物量的相关性由强变弱。

每个生育期提取的小波能量系数与生物量的相关性都高于原始高光谱与生物量的相关性,小波能量系数与生物量的相关系数绝对值最大值较原始光谱与生物量的相关系数绝对值最大值分别提高了 0.06、0.1、0.07 和 0.03,这是因为高光谱数据采集过程中,由于受土壤背景、环境因素等影响,导致原始光谱存在噪声,影响敏感光谱信息提取,利用小波变换技术可以细化光谱信息,深度挖掘潜在的敏感光谱信息,凸显出更多与生物量相关的有效信息,使得筛选出的小波能量系数与生物量的相关性更高。

利用原始高光谱数据估算生物量的困难在于提取原始高光谱数据中对生物量较敏感的光谱波段,以往基于高光谱数据的生物量估算模型大多基于植被指数。本研究采用 SVR、RF 和 GPR 3 种算法结合小波能量系数、小波能量系数耦合叶面积指数构建生物量估算模型,实现对冬小麦生物量的估算。研究结果显示,4 个生育期利用小波能量系数耦合叶面积指数基于 GPR 算法构建的模型对生物量估算效果最好,建模和验证精度均较高,4 个生育期建模 R^2 分别为 0.86、0.89、0.93、0.91,验证 R^2 分别为 0.77、0.84、0.94、0.91,这与增加叶面积指数作为输入变量有关,叶面积指数包含作物冠层结构中非常重要的信息,与生物量高度相关,是决定作物光合作用强弱的重要变量之一,研究结果表明耦合叶面积指数可以提高生物量估算精度,与文献[30-32]研究结果相一致。另外,利用 GPR 算法估算冬小麦生物量精度较高,原因可能是 GPR 作为一种概率方法,训练模型通过拟合均值和协方差函数找到所有训练数据的单独函数,与其他算法学习方法相比,更适合训练小样本模型,且受生育期的影响较其他两种算法小。

4 结束语

基于小波能量系数,同时耦合叶面积指数数据,利用支持向量回归、随机森林回归、高斯过程回归方法,构建冬小麦生物量估算模型,并对模型精度进行验证。结果表明:4 个生育期,开花期生物量估算精度最高;利用支持向量回归、随机森林回归、高斯过程回归方法估算生物量,高斯过程回归算法估算生物量效果最好;小波能量系数耦合叶面积指数与仅使用叶面积指数为模型因子相比,生物量估算精度得到提高,构建的验证模型拟合性和稳定性都较好。

参 考 文 献

[1] CLEMENT A. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs[J]. Remote Sensing, 2013, 5(2):949-981.

[2] 刘明星, 李长春, 李振海, 等. 基于高光谱遥感与 SAFY 模型的冬小麦地上生物量估算[J/OL]. 农业机械学报, 2020, 51(2): 192-202,220.
LIU Mingxing, LI Changchun, LI Zhenhai, et al. Estimation of dry aerial mass of winter wheat based on coupled hyperspectral remote sensing and SAFY model[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 192-202,220. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag=1&file_no=2002021&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2020.02.021. (in Chinese)

[3] 陈仲新, 任建强, 唐华俊, 等. 农业遥感研究应用进展与展望[J]. 遥感学报, 2016, 20(5): 748-767.
CHEN Zhongxin, REN Jianqiang, TANG Huajun, et al. Progress and prospectives on agricultural remote sensing research and applications in China[J]. National Remote Sensing Bulletin, 2016, 20(5): 748-767. (in Chinese)

[4] 李金帅. 遥感技术在农业中的应用[J]. 农业与技术, 2021, 41(11): 61-64.
LI Jinshuai. Application of remote sensing technology in agriculture[J]. Agriculture and Technology, 2021, 41(11): 61-64. (in Chinese)

[5] 孙奇, 关琳琳, 焦全军, 等. 基于植被指数融合的冬小麦生物量反演研究[J]. 遥感技术与应用, 2021, 36(2): 391-399.
SUN Qi, GUAN Linlin, JIAO Qunjun, et al. Research on retrieving biomass of winter wheat based on fusing vegetation index

[J]. Remote Sensing Technology and Application, 2021, 36(2): 391 – 399. (in Chinese)

[6] 李岚涛, 郭宇龙, 韩鹏, 等. 基于高光谱的冬小麦不同生育时期地上部生物量监测[J]. 麦类作物学报, 2021, 41(7): 904 – 913.
LI Lantao, GUO Yulong, HAN Peng, et al. Estimation of shoot biomass at different growth stages of winter wheat based on hyperspectral reflectance[J]. Journal of Triticeae Crops, 2021, 41(7): 904 – 913. (in Chinese)

[7] 郭超凡, 陈泽威, 张志高. 基于最优模型选择的牧草地上生物量遥感估算研究[J]. 草地学报, 2021, 29(5):946 – 955.
GUO Chaofan, CHEN Zewei, ZHANG Zhigao. Research on remote sensing estimation of forage above-ground biomass based on optimal model selection[J]. Acta Agrestia Sinica, 2021, 29(5):946 – 955. (in Chinese)

[8] 陶惠林, 冯海宽, 徐良骥, 等. 基于无人机高光谱遥感数据的冬小麦生物量估算[J]. 江苏农业学报, 2020, 36(5): 1154 – 1162.
TAO Huilin, FENG Haikuan, XU Liangji, et al. Winter wheat biomass estimation based on hyperspectral remote sensing data of unmanned aerial vehicle(UAV)[J]. Jiangsu Journal of Agricultural Sciences, 2020, 36(5): 1154 – 1162. (in Chinese)

[9] 刘杨, 孙乾, 冯海宽, 等. 基于小波分析的马铃薯地上生物量估算[J]. 光谱学与光谱分析, 2021, 41(4): 1205 – 1212.
LIU Yang, SUN Qian, FENG Haikuan, et al. Estimation of above-ground biomass of potato based on wavelet analysis[J]. Spectroscopy and Spectral Analysis, 2021, 41(4): 1205 – 1212. (in Chinese)

[10] 李长春, 施锦锦, 马春艳, 等. 基于小波变换和分数阶微分的冬小麦叶绿素含量估算[J/OL]. 农业机械学报, 2021, 52(8):172 – 182.
LI Changchun, SHI Jinjin, MA Chunyan, et al. Estimation of chlorophyll content in winter wheat based on wavelet transform and fractional differential[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(8): 172 – 182. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20210817&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2021.08.017. (in Chinese)

[11] 梁栋, 杨勤英, 黄文江, 等. 基于小波变换与支持向量机回归的冬小麦叶面积指数估算[J]. 红外与激光工程, 2015, 44(1): 335 – 340.
LIANG Dong, YANG Qinying, HUANG Wenjiang, et al. Estimation of leaf area index based on wavelet transform and support vector machine regression in winter wheat[J]. Infrared and Laser Engineering, 2015, 44(1): 335 – 340. (in Chinese)

[12] 刘杨, 冯海宽, 孙乾, 等. 基于无人机高光谱分数阶微分的马铃薯地上生物量估算[J/OL]. 农业机械学报, 2020, 51(12): 202 – 211.
LIU Yang, FENG Haikuan, SUN Qian, et al. Estimation of potato above-ground biomass based on fractional differential of UAV hyperspectral[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(12): 202 – 211. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20201222&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2020.12.022. (in Chinese)

[13] 茹克亚·萨吾提, 阿不都艾尼·阿不里, 尼加提·卡斯木, 等. 基于分数阶微分的春小麦叶绿素含量高光谱估算[J]. 麦类作物学报, 2019, 39(6): 738 – 746.
RUKEYA Sawuti, ANUAINI Abuli, NIJATI Kasimu, et al. Spectral estimation of chlorophyll content in spring wheat leaves based on fractional differential[J]. Journal of Triticeae Crops, 2019, 39(6): 738 – 746. (in Chinese)

[14] 吾木提·艾山江, 买买提·沙吾提, 马春玥. 基于分数阶微分和连续投影算法–反向传播神经网络的小麦叶片含水量高光谱估算[J]. 激光与光电子学进展, 2019, 56(15): 251 – 259.
UMUT Hasan, MAMAT Sawut, MA Chunyue. Hyperspectral estimation of wheat leaf water content using fractional differentials and successive projection algorithm back propagation neural network[J]. Laser & Optoelectronics Progress, 2019, 56(15): 251 – 259. (in Chinese)

[15] LUO Shanjun, HE Yingbin, WANG Zhuozhuo, et al. Comparison of the retrieving precision of potato leaf area index derived from several vegetation indices and spectral parameters of the continuum removal method[J]. Taylor & Francis, 2019, 52(1): 155 – 168.

[16] 郑煜, 常庆瑞, 王婷婷, 等. 基于连续统去除和偏最小二乘回归的油菜 SPAD 高光谱估算[J]. 西北农林科技大学学报(自然科学版), 2019, 47(8): 37 – 45.
ZHENG Yu, CHANG Qingrui, WANG Tingting, et al. Hyperspectral estimation of SPAD value in oilseed rape based on continuum removal and partial least squares regression[J]. Journal of Northwest A&F University(Natural Science Edition), 2019, 47(8): 37 – 45. (in Chinese)

[17] 李旭青, 李龙, 庄连英, 等. 基于小波变换和 BP 神经网络的水稻冠层重金属含量反演[J/OL]. 农业机械学报, 2019, 50(6):226 – 232.
LI Xuqing, LI Long, ZHUANG Lianying, et al. Inversion of heavy metal content in rice canopy based on wavelet transform and BP neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(6):226 – 232. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20190625&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2019.06.025. (in Chinese)

[18] 彭咏石, 陈水森, 陈金月, 等. 基于连续小波系数的叶绿素 a 浓度估测模型[J]. 激光与光电子学进展, 2021, 58(8): 431 – 439.

PENG Yongshi, CHEN Shuisen, CHEN Jinyue, et al. Estimation model of chlorophyll a concentration based on continuous wavelet coefficients[J]. Laser & Optoelectronics Progress, 2021, 58(8): 431 – 439. (in Chinese)

[19] 苗梦珂, 王宝山, 李长春, 等. 基于连续小波变换的冬小麦叶片最大净光合速率遥感估算[J]. 江苏农业学报, 2020, 36(3): 544 – 552.
MIAO Mengke, WANG Baoshan, LI Changchun, et al. Remote sensing estimation of the maximum net photosynthetic rate of winter wheat leaves based on continuous wavelet transform[J]. Jiangsu Journal of Agricultural Sciences, 2020, 36(3): 544 – 552. (in Chinese)

[20] 章涛, 于雷, 易军, 等. 高光谱小波能量特征估测土壤有机质含量[J]. 光谱学与光谱分析, 2019, 39(10): 3217 – 3222.
ZHANG Tao, YU Lei, YI Jun, et al. Determination of soil organic matter content based on hyperspectral wavelet energy features[J]. Spectroscopy and Spectral Analysis, 2019, 39(10): 3217 – 3222. (in Chinese)

[21] RASMUSSEN C E, WILLIAMS C K I. Gaussian processes for machine learning[M]. The MIT Press, 2018.

[22] 王娇娇, 宋晓宇, 梅新, 等. 基于高斯回归分析的水稻氮素敏感波段筛选及含量估算[J]. 光谱学与光谱分析, 2021, 41(6): 1722 – 1729.
WANG Jiaojiao, SONG Xiaoyu, MEI Xin, et al. Sensitive bands selection and nitrogen content monitoring of rice based on Gaussian regression analysis[J]. Spectroscopy and Spectral Analysis, 2021, 41(6): 1722 – 1729. (in Chinese)

[23] 赵梦恩. 改进高斯过程回归算法及其应用研究[D]. 杭州: 浙江理工大学, 2019.
ZHAO Meng'en. Improved Gaussian process regression algorithm and its application[D]. Hangzhou: Zhejiang Sci-Tech University, 2019. (in Chinese)

[24] VLADIMIR N. The nature of statistical learning theory[M]. New York: Springer, 2000.

[25] 林卉, 梁亮, 张连蓬, 等. 基于支持向量机回归算法的小麦叶面积指数高光谱遥感反演[J]. 农业工程学报, 2013, 29(11):139 – 146.
LIN Hui, LIANG Liang, ZHANG Lianpeng, et al. Wheat leaf area index inversion with hyperspectral remote sensing based on support vector regression algorithm[J]. Transactions of the CSAE, 2013, 29(11):139 – 146. (in Chinese)

[26] 方匡南, 吴见彬, 朱建平, 等. 随机森林方法研究综述[J]. 统计与信息论坛, 2011, 26(3): 32 – 38.
FANG Kuangnan, WU Jianbin, ZHU Jianping, et al. A review of technologies on random forests[J]. Journal of Statistics and Information, 2011, 26(3): 32 – 38. (in Chinese)

[27] KYOKO O. Machine learning[M]. Tritech Digital Media, 2018.

[28] 马立平. 回归分析[M]. 北京: 机械工业出版社, 2014.

[29] 侯学会, 牛铮, 黄妮, 等. 小麦生物量和真实叶面积指数的高光谱遥感估算模型[J]. 国土资源遥感, 2012, 24(4): 30 – 35.
HOU Xuehui, NIU Zheng, HUANG Ni, et al. The hyperspectral remote sensing estimation models of total biomass and true LAI of wheat[J]. Remote Sensing for Land & Resources, 2012, 24(4): 30 – 35. (in Chinese)

[30] ROCIO B, JOSE F O, DAVID H, et al. Onion biomass monitoring using UAV-based RGB imaging[J]. Precision Agriculture, 2018, 19(5):840 – 857.

[31] ANGELA K, HEATHER M N, DAVID L, et al. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops[J]. International Journal of Applied Earth Observations and Geoinformation, 2015, 34: 235 – 248.

[32] JIN Xiuliang, LI Zhenhai, FENG Haikuan, et al. Deep neural network algorithm for estimating maize biomass based on simulated Sentinel 2A vegetation indices and leaf area index[J]. The Crop Journal, 2020, 8(1):87 – 97.