DOI: 10.3724/SP.J.1006.2020.01004

MDBPSO 算法优化的全波段光谱数据协同冠层 SIF 监测小麦条锈病

白宗藩¹ 竞霞^{1,*} 张腾¹ 董莹莹²

¹ 西安科技大学测绘科学与技术学院, 陕西西安 710054; ² 中国科学院遥感与数字地球研究所数字地球重点实验室, 北京 100094

摘要: 为了从全波段光谱数据中提取对小麦条锈病敏感的特征参量, 提高小麦条锈病遥感探测模型的运行效率和精度, 本文首先从惯性权重和粒子更新方式两个方面对传统离散粒子群算法(discrete binary particle swarm optimization, DBPSO)进行改进, 利用改进离散粒子群算法(modified discrete binary particle swarm optimization, MDBPSO)从全波段光谱数据中优选遥感探测小麦条锈病严重度的特征变量, 然后与冠层日光诱导叶绿素荧光(solar-induced chlorophyll fluorescence, SIF)数据相结合作为自变量分别利用随机森林(random forest, RF)和后向传播(back propagation, BP)神经网络算法构建小麦条锈病遥感探测模型, 并将其与相关系数(correlation coefficient, CC)分析法和 DBPSO 算法提取特征参量构建模型的精度进行对比分析。结果表明: (1) MDBPSO 算法比传统 DBPSO 算法具有更快的收敛速度和更高的寻优精度, 改进前后其迭代次数从 395 次减少到 156 次, 最优适应度函数(optimum fitness value, OFV)值从 0.145 减小到 0.127。(2)采用 MDBPSO 算法选择特征变量时, RF 和 BP 神经网络两种方法构建的模型精度均高于 CC 分析法和 DBPSO 算法, 其中 RF 算法预测病情指数(disease index, DI)值和实测 DI 值间的检验集决定系数(validation set determination coefficient, R^2_v)比 CC 分析法和 DBPSO 算法分别提高了 9%和 3%, 均方根误差(validation set root mean square error, RMSE_v)分别降低了 28%和 11%, BP 神经网络算法预测 DI 值和实测 DI 值间的 R^2_v 比 CC 分析法和 DBPSO 算法分别提高了 13%和 6%, RMSE_v 分别降低了 21%和 10%, 利用 MDBPSO 算法优选特征参量能够提高小麦条锈病的遥感探测精度。(3)在 MDBPSO、DBPSO 和 CC 分析法 3 种特征选择算法中, RF 算法构建的模型精度均高于 BP 神经网络算法, 其中 RF 模型预测 DI 值和实测 DI 值间的 R^2_v 比 BP 神经网络算法至少提高了 7%, 平均提高了 9%, RMSE_v 至少降低了 15%, 平均降低了 20%。以 MDBPSO 算法优选的特征参量为自变量利用 RF 方法构建的小麦条锈病遥感探测的 MDBPSO-RF 模型是小麦条锈病遥感探测适宜模型, 该研究结果为进一步实现作物健康状况大面积高精度遥感监测提供了新的思路。

关键词: 全波段反射光谱; 改进离散粒子群; 日光诱导叶绿素荧光; 小麦条锈病; 特征波段

Canopy SIF synergize with total spectral reflectance optimized by the MDBPSO algorithm to monitor wheat stripe rust

BAI Zong-Fan¹, JING Xia^{1,*}, ZHANG Teng¹, and DONG Ying-Ying²

¹ College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, Shaanxi, China; ² Institute of Remote Sensing and Digital Earth, Chinese Academy of Science, Beijing 100094, China

Abstract: In order to extract the characteristic parameters sensitive to wheat stripe rust from total spectral reflectance and improve the operation efficiency and estimation accuracy of the wheat stripe rust remote sensing detection model, this paper improved the traditional discrete binary particle swarm optimization (DBPSO) algorithm from two aspects: inertia weight and particle update method. The modified discrete binary particle swarm optimization (MDBPSO) algorithm was used to select the characteristic parameters for the severity of wheat stripe rust from the total spectral reflectance. The selected characteristic variables were combined with canopy solar-induced chlorophyll fluorescence (SIF) data as independent variables to construct wheat stripe rust esti-

本研究由国家自然科学基金项目(41601467)资助。

This study was supported by the National Natural Science Foundation of China (41601467).

* 通信作者(Corresponding author): 竞霞, E-mail: jingxiastust@163.com

第一作者联系方式: E-mail: bzfl529@163.com

Received (收稿日期): 2020-01-09; Accepted (接受日期): 2020-03-24; Published online (网络出版日期): 2020-04-17.

URL: <http://kns.cnki.net/kcms/detail/11.1809.S.20200417.1057.002.html>

mation model. Random forest (RF) algorithm and back propagation (BP) neural network algorithm were used as model construction method to compare and analyze the accuracy of the feature parameter construction model extracted by the correlation coefficient (CC) analysis method and the DBPSO algorithm. The MDBPSO algorithm had faster convergence speed and higher optimization accuracy than the DBPSO algorithm. The number of iterations before and after the improvement was reduced from 395 to 156. The optimal fitness function (OFV) value decreased from 0.145 to 0.127. When the MDBPSO algorithm was used to select feature variables, the accuracy of the models constructed by the two methods of RF and BP neural networks was higher than that by CC analysis and DBPSO. The validation set determination coefficient (R^2_v) between the predicted disease index (DI) value and the measured DI value of the RF algorithm was 9% and 3% higher than that of the CC analysis method and the DBPSO algorithm, the validation set root mean square error (RMSE_v) was reduced by 28% and 11%, respectively. The R^2_v between the predicted disease index (DI) value and the measured DI value of the BP neural network algorithm was 13% and 6% higher than that of the CC analysis method and DBPSO algorithm, respectively, and the RMSE_v was reduced by 21% and 10% respectively. The MDBPSO algorithm can improve the remote sensing detection accuracy of wheat stripe rust. Among the three feature selection algorithms of MDBPSO, DBPSO, and CC analysis, the accuracy of the model constructed by the RF algorithm was higher than that by the BP neural network algorithm. The R^2_v between the predicted DI value and the measured DI value of the RF model was at least 7% higher than that of BP neural network algorithm, with an average increase of 9%; RMSE_v had reduced by at least 15%, with an average reduction of 20%. The MDBPSO-RF model for wheat stripe rust remote sensing detection constructed by RF method using the characteristic parameters selected by the MDBPSO algorithm as independent variables is an appropriate model for wheat stripe rust remote sensing detection. The research results provide new ideas for further realizing large-area high-precision remote sensing monitoring of crop health.

Keywords: total spectral reflectance; modified particle swarm optimization; solar-induced chlorophyll fluorescence; stripe rust of wheat; feature band

小麦条锈病(*Puccinia striiformis*)是一种气流传播病害,一旦爆发会造成小麦大面积减产^[1]。因此,小麦条锈病的及时监测和防治对我国农业发展具有重要意义。与传统田间取样的病害调查方法相比,遥感技术具有宏观、快速、无损等优势,已被广泛应用于作物病害的探测^[1-8],尤其是高光谱遥感数据包含丰富的光谱信息,能够很好地反映作物的生化特性^[9],提高病害监测精度^[10]。但已有的研究主要集中于利用少量波段信息计算反射率光谱指数^[5-7]或吸收特征^[1]等算法进行作物病害监测,这些方法在一定程度上丢失了对小麦病害遥感探测的有用信息,且未考虑作物冠层几何结构及大气等因素导致的冠层反射光谱的时空变化^[11],因此针对不同时空下的小麦条锈病,如何利用全波段信息获取动态的冠层光谱特征变量显得尤为重要。刘琦等^[10,12]在 325~1075 nm 全波段范围内成功建立模拟识别小麦条锈病的模型。但直接使用全波段数据,波段之间具有较高的相关性^[13],数据冗余不仅加大了计算机的运行时间,而且对模型精度具有一定影响,因此从全波段数据中提取有效的特征变量成为模型构建过程中至关重要的一步。粒子群算法(particle swarm optimization, PSO)是由 Eberhart 和 Kennedy^[14]于 1995 年提出的一种群体智能随机搜索算法,具有鲁棒性强、参数设置少及收敛速度快的优点^[15]。已有学者将粒子群算法应用于特征变量选择中并取得了理想的结果^[16-17]。但传统的粒子群算法在处理高光谱的全波段数据时易陷入局部最优,产生“早熟”收敛^[18],

为了解决这一问题,本文从惯性权重和粒子更新方式两方面对传统离散粒子群算法进行改进,以提高小麦条锈病遥感探测精度。

模型精度除受所选特征参量的影响外,建模算法也是影响小麦条锈病遥感探测精度的重要因素,选择合适的模型算法有助于提高小麦条锈病严重度的遥感估测精度。与传统的经验统计模型相比,机器学习模型具有良好的非线性拟合能力和泛化能力,能够通过不断训练样本数据使目标达到最优^[19]。在众多机器学习算法中 BP 神经网络算法具有通过学习获取知识并解决问题的能力^[20],RF 算法具有计算量小、能有效避免过拟合的发生以及良好的噪声容忍性等优点^[21],为此本文拟在利用 MDBPSO 算法优选全波段光谱特征的基础上分别采用 RF 和 BP 神经网络算法构建小麦条锈病遥感探测模型。

反射率光谱数据能够敏感地反映植被生化组分信息,但对作物光合活性不敏感,难以揭示植被光合生理状态^[22],而冠层 SIF 数据则能够直接探测到植被的光合作用状态^[23],小麦受条锈病菌侵染后,其水分及叶绿素含量、光合速率和光能转换率等一些生理生化指标均会发生变化^[13],综合利用反射率光谱在生化参数探测方面的优势和叶绿素荧光在光合生理诊断的优势可以提高小麦条锈病的探测精度^[24-25]。基于此,本文利用 MDBPSO 算法从全波段反射率光谱数据中选择对小麦条锈病敏感的特征因子,并将其与冠层 SIF 数据结合作为模型的输入变量分别通过 RF 和 BP 神经网络 2 种机器学习算法构

建小麦条锈病严重度的遥感估测模型,并将其与传统的DBPSO算法及CC分析法提取特征参量所建模型的精度进行对比分析,以确定遥感探测小麦条锈病的特征,以及模型构建的适宜算法。

1 材料与方法

1.1 试验设计

试验区位于河北省廊坊市中国农业科学院试验站(39°30'40"N, 11°63'62"E),小麦品种为对条锈病比较敏感的铭贤 169 号。试验区小麦平均种植密度为 113 棵 m^{-2} 。2018 年 4 月 7 日给试验田灌溉充足的水分,并于 4 月 9 日采用浓度为 9 mg 100 mL^{-1} 的孢子溶液对小麦进行条锈病接种。将试验区小麦分为健康组(编号为 A、D)和染病组(编号为 B、C),每个试验组的面积为 220 m^2 。将每个组分为 8 个样方($A_1 \sim A_8$ 、 $B_1 \sim B_8$ 、 $C_1 \sim C_8$ 、 $D_1 \sim D_8$)。

1.2 数据获取

1.2.1 冠层光谱测定与处理 分别于 2018 年 5 月 18 日、5 月 24 日和 5 月 30 日 3 个时期测定小麦条锈病不同病情指数下的冠层光谱数据,测量所使用的仪器为 ASD Field Spec 4 光谱仪和 QE 65 pro 光谱仪。其中 ASD Field Spec 4 测量的光谱范围为 350~2500 nm,在 350~1000 nm 的波段范围内光谱分辨率为 3 nm,采样间隔为 1.4 nm;在 1000~2500 nm 波段范围内光谱分辨率为 10 nm,采样间隔为 2 nm。QE 65 pro 光谱仪的光谱分辨率为 0.31 nm,采样间隔为 0.155 nm,光谱范围为 685~805 nm。冠层光谱测量时间为北京时间 11:00~12:30,探头视场角为 25°,测量高度距离地面 1.3 m,每次测量前后均用标准 BaSO_4 参考板进行校正。然后由公式(1)计算冠层反射率值。

为了消除和减弱样本间因散射引起的光谱误差,利用标准正态校正(stand normal variate, SNV)方法对原始光谱数据进行预处理,假设每个波段的光谱吸收值均满足正态分布,并利用此假设对每条光谱进行校正。SNV 的变换公式^[26]如式(2)所示。

$$R = \frac{L_{\text{target}}}{L_{\text{board}}} \times R_{\text{board}} \quad (1)$$

式中, R 为原始冠层反射率, L_{target} 为目标辐亮度,

$$\bar{F}_{\text{in}} = \frac{(\omega_{\text{left}} \times I_{\text{left}} + \omega_{\text{right}} \times I_{\text{right}}) \times L_{\text{in}} - I_{\text{in}} \times (\omega_{\text{left}} \times L_{\text{left}} + \omega_{\text{right}} \times L_{\text{right}})}{(\omega_{\text{left}} \times I_{\text{left}} + \omega_{\text{right}} \times I_{\text{right}}) - I_{\text{in}}} \quad (6)$$

其中, $\omega_{\text{left}} = \frac{\lambda_{\text{right}} - \lambda_{\text{in}}}{\lambda_{\text{right}} - \lambda_{\text{left}}}$, $\omega_{\text{right}} = \frac{\lambda_{\text{in}} - \lambda_{\text{left}}}{\lambda_{\text{right}} - \lambda_{\text{left}}}$

L_{board} 为参考板辐亮度, R_{board} 为参考板反射率。

$$x = \frac{R - \bar{R}}{\sqrt{\frac{\sum_{k=1}^m (R_k - \bar{R})^2}{(m-1)}}} \quad (2)$$

式中, x 为预处理后的光谱数据, m 为波段的个数, $\bar{R}$ 为光谱数据平均值, $\bar{R} = \frac{\sum_{k=1}^m x_k}{m}$ 。

1.2.2 病情指数调查 冠层病情指数的调查采用 5 点取样法,在每块小区内选取对称的 5 点,每点约 1 m^2 ,随机选取 30 株小麦分别调查其发病情况。参照国家标准“小麦条锈病测报技术规范”(GB/T15795)量化病情指数。单叶严重度分为 9 个梯度,即 0、1%、10%、20%、30%、45%、60%、80%和 100%的叶片病斑覆盖,分别记录不同病情严重度下的小麦叶片数,按式(3)计算测试群体的病情指数^[27]。

$$\text{DI} = \frac{\sum (x \times f)}{n \times \sum f} \quad (3)$$

式中, DI 为病情指数; x 为梯度级值; n 为最高梯度等级值; f 为各梯度的叶片数。

1.2.3 日光诱导叶绿素荧光提取 利用辐亮度基于夫琅和费暗线原理的日光诱导叶绿素荧光算法,主要包括 FLD (fraunhofer line discrimination)和 3FLD (3 bands FLD)等^[28]。标准 FLD 算法是基于吸收线内外反射率和透过率相等的假设来估测日光诱导叶绿素荧光强度,但由于吸收线内外波段的反射率和荧光值实际上存在差异,影响了荧光的估测精度^[29]。3FLD 荧光预测算法认为在吸收线波段周围叶绿素荧光和反射率光谱是线性变化的,利用吸收线左右各一个波段的加权平均值代替标准 FLD 算法中的单一波段值,从而在一定程度上减小标准 FLD 方法中荧光和反射率恒定假设所带来的误差,提高了冠层 SIF 的预测精度^[30],而且已有研究表明 3FLD 算法是估测 SIF 最鲁棒性算法^[31],基于此本文采用 3FLD 算法估测叶绿素荧光强度,计算公式如式(4)~(6)所示。

$$I_{\text{out}} = \omega_{\text{left}} \times I_{\text{left}} + \omega_{\text{right}} \times I_{\text{right}} \quad (4)$$

$$L_{\text{out}} = \omega_{\text{left}} \times L_{\text{left}} + \omega_{\text{right}} \times L_{\text{right}} \quad (5)$$

式中, λ_{in} 、 λ_{left} 、 λ_{right} 为吸收线内、左、右波段的波长; ω_{left} 、 ω_{right} 为吸收线左右 2 个参考波段所占的权

重; I_{left} 、 I_{right} 为吸收线左右的太阳辐照度光谱强度 ($\mu\text{W cm}^{-2} \text{ nm}^{-1}$); L_{left} 、 L_{right} 为吸收线左右的植被冠层反射辐亮度光谱强度 ($\mu\text{W cm}^{-2} \text{ nm}^{-1} \text{ sr}^{-1}$)。

为了减少太阳光强度对冠层 SIF 估测结果的影响, 将计算得到的 SIF 强度除以夫琅和费吸收线内参考板的太阳入射辐照度, 得到该吸收线处的冠层 SIF 相对强度^[32]。如式(7)所示。

$$F_{\text{相对}} = \bar{F}_{\text{in}} / I_{\text{in}} \quad (7)$$

式中, $F_{\text{相对}}$ 为日光诱导叶绿素荧光相对强度, $\bar{F}_{\text{in}}$ 为日光诱导叶绿素荧光绝对强度, I_{in} 为参考板获取的夫琅和费吸收线内的太阳入射辐照度。

1.3 改进离散粒子群算法

1.3.1 离散粒子群算法 离散粒子群算法是在粒子群算法的基础上提出的一种离散群体智能随机搜索算法^[14], 具有结构简单, 参数较少及易于实现等优点^[33]。假定 N 维搜索空间中第 i 个粒子的速度和位置分别为 $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$, $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, 则粒子 i 在第 d 维的速度和位置的更新公式如式(8)~(10)所示。

$$v_{id}^{k+1} = wv_{id}^k + c_1 r_1 (p_{id}^k - x_{id}^k) + c_2 r_2 (p_{gd}^k - x_{id}^k) \quad (8)$$

$$x_{id}^{k+1} = \begin{cases} 1, \text{rand} < S(v_{id}^k) \\ 0, \text{其他} \end{cases} \quad (9)$$

$$S(v_{id}^{k+1}) = 1 / [1 + \exp(-v_{id}^{k+1})] \quad (10)$$

式中, w 是权重系数; c_1 和 c_2 为学习因子; r_1 和 r_2 是随机数, 服从 $[0,1]$ 上的均匀分布; v_{id}^k 和 v_{id}^{k+1} 分别表示粒子更新 k 次和 $k+1$ 次时的速度; x_{id}^k 和 x_{id}^{k+1} 分别表示粒子更新 k 次和 $k+1$ 次时的位置; p_{gd}^k 和 p_{id}^k 分别表示粒子第 k 次搜索时历史最优解和全局最优解所对应的位置。

1.3.2 离散粒子群改进策略 传统的离散粒子群算法一旦达到了收敛, 所有的粒子都会朝着最优解靠拢, 失去了粒子的多样性, 影响了算法的收敛速度及精度。而且传统离散粒子群算法缺乏对速度的动态调整容易发生“早熟”收敛现象, 从而导致算法的收敛精度较低, 基于此, 本文分别从惯性权重和粒子更新方式两个方面对传统粒子群算法进行改进。

1) 惯性权重 w : 惯性权重 w 值越大越有利于加快收敛速度, 提高全局搜索能力, 但难以得到精确解, 反之则有利于局部搜索且能够得到更为精确的解, 但收敛速度慢, 易陷入局部极值。为了

提高离散粒子群算法的全局搜索能力, 在搜索早期将 w 设置为相对较大的值, 在搜索过程中, 随着迭代次数的增加使 w 值逐渐减小, 从而得到精确的解^[34]。基于此, 本研究对 w 做了如式(11)所示的动态调整。

$$w = w_{\text{max}} - t(w_{\text{max}} - w_{\text{min}}) / t_{\text{max}} \quad (11)$$

式中 w_{max} 和 w_{min} 分别为 w 的最大值和最小值, 本文中 w_{max} 和 w_{min} 分别取 0.9 和 0.4^[35], t 为当前迭代次数, t_{max} 为最大迭代次数。

2) 粒子更新方式: 将粒子的速度限定在一定的区间范围内, 可以避免函数 $S(v_{id}^{k+1})$ 出现饱和。参考已有的研究, 本文将粒子的速度限定在 $[-4, 4]$ 之间^[17]。当粒子速度处于 $[-4, 4]$ 时, $S(v_{id}^{k+1})$ 的值处于 $[0, 1]$ 。基于此, 粒子的速度和位置的更新公式如式(12)~(13)所示。

$$S(v_{id}^{k+1}) = \begin{cases} 1 - \frac{1}{1 + \exp(-v_{id}^{k+1})} \\ 1 \\ 1 + \exp(-v_{id}^{k+1}) \end{cases}, -4 \leq v_{id}^{k+1} \leq 4 \quad (12)$$

$$x_{id}^{k+1} = \begin{cases} x_{id}^k, \text{rand}() < S(v_{id}^{k+1}) \\ 0, \text{rand}() \leq S(v_{id}^{k+1}), -4 \leq v_{id}^{k+1} \leq 0 \\ 1, \text{rand}() \leq S(v_{id}^{k+1}), 0 \leq v_{id}^{k+1} \leq 4 \end{cases} \quad (13)$$

式中各变量的含义与公式(8)~(10)一致。

MDBPSO 算法具有较好的局部搜索能力, 但是全局搜索能力相对较差, 为了使算法同时具有较好的局部搜索和全局搜索能力, 本文在迭代前期($t < 30\% t_{\text{max}}$)通过式(8)~(10)对粒子进行更新, 在迭代中后期($30\% t_{\text{max}} < t < t_{\text{max}}$)通过式(8)、(12)、(13)对粒子进行更新。

1.3.3 适应度函数设计 适应度函数是粒子群算法中判断粒子位置优劣的重要依据, 适应度函数的设计直接影响到算法的收敛速度以及能否找到最优解。本研究基于均方根误差 RMSE [式(14)]构建适应度函数, 并以此作为各粒子优胜劣汰的评判标准。将各粒子当前的适应度值与其个体历史最佳位置对应的适应度作比较, 如果当前位置的适应度值更小, 则用当前位置更新全局最佳位置。

$$F = \sqrt{\frac{1}{m} \sum_{i=1}^m (y - \hat{y})^2} \quad (14)$$

式中, m 为样本个数, y 为样本的实测值, $\hat{y}$ 为模型预测值。

1.4 模型构建与精度评价

以 CC 分析法、DBPSO 和 MDBPSO 三种算法提取的全波段光谱特征参量协同冠层 SIF 数据作为模型的输入变量构建小麦条锈病严重度估测模型, 并通过保留样本交叉检验的方式对模型精度进行评价, 以确定遥感探测小麦条锈病的特征变量优选算法及适宜模型。

1.4.1 模型构建方法 分别采用 RF 和 BP 神经网络两种算法构建小麦条锈病严重度估测模型, 其中 RF 算法的核心思想是通过 bootstrap 重采样的方法在原始训练集中抽取多个样本, 对每个抽取出的样本都进行决策树建模, 最后通过多数投票法得到最终的预测结果^[36]。通过对训练样本的多次仿真确定 RF 算法中决策树的数量(ntree)为 500, 内部节点随机选择属性个数(mtry)取默认值。BP 神经网络算法的核心思想是通过调整各神经元之间的权值, 将误差由隐含层向输入层逐层反传, 对误差函数进行“链式求导”, 使误差逼近最小值^[37], 本研究中 BP 神经网络算法采用 3 层网络标准结构, 其中隐含层设置为 5 个神经元, 最大训练次数为 5000, 训练间隔为 10, 最小均方根误差为 0.001, 学习步长为 0.1。

1.4.2 模型精度评价 为了能够充分利用样本中所有数据, 在有限样本容量下尽可能减少“过拟合”问题, 本文采用保留样本交叉检验的方式对模型进行构建和检验, 将 52 个样本数据分为训练集和检验集, 其中 39 个数据(35 个染病样本, 4 个健康样本)作为训练样本用于模型构建, 剩余的 13 个数据(12 个染病样本, 1 个健康样本)作为验证样本用以模型评价。选择病情严重度估测值与实测值之间的 R^2 和 RMSE 两个模型精度评价指标[式(15), (16)], 其中 R^2 越高, RMSE 越低, 则表示模型的估测精度越高。

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (15)$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (16)$$

式中 m 为样本个数, y 为样本的实测值, $\hat{y}$ 为模型预测值, $\bar{y}$ 为样本平均值。

2 结果与分析

2.1 特征变量优选

2.1.1 CC 分析法优选特征参量 在对小麦冠层原始反射率光谱数据进行 SNV 变换处理的基础上利用 CC 分析法对光谱反射率数据与小麦条锈病严重

度进行相关性分析, 得到 SNV 变换后不同波长下反射率——病情严重度间相关系数的变化曲线(图 1), 选择与病情严重度达到极显著相关(相关系数的绝对值高于 0.443)波段范围内的极值点作为小麦条锈病遥感探测的敏感波段。由图 1 可以看出, 400~540 nm、742~868 nm 范围内的反射率与病情指数呈极显著负相关, 561~731 nm、905~1000 nm 范围内的反射率与病情指数呈极显著正相关, 通过 CC 法选择出 50 个波段作为小麦条锈病遥感探测的特征变量。

图 1 SNV 变换后反射率-病情指数相关系数曲线图($n = 52$)

Fig. 1 Curve of correlation coefficient between reflectance and disease index after SNV transformation ($n = 52$)

图中虚线表示 0.1% 显著水平, $R_{0.001[52]} = 0.443$ 。

The dotted line in the figure indicates the significant at the 0.1% probability level, and $R_{0.001[52]} = 0.443$.

2.1.2 MDBPSO 算法优选特征参量 为了从全波段反射率光谱数据中确定小麦条锈病遥感探测的特征变量, 分别基于 MDBPSO 和 DBPSO 算法提取了小麦条锈病遥感探测的特征波段。首先通过对样本的多次仿真训练确定 DBPSO 和 MDBPSO 的最大迭代次数 $t_{\max}$ 分别为 400 和 200, MDBPSO 速度限定的最大值 $V_{\max}$ 和最小值 $V_{\min}$ 分别为 4 和 -4。然后在上述参数确定的基础上对比分析了 DBPSO 和 MDBPSO 两种算法下适应度函数的收敛情况以确定 OFV 值(图 2)。从图 2 可以看出, DBPSO 算法迭代 395 次可以得到最优解, OFV 最小值为 0.145。MDBPSO 算法得到的 OFV 最小值和迭代次数分别减少到了 0.127 和 156, MDBPSO 算法比 DBPSO 算法具有更快的收敛速度和更高的寻优精度。这是由于 MDBPSO 算法在迭代过程中惯性权重发生了相应的动态变化, 从而使得 OFV 的值变化较快, 在提高效率的同时也增加了寻优的精度。

依据已确定的 OFV 的最小值和迭代次数, 分别通过 DBPSO 和 MDBPSO 两种算法从全波段反射率光谱数据中提取小麦条锈病遥感探测的特征变量(图 3)。从图 3 中可以看出, DBPSO 和 MDBPSO 两

种算法提取特征变量的个数分别为 320 和 56, MDBPSO 算法极大减少了特征参量的个数, 降低了算法运行的时间。此外, DBPSO 算法选择的特征波

段分布集中, 波段之间仍然存在较高的相关性, 波段信息重叠的问题依然存在, 特征波段冗余度较高的问题依然没有得到较好的解决。

图 2 DBPSO 和 MDBPSO 算法的适应度函数收敛曲线

Fig. 2 Convergence curves of fitness functions for DBPSO and MDBPSO algorithms

图 3 DBPSO 和 MDBPSO 算法选择的特征波段

Fig. 3 Feature bands selected by the DBPSO and MDBPSO algorithms

2.2 构建及精度评价

分别利用 CC 分析法、传统 DBPSO 和 MDBPSO 法从全波段反射率光谱数据中选择特征波段作为遥感探测小麦条锈病反射率数据的模型输入变量, 并将其与冠层 SIF 数据相结合作为模型输入的自变量, 通过 RF 算法和 BP 神经网络算法构建小麦条锈病遥感探测模型(图 4), 图中横坐标为小麦条锈病严重度的估测值, 纵坐标为小麦条锈病严重度的实测值, 实线为 1:1 关系线。在病情严重度实测值-估测值构成的二维特征空间中, 散点的分布越接近于 1:1 关系线, 检验集和训练集的决定系数 R^2_V 和 R^2_T 越高, 均方根误差 $RMSE_V$ 和 $RMSE_T$ 越小, 模型的估测精度越高。

由图 4 的模型训练集估测结果可以看出, 无论是采用 RF 算法还是 BP 神经网络算法构建模型, 以 MDBPSO 算法优选的特征参量为自变量构建的模型精度较 CC 分析法和 DBPSO 算法均有一定程度的提高。此外在 CC、DBPSO 和 MDBPSO 这 3 种特征优

选算法中, RF 模型的预测精度均高于对应的 BP 神经网络模型, RF 算法构建模型的散点分布与对应的 BP 神经网络算法相比更加贴合于 1:1 关系线, 所以 RF 算法更适合小麦条锈病的遥感探测。

从模型检验集的数据中可以看出, 无论采用 BP 神经网络算法还是 RF 算法构建小麦条锈病的探测模型, MDBPSO 算法提取特征变量构建的模型精度均高于 CC 分析法和 DBPSO 算法, 当使用 RF 算法构建模型时 R^2_V 与 CC 分析法和 DBPSO 算法相比分别提高了 9% 和 3%, $RMSE_V$ 分别减少了 28% 和 11%; 当使用 BP 神经网络算法构建模型时 R^2_V 与 CC 分析法和 DBPSO 算法相比分别提高了 13% 和 6%, $RMSE_V$ 分别减少了 21% 和 10%。无论采用 CC 法、DBPSO 算法还是 MDBPSO 算法进行特征选择, RF 算法构建的模型精度均高于对应的 BP 神经网络算法, CC、DBPSO、MDBPSO 三种特征选择算法中 RF 模型的 R^2_V 比 BP 神经网络模型至少提高了 7%, 平均提高了 9%, $RMSE_V$ 比 BP 神经网络模型至少减

图 4 小麦条锈病病情指数实测值与预测值分布

Fig. 4 Distribution of measured and predicted values of wheat stripe rust disease index

— 1:1 关系线; ■ 训练集数据; △ 检验集数据。
— 1:1 relationship line; ■ training set; △ test set.

少了 15%，平均减少了 20%。综上所述，在所有的模型构建方法中，MDBPSO-RF 模型精度最高，更适合小麦条锈病病情指数的估测。

3 讨论

本文综合利用反射率光谱数据在作物生化参数探测方面的优势和叶绿素荧光在光合生理诊断方面的优势，分别采用 CC 分析法、DBPSO 和 MDBPSO 三种方法从全波段光谱数据中提取特征变量，并将其与冠层 SIF 数据相融合构建小麦条锈病遥感探测模型，分析不同特征变量提取方法对小麦条锈病遥感探测精度的影响。

MDBPSO 算法在模型效率方面较 DBPSO 算法具有一定提高，其中 OFV 最小值从 0.145 减小到 0.127，迭代次数从 395 次减小到 156 次。无论建模方法采用 BP 神经网络算法还是 RF 算法，MDBPSO 算法提取特征变量构建的模型精度均高于 CC 分析法和 DBPSO 算法，其中 BP 神经网络方法构建的 MDBPSO-BP 模型预测 DI 值和实测 DI 值间的 R^2_v 比 CC-BP 和 DBPSO-BP 分别提高了 13% 和 6%， $RMSE_v$

分别减小了 21% 和 10%，RF 方法构建的 MDBPSO-RF 模型预测 DI 值和实测 DI 值间的 R^2_v 比 CC-RF 和 DBPSO-RF 分别提高了 9% 和 3%， $RMSE_v$ 分别减小了 28% 和 11%。这是因为 CC 分析法依据小麦条锈严重度与光谱反射率的相关性选取的特征变量之间存在较高的相关性和共线性影响了模型的精度。与 CC 法相比，DBPSO 算法以模型的 RMSE 作为特征波段适用性的评判标准，选择的特征变量构建的模型精度得到了一定的改善。但与 MDBPSO 算法相比，DBPSO 算法的 OFV 较大，模型运行的效率较低，这是因为 MDBPSO 算法在 DBPSO 算法的基础上对惯性权重和粒子更新方式两个方面做出了改进，在选择特征变量时考虑了各个波段的光谱信息对模型贡献率的问题^[17]，提高了模型的反演精度。

分别以 CC 分析法、DBPSO 和 MDBPSO 3 种算法筛选的特征因子为自变量构建小麦条锈病严重度估测模型时，RF 算法构建的模型精度均高于 BP 神经网络算法，其中 CC-RF、DBPSO-RF、MDBPSO-RF 模型预测 DI 值和实测 DI 值之间的 R^2_v 比 CC-BP、DBPSO-BP、MDBPSO-BP 分别提高了 10%、9%、

7%, $RMSE_V$ 分别减少了 15%、22%、23%。RF 算法比 BP 神经网络算法更适用于小麦条锈病的遥感探测, 这是因为 RF 算法对噪声有较好的容忍度, 反演精度比较高^[38]。而 BP 神经网络算法的网络权值初始化是随机的, 而且程序的运行结果和选用的输入参数均存在差异, 致使其模型精度较低^[39], 影响了模型的估测精度。

本文在基于 DBPSO 和 MDBPSO 算法利用全波段反射率光谱数据提取小麦条锈病遥感探测特征因子时, 并没有考虑因全波段反射率光谱数据量大而导致的运行速度缓慢问题, 如何在利用上述两种算法进行特征提取前对全波段数据进行有效降维处理是下一步要研究的主要内容。

冠层光谱数据不仅受到病害胁迫的影响, 也受到冠层几何结构等因素的影响, 本文在利用实验数据建立小麦条锈病遥感探测模型时并未考虑这些问题, 这是本文的不足之处。如何消除叶面积指数以及冠层几何结构等因素对冠层光谱数据的影响, 提高小麦条锈病的遥感探测精度是一个值得探讨的问题。

本研究只比较了 CC、DBPSO、MDBPSO 三种特征选择方法和 RF、BP 神经网络两种建模算法在小麦条锈病遥感探测中的适用性, 如果使用更多的特征选择算法或建模算法, MDBPSO-RF 是否仍为遥感探测小麦条锈病的最适宜方法还需要进一步的比较。

4 结论

以 MDBPSO 算法提取的特征因子为自变量利用 BP 神经网络算法构建的小麦条锈病遥感探测模型预测 DI 值和实测 DI 值间的 $RMSE_V$ 比 CC 分析法和 DBPSO 算法分别减小了 21% 和 10%, 利用 RF 算法构建模型预测 DI 值和实测 DI 值间的 $RMSE_V$ 比 CC 法和 DBPSO 算法分别减小了 28% 和 11%。利用文中 3 种算法筛选的特征因子为自变量构建小麦条锈病严重度估测模型时, RF 算法构建的模型精度均高于 BP 神经网络算法, 其估测 DI 值与实测 DI 值间的 $RMSE_V$ 比 BP 神经网络模型平均减少了 20%, 最少减少了 15%。因此在构建小麦条锈病遥感探测模型时, MDBPSO 算法为特征选择的适宜算法, RF 算法为模型构建的适宜算法。

References

[1] 董锦绘, 杨小冬, 杨贵军, 王宝山. 基于近地高光谱信息的小

麦条锈病情指数反演. 麦类作物学报, 2016, 36: 1674–1680.

Dong J H, Yang X D, Yang G J, Wang B S. Inversion of wheat stripe rust disease index based on near ground hyperspectral data. *J Triticeae Crops*, 2016, 36: 1674–1680 (in Chinese with English abstract).

[2] 赵叶, 竞霞, 黄文江, 董莹莹, 李存军. 日光诱导叶绿素荧光与反射率光谱数据监测小麦条锈病严重度的对比分析. 光谱学与光谱分析, 2019, 39: 2739–2745.

Zhao Y, Jing X, Huang W J, Dong Y Y, Li C J. Comparison of sun-induced chlorophyll fluorescence and reflectance data on estimating severity of wheat stripe rust. *Spectrosc Spect Anal*, 2019, 39: 2739–2745 (in Chinese with English abstract).

[3] Liang D, Liu N, Zhang D Y, Zhao J L, Ding Y W. Discrimination of powdery mildew and yellow rust of winter wheat using high-resolution hyperspectra and imageries. *Infrared Laser Eng*, 2017, 46: 50–58.

[4] Zheng Q, Huang W J, Cui X M, Shi Y, Liu L Y. New spectral index for detecting wheat yellow rust using sentinel-2 multispectral imagery. *Sensor (Basel)*, 2018, 18: 868–886.

[5] 梁琨, 张夏夏, 丁静, 徐剑宏, 韩东荣, 沈明霞. 傅里叶中红外光谱结合稀疏表示分类方法鉴别小麦赤霉病感染等级. 光谱学与光谱分析, 2019, 39: 3251–3255.

Liang K, Zhang X X, Ding J, Xu J H, Han D S, Shen M X. Discrimination of wheat scab infection level by Fourier mid-infrared technology combined with sparse representation based classification method. *Spectrosc Spect Anal*, 2019, 39: 3251–3255 (in Chinese with English abstract).

[6] Shi Y, Huang W J, González-Moreno P, Luke B, Dong Y Y, Zheng Q, Ma H Q, Liu L Y. Wavelet-based rust spectral feature set (WRSFs): a novel spectral feature set based on continuous wavelet transformation for tracking progressive host-pathogen interaction of yellow rust on wheat. *Remote Sens*, 2018, 252: 1–19.

[7] Shi Y, Huang W J, Zhou X F. Evaluation of wavelet spectral features in pathological detection and discrimination of yellow rust and powdery mildew in winter with hyperspectral reflectance data. *J Appl Remote Sens*, 2017, 11: 1–18.

[8] 蒋金豹, 陈云浩, 黄文江, 李京. 冬小麦条锈病严重度高光谱遥感反演模型研究. 南京农业大学学报, 2007, 30(3): 63–67.

Jiang J B, Chen Y H, Huang W J, Li J. Study on hyperspectral remote sensing retrieval models about winter wheat stripe rust severity. *J Nanjing Agric Univ*, 2007, 30(3): 63–67 (in Chinese with English abstract).

[9] Wang L, Qu J J. Satellite remote sensing applications for surface soil moisture monitoring: a review. *Front Earth Sci China*, 2009, 3: 237–247.

[10] 刘琦, 谷医林, 王翠翠, 王睿, 李薇, 马占鸿. 基于偏最小二乘法的小麦条锈病潜育期冠层高光谱分析. 植物保护学报, 2018, 45: 138–145.

Liu Q, Gu Y L, Wang C C, Wang R, Li W, Ma Z H. Canopy hyperspectral features analysis of latent period wheat stripe rust based on discriminant partial least squares. *J Plant Prot*, 2018, 45: 138–145 (in Chinese with English abstract).

[11] 黄木易, 黄文江, 刘良云, 黄义德, 王纪华, 赵春江, 万安民. 冬小麦条锈病单叶光谱特性及严重度反演. 农业工程学报, 2004, 20(1): 176–180.

- Huang M Y, Huang W J, Liu L Y, Huang Y D, Wang J H, Zhao C J, Wan A M. Spectral reflectance feature of winter wheat single leaf infected with stripe rust and severity level inversion. *Trans CSAE*, 2004, 20(1): 176–180 (in Chinese with English abstract).
- [12] 刘琦, 王翠翠, 王睿, 谷医林, 李薇, 马占鸿. 潜育期小麦条锈菌的高光谱定性识别. *植物保护学报*, 2018, 45: 153–160.
- Liu Q, Wang C C, Wang R, Gu Y L, Li W, Ma Z H. Hyperspectral qualitative identification on latent period of wheat stripe rust. *J Plant Prot*, 2018, 45: 153–160 (in Chinese with English abstract).
- [13] Yang C, Tan Y L, Bruzzone L, Lu L J, Guan R C. Discriminative feature metric learning in the affinity propagation model for band selection in hyperspectral images. *Remote Sens*, 2017, 9: 782–798.
- [14] Kennedy J, Eberhart R. A discrete binary version of the particle swarm algorithm. In: IEEE, eds. International Conference on Systems, Man, and Cybernetics. Florida, USA: Computational Cybernetics and Simulation, 1997. pp 4104–4108.
- [15] 沈林成, 霍霄华, 牛铁峰. 离散粒子群优化算法研究现状综述. *系统工程与电子技术*, 2008, 30: 1986–1990.
- Shen L C, Huo X H, Niu Y F. Survey of discrete particle swarm optimization algorithm. *J Syst Eng Electron*, 2008, 30: 1986–1990 (in Chinese with English abstract).
- [16] Yin C, Ye Y, Zhao H, Jiang Y Z, Wang H, Shang Y Z, Wang J F. Remote sensing of water quality based on HJ-1A HSI imagery with modified discrete binary particle swarm optimization-partial least squares (MDBPSO-PLS) in inland waters: a case in Weishan Lake. *Ecol Inf*, 2018, 44: 21–32.
- [17] 张珏, 田海清, 赵志宇, 张丽娜, 张晶, 李斐. 基于改进离散粒子群算法的青贮玉米原料含水率高光谱检测. *农业工程学报*, 2019, 35(1): 285–293.
- Zhang J, Tian H Q, Zhao Z Y, Zhang L N, Zhang J, Li F. Moisture content detection in silage maize raw material based on hyperspectrum and improved discrete particle swarm. *Trans CSAE*, 2019, 35(1): 285–293 (in Chinese with English abstract).
- [18] Yang J, Zhang H S, Ling Y, Cheng P. Task allocation for wireless sensor network using modified binary particle swarm optimization. *IEEE Sensor J*, 2014, 14: 882–891.
- [19] 王丽爱, 周旭东, 朱新开, 郭文善. 基于 HJ-CCD 数据和随机森林算法的小麦叶面积指数反演. *农业工程学报*, 2016, 32(3): 149–154.
- Wang L A, Zhou X D, Zhu X K, Guo W S. Inverting wheat leaf area index based on HJ-CCD remote sensing data and random forest algorithm. *Trans CSAE*, 2016, 32(3): 149–154 (in Chinese with English abstract).
- [20] 胡小平, 杨之为, 李振岐, 邓志勇, 柯长华. 汉中地区小麦条锈病的 BP 神经网络预测. *西北农业学报*, 2000, 9(3): 28–31.
- Hu X P, Yang Z W, Li Z Q, Deng Z Y, Ke C H. BP neural network prediction of wheat stripe rust in Huazhong region. *Acta Agric Boreali-Occident Sin*, 2000, 9(3): 28–31 (in Chinese with English abstract).
- [21] Rodriguez-Gailano V, Mendes M P, Garcia-Soldado M J, Chica-Olmo M, Ribeiro L. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: a case study in an agricultural setting (Southern Spain). *Sci Total Environ*, 2014, 476: 189–206.
- [22] Ashour loo D, Mobasheri M R, Huete A. Developing two spectral disease indices for detection for wheat leaf rust (*Puccinia triticina*). *Remote Sens*, 2014, 6: 4723–4740.
- [23] Song L, Guanter L, Guan K, You L, Huete A, Ju W, Zhang Y. Satellite sun-induced chlorophyll fluorescence detects early response of winter wheat to heat stress in the Indian Indo-Gangetic Plains. *Global Change Biol*, 2018, 24: 4023–4037.
- [24] 竞霞, 白宗璠, 高媛, 刘良云. 利用随机森林法协同 SIF 和反射率光谱监测小麦条锈病. *农业工程学报*, 2019, 35(13): 154–161.
- Jing X, Bai Z F, Gao Y, Liu L Y. Wheat stripe rust monitoring by random forest algorithm combined with SIF and reflectance spectrum. *Trans CSAE*, 2019, 35(13): 154–161 (in Chinese with English abstract).
- [25] 陈思媛, 竞霞, 董莹莹, 刘良云. 基于日光诱导叶绿素荧光与反射率光谱的小麦条锈病探测研究. *遥感技术与应用*, 2019, 34: 511–520.
- Chen S Y, Jing X, Dong Y Y, Liu L Y. Detection of wheat stripe rust using solar-induced chlorophyll fluorescence and reflectance spectral indices. *Remot Sens Technol Appl*, 2019, 34: 511–520 (in Chinese with English abstract).
- [26] 孙红, 郑涛, 刘宁, 程明, 李民赞, Zhang Q. 高光谱图像检测马铃薯植株叶绿素含量垂直分布. *农业工程学报*, 2018, 34(1): 149–156.
- Sun H, Zheng T, Liu N, Cheng M, Li M Z, Zhang Q. Vertical distribution of chlorophyll in potato plants based on hyperspectral imaging. *Trans CSAE*, 2018, 34(1): 149–156 (in Chinese with English abstract).
- [27] 鲁军景, 黄文江, 蒋金豹, 张竞成. 小波特征与传统光谱特征预测冬小麦条锈病病情严重度的对比研究. *麦类作物学报*, 2015, 35: 1456–1461.
- Lu J J, Huang W J, Jiang J B, Zhang J C. A comparative study on the severity of stripe rust disease in winter wheat estimated by little baud sign and traditional spectral features. *J Triticeae Crops*, 2015, 35: 1456–1461 (in Chinese with English abstract).
- [28] 刘新杰. 日光诱导叶绿素荧光的遥感反演研究. 中国科学院大学博士论文, 北京, 2016. pp 17–19.
- Liu X J. Retrieval of Solar-Induced Chlorophyll Fluorescence by Remote Sensing. PhD Dissertation of University of Chinese Academy of Sciences, Beijing, China, 2016. pp 17–19 (in Chinese with English abstract).
- [29] 刘良云. 植被定量遥感原理与应用. 北京: 科学出版社, 2014. pp 146–153.
- Liu L Y. Principles and Applications of Quantitative Remote Sensing for Vegetation. Beijing: Science Press, 2014. pp 146–153 (in Chinese).
- [30] Maier S W, Günther K P, Stellmes M. Sun-induced fluorescence: a new tool for precision farming. In: McDonald M, Schepers J, Tартly L, eds. Digital Imaging Spectral Techniques: Applications to Precision Agriculture Crop Physiology. Madison, USA: American Society of Agronomy, 2003. pp 209–222.
- [31] Liu X J, Liu L Y. Improving chlorophyll fluorescence retrieval using reflectance reconstruction based on principal components analysis. *IEEE Geosci Remote Sens Lett*, 2015, 12: 1645–1649.
- [32] Cordon G, Lagorio M G, Paruelo J M. Chlorophyll fluorescence, photochemical reflective index and normalized difference

- vegetative index during plant senescence. *J Plant Physiol*, 2016, 199: 100–110.
- [33] 曹引, 冶运涛, 赵红莉, 蒋云钟, 王浩, 严登明. 基于离散粒子群和偏最小二乘的水源地浊度高光谱反演. *农业机械学报*, 2018, 49(1): 173–182.
- Cao Y, Ye Y T, Zhao H L, Jiang Y Z, Wang H, Yan D M. Satellite hyperspectral retrieval of turbidity for water source based on discrete particle swarm and partial least squares. *Trans CSAM*, 2018, 49(1): 173–182 (in Chinese with English abstract).
- [34] Yang H C, Zhang S B, Deng K Z, Du P J. Research into a feature selection method for hyperspectral imagery using PSO and SVM. *J China Univ Min Technol*, 2007, 17: 473–478.
- [35] 张绘娟, 张达敏, 闫威, 陈忠云, 辛梓芸. 异构网络中基于吞吐量优化的资源分配机制. *计算机科学*, 2019, 46(10): 109–115.
- Zhang H J, Zhang D M, Yan W, Chen Z Y, Xin Z Y. Throughput optimization based resource allocation mechanism in heterogeneous networks. *Compute Sci*, 2019, 46(10): 109–115 (in Chinese with English abstract).
- [36] 李健丽, 董莹莹, 师越, 朱溢俊, 黄文江. 基于随机森林模型的小麦白粉病遥感监测方法. *植物保护学报*, 2018, 45: 395–396.
- Li J L, Dong Y Y, Shi Y, Zhu Y N, Huang W J. Remote sensing monitoring of wheat powdery mildew based on random forest model. *J Plant Prot*, 2018, 45: 395–396 (in Chinese with English abstract).
- [37] 袁冰清, 程功, 郑柳刚. BP 神经网络基本原理. *数字通信世界*, 2008, 8(17): 28–29.
- Yuan B Q, Cheng G, Zheng L G. Basic principle of BP neural networks. *Digital Commun World*, 2008, 8(17): 28–29 (in Chinese with English abstract).
- [38] 姚雄, 余坤勇, 杨玉洁, 曾琪, 陈璋昊, 刘健. 基于随机森林模型的林地叶面积指数遥感估算. *农业机械学报*, 2017, 48(5): 159–166.
- Yao X, Yu K Y, Yang Y J, Zeng Q, Chen Z H, Liu J. Estimation of forest leaf area index based on random forest model and remote sensing data. *Trans CSAM*, 2017, 48(5): 159–166 (in Chinese with English abstract).
- [39] 依尔夏提·阿不来提, 买买提·沙吾提, 白灯莎·买买提艾力, 安申群, 马春玥. 基于随机森林法的棉花叶片叶绿素含量估算. *作物学报*, 2019, 45: 81–90.
- Ershat A, Mamat S, Baidengsha M, An S Q, Ma C Y. Estimation of leaf chlorophyll content in cotton based on the random forest approach. *Acta Agron Sin*, 2019, 45: 81–90 (in Chinese with English abstract).