

基于无人机高光谱遥感的冬小麦全氮含量反演

杨欣^{1,2}, 袁自然^{1,2}, 叶寅^{1,2*}, 王道中^{1,2}, 花可可^{1,2}, 郭志彬^{1,2}

1. 安徽省农业科学院土壤肥料研究所, 安徽 合肥 230031

2. 养分循环与资源环境安徽省重点实验室, 安徽 合肥 230031

摘要 氮素是作物生长发育必需的营养元素之一, 作物的全氮含量是表征其氮素状况的主要指标。田块尺度的冬小麦全氮含量空间分布监测可以辅助其精准定量追肥, 减少环境污染。无人机高光谱遥感具有分辨率高、时效性高、成本低等优势, 可为作物长势信息反演提供重要数据源。XGBoost(extreme gradient boosting)作为一种新兴集成学习算法, 运行效率高, 泛化能力强, 可以有效的应用于构建冬小麦全氮含量遥感反演模型, 预测田块尺度冬小麦全氮含量空间分布。以农业部蒙城砂姜黑土生态环境站内拔节期冬小麦为研究对象, 开展以下工作: (1)以低空无人机搭载高光谱成像仪获取冬小麦拔节期冠层成像光谱影像, 结合地面采样数据, 获取 126 个样点全氮含量数据; (2)分析拔节期冬小麦冠层光谱特征, 并根据 Person 相关系数分析 176 个波段的光谱反射率与全氮含量之间的相关性; (3)构建基于 XGBoost 算法的不同土壤肥力条件下拔节期冬小麦全氮含量无人机高光谱反演模型。结果表明: (1)176 个波段(400~1 000 nm)的光谱反射率与冬小麦全氮含量之间具有较强的相关性, 除了 735.5 nm 外其他波段光谱反射率与全氮含量之间的相关系数均大于 0.5; (2)基于 XGBoost 算法构建的拔节期冬小麦全氮含量无人机高光谱遥感反演模型具有较高的反演精度($R^2=0.76$, $RMSE=2.68$); (3)基于 XGBoost 算法的冬小麦全氮含量反演模型可以获取不同土壤肥力条件下田块尺度的全氮含量空间分布图, 总体上呈现较为显著的空间差异。该研究可为冬小麦精准定量追肥提供一定的科学依据, 也为发展无人机高光谱遥感的精准农业应用提供了参考。

关键词 冬小麦全氮含量; 无人机高光谱; XGBoost; 遥感反演

中图分类号: TP79

文献标识码: A

DOI: 10.3964/j.issn.1000-0593(2022)10-3269-06

引言

氮素是作物生长发育最需要的营养元素之一, 对作物的光合作用和产量具有显著影响^[1-3]。植株的全氮含量是表征作物氮素状况的主要指标^[4], 因此有必要从植株氮素含量来评价作物氮素营养状况。

作物氮素状况的有效诊断需建立在对作物生长特性和氮素水平实时监测的基础上。目前, 测定作物氮含量的传统方法主要依靠田间取样和实验室化学分析^[5]。近年来, 人们提出了利用叶绿素仪、叶绿素荧光等方法对植物氮素状况进行无损估测^[6-7]。然而, 这些方法仅针对单个植株单个叶片对点的监测, 在实际应用中难以反映作物植株的种群状况。相比之下, 无人机遥感近年来在精准农业领域得到了广泛的

应用^[8-9], 能够实现对作物生长状况由点到面的监测, 并能够快速评估农作物的空间差异性。无人机遥感平台可以携带多光谱和高光谱传感器, 近十几年来, 高光谱遥感以其高分辨率、强连续性和海量信息获取的优势, 进一步推动了定量农业遥感的发展^[10]。结合无人机高光谱遥感技术提取小麦长势信息已经逐渐成为国内外学者研究热点。Xia 等利用该技术从高光谱数据中选取归一化差异植被指数(NDVI)和比值植被指数(RVI)构建了冬小麦生物量遥感反演模型^[11]。从无人机影像提取出来的植被指数可以提取小麦籽粒产量信息, 支持向量回归(SVR)结合顺序前向选择(SFS)技术能够较好的用于提取籽粒信息^[12]。Liu 等利用无人机高光谱数据和实测高光谱数据计算光谱指标, 构建了用于反演不同生育期的冬小麦叶片含氮量(LNC)的反演模型并分析了 LNC 与光谱反射率的相关性^[8]。Wang 等针对江苏北部某实验点小

收稿日期: 2021-08-18, 修订日期: 2022-02-21

基金项目: 国家自然科学基金项目(31901154), 安徽省自然科学基金项目(2008085MD108), 安徽省农业科学院团队项目(2021YL0085)资助

作者简介: 杨欣, 1994年生, 安徽省农业科学院土壤肥料研究所实习员 e-mail: 1792502127@qq.com

* 通讯作者 e-mail: yeyin1218@163.com

麦拔节、孕穗期和开花 3 个生育阶段的叶片叶绿素进行了反演并发现与支持向量回归 (SVR) 和反向传播神经网络 (BPNN) 相比, 基于随机森林 (RF) 算法的叶绿素反演模型有较好的预测效果^[13]。

在同一生育期, 不同施肥处理的冬小麦光谱特征存在显著差异^[14], 因此, 有必要引入适用于不同肥力条件下的具有较强泛化能力的算法构建冬小麦全氮含量遥感反演模型。近年来, 集成学习 (ensemble learning) 作为一种新的机器学习范式, 通过增加学习器数目提高了其化能力并被广泛应用于解决各类回归问题^[15-16]。本研究以位于安徽蒙城的农业部蒙城砂姜黑土生态环境站内冬小麦为研究对象, 基于无人机高光谱遥感数据和田间采样数据, 采用梯度提升树类集成学习算法 (extreme gradient boosting (XGBoost) 构建不同土壤肥力条件下拔节期冬小麦全氮含量遥感反演模型, 提取不同肥力条件的田块冬小麦全氮含量信息, 以期为其田间施肥管理提供科学参考。

1 实验部分

1.1 研究区及概况

以安徽蒙城的农业部蒙城砂姜黑土生态环境站 (116°37' E, 33°13' N) 为研究区 (如图 1)。本站位于安徽北部平原区, 属于暖温带半湿润季风气候, 年平均气温 14.8 °C, 无霜期 212 d, 年降水量 600~900 mm。试验土为常见的普通砂姜黑土。试验站始于 1982 年。试验开始时耕层土壤 (0~20 cm) 的肥力性状为: 有机质 10.4 g · kg⁻¹, 全氮 0.96 g · kg⁻¹, 碱解氮 84.5 mg · kg⁻¹, 全磷 0.28 g · kg⁻¹, 有效磷 9.8 mg · kg⁻¹。设置 7 种不同的施肥处理, 分别为撂荒、不施肥、施氮磷钾化肥、氮磷钾+低量麦秸、氮磷钾+全量麦秸、氮磷钾+猪粪、氮磷钾+牛粪。全部肥料于秋季小麦种植整地前一次性施入各处理小区, 人工或机械耕翻, 与土壤充分混匀。长期以来, 形成了具有较大肥力差异的小麦种植田块。

1.2 数据获取

使用的数据包括无人机高光谱成像数据和冬小麦全氮含量数据。所有数据都于 2021 年 3 月 26 日 (拔节期, 无云无风) 获取。无人机高光谱成像系统 (GaiaSky-mini2-VN) 由国内大疆公司的六旋翼无人机和四川双利合谱公司生产的高光谱成像仪 Gaisky2 组成。该系统可获取 176 个波段的光谱信息, 包括可见光到近红外 (400~1 000 nm) 光谱范围, 光谱分辨率为 3.5 nm, 空间分辨率为 4.7 cm。数据采集前用参考板进行数据校准, 飞行高度设置为 120 m, 移动速度设置为 7 m · s⁻¹, 航向重叠度和旁向重叠度分别设置为 50% 和 55%。

如图 1, 在研究区域试验田内均匀地布设 126 个采样点, 每个采样点取 3~5 株小麦, 封装保鲜带回实验室。将植株样品置于 70 °C 干燥箱内恒温干燥, 粉碎, 使用凯氏定氮法^[17]测定小麦全氮含量。对于每个采样点, 取 3~5 株小麦的均值作为该点全氮含量 (total nitrogen content, TNC)。计算公式如式 (1)

$$N = (V \times 0.05 \times 14 \times 100) / (1\ 000 \times M) \quad (1)$$

式 (1) 中: N 为全氮含量 (%); V 为盐酸体积变量 (mL); M

为样品质量 (g)。

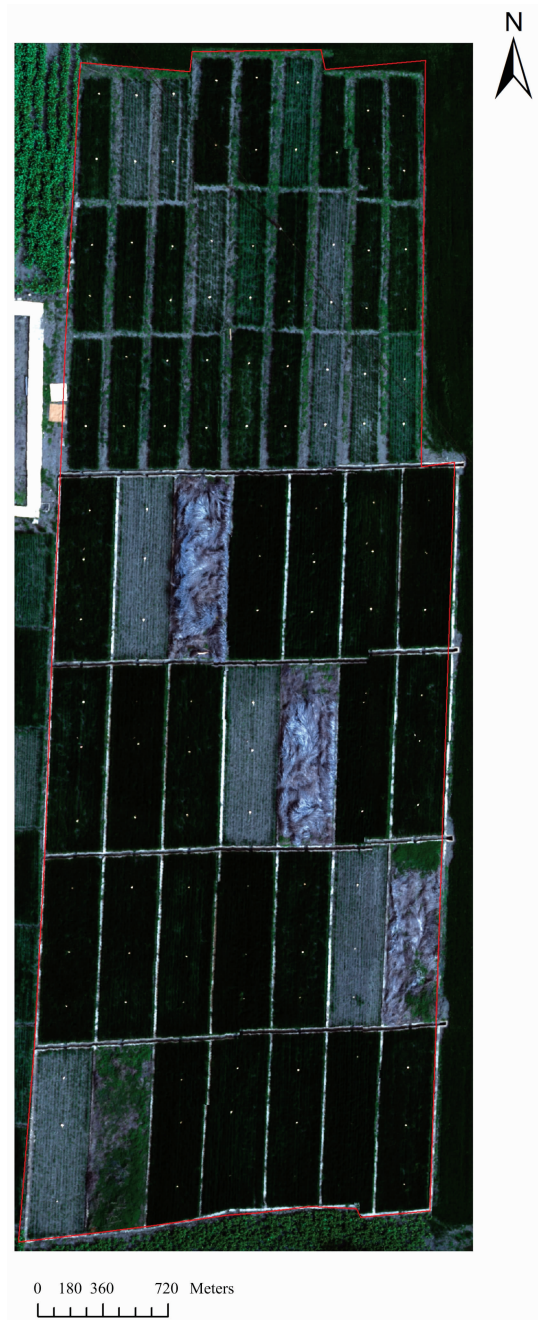


图 1 研究区域试验田分布

Fig. 1 Distribution of experimental plots the study site

1.3 数据预处理

无人机高光谱图像的处理包括以下步骤: 第一步是镜头校正, 消除图像畸变; 第二步, 辐射定标, 将数字灰度值 (digital number, DN) 转换为反射率; 然后进行大气校正, 以消除大气、水汽等因素的影响。使用 Hispectralstitch 软件 (四川双利合谱公司开发提供) 对多景遥感影像进行拼接; 最后, 从每个田块中提取感兴趣区域 (regions of interest, ROI) 的平均光谱。因此, 每个田块在 176 个波段都有对应的平均光谱, 作为模型输入的自变量, 这里将整个大田均匀地划分成

了 126 个田块。对于小麦全氮含量实测数据，为避免异常值对建模产生影响，去除了异常实测值。

1.4 方法

XGBoost 是由陈天奇等改进的一种 GBDT 算法^[17]，它与 GBDT 算法有较多不同点，比如在优化时，GBDT 仅仅用到了二阶导数，XGBoost 则另外引入了二阶导数对损失函数进行调优；在目标函数里，XGBoost 算法为了避免过拟合现象，将树模型复杂度作为正则项。此外，XGBoost 在训练过程中参考了随机森林的思想，每次迭代过程中不使用全数样本，且不采用样本的所有特征，而是有选择地采取部分样本的部分特征进行训练，这样可有效提高模型的泛化能力，削弱模型欠拟合和过拟合现象，而且 XGBoost 算法支持并行计算以提高运行速度。该算法原理是将原始数据集分割成多个数据子集，将每个子集随机地分配给基学习器进行预测，然后将基学习器的结果按照一定的权重进行计算，从而预测最后的结果。

图 2 为整体实验的技术路线图，分为 3 个步骤。第一步，构建模型训练集，将 176 个波段的光谱反射率信息作为模型输入的自变量，实测全氮含量数据作为模型因变量，随机选取 70% 的采样点的数据作为训练集并对自变量进行归一化处理；第二步，反演模型构建与验证，采用贝叶斯优化算法对模型进行超参数调优(<https://www.kaggle.com/tiliu7/>

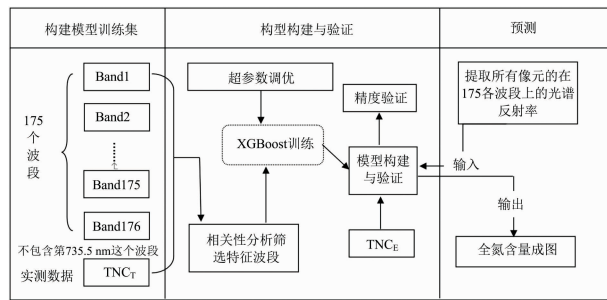


图 2 基于 XGBoost 模型的冬小麦全氮含量遥感预测技术路线图

Fig. 2 Technical flow chart for the winter wheat total nitrogen content (TNC) prediction based on XGBoost

表 1 XGBoost 模型中一些超参数的含义及其最优值

Table 1 The meaning and optimal values of some hyper-parameters of the XGBoost model

| 部分超参数 | 含义(默认值) | 最优值 |
|------------------|--------------------------|-----|
| learning rate | 为了防止过拟合，更新程中用到的收缩步长(0.3) | 0.1 |
| min_child_weight | 孩子节点中最小的样本权重和(1) | 5 |
| max_depth | 树的最大深度(6) | 9 |
| colsample_bytree | 在建树时对特征采样的比例(1) | 0.8 |
| subsample | 用于训练模型的子样本占整个样本集合的比例(1) | 1 |
| gamma | 在树的叶节点上进行进一步分区所需的最小损失(0) | 0 |
| reg_alpha | L1 正则化系数(0) | 0.1 |
| reg_lambda | L2 正则化系数(1) | 1 |

bayesian-optimizing-of-xgboost-parameters/notebook)，该优化方法可以看做一个高斯过程，考虑了先验参数信息且不断更新知识而传统的网格搜索或随机不考虑先验参数信息。此外，贝叶斯优化方法迭代次少运行速度快，可以优化 XGBoost 这种具有多个参数的算法，表 1 为本研究中 XGBoost 遥感反演模型中关键超参数的含义及其调优后的结果，其他参数设置为默认值。利用 30% 的验证集数据进行模型验证和精度评估，第三步，基于第二步构建好的全氮含量遥感反演模型预测大田小麦的全氮含量，提取无人机高光谱影像上所有像元的光谱反射率信息作为模型的输入，输出预测的全氮含量并成图。

2 结果与讨论

2.1 模型反演精度分析

在长期定位试验下，不同的肥料处理导致不同田块上的土壤肥力存在显著差异，田块间冬小麦的长势存在显著差异，其叶片光合作用性能不同，故其对光的吸收和反射也不同，如图 3 所示，不同田块冬小麦冠层光谱特征呈现出显著差异。为了筛选出用于构建该区域冬小麦全氮含量遥感反演的敏感波段，首先对光谱反射率和实测全氮含量进行了相关性分析，计算了全氮含量与各波段光谱反射率之间的皮尔逊相关系数。由图 4 可知，总体上各波段反射率与冬小麦全氮含量之间存在较高的相关性，在各波段的绝对相关系数均值达 0.732。在波长为 795.3 nm 时，全氮含量与该波段相关性达到最大值，为 0.794。波长小于 735.5 nm 时，各波段的光

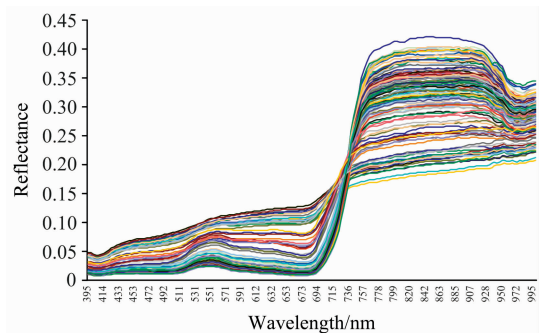


图 3 不同田块冬小麦冠层光谱图

Fig. 3 Canopy spectra of winter wheat at different fields

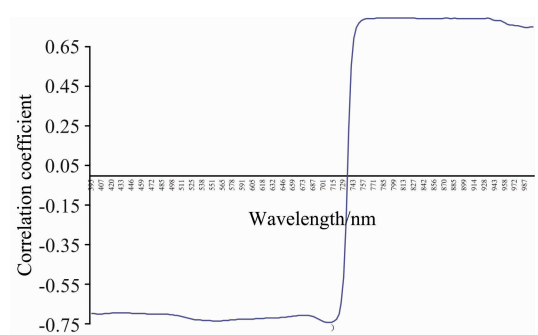


图 4 光谱反射率与全氮含量之间的相关系数

Fig. 4 Correlation coefficient of spectral reflectance and TNC

谱反射率与全氮含量均成负相关；波长大于等于 735.5 nm 时，各波段的光谱反射率与全氮含量均成正相关。当波长为 735.5 nm 时，反射率与全氮含量之间的相关性最低，为 0.232，其余波段的相关系数绝对值均高于 0.5，这可能是由于各田块在该波段上的光谱反射率几乎一致(图 3)，故各田块在该波段上无光谱差异性。

因此，本研究不采用全波段作为建模的自变量，采用了除 735.5 nm 波段以外的所有波段作为模型输入自变量，使用 XGBoost 方法构建冬小麦全氮含量遥感反演模型，模型验证精度使用决定系数 R^2 、均方根误差 (root mean absolute error, RMSE) 和平均绝对百分比误差 (mean absolute percent error, MAPE) 作为评价指标，其中，MAPE 用来评价模型预测结果的相对误差。由图 5 可知，XGBoost 模型预测的和实测的小麦全氮含量较为接近，模型反演精度较高， R^2 达到 0.76，RMSE 值较低，仅为 2.68，MAPE 值也较低，为 11.08%，这说明 XGBoost 算法具有相对较强的学习性能，可以应用于构建冬小麦全氮含量遥感反演模型。

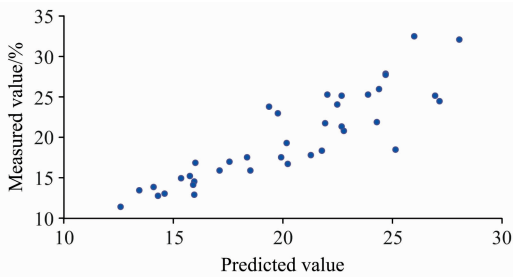


图 5 XGBoost 模型反演的与实测全氮含量之间的散点图

Fig. 5 Scatter plots between the predicted TNC and measured TNC

2.2 TNC 预测结果分析

基于 XGBoost 算法的拔节期冬小麦全氮含量反演模型具有较好的验证精度，现将该模型应用到整片小麦大田上，提取出不同土壤肥力条件下的小麦大田影像上所有像元在 176 个波段的反射率，输入该反演模型进行预测整片小麦大田的全氮含量 (TNC)。

由图 6 可知，XGBoost 模型预测的拔节期冬小麦全氮含量总体上呈现较为显著的空间分布差异，这是由长期试验不同田块不同的土壤肥力导致，该算法具有较强的泛化能力，构建的遥感反演模型适用于预测不同土壤肥力条件的整片大田的冬小麦全氮含量空间分布。模型预测的冬小麦 (拔节期) 全氮含量范围介于 14%~29% 之间，平均全氮含量为 19.66%。

2.3 冬小麦全氮含量与土壤全氮含量相关性分析

针对长期定位试验下，不同土壤肥力的田块上的冬小麦全氮含量进行了遥感反演研究，由 2.2 节研究发现，在不同土壤肥力的田块，冬小麦的全氮含量存在显著空间分布差异性，这是由于土壤养分对冬小麦长势有着一定的影响。因此，本节针对冬小麦全氮含量与土壤全氮含量进行了相关性

分析。

如图 7 所示，冬小麦全氮含量与土壤全氮含量之间存在着较为显著的相关性，其中，冬小麦全氮含量范围为 12%~32%，土壤全氮含量范围为 29%~55%，两者之间的决定系数 (R^2) 为 0.47。

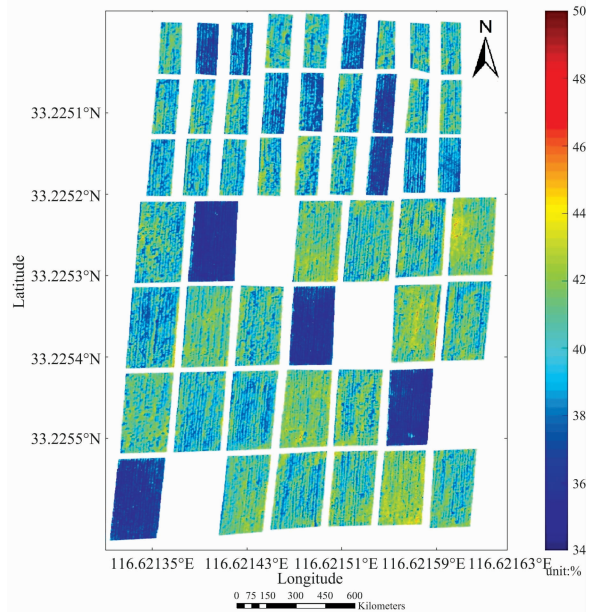


图 6 XGBoost 反演模型预测的 TNC 空间分布图

Fig. 6 Spatial distributions of TNC predicted by XGBoost

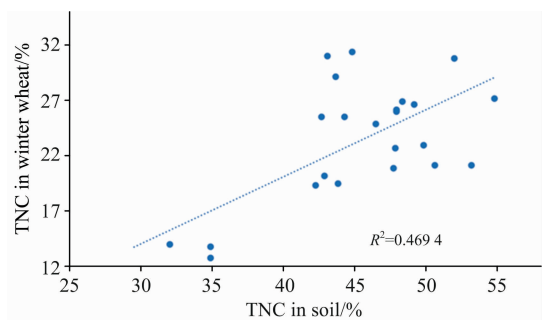


图 7 冬小麦全氮含量与土壤全氮含量之间的相关性

Fig. 7 Correlaion between TNC in winter wheat and TNC in soil

3 结 论

(1) 基于 XGBoost 算法，结合无人机高光谱遥感数据和田间采样冬小麦全氮含量数据可以构建全氮含量遥感反演模型，模型验证精度较高。

(2) 在不同土壤肥力的田块，XGBoost 模型预测的拔节期冬小麦全氮含量呈现较为显著的空间分布差异，平均全氮含量为 19.66%。

(3) 冬小麦全氮含量与土壤全氮含量之间存在着较为显著的相关性，两者之间的决定系数 (R^2) 为 0.47。

References

- [1] Hucklesby D P, Brown C M, Howell S E, et al. *Agronomy Journal*, 1971, 63(2): 274.
- [2] Scheromm P, Martin G, Bergoin A, et al. *Cereal Chemistry*, 1993, 69: 664.
- [3] Harper L A, Sharpe R R, Langdale G W, et al. *Agronomy Journal*, 1987, 79(6): 965.
- [4] Daigger L A, Sander D H, Perterson G A. *Agronomy Journal*, 1976, 68(5): 815.
- [5] Roth G W, Fox R H, Marshall H G. *Agronomy Journal*, 1989, 81(3): 502.
- [6] Turner F T, Jund M E. *Australian Journal of Experimental Agriculture*, 1994, 34(7): 1001.
- [7] WANG Lei, BAI You-lu(王 磊, 白由路). *Journal of Plant Nutrition and Fertilizers* (植物营养与肥料学报), 2006, 12(6): 902.
- [8] Liu H, Zhu H, Wang P. *International Journal of Remote Sensing*, 2017, 38(8-10): 2117.
- [9] Kaivosoja J, Pesonen L, Kleemola J, et al. *Proc. SPIE*, 2013, 8887: 88870H.
- [10] Zhang J, Cheng T, Guo W, et al. *Plant Methods*, 2021, 17(1): 49.
- [11] Xia L, Zhang R R, Chen L P, et al. *Advances in Animal Biosciences*, 2017, 8(2): 833.
- [12] Shafiee S, Lied L M, Burud I, et al. *Computers and Electronics in Agriculture*, 2021, 183(4): 106036.
- [13] WANG Li-ai, MA Chang, ZHOU Xu-dong, et al(王丽爱, 马 昌, 周旭东, 等). *Transactions of the Chinese Society for Agricultural Machinery*(农业机械学报), 2015, 46(1): 259.
- [14] ZHANG Yue, TIAN Yuan-sheng, SUN Wen-yi, et al(张 玥, 田园盛, 孙文义, 等). *Spectroscopy and Spectral Analysis* (光谱学与光谱分析), 2020, 40(1): 209.
- [15] Su H, Yang X, Lu W, et al. *Remote Sensing*, 2019, 11(13): 1598.
- [16] Xia Y F, Liu C Z, Li Y Y, et al. *Expert Systems with Applications*, 2017, 78: 225.
- [17] Chen T, Guestrin C. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, 785(<https://doi.org/10.1145/2939672.2939785>).

Winter Wheat Total Nitrogen Content Estimation Based on UAV Hyperspectral Remote Sensing

YANG Xin^{1, 2}, YUAN Zi-ran^{1, 2}, YE Yin^{1, 2*}, WANG Dao-zhong^{1, 2}, HUA Ke-ke^{1, 2}, GUO Zhi-bin^{1, 2}

1. Institute of Soil and Fertilizer, Anhui Academy of Agricultural Sciences, Hefei 230031, China

2. Anhui Key Laboratory of Nutrient Cycling, Resources and Environment, Hefei 230031, China

Abstract Nitrogen is one of the necessary nutrient elements for crops' growth and development, and crops' total nitrogen content is the main index to characterize its nitrogen status. Monitoring the spatial distribution of winter wheat total nitrogen content at the field scale can assist in accurate and quantitative topdressing and reduce environmental pollution. UAV (Unmanned aerial vehicle) hyperspectral data can provide an important data source for crop growth information inversion due to its high resolution, high timeliness and low cost. XGBoost (Extreme Gradient Boosting), a new ensemble learning algorithm with high efficiency and strong generalization ability, can be effectively applied to build a winter wheat total nitrogen content estimation model based on remote sensing data and predict the spatial distribution of winter wheat total nitrogen content at field scale. Therefore, this study selected the winter wheat at the jointing stage in the national soil quality observation and experimental station as the study object and carried out the following work: (1) we obtained the canopy imaging spectral image of winter wheat at the jointing stage with a hyperspectral imager mounted on a low-altitude UAV, and total nitrogen content data of 126 samples combined with ground sampling data. (2) The spectral characteristics of the winter wheat canopy at the jointing stage were analyzed, and the correlation between spectral reflectance of 176 bands and total nitrogen content was analyzed according to the Person correlation coefficient. (3) A winter wheat total nitrogen content estimation model based on UAV hyperspectral at the jointing stage was built with the XGBoost algorithm under different soil fertility conditions. The results showed that: (1) there was a strong correlation between spectral reflectance and total nitrogen content of winter wheat in 176 bands, and the correlation coefficients between spectral reflectance and total nitrogen content in all bands except 735.5 nm were greater than 0.5; (2) The UAV hyperspectral remote sensing estimation model of winter wheat total nitrogen content at jointing stage based on XGBoost algorithm shows high accuracy ($R^2=0.76$, RMSE=2.68); (3) The estimation model of winter wheat total nitrogen content based on XGBoost algorithm can obtain the spatial distribution map of total nitrogen content at field scale

under different soil fertility conditions, which shows a significant spatial difference on the whole. This study can provide a scientific basis for the accurate and quantitative topdressing of winter wheat and also provide a reference for the application of UAV hyperspectral remote sensing in precision agriculture.

Keywords Winter wheat total nitrogen content (TNC); UAV hyperspectral data; XGBoost; Remote sensing estimation

(Received Aug. 18, 2021; accepted Feb. 21, 2022)

* Corresponding author

本 刊 声 明

近期以来,一些不法分子假冒《光谱学与光谱分析》期刊社名义,以虚假网站等形式欺骗广大作者、读者。这些虚假网站公然假冒《光谱学与光谱分析》期刊名义进行大肆的征稿并骗取作者的审稿费和版面费。经部分作者及读者举报,现有有关部门已就此介入调查。本刊将通过法律途径向假冒者追究相应的责任,维护本刊权益。

本刊官方网站已正式开通,网址为

<http://www.gpxygpx.com/>

在此郑重声明,本网址为《光谱学与光谱分析》期刊唯一开通运行的官方网站。本刊从未授权任何单位或个人以任何形式(包括网上网下)代理本刊征稿、审稿等项业务。

希望广大读者和作者切实维护好自身的合法权益,防止受骗上当。

《光谱学与光谱分析》期刊社

2019年3月15日