

基于 XGBoost 的无人机多光谱遥感监测樟树叶 积指数和地上生物

鲁向晖^{1,2†}, 王倩^{1,2}, 杨宝城^{1,2}, 张海娜^{1,2}, 张 杰^{1,2}

(1.南昌工程学院江西省樟树繁育与开发利用工程研究中心, 南昌, 330099;

2.南昌工程学院水土保持学院, 南昌, 330099)

摘要: 利用无人机多光谱技术获取樟树叶面积指数和生物量信息, 及时快速诊断樟树矮林生长状况, 有助于提升樟树矮林精准管理水平。多光谱遥感图像提供大面积的植物光谱信息, 可对叶面积指数、地上生物量和其他植物理化参数进行详细的定量评估, 是评估分析植物健康非常有效的工具。本文以红壤地区的樟树矮林为研究对象, 利用无人机多光谱相机获取樟树矮林冠层遥感影像, 采用 XGBoost、GBDT、RF、RFNN、SVR 算法, 研究原始波段反射率、光谱指数与樟树 LAI、AGB 之间的相关性及其估测精度。结果表明: 不同模型输入量, 基于 RF 估测模型精度有较明显的差异, 其余 4 个模型有差异但较小。不同建模方法, 基于 XGBoost 估测樟树 LAI 模型精度最高, 以波段反射率为模型输入量, 模型测试集 R^2 为 0.862, RMSE 为 0.390。基于 XGBoost 估测樟树 AGB 模型精度最高, 以光谱指数为模型输入量, 模型测试集 R^2 为 0.929, RMSE 为 587.746 kg·hm⁻²。XGBoost 模型为最佳樟树 LAI 和 AGB 估测模型, 其次为 GBDT、RF、RFNN、SVR。本研究成果可为基于无人机多光谱技术的樟树矮林监测提供理论依据, 为快速估测樟树生长参数提供了参考依据。

关 字: 樟树; 叶面积指数; 地上生物量; 多光谱; 波段反射率; 光谱指数; XGBoost

1.引言

樟树 (*Cinnamomum camphora* (L.) presl), 属樟科常绿阔叶乔木, 主要分布于中国南部、韩国、日本、越南等亚热带和热带地区^[1]。樟科植物的各个部位(叶、枝条、树干、果等)含有丰富的香料物质, 是香精香料、药用卫生和食品等领域的主要原材料^[2-4], 樟树精油产业已成为南方地区林业的主导产业之一^[5]。叶面积指数(leaf area index, LAI)和地上生物量(aboveground biomass, AGB)是评价植物长势、养分状况的重要指标, 其中叶面积指数主要表征作物的冠层结构, 与植物的光合作用、呼吸作用、蒸腾作用等过程密切相关^[6]; 生物量是决定植物的光能利用、生长趋势以及产量的关键参数^[7]。根据叶面积指数和生物量, 可监测植物的健康状况和养分状况, 及时进行水肥管理。樟树长势状况影响叶片出油率, 进而影响樟树产业优质高效生产。因而, 快速、无损、准确监测樟树叶面积指数和地上部生物量, 对于指导樟树矮林种植与管理具有重要意义。

当前获取植物的叶面积指数和生物量主要采用田间取样法, 通过人工实地测量、破坏性取样等, 往往耗时费力, 时效性较差, 限制了其大规模应用^[8]。近年来, 遥感技术广泛应用于农业领域。已有研究利用卫星遥感数据较好地估测了多种作物的叶面积指数和生物量, Zhao 等^[9]利用 HJ-1A 卫星遥感影像对江苏地

[†]基金项目: 国家自然科学基金项目(52269013)、江西省自然科学基金面上项目(20232BAB205031)、江西省科技厅重大科技专项(20203AB C28W016-01-04)和江西省林业局樟树研究专项(202007-01-04)

作者简介: 鲁向晖(1976—), 男, 教授, 博士, 主要从事农林水文遥感技术研究, E-mail: xianghuilu@nit.edu.cn

区水稻产量进行估测; Morain 等^[10]通过 ERTS-1 卫星获取遥感图像, 对冬小麦产量进行估计; 此外估测作物还有棉花^[11]、大豆^[12]等。然而, 卫星遥感影像易受大气影响, 且受时空分辨率、实时性的限制, 导致无法获取适用于精准农业的高质量卫星遥感图像^[13,14]。例如, Landsat 系列卫星重复周期最短为 16 天, 主要利用光学传感器获取遥感影像, 当信号传播路线被云雾或者降雨影响时, 影像将无法应用到植物生理参数的精准监测; Sentinel-2A/B 卫星具有更好的时空分辨率, 但也易受天气影响^[15]。高光谱成像仪具有光谱波段范围广、分辨率高和窄带多的优点, 已有研究运用高光谱遥感对作物(小麦^[16]、玉米^[17]等)生理生化参数进行预测, 但由于其所含数据较大、信息冗余不易处理且设备价格高, 制约其普及化应用。相比于卫星遥感和高光谱仪, 无人机多光谱遥感成本相对较低、灵活性更高。多光谱遥感有 3 个及以上特定的光谱频带, 可包含监测农学参数信息重要的红边波段, 因而备受农业定量遥感领域的关注。

诸多学者基于无人机多光谱遥感对作物生理生化参数监测进行了大量研究, 例如 Qi 等^[18]基于无人机多光谱图像选取 8 个植被指数, 建立估测花生叶面积指数的简单回归和人工神经网络模型, 结果表明人工神经网络估测结果最优。Zheng 等^[19]利用 5 波段多光谱相机, 以 19 个植被指数作为模型自变量, 评估 13 种模型对冬小麦叶片氮含量的估测效果, 结果发现随机森林算法模型估测精度最高。Su 等^[20]利用无人机搭载 5 波段多光谱相机获取小麦多光谱遥感图像, 将比值植被指数(RVI)、归一化差异植被指数(NDVI)和优化土壤调节植被指数(OSAVI)应用于小麦黄锈病的监测, 取得了良好的监测效果。Feng 等^[21]选用植被指数、冠层盖度和植物高度等模型输入量, 估测不同时期的棉花产量。分析已有研究成果发现, 多种类型的变量如光谱波段、光谱指数和植物高度等可以作为建模的输入变量, 研究证明选择合适的输入量对 LAI 和 AGB 建模估测的重要性。多光谱技术能同时获取多个波段信息, 可获得几张不同光谱带的影像。与土壤、水体和岩石等相比, 绿色植物具有明显的光谱反射特征, 其特征主要由其化学和形态学特征决定的, 与植被健康状况密切相关。在可见光波段, 其光谱反射率主要受色素影响, 大部分可见光被植物叶片吸收, 反射率较小; 在近红外区域, 主要受植物叶片内部构造的控制, 反射率较大且吸收率低^[22]。在可见光波段与近红外波段之间, 反射率急剧上升, 形成“红边”, 这是植物曲线最明显的光谱特征, 也是植被遥感关注的一个焦点。光谱指数是基于不同波段信息的数学组合, 主要反映植被在可见光、近红外波段反射与土壤背景之间差异, 在一定条件下能用来定量说明植被的生长状况, 通常用于分析植物的含水量^[23]、叶面积指数和叶绿素^[24]等参数。在估测植物的叶面积指数或生物量时, 模型输入量选择是建模的重要步骤之一, 估测模型的自变量大多数选取植被光谱指数, 而未结合原始光谱反射率进行讨论, 未见有光谱波段发射率与光谱指数在同一种反演模型的模拟精度的比较。因此, 在植物生理生化指标估测建模中, 分析波段反射率与光谱指数为模型输入量对估测精度有待进一步探究。

除了模型输入量的选择, 模型的选择也至关重要。近年来, 机器学习算法凭借自身优越性广泛应用于精准农业领域, 对植物的理化参数进行反演。Pham 等^[25]通过对比基于支持向量回归、随机森林回归、Cat Boost 回归、梯度增强回归树和极度梯度增强决策树 5 种模型, 估测红树林地上生物量, 结果表明基于极度梯度增强决策树和遗传算法优化模型最优。Yuan 等^[26]估算热带濒危树种坡垒 SPAD 值, 通过比较不同算法得出随机森林模型的预测精度最高。Siegmann 等^[27]对比分析了支持向量回归、随机森林回归和偏最小二乘回归 3 种算法在估测小麦 LAI 的模型质量, 结果表明在单独年份以及整个数据集的交叉验证情况下, 基于支持向量回归估测模型结果最优。已有研究表明, 基于随机森林、支持向量回归、梯度提升决策树和极度梯度增强决策树估测植物 LAI 和 AGB 模型性能较好, 不同模型应用于不同植物理化参数的估测质量不同^[25, 28, 29]。此外, 人工神经网络是估测植物叶面积指数和地上生物量常用机器学习算法之一, 径向基函

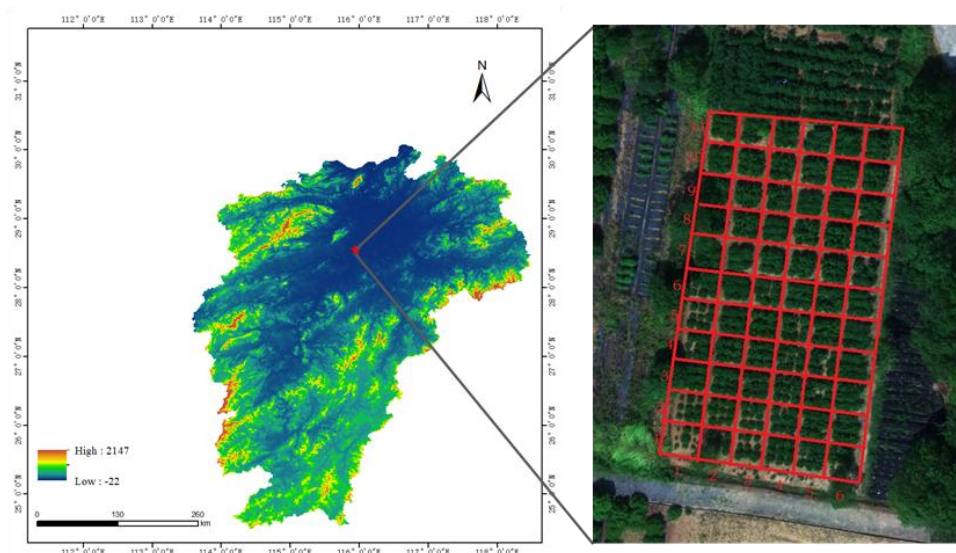
数神经网络是其中一种。径向基神经网络学习过程收敛速度快，具有很强的鲁棒性。因此，获取樟树冠层无人机多光谱遥感影像，研究不同模型与模型输入的选取，对樟树长势参量反演精度的影响是非常必要的。此外，利用无人机多光谱遥感技术对樟树生长状况的监测也鲜有报道。

本研究以樟树矮林为研究对象，利用无人机 6 波段多光谱相机获取樟树冠层多光谱影像，研究原始波段反射率、光谱指数与樟树 LAI、AGB 之间的相关关系及估测精度，以决定系数为依据分别构建基于极致梯度提升 (eXtreme Gradient Boosting, XGBoost)、梯度提升决策树 (Gradient Boosting Decision Tree, GBDT)、随机森林 (Random forest, RF)、径向基函数神经网络 (Radial Basis Function Neural Network, RBFNN)、支持向量回归 (Support Vector Regression, SVR) 估测模型，旨在探明不同模型及不同模型输入量对樟树 LAI 和 AGB 估测中的影响，寻求樟树矮林 LAI 和 AGB 最佳反演模型。通过精准估测樟树 LAI、AGB，可实现樟树矮林生长快速监测并及时指导林间种植与管理，为促进樟树矮林产业发展提供科学参考。

2.材料与方法

2.1 研究区域概况

樟树矮林试验田位于在江西省南昌市南昌工程学院生物技术实验基地 (28°41'33"N, 116°1'19.37"E) (图 1)。该地属亚热带湿润季风气候，年平均降雨量 1600mm，年平均气温在 17°C 左右，极端高温 42°C，极端低温 -10°C，年平均日照时间 1800h，光照充足。该试验区土壤质地为红壤土，pH 值约为 5.47，有机质含量 6.39 g/kg，全氮含量 0.62 g/kg，全磷含量 0.30 g/kg，全钾含量 13.00 g/kg，碱解氮含量 47.74 mg/kg，速效磷含量 1.49 mg/kg，速效钾含量 61.10 mg/kg。试验樟树采用矮林栽培，品种为“赣芳 1 号”，来自同一无性系扦插苗，2020 年 3 月份移栽，于 2021 年 3 月份定植，2022 年 9 月底进行伐桩，此时芳樟矮林生长基本处于收获状态，其干物质积累达到峰值，生长状况具有代表性。样本小区总计 66 个，每个样



本范围为 3m×3m，栽植行距、株距为 1m×1m。

图 1 研究区示意图

2.2 数据获取

2.2.1 无人机遥感影像采集

无人机遥感影像获取时间为 2022 年 9 月 26 日 12:00 左右,采用大疆经纬 Matrice 300 RTK 型四旋翼无人机搭载 MS600 PRO 型多光谱相机,该设备集成 6 个多光谱传感器通道:蓝光(中心波长 450nm,带宽 35nm)、绿光(中心波长 555nm,带宽 25nm)、红光(中心波长 660nm,带宽 20nm)、红边光 1(中心波长 720nm,带宽 10nm)、红边光 2(中心波长 750nm,带宽 15nm)和近红外光(中心波长 840nm,带宽 35nm)。该无人机具备六向定位避障与可视化飞行辅助界面,多光谱相机包含“双红边”植被敏感波段,有效提升植被光谱特性与形态特征。根据试验区的范围设置飞行航线,并进行白板校正,飞行高度设置为 30m,设置速度为 2.5m/s,选用自动捕获模式,航向、旁向的重叠度分别为 75%、65%。

2.2.2 地面实测数据获取

樟树 LAI 数据测定采用美国 LAI-2200C 植物冠层分析仪,每个样本小区选取 3 株具有代表性的樟树,每棵树选择 4 个方向(东西南北)测量并记录 LAI 值,取 4 次测量平均值作为该树的 LAI 值,最后取 3 株树的平均值作为该小区的 LAI 值。将每个样本小区全部樟树带回实验室后,先放置在烘箱中 105℃杀青 30min,然后 80℃烘干 48h 以上,直至恒重,最后称取样本质量,每个小区取均值,测得地上生物量。地面测量时间均采用与遥感影像成像时间相同步的日期。

本次试验分别获取 66 个樟树 LAI、AGB 样本数据,LAI 的最大值、最小值、均值、标准差、变异系数分别为 4.29、1.26、2.91、0.61、0.21;AGB 的最大值、最小值、均值、标准差、变异系数分别为 8720.09 kg/hm²、2095.28 kg/hm²、6398.92 kg/hm²、1.63、0.25。

2.3 数据处理

2.3.1 多光谱数据处理

本文采用 MS600 PRO 多光谱相机定制数据处理软件 Yusesense Map 处理樟树矮林冠层遥感影像。通过 Yusesense Map 进行影像波段校准、拼接合成、辐射校正等,生成 tif 格式文件。将 tif 文件导入 ENVI5.3 软件,在 Toolbox 中,使用 Region of interest 功能进行感兴趣区域截取,通过冠层信息提取、掩膜去除阴影和土壤背景等步骤,最后可分别获取 6 个波段下感兴趣区域内所有植被冠层像素点的平均反射率,取得该样本小区的樟树冠层原始波段反射率。

2.3.2 光谱指数的构建与选取

基于无人机多光谱图像反射率计算植被光谱指数,根据已有文献^[30],分别计算任意两波段之间的差值光谱指数(Difference Spectral Index, DSI),计算公式如(1)所示:

$$DSI=R_i-R_j \quad (1)$$

式中: R_i 、 R_j 分别代表蓝、绿、红、红边 1、红边 2 和近红外波段任意两不同波段的波段反射率。通过上述光谱指数与樟树 LAI、AGB 进行相关性分析,最终选取 3 个差值光谱指数,分别为 $DSI_{NIR,R}$ (近红外光-红光)、 $DSI_{RE1,B}$ (红边光 1-蓝光)、 $DSI_{RE2,B}$ (红边光 2-蓝光)。DSI 为两波段反射率之差,主要用来区分植被与土壤背景信息^[31]。此外,本文通过将大量的光谱指数与樟树 LAI 和 AGB 进行相关性分析,综合对比分析选取了 3 个经验植被指数,分别为土壤调节植被指数(SAVI)^[32]、三角植被指数(TVI)^[24]、过绿植被指数(EXG)^[33]。

2.3.3 数据分析及模型准确性评价

本文采用 Python3.7 软件分别对樟树 LAI、AGB 与原始波段反射率、光谱指数进行相关性分析、并作相关性热力图。用 Python3.7 软件平台搭建基于 XGBoost、GBDT 的樟树 LAI、AGB 估测模型,利用 Matlab2020a 软件平台建立基于 RF、RBF、SVR 的樟树 LAI、AGB 估测模型。采用 Origin2021 软件绘制各估

测模型的拟合图。

样本数据共 66 组，将数据随机分成 50 组作为模型训练集，其余 16 组作为测试集。采用决定系数 R^2 、均方根误差 RMSE 来评估不同算法模型的预测能力：

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

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - \bar{y})^2} \tag{3}$$

式中： y_i 为樟树 LAI 或 AGB 的实测值， $\hat{y}_i$ 为樟树 LAI 或 AGB 的模型估测值， $\bar{y}$ 为所有樟树 LAI 或 AGB 实测值的算术平均值。 R^2 值越高，RMSE 值越小，表明模型估测性能越好。

3 结果

3.1 原始波段反射率与樟树 LAI、AGB 的相关性分析

对多光谱影像处理分析得到 66 个样本小区的樟树冠层原始波段反射率，分别与对应的 66 个樟树 LAI、AGB 组成一个数据集，并进行相关性分析。绘制相关性热力图，结果显示：蓝光、绿光、红光、红边 1、红边 2 和近红外波段反射率与 LAI、AGB 均存在正相关关系（图 2），相关系数均在 0.6 以上。其中，对于樟树 LAI，红边 2 波段反射率与 LAI 相关性最高（0.731），绿光波段反射率与 LAI 相关性最低（0.659），按相关系数由高到低排序：RE2>NIR>B>R>RE1>G；对于樟树 AGB，红边 1 波段反射率与 AGB 相关性最高（0.761），近红外波段反射率与 AGB 相关性最低（0.664），按相关系数由高到低排序：RE1>B>RE2>G>R>NIR。此外，樟树叶面积指数与地上生物量相关性显著，相关系数达到 0.731。樟树叶面积指数与原始波段反射率的相关性总体低于 AGB 与原始波段反射率的相关性。基于上述分析表明，原始波段反射率与樟树 LAI、AGB 具有显著稳定的相关关系。

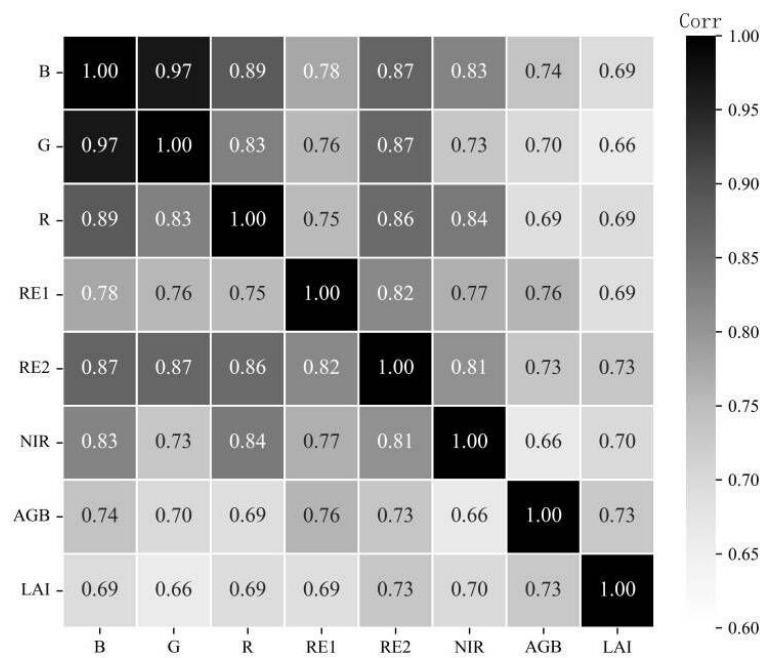


图 2 波段反射率与樟树 LAI、AGB 的相关性

3.2 光谱指数与樟树 LAI、AGB 的相关性分析

绘制 6 个选定的光谱指数与樟树 LAI、AGB 相关性热力图，结果显示：DSI_{NIR,R}，DSI_{RE1,B}，DSI_{RE2,B}，TVI，SAVI，EXG 与 LAI、AGB 均呈极显著正相关（图 3），相关系数均在 0.6 以上。其中，对于樟树 LAI，差值光谱指数 DSI_{RE2,B} 与 LAI 之间的相关系数最高（0.728），EXG 与 LAI 之间的相关系数最低（0.603），按相关系数由高到低排序：DSI_{RE2,B}>TVI>SAVI>DSI_{NIR,R}>DSI_{RE1,B}>EXG；对于樟树 AGB，差值光谱指数 DSI_{RE1,B} 与 AGB 之间的相关系数最高（0.737），SAVI 与 AGB 之间的相关系数最低（0.626），按相关系数由高到低排序：DSI_{RE1,B}>DSI_{RE2,B}>TVI>EXG>DSI_{NIR,R}>SAVI。樟树 LAI 和 AGB 分别与光谱指数的相关水平相差不大。基于上述分析表明，选定的光谱指数与樟树 LAI、AGB 具有显著稳定的相关关系。

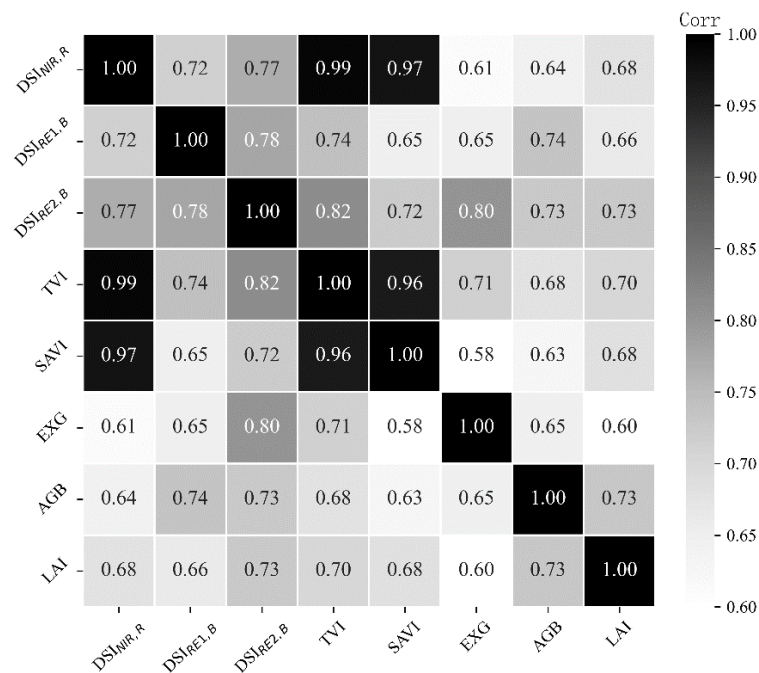


图 3 选定光谱指数与樟树 LAI、AGB 的相关性

3.3 基于波段反射率与光谱指数的樟树 LAI、AGB 估测模型的构建

本文分别以 6 个原始波段反射率和 6 个光谱指数作为樟树 LAI、AGB 反演模型的输入量，具有一定的可行性（详见图 2、图 3）。分别使用 XGBoost、GBDT、RF、RBFNN 和 SVR 算法进行建模，基于 XGBoost 模型结果如图 4，模型预测精度对比结果如表 1。

本文采用网络搜索和 5 折交叉验证进行调整 XGBoost、GBDT 和 RF 的模型参数，使用 for 循环和交叉验证的策略对 RBFNN 和 SVR 模型进行手动调参。基于 XGBoost 算法估测樟树 LAI、AGB，通过网格搜索找到最佳参数，再划定更详细的区间，将最优参数精细化，最终模型参数弱学习器的数量 $n_estimators$ 均设置为 100，收缩步长 $learning_rate$ 分别设置为 0.03、0.02，决策树最大深度 max_depth 均设置为 5。基于 GBDT 算法估测樟树 LAI、AGB，模型参数 $n_estimators$ 均设置为 100， $learning_rate$ 分别设置为 0.03、0.02， max_depth 均设置为 3。基于 RF 算法估测樟树 LAI、AGB，模型参数决策树树目 $trees$ 均设置为 100，最小叶子数 $leaf$ 分别设置为 1、2。基于 RBFNN 估测樟树 LAI、AGB，以 LAI 光谱指数、波段反射率为自变量，径向基函数的扩展速度 rbf_spread 分别设置为 90、120；以 AGB 光谱指数、波段反射率为自变量，径向基函数的扩展速度分别设置为 50、70。基于 SVR 估测樟树 LAI、AGB，以 LAI 光谱指数、波段反射率为自变量，惩罚因子 C 分别设置为 4、8，径向基参数 g 分别设置为 0.0625、0.0313；以 AGB 光

谱指数、波段反射率为自变量，惩罚因子 C 分别设置为 1、16，径向基参数 g 分别设置为 4、2。

在 5 个模型中，估测樟树 LAI 和 AGB 模型的训练集和测试集 R^2 均在 0.6 以上，最高可达 0.927，表明模型可用于樟树 LAI 和 AGB 的估测（详见表 1）。基于 RF 估测樟树 LAI 和 AGB，以光谱指数为自变量的模型测试集 R^2 高于波段反射率，在此模型下，光谱指数包含更多与樟树 LAI 和 AGB 相关的光谱特征信息。估测樟树 LAI 其余 4 个模型不同输入量，模型测试集的 R^2 差别不大，RMSE 均较低。整体来看，不同模型不同输入量来估测同一个指标（LAI 或 AGB），模型估测精度相差不大。

比较可知，XGBoost 模型在估测樟树叶面积指数和生物量中优于其他算法模型。XGBoost 模型以波段反射率为输入量，估测樟树 LAI 效果最好，模型训练集和测试集 R^2 分别为 0.899、0.872，RMSE 分别为 0.260、0.506。XGBoost 模型以光谱指数为输入量，估测樟树 AGB 精度最高，模型训练集和测试集 R^2 分别为 0.954、0.929，RMSE 分别为 550.041、587.746，说明 XGBoost 模型是估测樟树 LAI 和 AGB 的最佳模型。此外，GBDT 模型在 Python3.7 下建模速度极快（0.037s 以内），优于其余 4 个模型。

表 1 模型估测精度对比表

项目	模型	输入量	训练集		测试集	
			R^2	RMSE	R^2	RMSE
LAI	XGBoost	光谱指数	0.918	0.257	0.862	0.390
		波段反射率	0.899	0.260	0.872	0.506
	GBDT	光谱指数	0.922	0.193	0.856	0.347
		波段反射率	0.915	0.265	0.852	0.264
	RF	光谱指数	0.918	0.191	0.855	0.400
		波段反射率	0.927	0.184	0.823	0.445
	RBFNN	光谱指数	0.724	0.331	0.701	0.695
		波段反射率	0.778	0.270	0.716	0.522
	SVR	光谱指数	0.621	0.375	0.619	0.440
		波段反射率	0.628	0.379	0.609	0.453
AGB	XGBoost	光谱指数	0.954	550.041	0.929	587.746
		波段反射率	0.947	576.856	0.913	557.974
	GBDT	光谱指数	0.917	597.758	0.894	509.739
		波段反射率	0.915	577.032	0.909	725.561
	RF	光谱指数	0.873	598.332	0.862	827.374
		波段反射率	0.873	628.086	0.845	667.928
	RBFNN	光谱指数	0.749	845.087	0.717	941.614
		波段反射率	0.751	863.428	0.729	708.960
	SVR	光谱指数	0.751	844.610	0.749	712.076
		波段反射率	0.763	801.306	0.741	801.233

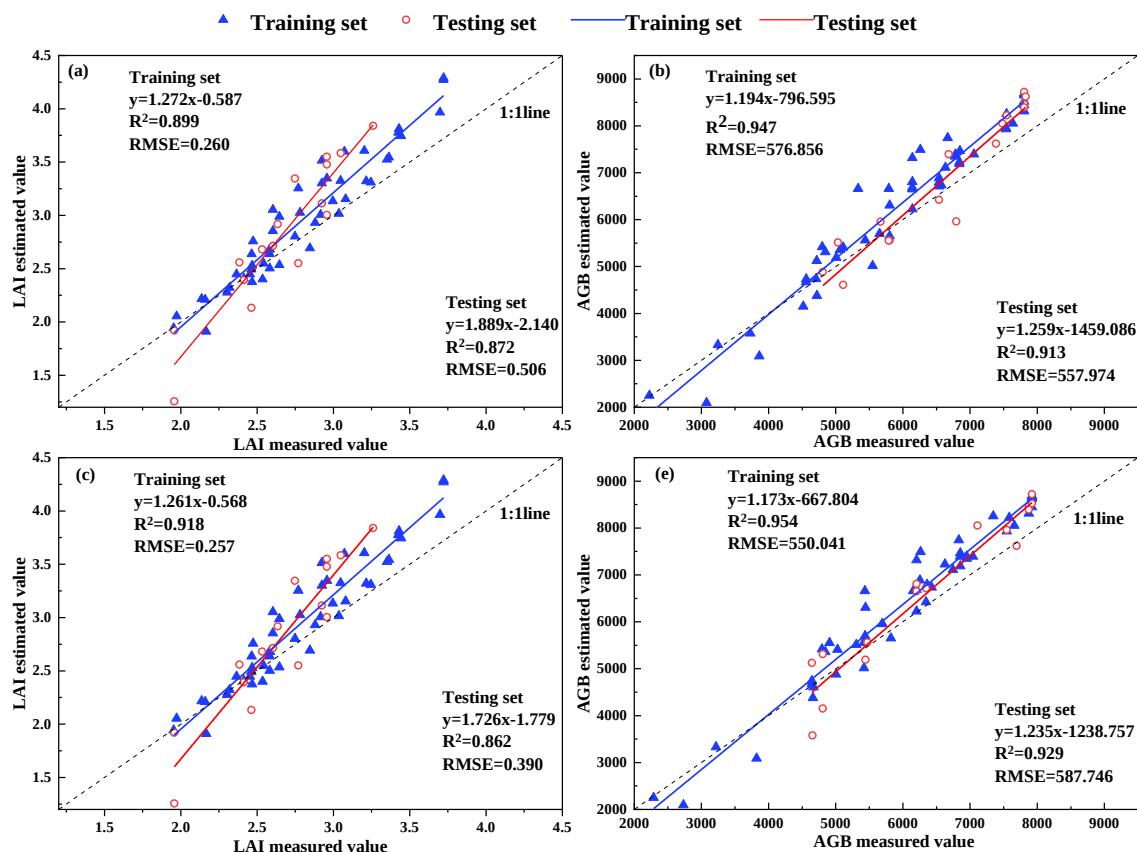


图4 基于XGBoost的樟树LAI和AGB估测模型训练集和测试集预测结果

(a)估测 LAI 输入变量为波段反射率；(b)估测 AGB 输入变量为波段反射率；(c)估测 LAI 输入变量为光谱指数；(d) 估测 AGB 输入变量为光谱指数。

4 讨论

叶面积指数和地上生物量是反映樟树矮林长势的重要指标，确定两个参数的敏感光谱指数是实现光谱快速无损监测的关键。本研究对蓝光、绿光、红光、红边 1、红边 2、近红外波段反射率与樟树叶面积指数和地上生物量进行相关性分析，发现樟树叶面积指数和地上生物量与红边波段相关性均最高，Fan 等^[34]研究认为不同生育期的油菜叶面积指数与近红外波段反射率相关更显著，与本研究结果不同，说明不同植物对光谱波段的敏感程度不同。红边波段是健康绿色植被在红光波段强吸收与近红外波段强反射，在 $0.67 \mu\text{m} \sim 0.74 \mu\text{m}$ 之间形成的窄条带区，与植物叶绿素、生物量和叶面积指数等生长生理参数紧密相关^[35]。在已有的研究中，大量的光谱指数应用于农业定量遥感，监测植物生长状况，如小麦^[36]、花生^[18]和玉米^[37]等。本文选取的差值光谱指数与樟树叶面积指数和生物量相关性显著，尤其是加入红边波段反射率的光谱指数，Dong 等^[38]研究了基于红边反射率和基于可见反射率的植被指数用于春小麦和油菜叶面积指数估测效果，发现基于红边反射率的光谱指数可用于开发针对不同作物的更通用的 LAI 估测模型，与本文研究结果相似。Mutanga 等^[39]基于红边和近红外波段计算的窄带光谱指数预测植被密集湿地地区生物量，有效改善估测精度。由此表明，红边波段反射率和基于红边波段反射率的光谱指数与樟树 LAI 和 SPAD 具有较强相关性。

在本研究采用的 5 个模型中，估测樟树 AGB 的模型测试集 R^2 均高于 LAI，说明在同一建模方法下，樟树 AGB 估测模型的精度普遍高于 LAI 估测模型，原因可能在于地上生物量更为直观，这与 BASCON 等^[40]的研究结果一致。估测樟树 AGB 模型的 RMSE 均远大于 LAI，原因在于 RMSE 对量纲较敏感，AGB

的数量级较大。与樟树 AGB 样本数据范围 2096~8721 相比,模型的 RMSE 数值较小。

本文通过相关性分析,选取 6 个光谱指数 ($DSI_{RE2,B}$ 、TVI、SAVI、 $DSI_{NIR,R}$ 、 $DSI_{RE1,B}$ 、EXG) 和 6 个波段反射率作为模型自变量,对于樟树 LAI,基于 GBDT、RF、SVR 估测模型,以光谱指数为模型自变量可以更好地估测樟树 LAI;基于 XGBoost、RBFNN 估测模型,以波段反射率为模型输入量估测 LAI 精度更优。对于樟树 AGB,基于 XGBoost、GBDT、RF、SVR 估测模型,以光谱指数为模型自变量可以更好地估测樟树 AGB;基于 RFNN 估测模型,较优模型自变量为波段反射率。综合来看,在同一模型中,多数模型估测 LAI 和 AGB 最佳自变量为光谱指数,这可能是由于光谱指数为多个波段反射率的线性或非线性组合,更能表征樟树 LAI 和 AGB 的光谱信息特征。此外,基于 RF 估测模型,不同自变量差异较明显,其余 4 个模型有差异但较小,说明变量选取对模型估测精度产生不同的影响。

估测樟树叶面积指数,RFNN 模型优于 SVR 模型;估测樟树地上生物量,SVR 模型优于 RFNN 模型,两个模型测试集的 R^2 在 0.60~0.75 之间。SVR 估测樟树 AGB 准确度高于 LAI,但模型的 R^2 仍较低,这可能是模型核函数和惩罚因子 C 等参数自身的局限性^[36]。RFNN 估测樟树 AGB 模型精度较低,原因可能在于模型在樟树 LAI 或 AGB 输入样本集中,初始中心过多,选取隐层基函数中心在大多数情况下难以反映系统真正的输出关系^[41]。基于 RF、GBDT、XGBoost 的樟树叶面积指数和地上生物量估测模型 R^2 均在 0.8 以上,对估测说明三种模型具有较好的估测能力。基于 RF、GBDT、XGBoost 的樟树 AGB 估测模型精度普遍高于 LAI 估测模型,这三种算法均属于基于决策树的机器学习算法,已有研究表明基于决策树的机器学习算法在估测植物生物量上有较明显优势^[42, 43]。研究发现,基于 RF 模型估测精度仅次于 XGBoost 和 GBDT,也具有较好的估测效果 (R^2 为 0.82 以上),RF 不易过拟合,有很好的抗噪声能力,由于其可调参数少、速度快、效率高、能够进行变量重要度计算和排列等优点,但由于其本身的复杂性,训练时间比其他类似算法更长^[44]。基于 GBDT 模型估测精度略次于 XGBoost,模型 R^2 为 0.85 以上,GBDT 的性能在 RF 的基础上又有一步提升,较适合本文所用的低维数据,并且在相对较少的调参时间下,估测准确度较高^[45]。基于 XGBoost 模型估测樟树叶面积指数和地上生物量的精度最高(表 1),优于其它算法模型,这与 Juan 等^[46]人利用无人机高光谱图像估算冬小麦叶面积指数、Li 等^[47]人利用卫星影像估算湖南省湘江流域不同类型森林生物量的研究结果一致。基于 GBDT 的模型估测精度略低于 XGBoost,原因可能是 XGBoost 对 GBDT 进一步优化了损失函数,XGBoost 可以自动地运用 CPU 的多线程进行并行计算,简化了模型,同时提高了预测准确度^[48]。XGBoost 算法是目前最快开源 boosting tree 工具包,对于中低维数据具有很好的处理速度和精度,已被许多领域的研究者广泛应用^[49-52],但其潜力在林业领域尚未得到充分发挥。

此外,本试验栽植的“赣芳 1 号”是江西省樟树繁育与开发利用工程研究中心自主申报的香料樟树优良品种,出油率、芳樟醇含量等关键指标方面表现良好。本研究结论不能直接应用不同时空的樟树矮林种植中,因为不同品种、不同时间和同地理条件的影响下,LAI 和 AGB 在多光谱遥感影像中可能会反映出不同的光谱特征。因此,后续研究中可在不同条件下分别进行建模估测,提高模型的广泛适用性。

5 结论

在本研究中,通过使用 Pearson 相关性分析,分别计算光谱指数、波段反射率与樟树叶面积指数和生物量之间的相关系数,结果表明相关性较显著。在选取的波段反射率和光谱指数中,樟树 LAI 分别与红边 2 波段反射率、 $DSI_{RE2,B}$ 相关性最高,樟树 AGB 分别与红边 1 波段反射率、 $DSI_{RE2,B}$ 相关性最高。基于红边波段的光谱指数与樟树 LAI 和 AGB 相关性较优。采用 5 种算法分别以光谱指数、原始波段反射率为模型输入量,估测樟树叶面积指数和地上生物量。通过综合比较分析,不同模型输入量,基于 RF 估测模型

精度有较明显的差异, 其余 4 个模型有差异但较小。在同一建模方法下, 樟树 AGB 估测模型的精度普遍高于 LAI 估测模型。不同建模方法, 基于 XGBoost 模型估测精度最高, 以波段反射率为模型输入量, 模型测试集 R^2 为 0.862, RMSE 为 0.390; 以光谱指数为模型输入量, 模型测试集 R^2 为 0.929, RMSE 为 587.746 $\text{kg}\cdot\text{hm}^{-2}$ 。研究表明基于无人机多光谱遥感图像, 采用 XGBoost 对樟树矮林 LAI 和 AGB 进行估测, 具有较好的预测效果, 对樟树矮林的管理具有一定指导作用。

参考文献:

- [1] LIU Z, LI H, ZHU Z, et al. Cinnamomum camphora fruit peel as a source of essential oil extracted using the solvent-free microwave-assisted method compared with conventional hydrodistillation [J]. LWT, 2022, 153: 112549.
- [2] GUO S, GENG Z, ZHANG W, et al. The chemical composition of essential oils from Cinnamomum camphora and their insecticidal activity against the stored product pests [J]. International journal of molecular sciences, 2016, 17(11): 1836.
- [3] LEE H J, HYUN E-A, YOON W J, et al. In vitro anti-inflammatory and anti-oxidative effects of Cinnamomum camphora extracts [J]. Journal of Ethnopharmacology, 2006, 103(2): 208-16.
- [4] SOKOLOVA A S, YAROVAYA O I, SHERNYUKOV A V, et al. Discovery of a new class of antiviral compounds: Camphor imine derivatives [J]. European Journal of Medicinal Chemistry, 2015, 105: 263-73.
- [5] YING-JIAO L, LI-JUAN X, LAN X, et al. Authentication of Two Different Chemical Types of Cinnamomum Camphora Leaves by Microscopic Technique with GC-MS and GC Analysis [J]. Pharmaceutical Chemistry Journal, 2020, 54: 154-61.
- [6] GOWER S T, KUCHARIK C J, NORMAN J M. Direct and Indirect Estimation of Leaf Area Index, fAPAR, and Net Primary Production of Terrestrial Ecosystems [J]. Remote Sensing of Environment, 1999, 70(1):29-51.
- [7] CHE S, DU G, WANG N, et al. Biomass estimation of cultivated red algae Pyropia using unmanned aerial platform based multispectral imaging [J]. Plant Methods, 2021,17(1):1-13.
- [8] WALTER J, EDWARDS J, MCDONALD G, et al. Photogrammetry for the estimation of wheat biomass and harvest index [J]. Field Crops Research, 2018,216:165-74.
- [9] ZHAO L H. Estimating Rice Yield by HJ-1A Satellite Images [J]. Rice Science, 2011.
- [10] MORAIN S A, WILLIAMS D L. Wheat Production Estimates Using Satellite Images [J]. Agronomy Journal, 1975,67(3):361-4.
- [11] SILVA M. Determination of biomass production of cotton using satellite images and spectral indexes [J]. 2017,(2).
- [12] XIN Q, PENG G, YU C, et al. A Production Efficiency Model-Based Method for Satellite Estimates of Corn and Soybean Yields in the Midwestern US [J]. Remote Sensing,2013,5(11):5926-43.
- [13] VON BUEREN S, BURKART A, HUENI A, et al. Comparative validation of UAV based sensors for the use in vegetation monitoring [J]. Biogeosciences Discussions, 2014,11(3):3837-64.
- [14] LU D, CHEN Q, WANG G, et al. Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates [J]. International Journal of Forestry Research,2012, 2012.
- [15] YU L, SHANG J, CHENG Z, et al. Assessment of Cornfield LAI Retrieved from Multi-Source Satellite Data Using Continuous Field LAI Measurements Based on a Wireless Sensor Network [J]. Remote Sensing, 2020,12(20):3304.
- [16] HANSEN P M, SCHJOERRING J K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression [J]. Remote Sensing of Environment, 2003, 86(4):542-53.
- [17] WANG C, NIE S, XI X, et al. Estimating the biomass of maize with hyperspectral and LiDAR data [J]. Remote Sensing, 2016,9(1):11.
- [18] QI H, ZHU B, WU Z, et al. Estimation of Peanut Leaf Area Index from Unmanned Aerial Vehicle Multispectral Images [J].Sensors, 2020,20(23):6732.

- [19] ZHENG H, LI W, JIANG J, et al. A comparative assessment of different modeling algorithms for estimating leaf nitrogen content in winter wheat using multispectral images from an unmanned aerial vehicle [J]. *Remote Sensing*, 2018,10(12):2026.
- [20] SU J, LIU C, COOMBES M, et al. Wheat Yellow Rust Monitoring by Learning from Multispectral UAV Aerial Imagery [J]. *Computers and Electronics in Agriculture*, 2018, 155.
- [21] FENG A, ZHOU J, VORIES E D, et al. Yield estimation in cotton using UAV-based multi-sensor imagery [J]. *Biosystems Engineering*, 2020,193:101-14.
- [22] SASAKI Y, OKAMOTO T, IMOU K, et al. Automatic diagnosis of plant disease - spectral reflectance of healthy and diseased leaves [J]. *IFAC Proceedings Volumes*, 1998, 31(5):145-50.
- [23] PEÑUELAS J, FILELLA I, BIEL C, et al. The reflectance at the 950–970 nm region as an indicator of plant water status [J]. *International journal of remote sensing*, 1993, 14(10):1887-905.
- [24] BROGE N H, LEBLANC E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density [J]. *Remote sensing of environment*, 2001,76(2):156-72.
- [25] PHAM T D, YOKOYA N, XIA J, et al. Comparison of Machine Learning Methods for Estimating Mangrove Above-Ground Biomass Using Multiple Source Remote Sensing Data in the Red River Delta Biosphere Reserve, Vietnam [J]. *Remote Sensing*, 2020, 12(8).
- [26] YUAN Y, WANG X, SHI M, et al. Performance comparison of RGB and multispectral vegetation indices based on machine learning for estimating *Hopea hainanensis* SPAD values under different shade conditions [J]. *Frontiers in plant science*, 2022: 2615.
- [27] SIEGMANN B, JARMER T. Comparison of different regression models and validation techniques for the assessment of wheat leaf area index from hyperspectral data [J]. *International Journal of Remote Sensing*, 2015,36(18):4519-34.
- [28] BLACKARD J, FINCO M, HELMER E, et al. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information [J]. *Remote Sensing of Environment*, 2008.
- [29] CARREIRAS J M B V, M.J.LUCAS, R.M. Understanding the relationship between aboveground biomass and ALOS PALSAR data in the forests of Guinea-Bissau (West Africa) [J]. *Remote Sensing of Environment: An Interdisciplinary Journal*, 2012, 121(Null).
- [30] LIU T, SHI T, ZHANG H, et al. Detection of Rise Damage by Leaf Folder (*Cnaphalocrocis medinalis*) Using Unmanned Aerial Vehicle Based Hyperspectral Data [J]. *Sustainability*, 2020, 12.
- [31] RICHARDSON A J, WIEGAND C L. Distinguishing vegetation from soil background information [J]. *Pe & Rs*, 1977, 43(12).
- [32] HUETE A R. A soil-adjusted vegetation index (SAVI) [J]. *Remote sensing of environment*, 1988,25(3): 295-309.
- [33] WOEBBECKE D M, MEYER G E, VON BARGEN K, et al. Color indices for weed identification under various soil, residue, and lighting conditions [J]. *Transactions of the ASAE*, 1995,38(1):259-69.
- [34] FAN H, LIU S, LI J, et al. Early prediction of the seed yield in winter oilseed rape based on the near-infrared reflectance of vegetation (NIRv) [J]. *Computers and Electronics in Agriculture*, 2021, 186: 106166.
- [35] FILELLA I, PENUELAS J. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status [J]. *International journal of remote sensing*, 1994,15(7):1459-70.
- [36] TANG Z, GUO J, XIANG Y, et al. Estimation of Leaf Area Index and Above-Ground Biomass of Winter Wheat Based on Optimal Spectral Index [J]. *Agronomy*, 2022,12(7):1729.
- [37] LIU S, JIN X, NIE C, et al. Estimating leaf area index using unmanned aerial vehicle data: shallow vs. deep machine learning algorithms [J]. *Plant Physiology*, 2021,187(3):1551-76.
- [38] DONG T, LIU J, SHANG J, et al. Assessment of red-edge vegetation indices for crop leaf area index estimation [J]. *Remote Sensing of Environment*, 2019,222:133-43.
- [39] MUTANGA O, ADAM E, CHO M A. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2012,18:399-406.
- [40] BASCON M V, NAKATA T, SHIBATA S, et al. Estimating Yield-Related Traits Using UAV-Derived Multispectral Images to Improve Rice Grain Yield Prediction [J]. *Agriculture*, 2022,12(8):1141.

- [41] MASHOR M Y. Hybrid training algorithm for RBF network [J]. International Journal of the computer, the Internet and Management, 2000,8(2):50-65.
- [42] BLACKARD J, FINCO M, HELMER E, et al. Mapping US forest biomass using nationwide forest inventory data and moderate resolution information [J]. Remote sensing of Environment, 2008,112(4): 1658-77.
- [43] CARREIRAS J M, VASCONCELOS M J, LUCAS R M. Understanding the relationship between aboveground biomass and ALOS PALSAR data in the forests of Guinea-Bissau (West Africa) [J]. Remote Sensing of Environment, 2012,121:426-42.
- [44] JORDAN M I, MITCHELL T M. Machine learning: Trends, perspectives, and prospects [J]. Science, 2015, 349(6245):255-60.
- [45] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine [J]. Annals of statistics, 2001:1189-232.
- [46] ZHANG J, CHENG T, GUO W, et al. Leaf area index estimation model for UAV image hyperspectral data based on wavelength variable selection and machine learning methods [J]. Plant Methods, 2021, 17(1): 49.
- [47] LI Y, LI C, LI M, et al. Influence of variable selection and forest type on forest aboveground biomass estimation using machine learning algorithms [J]. Forests, 2019, 10(12): 1073.
- [48] CHEN T, GUESTRIN C. Xgboost: A scalable tree boosting system; proceedings of the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, F, 2016 [C].
- [49] TORLAY L, PERRONE-BERTOLOTI M, THOMAS E, et al. Machine learning–XGBoost analysis of language networks to classify patients with epilepsy [J]. Brain informatics, 2017,4(3):159-69.
- [50] NOBRE J, NEVES R F. Combining Principal Component Analysis, Discrete Wavelet Transform and XGBoost to trade in the financial markets [J]. Expert Systems with Applications, 2019,125:181-94.
- [51] SONG K, YAN F, DING T, et al. A steel property optimization model based on the XGBoost algorithm and improved PSO [J]. Computational Materials Science, 2020,174:109472.
- [52] LI J, AN X, LI Q, et al. Application of XGBoost algorithm in the optimization of pollutant concentration [J]. Atmospheric Research, 2022, 276:106238.