



基于Sentinel-2A数据的东北森林植物多样性监测方法研究

周楷玲^{1,2} 赵玉金^{1*} 白永飞^{1,3*}

¹中国科学院植物研究所植被与环境变化国家重点实验室, 北京 100093; ²中国科学院大学生命科学学院, 北京 100049; ³中国科学院大学资源与环境学院, 北京 100049

摘要 植物多样性监测是开展生物多样性评估, 制定生物多样性保护政策的基础。传统的森林植物多样性监测以实地调查为主, 难以快速获取森林植物多样性的空间分布及其动态变化信息。遥感技术的发展为评估区域尺度森林植物多样性提供了重要工具。该研究选取凉水、丰林和珙春3个国家级自然保护区, 利用Sentinel-2A卫星影像和野外实测数据, 探讨了基于像元和聚类的光谱多样性直接估算方法, 以及基于随机森林回归的森林植物多样性反演方法。研究结果表明: (1)在像元尺度, 基于凸包面积计算的光谱多样性指数对Shannon-Wiener多样性指数的估算精度($R^2 = 0.74$)优于基于变异系数的方法($R^2 = 0.60$); (2)基于像元的光谱多样性估算方法对Shannon-Wiener多样性指数的估算精度优于聚类分析方法($R^2 = 0.59$); (3)基于6个特征变量, 利用随机森林回归算法对Shannon-Wiener多样性指数的估算精度最高($R^2 = 0.79$); (4)上述方法均不能精确估算Simpson多样性指数和物种丰富度。研究发现基于Sentinel-2A卫星影像能较好地反演Shannon-Wiener多样性指数, 为下一步能在大尺度上进行森林植物多样性估算提供了参考和依据。

关键词 森林植物多样性; Sentinel-2A; 光谱多样性; 聚类分析; 随机森林回归

周楷玲, 赵玉金, 白永飞 (2022). 基于Sentinel-2A数据的东北森林植物多样性监测方法研究. 植物生态学报, 46, 1251-1267. DOI: 10.17521/cjpe.2021.0373

Study on forest plant diversity monitoring based on Sentinel-2A satellite data in northeast China

ZHOU Kai-Ling^{1,2}, ZHAO Yu-Jin^{1*}, and BAI Yong-Fei^{1,3*}

¹State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Science, Beijing 100093, China; ²College of Life Science, University of Chinese Academy of Sciences, Beijing 100049, China; and ³College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract

Aims Plant diversity monitoring is the basis of biodiversity assessment and developing conservation policy. Traditional forest plant diversity monitoring is mainly based on field surveys, which is difficult to quickly obtain the spatial distribution and dynamic change of forest plant diversity. The development of remote sensing technology provides an important tool for assessing forest plant diversity at the regional scale. In this study, we explored two methods of forest plant diversity estimation based on Sentinel-2A satellite images and field data in three selected national nature reserves (Liangshui, Fenglin, and Hunchun).

Methods We used two methods to estimate forest plant diversity: (1) Direct estimation based on spectral diversity at the pixel and cluster scales, respectively; (2) Indirect estimation based on random forest regression. The spectral diversity was calculated based on the coefficient of variation and convex hull area at the pixel scale, respectively. K-means clustering method was used for cluster analysis to calculate the spectral diversity between clusters. For the indirect estimation, we used 10-fold cross validation to select characteristic variables for later diversity calculation.

Important findings Our results showed that: (1) At the pixel scale, the estimation accuracy of Shannon-Wiener diversity index based on convex hull area ($R^2 = 0.74$) was better than that of coefficient of variation ($R^2 = 0.60$); (2) The pixel-based estimation accuracy of Shannon-Wiener diversity index outperformed clustering basis ($R^2 = 0.59$); (3) Based on six feature variables, the Shannon-Wiener diversity index was best estimated using the random

收稿日期Received: 2021-10-15 接受日期Accepted: 2022-01-14

基金项目: 中国科学院战略性先导科技专项(A类)(XDA23080303)。Supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDA23080303).

* 通信作者Corresponding authors (Zhao YJ, zhaoyj@ibcas.ac.cn; Bai YF, yfbai@ibcas.ac.cn)

forest regression algorithm ($R^2 = 0.79$); (4) Both the Simpson diversity index and species richness could not be accurately estimated by the above methods. Our findings indicate the capability of Sentinel-2A satellite images to estimate the Shannon-Wiener diversity index, providing reference and basis for forest plant diversity estimation at a large scale.

Key words forest plant diversity; Sentinel-2A; spectral diversity; cluster analysis; random forest regression

Zhou KL, Zhao YJ, Bai YF (2022). Study on forest plant diversity monitoring based on Sentinel-2A satellite data in northeast China. *Chinese Journal of Plant Ecology*, 46, 1251-1267. DOI: 10.17521/cjpe.2021.0373

森林是陆地生态系统的重要组成部分, 占全球陆地面积的30%, 在维持生物多样性和生物圈功能方面发挥着重要作用(Pan *et al.*, 2013; Harrison *et al.*, 2014)。然而, 由于人类活动和气候变化的影响, 生物多样性正在以一种前所未有的速度丧失(Ceballos *et al.*, 2015)。及时掌握森林生物多样性的现状、格局、变化趋势和受到的威胁, 是制定生物多样性保护政策和措施的前提。长久以来, 森林植物多样性监测依赖于大量的野外调查, 费时费力, 调查结果也受样地选择、调查方法、抽样调查的力度、参与调查人员的专业知识等诸多因素限制(Gotelli & Colwell, 2001; Graham & Hijmans, 2006)。同时, 野外调查集中在物种和样地水平, 难以在景观、区域乃至全球尺度上实现对森林植物多样性在时空上的连续监测(Kerr & Ostrovsky, 2003; Nicholson *et al.*, 2009; Levrel *et al.*, 2010; Pereira *et al.*, 2010)。

近年来, 随着遥感技术的迅速发展, 使得在大尺度、多时空上评估生物多样性成为可能。遥感数据具有覆盖范围广、可持续性较强以及可重复的特点(Nagendra, 2001; Duro *et al.*, 2007), 有助于迅速地揭示大面积生物多样性丢失状况, 对大尺度、长时间序列的生物多样性评估至关重要(Turner, 2014)。生物多样性遥感监测方法可总结为直接法和间接法两种(Turner *et al.*, 2003)。直接法是利用高空间分辨率和高光谱分辨率的卫星传感器直接识别物种、种群或群落。Ke等(2010)利用QuickBird高分辨率多光谱数据和激光雷达数据, 通过图像分割和基于目标的分类方法对森林物种进行识别, 发现两种数据均有较好的结果, 但是结合两种数据源能获取更高的精度, Kappa系数最高可达91.6%。Cross等(2019)利用WorldView-3卫星影像对热带森林里的6个森林物种进行了分类, 分类准确率达85.37%。间接法是通过遥感数据反演与生物多样性密切相关的指标或参数, 再与野外实测数据构建统计模型以反演生物多样性。植被指数、植物生化组分和植被结构等都

可以作为植物多样性研究的相关指标。归一化植被指数(NDVI)经常被用来指示区域的物种多样性(Fairbanks & Mcgwire, 2004; Feeley *et al.*, 2005), 而增强型植被指数(EVI)则可以更好地在植被茂密的地区代替NDVI(Waring *et al.*, 2006)。此外, 与光谱变异相关的叶片叶绿素、氮含量、色素、水分等生化组分的变化也可以用来估算森林物种多样性(Carlson *et al.*, 2007)。

光谱变异假说是生物多样性直接遥感监测的主要理论基础, 即遥感图像光谱多样性与物种多样性有关。光谱多样性, 有时也被称为光学多样性(Palmer *et al.*, 2002; Ustin & Gamon, 2010)、光谱异质性或光谱变异性(Rocchini *et al.*, 2010), 表征光谱反射率的空间变化, 这种变化与植物多样性相关的植物光学性状的差异有关, 可以用来反映物种多样性(Gillespie *et al.*, 2008; Nagendra & Rocchini, 2008)。基于光谱变异假说的光谱多样性与物种多样性的研究已经有很多, 在温带森林(Laliberté *et al.*, 2020), 亚热带森林(Kalacska *et al.*, 2007), 热带森林(Asner & Martin, 2009; Féret & Asner, 2014; Schäfer *et al.*, 2016), 草原(Oldeland *et al.*, 2010; Wang *et al.*, 2016, 2018; Gholizadeh *et al.*, 2020)均有报道。常用光谱多样性指数包括变异系数(CV)(Somers *et al.*, 2015; Wang *et al.*, 2016), 凸包体积(CHV)(Dahlin, 2016; Gholizadeh *et al.*, 2018), 凸包面积(CHA)(Gholizadeh *et al.*, 2018), 光谱角度制图(SAM)(Zhang *et al.*, 2006; Gholizadeh *et al.*, 2018), 光谱信息散度(SID)(Chang, 2000)等。

然而, 由于物种内的光谱变异会混淆物种间的光谱多样性, 进一步影响植物多样性的预测精度。聚类分析作为一种非监督分类绘制植物多样性图的有效方法, 可以基于聚类结果进一步计算光谱异质性, 能够在一定程度上减少物种内光谱变异(Medina *et al.*, 2013; Féret & de Boissieu, 2020)。目前使用非监督分类方法的森林植物多样性研究还较少。Féret

和de Boissieu (2020)基于Sentinel-2光谱影像, 利用K均值聚类生成光谱物种图, 再以此绘制了亚马孙某片森林的 α 和 β 多样性。也有研究基于分辨率更高的高光谱影像, 通过不同聚类方法反演植物多样性。Medina等(2013)基于AISA机载高光谱数据, 利用光谱方差估算物种丰富度, 发现基于层次聚类得到的Shannon-Wiener多样性指数与野外数据之间的相关性是可变的, 甚至在某些情况下是负的, 但在聚类后进行光谱解混可以显著改善。Zhao等(2018)利用机载高光谱和激光雷达数据对神农架自然保护区森林冠层的单木进行分割, 利用自适应模糊C均值聚类方法, 估算了该区域的树种丰富度和Shannon-Wiener多样性指数($R^2 = 0.83$, 均方根误差(RMSE) = 0.25)。

除了基于光谱多样性的方法, 基于机器学习的方法在森林植被分类和植物多样性监测中也得到了广泛的关注。随机森林(Breiman, 2001)是一种广泛被用于预测的非参数的机器学习算法, 它不假设数据的正态分布, 具有对噪声不敏感、分类更快更稳定的特点(Clark & Roberts, 2012; Miao *et al.*, 2012), 是用于森林多样性建模的最优方法之一(Mallinis *et al.*, 2020; Gyamfi-Ampadu *et al.*, 2021)。随机森林里包括用于评估模型准确性的集成交叉验证, 能对重要变量进行排序, 并以非线性的方式处理交互数据。Erinjeri等(2018)采用最大似然法和随机森林分类的方法, 基于Sentinel-2光谱波段及其衍生的NDVI和纹理、Sentinel-1 SAR波段及其纹理对高海拔热带雨林不同植被类型图进行了绘制, 分类精度可达75%以上。Mallinis等(2020)基于4种不同分辨率的卫星数据, 利用随机森林回归的方法对森林植物多样性进行了评估, 结果发现基于WorldView-2图像生成的模型精度最高, Shannon-Wiener多样性指数估算精度为0.44, Simpson多样性指数为0.37。

由于遥感数据的成本效益和区域可用性, 已经成为评估生态系统生物多样性的工具。尽管利用无人机、激光雷达、高分影像或高光谱可提高多样性的预测精度, 但其获取困难、计算复杂、花费相对较高且在大区域上很难实现。基于多光谱数据反演生物多样性在大区域植被和森林覆盖中有着更广泛的应用。Sentinel-2号卫星数据是目前可免费获取的最高空间分辨率的多光谱遥感数据, 也是唯一一个有3个红边带的卫星遥感图像, 它里面包含的

叶绿素信息对植被的变化敏感, 对植被的健康监测有重要作用。本研究拟利用Sentinel-2A影像, 在研究区内分别测试基于像元和聚类分析的光谱多样性直接估算方法以及随机森林回归两种目前主流的反演森林植物多样性的方法, 以期发现Sentinel-2A数据在估算森林植物多样性方面的潜力, 为能在更大的尺度上进行森林植物多样性估算及相关研究提供可行的方法参考和依据, 为森林生态学研究 and 森林管理提供参考。

1 材料和方法

1.1 研究区概况

研究区选取东北区域涵盖针阔混交林、落叶阔叶林及落叶针叶林等不同森林类型的3个国家级自然保护区, 分别是凉水(128.80°–128.93° E, 47.12°–47.24° N)、丰林(128.98°–129.25° E, 48.03°–48.20° N)和珲春(130.29°–131.24° E, 42.41°–43.47° N)国家级自然保护区, 3个保护区属同一片区, 地理位置接近, 气候土壤条件相似, 主要森林植被类型有一定差异, 以提升模型的可移植性(图1)。

凉水国家级自然保护区总面积为6 394 hm^2 。气候类型为典型的温带大陆性季风气候, 年平均气温 $-0.3\text{ }^{\circ}\text{C}$, 年降水量676 mm, 积雪期130–150天。保护区内地形条件较为复杂, 属典型丘陵地区, 海拔275–720 m。区内土壤为暗棕壤, 顶极森林群落类型为阔叶红松林, 主要树种有红松(*Pinus koraiensis*)、臭冷杉(*Abies nephrolepis*)、白桦(*Betula platyphylla*)、紫椴(*Tilia amurensis*)、红皮云杉(*Picea koraiensis*)、水曲柳(*Fraxinus mandshurica*)等, 保护区周围的人工林类型主要是落叶松(*Larix gmelinii*)、红松和樟子松(*Pinus sylvestris* var. *mongolica*)人工林。

丰林国家级自然保护区总面积为18 165 hm^2 , 海拔约为300–450 m, 最高海拔为688 m。大陆性季风气候, 年平均气温为 $-0.5\text{ }^{\circ}\text{C}$, 年降水量640.5 mm, 主要集中在6–9月。丰林保护区的主要保护对象为原始阔叶红松林, 是全球原始红松林分布的中心地带, 是世界上最典型和完整的以红松为主的北温带针阔叶混交林生态系统。

珲春国家级自然保护区总面积为108 700 hm^2 , 保护区外围有因保护对象活动范围较大而设立的外围保护带, 面积为41 778 hm^2 。保护区气候类型为近海中温带海洋性季风气候, 年平均气温 $5.65\text{ }^{\circ}\text{C}$ 。保

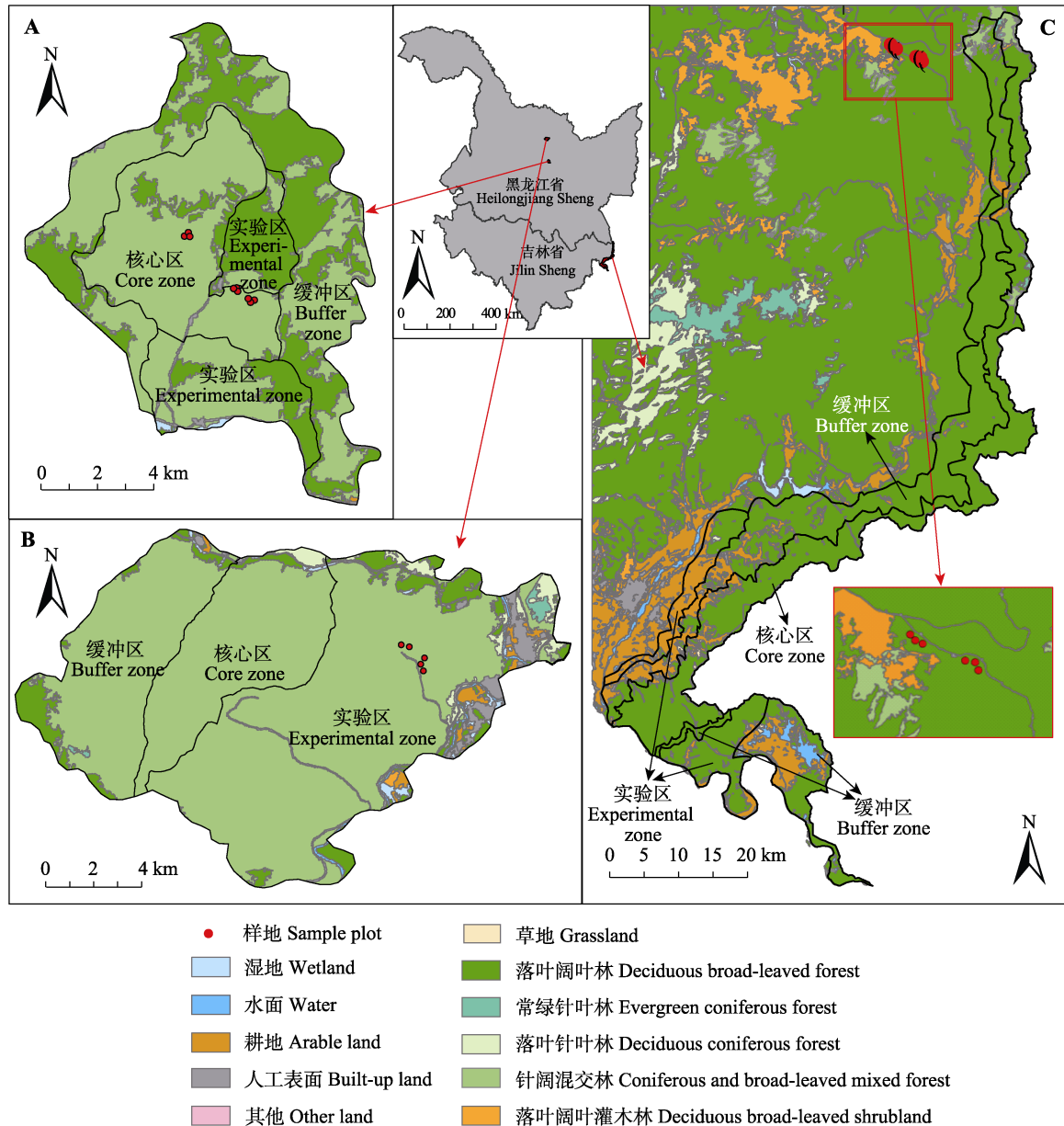


图1 东北森林研究区地理位置。A, 凉水国家级自然保护区。B, 丰林国家级自然保护区。C, 珲春国家级自然保护区。

Fig. 1 Location of the northeast forest study area. A, Liangshui National Nature Reserve. B, Fenglin National Nature Reserve. C, Hunchun National Nature Reserve.

保护区南北跨越珲春全境, 山地主要分布于北、中和南部的部分地区, 北部全区最高点海拔达973 m, 南部最低点海拔5 m, 高差大。植被类型以阔叶林为主, 北部有少量针叶林和针阔混交林。主要保护树种为东北红豆杉(*Taxus cuspidata*)、红松、紫椴、黄檗(*Phellodendron amurense*)、水曲柳等。区内野生动物资源丰富, 其中东北虎(*Panthera tigris* ssp. *altaica*)、东北豹(*Panthera pardus orientalis*)、梅花鹿(*Cervus nippon*)等9种动物被列为国家I级重点保护野生动物。

自然保护区可分为核心区、缓冲区和实验区3个功能区。核心区是指保护区内未经或很少经过人为干扰的区域, 保护效果最好, 在核心区外围可划分出一定面积的允许从事科研活动的缓冲区, 在缓冲区周围的为实验区, 为人类活动和规划发展的区域(Ma *et al.*, 2009; Xu *et al.*, 2016)。合理的功能区划可以在促进生物多样性保护的同时为当地的经济文化发展作出贡献, 对整个自然保护区的长远发展有重要意义。本研究所选保护区的功能区划概况如图1所示。

1.2 数据及预处理

1.2.1 野外数据

于2019年7–8月在凉水、珲春、丰林国家级自然保护区开展群落和物种调查,然而,在选取珲春自然保护区样点时由于保护区规划和安全问题,研究点选在保护区北部东北虎豹国家公园。每个保护区各挑选2–4种具有代表性的林分类型,每种林分类型下各设置3个30 m × 30 m样地,共获得了20个调查样地,记录样地4角的GPS坐标。在每个样地内进行每木检尺,记录样方内出现的所有树种名称、树高、胸径(≥5 cm)、枝下高及冠幅等参数。同时记录样地土壤类型和郁闭度等信息,样地基本情况如表1所示。其中,1–5号样地位于丰林保护区内,是以红松、青楷槭(*Acer tegmentosum*)、紫椴、硕桦(*Betula costata*)等为主要树种的阔叶红松林。6–11号样地属于珲春保护区范围,主要植被类型是以红松、紫椴、白桦、臭冷杉等为主要树种的针阔混交林。12–20号样地位于凉水保护区内,12–14号样地是以臭冷杉、鱼鳞云杉(*Picea jezoensis* var. *microsperma*)为主的混交林,伴生着白桦、花楷槭(*Acer*

ukurunduense)、青楷槭、春榆(*Ulmus davidiana* var. *japonica*)等多种阔叶树种,16–20号样地是以红松、鱼鳞云杉、白桦、水曲柳为主要树种的针阔混交林。

根据样方调查结果,计算样方内胸径≥5 cm的树种的物种丰富度(S)、Shannon-Wiener多样性指数(H')和Simpson多样性指数(D)。物种丰富度用样方中的乔木物种数表示。计算公式如下:

$$H' = - \sum_{i=1}^n P_i \ln P_i \quad (1)$$

$$D = 1 - \sum_{i=1}^n P_i^2 \quad (2)$$

式中, n 为样方中的物种数, P_i 为第 i 个物种的个体数占所有物种总个体数的比例(Pielou, 1966)。

1.2.2 遥感数据及数据预处理

同步于地面调查,选取2019年7–8月的Sentinel-2A卫星影像开展森林植物多样性遥感监测。该卫星携带多光谱成像仪(MSI),轨道高度为786 km,幅宽为290 km,卫星的重访周期为10 d,与Sentinel-2B互补,重访周期缩短至5 d。影像覆盖了13个光谱波段,其中包括4个可见光波段(蓝色、绿色、红色和

表1 东北森林样地调查信息表

Table 1 Sample plot survey information of northeast forest

样地 Sample plot	地理位置 Geographical position	郁闭度 Crown density	物种丰富度 Number of species	Shannon-Wiener多样性指数 Shannon-Wiener diversity index	Simpson多样性指数 Simpson diversity index
1	129.19° E, 48.12° N	0.85	6	1.56	0.76
2	129.19° E, 48.12° N	0.75	10	1.89	0.82
3	129.19° E, 48.12° N	0.80	7	1.67	0.79
4	129.18° E, 48.13° N	0.85	10	1.92	0.82
5	129.18° E, 48.13° N	0.75	10	1.96	0.83
6	131.11° E, 43.44° N	0.70	11	0.87	0.83
7	131.11° E, 43.44° N	0.60	10	0.88	0.85
8	131.11° E, 43.44° N	0.65	8	0.77	0.78
9	131.07° E, 43.47° N	0.60	6	0.65	0.74
10	131.07° E, 43.47° N	0.70	4	0.44	0.56
11	131.07° E, 43.47° N	0.60	8	0.73	0.74
12	128.90° E, 47.18° N	0.60	5	1.34	0.80
13	128.89° E, 47.18° N	0.70	7	1.27	0.66
14	128.89° E, 47.18° N	0.70	6	1.56	0.82
15	128.89° E, 47.18° N	0.65	7	1.01	0.73
16	128.89° E, 47.19° N	0.60	6	1.17	0.72
17	128.89° E, 47.18° N	0.60	4	0.96	0.67
18	128.86° E, 47.20° N	0.70	10	2.06	0.86
19	128.86° E, 47.20° N	0.70	11	2.17	0.84
20	128.86° E, 47.20° N	0.70	9	1.82	0.79

近红外波段), 空间分辨率为10 m, 短波红外波段(SWIR1和SWIR2)、卷云波段(SWIR-Cirrus)和红边波段(红边波段1-4), 空间分辨率为20 m, 其余两个波段的分辨率为60 m。Sentinel-2A数据有3个波段数据在红边范围内, 这对植被健康的监测十分有效。

Sen2Cor (<http://step.esa.int/main/snap-supported-plugins/sen2cor/>)是Sentinel-2 Level 2A产品生成和格式化的工具集, 利用安装的Sen2Cor插件对下载的L1C级遥感影像进行辐射定标和大气校正, 大气校正后的影像去除了B10卷云波段。对于经过大气校正后的影像在SNAP (<http://step.esa.int/main/download/snap-download/>)中进行重采样, 分辨率设置为10 m, 最后将处理后的影像在ENVI软件中进行波段的调整、合并、镶嵌和裁剪等处理。对裁剪之后的影像进行辐射滤波, 通过选取NDVI的阈值大于0.3以去除土壤、建筑等非植被区域的影响, 设置近红外和蓝色光谱波段的阈值分别小于1 500和500以去掉阴影区域和云的影响。

1.3 研究方法

本研究测试了两种目前主流的森林植物多样性的反演方法, 包括基于原始影像像元和基于聚类分析的光谱多样性方法和基于随机森林的植物多样性回归的估算方法。基于像元的方法是通过计算变异系数和凸包面积两种光谱多样性指数以反演森林植物多样性。聚类分析主要采用K均值聚类和自定义模糊C均值聚类, 然而自定义模糊C均值聚类预测结果并不理想, 因此本研究主要采用K均值聚类方法。随机森林是目前比较流行的机器学习方法, 本研究选取了与植物多样性相关的30个植被指数和12个特征波段, 通过随机森林回归估算森林植物多样性。考虑到样方的定位误差和边缘效应, 本研究选取5 × 5移动窗口(即50 m × 50 m)计算了光谱变异系数和光谱特征值。具体的流程如图2所示。

1.3.1 变异系数

光谱变异系数是 α 多样性的一种常用度量(Wang *et al.*, 2016)。根据光谱变异假说, CV 随着光谱

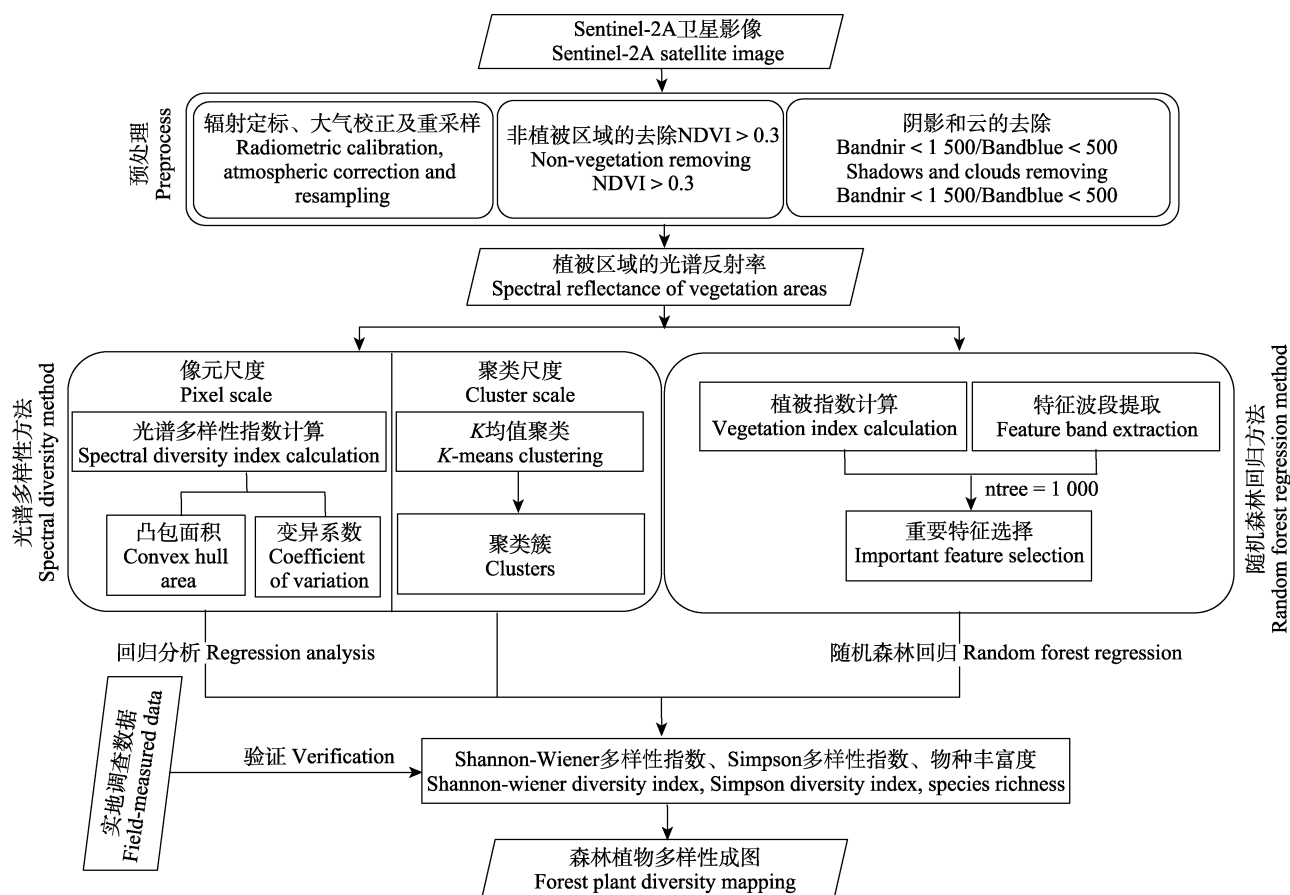


图2 东北森林植物多样性研究方法流程图。NDVI, 归一化植被指数。

Fig. 2 Flow chart of the research methods for forest plant diversity in northeast China. NDVI, normalized difference vegetation index.

变异性的增加而增加。本研究基于重采样后的Sentinel-2A卫星数据的12个波段计算了 5×5 移动窗口的变异系数。变异系数的计算公式如下:

$$CV = \frac{\sum_{\lambda=443}^{2190} \frac{\text{std}(\rho_{\lambda})}{\text{mean}(\rho_{\lambda})}}{n} \quad (3)$$

式中, ρ_{λ} 表示波长 λ 处的反射率, $\text{std}(\rho_{\lambda})$ 和 $\text{mean}(\rho_{\lambda})$ 表示所有像元波长 λ 处的标准差和平均值, n 为波段数量。

1.3.2 凸包面积

凸包面积是一种新的光谱多样性指标(Deng *et al.*, 2016; Gholizadeh *et al.*, 2018)。CHA是样方 L 的平均光谱和该样方内第 V 个像元的光谱所组成的二维空间中的凸包面积。如果一个像元的平均光谱和该区域的平均光谱相似或与其高度相关, 这些波段在图里的位置会近乎在一条直线上, 凸包面积趋于0。样方平均CHA的表达式如下:

$$\overline{\text{CHA}}_L = \frac{\sum_{V=1}^m \text{CHA}(R_{V,L}, \bar{R}_L)}{m} \quad (4)$$

式中, m 为该样方中所有像元的数量, \bar{R}_L 是样方 L 的平均光谱, $R_{V,L}$ 是样方 L 内第 V 个像元的光谱。物种丰富度越高的样地, 与光谱的平均值偏差越大, CHA越大。计算CHA时需要将波段值进行归一化处理。

1.3.3 K均值聚类

K均值聚类是一种被广泛应用的非监督分类方法, 它的主要思想为: 将数据自动划分为 K 个组, 随机创建一个初始的聚类中心, 然后采用迭代的方法, 通过不断移动聚类中心不断尝试改进划分。具体的处理流程是: 首先选择 K 个中心点, 再将每个数据点按与聚类中心的距离重新划分到离其最近的中心点, 计算每个聚类中的点到该类中心点距离的平均值, 用平均值取代每个聚类中心, 再重复前两个步骤, 直到所有的观测值不再被分配或是达到最大的迭代次数。

本研究利用biodivMapR软件包, 基于Sentinel-2A卫星影像的多光谱信息, 根据每个像元反射率对应的光学特征区分物种(Ustin & Gamon, 2010)。首先需要确定聚类窗口大小, 以窗口内能够计算多样性所需要的像元最少为原则, 综合考虑影像分辨率和现场调查的样地大小等因素, 通过测试 3×3 、 5×5 、 7×7 以及 9×9 等不同移动窗口(分辨率为10 m), 最

终选取 5×5 的移动窗口为计算植物多样性最好的聚类窗口; 然后, 将预处理后的影像进行主成分分析, 选择2–6个能够突出植物属性、噪声较少的主成分, 再利用K均值聚类绘制波谱物种图, 最后基于每个光谱物种包含的像元数, 计算多样性指数。

1.3.4 随机森林回归

随机森林是Leo Breiman在2001年提出的一种基于分类和回归树的集成学习算法, 可以解释若干个自变量对因变量的作用。本研究使用随机森林回归算法构建了基于图像的森林多样性反演模型, 算法的基本流程是: (1)利用bootstrap重采样, 从原始样本 n 中随机抽取 m 个样本点, 得到新的训练集, 再利用每个训练集生成对应的CART回归树, 测试样本为每次重采样未被抽到的样本所构成的袋外数据(OOB); (2)随机抽取 k 个特征作为每个节点的分裂特征集, 再从 k 个特征中选择最优的分裂方式对该节点进行分裂; (3)每棵回归树自顶向下递归分枝, 直到满足分割终止条件, 让获得的CART回归树都能得到最大限度的生长; (4)生成的 m 棵回归树构成了随机森林回归模型, 每一棵回归树最终的预测结果为该样本点所到叶节点的均值, 最终的预测结果为所有CART回归树估算结果的均值。回归的精度评价采用OOB预测的残差均方(MSE)表示如下:

$$MSE_{\text{OOB}} = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i^{\text{OOB}})^2 \quad (5)$$

$$R_{\text{RF}}^2 = 1 - \frac{MSE_{\text{OOB}}}{\hat{\sigma}_y^2} \quad (6)$$

式中, y_i 表示袋外数据中因变量的实际值, \hat{y}_i 表示对袋外数据的预测值, $\hat{\sigma}_y^2$ 表示袋外数据预测值的方差, R_{RF}^2 表示拟合优度。

随机特征的选取是基于影像的光谱信息以及衍生的植被指数。本研究采用了前面计算的两种光谱多样性指数、Sentinel-2A影像的12个波段信息和30个与森林多样性相关的植被指数作为特征波段, 相关植被指数的计算公式如表2所示。计算样地内的光谱变异系数, 以作为森林植被多样性的度量(Torresani *et al.*, 2019; Mallinis *et al.*, 2020)。

利用44个变量初步构建随机森林回归模型, 对这些变量进行重要性评估。变量的重要性排序按“%IncMSE”的大小表示, “%IncMSE”即increase in mean squared error, 是基于每一棵决策树的袋外数

据误差,通过对每一个预测变量随机赋值,如果袋外数据误差增幅越大,其值被随机替换后模型估算的误差就越大,说明该变量的重要性越高。但某些变量可能对多样性的估算作用很小,或者变量间存

在共线性,从而影响回归结果。为进一步提高随机森林的估算精度,减少噪声和冗余信息的干扰,本研究基于十折交叉验证对变量进行筛选,选取重要性高且使交叉验证误差最小的特征变量作为模型的

表2 植被指数计算公式
Table 2 Formula of calculating vegetation index

植被指数 Vegetation index	计算公式 Calculate formula	参考文献 Reference
TCARI	$3[(R_{699.19} - R_{668.98}) - 0.2(R_{699.19} - R_{550.67})(R_{699.19}/R_{668.98})]$	Kim <i>et al.</i> , 1994
OSAVI	$(1 + 0.16)(R_{750} - R_{705})/(R_{750} + R_{705} + 0.16)$	Wu <i>et al.</i> , 2008
OSAVI2	$(1 + 0.16)(R_{800} - R_{670})/(R_{800} + R_{670} + 0.16)$	Rondeaux <i>et al.</i> , 1996
DATT	$(R_{850} - R_{710})/(R_{850} - R_{680})$	Datt, 1999
DATT2	R_{850}/R_{710}	Datt, 1999
Gitelson	$1/R_{700}$	Gitelson <i>et al.</i> , 1999
SR1	R_{750}/R_{700}	Gitelson & Merzlyak, 1997
SR2	R_{700}/R_{670}	McMurtrey III <i>et al.</i> , 1994
SR3	R_{730}/R_{706}	Zarco-Tejada <i>et al.</i> , 2003
SR4	R_{675}/R_{700}	Gitelson <i>et al.</i> , 2003
MSI	R_{1600}/R_{819}	Hunt & Rock, 1989
NDII	$(R_{819} - R_{1649})/(R_{819} + R_{1649})$	Hardisky <i>et al.</i> , 1983
CRI1	$1/R_{510} - 1/R_{550}$	Gitelson <i>et al.</i> , 2002
CRI2	$1/R_{510} - 1/R_{700}$	Gitelson <i>et al.</i> , 2002
ARI	$1/R_{550} - 1/R_{700}$	Sims & Gamon, 2002
PSRI	$(R_{680} - R_{500})/R_{750}$	Merzlyak <i>et al.</i> , 1999
NDVI	$(R_{750.66} - R_{704.6})/(R_{750.66} + R_{704.6})$	Gitelson & Merzlyak, 1994
GNDVI	$(R_{783} - R_{560})/(R_{783} + R_{560})$	Rozenstein <i>et al.</i> , 2019
TNDVI	$((R_{842} - R_{665})/(R_{842} + R_{665}) + 0.5)^{0.5}$	Rozenstein <i>et al.</i> , 2019
WDVI	$R_{842} - R_{665} \times 0.5$	Rozenstein <i>et al.</i> , 2019
NDI45	$(R_{705} - R_{665})/(R_{705} + R_{665})$	Delegido <i>et al.</i> , 2011
SAVI	$(1 + L) \times (R_{799.09} - R_{680.045})/(R_{799.09} + R_{680.045} + L) (L = 0.5)$	Huete, 1988
SAVI2	$R_{799.09}/(R_{680.045} + b/a) (a = 0.969, b = 0.084, 726)$	Major <i>et al.</i> , 1990
ARVI	$RB = R_{680.045} - r(R_{444.5} - R_{680.045}) (r = 1)$ $ARVI = (R_{799.09} - RB)/(R_{799.09} + RB)$	Kaufman & Tanre, 1992
SARVI	$RB = R_{680.045} - r(R_{444.5} - R_{680.045}) (r = 1, L = 0.5)$ $SARVI = (1 + L)(R_{799.09} - RB)/(R_{799.09} + RB + L)$	Kaufman & Tanre, 1992
EVI	$G(R_{799.09} - R_{680.045})/(R_{799.09} + C_1R_{680.045} - C_2R_{444.5} + L)$ $(G = 2.5, C_1 = 6, C_2 = 7.5, L = 1)$	Huete <i>et al.</i> , 1997
IRECI	$(R_{783} - R_{665})/(R_{705}/R_{740})$	Frampton <i>et al.</i> , 2013
IPVI	$R_{842}/(R_{842} + R_{665})$	Rozenstein <i>et al.</i> , 2019
PSSRA	R_{783}/R_{665}	Rozenstein <i>et al.</i> , 2019
RVI	R_{842}/R_{665}	Rozenstein <i>et al.</i> , 2019

ARI, 花青素反射指数; ARVI, 耐大气植被指数; CRI, 类胡萝卜素反射指数; DATT, DATT植被指数; EVI, 增强型植被指数; Gitelson, Gitelson植被指数; GNDVI,绿色归一化差异植被指数; IPVI, 红外植被百分比指数; IRECI, 倒红边叶绿素指数; MSI, 水分胁迫指数; NDI45, 归一化差异指数; NDII, 归一化红外指数; NDVI, 归一化植被指数; OSAVI, 优化型土壤调节植被指数; PSRI, 植物衰老反射指数; PSSRA, 特征色素简单比值指数; RVI, 比值植被指数; SARVI, 土壤大气阻抗植被指数; SAVI, 土壤调节植被指数; SR, 比值植被指数; TCARI, 转换型叶绿素吸收植被指数; TNDVI, 转化后的归一化植被指数; WDVI, 加权差分植被指数。公式中的R及右下角数字代表位于该波长处的反射值。
ARI, anthocyanin reflectance index; ARVI, atmospherically resistant vegetation index; CRI, carotenoid reflectance index; DATT, DATT vegetation index; EVI, enhanced vegetation index; Gitelson, Gitelson vegetation index; GNDVI, green normalizad difference vegetation index; IRECI, inverted red-edge chlorophyll index; IPVI, infrared percentage vegetation index; MSI, moisture stress index; NDI45, normalized difference index 45; NDII, normalized difference infrared index; NDVI, vormalizad difference vegetation index; OSAVI, optimization soil-adjusted vegetation index; PSRI, plant senescence reflectance index; PSSRa, pigment specific simple ratio; RVI, ratio vegetation index; SARVI, soil atmospherically resistant vegetation index; SAVI, soil adjusted vegetation index; SR, simple ratio index; TCARI, transformed chlorophyll-absorbing vegetation index; TNDVI, transformed normalized difference vegetation index; WDVI, weighted difference vegetation index. *R* and the lower right number in the formula represent the reflection value at this wavelength.

输入变量(Genuer *et al.*, 2010; Edwards *et al.*, 2018)。基于所选择的变量构建最终的随机森林回归模型, 使用留一法验证模型精度。输入变量选择在R语言的“randomforest”包中实现, 留一法交叉验证在python中的sklearn库中实现。

2 结果

2.1 基于光谱多样性估算森林植物多样性

图3展示了基于全波段影像像元所计算的 CV 和 CHA 与实测的Shannon-Wiener多样性指数、Simpson多样性指数和物种丰富度的关系, 两种光谱多样性指数均能很好地估算Shannon-Wiener多样性指数, CHA 对Shannon-Wiener多样性指数的估算精度($R^2 =$

0.74, $RMSE = 0.07$)高于 CV 对其的估算精度($R^2 = 0.60$, $RMSE = 0.11$)。但两种光谱多样性指数不能估算Simpson多样性指数和物种丰富度, CV 和 CHA 与实测的Simpson多样性指数的估算精度均为0.12, 与实测的物种丰富度的精度分别为0.11和0.05, 结果均不显著。

与基于像元的光谱多样性相比, 基于聚类的森林植物多样性估算结果略有降低(图4)。基于聚类分析可以较好地估算Shannon-Wiener指数($R^2 = 0.59$, $RMSE = 0.12$, $p < 0.01$), 但Simpson多样性指数与物种丰富度估算精度较差, 结果均不显著(Simpson多样性指数: $R^2 = 0.02$, $RMSE = 0.08$; 物种丰富度: $R^2 = 0.12$, $RMSE = 7.94$)。

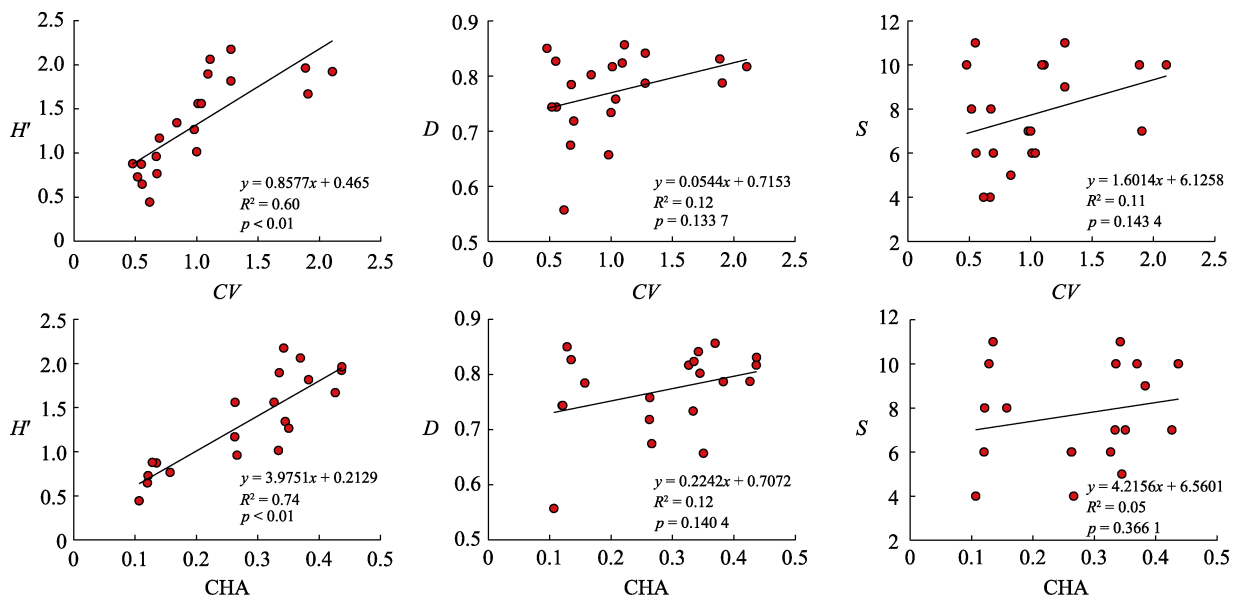


图3 基于原始波段的变异系数(CV)与凸包面积(CHA)与实测植物多样性(Shannon-Wiener多样性指数(H')、Simpson多样性指数(D)和物种丰富度(S))的关系。

Fig. 3 Relationship between coefficient of variation (CV) and convex hull area (CHA) based on original bands and measured plant diversity (Shannon-Wiener diversity index (H'), Simpson diversity index (D) and species richness (S)).

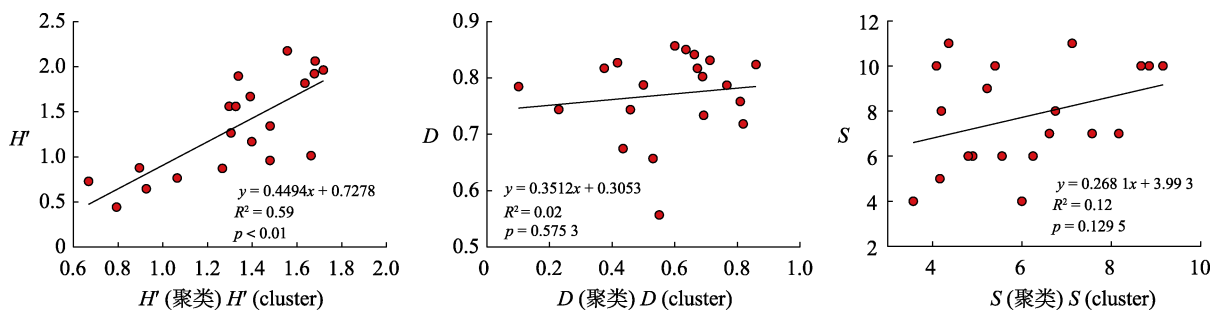


图4 基于聚类的植物多样性指数(Shannon-Wiener多样性指数(H')、Simpson多样性指数(D)和波谱物种数(S))与实测值的关系。

Fig. 4 Relationship between plant diversity index based on clustering (Shannon-Wiener diversity index (H'), Simpson diversity index (D) and the number of spectral species (S)) and measured value.

2.2 基于随机森林回归估算森林植物多样性

基于已有的44个特征变量进行随机森林回归, 基于%IncMSE对输入变量的重要性程度进行排序, 重要性排序前30的变量如图5所示, 可以发现CV与CHA的重要性程度远高于其余变量。通过十折交叉验证对输入变量进行筛选, 验证误差随着特征变量数量的增加先下降再增加, 根据验证结果可知当保留6个变量时可以获得最理想的回归结果。用筛选出的特征变量构建最终的随机森林回归模型, 使用留一法对模型精度进行检验, 结果发现对Shannon-Wiener多样性指数的解释度最高($R^2 = 0.79$, $RMSE = 0.06$, $p < 0.01$), 对物种丰富度($R^2 = 0.23$, $RMSE = 3.75$)和Simpson多样性指数($R^2 = 0.10$, $RMSE = 0.005$)的解释度较低(图6)。

2.3 森林植物多样性成图

通过以上结果可以发现, 两种方法均能较好地估算Shannon-Wiener多样性指数, 精度为0.59–0.79。在分析比较之后, 采用随机森林回归的方法对研究区域的森林植物多样性进行区域成图(图7)。凉水自

然保护区中间的核心区部分的主要林分类型为原始状态的阔叶红松林, 核心区的Shannon-Wiener多样性指数明显高于保护区南边的实验区以及保护区周围的缓冲区部分, 这可能是因为核心区多原始林, 人为干扰较少, 保护较好, 而保护区四周的缓冲区部分大多以次生林为主。丰林保护区的核心区和与其相连的缓冲区东部的Shannon-Wiener多样性值较高, 实验区的Shannon-Wiener多样性指数值较低, 在农田和建筑用地周围的多样性值最低。珲春自然保护区中部的实验区和南部的缓冲区部分有较多的农田、湖泊分布, 人类活动的影响使得其周围的Shannon-Wiener多样性指数值偏低, 北部的核心区和缓冲区部分多样性较高。整体来看, 丰林和凉水保护区的Shannon-Wiener多样性指数略高于珲春自然保护区, 前两者的森林类型主要是以红松为主的原始林和针阔混交林, 珲春保护区以阔叶林为主, 北部有少量针叶林和针阔混交林。

3 讨论

3.1 光谱多样性估算森林植物多样性

3.1.1 像元尺度

本研究基于像元光谱反射率计算了变异系数和凸包面积两种光谱多样性指数, 结果发现两种指数均能很好地估算Shannon-Wiener多样性指数($p < 0.01$)。这一结果与之前的研究(Lucas & Carter, 2008; Somers *et al.*, 2015; Gholizadeh *et al.*, 2018)一致, 表明CV和CHA是反映 α 多样性的有用指标。基于波段选择的凸包面积最早应用于草地生物多样性监测(Gholizadeh *et al.*, 2018), 但在森林中还没有进一步的应用。此外, 本研究发现基于影像像元计算的CHA对实测的Shannon-Wiener多样性指数估算精度最高, 进一步证实了Gholizadeh的研究结论, 该研究发现随着分辨率降低, CHA优于包括CV在内的所有其他指标。一般来说, 空间分辨率的降低会产生混合像元, 进一步导致光谱异质性(Rocchini, 2007; Xi *et al.*, 2019)。在所用影像都为中分辨率的情况下, CV的计算是用每个波段所有像元的变异系数之和除以波段数, 而CHA是先求出所有波段的变异再除以像元数。前者是整合所有波段值的空间变异, 可能反映所有物种或群落波谱变异的平均值, 而后者是每个波段空间变异的平均值, 可能反映所有物种或群落波谱变异的异质性, 更能体现物种或群

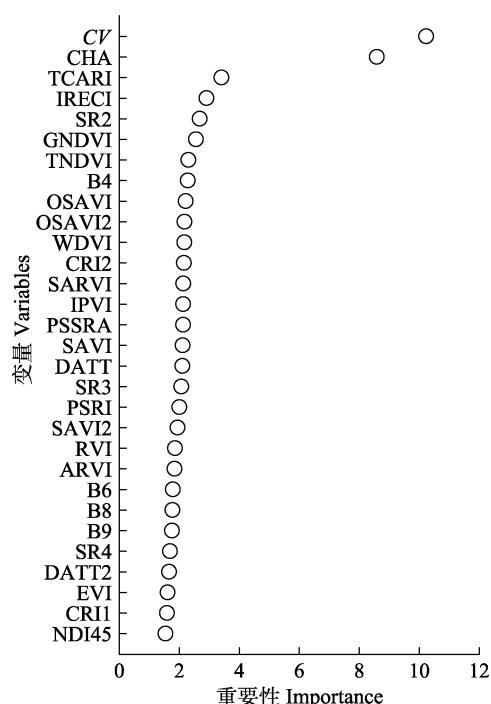


图5 随机森林回归中重要性排序前30的变量。CV, 变异系数; CHA, 凸包面积; B4、B6、B8、B9分别表示Sentinel-2A的第4、6、8、9波段; 其余变量为植被指数, 具体含义见表2。
Fig. 5 Top 30 variables of importance in random forest regression. CV, coefficient of variation; CHA, convex hull area; B4, B6, B8 and B9 represent the fourth, sixth, eighth and ninth bands of Sentinel-2A, respectively; the other variables are vegetation indexes, with specific meanings shown in Table 2.

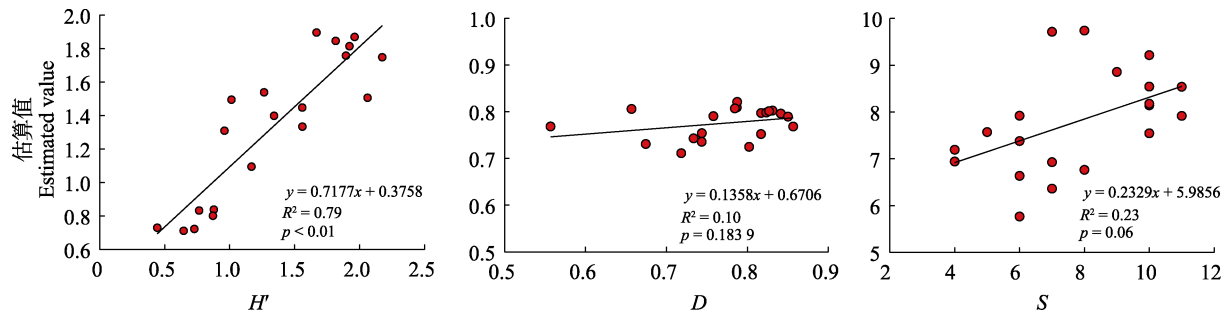


图6 基于随机森林回归估算植物多样性指数(Shannon-Wiener多样性指数(H')、Simpson多样性指数(D)和物种丰富度(S))。

Fig. 6 Estimation of plant diversity index (Shannon-Wiener diversity index (H'), Simpson diversity index (D) and species richness(S)) based on random forest regression.

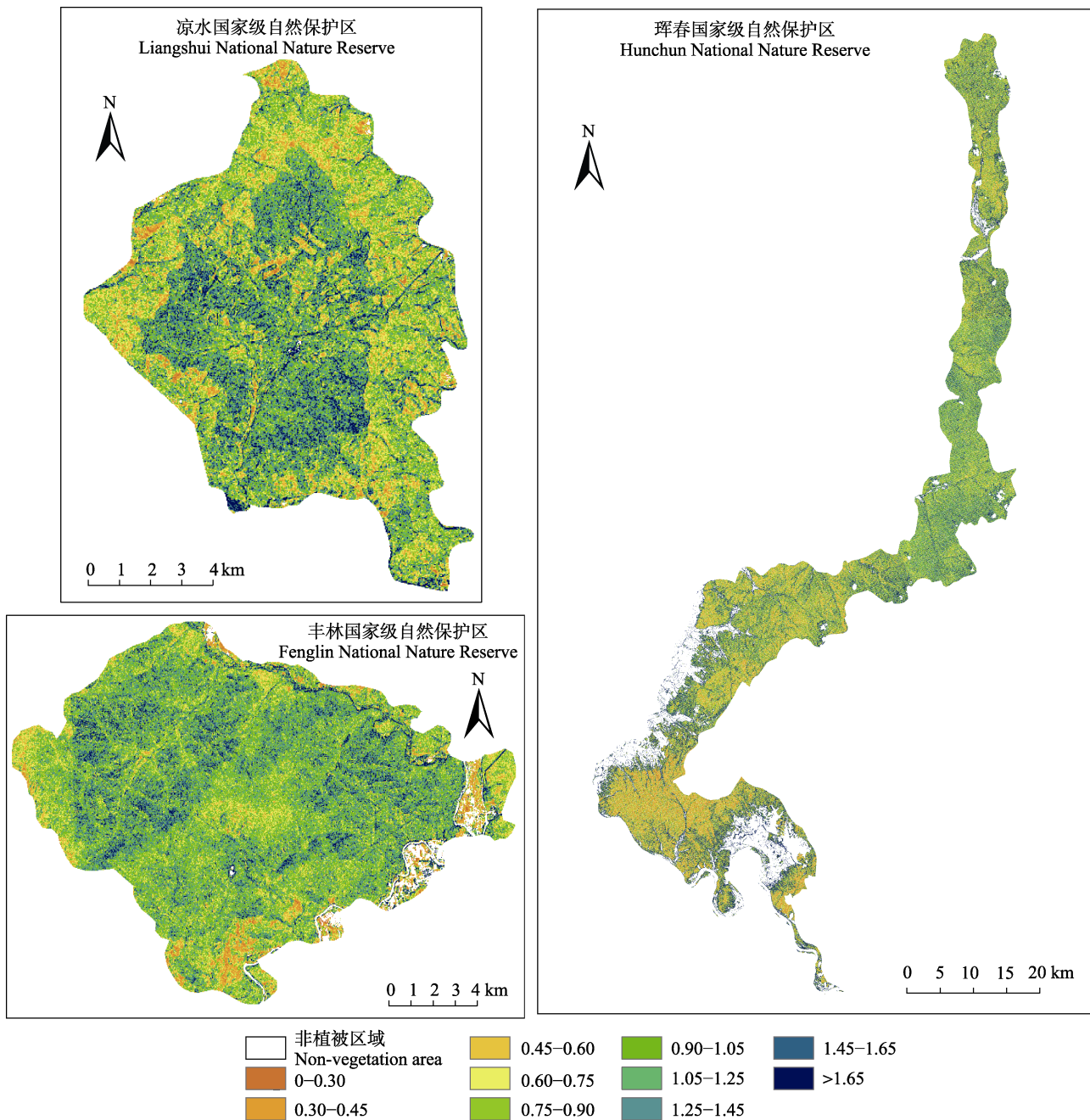


图7 三个国家级自然保护区Shannon-Wiener多样性指数空间分布图。

Fig. 7 Spatial distribution of Shannon-Wiener diversity index in three national nature reserves.

落间的多样性。

但是, 研究发现两种光谱多样性指数均不能估算Simpson多样性指数和物种丰富度。这主要是由于遥感估算多样性是基于面上的信息, 物种所占的面积比例对影像光谱的影响更大, 而非物种的个数。Shannon-Wiener多样性指数考虑了物种的数量和每个物种的相对丰度用以衡量异质性(Pielou, 1966), 所以估算精度较高, 而Simpson多样性指数则是反映群落中少数物种的优势度(Fauvel *et al.*, 2020)。此外, 多样性指数的估算比物种丰富度指数的估算更准确, 可能是因为丰富度仅基于物种的存在或缺失, 稀有物种的存在会影响物种丰富度, 但在影像上可能难以监测。在森林中稀有物种的冠幅占比较小, Sentinel-2A卫星的分辨率有限, 传感器很难检测到细微的单个物种变化, 而更多地指示像元内优势物种变异, 因此Shannon-Wiener多样性指数更适合区域尺度森林植物多样性的遥感估算(Oldeland *et al.*, 2010)。

3.1.2 聚类尺度

聚类分析是一种根据数据相似性划分分组的无监督分类方法, 这种方法与训练数据收集和数据分布无关, 可以直接映射植物多样性而并不单独辨别树种(Féret & Asner, 2014; Féret & de Boissieu, 2020)。聚类的方法很多, 本研究用到的是K均值聚类, 在尝试了 3×3 、 5×5 、 7×7 和 9×9 的聚类窗口后, 最终选定了聚类效果最好的 5×5 的聚类窗口。原因主要是受影像分辨率的限制, 更小的窗口导致聚类效果受影响, 无法聚类, 扩大窗口范围则不能与实地野外调查的样地大小相匹配, 估算结果误差较大。当所选的移动窗口为 5×5 时, 根据聚类结果计算得到的Shannon-Wiener多样性指数的估算结果最好($R^2 = 0.58$, $p < 0.01$), 7×7 的移动窗口估算结果次之($R^2 = 0.33$, $p < 0.01$), 9×9 的移动窗口效果最差($R^2 = 0.26$, $p < 0.05$)。通过聚类的方法获得的其他指数与实测值均无显著关系, 可能是因为通过聚类得到的物种数是基于光谱的波谱物种数, 并不是实际的物种数, 以此计算的结果与实测值有一定差异。

基于聚类方法的估算结果略低于基于像元光谱多样性方法, 这可能是受影像分辨率的限制。影像的分辨率对估算生物多样性有直接的影响, 分类的精度主要依赖于种间的光谱变异性(Zhang *et al.*, 2006)。现有学者利用无人机和激光雷达数据, 将分

辨率提高到单木水平, 通过单木分割并结合了植物叶片的生化组分以及结构参数等特征, 大大提高了分类精度(董文雪等, 2018; Zhao *et al.*, 2018; 衣海燕等, 2020)。在预测试的时候本研究也尝试其他聚类方法, 自定义模糊C均值聚类可以根据聚类有效性函数自动获取聚类数目, 解决了在物种数量未知时的初始值选择的问题(Zhao *et al.*, 2018; 衣海燕等, 2020), 可以较好地解决遥感信息的不确定性及多解性。但与前人的研究结果相比, 本研究使用该方法得到的结果精度并不高, 这也主要是受空间分辨率的限制, 不能从单木尺度进行聚类。

3.2 随机森林回归估算森林植物多样性

除了基于光谱多样性方法, 本研究也采用了机器学习的方法对森林植物多样性进行了估算, 结果发现基于所选特征指数的随机森林回归对Shannon-Wiener多样性指数有较好的反演结果($R^2 = 0.79$, $p < 0.01$), 而对其他多样性指数的估算效果较低。随机森林估算结果的准确性主要受两个方面因素的影响。一方面是受影像分辨率的影响, 影像分辨率不同对研究结果会产生一定影响。在草地生态系统中的大规模模拟研究表明, α 多样性的估算精度随着空间分辨率的降低而下降, 最佳的估算精度是像元大小与单个植株大小相接近时的精度(Wang *et al.*, 2018)。这是因为遥感反映的是面上的信息, 利用遥感数据计算的不是物种间的光谱变异, 而是用像元代表的群落间的光谱变异间接指示物种间的变异, 其中存在一定的误差, 而将分辨率提升到单木尺度会提高估算精度。另一方面受输入特征变量的影响, 本研究在44个输入变量中通过十折交叉验证对变量进行了筛选, 最后挑选了使估算结果最好的6个特征变量, 去除了变量间共线性和冗余信息的干扰。在变量挑选时发现输入的两种光谱多样性指数的重要性排序明显高于其余的特征变量, 说明这两种指数可以作为随机森林估算植物多样性的重要指标, 间接证明了前面直接利用两种指标进行植物多样性估算的可行性和合理性。但是, 重要性排序中的CV的重要性程度略高于CHA, 这可能是因为CHA与其他植被指数间存在一定相关性, 会对重要性结果产生一定影响。此外, 在森林生态系统类型中, 森林的空间分布和结构参数也会影响估算结果(Marceau *et al.*, 1994; Gholizadeh *et al.*, 2019), 即使采用相同分辨率的影像, 在不同地区或不同森林类型中结果可

能存在一定差异(Mallinis *et al.*, 2020; Gyamfi-Ampadu *et al.*, 2021)。

3.3 方法的局限性

本研究目前只是在温带森林中开展, 因为气候地理条件差异, 如果要将此方法应用在不同地区的不同类型的森林中, 还需要进行进一步的验证和评估。国外相关学者基于Sentinel-2A影像, 以地中海地区的落叶阔叶林和针阔混交林为研究对象, 利用随机森林回归方法对森林植物多样性进行评估, 结果发现估算的Shannon-Wiener多样性指数精度在0.29–0.31之间, 而Simpson多样性指数的估算精度可达0.31–0.37 (Chrysafis *et al.*, 2020; Mallinis *et al.*, 2020)。本研究的结果与这些结论略有差异。本研究除了基于波段本身, 还加入了与植物多样性相关的指数作为随机森林的输入变量, 据此估算的Shannon-Wiener多样性指数精度较高, 但未发现光谱多样性与Simpson多样性指数有显著关系($R^2 = 0.1$), 这可能是受限于东北森林类型及其物种分布格局。今后的研究在应用到其他区域时, 可以在更大尺度上利用更多样方数据来进一步验证。

此外, 在所测试的两种方法中, 随机森林回归反演植物多样性的方法精度略高, 但该方法不能对超越训练数据集的范围进行估算, 而且它是一种黑箱的方法, 其内部的运行也无法控制。此外, 由于森林的复杂性和影像分辨率的限制, 目前的研究只是基于影像的光谱特征反演多样性, 而忽略了其他因素如叶片化学成分、植被结构参数、林下灌木或草本光谱的影响以及土壤反射率等的影响(Wang & Gamon, 2019), 方法上还有一定的局限性。如果要在全国范围内推广, 在后续研究中还需要考虑环境数据, 如地形、海拔、降水等环境因素的影响, 以进一步验证和完善方法, 为开展全国的森林健康评估提供较为全面可行的方法与参考。

除了考虑与植物多样性的关系, 植物功能多样性也是生物多样性的重要组成部分, 它包括植物的化学、生理和形态结构的多样性, 反映了群落、景观、甚至大空间尺度内功能性状的变异性, 决定了生态系统的功能和稳定性(Tilman *et al.*, 1997; Ruiz-Benito *et al.*, 2014)。利用卫星遥感影像和野外实测功能性状的关系建模, 反演植物功能多样性对实现全球生物多样性快速监测具有重要意义(Jetz *et al.*, 2016)。但由于计算植物功能多样性的可用区

域很少且不可重复测量, 传统的计算植物功能多样性的方法不适合用于长期和大规模的监测。遥感反演功能多样性的原理是基于系统发育差异和资源限制导致的植物性状不同(Schweiger *et al.*, 2018), 这些性状主要包括叶片氮浓度、叶片碳浓度、比叶面积、叶片干物质含量和叶面积等叶片性状, 以及株高、树冠横截面积和胸径等全株性状(Ma *et al.*, 2019), 性状上的差异进而影响光谱反射率(Wang *et al.*, 2018)。已经有很多研究基于机载高光谱数据绘制高空间分辨率的植物功能多样性空间分布图(Asner *et al.*, 2017; Schneider *et al.*, 2017; Zheng *et al.*, 2021), 但在更大的区域范围内由于获取数据难以实现且成本昂贵, 研究还较少。Sentinel-2号卫星等多光谱空间遥感的发展有利于克服高空间分辨率和监测范围的问题(Rossi *et al.*, 2020; Hauser *et al.*, 2021)。目前相关的研究还较少且受很多因素影响, 比如野外数据和卫星影响时间不匹配, 物种的区域差异与方法的适用性等等, 这都有待进一步补充和验证, 可以考虑作为接下来的研究方向。

4 结论

本研究依托我国东北的3个森林生态系统国家级保护区, 基于遥感数据和野外调查数据, 探讨了基于像元和聚类分析的光谱多样性直接估算和基于随机森林回归的森林植物多样性反演。结果发现两种方法均能较好地估算研究区内Shannon-Wiener指数: (1)在像元尺度, 基于凸包面积计算的光谱多样性指数估算森林植物多样性精度($R^2 = 0.74$)优于基于变异系数方法($R^2 = 0.60$); (2)基于像元的光谱多样性估算方法优于聚类分析方法($R^2 = 0.59$); (3)利用随机森林回归算法估算精度($R^2 = 0.79$)最高。Sentinel-2号卫星影像是目前可以免费获取的最高分辨率的影像, 本研究为在大尺度上进行森林多样性估算提供了可行的方法和依据, 这对保护区森林管理和自然资源保护有重要意义。

参考文献

- Asner GP, Martin RE (2009). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment*, 7, 269–276.
- Asner GP, Martin RE, Knapp DE, Tupayachi R, Anderson CB, Sinca F, Vaughn NR, Llactayo W (2017). Airborne

- laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science*, 355, 385-389.
- Breiman L (2001). Random forests. *Machine Learning*, 45, 5-32.
- Carlson KM, Asner GP, Hughes RF, Ostertag R, Martin RE (2007). Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. *Ecosystems*, 10, 536-549.
- Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM (2015). Accelerated modern human-induced species losses: entering the sixth mass extinction. *Science Advances*, 1, e1400253. DOI: 10.1126/sciadv.1400253.
- Chang CI (2000). An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory*, 46, 1927-1932.
- Chrysafis I, Korakis G, Kyriazopoulos AP, Mallinis G (2020). Predicting tree species diversity using geodiversity and Sentinel-2 multi-seasonal spectral information. *Sustainability*, 12, 9250. DOI: 10.3390/su12219250.
- Clark ML, Roberts DA (2012). Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. *Remote Sensing*, 4, 1820-1855.
- Cross M, Scambos T, Pacifici F, Vargas-Ramirez O, Moreno-Sanchez R, Marshall W (2019). Classification of tropical forest tree species using meter-scale image data. *Remote Sensing*, 11, 1411. DOI: 10.3390/rs11121411.
- Dahlin KM (2016). Spectral diversity area relationships for assessing biodiversity in a wildland-agriculture matrix. *Ecological Applications*, 26, 2756-2766.
- Datt B (1999). Visible/near infrared reflectance and chlorophyll content in *Eucalyptus* leaves. *International Journal of Remote Sensing*, 20, 2741-2759.
- Delegido J, Verrelst J, Alonso L, Moreno J (2011). Evaluation of Sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors*, 11, 7063-7081.
- Deng S, Katoh M, Yu X, Hyypä J, Gao T (2016). Comparison of tree species classifications at the individual tree level by combining ALS data and RGB images using different algorithms. *Remote Sensing*, 8, 1034. DOI: 10.3390/rs8121034.
- Dong WX, Zeng Y, Zhao YJ, Zhao D, Zheng ZJ, Yi HY (2018). Forest species diversity mapping using airborne LiDAR and hyperspectral data. *Journal of Remote Sensing*, 22, 833-847. [董文雪, 曾源, 赵玉金, 赵旦, 郑朝菊, 衣海燕 (2018). 机载激光雷达及高光谱的森林乔木物种多样性遥感监测. *遥感学报*, 22, 833-847.]
- Duro DC, Coops NC, Wulder MA, Han T (2007). Development of a large area biodiversity monitoring system driven by remote sensing. *Progress in Physical Geography: Earth and Environment*, 31, 235-260.
- Edwards JA, Santos-Medellin CM, Liechty ZS, Nguyen B, Lurie E, Eason S, Phillips G, Sundareshan V (2018). Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLOS Biology*, 16, e2003862. DOI: 10.1371/journal.pbio.2003862.
- Erinjeri JJ, Singh M, Kent R (2018). Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery. *Remote Sensing of Environment*, 216, 345-354.
- Fairbanks DHK, McGwire KC (2004). Patterns of floristic richness in vegetation communities of California: regional scale analysis with multi-temporal NDVI. *Global Ecology and Biogeography*, 13, 221-235.
- Fauvel M, Lopes M, Dubo T, Rivers-Moore J, Frison PL, Gross N, Ouin A (2020). Prediction of plant diversity in grasslands using Sentinel-1 and -2 satellite image time series. *Remote Sensing of Environment*, 237, 111536. DOI: 10.1016/j.rse.2019.111536.
- Feeley KJ, Gillespie TW, Terborgh JW (2005). The utility of spectral indices from Landsat ETM+ for measuring the structure and composition of tropical dry forests. *Biotropica*, 37, 508-519.
- Féret JB, Asner GP (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications*, 24, 1289-1296.
- Féret JB, de Boissieu F (2020). biodivMapR: an R package for α - and β -diversity mapping using remotely sensed images. *Methods in Ecology and Evolution*, 11, 64-70.
- Frampton WJ, Dash J, Watmough G, Milton EJ (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 83-92.
- Genuer R, Poggi JM, Tuleau-Malot C (2010). Variable selection using Random forests. *Pattern Recognition Letters*, 31, 2225-2236.
- Gholizadeh H, Gamon JA, Helzer CJ, Cavender-Bares J (2020). Multi-temporal assessment of grassland α - and β -diversity using hyperspectral imaging. *Ecological Applications*, 30, e02145. DOI: 10.1002/eap.2145.
- Gholizadeh H, Gamon JA, Townsend PA, Zygielbaum AI, Helzer CJ, Hmimina GY, Yu R, Moore RM, Schweiger AK, Cavender-Bares J (2019). Detecting prairie biodiversity with airborne remote sensing. *Remote Sensing of Environment*, 221, 38-49.
- Gholizadeh H, Gamon JA, Zygielbaum AI, Wang R, Schweiger AK, Cavender-Bares J (2018). Remote sensing of biodiversity: soil correction and data dimension reduction methods assessment of α -diversity (species richness) in prairie ecosystems. *Remote Sensing of Environment*, 206, 240-253.
- Gillespie TW, Foody GM, Rocchini D, Giorgi AP, Saatchi S (2008). Measuring and modelling biodiversity from space.

- Progress in Physical Geography: Earth and Environment*, 32, 203-221.
- Gitelson A, Merzlyak MN (1994). Quantitative estimation of chlorophyll-*a* using reflectance spectra: experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology*, 22, 247-252.
- Gitelson AA, Buschmann C, Lichtenthaler HK (1999). The chlorophyll fluorescence ratio F_{735}/F_{700} as an accurate measure of the chlorophyll content in plants. *Remote Sensing of Environment*, 69, 296-302.
- Gitelson AA, Gritz Y, Merzlyak MN (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160, 271-282.
- Gitelson AA, Merzlyak MN (1997). Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, 18, 2691-2697.
- Gitelson AA, Zur Y, Chivkunova OB, Merzlyak MN (2002). Assessing carotenoid content in plant leaves with reflectance spectroscopy. *Photochemistry and Photobiology*, 75, 272-281.
- Gotelli NJ, Colwell RK (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4, 379-391.
- Graham CH, Hijmans RJ (2006). A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography*, 15, 578-587.
- Gyamfi-Ampadu E, Gebreslasie M, Mendoza-Ponce A (2021). Evaluating multi-sensors spectral and spatial resolutions for tree species diversity prediction. *Remote Sensing*, 13, 1033. DOI: 10.3390/rs13051033.
- Hardisky M, Klemas V, Smart RM (1983). The influence of soil-salinity, growth form, and leaf moisture on the spectral radiance of *Spartina alterniflora* canopies. *Photogrammetric Engineering and Remote Sensing*, 48, 77-84.
- Harrison PA, Berry PM, Simpson G, Haslett JR, Blicharska M, Bucur M, Dunford R, Egoh B, Garcia-Llorente M, Geamănă N, Geertsema W, Lommelen E, Meiresonne L, Turkelboom F (2014). Linkages between biodiversity attributes and ecosystem services: a systematic review. *Ecosystem Services*, 9, 191-203.
- Hauser LT, Féret JB, An Binh N, van der Windt N, Sil AF, Timmermans J, Soudzilovskaia NA, van Bodegom PM (2021). Towards scalable estimation of plant functional diversity from Sentinel-2: *in-situ* validation in a heterogeneous (semi-)natural landscape. *Remote Sensing of Environment*, 262, 112505. DOI: 10.1016/j.rse.2021.112505.
- Huete AR (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295-309.
- Huete AR, Liu HQ, Batchily K, Vanleeuwen W (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59, 440-451.
- Hunt Jr ER, Rock BN (1989). Detection of changes in leaf water content using Near- and Middle-Infrared reflectances. *Remote Sensing of Environment*, 30, 43-54.
- Jetz W, Cavender-Bares J, Pavlick R, Schimel D, Davis FW, Asner GP, Guralnick R, Kattge J, Latimer AM, Moorcroft P, Schaepman ME, Schildhauer MP, Schneider FD, Schrod F, Stahl U, Ustin SL (2016). Monitoring plant functional diversity from space. *Nature Plants*, 2, 16024. DOI: 10.1038/NPLANTS.2016.24.
- Kalacska M, Sanchez-Azofeifa GA, Rivard B, Caelli T, White HP, Calvo-Alvarado JC (2007). Ecological fingerprinting of ecosystem succession: estimating secondary tropical dry forest structure and diversity using imaging spectroscopy. *Remote Sensing of Environment*, 108, 82-96.
- Kaufman YJ, Tanre D (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 30, 261-270.
- Ke YH, Quackenbush LJ, Im J (2010). Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment*, 114, 1141-1154.
- Kerr JT, Ostrovsky M (2003). From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution*, 18, 299-305.
- Kim MS, Daughtry CST, Chappelle EW, McMurtrey JE, Walthall CL (1994). The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (a Par)/Proceedings of the 6th International Symposium on Physical Measurements and Signatures in Remote Sensing. Val d'Isère, France. 299-306.
- Laliberté E, Schweiger AK, Legendre P (2020). Partitioning plant spectral diversity into alpha and beta components. *Ecology Letters*, 23, 370-380.
- Levrel H, Fontaine B, Henry PY, Jiguet F, Julliard R, Kerbiriou C, Couvet D (2010). Balancing state and volunteer investment in biodiversity monitoring for the implementation of CBD indicators: a French example. *Ecological Economics*, 69, 1580-1586.
- Lucas KL, Carter GA (2008). The use of hyperspectral remote sensing to assess vascular plant species richness on Horn Island, Mississippi. *Remote Sensing of Environment*, 112, 3908-3915.
- Ma X, Mahecha MD, Migliavacca M, van der Plas F, Benavides R, Ratcliffe S, Kattge J, Richter R, Musavi T, Baeten L, Barneaiea I, Bohn FJ, Bouriaud O, Bussotti F, Coppi A, *et al.* (2019). Inferring plant functional diversity from space: the potential of Sentinel-2. *Remote Sensing of Environment*, 233, 111368. DOI: 10.1016/j.rse.2019.111368.
- Ma Z, Li B, Li WJ, Han NY, Chen JK, Watkinson AR (2009). Conflicts between biodiversity conservation and development

- in a biosphere reserve. *Journal of Applied Ecology*, 46, 527-535.
- Major DJ, Baret F, Guyot G (1990). A ratio vegetation index adjusted for soil brightness. *International Journal of Remote Sensing*, 11, 727-740.
- Mallinis G, Chrysafis I, Korakis G, Pana E, Kyriazopoulos AP (2020). A random forest modelling procedure for a multi-sensor assessment of tree species diversity. *Remote Sensing*, 12, 1210. DOI: 10.3390/rs12071210.
- Marceau DJ, Gratton DJ, Fournier RA, Fortin JP (1994). Remote sensing and the measurement of geographical entities in a forested environment. 2. The optimal spatial resolution. *Remote Sensing of Environment*, 49, 105-117.
- McMurtrey III JE, Chappelle EW, Kim MS, Meisinger JJ, Corp LA (1994). Distinguishing nitrogen fertilization levels in field corn (*Zea mays* L.) with actively induced fluorescence and passive reflectance measurements. *Remote Sensing of Environment*, 47, 36-44.
- Medina O, Manian V, China JD (2013). Biodiversity assessment using hierarchical agglomerative clustering and spectral unmixing over hyperspectral images. *Sensors*, 13, 13949-13959.
- Merzlyak MN, Gitelson AA, Chivkunova OB, Rakitin VY (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*, 106, 135-141.
- Miao X, Heaton JS, Zheng S, Charlet DA, Liu H (2012). Applying tree-based ensemble algorithms to the classification of ecological zones using multi-temporal multi-source remote-sensing data. *International Journal of Remote Sensing*, 33, 1823-1849.
- Nagendra H (2001). Using remote sensing to assess biodiversity. *International Journal of Remote Sensing*, 22, 2377-2400.
- Nagendra H, Rocchini D (2008). High resolution satellite imagery for tropical biodiversity studies: the devil is in the detail. *Biodiversity and Conservation*, 17, 3431-3442.
- Nicholson E, Mace GM, Armsworth PR, Atkinson G, Buckle S, Clements T, Ewers RM, Fa JE, Gardner TA, Gibbons J, Grenyer R, Metcalfe R, Mourato S, Muûls M, Osborn D, *et al.* (2009). Priority research areas for ecosystem services in a changing world. *Journal of Applied Ecology*, 46, 1139-1144.
- Oldeland J, Wesuls D, Rocchini D, Schmidt M, Jürgens N (2010). Does using species abundance data improve estimates of species diversity from remotely sensed spectral heterogeneity? *Ecological Indicators*, 10, 390-396.
- Palmer MW, Earls PG, Hoagland BW, White PS, Wohlgemuth T (2002). Quantitative tools for perfecting species lists. *Environmetrics*, 13, 121-137.
- Pan Y, Birdsey RA, Phillips OL, Jackson RB (2013). The structure, distribution, and biomass of the world's forests. *Annual Review of Ecology, Evolution, and Systematics*, 44, 593-622.
- Pereira HM, Belnap J, Brummitt N, Collen B, Ding H, Gonzalez-Espinosa M, Gregory RD, Honrado J, Jongman RHG, Julliard R, McRae L, Proenca V, Rodrigues P, Opige M, Rodriguez JP, *et al.* (2010). Global biodiversity monitoring. *Frontiers in Ecology and the Environment*, 8, 459-460.
- Pielou EC (1966). The measurement of diversity in different types of biological collections. *Journal of Theoretical Biology*, 13, 131-144.
- Rocchini D (2007). Effects of spatial and spectral resolution in estimating ecosystem α -diversity by satellite imagery. *Remote Sensing of Environment*, 111, 423-434.
- Rocchini D, Balkenhol N, Carter GA, Foody GM, Gillespie TW, He KS, Kark S, Levin N, Lucas K, Luoto M, Nagendra H, Oldeland J, Ricotta C, Southworth J, Neteler M (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. *Ecological Informatics*, 5, 318-329.
- Rondeaux G, Steven M, Baret F (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55, 95-107.
- Rossi C, Kneubüler M, Schütz M, Schaepman ME, Haller RM, Risch AC (2020). From local to regional: functional diversity in differently managed alpine grasslands. *Remote Sensing of Environment*, 236, 111415. DOI: 10.1016/j.rse.2019.111415.
- Rozenstein O, Haymann N, Kaplan G, Tanny J (2019). Validation of the cotton crop coefficient estimation model based on Sentinel-2 imagery and eddy covariance measurements. *Agricultural Water Management*, 223, 105715. DOI: 10.1016/j.agwat.2019.105715.
- Ruiz-Benito P, Gómez-Aparicio L, Paquette A, Messier C, Kattge J, Zavala MA (2014). Diversity increases carbon storage and tree productivity in Spanish forests. *Global Ecology and Biogeography*, 23, 311-322.
- Schäfer E, Heiskanen J, Heikinheimo V, Pellikka P (2016). Mapping tree species diversity of a tropical montane forest by unsupervised clustering of airborne imaging spectroscopy data. *Ecological Indicators*, 64, 49-58.
- Schneider FD, Morsdorf F, Schmid B, Petchey OL, Hueni A, Schimel DS, Schaepman ME (2017). Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nature Communications*, 8, 1441. DOI: 10.1038/s41467-017-01530-3.
- Schweiger AK, Cavender-Bares J, Townsend PA, Hobbie SE, Madritch MD, Wang R, Tilman D, Gamon JA (2018). Plant spectral diversity integrates functional and phylogenetic components of biodiversity and predicts ecosystem function. *Nature Ecology & Evolution*, 2, 976-982.
- Sims DA, Gamon JA (2002). Relationships between leaf

- pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, 81, 337-354.
- Somers B, Asner GP, Martin RE, Anderson CB, Knapp DE, Wright SJ, van de Kerchove R (2015). Mesoscale assessment of changes in tropical tree species richness across a bioclimatic gradient in Panama using airborne imaging spectroscopy. *Remote Sensing of Environment*, 167, 111-120.
- Tilman D, Knops J, Wedin D, Reich P, Ritchie M, Siemann E (1997). The influence of functional diversity and composition on ecosystem processes. *Science*, 277, 1300-1302.
- Torresani M, Rocchini D, Sonnenschein R, Zebisch M, Marcantonio M, Ricotta C, Tonon G (2019). Estimating tree species diversity from space in an alpine conifer forest: the Rao's Q diversity index meets the spectral variation hypothesis. *Ecological Informatics*, 52, 26-34.
- Turner W (2014). Sensing biodiversity. *Science*, 346, 301-302.
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, 18, 306-314.
- Ustin SL, Gamon JA (2010). Remote sensing of plant functional types. *New Phytologist*, 186, 795-816.
- Wang R, Gamon JA (2019). Remote sensing of terrestrial plant biodiversity. *Remote Sensing of Environment*, 231, 111128. DOI: 10.1016/j.rse.2019.111218.
- Wang R, Gamon JA, Emmerton CA, Li H, Nestola E, Pastorello GZ, Menzer O (2016). Integrated analysis of productivity and biodiversity in a southern Alberta prairie. *Remote Sensing*, 8, 214. DOI: 10.3390/rs8030214.
- Wang R, Gamon JA, Schweiger AK, Cavender-Bares J, Townsend PA, Zygielbaum AI, Kothari S (2018). Influence of species richness, evenness, and composition on optical diversity: a simulation study. *Remote Sensing of Environment*, 211, 218-228.
- Waring RH, Coops NC, Fan W, Nightingale JM (2006). MODIS enhanced vegetation index predicts tree species richness across forested ecoregions in the contiguous USA. *Remote Sensing of Environment*, 103, 218-226.
- Wu CY, Niu Z, Tang Q, Huang WJ (2008). Estimating chlorophyll content from hyperspectral vegetation indices: modeling and validation. *Agricultural and Forest Meteorology*, 148, 1230-1241.
- Xi YB, Ren CY, Wang ZM, Wei SQ, Bai JL, Zhang B, Xiang HX, Chen L (2019). Mapping tree species composition using OHS-1 hyperspectral data and deep learning algorithms in Changbai Mountains, northeast China. *Forests*, 10, 818. DOI: 10.3390/f10090818.
- Xu W, Li X, Pimm SL, Hull V, Zhang J, Zhang L, Xiao Y, Zheng H, Ouyang Z (2016). The effectiveness of the zoning of China's protected areas. *Biological Conservation*, 204, 231-236.
- Yi HY, Zeng Y, Zhao YJ, Zheng ZJ, Xiong J, Zhao D (2020). Forest species diversity mapping based on clustering algorithm. *Chinese Journal of Plant Ecology*, 44, 598-615. [衣海燕, 曾源, 赵玉金, 郑朝菊, 熊杰, 赵旦 (2020). 利用聚类算法监测森林乔木物种多样性. 植物生态学报, 44, 598-615.]
- Zarco-Tejada PJ, Pushnik JC, Dobrowski S, Ustin SL (2003). Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. *Remote Sensing of Environment*, 84, 283-294.
- Zhang J, Rivard B, Sánchez-Azofeifa A, Castro-Esau K (2006). Intra- and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: implications for species identification using HYDICE imagery. *Remote Sensing of Environment*, 105, 129-141.
- Zhao YJ, Zeng Y, Zheng ZJ, Dong WX, Zhao D, Wu BF, Zhao QJ (2018). Forest species diversity mapping using airborne LiDAR and hyperspectral data in a subtropical forest in China. *Remote Sensing of Environment*, 213, 104-114.
- Zheng Z, Zeng Y, Schneider FD, Zhao Y, Zhao D, Schmid D, Schaepman ME, Morsdorf F (2021). Mapping functional diversity using individual tree-based morphological and physiological traits in a subtropical forest. *Remote Sensing of Environment*, 252, 112170. DOI: 10.1016/j.rse.2020.112170.

责任编辑: 苏艳军 编辑: 赵航