

基于传统建模及机器学习的马尾松人工林碳储量模型构建

柴宗政

(贵州大学林学院, 贵阳 550025)

摘 要: 森林是陆地生态系统最大的植被碳库, 碳储量约占陆地生态系统的 80%, 在减缓气候变化和抵减工业温室气体排放中发挥着重要的作用。森林碳储量的精准估算已成为现代林业生产实践和科学研究的热点。对于特定区域和特定树种, 应当建立有效的碳储量模型, 以提高森林碳汇计量或价值核算的精准性。本研究以贵州省浪风关国有林场马尾松人工林为研究对象, 基于 86 块样地调查数据, 采用 3 种传统建模和 8 种机器学习算法构建林分碳储量模型并对比分析确定最优模型, 以期马尾松人工林碳储量精准预测及科学经营提供理论依据。主要结论如下: (1) 选取单分子式 (Mitscherlich)、舒马赫 (Schumacher)、理查德 (Richards)、坎派兹 (Gompertz)、韦布尔 (Weibull)、逻辑斯蒂 (Logistic)、考尔夫 (Korf) 7 个常见的理论生长方程对马尾松人工林优势高、平均高、平均胸径、林分密度指数、林分断面积、林分蓄积、林分碳储量等林分因子进行基础生长模型拟合, 其中, 单分子式对建模和验证数据均表现较好且稳定性强, 确定为最优基础生长模型并以此模型编制了研究区马尾松人工林生长过程表。(2) 在建立的单分子式碳储量模型 ($R^2=0.395$) 的基础上, 引入地位指数与林分密度指数构建了碳储量广义模型 ($R^2=0.922$), 有效提升了模型精度; 此外, 使用联立方程组模型将平均高、林分断面积生长模型和林分碳储量模型 3 个方程进行联合估计, 建立的林分碳储量模型系精度更高 ($R^2=0.956$)。表明林分碳储量的生长与林龄 (A)、林分密度指数 (SDI)、立地指数 (SI)、林分断面积 (Gha)、平均高 (Ph) 等林分因子具有密切联系, 林分碳储量随林分断面积和平均高的增加而增加。采用联立方程组方法是建立碳储量模型的有效方法, 可对研究区马尾松人工林碳储量动态进行有效预测。(3) 采用多元自适应回归样条、人工神经网络、支持向量机、随机森林、增强回归树、Bagging 回归、回归树、k-最近邻 8 种机器学习算法, 与不同自变量组合的 32 个模型比较来看, 自变量组合为 A-SI-SDI-Ph-Gha 的模型拟合效果明显优于其他模型。8 种模型的综合表现排序为: 多元自适应回归样条>人工神经网络>支持向量机>随机森林>增强回归树> Bagging 回归>回归树>k-最近邻。其中, 多元自适应回归样条、人工神经网络、支持向量机、随机森林 4 种机器学习算法对马尾松人工林碳储量模型的构建表现较佳, 无论建模数据和验证数据其 R^2 均在 0.920 以上。此外, 利用随机森林算法计算 5 个自变量对碳储量模型构建的重要性排序: Gha (27.467%)>SDI (22.142%)>Ph (20.204%)>SI (15.587%)>A (14.600%), 表明林分断面积与林分碳储量的关系更为密切。(4) 比较传统建模与机器学习算法的建模效果, 多元自适应回归样条>人工神经网络>支持向量机>随机森林>联立方程组模型>广义生长模型>增强回归树> Bagging 回归>回归树> k-最近邻>基础生长模型。除 k-最近邻模型, 其他机器学习模型在马尾松人工林碳储量模型构建中表现出了很高的预测准确性。其中, 多元自适应回归样条、人工神经网络、支持向量机、随机森林 4 种机器学习算法优于传统建模方法, 表明机器学习算法在构建马尾松人工林模型方面具有较大优势。综上, 马尾松人工林碳储量生长与林龄、林分密度指数、立地指数、林分断面积、平均高等林分因子具有密切联系。传统建模中联立方程组方法是建立碳储量模型的有效方法; 机器学习算法中多元自适应回归样条、人工神经网络、支持向量机、随机森林 4 种方法是建立碳储量模型的有效方法。总体来看, 机器学习算法在构建研究区马尾松人工林碳储量模型中比传统建模方法精度更高; 本研究建立的碳储量模型可以对研究区马尾松人工林碳储量动态进行有效预测, 为掌握林分碳储量的生长过程和森林碳汇评估提供了有效工具。