

引用格式:郭云开,张晓炯,许敏,等.路域植被等效水厚度估算模型研究[J].地球信息科学学报,2020,22(2):308-315. [Guo Y K, Zhang X J, Xu M, et al. Estimation model of equivalent water thickness in the road area[J]. Journal of Geo-information Science, 2020,22(2):308-315.] DOI:10.12082/dqxxkx.2020.190254

路域植被等效水厚度估算模型研究

郭云开^{1,2}, 张晓炯^{1,2*}, 许敏^{1,2}, 刘雨玲¹, 钱佳¹, 章琼¹

1. 长沙理工大学交通运输工程学院, 长沙 410014; 2. 长沙理工大学测绘遥感应用技术研究所, 长沙 410076

Estimation Model of Equivalent Water Thickness in the Road Area

GUO Yunkai^{1,2}, ZHANG Xiaojiong^{1,2*}, XU Min^{1,2}, LIU Yuling¹, QIAN Jia¹, ZHANG Qiong¹

1. School of Traffic and Transportation Engineering, Changsha University of Science & Technology, Changsha 410014, China;

2. Institute of Surveying and Mapping Remote Sensing Application Technology, Changsha University of Science & Technology, Changsha 410076, China

Abstract: Vegetation water content is an important evaluation index for vegetation health monitoring. The Equivalent Water Thickness (EWT) of vegetation is of great significance for the monitoring and evaluation of the ecological conditions in the road area, it could provide a guideline in road area environment management. Taking the Litan highway in Hunan Provinces as an example, this research used field data of canopy reflectance and equivalent water thickness of vegetation on the ground, and simulated reflectance and simulated equivalent water thickness established by PRO4SAIL. In total, 12 kinds of water indices were established by using the simulated reflectance of the PRO4SAIL canopy model and the ground measured reflectance. The random forest algorithm (RF) was introduced to analyze the importance of the 12 water indices and equivalent water thickness. We determined the ordination between water indices and equivalent water thickness as well as the optimal number of input water index in the equivalent water thickness estimation model by using the adjusted coefficient of determination. Based on the selected water index, the water index and equivalent water thickness were calculated by the PRO4SAIL simulation reflectance as the training set. Three equivalent water thickness estimation models were constructed: Random Forest Coupled Partial Least Squares (RF-PLS), Random Forest Coupled Support Vector Machine (RF-SVM) model, and Random Forest coupled Genetic Algorithm to optimize the Support Vector Machine (RF-GA-SVM) model. The applicability of 12 water indices in the estimation of equivalent water thickness in road-area of vegetation was also analyzed. The accuracy of the model was validated by measured equivalent water thickness on the ground. The experimental results show: (1) The adjusted determination coefficient of RF-SVM model was the highest, established by Normalized Difference Water Index (NDWI), Normalized Multi-band Drought Index (NMDI), Simple Ratio Water Index (SRWI), Simple Ratio (SR), Normalized Difference Infrared Index (NDII), Water Index (WI), Dattwater Index (DWI), Moisture Stress Index (MSI) and Soil Adjusted Vegetation Index (SAVI), with the determination coefficient of verification set reaching 0.8877. (2) The RF-PLS and RF-GA-SVM models with the four water indices of NDWI, NMDI, SRWI, and SR had the highest adjusted determination coefficient, with the validation set's determination coefficients reaching

收稿日期:2019-05-22;修回日期:2019-09-20.

基金项目:国家自然科学基金面上项目(41671489)。[**Foundation items:** National Natural Science Foundation of China, No.41671489.]

作者简介:郭云开(1958—),男,湖南常德人,教授,主要从事高等级路域环境遥感研究。E-mail: guoyunkai226@163.com

*通讯作者:张晓炯(1993—),男,甘肃定西人,硕士生,主要从事高等级路域环境遥感研究。E-mail: 1183582609@qq.com

0.8053 and 0.8952, respectively. (3) Among them, the RF- GA- SVM model was the best for estimating equivalent water thickness, which met the requirements of vegetation equivalent water thickness monitoring in road area. Our findings provide an effective and accurate method for the estimation of equivalent water thickness, and provide support for road area environment monitoring based on hyper-spectral remote sensing.

Key words: Equivalent water thickness; random forest; PRO4SAIL model; water index; machine learning; vegetation of road area; Litan highway in Hunan Province

***Corresponding author:** ZHANG Xiaojiong, E-mail: 1183582609@qq.com

摘要: 植被等效水厚度对路域生态环境的监测评估具有重要意义。本研究以湖南醴潭高速一段为研究对象,以地面实测光谱和等效水厚度以及PRO4SAIL模拟光谱和模拟等效水厚度为数据源,利用PRO4SAIL冠层模型模拟光谱与地面实测光谱建立12种常用水分指数,引入随机森林算法对水分指数与等效水厚度进行重要性分析,得到12种水分指数的重要性排序;利用调整 R^2 确定建立等效水厚度估算模型中输入水分指数的最佳个数;在优选水分指数基础上,以PRO4SAIL模拟光谱计算得到水分指数和等效水厚度为训练集,分别构建随机森林耦合偏最小二乘(RF-PLS)、随机森林耦合支持向量机(RF-SVM)和随机森林耦合遗传算法优化支持向量机(RF-GA-SVM)等效水估算模型,并用地面实测等效水厚度对估算模型进行精度验证与分析。结果表明:RF-SVM估算模型中输入重要性前9的水分指数(NDWI、NMDI、SRWI、SR、NDII、WI、DWI、MSI、SAVI)时,调整 R^2 最高,验证集决定系数为0.8877;RF-PLS和RF-GA-SVM估算模型中输入重要性前4的水分指数(NDWI、NMDI、SRWI、SR)时,调整 R^2 最高,验证集决定系数分别为0.8053、0.8952,其中RF-GA-SVM模型估算等效水厚度效果最佳,其精度满足路域植被等效水厚度监测要求。本文研究成果为等效水厚度估算提供一种有效且精确的方法,同时为发展基于高光谱遥感的路域环境监测提供重要支撑。

关键词: 等效水厚度;随机森林;PRO4SAIL模型;水分指数;机器学习;路域植被;湖南醴潭高速

1 引言

高等级公路在施工建设、运营阶段会影响道路两侧植被覆盖和植被类型,监测路域植被的健康情况是衡量路域生态环境的关键,对指导路域生态环境保护具有重要意义。植被含水量是进行植被健康监测的一种重要评价指标,能在一定程度上反映植被长势和健康状况,是表征植被形态结构和生理状态的重要参数。描述植被含水量常用指标为冠层含水量(Canopy Water Content, CWC)、叶片等效水厚度(Equivalent Water Thickness, EWT)、活体可燃物湿度(Live Fuel Moisture Content, LFMC)和相对含水量(Relative Water Content, RWC),而叶片等效水厚度被认为是监测植被水分有效指标^[1-3]。传统等效水厚度(Equivalent Water Thickness, EWT)测量是实地采集植被样本,实验室干重法测得。该方法虽准确度较高,但会破坏植被,对于大面积获取成本高、时效性差^[4-5]。

遥感技术以其范围广、数据更新快,信息丰富等优势在等效水厚度定量反演中应用广泛^[6-7]。植被等效水厚度遥感估算方法有辐射传输模型反演法、植被指数法、查表法等^[8-10],其中建立植被指数与植被冠层水分之间的统计方法模型是植物水含量

估算的常用方法,具有一定的物理解释基础,计算方式简单、明确^[11-12]。例如,侯学会等^[9]针对地面实测数据较少时利用PROSAIL模型和Landsat8数据进行冠层含水量反演,PROSAIL和Landsat8构建的NDWI5(Normalized Difference Water Index5)估算EWT精度优于NDVI(Normalized Difference Vegetation Index)、EVI(Enhanced Vegetation Index)和NDWI7(Normalized Difference Water Index7);郑兴明等^[13]利用PROSAIL和Landsat OLI数据比较了12种水分指数的相关性,得出了MSI2(Moisture Stress Index2)与植被含水量的拟合关系最好;程晓娟等^[12]引入增强型植被指数EVI,构建的新水分指数NDWI⁺提高了冬小麦含水量的估算精度。以上研究均建立了单个植被指数与等效水厚度的拟合方程,其建模过程简单快速,取得了较好的估算效果,但当植被覆盖度较高时,利用单个植被指数估算等效水厚度易出现过拟合^[13],并且不同的植被指数对等效水的估算效果不同,适用性较弱。李丹娜^[8]等提出利用PRO4SAIL模拟光谱计算得到的NDWI、SRWI、GVMI与支持向量机回归对等效水厚度进行反演的方法,结果具有较好的精度。但选用多个植被指数综合进行等效水厚度估算时,是计算每个植被指数与等效水厚度之间相关性,没有同

时考虑多个植被指数对等效水厚度的综合影响。采用多个植被指数与机器学习算法进行等效水厚度估算,因不同的植被指数在不同的土壤背景和环境因素下适用性不同,选取的植被指数必须具有代表性,随机森林可以衡量植被指数对等效水厚度的贡献率,预先优选等效水厚度估算模型中输入的植被指数;其次,虽然利用PROSAIL与机器学习组合方法估算叶面积指数等其他植被参数研究成果丰富,但在等效水厚度研究方面研究较少^[13-15]。

综合以上考虑,本文利用随机森林算法对常见的12种水分指数优选,构建随机森林与遗传算法优化支持向量机(RF-GA-SVM)的组合估算模型,同时建立随机森林与支持向量机(RF-SVM)估算模型、随机森林与偏最小二乘(RF-PLS)估算模型进行对比分析,利用评价指标验证3种等效水估算模型的可靠性,分析优选之后的植被指数在路域植被等效水估算中的适用性。

2 数据获取与研究方法

2.1 冠层光谱与等效水厚度测量

本研究实验区域位于湖南省醴陵高速湘潭市至醴陵市之间(27.8°N-27.95°N, 113.05°E-113.35°E),研究区全长36.4 km,研究区域范围包括道路中线两侧1 km,研究区植被覆盖程度较高,植被类型主要为阔叶林,少部分为针叶和低矮植被。根据研究区植被类型、生长情况的不同特点,选择具有代表性的地面样方,每个样方以4个角点和中心点为测点,使用AvaField-3地物光谱仪对各个测点植被进行冠层光谱测量。对每个测点测量10条光谱,剔除异常曲线,采用移动窗口最小二乘多项式平滑模型对实测光谱进行平滑处理,最后取光谱曲线的平均值作为最终的样点实测光谱,共采集50组冠层光谱。采集样方内典型植被叶片,使用分析天平进行叶片鲜重测量。用YMJ-A手持叶面积测定仪测量叶片面积,将新鲜叶片带回实验室进行下一步处理。叶片干重采用烘箱在80~100 °C的条件下烘干,再获取叶片干重。叶片等效水厚度由(叶片鲜重-叶片干重)/叶片面积计算得到^[11-12],共采集得到50组等效水厚度数据。

2.2 PRO4SAIL模型模拟冠层光谱

PRO4SAIL模型由叶片光学特性模型PROSPECT5和冠层反射模型4SAIL耦合而成。PROS-

PECT5模型在PROSPECT模型的基础上增加了类胡萝卜素含量和褐色素含量。4SAIL模型增加了热点效应的影响和土壤二向反射特性。通过PROSPECT5叶片光学特性模型模拟得到的叶片反射率和透射率作为4SAIL模型的输入参数,结合太阳天顶角、LAI(Leaf Area Index)太阳方位角等参数,可模拟得到植被冠层400~2500 nm的反射率^[16-17]。本文在PRO4SAIL耦合模型的基础上通过一定步长的叶面积指数、等效水厚度变化进行植被冠层光谱的模拟,耦合模型所需输入的叶片生化参数等数据来自LOPEX'93数据库^[8]和野外实测数据,模型的输入参数如表1所示,共得到200组冠层模拟光谱。

表1 PRO4SAIL模型参数设置

Tab. 1 Input parameters of the PRO4SAIL model

输入参数	输入值
结构参数(N)	1.22
叶绿素含量(Cab)/($\mu\text{g}/\text{cm}^2$)	42.93
类胡萝卜素含量(Car)/($\mu\text{g}/\text{cm}^2$)	8
褐色素含量($Cbrown$)/($\mu\text{g}/\text{cm}^2$)	0
干物质重量(cm)/($\mu\text{g}/\text{cm}^2$)	0.02316
平均叶倾角(LAD)	30
叶面积指数(LAI)	4~6,步长:0.2
等效水厚度(EWT)	0.01~0.03,步长:0.001
热点参数($hspot$)	0.15
太阳天顶角(θ_s)/°	23.9
观测天顶角(θ_v)/°	0
土壤反射率($rsoil$)	0.2

2.3 水分指数

水分指数是基于对水分敏感的近红外和短红外波段通过组合运算建立的,水分指数可以有效地减少单波段的散射效应,也可以在一定程度上消除环境背景等对光谱特性的影响,并且比单波段更为敏感^[18]。参考已有的研究成果^[19-20],选取以下常用的12种水分指数建立分析研究,这12种水分指数计算公式见表2。

2.4 等效水厚度估算模型构建方法

2.4.1 遗传算法优化支持向量机

支持向量机(Support Vector Machine, SVM)是由Corinna Cortes和Vapnik等^[21-22]于1995年首先提出的。支持向量机是一种针对有限样本的,实现结构风险最小化思想的机器学习算法,它将实际问题通过非线性转换到高维的特征空间,在高维空间中

表2 12种水分指数的计算方法列表

Tab. 2 Twelve different water indices

水分指数	计算公式	公式编号	参考文献
新近红外肩部区域光谱比值指数 (Spectral Ratio Index in the NIR Shoulder Region, NSRI)	$NSRI = \frac{\rho_{890}}{\rho_{780}}$	(1)	[15]
水分指数 (Water Index, WI)	$WI = \frac{\rho_{900}}{\rho_{970}}$	(2)	[15]
归一化差值植被指数 (Normalized difference vegetation index, NDVI)	$NDVI = \frac{(\rho_{860} - \rho_{680})}{(\rho_{860} + \rho_{680})}$	(3)	[15]
归一化差值红外指数 (Normalized Difference Infrared Index, NDII)	$NDII = \frac{(\rho_{850} - \rho_{1650})}{(\rho_{850} + \rho_{1650})}$	(4)	[15]
Datt水分指数 (Datt Water Index, DWI)	$DWI = \frac{(\rho_{816} - \rho_{2218})}{(\rho_{816} + \rho_{2218})}$	(5)	[15]
简单水比指数 (Simple Ratio Water Index, SRWI)	$SRWI = \frac{\rho_{860}}{\rho_{1240}}$	(6)	[15],[20]
湿度胁迫指数 (Moisture Stress Index, MSI)	$MSI = \frac{\rho_{1600}}{\rho_{819}}$	(7)	[15],[19]
水比指数 (Simple Ratio, SR)	$SR = \frac{\rho_{895}}{\rho_{675}}$	(8)	[15]
归一化差值水分指数 (Normalized Difference Water Index, NDWI)	$NDWI = \frac{(\rho_{860} - \rho_{1240})}{(\rho_{860} + \rho_{1240})}$	(9)	[15]
归一化多波段干旱指数 (Normalized Multi-band Drought Index, NMDI)	$NMDI = \frac{\rho_{860} - (\rho_{1640} - \rho_{2130})}{\rho_{860} + (\rho_{1640} + \rho_{2130})}$	(10)	[15]
全球植被水分指数 (Global Vegetation Moisture Index, GVWI)	$GVWI = \frac{(\rho_{820} + 0.1) - (\rho_{1600} - 0.02)}{(\rho_{820} + 0.1) + (\rho_{1600} + 0.02)}$	(11)	[18]
土壤调整植被指数 (Soil Adjusted Vegetation Index, SAVI)	$SAVI = \frac{(\rho_{860} - \rho_{680}) \times (1 + L)}{(\rho_{860} + \rho_{680} + L)}, L = 0.5$	(12)	[18]

注: 计算公式中 ρ_i 表示为 i nm 处的反射率, 式(12)中 L 表示土壤调整指数。

构造线性决策函数实现原空间中非线性决策函数, 解决数据维度的同时保持了良好的泛化能力。

遗传算法 (Genetic Algorithm, GA) 是模拟达尔文的遗传选择和自然淘汰的生物进化过程的计算模型。遗传算法以决策变量编码作为运算对象, 可以更好应用遗传操作算子, 并且无需辅助信息和具有隐含并行性^[22]。遗传算法优化的支持向量机在等效水厚度建模预测过程中, 需先确定最适宜的支持向量机参数, 提高所建立模型的预测精度, 而参数优化的问题可以描述为式(13)。

$$\min f(C, \sigma) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

$$\begin{cases} C \in [C_{\min}, C_{\max}] \\ \sigma \in [\sigma_{\min}, \sigma_{\max}] \end{cases} \quad (14)$$

式中: C 表示支持向量机的惩罚系数; σ 表示核函数的宽度; y_i 表示 EWT 的观测值; \hat{y}_i 表示 EWT 的预测值。

将反演 EWT 训练数据输入到支持向量机建立模型, 采用遗传算法寻找最优参数 C 和 σ , 以式(14)为指标, 通过不断进化得到最优的参数。

2.4.2 偏最小二乘原理

偏小二乘 (Partial Least Squares, PLS) 提供一种多对多线性回归建模的方法, 特别当两组变量的个数很多, 且都存在多重相关性, 而观测数据的数量 (样本量) 又较少时, 用偏小二乘回归建立的模型具有传统的经典回归分析等方法所没有的优点。偏小二乘回归分析在建模过程中集中了主成分分析, 典型相关分析和线性回归分析方法的特点, 有效地提取对应变量解释能力最好的自变量, 删除多重共线信息和无解释物理变量, 能够避免多重共线的影响。具有较好的拟合效果, 以及鲁棒性^[25]。

2.4.3 水分指数与 EWT 重要性分析

随机森林是以决策树为基学习器的集成学习算法^[23], 用训练得到的随机森林对测试样本进行预测, 并用票选法决定预测的结果。随机森林可以度量每个特征 (水分指数) 在建立的决策树上所做贡献, 取其平均值, 来比较特征之间贡献的大小, 通常用基尼指数 (Gini Index) 或者袋外数据 (OOB) 错误率作为评价指标。本文利用水分指数与 PRO4SAIL 模拟 EWT 和实测的 EWT 进行特征重要性分析, 随着输入的水分指数的改变会产生估计误差, 随机森

林算法可以给出 OOB 变量误差,根据 OOB 变量误差计算输入水分指数的重要性,重要性评分越高说明该水分指数对于估算 EWT 越重要。

2.5 精度评价指标

对 RF-PLS、RF-SVM、RF-GA-SVM 估算模型进行精度验证,常用指标有决定系数(Coefficient of Determination, R^2),调整 R^2 (Adjusted Coefficient of Determination, $R^2_{Adjusted}$) 均方根误差 (Root Mean Square Error, $RMSE$),其计算公式为:

$$R^2 = 1 - \frac{RSS}{TSS} \tag{15}$$

$$R^2_{Adjusted} = 1 - \frac{\frac{RSS}{n-P-1}}{\frac{TSS}{n-1}} \tag{16}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \tag{17}$$

式中: $RSS = \sum_{i=1}^n (y_i - \bar{y})^2$, $TSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$; y_i 表示测量值; \bar{y} 表示测量值的平均值; \hat{y}_i 表示预测值; n 为总样本个数; P 为自变量个数。决定系数越大,拟合效果越好,均方根误差越小,表明估算模型的拟合精度越高。

3 数据分析及模型建立

3.1 水分指数重要性分析

在随机森林算法中输入水分指数与 EWT,设置决策树数目 N 为 1000,随机特征变量 m 的取值为 6,得到具体重要性评分见表 3。从表 3 可看出,重要性评分大于 0.15 的水分指数有 3 个,重要性评分最高的是 NDWI,为 0.224 127,重要性评分最低的是 GVWI,为 0.000 273。

3.2 植被水分指数建模分析

在以上分析的基础上按照表 3 重要性评分排序依次输入 PLS、SVM、GA-SVM 模型中进行回归分析,其目的在于根据调整 R^2 确定建立等效水厚度估算模型中输入的水分指数的个数与类别。第一次输入一个,第 n 次输入 n 个参数($n=12$),最终得到调整 R^2 (图 1)。结果表明,基于 PLS 建模时,输入水分指数为 NDWI、NMDI、SRWI、SR 时,建模的调整 R^2 最高,调整 R^2 为 0.7793;基于 SVM 建模时,输入水分指数为 NDWI、NMDI、SRWI、SR、NDII、WI、DWI、MSI、SAVI 时,建模的调整 R^2 最高,调整 R^2 为 0.8812;基于 GA-SVM 建模时,输入水分指数为

表 3 水分指数与 EWT 重要性分析及排序

Tab. 3 Importance analysis and ordination between the water indices and EWT

特征编号	植被指数	重要性	特征编号	植被指数	重要性
1	NDWI	0.2241	7	DWI	0.0678
2	NMDI	0.1932	8	MSI	0.0571
3	SRWI	0.1581	9	SAVI	0.0491
4	SR	0.0734	10	NSRI	0.0301
5	NDII	0.0723	11	NDVI	0.0033
6	WI	0.0711	12	GVWI	0.0003

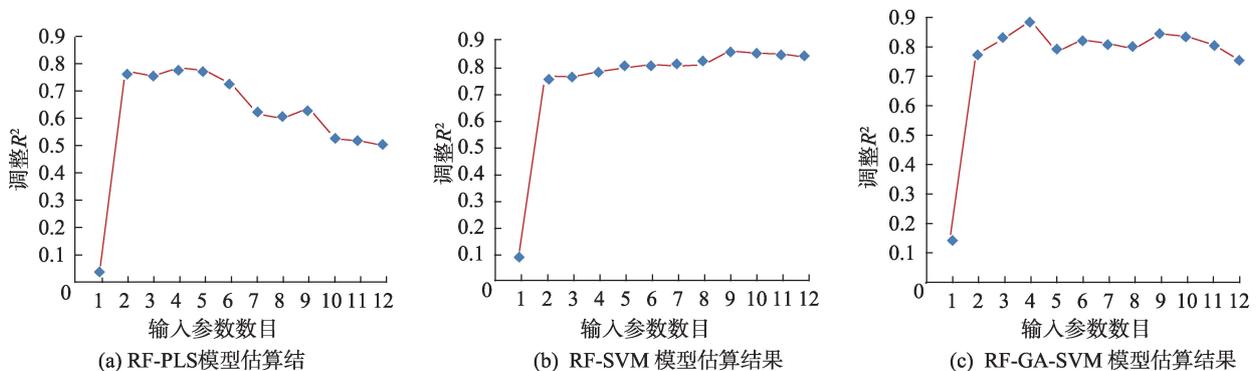


图 1 3 种模型估算结果对比

Fig. 1 Comparison of estimated EWT by three models

NDWI、NMDI、SRWI、SR时,建模的调整 R^2 最高,调整 R^2 为0.8524。3种估算模型的调整 R^2 总体变化趋势相近,总体趋势均为先增大后减少。

3.3 植被EWT估算模型

本文基于随机森林选择最优水分指数建立估算植被等效水厚度的模型。以植被指数为自变量, EWT为因变量,构建RF-PLS、RF-SVM、RF-GA-SVM等效水厚度遥感估算模型,并采用上述评价指标对3种模型进行精度验证。为不失一般性,将PRO4SAIL模拟的水分指数和EWT随机抽取70组作为训练样本,实测的25组EWT作为验证数据。EWT估算模型样本构建表如表4所示。本文SVM算法中核函数选用高斯核函数,惩罚参数C为1,核函数参数 σ 取1。在GA-SVM回归算法中设置遗传算法寻优初始参数C、 σ 的范围为0-100,种群数量为20,最大寻优次数为1000,最终得到的best C=0.48075, best σ =8.4211。基于以上3种算法,得到3个模型的EWT预测数据,通过与对应的EWT实测值比较,评价不同模型预测能力,同时对EWT实测值和预测值进行回归分析,绘制3种模型EWT实测值和预测值的1:1线。

4 结果及分析

4.1 3种等效水模型精度评价

由图2可知,RF-PLS模型验证精度分别为 $R^2=0.8053$ 、 $RMSE=0.00303$;RF-SVM模型验证精度分别为 $R^2=0.8877$ 、 $RMSE=0.002386$;RF-GA-SVM模型验证精度为 $R^2=0.8952$ 、 $RMSE=0.002313$ 。可以看出,RF-SVM和RF-GA-SVM验证精度相近,但RF-GA-SVM估算模型只需要输入4个水分指数,可认为本文建立的3种组合方法中RF-GA-SVM对于反演等效水厚度效果最好,表明该模型具有较好的预测能力。

4.2 等效水估算结果讨论

建模方法对等效水估算精度至关重要,本文引入遗传算法优化的支持向量机,偏最小二乘,支持向量机3种机器学习算法,构建3种等效水估算模型,利用实测EWT对估算模型进行验证。RF-PLS估算模型验证决定系数为0.8053;RF-SVM估算模型验证决定系数为0.8877;RF-GA-SVM估算模型验证决定系数为0.8952。RF-GA-SVM估算模型精

表4 建模样本和验证样本统计

Tab. 4 Descriptive statistics of the data samples for modeling and validation

反演模型	建模样本		预测样本	
	自变量	应变量	自变量	应变量
PLS	PRO4SAIL 模拟NDWI、NMDI、SRWI、SR	PRO4SAIL 模拟EWT	实测NDWI、NMDI、SRWI、SR	实测EWT
SVM	PRO4SAIL 模拟NDWI、NMDI、SRWI、SR、NDII、WI、DWI、MSI、SAVI	PRO4SAIL 模拟EWT	实测NDWI、NMDI、SRWI、SR、NDII、WI、DWI、MSI、SAVI	实测EWT
GA-SVM	PRO4SAIL 模拟NDWI、NMDI、SRWI、SR	PRO4SAIL 模拟EWT	实测NDWI、NMDI、SRWI、SR	实测EWT

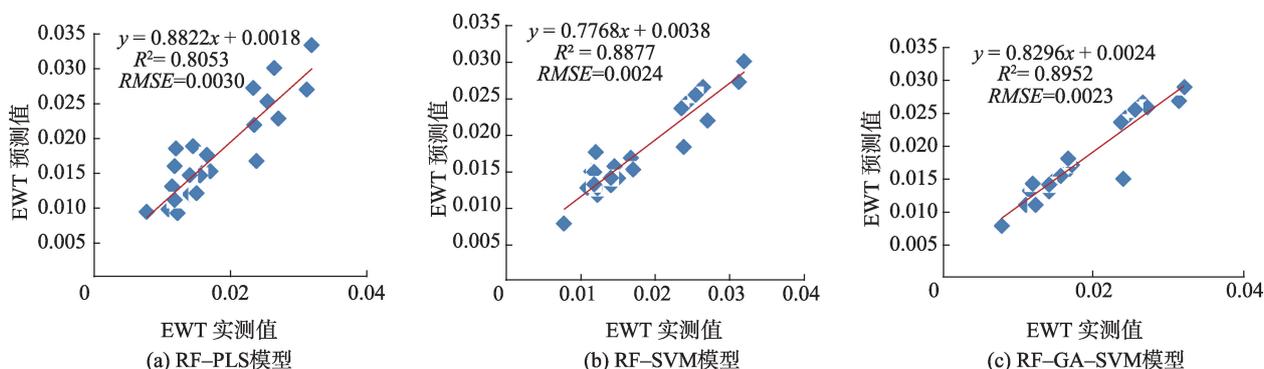


图2 EWT预测值与实测值拟合关系

Fig. 2 Fitting relationship between predicted and measured values of EWT

度优于RF-PLS和RF-SVM估算模型,原因可能在于RF-PLS估算模型中偏最小二乘回归为线性模型,其生成的特征值不具物理可解释性,降低了模型泛化能力,在RF-SVM估算模型中支持向量机中惩罚参数和核参数是人为确定的,其具有随机性,这一因素限制了该模型估算能力,而RF-GA-SVM中,遗传算法能够通过迭代搜索算法最优的惩罚参数和核参数^[21-22,24],使RF-GA-SVM模型精度明显改善,表明此模型能够满足生态环境对路域植被等效水厚度精度需求。

单一水分指数在等效水估算中由于包含波段信息较少,植被密度较高时,易造成不同程度的饱和现象,因此限制了采用单一水分指数估算等效水厚度的应用。本文利用随机森林算法对12种水分指数进行优选,得出RF-PLS、RF-GA-SVM估算模型选择NDWI、NMDI、SRWI、SR做为自变量模型估算效果最佳,而RF-SVM估算模型选择NDWI、NMDI、SRWI、SR、NDII、WI、DWI、MSI、SAVI作为自变量模型估算效果最佳。本文选择的这些水分指数可以对其他植被覆盖度较高的地区估算等效水厚度提供借鉴。

5 结论与讨论

5.1 结论

等效水厚度是衡量植被生长状况的重要指标。以辐射传输模型模拟数据和实地测量数据为数据来源,利用随机森林特征重要性分析与常见的机器学习算法结合,构建路域植被EWT遥感估算模型。主要研究结果如下:

(1)利用随机森林算法对12水分指数与实测等效水厚度和模拟等效水厚度进行重要性分析,得出12种水分指数的重要性排序,重要性前4的水分指数为NDWI、NMDI、SRWI、SR,其中NDWI的重要性评分为0.2241。

(2)优选建立了等效水厚度估算模型的水分指数。按水分指数重要性评分排序依次输入PLS、SVM、GA-SVM这3种回归模型中,按调整 R^2 确定每一种估算模型中输入的最佳水分指数。得出基于PLS和GA-SVM建模时,输入的最佳水分指数为NDWI、NMDI、SRWI、SR;基于SVM建模时,输入的最佳水分指数为NDWI、NMDI、SRWI、SR、NDII、WI、DWI、MSI、SAVI。

(3)建立了等效水厚度估算模型并进行对比分析。利用随机森林选择最优的水分指数后,建立了3种等效水厚度估算模型(RF-PLS、RF-SVM和RF-GA-SVM)利用实测EWT进行精度分析评价,RF-GA-SVM估算等效水厚度的效果最佳,决定系数为0.8952,均方根误差为0.002313,RF-PLS估算等效水精度较差。RF-GA-SVM等效水估算模型可满足路域植被等效水厚度监测要求。

5.2 讨论

对于已有的采用相关性分析方法选择水分指数,本文采用随机森林算法对水分指数预先进行优选,能够判断对路域植被等效水厚度估算重要的植被指数,选择最佳的水分指数建立等效水估算模型,而且可以减少估算模型中输入数据的维度。本文引入的遗传算优化的支持向量机在估算等效水厚度精度较高,与已有的方法相比,其精度有很大改善,选择的植被指数以及建立的EWT估算模型可为路域植被以及其他区域等效水厚度估算提供参考。本文选取的12种水分指数为红光波段与近红外波段组合,鉴于不同的水分指数在抗饱和能力和适用性等方面各有特点,是否还有其他的水分指数对于路域植被等效水厚度估算更为有效,是今后研究的一个重要方向。此外,路域环境中植被覆盖度较高,本文中的方法在其他植被覆盖稀疏地区应用时,其适用性还需进一步研究。

参考文献(References):

- [1] 张峰,周广胜.植被含水量高光谱遥感监测研究进展[J].植物生态学报,2018,42(5):517-525. [Zhang F, Zhou G S. Research progress on monitoring vegetation water content by using hyperspectral remote sensing[J]. Chinese Journal of Plant Ecology, 2018,42(5):517-525.]
- [2] Dziki S, Verreynne J S, Stuckens J, et al. Determining the water status of Satsuma mandarin trees "citrus unshiu-umarcovitch" using spectral indices and by combining hyperspectral and physiological data[J]. Agricultural and Forest Meteorology, 2009,150(3):369-379.
- [3] De Jong S M, Addink E A, Hoogenboom P, et al. The spectral response of Buxus sempervirens to different types of environmental stress: A laboratory experiment[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2012,74:56-65.
- [4] 刘良云,王纪华,张永江,等.叶片辐射等效水厚度计算与叶片水分定量反演研究[J].遥感学报,2007,11(3):289-295. [Liu L Y, Wang J H, Zhang Y J, et al. Detection of

- leaf EWT by calculating REWT from reflectance spectra [J]. *Journal of Remote Sensing*, 2007,11(3):289-295.]
- [5] 赵英时等. 遥感应用分析原理与方法[M]. 北京: 科学出版社, 2003. [Zhao Y S, et al. Principles and methods of remote sensing application analysis[M]. Beijing: Science Press, 2003.]
- [6] Seelig H D, Hoehn A, Stodieck L S, et al. Relations of remote sensing leaf water indices to leaf water thickness in cowpea, bean, and sugarbeet plants[J]. *Remote Sensing of Environment*, 2007,112(2):445-455.
- [7] Clevers J G P W, Kooistra L, Schaepman M E. Using spectral information from the NIR water absorption features for the retrieval of canopy water content[J]. *International Journal of Applied Earth Observations and Geo-information*, 2008,10(3):388-397.
- [8] 李丹娜, 郭云开, 朱善宽, 等. 利用PRO4SAIL与支持向量机回归的组合模型反演植被等效水厚度[J]. *测绘通报*, 2017(8):50-55. [Li D N, Guo Y K, Zhu S K, et al. Quantitative inversion vegetation equivalent water thickness by combined model of PRO4SAIL and the support vector regression[J]. *Bulletin of Surveying and Mapping*, 2017(8): 50-55.]
- [9] 侯学会, 王猛, 刘思含, 等. 基于Prosail模型和Landsat 8数据的小麦冠层含水量反演比较[J]. *麦类作物学报*, 2018, 38(4):493-497. [Hou X H, Wang M, Liu S H, et al. Comparison between prosailmodel and Landsat 8 images in inversion of water content of wheat canopy[J]. *Journal of Triticeae Crops*, 2018,38(4):493-497.]
- [10] Liu L L, Zhang S, Zhang B. Evaluation of hyperspectral indices for retrieval of canopy equivalent water thickness and gravimetric water content[J]. *International Journal of Remote Sensing*, 2016,37(14):3384-3399.
- [11] Gao B C. NDWI: A normalized difference water index for remote sensing of vegetation liquid water from space[J]. *Remote Sensing of Environment*, 1996,58(3):257-266.
- [12] 程晓娟, 杨贵军, 徐新刚, 等. 新植被水分指数的冬小麦冠层水分遥感估算[J]. *光谱学与光谱分析*, 2014,34(12): 3391-3396. [Cheng X J, Yang G J, Xu X G, et al. Estimating canopy water content in wheat based on new vegetation water index[J]. *Spectroscopy and Spectral Analysis*, 2014,34(12):3391-3396.]
- [13] 郑兴明, 丁艳玲, 赵凯, 等. 基于Landsat 8 OLI数据的玉米冠层含水量反演研究[J]. *光谱学与光谱分析*, 2014,34(12):3385-3390. [Zhen X M, Ding Y L, Zhao K, et al. Estimation of vegetation water content from Landsat 8 OLI data[J]. *Spectroscopy and Spectral Analysis*, 2014,34(12): 3385-3390.]
- [14] Liang L, Di L P, Zhang L P, et al. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method[J]. *Remote Sensing of Environment*, 2015, 165:123-134.
- [15] 张春兰, 杨贵军, 李贺丽, 等. 基于随机森林算法的冬小麦叶面积指数遥感反演研究[J]. *中国农业科学*, 2018,51(5): 855-867. [Zhang C L, Yang G J, Li H L, et al. Remote sensing inversion of leaf area index of winter wheat based on random forest algorithm[J]. *Scientia Agricultura Sinica*, 2018,51(5):855-867.]
- [16] Jacquemoud S, Verhoef W. PROSPECT+SAIL models: A review of use for vegetation characterization[J]. *Remote Sensing of Environment*, 2009,113:56-66.
- [17] Jacquemoud S, Baret F. Prospect: A model of leaf optical properties spectra[J]. *Remote Sensing of Environment*, 1990,34:75-91.
- [18] Hedja C, Florent M, Abdelaziz H. Regional equivalent water thickness modeling from remote sensing across a tree Cover/LAI gradient in mediterranean forests of northern Tunisia[J]. *Molecular Diversity Preservation International (MDPI)*, 2015,7(2):1937-1961.
- [19] Pietro C, Stéphane F, Stefano T, Set al. Detecting vegetation leaf water content using reflectance in the optical domain [J]. *Remote Sensing of Environment*, 2001,77(1):22-33.
- [20] Zarco-Tejada P J, Rueda C A, Ustin S A. Water content estimation in vegetation with MODIS reflectance data and model inversion methods[J]. *Remote Sensing of Environment*, 2003,85(1):109-124.
- [21] 牛亚莉. 遗传优化支持向量机的交通流量预测模型[J]. *微型电脑应用*, 2017,33(8):72-74. [Niu Y L. Traffic flow forecasting model by Support Vector Machine optimized by genetic algorithm[J]. *Microcomputer Applications*, 2017,33(8):72-74.]
- [22] 付俊娥, 苏庆珣, 潘世兵, 等. 基于支持向量机理论的地下水动态遥感监测模型与应用[J]. *地球信息科学学报*, 2010,12(4):466-472. [Fu J E, Su Q X, Pan S B, et al. Support vector machine based ground water level monitoring model by using remote scence images[J]. *Journal of Geo-information Science*, 2010,12(4):466-472.]
- [23] 沈润平, 郭佳, 张婧娴, 等. 基于随机森林的遥感干旱监测模型的构建[J]. *地球信息科学学报*, 2017,19(1):125-133. [Shen R J, Guo J, Zhang J X, et al. Construction of a drought monitoring model using the randform forest based remote sensing[J]. *Journal of Geo-information Science*, 2017,19(1):125-133.]
- [24] 李鹏霞. 基于偏最小二乘和支持向量机的生物分子的定量结构—活性/性质关系研究[D]. 兰州: 西北师范大学, 2011. [Li X P. Quantitative structure-activity/property relationship studies in biomolecules based on partial least squares and support vector machine[D]. Lanzhou: Northwest Normal University, 2011.]