

doi:10.6041/j.issn.1000-1298.2020.07.020

基于随机森林偏差校正的农业干旱遥感监测模型研究

刘冀^{1,2} 张特^{1,2} 魏榕^{1,2} 张茜^{1,2} 刘艳丽^{3,4} 董晓华^{1,2}

(1. 三峡大学水利与环境学院, 宜昌 443002; 2. 三峡库区生态环境教育部工程研究中心, 宜昌 443002;
3. 南京水利科学研究院水文水资源与水利工程科学国家重点实验室, 南京 210098;
4. 水利部应对气候变化研究中心, 南京 210029)

摘要: 以3个月尺度的标准化降水蒸散指数(SPEI3指数)为因变量,采用融合多源遥感数据的随机森林(RF)算法构建淮河流域2001—2014年作物生长季(4—10月)的农业干旱监测模型,采用简单线性回归、偏差估算法、旋转残差法和最优角度残差旋转法4种方法进行模型结果校正,以决定系数(R^2)、均方根误差(RMSE)及干旱等级监测准确率对模型监测能力进行评估。选取最优校正方法,构建随机森林偏差校正干旱监测模型(Bias-correcting random forest drought condition, BRFDC),通过站点实测土壤相对湿度及干旱事件记录对模型干旱监测能力进行验证。结果表明:采用最优角度残差旋转法校正后,模型模拟精度指标 R^2 和RMSE分别为0.897、0.874和0.335、0.362,优于其他校正方法;偏差估算法对各类干旱等级监测更为准确,尤其是对极端干旱的监测准确率最高,达到33.3%~50.0%,最终采用偏差估算法作为最优校正方法,构建BRFDC模型;相比SPEI3,BRFDC模型计算指数与大部分站点土壤相对湿度的相关性更加显著($P<0.01$),适于农业干旱监测;BRFDC模型能够准确监测淮河流域2001年严重干旱事件的时空演变过程,并能有效识别极端旱情。该模型可为淮河流域农业抗旱工作的有效开展提供科学依据。

关键词: 农业干旱监测; 遥感; 随机森林; 偏差校正

中图分类号: P426.616; S127 文献标识码: A 文章编号: 1000-1298(2020)07-0170-08 OSID: 

Development of Agricultural Drought Monitoring Model Using Remote Sensing Based on Bias-correcting Random Forest

LIU Ji^{1,2} ZHANG Te^{1,2} WEI Rong^{1,2} ZHANG Qian^{1,2} LIU Yanli^{3,4} DONG Xiaohua^{1,2}

(1. College of Hydraulic and Environmental Engineering, China Three Gorges University, Yichang 443002, China
2. Engineering Research Center of Eco-environment in Three Gorges Reservoir Region, Ministry of Education, Yichang 443002, China
3. State Key Laboratory of Hydrology, Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210098, China
4. Research Center for Climate Change, Ministry of Water Resources, Nanjing 210029, China)

Abstract: Drought is a frequent natural hazard in the Huaihe River Basin (HRB). Traditional agricultural drought monitoring methods have defects in spatial continuity, so developing an accurate agricultural drought monitoring model at regional scale is necessary. As a popular method, random forest (RF) is widely used due to its high prediction accuracy. However, RF may have significant bias in regression at times, especially for extreme values. The standardized precipitation evapotranspiration index for the 3-month time scale (SPEI3) was used as the dependent variable, and the multi-source satellite product from tropical rainfall measure mission (TRMM) and moderate-resolution imaging spectroradiometer (MODIS) was fused by RF to construct agricultural drought monitoring model in two regions of the HRB from April to October in 2001—2014. The accuracy of four bias-correcting methods, including simple linear regression (SLR), bias corrected method (BC), residual rotation method (RR) and best-angle residual rotation method (BRR) were assessed by determination coefficient (R^2), root mean square error (RMSE) and correct percentage of drought grades. The best bias-correcting method was used to establish agricultural drought monitoring model, which was called bias correcting random

forest drought condition model (BRFDC). The relative soil humidity data and drought records were applied to test the monitoring capacity of BRFDC model. The results showed that all of four bias-correcting methods improved the performance compared with original RF. The BRR method performed better with R^2 were 0.897 and 0.874, and RMSE were 0.335 and 0.362, which reduced the residuals efficiently. Additionally, the BC method performed better by the accuracy rate of different ranks of drought, especially the accuracy of extreme drought was between 33.3% and 50.0%. The BC method was applied to construct BRFDC at last. Compared with SPEI3, the outputs of BRFDC model had more significant correlation with soil relative humidity at most stations. Finally, the drought maps during the period from May to October in 2001 were produced by inverse distance weighting method (IDW), original RF and BRFDC model, and all of them showed a strong visual agreement. In particular, the extreme drought conditions were successfully monitored by BRFDC model.

Key words: agricultural drought monitoring; remote sensing; random forest; bias-correcting

0 引言

干旱是一种普遍的自然灾害,对农业、生态环境、社会经济产生重大影响^[1]。在全球气候变暖的大背景下,未来全球干旱频率将逐渐上升^[2]。淮河流域是我国重要的农业基地,亦是我国的干旱多发区^[3]。因此,针对淮河流域开展农业干旱监测模型研究对抗旱工作的开展具有重要的现实意义。

帕尔默指数 (Palmer drought severity index, PDSI)、标准化降水指数 (Standardized precipitation index, SPI)、标准化降水蒸散指数 (Standardized precipitation evapotranspiration index, SPEI) 被广泛应用于干旱监测,并取得了较好的监测效果^[4-6],尤其是 SPEI 指数与土壤水分之间有较强的相关性^[7-8]。干旱是一种区域性现象,这些指数通过站点观测数据计算得出,仅能准确反映站点附近的旱情。传统方法通过空间插值实现区域干旱监测,其结果受地理因素与插值方法等影响,存在较大的不确定性^[9]。遥感技术能够提供降水量、地表温度以及植被长势等空间分布信息,为实现区域农业干旱监测提供了有力支撑。一些研究尝试应用多源遥感数据构建综合干旱监测指数,并证实了多源遥感数据在干旱监测中的应用潜力,但这些研究采用权重方法结合多源遥感干旱因子,难以描述不同干旱因子间的非线性关系^[10-12]。近年来,一些学者采用支持向量机、人工神经网络、分类回归树、深度人工神经网络等机器学习方法,融合多源遥感信息构建模型,预测旱情空间分布,为实现区域干旱监测提供了新思路^[13-14]。

随机森林 (Random forest, RF) 是基于集成思想的机器学习算法,较单一分类器的误差更加稳定,具有较强的抗过拟合能力^[15]。董婷等^[16]采用 RF、Cubist 和 Bagging 3 种方法在黄淮地区构建了综合干旱指数,结果表明,RF 算法的拟合能力更强。PARK 等^[9]研究表明,相比增强回归树、Cubist 方

法,RF 算法对旱情预测更加准确。RF 算法基于多棵决策树的平均结果进行预测,其结果相对准确可信,但同时也会导致一定的偏差,尤其是对极值的预测能力较弱^[17]。对于干旱灾害而言,极端情况往往带来更大的损失,更应予以充分重视。相关研究表明,对 RF 算法进行偏差校正能有效提升预测结果的准确性^[18]。

本文在综合比较多种 RF 偏差校正方法适用性的基础上,合理选取最优偏差校正方法,构建基于多源遥感数据的淮河流域农业干旱监测模型,并分析模型在农业干旱监测中的适用性,以期提高空间旱情监测的准确性,为抗旱工作的开展提供科学支撑。

1 研究区概况

淮河流域位于我国东部 ($111^{\circ}55' \sim 121^{\circ}25'E$, $30^{\circ}55' \sim 36^{\circ}36'N$),地跨湖北、河南、安徽、江苏及山东五省,是我国重要的农业基地。流域面积约 $270\,000\text{ km}^2$,其中约有 $2/3$ 的面积为平原和 $1/3$ 的面积为丘陵山地。淮河流域地跨我国南北气候带,淮河以北属暖温带区,淮河以南为北亚热带区,流域内气候温和,年平均降雨量约 $1\,040\text{ mm}$,年平均气温 $11 \sim 16^{\circ}\text{C}$,年内降雨集中,汛期降水占全年降水量的 $60\% \sim 80\%$,其独特的气候条件和地理位置,导致流域内的旱涝灾害频发。流域位置、高程及土地利用概况如图 1 所示。

2 数据与方法

2.1 研究数据

研究数据主要包括 1970—2014 年的气象数据以及 2001—2014 年的遥感数据两部分。气象数据主要包括淮河流域内 31 个气象站的逐日气象数据,含 14 个农业气象站的逐旬土壤相对湿度数据,统一处理至月尺度。遥感数据包括 TRMM (Tropical rainfall measure mission) 3B43 降雨数据、MODIS (Moderate-resolution imaging spectroradiometer) 的

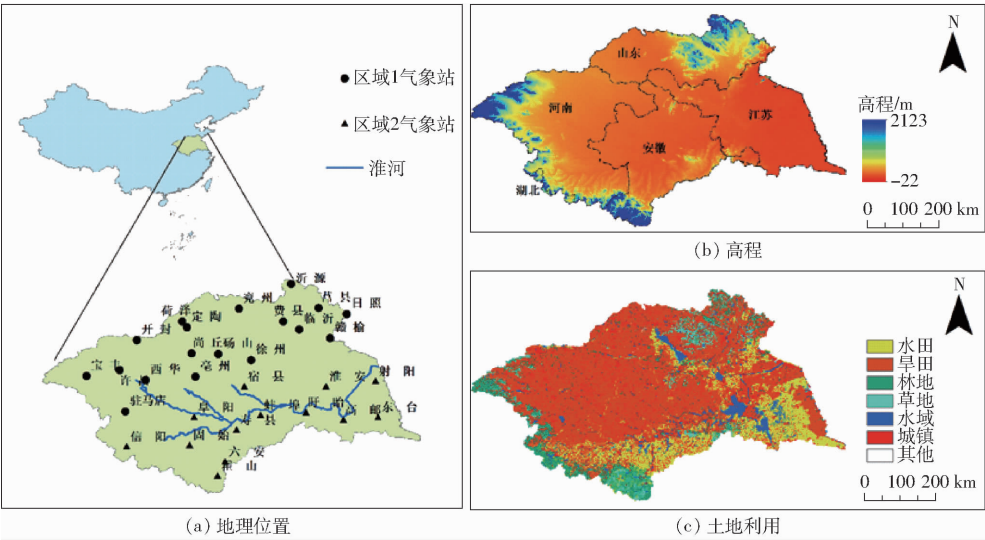


图1 淮河流域地理位置、高程及土地利用情况

Fig.1 Geographic location, elevation and land use type of Huaihe River Basin

MOD11A3 地表温度数据和 MOD13A3 归一化植被指数 (Normalized vegetation index, NDVI)、STRM (Shuttle radar topography mission) 高程数据。此外还包括土壤有效持水量 (Available water capacity of soil,AWC) 栅格数据,统一处理至 1 km 空间分辨率。各项数据来源信息如表 1 所示。

表 1 研究数据来源及说明

Tab.1 Sources and instructions of data

数据名称	时空分辨率	数据用途	来源
TRMM 3B43	25 km、月	获取流域降水	NASA (http://mirador.gsfc.nasa.gov)
MOD11A3	1 km、月	获取流域地表温度	地理空间数据云
MOD13A3	1 km、月	获取流域植被信息	NASA (http://mirador.gsfc.nasa.gov)
DEM	90 m	获取流域高程	地理空间数据云
AWC	1 km	获取流域有效持水量	http://globalchange.bnu.edu.cn
气象数据	日	计算参考蒸发量	中国气象数据网
土壤湿度	旬	获取站点土壤湿度	中国气象数据网

2.2 研究方法

2.2.1 SPEI 指数

SPEI 指数采用降水与潜在蒸发的亏缺程度来量化干旱程度^[19],计算式为

$$D_j^k = \sum_{i=j-k+1}^j (P_i - PET_i) \tag{1}$$

式中 D_j^k ——水分亏缺量,mm
 P_i ——降雨量,mm
 PET_i ——潜在蒸发量,mm
 k ——时间尺度
 i ——分析时段起始时间
 j ——分析时段终止时间

SPEI 指数可有效评估不同时间尺度的干旱情况,其中 3 个月尺度的 SPEI 指数与土壤湿度有较好的相关性^[8]。通常采用 Thornthwaite 公式计算 PET ,但该公式受温度影响较大,因此本文采用 Penman-Monteith 公式计算的参考蒸发量 ET_0 代替 PET ,并计算 1970—2014 年各站点 SPEI3 指数。当

$SPEI < -2$ 时,定义为极端干旱(极旱);当 $-2.0 \leq SPEI < -1.5$ 时,定义为严重干旱(重旱);当 $-1.5 \leq SPEI < -1.0$ 时,定义为中等干旱(中旱);当 $SPEI \geq -1.0$ 时,定义为无旱。

2.2.2 遥感指数

对各像元 1、3 个月的平均降水量、地表温度、NDVI 进行归一化处理,即可计算出不同时间尺度(1、3 个月)的降水状态指数 (Precipitation condition index, PCI)、地表温度状态指数 (Temperature condition index, TCI) 以及植被状态指数 (Vegetation condition index, VCI),计算方法如表 2 所示。

表 2 遥感指数计算方法

Tab.2 Computation of different remote sensing drought monitoring indexes

遥感指数	计算方法
降水状态指数 (PCI)	$(P - P_{min}) / (P_{max} - P_{min})$
地表温度状态指数 (TCI)	$(LST_{max} - LST) / (LST_{max} - LST_{min})$
植被状态指数 (VCI)	$(NDVI - NDVI_{min}) / (NDVI_{max} - NDVI_{min})$

2.2.3 随机森林算法

随机森林是一种基于决策树的集成算法,在回归及分类问题中应用广泛。该方法通过 Bootstrap 自助法在原数据集中有放回地随机抽取 n 个样本集,每次抽取 $2/3$ 样本容量的数据作为袋内数据,建立 n 棵决策树构建随机森林,并以 n 棵决策树回归结果的平均值进行预测。该方法的优点在于训练速度相对较快且不需要进交叉验证,同时抽样随机性和特征选择随机性使得随机森林不易陷入过拟合^[20]。随机森林方法中的两个重要参数分别为决策树的数量 n_{tree} 及随机选取特征变量的数量 m_{try} ,算法详见文献[17]。本文采用 R 语言中的 Random forest 程序包 (http://cran.at.r-project.org/src/contrib/Archive/randomForest/randomForest_4.6-7.tar.gz) 进行计算。

2.2.4 偏差校正方法

为提高 RF 模型精度,采用 4 种偏差校正方法对 RF 的回归结果进行校正。

(1)简单线性回归法

简单线性回归法 (Simple linear regression, SLR) 主要通过建立袋内数据预测值与实际值间的线性关系式,并应用该线性关系式对袋外样本数据的预测值进行修正,以达到偏差校正效果。计算公式为

$$y_{obs} = a + b\hat{y}_{pre} \tag{2}$$

式中 y_{obs} ——实际值 $\hat{y}_{pre}$ ——预测值
 a 、 b ——系数

(2)偏差估算法

ZHANG 等^[21] 提出的偏差估算法 (Bias corrected, BC) 以预测值和实际值之间的残差为因变量,以原样本数据及预测值为自变量,通过 RF 构建估计残差的模型,在原始预测结果上进行偏差校正。计算式为

$$r = y_{obs} - \hat{y}_{pre} \tag{3}$$

$$\hat{r} = RF_{res}(X, r) \tag{4}$$

$$\hat{y}_{cor} = \hat{y}_{pre} + \hat{r} \tag{5}$$

式中 r ——实际残差 $\hat{r}$ ——预测残差
 RF_{res} ——残差计算函数
 X ——样本数据
 $\hat{y}_{cor}$ ——校正后的预测值

(3)残差旋转法

SONG^[18] 提出基于残差旋转 (Residual rotation, RR) 的偏差校正方法,能有效降低原始 RF 的均方误差。该方法通过建立预测值与残差间的线性关系,将拟合直线旋转至与 X 轴重合位置,通过旋转矩阵估计残差对原始预测值进行校正。

(4)最优角度残差旋转法

最优角度残差旋转法 (Best-angle residual rotation, BRR) 基于 RR 法,对旋转角进行寻优,使整体均方误差最小。例如,将拟合曲线旋转 α 可与 X 轴重合,在 $(\alpha - \beta, \alpha + \beta)$ 区间寻找一个最优角度,采用该角度下的旋转矩阵对预测值进行校正。

2.2.5 模型精度评估方法

以站点 SPEI3 为因变量,以对应站点的 PC11、TC11、VCI1、PC13、TC13、VCI3 (1、3 分别代表 1、3 个月时间尺度)、高程、AWC 共 8 个变量为自变量,建立 2001—2014 年作物生长季 (4—10 月) 的 RF 模型,采用上述 4 种偏差校正方法对 RF 预测结果进行校正,并选取最优的校正方法构建随机森林偏差校正干旱监测模型 (Bias-corrected random forest drought condition, BRFDC)。采用均方根误差 (Root mean square error, RMSE) 最小原则确定 RF 模型参数 m_{try} ,并选取决定系数 (Determination coefficient, R^2)、RMSE 以及干旱等级预测准确度评估 BRFDC 模型精度,综合比较后确定最优校正方法,最后通过站点土壤相对湿度数据及流域典型干旱事件记录对 BRFDC 模型的监测能力进行评估。此外,考虑到淮河流域特殊的气候分区特征,采用 C 均值模糊聚类法将淮河流域划分为两个气候区分别研究,计算流程如图 2 所示。

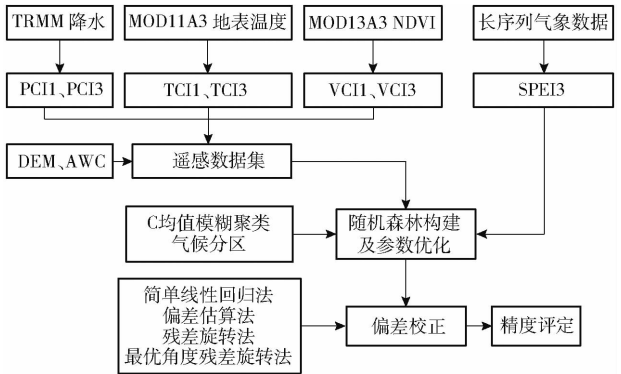


图 2 计算流程图
Fig. 2 Flowchart of procedures used

3 结果与分析

3.1 流域分区及 RF 参数优化

以各气象站年降雨量与参考蒸发量的差值 ($P - ET_0$) 为依据,将淮河流域各气象站划归为北 (区域 1)、南 (区域 2) 两个分区,站点聚类结果如图 1 所示,各分区气象站特征如表 3 所示。由表 3 可知,区域 1、区域 2 的多年平均降雨量分别为 755、1 018 mm,其气候差异主要体现在年降雨量,而年平均气温 T 和 ET_0 的差异较小。

本文构建的 RF 模型包含 8 个变量,从 1 ~ 8 逐一取值以确定最优参数 m_{try} ,而 n_{tree} 参数统一设置为

表 3 不同区域气候特征

分区	P/mm	$T/^{\circ}\text{C}$	ET_0/mm
区域 1	755	14.1	1 161
区域 2	1 018	15.2	1 088

1 000。将不同参数 m_{try} 的 RF 模型独立运行 100 次,各参数取值下的平均 RMSE 如图 3 所示。由图 3 可知,当 RMSE 最小时,区域 1 中 4—10 月各模型的 m_{try} 取值分别为 4、7、4、3、5、4、5,区域 2 的 m_{try} 取值分别为 5、5、7、6、5、7、5。

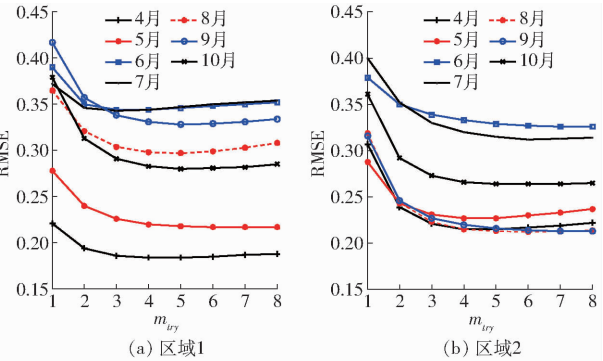


图 3 不同区域各参数取值对模型 RMSE 的影响
Fig. 3 RMSE of different parameter values on model of region 1 and region 2

3.2 偏差校正方法精度分析

采用 SLR、BC、RR 及 BRR 共 4 种方法对 RF 模型监测结果进行偏差校正,不同方法校正后监测结果与站点 SPEI3 的散点图如图 4 所示。由图 4 可知,区域 1 中 RF 监测值与站点 SPEI3 的 R^2 和 RMSE 分别为 0.719 和 0.534,经偏差校正后的 R^2 为 0.719 ~ 0.897, RMSE 为 0.335 ~ 0.533;区域 2 中

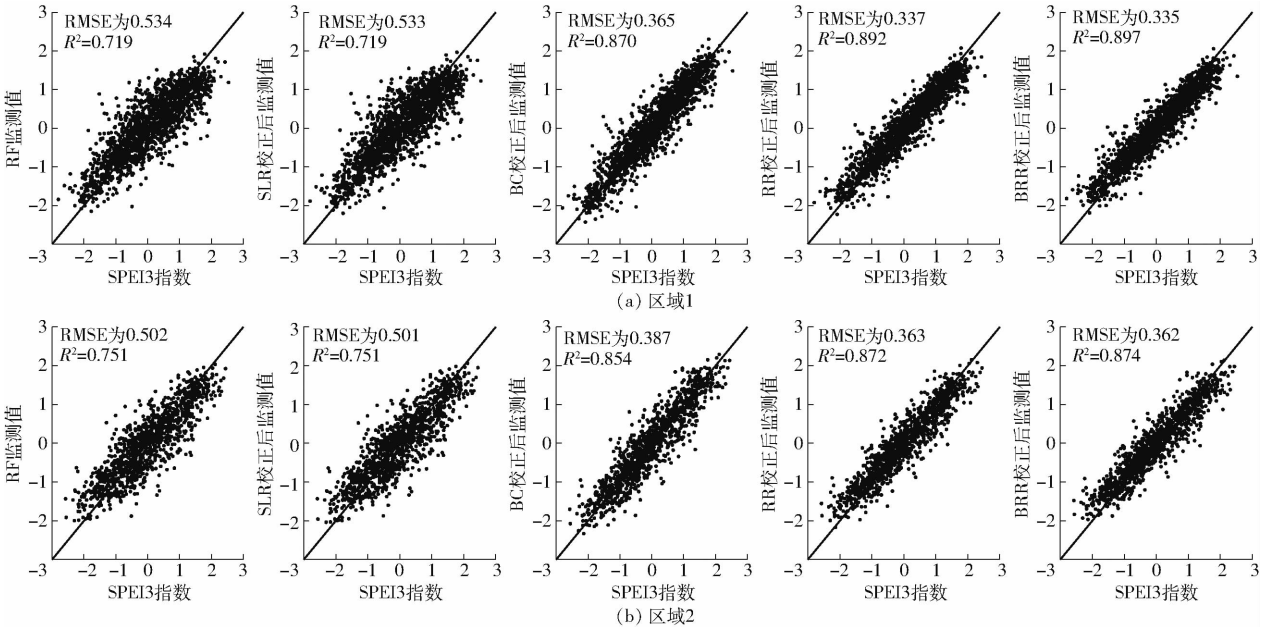


图 4 不同区域各校正方法的精度评估
Fig. 4 Scatter diagrams of drought index estimated by different bias-correcting methods and SPEI for three-month time scale

RF 监测值与站点 SPEI3 的 R^2 和 RMSE 分别为 0.751 和 0.502,经偏差校正后的 R^2 为 0.751 ~ 0.874, RMSE 为 0.362 ~ 0.501。由此可见,偏差校正后较校正前精度均有一定程度上的提高,其中 SLR 法的校正效果并不明显,而 BC、RR 以及 BRR 法对模型精度有显著提升。整体而言, BRR 方法的校正效果最优,校正后各区域的 R^2 分别为 0.897 和 0.874, RMSE 分别为 0.335 和 0.362。

不同校正方法对于干旱等级监测的准确率如表 4 所示。由表 4 可知, RF 对区域 1 中旱、重旱和极旱的监测准确率分别为 45.1%、40.2% 和 0,经校正后的中旱、重旱和极旱监测准确率分别为 44.4% ~ 69.3%、39.2% ~ 74.2% 和 0 ~ 50%; RF 对区域 2 中旱、重旱和极旱的监测准确率分别为 36.6%、18.6% 和 0,经校正后的中旱、重旱和极旱监测准确率分别为 39.2% ~ 64.1%、18.6% ~ 43.3%、0 ~ 33.3%。整体而言, SLR 法的校正效果较弱,而 BC、RR 及 BRR 法则有效提升了干旱等级监测的准确率,其中 BC 方法的校正效果最优,尤其是对于极端干旱监测具有显著的优势。综合考虑 R^2 、RMSE 以及干旱等级监测准确率的评估结果,最终采用 BC 方法构建 BRFD 模型。

3.3 干旱监测能力评估

为评估 BRFD 模型在农业干旱监测中的适用性,采用农业气象站相对土壤湿度及流域典型干旱事件对 BRFD 模型的监测能力进行验证。

采用 BRFD 模型计算指数、站点 SPEI3 分别与农业气象站 10、20 cm 的土壤相对湿度进行相关性分析,相关系数如表 5 所示。由表 5 可知,多数站点

表 4 各校正方法在不同区域对各干旱等级监测准确率
Tab.4 Accuracy of different drought conditions monitored
by all bias-correcting methods drought in two regions

分区	干旱级别及 样本数量	RF	校正方法			
			SLR	BC	RR	BRR
区域 1	中旱 ($n = 153$)	45.1	44.4	69.3	64.7	60.1
	重旱 ($n = 97$)	40.2	39.2	74.2	61.9	53.6
	极旱 ($n = 24$)	0	0	50.0	8.3	8.3
区域 2	中旱 ($n = 130$)	36.6	39.2	64.1	57.5	57.5
	重旱 ($n = 64$)	18.6	18.6	43.3	28.9	26.8
	极旱 ($n = 20$)	0	0	33.3	20.8	16.7

表 5 BRFDC 模型模拟干旱指数与不同深度土壤
相对湿度相关性分析
Tab.5 Correlation analysis between drought index simulated
by BRFDC and soil relative moisture in different depths

站点	10 cm		20 cm	
	SPEI3	BRFDC	SPEI3	BRFDC
驻马店	0.57	0.60	0.67	0.74
菏泽	0.44	0.47	—	—
莒县	0.41	0.44	0.38	0.39
临沂	0.54	0.52	0.61	0.57
西华	0.53	0.51	0.61	0.56
商丘	0.50	0.53	0.55	0.57
赣榆	0.57	0.53	0.60	0.51
亳州	0.38	0.40	0.44	0.48
淮安	0.39	0.46	—	0.36
宿县	0.55	0.56	0.56	0.63
盱眙	0.50	0.51	0.48	0.50
阜南	—	—	—	0.34
固始	0.59	0.60	0.54	0.55
寿县	0.61	0.62	0.69	0.66

注：“—”表示未通过 0.01 水平显著性检验。

的 SPEI3、BRFDC 模型计算指数与 10、20 cm 的土壤湿度显著相关 ($P < 0.01$), 二者都能有效表征土壤

墒情的变化。相比 SPEI3, 共有 10 个站点 10 cm 土壤相对湿度与 BRFDC 模型计算指数相关性更强, 此外, 9 个站点 20 cm 土壤相对湿度与 BRFDC 模型计算指数相关性更显著。由于土壤水分是农业干旱的决定性因素, 因此 BRFDC 模型计算的指数较 SPEI3 更适宜监测农业干旱。土壤水分变化受诸多因素影响, SPEI3 仅采用气象要素进行计算, 在表征土壤水变化过程上略有不足, 而 BRFDC 模型综合考虑降水、地表温度、植被长势及环境因素, 其计算指数更贴近于真实土壤水分的变化过程。在构建农业干旱监测模型时, 需综合考虑各种潜在影响因素。

根据文献[22]记载, 2001 年淮河流域发生了多年罕见的春、夏、秋连旱。河南省 3—5 月无有效降水, 8 月信阳地区创 50 多年降水最低纪录; 安徽省 5—9 月降水量较常年偏少 50% ~ 60%; 江苏省淮北地区 5—6 月降水量偏少, 7—8 月降水与往年基本持平, 旱情得到缓解, 9—11 月降水较同期偏少 81%, 旱情进一步发展; 山东省淮河流域汛期降水偏少, 且主要集中在 7—8 月。

采用反距离平方 (Inverse distance weighting, IDW) 法对站点 SPEI3 进行空间插值 (简称 IDW 模型), 并应用 RF 模型和 BRFDC 模型模拟了 2001 年 5—10 月淮河流域旱情空间分布, 结果如图 5 所示。由图 5 可知, 3 种方法对旱情发展过程的监测结果与真实旱情基本一致, 并且在空间上具有较好的一致性。IDW 模型通过简单的数学方式即可实现空间旱情监测, 能大致反映干旱的时空特征, 但受站点个数及插值方法限制, 无站点区域的旱情监测结果存在不确定性。例如 9 月流域西北部出现中度旱情, 由于该区域仅有开封站一个站点, 而未能监测到该区域的旱情; 10 月流域南部旱情减轻, 而该区域

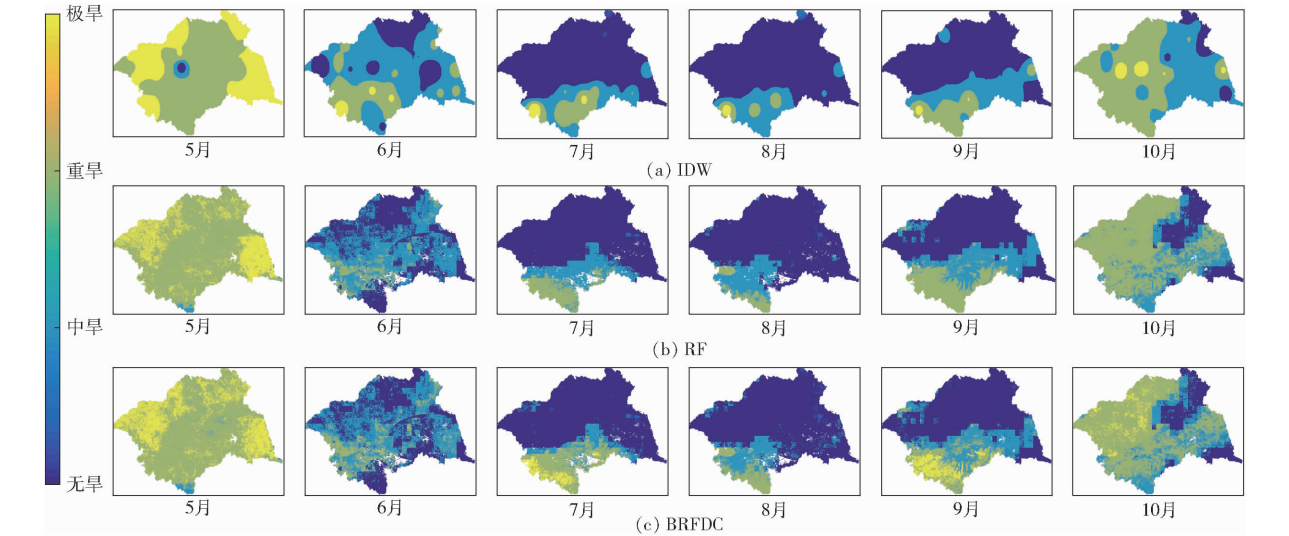


图 5 淮河流域 2001 年 5—10 月 IDW、RF、BRFDC 模型干旱监测结果
Fig. 5 Drought maps of Huaihe River Basin from May to October in 2001

仅霍山站一个站点,对旱情等级有所误判,而 RF 和 BRFD C 模型的监测结果则有效反映了这些区域的旱情。RF 模型以多源遥感数据为输入,考虑了降水、地表温度、植被状态、高程、AWC 等因素的空间异质性,其空间监测结果更加可靠。通常极端干旱将带来更大的损失,需要重视,但 RF 对 6、9 月流域南部以及 10 月流域中部极端干旱的监测效果并不理想,而 BRFD C 模型则准确识别了流域内的极端旱情。

整体上看,BRFD C 模型与土壤湿度具有较好的相关性,并且能有效反映淮河流域 2001 年 5—10 月的干旱事件的时空演变过程,对于极端旱情的监测结果也相对准确,表明该模型可用于淮河流域农业干旱监测。

3.4 讨论

农业干旱涉及降水、植被、温度以及环境因素影响,其过程非常复杂。结合多源遥感数据进行监测,能有效解释农业干旱的复杂性。相比原始 RF 而言,本文构建的 BRFD C 模型具有更准确的干旱监测能力,对极端干旱监测的准确性从 0 提升到 33.3%~50%,模型在淮河流域农业干旱监测中取得了较好的效果。BRFD C 模型依据免费的遥感数据即可实现区域农业干旱监测,能有效减少开展干旱监测业务的成本,为缺资料地区提供可靠的旱情空间分布信息。通常降水对干旱程度起决定性作

用,而本文仅采用了 TRMM 卫星的降水产品,其误差势必会对模型精度产生一定影响,提高降水产品的精度能显著改善监测结果的准确性。由于不同的卫星降水产品在不同区域、时间的精度有所差异^[23],融合多种卫星降水产品是提高干旱监测准确性的重要途径。此外,本文在构建模型时,考虑了降水、地表温度、植被长势、高程及 AWC 等影响因子,但农业干旱的复杂程度远不止如此。在后续的研究中,仍有一系列其他遥感数据需要考虑,以期提高模型的监测性能,例如 MODIS 蒸发数据^[24]、SMOS (Soil moisture and ocean salinity)土壤水分数据^[25]以及土地利用类型等数据。

4 结论

(1)BRR 法校正后,监测结果的 R^2 和 RMSE 分别为 0.897、0.874 和 0.335、0.362,优于其他方法;BC 法校正后,预测结果的 R^2 和 RMSE 分别为 0.870、0.854 和 0.365、0.387,同时显著提高了模型对极端干旱监测的准确率,达到 33.3%~50.0%。

(2)BRFD C 模型计算指数与站点实测 10、20 cm 的土壤相对湿度具有显著相关性,且 BRFD C 模型比 SPEI3 更适于农业干旱监测。

(3)相比 IDW 空间插值法及原始 RF 模型,BRFD C 模型模拟结果能更加准确地反映淮河流域 2001 年典型干旱事件的时空演变过程。

参 考 文 献

[1] ZHANG Qiang, GU Xihui, SINGH V P, et al. Spatiotemporal behavior of floods and droughts and their impacts on agriculture in China[J]. Global and Planetary Change, 2015, 131:63-72.

[2] LIANG Liang, ZHAO Shuhe, QIN Zhihao, et al. Drought change trend using MODIS TVDI and its relationship with climate factors in China from 2001 to 2010[J]. Journal of Integrative Agriculture, 2014, 13(7):1501-1508.

[3] 黄茹. 淮河流域旱涝急转事件演变及应对研究[D]. 北京:中国水利科学研究院,2015.

HUANG Ru. Research on evolution and countermeasures of droughts-floods abrupt alternation events in Huaihe River Basin [D]. Beijing: China Institute of Water Resource and Hydropower Research, 2015. (in Chinese)

[4] 赵维军, 马欢, 朱清科, 等. 陕北黄土区微地形土壤水分对降水特征的响应 [J/OL]. 农业机械学报, 2014, 45(5):118-124.

ZHAO Weijun, MA Huan, ZHU Qingke, et al. Response of soil moisture of microtopography on precipitation characteristics in the Loess Region of Northern Shaanxi [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(5): 118-124. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20140519&flag=1. DOI:10.6041/j.issn.1000-1298.2014.05.019. (in Chinese)

[5] 刘小刚,冷险险,孙光照,等. 基于 1961—2100 年 SPI 和 SPEI 的云南省干旱特征评估 [J/OL]. 农业机械学报,2018, 49(12):236-245,299.

LIU Xiaogang, LENG Xianxian, SUN Guangzhao, et al. Assessment of drought characteristics in Yunnan Province based on SPI and SPEI from 1961 to 2100 [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(12):236-245,299. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20181229&flag=1. DOI:10.6041/j.issn.1000-1298.2018.12.029. (in Chinese)

[6] 尹文杰, 张梦琳, 胡立堂. 柴达木盆地干旱时空变化特征 [J]. 干旱区研究, 2018, 35(2):387-394.

YIN Wenjie, ZHANG Menglin, HU Litang. Spatio temporal variation of drought in the Qaidam basin [J]. Arid Zone Research, 2018, 35(2):387-394. (in Chinese)

[7] WANG Hongshuo, ROGERS J C, MUNROE D K. Commonly used drought indices as indicators of soil moisture in China [J]. Journal of Hydrometeorology, 2015, 16(3):1397-1408.

- [8] WU Zhiyong, YUN Mao. Exploring spatiotemporal relationships among meteorological, agricultural, and hydrological droughts in Southwest China [J]. *Stochastic Environmental Research and Risk Assessment*, 2016, 30(3):1033–1044.
- [9] PARK S, IM J, JANG E, et al. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions [J]. *Agricultural and Forest Meteorology*, 2016, 216:157–169.
- [10] BENTO V A, GOUVEIA C M, DACAMARA C C, et al. A climatological assessment of drought impact on vegetation health index [J]. *Agricultural and Forest Meteorology*, 2018, 259:286–295.
- [11] GUO Hao, BAO Anming, LIU Tie, et al. Determining variable weights for an optimal scaled drought condition index (OSDCI): evaluation in Central Asia [J]. *Remote Sensing of Environment*, 2019, 231:111220.
- [12] RHEE J, IM J, CARBONE G J. Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data [J]. *Remote Sensing of Environment*, 2010, 114:2875–2887.
- [13] 张婧娴, 沈润平, 郭佳. 不同数据挖掘方法在综合干旱监测模型构建中的应用研究 [J]. *江西农业大学学报*, 2017, 39(5):1045–1056.
- ZHANG Jingxian, SHEN Runping, GUO Jia. A study of application of different data mining methods in integrated drought monitoring [J]. *Acta Agriculturae Universitatis Jiangxiensis*, 2017, 39(5):1045–1056. (in Chinese)
- [14] SHEN Ruiping, HUANG Anqi, LI Bolun, et al. Construction of a drought monitoring model using deep learning based on multi-source remote sensing data [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2019, 79:48–57.
- [15] 侯俊雄, 李琦, 朱亚杰, 等. 基于随机森林的 PM2.5 实时预报系统 [J]. *测绘科学*, 2017, 42(1):1–6.
- HOU Junxiong, LI Qi, ZHU Yajie, et al. Real-time forecasting system of PM2.5 concentration based on spark framework and random forest model [J]. *Science of Surveying and Mapping*, 2017, 42(1):1–6. (in Chinese)
- [16] 董婷, 任东, 邵攀, 等. 基于多源遥感数据和随机森林的综合旱情指标构建 [J/OL]. *农业机械学报*, 2019, 50(8):200–212.
- DONG Ting, REN Dong, SHAO Pan, et al. Construction of integrated drought condition index based on multi-sensor remote sensing and random forest [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(8):200–212. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190822&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2019.08.022. (in Chinese)
- [17] BREIMAN L. Random forests [J]. *Machine Learning*, 2001, 45(1):5–32.
- [18] SONG J. Bias corrections for random forest in regression using residual rotation [J]. *Journal of the Korean Statistical Society*, 2015, 44(2):321–326.
- [19] VICENTE S M, BEGUREIA S, LOPEZ-MORENO J I, et al. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index [J]. *Journal of Climate*, 2010, 23(7):1696–1718.
- [20] 张雷, 王琳琳, 张旭东, 等. 随机森林算法基本思想及其在生态学中的应用——以云南松分布模拟为例 [J]. *生态学报*, 2014, 34(3):650–659.
- ZHANG Lei, WANG Linlin, ZHANG Xudong, et al. The basic principle of random forest and its applications in ecology: a case study of *Pinus yunnanensis* [J]. *Acta Ecologica Sinica*, 2014, 34(3):650–659. (in Chinese)
- [21] ZHANG Guoyi, LU Yan. Bias-corrected random forests in regression [J]. *Journal of Applied Statistics*, 2012, 39(1):151–160.
- [22] 水利部淮河水利委员会. 治淮汇刊年鉴 [J]. 蚌埠:《治淮汇刊》编辑部, 2001–2014.
- [23] 廖荣伟, 张冬斌, 沈艳. 6 种卫星降水产品在中国区域的精度特征评估 [J]. *气象*, 2015(8):970–979.
- LIAO Rongwei, ZHANG Dongbin, SHEN Yan. Validation of six satellite-derived rainfall estimates over China [J]. *Meteorological Monthly*, 2015(8):970–979. (in Chinese)
- [24] MU Qiaozhen, HEINSCH F A, ZHAO Maosheng, et al. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data [J]. *Remote Sensing of Environment*, 2007, 111(4):519–536.
- [25] DALL'AMICO J T, SCHLENZ F, LOEW A, et al. The SMOS validation campaign 2010 in the upper Danube catchment: a data set for studies of soil moisture, brightness temperature, and their spatial variability over a heterogeneous land surface [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(1):364–377.