

邢雅洁¹ 沈润平¹ 黄安奇¹ 梁宇靖¹ 王云宇¹ 谢昭颖¹ 师春香² 孙帅²

基于 MODIS 和 CLDAS 的综合干旱监测模型研究

摘要

传统的干旱监测指数主要考虑单一影响因子,往往无法全面综合反映干旱状况。基于 MODIS 数据和 CLDAS 数据,选取多个影响因子和能够直接反映干旱程度的干旱指数作为自变量,以综合气象干旱指数(CI)为因变量,通过梯度提升机(GBM)机器学习算法建立日尺度综合干旱监测模型,并以 2015—2018 年华北地区干旱为例进行了研究。结果表明模型监测结果与站点 CI 计算值具有显著的相关性,训练集和测试集决定系数分别达到 0.945 和 0.655,均方根误差(RMSE)分别为 0.033 和 0.082,综合干旱监测模型具有较高的精度。且模型监测与 CI 监测各月等级一致率均在 65% 以上,并与标准化降水蒸散指数(SPEI)和土壤相对湿度(RSM)相关系数分别为 0.68 和 0.60,能较好地反映气象干旱和农业干旱状况。典型干旱情况监测表明,综合干旱监测模型综合考虑多种干旱影响因素,能较准确地识别出干旱的发生,表征综合干旱发生状况。

关键词

CLDAS; 综合干旱监测; 梯度提升机; MODIS

中图分类号 TP181; S423; P426.616

文献标志码 A

收稿日期 2023-02-28

资助项目 国家重点研发计划(2018YFC1506602); 国家自然科学基金重点项目(91437220)

作者简介

邢雅洁,女,硕士生,主要从事陆面过程遥感研究.xjxyj1113@163.com

沈润平(通信作者),男,博士,教授,主要从事陆面过程模拟与遥感研究.rpshen@nuist.edu.cn

1 南京信息工程大学 地理科学学院,南京,210044

2 国家气象信息中心,北京,100081

0 引言

干旱是指水分收支或供求不平衡所造成的水分短缺现象^[1]。据统计,气象灾害引起的损失占各类自然灾害的 85%,而干旱又占其损失的 50%^[2]。独特的地理环境与气候特点导致我国易受干旱影响,且易造成损失^[3]。干旱指数作为可以表征干旱发生的重要指标,能够将干旱的严重性进行等级区分和量化。目前为止,已经发展了多种干旱监测指数,如基于气象站点数据计算的标准化降水指数(Standardized Precipitation Index, SPI)、标准化降水蒸散指数(Standardized Precipitation Evapotranspiration Index, SPEI)和帕默尔干旱指数(Palmer Drought Severity Index, PDSI)^[4-6]等,但由于站点数量和位置分布不均,干旱监测缺乏空间连续性。随着遥感技术的发展,不同类型传感器可获取大范围近实时的降水、植被和土壤湿度等信息,为区域干旱监测提供了方法,提高了干旱监测的空间连续性。一些遥感干旱指数,例如温度状况指数(Temperature Condition Index, TCI)、降水状况指数(Precipitation Condition Index, PCI)、植被状况指数(Vegetation Condition Index, VCI)、温度植被干旱指数(Temperature Vegetation Dryness Index, TVDI)和条件温度植被指数(Vegetation Temperature Condition Index, TVCI)^[7-9]等广泛应用于大范围干旱监测,但遥感数据往往因其时间分辨率低影响干旱监测效果。

由于干旱与降水、地形、潜在蒸散发和土壤湿度等多种变量有着密切的关系,导致干旱成因复杂。虽然单一影响因子或干旱监测指数在有关行业和领域有着较好的应用,但会出现对于干旱因子的影响考虑不足或受限于应用条件,往往无法充分反映区域内的综合干旱发生状况。为此,许多学者利用多种数据及多种方法,开展了综合干旱监测研究^[10-12]。目前,综合干旱指数主要通过权重组合、联合分布和机器学习^[13]等 3 类方法进行构建。权重组合法由于指数的选择以及权重的分配限制了指数的应用。联合分布方法在需要综合的干旱因子或指数较多时,不同因子或指数之间关系的复杂性会加大多元分布估计的难度,影响监测效果。机器学习方法因其具有处理非线性问题的能力,能够对多种干旱因子进行信息挖掘,在综合干旱监测上被广泛应用。然而,由于众多不同来源的遥感数据精度不同,单纯依靠遥感数据构建的综合干旱监测指数对于空间连续的干旱监测还存在不足。考虑到综合干旱监测在日常农业生产和抢险救灾等方面的重要性,需

要更合适的方法和数据以构建日尺度的综合干旱监测模型。

中国气象局陆面数据同化系统(CMA Land Data Assimilation System, CLDAS)是我国自主建立的陆面数据同化系统,其通过融合多种气象站点观测、卫星遥感观测和再分析资料,输出多种高质量的气象及陆面数据,为日尺度气象监测提供重要的数据支撑^[14-16]。然而,当前缺少将CLDAS数据应用于综合干旱监测的研究,尚不清楚CLDAS应用于综合干旱监测的效果。因此,本文通过引入2015—2018年连续日尺度CLDAS数据,再辅以相应的中等分辨率成像光谱仪(Moderate-resolution Imaging Spectroradiometer, MODIS)数据,通过梯度提升机(Gradient Boosting Machine, GBM)方法构建日尺度综合干旱监测模型,统筹考虑与干旱发生的多种相关因子,以获得区域日尺度下干旱发生的整体状况,通过评估模型对于日尺度综合干旱的监测效果,以期获得更加精细合理的干旱监测结果。

1 研究区与研究数据

1.1 研究区

选取华北地区为研究区(图1)^[17],经纬度范围为 $110^{\circ}\sim 120^{\circ}\text{E}$, $34^{\circ}\sim 43^{\circ}\text{N}$,包括山西省、河北省、北京市、天津市以及内蒙古自治区中部部分地区。研究区南部临近黄河,北部为内蒙古高原,东部有海河平原,西部为吕梁山脉,地势总体西高东低,下垫面情况复杂。华北地区是我国重要的农业主产区和粮食生产基地,主要种植冬小麦和玉米,研究区域大部分属暖温带半湿润半干旱季风气候区,年降雨量大约

为 $400\sim 500\text{ mm}$,雨季主要集中在夏季,雨热同期,干湿期分隔明显。在全球气候变化的背景下,该区域年际降水分配不均,气候多变,干旱灾害事件频繁发生^[18]。

1.2 研究数据

1.2.1 MODIS 数据

反射率数据来自于美国国家航空航天局(<https://ladsweb.modaps.eosdis.nasa.gov>)的Terra MODIS表面反射率产品MOD09GA,该产品为表面波谱反射估计,数据经过大气和气溶胶的校正,空间分辨率为 500 m ,时间分辨率为 1 d ,提供 $1\sim 7$ 波段每日栅格化的数据。本文所用到的波段为红波段($620\sim 670\text{ nm}$)和近红外波段($841\sim 876\text{ nm}$)反射率数据,研究时段为2015—2018年。将地表反射率数据进行投影一致性转换、研究区的矢量边界裁剪等预处理后,通过双线性内插法重采样至 $0.0625^{\circ}\times 0.0625^{\circ}$ 分辨率,使其与CLDAS空间分辨率一致。

1.2.2 CLDAS 数据

CLDAS大气驱动场产品V2.0覆盖亚洲区域($60^{\circ}\sim 160^{\circ}\text{E}$, $0^{\circ}\sim 65^{\circ}\text{N}$),空间分辨率为 $0.0625^{\circ}\times 0.0625^{\circ}$,时间分辨率为逐小时,包括 2 m 气温、 2 m 比湿、 10 m 风速、地面气压、降水和短波辐射等6个要素的等经纬网格融合分析产品。CLDAS土壤体积含水量分析产品V2.0覆盖区域和时空分辨率与大气驱动场产品一致,垂直分为5层,分别为 $[0\text{ cm}, 5\text{ cm}]$ 、 $[0\text{ cm}, 10\text{ cm}]$ 、 $[10\text{ cm}, 40\text{ cm}]$ 、 $[40\text{ cm}, 100\text{ cm}]$ 、 $[100\text{ cm}, 200\text{ cm}]$ 。CLDAS地表温度数据分析产品V2.0与CLDAS大气驱动和土壤体积含水量产品覆盖区域与时空分辨率相同^[15,19]。本研究使用

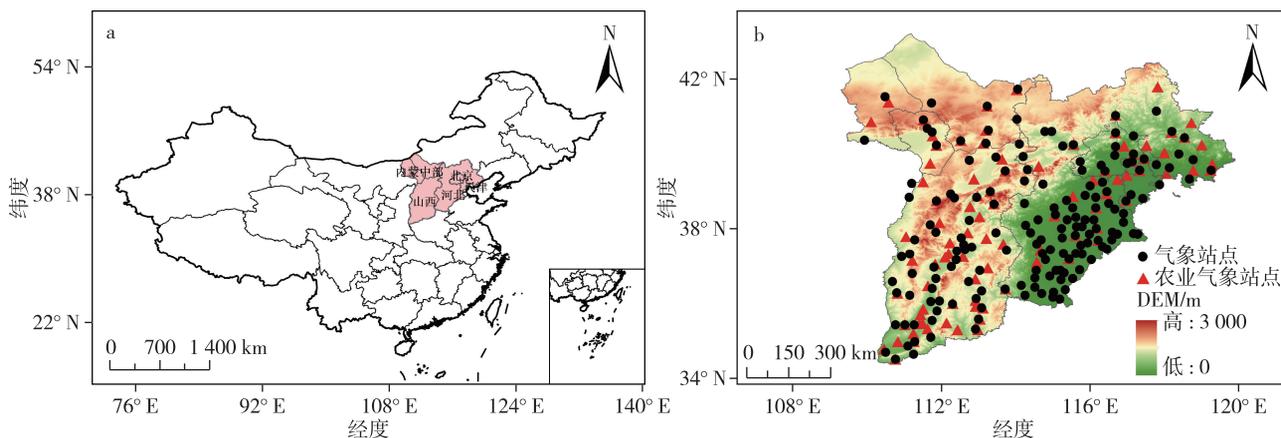


图1 研究区概况 a.华北地区区位;b.华北地区气象站点与农业气象站点分布

Fig. 1 Overview of the study area, (a) location map of North China, and (b) location of meteorological stations and agricultural meteorological stations in North China

2015—2018 年大气驱动产品中的地面气压、短波辐射、2 m 气温、降水要素,以及 0~10 cm 土壤体积含水量和地表温度数据分析产品,并将逐小时数据通过求平均或者加和合成为日尺度数据。

1.2.3 气象站点数据

气象站资料来源于国家气象信息中心 2015—2018 年中国地面气象站逐小时的气温和降水数据 (<http://data.cma.cn/>),选取研究区域内 158 个气象站点和 68 个农业气象站点,将逐小时降水加和、逐小时的气温求取平均值,分别得到研究区逐日的降水数据和气温数据。将通过气象站点数据计算得出的 CI 监测数据,75%作为训练集,25%作为测试集,将其作为综合干旱监测模型的建模和验证输入数据。

1.2.4 其他辅助数据

1) DEM 数据. 数字高程模型 (Digital Elevation Model, DEM) 为美国“奋进号”航天飞机的雷达地形测绘 (Shuttle Radar Topography Mission, SRTM) 获得的 DEM 数据,来源于中国科学院资源环境科学数据中心 (<http://www.resdc.cn>),范围为 60°N ~ 60°S. 采用基于 4.1 版本的 DEM 数据,利用新的插值算法得到,空间分辨率为 1 km. 将高程数据进行投影一致性转换、研究区的矢量边界裁剪等预处理后,通过双线性内插法重采样至 0.062 5°×0.062 5°分辨率,使其与 CLDAS 空间分辨率一致。

2) 土壤水文参数数据. 土壤水分亏缺指数 (Soil Water Deficit Index, SWDI) 的计算需要利用土壤水文参数数据得出. 本文所用的土壤水文参数数据来源于北京师范大学发布的面向陆面过程模型的土壤水文数据集^[20],分辨率为 30 角秒. 利用此数据的 -33 kPa 和 -1 500 kPa 土壤水势数据,作为田间持水量数据 (Field Capacity, FC) 和土壤凋萎湿度 (Wilting Point, WP). 土壤数据垂直剖面分为 7 层,分别为 [0 m, 0.045 m)、[0.045 m, 0.091 m)、[0.091 m, 0.166 m)、[0.166 m, 0.289 m)、[0.289 m, 0.493 m)、[0.493 m, 0.829 m)、[0.829 m, 1.383 m], 最深为 1.383 m. 研究首先利用双线性内插法将其重采样至 0.062 5°×0.062 5°,然后再插值到研究需要的 0~10 cm 土层深度,以田间持水量 (FC) 数据为例,具体计算方法如下:

$$C_F = \frac{C_{F,1} \times Z_1 + C_{F,2} \times Z_2 + C_{F,3} \times (Z - Z_1 - Z_2)}{Z} \quad (1)$$

式中: C_F 为重新插值后 (0~10 cm) 土壤深度的田间持水量 ($\text{cm}^3 \cdot \text{cm}^{-3}$); $C_{F,i}$ 为 0~10 cm 内第 i 层田间持水量平均 ($\text{cm}^3 \cdot \text{cm}^{-3}$); Z_i 为第 i 层田间持水量的土层深度 (cm); Z 为所需的田间持水量土层深度即 10 cm。

2 研究方法

2.1 综合干旱监测模型的构建

干旱受到降水、地表温度 (或气温)、土壤、地表植被状况和地形等关键因素的综合影响. 干旱主要是由于大气水分环流异常,导致降水时空分布不均,从而出现降水不足造成的. 本研究利用 CLDAS 降水数据 (PRE) 反映研究区的降水异常状况,同时通过降水与蒸散的差值 (AWD) 反映地表水分的亏缺状况. 土壤湿度则会受到降水不足和土壤性质的影响,土壤水分亏缺引起干旱. 本研究选用 CLDAS 土壤湿度 (SM) 反映土壤水分不足的状况^[21],同时利用土壤水分亏缺指数 (SWDI) 反映不同土壤类型下土壤湿度对干旱的响应. 当干旱发生时,植物受到水分胁迫影响其生长过程,通过 CLDAS 地表温度 (Land Surface Temperature, LST) 反映地表温度对植被生长的制约情况,植被供水指数 (Vegetation Supplication Water Index, VSWI) 反映植被生长受到水分胁迫的严重程度. 海拔 (DEM) 以及坡度 (slope) 所导致的下垫面状况差异对于区域干旱的发展同样有重要的影响^[22]. 为此,利用梯度提升机 (GBM), 将能够综合反映降水量与潜在蒸散发的综合气象干旱指数值 (CI) 作为因变量,能够反映干旱状况的多种因子和指数作为自变量,构建以 $CI = f(\text{PRE}, \text{SM}, \text{LST}, \text{AWD}, \text{SWDI}, \text{VSWI}, \text{DEM}, \text{slope})$ 为形式的综合干旱监测模型,技术路线如图 2 所示。

2.1.1 干旱指数

1) 大气水分亏缺指数 (Atmospheric Water Deficit, AWD)

降水与蒸散发都是影响干旱发生的重要因子,尤其是降水量的减少,往往会直接引起气象干旱的发生. 大气水分亏缺指数 (AWD)^[23] 为降水与潜在蒸散发的差值,选用其作为构建综合干旱指数的因子来反映大气的水分状况,具体计算方法^[24-25] 如下:

$$i_{\text{AWD},i} = P_i - m_{\text{PET},i} \quad (2)$$

式中: $i_{\text{AWD},i}$ 为某天的大气水分亏缺量; P_i 为降雨量; $m_{\text{PET},i}$ 为潜在蒸散量. 当 $\text{AWD} < 0$ 时,说明大气出现了水分亏缺的状况。

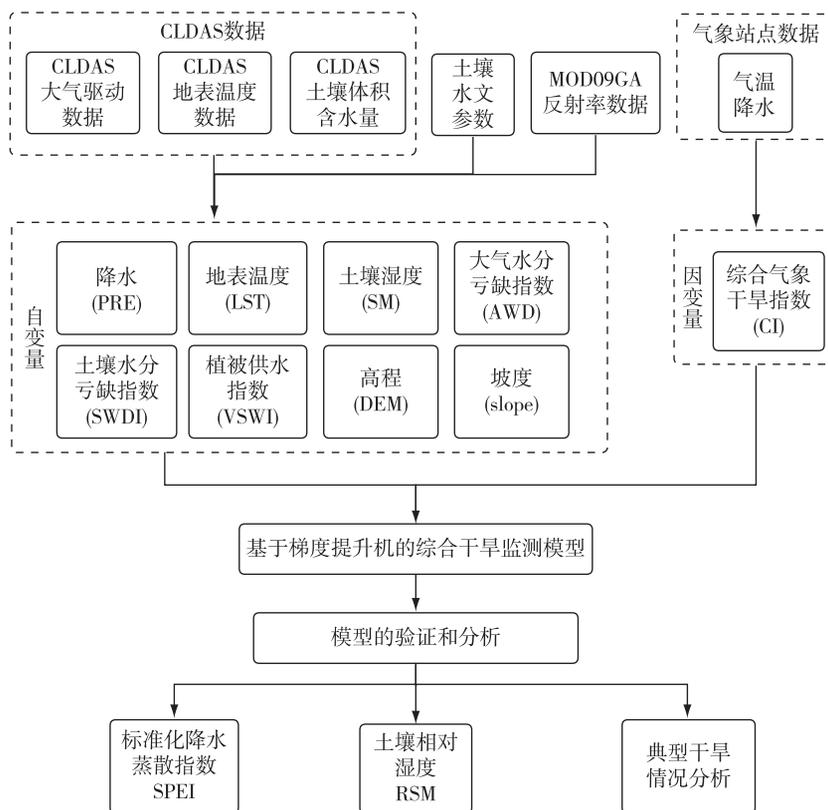


图 2 干旱监测模型构建流程

Fig. 2 Flowchart of the integrated drought monitoring model

2) 土壤水分亏缺指数 (Soil Water Deficit Index, SWDI)

土壤干湿状况是反映农业干旱发生的重要指标,而土壤水分亏缺指数 (SWDI)^[26]能够表征土壤水分缺失的严重程度,选取其作为综合干旱指数构建的一部分,来表征土壤状况,具体计算方法如下:

$$i_{SWD} = \frac{\theta - \theta_{FC}}{\theta_{AWC}} \quad (3)$$

式中: θ 是土壤含水量; θ_{FC} 表示田间持水量; θ_{AWC} 表示土壤有效含水量; θ_{WP} 表示土壤凋萎湿度; θ_{AWC} 是 θ_{FC} 与 θ_{WP} 的差值,即 $\theta_{AWC} = \theta_{FC} - \theta_{WP}$.

3) 植被供水指数 (Vegetation Supplication Water Index, VSWI)

植被供水指数 (VSWI) 通过植被生长受到制约的状况反映干旱程度,将其作为构建综合干旱指数的因子^[27],来表征干旱发生与植被变化的响应关系,VSWI 值越小,说明干旱发生状况越严重,VSWI 值越大,说明干旱发生程度轻.具体计算方法如下:

$$i_{VSW} = \frac{i_{NDV}}{T_{LS}} \quad (4)$$

式中: T_{LS} 为地表温度; i_{NDV} 为归一化植被指数.日尺度

NDVI 的计算使用 MOD09GA 产品的红波段 (620 ~ 670 nm) 和近红外波段 (841 ~ 876 nm),计算公式如下:

$$i_{NDV} = \frac{(R_{NI} - R)}{(R_{NI} + R)} \quad (5)$$

式中: R_{NI} 表示近红外波段反射率; R 表示红波段反射率.

4) 综合气象干旱指数 (CI)

综合气象干旱指数 (CI) 广泛应用于气象干旱监测,在我国具有广泛的适用性.综合气象干旱指数是通过近 30 天和近 60 天的标准化降水指数 (SPI),以及近 30 天的相对湿润度指数 (MI) 综合计算得到的,其具体计算公式^[28]如下:

$$I_C = aI_{SP,30} + bI_{SP,60} + cI_{M,30} \quad (6)$$

式中: a 为近 30 天标准化降水系数,取 0.4; b 为近 60 天标准化降水系数,取 0.4; c 为近 30 天相对湿润系数,取 0.8; $I_{SP,30}$ 为近 30 天的标准化降水指数值; $I_{SP,60}$ 为近 60 天的标准化降水指数值; $I_{M,30}$ 为近 30 天的相对湿润度指数.CI 指数的划分标准^[28]如表 1 所示.

表 1 综合气象干旱指数(CI)干旱等级分类标准
Table 1 Classification scales of drought with respect to CI

等级	类型	CI 值
1	无旱	$-0.6 < CI$
2	轻旱	$-1.2 < CI \leq -0.6$
3	中旱	$-1.8 < CI \leq -1.2$
4	重旱	$-2.4 < CI \leq -1.8$
5	特旱	$CI \leq -2.4$

2.1.2 梯度提升机(GBM)

在机器学习中,提升方法(Boosting)是一种通过组合一群复杂程度和训练成本程度低、不容易过拟合的弱分类器(Weak Learner),建立 N 个模型,并尝试在每次分类中都将在上一次分错的数据权重提高一点再进行分类,来获得一个强分类器(Strong Learner)的方法^[21,29].梯度提升机(GBM)属于 Boosting 的一种,是主要用于解决分类和回归问题的机器学习算法,通过集成多个弱学习器,每个弱学习器拟合先前累加模型的损失函数的负梯度,使加上该弱学习器后的累积模型损失往负梯度的方向减少,通过优化损失函数来生成最终强预测模型.梯度提升机机器学习算法对输入变量的相关性没有严格要求,其每轮训练目标是使预测值与真实值的平方误差最小.在利用梯度提升机进行模型构建时,考虑到可能出现的过拟合现象,所以在进行模型构建时采用交叉验证的方法来尽可能避免模型出现过拟合.

本研究首先将基于 CLDAS 与 MODIS 数据日尺度的干旱因子和指数值提取至站点并进行归一化后,与基于站点的日尺度 CI 指数共同作为输入数据,建立梯度提升机回归模型.之后,将基于 CLDAS 和 MODIS 数据的干旱因子和指数输入建立的回归模型中,生成对研究区内综合干旱监测模型的预测值,用于干旱监测.

2.2 模型的验证及相关性分析

参照综合气象干旱指数(CI)的旱情等级划分标准,将模型预测得到的干旱监测指数值分为特旱、重旱、中旱、轻旱和无旱 5 个等级,据此,将 CI 指数与模型监测的干旱等级进行对比,计算总体一致率,具体计算方法如下:

$$\text{总体一致率} = \frac{\text{等级一致率的站点数}}{\text{总站点数}} \times 100\%. \quad (7)$$

为了了解构建的综合干旱监测模型的监测效果,研究将其与标准化降水蒸散指数(Standardized

Precipitation Evapotranspiration Index, SPEI) 和土壤相对湿度(Relative Soil Moisture, RSM)进行相关性分析,具体计算方法如下:

$$R_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (8)$$

式中: x 和 y 分别为进行相关性分析的两个因子; R_{xy} 为因子相关系数; x_i 和 y_i 分别为两个因子的第 i 个值; \bar{x} 和 \bar{y} 分别为两个因子的平均值.

3 结果与分析

3.1 综合干旱监测模型验证

表 2 所示 75%训练集综合干旱监测值与实测综合气象干旱指数(CI)间的决定系数(R^2)达到 0.945,25%测试集综合干旱监测值与实测综合气象干旱指数(CI)的 R^2 为 0.655,两者有着较好的相关性,测试集的均方根误差(RMSE)和平均绝对误差(MAE)分别为 0.082、0.062,误差较小,说明所构建的综合干旱监测模型具有较高的精度.

表 2 综合干旱监测模型精度分析(2015—2018 年)

Table 2 Precision analysis of the proposed model for monitoring period of 2015–2018

数据	决定系数 (R^2)	均方根误差 (RMSE)	平均绝对误差 (MAE)
训练集(75%)	0.945	0.033	0.021
测试集(25%)	0.655	0.082	0.062

综合干旱监测模型与 CI 指数监测的干旱等级一致率比较结果表明(表 3),模型监测值与 CI 值的等级一致率达到 65.1%~89.4%,7、8、9 月等级一致率较高,在 82% 以上,其余月份等级一致率也都在 65% 以上,总体上两者监测结果表现出较好的一致率.

表 3 综合干旱监测模型结果与 CI 指数监测等级一致率

Table 3 Consistency rate between model monitored CI and calculated CI

指标	4 月	5 月	6 月	7 月	8 月	9 月	10 月
等级一致率/%	65.1	66.5	72.3	89.4	85.3	82.9	77.5

3.2 与标准化降水蒸散指数(SPEI)的相关性

标准化降水蒸散指数(SPEI)是重要的气象干旱监测指数,可采用它来验证构建的综合干旱监测模型对气象干旱的监测能力^[30].2015—2018 年日尺

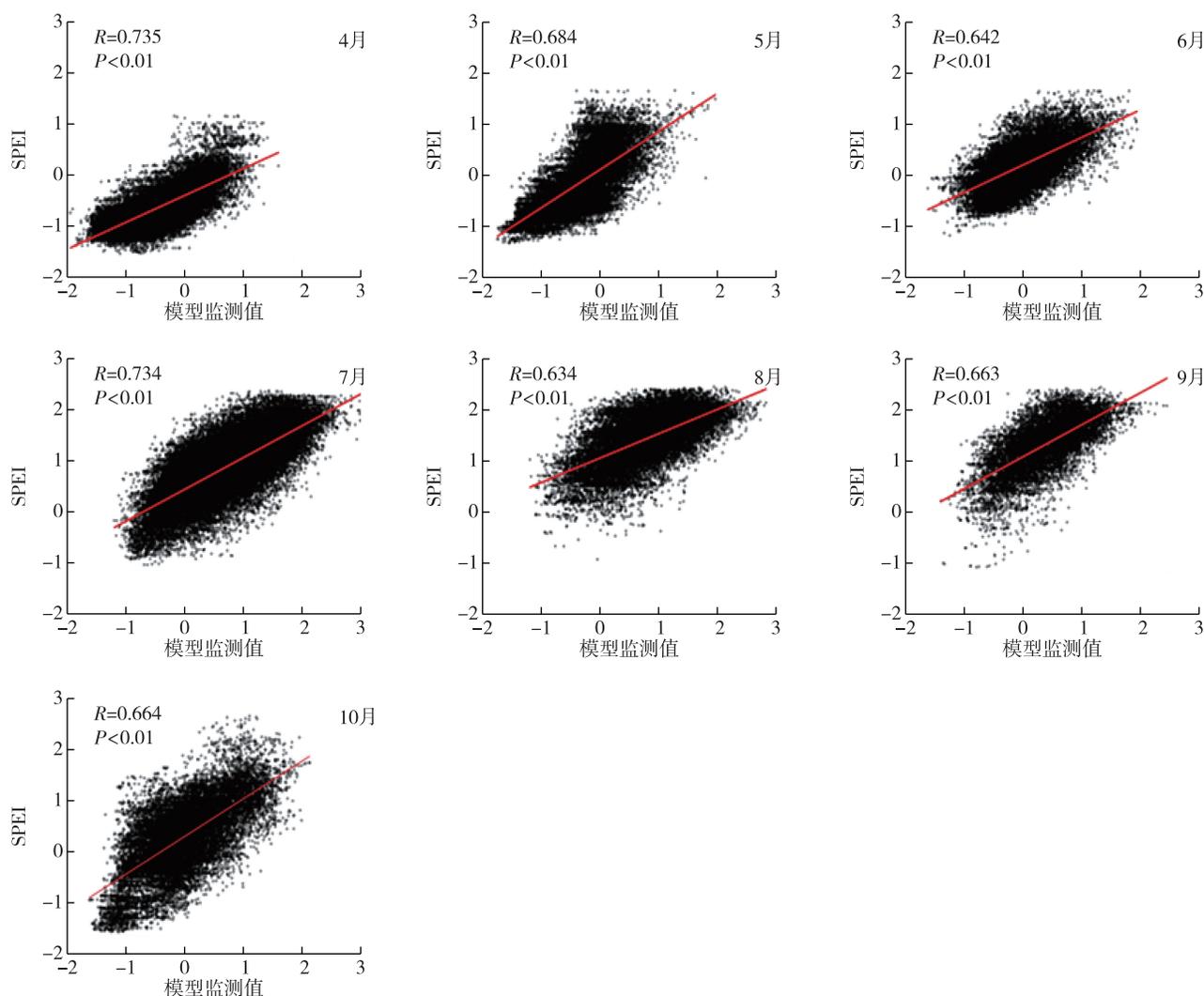


图3 模型监测与SPEI相关性散点图

Fig. 3 Scatter diagrams of correlation between model monitored SPEI and observed SPEI

度 SPEI 指数与综合干旱监测模型监测结果的相关分析表明,两者具有较好的相关性(图 3),各月相关系数在 0.634~0.735 之间,且各月份间相关系数相差较小,其中 4 月和 7 月相对较高,分别为 0.735 和 0.734,显著性检验 p 值均小于 0.01,相关性结果都达到极显著水平,说明综合干旱监测模型能够较好地反映气象干旱的发生,具有较好的气象干旱监测能力。

3.3 与土壤相对湿度(RSM)指标的相关性

土壤相对湿度是土壤含水量与田间持水量的百分比,是反映土壤水分平衡的重要指标,通常用来反映干旱对于农业的影响^[31]。研究分析了综合干旱监测模型与 0~10 cm 深度土壤相对湿度的相关性(图 4),4 月和 5 月的相关系数分别为 0.574 和 0.51,这可能与春季华北地区春小麦播种时农田灌溉,对土

壤湿度状况产生影响有关,其余月份相关系数均大于 0.6,显著性检验 p 值均小于 0.01,相关性结果都达到极显著水平,模型能够较好地反映土壤相对湿度变化对于干旱的产生的影响。

3.4 典型干旱情况监测分析

华北地区春季升温较快,当蒸发量较大,且尚未进入雨季时,极易导致春旱发生。我国 2015—2018 年均有春旱发生,尤其 2016 年北方冬麦区春旱^[32],3 月中旬,山西、河北等省部分地区开始发生旱情,4 月上旬,旱情快速发展,作物因旱大面积受损。本文选取 4 月作为研究月份,对综合干旱监测模型与利用克里金进行 CI 插值的监测结果进行了比较,计算出每月发生干旱的天数。

图 5a 结果显示,2015 年 4 月华北地区干旱天数较多,如河北省西北部、山西中部和北部,以及内蒙

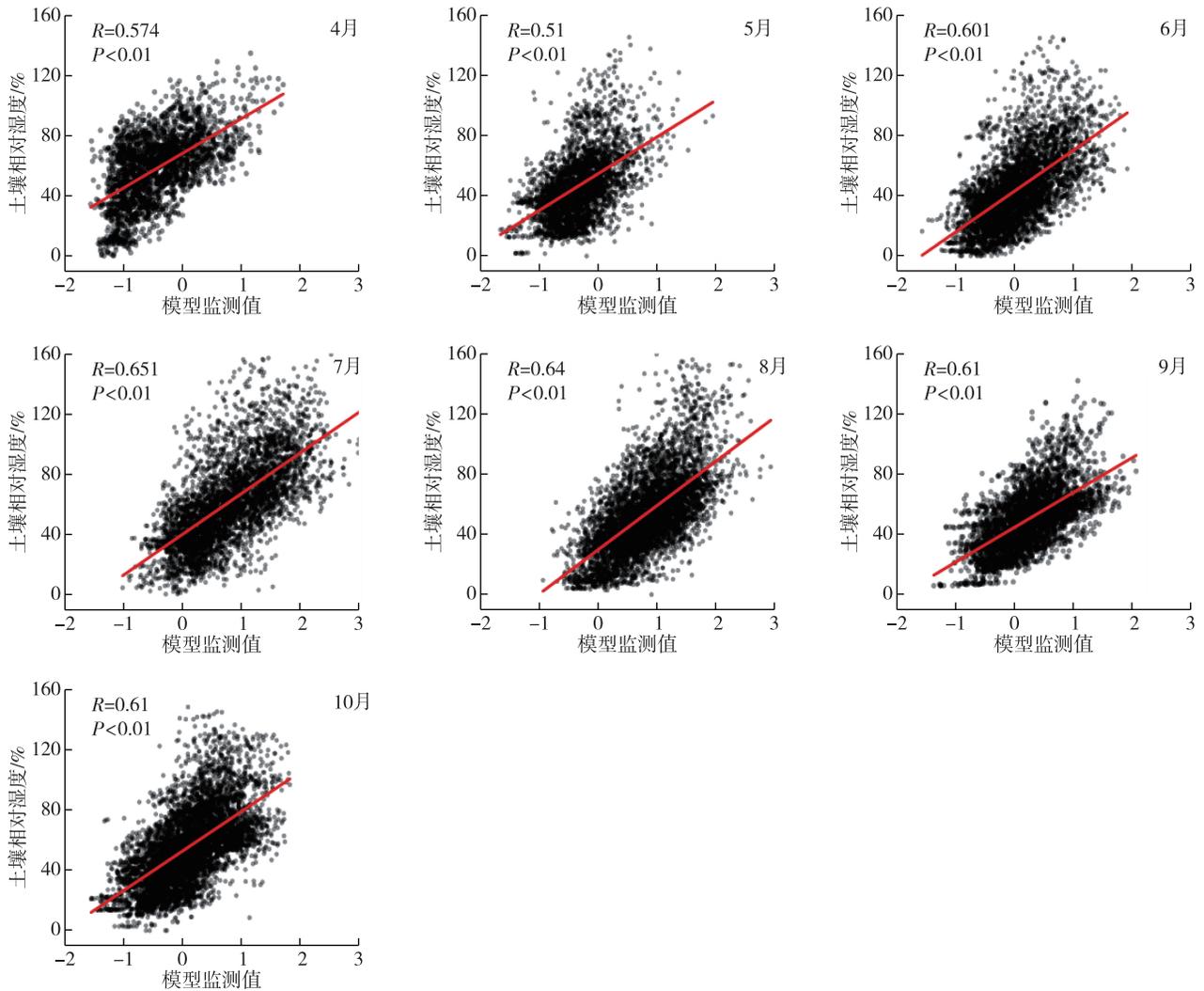


图4 模型监测与土壤相对湿度相关性散点图

Fig. 4 Scatter diagrams of correlation between model monitored RSM and measured RSM

古中部地区,干旱发生都超过了 15 d,而山西南部、河北南部干旱天数相对较少,低于 10 d.2016 年 4 月华北地区干旱发生如图 5b 所示,除山西南部部分地区外,华北大部干旱发生天数大于 15 d.如图 5c 所示,2017 华北地区 4 月干旱发生较多地区主要在河北中北部以及内蒙古中部地区,干旱发生天数约在 20 d 以上,而河北南部、山西中南部干旱发生相对较少.如图 5d 所示,2018 年 4 月干旱发生较多的地区主要集中在内蒙古中部、山西大部、河北少部分地区,天数约在 15 d 以上.

此外,相较于模型监测结果,CI 指数通过插值由区域尺度转化为空间尺度时,较少考虑地表温度、地表土壤或植被等下垫面分布状况对干旱发生的影响.CI 监测结果(图 5f)显示山西中部地区 2016 年 4 月干旱发生小于 10 d,与综合干旱监测模型结果相

比较低(图 5b),可能是由于此处的吕梁山脉黄土广布同时植被覆盖度较低,产生的降水难以下渗或被截留,水土流失严重,所以干旱发生状况较为严重,CI 监测结果低估.2018 年 4 月模型监测结果(图 5d)显示河北西北部与中南部干旱状况明显,这是由于河北地形的差异,河北西北部山区高原与谷地错落分布,以及林地广布使得下垫面状况较为复杂,而中南大部为平原地区,地形平坦且多为粮食种植区.但由于 CI 监测对于下垫面差异的因素考虑较少(图 5h),所以 CI 监测结果显示河北 2018 年 4 月干旱天数差异较少,未能充分反映河北省西北部和中南部干旱分布差异的细节.

本文使用克里金插值将由站点计算的 CI 值从区域尺度转换为空间尺度(图 5e—h)的结果表明,虽然站点插值结果基本反映了干旱发展过程,然而

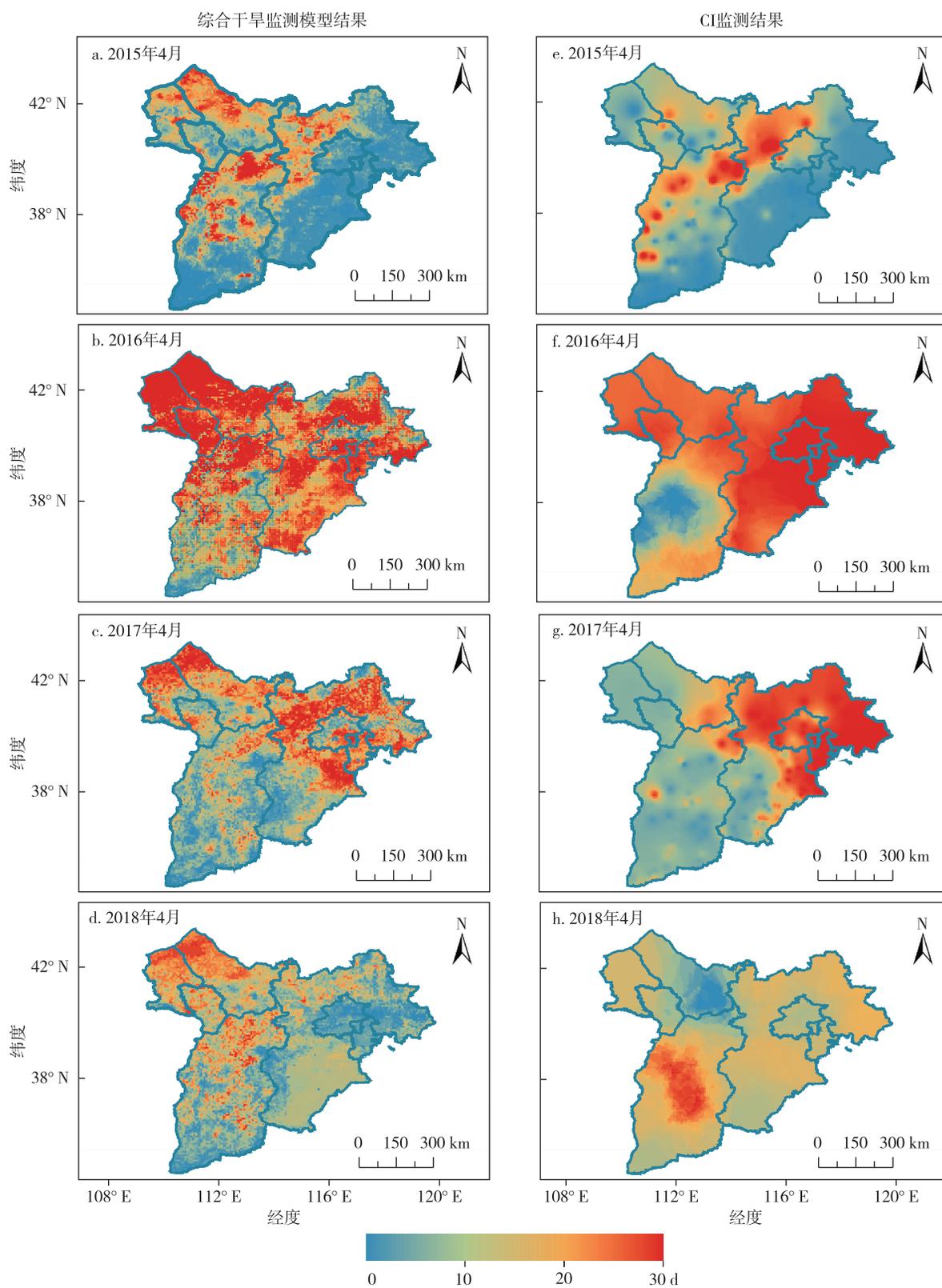


图5 华北地区2015—2018年4月干旱天数空间分布

a—d.为综合干旱监测模型结果;e—h.为CI监测结果

Fig. 5 Spatial distribution of drought days in April in North China during 2015–2018, (a)–(d) for results of integrated drought monitoring model, and (e)–(h) for results of CI monitoring

由于站点数量有限同时站点分布不均,结果无法有效反映部分干旱状况,所以其结果在没有气象站点分布的地方缺乏可信度.例如内蒙古中部地区 2015—2018 年存在较严重的干旱状况,干旱发生天数在 20 d 以上,但 CI 站点监测结果则显示此处干旱状况偏轻,干旱严重程度低于 CI 监测结果,这是由于插值方法为单纯的数学过程,无法顾及到站点数量较少或位置分布不均所导致的监测结果不准确的问题,从而影响监测效果.

4 结论与讨论

本文通过综合考虑能够影响或反映干旱的影响因子,充分利用 MODIS 卫星遥感资料和 CLDAS 数据的优势,选取相应降水、地表温度、土壤湿度、大气水分亏缺指数、植被供水指数、土壤水分亏缺指数、高程和坡度等监测指标作为自变量,以综合气象干旱指数(CI)作为因变量,利用梯度提升机(GBM)方法,构建了综合干旱日尺度监测模型,并进行了验证和相关性分析,且通过 2015—2018 年华北地区的干旱发生情况,对模型的适用性进行分析,结果表明:

1)综合干旱监测模型监测值与 CI 值的训练集的相关决定系数(R^2)为 0.945,与测试集 R^2 为 0.665,有较好的相关性,训练集均方根误差(RMSE)与平均绝对误差(MAE)分别为 0.033 与 0.021,测试集 RMSE 与 MAE 分别为 0.082 与 0.062,误差较小.生长季(4—10 月)模型监测结果与 CI 干旱监测等级一致率均在 65%以上,说明模型预测有较高的精度.

2)2015—2018 年生长季(4—10 月)逐天的综合干旱监测模型监测结果与基于站点的标准化降水蒸散指数(SPEI)和土壤相对湿度(RSM)相关系数也大多在 0.6 以上,呈现较好的相关性.说明模型不仅对气象干旱具有一定的监测能力,同时对于土壤相对湿度的变化及区域农业干旱也有一定的监测能力.

3)典型干旱情况监测对比结果表明,与站点 CI 监测方法相比,综合干旱监测模型能够较为准确地识别出干旱情况.同时,模型监测结果能够综合考虑各种干旱影响因素,如下垫面地形地势以及植被生长状况等,较好地表征综合干旱状况且不受站点数量和位置分布的影响.

此外,本文综合干旱监测模型仍存在一定的局限性,需要在今后的研究工作中改进:

1)本文主要利用机器学习中梯度提升机(GBM)方法从 MODIS 数据以及 CLDAS 数据中提取干旱发生的相关信息,探索利用多源数据进行区域的综合干旱监测方法,但对于更多其他构建综合干旱监测模型的方法,还需进一步研究.另外,本文提出的综合干旱监测模型,在构建时仍依赖研究区站点观测数据,这会影响模型的建立和模型的应用和推广.因此,如何不受站点观测数据限制建立综合干旱监测模型,仍需要进一步研究.

2)本文从大气、土壤和植被多方面综合考虑多种影响干旱因子,进行综合干旱监测模型的构建,并利用 CLDAS 数据,对日尺度综合干旱监测进行了初步探索,并取得了较好的效果.但是由于干旱成因复杂、致旱因子复杂多样,如植被类型、地表反照率或人类活动等都会对干旱发生产生影响,另外,由于所选择的降水、植被等自变量的前期状况也会对长历时或现时干旱的发生产生影响,而本文并未考虑到降水、植被等变量影响的滞后效应对于长历时或现时干旱的影响.因此,紧密结合区域特性,更精细地探究不同致旱因子在干旱发生中的作用,引入更长时间序列的 CLDAS 精确数据,辅以更加优越精确的干旱监测模型,例如处理延迟序列具有优势的长短期记忆网络(LSTM)模型,将滞后效应的影响纳入分析,进行综合干旱监测的研究,这些仍是需要进一步探讨的重要问题.

参考文献

References

- [1] 王劲松,李耀辉,王润元,等.我国气象干旱研究进展评述[J].干旱气象,2012(4):497-508
WANG Jinsong, LI Yaohui, WANG Runyuan, et al. Preliminary analysis on the demand and review of progress in the field of meteorological drought research[J]. Journal of Arid Meteorology, 2012(4):497-508
- [2] 李新周,刘晓东,马柱国.近百年来全球主要干旱区的干旱化特征分析[J].干旱区研究,2004,21(2):97-103
LI Xinzhou, LIU Xiaodong, MA Zhuguo. Analysis on the drought characteristics in the main arid regions in the world since recent hundred-odd years[J]. Arid Zone Research, 2004, 21(2):97-103
- [3] 屈艳萍,吕娟,苏志诚,等.抗旱减灾研究综述及展望[J].水利学报,2018,49(1):115-125
QU Yanping, LÜ Juan, SU Zhicheng, et al. Research review and perspective of drought mitigation[J]. Journal of Hydraulic Engineering, 2018, 49(1):115-125
- [4] McKee T B, Doesken N J, Kleist J. The relationship of drought frequency and duration to time scales[C]//Pro-

- ceedings of the 8th Conference on Applied Climatology. January 17-22, 1993, Anaheim, California, USA. 1993, 17(22): 179-183
- [5] Vicente-Serrano S, Beguería S, López-Moreno J I. A multi-scalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index [J]. *Journal of Climate*, 2010, 23(7): 1696-1718
- [6] Palmer W C. Keeping track of crop moisture conditions, nationwide: the new crop moisture index [J]. *Weatherwise*, 1968, 21: 156-161
- [7] Kogan F N. Application of vegetation index and brightness temperature for drought detection [J]. *Advances in Space Research*, 1995, 15(11): 91-100
- [8] Kogan F, Sullivan J. Development of global drought-watch system using NOAA/AVHRR data [J]. *Advances in Space Research*, 1993, 13(5): 219-222
- [9] Zhang L F, Jiao W Z, Zhang H M, et al. Studying drought phenomena in the continental United States in 2011 and 2012 using various drought indices [J]. *Remote Sensing of Environment*, 2017, 190: 96-106
- [10] 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): 1047-1056
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): 1047-1056
- [11] 牛乾坤, 刘浏, 程淑雅, 等. 基于多源遥感数据的河套灌区干旱时空演变特征 [J]. *干旱地区农业研究*, 2020, 38(4): 266-277
NIU Qiankun, LIU Liu, CHENG Qiuya, et al. Spatiotemporal characteristics of drought in the Hetao irrigation district based on remote sensing data from multi-sources [J]. *Agricultural Research in the Arid Areas*, 2020, 38(4): 266-277
- [12] 秦大辉, 杨灵, 谌伦超, 等. 基于多源数据的新疆干旱特征及干旱模型研究 [J]. *自然资源遥感*, 2022(1): 151-157
QIN Dahui, YANG Ling, CHEN Lunchao, et al. A study on the characteristics and model of drought in Xinjiang based on multi-source data [J]. *Remote Sensing for Natural Resources*, 2022(1): 151-157
- [13] 吴志勇, 程丹丹, 何海, 等. 综合干旱指数研究进展 [J]. *水资源保护*, 2021, 37(1): 36-45
WU Zhiyong, CHENG Dandan, HE Hai, et al. Research progress of composite drought index [J]. *Water Resources Protection*, 2021, 37(1): 36-45
- [14] 刘莹, 师春香, 王海军, 等. CLDAS 气温数据在中国区域的适用性评估 [J]. *大气科学学报*, 2021, 44(4): 540-548
LIU Ying, SHI Chunxiang, WANG Haijun, et al. Applicability assessment of CLDAS temperature data in China [J]. *Transactions of Atmospheric Sciences*, 2021, 44(4): 540-548
- [15] 师春香, 姜立鹏, 朱智, 等. 基于 CLDAS 2.0 驱动数据的中国区域土壤湿度模拟与评估 [J]. *江苏农业科学*, 2018, 46(4): 231-236
SHI Chunxiang, JIANG Lipeng, ZHU Zhi, et al. Simulation and evaluation of soil moisture in China based on CLDAS 2.0 driving data [J]. *Jiangsu Agricultural Sciences*, 2018, 46(4): 231-236
- [16] Huang A, Shen R, Shi C, et al. Effects of satellite LAI data on modelling land surface temperature and related energy budget in the Noah-MP land surface model [J]. *Journal of Hydrology*, 2022, 613: 128351
- [17] 许金涛. 中国大陆 IMERG 降水产品精度评价与融合校正研究 [D]. 杭州: 浙江大学, 2020
XU Jintao. Precision evaluation and fusion correction of IMERG precipitation products in Chinese mainland [D]. Hangzhou: Zhejiang University, 2020
- [18] 安雪丽, 王前锋, 莫新宇, 等. 华北地区农业干旱灾害变化特征 [J]. *北京师范大学学报(自然科学版)*, 2016, 52(5): 591-596
AN Xueli, WANG Qianfeng, MO Xinyu, et al. Characteristics of agricultural drought disaster in Northern China [J]. *Journal of Beijing Normal University (Natural Science)*, 2016, 52(5): 591-596
- [19] 孙帅, 师春香, 梁晓, 等. 不同陆面模式对我国地表温度模拟的适用性评估 [J]. *应用气象学报*, 2017, 28(6): 737-749
SUN Shuai, SHI Chunxiang, LIANG Xiao, et al. Assessment of ground temperature simulation in China by different land surface models based on station observations [J]. *Journal of Applied Meteorological Science*, 2017, 28(6): 737-749
- [20] 国家青藏高原科学数据中心. 面向陆面过程模型的中国土壤水文数据集(1980) [DS/OL]. [2023-01-01]. <https://data.tpdc.ac.cn/zh-hans/data/205da4ae-63cd-48e1-994e-0b5d8830812a/?q=>
National Tibetan Plateau/Third Pole Environment Data Center. A China dataset of soil hydraulic parameters pedotransfer functions for land surface modeling (1980) [DS/OL]. [2023-01-01]. <https://data.tpdc.ac.cn/zh-hans/data/205da4ae-63cd-48e1-994e-0b5d8830812a/?q=>
- [21] 韩慧敏, 沈润平, 黄安奇, 等. 基于集成学习方法的 CLDAS 土壤湿度降尺度研究 [J]. *南京信息工程大学学报(自然科学版)*, 2021, 13(6): 693-706
HAN Huimin, SHEN Runping, HUANG Anqi, et al. Downscaling of CLDAS soil moisture based on ensemble learning method [J]. *Journal of Nanjing University of Information Science & Technology (Natural Science Edition)*, 2021, 13(6): 693-706
- [22] 杜灵通. 基于多源空间信息的干旱监测模型构建及其应用研究 [D]. 南京: 南京大学, 2013
DU Lingtong. Construction and application of drought monitoring model based on multi-source spatial information [D]. Nanjing: Nanjing University, 2013
- [23] Purcell L C, Sinclair T R, McNew R W. Drought avoidance assessment for summer annual crops using long-term weather data [J]. *Agronomy Journal*, 2003, 95(6): 1566-1576
- [24] Mishra A, Vu T, Veetil A V, et al. Drought monitoring with soil moisture active passive (SMAP) measurements [J]. *Journal of Hydrology*, 2017, 552(9): 620-632
- [25] Makkink G F. Testing the Penman formula by means of

- lysimeters[J]. Journal of the Institution of Water Engineers, 1957, 11: 277-288
- [26] Martínez-Fernández J, González-Zamora A, Sánchez N, et al. A soil water based index as a suitable agricultural drought indicator[J]. Journal of Hydrology, 2015, 522: 265-273
- [27] Carlson T N, Gillies R R, Perry E M. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover[J]. Remote Sensing Reviews, 1994, 9(1/2): 161-173
- [28] 国家质量监督检验检疫总局, 中国国家标准化管理委员会. 气象干旱等级: GB/T 20481—2006[S]. 北京: 中国标准出版社, 2007
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Classification of meteorological drought: GB/T 20481—2006[S]. Beijing: Standards Press of China, 2007
- [29] 徐继伟, 杨云. 集成学习方法: 研究综述[J]. 云南大学学报(自然科学版), 2018, 40(6): 1082-1092
XU Jiwei, YANG Yun. A survey of ensemble learning approaches[J]. Journal of Yunnan University (Natural Sciences Edition), 2018, 40(6): 1082-1092
- [30] Shen R P, Huang A Q, Li B L, 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
- [31] 郭佳. 基于随机森林的遥感干旱监测模型及其应用研究[D]. 南京: 南京信息工程大学, 2016
GUO Jia. Remote sensing drought monitoring model based on random forest and its application[D]. Nanjing: Nanjing University of Information Science & Technology, 2016
- [32] 中华人民共和国水利部. 中国水旱灾害公报 2016[R]. 北京: 中国水利水电出版社, 2017

Integrated drought monitoring model based on MODIS and CLDAS

XING Yajie¹ SHEN Runping¹ HUANG Anqi¹ LIANG Yujing¹

WANG Yunyu¹ XIE Zhaoying¹ SHI Chunxiang² SUN Shuai²

1 School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China

2 National Meteorological Information Center, Beijing 100081, China

Abstract Traditional drought indices mainly consider a single factor and often cannot comprehensively reflect the drought condition. Based on data of MODIS and CLDAS (CMA Land Data Assimilation System), a daily scale integrated drought monitoring model was established by Gradient Boosting Machine (GBM) with multiple influencing factors and drought index as independent variables and comprehensive meteorological drought index (CI) as dependent variable. It was researched by taking drought in North China from 2015 to 2018 as a case. The results show that the model monitoring results are significantly correlated with the calculated CI values of the observation stations. The coefficient of determination is 0.945 and 0.655, and the Root Mean Square Error (RMSE) is 0.033 and 0.082 for training and test sets, respectively, indicating the high accuracy of the proposed integrated drought monitoring model. The consistency rate between the model monitored CI and calculated CI values is above 65%, and the correlation coefficient with Standard Precipitation Evapotranspiration Index (SPEI) and Relative Soil Moisture (RSM) is 0.68 and 0.6, respectively, showing its capacity to reflect both the meteorological drought and the agricultural drought. Monitoring of typical drought condition shows that the integrated drought monitoring model can accurately identify the drought occurrence, and represent the situation of comprehensive drought via considering various drought influencing factors.

Key words CLDAS; integrated drought monitoring; gradient boosting machine (GBM); MODIS