

DOI: 10.20103/j.stxb.202210082851

何锐, 陆恒, 晋子振, 秦艳, 杨弘毅, 刘治银, 杨广睿, 徐进滢, 龚雪, 赵求东. 基于随机森林算法的中国西南地区林火发生预测模型构建及驱动因子. 生态学报, 2023, 43(22): 9356-9370.

He R, Lu H, Jin Z Z, Qin Y, Yang H Y, Liu Z Y, Yang G R, Xu J Y, Gong X, Zhao Q D. Construction of forest fire prediction model and driving factors analysis based on random forests algorithm in Southwest China. Acta Ecologica Sinica, 2023, 43(22): 9356-9370.

基于随机森林算法的中国西南地区林火发生预测模型构建及驱动因子

何 锐^{1,3}, 陆 恒², 晋子振^{3,4}, 秦 艳⁵, 杨弘毅³, 刘治银³, 杨广睿⁶, 徐进滢⁶, 龚 雪², 赵求东^{1,3,*}

1 中国科学院 西北生态环境资源研究院 内陆河流域生态水文重点实验室, 兰州 730000

2 四川师范大学地理与资源科学学院, 成都 610000

3 中国科学院大学, 北京 100049

4 中国科学院 西北生态环境资源研究院 冰冻圈科学国家重点实验室, 兰州 730000

5 新疆大学资源与环境科学学院, 乌鲁木齐 830046

6 南京师范大学地理科学学院, 南京 210023

摘要: 林火直接破坏森林资源, 改变森林的结构与功能, 影响局地甚至全球气候状况并威胁人类生命和财产安全, 在气候变暖背景下林火将更加频发, 因此开展林火预测/预报研究至关重要。利用 MODIS (Moderate-resolution Imaging Spectroradiometer) 的温度异常/火产品 (MOD14A1) 获取逐日林火数据, 分析了 2001—2018 年中国西南地区林火时空分布特征; 采用随机森林算法, 综合考虑气象、地形、可燃物状况及植被等林火驱动因子, 构建了中国西南地区干、湿季林火发生预测模型, 系统分析了西南地区干湿季林火发生的主要驱动因子。结果表明: (1) 中国西南地区林火主要集中分布于云南大部、四川西南部及贵州南部地区, 并呈集聚分布特征; 林火多发于干季, 占林火发生总次数的 96.5%, 年林火发生次数呈阶段性变化特征, 2001—2014 年呈现显著增加趋势, 随后表现为不显著减少趋势; (2) 构建的干、湿季林火发生预测模型能较准确地模拟林火发生状况: 训练期模型准确率分别处于 82.94%—83.99% 与 85.12%—90.31% 之间, AUC (Area Under Curve) 值分别处于 0.908—0.914 与 0.922—0.965 之间; 测试期模型准确率分别为 79.73% 和 83.27%, AUC 值分别为 0.886 和 0.855; (3) 海拔是西南地区林火发生最关键的限制因子, 林火多集中于中海拔区, 而在低海拔和高海拔地区林火不易发生, 这与人类活动密切相关。当日的气象条件是干季林火发生次重要的驱动因子, 可燃物的温湿度状况则是湿季林火发生次重要的驱动因子。FWI 系统指标 (Fire Weather Index) 在西南地区有较好的适用性且对于区域干湿季林火发生均有重要的影响, 因此在西南地区林火预测/预报工作中有必要引入 FWI 系统指标。

关键词: 林火预测模型; 随机森林算法; 林火驱动因子; FWI 系统; 西南地区

Construction of forest fire prediction model and driving factors analysis based on random forests algorithm in Southwest China

HE Rui^{1,3}, LU Heng², JIN Zizhen^{3,4}, QIN Yan⁵, YANG Hongyi³, LIU Zhiyin³, YANG Guangrui⁶, XU Jinying⁶, GONG Xue², ZHAO Qiudong^{1,3,*}

1 Key Laboratory of Ecohydrology of Inland River Basin, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China

基金项目: 教育部人文社科青年项目 (21YJCZH099); 四川师范大学地理与资源科学学院创新训练项目 (202010636326)

收稿日期: 2022-10-08; **网络出版日期:** 2023-09-13

* 通讯作者 Corresponding author. E-mail: Zhaoqd@lzb.ac.cn

2 College of Geography and Resources Sciences, Sichuan Normal University, Chengdu 610000, China;

3 University of Chinese Academy of Sciences, Beijing 100049, China

4 State Key Laboratory of Cryospheric Science, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China

5 College of Resources and Environment Sciences, Xinjiang University, Urumqi 830046, China

6 School of Geographical Science, Nanjing Normal University, Nanjing 210023, China

Abstract: Forest fires directly destroy forest resources, change the structure and function of forests, affect local and even global climate conditions and threaten human life and property. And forest fire will be more frequent in the context of global warming, so the study of forest fire prediction/forecast is vital. We obtained the daily forest fire data from the Thermal Anomalies/Fire Daily L3 product (MOD14A1) of Moderate-resolution Imaging Spectroradiometer (MODIS) and analyzed the spatial and temporal distribution characteristics in Southwest China during 2001—2018. The random forests algorithm was adopted to construct the dry and wet season's forest fire prediction models with full consideration of driving factors (contained meteorological elements, topography, fuel and vegetation). Then, we identified the main driving factors of forest fire occurrences through simulation in dry and wet seasons in Southwest China. The following major conclusions were drawn: (1) The forest fires were mainly concentrated in the most regions of Yunnan Province, southwestern Sichuan Province and southern Guizhou Province, and showed a significant agglomeration distribution. The forest fires mostly occurred in the dry season accounting for 96.5% of the total forest fires. The number of annual forest fires showed a staggered transformation with a significant increase trend during 2001—2014 and non-significant decrease trend after 2014. (2) The developed random forests algorithm-based forest fire prediction model archived good performances. The accuracy of the model in the training period were between 82.94%—83.99% and 85.12%—90.31%, and the Area Under Curve (AUC) values were between 0.908—0.914 and 0.922—0.965, respectively; the accuracy of the model in the testing period were 79.73% and 83.27%, and the AUC values were 0.886 and 0.855, respectively. (3) The elevation was the most important limiting factor to forest fire occurrences for both dry and wet seasons in Southwest China. The forest fires were concentrated in the mid-altitude areas, while they were less likely to occur in the low and high altitude areas, which was most likely related to human activities. The meteorological conditions on the day of fire occurrences were the second important driving factors to forest fires in dry season, while the fuel moisture and temperature conditions were the second important driving factors to forest fires in wet season. The Fire Weather Index (FWI) system had good applicability in the Southwest China and had a significant impact on the occurrence of forest fires in both the wet and dry seasons, therefore it was necessary to consider the FWI system index in forest fire prediction/forecast in the region.

Key Words: forest fire prediction model; random forests algorithm; driving factors; Fire Weather Index system; Southwest China

林火是破坏森林资源的主要自然灾害之一^[1],改变了森林的结构和功能,降低了森林在生态环境保护中的作用,增加了大气中温室气体的含量^[2],从而影响局地乃至全球气候^[3],对人类生命和财产安全造成巨大威胁。气候变暖背景下,全球和中国的林火发生频次、规模及导致的经济损失呈现明显的增加趋势^[4-6]。据IPCC(Intergovernmental Panel on Climate Change)报告显示,未来极端干旱高温事件将更加频发^[7],可能导致林火发生频次随之增加^[8]。因此,开展林火预测/预报及风险评估研究对林火防治、维持社会经济和森林生态可持续发展具有重要科学意义。

林火发生预测模型是目前开展林火风险评估及预测/预报研究最广泛使用的方法,代表性模型有加拿大森林火险等级系统(CFFDRS)^[9]、美国国家火灾风险等级系统(NFDRS)^[10]等,是基于历史数据构建起林火与驱动因子间的数学关系,其模拟和预测精度主要取决于驱动因子和数学方法的选择,目前林火发生驱动因子的选择已从单纯的气象要素(也称气象火险预报)发展到综合考虑影响林火发生的众多要素(主要包括地形、

可燃物、人类活动、气象要素及火源等),使得林火模拟和预测精度得到极大提高。林火发生预测模型中常用的数学方法主要有逻辑斯蒂回归(LR)^[11]、地理加权回归(GWR)^[12]和 Gompit 回归等^[13],这三种统计学算法假设解释变量及因变量是线性关系,同时无法去除变量间的多重共线性,从而影响模拟和预测精度,因此在实际使用中受到一定限制。

近年来机器学习方法得到快速发展,这类方法能够快速、有效处理高维数据,因此有部分学者将其应用于林火发生预测研究中。机器学习模型将林火与驱动因子的关系转化为非线性关系,且在不同区域都显示出较高精度^[14-16],但机器学习模型无法明晰林火驱动因子与林火发生的具体关系,在保证准确度的同时一定程度上丢失了可解释性。而随机森林算法作为机器学习的典型代表,较之其他机器学习模型计算成本更低,能够对缺失值进行有效估计和处理,从而保证模型精度且不易陷入过拟合,并且能够很好克服变量间的共线性^[17],其良好的学习性能在与传统统计模型的比较中被凸显,在林火预测研究中的准确性也得到广泛验证^[18-20]。此外,我国气候存在明显的干湿季节性特征^[21],而以往研究中多将全年林火事件作为整体研究,忽略林火驱动因子的干湿季节差异,这可能导致对林火驱动因子的认识存在一定问题,且目前基于随机森林算法进行林火预测的研究多集中在较小区域内,模型泛化能力有待进一步研究。

因此,本研究以我国林火发生相对频繁的西南地区为研究对象,采用 EOS/MODIS 火点数据,基于随机森林算法,结合气象、地形、可燃物状况等构建中国西南地区林火发生预测模型,明晰中国西南地区干、湿季节林火发生的主要驱动因子及评估区域林火风险,为相关部门进行林火预测和制定林火防控政策提供科学支撑。

1 数据与方法

1.1 研究区概况

中国西南地区位于 21.1°—34.2°N, 97.3°—110.1°E,面积约 $112.49 \times 10^4 \text{ km}^2$ 。主要包括四川省,云南省,重庆市,贵州省和西藏自治区^[22],由于西藏自治区地处青藏高原,自然环境极为特殊且非林火频发区域,故本研究未将西藏自治区纳入考量(图 1)。研究区西部经横断山脉与青藏高原东南部接壤,南部为云贵高原,北

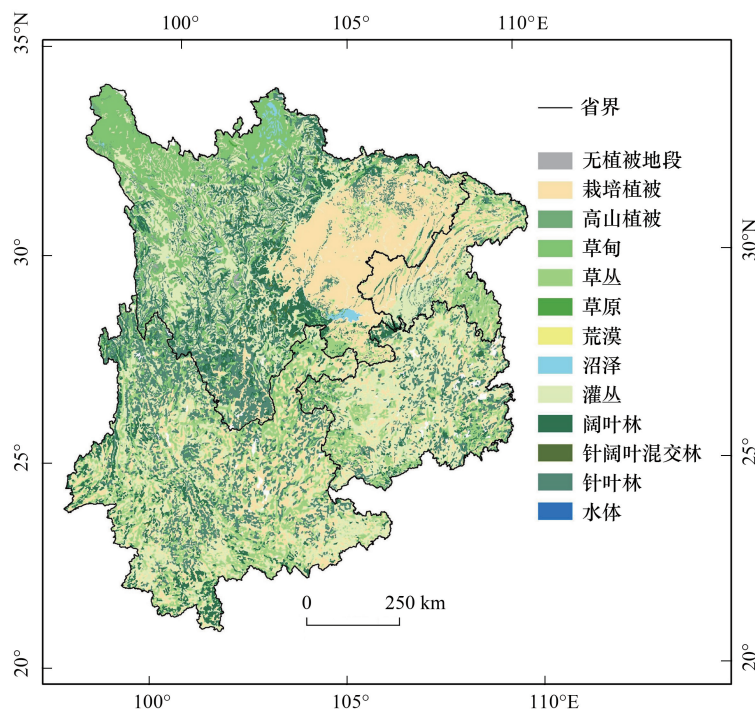


图 1 西南地区位置及植被情况

Fig.1 Location and vegetation distribution of Southwest China

部为四川盆地及川东方山丘陵,地势总体西高东低,山地、高原、盆地、平原和丘陵地形均有分布,区域高差达五千余米。该区域处于东亚季风,南亚季风以及青藏高原的交汇地带^[23],气候类型主要包括亚热带季风气候、高原山地气候和高原季风气候,干湿季分明^[24]。西南地区多样的气候和复杂的地形导致区域植被类型丰富,垂直带谱相对完整,森林面积约 $50.72 \times 10^4 \text{ km}^2$,森林覆盖率约 45.09%,森林蓄积量达 $44.32 \times 10^8 \text{ m}^3$ ^[4],是我国第二大天然林区,同时也是我国林火高发地区之一,2004—2018 年西南地区森林过火面积高达 $3.21 \times 10^3 \text{ km}^2$ ^[25],林火风险较大,区域森林资源管理面临严峻挑战。

1.2 数据和方法

1.2.1 林火数据

2001—2018 年林火数据来自美国国家航空航天局提供的 MODIS 日 1km 温度异常/火 L3 级产品 (MOD14A1, <https://ladsweb.modaps.eosdis.nasa.gov/>),该数据已被广泛运用于国内外林火研究中^[26—28]。此前多数研究将一个高可信度点 ($80 \leq \text{可信度} \leq 100$) 视为一次火灾过程^[29—30],这可能造成对火灾次数的高估。为避免一些为非火点的高可信度点被误识或传感器多次扫描到同一场火灾而生成多个火点,本文参考曾爱聪等^[31]、Guo^[32] 等方法对 MOD14A1 数据做以下预处理:(1)筛选出可信度 ≥ 80 的点作为真实火点;(2)利用土地覆被数据提取林地覆盖下的火点作为林火点;(3)将同时满足时间间隔 $\leq 1\text{d}$ 且距离 $\leq 1\text{km}$ 条件的火点归为同一次火灾过程,且仅保留最早出现的林火点的相关信息。为评价 MOD14A1 火点的可靠性,在中国科学院遥感与数字地球研究所获取了 2014—2018 年的 Landsat 8 OLI 火点数据 (<http://satsee.radi.ac.cn:8080/index.html>),经相同方法预处理后与 2014—2018 年 MOD14A1 数据作比对。图 2、图 3 显示 Landsat 8 OLI 能探测到更多的火点,这可能是由于其具有更高的空间分辨率 (30m) 所致,但两类火点数据年变化趋势相对一致,空间分布也极为吻合,因此认为 MOD14A1 火点数据较为可靠,可用于本研究。并遵循时空双随机原则,按照 1:1 的比例创建对照点。

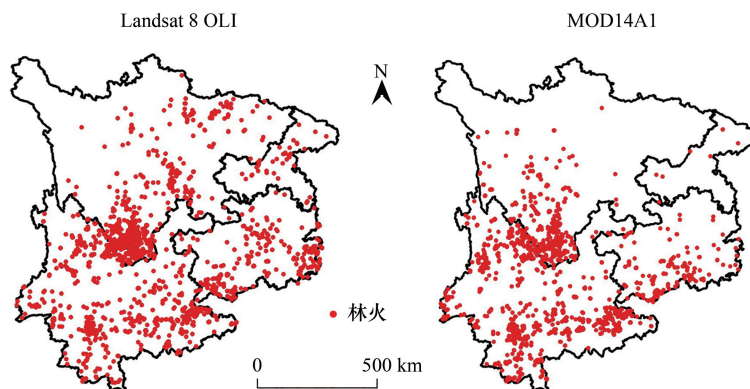


图 2 2014—2018 年 Landsat 8 OLI 与 MOD14A1 林火空间分布

Fig.2 Spatial distribution of Landsat 8 OLI and MOD14A1 forest fires during 2014—2018

1.2.2 气象及下垫面数据

2001—2018 年气象数据来自国家青藏高原科学数据中心提供的中国区域地面气象要素驱动数据集 (CMFD, <http://data.tpcd.ac.cn/zh-hans/>),该数据以 Princeton 再分析资料、GLDAS 气象资料、GEWEX-SRB 辐射资料及 TRMM 降水资料为背景场,融合气象站点观测数据而制作,数据整体准确度较高,已被广泛应用于水文和气候模拟研究。由于 CMFD 空间分辨率较低 (0.1°),导致火点处高程与格网中心高程存在明显差异,故本文利用逐月垂直梯度 (表 1) 将格网的气象要素 (气温、降水、比湿等) 订正到火点处。

地形数据采用空间分辨率 30m 全球数字高程模型 (ASTER GDEM V003, <https://search.earthdata.nasa.gov/>)。土地覆被数据采用空间分辨率 30m 全球地表覆盖数据 (GlobeLand30, <https://www.webmap.cn/>)。

1.2.3 林火驱动因子选取

在参考前人指标选取的基础上^[13,33],结合西南地区气候年内差异较大以及地形复杂多样的自然环境特点,共选取林火发生日的气象、地形、植被类型、可燃物含水率和火行为指数等 15 个潜在驱动因子(表 2),并对各指标进行归一化处理。其中可燃物含水率和火行为指数采用加拿大火灾天气指数(FWI)系统的指标,FWI 系统是加拿大森林火险等级系统(CFFDRS)的子系统,以时滞—平衡含水率理论为基础,通过气象要素计算与林火发生条件相关的可燃物含水率、火行为等综合指标,这些指标目前已被广泛应用于国内外林火监测和预报系统中且取得较好结果^[34—36]。本文利用 CMFD 逐日气象数据计算区域 2001—2018 年 FWI 系统内各指标,各指标计算公式具体参考文献^[37]。

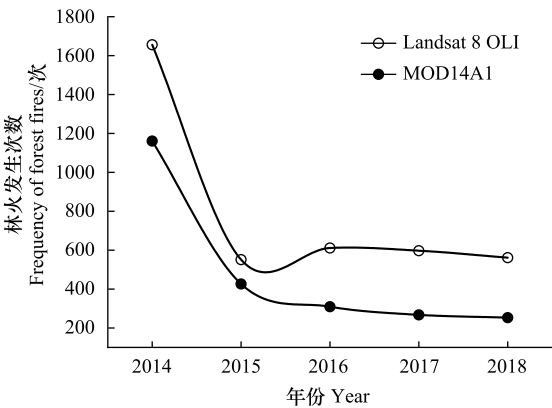


图 3 2014—2018 年 Landsat 8 OLI 与 MOD14A1 林火年际变化
Fig.3 Annual changes of Landsat 8 OLI and MOD14A1 forest fires during 2014—2018

表 1 气象要素月平均垂直梯度

Table 1 Monthly altitude gradient of meteorological elements

月份 Month	气温 Temperature/ (℃/100m)	降水 Precipitation/ (mm/100m)	比湿 Specific humidity/ ((kg/kg)/100m)	月份 Month	气温 Temperature/ (℃/100m)	降水 Precipitation/ (mm/100m)	比湿 Specific humidity/ ((kg/kg)/100m)
1	-0.46	-0.42	-0.001	7	-0.5	-1.86	-0.003
2	-0.44	-0.4	-0.001	8	-0.51	-1.59	-0.003
3	-0.46	-0.49	-0.001	9	-0.48	-1.47	-0.003
4	-0.49	-1.16	-0.002	10	-0.48	-1.06	-0.002
5	-0.48	-1.98	-0.002	11	-0.51	-1.01	-0.002
6	-0.47	-1.51	-0.002	12	-0.49	-0.46	-0.001

表 2 潜在驱动因子

Table 2 Potential driving factors

变量类型 Types of variables	变量名 Names of variables	代码 Code	单位 Unit	注释 Note
气象 Meteorology	正午气温	TEMP	℃	
	正午风速	WIND	m/s	
	日降水量	PREC	mm	
	正午相对湿度	RH	%	
可燃物 Fuel	细小可燃物湿度码	FFMC		凋落层 0—2cm 可燃物湿度状况
	半腐层湿度码	DMC		地被物 5—10cm 可燃物湿度状况
	干旱码	DC		地被物 10—20cm 可燃物湿度状况
火行为指数 Fire behavior index	初始蔓延速度	ISI		无可燃物数量影响下的蔓延速度
	可燃物累积指数	BUI		用于蔓延燃烧的可燃物总量
	火灾天气指数	FWI		单位火线长度的火强度
	日严重等级	DSR		每日潜在火险
地形 Topography	海拔	ELEVATION	m	
	坡度	SLOPE	(°)	
	坡向	ASPECT	(°)	
植被 Vegetation	植被类型	VEGETATION		反映不同林地的易燃性类别

1.3 随机森林算法

随机森林(Random Forests)是由 Leo Breiman 和 Adele Cutler 于 2001 年提出^[38]的一种基于集成学习思想的机器学习算法,可有效地估计和处理缺失值且能够保证模型准确性^[39]。基本单元是决策树(Decision Tree),其生长是通过有放回的自主抽样对输入样本进行判别^[40],单棵决策树是一个弱分类器,随机森林将多个决策树的分类结果集合形成一个强分类器,本研究中林火是否发生是一个简单的二分类问题,即依据所有决策树投票结果的众数进行分类。

2 结果

2.1 林火时空分布特征

2001—2018 年西南地区林火共发生 10088 次,年内分布极不均匀(图 4),多发于干季,18 年间共发生林火 9735 次,占林火总次数的 96.5%,其中 3 月林火发生总次数达到最高的 2446 次,占林火总次数的 24.25%;湿季林火发生次数相对较少,占林火总次数的 3.5%,其中 7 月林火发生总次数最少,为 28 次,仅占林火总次数的 0.28%。

2001—2018 年西南地区年林火发生次数呈现阶段性变化特征(图 4),2001—2014 年是西南地区林火高发时段且呈现显著上升趋势(P 值=0.015),共发生林火 8833 次,占林火总次数的 87.56%。其中 2010 年和 2014 年是林火高发年,2 年共发生林火 2476 次,占林火总次数的 24.54%,由图 4 可知主要是由于干季林火发生次数显著增加所致。2014 年以后林火发生次数处于低值状态,呈现不显著下降趋势(P 值=0.096)。

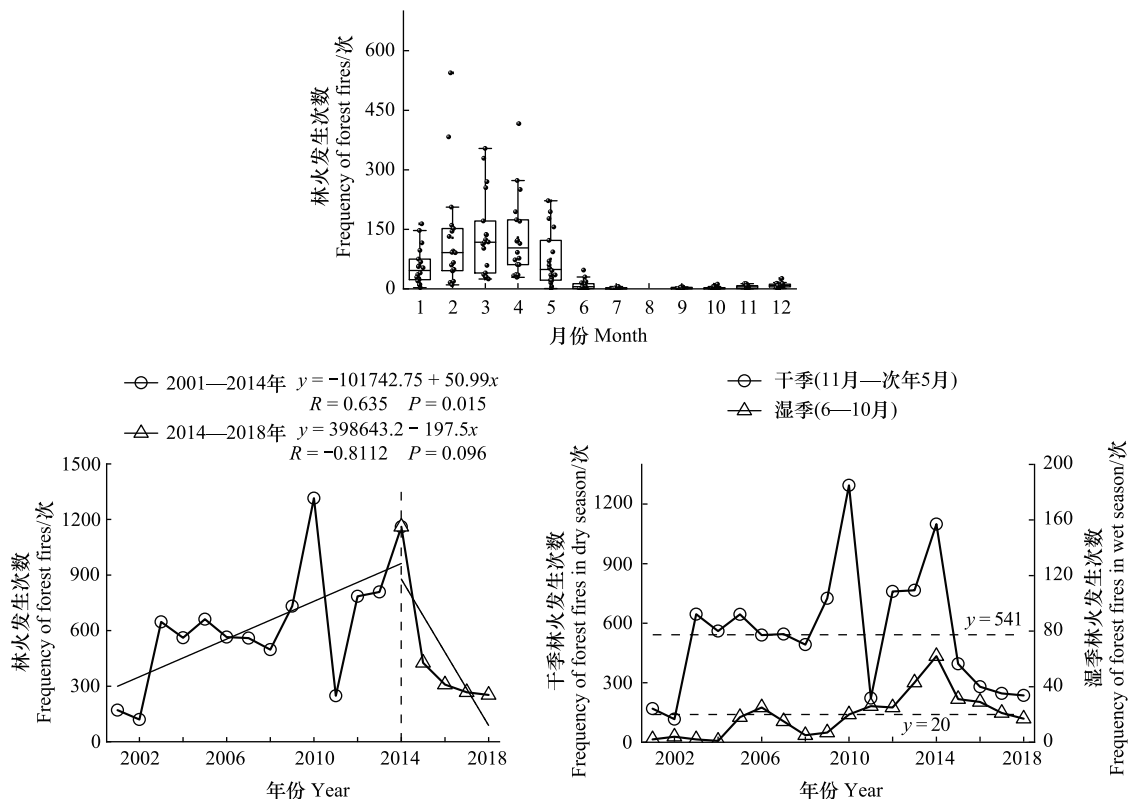


图 4 2001—2018 年林火年发生次数及年内分布

Fig.4 Annual and intra-annual variations of forest fires during 2001–2018

如图 5 所示,2001—2018 年西南地区林火主要集中分布于云南大部、四川西南部及贵州南部地区,而四川东部、重庆大部及贵州北部林火发生较少,林火总体上呈西多东少,南多北少的空间分布规律。运用多距离

空间聚类分析方法(Ripley's K 函数)对中国西南地区林火空间分布格局进行 5 年滑动分析,结果显示 K 的观测值均大于 K 的预期值、高值置信区间以及低值置信区间,表明该区域 2001—2018 年林火空间上呈集聚分布特征,即林火发生具有显著的区域差异且林火高风险区存在于部分区域(图 6)。

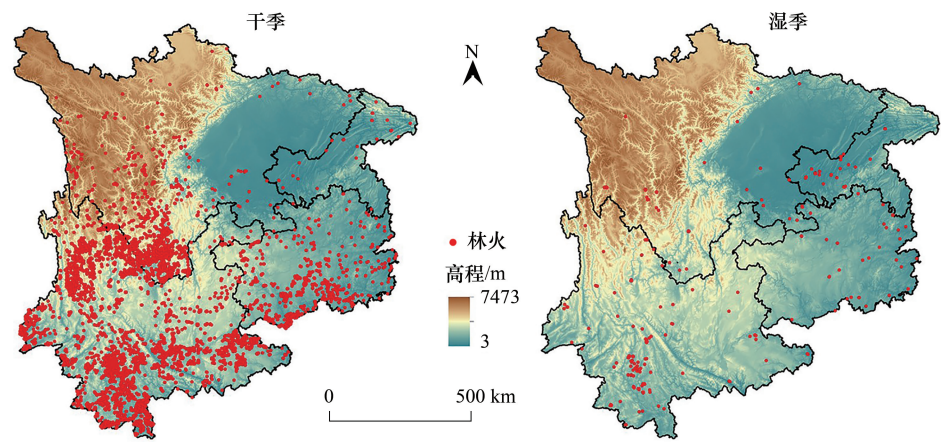


图 5 2001—2018 年林火空间分布
Fig.5 Spatial distribution of forest fires during 2001—2018

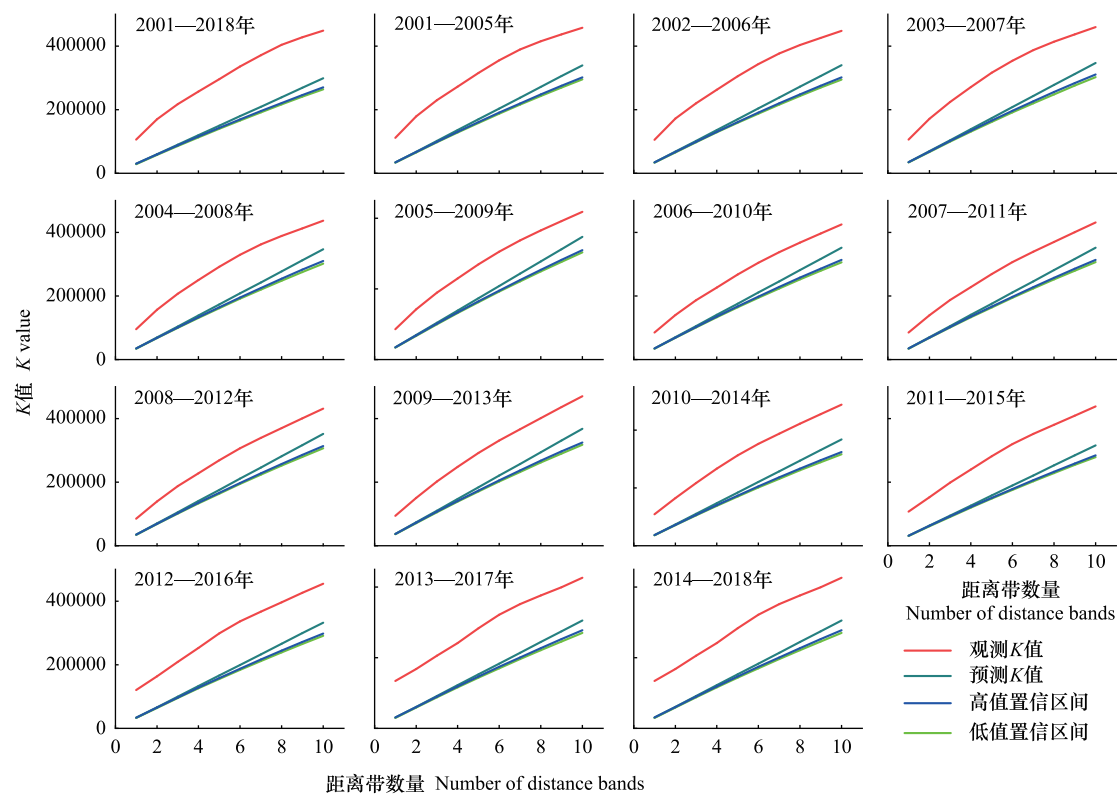


图 6 2001—2018 年林火五年滑动空间分析
Fig.6 Five-year sliding space analysis of forest fires during 2001—2018

2.2 林火发生预测模型评价

本文根据西南地区林火的时间分布特征及驱动因子可能存在的季节性差异,将火点与对照点划分为干季(11 月—次年 5 月)和湿季样本(6—10 月),按年份划分为训练期(2001—2015 年)和测试期(2016—

2018 年),训练期分别按照 60%和 40%的比例重复 5 次划分训练和验证样本建立中间模型并进行变量选择,选择 5 个中间模型中出现 3 次及以上的驱动因子建立指标体系以构建最终预测模型,并基于测试期(2016—2018 年)的数据检验和评价模型性能。模型中决策树数量(ntree)和每次抽样时选取的变量数(mtry)两个参数参考前人研究^[18],ntree 取值 2000,对于干、湿季 mtry 分别取值 4、3。

如表 3 所示,训练期干季中间模型准确率处于 82.94%—83.99%之间,AUC 值处于 0.908—0.914 之间;湿季中间模型准确率处于 85.12%—90.31%之间,AUC 值处于 0.922—0.965 之间。经变量选择后由 11、10 个驱动因子组成指标体系分别构建干、湿季林火预测模型。测试期干季林火预测模型准确率为 79.73%,AUC 值为 0.886;湿季林火预测模型准确率为 83.27%,AUC 值为 0.855。两个最终模型与其中间模型表现一致且准确度较高,与在该区域基于森林火险气象等级^[41—43]、逻辑斯蒂回归^[30]等建立的林火预测模型相比,准确度有一定程度提高,本文基于随机森林算法构建的林火预测模型在大范围模拟中精度较高,因此可应用于西南地区林火预测工作中。

表 3 模型表现

Table 3 Performances of the models

数据划分 Data segmentation	样本 Sample	干季 Dry season		湿季 Wet season	
		准确率	AUC	准确率	AUC
训练期 Training period	样本 1	83.41%	0.91	89.97%	0.953
	样本 2	83.99%	0.914	85.12%	0.922
	样本 3	82.94%	0.908	85.49%	0.926
	样本 4	83.72%	0.91	90.31%	0.965
	样本 5	83.53%	0.914	86.51%	0.932
测试期 Testing period		79.73%	0.886	83.27%	0.855

AUC:曲线下面积 Area under curve

2.3 林火驱动因子分析

本文基于随机森林模型对西南地区干、湿季林火驱动因子进行重要性排序以明确林火发生的主要驱动因子(图 7),并通过偏依赖图明确各驱动因子与林火发生概率的关系(图 8)。

干季和湿季林火发生预测模型中,海拔均是最重要的限制因子,这表明海拔对林火发生具有主要作用。海拔 1000—1800m 是干季林火发生概率最高的区域,海拔 1000—1600m 是湿季林火发生概率最高的区域,而其余海拔范围内林火发生概率较低。

对于干季林火来说,当正午气温小于 0℃时林火发生概率较小,高于 0℃后,林火发生概率呈现快速增加趋势;而随降水增加林火发生概率呈现先减后增的趋势;相对湿度小于 20%时,林火发生概率较高,大于 20%时,林火发生概率迅速下降;风速小于 4m/s 时林火发生概率较小,大于 4m/s 时林火发生概率迅速增大,大于 5m/s 时林火发生概率较高且保持稳定。细小可燃物湿度码(FFMC)、干旱码(DC)和半腐层湿度码(DMC)对干季林火发生影响较大,当 FFMC、DC 及 DMC 大于 70、100 和 5 之后,林火发生概率急剧增加。DC 和 DMC 大于 400 和 100 后林火发生概率出现一定程度下降,此外日严重等级(DSR)、火灾天气指数(FWI)及初始蔓延速度(ISI)这些指标对于干季林火也有一定影响,FWI、DSR 及 ISI 与林火发生概率呈现正相关,且当其分别等于 10、20、10 左右时林火发生概率达到最大。

对于湿季林火来说,当正午气温小于 13℃时林火发生概率较小,高于 13℃后,林火发生概率随气温升高呈现快速增加趋势;当相对湿度小于 40%时林火发生概率较大,后随相对湿度增加而迅速下降。FWI 系统各指标均进入模型中,对湿季林火发生产生一定影响,当 DC、DMC、BUI、FWI、DSR、ISI 分别小于 400、100、120、27、10 和 7 时林火发生概率与各指标值呈正相关关系,之后林火发生概率达到最大且保持相对稳定;FFMC 小于 80 时,林火发生概率较低,之后随着值的增大,林火发生概率迅速升高。

2.4 林火发生风险评估

利用经验证后的干、湿季林火发生预测模型分别对西南地区干、湿季林火发生概率进行预测,再利用克里

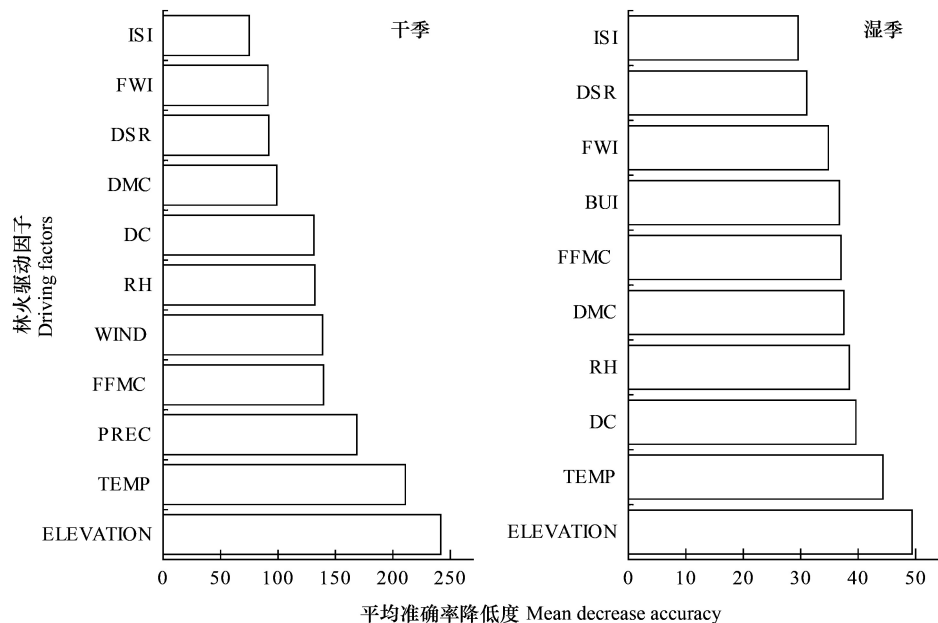


图7 干季、湿季林火驱动因子重要性排序

Fig.7 Importance ranking for driving factors of forest fires in dry and wet seasons

ELEVATION:海拔;TEMP:正午气温;PREC:日降水量;FPMC:细小可燃物湿度码;WIND:正午风速;RH:正午相对湿度;DC:干旱码;DMC:半腐层湿度码;DSR:日严重等级;FWI:火灾天气指数;ISI:初始蔓延速度;BUI:可燃物累积指数

金法对模型输出的林火发生概率进行空间插值,并参考马文苑等的研究(表4)^[44]对区域林火风险进行等级划分。如图9所示,干季云南、四川南部与贵州西部部分地区林火风险等级较高;此外四川西部有零星分布的林火高风险地区。湿季云南西部、四川西南部、贵州东部与重庆西部林火风险等级较高。综上,干季林火风险明显大于湿季,云南西部和四川西南部全年林火风险都较高,林火极易发生;而重庆、贵州干湿季林火风险分布具有明显空间和季节差异。

表4 火险等级区划

Table 4 Forest fire risk distribution

林火发生概率 Probability of forest fire	火险等级划分 Division of forest fire risk	可能火情 Possible fire condition	林火发生概率 Probability of forest fire	火险等级划分 Division of forest fire risk	可能火情 Possible fire condition
$0 \leq P \leq 0.2$	一级火险区	基本无火情	$0.6 < P \leq 0.8$	四级火险区	易发生
$0.2 < P \leq 0.4$	二级火险区	不易发生	$0.8 < P \leq 1$	五级火险区	极易发生
$0.4 < P \leq 0.6$	三级火险区	可能发生			

3 讨论

本文基于 MOD14A1 数据,综合考虑气象、地形、植被类型及 FWI 系统指标,利用随机森林算法构建西南地区的干、湿季林火预测模型,并系统评估西南林火发生的驱动因子。

为探究划分干湿季林火分别构建预测模型的必要性,尝试在不划分干湿季林火的情况下,采用相同方法构建林火预测模型。如图10所示,训练期,不划分干湿季林火构建的模型(全年模型)对于季林火的预测准确率处于82.13%—83.02%之间,对湿季林火的预测准确率在80.53%—83.87%之间;验证期,全年模型对于季林火的预测准确率为77.5%,对湿季林火的预测准确率为80.28%。结果表明无论在训练期还是验证期,全年模型对林火的预测精度均低于划分干湿季林火分别构建的预测模型,在湿季这一差异更为显著(P值=

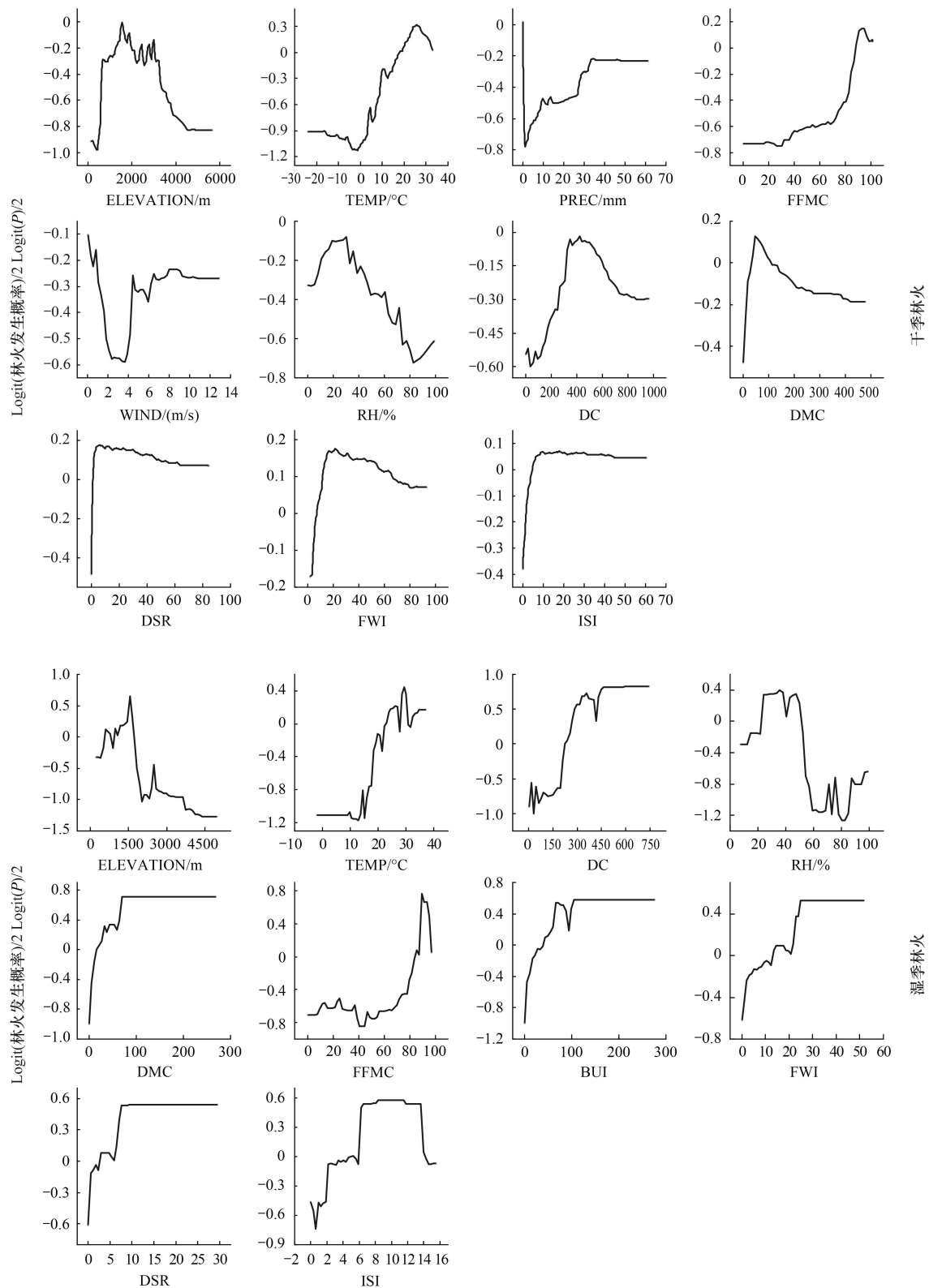


图8 干季、湿季林火发生概率与驱动因子偏依赖关系

Fig.8 The partial dependence relationships between forest fire probability and driving factors in dry and wet seasons

P : 林火发生概率; ELEVATION: 海拔 Elevation; TEMP: 正午气温 Temperature; PREC: 日降水量 Precipitation; FPMC: 细小可燃物湿度码 Fine fuel moisture code; WIND: 正午风速 Wind speed; RH: 正午相对湿度 Relative humidity; DC: 干旱码 Drought code; DMC: 半腐层湿度码 Duff moisture code; DSR: 日严重等级 Daily severity rating; FWI: 火灾天气指数 Fire weather index; ISI: 初始蔓延速度 Initial Spread index; DMC: 半腐层湿度码 Duff moisture code; BUI: 可燃物累积指数 Buildup index

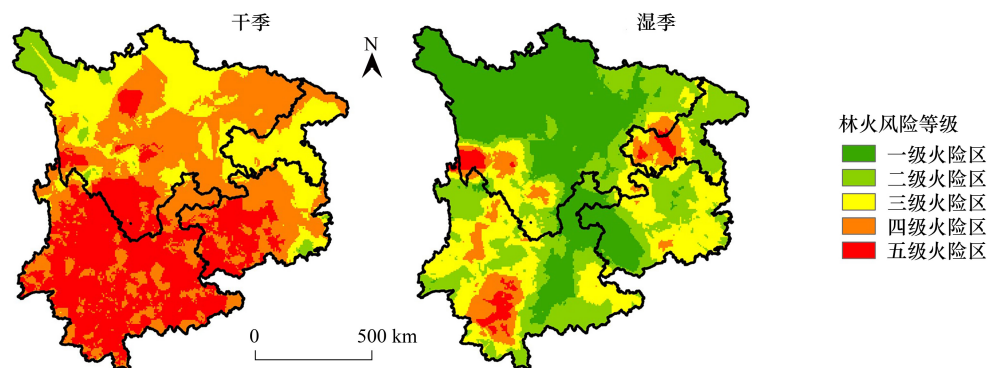


图9 干季、湿季林火风险等级区划

Fig.9 Forest fire risk distribution in dry and wet seasons

0.003)。虽西南地区 96.5%的林火发生于干季,全年模型对干季林火预测精度总体上尚可,但对于湿季林火预测精度明显低于利用湿季林火样本构建的模型,且林火主要驱动因子也存在干湿季节差异(图 7),因此对干湿季分别开展林火预测/预报能有效提高精度及明晰其发生机制。

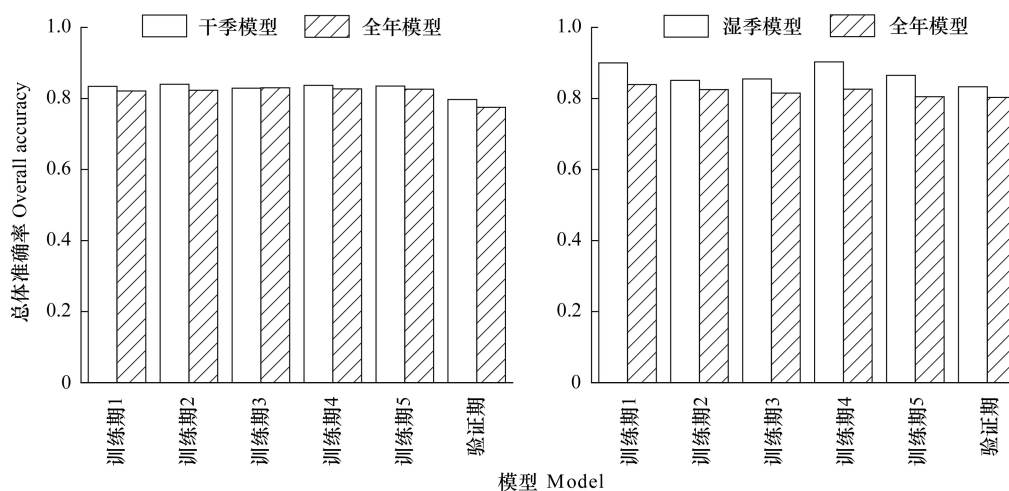


图10 干季、湿季与全年林火预测模型精度比较

Fig.10 Comparison of dry season, wet season and year-round forest fire prediction models

本文引入 FWI 系统指标来构建林火预测模型,该系统在我国西南地区的适用性则成为一个关键问题。为此我们分别提取并计算西南地区 2001—2018 年间干、湿季林火发生频次与 FWI 系统各指标平均值,并做皮尔逊相关性分析。由图 11 可知,FWI 系统各指标与西南地区林火发生频次有较好对应关系,即 FWI 系统各指标对西南地区林火发生有较强的响应,这与一些学者的研究结论相吻合^[45-48]。在利用随机森林构建西南地区林火预测模型时,FWI 系统中的全部和大部分指标也均进入干、湿季林火发生预测模型中,且是除高程、气温之外最为关键的驱动因子,尤其是反映林下可燃物湿度状况的因子(FPMC、DC、DMC)在预测模型中因子排序靠前。综上所述 FWI 系统指标适用于西南林区,在林火预报预测研究中,引入 FWI 系统指标能有效提高预测/预报精度。

火源是林火的重要诱发因素,总体上分为人为火源和自然火源,但囿于 MOD14A1 火点数据难以确定火源,故本研究中未将火源这一因子纳入模型中。本研究发现海拔是干季林火发生的最关键因子,为此我们初步探讨了其可能原因。利用 CMFD 数据计算 1979—2018 年干季、湿季垂直降水梯度,如图 12 所示,该区域

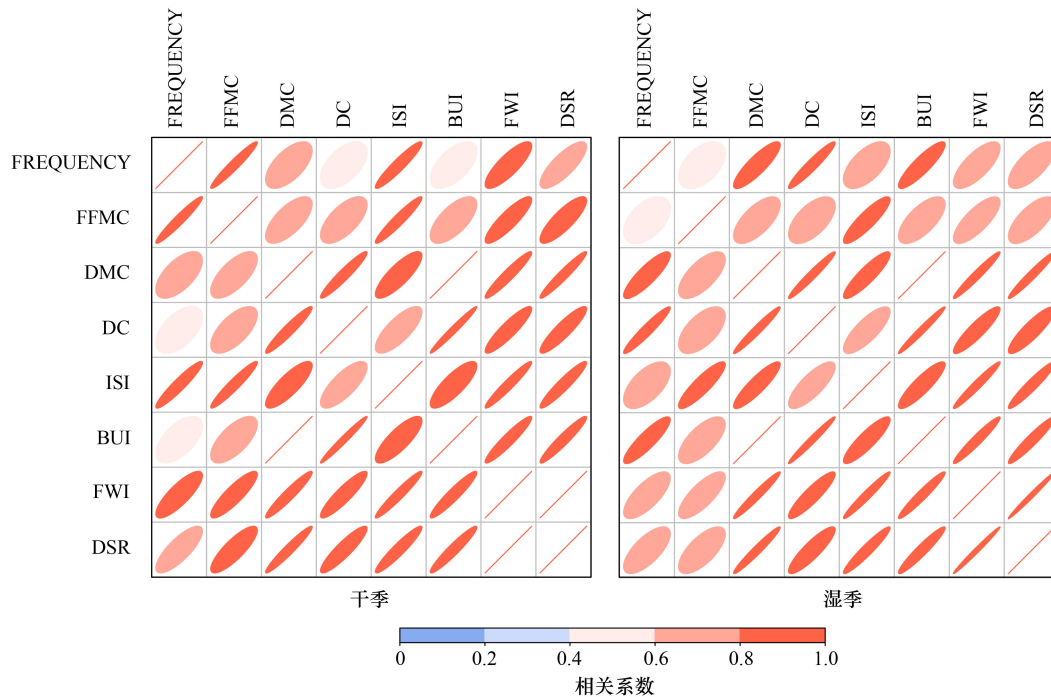


图 11 林火发生次数与 FWI 系统各指标相关系数矩阵

Fig.11 Matrix of correlation coefficients between the frequency of forest fires and the indexes of the FWI system

FREQUENCY: 林火发生次数

干季、湿季降水均随海拔抬高而减少^[49],受降水影响,山麓地带较湿润,应为林火低发区域,林火发生概率随海拔增加应逐渐增大,故降水的垂直分布规律无法完全解释林火的垂直分布特征,因此进一步分析了人类活动情况与海拔的关系。对林火发生的海拔范围(241—4773m)进行垂直分带(步长为 100m,共划分 46 个海拔带),为排除林火发生较少区域的影响,本文选取分析林火发生次数位于下四分位数以上范围内的林火发生情况,对各海拔范围内的林火发生次数与对应的单位面积人口数和 GDP 总量分别做皮尔逊相关性分析,结果表明西南地区林火发生次数与人类活动呈较强的正相关关系,相关系数分别为 0.725 和 0.629(图 12),表明人类集聚和人类活动加强对区域林火有明显的诱发作用,即人为火是 2001—2018 年间林火发生的主要火源。先前的研究也表明西南地区人为火源约占林火总数的 98%左右,自然火源的占比极小^[50],这与本文分析结果一致。因此相关部门应加强对中海拔地区居民的林火防控宣传工作以及合理调配该区域消防资源,尤需重视上述林火发生高风险地区对应高火险时段的林区防火工作,在消防人员和消防资源的调配中,可向这些区域适当倾斜。

4 结论

本文基于 MOD14A1 数据对西南地区 2001—2018 年林火时空变化进行分析,综合考虑地形、气象、FWI 系统中的可燃物状况及火行为指标,利用随机森林算法构建了西南地区干、湿季林火发生预测模型。主要结论如下:

(1) 时间上,2001—2018 年林火次数总体呈现先增加后减少的阶段变化特征,但年内分布不均,干季林火多发,湿季林火发生频次相对较少;空间上,中国西南地区林火呈集聚分布特点,林火集中分布于云南大部、四川西南部及贵州南部地区,而四川东部、重庆大部及贵州北部林火分布较少,总体上呈西多东少,南多北少的空间分布特征。

(2) 根据干、湿季林火分别构建的林火预测模型能够很好地模拟林火发生情况,训练期干、湿季中间模型

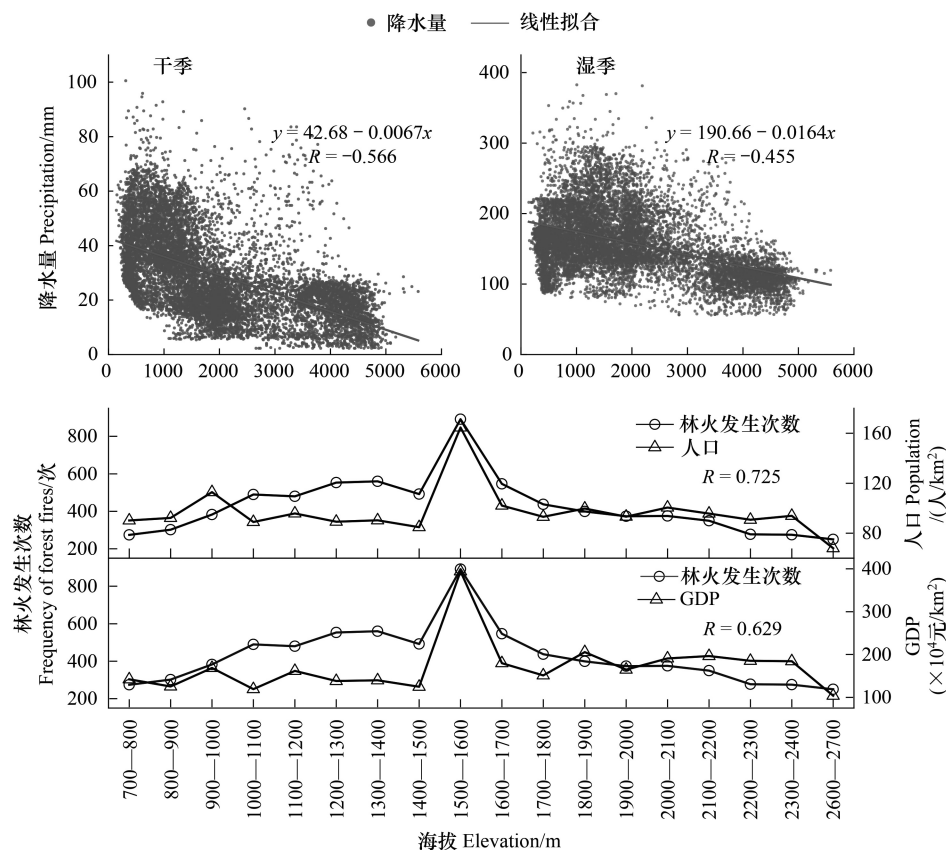


图 12 干季、湿季降水、林火、人口、GDP 的海拔分布特征

Fig.12 Elevation-distribution characteristics of precipitation, forest fires, population and GDP in dry and wet seasons

准确率分别处于 82.94%—83.99% 与 85.12%—90.31% 之间, AUC 值分别处于 0.908—0.914 与 0.922—0.965 之间; 验证期干季、湿季模型准确率分别为 79.73% 和 83.27%, AUC 值分别为 0.886 和 0.855; 两个最终模型与其中间模型表现一致且准确度较高。

(3) 海拔是西南地区林火发生最为关键的限制因子, 这与人类活动密切相关。当日的气象条件是干季林火发生次重要的驱动因子, 可燃物的温湿度状况则是湿季林火发生次重要的驱动因子。FWI 系统指标对于干季林火发生均有重要的影响, 且在西南地区林火预测中有较好适用性。

参考文献 (References):

- [1] Huang F, Liu X N, Yuan J G. Study on forest fire danger model with remote sensing based on GIS. Chinese Geographical Science, 2000, 10(1): 61-67.
- [2] 刘国华, 傅伯杰, 方精云. 中国森林碳动态及其对全球碳平衡的贡献. 生态学报, 2000(5): 733-740.
- [3] Houghton R A. Land-use change and the carbon cycle. Global Change Biology, 1995, 1(4): 275-287.
- [4] 胡汉舟, 孙守亮. 中国环境统计年鉴, 中国统计出版社, 2021, 97-99.
- [5] Wotton B M, Nock C A, Flannigan M D. Forest fire occurrence and climate change in Canada. International Journal of Wildland Fire, 2010, 19(3): 253.
- [6] Camia A, Amatulli G, San-Miguel-Ayaz J. Past and future trends of forest fire danger in Europe. European Commission, Joint Research Centre, 2008.
- [7] Zhou T J, Chen Z M, Chen X L, Zuo M, Jiang J, Hu S. Interpreting IPCC AR6: Future global climate based on projection under scenarios and on near-term information. Advances in Climate Change Research, 2021, 17(6): 652-663.
- [8] Hu Y L, Ji G X, Li J H, Ganjurjav H, HU G Z, GAO Q Z. Interpretation of IPCC AR6: terrestrial and freshwater ecosystems and their services.

- Advances in Climate Change Research, 2022, 18(4): 395-404.
- [9] Jain P, Castellanos-Acuna D, Coogan S C P, Abatzoglou J T, Flannigan M D. Observed increases in extreme fire weather driven by atmospheric humidity and temperature. *Nature Climate Change*, 2022, 12(1): 63-70.
- [10] Walding N G, Williams H T P, McGarvie S, Belcher C M. A comparison of the US National Fire Danger Rating System (NFDRS) with recorded fire occurrence and final fire size. *International Journal of Wildland Fire*, 2018, 27(2): 99-113.
- [11] 蔡奇均, 曾爱聪, 苏漳文, 郭福涛. 基于 Logistic 回归模型的浙江省林火发生驱动因子分析. *西北农林科技大学学报: 自然科学版*, 2020, 48(2): 102-109.
- [12] 梁慧玲, 王文辉, 郭福涛, 林芳芳, 林玉蕊. 比较逻辑斯蒂与地理加权逻辑斯蒂回归模型在福建林火发生的适用性. *生态学报*, 2017, 37(12): 4128-4141.
- [13] 苏漳文, 曾爱聪, 蔡奇均, 胡海清. 基于 Gompit 回归模型的大兴安岭林火预测模型及驱动因子研究. *林业工程学报*, 2019, 4(4): 135-142.
- [14] Jafarzadeh A A, Mahdavi A, Jafarzadeh H. Evaluation of forest fire risk using the Apriori algorithm and fuzzy c-means clustering. *Journal of Forest Science*, 2017, 63(8): 370-380.
- [15] Syafitri N, Labellapansa A, Kadir E A, Saian R, Zahari N N A, Anwar N H K, Shaharuddin N E M. Early detection of fire hazard using fuzzy logic approach. *International Journal of Advanced Computer Research*, 2019, 9(43): 252-259.
- [16] Singh K R, Neethu K P, Madhurekaa K, Harita A, Mohan P. Parallel SVM model for forest fire prediction. *Soft Computing Letters*, 2021, 3: 100014.
- [17] Pham B T, Jaafari A, Avand M, Al-Ansari N, Du T D, Yen H P H, Phong T V, Nguyen D H, Le H V, Mafi-Gholami D, Prakash I, Thuy H T, Tuyen T T. Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry*, 2020, 12(6): 1022.
- [18] 梁慧玲, 林玉蕊, 杨光, 苏漳文, 王文辉, 郭福涛. 基于气象因子的随机森林算法在塔河地区林火预测中的应用. *林业科学*, 2016, 52(1): 89-98.
- [19] Milanović S, Marković N, Pamučar D, Gigović L, Kostić P, Milanović S D. Forest fire probability mapping in eastern Serbia: Logistic regression versus random forest method. *Forests*, 2020, 12(1): 5.
- [20] Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A, Pereira J. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecology and Management*, 2012, 275: 117-129.
- [21] 张琪, 李跃清. 近 48 年西南地区降水量和雨日的气候变化特征. *高原气象*, 2014, 33(2): 372-383.
- [22] Ma W Y, Feng Z K, Cheng Z X, Chen S L, Wang F G. Identifying forest fire driving factors and related impacts in China using random forest algorithm. *Forests*, 2020, 11(5): 507.
- [23] Zhou R W, Zhang Y P, Peng M C, Jin Y Q, Song Q H. Effects of climate change on the carbon sequestration potential of forest vegetation in Yunnan Province, southwest China. *Forests*, 2022, 13(2): 306.
- [24] 贾孜拉·拜山, 李维京, 孙丞虎, 左金清, 张若楠, 刘景鹏. 1961—2014 年我国西南地区干湿季变化特征. *气候变化研究进展*, 2017, 13(2): 103-116.
- [25] 田晓瑞, 赵凤君, 舒立福, 王明玉. 西南林区卫星监测热点及森林火险天气指数分析. *林业科学研究*, 2010, 23(4): 523-529.
- [26] Masocha M, Dube T, Mpofu N T, Chimunhu S. Accuracy assessment of MODIS active fire products in southern African savannah woodlands. *African Journal of Ecology*, 2018, 56(3): 563-571.
- [27] Yin S. Biomass burning spatiotemporal variations over South and Southeast Asia. *Environment International*, 2020, 145: 106153.
- [28] Huo L Z, Boschetti L, Sparks A M. Object-based classification of forest disturbance types in the conterminous United States. *Remote Sensing*, 2019, 11(5): 477.
- [29] 曾爱聪, 蔡奇均, 苏漳文, 郭新彬, 靳全锋, 郭福涛. 基于 MODIS 卫星火点的浙江省林火季节变化及驱动因子. *应用生态学报*, 2020, 31(2): 399-406.
- [30] 张翔. 四川凉山州森林火灾预测模型和火险等级区划研究[D]. 成都: 成都理工大学, 2020.
- [31] 曾爱聪, 郭新彬, 郑文霞, 魏帽, 靳全锋, 郭福涛. 基于 MODIS 卫星火点数据的浙江省林火时空动态变化特征. *北京林业大学学报*, 2020, 42(11): 39-46.
- [32] Guo F T, Wang G Y, Su Z W, Liang H L, Wang W H, Lin F F, Liu A Q. What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests. *International Journal of Wildland Fire*, 2016, 25(5): 505-519.
- [33] 田超, 苏漳文, 杨夏捷, 郭福涛. 福建西部地区林火发生格局及驱动因子分析. *森林与环境学报*, 2017, 37(1): 88-94.
- [34] 谢雯, 寇卫利, 王秋华, 徐伟恒, 谢长波. 楚雄州森林火险综合预报模型研究. *西南林业大学学报: 自然科学*, 2022, 42(1): 151-158.
- [35] Ying L X, Han J, Du Y S, Shen Z H. Forest fire characteristics in China: Spatial patterns and determinants with thresholds. *Forest Ecology and Management*, 2018, 424: 345-354.

- [36] Carvalho A, Flannigan M D, Logan K, Miranda A I, Borrego C. Fire activity in Portugal and its relationship to weather and the Canadian Fire Weather Index System. *International Journal of Wildland Fire*, 2008, 17(3): 328-338.
- [37] 王超. 吉林省林火天气指标系统研究[D]. 哈尔滨: 东北林业大学, 2010.
- [38] Breiman L. Random forests. *Machine Learning*, 2001, 45: 5-32.
- [39] 方匡南, 吴见彬, 朱建平, 谢邦昌. 随机森林方法研究综述. *统计与信息论坛*, 2011, 26(3): 32-38.
- [40] Breiman L. Bagging predictors. *Machine Learning*, 1996, 24: 123-140.
- [41] 赵璠, 舒立福, 杨征夺, 张明远. 西南林区森林火灾火险预报模型评价研究. *林业机械与木工设备*, 2017, 45(5): 15-21.
- [42] 郑忠, 高阳华, 杨庆媛, 唐云辉, 徐永进, 陈艳英. 西南山地区域森林火险综合预报模型研究——以重庆市为例. *自然灾害学报*, 2020, 29(1): 152-161.
- [43] 朱政, 赵璠, 李世友, 秦明明, 邓小凡, 黄鹏桂. 滇中地区林火发生预报模型构建研究. *西北林学院学报*, 2022, 37(5): 146-154.
- [44] 马文苑, 冯仲科, 成竺欣, 王凤阁. 山西省林火驱动因子及分布格局研究. *中南林业科技大学学报*, 2020, 40(9): 57-69.
- [45] 王明玉. 气候变化背景下中国林火响应特征及趋势[D]. 北京: 中国林业科学研究院, 2009.
- [46] 赵凤君, 舒立福, 田晓瑞, 王明玉. 1957—2007年云南省森林火险变化. *生态学杂志*, 2009, 28(11): 2333-2338.
- [47] 杨美清, 姚启超, 方克艳, 董志鹏. 加拿大森林火险天气指数系统在全球及中国的应用. *亚热带资源与环境学报*, 2021, 16(1): 48-54.
- [48] Tian X R, Zhao F J, Shu L F, Wang M Y. Changes in forest fire danger for south-western China in the 21st century. *International Journal of Wildland Fire*, 2014, 23(2): 185-195.
- [49] 陶健, 董建新, 刘光亮, 张戈丽, 朱军涛, 宋文静, 王程栋, 陈爱国, 王树声. 西南不同农业区气温和降水量沿海拔梯度的变化特征. *中国农业气象*, 2016, 37(4): 379-389.
- [50] 张运林, 郭妍, 胡海清. 2001—2017年西南地区森林火灾数据特征分析. *西北林学院学报*, 2021, 36(1): 179-186.