

DOI:10.12171/j.1000-1522.20230161

内蒙古大兴安岭林火驱动因素识别及预测模型

周庆^{1,2,3,4} 张恒^{1,2} 张秋良^{1,2} 赵鹏武¹ 诺敏⁵ 王嘉夫⁶
高健⁷ 赵梦玉⁸ 杨泽华⁸

(1. 内蒙古农业大学林学院, 内蒙古呼和浩特 010019; 2. 内蒙古大兴安岭森林生态系统国家野外科学观测研究站, 内蒙古根河 022350;
3. 南京大学国际地球系统科学研究所, 江苏南京 210023; 4. 江苏省地理信息技术重点实验室, 江苏南京 210023; 5. 克一河森工公司,
内蒙古呼伦贝尔 021000; 6. 内蒙古大兴安岭林业生态研究院, 内蒙古牙克石 022150; 7. 北京市园林绿化局森林防火事务中心,
北京 102100; 8. 内蒙古自治区呼和浩特市气象局, 内蒙古呼和浩特 010020)

摘要:【目的】选择和验证适合研究区的林火预测模型, 明确火灾发生的关键驱动因素并绘制火险区划图, 为森林火灾预防和管理工作提供科学依据和决策支持。【方法】基于 1981—2020 年历史火灾数据, 融合多源(气象条件、地形、植被、人类活动和社会经济)数据, 比较 4 种机器学习方法在内蒙古大兴安岭林火预测中的适用性, 同时, 基于显著影响火灾发生的驱动因素绘制火灾发生可能性地图和火险区划图。【结果】(1)增强回归树模型接受者操作特性曲线下的面积值为 0.967, 随机森林模型的 AUC 为 0.947, 均表现出优异的预测性能。Logistic 回归模型和 Gompit 回归模型的预测准确率较上两种略低, AUC 分别为 0.852、0.851, 也满足研究区的基本预测要求。(2)气象因素气温日较差、日最小相对湿度是影响内蒙古大兴安岭林火发生的主导因素; 海拔在驱动因素的相对重要性排序中位居前列; 人类活动和社会经济因素(如距公路的距离、距火灾瞭望塔的距离、人均 GDP 等)对林火发生也有一定影响。(3)内蒙古大兴安岭东部和东南部存在大面积火灾中、高风险区, 北部中俄边境和西南部中蒙边境也有较高的火灾风险。火灾发生前一年秋季防火期的平均气温、平均地表温度等因素会影响第 2 年森林火灾的发生。【结论】与其他 3 种模型相比, 增强回归树模型是最适合内蒙古大兴安岭林火发生的预测模型。气象因子、海拔显著影响内蒙古大兴安岭林火发生, 人类活动和社会经济因素对火灾发生也有一定的影响。研究区的中高火险区域主要集中在东部和东南部, 北部和西南部也有一定的火灾风险。

关键词: 大兴安岭; 模型比较; 林火发生概率; 林火驱动因素; 火险区划

中图分类号: S762 **文献标志码:** A **文章编号:** 1000-1522(2024)12-0114-12

引文格式: 周庆, 张恒, 张秋良, 等. 内蒙古大兴安岭林火驱动因素识别及预测模型 [J]. 北京林业大学学报, 2024, 46(12): 114-125. Zhou Qing, Zhang Heng, Zhang Qiuliang, et al. Identification and prediction models of driving factors for forest fires in Daxing'an Mountains of Inner Mongolia, northern China [J]. Journal of Beijing Forestry University, 2024, 46(12): 114-125.

Identification and prediction models of driving factors for forest fires in Daxing'an Mountains of Inner Mongolia, northern China

Zhou Qing^{1,2,3,4} Zhang Heng^{1,2} Zhang Qiuliang^{1,2} Zhao Pengwu¹ Nuo Min⁵ Wang Jiafu⁶
Gao Jian⁷ Zhao Mengyu⁸ Yang Zehua⁸

(1. College of Forestry, Inner Mongolia Agricultural University, Hohhot 010019, Inner Mongolia, China;

2. Forest Ecosystem National Observation and Research Station of Greater Khingan Mountains,
Genhe 022350, Inner Mongolia, China;

3. International Institute for Earth System Science, Nanjing University, Nanjing 210023, Jiangsu, China;

4. Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing 210023, Jiangsu, China;

5. Keyihe Forestry Company, Hulunbuir 021000, Inner Mongolia, China;

收稿日期: 2023-06-27 **修回日期:** 2023-10-24

基金项目: 国家自然科学基金项目(32060344), 内蒙古自治区高等学校青年科技英才支持计划(NJYT24042), 内蒙古自治区科技计划(2022YFSH0077), 中央高校基本科研业务费专项(BFUKF202217), 内蒙古自治区科技计划(2023KYPT0001)。

第一作者: 周庆。主要研究方向: 森林草原火灾生态与管理。Email: wilsonzhou1224@emails.imau.edu.cn 地址: 010019 内蒙古呼和浩特市学苑东街 275 号。

责任作者: 张恒, 博士, 副教授。主要研究方向: 森林草原火灾生态与管理。Email: zhangheng_nefu@imau.edu.cn 地址: 同上。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

6. Inner Mongolia Daxing'an Mountains Forestry Science and Technology Research Institute,

Yakeshi 022150, Inner Mongolia, China;

7. Beijing Municipal Forest and Parks Bureau, Beijing 102100, China;

8. Hohhot Meteorological Service in Inner Mongolia Autonomous Region, Hohhot 010020, Inner Mongolia, China)

Abstract: [Objective] This paper aims to select and validate suitable forest fire prediction models for the study area, identify key driving factors of fire occurrence, and map fire risk zoning, then providing scientific basis and decision support for forest fire prevention and management. [Method] Using historical fire data from 1981 to 2020 and integrating multi-source data (meteorological conditions, topography, vegetation, human activities, and socio-economic factors), the applicability of four machine learning methods in predicting forest fires in the Daxing'an Mountains of Inner Mongolia of northern China was compared. Based on the significant factors influencing fire occurrence, maps of fire occurrence probability and fire risk zoning were generated. [Result] (1) The boosted regression tree model (BRT) showed an area under the curve (AUC) value of 0.967, and the random forest model (RF) achieved an AUC of 0.947, both demonstrating excellent predictive performance. The predictive accuracy of the Logistic regression model (LR) and the Gompit regression model (GR) was slightly lower than former two models, but still met the basic predictive requirements for the study area, with AUC values of 0.852 and 0.851, respectively. (2) Meteorological factors, such as diurnal temperature range and daily minimum relative humidity, were the dominant factors influencing forest fires in the Daxing'an Mountains of Inner Mongolia. Elevation also ranked high in the relative importance of driving factors. Human activities and socio-economic factors, such as distance to roads, distance to fire lookout towers, and per capita GDP, also had some influence on fire occurrence. (3) Large areas of medium to high fire risk were present in the eastern and southeastern parts of the Daxing'an Mountains of Inner Mongolia, while the northern China-Russia border and the southwestern China-Mongolia border also exhibited elevated fire risk. Factors such as average temperature and average surface temperature during fire prevention period in autumn of previous year influenced forest fire occurrences in the following year. [Conclusion] Among the four models compared, the BRT was identified as the most suitable one for predicting forest fire occurrence in the Daxing'an Mountains of Inner Mongolia. Meteorological factors and elevation significantly influence fire occurrence, while human activities and socio-economic factors also have a certain impact on the occurrence of fires. The high and medium fire risk areas are primarily concentrated in the eastern and southeastern parts of the study area, with some fire risks presenting in the northern and southwestern regions.

Key words: Daxing'an Mountains; model comparison; probability of forest fires; driving factors of forest fires; fire risk zoning

森林火灾会导致生态系统的结构和功能发生变化^[1],对物种组成、结构、演替和再生具有长期影响。同时,火灾会对社会系统构成严重威胁,造成人员伤亡和财产损失。森林火灾的发生机制对人类活动和其他变化的驱动因素非常敏感^[2],所以森林火灾管理需要了解火灾发生模式的空间特征,并采用量化的方法评估区域范围内各种驱动因素的相对重要性^[3]。建立森林火灾预测模型,可以预测火灾发生的概率,进而对可能发生火灾的区域进行管理,这对公共安全、森林管理和扑救规划具有重要意义。

林火发生受多种因素影响,具有典型的非线性和复杂特征。选择合适的林火预测模型是目前一项艰巨的任务^[4-5]。Logistic 回归模型(Logistic regression, LR)是一种广泛应用于森林火灾预测的统计方法,可以预测不同地理变量下火灾发生的概率^[6]。在日常

应用中,Gompit 回归模型(Gompit regression, GR)^[7]也称互补对数-对数模型,是处理非对称结构数据的一个有效选择。苏漳文等^[8]首次使用 GR 模型对中国大兴安岭地区的林火发生进行研究,而高超等^[9]指出,目前还没有对 Logistic 回归模型和 Gompit 回归模型预测效果的对比研究。随机森林模型(random forest model, RF)被认为是一种灵活的方法,能评估变量之间的复杂交互性,还可以克服过度拟合的问题^[10]。同时也有多项研究表明,随机森林模型比其他林火预测模型的预测效果更好^[11-12]。增强回归树模型(boosted regression tree model, BRT)在多项关于林火预测的研究中也取得了不错的建模效果^[13-14]。尽管这几种模型在林火预测中得到了很好的应用,但目前对这几种模型在林火预测适用性方面的对比研究还相对较少。

Massada 等^[15]在构建美国密歇根州的林火预测模型时发现,随机森林模型和最大熵模型的预测准确度略高于广义线性模型;Eskandari 等^[11]在伊朗戈列斯坦省的研究中指出,随机森林模型和一种新的混合模型(RF-CART-BRT)是该地区林火预测的最佳模型。我国许多地区也开展了林火预测模型的研究,但主要集中在南方林区和黑龙江大兴安岭,关于内蒙古大兴安岭的研究较少。除此之外,最适合该地区的林火预测模型也没有确定。

内蒙古大兴安岭森林面积为 8.37 万 km², 占整个大兴安岭面积的 46%, 森林覆盖率达 78.39%, 均居全国国有林区之首^[16], 区内林草、林木相连, 林农交错, 火源管理难度大, 极易发生森林火灾。鉴于此, 本文基于历史火灾数据(1981—2020 年), 综合考虑气候、植被、地形等多种可能影响林火发生的驱动因素, 构建 4 种预测模型。对这些模型的拟合结果进行比较分析, 选择最适合内蒙古大兴安岭的林火预测模型, 并识别及排序影响林火发生的主要驱动因素, 旨在为林火预测模型的选择提供理论支撑, 为森林防火部门和应急管理部门优化林火管理策略提供科学依据。

1 研究区概况与研究方法

1.1 研究区概况

内蒙古大兴安岭位于中国东北部, 总面积约 1 773 万 hm², 海拔 165 ~ 1 741 m(图 1)。属于寒温带大陆性季风气候, 冬季寒冷漫长(9 个月以上), 夏季温暖短暂(不到 1 个月), 无霜期短(70 ~ 100 d)。年平均降水量 442 mm, 其中 85% ~ 90% 的降水发生在夏季(6—8 月), 10% 的降水发生在 10 月底至次年 4 月初, 通常为降雪。年平均相对湿度为 70%。

主要树种有兴安落叶松(*Larix gmelinii*)、白桦(*Betula platyphylla*)、山杨(*Populus davidiana*)和樟子松(*Pinus sylvestris* var. *mongolica*)等^[17]。内蒙古大兴安岭在不同历史时期, 一直是森林火灾高发地区。据统计, 1962—1980 年共发生森林火灾 1 344 起, 受害森林面积 124.35 万 hm²^[18]。1980 年之后的火灾发生情况也较为严重。内蒙古大兴安岭一年内共有两个防火期, 春季防火期: 3 月 15 日—6 月 15 日; 秋季防火期: 9 月 15 日—11 月 15 日。

1.2 森林火灾数据

森林火灾数据由内蒙古自治区林业和草原局提供, 包含了每次森林火灾发生的详细信息, 包括地理位置、起火时间、火灾规模、火灾原因、伤亡人数等。1981—2020 年共记录了 2 552 起火灾, 其中 211 起火灾没有地理坐标信息, 未列入本文的研究。火点的地理分布格局如图 2a 所示。使用二进制目标变量, 将已知火点赋值为 1, 在研究区域内利用 ArcGIS10.7 创建随机点并赋值为 0(为了防止数据过分离散, 通常要以接近或多于火点的比例创建随机点, 参考前人研究, 随机点设置为火点的 1.5 倍^[19])。为了避免创建的随机点与已知火点重合或者邻近, 设置任意两个随机点的最小距离为 1 km, 并且根据实际火烧的时间分布概率选取随机点的时间。最终得到 2 341 个火点与 3 519 个随机点(图 2)。

1.3 驱动因子数据

(1)气象数据。火灾发生前 1 年防火期的气象因素显著影响未来火灾发生^[20]。本文选取火灾发生前 1 年春季防火期和秋季防火期的平均气温、平均地表温度等指标(表 1)。气象数据来源于温室数据共享平台。使用 ArcGIS10.7 匹配样本点和与其距离最近的气象站点(图 1), 将火点和非火点与气象站点

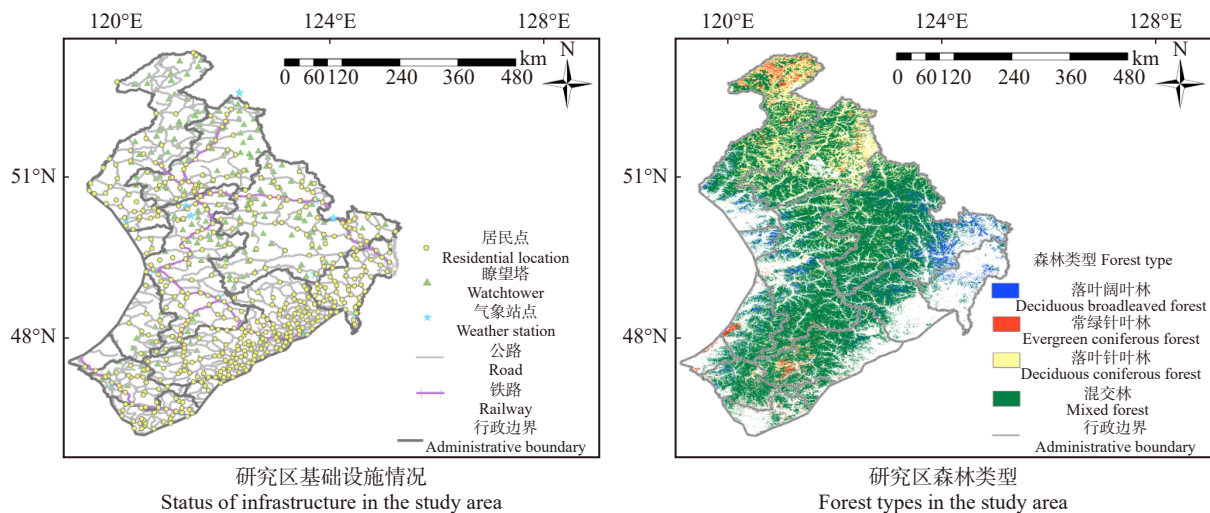


图 1 研究区的地理位置及森林类型

Fig. 1 Geographical location of the study area and forest types

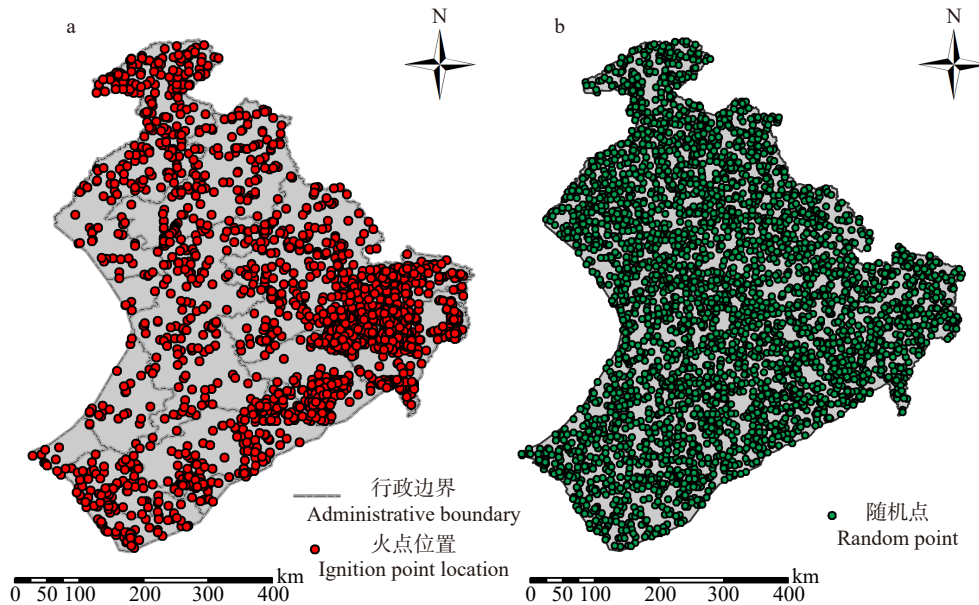


图 2 历史火点位置(a)与随机点的位置(b)

Fig. 2 Historical fire point location (a) and random point location (b)

相对应,在 SQL Sever 数据库中根据样本点对应的气象站点和日期匹配气象数据。

(2) 植被数据。植被类型是导致火灾发生的主要因素之一^[21]。从数据共享服务系统网站(<https://data.casearth.cn/>)下载了 1985—2020 年每 5 年全球 30 m 地表覆盖精细分类产品。为确保随机点均位于森林范围内,在建立随机点时,选择与地表分类数据年份相近的随机点,并提取其对应的植被类型。归一化植被指数(normalized difference vegetation index, NDVI),本文使用空间分辨率 8 km 的 GIMMS NDVI 3g v1.0 数据(NASA, <https://www.nasa.gov/nex>, 时间分辨率为 15 d),并使用 R 语言对这些数据进行月最大值合成。由于最早的 NDVI 数据是从 1981 年 7 月开始的 GIMMS NDVI 数据,因此对于 1981 年 1—6 月缺失的数据,使用 1982 年同月数据替代,并假设一年内的 NDVI 变化不大。

(3) 地形数据。地形数据来源于地理空间数据云(<https://www.gscloud.cn/>)的 GDEM V2 30 m 分辨率的数字高程模型(digital elevation model, DEM)。将 DEM 进行空间校正、拼接、剪辑和重采样等预处理。使用 ArcGIS 10.7 软件中的 Spatial Analyst 工具包,利用“Surface”工具提取研究区的坡度和坡向数据。根据公式计算坡向指数 i ^[21]。

$$i = \cos(2\theta \times \pi/360) \quad (1)$$

式中: θ 是由 ArcGIS 10.7 “Aspect”函数得出的坡向,范围为 $0^\circ \sim 360^\circ$ 。坡向指数范围 $-1 \sim 1$,数值越高,表示潜在的太阳辐射越高。最后,用“Extract”工具来提取样本点的海拔、坡度和坡向。

(4) 基础设施数据。从全国地理信息资源目录服务系统下载了 1:250 000 数字线图(<https://www.webmap.cn/main.do?method=index>),从地图中检索火点到最近铁路、公路、居民点和瞭望塔的距离。基础设施的分布如图 1 所示。

(5) 人口和社会经济数据。1981—2008 年人口和社会经济数据来自内蒙古自治区统计局出版的统计年鉴^[22],2009—2020 年数据来源于内蒙古自治区统计局官网(http://tj.nmg.gov.cn/tjyw/jpsj/index_1.html),得到 1981—2020 年内蒙古大兴安岭区域以县级为单位(共 13 个县)的人口密度与人均 GDP。

1.4 模型

选用 Logistic 回归模型(LR)^[23]、Gompit 回归模型(GR)、随机森林模型(RF)和增强回归树模型(BRT)进行林火预测与精度比较。在建模时,将全部样本随机划分为训练数据集(70%)和测试数据集(30%),并将训练数据集再次随机划分为 70% 的模型训练样本和 30% 的模型检验样本,4 种模型均重复 5 次,建立 5 个中间模型以减少随机划分对模型的影响。4 种模型使用同样的训练数据集和测试数据集,但每种模型在建模时重新划分训练样本和模型检验样本。本文基于 R studio 4.1.3 的“randomForest”包和“gbm”包构建了 RF 模型和 BRT 模型。BRT 模型涉及的相关参数:决策树复杂度(tree complexity)为 5,学习速率(learning rate)为 0.01,袋子分数(bag fraction)为 0.05,树数(number of trees)为 4 750。

1.5 识别重要变量

本文使用方差膨胀因子(variance inflation factor, VIF)对 LR、GR 模型的自变量进行多重共线

表1 变量的简写及其单位

Tab. 1 Abbreviation of variables and their units

因素 Factor	变量 Variable	简写 Abbreviation	单位 Unit
气象因素 Meteorological factor	日平均气温 Average daily temperature	Temp	°C
	日最高气温 Maximum daily temperature	Max_temp	°C
	气温日较差 Daily temperature range	Temp_diff	°C
	日均相对湿度 Daily average relative humidity	Hum	%
	日最小相对湿度 Daily minimum relative humidity	Minhum	%
	日均地表温度 Daily average surface temperature	G_temp	°C
	日最高地表温度 Daily maximum surface temperature	Maxg_temp	°C
	日均降水量 Daily mean precipitation	Prec	mm
	日均风速 Daily mean wind speed	Win	m/s
	日照时数 Sunshine hours	Sun	h
	月均气温 Monthly mean temperature	Meantemp	°C
	月均降水量 Monthly mean precipitation	Meanprec	mm
	月均日照时数 Monthly mean sunshine hours	Meansun	h
	月均湿度 Monthly mean humidity	Meanhum	%
	月均地表温度 Monthly mean surface temperature	Mg_temp	°C
	春季防火期平均气温 Average temperature during spring fire season	TempSpr	°C
	春季防火期平均湿度 Average humidity during spring fire season	HumSpr	%
	春季防火期平均地表温度 Average surface temperature during spring fire season	G_tempSpr	°C
	春季防火期平均降水量 Average precipitation during spring fire season	PrecSpr	mm
	春季防火期平均日照时数 Average sunshine hours during spring fire season	SunSpr	h
人为因素 Human factor	秋季防火期平均气温 Average temperature during fall fire season	TempAut	°C
	秋季防火期平均地表温度 Average surface temperature during fall fire season	G_tempAut	°C
	秋季防火期平均湿度 Average humidity during fall fire season	HumAut	%
	秋季防火期平均降水量 Average precipitation during fall fire season	PrecAut	mm
	秋季防火期平均日照时数 Average sunshine hours during fall fire season	SunAut	h
	与最近居民点的距离 Distance to the nearest settlement	Dis_res	km
	与最近道路的距离 Distance to the nearest road	Dis_road	km
	与最近铁路的距离 Distance to the nearest railroad	Dis_rail	km
植被状况 Vegetation status	与最近瞭望塔的距离 Distance to the nearest watchtower	Dis_watch	km
	植被类型 Vegetation type	Vegetation type	
地形信息 Topographic information	归一化植被指数 Normalized difference vegetation index	NDVI	
	海拔 Altitude	Dem	m
	坡向 Aspect	Aspect	
社会经济因素 Socio-economic factor	坡度 Slope	Slope	°
	人均 GDP GDP per capita	GDP	万元/km ² 10 ⁴ CNY/km ²
	人口密度 Population density	Pop	人/km ² person/km ²

注: 春季、秋季防火期相关因子指火灾发生前一年的数据。Note: spring and fall fire season weather factors refer to data from the year prior to the fire.

性检验, 以 $VIF = 10$ 为标准, 逐步剔除 $VIF > 10$ 的具有显著共线性的自变量。在 LR、GR 回归模型建模时, 标准化系数常被作为量化变量重要性的一种方法^[6,8], 而本文在构建 LR、GR 模型时使用了充足水平(Adequacy)^[24]来识别重要预测变量。充足水平

是指在统计模型中, 每个预测变量单独对模型对数似然值的贡献比例。它量化了每个预测变量在解释模型整体拟合度方面的相对重要性。

RF 模型在量化变量重要性时使用了 R studio 4.1.3 “randomForest”包中的“importance”功能, 得到每

个变量的平均准确率降低度。BRT 模型使用“gbm”包中的“summary”功能求得变量的相对重要性。

1.6 模型精度评价

采用接受者操作特性曲线(receiver operating characteristic curve, ROC)进行模型精度评价。ROC 曲线下的面积值(area under the curve, AUC)用于评价模型预测的准确性。AUC 为 0.5~0.7 表示拟合效果差, 0.7~0.9 表示拟合效果一般, 0.9 以上表示模型的拟合效果非常好。同时, 通过 ROC 的敏感性和特异性计算 Yoden 指数(灵敏度+特异性-1), 确定最佳临界值, 也就是最佳阈值(cut-off)。这一指数可以作为林火预测模型判断林火发生概率的阈值^[6]。

1.7 边际效应

为了量化变量对森林火灾发生概率的影响, 并确定其影响区间, 使用 R 语言中 gbm.plot 函数绘制边际效应图, 展示森林火灾发生概率(因变量)与显著影响林火发生的驱动因素(自变量)之间的关系。边际效应曲线可以直观展示自变量对因变量的非线性影响, 边际效应曲线递增表明随着自变量取值增加, 因变量预测值也会相应增加。相反, 如果边际效应曲线呈递减趋势, 则表明随着自变量值的增加, 因

变量的预测值会减少。

2 结果与分析

2.1 4 种林火模型对重要驱动因素的选择及模型精度比较

经过多重共线性诊断和显著性检验后, LR、GR 和 RF 模型选择在 5 个中间模型中出现 3 次或 3 次以上的自变量进入全样本数据进行拟合。对于 BRT 模型, 对 5 个训练样本进行变量重要性测试, 放弃不重要变量, 调整参数使模型精度达到最高, 去除 5 个中间模型中被放弃 4 次或者 5 次的变量。选择出的具体因子见图 3。

RF、BRT 模型的 AUC 值均在 0.90 以上, 预测准确率均在 0.85 以上, 其 AUC 值和预测准确率比其他 2 种回归模型高。LR、GR 模型的 AUC 值在 0.85 左右, 预测准确率均在 0.75 以上, 满足林火预测建模的基本要求(表 2)。

2.2 驱动因子重要性排序

在 LR 和 GR 模型中, 日最小相对湿度的重要性均排在首位(图 3), RF 和 BRT 模型中, 日平均相对湿度的综合排序相对靠前, 表明日相对湿度是影响

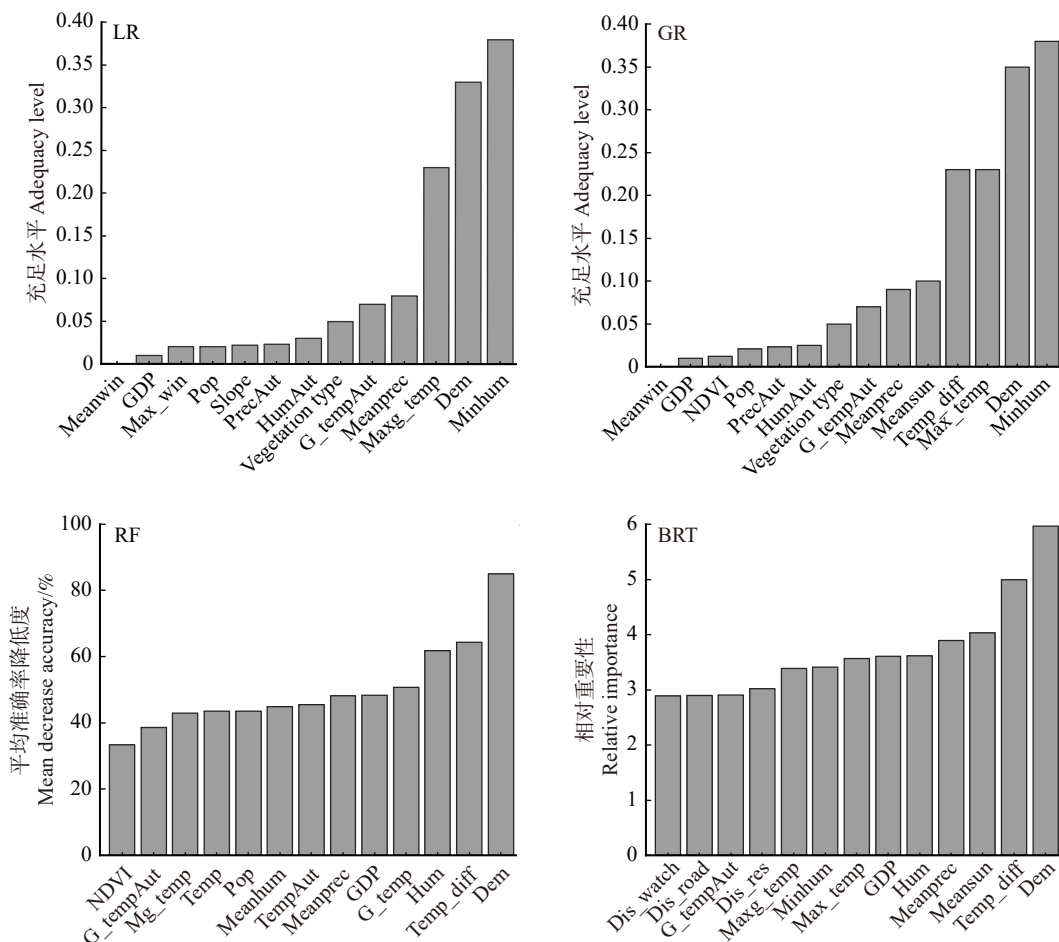


图 3 4 种模型重要因子的排序

Fig. 3 Ranking of important factors of four models

表2 4种模型预测准确性和拟合度的比较
Tab. 2 Comparison of prediction accuracy and fit of four models

样本 Sample	模型 Model	最佳临界值 Optimal critical value	AUC值 AUC value	预测准确率 Prediction accuracy	
				训练数据集 Training data set	测试数据集 Test data set
样本1 Sample 1	LR	0.378	0.855	0.776	0.842
	GR	0.307	0.847	0.772	0.839
	RF	0.338	0.949	0.874	0.876
	BRT	0.426	0.959	0.895	0.897
样本2 Sample 2	LR	0.318	0.854	0.758	0.854
	GR	0.352	0.849	0.771	0.852
	RF	0.320	0.943	0.867	0.894
	BRT	0.431	0.969	0.904	0.921
样本3 Sample 3	LR	0.345	0.843	0.760	0.863
	GR	0.313	0.850	0.758	0.837
	RF	0.320	0.943	0.867	0.876
	BRT	0.358	0.970	0.901	0.911
样本4 Sample 4	LR	0.361	0.856	0.770	0.857
	GR	0.347	0.857	0.774	0.841
	RF	0.292	0.946	0.867	0.892
	BRT	0.405	0.967	0.903	0.905
样本5 Sample 5	LR	0.346	0.858	0.773	0.872
	GR	0.350	0.855	0.777	0.857
	RF	0.300	0.939	0.866	0.887
	BRT	0.388	0.969	0.900	0.900
全样本 Full samples	LR	0.388	0.852	0.755	
	GR	0.400	0.851	0.769	
	RF	0.281	0.947	0.863	
	BRT	0.349	0.967	0.899	

火灾发生的关键驱动因素。气温日较差在 GR、RF 和 BRT 模型中的重要性排序位于前列,月平均日照时数和月平均降水量在 4 种模型中至少被选中的其中一种,且重要性也相对较高。前一年秋季防火期的平均气温和秋季防火期平均地表温度均对火灾事件具有显著影响。值得注意的是,海拔这一因素同时被 4 种模型识别为重要的预测变量之一,且排序靠前,表明海拔对火灾风险也具有关键性的影响。此外,人口密度和人均 GDP 在各模型中被选为重要变量,这两个指标反映了人类活动强度与经济发展水平,强调了社会经济因素在火灾风险评估中的重要性。

2.3 火灾发生概率图和火灾风险等级图

基于 4 种模型选择的重要变量,绘制火灾发生概率图(图 4)。林火发生概率为 0~1,网格值越大,表示火灾发生的概率越高。根据火灾发生概率图进一步绘制火险等级图(图 5)。将火险区域划分为 5 个等级(0~0.2 为极低风险,0.2~0.4 为低风险,0.4~0.6 为中风险,0.6~0.8 为高风险,0.8~1 为极

高风险)。4 种模型的火灾中高风险区基本相同,主要集中在研究区的东部和东南部。北部中俄边境区域和西南部中蒙边境区域也存在部分中高火险区,这表明入境火对内蒙古大兴安岭地区构成一定威胁。

Logistic 和 Gompit 回归模型预测结果显示,每个火险等级面积占比大致相同。而 RF 和 BRT 模型预测结果较为相似。然而,RF 和 BRT 模型预测的极低风险区面积大于 2 种回归模型的预测结果,而它们的极高风险区面积略高于 2 种回归模型的预测结果。

为更好地评估模型的拟合优度,比较预测概率与实际观测值,本研究计算了每个模型的残差,并绘制了残差分布图(图 6)。RF 和 BRT 模型的总体残差比 LR 和 GR 模型小,表明这两种模型具有更好的拟合优度。LR 和 GR 回归模型的残差分布主要集中在阴性残差区域,这意味着这两个模型倾向于过度预测(即模型预测的火灾发生概率高于实际发生概率)。在两种回归模型中,LR 的总体残差略小于 GR。此

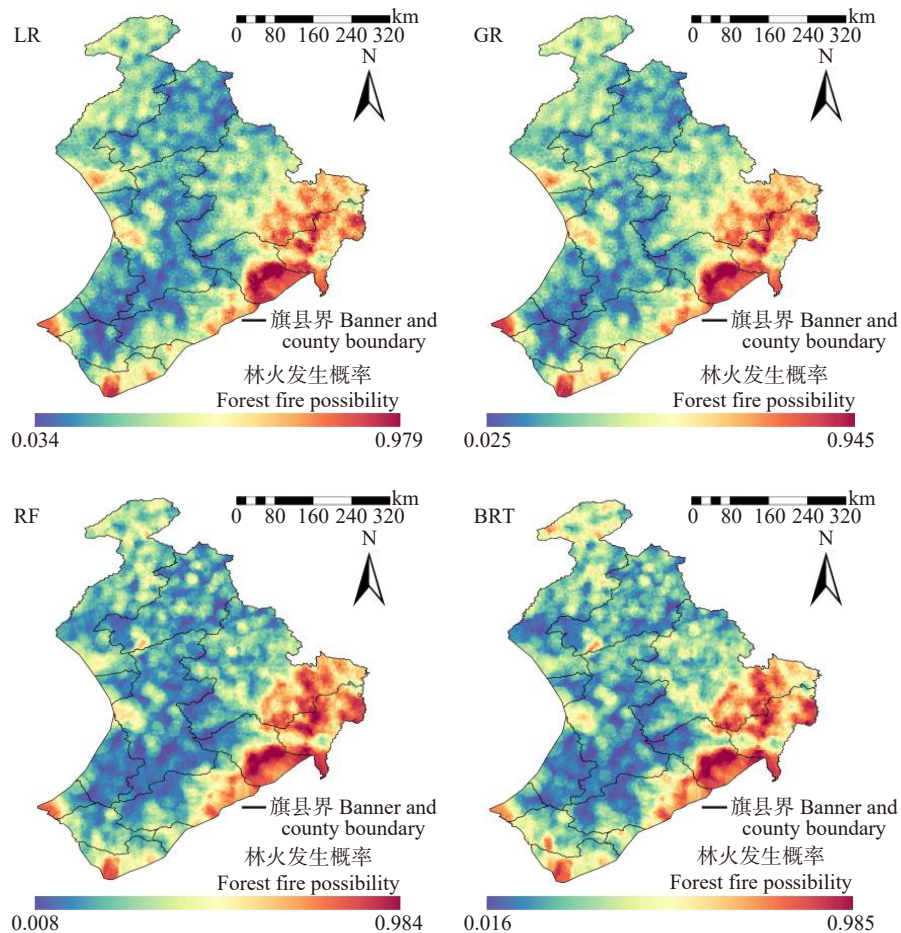


图 4 基于 4 种模型绘制的火灾发生概率

Fig. 4 Fire occurrence probability map based on four models

外, 4 种模型均显示研究区的南部存在一个阳性残差聚集区域, 表明在该区域, 模型预测的火灾发生概率低于实际发生概率, 即存在预测不足的现象。

2.4 林火驱动因素的边际效应

在全样本数据条件下, 本研究使用预测精度最高的 BRT 模型的拟合结果, 绘制了各重要驱动因素与林火发生之间的边际效应图(图 7)。在这些图中, 纵坐标代表边际效应, 其值越高, 表明火灾发生概率越大。海拔是显著影响火灾发生的一个重要因素, 从 0 到 900 m 递增时, 火灾发生概率显著下降, 低海拔地区更容易发生森林火灾。当气温日较差(Temp_diff)在 0~10℃ 时, 林火发生概率变化不大, 当其超过 10℃, 林火发生概率会显著增加。日均相对湿度(Hum)和日最小相对湿度(MinHum)影响林火发生概率有相同的趋势, 这两个指标在超过 20% 时, 火灾发生概率逐步下降。日最高气温(Max_temp)和日最高地表温度(Maxg_temp)影响林火发生概率也有相同的趋势, 日最高气温超过 15℃、日最高地表温度超过 20℃ 后, 林火发生概率显著上升。基础设施方面如距最近道路距离(Dis_road)在 0~3 km 内, 火灾发生概率逐渐下降,

3~10 km 开始上升, 10 km 以上保持稳定。距最近瞭望塔距离(Dis_watch)在 0~70 km 时, 火灾发生概率虽然有所波动, 但总体上保持逐步上升的趋势。

3 讨 论

3.1 影响林火发生的驱动因素

本研究发现, 低海拔地区森林火灾的发生概率更高(图 7), 主要原因是人类的密集活动往往集中在低海拔地区, 从而会增加人为引起火灾的可能性。随着海拔的升高, 植被覆盖和土壤湿度发生变化, 不利于火灾发生^[25]。有研究表明, 一天中气温日较差越大, 气候越干燥, 森林火险也随之增加^[26], 这与本文研究的结果一致。当环境中的相对湿度(日均相对湿度、日最小相对湿度)较低时, 气候会比较干燥, 可燃物含水率降低, 火灾发生的概率增加。本文还发现, 火灾发生概率会随着日最高气温、日最高地表温度的增加而增加。地表温度由气温和日照时数共同影响, 气温高或日照时数长都会导致地表温度增加, 从而使火灾发生概率增加^[27]。本研究发现, 前一年秋季防火期气象要素是影响林火发生的重要因素。RF 模型将前一年秋季防火期平均气温, LR、

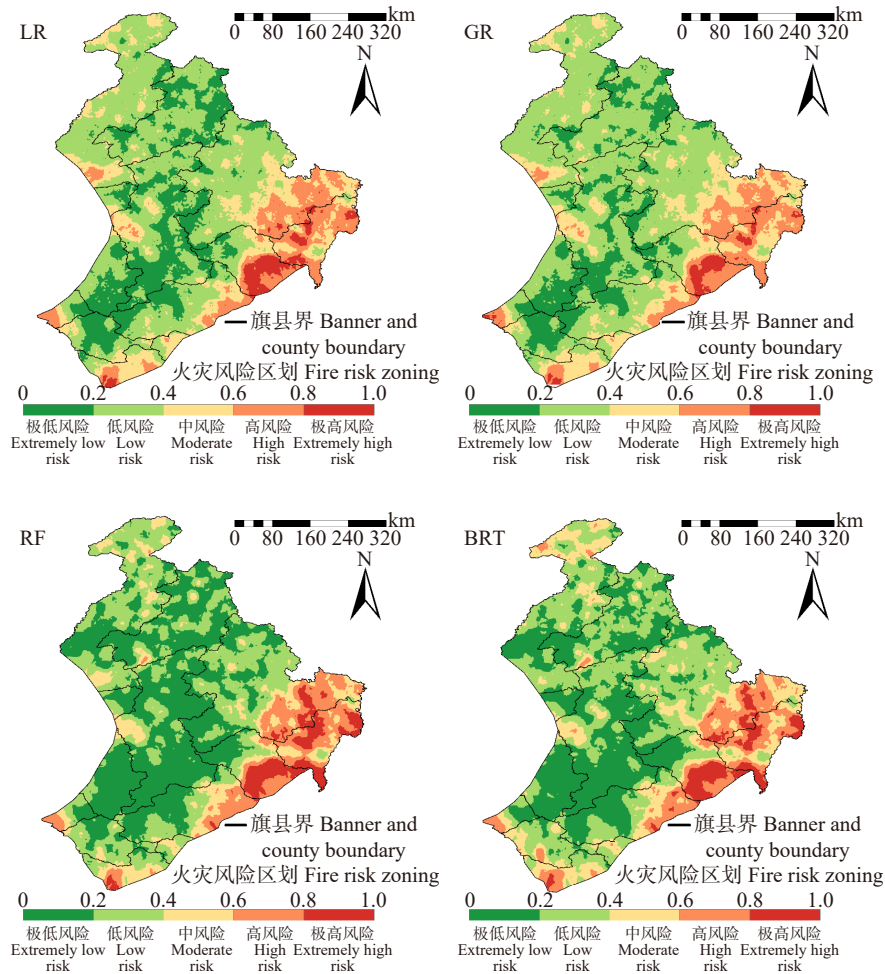


图5 基于4种模型绘制的火险等级

Fig. 5 Fire risk level map based on four models

GR 模型将前一年秋季防火期平均湿度、平均降水量识别为重要变量。以前一年秋季防火期平均地表温度为例,随着地表温度的增加,火灾发生概率也在逐步上升。舒展^[28]指出,如果在冬季非防火期气温比同期显著升高,可能会导致可燃物干燥程度的升高和第二年春季融雪日期的提前,从而使有利于森林火灾发生的风险条件提前出现。在制定未来森林防火策略时,林业管理部门不仅要关注当年的防火工作,还应密切关注上一年秋季防火期的相关气象指标。距瞭望塔的距离是驱动火灾发生的重要因素,多数火点发生的位置距瞭望塔距离较远,这也从侧面反映了一些人类活动稀少的地区,消防资源配置程度较低,导致这些区域的火灾预防和响应能力相对较弱。因此,瞭望塔的合理布局对于提前发现并预防林火的发生极为关键。当火灾不幸发生并蔓延至这些偏远森林时,由于消防资源的不足,林火可能会持续更长时间,造成更大的破坏^[29]。

3.2 不同模型的拟合优度比较及适用性

大量研究表明,RF 模型在林火发生概率预测上具有一定的优势,并且具有广泛的适用性^[9]。

Oliveira 等^[10]研究表明,RF 模型的可解释方差百分比高达 96.3%,而 LR 模型仅为 44%。Guo 等^[30]研究显示,在林火预测的总体准确率上,LR 模型为 60.8%,而 RF 模型达到了 70.1%。在我们的研究中,RF 模型在预测能力和拟合优度上也优于其他两种回归模型,这与前述学者的研究结果相一致。进一步比较 BRT 和其他 3 种模型发现,BRT 模型在预测精度和拟合优度上比其他 3 种模型表现更好。LR 模型与 GR 模型的预测精度相差不大,但在总体残差方面,LR 模型略胜一筹。本文非火点设置为火点数的 1.5 倍,GR 模型的预测精度和拟合优度并没有高于 LR 模型。苏漳文等^[8]使用 Gompit 回归模型进行林火预测时,发现当火点与非火点的比例为 1:3 和 1:4 时,预测准确率最高,达到 81.8%。而比例为 1:1 和 1:2 时,预测准确率并没有提高,均为 79.9%。在日常应用中,我们常用的建模比例为 1:1 和 1:1.5,因此在这些情况下,使用 LR 模型可能更为合适。此外,火点在不同空间位置存在环境异质性,会导致影响火灾发生因素的相对重要性存在差异,这种现象称为空间非平稳性。GR 可以有效解决空

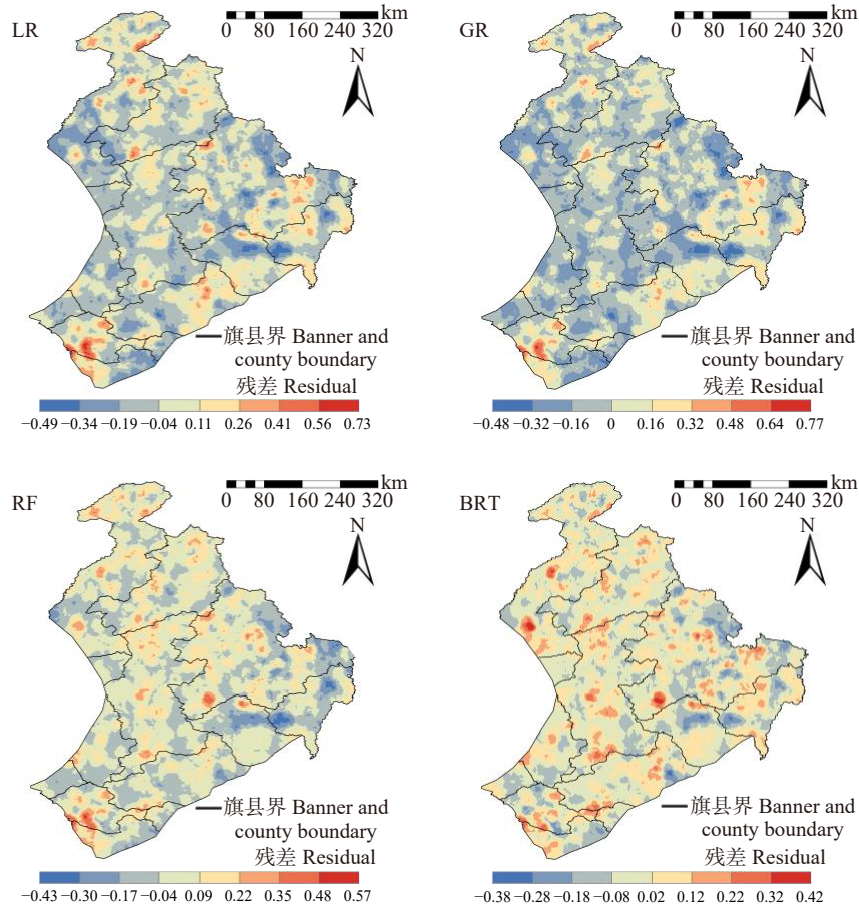


图 6 4 种模型的残差分布

Fig. 6 Residual distribution of four models

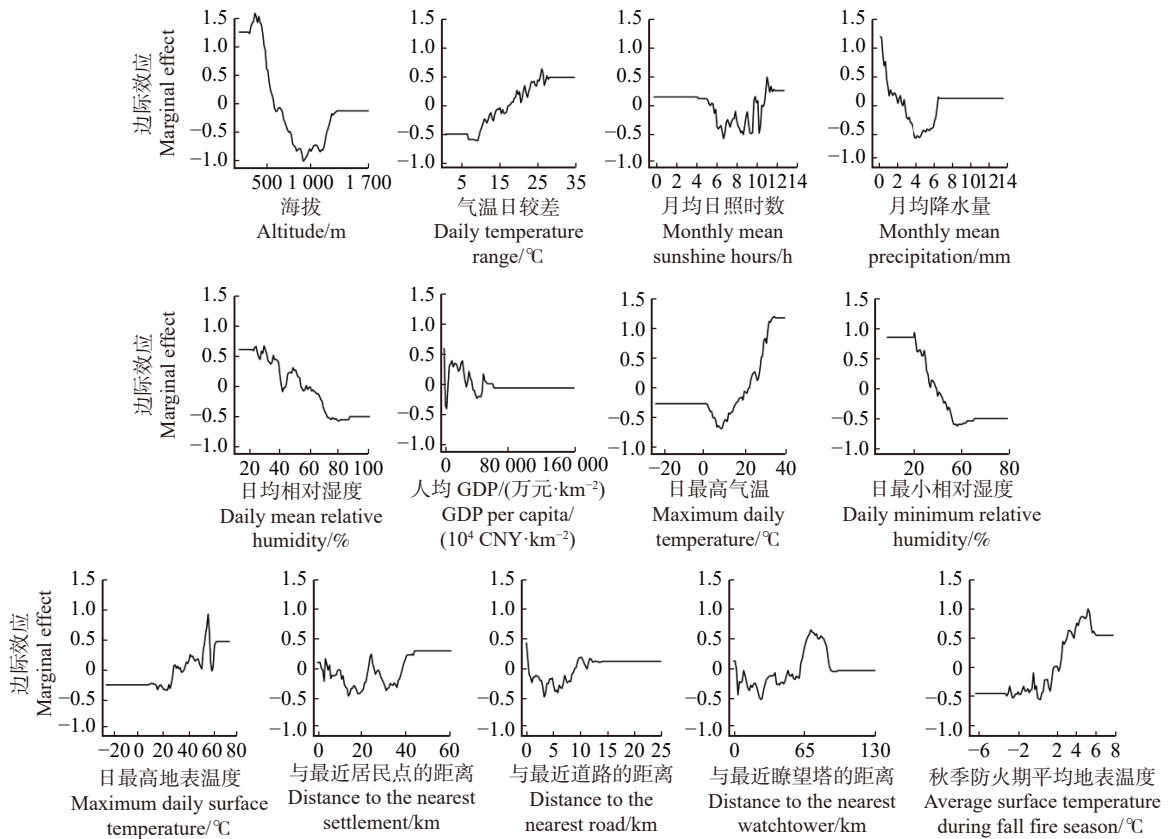


图 7 基于 BRT 模型全样本结果的边际效应

Fig. 7 Marginal effects based on full sample results of BRT model

间非平稳性问题^[31],它考虑了地理空间因素的影响,并能将解释变量的空间变化纳入模型中。下一步可考虑引入地理加权逻辑斯蒂回归(geographically weighted logistic regression, GWLR)模型,以分析林火驱动因素随空间位置变化对林火发生概率的影响。

3.3 内蒙古大兴安岭的林火发生概率

从图4和图5分析结果来看,内蒙古大兴安岭林火发生概率高的地区主要集中在东部与东南部。这些地区的人口相对密集,林草、林田交错复杂,导致火源管理难度较大,人为因素引发的火灾较多。1990年以后,该地区兴建了许多旅游景点及林内步道,道路密度增加,进一步增加了森林火灾发生的风险。研究区西南部和北部的火灾风险较高,因为这两个区分别是中蒙、中俄边境,不同国家对火灾管控的政策不同^[32],导致过境火在两个地区时常发生。此外,研究区北部为内蒙古大兴安岭原始林区,是雷击火的重灾区,具有多发频发的特点^[33],这也是该地区有较高火灾风险的原因之一。今后,在内蒙古大兴安岭的森林火灾防控需关注:(1)鉴于人口密集区及道路附近的火源管理难度较大,应持续加强对这些区域的用火管控。特别是对于近年来成为旅游热点的原生态原始林区,需要在景区开发和运营期间加强用火管控,以减少人为因素引发的火灾。(2)考虑到中蒙、中俄等边境地区越境火的风险,应加强对这些区域的监测与防控工作,严防境外火蔓延入境。(3)在雷击火高发地区,建议建立雷电监测系统,并加大对瞭望塔等基础设施的投入,以便更有效地监测和应对雷击引发的火灾。

4 结 论

本文利用1981—2020年的历史火灾数据,结合气象、社会经济等因素,使用4种模型对内蒙古大兴安岭森林火灾的发生概率及驱动因素进行了分析,并比较了4种模型的预测精度及拟合优度。

(1)气象因素是影响内蒙古大兴安岭林火发生的关键因素,主要包括气温日较差、日最小相对湿度等。前一年秋季防火期的气象要素如平均气温、平均地表温度会影响第2年森林火灾的发生。海拔、人类活动和社会经济因素对火灾发生也有一定的影响。(2)内蒙古大兴安岭中高火险区主要集中在东部和东南部,研究区北部(内蒙古大兴安岭北部原始林区)受境外火入侵(中俄边境)和雷击火影响,以及西南部受境外火(中蒙边境)影响的区域也存在一定的火灾风险。(3)在模型预测精度方面,全样本下的AUC值排序显示GR模型为0.851,LR模型为

0.852,RF模型为0.947,而BRT模型为0.967。这表明BRT模型和RF模型在预测效果和拟合优度上优于其他两种回归模型,其中BRT模型被认为是最适合内蒙古大兴安岭的林火预测模型。LR模型和GR模型的预测能力相近,在火点与非火点比例为1:1.5时,LR模型的拟合优度更高。

参 考 文 献

- [1] Thom D, Seidl R. Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests[J]. *Biological Reviews*, 2016, 91: 760–781.
- [2] Pang Y Q, Li Y D, Feng Z K, et al. Forest fire occurrence prediction in China based on machine learning methods[J/OL]. *Remote Sensing*, 2022, 14: 5546[2024-01-21]. <https://doi.org/10.3390/rs14215546>.
- [3] Mundo I A, Wiegand T, Kanagaraj R, et al. Environmental drivers and spatial dependency in wildfire ignition patterns of northwestern Patagonia[J]. *Journal of Environmental Management*, 2013, 123: 77–87.
- [4] Phelps N, Woolford D G. Comparing calibrated statistical and machine learning methods for wildland fire occurrence prediction: a case study of human-caused fires in Lac La Biche, Alberta, Canada[J]. *International Journal of Wildland Fire*, 2021, 30: 850–870.
- [5] Nguyen N T, Dang B T N, Pham X C, et al. Spatial pattern assessment of tropical forest fire danger at Thuan Chau area (Vietnam) using GIS-based advanced machine learning algorithms: a comparative study[J]. *Ecological Informatics*, 2018, 46: 74–85.
- [6] 曾爱聪,蔡奇均,苏漳文,等.基于MODIS卫星火点的浙江省林火季节变化及驱动因子[J].*应用生态学报*,2020,31(2):399–406.
Zeng A C, Cai Q J, Su Z W, et al. Seasonal variation and driving factors of forest fire in Zhejiang Province, China, based on MODIS satellite hot spots, Chinese[J]. *Journal of Applied Ecology*, 2020, 31(2): 399–406.
- [7] Guris S, Akya E C, Un T. Estimating of probability of homeownership in rural and urban areas: Logit, Probit and Gompit model[J]. *European Journal of Social Sciences*, 2011, 21: 405–411.
- [8] 苏漳文,曾爱聪,蔡奇均,等.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[J].*林业工程学报*,2019,4(4):135–142.
Su Z W, Zeng A C, Cai Q J, et al. Study on prediction model and driving factors of forest fire in Da Hinggan Mountains using Gompit regression method[J]. *Journal of Forestry Engineering*, 2019, 4(4): 135–142.
- [9] 高超,林红蕾,胡海清,等.我国林火发生预测模型研究进展[J].*应用生态学报*,2020,31(9):3227–3240.
Gao C, Lin H L, Hu H Q, et al. A review of models of forest fire occurrence prediction in China[J]. *Chinese Journal of Applied Ecology*, 2020, 31(9): 3227–3240.
- [10] Oliveira S, Oehler F, San-Miguel-Ayanz J, et al. Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest[J]. *Forest Ecology and Management*, 2012, 275: 117–129.

- [11] Eskandari S, Amiri M, Sadhasivam N, et al. Comparison of new individual and hybrid machine learning algorithms for modeling and mapping fire hazard: a supplementary analysis of fire hazard in different counties of Golestan Province in Iran[J]. *Natural Hazards*, 2020, 104: 305–327.
- [12] Chang C, Chang Y, Xiong Z P, et al. Predicting grassland fire-occurrence probability in Inner Mongolia Autonomous Region, China[J/OL]. *Remote Sensing*, 2023, 15(12): 2999[2024-01-23]. <https://doi.org/10.3390/rs15122999>.
- [13] Ruffault J, Mouillot F. Contribution of human and biophysical factors to the spatial distribution of forest fire ignitions and large wildfires in a French Mediterranean region[J]. *International Journal of Wildland Fire*, 2017, 26: 498–508.
- [14] Wu Z W, He H S, Keane R E, et al. Current and future patterns of forest fire occurrence in China[J]. *International Journal of Wildland Fire*, 2020, 29: 104–119.
- [15] Massada A B, Syphard A D, Stewart S I, et al. Wildfire ignition-distribution modelling: a comparative study in the Huron-Manistee National Forest, Michigan, USA[J]. *International Journal of Wildland Fire*, 2013, 22: 174–183.
- [16] 刘璞. 内蒙古大兴安岭典型森林积雪化学特征研究 [D]. 哈尔滨: 东北林业大学, 2022.
Liu P. Study up on chemical characteristics of snow cover in typical forests of Daxing'an Mountains, Inner Mongolia[D]. Harbin: Northeast Forestry University, 2022.
- [17] 赵鹏武, 葛晨赫, 王嘉夫, 等. 内蒙古大兴安岭森林火险等级评价与区划研究 [J]. *西南林业大学学报 (自然科学)*, 2021, 41(2): 151–158.
Zhao P W, Ge C H, Wang J F, et al. Study on rating and classification of forest fire risk in Daxing'an Mountains, Inner Mongolia[J]. *Journal of Southwest Forestry University (Natural Sciences)*, 2021, 41(2): 151–158.
- [18] 徐化成. 中国大兴安岭森林 [M]. 北京: 科学出版社, 1998: 40–41.
Xu H C. Da Hinggan Ling Mountains forests in China[M]. Beijing: Science Press, 1998: 40–41.
- [19] Chang Y, Zhu Z L, Bu R C, et al. Predicting fire occurrence patterns with logistic regression in Heilongjiang Province, China[J]. *Landscape Ecology*, 2013, 28: 1989–2004.
- [20] Littell J S, McKenzie D, Peterson D L, et al. Climate and wildfire area burned in western U. S. ecoprovinces, 1916–2003[J]. *Ecological Applications*, 2009, 19: 1003–1021.
- [21] Zhang H J, Qi P C, Guo G M. Improvement of fire danger modelling with geographically weighted logistic model[J]. *International Journal of Wildland Fire*, 2014, 23: 1130–1146.
- [22] 内蒙古自治区统计局. 改革开放 30 年的内蒙古 [M]. 北京: 中国统计出版社, 2008.
Inner Mongolia Autonomous Region Bureau of Statistics. Inner Mongolia in the 30 years of reform and opening up[M]. Beijing: China Statistics Press, 2008.
- [23] 白海峰, 刘晓东, 牛树奎, 等. 基于贝叶斯模型平均法的森林火灾预测模型构建研究以云南省大理州为例 [J]. *北京林业大学学报*, 2021, 43(5): 44–52.
- Bai H F, Liu X D, Niu S K, et al. Construction of forest fire prediction model based on Bayesian model averaging method: taking Dali Prefecture, Yunnan Province of southwestern China as an example[J]. *Journal of Beijing Forestry University*, 2021, 43(5): 44–52.
- [24] Frank E H. Regression modeling strategies-with applications to linear models, logistic regression, and survival analysis[M]. New York: Springer, 2001.
- [25] Guo F T, Selvalakshmi S, Lin F F, et al. Geospatial information on geographical and human factors improved anthropogenic fire occurrence modeling in the Chinese boreal forest[J]. *Canadian Journal of Forest Research*, 2016, 46: 582–594.
- [26] 陈锋, 林向东, 牛树奎, 等. 气候变化对云南省森林火灾的影响 [J]. *北京林业大学学报*, 2012, 34(6): 7–15.
Chen F, Lin X D, Niu S K, et al. Influence of climate change on forest fire in Yunnan Province, southwestern China[J]. *Journal of Beijing Forestry University*, 2012, 34(6): 7–15.
- [27] Wu Z W, He H S, Yang J, et al. Relative effects of climatic and local factors on fire occurrence in boreal forest landscapes of northeastern China[J]. *Science of the Total Environment*, 2014, 493: 472–480.
- [28] 舒展. 气候变化对大兴安岭塔河林业局森林火灾的影响研究 [D]. 哈尔滨: 东北林业大学, 2011.
Shu Z. Impacts research of climate change on forest fires in Tahe Forestry Bureau in Great Xing'an Mountain Region[D]. Harbin: Northeast Forestry University, 2011.
- [29] Guo F T, Su Z W, Wang G Y, et al. Understanding fire drivers and relative impacts in different Chinese forest ecosystems[J]. *Science of the Total Environment*, 2017, 605: 411–425.
- [30] Guo F T, Zhang L J, Jin S, et al. Modeling anthropogenic fire occurrence in the boreal forest of China using logistic regression and random forests[J/OL]. *Forests*, 2016, 7: 250[2023-01-23]. <https://www.mdpi.com/1999-4907/7/1/250>.
- [31] Su Z W, Hu H Q, Tigabu M, et al. Geographically weighted negative binomial regression model predicts wildfire occurrence in the Great Xing'an Mountains better than negative binomial model[J/OL]. *Forests*, 2019, 10: 377[2023-01-23]. <http://dx.doi.org/10.3390/f10050377>.
- [32] 周粉粉, 郭蒙, 钟超, 等. 呼伦贝尔草原火时空格局及特征分析 [J]. *地理科学*, 2022, 42(10): 1838–1847.
Zhou F F, Guo M, Zhong C, et al. Spatial-temporal patterns and characteristics of grassland fire in the Hulunbuir Grassland[J]. *Scientia Geographica Sinica*, 2022, 42(10): 1838–1847.
- [33] 杨少斌, 曹萌, 祝鑫海, 等. 2001—2019 年内蒙古大兴安岭北部原始林区森林火灾发生规律研究 [J]. *灾害学*, 2022, 37(3): 122–128.
Yang S B, Cao M, Zhu X H, et al. Research on occurrence law of forest fire in the northern primitive forest area of Greater Khingan Range, Inner Mongolia from 2001 to 2019[J]. *Journal of Catastrophology*, 2022, 37(3): 122–128.

(责任编辑 孟瑶 李 契
责任编辑 臧润国)