remote sensing

MDPI

Article

Explainable XGBoost–SHAP Modeling Coupled with GeoDetector for Forest-Fire Occurrence Susceptibility and Future Climate Sensitivity in the Dabie Mountains

Fanghao Wei [1,2,3], Xiaoli Huang [1,2,*], Meng Cao [4] and Yakui Shao [5]

[1] Anhui Province Key Laboratory of Physical Geographic Environment, Chuzhou University, Chuzhou 239000, China; wfh0914@bjfu.edu.cn
[2] Anhui Engineering Research Center of Remote Sensing and Geo-Informatics, Chuzhou University, Chuzhou 239000, China
[3] College of Forestry, Beijing Forestry University, Beijing 100083, China
[4] Office of Scientific Research, China Fire and Rescue Institute, Changping District, Beijing 102202, China; caomeng526@cfri.edu.cn
[5] Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China
* Correspondence: huangxl@chzu.edu.cn

Highlights

What are the main findings?

- XGBoost achieved the best performance for forest-fire occurrence susceptibility prediction, with an ROC-AUC of 0.850 under five-fold cross-validation.
- SHAP analysis revealed stable predictor contributions and nonlinear responses, with the log-transformed population showing the largest model-based contribution across validation analyses.
- The climate-scale XGBoost model indicated scenario-dependent susceptibility changes, with stronger late-century redistribution under SSP585 than SSP126.

What are the implications of the main findings?

- The explainable susceptibility model provides an integrated approach for regional fire-risk screening by integrating multi-source environmental information and interpretable machine learning.
- The explainable XGBoost framework supports relative susceptibility screening and climate-sensitivity assessment, while complementing rather than replacing operational fire-management information.

Abstract

Forest-fire occurrence susceptibility prediction remains challenging in mountainous regions because fire occurrence is controlled by nonlinear interactions among climatic, vegetation, topographic, and anthropogenic factors. Developing interpretable prediction models that integrate multi-source environmental information is essential for improving susceptibility assessment and understanding predictor contributions. Using 2003–2022 MODIS active-fire records and multi-source environmental data from the Dabie Mountains, this study developed an explainable XGBoost-based forest-fire occurrence susceptibility prediction framework. The predictive performance of XGBoost was compared with Random Forest and radial basis function support vector machine (RBF-SVM) using five-fold cross-validation, and TreeSHAP was applied to interpret predictor contributions, nonlinear responses, and interactions. KDE, standard deviational ellipses (SDEs), and GeoDetector were further used to characterize historical fire patterns and complementary spatial associations, while

Check for updates

Academic Editor: Nikolay S. Strigul

Received: 12 August 2026
Revised: 30 September 2026
Accepted: 5 October 2026
Published: 14 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

an independent climate-scale XGBoost model was applied to evaluate future susceptibility sensitivity under SSP126 and SSP585 scenarios. XGBoost achieved an ROC-AUC of 0.850, and log-transformed population showed the largest global SHAP contribution. GeoDetector identified precipitation and relative humidity as the strongest spatial associations with historical fire-density patterns across KDE bandwidths. The climate-scale model achieved a pooled five-fold ROC-AUC of 0.830 and indicated stronger late-century susceptibility-score redistribution under SSP585 than under SSP126. Together, these results demonstrate an explainable susceptibility modeling approach that integrates predictive modeling, model interpretation, spatial analysis, and future climate assessment. The outputs represent relative susceptibility and scenario-specific sensitivity responses rather than unconditional fire probabilities or deterministic forecasts.

Keywords: forest fire; occurrence susceptibility; GeoDetector; explainable machine learning; XGBoost; SHAP; CMIP6; climate sensitivity

1. Introduction

Forest-fire occurrence is shaped by complex interactions among climatic, vegetation, topographic, and anthropogenic factors, with climate variability influencing fuel conditions and fire-conducive environments [1,2]. Vegetation characteristics, terrain configuration, accessibility, population distribution, and land-use patterns further regulate ignition opportunities and fire exposure [3]. These controls are particularly heterogeneous in monsoon-influenced mountain regions, where steep topographic gradients, seasonal climate variability, and fragmented forest–human interfaces generate strong spatial differences in fire occurrence [3,4]. Therefore, developing accurate forest-fire occurrence susceptibility models in such complex environments remains challenging because fire responses often involve nonlinear relationships and interactions among multiple environmental factors [5].

Spatial analyses of historical fire records can reveal concentration patterns, directional structure, and environmental correlates, but these descriptions are sensitive to analytical scale and do not by themselves establish environmental effects [6]. KDE results vary with bandwidth and depend on spatial support and stratification [7–9]. GeoDetector provides a measure of spatially stratified association rather than causal effects; its q-value reflects spatial explanatory consistency rather than an effect size or direction, and results may depend on spatial representation and discretization [10]. Machine-learning models can estimate nonlinear relationships and complex interactions between fire occurrence and environmental conditions [11,12]. RF, SVM, gradient-boosted trees, and XGBoost are widely used for wildfire susceptibility modeling [11–13], but predictive performance alone does not explain predictor contributions and may be optimistic without separated development and evaluation [14]. SHAP improves interpretation of tree-based models by quantifying predictor contributions within model outputs, while the resulting attributions remain model-based rather than causal [15–17].

Spatial–statistical approaches and machine-learning interpretation methods address different analytical questions and should therefore be interpreted as complementary rather than interchangeable [18]. Full-grid spatial analyses characterize the organization of historical fire-density patterns, whereas event-level classifiers evaluate the ability of environmental predictors to distinguish observed fires from background conditions. However, predictive models require complementary spatial context because model-based predictor contributions do not necessarily represent long-term spatial associations. Therefore, integrating spatial association analysis with explainable machine-learning modeling can

provide a more comprehensive understanding of forest-fire susceptibility without directly comparing factor rankings across methods [3,17]. Furthermore, future climate sensitivity assessment requires consistency between historical model inputs and future scenario predictors, motivating the use of an independent climate-scale model based on temporally harmonized climatological variables [19–23].

Existing fire studies in the Dabie Mountains and surrounding regions have provided important insights into fire occurrence patterns and susceptibility assessment using different response variables, spatial extents, temporal supports, predictor sets, and modeling approaches. At the Anhui provincial scale, Zhang et al. analyzed forest-fire occurrence using 1 km annual grids and geographically and temporally weighted regression, emphasizing spatiotemporal heterogeneity and socioeconomic influences [24]. A province-wide susceptibility assessment integrating field-based fuel information, MODIS fire records, and multiple machine-learning algorithms identified relatively high-susceptibility clusters in southern and southwestern mountainous areas, including the Dabie Mountains [25]. By contrast, a park-scale study in Anhui using an optimized PSO-RF framework identified NDVI as the leading predictor [26]. These studies demonstrate the value of regional fire modelling, but also highlight the need for prediction frameworks that integrate heterogeneous environmental data, provide interpretable machine-learning outputs, and support evaluation under future climate scenarios.

The Dabie Mountains represent a typical monsoon-influenced mountainous forest region with pronounced environmental gradients, extensive forest cover, and complex forest–human interactions. Although previous studies have improved understanding of regional fire occurrence and susceptibility patterns, an explainable prediction framework that integrates heterogeneous environmental data, robust machine-learning modeling, and future climate assessment remains limited [18,21,22]. Using 2003–2022 MODIS active-fire records and multi-source environmental data [27], this study develops an explainable XGBoost-based forest-fire occurrence susceptibility prediction framework. Specifically, this study aims to: (1) develop and evaluate a machine-learning model for forest-fire occurrence susceptibility prediction using climatic, vegetation, topographic, and anthropogenic predictors; (2) interpret predictor contributions, nonlinear responses, and interactions using SHAP while assessing complementary spatial associations through GeoDetector; and (3) evaluate potential changes in modeled susceptibility under future climate scenarios using an independent climate-scale model with harmonized climatological predictors [21,22]. This framework provides an interpretable approach for regional forest-fire susceptibility assessment by integrating predictive modelling as the core approach with model explanation, spatial association analysis, and future climate sensitivity evaluation as complementary analyses [28,29].

2. Materials and Methods

2.1. Study Area

The Dabie Mountains (29°40′–32°45′N, 113°30′–117°50′E) extend across Anhui, Henan, and Hubei in central–eastern China [30]. Located in the northern subtropical monsoon zone, the region exhibits strong seasonal variability in temperature and precipitation [31,32]. Complex terrain, including extensive elevation and slope gradients, together with diverse forest communities and fragmented forest–human interfaces, creates heterogeneous environmental conditions relevant to forest-fire susceptibility assessment [33–37] (Figure 1).

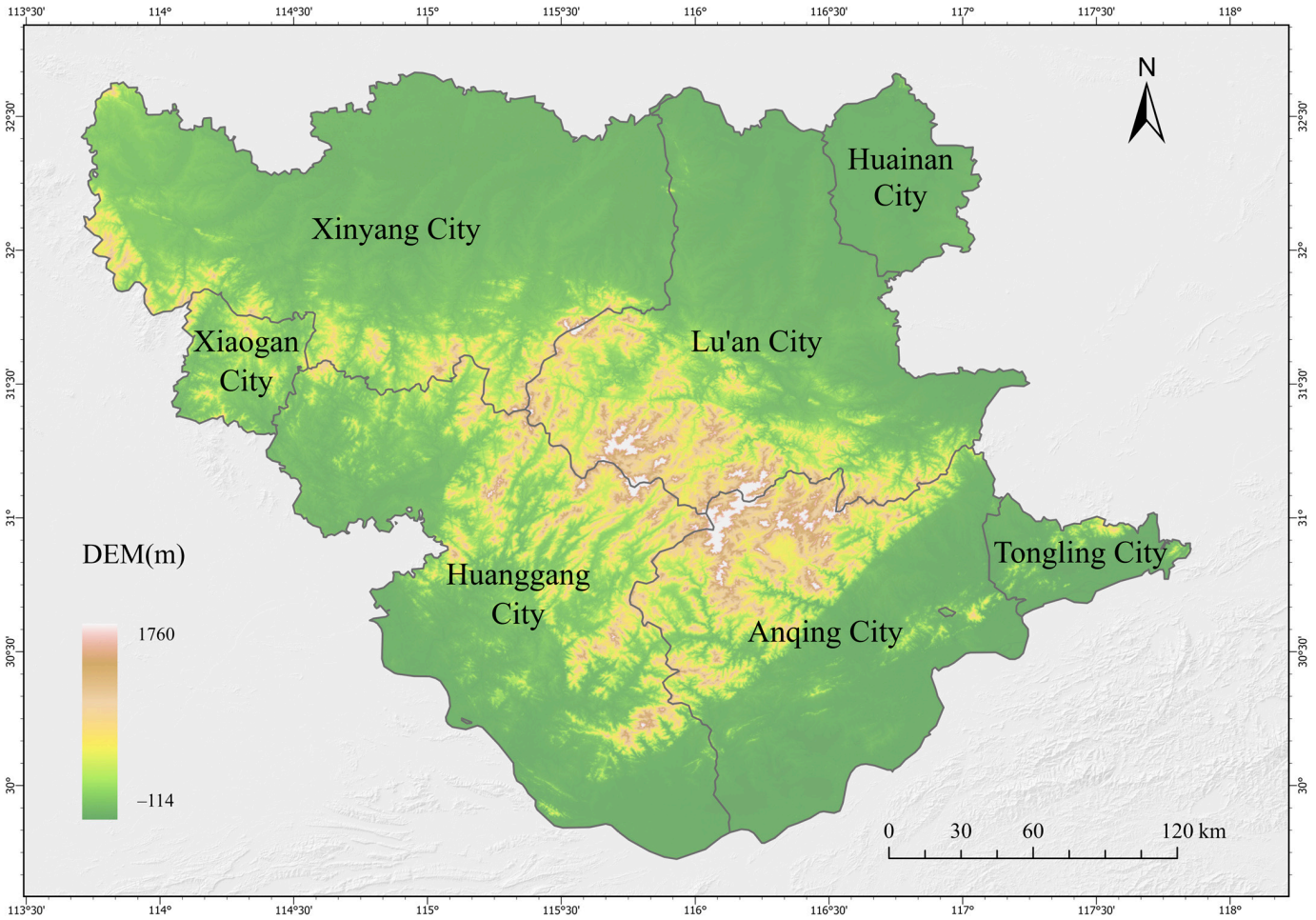


Figure 1. Location of the Dabie Mountains study area.

2.2. Multi-Source Environmental Data and Database Construction

Multi-source environmental datasets were integrated to construct the analytical basis for forest-fire occurrence susceptibility modeling, spatial association analysis, and future climate sensitivity assessment. These datasets included MODIS MCD14ML active-fire records [27], ERA5-Land meteorological variables [38], SRTM terrain data [39], MOD13A2 NDVI [40], annual MCD12Q1 land-cover products [41], WorldPop population data [42], and WorldClim CMIP6 BCC-CSM2-MR climate datasets [43,44]. The temporal coverage, native resolutions, and primary spatial characteristics of these datasets are summarized in Table 1.

MODIS MCD14ML was selected as the historical fire dataset for 2003–2022 because it provides continuous temporal coverage throughout the study period. Although VIIRS provides finer spatial resolution and greater sensitivity to small fires, its S-NPP record begins in 2012; therefore, integrating MODIS and VIIRS would introduce inconsistencies in fire-detection characteristics over the 2003–2022 study period. All environmental datasets were harmonized to a common analytical grid of approximately 1 km. The 1 km analytical grid was selected as a compromise between the spatial characteristics of MODIS fire observations and NDVI products and the integration requirements of higher-resolution terrain and population datasets and coarser climate variables. Detailed preprocessing procedures, including resampling, aggregation, categorical alignment, and predictor transformation, are provided in the Supplementary Materials (Section S1).

Table 1. Core datasets, temporal coverage, and spatial harmonization used in the analysis.

Data Category	Dataset Name	Variables/ Description	Temporal Resolution/Coverage	Native Resolution/Spatial Harmonization	Data Source/Platform
Historical Wildfires	MCD14ML	Fire records: coordinates and satellite acquisition date	Daily; 2003–2022	1 km nominal detection resolution; retained as point records	NASA FIRMS
Meteorological	ERA5-Land Reanalysis Dataset	Temperature (°C), precipitation (mm), wind speed (m/s), relative humidity (%)	Hourly; 2003–2022	Approx. 9 km; aligned with the common 1 km analysis grid without statistical downscaling	ECMWF CDS
Topographical	SRTM-DEM	Elevation (m), slope (°), aspect (°)	Static baseline; 2000	Approx. 30 m; harmonized with the common 1 km analysis grid	Google Earth Engine
Vegetation	MOD13A2	NDVI	16-day; 2003–2022	1 km; retained on the common 1 km analysis grid	Google Earth Engine
Socioeconomic	WorldPop Global Project Population Data	Population count aggregated within each fixed 1 km grid	Annual; 2003–2020; 2020 used as proxy for 2021–2022	100 m; native-pixel counts summed within each 1 km grid cell	Google Earth Engine
Land Cover	MCD12Q1	Forest cover mask	Annual; 2003–2022	500 m; nearest-neighbor alignment with the common 1 km analysis grid	Google Earth Engine
Future climate	WorldClim CMIP6 downscaled BCC-CSM2-MR	Temperature and precipitation under SSP126 and SSP585	20-year periods: 2021–2040, 2041–2060, 2061–2080, 2081–2100	2.5 arc-min; aligned with the common 1 km analysis grid	WorldClim/ CMIP6

Three analytical datasets were prepared for different analyses. The event-level dataset contained 182 positive fire observations and 182 pseudo-absence observations (364 records) and was used for susceptibility modeling, model comparison among RF, RBF-SVM, and XGBoost, five-fold validation, and model interpretation [45,46]. The full-coverage dataset consisted of 4463 stable forest grids and was used for GeoDetector-based spatial association analysis. The climate-scale dataset was constructed using temporally harmonized predictors to generate the historical susceptibility baseline and evaluate future scenario-specific susceptibility responses. Detailed procedures for database construction and predictor preprocessing are provided in the Supplementary Materials (Section S1). The overall workflow of this study is shown in Figure 2.

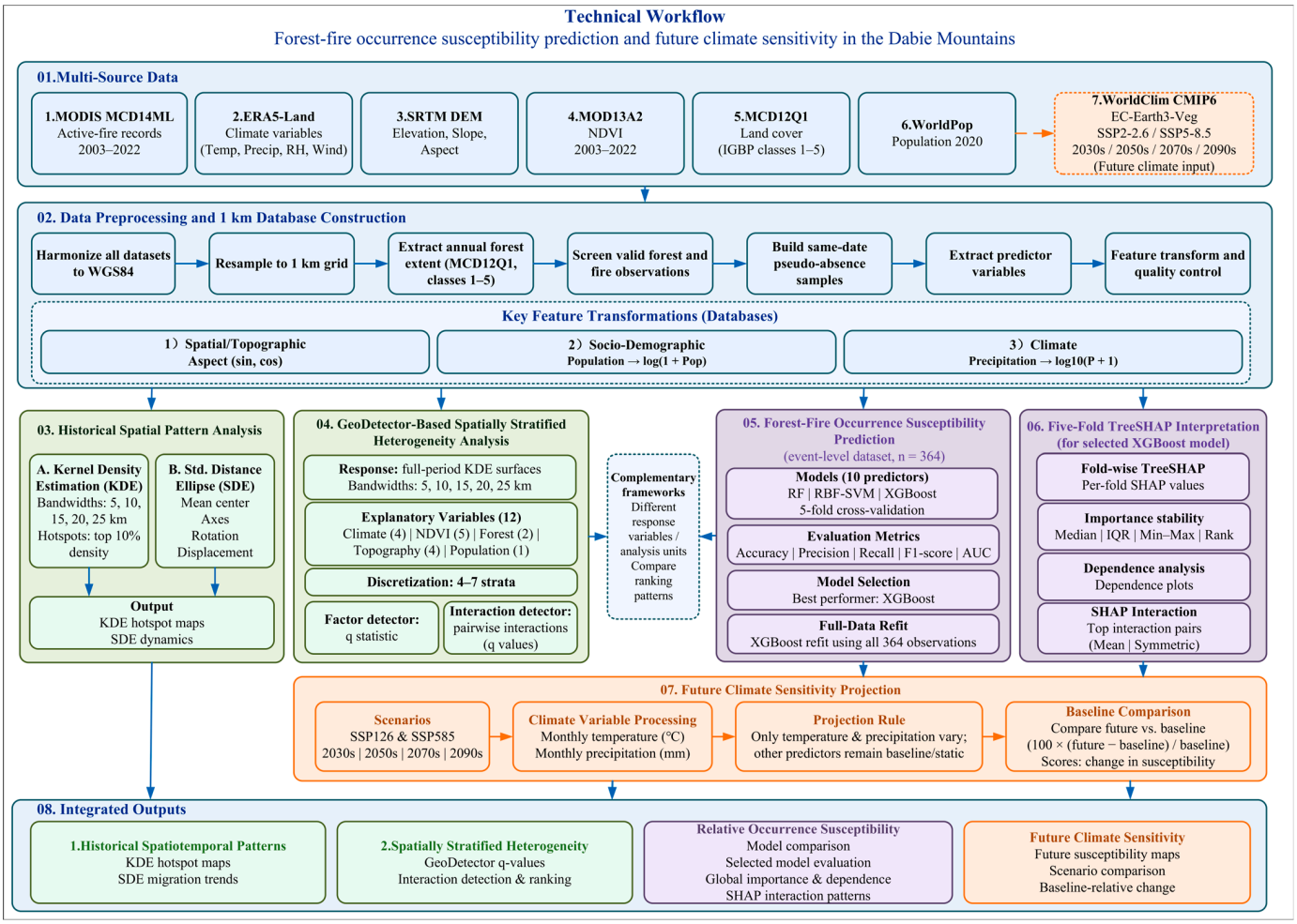


Figure 2. Overall workflow of the study.

2.3. Historical Spatial Pattern Analysis of Forest Fires

KDE and SDE were applied to the 182 positive active-fire observations to address complementary descriptive questions: KDE characterized where historical fire-density patterns were concentrated, whereas SDE tracked how the spatial direction and displacement of fire occurrence changed over time [47]. Using bandwidths of 5, 10, 15, 20, and 25 km, KDE produced continuous relative-density surfaces of historical fire occurrence. The highest-density areas were identified as bandwidth-specific hotspots. KDE was evaluated for the full period, the four temporal periods (2003–2007, 2008–2012, 2013–2017, and 2018–2022), and the four seasons (spring, summer, autumn, and winter) to assess the temporal and seasonal stability of fire-concentration patterns. These KDE surfaces supported historical spatial characterization and served as response variables in GeoDetector analysis; they were not used as predictors in the event-level susceptibility model [48].

SDE then characterized how the directional distribution and temporal displacement of historical fire occurrence changed across periods and seasons. For each period and season, the mean center, major and minor axes, ellipse area, orientation, and centroid migration were quantified to describe changes in spatial dispersion and directional patterns.

2.4. Spatial Association Analysis Using GeoDetector

GeoDetector was used to quantify spatial associations between environmental factors and historical fire-density patterns rather than to infer causal effects. The 4463 stable forest grids derived from the full-coverage database were used as spatial analysis units. Full-period KDE values under five bandwidth settings (5, 10, 15, 20, and 25 km) were

used as spatial response variables. Explanatory variables represented climatic, vegetation, topographic, and anthropogenic conditions consistent with the susceptibility modeling framework. Continuous variables were discretized using a predefined five-class quantile scheme, while alternative discretization strategies were evaluated in sensitivity analyses to examine the robustness of factor ranking.

Factor detector q-values were used to quantify spatial association strength, and significance was assessed using fixed-seed permutation tests Pairwise interactions were assessed by comparing interaction q-values with individual-factor q-values. Interaction robustness was evaluated based on joint-stratum support, and combinations with insufficient spatial support were interpreted cautiously. Detailed discretization sensitivity, q-value variation, and interaction results are provided in the Supplementary Materials.

2.5. Forest-Fire Occurrence Susceptibility Modeling

Random Forest (RF) and radial basis function support vector machine (RBF-SVM) were used as benchmark classifiers, while XGBoost was selected as the primary susceptibility modeling approach because of its ability to capture nonlinear relationships and complex interactions among environmental predictors [45–47]. All models were trained using the fixed event-level database with FireLabel as the binary response. The event-level predictors consisted of meteorological variables (temperature, ln(1 + P), wind speed, and relative humidity) and background environmental variables (elevation, slope, Northness, Eastness, NDVI, and ln(1 + Pop)).

To evaluate the influence of antecedent meteorological conditions, three XGBoost configurations were compared: M0 using acquisition-date weather, M7 using 7-day antecedent summaries, and M14 using 14-day antecedent summaries [48]. The six non-meteorological predictors, 364 observations, cross-validation partitions, and XGBoost specifications were kept identical among the three configurations. The event-level dataset consisted of observed fire locations and matched pseudo-absence samples, following commonly used presence-background modeling approaches [49,50]. Terrain predictors were derived from DEM-based geomorphometric attributes, with aspect represented by Northness and Eastness [51]. Potential multicollinearity among candidate predictors was assessed using variance inflation factor (VIF) diagnostics [52]. Model performance was evaluated using Accuracy, Precision, Recall, F1-score, and ROC-AUC based on pooled out-of-fold predictions. Detailed model comparison and antecedent-weather sensitivity results are provided in the Supplementary Materials (Tables S1 and S2).

The 364 observations were evaluated using five-fold cross-validation, with preprocessing and model fitting restricted to the development data to avoid information leakage. Held-out predictions were concatenated into pooled out-of-fold predictions for performance assessment. A complementary five-fold spatial-block cross-validation was conducted to evaluate spatial transferability, and XGBoost uncertainty was quantified using bootstrap-based model-refitting and OOF sampling procedures.

2.6. Five-Fold TreeSHAP Interpretation of the XGBoost Model

SHAP (SHapley Additive exPlanations) was used to interpret the selected XGBoost model under five-fold cross-validation. TreeSHAP values were calculated for held-out observations using models fitted without those observations, then pooled across folds while retaining fold identifiers for stability assessment. Global feature importance was quantified within each fold using mean absolute SHAP values. Fold-specific importance was summarized by median, interquartile range, minimum, maximum, and rank to evaluate contribution magnitude and stability across folds. Pooled held-out SHAP values were

visualized using beeswarm and absolute-SHAP distributions to characterize contribution direction, magnitude, and variability [15–17].

Sampling uncertainty in feature attribution was assessed using 1000 bootstrap resamples of the pooled held-out SHAP set. Mean absolute SHAP values were recalculated, normalized to relative importance, and converted into feature ranks for each bootstrap replicate. SHAP values and interaction terms were interpreted as model-based contributions rather than causal effects. Nonlinear predictor–response relationships were examined using SHAP dependence plots, and pairwise interactions were quantified using SHAP interaction values. Interaction strength was summarized using mean absolute SHAP interaction values [53].

2.7. Future Climate Sensitivity Assessment

A climate-scale application of XGBoost was developed to evaluate future susceptibility sensitivity using harmonized climatological predictors. The same historical fire observations and matched pseudo-absence locations as the event-level analysis were retained, while temporally varying predictors were represented by long-term climate conditions. Unlike the event-level model, which used acquisition-date weather variables, the climate-scale model used climatological predictors consistent with historical baseline and future scenario datasets [54].

Historical climate predictors were derived from ERA5-Land monthly aggregates for 2003–2022 [55]. Temperature, precipitation, relative humidity, and wind speed were summarized as long-term annual climatic variables, while NDVI and population were represented by historical baseline conditions. Precipitation was transformed using $\ln(1 + P)$ consistently across historical and future periods. The climate-scale XGBoost model was evaluated using the same five-fold cross-validation procedure and then applied to the 4463 stable forest grids to generate the historical susceptibility baseline [56].

Future climate predictors were obtained from the WorldClim CMIP6 downscaled BCC-CSM2-MR dataset under SSP126 and SSP585 for four future periods (2021–2040, 2041–2060, 2061–2080, and 2081–2100) [43,44]. Future changes were calculated relative to the historical baseline using susceptibility-score differences. Because only one GCM was considered and non-climatic predictors were maintained at baseline conditions, outputs were interpreted as scenario-specific relative susceptibility responses rather than deterministic fire predictions.

3. Results

3.1. Historical Fire Patterns and Spatial Context

KDE analysis identified a consistent south-central concentration of historical forest fires across bandwidths of 5–25 km, indicating stable large-scale spatial clustering patterns (Figure 3).

SDE analysis [57] showed temporal and seasonal variations in fire spatial distribution, with changes in centroid location and dispersion characteristics across periods and seasons (Figure 4).

3.2. XGBoost Model Selection and Spatial Validation

Model Comparison and Predictive Performance. Based on pooled out-of-fold predictions from five-fold cross-validation, XGBoost achieved the highest overall predictive performance, with Accuracy = 0.808, Precision = 0.792, Recall = 0.835, F1-score = 0.813, and ROC-AUC = 0.850. RF and RBF-SVM achieved lower ROC-AUC values of 0.825 and 0.788, respectively (Table 2). These results supported the selection of XGBoost as the primary model for subsequent interpretation. The supporting ROC curves for the model comparison are provided in the Supplementary Materials (Figure S1).

Across five cross-validation folds, XGBoost maintained balanced performance relative to RF and RBF-SVM, supporting stable discrimination across training–validation partitions.

Fold-level performance variations are provided in the Supplementary Materials (Figure S2). Spatial-block cross-validation resulted in lower performance estimates than conventional random five-fold validation for all evaluated models, reflecting the influence of spatial separation on model transferability. For XGBoost, ROC-AUC decreased from 0.850 under random cross-validation to 0.803 under spatial-block validation, while it remained higher than RF and RBF-SVM under spatially separated evaluation. These results indicate that spatial dependence contributed to performance variation, but the selected model retained comparatively robust discrimination ability across spatial domains (Table 3).

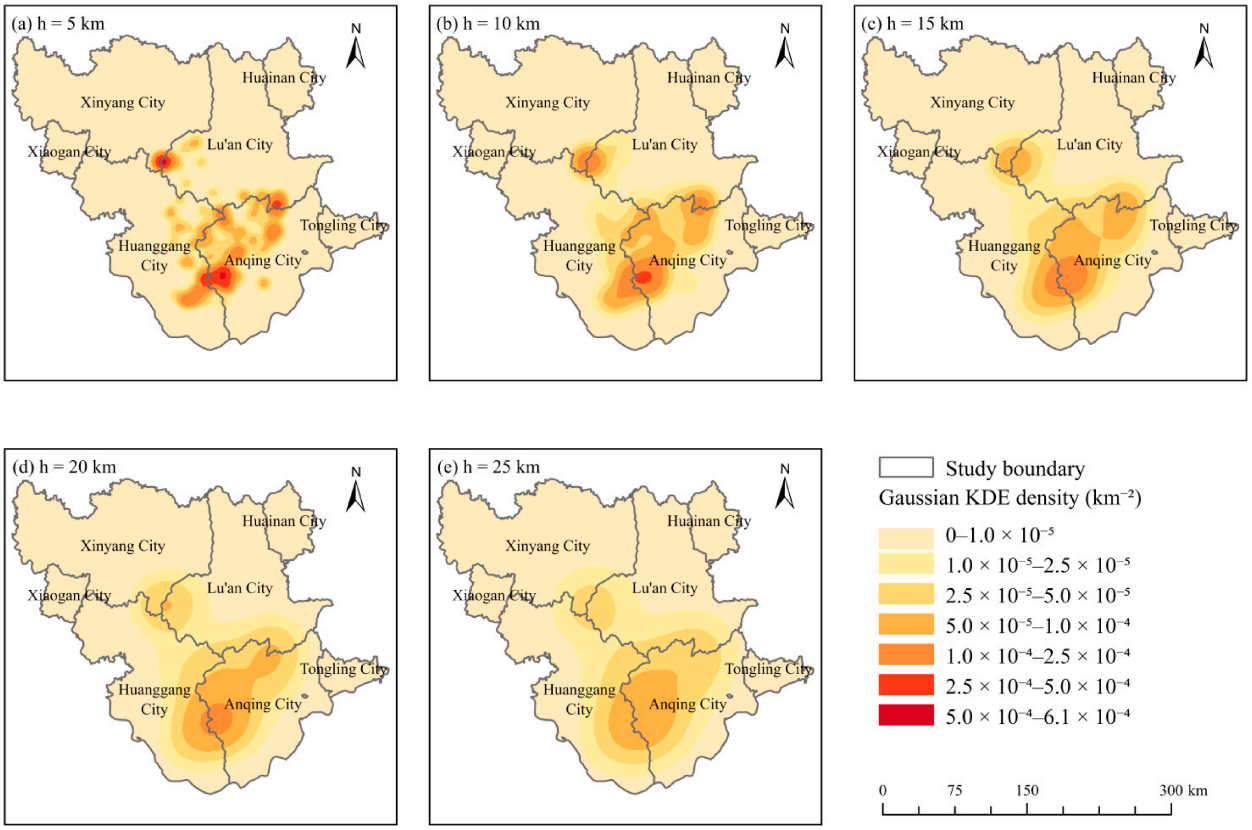


Figure 3. Gaussian kernel-density estimates of historical forest-fire density patterns at bandwidths of 5, 10, 15, 20, and 25 km. Values represent densities at 1 km grid centroids in EPSG:4548; each panel uses an independent color scale.

Table 2. Classification performance of XGBoost, Random Forest, and RBF-SVM based on pooled out-of-fold predictions from five-fold cross-validation.

Metric	XGBoost	Random Forest	RBF-SVM
Accuracy	0.808	0.745	0.717
Precision	0.792	0.713	0.698
Recall	0.835	0.819	0.764
F1-score	0.813	0.762	0.730
ROC-AUC	0.850	0.825	0.788

Bootstrap refitting and out-of-fold resampling produced ROC-AUC distributions consistent with the original estimate, indicating stability under sampling variation. Detailed bootstrap uncertainty distributions and SHAP rank–frequency summaries are provided in the Supplementary Materials (Figure S3 and Table S4). Antecedent-weather representations produced broadly comparable XGBoost performance, and detailed metric comparisons and meteorological attribution results are provided in the Supplementary Materials (Table S2).

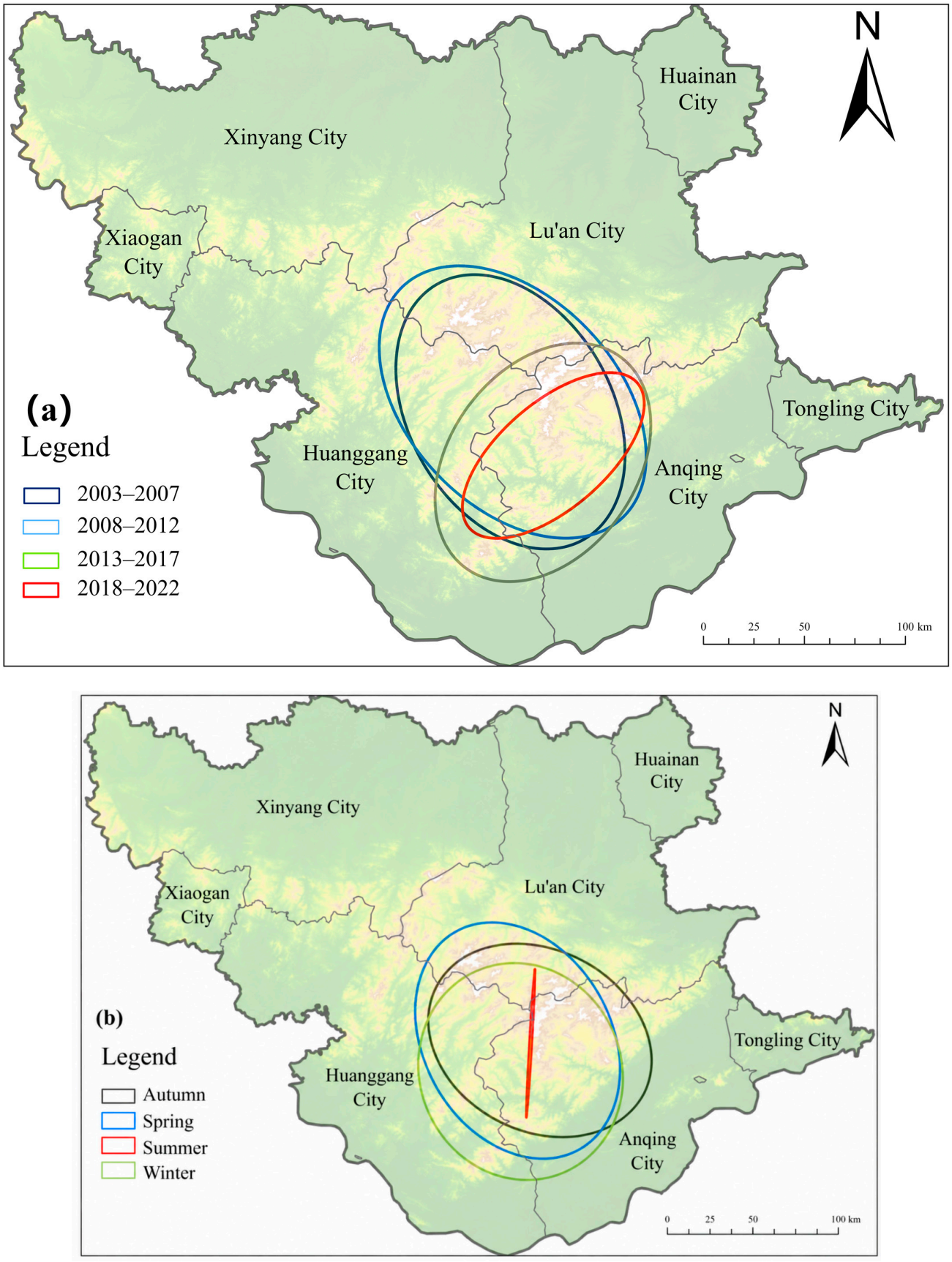


Figure 4. Standard deviational ellipse (SDE) analysis of the 182 historical forest–fire records in EPSG:4548. (**a**) One-standard-deviation ellipses and period mean-center migration for 2003–2007, 2008–2012, 2013–2017, and 2018–2022; (**b**) seasonal one-standard-deviation ellipses for spring, summer, autumn, and winter.

Table 3. Comparison of conventional and spatial-block cross-validation performance.

Model	Validation Scheme	Accuracy	F1-Score	ROC-AUC
RF	Random 5-fold CV	0.745	0.762	0.825
RF	Spatial-block 5-fold CV	0.722	0.745	0.783
RBF-SVM	Random 5-fold CV	0.717	0.730	0.788
RBF-SVM	Spatial-block 5-fold CV	0.674	0.692	0.751
XGBoost	Random 5-fold CV	0.808	0.813	0.850
XGBoost	Spatial-block 5-fold CV	0.751	0.783	0.803

3.3. Explainable XGBoost Interpretation Using SHAP

Global Feature Importance and Stability. SHAP analysis revealed a heterogeneous contribution structure among environmental predictors within the XGBoost model (Figure 5a,b). Log-transformed population exhibited the largest mean absolute SHAP contribution and the widest distribution of sample-level SHAP values, indicating that anthropogenic conditions contributed substantially to differentiating fire and non-fire observations. Wind speed, slope, NDVI, and relative humidity formed a secondary group of influential predictors, suggesting that atmospheric conditions, vegetation status, and terrain characteristics jointly contributed to model predictions. In contrast, elevation, Northness, Eastness, temperature, and precipitation showed relatively lower global SHAP magnitudes.

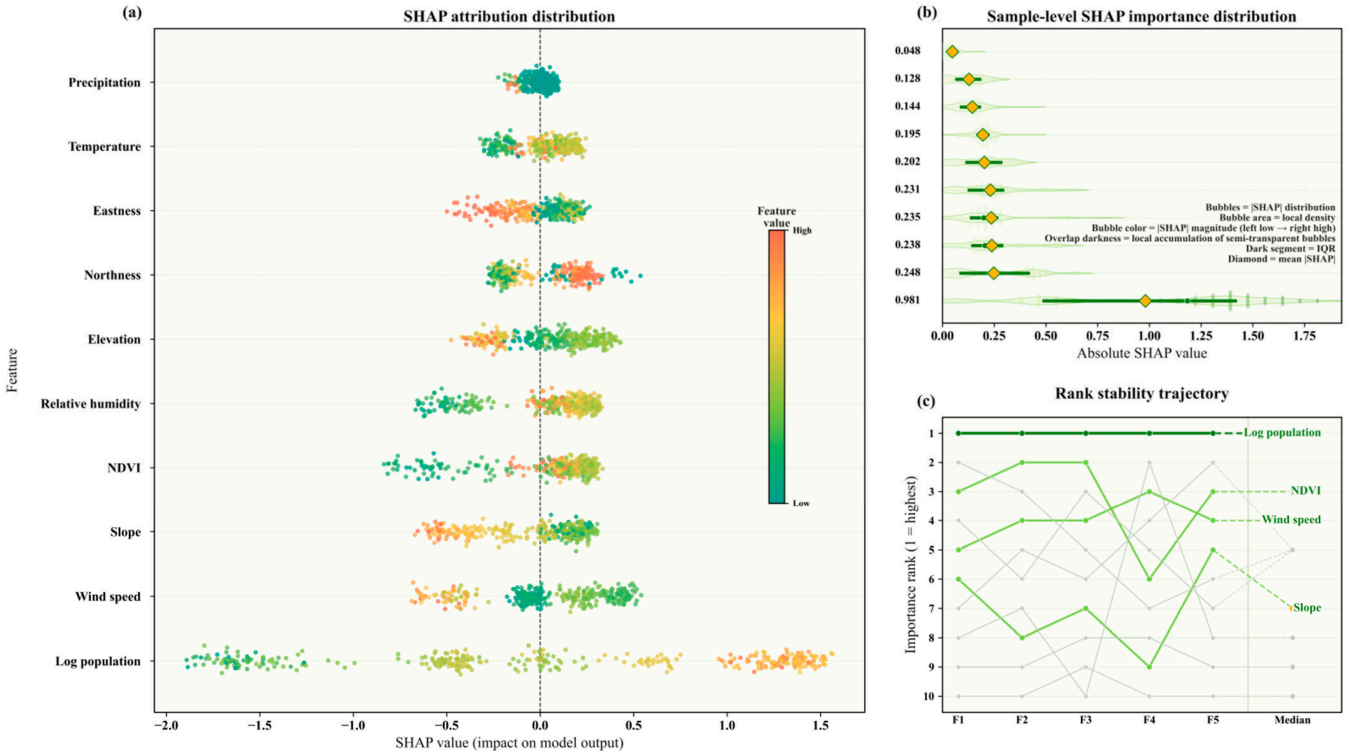


Figure 5. Global SHAP contribution patterns and rank stability of predictors in the selected XGBoost model. (**a**) SHAP attribution distribution; (**b**) sample-level absolute SHAP importance distribution; (**c**) rank stability trajectory across five cross-validation folds.

Log-transformed population ranked first across all folds, while NDVI and wind speed maintained stable intermediate positions; detailed fold-level ranking results are provided in the Supplementary Materials (Table S3).

Nonlinear Responses and Feature Interactions: SHAP dependence analysis revealed a pronounced nonlinear response of the log-transformed population within the XGBoost model (Figure 6a). Lower population values were generally associated with negative SHAP

contributions, whereas the contribution increased rapidly across an intermediate range and transitioned to consistently positive values at higher population levels. This threshold-like response indicates that the relationship between anthropogenic intensity and modeled fire susceptibility was nonlinear rather than monotonic within the prediction framework.

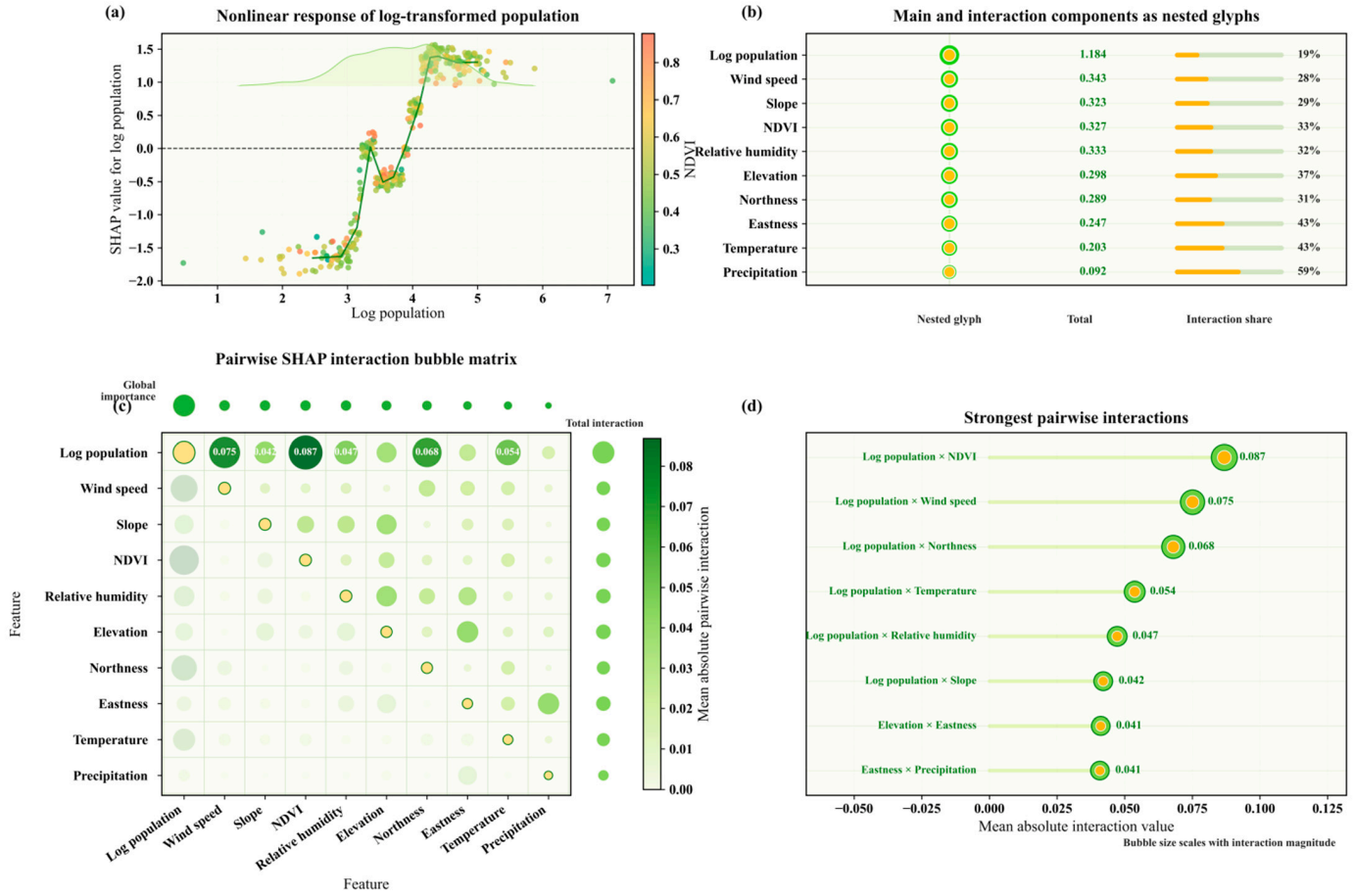


Figure 6. Nonlinear responses and feature interactions in the XGBoost model: (**a**) SHAP dependence of log-transformed population, with samples colored by NDVI and the fitted nonlinear trend overlaid; the horizontal dashed line denotes zero SHAP contribution; (**b**) mean absolute magnitudes of the main-effect and interaction components for individual predictors; (**c**) pairwise mean absolute SHAP interaction matrix; and (**d**) strongest pairwise SHAP interactions ranked by mean absolute interaction value.

Most SHAP contributions were dominated by main effects, although interaction components varied among predictors. The strongest pairwise interaction was observed between the log-transformed population and NDVI (mean absolute interaction value = 0.087), indicating that anthropogenic predictors were interpreted within vegetation contexts. Detailed interaction rankings are provided in the Supplementary Materials Section S3.1.

Complementary Interpretation across Analytical Scales: SHAP and GeoDetector provided complementary perspectives on forest-fire susceptibility because they addressed different analytical questions, response definitions, and temporal–spatial representations. SHAP quantified predictor contributions within the event-level XGBoost model, where individual fire observations and matched pseudo-absence samples were evaluated, whereas GeoDetector characterized spatial associations between environmental conditions and historical fire-density patterns derived from KDE surfaces. Therefore, the outputs from these approaches should not be interpreted as directly comparable rankings of environmental importance.

The contrast between the two approaches was particularly evident for precipitation and population-related variables. Precipitation showed strong spatial associations in GeoDetector because accumulated fire-season precipitation was closely associated with long-term KDE-based fire-density patterns. However, event-level precipitation showed a lower global SHAP contribution, although it participated in interaction structures within the XGBoost Model. Conversely, the log-transformed population exhibited the largest SHAP contribution but a relatively weak spatial association in GeoDetector.

These differences indicate that predictor contributions within a predictive model and regional environmental associations capture different dimensions of fire susceptibility. Event-level machine learning emphasizes the ability of environmental variables to discriminate fire occurrence from background conditions, whereas GeoDetector reflects the spatial organization of historical fire-density patterns. Combining explainable machine learning with spatial association analysis therefore provides complementary information for interpreting fire-occurrence patterns without implying causal relationships.

3.4. Spatial Association Patterns from GeoDetector

Environmental Spatial Associations across KDE Bandwidths: GeoDetector analysis revealed pronounced scale-dependent differences in the spatial associations between environmental variables and historical fire-density patterns across KDE bandwidths (Figure 7). Moisture-related variables exhibited the strongest and most consistent associations across spatial scales. Precipitation showed the highest q-values among all predictors, increasing from 0.163 at 5 km to 0.667 at 25 km, followed by relative humidity, which increased from 0.202 to 0.489. NDVI also maintained relatively strong spatial associations, with q-values increasing from 0.147 to 0.374 across bandwidths.

The increasing q-values of precipitation, relative humidity, and NDVI with expanding KDE bandwidths indicate that moisture availability and vegetation conditions were more strongly associated with regional fire-density organization than with fine-scale spatial patterns. In contrast, topographic variables, forest type, and population exhibited relatively weaker spatial associations across bandwidths, whereas wind speed showed intermediate but comparatively stable associations.

Notably, population exhibited limited spatial association with historical fire-density patterns despite its high contribution within the event-level XGBoost model.

Pairwise GeoDetector interactions showed stronger spatial associations than individual factors, and precipitation-related combinations remained prominent after the joint-stratum support criterion was applied; detailed interaction q-values and pairwise rankings are provided in Supplementary Materials (Figure S4).

Importantly, GeoDetector interaction q-values indicate enhanced spatial association rather than causal interaction effects. The robust interaction structure therefore supports the interpretation that precipitation, vegetation, atmospheric conditions, and anthropogenic context jointly contributed to the regional spatial patterning of forest-fire occurrence.

Sensitivity analyses showed that leading spatial-association rankings were generally stable across KDE bandwidths and response definitions; detailed q-value ranges and rank comparisons are provided in the Supplementary Materials (Tables S5–S7).

3.5. Future Climate Sensitivity of Forest-Fire Susceptibility

Historical Climate-scale Susceptibility Distribution: The climate-scale baseline represents model-derived relative susceptibility estimated from harmonized climatological, vegetation, topographic, and socioeconomic predictors across stable forest grids, rather than the KDE concentration of observed fire records. The independent climate-scale XG-

Boost model achieved a pooled five-fold ROC-AUC of 0.830 and was subsequently applied to 4463 stable forest grids to generate the historical susceptibility baseline.

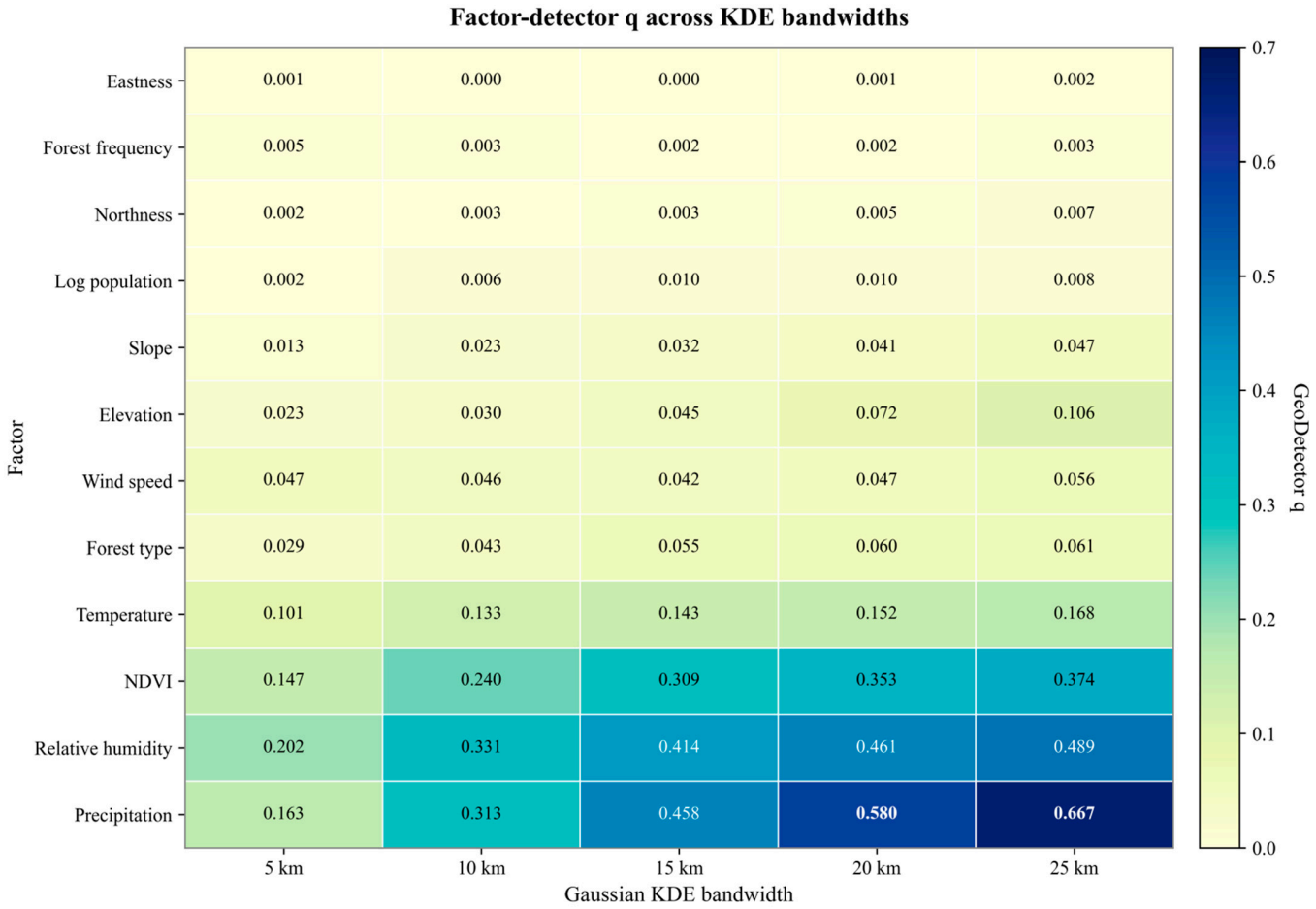


Figure 7. GeoDetector-based spatial associations of environmental variables with historical fire-density patterns across five KDE bandwidths. Continuous variables were discretized using a consistent classification scheme across bandwidths, whereas forest type retained original IGBP categories.

Under the 2003–2022 climatological baseline, higher modeled relative susceptibility was concentrated along the central and eastern sections of the main Dabie Mountains ridge, particularly within the mountainous transition between Lu'an and Anqing (Figure 8). Additional medium-to-high susceptibility patches occurred in surrounding hilly areas, whereas lower susceptibility values were more widespread in western and northwestern regions, including parts of Xinyang, Xiaogan, and Huanggang. Detailed environmental comparisons between high-susceptibility and remaining forest areas are provided in the Supplementary Materials (Table S9).

Future Scenario-Specific Susceptibility Patterns: Future susceptibility responses were evaluated relative to the historical climate-scale baseline by comparing modeled susceptibility patterns across future periods and climate scenarios. The resulting susceptibility patterns exhibited pronounced heterogeneity across future periods and climate scenarios (Figure 9). Positive susceptibility-score changes formed a spatially organized pattern, mainly along the central mountain ridge, northern mountainous slopes, and forest–human transition zones, whereas negative changes were primarily distributed in parts of the southern and central–southern regions.

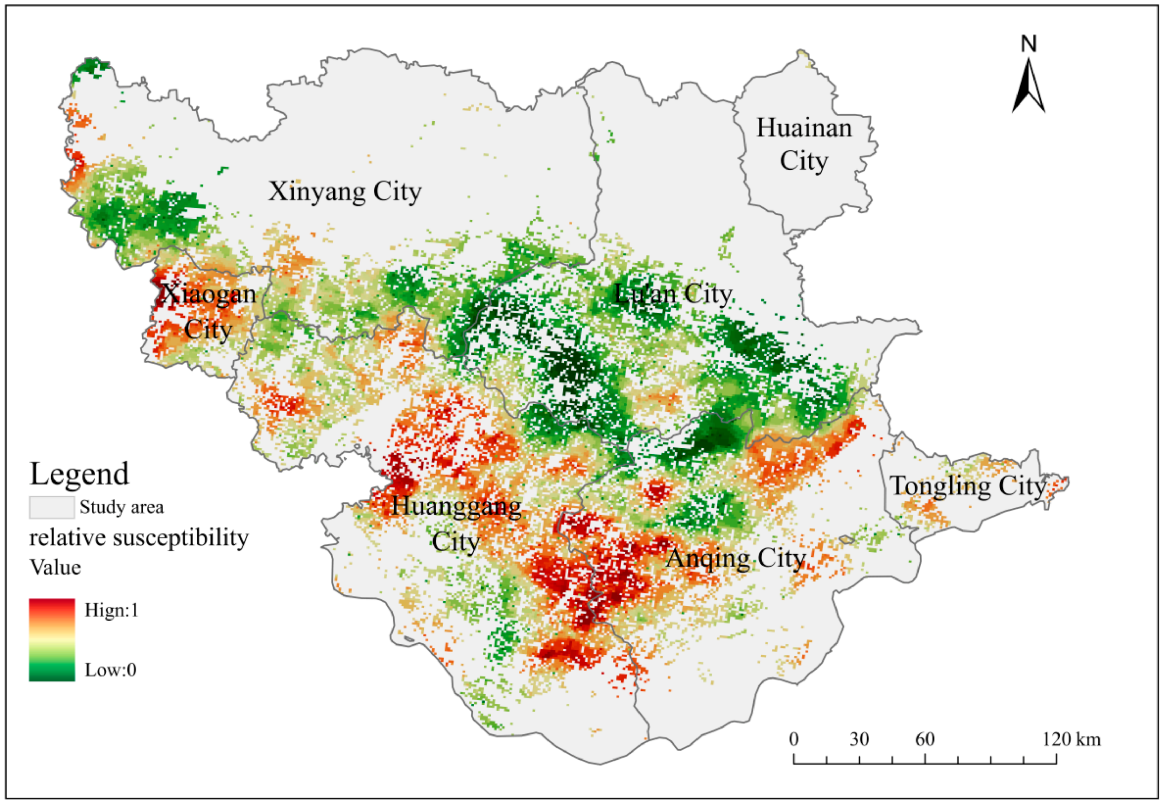


Figure 8. Baseline forest fire occurrence relative susceptibility in the Dabie Mountains under current climate conditions. Predictions were generated using the climate-scale XGBoost model based on the 2003–2022 climatological predictor set. Forest fire occurrence relative susceptibility was predicted only within forest-covered areas; gray areas indicate non-forest regions excluded by the forest mask.

Under SSP126, positive susceptibility changes remained relatively fragmented during the early and middle future periods, with localized increases mainly occurring along the main ridge and adjacent mountainous areas. By the late century, these positive-change areas became more spatially organized but remained relatively constrained. In contrast, SSP585 showed broader and more continuous positive susceptibility-score changes, particularly after the 2050s, with more extensive spatial expression along the main ridge and northern mountainous regions.

The temporal evolution of modeled susceptibility patterns indicated that future climate sensitivity was not characterized by a uniform regional increase. Instead, different areas exhibited contrasting responses, with some forest regions showing increased susceptibility while others experienced reduced modeled susceptibility. The increasingly distinct spatial patterns under SSP585 suggest stronger scenario-dependent reorganization of susceptibility distribution under the higher-emission pathway.

Threshold-based class proportions indicate stronger redistribution under SSP585 than SSP126, consistent with the observed scenario-dependent redistribution patterns; exact percentages are provided in the Supplementary Materials (Table S8). The class proportions summarize threshold-based redistribution of relative susceptibility and do not represent unconditional fire probabilities.

Scenario-specific relative-susceptibility maps further showed that, under SSP126, comparatively higher-susceptibility areas remained mainly associated with the main ridge and adjacent mountainous areas. Under SSP585, late-century patterns became more heterogeneous, with persistent higher-susceptibility belts along the main ridge and northern slopes but more extensive lower-susceptibility patches emerging in southern and south-

eastern forest areas. Together, these results indicate stronger scenario-dependent spatial reorganization of modeled susceptibility patterns under SSP585 than under SSP126.

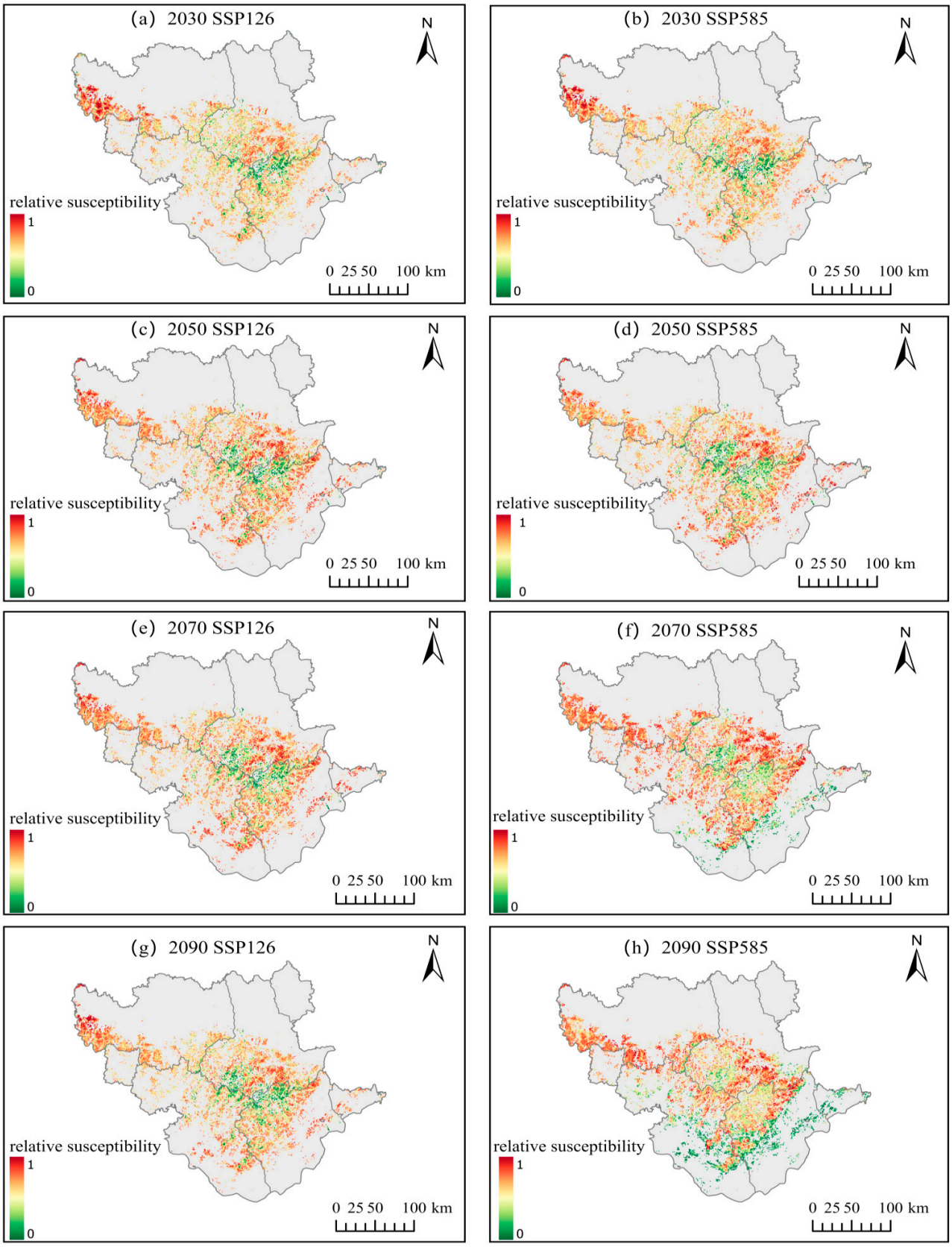


Figure 9. Model-based forest-fire occurrence susceptibility under SSP126 and SSP585 for the four future periods. Gray areas represent non-forest regions excluded from the susceptibility analysis.

4. Discussion

4.1. Historical Fire Patterns and Environmental Predictor Interpretation

Historical fire occurrence in the Dabie Mountains exhibited persistent spatial concentration in the south–central region, although local density and dispersion patterns varied with analytical scale and temporal period. KDE analysis showed that the dominant concentration area remained relatively stable across bandwidths, whereas local hotspot extent changed with spatial smoothing parameters. Previous studies have demonstrated that KDE-based fire-pattern analysis is sensitive to bandwidth selection and spatial support, and therefore hotspot interpretation should consider analytical scale rather than relying on a single spatial representation [58–60]. The SDE analysis further indicated directional persistence with temporal variation in fire distribution, suggesting that historical fire occurrence reflected both regional spatial constraints and temporal variability. Similar heterogeneous fire-occurrence patterns have been reported in Anhui and surrounding mountainous regions, although differences in temporal coverage, response definition, spatial resolution, and modeling frameworks limit direct comparison [24,25].

The environmental association analysis suggests that historical fire-density patterns were related to combined climatic, vegetation, and environmental conditions rather than a single dominant factor. GeoDetector identified precipitation and relative humidity as the strongest spatial associations across KDE bandwidths, followed by NDVI and temperature [61]. These findings indicate that moisture availability and vegetation conditions were important components of regional fire-density organization. However, GeoDetector q-values represent spatial association strength rather than causal effects, and should therefore be interpreted as indicators of spatial consistency rather than direct environmental controls [10]. Differences between studies identifying different dominant predictors, such as NDVI or anthropogenic variables, may result from variations in study extent, response definitions, predictor representations, and analytical scales [18,26]. Therefore, environmental associations should be interpreted within the context of the spatial and temporal framework used for analysis [62].

The contrasting patterns between GeoDetector and SHAP reflect their different analytical purposes: GeoDetector describes spatial associations in historical fire-density patterns, whereas SHAP explains predictor contributions within the event-level XGBoost model. Therefore, differences in variable rankings should be interpreted as complementary information rather than inconsistency [63]. In this study, precipitation showed stronger spatial association in GeoDetector, whereas population exhibited larger event-level SHAP contributions. These differences reflect the distinct analytical purposes of the two approaches rather than methodological inconsistency [64].

In addition, the nonlinear and interaction patterns identified by SHAP indicate that anthropogenic conditions were incorporated into model predictions through context-dependent relationships with vegetation, atmospheric, and terrain variables rather than acting independently.

4.2. Explainable Modeling Reliability and Uncertainty

The reliability of the susceptibility model was supported by comparative evaluation, conventional cross-validation, and spatially separated validation. XGBoost achieved higher predictive performance than RF and RBF-SVM, supporting its ability to capture nonlinear relationships and interactions among heterogeneous environmental predictors [65]. However, the reduction in performance under spatial-block validation indicates that spatial dependence contributed to predictive variability, emphasizing the importance of evaluating model transferability beyond random data partitioning [14,66,67]. The comparison of antecedent meteorological representations further indicated that temporal aggregation

influenced model performance, although the evaluated configurations showed broadly similar predictive ability.

Beyond predictive performance, the integration of SHAP provided a framework for examining predictor contributions and their stability across validation folds. The consistent ranking of major predictors, particularly the repeated high contribution of the log-transformed population, suggests that the dominant attribution pattern was relatively robust to variation in training–validation partitions. However, interpretation remains dependent on the observed data structure and modelling framework [15–17].

Several uncertainties should also be considered when interpreting the model outputs. The balanced case–control design and pseudo-absence construction support relative discrimination but do not represent regional fire incidence. In addition, MODIS active-fire products may omit small or short-duration fires, while the approximately 1 km analytical grid and coarser climate information limit fine-scale representation.

Future improvements should include higher-resolution and temporally consistent fire observations, multi-model CMIP6 ensembles, dynamic socioeconomic predictors, and additional spatial validation approaches to further evaluate the robustness of susceptibility patterns under changing environmental conditions.

4.3. Future Climate Sensitivity, Implications, and Future Directions

The climate-scale model provided a consistent framework for evaluating future susceptibility sensitivity relative to the historical baseline. Under the BCC-CSM2-MR climate scenarios, SSP585 produced broader and more spatially concentrated positive susceptibility-score changes than SSP126, while other areas exhibited negative changes, indicating heterogeneous spatial redistribution rather than a uniform regional response. The contrast between positive susceptibility-score changes and threshold-based susceptibility classes also highlights that future responses should be interpreted as spatial redistribution of modeled susceptibility rather than a simple expansion or contraction of high-susceptibility areas. These projected patterns should be interpreted as scenario-specific modeled susceptibility responses under the selected climate forcing, baseline conditions, and model assumptions, rather than deterministic predictions of future fire occurrence [21–23].

The mapped susceptibility patterns provide regional-scale information that may support identification of areas requiring further monitoring and assessment. However, these outputs should be integrated with local fire-weather conditions, fuel characteristics, field observations, accessibility information, and operational management practices rather than used as independent fire forecasts. The different analytical scales represented by KDE, GeoDetector, SHAP, and climate-scale susceptibility mapping should also be considered when combining these results for regional assessment [68].

5. Conclusions

This study developed an explainable XGBoost-based forest-fire occurrence susceptibility prediction framework, with SHAP model interpretation, GeoDetector spatial association analysis, and future climate sensitivity assessment as complementary components.

Historical fire occurrence in the Dabie Mountains showed persistent concentration in the south–central region, although hotspot extent and dispersion varied with spatial representation. GeoDetector revealed that precipitation, relative humidity, NDVI, and temperature were consistently associated with historical fire-density organization across KDE bandwidths, whereas SHAP identified different predictor contributions within the event-level model. These differences demonstrate that spatial associations and predictive contributions capture complementary dimensions of fire susceptibility.

The XGBoost model achieved strong predictive discrimination (ROC-AUC = 0.850) and retained useful discrimination under spatial-block validation and stable attribution patterns based on SHAP analysis. The independent climate-scale model further indicated that future climate scenarios may reorganize susceptibility patterns, with SSP585 showing broader late-century susceptibility-score changes than SSP126. These outputs should be interpreted as relative susceptibility patterns and scenario-specific climate responses rather than deterministic predictions of future fire occurrence.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18203521/s1, Figure S1: Receiver operating characteristic (ROC) curves for XGBoost, Random Forest, and RBF-SVM based on pooled out-of-fold predictions from five-fold cross-validation. Figure S2: Performance across five cross-validation folds of ROC-AUC, PR-AUC, accuracy, precision, recall, and F1-score for XGBoost, Random Forest, and RBF-SVM. Figure S3: Bootstrap uncertainty in XGBoost performance and SHAP ranking stability. (a1) Model-refitting uncertainty based on bootstrap resampling, model refitting, and out-of-bag evaluation; (a2) sampling uncertainty of pooled OOF ROC-AUC based on bootstrap resampling of fixed OOF prediction–label pairs; (b) bootstrap frequencies of SHAP feature-importance ranks. Dashed lines in (a1) and (a2) indicate the original pooled OOF ROC-AUC and the 2.5th and 97.5th percentiles. Figure S4. Spatial associations among environmental factor interactions identified by GeoDetector across KDE bandwidths. (a) Factor pairs with the highest median interaction q-values, including combinations with limited joint-stratum support; (b) robust interaction pairs satisfying the minimum joint-stratum support criterion across all bandwidths. Points indicate median interaction q-values, and horizontal ranges represent variation across bandwidths. Table S1: Predictive performance of XGBoost, Random Forest, and RBF-SVM based on pooled out-of-fold predictions from five-fold cross-validation. Table S2: Sensitivity of event-level XGBoost performance and meteorological attribution to different antecedent-weather representations. Table S3: Reported stability information for SHAP feature-importance rankings across the five validation folds. Table S4: Bootstrap summary of log-transformed population importance-rank stability. Table S5: Fixed-scheme q values and factor-ranking stability across GeoDetector discretization schemes. Table S6: Reported GeoDetector q-value ranges across the five KDE bandwidths. Table S7: Comparison between KDE-based and unsmoothed fire-count GeoDetector responses. Table S8: Future susceptibility-class proportions under SSP126 and SSP585 scenarios. Table S9: Environmental characteristics of high-susceptibility and remaining forest areas under the historical climate-scale baseline.

Author Contributions: Conceptualization, F.W., X.H., M.C. and Y.S.; methodology, F.W., X.H., M.C. and Y.S.; software, F.W. and Y.S.; validation, X.H., M.C. and Y.S.; formal analysis, F.W., M.C. and Y.S.; investigation, F.W. and M.C.; data curation, F.W. and Y.S.; writing—original draft preparation, F.W. and M.C.; writing—review and editing, F.W., X.H., M.C. and Y.S.; visualization, F.W. and Y.S.; supervision, X.H., M.C. and Y.S.; project administration, F.W., X.H. and Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Foundation of Anhui Province Key Laboratory of Physical Geographic Environment, grant number 2025PGE009.

Data Availability Statement: No new primary data were collected in this study. The data used in this study were obtained from publicly available datasets, including MODIS, ERA5-Land, SRTM, MOD13A2, MCD12Q1, WorldPop, and WorldClim CMIP6, as described and cited in the manuscript. The processed data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

Abbreviation	Full Name
AUC	Area Under the Receiver Operating Characteristic Curve
DEM	Digital Elevation Model
ERA5-Land	ECMWF Reanalysis v5 Land Dataset
GeoDetector	Geographical Detector
KDE	Kernel Density Estimation
RBF-SVM	Radial Basis Function Support Vector Machine
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index
RF	Random Forest
SDE	Standard Deviational Ellipse
VIF	Variance Inflation Factor
XGBoost	Extreme Gradient Boosting
BCC-CSM2-MR	Beijing Climate Center Climate System Model, Medium Resolution
CMIP6	Coupled Model Intercomparison Project Phase 6
SSP	Shared Socioeconomic Pathway
SHAP	SHapley Additive exPlanations

References

1. Jones, M.W. Global and Regional Trends and Drivers of Fire Under Climate Change. *Rev. Geophys.* **2022**, *60*, e2020RG000726. [CrossRef]
2. Lan, Y.; Wang, J.; Hu, W.; Kurbanov, E.; Cole, J.; Sha, J.; Jiao, Y.; Zhou, J. Spatial Pattern Prediction of Forest Wildfire Susceptibility in Central Yunnan Province, China Based on Multivariate Data. *Nat. Hazards* **2022**, *116*, 565–586. [CrossRef]
3. Liu, J.; Wang, Y.; Lu, Y.; Zhao, P.; Wang, S.; Yu, S.; Luo, Y. Application of Remote Sensing and Explainable Artificial Intelligence (XAI) for Wildfire Occurrence Mapping in the Mountainous Region of Southwest China. *Remote Sens.* **2024**, *16*, 3602. [CrossRef]
4. Cui, S.; Liu, J.; Tian, Y.; Chen, S. Explainable Machine Learning Identifies Anthropogenic Activity as a Key Driver of Forest Fire Severity in China's Temperate Monsoon Region. *GISci. Remote Sens.* **2025**, *62*, 2599489. [CrossRef]
5. Ji, X.; Li, X.; Zhang, D.; Liu, W.; Zhang, W.; Zhang, Z. Forecast Zoning of Forest Fire Occurrence: A Case Study in Southern China. *Forests* **2024**, *15*, 265. [CrossRef]
6. Koutsias, N.; Balatsos, P.; Kalabokidis, K. Fire Occurrence Zones: Kernel Density Estimation of Historical Wildfire Ignitions at the National Level, Greece. *J. Maps* **2014**, *10*, 630–639. [CrossRef]
7. Koutsias, N.; Kalabokidis, K.; Allgöwer, B. Fire Occurrence Patterns at Landscape Level: Beyond Positional Accuracy of Ignition Points with Kernel Density Estimation Methods. *Nat. Resour. Model.* **2004**, *17*, 359–375. [CrossRef]
8. Wang, J.; Zhang, T.; Fu, B. A Measure of Spatial Stratified Heterogeneity. *Ecol. Indic.* **2016**, *67*, 250–256. [CrossRef]
9. Guo, J.; Wang, J.; Xu, C.; Song, Y. Modeling of Spatial Stratified Heterogeneity. *GISci. Remote Sens.* **2022**, *59*, 1660–1677. [CrossRef]
10. Zhang, Z.; Song, Y.; Karunaratne, L.; Wu, P. Robust Interaction Detector: A Case of Road Life Expectancy Analysis. *Spat. Stat.* **2024**, *59*, 100814. [CrossRef]
11. Shmuel, A.; Heifetz, E. Global Wildfire Susceptibility Mapping Based on Machine Learning Models. *Forests* **2022**, *13*, 1050. [CrossRef]
12. Akıncı, H.A.; Akinci, H.; Zeybek, M. Comparison of Diverse Machine Learning Algorithms for Forest Fire Susceptibility Mapping in Antalya, Türkiye. *Adv. Space Res.* **2024**, *74*, 647–667. [CrossRef]
13. Yang, C.; Yao, P.; Wang, Q. XGBoost-Based Susceptibility Model Exhibits High Accuracy and Robustness in Plateau Forest Fire Prediction. *Forests* **2026**, *17*, 74. [CrossRef]
14. Roberts, D.R. Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure. *Ecography* **2017**, *40*, 913–929. [CrossRef]
15. Lundberg, S.M.; Erion, G.; Chen, H. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat. Mach. Intell.* **2020**, *929*, 56–67. [CrossRef] [PubMed]
16. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
17. Abohaia, Z.; Elkhouly, A.; El Barachi, M.; Al-Khatib, O. Explainable AI-Driven Wildfire Prediction in Australia: SHAP and Feature Importance to Identify Environmental Drivers in the Age of Climate Change. *Fire* **2025**, *8*, 421. [CrossRef]

18. Wang, J.; Huang, D.; Chen, C.; Liu, Y.; Wang, J.; Shao, Y.; Wang, A.; Li, X. Prediction of Forest-Fire Occurrence in Eastern China Utilizing Deep Learning and Spatial Analysis. *Forests* **2024**, *15*, 1672. [CrossRef]
19. Wu, Z.; He, H.S. Current and Future Patterns of Forest Fire Occurrence in China. *Int. J. Wildland Fire* **2020**, *29*, 104–119. [CrossRef]
20. Tian, X.; Cui, W.; Shu, L.; Zong, X.-Z. Effects of Climate Change on Burn Probability of Forests in Daxing'anling. *Forests* **2019**, *10*, 611. [CrossRef]
21. Li, J.; Huang, D.; Long, B.; Shao, Y. Prediction of Wildfire Occurrence in the Southern Forest Regions of China in the Future Scenario. *Forests* **2024**, *15*, 2029. [CrossRef]
22. Gallo, C.; Dieppois, B.; Quilcaille, Y. Future Impacts of Climate Change on Global Fire Weather: Insight from Weighted CMIP6 Multi-Model Ensembles. *J. Clim.* **2025**, *38*, 6445–6462. [CrossRef]
23. Peng, X.; Yu, M.; Chen, H.; Zhou, B.; Shi, Y.; Yu, L. Projections of Wildfire Risk and Activities under 1.5 °C and 2.0 °C Global Warming Scenarios. *Environ. Res. Commun.* **2023**, *5*, 031002. [CrossRef]
24. Zhang, X.; Lan, M.; Ming, J.; Zhu, J.; Lo, S. Spatiotemporal Heterogeneity of Forest Fire Occurrence Based on Remote Sensing Data: An Analysis in Anhui, China. *Remote Sens.* **2023**, *15*, 598. [CrossRef]
25. Zhang, J.; Zhou, H.; Zhang, B.; Song, Z.; Guo, Y.; Song, W. Integrated Geospatial Machine Learning Frameworks for Forest Fire Risk Prediction: A Data-Driven Approach Using Random Forest and Non-Linear Feature Transformation in Anhui Province. *Fire* **2026**, *9*, 291. [CrossRef]
26. Shi, C.; Zhang, F. A Forest Fire Susceptibility Modeling Approach Based on Integration Machine Learning Algorithm. *Forests* **2023**, *14*, 1506. [CrossRef]
27. Giglio, L.; Schroeder, W.; Justice, C.O. The Collection 6 MODIS Active Fire Detection Algorithm and Fire Products. *Remote Sens. Environ.* **2016**, *178*, 31–41. [CrossRef] [PubMed]
28. Warton, D.I.; Shepherd, L. Poisson Point Process Models Solve the "pseudo-Absence Problem" for Presence-Only Data in Ecology. *Ann. Appl. Stat.* **2010**, *4*, 1383–1402. [CrossRef]
29. Hanberry, B.B.; He, H.S. Prevalence, Statistical Thresholds, and Accuracy Assessment for Species Distribution Models. *Web Ecol.* **2013**, *13*, 13–19. [CrossRef]
30. Liu, Y.; Zhang, Z.; Li, J.; Zhang, Y.; Lu, S.; Ruan, X. A Survey of the Birds of the Dabie Shan Range, Central China. *Forktail* **2008**, *24*, 80–91.
31. Cai, Q.; Liu, Y.; Duan, B.; Sun, C. Regional Difference of the Start Time of the Recent Warming in Eastern China: Prompted by a 165-Year Temperature Record Deduced from Tree Rings in the Dabie Mountains. *Clim. Dyn.* **2018**, *50*, 2157–2168. [CrossRef]
32. Zheng, L.; Sun, J.; Qiu, X.; Yang, Z. Five-Year Climatology of Local Convections in the Dabie Mountains. *Atmosphere* **2020**, *11*, 1246. [CrossRef]
33. Cheng, S.; Li, Y.; Liu, S. Geomorphology Features of the Dabie Orogenic Belt Based on DEM Data. *J. East China Inst. Technol. (Nat. Sci. Ed.)* **2010**, *33*, 270–275. [CrossRef]
34. Tian, C.; Zhang, H.; Wang, J.; Li, J.; Zhang, Y.; Zha, T. Temporal and Spatial Changes of Soil Erosion and Its Response to Landscape Pattern in Dabie Mountains of Western Anhui in Recent 30 Years. *J. Soil Water Conserv.* **2024**, *38*, 37–44. [CrossRef]
35. Fang, Y.; Cai, S.; Xiang, J.; Zhong, Y.; Hu, Y.; Xia, X. Biodiversity of Dabieshan in Eastern Hubei Province and Its Protection Strategy. *J. Anhui Agric. Sci.* **2007**, *35*, 5246–5248. [CrossRef]
36. Fang, L.; Liu, Y.; Li, C.; Cai, J. Spatiotemporal Characteristics and Future Scenario Simulation of the Trade-Offs and Synergies of Mountain Ecosystem Services: A Case Study of the Dabie Mountains Area, China. *Chin. Geogr. Sci.* **2023**, *33*, 144–160. [CrossRef]
37. Huang, M.; Wang, Q.; Yin, Q.; Li, W.; Zhang, G.; Ke, Q.; Guo, Q. Analysis of Ecosystem Service Contribution and Identification of Trade-Off/Synergy Relationship for Ecosystem Regulation in the Dabie Mountains of Western Anhui Province, China. *Land* **2023**, *12*, 1046. [CrossRef]
38. Muñoz-Sabater, J. ERA5-Land: A State-of-the-Art Global Reanalysis Dataset for Land Applications. *Earth Syst. Sci. Data* **2021**, *13*, 4349–4383. [CrossRef]
39. Farr, T.G. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. [CrossRef]
40. Huete, A.; Didan, K.; Miura, T. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
41. Friedl, M.A. MODIS Collection 5 Global Land Cover: Algorithm Refinements and Characterization of New Datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [CrossRef]
42. Lloyd, C.T. High Resolution Global Gridded Data for Use in Population Studies. *Sci. Data* **2017**, *4*, 170001. [CrossRef] [PubMed]
43. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [CrossRef]
44. Wu, T.; Lu, Y.; Fang, Y.; Xin, X. The Beijing Climate Center Climate System Model (BCC-CSM): The Main Progress from CMIP5 to CMIP6. *Geosci. Model Dev.* **2019**, *12*, 1573–1600. [CrossRef]
45. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
46. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

47. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. *arXiv* **2016**. [CrossRef]
48. Ying, L.; Shen, Z.; Yang, M.; Piao, S. Wildfire Detection Probability of MODIS Fire Products under the Constraint of Environmental Factors: A Study Based on Confirmed Ground Wildfire Records. *Remote Sens.* **2019**, *11*, 3031. [CrossRef]
49. Barbet-Massin, M.; Jiguet, F.; Albert, C.H.; Thuiller, W. Selecting Pseudo-Absences for Species Distribution Models: How, Where and How Many? *Methods Ecol. Evol.* **2012**, *3*, 327–338. [CrossRef]
50. Phillips, S.J.; Dudík, M.; Elith, J.; Graham, C.H. Sample Selection Bias and Presence-Only Distribution Models: Implications for Background and Pseudo-Absence Data. *Ecol. Appl.* **2009**, *19*, 181–197. [CrossRef] [PubMed]
51. Florinsky, I.V. An Illustrated Introduction to General Geomorphometry. *Prog. Phys. Geogr. Earth Environ.* **2017**, *41*, 723–752. [CrossRef]
52. O'Brien, R.M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Qual. Quant.* **2007**, *41*, 673–690. [CrossRef]
53. Aas, K.; Jullum, M.; Løland, A. Explaining Individual Predictions When Features Are Dependent: More Accurate Approximations to Shapley Values. *Artif. Intell.* **2021**, *298*, 103502. [CrossRef]
54. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
55. Liu, W.; Wang, S.; Zhou, Y.; Wang, L.; Zhang, S. Spatial Distribution Patterns Analysis of Historical Forest Fires in DaXingAn Mountains of China. In Proceedings of the 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), Taiyuan, China, 22–24 October 2010.
56. Serra, L.; Juan, P.; Varga, D.; Mateu, J.; Saez, M. Spatial Pattern Modelling of Wildfires in Catalonia, Spain 2004–2008. *Environ. Model. Softw.* **2013**, *40*, 235–244. [CrossRef]
57. Yuill, R.S. The Standard Deviational Ellipse; An Updated Tool for Spatial Description. *Geogr. Ann. Ser. B Hum. Geogr.* **1971**, *53*, 28–39. [CrossRef]
58. Shao, Y.; Feng, Z.; Cao, M.; Wang, W.; Sun, L.; Yang, X.; Ma, T.; Guo, Z.; Fahad, S.; Liu, X.; et al. An Ensemble Model for Forest Fire Occurrence Mapping in China. *Forests* **2023**, *14*, 704. [CrossRef]
59. Wang, J.; Li, X.; Christakos, G.; Liao, Y.; Zhang, T.; Gu, X.; Zheng, X. Geographical Detectors-Based Health Risk Assessment and Its Application in the Neural Tube Defects Study of the Heshun Region, China. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 107–127. [CrossRef]
60. Cang, X.; Luo, W. Spatial Association Detector (SPADE). *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 2055–2075. [CrossRef]
61. Wang, J.; Haining, R.; Zhang, T.; Xu, C. Statistical Modeling of Spatially Stratified Heterogeneous Data. *Ann. Am. Assoc. Geogr.* **2024**, *114*, 499–519. [CrossRef]
62. Fawcett, T. An Introduction to ROC Analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
63. Steyerberg, E.W. Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures. *Epidemiology* **2010**, *21*, 128–138. [CrossRef] [PubMed]
64. Shao, Y.-K.; Li, W.-K.; Wang, M.-Y.; Du, Q.-Y.; Wang, J.; Shu, L.-F.; Si, L.-Q.; Zhao, F.-J.; Feng, Z.-K.; Sun, L.-H.; et al. Optimising Future Scenarios of Forest Fire Occurrence in Daxing'anling Using Long-Term Survey Data and Intelligent Modelling. *Adv. Clim. Change Res.* **2026**, *17*, 388–399. [CrossRef]
65. Yang, S.; Huang, Q.; Yu, M. Advancements in Remote Sensing for Active Fire Detection: A Review of Datasets and Methods. *Sci. Total Environ.* **2024**, *943*, 173273. [CrossRef] [PubMed]
66. Liu, C.; Newell, G.; White, M. The Effect of Sample Size on the Accuracy of Species Distribution Models: Considering Both Presences and Pseudo-Absences or Background Sites. *Ecography* **2019**, *42*, 535–548. [CrossRef]
67. Valavi, R.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. BlockCV: An R Package for Generating Spatially or Environmentally Separated Folds for k-Fold Cross-Validation of Species Distribution Models. *Methods Ecol. Evol.* **2019**, *10*, 225–232. [CrossRef]
68. Moreira, F. Wildfire Management in Mediterranean-Type Regions: Paradigm Change Needed. *Environ. Res. Lett.* **2020**, *15*, 011001. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.