People's Democratic Republic of Algeria

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF MILA

Faculty of Mathematics and Computer Sciences

Department of Computer Science

Doctoral Thesis

Prepared to obtain the

3rd cycle LMD Doctorate Degree in Computer Science

Specialty: Artificial Intelligence and Its Applications

Presented by **BOUZERAA Yehya**

An Intelligent Approach to Wild Fire Prediction: Case Study of Jijel Region

Jury:

Dr	Madjed BENCHEIKH LEHOCINE	University of Mila	President
Dr	Nardjes BOUCHEMAL	University of Mila	Supervisor
Dr	Salim DJAABOUB	University of Mila	Co-Supervisor
Dr	Mounira BOUZAHZAH	University of Mila	Examiner
Prof	Aissa BOULMERKA	National School of Artificial Intelligence Algiers	Examiner
Dr	Ahmed ALIOUA	University of Jijel	Examiner

Academic Year: 2025/2026

Acknowledgements

All praise is due to Allahthe Most Gracious, the Most Mercifulwho granted me faith, perseverance, and the opportunity to bring this work to completion.

I owe my deepest gratitude to my supervisor, Dr. Nardjes Bouchemal, for her steady guidance, rigorous standards, and generous support throughout my doctoral studies. Her patience, scientific insight, and constructive feedback shaped this thesis at every stage and greatly improved its quality.

I also extend my sincere thanks to the members of the jury for dedicating their time to read, evaluate, and enrich this manuscript. Their thoughtful comments and recommendations have been invaluable in strengthening the scientific contribution of this work.

My profound thanks go to my family, whose unconditional love, sacrifices, and prayers have been my constant source of strength and motivation. This achievement belongs to them as much as it does to me.

I am equally grateful to my friends, colleagues, and collaborators for their encouragement, technical assistance, and moral support, as well as to the administrative and technical staff who, often behind the scenes, helped make this journey smoother.

Finally, may Allah send His blessings upon His noble Messenger, his family, and his companions, and grant us goodness in our lives.

Abstract

Wildfires have intensified across the Mediterranean under compounding heatwaves, drought, and human pressure, making a shift from reaction to prevention essential. This thesis develops an operational wildfire susceptibility mapping (WSM) framework for Jijel Province (Algeria) that integrates Geographic Information Systems (GIS) and machine learning. Fourteen conditioning factorstopography (elevation, slope, aspect), climate (minimum, maximum, and average temperature, humidity, wind speed, precipitation), vegetation (Normalized Difference Vegetation Index, NDVI), and anthropogenic pressure (distance to roads and settlements)are harmonized on a 30,m grid and linked with civil protection fire records, yielding an analysis-ready matrix of approximately 2.6 million samples. Collinearity screening (maximum Variance Inflation Factor, VIF = 7.58; minimum tolerance = 0.11) and light k-means denoising of the *Low* class improve model stability.

We train Random Forest (RF), Extreme Gradient Boosting (XGBoost), a shallow Neural Network (NN), and Support Vector Machine (SVM) models with class balancing and probability calibration. On the internal holdout set, RF attains an Area Under the Receiver Operating Characteristic Curve (AUC) of 0.99 and an F1-score of 0.98, with XGBoost and NN performing similarly, and SVM showing solid performance (AUC = 0.94). Continuous susceptibility scores are calibrated and discretized into five classes (LowVery High). Independent time-forward validation overlays wildfire occurrences from 20242025 (Civil Protection records and Moderate Resolution Imaging Spectroradiometer, MODIS, data) on the susceptibility maps: 87.73

The outputscalibrated probability surfaces, clearly legended risk classes, and uncertainty layersare suitable for integration into web-based GIS dashboards to support resource prepositioning, fuel-break maintenance at the wildlandurban interface (WUI), and threshold-based early warning systems. Limitations (sparse meteorological stations and ignition geolocation uncertainty) and targeted improvements (denser in situ sensing and spatially blocked cross-validation) are discussed. The framework is transferable to similar Mediterranean environments.

Résumé

Les incendies de forêt se sont intensifiés dans le bassin méditerranéen sous leffet combiné des vagues de chaleur, des déficits hydriques et de la pression humaine, imposant un basculement de la réaction vers la prévention. Cette thèse propose un cadre opérationnel de *cartographie de la susceptibilité aux feux de forêt* (Wildfire Susceptibility Mapping, WSM) pour la wilaya de Jijel (Algérie), fondé sur lintégration des Systèmes dInformation Géographique (SIG) et de lapprentissage automatique. Quatorze facteurs de conditionnementtopographie (altitude, pente, exposition), climat (températures minimale, maximale et moyenne, humidité, vent, précipitations), végétation (Indice de Végétation par Différence Normalisée, NDVI) et pression anthropique (distance aux routes et aux zones habitées)sont harmonisés sur une grille de 30,m et croisés avec les enregistrements de la Protection Civile, produisant une matrice danalyse denviron 2,6 millions déchantillons. Le criblage de la colinéarité (Facteur dInflation de la Variance maximal, VIF = 7,58 ; tolérance minimale = 0,11) et un léger débruitage par k-means de la classe *Faible* améliorent la stabilité des modèles.

Nous entraînons des modèles de Forêt Aléatoire (Random Forest, RF), dExtreme Gradient Boosting (XGBoost), de réseau de neurones peu profond (NN) et de Machine à Vecteurs de Support (Support Vector Machine, SVM), avec équilibrage des classes et calibration probabiliste. En validation interne, le modèle RF atteint une Aire Sous la Courbe ROC (AUC) de 0,99 et un score F1 de 0,98 ; XGBoost et le réseau de neurones présentent des performances comparables, tandis que le SVM demeure robuste (AUC = 0,94). Les scores continus de susceptibilité sont calibrés puis discrétisés en cinq classes (Faible à Très Élevée). Une validation indépendante *time-forward* superpose les occurrences de feux de 20242025 (Protection Civile et données du radiomètre MODIS Moderate Resolution Imaging Spectroradiometer) aux cartes produites : 87,73,

Les sortiessurfaces de probabilité calibrées, classes de risque clairement légendées et couches dincertitudesont prêtes pour une intégration dans des tableaux de bord web-SIG afin de soutenir le pré-positionnement des ressources, lentretien des interfaces forêturbain (WildlandUrban Interface, WUI) et les systèmes dalerte précoce à seuils. Les limites (réseau de stations météorologiques clairsemé, incertitudes de géolocalisation des départs de feu) et les pistes damélioration (densification des capteurs in situ, validation croisée spatiale bloquée) sont discutées ; le cadre proposé est transférable à dautres contextes méditerranéens.

ملخص البحث

شهدت حرائق الغابات في حوض البحر الأبيض المتوسط تزايداً ملحوظاً نتيجة تضافر موجات الحر والجفاف المتزايد والضغوط البشرية، مما جعل الانتقال من منطق الاستجابة إلى استراتيجية الوقاية أمراً حتمياً. تقترح هذه الأطروحة إطاراً تشغيلياً لنمذجة قابلية التعرض لحرائق الغابات في ولاية جيجل بالجزائر، وذلك عبر دمج نظم المعلومات الجغرافية وتقنيات التعلم الآلي. تم الاعتماد على أربعة عشر عاملاً مؤثراً تشمل الخصائص الطبوغرافية من ارتفاع وانحدار وتوجيه، والبيانات المناخية المتعلقة بدرجات الحرارة والرطوبة وسرعة الرياح والتساقط المطري، بالإضافة إلى الحالة الخضرية والضغط البشري المتمثل في القرب من الطرق والتجمعات السكنية. تم توحيد هذه البيانات وفق شبكة بدقة 03 متراً وربطها بسجلات الحرائق الخاصة بالحماية المدنية، مما أتاح الحصول على قاعدة بيانات ضخمة تضم نحو مليونين وستمائة ألف عينة جاهزة للتحليل. وقد أظهرت الاختبارات الإحصائية استقرار البيانات بعد معالجة التداخل الخطي وتطبيق خوارزميات التصفية لتحسين دقة النتائج.

تضمنت الدراسة تدريب نماذج متقدمة شملت الغابات العشوائية وtsooBGX والشبكات العصبية وآلات المتجهات الداعمة، مع معايرة دقيقة للاحتمالات لضمان توازن النتائج. وقد أظهر نموذج الغابات العشوائية أداءً استثنائياً في مرحلة التحقق الداخلي، متبوعاً بنتائج متقاربة للنماذج الأخرى. ومن أجل اختبار القدرة التنبؤية للإطار المقترح، تم إسقاط حرائق الفترة الممتدة بين 4202 و5202 المستمدة من بيانات الحماية المدنية والاستشعار عن بعد على الخرائط المنتجة، حيث تبين أن الغالبية العظمى من الحرائق الجديدة وقعت ضمن المناطق المصنفة كذات خطورة متوسطة إلى عالية جداً، مما يؤكد دقة النموذج. كما أشارت تحليلات الأهمية إلى أن الأنشطة البشرية تمثل المحرك الأساسي لنشوب الحرائق، تليها الضغوط الهيدرومناخية، بينما تلعب الطبوغرافيا دوراً في توجيه انتشار النيران.

تعد المخرجات المتمثلة في خرائط الاحتمالات وطبقات المخاطر الجاهزة للاستخدام ضمن لوحات التحكم الجغرافية أداة هامة لدعم اتخاذ القرار، سواء في التوزيع المسبق لموارد الإطفاء أو صيانة حواجز النار عند الواجهات العمرانية الغابية وتفعيل أنظمة الإنذار المبكر. وتختتم الدراسة بمناقشة التحديات المتعلقة بتوزيع محطات الرصد ودقة تحديد المواقع، مع التأكيد على إمكانية نقل هذا الإطار المنهجي وتطبيقه في مناطق متوسطية أخرى ذات خصائص مشابهة.

الكلمات المفتاحية: خرائط قابلية الاشتعال، التعلم الآلي، التحليل الجغرافي المكاني، تحليل أهمية العوامل، الوقاية من حرائق الغابات.

Contents

List of Figures

List of Tables

List of Abbreviations

WSM	Wildfire Susceptibility Mapping
EWS	Early Warning System
WUI	WildlandUrban Interface
GIS	Geographic Information Systems
RS	Remote Sensing
IoT	Internet of Things
UAV	Unmanned Aerial Vehicle
WSN	Wireless Sensor Network
ML	Machine Learning
DL	Deep Learning
AI	Artificial Intelligence
RF	Random Forest
SVM	Support Vector Machine
NN	Neural Network
MLP	Multi-Layer Perceptron
MLP-NET	Multi-Layer Perceptron Network
DNN	Deep Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory

GBM	Gradient Boosting Machine
XGBoost	Extreme Gradient Boosting
LightGBM	Light Gradient Boosting Machine
CatBoost	Categorical Boosting
BRT	Boosted Regression Trees
KNN	k-Nearest Neighbors
DT	Decision Tree
LR	Logistic Regression
ANFIS	Adaptive Neuro-Fuzzy Inference System
MCDA	Multi-Criteria Decision Analysis
AHP	Analytic Hierarchy Process
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
VIKOR	Multi-criteria Optimization and Compromise Solution
FR	Frequency Ratio
GA	Genetic Algorithm
SHAP	Shapley Additive Explanations
AUC	Area Under the ROC Curve
ROC-AUC	Receiver Operating Characteristic Area Under the Curve
PR-AUC	PrecisionRecall Area Under the Curve
TPR	True Positive Rate
FPR	False Positive Rate
CV	Cross-Validation
VIF	Variance Inflation Factor
DEM	Digital Elevation Model
SRTM	Shuttle Radar Topography Mission

ALOS	Advanced Land Observing Satellite
MODIS	Moderate Resolution Imaging Spectroradiometer
VIIRS	Visible Infrared Imaging Radiometer Suite
NDVI	Normalized Difference Vegetation Index
NDMI	Normalized Difference Moisture Index
ERA5	ECMWF Reanalysis Version 5
WGS84	World Geodetic System 1984
UTM	Universal Transverse Mercator
NOM	National Office of Meteorology (Algeria)
GWIS	Global Wildfire Information System
EFFIS	European Forest Fire Information System
MSG	Meteosat Second Generation
RGB	Red Green Blue
CO	Carbon Monoxide
API	Application Programming Interface
GPU	Graphics Processing Unit
MQTT	Message Queuing Telemetry Transport
HTTP	Hypertext Transfer Protocol
SMS	Short Message Service
QA/QC	Quality Assurance and Quality Control
R-CNN	Region-based Convolutional Neural Network
YOLO	You Only Look Once
EfficientDet	Efficient Object Detection model family
AI (Aridity)	Aridity Index

General introduction

Wildfires have emerged as one of the most critical environmental hazards of the twenty-first century, posing severe threats to ecosystems, human lives, infrastructure, and socio-economic stability worldwide. In recent decades, their frequency, intensity, and spatial extent have increased markedly, transforming wildfires from episodic natural disturbances into recurrent disasters with long-lasting consequences. Recent global assessments report approximately 2,805 direct fatalities, 8,540 injuries, and impacts on more than 7 million people, while wildfires burn nearly 400 million hectares annually at the global scale. These alarming figures reflect the convergence of climatic change, environmental stress, and intensified human activities, which together amplify wildfire risk across regions and continents [1].

The environmental consequences of wildfires are profound. Large-scale burning contributes significantly to land degradation, biodiversity loss, and ecosystem fragmentation, while also accelerating soil erosion and altering hydrological regimes. From a climatic perspective, wildfires act as both a consequence and a driver of climate change by releasing substantial quantities of greenhouse gases and aerosols into the atmosphere. Globally, an estimated hundreds of millions of hectares burn each year, with Africa alone accounting for a substantial share of this total, highlighting the disproportionate vulnerability of certain regions. However, wildfire disasters are not confined to developing countries; recent catastrophic events in technologically advanced regions such as North America, Southern Europe, and Australia demonstrate that no nation is immune to the escalating wildfire crisis [2].

Among the most affected regions, the Mediterranean basin occupies a particularly critical position. Characterized by hot, dry summers, complex topography, and dense humanenvironment interactions, Mediterranean landscapes are inherently prone to wildfire ignition and spread. Over the past decades, the region has experienced some of the deadliest wildfire events on record, resulting in significant loss of life, extensive forest damage, and long-term ecological disruption. Algeria, located in the southern Mediterranean, exemplifies these challenges. Despite possessing vast forested and biodiverse areas, the country has suffered extensive wildfire damage, particularly in its northern and eastern regions. Historical records indicate that a substantial proportion of Algerias forest cover has been lost to wildfires over the past half-century, with recent years marked by increasingly destructive fire seasons and rising human casualties [3].

The drivers underlying this alarming increase in wildfire activity are multifactorial. Climate change plays a central role by altering temperature regimes, precipitation patterns, and atmospheric circulation. Rising temperatures, prolonged droughts, reduced relative humidity, and increased frequency of heatwaves create conditions that dry vegetation and enhance fuel flammability. These climatic stresses are often compounded by extreme weather events, such as strong winds, which can rapidly transform small ignitions into large, uncontrollable fires. At the same time, climate-driven shifts in vegetation structure and productivity can increase fuel continuity, further amplifying wildfire risk [4].

Equally important are anthropogenic factors, which have been consistently identified as the dominant source of wildfire ignitions in many regions. Human activities such as agricultural burning, land-use change, deforestation, infrastructure expansion, and increased accessibility to wildland areas significantly elevate ignition pressure. In Mediterranean contexts, the expansion of wild-landurban and wildlandagriculture interfaces has created landscapes where human presence and combustible fuels intersect, greatly increasing the likelihood of fire occurrence. Reports from European and international fire monitoring systems consistently indicate that the majority of wildfires are directly or indirectly linked to human actions, underscoring the need to explicitly integrate anthropogenic variables into wildfire risk assessments [5].

In light of these escalating threats, proactive wildfire management has become an urgent priority. Traditional reactive approaches focused solely on fire suppression are no longer sufficient in the face of increasing fire frequency and intensity. Instead, there is a growing emphasis on prevention, preparedness, and risk-informed decision-making. Within this paradigm, Wildfire Susceptibility Mapping (WSM) has emerged as a critical tool for identifying areas that are more likely to experience wildfire ignition and spread. By delineating zones of varying susceptibility, WSM supports strategic planning, resource allocation, early warning, and targeted prevention measures [6].

Wildfire susceptibility mapping aims to quantify the spatial likelihood of wildfire occurrence based on the analysis of conditioning factors that influence ignition and fire behavior. These factors typically include climatic variables, vegetation characteristics, topographic features, hydrological conditions, and indicators of human activity. By examining the relationships between historical fire occurrences and these explanatory variables, WSM enables the classification of landscapes into susceptibility levels ranging from low to very high. Such spatially explicit information is invaluable for fire management agencies, land-use planners, and policymakers seeking to reduce wildfire impacts [7].

However, producing reliable wildfire susceptibility maps presents significant methodological challenges. One of the primary difficulties lies in data availability and quality, particularly in regions where monitoring infrastructure is limited. Environmental datasets often vary in spatial resolution, temporal coverage, and accuracy, while fire occurrence records may contain positional errors or reporting biases. To address these challenges, many studies have adopted data integration strategies that combine global-scale datasets from international organizations with locally recorded information from regional agencies. This approach allows researchers to leverage the strengths of large-scale products while preserving the contextual specificity provided by local data sources [8].

Geographic Information Systems (GIS) play a central role in this integration process. GIS technologies provide the spatial framework necessary to collect, harmonize, analyze, and visualize heterogeneous datasets. Through GIS, environmental layers can be transformed into a common spatial reference, enabling pixel-based analysis and the exploration of spatial relationships between wildfire occurrences and conditioning factors. Beyond data management, GIS visualization capabilities facilitate the interpretation of results by highlighting spatial patterns, hotspots, and gradients of wildfire susceptibility that may not be apparent through numerical analysis alone [9].

Historically, wildfire susceptibility assessment has relied on traditional statistical and multi-criteria decision analysis (MCDA) methods, such as frequency ratio, analytic hierarchy process, and weighted overlay techniques. These methods are valued for their conceptual simplicity and interpretability, and they remain useful in data-limited contexts. However, wildfire systems are inherently complex and nonlinear, involving interactions among numerous environmental and anthropogenic factors. As the number of predictors and the volume of available data increase, conventional methods may struggle to capture these complex relationships effectively [10].

In response to these limitations, machine learning (ML) approaches have gained increasing prominence in wildfire susceptibility modeling. Techniques such as Random Forest, Support Vector Machines, gradient boosting algorithms, and neural networks have demonstrated strong performance in handling high-dimensional, nonlinear, and noisy datasets. These models are capable of learning intricate patterns from historical data and translating them into accurate predictive maps. Consequently, machine learning has become a cornerstone of modern WSM research, offering improved predictive accuracy and robustness compared to many traditional approaches [11].

Despite their advantages, machine learning models introduce new challenges, particularly regarding interpretability and operational trust. High predictive performance alone is insufficient for practical wildfire management if model outputs cannot be understood or justified by domain experts. Fire management agencies require not only accurate predictions but also clear explanations of why certain areas are classified as high risk. This need has driven the growing integration of explainable artificial intelligence (XAI) techniques, such as feature importance analysis and SHAP (SHapley Additive exPlanations), into wildfire modeling frameworks. These tools enable researchers to identify the most influential factors driving model predictions and to link statistical results with physical and human processes on the ground [12].

Another critical limitation of many existing WSM studies lies in model validation. While internal validation metrics such as accuracy and AUC are widely reported, relatively few studies perform independent, time-forward validation using wildfire events from subsequent years. As a result, high internal performance does not always translate into reliable real-world predictive skill. Incorporating independent post-model validation using newly observed fire occurrences is therefore essential to demonstrate model generalization and operational relevance [13].

In this context, we have contributed to advancing the field of wildfire management and susceptibility mapping through three publications. In our first work, presented at the 2022 NTIC International Conference, we provided a comprehensive survey on pre-disaster management based on machine learning, IoT, and big data [14]. This study highlighted how emerging technologies can enhance early warning systems, improve disaster preparedness, and provide rescue teams with more time to respond effectively to wildfire events.

Building on this foundation, our second publication, presented at the 2023 IEEE International Conference on Systems and Control, focused specifically on wildfire susceptibility mapping (WSM) [15]. We reviewed various environmental factors, GIS integration, and machine learning techniques, including statistical models and neural networks, emphasizing practical approaches for precise classification of areas according to fire risk. Finally, we also published a journal paper in Applied Sciences in 2025, which introduces the predictive study and methodology that form the basis of the research presented in this thesis [16]

The present thesis develops a comprehensive, data-driven wildfire susceptibility mapping framework for the province of Jijel, located in eastern Algeria. Jijel represents a particularly relevant case study due to its Mediterranean climate, complex topography, dense forest cover, and strong humanenvironment interactions. The region has experienced recurrent and severe wildfire events, yet remains underrepresented in the wildfire modeling literature, especially compared to northern Mediterranean countries.

The proposed framework integrates environmental, climatic, topographic, hydrological, and anthropogenic data from both global repositories and local agencies. GIS techniques are used to harmonize and analyze these datasets, while multiple machine learning models are trained to assess wildfire susceptibility. To improve data quality, preprocessing steps are applied to address class imbalance, multicollinearity, and label uncertainty. Model performance is evaluated using standard metrics and ROC analysis, followed by spatial mapping of susceptibility classes.

A key contribution of this work lies in its emphasis on interpretability and spatial reasoning. Feature importance is quantified using SHAP analysis, and susceptibility maps are systematically compared with environmental and human factor layers to assess physical coherence. Furthermore, the models robustness is evaluated through an independent validation using wildfire occurrences from 2024 and 2025, providing a rare example of time-forward validation in a southern Mediterranean context. This combination of statistical evaluation, spatial analysis, explainability, and real-world validation strengthens the credibility and operational relevance of the proposed approach.

Ultimately, the goal of this thesis is not only to produce accurate wildfire susceptibility maps, but also to deliver interpretable, GIS-ready outputs that can support decision-making by firefighters, civil protection authorities, and land-use planners. By enhancing understanding of the factors driving wildfire occurrence and identifying priority risk zones, this work aims to contribute to more effective prevention strategies, improved resource allocation, and reduced wildfire impacts on both ecosystems and communities.

The remainder of this thesis is organized as follows. Chapter 1 presents Pre-Wildfire Management, focusing on strategies and approaches for disaster prevention and preparedness. Chapter 2 reviews related work on wildfire susceptibility mapping, data sources, and modeling approaches. Chapter 3 describes the study area and dataset preparation, detailing the environmental and anthropogenic factors considered. Chapter 4 explains the methodology, including data processing, model development, and validation procedures. Chapter 5 presents the results and discussion, analyzing model performance, spatial patterns, and interpretability. Finally, the thesis concludes with a general conclusion summarizing key findings, limitations, and directions for future research."

Part I

Literature Review and Theoretical Background

Chapter 1

Pre-Wildfire Management

1.1 Introduction

Wildfires have become a major challenge for environmental protection and public safety, particularly in regions where climate variability, landscape conditions, and human pressure increase the likelihood of destructive events [17]. Recent seasons have shown that wildfire impacts extend beyond the burned area to include threats to human life, economic disruption, and long-term ecological degradation. This reality has motivated a stronger focus on approaches that reduce risk before disasters unfold, rather than relying mainly on actions taken after losses have already occurred [18].

Pre-wildfire management represents a strategic domain where decision-makers can actively shape outcomes by improving anticipation, preparedness, and coordination [19]. It involves building the capacity to understand risk patterns, identify critical conditions, and support timely action through structured planning and reliable information flows. In practice, this requires combining scientific knowledge with operational planning so that institutions and communities can move from uncertainty to informed readiness [20].

At the same time, the increasing availability of data and digital tools has opened new possibilities for strengthening pre-wildfire action. Sensing technologies, geospatial systems, and intelligent analysis methods can support more accurate risk assessment, more reliable monitoring, and more effective communication to relevant actors [21]. However, achieving real benefit depends on how well these elements are integrated into coherent and usable systems that match the needs of agencies and exposed communities.

This chapter establishes the conceptual and operational basis for proactive wildfire risk reduction. It clarifies the main pre-wildfire measures commonly used to anticipate and limit wildfire impacts, and it highlights how modern technological tools can support these measures through improved observation, analysis, and communication. By organizing the key concepts and system components, the chapter provides a clear foundation for linking pre-wildfire action to spatial decision-support approaches.

1.2 Wildfire Management

Wildfire management has become an essential priority as wildfire activity increases in frequency, severity, and societal consequences. Beyond emergency response, it represents a coordinated effort that combines prevention-oriented measures, planning, and preparedness with organized intervention and recovery mechanisms, supported by lessons learned from past wildfire events [22]. Establishing clear concepts and a structured management framework is therefore necessary to understand why wildfires produce growing impacts and how institutions can reduce risk, limit losses, and strengthen resilience in exposed landscapes and communities.

1.2.1 Wildfires in Recent Years

Over recent years, wildfire activity has increasingly been described as a new operational reality rather than an exceptional event. At the global scale, wildfire affects vast areas each year, with an annual burned area exceeding 400 million hectares, as illustrated in Figure 1.1, which presents the cumulative area burned by week [23]. Beyond the extent of burned land, wildfires represent a growing threat to ecosystems, they reportedly damage more than 3% of the worlds forests annually [24], which reinforces their status as a major concern for both environmental stability and human safety.

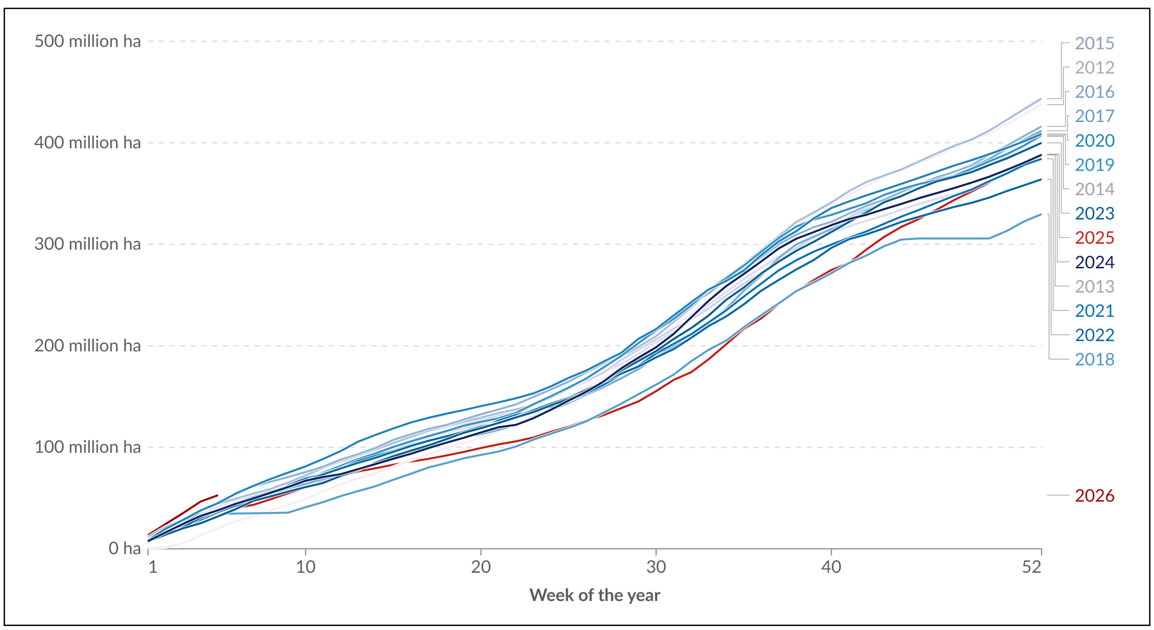

Figure 1.1: Global burned area accumulated by week.

This escalation is also reflected in the recurrence of high-impact wildfire seasons affecting large regions and generating long-lasting consequences. In several countries, recent wildfire episodes have caused severe disruption to livelihoods and economic activity. For example, the Australian bushfires (2020) affected more than 26% of the Australian economy [25], Figure 1.2. Similarly, Algeria has experienced multiple destructive wildfire seasons, notably in 2019 and 2021, with the 2021 wildfires associated with more than 90 fatalities [26], Figure 1.3. Such extreme events can burn extensive forest areas and generate social, environmental, and economic losses that may persist for years after the wildfire is contained.

Figure 1.2: Australian Bushfires (2020)

Figure 1.3: Algerian Wildfires (2021)

The impacts of wildfires can be described through direct effects, which occur during or immediately after the event. From a human perspective, wildfire disasters cause mortality, injuries, and disruptions to communities. According to the Georeferenced Emergency Events Database (EM-DAT), wildfires account for 2,805 direct fatalities, 8,540 injuries, and around 7 million people impacted [26]. Economically, wildfires directly damage houses, infrastructure, and productive areas, disrupting local activity and destroying property. Ecologically, direct impacts include the loss of vegetation cover, habitat destruction, and immediate degradation of air quality and environmental services.

Wildfires also generate indirect impacts, which are often less visible in the short term but can be more persistent and costly. For affected populations, post-disaster living conditionssuch as displacement, overcrowded shelters, or limited access to servicescan aggravate health risks and increase disease exposure. Economic consequences can extend well beyond the event itself through reduced tourism revenues, lower productivity in forest-related industries, depreciation of real estate values, increased health expenditures, insurance burdens, and the high costs of ecosystem and residential-area rehabilitation. In many cases, these longer-term and indirect costs can exceed the direct losses and tend to accumulate over time as recovery stretches across multiple seasons. Ecologically, indirect effects may include altered ecosystem trajectories, impacts on fauna and flora regeneration, and the intensification of other environmental hazards such as drought-related stress.

Several drivers are commonly cited to explain why wildfire impacts appear to be increasing. Climatic pressures contribute through longer dry seasons, more frequent or intense droughts, and a broader increase in aridity that dries vegetation and extends the wildfire season. At the same time, human activity amplifies risk by providing ignition sources, modifying land use and natural vegetation patterns, expanding exposure in wildfire-prone landscapes, and using wildfire as a land-management tool in agriculture and natural ecosystems. Together, these drivers help explain why recent wildfire disasters have attracted worldwide attention and why many countries are increasingly investing in stronger prevention, preparedness, and fire-fighting capacity [24].

1.2.2 Wildfire Management to Reduce Impacts

Wildfire management can be defined as a coordinated set of prevention and suppression activities undertaken by multiple agencies to reduce the likelihood of wildfire occurrence and to limit wildfire impacts when events happen [27]. It is often implemented through structured actions that may be sequential or distinct depending on institutional roles and operational needs. In addition, wildfire management relies on historical records of wildfire events and past practices to identify recurring patterns, evaluate what worked and what failed, and build stronger strategies that are adapted to local conditions and evolving risk.[28]

Wildfire management includes diverse operational and planning measures that support decision-making before and during wildfire events. These measures may include vegetation and fuel assessment, monitoring of fuel conditions, wildfire growth modeling or simulation to anticipate spread, evacuation planning and readiness, and prevention strategies aimed at reducing ignition probability and limiting potential damage. It also involves preparedness actions such as defining protection priorities, organizing resources, training and coordination protocols, and implementing measures that reduce vulnerability in exposed areas [28].

Historically, wildfire management in many regions focused primarily on suppression and emergency operations, with strong emphasis on response capacity and evacuation during wildfire crises. Over time, the increasing frequency and severity of wildfire events, together with improved operational knowledge and planning capacity, has encouraged a broader orientation toward prevention and mitigation. This shift reflects the recognition that suppression alone is insufficient under escalating risk conditions, and that reducing fuels, limiting ignitions, and improving preparedness can substantially reduce overall impacts [27].

The importance of wildfire management lies in its ability to reduce wildfire-related losses across multiple sectors. Effective wildfire management contributes to protecting human life and health, safeguarding economic activity and property, and limiting ecological degradation and long-term environmental damage. Beyond the immediate response, sustained wildfire management strengthens resilience by improving institutions, preparedness, and public safety, which aligns with long-term changes observed in wildfire-related mortality across income groups, where decreasing death rates in recent decades can be associated with improved management strategies and coordinated interventions [29].

1.2.3 Wildfire Management Cycle

Wildfire management cycle refers to a continuous, cyclical process used by governments, organizations, and communities to plan for and reduce the impacts of wildfires. It emphasizes proactive planning before, during, and after a wildfire occurrence, with the objective of minimizing losses, supporting affected populations, and restoring communities to normal or improved conditions[30]. Because Wildfire management is cyclical, it also encourages learning from historical wildfire events, integrating lessons learned into future planning, and strengthening societal resilience over time [28].

Similar to other disaster management frameworks, multiple wildfire management cycle models exist, and the differences between them mainly reflect how institutions choose to group actions and how they intend to use the framework. A commonly used approach is the five-phase model consisting of prevention, mitigation, preparedness, response, and recovery. A simpler four-phase model combines prevention and mitigation into a single phase. Another widely used practical framing is the three-phase model, where measures are classified by when they are applied: pre-disaster management (actions taken before the event), during-disaster management (immediate actions during the event), and post-disaster management (actions after the event) [30]. The conceptual differences across these models are summarized in the corresponding Figure 1.4.

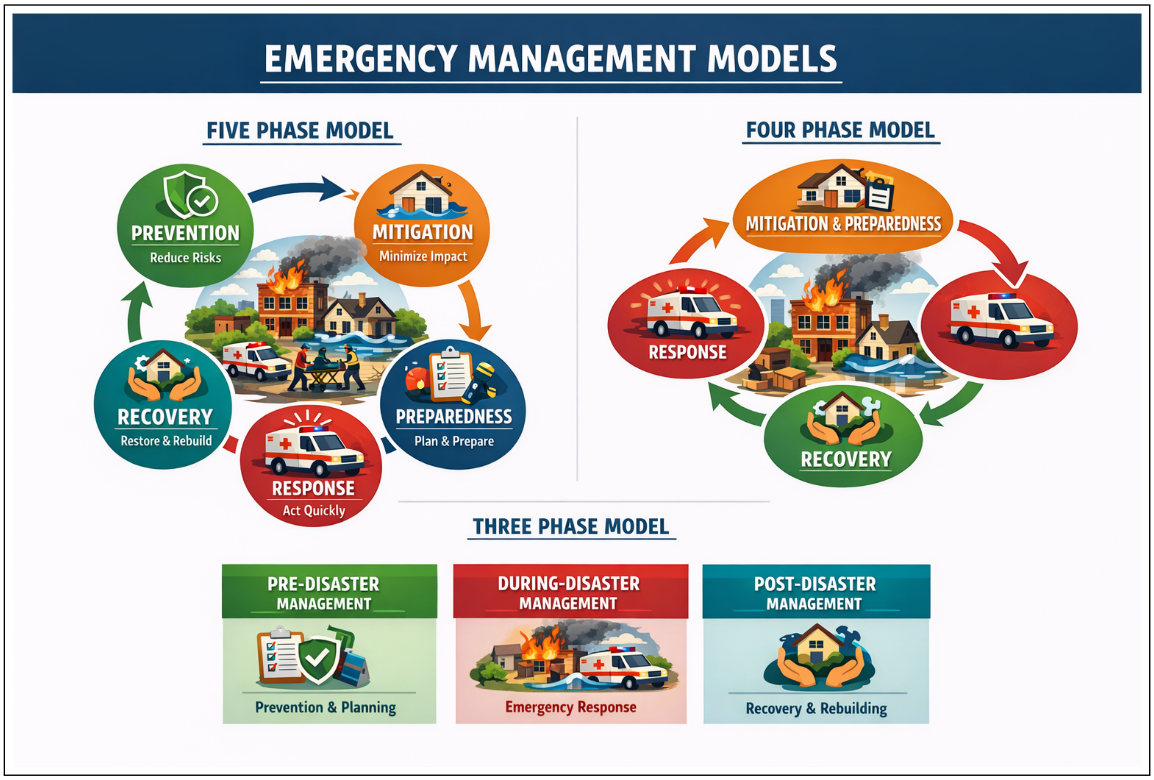

Figure 1.4: Comparison of Wildfire Management Cycle Models (3-, 4-, and 5-Phase model)

In general, more detailed cycle models are often preferred for planning and research because they help structure measures clearly and improve understanding of how actions connect across time. In contrast, simpler models may be more practical during operational periods because they support fast coordination and decision-making [29]. The choice of model can also depend on the hazard context, since some hazards require greater emphasis on particular periods or types of intervention [30]. For clarity in this part, the five-phase model is used to explain the main phases of the wildfire management cycle:

Prevention: focuses on actions taken to avoid wildfire incidents from occurring in the first place. This includes reducing human caused ignitions, controlling ignition sources, enforcing regulations, improving surveillance, and strengthening public awareness and education to reduce risky behaviors and improve compliance with safety measures [30].

Mitigation: refers to measures that reduce the impacts of wildfires that may still occur. These actions are typically long-term or structural, such as creating defensible space around homes by removing flammable materials, implementing fuel management strategies in exposed landscapes, and applying land-use measures that reduce vulnerability [28].

Preparedness: includes planning, organization, training, and equipping to ensure an effective response when a wildfire occurs. It involves developing evacuation plans, ensuring the readiness of wildfire-response equipment and personnel, establishing coordination procedures, and installing early warning and communication systems that support rapid decision-making [28].

Response: consists of immediate actions during and shortly after a wildfire to protect lives and reduce losses. Key response actions include activating emergency operations centers, evacuating exposed populations, conducting wildfire suppression operations, and opening shelters and support services for affected communities [29].

Recovery: includes actions that restore the affected community to its pre-wildfire condition or to an improved state. Typical recovery activities include debris clearance, financial and social assistance to victims, repair and reconstruction of infrastructure, and longer-term rehabilitation planning to reduce future vulnerability [28].

Although this chapter introduces the general wildfire management cycle, the remainder of the chapter and the rest of the study focus mainly on measures implemented before wildfire occurrence. For that reason, the three-phase model is adopted as the operational framing for the subsequent sections, with emphasis placed on pre-wildfire management.

1.3 Pre-Wildfire Management

Despite continuous advances in technologies, strategies, and operational capacity for wildfire suppression, it remains difficultand in some situations impossibleto fully control wildfires once they start. This challenge becomes more critical under conditions that favor ignition and rapid spread, such as heatwaves, very dry fuels, and strong winds, where wildfire behavior can exceed response capacity and quickly escalate into large, destructive events [31].

For this reason, increasing attention has shifted toward actions implemented before wildfire occurrence, with the primary goal of preventing ignition where possible and reducing potential impacts where prevention is not sufficient. Pre-wildfire management is therefore viewed as a strategic integration of measures designed to limit wildfire occurrence, reduce the potential severity and spread of future wildfires, and ensure readiness for an effective response when an event begins [32].

Pre-wildfire management measures can be broadly grouped into two complementary categories. The first is risk reduction, which aims to minimize the likelihood of wildfire occurrence and decrease the potential severity of impacts on landscapes and assets through actions such as burn bans and restrictions, public awareness and education, infrastructure safeguards, fuel management, and predictive planning. The second category is pre-response preparedness, which ensures that communities and responders can act immediately and effectively when a wildfire starts, including evacuation mapping and planning, early warning procedures, and operational wildfire detection systems [33].

Within this perspective, and considering the growing role of data and analytical tools, this section focuses on three key pre-wildfire measures that support decision-making and rapid action: wildfire prediction, wildfire detection, and early warning systems. These measures rely on collecting and analyzing information about factors influencing wildfire occurrence and behavior, enabling timely communication to relevant agencies and supporting proactive planning, situational awareness, and coordinated response readiness.

1.3.1 Wildfire Prediction

Wildfire prediction is the scientific and computational discipline of forecasting the probability, location, intensity, and potential behavior of wildfires across space and time [34]. It combines data-driven methods (e.g., machine learning and deep learning), remote sensing, geospatial analysis, and process-based ecological or physical modeling to estimate future wildfire conditions and support decision making before, during, and after wildfire events [35].

Wildfire prediction is essential because it shifts wildfire management from reactive suppression to proactive risk reduction. By anticipating hazardous conditions and likely wildfire outcomes, agencies can allocate personnel and equipment more efficiently, prioritize prevention actions (e.g., fuel treatments, prescribed burning, community preparedness), and issue earlier, more targeted warnings that protect lives, infrastructure, ecosystems, and public healthespecially as climate change increases the frequency of extreme wildfire conditions [35].

A major direction in the literature is ignition prediction (also called susceptibility or occurrence prediction), which aims to identify where wildfires are more likely to occur in the future under similar conditions. This type relies primarily on historical wildfire records and long-term environmental and human related factors to learn patterns associated with ignition likelihood and spatial risk [22].

In terms of architecture, ignition prediction typically starts with multi-source historical datasets: wildfire inventories (locations and dates), meteorological summaries, topography, vegetation and fuel proxies, land use/land cover, and human pressure indicators (e.g., roads, population density). After preprocessing (cleaning, spatial alignment, handling class imbalance), models are trained to output a risk classification or probability surface, commonly delivered as susceptibility/risk maps [36], see Figure 1.5.

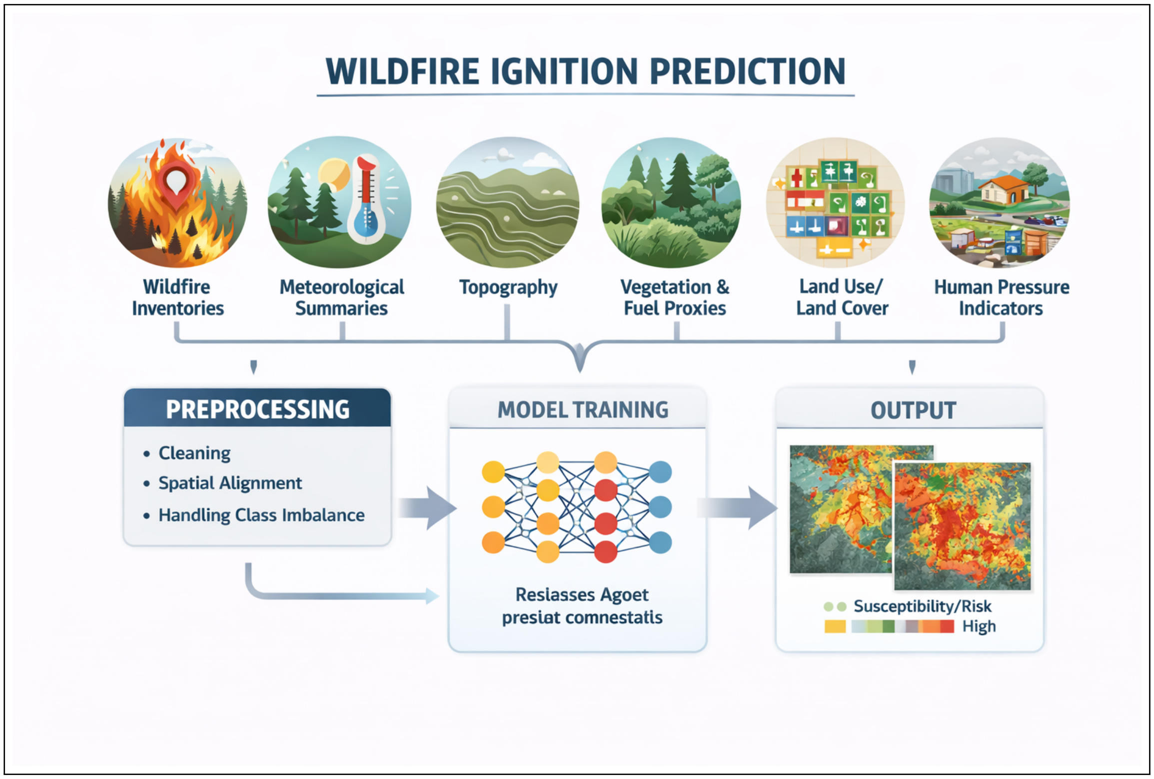

Figure 1.5: Architecture of wildfire ignition (susceptibility) prediction

Outputs from ignition prediction are often used for planning and mitigation, such as prioritizing high-risk zones for prevention, improving patrol routes, optimizing sensor/Unmanned Aerial Vehicle (UAV) deployment, and supporting policy decisions. Importantly, the goal is not to predict the exact ignition time, but to provide actionable spatial intelligence about likely hotspots and the relative contribution of driving factors [22].

Researchers in [36] conducted a comparative study applying multiple machine learning methods to build an ignition prediction model for Portugals Montesinho Natural Park using spatial and temporal predictors. Researchers in [37] represented environmental factorscombining climate data and wildfire recordswithin a graph structure to enable the use of graph convolutional networks alongside Long Short-Term Memory (LSTM) models, aiming to capture spatial dependencies and temporal dynamics simultaneously.

A second direction is wildfire spread prediction, which focuses on forecasting how a wildfire will evolve after ignitionincluding direction, rate of spread, and potentially affected area. Unlike ignition prediction, spread prediction is dynamic and time-sensitive, and it is designed to support operational decisions during active wildfire events [38].

The architecture for spread prediction typically requires (1) an ignition starting point or perimeter, (2) near real time and short term forecast variables (wind speed/direction, temperature, humidity), (3) terrain and fuels information (topography, vegetation state), and (4) historical context to stabilize learning, Figure 1.6. Models ingest recent temporal windows (hours/days) and produce forecasts such as next-step perimeter growth, spread probability maps, or affected-surface estimates over a defined horizon [20].

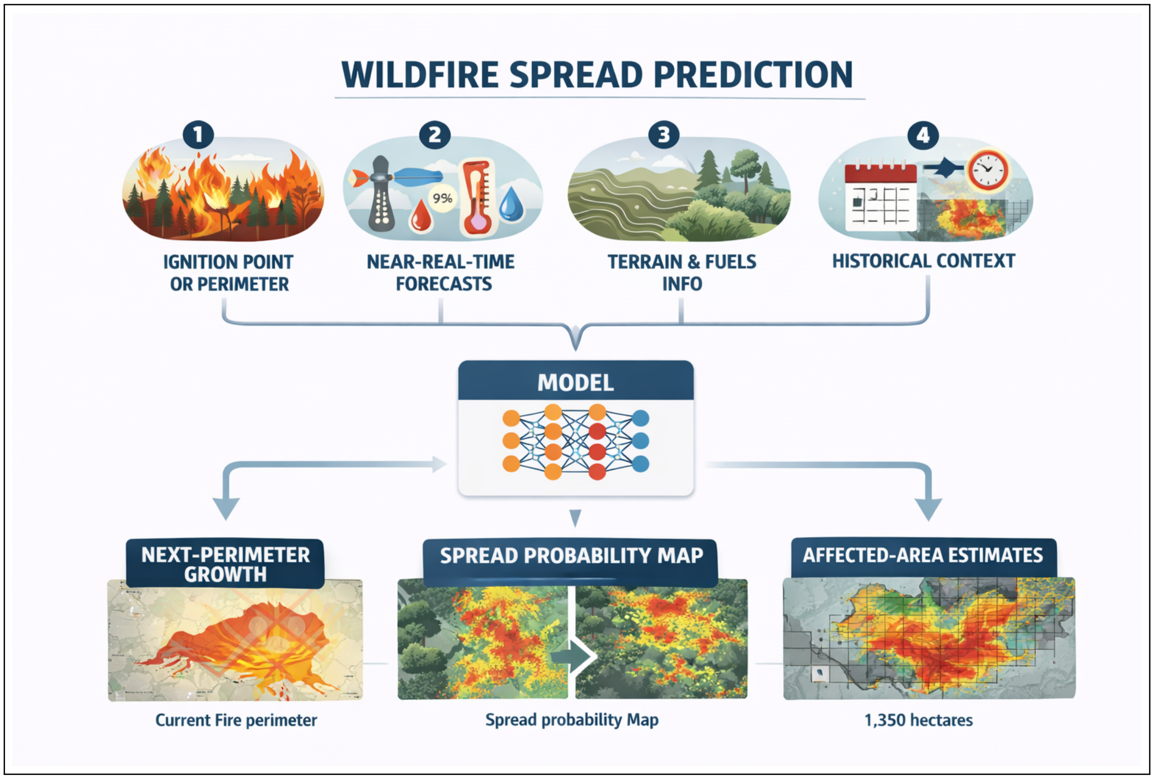

Figure 1.6: Architecture of wildfire spread prediction

Spread prediction is especially valuable for incident response, because it supports decisions like evacuation planning, line placement, resource dispatch, and protection of critical assets. It can be implemented using physics-based simulators, data-driven deep models, or hybrid approaches that combine both, depending on the required speed, scale, and interpretability [39].

Researchers in [20] proposed a deep learning model for large-scale wildfire spread forecasting based on an Multi-Attention Network (MA-Net) architecture, considering prediction horizons from 1 to 5 days and using environmental and climate inputs. Researchers in [40] developed a Convolutional Neural NetworkBidirectional Long Short-Term Memory (CNNBiLSTM) model for wildfire spread prediction in Canada using predictors such as topography, temperature, wind information, and vegetation indices.

Wildfire prediction still faces major challenges due to environmental complexity and data limitations. Key issues include sparse real-time ground observations (e.g., fuel moisture in remote areas), difficulty capturing wildfireatmosphere feedbacks where large wildfires generate their own local weather, uncertainty from ember spotting that triggers new ignitions ahead of the wildfire front, and novel climate conditions that reduce the reliability of models trained on historical regimes [36]. Additionally, there is often a critical trade-off between speed and accuracy, where operational needs demand fast forecasts while high-fidelity modeling can be computationally expensive during time-critical wildfire emergencies [38].

1.4 Wildfire Detection

Wildfire detection represents the operational bridge between a state of readiness and active response. It can be defined as the process of identifying and locating the earliest signs of wildfire ignitionsuch as smoke plumes, thermal anomalies, or gas emissionsin order to trigger rapid intervention while the event is still small and localized [41].

The importance of wildfire detection lies in exploiting the critical time window between ignition and escalation. When detection occurs early, responders can mobilize while the wildfire remains limited in size and intensity, which increases the probability of successful containment and reduces the likelihood that it becomes uncontrollable under adverse conditions. In this way, early detection contributes directly to reducing wildfire impacts and threats on human safety, economic assets, and ecosystems [42].

In general, the architecture of wildfire detection systems begins with real-time data acquisition from one or more sources, such as satellite imagery, fixed cameras, in situ sensors, aerial platforms, or even human-generated signals such as social media reports [21]. The incoming data are then processed using automated methodsoften based on computer vision, signal processing, or real-time classification modelsto identify wildfire signatures (smoke, flames, heat, or gases) and generate alerts with location estimates, as summarized in the corresponding Figure 1.7 .

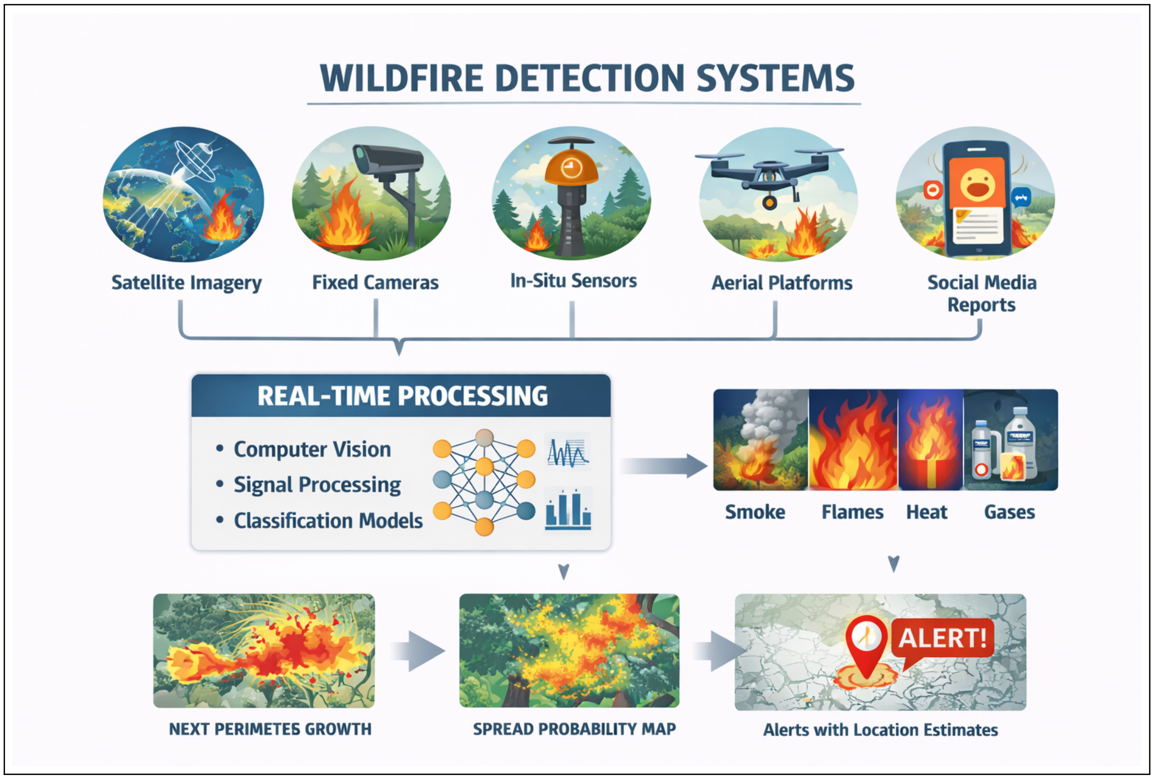

Figure 1.7: General architecture of wildfire detection

A first major category in wildfire detection is satellite-based systems, where spaceborne observations are used to identify potential ignition points through thermal or radiative anomalies. This approach is valuable for monitoring large territories and supporting regional-scale situational awareness. However, it is often constrained by revisit time, spatial resolution trade-offs, and atmospheric limitations such as cloud cover, which can delay or obscure early ignition signals. Researchers in [43] provided an example using a Convolutional Neural Network (CNN)-based approach trained on spaceborne data for wildfire detection.

A second category is ground-based camera networks, which rely on fixed cameras deployed by agencies or authorities, especially in wildfire-prone zones such as forests and agricultural interfaces. These systems analyze video streams to detect smoke or flames using image-processing and deep learning methods, often providing high temporal resolution and fast alerting. In [44] researchers developed a prototype for smoke detection by leveraging networks of cameras originally installed for surveillance and alerting purposes.

A third category is Internet of Things (IoT)-based sensing, where sensors are distributed in the landscape to capture early physical or chemical signals associated with ignition. This may include gas sensors capable of detecting emissions during smoldering phases before smoke becomes clearly visible, as well as other sensing modalities. For instance, [45] proposed a detection approach based on sound-spectrum sensing, using acoustic signatures to support wildfire detection.

A fourth category is aerial surveillance, which uses UAVs (drones) equipped with optical, thermal, or chemical sensors to identify wildfire hotspots and confirm ignition locations. Aerial platforms can provide flexible deployment, detailed close-range observation, and improved detection under some visibility constraints, including heavy smoke, depending on the sensor payload. Researchers in [46] developed a deep ensemble learning approach to classify UAV-collected imagery for wildfire detection.

Despite their promise, wildfire detection systems face important challenges, particularly the risk of false alarms and missed detections. False alarms, or false positives, occur when the system signals a fire where none exists, for example due to industrial smoke, dust, fog, sun glare, or red roofs, leading to unnecessary operational responses [21]. Conversely, false negatives, where an actual wildfire is not detected, can be catastrophic as they delay response and allow fires to spread uncontrolled. In addition, operational performance may be affected by computational latency, and deploying and maintaining sensing infrastructure, communication networks, and real-time processing can be costly, especially for wide geographic coverage [45].

1.5 Wildfire Early Warning Systems

Early warning systems (EWS) for wildfires can be defined as an integrated set of capacities designed to provide timely and actionable information that enables individuals, communities, and institutions to prepare and act appropriately to reduce harm [22]. Unlike wildfire detection, which is primarily a technical task focused on identifying the presence of a wildfire (or the earliest ignition signals), an EWS is a holistic process: it incorporates detection but extends it with risk analysis, forecasting, formal communication procedures, and organized community and institutional response mechanisms [47].

The architecture of a wildfire EWS typically begins with risk knowledge, where multi-source data are collected and analyzed to understand wildfire occurrence, exposure, and vulnerability patterns. This is followed by monitoring and forecasting, using observational systems and analytical models to track current conditions and anticipate potential wildfire development. The next component is warning dissemination, where alerts are communicated through predefined channels to responsible agencies, responders, and at-risk populations [22]. Finally, preparedness and response capacities ensure that warnings translate into effective action, meaning that institutions and communities are trained, coordinated, and equipped to respond based on updated information and clear procedures [47], as illustrated in the associated Figure 1.8.

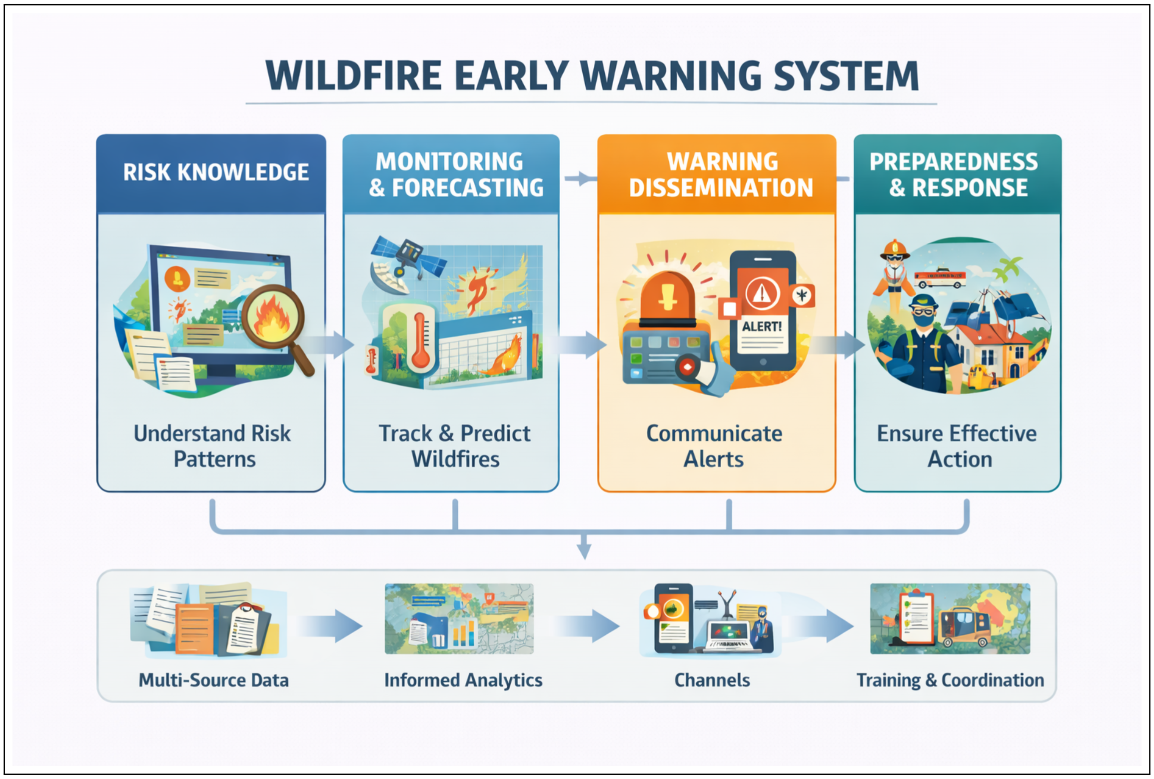

Figure 1.8: Architecture of a Wildfire Early Warning System

The importance of wildfire early warning systems lies in their capacity to transform information into protection. By combining understanding, prediction, detection, and communication in a single coordinated system, EWS can reduce fatalities, injuries, and economic losses by enabling earlier and more targeted preventive actions and evacuations. Well-designed EWS also contribute to resilience by strengthening coordination across agencies, improving public awareness and trust, and ensuring that communities can respond effectively under rapidly evolving wildfire conditions [22].

Examples in the literature illustrate different implementation approaches. Researchers in [48] developed a forest wildfire early warning system that integrated data collection using UAVs and installed sensors, used these inputs to support wildfire prediction and detection, and then transmitted alerts autonomously through applications installed on citizens devices. In another example, [19] proposed a system based on a network of carbon dioxide (CO) sensors used for early wildfire detection, where sensor data were aggregated and analyzed through cloud technologies to support warning generation and information sharing.

1.6 Technologies Supporting Pre-Wildfire Management

Pre-wildfire management requires highly effective systems capable of operating under complex environmental conditions and strict time constraints. To meet these demands, agencies increasingly rely on a combination of available technologies to improve situational awareness, strengthen risk analysis, and support timely decision-making. Tools such as IoT sensing networks, UAV-based surveillance, remote sensing, GIS, cloud computing, artificial intelligence methods, and operational platforms (including web and mobile applications) contribute to building high-performance capabilities for understanding wildfire drivers, predicting risk, detecting early signals, and communicating actionable information to responders and communities.

1.6.1 IoT for Pre-Wildfire Management

The Internet of Things (IoT) can be viewed as a high-resolution nervous system for wildfire management, enabling granular, real-time monitoring that is difficult to achieve through satellite observations or manual patrols alone. By deploying dense networks of solar-powered or energy-harvesting sensor nodes across large and remote landscapes, agencies can continuously track micro-climatic conditions and local risk thresholds such as temperature, humidity, and wind speed [49]. These networks are particularly valuable for early wildfire detection, since they can identify abnormal signals and provide precise location information (including GPS coordinates) at very early stages, sometimes before smoke becomes visible to observers or detectable by conventional optical systems [50].

IoT also plays a central role in data transmission for wide-area monitoring by using low-power and long-range communication solutions such as Long Range (LoRa)-based networks, LoRaWAN, and, when necessary, direct-to-satellite connectivity. These communication capabilities make it possible to deliver alerts from areas without cellular coverage, improving detection reliability in deep wilderness regions. In addition, modern IoT nodes increasingly incorporate edge computing, allowing preliminary processing to be performed locally so that only high-probability threats trigger communication and alerts, reducing bandwidth requirements and improving system responsiveness [51].

Beyond detection, IoT supports wildfire prevention through continuous monitoring of vulnerable infrastructure. For example, sensors installed along utility networks can detect mechanical anomalies in real time (such as line faults or abnormal movement), enabling rapid intervention and, when applicable, automated shut-offs that reduce the likelihood of ignition sources reaching dry fuels [52]. In this way, IoT contributes not only to early warning but also to reducing human-caused ignition risk through infrastructure-level safeguards, see Figure 1.9.

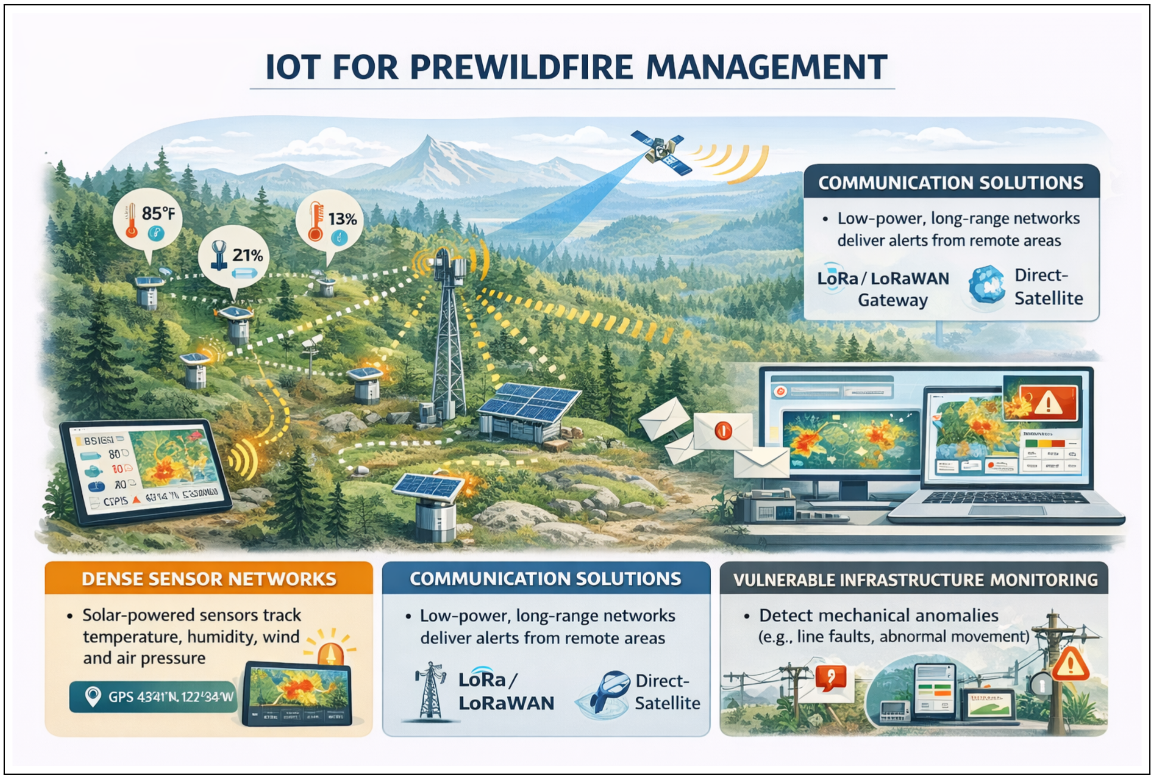

Figure 1.9: Role of IoT in Pre-Wildfire Management: Real-Time Monitoring, Early Detection, and Alert Transmission

A scalable approach combining IoT connectivity and machine learning has been proposed in [53], where a forest wildfire detection solution based on LoRa technology and wireless sensor networks was integrated with a feed-forward neural network. In that system, data analysis and alert management were supported through a Social Internet of Things (SIoT) paradigm, and the overall configuration was validated in a real forest scenario, demonstrating its applicability for early detection.

Despite these advantages, deploying IoT for wildfire management faces substantial economic and logistical challenges. Achieving reliable coverage across rugged forest landscapes may require thousands of sensor units, in addition to gateways and, in some cases, satellite backhaul subscriptions, creating high initial investment costs and ongoing operational expenses for many local authorities [54]. Maintenance is also demanding, because sensors can be damaged by extreme weather, wildlife activity, or even the wildfires they are designed to detect, requiring hazardous field interventions in remote areas. Although energy-harvesting solutions are improving, many systems still depend on batteries that must be replaced periodically. Finally, rapid hardware obsolescence and evolving communication standards can complicate long-term interoperability, making sustained operation and upgrades a persistent technical and financial burden [51].

1.6.2 UAVs for Pre-Wildfire Management

Unmanned Aerial Vehicles (UAVs) provide a flexible aerial intelligence layer for pre-wildfire management, bridging the scale gap between ground sensing and satellite observation through high-resolution, on-demand data collection. Their main advantage is the ability to operate over complex terrain and hazardous environmentssuch as steep valleys, remote mountain areas, and dense forest canopieswhere ground access is difficult and manned operations can be costly or constrained [55]. Equipped with payloads such as optical cameras, thermal infrared sensors, Light Detection and Ranging (LiDAR), or hyperspectral sensors, UAVs can be used proactively to generate detailed maps of fuel structure and vegetation condition, supporting the identification of zones where fuels are dry and continuous enough to favor ignition and rapid spread. This type of granular information helps agencies plan targeted mitigation actions, including prioritizing fuel treatments, defining strategic firebreak locations, and improving monitoring coverage over high-risk landscapes [18].

UAVs also support the operational transition from detection to response through rapid verification and real-time surveillance. When an IoT network or satellite observation indicates a potential hotspot, UAVs can be dispatched to the reported coordinates to provide visual confirmation, reducing false alarms caused by non-wildfire sources such as dust or industrial emissions. In early wildfire stages, thermal sensors can help locate active hotspots and assess initial spread patterns under low visibility conditions (e.g., smoke or nighttime), delivering live situational awareness that supports faster and more precise decisions, including resource staging and targeted evacuation planning [56], see Figure 1.10. Recent research has also adapted deep learning for UAV-based wildfire detection, including an enhanced YOLOv8 approach designed for UAV imagery that improves localization accuracy through Wise-IoU (WIoU) v3 loss, reduces model complexity via Ghost Shuffle Convolution, and strengthens smoke-feature extraction in complex wooded scenes using BiFormer attention [57].

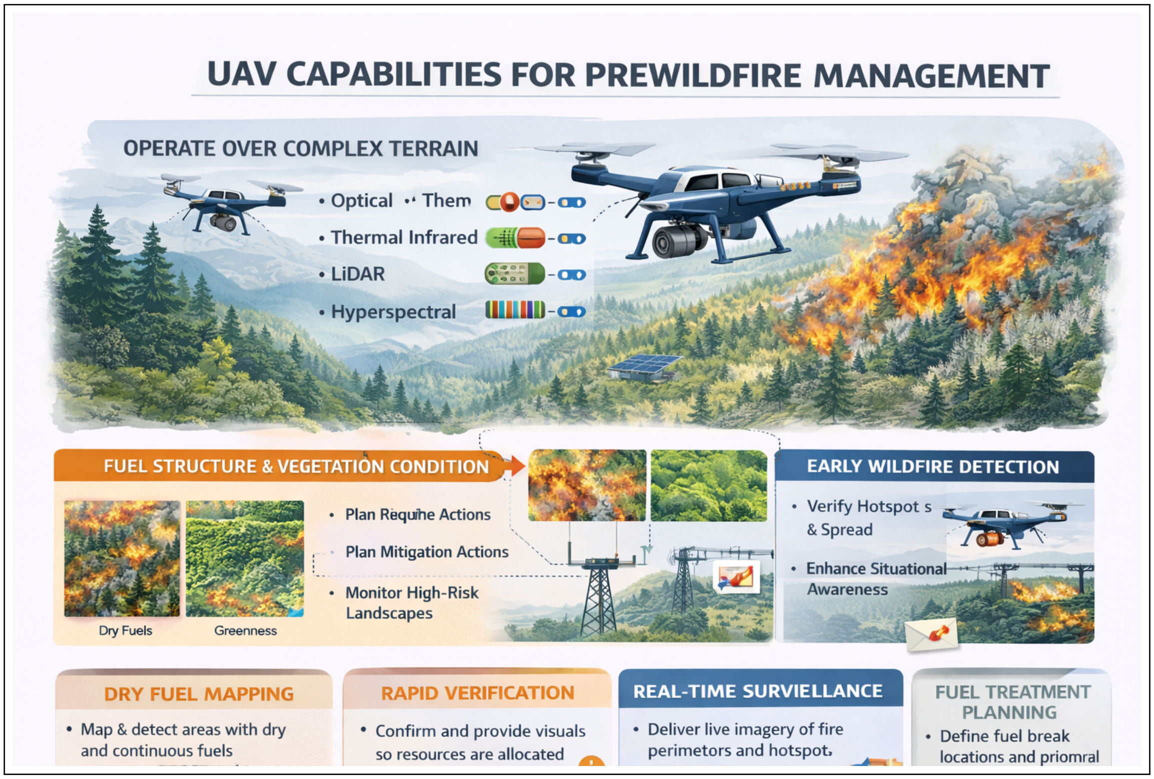

Figure 1.10: Role of UAVs in Pre-Wildfire Management: Fuel Mapping, Hotspot Verification, and Real-Time Surveillance

At the same time, UAV deployment faces technical, operational, and regulatory challenges that can limit sustained coverage and reliable real-time use. Battery endurance is often restricted, especially for multi-sensor payloads, requiring frequent recharging or coordinated rotations for persistent monitoring. Performance can degrade under harsh conditions such as strong winds, turbulence, and high temperatures, while reliable data transmission remains difficult in mountainous regions or remote forests with limited connectivity, and smoke can reduce visibility unless advanced (and costly) thermal or LiDAR sensors are used. Regulatory constraintsparticularly restrictions on beyond-visual-line-of-sight operationscan reduce deployment flexibility, and safety concerns arise when unauthorized drone activity interferes with official aerial operations, creating operational risk and occasionally forcing mission interruptions [55].

1.6.3 Remote Sensing and GIS for Pre-Wildfire Management

Remote sensing and Geographic Information Systems (GIS) form a foundational intelligence layer for pre-wildfire management by providing wide-area spatial observation and the analytical framework needed to understand landscape-scale risk. Remote sensing acts as a primary data source by using satellite missions (e.g., Sentinel and Landsat families) and additional sensing technologies such as

LiDAR and Synthetic Aperture Radar (SAR) to monitor environmental conditions without physical contact [58]. These observations support the derivation of key indices and variables used in wildfire analysis, including vegetation-condition indicators such as Normalized Difference Vegetation Index (NDVI) and severity-related metrics such as Normalized Burn Ratio (NBR).. In addition, SAR observations are valuable because they can maintain coverage under cloud conditions and some visibility constraints, while thermal sensors can support hotspot identification and help flag potential ignitions in remote areas [59], see Figure 1.11.

GIS complements remote sensing by integrating multi-source information into a unified decision-support environment. By combining remote sensing products with terrain variables (elevation, slope, aspect), historical wildfire occurrence, weather information, and indicators of human presence (e.g., proximity to roads and settlements), GIS enables the production of wildfire susceptibility maps and other risk layers that are directly usable for planning [59]. In the pre-wildfire phase, these outputs help prioritize areas for prevention and mitigation actions such as fuel treatment planning, targeted monitoring, and risk-informed land-management decisions. More broadly, the integration of remote sensing and GIS supports a shift toward data-driven risk reduction by improving spatial understanding of wildfire drivers and enabling systematic prioritization across large territories [**ref45**], see Figure 1.11.

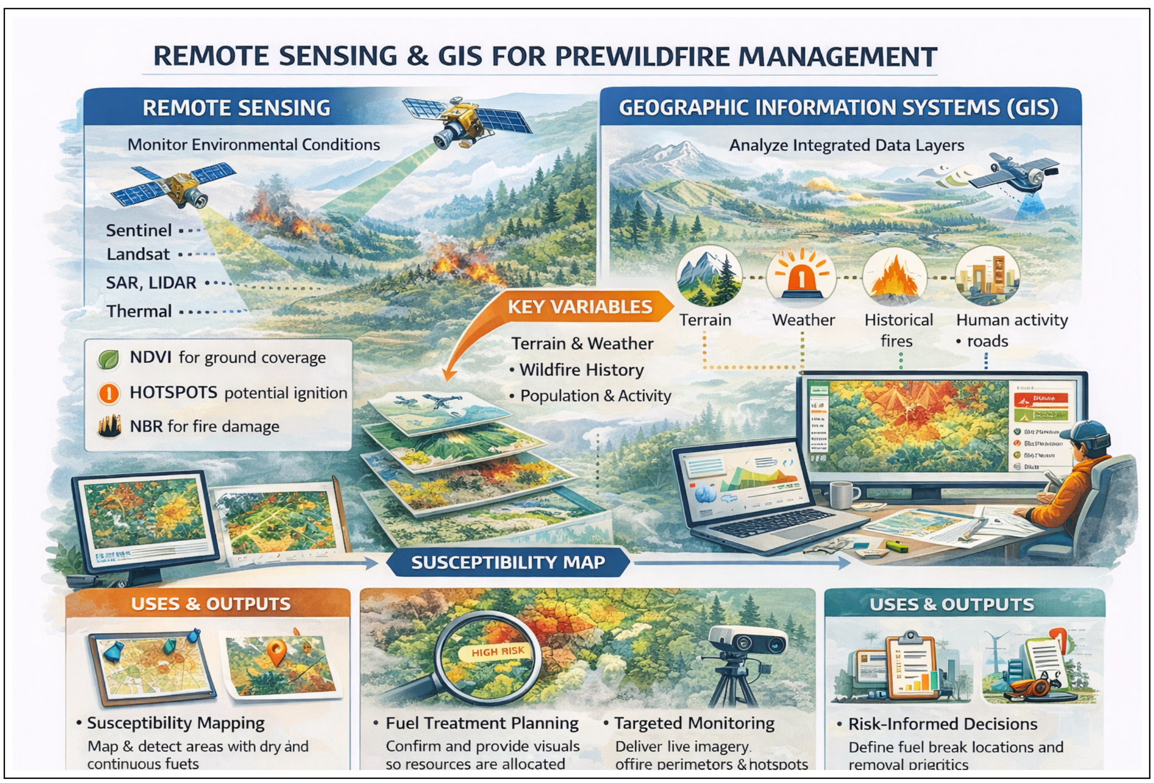

Figure 1.11: Role of Remote Sensing and GIS in Pre-Wildfire Management: Landscape Monitoring, Risk Mapping, and Decision Support

A representative example of this integration is reported in [60], where remote sensing and spatial statistical methods were used to analyze wildfire occurrence patterns and their relationships with multiple environmental and anthropogenic factors. Spatial distribution was explored using techniques such as kernel density estimation and spatial autocorrelation/hotspot analyses (e.g., Morans I and Getis-Ord Gi), while regression-based models (including geographically weighted approaches) were applied to quantify how variables such as elevation, slope, relative humidity, temperature, and land-use/deforestation indicators relate to wildfire frequency across heterogeneous regions. Such studies illustrate how remote sensing products and GIS-based spatial analysis can jointly support robust understanding of where and why wildfire activity concentrates.

Despite their importance, remote sensing and GIS approaches face several practical challenges. Processing high-resolution SAR or LiDAR data over large areas can require substantial storage, computational power, and advanced preprocessing workflows, which may exceed the capacity of many local agencies. Access to very high-resolution imagery can also be limited by cost, creating uneven data availability across regions. Finally, temporal constraints remain significant: high-resolution satellite observations may have revisit intervals of days, while sensors with frequent updates often provide coarser resolution, which can limit the detection of small or early-stage ignitions and reduce responsiveness in fast-evolving conditions [**ref45**].

1.6.4 AI Methods for Pre-Wildfire Management

Artificial intelligence (AI) can be considered the cognitive core of pre-wildfire management because it enables rapid, autonomous analysis of large and heterogeneous datasets that would be difficult to process manually. In ignition prediction (susceptibility/occurrence prediction), machine learning (ML) methods such as Random Forest, Support Vector Machines (SVM), and gradient boosting models (e.g., XGBoost) are widely used to learn non-linear relationships from historical archives that combine meteorological conditions, topography, vegetation proxies, and human-activity indicators. These models are particularly valuable because, in addition to producing susceptibility maps, they can support factor interpretation through feature importance analysis, helping agencies identify which driverssuch as drought-related indicators or proximity to infrastructurecontribute most strongly to wildfire likelihood in a given region [61], see Figure 1.12.

Deep learning (DL) has further increased automation in wildfire detection and behavior-related forecasting. Convolutional Neural Networks (CNNs) are commonly used in camera networks and satellite-image workflows to detect smoke signatures or thermal anomalies quickly and consistently, including under conditions where human observation may be constrained (e.g., low visibility or continuous monitoring needs) [62]. For short-term wildfire behavior and spread-related prediction, time-series models such as recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks can integrate evolving meteorological inputs to estimate near-future wildfire dynamics over operational horizons. More recent directions also include generative approaches and physics-informed learning, which can support scenario generation, training, and faster approximations of coupled wildfireatmosphere effects within simulation or decision-support contexts [63], see Figure 1.12.

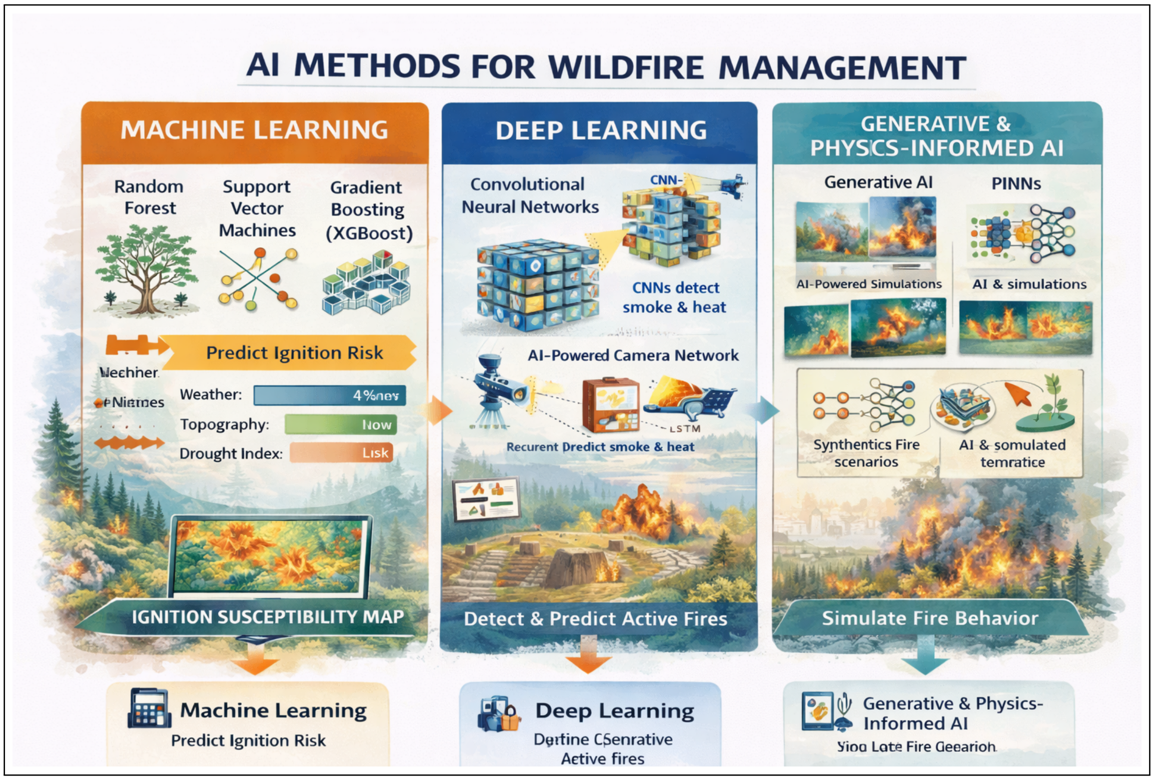

Figure 1.12: Role of Artificial Intelligence in Pre-Wildfire Management: Susceptibility Modeling, Automated Detection, and Behavior Forecasting

Various ML approaches, including Support Vector Regression (SVR), Gaussian Process Regression (GPR), regression trees, and neural networks, have also been evaluated for predicting wildfire rate of spread in grassland contexts using datasets that combine wildfire and experimental records, with results showing strong fits on model development data [64].

However, important limitations remain for AI-enabled wildfire management. A key concern is interpretability: while DL models can achieve high predictive performance, their black-box nature may reduce trust and hinder adoption in high-stakes operational decisions where decision-makers need understandable justifications. AI performance also depends strongly on training data quality and representativeness; models trained mainly on historical conditions may generalize poorly under novel climate extremes, increasing the risk of failure when wildfire behavior departs from past patterns. In addition, the computational demands of DLespecially for large-scale, near-real-time processingoften require cloud or high-performance infrastructure, which can become a vulnerability when connectivity or power is disrupted during wildfire emergencies [65].

1.6.5 Operational Platforms for Pre-Wildfire Management

Operational platforms for wildfire management have increasingly evolved into common operational pictures that unify fragmented information into intuitive web and mobile interfaces for both professionals and the public. At the agency level, these platforms function as coordination and decision-support environments by integrating heterogeneous streams such as satellite-based detections, GIS layers, weather information, analytical outputs, and GPS-based tracking of operational resources. By providing a synchronized view across institutions, they reduce communication gaps and support faster and more consistent decisions during rapidly evolving wildfire situations [66], see Figure 1.13.

For communities, high-accessibility mobile applications play a complementary role by bridging the last mile of emergency communication. These tools translate operational information into actionable guidance by providing perimeter visualization, evacuation updates, and practical safety instructions, while also supporting public health awareness through smoke and air-quality information. Through connected workflows, the same ecosystem can support a continuous information chainfrom the moment an ignition signal is detected to the moment alerts and guidance reach exposed populationsthereby improving timeliness and coordination [67], see Figure 1.13.

Figure 1.13: Operational Web and Mobile Platforms for Pre-Wildfire Management

A relevant approach is presented in [68], where a decision support system adapted to the Portuguese context combines multi-sensor technologies with GIS functionalities through three components: an ArcGIS Online feature service for near-real-time collaborative operational data sharing, a mobile client application providing GIS-based visualization dashboards, and a multi-sensor device collecting field data to enrich the information flow for wildfire management.

Even the most advanced operational platforms face significant hurdles that can limit effectiveness during wildfire crises. A central challenge is interoperability, since different agencies may rely on incompatible software and data standards, producing fragmented situational awareness and conflicting versions of key layers such as wildfire perimeters. The digital divide also remains critical: remote or low-income populations may lack connectivity or digital literacy, reducing access to timely alerts. In addition, operational trust can be affected when AI-driven outputs are difficult to interpret, while latency and uncertainty in data streams can reduce usefulness in fast-moving events where minutes matter. Finally, when platforms incorporate crowdsourced information, maintaining data quality becomes difficult, and unverified reports can introduce confusion for both citizens and responders [69].

1.7 Conclusion

Pre-wildfire management provides a practical pathway to reducing wildfire losses by strengthening anticipation, preparedness, and coordination before critical conditions escalate. When combined with reliable information, efficient monitoring, and effective communication, it supports better prioritization of prevention actions and improves readiness for rapid intervention, while acknowledging that performance is shaped by data limitations, system costs, and operational constraints.

Building on this foundation, the next chapter reviews wildfire susceptibility mapping (WSM) literature as a key risk-assessment component that translates multi-source factors into actionable spatial evidence for prevention planning and decision support.

Chapter 2

Wildfire Susceptibility Mapping

2.1 Introduction

Over the past two decades, the literature on wildfire susceptibility mapping has expanded significantly, reflecting advances in Earth observation technologies, geographic information systems, and data-driven modeling techniques. Researchers have explored a wide range of conditioning factors, methodological approaches, and analytical tools, resulting in diverse workflows and modeling philosophies. While this growing body of work has led to notable improvements in predictive accuracy, it has also introduced heterogeneity in data sources, factor selection strategies, modeling methods, and validation practices, making it difficult to identify best practices or transferable solutions.

The objective of this chapter is to synthesize and critically review existing research on wildfire susceptibility mapping in order to establish a structured foundation for the proposed methodology. The chapter examines key concepts and system architectures, commonly used environmental and anthropogenic factors, the role of GIS in data integration and analysis, modeling techniques ranging from traditional methods to advanced machine learning and deep learning approaches, and recent efforts toward interpretability. By highlighting methodological trends, strengths, and limitations, this review identifies persistent research gaps and motivates the design choices adopted in the subsequent chapters of this thesis.

2.2 Wildfire Susceptibility Mapping: Definition and Architecture

Wildfire Susceptibility Mapping (WSM) refers to the spatial assessment of the likelihood that a wildfire may occur in a given location based on its inherent environmental and anthropogenic characteristics [70]. Unlike short-term fire danger or hazard indices that fluctuate daily with meteorological conditions, susceptibility mapping focuses on relatively stable or slowly varying factors such as terrain, vegetation structure, climate trends, and human presence. The outcome is typically a spatially explicit map that classifies the landscape into different susceptibility levels, providing a long-term perspective on where wildfire ignition is more likely to occur [71].

The importance of WSM lies in its role as a proactive decision-support tool for wildfire risk reduction. By identifying areas that are structurally prone to wildfire occurrence, susceptibility maps support long-term land-use planning, fuel management strategies, and prevention policies. They allow authorities to prioritize mitigation measures such as fuel treatments, firebreak construction, and regulation enforcement in the wildlandurban interface. In recent years, WSM products have also gained relevance for emergency preparedness, insurance risk assessment, and public awareness, as they provide a scientifically grounded basis for anticipating wildfire exposure under typical environmental conditions rather than reacting to crises after ignition [72].

From a systems perspective, WSM is commonly implemented through a structured data-to-decision workflow that integrates heterogeneous spatial information into a coherent analytical framework. This workflow generally begins with the acquisition of conditioning factors representing topographic, vegetation, climatic, and human influences, together with historical wildfire records. These inputs are harmonized and processed within a geospatial environment before being analyzed using statistical or machine learning techniques to estimate susceptibility patterns. The final outputs are spatial productsoften continuous probability surfaces or classified mapsthat are designed for integration into geographic information systems and operational planning platforms [73], see Figure 2.1. This general architecture provides the foundation upon which diverse modeling approaches in the literature are built and compared in this chapter.

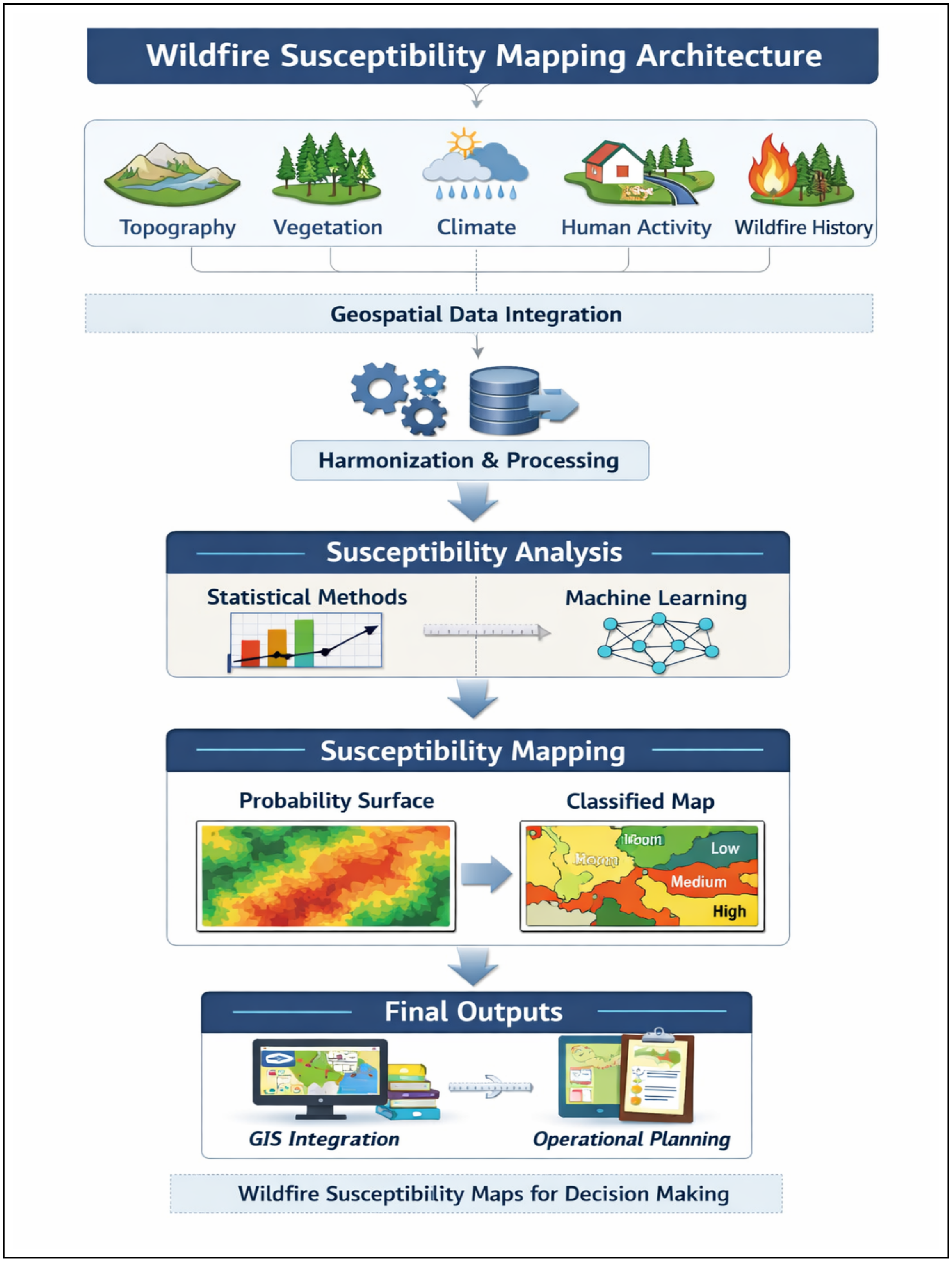

Figure 2.1: General workflow of wildfire susceptibility mapping system.

2.3 Factors Influencing Wildfire Occurrence

The development of a Wildfire Susceptibility Mapping system begins with the identification and collection of environmental and anthropogenic factors that influence wildfire occurrence. Researchers first focus on assembling spatial data describing the physical environment, vegetation characteristics, climatic conditions, and human activities from all available sources, including remote sensing products, meteorological records, field surveys, and geospatial databases. This step is fundamental not only for building susceptibility models but also for understanding how different factors contributeindividually and collectivelyto wildfire ignition and spatial patterns. By analyzing these relationships, WSM studies aim to move beyond simple prediction toward a more explanatory understanding of wildfire dynamics across landscapes[74]

In recent years, advances in Earth observation, data availability, and machine learning techniques have led to the use of an increasingly large and diverse set of conditioning factors in wildfire susceptibility studies. While this richness of data has improved modeling capabilities, it has also introduced new challenges related to redundancy, scale mismatch, and model interpretability. Consequently, selecting appropriate factors has become a critical methodological step, requiring careful consideration of the study areas environmental context, data quality, and research objectives. Establishing clear criteria for factor selection helps ensure that susceptibility models remain robust, transferable, and physically meaningful rather than overly complex or data-driven without environmental justification [74].

2.3.1 Environmental and Anthropogenic Factors

Environmental factors are widely recognized as the fundamental elements that precondition landscapes for wildfire ignition and spread. Among them, climatic and meteorological variables play the most dynamic and influential role, as they directly control fuel moisture and fire behavior. Rising temperatures, prolonged droughts, and declining relative humidity accelerate vegetation desiccation, increasing flammability across large regions. High-risk fire conditions are often associated with the combined effects of heat, dryness, and strong winds, where temperature and humidity primarily govern ignition probability, while wind speed and direction strongly influence fire propagation and intensity. Because of their temporal variability, climatic variables are incorporated into WSM studies using different temporal representations, ranging from long-term averages that reflect structural susceptibility to seasonal or lagged indicators that capture pre-fire drying conditions [75].

Topography provides the physical framework within which fires evolve and is therefore among the most consistently used groups of factors in susceptibility modeling. Elevation, slope, and aspect influence local microclimates, fuel distribution, and wind exposure [76]. Slope is particularly significant, as fires tend to spread more rapidly uphill due to flame tilt and preheating of fuels, while aspect controls solar radiation and moisture retention, creating contrasts between drier sun-exposed slopes and cooler shaded ones. These variables are favored in regional and national-scale studies because they are stable over time, globally available through digital elevation models, and easily interpretable by decision-makers, making them a reliable foundation for comparative WSM analyses [77].

Vegetation-related factors represent the fuel component of the fire triangle and are essential for sustaining wildfire activity. Fuel load, continuity, and condition strongly influence susceptibility, and are commonly represented through remote sensingderived indices such as NDVI, EVI, and land-cover classifications [72]. These proxies capture vegetation density, health, and moisture status, allowing researchers to distinguish between fuel-rich forested areas, shrublands, and fragmented agricultural or urban interfaces. Several studies report non-linear relationships between vegetation indices and fire occurrence, particularly in Mediterranean and semi-arid environments, where intermediate vegetation greenness often corresponds to abundant fine fuels that favor ignition and rapid spread. Species composition and the presence of dead or stressed vegetation further increase susceptibility by facilitating ladder fuels and crown fire development [76].

Hydrological and soil-related factors are incorporated less consistently but can be locally important in shaping fire occurrence patterns. Variables such as distance to rivers, wetlands, and water bodies are often used as proxies for moisture availability and landscape accessibility, with effects that vary depending on regional climate and land use [76]. Soil moisture, soil type, and texture influence water retention and vegetation growth, thereby indirectly affecting fuel conditions. Advances in satellite-based soil moisture products have increased the inclusion of these variables in large-scale studies, while higher-resolution national soil datasets are typically employed in localized case studies [78].

Anthropogenic factors are widely acknowledged as the dominant catalysts of wildfire ignitions, accounting for the majority of recorded fire events worldwide. Proximity to roads, settlements, and infrastructure is consistently identified as a strong predictor of fire occurrence, reflecting increased human presence, accessibility, and accidental or negligent ignition sources [76]. The expansion of the WildlandUrban Interface intensifies this pressure by introducing artificial fuels and increasing interactions between human activities and fire-prone ecosystems. Power lines, recreational areas, agricultural boundaries, and population density are also frequently used proxies, although many studies deliberately limit the number of human-related variables to reduce redundancy and maintain model interpretability [77].

Finally, historical wildfire data play a dual role in susceptibility studies. Past fire occurrences provide essential labels for supervised learning models, linking observed ignitions to their environmental and anthropogenic context [76]. In some cases, smoothed fire density or hotspot layers are also used to represent persistent ignition patterns driven by recurring human activity. However, the literature emphasizes the need for careful temporal separation between training and validation data to avoid circularity and over-optimistic model performance. When used appropriately, historical fire records significantly enhance the robustness and explanatory power of wildfire susceptibility models [79].

2.3.2 Selection of Conditioning Factors

The selection of conditioning factors is a critical step in wildfire susceptibility mapping, as it directly influences model performance, interpretability, and generalization. Although a large number of environmental and anthropogenic variables have been reported in the literature, not all factors contribute equally to wildfire occurrence, and inappropriate selection may introduce redundancy, multicollinearity, or noise into susceptibility models

To understand how researchers select conditioning factors for wildfire susceptibility modeling, 23 peer-reviewed studies were analyzed. The factors employed in each study were identified and compiled, and their frequencies of use were summarized in Figure 2.2. Across the selected studies, a total of 28 distinct conditioning factors were used. The analysis reveals substantial variability in factor selection, reflecting differences in study objectives, geographic context, data availability, and methodological choices. Some factors are widely adopted across the literature, while others appear only sporadically.

From Figure 2.2, it is evident that certain variables are consistently used in the majority of studies. Topographic factors (particularly slope, which appears in all 23 studies) are among the most frequently adopted variables, followed by elevation and aspect. Climatic variables such as temperature and precipitation are also widely used due to their direct influence on fuel moisture and ignition probability. Vegetation-related indicators, especially the Normalized Difference Vegetation Index (NDVI), are commonly employed as proxies for fuel availability and vegetation health. In contrast, other factors, such as drought indices, soil moisture, or specific hydrological variables, are used far less frequently, often appearing only in studies focused on particular climatic or ecological contexts.

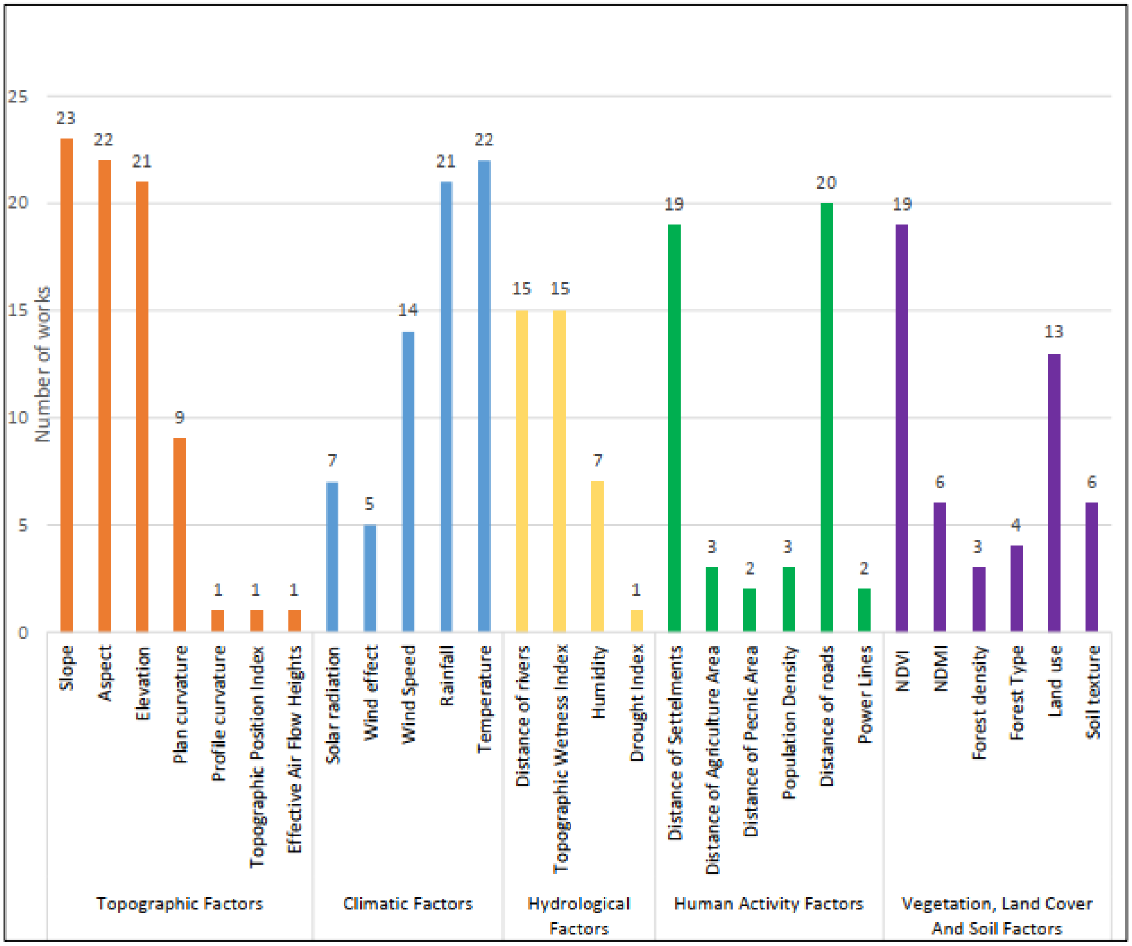

Figure 2.2: Frequency of conditioning factors reported in recent WSM studies by category.

A key observation from the reviewed literature is that factor selection is strongly dependent on the objectives of the study. Some authors aim to assess the combined influence of a wide range of environmental and anthropogenic drivers on wildfire occurrence and therefore include a large number of factors; for example, some studies employ more than 20 variables to capture all potential influences [77]. Conversely, other studies deliberately limit the number of factors (sometimes to only two or three) in order to focus on the role of specific phenomena, such as vegetation condition or soil moisture, and to better interpret their individual effects [78]. Nevertheless, the majority of studies adopt an intermediate strategy, typically selecting between 10 and 14 factors [71, 79]. This range appears to provide a balance between model comprehensiveness and methodological robustness, while reducing redundancy, multicollinearity, and the inclusion of irrelevant variables.

Practical considerations also play a significant role in factor selection. Many of the most frequently used variables are derived from globally available datasets, such as digital elevation models for slope and aspect [79]or satellite products for NDVI [80], which facilitates their use across different regions. When multiple variables represent similar processes (such as different wind) related metrics or overlapping proxies of human activity, researchers often retain only one to avoid redundancy [72].

Human-related factors illustrate this selection trade-off particularly well. While proximity to roads and settlements is almost universally included as a proxy for anthropogenic ignition pressure, some authors extend their analysis by incorporating additional variables such as population density, agricultural areas, recreational zones, or power lines to examine their individual contributions in more detail [81, 82]. When multiple variables represent similar processes, researchers often retain only the most interpretable or best-curated indicators to avoid conceptual and statistical overlap.

Several studies further support factor selection through statistical screening methods. Collinearity diagnostics, such as correlation analysis and Variance Inflation Factor (VIF), are frequently applied to identify and remove redundant predictors. For instance, authors in [83] excluded variables such as water vapor pressure, gross primary productivity (GPP), potential evapotranspiration (PET), and aridity index (AI) due to high VIF values, thereby improving model stability and interpretability.

Despite these practices, explicit justification of factor choice remains uneven across the literature. Only a limited number of studies [84, 85] thoroughly discuss the physical relevance of their selected factors, grounding their decisions in expert knowledge, fire behavior theory, and domain-specific reports. This approach provides deeper insight into the mechanisms linking conditioning factors to wildfire occurrence [73, 86]. In contrast, many studies rely primarily on commonly used variables without a detailed examination of their suitability for the specific study area, often driven by data availability rather than contextual relevance.

Overall, the literature indicates that conditioning factor selection in WSM is governed by a combination of scientific relevance, study objectives, expert judgment, methodological constraints, and data availability. This diversity of approaches highlights the absence of a universal factor set and underscores the need for context specific selection strategies that balance physical relevance, data quality, and model interpretability.

2.4 Role of GIS in Wildfire Susceptibility Mapping

Geographic Information Systems (GIS) constitute a core component of wildfire susceptibility mapping due to their capacity to manage, analyze, and integrate spatial data from heterogeneous sources 2.3. One of the primary functions of GIS in WSM is spatial data extraction. GIS tools are widely used to derive relevant variables from satellite imagery and other geospatial datasets, including topographic attributes, vegetation indices, climatic parameters, and land use or land cover information. These spatial layers are converted into quantitative representations that can be directly incorporated into statistical and machine learning models for susceptibility assessment [87, 88, 70, 89].

Another fundamental contribution of GIS lies in spatial scale harmonization. Data used in wildfire studies are typically collected from multiple sources that differ in spatial resolution, extent, and projection systems. GIS enables the resampling, reprojection, and alignment of these datasets to a common spatial resolution and coordinate framework. This unification process ensures spatial consistency across all layers, which is essential for accurate comparison, integration, and modeling of wildfire susceptibility [90].

GIS also plays a critical role in dataset construction and integration. Multiple spatial layers representing environmental conditions and human-related factors are combined within a GIS environment to build comprehensive analysis-ready datasets. Each conditioning factor is organized as a raster or vector layer and structured into a geographic matrix, allowing pixel-level or unit-based analysis. This systematic integration provides the foundation upon which machine learning and other modeling techniques operate [74, 91].

Finally, GIS is indispensable for visualization and interpretation of results. Its cartographic capabilities enable the effective representation of spatial relationships and patterns through thematic maps. By overlaying and comparing different layerssuch as elevation, vegetation cover, and proximity to human infrastructureresearchers can visually explore the interactions among conditioning factors and better interpret their influence on wildfire occurrence. These visualization tools are also essential for communicating susceptibility outcomes to decision-makers and stakeholders [92, 89].

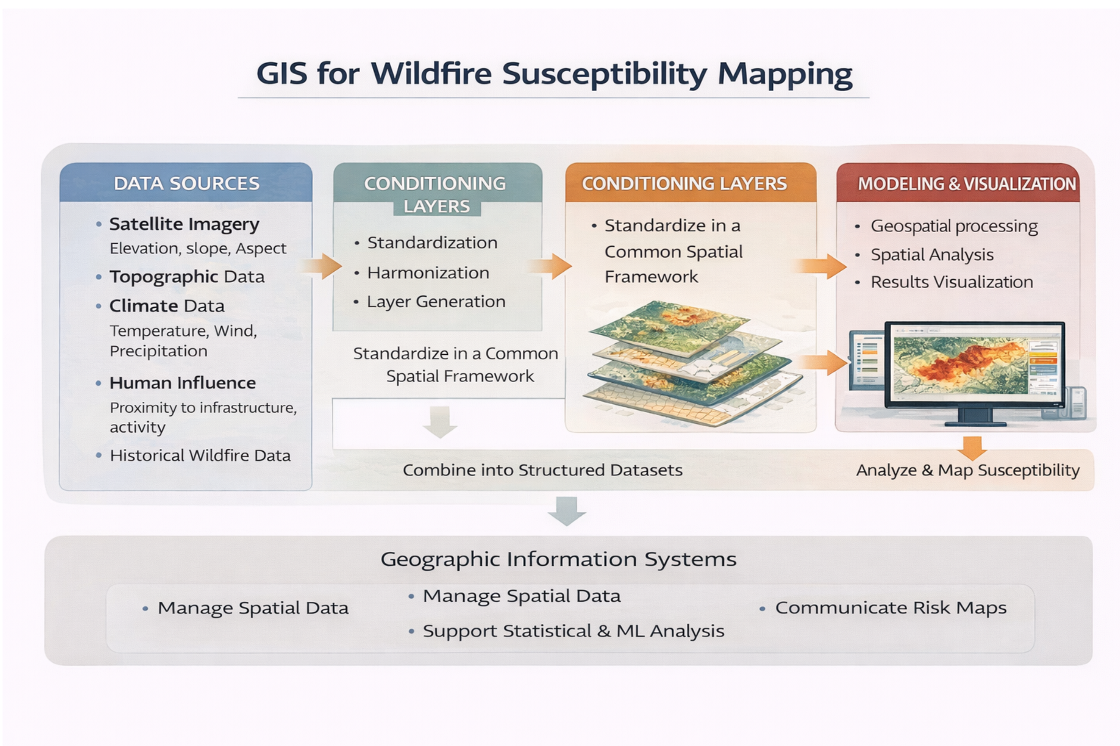

Figure 2.3: Role of GIS in Data Integration, Analysis, and Visualization for Wildfire Susceptibility Mapping

2.5 Methods Used for Wildfire Susceptibility Modeling

Wildfire susceptibility modeling (WSM) aims to explore and quantify the relationships between historical wildfire occurrences and the environmental and anthropogenic factors that influence them, with the ultimate goal of classifying areas according to their likelihood of experiencing future fires. Over time, a wide range of methodological approaches has been developed for this purpose, evolving from traditional prescriptive and rule-based techniques to more advanced data-driven methods, including machine learning (ML) and deep learning (DL) algorithms [84]. This methodological evolution reflects the increasing availability of spatial data, advances in computational capacity, and the need for more accurate and scalable predictive models.

Early WSM studies predominantly relied on statistical and multi-criteria decision analysis (MCDA) methods to assess wildfire risk. Techniques such as the Analytic Hierarchy Process (AHP),Vlse Kriterijumska Optimizacija Kompromisno Resenje (VIKOR), and Technique for Order Preference by Similarity to Ideal Solution(TOPSIS) have been widely employed to classify regions based on their susceptibility to wildfires by aggregating multiple conditioning factors using predefined weights. For example, researchers in [93] applied VIKOR and TOPSIS to generate a wildfire susceptibility map for Mula Province in Turkey, while authors in [94] combined AHP and TOPSIS to assess wildfire risk in West Sikkim, India. These methods provide transparent and interpretable decision-making frameworks and are relatively easy to implement. However, their reliance on expert-defined weights, static assumptions, and linear relationships, as well as their limited predictive capability and scalability, restrict their effectiveness for large-scale and complex wildfire susceptibility assessments.

In recent years, machine learning methods have gained substantial popularity in WSM due to their ability to process large and heterogeneous datasets, capture nonlinear interactions among variables, and generate robust predictive outputs. Many studies adopt a comparative approach, evaluating multiple ML algorithms to identify the most suitable model for a given study area. For instance, decision treebased algorithms were applied in [84] to produce a WSM for the Liguria region in Italy. Among ML techniques, Random Forest (RF) has emerged as one of the most frequently used models, owing to its high predictive accuracy, robustness to overfitting, and capacity to handle both linear and nonlinear relationships. In [83], RF outperformed several alternative ML algorithms according to multiple evaluation metrics. Gradient Boosting Machine (GBM) models, including CatBoost, XGBoost, and LightGBM, have also been extensively applied in WSM. A comparative study in [77] demonstrated that CatBoost achieved an accuracy of 95.47% in a wildfire susceptibility assessment in Turkey, highlighting the strong performance of boosting-based approaches.

To further enhance model interpretability and predictive power, some researchers have incorporated temporal dimensions into WSM. By segmenting datasets according to seasonal or temporal patterns, these studies examine how the influence of conditioning factors varies throughout the year.

For example, authors in [70, 90] applied ML algorithms separately to summer and winter datasets, revealing distinct seasonal dependencies in wildfire occurrence.

Deep learning approaches have also been explored in wildfire susceptibility modeling, particularly with the growing availability of high-resolution satellite imagery and gridded environmental data. Convolutional Neural Networks (CNNs) are commonly used to extract spatial patterns from imagery, while Deep Neural Networks (DNNs) model complex nonlinear relationships among multiple predictors. For instance, a CNN-based approach achieved an accuracy of approximately 95% for wildfire susceptibility mapping in the Biobío and Ñuble regions of Chile [73]. Similarly, a DNN model applied in the Gippsland region of Australia reached an accuracy of 0.92, underscoring the potential of DL models to capture complex spatial and environmental interactions in WSM [75].

Beyond CNN- and DNN-based frameworks, recent studies have increasingly framed wildfire risk assessment as a spatiotemporal learning problem. In this context, transformer-based architectures have been introduced to jointly model spatial dependencies and temporal evolution in Earth system variables. These models are particularly effective in capturing long-range interactions and dynamic risk patterns, thereby complementing static susceptibility mapping approaches with time-aware representations and nonlocal relationships [95].

To improve prediction accuracy and model robustness, several studies have proposed hybrid and ensemble modeling strategies. For example, authors in [96] introduced a hybrid framework integrating Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Support Vector Machines (SVM) with Genetic Algorithms and meta-heuristic optimization techniques. Similarly, [76] employed Genetic Algorithms and other meta-heuristic methods to optimize WSM in the Zagros Mountains of Iran. Ensemble learning approaches have also been widely investigated. In [82], three base modelsSVM, Multilayer Perceptron (MLP), and XGBoostwere combined in a stacked architecture, followed by meta-learning and hyperparameter optimization in a second layer to reduce subjectivity and enhance overall predictive performance.

Table **??** provides a summary of representative methods employed in selected WSM studies, without repetition, along with their reported accuracies. Collectively, these developments illustrate a clear transition from traditional rule-based and statistical methods toward advanced ML and DL approaches in wildfire susceptibility mapping, driven by the increasing demand for higher accuracy, scalability, and methodological robustness.

Table 2.1: Methods used for wildfire susceptibility mapping

Reference	**Methods**	**Accuracy**	**Validation Method**
[97]	Fuzzy AHP	0.83	MCDA ranking validation
	AHP	0.81	MCDA ranking validation
	Frequency Ratio	0.77	MCDA ranking validation
[93]	VIKOR	0.83	MCDA ranking validation
	TOPSIS	0.77	MCDA ranking validation
[77]	MLP-NET	0.88	ROCAUC
	Random Forest	0.93	ROCAUC
	Logistic Regression	0.79	ROCAUC
[85]	Recurrent Neural Network	0.97	ROC/AUC
	LSTM	0.96	ROC/AUC
[98]	KNearest Neighbor	0.88	kfold
	Logistic Regression	0.87	assumed kfold
	Decision Tree	0.81	kfold
[73]	SVM	0.89	ROCAUC
	Boosted Regression Tree	0.91	ROCAUC
[99]	CatBoost	0.95	ROCAUC
	XGBoost	0.92	ROCAUC
	LightGBM	0.94	T ROCAUC

2.6 Interpretability in WSM

Interpretability in Wildfire Susceptibility Mapping (WSM) refers to the ability to understand and explain how a model arrives at its predictions. Rather than only assigning a susceptibility score to each spatial unit, an interpretable model clarifies the contribution of individual environmental and anthropogenic factors to that prediction [35]. This allows researchers and practitioners to link model outputs to physical wildfire processes, such as the influence of terrain, vegetation, climate, or human activity, thereby bridging computational modeling and domain knowledge[38].

In wildfire risk management, predictive performance metrics alone are not sufficient for validating models used in decision-making. High accuracy does not guarantee that a model relies on meaningful or physically plausible relationships, and models may inadvertently exploit spurious correlations. Interpretability is therefore essential for building trust, ensuring accountability, and supporting informed decisions, particularly when susceptibility maps guide land-use planning, prevention strategies, or emergency response actions that must be justified to stakeholders [75].

Interpretability in WSM begins with transparent justification of conditioning factor selection. Factors should be chosen based on their documented relevance to wildfire occurrence, supported by previous studies and expert knowledge, rather than data availability alone. Clear reporting of factor selection, correlation analysis, and multicollinearity assessment helps ensure that predictors provide complementary information and that model behavior reflects known wildfire dynamics rather than redundant or irrelevant signals [86].

For complex machine learning models commonly used in WSM, SHAP (SHapley Additive exPlanations) has become a standard tool for interpretability. SHAP quantifies the contribution of each factor to model predictions, enabling both global interpretation (by ranking the most influential variables across the study area) and local interpretation (by explaining why a specific location is classified as high risk). This dual perspective enhances understanding of model behavior and supports targeted wildfire prevention measures [86].

Spatial visualization is a key component of interpretable WSM. Beyond susceptibility maps, explanatory maps such as dominant-factor maps or spatial SHAP value maps provide insight into the drivers of wildfire risk across different areas. These visualizations transform numerical explanations into intuitive spatial patterns, helping decision-makers distinguish between regions dominated by climatic stress, vegetation conditions, or human activity, and enabling more effective, location-specific management strategies [91].

2.7 Research Gaps in Wildfire Susceptibility Mapping

A primary challenge in wildfire susceptibility mapping lies in data availability and quality. In many regions, particularly in developing and low-resource countries, local ground-based observations of climate, fuel conditions, and human activities are sparse or entirely absent. This limitation forces researchers to rely heavily on global datasets derived from satellites or reanalysis products [73]. While these sources provide valuable and consistent spatial coverage, they often lack the spatial resolution, temporal frequency, or local calibration needed to accurately represent microclimatic conditions and human-driven ignition processes, thereby constraining the reliability and scope of WSM studies, see Figure 2.4.

Another important research gap concerns the selection of conditioning factors. Although a large number of environmental and anthropogenic variables have been proposed in the literature, there is no standardized or universally accepted framework for selecting the most relevant factors for a given study area [78]. Many studies adopt commonly used variables without systematically evaluating their physical relevance, redundancy, or regional specificity. As a result, models may suffer from multicollinearity, overfitting, or the inclusion of non-informative predictors, which reduces both predictive robustness and interpretability, see Figure 2.4.

The methodology used for labeling wildfire occurrence data also presents significant challenges. Most WSM studies rely on historical fire records to define fire and non-fire samples; however, the criteria used to generate absence or low-susceptibility labels vary widely across studies [84]. Inconsistent buffer distances, temporal windows, and sampling strategies can introduce bias, mislabeling, or spatial autocorrelation, ultimately affecting model learning and evaluation. The lack of standardized and transparent labeling protocols remains a major obstacle to comparability and reproducibility in WSM research, see Figure 2.4.

Model selection and data treatment strategies represent another unresolved issue. While machine learning and deep learning methods have demonstrated strong predictive capabilities, there is no consensus on which models are most appropriate under different data conditions, spatial scales, or objectives [86]. Many studies prioritize performance comparisons without sufficiently addressing data preprocessing, class imbalance handling, or model complexity relative to dataset size. This often leads to the use of overly complex models where simpler approaches may suffice, or to limited exploration of hybrid and spatiotemporal architectures, see Figure 2.4.

Validation practices in WSM also reveal notable limitations. The majority of studies rely on traditional evaluation metrics such as accuracy, AUC, or F1-score, which may not adequately capture spatial reliability, uncertainty, or operational usefulness [77]. These metrics can be insensitive to spatial clustering, class imbalance, and threshold selection, potentially overstating model performance. The absence of spatially explicit validation, calibration analysis, and uncertainty quantification restricts the confidence with which susceptibility maps can be used in real-world decision-making, see Figure 2.4.

Finally, the interpretation of model results remains an underdeveloped aspect of WSM research. Despite increasing model complexity, many studies provide limited insight into the drivers of predicted wildfire susceptibility, focusing primarily on performance outcomes [91]. Insufficient use of explainability techniques and explanatory maps reduces the transparency and practical applicability of susceptibility models. This gap hinders the translation of WSM outputs into actionable wildfire management strategies and underscores the need for systematic integration of interpretability into future research, see Figure 2.4.

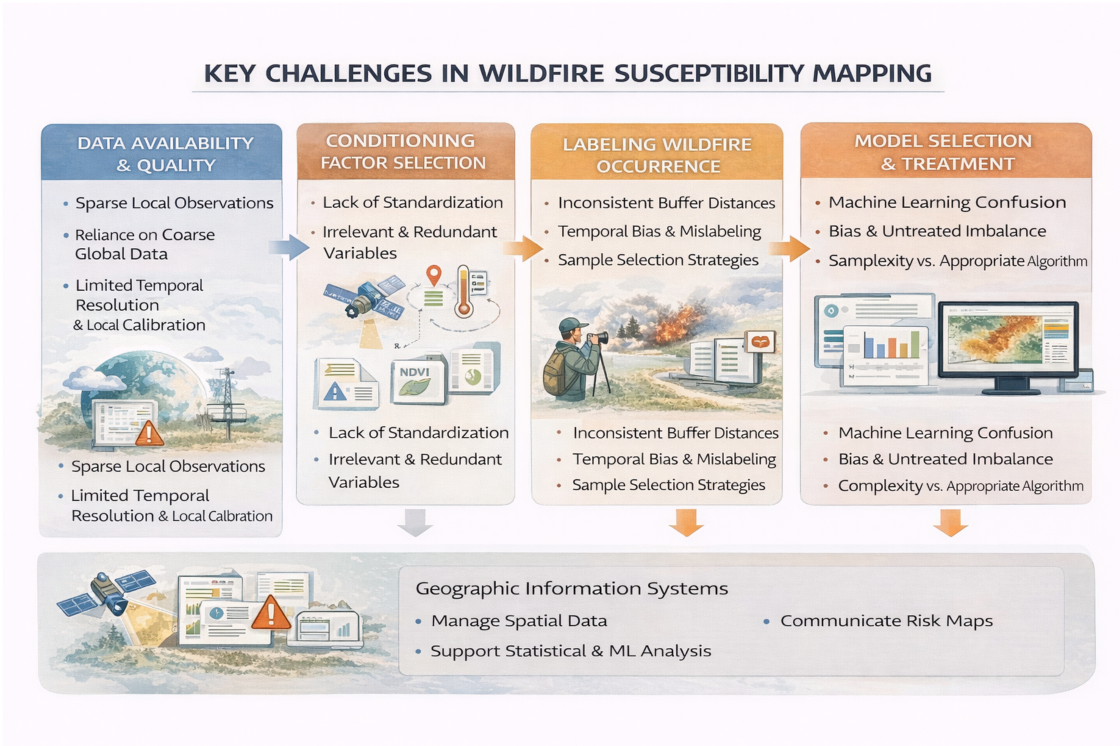

Figure 2.4: Key Challenges in Wildfire Susceptibility Mapping

2.8 Conclusion

This chapter has reviewed the state of the art in wildfire susceptibility mapping, emphasizing how researchers conceptualize susceptibility, construct spatial datasets, select conditioning factors, and apply increasingly sophisticated modeling techniques. The literature demonstrates a clear shift from rule-based and expert-driven approaches toward data-driven machine learning and deep learning frameworks, supported by GIS and remote sensing technologies. At the same time, the review reveals substantial variability in factor selection, labeling strategies, validation methods, and levels of interpretability, reflecting both regional constraints and methodological trade-offs.

Despite significant progress, the analysis highlights several unresolved challenges, including limited local data availability, the absence of standardized factor selection and labeling protocols, reliance on conventional evaluation metrics, and insufficient attention to model interpretability. These gaps underscore the need for susceptibility models that are not only accurate, but also transparent, context-aware, and operationally meaningful. Building on the insights derived from this review, the next chapter focuses on the literature and methodological foundations of wildfire susceptibility modeling in greater depth, providing the basis for the proposed framework and experimental design adopted in this research.

Part II

Methodology and Experimental Framework

Chapter 3

Study Area and Dataset Preparation

3.1 Introduction

A reliable wildfire susceptibility model depends fundamentally on the quality, consistency, and representativeness of the input dataset. Before any learning stage, it is therefore essential to construct a well-structured dataset that accurately captures the environmental and anthropogenic conditions influencing wildfire occurrence. Following the analysis of the study area and the identification of relevant wildfire-driving factors, this chapter details the process of transforming heterogeneous spatial information into a coherent and learning-ready dataset.

The dataset preparation workflow begins with the collection of multi-source data describing topographic, climatic, vegetation-related, and human-related factors. These data are then harmonized and preprocessed to ensure spatial consistency and numerical reliability. Special attention is given to data cleaning, multicollinearity assessment, and factor suitability to avoid bias and instability in subsequent modeling stages. Finally, an effective data labeling strategy based on historical wildfire records is applied to convert continuous spatial information into discrete susceptibility classes suitable for supervised learning. Through this structured pipeline, the chapter establishes a robust foundation for the modeling and evaluation phases that follow.

3.2 Study Area: Jijel Province

The Jijel Province is located in the northeastern part of Algeria, along the Mediterranean coastline. It is bordered by the Mediterranean Sea to the north, Skikda Province to the east, Mila and Constantine Provinces to the south, and Béjaïa Province to the west. Administratively, Jijel is composed of several districts (daïras) and municipalities, covering a total area of approximately 2,400 kmš [100]. The province is characterized by a complex terrain that includes coastal plains, mountainous areas belonging to the Tell Atlas range, and dense forested landscapes. Its geographic position and physical diversity make Jijel a representative case for wildfire studies in Mediterranean environments, see Figure 3.1.

Figure 3.1: Study area map (Jijel province, Algeria).

The selection of Jijel as the study area is motivated by its climatic and environmental characteristics, which are typical of the Mediterranean region. The province experiences a Mediterranean climate, with hot and dry summers and mild, wet winters, conditions that are well known to favor wildfire ignition and spread. In addition, Jijel shares similar climatic patterns, topographic features, and vegetation cover with the broader eastern northern Algeria region, including neighboring provinces. Dense forests, shrublands, steep slopes, and seasonal droughts are common across this area, making Jijel a suitable proxy for studying wildfire susceptibility in a larger regional context [100].

Furthermore, Jijel has been severely affected by recurrent wildfires in recent years, causing significant threats to human lives, residential areas, and critical infrastructure, as well as extensive damage to forest ecosystems and biodiversity. Several wildfire events have resulted in the loss of agricultural land, forest resources, and ecological services, highlighting the increasing vulnerability of the region to fire hazards. These repeated incidents underline the urgent need for advanced wildfire management strategies, including susceptibility assessment and early warning systems. Consequently, Jijel represents a critical and relevant case study where research efforts can contribute to risk reduction, sustainable land management, and improved disaster preparedness [100].

3.2.1 Wildfire Situation in Jijel Province

Jijel Province has historically been recognized as an active wildfire-prone area due to the combination of unfavorable natural conditions. The region experiences hot and dry summers, dense forest and shrub vegetation, and complex mountainous topography, all of which contribute to high wildfire susceptibility [100]. These conditions have repeatedly facilitated fire ignition and rapid spread, particularly during prolonged dry periods. One of the earliest major wildfire events occurred in 1994, which affected large forested areas; however, precise official statistics for this event remain unavailable, reflecting limitations in historical fire documentation [101].

In recent years, wildfire activity in Jijel and across northern Algeria has intensified significantly, largely driven by extreme temperature waves and prolonged drought conditions. Exceptional summer temperatures have exceeded 47 řC in some cases, creating critical fire weather conditions. In 2017 alone, more than 3,000 ha were burned in Jijel Province. The situation worsened in 2019, when catastrophic wildfires across 14 provinces, including Jijel, resulted in more than 90 fatalities nationwide. Subsequent years continued to record substantial burned areas, with approximately 1,200 ha affected in 2020 and 1,300 ha in 2023. Furthermore, analysis of wildfire statistics between 2012 and 2021 reveals a clear increasing trend in fire severity, with the average burned area per wildfire rising from approximately 5 ha to nearly 40 ha. This trend, illustrated in the corresponding Figure 3.2, highlights the escalating wildfire problem in the region [102].

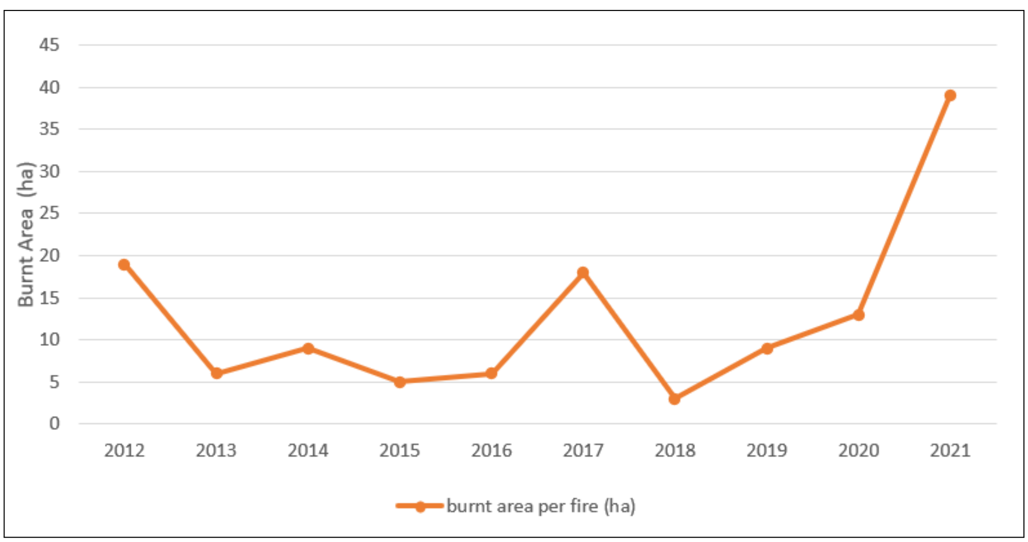

Figure 3.2: Burned area per fire in recent years in Jijel province.

The socio-economic and environmental impacts of these wildfires have been severe. Summer wildfire events pose a direct threat to human lives, particularly for farmers and rural communities living near forest boundaries, often requiring emergency evacuations. Agricultural areas have been repeatedly damaged, leading to significant economic losses. For instance, in 2019, more than 287 farmers were compensated for wildfire-related damages that affected over 150 ha of fruit tree plantations. These losses extend beyond immediate agricultural damage, encompassing long-term economic impacts related to tree destruction, insurance claims, and compensation mechanisms managed by national disaster risk agencies, thereby placing additional strain on the regional economy [101].

To cope with wildfire threats, Algerian authorities have implemented substantial firefighting and mitigation efforts. The primary institutions responsible for wildfire management are the Algerian Civil Protection and the Directorate of Forests, supported during the fire season by a multi-agency fire-risk committee that includes the army, hydrology services, police, and other authorities. Firefighting capacity has been strengthened through the acquisition of four Beriev Be-200 amphibious aircraft, twelve Air Tractor planes (three stationed in Jijel), and six helicopters dedicated to surveillance within the province. Additional infrastructure includes the construction of small airfields for refueling and water loading, the development of approximately 210 km of forest access roads, and the creation of firebreak trenches to limit fire spread. These operational measures are complemented by continuous training programs for Civil Protection agents, conducted in collaboration with international partners such as France, Russia, Tunisia, and Germany [102].

Despite these considerable efforts, wildfire management remains a major challenge, particularly under extreme fire weather conditions. As demonstrated in previous chapters, suppression alone is increasingly insufficient to address the growing intensity and frequency of wildfires. While Algerias Civil Protection strategy has made significant progress in terms of resources and operational response, limitations persist in preventive planning, environmental analysis, and early detection capabilities. This study directly addresses these gaps by contributing to research-oriented wildfire management, with the objective of improving understanding of wildfire occurrence conditions and behavior, and supporting more effective prevention and control strategies [102].

3.2.2 Wildfire Conditioning Factors in the Jijel Region

Wildfire occurrence in Jijel Province is governed by a complex interaction between environmental and anthropogenic factors that collectively shape ignition probability, fire behavior, and spatial fire patterns. The province exhibits a combination of rugged topography, Mediterranean climatic conditions, dense and diverse vegetation cover, and significant human pressure, all of which contribute to its high wildfire susceptibility. Understanding how each factor operates within the specific geographic and ecological context of Jijel is essential for interpreting wildfire dynamics and justifying their inclusion in susceptibility modeling [103].

3.2.2.1 Topographic Factors

Topography plays a fundamental role in controlling wildfire behavior in Jijel. The province is characterized by rugged mountainous terrain dominated by forested massifs such as the Salma, Bouazza, and Al-Afroun ranges, which together cover more than 80% of the territory. Elevation rises sharply from sea level along the Mediterranean coast to interior highlands, creating steep slopes, narrow valleys, and highly dissected relief. These features influence local microclimates, wind exposure, fuel distribution, and fire spread dynamics, while also constraining accessibility for ground firefighting operations [102].

Elevation acts as a background environmental gradient rather than a direct driver of fire occurrence. In Jijel, lower elevations near the coast and valley bottoms experience warmer conditions and host agricultural mosaics and peri-urban areas, which tend to increase ignition pressure. In contrast, higher elevations may retain moisture longer and support different vegetation types [104]. However, mid- to high-altitude ridgelines are often exposed to strong winds and summer drought, making them susceptible despite cooler average conditions, as shown in Figure 3.3. Thus, elevation interacts with climate, vegetation, and wind rather than dictating wildfire risk independently [105].

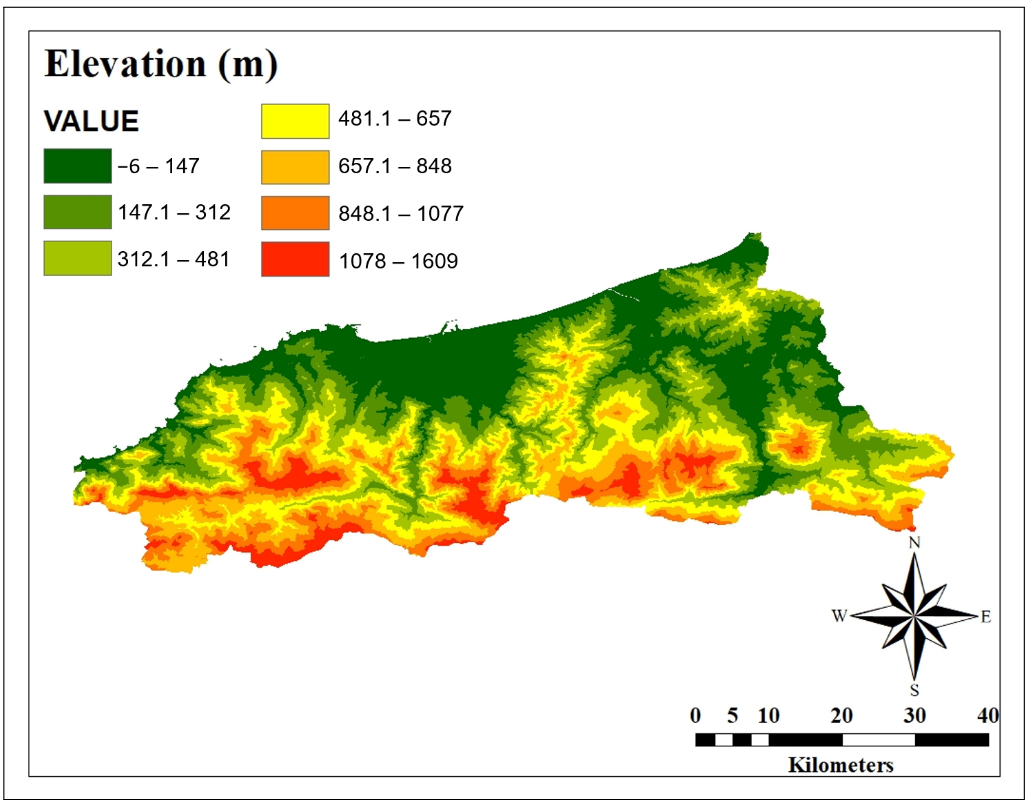

Figure 3.3: Elevation map of Jijel Province.

Slope is one of the most influential topographic controls on wildfire behavior in Jijels mountainous landscapes. Steep slopes enhance flame attachment and convective heat transfer, accelerating uphill fire spread and increasing fire intensity [106]. In many parts of the province, narrow valleys and abrupt gradients allow fires to rapidly transition from surface to fast-moving fronts, especially under aligned wind conditions, see Figure 3.4. The influence of slope is strongly nonlinear: while gentle slopes have limited effect, steep terrain can dramatically increase spread rates and reduce suppression effectiveness [107].

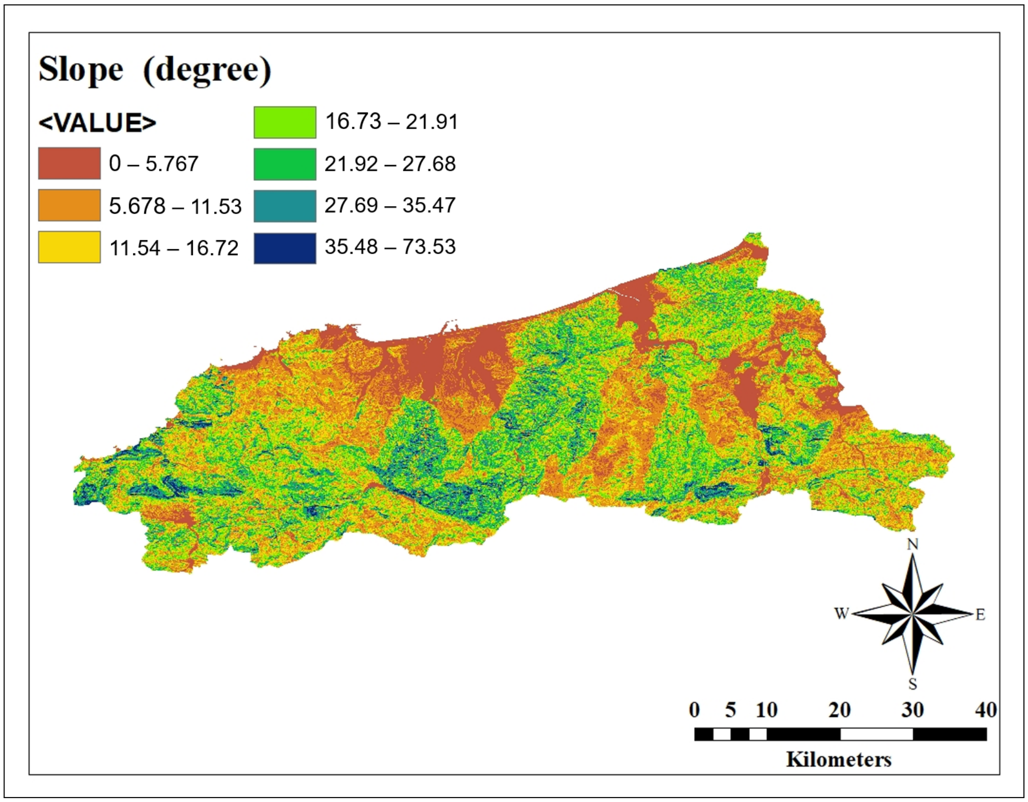

Figure 3.4: Slope distribution of Jijel Province.

Aspect further modulates wildfire susceptibility by controlling solar radiation exposure and fuel moisture. South- and southwest-facing slopes in Jijel receive higher solar input during summer, leading to faster drying of litter and fine fuels and reduced nighttime moisture recovery. North-facing slopes, shaded gullies, and sheltered aspects generally maintain higher moisture levels and lower susceptibility [108]. These contrasts are particularly pronounced in open shrublands and maquis formations, where canopy shading is limited and solar exposure directly affects fuel receptivity, see Figure 3.5.

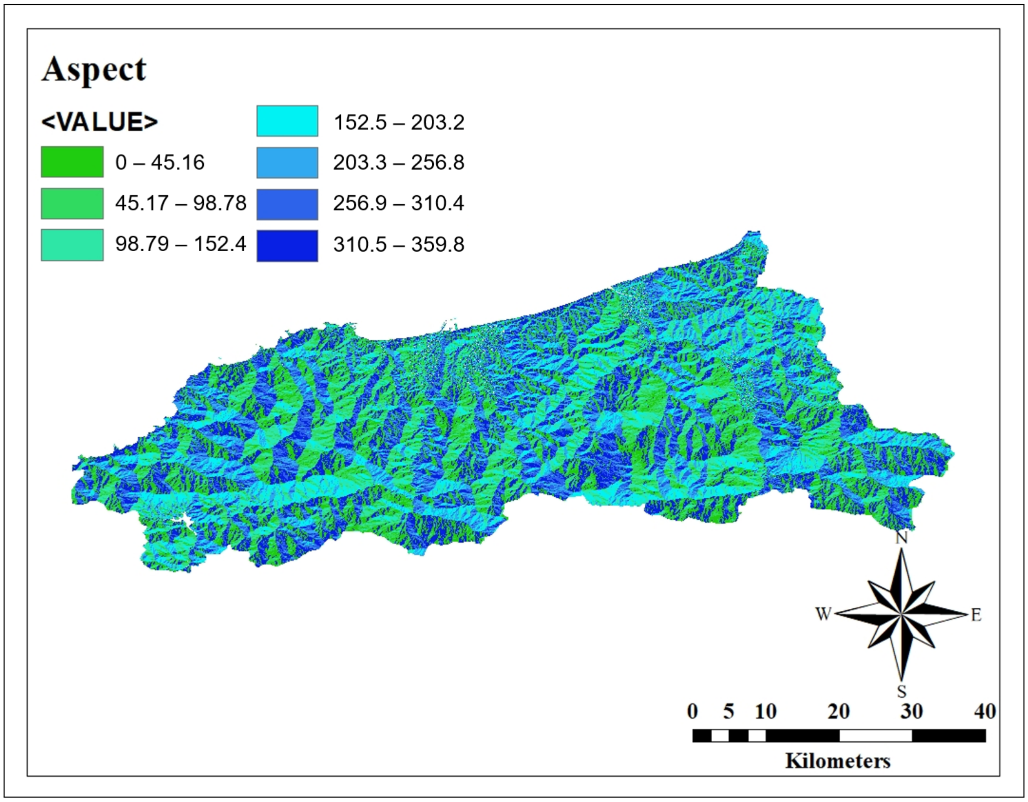

Figure 3.5: Aspect orientation map of Jijel Province.

3.2.2.2 Climatic Factors

Jijel experiences a typical Mediterranean climate, characterized by wet and mild winters and hot, dry summers. Annual precipitation generally ranges between 800 and 1200 mm, concentrated in autumn and winter, while summers are marked by prolonged dry periods[109]. Seasonal temperature contrasts strongly influence vegetation phenology and fuel moisture dynamics, resulting in a pronounced summer peak in wildfire occurrence.

Temperature is a key driver of fuel desiccation and ignition potential. Average and maximum summer temperatures promote progressive drying of fine fuels, while extreme heatwavesfrequently recorded in recent yearscan push fuels beyond critical moisture thresholds [110], see Figure 3.6. In Jijel, exceptional summer conditions have included temperatures exceeding 4750 řC, dramatically increasing fire risk (Figure 3.7). Minimum temperatures also play an important role by limiting overnight fuel moisture recovery; warm nights maintain fuel receptivity into the early morning hours, extending daily fire windows [111] (Figure 3.7).

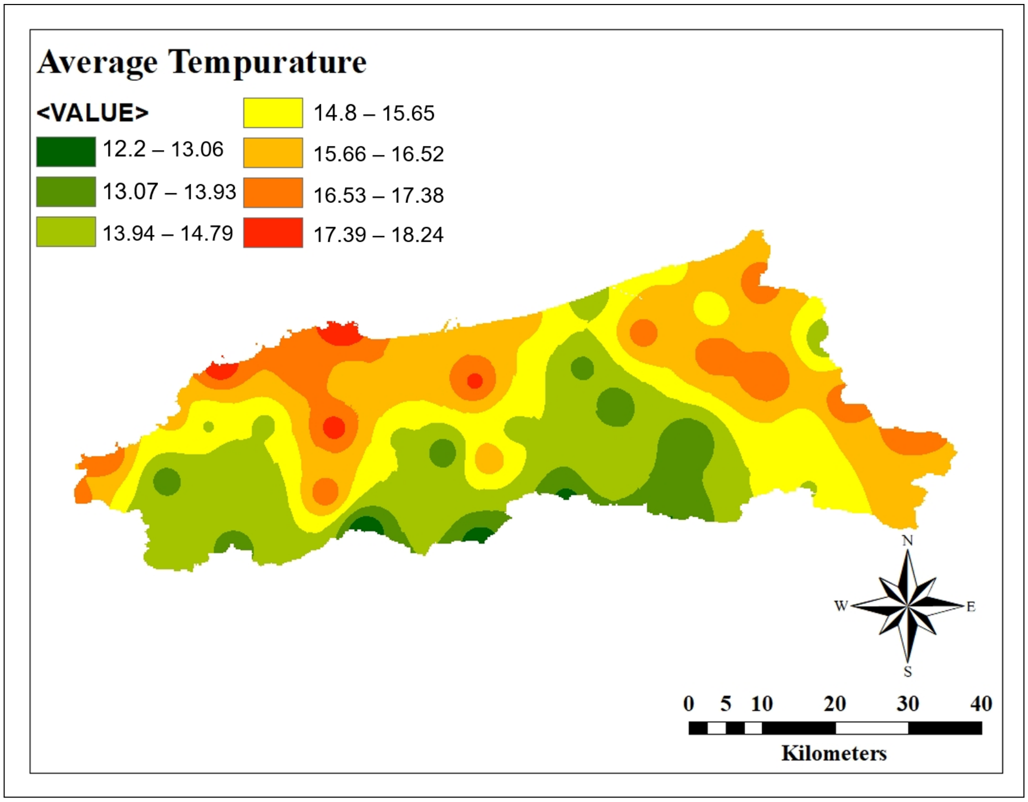

Figure 3.6: Spatial distribution of average temperature in Jijel Province.

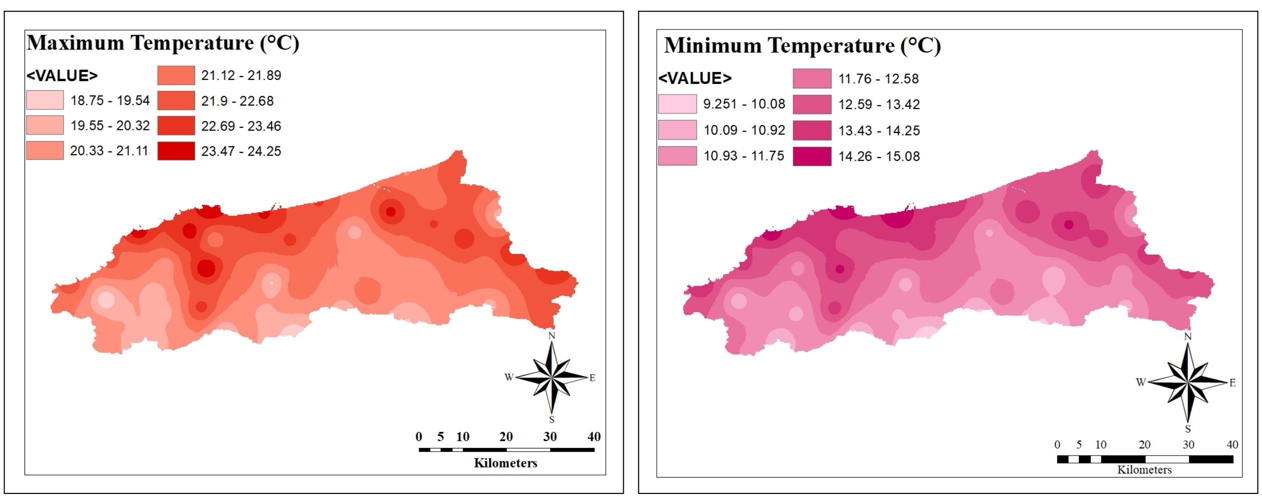

Figure 3.7: Spatial distribution of minimum and maximum temperature in Jijel Province.

Relative humidity directly governs fine fuel moisture content and is among the strongest short-term predictors of ignition potential during the fire season. Low humidity conditions, particularly when combined with high temperatures, create highly flammable environments [112], as shown in Figure 3.8. Wind acts as a dominant control on fire behavior by supplying oxygen, enhancing convective drying, tilting flames toward unburned fuels, and enabling ember transport. In Jijels complex terrain, wind effects are amplified along ridgelines and within valleys that channel airflow, allowing fires to spread rapidly and cross natural barriers (Figure 3.8).

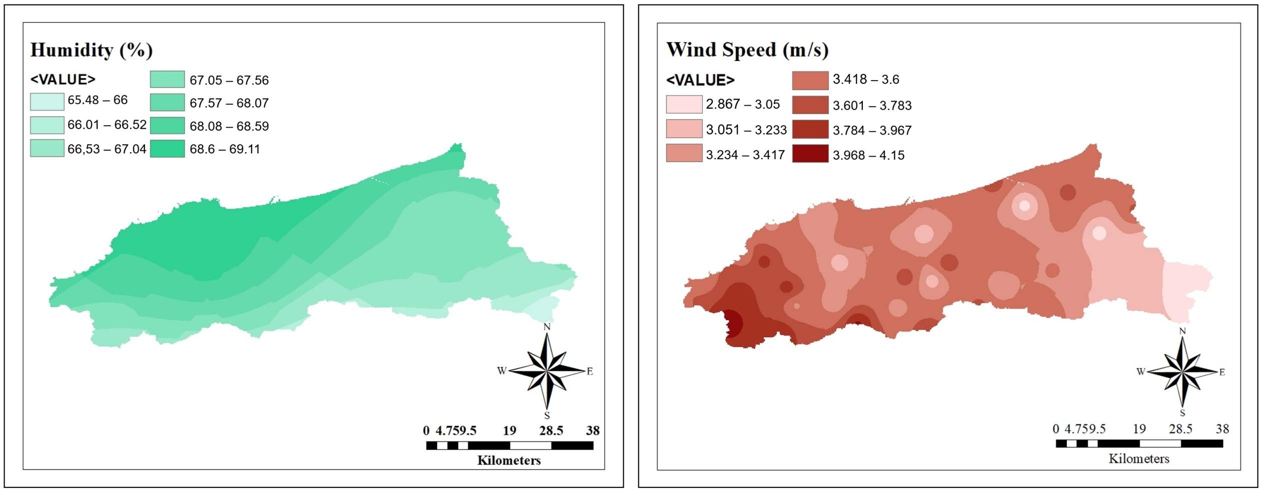

Figure 3.8: Spatial distribution of humidity and wind speed in Jijel Province.

Precipitation influences wildfire susceptibility over both short and medium timescales. Recent rainfall can temporarily suppress fire risk by increasing fuel moisture, but this effect decays rapidly during hot summer conditions. Conversely, prolonged precipitation deficits create cumulative drought stress that raises baseline susceptibility. The response time varies by fuel type, with grasses and shrubs reacting quickly to drying and woody fuels integrating moisture conditions over longer periods [112], see Figure 3.9.

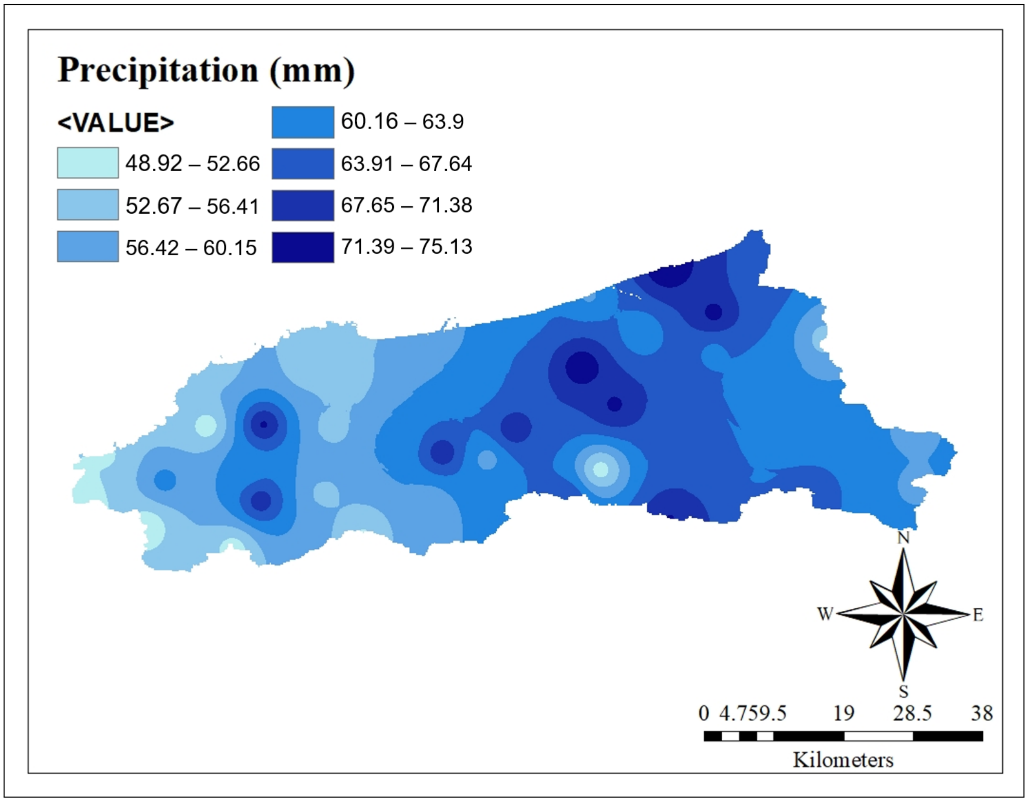

Figure 3.9: Spatial distribution of precipitation in Jijel Province.

3.2.2.3 Vegetation and Land Cover

Vegetation represents the fuel component of the fire triangle and is a dominant determinant of wildfire behavior in Jijel [103]. Approximately 60% of the province is covered by forests, maquis, and scrub, hosting rich biodiversity including cedar, oak, pine, and olive species. Evergreen and sclerophyllous vegetation is widespread, forming continuous or semi-continuous fuel beds across slopes, foothills, and transitional zones (Figure 3.10).

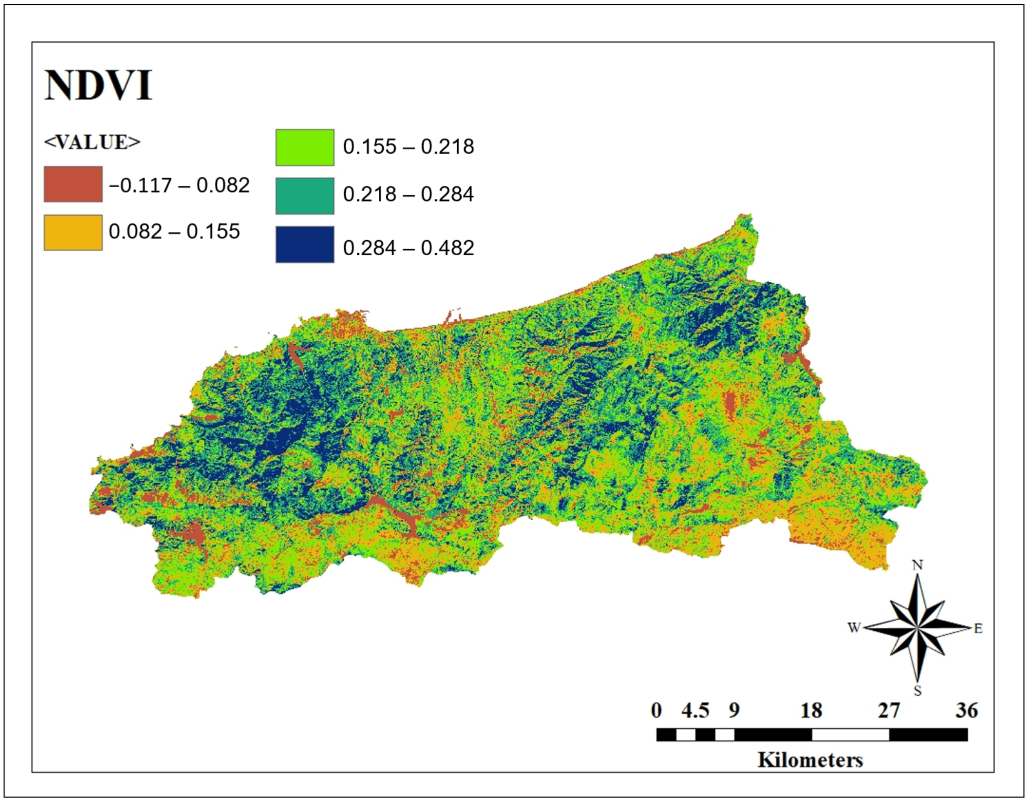

Figure 3.10: Normalized Difference Vegetation Index (NDVI) map of Jijel Province.

Vegetation condition influences both fuel quantity and continuity. Intermediate vegetation densitytypical of shrublands, grasslands, and forest edgesoften corresponds to the highest wildfire susceptibility, as it provides abundant fine fuels that ignite easily and carry fire rapidly under wind [113]. Very sparse vegetation limits fire spread due to fuel discontinuity, while very dense, moist vegetation may resist ignition unless prolonged drought conditions prevail. In Jijel, the mosaic of forests, agricultural lands, and grazing areas creates heterogeneous fuel structures that favor ignition at interfaces and along land-use boundaries [114].

Seasonal vegetation dynamics are critical. During spring, vegetation growth may increase fuel continuity, while summer curing transforms once-green biomass into highly receptive fine fuels. Abrupt vegetation stress during heatwaves further elevates flammability, particularly in maquis and scrub ecosystems that dominate large portions of the province [115].

3.2.2.4 Hydrological Influences

Hydrological features introduce localized moisture gradients that can moderate wildfire occurrence. Jijels hydrographic network consists of short, steep catchments draining from the mountains toward the Mediterranean Sea. River corridors and riparian zones often retain higher moisture levels and create cooler microclimates, which can hinder ignition and slow fire spread [115].

However, the protective effect of watercourses is not uniform. River valleys frequently coincide with roads, agricultural activity, and settlements, increasing human presence and ignition pressure. In steep terrain, valleys may trap cool, humid air at night but rapidly ventilate under daytime upslope winds, shifting their influence across the diurnal cycle. As a result, proximity to rivers acts as a moderating factor that can locally reduce susceptibility, with important exceptions driven by human activity and wind alignment [116], see Figure 3.11.

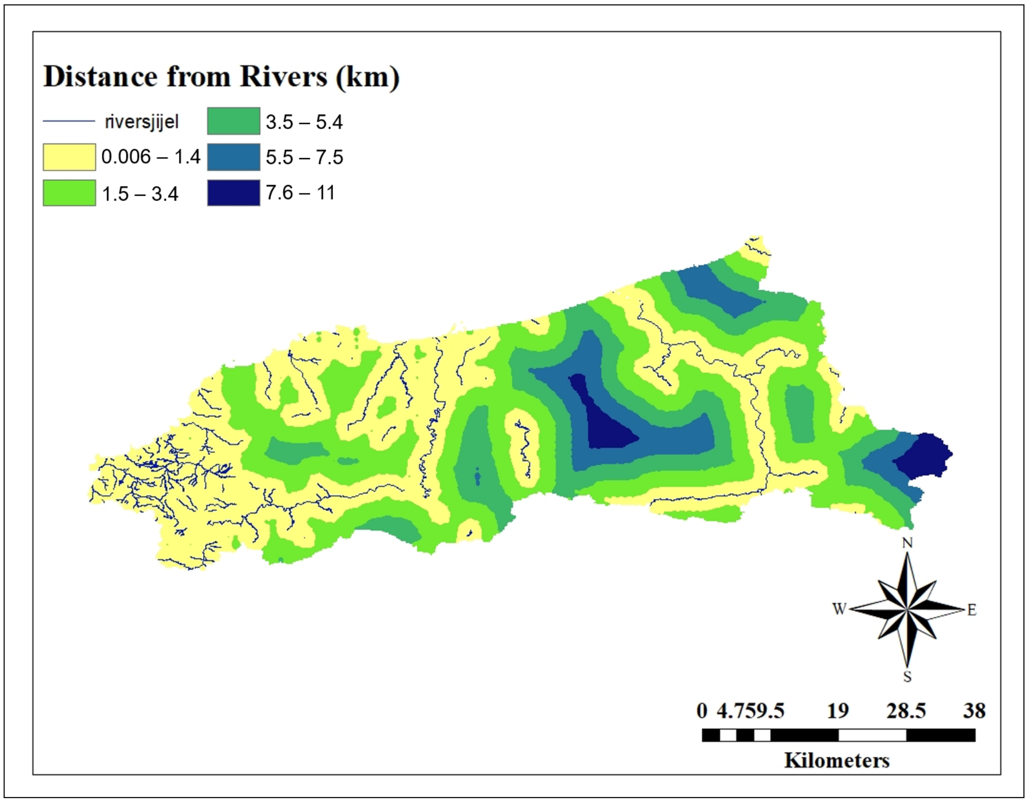

Figure 3.11: Distance to rivers map of Jijel Province.

3.2.2.5 Anthropogenic Factors

Anthropogenic factors are central to the occurrence of wildfires in Jijel, as human activities account for the majority of the ignitions. Proximity to roads reflects accessibility for vehicles, maintenance operations, recreation, and daily movement, all of which increase the likelihood of accidental ignitions. Roadside vegetation management and linear infrastructure further contribute to ignition risk [117].

Distance to settlements and zones of human activity captures sustained presence related to agriculture, grazing, tourism, and peri-urban expansion. Jijel exhibits a pronounced wildlandurban and wildlandagriculture interface along coastal areas and valley bottoms, where human activities intersect with continuous natural fuels [118]. These interfaces are consistently associated with elevated ignition density, especially during summer when agricultural practices and outdoor activities intensify (Figure 3.12).

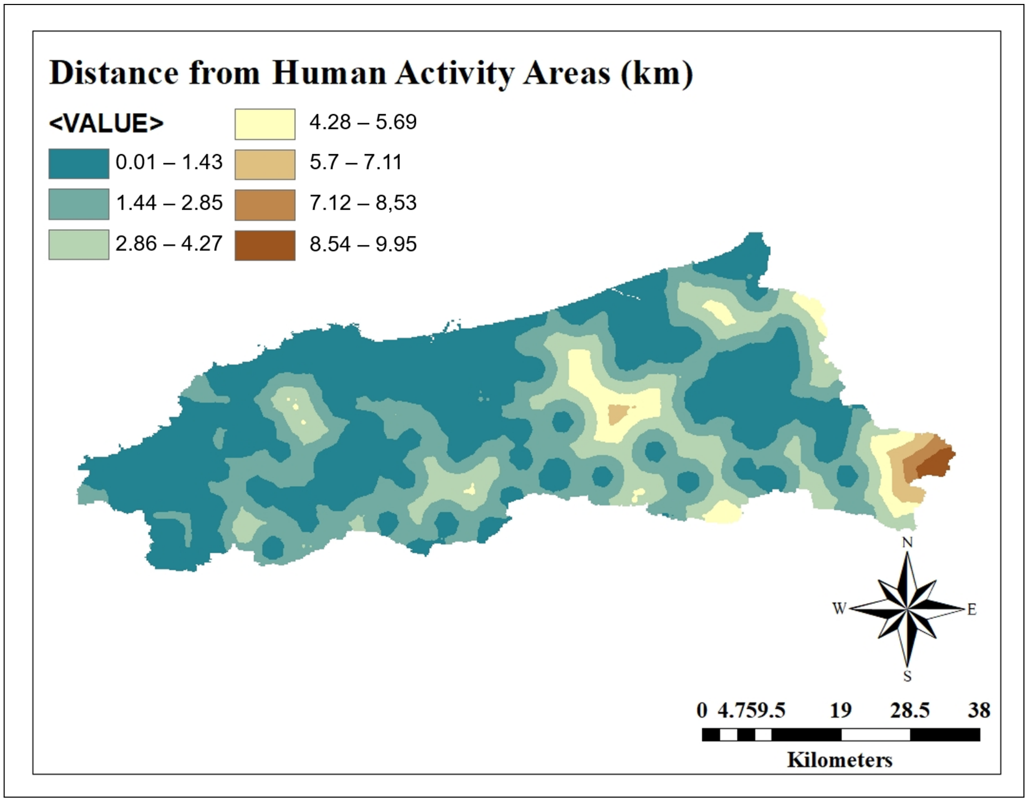

Figure 3.12: Distance to settlements and human activity areas in Jijel Province.

The influence of human-related factors is strongest at short distances and diminishes with remoteness, but it is heavily modulated by terrain, land ownership patterns, and accessibility [119]. Roads provide linear access corridors, while settlements represent diffuse activity centers; together, they offer complementary perspectives on anthropogenic fire pressure, see Figure 3.13.

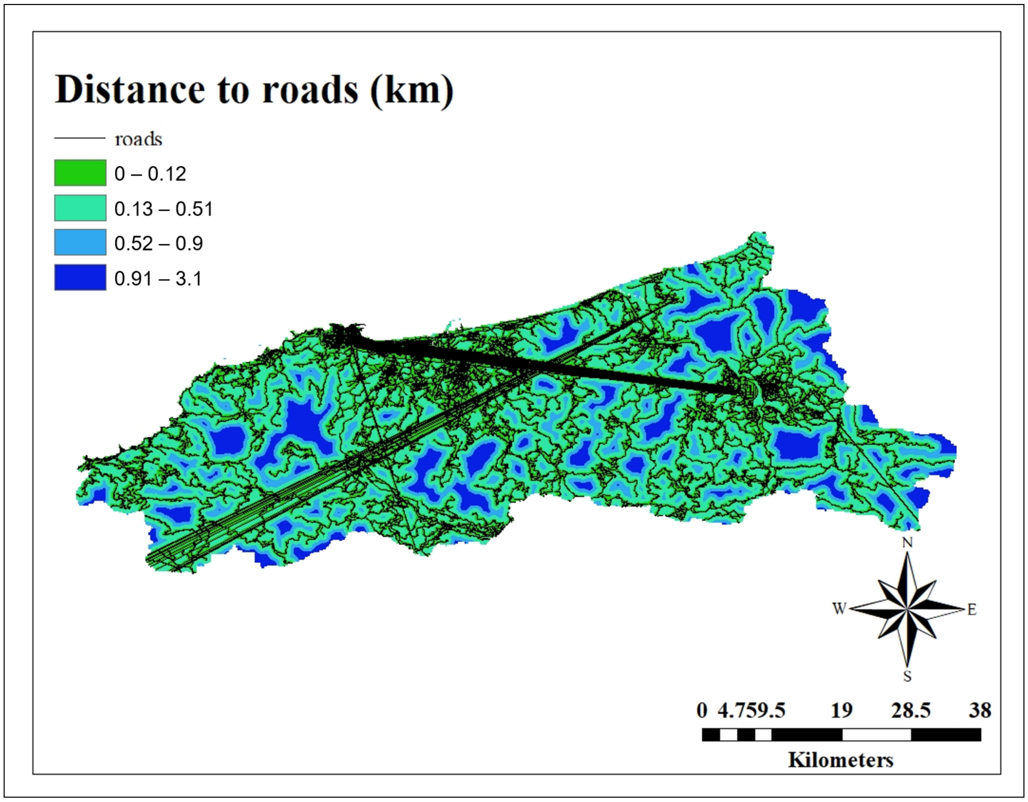

Figure 3.13: Distance to roads map of Jijel Province.

3.2.2.6 Historical Fire Occurrence

Past wildfire occurrence reflects persistent spatial patterns driven by stable combinations of fuels, access, and microclimate. Areas that have burned repeatedly in Jijel often do so because ignition sources remain active and landscape configurations continue to favor fire spread [120]. Historical fire patterns therefore provide empirical evidence of underlying susceptibility rather than isolated events (Figure 3.14).

At the same time, recent burns can temporarily reduce local fuel loads, creating short-term negative feedbacks [121]. This dual role highlights the need to interpret historical fire influence over appropriate temporal horizons, capturing long-term propensity while acknowledging short-term fuel depletion effects.

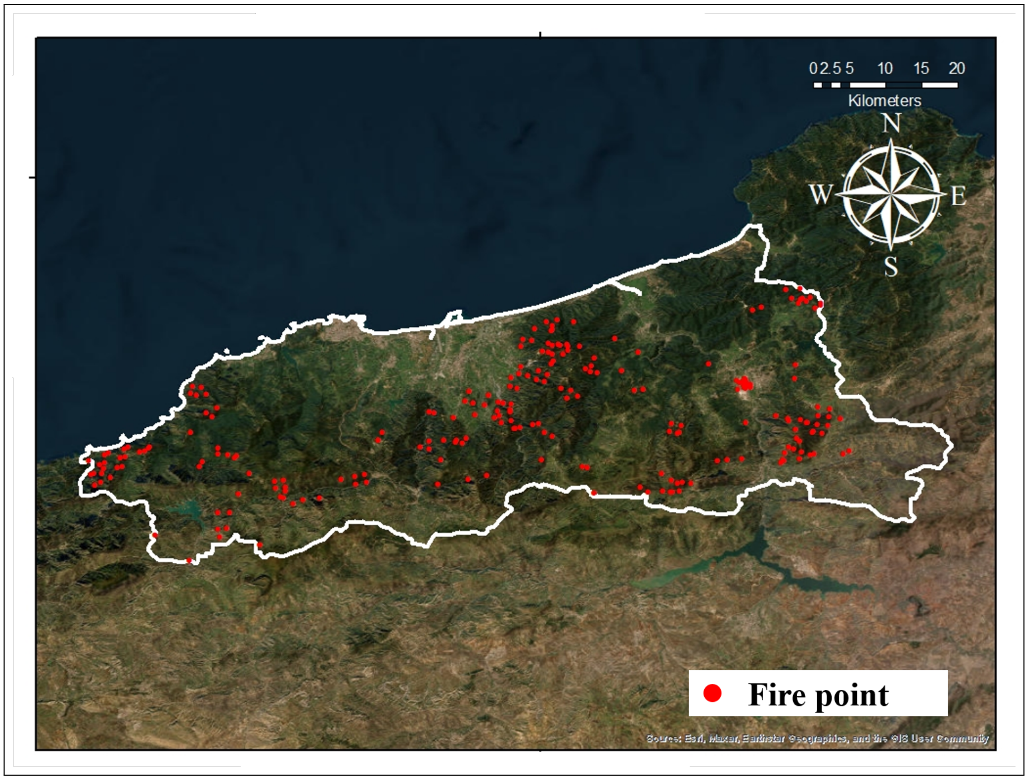

Figure 3.14: Spatial distribution of historical wildfire occurrences in Jijel Province.

3.2.2.7 Interactions and Combined Effects

While each factor contributes a distinct mechanism, wildfire susceptibility in Jijel emerges primarily through interactions and thresholds. Wind aligns with slope and aspect to accelerate fires on sun-exposed hillsides; temperature and humidity jointly regulate fuel moisture; vegetation condition interacts with drought to define ignition windows; and human presence amplifies risk under adverse environmental conditions. These coupled effects explain why extreme wildfire events often occur during short periods when multiple drivers align.

Together, the environmental and anthropogenic characteristics of Jijel create a landscape highly sensitive to wildfire ignition and spread. The provinces dense vegetation, rugged terrain, Mediterranean climate, and extensive humanenvironment interfaces justify the integrated consideration of topography, climate, vegetation, hydrology, and human activity in wildfire susceptibility analysis and support its selection as a representative and critical case study for wildfire management research.

3.3 Dataset Preparation

After characterizing the study area and examining the environmental and anthropogenic drivers of wildfire occurrence, the next step in developing a wildfire susceptibility mapping framework is the construction of a reliable and analysis-ready dataset[16]. This phase begins with the collection of spatial data describing the selected conditioning factors from multiple available sources. The acquired datasets are then subjected to a series of preprocessing operations aimed at cleaning, harmonizing, and structuring the data to ensure consistency and quality. Finally, an appropriate data labeling strategy is applied to associate environmental conditions with wildfire occurrence patterns, resulting in a coherent dataset that is suitable for machine learning and analytical modeling.

3.3.1 Data Collection

After defining the conditioning factors based on the characteristics of the study area and the wildfire phenomenon, the data collection process was initiated. Priority was given to locally collected datasets whenever available, as local data generally provide higher spatial detail, better temporal representativeness, and greater reliability compared to global products [122]. However, when local data were unavailable or incomplete, well-established global datasets were used as complementary sources to ensure full spatial coverage of the study area.

In the data collection phase, a total of 14 variables were gathered to support wildfire susceptibility mapping. These variables comprised topographical factors (elevation, slope, and aspect), climatic factors (minimum, maximum, and average temperature, wind speed, relative humidity, and precipitation), vegetation data represented by the Normalized Difference Vegetation Index (NDVI), hydrological factors (distance to rivers), anthropogenic factors (distance to roads and residential areas), and historical wildfire occurrence data [123].

Topographical data, including elevation, slope, and aspect, were derived from a Digital Elevation Model obtained from the Shuttle Radar Topography Mission (SRTM) through the EarthExplorer platform, with a spatial resolution of 30 m [124]. Climatic variables were provided by the Algerian National Office of Meteorology (NOM) based on observations from the Jijel meteorological station, with a spatial resolution of approximately 1 km. These data included minimum, maximum, and average temperatures, wind speed, relative humidity, and precipitation.

Vegetation data were collected using MODIS imagery from NASAs Terra platform. NDVI layers were generated at a 30 m resolution to ensure consistency with other spatial datasets [125]. Anthropogenic and hydrological datasets were generated using GIS tools by processing base spatial layers such as roads, rivers, settlements, and land-use maps obtained from Google Earth and national cartographic sources. Distance-to-road, distance-to-residential areas, and distance-to-river layers were produced to represent human activity and hydrological features within the study area.

Historical wildfire data were collected from official records provided by the Algerian Civil Protection authorities. In Jijel Province, 244 forest wildfire events were recorded over the period 20212023, including 121 fires in 2021, 34 fires in 2022, and 89 fires in 2023. These records were spatially mapped to generate wildfire occurrence layers. To enhance data completeness, wildfire occurrences from 2024 and 2025 were also collected and complemented with satellite-based fire detections from MODIS active fire products (Terra and Aqua). A total of 2348 fire points were retrieved and cross-validated with Civil Protection records to ensure consistency and reliability [126].

All collected datasets, their sources, and value ranges are summarized in Table 3.1. Together, these spatial layers constitute the core database used in this study and provide the foundation for subsequent preprocessing, labeling, and modeling stages. The final dataset comprises more than 2.6 million records described by 14 factors, as illustrated in Figure 3.15.

Table 3.1: Data used for wildfire susceptibility.

Data type	**Data**	**Range**	**Source**
Topography	Elevation	$[-6, 1906]$ m	SRTM (earthexplorer.usgs.gov)
	Aspect	$[0, 359.8]$	
	Slope	$[0, 73.53]$	
Climate	Avg. Temperature	$[12.2, 18.24]$ °C	NOM (Jijel)
	Max Temperature	$[18.75, 24.25]$ °C	
	Min Temperature	$[9.25, 15.08]$ °C	
	Humidity	$[65.48, 69.11]$%	
	Wind Speed	$[2.86, 4.15]$ m/s	
Vegetation	NDVI	$[-0.11, 0.48]$	MODIS/Terra
Hydrological	Annual Precip.	$[48.92, 75.13]$ mm	NOM
	Dist. to Rivers	$[0, 11]$ km	Generated
Human Activity	Dist. to Roads	$[0, 3.1]$ km	Generated
	Dist. to Residential	$[0, 9.95]$ km	
Wildfire data	Historical Fires	/	Algerian Civil Protection

	Aspect	Elevation	Slope	Humidity	Max_Temperature	Min_Temperature	Avg_Temperature
0	192.81	311.902	17.3014	67.7474	21.6892	12.7412	15.7886
1	65.662	149.557	15.7934	68.5932	22.0966	13.2039	15.4267
2	241.269	198.156	17.8403	68.4722	21.923	12.9592	15.4385
3	170.184	6.6151	5.9992	68.5035	22.1388	13.3527	16.2818
4	189.945	337.386	18.6767	68.5014	22.0118	13.0871	15.47

	Precipitation	Wind_Speed	D_Rivers	NDVI	D_Rods	D_HumanActivity	D_FireOccurrence
0	62.2248	3.4627	0.0411	0.1129	0.0429	0.0423	9718.30
1	63.787	3.4872	0.0167	0.2635	0.0171	0.0165	3239.43
2	64.588	3.4402	0.0416	0.1992	0.0147	0.0148	4859.15
3	66.0397	3.5174	0.0055	0.1006	0.0034	0.003	0
4	64.2669	3.4531	0.0329	0.1608	0.0215	0.0215	4859.15

Figure 3.15: derived from GIS layers (2.6 M×14).

3.3.2 Data Preprocessing

The data preprocessing phase aimed to harmonize heterogeneous spatial layers and transform them into a consistent, analysis-ready dataset suitable for wildfire susceptibility modeling. All raster layers were first standardized onto a common spatial framework to ensure pixel-wise comparability across variables [127]. Specifically, all datasets were projected to the WGS 84 / UTM Zone 31N coordinate system, clipped to the administrative boundary of Jijel Province, and resampled to a uniform spatial resolution of 30 m Œ 30 m.

Continuous variables such as temperature, humidity, precipitation, wind speed, and NDVI were resampled using bilinear interpolation to preserve spatial continuity, while discrete layers and proximity-based variables (e.g., distance to roads, rivers, and residential areas) were resampled using nearest-neighbor interpolation to avoid artificial smoothing effects. NoData values were masked, and a landsea mask was applied along the coastal boundary to remove offshore pixels and ensure spatial consistency within the terrestrial study area [128].

Following spatial harmonization, basic data integrity checks were performed. The GIS-derived layers exhibited high completeness and consistency, and after clipping, masking, and resampling, the final datasetcomprising more than 2.6 million records and 14 conditioning factorscontained no missing or duplicated values. Minor spatial artifacts, such as isolated single-pixel anomalies introduced during raster processing, were removed through light de-noising operations. In addition, each variable was constrained to physically meaningful ranges to prevent the propagation of invalid values into the modeling stage [129].

To ensure the robustness of the dataset, multicollinearity among conditioning factors was evaluated prior to model construction. Multicollinearity can negatively affect susceptibility modeling by inflating parameter estimates and reducing model stability. The Variance Inflation Factor (VIF) and its reciprocal, tolerance, were employed to quantify interdependencies among predictors. For each factor, VIF measures the increase in variance caused by correlations with other variables, while tolerance represents the proportion of variance not explained by the remaining predictors. Commonly accepted thresholds (VIF > 10 or tolerance < 0.1) indicate problematic collinearity [129]. The analysis confirmed that all factors fell within acceptable limits, indicating that no variable exerted undue influence on the dataset.

Temporal consistency between predictors and wildfire occurrence was also considered. Climatic variables were aggregated to reflect summer conditions, corresponding to the primary wildfire season in Jijel. Temperature, humidity, and wind speed were represented using seasonal averages, while precipitation was aggregated over the same period. Vegetation condition, represented by NDVI, was composited for the summer window to reduce short-term noise and cloud contamination. In contrast, terrain derivatives and proximity-based anthropogenic factors were treated as time-invariant and used directly after rasterization [130].

Finally, all harmonized raster layers were transformed into numerical matrices and merged into a single spatial database using GIS-based geographic indexing. This process resulted in a comprehensive and consistent dataset that serves as the foundation for subsequent labeling and machine learning stages. All preprocessing steps were documented to ensure traceability and reproducibility from raw spatial inputs to the final analysis-ready dataset [131].

3.3.3 Dataset Labeling

After completing data preprocessing, the next step in building the wildfire susceptibility dataset consisted of labeling each spatial unit based on historical wildfire occurrence records. The objective of this stage was to convert heterogeneous spatial information into a robust categorical target suitable for supervised learning, while minimizing label noise and avoiding information leakage. Each pixel in the analysis grid was assigned to one of four wildfire susceptibility classes: Low, Medium, High, or Very High.

Labeling was performed using historical wildfire ignition records provided by the Algerian Civil Protection. These records were geocoded and verified using GIS tools to ensure spatial consistency. To reduce duplicate reporting and positional uncertainty, ignition points occurring within short spatial distances on the same day were consolidated into single representative locations, and obvious outliers (e.g., offshore or incorrectly positioned points) were removed. Label construction was restricted to the active wildfire season and to a predefined temporal window used exclusively for training, while more recent fire occurrences were reserved for independent validation [132].

To represent the spatial propensity of wildfire occurrence, two complementary indicators were derived from the historical records: (i) distance to previous wildfire ignitions and (ii) wildfire density. A smoothed fire-density surface was generated using kernel density estimation (KDE) on the ignition points, producing a continuous raster that captures persistent spatial patterns of fire activity. In parallel, a distance-to-nearest-fire raster was computed to reflect the localized influence of historical ignition corridors. These two surfaces were combined into a composite susceptibility score that integrates both local proximity and broader spatial persistence of wildfire occurrence [133].

The resulting continuous susceptibility surface was then classified into four ordinal classes using calibrated thresholds. The Very High class corresponds to areas with the highest concentration of historical ignitions, while High and Medium classes represent progressively lower but still elevated propensities. The Low class includes areas with minimal historical fire influence. Thresholds were selected to ensure stable spatial patterns and meaningful class prevalence across the study area.

While the Medium, High, and Very High classes are directly supported by historical evidence, labeling the Low susceptibility class presents an inherent challenge. Some areas may share similar environmental and anthropogenic characteristics with high-risk zones but remain unburned during the observation period due to chance rather than true low susceptibility. Such cases introduce label noise that can negatively affect model learning [134].

To mitigate this issue, an unsupervised refinement step was applied specifically to the Low class. K-means clustering ($k = 2$) was performed in the 14-factor feature space on paired subsets (Very HighLow, HighLow, and MediumLow). Low-susceptibility samples that clustered closer to higher-risk centroids in any pairing were flagged as ambiguous and removed from the training dataset. This procedure eliminated approximately 1.9% of Low class samples. Sensitivity analyses using different initialization counts, random seeds, and light feature reweighting produced nearly identical results, confirming the robustness of this denoising step [135].

The final labeled dataset thus represents a balanced compromise between historical evidence and environmental similarity, reducing the risk of teaching models that clearly fire-prone configurations correspond to low susceptibility. The four classes were encoded as an ordinal target with integer labels 1, 2, 3, 4, corresponding to Low, Medium, High, Very High. Class imbalance was addressed during model training using stratified sampling and class-weighted loss functions rather than synthetic oversampling, thereby preserving the spatial structure of the data.

Overall, this labeling strategy provides a defensible and interpretable target for wildfire susceptibility modeling. It integrates local and regional fire history, limits noise in the Low class, respects temporal separation between predictors and labels, and enhances the generalization capacity of the subsequent learning models. Figure 3.16 illustrates the clustering-assisted refinement applied during the labeling process.

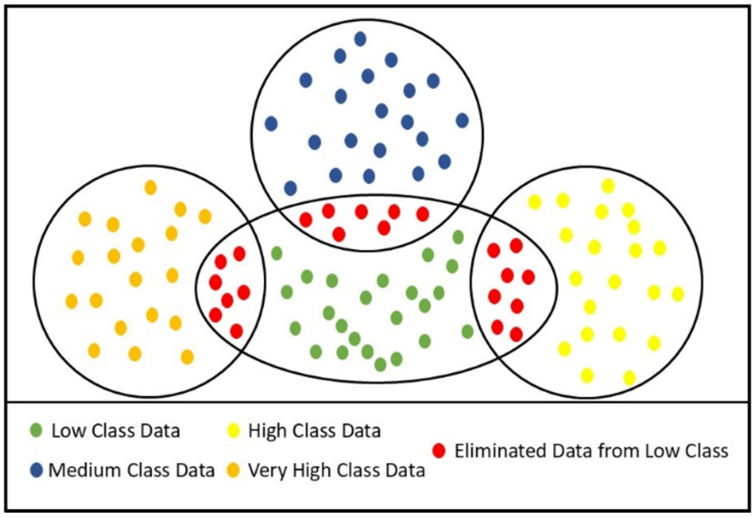

Figure 3.16: Clustering-assisted refinement of labels.

3.4 Conclusion

This chapter presented the complete dataset preparation pipeline adopted for wildfire susceptibility modeling in the study area. Starting from multi-source data collection, the workflow ensured the integration of environmental and anthropogenic factors into a unified spatial database comprising over 2.6 million records and 14 conditioning variables. Rigorous preprocessing steps confirmed the high quality of the extracted data and addressed potential issues such as multicollinearity, thereby safeguarding the reliability of subsequent analyses.

A carefully designed labeling strategy was then introduced to generate a four-class wildfire susceptibility target. By combining distance-based information with historical fire density and applying clustering-assisted refinement to reduce noise in the low-susceptibility class, the resulting labels achieved both interpretability and robustness. The adopted procedures minimized information leakage, reduced class ambiguity, and improved the learning conditions for supervised models.

Overall, this chapter established a strong and defensible dataset that accurately reflects wildfire dynamics in the region. The resulting data foundation enables effective model training and fair performance evaluation, which are presented and discussed in the following chapters.

Chapter 4

Proposed Methodology and Experimental Framework

4.1 Introduction

This chapter presents the methodological backbone of our wildfire susceptibility mapping framework and explains how it addresses the main limitations identified in the related work. Prior studies often face four recurring challenges: inconsistent and weakly justified factor selection; fragmentation of multi-source data with heterogeneous scales and projections; label noise and information leakage that inflate apparent model skill; and an emphasis on accuracy metrics with limited interpretability and scarce out-of-period validation. Our architecture is designed to meet these challenges head-on by linking defensible factor curation with a reproducible GIS-to-ML pipeline, careful label construction and denoising, calibrated learning under spatially aware validation, independent time-forward testing, and interpretable outputs delivered as actionable maps.

4.2 Methodology Workflow

Following the construction of the wildfire susceptibility dataset described in the previous chapter, the methodological workflow focuses on analyzing this dataset and extracting reliable, interpretable predictive knowledge. The overall objective is to evaluate the capability of selected machine learning (ML) methods to model wildfire susceptibility and to understand the driving factors behind their predictions [16].

Once the dataset was finalized and labeled, the first step of this chapter involved exploratory analysis of the data to understand class distributions, feature behavior, and potential learning challenges such as imbalance and nonlinearity. Based on an extensive review of recent wildfire susceptibility studies, a set of machine learning algorithms known for their strong performance in similar contexts was selected. These models represent different learning paradigms, enabling a balanced comparison between tree-based ensembles, kernel-based methods, and neural models.

The selected ML models were then trained using the prepared dataset under consistent experimental conditions. Appropriate training strategies, including stratified sampling and class weighting, were applied to account for class imbalance. Hyperparameters were optimized to ensure fair performance comparison, and the same evaluation protocol was used across all models. Model outputs were assessed using multiple performance metrics to capture both discrimination ability and classification reliability.

After model training, a comparative analysis was conducted to evaluate the relative strengths and weaknesses of each method. This comparison aimed not only to identify the best-performing model but also to understand how different algorithms respond to the same environmental and anthropogenic conditioning factors.

Beyond performance evaluation, particular attention was given to model interpretation and explanation. Since wildfire susceptibility mapping is intended to support decision-making, it is essential that model predictions are explainable and scientifically defensible. Therefore, interpretation techniques were applied to analyze feature contributions and to link model behavior with known wildfire-driving mechanisms in the study area.

Figure 4.1 illustrates the complete methodological workflow, from dataset utilization to modeling, comparison, and interpretation. The following sections provide detailed descriptions of the selected machine learning models, evaluation metrics, and interpretability methods employed in this study.

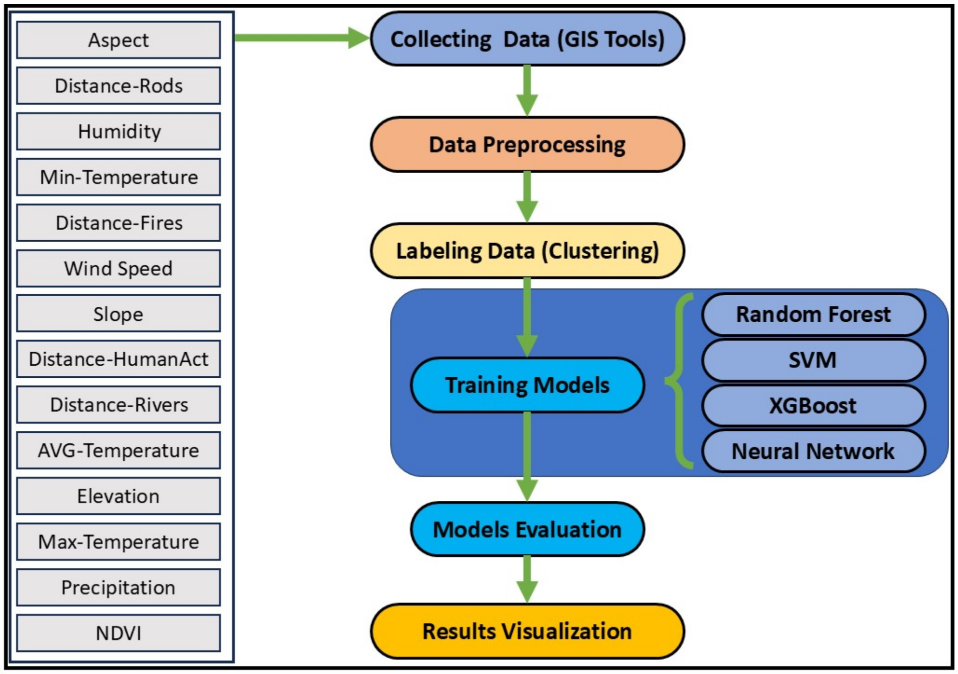

Figure 4.1: End-to-end framework for wildfire susceptibility mapping: from multi-source data fusion and defensible labeling to calibrated learning, independent validation, interpretability, and GIS delivery.

4.3 Theoretical Background

To model wildfire susceptibility, we employ a set of widely used and well-established classifiers, namely Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XG-Boost), and Artificial Neural Networks (NN), which have demonstrated strong performance in previous wildfire prediction and environmental susceptibility studies. These methods were selected for their complementary learning mechanisms, robustness to nonlinear relationships, and ability to handle high-dimensional spatial data. Model performance is assessed using standard and widely accepted evaluation metrics, including the Area Under the Receiver Operating Characteristic Curve (AUC) and the F1-score, which jointly capture ranking ability and class-wise discrimination under class imbalance. In addition, cross-validation is applied to ensure model stability and generalization, particularly in the presence of spatial heterogeneity inherent to wildfire susceptibility mapping.

4.3.1 Machine Learning Methods for WSM

4.3.1.1 Random Forest

Random Forest (RF) is an ensemble machine learning method that constructs a large number of decision trees during the training phase and aggregates their predictions to produce a final output. In classification problems such as wildfire susceptibility mapping (WSM), the final class is typically determined by majority voting or by averaging class probabilities across all trees [57]. The core idea behind RF is to reduce model variance and improve generalization by combining multiple weak learners that are trained on different bootstrap samples of the data and different subsets of features [136].

For a given sample with predictor vector $\mathbf{x}$ and class c, the estimated class probability produced by RF is defined as:

$$\hat{p}(c \mid \mathbf{x}) = \frac{1}{n_{\text{trees}}} \sum_{t=1}^{n_{\text{trees}}} p_t(c \mid \mathbf{x}), \tag{4.1}$$

where $p_t(c \mid \mathbf{x})$ denotes the class probability estimated by the t-th decision tree. Each tree is grown by recursively splitting the data to maximize node purity, commonly measured using the Gini impurity criterion:

$$G = \sum_{k} p_k(1 - p_k), \tag{4.2}$$

where p_k represents the proportion of samples belonging to class k within a node. The optimal split is selected by minimizing the weighted impurity after the split.

RF offers several advantages that make it particularly suitable for environmental and wildfire-related applications. First, it can effectively handle large datasets with high dimensionality and complex nonlinear relationships, which are typical characteristics of wildfire conditioning factors such as topography, climate, vegetation indices, and human-related variables [137]. Second, the use of bootstrap aggregation and random feature selection at each split enhances robustness against overfitting, even when predictors are noisy or partially correlated. This property is especially valuable in WSM, where environmental variables often exhibit strong interdependencies. Moreover, RF naturally accommodates mixed data types and does not require strict assumptions about data distributions or feature scaling [137].

Numerous studies have demonstrated the strong performance of RF in wildfire prediction and susceptibility mapping. For example, RF has been successfully applied to model fire occurrence using meteorological variables such as temperature, precipitation, wind speed, and humidity, achieving low prediction errors after hyperparameter tuning [138]. Other studies integrating topographic, climatic, and vegetation-related factors (e.g., elevation, NDVI, precipitation, land surface temperature) reported that RF outperformed single-tree models by providing higher accuracy and more nuanced spatial risk patterns. These findings highlight the models ability to capture complex environmental interactions and produce reliable susceptibility maps that are useful for operational fire management and ecological conservation [138].

In addition to predictive performance, RF provides several tools that enhance interpretability, which is a critical requirement in environmental decision-making. Feature importance measures, such as mean decrease in impurity and permutation importance, allow identification of the most influential wildfire drivers. Partial dependence plots further help in understanding the marginal effect of individual factors while averaging out the influence of others. These properties support transparent interpretation of model outputs and facilitate communication with stakeholders and policymakers [139].

Despite its strengths, RF also has certain limitations. The model can become computationally expensive and memory-intensive when a large number of trees are used, particularly with high-resolution spatial datasets. RF predictions may exhibit stepwise probability surfaces due to the tree-based structure, which can require additional smoothing or calibration for operational mapping. Furthermore, RF generally performs poorly when extrapolating beyond the range of the training data and may distribute importance across correlated predictors, making interpretation less straightforward unless grouped or conditional importance measures are applied [139].

Random Forest represents a robust and widely validated machine learning approach for wildfire susceptibility mapping. Its balance between predictive accuracy, resistance to overfitting, and interpretability makes it particularly well suited for modeling wildfire risk in complex environmental systems, where interactions between climatic, topographic, vegetation, and anthropogenic factors play a dominant role.

4.3.1.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm grounded in statistical learning theory and is widely used for classification and regression problems. In wildfire susceptibility mapping (WSM), SVM aims to identify an optimal decision boundary that maximizes the separation margin between classes representing different levels of fire risk. Its strong theoretical foundation and ability to model complex, nonlinear relationships make it a popular choice for environmental hazard prediction [140].

Given a training dataset $(\mathbf{x}i, y_i)i = 1^N$, where $\mathbf{x}i \in \mathbb{R}^d$ denotes the feature vector and $y_i \in -1, +1$ the class label, SVM seeks to find a hyperplane defined by $(\mathbf{w}, b)$ that maximizes the margin while minimizing classification error. This optimization problem is formulated as:

$$\min \mathbf{w}, b, \xi \ \frac{1}{2}|\mathbf{w}|^2 + C \sum_{i=1}^{N} \xi_i \tag{4.3}$$

subject to

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \tag{4.4}$$

where ξ_i are slack variables allowing misclassification and C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing training errors.

To handle nonlinear separability, SVM employs kernel functions that implicitly map the input data into a higher-dimensional feature space. Commonly used kernels in wildfire studies include the radial basis function (RBF), polynomial, and sigmoid kernels. The RBF kernel, defined as

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma|\mathbf{x}_i - \mathbf{x}_j|^2\right), \tag{4.5}$$

is particularly effective for capturing nonlinear interactions among environmental variables such as temperature, humidity, vegetation indices, and terrain attributes [140].

SVM presents several advantages for wildfire prediction and susceptibility modeling. It is highly effective in high-dimensional feature spaces and performs well even when the number of predictors exceeds the number of observations. The margin maximization principle provides good generalization capability, making SVM relatively resistant to overfitting when properly regularized [141]. In environmental applications, SVM has demonstrated strong predictive performance using heterogeneous data sources, including meteorological parameters, vegetation characteristics, topography, and socio-economic indicators. Studies conducted in regions such as China, India, and Portugal reported classification accuracies exceeding 75% and AUC values above 0.84, confirming the suitability of SVM for large-scale wildfire risk mapping [142].

However, SVM also exhibits several limitations in the context of wildfire susceptibility mapping. The method is sensitive to the choice of kernel and hyperparameters (e.g., C and γ), which requires careful tuning and cross-validation. Conventional SVM implementations may struggle with very large datasets, both in terms of computational time and memory usage, especially when using nonlinear kernels [143]. This limitation has motivated the development of parallel and distributed SVM variants using frameworks such as SPARK and PySpark, which significantly reduce computation time and storage requirements while preserving predictive accuracy. Additionally, SVM models tend to be less interpretable than tree-based approaches, as the influence of individual variables on the decision boundary is not directly observable [144].

From an operational perspective, SVM has shown higher efficiency in predicting small to moderate fire events but may exhibit reduced sensitivity to large-scale fire occurrences unless trained on sufficiently balanced datasets. Class imbalance, which is common in wildfire data, can further bias SVM toward dominant classes if not adequately addressed through class weighting or sampling strategies [145].

In summary, Support Vector Machine constitutes a powerful and theoretically sound approach for wildfire susceptibility mapping, particularly when modeling nonlinear relationships among diverse environmental and anthropogenic factors. While its computational demands and limited interpretability pose challenges for large-scale, high-resolution studies, advancements in parallel computing and careful model tuning make SVM a valuable component of comparative WSM frameworks, complementing ensemble and deep learning methods.

4.3.1.3 Neural Networks (NN)

Neural Networks (NN), inspired by the structure and functioning of the human brain, are powerful machine learning models capable of capturing complex, nonlinear relationships among multiple variables. In wildfire susceptibility mapping (WSM), NNs are particularly suitable due to the inherently nonlinear interactions between climatic conditions, vegetation characteristics, topography, and human activities that govern wildfire occurrence [133].

The most commonly used NN architecture in environmental modeling is the multilayer perceptron (MLP), which consists of an input layer, one or more hidden layers, and an output layer. Each neuron computes a weighted sum of its inputs followed by a nonlinear activation function. For a given neuron, the output is expressed as:

$$z = f\left(\sum_{i=1}^{d} w_i x_i + b\right), \tag{4.6}$$

where x_i are the input features, w_i the associated weights, b the bias term, and $f(\cdot)$ a nonlinear activation function such as ReLU, sigmoid, or tanh. For classification problems, the output layer commonly employs a softmax or sigmoid function to estimate class probabilities.

Training NNs involves minimizing a loss function, such as cross-entropy for classification, through iterative optimization using backpropagation and gradient-based algorithms. The backpropagation process updates the network weights by propagating the prediction error backward through the layers, allowing the model to learn hierarchical feature representations. In deep neural networks (DNNs), multiple hidden layers enable the automatic extraction of high-level abstractions from raw input data, reducing the need for extensive manual feature engineering [146].

Neural Networks offer several advantages for wildfire susceptibility assessment. Their ability to model highly nonlinear and multivariate relationships makes them well suited for complex environmental systems. Recent studies have demonstrated that optimized deep neural networks can achieve very high predictive performance in wildfire risk mapping, with reported AUC values exceeding 0.95 in regions such as Australia [147]. NN-based frameworks are also highly adaptable; once trained, they can be transferred to new geographic regions with minimal adjustment, particularly when using standardized input layers and region-specific weighting schemes. Advanced NN architectures, such as sparse autoencoders, have been shown to improve feature learning and enhance predictive accuracy, especially in the presence of noisy or redundant variables [148].

Furthermore, Neural Networks can effectively integrate heterogeneous data sources, including meteorological variables, vegetation indices, topographic features, and socio-economic indicators. This flexibility allows NNs to capture subtle interactions and thresholds that are often difficult to represent using traditional statistical or rule-based models. When combined with data balancing strategies, NNs have also demonstrated improved performance in predicting rare but high-impact wildfire events, which are typically underrepresented in historical datasets [149].

Despite these strengths, Neural Networks present notable challenges in wildfire susceptibility modeling. They are sensitive to the quality and quantity of training data, and their performance can be significantly affected by noise, missing values, and class imbalance. Hyperparameter tuning, including the selection of network depth, number of neurons, learning rate, and regularization terms, is computationally demanding and can introduce uncertainty if not carefully managed. Overfitting remains a critical concern, particularly when deep architectures are trained on limited or spatially correlated data [150].

Another major limitation of NN models is their limited interpretability. Unlike tree-based methods, NNs are often considered black-box models, making it difficult to directly assess the influence of individual factors on wildfire susceptibility. This limitation can reduce their acceptance in operational wildfire management contexts, where transparency and explainability are essential for decision-making. Recent advances in explainable AI (XAI), such as SHAP and sensitivity analysis, partially address this issue but add further computational complexity [151].

Neural Networks represent a highly flexible and powerful approach for wildfire susceptibility mapping, capable of achieving superior predictive performance by capturing complex nonlinear relationships across diverse environmental factors. While their computational cost, sensitivity to data quality, and limited interpretability pose challenges, carefully designed and optimized NN frameworks can provide valuable insights and high-accuracy susceptibility maps, particularly when used alongside more interpretable machine learning models in a comparative or ensemble setting [152].

4.3.1.4 Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is an advanced ensemble learning algorithm based on the gradient boosting decision tree (GBDT) framework. It builds predictive models by sequentially adding decision trees, where each new tree is trained to correct the residual errors of the previous ensemble. Due to its high predictive accuracy, scalability, and built-in regularization mechanisms, XGBoost has become one of the most widely used machine learning methods for wildfire susceptibility mapping (WSM) and related environmental risk assessments [153].

Given a dataset $(\mathbf{x}i, y_i)i = 1^N$, XGBoost models the prediction as an additive function of K regression trees:

$$\hat{y}i = \sum k = 1^K f_k(\mathbf{x}i), \quad f_k \in \mathcal{F}, \tag{4.7}$$

where $\mathcal{F}$ denotes the space of decision trees. The model is optimized by minimizing a regularized objective function:

$$\mathcal{L} = \sum i = 1^N \ell(y_i, \hat{y}i) + \sum k = 1^K \Omega(f_k), \tag{4.8}$$

where $\ell(\cdot)$ is a differentiable loss function (e.g., logistic loss for classification), and $\Omega(f_k)$ is a regularization term that penalizes model complexity, typically defined as:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \tag{4.9}$$

with T being the number of leaves in the tree, w_j the leaf weights, and γ and λ regularization parameters. This formulation enables XGBoost to control overfitting while maintaining high learning capacity [154].

One of the key advantages of XGBoost in wildfire susceptibility modeling lies in its ability to capture complex nonlinear relationships and high-order interactions among environmental and anthropogenic factors, such as temperature, vegetation indices, slope, proximity to roads, and human settlements. Its tree-based structure allows it to naturally handle mixed data types and missing values, which are common in large-scale geospatial datasets. Moreover, XGBoost is computationally efficient due to parallel processing, cache-aware access, and optimized tree construction, making it suitable for large raster-based susceptibility datasets [150].

Several studies have demonstrated the strong performance of XGBoost in fire risk and susceptibility assessment. In heritage and urban fire risk studies, XGBoost has been shown to outperform other ensemble and boosting methods, including GBDT, AdaBoost, CatBoost, and SVM, particularly when historical fire records are sparse [155]. By integrating spatially unique features and risk indices, XGBoost has proven effective in predicting fire risk levels at fine temporal scales, such as monthly assessments. These characteristics are directly transferable to wildfire susceptibility mapping, where fire occurrence data are often limited or unevenly distributed in space and time [141].

Another major strength of XGBoost is its compatibility with explainable artificial intelligence (XAI) frameworks. Feature attribution methods such as SHAP (SHapley Additive exPlanations) can be seamlessly applied to XGBoost models, providing both global and local interpretations of factor contributions [156]. This capability is particularly valuable in environmental and wildfire studies, as it allows researchers and decision-makers to understand why certain areas are classified as high-risk and to relate model outputs to physical drivers and management actions [146].

Despite its advantages, XGBoost has some limitations. The algorithm requires careful hyperparameter tuning, including learning rate, tree depth, subsampling ratios, and regularization terms, to achieve optimal performance. Inadequate tuning may lead to overfitting or overly complex models. Additionally, while XGBoost is more interpretable than deep neural networks, its boosted ensemble structure can still be less transparent than single decision trees, especially when many trees are involved [141].

In summary, XGBoost offers a powerful and flexible framework for wildfire susceptibility mapping by combining high predictive accuracy, robustness to data imperfections, and advanced regularization techniques. Its ability to integrate spatial features, handle sparse fire occurrence data, and support explainable modeling makes it particularly well suited for operational wildfire risk assessment and decision support systems in complex environmental settings [157].

4.3.2 Evaluation Methods

To objectively assess and compare the performance of the applied machine learning models in wildfire susceptibility mapping, three complementary evaluation metrics were employed: the Area Under the Receiver Operating Characteristic Curve (AUC), the F1 score, and cross-validation. These metrics were selected to capture different aspects of model performance, including discrimination capability, classification balance, and generalization robustness.

4.3.2.1 Area Under the Curve (AUC)

The Area Under the Curve (AUC) is a threshold-independent metric that measures a models ability to discriminate between different susceptibility classes. It summarizes the Receiver Operating Characteristic (ROC) curve, which represents the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR) across all possible decision thresholds. AUC values range from 0 to 1, where higher values indicate stronger discriminatory power and better overall model performance [158].

The AUC is defined as:

$$AUC = \int_0^1 TPR(FPR)\, d(FPR) \tag{4.10}$$

where TPR (True Positive Rate) represents the proportion of correctly identified positive samples, and FPR (False Positive Rate) denotes the proportion of incorrectly classified negative samples [158]. In wildfire susceptibility studies, AUC is particularly valuable because it evaluates ranking quality independently of class thresholds.

4.3.2.2 F1 Score

The F1 score is the harmonic mean of precision and recall, providing a single metric that balances omission and commission errors. It is especially suitable for wildfire susceptibility mapping, where class imbalance is common and high-risk fire areas typically occupy a smaller spatial extent than low-risk areas. The F1 score ranges from 0 to 1, with higher values indicating better classification performance [159].

The F1 score is computed as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4.11}$$

where precision measures the proportion of correctly predicted positive instances, and recall measures the proportion of actual positives that are correctly identified [157].

4.3.2.3 Cross-Validation

Cross-validation is a resampling technique used to evaluate how well a model generalizes to unseen data. It reduces the risk of overfitting by ensuring that model performance is not dependent on a single train–test split. In this study, k-fold cross-validation was applied, where the dataset is partitioned into k equally sized subsets [160].

The general procedure of k-fold cross-validation is as follows:

- Divide the dataset into k equally sized folds.
- For each fold:
  - Train the model using $k - 1$ folds.
  - Validate the model on the remaining fold.
- Compute the average performance across all k folds.

4.4 Workflow Implementation and Modeling Procedures

After data collection and the application of preprocessing and labeling procedures described in the previous chapter, the analysis proceeded to the modeling stage. The finalized labeled dataset was first partitioned into two subsets to enable model training and evaluation. Specifically, 70% of the data were allocated for training, while the remaining 30% were reserved for validation. To preserve the original distribution of wildfire susceptibility classes and avoid class imbalance bias, the split was performed using stratified sampling. This strategy ensured that each susceptibility class (low, medium, high, and very high) was proportionally represented in both subsets, providing a reliable basis for learning model parameters and assessing predictive performance.

Random Forest (RF)

The Random Forest (RF) model was applied to the preprocessed wildfire susceptibility dataset to predict the four-class labels. The dataset was loaded, and unnecessary columns, such as identifiers or GIS coordinates used only for spatial referencing, were removed. Predictor variables, including topographical, climatic, vegetation, and anthropogenic factors, were assigned to X, while the wildfire susceptibility class was assigned to y. The target variable was encoded to integer labels to ensure compatibility with the classifier.

The dataset was then split into training and testing sets using a 70/30 stratified partition to preserve the distribution of susceptibility classes. The RF model was trained on the training subset, constructing an ensemble of decision trees with bootstrap sampling and random feature selection at each split.

Parameters:

- Number of trees (‘n_estimators‘) = 100
- Random seed (‘random_state‘) = 42 for reproducibility
- Split criterion (‘criterion‘) = Gini impurity
- Maximum depth of trees (‘max_depth‘) = None (trees grow until all leaves are pure)
- Minimum samples to split a node (‘min_samples_split‘) = 2
- Minimum samples per leaf (‘min_samples_leaf‘) = 1
- Maximum features considered for splitting (‘max_features‘) = sqrt(total features)
- Bootstrap sampling (‘bootstrap‘) = True

After training, predictions were generated for the testing subset, and model performance was evaluated using accuracy, F1-score, and AUC, see Figure 4.2.

Algorithm Random Forest Training Algorithm

1. Load dataset
2. Remove unnecessary columns
3. Split dataset into features X and target y
4. Encode target classes (if categorical)
5. Split X and y into training and testing sets
6. Train Random Forest model on training set
7. Predict classes for testing set
8. Evaluate model accuracy

Figure 4.2: Algorithm: Random Forest Training

Neural Network (MLP)

The neural network model was applied to the processed wildfire susceptibility dataset to predict the four-class labels. Predictor variables were assigned to X, while the wildfire susceptibility class was assigned to y, and the target classes were numerically encoded. The dataset was then split into training and testing subsets using a 70/30 stratified partition to maintain class proportions. Feature values were normalized using standard scaling prior to training.

Parameters and architecture:

- Input layer = number of features
- Hidden layer 1 = 64 neurons, ReLU activation
- Hidden layer 2 = 32 neurons, ReLU activation
- Output layer = 4 neurons, Softmax activation
- Loss function = categorical cross-entropy
- Optimizer = Adam
- Batch size = 32
- Number of epochs = 100
- Validation split = 20% of training set

The network was trained on the training set, predictions were generated on the testing set, and model performance was evaluated using accuracy, F1-score, and AUC, see Figure 4.3.

Algorithm Neural Network Model

Require: Dataset maindataaf.csv

Output: Predicted class labels

1: Load dataset (maindataaf.csv)
2: Remove unnecessary columns (Unnamed: 0, firej, Cluster)
3: Split dataset into features X and target y (**fireclasse**)
4: Encode target classes to numerical labels (if categorical)
5: Split X and y into training and testing sets
6: Normalize features using standard scaling
7: Define neural network architecture (input layer, hidden layers, output layer with softmax)
8: Compile model with optimizer, loss function, and metrics
9: Train model on training set with validation split
10: Evaluate model accuracy on testing set
11: Predict classes for testing set
12: Convert predicted class indices back to original labels

Figure 4.3: Algorithm: Neural Network Training steps

XGBoost

The XGBoost classifier was applied to the processed dataset following a similar workflow. Predictor variables were assigned to X, the target variable y was numerically encoded, and the dataset was split into 70/30 training/testing subsets using stratified sampling.

Parameters:

- Objective = ‘multi:softmax‘ (for multi-class classification)
- Number of classes (‘num_class‘) = 4
- Evaluation metric (‘eval_metric‘) = ‘mlogloss‘
- Use label encoder (‘use_label_encoder‘) = False
- Random state = 42
- Test size = 0.3 (30% for testing)

The model was trained on the training set, predictions were made on the testing set, and performance was evaluated using accuracy, F1-score, and AUC, see Figure 4.4.

Algorithm XGBoost Model Training Steps

Require: Preprocessed dataset ready for training

Ensure: Model accuracy, classification report, AUC, F1 score

1. Load dataset
2. Remove unnecessary columns
3. Rename columns if needed and strip whitespace
4. Split dataset into features X and target y
5. Encode target classes (if categorical)
6. Split X and y into training and testing sets
7. Initialize XGBoost classifier with appropriate parameters
8. Train XGBoost model on training set
9. Predict classes for testing set
10. Evaluate model accuracy, classification report, AUC, F1 score

Figure 4.4: Algorithm: XGBoost Training steps

Support Vector Machine (SVM)

The SVM model was applied to the dataset using a consistent workflow. Predictor variables were assigned to X, the target variable y was numerically encoded, and the dataset was split using a

70/30 stratified partition.

Parameters:

- Kernel = RBF (default)
- C (regularization) = 1.0
- Gamma = ‘scale‘
- Multi-class strategy = One-vs-Rest (OvR)
- Random state = 42 (for reproducibility in train-test split)

The model was trained on the training subset, predictions were made on the testing set, and performance was evaluated using accuracy, F1-score, and AUC, see Figure 4.5.

Algorithm SVM Model Training Steps

```
Load dataset
Remove unnecessary columns
Rename columns if needed and strip whitespace
Split dataset into features X and target y
Encode target classes (if categorical)
Split X and y into training and testing sets
Initialize SVM classifier with appropriate parameters
Train SVM model on training set
Predict classes for testing set
Evaluate model accuracy, classification report, AUC, F1 score
```

Figure 4.5: Algorithm: SVM Training steps

4.5 Conclusion

In this chapter, we detailed the complete workflow for wildfire susceptibility mapping, starting from dataset preprocessing and labeling, through the application of multiple machine learning models, to evaluation and interpretation. The rigorous preparation of a high-quality, 2.6 millionentry dataset ensured reliable model inputs, while stratified training and validation allowed fair assessment of model performance. Random Forest, SVM, XGBoost, and Neural Networks were applied and compared using standard metrics (AUC, F1 score, and cross-validation), with results visualized both as charts and spatial maps for expert analysis.

Independent validation using 20242025 fire occurrences demonstrated the generalization capability of the selected model, and overlaying predicted high-risk areas on feature-specific maps provided valuable insights into the link between environmental and anthropogenic factors and wildfire occurrence. This methodology establishes a robust, interpretable, and actionable framework for wildfire risk assessment, supporting both operational decision-making and further research on factor-driven fire dynamics.

Chapter 5

Results and Discussion

5.1 Introduction

This section presents and discusses the results of the wildfire susceptibility modeling framework developed for the study area. The analysis goes beyond reporting predictive accuracy and focuses on understanding model behavior, spatial consistency, and real-world relevance. First, multiple machine learning models are compared using standard performance metrics and ROC analysis to identify the most reliable predictors of wildfire occurrence. Second, susceptibility maps are examined spatially and contrasted with environmental and anthropogenic factors to assess physical realism. Third, model interpretability is addressed through SHAP-based feature importance and spatial overlays, enabling an explicit link between predictors, risk patterns, and underlying ignition mechanisms. Finally, the robustness of the selected model is evaluated using independent wildfire occurrences from 20242025, providing a time-forward validation that strengthens confidence in the results. Together, these analyses ensure that the proposed framework delivers not only high statistical performance but also transparent, spatially coherent, and decision-relevant wildfire risk information.

5.2 Data Treatment Results: Correlation Structure and Multicollinearity

Before constructing predictive models for wildfire susceptibility, it is essential to examine the relationships among the predictor variables to ensure their independence and reduce potential multicollinearity. Multicollinearity occurs when independent variables are highly correlated with one another, which can inflate the variance of coefficient estimates and reduce the interpretability and stability of the model. To address this, two complementary analyses were conducted: the Variance

Inflation Factor (VIF) table and the Pearson correlation matrix. Together, these provide both a global view of multicollinearity and the specific pairwise relationships between predictors.

The VIF table (Table 5.1) indicates that most variables in the dataset have acceptable levels of multicollinearity, confirming that the dataset is suitable for initial model training. VIF measures the extent to which the variance of a regression coefficient is increased due to correlations with other variables, while Tolerancethe inverse of VIFrepresents the proportion of variability in a predictor that is not explained by other predictors. In this study, the highest VIF values were observed for Wind-Speed (7.58) and Average Temperature (6.27).

Wind-Speeds low tolerance of 0.11 suggests that a substantial portion of its variance is explained by other features, particularly topographic and hydrological variables. Similarly, Average Temperature, which is derived from the maximum and minimum temperature, shows considerable overlap with other climatic variables, explaining its elevated VIF. Variables such as Max-Temperature, Min-Temperature, and Humidity also exhibit moderate VIF values between 2.6 and 2.9, reflecting their partial redundancy but still remaining within acceptable thresholds for environmental modeling. On the other hand, NDVI and Aspect demonstrate very low VIFs (1.02 and 1.10, respectively), highlighting their independence from other predictors and their unique contribution to the model.

Table 5.1: VIF and tolerance values.

Feature	**VIF**	**Tolerance**
Aspect	1.10	0.90
Elevation	2.29	0.43
Slope	1.39	0.71
Humidity	2.63	0.37
Max-Temperature	2.76	0.34
Min-Temperature	2.90	0.16
Avg-Temperature	6.27	0.43
Precipitation	2.28	0.44
Wind-Speed	7.58	0.11
Distance to Rivers	2.04	0.48
NDVI	1.02	0.97
Distance to Human Activity	2.01	0.16
Distance to Roads	2.08	0.17

The correlation matrix further clarifies the relationships between variables, especially in relation to wildfire occurrence, Figure 5.1. The last column of the matrix, corresponding to D-

FireOccurrence, reveals that human activity is the primary driver of fire events in the study area, with a high positive correlation of 0.78. Elevation also shows a strong positive relationship with fire occurrence, suggesting that higher altitudes, possibly due to specific vegetation types or reduced accessibility for firefighting, are more prone to fires.

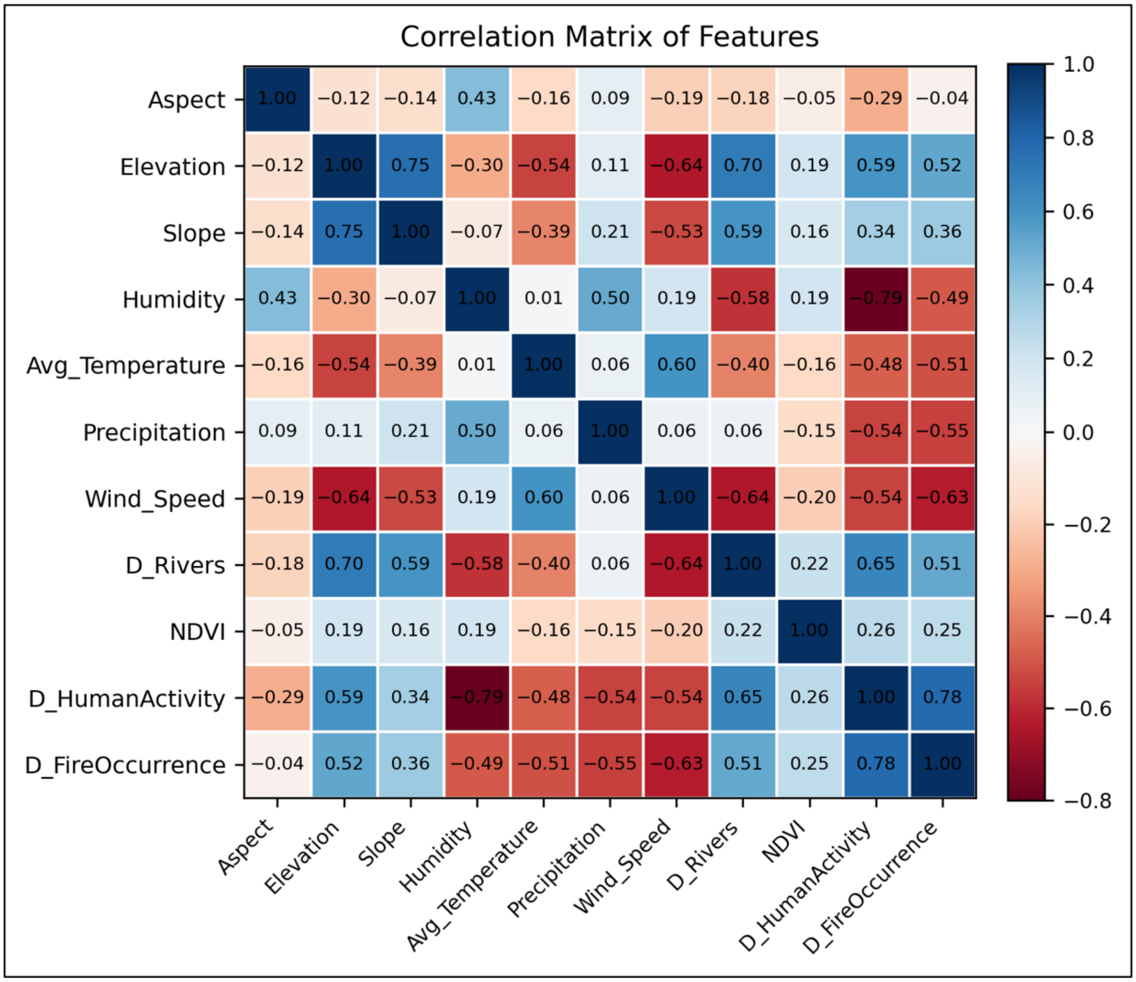

Figure 5.1: Correlation matrix across conditioning factors.

Environmental variables such as Wind-Speed and Precipitation exhibit negative correlations with fire occurrence, indicating that higher winds and rainfall reduce wildfire risk, likely through moisture retention and climatic moderation. The matrix also identifies clusters of correlated predictors, revealing underlying redundancy. For instance, Elevation and Slope (r = 0.75) form a terrain cluster, while Distance to Human Activity, Distance to Rivers, and Humidity show strong interrelations, reflecting the tendency of human settlements to occur in specific geographic and climatic zones. Interestingly, NDVI maintains very low correlations with all other variables, confirming its role as an independent indicator of vegetation conditions that contributes unique information to the model.

By linking the correlation matrix with the VIF table, the sources of multicollinearity become clearer. High VIF values for Wind-Speed and Average Temperature arise from their strong correlations with topographic and climatic features. Wind-Speed, for example, is closely linked to Elevation and Distance to Rivers, which explains its elevated VIF, while Average Temperatures relationship with other temperature variables and geographic factors accounts for its redundancy. Conversely, NDVI and Aspect remain largely independent, providing valuable, orthogonal information for wildfire prediction. Human-related factors, such as Distance to Human Activity, show moderate VIF values due to their overlap with environmental features, highlighting the interplay between human settlement patterns and the surrounding geography.

the analysis confirms that the primary training dataset is stable and suitable for model development. While certain variables exhibit multicollinearity, none exceed critical thresholds, and the independent predictors offer strong, unique contributions. These results provide confidence in the datasets robustness and should guide further model refinement, such as considering variable transformations or selective feature reduction, to enhance predictive accuracy in subsequent analyses.

5.3 Methods, Accuracy, and Comparative Performance

The ROC curves (Figure 5.2) illustrate the discriminative power of the models across all classification thresholds. The Random Forest model achieved a perfect AUC of 1.00, indicating near-ideal separation between fire and non-fire locations within the dataset. Gradient boosting approachesXG-Boost, LightGBM, and CatBoostclosely followed, each attaining an AUC of 0.99, confirming the effectiveness of ensemble learning methods for wildfire susceptibility mapping. The ROC curves for all models cluster near the top-left corner of the ROC space, representing optimal classifier behavior. This outcome demonstrates that the selected predictors, particularly Distance to Human Activity and Maximum Temperature, provide strong and distinguishable signals for wildfire occurrence. Although a perfect AUC can raise concerns regarding overfitting, subsequent forward-looking validation using independent wildfire events confirms that the models generalize well beyond the training data.

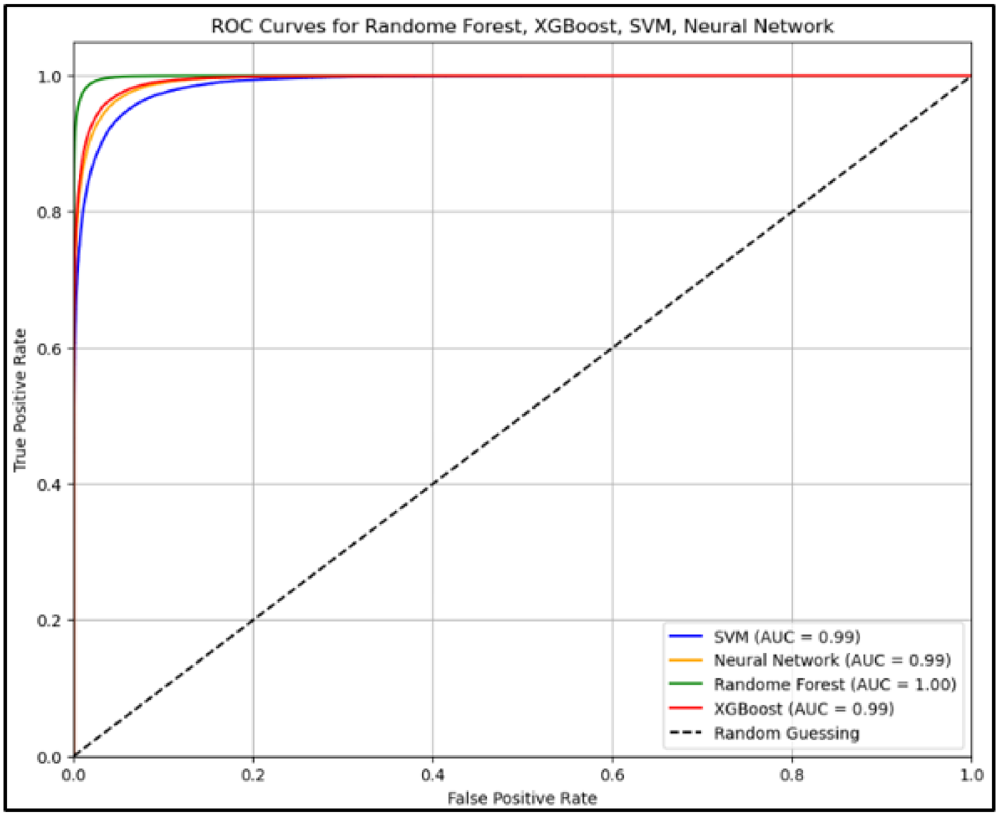

Figure 5.2: ROC curves for the evaluated learners on the stratified holdout split.

The comparative performance metrics (Figure 5.3) further reinforce these findings. Random Forest consistently outperforms the other models, achieving Accuracy, Precision, Recall, and F1 Score values of 0.98, alongside an AUC of 1.00. XGBoost ranks second, with strong yet slightly lower metrics (Accuracy and F1 Score of 0.94, AUC of 0.99), followed by the Neural Network and SVM models, which still exhibit high performance (F1 Scores of 0.93 and 0.91, respectively), Table 5.2. Across all models, the close alignment between Precision and Recall indicates balanced classification behavior, minimizing both false alarms and missed fire eventsan essential requirement for safety-critical applications such as wildfire risk management. The consistently high AUC values across models confirm that the predictor set is both informative and robust, capturing the underlying drivers of wildfire occurrence effectively.

Table 5.2: Test accuracy comparison across learners.

Model	**AUC**	**F1**
Random Forest	0.99	0.98
SVM	0.94	0.94
Neural Network	0.95	0.96
XGBoost	0.96	0.95

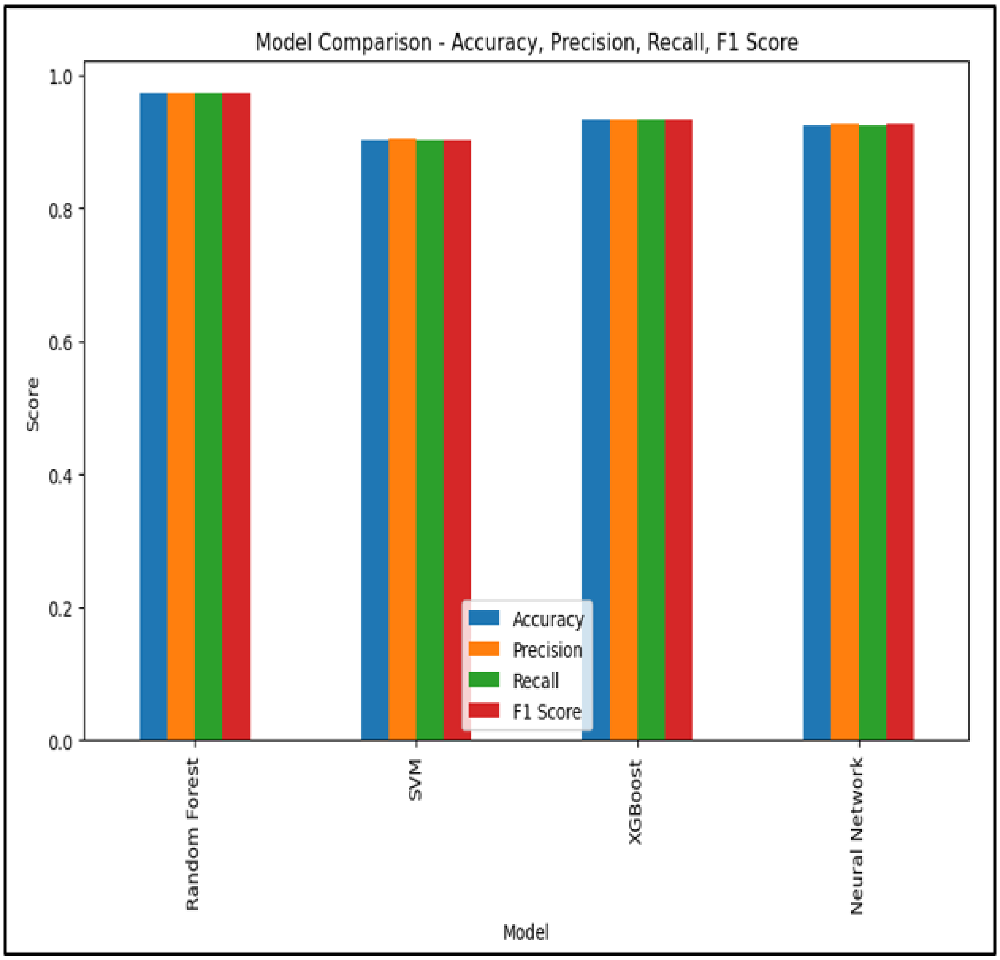

Figure 5.3: Overall model accuracy comparison at the chosen operating point.

To translate these statistical results into actionable spatial insights, each model was applied to the entire study area to generate wildfire susceptibility maps (Figure 5.4). These maps employ a standardized color scheme, where red represents high to very high risk, orange and yellow indicate moderate risk, and green denotes low or negligible risk. All models demonstrate strong spatial agreement, consistently identifying major wildfire hotspots in the central-northern and eastern inland regions. Tree-based models such as Random Forest and XGBoost produce more granular and spatially detailed risk patterns, capturing localized variability in susceptibility, whereas the SVM generates smoother, more generalized risk zones due to its global decision boundary. The Neural Network results lie between these two extremes, preserving major hotspots while slightly smoothing risk transitions.

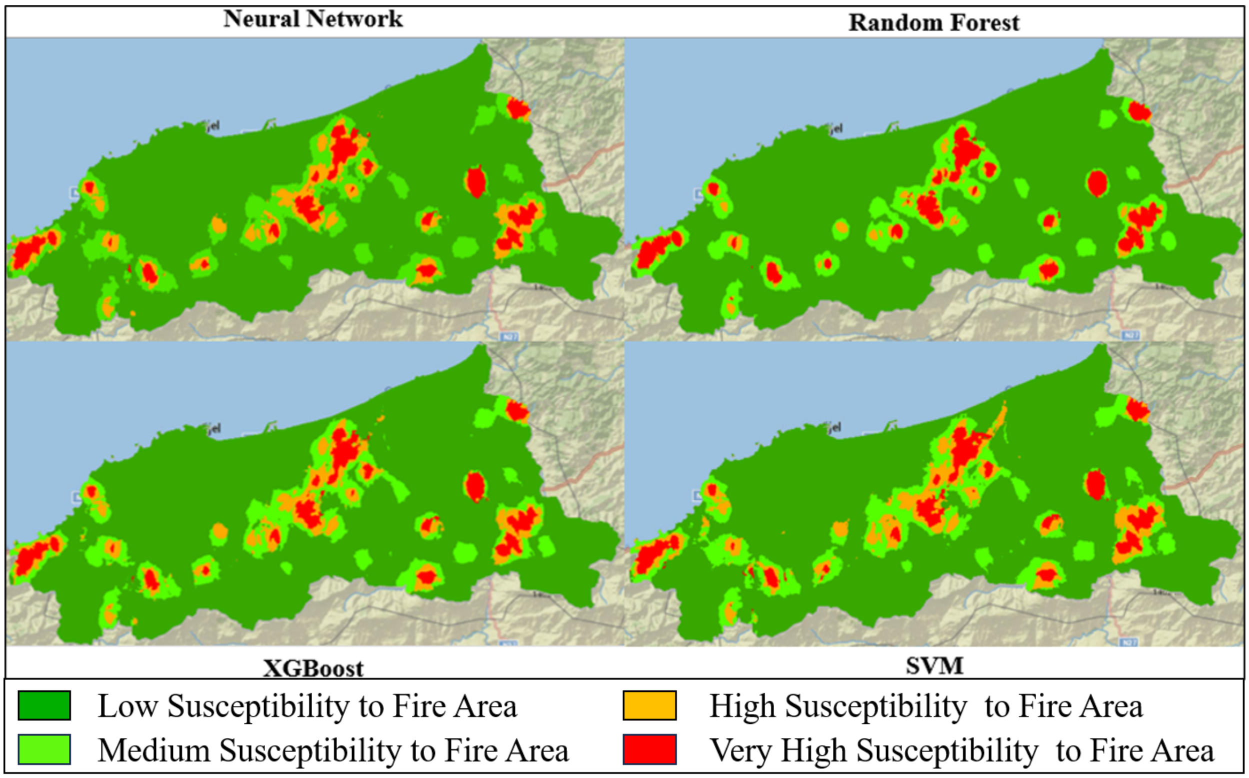

Figure 5.4: Wildfire susceptibility map for the study area.

Despite Random Forest achieving the highest internal performance metrics, XGBoost was selected as the final operational model following consultations with civil protection engineers and fire management authorities. This decision was guided by the principle of safety, as the XGBoost susceptibility map delineates the largest extent of high and very high risk areas, thereby ensuring conservative coverage for prevention planning, early intervention, and resource allocation. From a risk management perspective, prioritizing broader identification of potentially vulnerable zones is preferable to underestimation, particularly in regions with complex humanenvironment interactions.

The robustness of the selected model was further evaluated through spatial validation using independent wildfire occurrences from 20242025. The results show that 87.73% of recorded fire events fall within areas classified as medium to very high susceptibility, demonstrating strong predictive and spatial generalization capability. Specifically, 45.32% of fires occurred in very high-risk zones, 26.76% in high-risk zones, and 15.65% in medium-risk zones, while only 12.27% were located outside the predicted risk areas, see Figure 5.5. This high correspondence between predicted susceptibility and observed fire events confirms the models ability to anticipate future wildfire patterns and highlights its practical utility as a decision-support tool for wildfire prevention and mitigation.

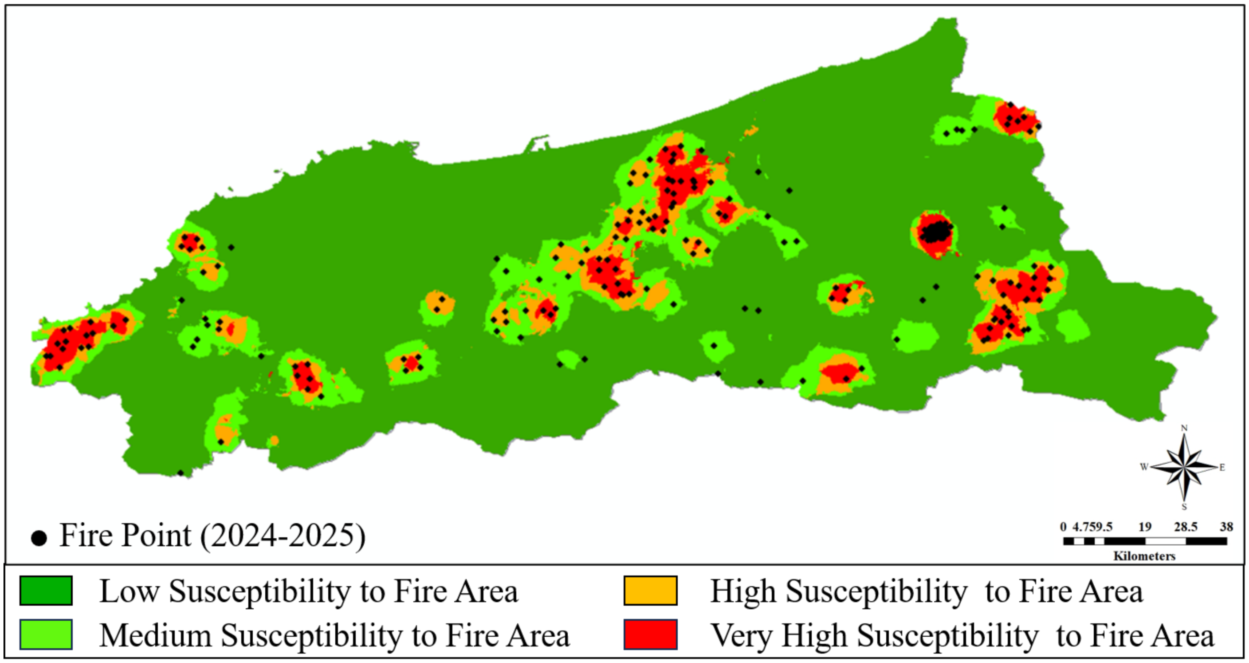

Figure 5.5: Spatial validation using independent 2024–2025 fire occurrences overlaid on susceptibility classes.

Only a limited number of previous studies have incorporated real and independent wildfire occurrences for post-model validation, underscoring both the originality and the strength of the present framework. For instance, Study [44] relied on qualitative visual comparison without reporting quantitative accuracy measures, while Study [22] achieved approximately 76% correctly classified fire events. In contrast, the proposed approach attains a substantially higher validation accuracy of 87.73%, demonstrating an enhanced capacity to capture both the spatial and temporal dynamics of wildfire occurrence. This improvement can be attributed to the integration of a diverse set of environmental and anthropogenic predictors, a refined data preprocessing workflow, and the application of advanced machine learning algorithms supported by explainable AI techniques, particularly SHAP analysis.

In line with prior research reporting AUC values above 0.93 for wildfire susceptibility mapping tasks [47, 60], the present study confirms that high internal performance metrics alone are insufficient to demonstrate true predictive skill. However, the forward-looking validation on unseen 20242025 fire events provides compelling evidence of out-of-sample robustness, rather than overfitting. Furthermore, susceptibility maps generated by RF, XGBoost, NN, and SVM show strong cross-model consistency, with differences confined mainly to boundary cells between adjacent risk classes. The spatial patterns are physically coherent, aligning with known ignition mechanisms such as agricultureforest interfaces, proximity to roads and settlements, and regional climatic constraints. SHAP-based feature importance rankings consistently emphasize human activity proxies, hydro-meteorological variables, and topography as dominant drivers, reinforcing the interpretability and regional relevance of the models.

Overall, these results demonstrate that the proposed framework not only achieves state-of-the-art statistical performance, but also delivers spatially and temporally validated wildfire susceptibility predictions, supporting its suitability for operational implementation in wildfire risk assessment and management systems in the Jijel province and similar Mediterranean environments.

5.4 Interpretability and Spatial Visualization of Results

Global feature contributions were quantified using SHAP values derived from the calibrated XGBoost model (Figure 5.6). The results clearly indicate that human-activity-related variables, particularly proximity to roads and settlements, are the dominant drivers of wildfire susceptibility in the study area. These are followed by hydro-meteorological variables, namely precipitation, maximum temperature, and humidity, while topographic factors such as slope and elevation play a secondary, modulating role rather than acting as primary ignition drivers. This ordering reflects the ignition reality of Jijel Province, where a large proportion of fires originate near wildlandagriculture interfaces and peri-urban areas, and where climatic stress conditions determine whether ignition develops into a sustained fire event.

Importantly, SHAP values also provide directionality: shorter distances to roads and residential areas increase risk; lower precipitation and humidity elevate susceptibility; higher temperatures amplify ignition probability; and steeper slopes slightly increase risk when combined with favorable climatic and anthropogenic conditions. Because predictor collinearity was addressed during preprocessing, these attributions are stable and physically meaningful rather than artifacts of redundant inputs.

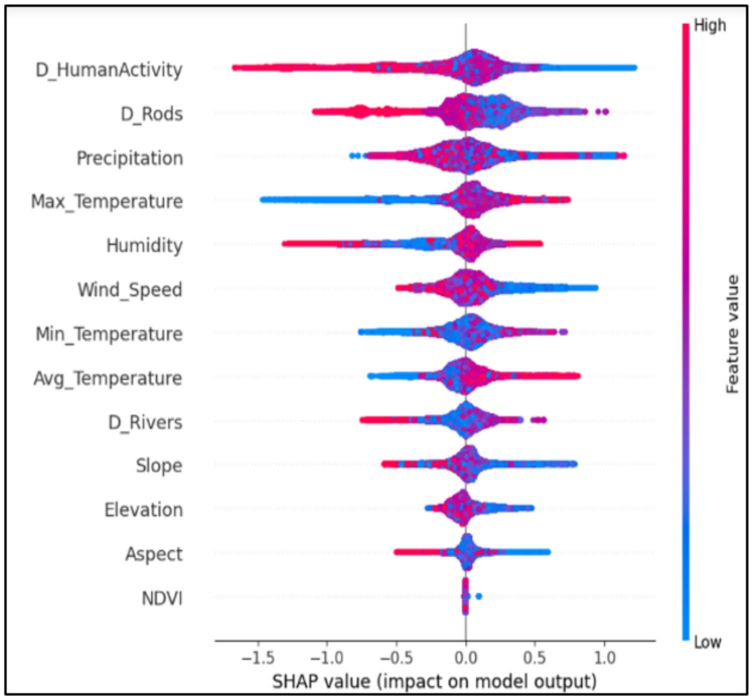

Figure 5.6: SHAP summary showing global feature importance for the calibrated XGBoost model. Higher absolute values indicate stronger contributions to predicted susceptibility.

The spatial manifestation of these relationships is clearly visible in the wildfire susceptibility map (Figure 5.4). High and very high susceptibility zones form coherent spatial patterns aligned with wildlandagriculture transitions, peri-urban fringes, and road-connected mid-slope corridors, where ignition pressure and fuel continuity coexist. Conversely, low-susceptibility areas are concentrated in wetter valley bottoms and higher-elevation zones with shaded aspects, where fuel moisture remains relatively high. Feature-overlay analyses further reinforce this interpretation. When susceptibility is examined alongside precipitation patterns (Figure 5.7), high-risk belts correspond predominantly to drier micro-regions during the reference period. Similarly, overlays with human activity and road proximity (Figure 5.8) reveal clustering of high-risk pixels along access corridors and settlement interfaces. Local pixel-level inspections confirm this logic: near roadforest edges, positive SHAP contributions from human-access variables dominate and are amplified by low humidity and recent precipitation deficits, whereas in remote, wetter uplands, moisture-related variables suppress susceptibility despite steep slopes.

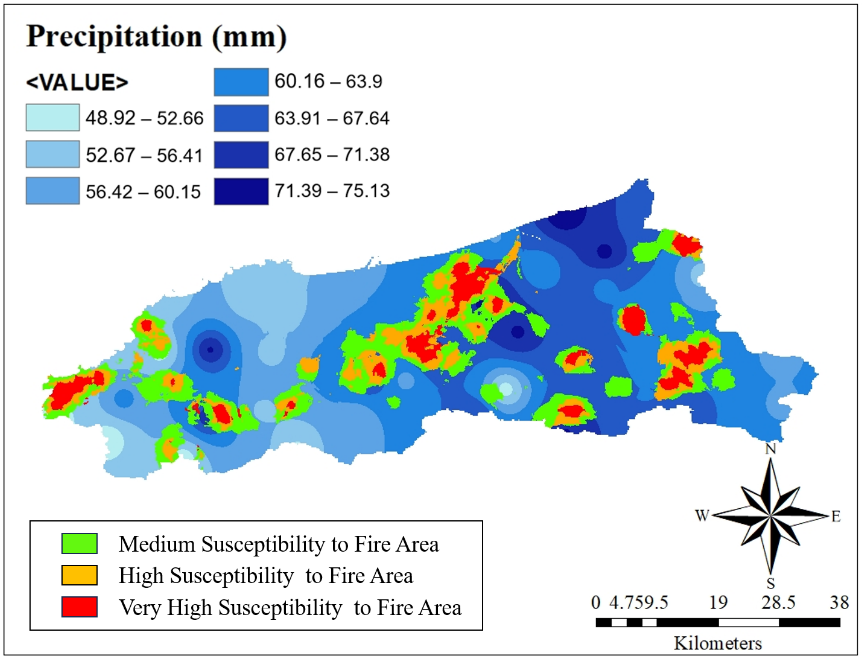

Figure 5.7: High-risk areas relative to precipitation patterns during the reference period.

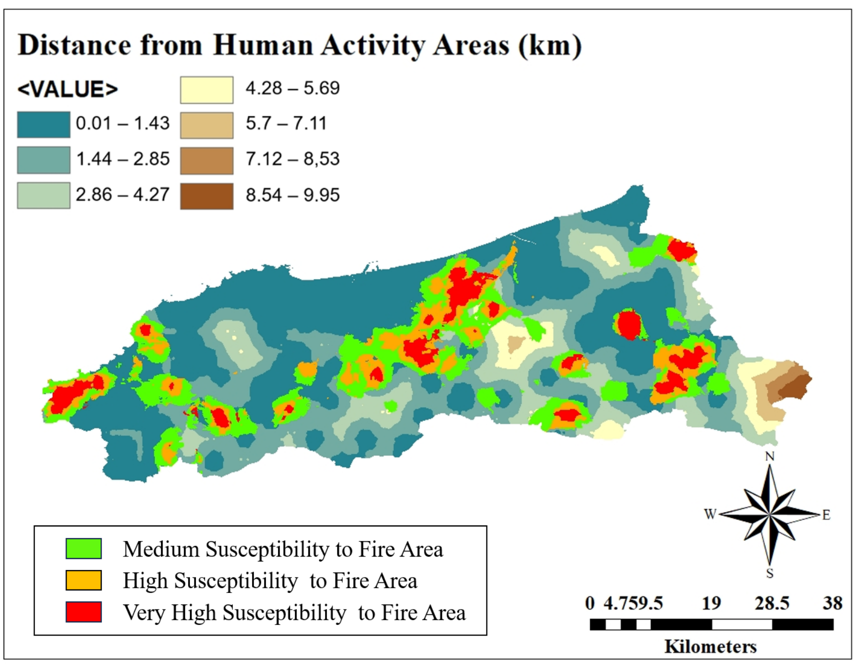

Figure 5.8: High-risk areas relative to human activity and road proximity.

While internal performance metrics demonstrate predictive strength, they do not guarantee spatial realism. To address this, an independent time-forward validation was conducted using wildfire occurrences from 20242025 obtained from Protection Civile records and MODIS hot spots. The majority of these new fire events fall within areas previously classified as medium to very high susceptibility, with a graded distribution that mirrors predicted risk levels rather than clustering randomly. This confirms that the learned susceptibility ranking generalizes to unseen fire seasons and that the spatial patterns produced by the model reflect genuine fire-driving mechanisms rather than overfitting to historical data.

Interpretability was further strengthened by examining model uncertainty in relation to spatial patterns. Although not explicitly mapped, ensemble variability from RF and XGBoost increases mainly along class boundaries and in topographically complex or data-sparse areaslocations where SHAP contributions from meteorological variables fluctuate most. These zones represent logical candidates for targeted field verification, improved monitoring, or enhanced data collection. In contrast, areas where high susceptibility coincides with low model spread and strong SHAP support from multiple drivers provide the highest confidence for prioritizing preventive interventions.

From an operational perspective, the interpretability results directly inform wildfire management strategies. High-risk zones concentrated near roads and settlements can be prioritized for fuel-break maintenance, targeted patrols during heat episodes, and community awareness campaigns. Medium-risk belts function as adaptive buffer zones, where preparedness levels can be escalated under forecasted high temperatures and low humidity. At the municipal scale, SHAP-based factor rankings offer an interpretable rationale for intervention design: where human-access variables dominate, ignition prevention and access control are most effective, whereas areas driven primarily by hydro-meteorological stress require heightened readiness during adverse weather conditions.

Overall, the integration of feature importance analysis, spatial feature overlays, and independent fire-event validation grounds the susceptibility maps in understandable physical and human processes. This interpretability framework not only reinforces trust in the model outputs but also bridges the gap between predictive accuracy and practical decision-making, enabling wildfire susceptibility mapping to function as a reliable tool for prevention planning, resource allocation, and risk communication.

5.5 Challenges and Future Work

Developing reliable wildfire susceptibility maps that are both scientifically robust and operationally useful presents a series of interconnected methodological and practical challenges. These challenges extend beyond model selection and encompass data availability, feature justification, training and validation strategies, interpretability, and operational integration. While the proposed framework addresses many of these issues, recognizing its current limitations is essential to guide future improvements and ensure sustainable applicability in wildfire management.

One of the most persistent challenges lies in the scarcity, uneven distribution, and heterogeneity of input data, particularly at the local scale. Meteorological stations are sparse and often concentrated near settlements or transportation corridors, which can bias humidity, temperature, and wind measurements relative to forested and mountainous areas. Historical fire records combine multiple sources with varying spatial accuracy, temporal resolution, and reporting reliability, leading to potential positional errors and underrepresentation of small or rapidly suppressed fires. Similarly, anthropogenic layers derived from road and settlement databases often omit informal tracks and seasonal access routes that significantly influence ignition pressure.

These limitations propagate uncertainty throughout the modeling pipeline and may distort feature attribution, especially in topographically complex or data-poor regions. Future work should prioritize densifying the observational backbone through low-cost micro-meteorological sensors deployed in wildlandurban interfaces and canyon corridors, complemented by camera-based systems for early smoke detection. Data assimilation techniques that fuse local observations with reanalysis products (e.g., ERA5) can further refine near-surface meteorological fields. Regular updates of human-access layers using recent imagery and local authority records are also essential. In parallel, establishing lightweight data governance protocolscovering coordinate validation, duplicate removal, and metadata consistencywould significantly improve the reliability of fire incident archives used for model training and validation.

The rapid growth of open-access geospatial datasets has expanded the pool of potential predictors for wildfire modeling, but this abundance introduces a critical challenge: determining which variables are truly relevant, how they should be represented, and how redundancy should be managed. Climatic variables, for example, can be expressed as means, extremes, anomalies, or seasonal aggregates, each capturing different aspects of fire risk. Without careful justification, feature-rich models risk instability, reduced interpretability, and overfitting.

Future research should continue to combine expert knowledge with data-driven selection strategies, including multicollinearity analysis, feature reduction, and importance screening across multiple model families. Beyond static selection, investigating how feature importance evolves over timeparticularly under changing climate conditionscan provide insight into shifting ignition regimes. Such longitudinal analyses may reveal whether anthropogenic drivers remain dominant or whether climatic stressors gain relative influence in future decades.

Machine learning models for wildfire susceptibility are particularly vulnerable to inflated performance estimates due to spatial autocorrelation and temporal leakage. Neighboring pixels often share similar environmental and anthropogenic characteristics, meaning that random data splits can unintentionally transfer information between training and testing samples. Similarly, models evaluated on years that closely resemble training periods may fail to generalize to anomalous fire seasons.

Addressing these issues requires more rigorous validation designs. Spatially blocked cross-validation, combined with time-forward evaluation, should become standard practice to better reflect operational deployment. Performance should also be disaggregated across landscape contextssuch as coastal versus inland areas or near-access versus remote zonesto identify where predictive skill is strongest or weakest. Monitoring probability calibration over time and implementing periodic recalibration or rolling retraining will be increasingly important as land use and climate patterns evolve.

Although advanced machine learning models can achieve high predictive accuracy, their adoption in operational wildfire management depends heavily on transparency and interpretability. Decision-makers require clear explanations of why specific areas are classified as high risk and which factors dominate local susceptibility. Without such insights, even high-performing models may be perceived as black boxes and remain underutilized.

Future work should expand the interpretability toolkit beyond global feature rankings. Partial dependence or accumulated local effects can clarify nonlinear responses, while interaction analyses can reveal coupled effects such as slopewind alignment. Local explanations at selected high-risk sites should be routinely examined and their stability tested across model variants to avoid brittle narratives. When image-based or deep learning components are introduced, saliency methods and concept-based interpretability tests will be essential to ensure that models focus on physically meaningful patterns rather than artifacts. Counterfactual analysesquantifying how much a variable must change to shift susceptibility classcan further support policy design and risk communication.

A fundamental limitation of static susceptibility maps is that they represent long-term predisposition rather than immediate danger. While they are invaluable for prevention planning and resource allocation, operational wildfire management requires dynamic systems that integrate both predictive (leading) indicators and reactive (trailing) signals.

Future research should therefore focus on integrating susceptibility outputs with real-time and forecasted data streams. Predictive features such as temperature, humidity, vegetation dryness, and human activity provide early warning potential, while reactive featuresthermal anomalies, gas emissions, smoke aerosols, and surface changesenable rapid detection and confirmation of active fires. Hybrid architectures that combine these two information streams can bridge the gap between anticipation and response.

Spatio-temporal deep learning frameworks, such as ConvLSTM-based architectures, offer promising avenues for modeling the joint evolution of environmental conditions and fire activity. A practical direction is a hybrid modeling strategy, where deep networks extract spatio-temporal representations from imagery and sensor data, while calibrated machine learning models integrate these representations with tabular predictors to preserve interpretability and robustness. Extensive ablation studies will be required to quantify the added value of such complexity and to identify the minimal architecture that delivers meaningful operational gains.

For susceptibility mapping to deliver real-world impact, models must be embedded within operational workflows. This includes publishing maps as versioned web services, defining thresholds linked to decision costs, integrating outputs into dashboards with live weather and fire feeds, and establishing clear standard operating procedures for alert escalation and deactivation. Training and continuous engagement with civil protection agencies are essential to align expectations and incorporate practitioner feedback.

Equally important is the explicit communication of uncertainty. Future systems should expose both aleatoric uncertainty (from noisy and incomplete data) and epistemic uncertainty (from limited knowledge), using ensemble methods, Bayesian approximations, or conformal prediction. Sensitivity analyses and stress tests under extreme meteorological scenarios can reveal failure modes and guide conservative decision thresholds. Ethical considerations, including privacy protection in human-activity layers and transparent communication of false-alarm risks, must also be addressed to ensure equitable and responsible use.

In summary, advancing wildfire susceptibility mapping requires progress on two complementary fronts: raising the scientific standard through improved data, validation, multimodal modeling, and explainability; and completing the operational bridge that connects predictions, uncertainties, and explanations to everyday decision-making. By addressing these challenges, susceptibility mapping can evolve from a static analytical product into a core component of proactive wildfire management and early warning systems, supporting more effective prevention, preparedness, and response in Jijel and comparable Mediterranean environments.

5.6 Conclusion

The results demonstrate that the proposed machine learning framework provides a robust and interpretable approach to wildfire susceptibility mapping. Comparative evaluation confirms that ensemble tree-based models, particularly XGBoost, achieve superior performance while maintaining spatial consistency and generalization capability. The integration of SHAP analysis reveals a clear and region-specific hierarchy of influential factors, with human activity proxies and hydro-meteorological conditions emerging as primary drivers, and topography acting as a secondary modulator. Susceptibility maps exhibit physically meaningful patterns that align with known ignition contexts, such as wildlandagriculture interfaces and road-adjacent corridors, and their validity is further reinforced by the strong correspondence with independent fire events from 20242025. By combining quantitative metrics, explainable AI, spatial analysis, and real-world validation, this study bridges the gap between predictive modeling and operational understanding, demonstrating that accurate wildfire susceptibility maps can also be transparent, interpretable, and actionable for expert analysis and fire management planning.

General conclusion

This thesis addressed the growing challenge posed by wildfires, a phenomenon that has intensified worldwide under the combined influence of climate change, environmental degradation, and expanding human activities. As highlighted in the introduction, the increasing frequency, severity, and socio-environmental impacts of wildfires demand a shift from reactive fire suppression toward proactive, data-driven risk assessment and prevention strategies. Within this context, Wildfire Susceptibility Mapping (WSM) emerges as a critical decision-support tool, particularly in Mediterranean regions such as northern Algeria, where wildfire drivers are complex, spatially heterogeneous, and strongly influenced by humanenvironment interactions.

Focusing on the province of Jijel, this work developed and validated an integrated framework for wildfire susceptibility mapping that combines Geographic Information Systems, advanced machine learning models, and explainable artificial intelligence techniques. By merging environmental, climatic, topographic, hydrological, and anthropogenic datasets from both global repositories and local agencies, the study responded directly to the data availability challenges outlined in the introduction. The resulting harmonized database provided a comprehensive and regionally relevant representation of wildfire conditioning factors, enabling robust modeling and spatial analysis.

Multiple machine learning algorithmsRandom Forest, XGBoost, Neural Networks, and Support Vector Machineswere implemented and systematically compared to capture nonlinear relationships between wildfire occurrences and their driving factors. The consistently high performance achieved across models, as reflected by accuracy, F1 score, and AUC metrics, confirms the suitability of machine learning approaches for wildfire susceptibility assessment in complex Mediterranean landscapes. Importantly, the comparative analysis revealed strong agreement in spatial patterns among models, reinforcing confidence in the identified high-risk zones and demonstrating that the predictive signal is not model-specific but rooted in meaningful environmental and anthropogenic processes.

A central contribution of this thesis lies in moving beyond internal model validation to address one of the major limitations identified in prior wildfire susceptibility studies: the lack of independent, real-world validation. By incorporating newly observed wildfire events from 2024 and 2025, the proposed framework demonstrated strong time-forward predictive capability, with 87.73% of recorded fires falling within areas classified as medium to very high susceptibility. This result confirms that the model captures not only historical patterns but also the evolving spatial dynamics of wildfire risk, providing compelling evidence of generalization rather than overfitting. In doing so, this study strengthens the operational credibility of wildfire susceptibility mapping and aligns with the growing call for validation strategies that reflect real decision-making conditions.

Equally important, this research addressed the interpretability gap often associated with machine learning models. Through the integration of SHAP-based explainability, the thesis translated statistical predictions into understandable insights, clarifying why certain areas are classified as high risk. Feature importance analysis consistently identified human activity proxiessuch as proximity to roads and settlementsas dominant drivers of wildfire susceptibility, followed by hydro-meteorological variables including precipitation, temperature, and humidity. Topographic factors played a secondary, modulatory role. This ordering aligns closely with the regional ignition context of Jijel and validates the thesiss emphasis, introduced at the outset, on the critical role of anthropogenic pressure in Mediterranean wildfire regimes.

The spatial interpretation of susceptibility maps further reinforced this understanding. High-risk areas were shown to concentrate in physically and socially coherent zones, particularly along wildlandagriculture interfaces, peri-urban fringes, and transportation corridors. Overlay analyses with environmental and human layers revealed clear links between susceptibility patterns and known fire-prone conditions, while local-scale inspections confirmed the consistency of model behavior at the pixel level. These results demonstrate that the produced maps are not abstract statistical outputs, but spatially meaningful representations of wildfire risk that reflect real-world processes.

From an operational perspective, the framework developed in this thesis offers tangible benefits for wildfire management and land-use planning. The identification of high- and very-high-risk zones provides actionable guidance for prioritizing fuel management, surveillance, and community awareness efforts, particularly in areas where human access and environmental stress converge. Medium-risk zones, when interpreted alongside weather forecasts, can serve as dynamic buffers for escalating preparedness. Moreover, the explainable nature of the model outputs enhances trust among decision-makers by providing clear, factor-based rationales for intervention strategies.

Beyond its practical implications, this thesis contributes to the scientific literature by addressing a clear regional gap in wildfire susceptibility research. Southern Mediterranean and North African regions remain underrepresented in comparative WSM studies, despite exhibiting fire regimes that differ markedly from those of northern Mediterranean countries. By emphasizing locally relevant drivers, integrating independent validation, and demonstrating the region-specific nature of factor importance, this work underscores the necessity of context-aware modeling rather than the uncritical transfer of models and assumptions across regions.

Nevertheless, certain limitations remain. Spatial resolution constraints, the availability of dense meteorological observations, and potential positional uncertainty in fire records may influence local-scale accuracy. While several mitigation strategies were applied, including stratified sampling and time-forward validation, future work would benefit from denser climate monitoring, multi-temporal predictors, and spatially blocked cross-validation schemes. Additionally, coupling static susceptibility maps with near-real-time meteorological forecasts could further enhance early warning capabilities.

In conclusion, this thesis demonstrates that wildfire susceptibility mapping, when grounded in high-quality data, advanced machine learning, explainable modeling, and independent validation, can provide reliable, interpretable, and operationally relevant insights for wildfire risk management. By responding directly to the challenges outlined in the introductionrising wildfire risk, complex drivers, data limitations, and the need for proactive toolsthis work contributes a robust framework that can be adapted to other Mediterranean and fire-prone regions. Ultimately, the integration of predictive accuracy, spatial coherence, and interpretability represents a meaningful step toward more effective wildfire prevention, preparedness, and resilience in an era of increasing environmental uncertainty.

Bibliography

[1] Amila Akagic and Emir Buza. “Lw-fire: A lightweight wildfire image classification with a deep convolutional neural network.” In: *Applied Sciences* 12.5 (2022), p. 2646.

[2] Minsoo Park, Dai Quoc Tran, Seungsoo Lee, and Seunghee Park. “Multilabel image classification with deep transfer learning for decision support on wildfire response.” In: *Remote Sensing* 13.19 (2021), p. 3985.

[3] Omer Kantarcioglu, Konrad Schindler, and Sultan Kocaman. “Forest fire susceptibility assessment with machine learning methods in north-east Turkiye.” In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 48 (2023), pp. 161–167.

[4] M Seddouki, M Benayad, Z Aamir, M Tahiri, M Maanan, and H Rhinane. “Using machine learning coupled with remote sensing for forest fire susceptibility mapping. Case study Tetouan province, Northern Morocco.” In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 48 (2023), pp. 333–342.

[5] Muhammad Hasif bin Azami, Necmi Cihan Orger, Victor Hugo Schulz, Mengu Cho, et al. “Demonstration of wildfire detection using image classification onboard CubeSat.” In: (2021), pp. 5413–5416.

[6] Mohan Prakash, S Neelakandan, M Tamilselvi, S Velmurugan, S Baghavathi Priya, and Eric Ofori Martinson. “Deep learning-based wildfire image detection and classification systems for controlling biomass.” In: *International journal of intelligent systems* 2023.1 (2023), p. 7939516.

[7] Cristian Vidal-Silva, Roberto Pizarro, Miguel Castillo-Soto, Ben Ingram, Claudia de la Fuente, Vannessa Duarte, Claudia Sangüesa, and Alfredo Ibañez. “A comparative study of a deep reinforcement learning solution and alternative deep learning models for wildfire prediction.” In: *Applied Sciences* 15.7 (2025), p. 3990.

[8] Emilio Chuvieco, Inmaculada Aguado, Marta Yebra, Héctor Nieto, Javier Salas, M Pilar Martín, Lara Vilar, Javier Martínez, Susana Martín, Paloma Ibarra, et al. "Development of a framework for fire risk assessment using remote sensing and geographic information system technologies." In: *Ecological modelling* 221.1 (2010), pp. 46–58.

[9] Shahfahad, Swapan Talukdar, Tanmoy Das, Mohd Waseem Naikoo, Mohd Rihan, and Atiqur Rahman. "Forest fire susceptibility mapping by integrating remote sensing and machine learning algorithms." In: *Advances in remote sensing for forest monitoring* (2022), pp. 179–195.

[10] Haoyuan Hong, Paraskevas Tsangaratos, Ioanna Ilia, Junzhi Liu, A-Xing Zhu, and Chong Xu. "Applying genetic algorithms to set the optimal combination of forest fire related variables and model forest fire susceptibility based on data mining models. The case of Dayu County, China." In: *Science of the total environment* 630 (2018), pp. 1044–1056.

[11] Muhammad Hasif bin Azami, Necmi Cihan Orger, Victor Hugo Schulz, Takashi Oshiro, and Mengu Cho. "Earth observation mission of a 6U CubeSat with a 5-meter resolution for wildfire image classification using convolution neural network approach." In: *Remote Sensing* 14.8 (2022), p. 1874.

[12] Dieu Tien Bui, Kim-Thoa Thi Le, Van Cam Nguyen, Hoang Duc Le, and Inge Revhaug. "Tropical forest fire susceptibility mapping at the Cat Ba National Park Area, Hai Phong City, Vietnam, using GIS-based kernel logistic regression." In: *Remote Sensing* 8.4 (2016), p. 347.

[13] Sandra Oliveira, Friderike Oehler, Jesús San-Miguel-Ayanz, Andrea Camia, and José MC Pereira. "Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest." In: *Forest Ecology and Management* 275 (2012), pp. 117–129.

[14] Yehya Bouzeraa, Nardjas Bouchemal, and Nada Zendaoui. "Pre-disaster Management based Machine Learning, IoT and Big Data: Survey and future direction." In: *2022 2nd International Conference on New Technologies of Information and Communication (NTIC)*. IEEE. 2022, pp. 1–7.

[15] Yehya Bouzreraa, Nardjas Bouchemal, and Salim Djaaboub. "GIS, AI, and environmental data in fire susceptibility mapping: A survey." In: *2023 IEEE 11th International Conference on Systems and Control (ICSC)*. IEEE. 2023, pp. 159–164.

[16] Yehya Bouzeraa, Nardjes Bouchemal, Salim Djaaboub, Georgi Hristov, and Plamen Zahariev. "Machine Learning-Based Wildfire Susceptibility Mapping: A GIS-Integrated Predictive Framework." In: *Applied Sciences* 15.22 (2025), p. 12188.

[17] *EM-DAT: The International Disaster Database*. URL: https://www.emdat.be/ (visited on 08/31/2025).

[18] Rafael Bailon-Ruiz, Arthur Bit-Monnot, and Simon Lacroix. "Real-time wildfire monitoring with a fleet of UAVs." In: *Robotics and Autonomous Systems* 152 (2022), p. 104071.

[19] Alessio De Rango, Luca Furnari, Fabio Cortale, Alfonso Senatore, and Giuseppe Mendicino. "Wildfire early warning system based on a smart CO2 sensors network." In: *Sensors* 25.7 (2025), p. 2012.

[20] Dmitrii Shadrin, Svetlana Illarionova, Fedor Gubanov, Ksenia Evteeva, Maksim Mironenko, Ivan Levchunets, Roman Belousov, and Evgeny Burnaev. "Wildfire spreading prediction using multimodal data and deep neural network approach." In: *Scientific reports* 14.1 (2024), p. 2606.

[21] Viktor Slavkovikj, Steven Verstockt, Sofie Van Hoecke, and Rik Van de Walle. "Review of wildfire detection using social media." In: *Fire safety journal* 68 (2014), pp. 109–118.

[22] Rohan T Bhowmik, Youn Soo Jung, Juan A Aguilera, Mary Prunicki, and Kari Nadeau. "A multi-modal wildfire prediction and early-warning system based on a novel machine learning framework." In: *Journal of environmental management* 341 (2023), p. 117908.

[23] *Statista*. URL: https://www.statista.com/ (visited on 08/31/2025).

[24] *UN DESA Policy Brief No. 111: Wildfires: A growing concern for sustainable development*. URL: https://policy.desa.un.org/publications/undesa-policy-brief-111-wildfires-a-growing-concern-for-sustainable-development (visited on 08/31/2025).

[25] *BBC News*. URL: https://www.bbc.com/news (visited on 08/31/2025).

[26] *Algeria forest fires (Al Jazeera, 2021-08-10)*. URL: https://www.aljazeera.com/news/2021/8/10/algeria-forest-fires (visited on 08/31/2025).

[27] *FIREURISK: Developing a holistic, risk-wise strategy for European wildfire management*. Project website. URL: https://www.fireurisk.eu/ (visited on 08/31/2025).

[28] Cordy Tymstra, Brian J Stocks, Xinli Cai, and Mike D Flannigan. "Wildfire management in Canada: Review, challenges and opportunities." In: *Progress in Disaster Science* 5 (2020), p. 100045.

[29] Sayed Pedram Haeri Boroujeni, Abolfazl Razi, Sahand Khoshdel, Fatemeh Afghah, Janice L Coen, Leo ONeill, Peter Fule, Adam Watts, Nick-Marios T Kokolakis, and Kyriakos G Vamvoudakis. "A comprehensive survey of research towards AI-enabled unmanned aerial systems in pre-, active-, and post-wildfire management." In: *Information Fusion* 108 (2024), p. 102369.

[30] Ihab H. Sawalha. "A contemporary perspective on the disaster management cycle." In: *Foresight* 22.4 (2020), pp. 469–482.

[31] Annie Singla and Rajat Agrawal. "iStage: a deep learning based framework to determine the stage of disaster management cycle from a social media message." In: *Global Knowledge, Memory and Communication* 74.1-2 (2025), pp. 198–219.

[32] Melanie Colavito. "The human dimensions of spatial, pre-wildfire planning decision support systems: A review of barriers, facilitators, and recommendations." In: *Forests* 12.4 (2021), p. 483.

[33] Gabriela Huidobro, Lukas Giessen, and Sarah L Burns. "And it burns, burns, burns, the ring-of-fire: Reviewing and harmonizing terminology on wildfire management and policy." In: *Environmental Science & Policy* 157 (2024), p. 103776.

[34] H. S. Andrianarivony et al. "Machine Learning and Deep Learning for Wildfire Spread Prediction: A Review." In: *Fire* 7.12 (2024), p. 482.

[35] Abira Sengupta and Brendon J Woodford. "Recent advances in explainable machine learning models for wildfire prediction." In: *Applied computing and geosciences* (2025), p. 100266.

[36] Sibo Cheng, Yufang Jin, Sandy P Harrison, César Quilodrán-Casas, Iain Colin Prentice, Yi-Ke Guo, and Rossella Arcucci. "Parameter flexible wildfire prediction using machine learning techniques: Forward and inverse modelling." In: *Remote Sensing* 14.13 (2022), p. 3228.

[37] Hao Dong, Han Wu, Pengfei Sun, and Yunhong Ding. "Wildfire prediction model based on spatial and temporal characteristics: A case study of a wildfire in Portugals Montesinho Natural Park." In: *Sustainability* 14.16 (2022), p. 10107.

[38] Dayou Chen, Yi Yang, Jinwei Hu, Matthew Kasoar, Rossella Arcucci, and Sibo Cheng. *Explainable Global Wildfire Prediction Model using Graph Neural Networks*. 2026.

[39] Nicolas Caron, Hassan N Noura, Lise Nakache, Christophe Guyeux, and Benjamin Aynes. "AI for Wildfire Management: From Prediction to Detection, Simulation, and Impact AnalysisBridging Lab Metrics and Real-World Validation." In: *Ai* 6.10 (2025), p. 253.

[40] Mohammad Marjani, Masoud Mahdianpari, and Fariba Mohammadimanesh. "CNN-BiLSTM: A novel deep learning model for near-real-time daily wildfire spread prediction." In: *Remote Sensing* 16.8 (2024), p. 1467.

[41] Ankita Mohapatra and Timothy Trinh. "Early wildfire detection technologies in practicea review." In: *Sustainability* 14.19 (2022), p. 12270.

[42] George L James, Ryeim B Ansaf, Sanaa S Al Samahi, Rebecca D Parker, Joshua M Cutler, Rhode V Gachette, and Bahaa I Ansaf. "An efficient wildfire detection system for ai-embedded applications using satellite imagery." In: *Fire* 6.4 (2023), p. 169.

[43] Kathiravan Thangavel, Dario Spiller, Roberto Sabatini, Stefania Amici, Sarathchandrakumar Thottuchirayil Sasidharan, Haytham Fayek, and Pier Marzocca. "Autonomous satellite wildfire detection using hyperspectral imagery and neural networks: A case study on australian wildfire." In: *Remote Sensing* 15.3 (2023), p. 720.

[44] Kinshuk Govil, Morgan L Welch, J Timothy Ball, and Carlton R Pennypacker. "Preliminary results from a wildfire detection system using deep learning on remote camera images." In: *Remote Sensing* 12.1 (2020), p. 166.

[45] Shuo Zhang, Demin Gao, Haifeng Lin, and Quan Sun. "Wildfire detection using sound spectrum analysis based on the internet of things." In: *Sensors* 19.23 (2019), p. 5093.

[46] Rafik Ghali, Moulay A Akhloufi, and Wided Souidene Mseddi. "Deep learning and transformer approaches for UAV-based wildfire detection and segmentation." In: *Sensors* 22.5 (2022), p. 1977.

[47] William J De Groot and Michael D Flannigan. "Climate change and early warning systems for wildland fire." In: *Reducing disaster: Early warning systems for climate change*. Springer, 2014, pp. 127–151.

[48] Songsheng Li. "Wildfire early warning system based on wireless sensors and unmanned aerial vehicle." In: *Journal of Unmanned Vehicle Systems* 7.1 (2018), pp. 76–91.

[49] Harkiran Kaur and Sandeep Kumar Sood. "Energy-efficient IoT-fog-cloud architectural paradigm for real-time wildfire prediction and forecasting." In: *IEEE Systems Journal* 14.2 (2019), pp. 2003–2011.

[50] Nikos Kalatzis, George Routis, Yiorgos Marinellis, Marios Avgeris, Ioanna Roussaki, Symeon Papavassiliou, and Miltiades Anagnostou. "Semantic interoperability for IoT platforms in support of decision making: An experiment on early wildfire detection." In: *Sensors* 19.3 (2019), p. 528.

[51] Chiu Chun Chan, Sheeraz A Alvi, Xiangyun Zhou, Salman Durrani, Nicholas Wilson, and Marta Yebra. "A survey on IoT ground sensing systems for early wildfire detection: technologies, challenges, and opportunities." In: *IEEe Access* 12 (2024), pp. 172785–172819.

[52] Osama M Bushnaq, Anas Chaaban, and Tareq Y Al-Naffouri. "The role of UAV-IoT networks in future wildfire detection." In: *IEEE Internet of Things Journal* 8.23 (2021), pp. 16984–16999.

[53] Martina Putzu, Daniele Loru, Francesco Carta, Angelo Ledda, Alessio Chirigu, Mariella Sole, Matteo Anedda, and Daniele Giusto. "Enhancing Wildfire Risk Management Through Sensor-Based AI Integration in Social IoT Frameworks." In: vol. 78. 1. MDPI, 2024, p. 4.

[54] Mounir Grari, Idriss Idrissi, Mohammed Boukabous, Omar Moussaoui, Mostafa Azizi, and Mimoun Moussaoui. "Early wildfire detection using machine learning model deployed in the fog/edge layers of IoT." In: *Indones. J. Electr. Eng. Comput. Sci* 27.2 (2022), pp. 1062–1073.

[55] Esmaeil Seraj, Andrew Silva, and Matthew Gombolay. "Multi-UAV planning for cooperative wildfire coverage and tracking with quality-of-service guarantees." In: *Autonomous Agents and Multi-Agent Systems* 36.2 (2022), p. 39.

[56] Abdelmalek Bouguettaya, Hafed Zarzour, Amine Mohammed Taberkit, and Ahmed Kechida. "A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms." In: *Signal Processing* 190 (2022), p. 108309.

[57] Saydirasulov Norkobil Saydirasulovich, Mukhriddin Mukhiddinov, Oybek Djuraev, Akmalbek Abdusalomov, and Young-Im Cho. "An improved wildfire smoke detection based on YOLOv8 and UAV images." In: *Sensors* 23.20 (2023), p. 8374.

[58] Xuhong Yang, Xiaobin Jin, and Yinkang Zhou. "Wildfire risk assessment and zoning by integrating Maxent and GIS in Hunan province, China." In: *Forests* 12.10 (2021), p. 1299.

[59] J. V. N. Ramesh, S. Suneel, and J. A. Dhanke. "Integrating Remote Sensing and GIS Techniques for Effective Wildfire Risk Mapping and Management." In: *Remote Sensing Earth System Sciences* 7 (2024), pp. 271–282.

[60] C. Ma, R. Pu, J. Downs, and H. Jin. "Characterizing Spatial Patterns of Amazon Rainforest Wildfires and Driving Factors by Using Remote Sensing and GIS Geospatial Technologies." In: *Geosciences* 12 (2022), p. 237.

[61] Naveed Ejaz and Salimur Choudhury. "A comprehensive survey of the machine learning pipeline for wildfire risk prediction and assessment." In: *Ecological Informatics* 90 (2025), p. 103325. DOI: 10.1016/j.ecoinf.2025.103325.

[62] Fantine Huot, Yi-Fan Chen, and Matthias Ihme. "Next Day Wildfire Spread: A Machine Learning Dataset to Predict Wildfire Spreading From Remote-Sensing Data." In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–13.

[63] Qian He, Ziyu Jiang, Ming Wang, and Kai Liu. "Landslide and wildfire susceptibility assessment in Southeast Asia using ensemble machine learning methods." In: *Remote Sensing* 13.8 (2021), p. 1572.

[64] Sadegh Khanmohammadi, Mehrdad Arashpour, Emadaldin Mohammadi Golafshani, Miguel G Cruz, Abbas Rajabifard, and Yu Bai. "Prediction of wildfire rate of spread in grasslands using machine learning methods." In: *Environmental Modelling & Software* 156 (2022), p. 105507.

[65] Sanaz Moghim and Majid Mehrabi. "Wildfire assessment using machine learning algorithms in different regions." In: *Fire Ecology* 20.1 (2024), p. 104.

[66] Triantafyllos Falaras, Ioanna Tselka, Ioannis Papadopoulos, Maria Nikolidaki, Andreas Karavias, Despoina Bafi, Aliki Petani, Pavlos Krassakis, and Issaak Parcharidis. "Operational Mapping and Post-Disaster Hazard Assessment by the Development of a Multiparametric Web App Using Geospatial Technologies and Data: Attica Region 2021 Wildfires (Greece)." In: *Applied Sciences* 12.14 (2022), p. 7256.

[67] Rony Teguh, Fengky F Adji, Benius Benius, and Mohammad Nur Aulia. "Android mobile application for wildfire reporting and monitoring." In: *Bulletin of Electrical Engineering and Informatics* 10.6 (2021), pp. 3412–3421.

[68] Miguel Lourenço, Luís B Oliveira, João P Oliveira, André Mora, Henrique Oliveira, and Rui Santos. "An integrated decision support system for improving wildfire suppression management." In: *ISPRS International Journal of Geo-Information* 10.8 (2021), p. 497.

[69] A Kamilaris, JB Filippi, C Padubidri, R Koole, and S Karatsiolis. "Examining the potential of mobile applications to assist people to escape wildfires in real-time." In: *Fire Safety Journal* 136 (2023), p. 103747.

[70] Zohre Sadat Pourtaghi, Hamid Reza Pourghasemi, and Mauro Rossi. "Forest fire susceptibility mapping in the Minudasht forests, Golestan province, Iran." In: *Environmental earth sciences* 73.4 (2015), pp. 1515–1533.

[71] Olga Nikolaychuk, Julia Pestova, and Aleksandr Yurin. "Wildfire susceptibility mapping in baikal natural territory using random forest." In: *Forests* 15.1 (2024), p. 170.

[72] Saeedeh Eskandari, Hamid Reza Pourghasemi, and John P Tiefenbacher. "Fire-susceptibility mapping in the natural areas of Iran using new and ensemble data-mining models." In: *Environmental Science and Pollution Research* 28.34 (2021), pp. 47395–47406.

[73] Bahareh Kalantar, Naonori Ueda, Mohammed O Idrees, Saeid Janizadeh, Kourosh Ahmadi, and Farzin Shabani. "Forest fire susceptibility prediction based on machine learning models with resampling algorithms on remote sensing data." In: *Remote Sensing* 12.22 (2020), p. 3682.

[74] Akif Al-Fugara, Ali Nouh Mabdeh, Mohammad Ahmadlou, Hamid Reza Pourghasemi, Rida Al-Adamat, Biswajeet Pradhan, and Abdel Rahman Al-Shabeeb. "Wildland fire susceptibility mapping using support vector regression and adaptive neuro-fuzzy inference system-based whale optimization algorithm and simulated annealing." In: *ISPRS International Journal of Geo-Information* 10.6 (2021), p. 382.

[75] Arnick Abdollahi and Biswajeet Pradhan. "Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model." In: *Science of the Total Environment* 879 (2023), p. 163004.

[76] Seyed Vahid Razavi-Termeh, Abolghasem Sadeghi-Niaraki, and Soo-Mi Choi. "Ubiquitous GIS-based forest fire susceptibility mapping using artificial intelligence methods." In: *Remote Sensing* 12.10 (2020), p. 1689.

[77] R Eslami, M Azarnoush, A Kialashki, and F Kazemzadeh. "GIS-based forest fire susceptibility assessment by random forest, artificial neural network and logistic regression methods." In: *Journal of Tropical Forest Science* 33.2 (2021), pp. 173–184.

[78] Kyriakos Chaleplis, Avery Walters, Bin Fang, Venkataraman Lakshmi, and Alexandra Gemitzi. "A soil moisture and vegetation-based susceptibility mapping approach to wildfire events in Greece." In: *Remote Sensing* 16.10 (2024), p. 1816.

[79] Khalil Gholamnia, Thimmaiah Gudiyangada Nachappa, Omid Ghorbanzadeh, and Thomas Blaschke. "Comparisons of diverse machine learning approaches for wildfire susceptibility mapping." In: *Symmetry* 12.4 (2020), p. 604.

[80] Tran Thi Tuyen, Abolfazl Jaafari, Hoang Phan Hai Yen, Trung Nguyen-Thoi, Tran Van Phong, Huu Duy Nguyen, Hiep Van Le, Tran Thi Mai Phuong, Son Hoang Nguyen, Indra Prakash, et al. "Mapping forest fire susceptibility using spatially explicit ensemble models based on the locally weighted learning algorithm." In: *Ecological Informatics* 63 (2021), p. 101292.

[81] Ana Novo, Hurem Dutal, and Saeedeh Eskandari. "Fire susceptibility modeling and mapping in Mediterranean forests of Turkey: a comprehensive study based on fuel, climatic, topographic, and anthropogenic factors." In: *Euro-Mediterranean Journal for Environmental Integration* 9.2 (2024), pp. 655–679.

[82] Ling Hu, Volker Hochschild, Harald Neidhardt, Michael Schultz, Pegah Khosravani, and Hadi Shokati. "BIPE: A bi-layer predictive ensemble framework for forest fire susceptibility mapping in Germany." In: *Remote Sensing* 17.1 (2024), p. 7.

[83] Jayshree Das, Susanta Mahato, Pawan Kumar Joshi, and Yuei-An Liou. "Forest fire susceptibility zonation in Eastern India using statistical and weighted modelling approaches." In: *Remote Sensing* 15.5 (2023), p. 1340.

[84] Farzaneh Noroozi, Gholamabbas Ghanbarian, Roja Safaeian, and Hamid Reza Pourghasemi. "Forest fire mapping: a comparison between GIS-based random forest and Bayesian models." In: *Natural Hazards* 120.7 (2024), pp. 6569–6592.

[85] Nazanin Bahadori, Seyed Vahid Razavi-Termeh, Abolghasem Sadeghi-Niaraki, Khalifa M Al-Kindi, Tamer Abuhmed, Behrokh Nazeri, and Soo-Mi Choi. "Wildfire susceptibility mapping using deep learning algorithms in two satellite imagery dataset." In: *Forests* 14.7 (2023), p. 1325.

[86] Muzaffer Can Iban and Oktay Aksu. "SHAP-driven explainable artificial intelligence framework for wildfire susceptibility mapping using MODIS active fire pixels: An in-depth interpretation of contributing factors in Izmir, Türkiye." In: *Remote Sensing* 16.15 (2024), p. 2842.

[87] Marj Tonini, Mirko DAndrea, Guido Biondi, Silvia Degli Esposti, Andrea Trucchia, and Paolo Fiorucci. "A machine learning-based approach for wildfire susceptibility mapping. The case study of the Liguria region in Italy." In: *Geosciences* 10.3 (2020), p. 105.

[88] Abdel Rahman Al-Shabeeb, Ibraheem Hamdan, Sedigheh Meimandi Parizi, Akif Al-Fugara, Sanaa Odat, Ismail Elkhrachy, Tongxin Hu, and Saad Sh Sammen. "A comparative study of genetic algorithm-based ensemble models and knowledge-based models for wildfire susceptibility mapping." In: *Sustainability* 15.21 (2023), p. 15598.

[89] Hazem Ghassan Abdo, Hussein Almohamad, Ahmed Abdullah Al Dughairi, and Motirh Al-Mutiry. "GIS-based frequency ratio and analytic hierarchy process for forest fire susceptibility mapping in the western region of Syria." In: *Sustainability* 14.8 (2022), p. 4668.

[90] Andrea Trucchia, Giorgio Meschi, Paolo Fiorucci, Andrea Gollini, and Dario Negro. "Defining wildfire susceptibility maps in Italy for understanding seasonal wildfire regimes at the national level." In: *Fire* 5.1 (2022), p. 30.

[91] Bin Liao, Tao Zhou, Yanping Liu, Min Li, and Tao Zhang. "Tackling the wildfire prediction challenge: an explainable artificial intelligence (XAI) model combining extreme gradient boosting (XGBoost) with SHapley additive exPlanations (SHAP) for enhanced interpretability and accuracy." In: *Forests* 16.4 (2025), p. 689.

[92] Tuyet Anh Thi Do, Hau Duc Tran, Anh Ngoc Thi Do, et al. "Classifying forest cover and mapping forest fire susceptibility in Dak Nong province, Vietnam utilizing remote sensing and machine learning." In: *Ecological Informatics* 79 (2024), p. 102392.

[93] Anuj Tiwari, Mohammad Shoab, and Abhilasha Dixit. "GIS-based forest fire susceptibility modeling in Pauri Garhwal, India: a comparative assessment of frequency ratio, analytic hierarchy process and fuzzy modeling techniques." In: *Natural hazards* 105.2 (2021), pp. 1189–1230.

[94] Panagiotis Symeonidis, Thanasis Vafeiadis, Dimosthenis Ioannidis, and Dimitrios Tzovaras. "Wildfire susceptibility mapping in Greece using ensemble machine learning." In: vol. 6. 3. MDPI, 2025, p. 75.

[95] Jiankai Zhu, Xiaodong Liu, Pengle Cheng, Mingyu Wang, and Ying Huang. "Unveiling spatiotemporal patterns of wildfire risk: a transformer-based earth system analysis." In: *Climate Dynamics* 63.1 (2025), p. 21.

[96] Ali Nouh Mabdeh, Akif Al-Fugara, Khaled Mohamed Khedher, Muhammed Mabdeh, Abdel Rahman Al-Shabeeb, and Rida Al-Adamat. "Forest fire susceptibility assessment and mapping using support vector regression and adaptive neuro-fuzzy inference system-based evolutionary algorithms." In: *Sustainability* 14.15 (2022), p. 9446.

[97] Fatih Sari. "Forest fire susceptibility mapping via multi-criteria decision analysis techniques for Mugla, Turkey: A comparative analysis of VIKOR and TOPSIS." In: *Forest Ecology and Management* 480 (2021), p. 118644.

[98] Alexandra Bjånes, Rodrigo De La Fuente, and Pablo Mena. "A deep learning ensemble model for wildfire susceptibility mapping." In: *Ecological Informatics* 65 (2021), p. 101397.

[99] Sohaib KM Abujayyab, Moustafa Moufid Kassem, Ashfak Ahmad Khan, Raniyah Wazirali, Mücahit Cokun, Enes Taolu, Ahmet Öztürk, and Ferhat Toprak. "Wildfire susceptibility mapping using five boosting machine learning algorithms: The case study of the Mediterranean region of Turkey." In: *Advances in Civil Engineering* 2022.1 (2022), p. 3959150.

[100] Omayma Chabou, Youcef Lazri, and Mohammed Chérif Lehtihet. "Application of Flood Hazard Index and Multi-Criteria Decision Analysis to Determine the effect of Urbanization on Flood Hazard areas in the City of Jijel, Algeria." In: *Indones J Soc Sci Res* 5 (2024), pp. 427–43.

[101] Abdelmalek Drouiche, Faouzi Zahi, Taha-Hocine Debieche, Abdelmalek Lekoui, and Souhil Mahdid. "Assessment of surface water quality: a case of Jijel region, North-East Algeria." In: *Arabian Journal of Geosciences* 15.3 (2022), p. 252.

[102] Omayma Chabou, Youcef Lazri, Simona Mannucci, Adriana Ciardiello, Federica Rosso, and Marco Ferrero. "Advancing flood resilience in North African coastal Cities: A contextual analysis of Jijel, Algeria." In: *City and Environment Interactions* 28 (2025), p. 100233.

[103] Toufik Boutellis and Ammar Bouchair. "Predictive capacity analysis for outdoor thermal comfort assessments: a case study of Jijel City, Algeria." In: *J. Adv. Res. Fluid Mech. Therm. Sci* 98.1 (2022), pp. 18–41.

[104] Zheng Zhu, Xiaofan Deng, Fan Zhao, Shiyou Li, and Leiguang Wang. "How environmental factors affect forest fire occurrence in yunnan forest region." In: *Forests* 13.9 (2022), p. 1392.

[105] Nicoletta D'Angelo, Alessandro Albano, Andrea Gilardi, and Giada Adelfio. "Spatio-temporal point process modelling of fires in Sicily exploring human and environmental factors." In: *arXiv preprint arXiv:2402.10859* (2024).

[106] Christos Bountzouklis, Dennis M Fox, and Elena Di Bernardino. "Environmental factors affecting wildfire-burned areas in southeastern France, 1970–2019." In: *Natural Hazards and Earth System Sciences* 22.4 (2022), pp. 1181–1200.

[107] Julia Miller, Danielle Touma, and Manuela I Brunner. "Compounding preconditions of wildfires vary in time and space within Europe." In: *Communications Earth & Environment* 6.1 (2025), p. 1005.

[108] Randall P Benson, John O Roads, and David R Weise. "Climatic and weather factors affecting fire occurrence and behavior." In: *Developments in environmental science* 8 (2008), pp. 37–59.

[109] Wu Rihan, Hongyan Zhang, Jianjun Zhao, Yu Shan, Xiaoyi Guo, Hong Ying, Guorong Deng, and Hui Li. "Promote the advance of the start of the growing season from combined effects of climate change and wildfire." In: *Ecological Indicators* 125 (2021), p. 107483.

[110] Anthony L Westerling, Alexander Gershunov, Timothy J Brown, Daniel R Cayan, and Michael D Dettinger. "Climate and wildfire in the western United States." In: *Bulletin of the American Meteorological Society* 84.5 (2003), pp. 595–604.

[111] John T Abatzoglou and A Park Williams. "Impact of anthropogenic climate change on wildfire across western US forests." In: *Proceedings of the National Academy of sciences* 113.42 (2016), pp. 11770–11775.

[112] X Yue, LJ Mickley, JA Logan, RC Hudman, M Val Martin, and RM Yantosca. "Impact of 2050 climate change on North American wildfire: consequences for ozone air quality." In: *Atmospheric Chemistry and Physics* 15.17 (2015), pp. 10033–10055.

[113] Juli G Pausas and Susana Paula. "Fuel shapes the fire–climate relationship: evidence from Mediterranean ecosystems." In: *Global Ecology and Biogeography* 21.11 (2012), pp. 1074–1082.

[114] Jon E Keeley and Alexandra D Syphard. "Climate change and future fire regimes: examples from California." In: *Geosciences* 6.3 (2016), p. 37.

[115] David MJS Bowman, Jennifer K Balch, Paulo Artaxo, William J Bond, Jean M Carlson, Mark A Cochrane, Carla M dAntonio, Ruth S DeFries, John C Doyle, Sandy P Harrison, et al. “Fire in the Earth system.” In: *science* 324.5926 (2009), pp. 481–484.

[116] Henri D Grissino Mayer and Thomas W Swetnam. “Century scale climate forcing of fire regimes in the American Southwest.” In: *The Holocene* 10.2 (2000), pp. 213–220.

[117] Jeremy S Littell, Donald McKenzie, David L Peterson, and Anthony L Westerling. “Climate and wildfire area burned in western US ecoprovinces, 1916–2003.” In: *Ecological applications* 19.4 (2009), pp. 1003–1021.

[118] Chris Stockdale, Mike Flannigan, and Ellen Macdonald. “Is the END (emulation of natural disturbance) a new beginning? A critical analysis of the use of fire regimes as the basis of forest ecosystem management with examples from the Canadian western Cordillera.” In: *Environmental Reviews* 24.3 (2016), pp. 233–243.

[119] Marc-André Parisien, Susan Snetsinger, Jonathan A Greenberg, Cara R Nelson, Tania Schoennagel, Solomon Z Dobrowski, and Max A Moritz. “Spatial variability in wildfire probability across the western United States.” In: *International Journal of Wildland Fire* 21.4 (2012), pp. 313–327.

[120] Max A Moritz, Marc-André Parisien, Enric Batllori, Meg A Krawchuk, Jeff Van Dorn, David J Ganz, and Katharine Hayhoe. “Climate change and disruptions to global fire activity.” In: *Ecosphere* 3.6 (2012), pp. 1–22.

[121] Florent Mouillot, Serge Rambal, and Richard Joffre. “Simulating climate change impacts on fire frequency and vegetation dynamics in a Mediterranean-type ecosystem.” In: *Global Change Biology* 8.5 (2002), pp. 423–437.

[122] Boris Thies. “Machine learning wildfire susceptibility mapping for Germany.” In: *Natural Hazards* 121.11 (2025), pp. 12517–12530.

[123] Uro Durlevi, Velibor Ili, and Aleksandar Valjarevi. “Wildfire susceptibility mapping using deep learning and machine learning models based on multi-sensor satellite data fusion: a case study of Serbia.” In: *Fire* 8.10 (2025), p. 407.

[124] Seyd Teymoor Seydi, Yousef Kanani-Sadat, Mahdi Hasanlou, Roya Sahraei, Jocelyn Chanussot, and Meisam Amani. “Comparison of machine learning algorithms for flood susceptibility mapping.” In: *Remote Sensing* 15.1 (2022), p. 192.

[125] Mehmet smail Gürsoy, Osman Orhan, and Senem Tekin. "Creation of wildfire susceptibility maps in the Mediterranean Region (Turkey) using convolutional neural networks and multilayer perceptron techniques." In: *Forest Ecology and Management* 538 (2023), p. 121006.

[126] Omid Ghorbanzadeh, Khalil Valizadeh Kamran, Thomas Blaschke, Jagannath Aryal, Amin Naboureh, Jamshid Einali, and Jinhu Bian. "Spatial prediction of wildfire susceptibility using field survey GPS data and machine learning approaches." In: *Fire* 2.3 (2019), p. 43.

[127] Marcos Rodrigues and Juan De la Riva. "An insight into machine-learning algorithms to model human-caused wildfire occurrence." In: *Environmental Modelling & Software* 57 (2014), pp. 192–201.

[128] Ehsan Masoudian, Ali Mirzaei, and Hossein Bagheri. "Assessing wildfire susceptibility in Iran: Leveraging machine learning for geospatial analysis of climatic and anthropogenic factors." In: *Trees, Forests and People* 19 (2025), p. 100774.

[129] Udaya Bhasker Cheerala, Varun Teja Chirukuri, Venkata Akhil Kumar Gummadi, Jintu Moni Bhuyan, and Praveen Damacharla. "Probabilistic Wildfire Susceptibility from Remote Sensing Using Random Forests and SHAP." In: *arXiv e-prints* (2025), arXiv–2511.

[130] Shruti Sachdeva, Tarunpreet Bhatia, and etAK Verma. "GIS-based evolutionary optimized Gradient Boosted Decision Trees for forest fire susceptibility mapping." In: *Natural Hazards* 92.3 (2018), pp. 1399–1418.

[131] Weiting Yue, Chao Ren, Yueji Liang, Jieyu Liang, Xiaoqi Lin, Anchao Yin, and Zhenkui Wei. "Assessment of wildfire susceptibility and wildfire threats to ecological environment and urban development based on GIS and multi-source data: A case study of Guilin, China." In: *Remote Sensing* 15.10 (2023), p. 2659.

[132] Sepideh Tavakkoli Piralilou, Golzar Einali, Omid Ghorbanzadeh, Thimmaiah Gudiyangada Nachappa, Khalil Gholamnia, Thomas Blaschke, and Pedram Ghamisi. "A Google Earth Engine approach for wildfire susceptibility prediction fusion with remote sensing data of different spatial resolutions." In: *Remote sensing* 14.3 (2022), p. 672.

[133] Khabat Khosravi, Mahdi Panahi, Ali Golkarian, Saskia D Keesstra, Patricia M Saco, Dieu Tien Bui, and Saro Lee. "Convolutional neural network approach for spatial prediction of flood hazard at national scale of Iran." In: *Journal of Hydrology* 591 (2020), p. 125552.

[134] Joaqun Bedia, Sixto Herrera, Andrea Camia, Jose Manuel Moreno, and Jose Manuel Gutiérrez. "Forest fire danger projections in the Mediterranean using ENSEMBLES regional climate change scenarios." In: *Climatic change* 122.1 (2014), pp. 185–199.

[135] Jiehang Deng, Weiming Wang, Guosheng Gu, Zhiqiang Chen, Jing Liu, Guobo Xie, Shaowei Weng, Lei Ding, and Chuan Li. "Wildfire susceptibility prediction using a multisource and spatiotemporal cooperative approach." In: *Earth Science Informatics* 16.4 (2023), pp. 3511–3529.

[136] Fahed El Amarty, Afaf Chakir, Sanju Purohit, Mohamed Elhag, Ferdaouss Lakhili, Lahcen Benaabidate, and Abderrahim Lahrach. "Comparative evaluation of machine learning models and remote sensing data for wildfire prediction in the Atlas middle forest, Morocco." In: *Earth Systems and Environment* (2025), pp. 1–20.

[137] R Subash, M Vijayabhama, M Kalpana, K Baranidharan, K Arunadevi, and M Sankavi. "Machine learning techniques for forest fire prediction: Current trends, challenges and future directions." In: *International System for Agricultural Science and Technology* (2025), pp. 1–20.

[138] Ivan P Malashin, Igor Masich, Vladimir Nelyub, Aleksei Borodulin, Andrei Gantimurov, and Vadim Tynchenko. "Assessing wildfire extents in Siberian forests using machine learning." In: *Scientific Reports* 15.1 (2025), p. 32834.

[139] Di Fan, Ayan Biswas, and James Paul Ahrens. *Explainable ai integrated feature engineering for wildfire prediction.* 2024.

[140] Shashank Bhardwaj. *Predicting the Containment Time of California Wildfires Using Machine Learning.* 2025.

[141] Suwei Yang, Massimo Lupascu, and Kuldeep S Meel. "Predicting forest fire using remote sensing data and machine learning." In: 35.17 (2021), pp. 14983–14990.

[142] rem Üstek, Miguel Arana-Catania, Alexander Farr, and Ivan Petrunin. *Deep autoencoders for unsupervised anomaly detection in wildfire prediction.* 2024.

[143] JD Vilaças Nogueira, EJ Solteiro Pires, Arsénio Reis, PB de Moura Oliveira, António Pereira, and João Barroso. "Forest fire risk prediction using machine learning." In: (2024), pp. 221–230.

[144] Jorge Caiado and Mariana Marques. "Predicting wildfire occurrences in Portugal using machine learning classification models." In: *Ecological Informatics* (2025), p. 103455.

[145] Fahad Shahzad, Kaleem Mehmood, Khadim Hussain, Ijlal Haidar, Shoaib Ahmad Anees, Sultan Muhammad, Jamshid Ali, Muhammad Adnan, Zhichao Wang, and Zhongke Feng. "Comparing machine learning algorithms to predict vegetation fire detections in Pakistan." In: *Fire Ecology* 20.1 (2024), p. 57.

[146] Ying Xie and Minggang Peng. "Forest fire forecasting using ensemble learning approaches." In: *Neural Computing and Applications* 31.9 (2019), pp. 4541–4550.

[147] JNS Rubí and Paulo RL Gondim. "A performance comparison of machine learning models for wildfire occurrence risk prediction in the Brazilian Federal District region." In: *Environment Systems and Decisions* 44.2 (2024), pp. 351–368.

[148] Guoli Zhang, Ming Wang, and Kai Liu. "Deep neural networks for global wildfire susceptibility modelling." In: *Ecological Indicators* 127 (2021), p. 107735.

[149] Abdelilah Et-taleby, Yassine Chaibi, Amine Allouhi, Mohammed Boussetta, and Mohamed Benslimane. "A combined convolutional neural network model and support vector machine technique for fault detection and classification based on electroluminescence images of photovoltaic modules." In: *Sustainable Energy, Grids and Networks* 32 (2022), p. 100946.

[150] Mohamad Alipour, Inga La Puma, Joshua Picotte, Kasra Shamsaei, Eric Rowell, Adam Watts, Branko Kosovic, Hamed Ebrahimian, and Ertugrul Taciroglu. "A multimodal data fusion and deep learning framework for large-scale wildfire surface fuel mapping." In: *Fire* 6.2 (2023), p. 36.

[151] Sally S-C Wang and Yuxuan Wang. "Quantifying the effects of environmental factors on wildfire burned area in the south central US using integrated machine learning techniques." In: *Atmospheric Chemistry and Physics* 20.18 (2020), pp. 11065–11087.

[152] Samiksha Bharti, Prokriti Das, and Bappaditya Das. "The Robustness Gap: Real-World Challenges in AI-Based Flood Severity Assessment." In: *Asian Journal of Research in Computer Science* 19.1 (2026), pp. 51–70.

[153] Brent Kong. "A comparative analysis of machine learning models for wildfire prediction." In: *The National High School Journal of Science* (2024).

[154] Mohammad Khaled Al-Bashiti and MZ Naser. "Machine learning for wildfire classification: Exploring blackbox, eXplainable, symbolic, and SMOTE methods." In: *Natural Hazards Research* 2.3 (2022), pp. 154–165.

[155] Caisu Meng and Hailiang Jin. “A comparison of machine learning models for predicting flood susceptibility based on the enhanced NHAND method.” In: *Sustainability* 15.20 (2023), p. 14928.

[156] Svetlana Illarionova, Dmitrii Shadrin, Fedor Gubanov, Mikhail Shutov, Usman Tasuev, Ksenia Evteeva, Maksim Mironenko, and Evgeny Burnaev. “Exploration of geo-spatial data and machine learning algorithms for robust wildfire occurrence prediction.” In: *Scientific Reports* 15.1 (2025), p. 10712.

[157] Satyanarayana Raju Vuddaraju, Sri Hari Sai Saran Polisetty, and J Shanthini. “Forest Fire Prediction: A Hybrid Approach Using Context Based Learning, LSTM, and XGBoost.” In: (2025), pp. 394–407.

[158] A Emily Jenifer and Sudha Natarajan. “DeepFlood: A deep learning based flood detection framework using feature-level fusion of multi-sensor remote sensing images.” In: *Journal of Universal Computer Science* 28.3 (2022), p. 329.

[159] Angga Bayu Santoso, Imas Sukaesih Sitanggang, and Medria Kusuma Dewi Hardhienata. “Development of a Prediction Model for Potential Forest and Land Fires using Machine Learning Algorithms Based on Patrol Data.” In: *JURNAL INFOTEL* 16.3 (2024), pp. 642–659.

[160] Maryam Nasourinia and Kalpdrum Passi. “Wildfire Prediction in British Columbia Using Machine Learning and Deep Learning Models: A Data-Driven Framework.” In: *Big Data and Cognitive Computing* 9.11 (2025), p. 290.