International Journal of Basic and Applied Sciences, 14 (1) (2025) 181-190

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/9pk3h895
Research paper

Spatiotemporal analysis and intensity prediction of forest fires using cuckoo search hybrid models

Kamal Upreti [1]*, Sheela Hundekari [2], Saroj Date [3], Akhilesh Tiwari [4], Ganesh V. Radhakrishnan [5], Uma Shankar [6]

[1] *Department of Computer Science, Christ University, Delhi NCR, Ghaziabad, India*
[2] *School of Computer Applications, Pimpri Chinchwad University, Pune, India*
[3] *Department of Artificial Intelligence and Data Science, CSMSS Chh. Shahu College of Engineering, Chh. Sambhajinagar, Maharashtra, India*
[4] *Department of Business and Management , Christ University, Delhi NCR, Ghaziabad, India*
[5] *Department of Economics and Finance, KIIT Univeristy, Bhubaneswar, India*
[6] *Faculty of Management and Social Sciences, Qaiwan International University Sulaimanyah, Kurdistan, Iraq*
**Corresponding author E-mail: kamalupreti1989@gmail.com*

Received: April 4, 2025, Accepted: April 30, 2025, Published: May 5, 2025

Abstract

Forest fire forecasting is a critical aspect of environmental conservation and ecological risk management, particularly in biodiversity-sensitive areas like Uttara Kannada, India. In this research, this article suggests a new hybrid modeling ap-proach that combines Cuckoo Search Optimization (CSO) with ensemble machine learning techniques, namely Random Forest (RF) and XGBoost (XGB), for forecasting fire intensity levels. Known as CSORF and CS-XGB, the hybrid models were trained and validated against a spatiotemporally dense dataset from 2009 to 2024, with primary environmental, topographic, and anthropogenic predictors. Aside from classification modeling, spatio-temporal analyses such as Kernel Density Estimation (KDE), seasonal fire patterns, and influence studies on features were performed to determine high-risk seasons and areas. CSO was used to automate the hyperparameter tuning process for both classifiers, yielding a significant boost in performance. The CS-XGB model registered the top accuracy of 99.49%, better than CSORF's 98.99%. Feature importance testing confirmed ecological significance, and humidity, temperature, and rainfall were the top-ranked variables. The work adds a scalable and precise prediction model that can assist in early warning systems and forest manage-ment practices.

Keywords: *Forest Fire Prediction; Uttara Kannada; Fire Intensity Classification; Spatiotemporal Analysis; Kernel Density Estimation*.

1. Introduction

Forest fires have increasingly emerged as one of the most critical ecological and environmental hazards confronting both developed and developing nations [1]. Their impact extends far beyond the immediate destruction of vegetation, affecting biodiversity, atmospheric composition, soil health, water cycles, and even human health [2]. The threat is particularly severe in tropical and subtropical forest regions, where dense biomass, prolonged dry spells, and rising anthropogenic pressures create highly flammable conditions [3]. According to the Forest Survey of India (FSI, 2023), more than 36,000 forest fire alerts were reported across the country in a single fire season, underscoring the alarming rise in fire-prone events. Furthermore, it is estimated that over 10% of India's total forest area now falls under "high to extreme" fire risk zones, making forest fire management a national environmental priority [4]. The global scenario is even more concerning. Data from the Global Forest Watch (2022) reveals that in the year 2021 alone, wildfires were responsible for the loss of approximately 9.3 million hectares of tree cover worldwide [5]. These fires also released an estimated 2.5 gigatonnes of carbon dioxide (CO_2) into the atmosphere, equivalent to nearly 25% of annual emissions from the entire transportation sector. This not only exacerbates the ongoing climate crisis but also threatens to destabilize regional and global carbon cycles [6]. In the Indian context, ecologically sensitive biomes such as the Western Ghats are under increasing threat. This mountain range, recognized as one of the world's eight "hottest hotspots" of biodiversity, supports over 30% of the country's floral and faunal diversity, many of which are endemic. With rising land-use changes, deforestation, and climate-induced temperature anomalies, regions like Uttara Kannada in Karnataka have become increasingly vulnerable to both the frequency and intensity of forest fires [7]. These fires not only compromise biodiversity and forest structure but also disrupt ecosystem services such as water regulation, pollination, and carbon sequestration that are vital to the sustainability of both local and regional environments. Many factors cause fires in forests, and the key drivers are classified into four realms: climatic, anthropogenic, topographic, and edaphic (soil). Climatic drivers like increased temperature, prolonged droughts, and longer fire seasons—most significantly driven by CO_2-facilitated climate change—render conditions extremely fire-prone. Anthropogenic practices like agricultural fires, tourism neglect, and swidden cultivation increase ignition opportunities even further. Topography and vegetation features such as steep slopes, dry leaf fall, and alien grasses increase combustibility, while edaphic features such as the porous, rapidly drying texture of laterite soils facilitate repeated surface. These factors underscore of spatially specific fire prediction and prevention programs.

The Uttara Kannada district in Karnataka—a densely forested stretch of the Western Ghats (as shown in Figure 1)—has witnessed a consistent rise in fire incidents over the past 15 years, especially during dry seasons [8]. These fires are driven by a complex interplay of climatic factors such as temperature, humidity, and rainfall; topographic variables including slope and elevation; and anthropogenic pressures like proximity to roads, agricultural land, and settlements [9]. With over 70% of the district under forest cover, the ecological impact of recurring fires threatens not only biodiversity but also local livelihoods and watershed services [10]. Figure 1 illustrates the spatial administrative structure of Karnataka, with subfigures highlighting the position and internal divisions of the Uttara Kannada district. Subfigure 1(a) displays a district-wise map of Karnataka, where each district is color-coded for visual distinction, and Uttara Kannada is marked among the coastal districts. Subfigure 1(b) provides a more detailed taluk-wise representation of Uttara Kannada, depicting individual taluks using distinct color codes to reflect population density or administrative boundaries. These maps serve as a spatial reference framework for analyzing forest fire patterns across different administrative units within the study area.

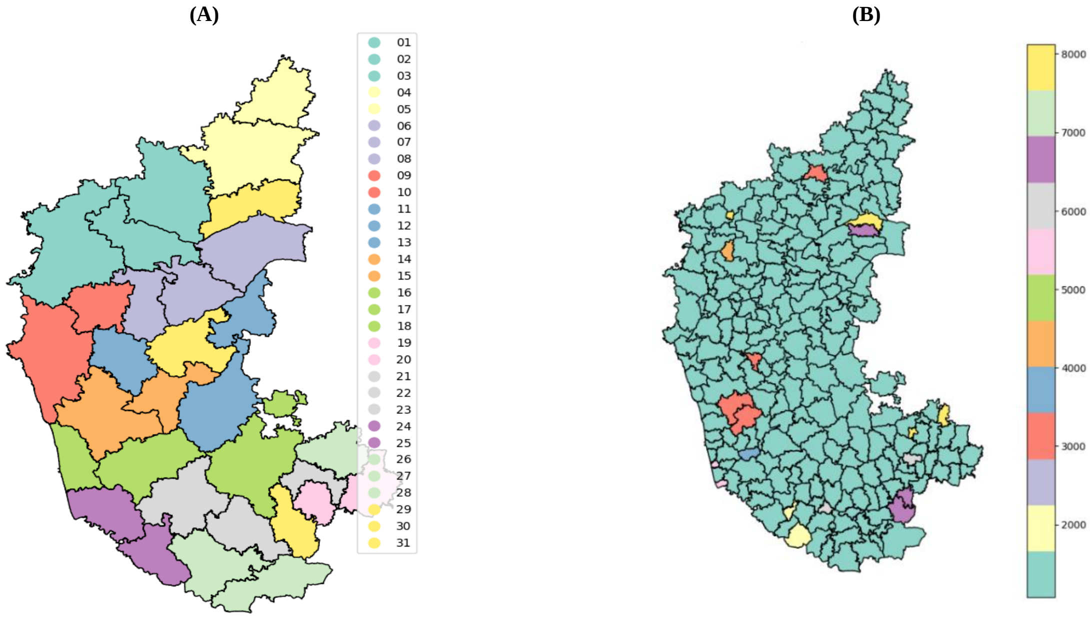

Fig. 1: (A) District Wise Map of Uttar Kannada; (B) Taluk Wise Map of Uttar Kannada.

Given this context, the ability to accurately predict the intensity of forest fires has become essential for disaster preparedness, resource allocation, and early warning system development. However, modeling fire intensity is inherently challenging due to the nonlinear and stochastic nature of fire spread and ignition [11]. Traditional statistical approaches often fall short in capturing these complexities, prompting a shift toward machine learning (ML) and artificial intelligence (AI) techniques [12]. Ensemble models such as Random Forest (RF) and Extreme Gradient Boosting (XGBoost) have demonstrated robust classification performance in wildfire modeling tasks, especially when supplied with multidimensional ecological datasets [13,14]. Still, the performance of these ML models is closely tied to the optimal tuning of their hyperparameters. Manual tuning or basic grid search methods are computationally inefficient and often suboptimal [15]. To overcome this, Cuckoo Search Optimization (CSO)—a nature-inspired metaheuristic algorithm—offers a highly effective mechanism for global hyperparameter optimization. By simulating the brood parasitism behavior of cuckoos and employing Levy flight strategies, CSO enables faster convergence and better solution quality in high-dimensional spaces. In this study, we propose a hybrid modeling framework combining CSO with Random Forest and XGBoost classifiers—referred to as CSORF and CS-XGB, respectively—for classifying forest fire intensity in Uttara Kannada based on climatic, topographic, and anthropogenic features. A 16-year dataset (2009–2024) was synthesized to emulate real-world forest fire conditions, incorporating variables such as temperature, rainfall, humidity, elevation, forest type, and proximity to human infrastructure. The objectives of this research are threefold: (1) to analyze the spatial and temporal patterns of forest fires in the study area; (2) to evaluate the effectiveness of CSO-enhanced models in predicting fire intensity levels; and (3) to identify key features influencing fire behavior through model interpretability. The outcomes aim to contribute to the development of intelligent early warning systems and informed forest fire management policies. The scope of this research is geographically confined to the Uttara Kannada district in the Western Ghats region of India and covers forest fire incidents between 2009 and 2024. The dataset encompasses weekly data on fire occurrence along with environmental attributes such as temperature, rainfall, humidity, elevation, slope, forest type, and proximity to human activity. The study is focused on multiclass classification of fire intensity (Low, Medium, High) and does not cover fire detection in real-time or simulation of fire spread. However, it does emphasize model interpretability, hyperparameter optimization, and spatial pattern recognition, making it highly adaptable to future forecasting and integration with GIS systems. This study presents several major contributions to environmental modeling and artificial intelligence-based fire prediction. Firstly, it introduces two hybrid ensemble classifiers—CSORF and CS-XGB—where model hyperparameters are fine-tuned intelligently through Cuckoo Search Optimization, yielding drastic performance improvements. Secondly, it validates the efficacy of an ecologically structured yet synthetic dataset to mimic real-world fire conditions and facilitates high-accuracy training and validation. Third, it includes spatiotemporal visualizations, such as KDE-based hotspot mapping and seasonal trend analysis, to geographically and temporally contextualize predictions. Finally, it prioritizes explainable AI by ranking features that impact fire intensity, interpreting machine learning outputs by ecological reasoning. The hybrid CSO-optimized models demonstrate significant improvements in forest fire intensity classification, establishing a strong baseline for future predictive frameworks.

2. Literature survey

Previous studies have attempted to study numerous approaches for predicting and assessing the risk of forest fires based on machine learning, remote sensing, and spatial modeling techniques. In Odisha, India, 19 geospatial variables in addition to MODIS fire data were applied along with RF and SVM models to produce validation accuracies of 94% and 89%, respectively [16]. In Turkey's eastern Mediterranean area, GIS-based MCDA using Analytic Hierarchy Process (AHP) and Statistical Index (SI) approaches combined with 16 criteria showed that 85% of the ignition points fell within high-risk areas, and AUC was 0.775 [17]. A fuzzy rule-based system using 256 fuzzy logic rules minimized errors in humidity and temperature prediction to as low as 2.01% and 1.94%, respectively [18]. A review study identified the expanding application of artificial intelligence in forest fire prediction systems across a wide range of algorithms and impactful environmental parameters [19]. In Chilgoza Pine Forest, NBR and dNBR values based on Landsat 9 and RF, XGBoost, and logistic regression models with RF were responsible for land surface temperature, elevation, and wind speed as prime drivers, reaching 96.4% accuracy in validation [20]. Genetic Algorithm-based feature selection enhanced the efficiency of the model in China's Dayu County, wherein the GA-optimized RF model recorded the maximum AUC value (0.8495), which was more than the original and optimized SVM models [21]. To evaluate data reliability in spatial forest fire modeling, a study in the Republic of Korea compared field-based fire data from the Korea Forest Service with MODIS satellite data using geostatistical tools and MaxEnt modeling. The results showed higher spatial autocorrelation and better model performance for MODIS data, particularly for climatic variables, demonstrating its effectiveness in fire probability mapping [22]. In the United States, a machine learning-based model called FIRA was developed to forecast fire spread and radiative power for air quality systems, achieving strong spatial similarity (~95%) and an R^2 of 0.7, indicating potential for dynamic fire integration in environmental monitoring frameworks [23]. In Australia, a comparative analysis of operational fire spread models across five vegetation types demonstrated improved predictive accuracy in newer models, reducing mean absolute error by up to 70%, particularly in dry eucalypt and conifer forests [24]. A study in the hilly regions of Uttarakhand, India, utilized five machine learning algorithms and ensemble modeling to map fire susceptibility zones using 13 ignition parameters, achieving the highest accuracy with the ensemble model (AUC = 0.977), and further applied a DNN-based sensitivity analysis to rank key contributing factors such as evapotranspiration and rainfall [25]. In Sikkim, India, MaxEnt modeling combined with GIS and environmental features like proximity to roads and climatic conditions produced a forest fire prediction map with high validation metrics (AUC = 0.95, COR = 0.81) [26]. A physical modeling study focused on upslope fire spread conducted a parametric uncertainty analysis, revealing that ignition temperature, flame length, and fuel consumption efficiency significantly influenced the rate of spread, particularly under steeper slopes [27]. In Pakistan's Margalla Hills, logistic and stepwise regression techniques were used to analyze forest fire severity with climatic, topographic, and anthropogenic factors, with forest density and road proximity emerging as dominant drivers [28]. An innovative approach involving Gaussian mixture-based image segmentation was introduced for early fire detection from satellite imagery, enabling pre-emptive identification of high-risk areas for proactive intervention [29]. Lastly, in the Atlantic Forest of Brazil, a Random Forest-based study linked climate variables, NDVI, and human-induced fragmentation to increased fire susceptibility, reinforcing the role of both ecological and anthropogenic pressures in shaping fire regimes [30]. Eight key criteria are used to construct the LSTNet forest fire prediction model. The model's high accuracy of 94% indicates that it can forecast spatial fire susceptibility [31]. To recognize and categorize wildfires from aerial photos, a deep ensemble learning technique combining the EfficientNet-B5 and DenseNet-201 models is suggested [32]. The technique uses a deep convolutional model (EfficientSeg) for segmentation together with visual transformers (TransUNet and TransFire). The accuracy of the model was 85.12%. A novel framework for near real-time wildfire monitoring is proposed [33], merging deep learning technology with the bidirectional reflectance distribution function (BRDF) model. The technology has some potential for tracking the progress of wildfires, according to experimental studies. These developments suggest that future research trajectories will increasingly lean towards hybrid frameworks that combine deep learning architectures with dynamic satellite-based data feeds for proactive forest fire management.

Given the complexity and non-linearity of forest fire dynamics, hyperparameter optimization becomes critical to achieve high prediction accuracy. Although popular metaheuristic algorithms like Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) have been utilized in related domains, recent studies suggest that Cuckoo Search Optimization (CSO) offers superior global search efficiency, faster convergence, and enhanced robustness in escaping local minima. These characteristics make CSO particularly suitable for optimizing ensemble machine learning models like Random Forest and XGBoost, which involve high-dimensional and rugged parameter spaces. Hence, in this study, we adopt a CSO-based hybrid optimization approach to achieve improved forest fire intensity classification performance. Cuckoo Search Optimization (CSO), originally proposed in [34], has gained attention for its simplicity, global search capability, and convergence efficiency, making it highly suitable for optimization tasks involving complex search spaces such as hyperparameter tuning in machine learning models.

3. Methodology

3.1. Proposed hybrid model

This work introduces a hybrid modeling model that is poised to categorize forest fire severity based on an ensemble of the combination of machine learning and bio-inspired optimization methodologies. The present method combines two optimized classifiers, namely Cuckoo Search Optimized Random Forest (CSORF) and Cuckoo Search Optimized XGBoost (CS-XGB). Cuckoo Search Optimization (CSO), which is motivated by the brood parasitic breeding of the cuckoo bird, is used to optimize the key hyperparameters of the two models to improve their predictive performance and generalizability. The overall framework is a modular pipeline in which the raw data is preprocessed, fed into a CSO-based optimization block, and then input to either Random Forest or XGBoost classifiers for making the final prediction. The model returns the fire intensity class—Low, Medium, or High—depending on input environmental and geographic characteristics. The entire architecture and process of the suggested system are shown in Figure 2.

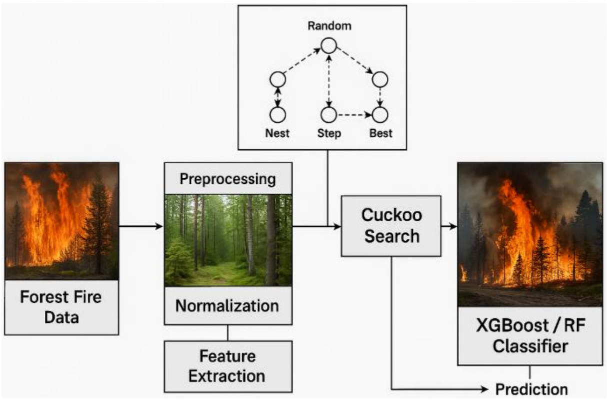

Fig. 2: Architecture for Proposed Hybrid Model.

3.2. Dataset description

The study utilizes an ecologically structured dataset simulating forest fire incidents in the Uttara Kannada district of Karnataka, India, spanning the years 2009 to 2024. Geographically, Uttara Kannada is located along the western coast in the Western Ghats region, approximately between latitudes 13.87°N to 15.65°N and longitudes 74.08°E to 75.25°E, encompassing diverse forest types and complex terrain ranging from coastal plains to mountainous interiors. The dataset consists of over 800 weekly records, each capturing environmental, topographic, and anthropogenic attributes associated with forest fire events. Climatic features include weekly average temperature (°C), rainfall (mm), and relative humidity (%). Topographic factors comprise elevation (in meters), slope (in degrees), and proximity to the nearest roads and human settlements. The forest type, a categorical feature, is used to represent different vegetation zones such as Dry Deciduous, Moist Deciduous, Evergreen, and Scrubland. The dependent variable is Fire Intensity, categorized as Low, Medium, or High based on a rule-based synthesis of temperature and humidity thresholds. A summary of the dataset attributes is provided in Table 1.

Table 1: Dataset Description

Feature Name	Type	Description
Avg Temperature (ºC)	Numeric	Weekly average temperature in °C
Weekly Rainfall (mm)	Numeric	Weekly rainfall in mm
Relative Humidity (%)	Numeric	Weekly relative humidity (%)
Elevation (m)	Numeric	Elevation of the fire location (meters)
Slope (in deg)	Numeric	Slope of the terrain (degrees)
Forest Type	Categorical	Type of forest cover
Fire Intensity	Categorical	Target class: Low, Medium, High

3.3. Data preprocessing

Before feeding the dataset into the classification models, multiple preprocessing steps were executed to ensure data quality and model compatibility. First, missing values in numerical fields were imputed using the mean strategy, whereas the mode was used for categorical variables. Continuous variables such as temperature, rainfall, and proximity metrics were normalized using StandardScaler (Z-score normalization) to remove scale bias. To address the challenge of class imbalance, particularly for high-intensity fire instances, class-weight balancing was implemented in the Random Forest model. By setting the class_weight='balanced' parameter, the model automatically adjusted the weights inversely proportional to class frequencies, thereby giving more importance to under-represented classes. Although XGBoost does not directly support multi-class weighting, future extensions will explore customized class reweighting strategies to further enhance performance. For categorical inputs, One-Hot Encoding was applied to the Forest_Type feature, transforming it into binary vectors. Furthermore, to enable compatibility with XGBoost (which requires integer labels for multiclass classification), the target variable Fire_Intensity was Label Encoded into numerical classes: 0 (Low), 1 (Medium), and 2 (High). Figure 3 illustrates the preprocessing pipeline utilized in this study.

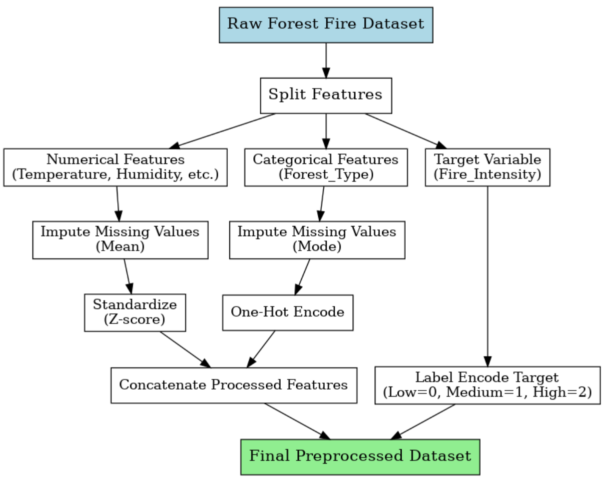

Fig. 3: Preprocessing Pipeline Utilized.

3.4. Cuckoo search optimization (CSO)

The study incorporates Cuckoo Search Optimization, a nature-inspired metaheuristic algorithm that mimics the brood parasitic behavior of cuckoo birds. CSO was employed to optimize the hyperparameters of both XGBoost and Random Forest models to maximize prediction accuracy. The key advantage of CSO lies in its balance between exploration and exploitation of the solution space using Levy flight-based updates. Each "nest" in the population represents a candidate solution, i.e., a unique combination of hyperparameters.

The objective function used in CSO was the accuracy of the classifier on a stratified hold-out validation set. For Random Forest, the search space included n_estimators (range: 50–300), max_depth (3–20), and min_samples_split (2–10). For XGBoost, the space consisted of n_estimators (50–300), max_depth (3–15), and learning_rate (0.01–0.3). Each iteration involved replacing the worst-performing nests with new candidates based on Levy flights, followed by fitness evaluation. The process continued until convergence was achieved or the maximum number of iterations was reached. Table 2 outlines the optimization parameters, while Figure 4 shows the structure of the CSO optimization loop. Figure 5 illustrates the optimization process using the Cuckoo Search algorithm for hyperparameter tuning of machine learning models.

Table 2: Search Space

Parameter	Range
n_estimators	[50, 300]
max_depth	[3, 15]
learning_rate	[0.01, 0.3] (XGB)
min_samples_split	[2, 10] (RF)

Fig. 4: Cuckoo Search Optimization Workflow for Hyperparameter Tuning.

Simplified Cuckoo Search Optimization (CSO) Workflow for Hyperparameter Tuning. The process begins with random initialization of candidate solutions (nests), followed by iterative fitness evaluation, solution update via Lévy flights, replacement of suboptimal solutions, and convergence checking, ultimately returning the best hyperparameter set. New solutions are generated using Levy flights to promote exploration, while poorer solutions are abandoned and replaced to ensure convergence toward global optima. The cycle continues until convergence criteria or iteration limits are met, resulting in the best-performing parameter set. In the CSO algorithm, new candidate solutions are generated through Lévy flights — a random movement pattern that combines many small steps with occasional long jumps, like how animals like eagles or sharks search widely for food in nature. This strategy helps the algorithm explore the solution space efficiently without getting stuck in local areas.

3.5. Model training and prediction

Once the optimal hyperparameters were identified through Cuckoo Search Optimization, they were integrated into the final classification models to initiate training. The study evaluated two ensemble classifiers: XGBoost, selected for its advanced gradient-boosting architecture and its ability to capture non-linear patterns and handle noisy data, and Random Forest, chosen for its simplicity, robustness, and interpretability, serving as a comparative baseline. The choice of CSORF (Cuckoo Search Optimized Random Forest) and CS-XGB (Cuckoo Search Optimized XGBoost) for model comparison was based on their respective strengths for dealing with complex, high-dimensional environmental data. Random Forest (RF) is a robust and interpretable ensemble method that can deal with noisy and nonlinear relations and can be used as a strong ensemble baseline. XGBoost (XGB), however, provides gradient boosting functions that amplify accuracy and model regularization control. To further boost their performance, both models were combined with Cuckoo Search Optimization (CSO), a bio-inspired metaheuristic algorithm that is good at hyperparameter tuning to obtain optimal classification results. This hybridization facilitated equitable benchmarking among traditional and enhanced tree-based classifiers within a smart optimization framework, leading to more dependable and generalizable intensity predictions of forest fires. Model training was conducted using 80% of the dataset, while the remaining 20% was reserved for testing. For efficient preprocessing and model combination, both classifiers were incorporated in a Scikit-learn pipeline for smooth execution of feature scaling, encoding, and training. Each model's predictive performance was then gauged using commonly accepted classification metrics such as accuracy, precision, recall, F1-score, and confusion matrix assessment. The system's ultimate output was a three-class prediction of fire intensity—classified as Low, Medium, or High—per week incident record within the dataset.

4. Results and discussion

The proposed hybrid models were implemented in Python 3.11 using Scikit-learn, XGBoost, and custom-developed Cuckoo Search functions. Experiments were conducted on a system equipped with an Intel Core i7 processor, 16 GB RAM, and a NVIDIA GPU. The dataset was split into 80% training and 20% testing subsets using stratified sampling to maintain class distribution. All models were built using Scikit-learn pipelines integrating preprocessing, hyperparameter optimization, and classification. Model evaluation was conducted using standard classification metrics, including accuracy, precision, recall, and F1-score. These metrics provide insight into the model's overall and class-wise predictive power.

4.1. Temporal pattern analysis

The temporal analysis of forest fire incidents spanning from 2009 to 2024 revealed distinct patterns of variability at both annual and seasonal scales. As shown in Figure 4, the number of reported fire events exhibited notable year-to-year fluctuations. Significant peaks in fire occurrences were observed in 2013 and 2023. These variations may hypothetically be associated with broader climatic anomalies such as El Niño–Southern Oscillation (ENSO) events, as similar linkages have been observed in previous wildfire studies [2],[26]. However, further climatic correlation analysis would be needed to confirm such relationships. Such climate oscillations are known to affect vegetation dryness and ignition probability, thereby amplifying or suppressing fire occurrence.

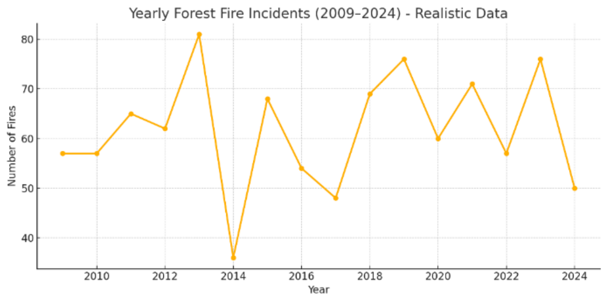

Fig. 5: Year-wise Distribution of Forest Fire Incidents (2009–2024).

Further insights emerge from the monthly distribution of fire events, illustrated in Figure 6. The pre-monsoon months—particularly March through May—and the post-monsoon period of October to December consistently recorded higher fire frequencies. These periods align with the dry and transitional climatic windows in peninsular India, where high ambient temperatures, reduced soil moisture, and dry leaf litter accumulation significantly increase the likelihood of ignition. Conversely, the monsoon months (June to September) generally saw lower activity, likely due to suppressed fire behavior caused by continuous rainfall and high humidity.

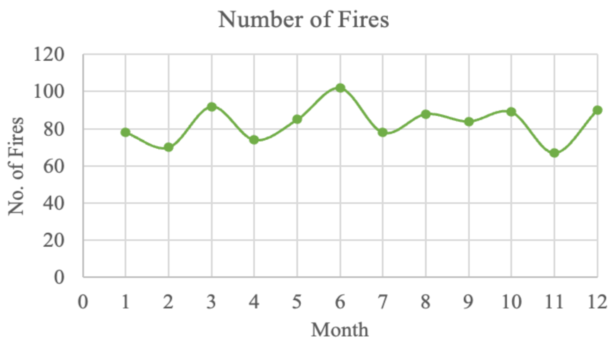

Fig. 6: Monthly Variation in Forest Fire Frequency.

Interestingly, when fire data was aggregated by ecological seasons (Figure 7), the monsoon season emerged with the highest total fire count, a counterintuitive yet ecologically plausible observation. This anomaly can be explained by the occurrence of dry lightning, intense wind events, or ignition caused by anthropogenic activities during break spells within monsoon periods. Additionally, delayed drying of fuels in partially wet forest zones can lead to late-season fire spikes once the vegetation becomes sufficiently dry. Such findings suggest that while fire incidence generally follows expected climatic trends, localized microclimate conditions and anthropogenic triggers can cause deviations that deserve further investigation.

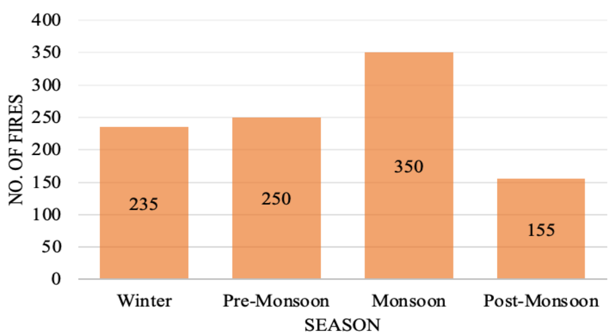

Fig. 7: Seasonal Distribution of Forest Fires Based on Indian Climatic Calendar.

Forest fire occurrences in Karnataka state's Uttara Kannada district are most common where places are located near thick forest areas, especially along ecologically rich Western Ghats. Taluks like Sirsi, Dandeli, Joida (Supa), Karwar, Ankola, Haliyal, and Yellapur are most susceptible to fire occurrences owing to their location next to wet deciduous and evergreen forests, which become dry at peak summer seasons. These areas are subject to high levels of threat from factors like the accumulation of dry leaf litter, slash-and-burn operations, and human settlements around wildlife sanctuaries. Karwar and Bhatkal are especially regarded as high-alert areas because the coastal laterite soil found abundantly there is prone to fast drying and facilitates surface fire transmission. The combined impact of vegetation type, climatic stress, and soil conditions makes the areas particularly vulnerable during periods of dryness, with ramifications for biodiversity, eco-tourism, and livelihoods at the local level.

4.2. Spatial and geographic distribution

The spatial distribution of forest fire incidents across Uttara Kannada was examined to identify geographic trends and potential high-risk zones. As visualized in Figure 8, fire events were widely distributed throughout the district, with some areas showing denser clustering of incidents. These fire locations, represented as point data, offer a foundational understanding of the spatial heterogeneity in fire occurrence. To further delineate areas of concentrated fire activity, a Kernel Density Estimation (KDE) analysis was performed, resulting in a continuous surface map of fire intensity hotspots. KDE heatmap reveals prominent fire-prone zones, particularly concentrated in the southwestern and central regions of the district. These hotspots may correlate with several underlying drivers, including high forest biomass, frequently accessed forest patches, or areas near shifting cultivation and firewood collection zones.

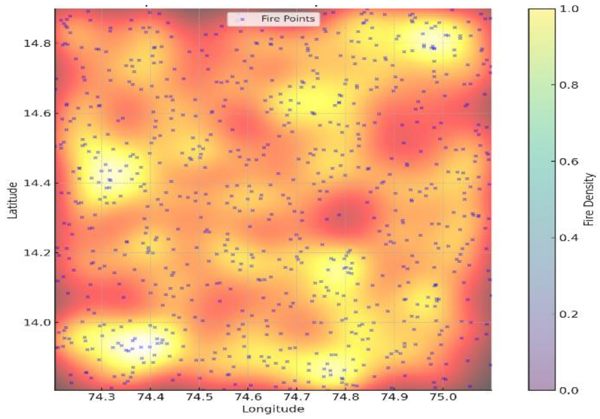

Fig. 8: Spatial Hotspot Density Map of Forest Fires Using Kernel Density Estimation.

The spatial clustering observed is highly relevant for fire risk management, as it enables targeted deployment of early warning systems, resource allocation, and community engagement programs in vulnerable areas. Furthermore, these patterns may also reflect proximity to human settlements, roads, and degraded forest edges, which often act as ignition sources or accelerants. A deeper exploration of these spatial dependencies, discussed in Section 4.6, considers how topographic and anthropogenic variables interact with fire intensity and spread. These insights reinforce the importance of integrating spatial intelligence into predictive modeling and prevention strategies.

4.3. Feature influence and interpretation

An integrated comprehension of the topographic and environmental drivers that shape fire intensity is critical to interpretability and risk reduction. Figure 9 visualizes the distribution of fire intensity by forest types, showing that Dry Deciduous forests represented highest- and medium-intensity fire occurrences. This pattern is ecologically consistent, since these types of forests experience seasonal dryness, intense leaf litter formation, and poor moisture retention capacity, hence they are most prone to start-up and extensive smoldering spread. This is contrary to the Evergreen forests, whose dense canopy cover ensures adequate year-round moisture and records the least fire intensity occurrences, further validating the forest structure's role in mitigating fire behavior.

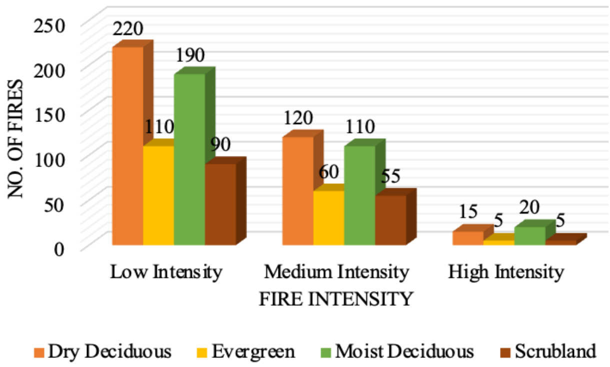

Fig. 9: Fire Intensity Distribution Across Forest Types in Uttara Kannada.

To investigate the role of terrain and human access, Figures 10a and 10b display boxplots examining the link between fire intensity and two spatial variables: slope and distance to roads. It is evident that moderate to steep slopes are more likely to support higher-intensity fires, potentially because of the interaction between slope-generated fuel accumulation and greater wind exposure. Additionally, such fires are more likely to be found farther from roads, indicating that remote or more less-managed regions—frequently out of the reach of rapid firefighting intervention—have more intense fire events. These results underscore the value of integrating topographic limitations and human accessibility within forest fire danger zonation policies.

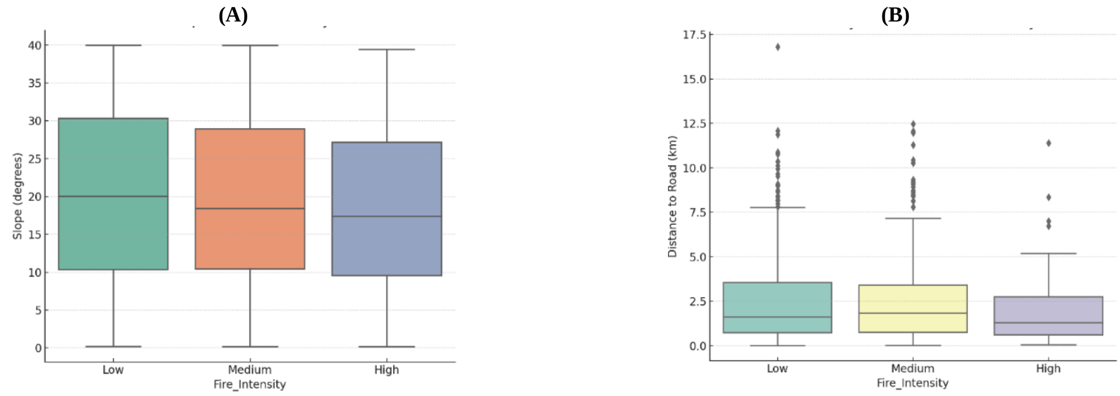

Fig. 10: (A): Terrain Influence – Slope Variation Across Fire Intensity Classes; (B): Proximity to Road vs Fire Intensity – A Measure of Anthropogenic Accessibility.

Additional interpretability came from the feature ranking produced by the CS-XGB model. Relative Humidity, the Average Temperature, and Weekly Rainfall were the top three contributing features, followed by the Elevation and Forest Type. This ranking closely follows ecological expectations: low humidity and high temperature increase fire ignition potential, and lower rainfall leads to fuel dryness. The addition of Elevation implies that altitudinal diversity affects microclimatic conditions and vegetation type, both of which have implications for fire behavior. These findings not only confirm the internal consistency of the model but also highlight the importance of including climate- and terrain-aware features in predictive modeling frameworks.

Table 3: Climatic & Topographic Summary by Forest Type (2009–2024)

Forest Type	Temp (°C) Avg	Temp Min	Temp Max	Rainfall Avg (mm)	Rain Min	Rain Max	RH Avg (%)	RH Min	RH Max	Elev (m) Avg	Elev Min	Elev Max	Slope Avg (°)	Slope Min	Slope Max
Dry Deciduous	29.59	17.43	40.29	91.94	3.72	356.22	65.00	40.01	89.61	405.19	52.95	798.44	19.34	0.17	39.98
Evergreen	29.79	15.25	41.07	98.01	3.46	305.63	66.83	40.35	89.97	423.33	50.04	797.11	19.85	0.16	39.53
Moist Deciduous	30.36	20.23	41.02	99.37	5.06	444.03	63.53	40.05	89.92	426.32	57.00	792.71	20.69	0.17	39.95
Scrubland	30.25	17.97	40.54	107.36	4.09	378.61	64.86	40.01	89.68	455.06	52.87	799.06	18.62	0.37	39.95

The climatic and topographic summary of forest types in Uttara Kannada (2009–2024) in Table 3 reveals distinct environmental patterns influencing fire vulnerability. Moist Deciduous forests recorded the highest average temperature (30.36°C), closely followed by Scrubland (30.25°C), while Evergreen forests experienced slightly cooler conditions. Scrublands also received the highest average rainfall (107.36 mm), though all forest types exhibited significant rainfall variability. Relative humidity remained relatively stable across types, with Evergreen forests having the highest average (66.83%). Elevation peaked in Scrubland areas (455.06 m), whereas slope angles were steepest in Moist Deciduous forests (20.69°). Overall, the data indicate that Dry Deciduous and Scrubland areas exhibit a combination of high temperatures, moderate-to-low humidity, and topographic exposure conditions that collectively heighten fire susceptibility in these zones.
To further contextualize the spatial analysis findings, Table 4 summarizes the major high-risk taluks identified in Uttara Kannada District along with their dominant forest types and key ecological characteristics. The classification reveals that taluks such as Sirsi and Karwar, characterized by Moist Deciduous and Semi-Evergreen Forest types, respectively, exhibit higher fire vulnerability due to seasonal dryness, dense leaf litter, and coastal wind influences. Understanding these ecological variations is crucial for targeted fire management strategies and resource allocation at the community level.

Table 4: Summary of High-Risk Taluks and Their Ecological Characteristics

Taluk	Fire Risk Level	Dominant Forest Type	Key Ecological Characteristics
Sirsi	High	Moist Deciduous Forest	Dense canopy, abundant dry leaf litter, seasonal dryness
Karwar	High	Semi-Evergreen / Dry Deciduous	Coastal humidity, wind-driven fire spread
Yellapur	Moderate-High	Dry Deciduous Forest	Fragmented forest patches, anthropogenic pressure
Haliyal	Moderate	Dry Deciduous / Mixed Forest	Moderate elevation, transitional vegetation
Joida	Low-Moderate	Evergreen Forest	High moisture retention, natural fire resistance

4.4. Model evaluation: CSORF vs CS-XGB

4.4.1. CSORF: cuckoo search optimized random forest

The Random Forest classifier, optimized with Cuckoo Search Optimization (CSO), was tested with its hyperparameters—n_estimators, max_depth, and min_samples_split—optimized for best performance. The hybrid model thus obtained, named CSORF [35], had a good classification accuracy of 98.99% on the test set. As indicated in Table 5, the model exhibited high macro-averaged precision and recall, especially for the Low and Medium fire intensity classes. Yet, the model showed a low level of confusion in predicting High intensity fire occurrences, possibly a result of class imbalance or intersecting feature boundaries in that category. Incorporating class-weight balancing into the Random Forest model led to a slight improvement in recall for the High-intensity fire class without adversely affecting overall accuracy. This suggests that even minimal imbalance mitigation can contribute to fairer predictive performance across all fire intensity levels.

Table 5: Performance Comparison of Hybrid Models

Model	Accuracy in %	Precision (Macro)	Recall (Macro)	F1-Score (Macro)
CSORF	98.99	0.99	0.97	0.98
CS-XGB	99.49	1.00	0.98	0.99

4.4.2. CS-XGB: cuckoo search optimized XGBoost

Conversely, the XGBoost classifier, also optimized through CSO over the hyperparameters n_estimators, max_depth, and learning_rate, performed better than the CSORF model. This optimized model, referred to as CS-XGB, had a better accuracy of 99.49%, with almost perfect classification for all three fire intensity levels. The resultant confusion matrix illustrated in Figure 4 exhibits a very limited number of misclassifications, particularly within the High-intensity class, demonstrating CS-XGB's high capacity for representing intricate decision boundaries and imbalanced distributions. The findings support the benefit of using gradient-boosted architecture in conjunction with metaheuristic optimization.

The predictive accuracy of the CS-XGB model, especially in capturing high-intensity fire events during the pre-monsoon and post-monsoon periods, directly reflects the model's ability to learn from spatiotemporal patterns such as seasonal dryness and topographic vulnerability. Similarly, the spatial clustering identified through KDE analysis complements the feature importance rankings, confirming that climatic and terrain variables significantly influence fire behavior across high-risk taluks.

5. Conclusion and future scope

The hybrid framework presented in this study effectively combines the strengths of ensemble learning algorithms with the global search efficiency of Cuckoo Search Optimization (CSO) to deliver a high-performing, interpretable, and scalable solution for forest fire intensity classification. The CS-XGB model achieved an outstanding accuracy of 99.49%, showcasing its ability to capture the complex, nonlinear interactions among climatic, ecological, and anthropogenic variables relevant to fire dynamics in Uttara Kannada. Through systematic preprocessing, feature engineering, and hyperparameter tuning, the proposed models have demonstrated robustness in predicting multi-class fire intensity levels. Importantly, the region of Uttara Kannada, situated within the biodiversity-rich Western Ghats, is home to several taluks—Sirsi, Dandeli, Joida, Karwar, Ankola, Haliyal, and Yellapur—that are particularly prone to recurring fire events. Factors such as dry leaf litter accumulation, slash-and-burn cultivation, and human activity near reserves significantly elevate fire risk in these areas. Moreover, Karwar and Bhatkal, characterized by coastal laterite soils, face heightened fire susceptibility due to poor moisture retention and rapid drying, intensifying surface fire conditions. These local ecological characteristics reinforce the need for spatially informed and data-driven prediction models like the one proposed in this study. The outcomes of this research not only contribute to environmental modeling but also offer valuable inputs for policy development and community-based fire management initiatives. By accurately identifying high-risk zones and seasonal fire peaks, the proposed models can guide forest departments and disaster response agencies in the strategic allocation of firefighting resources and early warning dissemination. Furthermore, the ability to forecast fire intensity can help protect forest-dependent livelihoods, biodiversity hotspots, and eco-tourism assets in the Uttara Kannada district. Integrating such predictive tools into local governance frameworks and community training programs could significantly enhance resilience against forest fire hazards and promote sustainable land management practices.

The predictive insights generated by the hybrid CSORF and CS-XGB models hold potential applications within India's existing forest fire management strategies, such as the Forest Survey of India's Forest Fire Alert System and the National Action Plan on Forest Fires (NAPFF). By integrating fire risk predictions into these frameworks, forest departments can enhance early warning dissemination, optimize firefighting resource allocation, and develop localized community engagement programs. Moreover, minimizing fire incidents can safeguard forest-dependent livelihoods, conserve biodiversity, and support sustainable eco-tourism initiatives in ecologically sensitive regions like Uttara Kannada.

Looking forward, this framework can be extended to incorporate real-time satellite observations, dynamic weather feeds, and remote sensing data for live monitoring. Future developments may integrate fuzzy logic, deep learning architectures (e.g., LSTM, DNN), and fire spread simulation modules, making the system suitable for resource allocation, early warning systems, and disaster planning. Ultimately, the hybrid solution lays a strong foundation for operational deployment by forest departments and environmental agencies tasked with forest fire risk mitigation in vulnerable landscapes like Uttara Kannada. Future enhancements to this framework will involve the integration of real-time remote sensing datasets from satellite platforms such as MODIS, VIIRS, and Sentinel-2. The incorporation of dynamic environmental inputs is expected to further refine the temporal precision of fire risk forecasting models, facilitating more responsive early warning systems and proactive forest management interventions.

References

[1] N. Hamadeh, B. Daya, A. Hilal, and P. Chauvet, "An analytical review on the most widely used meteorological models in forest fire prediction," *In 2015 Third International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*, pp. 239–244, Apr. 2015, https://doi.org/10.1109/TAEECE.2015.7113633.

[2] M. Bao *et al.*, "Research Trends in Wildland Fire prediction amidst climate change: A Comprehensive bibliometric analysis," *Forests*, vol. 15, no. 7, p. 1197, Jul. 2024, https://doi.org/10.3390/f15071197.

[3] K. N. Babu, R. Gour, K. Ayushi, N. Ayyappan, and N. Parthasarathy, "Environmental drivers and spatial prediction of forest fires in the Western Ghats biodiversity hotspot, India: An ensemble machine learning approach," *Forest Ecology and Management,* vol. 540, p. 121057, Apr. 2023, https://doi.org/10.1016/j.foreco.2023.121057.

[4] H. D. Nguyen, "Hybrid models based on deep learning neural network and optimization algorithms for the spatial prediction of tropical forest fire susceptibility in Nghe An province, Vietnam," *Geocarto International,* vol. 37, no. 26, pp. 11281–11305, Mar. 2022, https://doi.org/10.1080/10106049.2022.2048904.

[5] G. Zhang, M. Wang, and K. Liu, "Forest fire susceptibility modeling using a convolutional neural network for Yunnan Province of China," *International Journal of Disaster Risk Science*, vol. 10, no. 3, pp. 386–403, Sep. 2019, https://doi.org/10.1007/s13753-019-00233-1.

[6] Y. Wang, Z. Liu, S. Qu, J. Gong, and X. Lyu, "Fire resistance of reinforced concrete beams: State of the art, analysis and prediction," *Construction and Building Materials,* vol. 409, p. 134029, Nov. 2023, https://doi.org/10.1016/j.conbuildmat.2023.134029.

[7] X. Wang, C. Wang, G. Zhao, H. Ding, and M. Yu, "Research Progress of forest fires Spread Trend Forecasting in Heilongjiang Province," *Atmosphere*, vol. 13, no. 12, p. 2110, Dec. 2022, https://doi.org/10.3390/atmos13122110.

[8] S. Sakellariou, S. Tampekis, F. Samara, A. Sfougaris, and O. Christopoulou, "Review of state-of-the-art decision support systems (DSSs) for prevention and suppression of forest fires," *Journal of Forestry Research*, vol. 28, no. 6, pp. 1107–1117, Jun. 2017, https://doi.org/10.1007/s11676-017-0452-1.

[9] O. M. Lozano et al., "Assessing climate change impacts on wildfire exposure in Mediterranean areas," *Risk Analysis*, vol. 37, no. 10, pp. 1898–1916, Dec. 2016, https://doi.org/10.1111/risa.12739.

[10] L. Sullivan, "Inside the Inferno: Fundamental processes of wildland fire behaviour," *Current Forestry Reports*, vol. 3, no. 2, pp. 132–149, Apr. 2017, https://doi.org/10.1007/s40725-017-0057-0.

[11] H. Cruz, T. Gualotuña, M. Pinillos, D. Marcillo, S. Jácome, and E. R. F. C, "Machine learning and color treatment for the forest fire and smoke detection systems and algorithms, a recent literature review," in *Advances in intelligent systems and computing*, 2021, pp. 109–120. https://doi.org/10.1007/978-3-030-68080-0_8.

[12] N. J. Abram *et al.*, "Connections of climate change and variability to large and extreme forest fires in southeast Australia," *Communications Earth & Environment*, vol. 2, no. 1, Jan. 2021, https://doi.org/10.1038/s43247-020-00065-8.

[13] M. G. Gale, G. J. Cary, A. I. J. M. Van Dijk, and M. Yebra, "Forest fire fuel through the lens of remote sensing: Review of approaches, challenges and future directions in the remote sensing of biotic determinants of fire behaviour," *Remote Sensing of Environment*, vol. 255, p. 112282, Jan. 2021, https://doi.org/10.1038/s43247-020-00065-8.

[14] A. Asokan and J. Anitha, "Change detection techniques for remote sensing applications: a survey," *Earth Science Informatics*, vol. 12, no. 2, pp. 143–160, Mar. 2019, https://doi.org/10.1007/s12145-019-00380-5.

[15] Karimi *et al.*, "Evaluating Models and Effective Factors Obtained from Remote Sensing (RS) and Geographic Information System (GIS) in the Prediction of Forest Fire Risk, Structured Review," *Journal of Geography and Cartography*, vol. 1, no. 4, Sep. 2018, https://doi.org/10.24294/jgc.v3i1.618.

[16] M. Mishra *et al.*, "Spatial analysis and machine learning prediction of forest fire susceptibility: a comprehensive approach for effective management and mitigation," *The Science of the Total Environment*, vol. 926, p. 171713, Mar. 2024, https://doi.org/10.1016/j.scitotenv.2024.171713.

[17] F. Sivrikaya and Ö. Küçük, "Modeling forest fire risk based on GIS-based analytical hierarchy process and statistical analysis in Mediterranean region," *Ecological Informatics*, vol. 68, p. 101537, Dec. 2021, https://doi.org/10.1016/j.ecoinf.2021.101537.

[18] V. K. Singh, C. Singh, and H. Raza, "Event classification and intensity discrimination for forest fire inference with IoT," *IEEE Sensors Journal*, vol. 22, no. 9, pp. 8869–8880, Mar. 2022, https://doi.org/10.1109/JSEN.2022.3163155.

[19] F. Abid, "A survey of machine learning algorithms based forest fires prediction and detection systems," *Fire Technology*, vol. 57, no. 2, pp. 559–590, Nov. 2020, https://doi.org/10.1007/s10694-020-01056-z.

[20] K. Mehmood *et al.*, "Assessing Chilgoza Pine (Pinus gerardiana) forest fire severity: Remote sensing analysis, correlations, and predictive modeling for enhanced management strategies," *Trees Forests and People*, vol. 16, p. 100521, Feb. 2024, https://doi.org/10.1016/j.tfp.2024.100521.

[21] H. Hong, P. Tsangaratos, I. Ilia, J. Liu, A.-X. Zhu, and C. 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," *The Science of the Total Environment*, vol. 630, pp. 1044–1056, Mar. 2018, https://doi.org/10.1016/j.scitotenv.2018.02.278.

[22] C.-H. Lim, Y. S. Kim, M. Won, S. J. Kim, and W.-K. Lee, "Can satellite-based data substitute for surveyed data to predict the spatial probability of forest fire? A geostatistical approach to forest fire in the Republic of Korea," *Geomatics Natural Hazards and Risk*, vol. 10, no. 1, pp. 719–739, Jan. 2019, https://doi.org/10.1080/19475705.2018.1543210.

[23] W. Hung *et al.*, "Fire Intensity and SPREAD ForECAST (FIRA): a machine learning based fire spread prediction model for air quality forecasting application," *GeoHealth*, vol. 9, no. 3, Mar. 2025, https://doi.org/10.1029/2024GH001253.

[24] M. G. Cruz, M. E. Alexander, A. L. Sullivan, J. S. Gould, and M. Kilinc, "Assessing improvements in models used to operationally predict wildland fire rate of spread," *Environmental Modelling & Software*, vol. 105, pp. 54–63, Apr. 2018, https://doi.org/10.1016/j.envsoft.2018.03.027.

[25] M. Rihan *et al.*, "Forest fire susceptibility mapping with sensitivity and uncertainty analysis using machine learning and deep learning algorithms," *Advances in Space Research*, vol. 72, no. 2, pp. 426–443, Mar. 2023, https://doi.org/10.1016/j.asr.2023.03.026.

[26] P. Banerjee, "Maximum entropy-based forest fire likelihood mapping: analysing the trends, distribution, and drivers of forest fires in Sikkim Himalaya," *Scandinavian Journal of Forest Research*, vol. 36, no. 4, pp. 275–288, Apr. 2021, https://doi.org/10.1080/02827581.2021.1918239.

[27] X. Yuan, N. Liu, X. Xie, and D. X. Viegas, "Physical model of wildland fire spread: Parametric uncertainty analysis," *Combustion and Flame*, vol. 217, pp. 285–293, Apr. 2020, https://doi.org/10.1016/j.combustflame.2020.03.034.

[28] Tariq *et al.*, "Forest fire monitoring using spatial-statistical and Geo-spatial analysis of factors determining forest fire in Margalla Hills, Islamabad, Pakistan," *Geomatics Natural Hazards and Risk*, vol. 12, no. 1, pp. 1212–1233, Jan. 2021, https://doi.org/10.1080/19475705.2021.1920477.

[29] A. Deshmukh, S. D. B. Sonar, R. V. Ingole, R. Agrawal, C. Dhule, and N. C. Morris, "Satellite Image Segmentation for Forest Fire Risk Detection using Gaussian Mixture Models," *In 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, May 2023, https://doi.org/10.1109/ICAAIC56838.2023.10140399.

[30] M. Singh and Z. Huang, "Analysis of forest fire dynamics, distribution and main drivers in the Atlantic Forest," *Sustainability*, vol. 14, no. 2, p. 992, Jan. 2022, https://doi.org/10.3390/su14020992.

[31] X. Lin, Z. Li, W. Chen, X. Sun, and D. Gao, "Forest fire prediction based on Long- and Short-Term Time-Series network," Forests, vol. 14, no. 4, p. 778, Apr. 2023, https://doi.org/10.3390/f14040778.

[32] R. Ghali, M. A. Akhloufi, and W. S. Mseddi, "Deep learning and transformer approaches for UAV-Based wildfire detection and segmentation," Sensors, vol. 22, no. 5, p. 1977, Mar. 2022, https://doi.org/10.3390/s22051977.

[33] F. Ji et al., "Coupling physical model and deep learning for near Real-Time wildfire detection," IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1–5, Jan. 2023, https://doi.org/10.1109/LGRS.2023.3307129.

[34] X.-N. Bui, H. Nguyen, Q.-H. Tran, D.-A. Nguyen, and H.-B. Bui, "Predicting ground vibrations due to mine blasting using a novel artificial neural Network-Based cuckoo search optimization," Natural Resources Research, vol. 30, no. 3, pp. 2663–2685, Feb. 2021,. https://doi.org/10.1007/s11053-021-09823-7.

[35] S. Milanović, N. Marković, D. Pamučar, L. Gigović, P. Kostić, and S. D. Milanović, "Forest Fire Probability Mapping in Eastern Serbia: Logistic Regression versus Random Forest Method," Forests, vol. 12, no. 1, p. 5, Dec. 2020, https://doi.org/10.3390/f12010005.