FRI-2G.303-1-CCT1-01

ADVANCING WILDFIRE SUSCEPTIBILITY ANALYSIS THROUGH TEMPORAL FEATURE INTERPRETATION AND MODEL EXPLAINABILITY [1]

Yehya Bouzeraa, PhD Student
Abdelhafid Boussouf University Center of Mila, Algeria
LIRE Laboratory of Constantine2, Constantine, Algeria
E-mail: bouzeraa.yehya@centre-univ-mila.dz

Prof. Nardjes Bouchemal, PhD
Abdelhafid Boussouf University Center of Mila, Algeria
LISI Laboratory of Intelligent Systems and Informatics,
Abdelhafid Boussouf University Center of Mila
E-mail: n.bouchemal.dz@ieee.org

Assoc. Prof. Svetlana Stefanova, PhD
Department of Computer Science,
University of Ruse “Angel Kanchev”
E-mail: sstefanova@ecs.uni-ruse.bg

Abstract: *Wildfire susceptibility mapping (WSM) is essential for anticipating where fires are likely to occur and for planning prevention and resources. Many existing studies still rely on subjective factor choice, limited model optimization, conventional metrics only, and weak interpretability. This paper proposes a WSM framework that combines data-driven feature selection, spatially aware model training with hyperparameter tuning, validation using recent fire events, and explainable analysis of model behaviour. Designed for the Jijel region, the framework aims to produce more accurate and stable susceptibility maps while clarifying the roles of climatic, topographic, vegetation, and human factors. The ultimate goal is to provide outputs that are both operationally useful and scientifically transparent.*

INTRODUCTION

Wildfire Susceptibility Mapping (WSM) is a predictive approach that estimates the likelihood of fire occurrence across a given territory. It combines spatial data—such as topography, vegetation, climate, and human activity—with past fire records to produce maps that classify areas into different susceptibility levels (e.g. low, medium, high, very high). These maps allow authorities to anticipate where wildfires are most likely to ignite before they occur (Abdollahi, A., & Pradhan, B., 2023).

Beyond its scientific interest, WSM is a critical decision-support tool for operational agencies. Firefighting services, civil protection units, and land-use planners rely on susceptibility maps to design prevention strategies, allocate resources, and define early warning zones. By clearly visualizing high-risk areas, WSM supports more efficient surveillance, infrastructure protection, and long-term land management policies (Abujayyab, S. K. M. et al, 2022).

However, developing accurate and reliable wildfire susceptibility maps remains challenging. The increasing availability of open-access geospatial data, the complexity of machine learning models, and the need for robust validation and interpretability all introduce methodological and practical difficulties (Abdollahi, A., & Pradhan, B., 2023).

---

[1] The paper was presented on 24 October 2025 in section “Communication and Computer Technologies” with original title in English: ADVANCING WILDFIRE SUSCEPTIBILITY ANALYSIS THROUGH TEMPORAL FEATURE INTERPRETATION AND MODEL EXPLAINABILITY

This work aims to highlight the main challenges faced by researchers when building wildfire susceptibility maps and to propose a structured framework to address them. The key challenges considered in this study are:

- Feature selection and data justification
- Model training and optimization
- Model evaluation and real-world validation
- Interpretation and explainability of results

Based on these challenges, we propose an improved wildfire susceptibility mapping framework that integrates data-driven feature selection, rigorous training procedures, real-data validation, and explainable analysis to support reliable and operational wildfire management.

Challenges in Developing Reliable Wildfire Susceptibility Maps

Wildfire susceptibility mapping involves several methodological choices that can strongly influence the quality and usefulness of the resulting maps. Beyond data availability and model choice, researchers must carefully decide which variables to include, how to train and validate their models, and how to interpret the predicted patterns. In this section, we discuss the main challenges encountered in developing reliable susceptibility maps, starting with feature selection and data justification.

Feature Selection and Data Justification

In most wildfire susceptibility mapping studies, the selection of predictive factors has traditionally relied on previous literature, expert reports, or fire management guidelines. With the emergence of numerous open-access geospatial platforms and big data sources—such as ERA5, TerraClimate, MODIS, Sentinel, and Copernicus—researchers now have access to a very large variety of environmental and anthropogenic variables. Each factor can also be represented in multiple ways (for example, temperature as mean, maximum, seasonal average, or anomaly), creating a complex pool of potential predictors (Abdollahi, A., & Pradhan, B., 2023).

This abundance of data, while valuable, poses a central challenge: deciding which features are truly relevant, how to represent them effectively, and how to avoid redundancy or bias. Without a well-justified selection strategy, models risk becoming unstable, overfitted, or difficult to interpret. It is therefore crucial to combine expert knowledge with data-driven methods—such as correlation analysis, multicollinearity testing, and feature reduction—to ensure that each chosen variable is scientifically justified and meaningfully related to wildfire occurrence (Iban, M. C., & Sekertekin, A., 2022).

Model Training and Optimization

Machine learning algorithms such as Random Forest, Support Vector Machine, XGBoost, and other ensemble methods are widely used to model the relationship between environmental factors and fire occurrence. In many studies, however, models are trained using a single dataset or a limited set of hyperparameter configurations, without systematically exploring how model structure and hyperparameters affect performance and stability (Jain, P. et al, 2020).

This limited training strategy often produces models that perform well on specific data samples but fail to generalize to new spatial or temporal conditions. In addition, the spatial nature of wildfire data—where neighboring pixels tend to share similar characteristics—introduces spatial autocorrelation, which can increase the risk of overfitting and artificially inflated performance metrics. Addressing these issues requires more rigorous training procedures, including appropriate cross-validation schemes, hyperparameter tuning, and spatially-aware sampling (Abujayyab, S. K. M. et al, 2022).

Model Evaluation and Real-World

In many wildfire susceptibility mapping studies, model performance is evaluated only with traditional statistical metrics such as AUC, accuracy, precision–recall, or F1-score. While these indicators are useful for quantifying predictive power on historical data, they are often insufficient for critical environmental applications, where the ultimate objective is to anticipate real fire occurrences rather than to maximize a metric (Iban, M. C., & Sekertekin, A., 2022).

For operational wildfire management, authorities need models that not only perform well on training and testing datasets but also demonstrate reliability under real conditions. However, only a limited number of studies go beyond conventional validation by testing their models on recent or independent fire events. Few works overlay model outputs with actual fires that occurred later in the same region, which is essential to assess spatial and temporal generalization. This gap limits the confidence that decision-makers can have in susceptibility maps when using them for planning and resource allocation (Valavi, R. et al, 2019).

Validation Interpretation and Explainability of Results

Many of the machine learning models used in wildfire susceptibility mapping—especially ensemble and deep learning approaches—are treated as black boxes. They can achieve high predictive performance but often provide limited insight into why specific areas are classified as high-risk. This lack of interpretability reduces trust among decision-makers and restricts the adoption of these models in operational wildfire management (Wang, Y. et al, 2023).

In critical applications such as fire prevention and emergency planning, understanding model reasoning is as important as accuracy. Authorities need to know which factors—such as temperature, vegetation dryness, topography, or proximity to roads—contribute most to ignition risk and how these contributions may change over time. Without clear explanations and visual tools, even high-performing models may not be considered transparent or robust enough for real-world use (Zakari, R. et al, 2025).

Proposed Solutions and Framework Development

After identifying the main challenges in wildfire susceptibility mapping, we propose a framework designed to improve accuracy, robustness, and interpretability by integrating data-driven feature selection, optimized model training, real-world validation, and explainable analysis of results, with the overall goal of obtaining a susceptibility mapping system that is both scientifically sound and operationally useful for wildfire management.

Data-Driven Feature Selection and Justification

To reduce subjectivity in factor choice and avoid redundant predictors, the framework adopts a data-driven feature selection strategy. All potentially relevant variables are first collected from open-access global and local datasets, including climate, topography, vegetation, land cover, and human activity. These may include reanalysis products, satellite-derived indices, and proximity-based indicators of human pressure.

Pairwise correlation analysis and multicollinearity diagnostics are then applied to detect highly correlated variables, and redundant features are removed or aggregated in order to prevent instability and overfitting. In a subsequent step, preliminary machine learning models are trained to estimate feature importance scores, and the most influential variables are retained while low-impact factors are discarded. This process ensures that the final feature set is scientifically justified, non-redundant, and representative of the main drivers of wildfire occurrence in the study area.

Rigorous Model Training and Optimization

To improve model reliability and generalization, the framework emphasizes robust training procedures. Systematic cross-validation is combined with automated hyperparameter search, such as grid search or Bayesian optimization, to identify stable model configurations and reduce overfitting. In addition, training and testing samples are drawn using spatially-aware strategies to limit the impact of spatial autocorrelation between neighboring pixels, thereby providing more realistic performance estimates.

Several machine learning algorithms, such as Random Forest, XGBoost, and Gradient Boosting, are trained and compared under the same conditions. The final model is selected not only on the basis of conventional accuracy metrics but also according to its stability and generalization ability across different spatial and temporal settings (Fig.1). These steps lead to a more robust susceptibility model that can better adapt to varying environmental and climatic conditions.

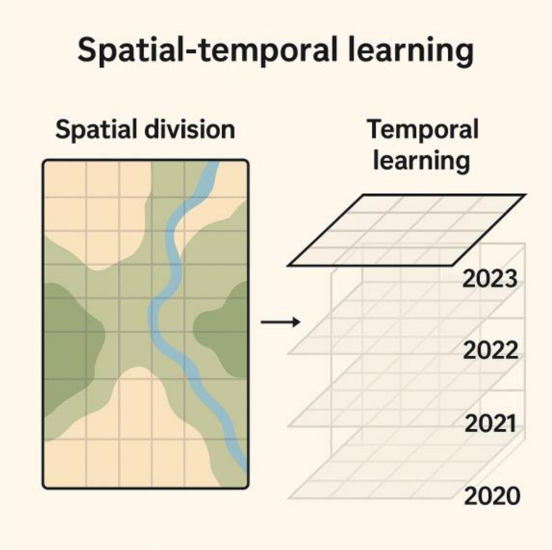

Fig. 1. Conceptual illustration of the spatial–temporal learning setup.

Real-World Validation with Recent Fire

To move beyond purely metric-based evaluation, the framework incorporates a real-world validation stage using recent fire occurrences. Actual fire locations, obtained from satellite products such as MODIS detections and from national civil protection records, are overlaid on the generated susceptibility map. On this basis, operational performance indicators are computed, notably the percentage of fires that fall within the medium, high, and very high susceptibility zones; a high proportion of events in these classes is interpreted as evidence of good spatial predictive ability.

When possible, the results are also compared with earlier susceptibility maps produced for the same region in order to assess improvements in spatial and temporal reliability. This validation stage ensures that the model is not only statistically performant but also operationally meaningful and consistent with observed fire patterns.

Events Explainability and Interpretation of Model Results

To transform the model from a black box into a transparent analytical tool, the framework integrates explainable AI techniques. Global feature importance measures, correlation matrices, and related visualizations are used to clarify the relative contribution of each variable to fire susceptibility. Partial dependence plots or similar response analysis tools are applied to show how variations in key predictors, such as temperature, vegetation dryness, or distance to roads, influence the predicted susceptibility levels.

The evolution of feature importance is further analyzed across multiple years, for example from 2020 to 2025, in order to detect potential changes in the roles of climatic, topographic, and anthropogenic factors over time. Finally, the results are summarized in interpretable spatial outputs that classify areas by susceptibility level and highlight the dominant drivers in different zones. These explainability components increase trust among decision-makers and provide additional scientific insight into wildfire dynamics in the study area.

Together, these four components form a comprehensive architecture for wildfire susceptibility mapping that is transparent, validated, and scientifically grounded. The proposed framework is (as shown in Fig.2) designed to be applied to the Jijel region but can also be extended to other fire-prone areas, thereby supporting authorities in the design of more effective prevention strategies, surveillance systems, and resource allocation plans.

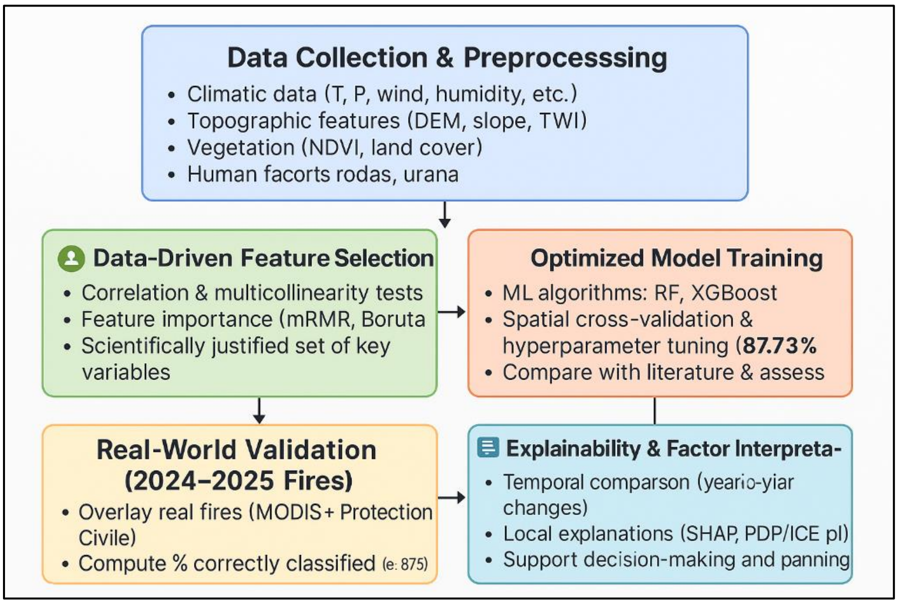

Fig. 2. Proposed framework for wildfire susceptibility mapping and evaluation.

Expected Results

The proposed framework is expected to improve the accuracy and reliability of wildfire susceptibility maps by combining a richer and more appropriately selected set of features with rigorous model training procedures. Real-world validation using fire events from the 2024–2025 seasons should confirm strong spatial predictive performance, particularly in correctly identifying medium, high, and very high susceptibility zones. In parallel, the integration of explainability methods is anticipated to clarify the contribution of each factor to fire occurrence, thereby producing susceptibility maps that are not only more precise but also easier to interpret and more useful for operational decision-making and planning.

Conclusion

This work proposes an improved and interpretable framework for wildfire susceptibility mapping that directly addresses four major challenges: feature selection, model training, model evaluation, and result interpretation. By combining data-driven feature selection with rigorous spatial cross-validation, hyperparameter optimization, and real-world validation using recent fire events, the framework aims to produce susceptibility maps that are both accurate and robust. The integration of explainability techniques further transforms the model into a transparent analytical tool, clarifying the role of climatic, topographic, vegetation, and human factors in driving fire occurrence. Applied to the Jijel region, the framework is intended to support fire prevention, surveillance, and resource planning, and it can be adapted to other fire-prone areas facing similar data and modeling constraints. Future work may extend this approach by incorporating near real-time data streams and by coupling susceptibility mapping with early warning and response systems.

Acknowledgement

This paper is supported by project 2025-FEEA-02 “Application of AI in various professional fields: opportunities, challenges and ethical issues”, funded by the Research Fund of the “Angel Kanchev” University of Ruse, Bulgaria

REFERENCES

Abdollahi, A., & Pradhan, B. (2023). Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model. *Science of the Total Environment, 879*, 163004.

Abujayyab, S. K. M., Kassem, M. M., Khan, A. A., Wazirali, R., Coşkun, M., Taşoğlu, E., Öztürk, A., & Toprak, F. (2022). Wildfire susceptibility mapping using five boosting machine learning algorithms: The case study of the Mediterranean region of Turkey. *Advances in Civil Engineering, 2022*, 3959150.

Iban, M. C., & Sekertekin, A. (2022). Machine learning based wildfire susceptibility mapping using remotely sensed fire data and GIS: A case study of Adana and Mersin provinces, Turkey. *Ecological Informatics, 69*, 101647.

Jain, P., Coogan, S. C. P., Subramanian, S. G., Crowley, M., Taylor, S., & Flannigan, M. D. (2020). A review of machine learning applications in wildfire science and management. *Environmental Reviews, 28*(4), 478–505.

Nikolaychuk, O., Pestova, J., & Yurin, A. (2024). Wildfire susceptibility mapping in Baikal natural territory using random forest. *Forests, 15*(1), 170. (Nikolaychuk et al 2024)

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). blockCV: An R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. *Methods in Ecology and Evolution, 10*(2), 225–232.

Wang, Y., Khodadadzadeh, M., & Zurita-Milla, R. (2023). Spatial+: A new cross-validation method to evaluate geospatial machine learning models. *International Journal of Applied Earth Observation and Geoinformation, 121*, 103364.

Zakari, R. Y., Malik, O. A., & Ong, W.-H. (2025). Machine learning-driven wildfire susceptibility mapping in New South Wales, Australia using remote sensing and explainable artificial intelligence. *Natural Hazards, 121*(13), 15331–15357.