

Spatio-Temporal Forest and Land Fire Risk Modeling in Sumatra Using Atmospheric-Edaphic Integration via Bivariate Fuzzy C-Means

Ade Firmansyah¹, Dedi Darwis²

^{1,2}Faculty of Engineering and Computer Science, Universitas Teknokrat Indonesia, Bandar Lampung, Indonesia

Article Info

Article history:

Received Mar 07, 2026

Revised May 25, 2026

Accepted July 16, 2026

Keywords:

VLIF-Model

Bivariate Fuzzy C-Means

Vapor Pressure Deficit

Forest and Land Fire

Sumatra

ABSTRACT

The study focused on the persistent forest and land fires in Sumatra and the conventional early warning systems that showed limitations in capturing the dynamics of the soil-atmosphere and tended to overestimate risk zones during the peak periods of fires. Existing multivariate systems also showed limitations when there was a lack of integration between atmospheric water demand and edaphic (soil) vulnerability, resulting in lower spatial efficiency in the risk area. This paper presented a new bivariate physics-based model called the VPD-LVI Integrated Fuzzy (VLIF) Model, which integrated Vapor Pressure Deficit (VPD) and Land Vulnerability Index (LVI). For the VLIF-Model, to ensure a high level of spatial detail, the model processed 2,614,056 spatiotemporal observations, representing a grid resolution of 0.1° over a two-year period (2023–2024). The model showed acceptable structural stability at $k=3$ (FPC = 0.6137), with distinct risk zoning contrasts. Furthermore, validation against 61,015 VIIRS thermal anomalies indicated the model's predictive reliability with an ROC-AUC of 0.8288 and a Brier Score of 0.0169, suggesting good discrimination and calibration. In the spatial efficiency analysis, 17.80% of the area categorized as 'high' risk captured 62.28% of actual thermal anomaly grids, achieving a lift factor of 3.50 and a recall of 88.91%. The VLIF-Model provided clearer contouring of peak fire zones and enabled a more targeted spatial risk intelligence layer for specific tropical peatland mitigation efforts.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Dedi Darwis,

Faculty of Engineering and Computer Science,

Universitas Teknokrat Indonesia,

Jl. Zainal Abidin Pagaram No. 9-11, Labuhan Ratu, Bandar Lampung.

Email: darwisdedi@teknokrat.ac.id

1. INTRODUCTION

Fires in forests and on land in tropical ecosystems are an extremely complicated environmental disaster, especially on the island of Sumatra. They cause a tremendous amount of biodiversity loss and carbon emission, as well as causing a public health crisis because of the impact of trans-boundary haze caused by fire [1]. The economic and ecological impacts from fires are across several sectors such as agriculture, plantations, transportation and tourism, putting the targets for reducing emissions in jeopardy. The past decade has seen a substantial increase in both frequency and intensity of fires in Sumatra, including on peatlands, which is a major carbon sink and are particularly vulnerable to subsurface burning [2].

The role of human activity in explaining fire behaviour in technologies employed for agriculture or plantations in tropical rainforests is limited. Existing literature in tropical fire ecology suggests that soil-atmosphere coupling plays a major role in creating fire-prone conditions [3],[4]. Although our VLIF-Model is conceptually inspired by this theory, it differs by integrating

anthropogenic factors (HAI) into the edaphic component, thus representing an extended edaphic-atmospheric integration rather than a pure biophysical coupling. Furthermore, its associated positive feedback mechanism is one where dry land conditions will limit surface cooling due to the evaporation process which will in turn cause an increase in surface temperature and VPD of the atmosphere [3]. A growing VPD as established by Clarke et al. will therefore also lead to an increase in drying of fine and dead fuels, causing severe conditions for the occurrence of forest fires [5]. In addition, research into the role of soil moisture feedback indicates that 90% of the variability in terrestrial carbon storage can be attributed to this phenomenon; furthermore, hydrologic drought has been found to generate more forest fires in tropical rainforests than weather-based approaches have found [6].

The existing operational early warning systems implemented in Indonesia are still dominated by the Fire Danger Rating System (FDRS), which depends on conventional meteorological indices [7]. There are several basic shortcomings of such an approach. First, the static Fire Danger Rating System does not take anthropogenic factors and edaphic heterogeneity, such as peatlands and types of land cover that indirectly indicate anthropogenic influences, into account [8],[9]. Second, it ignores extreme variations in data aggregated by 1-day intervals (24 h), although the daytime period from 09:00 to 15:00 WIB is critical for fire ignition and Fire Radiation Power peaks, which are often not captured by average data [10]. Third, multivariate baselines using multiple meteorological parameters simultaneously are often confronted with non-selective expansion of risk areas during peak fire season, thus compromising their spatial efficiency.

It is important to clarify that the proposed VLIF-Model is not intended to replace the existing Fire Danger Rating System (FDRS). Rather, this model serves as a complementary spatial decision-support layer. While the FDRS provides a general fire hazard index (e.g., “warning” or “red alert”) across broad administrative regions, VLIF translates the underlying physical drivers into spatially explicit risk clusters at a 0.1° resolution. This enables operational agencies to shift from area-wide warnings to targeted preventive actions, such as prioritizing patrol routes or allocating water-bombing resources to specific high-risk grids. In other words, FDRS indicates when hazardous conditions exist; VLIF responds precisely where those hazards are most concentrated.

Various attempts to develop alternative models have been made using a soft-computing approach. Å. Nebot and F. Mugica demonstrated a fuzzy model for fire forecasting by considering the uncertainty of environmental variables [11]. A. Jafarzadeh combined the Apriori and fuzzy c-means algorithms for risk evaluation, demonstrating the ability of the fuzzy cluster method in delineating vulnerable areas [12]. In Indonesia, F. Rohimahastuti developed a Bayesian spatio-temporal model for hotspot estimation in Kalimantan [13], meanwhile, Vatesia et al. implemented ST-DBSCAN to analyze hotspot distribution patterns in Sulawesi, with the finding that the July–October period was the peak of hotspot occurrence [14]. However, a significant Research Gap and Contribution of this study lies in addressing the limitations of these multivariate approaches which often face challenges in maintaining the sharpness of risk zoning when entering a critical period [15]. There is still a lack of integration between atmospheric water demand and edaphic factors into a single physically grounded indicator, especially in peat ecosystems that contribute significantly to atmospheric emissions [16], [17].

Unlike previous studies that relied on baseline multivariate meteorological indices with a tendency for risk zone expansion during peak periods, this study introduced a high-resolution (0.1°) bivariate modeling framework that explicitly integrated atmospheric water demand (VPD) and a multi-dimensional Land Vulnerability Index (LVI). The key innovation lay in the enhancement of the LVI by incorporating the Human Accessibility Index (HAI), which quantifies proximity to roads and settlements, alongside edaphic factors to capture anthropogenic influences. This integration was operationalized through the VPD-LVI Integrated Fuzzy (VLIF) Model, which coupled both components using feature engineering, thus synthetically replicating the positive feedback mechanism at the core of soil-atmosphere coupling theory [18]. The bivariate Fuzzy C-Means (FCM) approach was then applied to produce objective risk zoning, overcoming the weakness of sharp boundaries in crisp clustering methods such as K-Means [12],[19],[20]. By

utilizing a massive spatio-temporal dataset covering 2,614,056 observations (2023–2024) derived from ERA5-Land at 0.1° resolution, a period deliberately selected because 2023 experienced one of the strongest El Niño events in the last decade, creating extreme fire-prone conditions, while 2024 served as a transition year, allowing the model to be rigorously tested under both extreme drought and post-drought recovery phases, and validation using VIIRS thermal anomalies with a resolution of 375 m [21], this study demonstrated robust predictive reliability with an ROC-AUC of 0.8288 and a Brier Score of 0.0169. Within the high-risk zone, the model achieved a hit rate of 62.28% and a lift factor of 3.50 while only covering 17.80% of the total area, providing a more precise mitigation tool for land management authorities in Sumatra.

2. RESEARCH METHOD

This section describes a structured, step-by-step process for the spatial-temporal clustering of forest and land fire risks in Sumatra using the Bivariate Fuzzy C-Means algorithm. This model utilizes the bivariate integration of two core physical drivers: Vapor Pressure Deficit (VPD), which represents atmospheric water demand, and the Land Vulnerability Index (LVI), which represents edaphic vulnerability. These variables are synthesized to provide a robust measure of spatial-temporal risk, with the overall research workflow illustrated in Figure 1.

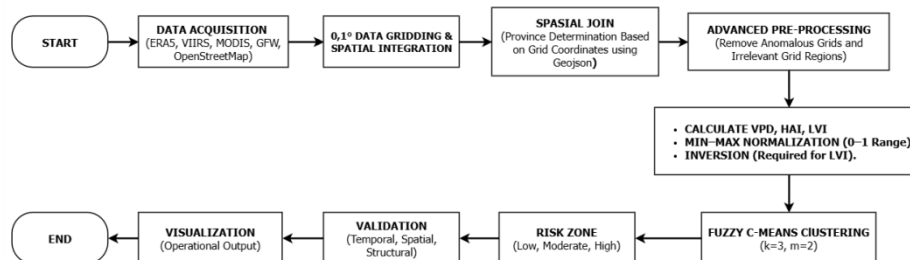


Figure 1. Research Flow

2.1. Data Acquisition and Integration

This study utilized an extensive spatio-temporal dataset integrated from heterogeneous sources, including atmospheric physics, edaphic biophysics, peatland distribution, land cover, and anthropogenic accessibility markers. The final integrated dataset comprised 2,614,056 rows of spatial observations in daily aggregates for the period January 2023 to December 2024. The data acquisition and integration process was conducted as follows:

1. Hotspots (Ground Truth): The datasets of hotspots were obtained from the NASA FIRMS portal using S-NPP VIIRS sensor data with 375 m spatial resolution (FIRMS) [22],[23].
2. Meteorology and Soil: Meteorological and soil data from the ECMWF ERA5-Land dataset were utilized for obtaining temperature, relative humidity (RH), and soil moisture information [24].
3. Land Cover: MODIS MCD12Q1 (version 6.1: Terra and Aqua) provided the data for years of land cover classification globally at 500 m resolution. This land cover data was obtained from NASA Earthdata website and was in GeoTiff format. These data were used to classify forest, shrubland and open land use of International Geosphere-Biosphere Programme [25].
4. Peatland: Peatland data was obtained from Global Forest Watch (GFW) services which provides a peatland map of Indonesia at 500 m resolution.
5. Roads and Settlements (Anthropogenic Accessibility): Road and settlement/building data were downloaded from the Geofabrik download server (<https://download.geofabrik.de>), which provides free OpenStreetMap (OSM) extracts for the Sumatra region. These datasets were used to calculate the Human Accessibility Index (HAI) based on the distance from each grid to the nearest road and to the nearest settlement or building.
6. Administrative Boundaries: GeoJSON files of the administrative boundaries of the provinces of Sumatra, for the aggregation of data per province and the visualization of risk maps, were created by the Nusa Map Team and are based on the Decree of the Minister of Home Affairs Number 300.2.2-2138 of 2025.

Dataset harmonization was carried out through spatial grouping to integrate the scales of atmospheric and land surface datasets. The first important step in data harmonization was to convert the different observation coordinates from several sources into a single coordinate scale with a resolution of $0.1^\circ \times 0.1^\circ$ (approximately 11 km), with rounded latitude and longitude values. The use of a resolution of 0.1° , which is the resolution of the ERA5-Land meteorological dataset, is an optimal solution in balancing the spatial resolution and temporal stability, while enabling a more detailed representation of edaphic and anthropogenic factors. The coarsest resolution was used as a benchmark for the spatial resolution since it is the lowest resolution of the datasets used for integration. This is also a precautionary measure for interpolation errors and false precision that usually arise when downsampling meteorological datasets to match the higher resolution of satellite imagery such as VIIRS and MODIS. At a regional level such as Sumatra, this resolution is both efficient and accurate for the dynamics of bivariate edaphic-atmosphere interactions (VPD and LVI_norm), where LVI_norm now integrates soil moisture, peat score, land cover score, and the newly developed Human Accessibility Index (HAI) to account for anthropogenic drivers. This is essential for the integration of high-precision datasets from ERA5-Land, VIIRS, MODIS, GFW, and hotspot validation data with nominal (n) and high (h) confidence levels for objective accumulation per grid, thus ensuring consistency in the granulation of the data for spatial-temporal modeling of forest and land fire risk in Sumatra [20].

2.2. Computing Environment and Advanced Data Preprocessing

Initial data inspection and basic administrative checks were performed using spreadsheet software (Microsoft Excel) solely for observational purposes, as Excel is incapable of handling the full dataset of 2,614,056 rows. All technical preprocessing, feature engineering, and model execution were conducted in Google Colab [26] using Python (v3.10). The primary dataset was stored in CSV format, not Excel, and was loaded directly into the Python environment. Data manipulation employed Pandas (v2.0.3), numerical operations used NumPy (v1.23.5), and fuzzy logic implementation relied on Scikit-Fuzzy (v0.4.2). Additional libraries included Scikit-Learn (v1.2.2) for data scaling and Statsmodels (v0.14.0) for multicollinearity analysis. A key preprocessing innovation was the selection of meteorological data within the 09:00–15:00 WIB time window to avoid the masking effect of 24-hour daily averages, which tend to obscure extreme fire-risk conditions due to nighttime cooling and high humidity [6]. Meteorological variables were derived from the ERA5-Land dataset at 0.1° spatial resolution (≈ 11 km), aggregated daily per grid cell, and served as the master grid for harmonizing other datasets. Advanced preprocessing further included spatial filtering across eight mainland provinces of Sumatra and cleaning of anomalous records. To ensure clustering stability, input features were normalized using StandardScaler (Z-score normalization) prior to feature engineering. The bivariate Fuzzy C-Means (FCM) algorithm was configured with a fuzziness exponent (m) of 2, a maximum of 1,000 iterations, and a stopping error (ϵ) of 0.005 to guarantee convergence and objective risk partitioning.

2.3. Feature Engineering: Construction of VPD, HAI, and LVI

At this stage, the raw variables extracted from critical hours are transformed into standardized risk indicators through feature engineering. The key innovation of this approach is the integration of atmospheric water demand (VPD) and land vulnerability (LVI) components as bivariate inputs for the Fuzzy C-Means clustering algorithm, enabling the quantification of land-atmosphere coupling strength.

2.3.1 Calculation of Vapor Pressure Deficit (VPD)

Vapor Pressure Deficit (VPD) is a standard parameter in forest and land fire studies that represents the atmosphere's ability to dry out fuel [27]. VPD is calculated from air temperature (T) and relative humidity (RH) [28], using the Magnus–Tetens equation, as shown in Equation (1):

$$VPD = e_s - e_a \quad (1)$$

Where e_s is the saturation vapor pressure (kPa) and e_a is the actual vapor pressure (kPa). The saturation vapor pressure is derived from temperature:

$$e_s = 0.6108 \times \exp\left(\frac{17.27 \times T}{T + 237.3}\right) \quad (2)$$

And the actual vapor pressure is obtained by multiplying e_s by the relative humidity:

$$e_a = e_s \times \left(\frac{RH}{100}\right) \quad (3)$$

Thus, VPD can be expressed in a single equation:

$$VPD = 0.6108 \times \exp\left(\frac{17.27 \times T}{T + 237.3}\right) \times \left(1 - \frac{RH}{100}\right) \quad (4)$$

The VPD values obtained are then normalized to the range (0–1) for use as input in the clustering algorithm.

2.3.2 Human Accessibility Index (HAI)

The Human Accessibility Index (HAI) was introduced to represent anthropogenic pressure on fire-prone areas. The HAI is derived from two distance-based components: distance to the nearest road and distance to the nearest settlement or building. Each distance value (in kilometers) is first converted into a score based on a piecewise linear decay function, as shown in Table 1. The final HAI is then obtained by summing these two scores and normalizing the result to the range [0,1].

Table 1. Scoring function for distance to roads and settlements (HAI)

Distances (km)	Score
<2	1.0
2-4	0.8
4-6	0.6
6-8	0.4
8-10	0.2
>10	0

The rationale is that areas closer to human infrastructure (≤ 2 km) have the highest accessibility and thus a higher risk of ignition, while areas beyond 10 km are considered effectively inaccessible. Raw HAI scores are normalized to [0,1] to ensure comparability with other indices.

2.3.3 Calculation of Land Vulnerability Index (LVI)

The weighting scheme is based on empirical evidence that the hydrological conditions of peatlands have a greater influence on fire persistence than variations in land cover in tropical ecosystems [29]. In Equation (5), a multiplier of 2 is applied to Soil Moisture (S_M) to amplify sensitivity, ensuring that dynamic moisture signals remain the primary factor and are not overshadowed by static landscape variables. The edaphic component further differentiates vulnerability through the integrated weighting of Peat Risk (P_R), Land Cover Risk (LC_R), and the Human Accessibility Index (HAI) with coefficients of 0.4, 0.3, and 0.3, respectively.

$$LVI_{raw} = 2S_M - (0.4P_R + 0.3LC_R + 0.3HAI_R) \quad (5)$$

Where S_M is soil moisture, P_R is the peat risk indicator, LC_R is the land cover risk score, and HAI is the Human Accessibility Index. To ensure the robustness of these weights, a sensitivity analysis was conducted by varying each weight by ± 5 –10%. The results showed that the model remained stable, with no significant fluctuations in the Fuzzy Partition Coefficient (FPC), confirming that the selected weights consistently represent the landscape's physical vulnerability. Furthermore, stepwise validation confirmed that the integration of these specific weights

significantly improved the model's Pearson correlation (r) and R^2 compared to a model without edaphic factors, validating its accuracy in representing field conditions.

The resulting LVI_{raw} value is negatively correlated with fire risk. To integrate it with the VPD, which is positively correlated with risk, the LVI is inverted to produce the LVI_{norm} , as described in (6).

$$LVI_{scale} = \frac{(LVI_{raw} - LVI_{min})}{(LVI_{max} - LVI_{min})} \quad (6)$$

Next, an inversion is performed to align the direction of the correlation so that higher index values represent greater risk, as explained in (7).

$$LVI_{norm} = 1 - LVI_{scale} \quad (7)$$

This transformation ensures that LVI_{norm} close to 1 represents the most extreme land drought conditions and has a positive impact on fire risk.

It is worth emphasizing that the VLIF-Model deliberately reduces complexity by focusing solely on the bivariate interaction between atmospheric water demand (VPD) and edaphic vulnerability (LVI). The goal is not to fully represent all mechanisms, such as lightning ignitions, agricultural burning schedules, or peat-specific combustion chemistry. Rather, the objective is to design a parsimonious bivariate index that operationalizes the logic of soil-atmosphere coupling, which existing literature suggests plays a major role in creating fire-prone conditions in tropical ecosystems [3],[4]. We do not claim that this empirical causal relationship is more dominant than other factors. Rather, we argue that this bivariate design enhances interpretability, reduces multicollinearity, and preserves spatial resolution during the peak fire season, as will be demonstrated in the clustering results.

2.4 Bivariate Fuzzy C-Means Clustering for Risk Zonation

After calculating the VPD and LVI indices, the next step is to group the data into three clusters using the Fuzzy C-Means (FCM) algorithm. The FCM algorithm was chosen because it is well-suited for capturing gradual transitions in fire-related conditions, where cluster boundaries are not strictly discrete [12]. In this study, the algorithm generated unlabeled numerical clusters (e.g., Cluster 0, Cluster 1, and Cluster 2), and these clusters were interpreted as Low, Moderate, and High risk only after analyzing the cluster centroids and their relationship with the observed hotspot patterns. The computational workflow begins with feature scaling using Z-score standardization, $z = (x - \mu)/\sigma$, to ensure that distance-based clustering is not biased by the magnitude of the input values. This process is performed using two input features (normalized VPD values and normalized LVI values) by minimizing the objective function described in Equation (8):

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2 \quad (8)$$

with the following parameter configurations derived from the experimental setup:

- Number of Clusters (C) = 3; The algorithm divides the data into three numerical clusters (0, 1, 2). Risk labels (Low, Moderate, and High) are then assigned based on the relative position of the cluster centroids in the bivariate feature space.
- Fuzziness Exponent (m) = 2; this adjusts the sharpness and ambiguity of a particular cluster.
- Convergence threshold (ϵ) = 0.005; this is the rule that ensures the iteration stops when the change in the maximum membership degree is less than this value.
- Maximum iterations = 1000, This is the upper bound that can ensure convergence, given the 2,614,056 data points available for clustering (a large dataset). The maximum bound applied here was chosen based on established practice in empirical studies [30].

This algorithm iteratively updates the cluster centers (c_j) and the membership matrix (u_{ij}) until the convergence criteria are met. The final output is the membership degree for each grid cell in the three clusters. For mapping purposes, a defuzzification step is performed by assigning each

grid cell to the cluster with the highest membership degree ($\max(u_{ij})$). However, continuous membership values are retained to represent the risk gradient, which reflects the natural transitions between fire risk categories.

2.5. Spatial and Statistical Validation

To validate the reliability of the forest and land fire risk classification model, three validation approaches were employed. Before detailing these approaches, it is important to explain the rationale for selecting the validation variables. This study used VIIRS 375-m active fire detection as the primary ground truth because it is the most consistent and near-real-time operational indicator of active fires in Indonesia. Although it has limitations as an ignition proxy and does not represent burned area, the VIIRS hotspot has been widely used in tropical fire research due to its high temporal resolution (daily) and its ability to detect small fires (down to ~375 m) [23]. Additionally, Fire Radiative Power (FRP) was used as a secondary validation metric. FRP provides a physical measure of fire intensity (in MW), so the model not only assesses whether a fire is occurring but also how intense it is. This two-metric approach (hotspot presence for fire occurrence, FRP for intensity) ensures that high-risk zones are evaluated both in terms of the likelihood of ignition and the potential severity of the fire [31].

First, internal cluster validation was performed using the Fuzzy Partition Coefficient (FPC) obtained from the convergence process of the FCM algorithm, along with the elbow method to determine the optimal number of clusters ($C=3$). Cluster centroids (normalized VPD and LVI) were used to interpret the FCM output, with Low, Moderate, and High labels assigned post hoc based on centroid ordering and supported by hotspot distribution. Second, spatial efficiency validation was performed by calculating the area coverage percentage, the percentage of detected hotspots, hotspot density per grid, and the lift factor for each risk zone (Low, Moderate, High). The lift factor was obtained from the ratio of a zone's hotspot density to the average density of the entire region. Third, temporal validation was performed using a density scatter plot to test the relationship between increases in LVI values and hotspot density over time, including during wet and dry periods.

3. RESULTS AND DISCUSSION

The results and discussion section of this paper describes how well the bivariate VPD-LVI clustering model performs with respect to validation using VIIRS hotspot data. Four different analyses are performed: cluster partition stability, spatial efficiency, temporal statistical validation, and comparison with conventional multivariate methods (using more than three individual variables).

3.1. Cluster Stability and Optimal Number of Clusters

The structural integrity of the bivariate FCM algorithm (using normalized VPD and LVI) was validated using the Fuzzy Partition Coefficient (FPC) across various numbers of clusters ($k = 2$ to 6). This analysis was performed on the entire dataset ($N = 2,614,056$) to ensure overall population validity. As shown in Figure 2, the FPC values are 0.7379 for $k = 2$, 0.6137 for $k = 3$, 0.5443 for $k = 4$, 0.5149 for $k = 5$, and 0.4873 for $k = 6$. Although $k = 2$ yields the highest FPC value (0.7379), indicating the most compact cluster structure, the two-cluster solution is insufficient for operational fire risk management, which requires three distinguishable partitions. Therefore, the three-cluster solution ($k = 3$) was selected as the optimal configuration, as it provides a balance between statistical validity and practical interpretability. These three numerical clusters (clusters 0, 1, and 2) were then interpreted as Low, Moderate, and High risk based on the order of their cluster centers in the bivariate VPD-LVI space. The FPC value of 0.6137 indicates that the three risk levels can be statistically distinguished and that the cluster structure is robust despite the high complexity of the environmental dataset.

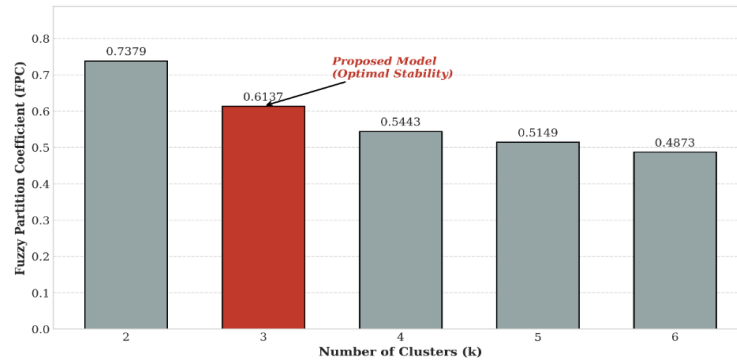


Figure 2. Fuzzy Partition Coefficient (FPC) Across Different Numbers of Clusters ($k=2-6$).

Based on this structural validation, the FCM algorithm generates cluster centroids in a two-dimensional feature space comprising normalized VPD and normalized LVI. Figure 3 presents a scatter plot of all grid points colored according to the FCM output clusters, along with the cluster centroids (marked with stars). The cluster with the lowest centroid values in normalized VPD and LVI was interpreted as Low risk (VPD_norm = 0.161, LVI_norm = 0.239). The intermediate cluster was interpreted as Moderate risk (VPD_norm = 0.198, LVI_norm = 0.392). The cluster with the highest centroid values was interpreted as High risk (VPD_norm = 0.344, LVI_norm = 0.558). This ranking reflects the post hoc operational labeling of the numerical clusters rather than a priori class supervision. The clear separation between these clusters indicates that each risk category has statistically distinct characteristics, reflecting a dual gradient of critical conditions (atmospheric demand and land vulnerability) in the study area.

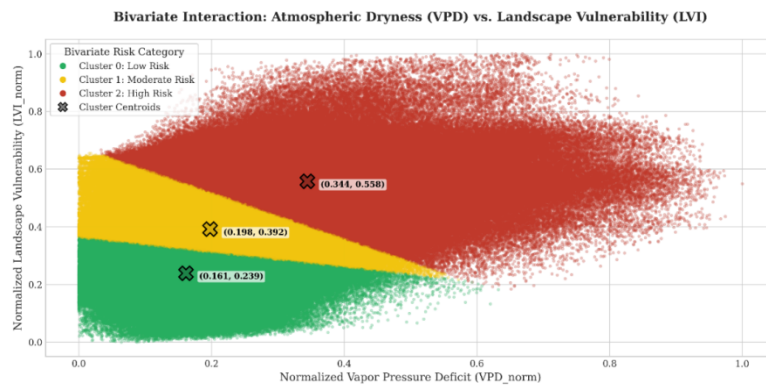


Figure 3. Cluster Centroids of the Bivariate VPD-LVI FCM Output and Post Hoc Risk Labels

3.2. Spatial Efficiency of Risk Clusters

The spatial efficiency of the bivariate FCM VPD-LVI model is quantified in Table 2, which reports the proportion of area, hotspot capture rate, lift factor, and average fire radiative power (FRP) for each risk cluster. The cluster interpreted as Low risk occupies 40.59% of the study area but contains only 11.09% of the hotspots (lift = 0.27, mean FRP = 7.24), indicating that the model effectively excludes non-fire areas from High-risk classifications. The moderate-risk cluster covers a similar portion of the area (41.61%) and captures 26.62% of the hotspots (lift = 0.64, mean FRP = 8.44), acting as a transition zone with moderate fire activity.

Table 2. Spatial Efficiency of FCM Result Risk Clusters

Cluster	Area (%)	Hotspot (%)	Lift Factor	Mean FRP
Low	40.59%	11.0931	0.27	7.24
Moderate	41.61%	26.6227	0.64	8.44
High	17.80%	62.2842	3.50	21.46

The High-Risk cluster demonstrates the operational value of this model: this cluster covers only 17.80% of the total area but accounts for 62.28% of all hotspots, resulting in a lift factor of 3.50, which means the hotspot density is 3.5 times higher than the regional average. Furthermore, the mean FRP in the High-Risk cluster (21.46) is substantially higher than in the Low (7.24) and Moderate (8.44) clusters, indicating that fires in this zone are not only more frequent but also more intense (higher radiative power). This combination of high hotspot concentration and high FRP validates that the bivariate VPD-LVI interaction effectively isolates the most hazardous areas. Consequently, authorities can focus mitigation efforts on just 17.8% of the landscape, thereby significantly improving operational efficiency.

3.3. Temporal Validation and Predictive Accuracy

To evaluate the predictive accuracy of the bivariate VPD-LVI classification model in distinguishing fire-prone areas from non-fire-prone areas, a Receiver Operating Characteristic (ROC) curve analysis was conducted. The ROC curve is a standard tool for binary classification tasks, plotting True Positive Rate (Sensitivity) against False Positive Rate ($1 - \text{Specificity}$) at various probability thresholds. Here, true positivity is defined as the presence of at least one VIIRS hotspot ($n=61,015$) within each spatio-temporal grid cell, while the model's predicted probability for high risk is derived from the normalized membership degree of the High-risk cluster (cluster 2) from the FCM output. The analysis was conducted on the full dataset ($N = 2,614,056$ grid observations from 2023–2024), ensuring a robust evaluation of temporal and spatial prediction performance.

Figure 4 presents the ROC curve, which shows a steep initial rise toward the upper-left corner, indicating that the model achieves a high true positive rate while maintaining a low false positive rate. The area under the curve (AUC) is 0.8288, which is significantly above the baseline for a random classifier (0.5). This value is generally interpreted as “good” to “very good” discriminatory ability in environmental risk modeling. In the context of fire risk zoning, an ROC-AUC of 0.8288 means that if a random High-risk grid and a random low-risk grid are selected, the model will assign a higher risk probability to the actual fire-affected grid with a probability of approximately 82.9%. This result indicates that the bivariate VPD-LVI interaction captures important temporal patterns of fire events across Sumatra's seasonal cycle, without relying on monthly aggregate regression.

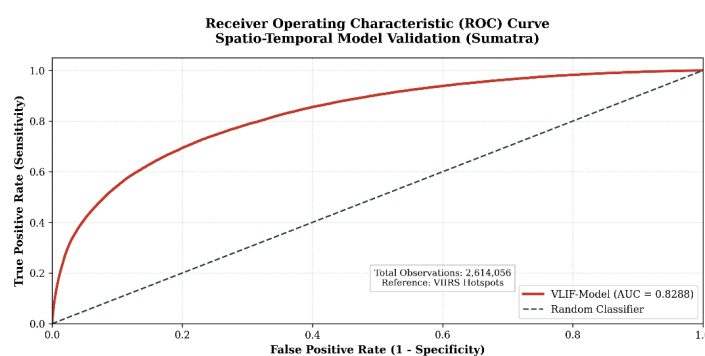


Figure 4. ROC-AUC Curve (Accuracy Validation)

In addition to ROC-AUC, additional metrics derived from the confusion matrix and cluster validity analysis are summarized in Table 3. These metrics provide a multifaceted assessment of the model's reliability. A recall value of 0.8890 indicates that the model successfully detected nearly 89% of actual fire points within a grid when the positive class was defined as the combination of moderate- and High-risk clusters, demonstrating its ability to capture the majority of fire incidents. The Brier value of 0.0169 is very low (close to 0), reflecting minimal probabilistic calibration error, meaning that the predicted membership probabilities align with the observed fire

frequencies. A Fuzzy Partition Coefficient (FPC) of 0.6137 indicates stable cluster separation for the three-cluster solution, while a High Cluster Efficiency Ratio (Lift) of 3.50 indicates that High-risk zones have a fire point density 3.5 times higher than the regional average. Collectively, these indicators validate that the bivariate FCM VPD-LVI model provides temporal predictive accuracy and spatial efficiency. The VLIF-Model thus offers a precise, operationally efficient zoning tool that can support year-round spatial targeting as a complementary decision-support layer for fire prevention. Future work should focus on downscaling to village-level resolution, integrating real-time socio-economic triggers, and expanding to near-real-time forecasting to further enhance its utility as a comprehensive decision-support system for disaster mitigation in Sumatra and surrounding areas.

Table 3. Summary of Statistical Evaluation and Predictive Reliability of VLIF-Model.

Statistical Indicator	Statistical Metric	Value	Interpretation
Predictive Accuracy	ROC-AUC	0.8288	Robust Discriminative
Probabilistic Calibration	Brier Score	0.0169	Minimal Calibration Error
Detection Capability	Recall	0.8890	Optimal Detection Sensitivity
Structural Stability	Fuzzy Partition Coefficient (FPC)	0.6137	Robust Cluster Stability
Operational Impact	High Cluster Efficiency Ratio (ER)	3.4989	Efficient Spatial Isolation

3.4. Comparative Performance: Multivariate Baselines vs. VLIF-Model

The multivariate baseline model was constructed using four input variables: dew point temperature at 2 meters (d2m), air temperature at 2 meters (t2m), volumetric soil water layer 1 (swv11), and the Human Accessibility Index (HAI). Structural integrity analysis revealed significant multicollinearity among these input variables. The conventional multivariate model showed an extremely high Average Variance Inflation Factor (VIF) of 50.8990, indicating severe information redundancy, particularly between the temperature and dew point variables. In contrast, the VLIF-Model successfully reduced the VIF value to 4.8355 (below the critical threshold of 5.0). This indicates that physical integration via a bivariate scheme simplifies the model architecture and significantly reduces the interdependence among variables that often undermine the stability of traditional forest fire monitoring models.

Table 4. Performance Comparison: Multivariate Baseline vs. VLIF-Model

Evaluation Criteria	Multivariate Baseline	VLIF-Model (Proposed)
Number of Input Variables	4 Features	2 Features (VPD, LVI)
Average VIF (Multicollinearity)	50.8990	4.8355
High-Risk Area Coverage (%)	50.36%	17.80%
High-Risk Hotspots Captured (%)	58.79%	62.28%
Efficiency Ratio (ER)	1.1673	3.4989
High Risk Mean FRP (Fire Intensity)	15.2810	21.4593

Operational efficiency is the key differentiator between the two approaches. Based on the results presented in Table 4, the multivariate model tends to be “area-greedy,” classifying 50.36% of the total area of Sumatra as a High-risk zone. Although it covers a very large area, this model is only able to capture 58.79% of the total thermal anomalies present. This phenomenon results in a low Efficiency Ratio (ER) of 1.1673, meaning the model is barely better than randomly designating risk areas. This places a heavy operational burden on authorities due to the vast extent of the area that must be prioritized.

In contrast, the VLIF-Model demonstrated significantly higher spatial accuracy by classifying only 17.80% of the area as a High-risk zone, yet it was able to detect a higher percentage of hotspots, namely 62.28%. This improvement in performance is reflected in the ER value, which increased to 3.4989, indicating that the VLIF-Model is significantly more efficient at identifying fire-prone areas compared to the baseline model. Physical validation using Mean Fire Radiative Power (FRP) further supports these findings. High-risk zones in the VLIF-Model have an

average fire intensity of 21.4593, higher than that of the multivariate model (15.2810). This indicates that the integration of VPD-LVI not only incorporates fire events but also helps identify landscapes with more vulnerable fuel characteristics and higher potential for intense fires.

3.5. Spatio-Temporal Evolution: Wet vs. Dry Season Analysis

To evaluate the consistency of the VPD-LVI bivariate model across seasonal cycles and to understand how fire risk patterns evolve spatially, a quarterly performance analysis was conducted using two key metrics: recall (the proportion of hotspots captured by the combined High and Moderate clusters) and efficiency ratio (the lift factor of the High cluster alone). Data were collected over eight quarters (Q1 2023 to Q4 2024) from a total of 2,614,056 spatio-temporal grid cells. Figure 5 presents the dynamics of these two metrics side by side, allowing readers to observe the relationship between the model's detection capability and its spatial efficiency across seasons.

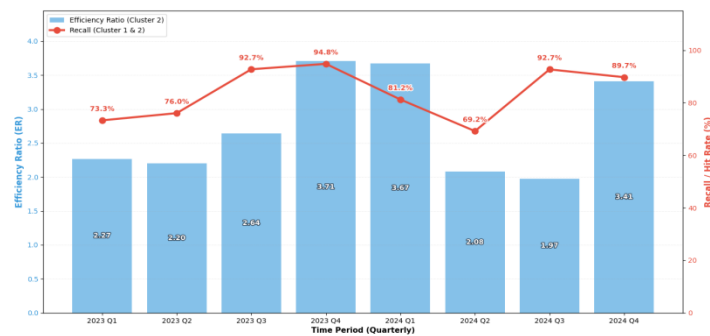


Figure 5. Seasonal Performance Dynamics: Recall vs Efficiency (2023-2024)

As shown in Figure 5, the model's recall remains consistently high throughout the study period, always exceeding 69% and peaking during the dry seasons, particularly in Q3 of both years and Q4 of 2023. The highest recall occurred in Q4 2023 (94.8%), followed closely by Q3 2024 (92.7%) and Q3 2023 (92.7%). Even during wet quarters, recall never fell below 69% – the lowest observed value was 69.2% in Q2 2024 – while other wet quarters such as Q1 2023 (73.3%) and Q1 2024 (81.2%) were higher. This demonstrates that the model does not lose predictive capability under high rainfall conditions. The notably higher recall in Q1 2024 compared to Q1 2023 (81.2% vs. 73.3%) suggests inter-annual variability in fire distribution, but the model adapts well.

The efficiency ratio (lift of the High cluster) exhibits a different, complementary pattern. The highest efficiency ratios occur during transitional quarters: Q4 2023 (3.71), Q1 2024 (3.67), and Q4 2024 (3.41). In contrast, at the peak of the dry season, the lift is substantially lower: Q3 2023 (2.64) and Q3 2024 (1.97). The lowest efficiency ratio (1.97 in Q3 2024) coincides with the second-highest recall (92.7%), clearly illustrating the trade-off between detection coverage and spatial precision. When fires become widespread during extreme dry conditions, the High zone must expand to capture nearly all active fires, which dilutes hotspot concentration and reduces the lift. Conversely, during wet or transitional seasons, fires are more localized, allowing the model to draw a narrower High zone that remains highly efficient (lift >3.4).

This inverse relationship between recall and efficiency ratio is physically intuitive and operationally valuable. The model intelligently adjusts its spatial coverage to seasonal threat levels: it expands the High zone to maximize detection during peak fire seasons (high recall, moderate lift) and contracts it to focus resources precisely when fires are rare and clustered (lower recall, high lift). Such adaptive behavior is essential for a spatial decision-support framework that must balance detection coverage and spatial precision across contrasting climatic regimes.

To complement this temporal analysis, Figure 6 displays maps of the spatial distribution of fire risk by quarter (Q1–Q4) compiled from 2023–2024. Each grid cell is classified into three categories: Low (green), Moderate (yellow), and High (red). For each quarterly map, the risk class assigned to a grid cell is determined by the mode (most frequent cluster) of daily risk classifications within that quarter, ensuring that the dominant vulnerability condition is represented. These maps visually indicate that the High-risk zones expand and contract seasonally, with the largest

contiguous High-risk areas appearing in Q3 and Q4, particularly in Riau, Jambi, and South Sumatra.

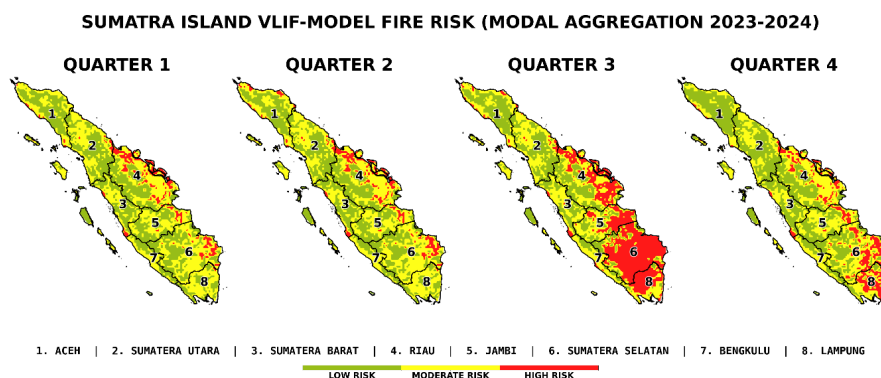


Figure 6. Spatio-Temporal Monitoring of Sumatera Forest And Land Fire Risk

Figure 6 clearly shows the spatial evolution. In the first quarter (rainy season, January–March), the High zone was very limited and fragmented, primarily located in eastern Riau and small pockets in South Sumatra, only about 7.3% falls into the High category. Entering the second quarter (April–June), a moderate expansion occurs: the High zone begins to spread across Riau, Lampung, and parts of Jambi, though the proportion of High remains low (7.7%). The peak occurs in the third quarter (July–September), which is the dry season. The High zone expanded dramatically to cover large, contiguous areas in Riau, Jambi, South Sumatra, and even reached the eastern coast of North Sumatra. The proportion of the High zone surged to 32.2%, while the Moderate zone also expanded. This aligns with the peak in recall and the decline in lift shown in Figure 5. Entering the fourth quarter (October–December), the High zone began to shrink (12.6%) but persisted in Riau, South Sumatra, and Lampung, indicating that the southern and central regions of Sumatra have a longer period of vulnerability compared to the northern and western regions.

More specifically, the spatial aggregation by province shows that South Sumatra has the highest proportion of High-risk grids (30.1%), followed by Lampung (26.3%) and Riau (21.8%). In contrast, provinces along the Bukit Barisan range, such as Aceh (3.8%), North Sumatra (2.4%), West Sumatra (3.9%), and Bengkulu (1.3%), have almost no Hazard zones. This pattern aligns closely with the distribution of peatlands, higher levels of atmospheric aridity, and greater human accessibility (HAI) in the eastern lowlands.

Overall, Figure 6 demonstrates that the VPD-LVI bivariate FCM model produces risk zonation that is spatially and temporally robust. The model maintains a very high detection rate (recall >92%) during the peak dry season (Q3) while showing high efficiency (lift often exceeding 3.4) during the transition season. It never produced a zero hazard rating, even during the rainy season, demonstrating its ability to detect local fires triggered by short-term drought or human activity. Thus, this model is well-suited for year-round spatial targeting as a complementary decision-support layer for forest and land fire prevention in Sumatra.

3.6. Practical Implications and Institutional Linkage

The practical implications of the VLIF-Model lie in its ability to provide targeted spatial risk intelligence for fire management in Indonesia. Unlike the conventional Fire Danger Rating System (FDRS) operated by BMKG, which issues broad-area warnings (e.g., “red alert” for entire provinces), VLIF translates the same underlying physical drivers into actionable grid-level information. Specifically, the model shows that 62.28% of thermal anomalies occur in “High Risk” areas, which cover only 17.80% of the mapped area (lift factor = 3.50). Thus, this model reduces resource imbalances while maintaining high detection performance during peak vulnerability periods. In addition, VLIF adds scientifically-based spatial discrimination to the FDRS, particularly in peatland-dominated provinces such as Riau and South Sumatra, answering not only

when conditions are dangerous (FDRS) but exactly where that danger is most concentrated. This supports a precise resource allocation strategy for the National Disaster Management Agency (BNPB) and the Ministry of Environment and Forestry, enabling more targeted operations, from prioritizing patrol routes to allocating water-bombing assets. In areas where water resources and technical assets are unevenly distributed, maintaining a stable efficiency ratio under extreme drought conditions minimizes false alarms, prevents excessive expansion of high-risk zones, and ensures cost-efficient mitigation management. Therefore, the VLIF-Model functions not only as a predictive model but also as a targeted spatial risk intelligence layer that can be scaled within the national early warning ecosystem, not replacing FDRS but adding precision where it is most needed.

3.7. Limitations and Future Work

The 0.1° resolution (approximately 11 km) used in this study was deliberately chosen as a mesoscale operational layer, designed for provincial and regional level priorities rather than village level tactical interventions. Consequently, the VLIF-Model is not intended to replace local scale decision making systems. Instead, it serves as a strategic screening layer, namely a pre-operational decision support and spatial pre-emption tool that narrows down broad risk areas before more detailed and resource intensive assessments are applied.

One inherent limitation of this study is the use of VIIRS thermal anomalies as a proxy for fire occurrence. VIIRS hotspots represent ignition proxies rather than directly measured burned area or fire intensity. False positives can arise from agricultural waste burning or industrial heat sources, while false negatives may occur due to cloud cover or low intensity surface fires, particularly in peatlands. However, VIIRS 375 m active fire data remain the most widely accepted operational indicator for near-real-time fire detection in Indonesia, and their consistency across the study period (61,015 hotspots) provides a robust basis for validation. Future work could integrate burned area products (e.g., MODIS MCD64A1) or higher resolution Sentinel-2 imagery to refine ground truth labeling.

Although the 2023–2024 dataset includes the extreme El Niño event of 2023, one of the strongest in the last decade, it does not capture the full spectrum of ENSO and Indian Ocean Dipole (IOD) variability. Extreme La Niña or negative IOD phases, which may alter fire risk patterns differently, are not represented. Consequently, model performance under contrasting climatic extremes (e.g., very wet versus very dry) remains to be further validated. Extending the analysis to multiple El Niño and La Niña cycles would strengthen the temporal robustness claim.

Several directions remain open for further enhancement. First, higher resolution downscaling techniques, such as machine learning based disaggregation or fusion with Sentinel-2 and Landsat imagery, could enable operational implementation at the village or sub-district level. Second, the bivariate structure (VPD and LVI) successfully reduces multicollinearity and maintains spatial sharpness, but incorporating additional dynamic socio-economic indicators, such as real-time land clearing alerts, plantation activity schedules, or economic drivers of fire use, could further improve predictive accuracy and contextual relevance. Third, the model currently relies on historical data from 2023 to 2024. Extending it to near-real-time forecasting by integrating seasonal climate forecasts and daily VIIRS hotspot monitoring would enable operational early warning systems. Finally, the VLIF framework offers a scalable and adaptable architecture for other fire-prone regions, such as Kalimantan and Papua. However, its successful implementation will require recalibration of local parameters, particularly adjustments to edaphic weights for variables such as peat depth and land cover, to ensure that the model captures the specific environmental nuances of these regions.

4. CONCLUSION

This study developed the VPD-LVI Integrated Fuzzy (VLIF) Model, a physically based bivariate approach for assessing forest and land fire risk in tropical ecosystems. Its main innovations include: (1) the integration of atmospheric water demand (VPD) with a multidimensional Land Vulnerability Index (LVI) that combines soil moisture, peat risk, land cover

risk, and the newly developed Human Accessibility Index (HAI); (2) the use of high resolution ERA5-Land data at a spatial resolution of 0.1°; and (3) the application of bivariate Fuzzy C-Means clustering to generate an objective three tier risk zoning (Low, Moderate, High). This model helps address the issue of overestimated risk in the "High" zone during peak drought periods, reducing the high risk area to just 17.80% of the landscape while capturing 62.28% of actual fire hotspots, achieving an enhancement factor of 3.50. Validation against 61,015 VIIRS thermal anomalies indicated good prediction reliability, with an ROC-AUC of 0.8288 and a Brier score of 0.0169. The VLIF-Model provides a spatially selective and operationally efficient zoning tool that can support year-round spatial targeting as a complementary decision support layer for fire prevention. Unlike conventional indices that issue broad area alerts, VLIF delivers grid level intelligence, making it a targeted spatial risk intelligence framework for pre-emptive mitigation. Future work should focus on downscaling to village level resolution, integrating real-time socio-economic triggers, and expanding to near-real-time forecasting to enhance the model's utility as a comprehensive decision support system for disaster mitigation in Sumatra and the surrounding region.

ACKNOWLEDGEMENTS

The author would like to thank NASA Earth Observing System Data and Information System (EOSDIS), European Centre for Medium-Range Weather Forecasts, Global Forest Watch, and the Peta Nusa Team for providing access to the open data used in this study, as described in subsection 2.1. The author also thanks Geofabrik GmbH (Karlsruhe, Germany), founded by Jochen Topf and Frederik Ramm, for providing free OpenStreetMap data extracts for roads and settlements in the Sumatra region.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

DATA OR SOFTWARE RESOURCES

The VLIF-Model source code, spatial dataset, and implementation protocol are permanently stored and accessible via GitHub at <https://github.com/adefirmansyahuti/VLIF-MODEL-SUMATRA> [26].

REFERENCES

- [1] S. Nurdianti *et al.*, "The impact of El Niño southern oscillation and Indian Ocean Dipole on the burned area in Indonesia," *Terr. Atmos. Ocean. Sci.*, vol. 33, no. 15, 2022, doi: 10.1007/S44195-022-00016-0.
- [2] L. Kiely *et al.*, "Assessing costs of Indonesian fires and the benefits of restoring peatland," *Nat. Commun.*, vol. 12, no. 1, pp. 1–11, 2021, doi: 10.5285/fdae44ed-8b22-4935-b889-b4b271138385.
- [3] S. Mezbahuddin *et al.*, "Accuracy of tropical peat and non-peat fire forecasts enhanced by simulating hydrology," *Sci. Rep.*, vol. 13, no. 1, pp. 1–10, 2023, doi: 10.1038/s41598-022-27075-0.
- [4] L. Mareta, A. L. Latifah, and R. Hidayat, "The Influence of Natural and Anthropogenic Factors on the Area of Forest and Land Fires in Kalimantan," *J. Tanah dan Iklim*, vol. 43, no. 2, pp. 143–151, 2019.
- [5] H. Clarke *et al.*, "Forest fire threatens global carbon sinks and population centres under rising atmospheric water demand," *Nat. Commun.*, vol. 13, no. 1, pp. 1–10, 2022, doi: 10.1038/s41467-022-34966-3.
- [6] P. A. Levine *et al.*, "Water Stress Dominates 21st-Century Tropical Land Carbon Uptake," *Global Biogeochem. Cycles*, vol. 37, no. 12, pp. 1–17, 2023, doi: 10.1029/2023GB007702.
- [7] A. D. Nurhayati, B. Hero Saharjo, L. Sundawati, Syartinilia, M. A. Cochrane, and C. Hély, "Anthropogenic factors in a forest and peatland fire danger rating system in South Sumatra," *Environ. Res. Lett.*, vol. 20, no. 6, 2025, doi: 10.1088/1748-9326/add609.
- [8] L. L. B. Graham, V. Arifanti, I. Malik, M. A. Pribadi, and G. B. Applegate, "Steps towards

- the development of a Peat Fire Danger Rating System in Indonesia,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 874, no. 1, p. 012010, 2021, doi: 10.1088/1755-1315/874/1/012010.
- [9] N. Itsnaini, B. Sasmito, A. Sukmono, and I. Prasarti, “Analisis Hubungan Curah Hujan Dan Parameter Sistem Peringkat Bahaya Kebakaran (SPBK) Dengan Kejadian Kebakaran Hutan dan Lahan untuk Menentukan Nilai Ambang Batas Kebakaran,” *J. Geod. Undip*, vol. 6, no. 2, pp. 62–70, 2017, doi: 10.14710/jgundip.2017.16257.
- [10] H. Hayasaka, “Peatland Fire Weather Conditions in Sumatra, Indonesia,” *Climate*, vol. 11, no. 5, 2023, doi: 10.3390/cli11050092.
- [11] À. Nebot and F. Mugica, “Forest fire forecasting using fuzzy logic models,” *Forests*, vol. 12, no. 8, 2021, doi: 10.3390/f12081005.
- [12] A. A. Jafarzadeh, A. Mahdavi, and H. Jafarzadeh, “Evaluation of forest fire risk using the Apriori algorithm and fuzzy c-means clustering,” *J. For. Sci.*, vol. 63, no. 8, pp. 370–380, 2017, doi: 10.17221/7/2017-JFS.
- [13] F. Rohimahastuti, A. Djuraidah, and H. Wijayanto, “Spatio-temporal Bayes model for estimating the number of hotspots as an indicator of forest and land fires in Kalimantan Island, Indonesia,” *J. Agrometeorol.*, vol. 27, no. 1, pp. 27–32, 2025, doi: 10.54386/jam.v27i1.2761.
- [14] A. Vatesia, R. R. Rais, and R. S. Miliana, “Spatio-Temporal Clustering for Fire Hot Spots Occurrence Over Sulawesi Island,” *2020 FORTEI-International Conf. Electr. Eng.*, pp. 156–161, 2020, doi: 10.1109/FORTEI-ICEE50915.2020.9249779.
- [15] Louis Anthony Cox Jr., “Implications of nonlinearity, confounding, and interactions for estimating exposure concentration-response functions in quantitative risk analysis,” *Environ. Res.*, vol. 182, p. 109026, 2020, doi: 10.1016/j.envres.2019.109026.
- [16] C. P. R. McCarter *et al.*, “Peat fires and legacy toxic metal release: An integrative biogeochemical and ecohydrological conceptual framework,” *Earth-Science Rev.*, vol. 256, no. December 2023, p. 104867, 2024, doi: 10.1016/j.earscirev.2024.104867.
- [17] A. J. Kohlenberg, M. R. Turetsky, D. K. Thompson, B. A. Branfireun, and C. P. J. Mitchell, “Controls on boreal peat combustion and resulting emissions of carbon and mercury,” *Environ. Res. Lett.*, vol. 13, no. 3, 2018, doi: 10.1088/1748-9326/aa9ea8.
- [18] A. J. Rigden, R. S. Powell, A. Trevino, K. A. McColl, and P. Huybers, “Microwave Retrievals of Soil Moisture Improve Grassland Wildfire Predictions,” *Geophys. Res. Lett.*, vol. 47, no. 23, pp. 1–8, 2020, doi: 10.1029/2020GL091410.
- [19] T. Yulianto, A. F. Rahmah, F. Faisol, and R. Amalia, “Clustering Daerah Bencana Alam di Indonesia Menggunakan Metode Fuzzy C-Means,” *Unisda J. Math. Comput. Sci.*, vol. 9, no. 2, pp. 29–39, 2023, doi: 10.52166/ujmc.v9i2.4776.
- [20] N. A. Khairani and E. Sutoyo, “Application of K-Means Clustering Algorithm for Determination of Fire-Prone Areas Utilizing Hotspots in West Kalimantan Province,” *Int. J. Adv. Data Inf. Syst.*, vol. 1, no. 1, pp. 9–16, 2020, doi: 10.25008/ijadis.v1i1.13.
- [21] W. Schroeder, P. Oliva, L. Giglio, and I. A. Csiszar, “The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment,” *Remote Sens. Environ.*, vol. 143, pp. 85–96, 2014, doi: 10.1016/j.rse.2013.12.008.
- [22] A. S. Nur, Y. J. Kim, J. Lee, and C. W. Lee, “Spatial Prediction of Wildfire Susceptibility Using Hybrid Machine Learning Models Based on Support Vector Regression in Sydney, Australia,” *Remote Sens.*, vol. 15, no. 3, 2023, doi: 10.3390/rs15030760.
- [23] P. Sofan, D. Bruce, W. Schroeder, E. Jones, and J. Marsden, “Assessment of VIIRS 375 m active fire using tropical peatland combustion algorithm applied to Landsat-8 over Indonesia’s peatlands,” *Int. J. Digit. Earth*, vol. 13, pp. 1695–1716, 2020, doi: 10.1080/17538947.2020.1791268.
- [24] H. Hersbach *et al.*, “The ERA5 global reanalysis,” *Q. J. R. Meteorol. Soc.*, vol. 146, no. 730, pp. 1999–2049, 2020, doi: 10.1002/qj.3803.
- [25] N. Sidhu, E. Pebesma, and Y. C. Wang, “Usability study to assess the IGBP land cover classification for Singapore,” *Remote Sens.*, vol. 9, no. 10, pp. 1–15, 2017, doi: 10.3390/rs9101075.
- [26] A. Firmansyah, “VLIF-Model Sumatra: Spatio-Temporal Forest and Land Fire Risk

- Modeling,” 2026, *GitHub Repository*. [Online]. Available: <https://github.com/adefirmansyahuti/VLIF-MODEL-SUMATRA>
- [27] R. Seager, A. Hooks, A. P. Williams, B. Cook, J. Nakamura, and N. Henderson, “Climatology, variability, and trends in the U.S. Vapor pressure deficit, an important fire-related meteorological quantity,” *J. Appl. Meteorol. Climatol.*, vol. 54, no. 6, pp. 1121–1141, 2015, doi: 10.1175/JAMC-D-14-0321.1.
- [28] J. J. Brown, M. Pascual, M. C. Wimberly, L. R. Johnson, and C. C. Murdock, “Humidity – The overlooked variable in the thermal biology of mosquito-borne disease,” *Ecol. Lett.*, vol. 26, no. 7, pp. 1029–1049, 2023, doi: 10.1111/ele.14228.
- [29] Prayoto, M. I. Ishihara, R. Firdaus, and N. Nakagoshi, “Peatland Fires in Riau, Indonesia, in Relation to Land Cover Type, Land Management, Landholder, and Spatial Management,” *J. Environ. Prot. (Irvine, Calif.)*, vol. 08, no. 11, pp. 1312–1332, 2017, doi: 10.4236/jep.2017.811081.
- [30] F. Mei, T. Boy, R. Hidayat, R. I. Hidayat, and M. H. Fasya, “A Fuzzy C-Means – Based Clustering Model for Analyzing TOEFL Prediction Scores in Higher Education,” vol. 10, no. 1, pp. 115–124, 2026, doi: <https://doi.org/10.30871/jaic.v10i1.11468>.
- [31] G. Roberts, M. J. Wooster, and T. Strydom, “Assessment and validation of Meteosat SEVIRI fire radiative power (FRP) retrievals over Kruger National Park,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 136, no. October 2024, p. 104375, 2025, doi: 10.1016/j.jag.2025.104375.