

Article

Underutilized Feature Extraction Methods for Burn Severity Mapping: A Comprehensive Evaluation

Linh Nguyen Van  and Giha Lee *School of Advanced Science and Technology Coverage, Kyungpook National University,
Sangju 37224, Republic of Korea

* Correspondence: leegiha@knu.ac.kr; Tel.: +82-10-4057-5032

Abstract: Wildfires increasingly threaten ecosystems and infrastructure, making accurate burn severity mapping (BSM) essential for effective disaster response and environmental management. Machine learning (ML) models utilizing satellite-derived vegetation indices are crucial for assessing wildfire damage; however, incorporating many indices can lead to multicollinearity, reducing classification accuracy. While principal component analysis (PCA) is commonly used to address this issue, its effectiveness relative to other feature extraction (FE) methods in BSM remains underexplored. This study aims to enhance ML classifier accuracy in BSM by evaluating various FE techniques that mitigate multicollinearity among vegetation indices. Using composite burn index (CBI) data from the 2014 Carlton Complex fire in the United States as a case study, we extracted 118 vegetation indices from seven Landsat-8 spectral bands. We applied and compared 13 different FE techniques—including linear and nonlinear methods such as PCA, t-distributed stochastic neighbor embedding (t-SNE), linear discriminant analysis (LDA), Isomap, uniform manifold approximation and projection (UMAP), factor analysis (FA), independent component analysis (ICA), multidimensional scaling (MDS), truncated singular value decomposition (TSVD), non-negative matrix factorization (NMF), locally linear embedding (LLE), spectral embedding (SE), and neighborhood components analysis (NCA). The performance of these techniques was benchmarked against six ML classifiers to determine their effectiveness in improving BSM accuracy. Our results show that alternative FE techniques can outperform PCA, improving classification accuracy and computational efficiency. Techniques like LDA and NCA effectively capture nonlinear relationships critical for accurate BSM. The study contributes to the existing literature by providing a comprehensive comparison of FE methods, highlighting the potential benefits of underutilized techniques in BSM.

Keywords: burn severity mapping; feature extraction Landsat; machine learning



Citation: Nguyen Van, L.; Lee, G. Underutilized Feature Extraction Methods for Burn Severity Mapping: A Comprehensive Evaluation. *Remote Sens.* **2024**, *16*, 4339. <https://doi.org/10.3390/rs16224339>

Academic Editors: Xia Xu, Sen Lei, Yuanchao Su and Xuanwen Tao

Received: 15 September 2024
Revised: 12 November 2024
Accepted: 19 November 2024
Published: 20 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Wildfires are a recurring and increasingly severe threat to ecosystems, human infrastructure, and water resources. Accurate burn severity mapping (BSM) is essential for effective disaster response [1], ecological restoration [2], and long-term water quality management [3]. Optical satellite remote sensing has become a critical tool in wildfire severity assessment [4]. Reflectance changes are a key indicator used in BSM to distinguish between burnt and unburnt areas [5,6]. When a wildfire occurs, the vegetation and soil undergo physical and chemical changes, which alter how they reflect light across different parts of the electromagnetic spectrum [7]. Burnt vegetation absorbs more visible light and reflects less of the near-infrared (NIR) spectrum than healthy, unburned vegetation [8]. These changes in reflectance can be captured by optical satellite sensors and are used to identify and map burnt areas. Indices like the Normalized Burn Ratio (NBR) utilize these reflectance differences to assess the severity of the burn [9] and help monitor ecosystem recovery following a wildfire [10].

Implementing machine learning (ML) techniques has revolutionized remote sensing and environmental monitoring [11–17]. ML models can handle vast volumes of data and identify complex patterns that traditional empirical methods might overlook [18]. In BSM studies, ML applications have demonstrated superior accuracy [19–22]. Vegetation indices, commonly calculated from combinations of spectral bands, are crucial predictors in assessing the extent and intensity of wildfire damage [23]. These indices are derived from specific mathematical forms between spectral bands that capture variations in reflectance characteristics associated with vegetation health, such as chlorophyll content and water stress. For instance, indices like the Burn Area Index (BAI), Normalized Difference Vegetation Index (NDVI), and Normalized Burn Ratio (NBR) leverage the visible, near-infrared (NIR), and short-wave infrared (SWIR) bands to highlight differences between healthy and burned vegetation. Incorporating more indices typically improves the classification of burn severity levels. However, when too many indices derived from the same spectral bands are included in ML models, multicollinearity—a condition where predictors are highly correlated—can arise. This leads to redundant information and a subsequent reduction in classification accuracy [24].

To address this issue, techniques such as feature reduction (FR), which involves selecting a subset of the original variables, and feature extraction (FE), which transforms data into a lower-dimensional space while preserving essential characteristics, are commonly employed. Several studies have utilized random forest (RF) feature importance for dimensionality reduction in BSM. By ranking features based on their contribution to model accuracy, RF can help identify the most influential variables while reducing the dimensionality of the dataset. For instance, Quintano et al. (2020) employed RF to select key variables for burn severity classification, demonstrating that focusing on the most significant features improved overall model performance [25]. Similarly, Liu et al. (2023) applied RF to identify the spectral indices that contributed most significantly to burn severity prediction, which led to more accurate mapping results by reducing redundancy in the data [23]. In another study, Van et al. (2024) introduced a novel FR approach using an eXplainable AI (XAI) framework [24]. Their methodology incorporates qualitative and quantitative feature selection, a crucial step in the broader feature extraction process, to improve the accuracy and interpretability of ML models. By applying an RF algorithm optimized through the SHapley Additive exPlanation (SHAP) and a forward stepwise selection approach, the study narrowed down 235 potential predictors to the 12 most critical.

Principal component analysis (PCA) is one of the most widely used FE methods due to its ability to transform large sets of correlated variables into smaller, uncorrelated components. By reducing dimensionality, PCA enhances the performance of ML classifiers and other predictive models [26,27]. Numerous studies have demonstrated its effectiveness, particularly in improving the accuracy of models used in burn severity assessments, where datasets often contain complex, interrelated features. For instance, Richards (1984) found that PCA enhances the detection of fire-related changes and reduces classification errors, leading to more accurate thematic maps that distinguish between static and dynamic cover types [28]. Siegert and Ruecker (2000) found that PCA enhances the detection of burned scars in tropical rainforests by emphasizing changes in radar backscatter, making it easier to differentiate between burned and unburned areas, especially under dry conditions, leading to the more accurate identification and mapping of fire-affected regions [29]. Nielsen et al. (2002) found that PCA effectively identifies burned areas by capturing changes in vegetation index and surface temperature, leading to a more accurate classification of fire-affected pixels [30]. The study demonstrated that PCA simplifies the identification of burned areas by reducing atmospheric variability and decorrelating the data, improving the accuracy of burned area estimation from AVHRR data. Koutsias et al. (2009) applied PCA to enhance the visibility of burned areas in Landsat-7 imagery [31]. By transforming the multivariate data, they demonstrated that PCA effectively distinguishes burnt surfaces by reducing the dimensionality of the dataset, which in turn highlights the most significant changes pertinent to identifying burned areas. Chen et al. (2015) employed PCA to streamline

the handling of extensive spectral data from high spectral resolution MASTER airborne images by reducing the number of spectral bands from 50 to a smaller number of principal components, reducing the computational load and enhancing the efficiency of the analysis process [32]. Similarly, in the recent study by Kulinan et al. (2024), PCA was used as a critical step to enhance classification efficiency [20]. PCA was applied after extracting textural features from Sentinel-2 imagery to identify the most relevant features for burned area training data.

Despite the widespread use of PCA as a dimensionality reduction technique in BSM, very few studies have critically evaluated its performance relative to other FE methods. Most research has defaulted to PCA due to its simplicity and well-established status in the field. This narrow focus overlooks the potential benefits that alternative methods could offer. As advances in ML and remote sensing continue to evolve, FE techniques have shown promise in different fields but remain underexplored in the context of wildfire mapping (Table 1). The lack of comparative studies examining a broader array of FE techniques suggests a gap in the wildfire mapping literature.

Table 1. Recent research using feature extraction methods for dimension reduction.

Method	Application
t-NSE	Air quality [33] and bioinformatics [34]
LDA	Land cover mapping [35]
Isomap	Groundwater quality assessment [36], vegetation mapping [37,38], and fluid mechanics [39]
UMAP	Species distribution modeling [40] and air quality [33]
FA	Medical [41] and biological applications [42]
ICA	Mechanical fault diagnosis [43,44]
MDS	Pattern recognition [45]
TSVD	Reservoir simulation [46]
NMF	Information retrieval [47]
LLE	Fluid mechanics [39], bioinformatics [34], and medical applications [48]
SE	Bioinformatics [34]
NCA	Pattern recognition [49] and data science [50]

This study aims to enhance the accuracy of ML classifiers in BSM by evaluating FE techniques that mitigate multicollinearity in vegetation indices. Using composite burn index (CBI) data from the 2014 Carlton Complex fire that occurred in the United States as a case study, we used 112 vegetation indices from seven Landsat-8 bands. We then applied and compared 13 different FE techniques, including linear and nonlinear methods such as PCA, t-distributed stochastic neighbor embedding (t-SNE), linear discriminant analysis (LDA), Isomap, uniform manifold approximation and projection (UMAP), factor analysis (FA), independent component analysis (ICA), multidimensional scaling (MDS), truncated singular value decomposition (TSVD), non-negative matrix factorization (NMF), locally linear embedding (LLE), spectral embedding (SE), and neighborhood components analysis (NCA). The performance of these techniques is benchmarked against six ML classifiers.

In the following sections, we will provide a detailed description of our methodology, present the results of our comparative analysis, and discuss the implications of our findings for future BSM and management practices.

2. Methodology

2.1. Feature Extraction Methods

2.1.1. Principal Component Analysis (PCA)

PCA is a linear dimensionality reduction technique that transforms the original data into a set of linearly uncorrelated components called principal components [51]. These components maximize the variance in the data. The transformation is defined as:

$$Z = XW \quad (1)$$

where X is the data matrix of shape $n \times p$, with n samples (CBI data) and p features (vegetation indices), and W is the matrix of eigenvectors corresponding to the k largest eigenvalues of the covariance matrix of X .

2.1.2. Linear Discriminant Analysis (LDA)

LDA is a supervised linear method that maximizes the ratio of between-class variance to within-class variance, providing optimal class separability [52]. It projects the data into a lower-dimensional space using the following criterion:

$$W_{LDA} = \operatorname{argmax}_W \frac{W^T S_B W}{W^T S_W W} \quad (2)$$

where S_B is the between-class scatter matrix, and S_W is the within-class scatter matrix. W_{LDA} represents the projection matrix.

2.1.3. t-Distributed Stochastic Neighbor Embedding (t-SNE)

t-SNE is a nonlinear technique that embeds high-dimensional data into a lower-dimensional space by minimizing the divergence between probability distributions of pairwise point distances in high- and low-dimensional spaces [53]. Given the pairwise similarities P_{ij} in high-dimensional space and Q_{ij} in the lower-dimensional space, t-SNE minimizes:

$$C = \sum_i \sum_j P_{ij} \log \frac{P_{ij}}{Q_{ij}} \quad (3)$$

2.1.4. Isomap

Isomap is a nonlinear extension of classical multidimensional scaling (MDS) that seeks to preserve the geodesic distances between points on a manifold [54]. The geodesic distance $d_G(i, j)$ is computed as the shortest path between points in the neighborhood graph. The final embedding is found by applying classical MDS to the geodesic distance matrix.

2.1.5. Uniform Manifold Approximation and Projection (UMAP)

UMAP is a manifold learning technique that optimizes a fuzzy set representation of the local neighborhood in high-dimensional space and attempts to preserve local and global data structures [55]. UMAP seeks to minimize the cross-entropy between the fuzzy simplicial sets in high- and low-dimensional spaces:

$$C = \text{cross entropy} \left(A^{(\text{high})}, B^{(\text{low})} \right) \quad (4)$$

where $A^{(\text{high})}$ and $B^{(\text{low})}$ are the high- and low-dimensional fuzzy simplicial sets, respectively.

2.1.6. Independent Component Analysis (ICA)

ICA is a linear method that separates a multivariate signal into additive, independent components [56]. Unlike PCA, which focuses on variance, ICA assumes that the source components are non-Gaussian and statistically independent. The mixing model is defined as:

$$X = AS \quad (5)$$

where A is the mixing matrix and S is the source matrix of independent components.

2.1.7. Factor Analysis (FA)

FA is a statistical method that assumes the observed data are generated by fewer unobserved latent variables (factors) plus some noise [57]. The model is represented as:

$$X = LF + \epsilon \quad (6)$$

where L is the loading matrix, F is the latent factors, and ϵ is the noise.

2.1.8. Multidimensional Scaling (MDS)

MDS aims to preserve pairwise Euclidean distances in the low-dimensional embedding [58]. It minimizes the stress function:

$$\text{Stress}(X', X) = \sqrt{\sum_{i < j} (d_{ij}(X') - d_{ij}(X))^2} \quad (7)$$

where $d_{ij}(X')$ is the distance between points i and j in the lower-dimensional space, and $d_{ij}(X)$ is the distance in the original space.

2.1.9. Spectral Embedding (SE)

SE, or Laplacian eigenmaps, is a nonlinear method that uses the eigenvectors of the graph Laplacian to embed the data in a lower-dimensional space [59]. The graph Laplacian L is constructed from the adjacency matrix of the nearest-neighbor graph and the embedding is derived by solving the generalized eigenvalue problem:

$$Ly = \lambda Dy \quad (8)$$

where D is the degree matrix, $L = D - A$ is the graph Laplacian, and y is the embedding coordinates.

2.1.10. Locally Linear Embedding (LLE)

LLE is a nonlinear dimensionality reduction technique that preserves the local structure of the data [60]. It computes weights, W_{ij} , that best reconstruct each data point from its neighbors, and then minimizes the reconstruction error in the lower-dimensional space:

$$\Phi = \sum_i \left| X_i - \sum_j W_{ij} X_j \right|^2 \quad (9)$$

where X_i is the i -th data point and W_{ij} are the reconstruction weights.

2.1.11. Non-Negative Matrix Factorization (NMF)

NMF is a linear method that decomposes the data matrix X into two non-negative matrices W and H , such that

$$X \approx WH \quad (10)$$

where W represents the basis vectors and H the coefficients. This method is useful when the data components, such as spectral indices, are expected to be non-negative.

2.1.12. Truncated Singular Value Decomposition (TSVD)

TSVD is a variant of PCA where only the top k singular values and corresponding vectors are retained [61]. The data matrix X is decomposed as:

$$X = U\Sigma V^T \quad (11)$$

where U and V are orthogonal matrices, and Σ is the diagonal matrix of singular values. Only the top k singular values are used to reconstruct the data.

2.1.13. Neighborhood Components Analysis (NCA)

NCA is a supervised, nonlinear method that learns a linear transformation that maximizes the accuracy of nearest neighbor classification [62]. The transformation matrix A is learned by minimizing the stochastic loss function:

$$L(A) = - \sum_i \log \sum_{j \neq i} \exp(-\|Ax_i - Ax_j\|^2) \quad (12)$$

where x_i and x_j are data points and A is the learned transformation.

2.2. Satellite Data Acquisition and Preprocessing

Landsat-8 is a satellite in the Landsat program that captures Earth observation data, specifically designed for monitoring changes in the environment and land use. Launched in 2013, it carries two key instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The OLI captures images in nine spectral bands, including visible, NIR, and SWIR, allowing for the detailed analysis of vegetation, water bodies, and urban areas. TIRS measures thermal radiation in two bands, which is particularly useful for monitoring surface temperatures, heat patterns, and water stress. Landsat-8's data, with a spatial resolution of 30 m for most bands and 100 m for thermal bands, offers high-quality imagery used in various applications such as disaster monitoring, agriculture, forestry, and climate change research. With a revisit time of 16 days, Landsat-8 provides consistent, repeatable coverage of Earth's surface, making it an essential tool for long-term environmental monitoring.

We used Google Earth Engine to acquire and pre-process Landsat-8 data due to its comprehensive tools and datasets [63]. To mitigate seasonal variations and accurately represent post-wildfire conditions, we selected images from one month before and one month after the field survey, which was conducted one year after the fire. We chose atmospherically corrected surface reflectance products for their precision in measuring ground reflectance, unaffected by atmospheric distortions such as scattering and absorption [64]. We only considered images with a maximum cloud coverage of 25%, resulting in five viable scenes. Finally, we systematically applied the median filter and band selection functions across the entire dataset to generate a composite image incorporating seven high-resolution bands (Ultra blue, Red, Green, Blue, NIR, SWIR1, and SWIR2).

2.3. Data Structuring

We organized our input data into a numerical matrix format for pixel-based classification. In this structure, each column represents a different feature, such as an optical index, while the rows correspond to labeled data.

2.3.1. Optical Index Generation

The selection of optical indices was based on the premise that the distinctive characteristics of burned areas differ from those of unburned sites. We selected nine commonly used indices from BSM research (Table 2). Additionally, we generated synthetic optical indices through custom calculations proposed by Van et al. (2024) [24], aimed to provide a more nuanced and effective means of assessing the varying levels of burn severity across the landscape. These formulations included normalized difference $(a - b)/(a + b)$, ratio-based (a/b) , algebraic sum $(a + b)$, algebraic difference $(a - b)$, and product-based $(a \times b)$ calculations, where a and b represent any two of the seven Landsat-8 bands. As a result, a total of 118 indices, including the 7 original bands, were used in this study.

Table 2. Lists of nine established indices were used in this study. Acronyms: BAI—Burn Area Index; NBR—Normalized Burn Ratio; SAVI—Soil-Adjusted Vegetation Index; EVI—Enhanced Vegetation Index; GEMI—Global Environment Monitoring Index; NIRv—Near-Infrared Reflectance of Vegetation; VARI—Visible Atmospherically Resistant Index; NDVI—Normalized Difference Vegetation Index; CSI—Char Soil Index.

No.	Index	Formula	References
1	BAI	$\frac{1}{(0.1 - \text{Red})^2 + (0.06 - \text{NIR})^2}$	[65]
2	NBR	$\frac{\text{NIR} - \text{SWIR2}}{\text{NIR} + \text{SWIR2}}$	[66]
3	SAVI	$\frac{1.5 \times (\text{NIR} - \text{Red})}{\text{NIR} + \text{Red} + 0.5}$	[67]
4	EVI	$\frac{2.5 \times (\text{NIR} - \text{Red})}{\text{NIR} + 6 \times \text{Red} - 7.5 \times \text{Blue} + 1}$	[68]
5	GEMI	$\text{eta} \times (1 - 0.25 \times \text{eta}) - \left(\frac{\text{Red} - 0.125}{1 - \text{Red}} \right)$, where $\text{eta} = \frac{2 \times (\text{NIR}^2 - \text{Red}^2) + 1.5 \times \text{NIR} + 0.5 \times \text{Red}}{\text{NIR} + \text{Red} + 0.5}$	[69]
6	NIRv	$\text{NIR} \times \text{NDVI}$	[70]
7	NDVI	$\frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}}$	[71]
8	VARI	$\frac{\text{Green} - \text{Red}}{\text{Green} + \text{Red} - \text{Blue}}$	[72]
9	CSI	$\frac{\text{NIR}}{\text{SWIR1}}$	[73]

2.3.2. Data Labeling: Composite Burn Index (CBI)

The CBI method, developed by Key and Benson (2006), is a widely used technique to assess the severity of wildfires across different vegetation strata [66]. The method evaluates fire effects in five distinct layers: substrate (soil and ground cover), herbaceous vegetation, shrubs, intermediate tree layers, and canopy trees. For each layer, a trained observer assigns a burn severity rating based on criteria, such as vegetation mortality, scorch height, and changes in surface fuels. CBI is typically applied within predefined plots, allowing for calculating an overall burn severity score by averaging the ratings across strata. This method provides a standardized and comprehensive approach to quantify fire impacts, aiding in ecological recovery assessments and landscape-level fire management strategies.

In our study, we used the dataset titled ‘Composite Burn Index (CBI) Data for the Conterminous US, Burned Areas Boundaries, Collected Between 1994 and 2018 [74]’ because it provides a comprehensive and standardized assessment of wildfire burn severity across a wide geographical and temporal range. The dataset is published by the U.S. Geological Survey, and labeling involves categorizing various burn severity levels across different regions of the U.S. using satellite imagery and ground measurements [75]. This dataset spans more than two decades and provides valuable information on the extent of burned areas, aiding in analyzing fire impact on ecosystems. The 2014 Carlton Complex fire was selected for this study because it is the largest wildfire in Washington State’s history, burning approximately 103,643 hectares [76]. Its scale presents a unique opportunity to study various burn severities across diverse vegetation types and landscapes, making it an ideal case for advancing BSM techniques. Additionally, the fire contains the highest number of CBI measurements—a total of 328 plots—offering a rich and comprehensive dataset for ML modeling (Table 3) [77].

Table 3. Burn severity category definition used in this study, proposed by Miller and Thode (2009) [78].

Severity Category	CBI Values	Number of Data
No burn	0.00–0.1	110
Low	0.1–1.24	105
Moderate	1.25–2.24	54
High	2.25–3.00	59

2.4. Machine Learning

Six different ML models were selected to ensure the robustness and generalization of data classification. Each model has unique strengths and weaknesses, enhancing our ability to test our methodology comprehensively. Additionally, employing multiple models facilitates the validation of results, with consistent findings across different models improving the reliability of the research. This section briefly outlines ML models chosen for their prominence in the literature and practical BSM applications.

2.4.1. Random Forest (RF)

RF is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes from individual trees [79]. Each tree is trained on a bootstrap sample from the dataset, and at each node, only a random subset of features is considered for splitting, reducing overfitting. Given a set of decision trees T_1, T_2, \dots, T_M , the final predicted class \hat{y} for an input x is determined by majority voting:

$$\hat{y} = \text{mode}(T_1(x), T_2(x), \dots, T_M(x)) \quad (13)$$

The individual decision trees T_i are built by recursively splitting the data based on feature values to maximize a certain criterion, such as Gini impurity or information gain. This process is naturally extended to multi-class classification, where the RF model predicts the class label that receives the most votes from all decision trees, allowing it to handle problems involving multiple classes effectively.

2.4.2. Support Vector Machines (SVM)

SVM is a supervised learning model that aims to find the hyperplane that best separates classes in the feature space [80]. For linearly separable data, the hyperplane is defined by:

$$w^T x + b = 0 \quad (14)$$

where w is the weight vector, x is the input vector, and b is the bias term. SVM seeks to maximize the margin γ , the distance between the hyperplane and the closest data points (support vectors). This is achieved by solving the optimization problem:

$$\min_w \frac{1}{2} \|w\|^2 \text{ subject to } y_i (w^T x_i + b) \geq 1 \quad (15)$$

where y_i represents the class labels and x_i are the feature vectors. For nonlinearly separable data, SVM can apply kernel functions $K(x_i, x_j)$ to transform the input space into a higher-dimensional feature space, allowing for a more flexible decision boundary. In this study, we extended SVM to handle multi-class classification by employing one-vs-rest strategies, which allow SVM to effectively differentiate between multiple classes. These approaches involve decomposing the multi-class problem into several binary classification problems, enabling SVM to be used for distinguishing the different burn severity levels effectively.

2.4.3. K-Nearest Neighbors (KNN)

KNN is a simple, non-parametric classifier that assigns a class to an input x based on the majority class of its k -nearest neighbors in the feature space [81]. The distance metric, typically Euclidean distance, is used to find the k closest data points to x . The predicted class \hat{y} is:

$$\hat{y} = \text{mode}(\{y_i | x_i \in N_k(x)\}) \quad (16)$$

where $N_k(x)$ presents the set of the k -nearest neighbors of x in the training data and y_i is the label of neighbor i . The number of neighbors k is a hyperparameter that controls the classifier's behavior, balancing local sensitivity (low k) and global generalization (high k).

In the case of multi-class classification, KNN can effectively handle multiple classes by determining the majority class among the k -nearest neighbors, making it suitable for

distinguishing different levels of burn severity. The mode function used in KNN naturally extends to handle multiple classes, allowing it to predict the most common class label among its neighbors.

2.4.4. Logistic Regression (LR)

Logistic regression is a linear classifier used for binary or multiclass classification problems [82]. It models the probability of a sample belonging to a particular class using the logistic sigmoid function. For binary classification, the probability $P(y = 1|x)$ is given by:

$$P(y = 1|x) = \frac{1}{1 + e^{-w^T x}} \quad (17)$$

where w is the weight vector, and x is the input feature vector. The decision rule for classification is:

$$\hat{y} = \begin{cases} 1 & \text{if } P(y = 1|x) \geq 0.5 \\ 0 & \text{if } P(y = 1|x) < 0.5 \end{cases} \quad (18)$$

The model parameter w is estimated by maximizing the log-likelihood function:

$$LL(w) = \sum_{i=1}^n [y_i \log P(y_i|x_i) + (1 - y_i) \log(1 - P(y_i|x_i))] \quad (19)$$

For multi-class classification, LR can be extended using the one-vs-res strategy, where multiple binary classifiers are trained to distinguish one class from the rest. Alternatively, the SoftMax regression (or multinomial logistic regression) approach can be used, which allows LR to handle multiple classes simultaneously by modeling the probability distribution over all possible classes. This multi-class adaptation is particularly useful in scenarios like burn severity classification, where multiple burn severity levels must be differentiated.

2.4.5. Multi-Layer Perceptron (MLP)

MLP is an artificial neural network (ANN) consisting of layers of multiple neurons (nodes) [83]. Each neuron in a layer receives input from the previous layer, applies a weighted sum followed by a nonlinear activation function, and passes the result to the next layer. For a single hidden layer MLP, the output of the network is:

$$\hat{y} = f(W_2 f(W_1 x + b_1) + b_2) \quad (20)$$

where W_1 and W_2 are weight matrices, b_1 and b_2 are bias vectors, and f is the activation function, commonly the sigmoid function or rectified linear unit (ReLU). The MLP is trained using backpropagation to minimize a loss function, such as cross-entropy, for classification tasks.

In the case of multi-class classification, MLP uses a SoftMax activation function in the output layer to generate a probability distribution over all possible classes. This allows the network to effectively differentiate between multiple burn severity levels by assigning probabilities to each class, with the final predicted class being the one with the highest probability. The use of cross-entropy loss is ideal for training MLPs in multi-class scenarios, as it penalizes incorrect predictions based on the predicted probability for each class.

2.4.6. Adaptive Boosting (AB)

AB is an ensemble method that combines weak classifiers to create a strong classifier [84]. AdaBoost trains a series of weak learners (e.g., decision stumps) iteratively, giving more weight to misclassified samples at each step. The final prediction is a weighted sum of the weak classifiers' predictions:

$$\hat{y} = \text{sign} \left(\sum_{t=1}^T \alpha_t h_t(x) \right) \quad (21)$$

where $h_t(x)$ is the t -th weak classifier and α_t is its corresponding weight based on its accuracy. The weights α_t are updated at each iteration according to the classification error of h_t . Samples misclassified by h_t receive higher weights in the next iteration, forcing the subsequent classifier to focus on difficult examples.

For multi-class classification, AB can be extended using methods such as SAMME. These approaches adapt the original boosting framework to handle more than two classes by modifying the weighting scheme and decision process to accommodate multiple class labels. Specifically, in the SAMME variant, the weak classifiers output class probabilities rather than binary decisions, and the algorithm then aggregates these to make a final prediction across multiple classes. This adaptation makes AB well-suited for distinguishing between different burn severity levels, as it emphasizes challenging samples and iteratively improves classification accuracy.

2.4.7. Particle Swarm Optimization (PSO)

PSO, a hyper-parameter optimization technique inspired by the social behavior of birds flocking, is used in this study (Table 4). We selected PSO because of its proven effectiveness in navigating large hyperparameter spaces and finding globally optimal solutions with a reasonable computational cost [85]. In PSO, a group of particles (candidate solutions) explore the search space, where each particle adjusts its position based on its own experience (personal best) and the experience of the entire swarm (global best) [86]. The velocity of each particle is updated using a combination of inertia, cognitive influence (how close the particle is to its own best solution), and social influence (how close it is to the swarm's best solution). Mathematically, the velocity update is given by:

$$v_i^{t+1} = wv_i^t + c_1r_1(p_i^t - x_i^t) + c_2r_2(g^t - x_i^t) \quad (22)$$

where w is the inertia, c_1 and c_2 are the cognitive and social coefficients, and r_1 and r_2 are random factors. The particle's position is then updated based on the new velocity. The process repeats iteratively until a stopping criterion is met in order to converge on an optimal solution by exploiting the shared knowledge of the swarm.

In our hyperparameter tuning process, we employed PSO from the Python `pyswarm` library to efficiently determine optimal model settings. Specifically, we used a swarm of 30 particles, which allowed for a balance between diversity in exploring the hyperparameter space and computational efficiency. The number of iterations was set to 100, ensuring the optimization process had enough time to converge on a suitable solution while avoiding unnecessary computation time. For parameter settings, the inertia weight (w) was set to 0.5 to control the influence of a particle's previous velocity on its new velocity, thereby balancing exploration and exploitation in the search space. We set the cognitive parameter (c_1) and the social parameter (c_2) to 1.5 each. This configuration ensures that each particle was influenced by its best historical position and the best position found by the swarm, maintaining a balance between individual search and social learning. Velocity clamping was not employed in our approach since we found that the dimensionality and scaling of the hyperparameters were well-handled without needing to restrict particle velocity. Additionally, we defined specific bounds for each hyperparameter to maintain the optimization within feasible and meaningful limits. These bounds were established based on prior experience and best practices in the field. The stopping criteria for the optimization were either the maximum number of iterations (100) or an improvement in the fitness function that fell below a threshold (epsilon) of 1×10^{-5} for 10 consecutive iterations, which indicated that convergence had been achieved.

Table 4. Optimal values of hyperparameters determined using the PSO method.

Models	Hyper-Parameters	Optimal Values
SVM	C	1.261
	kernel	'poly'
	gamma	'scale'
	degree	2
	coef0	5.846
RF	n_estimators	199
	max_depth	25
	min_samples_split	12
	min_samples_leaf	2
	max_features	'sqrt'
	criterion	'log_loss'
MLP	hidden_layer_sizes	(100, 100)
	activation	'tanh'
	solver	'adam'
	alpha	0.00342
	learning_rate	'constant'
AB	max_depth	8
	min_samples_split	10
	min_samples_leaf	2
	n_estimators	168
	learning_rate	0.623
	algorithm	'SAMME'
LR	C	2.437
	solver	lbfgs
KNN	n_neighbors	12
	weights	'uniform'
	algorithm	'auto'

2.5. Experimental Setups

This research compares 13 FE methods to determine which performs best across multiple classifiers. The primary goal is identifying the methods that help classifiers maintain high BSM performance. Each FE method is evaluated using three components to ensure a fair and comprehensive comparison. This choice is driven by the limitations of certain methods like t-SNE and LDA, which constrain the maximum number of components to the number of classes minus one. Figure 1 presents 3D scatter plots showing the results of FE methods applied to our BSM dataset with four burn severity levels, namely no burn, low, moderate, and high severity, and the goal is to visually assess how well the methods separate the burn severity classes. It is noted that NCA and LDA, which are supervised techniques, exhibit the clearest separation, as they are designed to maximize class distinction, making them particularly effective at clustering the burn severity categories into distinct groups.

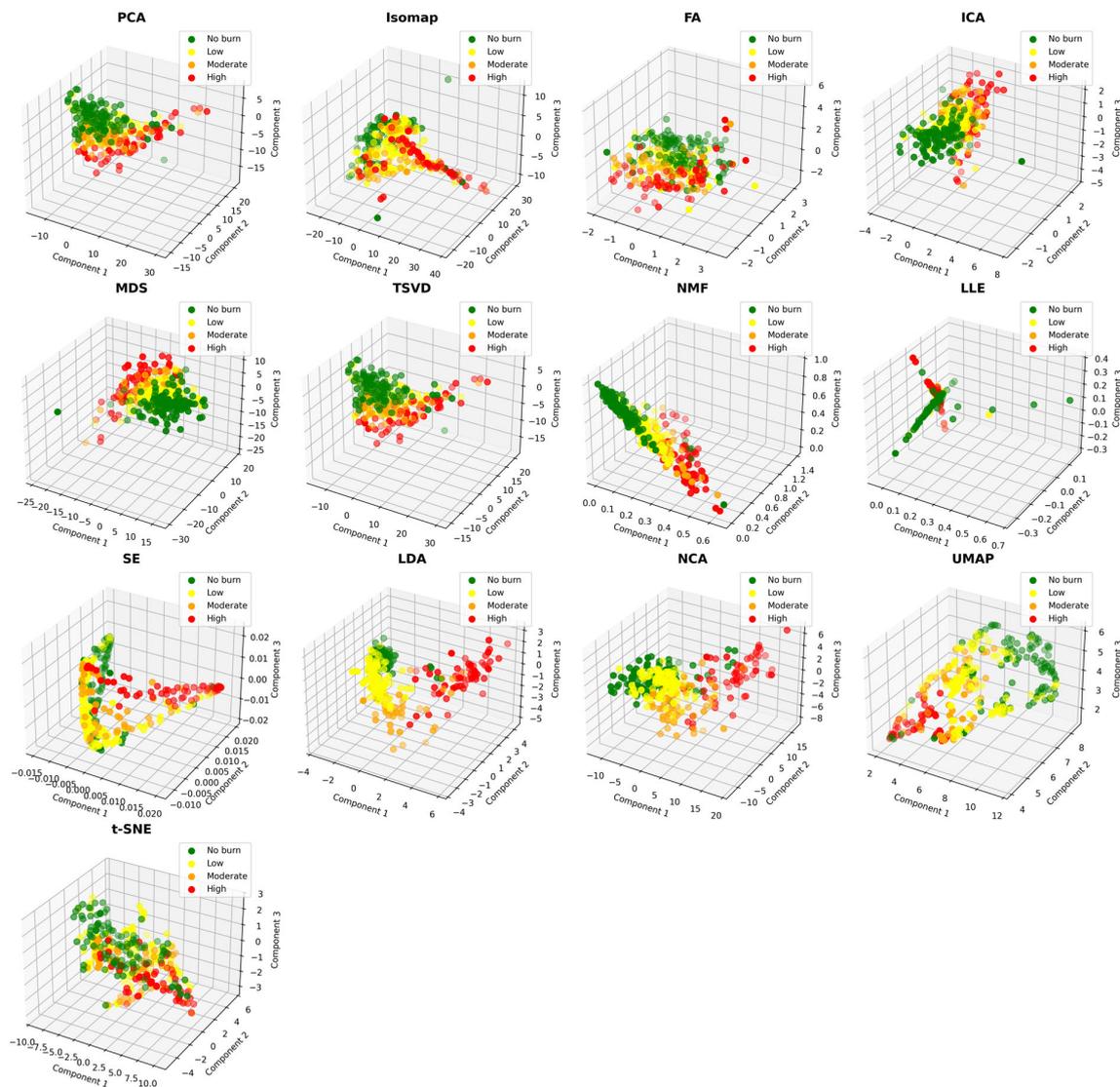


Figure 1. Visualization of dimensionality reduction techniques. Each plot represents a 3D data projection using three main components.

Other methods, such as PCA, Isomap, and FA, show a fair amount of overlap between classes, with the limited separation of the burn severity groups. Similarly, TSVD, LLE, SE, and t-SNE exhibit overlapping clusters, suggesting they are less effective in distinguishing between the burn classes. In contrast, ICA, MDS, NMF, and UMAP show better-structured clustering, with MDS and UMAP providing moderate separation between classes, particularly between the no-burn and other categories.

For the next phase, we conducted a comprehensive benchmarking of thirteen FE methods using six ML models to evaluate their performance in predicting fire severity levels. To ensure that our evaluation is statistically robust and reliable, we incorporated a bootstrapping method into our experimental design. The evaluation process involves the following steps.

- **Step 1: Balanced Sample Selection:** We randomly selected 54 training samples for each fire severity level, as detailed in Table 3. This selection ensures a balanced dataset, which is crucial for the performance of classification algorithms. A balanced sample mitigates the risk of model bias toward any particular class, enhancing the fairness and accuracy of the evaluation [77].

- Step 2: Dataset Splitting. The selected data ($n = 216$) are then randomly split into training and testing sets using a 70/30 ratio. This standard practice in machine learning balances the need for sufficient data to train the model (70%) with enough data to validate its performance (30%). The training set is used to build the ML models, while the testing set provides an unbiased assessment of how well the models generalize to unseen data.
- Step 3: Bootstrapping for Statistical Robustness: To enhance our results' reliability and statistical robustness, steps 1 and 2 are repeated 1000 times, mitigating any anomalies that might influence the outcomes due to variability in data splits.

For model accuracy assessment, overall accuracy (OA), precision, recall, and F1-score are used in this study (Table 5). OA generally measures how well the model correctly classifies instances across all categories. Precision focuses on the model's ability to correctly identify positive instances without misclassifying negative ones, while recall assesses how well the model captures all relevant positive instances in the dataset. The F1-score, a harmonic mean of precision and recall, offers a balanced assessment of the model's performance by considering both false positives and false negatives, making it especially useful in cases of an imbalance between classes. These metrics together provide a robust evaluation of the model's effectiveness in the study.

Table 5. Statistical metrics used in this study. Acronyms: TP—true positive; TN—true negative; FP—false positive; FN—false negative.

Metrics	Formula	Range	Optimal Value
OA	$\frac{TP+TN}{TP+TN+FP+FN}$	0.0–1.0	1.0
Precision	$\frac{TP}{TP+FP}$	0.0–1.0	1.0
Recall	$\frac{TP}{TP+FN}$	0.0–1.0	1.0
F1-score	$2 \times \frac{Precision \times Recall}{Precision+Recall}$	0.0–1.0	1.0

2.6. Case Study

The Carlton Complex fire (Figure 2), which ignited in July 2014 due to lightning strikes amidst severe drought conditions, became the largest and one of the most destructive wildfires in Washington State's history [87]. Located in north-central Washington's Okanogan County, the fire began as four blazes—the Stokes, Gold Hikes, French Creek, and Cougar Flat fires—merged into a massive inferno fueled by high temperatures and strong winds [88]. Over its duration, the fire scorched approximately 103,643 hectares of land, destroyed over 300 homes and numerous other structures, and caused an estimated USD 98 million in damages. Thousands of residents were forced to evacuate as the fire damaged infrastructure, including power lines and communication networks. Despite the efforts of over 3000 firefighters from various agencies, the fire was not fully contained until late August 2014. The Carlton Complex fire left lasting scars on the landscape, leading to environmental issues like soil erosion and habitat loss, and profoundly impacted local communities, underscoring the challenges of wildfire management and the need for improved preparedness in the face of changing climate conditions.

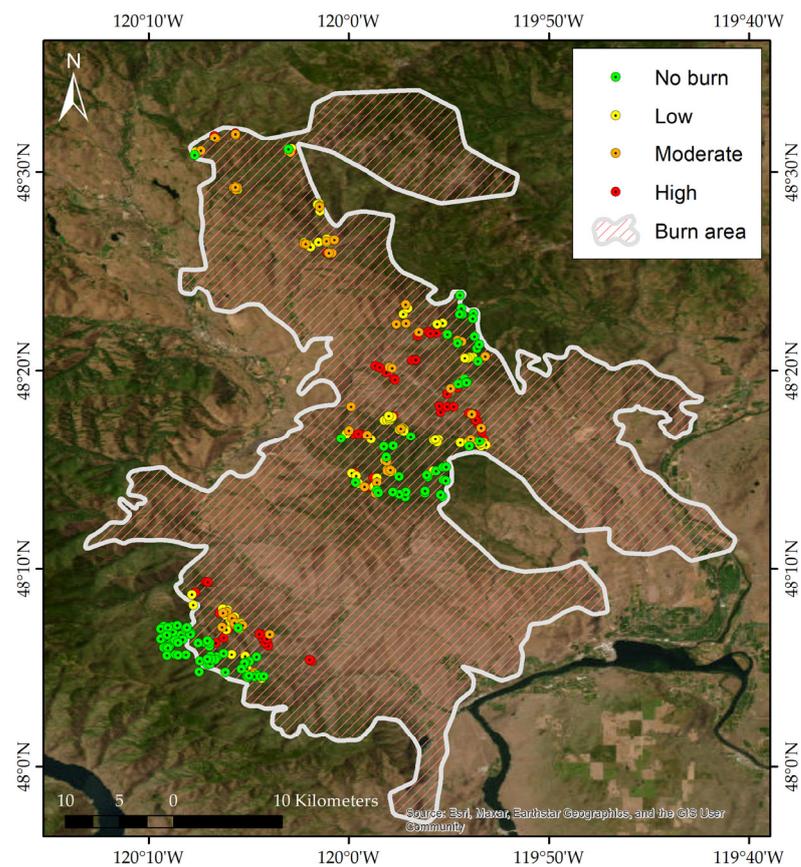


Figure 2. Location of the 2014 Carlton Complex wildfire used in this study.

The aftermath of the Carlton Complex fire has left enduring ecological scars, including soil erosion and habitat loss, highlighting the ongoing challenges of wildfire management and the imperative for enhanced preparedness amidst evolving climate conditions. The selected imagery for this study includes data from one month before and one month between the field survey date in August 2015. By capturing the post-fire effects close to the field survey date, we aimed to minimize the impact of seasonal variations and provide a robust basis for evaluating burn severity levels across the affected area.

3. Results

3.1. Comparative Performance Analysis of Feature Extraction Methods

Figure 3 presents the classification results by different ML models when applied to various RF methods. Overall, LDA consistently emerges as the top-performing RF method across various classifiers and performance metrics. When evaluating the OA, LDA achieves superior results with high values across all classifiers, reaching 0.84 with SVM, RF, and MLP and 0.82 with AB. Notably, even for the weaker classifiers like KNN and LR, LDA still delivers a relatively strong accuracy of 0.85, outperforming most other methods. Regarding precision, LDA also excels, showing a high precision value of 0.84 with SVM and 0.85 with RF, MLP, LR, and KNN, indicating its ability to minimize false positives. LDA's recall performance follows a similar trend, maintaining consistently high recall values, such as 0.84 with SVM and RF and 0.85 with MLP, RF, and KNN, showcasing its ability to capture relevant instances effectively. The F1-score heatmap further cements LDA's dominance, with high F1-scores across classifiers, including 0.84 for KNN and 0.85 for LR, reflecting its balanced strength between precision and recall.

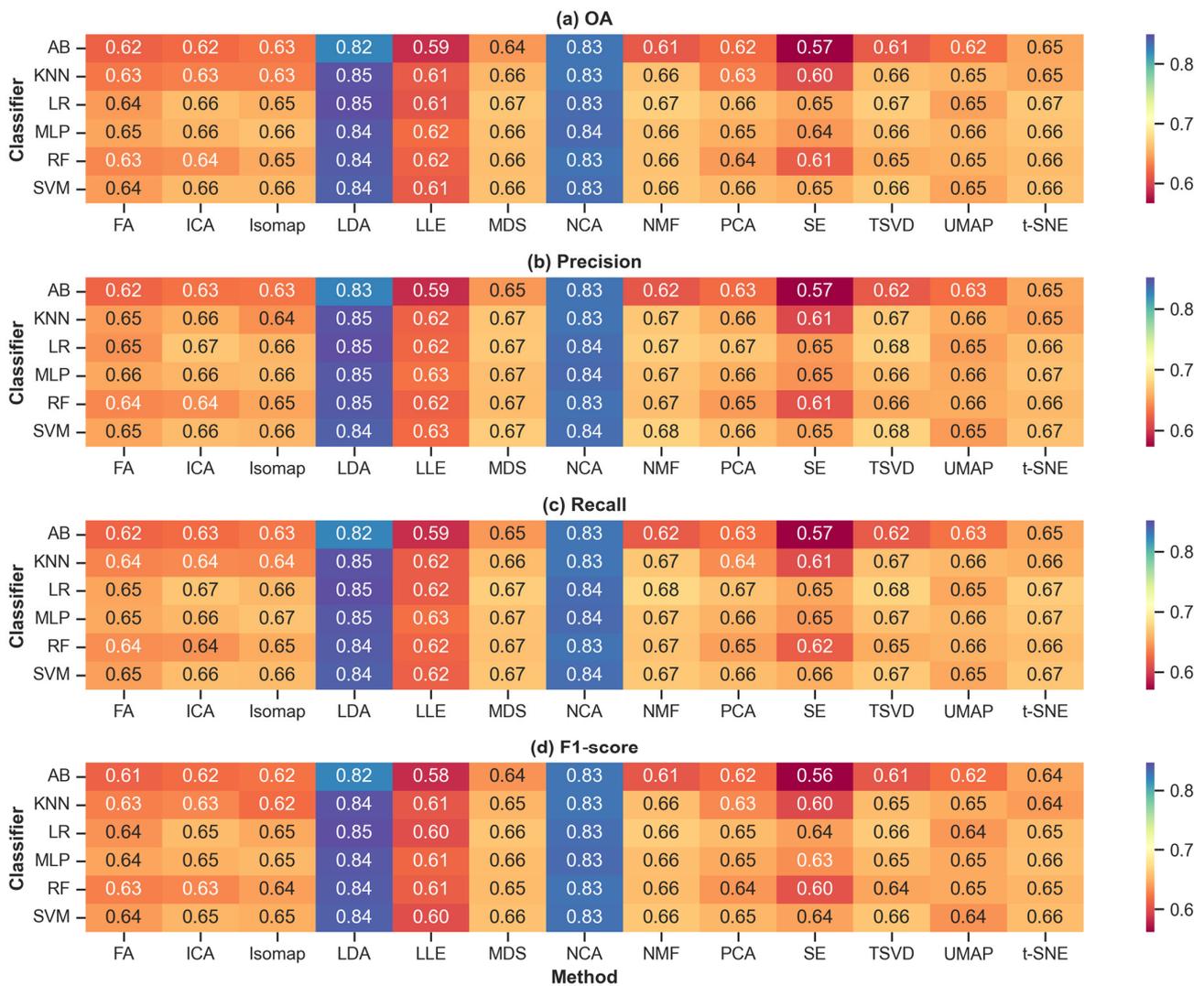


Figure 3. Heatmaps representing the performance of thirteen feature extraction methods on four different metrics, namely (a) overall accuracy (OA), (b) precision, (c) recall, and (d) F1-score, across six machine learning classifiers. The x-axis of each heatmap lists the FR methods, while the y-axis lists the classifiers. The color intensity in each heatmap indicates the mean performance score of 1000 simulations.

While LDA leads overall, NCA closely follows, occasionally surpassing LDA in specific cases, especially with the AB model. For instance, when using AB, NCA achieves a slightly higher F1-score of 0.83 compared to LDA’s 0.82 and achieves the same 0.83 for OA, precision, and recall. This demonstrates NCA’s ability to enhance performance slightly with this classifier, even though it tends to be less effective than LDA in most other scenarios. In other cases, such as with SVM, MLP, and LR, NCA performs closely to LDA but does not surpass it, typically trailing by only a small margin, such as achieving 0.84 in the F1-score for SVM, which is just slightly lower than LDA’s value. These small differences highlight that, while NCA is a robust dimensionality reduction method, particularly for AB, LDA remains the most reliable and high-performing method across different classifiers, making it the optimal choice for dimensionality reduction in this context.

3.2. Impact of Feature Extraction Components on Classifier Performance

This analysis explores the relationship between the number of components used in various FE methods and their impact on OA (Figure 4). To approach this analysis more effectively, we grouped the FE methods based on their behavior as the number of components increases and then analyzed each group. This allowed for a more structured and insightful analysis by focusing on methods that share similar patterns in performance. The methods can be grouped into three main groups based on their behaviors. Understanding these groupings allows for more informed decisions when selecting the number of components for each FE method, maximizing classification performance across models.

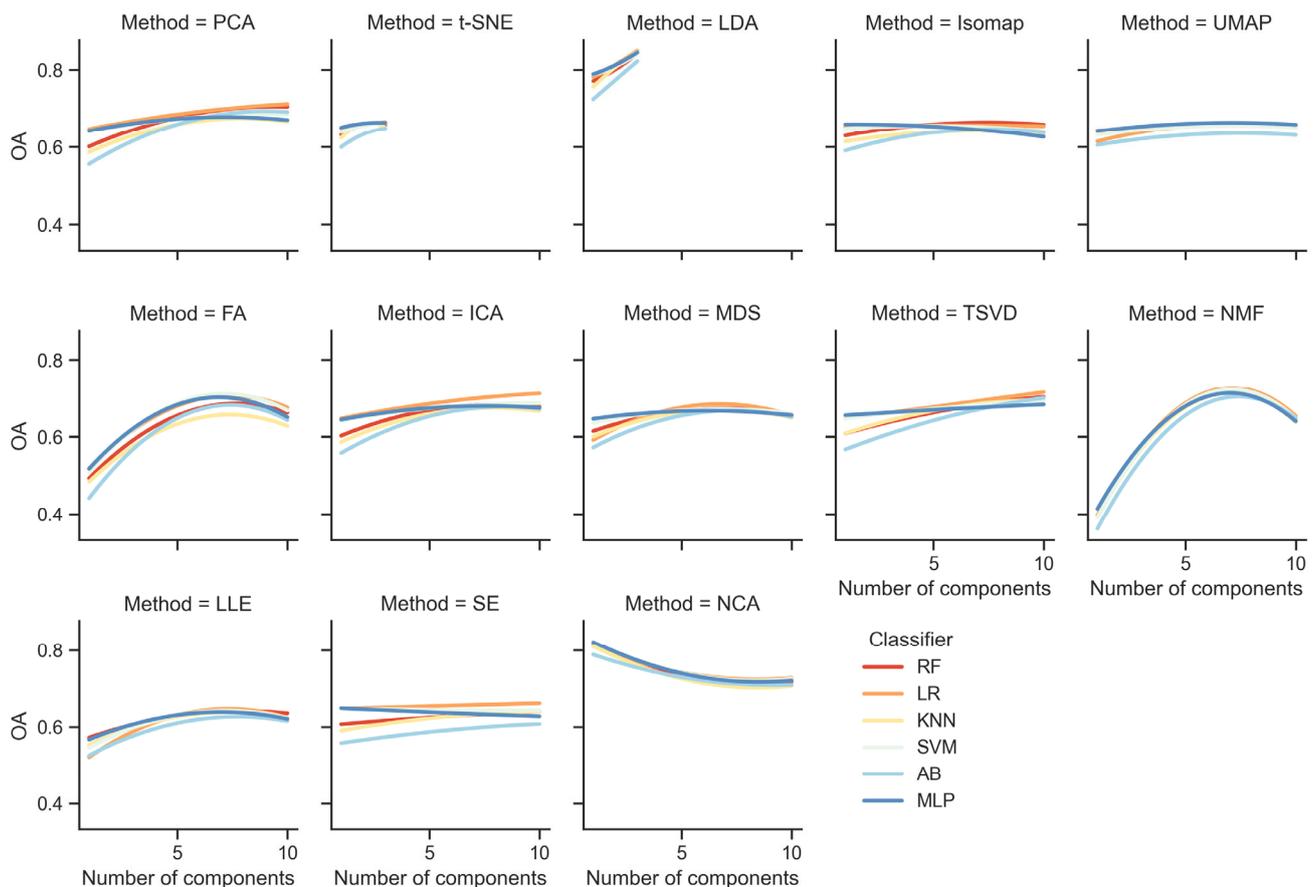


Figure 4. Relationship between the number of components used in thirteen feature reduction methods and the performance (overall accuracy, OA) of six classifiers—RF, LR, KNN, SVM, AB, and MLP. The x-axis in each plot shows the number of components, while the y-axis represents the OA. Each line corresponds to one of the classifiers fitted by quadratic polynomial regression models.

Group 1 consists of methods that exhibit a steady increase in performance as the number of components increases. PCA, t-SNE, LDA, ICA, and TSVD fall into this category. PCA shows a clear improvement as components increase, with the OA rising from approximately 0.6 at one component to nearly 0.7 or more by ten components. This behavior suggests that more components allow PCA to retain more variance and preserve critical information, which improves classification performance. For example, classifiers like RF, LR, and MLP perform consistently better with 10 components, reaching accuracy levels close to 0.7. Similarly, t-SNE and LDA demonstrate a gradual increase in accuracy as the number of components increases. From around 0.75 at one component, LDA reaches approximately 0.85 at three components. ICA and TSVD also follow this trend, with accuracy increasing steadily from 0.6 at one component to over 0.65 at ten components for classifiers like SVM and LR.

Group 2 includes methods that reach peak performance at a certain number of components and then either plateau or show a decline in accuracy with further increases in components. NMF, FA, MDS, and LLE exhibit this behavior. NMF shows a significant increase in accuracy up to around seven components, reaching 0.70 for most classifiers, but beyond this point, accuracy starts to decline, dropping back to around 0.63 at ten components. This indicates that NMF performs best with seven components, and increasing beyond this introduces unnecessary complexity, leading to decreased performance. This pattern highlights the risk of overfitting with too many components. FA also peaks around seven components, where accuracy reaches approximately 0.68 across classifiers and then plateaus or slightly declines with further increases. MDS and LLE behave similarly, with accuracy gradually improving to about eight components, reaching 0.65, but then plateauing or slightly declining as the number of components increases. For methods in this group, selecting an optimal number of components (around seven) is advisable, as adding more components beyond this point does not yield significant performance improvements and may even degrade accuracy.

The final group, Group 3, consists of methods that show little to no sensitivity to the number of components, meaning their performance remains stable, improves minimally, or decreases as components increase. Isomap, UMAP, SE, and NCA fall into this category. Isomap and UMAP show stable performance across most classifiers, with accuracy hovering between 0.60 and 0.62 regardless of the number of components. Similarly, SE shows a small increase in accuracy as the number of components increases, but the improvement is minimal, with accuracy hovering around 0.60 to 0.63. This suggests that those models capture most of the important structures with very few components, and adding more components does not impact classification performance. In contrast to the other methods, NCA shows a slight decline in performance as the number of components increases. Accuracy starts around 0.8 with one component but drops to around 0.75 as the number of components increases. This indicates that NCA performs best with fewer components. In conclusion, for methods in this group, there is little benefit in increasing the number of components beyond three, as performance remains relatively stable.

4. Discussion

4.1. Limitations of PCA

PCA is a well-established FE technique in BSM due to its ability to reduce dimensionality by transforming correlated variables into orthogonal components [20,28–32]. While PCA effectively simplifies data structure and enhances computational efficiency, its ability to capture underlying class separability is limited, particularly in complex, high-dimensional datasets such as those used for BSM. Specifically, PCA's focus on maximizing variance without incorporating class label information often results in suboptimal performance distinguishing different burn severity levels [89]. The transformation performed by PCA does not account for the inherent relationships between classes, which may lead to overlapping distributions that make classification challenging. This study demonstrates that, while PCA can effectively mitigate multicollinearity, its utility is constrained by its inability to fully exploit the discriminative power of the feature space. To quantify these observations, our experiments revealed that the overall accuracy (OA) of models utilizing PCA as an FE method averaged around 64% across different classifiers, with significant variability depending on the complexity of the dataset and the classifier used (Figure 3). These results highlight PCA's limitations in discriminative performance, particularly in comparison to supervised methods that leverage class label information. Another notable limitation of PCA is its linear nature, which restricts its ability to capture nonlinear relationships between features [90]. BSM data often involves complex interactions between spectral indices, which are better represented by nonlinear transformations [91]. Consequently, PCA's linear transformation approach can result in a loss of information when dealing with nonlinear patterns inherent in the dataset. This is particularly problematic in BSM, where accurately distinguishing between different burn severity levels is critical for effective wild-

fire management and ecological restoration. Therefore, reliance on PCA may hinder the effectiveness of ML models in capturing subtle but important differences in burn severity.

4.2. Effectiveness of LDA and NCA

In our experiments, LDA achieved an average OA of 84% across all classifiers, significantly improving PCA and underscoring LDA's effectiveness in preserving class information and improving overall model performance. Unlike PCA, LDA leverages class label information during the dimensionality reduction process, which enables it to prioritize the most informative features for distinguishing between different burn severity levels. This characteristic makes LDA particularly suitable for BSM, where accurate discrimination among varying levels of burn severity is crucial. Moreover, LDA is computationally efficient, making it a practical choice for large datasets commonly encountered in remote sensing applications (Figure 4). These findings align with prior research, emphasizing the effectiveness of LDA in classification tasks [35,92–95]. NCA also demonstrated promise, achieving an average OA of 83% across classifiers. Although NCA did not consistently outperform LDA across all classifiers (except for the AdaBoost model), it exhibited significant advantages in scenarios where local class separation was crucial. Similarly to LDA, NCA leverages class label information, but it does so by optimizing the local distance metrics to improve classification outcomes.

Figure 5 compares the performance of PCA, LDA, and NCA using six classifiers. The F1-scores are used to assess four fire severity levels: no burn, low, moderate, and high severity. LDA and NCA consistently outperform PCA across all burn severity levels. For example, for the no-burn to moderate-severity classes, LDA and NCA achieve F1 scores between 0.75 and 0.9, while PCA lags behind with F1 scores around 0.5 to 0.7. This represents a significant performance gap, with LDA and NCA providing a clear advantage in these more complex scenarios. However, the performance gap between these FE methods narrows in the high-severity class, making PCA more competitive. For this class, LDA and NCA still maintain slightly higher F1-scores, typically between 0.9 and 0.93, but PCA achieves comparable results with F1-scores in the range of 0.80 to 0.87. This smaller gap indicates that while LDA and NCA generally perform better, PCA can hold its own in the high-severity class, where the three methods are much closer in effectiveness.

Classifiers combined with LDA and NCA deliver much more consistent performance than those paired with PCA, where results show significant fluctuations across models. With LDA and NCA, F1-scores remain stable across different classifiers, typically ranging from 0.8 to 0.93, depending on the fire severity level, highlighting the robustness of these FE methods. In contrast, PCA shows much greater variability, especially in low- and moderate-severity classes where F1-scores for some classifiers (i.e., KNN, LR, and SVM) fluctuate. This inconsistency suggests that PCA is more sensitive to the choice of classifier, leading to greater fluctuations in performance.

Our findings further indicate that LDA or NCA combined with LR or MLP achieve the highest F1-scores, often nearing 0.92, making this combination the most effective for the no-burn class. In the low-severity class, where classification is more challenging, LDA-LR or NCA-KNN also performs best, with F1-scores ranging between 0.8 and 0.82. For the moderate-severity class, the optimal combination is LDA with RF, yielding F1-scores of 0.78. While the gap between methods narrowed for the high-severity class, LDA provided the best performance, with F1-scores of approximately 0.93 when paired with KNN.

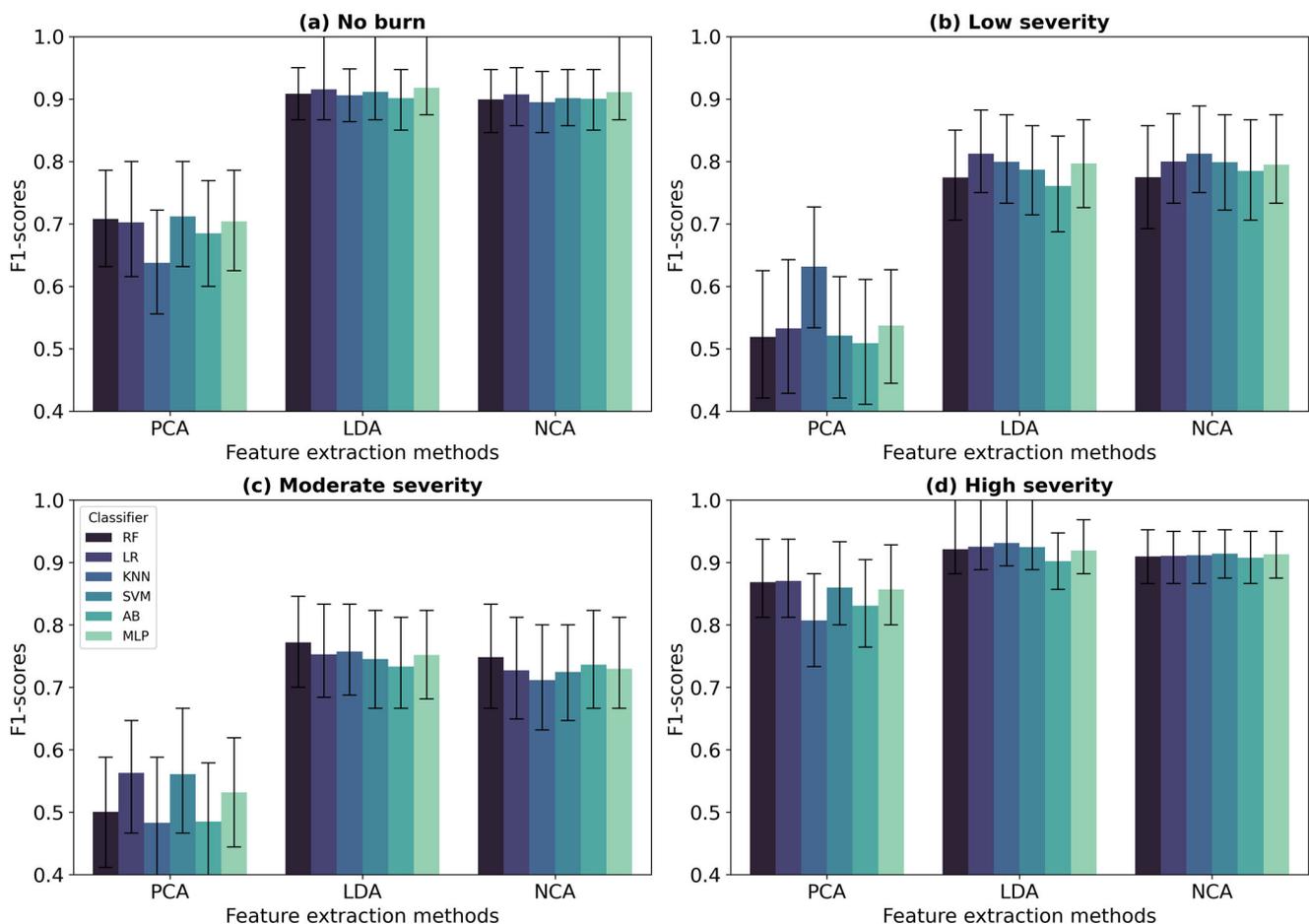


Figure 5. Performance comparison of PCA, LDA, and NCA across four wildfire severity categories: (a) no burn, (b) low, (c) moderate, and (d) high severity. The performance is evaluated using six classifiers, and the y-axis shows the F1-score value. Error bars representing interquartile ranges indicate the variability in model performance across each severity level.

4.3. Research Limitations and Future Directions

Despite the promising results, several limitations should be acknowledged. First, the study focused on a single fire event (the 2014 Carlton Complex fire), limiting the generalizability of the findings to other fire events or regions. Additionally, the vegetation indices generated were based solely on Landsat-8 data, which, while widely used, might not capture the full spectral information necessary for optimal classifier performance. Other data sources like hyperspectral sensors may offer additional insights [96]. Future research could address these limitations by testing the effectiveness of LDA and other FE techniques across a broader range of fire events and regions.

The lack of consideration for the influence of vegetation types on the accuracy of burn severity mapping is also a limitation of this study. Different vegetation types, such as forests, shrublands, and grasslands, can exhibit varying responses to fire, potentially affecting the accuracy of burn severity assessments [97]. While our analysis focused on using the CBI dataset, which provides a comprehensive overview of burn severity across different landscapes, we did not specifically account for vegetation type variability. This could be an important factor influencing burn severity detection and mapping accuracy. As such, future research will aim to address this limitation by incorporating vegetation type data to better understand how different ecosystems respond to fire. This will allow for more precise and vegetation-specific BSM, enhancing the robustness of our assessments.

Another avenue of investigation could involve ensemble methods that combine multiple FE techniques to handle multicollinearity and nonlinear relationships simultaneously.

Finally, expanding this research to include real-time monitoring applications could enhance the utility of BSM for on-the-ground disaster management efforts.

5. Conclusions

This study evaluated 13 feature extraction (FE) techniques to enhance machine learning classifiers for burn severity mapping (BSM) by mitigating multicollinearity among vegetation indices. Using data from the 2014 Carlton Complex fire and 118 vegetation indices from Landsat-8, we found that linear discriminant analysis (LDA) consistently outperformed other FE methods across multiple classifiers, making it a reliable choice for dimensionality reduction in BSM. Notably, neighborhood components analysis (NCA) occasionally surpassed LDA under specific conditions—particularly with the adaptive boosting classifier and when using only one component—highlighting the importance of tailoring FE methods to the classifier and data characteristics. The number of components impacted classifier performance, with different FE methods exhibiting varying behaviors as components increased. These findings underscore the need for careful selection and tuning of FE methods and the number of components to optimize BSM models, ultimately contributing to more accurate and efficient wildfire management strategies.

By improving the handling of complex, high-dimensional datasets, this study provides a framework to enhance decision-making and data analysis in diverse applications beyond wildfire management. The FE and classification techniques used here, such as LDA and NCA, can be applied to other types of natural disaster mapping, such as landslide, flood, and drought assessment, enhancing their precision and efficiency.

Author Contributions: Conceptualization, L.N.V. and G.L.; methodology, L.N.V.; formal analysis, L.N.V. and G.L.; writing—original draft preparation, L.N.V.; visualization, L.N.V.; data curation, L.N.V.; writing—review and editing, G.L.; supervision, G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Disaster-Safety Platform Technology Development Program of the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT (no. 2022M3D7A1090338).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: We would like to thank the four anonymous reviewers for their insightful comments and suggestions, which have greatly improved the quality of this manuscript.

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

References

1. Burnett, J.T.; Edgeley, C.M. Factors Influencing Flood Risk Mitigation after Wildfire: Insights for Individual and Collective Action after the 2010 Schultz Fire. *Int. J. Disaster Risk Reduct.* **2023**, *94*, 103791. [[CrossRef](#)]
2. Souza-Alonso, P.; Saiz, G.; García, R.A.; Pauchard, A.; Ferreira, A.; Merino, A. Post-Fire Ecological Restoration in Latin American Forest Ecosystems: Insights and Lessons from the Last Two Decades. *For. Ecol. Manag.* **2022**, *509*, 120083. [[CrossRef](#)]
3. Pacheco, F.A.L.; Sanches Fernandes, L.F. Hydrology and Stream Water Quality of Fire-Prone Watersheds. *Curr. Opin. Environ. Sci. Health* **2021**, *21*, 100243. [[CrossRef](#)]
4. Chuvieco, E.; Mouillot, F.; van der Werf, G.R.; San Miguel, J.; Tanase, M.; Koutsias, N.; García, M.; Yebra, M.; Padilla, M.; Gitas, I.; et al. Historical Background and Current Developments for Mapping Burned Area from Satellite Earth Observation. *Remote Sens. Environ.* **2019**, *225*, 45–64. [[CrossRef](#)]
5. Giglio, L.; Csiszar, I.; Restás, Á.; Morissette, J.T.; Schroeder, W.; Morton, D.; Justice, C.O. Active Fire Detection and Characterization with the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *Remote Sens. Environ.* **2008**, *112*, 3055–3063. [[CrossRef](#)]
6. Kumar, S.S.; Roy, D.P. Global Operational Land Imager Landsat-8 Reflectance-Based Active Fire Detection Algorithm. *Int. J. Digit. Earth* **2018**, *11*, 154–178. [[CrossRef](#)]
7. Koutsias, N.; Karteris, M. Burned Area Mapping Using Logistic Regression Modeling of a Single Post-Fire Landsat-5 Thematic Mapper Image. *Int. J. Remote Sens.* **2000**, *21*, 673–687. [[CrossRef](#)]
8. van Wagendonk, J.W.; Root, R.R.; Key, C.H. Comparison of AVIRIS and Landsat ETM+ Detection Capabilities for Burn Severity. *Remote Sens. Environ.* **2004**, *92*, 397–408. [[CrossRef](#)]

9. Boucher, J.; Beaudoin, A.; Hébert, C.; Guindon, L.; Bauce, E. Assessing the Potential of the Differenced Normalized Burn Ratio (dNBR) for Estimating Burn Severity in Eastern Canadian Boreal Forests. *Int. J. Wildland Fire* **2017**, *26*, 32–45. [[CrossRef](#)]
10. Zahura, F.T.; Bisht, G.; Li, Z.; McKnight, S.; Chen, X. Impact of Topography and Climate on Post-Fire Vegetation Recovery across Different Burn Severity and Land Cover Types through Random Forest. *Ecol. Inform.* **2024**, *82*, 102757. [[CrossRef](#)]
11. Le, X.-H.; Nguyen Van, L.; Hai Nguyen, D.; Nguyen, G.V.; Jung, S.; Lee, G. Comparison of Bias-Corrected Multisatellite Precipitation Products by Deep Learning Framework. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *116*, 103177. [[CrossRef](#)]
12. Nguyen, G.V.; Le, X.H.; Nguyen Van, L.; Do, T.T.M.; Jung, S.; Lee, G. Machine Learning Approaches for Reconstructing Gridded Precipitation Based on Multiple Source Products. *J. Hydrol. Reg. Stud.* **2023**, *48*, 101475. [[CrossRef](#)]
13. Nguyen, G.V.; Le, X.-H.; Van, L.N.; Jung, S.; Yeon, M.; Lee, G. Application of Random Forest Algorithm for Merging Multiple Satellite Precipitation Products across South Korea. *Remote Sens.* **2021**, *13*, 4033. [[CrossRef](#)]
14. Nguyen, G.V.; Le, X.-H.; Van, L.N.; Jung, S.; Choi, C.; Lee, G. Evaluating the Performance of Light Gradient Boosting Machine in Merging Multiple Satellite Precipitation Products over South Korea. In Proceedings of the 4th International Conference on Sustainability in Civil Engineering, Hanoi, Vietnam, 25–27 November 2022; Springer: Singapore, 2024; pp. 513–522.
15. Le, X.-H.; Van, L.N.; Nguyen, G.V.; Nguyen, D.H.; Jung, S.; Lee, G. Towards an Efficient Streamflow Forecasting Method for Event-Scales in Ca River Basin, Vietnam. *J. Hydrol. Reg. Stud.* **2023**, *46*, 101328. [[CrossRef](#)]
16. Nguyen, G.V.; Nguyen Van, L.; Jung, S.; Hyunuk, A.; Lee, G. Exploring the Power of Physics-Informed Neural Networks for Accurate and Efficient Solutions to 1D Shallow Water Equations. *J. Korean Water Resour. Assoc.* **2023**, *56*, 939–953.
17. Tran, V.; Ivanov, V.; Kim, J. Data Reformation—A Novel Data Processing Technique Enhancing Machine Learning Applicability for Predicting Streamflow Extremes. *Adv. Water Resour.* **2023**, *182*, 104569. [[CrossRef](#)]
18. Collins, L.; Griffioen, P.; Newell, G.; Mellor, A. The Utility of Random Forests for Wildfire Severity Mapping. *Remote Sens. Environ.* **2018**, *216*, 374–384. [[CrossRef](#)]
19. Holden, Z.A.; Morgan, P.; Evans, J.S. A Predictive Model of Burn Severity Based on 20-Year Satellite-Inferred Burn Severity Data in a Large Southwestern US Wilderness Area. *For. Ecol. Manag.* **2009**, *258*, 2399–2406. [[CrossRef](#)]
20. Kulinan, A.S.; Cho, Y.; Park, M.; Park, S. Rapid Wildfire Damage Estimation Using Integrated Object-Based Classification with Auto-Generated Training Samples from Sentinel-2 Imagery on Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *126*, 103628. [[CrossRef](#)]
21. Badola, A.; Panda, S.K.; Roberts, D.A.; Waig, C.F.; Jandt, R.R.; Bhatt, U.S. A Novel Method to Simulate AVIRIS-NG Hyperspectral Image from Sentinel-2 Image for Improved Vegetation/Wildfire Fuel Mapping, Boreal Alaska. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102891. [[CrossRef](#)]
22. Seydi, S.T.; Hasanlou, M.; Chanussot, J. Burnt-Net: Wildfire Burned Area Mapping with Single Post-Fire Sentinel-2 Data and Deep Learning Morphological Neural Network. *Ecol. Indic.* **2022**, *140*, 108999. [[CrossRef](#)]
23. Liu, P.; Liu, Y.; Guo, X.; Zhao, W.; Wu, H.; Xu, W. Burned Area Detection and Mapping Using Time Series Sentinel-2 Multispectral Images. *Remote Sens. Environ.* **2023**, *296*, 113753. [[CrossRef](#)]
24. Van, L.N.; Tran, V.N.; Nguyen, G.V.; Yeon, M.; Do, M.T.-T.; Lee, G. Enhancing Wildfire Mapping Accuracy Using Mono-Temporal Sentinel-2 Data: A Novel Approach through Qualitative and Quantitative Feature Selection with Explainable AI. *Ecol. Inform.* **2024**, *81*, 102601. [[CrossRef](#)]
25. Quintano, C.; Fernández-Manso, A.; Roberts, D.A. Enhanced Burn Severity Estimation Using Fine Resolution ET and MESMA Fraction Images with Machine Learning Algorithm. *Remote Sens. Environ.* **2020**, *244*, 111815. [[CrossRef](#)]
26. Howley, T.; Madden, M.G.; O’Connell, M.-L.; Ryder, A.G. The Effect of Principal Component Analysis on Machine Learning Accuracy with High-Dimensional Spectral Data. *Knowl.-Based Syst.* **2006**, *19*, 363–370. [[CrossRef](#)]
27. Buatoom, U.; Jamil, M.U. Improving Classification Performance with Statistically Weighted Dimensions and Dimensionality Reduction. *Appl. Sci.* **2023**, *13*, 2005. [[CrossRef](#)]
28. Richards, J.A. Thematic Mapping from Multitemporal Image Data Using the Principal Components Transformation. *Remote Sens. Environ.* **1984**, *16*, 35–46. [[CrossRef](#)]
29. Siegert, F.; Ruecker, G. Use of Multitemporal ERS-2 SAR Images for Identification of Burned Scars in South-East Asian Tropical Rainforest. *Int. J. Remote Sens.* **2000**, *21*, 831–837. [[CrossRef](#)]
30. Nielsen, T.T.; Mbow, C.; Kane, R. A Statistical Methodology for Burned Area Estimation Using Multitemporal AVHRR Data. *Int. J. Remote Sens.* **2002**, *23*, 1181–1196. [[CrossRef](#)]
31. Koutsias, N.; Mallinis, G.; Karteris, M. A Forward/Backward Principal Component Analysis of Landsat-7 ETM+ Data to Enhance the Spectral Signal of Burnt Surfaces. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 37–46. [[CrossRef](#)]
32. Chen, G.; Metz, M.R.; Rizzo, D.M.; Dillon, W.W.; Meentemeyer, R.K. Object-Based Assessment of Burn Severity in Diseased Forests Using High-Spatial and High-Spectral Resolution MASTER Airborne Imagery. *ISPRS J. Photogramm. Remote Sens.* **2015**, *102*, 38–47. [[CrossRef](#)]
33. Russo, A.; Borrás, A. Comparison of Dimension Reduction Techniques Applied to the Analysis of Airborne Radionuclide Activity Concentration. *J. Environ. Radioact.* **2022**, *244–245*, 106813. [[CrossRef](#)] [[PubMed](#)]
34. Seok, H.-S. Performance Comparison of Dimensionality Reduction Methods on RNA-Seq Data from the GTEx Project. *Genes Genom* **2020**, *42*, 225–234. [[CrossRef](#)] [[PubMed](#)]
35. Frat, H.; Asker, M.E.; Hanbay, D. Classification of Hyperspectral Remote Sensing Images Using Different Dimension Reduction Methods with 3D/2D CNN. *Remote Sens. Appl. Soc. Environ.* **2022**, *25*, 100694. [[CrossRef](#)]

36. Böttcher, S.; Merz, C.; Lischeid, G.; Dannowski, R. Using Isomap to Differentiate between Anthropogenic and Natural Effects on Groundwater Dynamics in a Complex Geological Setting. *J. Hydrol.* **2014**, *519*, 1634–1641. [[CrossRef](#)]
37. Feilhauer, H.; Faude, U.; Schmidlein, S. Combining Isomap Ordination and Imaging Spectroscopy to Map Continuous Floristic Gradients in a Heterogeneous Landscape. *Remote Sens. Environ.* **2011**, *115*, 2513–2524. [[CrossRef](#)]
38. Harris, A.; Charnock, R.; Lucas, R.M. Hyperspectral Remote Sensing of Peatland Floristic Gradients. *Remote Sens. Environ.* **2015**, *162*, 99–111. [[CrossRef](#)]
39. Wang, Z.; Zhang, G.; Xing, X.; Xu, X.; Sun, T. Comparison of Dimensionality Reduction Techniques for Multi-Variable Spatiotemporal Flow Fields. *Ocean Eng.* **2024**, *291*, 116421. [[CrossRef](#)]
40. Lopez-Collado, J.; Jacinto-Padilla, J.; Rodríguez-Aguilar, O.; Hidalgo-Contreras, J.V. Bioclimatic Similarity between Species Locations and Their Environment Revealed by Dimensionality Reduction Analysis. *Ecol. Inform.* **2024**, *79*, 102444. [[CrossRef](#)]
41. Kaya, Y.; Kuncan, F. A Hybrid Model for Classification of Medical Data Set Based on Factor Analysis and Extreme Learning Machine: FA + ELM. *Biomed. Signal Process. Control* **2022**, *78*, 104023. [[CrossRef](#)]
42. Gündoğdu, Y.; Karabağlı, P.; Alptekin, H.; Şahin, M.; Kılıç, H.Ş. Comparison of Performances of Principal Component Analysis (PCA) and Factor Analysis (FA) Methods on the Identification of Cancerous and Healthy Colon Tissues. *Int. J. Mass Spectrom.* **2019**, *445*, 116204. [[CrossRef](#)]
43. Guo, Y.; Na, J.; Li, B.; Fung, R.-F. Envelope Extraction Based Dimension Reduction for Independent Component Analysis in Fault Diagnosis of Rolling Element Bearing. *J. Sound Vib.* **2014**, *333*, 2983–2994. [[CrossRef](#)]
44. Wang, J.; Zhang, Y.; Cao, H.; Zhu, W. Dimension Reduction Method of Independent Component Analysis for Process Monitoring Based on Minimum Mean Square Error. *J. Process Control* **2012**, *22*, 477–487. [[CrossRef](#)]
45. Ma, Y.; He, K.; Hopcroft, J.; Shi, P. Neighbourhood-Preserving Dimension Reduction via Localised Multidimensional Scaling. *Theor. Comput. Sci.* **2018**, *734*, 58–71. [[CrossRef](#)]
46. Shirangi, M.G.; Emerick, A.A. An Improved TSVD-Based Levenberg–Marquardt Algorithm for History Matching and Comparison with Gauss–Newton. *J. Pet. Sci. Eng.* **2016**, *143*, 258–271. [[CrossRef](#)]
47. Tsuge, S.; Shishibori, M.; Kuroiwa, S.; Kita, K. Dimensionality Reduction Using Non-Negative Matrix Factorization for Information Retrieval. In Proceedings of the 2001 IEEE International Conference on Systems, Man and Cybernetics, Tucson, AZ, USA, 7–10 October 2001; e-Systems and e-Man for Cybernetics in Cyberspace (Cat.No.01CH37236); Volume 2, pp. 960–965.
48. Qi, N.; Zhang, Z.; Xiang, Y.; Harrington, P. de B. Locally Linear Embedding Method for Dimensionality Reduction of Tissue Sections of Endometrial Carcinoma by near Infrared Spectroscopy. *Anal. Chim. Acta* **2012**, *724*, 12–19. [[CrossRef](#)]
49. Manit, J.; Youngkong, P. Neighborhood Components Analysis in sEMG Signal Dimensionality Reduction for Gait Phase Pattern Recognition. In Proceedings of the 7th International Conference on Broadband Communications and Biomedical Applications, Bundoora, VIC, Australia, 21–24 November 2011; pp. 86–90.
50. Siddique, M.A.B.; Sakib, S.; Rahman, M.A. Performance Analysis of Deep Autoencoder and NCA Dimensionality Reduction Techniques with KNN, ENN and SVM Classifiers. In Proceedings of the 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET), Dhaka, Bangladesh, 23–24 December 2019; pp. 1–6.
51. Maćkiewicz, A.; Ratajczak, W. Principal Components Analysis (PCA). *Comput. Geosci.* **1993**, *19*, 303–342. [[CrossRef](#)]
52. Martinez, A.M.; Kak, A.C. PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 228–233. [[CrossRef](#)]
53. Balamurali, M.; Silversides, K.L.; Melkumyan, A. A Comparison of T-SNE, SOM and SPADE for Identifying Material Type Domains in Geological Data. *Comput. Geosci.* **2019**, *125*, 78–89. [[CrossRef](#)]
54. Tenenbaum, J.B.; de Silva, V.; Langford, J.C. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* **2000**, *290*, 2319–2323. [[CrossRef](#)]
55. McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* **2020**, arXiv:1802.03426v3.
56. Comon, P. Independent Component Analysis, A New Concept? *Signal Process.* **1994**, *36*, 287–314. [[CrossRef](#)]
57. Jöreskog, K.G. Factor Analysis as an Errors-in-Variables Model. In *Principals of Modern Psychological Measurement*; Routledge: London, UK, 1983; ISBN 978-0-203-05665-3.
58. Bronstein, A.M.; Bronstein, M.M.; Kimmel, R. Generalized Multidimensional Scaling: A Framework for Isometry-Invariant Partial Surface Matching. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1168–1172. [[CrossRef](#)] [[PubMed](#)]
59. Shi, J.; Malik, J. Normalized Cuts and Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 888–905. [[CrossRef](#)]
60. Belkina, A.C.; Ciccolella, C.O.; Anno, R.; Halpert, R.; Spidlen, J.; Snyder-Cappione, J.E. Automated Optimized Parameters for T-Distributed Stochastic Neighbor Embedding Improve Visualization and Analysis of Large Datasets. *Nat. Commun.* **2019**, *10*, 5415. [[CrossRef](#)]
61. Vu, T.; Chunikhina, E.; Raich, R. Perturbation Expansions and Error Bounds for the Truncated Singular Value Decomposition. *Linear Algebra Its Appl.* **2021**, *627*, 94–139. [[CrossRef](#)]
62. Goldberger, J.; Hinton, G.E.; Roweis, S.; Salakhutdinov, R.R. Neighbourhood Components Analysis. In Proceedings of the Advances in Neural Information Processing Systems; MIT Press: Cambridge, MA, USA, 2004; Volume 17.
63. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [[CrossRef](#)]
64. Vermote, E.; Justice, C.; Claverie, M.; Franch, B. Preliminary Analysis of the Performance of the Landsat 8/OLI Land Surface Reflectance Product. *Remote Sens. Environ.* **2016**, *185*, 46–56. [[CrossRef](#)]

65. Chuvieco, E.; Martín, M.P.; Palacios, A. Assessment of Different Spectral Indices in the Red-near-Infrared Spectral Domain for Burned Land Discrimination. *Int. J. Remote Sens.* **2002**, *23*, 5103–5110. [CrossRef]
66. Key, C.H.; Benson, N.C. Landscape Assessment (LA). 2006. Available online: https://www.fs.usda.gov/rm/pubs_series/rmrs/gtr/rmrs_gtr164/rmrs_gtr164_13_land_assess.pdf (accessed on 18 November 2024).
67. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
68. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
69. Pinty, B.; Verstraete, M.M. GEMI: A Non-Linear Index to Monitor Global Vegetation from Satellites. *Vegetatio* **1992**, *101*, 15–20. [CrossRef]
70. Badgley, G.; Field, C.B.; Berry, J.A. Canopy Near-Infrared Reflectance and Terrestrial Photosynthesis. *Sci. Adv.* **2017**, *3*, e1602244. [CrossRef] [PubMed]
71. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
72. Gitelson, A.A.; Stark, R.; Grits, U.; Rundquist, D.; Kaufman, Y.; Derry, D. Vegetation and Soil Lines in Visible Spectral Space: A Concept and Technique for Remote Estimation of Vegetation Fraction. *Int. J. Remote Sens.* **2002**, *23*, 2537–2562. [CrossRef]
73. Smith, A.M.S.; Wooster, M.J.; Drake, N.A.; Dipotso, F.M.; Falkowski, M.J.; Hudak, A.T. Testing the Potential of Multi-Spectral Remote Sensing for Retrospectively Estimating Fire Severity in African Savannahs. *Remote Sens. Environ.* **2005**, *97*, 92–115. [CrossRef]
74. Composite Burn Index (CBI) Data for the Conterminous US, Burned Areas Boundaries, Collected Between 1994 and 2018 | USGS Science Data Catalog. Available online: <https://data.usgs.gov/datacatalog/data/USGS:62e968e5d34e749ac04cc10a> (accessed on 18 October 2024).
75. Picotte, J.J.; Bhattarai, K.; Howard, D.; Lecker, J.; Epting, J.; Quayle, B.; Benson, N.; Nelson, K. Changes to the Monitoring Trends in Burn Severity Program Mapping Procedures and Data Products. *Fire Ecol.* **2020**, *16*, 16. [CrossRef]
76. Edgeley, C.M.; Paveglio, T.B. Community Recovery and Assistance Following Large Wildfires: The Case of the Carlton Complex Fire. *Int. J. Disaster Risk Reduct.* **2017**, *25*, 137–146. [CrossRef]
77. Collins, L.; McCarthy, G.; Mellor, A.; Newell, G.; Smith, L. Training Data Requirements for Fire Severity Mapping Using Landsat Imagery and Random Forest. *Remote Sens. Environ.* **2020**, *245*, 111839. [CrossRef]
78. Miller, J.D.; Knapp, E.E.; Key, C.H.; Skinner, C.N.; Isbell, C.J.; Creasy, R.M.; Sherlock, J.W. Calibration and Validation of the Relative Differenced Normalized Burn Ratio (RdNBR) to Three Measures of Fire Severity in the Sierra Nevada and Klamath Mountains, California, USA. *Remote Sens. Environ.* **2009**, *113*, 645–656. [CrossRef]
79. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
80. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support Vector Machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]
81. Cover, T.; Hart, P. Nearest Neighbor Pattern Classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [CrossRef]
82. Hosmer, D.W.; Hosmer, T.; Le Cessie, S.; Lemeshow, S. A Comparison of Goodness-of-Fit Tests for the Logistic Regression Model. *Stat. Med.* **1997**, *16*, 965–980. [CrossRef]
83. Kelley, H.J. Gradient Theory of Optimal Flight Paths. *ARS J.* **1960**, *30*, 947–954. [CrossRef]
84. Friedman, J.; Hastie, T.; Tibshirani, R. Additive Logistic Regression: A Statistical View of Boosting (with Discussion and a Rejoinder by the Authors). *Ann. Stat.* **2000**, *28*, 337–407. [CrossRef]
85. Tani, L.; Veelken, C. Comparison of Bayesian and Particle Swarm Algorithms for Hyperparameter Optimisation in Machine Learning Applications in High Energy Physics. *Comput. Phys. Commun.* **2024**, *294*, 108955. [CrossRef]
86. Bonyadi, M.R.; Michalewicz, Z. Particle Swarm Optimization for Single Objective Continuous Space Problems: A Review. *Evol. Comput.* **2017**, *25*, 1–54. [CrossRef]
87. Halofsky, J.E.; Peterson, D.L.; Harvey, B.J. Changing Wildfire, Changing Forests: The Effects of Climate Change on Fire Regimes and Vegetation in the Pacific Northwest, USA. *Fire Ecol.* **2020**, *16*, 4. [CrossRef]
88. Ma, M. Thinning, Prescribed Burns Protected Forests During the Massive Carlton Complex Wildfire. Available online: <https://phys.org/news/2020-02-thinning-forests-massive-carlton-complex.html#:~:text=In%20the%20first%20major%20study%20following%20the%20devastating,and%20prescribed%20burns%20helped%20forests%20survive%20the%20fire> (accessed on 14 September 2024).
89. Chen, S.; Sun, T. Class-Information-Incorporated Principal Component Analysis. *Neurocomputing* **2005**, *69*, 216–223. [CrossRef]
90. Kerschen, G.; Golinval, J.-C. Non-Linear Generalization of Principal Component Analysis: From a Global to a Local Approach. *J. Sound Vib.* **2002**, *254*, 867–876. [CrossRef]
91. Spiegel, C.J.; Mladenov, N.; Wall, C.B.; Hollman, K.; Tran, C.H.; Symons, C.C.; Shurin, J.B. Life after a Fiery Death: Fire and Plant Biomass Loading Affect Dissolved Organic Matter in Experimental Ponds. *Glob. Chang. Biol.* **2024**, *30*, e17061. [CrossRef] [PubMed]
92. Roth, K.L.; Roberts, D.A.; Dennison, P.E.; Alonzo, M.; Peterson, S.H.; Beland, M. Differentiating Plant Species within and across Diverse Ecosystems with Imaging Spectroscopy. *Remote Sens. Environ.* **2015**, *167*, 135–151. [CrossRef]
93. Kluger, D.M.; Wang, S.; Lobell, D.B. Two Shifts for Crop Mapping: Leveraging Aggregate Crop Statistics to Improve Satellite-Based Maps in New Regions. *Remote Sens. Environ.* **2021**, *262*, 112488. [CrossRef]

94. Mudereri, B.T.; Dube, T.; Niassy, S.; Kimathi, E.; Landmann, T.; Khan, Z.; Abdel-Rahman, E.M. Is It Possible to Discern Striga Weed (*Striga Hermonthica*) Infestation Levels in Maize Agro-Ecological Systems Using In-Situ Spectroscopy? *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *85*, 102008. [[CrossRef](#)]
95. Li, N.; Huo, L.; Zhang, X. Classification of Pine Wilt Disease at Different Infection Stages by Diagnostic Hyperspectral Bands. *Ecol. Indic.* **2022**, *142*, 109198. [[CrossRef](#)]
96. Quintano, C.; Calvo, L.; Fernández-Manso, A.; Suárez-Seoane, S.; Fernandes, P.M.; Fernández-Guisuraga, J.M. First Evaluation of Fire Severity Retrieval from PRISMA Hyperspectral Data. *Remote Sens. Environ.* **2023**, *295*, 113670. [[CrossRef](#)]
97. Stow, D.; Petersen, A.; Rogan, J.; Franklin, J. Mapping Burn Severity of Mediterranean-Type Vegetation Using Satellite Multispectral Data. *GIScience Remote Sens.* **2007**, *44*, 1–23. [[CrossRef](#)]

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