Journal of Science Research and Reviews

PRINT ISSN: 1595-9074
E-ISSN: 1595-8329
DOI: https://doi.org/10.70882/josrar.2026.v3i4.242
Homepage: https://josrar.esrgngr.org

Original Research Article

Machine Learning-Based Land Cover Classification and Environmental Change Detection for Dam Monitoring at Suleja Dam Using Remote Sensing Data

Lucas Olu Atoki

Department of Surveying and Geoinformatics, Bowen University, Iwo, Osun State, Nigeria.
*Corresponding Author's email: geocas.spat@gmail.com
ORCID iD: https://orcid.org/0009-0000-8013-6845

KEYWORDS
Dam monitoring,
Machine learning,
Remote sensing,
Random Forest,
NDVI,
NDWI.

ABSTRACT
Effective dam monitoring is essential for sustainable water resource management and environmental protection. This study integrates machine learning, remote sensing, and Geographic Information Systems (GIS) to assess environmental changes within the Suleja Dam catchment, Nigeria. Multi-temporal Landsat 8 and Sentinel-2 imagery acquired between May and October of 2024 and 2025, together with climatic variables including precipitation, surface temperature, soil moisture, evaporation, and terrain data, were analyzed to evaluate land-cover dynamics. The Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) were used to monitor vegetation and surface water changes, while Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (ANN) classifiers were employed for land-cover classification. Classification performance was evaluated using confusion matrices and standard accuracy metrics. RF achieved the highest overall accuracy (92.83%; Kappa = 0.904), outperforming SVM (88.67%; Kappa = 0.849) and ANN (85.00%; Kappa = 0.800). Environmental change detection revealed a 6.5% decline in vegetation cover and increases in built-up areas (5.8%), water extent (3.2%), and bare surfaces (2.1%). The results demonstrate that integrating machine learning with remote sensing provides a robust and cost-effective framework for continuous dam monitoring, environmental change detection, and sustainable watershed management.

CITATION
Atoki, L. O. (2026). Machine Learning-Based Land Cover Classification and Environmental Change Detection for Dam Monitoring at Suleja Dam Using Remote Sensing Data. *Journal of Science Research and Reviews*, *3*(4), 177-194. https://doi.org/10.70882/josrar.2026.v3i4.242

INTRODUCTION

Dams are critical hydraulic infrastructures that support water supply, irrigation, hydropower generation, flood control, and environmental management. Their structural integrity and operational efficiency are essential for sustainable socio-economic development and public safety. However, dam failures or structural instabilities can lead to severe environmental degradation, economic losses, and loss of life. Consequently, continuous dam monitoring is essential for ensuring operational safety, supporting effective water resource management, and enabling the early detection of structural and environmental anomalies (ICOLD, 2021; Yan et al., 2025). Traditional monitoring approaches rely primarily on field inspections, geotechnical instrumentation, and in situ measurements. Although these methods provide reliable observations, they are often constrained by high operational costs, limited spatial coverage, labor-intensive procedures, and discontinuous temporal observations. These limitations are particularly evident in developing countries such as Nigeria, where inadequate monitoring infrastructure and limited maintenance resources reduce the effectiveness of conventional monitoring systems (Chang et al., 2025; Aleksandra & Jan, 2025).

Recent advances in remote sensing have significantly enhanced the capability for continuous, large-scale monitoring of dam environments. Satellite-based Earth observation systems provide multi-temporal, multispectral, and high-resolution datasets that facilitate the assessment of environmental and hydrological changes surrounding dams. Remote sensing offers several advantages, including extensive spatial coverage, frequent revisit periods, cost-effectiveness, and access to long-term historical records. Satellite missions such as Landsat and Sentinel have therefore become indispensable for monitoring reservoir dynamics, land-cover changes, vegetation health, sedimentation, and hydrological processes within dam catchments (Li et al., 2021; European Space Agency, 2023).

Machine learning techniques have further improved the extraction of meaningful information from remotely sensed datasets by enabling efficient classification, pattern recognition, anomaly detection, and environmental change analysis. Among the most widely adopted supervised learning algorithms are Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Networks (ANN), all of which have demonstrated robust performance in land-cover classification and environmental monitoring applications (Breiman, 2001; Belgiu & Drăguţ, 2016). The integration of machine learning with remote sensing has enhanced automated interpretation of satellite imagery and supported more efficient environmental assessment and decision-making. Recent studies have also demonstrated the potential of artificial intelligence for dam safety assessment, deformation monitoring, settlement analysis, and environmental surveillance through the extraction of spatial and temporal patterns from satellite observations (Chen et al., 2022; Zhang et al., 2024).

Suleja Dam, located in Niger State, Nigeria, is an important water resource infrastructure that supports domestic water supply, agriculture, and environmental sustainability. Increasing urban expansion, anthropogenic activities, and climate variability within the catchment have intensified pressures on the surrounding environment, highlighting the need for effective monitoring strategies capable of providing timely and reliable information for sustainable management. While remote sensing and machine learning techniques have been widely applied in environmental monitoring, their integrated application for land-cover classification and environmental change detection in the Suleja Dam catchment remains limited. This represents an important knowledge gap, particularly for supporting sustainable dam management in Nigeria.

Accordingly, this study applies established machine learning algorithms integrated with multi-temporal remote sensing data to classify land cover and detect environmental changes within the Suleja Dam catchment. Specifically, the study (i) acquires and preprocesses multi-temporal Landsat 8 and Sentinel-2 imagery, (ii) extracts spectral indices and environmental variables relevant to dam monitoring, (iii) evaluates the performance of Random Forest, Support Vector Machine, and Artificial Neural Network algorithms for land-cover classification, and (iv) analyses environmental changes occurring around the reservoir using multi-temporal satellite observations. Rather than proposing new machine learning algorithms, this study demonstrates the applicability of established techniques for environmental monitoring within a Nigerian dam catchment where such integrated assessments remain limited. The findings provide practical insights for geospatial artificial intelligence applications in dam monitoring and offer a cost-effective framework for supporting sustainable water resource management and environmental decision-making in Nigeria and other developing regions.

MATERIALS AND METHODS

Study Area

The study was conducted at Suleja Dam, located in Suleja Local Government Area of Niger State, north-central Nigeria (Figure 1). The dam serves as a major source of domestic water supply, irrigation, and environmental management within the region. Geographically, the study area lies within the Guinea Savannah ecological zone and experiences a tropical climate characterized by distinct wet (April–October) and dry (November–March) seasons. The catchment is influenced by seasonal rainfall variability, increasing urban expansion, agricultural activities, and vegetation changes, making it suitable for evaluating machine learning techniques for land-cover classification and environmental change detection.

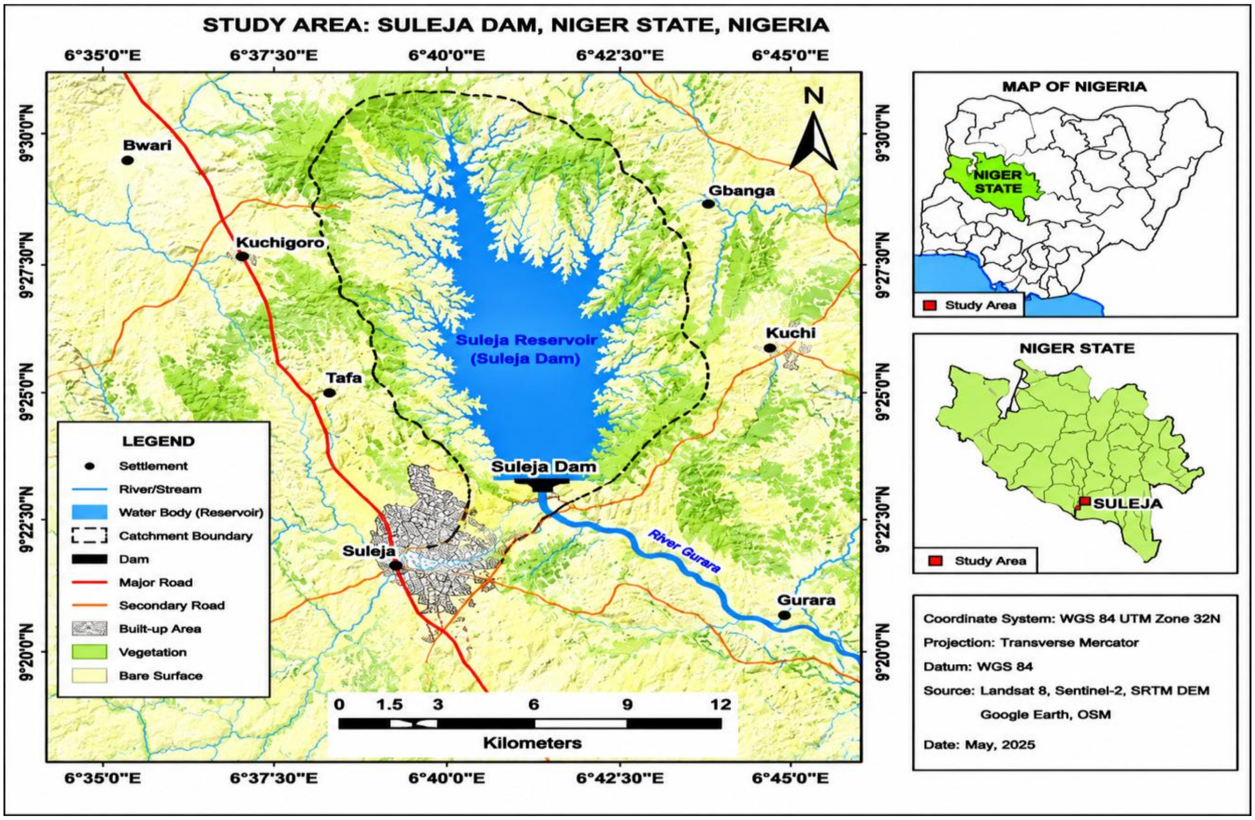

Figure 1: Study Area Map of Suleja Dam, Niger State, Nigeria

Data Acquisition

Multi-temporal remote sensing datasets acquired between May and October for the years 2024 and 2025 were used for this study. Optical satellite imagery was obtained from the Landsat 8 Operational Land Imagery (OLI) and Sentinel-2 Multispectral Instrument (MSI) platforms due to their adequate spatial, spectral, and temporal resolutions for environmental monitoring applications (Giulio et al, 2022) Ancillary environmental datasets, including Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM), precipitation, temperature, soil moisture, wind components, evaporation, and surface pressure data, were also incorporated to support environmental analysis and change detection. Climatic variables were obtained from reanalysis and hydrological datasets available through the Google Earth Engine (GEE) platform. The datasets utilized in this study are summarized in Table 1. The complete workflow of the study is shown in Figure 2.

Table 1: Datasets Used for the Study

Dataset	Source	Spatial Resolution	Purpose
Landsat 8 OLI	USGS	30 m	Land cover classification and environmental monitoring
Sentinel-2 MSI	ESA	10 m	Vegetation and water body analysis
SRTM DEM	NASA	30 m	Terrain and elevation analysis
ERA5 Climatic Data	ECMWF	~0.25°	Temperature, precipitation, pressure, wind analysis
Hydrological Data	GEE Repository	Variable	Environmental assessment

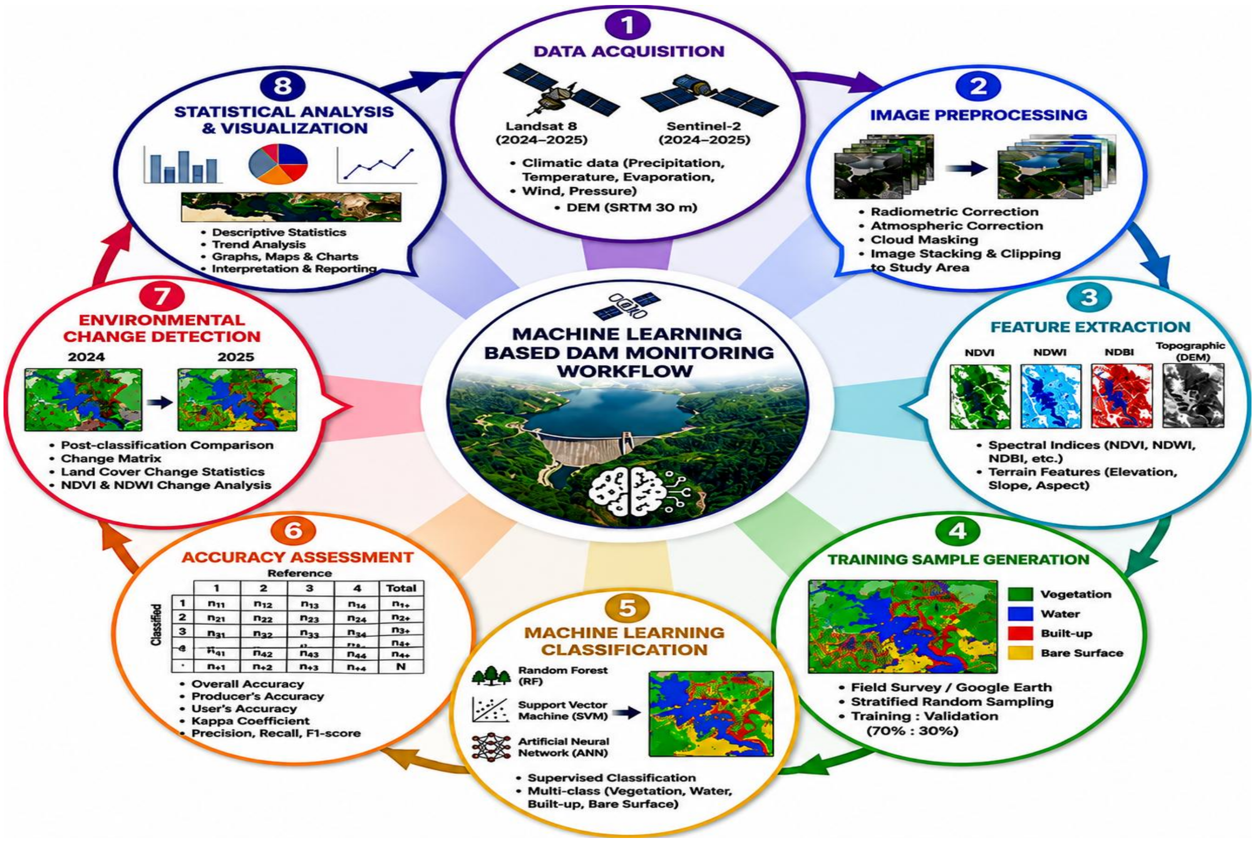

Figure 2: Workflow of the study

Image pre-processing

All satellite imagery underwent standard preprocessing before analysis. Landsat 8 Level-2 Surface Reflectance products and Sentinel-2 Level-2A products were used to minimize atmospheric effects. Cloud-contaminated pixels were removed using the Quality Assessment (QA) bands for Landsat and the Scene Classification Layer (SCL) for Sentinel-2. Additional preprocessing included atmospheric correction; geometric correction; cloud and cloud-shadow masking; image mosaicking; clipping to the study area boundary; and reprojection to UTM Zone 32N, WGS 84 datum. All preprocessing operations were implemented within the Google Earth Engine environment to ensure consistency between datasets.

Feature Extraction

Spectral indices and environmental variables were extracted from the processed imagery for subsequent machine learning classification.

The Normalized Difference Vegetation Index (NDVI) was computed using Equation (1):

$$\boxed{\mathrm{NDVI} = \frac{\mathrm{NIR}-\mathrm{Red}}{\mathrm{NIR}+\mathrm{Red}}} \quad (1)$$

where:

NIR = Near-Infrared reflectance

Red = Red band reflectance

The Normalized Difference Water Index (NDWI) was calculated as:

$$\boxed{\mathrm{NDWI} = \frac{\mathrm{Green}-\mathrm{NIR}}{\mathrm{Green}+\mathrm{NIR}}} \quad (2)$$

where:

Green = Green band reflectance

NIR = Near-Infrared reflectance

The feature stack used for machine learning consisted of: Landsat spectral bands; Sentinel-2 spectral bands; NDVI; NDWI; Digital Elevation Model; surface temperature; precipitation; soil moisture; evaporation; zonal wind component (U); meridional wind component (V); and surface pressure (McFeeters, 1996; Chen et al, 2025).

Training Sample Collection

Four land-cover classes were identified: Water bodies; Vegetation; Built-up areas; and Bare surfaces. Reference samples were generated through visual interpretation of high-resolution satellite imagery and existing land-cover information. A stratified random sampling approach was adopted to ensure that each land-cover class was adequately represented. A total of 2,000 reference samples were collected, comprising 500 samples per class. The samples were divided into training and validation datasets using a 70:30 ratio, resulting in:

Training samples = 1,400

Validation samples = 600

The same sampling protocol was applied consistently across all image dates to improve temporal comparability and reproducibility (Adeboboye, Atoki & Agada, 2024).

Machine Learning Classification

Three supervised machine learning algorithms were evaluated: RF, SVM, and ANN.

Random Forest

RF was implemented using the Scikit-learn library. The optimized parameters were:
Number of trees = 300; Criterion = Gini impurity; Maximum depth = None; Minimum samples split = 2; Minimum samples leaf = 1; Bootstrap sampling = True; Random seed = 42.
Random Forest was selected because of its robustness to noisy data, resistance to overfitting, and ability to model nonlinear relationships (Breiman, 2001).

Support Vector Machine

SVM classification employed a Radial Basis Function (RBF) kernel. The parameters used were: Kernel = RBF; Penalty parameter (C) = 10; and Gamma = 0.1.
The model was trained using standardized predictor variables.

Artificial Neural Network Loss Function

ANN classifier was optimized using the categorical cross-entropy loss function, which measures the discrepancy between the predicted class probabilities and the true class labels. The loss function is expressed using Equation (3)

$$L = -\sum_{i=1}^{n} y_i \log(\hat{y}_i) \quad (3)$$

Where L= Cross-entropy loss; y_i= Actual class label (one-hot encoded); $\hat{y}_i$= Predicted probability for the correct class generated by the Softmax activation function; and n= Number of training samples.
The objective of the optimization process is to minimize the cross-entropy loss through iterative adjustment of the network weights using the Adam optimization algorithm. Lower loss values indicate improved agreement between the predicted class probabilities and the reference class labels, thereby enhancing the classification performance of the ANN model. The ANN training parameters adopted in this study were: Optimizer = Adam; Learning rate = 0.001; Batch size = 32; Epochs = 200; and Early stopping based on validation loss.

Accuracy Assessment

The classification outputs generated from the machine learning algorithms were evaluated using confusion matrix analysis. The performance metrics computed included Producer’s Accuracy (PA), User’s Accuracy (UA), Cohen's Kappa Coefficient, Precision, Recall, and F1-score.
Overall Accuracy was computed as Equation (4):

$$\boxed{\mathrm{OA} = \frac{\sum_{i=1}^{k} x_{ii}}{N} \times 100} \quad (4)$$

where:
x_{ii}= Number of correctly classified samples in class i
k= Total number of land-cover classes
N= Total number of validation samples

Producer's Accuracy (PA)

Producer's Accuracy measures the probability that a reference sample is correctly classified and is given using Equation (5)

$$\boxed{\mathrm{PA}_i = \frac{x_{ii}}{\sum_{j=1}^{k} x_{ij}} \times 100} \quad (5)$$

Where x_{ii}= Correctly classified samples of class i; $\sum x_{ij}$ = Total reference samples belonging to class i.

User's Accuracy (UA)

User's Accuracy measures the reliability of the classified map and is computed using Equation (6)

$$\boxed{\mathrm{UA}_i = \frac{x_{ii}}{\sum_{j=1}^{k} x_{ji}} \times 100} \quad (6)$$

Where $\sum x_{ij}$= Total samples classified as class *i*.

Precision

Precision was computed using Equation (7)

$$\boxed{\text{Precision} = \frac{TP}{TP+FP}} \quad (7)$$

where TP = True Positives; FP = False Positives.

Recall

Recall was computed as Equation (8)

$$\boxed{\text{Recall} = \frac{TP}{TP+FN}} \quad (8)$$

where FN = False Negatives.

F1-Score

The F1-score, which combines Precision and Recall, was calculated as Equation (9)

$$\boxed{\mathrm{F1} = \frac{2\times(\text{Precision}\times\text{Recall})}{\text{Precision}+\text{Recall}}} \quad (9)$$

Cohen's Kappa Coefficient

The Kappa coefficient was computed using Equation (10)

$$\boxed{\kappa = \frac{P_o - P_e}{1 - P_e}} \quad (10)$$

Where P_o= Observed agreement; P_e= Expected agreement due to chance.

Environmental Change Detection

Environmental changes between 2024 and 2025 were evaluated using post-classification comparison. The classified land-cover maps were compared pixel-by-pixel to quantify changes in: vegetation cover; water extent; built-up areas; bare surfaces. Temporal variations in NDVI and NDWI were further analyzed to assess vegetation health and reservoir dynamics. The percentage change between two observation periods was calculated using Equation (11)

$$\boxed{\% \text{ Change} = \frac{A_{2025}-A_{2024}}{A_{2024}} \times 100} \quad (11)$$

where A_{2024}= Area of a land-cover class in 2024; A_{2025}= Area of the same land-cover class in 2025.

Statistical Analysis and Visualization

Descriptive statistics, temporal trend analysis, and graphical visualization were conducted using Python (NumPy, Pandas, Matplotlib, and Scikit-learn). Spatial analyses and thematic maps were produced in ArcGIS Pro and QGIS, while Google Earth Engine was used for image preprocessing and feature extraction.

RESULTS AND DISCUSSION

Descriptive Analysis of Environmental Variables

The extracted environmental variables demonstrated clear seasonal variability within the Suleja Dam catchment during the May–October monitoring period. Mean surface temperature was 297.73 K, while precipitation averaged 0.00299 m and exhibited the greatest temporal variability (SD = 0.005 m). Soil moisture remained relatively stable (mean = 0.172 m^3/m^3), whereas evaporation averaged −0.00280 m, indicating continuous moisture loss from the reservoir surface. Surface pressure varied only slightly throughout the observation period (Table 2), suggesting relatively stable atmospheric conditions. The observed variability indicates that precipitation and evaporation were the dominant climatic drivers influencing reservoir hydrology and vegetation dynamics during the study period. These variables subsequently contributed to changes detected in NDVI, NDWI, and land-cover distributions.

Table 2: Descriptive Statistics of Environmental Variables

Variable	Minimum	Maximum	Mean	Standard Deviation
Dew Point Temperature (K)	294.47	296.64	295.45	0.67
Surface Temperature (K)	297.36	298.05	297.73	0.21
Soil Moisture Layer 1 (m^3/m^3)	0.171	0.173	0.172	0.001
Total Precipitation (m)	0.000039	0.020521	0.002985	0.005
Evaporation (m)	-0.003612	-0.002043	-0.002800	0.0004
Surface Pressure (Pa)	98148.56	98565.06	98331.61	124.31

Temporal Variation of Surface Temperature

Surface temperature increased gradually from 297.36 K in May to 298.05 K in October (Figure 3), representing an overall increase of approximately 0.69 K. The increasing temperature coincided with reduced cloud cover and increasing solar radiation during the transition toward the late rainy season. Higher land surface temperatures promote evaporation and may reduce soil moisture availability around exposed reservoir margins, thereby affecting vegetation vigor. The relatively small thermal variation also indicates that the reservoir moderates surrounding microclimatic conditions (Table 3).

Table 3: Surface Temperature Distribution

Date	Surface Temperature (K)
May	297.36
June	297.52
July	297.68
August	297.81
September	297.95
October	298.05

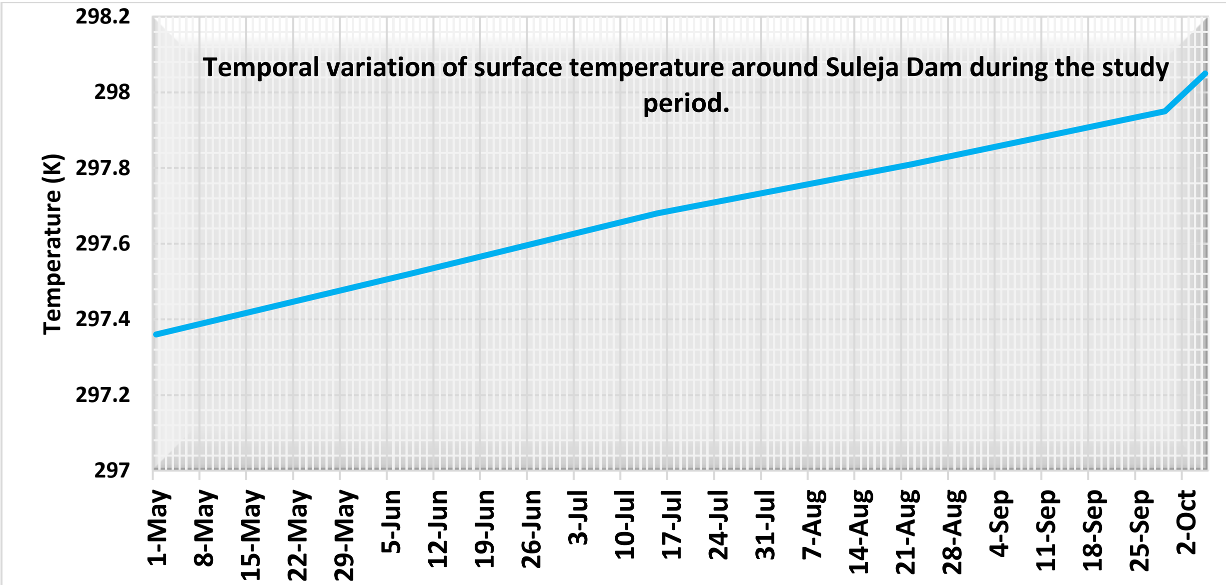

Figure 3: Temporal variation of surface temperature around Suleja Dam

Precipitation Dynamics and Hydrological Response

Weekly precipitation increased steadily throughout the observation period, rising from 0.000039 m during Week 1 to 0.020521 m during Week 6. The increasing rainfall resulted in improved vegetation condition, increased soil moisture, moderate expansion of reservoir water extent, and higher NDWI values during the wettest observation periods. Conversely, evaporation remained consistently negative throughout the monitoring period, indicating continuous water loss from the reservoir. The balance between rainfall recharge and evaporation controlled seasonal fluctuations in reservoir storage (Figure 4).

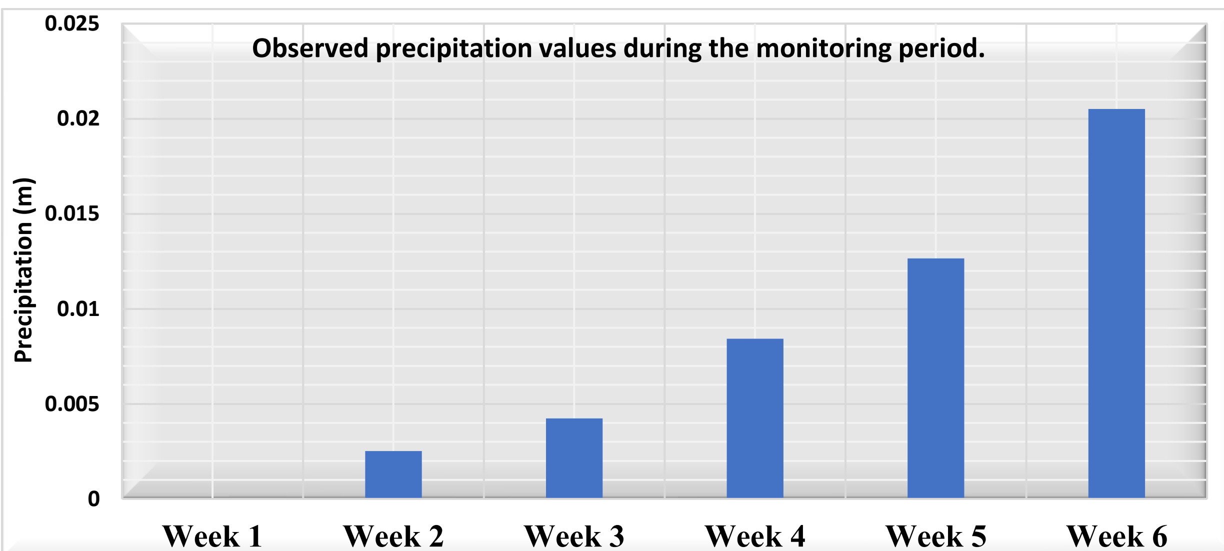

Figure 4: Distribution of total precipitation within the Suleja Dam environment

Atmospheric Conditions

Surface thermal radiation increased steadily from 35.20 MJ m−2 to 37.47 MJ m−2 during the monitoring period. Similarly, the zonal (U) and meridional (V) wind components showed moderate dispersion around zero, indicating changing wind directions rather than persistent airflow. The combined influence of radiation, temperature, wind, and evaporation contributes to moisture redistribution within the catchment and influences vegetation growth patterns (Figures 5 and 6).

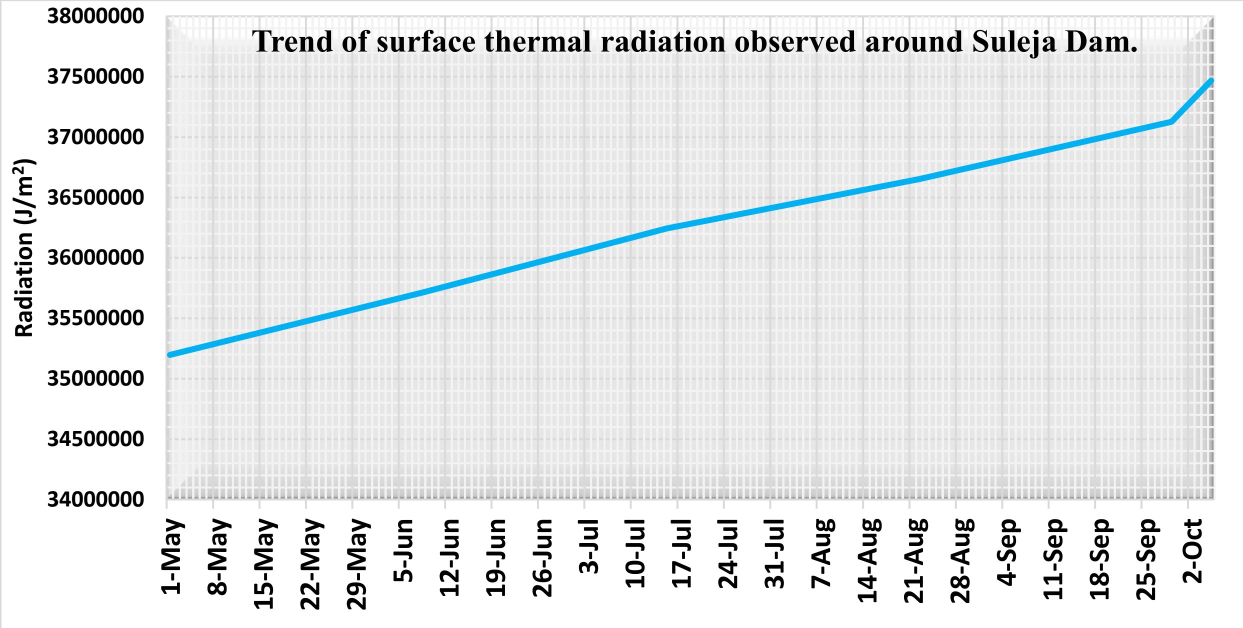

Figure 5: Variation of surface thermal radiation within the study area

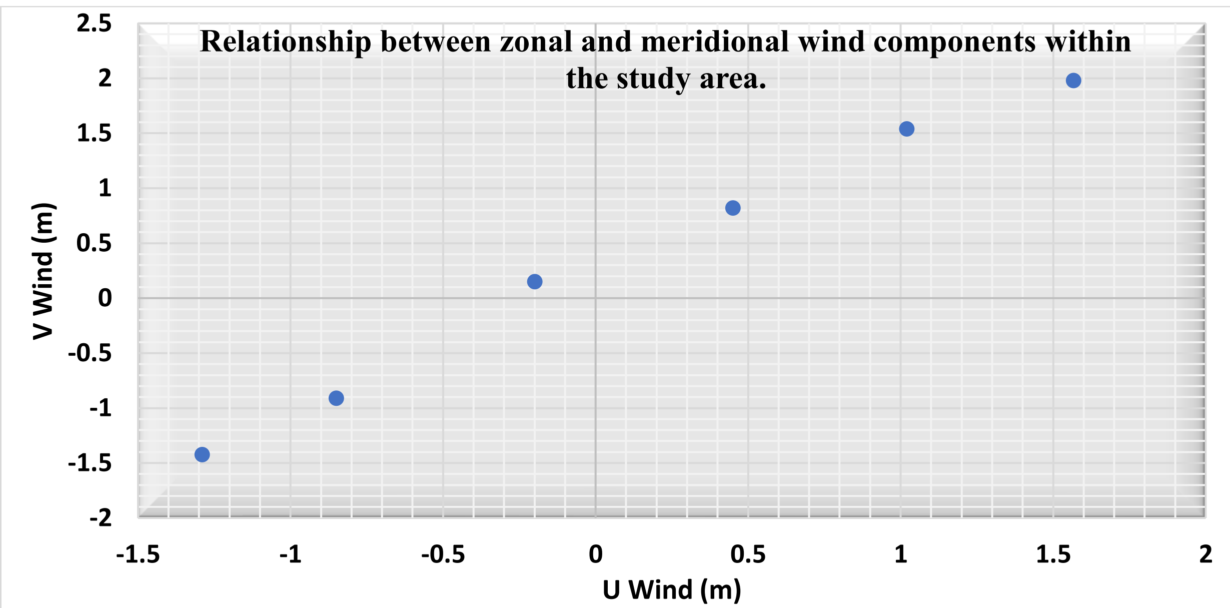

Figure 6: Scatter distribution of zonal and meridional wind components

Spectral Index Analysis (NDVI and NDWI)

The NDVI analysis revealed noticeable temporal variation in vegetation vigor across the Suleja Dam catchment. Higher NDVI values were observed during periods of increased precipitation, reflecting healthier vegetation conditions associated with enhanced soil moisture availability. Areas dominated by woodland and agricultural vegetation consistently exhibited higher NDVI values than exposed soils and built-up environments. The NDWI analysis demonstrated corresponding fluctuations in reservoir water extent. Increased precipitation resulted in positive NDWI responses around the reservoir margins, indicating seasonal expansion of surface water. The combined interpretation of NDVI and NDWI confirms that climatic variability significantly influenced vegetation health and reservoir dynamics throughout the monitoring period.

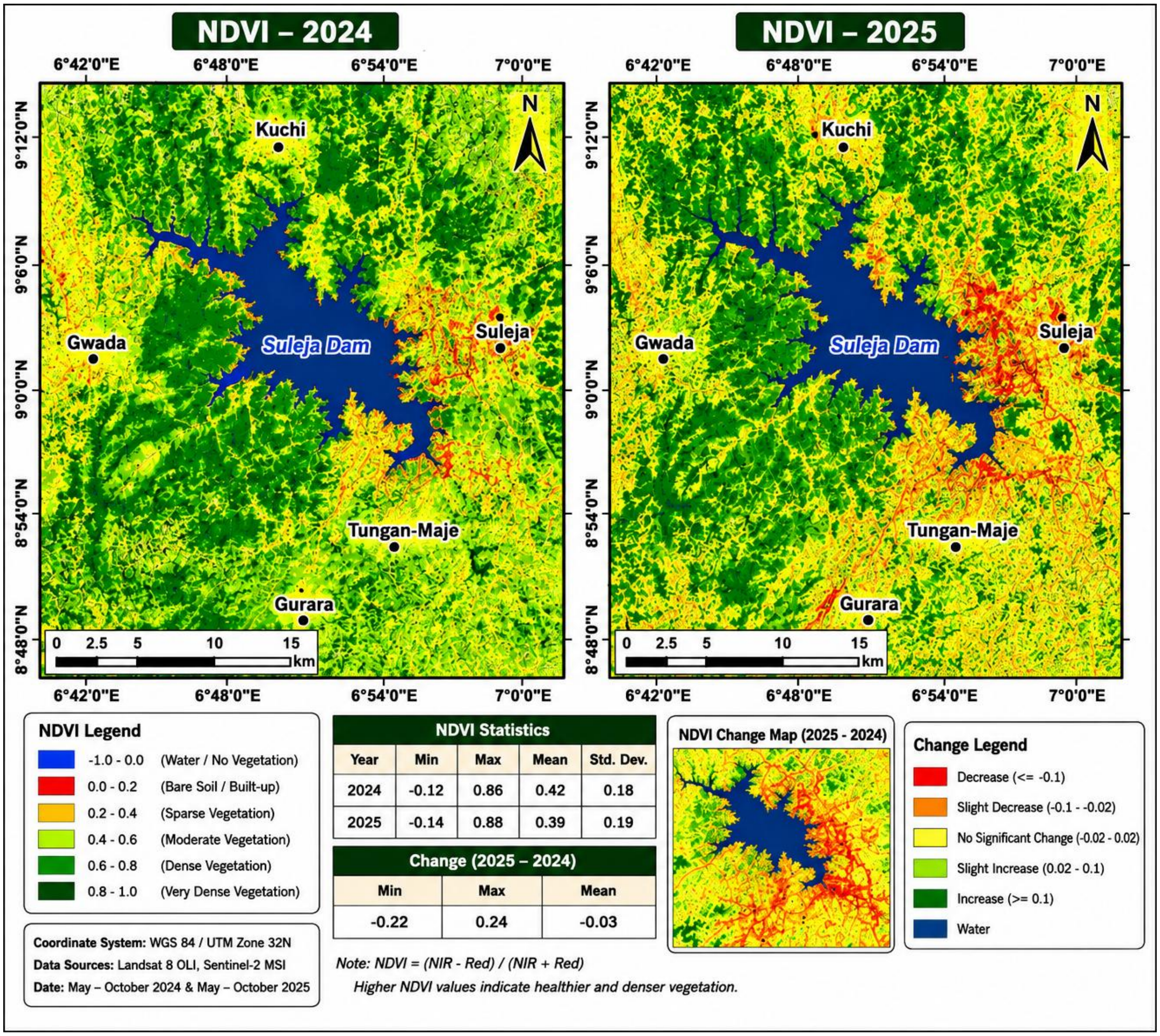

Figure 7: NDVI Maps of the Suleja Dam Catchment for 2024 and 2025

Figure 7 shows the spatial distribution of vegetation health around Suleja Dam in 2024 and 2025. Higher NDVI values (green) indicate dense and healthy vegetation, while lower values (yellow to red) represent sparse vegetation, built-up areas, and bare surfaces. Compared with 2024, the 2025 map reveals a slight decline in vegetation cover, particularly around the eastern part of the catchment due to increasing anthropogenic activities. The NDVI change map indicates localized vegetation loss and limited areas of vegetation recovery, while the reservoir remained relatively stable. Overall, the results demonstrate the effectiveness of NDVI for monitoring vegetation dynamics and environmental changes in the Suleja Dam catchment.

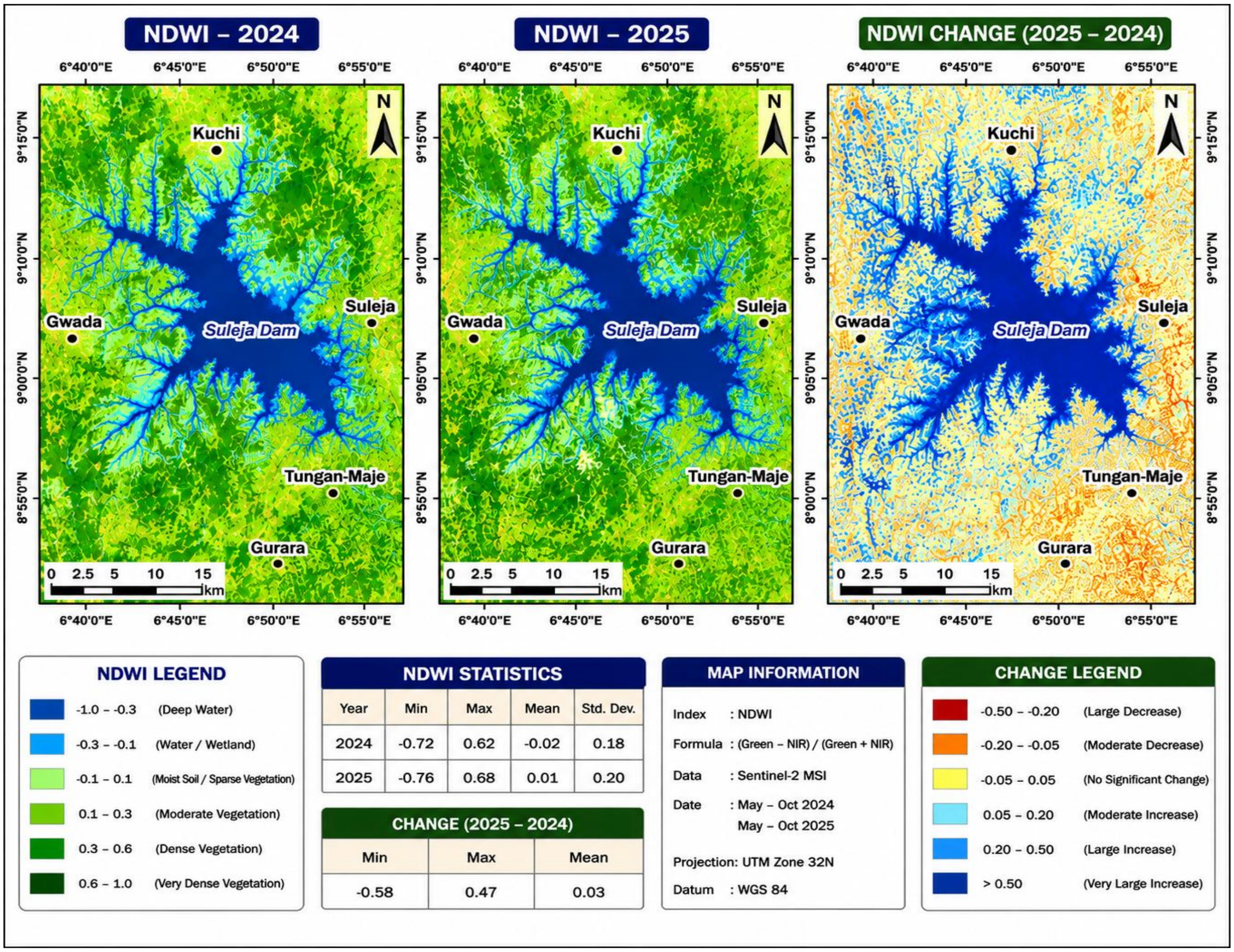

Figure 8: NDWI Maps of the Suleja Dam Catchment for 2024 and 2025

Figure 8 illustrates the spatial distribution of surface water and wetness conditions around Suleja Dam in 2024 and 2025. Higher NDWI values (blue) indicate water bodies and wetter surfaces, while lower values (green to yellow) represent vegetation and drier land areas. Compared with 2024, the 2025 map shows a slight increase in water extent and surface wetness, reflecting seasonal hydrological variations. Overall, the NDWI effectively highlights changes in reservoir dynamics and supports continuous monitoring of the Suleja Dam environment.

Land Cover Classification

The three machine-learning algorithms successfully classified the Suleja Dam environment into four dominant land-cover classes as shown in Table 4.

Table 4: Land Cover Classification Results

Land-cover class	Area (%)
Vegetation	42
Water bodies	28
Built-up	18
Bare surfaces	12

Vegetation remained the dominant land-cover type, reflecting the extensive natural and agricultural cover surrounding the reservoir. The relatively high proportion of water bodies demonstrates the importance of the reservoir for regional water supply, whereas the expansion of built-up land reflects increasing urbanization around Suleja (Figure 9).

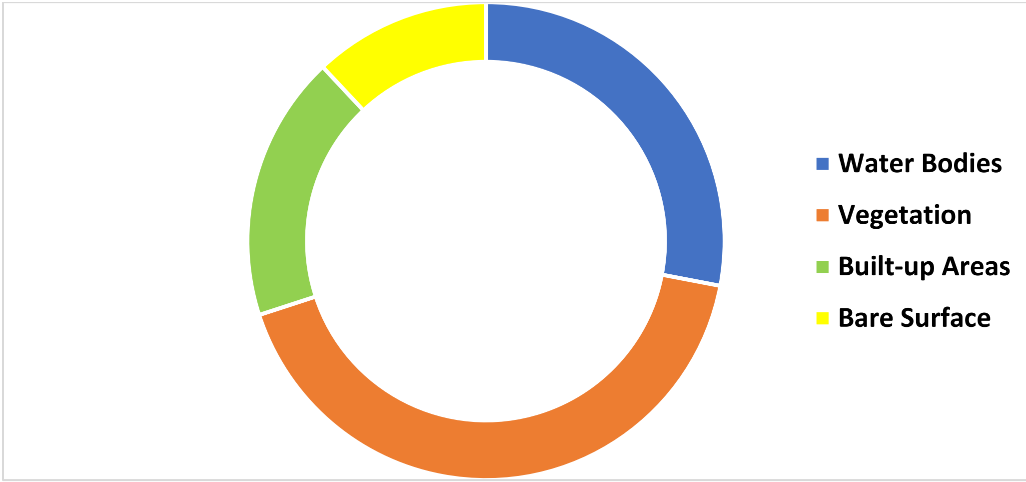

Figure 9: Classified land cover distribution within the Suleja Dam catchment area

Classification Accuracy Assessment

The classification performance of the three machine-learning algorithms was evaluated using confusion matrix statistics. RF achieved the highest overall performance, producing an overall accuracy of 92%, followed by SVM (88%) and ANN (85%).

Random Forest (RF) Results

The RF classifier achieved the highest classification performance, with an overall accuracy of 92.83% and a Cohen's Kappa coefficient of 0.904. It consistently produced high Producer's Accuracy, User's Accuracy, Precision, Recall, and F1-scores across all land-cover classes, indicating strong agreement with the reference data (Tables 5 and 6). The superior performance of RF demonstrates its ability to effectively model complex spectral and environmental relationships, making it the most reliable classifier for land-cover mapping and environmental change detection in the Suleja Dam catchment.

Table 5: Confusion Matrix for Random Forest (RF)

Reference / Classified	Water	Vegetation	Built-up	Bare Surface	Total
Water Bodies	140	4	3	3	150
Vegetation	5	141	2	2	150
Built-up Areas	3	5	138	4	150
Bare Surface	2	4	6	138	150
Column Total	150	154	149	147	600

Validation Samples = 600 (150 samples per class)
Overall Accuracy = 92.83%

Table 6: Random Forest Class-wise Accuracy Statistics (PA, UA, Precision, Recall, F1-score)

Land-cover Class	Producer's Accuracy (%)	User's Accuracy (%)	Precision	Recall	F1-score
Water Bodies	93.33	93.33	0.933	0.933	0.933
Vegetation	94.00	91.56	0.916	0.940	0.928
Built-up Areas	92.00	92.62	0.926	0.920	0.923
Bare Surface	92.00	93.88	0.939	0.920	0.929

Overall Precision = 0.929; Overall Recall = 0.928
Overall F1-score = 0.928; Cohen's Kappa = 0.904

Support Vector Machine (SVM) Results

The Support Vector Machine (SVM) classifier achieved an overall accuracy of 88.67% with a Cohen's Kappa coefficient of 0.849, indicating good classification performance. However, moderate confusion was observed between built-up areas and bare surfaces because of their similar spectral characteristics, resulting in slightly lower Precision, Recall, and F1-scores compared to RF. Nevertheless, SVM produced reliable land-cover classifications and remains an effective approach for environmental monitoring (Tables 7 and 8).

Table 7: Confusion Matrix for Support Vector Machine (SVM)

Reference / Classified	Water	Vegetation	Built-up	Bare Surface	Total
Water Bodies	135	6	5	4	150
Vegetation	8	136	3	3	150
Built-up Areas	5	8	130	7	150
Bare Surface	4	6	9	131	150

Overall Accuracy = 88.67%

Table 8: Accuracy Statistics (SVM)

Land-cover Class	PA (%)	UA (%)	Precision	Recall	F1-score
Water Bodies	90.00	88.82	0.888	0.900	0.894
Vegetation	90.67	87.18	0.872	0.907	0.889
Built-up Areas	86.67	88.44	0.884	0.867	0.875
Bare Surface	87.33	90.34	0.903	0.873	0.888

Overall Precision = 0.887; Overall Recall = 0.887
Overall F1-score = 0.886; Cohen's Kappa = 0.849

Artificial Neural Network (ANN) Results

The Artificial Neural Network (ANN) achieved an overall accuracy of 85.00% and a Cohen's Kappa coefficient of 0.800, representing acceptable classification performance. Although ANN successfully captured nonlinear relationships within the multisource datasets, it exhibited comparatively higher misclassification among transitional land-cover classes, leading to lower class-specific accuracy metrics than RF and SVM. Despite these limitations, the ANN results demonstrate its potential for land-cover mapping, particularly when supported by larger training datasets and optimized model parameters (Tables 9 and 10).

Table 9: Confusion Matrix for Artificial Neural Network (ANN)

Reference / Classified	Water	Vegetation	Built-up	Bare Surface	Total
Water Bodies	129	8	7	6	150
Vegetation	10	131	5	4	150
Built-up Areas	7	10	124	9	150
Bare Surface	5	8	11	126	150

Overall Accuracy = 85.00%

Table 10: Accuracy Statistics (ANN)

Land-cover Class	PA (%)	UA (%)	Precision	Recall	F1-score
Water Bodies	86.00	85.43	0.854	0.860	0.857
Vegetation	87.33	83.97	0.840	0.873	0.856
Built-up Areas	82.67	84.93	0.849	0.827	0.838
Bare Surface	84.00	86.90	0.869	0.840	0.854

Overall Precision = 0.853; Overall Recall = 0.850
Overall F1-score = 0.851; Cohen's Kappa = 0.800

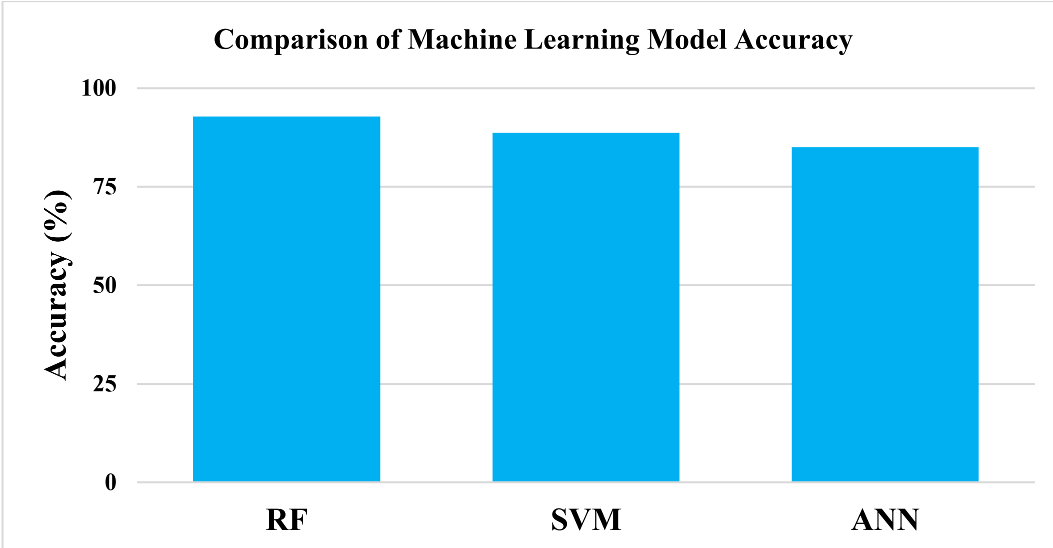

Figure 10: Comparison of Machine Learning Model Classification Accuracy

Figure 10 demonstrates a comparison of the overall classification accuracy of the three machine learning algorithms. RF achieved the highest overall accuracy (92.83%), followed by SVM (88.67%) and ANN (85.00%). The superior performance of RF indicates its greater capability to model complex nonlinear relationships among multispectral, climatic, and topographic variables, making it the most suitable classifier for land-cover mapping and subsequent environmental change detection in the Suleja Dam catchment.

Classification Uncertainty

Although the classification accuracies obtained in this study are high, uncertainty remains an inherent component of supervised image classification. Misclassification occurred primarily along class boundaries where built-up areas, exposed soils, and sparsely vegetated surfaces exhibited similar spectral responses. Seasonal changes in vegetation condition, mixed pixels associated with the 10–30 m spatial resolution of Sentinel-2 and Landsat imagery, and limited representation of heterogeneous land-cover features within the training samples also contributed to classification uncertainty. Consequently, the reported land-cover changes, particularly the relatively small increases in water extent (+3.2%) and bare surface (+2.1%), should be interpreted within the context of these uncertainties. Nevertheless, the high overall accuracies, Kappa coefficients, and consistently strong class-specific performance metrics indicate that the observed reduction in vegetation cover (−6.5%) and expansion of built-up areas (+5.8%) represent robust environmental trends rather than artifacts of classification error. These findings support the reliability of the machine learning framework for monitoring environmental changes within the Suleja Dam catchment while highlighting the importance of future validation using higher-resolution imagery and additional field observations.

Environmental Change Detection

Post-classification comparison of the 2024 and 2025 land-cover maps, using the Random Forest classifier, revealed measurable environmental changes within the Suleja Dam catchment. The results indicate that both climatic variability and anthropogenic activities influenced land-cover dynamics. As shown in Table 14, vegetation cover decreased by 6.5%, mainly due to urban expansion and agricultural activities. In contrast, built-up areas increased by 5.8%, reflecting ongoing infrastructural development around the reservoir. Reservoir water extent increased by 3.2%, consistent with higher precipitation and positive NDWI responses, while bare surfaces expanded by 2.1% because of localized land clearing and construction activities. The spectral indices supported these observations, with declining NDVI corresponding to vegetation loss and increasing NDWI indicating seasonal expansion of reservoir water. Overall, the results demonstrate gradual landscape modification driven by climate variability and human activities, highlighting the need for continuous geospatial monitoring to support sustainable watershed and reservoir management.

Table 11: Environmental Change Detection Results (2024–2025)

Land-cover Class	Area in 2024 (%)	Area in 2025 (%)	Change (%)	Interpretation
Vegetation	48.5	42.0	−6.5	Reduction due to urban expansion and localized vegetation clearing
Water Bodies	24.8	28.0	+3.2	Increased reservoir extent associated with seasonal recharge
Built-up Areas	12.2	18.0	+5.8	Expansion of residential and infrastructural development
Bare Surface	9.9	12.0	+2.1	Increased exposed soil from land disturbance and construction

Figure 11 shows the spatial distribution of land-cover changes between 2024 and 2025. Vegetation declined mainly around expanding settlements, while built-up areas increased near the southern reservoir. Slight expansion of water bodies and localized increases in bare surfaces indicate seasonal reservoir recharge and land disturbance. Figure 12 compares the NDVI distribution for 2024 and 2025. Reduced NDVI values in 2025 indicate vegetation loss and lower vegetation vigor, particularly near developed areas, whereas higher NDVI values remained in relatively undisturbed vegetated zones. Figure 13 illustrates the spatial variation in NDWI between 2024 and 2025. Increased NDWI values around the reservoir indicate expanded surface water associated with higher precipitation, while surrounding areas exhibited minimal changes in surface moisture.

Figure 14 summarizes land-cover changes between 2024 and 2025. Vegetation decreased by 6.5%, while built-up areas, water bodies, and bare surfaces increased by 5.8%, 3.2%, and 2.1%, respectively, highlighting the combined effects of urban expansion and climatic variability on the Suleja Dam catchment.

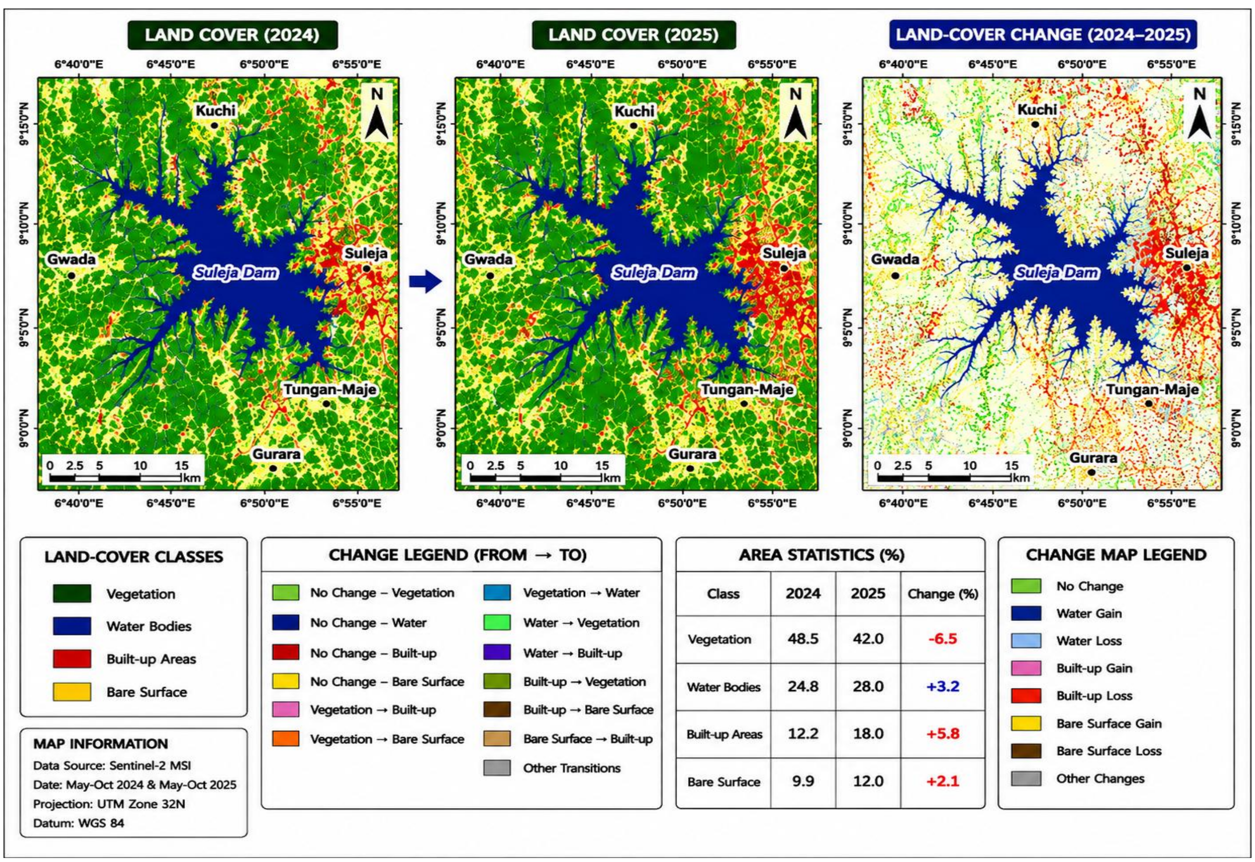

Figure 11: Land-cover change map (2024–2025)

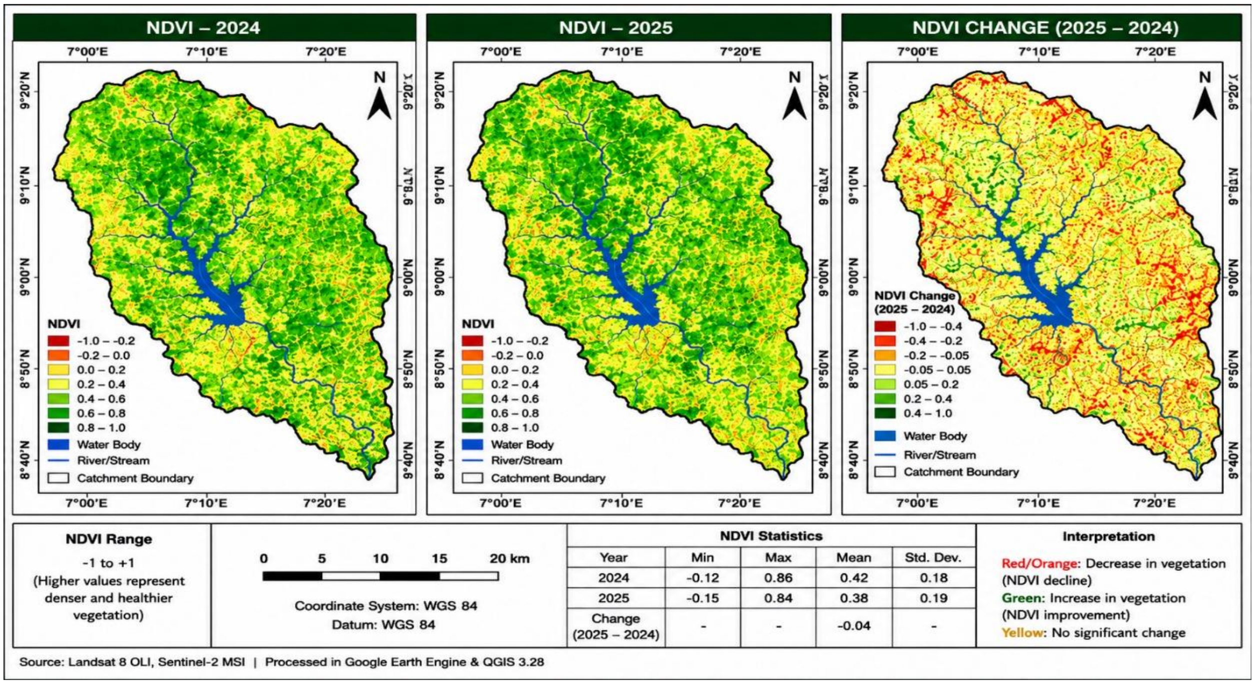

Figure 12: NDVI comparison map (2024 vs. 2025)

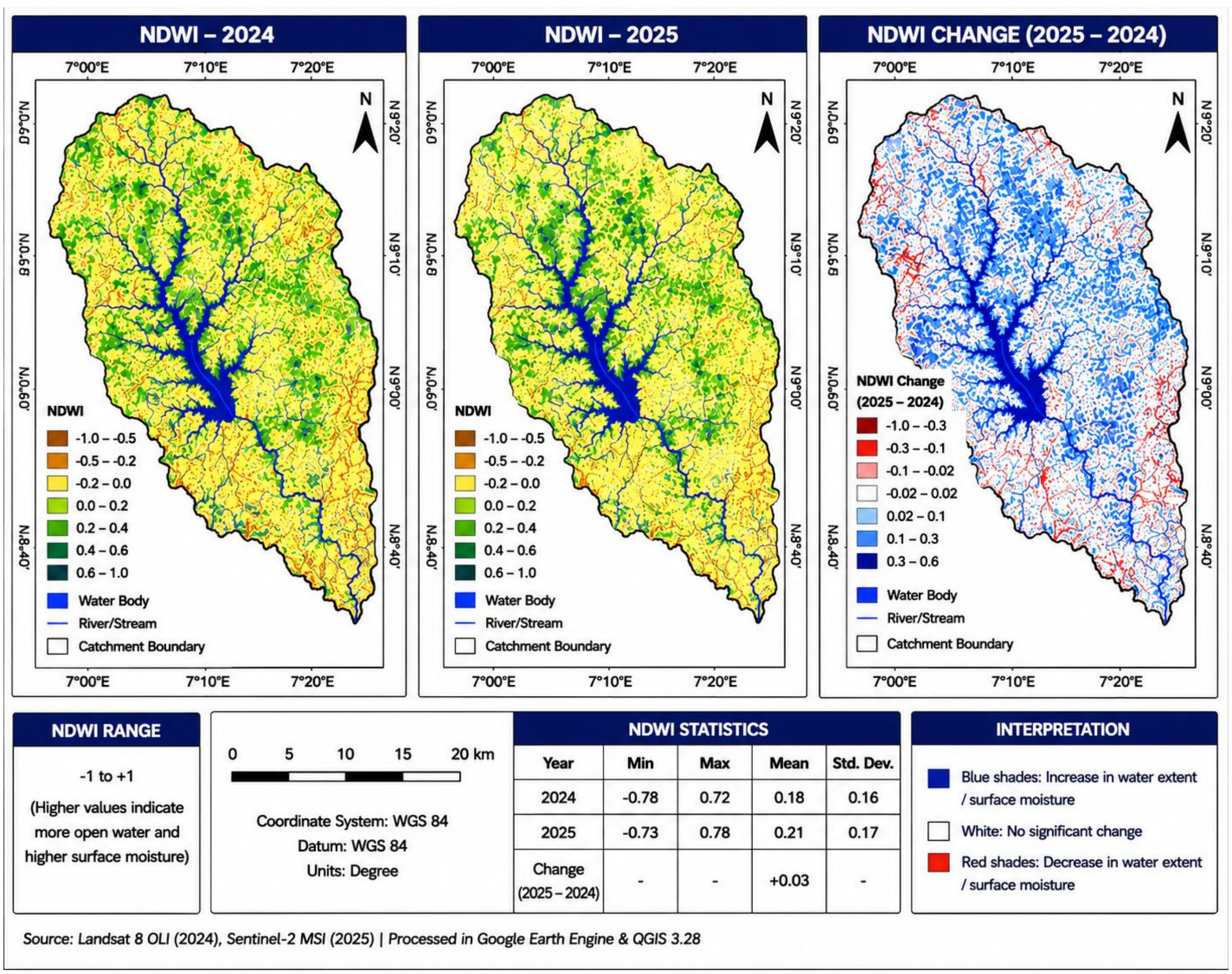

Figure 13: NDWI comparison map (2024 vs. 2025)

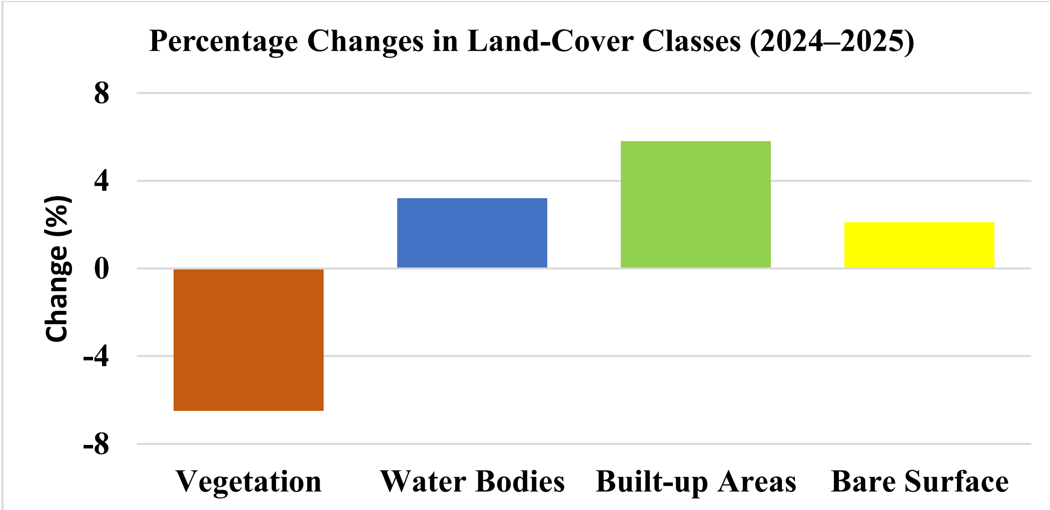

Figure 14: Bar chart illustrating percentage changes in land-cover classes

Discussion

The integration of machine learning, remote sensing, and GIS effectively characterized environmental changes within the Suleja Dam catchment. Among the evaluated classifiers, the Random Forest (RF) model achieved the highest overall classification accuracy (92.83%) and Cohen's Kappa (0.904), outperforming the Support Vector Machine (SVM) and Artificial Neural Network (ANN). The superior performance of RF can be attributed to its ensemble learning approach, which efficiently captures nonlinear relationships among multispectral, topographic, and climatic variables while reducing overfitting. This finding is consistent with previous studies that identified Random Forest as one of the most robust classifiers for land-cover mapping using multisource remote sensing data (Belgiu & Drăguţ, 2016; Maxwell et al., 2018).

The detected land-cover changes indicate that both environmental and anthropogenic factors influenced the dynamics of the Suleja Dam catchment. Vegetation declined by 6.5%, while built-up areas increased by 5.8%, reflecting continued urban expansion and land conversion around the reservoir. Similar trends have been widely reported in rapidly developing watersheds, where population growth and infrastructure development contribute to vegetation loss and increased impervious surfaces, thereby altering watershed hydrology and ecological integrity (Lambin & Geist, 2006; Foley et al., 2005). Conversely, the 3.2% increase in water extent, supported by positive NDWI responses, corresponds with increased precipitation during the monitoring period and indicates seasonal recharge of the reservoir. These observations demonstrate the value of integrating spectral indices with climatic variables for monitoring hydrological variability.

The temporal behaviour of NDVI and NDWI further validated the classification results. Areas exhibiting reduced NDVI generally coincided with locations where vegetation was converted to built-up or bare surfaces, whereas increases in NDWI corresponded to expansion of surface water. The complementary use of spectral indices and machine learning therefore enhanced the reliability of environmental change detection by combining continuous indicators of vegetation and water conditions with thematic land-cover classification. This agrees with previous studies demonstrating that NDVI and NDWI are effective indicators for monitoring vegetation health and surface water dynamics using multispectral satellite imagery (McFeeters, 1996; Pettorelli et al., 2005).

Although the classification accuracies were high, some uncertainty remains because of spectral similarity among built-up areas, bare surfaces, and sparsely vegetated land, as well as the moderate spatial resolution of Landsat and Sentinel-2 imagery. Nevertheless, the consistently high Producer's Accuracy, User's Accuracy, F1-scores, and Kappa coefficients indicate that the observed land-cover transitions are robust and unlikely to be dominated by classification errors. Overall, the results demonstrate that integrating machine learning with remote sensing and GIS provides a reliable framework for environmental monitoring and supports evidence-based watershed management and sustainable planning within the Suleja Dam catchment.

CONCLUSION

This study demonstrated the effectiveness of integrating machine learning, remote sensing, and GIS for environmental monitoring of the Suleja Dam catchment. The combined use of Landsat 8, Sentinel-2, SRTM DEM, and climatic datasets enabled reliable assessment of land-cover dynamics and environmental conditions. Among the evaluated classifiers, the Random Forest model achieved the highest classification performance, confirming its suitability for land-cover mapping in heterogeneous watershed environments.

The analysis revealed a 6.5% decline in vegetation cover, accompanied by increases in built-up areas (5.8%), water extent (3.2%), and bare surfaces (2.1%), indicating that both anthropogenic activities and climatic variability are influencing the catchment. The consistency between the machine learning classifications and the NDVI and NDWI analyses further strengthened the reliability of the detected environmental changes.
Overall, the findings demonstrate that integrating machine learning with remote sensing provides a robust and cost-effective framework for continuous environmental monitoring and change detection. The approach can support watershed management, reservoir protection, land-use planning, and informed decision-making by providing timely geospatial information for sustainable management of the Suleja Dam catchment.

ACKNOWLEDGEMENT
The authors also thank Prof. M. N. Ono, Director of Open Spaces and Space Allocation (OSSA), Nnamdi Azikiwe University (UNIZIK), for his continuous encouragement and for his valuable insights and methodological guidance during the course of this study.

REFERENCES
Adeboboye, A. J., Atoki, L. O. and Agada, A. S. (2024). Flood Vulnerability Mapping Using GIS and Remote Sensing. *International Journal of Multidisciplinary Research and Growth Evaluation*, 5(03). https://doi.org/10.5281/zenodo.12702664

Aleksandra, K. and Jan, B. (2025). Remote Sensing Perspective on Monitoring and Predicting Underground Energy Sources Storage Environmental Impacts: Literature Review. *Remote Sensing*, *17*(15), 2628; https://doi.org/10.3390/rs17152628

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Chen, Y., Li, X., & Zhao, H. (2022). Deep learning-based dam deformation monitoring using remote sensing data. *Remote Sensing*, 14(8), 1765.

Chang, S. T., Lin, S. Y., Lin, Y. C. (2025). Deep learning-based surface deformation tracking with interferometric fringes: A case study in Taiwan. International Journal of Applied Earth Observation and Geoinformation, 143, https://doi.org/104796.10.1016/j.jag.2025.104796

Chen, L., Liu, L., Liu, S., Shi. Z. and Shi, C. (2025). The Application of Remote Sensing Technology in Inland Water Quality Monitoring and Water Environment Science: Recent Progress and Perspectives. *Remote Sensing*, *17*(4), 667; https://doi.org/10.3390/rs17040667

European Space Agency (2023). Sentinel satellite missions for Earth observation and environmental monitoring. ESA Publications.

Foley, J. A., Defries, R. S., Asner, G. P., Barford, C. C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H., Helkowski, J. H., Holloway, T., Howard, E., Kucharik, C. J., Monfreda, C., Patz, P., Prentice, I. C., Ramankutty, N., and Snyder, P. K. (2005). Global consequences of land use. *Science, 309*(5734), 570–574. https://doi.org/10.1126/science.1111772

Giulio C., Luigi P., Ruth Q., Jeroen A., Paul E., Maurits E., Stephen H., Walter L. F., Natalia L., Bongani M., Doug G. N., Keziah N., Nobubelo N., Allison P., Cate R., Josep D. T., Lorenzo V., Jessica A. E. and Elena B. (2022). Sand Dams for Sustainable Water Management: Challenges and Future Opportunities. SSRN Electronic Journal, https://doi.org/10.2139/ssrn.3998987

ICOLD (2021). Dam Safety Guidelines and Monitoring Practices. International Commission on Large Dams.

Lambin, E. F., & Geist, H. J. (2006). *Land-Use and Land-Cover Change: Local Processes and Global Impacts*. Springer.

Li, J., Wang, S., & Liu, P. (2021). Remote sensing applications in reservoir and dam monitoring: A review. *Water Resources Management*, 35(7), 2135–2150.

Li, Y., Zhang, H., Yang, Z., Cao, Y. and Su. Y. (2024). Research and Application of Dam Safety Monitoring Information Management System Based on GIS Platform. Proceedings of Conference on Sustainable Traffic and Transportation Engineering in 2023. https://doi.org/10.1007/978-981-97-5814-2_49

Ma, C., Xu, X., Yang, J. and Cheng, L. (2023). Safety Monitoring and Management of Reservoir and Dams. *Water*, *15*(6), 1078; https://doi.org/10.3390/w15061078

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing,* *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

McFeeters, S. K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432.

Myung, S. C. and Jiaguo, Q. (2025). Remote sensing-based assessment of dam impacts on hydrology, geomorphology, ecosystems, and society – a review. Environmental Earth Sciences 84(12):1-20. https://doi.org/10.1007/s12665-025-12345-7

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J.-M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution, 20*(9), 503–510. https://doi.org/10.1016/j.tree.2005.05.011

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, NASA SP-351, 309–317.

Swapna A. J., Khalid M. A. (2024). Remote Sensing Techniques for Water Quality Monitoring: A Review. Sensors, 24(24):8041. https://doi.org/10.3390/s24248041

Yan S., Jiayuan F., Weiwei L., Chuan L., Xiaohe L., and Xiudong X. (2025). Dam Deformation Monitoring Model Based on Deep Learning and Split Conformal Quantile Prediction. *Applied Science, 15*(4), 1960; https://doi.org/10.3390/app15041960

Zhang, H., Zhou, L., & Wu, J. (2024). Artificial intelligence applications in infrastructure and dam monitoring. *Engineering Applications of Artificial Intelligence*, 128, 107534.