Journal of Spatial Information Sciences

... JSIS



MODELLING VEGETATION PATTERNS IN SOUTHERN KADUNA NIGERIA USING XGBOOST AND SHAP

JACOB REUBEN JOBIEN, YAKUBU ISMA'IL YUNISA, EKELE JOHN IFENE



JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 DOI: https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026

www.journals.unizik.edu.ng/jsis

MODELLING VEGETATION PATTERNS IN SOUTHERN KADUNA NIGERIA USING XGBOOST AND SHAP

[1]Jacob Reuben Jobien, [2]Yakubu Isma'il Yunisa, [3]Ekele John Ifene

[1]Department of Surveying and Geoinformatics, Confluence University of Science and Technology, Osara, Kogi State, Nigeria
[2]Institute of Maritime Studies, University of Lagos
[3]School of Postgraduate Studies, Nasarawa State University, Keffi, Nigeria
Corresponding Author's Email: email: jobienjacob@gmail.com, 08067186726

DOI: https://doi.org/10.5281/zenodo.20764537

Abstract

Understanding the drivers of vegetation dynamics is critical for sustainable land management and ecological conservation. This study applied the XGBoost machine learning algorithm to model NDVI-based vegetation patterns in Southern Kaduna, Nigeria, using hydrological, climatic, topographic, and anthropogenic predictors within the R statistical computing environment. The model demonstrated excellent predictive performance, with training $R^2 = 0.9978$, RMSE = 0.0022, MAE = 0.0016; spatial cross-validation $R^2 = 0.9054$, RMSE = 0.0142, MAE = 0.0100; and independent testing $R^2 = 0.9523$, RMSE = 0.0101, MAE = 0.0076, highlighting its ability to capture complex nonlinear interactions. Global and local SHAP analyses revealed that NDWI, representing surface moisture and hydrological conditions, was the most influential predictor, followed by NDBI, reflecting urbanization and land-use change. Climatic variables, slope, elevation, evapotranspiration, and population density exerted secondary but meaningful effects. SHAP dependence plots further illustrated the directional relationships and interactions among predictors, confirming that vegetation greenness is highly sensitive to moisture availability and urban development pressures. These findings emphasize the dominant role of hydrological and anthropogenic factors in controlling vegetation variability in the region. The study provides actionable insights for ecological management, urban planning, and sustainable development in Southern Kaduna.

Keywords: Vegetation (NDVI), XGBoost, SHAP Analysis, Machine Learning, Southern Kaduna

1.0 INTRODUCTION

Vegetation is a fundamental component of the Earth's system, playing a critical role in regulating global climate, maintaining biodiversity, supporting ecosystem services, and sustaining human livelihoods [1], [2], [3]. It influences key processes such as carbon cycling, energy balance, and hydrological regulation, making it essential for both environmental stability and socio-economic development [2], [4]. At the global scale, vegetation dynamics are increasingly affected by climate change, land use transformation, and population growth, leading to significant alterations in ecosystem structure and function [5], [6]. Understanding and predicting vegetation patterns have therefore become central to addressing global environmental challenges, including land degradation, food security, and climate resilience [7], [8]. In regions such as Southern Kaduna, Nigeria, these global pressures manifest locally through a combination of climatic variability, topographic

diversity, and increasing anthropogenic activities, necessitating robust approaches for vegetation monitoring and prediction.

Remote sensing has become an indispensable tool for vegetation assessment due to its ability to provide consistent and spatially extensive observations [9], [10]. The Normalized Difference Vegetation Index (NDVI) is a widely used indicator of vegetation greenness, density, and health, sensitive to chlorophyll content and capable of capturing temporal and spatial variations [11], [12]. Vegetation dynamics reflected by NDVI are influenced by both environmental and human-related factors, including hydrological conditions measured by NDWI and urbanization or land-use change captured by NDBI [13], [14]. Climatic variables such as rainfall, temperature, and evapotranspiration regulate plant growth and water balance [15], [16]. In addition, topographic variables such as slope and elevation influence microclimatic conditions, soil properties, and water distribution [17], [18]. Population density further represents human pressure on vegetation through land conversion and resource exploitation [19], [20].

Previous studies [21], [22], [23], [24] have explored the relationship between NDVI and environmental drivers using both statistical and machine learning approaches. Traditional regression models have been applied to examine linear relationships, while machine learning techniques such as Extreme Gradient Boosting (XGBoost) have demonstrated improved performance in capturing nonlinear interactions [25], [26], [27], [28], [29]. In environmental modelling, XGBoost has demonstrated superior predictive performance compared with several conventional methods, especially when applied to heterogeneous datasets involving climatic, hydrological, topographic, and anthropogenic factors [30], [31].

Despite these advancements, important research gaps remain. Many studies have considered a limited set of predictors, often focusing only on climatic or land use variables, without integrating hydrological, topographic, and socio-economic factors in a unified modelling framework. In addition, there is limited application of advanced machine learning techniques, particularly XGBoost, in vegetation studies within Nigeria, especially at finer spatial scales such as Southern Kaduna. Furthermore, while machine learning models can achieve high predictive accuracy, their interpretability remains a challenge. Few studies have incorporated explainable artificial intelligence techniques to understand the contribution of predictors influencing vegetation patterns. This limits the ability to translate model outputs into actionable environmental insights.

Therefore, this study integrated multiple variables, including NDWI, NDBI, temperature, slope, elevation, rainfall, evapotranspiration, and population density, to predict vegetation dynamics using NDVI as the response variable. XGBoost was employed to capture complex nonlinear relationships among these predictors and to assess model performance. SHapley Additive exPlanations (SHAP) were used to quantify the contribution of each predictor and reveal interaction effects. The study aimed to develop a robust predictive model for vegetation distribution in Southern Kaduna, evaluate predictor importance, and provide interpretable insights to support evidence-based land-use planning, vegetation management, and environmental sustainability.

2.0 MATERIAL AND METHODS

2.1 Study Area

Southern Kaduna is located between Latitudes 09º00'00'' and 10º45'00'' of the Equator and between Longitudes 7º15'00'' and 8º40'00'' of the Greenwich meridian as shown in Figure 1. It consists of Jaba, Jema'a, Kachia, Kagarko, Kaura, Sanga, Kauru and Zangon Kataf Local Government Areas (LGAs). The average annual air temperature in Southern Kaduna is about 25.2 °C [32]. March to May is the hottest period, with surface temperatures around 29°C and warm air temperatures [33]. In the southern regions of Kaduna state, precipitation can reach 500mm per month during the peak of the wet season, which runs from April to September. This rainy period typically spans 160–180 days, reaching its maximum intensity in August. Kafanchan records the highest mean annual rainfall at approximately 1659mm. Southern Kaduna lies on an undulating plateau dissected by River Kaduna and its tributaries [34]. River Kaduna is a long plateau river with single-peak flow, receiving many short tributaries [35].

2.2 Types and Sources of Data

All datasets were downloaded from Google Earth Engine (GEE) as presented in Table 1. Vegetation and land cover indices, including the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Normalized Difference Built-up Index (NDBI), were derived from Landsat 8 imagery for 2025. Climatic variables, including temperature, rainfall, and evapotranspiration for the period January–December 2025, were obtained from ERA5 via GEE. Population density data corresponding to 2020 were included, and Shuttle Radar Topography Mission (SRTM) data were used to generate elevation. The slope layer was derived from the DEM in GEE, and all variables were subsequently mapped in ArcGIS as shown in Figure 2. All raster datasets were projected to UTM Zone 32 North, resampled to 100m resolution using bilinear interpolation, and clipped to the study area boundary in ArcGIS to ensure spatial alignment and

JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 DOI: https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026
www.journals.unizik.edu.ng/jsis

uniform resolution. Pixels with missing or invalid values were removed. From the resulting clean dataset, 10,000 valid pixels were randomly sampled for model training and validation.

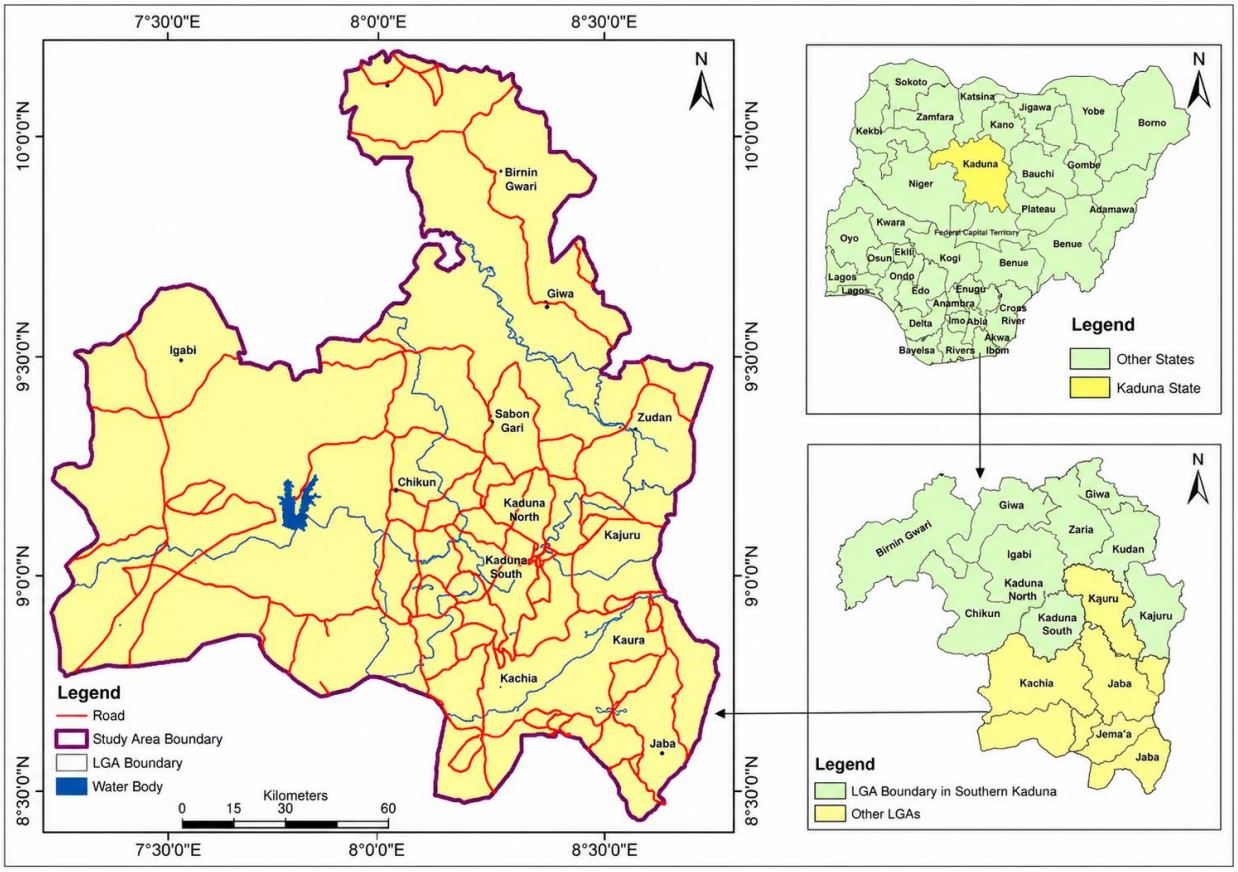


Figure 1: The study Area

Table 1: Types and Sources of Data

SN	Name	Resolution	Source	Processing Platform	Purpose
1	Landsat 8 (2025)	30m	LandSat 8	Google Earth Engine (GEE)	Response variable and predictors
2	Temperature, rainfall, and evapotranspiration	4638.3m	ERA5 (2025)	GEE	Predictors
3	Elevation and slope	30m	SRTM DEM	GEE	Predictors
4	Population Density (2020)	927.67m	GPWv4	GEE	Predictor

GPWv4: Gridded Population of World Version 4

2.3 Computation of NDVI, NDWI and NBDI

2.3.1 Computation of Normalized Difference Water Index (NDWI)

The NDVI for the study area was derived using Google Earth Engine and is calculated using the following formula [36]:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

where, NIR and RED are the near infrared and red bands of Landsat 8, respectively

2.3.2 Computation of Normalized Difference Water Index (NDWI)

The NDWI for the study area was derived using Google Earth Engine and is computed according to the following formula [36]:

$$NDWI = \frac{G - NIR}{G + NIR} \quad (2)$$

where, G and NIR are the green and near infrared bands of Landsat 8, respectively.

2.3.3 Computation of Normalized Difference Built-up Index (NDBI)

The NDBI for the study area was calculated using Google Earth Engine and is mathematically expressed using the following formula [37]:

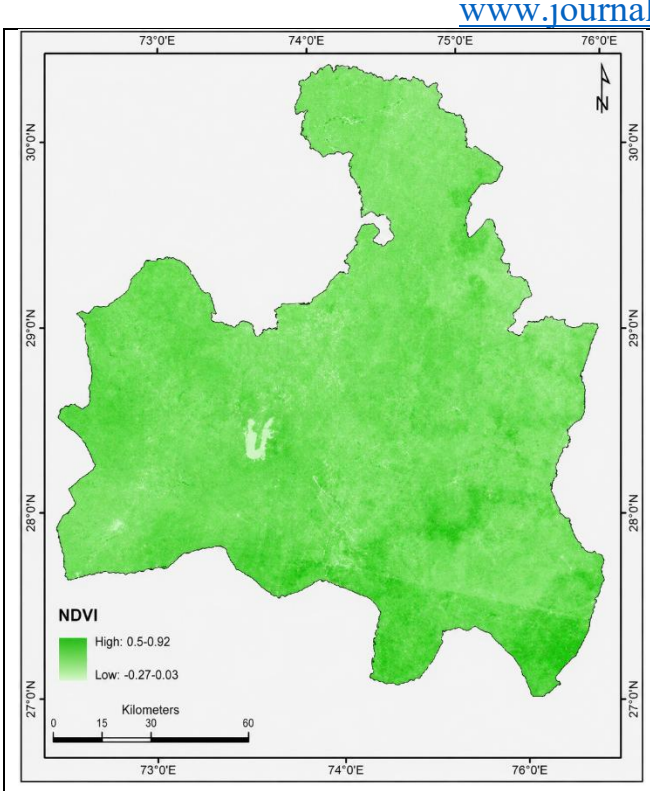
$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \quad (3)$$

where, SWIR and NIR are the short wave and near infrared bands of Landsat 8.

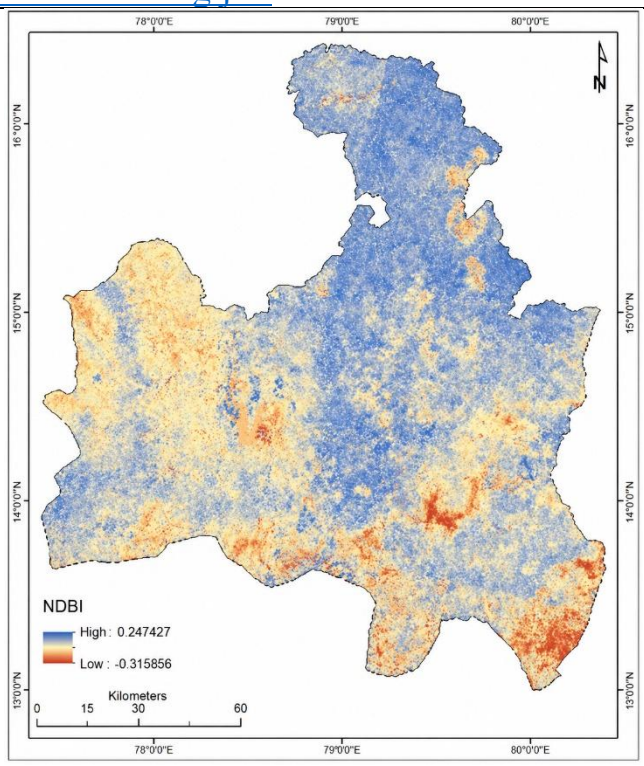
2.4 Mapping of Variables

The dependent variable (NDVI) and the independent variables were processed and spatially mapped in ArcGIS to visualize their distribution across the study area, as shown in Figure 2.

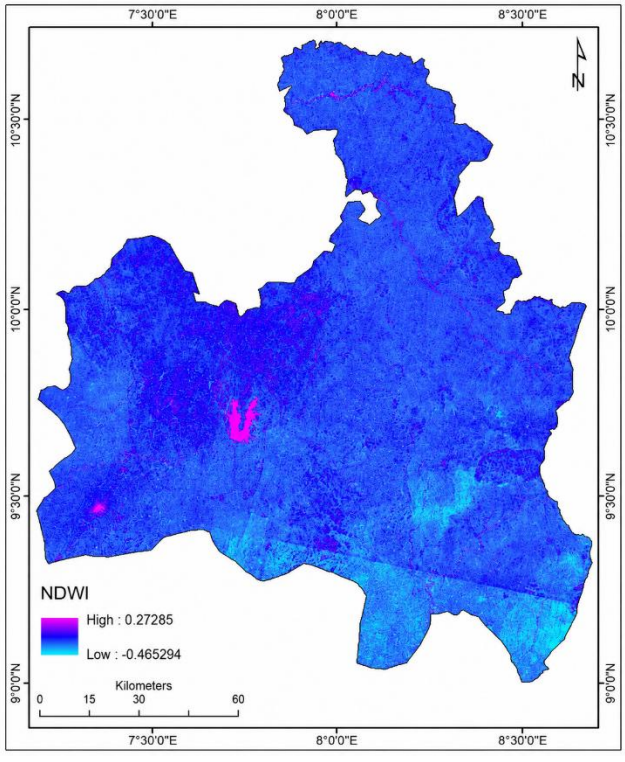
JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 **DOI:** https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026
www.journals.unizik.edu.ng/jsis



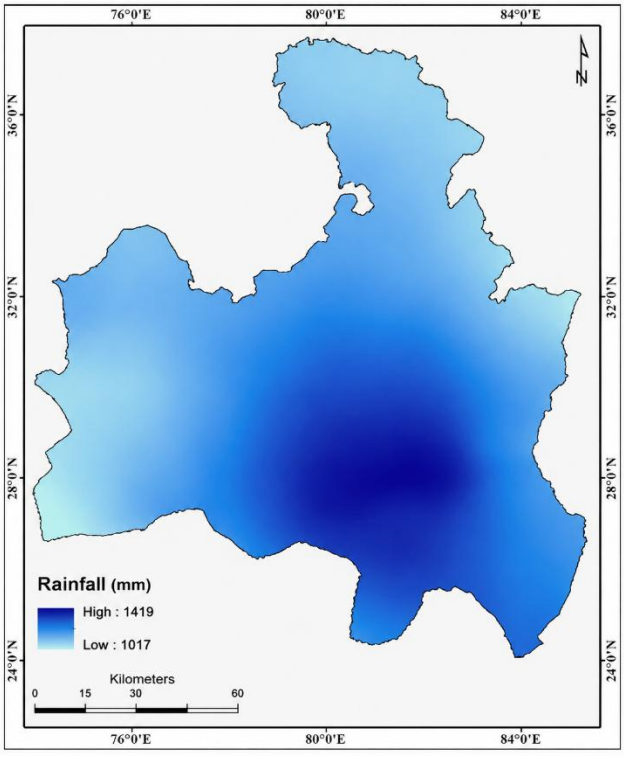
NDVI



NDBI

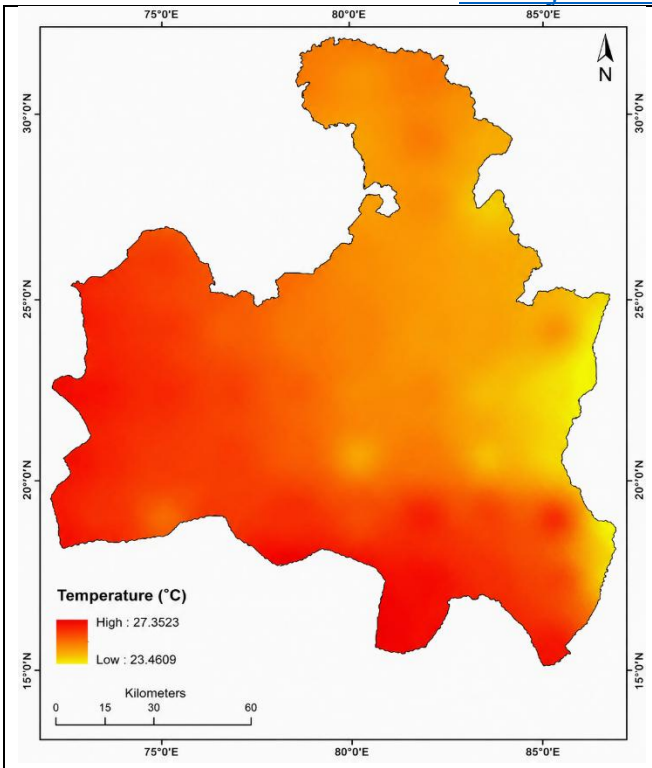


NDWI

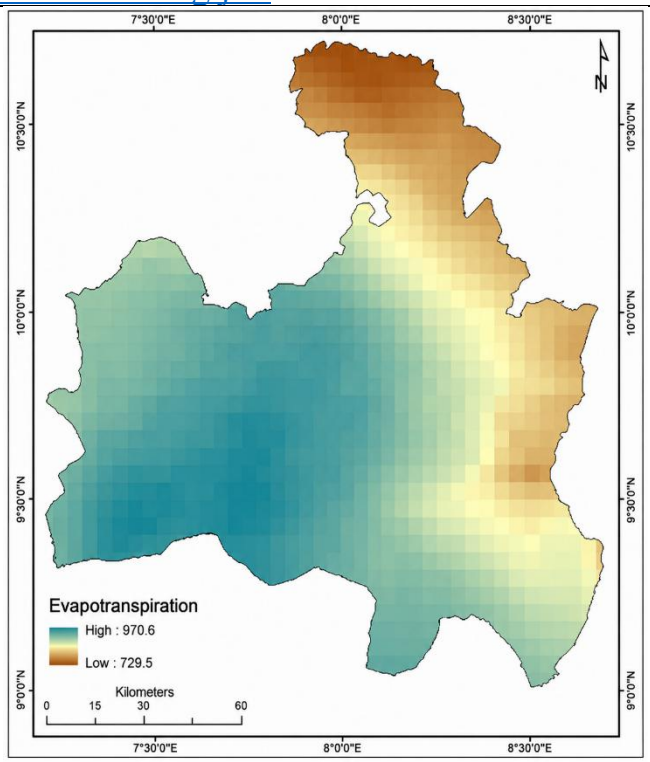


Rainfall

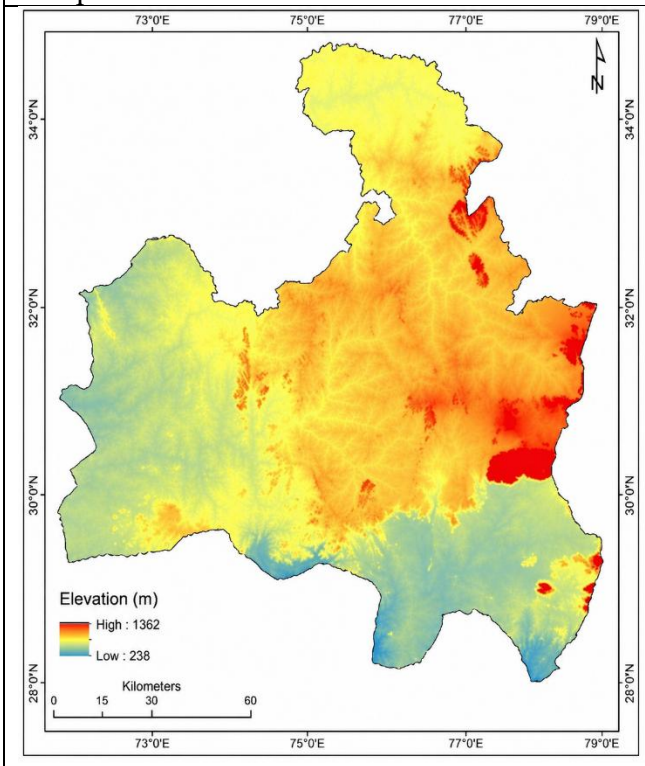
JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 DOI: https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026
www.journals.unizik.edu.ng/jsis



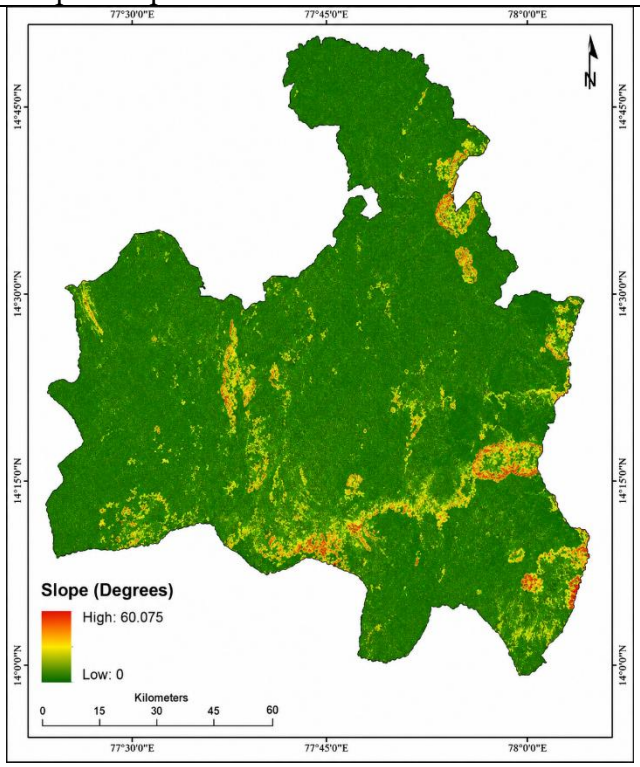
Temperature



Evapotranspiration

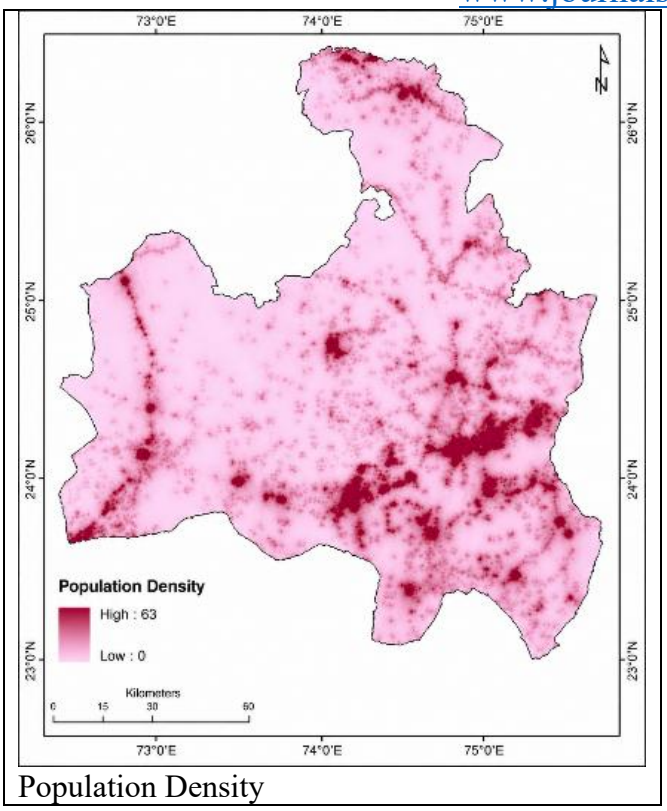


Elevation



Slope

JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 DOI: https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026
www.journals.unizik.edu.ng/jsis



Population Density
Figure 2: Maps of Variables

2.5 Spatial Cross-Validation, XGBoost Modelling and Hyperparameter Tuning

An Extreme Gradient Boosting (XGBoost) regression model was employed to predict NDVI using the predictor variables: NDWI, NDBI, Temperature, Slope, Elevation, Rainfall, Evapotranspiration, and Population Density. The dataset was divided into 70% training and 30% test sets. To account for spatial autocorrelation, 5-fold spatial cross-validation with hexagonal blocking was applied to the training dataset. Hyperparameter tuning was conducted to optimize model performance, adjusting the maximum tree depth (max_depth), learning rate (eta), and number of boosting rounds (nrounds). The final parameters used were max_depth = 6, eta = 0.1, and nrounds = 1000. These settings ensured that the model accurately captured the nonlinear relationships between predictors and vegetation indices.

2.6 Model Training and Evaluation

The XGBoost model was trained on the full training dataset using the selected hyperparameters. Model performance was evaluated using coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) for the training set, spatial cross-validation folds, and independent test dataset. Scatter plots of observed versus predicted values were generated to visually assess model accuracy and bias. RMSE, MAE, and R^2 were calculated using the following formula [38]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \quad (4)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2} \quad (6)$$

where,

n is the total number of observations/data points, i is the index of each observation (from 1 to n), y_i is the actual value of NDVI for observation i, $\hat{y}_i$ is the predicted value of NDVI for observation i, and $\bar{y}$ is the mean of the observed values of NDVI.

The trained XGBoost model was applied to the full raster grid to generate predicted NDVI values across the study area.

2.7 SHAP-Based Model Interpretation

To interpret the contribution of predictor variables, SHapley Additive exPlanations (SHAP) were computed using the fastshap and shapviz packages. SHapley Additive exPlanations (SHAP) analysis was applied to quantify the relative contribution of each predictor variable to NDVI predictions, allowing interpretation of both the importance and direction of influence of environmental, topographic, and anthropogenic factors. SHAP-based analyses produced global feature importance rankings and dependence plots for each predictor, providing interpretable insights into the drivers of vegetation patterns. The SHAP value for feature j is defined as [39]:

$$\phi_j(v) = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!\left(|N| - |S| - 1\right)!}{|N|!}\left[v\left(S \cup \{j\}\right) - v(S)\right] \quad (7)$$

where, N is the set of all features

S is a subset of features excluding feature j

v(S) is the model trained on feature subset S.

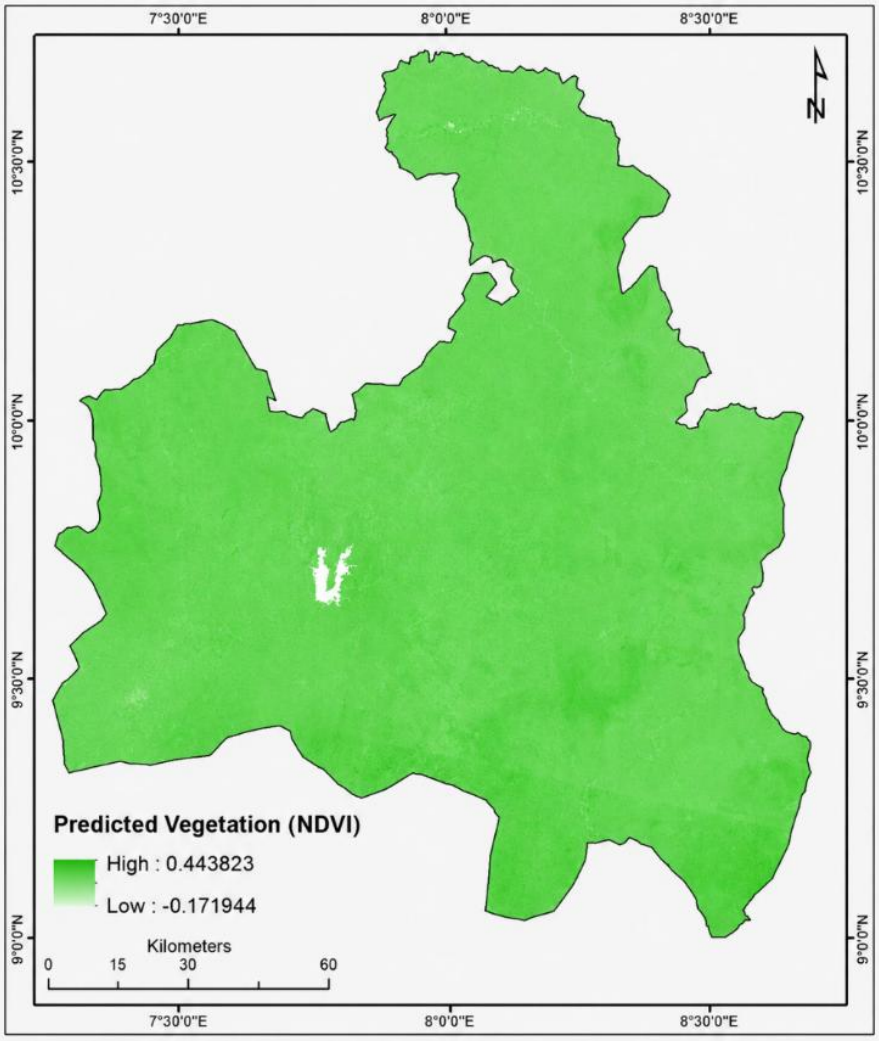
v(S U{j}) - v(S) isolates the marginal contribution that feature j adds to that specific sub-coalition.

JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 DOI: https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026
www.journals.unizik.edu.ng/jsis

3.0 Results and Discussion

3.1 Performance Metrics

The XGBoost model was generated and the results are predicted vegetation map and the scatter plots of the performance metrics are shown in Figures 3 and 4 respectively. During training, the model achieved a very high R^2 of 0.9978, with RMSE = 0.0022 and MAE = 0.0016, indicating that nearly all variability in NDVI was captured. Spatial cross-validation yielded R^2 = 0.9054, RMSE = 0.0142, and MAE = 0.0100, demonstrating strong generalization across heterogeneous landscapes despite minor overfitting to local spatial patterns. Independent testing further confirmed reliability with R^2 = 0.9523, RMSE = 0.0101, and MAE = 0.0076. These results indicate that XGBoost effectively models complex nonlinear relationships among hydrological, climatic, topographic, and anthropogenic drivers of vegetation, consistent with [29], who reported superior XGBoost performance in forest degradation modelling due to its ability to capture intricate environmental interactions.



JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 DOI: https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026
www.journals.unizik.edu.ng/jsis

Figure 3: Predicted Vegetation

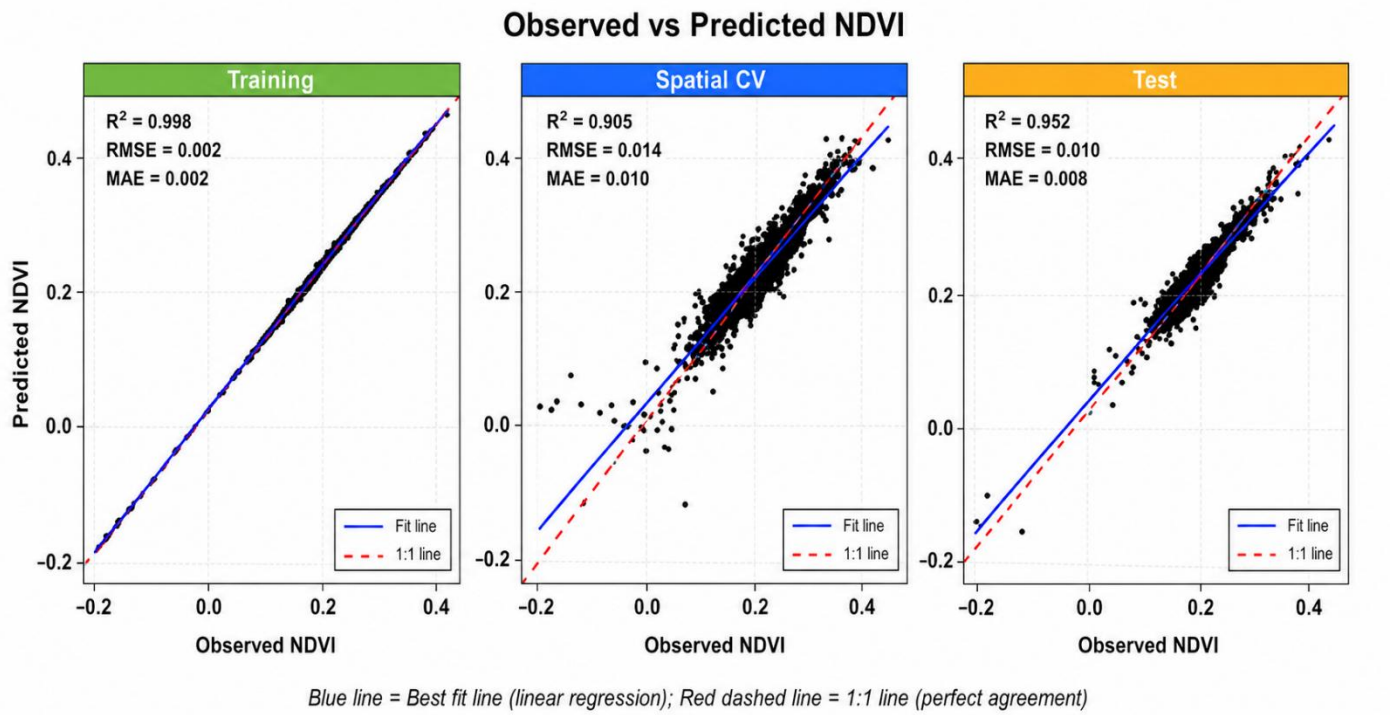


Figure 4: Scatter Plots for Training, Spatial CV and Test Datasets

3.2 Global SHAP Importance Results

The global SHAP importance of each predictor, representing the average impact of each variable on the XGBoost model's predictions, was calculated using the fastshap package in R and is presented in Table 2.

Table 2: Mean Absolute SHAP Values for XGBoost

Feature	Mean Absolute SHAP Value	Percentage
NDWI	0.0217	40.56
NDBI	0.0174	32.52
Evapotranspiration	0.0039	7.29
Population Density	0.0033	6.17
Temperature	0.0017	3.18
Rainfall	0.0018	3.36
Slope	0.0021	3.93
Elevation	0.0016	2.99
Total	0.0535	100.00

The SHAP importance results in Table 2, indicate that NDWI was the most influential predictor, with a mean absolute SHAP value of 0.0217 (40.56%), emphasizing the critical role of surface water and soil moisture in supporting vegetation greenness. NDBI was the second most influential feature (0.0174, 32.52%), indicating

that urbanization and land-use conversion negatively affect NDVI. Evapotranspiration (0.0039, 7.29%) and population density (0.0033, 6.17%) contributed moderately, reflecting ecosystem water use and human pressure on vegetation. Climatic variables such as temperature (0.0017, 3.18%) and rainfall (0.0018, 3.36%) had smaller effects, suggesting that their influence is largely mediated through hydrological variables, while topographic factors, slope (0.0021, 3.93%) and elevation (0.0016, 2.99%), exert minor but localized effects. These findings are consistent with [22] and [24], who emphasized the strong control of hydrological conditions on vegetation, while human activities and terrain modulate local NDVI patterns.

3.3 SHAP Dependance Plots

SHAP dependence plots in Figure 5 reveal predictor interactions and effect directions, with the NDWI plot showing a strong positive relationship with NDVI, indicating that areas with higher surface moisture substantially enhance vegetation greenness. In contrast, NDBI exhibits a negative relationship, indicating vegetation loss with increasing built-up area, consistent with urbanization-driven degradation reported by [21]. Climatic variables such as temperature and rainfall display more subtle nonlinear effects, suggesting context-dependent responses, in agreement with [23], who observed complex temperature-vegetation interactions in Jilin Province. Topographic variables, including slope and elevation, exert localized influences on NDVI, reflecting the role of terrain in soil moisture retention and vegetation distribution, as also noted by [22]. Evapotranspiration shows a positive association with NDVI, indicating that higher water use by vegetation corresponds to higher predicted greenness, while population density shows a modest negative effect, highlighting the indirect impact of human pressure.

JOURNAL OF SPATIAL INFORMATION SCIENCES ISSN: 2354-3361
VOL. 3, ISSUE 1, PP 259–276, 2026 DOI: https://doi.org/10.5281/zenodo.20764537
PUBLISHED 19-06-2026
www.journals.unizik.edu.ng/jsis

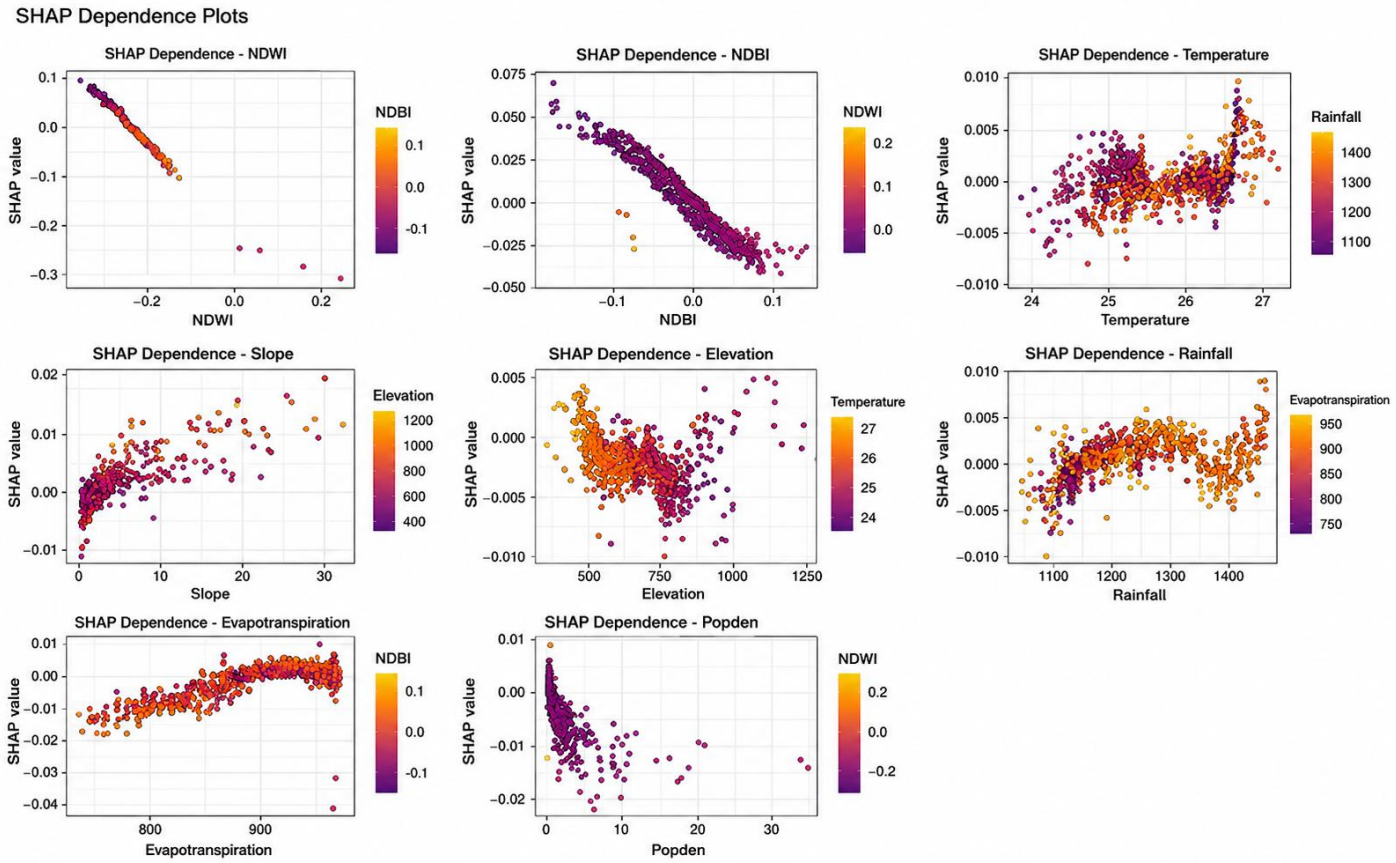


Figure 5: SHAP Dependence Plots

Conclusion

This study shows that vegetation dynamics in Southern Kaduna, Nigeria, are mainly controlled by hydrological and land-surface factors, with NDWI being the most influential predictor of NDVI. Urbanization and land-use changes, captured by NDBI, also significantly affect vegetation, while climatic and topographic variables influence patterns locally. The XGBoost model demonstrated excellent predictive performance across training, spatial cross-validation, and independent testing, with consistently low RMSE and MAE values. SHAP analysis provided interpretable insights, revealing that vegetation greenness is highly sensitive to moisture availability and negatively impacted by urban development.

Based on these findings, land management should focus on conserving areas with high NDWI to maintain soil moisture and ecosystem health. Urban growth should be carefully monitored and regulated to minimize its impact on vegetation, especially in peri-urban zones. Water resources, including natural water bodies and wetlands, should be preserved or restored to support vegetation productivity. The SHAP-informed model outputs can guide targeted monitoring and interventions in areas most sensitive to hydrological or

anthropogenic pressures. Integrating these results into land-use policies and sustainable development planning can balance ecological conservation with urban expansion, while future studies should examine seasonal NDVI variations and finer-scale climate variables to improve understanding of short-term vegetation responses.

References

[1] Miralles D., De Arellano V., McVicar T. & Mahecha, M. (2024). *Vegetation–climate feedbacks across scales*. Annals of the New York Academy of Sciences, 1544, 27–41. https://doi.org/10.1111/nyas.15286

[2] Li W., Duveiller G., Wieneke S., Forkel M., Gentine P., Reichstein M., Niu S., Migliavacca M. & Orth R. (2024). *Regulation of the global carbon and water cycles through vegetation structural and physiological dynamics*. Environmental Research Letters, 19. https://doi.org/10.1088/1748-9326/ad5858

[3] Chen J., Shao Z., Deng X., Huang X. & Dang C. (2023). *Vegetation as the catalyst for water circulation on global terrestrial ecosystem*. The Science of the Total Environment, 165071. https://doi.org/10.1016/j.scitotenv.2023.165071

[4] Zhang G. & Wu G. (2024). Factors influencing carbon and water use efficiency in changing environments. Frontiers in Environmental Science. https://doi.org/10.3389/fenvs.2024.1481082

[5] Sun G., Li L., Li J., Liu C., Wu Y.-P., Gao S., Wang Z. & Feng G.-L. (2022). *Impacts of climate change on vegetation pattern: Mathematical modeling and data analysis*. Physics of Life Reviews, 43, 239–270. https://doi.org/10.1016/j.plrev.2022.09.005

[6] Chen Y., Zhang T.-B., Zhu X., Yi G.-H., Li J.-J., Bie X.-J., Hu J. & Liu X. (2024). *Quantitatively analyzing the driving factors of vegetation change in China: Climate change and human activities*. Ecol. Informatics, 82, 102667. https://doi.org/10.1016/j.ecoinf.2024.102667

[7] Overpeck J. & Breshears D. (2021). *The growing challenge of vegetation change*. Science, 372, 786–787. https://doi.org/10.1126/science.abi9902

[8] Gong Z., Ge W., Guo J. & Liu J. (2024). *Satellite remote sensing of vegetation phenology: Progress, challenges, and opportunities*. ISPRS Journal of Photogrammetry and Remote Sensing. https://doi.org/10.1016/j.isprsjprs.2024.08.011

[9] Yan K., Gao S., Yan G., Xuanlong Chen X., Zhu P., Li J., Gao S., Gastellu-Etchegorry J., Myneni R. & Wang Q. (2025). *A global systematic review of the remote sensing vegetation indices*. Int. J. Appl. Earth Obs. Geoinformation, 139, 104560. https://doi.org/10.1016/j.jag.2025.104560

[10] Mutanga O., Masenyama A. & Sibanda M. (2023). *Spectral saturation in the remote sensing of high-density vegetation traits: A systematic review of progress, challenges, and prospects*. ISPRS Journal of Photogrammetry and Remote Sensing. https://doi.org/10.1016/j.isprsjprs.2023.03.010

[11] Jiménez R., Lane K., Hutyra L. & Fabian M. (2022). *Spatial resolution of Normalized Difference Vegetation Index and greenness exposure misclassification in an urban cohort*. Journal of Exposure Science & Environmental Epidemiology, 32, 213–222. https://doi.org/10.1038/s41370-022-00409-w

[12] Zhao Q. & Qu Y. (2024). *The Retrieval of Ground NDVI (Normalized Difference Vegetation Index) Data Consistent with Remote-Sensing Observations*. Remote Sensing, 16(7), 1212. https://doi.org/10.3390/rs16071212

[13] Dapke P. P., Nagare S. M., Quadri S. A., Bandal S. B., Gaikwad R. M. & Baheti M. R. (2025). *Seasonal Analysis of Vegetation, Moisture, Urbanization, and Land Surface Temperature (LST) Using NDVI, NDMI, NDWI, and NDBI Indices: A Case Study of Sillod, Maharashtra*. 2025 International Conference on Computational, Communication and Information Technology (ICCCIT), 753–760. https://doi.org/10.1109/ICCCIT62592.2025.10928110

[14] Taiwo B. E., Kafy A.- Al, Samuel A. A., Rahaman Z. A., Ayowole O. E., Shahrier M., Duti B. M., Rahman M. T., Peter O. T. & Abosede O. O. (2023). *Monitoring and predicting the influences of land use/land cover change on cropland characteristics and drought severity using remote sensing techniques*. Environmental and Sustainability Indicators, 18, 100248. https://doi.org/10.1016/j.indic.2023.100248

[15] Parada-Molina P. C., Cerdán-Cabrera C., Cervantes-Pérez J., Barradas V. & Ortiz-Ceballos G. (2025*). Impact of climate on water status, growth, yield, and phenology of coffee (Coffea arabica) plants in the central region of the state of Veracruz, Mexico*. PLOS One, 20. https://doi.org/10.1371/journal.pone.0319670

[16] Wang H., Yan S., Ciais P., Wigneron J., Liu L., Li Y., Fu Z., Hongliang Liang Z., Wei F., Wang Y. & Li S. (2022). *Exploring complex water stress–gross primary production relationships: Impact of climatic drivers, main effects, and interactive effects*. Global Change Biology, 28, 4110–4123. https://doi.org/10.1111/gcb.16201

[17] Wang H., Guan H., Xu X., Gao L., Gutiérrez-Jurado H. & Simmons C. (2024). *Topographic regulations on ecohydrological dynamics in a montane forest catchment and the implications for plant adaptation to environment*. Journal of Hydrology. https://doi.org/10.1016/j.jhydrol.2024.131412

[18] Madugundu R., Al-Gaadi K., Tola E., Zeyada A., Alameen A., Edrris M., Edrees H. & Mahjoop O. (2022). *Impact of Field Topography and Soil Characteristics on the Productivity of Alfalfa and Rhodes Grass: RTK-GPS Survey and GIS Approach*. Agronomy. https://doi.org/10.3390/agronomy12122918

[19 Zhang C.-F., Wang Z.-Y., Wang Q. & Yang C. (2025). *Interaction of population density and slope will exacerbate spatiotemporal changes in land use and landscape patterns in mountain city*. Scientific Reports, 15. https://doi.org/10.1038/s41598-025-87550-2

[20] Hunde F. M., Benti A. A. & Kapula T. J. (2026). *Impact of population pressure on forest resources depletion in Yayo coffee forest Biosphere Reserve, Southwest Ethiopia*. PLOS One, 21. https://doi.org/10.1371/journal.pone.0324407

[21] Yang L., Shen F., Zhang L., Cai Y., Yi F. & Zhou C. (2021). *Quantifying influences of natural and anthropogenic factors on vegetation changes using structural equation modeling: A case study in Jiangsu Province, China*. Journal of Cleaner Production, 280, 124330. https://doi.org/10.1016/j.jclepro.2020.124330

[22] Zhang Y., He Y., Li Y. & Jia L. (2022). *Spatiotemporal variation and driving forces of NDVI from 1982 to 2015 in the Qinba Mountains, China*. Environmental Science and Pollution Research, 29, 52277–52288. https://doi.org/10.1007/s11356-022-19502-6

[23] Ren Y., Zhang F., Zhao C. & Cheng, Z. (2023). *Attribution of climate change and human activities to*

vegetation NDVI in Jilin Province, China during 1998–2020. Ecological Indicators. https://doi.org/10.1016/j.ecolind.2023.110415

[24] Mehmood K., Anees S., Rehman A., Pan S., Tariq A., Zubair M., Liu Q., Rabbi F., Khan K. A. & Luo M. (2024). *Exploring spatiotemporal dynamics of NDVI and climate-driven responses in ecosystems: Insights for sustainable management and climate resilience*. Ecol. Informatics, 80, 102532. https://doi.org/10.1016/j.ecoinf.2024.102532

[25] Liang D., Frederick D., Lledo E., Rosenfield N., Berardi V., Linstead E. & Maoz U. (2022). *Examining the utility of nonlinear machine learning approaches versus linear regression for predicting body image outcomes: The U.S. Body Project I*. Body Image, 41, 32–45. https://doi.org/10.1016/j.bodyim.2022.01.013

[26] Kumar A., Devi K., Katoch M., Sharma A., Angra P., Bose S., Farhan M. & Singh, P. (2023). *Comparison of Regression Techniques and Integration Models Using Machine Learning*. 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), 6, 1349–1353. https://doi.org/10.1109/ic3i59117.2023.10397945

[27] Ali Z. A., Abduljabbar Z., Tahir H., Sallow A. B. & Almufti, S. (2023*). eXtreme Gradient Boosting Algorithm with Machine Learning: a Review*. Academic Journal of Nawroz University. https://doi.org/10.25007/ajnu.v12n2a1612

[28] Niazkar M., Menapace A., Brentan B., Piraei R., Jimenez D., Dhawan P. & Righetti, M. (2024). *Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018-May 2023)*. Environ. Model. Softw., 174, 105971. https://doi.org/10.1016/j.envsoft.2024.105971

[29] Abdullah M. H. & Mustafa Y. T. (2025). *Machine Learning and SHAP-Based Analysis of Deforestation and Forest Degradation Dynamics Along the Iraq–Turkey Border*. Earth, 6(2), 49. https://doi.org/10.3390/earth6020049

[30] Yang C., Yao P., Wang Q., Wang S., Xing D., Wang Y. & Zhang J. (2026). *XGBoost-Based Susceptibility Model Exhibits High Accuracy and Robustness in Plateau Forest Fire Prediction*. Forests. https://doi.org/10.3390/f17010074

[31] Zhang J., Ma X., Zhang J., Sun D., Zhou X., Mi C. & Wen H. (2023). *Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model*. Journal of Environmental Management, 332. https://doi.org/10.1016/j.jenvman.2023.117357

[32] Adebajo A., Atobatele A., Bello O. W. & Dele-Dada M. (2025). *Gender, Climate Change and Herder Farmer Conflicts: How Far Has SDGs 1,2,3,13,15 & 16 Addressed the Crises in Southern Kaduna, Nigeria?* Journal of Lifestyle and SDGs Review. https://doi.org/10.47172/2965-730x.sdgsreview.v5.n03.pe02737

[33] Musa I., Olusola A. T. & Magaji, S. (2025). *Effects of Climate Change on Environmental Security among Vulnerable Groups in Zango Kataf Local Government Area of Kaduna State*. Loka: Journal of Environmental Sciences, 2(1), 169–191. https://doi.org/10.38142/ljes.v2i1.251

[34] Garba H. & Abubakar Z. (2023). *Effect of Slope and Runoff Trends on the Hydrological Response of River Kaduna*. ABUAD Journal of Engineering Research and Development (AJERD), 6(2), 183–191. https://doi.org/10.53982/ajerd.2023.0602.18-j

[35] Daramola J., Adepehin E., Ekhwan T., Choy L., Mokhtar J. & Tabiti T. (2022). *Impacts of Land-Use Change, Associated Land-Use Area and Runoff on Watershed Sediment Yield: Implications from the Kaduna Watershed*. Water. https://doi.org/10.3390/w14030325

[36] Hussein S. J. & Naji R. (2023). *Modelling the Behaviour of Vegetation Indicators (NDVI - NDWI - IPVI) in Busaiya District Using GIS-RS*. IOP Conference Series: Earth and Environmental Science, 1225. https://doi.org/10.1088/1755-1315/1225/1/012013

[37] Mansourmoghaddam M., Rousta I., Ghafarian Malamiri H., Sadeghnejad M., Krzyszczak, J. & Ferreira C. S. S. (2024). *Modeling and Estimating the Land Surface Temperature (LST) Using Remote Sensing and Machine Learning (Case Study: Yazd, Iran)*. Remote Sensing, 16(3), 454. https://doi.org/10.3390/rs16030454

[38] Kushal K. C., Romanko M., Perrault A. & Khanal S. (2024). *On-farm cereal rye biomass estimation using machine learning on images from an unmanned aerial system*. Precision Agriculture, 25, 2198–2225. https://doi.org/10.1007/s11119-024-10162-9

[39] Qin L., Zhu Y., Liu S., Zhang X. & Zhao Y. (2025). *The Shapley Value in Data Science: Advances in Computation, Extensions, and Applications*. Mathematics, 13(10), 1581. https://doi.org/10.3390/math13101581