Check for updates

Ecology and Evolution

WILEY

Ecology and Evolution Open Access

RESEARCH ARTICLE OPEN ACCESS

Machine Learning and Spatio Temporal Analysis for Assessing Ecological Impacts of the Billion Tree Afforestation Project

Kaleem Mehmood[1,2,3] | Shoaib Ahmad Anees[4] | Sultan Muhammad[3] | Fahad Shahzad[5] | Qijing Liu[1,2] | Waseem Razzaq Khan[6] | Mansour Shrahili[7] | Mohammad Javed Ansari[8] | Timothy Dube[9]

[1]College of Forestry, Beijing Forestry University, Beijing, China | [2]Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing, China | [3]Institute of Forest Science, University of Swat, Swat, Pakistan | [4]Department of Forestry, The University of Agriculture, Dera Ismail Khan, Pakistan | [5]Precision Forestry Key Laboratory of Beijing, Beijing Forestry University, Beijing, China | [6]Department of Forestry Science and Biodiversity, Faculty of Forestry and Environment, Universiti Putra Malaysia, Serdang, Malaysia | [7]Department of Statistics and Operations Research, College of Science, King Saud University, Riyadh, Saudi Arabia | [8]Department of Botany, Hindu College Moradabad (Mahatma Jyotiba Phule Rohilkhand University Bareilly), Moradabad, India | [9]Institute for Water Studies, Faculty of Science, University of the Western Cape, Cape Town, South Africa

Correspondence: Shoaib Ahmad Anees (anees.shoaib@gmail.com) | Waseem Razzaq Khan (khanwaseem@upm.edu.my)

Received: 29 September 2024 | **Revised:** 28 November 2024 | **Accepted:** 3 December 2024

Funding: This research was funded by the Universiti Putra Malaysia, Malaysia, Vote No. 9750500.

Keywords: afforestation | land-use change | machine learning | NDVI | remote sensing | Sentinel-2

ABSTRACT

This study evaluates the Billion Tree Afforestation Project (BTAP) in Pakistan's Khyber Pakhtunkhwa (KPK) province using remote sensing and machine learning. Applying Random Forest (RF) classification to Sentinel-2 imagery, we observed an increase in tree cover from 25.02% in 2015 to 29.99% in 2023 and a decrease in barren land from 20.64% to 16.81%, with an accuracy above 85%. Hotspot and spatial clustering analyses revealed significant vegetation recovery, with high-confidence hotspots rising from 36.76% to 42.56%. A predictive model for the Normalized Difference Vegetation Index (NDVI), supported by SHAP analysis, identified soil moisture and precipitation as primary drivers of vegetation growth, with the ANN model achieving an R^2 of 0.8556 and an RMSE of 0.0607 on the testing dataset. These results demonstrate the effectiveness of integrating machine learning with remote sensing as a framework to support data-driven afforestation efforts and inform sustainable environmental management practices.

1 | Introduction

Afforestation is important to global climate change mitigation, land rehabilitation, and biodiversity enhancement strategies. It has recently been announced that the BTAP Project in Pakistan is one of seven ambitious global initiatives and policy tools emphasizing scaling up forest landscape restoration (Kamal, Ali, and Yingjie 2018; Ullah et al. 2020). This afforestation activity removes large volumes of CO_2 from the atmosphere, which is crucial in combating global warming (Haider et al. 2017; Jallat et al. 2021; Khan et al. 2020). It also enhances ecosystem services related to carbon sequestration and wildlife conservation (Chen and Zhang 2023; Wang et al. 2022; Wang et al. 2024). Remote sensing technologies have dramatically changed how afforestation programs are monitored and evaluated, predominantly through high-resolution satellite imagery (Shawky

Kaleem Mehmood and Shoaib Ahmad Anees contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Ecology and Evolution* published by John Wiley & Sons Ltd.

et al. 2023; Zha et al. 2020). Such technological advancement now allows detailed global land observation, giving open access to imagery of all spatial scales from local to global (Galidaki et al. 2017; Zhang et al. 2024a, 2024b). The technology enables large-scale, comprehensive analyses to increase the extent and resolution of studies to a new level (Colomina and Molina 2014; Huang et al. 2018).

Furthermore, when integrated with cloud computing platforms such as Google Earth Engine (GEE), giving advanced extensive data processing capabilities, remote sensing becomes an even more potent tool for the evaluation and monitoring of afforestation activities on a global scale (Anees et al. 2022a; Mehmood et al. 2024b). For instance, two panels of Landsat program-derived, time series surface reflectance data from satellites (Fassnacht et al. 2019; Furniss et al. 2020), spanning four decades, have added much insight into land-use and forest cover changes on the surface of the Earth (Smith et al. 2021). They are beneficial data to test afforestation and deforestation criteria and serve as an example of metrics in gauging the success or failure of projects like the BTAP. Advanced classification algorithms, such as RF, can successfully manage and analyze this vast reference database through platforms such as GEE (Cheng et al. 2022; Qasimi et al. 2023). RF is beneficial in this regard, as over the years, it has shown very high accuracy and, at the same time, meager rates of overfitting due to its ability to model intricate, non-Gaussian relationships among variables (Ma et al. 2020; Mehmood et al. 2024e; Pang, Chang, and Chen 2022).

The Landsat program has been acquiring multispectral imagery for the last several decades (Irons, Dwyer, and Barsi 2012), which has led to spatially continuous and extensive data records of long-term observations of land surface reflectance (Ouchra, Belangour, and Erraissi 2023; USGS 2022). Such records are essential for assessing changes in land use and modifications in the cover of forests from the past. They are irreplaceable in estimating the impacts of afforestation projects like the BTAP. This is a significant strength of the GEE platform: it can process and analyze this massive amount of data through advanced classification methods, such as RF. The RF algorithm, an up-to-date solution for better predictive power and generalization capacity, provides a good measure for preventing overfitting and increases accuracy in environmental assessments.

Remote sensing and machine learning have been increasingly highlighted in studies to monitor large-scale afforestation, mainly through commercial plantations (Kupssinskü et al. 2020; Zheng, Abd-elrahman, and Whitaker 2021; Anees et al. 2024a). The extensive integration of GEE and Landsat data has widely been applied to assess changes in forest cover and the sustainability of afforestation initiatives. Among different machine-learning algorithms, RF and ANN performed well for land-cover classification and afforestation monitoring (Mehmood et al. 2024a; Shahzad et al. 2024). RF has demonstrated high success in these types of studies at large scale and satisfactory accuracy, as it is typically able to separate well between different land-cover types, with strong lines of evidence from numerous empirical case examples worldwide. In addition, the Random Forest's ability to model complex and nonlinearities in ecological data provides more efficient information for assessing afforestation programs (Hussain et al. 2024b).

In addition, combining remote sensing methods with biodiversity assessments has significantly increased biodiversity, particularly in degraded ecosystems (Zhang et al. 2024a, 2024b; Chen et al. 2023; Xie et al. 2023; Jiang et al. 2023). Evidence from numerous studies indicates that combining vegetation indices, such as NDVI, with machine-learning models can effectively assess the ecological impacts of afforestation and support more strategic conservation efforts (Brieva et al. 2023; Mehmood et al. 2024c; Roy 2021; Anees et al. 2024a). Sustained, long-term monitoring is increasingly recognized as crucial, as these processes often become observable only over decades. Scholars advocate for using temporal datasets to detect the frequently gradual or delayed impacts of afforestation initiatives, which are usually underestimated in short-term studies (Nazir et al. 2019; Ullah et al. 2023; Zheng et al. 2024).

The applications of remote sensing to massive afforestation efforts face significant challenges and offer great opportunities, but there is a need for continuous innovation and complementary data sources. Coupling in situ validation with the integration of ecological models and remote sensing datasets enables long-term prediction of ecosystem processes with a high level of reliability, together with an enhancement in accuracy when undertaking assessments (Doelman et al. 2020). Besides, proper monitoring of the effects of afforestation on the storage of soil carbon, diversity in microorganisms, and general health status of an ecosystem is essential in the provision of accurate information, as noted by studies focusing on the impacts of afforestation in different regions (Burke et al. 2023; Cao et al. 2011; Kong et al. 2022; Nave et al. 2013).

Approaches to integrating remote sensing data with biodiversity assessments have been demonstrated to improve our understanding of ecosystems massively (Bunce et al. 2014). Researchers have shown the efficacy of coupling vegetation indices, such as the NDVI, with machine-learning methods for characterizing the ecological impacts of afforestation, thereby facilitating better conservation planning (Xiao, Xiao, and Sun 2020; Yao, Xiao, and Ma 2021; Anees et al. 2024a, 2024b). Regional-based strategies benefit from more grounded policy support, while location specific investments are customized to local conditions (João et al. 2018).

The need for long-term, consistent monitoring has gained recognition as necessary to capture the dynamic nature of forest ecosystems over time. Scholars propose the addition of temporal datasets to detect delayed or typically slow effects, often underestimated in brief-term studies of initiatives like reafforestation (Hao et al. 2022; Zhang et al. 2022). Although implementing remote sensing within large-scale afforestation projects is filled with limitations and potential, continuous change through multisensor data sourcing is essential. The coupling of ecological models with remote sensing data can provide improved predictions for long-term outcomes, as this approach is validated by ground surveys, thereby increasing precision, reliability, and predictive accuracy (Pan, Harrou, and Sun 2023; Pan et al. 2023; Wang and Fan 2021).

While considerable progress has been made in applying remote sensing and machine learning to monitor afforestation efforts, existing work often lacks detailed long-term examination of

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

localized ecological outcomes (Anees et al. 2024a). Additionally, advanced techniques such as ANN and spatial analysis methods like hotspot analysis and Moran's *I* are not well exploited in evaluating vegetation dynamics associated with afforestation projects. This study aims to (1) review the ecological impact on land-use and land-cover (LULC) changes in response to the BTAP from 2015 to 2023, mainly focusing on rapid area changes using Sentinel-2 imagery combined with the RF algorithm in the loess lands of Khyber Pakhtunkhwa (KPK) province. The research also measured major land-cover type changes, such as tree cover, barren lands, and shrublands. (2) Conduct buffer analysis, hotspot analysis, and Moran's *I* spatial autocorrelation within the BTAP plantation buffer zones to study vegetation change over time. (3) Utilize machine-learning methods, particularly ANNs, to forecast NDVI from biotic and abiotic factors, incorporating SHAP analysis to determine the primary drivers of vegetation change. This study explored how to effectively study large-scale afforestation initiatives while drawing out key lessons applicable to forest cover and ecological restoration efforts more broadly.

2 | Methods and Materials

2.1 | Study Area

The study was conducted in the Khyber Pakhtunkhwa (KPK) province of Pakistan, located between latitudes 33.0°N and 36.0°N and longitudes 70.5°E and 73.5°E (Ahmed 2011; Mohiuddin 2007). KPK is a province with substantial geographic and ecological variation, making it an ideal location for the BTAP. Situated in the northwestern region of Pakistan, KPK is bordered by the Hindu Kush mountains to the north and the Indus River to the south, encompassing a range of climates from subtropical plains in the south to temperate and alpine zones in the north (Sohail et al. 2023) (Figure 1). The study focused on nine districts within KPK: Bajaur, Buner, Dir Lower, Dir Upper, Kalam, Khurum, Khyber, Malakand, and Mohmand. These districts were selected due to their critical role in the BTAP, which aims to rehabilitate degraded lands and enhance forest cover across the province.

The climate in KPK varies significantly across these districts. Southern regions, such as Buner and Khyber, experience hot summers and mild winters, while northern areas, including Kalam and Dir Upper, are characterized by cold winters with significant snowfall (Ali, Khan, and Ahmad 2018; Muhammad 2023). This climatic variability, combined with the province's complex topography, ranging from lowland plains to high mountain peaks, significantly influences the types of vegetation that thrive in the region (Bacha et al. 2021; Ul-Haq et al. 2019; Anees et al. 2024b). This study leverages the diverse landscape and climate to investigate the effects of afforestation initiatives, like the BTAP, under different environmental conditions. The study sites were chosen based on their ecological significance, existing forest cover, and strategic importance within the BTAP framework. Additionally, the socioeconomic context of KPK, where local communities depend heavily on natural resources for their livelihoods, must be considered (Khan et al. 2019). The success of the BTAP is contingent on not only achieving ecological objectives but also promoting sustainable land-use practices that benefit these communities.

2.2 | Dataset and Preprocessing

Our study focuses on 344 plantation sites located in nine districts of KPK, Pakistan, which are strategically selected for the assessment of the BTAP. These sites are essential reference points for evaluating the project's impact on land-cover changes. The geographical distribution of these plantation sites is categorized by district (Table 1).

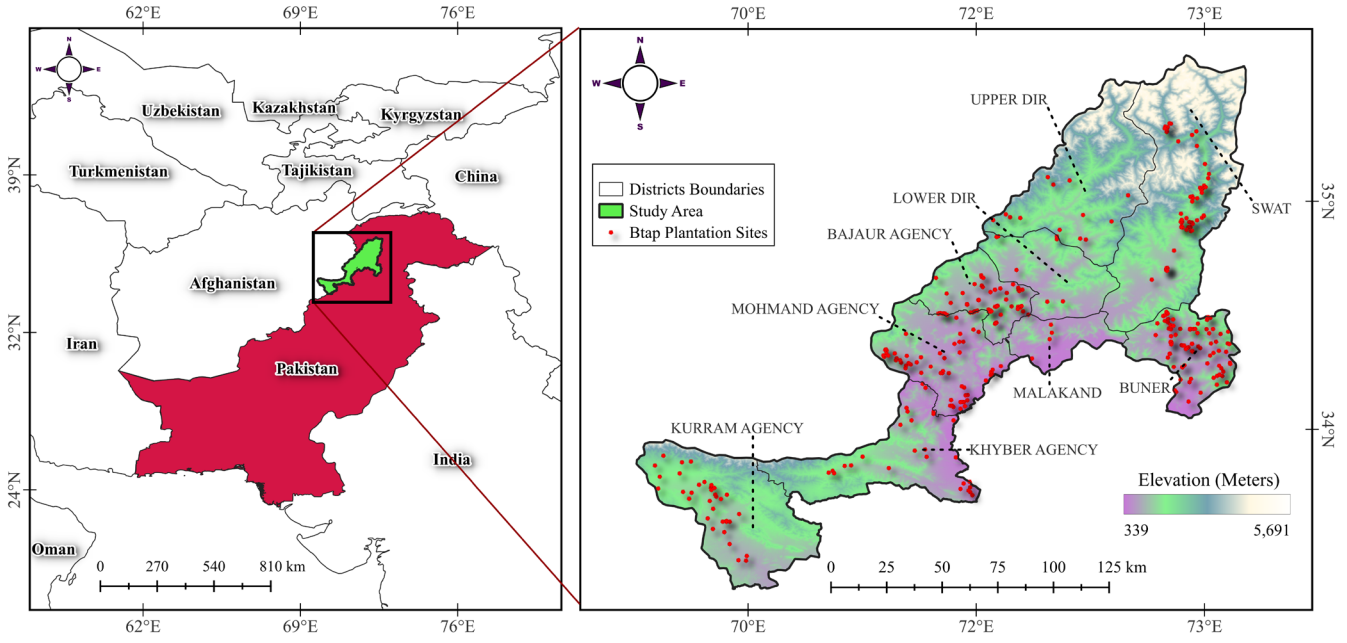

FIGURE 1 | Map of the study area.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

This dataset offers an in-depth view of the plantation sites, enabling a thorough evaluation of the BTAP and its impact on ecological restoration and sustainable land use across KPK. The spatial variability of the sites across different districts ensures a comprehensive assessment of the project's impact under diverse regional ecological and geographical conditions.

This study used Sentinel-2 satellite imagery to assess LULC changes from 2015 to 2023, specifically focusing on 2015, 2019, and 2023. The imagery was acquired for October and November, a post-monsoon period characterized by significant vegetation growth and minimal cloud cover (Qiu et al. 2019; Valero et al. 2021). These periods were selected to exploit optimal conditions for clear, cloud-free habitats necessary for accurate LULC mapping and vegetation assessment (Yan et al. 2022). Sentinel-2 imagery was preprocessed in the GEE environment using the Sen2Cor algorithm to create cloud and shadow masks, ensuring the selection of cloud-free pixels (Bui et al. 2022; Li et al. 2018; Xu, Li, and Chen 2022). To further enhance image quality, monthly composites for October and November were generated using median reflectance values, effectively mitigating residual clouds' impact or atmospheric disturbances (Castaldi et al. 2023; Rumora, Miler, and Medak 2020). Additionally, NDVI rasters for the respective years were derived from these processed images, providing a solid foundation for further analysis. The detailed schematic diagram of the research methodology is explained in Figure 2.

TABLE 1 | Distribution of plantation sites across districts in Khyber Pakhtunkhwa for the BTAP afforestation project (2015–2023).

S.No.	District	Number of plantation sites
1	Bajaur	57
2	Buner	84
3	Dir Lower (DirL)	5
4	Dir Upper (DirU)	15
5	Kalam	60
6	Khurum	34
7	Khyber	29
8	Malakand	4
9	Mohmand	56

Buffer and autocorrelation analyses were performed to assess vegetation health and spatial distribution patterns around the plantation locations using these NDVI rasters as inputs. Alongside NDVI, a wide range of environmental variables, including temperature (temp), precipitation (Ppt), solar radiation (SR), soil moisture (SM), elevation, slope, aspect, and population density (POP) at the nearest second-level administrative unit level to site locations, were considered as drivers for predictive NDVI models using machine-learning approaches (Anees et al. 2024a) (Table 2). Utilizing the full-scale computational

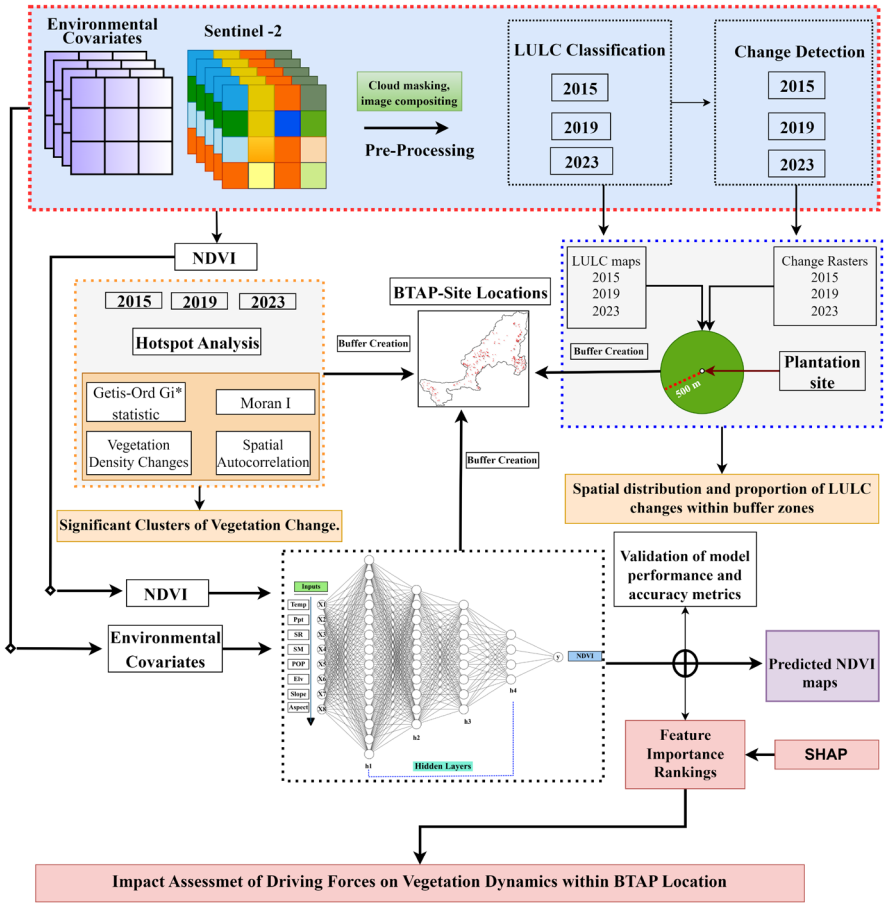

FIGURE 2 | Schematic diagram of the research methodology.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Data sources, resolutions, and descriptions of environmental variables used in the study.

Data type	Source	Spatial resolution	Temporal resolution	Description	References
Sentinel-2	European Space Agency (ESA)	10 m	5 days	Bands: Blue, green, red, NIR, SWIR	Tamiminia et al. (2024), Wang et al. (2019)
NDVI	ESA	10 m	5 days	NDVI derived from Sentinel-2	Karlsen et al. (2021), Yang et al. (2022)
Temp	ERA5	9 km	Hourly/monthly	Surface temperature data	Gomis-Cebolla et al. (2023), Muñoz-Sabater et al. (2021)
Ppt	CHIRPS	5 km	Pentad	Climate Hazards Group InfraRed Precipitation with Station data	de Sousa et al. (2020), Vargas Godoy and Markonis (2023)
SR	TerraClimate: University of Idaho	4.5 m	Monthly	Downward surface shortwave radiation, used for climate and water balance studies	Abatzoglou et al. (2018), Abdi (2019), TerraClimate (2022), University of Idaho (2019)
SM				Soil moisture, derived using a one-dimensional soil water balance model	
POP	World Pop	100 m	Annually	Gridded population density data	Gaughan et al. (2013), Stevens et al. (2015)
Topography (DEM)	ALOS World 3D (AW3D30)	30 m	—	Digital Elevation Model (DEM) data	JAXA (2019), Li and Zhao (2018)
Slope		30 m	—	Slope calculated from DEM	
Aspect		30 m	—	Aspect calculated from DEM	

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

capabilities of GEE, the processing workflow resulted in high-quality, low-cloud cover composites, ideally suited for LULC classification and change detection analyses. The carefully compiled datasets offer a strong and dependable foundation for assessing the ecological impacts of the BTAP plantation sites in Pakistan.

2.3 | LULC Classification Using the Random Forest Algorithm

To classify land use and land cover (LULC) for the years 2015, 2019, and 2023, we employed the Random Forest (RF) algorithm, a robust and widely recognized machine-learning technique known for its exceptional accuracy in remote sensing applications (Badshah et al. 2024; Kumar and Agrawal 2023). The classification process was designed to categorize the landscape into seven distinct classes: Trees, Grassland, Swamp Vegetation, Arable Land, Shrubland, Built-up Areas, and Barren Land. Training samples for each LULC class were meticulously collected from high-resolution satellite imagery and validated against ground truth data where available (Hussain et al. 2024a). These samples were carefully distributed across the study area to ensure each class's balanced and representative dataset.

The spectral bands from Sentinel-2 imagery, including visible, near-infrared, and short-wave infrared bands, were utilized as input features for the classification. In addition to these bands, vegetation indices such as the NDVI were calculated and incorporated to enhance the discrimination of vegetation related classes (Marino 2023; Marlina 2022; Anees et al. 2024b). The RF classifier was trained using thoroughly gathered training samples (Lange et al. 2017). The RF algorithm operates by constructing an ensemble of decision trees during the training phase, subsequently deriving its classification output based on the mode of the individual trees' predictions (Jun 2021). Key parameters of the RF classifier, including the number of trees and the "mtry" parameter, were systematically optimized to enhance overall classification accuracy. We used 500 decision trees, which optimized the balance between performance and computational efficiency. Additionally, the "mtry" parameter, controlling the number of variables available for splitting at each node, was tuned through a fivefold cross-validation approach. This cross-validation framework provided a robust evaluation process, ensuring that performance metrics were reliable and not biased by any single data partition. This parameter tuning was essential to ensure both model accuracy and stability (Anees et al. 2022b; Khan et al. 2024).

The refined RF model was subsequently applied to preprocessed Sentinel-2 imagery for 2015, 2019, and 2023, producing classified LULC maps for each period. Each pixel was categorized into one of the seven LULC classes based on the model's predictions during this process. Postprocessing of these maps included general cleanup tasks like removing isolated pixels and smoothing out class boundaries to improve the overall consistency and accuracy of the classification results (Gupta, Kanga, and Mishra 2024; Wijaya, Munir, and Utama 2023). An independent set of validation samples was employed to evaluate classification accuracy. To ensure spatial independence and reduce potential overlap with the training data, these validation samples were selected from separate spatial regions. This spatial separation minimizes the risk of data leakage, providing a rigorous and unbiased assessment of model performance (Pomme et al. 2022). Feature importance scores were computed to identify the most influential input variables for LULC classification, with Sentinel-2's near-infrared and short-wave infrared bands ranking among the highest. This analysis provides insight into the factors most critical for distinguishing between LULC classes and supporting model transparency.

To determine the BTAP's impact on land-cover dynamics, a change detection analysis was performed using the LULC maps from 2015, 2019, and 2023. This analysis involved comparing the LULC maps from these time points, focusing on transitions to and from the "Trees" class, which indicate afforestation or deforestation activities (Vujović 2021). A change matrix was developed to quantify the transitions between different LULC classes over time, providing detailed insights into the extent and nature of land-cover changes and identifying areas of significant change (Larbi 2023; Sadhwani, Eldho, and Karmakar 2023).

Spatial analysis was employed to map areas of significant land-cover change, utilizing hotspot analysis and Moran's *I* autocorrelation to detect clusters of change and to evaluate the spatial patterns underlying these dynamics (Anees et al. 2020; Gomes et al. 2021). In addition, buffer analysis was conducted to examine LULC changes within the defined plantation buffer zones (Ziaul Hoque et al. 2022), thereby providing a more nuanced understanding of the localized impacts of the BTAP. The combined classification and change detection analyses yielded a comprehensive knowledge of LULC dynamics across Khyber Pakhtunkhwa, effectively elucidating the influence of the BTAP plantation sites over the study period. These analyses are pivotal in informing future conservation and afforestation strategies in the region, ensuring that efforts are targeted and effective.

2.4 | Buffer Analysis

One of the objectives of this study was to assess the spatial–temporal variations in land-cover types across various districts by analyzing land-cover data within defined buffer zones around specific locations. This analysis was performed for 2015, 2019, and 2023 using high-resolution classified LULC. A spatial point dataset containing particular locations within the districts was used to create buffer zones (Feng et al. 2021). A buffer radius of 500 m was selected to balance ecological and practical considerations, ensuring the inclusion of diverse land-cover types within each buffer (Equation 1). A buffer distance of 500 m was selected for this study to capture the immediate ecological impact zone around the plantation sites (Chapagain et al. 2021; Yana and Rahayu 2017).

Preliminary analysis of multiple buffer sizes (100, 250, 500, and 1000 m) indicated that a 500-m buffer provided the best balance between representing local landscape heterogeneity and capturing the spatial scale of plantation impacts (Horton et al. 2018; Liu et al. 2023). This distance is commonly used in vegetation

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and ecological studies to achieve a balance between detailed local analysis and broader generalization. A 500-m buffer effectively includes the surrounding vegetation and land-cover types likely influenced by plantation activities, as supported by previous studies using buffer ranges of 100 to 1000 m to assess afforestation impacts (Janeczko et al. 2019; Wibowo et al. 2015). Additionally, the 500-m buffer maintains computational feasibility while providing meaningful insights into the direct influence of the plantation efforts.

$$\text{Buffer zone area} = \pi \times (500\ m)^2 = \ = 785{,}398\ m^2 \quad (1)$$

2.4.1 | Spatial Analysis

Land-cover data were extracted from the corresponding LULC for each buffer zone. This involved overlaying the buffer zones on the raster data and calculating the proportion of each land-cover type within each buffer (Wang et al. 2024). The proportion of each land-cover type within each buffer zone was computed (Equation 2)

$$\text{Propostion}_{i,j} = \frac{\text{Area of land} - \text{cover type } i \text{ in buffer } j}{\text{Total area of buffer } j} \quad (2)$$

where i denotes the land-cover type and j denotes the buffer zone.

The proportions of different land-cover types were aggregated by district and year. This aggregation helped us to understand the overall land-cover distribution and changes at the district level over time. The buffer analysis revealed significant improvements in land cover across various districts in KPK from 2015 to 2023, highlighting the success of the BTAP afforestation project.

2.5 | Hotspot Analysis (Getis-Ord Gi* Statistic)

In this study, we conducted a hotspot analysis to identify regions exhibiting significant variations in vegetation density. These regions display distinct positive or negative NDVI anomalies, indicating extensive spatial differences in vegetation. Utilizing remote sensing capabilities, we applied the Getis-Ord Gi* statistic (Getis and Ord 2010, 1992; Ord and Getis 1995), a well-established spatial statistical method, to pinpoint regions with significant vegetation changes. These regions, marked by notable NDVI values, show discernible positive or negative deviations in NDVI patterns. The Getis-Ord Gi* statistic is crucial for locally identifying statistically significant spatial clusters characterized by high (hot-spots) and low NDVI values (cold spots). (Equation 3).

$$G_I^* \frac{\sum_{j=1}^{n} W_{i,j}X_j - \overline{X}\sum_{j-1}^{n} W_{i,j}}{\sqrt{\frac{\left[n\sum_{j=1}^{n} W_{i,j}^2 - \left(\sum_{j=1}^{n} W_{i,j}\right)^2\right]}{n-1}}} \quad (3)$$

wherein x_j delineates the attribute value associated with feature j: w_i, j represents the spatial weight between features i and j, typically derived from their spatial relationship. n is the aggregate count of features.

Furthermore, the mean X and variance S of the attribute values are defined as Equations (4) and (5):

$$\overline{X} = \frac{\sum_{j=1}^{n} W_{i,j}}{n} \quad (4)$$

$$S = \sqrt{\frac{\sum_{J=1}^{N} X_j^2}{n} - (\overline{x})^2} \quad (5)$$

The fundamental characteristic of the Gi* statistic is its ability to calculate the local aggregation of attribute values for a specific feature compared to its neighboring features, contrasting it with the aggregation across all features (Baldo et al. 2023). The presence of a statistically significant cluster is indicated by a notable departure from the expected local sum that exceeds the thresholds of random chance. This mathematical framework allows for a systematic and comprehensive investigation of geographical patterns, enabling researchers to accurately identify regions characterized by significant attribute concentration (Zhou et al. 2023).

Using ArcGIS Pro's Hotspot Analysis (Getis-Ord Gi*) tool, we calculated spatial clustering metrics based on the tool's standardized procedures, which include automatic z-score and p-value calculations and corrections for spatial dependency using fixed distance bands. The Getis-Ord Gi* statistic provides measurements of statistical significance for individual spatial features or regions. Two main metrics are calculated: the Gi* p value, which measures the likelihood, and the Gi* z score, which assesses the standard deviation (Garik 2021). The z score evaluates the level of concentration or dispersion within the features or regions. At the same time, the p value offers a probabilistic assessment to determine if the observed hotspot patterns could result from random spatial distributions. A substantial z score with a small p value indicates a statistically significant hotspot. In contrast, a significantly negative z score with a low p value indicates a statistically substantial cold spot. The size of the z score is directly proportional to the degree of clustering, with larger absolute values indicating more prominent clustering patterns (ESRI 2022, 2013).

Cold spot: Regions with a significant clustering of lower NDVI values, characterized by a Gi* z score less than −1.65. Hotspot: Areas showcasing a substantial aggregation of elevated NDVI values, marked by a Gi* z score greater than 1.65. Neutral areas: Regions that do not exhibit significant spatial correlation, falling within the z score range of −1.65 to 1.65. Confidence criteria were utilized to determine the statistical significance of the detected vegetation zones, with thresholds set at 90%, 95%, and 99%. Regions exhibiting pronounced Gi* z scores (either exceeding 2.58 or below −2.58) and minimal Gi* p values (< 0.01) are classified as "high confidence category," signifying areas with extreme vegetation values at a 99% confidence level. The classification of vegetation zones based on confidence and significance levels is summarized in Table 3.

Inverse distance weighted (IDW) interpolating the Getis-Ord Gi* results generated raster maps of hot and cold spots aligned with Sentinel 2 resolutions (Sanusi et al. 2024). These were subsequently vectorized to produce polygonal representations of

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 | Classification of vegetation hotspots and cold spots based on Getis-Ord Gi statistic with corresponding confidence and significance levels*.

Classification	Confidence threshold	*p*	*z*
High-confidence cold spots	99%	< 0.01	< −2.58
Moderate-confidence cold spots	95%	< 0.05	< −1.96
Low-confidence cold spots	90%	< 0.10	< −1.65
Nonsignificant	Not significant	N/A	−1.65 < *z* score < 1.65
Low-confidence hotspots	90%	< 0.10	> 1.65
Moderate-confidence hotspots	95%	< 0.05	> 1.96
High-confidence hotspots	99%	< 0.01	> 2.58

the vegetation anomalies. Following the identification and classification of vegetation hotspots and cold spots, the subsequent analysis focused on understanding the vegetation properties characteristic of these zones. This is critical due to the substantial contribution of vegetation to regional sustainability and the mitigation of environmental impacts.

2.6 | Spatial Autocorrelation

To further evaluate the spatial autocorrelation of NDVI values within the buffer zones of the BTAP plantation sites for the years 2015, 2019, and 2023, we employed Moran's *I* statistic. This widely used measure for spatial autocorrelation in geographical data helps to understand the degree to which NDVI values are clustered or dispersed within the study area (Mielke et al. 2020). NDVI data points were obtained from a spatial dataset that included coordinates and NDVI values for each observation point within the BTAP buffer zones. Spatial weights were defined based on the k-nearest neighbors' approach (Chen 2023).

Specifically, each observation was linked to its four nearest neighbors. This choice was informed by preliminary spatial analyses and validated through a sensitivity test over $k = 3$ to $k = 8$, which confirmed the robustness of observed spatial autocorrelation patterns. This was achieved using a k-nearest neighbor algorithm, which identifies the eight closest points for each data point in the dataset (Okunev and Kushnareva 2023). The spatial weights matrix W was constructed such that $\omega_{ij} = 1$ if locations i and j are neighbors and $\omega_{ij} = 0$ otherwise. The weights were row standardized to ensure that the sum of weights for each row equals one, which normalizes the influence of neighbors (Chen 2021) (Equation 6).

$$I = \frac{N}{S_0} \bullet \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \omega_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \quad (6)$$

where N is the number of spatial units indexed by i and j. x_i is the NDVI value at location i. $\bar{x}$ represents the mean of the NDVI values. ω_{ij} are the elements of the spatial weights' matrix W. $S_0 = \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_{ij}$ Corresponds to the sum of all spatial weights.

The significance of Moran's *I* was tested using a randomization approach (Chen 2012). The null hypothesis posits no spatial autocorrelation (i.e., the observed spatial pattern is random). A low *p* value (typically less than 0.05) indicates significant spatial autocorrelation, confirming that the observed clustering or dispersion of NDVI values within the buffer zones is unlikely to be due to random chance (Gaspard, Kim, and Chun 2019). By integrating the Getis-Ord Gi* hotspot analysis with Moran's *I* spatial autocorrelation, this study provides a comprehensive understanding of the spatial patterns and vegetation dynamics within the BTAP plantation buffer zones, highlighting areas of significant vegetation change and their spatial relationships.

2.7 | Machine-Learning Analysis for NDVI Prediction

This study employed ANN to predict NDVI based on a suite of climatic, topographic, and demographic variables. This method was chosen for capturing complex, nonlinear relationships within large datasets, making it ideal for ecological and environmental modeling tasks. The dataset included temp, Ppt, SR, SM, elevation, slope, aspect, and POP variables. These variables were carefully selected to reflect the key drivers of vegetation dynamics and served as predictors for NDVI, the response variable. The dataset was divided into training and evaluation subsets, with 70% allocated for model training and 30% reserved for model evaluation (Bradshaw et al. 2023; Gerber and Nychka 2021; Anees et al. 2024a; Luo et al. 2024). This split ensured that the model could be adequately trained while maintaining enough data to assess its predictive performance independently.

2.7.1 | ANN Model

In this study, we implemented an ANN to predict NDVI using a comprehensive set of environmental and demographic variables. The ANN architecture consists of an input layer, four hidden layers, and an output layer inspired by the computational processes of the human brain (Lot et al. 2020; Madhiarasan and Louzazni 2022; Zhang et al. 2023; Huang et al. 2024; Zhang et al. 2022). To assess multicollinearity among predictors, we calculated the VIF for each variable, with all values provided in the supporting information (refer to Figure S2 and Table S1). To ensure robust model performance, a fivefold cross-validation approach was employed during hyperparameter tuning. This method allowed us to assess model stability across different data splits, providing a more reliable estimate of predictive accuracy while minimizing the risk of overfitting. The decision to use four hidden layers was based on both empirical testing and existing

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

literature, which suggests that deeper networks can capture complex, hierarchical patterns effectively, especially in ecological data modeling (Haq et al. 2022). Each hidden layer progressively reduced the number of neurons, 128 in the first layer, 64 in the second, 32 in the third, and 16 in the fourth, to enable the network to build increasingly abstract representations and reduce computational complexity without sacrificing performance (see Figure 3).

The ReLU (rectified linear unit) activation function was employed across all hidden layers, introducing nonlinearity into the model and enabling it to learn complex patterns in the data (Shahade et al. 2023). The Adam optimizer was chosen for its efficient handling of large datasets and adaptive learning rate, which facilitates faster convergence during training (Zhang 2019). The learning rate of 0.001, batch size of 64, and 200 epochs were selected based on extensive tuning and validation, as these values provided a balance between convergence speed and model accuracy. A systematic approach was adopted to tune the key hyperparameters, including the number of neurons in the hidden layers, the learning rate, and the number of epochs. This process was critical in refining the model's performance and ensuring accurate NDVI predictions. Table 4 outlines the final hyperparameters selected after extensive tuning, including a learning rate of 0.001, a batch size of 64, and an epoch count of 200.

To prevent overfitting, we incorporated dropout with a rate of 0.2 and L2 regularization with a strength of 0.01. Dropout randomly deactivates neurons during training, reducing the model's reliance on specific neurons and enhancing generalization (Ben Hamida et al. 2024). L2 regularization, which penalizes large weight magnitudes, was applied based on literature supporting its efficacy in improving model robustness, especially for regression tasks in environmental applications (Kim 2024). A validation set comprising 10% of the training data was also used to monitor performance and detect overfitting or underfitting. These hyperparameter choices, informed by empirical tuning and supported by relevant literature, enabled the model to achieve accurate NDVI predictions while maintaining generalization capabilities. To assess the uncertainty in model predictions, we conducted a Monte Carlo simulation involving 1000 iterations, where random noise was introduced to key predictor variables (e.g., temperature, precipitation, and elevation). This process generated a distribution of predicted NDVI values for each observation, reflecting the variability introduced by uncertainty in predictor measurements. We then calculated 95% confidence intervals (CI) for each prediction to quantify the expected range of true values. Furthermore, observations with high uncertainty (i.e., CI width exceeding a predefined threshold) were flagged to identify potential areas requiring additional data or model refinement.

2.7.2 | SHAP Feature Importance Analysis

We used SHapley Additive exPlanations (SHAP) as an interpretability framework to analyze the significance of input features in our NDVI prediction model. SHAP is model agnostic and provides a single measure of feature importance that can be used across all machine-learning models (including ANN) (Younisse, Ahmad, and Abu Al-Haija 2022). Traditional methods, such as feature importance in decision tree-based models, are model specific. In contrast, SHAP values provide a general and theoretical approach to estimating the contributions of each feature. SHAP is based on game theory, and the SHAP value of a feature is what the model contributes to its output (Kim et al. 2023; Nordin et al. 2023). We average this contribution over all such combinations to provide an overall

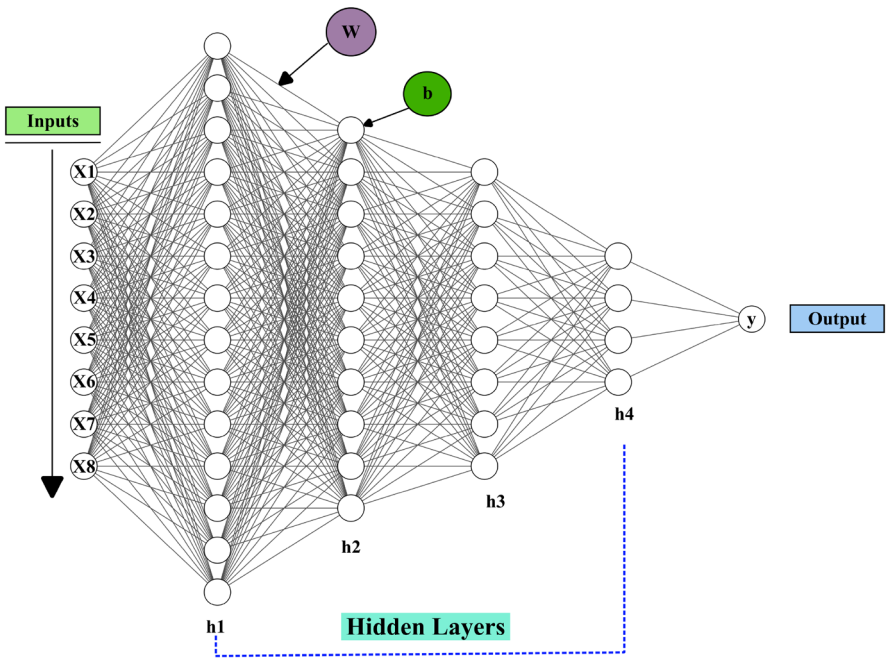

FIGURE 3 | Schematic representation of a deep neural network used for regression analysis. The network consists of eight input features (X1 to X8), multiple hidden layers (h1–h4), and a single output node (y). Weights (w) and biases (b) are applied to each layer's connection. The hidden layers transform the input data through nonlinear activation functions, ultimately predicting the output value.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Hyperparameters used for training the ANN model.

Hyperparameter	Value
Number of hidden layers	4
Learning rate	0.001
Batch size	64
Number of epochs	200
Activation function	ReLU
Optimizer	Adam
Dropout rate	0.2
L2 regularization	0.01

feature importance measure. The SHAP value for a given feature j is defined as (Equation 7):

$$\varphi_j = \frac{1}{|N|!} \sum_{S \subseteq Nleft\{j\}} |S|!\,(|N| - |S| - 1)!\,\left[f(S \cup \{J\}) - F(S)\right] \quad (7)$$

where F is the set of all features. S is any subset of the features F that does not include feature j. $|S|$ is the number of features in subset S. $f(S)$ is the prediction function applied to the subset S. $f(S\cup\{j\})$ is the prediction function applied to the subset S with the addition of feature j. φ_j represents the SHAP value for feature j. This formulation ensures that the SHAP value captures the average marginal contribution of a feature, considering all possible interactions with other features (Gebreyesus et al. 2023). The computation of SHAP values requires evaluating the model on different subsets of features, which can be computationally intensive but provides a robust measure of feature importance.

In our study, SHAP values were computed using the empirical approach implemented in the shapr package (Aas, Jullum, and Løland 2021; Kelemen et al. 2024), which ensures robustness in the estimation of feature contributions. After training the RF model to predict NDVI, an explainer object was created using the training dataset. The baseline prediction was set as the mean NDVI value of the training data. SHAP values were then calculated by evaluating the model's predictions across various subsets of input features from the evaluation dataset, simulating the contribution of each feature to the predicted NDVI. This analysis was performed in two stages: first, for the overall study area, and second, specifically for the plantation sites associated with the BTAP. SHAP values were calculated for each feature for the entire study area to understand their contribution to NDVI predictions across different environmental and topographic conditions in KPK. A separate SHAP analysis was conducted for the BTAP plantation sites to identify the key drivers of vegetation recovery within these afforestation zones. The SHAP values provided insights into the factors most critical in improving NDVI within these areas. The dual application of SHAP for the overall area and plantation sites provides a clear understanding of factors influencing vegetation dynamics, aiding informed decisions in managing and optimizing afforestation strategies in KPK.

2.8 | Accuracy Assessment

The accuracy of both the RF and ANN models was rigorously evaluated using several statistical metrics: root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), mean squared error (MSE), and the relative root mean square error (RRMSE%). These metrics were selected to assess the models' predictive performance across different dimensions comprehensively (Mehmood et al. 2024d; Xinde et al. 2023). The RMSE was used to quantify the average magnitude of the error between the predicted and actual NDVI values. As a metric sensitive to significant errors, RMSE provides valuable insights into the model's accuracy by measuring how well the model predicts NDVI across all data points (Doulah Md and Islam Md 2023; Pan, Harrou, and Sun 2023) (Equation 8).

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{\left(y_{i-}\hat{y}_i\right)^2}{n}} \quad (8)$$

where n represents the number of observations, y_i the observed values, and $\hat{y}$ the predicted values.

The MAE was also calculated to complement RMSE. This metric measures the average absolute difference between predicted and actual values, providing a straightforward and easily interpretable assessment of prediction accuracy (Ağbulut, Gürel, and Biçen 2021; Nadakinamani et al. 2022; Guo et al. 2024; Zhang et al. 2024a, 2024b). Unlike RMSE, MAE treats all errors equally, making it useful for assessing the model's performance (Equation 9).

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{n} \quad (9)$$

The R^2 value, or the coefficient of determination, was used to determine the proportion of variance in NDVI that is predictable from the independent variables. This metric is beneficial for understanding the goodness of fit of the model, indicating how well the model captures the variability in the data (Nihar, Patel, and Danodia 2022; Suarez, Robson, and Brinkhoff 2023) (Equation 10).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \quad (10)$$

where $\bar{y}$ represents the mean of the observed values.

In addition, the MSE was computed, which represents the average of the squared differences between predicted and actual values (Equation 11), helps identify the variance of residuals, offering another perspective on the model's prediction errors (Ahmad et al. 2023).

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \quad (11)$$

Finally, the RRMSE% was calculated to provide a normalized measure of prediction error relative to the mean observed value.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

RRMSE% (Equation 12), according to Shoko, Mutanga, and Dube (2018), is particularly useful in making the error magnitude more interpretable in the context of the scale of the data. This metric allows for an easier comparison across different datasets or studies.

$$\text{RMSE} = \left(\frac{\text{RMSE}}{\bar{y}}\right) \times 100 \tag{12}$$

By computing these metrics for the RF and ANN models, we ensured a robust and thorough evaluation of their predictive accuracy. This comprehensive approach allows for a detailed comparison between the models, highlighting their strengths and weaknesses in predicting NDVI based on the chosen environmental and demographic variables.

3 | Results

3.1 | LULC Classification and Change Detection Analysis

The LULC classification results for Khyber Pakhtunkhwa (KPK), Pakistan, across the years 2015, 2019, and 2023, reveal significant changes in land cover (Table 5). The area under tree cover increased from 5153.57 km^2 (25.02%) in 2015 to 6835.66 km^2 (29.99%) in 2023. Grassland areas expanded from 106.22 km^2 (0.52%) in 2015 to 483.78 km^2 (2.12%) in 2023. This growth suggests successful natural succession and practical afforestation efforts, contributing to improved ecosystem health. Arable land showed a slight fluctuation, increasing from 3164.76 km^2 (15.36%) in 2015 to 3326.47 km^2 (15.45%) in 2019 and then decreasing to 3024.83 km^2 (13.27%) in 2023 (Figure 4). Built-up areas expanded significantly from 1641.16 km^2 (7.97%) in 2015 to 2483.85 km^2 (10.90%) in 2023, reflecting urban growth and development. Barren land decreased from 4251.24 km^2 (20.64%) in 2015 to 3830.70 km^2 (16.81%) in 2023. This reduction indicates successful land rehabilitation and increased vegetation cover, further supporting the positive outcomes of the BTAP.

The change detection analysis for Khyber Pakhtunkhwa (KPK) between 2015 and 2023 provides valuable insights into the impacts of the BTAP, focusing on tree cover, shrubland, and barren land classes. Tree cover experienced significant growth throughout the study period. In the initial phase from 2015 to 2019, tree cover increased from 4983.2 km^2 to 5898.8 km^2, reflecting the early success of the afforestation efforts. This upward trend continued from 2019 to 2023, with tree cover expanding to 6835.66 km^2. Tree cover saw a net increase of 5687.1 km^2 over the entire period. The gains in tree cover amounted to 6352.1 km^2, with a loss of only 665 km^2, resulting in a 24.1% net change. These figures highlight the project's effectiveness in significantly enhancing forest cover in the region.

Shrubland, a crucial vegetation type in the area, underwent notable changes, particularly transitioning to tree cover. From 2015 to 2019, about 475.3 km^2 of shrubland was converted to tree cover. This trend persisted from 2019 to 2023, with an additional 857 km^2 of shrubland transitioning to tree cover. Across the period, 889.2 km^2 of shrubland was converted to tree cover (Figure 5). The net change in shrubland was 5232.7 km^2, with 6928.9 km^2 gained and 1696.2 km^2 lost, resulting in a net change percentage of 22.1%. These transitions underscore the successful implementation of afforestation practices to increase forest cover while effectively managing shrub ecosystems. Barren land saw the most dramatic reduction, reflecting successful land restoration efforts by the BTAP. The area of barren land decreased from 2734.7 km^2 in 2015 to 3022.7 km^2 in 2019 and further to 2293.9 km^2 in 2023. Over the entire study period, barren land was reduced by 1847.7 km^2, with total gains of 3326.9 km^2 and losses of 2154.8 km^2, resulting in a net change of 1172.1 km^2 and a net change percentage of 5%. This significant decrease in barren land highlights the project's effectiveness in converting degraded lands into productive vegetated areas, contributing to ecological restoration. The BTAP has markedly increased tree cover and effectively converted shrubland and barren land into forested areas, significantly improving the region's environmental health. The net increases in tree cover and the substantial reduction in barren land underscore the success of the BTAP's comprehensive approach to sustainable land use and ecological restoration in Khyber Pakhtunkhwa.

3.2 | Buffer Analysis of BTAP Afforestation Project Within the Plantation Sites

The BTAP afforestation project's buffer analysis reveals significant land-cover improvements across various districts in Khyber Pakhtunkhwa (KPK) from 2015 to 2023. The project has successfully increased tree cover, a critical indicator of ecological

TABLE 5 | LULC classification results for 2015, 2019, and 2023.

Class	Area (km^2)	(%)	Area (km^2)	(%)	Area (km^2)	(%)
	2015		2019		2023	
Tree	5153.57	25.02	5729.32	26.62	6835.66	29.99
Grassland	106.22	0.52	398.17	1.85	483.78	2.12
Swamp vegetation	473.13	2.30	206.19	0.96	264.78	1.16
Arable land	3164.76	15.36	3326.47	15.45	3024.83	13.27
Shrubland	5811.06	28.21	6123.43	28.45	5866.51	25.74
Built up	1641.16	7.97	1989.04	9.24	2483.85	10.90
Barren land	4251.24	20.64	3752.29	17.43	3830.70	16.81

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

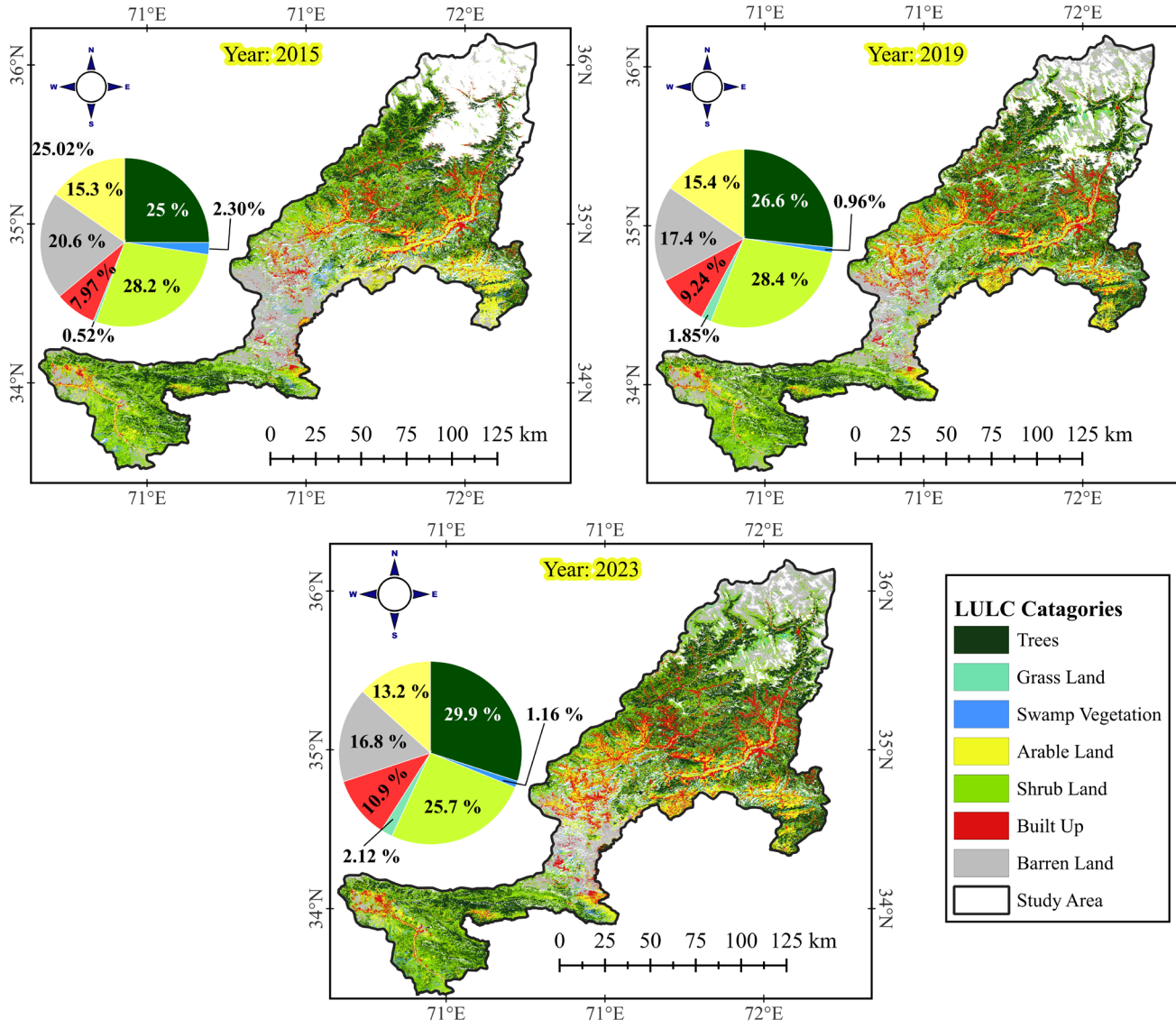

FIGURE 4 | LULC classification maps for the years 2015, 2019, and 2023 within the study area, illustrating the spatial distribution and changes in LULC categories over time. The maps display eight LULC categories: Trees, Grassland, Swamp Vegetation, Arable Land, Shrubland, Built-Up Areas, Barren Land, and the Study Area boundary. The accompanying pie charts show the percentage composition of each LULC category for each respective year, highlighting shifts in land-cover types.

restoration. For instance, tree cover in Bajaur expanded from 10.10% in 2015 to 19.20% in 2023, reflecting a remarkable 90.28% growth. This positive trend is reflected in other districts, such as Khyber, where tree cover rose from 8.25% to 14.75% over the same period, marking a 78.79% increase. Similarly, in Mohmand, tree cover increased from 9.87% to 18.33%, an 85.74% growth, indicating the effectiveness of the afforestation efforts across multiple regions (Figure 6). In addition to tree cover, changes in shrubland and other land-cover types provide critical insights into the ecological impact of the afforestation project. Shrubland proportions have remained relatively stable in most districts, with slight variations. For example, in Bajaur, shrubland changed from 36.91% in 2015 to 37.22% in 2023, and in Khyber, it remained around 35%. In Mohmand, however, shrubland increased from 4.80% to 7.34%, suggesting some regions may be experiencing shifts in land-cover types due to the project.

The status of afforestation, as indicated by the increased tree cover across all districts, is highly encouraging. However, the data also highlights potential areas for further plantation efforts. Grassland proportions, which have increased in many districts, indicate areas that may still be in transition and could benefit from additional afforestation activities. For instance, grassland in Bajaur rose from 1.11% in 2015 to 6.25% in 2023, while in Mohmand, similar increases were noted, suggesting these areas have potential for further forest cover expansion. Moreover, arable land, which showed minor changes, represents another opportunity for future afforestation. In Khyber, arable land remained stable, while Bajaur slightly decreased from 16.42% to 17.95%. These areas could be targeted for conversion to forested regions, thereby enhancing the overall ecological impact of the BTAP project. Barren land areas have shown noteworthy changes as well. For example, in Bajaur, barren land decreased from 11.35% in 2015 to 7.82% in 2023, indicating successful

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

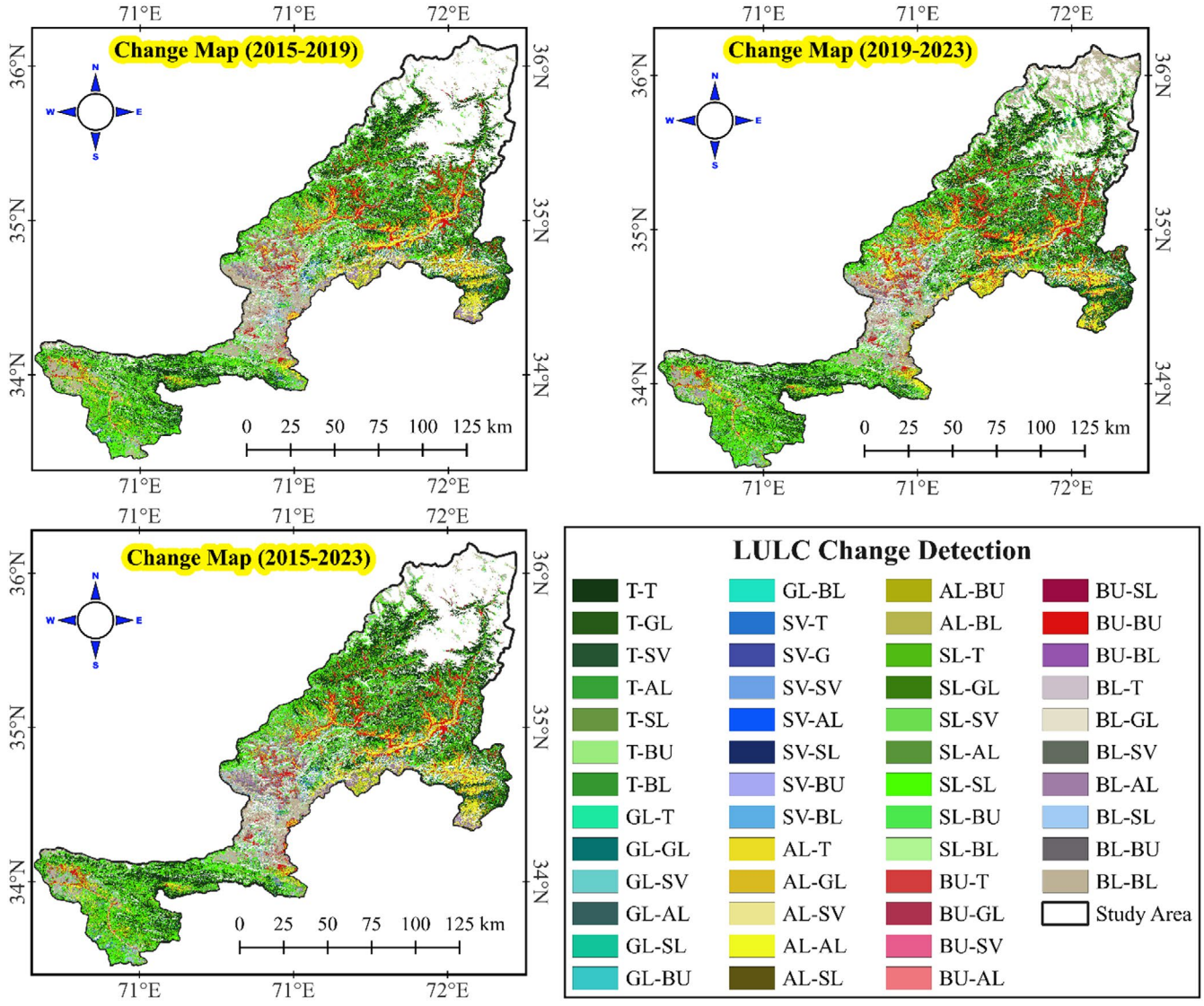

FIGURE 5 | LULC change detection maps for the periods 2015–2019, 2019–2023, and 2015–2023, highlighting transitions between different LULC categories within the study area. Each color-coded category in the legend represents a specific type of transition, such as forest-to-grassland, arable-to-built up, or shrubland-to-swamp, capturing the direction and nature of land-cover changes over time. These maps reveal areas of significant land transformation, indicating patterns of urban expansion, deforestation, and vegetation shifts, which are essential for assessing the ecological impact of land-use changes in the region.

conversion to vegetated land. In Khyber, barren land also reduced from 9.51% to 6.48% during the same period. These decreases reflect the project's success in transforming previously barren areas into productive, vegetated land. Continued efforts to convert remaining barren regions into forested areas is crucial in sustaining the momentum of the BTAP project.

Overall, within the buffer zones of the plantation sites, significant land-cover changes were observed from 2015 to 2023. In 2015, barren land accounted for 28.94% of the area, which reduced to 26.07% in 2019 and 17.48% in 2023 (Figure S1). Meanwhile, tree cover increased from 15.47% in 2015 to 19.20% in 2019, reaching 24.07% in 2023, indicating substantial afforestation progress. Additionally, shrubland proportions slightly decreased from 29.80% in 2015 to 25.79% in 2023. These changes reflect the effectiveness of the BTAP project in transforming barren lands into vegetated areas and enhancing overall ecological restoration.

3.2.1 | Analysis of LULC Transitions Within Plantation Buffer Zones

The analysis of LULC changes within 500-m buffer zones around plantation sites reveals significant dynamics in land-cover transitions, highlighting the impact of the BTAP. Notably, there were 16,057 transitions from barren to barren land, covering 33.4 km^2, constituting 29.71% of all transitions (Figure 7). This indicates a high degree of persistence within barren land areas. However, significant transitions from barren land to other uses, such as 11.7 km^2 to crops (17.23%) and 1.02 km^2 to trees (22.57%), suggest a shift toward productive land-use and potential ecological rehabilitation. The built-up area experienced considerable changes, with 445,497 counts remaining built-up (13.5 km^2, 8.21%). There were notable transitions from built-up areas to crops (1.19 km^2, 16.19%) and shrubland (0.968 km^2, 27.48%), reflecting ongoing urbanization pressures and the potential expansion of suburban areas. The

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

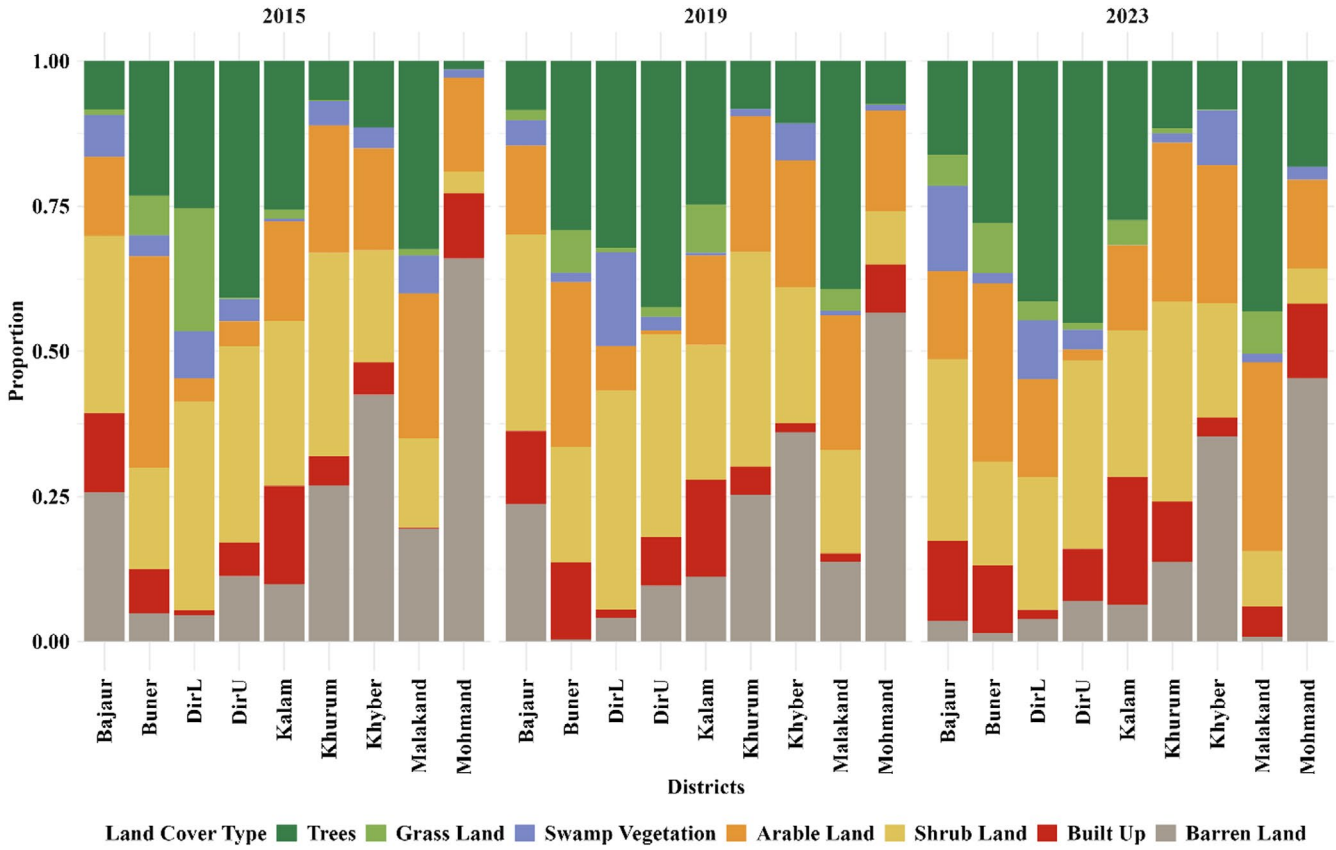

FIGURE 6 | LULC changes in BTAP afforestation project buffer zones from 2015 to 2023. This figure illustrates the proportional changes in different LULC types across districts in the BTAP buffer zones over three time periods (2015, 2019, and 2023). The increase in tree cover (green bars) is evident across most districts, highlighting the success of afforestation efforts under the BTAP initiative. Concurrently, a decrease in barren land (brown bars) and built-up areas (red bars) can be observed, signifying reduced land degradation and urban sprawl in some districts. Notable shifts in grassland (light green) and arable land (orange) proportions suggest varying impacts of land-use policies and local ecological conditions across the districts. These results emphasize the regional variability in land-cover changes and the critical role of afforestation in improving vegetation cover.

stability of built-up land cover shows continued urban and infrastructural development, impacting land-use patterns.

Trees remained stable over 30.7 km^2, accounting for 22.57% of the total transitions, suggesting effective preservation efforts and natural resilience. However, transitions from trees to shrubland (2.22 km^2, 27.48%) and built-up areas (0.676 km^2, 8.21%) highlight urban expansion and infrastructure development pressures. Conversely, the transition from shrubland to trees amounted to 7.21 km^2, indicating successful restoration and afforestation efforts to convert degraded lands into forested areas. This transition underscores the positive impact of BTAP in increasing tree cover within degraded regions. The dynamic nature of agricultural land use is evident, with 24.9 km^2 of arable land (21.59%) remaining in crops, while transitions from crops to barren land accounted for 0.493 km^2 or 16.19% of transitions. Such shifts reflect changes in agricultural practices or land abandonment. Notably, arable land also transitioned to tree cover over 5.37 km^2, representing 22.57%, highlighting efforts to increase forest cover through afforestation.

The chi-square test results, with a highly significant p value (< 0.001), suggest that the observed LULC changes are not random but influenced by underlying factors likely related to the impacts of plantation activities. Key insights include potential deforestation near plantation sites, as indicated by transitions from trees to shrubland and built-up areas, highlighting pressures from urbanization and infrastructure development. Conversely, the significant transition from shrubland to trees underscores successful afforestation efforts to convert degraded lands into forested areas. Furthermore, the stability in certain land-cover types, such as shrubland and tree cover, implies effective management practices or inherent resilience of these ecosystems.

3.3 | Hotspot Analysis of NDVI Data Within the Buffer Zone of Plantation Sites (2015, 2019, and 2023)

The hotspot analysis performed on the NDVI data for 2015, 2019, and 2023 provides significant insights into the spatial and temporal dynamics of vegetation in the study area. This section delves into the key findings, discussing the observed changes and their implications for the BTAP. In 2015, significant portions of the study region were classified as cold spots, particularly in the northern and central parts, indicating areas of lower vegetation density. By 2019, there was a noticeable increase in hotspots, especially

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

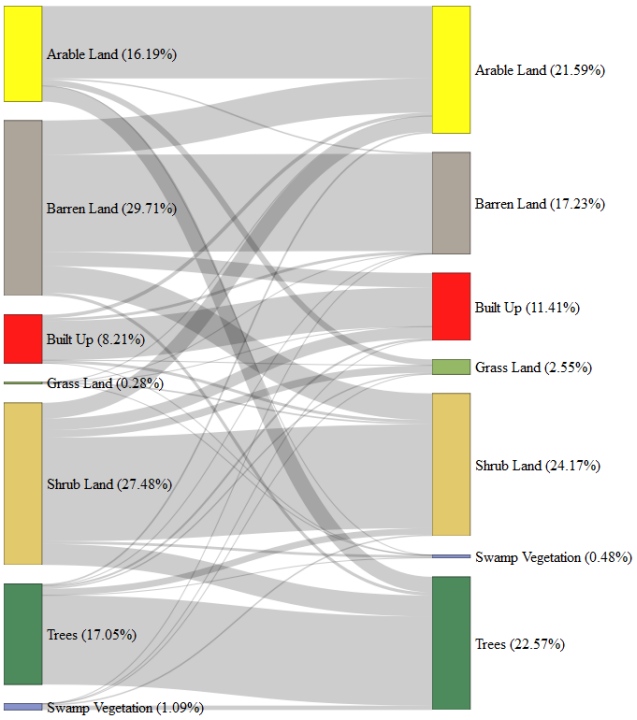

FIGURE 7 | Sankey diagram illustrating the transitions between different LULC classes within 500-m buffer zones around plantation sites from 2015 to 2023.

in the southern and central regions, suggesting an improvement in vegetation density, likely due to the afforestation efforts under BTAP (Figure 8). This positive trend continued into 2023, with more areas in the northern region also showing high vegetation density, indicating a sustained impact of the afforestation project. The spatial distribution of hotspots and cold spots suggests that the southern areas have consistently improved vegetation density over the years, potentially due to targeted afforestation efforts. Initially, the northern regions displayed significant cold spots in 2015 but transitioned to include more hotspots by 2023, reflecting effective vegetation recovery efforts in these areas.

The hotspot analysis revealed significant changes in vegetation density over the years. In 2015, 36.76% of the study area was identified as high-confidence hotspots, while 28.03% were high-confidence cold spots. By 2019, high-confidence hotspots increased to 39.59%, and high-confidence cold spots decreased to 25.44%. In 2023, high-confidence hotspots rose to 42.56%, while high-confidence cold spots reduced to 21.34%. Moderate- and low-confidence hotspots and cold spots showed slight variations, reinforcing the positive impact of afforestation efforts (Table 6). In addition, Moran's *I* analysis for the years 2015, 2019, and 2023 revealed an increasing positive spatial autocorrelation, supporting the hotspot analysis findings. In 2015, the Moran's *I* value was 0.929, indicating strong clustering of NDVI values ($p < 0.001$). By 2019, Moran's *I* value slightly increased to 0.931, further clustering NDVI values ($p < 0.001$). In 2023, Moran's *I* value reached 0.933, suggesting an even stronger clustering of high NDVI values ($p < 0.001$) (Figure 9). The positive correlation further confirms the spatial clustering of NDVI, supporting the observed increase in vegetation density.

3.4 | Interannual NDVI Trends and Distribution Analysis

This study examines the interannual trends of NDVI across 343 plantation sites involved in the Billion Tree Afforestation Project (BTAP). Utilizing linear regression, the analysis offers insights into temporal changes in NDVI, thereby reflecting the ecological impacts of the afforestation efforts. The results indicate a statistically significant positive trend in NDVI across all plantation sites, with an estimated slope of 0.0030 ($p < 0.01$), signifying a consistent annual increase in vegetation density. This trend underscores the success of the BTAP in enhancing vegetation cover and improving ecosystem health at these plantation sites (Figure 10A).

The density plot of NDVI slopes reveals a right-skewed distribution, suggesting that while most sites experienced moderate improvements, a subset achieved exceptionally high growth rates (Figure 10B). The peak of this distribution centers around a slope of approximately 0.003, further corroborating the overall positive trend. Notably, many sites exhibit slopes clustered near zero, indicating stable NDVI values with minimal change over time (Figure 10C). These observed variations may be influenced by local environmental conditions, management practices, or species selection, underscoring the need for tailored afforestation strategies to optimize outcomes across diverse ecological contexts.

Among the districts analyzed, Bajaur and Mansehra demonstrated the most considerable NDVI improvement, with average slopes of 0.0063 and 0.0061, respectively, reflecting highly effective afforestation efforts. Abbottabad and Swat showed intermediate performance, with slopes of 0.0040 and 0.0038, suggesting moderate success that could benefit from enhanced management strategies. Conversely, Lower Dir exhibited the lowest improvement, with an average slope of 0.0015, highlighting significant challenges in this region.

3.5 | Model Performance and Validation

The ANN model's architecture, optimized through careful hyperparameter tuning, was subjected to training and validation to evaluate its predictive performance. The training process spanned 200 epochs, which allowed the model to converge effectively without imposing excessive computational demands (Figure 11A), the training and validation loss curves converged smoothly, indicating that the model learned effectively from the data without overfitting. This result underscores the success of the chosen hyperparameters in balancing model complexity with generalization ability.

Further evaluation of the model's performance is depicted in Figure 11B, which shows a scatter plot comparing the predicted NDVI values against the observed NDVI values. The strong alignment of the data points along the diagonal line suggests a high degree of accuracy in the model's predictions, with minimal deviation from the actual NDVI values. The R^2 value of 0.8556 further confirms that the model explains approximately 85.56% of the variance in the observed NDVI values. Additionally, the

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

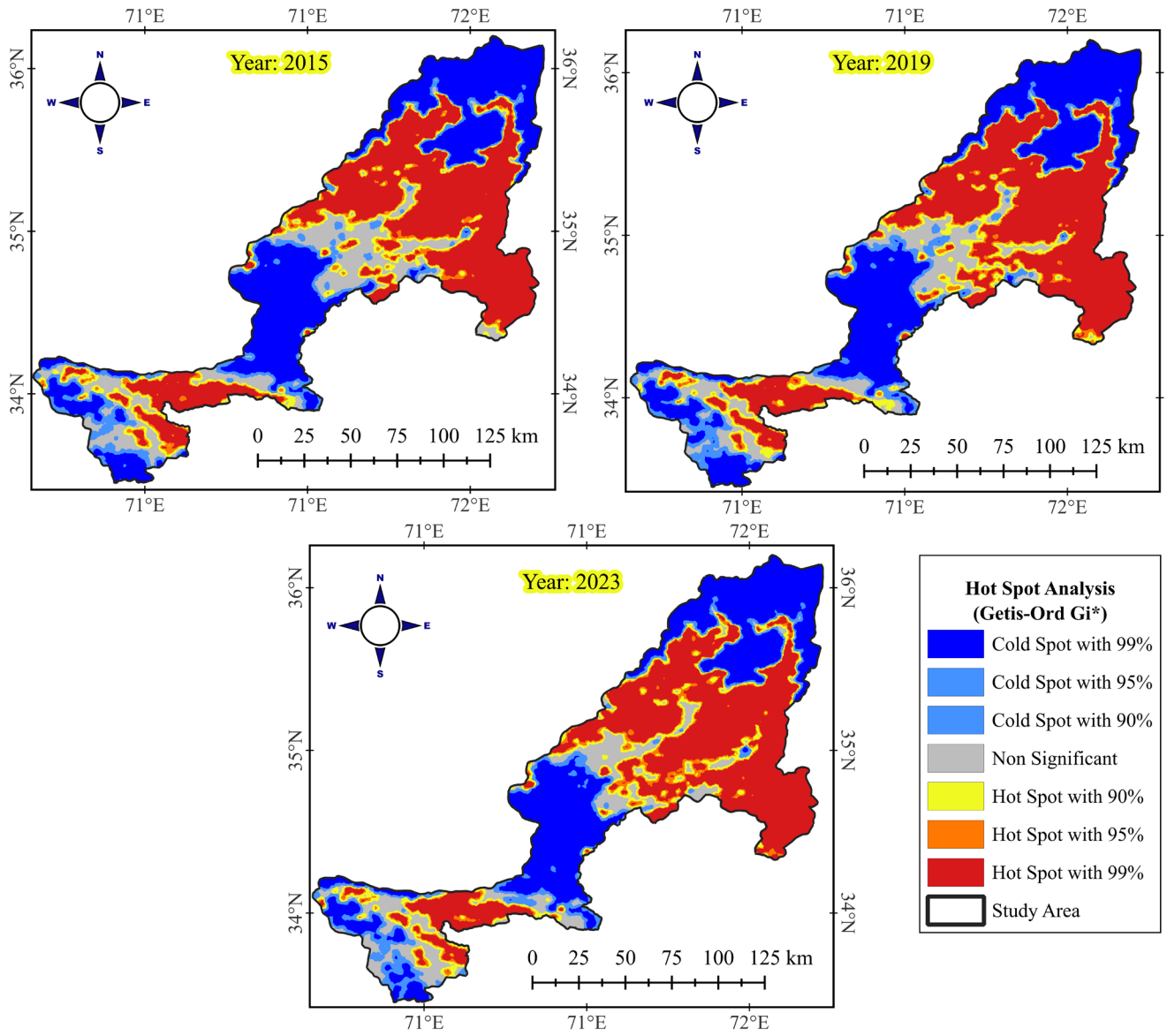

FIGURE 8 | Spatial distribution of hotspots and cold spots in NDVI data within BTAP plantation buffer zones (2015, 2019, and 2023). This figure shows the spatial pattern of hotspots and cold spots of NDVI values within the BTAP plantation buffer zones for the years 2015, 2019, and 2023, using the Getis-Ord Gi* analysis. Red areas represent statistically significant hotspots, indicating regions with consistently high NDVI values, while blue areas represent cold spots, indicating consistently low NDVI values. The spatial distribution of these areas provides insights into regions with high vegetation productivity and regions potentially affected by degradation or other stressors.

RMSE of 0.0607 indicates a low average prediction error, further validating the model's robustness (Table 7). The testing phase of the model also demonstrated strong performance metrics. The MSE was 0.0037, with an RMSE of 0.0607 and an R^2 value of 0.8556. The RRMSE% was calculated at 17.87%, and the MAE was 0.0461, reflecting the model's accuracy in predicting NDVI values based on the input variables. The validation phase involved testing the model on new sites of plantation across three districts (Mardan, Charsada, and Peshawar), outside the area where the model was trained. This external validation highlights the model's ability to generalize to unseen data. The validation results include an MSE of 0.0057, RMSE of 0.0758, and an R^2 value of 0.7818, indicating that the model explains approximately 78.18% of the variance in the observed NDVI values (Figure S4). The MAE of 0.0606 and RMSE% of 21.80% suggest a moderate level of accuracy when applied to these new sites, demonstrating the model's practical utility in predicting NDVI for regions outside the training area.

In addition to the testing results, the training data exhibited strong performance, indicating that the model was well-tuned during the training phase. The training data results are as follows: MSE of 0.0006, RMSE of 0.0255, and an R^2 value of 0.8765. The MAE for the training data was 0.0191, indicating the model's effectiveness in learning from the training dataset. Additionally, the RRMSE% for the training data was calculated to be 7.42%, which reflects the model's accuracy relative to the average magnitude of the observed NDVI values. The systematic

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 | Temporal hotspot analysis of NDVI data within BTAP plantation buffer zones (2015, 2019, and 2023).

Year	High-confidence hotspots (%) (*n*)	Moderate-confidence hotspots (%) (*n*)	Low-confidence hotspots (%) (*n*)	Nonsignificant (%) (*n*)	Low-confidence cold spots (%) (*n*)	Moderate-confidence cold spots (%) (*n*)	High-confidence cold spots (%) (*n*)
2015	36.76 (18,380)	5.90 (2950)	2.51 (1256)	23.09 (11,544)	1.43 (715 points)	2.50 (1250)	28.03 (14,015)
2019	39.59 (19,846)	5.44 (2720)	2.20 (1098)	22.29 (11,146)	2.21 (1104)	3.16 (1580)	25.44 (12,720)
2023	42.56 (21,168)	6.06 (3029)	2.26 (1128)	24.04 (12,008)	2.04 (1018)	2.94 (1468)	21.34 (10,668)

Note: n represents the Pixel counts.

selection of hyperparameters, combined with regularization and continuous monitoring during training, resulted in a strong model capable of accurately predicting NDVI across diverse environmental conditions. The selected hyperparameters minimized the validation loss, ensuring the model could generalize well to unseen data.

The model's spatial predictions of NDVI, shown in the predicted NDVI raster Figure 12A, provide a detailed distribution of predicted values across the study area, ranging from −0.0462 to 0.6349. Residual and RMSE rasters were generated to assess spatial accuracy. The residual raster Figure 12B displays prediction errors, with values from −0.2713 to 0.2672, indicating areas of underestimation and overestimation. The RMSE raster Figure 12C shows average prediction errors, with values from 0.0000007 to 0.2713, highlighting regions of high accuracy and areas where the model's performance was less reliable. The classified RMSE accuracy map further categorizes the study area into high-, moderate-, and low-accuracy zones, covering 15,848.26 km^2, 6912.53 km^2, and 353.23 km^2, respectively Figure 12D. This classification clearly visualizes the model's performance across the region, indicating where predictions are most reliable.

An uncertainty analysis using Monte Carlo simulations was conducted to evaluate the reliability of the model's predictions. As visualized in Figure S3, predicted NDVI values are shown alongside their 95% confidence intervals, which represent the range of likely true values for each observation. Observations with high uncertainty, defined as having confidence interval widths exceeding 0.05, are highlighted in red. These observations indicate areas where the model may require refinement or where additional input data could enhance reliability. This analysis not only highlights the variability in predictions but also underscores the model's overall stability, with the majority of predictions exhibiting narrow confidence intervals and high reliability.

3.6 | SHAP Analysis of NDVI Predictors in the Study Area and Plantation Sites

A deeper understanding of the factors influencing NDVI predictions was achieved by employing SHAP to evaluate the relative importance of various environmental and demographic variables. The analysis was performed separately for the entire study area and the specific plantation sites associated with the BTAP. Elevation emerged as a key factor in the broader study area, consistently showing a strong negative influence on NDVI. This negative impact was particularly evident at elevations above 2000 m, where harsh environmental conditions, such as lower temperatures and shorter growing seasons, significantly limit vegetation growth. The average SHAP value for elevation was approximately −0.0506, highlighting its critical role in reducing NDVI at higher altitudes Figure 13C, where a decline in NDVI with increasing elevation is apparent. Precipitation showed a mixed influence, with SHAP values indicating positive and negative effects depending on specific regional conditions. The temperature had a more balanced impact, with average SHAP values around 0.0199, reflecting its varying influence on NDVI based on the local climate. On the other hand, soil moisture contributed

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

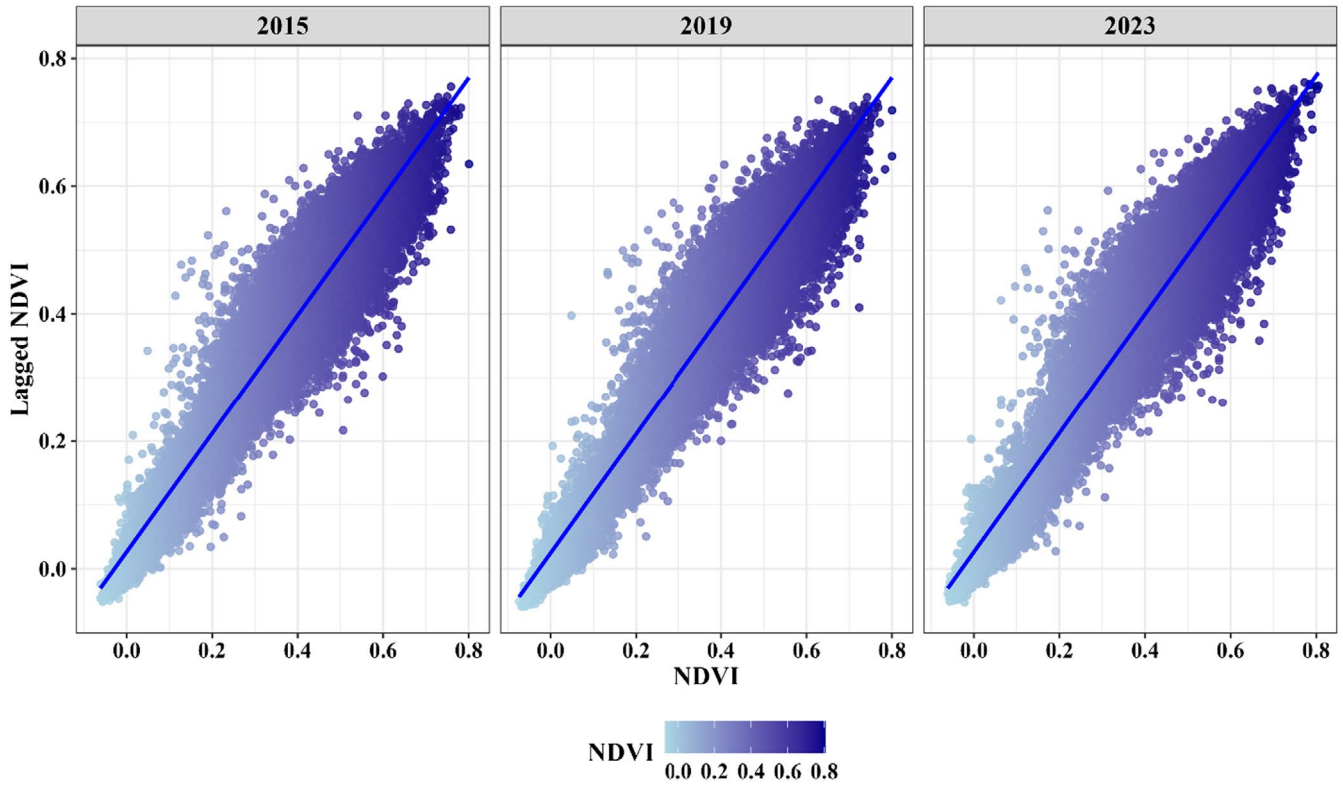

FIGURE 9 | Moran's *I* scatterplots for NDVI values in 2015, 2019, and 2023 within BTAP plantation buffer zones.

positively to NDVI, with an average SHAP value of 0.0326, underscoring its importance in supporting vegetation across the study area Figure 13B.

In the plantation sites associated with the BTAP, the SHAP analysis identified soil moisture as the most influential variable, with a strong positive impact on NDVI. The average SHAP value for soil moisture was approximately 0.0498, indicating its vital role in promoting vegetation growth within these afforestation zones. Precipitation was also a significant factor, with an average SHAP value of 0.0398, emphasizing the importance of adequate rainfall for the success of plantation efforts. Additionally, SR and POP positively influenced NDVI, with average SHAP values of 0.0264 and 0.0147, respectively, reflecting their roles in vegetation recovery within these managed environments Figure 13A. Elevation in the plantation sites had a less noticeable but relevant impact, with an average SHAP value of 0.0117. Unlike the entire study area, elevation's influence here was more balanced, suggesting that the controlled conditions within the plantation sites helped mitigate some adverse effects typically associated with higher altitudes. This finding reflects the targeted management strategies implemented in these areas to counteract the challenges posed by elevation.

4 | Discussion

4.1 | BTAP Impact on Forest Cover and Land-Use Dynamics

The results of this study demonstrate the significant positive impact of the BTAP in KPK on forest cover and overall land-use dynamics between 2015 and 2023. Tree cover increased from 25.02% (5153.57 km^2) in 2015 to 29.99% (6835.66 km^2) in 2023, while barren land decreased from 20.64% (4251.24 km^2) to 16.81% (3830.70 km^2), indicating the success of the afforestation efforts in promoting ecological restoration. Similar afforestation projects have shown comparable outcomes; for instance, the "One Million-Mu Plain Afforestation Project" in Beijing led to increased forest cover, though with varying greenness levels (Chen, Wang, and Jin 2021; Yu et al. 2018).

Additionally, the natural afforestation observed on abandoned agricultural lands in Russia and Belarus during the post-Soviet period further underscores the positive impact of such initiatives on forest recovery and ecological restoration (Ershov et al. 2022). These comparative studies reinforce the success of the BTAP in enhancing forest cover and environmental health in the region. Using Sentinel-2 imagery and the RF algorithm resulted in high-accuracy LULC classifications, with an overall accuracy exceeding 85% for the years analyzed. Comparable studies have demonstrated similar efficacy in using RF for land-cover classification. For instance, a method combining Landsat time series data with RF achieved an accuracy of 87% in predicting afforestation areas, underscoring the effectiveness of these techniques in large-scale afforestation monitoring (Avci et al. 2023; Cavalli et al. 2023).

Additionally, research mapping forest changes in Guangdong Province using Landsat and PALSAR data reported classification accuracies between 75% and 85%, further validating the reliability of these methods for accurate LULC classification (Shen et al. 2019). These studies highlight the robustness of the approach used in this analysis. This allowed for a detailed analysis of changes over time, showcasing the effectiveness of BTAP in reversing deforestation and enhancing vegetation density.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

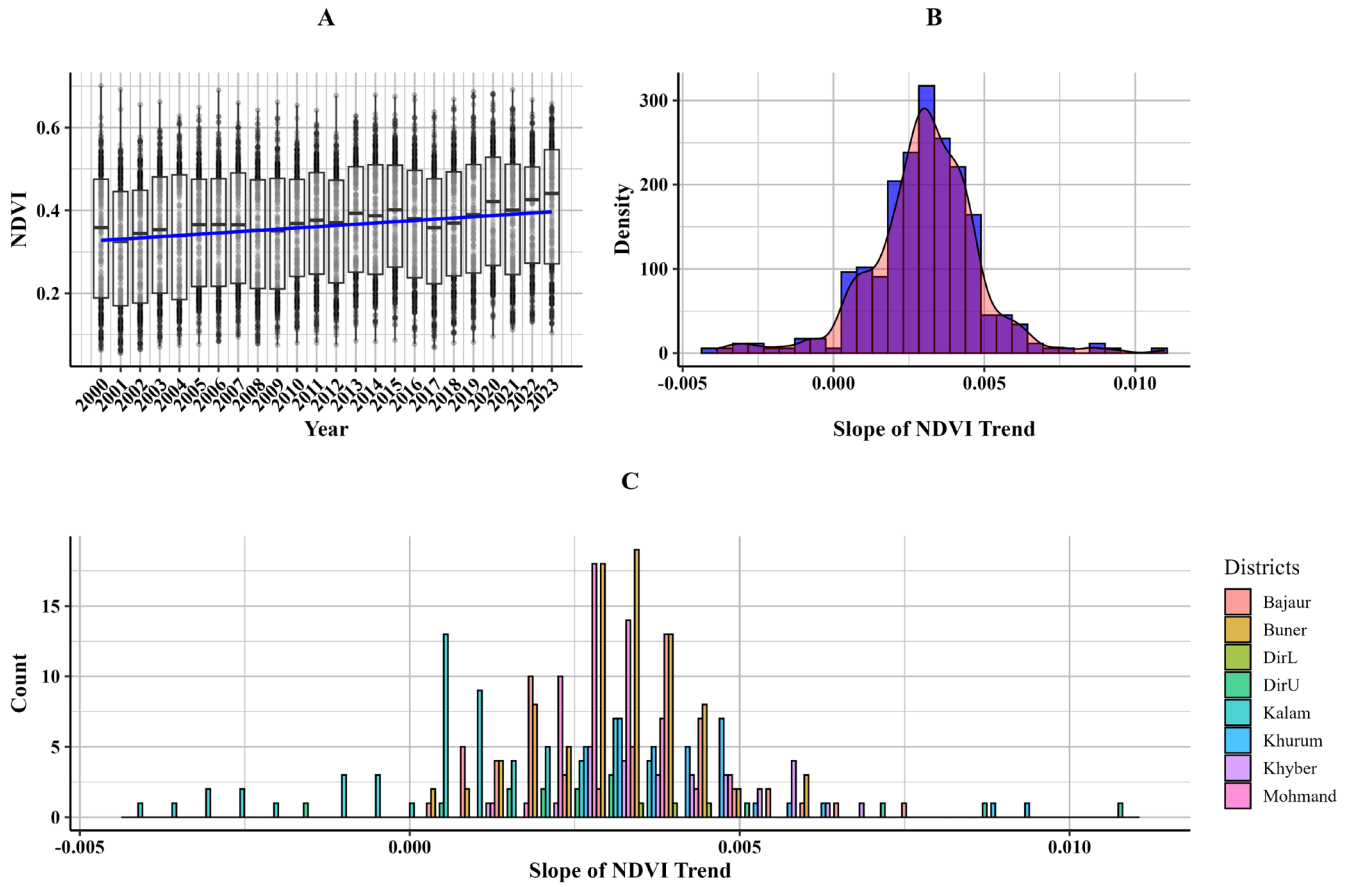

FIGURE 10 | NDVI analysis across plantation sites. (A) Boxplot depicting the annual distribution of NDVI values from 2000 to 2023, with a trend line indicating the overall increase in vegetation greenness over time. (B) Histogram with density plot showing the distribution of the slope of NDVI trends across all plantation sites, highlighting the generally positive trend in NDVI. (C) Histogram illustrating the count of NDVI trend slopes for different districts, with each color representing a specific district, indicating variability in vegetation trends across the study area.

Similar outcomes were observed in a World Bank project in Nigeria and the Three-North Afforestation Program in China, which significantly increased vegetation cover and reversed deforestation trends (Zhu et al. 2017). These studies affirm the positive impact of BTAP on ecological restoration.

4.2 | Localized Impacts and Spatial Variability in Afforestation

Buffer analysis within the plantation zones revealed significant localized improvements, with tree cover in districts like Bajaur, Khyber, and Mohmand increasing by 90.28%, 78.79%, and 85.74%, respectively, from 2015 to 2023, like other successful afforestation projects that enhanced forest cover through targeted efforts (Ullah et al. 2021). This localized success is likely attributable to targeted afforestation strategies and favorable environmental conditions within these buffer zones, as shown in studies where strategic planning and species selection significantly improved outcomes (Qiu et al. 2019). However, the persistence of barren land in some districts, such as the 17.48% observed in 2023 within the plantation buffer zones, suggests that additional efforts are required to fully rehabilitate these areas, consistent with findings that ongoing interventions are necessary in arid regions (Liu et al. 2018; Tajik, Ayoubi, and Zeraatpisheh 2020). The results of the hotspot analysis, which showed an increase in high-confidence hotspots from 36.76% in 2015 to 42.56% in 2023 and a corresponding decrease in high-confidence cold spots from 28.03% to 21.34%, further confirm the positive vegetation trends resulting from BTAP interventions, aligning with studies demonstrating satisfactory spatial variability in afforestation outcomes (Wu et al. 2021).

4.3 | Predictive Performance and Spatial Consistency in Afforestation

The machine-learning-based NDVI predictions using an ANN model yielded an R^2 value of 0.8556, with an RMSE of 0.0607, demonstrating strong predictive performance, like the high accuracies reported in studies using ANNs for environmental predictions (Celik et al. 2022; Emadi et al. 2020). SHAP analysis identified soil moisture (average SHAP value of 0.0498) and precipitation (average SHAP value of 0.0398) as the most influential variables driving vegetation growth, consistent with findings that emphasize the critical role of these factors in environmental modeling (Ren, Ling, and Wang 2023; Zhu et al. 2021). These findings suggest that successful afforestation in this region heavily depends on adequate water availability, aligning with previous research highlighting the importance of moisture for vegetation recovery (Akram et al. 2022; Fernández 2023; Otkin et al. 2019; Anees, Yang, and Mehmood 2024; Pan et al. 2023).

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

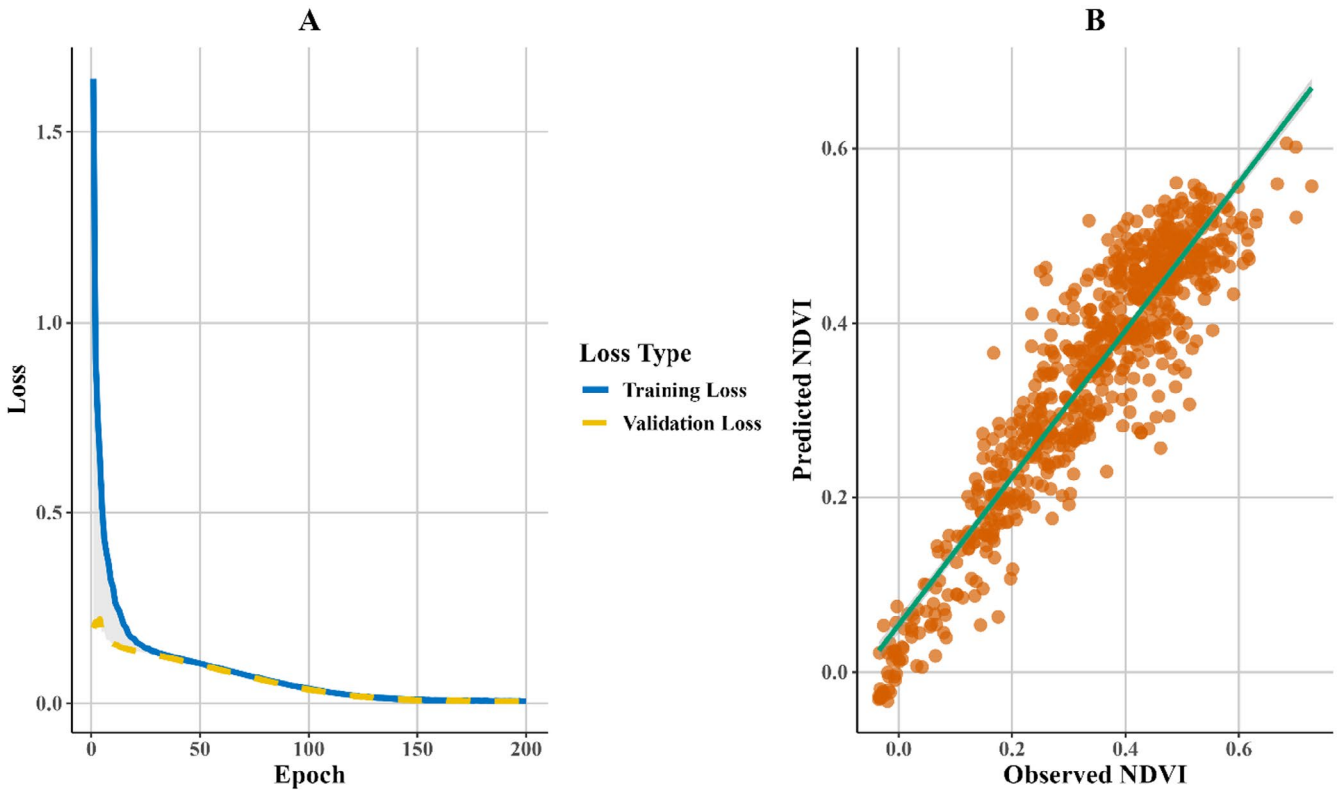

FIGURE 11 | (A) Training and validation losses of the ANN model. (B) Scatter plot of predicted versus observed NDVI values.

TABLE 7 | Performance metrics of the ANN model.

Metric	Training data	Testing data	Validation data (new sites)
MSE	0.0006	0.0037	0.0057
RMSE	0.0255	0.0607	0.0758
R^2	0.8765	0.8556	0.7818
MAE	0.0191	0.0461	0.0606
RMSE%	7.42%	17.87%	21.80%

The identification of soil moisture and precipitation as key predictors offers actionable insights for afforestation management (Andreevich et al. 2020; Usoltsev et al. 2020; Shobairi et al. 2022; Usoltsev et al. 2022; Gong et al. 2024). For example, areas with low soil moisture could benefit from interventions such as soil moisture retention techniques, including mulching and soil amendments, to enhance vegetation survival and growth (Aslam et al. 2022). Additionally, aligning plantation schedules with periods of adequate rainfall can maximize the establishment success of young vegetation. These strategies, guided by SHAP findings, enable targeted resource allocation and improved decision making in afforestation initiatives, ultimately enhancing the resilience and effectiveness of such projects.

The spatial autocorrelation results, measured by Moran's *I* statistic, showed a consistent increase in positive spatial autocorrelation from 0.929 in 2015 to 0.933 in 2023 ($p < 0.001$), further supporting the observed clustering of NDVI values and indicating a sustained and nonrandom distribution of vegetation improvements over the study period. This finding aligns with studies on vegetation patterns in earthquake-affected Southwest China, where Moran's *I* revealed significant clustering of damaged vegetation (Li et al. 2019). Another study using Moran's *I* to detect land-cover change patterns in a large-scale remote sensing imager demonstrates this statistic's effectiveness in identifying spatial clustering (Kiani et al. 2023; Self et al. 2023). This spatial consistency underscores the effectiveness of BTAP not only in isolated pockets but across the broader landscape of KPK, reflecting the broad-scale impact of the project.

While the BTAP has achieved substantial gains in forest cover, the findings also highlight areas needing further attention, particularly in districts where barren land remains high or where tree cover improvements have been minimal. Similar challenges in reforestation, such as those related to seedling survival and land degradation, have been observed in other regions, underscoring the need for more intensive efforts in these areas (Flores et al. 2021; Román-Dañobeytia et al. 2015). The persistence of barren land at 17.48% within the buffer zones suggests the need for more intensive reforestation efforts, a necessity echoed in studies emphasizing continuous forest management to address land-use challenges and optimize reforestation outcomes (Song et al. 2023; Warner et al. 2022). Moreover, the transition of some areas from tree cover to shrubland and built-up land, such as the 2.22 km^2 transition to shrubland and 0.676 km^2 to built-up areas, suggests potential pressures from urbanization and land-use change, which could undermine the long-term success of the project. This pattern is consistent with research highlighting the negative impacts of urbanization on forest ecosystems and

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

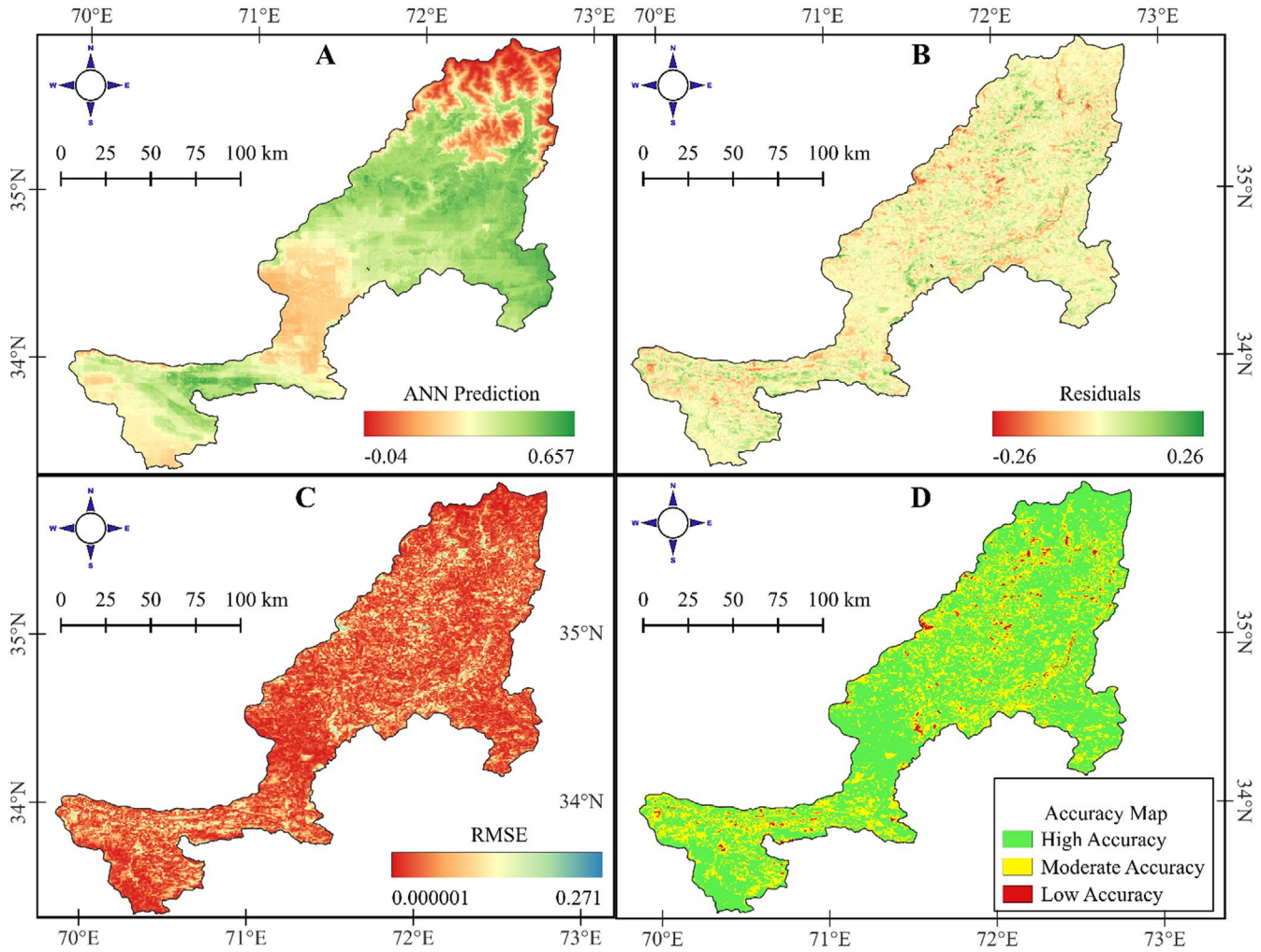

FIGURE 12 | Spatial analysis of NDVI prediction accuracy using an ANN model. (A) ANN prediction map displaying predicted NDVI values across the study area. (B) Residual map illustrating the difference between observed and predicted NDVI values, highlighting areas of over- and underestimation. (C) RMSE map shows the RMSE of the predictions, with lower values indicating better model performance. (D) Accuracy map categorizing the study area into zones of high, moderate, and low accuracy based on the residuals and RMSE, providing insight into the model's performance across different regions.

the challenges it poses to maintaining forest cover (Miroshnyk et al. 2022; Rai et al. 2023).

4.4 | Limitations

This research adopted an all-around approach by running hotspot and interannual analyses from 2000 to 2023 and considering the ground validation of the LULC classification for improved accuracy in the results. Yet, a few limitations exist. Although ground truth and site visits were done for rigorous verification of the LULC classification and effectiveness of BTAP, inherent challenges in remote sensing, such as resolution constraints and the potential influence of atmospheric conditions on satellite data, can never be fully resolved. Moreover, while the study period's length allows for capturing long-term trends, the dynamism and complexity of most ecological processes may mean that some subtle or delayed effects from afforestation are not entirely captured within this study. The integration of ground-truthing and site visits further strengthens the validity of results, reducing the limitations presented by this method.

4.5 | Future Implications

Such future studies could benefit from further integrating remote sensing with ground-based data collection, potentially through higher resolution or hyperspectral imagery, LiDAR, and UAV data. These technologies offer enhanced detail on vegetation dynamics, canopy structure, and biomass, which could complement Sentinel-2 imagery to provide a more comprehensive understanding of afforestation impacts. Extending the study duration beyond 2023 and increasing the frequency of temporal observations would also help capture long-term ecological changes and potential lag effects resulting from afforestation activities. Moreover, employing more sophisticated machine-learning models could improve predictive accuracy, accommodating the complexities of ecological data and the interactions between environmental and socioeconomic variables. These advancements

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

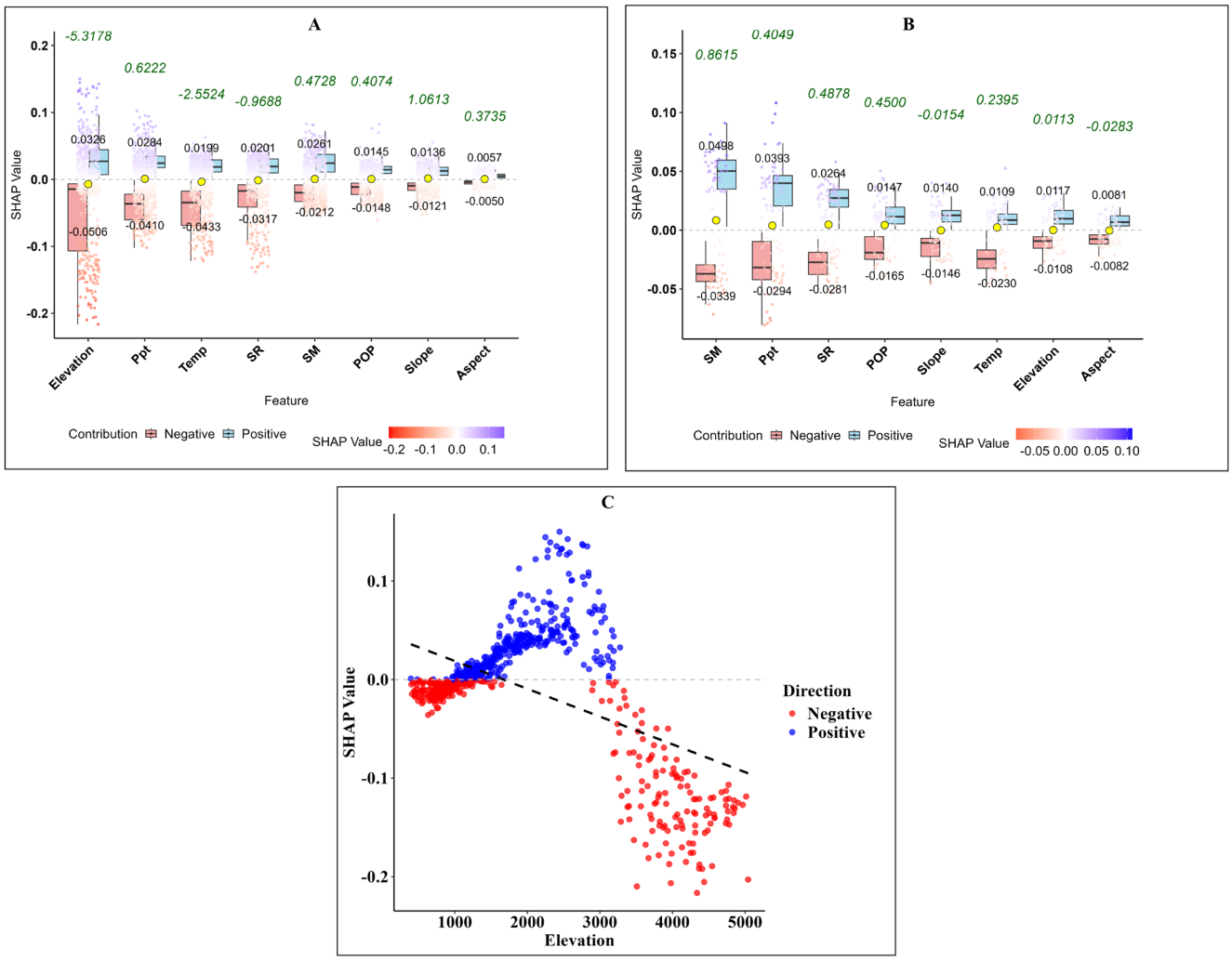

FIGURE 13 | SHAP analysis of NDVI prediction model. (A) SHAP summary plot for the entire study area, illustrating the influence of various environmental and demographic variables on NDVI predictions. The magnitude and direction of SHAP values indicate the contribution of each variable, with positive values boosting predictions and negative values reducing them. (B) SHAP summary plot for the BTAP plantation sites, focusing on critical variables affecting NDVI predictions within these zones. Notably, soil moisture (SM) and precipitation (Ppt) emerge as dominant predictors. (C) Scatter plot depicting the relationship between SHAP values and elevation, revealing a nonlinear influence of elevation on NDVI predictions. Red dots represent negative contributions, while blue dots signify positive contributions, emphasizing the varying impact of elevation across the landscape.

would support the development of sensitive, adaptive, and sustainable afforestation strategies, ensuring that projects like BTAP continue to make positive contributions to environmental restoration. Our findings demonstrate the effectiveness of large-scale afforestation in landscape restoration, advocating for continuous monitoring and adaptive strategies to address emerging challenges. The integration of remote sensing and machine learning offers a robust framework for evaluating and guiding these efforts, providing data-driven insights that are essential for sustainable land management and conservation.

5 | Conclusion

The BTAP assessment conducted in this study demonstrates the effectiveness of large-scale afforestation efforts in Pakistan using advanced remote sensing and machine-learning techniques. Our approach enabled accurate LULC classification, revealing a substantial increase in tree cover from 25.02% in 2015 to 29.99% in 2023, and a corresponding decrease in barren land. Additionally, the hotspot analysis and spatial autocorrelation confirmed positive clustering in vegetation recovery, while the ANN model's predictive accuracy underscored the critical influence of soil moisture and precipitation on vegetation health. While these findings affirm the success of the BTAP, they also highlight the need for continuous monitoring and adaptive management to address challenges such as the persistence of barren land and transitions to shrubland or built-up areas in some regions. We recommend that policymakers and project managers prioritize ongoing monitoring frameworks and adaptive strategies to ensure that afforestation efforts respond effectively to environmental changes. Engaging local communities in these efforts is essential, as their involvement can enhance land stewardship and contribute to sustainable management practices. Furthermore, implementing soil moisture retention techniques and optimizing planting schedules to

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

align with precipitation patterns can support vegetation resilience in arid regions. By integrating data-driven insights and fostering community partnerships, afforestation programs like the BTAP can achieve more sustainable and lasting ecological outcomes.

5.1 | Recommendations

To enhance the effectiveness and sustainability of afforestation efforts, the following actions are recommended:

1. Intensify reforestation efforts in persistently barren areas by utilizing locally adapted species.
2. Implement enhanced monitoring and adaptive management strategies to ensure continuous sustainability.
3. Strengthen community engagement to align afforestation initiatives with socioeconomic needs.
4. Balance urbanization with ecological restoration through stringent land-use planning.
5. To refine afforestation strategies, integrate higher resolution satellite data and advanced machine-learning models in future research.

Author Contributions

Kaleem Mehmood: conceptualization (equal), data curation (equal), formal analysis (equal), investigation (equal), methodology (equal), software (equal), validation (equal), visualization (equal), writing – original draft (equal), writing – review and editing (equal). **Shoaib Ahmad Anees:** conceptualization (equal), data curation (equal), formal analysis (equal), investigation (equal), methodology (equal), software (equal), supervision (equal), validation (equal), visualization (equal), writing – original draft (equal), writing – review and editing (equal). **Sultan Muhammad:** writing – review and editing (equal). **Fahad Shahzad:** writing – review and editing (equal). **Qijing Liu:** writing – review and editing (equal). **Waseem Razzaq Khan:** formal analysis (equal), investigation (equal), writing – review and editing (equal). **Mansour Shrahili:** writing – review and editing (equal). **Mohammad Javed Ansari:** writing – review and editing (equal). **Timothy Dube:** writing – review and editing (equal).

Acknowledgments

We are grateful to the Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing 100083, China, for providing assistance and platforms for this research. We are also grateful to the Department of Forestry, The University of Agriculture, Dera Ismail Khan, 29050, Pakistan, for providing assistance and platforms for this research. This project was supported by Researchers Supporting Project number. RSP2025R464, King Saud University, Riyadh, Saudi Arabia. Authors also acknowledge the support of the Universiti Putra Malaysia.

Ethics Statement

The authors have nothing to report.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors confirm that the data links supporting the findings of this study are available within the article.

References

Aas, K., M. Jullum, and A. Løland. 2021. "Explaining Individual Predictions When Features Are Dependent: More Accurate Approximations to Shapley Values." *Artificial Intelligence* 298: 103502. https://doi.org/10.1016/j.artint.2021.103502.

Abatzoglou, J. T., S. Z. Dobrowski, S. A. Parks, and K. C. Hegewisch. 2018. "TerraClimate, a High-Resolution Global Dataset of Monthly Climate and Climatic Water Balance From 1958–2015." *Scientific Data* 5: 170191. https://doi.org/10.1038/sdata.2017.191.

Abdi, O. 2019. "Climate-Triggered Insect Defoliators and Forest Fires Using Multitemporal Landsat and TerraClimate Data in NE Iran: An Application of GEOBIA TreeNet and Panel Data Analysis." *Sensors* 19: 3965.

Ağbulut, Ü., A. E. Gürel, and Y. Biçen. 2021. "Prediction of Daily Global Solar Radiation Using Different Machine Learning Algorithms: Evaluation and Comparison." *Renewable and Sustainable Energy Reviews* 135: 110114. https://doi.org/10.1016/j.rser.2020.110114.

Ahmad, R., B. Yang, G. Ettlin, A. Berger, and P. Rodríguez-Bocca. 2023. "A Machine-Learning Based ConvLSTM Architecture for NDVI Forecasting." *International Transactions in Operational Research* 30: 2025–2048. https://doi.org/10.1111/itor.12887.

Ahmed, S. 2011. "Does Economic Geography Matter for Pakistan? A Spatial Exploratory Analysis of Income and Education Inequalities." *Pakistan Development Review* 50: 929–953. https://doi.org/10.30541/v50i4iipp.929-953.

Akram, M., U. Hayat, J. Shi, and S. A. Anees. 2022. "Association of the Female Flight Ability of Asian Spongy Moths (*Lymantria dispar asiatica*) With Locality, Age and Mating: A Case Study From China." *Forests* 13, no. 8: 1158. https://doi.org/10.3390/f13081158.

Ali, A., T. A. Khan, and S. Ahmad. 2018. "Analysis of Climate Data of Khyber Pakhtunkhwa, Pakistan." *International Research Journal of Engineering and Technology* 5: 4266–4283.

Andreevich, U. V., S. S. O. Reza, T. I. Stepanovich, et al. 2020. "Are There Differences in the Response of Natural Stand and Plantation Biomass to Changes in Temperature and Precipitation? A Case for Two-Needled Pines in Eurasia." *Journal of Resources and Ecology* 11, no. 4: 331. https://doi.org/10.5814/j.issn.1674-764x.2020.04.001.

Anees, M. M., D. Mann, M. Sharma, E. Banzhaf, and P. K. Joshi. 2020. "Assessment of Urban Dynamics to Understand Spatiotemporal Differentiation at Various Scales Using Remote Sensing and Geospatial Tools." *Remote Sensing* 12: 1306. https://doi.org/10.3390/RS12081306.

Anees, S. A., K. Mehmood, W. R. Khan, et al. 2024a. "Integration of Machine Learning and Remote Sensing for Above Ground Biomass Estimation Through Landsat-9 and Field Data in Temperate Forests of the Himalayan Region." *Ecological Informatics* 82: 102732. https://doi.org/10.1016/j.ecoinf.2024.102732.

Anees, S. A., K. Mehmood, A. Rehman, et al. 2024b. "Unveiling Fractional Vegetation Cover Dynamics: A Spatiotemporal Analysis Using MODIS NDVI and Machine Learning." *Environmental and Sustainability Indicators* 24: 100485. https://doi.org/10.1016/j.indic.2024.100485.

Anees, S. A., X. Yang, and K. Mehmood. 2024. "The Stoichiometric Characteristics and the Relationship With Hydraulic and Morphological Traits of the Faxon Fir in the Subalpine Coniferous Forest of Southwest

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

China." *Ecological Indicators* 159: 111636. https://doi.org/10.1016/j.ecolind.2024.111636.

Anees, S. A., X. Zhang, K. A. Khan, M. Abbas, H. A. Ghramh, and Z. Ahmad. 2022a. "Estimation of Fractional Vegetation Cover Dynamics and Its Drivers Based on Multi-Sensor Data in Dera Ismail Khan, Pakistan." *Journal of King Saud University, Science* 34, no. 6: 102217. https://doi.org/10.1016/j.jksus.2022.102217.

Anees, S. A., X. Zhang, M. Shakeel, et al. 2022b. "Estimation of Fractional Vegetation Cover Dynamics Based on Satellite Remote Sensing in Pakistan: A Comprehensive Study on the FVC and Its Drivers." *Journal of King Saud University, Science* 34, no. 3: 101848. https://doi.org/10.1016/j.jksus.2022.101848.

Aslam, M. S., P. Huanxue, S. Sohail, M. T. Majeed, S. U. Rahman, and S. A. Anees. 2022. "Assessment of Major Food Crops Production-Based Environmental Efficiency in China, India, and Pakistan." *Environmental Science and Pollution Research* 29: 1–10. https://doi.org/10.1007/s11356-021-16161-x.

Avci, C., M. Budak, N. Yagmur, and F. B. Balcik. 2023. "Comparison Between Random Forest and Support Vector Machine Algorithms for LULC Classification." *International Journal of Engineering and Geosciences* 8: 1–10. https://doi.org/10.26833/ijeg.987605.

Bacha, M. S., M. Muhammad, Z. Kılıç, and M. Nafees. 2021. "The Dynamics of Public Perceptions and Climate Change in Swat Valley, Khyber Pakhtunkhwa, Pakistan." *Sustainability* 13: 4464.

Badshah, M. T., K. Hussain, A. U. Rehman, et al. 2024. "The Role of Random Forest and Markov Chain Models in Understanding Metropolitan Urban Growth Trajectory." *Frontiers in Forests and Global Change* 7: 47. https://doi.org/10.3389/ffgc.2024.1345047.

Baldo, M., F. Buldrini, A. Chiarucci, et al. 2023. "Remote Sensing Analysis on Primary Productivity and Forest Cover Dynamics: A Western Ghats India Case Study." *Ecological Informatics* 73: 101922. https://doi.org/10.1016/j.ecoinf.2022.101922.

Ben Hamida, S., H. Mrabet, F. Chaieb, and A. Jemai. 2024. "Assessment of Data Augmentation, Dropout With L2 Regularization, and Differential Privacy Against Membership Inference Attacks." *Multimedia Tools and Applications* 83, no. 15: 44455–44484.

Bradshaw, T. J., Z. Huemann, J. Hu, and A. Rahmim. 2023. "A Guide to Cross-Validation for Artificial Intelligence in Medical Imaging." *Radiology: Artificial Intelligence* 5: e220232.

Brieva, C., P. M. Saco, S. G. Sandi, S. Mora, and J. F. Rodríguez. 2023. "NDVI Response to Satellite-Estimated Antecedent Precipitation in Dryland Pastures." *Remote Sensing* 15: 615. https://doi.org/10.3390/rs15143615.

Bui, Q. T., C. Jamet, V. Vantrepotte, X. Mériaux, A. Cauvin, and M. A. Mograne. 2022. "Evaluation of Sentinel-2/MSI Atmospheric Correction Algorithms Over Two Contrasted French Coastal Waters." *Remote Sensing* 14: 99. https://doi.org/10.3390/rs14051099.

Bunce, R. G. H., C. M. Wood, S. M. Smart, et al. 2014. "The Landscape Ecological Impact of Afforestation on the British Uplands and Some Initiatives to Restore Native Woodland Cover." *Journal of Landscape Ecology(Czech Republic)* 7: 5–24. https://doi.org/10.2478/jlecol-2014-0013.

Burke, T., C. S. Rowland, J. D. Whyatt, G. A. Blackburn, and J. Abbatt. 2023. "Spatially Targeting National-Scale Afforestation for Multiple Ecosystem Services." *Applied Geography* 159: 103064. https://doi.org/10.1016/j.apgeog.2023.103064.

Cao, S., L. Chen, D. Shankman, C. Wang, X. Wang, and H. Zhang. 2011. "Excessive Reliance on Afforestation in China's Arid and Semi-Arid Regions: Lessons in Ecological Restoration." *Earth-Science Reviews* 104: 240–245. https://doi.org/10.1016/j.earscirev.2010.11.002.

Castaldi, F., M. Halil Koparan, J. Wetterlind, et al. 2023. "Assessing the Capability of Sentinel-2 Time-Series to Estimate Soil Organic Carbon and Clay Content at Local Scale in Croplands." *ISPRS Journal of Photogrammetry and Remote Sensing* 199: 40–60. https://doi.org/10.1016/j.isprsjprs.2023.03.016.

Cavalli, A., S. Francini, R. E. McRoberts, et al. 2023. "Estimating Afforestation Area Using Landsat Time Series and Photointerpreted Datasets." *Remote Sensing* 15: 923. https://doi.org/10.3390/rs15040923.

Celik, M. F., M. S. Isik, O. Yuzugullu, N. Fajraoui, and E. Erten. 2022. "Soil Moisture Prediction From Remote Sensing Images Coupled With Climate, Soil Texture and Topography via Deep Learning." *Remote Sensing* 14: 584. https://doi.org/10.3390/rs14215584.

Chapagain, U., B. P. Chapagain, S. Nepal, and M. Manthey. 2021. "Impact of Disturbances on Species Diversity and Regeneration of Nepalese Sal (*Shorea robusta*) Forests Managed Under Different Management Regimes." *Earth (Switzerland)* 2: 826–844. https://doi.org/10.3390/earth2040049.

Chen, W. Y., C. Wang, and J. Jin. 2021. "Beijing Afforestation Project." In *Nature-Based Solutions for More Sustainable Cities – A Framework Approach for Planning and Evaluation*. Beijing: Emerald Publishing Limited. https://doi.org/10.1108/978-1-80043-636-720211027.

Chen, X., D. Xie, Z. Zhang, et al. 2023. "Compatible Biomass Model With Measurement Error Using Airborne LiDAR Data." *Remote Sensing* 15, no. 14: 3546. https://doi.org/10.3390/rs15143546.

Chen, X., and Y. Zhang. 2023. "Impacts of Climate, Phenology, Elevation and Their Interactions on the Net Primary Productivity of Vegetation in Yunnan, China Under Global Warming." *Ecological Indicators* 154: 110533. https://doi.org/10.1016/j.ecolind.2023.110533.

Chen, Y. 2012. "On the Four Types of Weight Functions for Spatial Contiguity Matrix." *Letters in Spatial and Resource Sciences* 5: 65–72. https://doi.org/10.1007/s12076-011-0076-6.

Chen, Y. 2021. "An Analytical Process of Spatial Autocorrelation Functions Based on Moran's Index." *PLoS One* 16: 589. https://doi.org/10.1371/journal.pone.0249589.

Chen, Y. 2023. "Spatial Autocorrelation Equation Based on Moran's Index." *Scientific Reports* 13: 19296. https://doi.org/10.1038/s41598-023-45947-x.

Cheng, L., C. Zhong, X. Li, M. Jia, Z. Wang, and D. Mao. 2022. "Rapid and Automatic Classification of Intertidal Wetlands Based on Intensive Time Series Sentinel-2 Images and Google Earth Engine." *National Remote Sensing Bulletin* 26: 348–357. https://doi.org/10.11834/jrs.20211311.

Colomina, I., and P. Molina. 2014. "Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review." *ISPRS Journal of Photogrammetry and Remote Sensing* 92: 79–97. https://doi.org/10.1016/j.isprsjprs.2014.02.013.

de Sousa, K., A. Sparks, W. Ashmall, J. van Etten, and S. Solberg. 2020. "Chirps: API Client for the CHIRPS Precipitation Data in R." *Journal of Open Source Software* 5: 419. https://doi.org/10.21105/joss.02419.

Doelman, J. C., E. Stehfest, D. P. van Vuuren, et al. 2020. "Afforestation for Climate Change Mitigation: Potentials, Risks and Trade-Offs." *Global Change Biology* 26: 1576–1591. https://doi.org/10.1111/gcb.14887.

Doulah Md, S. U., and N. Islam Md. 2023. "Performance Evaluation of Machine Learning Algorithm in Various Datasets." *Journal of Artificial Intelligence, Machine Learning and Neural Network* 3: 14–32. https://doi.org/10.55529/jaimlnn.32.14.32.

Emadi, M., R. Taghizadeh-Mehrjardi, A. Cherati, M. Danesh, A. Mosavi, and T. Scholten. 2020. "Predicting and Mapping of Soil Organic Carbon Using Machine Learning Algorithms in Northern Iran." *Remote Sensing* 12: 2234. https://doi.org/10.3390/rs12142234.

Ershov, D. V., E. A. Gavrilyuk, N. V. Koroleva, et al. 2022. "Natural Afforestation on Abandoned Agricultural Lands During Post-Soviet Period: A Comparative Landsat Data Analysis of Bordering Regions in Russia and Belarus." *Remote Sensing* 14: 322. https://doi.org/10.3390/rs14020322.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ESRI. 2013. *How Hot Spot Analysis: Getis-Ord Gi (Spatial Statistics) Works.* ArcGIS Resources*. Redlands, CA: Environmental Systems Research Institute, Inc.

ESRI. 2022. *How Emerging Hot Spot Analysis Works [WWW Document]*. Redlands, CA: Environmental Systems Research Institute, Inc.

Fassnacht, F. E., C. Schiller, T. Kattenborn, X. Zhao, and J. Qu. 2019. "A Landsat-Based Vegetation Trend Product of the Tibetan Plateau for the Time-Period 1990–2018." *Scientific Data* 6: 78. https://doi.org/10.1038/s41597-019-0075-9.

Feng, D., W. Bao, Y. Yang, and M. Fu. 2021. "How Do Government Policies Promote Greening? Evidence From China." *Land Use Policy* 104: 105389. https://doi.org/10.1016/j.landusepol.2021.105389.

Fernández, C. 2023. "Effects of Post-Fire Application of Straw Mulch Strips on Soil erosion, Soil Moisture and Vegetation Regeneration in European Dry Heathlands in NW Spain." *Ecological Engineering* 196: 107095. https://doi.org/10.1016/j.ecoleng.2023.107095.

Flores, A., M. E. Romero-Sánchez, R. Pérez-Miranda, T. Pineda-Ojeda, and F. Moreno-Sánchez. 2021. "Potential of Restoration of Coniferous Forests From Germplasm Transfer Zones in Mexico." *Revista Mexicana De Ciencias Forestales* 12: 813. https://doi.org/10.29298/rmcf.v12i63.813.

Furniss, T. J., V. R. Kane, A. J. Larson, and J. A. Lutz. 2020. "Detecting Tree Mortality With Landsat-Derived Spectral Indices: Improving Ecological Accuracy by Examining Uncertainty." *Remote Sensing of Environment* 237: 111497.

Galidaki, G., D. Zianis, I. Gitas, et al. 2017. "Vegetation Biomass Estimation With Remote Sensing: Focus on Forest and Other Wooded Land Over the Mediterranean Ecosystem." *International Journal of Remote Sensing* 38: 1940–1966.

Garik, G. 2021. *Esri Releases New 2020 Global Land Cover Map – Esri*. Redlands, CA: Esri.

Gaspard, G., D. Kim, and Y. Chun. 2019. "Residual Spatial Autocorrelation in Macroecological and Biogeographical Modeling: A Review." *Journal of Ecology and Environment* 43: 19. https://doi.org/10.1186/s41610-019-0118-3.

Gaughan, A. E., F. R. Stevens, C. Linard, P. Jia, and A. J. Tatem. 2013. "High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015." *PLoS One* 8: e55882. https://doi.org/10.1371/journal.pone.0055882.

Gebreyesus, Y., D. Dalton, S. Nixon, D. De Chiara, and M. Chinnici. 2023. "Machine Learning for Data Center Optimizations: Feature Selection Using Shapley Additive exPlanation (SHAP)." *Future Internet* 15: 88. https://doi.org/10.3390/fi15030088.

Gerber, F., and D. W. Nychka. 2021. "Parallel Cross-Validation: A Scalable Fitting Method for Gaussian Process Models." *Computational Statistics and Data Analysis* 155: 107113.

Getis, A., and J. K. Ord. 1992. "The Analysis of Spatial Association by Use of Distance Statistics." *Geographical Analysis* 24: 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x.

Getis, A., and J. K. Ord. 2010. "The Analysis of Spatial Association by Use of Distance Statistics." In *Advances in Spatial Science*, edited by L. Anselin and S. J. Rey, 127–145. Berlin: Springer. https://doi.org/10.1007/978-3-642-01976-0_10.

Gomes, E., M. Inácio, K. Bogdzevič, M. Kalinauskas, D. Karnauskaitė, and P. Pereira. 2021. "Future Scenarios Impact on Land Use Change and Habitat Quality in Lithuania." *Environmental Research* 197: 111101. https://doi.org/10.1016/j.envres.2021.111101.

Gomis-Cebolla, J., V. Rattayova, S. Salazar-Galán, and F. Francés. 2023. "Evaluation of ERA5 and ERA5-Land Reanalysis Precipitation Datasets Over Spain (1951–2020)." *Atmospheric Research* 284: 106606. https://doi.org/10.1016/j.atmosres.2023.106606.

Gong, H., J. Sardans, H. Huang, Z. Yan, Z. Wang, and J. Peñuelas. 2024. "Global Patterns and Controlling Factors of Tree Bark C: N: P stoichiometry in Forest Ecosystems Consistent With Biogeochemical Niche Hypothesis." *New Phytologist* 244, no. 4: 1303–1314.

Guo, S., B. Deng, C. Chen, et al. 2024. "Seeking in Ride-On-Demand Service: A Reinforcement Learning Model With Dynamic Price Prediction." *IEEE Internet of Things Journal* 11: 29890–29910.

Gupta, P., S. Kanga, and V. N. Mishra. 2024. "A Comparative Study and Machine Learning Enabled Efficient Classification for Multispectral Data in Agriculture." *Baghdad Science Journal* 21: 2462. https://doi.org/10.21123/bsj.2023.8952.

Haider, K., M. F. Khokhar, F. Chishtie, W. RazzaqKhan, and K. R. Hakeem. 2017. "Identification and Future Description of Warming Signatures Over Pakistan With Special Emphasis on Evolution of CO_2 Levels and Temperature During the First Decade of the Twenty-First Century." *Environmental Science and Pollution Research* 24: 7617–7629.

Hao, B., X. Xu, F. Wu, and L. Tan. 2022. "Long-Term Effects of Fire Severity and Climatic Factors on Post-Forest-Fire Vegetation Recovery." *Forests* 13: 883. https://doi.org/10.3390/f13060883.

Haq, M. A., A. Ahmed, I. Khan, et al. 2022. "Analysis of Environmental Factors Using AI and ML Methods." *Scientific Reports* 12: 13267. https://doi.org/10.1038/s41598-022-16665-7.

Horton, A. J., E. D. Lazarus, T. C. Hales, J. A. Constantine, M. W. Bruford, and B. Goossens. 2018. "Can Riparian Forest Buffers Increase Yields From Oil Palm Plantations?" *Earth's Future* 6: 1082–1096. https://doi.org/10.1029/2018EF000874.

Huang, Y., Z. X. Chen, T. Yu, X. Z. Huang, and X. F. Gu. 2018. "Agricultural Remote Sensing Big Data: Management and Applications." *Journal of Integrative Agriculture* 17: 1915–1931. https://doi.org/10.1016/S2095-3119(17)61859-8.

Huang, Z., K. Li, Y. Jiang, Z. Jia, L. Lv, and Y. Ma. 2024. "Graph Relearn Network: Reducing Performance Variance and Improving Prediction Accuracy of Graph Neural Networks." *Knowledge-Based Systems* 301: 112311.

Hussain, K., K. Mehmood, S. A. Anees, et al. 2024a. "Assessing Forest Fragmentation due to Land Use Changes From 1992 to 2023: A Spatio-Temporal Analysis Using Remote Sensing Data." *Heliyon* 10: e34710. https://doi.org/10.1016/j.heliyon.2024.e34710.

Hussain, K., K. Mehmood, S. Yujun, et al. 2024b. "Analysing LULC Transformations Using Remote Sensing Data: Insights From a Multilayer Perceptron Neural Network Approach." *Annals of GIS*: 1–28. https://doi.org/10.1080/19475683.2024.2343399.

Irons, J. R., J. L. Dwyer, and J. A. Barsi. 2012. "The Next Landsat Satellite: The Landsat Data Continuity Mission." *Remote Sensing of Environment* 122: 11–21. https://doi.org/10.1016/j.rse.2011.08.026.

Jallat, H., M. F. Khokhar, K. A. Kudus, et al. 2021. "Monitoring Carbon Stock and Land-Use Change in 5000-Year-Old Juniper Forest Stand of Ziarat, Balochistan, Through a Synergistic Approach." *Forests* 12, no. 1: 51.

Janeczko, E., R. Dabrowski, J. Budnicka-Kosior, and M. Woźnicka. 2019. "Influence of Urbanization Processes on the Dynamics and Scale of Spatial Transformations in the Mazowiecki Landscape Park." *Sustainability (Switzerland)* 11: 3007. https://doi.org/10.3390/su11113007.

JAXA. 2019. "ALOS Global Digital Surface Model "ALOS World 3D—30m" (AW3D30)." JAXA Data Report.

Jiang, C., Y. Wang, Z. Yang, and Y. Zhao. 2023. "Do Adaptive Policy Adjustments Deliver Ecosystem-Agriculture-Economy Co-Benefits in Land Degradation Neutrality Efforts? Evidence From Southeast Coast of China." *Environmental Monitoring and Assessment* 195, no. 10: 1215. https://doi.org/10.1007/s10661-023-11821-6.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

João, T., G. João, M. Bruno, and H. João. 2018. "Indicator-Based Assessment of Post-Fire Recovery Dynamics Using Satellite NDVI Time-Series." *Ecological Indicators* 89: 199–212. https://doi.org/10.1016/j.ecolind.2018.02.008.

Jun, M.-J. 2021. "A Comparison of a Gradient Boosting Decision Tree, Random Forests, and Artificial Neural Networks to Model Urban Land Use Changes: The Case of the Seoul Metropolitan Area." *International Journal of Geographical Information Science* 35: 2149–2167.

Kamal, A., A. Ali, and M. Yingjie. 2018. "Significance of Billion Tree Tsunami Afforestation Project and Legal Developments in Forest Sector of Pakistan." *International Journal of Law and Society* 1: 157–165.

Karlsen, S. R., L. Stendardi, H. Tømmervik, L. Nilsen, I. Arntzen, and E. J. Cooper. 2021. "Time-Series of Cloud-Free Sentinel-2 Ndvi Data Used in Mapping the Onset of Growth of Central Spitsbergen, Svalbard." *Remote Sensing* 13: 3031. https://doi.org/10.3390/rs13153031.

Kelemen, M., E. Vigorito, L. Fachal, C. A. Anderson, and C. Wallace. 2024. "ShaPRS: Leveraging Shared Genetic Effects Across Traits or Ancestries Improves Accuracy of Polygenic Scores." *American Journal of Human Genetics* 111, no. 6: 1006–1017. https://doi.org/10.1016/j.ajhg.2024.05.012.

Khan, N., S. J. Shah, T. Rauf, M. Zada, C. Yukun, and J. Harbi. 2019. "Socioeconomic Impacts of the Billion Trees Afforestation Program in Khyber Pakhtunkhwa Province (KPK), Pakistan." *Forests* 10: 703.

Khan, W. R., M. Nazre, S. Akram, et al. 2024. "Assessing the Productivity of the Matang Mangrove Forest Reserve: Review of One of the Best-Managed Mangrove Forests." *Forests* 15: 747. https://doi.org/10.3390/f15050747.

Khan, W. R., F. Rasheed, S. Z. Zulkifli, et al. 2020. "Phytoextraction Potential of *Rhizophora apiculata*: A Case Study in Matang Mangrove Forest Reserve, Malaysia." *Tropical Conservation Science* 13: 1940082920947344.

Kiani, B., M. Fatima, N. H. Amin, and A. Hesami. 2023. "Comparing Geospatial Clustering Methods to Study Spatial Patterns of Lung cancer Rates in Urban Areas: A Case Study in Mashhad, Iran." *GeoJournal* 88: 1659–1669. https://doi.org/10.1007/s10708-022-10707-3.

Kim, M., D. Kim, D. Jin, and G. Kim. 2023. "Application of Explainable Artificial Intelligence (XAI) in Urban Growth Modeling: A Case Study of Seoul Metropolitan Area, Korea." *Land (Basel)* 12: 1–17. https://doi.org/10.3390/land12020420.

Kim, W. 2024. "A Random Focusing Method With Jensen–Shannon Divergence for Improving Deep Neural Network Performance Ensuring Architecture Consistency." *Neural Processing Letters* 56, no. 4: 199.

Kong, W., X. Wei, Y. Wu, et al. 2022. "Afforestation Can Lower Microbial Diversity and Functionality in Deep Soil Layers in a Semiarid Region." *Global Change Biology* 28: 6086–6101. https://doi.org/10.1111/gcb.16334.

Kumar, V., and S. Agrawal. 2023. "A Multi-Layer Perceptron–Markov Chain Based LULC Change Analysis and Prediction Using Remote Sensing Data in Prayagraj District, India." *Environmental Monitoring and Assessment* 195: 619.

Kupssinskü, L. S., T. T. Guimarães, E. M. De Souza, et al. 2020. "A Method for Chlorophyll-a and Suspended Solids Prediction Through Remote Sensing and Machine Learning." *Sensors (Switzerland)* 20: 2125. https://doi.org/10.3390/s20072125.

Lange, M., B. Dechant, C. Rebmann, M. Vohland, M. Cuntz, and D. Doktor. 2017. "Validating MODIS and Sentinel-2 NDVI Products at a Temperate Deciduous Forest Site Using Two Independent Ground-Based Sensors." *Sensors (Switzerland)* 17: 1855. https://doi.org/10.3390/s17081855.

Larbi, I. 2023. "Land Use-Land Cover Change in the Tano Basin, Ghana and the Implications on Sustainable Development Goals." *Heliyon* 9: e14859. https://doi.org/10.1016/j.heliyon.2023.e14859.

Li, H., and J. Zhao. 2018. "Evaluation of the Newly Released Worldwide AW3D30 DEM Over Typical Landforms of China Using Two Global DEMs and ICESat/GLAS Data." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11: 4430–4440. https://doi.org/10.1109/JSTARS.2018.2874361.

Li, J., J. He, Y. Liu, et al. 2019. "Spatial Autocorrelation Analysis of Multi-Scale Damaged Vegetation in the Wenchuan Earthquake-Affected Area, Southwest China." *Forests* 10, no. 2: 195.

Li, Y., J. Chen, Q. Ma, H. K. Zhang, and J. Liu. 2018. "Evaluation of Sentinel-2A Surface Reflectance Derived Using Sen2Cor in North America." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11: 1997–2021. https://doi.org/10.1109/JSTARS.2018.2835823.

Liu, X., T. Yang, Q. Wang, F. Huang, and L. Li. 2018. "Dynamics of Soil Carbon and Nitrogen Stocks After Afforestation in Arid and Semi-Arid Regions: A meta-Analysis." *Science of the Total Environment* 618: 1658–1664. https://doi.org/10.1016/j.scitotenv.2017.10.009.

Liu, Y., M. Zhang, Y. Li, et al. 2023. "Influence of Nitrogen Fertilizer Application on Soil Acidification Characteristics of Tea Plantations in Karst Areas of Southwest China." *Agriculture (Switzerland)* 13: 849. https://doi.org/10.3390/agriculture13040849.

Lot, R., F. Pellegrini, Y. Shaidu, and E. Küçükbenli. 2020. "PANNA: Properties From Artificial Neural Network Architectures." *Computer Physics Communications* 256: 107402. https://doi.org/10.1016/j.cpc.2020.107402.

Luo, M., S. A. Anees, Q. Huang, et al. 2024. "Improving Forest Above-Ground Biomass Estimation by Integrating Individual Machine Learning Models." *Forests* 15, no. 6: 975. https://doi.org/10.3390/f15060975.

Ma, W., Z. Feng, Z. Cheng, S. Chen, and F. Wang. 2020. "Identifying Forest Fire Driving Factors and Related Impacts in China Using Random Forest Algorithm." *Forests* 11: 507. https://doi.org/10.3390/F11050507.

Madhiarasan, M., and M. Louzazni. 2022. "Analysis of Artificial Neural Network: Architecture, Types, and Forecasting Applications." *Journal of Electrical and Computer Engineering* 2022: 1–23. https://doi.org/10.1155/2022/5416722.

Marino, S. 2023. "Understanding the Spatio-Temporal Behaviour of the Sunflower Crop for Subfield Areas Delineation Using Sentinel-2 NDVI Time-Series Images in an Organic Farming System." *Heliyon* 9: e19507. https://doi.org/10.1016/j.heliyon.2023.e19507.

Marlina, D. 2022. "Klasifikasi Tutupan Lahan pada Citra Sentinel-2 Kabupaten Kuningan dengan NDVI dan Algoritme Random Forest." *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)* 7: 41. https://doi.org/10.30998/string.v7i1.12948.

Mehmood, K., S. A. Anees, M. Luo, et al. 2024a. "Assessing Chilgoza Pine (*Pinus gerardiana*) Forest Fire Severity: Remote Sensing Analysis, Correlations, and Predictive Modeling for Enhanced Management Strategies." *Trees, Forests and People* 16: 100521. https://doi.org/10.1016/j.tfp.2024.100521.

Mehmood, K., S. A. Anees, S. Muhammad, et al. 2024b. "Analyzing Vegetation Health Dynamics Across Seasons and Regions Through NDVI and Climatic Variables." *Scientific Reports* 14: 11775. https://doi.org/10.1038/s41598-024-62464-7.

Mehmood, K., S. A. Anees, A. Rehman, et al. 2024c. "Exploring Spatiotemporal Dynamics of NDVI and Climate-Driven Responses in Ecosystems: Insights for Sustainable Management and Climate Resilience." *Ecological Informatics* 80: 102532. https://doi.org/10.1016/J.ECOINF.2024.102532.

Mehmood, K., S. A. Anees, A. Rehman, et al. 2024d. "Assessment of Climatic Influences on Net Primary Productivity Along Elevation Gradients in Temperate Ecoregions." *Trees, Forests and People* 18: 100657. https://doi.org/10.1016/j.tfp.2024.100657.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Mehmood, K., S. A. Anees, A. Rehman, et al. 2024e. "Assessing Forest Cover Changes and Fragmentation in the Himalayan Temperate Region: Implications for Forest Conservation and Management." *Journal of Forestry Research (Harbin)* 35: 82. https://doi.org/10.1007/s11676-024-01734-6.

Mielke, K. P., T. Claassen, M. Busana, et al. 2020. "Disentangling Drivers of Spatial Autocorrelation in Species Distribution Models." *Ecography* 43: 1741–1751. https://doi.org/10.1111/ecog.05134.

Miroshnyk, N. V., A. F. Likhanov, T. O. Grabovska, I. K. Teslenko, and H. Roubík. 2022. "Green Infrastructure and Relationship With Urbanization – Importance and Necessity of Integrated Governance." *Land Use Policy* 114: 105941. https://doi.org/10.1016/j.landusepol.2021.105941.

Mohiuddin, Y. N. 2007. *Pakistan: A Global Studies Handbook*. USA: ABC-CLIO.

Muhammad, S. 2023. "Analyzing the Impact of Forest Harvesting Ban in Northern Temperate Forest. A Case Study of Anakar Valley, Kalam Swat Region, Khyber-Pakhtunkhwa, Pakistan." *Pure and Applied Biology* 12: 143. https://doi.org/10.19045/bspab.2023.120143.

Muñoz-Sabater, J., E. Dutra, A. Agustí-Panareda, et al. 2021. "ERA5-Land: A State-Of-The-Art Global Reanalysis Dataset for Land Applications." *Earth System Science Data* 13: 4349–4383.

Nadakinamani, R. G., A. Reyana, S. Kautish, et al. 2022. "Clinical Data Analysis for Prediction of Cardiovascular Disease Using Machine Learning Techniques." *Computational Intelligence and Neuroscience* 2022: 1–13. https://doi.org/10.1155/2022/2973324.

Nave, L. E., C. W. Swanston, U. Mishra, and K. J. Nadelhoffer. 2013. "Afforestation Effects on Soil Carbon Storage in the United States: A Synthesis." *Soil Science Society of America Journal* 77: 1035–1047. https://doi.org/10.2136/sssaj2012.0236.

Nazir, N., A. Farooq, S. Ahmad Jan, and A. Ahmad. 2019. "A System Dynamics Model for Billion Trees Tsunami Afforestation Project of Khyber Pakhtunkhwa in Pakistan: Model Application to Afforestation Activities." *Journal of Mountain Science* 16: 2640–2653. https://doi.org/10.1007/s11629-018-5076-1.

Nihar, A., N. R. Patel, and A. Danodia. 2022. "Machine-Learning-Based Regional Yield Forecasting for Sugarcane Crop in Uttar Pradesh, India." *Journal of the Indian Society of Remote Sensing* 50: 1519–1530. https://doi.org/10.1007/s12524-022-01549-0.

Nordin, N., Z. Zainol, M. H. Mohd Noor, and L. F. Chan. 2023. "An Explainable Predictive Model for Suicide Attempt Risk Using an Ensemble Learning and Shapley Additive Explanations (SHAP) Approach." *Asian Journal of Psychiatry* 79: 103316. https://doi.org/10.1016/j.ajp.2022.103316.

Okunev, I. Y., and A. E. Kushnareva. 2023. "Alternative Spatial Weights Matrices: Methodology and Application in Calculating LISA." *Vestnik of Saint Petersburg University. Earth Sciences* 68: 210. https://doi.org/10.21638/spbu07.2023.210.

Ord, J. K., and A. Getis. 1995. "Local Spatial Autocorrelation Statistics: Distributional Issues and an Application." *Geographical Analysis* 27: 286–306. https://doi.org/10.1111/j.1538-4632.1995.tb00912.x.

Otkin, J. A., Y. Zhong, E. D. Hunt, et al. 2019. "Assessing the Evolution of Soil Moisture and Vegetation Conditions During a Flash Drought-Flash Recovery Sequence Over the South-Central United States." *Journal of Hydrometeorology* 20: 549–562. https://doi.org/10.1175/JHM-D-18-0171.1.

Ouchra, H., A. Belangour, and A. Erraissi. 2023. "Machine Learning Algorithms for Satellite Image Classification Using Google Earth Engine and Landsat Satellite Data: Morocco Case Study." *IEEE Access* 11: 71127–71142. https://doi.org/10.1109/ACCESS.2023.3293828.

Pan, Q., F. Harrou, and Y. Sun. 2023. "A Comparison of Machine Learning Methods for Ozone Pollution Prediction." *Journal of Big Data* 10: 63. https://doi.org/10.1186/s40537-023-00748-x.

Pan, S. A., S. A. Anees, X. Li, X. Yang, X. Duan, and Z. Li. 2023. "Spatial and Temporal Patterns of Non-Structural Carbohydrates in Faxon Fir (Abies Fargesii Var. Faxoniana), Subalpine Mountains of Southwest China." *Forests* 14, no. 7: 1438. https://doi.org/10.3390/f14071438.

Pang, A., M. W. L. Chang, and Y. Chen. 2022. "Evaluation of Random Forests (RF) for Regional and Local-Scale Wheat Yield Prediction in Southeast Australia." *Sensors* 22: 717. https://doi.org/10.3390/s22030717.

Pomme, L. E., R. Bourqui, R. Giot, and D. Auber. 2022. "Relative Confusion Matrix: Efficient Comparison of Decision Models." In *Proceedings of the International Conference on Information Visualisation*. Vienna, Austria: IEEE. https://doi.org/10.1109/IV56949.2022.00025.

Qasimi, A. B., V. Isazade, G. Kaplan, and Z. Nadry. 2023. "Spatiotemporal and Multi-Sensor Analysis of Surface Temperature, NDVI, and Precipitation Using Google Earth Engine Cloud Computing Platform." *Russian Journal of Earth Sciences* 22: 1–12. https://doi.org/10.2205/2022ES000812.

Qiu, C., L. Mou, M. Schmitt, and X. X. Zhu. 2019. "Local Climate Zone-Based Urban Land Cover Classification From Multi-Seasonal Sentinel-2 Images With a Recurrent Residual Network." *ISPRS Journal of Photogrammetry and Remote Sensing* 154: 151–162. https://doi.org/10.1016/j.isprsjprs.2019.05.004.

Rai, M. K., B. Paudel, Y. Zhang, et al. 2023. "Appraisal of Empirical Studies on Land-Use and Land-Cover Changes and Their Impact on Ecosystem Services in Nepal Himalaya." *Sustainability (Switzerland)* 15: 7134. https://doi.org/10.3390/su15097134.

Ren, Y., F. Ling, and Y. Wang. 2023. "Research on Provincial-Level Soil Moisture Prediction Based on Extreme Gradient Boosting Model." *Agriculture (Switzerland)* 13: 927. https://doi.org/10.3390/agriculture13050927.

Román-Dañobeytia, F., M. Huayllani, A. Michi, et al. 2015. "Reforestation With Four Native Tree Species After Abandoned Gold Mining in the Peruvian Amazon." *Ecological Engineering* 85: 39–46. https://doi.org/10.1016/j.ecoleng.2015.09.075.

Roy, B. 2021. "Optimum Machine Learning Algorithm Selection for Forecasting Vegetation Indices: MODIS NDVI & EVI." *Remote Sens Appl* 23: 100582. https://doi.org/10.1016/j.rsase.2021.100582.

Rumora, L., M. Miler, and D. Medak. 2020. "Impact of Various Atmospheric Corrections on Sentinel-2 Land Cover Classification Accuracy Using Machine Learning Classifiers." *ISPRS International Journal of Geo-Information* 9: 277. https://doi.org/10.3390/ijgi9040277.

Sadhwani, K., T. I. Eldho, and S. Karmakar. 2023. "Investigating the Influence of Future Landuse and Climate Change on Hydrological Regime of a Humid Tropical River Basin." *Environment and Earth Science* 82: 1–19. https://doi.org/10.1007/s12665-023-10891-6.

Sanusi, W., S. Sidjara, S. Patahuddin, and M. Danial. 2024. "A Comparison of Spatial Interpolation Methods for Regionalizing Maximum Daily Rainfall Data in South Sulawesi, Indonesia." *ITM Web of Conferences* 58: 4003. https://doi.org/10.1051/itmconf/20245804003.

Self, S., A. Overby, A. Zgodic, D. White, A. McLain, and C. Dyckman. 2023. "A Hypothesis Test for Detecting Distance-Specific Clustering and Dispersion in Areal Data." *Spatial Statistics* 55: 100757. https://doi.org/10.1016/j.spasta.2023.100757.

Shahade, A. K., K. H. Walse, V. M. Thakare, and M. Atique. 2023. "Multi-Lingual Opinion Mining for Social Media Discourses: An Approach Using Deep Learning Based Hybrid Fine-Tuned Smith Algorithm With Adam Optimizer." *International Journal of Information Management Data Insights* 3: 100182. https://doi.org/10.1016/j.jjimei.2023.100182.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Shahzad, F., K. Mehmood, K. Hussain, et al. 2024. "Comparing Machine Learning Algorithms to Predict Vegetation Fire Detections in Pakistan." *Fire Ecology* 20: 57. https://doi.org/10.1186/s42408-024-00289-5.

Shawky, M., M. R. Ahmed, E. Ghaderpour, et al. 2023. "Remote Sensing-Derived Land Surface Temperature Trends Over South Asia." *Ecological Informatics* 74: 101969. https://doi.org/10.1016/j.ecoinf.2022.101969.

Shen, W., M. Li, C. Huang, X. Tao, S. Li, and A. Wei. 2019. "Mapping Annual Forest Change due to Afforestation in Guangdong Province of China Using Active and Passive Remote Sensing Data." *Remote Sensing* 11: 490. https://doi.org/10.3390/rs11050490.

Shobairi, S. O. R., H. Lin, V. A. Usoltsev, et al. 2022. "A Comparative Pattern for *Populus* spp. and *Betula* spp. Stand Biomass in Eurasian Climate Gradients." *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering* 43, no. 2: 457–467. https://doi.org/10.5552/crojfe.2022.1340.

Shoko, C., O. Mutanga, and T. Dube. 2018. "Determining Optimal New Generation Satellite Derived Metrics for Accurate C3 and C4 Grass Species Aboveground Biomass Estimation in South Africa." *Remote Sensing* 10: 564. https://doi.org/10.3390/rs10040564.

Smith, B., N. Pahlevan, J. Schalles, et al. 2021. "A Chlorophyll-a Algorithm for Landsat-8 Based on Mixture Density Networks." *Frontiers in Remote Sensing* 1: 623678.

Sohail, M., K. Mehmood, S. A. Anees, et al. 2023. "Tourism, Threat, and Opportunities for the Forest Resources: A Case Study of Gabin Jabaa, District Swat, Khyber-Pakhtunkhwa, Pakistan." *International Journal of Forest Sciences* 3: 194–203.

Song, S., Y. Ding, W. Li, et al. 2023. "Mangrove Reforestation Provides Greater Blue Carbon Benefit Than Afforestation for Mitigating Global Climate Change." *Nature Communications* 14: 756. https://doi.org/10.1038/s41467-023-36477-1.

Stevens, F. R., A. E. Gaughan, C. Linard, and A. J. Tatem. 2015. "Disaggregating Census Data for Population Mapping Using Random Forests With Remotely-Sensed and Ancillary Data." *PLoS One* 10: 42. https://doi.org/10.1371/journal.pone.0107042.

Suarez, L. A., A. Robson, and J. Brinkhoff. 2023. "Early-Season Forecasting of citrus Block-Yield Using Time Series Remote Sensing and Machine Learning: A Case Study in Australian Orchards." *International Journal of Applied Earth Observation and Geoinformation* 122: 103434. https://doi.org/10.1016/j.jag.2023.103434.

Tajik, S., S. Ayoubi, and M. Zeraatpisheh. 2020. "Digital Mapping of Soil Organic Carbon Using Ensemble Learning Model in Mollisols of Hyrcanian Forests, Northern Iran." *Geoderma Regional* 20: e00256. https://doi.org/10.1016/j.geodrs.2020.e00256.

Tamiminia, H., B. Salehi, M. Mahdianpari, and T. Goulden. 2024. "State-Wide Forest Canopy Height and Aboveground Biomass Map for New York With 10 m Resolution, Integrating GEDI, Sentinel-1, and Sentinel-2 Data." *Ecological Informatics* 79: 102404.

TerraClimate. 2022. *TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces. Earth Engine Data Catalog.* Moscow, ID: University of Idaho.

Ul-Haq, Z., S. Fazid, N. Sultana, et al. 2019. "Khyber Pakhtunkhwa Integrated Population Health Survey (KP-IPHS) 2016–17." *Khyber Medical University Journal* 11: 254–257.

Ullah, A., M. Bavorova, A. A. Shah, and G. P. Kandel. 2023. "Community Participation in Development Programs: Key Lessons From the Billion Trees Afforestation Project (BTAP)." *Environmental Science & Policy* 150: 103581. https://doi.org/10.1016/j.envsci.2023.103581.

Ullah, A., A. S. Sam, A. R. Sathyan, N. Mahmood, A. Zeb, and H. Kächele. 2021. "Role of Local Communities in Forest Landscape Restoration: Key Lessons From the Billion Trees Afforestation Project, Pakistan." *Science of the Total Environment* 772: 145613. https://doi.org/10.1016/j.scitotenv.2021.145613.

Ullah, I., A. Saleem, L. Ansari, et al. 2020. "Growth and Survival of Multipurpose Species; Assessing Billion Tree Afforestation Project (BTAP), the Bonn Challenge Initiative." *Applied Ecology and Environmental Research* 18: 2057–2072. https://doi.org/10.15666/aeer/1802_20572072.

University of Idaho. 2019. *TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces. Earth Engine Data Catalog.* Moscow, ID: University of Idaho.

USGS. 2022. *What Is the Landsat Satellite Program and Why Is It Important?* Reston, VA: U.S. Geological Survey.

Usoltsev, V. A., B. Chen, S. O. R. Shobairi, I. S. Tsepordey, V. P. Chasovskikh, and S. A. Anees. 2020. "Patterns for Populus Spp. Stand Biomass in Gradients of Winter Temperature and Precipitation of Eurasia." *Forests* 11, no. 9: 906. https://doi.org/10.3390/f11090906.

Usoltsev, V. A., H. Lin, S. O. R. Shobairi, I. S. Tsepordey, Z. Ye, and S. A. Anees. 2022. "The Principle of Space-For-Time Substitution in Predicting Betula Spp. Biomass Change Related to Climate Shifts." *Applied Ecology and Environmental Research* 20, no. 4: 3683–3698. https://doi.org/10.15666/aeer/2004_36833698.

Valero, S., L. Arnaud, M. Planells, and E. Ceschia. 2021. "Synergy of Sentinel-1 and Sentinel-2 Imagery for Early Seasonal Agricultural Crop Mapping." *Remote Sensing* 13: 891. https://doi.org/10.3390/rs13234891.

Vargas Godoy, M. R., and Y. Markonis. 2023. "pRecipe: A Global Precipitation Climatology Toolbox and Database." *Environmental Modelling and Software* 165: 105711. https://doi.org/10.1016/j.envsoft.2023.105711.

Vujović, Ž. 2021. "Classification Model Evaluation Metrics." *International Journal of Advanced Computer Science and Applications* 12: 670. https://doi.org/10.14569/IJACSA.2021.0120670.

Wang, S., and F. Fan. 2021. "Analysis of the Response of Long-Term Vegetation Dynamics to Climate Variability Using the Pruned Exact Linear Time (PELT) Method and Disturbance Lag Model (DLM) Based on Remote Sensing Data: A Case Study in Guangdong Province (China)." *Remote Sensing* 13, no. 10: 1873.

Wang, Y., M. Li, and G. Jin. 2024. "Exploring the Optimization of Spatial Patterns for Carbon Sequestration Services Based on Multi-Scenario Land Use/Cover Changes in the Changchun-Jilin-Tumen Region, China." *Journal of Cleaner Production* 438: 140788. https://doi.org/10.1016/j.jclepro.2024.140788.

Wang, Y. B., M. Junaid, J. Y. Deng, et al. 2024. "Effects of Land-Use Patterns on Seasonal Water Quality at Multiple Spatial Scales in the Jialing River, Chongqing, China." *Catena (Amst)* 234: 107646. https://doi.org/10.1016/j.catena.2023.107646.

Wang, L., Y. Lu, and Y. Yao. 2019. "Comparison of Three Algorithms for the Retrieval of Land Surface Temperature From Landsat 8 Images." *Sensors* 19, no. 22: 5049.

Wang, Y.-R., D. O. Hessen, B. H. Samset, and F. Stordal. 2022. "Evaluating Global and Regional Land Warming Trends in the Past Decades With Both MODIS and ERA5-Land Land Surface Temperature Data." *Remote Sensing of Environment* 280: 113181.

Warner, E., O. T. Lewis, N. Brown, et al. 2022. "Does Restoring Native Forest Restore Ecosystem Functioning? Evidence From a Large-Scale Reforestation Project in the Scottish Highlands." *Restoration Ecology* 30: 530. https://doi.org/10.1111/rec.13530.

Wibowo, Y., A. S. D. Laksana, J. Mulyanto, M. A. Wicaksono, and A. Y. Purnomo. 2015. "Incidence of Malaria Is Clustered and Buffers Around Plantations: A Spatial Analysis." *Universa Medicina* 34: 138. https://doi.org/10.18051/univmed.2015.v34.138-148.

Wijaya, B. S., R. Munir, and N. P. Utama. 2023. "Curating Multimodal Satellite Imagery for Precision Agriculture Datasets With Google Earth Engine." Proceedings of the International Conference on Data Science and Official Statistics 2023. https://doi.org/10.34123/icdsos.v2023i1.399.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Wu, X., S. Wang, B. Fu, and J. Liu. 2021. "Spatial Variation and Influencing Factors of the Effectiveness of Afforestation in China's Loess Plateau." *Science of the Total Environment* 771: 144904. https://doi.org/10.1016/j.scitotenv.2020.144904.

Xiao, Y., Q. Xiao, and X. Sun. 2020. "Ecological Risks Arising From the Impact of Large-Scale Afforestation on the Regional Water Supply Balance in Southwest China." *Scientific Reports* 10: 4150. https://doi.org/10.1038/s41598-020-61108-w.

Xie, D., H. Huang, L. Feng, et al. 2023. "Aboveground Biomass Prediction of Arid Shrub-Dominated Community Based on Airborne LiDAR Through Parametric and Nonparametric Methods." *Remote Sensing* 15, no. 13: 3344. https://doi.org/10.3390/rs15133344.

Xinde, L. I., F. Dunkin, and J. Dezert. 2023. "Multi-Source Information Fusion: Progress and Future." *Chinese Journal of Aeronautics* 37, no. 7: 24–58.

Xu, H., Q. Li, and J. Chen. 2022. "Highlight Removal From A Single Grayscale Image Using Attentive GAN." *Applied Artificial Intelligence* 36, no. 1: 1988441. https://doi.org/10.1080/08839514.2021.1988441.

Yan, L., D. P. Roy, A. Promkhambut, J. Fox, and Y. Zhai. 2022. "Automated Extraction of Aquaculture Ponds From Sentinel-2 Seasonal Imagery – A Validated Case Study in Central Thailand." *Science of Remote Sensing* 6: 100063. https://doi.org/10.1016/j.srs.2022.100063.

Yana, Y., and S. R. Rahayu. 2017. "Analisis Spasial Faktor Lingkungan dan Distribusi Kasus Demam Berdarah Dengue." *HIGEIA (Journal of Public Health Research and Development)* 1: 106–116.

Yang, K., Y. Luo, M. Li, S. Zhong, Q. Liu, and X. Li. 2022. "Reconstruction of Sentinel-2 Image Time Series Using Google Earth Engine." *Remote Sensing* 14: 4395. https://doi.org/10.3390/rs14174395.

Yao, Z., J. Xiao, and X. Ma. 2021. "The Impact of Large-Scale Afforestation on Ecological Environment in the Gobi Region." *Scientific Reports* 11: 14383. https://doi.org/10.1038/s41598-021-93948-5.

Younisse, R., A. Ahmad, and Q. Abu Al-Haija. 2022. "Explaining Intrusion Detection-Based Convolutional Neural Networks Using Shapley Additive Explanations (SHAP)." *Big Data and Cognitive Computing* 6: 126. https://doi.org/10.3390/bdcc6040126.

Yu, Z., X. Liu, J. Zhang, D. Xu, and S. Cao. 2018. "Evaluating the Net Value of Ecosystem Services to Support Ecological Engineering: Framework and a Case Study of the Beijing Plains Afforestation Project." *Ecological Engineering* 112: 148–152. https://doi.org/10.1016/j.ecoleng.2017.12.017.

Zha, H., Y. Miao, T. Wang, et al. 2020. "Improving Unmanned Aerial Vehicle Remote Sensing-Based Rice Nitrogen Nutrition Index Prediction With Machine Learning." *Remote Sensing* 12: 215. https://doi.org/10.3390/rs12020215.

Zhang, B., H. Zhu, W. Song, et al. 2024a. "A Multi-Baseline Forest Height Estimation Method Combining Analytic and Geometric Expression of the RVoG Model." *Forests* 15, no. 9: 1496. https://doi.org/10.3390/f15091496.

Zhang, X., Y. Sun, W. Jia, F. Wang, H. Guo, and Z. Ao. 2022. "Research on the Temporal and Spatial Distributions of Standing Wood Carbon Storage Based on Remote Sensing Images and Local Models." *Forests* 13, no. 2: 346.

Zhang, X., M. Usman, A. U. R. Irshad, M. Rashid, and A. Khattak. 2024b. "Investigating Spatial Effects Through Machine Learning and Leveraging Explainable AI for Child Malnutrition in Pakistan." *ISPRS International Journal of Geo-Information* 13, no. 9: 330.

Zhang, Z. 2019. "Improved Adam Optimizer for Deep Neural Networks." In *2018 IEEE/ACM 26th International Symposium on Quality of Service, IWQoS 2018*. Canada: IEEE. https://doi.org/10.1109/IWQoS.2018.8624183.

Zhang, Y., Z. Gao, X. Wang, and Q. Liu. 2023. "Image Representations of Numerical Simulations for Training Neural Networks." *Computer Modeling in Engineering and Sciences* 134, no. 2: 821–833.

Zheng, C., A. Abd-Elrahman, and V. Whitaker. 2021. "Remote Sensing and Machine Learning in Crop Phenotyping and Management, With an Emphasis on Applications in Strawberry Farming." *Remote Sensing* 13, no. 3: 531.

Zheng, L., J. Lu, H. Liu, X. Chen, and H. Yesou. 2024. "Evidence of Vegetation Greening Benefitting From the Afforestation Initiatives in China." *Geo-spatial Information Science* 27: 683–702. https://doi.org/10.1080/10095020.2023.2238782.

Zhou, R., X. Ci, J. Hu, et al. 2023. "Transitional Areas of Vegetation as Biodiversity Hotspots Evidenced by Multifaceted Biodiversity Analysis of a Dominant Group in Chinese Evergreen Broad-Leaved Forests." *Ecological Indicators* 147: 110001. https://doi.org/10.1016/j.ecolind.2023.110001.

Zhu, J. J., X. Zheng, G. G. Wang, et al. 2017. "Assessment of the World Largest Afforestation Program: Success, Failure, and Future Directions." *bioRxiv*. https://doi.org/10.1101/105619.

Zhu, X., H. Liu, Y. Li, and B. Liang. 2021. "Quantifying the Role of Soil in Local Precipitation Redistribution to Vegetation Growth." *Ecological Indicators* 124: 107355. https://doi.org/10.1016/j.ecolind.2021.107355.

Ziaul Hoque, M., I. Islam, M. Ahmed, S. Shamim Hasan, and F. Ahmed Prodhan. 2022. "Spatio-Temporal Changes of Land Use Land Cover and Ecosystem Service Values in Coastal Bangladesh." *Egyptian Journal of Remote Sensing and Space Science* 25: 173–180. https://doi.org/10.1016/j.ejrs.2022.01.008.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

20457758, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ece3.70736 by National Institutes Of Health Malaysia, Wiley Online Library on [05/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License