Particle Swarm Optimization for Woody Vegetation Assessment in a Semi-Arid Savannah Ecosystem

Nesisa Analisa Nyathi[1,*], Juliane Krenz[1], Andrea Meier[1], Brigitte Kuhn[1], Nikolaus J. Kuhn[1]

[1]Physical Geography and Environmental Change Research Group,
Faculty of Philosophy and Natural Sciences,
University of Basel, Basel, 4056, Switzerland
[*]Corresponding Author: nesisa.nyathi@unibas.ch

Keywords: Particle Swarm Optimization, Vegetation Indices, Remote Sensing, Woody Vegetation, Semi-arid Savannah, Habitat Quality

Abstract

This study explores the application of Particle Swarm Optimization (PSO) to enhance vegetation indices (VIs) for the assessment of woody vegetation in a semi-arid savannah ecosystem. By optimizing VIs, the research aims to improve the discrimination between vegetated and non-vegetated areas, facilitating a more accurate random forest classification for habitat quality assessment. The optimization process preserves minimum VI values across different sensors to maintain lower bounds of reflectance, ensuring ecologically valid signals are represented, particularly in low-vegetated areas. Results indicate that maximum VI values increase post-optimization, enhancing sensitivity to canopy vigor, stress, health, and presence. The study highlights the effectiveness of UAV-derived indices, such as NDVI, NDRE, and SAVI, in capturing the dynamics of vegetation health and dryness, thereby contributing valuable insights into remote sensing methodologies for ecological monitoring.

1. Introduction

Protected areas in southern Africa are mostly characterized by semi-arid heterogeneous and complex woody vegetation (Madonsela et al., 2018). These natural environments, particularly within protected areas, are predominantly populated by large woody vegetation species that play a critical role in maintaining ecosystem services such as habitat quality (Muposhi et al., 2014) and carbon sequestration (Gebre et al., 2019). Thus, maintaining the heterogeneity and complexity of the woody vegetation composition and its functions towards the ecosystem is important (Fundisi et al., 2020). However, despite their ecological importance, these vegetation communities are increasingly threatened by intensive grazing from large herbivore populations (Chafota, 1998; Druce et al., 2008), the impacts of climate change, human settlements (Dube and Nhamo, 2020) and anthropogenic pressures near water sources. These pressures often lead to the degradation of micro biomes (Coleine et al., 2024), ultimately compromising the overall habitat health and quality of these areas. While efforts have been made to mitigate these impacts (Wolmer, 2003), effective management and conservation strategies remain heavily dependent on the availability of up-to-date woody vegetation dynamics data and comprehensive field inventories (Madonsela et al., 2017). Moreover, traditional field-based methods for assessing habitat quality over time are costly, labour intensive and limited in both spatial extent and temporal frequency (Cui et al., 2022).

Remote sensing methods have the potential for monitoring habitat quality dynamics across spatial, spectral and temporal resolution of the dataset (Senf, 2022). While these technologies enable large-scale assessments, their capacity for cost effective and detailed quantification of vegetation dynamics at species or structural level remains limited (Arvor et al., 2013). To this end, numerous machine learning algorithms such as random forest (Svoboda et al., 2022), neural networks (Flood et al., 2019), maximum likelihood (Zagajewski et al., 2021), support vector machine (Acevedo and Jones, 2012) and boosted regression tree (Batta, 2018), have been used to extract woody vegetation dynamics using remotely sensed data for habitat quality assessments. Although these approaches have demonstrated success in specific case studies, they are often complex and limited to mapping specific vegetation dynamics and habitat types as well as transitioning zones (Fakhri et al., 2022a) for ecological assessments. Consequently, these methods can underestimate the rate at which the ecosystem is degrading or fragmenting at woody vegetation level (Pfeifer et al., 2015). As an alternative, statistical modelling approaches that use vegetation indices (VI) (Storch et al., 2018; Shumi et al., 2021; Peña-Lara et al., 2022; Chapungu et al., 2020) have gained traction for assessing vegetation health and vigour, thus enabling habitat quality assessments across different earth observation scales (Levin et al., 2007).

Vegetation indices have been widely applied in evaluating vegetation conditions and monitoring ecosystem dynamics especially for large areas (Yue et al., 2025). Each vegetation index captures a finite amount of variation within a given image by incorporating a specific range of pixels that quantify key biophysical parameters (Fakhri et al., 2019). When applied in a multitemporal scale and context, these indices can effectively describe and monitor changes in vegetation dynamics over time. VIs are derived from spectral reflectance values, typically by combining reflectance measurements from two or more sections of the electromagnetic spectrum (Xue and Su, 2017). The most commonly used bands for analysis are the red, near-infrared (NIR) and shortwave infrared. These bands use the differences in how vegetation reflects and absorbs light in these specific regions; VIs enhance the detection and interpretation of vegetation characteristics (Darra et al., 2025). Additionally, they also help to minimize the influence of external factors such as soil, atmospheric conditions and sensor variations, leading to more accurate and consistent assessments of vegetation cover

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-3-2026-583-2026

and conditions (Delaney et al., 2025).

Studies (Mogashoa et al., 2021; Zhu et al., 2022; Taddeo et al., 2019; Xue and Su, 2017) have noted that most VIs improve the sensitivity of green vegetation in most cases; however, different environments have specific characteristics that need to be accounted for when using VIs. For example, (Fakhri et al., 2019) emphasized the importance of enhancing different satellite bands to improve the accuracy of the VIs in vegetation mapping in semi-arid areas. Likewise, (Fundisi et al., 2020) cautioned against the reliance on VIs in vegetation mapping as a low index value may be misinterpreted as a low vegetation diversity value. (Levin et al., 2007) used Landsat and Quickbird imagery to predict mountain plant richness and rarity using the normalized difference vegetation index (NDVI) and soil adjusted vegetation index (SAVI) on Mount Hermon, Israel. Here, NDVI values for the summer months revealed a strong relationship between the image composition and plant species richness as sampled in the field compared to the SAVI.

For different contexts, the effectiveness of a given vegetation index is highly dependent on the parameterization of its coefficients (Gao et al., 2023). This is important because variations in study area (De et al., 2024), vegetation type (Fan et al., 2015), canopy density (Broge and Mortensen, 2002), soil brightness (Jiang et al., 2023) and seasonal conditions (Song et al., 2023) can influence reflectance values. This changes the parameters of an index to account for the differences, ensuring that the index accurately represents the vegetation characteristics of each specific image and region. For example, reducing noise from non-vegetative elements and enhancing vegetation specific signals from VIs help identify and map woody vegetation habitats across different spatial resolutions (Fakhri et al., 2022b). A study by (Zhang et al., 2006a) shows that coarse-resolution indices derived from sensors like the moderate resolution imaging spectroradiometer (MODIS) can capture broad scale patterns in vegetation health and productivity over large regions, while finer-resolution images from Landsat and Sentinel-2 sensors as well as hyperspectral imagery enable more detailed assessments at plot and species level (Clark, 2020). This allows VIs to be used for habitat quality assessments and species distribution monitoring in complex heterogeneous environments.

Limitations of VIs and remote sensing techniques in monitoring woody vegetation dynamics include parameter tuning (Ochtyra et al., 2020) and generalization across different ecosystems and heterogeneous landscapes. Moreover, most widely used and recommended VIs are not adaptable to semi-arid heterogeneous landscapes. This is primarily because these environments are characterized by sparse, patchy vegetation cover, high background soil reflectance and seasonal changes. These are all variables that can interfere with the accurate detection and interpretation of vegetation spectral signals using standard VIs. To bridge this gap, a novel, nature inspired optimization algorithm such as the particle swarm optimization (PSO) has been used to enhance the optimal coefficients for vegetation classification tuned for specific regions and contexts (Abualigah et al., 2024). PSO is a population-based optimization technique that is inspired by the social behaviour of bird flocking and fish schooling (Cao et al., 2011). Initially proposed by (Kennedy and Eberhart, 1996), PSO mimics social dynamics to iteratively search for optimal solutions in both mono- and multi-dimensional problem spaces. In our study, PSO is applied to optimize the coefficients of five key vegetation indices enabling more accurate estimation and mapping of woody vegetation characteristics such as health, vigour, degradation, and fragmentation that influence habitat quality. The difference in the optimised VIs to the traditional way of calculating them is the ability of the algorithm to treat each index's spectral coefficients as variables within a defined context (Fakhri et al., 2022b). Here, PSO is used to identify the most effective combinations of these coefficients for detecting vegetation changes across different spatial and spectral resolutions.

PSO has been constantly highlighted as an effective optimization tool for enhancing vegetation analysis through feature selection, model parameter tuning and spectral band optimization. For example, in (Zhang et al., 2006b), although PSO was not explicitly implemented, the focus on correcting NDVI scaling for vegetation cover estimation sets the foundation for PSO based methods to refine such models. (Banerjee and Raval, 2021) applied PSO with a minimum estimated abundance covariance (MEAC) criterion to efficiently select optimal hyperspectral bands for UAV-based vegetation mapping, particularly valuable in environments with limited field data. Similarly, (Islam et al., 2024) employed PSO to optimize deep learning extracted features before classification with a support vector machine (SVM) algorithm, showing that PSO can boost model performance and reduce computational processing. Collectively, the authors agreed on PSO's benefits in reducing dimensionality, improving model generalization and enhancing classification accuracy. However, while (Banerjee and Raval, 2021) focused on pre-acquisition optimization for hyperspectral sensors, (Islam et al., 2024) used PSO post feature extraction for classification accuracy for deep learning, highlighting PSO's adaptability across workflows.

In this study the PSO technique is suggested for the reparameterization of optimal coefficients across five key VIs: NDVI, SAVI, normalized difference red edge (NDRE), modified soil-adjusted vegetation index (MSAVI) and green normalized difference vegetation index (GNDVI). These VIs are suitable for detecting woody vegetation health, vigor, fragmentation and degradation for large scale habitat quality mapping and assessment. The significance of this method lies in its novelty in semi-arid heterogeneous areas during a dry season, a challenge for habitat quality assessments at both spatial and spectral scales. The advantages of a dry season assessment are the improved spectral separability of woody vegetation, improved visibility of degraded areas compared to when the vegetation has full canopies, as well as higher quality imagery due to the clearer atmospheric conditions.

Thus, this study aims to use the PSO algorithm to help optimize VIs to improve their ability to discriminate between vegetated and non-vegetated areas for a random forest classification for assessing habitat quality across different remote sensing platforms in a semi-arid area. This work builds upon (Fakhri et al., 2022b) who used a multi-objective PSO approach to enhance plant greenness and classify vegetation using multiple VIs and a combination of high- and low-resolution imagery in a mountainous area. Our study is different in that we optimize five vegetation indices that are specifically suited to the characteristics of semi-arid regions. The optimization is applied across four different sites within different ecological zones using data from Sentinel-2, Landsat 9 and unmanned aerial vehicles (UAVs), another novel approach.

2. Vegetation Indices

VIs are derived from spectral reflectance values and are used to enhance the detection of vegetation features from remote sensing sources (value = -1 to 1). This study chose to optimise NDVI, that uses the red (RED) and near infrared (NIR) bands to assess vegetation density and vigour (Madonsela et al., 2018). This is beneficial in habitat quality assessments as the index helps highlight areas of high and low vegetation in a specific context. The index is expressed as:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

where NIR represents near-infrared and RED represents the red portion of the electromagnetic spectrum (Zhang et al., 2006b).

The NDRE index was chosen for its sensitivity to chlorophyll content, capturing early signs of woody vegetation stress and degradation (Boiarskii, 2019). This index works well in semi-arid areas as it is suited for detecting subtle stress and assessing phenological changes in dry seasons. The index is calculated as follows:

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge} \tag{2}$$

where NIR is the near-infrared and RedEdge represents the portion of the electromagnetic spectrum detected by the red-edge band.

The SAVI was selected specifically for its tolerance for high soil reflectance, making it adaptable to different levels of vegetation cover (Qi et al., 1994). The index is calculated as follows:

$$SAVI = \frac{(NIR - Red)(1 + L)}{NIR + Red + L} \tag{3}$$

where NIR is the near-infrared, RED is the red band, and L is a soil correction factor.

The MSAVI index was selected because it minimises soil influence in sparse environments that typically characterize semi-arid regions (Qi et al., 1994). It is also better suited for early season monitoring in dry areas. The index is calculated as follows:

$$MSAVI = \frac{1}{2}\Big(2 \cdot NIR + 1 - \sqrt{(2 \cdot NIR + 1)^2 - 8(NIR - RED)}\Big) \tag{4}$$

Finally, the GNDVI was chosen because its red band is substituted with the green band, making it more sensitive to chlorophyll concentration, better detecting water stress and photosynthetic activity than NDVI (Wang et al., 2007). This is good for assessing habitat quality in semi-arid areas where small changes in chlorophyll content are critical for early stress detection in woody vegetation. The index is calculated as follows:

$$GNDVI = \frac{NIR - Green}{NIR + Green} \tag{5}$$

where NIR is the near-infrared and Green represents the green band of the electromagnetic spectrum.

3. Particle Swarm Optimization

In PSO, each potential solution also called a particle represents a candidate set of index coefficients, for example, the weights applied to NIR, RED, Green and RedEdge bands (Wang et al., 2022). Each particle moves through the search space with an associated position and velocity that is guided by both its own best-known solution (p_{best}) and the best-known solution found by the swarm (g_{best}). These positions are updated through multiple iterations towards an optimal solution as described in the following equation according to (Fakhri et al., 2022b):

$$v_{id}^{(t+1)} = w \cdot v_{id}^{(t)} + c_1 \cdot r_1 \cdot (p_{id} - x_{id}) + c_2 \cdot r_2 \cdot (g_d - x_{id}) \tag{6}$$

where $v_{id}^{(t+1)}$ is the velocity of particle i in dimension d at iteration t, x_{id} is the current position and coefficient value, p_{id} is the personal best position, g_d is the global best across all particles, w is the inertia weight, c_1 and c_2 are the cognitive and social acceleration coefficients, and r_1 and r_2 are random numbers in $[0,\ 1]$.

This PSO formula ensures a balance between searching new regions in the image and refining known particles, which is important when optimizing heterogeneous semi-arid habitats with complex spectral signatures (Banerjee and Raval, 2021). For this study, based on (Fakhri et al., 2019), Table 1 summarizes the PSO parameterization settings that ensure algorithm stability and convergence of the swarm.

By optimizing the VIs for the selected study sites, PSO will improve the image adaptability to soil brightness, low vegetation density and spectral noise such as shadows that present as challenges in semi-arid savannah ecosystems during the dry season.

4. Materials and Methods

4.1 Study Area

The Kruger National Park (KNP) in South Africa is part of the Greater Limpopo Transfrontier Conservation Area (GLTFCA) that includes the Limpopo National Park (LNP) of Mozambique, and Gonarezhou National Park (GNP) in Zimbabwe (SANParks, 2021). This study focuses only on the southern part of the KNP (Figure 1) due to data availability. The KNP is one of the largest national parks in the world with an area of approximately 35,000 km^2. The park comprises a width of approximately 65 km and a length of 360 km (Grab and Knight, 2015). It is separated from Mozambique in the east through the Lebombo Mountains, and in the north and south by the Limpopo River and the Crocodile River respectively as natural boundaries. The KNP is situated in the north-eastern ‘Lowveld’ or low ground, which is part of the ‘lowveld geomorphic province’ of South Africa (Partridge et al., 2010). The altitude varies between 300 and 840 m (Grab and Knight, 2015).

The area has an average annual rainfall of 243–648 mm and the average minimum temperature of 25°C and a maximum of 38°C. The area has a semi-arid seasonal climate

Table 1. PSO parameterization settings.

Parameters	Value	Purpose
Inertia Weight	0.8	Maintains a balance between local and global search
Acceleration Coefficient	1.496	Based on the construction factor method
Construction factor	0.729	Ensures convergence of particle trajectories
Maximum velocity	*	Allows flexible movement based on each spectral coefficient range
Population	30	Suitable for moderate complexity and computational efficiency
Iterations	100–200	Convergence without excessive computation

*Defined per dimension based on spectral coefficient bounds. Adapted from (Fakhri et al., 2019).

with most of the rainfall between November and March, with May–September being the driest months (SANParks, 2021). Droughts are relatively common in the area with occasional severe occurrences, while parts of the park are prone to flooding (Malherbe et al., 2020). The park is home to many tourist attractions that include the “big 5”, making it a conservation hotspot (Hoogendoorn et al., 2019). The area has three dominating vegetation zones: mixed thorn and marula woodlands on granite, mopane dominated woodlands on granite, and mixed woodland and thorn thickets (SANParks, 2021).

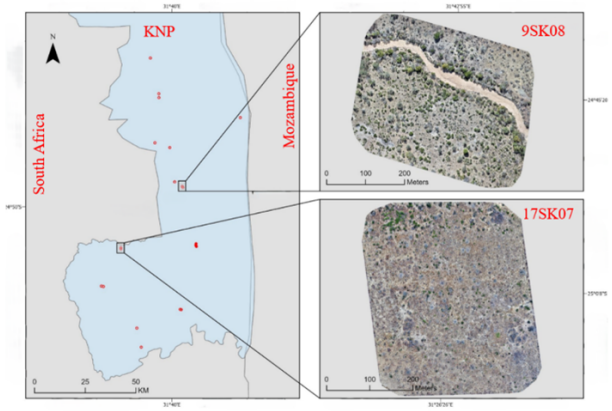

Figure 1. Map of the Southern part of the Kruger National Park and two of the plots flown with the UAV sensor. Source: Andrea Meier (2025).

4.2 Woody Vegetation Data (Ground Truthing)

Fieldwork was conducted over 11 days from the 20th of August until the 1st of September 2024. In total, approximately 800 woody vegetation species were mapped across 23 plots. For the purposes of this study, only four plots (most populated) and 131 species were used for the analysis. Additionally, these plots had the most variations for vegetated and non-vegetated areas, making them accessible for analysis. Furthermore, these plots could be identified in Sentinel-2 and Landsat 9 images with optimal atmospheric conditions. The selected plots had trees that were >5 metres tall and exhibited vegetational and geomorphological diversity. To capture the locations of these trees in the field, data was acquired using a predesigned survey employing the software Survey123 from the ESRI catalogue. This made it possible to record the GPS location of the trees, photos and additional allometric variables (diameter at breast height (DBH), stem height and tree height). Tree species were identified using *Trees of South Africa* (Van Wyk et al., 2012) and the expertise of our game guards.

4.3 Remotely Sensed Data

This study was conducted across multiple geographically distinct sites in the KNP, selected to represent a variety of vegetation types. Multispectral imagery was collected from three different remote sensing platforms, each offering different spatial resolutions.

High-resolution (2.4 cm) imagery was acquired using the DJI Mavic 3 Pro UAV equipped with a multispectral sensor. This platform provided fine spatial resolution imagery suitable for detailed, plot-level vegetation analysis. The UAV was flown at an average flight height of 60.44 metres and a speed of 5.39 m/s. Medium-resolution (10 m) imagery was obtained from the Sentinel-2 satellite. Coarser resolution imagery was sourced from the Landsat 9 satellite, offering broader spatial coverage with a 30 m resolution, useful for regional scale assessments.

Google Earth Engine (GEE) was used for extracting medium and coarse resolution remote sensing imagery. This is an Application Programming Interface (API) based data platform that leverages Google’s computational infrastructure to enable geospatial data processing to reduce computational time (Tamiminia et al., 2020). For both Sentinel-2 and Landsat 9, images with less than 2% cloud cover were selected and were atmospherically corrected for the same date range as the UAV flights (20th August 2024 to 1st September 2024). The chosen date range corresponds with the late dry season in the study area. This period was intentionally chosen to capture conditions where vegetation stress is most pronounced, enhancing the detectability of woody vegetation structures. A cloud mask was applied where applicable, and a single composite image was derived by calculating the mean image from the collected time frame.

Pre-processing steps were undertaken to ensure the consistency and comparability of data across platforms. For each dataset, radiometric calibration was applied to convert raw digital numbers (DNs) to surface reflectance values. Geometric corrections were performed to align imagery to a common coordinate reference system, ensuring spatial accuracy across datasets. For Sentinel-2 and Landsat 9 imagery, cloud masking and atmospheric correction were conducted using the Sen2Cor and LEDAPS algorithms, respectively. UAV imagery underwent additional pre-processing steps, including orthomosaic generation and geometric correction using DJI D-RTK 2 high precision GNSS mobile station ground control points (GCPs), to ensure high positional accuracy. Following pre-processing, the five above-mentioned vegetation indices were calculated across all datasets to evaluate vegetation condition and canopy vigour.

Table 2. Plots surveyed in the field including the number of trees per plot and their allometric variables.

Plot name	Tree Count	DBH Count	Tree height count
1PH01	30	29	30
8SK08	28	27	28
9SK08	37	37	37
17SK07	36	36	36

*where plot name is derived from the ecological zones found in the Kruger National Park.

4.4 PSO Procedure

To improve the sensitivity and performance of the VIs, PSO was used to estimate their optimal coefficient values. This approach targeted the optimization of three coefficients (a, b and c) (Table 1). The optimization objective was to maximize the mean VI value across each study site, thereby enhancing the index's ability to differentiate between vegetated areas, particularly under sparse canopy conditions typical of dryland areas (Lu et al., 2023). The optimization problem was framed as a minimization of the negative mean VI value. The algorithm was implemented in MATLAB R2022b using the Global Optimization Toolbox (Alam, 2016). The algorithm employed a swarm of 50 particles to explore the solution space defined by the coefficients a, b, and c.

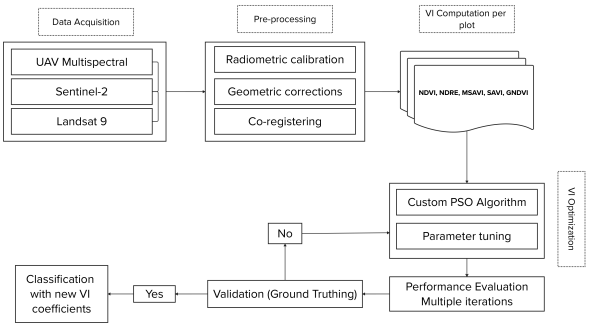

Figure 2. Particle Swarm Optimisation workflow.

Each particle's position was initialized within specified bounds $[1, 1, 1]$ to $[10, 10, 10]$ and updated iteratively over 100 iterations using standard PSO parameters: an inertia weight of 0.8 and cognitive and social acceleration coefficients set to 1.49618 (Fakhri et al., 2019). At each iteration, VIs were recalculated using the proposed parameters, and the solution yielding the highest mean VI values was selected for final use (Alam, 2016). This optimization process was integrated into a comprehensive remote sensing workflow for vegetation analysis, as illustrated in Figure 2.

4.5 Classification

Vegetated and non-vegetated areas post index optimisation were classified using the random forest classification algorithm in ArcGIS Pro. Random forest was selected because of its robustness (Sun et al., 2024), non-parametric nature and ability to handle high-dimensional input data without requiring assumptions about the distribution of the data (Svoboda et al., 2022). Additionally, random forest is effective in this context as it can manage non-linear relationships, is resistant to model overfitting, and provides internal measures of variable importance — important when working with multiple optimised indices such as NDVI, NDRE, MSAVI, SAVI and GNDVI for semi-arid regions (Xue and Su, 2017). For accuracy assessments, an error matrix was used to extract overall, user and producer accuracies as well as the kappa coefficient.

4.6 Statistical Analysis

To validate the model results and generalisation across different sensors for woody vegetation mapping, only the best performing indices were chosen for agreement analysis. The Bland-Altman plot was used to visualise agreement and detect systematic bias between the best performing indices. This approach plots the mean differences against the average value of indices for consistency and deviations (Giavarina, 2015) in assessing vegetation characteristics. Additionally, as an outlook, principal component analysis (PCA) was employed to validate the spatial and spectral consistency of the optimised indices for individual woody species identification and mapping for a habitat quality assessment.

5. Results

5.1 VIs Pre- and Post-Optimization

VI results from pre- and post-optimization (Appendix 1, Tables 4–9) show an increase in maximum index values per index per plot. For example, NDVI values in UAV multispectral imagery increased from a maximum of 0.7 to 0.8 in plot P-8SK08 and from 0.76 to 0.87 at P-17SK07. Similarly, MSAVI values increased from 0.29–0.47 pre-optimization to 0.38–0.62 post-optimization. GNDVI and SAVI followed similar trends across all platforms, with Sentinel-2 showing the most substantial post-optimization increase in GNDVI, reaching a maximum of 1.0 at plot P-1PH01. Post-optimization, UAV NDRE and MSAVI indices showed the most consistent modest increases across all plots, especially at P-9SK08 and P-1PH01, showing improved chlorophyll detection and soil discrimination in these areas. Sentinel-2 MSAVI and GNDVI (0.87–1) also showed strong increases across all plots with P-9SK08 reflecting the most improved (0.59–0.78) soil-vegetation discrimination and vegetation stress. For this reason, these indices and platforms were chosen for woody vegetation classification.

For sensor comparison (Figure 3), post-optimization boxplots show that UAV indices respond strongly to optimization for NDVI (0.79–0.87) and GNDVI (0.67–0.771). SAVI and MSAVI values showed improved upper values with MSAVI reaching a new high of 0.49. UAV NDRE, while values were low, showed the highest sensitivity to less dense woody vegetation. Sentinel-2 indices showed a more steady and consistent improvement with NDVI peaking at 0.23, GNDVI (1) and MSAVI (0.7) at plot P-9SK08, the third most vegetated plot.

5.2 Post-Optimization Classification

It should be noted that Landsat 9 was excluded from further analysis despite showing post-optimization improvements, due to inconsistencies in the VI values and its coarse spatial resolution at the plot scale. Therefore, all subsequent results are derived from UAV and Sentinel-2 datasets. Four distinct land use/land cover classes per plot were distinguished in the field: three vegetation classes (woodland consisting of large woody trees, shrubland, and grassland) and one non-vegetated areas class. Post-optimized UAV NDRE, UAV MSAVI, Sentinel-2

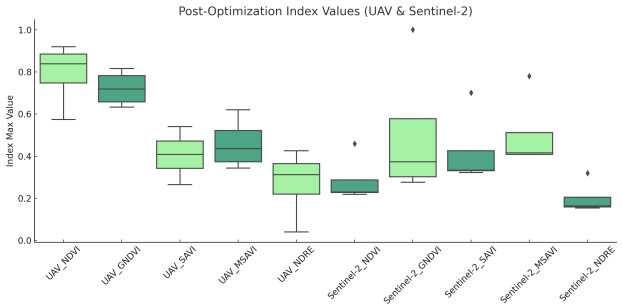

Figure 3. Post-optimization index values for UAV and Sentinel-2.

MSAVI and Sentinel-2 GNDVI classification results (Figure 4) show that the UAV derived indices indicate a high level of grasslands in plots P-8SK08 and P-1PH01, and the VIs effectively isolated the non-vegetated from grasses and shrubs in plot P-9SK08. The Sentinel-2 classification was consistent in isolating woodland and shrubs in plots P-8SK08 and P-17SK07. Shrubs were consistently classified on both platforms.

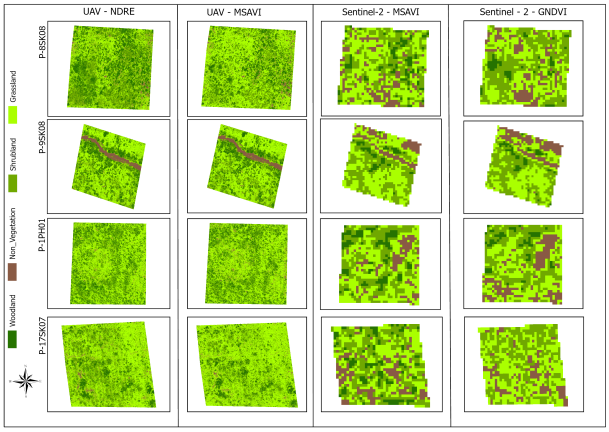

Figure 4. Post-optimization classification for UAV and Sentinel-2 MSAVI and GNDVI.

For classification accuracy assessment, four plots were chosen per index: plot P-8SK08 for UAV MSAVI, P-9SK08 for UAV NDRE, P-1PH01 for Sentinel-2 MSAVI and P-17SK07 for Sentinel-2 GNDVI. Overall classification accuracy for UAV MSAVI was 80% with a kappa coefficient of 0.71. UAV NDRE performed at 85% accuracy with a 0.79 kappa coefficient, with woodland (100%) and grassland (83%) correctly classified. Sentinel-2 MSAVI performed at 65% overall accuracy and 0.48 kappa coefficient, with woodland (71%) and shrubland (83%) correctly classified. Sentinel-2 GNDVI performed at 70% accuracy with a kappa coefficient of 0.59, with woodland and grassland correctly classified.

5.3 Model Validation

For comparability and consistency, the Bland-Altman agreement analysis compared UAV MSAVI and Sentinel-2 MSAVI post-optimization classification across matching sample points. Results show a consistent bias (mean difference = -0.157) between UAV and Sentinel-2 MSAVI values, with UAV-derived values being lower (Figure 5). Standard deviations (± 1.96) show a moderate and acceptable level of agreement. While the clustering of the ground-truthing points is on average low (0.2–0.3) for both sensors, UAV values are systematically lower, indicating that UAV MSAVI classification provided more accurate information of the different classes while Sentinel-2 overestimated vegetation structure. This is attributed to differences in both spatial and spectral resolution. While the model validation agreement is low, the agreement of the model shows that the methods can be applied across other indices and sources used in this study.

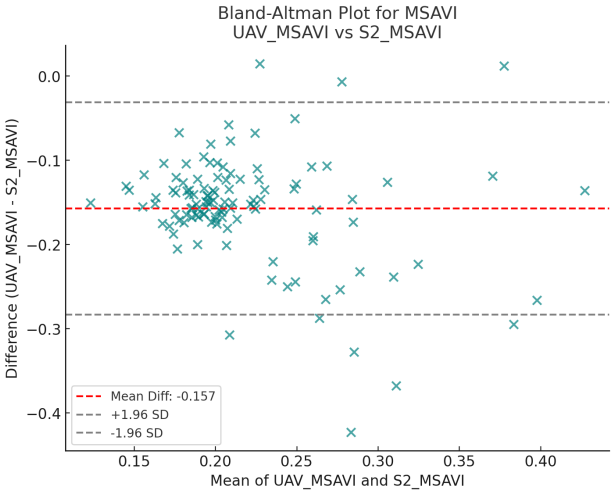

Figure 5. Bland-Altman Plot for MSAVI: a comparison of the mean UAV MSAVI and Sentinel-2 MSAVI and their difference.

6. Discussion

6.1 VIs and Particle Swarm Optimization

This study used the PSO algorithm to enhance VIs to help discriminate between vegetated and non-vegetated areas for a random forest classification to assess habitat quality. Regarding optimization behaviour and performance, minimum VI values across indices per sensor were preserved post-optimization to maintain the lower bounds of reflectance values and to avoid underestimating ecologically valid signals for low vegetated areas (Gobron et al., 2000). Additionally, this was intentionally done so that the minimum values represent the dryness of the area more realistically. Maximum values consistently increased across all sensors and plots, showing improved sensitivity to canopy vigor, stress, health and presence per vegetation index.

Across sensors, post-optimization, UAV-derived indices such as the NDVI, NDRE and SAVI overestimated vegetative features, particularly in areas around large woody vegetation with wide canopies, although the classification was 100% accurate. This problem is attributed to sensor resolution, which introduces shadows beneath large trees that are then misclassified into high values representing woody vegetation. In this regard, (Pons and Padró, 2019) suggested using empirical line correction through radiometric references to correct the cast shadows of the canopies of woody vegetation. Furthermore, to mitigate this problem, a separate class representing shadows may be created, although this introduces the risk of data loss or oversimplification as these areas fall within other classes such as grasses and shrubs. Our study did not apply this correction due to the primary focus being the enhancement of the VIs through PSO across spatial extents; incorporating radiometric

corrections would have been beyond the scope of this methodological framework. However, future studies can incorporate this step for shadow-specific classes to improve classification accuracy.

Regarding VIs, NDVI demonstrated a significant increase post-optimization across all sensors. The index overestimated vegetation greenness, which falsely suggested higher vegetation density than was present in the plots. Its high variance led to its exclusion from further analysis and classification. (Gao et al., 2023) also observed that the NDVI index enhances the contrast of the reflectivity of the NIR band in remote sensing images, leading to its inability to characterize between new vegetation growth and soils, especially in semi-arid regions. MSAVI and GNDVI, optimized for sparse vegetation environments, showed strong improvements in UAV and Sentinel-2 data, making them effective for detecting fragmented vegetation in semi-arid regions. Among all platforms, Sentinel-2 consistently outperformed others in post-optimization. This can be attributed to the sensor's higher spatial and spectral resolution allowing it to better capture the heterogeneity of woody vegetation (Qi et al., 1994).

As mentioned earlier, Landsat 9, while displaying consistent improvements post-optimization, was excluded from further analysis due to inconsistency found in the vegetation index values and its coarser spatial resolution at plot level. Although resampling Landsat 9 imagery could help mitigate this challenge, the process risks data loss and inflated VIs, leading to unreliable habitat quality assessments and low accuracy classification at plot level.

Post-optimization classification significantly improved ($+15\%$) the ability to differentiate vegetated from non-vegetated areas, especially in plots that are difficult to detect due to the spectral signatures and the heterogeneity of the land cover. Although the enhancement of the VIs helped to isolate woody vegetation more effectively, the classification of grasses was overestimated due to their spectral similarity to bare soils. Additionally, across all remote sensing platforms, the images were acquired during a dry season, making it even more challenging for differentiation post optimization. These observations corroborate with (Domínguez-Beisiegel et al., 2016) who, although in an arid study site, found that soil background had a high spectral reflectance making it challenging to map vegetation dynamics.

The classification results per plot per sensor show consistent patterns of habitat fragmentation with most areas classified as grasslands and less woody vegetation cover (Appendix 1, Figure 1). This is typical of semi-arid protected ecosystems (Bindajam et al., 2021) where climatic conditions and wildlife pressures lead to ecosystem degradation (Pueyo et al., 2006) and transitions in vegetation structure. However, the abundance of shrubs and grasses shows a potential decline in woody vegetation. Field validation and observed patterns in the classified imagery also support the understanding that wildlife pressures, fire regimes and drought contribute to habitat degradation and vegetation fragmentation. This emphasises the value of optimization-driven classification frameworks in improving the detection and interpretation of semi-arid vegetation dynamics.

6.2 Implications for Habitat Quality and Conservation

The results of this study show the reliability and potential of remote sensing techniques in quantifying habitat quality for biodiversity conservation in semi-arid ecosystems. Optimized VIs enhanced the ability and effectiveness to assess and monitor vegetation dynamics in this heterogeneous area where field-based methods are both expensive and challenging. The increased accuracy in distinguishing vegetation types using different sensors will equip management authorities, in this case, the South African National Parks (SANParks) to better understand habitat quality, its decline and degradation. Furthermore, this study also helps inform policies on vegetation restoration, particularly where bush encroachment, overgrazing and land use conversion are concerned. This study can be used, with the help of scalable sensors like Sentinel-2 combined with UAV imagery, for local and regional scale interventions, conservation planning and ecological monitoring.

6.3 Advantages and Limitations

Although the results of this study are conclusive and provide practical observations for both the field and image, there are several limitations observed. While Sentinel-2 imagery was the most suitable for this study area due to its spatial, spectral and temporal resolution, and UAV imagery had the higher classification accuracy, both images could be improved spatially. Furthermore, the computational power required to run the PSO algorithm in MATLAB R2022b was high, taking each image an average of 45 minutes per run for a single iteration. This is a limitation for large-scale area analysis over multiple VIs, especially for a transboundary protected area. The observed advantages of the PSO algorithm were its suitability in assessing multiple VI coefficients for optimal classification and accuracy, as well as its ability to incorporate multiple sensors for analysis. The disadvantages noted were the grass and soil confusion where optimization did not fully fix spectral similarities, as well as shadow effects in the analysis. Finally, we cannot be fully confident that the PSO results are transferable across all resolutions and sensors.

7. Conclusion

In conclusion, this study demonstrates that PSO-based optimization of vegetation indices can enhance the precision and scalability of habitat quality assessments across multiple spatial and spectral resolutions. The optimized indices (NDRE, GNDVI, MSAVI) proved especially valuable in heterogeneous, semi-arid ecosystems, where vegetation is dynamic, sparse, and vulnerable to both climatic variability and human activity. Among the evaluated sensors, Sentinel-2 emerged as the most effective platform for scalable vegetation mapping in these environments, offering a promising pathway for future ecological monitoring and land management strategies.

Acknowledgments

We would like to extend our deepest gratitude to the South African National Parks (SANParks) for working with us and keeping us safe in the field. To Mrs Samantha Mabuza and Mrs Chane Simms for facilitating our fieldwork. Miss Isabel Karst and Miss Sandra Reuther for their valuable assistance during the data collection phase of the fieldwork.

Supplementary Material

The appendix is available online https://doi.org/10.5281/zenodo.20344817.

References

Abualigah, L. *et al.*, “Particle swarm optimization algorithm: review and applications,” in *Metaheuristic Optimization Algorithms*, Elsevier, pp. 1–14, 2024.

Acevedo, S. and Jones, S., “An Evaluation of the utility of two classifiers for mapping woody vegetation using remote sensing,” 2012.

Alam, M.N., “Codes in MATLAB for Particle Swarm Optimization,” 2016.

Arvor, D., Durieux, L., Andrés, S. and Laporte, M.A., “Advances in Geographic Object-Based Image Analysis with ontologies: A review of main contributions and limitations from a remote sensing perspective,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 82, pp. 125–137, 2013.

Banerjee, B.P. and Raval, S., “A particle swarm optimization based approach to pre-tune programmable hyperspectral sensors,” *Remote Sens. (Basel).*, vol. 13, no. 16, 2021.

Batta, M., “Machine Learning Algorithms - A Review,” *International Journal of Science and Research (IJSR)*, vol. 18, no. 8, pp. 381–386, 2018.

Bindajam, A.A. *et al.*, “Assessing landscape fragmentation effects on ecosystem services in a semi-arid mountainous environment,” *Appl. Ecol. Environ. Res.*, vol. 19, no. 3, pp. 2519–2539, 2021.

Boiarskii, B., “Comparison of NDVI and NDRE Indices to Detect Differences in Vegetation and Chlorophyll Content,” *Journal of Mechanics of Continua and Mathematical Sciences*, vol. spl1, no. 4, 2019.

Broge, N.H. and Mortensen, J.V., “Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data,” 2002.

Cao, J., Zhang, N., Li, M.B. and Zhang, J.W., “Based on the PSO optimization wood fiber mechanical parameters control model research,” in *Procedia Engineering*, pp. 1093–1097, 2011.

Chafota, J., “Effects of Changes in Elephant Densities on the Environment and Other Species – How Much Do We Know,” *Cooperative Regional Wildlife Management in Southern Africa*, pp. 1–21, 1998.

Chapungu, L., Nhamo, L., Gatti, R.C. and Chitakira, M., “Quantifying changes in plant species diversity in a savanna ecosystem through observed and remotely sensed data,” *Sustainability (Switzerland)*, vol. 12, no. 6, 2020.

Clark, M.L., “Comparison of multi-seasonal Landsat 8, Sentinel-2 and hyperspectral images for mapping forest alliances in Northern California,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 26–40, 2020.

Coleine, C. *et al.*, “Dryland microbiomes reveal community adaptations to desertification and climate change,” *Oxford University Press*, 2024.

Cui, G. *et al.*, “Study of Spatiotemporal Changes and Driving Factors of Habitat Quality: A Case Study of the Agro-Pastoral Ecotone in Northern Shaanxi, China,” *Sustainability (Switzerland)*, vol. 14, no. 9, 2022.

Darra, N. *et al.*, “Spectral bands vs. vegetation indices: An AutoML approach for processing tomato yield predictions based on Sentinel-2 imagery,” *Smart Agricultural Technology*, vol. 10, 2025.

De, A., Sahani, N., Datta, A. and Maitra, A., “Spatiotemporal analysis of different vegetation indices and relation to meteorological parameters over a tropical urban location,” *Atmosfera*, vol. 38, pp. 911–934, 2024.

Delaney, B., Tansey, K. and Whelan, M., “Satellite Remote Sensing Techniques and Limitations for Identifying Bare Soil,” *MDPI*, 2025.

Domínguez-Beisiegel, M., Castañeda, C., Mougenot, B. and Herrero, J., “Analysis and mapping of the spectral characteristics of fractional green cover in saline wetlands (NE Spain),” *Remote Sens. (Basel).*, vol. 8, no. 7, 2016.

Druce, D.J., Shannon, G., Page, B.R., Grant, R. and Slotow, R., “Ecological thresholds in the Savanna landscape: Developing a protocol for monitoring the change in composition and utilisation of large trees,” *PLoS One*, vol. 3, no. 12, pp. 1–8, 2008.

Dube, K. and Nhamo, G., “Evidence and impact of climate change on South African national parks. Potential implications for tourism in the Kruger National Park,” *Environ. Dev.*, vol. 33, p. 100485, 2020.

Fakhri, S.A., Sayadi, S., Latifi, H. and Khare, S., “An optimized Enhanced Vegetation Index for Sparse Tree Cover Mapping across a Mountainous Region,” in *2019 IEEE International Workshop on Metrology for Agriculture and Forestry*, pp. 146–151, 2019.

Fakhri, S.A., Sayadi, S., Naghavi, H. and Latifi, H., “A novel vegetation index-based workflow for semi-arid, sparse woody cover mapping,” *J. Arid Environ.*, vol. 201, p. 104748, 2022.

Fakhri, S.A., Sayadi, S., Naghavi, H. and Latifi, H., “A novel vegetation index-based workflow for semi-arid, sparse woody cover mapping,” *J. Arid Environ.*, vol. 201, p. 104748, 2022.

Fan, H., Fu, X., Zhang, Z. and Wu, Q., “Phenology-based vegetation index differencing for mapping of rubber plantations using landsat OLI data,” *Remote Sens. (Basel).*, vol. 7, no. 5, pp. 6041–6058, 2015.

Flood, N., Watson, F. and Collett, L., “Using a U-net convolutional neural network to map woody vegetation extent from high resolution satellite imagery across Queensland, Australia,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 82, 2019.

Fundisi, E., Musakwa, W., Ahmed, F.B. and Tesfamichael, S.G., “Estimation of woody plant species diversity during a dry season in a savanna environment using the spectral and textural information derived from WorldView-2 imagery,” *PLoS One*, vol. 15, no. 6, pp. 1–19, 2020.

Gao, S. *et al.*, “Evaluating the saturation effect of vegetation indices in forests using 3D radiative transfer simulations and satellite observations,” *Remote Sens. Environ.*, vol. 295, 2023.

Gebre, A.B., Birhane, E., Gebresamuel, G., Hadgu, K.M. and Norgrove, L., “Woody species diversity and carbon stock under different land use types at Gergera watershed in eastern Tigray, Ethiopia,” *Agroforestry Systems*, vol. 93, no. 3, pp. 1191–1203, 2019.

Giavarina, D., “Understanding Bland Altman analysis,” *Biochem. Med. (Zagreb).*, vol. 25, no. 2, pp. 141–151, 2015.

Gobron, N., Pinty, B., Verstraete, M.M. and Widlowski, J.-L., “Advanced Vegetation Indices Optimized for Up-Coming Sensors: Design, Performance, and Applications,” 2000.

Grab, S. and Knight, J., “World Geomorphological Landscapes: Landscapes and Landforms of South Africa,” 2015.

Hoogendoorn, G., Kelso, C. and Sinthumule, I., “Tourism in the Great Limpopo Transfrontier Park: A review,” *African Journal of Hospitality, Tourism and Leisure*, vol. 8, no. 5, 2019.

Islam, M.T. *et al.*, “Medicinal Plant Classification Using Particle Swarm Optimized Cascaded Network,” *IEEE Access*, vol. 12, pp. 42465–42478, 2024.

Jiang, H. *et al.*, “Influence of different soil reflectance schemes on the retrieval of vegetation LAI and FVC from PROSAIL in agriculture region,” *Comput. Electron. Agric.*, vol. 212, 2023.

Kennedy, J. and Eberhart, R., “Particle Swarm Optimization,” in *Proceedings of IEEE International Conference on Neural Networks*, 1996.

Levin, N., Shmida, A., Levanoni, O., Tamari, H. and Kark, S., “Predicting mountain plant richness and rarity from space using satellite-derived vegetation indices,” *Divers. Distrib.*, vol. 13, no. 6, pp. 692–703, 2007.

Lu, J. *et al.*, “Lithology classification in semi-arid area combining multi-source remote sensing images using support vector machine optimized by improved particle swarm algorithm,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 119, 2023.

Madonsela, S. *et al.*, “Multi-phenology WorldView-2 imagery improves remote sensing of savannah tree species,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 58, pp. 65–73, 2017.

Madonsela, S., Cho, M.A., Ramoelo, A., Mutanga, O. and Naidoo, L., “Estimating tree species diversity in the savannah using NDVI and woody canopy cover,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 66, pp. 106–115, 2018.

Malherbe, J., Smit, I.P.J., Wessels, K.J. and Beukes, P.J., “Recent droughts in the Kruger National Park as reflected in the extreme climate index,” *Afr. J. Range Forage Sci.*, vol. 37, no. 1, pp. 1–17, 2020.

Mogashoa, R., Dlamini, P. and Gxasheka, M., “Grass species richness decreases along a woody plant encroachment gradient in a semi-arid savanna grassland, South Africa,” *Landsc. Ecol.*, vol. 36, no. 2, pp. 617–636, 2021.

Muposhi, V.K., Ndlovu, M., Gandiwa, E., Muvengwi, J. and Muboko, N., “Vegetation dynamics prior to wildlife reintroductions in southern umfurudzi park, Zimbabwe,” *J. Anim. Plant Sci.*, vol. 24, no. 6, pp. 1680–1690, 2014.

Ochtyra, A., Marcinkowska-Ochtyra, A. and Raczko, E., “Threshold- and trend-based vegetation change monitoring algorithm based on the inter-annual multi-temporal normalized difference moisture index series,” *Remote Sens. Environ.*, vol. 249, 2020.

Partridge, T.C., Dollar, E.S.J., Dollar, L.H. and Moolman, J., “The geomorphic provinces of South Africa, Lesotho and Swaziland,” *Transactions of the Royal Society of South Africa*, vol. 65, no. 1, pp. 1–47, 2010.

Peña-Lara, V.A. *et al.*, “Modelling Species Richness and Functional Diversity in Tropical Dry Forests Using Multispectral Remotely Sensed and Topographic Data,” *Remote Sens. (Basel).*, vol. 14, no. 23, pp. 1–15, 2022.

Pfeifer, M. *et al.*, “Deadwood biomass: An underestimated carbon stock in degraded tropical forests?,” *Environmental Research Letters*, vol. 10, no. 4, 2015.

Pons, X. and Padró, J.-C., *2019 IEEE International Geoscience & Remote Sensing Symposium*. IEEE, 2019.

Pueyo, Y., Alados, C.L. and Barrantes, O., “Determinants of land degradation and fragmentation in semiarid vegetation at landscape scale,” *Biodivers. Conserv.*, vol. 15, no. 3, pp. 939–956, 2006.

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H. and Sorooshian, S., “A Modified Soil Adjusted Vegetation Index,” 1994.

SANParks, “Great Limpopo Transfrontier Park,” 2021.

Senf, C., “Seeing the System from Above: The Use and Potential of Remote Sensing for Studying Ecosystem Dynamics,” *Ecosystems*, vol. 25, no. 8, pp. 1719–1737, 2022.

Shumi, G. *et al.*, “Woody plant species diversity as a predictor of ecosystem services in a social–ecological system of southwestern Ethiopia,” *Landsc. Ecol.*, vol. 36, no. 2, pp. 373–391, 2021.

Song, J. *et al.*, “Evaluating the Effect of Vegetation Index Based on Multiple Tree-Ring Parameters in the Central Tianshan Mountains,” *Forests*, vol. 14, no. 12, 2023.

Storch, F., Dormann, C.F. and Bauhus, J., “Quantifying forest structural diversity based on large-scale inventory data,” *For. Ecosyst.*, vol. 5, no. 1, pp. 1–14, 2018.

Sun, Z. *et al.*, “An improved random forest based on the classification accuracy and correlation measurement of decision trees,” *Expert Syst. Appl.*, vol. 237, 2024.

Svoboda, J., Štych, P., Laštovička, J., Paluba, D. and Kobliuk, N., “Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data,” *Remote Sens. (Basel).*, vol. 14, no. 5, 2022.

Taddeo, S., Dronova, I. and Depsky, N., “Spectral vegetation indices of wetland greenness: Responses to vegetation structure, composition, and spatial distribution,” *Remote Sens. Environ.*, vol. 234, p. 111467, 2019.

Tamiminia, H. *et al.*, “Google Earth Engine for geo-big data applications: A meta-analysis and systematic review,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 164, pp. 152–170, 2020.

Van Wyk, B., Coates Palgrave, K. and Van Wyk, P., “Field guide to trees of Southern Africa,” 2012.

Wang, F., Huang, J., Tang, Y. and Wang, X., “New Vegetation Index and Its Application in Estimating Leaf Area Index of Rice,” *Rice Sci.*, vol. 14, no. 3, pp. 195–203, 2007.

Wang, H., Yan, C., Yuan, J. and Lu, Q., “Hyperspectral Band Selections for Enhancing the Discrimination of Difficult Targets Using Local Band Index and Particle Swarm Optimization,” *Applied Sciences (Switzerland)*, vol. 12, no. 8, 2022.

Wolmer, W., “Transboundary conservation: The politics of ecological integrity in the Great Limpopo transfrontier park,” *J. South. Afr. Stud.*, vol. 29, no. 1, pp. 261–278, 2003.

Xue, J. and Su, B., “Significant remote sensing vegetation indices: A review of developments and applications,” *J. Sens.*, vol. 2017, 2017.

Xue, J. and Su, B., “Significant remote sensing vegetation indices: A review of developments and applications,” *J. Sens.*, vol. 2017, 2017.

Yue, Y., Zhao, W. and Liu, R., “Relationships between vegetation indices and surface reflectance: Implications for detecting and monitoring sandification in arid regions,” *Ecol. Indic.*, vol. 176, 2025.

Zagajewski, B. *et al.*, “Comparison of random forest, support vector machines, and neural networks for post-disaster forest species mapping,” *Remote Sens. (Basel).*, vol. 13, no. 13, 2021.

Zhang, X., Friedl, M.A. and Schaaf, C.B., “Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS),” *J. Geophys. Res. Biogeosci.*, vol. 111, no. 4, 2006.

Zhang, X., Yan, G., Li, Q., Li, Z.L., Wan, H. and Guo, Z., “Evaluating the fraction of vegetation cover based on NDVI spatial scale correction model,” *Int. J. Remote Sens.*, vol. 27, no. 24, pp. 5359–5372, 2006.

Zhang, Y., Wang, S. and Ji, G., “A Comprehensive Survey on Particle Swarm Optimization Algorithm and Its Applications,” *Hindawi Limited*, 2015.

Zhu, H. *et al.*, “Predicting plant diversity in beach wetland downstream of Xiaolangdi reservoir with UAV and satellite multispectral images,” *Science of the Total Environment*, vol. 819, p. 153059, 2022.