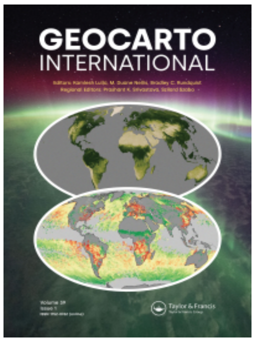

Geocarto International

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgei20

Utilising RGB drone imagery and vegetation indices for accurate above-ground biomass estimation: a case study of the cradle nature reserve, Gauteng Province, South Africa

Charles Matyukira & Paidamwoyo Mhangara

GEOCARTO INTERNATIONAL
2024, VOL. 39, NO. 1, 2390512
https://doi.org/10.1080/10106049.2024.2390512

Taylor & Francis
Taylor & Francis Group

Utilising RGB drone imagery and vegetation indices for accurate above-ground biomass estimation: a case study of the cradle nature reserve, Gauteng Province, South Africa

Charles Matyukira and Paidamwoyo Mhangara

School of Geography, Archaeological & Environmental Studies, Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

ABSTRACT

Accurate estimation of above-ground biomass (AGB) is a crucial aspect of ecosystem management, particularly in game reserves like the Cradle Nature Reserve. This study utilises 2018 AGB at a 50 m resolution to estimate AGB by integrating RGB drone imagery vegetation indices, normalised digital surface model (nDSM), and field-measured chlorophyll concentrations. Principal Component Analysis (PCA) and stepwise regression were employed to identify significant predictors of AGB. The final model for non-riparian forest area explained 66% ($R^2 = 0.66$), with nDSM and red-blue vegetation index (RBVI) as key predictors. In contrast, the riparian model explained 42% ($R^2 = 0.42$), with chlorophyll content and blue-green vegetation index (BGVI) as significant predictors. Our findings highlight the significant influence of nDSM and chlorophyll content on AGB. Vegetation indices such as the green-blue vegetation index (GBVI), green-red vegetation index (GRBI), and RBVI showed diverse correlations with chlorophyll levels, capturing plant greenness and vitality and directly influencing biomass. Comparative analyses of riparian and non-riparian vegetation models revealed that chlorophyll content and BGVI were significant predictors in riparian zones, reflecting interactions between water availability, nutrient flow, and vegetation health. In non-riparian areas dominated by tall trees, nDSM and RBVI were key predictors, underscoring the importance of plant height and stress indicators. This study underscores the effectiveness of drones in remote sensing applications for forest management. These findings contribute to the growing body of knowledge on drone-based biomass estimation, offering valuable insights for managing and conserving ecosystems such as the Cradle Nature Reserve, thereby providing practical implications for ecosystem management and conservation.

ARTICLE HISTORY
Received 7 June 2024
Accepted 5 August 2024

KEYWORDS
Cradle nature reserve; above-ground biomass; chlorophyll concentrations; vegetation indices; normalised digital surface model

CONTACT Paidamwoyo Mhangara paida.mhangara@wits.ac.za

1. Introduction

Precisely assessing the amount of biomass present above the ground (AGB) is crucial for efficient management of ecosystems (Flores et al. 2019; Song et al. 2024), especially in protected areas like game reserves and nature reserves such as the Cradle Nature Reserve. AGB provides critical data on plant health, carbon storage, and habitat quality, all of which are necessary for biodiversity conservation and sustainable management techniques (Numata et al. 2010; Gao et al. 2018; Bazzo et al. 2023; Song et al. 2024). Conventional techniques for estimating biomass, such as conducting field surveys and observations, are often characterised by their need for extensive manual effort, significant time investment, and potential negative impact on the environment (Bazzo et al. 2023; Nyamukuru et al. 2023; Song et al. 2024). As a result, remote sensing technologies have become a cost-effective and dependable option for monitoring plant biomass across large areas (Wang et al. 2023; Rapiya et al. 2023; Khumalo et al. 2024). Recent advancements in drone technology have transformed the discipline of remote sensing, providing high-resolution pictures and intricate 3D models of plant structures (Bazzo et al. 2023; Wang et al. 2023; Song et al. 2024). Researchers have demonstrated the usefulness of drone imaging using the RGB (Red, Green, and Blue) colour model in evaluating plant features (Yue et al. 2017; Agapiou 2020; Wang et al. 2023; Song et al. 2024). Methods such as photogrammetric imaging and structure from motion (SfM) can generate digital surface models (DSMs) and digital terrain models (DTMs) (Turner et al. 2012; Holata et al. 2018). These models can then be used to construct normalised digital surface models (nDSM), which are obtained by subtracting the DTM from the DSM Equation (1) (Rose 2021). The nDSM provide precise measurements of plant height, a critical component for determining aboveground biomass (AGB) (Viljanen et al. 2018; Rose 2021).

$$\mathrm{nDSM} = (\mathrm{DSM} - \mathrm{DTM}) \tag{1}$$

Multiple studies have shown that using drone-based RGB images and vegetation indices is useful for estimating biomass (Bazzo et al. 2023; Wang et al. 2023). For example, Bareth et al. (2015) used drone-based canopy height models (CHMs) to quantify grass height and biomass. They discovered that photogrammetric CHMs provided precise estimates of height throughout the advanced phases of plant growth. Nevertheless, these models exhibited lower precision in the first phases of development when the canopy was less dense. Possoch et al. (2016) also used, a low-cost RGB drone system to calculate canopy height models (CHMs) and RGB vegetation indices (RGBVIs) to estimate grassland biomass. The study emphasises the promise of using drones for remote sensing to measure aboveground biomass (AGB) (Possoch et al. 2016). However, they also emphasise the need for more studies to enhance accuracy across various development phases and kinds of vegetation (Possoch et al. 2016). Vegetation indices obtained from RGB drone footage are crucial for assessing vegetation characteristics and states. Indices such as BGVI, excess green index (ExG), GBVI, GRBI, normalised green-red difference index (NGRDI), and RBVI provide useful insights into the distribution and health of vegetation (Kazemi and Parmehr 2023; Lóránt et al. 2024). These indices use the spectral characteristics of vegetation to increase the difference between different vegetation types and emphasise vegetation's presence (Bareth et al. 2015; Possoch et al. 2016; Kazemi and Parmehr 2023). This makes them valuable for applications such as remote sensing and environmental monitoring. However, RGB imagery lacks the spectral richness of hyperspectral data, which can capture a broader range of wavelengths and provide more detailed spectral information for AGB estimation. Hyperspectral data can mitigate issues like spectral saturation and provide more accurate AGB estimates across different growth stages due to its ability to capture subtle variations in reflectance (Liu, Fan, Feng, et al. 2024; Liu, Feng, Fan, et al. 2024; Liu, Feng, Yue, et al. 2024). Nonetheless, hyperspectral sensors are often more expensive, require more complex data processing, and have higher computational demands, which can hinder their widespread adoption. The hyperspectral approach also faces challenges in achieving robustness across various growth stages and environmental conditions due to the dynamic nature of spectral features (Liu, Feng, Yue, et al. 2024).

While texture feature extraction from RGB images presents an alternative method that can capture detailed spatial patterns and structures of vegetation, research often emphasises height and spectral features due to their significant effectiveness in biomass estimation (Liu, Feng, Yue, Li, et al. 2022). Texture features, such as those derived from the Gray Level Co-occurrence Matrix (GLCM) and Gabor filters, can provide detailed information about vegetation characteristics, but they add complexity to the data processing pipeline and require additional computational resources and time (Mirzapour and Ghassemian 2015; Azeez et al. 2018; Liu, Feng, Yue, Jin et al. 2022; Liu, Feng, Yue, Li, et al. 2022; Liu et al. 2023). Previous research has shown that crop height directly correlates with the vertical structure of vegetation, making it a crucial indicator of above-ground biomass (AGB), while spectral indices derived from UAV images effectively capture vegetation health and vigor (Liu, Feng, Yue, et al. 2022; Cui et al. 2023; Liu et al. 2023). Studies have demonstrated that combining crop height with texture features such as GLCM-based and Gabor-based textures significantly improves the accuracy of biomass estimation models (Liu, Feng, Yue, et al. 2022; Liu et al. 2023). However, focusing on height and spectral features simplifies data processing and analysis, making it more efficient while maintaining high accuracy (Poley and McDermid 2020). Thus, while texture features offer some advantages, the increased complexity and resource demands often lead researchers to leverage the effectiveness and efficiency of height and spectral features in biomass estimation (Poley and McDermid 2020).

Principal component analysis (PCA) is a robust statistical method that reduces the number of variables in complex datasets while maintaining essential information, making it valuable for investigating correlations between above-ground biomass (AGB) and various vegetation indices, calculated nDSM, and chlorophyll content (Abonyi et al. 2022; Marukatat 2023). By identifying the primary components that explain the highest variation, PCA simplifies analysis and highlights the most relevant factors impacting AGB, which can be further refined through stepwise regression to pinpoint the most influential predictors for accurate biomass estimation (Ruengvirayudh and Brooks 2016). Stepwise regression, a process of adding and removing variables based on their statistical significance, has been shown in various studies to select the most predictive subset for the final model, balancing complexity and forecasting accuracy (Yamashita et al. 2007; Wagner and Shimshak 2007; Yuan et al. 2021). Research has demonstrated that stepwise regression produces a clear, comprehensible model suitable for biomass estimation and can achieve similar accuracy to other regression techniques like support vector machines (SVM) and partial least squares regression (PLSR) (Marabel and Alvarez-Taboada 2013; Shi et al. 2013; Ruengvirayudh and Brooks 2016; Kiala et al. 2016). By combining PCA with stepwise regression, researchers can leverage the strengths of both methods: PCA's ability to condense and summarise complex data and stepwise regression's capability to fine-tune the model by selecting the most pertinent predictors. This hybrid approach results in a more robust, parsimonious, and accurate model for estimating above-ground biomass, as it integrates comprehensive data representation with statistical rigour in variable selection (Shukla et al. 2004; Li et al. 2008; Liu, Feng, Yue, Li, et al. 2022).

The precise assessment of aboveground biomass (AGB) in the Cradle Nature Reserve is a scientific endeavour crucial for efficient ecosystem administration. Understanding the distribution of AGB is critical to effectively monitoring plant health, controlling the grazing patterns of wild animals, and preserving ecological equilibrium. This research aims to provide essential data to the reserve's management, enabling well-informed decisions about conservation policies, sustainable tourism, and habitat restoration initiatives. Additionally, precise AGB assessments will significantly evaluate carbon stocks, contributing to broader environmental and climate change mitigation objectives. To achieve this, we will utilise data from southern Africa's 2018 AGB at 50 m resolution, RGB drone-derived vegetation indices, vegetation heights obtained from nDSM, and field-measured chlorophyll concentration. We will employ PCA to condense and summarise complex data and stepwise regression to fine-tune the model by selecting the most pertinent predictors, providing a robust, parsimonious, and accurate model for estimating AGB. Given the complexity and resource demands associated with texture-based methods, we have chosen to focus on height and spectral features, which have proven to be highly effective and efficient in previous research.

2. Materials and method

2.1. Study area

The Cradle Nature Reserve is located in the Sterkfontein Valley, about 50 km northwest of Johannesburg, South Africa, and about 10 kilometres north of Krugersdorp (Lelliott 2016; Rogerson and van der Merwe 2016). The site occupies an estimated 8000 hectares within the Cradle of Humankind World Heritage Site (COHWHS) (Lelliott 2016). The reserve is located within the geographical coordinates of 27°42'58" and 27°52'57" longitude and 25°51'13" and 25°51'19" latitude (see Figure 1). In 1999, UNESCO designated this location as a world heritage site because of its profound significance in the study of paleoanthropology (Rogerson and van der Merwe 2016). The COHWHS covers a total area of 47,000 hectares in South Africa's Gauteng and North-West regions. The average annual rainfall in the nature reserve ranges from 650 to 750 mm, while temperatures may reach a maximum of 39 °C during summer and drop as low as −12 °C in winter (SA-Venues.com 2022).

In the Cradle Nature Reserve, the chemical weathering of limestone and dolomitic rocks forms karst landforms, resulting in a denudated pattern of rocky elements (Bradley et al. 2010). The distinctive landscape, shaped by a gradual process of disintegration, enhances the environmental diversity and promotes a high concentration of plant species in the area (Eloff 2010). The karst environment supports various plant life, such as grasslands and wooded regions along ravines and dolomitic sinkholes (FLOW Communications 2022). These places provide natural fire protection for trees, like white stinkwood and other plants (FLOW Communications 2022). We chose two plots for this study: a non-riparian forest region and a riparian forest area.

Plots 1 and 2, Figure 1, located within the rolling plains of the study area and part of the Magaliesberg Biosphere (designated an international Biosphere Reserve by UNESCO on June 9, 2015), showcase the diverse landscape of the savanna biome known as Moot Plains Bushveld (Magaliesberg Biosphere Home Page – Magaliesberg Biosphere 2024). Characterised by stony clay-loam soils, Plot 1 gently slopes to the northeast and features a mix of open to closed low-growing thorn savannas dominated by Acacia species, particularly in the bottomlands and plains. Open areas intersperse dense clusters of green

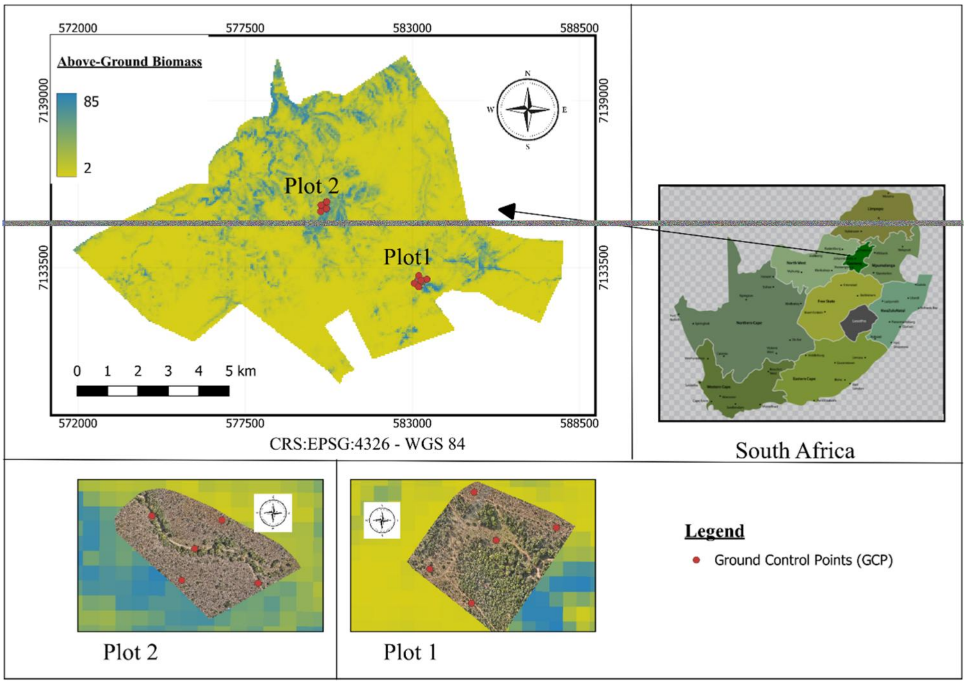

Figure 1. Study area-cradle nature reserve in South Africa, Gauteng Province.

vegetation, with higher tree concentrations in the central and lower parts of the plot transitioning to more open spaces towards the edges. Plot 2, with a northwest slope, includes a stream tributary of the Crocodile River, lined with tall trees reflecting riparian vegetation. The rest of Plot 2 features a varied vegetation density typical of savanna ecosystems. Both locations boast dense stands of native tall trees, making them ideal for accurately measuring aboveground biomass (AGB) and comparing various developed models. These plots provide a wide range of climatic variables for evaluating and confirming techniques for estimating AGB, highlighting the distinct bands of vegetation adapted to local soil and topographic conditions and emphasising the environmental gradients created by slope and water availability.

2.2. Procedure for doing random field measurements of chlorophyll using the Apogee MC-100 chlorophyll concentration meter and N3 IMU GNSS receiver

The aim of this approach was to quantify and document the chlorophyll levels in different species of plants at randomly selected field sites using the Apogee MC-100 Chlorophyll Concentration Meter and to establish the XY coordinates using the N3 IMU GNSS Receiver. This procedure's necessary equipment and supplies consisted of an Apogee MC-100 Chlorophyll Concentration Meter, an N3 IMU GNSS Receiver, a pole pruner, a moist towel, and a field notebook. Before conducting any tests, the Apogee MC-100 Chlorophyll Concentration Meter underwent calibration in accordance with the manufacturer's instructions to guarantee precise results (Apogee Instruments 2024). The data collection was conducted randomly in several field areas consisting of grass, bushes, and trees. This approach ensured an impartial sampling process. A pole pruner was used to reach the uppermost foliage of the towering trees. Before measuring the chlorophyll content, the gathered leaves were delicately wiped with a moist cloth to eliminate soil or debris. The chlorophyll concentration in the field was determined by putting the Apogee MC-100 sensor on a representative leaf, establishing proper contact, and waiting for the reading to reach a stable state. Each plot was considered as one field, with several measurements of chlorophyll concentration collected from various leaves to ensure accurate and representative outcomes. Plot 1 included 34 trees, 13 shrubs, and 15 grass samples, while Plot 2 included 45 trees, 12 shrubs, and 13 grass samples. The XY coordinates of the location were recorded using the N3 IMU GNSS Receiver, which was positioned to have an unobstructed view of the sky (ComNav Technology Ltd 2024). The vegetation types, tree (TR), shrub (SH), and grass (GR), were recorded as "Code" on the GPS controller, while the chlorophyll content was recorded as ‘Tag’. To assure accuracy, all data inputs on the GPS controller were cross-verified with the field notebook.

2.3. Construction of 3D models using the techniques of photogrammetry and UAV imaging

The process of creating a very precise 3D model using photogrammetry entails a series of crucial phases, beginning with the development of Ground Control Points (GCPs). Georeferenced Control Points (GCPs) are very accurate geographical markers that improve the spatial precision of the model. In this research, strategically positioned Ground Control Points (GCPs) were set inside the surveyed area covered by the DJI Mavic 3T drone. The DJI Mavic 3 T drone features a 48 MP wide-angle camera with a 1/2-inch CMOS sensor, a tele camera, and a thermal imaging camera to provide high-resolution RGB and thermal images. The camera has an equivalent focal length of 24 mm, ensuring a wide field of view for capturing large areas. The drone offers a maximum flight time of up to 46 min and a transmission range of up to 15 km, allowing extensive area coverage. Its omnidirectional obstacle sensing and avoidance systems ensure safe operation, while the foldable design provides portability and quick deployment. In this study, the flying height was set to 60 meters, resulting in a spatial resolution of approximately 3.5 cm per pixel. Each image had a resolution of 4000 × 3000 pixels, which confirms high-resolution data capture suitable for detailed analysis. These advanced characteristics make the DJI Mavic 3 T an ideal tool for detailed and precise data collection in environmental monitoring (MapperX.com 2023)).

To get precise coordinates, the GCPs were surveyed using high-precision GPS (N3 IMU RTK GNSS Receiver [34]) technology. During the aerial survey, the drone flew in altitude mode at 60metres above ground level and took pictures of the regions that overlapped with each other. This ensured that Ground Control Points (GCPs) were visible in numerous pictures. This would help align the photos accurately during the processing step in Agisoft Metashape Professional version 2.1.1 (Stöcker et al. 2017; Agisoft 2024). After aligning the photos, Agisoft Metashape’s Build Dense Cloud function produced a dense point cloud. The function was set to the maximum quality and was compatible with the hardware being used. This process greatly enhances the amount of 3D information compared to the original point cloud, effectively capturing the surface characteristics with high precision (Westoby et al. 2012). The compact point cloud was thereafter carefully refined to eliminate anomalies and incorrect points that do not adequately depict the surface. To maintain the accuracy of the dataset, any points that were found to be above the vegetation or below the real topography were detected and removed.

This is an important step, especially in places with diverse surface characteristics, such as forests and urban contexts (James and Robson 2012). After refining the point cloud, Digital Surface Models (DSM) and Digital Terrain Models (DTM) were generated. The DSM, including all visible structures such as buildings and plants, was immediately produced from the high-density point cloud (Remondino et al. 2014; Oliveira et al. 2024) and vehicles captured were manually removed from the DSM. Superfluous features were eliminated so the DTM could reflect the naked ground surface. Ultimately, an orthomosaic was produced by merging the georeferenced aerial photos into a smooth and accurate map that considers differences in topography and distortions caused by perspective (Dandois and Ellis 2010; Dandois and Ellis 2013). The orthomosaic underwent raster modification to extract distinct red, green, and blue bands, which accurately depict the colour information of the scene. The DTM, DSM, and RGB bands were exported as final outputs for viewing, analysis, and integration using Quantum GIS version 3.28.3-Firenze. This export ensured precise texture rendering and interoperability with 3D viewers (Verhoeven 2011).

2.4. Use of AGB 50 m resolution for Southern Africa (Bouvet et al. 2018)

The 2018 AGB dataset from the Theia African Biomass Map was chosen for this study mainly because of the Cradle Nature Reserve's consistent land use and plant cover. In recent years, the Cradle Nature Reserve has consistently maintained land use patterns that mostly revolve around animal grazing and browsing, with no notable changes in agricultural or urban growth. Research has shown that regions characterised by consistent land use often have limited fluctuations in biomass over a period of time (Numata et al. 2010; Bessou 2018; Erb et al. 2018; Baldocchi and Penuelas 2019). In addition, the plant cover, which includes forests and riparian landscapes, has remained mostly unaltered, save for isolated- occurrences of invasive weed incursions, which have had no substantial influence on the main study region. Historical AGB data, consistent with plant cover, is a reliable reference point for estimating current biomass levels (Zianis et al. 2005). We obtained the 50-meter-resolution map from Theia and cropped it using the research area shapefile in QGIS (QGIS Documentation documentation 2024).

In 2012, nine French government institutes specialising in earth observation and environmental sciences founded Theia, which focuses on the study of continental surfaces (Bouvet et al. 2018). Theia offers the scientific community and public policy a diverse array of pictures, goods, methodologies, and training pertaining to monitoring land surfaces from space (Bouvet et al. 2018). Theia also provides a dataset that estimates woody AGB at a resolution of 50 m $\times$ 50 m for savannahs and woods in continental Africa and Madagascar (Bouvet et al. 2018). These estimates are made by using 144 reference plots and are obtained through a Bayesian inversion of an L-band synthetic aperture radar (SAR) mosaic from ALOS PALSAR (horizontal transmit, horizontal receive (HH), and horizontal transmit vertical receive (HV) polarisations) (Bouvet et al. 2018). The mosaic has a resolution of 50 m x 50 m. In utilising the 2018 above-ground biomass map, this study recognises the constraints associated with using the information from 2018. The time difference between the data collection and the current research period creates uncertainty because environmental variables such as changes in climate might have influenced biomass quantities (Xiao et al. 2019; Bhan et al. 2021). Furthermore, the dataset's 50 m resolution may not be sufficient to capture small-scale changes that drone data with a higher resolution could detect. Although there are limits, historical AGB data may provide useful insights and act as a benchmark for identifying patterns and changes compared to more current, higher-resolution data (Zianis et al. 2005; Goetz et al. 2009). By recognising and considering these limitations, the research guarantees a fair and scientifically rigorous examination of the amount of biomass above the ground in the two Plots studied in the Cradle Nature Reserve.

2.5. Analysis of vegetation using QGIS and UAV imaging

The essential datasets for each plot, consisting of Red, Green, and Blue (RGB) photos, a Digital Surface Model (DSM), and a Digital Terrain Model (DTM), were imported into QGIS, an open-source Geographic Information System. As stated, these datasets were processed from drone photographs using Agisoft Metashape Professional. In addition, chlorophyll data obtained from field surveys utilising an N3 IMU GNSS Receiver were prepared for integration. The data layers were imported into QGIS as raster layers for raster data analysis (QGIS Documentation documentation 2024). The RGB pictures generated several vegetation indices, while the DSM and DTM were utilised to produce the nDSM.

The QGIS raster calculator tool was used to calculate several vegetation indices that are crucial for vegetation study (QGIS Documentation documentation 2024; Lóránt et al. 2024). The indices included the BGVI, ExG, BVI, GRBI, NGRDI, and RBVI. Table 1 provides the precise calculations applied to the RGB bands to derivate each index (Agapiou 2020; Kazemi and Parmehr 2023; Lóránt et al. 2024).

The vegetation indices in Table 1 were selected due to their effectiveness in various vegetation monitoring and biomass estimation applications. BGVI and BVI are sensitive to chlorophyll content, indicating plant health, while ExG differentiates green vegetation from other land cover types. GRBI and NGRDI assess vegetation vigour and are less affected by soil background effects, and RBVI enhances the contrast between vegetation and non-vegetation areas. These indices collectively provide a comprehensive assessment of vegetation condition, enhancing the accuracy of aboveground biomass estimation (Xue and Su 2017).

Next, we determined the nDSM by subtracting the DTM from the DSM (Turner et al. 2012; Holman et al. 2016). This stage was crucial in comprehending the structure and height of vegetation. The calculated indices and nDSM provided comprehensive information on vegetation attributes, which is crucial for further spatial analysis. To combine the chlorophyll data collected during field surveys with the calculated vegetation indices and nDSM, we imported the field data, consisting of XY coordinates, into QGIS as a point layer. Subsequently, the 'Sample Raster Values' tool (QGIS Documentation documentation 2024) was used to get the associated raster values (vegetation indices and nDSM) for every field data point. This integration facilitated the creation of a complete dataset by merging data collected in the field with data obtained *via* remote sensing. The obtained data was exported in CSV format, including chlorophyll readings, vegetation indices, and nDSM values. The CSV data were then used in Minitab version 18 for the Principal Component examination (PCA) and regression modelling. This allowed for a comprehensive statistical examination of the correlation between chlorophyll content, several

Table 1. Vegetation index equations using an RGB drone image.

Equations	Formula
1	BGVI=(Blue-Green)/(Blue + Green) (Yu et al. 2018)
2	ExG = 2 × Green – Red – Blue (Meyer and Neto 2008)
3	GBVI=(Green-Blue)/(Green + Blue) (Mehrotra and Srinivasan 2022)
4	GRBI=(Green-Red)/(Green + Blue)
5	NGRDI=(Green-Red)/(Green + Red) (Tucker 1979)
6	RBVI=(Red-Blue)/(Red + Blue) (Kawashima and Nakatani 1998)

vegetative factors, and vegetation heights (nDSM) (Wagner and Shimshak 2007; Ruengvirayudh and Brooks 2016).

2.6. Performing principal component analysis (PCA) and stepwise regression analysis using Minitab version 18

This approach utilised Minitab Version 18 to conduct PCA and Stepwise Regression Analysis on a dataset comprising variables such as AGB, normalised digital surface model (nDSM), chlorophyll concentration (Chloro), and vegetation indices derived from the RGB DJI Mavic 3 T for Plot 1 and Plot 2. First, the datasets are imported into Minitab, and PCA is performed to minimise the dimensionality of the data by finding the principal components that account for the bulk of the variation. This stage included selecting relevant factors, adjusting the number of principal components based on cumulative explained variance, and interpreting the outcomes to comprehend the contributions of each variable.

Afterwards, a stepwise regression analysis was conducted to determine the most accurate prediction model for AGB using nDSM, chloro, and vegetation indicators as predictors. This procedure involves putting up sequential configurations, executing the regression, and assessing the model's performance using R-squared values and Variance Inflation Factor (VIF) to tackle multicollinearity. Variables exhibiting high VIF values were eliminated, and the model was then rerun to check its trustworthiness. The completed model is verified by identifying relevant predictors and ensuring low VIF values. The validation process includes detailed recording and reporting of the results, which includes regression equations and insights into the dependability of the model.

3. Results and analysis

3.1. Comparison of variable statistics between non-riparian and riparian areas

The results in Table 1, non-riparian area, and Table 2, riparian area, show that, in general, the mean values for variables in Plot 2 (riparian area) with 70 observations are slightly lower than those in Plot 1 (non-riparian area) with 80 observations. For instance, the mean AGB decreases from 44.33 in Plot 1 to 23.63 in Plot 2, and the mean nDSM decreases from 7.21 in Plot 1 to 5.41 in Plot 2. Additionally, the standard deviations are generally lower in Plot 2, indicating less variability in the riparian area dataset. Notably, the mean chlorophyll (Chloro), field-observed, decreases from 314.52 in Plot 1 to 287.22 in Plot 2. The standard deviation of chlorophyll is also lower in Plot 2 (165.14) compared to Plot 1 (205.97), suggesting less variability in chlorophyll levels within the riparian area compared to the non-riparian area (Table 3).

Table 2. Summary statistics plot 1 (non-riparian).

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
Ndsm	80	0	80	0.00	19.00	7.21	5.91
Chloro	80	0	80	24.30	877.80	314.52	205.97
AGB	80	0	80	2.00	85.00	44.33	32.16
BGVI	80	0	80	−0.35	−0.01	−0.16	0.08
ExG	80	0	80	−17.00	71.00	20.06	17.25
GBVI	80	0	80	0.01	0.37	0.17	0.08
GRBI	80	0	80	0.45	0.75	0.56	0.06
NGRDI	80	0	80	−0.15	0.09	−0.03	0.05
RBVI	80	0	80	0.05	0.35	0.19	0.07

Table 3. Summary statistics plot 2 (riparian).

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
nDSM	70	0	70	0.00	16.00	5.41	4.38
Chloro	70	0	70	19.10	693.60	287.22	165.14
AGB	70	0	70	6.00	44.00	23.63	9.35
BGVI	70	0	70	−0.56	−0.03	−0.18	0.12
ExG	70	0	70	−10.00	54.00	17.49	15.70
GBVI	70	0	70	0.03	0.39	0.15	0.09
GRBI	70	0	70	−0.21	0.06	−0.05	0.05
NGRDI	70	0	70	−0.12	0.04	−0.03	0.04
RBVI	70	0	70	0.00	0.47	0.20	0.09

3.2. Comparison of principal components analysis (PCA) results

Tables A1 (a) and A2 (a), Appendix A both present the eigenvalues, variability percentages, and cumulative percentages obtained from principal components analysis (PCA), highlighting their importance in explaining the variability of the data. In Table A1 (a), a non-riparian area, F1 explains 68.07% of the overall variance, with F1 alone accounting for the highest proportion, followed by F2 and F3. The combination of F1 and F3 contributes to 96.35% of the variability. Similarly, Table A2(a), riparian area, shows that F1 has the greatest eigenvalue, explaining 39.27% of the total variance. The cumulative percentages indicate that F1 alone contributes 39.27%, with the combined impact of F1, F2, and F3 accounting for 70.62%. In both cases, most of the variability stems from the F1 factor, and we should disregard eigenvalues with smaller significance.

Furthermore, the eigenvectors in Tables A1 (b) and A2 (b), Appendix A, both reveal the individual impacts of several variables on Factor 1 (F1). In the non-riparian area, Table A1 (b), the significant contributors to F1 include nDSM (0.34), Chloro (0.34), AGB (0.30), ExG (0.38), GBVI (0.36), GRBI (0.38), and RBVI (0.23), with ExG, Chloro, and GRBI making notably significant contributions. Similarly, in the riparian area, Table A2 (b) shows that nDSM (0.39), AGB (0.38), ExG (0.38), GBVI (0.38), and Chloro (0.37) are significant contributors, indicating that F1 is a combination of these variables.

Finally, Tables A1(d) and A2(d), Appendix A, both show the variable percentage contributions to Factor 1 (F1). In the non-riparian area, plot 1, the significant contributors to F1 are nDSM (11.81%), Chloro (11.84%), AGB (9.10%), BGVI (13.42%), ExG (14.40%), GBVI (13.02%), GRBI (14.24%), NGRDI (7.08%), and RBVI (5.12%), with GRBI, ExG, BGVI, GBVI, and Chloro having notable impacts. AGB and NGRDI have less significant impacts, and RBVI has a limited influence. Similarly, Table A2 (b), riparian area, plot 2, indicates that the significant contributors to F1 are nDSM (15.29%), Chloro (13.42%), AGB (14.50%), BGVI (12.42%), ExG (14.46%), GBVI (14.58%), NGRDI (7.47%), and RBVI (7.40%), with GBVI, ExG, AGB, Chloro, and BGVI making significant contributions. While RBVI and NGRDI also play roles, their contributions are comparatively smaller, and GRBI makes very little contribution to F1.

3.3. Pairwise comparison of regression models for non-riparian and riparian plots in the Cradle Nature Reserve

The regression analysis for non-riparian (Plot 1) and riparian (Plot 2) areas in the Cradle Nature Reserve reveals distinct differences in model performance and predictor significance (Tables 4 and 5). The non-riparian model exhibits a strong correlation ($R = 0.81$) and explains 66% of the variance in Above Ground Biomass (AGB) (R Square = 0.66),

Table 4. Stepwise regression modelling Plot 1(non-riparian).

Model summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0.81	0.66	0.64	19.38	1.22

a. Predictors: (Constant), RBVI, nDSM, GRBI, GBVI
b. Dependent Variable: AGB

Coefficients

Model		Unstandardised coefficients B	Unstandardised coefficients Std. Error	Standardised ntercoefficients Beta	t	Sig.	Collinearity statistics Tolerance	Collinearity statistics VIF
1	(Constant)	17.66	46.38		0.38	0.71		
	nDSM	4.20	0.61	0.76	6.88	<.001	0.49	2.06
	GBVI	69.78	97.90	0.17	0.71	0.48	0.11	9.28
	GRBI	24.02	91.90	0.04	0.26	0.80	0.23	4.40
	RBVI	−141.32	69.38	−0.31	−2.04	0.05	0.26	3.79

Collinearity diagnostics

Model	Dimension	Eigenvalue	Condition Index	Variance proportions (Constant)	nDSM	GBVI	GRBI	RBVI
1	1	4.58	1.00	0.00	0.01	0.00	0.00	0.00
	2	0.30	3.91	0.00	0.48	0.00	0.00	0.00
	3	0.09	6.99	0.01	0.19	0.06	0.00	0.09
	4	0.03	13.48	0.00	0.29	0.34	0.01	0.54
	5	0.00	61.79	0.99	0.04	0.60	0.99	0.36

Residuals statistics

	Minimum	Maximum	Mean	Std. deviation	*N*
Predicted Value	1.06	97.73	43.68	26.12	62.00
Residual	−43.19	35.98	0.00	18.73	62.00
Std. Predicted Value	−1.63	2.07	0.00	1.00	62.00
Std. Residual	**−2.23**	**1.86**	**0.00**	**0.97**	**62.00**

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
AGB	18	2	84	46.56	33.06
Predicted _AGB	18	6.70	90.20	52.55	27.04
Valid N (listwise)	18				

compared to the riparian model's moderate correlation ($R = 0.57$) and 33% variance explanation (R Square = 0.33). The standard error of the estimate is higher for the non-riparian model (19.38) versus the riparian model (7.77), indicating greater variability in predictions in the non-riparian area. Additionally, both models show some positive autocorrelation in residuals, with the non-riparian model having a slightly higher Durbin-Watson value (1.22) than the riparian model (1.15). The critical predictor, nDSM, significantly influences AGB in both models, showing a more substantial impact in the non-riparian area (Beta = 0.76, $p < 0.001$) compared to the riparian area (Beta = 0.35, $p = 0.02$). In contrast, GBVI, GRBI, and RBVI are not significant predictors in either model, with GBVI showing high multicollinearity in the non-riparian model (VIF = 9.28).

The residual statistics further emphasise the differences between the two models. The non-riparian model displays a broader range of predicted values (1.06–97.73) and

Table 5. Stepwise regression modelling Plot 2 (riparian).

Model summary

Model	R	R square	Adjusted R square	Std. error of the estimate	Durbin-Watson
1	0.57	0.33	0.27	7.77	1.15

a. Predictors: (Constant). nDSM. RBVI. GRBI. GBVI
b. Dependent Variable: AGB

Coefficients

Model		Unstandardised coefficients		Standardised coefficients			Collinearity Statistics	Collinearity Statistics
		B	Std. Error	Beta	*t*	Sig.	Tolerance	VIF
1	(Constant)	15.51	3.30		4.70	<.001		
	GBVI	20.02	17.48	0.20	1.15	0.26	0.50	1.99
	GRBI	37.13	24.82	0.19	1.50	0.14	0.95	1.05
	RBVI	13.16	17.30	0.12	0.76	0.45	0.61	1.65
	nDSM	0.74	0.31	0.35	2.42	0.02	0.76	1.32

Collinearity diagnostics

Model	Dimension	Eigenvalue	Condition Index	Variance proportions				Variance proportions
				(Constant)	GBVI	GRBI	RBVI	nDSM
1	1	4.06	1.00	0.01	0.01	0.02	0.01	0.01
	2	0.51	2.82	0.00	0.01	0.42	0.00	0.27
	3	0.24	4.11	0.01	0.06	0.42	0.10	0.51
	4	0.13	5.59	0.45	0.46	0.12	0.01	0.00
	5	0.06	8.55	0.53	0.46	0.03	0.89	0.20

Residuals statistics

	Minimum	Maximum	Mean	Std. deviation	*N*
Predicted value	13.50	33.37	22.64	5.20	47
Residual	−12.37	17.01	0.00	7.43	47
Std. predicted value	−1.76	2.07	0.00	1.00	47
Std. residual	−1.59	2.19	0.00	0.96	47

Descriptive statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
AGB	23	31	13	44	25.65	9.80	95.96
Predicted_AGB	23	15.27	15.75	31.02	24.44	4.70	22.12
Valid N (listwise)	23						

residuals (−43.19 to 35.98), reflecting more significant prediction variability. In comparison, the riparian model shows a narrower range of predicted values (13.50–33.37) and residuals (−12.37 to 17.01). The descriptive statistics indicate that AGB in the non-riparian area has a higher mean (46.56) and greater variability (Std. Deviation = 33.06) compared to the riparian area, which has a lower mean AGB (25.65) and less variability (Std. Deviation = 9.80). The predicted AGB values are close to the actual values in both areas but show a slight overestimation in the non-riparian area. The non-riparian model demonstrates a stronger fit but faces significant multicollinearity and greater prediction variability. In contrast, despite explaining less variability in AGB, the riparian model benefits from lower multicollinearity and more consistent predictions. These findings highlight the different dynamics in AGB prediction for non-riparian and riparian plots, underscoring the importance of considering local environmental characteristics in predictive modelling.

3.4. Comparative performance of regression models for non-riparian and riparian plots in the Cradle Nature Reserve

The regression models for non-riparian and riparian plots in the Cradle Nature Reserve reveal distinct differences in their predictive performance and residual behaviour. The non-riparian model demonstrates a stronger correlation between predicted and observed AGB values ($R = 0.81$) and explains a higher proportion of variance (R Square = 0.66) compared to the riparian model ($R = 0.57$, R Square = 0.33). However, the non-riparian model exhibits more pronounced systematic bias, with residuals showing a discernible pattern where AGB is overestimated for predicted values around 60-80 and underestimated for values below 40 and above 80. This model also shows significant outliers, with residuals ranging from −60 to 40, indicating substantial prediction errors. Additionally, heteroscedasticity is more severe in the non-riparian model, with the spread of residuals increasing with higher predicted AGB values, suggesting that the model's reliability decreases for higher biomass estimates (Figures 2 and 3).

In contrast, while explaining less variance, the riparian model displays a more random distribution of residuals with fewer pronounced biases. The residual range is narrower (−10 to 20), indicating smaller deviations between observed and predicted values. This model exhibits slight heteroscedasticity, with residuals clustering around predicted AGB values of 25–30, but the effect is less pronounced than in the non-riparian model. Despite the lower R Square value, the riparian model's consistency and fewer significant outliers suggest better overall performance in terms of prediction accuracy. The significant

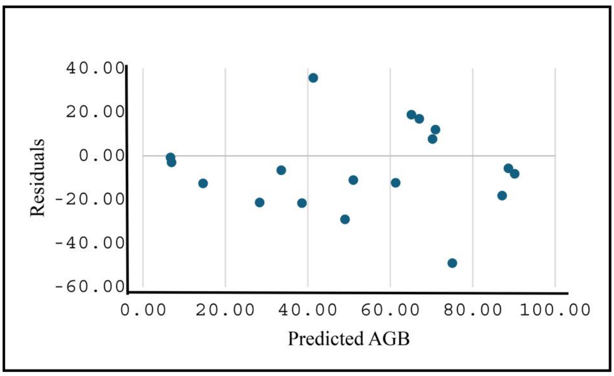

Figure 2. Scatter plot of residuals versus predicted AGB for Plot1 (non-riparian).

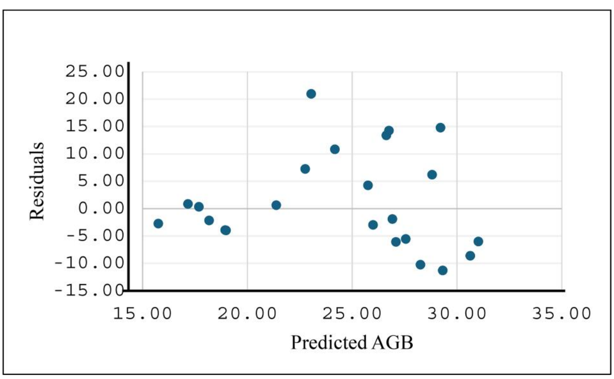

Figure 3. Scatter plot of residuals versus predicted AGB for Plot 2 (riparian).

predictors, such as nDSM, show a moderate impact on AGB in both models but with a stronger influence in the non-riparian area (Beta = 0.76) compared to the riparian area (Beta = 0.35). These differences highlight the need for tailored model refinements, such as addressing systematic biases and heteroscedasticity, to improve the robustness and reliability of AGB predictions in both non-riparian and riparian plots.

4. Discussion

4.1. Effectiveness of RGB drone imagery in estimating above-ground biomass

The study demonstrates mixed results regarding the effectiveness of RGB drone imagery in estimating above-ground biomass (AGB) for non-riparian and riparian forest areas. In the non-riparian area, Factor 1 (F1) explains 68.07% of the variance, with significant contributors including nDSM (0.34), Chloro (0.34), ExG (0.38), GBVI (0.36), and GRBI (0.38). In the riparian area, F1 explains 39.27% of the total variance, with significant contributors being nDSM (0.39), AGB (0.38), ExG (0.38), GBVI (0.38), and Chloro (0.37). Regression analysis reveals that the non-riparian model explains 66% of the variance in AGB ($R^2 = 0.66$) with a strong correlation ($R = 0.81$), while the riparian model explains only 33% of the variance ($R^2 = 0.33$) with a moderate correlation ($R = 0.57$). The critical predictor, nDSM, significantly influences AGB in both models, with a more substantial impact in the non-riparian area (Beta = 0.76, $p < 0.001$) compared to the riparian area (Beta = 0.35, $p = 0.02$). Despite the overall usefulness of RGB drone imagery, the lower R^2 values in the riparian model indicate that the indices derived from RGB images were less effective in explaining AGB variance in these areas. The non-riparian model shows greater prediction variability, with a standard error of the estimate at 19.38, compared to 7.77 in the riparian model. Additionally, the non-riparian model has a higher Durbin-Watson value (1.22) compared to the riparian model (1.15), suggesting more positive autocorrelation. Variance Inflation Factor (VIF) values for the non-riparian model are higher, especially for GBVI (9.28), indicating more severe multicollinearity. In contrast, the riparian model benefits from lower multicollinearity and more consistent predictions, as evidenced by the tighter clustering of residuals and lower VIF values (all below 5). These findings highlight that while RGB drone imagery can be effective, its indices may not always sufficiently capture the complexity of biomass in riparian zones, particularly where vegetation heights vary significantly from grass to tall trees (Nagler et al. 2018; Dormann et al. 2013; Richardson et al. 2001).

Similar studies have shown varying degrees of effectiveness when using RGB drone imagery for AGB estimation. Bareth et al. (2015) demonstrated the utility of drone-based canopy height models (CHMs) in accurately estimating biomass, particularly in advanced stages of plant growth. Possoch et al. (2016) found that low-cost RGB drones effectively estimate grassland biomass using CHMs and RGB vegetation indices. Further research by Bendig et al. (2014) confirmed the reliability of UAV-based RGB imaging for assessing crop biomass in uniform vegetation such as potatoes and rice, showing high accuracy where vegetation height is relatively consistent. A study by Lisein et al. (2013) also established the effectiveness of UAVs in forest biomass estimation, highlighting their ability to provide high-resolution data over large areas with uniform canopy. However, our study had varying vegetation heights from grass to tall trees, which might have contributed to the lower R^2 values in riparian areas. Moreover, research by Wallace et al. (2012) underscored the potential of UAVs in mapping and monitoring forest structure, supporting the current study's findings.

4.2. Differentiating predictive performance in riparian and non-riparian vegetation models

The riparian vegetation model has an R^2 value of 0.33, explaining 33% of the AGB variation. The standard error of the estimate is 7.77, indicating moderate prediction variability. Key predictors in this model include nDSM (Beta = 0.35, $p = 0.02$), which has a moderate positive impact on AGB. The Durbin-Watson value is 1.15, showing slight positive autocorrelation in residuals. The model indicates that significant predictors like nDSM, Chloro, and ExG contribute notably to AGB, while variables such as GBVI, GRBI, and RBVI do not show significant predictive power. The lower VIF values suggest minimal multicollinearity issues. The residuals exhibit a relatively narrow range (−12.37 to 17.01), indicating smaller prediction deviations. This model emphasises the importance of vegetation indices and height in biomass estimation in riparian zones, influenced by water availability and nutrient flow. Previous studies have similarly highlighted the role of vegetation indices such as NDVI, ExG, and chlorophyll concentration in assessing biomass in riparian areas (Richardson et al. 2001; Nagler et al. 2018). Furthermore, the integration of nDSM in biomass estimation has been validated by research demonstrating its effectiveness in capturing vegetation structure and height, crucial factors in riparian ecosystems (Richardson et al. 2001; Gao et al. 2018).

Conversely, the vegetation model for non-riparian areas, which primarily consists of tall trees and does not grow near a river, demonstrates a stronger ability to explain the data, achieving an R-value of 0.81 and an R^2 value of 0.66. This model finds nDSM and certain vegetation indices to be important predictors. The nDSM has a significant and favourable impact on AGB, with a beta coefficient of 0.76 and a p-value less than 0.001. High VIF values, particularly for GBVI (9.28), indicate significant multicollinearity in the model, necessitating careful interpretation of the coefficients. The Durbin–Watson statistic's closeness to 1.22 suggests that there is a slight concern about positive autocorrelation compared to the riparian model. The findings emphasise the critical role of vegetation height and specific stress indicators in clarifying biomass in non-riparian, wooded regions. Previous studies have validated the use of nDSM in biomass estimation, highlighting its ability to capture vegetation structure and height, which are crucial factors in non-riparian ecosystems (Gao et al. 2018; Tian et al. 2024). Additionally, research has shown that vegetation indices like GBVI, despite potential multicollinearity issues, can significantly contribute to biomass estimation in forested areas (Lu et al. 2016).

The contextual disparities between the two models illustrate the many environmental conditions and biological processes that affect biomass in riparian and non-riparian zones. Previous research has shown that Chloro has a beneficial effect and BGVI has a detrimental effect in riparian environments (Nagler et al. 2018). These findings support the idea that chlorophyll concentration and stress indices are important factors in these dynamic ecosystems (Kancheva et al. 2014; Nagler et al. 2018; Raddi et al. 2022). On the other hand, nDSM is a good predictor of above-ground biomass in places without water, where tall trees are the main type of vegetation (Wilkes et al. 2018; Gao et al. 2018; Zhang et al. 2024). This aligns with previous research in forest ecology that stresses how important plant height is (Andersen et al. 2006). Distinct stressors, separate from those in riparian zones, may exacerbate the adverse effects of RBVI in non-riparian regions, underscoring the intricate interaction of environmental elements in diverse biological settings (Nagler et al. 2018). These observations emphasise the need for customised remote sensing methods to precisely evaluate and control biomass and vegetation well-being in various settings (Yuan et al. 2016; Fagua et al. 2019; Abbas et al. 2020).

4.3. Limitations of RGB drone imagery, nDSM, and chlorophyll measurements in estimating above-ground biomass in riparian and non-riparian areas

Although there are important discoveries, there are various constraints that impact the effectiveness of RGB drone photography, nDSM, and chlorophyll measurements in predicting above-ground biomass (AGB) in riparian and non-riparian environments. RGB cameras only record the wavelengths of light that are visible to the human eye, which range from around 0.4 to 0.7 µm. However, they do not catch the crucial near-infrared (NIR) and middle-infrared wavelengths that are required for the precise study of vegetation and biomass determination (Jones and Vaughan 2010; Sweet et al. 2022; Alevizos et al. 2022). This constraint reduces the dependability of vegetation health assessments and the ability to identify the first indications of vegetation strain (Sweet et al. 2022; Feng et al. 2022). The presence of high reflectance saturation in places with thick vegetation might complicate the analysis and perhaps result in incorrect estimations of biomass (Gamon et al. 1995). The lack of NIR data makes it harder to accurately calculate commonly used vegetation indices like the Normalized Difference Vegetation Index (NDVI), which limits the study's ability to fully assess the health of the vegetation. In addition, the restricted range of spectral data limits the capacity to distinguish between various species of vegetation and precisely measure moisture content. These factors are crucial for effectively estimating biomass (Gamon et al. 1995; Sweet et al. 2022; Feng et al. 2022). The analysis reveals that the nDSM has high VIF values, which suggests the presence of significant multicollinearity. This makes it challenging to acquire precise coefficient estimates, especially in riparian regions with intricate plant structures (Chave et al. 2014; Bazzo et al. 2023). Plant height, which fluctuates considerably due to environmental variables like water availability and soil conditions, greatly influences the estimation of biomass in nDSM. This variability may lead to less reliable biomass estimations. Precise computation of normalized Digital Surface Model (nDSM) necessitates accurate Digital Surface Model (DSM) and Digital Terrain Model (DTM) data. Any inaccuracies in these models may have a direct impact on the precision of the nDSM and, subsequently, biomass estimations (Lefsky et al. 2002; Abdullah et al. 2021; Feng et al. 2022; Bazzo et al. 2023). The chlorophyll content is an important factor in predicting aboveground biomass (AGB), but its correlation with biomass may vary depending on the plant species, health, and climatic circumstances. This variability introduces uncertainty in biomass predictions (Gitelson et al. 2003). The precision of chlorophyll measurements derived from Apogee MC-100 sensor may be lower compared to measurements acquired from multispectral or hyperspectral sensors. This difference in precision could have an impact on the accuracy of chlorophyll-related vegetation indices and their association with biomass. Plant and water levels are closely connected in riparian zones, which makes it harder to figure out which predictor variables have the most effect on biomass (Naiman and Décamps 1997; Zarco-Tejada et al. 2001; Gitelson et al. 2003). Significant multicollinearity complicates the regression analysis and reduces the interpretability of the model, especially for variables like nDSM and chlorophyll (Li et al. 2008; Dormann et al. 2013; Bazzo et al. 2023). In locations that are not next to a river, the predictors GBVI and GRBI have high VIF values, which suggest the presence of strong multicollinearity. As a result, it is important to interpret the coefficients with caution. The negative impacts of RBVI in areas not near rivers, where tall trees are the dominant plant species, underscore the intricate interplay of environmental factors that the existing model may not adequately represent (Baldocchi 2008; Li et al. 2008; Dormann et al. 2013). It is critical to deal with multicollinearity and ensure the stability of the regression model in order to get trustworthy estimates of biomass. Combining RGB data with extra spectral data from multispectral or hyperspectral sensors can improve the accuracy of biomass estimates and get around the problems that come with using only RGB imaging (Li et al. 2008; Asner et al. 2012; Dormann et al. 2013; Bazzo et al. 2023). To get around these problems, future research might focus on improving data collection methods, such as using different kinds of sensors, and analytical models to make above-ground biomass estimates more accurate and reliable in a range of ecological settings (Poley and McDermid 2020).

5. Conclusion

The comparative analysis of riparian and non-riparian vegetation models underscores the distinct ecological processes and environmental factors influencing above-ground biomass (AGB) in these different contexts. The study's key findings can be summarised with quantitative indicators to highlight the main conclusions. In riparian zones, the model explained 33% of the variance in AGB ($R^2 = 0.33$) with a standard error of 7.77. Significant predictors included normalised Digital Surface Model (nDSM, Beta = 0.35, $p = 0.02$) and chlorophyll content, reflecting the dynamic interactions between water availability, nutrient flow, and vegetation health. The lower multicollinearity issues in the riparian model, with all Variance Inflation Factor (VIF) values below 5, indicate the robustness of these predictors. These findings align with existing research emphasising the importance of chlorophyll and vegetation indices in assessing riparian vegetation health and productivity.

Conversely, the non-riparian vegetation model demonstrated a stronger explanatory power, explaining 66% of the variance in AGB ($R^2 = 0.66$) with a standard error of 19.38. Key predictors included nDSM (Beta = 0.76, $p < 0.001$) and the red-blue vegetation index (RBVI). The model exhibited higher multicollinearity, particularly for GBVI (VIF = 9.28), necessitating careful interpretation of the coefficients. The Durbin-Watson statistic for the non-riparian model was 1.22, suggesting slight positive autocorrelation. Despite these challenges, the significant predictors provided valuable insights into the factors driving biomass in non-riparian environments, consistent with studies on forest ecology that highlight the critical role of structural metrics like vegetation height.

Overall, these findings illustrate the necessity for tailored remote sensing approaches and specific ecological considerations when assessing and managing biomass and vegetation health across varied landscapes. The study highlights that while RGB drone imagery can be effective, its indices may not always sufficiently capture the complexity of biomass in riparian zones, particularly where vegetation heights vary significantly. Future research should focus on improving data collection methods and incorporating additional spectral data to enhance the accuracy and reliability of biomass estimates in diverse ecological settings.

Acknowledgement

This project was financed by the University of the Witwatersrand School of Geography, Archaeology, and Environmental Studies in South Africa.

Author contributions

Conceptualisation: C.M., P.M, Methodology: C.M., P.M, Validation: C.M., pm, Formal analysis: C.M., pm, Investigation: C.M., pm, Writing-Original Draft: C.M., pm, Writing-Review & Editing: C.M., pm,

Disclosure statement

No potential conflict of interest was reported by the authors.

Data availability statement

The corresponding author can provide the datasets used in the current study upon reasonable request.

References

Abbas S, Wong MS, Wu J, Shahzad N, Irteza SM. 2020. Approaches of satellite remote sensing for the assessment of above-ground biomass across tropical forests: pan-tropical to national scales. Remote Sens (Basel). 12(20):3351. doi: 10.3390/rs12203351.

Abdullah MM, Al-Ali ZM, Abdullah MM, Srinivasan S, Assi AT, Al Atiqi S. 2021. Investigating the applicability of UAVs in characterizing desert shrub biomass and developing biological indicators for the selection of suitable revegetation sites. J Environ Manage. 288:112416. doi: 10.1016/j.jenvman.2021.112416.

Abonyi J, Czvetkó T, Kosztyán ZT, Héberger K. 2022. Factor analysis, sparse PCA, and Sum of Ranking Differences-based improvements of the Promethee-GAIA multicriteria decision support technique. PLoS One. 17(2):e0264277. doi: 10.1371/journal.pone.0264277.

Agapiou A. 2020. Vegetation extraction using visible-bands from openly licensed unmanned aerial vehicle imagery. Drones. 4(2):27. doi: 10.3390/drones4020027.

Agisoft. 2024. Agisoft metashape user manual professional edition. Version 2.1.

Alevizos E, Oikonomou D, Argyriou AV, Alexakis DD. 2022. Fusion of drone-based RGB and multi-spectral imagery for shallow water bathymetry inversion. Remote Sens (Basel). 14(5):1127. doi: 10.3390/rs14051127.

Andersen H-E, Reutebuch SE, Mcgaughey RJ. 2006. A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. Can J Remote Sensing. 32(5):355–366. doi: 10.5589/m06-030.

Apogee Instruments. 2024. MC-100 chlorophyll concentration meter | Apogee Instruments. [Internet]. [accessed May 21]. https://www.apogeeinstruments.com/mc-100-chlorophyll-concentration-meter/.

Asner GP, Clark JK, Mascaro J, Galindo García GA, Chadwick KD, Navarrete Encinales DA, Paez-Acosta G, Cabrera Montenegro E, Kennedy-Bowdoin T, Duque Á, et al. 2012. High-resolution mapping of forest carbon stocks in the Colombian Amazon. Biogeosciences. 9(7):2683–2696. doi: 10.5194/bg-9-2683-2012.

Azeez N, Al-Taie I, Yahya W, Basbrain A, Clark A. 2018. Regional agricultural land texture classification based on GLCMs, SVM and decision tree induction techniques. In 2018 10th Computer Science and Electronic Engineering (CEEC); p. 131–135. doi: 10.1109/CEEC.2018.8674193.

Baldocchi D. 2008. TURNER REVIEW No. 15. "Breathing" of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust J Bot. 56(1):1–26. doi: 10.1071/BT07151.

Baldocchi D, Penuelas J. 2019. The physics and ecology of mining carbon dioxide from the atmosphere by ecosystems. Glob Chang Biol. 25(4):1191–1197. doi: 10.1111/gcb.14559.

Bareth G, Bolten A, Hollberg J, Aasen H, Burkart A, Schellberg J. 2015. Feasibility study of using non-calibrated UAV-based RGB imagery for grassland monitoring: case study at the Rengen Long-term Grassland Experiment (RGE), Germany. www.icasd.org.

Bazzo COG, Kamali B, Hütt C, Bareth G, Gaiser T. 2023. A review of estimation methods for aboveground biomass in grasslands using UAV. Remote Sens (Basel). 15(3):639. doi: 10.3390/rs15030639.

Bendig J, Bolten A, Bennertz S, Broscheit J, Eichfuss S, Bareth G. 2014. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. Remote Sens. 6(11): 10395–10412.

Bessou C. 2018. Review of the impacts on soils of land-use changes induced by non-food biomass production. In: Réchauchère O, Bispo A, Gabrielle B, Makowski D, editors. Sustainable agriculture reviews 30: environmental impact of land use change in agricultural systems [Internet]. Cham: Springer International Publishing; p. 79–125. doi: 10.1007/978-3-319-96289-4_4.

Bhan M, Gingrich S, Matej S, Fritz S, Erb KH. 2021. Land use increases the correlation between tree cover and biomass carbon stocks in the global tropics. Land (Basel). 10(11):1217. doi: 10.3390/land10111217.

Bouvet A, Mermoz S, Le Toan T, Villard L, Mathieu R, Naidoo L, Asner GP. 2018. An above-ground biomass map of African savannahs and woodlands at 25 m resolution derived from ALOS PALSAR. Remote Sens Environ. 206:156–173. doi: 10.1016/j.rse.2017.12.030.

Bradley C, Cross J, Durand JF, Ellis R, Groenewald J, Grove A, Holland M, Jamison AA, Kenyon P, Krige G, et al. 2010. The Karst system of the cradle of humankind world heritage site. Water Res Comm. 410(1):88–101.

Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, et al. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob Chang Biol. 20(10):3177–3190. doi: 10.1111/gcb.12629.

ComNav Technology Ltd. 2024. N3 IMU GNSS Receiver. [Internet]. [accessed 2024 May 21]. https://www.comnavtech.com/new_N3.html.

Cui X, Han W, Dong Y, Zhai X, Ma W, Zhang L, Huang S. 2023. Estimating and mapping soil salinity in multiple vegetation cover periods by using unmanned aerial vehicle remote sensing. Remote Sens (Basel). 15(18):4400. doi: 10.3390/rs15184400.

Dandois JP, Ellis EC. 2010. Remote sensing of vegetation structure using computer vision. Remote Sens (Basel). 2(4):1157–1176. doi: 10.3390/rs2041157.

Dandois JP, Ellis EC. 2013. High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. Remote Sens Environ. 136:259–276. doi: 10.1016/j.rse.2013.04.005.

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography. 36(1):27–46. doi: 10.1111/j.1600-0587.2012.07348.x.

Eloff G. 2010. The phytosociology of the natural vegetation occuring in the cradle of humankind world heritage site, Gauteng. South Africa. Heritage:p. 141.

Erb KH, Kastner T, Plutzar C, Bais ALS, Carvalhais N, Fetzel T, Gingrich S, Haberl H, Lauk C, Niedertscheider M, et al. 2018. Unexpectedly large impact of forest management and grazing on global vegetation biomass. Nature. 553(7686):73–76. doi: 10.1038/nature25138.

Fagua JC, Jantz P, Rodriguez-Buritica S, Duncanson L, Goetz SJ. 2019. Integrating LiDAR, multispectral and SAR data to estimate and map canopy height in tropical forests. Remote Sens (Basel). 11(22):2697. doi: 10.3390/rs11222697.

Feng H, Tao H, Li Z, Yang G, Zhao C. 2022. Comparison of UAV RGB imagery and hyperspectral remote-sensing data for monitoring winter wheat growth. Remote Sens (Basel). 14(15):3811. doi: 10.3390/rs14153811.

Flores A, Herndon K, Thapa R, Cherrington E. 2019. SAR Handbook: comprehensive Methodologies for Forest Monitoring and Biomass Estimation. The SAR Handbook Comprehensive Methodologies for Forest Monitoring and Biomass Estimation [Internet]. [accessed 2021 Oct 4]:1–307. https://www.servirglobal.net/Global/Articles/Article/2674/sar-handbook-comprehensive-methodologies-for-forest-monitoring-and-biomass-estimation.

FLOW Communications. 2022. Maropeng and Sterkfontein Caves. https://www.maropeng.co.za/content/page/environment-and-climate. [Internet]. [accessed 2022 May 23]. https://www.maropeng.co.za/content/page/environment-and-climate.

Gamon JA, Field CB, Goulden ML, Griffin KL, Hartley AE, Joel G, Peñuelas J, Valentini R. 1995. Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types.

Gao Y, Lu D, Li G, Wang G, Chen Q, Liu L, Li D. 2018. Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. Remote Sens (Basel). 10(4):627. doi: 10.3390/rs10040627.

Gitelson AA, Gritz Y, Merzlyak MN. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. http://www.urbanfischer.de/journals/jpp.

Goetz SJ, Baccini A, Laporte NT, Johns T, Walker W, Kellndorfer J, Houghton RA, Sun M. 2009. Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. Carbon Balance Manag. 4(1):2. doi: 10.1186/1750-0680-4-2.

Holata L, Plzák J, Světlík R, Fonte J. 2018. Integration of low-resolution ALS and ground-based SfM photogrammetry data. A cost-effective approach providing an "enhanced 3D model" of the hound tor archaeological landscapes (Dartmoor, South-West England). Remote Sens (Basel). 10(9):1–28. doi: 10.3390/rs10091357.

Holman FH, Riche AB, Michalski A, Castle M, Wooster MJ, Hawkesford MJ. 2016. High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. Remote Sens (Basel). 8(12):1031. doi: 10.3390/rs8121031.

James MR, Robson S. 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application. J Geophys Res. 117(F3):1–17. doi: 10.1029/2011JF002289.
Jones HG, Vaughan RA. 2010. Remote sensing of vegetation: principles, techniques, and applications. USA: Oxford University Press.
Kancheva R, Borisova D, Georgiev G. 2014. Assessment of chlorophyll and stress detection through vegetation optical properties.
Kawashima S, Nakatani M. 1998. An algorithm for estimating chlorophyll content in leaves using a video camera.
Kazemi F, Parmehr EG. 2023. Evaluation of RGB vegetation indices derived from uav images for rice crop growth monitoring. ISPRS Ann Photogramm Remote Sens Spatial Inf Sci. X-4/W1-2022:385–390. doi: 10.5194/isprs-annals-X-4-W1-2022-385-2023.
Khumalo W, Odindi J, Mngadi M. 2024. Assessing above ground biomass of communal rangelands using 2 fused optical and synthetic aperture radar (SAR) data [Internet]. https://ssrn.com/abstract=4800031.
Kiala Z, Odindi J, Mutanga O, Peerbhay K. 2016. Comparison of partial least squares and support vector regressions for predicting leaf area index on a tropical grassland using hyperspectral data. J Appl Remote Sens. 10(3):036015. doi: 10.1117/1.JRS.10.036015.
Lefsky MA, Cohen WB, Parker GG, Harding DJ. 2002. Lidar remote sensing for ecosystem studies: lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. Bioscience. 52(1):19–30. [Mismatch doi: 10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2.
Lelliott A. 2016. Visitors' views of human origins after visiting the Cradle of Humankind World Heritage Site. S Afr J Sci. 112(1/2):8. doi: 10.17159/sajs.2016/20150210.
Li Y, Andersen HE, McGaughey R. 2008. A Comparison of statistical methods for estimating forest biomass from light detection and ranging data. West J Appl For. 23(4):223–231. doi: 10.1093/wjaf/23.4.223.
Lisein J, Pierrot-Deseilligny M, Bonnet S Lejeune P. 2013. A photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. Forests. 4(4):922–944.
Liu Y, Fan Y, Feng H, Chen R, Bian M, Ma Y, Yue J, Yang G. 2024. Estimating potato above-ground biomass based on vegetation indices and texture features constructed from sensitive bands of UAV hyperspectral imagery. Comput Electron Agric. 220:108918. doi: 10.1016/j.compag.2024.108918.
Liu Y, Feng H, Fan Y, Yue J, Chen R, Ma Y, Bian M, Yang G. 2024. Improving potato above ground biomass estimation combining hyperspectral data and harmonic decomposition techniques. Comput Electron Agric. 218:108699. doi: 10.1016/j.compag.2024.108699.
Liu Y, Feng H, Yue J, Fan Y, Bian M, Ma Y, Jin X, Song X, Yang G. 2023. Estimating potato above-ground biomass by using integrated unmanned aerial system-based optical, structural, and textural canopy measurements. Comput Electron Agric. 213:108229. doi: 10.1016/j.compag.2023.108229.
Liu Y, Feng H, Yue J, Jin X, Fan Y, Chen R, Bian M, Ma Y, Li J, Xu B, et al. 2024. Improving potato AGB estimation to mitigate phenological stage impacts through depth features from hyperspectral data. Comput Electron Agric. 219:108808. doi: 10.1016/j.compag.2024.108808.
Liu Y, Feng H, Yue J, Jin X, Li Z, Yang G. 2022. Estimation of potato above-ground biomass based on unmanned aerial vehicle red-green-blue images with different texture features and crop height. Front Plant Sci. 13:938216. doi: 10.3389/fpls.2022.938216.
Liu Y, Feng H, Yue J, Li Z, Yang G, Song X, Yang X, Zhao Y. 2022. Remote-sensing estimation of potato above-ground biomass based on spectral and spatial features extracted from high-definition digital camera images. Comput Electron Agric. 198:107089. doi: 10.1016/j.compag.2022.107089.
Lóránt B, Veronika K-B, József B. 2024. Comparison of RGB indices used for vegetation studies based on structured similarity index (SSIM). J Plant Sci Phytopathol. 8(1):007–012. doi: 10.29328/journal.jpsp.1001124.
Lu D, Chen Q, Wang G, Liu L, Li G, Moran E. 2016. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. Int J Digit Earth. 9(1):63–105. doi: 10.1080/17538947.2014.990526.
Magaliesberg Biosphere Home Page – Magaliesberg Biosphere. 2024. [accessed 2024 Jul 10]. https://magaliesbergbiosphere.org.za/magaliesberg-biosphere-home-page/.
MapperX.com. 2023. DJI Mavic 3 Thermal (M3T) Technical Specifications - MapperX. [Internet]. [accessed 2024 May 21]. https://mapperx.com/en/dji-mavic-3-thermal-m3t-technical-specifications/.
Marabel M, Alvarez-Taboada F. 2013. Spectroscopic determination of aboveground biomass in grasslands using spectral transformations, support vector machine and partial least squares regression. Sensors (Basel). 13(8):10027–10051. doi: 10.3390/s130810027.
Marukatat S. 2023. Tutorial on PCA and approximate PCA and approximate kernel PCA. Artif Intell Rev. 56(6):5445–5477. doi: 10.1007/s10462-022-10297-z.
Mehrotra N, Srinivasan S. 2022. Analysing drone and satellite imagery using vegetation indices.

Meyer GE, Neto JC. 2008. Verification of color vegetation indices for automated crop imaging applications. Comput Electron Agric. 63(2):282–293. doi: 10.1016/j.compag.2008.03.009.
Mirzapour F, Ghassemian H. 2015. Fast GLCM and gabor filters for texture classification of very high resolution remote sensing images. https://www.researchgate.net/publication/289790781.
Nagler PL, Jarchow CJ, Glenn EP. 2018. Remote sensing vegetation index methods to evaluate changes in greenness and evapotranspiration in riparian vegetation in response to the Minute 319 environmental pulse flow to Mexico. Proc IAHS. 380:45–54. doi: 10.5194/piahs-380-45-2018.
Naiman RJ, Décamps H. 1997. The ecology of interfaces: riparian zones. Annu Rev Ecol Syst. 28(1):621–658. doi: 10.1146/annurev.ecolsys.28.1.621.
Numata I, Cochrane MA, Roberts DA, Soares JV, Souza CM, Sales MH. 2010. Biomass collapse and carbon emissions from forest fragmentation in the Brazilian Amazon. J Geophys Res. 115(G3):1–10. doi: 10.1029/2009JG001198.
Nyamukuru A, Whitney C, Tabuti JRS, Esaete J, Low M. 2023. Allometric models for aboveground biomass estimation of small trees and shrubs in African savanna ecosystems. Trees, Forests and People. 11:100377. doi: 10.1016/j.tfp.2023.100377.
Oliveira ASB, Fernandes JVA, Figueiredo VLFA, Leonel LCPC, Bauman MMJ, Link MJ, Peris-Celda M. 2024. 3D models as a source for neuroanatomy education: a Stepwise white matter dissection using 3D images and photogrammetry scans. Brain Topogr. doi: 10.1007/s10548-024-01058-y.
Poley LG, McDermid GJ. 2020. A systematic review of the factors influencing the estimation of vegetation aboveground biomass using unmanned aerial systems. Remote Sens (Basel). 12(7):1–46. doi: 10.3390/rs12071052.
Possoch M, Bieker S, Hoffmeister D, Bolten AA, Schellberg J, Bareth G. 2016. Multi-Temporal crop surface models combined with the rgb vegetation index from UAV-based images for forage monitoring in grassland. Int Arch Photogramm Remote Sens Spatial Inf Sci. XLI-B1:991–998. doi: 10.5194/isprsarchives-XLI-B1-991-2016.
QGIS Documentation documentation. 2024. QGIS User Guide. [Internet]. [accessed 2024 May 21]. https://docs.qgis.org/3.34/en/docs/user_manual/index.html.
Raddi S, Giannetti F, Martini S, Farinella F, Chirici G, Tani A, Maltoni A, Mariotti B. 2022. Monitoring drought response and chlorophyll content in Quercus by consumer-grade, near-infrared (NIR) camera: a comparison with reflectance spectroscopy. New For (Dordr) [Internet]. 53(2):241–265. doi: 10.1007/s11056-021-09848-z.
Rapiya M, Ramoelo A, Truter W. 2023. Seasonal evaluation and mapping of aboveground biomass in natural rangelands using Sentinel-1 and Sentinel-2 data. Environ Monit Assess. 195(12):1544. doi: 10.1007/s10661-023-12133-5.
Remondino F, Spera MG, Nocerino E, Menna F, Nex F. 2014. State of the art in high density image matching. Photogramm Rec. 29(146):144–166. doi: 10.1111/phor.12063.
Richardson MD, Karcher DE, Purcell LC. 2001. Quantifying turfgrass cover using digital image analysis determination of color and fertility differences in Zea. Crop Sci. 41(6):1884–1888. doi: 10.2135/cropsci2001.1884.
Rogerson CM, van der Merwe CD. 2016. Heritage tourism in the global South: development impacts of the Cradle of Humankind World Heritage Site, South Africa. Local Economy. 31(1-2):234–248. doi: 10.1177/0269094215614270.
Rose N. 2021. nDSMs: how digital surface models and digital terrain models elevate your insights. [Internet]. [accessed 2024 May 23]. https://up42.com/blog/ndsms-how-digital-surface-models-and-digital-terrain-models-elevate-your.
Ruengvirayudh P, Brooks GP. 2016. Comparing stepwise regression models to the best-subsets comparing stepwise regression models to the best-subsets models, or, the art of stepwise models, or, the art of stepwise. General linear model journal [Internet]. https://digitalcommons.lmu.edu/gess_fac.
SA-Venues.com. 2022. Cradle Nature Reserve, Gauteng. [Internet]. [accessed 2022 May 23]. https://www.sa-venues.com/game-reserves/cradle.php.
Shi T, Cui L, Wang J, Fei T, Chen Y, Wu G. 2013. Comparison of multivariate methods for estimating soil total nitrogen with visible/near-infrared spectroscopy. Plant Soil. 366(1-2):363–375. doi: 10.1007/s11104-012-1436-8.
Shukla MK, Lal R, Ebinger M. 2004. Principal component analysis for predicting corn biomass and grain yields. Soil Sci [Internet]. 169(3):215–224. https://journals.lww.com/soilsci/fulltext/2004/03000/principal_component_analysis_for_predicting_corn.7.aspx.
Song E, Shao G, Zhu X, Zhang W, Dai Y, Lu J. 2024. Estimation of plant height and biomass of rice using unmanned aerial vehicle. Agronomy. 14(1):145. doi: 10.3390/agronomy14010145.

Stöcker C, Nex F, Koeva M, Gerke M. 2017. Quality assessment of combined IMU/GNSS data for direct georeferencing in the context of UAV-based mapping. Int Arch Photogramm Remote Sens Spatial Inf Sci. XLII-2/W6:355–361. doi: 10.5194/isprs-archives-XLII-2-W6-355-2017.

Sweet DD, Tirado SB, Springer NM, Hirsch CN, Hirsch CD. 2022. Opportunities and challenges in phenotyping row crops using drone-based RGB imaging. Plant Phenome J. 5(1):1–20. doi: 10.1002/ppj2.20044.

Tian X, Li J, Zhang F, Zhang H, Jiang M. 2024. Forest aboveground biomass estimation using multisource remote sensing data and deep learning algorithms: a case study over Hangzhou Area in China. Remote Sens (Basel). 16(6):1074. doi: 10.3390/rs16061074.

Tucker CJ. 1979. Red and photographic infrared linear combinations for monitoring vegetation.

Turner D, Lucieer A, Watson C. 2012. An automated technique for generating georectified mosaics from ultra-high resolution Unmanned Aerial Vehicle (UAV) imagery, based on Structure from Motion (SFM) point clouds. Remote Sens (Basel). 4(5):1392–1410. doi: 10.3390/rs4051392.

Verhoeven G. 2011. Taking computer vision aloft - archaeological three-dimensional reconstructions from aerial photographs with photoscan. Archaeol Prospect. 18(1):67–73. doi: 10.1002/arp.399.

Viljanen N, Honkavaara E, Näsi R, Hakala T, Niemeläinen O, Kaivosoja J. 2018. A novel machine learning method for estimating biomass of grass swards using a photogrammetric canopy height model, images and vegetation indices captured by a drone. Agriculture (Switzerland). 8(5):70. doi: 10.3390/agriculture8050070.

Wagner JM, Shimshak DG. 2007. Stepwise selection of variables in data envelopment analysis: procedures and managerial perspectives. Eur J Oper Res. 180(1):57–67. doi: 10.1016/j.ejor.2006.02.048.

Wallace L, Lucieer A, Watson C, Turner D. 2012. Development of a UAV-LiDAR system with application to forest inventory. Remote Sens. 4(6):1519–1543. doi: 10.3390/rs4061519.

Wang G, Li S, Huang C, He G, Li Y, Feng J, Tang F, Yan P, Qiu L. 2023. Mapping the spatial distribution of aboveground biomass in China's subtropical forests based on UAV LiDAR data. Forests. 14(8):1560. doi: 10.3390/f14081560.

Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. 2012. "Structure-from-Motion" photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology. 179:300–314. doi: 10.1016/j.geomorph.2012.08.021.

Wilkes P, Disney M, Vicari MB, Calders K, Burt A. 2018. Estimating urban above ground biomass with multi-scale LiDAR. Carbon Balance Manag [Internet]. 13(1):10. doi: 10.1186/s13021-018-0098-0.

Xiao J, Chevallier F, Gomez C, Guanter L, Hicke JA, Huete AR, Ichii K, Ni W, Pang Y, Rahman AF, et al. 2019. Remote sensing of the terrestrial carbon cycle: a review of advances over 50 years. Remote Sens Environ. 233:111383. doi: 10.1016/j.rse.2019.111383.

Xue J, Su B. 2017. Significant remote sensing vegetation indices: a review of developments and applications. J Sens. 2017:1–17. doi: 10.1155/2017/1353691.

Yamashita T, Yamashita K, Kamimura R. 2007. A stepwise AIC method for variable selection in linear regression. Commun Stat—Theory Methods. 36(13):2395–2403. doi: 10.1080/03610920701215639.

Yu X, Zhao G, Chang C, Yuan X, Heng F. 2018. BGVI: a new index to estimate street-side greenery using Baidu Street View Image. Forests. 10(1):3. doi: 10.3390/f10010003.

Yuan L, Li L, Zhang T, Chen L, Liu W, Hu S, Yang L. 2021. Modeling soil moisture from multisource data by stepwise multilinear regression: an application to the Chinese loess plateau. IJGI. 10(4):233. doi: 10.3390/ijgi10040233.

Yuan X, Li L, Tian X, Luo G, Chen X. 2016. Estimation of above-ground biomass using MODIS satellite imagery of multiple land-cover types in China. Remote Sens Lett [Internet]. 7(12):1141–1149. doi: 10.1080/2150704X.2016.1219458.

Yue J, Yang G, Li C, Li Z, Wang Y, Feng H, Xu B. 2017. Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. Remote Sens (Basel). 9(7):708. doi: 10.3390/rs9070708.

Zarco-Tejada PJ, Miller JR, Noland TL, Mohammed GH, Sampson PH. 2001. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data.

Zhang J, Zhou C, Zhang G, Yang Z, Pang Z, Luo Y. 2024. A novel framework for forest above-ground biomass inversion using multi-source remote sensing and deep learning. Forests. 15(3):456. doi: 10.3390/f15030456.

Zianis D, Muukkonen P, Mäkipää R, Mencuccini M. 2005. Biomass and stem volume equations for tree species in Europe. Silva Fenn, Monogr. 2005(4):1–63. doi: 10.14214/sf.sfm4.

Appendices

Appendix A.

Table A1. Multicollinearity and principal component analysis plot 1.

Principal component analysis:

(a) Eigenvalues:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Eigenvalue	6.13	1.55	1.00	0.24	0.07	0.01	0.00	0.00	0.00
Variability (%)	68.07	17.19	11.10	2.62	0.82	0.15	0.04	0.02	0.00
Cumulative %	68.07	85.26	96.35	98.97	99.79	99.94	99.98	100.00	100.00

(b) Eigenvectors:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	0.34	−0.12	0.45	−0.45	−0.08	−0.67	−0.06	−0.03	0.00
Chloro	0.34	−0.16	0.45	−0.35	0.10	0.72	0.07	0.05	0.00
AGB	0.30	−0.28	0.41	0.81	−0.08	−0.05	0.01	−0.01	0.00
BGVI	−0.37	−0.31	0.15	−0.02	0.24	−0.09	0.26	0.51	0.60
ExG	0.38	−0.08	−0.32	0.00	−0.16	0.03	−0.52	0.67	0.00
GBVI	0.36	0.29	−0.11	0.10	0.86	−0.11	0.05	0.02	0.00
GRBI	0.38	−0.06	−0.34	−0.02	−0.23	−0.05	0.79	0.18	−0.13
NGRDI	0.27	−0.51	−0.41	−0.06	0.04	0.03	−0.14	−0.51	0.47
RBVI	0.23	0.66	0.09	0.06	−0.31	0.05	0.02	−0.06	0.63

(c) Correlations between variables and factors:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	0.85	−0.15	0.45	−0.22	−0.02	−0.08	0.00	0.00	0.00
Chloro	0.85	−0.20	0.45	−0.17	0.03	0.08	0.00	0.00	0.00
AGB	0.75	−0.34	0.41	0.39	−0.02	−0.01	0.00	0.00	0.00
BGVI	−0.91	−0.39	0.15	−0.01	0.07	−0.01	0.02	0.02	0.00
ExG	0.94	−0.10	−0.32	0.00	−0.04	0.00	−0.03	0.03	0.00
GBVI	0.89	0.37	−0.11	0.05	0.23	−0.01	0.00	0.00	0.00
GRBI	0.93	−0.08	−0.34	−0.01	−0.06	−0.01	0.05	0.01	0.00
NGRDI	0.66	−0.63	−0.41	−0.03	0.01	0.00	−0.01	−0.02	0.00
RBVI	0.56	0.82	0.09	0.03	−0.08	0.01	0.00	0.00	0.00

(d) Contribution of the variables (%):

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	11.81	1.51	19.94	20.30	0.60	45.43	0.31	0.11	0.00
Chloro	11.84	2.47	19.99	12.54	0.92	51.51	0.49	0.24	0.00
AGB	9.10	7.62	17.07	65.24	0.67	0.28	0.01	0.01	0.00
BGVI	13.42	9.76	2.20	0.06	5.79	0.79	6.50	25.55	35.94
ExG	14.40	0.65	10.44	0.00	2.49	0.07	27.38	44.56	0.00
GBVI	13.02	8.62	1.18	1.02	74.61	1.28	0.22	0.06	0.00
GRBI	14.24	0.41	11.51	0.05	5.33	0.29	63.18	3.36	1.63
NGRDI	7.08	25.67	16.80	0.39	0.18	0.08	1.88	25.75	22.18
RBVI	5.12	43.28	0.88	0.40	9.41	0.27	0.03	0.36	40.25

(e) Squared cosines of the variables:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	0.72	0.02	0.20	0.05	0.00	0.01	0.00	0.00	0.00
Chloro	0.73	0.04	0.20	0.03	0.00	0.01	0.00	0.00	0.00
AGB	0.56	0.12	0.17	0.15	0.00	0.00	0.00	0.00	0.00
BGVI	0.82	0.15	0.02	0.00	0.00	0.00	0.00	0.00	0.00
ExG	0.88	0.01	0.10	0.00	0.00	0.00	0.00	0.00	0.00
GBVI	0.80	0.13	0.01	0.00	0.05	0.00	0.00	0.00	0.00
GRBI	0.87	0.01	0.11	0.00	0.00	0.00	0.00	0.00	0.00
NGRDI	0.43	0.40	0.17	0.00	0.00	0.00	0.00	0.00	0.00
RBVI	0.31	0.67	0.01	0.00	0.01	0.00	0.00	0.00	0.00

Table A2. Multicollinearity and principal component analysis Plot 2.

Principal component analysis:

(a) Eigenvalues:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
Eigenvalue	3.53	1.85	0.97	0.88	0.55	0.48	0.44	0.28	0.03
Variability (%)	39.27	20.55	10.80	9.73	6.08	5.32	4.86	3.11	0.28
Cumulative %	39.27	59.82	70.62	80.35	86.43	91.75	96.61	99.72	100.00

(b) Eigenvectors:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	0.39	−0.39	−0.39	−0.06	−0.02	−0.04	−0.10	−0.07	0.72
Chloro	0.37	−0.42	−0.42	−0.09	0.00	−0.01	−0.09	−0.12	−0.69
AGB	0.38	−0.06	0.30	−0.18	0.76	0.06	0.19	0.34	−0.01
BGVI	−0.35	−0.40	−0.21	0.15	−0.14	−0.01	0.36	0.71	−0.01
ExG	0.38	0.13	0.08	0.37	−0.25	0.77	−0.01	0.17	−0.02
GBVI	0.38	0.31	0.06	−0.04	−0.30	−0.45	−0.42	0.53	−0.05
GRBI	0.07	−0.40	0.55	−0.57	−0.44	0.11	0.07	−0.02	0.01
NGRDI	0.27	−0.24	0.38	0.62	−0.11	−0.42	0.32	−0.21	−0.02
RBVI	0.27	0.42	−0.27	−0.29	−0.22	−0.09	0.73	−0.06	0.00

(c) Correlations between variables and factors:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	0.74	−0.53	−0.39	−0.05	−0.01	−0.02	−0.07	−0.04	0.11
Chloro	0.69	−0.57	−0.42	−0.09	0.00	−0.01	−0.06	−0.06	−0.11
AGB	0.72	−0.08	0.30	−0.17	0.56	0.04	0.12	0.18	0.00
BGVI	−0.66	−0.54	−0.21	0.14	−0.10	−0.01	0.24	0.38	0.00
ExG	0.71	0.18	0.08	0.35	−0.19	0.54	−0.01	0.09	0.00
GBVI	0.72	0.42	0.05	−0.04	−0.22	−0.31	−0.28	0.28	−0.01
GRBI	0.13	−0.54	0.55	−0.53	−0.32	0.08	0.04	−0.01	0.00
NGRDI	0.51	−0.33	0.37	0.58	−0.08	−0.29	0.21	−0.11	0.00
RBVI	0.51	0.58	−0.26	−0.27	−0.16	−0.06	0.48	−0.03	0.00

(d) Contribution of the variables (%):

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	15.29	15.46	15.53	0.33	0.03	0.13	1.02	0.52	51.69
Chloro	13.42	17.57	17.83	0.87	0.00	0.01	0.84	1.46	48.01
AGB	14.50	0.35	9.07	3.37	57.48	0.31	3.55	11.37	0.00
BGVI	12.42	15.69	4.41	2.31	1.89	0.02	12.85	50.39	0.01
ExG	14.46	1.69	0.62	13.85	6.46	59.84	0.01	3.05	0.02
GBVI	14.58	9.63	0.31	0.15	8.93	20.46	17.42	28.30	0.21
GRBI	0.46	15.68	30.75	32.15	19.27	1.22	0.43	0.03	0.00
NGRDI	7.47	5.90	14.41	38.75	1.23	17.26	10.40	4.53	0.05
RBVI	7.40	18.04	7.08	8.21	4.70	0.76	53.47	0.34	0.00

(e) Squared cosines of the variables:

	F1	F2	F3	F4	F5	F6	F7	F8	F9
nDSM	0.54	0.29	0.15	0.00	0.00	0.00	0.00	0.00	0.01
Chloro	0.47	0.33	0.17	0.01	0.00	0.00	0.00	0.00	0.01
AGB	0.51	0.01	0.09	0.03	0.31	0.00	0.02	0.03	0.00
BGVI	0.44	0.29	0.04	0.02	0.01	0.00	0.06	0.14	0.00
ExG	0.51	0.03	0.01	0.12	0.04	0.29	0.00	0.01	0.00
GBVI	0.52	0.18	0.00	0.00	0.05	0.10	0.08	0.08	0.00
GRBI	0.02	0.29	0.30	0.28	0.11	0.01	0.00	0.00	0.00
NGRDI	0.26	0.11	0.14	0.34	0.01	0.08	0.05	0.01	0.00
RBVI	0.26	0.33	0.07	0.07	0.03	0.00	0.23	0.00	0.00

(f) Correlation matrix:

Variables	nDSM	Chloro	AGB	BGVI	ExG	GBVI	GRBI	NGRDI	RBVI
nDSM	1.00	0.97	0.44	−0.16	0.37	0.30	0.20	0.38	0.16
Chloro	0.97	1.00	0.41	−0.11	0.32	0.24	0.21	0.33	0.13
AGB	0.44	0.41	1.00	−0.48	0.39	0.38	0.21	0.35	0.25
BGVI	−0.16	−0.11	−0.48	1.00	−0.49	−0.65	0.06	−0.14	−0.51
ExG	0.37	0.32	0.39	−0.49	1.00	0.48	−0.05	0.39	0.35
GBVI	0.30	0.24	0.38	−0.65	0.48	1.00	−0.05	0.25	0.52
GRBI	0.20	0.21	0.21	0.06	−0.05	−0.05	1.00	0.15	−0.18
NGRDI	0.38	0.33	0.35	−0.14	0.39	0.25	0.15	1.00	−0.05
RBVI	0.16	0.13	0.25	−0.51	0.35	0.52	−0.18	−0.05	1.00

(g) Multicollinearity statistics:

Statistic	nDSM	Chloro	AGB	BGVI	ExG	GBVI	GRBI	NGRDI	RBVI
R^2	0.95	0.95	0.42	0.57	0.41	0.54	0.14	0.31	0.38
Tolerance	0.05	0.05	0.58	0.43	0.59	0.46	0.86	0.69	0.62
VIF	20.66	19.28	1.72	2.32	1.70	2.18	1.16	1.45	1.62