



An African Tree Crown–Stem Allometry Dataset for Monitoring Trees Outside Forests and Landscape Restoration

Tabitha A. Alimo¹, David L. Skole^{1*}, Jay H. Samek¹, Dan Ndalowa², Vincent Oeba³, Judith Kamoto⁴, Alfred Chioza⁴, Chad Papa¹, Justine Spore⁵, Darby Levin⁵, John Brandt⁵, Tangu Tumeo⁶

¹Department of Forestry, Michigan State University, East Lansing USA

²Forestry Research Institute of Malawi, Zomba, Malawi

³Kenya Forest Research Institute, Nairobi, Kenya

⁴Forestry Department, Lilongwe University of Agriculture and Natural Resources, Lilongwe, Malawi

⁵World Resources Institute, Washington D.C., USA

⁶International Union for Conservation of Nature, Lilongwe, Malawi

Correspondence to: David L. Skole (skole@msu.edu)

Abstract. Very high-resolution (VHR) satellite imagery and advances in machine learning increasingly enable individual tree crowns to be mapped across large landscapes, creating new opportunities for census-based estimation of tree biomass and carbon. A fundamental limitation, however, is that satellite imagery can measure crown dimensions but cannot directly observe diameter at breast height (DBH), the principal predictor used in conventional tree biomass allometry. This measurement gap is particularly important in African landscapes dominated by trees outside forests (TOF), where trees are spatially dispersed, conventional forest inventory sampling is difficult to scale, and ground datasets containing concurrent measurements of crown and stem dimensions remain scarce. Here, we present an African individual-tree crown–stem allometry dataset designed to provide the empirical link between remotely observable crown dimensions and the stem measurements required by established biomass allometric models. The dataset was collected using a standardized field protocol across landscapes associated with African Forest Landscape Restoration Initiative (AFR100) projects in Kenya, Ghana, Rwanda, and Malawi. The sampled landscapes encompass agroforestry, reforestation, assisted natural regeneration, and other restoration and land-management systems representative of TOF conditions. Following quality control, the dataset contains 3,493 individual trees with concurrent measurements of DBH, two perpendicular crown diameters, tree height, species identity, geographic location, and associated plot and landscape identifiers. Field measurements were designed for spatial linkage with individual tree crowns delineated from VHR satellite imagery, providing calibration observations for translating remotely sensed crown projected area (CA) into stem diameter.

We demonstrate this application by fitting power-law DBH–CA transfer functions and evaluating their predictive performance across sites and countries. All site-specific relationships were significant ($p < 0.001$), with R^2 values ranging from 0.44 to 0.85 and generally small prediction bias. Variation in model coefficients and predictive error among landscapes demonstrates substantial ecological and structural heterogeneity in crown–stem relationships and emphasizes the value of geographically



distributed calibration data rather than reliance on a single universal allometric relationship. A hierarchical synthesis of 19
 35 site-specific models further demonstrates how these data can be used to derive a generalized crown-to-stem relationship while
 retaining information on among-landscape variation. The principal contribution of this dataset is therefore to provide a
 measurement bridge between individual-tree remote sensing and conventional stem-based biomass allometry. VHR imagery
 and machine-learning methods can delineate individual crowns and measure their projected areas; crown-to-stem transfer
 functions calibrated with these field data can then estimate DBH; and established DBH-based allometric equations can
 40 subsequently be used to estimate biomass and carbon for each remotely detected tree. This framework extends individual-tree
 allometric scaling from field plots to landscape-scale tree censuses while maintaining compatibility with established forest
 mensuration and biomass estimation methods. The dataset has applications beyond the models demonstrated here, including
 development and validation of alternative crown–stem and crown–biomass models, analysis of tree structural and species-
 level variation, ecological scaling studies, and calibration of individual-tree remote-sensing products. It provides a reusable
 45 observational resource for advancing the inclusion of TOF in Forest and Landscape Restoration monitoring, carbon accounting,
 and measurement, reporting, and verification across African landscapes.

1 Introduction

Landscapes of trees outside forests (TOF) are ubiquitous across the continent of Africa. They include sparse tree ecosystems,
 such as savannas, and a variety of tree-based production systems such as agroforestry, ally cropping, small-holder plantations,
 50 energy farms, shelter belts, village or community woodlots, scattered individual trees and other woody perennial
 establishments in predominantly small holder agricultural landscapes. They generally have less than 10% canopy cover (De
 Foresta et al., 2013). They are recognized for their role in biomass accumulation and carbon sequestration (Mbow et al., 2014;
 Skole, et al., 2021). Recent studies estimate that nearly 30 % of tree cover across the continent exists outside traditional forest
 boundaries (Reiner et al., 2023; Zomer et al., 2009), highlighting the ecological and climate change mitigation importance of
 55 these systems (Skole, et al., 2021; Mbow et al. 2014; Quandt et al. 2023, Dmuchowski et al. 2024).

However, quantifying TOF's contribution to carbon stocks at the landscape or regional scale remains a challenge. This is
 particularly important as tropical tree-based ecosystems are experiencing reduced carbon stocks and increased carbon
 emissions due to the direct effect of accelerated land use and cover change (LUCC) and indirect effects from deforestation-
 induced climate change (Noble et al. 2000; DeFries et al 2002; Houghton 2007; Houghton and Castanho 2023; Friedlingstein
 60 et al., 2026). Precise and accurate quantification of emissions has been challenging due to the lack of large-scale monitoring
 systems (Huang et al. 2022). Moreover, policy actions to reduce emissions and increase carbon dioxide removals through
 forest and landscape restoration (FLR) have been limited by the lack of monitoring systems in TOF landscapes.



65 Ambitious regional initiatives such as the African Forest Landscape Restoration Initiative (AFR100) are being developed to
 restore degraded landscapes and mitigate desertification and climate change. AFR100 is a pan-African program with
 commitments from 34 countries to restore 100 million hectares of degraded land by 2030 through interventions that increase
 tree cover outside of forests and the regeneration of degraded natural forests. AFR100 supports Africa's contribution to the
 Bonn Challenge and the United Nations Decade on Ecosystem Restoration by promoting restoration in non-forest ecosystems
 70 and agroforestry systems (Parr et al., 2024). This urgency for monitoring of TOF is set against a backdrop of climate change
 mitigation policies that are mostly focused on closed forests, through frameworks such as REDD+ (Lewis et al. 2019). Given
 the increasing importance of TOF landscapes, which are found mostly in agricultural areas, there are sound reasons to raise
 the priority of the Agriculture, Forestry, and Other Land Use (AFOLU) sector rather than continue to focus on closed-canopy
 forests, thereby aligning the monitoring framework with the realities of African landscapes (Anyanwu et al. 2023; Frank et al.,
 75 2024; Mugabowindekwe et al., 2024).

Forest carbon monitoring systems, such as those designed for REDD+, rely heavily on national forest inventories that rarely
 account for isolated, dispersed, or agroforestry trees occurring in TOF landscapes (Mbow et al., 2014; Terasaki Hart et al 2026;
 Rosenstock et al. 2019; Minang et al., 2014; Mugabowindekwe et al., 2024). Additionally, given the spatial complexity
 80 inherent in the non-uniformity of tree distribution and size classes in TOF landscapes, it is difficult to effectively scale sparse
 plot-based sample frames to include dispersed small clusters of trees, especially for national-level assessments (Minarsch et
 al., 2025; Schnell et al. 2015). This small area estimation problem in TOF landscapes requires a different approach than limited
 sample-based inventories (Thapa et al 2023; Minarsch et al. 2025). Similarly, plot-based sample frames for closed forests have
 statistical assumptions and spatial logic that were designed for continuous forest stands, not fragmented tree distributions
 85 (Corona et al. 2011; Beckschafer et al. 2018; Brandt et al. 2020). For instance, stratification and the general uniformity in
 forest cover allows small biomass plots to be extrapolated on the basis of biomass density, while biomass of TOF landscapes
 is more accurately measured at the individual tree level through tree census methods (Graves et al. 2018). Further, inventory-
 based biomass estimation from small plots is often labor-intensive, costly, and impractical in remote areas (Brown, 1997). As
 a result, the contribution in TOF to national carbon assessments remains systematically underestimated, especially in countries
 90 where agroecological landscapes dominate (Kuyah et al., 2016; Skole et al., 2021).

Satellite remote sensing holds promise for large scale FLR-focused carbon monitoring in TOF landscapes because it can map
 individual tree crowns in a census approach across large areas (Brandt et al., 2020; Malkoç et al., 2021), thereby overcoming
 shortcomings that would arise using standard small sample plot-based forest inventories. Moreover, with the increasing
 95 availability of very high-resolution (VHR) satellite remote sensing (sub-1 m spatial resolution), opportunities now exist for
 measuring carbon stocks at the individual tree level using proper allometric scaling models (Skole et al., 2024; Tucker et al.,
 2023). There is thus an emerging need for ground calibration data to enable accurate allometry for remote sensing mapping of
 individual trees in TOF landscapes. Remote sensing methods that can derive parameters such as tree crown area (CA), stem



diameter at 1.3 m height (DBH), and tree height (H) have gained traction as a means to estimate whole tree biomass
100 complementary with standard mensuration methods. Advances in very high-resolution satellite imagery and machine learning
algorithms now enable individual tree crown detection and crown polygon delineation to get accurate crown dimensions for
use in allometric scaling models (Skole et al., 2021b; Skole et al. 2021).

Groundbreaking studies have demonstrated the feasibility of delineating billions of trees in drylands and agroforestry mosaics
105 using remote sensing-based crown metrics, (Brandt et al. 2020; Tucker et al., 2023). Although very high spatial resolution
satellite remote sensing can measure related allometric parameters such as crown dimensions, with some studies extending
crown allometric scaling to carbon stock estimation (Kuyah et al. 2012, Tucker et al. 2023), the availability of allometric
scaling models based on these variables alone are far less common than those using DBH. A key challenge when using satellite
remote sensing image data for carbon estimation lies in its obvious inability to directly measure DBH, the most common and
110 reliable allometric predictor of aboveground biomass. To routinely obtain large area estimates of individual tree biomass using
existing allometric scaling models, it will be necessary to use models that derive DBH from crown dimension
(Mugabowindekwe et al. 2023, Skole et al. 2024). Several studies using ground data have explored correlations between crown
and stem dimensions at the individual tree level (Cano et al., 2019; Jucker et al., 2022; Kuyah et al., 2012) but standardized,
validated protocols for TOF-specific crown-to-stem transfer functions (DBH-CA calibration) are still lacking, as are ground
115 datasets of simultaneously measured tree stem diameters and crown dimensions, especially for TOF systems (Kuyah and
Rosenstock 2015).

In this study we present a major new dataset which contains several thousand individual tree allometric measurements in FLR
projects in four countries in West and East Africa. This dataset represents a unique large-scale collection of simultaneous field
120 measurements of crown and stem dimensions, which can be used to calibrate DBH transfer models from CA measurements
and, subsequently, biomass allometric scaling models based on primary measurements from very high-resolution satellite-
based tree crown dimensions, in addition to other allometric scaling applications. We also provide a rigorous field protocol for
collection of these kinds of data. Using data from 24 AFR100 project sites across Ghana, Kenya, Rwanda, and Malawi, we
demonstrate the application of the dataset to very high-resolution remote sensing crown measures for estimating individual
125 tree carbon. The resulting models are intended to support large-scale carbon stock estimation in TOF systems and enhance
monitoring, reporting, and verification (MRV) for landscape restoration programs across the continent.



2 METHODS

2.1 Landscapes for Data Collection

Allometric tree data were collected across sampling sites in 24 landscapes associated with projects involved in the African
 130 Forest Landscape Restoration Initiative (AFR100), representing a variety of TOF landscapes across Ghana, Kenya, Rwanda,
 and Malawi (Fig. 1, Table 1, Table S1). Each project aligns with a national contribution to its country's participation in the
 AFR100 restoration commitment by 2030. Ghana aims to restore 2 million hectares, Kenya plans to restore 5.1 million
 hectares, Rwanda is committed to 2 million hectares, and Malawi plans to restore 4.5 million hectares (Mansourian and
 Berrahmouni 2021; Stanturf et al. 2019). The selected sites span a wide range of ecological zones and include different FLR
 135 interventions, such as reforestation, agroforestry, assisted natural regeneration, and sustainable land management practices

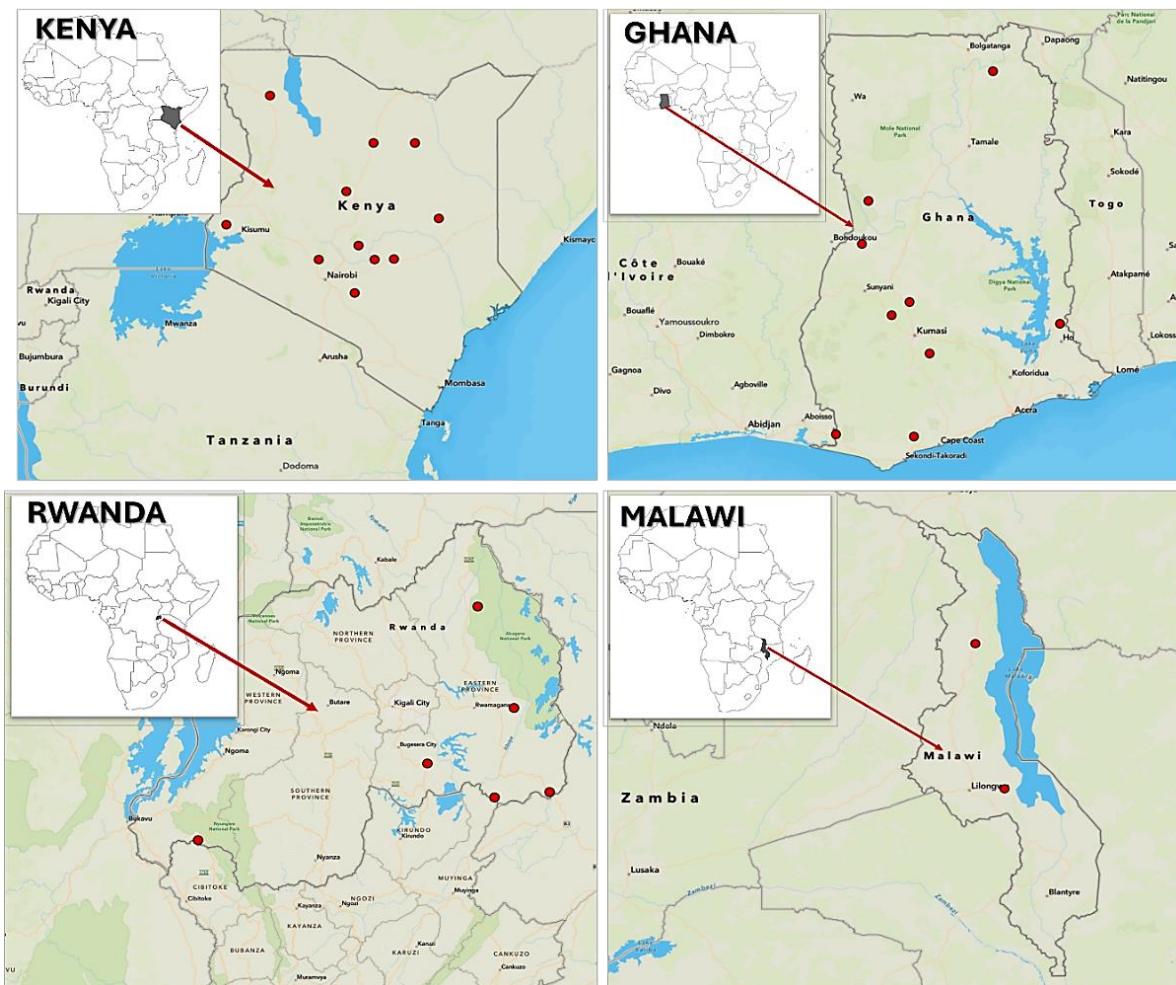


Figure 1: Study area map. The map illustrates the geographic distribution of sampling landscapes across all 24 sites in a) Kenya, b) Ghana, c) Rwanda, and d) Malawi in Africa. Basemap sources: ESRI, TomTom, Garmin, FAO, NOAA, USGS, OpenStreetMap contributors, and the GIS User Community | Powered by ESRI.



The dataset compiled represents a cross section of trees measured in 1 ha plots, all in landscapes dominated by trees outside
 140 of forests. The sampled landscapes are associated with projects supported by the *TerraFund for AFR100* project, a new FLR
 investment program led by the World Resources Institute and funded by the Bezos Earth Fund. Each sample landscape is
 assigned a unique identifier for easy classification and communication (Table 1). Table S1 in the Supplemental Materials
 provides a plot-level land-cover classification based on the FAO Land Cover Classification System (LCCS) (Di Gregorio &
 Leonardi, 2016).

145 Eleven (11) sample landscapes were located in Kenya. These FLR sites focus on restoring ecosystems that have been impacted
 by deforestation, overgrazing, and land degradation. Seven (7) sample landscapes were in Ghana across different ecosystem
 types. These landscapes feature cocoa-based agroforestry systems, forest rehabilitation, and dryland restoration efforts. They
 represent key restoration frontiers in both humid forest and savannah transition zones, where agricultural expansion, illegal
 150 logging, and charcoal production have reduced tree cover. In Rwanda, there were four (4) sample landscapes. The sites are
 mainly in steep terrain with intensive farming, where erosion causes serious land degradation following tree cover loss. These
 areas exemplify Rwanda's national Forest Landscape Restoration (FLR) policy, which integrates agroforestry, buffer-zone
 rehabilitation, and watershed protection. There were two (2) landscapes in Malawi. These areas face high pressures from
 deforestation, erosion, and overgrazing, with communities relying on smallholder farming and charcoal production.
 155 Incorporating TOF monitoring here supports Malawi's national restoration strategies and community-based forest
 management goals.

Table 1: Summary information on sample landscapes. The unique ID assigned to each project's sampling site, and the country where the sample was collected.

Sample Landscape and Associated Project Name	Landscape ID	Country
Sadhana Forest Kenya	SADHANA	Kenya
Drylands Natural Resources Centre	DNRC	Kenya
Fairtree, Kenya	FAIRTREE	Kenya
Kijabe Forest Trust	KFT	Kenya
Acacia Export Processing Zone	KITUI EPZ	Kenya
Acacia Export Processing Zone	TURKANA EPZ	Kenya
Mount Kenya Environmental Conservation	MKEC	Kenya
Kijabe Environment Volunteers	KEV	Kenya
Institute For Law and Environmental Governance	ILEG	Kenya
International Tree Foundation	ITF	Kenya
Development Response Kenya	DREK	Kenya
Institute of Green Growth Solutions	IGGS	Ghana
Institute of Nature and Environment Conservation	INEC	Ghana
Sustainable Agroforestry Initiative	SAFI	Ghana



Goshen Global Vision	GOSH	Ghana
CERATH Development Organization	CERATH	Ghana
PADO Investments Limited	PADO	Ghana
Herp-Ghana	HERP	Ghana
ARCOS	ACROS	Rwanda
Birdlife International	BIRDLIFE	Rwanda
World Agroforestry Center	ICRAF	Rwanda
Rwanda Environmental Conservation Organization	RECOR	Rwanda
Wells for Zoë	WFZ	Malawi
Wildlife Action Group	WAG	Malawi

160 2.2 Methodological Framework

A robust literature supports the use of stem-based allometric scaling models for estimating biomass and carbon stocks in forests; however, comparable approaches remain far less developed for TOF systems (Schnell et al 2015). The dataset presented here was prepared to support remote sensing-based individual tree allometric scaling. Emerging interest in carbon sequestration activities with landscape restoration and agroforestry expansion have created a growing demand for measurement platforms capable of quantifying tree resources across TOF landscapes. The spatial heterogeneity and discontinuous distribution of trees in such systems limit the effectiveness of traditional plot-based sampling strategies and motivate the development of census-focused landscape-scale measurement approaches. VHR satellite data captures individual tree allometric parameters for landscape-wide census of all trees without sampling. It also more closely aligns with standard individual tree allometric scaling, which would be an improvement over current medium and coarse resolution remote sensing mapping of biomass, which uses aggregate area-based methods (Tian et al. 2023; Papucci et al. 2026; Bossy et al. 2025; Baccini et al 2008; Dubayah et al. 2022; Saatchi et al. 2011; Mitchard et al. 2014).

Recent advances in VHR satellite remote sensing (0.35–1.0 m resolution) combined with deep learning tree detection algorithms are enabling direct mapping of individual trees across large areas (Brandt et al., 2020; Skole et al., 2021; Tucker et al., 2023; Skole et al. 2024). However, VHR-based measurement systems rely primarily on remotely observable crown attributes rather than stem diameter allometric scaling, which is traditionally the main parameter used to estimate forest biomass. One aim of collecting these data is to calibrate a stem diameter transfer function from remote sensing measurements.

VHR satellite image data are analyzed using deep learning algorithms for precise mapping of Crown Projected Area (CPA) at the individual-tree level (Skole et al 2024). CPA represents the vertical projection of the tree crown onto the ground surface and is commonly expressed in m². We use the notation of CA in our model formulations to represent CPA measured in the field. We note that some allometric scaling models may represent crown areas as the crown surface area, particularly those related to resource allocation, leaf area indexing and metabolic scaling. In this study CA = CPA. Although CPA is frequently



measured in field inventories, such measurements are often subject to uncertainty due to irregular crown geometry and observer variability. Also, CPA measured on the ground frequently uses a sample of only one or two canopy axes. Consequently, CPA-based biomass allometric models remain less widely used than those based on diameter at breast height (DBH).

Nonetheless, multiple studies have demonstrated strong empirical relationships between DBH and CPA (Asigbaase et al., 2023; Kuyah et al., 2012; Mugabowindekwe et al., 2023; Mugo et al., 2011; Skole et al., 2021b), and conceivably VHR remote sensing could provide more precise estimates of CPA than measurements on the ground. Other studies have reported relationships between DBH and crown diameter, CD (Loubota Panzou et al, 2021, Jucker et al 2022). Building on these relationships, the present study supports a hybrid crown-to-stem transfer-function methodology in which remotely sensed tree CPA is first used to estimate the associated tree DBH through empirically calibrated regression models. The estimated tree DBH is then applied within established stem-based biomass allometric equations (Skole et al. 2021b, Skole et al. 2024). This approach leverages the scalability and spatial completeness of crown remote sensing observations while maintaining compatibility with internationally recognized biomass estimation allometry derived from stem-based measurements.

2.3 Field Data Collection Protocol

A standard protocol for field sampling and data collection was created and tested. This protocol is found in Supplemental Materials. Data collection was conducted in sample landscapes associated with the FLR projects, focusing on areas with high TOF densities of isolated, non-overlapping tree crowns. To identify geographically a set of different areas for collection of tree data, we identified and delineated sample landscapes using satellite data. A sample landscape is defined as uniform ecological area encompassing most of the intervention sites in each of the 24 individual projects. This sample landscape is an area not less than 5×10^3 ha and not more than 1×10^5 ha. Individual tree allometric variables were measured concurrently on the ground and in satellite imagery at randomly located 1 ha circular plots mapped on VHR satellite imagery across each sample landscape.

These sample plots were mapped onto VHR imagery and prepared for field crews in advance of their deployment. For this, a machine learning model was used to delineate tree crown polygons at sample plots on concurrent VHR satellite data (0.35 m) based on the method described in Skole et al. (2024). Digital tree crown polygons from a machine learning model output layers were georeferenced with precise coordinates (Skole et al., 2024). Each mapped crown was subsequently matched to its corresponding tree in the field and labeled. VHR data were acquired from NASA's Commercial Smallsat Data Acquisition program (Neigh et al. 2013), mostly WorldView-2 and WorldView-3 satellite platforms.

The geographic coordinates of plot centers and all individual trees were recorded in the field using handheld GPS units. Coordinates were documented in decimal degrees (DD). The DD coordinate format was preferred over Universal Transverse Mercator (UTM) or degrees–minutes–seconds (DMS) formats because the DD system, referenced to WGS84, minimizes



geospatial inconsistencies and transformation errors during dataset integration with VHR imagery, thereby reducing potential spatial misalignments in the analytical and mapping system (Alcaras et al., 2020).

220 All individual trees with non-overlapping crowns inside the plot with a DBH ≥ 5 cm were measured. DBH was measured for all trees at 1.3 m above ground level using a calibrated diameter tape. For trees on slopes, measurements were taken from the uphill side. In cases of leaning stems, the diameter tape was positioned perpendicular to the stem's longitudinal axis. Forked trees were measured separately if branching occurred below 1.3 m; if branching occurred above this height, only the main stem was measured, unless a structural irregularity was present at the standard height.

225 Individual tree crown dimensions were measured in the field with a 50 m tape. To estimate the ellipsoidal CPA, two measurements were taken for each tree: the crown longest axis diameter and the crown diameter perpendicular to the long axis diameter. The crown area was computed using the elliptical formula (Kuyah et al. 2016). No minimum crown diameter threshold was set for measurements. Each of these trees also had an associated CPA polygon from the remote sensing analysis.

230 For each field measured tree, the attributes listed in Table 2 were recorded and cross-referenced with their spatial locations to their respective tree on the VHR map. These attributes were retained and transferred into the full database, along with additional attributes computed from field data, as shown in Table 2, which provides the data descriptor record of attribute fields found in the accessible dataset.

235 **Table 2: Description of recorded attributes. The table shows the attribute ID, the unit of measurement, the data type, description, and example of individual field measurement in the dataset.**

Attribute ID	Unit	Data Type	Description	Example
Site ID	—	String	Unique identifier for the sampling site	HERP
Plot ID	—	Integer	Sequential identifier for each 1-ha sampling plot within a site	3
Plot center (Lat)	DD	Float	Geographic latitude of the 1-ha plot centroid, referenced in WGS84	6.8618
Plot center (Lon)	DD	Float	Geographic longitude of the 1-ha plot centroid, referenced in WGS84	0.4493
Date	—	Date	Date of data collection, recorded in MM-DD -YYYY format	2/15/2024
Tree ID-matched on imagery.	—	Integer	Unique tree identifier, matched to remote sensing imagery	14
Tree location from GPS (Lat)	DD	Float	Geographic latitude of the individual tree location, referenced in WGS84	10.47772
Tree location from GPS (Long)	DD	Float	Geographic longitude of the individual tree location, referenced in WGS84	-0.50222
Tree Genus and species	—	String	Scientific name of the sampled tree	<i>Mangifera indica</i>
DBH	cm	Float	Trunk diameter at 1.3 m above ground	34.2
Crown_L	m	Float	Longest horizontal crown dimension	6.9
Crown_P	m	Float	Crown diameter perpendicular to the major axis	6.1



Mean crown diameter	m	Float	Mean of the long and perpendicular crown diameters	6.5
Tree height	m	Float	Vertical distance from the ground to the top of the tree	13.5

2.4 Testing and Refinement of the Data Collection Protocol

The data collection protocol was documented as a Field Guide (see Supplemental Materials) which includes plot allocation methods, data collection procedures, measurement procedures, and standard operating procedures. Field deployments occurred over the course of several campaigns. The first campaign was to AFR100 sample landscapes in Kenya and Ghana in June, 2023. The purpose of this campaign was field-testing the data collection protocol to: 1) determine refinements in procedures, 2) test consistency in data, and 3) identify potential logistical modifications. All equipment, including GPS units, diameter tapes, and clinometers, were calibrated before use to maintain measurement precision and reduce systematic error.

This initial implementation of the field protocol was conducted at five distinct sample landscapes (see Table 1 and Table S1): three in Kenya—DNRC, SADHANA, and FAIRTREE—and two in Ghana—INEC and IGGS. During this preliminary assessment phase, field crown diameter measurements were not included, following a method established in Skole et al. (2021b), since we assumed that CPA could be derived from remotely sensed polygons rather than field measurements. However, limitations, particularly the asynchrony between field data collection and remote sensing imagery acquisition, limited the concurrence of DBH and CPA measurements. Moreover, we decided to adhere to normal forest tree inventory practice, to closely align our method and data with standard datasets of this. This adjustment facilitated the integration of temporally concurrent field and remote-sensing data, thereby improving the reliability of structural estimates. Thus, all final data in this collection include field-measured crowns and users can add their own remote sensing CPA. Measurements were conducted in each 1-ha plot using a standardized procedure which starts at the first tree located north of the plot centroid, and all subsequent trees then sampled sequentially in a clockwise direction from due north (Fig. 2).

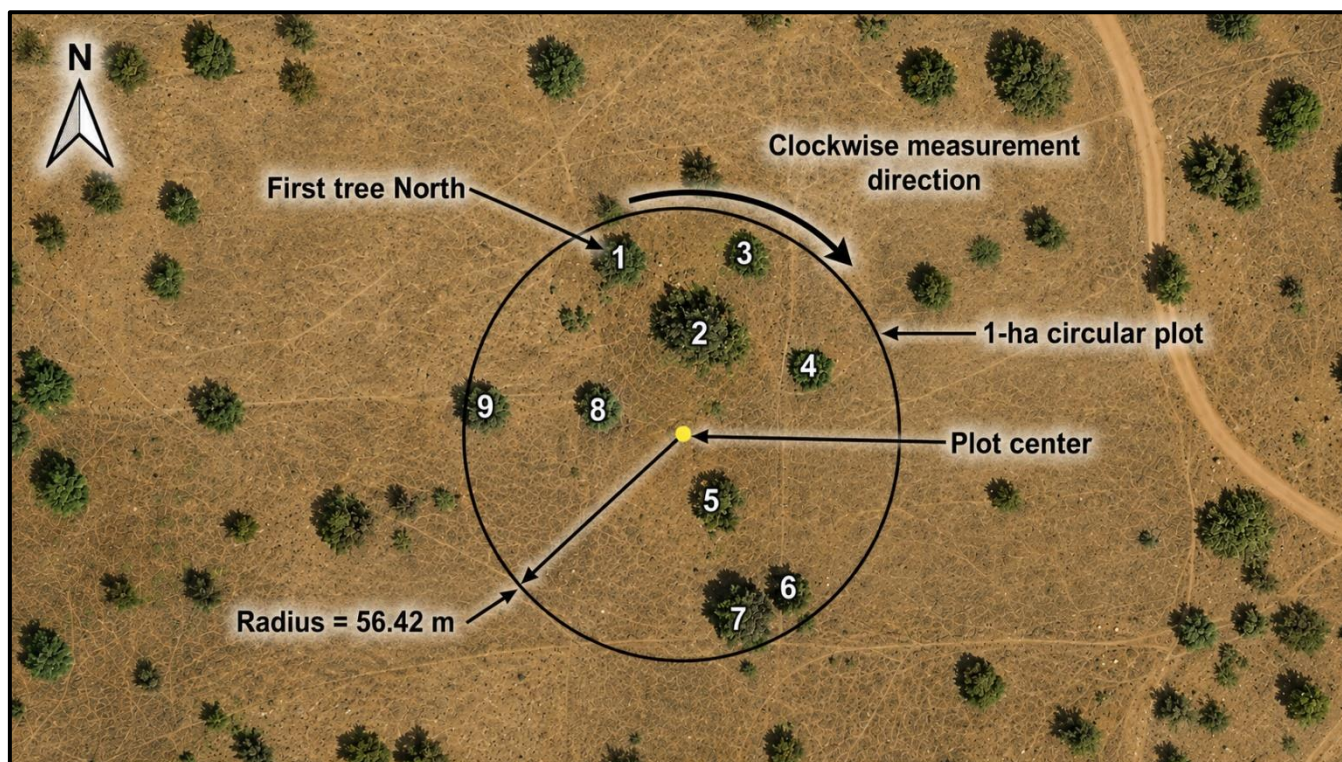


Figure 2: Illustration of the field measurement protocol. The figure depicts the 1-ha circular plot with its centroid, the first tree located to the north, and the clockwise measurement procedure. Imagery credit: Esri, Vantor, Earthstar Geographics, and the GIS User Community, Sources: Esri, TomTom, Garmin, FAO, NOAA, USGS, (c) OpenStreetMap contributors, and the GIS User Community | Powered by ESRI.

3 DATASET DESCRIPTION

3.1 Quality Control and Data Cleaning

Data processing, compilation and quality checks were first implemented following Chicco et al. (2022). Following this basic cleaning, all entries and measurement records were screened for completeness; individual records lacking key attributes, such as records without measurements of both DBH and crown dimensions, were excluded. We also ensured spatial consistency by removing individual tree entries with missing or incorrect geographic coordinates, as these could not be reliably integrated with remote sensing data. Using field notes and comments in the data file, we were able to check for structural anomalies; trees that were pruned, coppiced, had forked stems, dense or overlapping crowns, or experienced pronounced crown deformation from management practices (e.g., pollarding) were systematically excluded from this database and will be put into another database in the future. This exclusion criterion minimized bias and ensured that the dataset accurately reflects the typical tree form and structure required for allometric modeling and environmental analysis. After applying these quality-control filters,



the final dataset was reduced from an original collection of 5,218 to 3,493 trees. This rigorous curation ensured a consistent, high-quality dataset suitable for robust model calibration and validation.

3.2 Description

275 Table 2 presents the data dictionary for each record in the dataset. The dataset is grouped according to individual sample
landscape and country (Table S1 and Table 3). The distributions of key tree attribute parameters, including DBH, mean crown
diameter, and height, are illustrated using violin plots (Fig. 3). Descriptive statistics are used to summarize the main tree
attribute parameters in the dataset (Table 4). For each country, the mean, standard deviation, and range were calculated for
DBH, crown long diameter, crown perpendicular diameter, and tree height. These statistics provide an overview of the central
280 tendency and variation of tree attributes across the study regions.

Given the extensive diversity of tree species across sampling locations and countries, we present the distribution of the 10 most
abundant species per country using a species-abundance matrix (Fig. 4). The graph summarizes the abundance patterns of the
most frequently recorded tree species and allows direct comparison of shared and country-specific dominant taxa. Bubble size
285 represents the absolute abundance of each species, whereas bubble color represents the relative abundance. Relative abundance
was calculated as the abundance of each species divided by the total number of sampled individuals within each country. The
total number of sampled individuals by genus and species was 953 for Kenya, 895 for Ghana, 521 for Rwanda, and 527 for
Malawi.

290 **Table 3: Summary of dataset records. The table lists the countries where data was collected, the dates of collection, the number of
sites, 1-ha plots, individual trees, and different species sampled.**

Country	Data Date	Collection	Number of sites	of	Number of 1-ha plots sampled	Individual trees record	Number of different species sampled
Kenya	Oct-23, Mar-24		8		89	1,384	170
Ghana	Feb-24		5		31	895	69
Rwanda	Feb-24		4		46	526	40
Malawi	Dec-23		2		24	688	91

Table 4: Descriptive statistics of DBH, crown diameters, and tree height in the dataset for Kenya, Ghana, Rwanda, and Malawi. The values show the mean, standard deviation, and range of observations.

Country	DBH (cm)		Crown_L (m)		Crown_P (m)		Height (m)	
	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range
Kenya	24.7 ± 14.2	3.5–162.0	8.1 ± 3.7	1.5–27.1	7.2 ± 3.4	1.2–25.3	8.8 ± 3.6	1.5–26.7.7
Ghana	36.0 ± 12.5	8.5–93.7	8.9 ± 3.0	2.6–22.0	8.0 ± 2.6	2.0–22.0	10.2 ± 2.9	5.0–34.7
Rwanda	22.5 ± 9.2	5.3–72.0	6.1 ± 2.7	1.2–16.8	5.6 ± 2.6	1.0–15.5	9.8 ± 4.0	2.7–24.9
Malawi	29.5 ± 18.7	3.6–121.4	7.6 ± 3.8	1.5–29.5	8.0 ± 2.8	1.5–26.0	8.8–5.4	2.2–36.3

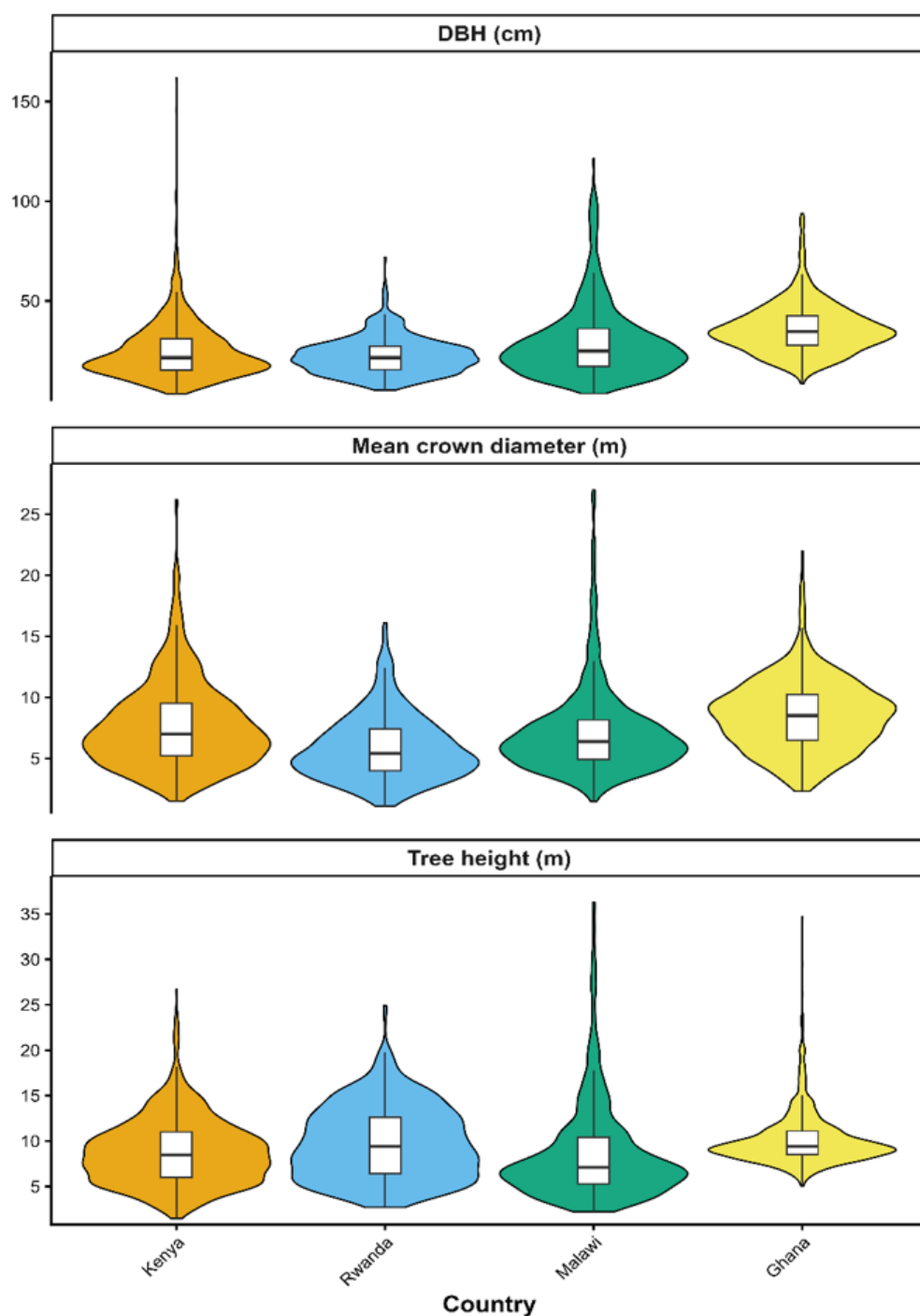
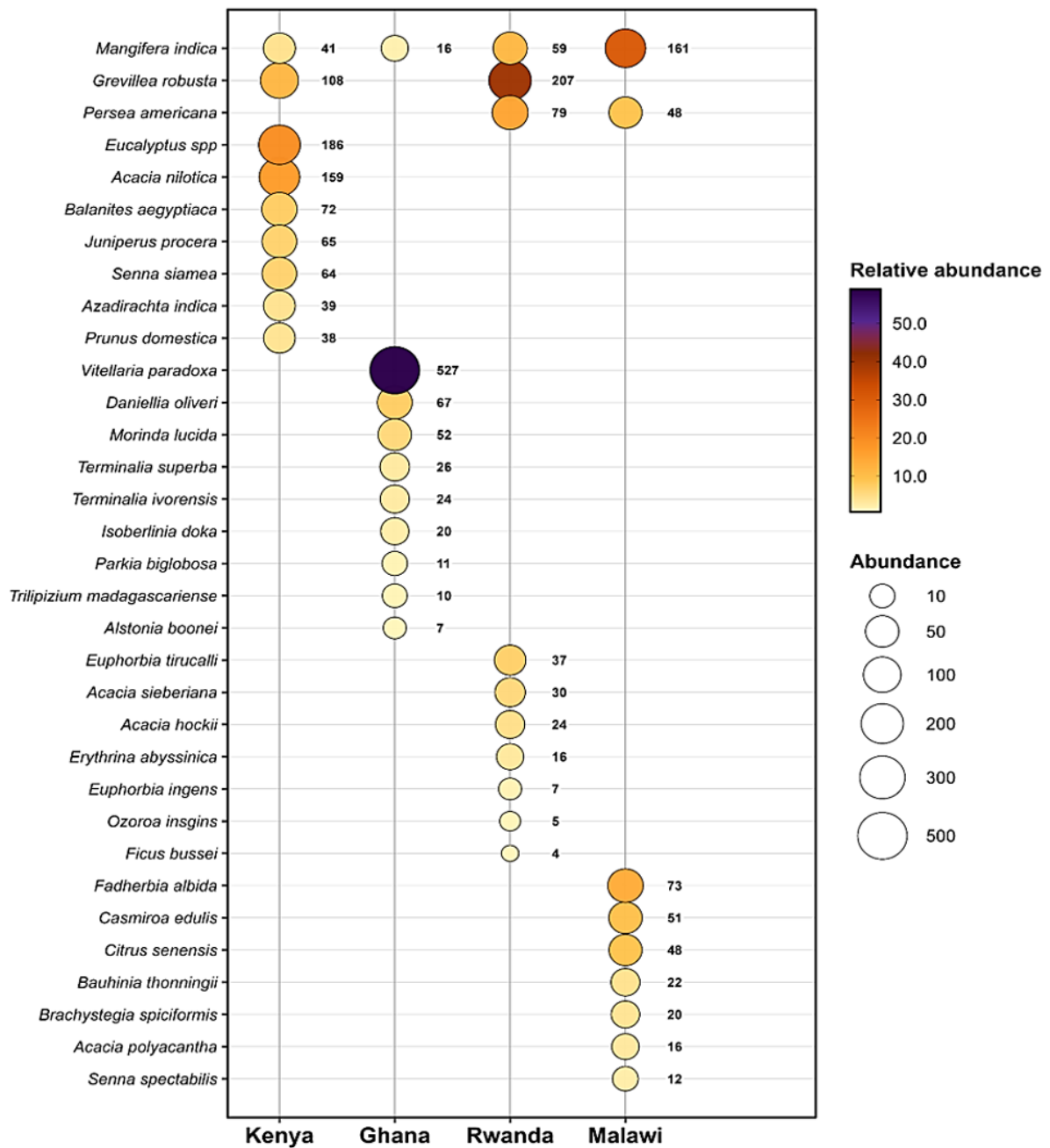


Figure 3: Probability density distribution of diameter at breast height (DBH), mean crown diameter, and tree height by country in the dataset. The width of each violin indicates the relative frequency of observations based on kernel density estimation, while the embedded boxplots display the median and interquartile range.



300 **Figure 4.** Comparative species-abundance matrix of dominant tree taxa across Kenya, Ghana, Rwanda, and Malawi. The bubble size and label show the absolute abundance of individuals of the species, and bubble color represents the relative abundance of individuals, and percentage of total number of sampled individuals within each country.



4 TECHNICAL EVALUATION

4.1 Statistical Metrics and Diameter-Crown Function

A pairwise scatter plot matrix was created to present the distributions and relationships among DBH, crown projected area (designated here as Crown area, CA), and tree height (Fig. 5). The diagonals show histograms of frequency distributions for the variable, and the cells above the histograms show the Pearson correlation. The Pearson correlation (r) quantifies the strength and direction of these linear relationships. An important core relationship for which these data were initially collected is between crown dimensions as the independent variable and stem DBH as the dependent variable. A scatter plot was used to evaluate the relationship between DBH and CA, as well as crown diameter, across each country's dataset, with a 95% bootstrap confidence interval, highlighting the dataset's robustness in capturing structural relationships among these attributes (Fig. 6).

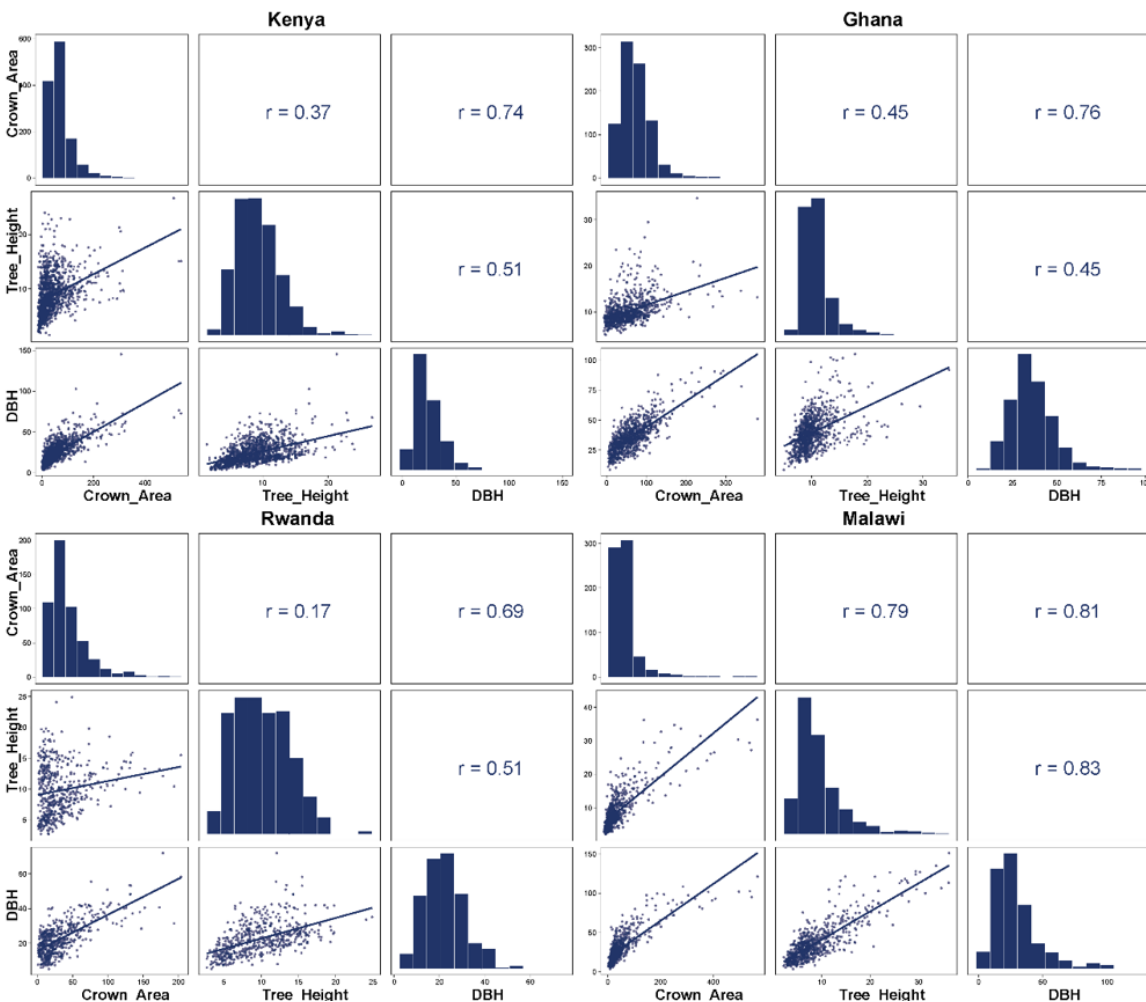
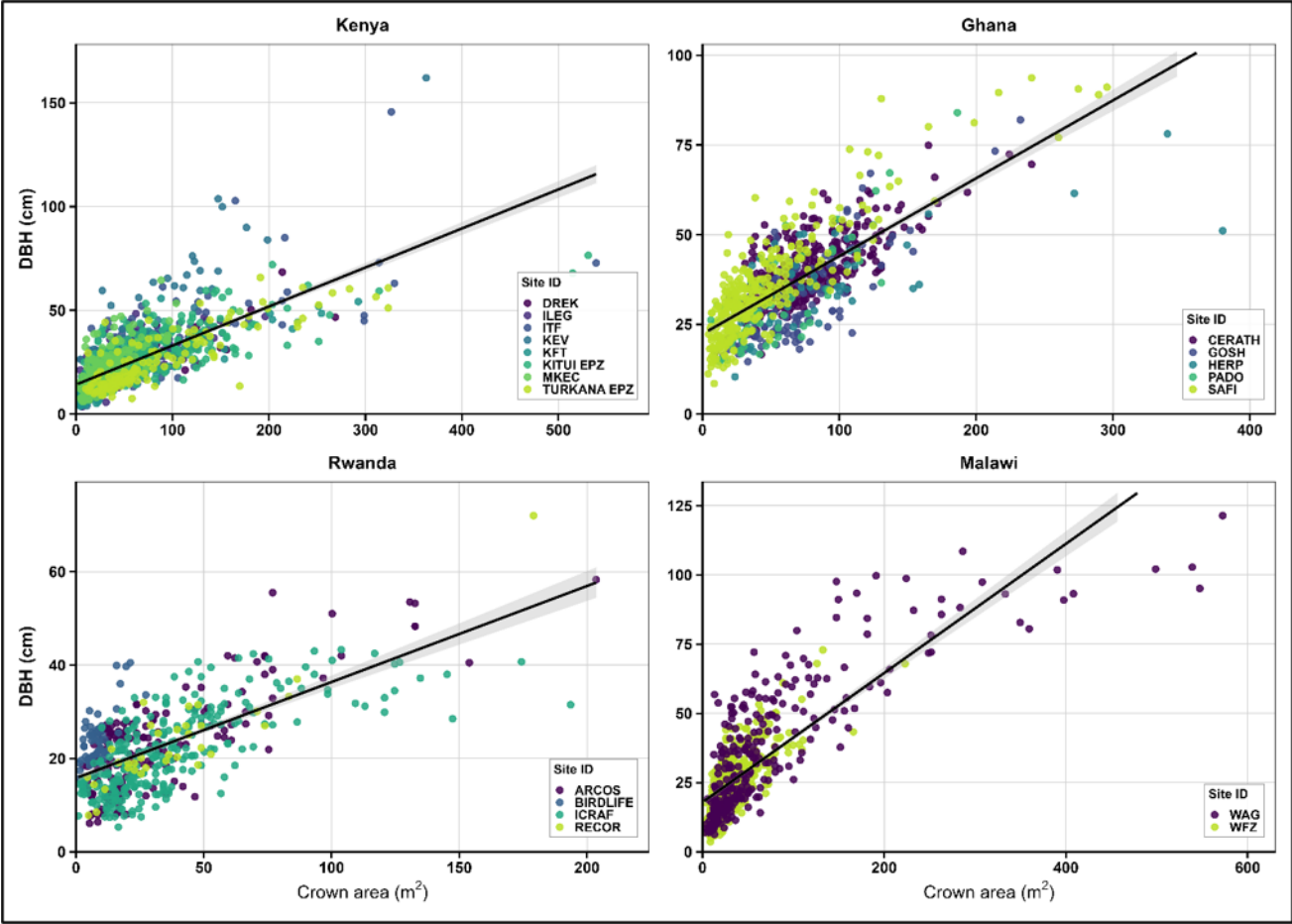


Figure 5: Pairwise scatterplot matrix of DBH, crown area and tree height. Diagonal panels show the univariate distributions of each variable, lower panels show bivariate scatter plots for all variable pairs, and upper panels show the Pearson correlation coefficients (r) for Kenya, Ghana, Rwanda, and Malawi



A



320

325



330

B

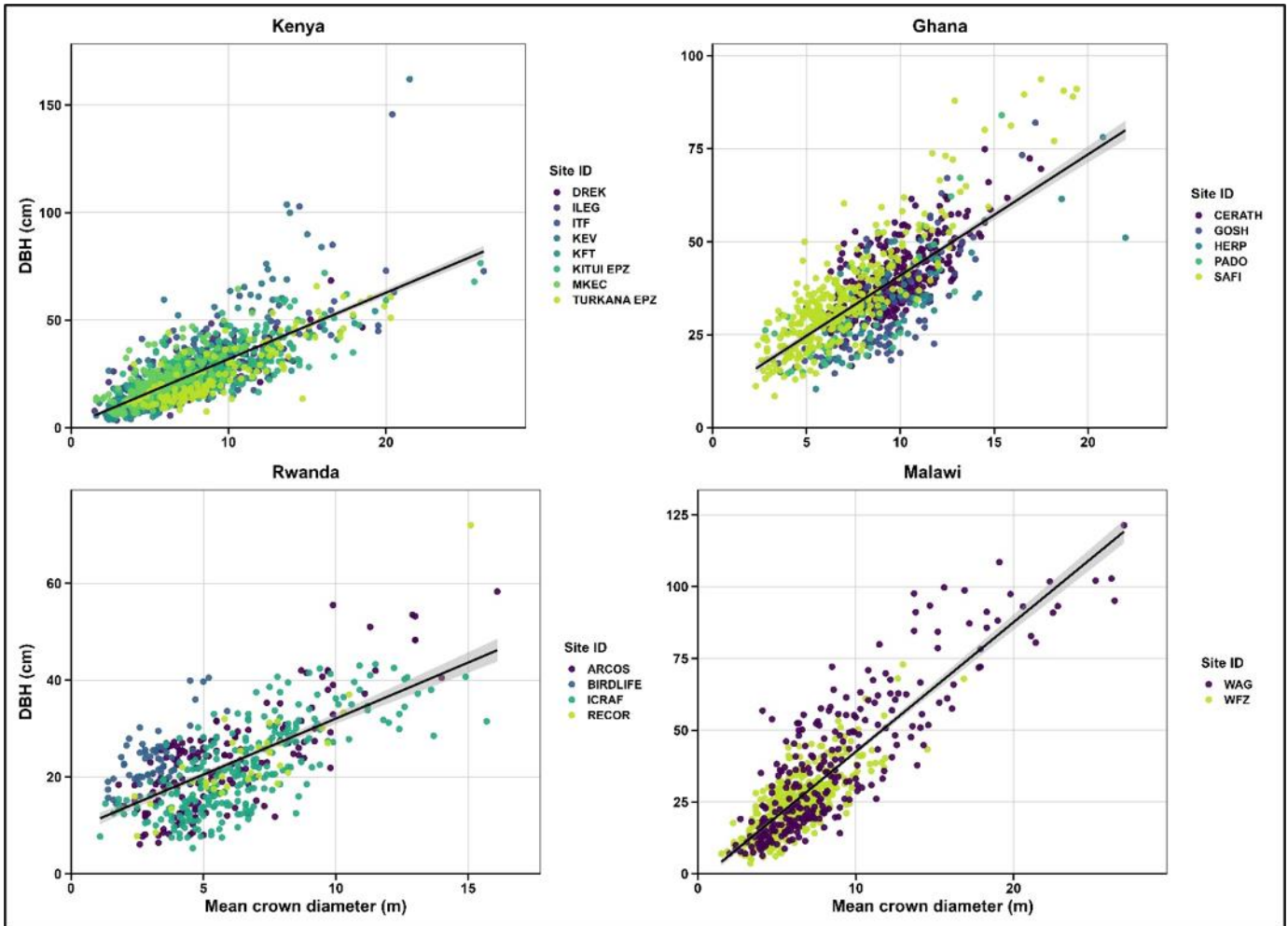


Figure 6: A scatter plot showing the relationship between measured DBH and crown area, CA (A) and crown diameter (B) across all sites in Kenya, Ghana, Rwanda, and Malawi. The different colour points represent site IDs, the black line indicates the fitted linear regression, and the grey shading displays the 95% confidence interval.

335 Based on the nonlinear relationship between DBH and crown area observed at the sample landscape level and in most
 allometric studies, we adopt a power-function to demonstrate the allometric model. We refer to this model formulation as
 DBH-CA. We fit a power function to the data collected at each landscape and country, and to the combined dataset. The power-
 law model, widely recognized for capturing allometric relationships in tree structure and biomass estimation (Kuyah et al.,
 2012a; Kuyah & Rosenstock, 2015; Sileshi et al., 2023), was formulated as Eq. (1):

340

$$DBH = a (CA)^b \quad (1)$$



Where a is the scaling coefficient, and b is the allometric exponent describing the rate of change. To fit the coefficients to this DBH-CA model, we linearize the relationship where both variables were transformed as Eq. (2):

$$\ln(DBH) = \ln(a) + b \cdot \ln(CA) \quad (2)$$

The transformation improves model linearity and homoscedasticity (Mahmood et al., 2019), enabling parameter estimation via ordinary least squares (OLS) regression using the `lm()` linear fitting function in R CRAN programming language package. The fitted parameters were subsequently back-transformed to the original scale to obtain the final power-function form. To correct for bias induced by logarithmic transformation, a correction factor (CF) was applied following (Baskerville, 1972), Eq. (3).

$$CF = \exp\left(\frac{SE^2}{2}\right) \quad (3)$$

Where SE is the standard error of the estimated regression. The back-transformed intercept was multiplied by CF to yield unbiased predictions (Chave et al., 2005).

Model performance was evaluated using multiple statistical metrics: the coefficient of determination (R^2), root mean square error (RMSE), relative RMSE (%RMSE), and absolute prediction bias (Bias) and its percentage (%Bias) (Kuyah et al., 2012; Fischer et al., 2020; Luck et al., 2020). Graphical prediction evaluation was conducted using scatterplots of observed DBH versus predicted DBH values, with a 1:1 reference line representing perfect prediction (Fig. 7, S1-S4).

The R^2 values ranged from 0.50 to 0.85 at the Kenya sites, 0.43 to 0.74 at the Ghana sites, 0.60 to 0.74 at the Malawi sites, and 0.44 to 0.83 at the Rwanda sites. Our model performance values reflect moderate to strong associations between DBH and CA, across all landscapes and are generally consistent with other reported DBH-CA relationships, such as Skole et al. (2021) and Kuyah et al. (2012), of 0.84 and 0.86, respectively. Our values here are somewhat lower than these but with more variance due to a higher number of agroforestry trees. So, for instance, Asigbaase et al. (2023) also reported R^2 values ranging from 0.27 to 0.87, demonstrating that tree species composition and structural differences due to management influence tree allometry across TOF landscapes. The large dataset from Loubota Panzou et al. (2021) reports R^2 values of 0.82 for trees in Kenya, 0.75 for trees in Ghana and 0.64 for trees in Malawi, with a higher mix of naturally occurring trees than our dataset. All models were highly significant ($p < 0.001$), with varying degrees of predictive accuracy, scaling coefficients, and overall model fit (Tables 5,6 and Table S3).

To evaluate the robustness of the allometric relationships across tree size classes, models were developed using the full diameter range as well as restricted subsets ($DBH \leq 60$ cm, ≤ 50 cm, and ≤ 40 cm) (Table S3). The primary analyses and



reported model coefficients are based on the full DBH range to ensure applicability across the entire structural gradient of trees within each site. However, additional models fitted to reduced DBH ranges are provided in the Supplementary Materials to assess potential scale-dependent effects on model performance.

Overall, model bias was small, with mean percentage error typically within $\pm 3\%$, and often within $\pm 1\text{--}2\%$ for the best-performing site–model combinations, indicating limited systematic over- or underestimation (Table 5). For many ecological applications, such bias levels are considered acceptable, as they imply that most of the uncertainty arises from random rather than systematic error (Aabeyir et al., 2020; Kuyah & Rosenstock, 2015). The RMSE% values range from 15.9 to 38.5%, reflecting differences in error magnitude relative to the mean DBH. In ecological modelling, RMSE should not be interpreted against a universal cutoff; instead, it should be assessed in relation to the scale of the response variable, model objective, and other diagnostic statistics (Chai & Draxler, 2014; Rykiel, 1996; Sileshi, 2014). These error levels are comparable to other evaluations of DBH–CA allometry, where RMSE% values around 40% have been reported (Skole et al., 2024; Mugabowindekwe et al., 2023), underscoring that relatively high uncertainty is not uncommon in structurally heterogeneous landscapes. In models with high RMSE%, improved performance may require additional predictors or species-specific equations.

Table 5: Site-specific allometric model performance across sample landscapes in Kenya, Ghana, Rwanda, and Malawi. DBH is modeled as a power function of DBH = $a(\text{CA})^b$. Models are assessed using R^2 , RMSE, %RMSE, Bias, and percent error (Bias%).

Site ID	Equation	n	P-value	R^2	RMSE	RMSE%	BIAS	BIAS%
DREK	DBH = $5.787 * \text{CA}^{0.376}$	32	< 0.001	0.50	7.89	26.32	-0.01	-0.03
ITF	DBH = $3.869 * \text{CA}^{0.498}$	157	< 0.001	0.58	11.72	38.54	0.10	0.32
KITUI EPZ	DBH = $6.075 * \text{CA}^{0.378}$	291	< 0.001	0.55	7.84	27.46	-0.04	-0.15
TURKANA EPZ	DBH = $2.654 * \text{CA}^{0.530}$	135	< 0.001	0.72	7.42	28.81	0.02	0.07
MKEC	DBH = $5.292 * \text{CA}^{0.434}$	204	< 0.001	0.55	6.89	31.05	-0.09	-0.41
ILEG	DBH = $4.354 * \text{CA}^{0.457}$	201	< 0.001	0.61	6.33	27.8	-0.25	-1.08
KEV	DBH = $2.424 * \text{CA}^{0.663}$	145	< 0.001	0.85	10.08	42.79	0.58	2.44
KFT	DBH = $4.633 * \text{CA}^{0.407}$	219	< 0.001	0.64	4.86	25.43	-0.14	-0.74
CERATH	DBH = $8.794 * \text{CA}^{0.354}$	333	< 0.001	0.46	6.39	15.87	0	-0.01
SAFI	DBH = $8.538 * \text{CA}^{0.390}$	325	< 0.001	0.74	7.33	21.32	0.03	0.08
HERP	DBH = $7.460 * \text{CA}^{0.366}$	58	< 0.001	0.51	8.15	22.55	-0.19	-0.51
GOSH	DBH = $5.163 * \text{CA}^{0.436}$	116	< 0.001	0.52	8.69	27.49	0.12	0.38
PADO	DBH = $8.736 * \text{CA}^{0.330}$	63	< 0.001	0.53	9.39	31.13	0.16	0.54
RECOR	DBH = $5.583 * \text{CA}^{0.411}$	35	< 0.001	0.71	5.92	24.96	0.08	0.34
ARCOS	DBH = $2.722 * \text{CA}^{0.511}$	165	< 0.001	0.83	4.26	18.44	-0.20	-0.86
BIRDLIFE	DBH = $15.799 * \text{CA}^{0.215}$	57	< 0.001	0.44	4.25	17.71	0.01	0.05
ICRAF	DBH = $6.780 * \text{CA}^{0.339}$	264	< 0.001	0.54	6.11	28.21	-0.05	-0.24
WAG	DBH = $3.833 * \text{CA}^{0.563}$	269	< 0.001	0.74	12.98	34.52	-0.69	-1.84
WFZ	DBH = $3.993 * \text{CA}^{0.533}$	419	< 0.001	0.60	6.89	28.26	-0.20	-0.84



395 Table 6. Country-wide allometric performance metrics for power function model of $DBH = a(CA)^b$

Country	<i>a</i>	<i>b</i>	<i>n</i>	R ²	MAE	RMSE	%RMSE	Bias	Bias%
Ghana	9.3285	0.339	895	0.55	6.36	8.39	23.30	0.01	0.02
Kenya	4.3136	0.4596	1384	0.58	6.20	9.24	37.35	-0.03	-0.11
Malawi	3.7609	0.5594	688	0.73	7.33	9.68	32.78	-0.25	-0.84
Rwanda	10.4729	0.242	521	0.39	5.81	7.22	32.05	-0.09	-0.38
All countries	5.6476	0.4304	3488	0.58	7.33	9.76	34.56	-0.12	-0.41

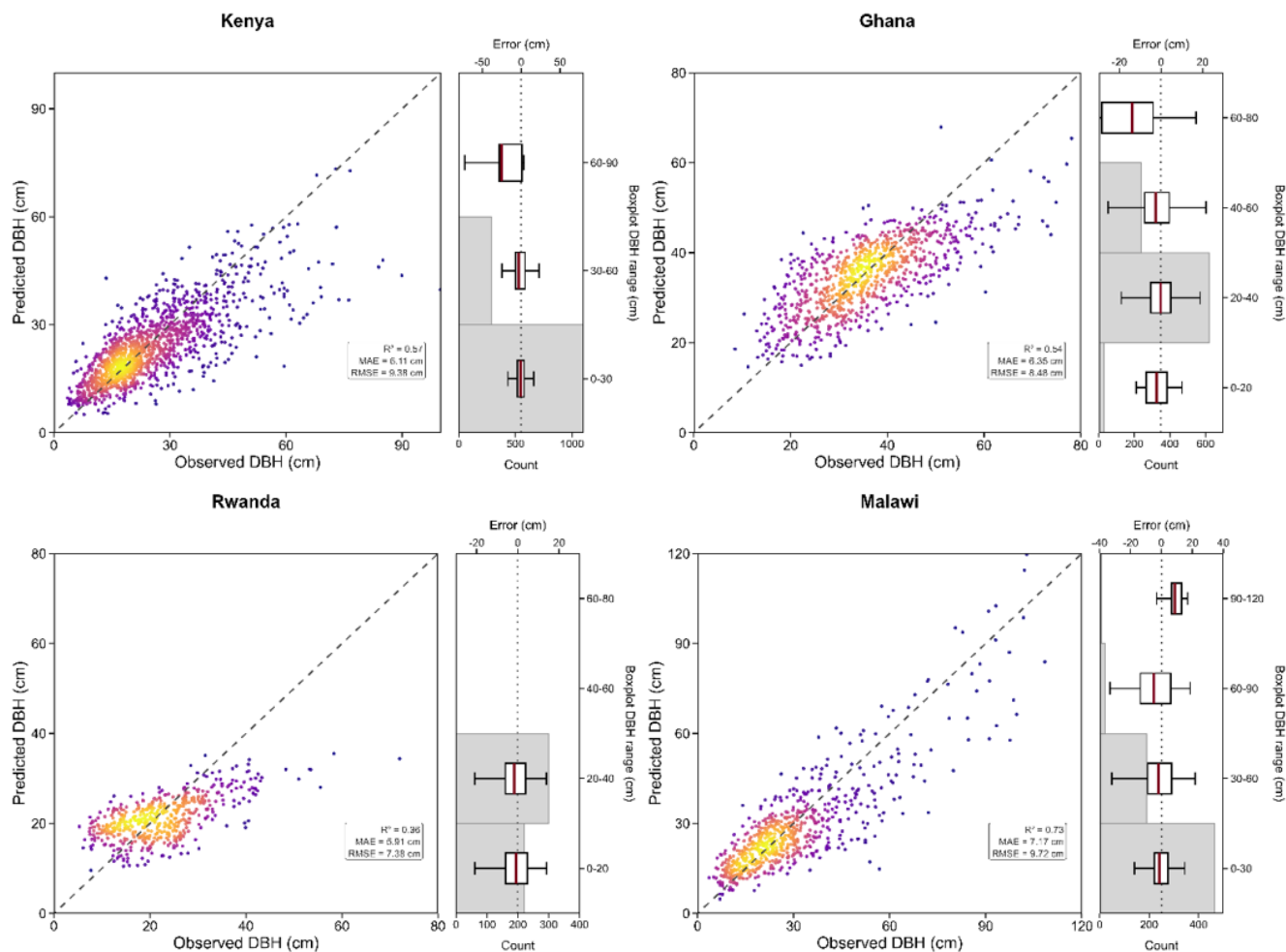


Figure 7. Performance of crown area-based DBH prediction models across Kenya, Ghana, Rwanda, and Malawi. Density scatter plots show the relationship between observed and predicted DBH for each country. The adjacent error panels summarize model error across predicted DBH ranges, and boxplots show the distribution of prediction errors for each range.

400



The fit and predictive accuracy of each allometric model were evaluated by comparing observed DBH versus predicted DBH values across all sample landscape sites (Fig. S1-S4). Across all sites, the plots are characterized by slopes less than 1 and positive intercepts. This pattern suggests an overestimation of smaller DBH values and an underestimation of larger ones. ILEG, SAFI, ARCOS, and WAG models displayed the closest alignment with the 1:1 line across the observed DBH range in the Kenya, Ghana, Rwanda, and Malawi sites, respectively. Also, a country-specific density scatter plot of the relationship between observed and predicted DBH, derived from crown area using power-function models, was created for each country dataset (Fig. 8). The error panels show how model error varies across predicted DBH ranges, the grey bars represent the frequency of observations within each DBH interval, and the boxplots summarize the spread of prediction errors.

Overall, prediction error decreased when models were fitted to lower DBH ranges, with improvements in RMSE, RMSE%, Bias, and Bias% (Table S3). This suggests that crown area predicts DBH more consistently for small- to medium-sized trees than for larger trees, where crown architecture, competition, and site-level structural heterogeneity may weaken the crown–stem relationship. However, R^2 values generally declined in the restricted models, likely because reducing the DBH range also reduced the variance available to explain. Thus, the reduced-range models improved prediction accuracy but captured less overall variation, whereas the full-range models provide broader applicability across the complete sampled tree-size distribution.

4.2 Hierarchical Synthesis of Crown-Stem Allometry Across Landscapes

The site-specific models demonstrate that the relationship between crown area (CA) and stem diameter (DBH) is consistently expressed across the sampled landscapes, but that the parameters of this relationship vary substantially among sites. Across the 19 site-specific models, the scaling exponent b ranges from 0.215 to 0.663, while model performance also varies with sample size, landscape structure, species composition, and other site-level characteristics (Table 5). Rather than treating this variation as statistical noise or assuming that a single pooled regression adequately represents all trees and landscapes, we interpret the distribution of site-level relationships as an important property of crown–stem allometry in heterogeneous TOF systems. The objective of the synthesis is therefore to estimate a generalized cross-landscape relationship while retaining information about the magnitude and structure of among-site variation.

A simple pooling of all individual-tree observations would implicitly give greatest influence to sites containing the largest numbers of measured trees and would treat observations from different landscapes as though they arose from a single homogeneous population. Similarly, calculating an arithmetic mean of independently fitted site coefficients would give equal influence to models that differ substantially in sample size and predictive performance. We therefore use a bivariate random-effects hierarchical synthesis in which the independently estimated site-specific power-law relationships are treated as realizations of an underlying cross-landscape crown–stem relationship.



To synthesize the site-specific relationships while retaining among-landscape variation, we extend the linearized form of the power-law model in Eq. (2) to a bivariate random effects hierarchical model. For each individual landscape, the power-law relationship $DBH = a(CA)^b$ can be expressed on the natural log scale as $\ln(DBH) = \ln(a) + b \ln(CA)$, where $\ln(a)$ is the intercept and b is the allometric scaling exponent. In the hierarchical formulation, these site-specific parameters are treated as deviations from an overall cross-landscape relationship. Specifically,

$$\ln(DBH) = (\beta_0 + u_i) + (\beta_1 + v_i) \ln(CA) + \epsilon_i. \quad (4)$$

Here, β_0 is the overall mean intercept on the natural log-scale, corresponding to $\ln(a)$ in Eq. (2), and β_1 is the overall mean scaling exponent, corresponding to b . The terms u_i and v_i represent landscape-specific departures from the overall intercept and exponent, respectively, and ϵ_{ij} represents residual variation among individual trees within landscapes. Thus, for landscape i , the corresponding power-law parameters are $a_i = \exp(\beta_0 + u_i)$ and $b_i = \beta_1 + v_i$. This formulation explicitly recognizes that both the scaling coefficient and scaling exponent can vary among landscapes, while estimating their central cross-landscape tendency and preserving the covariance between them.

The contribution of each landscape-specific model to the hierarchical synthesis was determined using an empirical information-quality weight that incorporated sample size, model fit, and prediction error. Specifically, w_i is the empirical information-quality weight and each model was assigned a weight according to Eq. (5), where n_i is the number of trees used to fit the site-specific model, R_i^2 is its coefficient of determination, and $RMSE_i$ is its root mean square prediction error. This weighting scheme gives greater influence to models supported by larger samples, stronger fit, and lower prediction error, while retaining all site-specific relationships in the synthesis.

$$w_i = \frac{n_i R_i^2}{RMSE_i^2} \quad (5)$$

Figure 7 illustrates both the resulting generalized relationship and the heterogeneity from which it is derived. The ensemble of individual site-specific curves (Fig. 7A) defines the empirical range of crown–stem relationships represented in the dataset, while the weighted hierarchical model describes their central cross-landscape tendency. The associated Monte Carlo prediction envelope propagates uncertainty from the joint distribution and covariance of $\ln(a)$ and b , thereby representing uncertainty in the generalized relationship rather than simply the dispersion of individual observations. The remaining panels show how the generalized model emerges from the distribution of site-level information: variation in scaling exponents and statistical weights (Fig. 7B), the relative contribution of individual landscapes to the synthesis (Fig. 7C), and the joint parameter space of the site-specific intercepts and exponents relative to their weighted mean (Fig. 7D).



An important result of this analysis is that there is no single site-specific curve that can be regarded as universally representative of African TOF landscapes. Instead, the ensemble reveals a family of related allometric relationships whose differences likely reflect variation in tree architecture, species composition, management history, environmental conditions, and tree-size distributions among landscapes. The hierarchical model should therefore not be interpreted as a universal African allometric equation. Rather, it represents a generalized, empirically weighted crown-to-stem transfer function derived from the range of TOF conditions sampled in Kenya, Ghana, Rwanda, and Malawi.

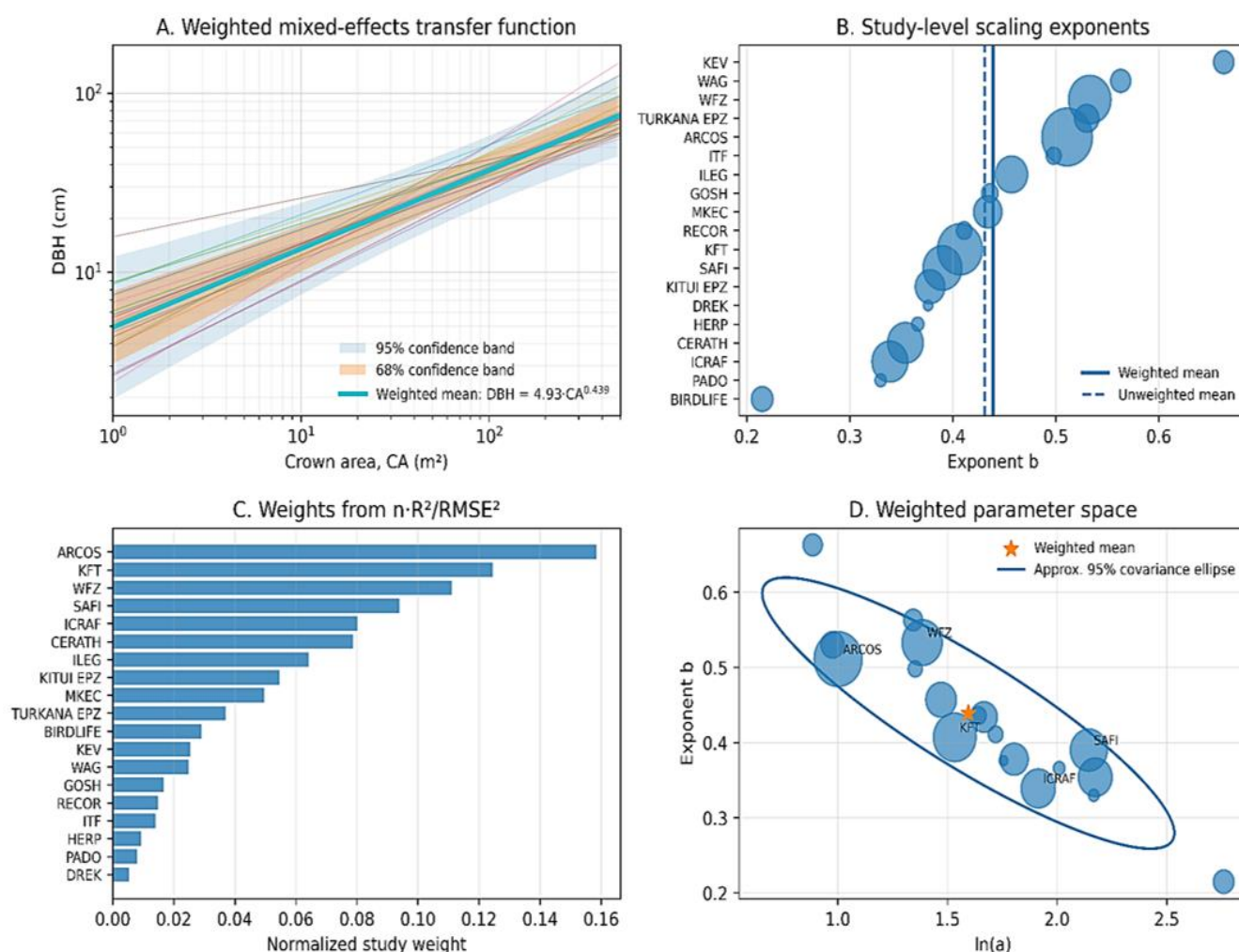


Figure 8. Hierarchical synthesis of site-specific crown area–stem diameter allometry across 19 African TOF landscapes. (A) Ensemble of site-specific DBH–CA power-law models (light grey) and the generalized weighted hierarchical relationship, with Monte Carlo uncertainty intervals propagated from the joint distribution and covariance of $\log(a)$ and b . The dispersion of individual curves represents observed among-landscape variation in crown–stem allometry rather than prediction error for individual trees. **(B)** Distribution of site-specific scaling exponents (b) and their relative statistical weights in the hierarchical synthesis. **(C)** Relative contribution of individual landscape models to the generalized relationship based on the weighting framework incorporating sample



480 **size, model fit, and prediction error. (D) Joint parameter space of log-transformed scaling coefficients $[\log(a)]$ and scaling exponents (b) for the 19 site-specific models, showing their distribution around the weighted hierarchical mean. Together, the panels illustrate how heterogeneous site-specific allometric relationships are retained and synthesized to estimate a generalized cross-landscape crown-to-stem transfer function and its uncertainty.**

This distinction has important implications for application of the dataset. Where sufficient local calibration data are available, site- or landscape-specific crown-to-stem relationships will generally provide the most appropriate basis for estimating DBH
 485 from remotely sensed crown area. Where local calibration data are limited or unavailable, however, the hierarchical relationship provides a defensible generalized transfer function together with an explicit representation of uncertainty. It can therefore serve both as a predictive model and as a prior expectation that can be progressively refined as additional crown–stem observations become available from new landscapes.

490 More broadly, the hierarchical synthesis demonstrates one of the principal advantages of the dataset: its value lies not only in the large number of individual tree measurements, but also in its distributed sampling across contrasting landscapes. The observed variation among site-level equations provides information about the transferability and limits of crown-based allometry that would be obscured by fitting a single pooled regression to all observations. As the dataset expands geographically, the same framework can accommodate additional landscapes and progressively characterize both the central
 495 tendency and geographic variability of crown–stem allometry across African TOF systems.

4.3 Cross-Validation: Model Generalization

Monte Carlo cross-validation was used to assess how well the site-specific DBH-CA allometric models generalize (Steyerberg & Harrell, 2015). The dataset was randomly partitioned into tree records for training (80%) and tree records for validation (20%) subsets, and the process was repeated 100 times to reduce sampling variability (Sileshi, 2014). For each iteration, the
 500 model was fitted to the training data and evaluated on the validation set. The resulting accuracy metrics were averaged across iterations to provide stable estimates of predictive performance (Mulatu et al., 2024). Validation performance was assessed using the coefficient of determination (R^2), root mean square error (RMSE, cm), relative RMSE (%RMSE), and prediction bias expressed both in absolute (cm) and relative (%) terms (Table 7).

505 Model performance on the validation datasets ranged from strong to weak, reflecting variability in sample size, ecological structure, and measurement consistency. The highest-performing models were observed in ARCOS, WAG, KEV, and SAFI, with R^2 values of 0.81, 0.70, 0.79, and 0.73, respectively, and low RMSE and bias. These performances reflect robust generalization and reliability of these models for biomass estimation in comparable TOF systems. Conversely, models such as HERP, PADO, RECOR, BIRDLIFE, and DREK yielded lower validation values, suggesting potential overfitting. This
 510 tendency was most pronounced at sites with small sample sizes ($n < 65$), which increases the likelihood of unstable parameter estimation during resampling. DREK_KE showed negative R^2 and the highest relative bias (+3.79%).



515

520

Table 7: Cross-validation results for site-specific DBH models based on the mean of 100 Monte Carlo iterations across the Kenya, Ghana, Rwanda, and Malawi sites. The validation performances were assessed using R^2 , RMSE, %RMSE, Bias, and percent error (Bias%).

Site ID	R ²	RMSE	RMSE%	Bias	Bias %
DREK	0.50	8.03	27.25	0.815	3.789
ITF	0.55	11.32	37.15	0.085	0.699
KITUI EPZ	0.52	7.85	27.72	0.234	0.99
TURKANA EPZ	0.7	7.44	29.06	0.275	1.365
MKEC	0.52	6.84	30.97	0.267	1.418
ILEG	0.6	6.31	27.93	0.345	1.677
KEV	0.79	10.17	42.8	-0.955	-2.97
KFT	0.62	4.8	24.93	0.164	0.965
CERATH	0.43	6.39	15.88	0.081	0.256
SAFI	0.73	7.26	21.17	-0.243	-0.637
HERP	0.22	8.54	23.82	0.019	0.501
GOSH	0.43	8.83	27.74	-0.45	-1.087
PADO	0.25	9.32	31.09	0.124	1.283
RECOR	0.36	5.58	22.77	-0.11	0.654
ARCOS	0.81	4.3	18.72	0.231	1.061
BIRDLIFE	0.24	4.39	18.15	-0.027	0.15
ICRAF	0.51	6.15	28.7	0.172	0.956
WAG	0.7	13.05	35.01	0.147	0.485
WFZ	0.59	6.89	28.41	0.227	1.029

5 APPLICATION

5.1 Measuring and Mapping Carbon in Trees Outside Forests

We can use the dataset to calibrate and validate an approach for measuring individual tree carbon in landscapes of trees outside of forests. These are landscapes in agricultural areas or severely degraded former forests which have not been readily measured due to the problems inherent in small plot sampling frames typically used in forest carbon inventories. The primary reason 10-30 meter resolution satellite data from platforms such as Landsat or Sentinel-2 are insufficient for mapping carbon in TOF



landscapes is the mismatch between the size of individual tree crowns and the size of the sensor's pixels. In many TOF landscapes, trees are scattered, linear, or isolated, meaning their individual crown areas are often much smaller than a single 100 m² or 900 m² pixel. Thus, when using medium-resolution data, the analysis may be able to predict the *presence* of tree cover as a probability, but cannot delineate the precise boundary of tree crowns. This leads to a significant overestimation of tree cover area; for example, in an earlier study in Odisha, India, Skole et al. (2024) note that Sentinel-2 analysis overestimated tree cover by 198% compared to very high resolution (VHR) data because it essentially "filled in" entire pixels wherever tree cover was detected.

Furthermore, accurate carbon estimation relies on allometric scaling, which requires specific structural parameters for individual trees, such as the crown or stem dimensions. Medium-resolution data are too coarse to perform the object-based segmentation needed to delineate these individual crowns. Because these sensors cannot distinguish individual "tree-objects," they must rely on generalized canopy cover averages for the whole landscape, which introduces significant errors and fails to account for the unique biomass profiles of scattered trees. In contrast, sub 1-m VHR imagery allows for the precise measurement of each tree crown, which can then be used with a ground-calibrated transfer function to estimate stem diameter and carbon stocks with high integrity. Without this sub-meter spatial resolution, the biomass of individual trees in agricultural settings remains largely invisible or incorrectly aggregated, preventing the development of precise measurement and verification systems required for modern carbon markets.

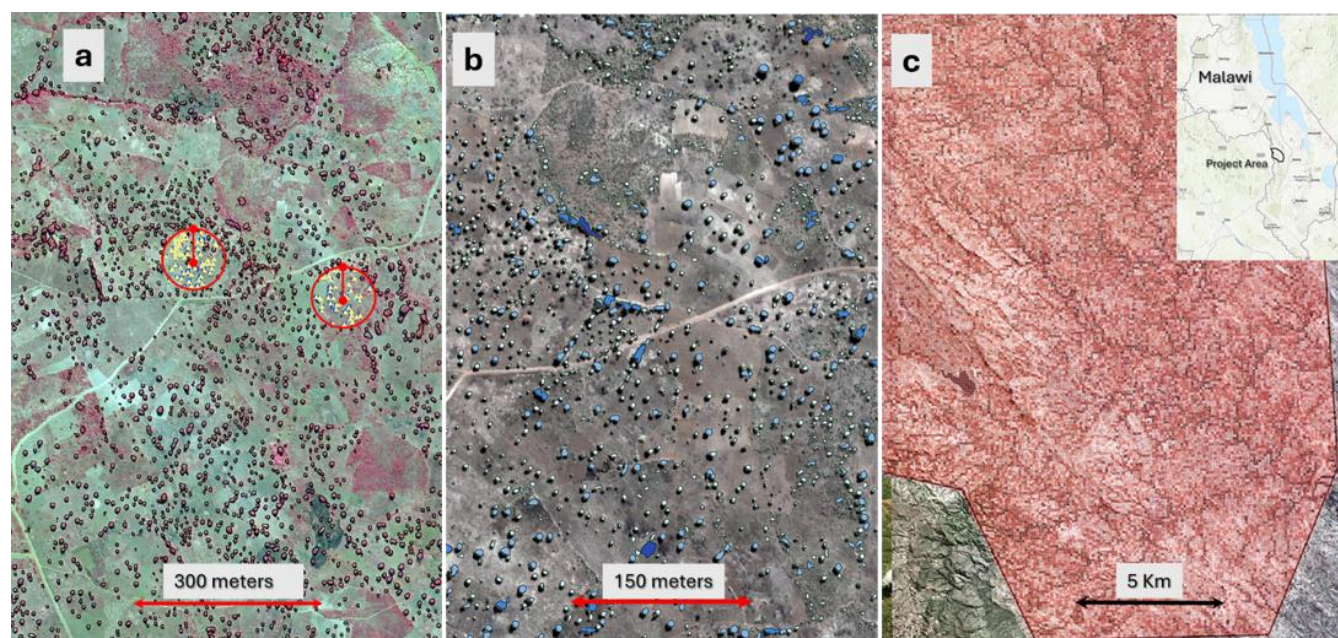


Figure 9. Example application of this framework in the TerraFund for AFR100 program using our model development landscape in the Areeca FLR project. (a) shows some of the sample plots used to calibrate the crown-to-stem transfer function along with the machine leaning derived polygons of CPA, while (b) shows the deep learning mapping of tree crowns and results of allometric scaling of carbon stocks in individual trees, which is represented as a color shade, with darker tones indicating higher carbon stocks.



550 **Individual trees across a large-area landscape are then mapped, shown here as a 1-ha grid aggregation across 50,000 hectares, which contains more than 50×10^6 trees (c).**

For mapping individual tree carbon at landscape scales, we begin with the collection of ground-based calibration data as presented here (Fig. 9a). These ground measurements are used to develop a crown-to-stem transfer function, specifically a canopy-based allometric scaling model that establishes a mathematical relationship between CPA and stem parameters such as DBH. This relationship is critical because it allows for the estimation of physical tree attributes that are otherwise invisible to medium resolution satellite sensors. In our recent studies (Skole et al. 2024), we apply this DBH-CA function across large areas using very high resolution VHR satellite imagery to provide the CA parameters derived from deep learning models with automated segmentation and precise delineation of individual tree crowns from the landscape background to derive their CPA. Once individual tree crowns are mapped and their CPAs are calculated, the previously derived crown-to-stem transfer function is applied to estimate the DBH for every identified tree. These estimated stem parameters are then integrated with standard and local biomass allometric scaling equations to calculate the carbon content of each individual tree (Fig. 9b). This multi-step process effectively bridges the gap between the accuracy of field-level accuracy and the coverage of satellite data, producing comprehensive project reporting (Table S4) of carbon stocks from a census of all trees. The approach provides a robust measurement, reporting, and verification (MRV) framework that can track the biomass in non-forest agricultural landscape across over large areas (Fig. 9c).

5.2 Allometric and Metabolic Scaling

The relationship between stem diameter and crown area in this study demonstrates allometric patterns with both theoretical and applied implications. Within our DBH-CA model, ($DBH = aCA^b$) the value of its exponent reflects patterns of biological growth and resource allocation to tree stems and tree crowns (Sileshi et al., 2023). Across all models of DBH predicted from CA presented here, the scaling exponent b ranges from 0.215 to 0.663, indicating that DBH scales sub-linearly with CA; increases in crown area are associated with proportionally smaller increases in stem diameter. For the inverse form, CA-DBH, the exponent ranges from 1.083 to 2.068 suggesting that CA scales much faster than DBH. Metabolic scaling theory (MST) suggests how various tree physical dimensions change relative to one another based on physiological resource transport networks and mechanical stability (Brown et al., 2004). MST following West et al., (2009) assumes that crown diameter (CD) should ordinarily scale to the 0.667 power of stem diameter (DBH). When applied to the relationship between crown diameter (CD) and stem diameter, $CD = aDBH^b$ (CD-DBH), the exponent, b , ranges from 0.54 to 1.02 in models presented here, which suggests an expected MST relationship which is sublinear and generally not isomorphic.

The Tallow tree dataset developed by Jucker et al. (2022) and the Pantropical Variability in Tree Crown Allometry database from Loubota Panzou et al., (2021) contain merged collections of other crown diameters similar to one we present here. A closer examination of all project sites in our dataset shows a mean value of the exponent of CD-DBH plus or minus the 95% confidence interval of 0.68 ± 0.063 . This is consistent with the Tallow database (0.62 ± 0.062) as a whole, as well as specifically



for its biogeographic categories of Tropical Dry Forest Angiosperms and Tropical Savanna Angiosperms. Our CD-DBH exponents are also consistent with data provided in Loubota Panzou et al. (2021) for Kenya, in which the exponent of DBH is 0.732. Our dataset CD-DBH is also similar but lower than Loubota Panzou et al. for Malawi, 0.746.

Our dataset provides scaling exponents close to an MST value of 0.667 but variations higher and lower occur. Our dataset reflects a sample from a range of landscapes with differing human management conditions and a mix of natural tree occurrences and agroforestry. The expectation was that allometric scaling of CD-DBH would not be isomorphic but would also not be sublinear. Open grown trees would be expected to increase crown dimensions more than stem dimensions. But since many of the trees and landscapes with TOF are drier or more degraded and thus experiencing water and nutrient stress, limits on CD relative to DBH could be present. These deviations from theoretical exponents may also indicate environmental stressors or adaptations specific to certain species or landscape types. Shifts in scaling exponents are expected based on biome and species traits. According to Sileshi et al., (2023) the typical exponent for tropical trees ranges from 0.5 to 0.7, while for savanna trees, it falls between 0.4 and 0.6 and is approximately 0.5 for trees in dry forests. In savanna environments, lower exponents are expected due to water stress and the presence. In arid regions, lower exponents reflect limited trunk growth relative to crown expansion, also driven by scarcity of water.

More work could be done with this dataset and other data to evaluate MST and general geometric similarity theory (GST) metrics. We see opportunities to expand models of crown-based biomass and carbon allometry using MST and GST parameters or perhaps provide overall constraints on the general DBH-CPA model predictions. Since MST suggests that CD is isomorphic with tree height, H ($CD \propto H^1$), the estimation of tree H would be a useful constraint as well. However we note that the prediction of H from CPA is not as tight as with DBH in our dataset. This approach to model constraints and the inclusion of additional predictors could be especially important in training more advanced deep machine learning models as described above.

6 COMMUNITY VALUE, ACCESSIBILITY, AND REUSABILITY

6.1 Community value. This data product was developed to address an important measurement gap in the monitoring of trees outside forests (TOF), agroforestry systems, and forest and landscape restoration. Very high-resolution satellite imagery and machine-learning methods increasingly permit individual tree crowns to be delineated over large landscapes, but satellite observations cannot directly measure diameter at breast height (DBH), which remains the principal predictor used in widely established tree biomass allometry. Concurrent field measurements of crown dimensions and stem diameter that can be used to calibrate this relationship remain comparatively scarce, particularly for heterogeneous African TOF landscapes. The dataset presented here provides a geographically distributed collection of individual-tree measurements designed specifically to address this gap. Following quality control, it contains 3,493 trees with concurrent measurements of DBH, crown dimensions, tree height, species identity, geographic location, and associated plot and landscape identifiers. The observations span



restoration and land-management systems in Kenya, Ghana, Rwanda, and Malawi, including agroforestry, reforestation, assisted natural regeneration, and other TOF settings. The dataset therefore provides a resource that would be costly and time-consuming for individual research groups to reproduce and can support a broad community working in Earth observation, forest mensuration, allometry, carbon accounting, agroforestry, landscape restoration, and measurement, reporting, and verification. In particular, it provides the empirical bridge needed to connect increasingly available individual-tree remote-sensing observations with established stem-based biomass estimation methods.

6.2 Accessibility. The data product is designed to permit independent users to understand and apply the observations without requiring access to the original field campaigns. The manuscript provides the geographic and ecological context of data collection, a standardized field measurement protocol, explicit quality-control and exclusion procedures, and a data dictionary defining the principal variables, units, identifiers, and spatial information. For example, the dataset records site and plot identifiers, tree geographic coordinates, species identity, DBH, two perpendicular crown diameters, mean crown diameter, and tree height, with geographic coordinates referenced to WGS84. The accompanying field protocol further documents plot allocation, measurement procedures, and standard operating procedures, thereby providing users with the information required to interpret both the measurements and their provenance. Depositing the complete dataset in a persistent open repository under a non-restrictive license will ensure that the observations and supporting documentation remain freely accessible and citable independently of the manuscript.

6.3 Reusability. The dataset has intentionally been structured as a general observational resource rather than as data restricted to the particular allometric models demonstrated in this manuscript. The concurrent measurements allow users to develop alternative DBH–crown area and DBH–crown diameter transfer functions, investigate relationships involving tree height and species, develop species- or landscape-specific allometric models, validate independent tree-structure products, and combine these observations with individual-tree crowns derived from new or existing very high-resolution imagery. This is especially important because the manuscript demonstrates substantial geographic and structural variability in crown–stem relationships. Site-specific DBH–crown area relationships are statistically significant but vary appreciably among landscapes, demonstrating both the utility of the observations and the importance of geographically distributed calibration rather than reliance on a single universal relationship. The dataset has also undergone explicit quality control: records lacking required measurements or reliable geographic coordinates and trees with structural conditions that could compromise conventional crown–stem allometry were excluded, reducing the original collection of 5,218 records to a curated dataset of 3,493 trees. The manuscript further quantifies model performance, prediction error, bias, and variation among sites, providing users with information needed to judge where and how the observations can appropriately be reused. Consequently, the data can serve not only to reproduce the analyses presented here but also as an independent calibration, validation, and comparative resource for future individual-tree remote-sensing, ecological scaling, biomass, carbon, and restoration-monitoring studies.



7 CONCLUSIONS

This study presents a standardized, georeferenced dataset of individual-tree structural measurements developed specifically to address a persistent gap between field-based tree allometry and emerging approaches for mapping individual trees using very high-resolution (VHR) remote sensing. Following quality control, the dataset contains 3,493 trees measured across landscapes in Kenya, Ghana, Rwanda, and Malawi and includes concurrent measurements of stem diameter at breast height (DBH), crown dimensions, tree height, species identity, geographic location, and associated plot and landscape information. The measurements were collected using a standardized field protocol designed to support consistent crown–stem allometry and spatial linkage with individual tree crowns delineated from VHR imagery. The resulting dataset therefore provides both a substantial collection of African tree measurements and a reproducible framework for expanding comparable observations to additional landscapes.

A principal motivation for developing the dataset was to overcome an important measurement discontinuity between remote sensing and conventional biomass allometry. VHR satellite imagery and machine-learning methods can increasingly identify individual trees and delineate their crown projected areas over large landscapes, but they cannot directly observe DBH, which remains the principal predictor in most established tree biomass equations. The paired crown and stem measurements presented here provide the empirical observations needed to bridge these two measurement domains. Crown-to-stem transfer functions calibrated from the dataset allow remotely observed crown area to be translated into estimates of DBH, which can subsequently be used with established stem-based biomass equations. This approach retains compatibility with conventional forest mensuration and biomass allometry while extending their application from field measurements and sample plots to remotely sensed individual-tree observations over much larger areas.

The site-specific analyses presented in this study demonstrate both the utility of this approach and the importance of recognizing heterogeneity in tree allometry. Significant relationships between crown area and DBH occur across the sampled landscapes, but the scaling coefficients, exponents, and predictive performance vary among sites. Rather than representing only statistical uncertainty, this variation contains information about differences in tree architecture and the ecological and management conditions under which trees grow. The hierarchical synthesis of site-specific models provides one approach for retaining this among-landscape variation while estimating a generalized cross-landscape crown-to-stem relationship. The resulting generalized model should not be interpreted as a universal African allometric equation; rather, it demonstrates how geographically distributed calibration observations can be combined while preserving information about the variability and uncertainty of crown–stem relationships among landscapes. Where adequate local observations are available, locally calibrated relationships remain preferable, whereas generalized relationships can provide an initial transfer function or prior expectation where local calibration data are limited.



The value of the dataset, however, extends beyond the particular DBH–crown area models demonstrated here. Because stem
685 diameter, two perpendicular crown dimensions, tree height, species identity, geographic location, and landscape context were
recorded for individual trees, the data provide a reusable observational resource for developing and testing alternative
allometric formulations. Researchers can use the measurements to investigate relationships among crown diameter, crown
projected area, DBH, and height; develop species-, functional-group-, site-, or landscape-specific models; evaluate multivariate
and hierarchical allometric approaches; and examine how crown–stem relationships vary with tree size, species composition,
690 and landscape conditions. The data can also contribute to studies of crown architecture, geometric and metabolic scaling,
agroforestry structure, and the structural characteristics of open-grown and managed trees. In this sense, the dataset should be
viewed not as a collection tied to a single fitted allometric equation, but as a set of primary tree-level observations from which
alternative models and scientific hypotheses can be developed and tested.

695 A second major area of reuse is the calibration and validation of individual-tree Earth observation products. The geographic
coordinates and tree-level structural measurements provide reference observations that can be linked with tree crowns derived
from VHR satellite imagery and potentially other remotely sensed measurements of vegetation structure. They can therefore
support evaluation of crown-delineation algorithms, comparison of remotely sensed and field-measured crown dimensions,
development of crown-based biomass models, and testing of new statistical and machine-learning approaches for predicting
700 tree structural attributes. As individual-tree remote sensing continues to develop, retaining the primary measurements rather
than only the fitted equations allows future investigators to reanalyse the observations using methods and Earth observation
datasets that were not available when the measurements were originally collected.

These capabilities are particularly important for Trees Outside Forests (TOF). Sparse, scattered, linear, agroforestry, and other
705 non-forest tree configurations remain difficult to characterize using conventional forest inventory designs and medium-
resolution remote sensing. Yet these trees constitute an important component of many African agricultural and restoration
landscapes. Combining VHR individual-tree mapping with ground-calibrated crown–stem relationships provides a pathway
toward census-based measurement in which individual trees can be identified, measured, and assigned biomass and carbon
estimates across large areas. Such an approach can complement rather than replace conventional forest inventories by
710 extending allometric measurement into landscape components that are frequently underrepresented in plot-based monitoring
systems.

Forest and Landscape Restoration (FLR) provides an immediate application of this capability. The dataset was developed from
landscapes associated with AFR100 restoration projects and encompasses agroforestry, reforestation, assisted natural
715 regeneration, and other land-management systems. Current restoration indicators frequently emphasize hectares under
restoration, trees planted, or changes in tree cover, but these measures alone do not quantify changes in tree structure, biomass,
or carbon stocks. Crown-based individual-tree observations calibrated with datasets such as this one can provide a quantitative



baseline of existing tree resources that can subsequently be remeasured using VHR remote sensing. Repeated observations could document changes in tree abundance, crown dimensions, biomass, and carbon stocks and thereby provide an additional quantitative dimension to restoration monitoring and measurement, reporting, and verification (MRV).

The longer-term scientific value of the dataset will increase as comparable observations are collected across additional ecological zones, species assemblages, management systems, and African regions. The standardized field protocol provided with the dataset is therefore an important component of the data product. Application of a common protocol would allow new crown–stem observations to be incorporated into an expanding distributed database, providing increasingly robust estimates of both the central tendencies and geographic variability of African tree allometry. Such expansion would also allow stronger testing of the extent to which crown–stem relationships can be transferred among landscapes and where local calibration remains necessary. More broadly, the dataset provides an empirical connection between three measurement domains that have often developed separately: field-based tree mensuration, individual-tree Earth observation, and landscape-scale biomass and carbon monitoring. Its principal contribution is therefore not a single allometric equation, but a standardized set of primary observations that enables these domains to be linked. By making concurrent crown, stem, height, species, and spatial measurements available for reuse, the dataset provides a foundation for improving crown-based allometry, testing ecological scaling relationships, developing and validating new individual-tree remote-sensing methods, and extending biomass and carbon measurement to TOF landscapes. This combination of standardized observations, demonstrated allometric application, and capacity for future reuse provides a basis for progressively improving individual-tree measurement and for more fully representing trees outside forests in African restoration monitoring, carbon accounting, and Earth system observations.

Code and data availability

The complete and open-access dataset has been published in Zendo at <https://doi.org/10.5281/zenodo.20862469> (Alimo et al. 2025) under Creative Commons Attribute 4.0 license. Data are provided in CSV format, accompanied by a data dictionary, field measurement protocol, site and plot identifiers, units, coordinate reference information and documentation of quality control procedures. All results from data analysis needed to evaluate this report are also available in the main text or in tables and figures. Code developed in R routine statistical code and graphical production libraries (e.g. ggplot2) are not archived in the Zendo but are available as open source from the R Project (<https://www.r-project.org/>). Satellite data were used to process tree crown data but these products were not essential to this field dataset, but can be obtained from the NASA Commercial Smallsat Data Acquisition Program (<https://www.earthdata.nasa.gov/esds/csda>). The remote sensing–derived products, spatial data layer products, and validation data, are also available from the Tropical Rain Forest Information Center, a NASA ESIP Data Center, now maintained by the Global Observatory for Ecosystem Services at Michigan State University (<https://goeslab.us/Indiadata>) and from the NASA Land Cover and Land Use Change Program data server (<https://lcluc.umd.edu/content/metadata>).



750 **Supplement link**

The supplementary material is provided online as separate file.

Author contributions

DS and TA developed the project, field protocols, analytical methods and design, and prepared the first draft of the manuscript. DS was the project lead and supervised the preparation of the manuscript. All authors participated in the preparation of the manuscript. JS supervised the remote sensing analysis and deployment of its ground calibration, as well as model development and testing and data analysis. TT provided the lead on content related to Forest Landscape Restoration requirements. TA, DN, JS and DS developed the field protocols. TA and DN lead the field collection methods overall and JK and AC led the field teams in Malawi and Rwanda. VO led the field team in Kenya. TA and DN led the field teams in Ghana. TA, DN, and CP led initial field testing in Ghana and Kenya. DS supervised all field work and quality control.

760 **Competing interests**

The authors declare no competing interests.

Disclaimer

Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

The authors are grateful to the AFR100 program, World Resources Institute, and One Tree Planted organizations for support and access to field sites. We also acknowledge the many project developers and communities, as named in this paper, who are doing very important restoration work on the ground in Africa. We acknowledge the dedicated work of the local participant field crews. The execution of the field work, international management, and general administration of this very complex operations and budget could not have been done without the cooperation and leadership of Katie James in the Department of Forestry at Michigan State University. Project leadership in the Land and Carbon Lab of WRI was extremely important for the guiding the development and deployment of the field data collection objectives, developing measurement aims and protocols, and designing modelling strategies, including Nancy Harris, Fred Stolle, and Katie Reyntar. We thank Sean Dewitt, Will



Anderson, Dow Martin and Edward Saenz in the Restoration Program and Peter Ndunda, Lawrence Eric Ng'ang'a, Spencer Ngoma, Duncan Okowa, and Diana Mawoko at WRI Africa.

Financial support

780 This project was supported by a grant from the National Aeronautics and Space Administration (NASA), under the Land Cover and Land Use Change Program (LCLUC) under NASA grant 80NSSC21K0315 and a grant from the World Resources Institute under P00634478, *TerraMatch Champions Project Training, Mapping, and Carbon Estimation*. This project is based on research that was partially supported by the AgBioResearch at Michigan State University with funding from the Hatch Act capacity funding program (Accession Number7001998) from the USDA National Institute of Food and Agriculture.

Review statement

785 The review statement will be added by Copernicus Publications listing the handling editor as well as all contributing referees according to their status anonymous or identified.

References

790

- Aabeyir, R., Adu-Bredu, S., Agyare, W. A., & Weir, M. J. (2020). Allometric models for estimating aboveground biomass in the tropical woodlands of Ghana, West Africa. *Forest Ecosystems*, 7(1), 41. <https://doi.org/10.1186/s40663-020-00250-3>
- 795 Alcaras, E., Parente, C., & Vallario, A. (2020). The importance of the coordinate transformation process in using heterogeneous data in coastal and marine geographic information system. *Journal of Marine Science and Engineering*, 8(9), 708. <https://doi.org/10.3390/jmse8090708>
- Alimo, T. A., Skole, D. L., Samek, J. H., Ndalowa, D., Oeba, V., Kamoto, J., Chioza, A., Papa, C., Spore, J., Levin, D., Brandt, J., & Tumeo, T. (2026). *Individual tree crown and stem allometric dataset for forest and landscape restoration monitoring* [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.20862469>
- 800 Anyanwu, C. N., Ojike, O., Emodi, N. V., Ekwe, E. B., Okereke, C., Diemuodeke, E. O., Elochukwu, A. E., & Nnamani, U. A. (2023). Deep decarbonization options for the agriculture, forestry, and other land use (AFOLU) sector in Africa: A systematic literature review. *Environmental Monitoring and Assessment*, 195(5), 565. <https://doi.org/10.1007/s10661-023-11184-y>



- Asigbaase, M., Dawoe, E., Abugre, S., Kyereh, B., & Ayine Nsor, C. (2023). Allometric relationships between stem diameter,
 805 height and crown area of associated trees of cocoa agroforests of Ghana. *Scientific Reports*, 13(1), 14897.
<https://doi.org/10.1038/s41598-023-42219-6>
- Baccini, A., Laporte, N., Goetz, S. J., Sun, M., & Dong, H. (2008). A first map of tropical Africa's above-ground biomass
 derived from satellite imagery. *Environmental Research Letters*, 3(4), 045011. <https://doi.org/10.1088/1748-9326/3/4/045011>
- Baskerville, G. L. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest*
 810 *Research*, 2(1), 49–53. <https://doi.org/10.1139/x72-009>
- Beckschäfer, P., Schnell, S., & Kleinn, C. (2018). Monitoring and assessment of trees outside forests (TOF). In J. C. Dagar &
 V. P. Tewari (Eds.), *Agroforestry: Anecdotal to modern science* (pp. 137–161). Springer Singapore.
https://doi.org/10.1007/978-981-10-7650-3_5
- Bossy, T., Ciais, P., Renaudineau, S., Wan, L., Ygorra, B., Adam, E., Barbier, N., Bauters, M., Delbart, N., Frappart, F., Gara,
 815 T. W., Hamunyela, E., Ifo, S. A., Jaffrain, G., Maisongrande, P., Mugabowindekwe, M., Mugiraneza, T., Normandin, C.,
 Obame, C. V., ... Wigneron, J.-P. (2025). State of the art in remote sensing monitoring of carbon dynamics in African tropical
 forests. *Frontiers in Remote Sensing*, 6, 1532280. <https://doi.org/10.3389/frsen.2025.1532280>
- Brandt, M., Tucker, C. J., Kariyaa, A., Rasmussen, K., Abel, C., Small, J., Chave, J., Rasmussen, L. V., Hiernaux, P., Diouf,
 A. A., & Kergoat, L. (2020). An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature*, 587(7832),
 820 78–82. <https://doi.org/10.1038/s41586-020-2824-5>
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology.
Ecology, 85(7), 1771–1789. <https://doi.org/10.1890/03-9000>
- Brown, S. (1997). *Estimating biomass and biomass change of tropical forests: A primer* (FAO Forestry Paper No. 134). Food
 and Agriculture Organization of the United Nations.
- 825 Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against
 avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T.,
 Lescure, J.-P., Nelson, B. W., Ogawa, H., Puig, H., Riéra, B., & Yamakura, T. (2005). Tree allometry and improved estimation
 830 of carbon stocks and balance in tropical forests. *Oecologia*, 145(1), 87–99. <https://doi.org/10.1007/s00442-005-0100-x>
- Corona, P., Fattorini, L., & Franceschi, S. (2011). Two-stage sector sampling for estimating small woodlot attributes. *Canadian*
Journal of Forest Research, 41(9), 1819–1826. <https://doi.org/10.1139/x11-101>
- De Foresta, H., Somarriba, E., Temu, A., Boulanger, D., Feuilly, H., Gauthier, M., & Taylor, D. (2013). *Towards the*
assessment of trees outside forests: A thematic report prepared in the framework of the Global Forest Resources Assessment
 835 *2010* (Forest Resources Assessment Working Paper No. 183). Food and Agriculture Organization of the United Nations.



- DeFries, R. S., Houghton, R. A., Hansen, M. C., Field, C. B., Skole, D., & Townshend, J. (2002). Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. *Proceedings of the National Academy of Sciences of the United States of America*, 99(22), 14256–14261. <https://doi.org/10.1073/pnas.182560099>
- Di Gregorio, A., & Leonardi, U. (2016). *Land cover classification system—User manual: Software version 3*. Food and Agriculture Organization of the United Nations.
- Dmuchowski, W., Baczewska-Dąbrowska, A. H., & Gworek, B. (2024). The role of temperate agroforestry in mitigating climate change: A review. *Forest Policy and Economics*, 159, 103136. <https://doi.org/10.1016/j.forpol.2023.103136>
- Dubayah, R., Armston, J., Healey, S. P., Bruening, J. M., Patterson, P. L., Kellner, J. R., Duncanson, L., Saarela, S., Ståhl, G., Yang, Z., Tang, H., Blair, J. B., Fatoyinbo, L., Goetz, S., Hancock, S., Hansen, M., Hofton, M., Hurtt, G., & Luthcke, S. (2022). GEDI launches a new era of biomass inference from space. *Environmental Research Letters*, 17(9), 095001. <https://doi.org/10.1088/1748-9326/ac8694>
- Fischer, F. J., Labrière, N., Vincent, G., Hérault, B., Alonso, A., Memiaghe, H., Bissiengou, P., Kenfack, D., Saatchi, S., & Chave, J. (2020). A simulation method to infer tree allometry and forest structure from airborne laser scanning and forest inventories. *Remote Sensing of Environment*, 251, 112056. <https://doi.org/10.1016/j.rse.2020.112056>
- Frank, S., Lessa Derci Augustynczyk, A., Havlík, P., Boere, E., Ermolieva, T., Fricko, O., Di Fulvio, F., Gusti, M., Krisztin, T., Lauri, P., Palazzo, A., & Wögerer, M. (2024). Enhanced agricultural carbon sinks provide benefits for farmers and the climate. *Nature Food*, 5(9), 742–753. <https://doi.org/10.1038/s43016-024-01039-1>
- Friedlingstein, P., O’Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Aas, K., Alin, S. R., ... Zeng, J. (2026). Global Carbon Budget 2025. *Earth System Science Data*, 18(5), 3211–3288. <https://doi.org/10.5194/essd-18-3211-2026>
- Graves, S. J., Caughlin, T. T., Asner, G. P., & Bohlman, S. A. (2018). A tree-based approach to biomass estimation from remote sensing data in a tropical agricultural landscape. *Remote Sensing of Environment*, 218, 32–43. <https://doi.org/10.1016/j.rse.2018.09.009>
- Houghton, R. A. (2007). Balancing the global carbon budget. *Annual Review of Earth and Planetary Sciences*, 35, 313–347. <https://doi.org/10.1146/annurev.earth.35.031306.140057>
- Houghton, R. A., & Castanho, A. (2023). Annual emissions of carbon from land use, land-use change, and forestry from 1850 to 2020. *Earth System Science Data*, 15(5), 2025–2054. <https://doi.org/10.5194/essd-15-2025-2023>
- Huang, C., Gong, W., & Pang, Y. (2022). Remote sensing and forest carbon monitoring—A review of recent progress, challenges and opportunities. *Journal of Geodesy and Geoinformation Science*, 5(2), 124–147. <https://doi.org/10.11947/j.JGGS.2022.0212>
- Jucker, T., Fischer, F. J., Chave, J., Coomes, D. A., Caspersen, J., Ali, A., Loubota Panzou, G. J., Feldpausch, T. R., Falster, D., Usoltsev, V. A., Adu-Bredu, S., Alves, L. F., Aminpour, M., Angoboy, I. B., Anten, N. P. R., Antin, C., Askari, Y., Muñoz,



- R., Ayyappan, N., ... Zavala, M. A. (2022). Tallo: A global tree allometry and crown architecture database. *Global Change Biology*, 28(17), 5254–5268. <https://doi.org/10.1111/gcb.16302>
- Kuyah, S., Mbow, C., Sileshi, G. W., van Noordwijk, M., Tully, K. L., & Rosenstock, T. S. (2016). Quantifying tree biomass carbon stocks and fluxes in agricultural landscapes. In T. S. Rosenstock, M. C. Rufino, K. Butterbach-Bahl, E. Wollenberg, & M. Richards (Eds.), *Methods for measuring greenhouse gas balances and evaluating mitigation options in smallholder agriculture* (pp. 119–134). Springer. https://doi.org/10.1007/978-3-319-29794-1_6
- 875 Kuyah, S., Muthuri, C., Jamnadass, R., Mwangi, P., Neufeldt, H., & Dietz, J. (2012). Crown area allometries for estimation of aboveground tree biomass in agricultural landscapes of western Kenya. *Agroforestry Systems*, 86(2), 267–277. <https://doi.org/10.1007/s10457-012-9529-1>
- Kuyah, S., & Rosenstock, T. S. (2015). Optimal measurement strategies for aboveground tree biomass in agricultural landscapes. *Agroforestry Systems*, 89(1), 125–133. <https://doi.org/10.1007/s10457-014-9747-9>
- 880 Lewis, S. L., Wheeler, C. E., Mitchard, E. T. A., & Koch, A. (2019). Restoring natural forests is the best way to remove atmospheric carbon. *Nature*, 568(7750), 25–28. <https://doi.org/10.1038/d41586-019-01026-8>
- Loubota Panzou, G. J., Fayolle, A., Jucker, T., Phillips, O. L., Bohlman, S., Banin, L. F., Lewis, S. L., Affum-Baffoe, K., Alves, L. F., Antin, C., Arets, E., Arroyo, L., Baker, T. R., Barbier, N., Beeckman, H., Berger, U., Bocko, Y. E., Bongers, F., Bowers, S., ... Feldpausch, T. R. (2021). Pantropical variability in tree crown allometry. *Global Ecology and Biogeography*, 30(2), 459–475. <https://doi.org/10.1111/geb.13231>
- 885 Luck, L., Hutley, L. B., Calders, K., & Levick, S. R. (2020). Exploring the variability of tropical savanna tree structural allometry with terrestrial laser scanning. *Remote Sensing*, 12(23), 3893. <https://doi.org/10.3390/rs12233893>
- Malkoç, E., Rüetschi, M., Ginzler, C., & Waser, L. T. (2021). Countrywide mapping of trees outside forests based on remote sensing data in Switzerland. *International Journal of Applied Earth Observation and Geoinformation*, 100, 102336. <https://doi.org/10.1016/j.jag.2021.102336>
- 890 Mansourian, S., & Berrahmouni, N. (2021). *Review of forest and landscape restoration in Africa 2021*. Food and Agriculture Organization of the United Nations.
- Martínez Cano, I., Muller-Landau, H. C., Wright, S. J., Bohlman, S. A., & Pacala, S. W. (2019). Tropical tree height and crown allometries for the Barro Colorado Nature Monument, Panama: A comparison of alternative hierarchical models incorporating interspecific variation in relation to life history traits. *Biogeosciences*, 16(4), 847–862. <https://doi.org/10.5194/bg-16-847-2019>
- 895 Mbow, C., Smith, P., Skole, D., Duguma, L., & Bustamante, M. (2014). Achieving mitigation and adaptation to climate change through sustainable agroforestry practices in Africa. *Current Opinion in Environmental Sustainability*, 6, 8–14. <https://doi.org/10.1016/j.cosust.2013.09.002>
- Mbow, C., Verstraete, M. M., Sambou, B., Diaw, A. T., & Neufeldt, H. (2014). Allometric models for aboveground biomass in dry savanna trees of the Sudan and Sudan–Guinean ecosystems of southern Senegal. *Journal of Forest Research*, 19(3), 340–347. <https://doi.org/10.1007/s10310-013-0414-1>
- 900



- Minang, P. A., Duguma, L. A., Bernard, F., Mertz, O., & van Noordwijk, M. (2014). Prospects for agroforestry in REDD+ landscapes in Africa. *Current Opinion in Environmental Sustainability*, 6, 78–82. <https://doi.org/10.1016/j.cosust.2013.10.015>
- Minarsch, E. M. L., Thoss, V. M., & Beule, L. (2025). Guidance on spatial sampling designs and data reusability in agroforestry research. *Agroforestry Systems*, 99, Article 127. <https://doi.org/10.1007/s10457-025-01225-5>
- Mitchard, E. T. A., Feldpausch, T. R., Brien, R. J. W., Lopez-Gonzalez, G., Monteagudo, A., Baker, T. R., Lewis, S. L., Lloyd, J., Quesada, C. A., Gloor, M., ter Steege, H., Meir, P., Alvarez, E., Araujo-Murakami, A., Aragão, L. E. O. C., Arroyo, L., Aymard, G., Banki, O., Bonal, D., ... Phillips, O. L. (2014). Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites. *Global Ecology and Biogeography*, 23(8), 935–946. <https://doi.org/10.1111/geb.12168>
- Mugabowindekwe, M., Brandt, M., Chave, J., Reiner, F., Skole, D. L., Kariryaa, A., Igel, C., Hiernaux, P., Ciais, P., Mertz, O., Tong, X., Li, S., Rwanyiziri, G., Dushimiyimana, T., Ndoli, A., Uwizeyimana, V., Lillesø, J. P. B., Gieseke, F., Tucker, C. J., ... Fensholt, R. (2023). Nation-wide mapping of tree-level aboveground carbon stocks in Rwanda. *Nature Climate Change*, 13(1), 91–97. <https://doi.org/10.1038/s41558-022-01544-w>
- Mugabowindekwe, M., Brandt, M., Mukuralinda, A., Ciais, P., Reiner, F., Kariryaa, A., Igel, C., Chave, J., Mertz, O., Hiernaux, P., Tong, X., Rwanyiziri, G., Gominiski, D., Li, S., Liu, S., Gasangwa, I., Hategekimana, Y., Ndoli, A., Nduwamungu, J., ... Fensholt, R. (2024). Trees on smallholder farms and forest restoration are critical for Rwanda to achieve net zero emissions. *Communications Earth & Environment*, 5, Article 113. <https://doi.org/10.1038/s43247-024-01278-x>
- Mugo, J. M., Njunge, J. T., Malimbwi, R. E., Kigomo, B. N., Mwasi, B. N., & Muchiri, M. N. (2011). Models for predicting stem diameter from crown diameter of open grown trees in Sondu-Nyando River catchment, Kenya. *Asian Journal of Agricultural Sciences*, 3(2), 119–126.
- Mulatu, A., Negash, M., & Asrat, Z. (2024). Species-specific allometric models for reducing uncertainty in estimating above ground biomass at Moist Evergreen Afromontane Forest of Ethiopia. *Scientific Reports*, 14(1), 1147. <https://doi.org/10.1038/s41598-023-51002-6>
- Noble, I. R., Apps, M. J., Houghton, R. A., Lashof, D. A., Makundi, W., Murdiyarso, D., Murray, B., Sombroek, W., Valentini, R., Amano, M., Fearnside, P., Frangi, J., Frumhoff, P., Goldberg, D., Higuchi, N., Janetos, A., Kirschbaum, M. U. F., Lasco, R., Nabuurs, G., ... Scholes, R. J. (2000). Implications of different definitions and generic issues. In R. T. Watson, I. R. Noble, B. Bolin, N. H. Ravindranath, D. J. Verardo, & D. J. Dokken (Eds.), *Land use, land-use change, and forestry* (pp. 53–126). Cambridge University Press.
- Papucci, E., Valbuena, R., Roberge, C., Appiah Mensah, A., & Ståhl, G. (2026). A review of forest biomass assessments based on remote sensing reveals progress in methodological quality—but major challenges remain. *Forestry: An International Journal of Forest Research*, 99(2), cpag007. <https://doi.org/10.1093/forestry/cpag007>
- Parr, C. L., te Beest, M., & Stevens, N. (2024). Conflation of reforestation with restoration is widespread. *Science*, 383(6684), 698–701. <https://doi.org/10.1126/science.adj0899>
- Quandt, A., Neufeldt, H., & Gorman, K. (2023). Climate change adaptation through agroforestry: Opportunities and gaps. *Current Opinion in Environmental Sustainability*, 60, 101244. <https://doi.org/10.1016/j.cosust.2022.101244>



- Reiner, F., Brandt, M., Tong, X., Skole, D., Kariryaa, A., Ciais, P., Davies, A., Hiernaux, P., Chave, J., Mugabowindekwe, M., Igel, C., Oehmcke, S., Gieseke, F., Li, S., Liu, S., Saatchi, S., Boucher, P., Singh, J., Taugourdeau, S., ... Fensholt, R. (2023). More than one quarter of Africa's tree cover is found outside areas previously classified as forest. *Nature Communications*, 14, Article 2258. <https://doi.org/10.1038/s41467-023-37880-4>
- 940 Rosenstock, T. S., Wilkes, A., Jallo, C., Namoi, N., Bulusu, M., Suber, M., Mboi, D., Mulia, R., Simelton, E., Richards, M., Gurwick, N., & Wollenberg, E. (2019). Making trees count: Measurement and reporting of agroforestry in UNFCCC national communications of non-Annex I countries. *Agriculture, Ecosystems & Environment*, 284, 106569. <https://doi.org/10.1016/j.agee.2019.106569>
- Rykiel, E. J., Jr. (1996). Testing ecological models: The meaning of validation. *Ecological Modelling*, 90(3), 229–244. [https://doi.org/10.1016/0304-3800\(95\)00152-2](https://doi.org/10.1016/0304-3800(95)00152-2)
- 945 Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., & Petrova, S. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences of the United States of America*, 108(24), 9899–9904. <https://doi.org/10.1073/pnas.1019576108>
- 950 Schnell, S., Kleinn, C., & Ståhl, G. (2015). Monitoring trees outside forests: A review. *Environmental Monitoring and Assessment*, 187(9), 600. <https://doi.org/10.1007/s10661-015-4817-7>
- Sileshi, G. W. (2014). A critical review of forest biomass estimation models, common mistakes and corrective measures. *Forest Ecology and Management*, 329, 237–254. <https://doi.org/10.1016/j.foreco.2014.06.026>
- Sileshi, G. W., Nath, A. J., & Kuyah, S. (2023). Allometric scaling and allocation patterns: Implications for predicting productivity across plant communities. *Frontiers in Forests and Global Change*, 5, 1084480. <https://doi.org/10.3389/ffgc.2022.1084480>
- 955 Skole, D. L., Mbow, C., Mugabowindekwe, M., Brandt, M. S., & Samek, J. H. (2021). Trees outside of forests as natural climate solutions. *Nature Climate Change*, 11(12), 1013–1016. <https://doi.org/10.1038/s41558-021-01230-3>
- Skole, D. L., Samek, J. H., Dieng, M., & Mbow, C. (2021). The contribution of trees outside of forests to landscape carbon and climate change mitigation in West Africa. *Forests*, 12(12), 1652. <https://doi.org/10.3390/f12121652>
- 960 Skole, D. L., Samek, J., Mehra, S., Bajaj, R., Tanmay, T., Suresh, S., Jindal, S., & Ndalowa, D. (2024). Measuring the extent of trees outside of forests: A nature-based solution for net zero emissions in South Asia. *Environmental Research Letters*, 19(11), 114092. <https://doi.org/10.1088/1748-9326/ad845d>
- Stanturf, J. A., Kleine, M., Mansourian, S., Parrotta, J. A., Madsen, P., Kant, P., Burns, J., & Bolte, A. (2019). Implementing forest landscape restoration under the Bonn Challenge: A systematic approach. *Annals of Forest Science*, 76, Article 50. <https://doi.org/10.1007/s13595-019-0833-z>
- Steyerberg, E. W., & Harrell, F. E., Jr. (2016). Prediction models need appropriate internal, internal-external, and external validation. *Journal of Clinical Epidemiology*, 69, 245–247. <https://doi.org/10.1016/j.jclinepi.2015.04.005>



- 970 Terasaki Hart, D. E., Yeo, S., Almaraz, M., Beillouin, D., Cardinael, R., Garcia, E., Kay, S., Lovell, S. T., Rosenstock, T. S.,
 Sprengle-Hyppolite, S., Stolle, F., Suber, M., Thapa, B., Wood, S., & Cook-Patton, S. C. (2023). Priority science can accelerate
 agroforestry as a natural climate solution. *Nature Climate Change*, 13(11), 1179–1190. <https://doi.org/10.1038/s41558-023-01810-5>
- 975 Thapa, B., Lovell, S., & Wilson, J. (2023). Remote sensing and machine learning applications for aboveground biomass
 estimation in agroforestry systems: A review. *Agroforestry Systems*, 97(6), 1097–1111. <https://doi.org/10.1007/s10457-023-00850-2>
- Tian, L., Wu, X., Tao, Y., Li, M., Qian, C., Liao, L., & Fu, W. (2023). Review of remote sensing-based methods for forest
 aboveground biomass estimation: Progress, challenges, and prospects. *Forests*, 14(6), 1086. <https://doi.org/10.3390/f14061086>
- 980 Tucker, C. J., Brandt, M., Hiernaux, P., Kariryaa, A., Rasmussen, K., Small, J., Igel, C., Reiner, F., Melocik, K., Meyer, J.,
 Sinno, S., Romero, E., Glennie, E., Fitts, Y., Morin, A., Pinzon, J., McClain, D., Morin, P., Porter, C., ... Fensholt, R. (2023).
 Sub-continental-scale carbon stocks of individual trees in African drylands. *Nature*, 615(7950), 80–86.
<https://doi.org/10.1038/s41586-022-05653-6>
- West, G. B., Enquist, B. J., & Brown, J. H. (2009). A general quantitative theory of forest structure and dynamics. *Proceedings
 of the National Academy of Sciences of the United States of America*, 106(17), 7040–7045.
<https://doi.org/10.1073/pnas.0812294106>
- 985 Zomer, R. J., Bossio, D. A., Trabucco, A., van Noordwijk, M., & Xu, J. (2022). Global carbon sequestration potential of
 agroforestry and increased tree cover on agricultural land. *Circular Agricultural Systems*, 2(1), 1–10.
<https://doi.org/10.48130/CAS-2022-0003>
- Zomer, R. J., Trabucco, A., Coe, R., & Place, F. (2009). *Trees on farm: Analysis of global extent and geographical patterns
 of agroforestry* (ICRAF Working Paper No. 89). World Agroforestry Centre. <https://doi.org/10.5716/WP16263.PDF>