https://doi.org/10.5194/essd-2026-102
Preprint. Discussion started: 18 March 2026

Earth System Science Data
Open Access
Discussions

FineKarstAGB: A 30 m resolution aboveground biomass dataset for Southwest China derived by upscaling plot-level inventory using sub-meter GaoFen satellite data

Yixiang Li[1,2], Yongqing Bai[1], Zhengchao Chen[1], Caixia Liu[1], Xuan Yang[1], Xiaowei Tong[3,4], Martin Brandt[5], Zhaoming Wu[1,2], Zeqing Wang[1,2], Huaming Gao[1,2], Xiaoyi Wang[6], Sizhuo Li[5], Xin Xu[3,5], and Siyu Liu[5]

[1]State Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha 410125, China
[4]Guangxi Key Laboratory of Karst Ecological Processes and Services, Huanjiang Observation and Research Station for Karst Ecosystems, Chinese Academy of Sciences, Huanjiang 547100, Guangxi, China
[5]Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark
[6]State Key Laboratory of Tibetan Plateau Earth System, Resources and Environment, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, China

Correspondence to: Yongqing Bai (baiyq@aircas.ac.cn)

Abstract. Southwest China has emerged as a key global carbon stock due to widespread forest expansion and aboveground biomass (AGB) increases driven by major ecological restoration since 2000, making accurate AGB estimations vital for assessing restoration efficacy. However, existing global and national-scale AGB products exhibit substantial limitations in this region, with little correlation with National Forest Inventory (NFI) plot and UAV LiDAR data, which is likely related to the pronounced spatial heterogeneity induced by Karst landscapes and large-scale restoration efforts that exacerbate mixed-pixel effects. To address these challenges, this study proposes a Canopy Structure-driven Multi-feature Fusion Network (CSMF-Net) designed for high-precision AGB estimation in complex regions. The method takes NFI plots data as ground truth and integrates GaoFen imagery, horizontal structure derived from tree crown segmentation and vertical structure represented by canopy height data. Based on this approach, we generated a fine-grained 30 m AGB dataset (FineKarstAGB) covering four provinces in Southwest China (Yunnan, Guizhou, Guangxi, and Hunan). Accuracy assessment against independent NFI plot data demonstrated the model's robust performance ($r = 0.83$, RMSE = 28.51 Mg/ha), showing no evidence of saturation in high-biomass regions. Furthermore, a structural consistency assessment using an independent UAV LiDAR-derived Canopy Height Model (CHM) confirmed that FineKarstAGB maintains high ecological consistency with the true forest vertical structure ($R^2 = 0.54$). Other public datasets show a weak correlation with both NFI ($r < 0.4$) and LiDAR data ($R^2 < 0.1$). Due to the tree-level segmentation, our dataset also quantifies AGB contributions from sparse trees outside forests, thus enabling more comprehensive and spatially explicit carbon accounting. This dataset provides critical support for regional carbon cycle assessments, fine-scale evaluations of ecological restoration outcomes, and progress toward national carbon neutrality targets. The dataset is available at https://doi.org/10.57760/sciencedb.33452 (Li et al., 2026).

https://doi.org/10.5194/essd-2026-102
Preprint. Discussion started: 18 March 2026

1 Introduction

Forests represent the largest carbon pool in terrestrial ecosystems and play a central role in regulating the global carbon cycle through long-term carbon sequestration in woody biomass (Pan et al., 2011; Cheng et al., 2024; Brandt et al., 2018; Chen et al., 2019). Accurate and spatially explicit estimates of forest aboveground biomass (AGB) are essential for quantifying terrestrial carbon stocks, supporting carbon cycle modeling, and informing climate change mitigation strategies at regional and global scales (Huang et al., 2013; Ma et al., 2024; Araza et al., 2022). High-quality AGB estimates provide crucial inputs for international reporting frameworks, including the Reducing Emissions from Deforestation and Degradation and Enhancement of Carbon Stocks program (REDD+), national greenhouse gas inventories, and the United Nations Sustainable Development Goals (SDGs), all of which rely on AGB datasets that are accurate, spatially detailed, and methodologically transparent to ensure reuse across ecological and policy applications (UNFCCC., 2013; Kashongwe et al., 2023; Zhang et al., 2023a).

Southwest China hosts the world's largest contiguous distribution of carbonate rocks and is also one of the regions facing the most intense human–land conflicts (Tong et al., 2018). However, long-term state-led ecological restoration programs in this region have demonstrated a significant carbon sink capacity and substantial potential (Tong et al., 2020). A series of initiatives targeting rocky desertification control and ecological restoration in these Karst areas, such as the Natural Forest Protection Program, the Grain for Green Program, integrated rocky desertification management, and the Yangtze River Shelterbelt Program, have fundamentally altered forest resources in the region (Tong et al., 2017; Trac et al., 2007), driving a sustained increase in AGB. Against this backdrop, developing accurate, high-spatial-resolution AGB datasets is particularly crucial for effectively evaluating the outcomes of these ecological restoration programs and for identifying and quantifying their ecological contributions. The vegetation in Southwest China exhibits an extremely complex, patchy, and fragmented distribution, as shown in Fig. 1. The landscape often features composite ecosystems like forest-grassland, agroforestry, and forest-shrubland systems. Numerous restoration projects where farmlands have been transformed to forests include natural regeneration and active tree plantations that often differ from field to field, creating a local patchwork of different forest systems. This complex structure makes it challenging for existing medium and low-resolution remote sensing imagery, and models built upon them, to accurately capture the region's fine-scale forest composition and high heterogeneity.

Methods for estimating aboveground biomass can be broadly classified into two main categories: field-based approaches and remote sensing–based methods. Field-based approaches typically involve destructive sampling or the application of allometric models that incorporate measurements such as tree height (H) and diameter at breast height (DBH) (Lu et al., 2016; Li et al., 2024b). Although these methods provide highly accurate and reliable estimates at local scales, their labor-intensive, time-consuming, and costly nature makes it challenging to upscale the plot-based data to heterogeneous regions, such as China Karst.

With the rapid advancement of remote sensing technology, mapping forest AGB across large spatial scales has become increasingly feasible (Yang et al., 2023; Yan et al., 2023b; Brovkina et al., 2017). The growing availability of diverse remote sensing datasets, such as optical imagery with high spatial, temporal, and spectral resolution (Zhu et al., 2025; Hu et al., 2020; Ma et al., 2023), an increasing number of spaceborne LiDAR missions (Duncanson et al., 2022; Markus et al., 2017), and

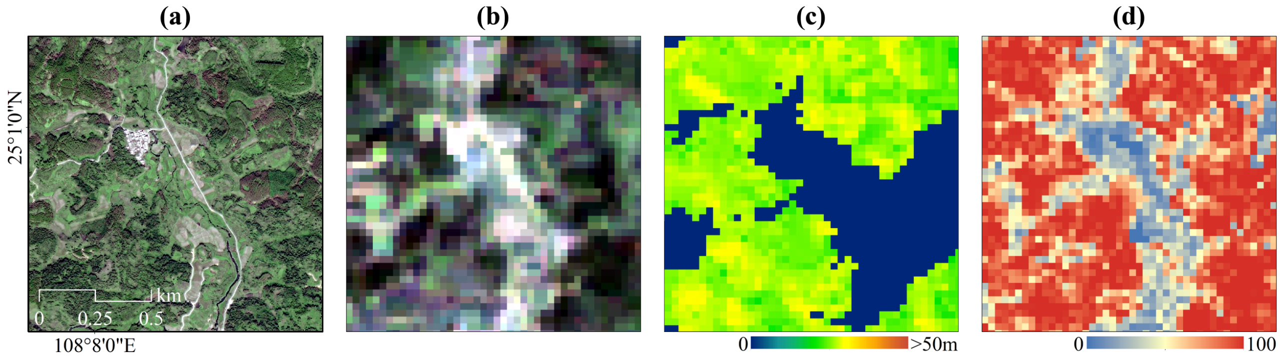


Figure 1. Comparison of remote sensing imagery and forest parameter products at different resolutions in Southwest China. (a) GaoFen imagery (0.8 m), (b) Landsat imagery (30 m), (c) Forest Canopy Height (Liu et al., 2022b) (30 m), (d) Tree Cover (Cai et al., 2024a) (30 m).

Synthetic Aperture Radar (SAR) observations (Lu et al., 2021) have greatly propelled the application of remote sensing in forestry and ecological monitoring. However, each remote sensing modality has inherent limitations. Optical sensors cannot penetrate forest canopies, often resulting in saturation effects in high-AGB regions (Mitchard et al., 2012). Although spaceborne LiDAR provides unique advantages for retrieving canopy height and AGB, its discrete sampling footprints and substantial geolocation uncertainties limit its large-scale applicability, in particular for heterogeneous landscapes. Consequently, multi-source data fusion has emerged as the mainstream approach for large-scale forest AGB estimation (Li et al., 2024a; Wang et al., 2023; Schwartz et al., 2023).

Numerous studies have explored the generation of high-resolution, large-scale forest AGB datasets in China (Chang et al., 2021; Huang et al., 2019; Su et al., 2016). These studies typically fuse field plot data with multi-source remote sensing information and conventionally incorporate auxiliary parameters as covariates to enhance model accuracy. For example, Yang et al. (2023) developed a forest AGB dataset for China for 2019 by combining field plot data, canopy height (Liu et al., 2022b) (Fig. 1d), and environmental variables (topography, climate, soil) within a Random Forest model. Zhang et al. (2023b) constructed a Random Forest model based on Landsat bands (Fig. 1b), vegetation indices, topography, and climate variables to produce a national forest AGB dataset for China spanning 2013–2021 (Yan et al., 2023a). Meanwhile, Cai et al. (2025) used GEDI data as a supervisory source, selected features from Landsat bands, tree cover (Cai et al., 2024a) (Fig. 1c), topography, and climate using Recursive Feature Elimination (RFE), and built a CNN model to generate a dynamic, annual forest AGB product for China covering 1985–2023 (Cai et al., 2024b).

Existing methods rely heavily on indirect environmental variables such as topography, climate, and soil in AGB estimation. Although these variables influence AGB distribution and can improve model performance, the relationships established through them often fail to directly reflect the contribution of trees' intrinsic structural attributes (e.g., individual tree height, crown width, DBH). Consequently, AGB patterns in these datasets may reflect environmental gradients rather than the actual distribution of forest structure. Furthermore, most datasets derived from these coarse-resolution and relatively static variables often fail to capture southern China's landscapes adequately and dynamics of AGB.

https://doi.org/10.5194/essd-2026-102
Preprint. Discussion started: 18 March 2026

Recent advances in remote sensing, particularly the availability of sub-meter satellite imagery (Zhang et al., 2023b), combined with the power of deep learning, offer unprecedented potential for detailed forest mapping and structural characterization. Deep learning models represented by U-Net (Ronneberger et al., 2015) and its variants are widely used for individual tree crown segmentation and counting, as well as canopy height and AGB estimation (Tong and Zhang, 2025; Li et al., 2025; Campbell et al., 2025; Wagner et al., 2024; Li et al., 2023b). For example, Li et al. (2023a) leveraged aerial imagery to train a U-Net model for nationwide individual tree crown mapping in Denmark, indicating that approximately 30% of the country's tree cover is contributed by trees outside forests. Extending to larger spatial scales, Liu et al. (2023) applied a U-Net–based approach to PlanetScope imagery to generate wall-to-wall canopy height estimates across Europe.

However, translating these capabilities into accurate AGB estimates, especially in structurally complex and highly heterogeneous landscapes like Southwest China's karst region, remains a significant hurdle. This raises several critical questions that current methodologies often fail to adequately address: (1) How can we overcome the inherent limitations of traditional pixel-based approaches that struggle with mixed pixels and fail to capture the fine-grained, patchy forest structure typical of karst environments? (2) Given that optical sub-meter data primarily capture spectral and textural information, how can we effectively integrate explicit three-dimensional structural information, both horizontal crown delineation and vertical height, to create more robust and accurate AGB models for these complex terrains? (3) Recognizing that significant biomass exists outside dense forest stands, how can we develop a methodology that reliably quantifies the AGB contribution of sparse trees, e.g., in agricultural lands or fragmented patches, which are often overlooked, to achieve more comprehensive regional carbon accounting? This study aims to answer these questions by developing a novel deep learning framework specifically designed for the challenges of AGB estimation in Southwest China.

2 Study area and data

2.1 Study area

The study area is located in Southwest China, encompassing the provinces of Yunnan, Guizhou, Guangxi and Hunan (20°–31°N, 97°–115°E, Fig. 2). Elevation ranges from near sea level in Guangxi to 6,740 m in Yunnan, reflecting substantial topographic variability. The region is characterized by karst landforms, including widespread peak–cluster depressions (Fig. 2c, d), trough valleys, basins, plateaus, and hills (Tong et al., 2017), which are associated with fragmented terrain, discontinuous surface hydrology, and strong spatial heterogeneity. The prevailing climate is subtropical monsoon with plateau and mountain influences, characterized by warm, humid conditions and abundant rainfall (Pei et al., 2022). Local microclimates and vertical differentiation in hydrothermal conditions provide favorable habitats for diverse vegetation types. Forests in the region account for over 20% of China's total forest area and include evergreen broadleaf forests, mixed evergreen–deciduous broadleaf forests, and a variety of plantations (e.g., Masson pine and eucalyptus) (Liu et al., 2022a; Qian et al., 2025). This unique combination of complex terrain and heterogeneous vegetation patterns provides contextual background for the spatial distribution of forest structure and remotely sensed forest variables.

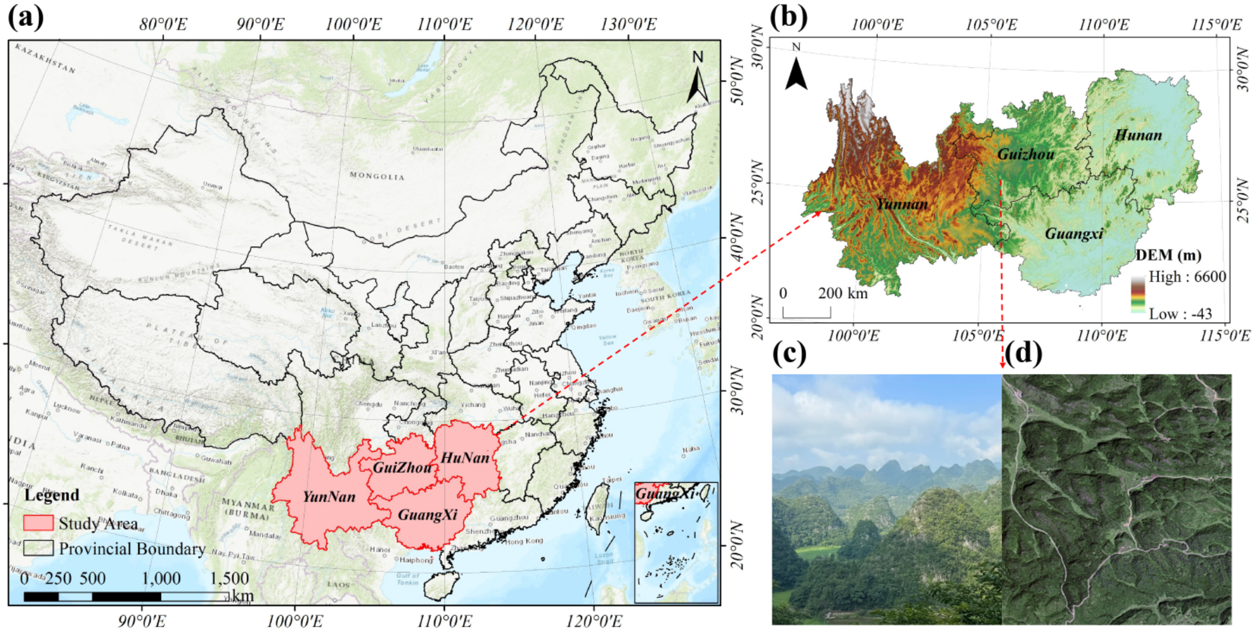


Figure 2. Overview and typical landscapes of the study area. (a) Location of the study area (red section; basemap: Esri, USGS, OpenStreetMap | powered by Esri), (b) Digital Elevation Model (DEM) of the study area derived from Copernicus DEM, (c) Field photo of the peak-cluster depression landscape in Southwest China, (d) The corresponding GaoFen imagery (RGB) of the same landscape.

2.2 Remote sensing data

2.2.1 Sub-meter GaoFen imagery

Sub-meter optical imagery covering Southwest China was collected from the Gaofen satellite constellation, including 10,330 GaoFen-2 (GF-2) scenes and 1,299 GaoFen-7 (GF-7) scenes. The panchromatic and multispectral (PMS) sensors onboard GF-2 provides a 0.8 m panchromatic band and four 3.2 m multispectral bands, while GF-7, as a stereo-mapping satellite, offers a similar spectral configurations with a 0.8 m panchromatic and four 3.2 m multispectral bands (Zhang et al., 2023b; Pan, 2015). The imagery was acquired between 2023 and 2024, with data from the summer of 2024 as the primary source and imagery from 2023 to fill spatial gaps. All raw images were processed using a comprehensive workflow, including dehazing and enhancement, geometric and radiometric correction, pan-sharpening (fusion of panchromatic and multispectral bands), cloud detection, mosaicking and tiling. The processed imagery was organized into tiles based on a hierarchical grid system analogous to the Google Maps zoom level framework. The final dataset consists of 13,648 image tiles capturing fine-scale surface heterogeneity across the study area. An overview mosaic is presented in Fig. S1. This imagery dataset served as the primary input for deriving the tree crown segmentation and canopy height datasets described in the following sections.

2.2.2 Tree crown segmentation dataset

To characterize the highly heterogeneous patterns of tree distribution across the study area, this study used a tree crown segmentation dataset derived from sub-meter GaoFen imagery and a U-Net model. The dataset originates from previous tree mapping efforts, in which Xu et al. (2026) delineated more than 5 billion trees in Guangxi Province with an overall R^2 of 0.82. Using the same methodological framework, the tree crown segmentation was extended to Southwest China, resulting in more than 21 billion mapped trees with a crown size larger than 0.64 m^2 (equivalent to one pixel) across four provinces, including approximately 5.2 billion trees in Guangxi, consistent with the estimates reported by Xu et al. For this study, the tree crown dataset provides tree–scale information that supports the analysis of highly heterogeneous forest patches and the spatial distribution of trees outside forests (TOF). The tree counts for Southwest China are summarized in Table S1.

2.2.3 Canopy height dataset

Canopy height is a key parameter of forest vertical structure and has been demonstrated to play an important role in large-scale AGB estimation (Yang et al., 2023; Liu et al., 2023). However, existing global or national canopy height datasets have been shown to exhibit relatively low correlations with airborne LiDAR canopy height model (CHM) within the study area. Therefore, this study applies a new canopy height dataset with a spatial resolution of 0.8 m, which was produced using 200,000 km^2 of airborne LiDAR from the study area, and the same sub-meter GaoFen imagery used in this study (Tong et al., 2025). The dataset achieved an accuracy of $r = 0.87$, MAE = 1.94 m, and RMSE = 2.76 m on an independent validation dataset, and no evident saturation was observed in the higher canopy height range. Overall, the errors are substantially lower than those reported for previously published canopy height products ($r \leq 0.22$). In addition, further accuracy assessment was conducted using independent UAV LiDAR data (described in Sect. 2.3.2), yielding results ($r = 0.82$, MAE = 1.50 m, RMSE = 2.82 m) consistent with those reported by Tong et al. The corresponding validation results are shown in Fig. S2. This canopy height dataset provides high-resolution information on forest vertical structure across Southwest China and serves as a critical explanatory variable for large-scale AGB inversion.

2.3 Reference and elevation Data

2.3.1 National Forest Inventory (NFI) data

Ground-truth reference data were derived from China's 9th National Forest Inventory (NFI) (Jian et al., 2022). As an integral component of the national forest resource monitoring system, the NFI provides forest metrics, including forest area, stock volume, and AGB, based on systematically repeated field surveys conducted at fixed sample plots nationwide at five-year intervals. The NFI plots used in this study encompass six provinces—Yunnan, Guizhou, Guangxi, Guangdong, Hunan, and Hubei, which collectively capture the complex topographic gradients and ecological diversity characteristic of Southwest China. NFI plots are typically square in shape with an area of 0.0667 ha (1 mu). Plot-level AGB is estimated from measured stand volume using biomass expansion factors (BEF) specific to tree species and age classes. The temporal coverage of the dataset spans 2014–2018; however, detailed timestamp information for individual plots is unavailable.

The temporal gap of up to one decade between the NFI data and the remote sensing imagery used in this study introduces potential uncertainty (Yang et al., 2023). In Southwest China, government-led land-use policies and large-scale ecological restoration initiatives, such as afforestation and reforestation programs, have driven extensive forest regeneration and expansion. Consequently, the region is dominated by young to middle-aged, fast-growing forests, where AGB exhibits rapid and substantial decadal dynamics (Tong et al., 2020). Under these conditions, direct use of temporally misaligned datasets for model training or validation may affect the accuracy of AGB estimation.

To mitigate the potential impacts of this temporal inconsistency, we designed an automated and systematic selection workflow to ensure that only matching training samples were used. First, an initial data cleaning step removed plots with incomplete records, missing AGB values, zero or negative AGB, or zero or negative mean tree height. Second, three simple features strongly related to AGB were derived from the high-resolution imagery for the remaining plots: mean NDVI, crown coverage, and mean canopy height. These features capture key spectral and structural information directly relevant to AGB while minimizing potential biases associated with more complex feature representations. Third, to identify and filter plots exhibiting strong inconsistencies between ground measurements and imagery, a low-parameter Random Forest model was trained using the aforementioned features. Plots with relative prediction errors exceeding 50% were excluded from subsequent analyses. This procedure aimed to reduce the influence of temporally inconsistent samples while maintaining the representativeness and diversity of the AGB value distribution. The final set of selected plots (n = 3,637) was used for model training (n = 2,545), validation (n = 364) and testing (n = 728), and their distribution is shown in Fig. S3.

2.3.2 UAV LiDAR data

The evaluation CHM data used in this study were obtained from the LIDARNET (http://lidar.pku.edu.cn) and covered typical forest scenes in Guangxi Province. The LiDAR point clouds were acquired using Riegl UVX-1, with an average point cloud density exceeding 10 $\mathrm{pts/m^2}$, meeting the requirements for depicting forest canopy structure. The raw UAV LiDAR data underwent a standardized processing pipeline, including digital terrain modeling, canopy height extraction, and plot-scale aggregation (i.e., aggregating canopy height by calculating the mean, maximum and 95th percentile, from AGB-aligned pixel with each plot). We examined all CHM samples and excluded those exhibiting non-natural growth-related changes (e.g., logging). In total, 3,982 CHM validation samples (3,591 in forests and 391 outside of forest) were obtained and used to evaluate the AGB estimates at the plot level. As shown in Figs. 7, 8, only samples with available AGB predictions were included in the validation, which explains the variation in sample size (n) among datasets. This variation is primarily related to differences in spatial coverage and data availability across the compared datasets.

https://doi.org/10.5194/essd-2026-102
Preprint. Discussion started: 18 March 2026

2.4 Comparison AGB datasets

To compare the performance of the AGB dataset developed in this study with existing datasets, we selected four global and national-scale AGB datasets with different spatial resolutions (Table 1). For each dataset, the year closest to this study was chosen to ensure temporal comparability.

1. **Yang Map**: This dataset integrates 4,789 national ground plot measurements of AGB with multi-source remote sensing data, including forest canopy height products, optical spectral indices, and topographic and climatic variables. A Random Forest regression model was employed to generate the AGB estimates. The resulting product provides a 30 m spatial resolution map of China's forest AGB for 2019 (Yang et al., 2023).

2. **Yan Map**: This approach utilized GaoFen and Landsat satellite data combined with a Random Forest model trained on 6,667 ground plots. Specifically, GF-1 and GF-2 imagery were used to map forest cover, while time-series Landsat data informed forest AGB estimation. The study aimed to assess China's forest status and trends in support of Sustainable Development Goal (SDG) evaluations (Zhang et al., 2023a). The dataset provides annual 30 m resolution maps of China's forest AGB from 2015 to 2021 (Yan et al., 2023a).

3. **ESA CCI**: The ESA CCI AGB dataset is a global-scale dataset generated using the BIOMASAR semi-empirical algorithm that estimates AGB from C- and L-band SAR data. LiDAR canopy height, external AGB statistics and land cover data are used to calibrate the model. The dataset provides global forest AGB maps at 100 m resolution for the years 2007, 2010 and 2015–2022 (Santoro and Cartus, 2025).

4. **Cai Map (CFATD)**: This dataset employed a deep learning framework combining multi-source remote sensing inputs with ResNet architectures. The model was trained on over 50,000 multi-temporal GEDI-derived AGB training samples and independently validated using multi-year field survey data (Cai et al., 2025). The resulting product delivers annual 30 m resolution forest AGB maps for China from 1985 to 2023 (Cai et al., 2024b).

Table 1. Overview of comparison AGB datasets

Dataset Name	Region	Temporal Scope	Spatial Resolution (m)	Comparison Year	Source
Yang Map	China	2019	30	2019	(Yang et al., 2023)
Yan Map	China	2013-2021	30	2021	(Yan et al., 2023a)
ESA CCI	Global	2007, 2010, 2015-2022	100	2022	(Santoro and Cartus, 2025)
Cai Map	China	1985-2023	30	2023	(Cai et al., 2024b)

3 Methods

To enable accurate AGB estimation in the complex terrain and highly heterogeneous landscapes of Southwest China, this study developed an AGB estimation framework that integrates sub-meter remote sensing imagery with tree crown segmentation and canopy height datasets. The overall workflow is illustrated in Fig. 3 and comprises four main stages:

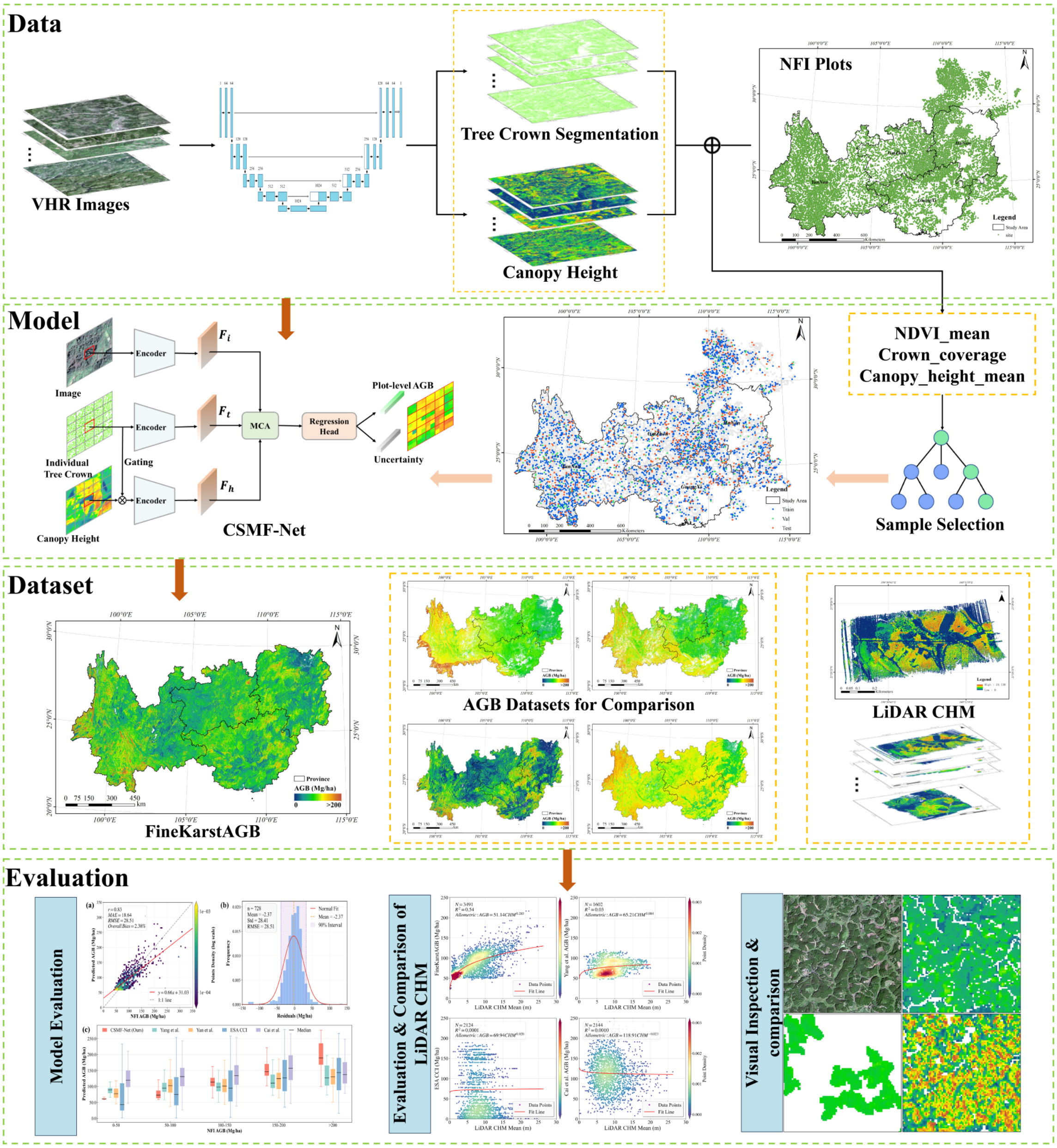


Figure 3. Workflow of plot-level aboveground biomass estimation.

(1) Data and Sample Processing: Remote sensing features were constructed for the NFI plots by combining sub-meter GaoFen imagery with tree crown segmentation and canopy height datasets. Subsequently, a Random Forest model was employed to perform quality control and filter the training samples, ensuring their representativeness and reliability. (2) Model Development: We developed an end-to-end, multi-feature fusion network capable of integrating multi-source remote sensing features with both horizontal and vertical forest structural information. The network regresses plot-level AGB and quantifies its associated uncertainty. (3) AGB Dataset Generation for Southwest China: The trained network was applied across the entire study area to produce a fine-grained AGB dataset for Southwest China. (4) Accuracy Assessment and Comparative Evaluation: Model performance was evaluated using an independent test set to assess accuracy. Further comparative evaluation was conducted, in which the generated AGB dataset and existing global and national-scale AGB datasets were compared against independent UAV LiDAR-derived CHM data.

3.1 Canopy Structure-driven Multi-feature Fusion Network

As illustrated in Fig. 4, this study proposes a Canopy Structure-driven Multi-feature Fusion Network (CSMF-Net) for plot-level aboveground biomass regression. The central concept of CSMF-Net is to achieve high-precision AGB estimation by effectively integrating sub-meter remote sensing imagery with tree crown and canopy height datasets, thereby fusing the forest's spectral information with its horizontal and vertical structural characteristics.

CSMF-Net adopts a *Patch to Point* regression strategy, in which a preprocessed plot-level (NFI) data patch is used as input to directly predict a single AGB value and its associated uncertainty for that plot. Unlike conventional pixel-wise regression approaches (e.g., Random Forest applied to 30 m Landsat pixels), this strategy explicitly models the spatial context and internal heterogeneity within each patch rather than treating pixels as independent samples.

In karst landscapes, vegetation distribution is extremely fragmented due to complex terrain, shallow soils, and strong lithological controls, causing individual medium-resolution pixels to frequently represent mixtures of vegetation, bare rock, and soil. Pixel-level models trained on such mixed signals tend to smooth local variability and are highly sensitive to sub-pixel composition. In contrast, the patch-to-point framework aggregates multi-pixel spectral–structural patterns and learns their joint relationship with plot-level AGB, thereby reducing the influence of mixed pixels and sub-pixel noise. As a result, CSMF-Net enables reliable AGB prediction at the plot level for areas with tree cover, facilitating a more comprehensive representation of the spatial distribution and variability of aboveground biomass across complex terrains.

3.1.1 CSMF-Net model architecture

The overall architecture of the proposed Canopy Structure-driven Multi-feature Fusion Network (CSMF-Net) is illustrated in Fig. 4. It consists of a multi-branch feature extractor, a multi-feature cross-attention module, and a regression head.

First, in the feature extraction stage, the GaoFen optical imagery, tree crown data, and canopy height data for a given plot patch are processed through three parallel encoder branches. A ResNet backbone is employed in each branch to extract deep features from the corresponding modality. Furthermore, prior to feature extraction on the canopy height data, a gating mechanism is applied. This mechanism performs pixel-wise multiplication of the canopy height data with the tree crown

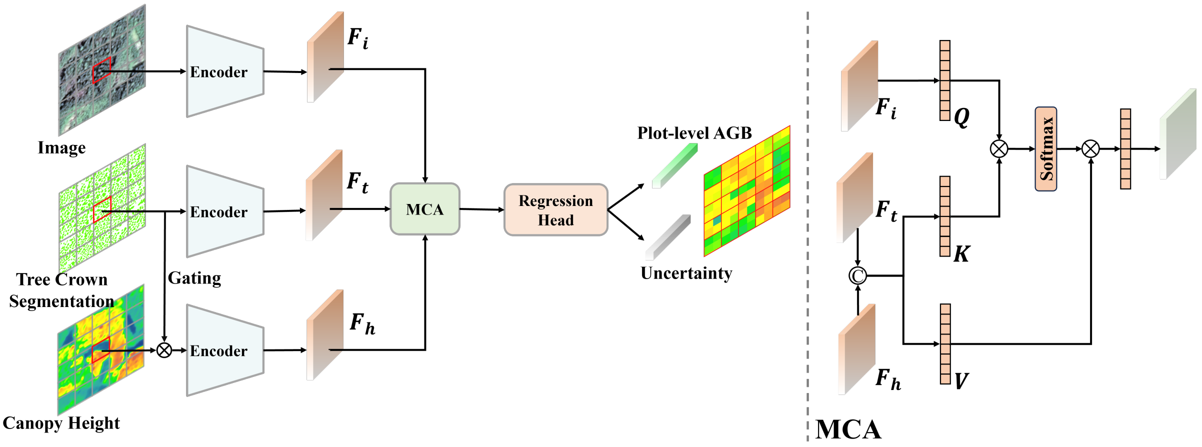


Figure 4. Architecture of the Canopy Structure-driven Multi-feature Fusion Network (CSMF-Net) and the Multi-feature Cross-Attention Module.

segmentation mask, effectively suppressing background interference and enabling the network to concentrate on the vertical structural information of the canopy.

Subsequently, the extracted features (F_i, F_t, F_h) from the different branches are input into the core Multi-feature Cross-Attention (MCA) module. This module is key to achieving deep fusion of cross-feature information within CSMF-Net. It adaptively learns the interdependencies between spectral and structural features to generate a fused feature representation that is richer in information and more expressive.

Finally, the fused feature is input into a regression head, which consists of multiple fully connected layers. This head outputs the plot-level AGB along with its corresponding heteroscedastic uncertainty.

3.1.2 Multi-feature Cross-attention module

To achieve effective interaction and informational complementarity between spectral and structural information, we designed a Multi-feature Cross-Attention (MCA) module, the detailed structure of which is illustrated on the right side of Fig. 4.

The features from the sub-meter GaoFen imagery provide spatial context, while the tree crown and canopy height features explicitly capture the horizontal and vertical structural attributes of the forest. Therefore, within the MCA module, we designated the imagery feature as the Query (Q), leveraging its global contextual information to aggregate relevant information from the other structural features. Simultaneously, we concatenated the canopy height feature with the tree crown feature, followed by separate linear projections to derive the Key (K) and the Value (V).

The cross-attention operation is formally expressed mathematically as follows:

$$Q = F_i W_Q \tag{1}$$

$$K = \text{Concat}\left(F_h, F_t\right) W_K \tag{2}$$

$$V = \text{Concat}\left(F_h, F_t\right) W_V \tag{3}$$

Where W_Q, W_K, and W_V are learnable weight matrices used for feature projection. The attention weights are computed by evaluating the similarity between the Query and the Key, followed by normalization using the Softmax function. These weights are then applied to the Value to generate the fused feature representation.

$$F_{fused} = \text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V \tag{4}$$

Where d_k is the dimension of the Key, used for scaling to stabilize the gradients during training.

3.1.3 Loss function

To jointly optimize the model's regression accuracy and its predictive uncertainty, we adopted a composite loss function comprising two components: the R^2 loss (L_{R^2}) and the heteroscedastic uncertainty loss (L_{unc}).

$$L_{total} = L_{R^2} + \lambda L_{unc} \tag{5}$$

Where λ is a hyperparameter that balances the contributions of the two loss components. In this study, we set λ to 0.1 based on empirical experiments.

The R^2 loss is formulated to directly optimize the model's goodness-of-fit for AGB by maximizing the R^2 between the predicted and true values. It is defined as follows:

$$L_{R^2} = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{6}$$

Where N is the number of samples in a batch, y_i and $\hat{y}_i$ represent the true and predicted AGB values for the i^{th} sample, respectively, and $\bar{y}$ is the mean of the true AGB values.

The heteroscedastic uncertainty loss is formulated to capture and explain the discrepancies between the imagery and the plot data, enabling the model to learn the prediction uncertainty associated with the input data. For samples where the model exhibits low prediction confidence, their contribution to the overall loss is proportionally reduced, thereby mitigating the negative impact of noisy or anomalous samples on model training. It is defined as follows:

$$L_{unc} = \frac{1}{N}\sum_{i=1}^{N}\left(e^{-\log\widehat{\sigma_i^2}} \cdot \left(y_i - \widehat{y_i}\right)^2 + \log\widehat{\sigma_i^2}\right) \tag{7}$$

The first term of this loss function is a Mean Squared Error (MSE) term, which is modulated by an adaptive weight; the second term is a regularization term.

3.1.4 Model training and implementation details

The model was implemented using the PyTorch deep learning framework. During training, we employed the AdamW optimizer (Loshchilov and Hutter, 2019), which effectively enhances the model's generalization capability. The initial learning rate was set to $1\mathrm{e}{-4}$, and a learning rate decay strategy was applied to achieve finer parameter adjustments in the later stages of training. Considering both computational constraints and convergence efficiency, the batch size was set to 32. All experiments were conducted on an NVIDIA RTX 3090 GPU.

3.2 Accuracy assessment and cross-dataset comparison

To evaluate the fine-scale AGB dataset developed for Southwest China in this study, we established an integrated assessment framework consisting of three components. (1) An independent test set derived from NFI plots was used to quantitatively evaluate model performance and dataset accuracy. (2) Independent UAV LiDAR–derived CHM data were employed to assess the reliability of the AGB dataset and its consistency with forest vertical structure. (3) Visual inspection and comparisons with representative existing AGB datasets were conducted over typical tree-covered landscapes to examine differences in spatial detail and structural representation among datasets. To ensure objectivity and comparability, the quantitative evaluation approaches described in (1) and (2) were also applied to the comparison datasets.

Four commonly used statistical metrics were employed to evaluate the consistency between the predicted AGB values and the ground plot reference values: the coefficient of determination (R^2), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Overall Bias.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \tag{8}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \widehat{y_i}\right| \tag{9}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i - \widehat{y_i}\right)^2}{N}} \tag{10}$$

$$Overall\ Bias = \frac{\left|\sum_{i=1}^{N}\left(\widehat{y_i} - y_i\right)\right|}{\left|\sum_{i=1}^{N} y_i\right|} \tag{11}$$

Where N is the total number of samples, y_i and $\hat{y}_i$ represent the true and predicted AGB values for the i^{th} sample, respectively, and $\bar{y}$ is the mean of the true AGB values.

4 Results

4.1 Evaluation with NFI data

Model performance was evaluated using an independent subset of 728 NFI plots (20% of the total plots), which were not used during the training and validation process or for hyperparameter optimization. The evaluation results are shown in Fig. 5. The CSMF-Net predictions exhibit strong agreement with the independent NFI observations, achieving correlation coefficient $r = 0.83$, MAE = 18.64 Mg/ha, RMSE = 28.51 Mg/ha and an Overall Bias of 2.38% (Fig. 5a). The residuals are approximately normally distributed, centered at -2.37 Mg/ha (σ = 28.41 Mg/ha), with over 90% of the absolute residuals falling below 40 Mg/ha, suggesting largely homoscedastic errors and no systematic regional bias across heterogeneous terrain (Fig. 5b).

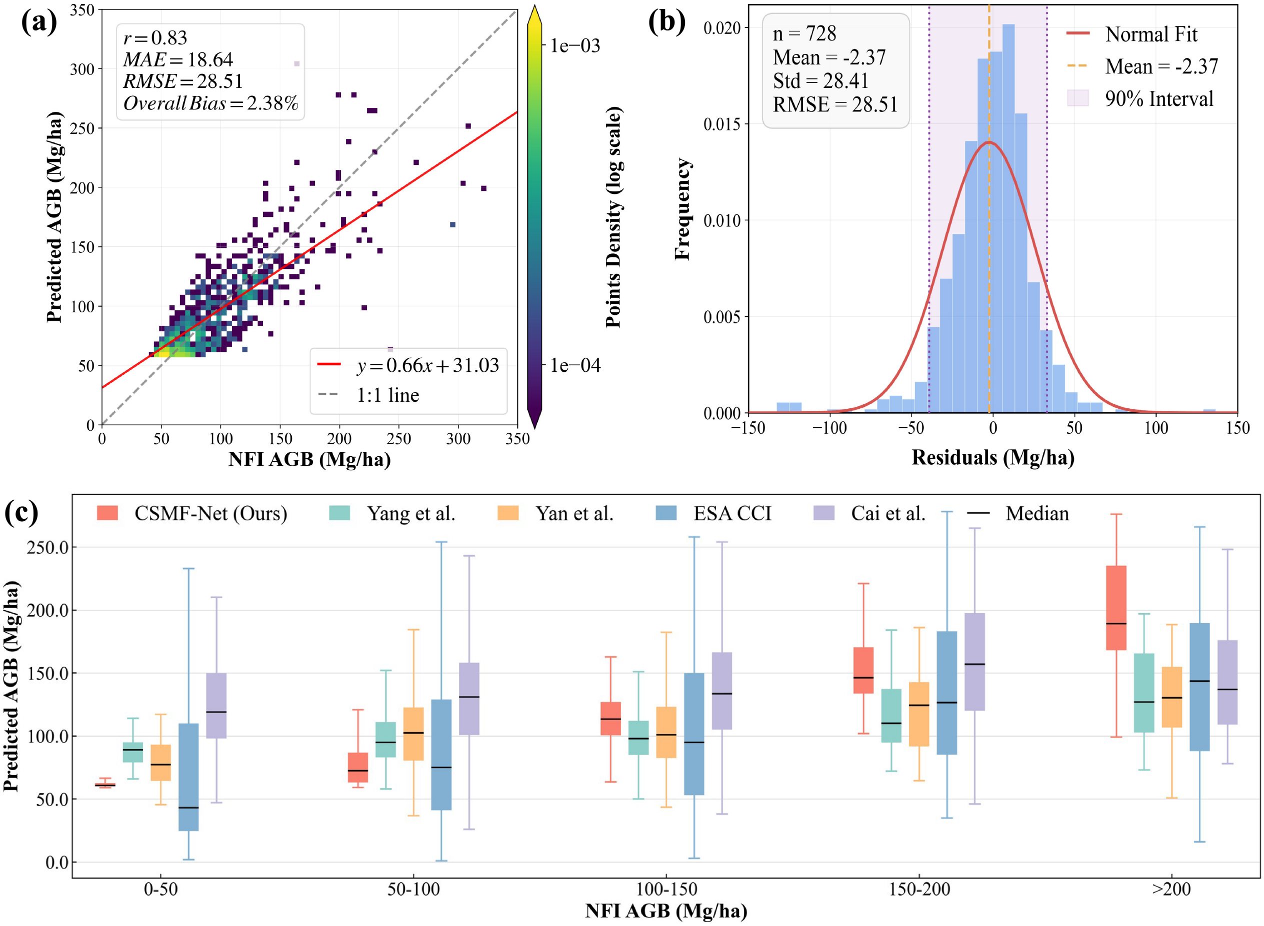


Figure 5. Evaluation of the CSMF-Net against independent NFI plots (n=728). (a) Scatter plot of predicted versus observed AGB for independent NFI test plots. (b) Distribution of residuals (predicted AGB minus observed AGB) for the NFI test plots. (c) Binned analysis of the AGB estimated by CSMF-Net (Ours) with the existing datasets.

To further evaluate model performance and uncertainty, we grouped the NFI plots in different classes of AGB ranges and CSMF-Net estimates were compared with multiple established AGB datasets across the classes (Fig. 5c). Overall, CSMF-Net shows reduced predictive uncertainty, with narrower interquartile ranges than existing datasets. Notably, CSMF-Net demonstrates a better capability in capturing both high and low values. In high-AGB regions (> 200 Mg/ha), CSMF-Net achieves a median prediction of approximately 190 Mg/ha, which is 50–60 Mg/ha higher than that of conventional products, highlighting its ability to retain high-AGB signals without saturation effects. At 150–200 Mg/ha AGB levels, CSMF-Net maintains a median of around 145 Mg/ha, aligning well with NFI observations, whereas other datasets underestimate field data by 25–30 Mg/ha. In lower AGB areas (50–100 Mg/ha), CSMF-Net exhibits a median of roughly 70 Mg/ha with an IQR < 25 Mg/ha, effectively suppressing high-noise fluctuations observed in products such as ESA CCI (IQR of about 90 Mg/ha). These results indicate that our map provides both accurate and stable AGB predictions across the full AGB range, offering a robust, uncertainty-aware basis for large-scale AGB mapping in the study area.

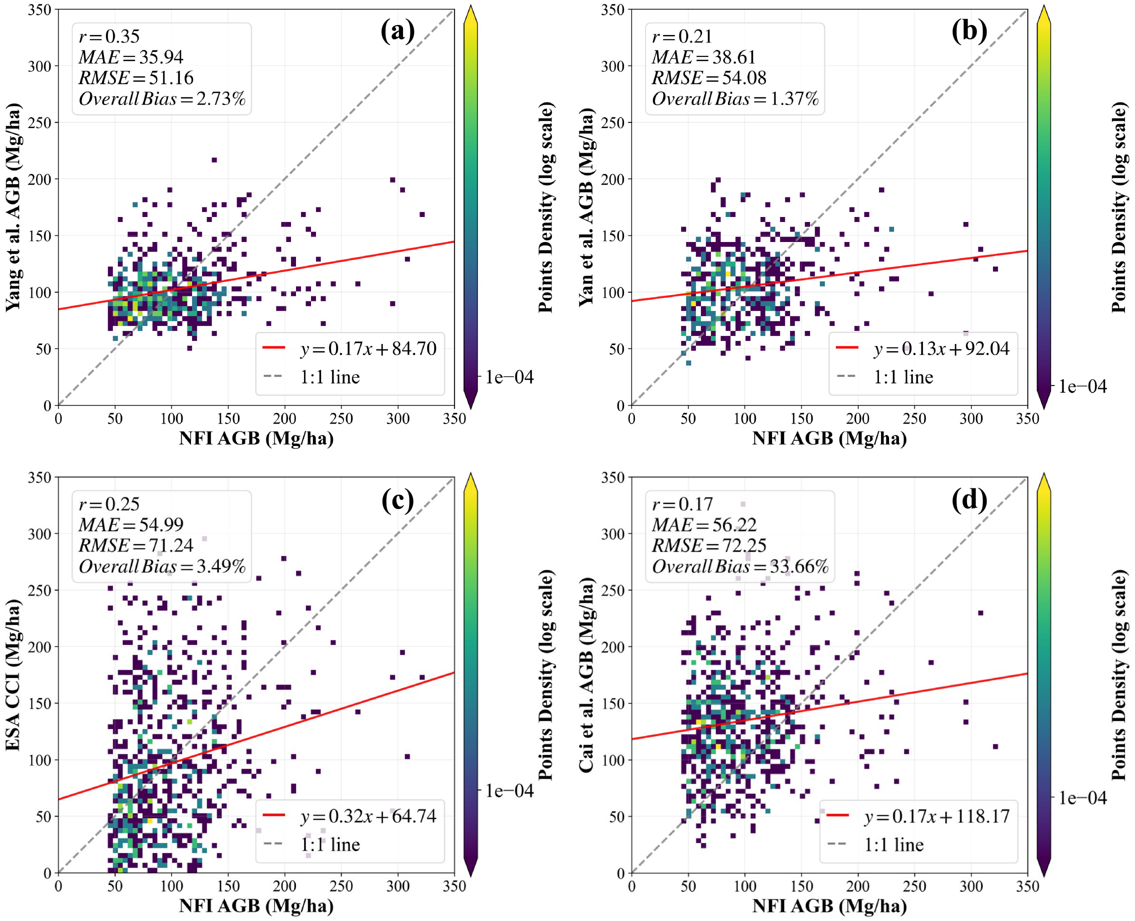


Figure 6. Comparison of the existing datasets with NFI data. (a) Yang et al. (2023), (b) Yan et al. (2023a), (c) ESA CCI and (d) Cai et al. (2024b).

The evaluation results for the comparison datasets are presented in Fig. 6. Existing global- or national-scale AGB products generally exhibit weaker performance in Southwest China, with relatively low correlations with NFI plot data ($r = 0.17$–0.35) and RMSE values ranging from 51.16 to 72.25 Mg/ha. Among them, the dataset developed by Cai et al. shows a pronounced systematic bias (Overall Bias = 33.66%). Moreover, linear regression analyses reveal consistently low slopes (0.13–0.32) for the comparison datasets, accompanied by substantial underestimation in high-AGB regions (> 200 Mg/ha), indicating varying degrees of saturation effects.

4.2 Evaluation with independent UAV LiDAR CHM data

The relationships between UAV LiDAR-derived canopy height model (CHM) metrics and the AGB estimates produced in this study (Fig. 7 for mean CHM, Fig. S4 for maximum CHM, and Fig. S5 for the 95th percentile CHM; Fitted using the allometric relationship $AGB = aH^b$) provide further evidence for the reliability of the proposed AGB dataset (FineKarstAGB). As shown in Fig. 7a, our AGB dataset exhibits strong structural consistency with CHM data at the plot level ($R^2 = 0.54$). Predicted AGB increases monotonically across the full canopy height spectrum, without an evident saturation effect at higher canopy height (> 20 m), indicating that our dataset effectively captures the physical relationship between forest canopy height and AGB in this complex terrain region.

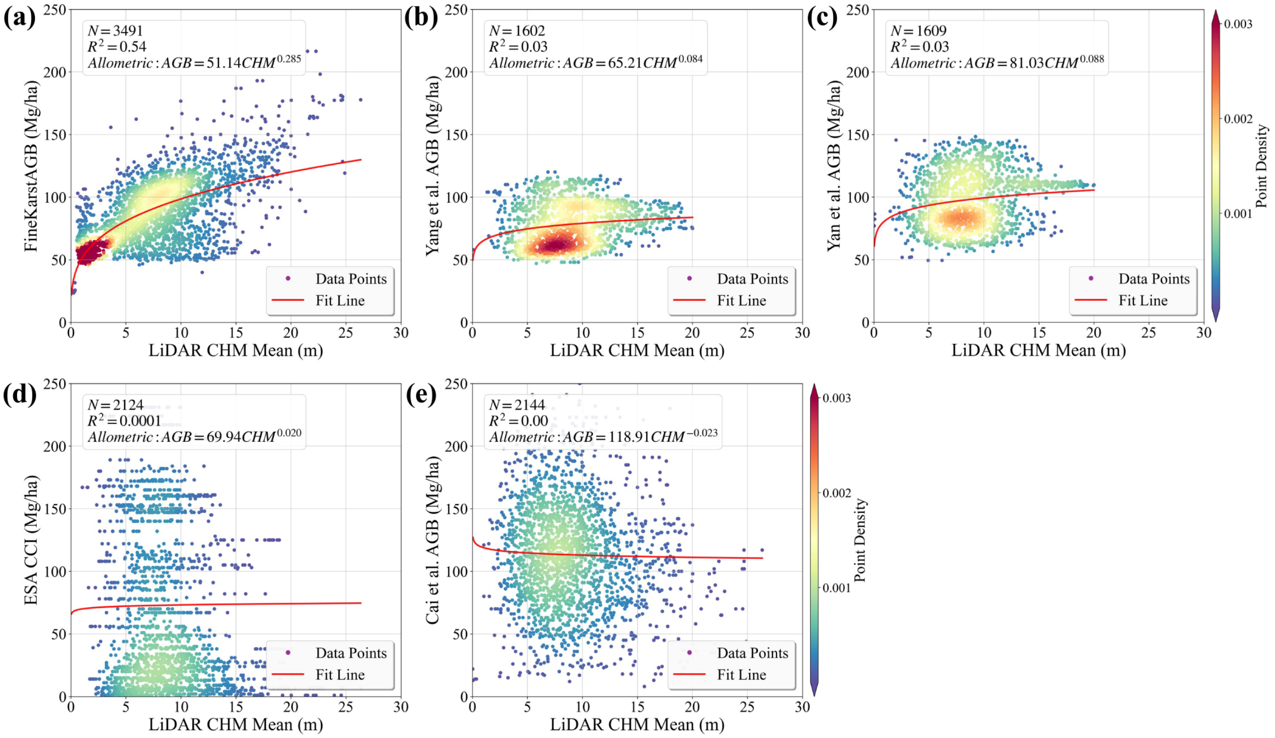


Figure 7. Evaluation of our AGB dataset comparing it against UAV LiDAR CHM data. Results are shown for (a) FineKarstAGB (Ours), (b) Yang et al. (2023), (c) Yan et al. (2023a), (d) ESA CCI and (e) Cai et al. (2024b).

In comparison, the relationships between LiDAR CHM and existing AGB datasets are generally weaker and less distinct. Both Yang et al. (2023) and Yan et al. (2023a) AGB datasets show very limited explanatory power, with R^2 values of 0.03, and low scaling exponents (b = 0.084 and 0.088), accompanied by broader dispersion and flatter fitted curves, suggesting reduced sensitivity of the estimated AGB to variations in canopy height. The ESA CCI AGB product exhibits the lowest correlation with LiDAR CHM ($R^2 = 0.0001$), which may partly be attributed to its coarser spatial resolution (100 m), potentially diluting height–AGB relationships at the plot scale. The Cai et al. (2024b) dataset shows a weak negative correlation with LiDAR CHM ($R^2 = 0.001$) and substantial dispersion across the CHM range, indicating that taller stands are not systematically associated with higher predicted AGB.

We further evaluated our dataset for areas outside contiguous forests (n = 391). Considering that most existing AGB products apply strict forest masks during their mapping procedures, while the dataset of Cai et al. (2024b) defines valid AGB values only for areas with tree cover exceeding 20%, this comparison was limited to the Cai et al. dataset. As illustrated in Fig. 8, the FineKarstAGB dataset maintains a significant positive relationship with LiDAR CHM in areas outside forests ($R^2 = 0.22$, $P < 0.0001$). Although scatter is greater than that observed in forests, AGB generally increases with canopy height. Given the higher spatial heterogeneity and structural complexity of trees outside forests, the results indicate that FineKarstAGB remains sensitive to variations in tree vertical structure in non-forest environments, supporting its applicability beyond contiguous forest regions. In contrast, the Cai et al. dataset exhibits only a weak negative relationship with LiDAR CHM in non-forest scenes ($R^2 = 0.02$), with highly scattered points and no clear trend in AGB with increasing canopy height.

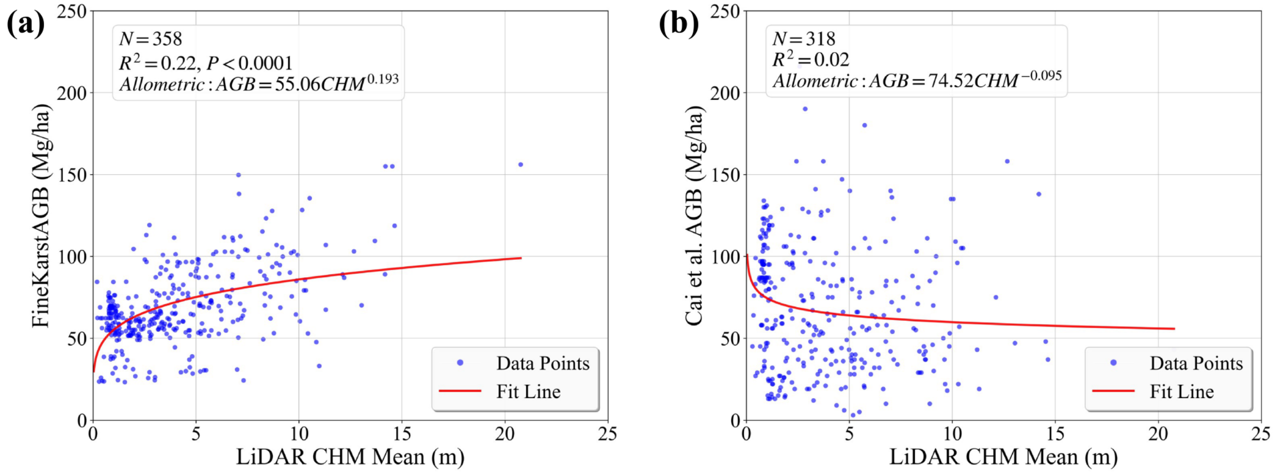


Figure 8. Evaluation of our dataset against LiDAR CHM data in non-forest areas. (a) results of FineKarstAGB (Ours), (b) results from Cai et al. (2024b).

4.3 Spatial pattern of AGB in Southwest China and typical scenarios

The high-spatial-resolution AGB map of Southwest China (Fig. 9a) reveals a distinct pattern of spatial heterogeneity. In general, high AGB density is predominantly concentrated in western and southern Yunnan, southeastern Guizhou, northern Guangxi, and western Hunan. Visual inspection using GaoFen imagery indicates that these high-value clusters correspond to mountainous regions characterized by rugged terrain, mature natural forests, and high canopy closure. Conversely, low AGB values exhibit a fragmented distribution, appearing sporadically in agro-forestry ecotones, peri-urban zones, and recent logging sites, reflecting areas with intense anthropogenic disturbance. At the provincial scale, spatial patterns vary significantly: mountainous regions in Yunnan and Hunan display continuous high-AGB clusters, whereas Guangxi and Guizhou are characterized by moderate AGB values with scattered high-value patches.

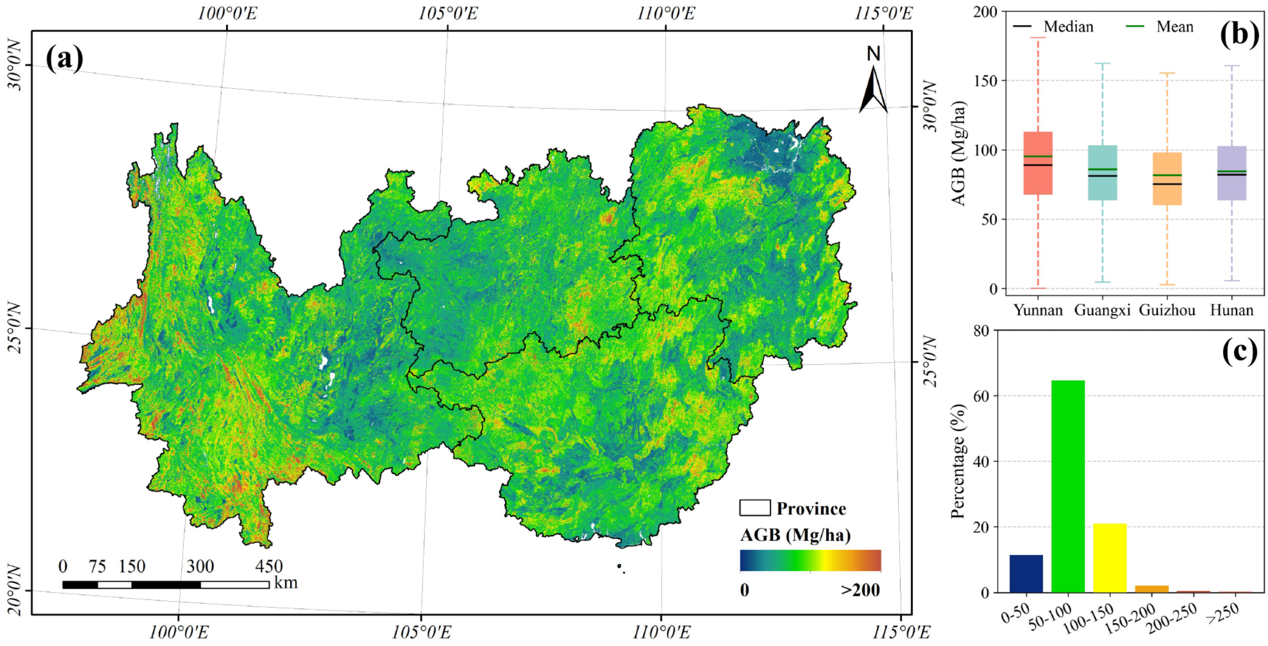


Figure 9. Spatial pattern and statistical characteristics of AGB in Southwest China. (a) Spatial distribution of AGB, (b) Boxplots of provincial AGB values for the four provinces, (c) Pixel-based frequency distribution of AGB for the study area.

Provincial statistics further quantify the spatial discrepancies (Fig. 9b). The mean AGB for the four provinces is 88.22 Mg/ha, with Yunnan exhibiting the highest average (95.32 Mg/ha) and Guizhou the lowest (81.59 Mg/ha). Notably, our estimate for Yunnan is highly consistent with the reference value of 95.92 Mg/ha, derived from the forest stock volume and area data of the Yunnan Provincial Results of the *National Forest, Grassland, Wetland and Desertification Survey (2025)* using the conversion method by Guo et al. (2013). The broad ranges observed in the boxplots indicate substantial intra-provincial variability, particularly in Yunnan and Hunan. This variability reflects the complex spatial heterogeneity driven by the Karst topography and diverse forest types inherent to the region.

Pixel-level frequency statistics (Fig. 9c) shows that the AGB in the study area is primarily concentrated in the 50–100 Mg/ha range, accounting for 64.6% of the total area. This dominance suggests a prevalence of young to middle-aged forests, likely attributable to active forest management and long-term ecological restoration projects (Tong et al., 2020). The 100–150 Mg/ha interval follows, comprising a substantial proportion (20.1%) of moderate-to-high biomass forests. Areas with AGB < 50 Mg/ha constitute a smaller fraction, primarily corresponding to sparse woodlands and trees outside forests. Although regions exceeding 150 Mg/ha (and those > 250 Mg/ha) cover a limited area, they contribute disproportionately to the regional carbon stock.

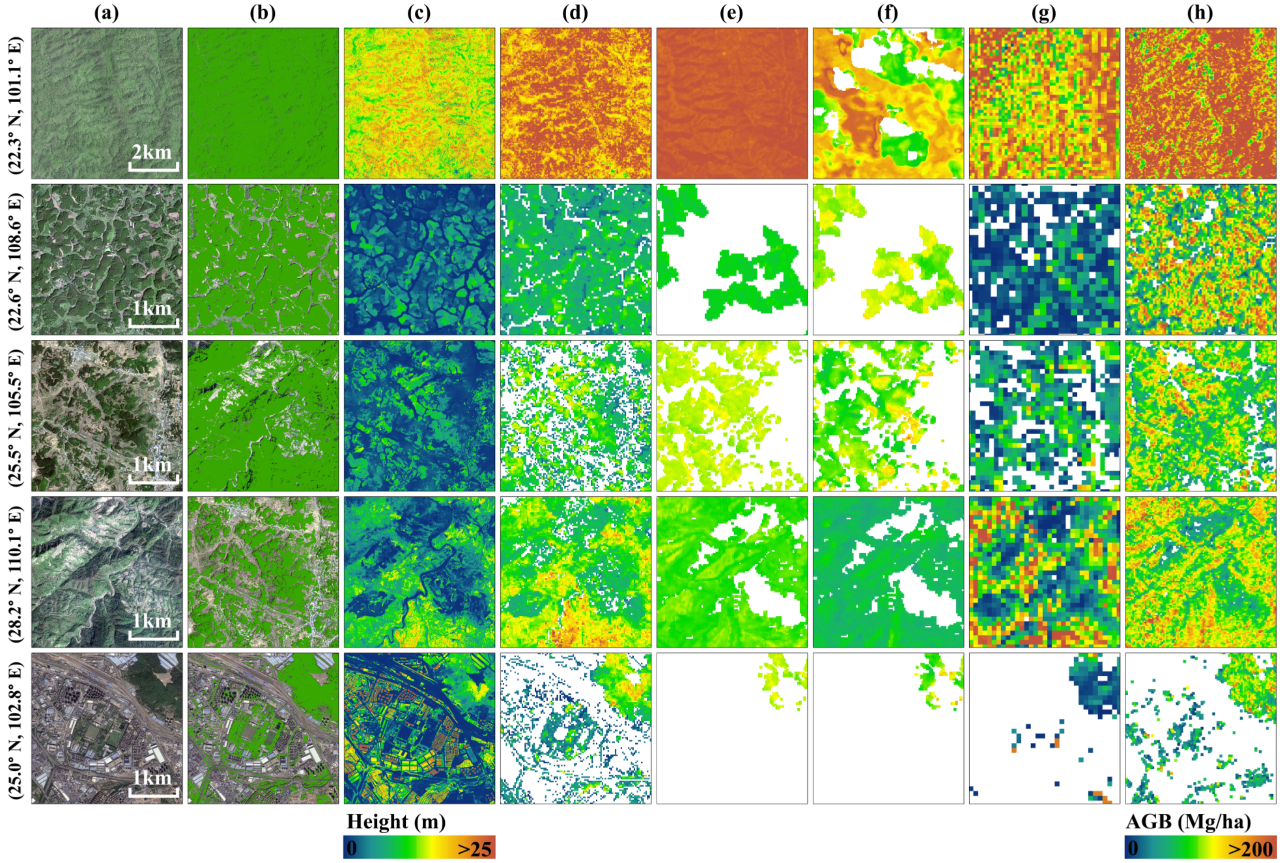


Figure 10. Comparison of different AGB datasets across typical landscapes. (a) GaoFen imagery, (b) Tree Cover, (c) Canopy Height, (d) FineKarstAGB (Ours), (e) AGB from Yang et al. (2023), (f) AGB from Yan et al. (2023a), (g) AGB from ESA CCI and (h) AGB from Cai et al. (2024b).

To qualitatively evaluate the capability of FineKarstAGB to delineate spatial heterogeneity, we selected five representative landscapes (Fig. 10). In these cases, we compared the performance of FineKarstAGB with existing datasets, using GaoFen imagery, tree cover and canopy height as references.

- **Natural Forest (Xishuangbanna, Yunnan)**: In dense natural forest areas, FineKarstAGB shows variations in AGB corresponding to topographic fluctuations and apparent differences in tree growth status, reflecting fine-scale spatial

https://doi.org/10.5194/essd-2026-102
Preprint. Discussion started: 18 March 2026

differentiation. Compared to existing 30 m resolution products, FineKarstAGB delineates the boundaries of local high- and low-AGB areas with greater details, reducing the smoothing or blocky artifacts observed in coarser products (Fig. 10, Row 1).

- **Plantation Forest (Nanning, Guangxi)**: The plantation scene in Nanning, Guangxi, demonstrates FineKarstAGB representation of AGB in managed stands. The dataset displays distinct AGB values between adjacent afforested plots, potentially reflecting differences in planting period, growth age and stand maturity. In comparison to existing datasets, FineKarst AGB shows more granular variations consistent with different growth stages within the plantations, which is relevant for assessing carbon sink potential and management practices (Fig. 10, Row 2).

- **Highly Heterogeneous Landscapes**: In landscapes with high tree cover heterogeneity, such as agroforestry mosaics, forest-shrubland complexes, and urban areas, FineKarstAGB represents AGB distribution at a fine spatial scale.

  - **Agroforestry System (Qianxinan, Guizhou)**: FineKarstAGB maps AGB for patchy, linear, and scattered trees within the agricultural matrix, delineating the interspersed distribution pattern of farmland and forest elements (Fig. 10, Row 3).

  - **Forest-Shrubland Mosaic (Xiangxi, Hunan)**: FineKarstAGB displays differing AGB values that appear associated with contributions from shrub and tree layers, showing detailed distinctions possibly related to variations in vegetation height and density (Fig. 10, Row 4).

  - **Urban Trees (Kunming, Yunnan)**: FineKarstAGB assigns biomass values to sporadic trees, such as those along streets or in parks, even when distributed among non-vegetated features like buildings and roads (Fig. 10, Row 5).

5 Discussion

5.1 Comparison with other AGB datasets

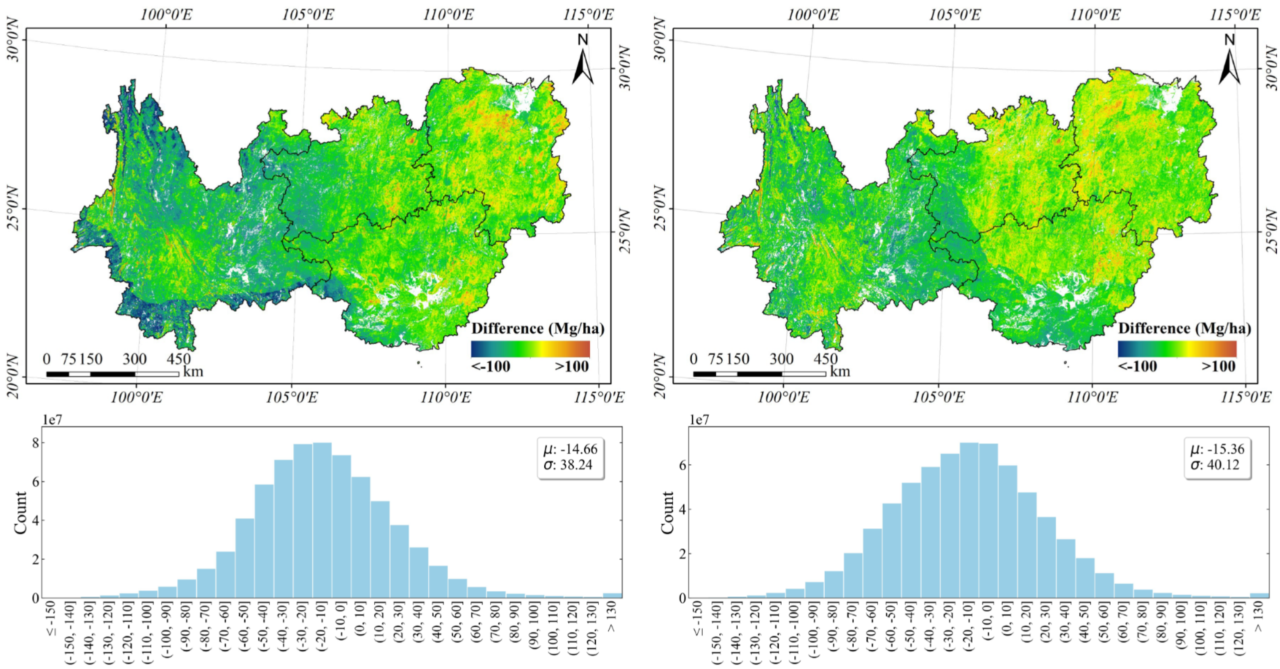
Pixel-wise comparisons between our AGB map and existing AGB datasets indicate that, at the overall statistical level, our estimates are closer to those of Yang et al. (2023) and Yan et al. (2023a), with mean differences of -14.66 Mg/ha and -15.36 Mg/ha, respectively (Fig. 11). In contrast, although the comparison with ESA CCI yields the smallest mean difference (μ = 6.40 Mg/ha), the substantially larger dispersion of differences (σ = 57.17 Mg/ha) indicates pronounced inconsistencies at local scales. Comparisons with the dataset of Cai et al. (2024b) reveal a much larger overall bias, with a mean difference of -40.98 Mg/ha. It is important to note that these discrepancies do not manifest as spatially homogeneous or systematic biases across the study region, but instead exhibit pronounced spatial heterogeneity.

The spatial patterns of the differences further highlight the influence of modeling strategies on the representation of AGB spatial structure. In the comparison with Yang et al. (2023) (Fig. 11a), the differences display a clear regional boundary, particularly evident in southwestern Yunnan, where our AGB map tends to be lower than the estimates of Yang et al. to the west of the boundary, and higher eastwards of the boundary. This abrupt spatial transition is closely associated with the vegetation-zone-based modeling strategy adopted by Yang et al. A similar boundary pattern is also observed in the comparison with Yan et al. (2023a) (Fig. 11b), notably in western Guizhou and northern Guangxi. These results suggest that while geographically or vegetation-stratified modeling approaches can improve internal coherence within individual zones, they may introduce discontinuities at zone boundaries. Such boundary effects are especially pronounced in Southwest China, where complex topography, steep ecological gradients, and highly heterogeneous vegetation amplify inconsistencies.

The comparison between our AGB map and the ESA CCI map shows pronounced differences (Fig. 11c), which is likely attributable to the global modeling framework and coarser spatial resolution of ESA CCI. In mountainous and fragmented forest landscapes, spatial aggregation effects tend to smooth fine-scale AGB variability, limiting the ability of such products to capture ecologically meaningful AGB heterogeneity. Although no clear zonal boundaries are evident in the comparison with the Cai et al. (2024b) dataset (Fig. 11d), the overall magnitude of the bias is substantially large, with very high values in western and southern Yunnan. This discrepancy may be linked to the strong reliance of the Cai et al. product on GEDI observations as primary reference data. Under complex terrain conditions, GEDI measurements are susceptible to slope effects and footprint geolocation uncertainties (Tang et al., 2023), which can substantially increase the uncertainty of AGB reference values and, in turn, amplify discrepancies between the two datasets.

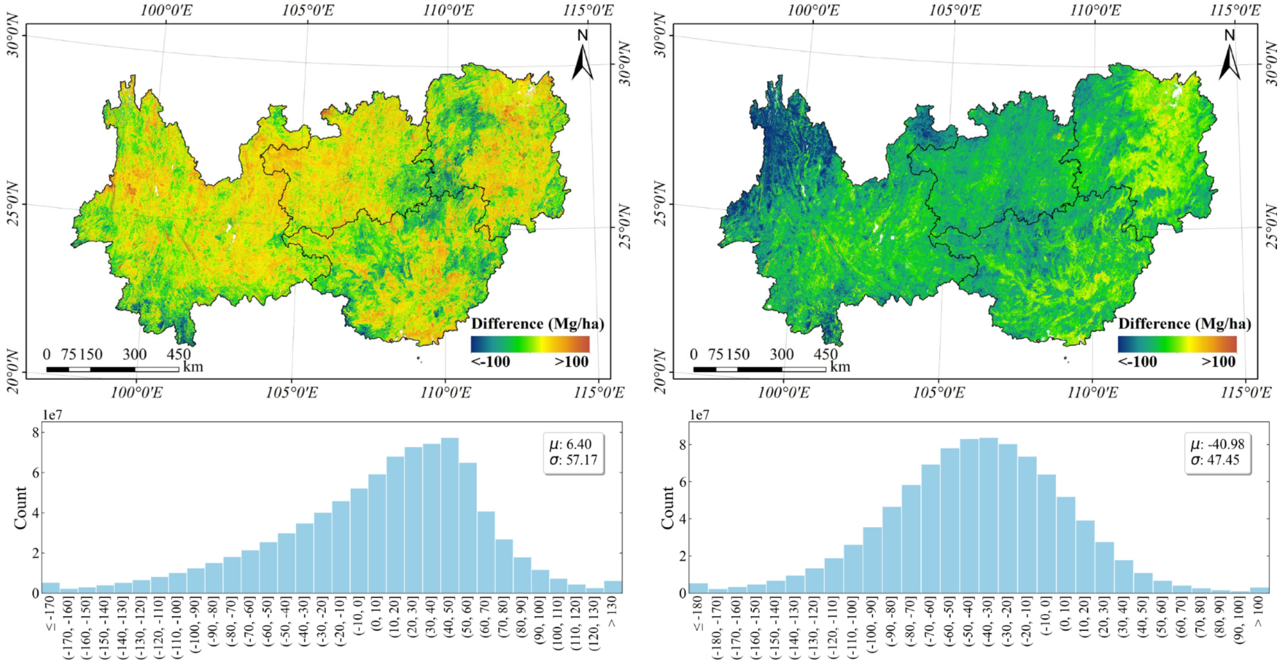
In addition to the dataset-specific differences discussed above, discrepancies between our AGB map dataset and the comparison datasets may arise from differences in time and spatial resolution. Our AGB map utilizes 2024 GaoFen imagery and derived canopy height, capturing the most recent forest conditions. In contrast, the comparison datasets are based on data from 2019, 2021, and 2023, respectively, and therefore do not reflect the latest forest dynamics, such as growth, harvesting, or recovery. This temporal lag is particularly critical in regions like Southwest China, where human disturbances drive rapid changes in forest landscapes. Furthermore, although all AGB datasets are at $\geq$ 30 m resolution, the difference between ours and the other datasets is that we leverage very-high-resolution images (0.8 m) as input, which preserves tree-level structural information. Our 30 m AGB map is derived from the 0.8 m data and thus retains detailed information on sub-pixel heterogeneity, whereas other products convert data with a native ground spacing of 30 m to AGB, without access to finer-scale spatial details. This high resolution enables a more precise characterization of heterogeneous landscapes, especially in complex scenes such as agroforestry systems, forest-shrub mosaics, and urban green spaces, and allows the inclusion of sparsely distributed trees that are missed in 30 m resolution datasets.

The comparison of representative regions with significant differences (Fig. 12) illustrates previously discussed causes of discrepancies among AGB datasets. In both the representative areas and their zoomed-in views, the AGB estimates from this study exhibit spatial patterns that are consistent with the canopy height and GaoFen imagery. In contrast, the results of Yang et al. (2023) and Yan et al. (2023a) clearly reflect the influence of geographic and vegetation zonation. The Pu’er region in Yunnan is located within tropical monsoon forest–rainforest zones (Fig. 12d, e, Row 1–2) and therefore shows generally higher AGB estimates, whereas the Tongren region in Guizhou lies within subtropical evergreen broadleaf forests (Fig. 12d, e, Row 3–4) and is characterized by predominantly moderate to high AGB values; differences among vegetation zones are thus amplified in these datasets. In addition, due to the application of a forest mask, the dataset of Yan et al. exhibits evident no-



(a) FineKarstAGB vs. Yang

(b) FineKarstAGB vs. Yan



(c) FineKarstAGB vs. ESA CCI

(d) FineKarstAGB vs. Cai

Figure 11. Spatial Distribution and Statistics of AGB Differences Between FineKarstAGB (ours) and Comparison datasets.(a) Yang et al. (2023), (b) Yan et al. (2023a), (c) ESA CCI and (d) Cai et al. (2024b). μ and σ represent the mean and standard deviation of the AGB differences, respectively.

data gaps. The discrepancies observed in the Cai et al. (2024b) (Fig. 12g) dataset are mainly reflected in areas with complex 465 terrain and dense forests, where AGB retrievals based solely on spectral features are less capable of capturing true variability associated with canopy vertical structure.

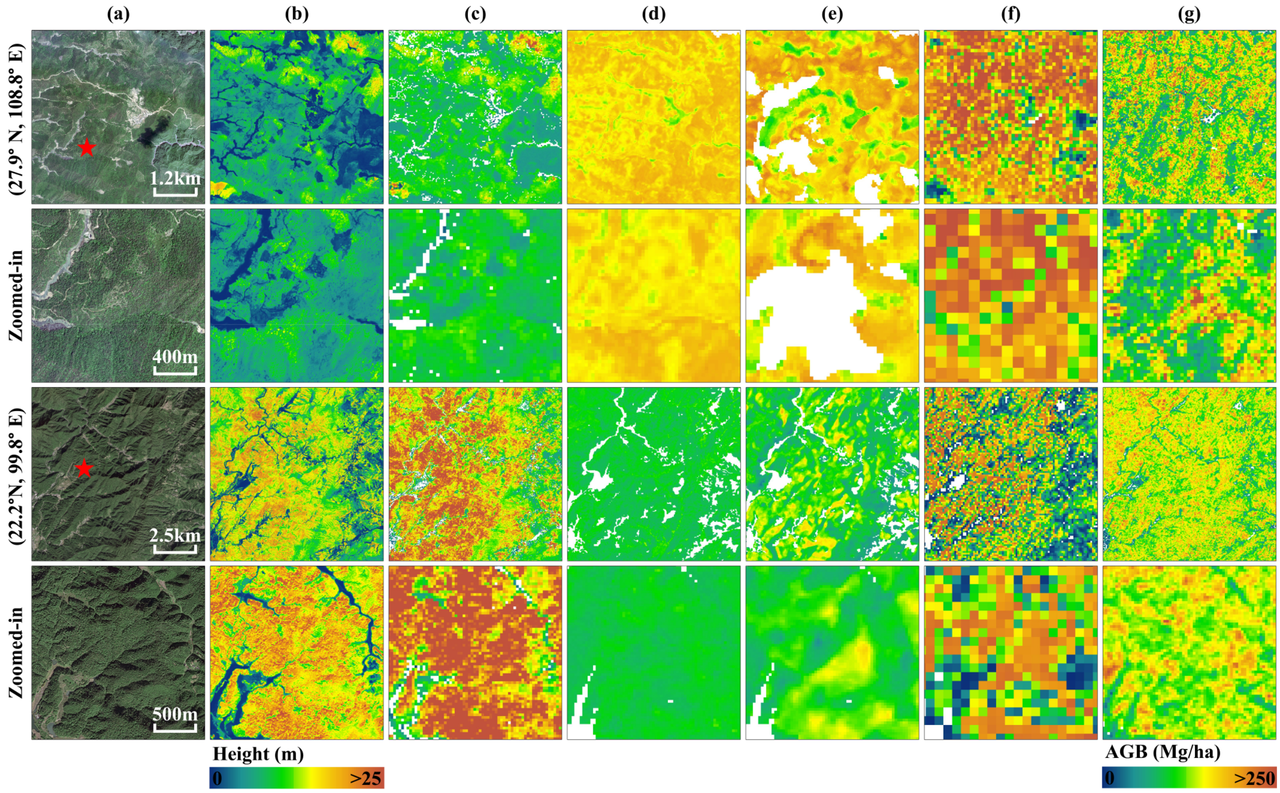


Figure 12. Illustration of typical scenarios and local details showing differences between FineKarstAGB (ours) and comparison AGB datasets. (a) GaoFen imagery, (b) Canopy height, (c) Our AGB map; (d) Yang et al. (2023), (e) Yan et al. (2023a), (f) ESA CCI and (g) Cai et al. (2024b).

5.2 Significance of fine-scale vertical structure in characterizing AGB heterogeneity

Forest vertical structure, particularly canopy height, is key for accurately estimating AGB. Ecologically, this relationship is governed by allometric scaling laws, where tree height acts as a primary proxy for accumulated biomass. In the context of 470 remote sensing, canopy height is also one of the most directly observable structural metrics. Numerous studies have constructed explicit models that utilize canopy height as a central proxy parameter, establishing direct functional relationships between vertical structure and AGB/carbon stocks (Liu et al., 2023; Zhu et al., 2025).

Our study further elucidates the profound impact of fine-scale vertical structural information on characterizing AGB heterogeneity in complex landscapes. As evidenced by the detailed comparisons in Figs. 10, 12, the inclusion of high-resolution canopy height enables FineKarstAGB to mirror subtle variations in vegetation structure. Whether in the dense natural forests of Xishuangbanna or the fragmented agro-forestry systems of Qianxinan, our dataset exhibits spatial patterns that are highly consistent with the canopy height model.

While existing AGB datasets demonstrate reliability at regional or national scales, they often exhibit significant deviations from ground truth at local scales. This phenomenon may be closely related to the production methods of the datasets. Both the datasets by Yang et al. (2023) and Yan et al. (2023a) used forest canopy height (Liu et al., 2022b). However, that canopy height product was derived from the interpolation of GEDI and ICESat-2 data. The inherent limitations of GEDI and ICESat-2 under the complex terrain and canopy conditions of Southwest China compromised the applicability of the retrieved canopy height data (Fu et al., 2025), which in turn resulted in a weak correlation between biomass and canopy height in this region. Meanwhile, the production process for the Cai et al. (2024b) dataset did not directly use canopy height information; its predictions relied more on non-structural remote sensing texture and spectral features. Consequently, it showed almost no correlation with the true CHM, although it is worth noting that the tree cover information utilized in such approaches retains unique advantages for long-term, large-scale monitoring of phenomena such as forest densification and degradation.

5.3 Limitations and future improvements

5.3.1 Temporal and spatial matching of plot data

The National Forest Inventory (NFI) plot data serve as an indispensable source of ground truth for remote sensing–based AGB estimation, yet their application often encounters two major challenges. First, the spatial extent of traditional NFI plots is typically comparable to the pixel size of commonly used medium-resolution remote sensing datasets (e.g., Landsat), which can easily result in incomplete or partial spatial overlap between plots and imagery (Fu et al., 2024; Ma et al., 2024). This spatial mismatch introduces substantial bias and uncertainty, particularly in regions with complex terrain and highly fragmented landscapes. In this study, by leveraging sub-meter GaoFen imagery, we effectively mitigated this spatial mismatch between field plots and remote sensing data, thereby reducing the uncertainty in AGB estimations. However, a considerable temporal gap between the NFI data collection and the remote sensing observation period remains a systemic challenge (Qin et al., 2025). Despite the implementation of rigorous quality control and filtering procedures, forest ecosystems are inherently dynamic, leading to model uncertainty arising from real AGB changes caused by growth, harvesting, or natural disturbances. In the future, where feasible, temporally synchronized or annually updated NFI datasets should be employed to minimize the influence of this temporal discrepancy on model development and validation.

5.3.2 Trade-off between VHR data acquisition and coverage

This study employed GaoFen imagery as the primary data source, whose high spatial and radiometric quality was fundamental to achieving accurate AGB estimation. The raw GF imagery was subjected to geometric and radiometric correction, as well as cloud masking and removal. Nevertheless, a few small areas remained affected by residual cloud contamination; their corresponding AGB estimates were excluded from the final dataset to ensure the reliability of the dataset. Furthermore, the acquisition and processing costs of VHR imagery remain substantial. Although this study achieved seamless coverage across the entire Southwest China region, extending such coverage to larger spatial domains (e.g., national scale) or for long-term temporal monitoring remains challenging. This limitation underscores the trade-off between spatial resolution and geographic or temporal coverage. Future research should therefore explore the integration of alternative high-resolution imagery sources, such as wide-swath GaoFen missions (e.g., GF-1, GF-6) or commercial constellations with high revisit frequency (e.g., PlanetScope), for regional AGB estimation (Mugabowindekwe et al., 2023). Expanding both spatial and temporal coverage while maintaining high estimation accuracy will be a key research priority moving forward.

5.3.3 Validation of non-forest areas and higher-resolution AGB estimation

This study extended AGB estimation to non-forest landscapes with high spatial heterogeneity, including scattered farmland trees, urban green spaces, and sporadic urban vegetation. By leveraging VHR GaoFen imagery, we achieved fine-scale mapping of AGB in these complex environments. It is important to note that accurate AGB estimation in non-forest scenes relies heavily on VHR data; only such imagery can effectively capture and distinguish these small, heterogeneous vegetation units. However, the validation of AGB estimates in non-forest areas remains a challenge, because NFI plots are mostly found in closed canopy forests. The absence of scale-matched, high-quality ground truth data for validating sub-meter or higher-resolution AGB products remains a critical limitation. Traditional NFI plots, constrained by their coarse spatial scale, cannot adequately represent AGB variation arising from fine-scale structural differences. Future research urgently needs to construct or acquire high-quality, scale-matched validation data specifically for non-forest scenes and VHR AGB products. Future work could involve conducting Terrestrial Laser Scanning (TLS) or Airborne Laser Scanning (ALS) data acquisition, such as LIDARNET, to directly measure AGB across urban trees, agroforestry landscapes, and other structurally complex environments at the sub-meter scale. These TLS and ALS datasets can generate high-density 3D point clouds that enable individual tree-level feature extraction, thereby supporting robust validation of VHR AGB products. Incorporating such high-quality, high-resolution reference data will significantly improve the quantitative accuracy and reliability of AGB estimation in non-forest areas, providing a stronger empirical foundation for urban carbon accounting, agroforestry carbon management, and precision forestry.

6 Conclusions

To accurately estimate aboveground biomass in the highly heterogeneous forests of Southwest China, a globally critical ecoregion, this study developed a Canopy Structure–driven Multi-feature Fusion Network (CSMF-Net). The model effectively integrates multi-feature information by fusing spectral features, horizontal structure, and vertical structure derived from sub-meter GaoFen imagery, tree crown segmentation results, and canopy height data. This synergistic fusion enables precise modeling of forest structural complexity and spatial variation. The proposed CSMF-Net achieved robust AGB estimation when evaluated against independent field plots ($r = 0.83$, RMSE = 28.51 Mg/ha). Based on this model, we produced the first fine-grained, high-spatial-resolution AGB dataset for the four provinces of Southwest China (2024). The dataset captures the spatial distribution of AGB across natural forests, plantations, and agroforestry systems, while also quantifying biomass from trees outside forests, including urban trees and scattered farmland trees, that are often overlooked in conventional large-scale AGB products. This inclusion promotes more comprehensive and spatially explicit carbon accounting. Furthermore, the AGB estimates demonstrate strong structural consistency with real forest 3D features; a comparative evaluation against a UAV LiDAR CHM shows a high degree of agreement ($R^2 = 0.54$). Overall, this research establishes a reliable benchmark dataset for assessing biomass and carbon storage in one of China's most ecologically important regions. It provides spatially detailed evidence to evaluate the outcomes of major national ecological restoration initiatives, such as the Grain for Green Program and Karst Rocky Desertification Control. Moreover, the dataset offers fine-scale data support for regional carbon cycle modeling, biodiversity conservation, and sustainable forest management, contributing directly to China's long-term carbon neutrality goals.

Data availability. The FineKarstAGB dataset presented in this study is publicly accessible through the Science Data Bank repository under the digital object identifier (DOI): https://doi.org/10.57760/sciencedb.33452 (Li et al., 2026).

Author contributions. YL and YB designed the research. YL conducted the research and drafted the manuscript. YB, ZC, CL, XT and MB reviewed and revised the paper. XY provided part of the related data. Z.Wu, Z.Wang, HG and XW provided advice. S.Li, XX and S.Liu provided technical support. All authors contributed to the final draft of the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. We gratefully acknowledge LIDARNET (http://lidar.pku.edu.cn) for providing the UAV LiDAR data, which were used to evaluate both the canopy height dataset employed in this study and the AGB dataset.

Financial support. This work was supported by the National Key Research and Development Program of China for Young Scientists (Grant No: 2023YFF1305700), the National Natural Science Foundation of China (Grant No: 42371129), the Science and Technology Innovation Program of Hunan Province (Grant No: 2024RC1067).

References

Araza, A., De Bruin, S., Herold, M., Quegan, S., Labriere, N., Rodriguez-Veiga, P., Avitabile, V., Santoro, M., Mitchard, E. T., Ryan, C. M., Phillips, O. L., Willcock, S., Verbeeck, H., Carreiras, J., Hein, L., Schelhaas, M.-J., Pacheco-Pascagaza, A. M., Da Conceição Bispo, P., Laurin, G. V., Vieilledent, G., Slik, F., Wijaya, A., Lewis, S. L., Morel, A., Liang, J., Sukhdeo, H., Schepaschenko, D., Cavlovic, J., Gilani, H., and Lucas, R.: A Comprehensive Framework for Assessing the Accuracy and Uncertainty of Global Above-Ground Biomass Maps, Remote Sensing of Environment, 272, 112 917, https://doi.org/10.1016/j.rse.2022.112917, 2022.

Brandt, M., Yue, Y., Wigneron, J. P., Tong, X., Tian, F., Jepsen, M. R., Xiao, X., Verger, A., Mialon, A., Al-Yaari, A., et al.: Satellite-Observed Major Greening and Biomass Increase in South China Karst during Recent Decade, Earth's Future, 6, 1017–1028, https://doi.org/10.1029/2018EF000890, 2018.

Brovkina, O., Novotny, J., Cienciala, E., Zemek, F., and Russ, R.: Mapping Forest Aboveground Biomass Using Airborne Hyperspectral and LiDAR Data in the Mountainous Conditions of Central Europe, Ecological Engineering, 100, 219–230, https://doi.org/10.1016/j.ecoleng.2016.12.004, 2017.

Cai, Y., Xu, X., Zhu, P., Nie, S., Wang, C., Xiong, Y., and Liu, X.: Unveiling Spatiotemporal Tree Cover Patterns in China: The First 30 m Annual Tree Cover Mapping from 1985 to 2023, ISPRS Journal of Photogrammetry and Remote Sensing, 216, 240–258, https://doi.org/10.1016/j.isprsjprs.2024.08.001, 2024a.

Cai, Y., Zhu, P., Xu, X., Li, X., Zhang, H., Nie, S., Wang, C., Wang, J., Shen, Q., Li, B., Wu, C., Liu, X., and Chen, Y.: CFATD: The First High-Spatiotemporal-Resolution Mapping of Forest Aboveground Biomass in China from 1985 to 2023 (Part VI: 2022-2023), https://doi.org/10.5281/zenodo.12747329, 2024b.

Cai, Y., Zhu, P., Li, X., Liu, X., Chen, Y., Shen, Q., Xu, X., Zhang, H., Nie, S., Wang, C., Wang, J., Li, B., Wu, C., and Zhuang, H.: Dynamics of China's Forest Carbon Storage: The First 30 m Annual Aboveground Biomass Mapping from 1985 to 2023, Earth System Science Data, 17, 6993–7018, https://doi.org/10.5194/essd-17-6993-2025, 2025.

Campbell, M. J., Eastburn, J. F., Brewer, S. C., and Dennison, P. E.: Deep Learning and Machine Learning Enable Broad-Scale Woodland Height, Cover, and Biomass Mapping, ISPRS Journal of Photogrammetry and Remote Sensing, 226, 187–203, https://doi.org/10.1016/j.isprsjprs.2025.05.016, 2025.

Chang, Z., Hobeichi, S., Wang, Y.-P., Tang, X., Abramowitz, G., Chen, Y., Cao, N., Yu, M., Huang, H., Zhou, G., Wang, G., Ma, K., Du, S., Li, S., Han, S., Ma, Y., Wigneron, J.-P., Fan, L., Saatchi, S. S., and Yan, J.: New Forest Aboveground Biomass Maps of China Integrating Multiple Datasets, Remote Sensing, 13, https://doi.org/10.3390/rs13152892, 2021.

Chen, C., Park, T., Wang, X., Piao, S., Xu, B., Chaturvedi, R. K., Fuchs, R., Brovkin, V., Ciais, P., Fensholt, R., Tømmervik, H., Bala, G., Zhu, Z., Nemani, R. R., and Myneni, R. B.: China and India Lead in Greening of the World through Land-Use Management, Nature Sustainability, 2, 122–129, https://doi.org/10.1038/s41893-019-0220-7, 2019.

Cheng, K., Yang, H., Tao, S., Su, Y., Guan, H., Ren, Y., Hu, T., Li, W., Xu, G., Chen, M., Lu, X., Yang, Z., Tang, Y., Ma, K., Fang, J., and Guo, Q.: Carbon Storage through China's Planted Forest Expansion, Nature Communications, 15, 4106, https://doi.org/10.1038/s41467-024-48546-0, 2024.

Duncanson, L., Kellner, J. R., Armston, J., Dubayah, R., Minor, D. M., Hancock, S., Healey, S. P., Patterson, P. L., Saarela, S., Marselis, S., Silva, C. E., Bruening, J., Goetz, S. J., Tang, H., Hofton, M., Blair, B., Luthcke, S., Fatoyinbo, L., Abernethy, K., Alonso, A., Andersen, H.-E., Aplin, P., Baker, T. R., Barbier, N., Bastin, J. F., Biber, P., Boeckx, P., Bogaert, J., Boschetti, L., Boucher, P. B., Boyd, D. S., Burslem, D. F., Calvo-Rodriguez, S., Chave, J., Chazdon, R. L., Clark, D. B., Clark, D. A., Cohen, W. B., Coomes, D. A., Corona, P., Cushman, K., Cutler, M. E., Dalling, J. W., Dalponte, M., Dash, J., de-Miguel, S., Deng, S., Ellis, P. W., Erasmus, B., Fekety, P. A., Fernandez-Landa, A., Ferraz, A., Fischer, R., Fisher, A. G., García-Abril, A., Gobakken, T., Hacker, J. M., Heurich, M., Hill, R. A., Hopkinson, C., Huang, H., Hubbell, S. P., Hudak, A. T., Huth, A., Imbach, B., Jeffery, K. J., Katoh, M., Kearsley, E., Kenfack, D., Kljun, N., Knapp, N., Král, K., Krůček, M., Labrière, N., Lewis, S. L., Longo, M., Lucas, R. M., Main, R., Manzanera, J. A., Martínez, R. V., Mathieu, R., Memiaghe, H., Meyer, V., Mendoza, A. M., Monerris, A., Montesano, P., Morsdorf, F., Næsset, E., Naidoo, L., Nilus, R., O'Brien, M., Orwig, D. A., Papathanassiou, K., Parker, G., Philipson, C., Phillips, O. L., Pisek, J., Poulsen, J. R., Pretzsch, H., Rüdiger, C., Saatchi, S., Sanchez-Azofeifa, A., Sanchez-Lopez, N., Scholes, R., Silva, C. A., Simard, M., Skidmore, A., Stereńczak, K., Tanase, M., Torresan, C., Valbuena, R., Verbeeck, H., Vrska, T., Wessels, K., White, J. C., White, L. J., Zahabu, E., and Zgraggen, C.: Aboveground Biomass Density Models for NASA's Global Ecosystem Dynamics Investigation (GEDI) Lidar Mission, Remote Sensing of Environment, 270, 112 845, https://doi.org/10.1016/j.rse.2021.112845, 2022.

Fu, L., Shu, Q., Yang, Z., Xia, C., Zhang, X., Zhang, Y., Li, Z., and Li, S.: Accuracy assessment of topography and forest canopy height in complex terrain conditions of Southern China using ICESat-2 and GEDI data., Frontiers in plant science, 16, 1547 688, 2025.

Fu, Y., Li, R., Zhu, Z., Xue, Y., Ding, H., Wang, X., Na, J., and Xia, W.: SCARF: A New Algorithm for Continuous Prediction of Biomass Dynamics Using Machine Learning and Landsat Time Series, Remote Sensing of Environment, 314, 114 348, https://doi.org/10.1016/j.rse.2024.114348, 2024.

Guo, Z., Hu, H., Li, P., Li, N., and Fang, J.: Spatio-Temporal Changes in Biomass Carbon Sinks in China's Forests from 1977 to 2008, Science China Life Sciences, 56, 661–671, https://doi.org/10.1007/s11427-013-4492-2, 2013.

Hu, Y., Xu, X., Wu, F., Sun, Z., Xia, H., Meng, Q., Huang, W., Zhou, H., Gao, J., Li, W., Peng, D., and Xiao, X.: Estimating Forest Stock Volume in Hunan Province, China, by Integrating in Situ Plot Data, Sentinel-2 Images, and Linear and Machine Learning Regression Models, Remote Sensing, 12, https://doi.org/10.3390/rs12010186, 2020.

Huang, H., Liu, C., Wang, X., Zhou, X., and Gong, P.: Integration of Multi-Resource Remotely Sensed Data and Allometric Models for Forest Aboveground Biomass Estimation in China, Remote Sensing of Environment, 221, 225–234, https://doi.org/10.1016/j.rse.2018.11.017, 2019.

Huang, W., Sun, G., Dubayah, R., Cook, B., Montesano, P., Ni, W., and Zhang, Z.: Mapping Biomass Change after Forest Disturbance: Applying LiDAR Footprint-Derived Models at Key Map Scales, Remote Sensing of Environment, 134, 319–332, https://doi.org/10.1016/j.rse.2013.03.017, 2013.

Jian, N., Zhao, H., and Liu, M.: Driving Force and Potential of Forest Resource Change: Based on the Results of the Ninth National Forest Resources Inventory, Journal of Forestry and Grassland Policy, 2, 64–71, https://doi.org/10.12344/lczcyj.2022.09.22.0001, 2022.

Kashongwe, H., Roy, D., and Skole, D.: Examination of the Amount of GEDI Data Required to Characterize Central Africa Tropical Forest Aboveground Biomass at REDD+ Project Scale in Mai Ndombe Province, Science of Remote Sensing, 7, 100 091, https://doi.org/10.1016/j.srs.2023.100091, 2023.

Li, H., Hiroshima, T., Li, X., Hayashi, M., and Kato, T.: High-Resolution Mapping of Forest Structure and Carbon Stock Using Multi-Source Remote Sensing Data in Japan, Remote Sensing of Environment, 312, 114 322, https://doi.org/10.1016/j.rse.2024.114322, 2024a.

Li, J., Bao, W., Wang, X., Song, Y., Liao, T., Xu, X., and Guo, M.: Estimating Aboveground Biomass of Boreal Forests in Northern China Using Multiple Datasets, IEEE Transactions on Geoscience and Remote Sensing, 62, 1–10, https://doi.org/10.1109/TGRS.2024.3408316, 2024b.

Li, Q., Yin, J., Zhang, X., Hao, D., Ferreira, M. P., Yan, W., Tian, Y., Zhang, D., Tan, S., Nie, S., An, T., Li, X., Huang, J., Su, W., and Zeng, Y.: Tree-Level Carbon Stock Estimations across Diverse Species Using Multi-Source Remote Sensing Integration, Computers and Electronics in Agriculture, 231, 109 904, https://doi.org/10.1016/j.compag.2025.109904, 2025.

Li, S., Brandt, M., Fensholt, R., Kariryaa, A., Igel, C., Gieseke, F., Nord-Larsen, T., Oehmcke, S., Carlsen, A. H., Junttila, S., Tong, X., d'Aspremont, A., and Ciais, P.: Deep Learning Enables Image-Based Tree Counting, Crown Segmentation, and Height Prediction at National Scale, PNAS Nexus, 2, pgad076, https://doi.org/10.1093/pnasnexus/pgad076, 2023a.

Li, S., Brandt, M., Tong, X., Oehmcke, S., Igel, C., Gieseke, F., Nord-Larsen, T., Fensholt, R., Chave, J., and Ciais, P.: Deep Learning Tree and Forest Biomass from Sub-Meter Resolution Images, https://doi.org/10.21203/rs.3.rs-3335298/v1, 2023b.

Li, Y., Bai, Y., and Chen, Z.: FineKarstAGB: A Fine-Grained, High-Resolution Dataset of Aboveground Biomass in Southwest China, https://doi.org/10.57760/sciencedb.33452, 2026.

Liu, S., Brandt, M., Nord-Larsen, T., Chave, J., Reiner, F., Lang, N., Tong, X., Ciais, P., Igel, C., Pascual, A., Guerra-Hernandez, J., Li, S., Mugabowindekwe, M., Saatchi, S., Yue, Y., Chen, Z., and Fensholt, R.: The Overlooked Contribution of Trees Outside Forests to Tree Cover and Woody Biomass across Europe, SCIENCE ADVANCES, 2023.

Liu, X., Liu, S., Bao, W., Li, X., Li, G., Zhu, W., Yu, L., Cai, L., Pan, H., Feng, Q., Li, H., Xu, Z., Hu, Z., Zhang, L., and Liu, Q.: Constructing Approaches and Ecological Countermeasures for Ecological Security Barrier in Southwest China, Terrestrial Ecosystem and Conservation, 2, 84–94, https://doi.org/10.12356/j.2096-8884.2022-0043, 2022a.

Liu, X., Su, Y., Hu, T., Yang, Q., Liu, B., Deng, Y., Tang, H., Tang, Z., Fang, J., and Guo, Q.: Neural Network Guided Interpolation for Mapping Canopy Height of China's Forests by Integrating GEDI and ICESat-2 Data, Remote Sensing of Environment, 269, 112 844, https://doi.org/10.1016/j.rse.2021.112844, 2022b.

Loshchilov, I. and Hutter, F.: Decoupled Weight Decay Regularization, 2019.

Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., and Moran, E.: A Survey of Remote Sensing-Based Aboveground Biomass Estimation Methods in Forest Ecosystems, International Journal of Digital Earth, 9, 63–105, https://doi.org/10.1080/17538947.2014.990526, 2016.

Lu, H., Zhang, H., Fan, H., Liu, D., Wang, J., Wan, X., Zhao, L., Deng, Y., Zhao, F., and Wang, R.: Forest Height Retrieval Using P-band Airborne Multi-Baseline SAR Data: A Novel Phase Compensation Method, ISPRS Journal of Photogrammetry and Remote Sensing, 175, 99–118, https://doi.org/10.1016/j.isprsjprs.2021.02.022, 2021.

Ma, T., Hu, Y., Wang, J., Beckline, M., Pang, D., Chen, L., Ni, X., and Li, X.: A Novel Vegetation Index Approach Using Sentinel-2 Data and Random Forest Algorithm for Estimating Forest Stock Volume in the Helan Mountains, Ningxia, China, Remote Sensing, 15, https://doi.org/10.3390/rs15071853, 2023.

Ma, T., Zhang, C., Ji, L., Zuo, Z., Beckline, M., Hu, Y., Li, X., and Xiao, X.: Development of Forest Aboveground Biomass Estimation, Its Problems and Future Solutions: A Review, Ecological Indicators, 159, 111 653, https://doi.org/10.1016/j.ecolind.2024.111653, 2024.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., Popescu, S., Shum, CK., Schutz, B. E., Smith, B., Yang, Y., and Zwally, J.: The Ice, Cloud, and Land Elevation Satellite-2 (Icesat-2): Science Requirements, Concept, and Implementation, Remote Sensing of Environment, 190, 260–273, https://doi.org/10.1016/j.rse.2016.12.029, 2017.

Mitchard, E. T. A., Saatchi, S. S., White, L. J. T., Abernethy, K. A., Jeffery, K. J., Lewis, S. L., Collins, M., Lefsky, M. A., Leal, M. E., Woodhouse, I. H., and Meir, P.: Mapping Tropical Forest Biomass with Radar and Spaceborne LiDAR in Lopé National Park, Gabon: Overcoming Problems of High Biomass and Persistent Cloud, Biogeosciences (Online), 9, 179–191, https://doi.org/10.5194/bg-9-179-2012, 2012.

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., Saatchi, S., and Fensholt, R.: Nation-Wide Mapping of Tree-Level Aboveground Carbon Stocks in Rwanda, Nature Climate Change, 13, 91–97, https://doi.org/10.1038/s41558-022-01544-w, 2023.

Pan, T.: Technical Characteristics of GF-2 Satellite, Aerospace China, pp. 3–9, 2015.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., and Hayes, D.: A Large and Persistent Carbon Sink in the World's Forests, Science, 333, 988–993, https://doi.org/10.1126/science.1201609, 2011.

Pei, J., Wang, L., Huang, H., Wang, L., Li, W., Wang, X., Yang, H., Cao, J., Fang, H., and Niu, Z.: Characterization and Attribution of Vegetation Dynamics in the Ecologically Fragile South China Karst: Evidence from Three Decadal Landsat Observations, Frontiers in plant science, 13, 1043 389, https://doi.org/10.3389/fpls.2022.1043389, 2022.

Qian, C., Qiang, H., and Li, M.: A Novel Multiscale Geographically and Temporally Gravity-Weighted Regression Model: Algorithm Principle and an Application in Assessment of Forest Biomass in Karst Region, IEEE Transactions on Geoscience and Remote Sensing, 63, 1–14, https://doi.org/10.1109/TGRS.2025.3537324, 2025.

Qin, S., Wang, H., Rogers, C., Bermúdez, J., Lourenço, R. B., Zhang, J., Li, X., Chau, J., Tompalski, P., and Gonsamo, A.: Aboveground Biomass Mapping of Canada with SAR and Optical Satellite Observations Aided by Active Learning, ISPRS Journal of Photogrammetry and Remote Sensing, 226, 204–220, https://doi.org/10.1016/j.isprsjprs.2025.05.022, 2025.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241, Springer, 2015.

Santoro, M. and Cartus, O.: ESA Biomass Climate Change Initiative (Biomass_cci): Global Datasets of Forest above-Ground Biomass for the Years 2007, 2010, 2015, 2016, 2017, 2018, 2019, 2020, 2021 and 2022, v6.0, https://doi.org/10.5285/95913FFB6467447CA72C4E9D8CF30501, 2025.

Schwartz, M., Ciais, P., De Truchis, A., Chave, J., Ottlé, C., Vega, C., Wigneron, J.-P., Nicolas, M., Jouaber, S., Liu, S., Brandt, M., and Fayad, I.: FORMS: Forest Multiple Source Height, Wood Volume, and Biomass Maps in France at 10 to 30 m Resolution Based on Sentinel-1, Sentinel-2, and Global Ecosystem Dynamics Investigation (GEDI) Data with a Deep Learning Approach, Earth System Science Data, 15, 4927–4945, https://doi.org/10.5194/essd-15-4927-2023, 2023.

Su, Y., Guo, Q., Xue, B., Hu, T., Alvarez, O., Tao, S., and Fang, J.: Spatial Distribution of Forest Aboveground Biomass in China: Estimation through Combination of Spaceborne Lidar, Optical Imagery, and Forest Inventory Data, Remote Sensing of Environment, 173, 187–199, https://doi.org/10.1016/j.rse.2015.12.002, 2016.

Tang, H., Stoker, J., Luthcke, S., Armston, J., Lee, K., Blair, B., and Hofton, M.: Evaluating and Mitigating the Impact of Systematic Geolocation Error on Canopy Height Measurement Performance of GEDI, Remote Sensing of Environment, 291, 113 571, https://doi.org/10.1016/j.rse.2023.113571, 2023.

Tong, F. and Zhang, Y.: Individual Tree Crown Delineation in High Resolution Aerial RGB Imagery Using StarDist-based Model, Remote Sensing of Environment, 319, 114 618, https://doi.org/10.1016/j.rse.2025.114618, 2025.

Tong, X., Wang, K., Yue, Y., Brandt, M., Liu, B., Zhang, C., Liao, C., and Fensholt, R.: Quantifying the Effectiveness of Ecological Restoration Projects on Long-Term Vegetation Dynamics in the Karst Regions of Southwest China, International Journal of Applied Earth Observation and Geoinformation, 54, 105–113, https://doi.org/10.1016/j.jag.2016.09.013, 2017.

Tong, X., Brandt, M., Yue, Y., Horion, S., Wang, K., Keersmaecker, W. D., Tian, F., Schurgers, G., Xiao, X., Luo, Y., Chen, C., Myneni, R., Shi, Z., Chen, H., and Fensholt, R.: Increased Vegetation Growth and Carbon Stock in China Karst via Ecological Engineering, Nature Sustainability, 1, 44–50, https://doi.org/10.1038/s41893-017-0004-x, 2018.

Tong, X., Brandt, M., Yue, Y., Ciais, P., Rudbeck Jepsen, M., Penuelas, J., Wigneron, J.-P., Xiao, X., Song, X.-P., Horion, S., Rasmussen, K., Saatchi, S., Fan, L., Wang, K., Zhang, B., Chen, Z., Wang, Y., Li, X., and Fensholt, R.: Forest Management in Southern China Generates Short Term Extensive Carbon Sequestration, Nature Communications, 11, 129, https://doi.org/10.1038/s41467-019-13798-8, 2020.

Tong, X., Xu, Q., Yue, Y., Wang, K., Du, H., Fensholt, R., Li, S., Xu, X., Bai, Y., Yang, X., et al.: Old-Growth Forest Patches Are Widespread Outside Nature Reserves in Southern China, https://doi.org/10.21203/rs.3.rs-7643250/v1, 2025.

Trac, C. J., Harrell, S., Hinckley, T. M., and Henck, A. C.: Reforestation Programs in Southwest China: Reported Success, Observed Failure, and the Reasons Why, Journal of Mountain Science, 4, 275–292, https://doi.org/10.1007/s11629-007-0275-1, 2007.

UNFCCC.: Warsaw Framework for REDD-plus, 2013.

Wagner, F. H., Roberts, S., Ritz, A. L., Carter, G., Dalagnol, R., Favrichon, S., Hirye, M. C., Brandt, M., Ciais, P., and Saatchi, S.: Sub-Meter Tree Height Mapping of California Using Aerial Images and LiDAR-informed U-Net Model, Remote Sensing of Environment, 305, 114 099, https://doi.org/10.1016/j.rse.2024.114099, 2024.

Wang, Y., Peng, Y., Hu, X., and Zhang, P.: Fine-Resolution Forest Height Estimation by Integrating Icesat-2 and Landsat 8 OLI Data with a Spatial Downscaling Method for Aboveground Biomass Quantification, Forests, 14, https://doi.org/10.3390/f14071414, 2023.

Xu, X., Brandt, M., Tong, X., Mugabowindekwe, M., Yue, Y., Li, S., Xu, Q., Liu, S., Reiner, F., Wang, K., Chen, Z., Bai, Y., and Fensholt, R.: Large-Scale Characterization of Horizontal Forest Structure from Remote Sensing Optical Images, Remote Sensing in Ecology and Conservation, p. rse2.70058, https://doi.org/10.1002/rse2.70058, 2026.

Yan, S., He, G., and Zhang, X.: Forest Aboveground Biomass Products in China, 2013-2021, https://doi.org/10.57760/sciencedb.07122, 2023a.

Yan, X., Li, J., Smith, A. R., Yang, D., Ma, T., Su, Y., and Shao, J.: Evaluation of Machine Learning Methods and Multi-Source Remote Sensing Data Combinations to Construct Forest above-Ground Biomass Models, International Journal of Digital Earth, 16, 4471–4491, https://doi.org/10.1080/17538947.2023.2270459, 2023b.

Yang, Q., Niu, C., Liu, X., Feng, Y., Ma, Q., Wang, X., Tang, H., and Guo, Q.: Mapping High-Resolution Forest Aboveground Biomass of China Using Multisource Remote Sensing Data, GIScience & Remote Sensing, 60, 2203 303, https://doi.org/10.1080/15481603.2023.2203303, 2023.

Zhang, X., He, G., Yan, S., Long, T., Peng, X., Zhang, Z., and Wang, G.: Forest Status Assessment in China with SDG Indicators Based on High-Resolution Satellite Data, International Journal of Digital Earth, 16, 1008–1021, https://doi.org/10.1080/17538947.2023.2190625, 2023a.

Zhang, Z., Lu, L., Zhao, Y., Wang, Y., Wei, D., Wu, X., and Ma, X.: Recent Advances in Using Chinese Earth Observation Satellites for Remote Sensing of Vegetation, ISPRS Journal of Photogrammetry and Remote Sensing, 195, 393–407, https://doi.org/10.1016/j.isprsjprs.2022.12.006, 2023b.

Zhu, N., Yang, B., Gong, W., Ying, S., Dai, W., and Dong, Z.: An Explicit Forest Carbon Stock Model and Applications, Geo-spatial Information Science, pp. 1–12, https://doi.org/10.1080/10095020.2025.2469889, 2025.