remote sensing

MDPI

Article

A Two-Stage Reference-Guided Workflow for Improving VIIRS Leaf Area Index Retrieval over Mixed Pixels

Tengqi Yue, Haiyong Ding * and Yuanfei Zhang

School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 202412350355@nuist.edu.cn (T.Y.); 202511770010@nuist.edu.cn (Y.Z.)
* Correspondence: hyongding@nuist.edu.cn

Highlights

What are the main findings?

- A GBOV-anchored two-stage workflow substantially improved 500 m VIIRS LAI retrieval over mixed pixels by correcting supervisory labels before final inversion.
- The largest performance gain came from label correction, while subpixel PFT information provided an additional and consistent improvement over both LOSO validation and same-site temporal-transfer evaluation.

What are the implications of the main findings?

- Mixed-pixel LAI retrieval should be treated not only as a predictor-side problem, but also as a supervisory-label quality problem when high-resolution products are aggregated to moderate resolution.
- Reference-guided label correction provides a practical way to improve operational VIIRS LAI retrieval in heterogeneous landscapes, especially where dense canopies and canopy-background mixing remain challenging.

Abstract

Moderate-resolution leaf area index (LAI) retrieval over heterogeneous landscapes is affected not only by unresolved subpixel composition in coarse-resolution predictors, but also by structural bias in supervisory labels aggregated from higher-resolution products. To address this issue, we developed a reference-guided two-stage workflow to improve LAI retrieval from the Visible Infrared Imaging Radiometer Suite (VIIRS). In the first stage, aggregated Sentinel-2 LAI was calibrated against Ground-Based Observations for Validation (GBOV) LP3 reference LAI using subpixel plant functional type (PFT) fractions and forest-sensitive hinge terms to generate corrected 500 m labels. In the second stage, a random-forest model was trained using VIIRS spectral reflectance, viewing geometry, vegetation indices, texture, and subpixel compositional variables. Model development was based on 2020–2021 data from 11 U.S. GBOV sites. Performance was evaluated by same-site temporal transfer to 2019 and 2022 and by strict leave-one-site-out (LOSO) validation. Label calibration improved agreement with GBOV from a coefficient of determination (R^2) of 0.752 and a root mean square error (RMSE) of 1.110 to an R^2 of 0.908 and an RMSE of 0.676. Under LOSO validation, the final model achieved an R^2 of 0.901 with an RMSE of 0.703. On the 2019/2022 overlap subset shared by the final VIIRS retrieval, the official VNP product, and the GBOV reference, the final model achieved an R^2 of 0.905 and an RMSE of 0.609, compared with 0.755 and 0.978 for the official VNP product. These results show that reference-guided label correction, combined with explicit subpixel compositional information, can substantially improve VIIRS LAI retrieval over mixed pixels within the evaluated study domain.

Check for updates

Academic Editor: Jochem Verrelst

Received: 10 May 2026
Revised: 2 July 2026
Accepted: 2 July 2026
Published: 6 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: leaf area index; mixed pixels; label calibration; plant functional type

1. Introduction

Leaf area index (LAI) is a fundamental descriptor of canopy structure and vegetation functioning, and it is widely used in studies of carbon cycling, water exchange, land–atmosphere interaction, and ecosystem dynamics [1,2]. As an essential climate variable, its routine monitoring over broad spatial extents has relied heavily on moderate-resolution optical sensors and operational product streams such as Moderate Resolution Imaging Spectroradiometer (MODIS) and Visible Infrared Imaging Radiometer Suite (VIIRS) [3]. The development of LAI retrieval methods has generally progressed through three broad stages: empirical statistical relationships, physically based radiative transfer models, and hybrid approaches that combine radiative transfer simulations with machine learning [4]. Although these approaches have substantially advanced large-scale vegetation monitoring, multiple validation and uncertainty-assessment studies have shown that moderate-resolution LAI products still retain non-negligible errors, especially under heterogeneous surface conditions and at biome transitions [5–10].

This challenge is closely linked to the mixed-pixel problem. At a 500 m scale, a single pixel commonly integrates multiple vegetation functional types together with non-vegetated background; whereas, many retrieval strategies still operate under assumptions that are closer to pixel-level homogeneity than to real landscape structure. Under such conditions, the nonlinear relationship between canopy reflectance and LAI can give rise to scale effects, retrieval instability, and systematic bias, particularly over heterogeneous terrain and mixed land-cover mosaics [11,12]. The problem is not confined to ambiguity in the coarse-resolution predictors alone. It can also affect the supervisory labels used in data-driven retrieval, because high-resolution optical LAI estimates may themselves retain structural bias when aggregated to coarse resolution without explicitly accounting for canopy composition, background effects, and high-LAI saturation [13].

Existing studies have addressed this issue mainly along two lines. One line improves coarse-resolution LAI retrieval through mixed-pixel correction or heterogeneity-aware adjustment schemes [11,12]. The other introduces subpixel or decametric information into moderate-resolution retrieval so that coarse-scale inversion can better represent within-pixel composition [14]. Both directions demonstrate the value of subpixel information for heterogeneous surfaces. However, they also expose an important limitation for machine-learning workflows: if the high-resolution LAI prior used for model training remains biased in dense forests or structurally complex canopies, part of the upstream error may simply be transferred into the final coarse-resolution model [13].

Recent VIIRS studies have further shown that machine-learning-based strategies, including domain adaptation and deep transfer learning, can improve LAI retrieval beyond standard operational products [15–17]. At the same time, recent analyses of MODIS and VIIRS LAI products from the perspective of long-term stability [18–20], spatiotemporal consistency, compositing quality, and sensor-independent reprocessing indicate that product accuracy and temporal coherence remain active research issues rather than solved problems [21–23]. However, most existing studies have treated mixed pixels mainly as a predictor-side problem, while paying less attention to the fact that supervisory labels can also inherit structural bias after aggregation from higher-resolution products. In heterogeneous landscapes, this distinction matters because retrieval error may arise not only from unresolved subpixel composition in coarse-resolution observations, but also from bias already embedded in the labels used for model training.

In this study, we propose a two-stage framework for 500 m VIIRS LAI retrieval over mixed pixels. In the first stage, aggregated Sentinel-2 LAI is calibrated against Ground-Based Observations for Validation (GBOV) LP3 reference LAI using subpixel plant functional type (PFT) fractions and forest-sensitive nonlinear correction terms. In the second stage, the corrected labels are used to train a VIIRS random-forest model driven by spectral, angular, texture, and subpixel compositional variables. The framework is evaluated over 11 selected U.S. GBOV sites using strict leave-one-site-out (LOSO) validation, and same-site temporal-transfer testing in 2019 and 2022, with additional comparison against the official VNP product on the overlapping subset. The goal of this study is not to claim broad global transferability, but to test whether jointly addressing supervisory-label bias and subpixel heterogeneity can improve VIIRS LAI retrieval within a mixed-pixel study domain under both site-holdout and temporal-holdout conditions.

2. Materials and Methods

2.1. Study Area and Dataset

GBOV sites that span a broad ecological gradient from sparse vegetation to dense forests, encompassing temperate forests, grasslands, arid shrublands, savannas, and croplands. Bartlett Experimental Forest (BART), Harvard Forest (HARV), Smithsonian Conservation Biology Institute (SCBI), and Talladega National Forest (TALL) represent relatively high-LAI forest environments. Central Plains Experimental Range (CPER), Santa Rita Experimental Range (SRER), and Woodworth (WOOD) provide sparse vegetation and strong soil-background conditions. Ordway-Swisher Biological Station (OSBS) represents a structurally complex woodland environment, while North Sterling (STER), Blandy Experimental Farm (BLAN), and Kona Experimental Forest (KONA) extend the domain toward agricultural or mixed agricultural landscapes. This site set was selected to expose the model to both canopy-dominated and background-dominated mixed-pixel conditions and to support testing across heterogeneous ecological settings [24,25]. The spatial distribution of the selected sites is shown in Figure 1. Across the selected sites, the matched scenes used for model development and evaluation spanned from April to October rather than a narrow seasonal window, allowing the experiments to reflect month-to-month variability in canopy condition within the study domain.

2.1.1. VIIRS Surface Reflectance

This study used the VNP09GA Version 2 surface reflectance product from the Visible Infrared Imaging Radiometer Suite as the fundamental data source for retrievals at the 500 m scale [26,27]. This product provides daily observations of surface reflectance after the application of atmospheric correction and quality control. With the central coordinates of the sites serving as the reference, the VNP09GA data undergoes spatial cropping to extract reflectance in the visible, near-infrared, and shortwave infrared bands, along with variables of observation geometry that include the solar zenith angle, the view zenith angle, and the relative azimuth angle. To construct a stable feature space, two processing steps are further implemented. First, based on flags of product quality, pixels containing clouds, cloud shadows, and abnormal observations are removed, and reasonable constraints of physical ranges are applied to the reflectance. Second, using the preprocessed reflectance, spectral features, including the normalized difference vegetation index and the near-infrared reflectance of vegetation, are calculated to enhance the signals of vegetation.

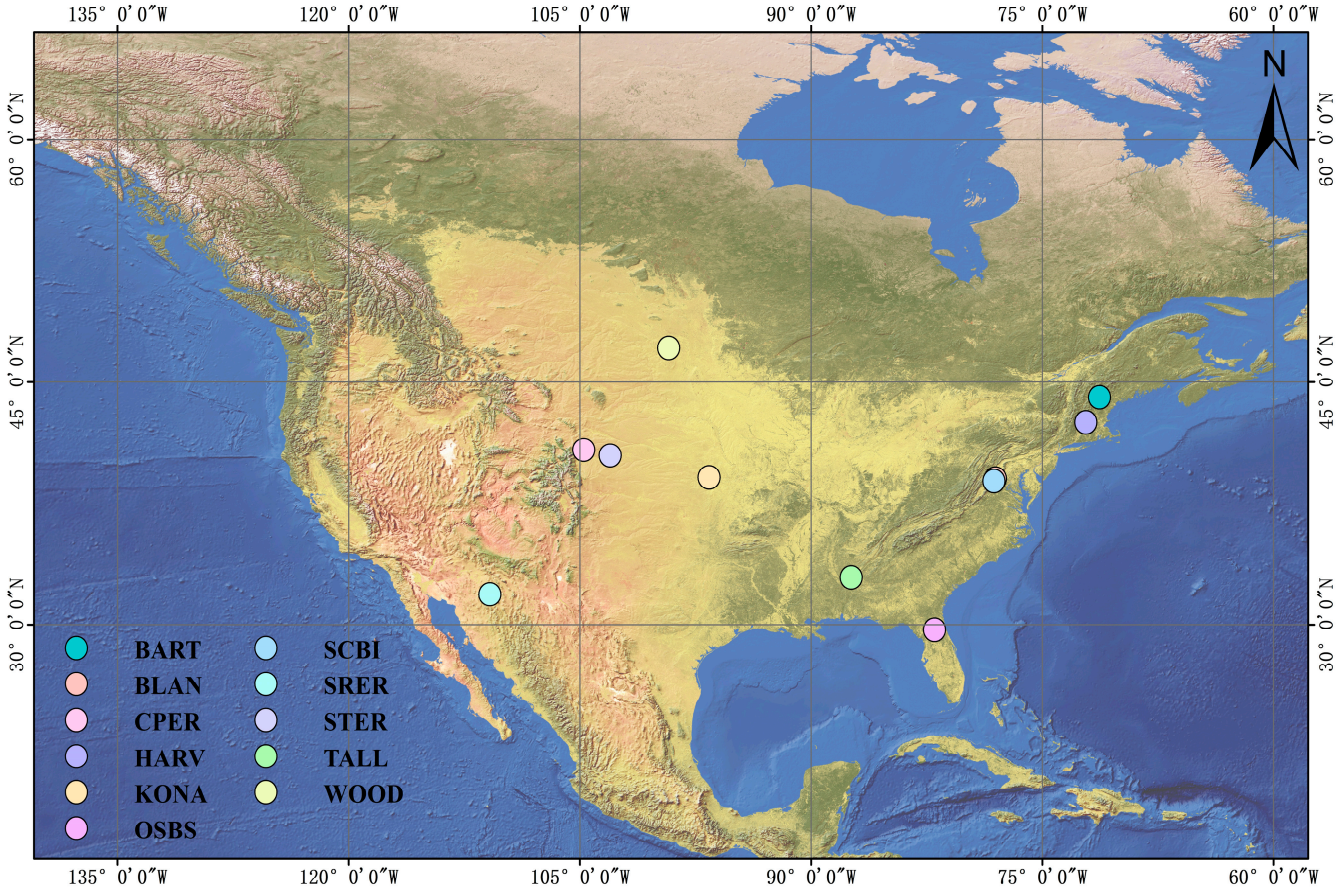

Figure 1. Distribution of the selected U.S. GBOV sites.

2.1.2. Sentinel-2 High-Resolution LAI and Subpixel PFT Composition

High-resolution LAI priors were generated from Sentinel-2 surface reflectance using the PFT-aware hybrid retrieval framework of Wan et al. [13], which applies differentiated canopy structural constraints for individual plant functional types. For each 500 m target grid cell used in a given stage, the area fractions of these PFTs were computed by aggregating the classifications. These subpixel PFT fractions serve two purposes in the framework: first, as components in the label-correction model; second, as explicit predictors in the final VIIRS LAI retrieval model.

Subpixel land-cover composition was derived from the GLC-FCS30D global 30 m annual land-cover product [28]. Following the International Geosphere-Biosphere Program (IGBP) classification scheme, the original classes were regrouped into nine vegetation plant functional types (PFTs)—cropland, shrubland, grassland, woody savanna, wetland, evergreen needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, and mixed forest—plus a non-vegetated class. For each 500 m VIIRS grid cell, the area fractions of these PFTs were computed by aggregating the 30 m classifications. These subpixel PFT fractions serve two purposes in the framework: first, as components in the label-correction model; second, as explicit predictors in the final VIIRS LAI retrieval model.

2.1.3. GBOV LP3 Reference LAI Data

LP3 reference LAI product [29] was adopted as reference targets for calibrating the aggregated Sentinel-2 LAI labels and as benchmark reference data for evaluating the final VIIRS LAI retrievals. This site-scale product, which covers approximately 3 × 3 km around each NEON site, is derived from upscaled in situ measurements and provides pixel-wise LAI estimates together with valid-pixel fractions and quality flags. All GBOV scenes were

reprojected to the UTM coordinate system of the corresponding site. A strict quality-control procedure was then applied: pixels with valid fractions below a prescribed threshold were discarded, observations flagged as abnormal were removed, and LAI values outside a physically plausible range of 0–10 were excluded. The quality-controlled GBOV pixels served two primary roles: as reference targets for correcting the aggregated Sentinel-2 LAI labels, and as independent ground-truth data for evaluating the final VIIRS LAI retrievals. In the GBOV documentation, LP3 is operationally referred to as LAI. However, because the underlying in situ measurements may also be sensitive to non-foliage canopy elements, the reference quantity can in some cases be conceptually closer to plant area index (PAI) than to a strictly foliage-only LAI definition. In this study, we retain the LP3 terminology used by GBOV and use it consistently as the high-confidence reference anchor for both label calibration and evaluation.

2.2. Method

2.2.1. Overall Framework

This study addresses VIIRS leaf area index (LAI) retrieval over mixed pixels at the 500 m scale through a two-stage framework consisting of high-resolution label correction and moderate-resolution retrieval modeling (Figure 2). The overall idea is to first combine high-resolution remote-sensing priors with a high-confidence ground-anchored reference in order to generate more reliable 500 m target labels, and then to use VIIRS moderate-resolution observations to train the final retrieval model. In this way, the framework is designed to reduce the mixed-pixel bias and high-LAI saturation effects that commonly affect conventional moderate-resolution LAI products over heterogeneous surfaces.

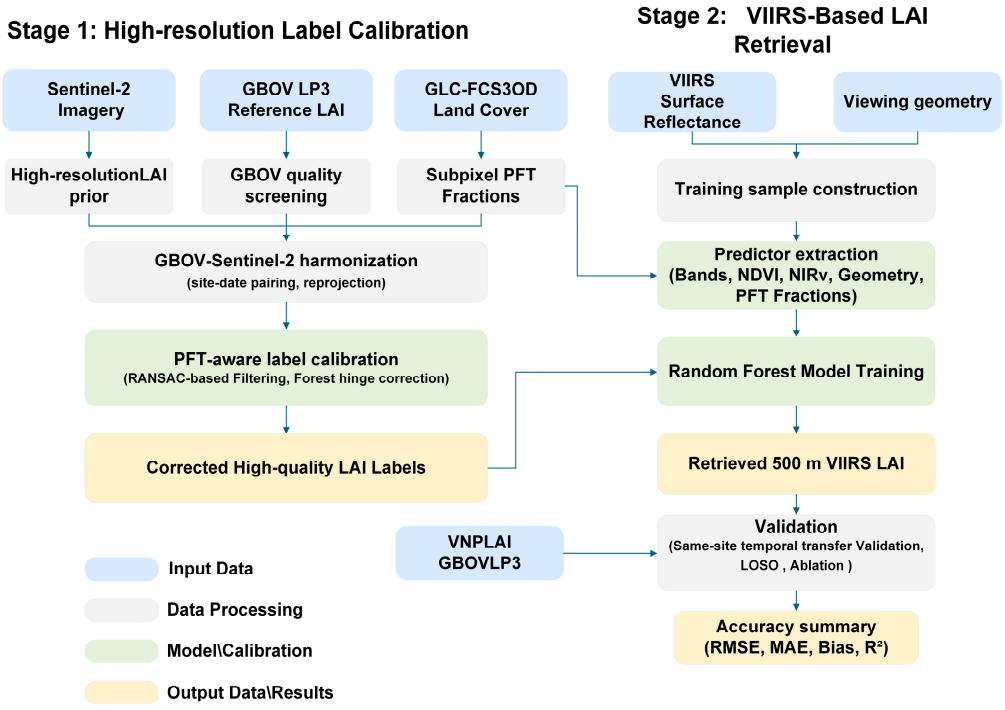

Figure 2. Overall workflow of the proposed two-stage framework.

In Stage 1, high-resolution LAI prior information is derived from Sentinel-2 multispectral imagery, and PFT fractions are quantified from land-cover data to characterize within-pixel composition and spatial heterogeneity. On this basis, the GBOV LP3 reference LAI product is introduced as a high-confidence anchor. Because GBOV is generated by upscaling coordinated ground observations together with high-resolution remote-sensing

information, it provides relatively strong spatial representativeness and reliability. It is therefore used for label correction and as the reference benchmark for subsequent accuracy assessment. After additional screening based on observation quality, spatial representativeness, and vegetation-cover conditions, a high-quality corrected LAI label set is produced for the next stage.

In Stage 2, the corrected LAI label set produced by Stage 1 is used as the target variable for VIIRS-scale retrieval modeling. The main inputs are derived from Suomi-NPP/VIIRS surface reflectance data and include key visible, near-infrared, and shortwave-infrared reflectance bands, together with solar zenith angle, view zenith angle, and relative azimuth angle, as well as subpixel compositional information. Samples from 2020 to 2021 are used for the main label-correction and model-training workflow; whereas, samples from 2019 and 2022 are reserved for cross-year independent evaluation.

2.2.2. Spatiotemporal Harmonization

Multi-source observations differ in projection, spatial resolution, and acquisition timing, and such inconsistencies can directly affect label quality and upscaled retrieval accuracy. To ensure strict sample correspondence, all datasets were harmonized at a common 500 m analysis scale, while the target grid was defined separately for each stage of the workflow according to the role of the data in that stage.

In Stage 1 label calibration, the aggregated Sentinel-2 LAI scene served as the target 500 m grid. The GBOV LP3 reference product was reprojected and aligned to that grid so that each calibration sample contained aggregated Sentinel-2 LAI, GBOV reference LAI, subpixel PFT composition, and associated quality information on the same statistical unit. In Stage 2 VIIRS retrieval, the VIIRS scene served as the target retrieval grid. Because the original VIIRS scenes were provided in a geographic coordinate system; whereas, the selected sites span different UTM zones, the corresponding UTM projection was determined automatically from the scene center, after which the VIIRS scene was reprojected to a 500 m UTM grid. The corrected labels from Stage 1 and the subpixel compositional variables were then co-registered to the same VIIRS grid so that all model inputs and target labels were spatially matched at the final retrieval scale.

For continuous variables, including aggregated Sentinel-2 LAI and GBOV reference LAI, scale harmonization was performed using area-weighted averaging in order to preserve the meaning of grid-cell means after reprojection. For categorical layers, such as land-cover classes and quality flags, nearest-neighbor resampling was used to preserve class boundaries and avoid artificial edge mixing. Subpixel PFT fractions were then computed on the target 500 m grid from the harmonized land-cover composition.

Temporal matching was performed with the VIIRS acquisition date as the reference. Only sample triplets for which the acquisition-time difference among Sentinel-2, VIIRS, and GBOV observations was within ± 3 days were retained, thereby reducing inconsistencies caused by phenological variation during periods of rapid vegetation growth. Samples that did not satisfy this temporal constraint were discarded.

After harmonization, a high-quality calibration sample set was constructed by applying several screening criteria: the valid-pixel fraction of aggregated Sentinel-2 LAI exceeded 0.7, the total vegetation fraction reached at least 0.7, the GBOV valid-pixel fraction exceeded 70%, and LAI values fell within the physically plausible range of 0–10.0. After screening, the label-calibration dataset contained 21,076 matched pixels. For the second-stage retrieval model, training samples were constructed from matched VIIRS predictors, subpixel PFT fractions, and corrected labels. The model-development workflow used 2020–2021 data and retained 643,597 training samples after the final sampling procedure. The strict LOSO summary was based on 29,865 blind-test pixels aggregated across all folds. Because the

matched scenes span all months with available observations, these datasets represent full-year conditions within the selected study domain rather than a limited seasonal subset.

2.2.3. GBOV-Anchored PFT-Aware Label Calibration

In heterogeneous landscapes, high-resolution LAI retrievals from Sentinel-2 are prone to optical saturation over dense forests and are also affected by understory and background interference, leading to systematic biases after aggregation to 500 m. To reduce these biases, Stage 1 calibrates the aggregated Sentinel-2 LAI against spatially representative GBOV LP3 references while explicitly accounting for subpixel plant functional type (PFT) composition.

Inspired by the mixed-pixel area-weighting concept of Dong et al. [12], but not directly retrieving pure-pixel LAI for each PFT followed by area-weighted averaging, we decomposed the pixel-scale aggregated LAI according to PFT area fractions and estimated PFT-specific correction coefficients to represent the differential contributions of subpixel land-cover components. Because evergreen needleleaf forest, evergreen broadleaf forest, deciduous broadleaf forest, and mixed forest are more susceptible to saturation under high-LAI conditions, an additional hinge term was introduced for the forest PFT subset. This term is activated only when the aggregated LAI exceeds a predefined threshold, thereby providing additional flexibility for dense forest pixels and alleviating compression and underestimation in the high-LAI range.

To reduce the influence of local outliers and noisy sample pairs, RANSAC was first used to identify inlier samples [30]. In the current implementation, RANSAC was configured with min_samples = 0.7, residual_threshold = 1.5, and max_trials = 500. The final correction coefficients were then estimated on the inlier set under nonnegative and range-constrained conditions to maintain numerical stability and physical interpretability. Previous studies have shown that optical vegetation indices tend to saturate when LAI exceeds approximately 3, and evaluations in forested environments further suggest that LAI around 4 can be regarded as an empirical transition from moderate to high canopy density. Zhao et al. [31] analyzed the mechanism of LAI inversion saturation using the PROSAIL model and reported that canopy reflectance tends to saturate when LAI exceeds 3, with red and near-infrared reflectance becoming less responsive to further vegetation growth. In addition, the CYCLOPES LAI product, which was generated using a neural network trained with PROSAIL radiative transfer simulations, also shows dynamic-range compression in the high-LAI domain, with saturation commonly occurring around LAI values of 4–5 [32]. Consistent with these findings, our diagnostic analysis of the uncalibrated aggregated Sentinel-2 LAI labels showed that the systematic bias was strongly LAI-dependent(Table 1). The original labels showed a slight overestimation for LAI < 1, with a Bias of 0.193, and nearly unbiased behavior in the 1–2 LAI interval. Underestimation started to emerge in the 2–3 interval, with a Bias of −0.335, and became stronger in the 3–4 interval, with a Bias of −0.562. In contrast, the Bias increased markedly to −1.033, −1.557, and −2.216 in the 4–5, 5–6, and 6–8 intervals, respectively. These results indicate that the pronounced systematic underestimation of the original aggregated labels mainly occurred when LAI exceeded 4.

Accordingly, we set T = 4.0 as the activation threshold for the forest-specific high-LAI compensation term, representing the transition from moderate underestimation to strong underestimation in dense forest conditions. This threshold is not intended to represent a universal forest saturation boundary, but a pragmatic calibration choice supported by both previous optical-LAI saturation studies and the observed label-error pattern. Therefore, the hinge threshold was set to T = 4.0. The corrected label is expressed as follows:

$$L_H = \alpha + \sum_{k\in\Omega} \beta_k L_{S2} f_k + \beta_o L_{S2} f_o + \sum_{k\in\Phi} \gamma_k \max(0, L_{S2} - T) f_k \quad (1)$$

where L_H is the corrected high-quality 500 m label, α is a global background intercept, Ω denotes the set of retained vegetation PFTs, $\Phi \subset \Omega$ specifically denotes the forest-community subset prone to optical saturation (evergreen needleleaf, evergreen broadleaf, deciduous broadleaf, and mixed forest), L_{S2} is the initial aggregated Sentinel-2 LAI within the pixel, and f_k is the subpixel area fraction of the vegetation type. The term f_o is the residual fraction not explicitly assigned to the retained vegetation PFTs; in practice, it mainly represents non-vegetated or background-related components, hence the corresponding term absorbs background interference rather than representing a vegetation LAI contribution. The coefficient β_k is the linear scaling factor associated with each land-cover class. The third term is the hinge piecewise constraint, parameterized by the saturation threshold T and the nonlinear compensation coefficient γ_k. When $L_{S2} < T$, this term evaluates to zero and the model follows a conventional component-wise linear mapping. When $L_{S2} \geq T$, the term becomes active and provides additional flexibility in the high-LAI range, which can help reduce saturation-related compression and underestimation in dense forest pixels. Through this procedure, the Stage 1 labels retain the spatial detail inherited from the high-resolution inputs while incorporating ground-reference constraints and subpixel compositional information, thereby providing a more stable and better-constrained supervisory target for the subsequent VIIRS retrieval.

Table 1. Stratified errors of the original aggregated LAI labels against GBOV reference LAI.

GBOV LAI Interval	RMSE	MAE	Bias
0–1	0.395	0.270	0.193
1–2	0.559	0.443	−0.024
2–3	0.688	0.576	−0.335
3–4	0.839	0.682	−0.562
4–5	1.181	1.048	−1.033
5–6	1.626	1.557	−1.557
6–8	2.259	2.216	−2.216

2.2.4. VIIRS LAI Retrieval with Subpixel Heterogeneity

After the corrected high-quality labels had been generated, Stage 2 established the nonlinear mapping between VIIRS observations and target LAI at the 500 m scale. Because substantial surface heterogeneity is common within a 500 m pixel, the predictor system was designed to represent both spectral canopy signals and subpixel compositional complexity.

The predictor set comprised four groups of variables, as summarized in Table 2. The first group included multispectral reflectance and derived vegetation indices, specifically I1_RED, I2_NIR, I3_SWIR, M3_BLUE, M4_GREEN, and M11_SWIR2 reflectance, together with normalized difference vegetation index (NDVI) and near-infrared reflectance of vegetation (NIRv), to characterize vegetation biochemical status and canopy radiative response. The second group included solar-view geometry variables and their cosine-transformed terms, including solar zenith angle, view zenith angle, and relative azimuth angle, in order to account for bidirectional reflectance effects. The third group comprised a local texture metric defined as the standard deviation of NDVI within a 3×3 NDVI standard-deviation texture metric was calculated to describe local contextual heterogeneity around each VIIRS pixel. Because the metric is computed over a 3×3 VIIRS neighborhood, it represents landscape-level variability at an approximate 1.5 km scale rather than the subpixel heterogeneity within an individual 500 m pixel. Within-pixel mixing was instead represented by the PFT fractional variables derived from the 30 m land-cover product. The fourth group comprised subpixel compositional descriptors, including the area fractions of all retained PFTs, total forest fraction, and dominant-component fraction, thereby explicitly representing within-pixel mixing structure and component complexity.

Table 2. Summary of Stage 2 predictor variables.

Group	Variables
VIIRS reflectance	I1_RED, I2_NIR, I3_SWIR, M3_BLUE, M4_GREEN, M11_SWIR2
Vegetation indices	NDVI, NIRv
Observation geometry	SZA, VZA, RAA, cos(SZA), cos(VZA), cos(RAA)
Texture	3 × 3 NDVI standard deviation
PFT composition	Frac_CRO, Frac_SHR, Frac_GRA, Frac_WSA, Frac_WET, Frac_ENF, Frac_EBF, Frac_DBF, Frac_MF, Dominant-component fraction

The retrieval model was implemented using random-forest regression. This algorithm is well suited to nonlinear relationships among multi-source remote-sensing predictors, is relatively robust to collinearity and noise, and is computationally practical at the present sample size. The final implementation used 350 trees, with max_features set to the square root of the predictor dimension, min_samples_leaf set to 3, and min_samples_split set to 4.

Training samples were constructed by pixel-wise expansion of matched VIIRS scenes, subpixel PFT information, and corrected LAI labels after temporal and site matching had been completed. In the main workflow, samples from 2020 to 2021 were used for Stage 1 label calibration and Stage 2 model training; whereas, samples from 2019 and 2022 were reserved for cross-year independent evaluation.

2.2.5. Evaluation Design

The performance of the proposed framework was assessed using two complementary evaluation protocols. First, temporal transferability was evaluated using data from 2019 and 2022, which were entirely excluded from model fitting and label calibration. These experiments were designed to assess interannual generalization at the same selected sites under different observation years. For comparison with the operational product, the final VIIRS retrieval and the official VNP LAI product were evaluated on the overlap subset where both predictions and GBOV reference data were available. Second, site transferability was assessed using strict leave-one-site-out validation. In each fold, one site was excluded from random-forest training and used only for blind testing. Accordingly, the LOSO results quantify the site generalization ability of the Stage 2 VIIRS retrieval model under unseen-site conditions.

3. Results

3.1. Ablation of Label Correction and PFT Information

To clarify the contribution of each key design within the two-stage framework to the performance of the model, this study conducts an ablation experiment. Three models are designed for comparison to verify the independent contributions of the correction of labels and the information regarding components of sub-pixel plant functional types to the retrieval of VIIRS LAI. The three models share identical rules for the screening of training samples, configurations of parameters for the random forest, and procedures for validation. Their differences lie solely in the source of training labels and the composition of input features. Specifically, model_v01 employs uncorrected 500 m aggregated Sentinel-2 LAI as training labels and inputs all final features to evaluate the limitation of performance imposed by the original aggregated labels. Model_v02 adopts corrected labels and inputs spectral features, vegetation indices, observation geometry, and features of local texture without introducing information regarding sub-pixel components. Model_v03 builds upon model_v02 by further incorporating sub-pixel components and their related features,

which serves as the final model of this study. Through these comparisons, the respective contributions of the correction of labels and the information of sub-pixel components are identified.

The first-stage calibration substantially improved the consistency between the aggregated Sentinel-2 labels and the GBOV reference (Figure 3). Relative to the raw aggregated labels, the calibrated labels reduced root mean square error (RMSE) from 1.110 to 0.676 and mean absolute error (MAE) from 0.817 to 0.497, while coefficient of determination (R^2) increased from 0.752 to 0.908. The bias magnitude decreased from −0.574 to −0.031. These changes indicate that the original aggregated labels contained strong systematic underestimation and that the GBOV-anchored, PFT-aware correction markedly improved the quality of the downstream training target.

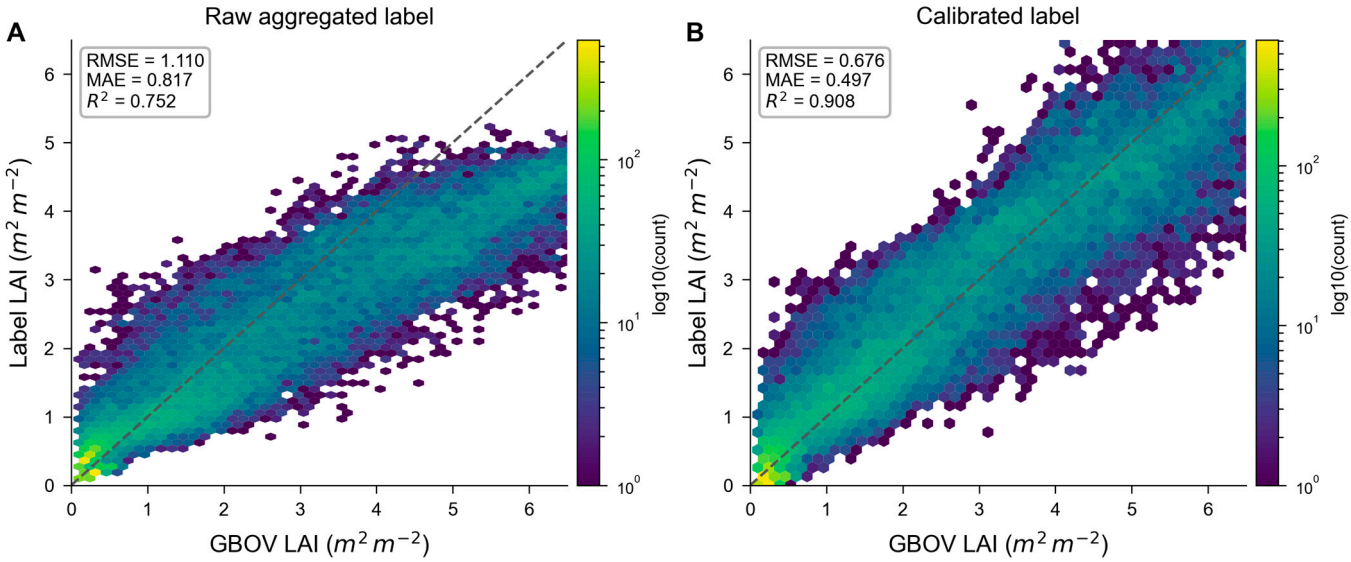

Figure 3. Effect of GBOV−anchored label calibration: panel (**A**) shows the agreement between the raw aggregated Sentinel−2 LAI labels and the GBOV reference, while panel (**B**) shows the corresponding relationship after GBOV-anchored PFT-aware label calibration. The dotted line represents the 1:1 reference line.

Across both strict leave-one-site-out (LOSO) validation and same-site temporal-transfer evaluation, the ablation experiments showed a clear and consistent pattern (Table 3; Figure 4). Under LOSO, the raw-label baseline (v01) achieved an RMSE of 1.457, an MAE of 0.932, and an R^2 of 0.573. Replacing the raw aggregated Sentinel-2 labels with GBOV-anchored calibrated labels, while still excluding subpixel PFT fractions from the final predictor set (v02), improved performance substantially to an RMSE of 0.828, an MAE of 0.518, and an R^2 of 0.862. Adding subpixel PFT fractions and related composition variables further improved the final model (v03) to an RMSE of 0.703, an MAE of 0.452, and an R^2 of 0.901.

Table 3. Overall performance of the key model variants.

Model	LOSO			Same-Site Temporal Transfer		
	RMSE	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**
v01	1.457	0.932	0.573	1.216	0.743	0.614
v02	0.828	0.518	0.862	0.710	0.452	0.878
v03	0.703	0.452	0.901	0.608	0.379	0.900

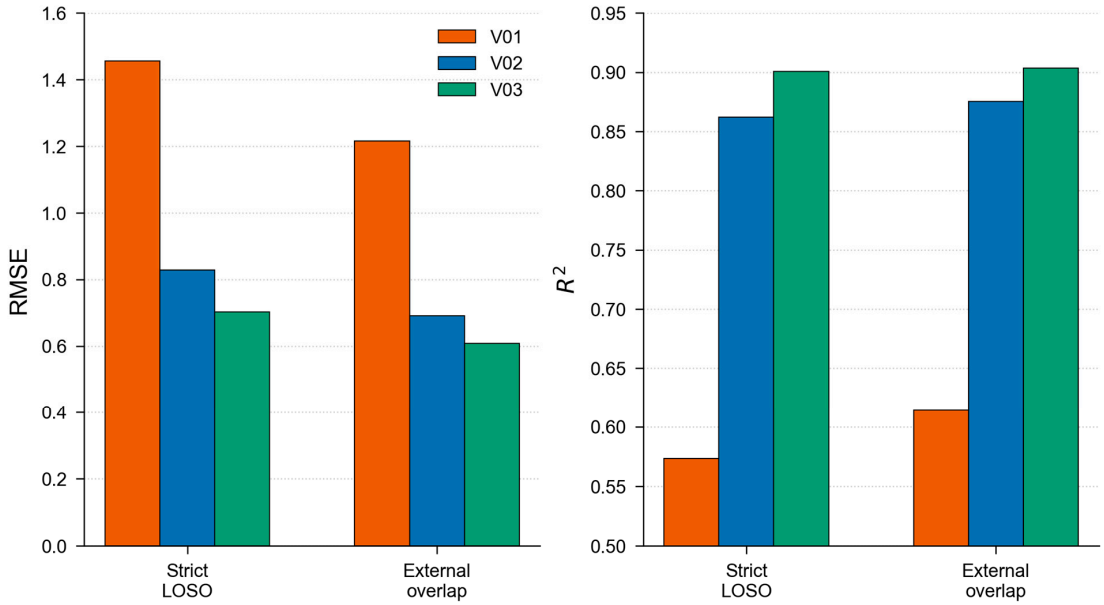

Figure 4. Ablation summary for label correction and subpixel PFT information.

A similar ablation pattern was observed in the same-site temporal-transfer evaluation. The raw-label baseline (v01) achieved an RMSE of 1.216, an MAE of 0.743, and an R^2 of 0.614; whereas, the calibrated-label model without subpixel PFT predictors (v02) improved to an RMSE of 0.710, an MAE of 0.452, and an R^2 of 0.878. After subpixel PFT fractions were incorporated, the final model (v03) achieved the best performance, with an RMSE of 0.608, an MAE of 0.379, and an R^2 of 0.900. Taken together, these results indicate that the dominant gain arose from improving the supervisory labels, while subpixel PFT information provided a smaller but consistent additional improvement under both spatial-transfer and temporal-transfer conditions. This ablation pattern suggests that mixed-pixel heterogeneity affects not only the final inversion step, but also the quality of the supervisory target used to train the model.

3.2. Feature Importance of the Final Retrieval Model

To better interpret the final random-forest retrieval model, we evaluated the relative importance of different predictor categories on the external-year validation subset, as shown in Figure 5. Specifically, predictors were organized into five categories, including PFT composition, VIIRS reflectance, vegetation indices, observation geometry, and texture/quality variables. The importance of each category was assessed by perturbing the corresponding predictors and quantifying the resulting increase in RMSE relative to the unperturbed model. The positive RMSE increases were then normalized to 100% to facilitate comparison among predictor categories. Therefore, the reported values should be interpreted as relative model reliance under the external-year validation setting, rather than as independent physical or causal contributions.

These results indicate that the improvement of the proposed workflow was mainly driven by the combination of subpixel compositional constraints and optical vegetation information. The dominant role of PFT composition supports the central assumption that mixed-pixel LAI retrieval at 500 m benefits from explicitly representing within-pixel land-cover structure. At the same time, the substantial contributions from vegetation indices and VIIRS reflectance confirm that the model still relies strongly on physically relevant spectral signals rather than on auxiliary variables alone. Because vegetation indices are derived from reflectance bands and are therefore correlated with the original spectral variables, the

relative contribution rates should be interpreted as group-level predictive sensitivity rather than independent physical variance partitioning.

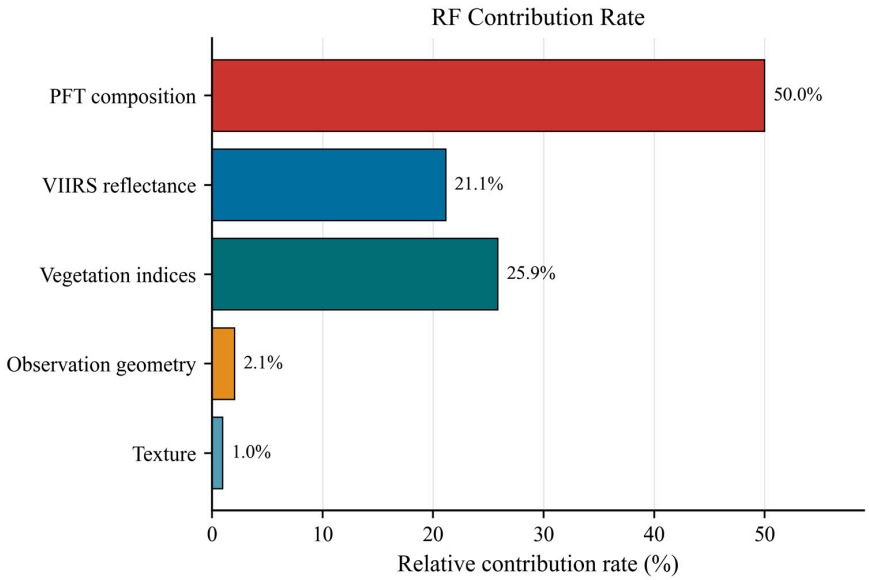

Figure 5. Feature importance of the final random-forest retrieval model.

3.3. Comparison with the Official VNP Product

The final model generalized well to temporally independent observations from 2019 and 2022 at the same study sites. The overall comparison of the shared overlap subset is shown in Figure 6. On that subset, where the final VIIRS model, the official VNP product, and the GBOV reference were all available, the final model achieved an R^2 of 0.905, an RMSE of 0.609, and an MAE of 0.380. On the same subset, the VNP product achieved an R^2 of 0.755, an RMSE of 0.978, and an MAE of 0.635. This corresponds to an RMSE reduction of about 37.7% and an MAE reduction of about 40.1% relative to the official product. The final model also showed near-zero bias (0.021); whereas, the VNP product retained a pronounced negative bias (−0.137).

The year-wise overlap results were stable across the two external years. In 2019, the final model achieved an RMSE of 0.615 with an R^2 of 0.917, and in 2022 it achieved an RMSE of 0.603 with an R^2 of 0.883. Paired site-level comparisons across all 11 study sites are shown in Figures 7 and 8. These results indicate that the model improved both high-LAI forest scenes and low-LAI sparse vegetation scenes. For example, on the external overlap subset, RMSE values at BART and HARV were about 0.964 and 0.988, respectively, while CPER and SRER remained near 0.136 and 0.122, indicating good control of both dense-canopy underestimation and soil-background interference.

3.4. LOSO Validation

The LOSO experiment provided the clearest test of site-transfer performance. The overall LOSO response is shown in Figure 9, and the site-wise summary is shown in Figure 10. The final model achieved an overall R^2 of 0.901 with an RMSE of 0.703 and an MAE of 0.452 across all left-out sites. This was substantially better than the raw-label baseline, which reached an R^2 of only 0.573 and an RMSE of 1.456. The calibrated-but-no-PFT model produced intermediate performance, again showing that both corrected supervision and explicit subpixel composition contribute to spatial generalization.

The site-wise LOSO results also clarify the remaining limitations of the model. Low-LAI sites such as CPER and STER showed very small absolute errors even when site-level R^2 values were weak or negative because the variance of the reference LAI was

extremely small. In these cases, RMSE, MAE, and bias are more informative than R^2. The most difficult site was OSBS, where LOSO RMSE remained high and positive bias was substantial, suggesting that the current training domain does not fully represent some structurally complex canopy-understory conditions. Taken together, these site-holdout results indicate that the model captures relationships that remain useful when transferred to unseen sites within the evaluated GBOV/NEON domain, although performance still degrades at structurally atypical sites.

3.5. Heterogeneity-Stratified Error Analysis

To further examine whether the improvement of the proposed VIIRS retrieval was maintained under different degrees of mixed-pixel heterogeneity, we conducted a stratified residual analysis using the forest-spectral-enhanced model and the official VNP product on the common GBOV overlap samples. The analysis was performed for both the same-site temporal-transfer evaluation in 2019 and 2022 and the strict leave-one-site-out evaluation. Absolute retrieval errors were summarized as a function of subpixel forest fraction and PFT entropy, which, respectively, describe the degree of forest dominance and the compositional complexity within each 500 m pixel.

The stratified results show that the proposed model consistently reduced errors relative to VNP across the evaluated heterogeneity gradients (Figure 11). In the temporal-transfer evaluation, the proposed model achieved an RMSE of 0.591 on the VNP-overlap subset, compared with 0.978 for VNP. In the strict LOSO subset where VNP values were available, the corresponding RMSE values were 0.693 and 1.018, respectively. The improvement was not limited to low-complexity pixels. When samples were grouped by forest fraction, the proposed model showed lower MAE than VNP in all forest-fraction intervals. The advantage became particularly clear in high-forest pixels: for forest fraction between 0.8 and 1.0, the mean absolute error was reduced from 1.368 to 0.763 in the temporal-transfer evaluation and from 1.450 to 0.900 in the strict LOSO evaluation. Similar patterns were observed along the PFT entropy gradient. Across all entropy intervals, the proposed retrieval maintained lower MAE than VNP, indicating that its improvement was retained not only in relatively homogeneous pixels but also in compositionally mixed pixels.

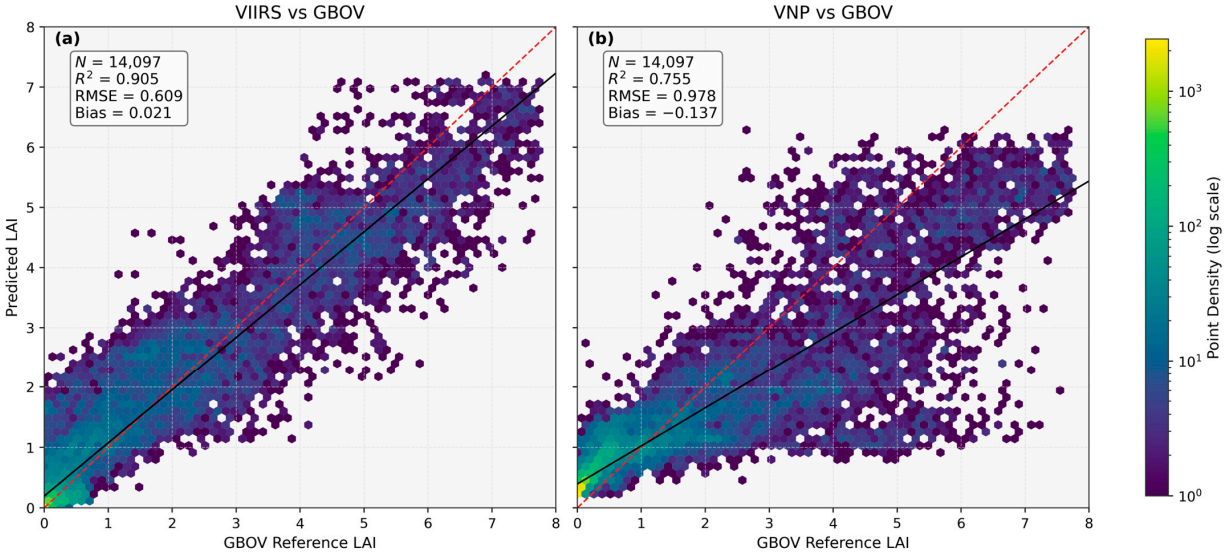

Figure 6. Overall external year comparison between the final VIIRS model and the official VNP product. Panel (**a**) shows the scatter comparison between the LAI retrieved by the final VIIRS model and the GBOV reference, while panel (**b**) shows the corresponding comparison between the official VNP LAI product and the GBOV reference. The red dashed line represents the 1:1 reference line, and solid black lines in site-level panels denote ordinary least-squares fits.

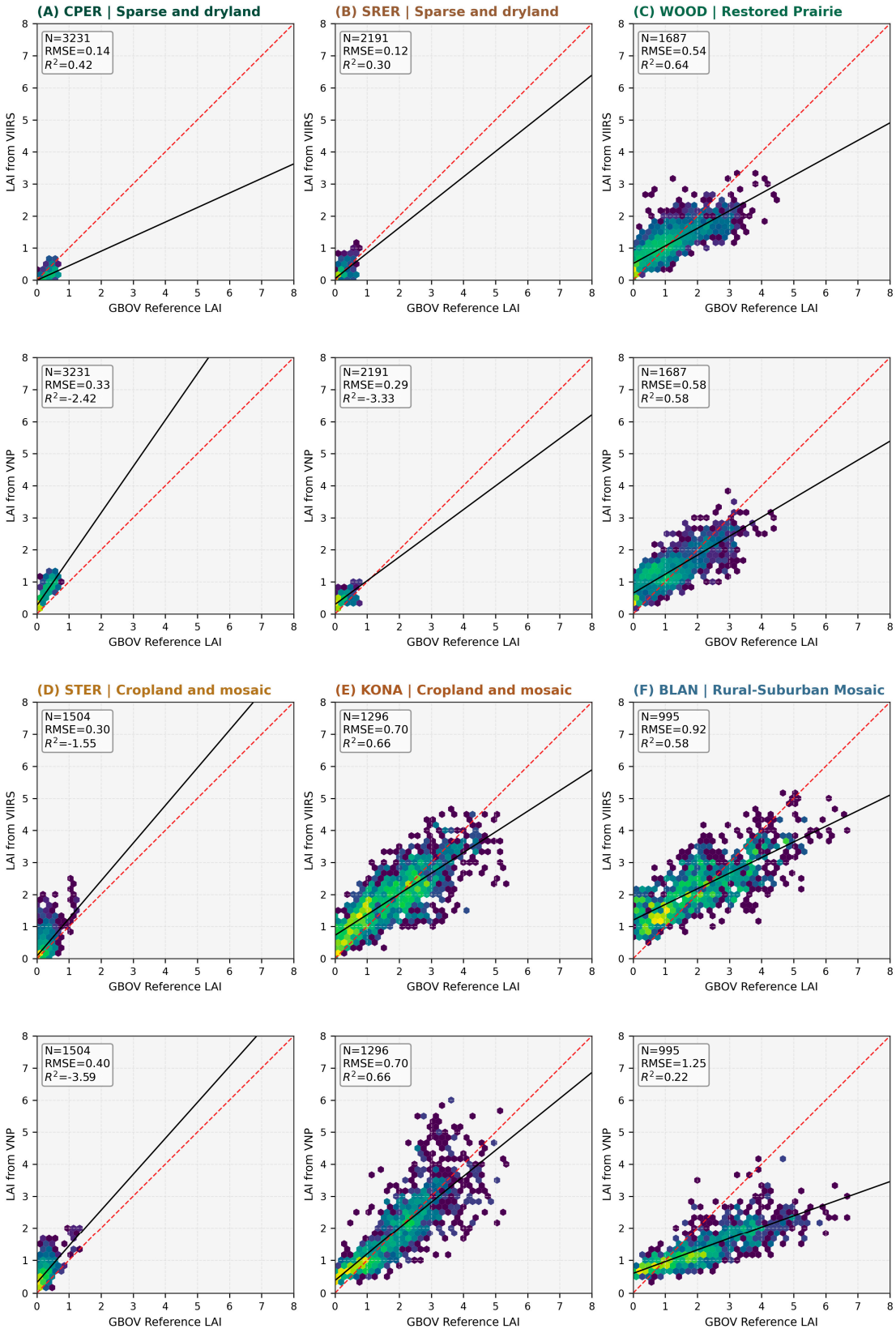

Figure 7. Site-level external year comparisons for CPER, SRER, WOOD, KONA, BLAN. The red dashed line represents the 1:1 reference line, and solid black lines in site-level panels denote ordinary least-squares fits.

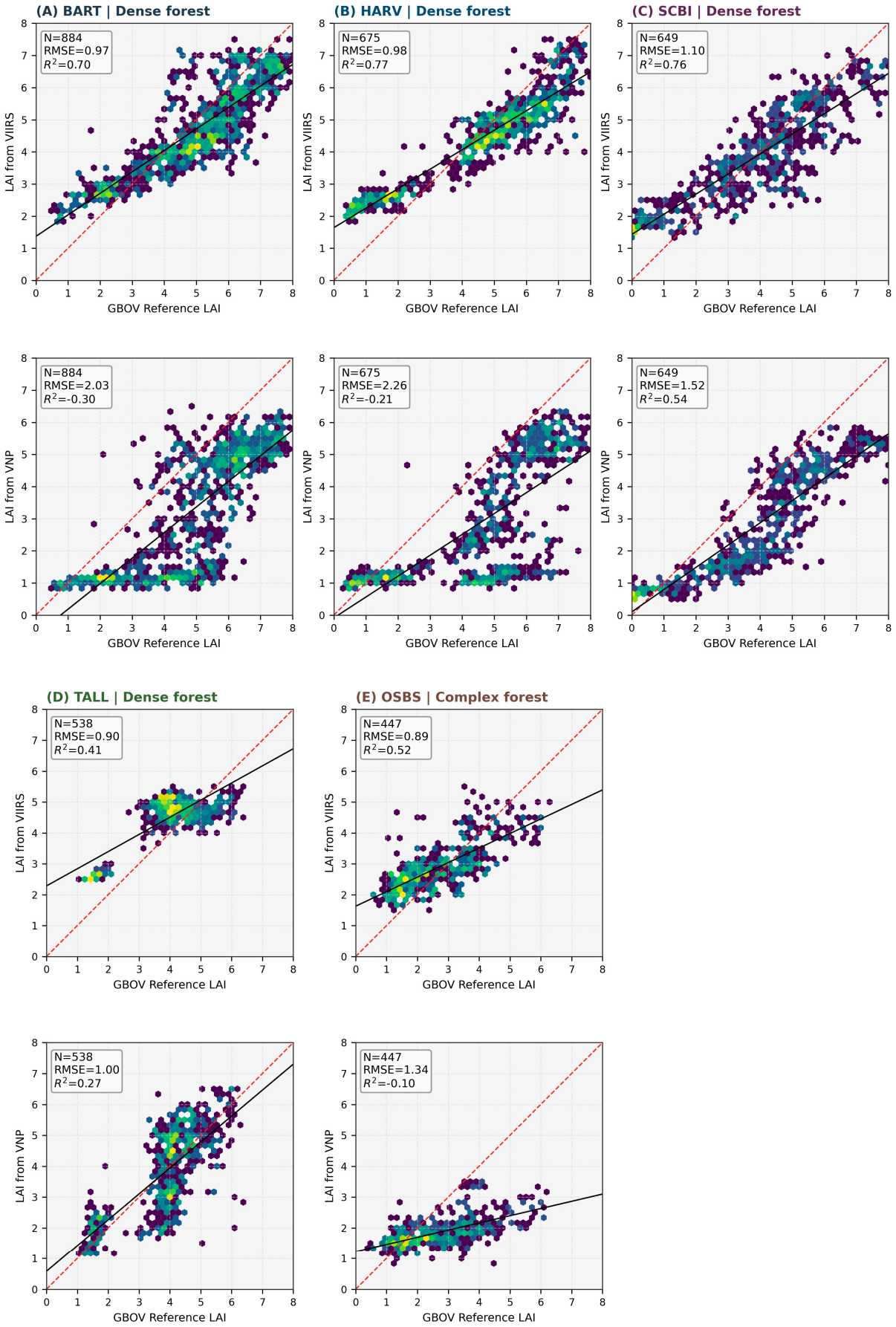

Figure 8. Site-level external year comparisons for BART, HARV, SCBI, TALL, OSBS. The red dashed line represents the 1:1 reference line, and solid black lines in site-level panels denote ordinary least-squares fits.

https://doi.org/10.3390/rs18132214

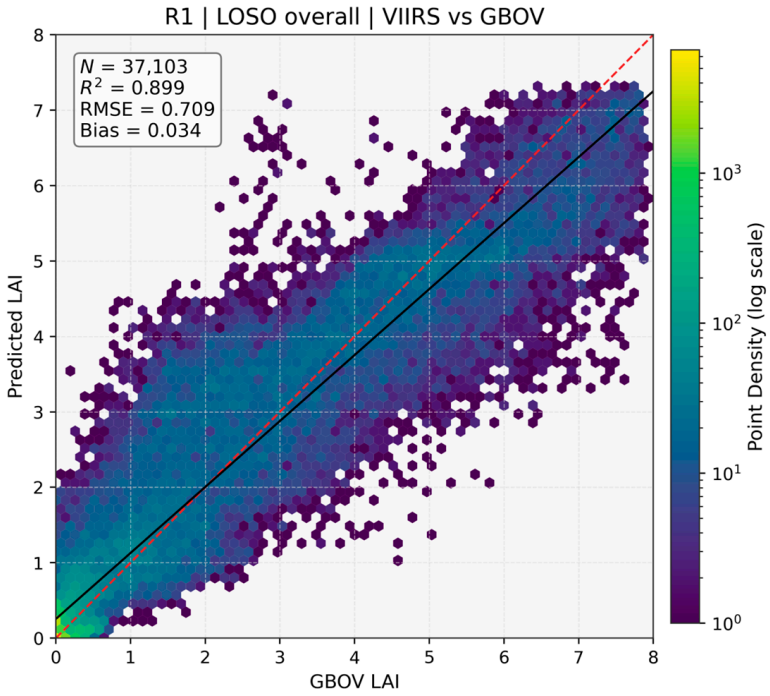

Figure 9. Overall strict leave-one-site-out validation scatter. The red dashed line represents the 1:1 reference line, and solid black lines in site-level panels denote ordinary least-squares fits.

Because sparse-canopy conditions may respond differently from dense vegetation regimes, we further stratified the GBOV-overlap samples by reference LAI (Figure 12). For LAI < 1, the proposed retrieval reduced MAE from 0.334 to 0.197 in the temporal-transfer evaluation and from 0.333 to 0.174 in the strict LOSO evaluation. The corresponding bias was also reduced from 0.327 to 0.122 and from 0.330 to 0.075, respectively. RMSE was nearly unchanged in the temporal-transfer subset for LAI < 1, but decreased from 0.390 to 0.318 in the strict LOSO subset. These results indicate that the proposed workflow did not degrade sparse-canopy retrievals and was able to reduce the positive bias of VNP under low-LAI conditions, although the small dynamic range of LAI < 1 limits the magnitude of RMSE improvement. The advantage of the proposed retrieval became more pronounced at moderate and high LAI levels. In the temporal-transfer evaluation, RMSE decreased from 1.700 to 0.823 for LAI between 3 and 5, and from 2.266 to 1.066 for LAI $\geq$ 5. In the strict LOSO evaluation, the corresponding RMSE decreased from 1.805 to 0.752 and from 1.716 to 1.167, respectively. The bias patterns further show that VNP exhibited increasingly negative bias as LAI increased; whereas, the proposed retrieval substantially reduced this underestimation. For LAI $\geq$ 5, the temporal-transfer bias was reduced from -1.891 for VNP to -0.784 for the proposed retrieval, and the strict LOSO bias was reduced from -1.359 to -0.828.

These results suggest that the combined use of GBOV-anchored label calibration, subpixel PFT constraints, and forest-sensitive spectral predictors improves not only the overall retrieval accuracy, but also the robustness of the retrieval under mixed-pixel conditions where operational LAI estimates are more prone to compositional and saturation-related errors.

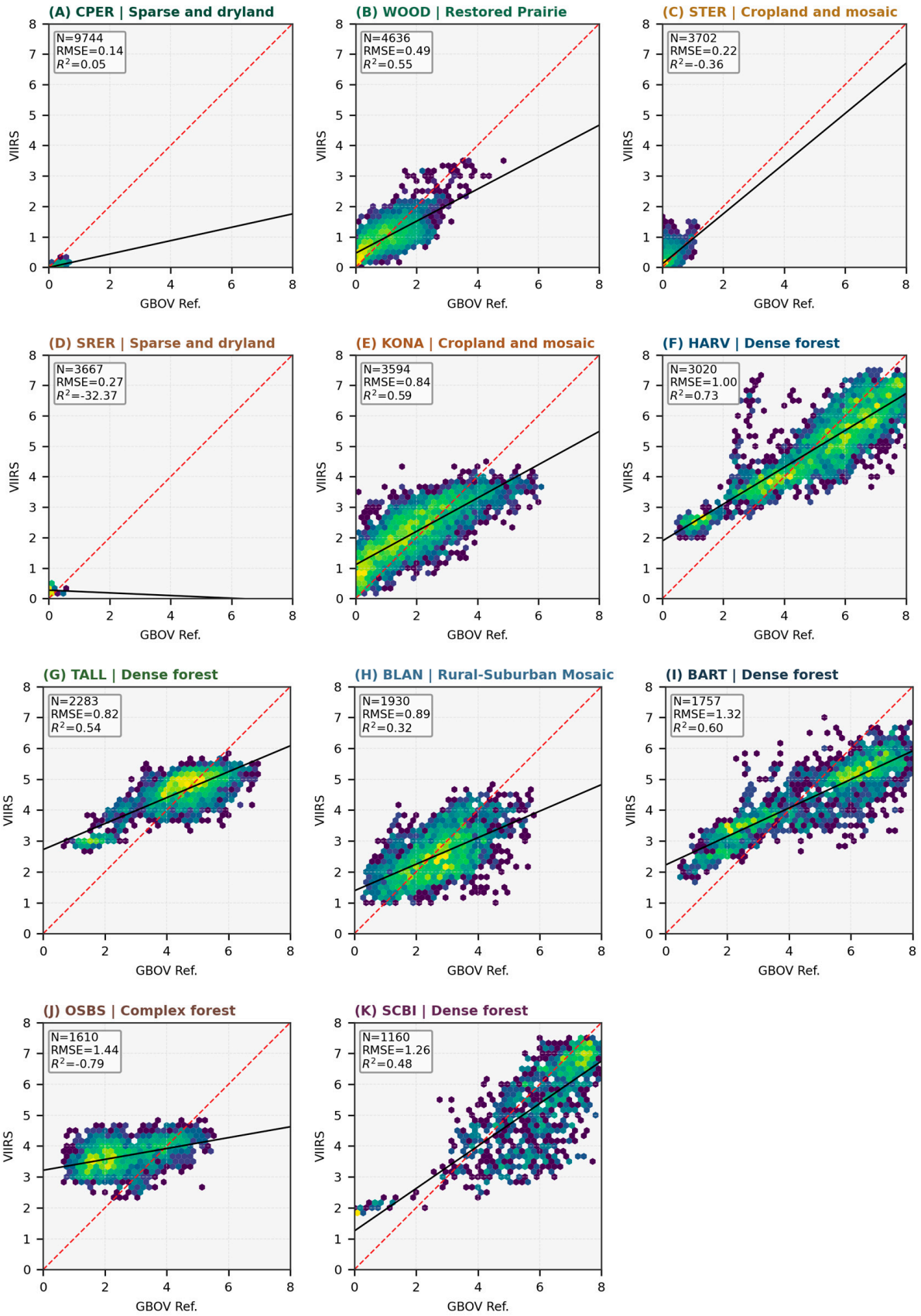

Figure 10. Site-level leave-one-site-out validation results across all 11 study sites. The red dashed line denotes the 1:1 reference line, and the black solid line denotes the ordinary least-squares fit.

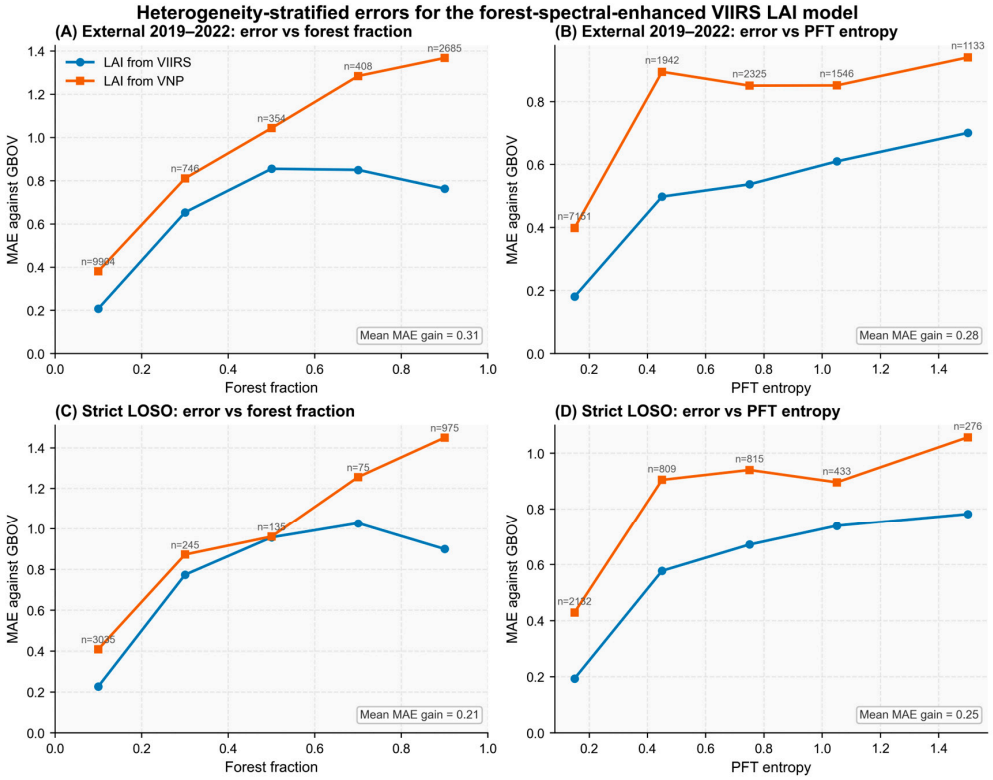

Figure 11. Heterogeneity-stratified error analysis for the forest-spectral-enhanced VIIRS LAI retrieval and the official VNP product. Panels (**A**,**C**) show mean absolute error as a function of subpixel forest fraction for the same site temporaltransfer evaluation and the strict LOSO evaluation, respectively. Panels (**B**,**D**) show mean absolute error as a function of PFT entropy. Points represent binned mean absolute errors against the GBOV reference, and sample counts are shown for each bin.

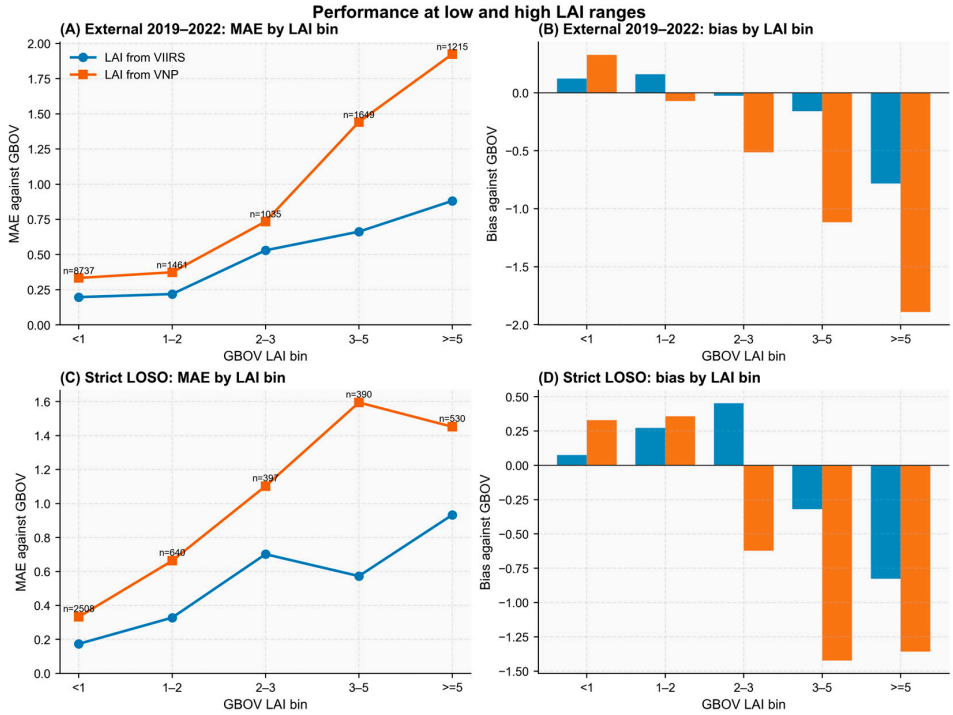

Figure 12. LAI-stratified error patterns for the proposed VIIRS retrieval and the official VNP product. In panels (**B**) and (**D**), blue bars indicate the proposed VIIRS retrieval, whereas orange bars indicate the official VNP product; positive and negative bias values represent overestimation and underestimation relative to the GBOV reference, respectively.

4. Discussion

4.1. Interpretation of the Two-Stage Improvement

The main methodological contribution of this study is not the use of a different regression algorithm in isolation. Rather, it lies in reformulating moderate-resolution mixed-pixel LAI retrieval as a coupled problem of supervisory-label bias and coarse-resolution inversion. In heterogeneous landscapes, unresolved subpixel structure can affect not only the mapping from coarse-resolution observations to LAI, but also the quality of the aggregated labels used for model training. This distinction motivated the two-stage design of the proposed workflow: Stage 1 improves the supervisory target through reference-guided label calibration; whereas, Stage 2 learns the final VIIRS retrieval with explicit subpixel compositional information.

The ablation experiments support this interpretation. The largest improvement was obtained by replacing raw aggregated Sentinel-2 labels with GBOV-anchored calibrated labels, indicating that mixed-pixel bias and high-LAI compression can be inherited through the training target itself. Thus, label calibration should be regarded as a core component of the workflow rather than as a secondary preprocessing step. The smaller but consistent gain from adding PFT fractions further shows that subpixel composition remains informative after label correction, helping both to construct a more reliable supervisory label and to provide structural information to the final VIIRS model.

4.2. Relationship to Existing Mixed-Pixel and VIIRS LAI Studies

The results of this study are closely related to previous work on scale effects and mixed-pixel uncertainty in moderate-resolution LAI retrieval. Multiscale validation studies of MODIS LAI products have shown that spatial heterogeneity, sampling strategy, and nonlinear canopy reflectance–LAI relationships can strongly affect product uncertainty when fine-scale vegetation structure is aggregated to coarser pixels [5–7]. These studies provide the theoretical basis for treating mixed pixels as more than a simple spatial-resolution mismatch. The present results are consistent with this view, but further indicate that heterogeneity can influence not only the coarse-resolution predictor space, but also the supervisory labels used to train retrieval models.

Recent studies have addressed mixed-pixel and structural effects by incorporating subpixel background components, mixed-biome correction factors, decametric-resolution information, or PFT-aware treatment of canopy and soil-background effects [11–14]. Compared with these approaches, the present workflow uses subpixel information in a complementary way. PFT fractions are not only added as predictors in the final VIIRS model; they are also used in the reference-guided label-calibration stage to adjust the aggregated high-resolution target. This design explicitly links subpixel composition to both the training label and the final inversion model.

This work is also related to recent efforts to improve VIIRS and MODIS LAI records through product evaluation, transfer learning, domain adaptation, temporal-consistency enhancement, and reprocessed high-quality LAI datasets [10,15–17,20–23]. These studies mainly address algorithm transfer, product stability, or time-series consistency. In contrast, the present study focuses on the quality of the training target used for moderate-resolution retrieval. This distinction is important because a model trained on biased aggregated labels can reproduce label-side errors even when the predictor variables and learning algorithm are improved.

Therefore, the proposed workflow should be viewed as complementary to existing mixed-pixel correction and VIIRS LAI retrieval approaches rather than as a replacement for them. The results suggest that moderate-resolution LAI retrieval over heterogeneous landscapes benefits from jointly addressing two coupled sources of uncertainty: predictor-

side mixing in VIIRS observations and label-side bias in aggregated high-resolution training targets. Nevertheless, direct numerical comparison with published VIIRS or MODIS LAI retrieval accuracies should be interpreted cautiously because of differences in study region, reference data, quality screening, temporal matching, and validation protocol. The purpose of this study is not to claim universal superiority over existing products or algorithms, but to demonstrate that explicit supervisory-label calibration can strengthen VIIRS LAI retrieval within a harmonized mixed-pixel evaluation framework.

4.3. Performance Across Heterogeneity and LAI Gradients

The heterogeneity-stratified analysis provides an important check on whether the proposed workflow improves retrieval only in relatively simple pixels or remains effective under mixed-pixel conditions. The results show that the performance gain was retained across both forest-fraction and PFT-entropy gradients, indicating that the improvement was not driven solely by homogeneous or low-complexity samples. This finding supports the central assumption of the study: mixed-pixel effects should be addressed through both a better supervisory target and an explicit representation of subpixel composition. The stronger advantage observed in high-forest and high-entropy pixels further suggests that PFT information helps compensate for structural and compositional variability that cannot be fully captured by VIIRS reflectance, angular variables, or contextual texture metrics alone. It should also be noted that the heterogeneity analysis in this study characterizes compositional heterogeneity using PFT fractions and PFT-based entropy. This metric is useful for quantifying land-cover mixture within coarse pixels, but it does not directly measure subpixel canopy structural heterogeneity.

The LAI-stratified analysis also clarifies the operating range of the workflow. Under sparse-canopy conditions, the proposed retrieval reduced bias and MAE without degrading overall performance, suggesting that the label-calibration and PFT-constrained design did not improve dense vegetation at the expense of low-LAI pixels. The largest gains occurred in moderate-to-high LAI ranges, where VNP showed stronger underestimation and where optical saturation and label compression are more likely to affect retrieval. Therefore, the main benefit of the workflow is not a uniform error reduction across all conditions, but a targeted improvement in the regimes where moderate-resolution LAI products are most vulnerable to mixed-pixel structure, canopy saturation, and background effects.

4.4. Limitations and Domain of Applicability

The proposed workflow should be interpreted within a clearly defined domain of applicability. First, the model was developed and evaluated using 11 U.S. GBOV/NEON sites, which cover useful gradients in vegetation density, canopy structure, and background conditions, but do not represent the full global diversity of tropical forests, equatorial savannas, boreal forests, drylands, and intensively managed agricultural systems. This limitation is particularly relevant for machine-learning retrieval because random forests are nonparametric ensemble models whose predictions remain dependent on the predictor–response space represented in the training data [33,34]. Applications to canopy states, management regimes, or biome conditions that are poorly represented in the current training domain should therefore be regarded as extrapolations beyond the present area of support, consistent with the area-of-applicability concept in spatial prediction [35].

This domain limitation is also relevant to the fixed hinge threshold of T = 4.0 and to high-LAI retrieval. The threshold provides a practical and interpretable way to introduce additional flexibility in the high-LAI range and is motivated by the known tendency of optical vegetation signals to saturate at moderate-to-high LAI values [2,4,13]. However, saturation behavior varies with canopy architecture, leaf angle distribution, under-

story contribution, background brightness, and illumination-viewing geometry. Therefore, T = 4.0 should be interpreted as a domain-specific calibration choice rather than as a universal biophysical threshold. Although the hinge-based correction improves the supervisory labels under high-LAI conditions within the calibrated domain, it does not provide the final random-forest model with unrestricted extrapolation capability. If validation or application samples contain canopy states with LAI values higher than those represented by the corrected training labels, the model may tend to constrain predictions toward the upper end of the learned target distribution, leading to residual underestimation under extreme high-LAI conditions.

Second, the workflow also depends on the quality of the reference and auxiliary datasets used to construct labels and predictors. Errors in the Sentinel-2 LAI prior can influence the initial aggregated label, while the GBOV LP3 benchmark reference may contain residual measurement and upscaling errors, including its PAI-like measurement characteristics in some cases [9,29]. In addition, the GLC-FCS30D land-cover product is used to derive PFT fractions, and misclassification in this product may affect both Stage 1 label calibration and Stage 2 retrieval [28]. For example, an incorrect PFT fraction may influence the estimated label correction and simultaneously provide a misleading compositional predictor to the final VIIRS model. The heterogeneity analysis also relies mainly on PFT-derived forest fraction and PFT entropy. These metrics provide interpretable descriptions of compositional mixing, but they do not fully capture within-class structural heterogeneity, canopy clumping, topographic effects, understory variability, or subpixel variation in soil background and management conditions.

Third, sampling dependence should also be considered when interpreting the reported accuracy. Although the dataset contains a large number of pixel-level samples, neighboring pixels and repeated seasonal observations from the same sites are not fully independent. Such spatial and temporal dependence can lead to optimistic accuracy estimates in spatial predictive modeling [36]. In this study, same-site temporal-transfer experiments and strict leave-one-site-out validation were used to reduce this risk by separating years and sites. These protocols provide a more conservative assessment than random pixel-level validation alone, but they cannot fully remove all dependence among neighboring pixels or repeated observations. The $\pm$3-day temporal-matching window also reduces, but does not eliminate, phenological mismatch, especially during rapid spring growth, harvesting, irrigation, disturbance, or early-season vegetation development.

Finally, the multi-stage and multi-feature design improves retrieval accuracy but may require additional evaluation before broad operational transfer, especially compared with a simpler VIIRS-only retrieval. Although the main predictor-side datasets used in this study are publicly available, the framework still requires high-resolution LAI priors, reliable benchmark references, land-cover composition information, and careful spatiotemporal matching. These requirements define the practical conditions under which the method is most suitable. Future applications should therefore test the workflow across broader biome and management gradients, evaluate its sensitivity to missing or misclassified PFT information, and explore whether simplified predictor configurations can retain most of the observed accuracy gains.

4.5. Implications and Future Work

The results suggest that reference-guided label correction is a practical route for improving moderate-resolution LAI retrieval over heterogeneous landscapes. For product-oriented applications, the workflow is most useful where high-resolution LAI priors, reliable benchmark references, and land-cover composition information are available. Although most predictor-side datasets used in this study are publicly accessible, the availabil-

ity and quality of benchmark reference data remain important constraints for transferring the workflow to regions without dense validation networks. Therefore, the proposed framework should be viewed as a reference-guided methodological strategy rather than an immediately universal operational product. More broadly, this study highlights that the quality of the supervisory target is a central issue in coarse-resolution vegetation retrieval, especially when training labels are derived by aggregating finer-resolution products over heterogeneous surfaces.

Future work should extend the framework along the three uncertainty pathways identified above. First, broader testing is needed across tropical rainforests, equatorial savannas, boreal forests, drylands, and intensively managed agricultural systems to determine the area of applicability of the current calibration strategy. Second, the reference-feature chain could be strengthened by incorporating additional structural information, such as canopy height, SAR backscatter, or LiDAR-derived metrics, which may improve robustness in dense or vertically complex canopies. Third, adaptive label-calibration strategies, biome-specific saturation thresholds, and uncertainty-aware learning frameworks should be explored to explicitly account for reference uncertainty, land-cover classification uncertainty, and spatial-temporal dependence. These extensions would help transform the current domain-bounded proof of concept into a more generalizable framework for operational moderate-resolution LAI retrieval.

5. Conclusions

This study developed a two-stage framework for improving 500 m VIIRS LAI retrieval over mixed pixels through GBOV-anchored label calibration and subpixel PFT constraints. The central idea is that mixed-pixel LAI retrieval should be treated as a coupled problem of supervisory-label bias and coarse-resolution inversion, rather than as a predictor-side problem alone. By explicitly separating these two steps, the proposed framework addresses bias embedded in aggregated high-resolution labels before learning the final VIIRS retrieval model. Within the evaluated GBOV study domain, the results show that label calibration was the primary source of performance improvement, while subpixel PFT information provided an additional and consistent gain in the final inversion stage. The framework performed well in same-site temporal transfer to 2019 and 2022, and remained robust under strict leave-one-site-out blind testing. Taken together, these findings support a workflow-level innovation: mixed-pixel heterogeneity is most effectively addressed when supervisory-label debiasing and final coarse-resolution retrieval are treated as connected but distinct steps. Nevertheless, the current evaluation is bounded by 11 U.S.-based GBOV/NEON sites and should be regarded as a methodological proof of concept rather than evidence of global transferability. Future work should test the framework in highly distinct global biomes, including tropical rainforests, equatorial savannas, boreal forests, drylands, and intensively managed agricultural systems, where canopy-background interactions, canopy architecture, and optical saturation behavior may differ substantially from those represented in the present study.

Author Contributions: Conceptualization, T.Y. and H.D.; methodology, T.Y.; software, T.Y.; validation, T.Y. and H.D.; formal analysis, T.Y.; investigation, T.Y.; resources, H.D.; data curation, T.Y. and Y.Z.; writing—original draft preparation, T.Y. and Y.Z.; writing—review and editing, T.Y. and H.D.; visualization, T.Y. and Y.Z.; supervision, H.D.; project administration, T.Y.; funding acquisition, H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2022YFC3004200.

Data Availability Statement: The GBOV LP3 LAI reference data are publicly available from the Copernicus Ground-Based Observations for Validation (GBOV) service (https://land.copernicus.eu/en/products/GBOV (accessed on 24 April 2026)). VIIRS surface reflectance data were obtained from the NASA LP DAAC VNP09GA Version 2 product (https://doi.org/10.5067/VIIRS/VNP09GA.002 (accessed on 24 April 2026)) and accessed through Google Earth Engine. The GLC_FCS30 global land-cover product used to derive subpixel PFT fractions is publicly available from its original data providers and was accessed through Google Earth Engine.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (GPT-5.4, OpenAI) for the purposes of translating and polishing selected manuscript sections. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Meteorological Organization. Leaf Area Index. Available online: https://space.oscar.wmo.int/variables/view/leaf_area_index_lai (accessed on 24 April 2026).
2. Zheng, G.; Moskal, L.M. Retrieving leaf area index (LAI) using remote sensing: Theories, methods and sensors. *Sensors* **2009**, *9*, 2719–2745. [CrossRef] [PubMed]
3. Myneni, R.B.; Hoffman, S.; Knyazikhin, Y.; Privette, J.L.; Glassy, J.; Tian, Y.; Wang, Y.; Song, X.; Zhang, Y.; Smith, G.R.; et al. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* **2002**, *83*, 214–231. [CrossRef]
4. Mthembu, N.; Lottering, R.; Kotze, H. Forest, crop and grassland leaf area index estimation using remote sensing: A review of current research methods, sensors, estimation models and accomplishments. *Appl. Sci.* **2023**, *13*, 4005. [CrossRef]
5. Tian, Y.; Woodcock, C.E.; Wang, Y.; Privette, J.L.; Shabanov, N.V.; Zhou, L.; Zhang, Y.; Buermann, W.; Dong, J.; Veikkanen, B.; et al. Multiscale analysis and validation of the MODIS LAI product: I. Uncertainty assessment. *Remote Sens. Environ.* **2002**, *83*, 414–430. [CrossRef]
6. Tian, Y.; Wang, Y.; Zhang, Y.; Knyazikhin, Y.; Bogaert, J.; Myneni, R.B. Multiscale analysis and validation of the MODIS LAI product: II. Sampling strategy. *Remote Sens. Environ.* **2002**, *83*, 431–441. [CrossRef]
7. Garrigues, S.; Allard, D.; Baret, F.; Weiss, M. Influence of landscape spatial heterogeneity on the non-linear estimation of leaf area index from moderate spatial resolution remote sensing data. *Remote Sens. Environ.* **2006**, *105*, 286–298. [CrossRef]
8. Fang, H.; Wei, S.; Liang, S. Theoretical uncertainty analysis of global MODIS, CYCLOPES, and GLOBCARBON LAI products using a triple collocation method. *Remote Sens. Environ.* **2012**, *124*, 610–621. [CrossRef]
9. Brown, L.A.; Meier, C.; Morris, H.; Pastor-Guzman, J.; Bai, G.; Lerebourg, C.; Gobron, N.; Lanconelli, C.; Clerici, M.; Dash, J. Evaluation of global leaf area index and fraction of absorbed photosynthetically active radiation products over North America using Copernicus Ground Based Observations for Validation data. *Remote Sens. Environ.* **2020**, *247*, 111935. [CrossRef]
10. Yan, K.; Park, T.; Yan, G.; Chen, C.; Yang, B.; Liu, Z.; Nemani, R.R.; Knyazikhin, Y.; Myneni, R.B. Performance stability of the MODIS and VIIRS LAI algorithms inferred from analysis of long time series of products. *Remote Sens. Environ.* **2021**, *260*, 112438. [CrossRef]
11. Xu, B.; Li, J.; Park, T.; Liu, H.; Xiao, Z.; Knyazikhin, Y. Improving leaf area index retrieval over heterogeneous surface mixed with water. *Remote Sens. Environ.* **2020**, *240*, 111700. [CrossRef]
12. Dong, Y.; Li, J.; Jiao, Z.; Liu, Q.; Zhao, J.; Xu, B.; Zhang, H.; Zhang, Z.; Liu, C.; Knyazikhin, Y.; et al. A method for retrieving coarse-resolution leaf area index for mixed biomes using a mixed-pixel correction factor. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4400317. [CrossRef]
13. Wan, L.; Ryu, Y.; Dechant, B.; Hwang, Y.; Feng, H.; Kang, Y.; Jeong, S.; Lee, J.; Choi, C.; Bae, J. Correcting confounding canopy structure, biochemistry and soil background effects improves leaf area index estimates across diverse ecosystems from Sentinel-2 imagery. *Remote Sens. Environ.* **2024**, *309*, 114224. [CrossRef]
14. Jin, W.; Zhang, Z.; Wu, T.; Meng, K.; Wang, Q.; Tong, W.; Wang, C.; Yin, G.; Xu, B. Retrieval of leaf area index from MODIS surface reflectance by incorporating the subpixel information from decametric-resolution data. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–17. [CrossRef]
15. Zou, D.; Yan, K.; Pu, J.; Gao, S.; Li, W.; Mu, X.; Knyazikhin, Y.; Myneni, R.B. Revisit the performance of MODIS and VIIRS leaf area index products from the perspective of time-series stability. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 8958–8973. [CrossRef]
16. Li, J.; Xiao, Z.; Sun, R.; Song, J. Retrieval of the leaf area index from Visible Infrared Imaging Radiometer Suite (VIIRS) surface reflectance based on unsupervised domain adaptation. *Remote Sens.* **2022**, *14*, 1826. [CrossRef]

17. Li, J.; Xiao, Z.; Sun, R.; Song, J. A method to estimate leaf area index from VIIRS surface reflectance using deep transfer learning. *ISPRS J. Photogramm. Remote Sens.* **2023**, *202*, 512–527. [CrossRef]
18. Peng, R.; Yan, K.; Liu, J.; Gao, S.; Pu, J.; Maeda, E.E.; Zhu, P.; Knyazikhin, Y.; Myneni, R.B. Revisiting the consistency of MODIS LAI products from a new perspective of spatiotemporal variability. *Int. J. Digit. Earth* **2024**, *17*, 2407045. [CrossRef]
19. Zhang, X.; Liu, J.; Yang, K.; Pu, J.; Yan, G.; Heiskanen, J.; Zhu, P.; Knyazikhin, Y.; Myneni, R.B. An insight into the internal consistency of MODIS global leaf area index products. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 4411716. [CrossRef]
20. Pu, J.; Yan, K.; Roy, S.; Zhu, Z.; Rautiainen, M.; Knyazikhin, Y.; Myneni, R.B. Sensor-independent LAI/FPAR CDR: Reconstructing a global sensor-independent climate data record of MODIS and VIIRS LAI/FPAR from 2000 to 2022. *Earth Syst. Sci. Data* **2024**, *16*, 15–34. [CrossRef]
21. Pu, J.; Yan, K.; Gao, S.; Zhang, Y.; Park, T.; Sun, X.; Weiss, M.; Knyazikhin, Y.; Myneni, R.B. Improving the MODIS LAI compositing using prior time-series information. *Remote Sens. Environ.* **2023**, *287*, 113493. [CrossRef]
22. Wang, J.; Yan, K.; Gao, S.; Pu, J.; Liu, J.; Park, T.; Bi, J.; Maeda, E.E.; Heiskanen, J.; Knyazikhin, Y.; et al. Improving the quality of MODIS LAI products by exploiting spatiotemporal correlation information. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4402319. [CrossRef]
23. Yan, K.; Wang, J.; Peng, R.; Yang, K.; Chen, X.; Yin, G.; Dong, J.; Weiss, M.; Pu, J.; Myneni, R.B. HiQ-LAI: A high-quality reprocessed MODIS leaf area index dataset with better spatiotemporal consistency from 2000 to 2022. *Earth Syst. Sci. Data* **2024**, *16*, 1601–1622. [CrossRef]
24. National Ecological Observatory Network. Explore Field Sites. Available online: https://www.neonscience.org/field-sites/explore-field-sites (accessed on 24 April 2026).
25. Copernicus Land Monitoring Service. Ground-Based Observations for Validation (GBOV). Available online: https://land.copernicus.eu/en/products/GBOV (accessed on 24 April 2026).
26. Vermote, E.; Franch, B.; Claverie, M. *VIIRS/NPP Surface Reflectance Daily L2G Global 1 Km and 500 M SIN Grid V002*; NASA EOSDIS Land Processes Distributed Active Archive Center: Sioux Falls, SD, USA, 2023. [CrossRef]
27. NASA Land Processes Distributed Active Archive Center. Suomi-NPP VIIRS Surface Reflectance User's Guide: V2 Reprocessing, Version 2.0. Available online: https://lpdaac.usgs.gov/documents/1657/VNP09_User_Guide_V2.pdf (accessed on 24 April 2026).
28. Zhang, X.; Liu, L.; Chen, X.; Gao, Y.; Xie, S.; Mi, J. GLC_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery. *Earth Syst. Sci. Data* **2021**, *13*, 2753–2776. [CrossRef]
29. Copernicus Land Monitoring Service. Algorithm Theoretical Basis Document for Vegetation Products: LP3 LAI, LP4 FAPAR and LP5 FCOVER, Version 3.2. Available online: https://land.copernicus.eu/en/technical-library/algorithm-theoretical-basis-document-for-vegetation-products-lp-3-lai-lp-4-fapar-and-lp-5-fcover/%40%40download/file (accessed on 24 April 2026).
30. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]
31. Zhao, J.; Li, J.; Liu, Q.; Yang, L. A preliminary study on mechanism of LAI inversion saturation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *XXXIX-B1*, 77–81. [CrossRef]
32. Baret, F.; Hagolle, O.; Geiger, B.; Bicheron, P.; Miras, B.; Huc, M.; Berthelot, B.; Niño, F.; Weiss, M.; Samain, O.; et al. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. *Remote Sens. Environ.* **2007**, *110*, 275–286. [CrossRef]
33. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
34. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
35. Meyer, H.; Pebesma, E. Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods Ecol. Evol.* **2021**, *12*, 1620–1633. [CrossRef]
36. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **2017**, *40*, 913–929. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.